



An Individual Finger Gesture Recognition System Based on Motion-Intent Analysis Using Mechanomyogram Signal

Huijun Ding¹, Qing He¹, Yongjin Zhou¹, Guo Dan^{1,2*} and Song Cui³

¹Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, School of Biomedical Engineering, Health Science Center, Shenzhen University, Guangdong, China, ²Center for Neurorehabilitation, Shenzhen Institute of Neuroscience, Guangdong, China, ³Institute of High Performance Computing, Singapore, Singapore

OPEN ACCESS

Edited by:

Ping Zhou,
University of Texas Health Science
Center at Houston, United States

Reviewed by:

Hongbo Xie,
Queensland University of
Technology, Australia
Keum-Shik Hong,
Pusan National University,
South Korea

*Correspondence:

Guo Dan
danguo@szu.edu.cn

Specialty section:

This article was submitted to Stroke,
a section of the journal *Frontiers in
Neurology*

Received: 12 August 2017

Accepted: 12 October 2017

Published: 08 November 2017

Citation:

Ding H, He Q, Zhou Y, Dan G and Cui S (2017) An Individual Finger Gesture Recognition System Based on Motion-Intent Analysis Using Mechanomyogram Signal. *Front. Neurol.* 8:573. doi: 10.3389/fneur.2017.00573

Motion-intent-based finger gesture recognition systems are crucial for many applications such as prosthesis control, sign language recognition, wearable rehabilitation system, and human-computer interaction. In this article, a motion-intent-based finger gesture recognition system is designed to correctly identify the tapping of every finger for the first time. Two auto-event annotation algorithms are firstly applied and evaluated for detecting the finger tapping frame. Based on the truncated signals, the Wavelet packet transform (WPT) coefficients are calculated and compressed as the features, followed by a feature selection method that is able to improve the performance by optimizing the feature set. Finally, three popular classifiers including naive Bayes (NBC), K-nearest neighbor (KNN), and support vector machine (SVM) are applied and evaluated. The recognition accuracy can be achieved up to 94%. The design and the architecture of the system are presented with full system characterization results.

Keywords: mechanomyogram, inertial sensor, finger gesture recognition, motion intent, feature selection

1. INTRODUCTION

Modeling and recognizing human hand gesture is an extremely important research topic, and it is the core of any intelligent human-computer interaction system with applications in automatic control, virtual reality, augmented reality, human-robotic interaction, and computer animation. In addition, it attracts more attentions and interests in biomedical engineering recently, e.g., prosthesis control and wearable rehabilitation system. As it is difficult to model the gesture accurately from images and the gesture appearance varies a lot, it is still a difficult task to recognize and track hand gestures. On the other hand, stroke is the leading cause of disability in adults worldwide. Exercise disorders as the most common sequelae of stroke can seriously affect normal activity and quality of life. Exercise and training have long been used to restore motor function after stroke. Among that, hand movement as a kind of fine action is more difficult to recover. Finger gesture recognition technique can be used to train and lead the patients to do the rehabilitation exercises, which normally requires real-time interaction and long-term monitoring. Thus, portable system with real-time processing technique is a valuable topic to pay attention and be explored in the stroke application area.

Currently, most hand-gesture recognition systems are using either hand-motion-based (HMB) technologies or motion-intent-based (MIB) technologies. HMB methods are capable of directly detecting the tracking of hand motions. Non-skin attached sensors such as optical sensors, inertial sensors, and force sensors (1) are usually deployed in HMB technologies. Many commercial

products, such as Kinect and Data Glove, are built on HMB technologies. However, the performance of HMB-based hand-gesture recognition system degrades when the lighting variations present. The recognition accuracy will also drop significantly if skin colors are too similar to background colors or the usage is in dark environment. In addition, the coverage of the system is limited by the sensors' measurement range and attaching sensors to fingers and hands make user hand movements unsmoothly and uncomfortable.

MIB methods measure the motion intentions instead of the actual hand motions, which is critical in certain applications such as prosthesis control, virtual reality, and motor rehabilitation. Hand-gesture recognition systems relying on MIB technologies are capable of recognizing the hand gestures based on the forearm muscle activities (2) or the signal decoding from brain (3). The brain related methods are driven by the neuroplasticity for stroke rehabilitation and are explained more from the nervous system. Brain cortex activities are analyzed by imaging techniques, e.g., electroencephalography (4, 5), electrocorticography (6), near-infrared spectroscopy (7, 8), magnetic resonance imaging (9), and optical tomography (10). A review of hybrid brain-computer interface techniques can also be found in Ref. (11). All these methods by using imaging technologies are with high cost and sensitive to the experimental environment and setup. For example, electroencephalography is not reliable under exposure to high-intensity magnetic fields and also cannot be applied to the participants having metal implants in their body (12).

The most commonly used forearm muscle analysis solutions relying on MIB approach are the surface electromyography (sEMG) method and the mechanomyogram (MMG) method. This kind of methods explains more about the efforts made by muscles. sEMG uses the surface electrodes to record the electrical currents signal produced by the muscular contraction and many sEMG-based systems have been proposed in the past decades (13). On the other hand, the MMG approach becomes an active topic recently because these sensors can be made into small sizes with reduced cost and good performances. MMG is a low-frequency mechanical signal which can be detected during the muscular contracting period. It is transmitted from the muscle to the surface skin by the soft tissue, and it can be detected by the motion sensors, i.e., inertial sensors (14, 15), laser sensors (16, 17), and microphones (18, 19) attached to the skin. Among all the sensors, inertial sensors are widely used. They are cheaper and more wearable than laser sensors (20). For the microphone-based MMG acquisition system, an air chamber is normally required to be placed between the condenser microphone and the surface of the skin to improve the quality of the acoustic signal. The design of the air chamber is a key factor which affects the frequency response of the acquisition system (19). In addition, the microphone-based MMG acquisition system is less wearable than inertial sensor-based system. Since different hand gestures are due to different modes of voluntary isometric contraction (17), it is possible to recognize the hand gestures based on the MMG signal detected from the forearm muscle group (14, 21). Compared with the sEMG approaches, MMG approaches have some advantages. First of all, MMG provides flexibility in setting up the sensors. Since the

MMG signal can be detected on the distal of soft tissue during the muscle contraction (14), the location for placing the sensors are more flexible in contrast to the sEMG, which normally requires an experienced technician to find out the good positions for the sEMG sensors. This flexibility will give MMG users better user experiences compared with sEMG users. Second, the MMG signal is independent of the skin impedance. On the other hand, the sEMG signal is easier to be affected by the skin impedance, which is one of the shortcomings of sEMG-based approach (22). This drawback makes the sEMG sensors unstable due to the reasons that the skin impedance is sensitive to many factors such as temperature and humidity. Finally, systems with MMG have low cost and have less computational loads. MMG is a low-frequency signal with the range of 10–22 Hz (23) while sEMG signal is in the range of 50–900 Hz (15). In that case, there is no need for high-frequency electronic components in the MMG acquisition system and lower sampling rate produces less data for processing thus reduces the computational load.

Due to these advantages, many hand-gesture identification systems are built using MMG–MIB approach. A recognition system is built for identifying the flexion and extensor of the wrist in 1986 (14). The MMG signal was collected by a microphone from the flexion digitorum and extensor digitorum in the system. The amplitude of the MMG signal is used to classify these two gestures. Based on a similar idea, a recognition system is built and capable of identifying additional hand gestures including wrist flex, wrist extensor, hand open, and hand close (24). In this system, the MMG capture system is deploying acceleration sensors. Then, wavelet packet transform (WPT) is used to process the MMG raw signal, followed by the singular value decomposition (SVD) to reduce dimensions of features for gesture recognition. Based on a linear discriminant analysis classifier, the identification algorithm gives an average accuracy rate of 89.7%. After that, many researchers (25–27) extend the study for the recognition of different hand gestures.

There are only few reported MMG–MIB systems on finger gesture recognition (FGR) in the literature. Three types of finger gestures including thumb flexion, pinkie flexion, and middle three finger flexions are classified in an FGR system based on the microphone-accelerometer (28). This system recognizes these finger gestures by root mean square (RMS) of the MMG signal amplitudes, and its average identification accuracy rate is 76.2%. However, individual flexion of the middle three fingers cannot be distinguished in this FGR system. Subsequently, a system is presented in Ref. (29) to identify the middle three finger gestures, namely, index tapping, middle tapping, and ring finger tapping. Two tri-axial accelerometers are used to record the MMG signal from the forearm muscles. Three types of finger gestures can be recognized, and the average accuracy rate is 75%. At the same time, some researchers dedicate to recognize more motion patterns of a single finger. For instance, a system proposed in Ref. (30) is capable of classifying four thumb motion patterns, which includes flexion, extension, abduction, and adduction. The features are extracted by a hybrid algorithm combing the mean absolute value of MMG signal, RMS of amplitude, mean frequency, etc. The average accuracy rate is 81.5% achieved by the quadratic discriminant analysis.

Among all the existing FGR systems, none of them can recognize the individual movement of five fingers, which is the goal of our efforts. Our previous work (31) proves that the individual finger tapping can be recognized based on the analysis of forearm muscle contractions, and a support vector machine (SVM) classifier shows a good performance on classification and recognition. In this article, we further present a FGR system capable of recognizing thumb tapping, index finger tapping, middle finger tapping, ring finger tapping and little finger tapping based on MMG-MIB approach. Two auto-event annotation algorithms are applied and evaluated for detecting the finger tapping frame. They are better for real-time processing system than the manual cutting approach although the latter one is more accurate. Based on the analysis of MMG frame, the wavelet packet transform (WPT) coefficients are calculated and compressed as the feature, followed by a feature selection method which is able to improve the performance of only one kind of feature. Finally, three popular classifiers including NBC, KNN, and SVM are applied and evaluated by their performances on finger gesture recognition.

In the current stage, only healthy participants are invited to attend our experiments to prove the feasibility of recognition for individual finger movement. For stroke patients, the situation is extremely complicated. The applicable subjects for the experiments of the proposed system are limited to the group of who is capable of controlling the contractions of the forearm muscle. Finding such subjects and to some extent further studying how the proposed system can be adaptive and scalable for subjects with different degrees of the capability of controlling forearm muscle is a much broader topic demanding more effort which will be further explored in future work.

2. SYSTEM ARCHITECTURE

The proposed system relies on two MMG signal channels for identifying movements of five fingers. The architect of the proposed FGR system is shown in **Figure 1**. The first module contains the MMG acquisition system obtaining the two-channel MMG signals from the forearm muscle by the inertial sensor. Then, the detected MMG signals are going through a band-pass filter to reduce noise distortion. The third stage of the system consists of the tapping event detection (TED) algorithm to extract the MMG signal segments with the muscle activity information. The next stage is the feature extraction process on the obtained MMG signal segments. Finally, different classifiers including SVM, KNN, and NBC are built for the recognition purpose based on the features extracted. A detailed description of each sub-system is given below.

2.1. MMG Acquisition Module

The MMG signal is detected by the inertial sensor MUP6050 (InvenSense, USA), which is capable of capturing the information of acceleration and angular velocity. In our system, the acceleration signal from its Z-axis is recorded for FGR. Since Z-axis is perpendicular to the skin surface, it is the most important direction for studying the motions of muscular contraction (24). As mentioned previously, the MMG signals measuring

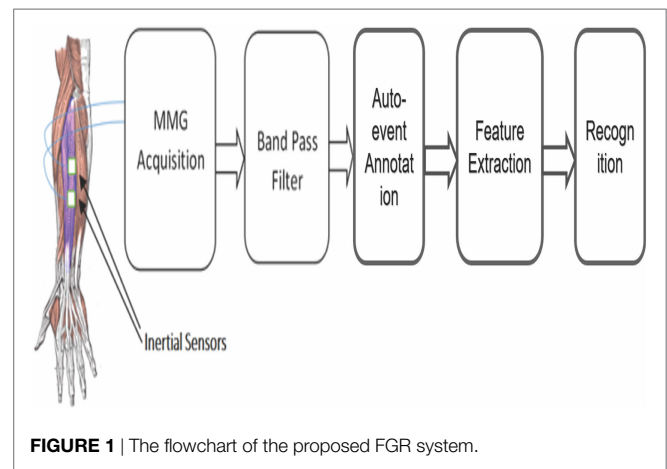


FIGURE 1 | The flowchart of the proposed FGR system.

the mechanical activities are characterized by low-frequency vibrations (<50 Hz) (32). The sampling frequency is chosen to be 1 kHz, and the cutoff frequency is chosen to be 200 Hz for the embedded analog to digital converter and low-pass filter in MUP6050, respectively. Since the contractions of the extensor digitorum muscle are the main muscle activities when finger taps, the MMG signal is detected from the belly of extensor digitorum muscle. Two sensors are attached on the skin surface within the above mentioned area to have a multichannel understanding of the extensor digitorum muscle. This is the reason why we have two-channel MMG signal measurements.

The relatively high movement artifact is the major deficiency of MMG acquisition, which can significantly degrade the recognition accuracy. Thus, all the subjects were asked to put their arms on the desk to keep their forearms as motionless as possible during the experiments.

2.2. Band-Pass Filtering Module

As reported in Ref. (17, 33), the predominant power in MMG signals detected is within 10–22 Hz, and the acoustic frequency contents of MMG signal produced by contracting muscles are within 5–50 Hz (32). Thus, a 4th-order band-pass Butterworth filter in the range of 0.1–50 Hz is used in preprocessing to retain the useful information and reduce noises.

2.3. Auto-Event Annotation Module

The MMG signals after band-pass filters need to go through our auto-event annotation process. Auto-event annotation is designed to detect the finger tapping event and extract the activity segment in this article. Thus it is also called tapping event detection (TED) in the following. This module is also considered as the first stage of finger gesture recognition. In the literature, the MMG signal segments are obtained by manual cut (14) or predefined time slot when the participants are asked to make each movement (24, 29). These approaches introduce the inconvenience in either the data preprocessing stage or in the data acquisition process. Furthermore, they do not have the capability of real-time processing. To make the proposed FGR system with real-time processing capability, an automatic TED algorithm is required. In our system,

two different TED algorithms are tested. The first one is the root mean square (RMS) algorithm (30), and the other algorithm is the difference-template filtering (DTF). Their performances are compared in our experiments.

2.3.1. Root Mean Square (RMS) Algorithm

The RMS algorithm has been used in many MMG signal processing applications such as the assessment of muscle function (32, 34) and automatic detection of muscle activities (35). It can be described by the following equation:

$$Z_r(t) = \begin{cases} 1, & \text{if } (g_r[t] > TH_r) \\ 0, & \text{otherwise} \end{cases}, \quad (1)$$

where t is the time index. $Z_r[t]$ is the output of this algorithm. Its value will be set to 1 when the finger motion event is detected, otherwise it will be zero. $g_r[t]$ and TH_r are defined by the following equations:

$$g_r[t] = \sqrt{\frac{1}{W} \sum_{i=t}^{t+W} x[i]^2}, \quad (2)$$

$$TH_r = \alpha_r * \frac{\sum_{i=T_{s1}}^{T_{s2}} g_r[i]}{T_{s2} - T_{s1}}, \quad (3)$$

where the $g_r[t]$ is the RMS value from $x[t]$ to $x[t + W]$ and the parameter W is the window size. As reported in Ref. (35), an appropriate analysis window size for MMG signal is from 100 to 400 ms. In our system, the window size is fixed to 400 ms, which has a superior performance in contrast to other smaller sizes. The parameters, T_{s1} and T_{s2} , are the start and end points of time during which the forearm muscle group of participants are required to keep relaxed. The parameter, α_r , is the threshold scaling factor which is set to 2.15 in our system.

2.3.2. Difference-Template Filtering (DTF)

A DTF-based TED algorithm is proposed to detect the muscle activities in our system. The vector $D = [-1, -1, -1, -1, -1, -1, -1, -1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1]$ is designed as the difference template for convolution with MMG signal. The difference template-based TED can be described by equations (4)–(6):

$$Z_d(t) = \begin{cases} 1, & \text{if } (g_d[t] > TH_d) \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

where $g_d[t]$ is defined by equation (5) and the parameter TH_d is defined by equation (6).

$$g_d = D \otimes x. \quad (5)$$

The \otimes is the convolution operator, and x is the input MMG signal.

$$TH_d = \alpha_d * \frac{\sum_{i=T_{s1}}^{T_{s2}} g_d[i]}{T_{s2} - T_{s1}}. \quad (6)$$

The parameter α_d in above equation is the threshold scaling factor, which is set to 2 in our system. T_{s1} and T_{s2} have same definitions as mentioned above.

2.4. Feature Extraction Module

Feature extraction is an extremely important process in our system as the recognition accuracy is heavily dependent on whether indicative and relevant features to finger movements can be found. In our system, the MMG signals are transformed to wavelet domain for wavelet packet transform (WPT) coefficients. Then singular value decomposition (SVD) is applied to reduce the dimension of the coefficients for computational efficiency. Then, WPT features will be sent to our classification module for the recognition purpose.

2.4.1. Feature Matrix Extraction

Wavelet transform (WT) is proposed for multi-resolution analysis developed from the Fourier transform. It is able to represent the local signal characteristics in time-frequency domain. However, its resolution is decreasing when the signal frequency is increasing. In other words, the resolution in high-frequency region is very poor in the WT analysis. As an extension of the standard WT, WPT is able to provide an arbitrary time-frequency resolution (36). Therefore, WPT is used to extract the signal characteristics from the MMG signal here for analyzing the muscle activities.

The **Figure 2** shows a 5-level wavelet packet decomposition which is also known as optimal subband tree structuring. Each node of the tree is marked as $x_j^p(t)$, where j is the level with the range from 1 to 5, and $p = 1, \dots, 2^j$ is the number of the packet in the j th level. Each node, $x_j^p(t)$, can be decomposed to two nodes, $x_{j+1}^{2p-1}(t)$ and $x_{j+1}^{2p}(t)$, which is described in equation (7). The reconstructions are given in equations (8) and (9) as below:

$$X_j^p = X_{j+1}^{2p-1} \oplus X_{j+1}^{2p}, \quad (7)$$

$$x_j^{2p-1}(t) = \sqrt{2} \sum h(n)x_{j-1}^p(n - 2t), \quad (8)$$

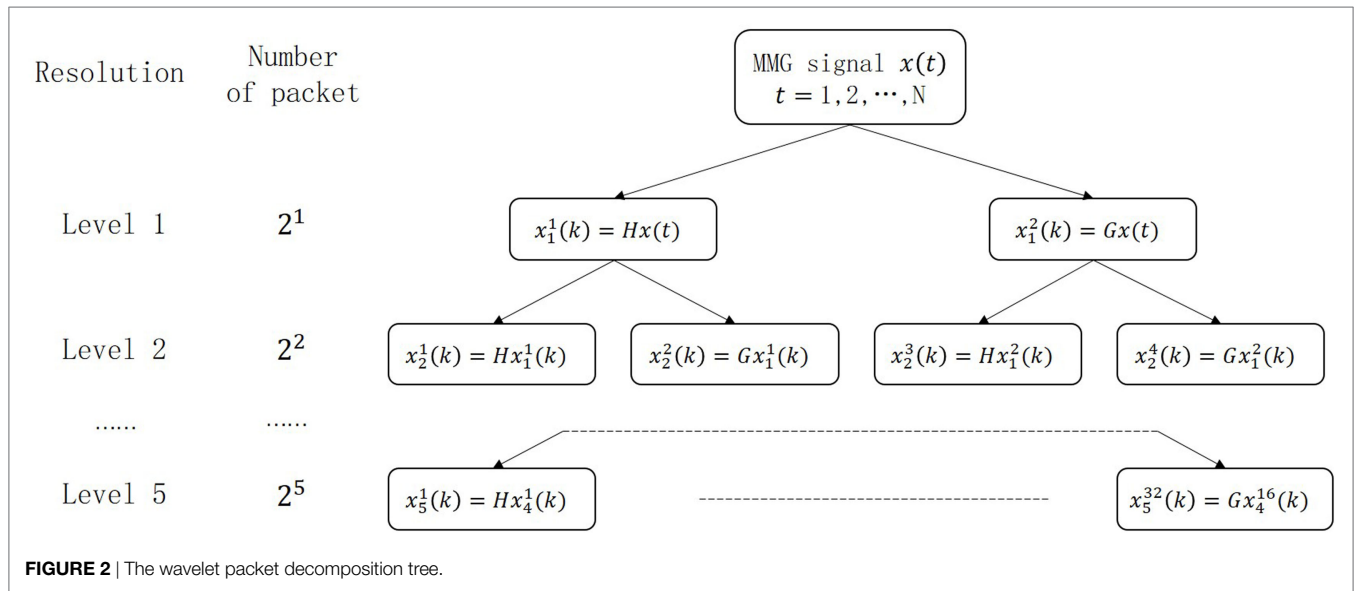
$$x_j^{2p}(t) = \sqrt{2} \sum g(n)x_{j-1}^p(n - 2t), \quad (9)$$

where n and t are both time indexes, $h(n)$ in equation (8) is the scaling function, $g(n)$ in equation (9) is the wavelet filter, and the details of $h(n)$ and $g(n)$ can be referred to Ref. (37). In our system, every MMG signal segment is decomposed into the 5th level shown in **Figure 2**. Each node of the 5th level wavelet is a column vector containing $LN/32$ coefficients, where LN is the size of the processed MMG signal. All the nodes of the 5th level construct a $(LN/32) \times 32$ feature matrix, X , shown in the following equation:

$$X = [x_5^1(k), x_5^2(k), \dots, x_5^{32}(k)] (k = 1, 2, 3, \dots, LN/32). \quad (10)$$

2.4.2. Feature Compression

WPT normally generates a large number of features for classification. To reduce the workload of following classifier and prompt the proposed system to be real-time processing, the SVD technique is used in our system to extract the new feature from the X for reducing dimension. As a matrix factorization technique, SVD is proven to be reliable and efficient (38, 39). The feature matrix, $X_{L \times M}$



($L = LN/32, M = 32$), given by equation (10) is decomposed by the SVD technique given by the following equation:

$$X_{L \times M} = U_{L \times L} \Lambda_{L \times M} V_{M \times M}^T, \tag{11}$$

where U and V are the $L \times L$ and $M \times M$ orthogonal matrixes, and Λ is a $L \times M$ nonnegative diagonal matrix described as follows:

$$\Lambda = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}, S = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r), \tag{12}$$

where r is the rank of the matrix, and $\sigma_1, \sigma_2, \dots, \sigma_r$ are the singular values of matrix X , which are used as the extracted features. Since ($LN/32$) is normally larger than 32 in this article, the extracted features are with the size of 32.

2.4.3. Feature Selection

There are 32 features from each MMG channel after the feature compression stage. However, these features have different degrees of relevance for the hand-motion classification. The MMG signal obtained by the MMG acquisition system contains both the mechanical signal generated by the hand motions and noises from different sources such as electrical noise and the vibration noise from the surroundings. The features containing the noise usually have an adverse effect for the recognition. It should be noted that a larger number of SVD features do not always lead to more effective classification results (24). Therefore, a feature ranking algorithm is required to select the most relevant features to further improve the recognition results.

In the proposed system, the methodology proposed in Ref. (40) is utilized for features ranking. The parameter F -value is calculated from the features extracted by WPT-SVD algorithm. For each feature, the bigger value of the F indicates higher importance for classification. The details of how to calculate the F are described following.

First of all, the average distance of each kind of feature from the same finger gesture is calculated by the following equation:

$$d_{i,j} = \frac{1}{N \times (N-1)} \sum_{m,n=1}^N |p_{i,j}(m) - p_{i,j}(n)|; \tag{13}$$

$(m, n = 1, 2, \dots, N, m \neq n),$

where N is the number of the samples of the same finger gestures, $p_{i,j}(m)$ is the m th sample of the i th feature of the j th gesture. For each feature, its average distance between 5 finger gestures can be defined as follows:

$$d_{ai} = \frac{1}{M} \sum_{j=1}^M d_{i,j}. \tag{14}$$

The ai is the number of the gestures. The value of M is 5 since there are five gestures needed to be recognized. Then, we calculate the average value of each feature in the same gesture as follows:

$$p_{ai,j} = \frac{1}{N} \sum_{n=1}^N p_{i,j}(n) (n = 1, 2, \dots, N). \tag{15}$$

The average distance of each feature between different finger gestures can be defined by the following equation:

$$d'_{ai,j} = \frac{1}{M \times (M-1)} \sum_{m,n=1}^M |p_{ai,m} - p_{ai,n}| \tag{16}$$

$(m, n = 1, 2, \dots, N; m \neq n).$

Finally, the F -values can be calculated as follows: $F = d_{ai}/d'_{ai}$. Following this algorithm, the 32 features are ranked in descending order according to its F -value.

2.5. Classification Module

The last module of our FGR system is the classifier to recognize the finger movement based on the features selected. In many

reported MMG-based hand-gesture recognition systems, various pattern recognition algorithms were used such as linear classifier (24), multilayer perception (26), and neural networks (21). In our system, the classifiers, NBC, KNN, and SVM are selected, and the performance of three algorithms is compared.

2.5.1. NBC

NBC, as a subclass of Bayes classification algorithm, is widely applied in the pattern recognition because of its simplicity and effectiveness. In some applications, its performance is comparable with other classifiers with increased computational complexity (41). Naive Bayes classification model assumes that each feature is independent for the classification. Although this unrealistic assumption limits its scope of applications, the time and space complexities are reduced. Therefore, the NBC is still a popular choice in various applications (42, 43). The implementation of NBC is based on the reference (44, 45).

2.5.2. KNN

KNN is a classification model based on statistical analysis. The KNN algorithm classifies the sample according to the class of its k -nearest neighbor samples in the feature space. And the sample is assigned to the class where most of its neighbor samples belong to. KNN is a simple classifier without priori statistical knowledge. Thus, it is widely used in applications with non-normal or unknown sample distributions (46). However, its performance will be degraded when the number of samples is large (47). The implementation of KNN is based on the reference (48).

2.5.3. SVM

SVM is a machine learning algorithm based on the statistical learning theory. It is good to deal with the situations when small samples are available and high dimension pattern recognition task is given (49). The libSVM3.12 package (50) is adopted for programming implementation in our system. The polynomial kernel function, $G(x)$, given by equation (17) is utilized in SVM classifier,

$$G(x) = (\gamma \times x \times z + \text{coef}0)^{\text{degree}}, \quad (17)$$

where x is the vector of input features, z is the support vector produced by SVM after training, and the parameters $\text{coef}0$, γ , and degree are set to 0, 1, and 3, respectively, in our system.

3. SYSTEM CHARACTERIZATION RESULTS

In this section, we will describe the detailed experimental protocol for our FGR system characterizations. The characterization results are presented in this section as well with discussions. First of all, the performance of two TED algorithms introduced is compared with the manual event detection method. After that, the effectiveness of the features ranking and the performances of different classifiers are presented and discussed. At the end of this section, possible factors affecting the recognition results such as power of grip (PG) and body mass index (BMI) are discussed.

3.1. Participants

To fully characterize our system, twelve healthy participants (8 males and 4 females, 1 left-hander and 11 right-handers, age: 23 ± 3.21 , height: 164.2 ± 7.65 cm, weight 66.2 ± 16.51 kg) participate in our experiments voluntarily. No medical history of neuromuscular disorders is reported. All the participants understand and agree with the experiment protocol before joining in the test for MMG signal acquisition.

3.2. Experimental Procedure

All participants are instructed to use five fingers of their dominant hand to tap, including thumb tapping, index finger tapping, middle finger tapping, ring finger tapping, and little finger tapping. Following a metronome with 30 beats per minute, every participant is asked to tap five fingers one by one from thumb to little finger, and each finger gesture is repeated five times before moving to next finger. The above process is repeated five times with 5 min interval to avoid the unreliable samples due to large number of consecutive tapings on the same finger. Therefore, there are totally 125 finger-gesture movements (5 finger gestures \times 5 times \times 5 rounds) for every participant. A total of 25 movements on each finger are obtained from tapping with either short time interval or long time interval which increases the variety of data set.

3.3. Characterization of Auto-Event Annotation Module

The auto-event annotation module in our system is tested. In details, the performance of two automatic TED algorithms is compared with the manual approach which is supposed to be optimal. The false detection events ratio (FDER) of the RMS-based and the DTF-based automatic TED algorithms are calculated which relates to the error rate of the muscle activity event detection. In addition, the accuracies of the MMG signal segmentation by different methods including DTF, RMS, and manual cutting are calculated and discussed.

3.3.1. The Error Rate of Automatic TED

In this experiment, the false detection events in two scenarios are measured. The first one is the false negative event shown on the top of **Figure 3**. The muscular contraction event occurs but the TED algorithm cannot detect it. In this situation, the proposed system is not able to recognize finger-gesture movements. The other one is the false positive event shown in the bottom of **Figure 3**. The muscular contraction events can be detected by the TED algorithm while this signal segment contains irrelevant information. As a result, the wrong recognition results will be obtained by the proposed system.

FDER is a suitable indicator to measure and compare the performance between the RMS- and DTF-based automatic TED algorithms. The FDER is defined by the equation (18):

$$FDER = \frac{N_{FP} + N_{FN}}{N_{all}}, \quad (18)$$

where the N_{all} is the number of all events detected by the algorithm. The N_{FP} and N_{FN} are the numbers of the false positive event and the false negative event, respectively.

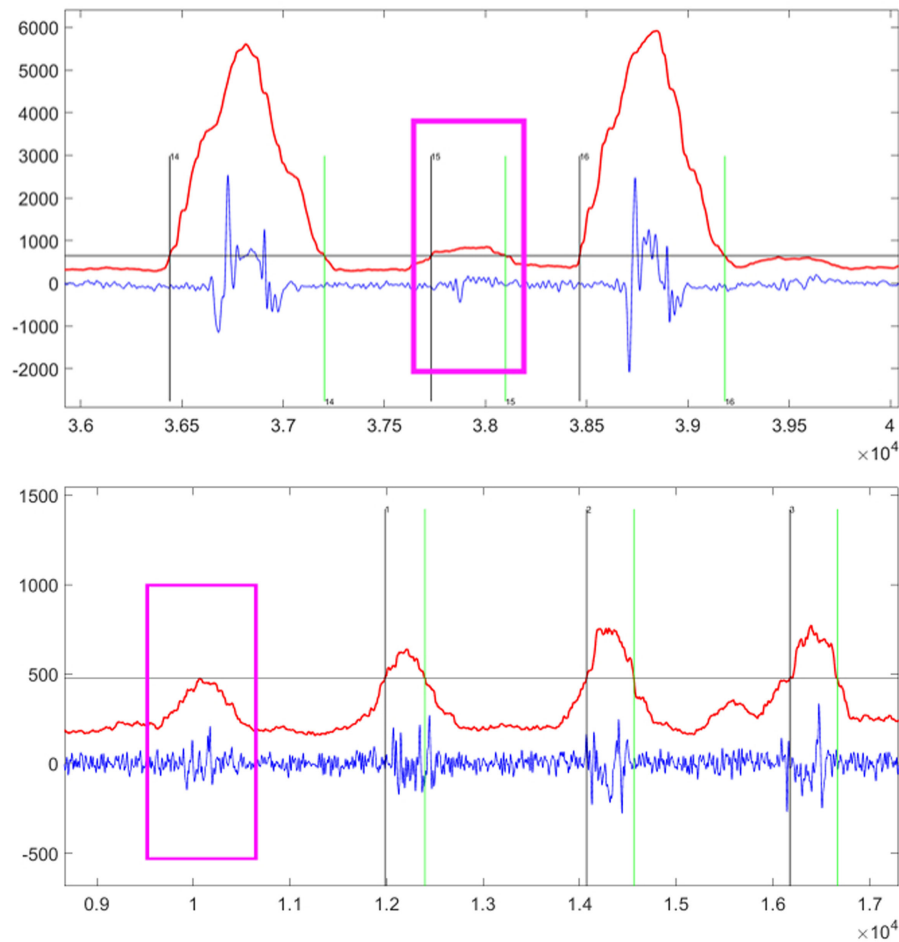


FIGURE 3 | The examples of false negative event (top) and false positive event (bottom) highlighted by the pink windows. The vertical axis title is “Amplitude” and the horizontal axis title is “Number of Sampling Points.” The sampling rate is 1 kHz, and then the above chart shows segments lasting approximately 4 s (top) and 8 s (bottom). Within this chart, the blue line is the original MMG signal, the red line is the automatic TED result, and the horizontal gray line is the threshold calculated by equation (3) or equation (6). Thus, the vertical black line indicates the start point of detected event after which the red line goes beyond the gray line, while the vertical green line indicates the end point of detected event after which the red line goes under the gray line.

In our experiment, all the data are processed separately by the RMS- and DTF-based automatic TED algorithms. The FDER values are 1.5 and 2.2%, respectively, for RMS-based method and the DTF-based method.

3.3.2. The Accuracy of MMG Segmentation

There are two tasks for the TED algorithms. The first task is to detect the fingers activity events correctly. The second task is to extract the MMG signal segment accurately. The FDER is the index to measure the performance of the first task. Next, we discuss how to measure the performance of the second task. There are two scenarios when the MMG signal segment is extracted inaccurately. As shown in **Figure 4**, the segment of the MMG signal shown in the top is very weak where the starting point of the event is too close to the end point, resulting in loss of the key information. On the other hand, the sub-figure in the bottom part of **Figure 4** shows that the MMG signals contain some noises which may introduce some interferences. These inaccurate segmentations are the two factors

that can degrade the recognition accuracy of the proposed FGR system.

Currently, there is no quantitative index to directly measure the quality of the MMG signal segment extraction. As our system focuses on the final finger-gesture recognition accuracy, we only compare the performance of RMS- and DTF-based TED algorithms with manual segmentation in achieving the final recognition rate. Among that, manual segmentation is supposed to be the accurate and optimal method for extracting the finger tapping event. In this comparison experiment, only the first 4 features after feature ranking process are used, and the NBC classifier is used. The results are reported in the **Figure 5**. This figure clearly shows that the manual segmentation has the best performance for the proposed FGR system, followed by the RMS- and DTF-based automatic TED algorithms.

3.3.3. Discussion of TED Algorithm Evaluation

From the results in Section 3.3.1 and 3.3.2, the manual segmentation has slightly better performance compared with

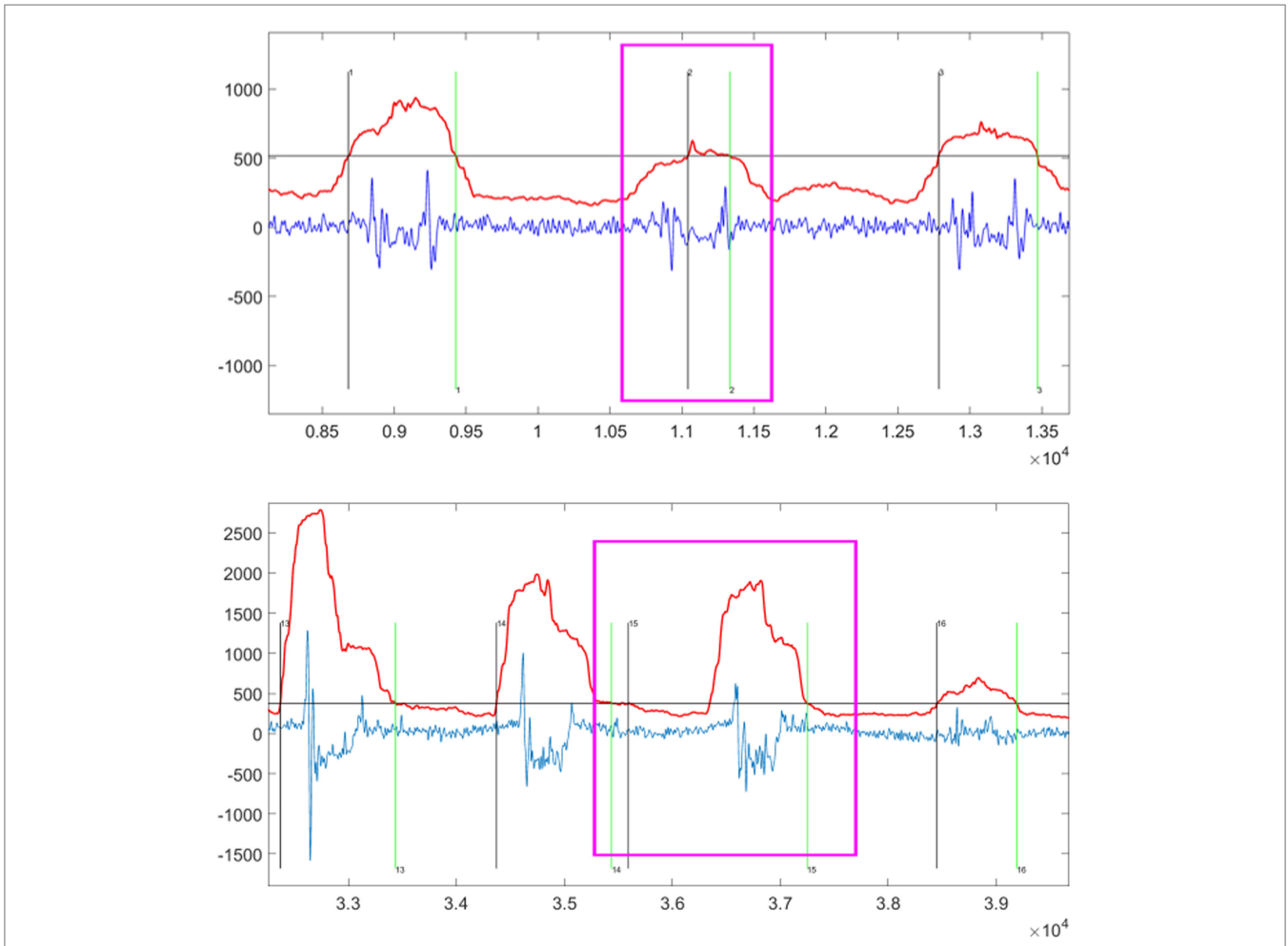


FIGURE 4 | The examples of inaccurate extraction by automatic TED Algorithms highlighted by the pink windows. The vertical axis title is “Amplitude” and the horizontal axis title is “Number of Sampling Points.” The sampling rate is 1 kHz, and then the above chart shows segments lasting approximately 5 s (top) and 6 s (bottom). Within this chart, the blue line is the original MMG signal, the red line is the automatic TED result, and the horizontal gray line is the threshold calculated by equation (3) or equation (6). Thus, the vertical black line indicates the start point of detected event after which the red line goes beyond the gray line, while the vertical green line indicates the end point of detected event after which the red line goes under the gray line.

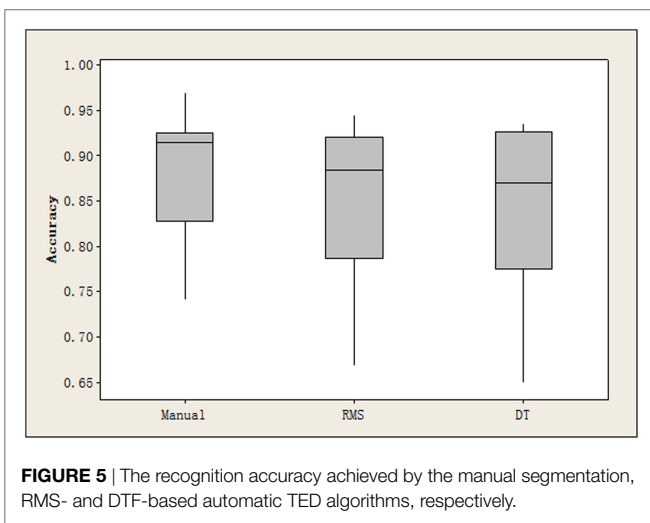
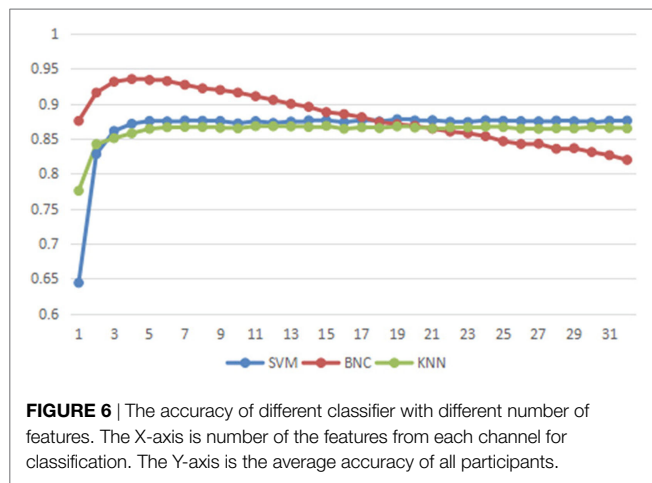


FIGURE 5 | The recognition accuracy achieved by the manual segmentation, RMS- and DTF-based automatic TED algorithms, respectively.

two automatic TED algorithms. If the performance between two automatic TED algorithms is compared, the RMS-based approach has a slightly better performance. However, the DTF-based approach has less computational complexity compared with the RMS-based method. Thus, the DTF-based approach is preferred in some low cost and real-time processing applications.

3.4. Classification Evaluation

The recognition accuracies of the classifiers SVM, KNN, and NBC are computed, and the results are compared and discussed. There are in total of 64 features from two-channel recorded MMG signals and 32 features are extracted from each channel. During the system characterization evaluation, the appropriate number of features used for recognition is also studied. Different numbers of features are selected by the ranking algorithm and the resulting recognition accuracies are compared. In this section, the



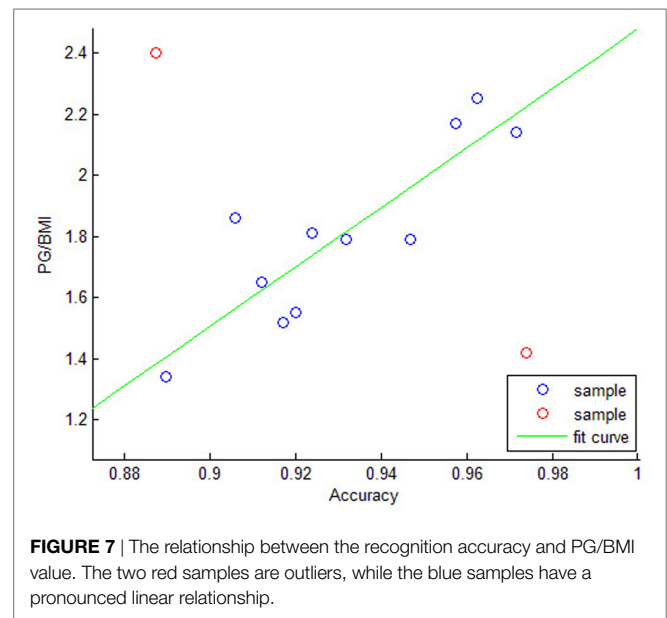
manual segmentation for TED is selected to reduce unnecessary interference.

The experimental results are shown in **Figure 6**. The accuracy rate of identification is up to 94.0%, and the average accuracy rate is 87.9% by the NBC. The highest recognition results (94%) in NBC classifier are obtained when the first 4 features of each channel are used. The accuracy is reduced with the increased number of features as NBC assumes that every feature is dependent. In other words, each feature has the same importance impact for the NBC classifier. According to the ranking algorithm introduced in Section 2.4.3, the feature with lower rank means that it is less useful because of more noises involved. Thus, when less irrelevant features are used in the classification process, more noises are introduced, resulting in less accuracy of the NBC classifier. In addition, this result also shows that the feature ranking algorithm is critical for the NBC-based classifier, because it can remove the irrelevant features and improve the NBC performance.

The recognition accuracies from the KNN and SVM classifiers are improved slightly when the number of features per channel is larger than 3. The possible reasons of this observation are explained as follows. KNN is a classification algorithm based on the distance of the feature space. In that case, it is insensitive to the weak features. Therefore, the classification accuracy is not degraded when the number of weak features is increased (46). The SVM classifier aims to optimize the performance based on all features, which definitely includes the weak features. The best linear combination of all features is chosen by the SVM classification. Thus, the SVM performance is becoming slightly better when more features are used in **Figure 6**. In addition, the performance of SVM is also dependent on the used kernel function. The proposed system only considers the polynomial kernel and nonlinear kernel functions may improve the performance which will be explored in the future work.

3.5. Influencing Factors

As the MMG signals are produced by the muscle and transmitted by the soft tissue (14), intuitively, our FGR system might have better performance on participants who have well developed forearm muscles and thinner hypodermic fat. To verify this hypothesis, the power of grip (PG) is measured as the index of muscle strength,



and the body mass index (BMI) is recorded as the index of hypodermic fat for each participant in this experiment. The recognition performance on each participant is then compared.

The **Figure 7** shows the relationship between (PG/BMI) and the recognition accuracy rate. The X-axis is the recognition accuracy based on every participant data where the NBC classifier with 8 features (4 features for each channel) is adopted for the FGR system. The Y-axis is the value of PG divided by BMI, namely (PG/BMI). There is a linear relationship between the (PG/BMI) and the recognition accuracy in our experiments, and the cross-correlation coefficient is 0.728. The above results show that the recognition accuracy rate of FGR system is proportional to the muscle strength, while has a inverse proportion to the hypodermic fat. Although more testing data are needed to reach a firm conclusion instead of the above results from 12 participants, it is highly possible that the recognition accuracy of FGR system and the value of (PG/BMI) have a linear relationship. The MMG-based FGR system may have higher recognition accuracy on participants with higher ratio of PG and BMI. This finding also suggests that the proposed FGR system users should take more exercises to strengthen the forearm muscles and reduce the hypodermic fat to well control the system. This system may be adopted in different applications such as prosthetic control.

4. CONCLUSION

In this article, a novel motion-intend-based finger gesture recognition system based on mechanomyogram (MMG) signal is presented to recognize each finger tapping. This is the first system capable of detecting every finger tapping. The architect and the design of the system are presented. A full system characterization is also evaluated and the recognition accuracies of the system are compared under different experimental settings. Our system is able to achieve up to 94% accuracy. How different factors such as PG and BMI of the participants affect the system performances are also discussed, and the obtained results suggest that the proposed

FGR system users should take more exercises to strengthen their muscles and reduce the hypodermic fat.

The proposed system focuses on identifying motion intentions instead of actual finger movements. It is therefore suitable for finger rehabilitation training of stroke patients. In some cases, when stroke patients try to move their fingers, they cannot find corresponding action. With the proposed system, we can provide guidance to the patients by correcting the effort made toward their intentions of finger movement and can further quantitatively measure the accuracy by exploring the enhancement of the system design. The quantified result is beneficial to further understanding of training intensity, contrasting training effect, and so on. Therefore, applying the proposed system to the stroke rehabilitation is worthy of continuous exploration and development.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Ethical Committee of Health Science Center, Shenzhen University with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the School of Biomedical Engineering, Health Science Center, Shenzhen University, China.

REFERENCES

1. Ibraheem NA, Khan R. Survey on various gesture recognition technologies and techniques. *Int J Comput Appl* (2012) 50(7):38–44. doi:10.5120/7786-0883
2. Brunelli D, Tadesse AM, Vodermayr B, Nowak M, Castellini C. Low-cost wearable multichannel surface EMG acquisition for prosthetic hand control. *Proceedings of the 6th International Workshop on Advances in Sensors and Interfaces*. Gallipoli: IEEE (2015). p. 94–9.
3. Beisteiner R, Windischberger C, Lanzenberger R, Edward V, Cunningham R, Erdler M, et al. Finger somatotopy in human motor cortex. *Neuroimage* (2001) 13(6):1016–26. doi:10.1006/nimg.2000.0737
4. Liao K, Xiao R, Gonzalez J, Ding L. Decoding individual finger movements from one hand using human EEG signals. *PLoS One* (2014) 9(1):e85192. doi:10.1371/journal.pone.0085192
5. Khan MJ, Hong MJ, Hong K-S. Decoding of four movement directions using hybrid NIRS-EEG brain-computer interface. *Front Hum Neurosci* (2014) 8:244. doi:10.3389/fnhum.2014.00244
6. Hotson G, McMullen DP, Fifer MS, Johannes MS, Katyal KD, Para MP, et al. Individual finger control of a modular prosthetic limb using high-density electrocorticography in a human subject. *J Neural Eng* (2016) 13(2):026017. doi:10.1088/1741-2560/13/2/026017
7. Nguyen H-D, Hong K-S, Shin Y-I. Bundled-optode method in functional near-infrared spectroscopy. *PLoS One* (2016) 11(10):e0165146. doi:10.1371/journal.pone.0165146
8. Hong K-S, Bhutta MR, Liu X, Shin Y-I. Classification of somatosensory cortex activities using fNIRS. *Behav Brain Res* (2017) 333:225–34. doi:10.1016/j.bbr.2017.06.034
9. Olman CA, Pickett KA, Schallmo M-P, Kimberley TJ. Selective bold responses to individual finger movement measured with fMRI at 3t. *Hum Brain Mapp* (2012) 33(7):1594–606. doi:10.1002/hbm.21310
10. Habermehl C, Holtze S, Steinbrink J, Koch SP, Obrig H, Mehnert J, et al. Somatosensory activation of two fingers can be discriminated with ultrahigh-density diffuse optical tomography. *Neuroimage* (2012) 59(4):3201–11. doi:10.1016/j.neuroimage.2011.11.062

AUTHOR CONTRIBUTIONS

All the authors listed in this article have participated in the following work: substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; drafting the work or revising it critically for important intellectual content; final approval of the version to be published; and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. HD and QH designed and implemented the simulation model, and prepared the manuscript; YZ built initial constructs and GD supervised the project; SC supervised the analysis and edited the manuscript; all the authors discussed the results and implications and commented on the manuscript at all stages.

FUNDING

This work is supported by the Shenzhen Science and Technology Innovation Council (JCYJ20160608173106220, KQTD2015033016104926) and Natural Science Foundation of SZU (827/000088). This work is also supported in part by the Science and Technology Planning Project of Guangdong Province (2015B020214007, 2016A020220002). SC is only financially supported by Institute of High Performance and Computing, Agency for Science, Technology and Research, Singapore.

11. Hong K-S, Khan MJ. Hybrid brain-computer interface techniques for improved classification accuracy and increased number of commands: a review. *Front Neurobot* (2017) 11:35. doi:10.3389/fnbot.2017.00035
12. Schenck JF. The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds. *Med Phys* (1996) 23(6):815–50. doi:10.1118/1.597854
13. Al-Mulla MR, Sepulveda F, Colley M. A review of non-invasive techniques to detect and predict localised muscle fatigue. *Sensors (Basel)* (2011) 11(4):3545–94. doi:10.3390/s110403545
14. Barry DT, Leonard J Jr, Gitter AJ, Ball RD. Acoustic myography as a control signal for an externally powered prosthesis. *Arch Phys Med Rehabil* (1986) 67(4):267–9.
15. Balbinot A, de Freitas JCR, Córrea DS. Use of inertial sensors as devices for upper limb motor monitoring exercises for motor rehabilitation. *Health Technol* (2015) 5(2):91–102. doi:10.1007/s12553-015-0110-6
16. Tosovic D, Seidl L, Ghebremedhin E, Brown MJ. Determining minimal stimulus intensity for mechanomyographic analysis. *J Electromyogr Kinesiol* (2015) 25(5):749–53. doi:10.1016/j.jelekin.2015.06.003
17. Orizio C, Perini R, Diemont B, Figini MM, Veicsteinas A. Spectral analysis of muscular sound during isometric contraction of biceps brachii. *J Appl Physiol* (1990) 68(2):508–12.
18. Posatskiy A, Chau T. Design and evaluation of a novel microphone-based mechanomyography sensor with cylindrical and conical acoustic chambers. *Med Eng Phys* (2012) 34(8):1184–90. doi:10.1016/j.medengphy.2011.12.007
19. Watakabe M, Mita K, Akataki K, Itoh Y. Mechanical behaviour of condenser microphone in mechanomyography. *Med Biol Eng Comput* (2001) 39(2):195–201. doi:10.1007/BF02344804
20. Islam MA, Sundaraj K, Ahmad R, Ahamed NU, Ali M. Mechanomyography sensor development, related signal processing, and applications: a systematic review. *IEEE Sens J* (2013) 13(7):2499–516. doi:10.1109/JSEN.2013.2255982
21. Alves-Kotzev NL. *Mechanomyography as an Access Pathway for Binary and Multifunction Control [PhD thesis]*. Citeseer. Toronto: Doctoral dissertation, University of Toronto (2010).

22. Silva J, Chau T, Goldenberg A. MMG-based multisensor data fusion for prosthesis control. *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. (Vol. 3), Cancun: IEEE (2003). p. 2909–12.
23. Ibitoye MO, Hamzaid NA, Zuniga JM, Wahab AKA. Mechanomyography and muscle function assessment: a review of current state and prospects. *Clin Biomech* (2014) 29(6):691–704. doi:10.1016/j.clinbiomech.2014.04.003
24. Xie H-B, Zheng Y-P, Guo J-Y. Classification of the mechanomyogram signal using a wavelet packet transform and singular value decomposition for multi-function prosthesis control. *Physiol Meas* (2009) 30(5):441. doi:10.1088/0967-3334/30/5/002
25. Alves N, Chau T. Uncovering patterns of forearm muscle activity using multi-channel mechanomyography. *J Electromyogr Kinesiol* (2010) 20(5):777–86. doi:10.1016/j.jelekin.2009.09.003
26. Sasidhar S, Panda SK, Xu J. A wavelet feature based mechanomyography classification system for a wearable rehabilitation system for the elderly. *Inclusive Society: Health and Wellbeing in the Community, and Care at Home*. Berlin, Heidelberg: Springer (2013). p. 45–52.
27. Hong-liu Y, Sheng-nan Z, Jia-hua H. MMG signal and its applications in prosthesis control. *Proceedings of the 4th International Convention on Rehabilitation Engineering & Assistive Technology*. Shanghai: Singapore Therapeutic, Assistive & Rehabilitative Technologies (START) Centre (2010). 58 p.
28. Grossman A, Silva J, Chau T. Functional mapping of multiple mechanomyographic signals to hand kinematics. *Proceedings of the Canadian Conference on Electrical and Computer Engineering*. (Vol. 1), Niagara Falls: IEEE (2004). p. 493–6.
29. Yu W, Kishi T, Acharya UR, Horiuchi Y, Gonzalez J. Finger motion classification by forearm skin surface vibration signals. *Open Med Inform J* (2010) 4:31. doi:10.2174/1874431101004020031
30. Saha DP. *Design of a Wearable Two-Dimensional Joystick as a Muscle-Machine Interface Using Mechanomyographic Signals [Master's thesis]*. Blacksburg: Virginia Polytechnic Institute and State University (2013).
31. Ding H, He Q, Zeng L, Zhou Y, Shen M, Dan G. Motion intent recognition of individual fingers based on mechanomyogram. *Pattern Recognit Lett* (2017) 88:41–8. doi:10.1016/j.patrec.2017.01.012
32. Silva J, Heim W, Chau T. A self-contained, mechanomyography-driven externally powered prosthesis. *Arch Phys Med Rehabil* (2005) 86(10):2066–70. doi:10.1016/j.apmr.2005.03.034
33. Rhatigan BA, Mylrea KC, Lonsdale E, Stern LZ. Investigation of sounds produced by healthy and diseased human muscular contraction. *IEEE Trans Biomed Eng* (1986) BME-33(10):967–71. doi:10.1109/TBME.1986.325668
34. Madeleine P, Farina D, Merletti R, Arendt-Nielsen L. Upper trapezius muscle mechanomyographic and electromyographic activity in humans during low force fatiguing and non-fatiguing contractions. *Eur J Appl Physiol* (2002) 87(4–5):327–36. doi:10.1007/s00421-002-0655-8
35. Alves N, Chau T. Automatic detection of muscle activity from mechanomyogram signals: a comparison of amplitude and wavelet-based methods. *Physiol Meas* (2010) 31(4):461. doi:10.1088/0967-3334/31/4/001
36. Sun Z, Chang C. Structural damage assessment based on wavelet packet transform. *J Struct Eng* (2002) 128(10):1354–61. doi:10.1061/(ASCE)0733-9445(2002)128:10(1354)
37. Walczak B, Massart D. Noise suppression and signal compression using the wavelet packet transform. *Chemom Intell Lab Syst* (1997) 36(2):81–94. doi:10.1016/S0169-7439(96)00077-9
38. Selvan S, Ramakrishnan S. SVD-based modeling for image texture classification using wavelet transformation. *IEEE Trans Image Process* (2007) 16(11):2688–96. doi:10.1109/TIP.2007.908082
39. Brenner MJ. Non-stationary dynamics data analysis with wavelet-svd filtering. *Mech Syst Signal Process* (2003) 17(4):765–86. doi:10.1006/mssp.2002.1512
40. Yang B-S, Han T, An JL. ART-KOHONEN neural network for fault diagnosis of rotating machinery. *Mech Syst Signal Process* (2004) 18(3):645–57. doi:10.1016/S0888-3270(03)00073-6
41. Rish I. An empirical study of the naive Bayes classifier. *Proceedings of the IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*. (Vol. 3), New York: IBM (2001). p. 41–6.
42. Domingos P, Pazzani M. On the optimality of the simple Bayesian classifier under zero-one loss. *Mach Learn* (1997) 29(2–3):103–30. doi:10.1023/A:1007413511361
43. Hellerstein JL, Jayram T, Rish I. *Recognizing End-User Transactions in Performance Management*. Hawthorne, NY: IBM Thomas J. Watson Research Division (2000).
44. Mitchell TM. *Machine Learning*. New York: McGraw-Hill (1997).
45. Metsis V, Androutsopoulos I, Paliouras G. Spam filtering with naive Bayes— which naive Bayes? *CEAS*. (Vol. 17), Mountain View, CA (2006). p. 28–69.
46. Sang Y. *Research of Classification Algorithm Based on K-Nearest Neighbor*. Chongqing: Chongqing University (2009).
47. Angiulli F. Fast condensed nearest neighbor rule. *Proceedings of the 22nd International Conference on Machine Learning*. Bonn: ACM (2005). p. 25–32.
48. Friedman JH, Bentley JL, Finkel RA. An algorithm for finding best matches in logarithmic expected time. *ACM Trans Math Software* (1977) 3(3):209–26. doi:10.1145/355744.355745
49. Xuegong Z. Introduction to statistical learning theory and support vector machines. *Acta Autom Sinica* (2000) 26(1):32–42.
50. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* (2011) 2:27. doi:10.1145/1961189.1961199

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ding, He, Zhou, Dan and Cui. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.