A Multimodal Framework Based on Integration of Cortical and Muscular Activities for Decoding Human Intentions About Lower Limb Motions

Chengkun Cui, Gui-Bin Bian, *Member, IEEE*, Zeng-Guang Hou, *Senior Member, IEEE*, Jun Zhao, and Hao Zhou

***Abstract*—In this study, a multimodal fusion framework based on three different modal biosignals is developed to recognize human intentions related to lower limb multi-joint motions which commonly appear in daily life. Electroencephalogram (EEG), electromyogram (EMG) and mechanomyogram (MMG) signals were simultaneously recorded from twelve subjects while performing nine lower limb multi-joint motions. These multimodal data are used as the inputs of the fusion framework for identification of different motion intentions. Twelve fusion techniques are evaluated in this framework and a large number of comparative experiments are carried out. The results show that a support vector machine-based three-modal fusion scheme can achieve average accuracies of 98.61%, 97.78% and 96.85%, respectively, under three different data division forms. Furthermore, the relevant statistical tests reveal that this fusion scheme brings significant accuracy improvement in comparison with the cases of two-modal fusion or only a single modality. These promising results indicate the potential of the multimodal fusion framework for facilitating the future development of human-robot interaction for lower limb rehabilitation.**

***Index Terms*—Electroencephalogram (EEG), electromyogram (EMG), human-robot interaction, mechanomyogram (MMG), motion intention recognition, multimodal fusion.**

Manuscript received October 28, 2016; revised February 13, 2017; accepted April 4, 2017. Date of publication July 18, 2017; date of current version July 26, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61533016, Grant 61611130217, Grant 61603386, and Grant 61421004, in part by the Strategic Priority Research Program of CAS under Grant XDB02080000, in part by the International Innovation Team of CAS (20140491524), and in part by the Beijing Science and Technology Project (Z161100001516004). This paper was recommended by Associate Editor M. Sahin. *(Corresponding author: Zeng-Guang Hou.)*

C. Cui is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: cuichengkun2014@ia.ac.cn).

G.-B. Bian is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: guibin.bian@ia.ac.cn).

Z.-G. Hou is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and University of Chinese Academy of Sciences, Beijing 100049, China, and also with the CAS Center for Excellence in Brain Science and Intelligence Technology, Beijing 100190, China (e-mail: zengguang.hou@ia.ac.cn).

J. Zhao and H. Zhou are with the Department of Neurology, Beijing Bo'ai Hospital, China Rehabilitation Research Center, Beijing 100068, China (e-mail: zaojun@aliyun.com; halfsunny@foxmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBCAS.2017.2699189

I. Introduction

THERE have been an increasing number of biorobots designed for improving the quality of human life over the last decades. Some devices, like exoskeletons [1], [2] and prostheses [3]–[5] can provide effective help for paralyzed or disabled people. Up to now, existing biorobots are still facing a lot of challenges. One of the major problems is that most robots cannot sufficiently distinguish users' various motion intentions [6]. Therefore, it is very difficult for robots to make accurate adjustments in time according to users' intentions, which vastly limits their practical use.

To address this issue, many advanced human-robot interfaces (HRI) have been developed to capture human motion intentions. Some of them were designed based on the biosignals recorded from human body [7]–[16], such as electroencephalogram (EEG), electromyogram (EMG) and mechanomyogram (MMG), all of which contain rich information about limb motion. EEG is the measurement of brain electrophysiological activity from the scalp [17]. The EEG recorded from the motor cortex has a strong relationship with the movement control of human body. EMG is the recording of the electrical activity produced by a skeletal muscle, which can reflect the activation level of the muscle [18]. Specially, surface EMG acquired from the skin surface is more practical due to its non-invasiveness. MMG is a low frequency vibration recorded from the surface of the skin above a skeletal muscle, which embodies the mechanical activity of the muscle [19].

Recently, there has been a propulsion to fuse multimodal biosignals to achieve a more accurate and complete description of human motion intentions and to improve recognition performance of the systems that only use a single modal biosignal [6]. Since different modalities contain different human motion information, proper combination strategies can take the advantage of the complementarity between them to bring better results. For instance, on the one hand, by considering the influence of muscle weakness or fatigue, EEG can be used as an effective supplement for EMG-based or MMG-based approaches. On the other hand, EMG or MMG can contribute to relieve the mental fatigue which is from intensive concentration when using EEG-based devices. Moreover, the complementary information about electrical and mechanical activity of a muscle can be employed by effectively fusing EMG and MMG. As the state of the art is extensive, we narrow the focus on the multimodal fusion

methods based on EEG, EMG or MMG. Therefore, the fusion approaches using other signals (e.g., force signals) will not be considered in this paper.

Some researches focused on the motion intention recognition methods based on the fusion of EEG and EMG. Leeb *et al.* [20] designed a hybrid brain-computer interface to recognize two kinds of hand movements by using 16-channel EEG and 4-channel EMG. Two decision fusion algorithms, the average-rule-based fusion and Bayesian fusion were used to combine the outputs of EEG classifier and EMG classifier. For single EEG and single EMG, the recognition accuracies were 73% and 87%, respectively. While, the accuracy increased to 93% after using Bayesian fusion. Kiguchi and Hayashi [21] studied a judgment approach for the perception-assist of an upper-limb exoskeleton based on the combination of 256-channel EEG and 16-channel EMG. The results indicated that the accuracy was better improved by integrating EEG and EMG compared with just employing EMG in a two-class problem. Moreover, for a few subjects in this experiment, the outcomes after combination were enhanced greatly (e.g., from 60% to 85%). Kirchner and Tabie [22] studied the hand motion prediction by integrating EEG and EMG. 128-channel EEG and 4-channel EMG were synchronously recorded and two action events were labeled. The combination of multimodal data was implemented by rule-based decision fusion approaches (e.g., AND rule). The results indicated that the true positive rate or false positive rate was improved under different fusion rules. Xie *et al.* [23] applied a feature fusion method in identifying the knee extension and flexion, which combined local-mean-decomposition features and multi-scale-entropy features extracted from 14-channel EEG and 2-channel EMG, respectively. This study showed that the recognition rate by using feature fusion increased by about 5% than using EMG alone. Song *et al.* [24] investigated the multi-domain feature extraction of EEG and EMG to separately recognize three upper limb motion forms. Each motion forms contained two patterns. The features extracted from 2-channel EEG and 2-channel EMG were fused by nonnegative tensor factorization. The accuracies of the three motion forms by applying feature fusion were 92.14%, 89.78% and 93.75%, respectively, all of which were higher than the cases of only using a single modality. Shusharina *et al.* [25] developed a multifunctional neurodevice to measure 10-channel EEG and 6-channel EMG for recognizing hip flexion and extension. The linear discriminant analysis method was used for classification. Comparison experiments were conducted by only using EMG and combining EEG and EMG, respectively. The results showed that the fusion of the two signals could reach an accuracy of 86.8% while for using only EMG was 74.3%.

Several studies explored the motion intention identification approaches by combining EMG and MMG. Prociow *et al.* [26] presented a feature fusion method for EMG and MMG to recognize seven kinds of hand movements. Different time and frequency domain features extracted from 3-channel EMG and 3-channel MMG were combined as the input of a neural network. The results demonstrated that the recognition rate by using fused features increased by almost 2% than only using EMG features. Guo *et al.* [27] designed a multimodal signal acquisition system to recognize five hand gestures through 1-channel EMG and 1-channel MMG. The comparison and analysis were conducted on the time-frequency responses of EMG and MMG. Then, the features extracted from the two modalities were combined together. The accuracy was 88.7% for using fused features while for using EMG alone was 74.1%. Kurzynski *et al.* [28] applied a multi-classifier system based on the integration of EMG and MMG to identify users' intentions about hand motions. 8-channel EMG and 8-channel MMG were synchronously measured during five hand grasping actions. The results showed that the approach based on the combination of EMG and MMG obtained a higher accuracy of 95.1% compared with 87.0% for EMG alone and 60.6% for MMG alone. And in [29], they further improved the recognition performance by combining the meta-Bayes concept and Markov model into the multi-classifier system. The number of grasping actions increased to six and the accuracy increased to 95.8%. It is worth noting that few researches about motion intention recognition based on fusing EEG and MMG have been carried out at present.

According to the above literature, there exist some problems in the studies about two-modal fusion. Although the accuracy is enhanced compared with the case of only using a single modality, there are still ample rooms for accuracy improvement. Furthermore, more relevant experiments by combining EEG and MMG need to be conducted. Finally, and most importantly, there are still rare researches studying the multimodal-biosignal-fusion methods for identifying human intentions about lower limb multi-joint movements. Considering the complementarity between different modalities, we suppose that appropriate three-modal fusion schemes could achieve more accurate results than two-modal fusion. It is worth mentioning that literature [30] proposed a multi-classifier structure which can fuse EMG, MMG, EEG and can be considered for the application in prosthetic hands. However, this literature just presented a concept without further studies on specific methods and experiments. Moreover, to the best of our knowledge, the fusion framework based on the above three signals has never been designed to classify different lower limb movement patterns. Consequently, it is necessary to develop the technologies of combining EEG, EMG and MMG for decoding various motion intentions of human lower extremity.

In this paper, an enhanced multimodal fusion framework is proposed for recognizing human intentions about lower limb multi-joint movements. This framework, based on the combination of EEG, EMG and MMG, consists of two layers. In the first layer, three individual classifiers (i.e., EEG, EMG and MMG classifiers) are utilized to produce local decisions. The second layer consists of a decision fusion model which combines local decisions to achieve the final recognition result. In comparison with feature fusion methods in [23], [24], [26], [27], decision fusion has better scalability, more flexibility and is easier to be implemented. Compared with multi-classifier systems in [28], [29], decision fusion has a simpler structure and faster processing speed. Although some previous studies (e.g., [20], [22]) have explored the decision fusion methods of two modalities, most of them were limited to several simple fusion algorithms and many other methods have not been adopted.

The main contributions of this research include: 1) An enhanced multimodal fusion framework based on EEG, EMG and MMG is designed to decode human intentions about lower limb multi-joint motions. 2) Twelve decision fusion algorithms are evaluated in this framework. To the best of our knowledge, some of them have never been studied in the recognition of human motion intentions. 3) Extensive comparative experiments indicate that the best three-modal fusion scheme can obtain average accuracies of 98.61%, 97.78% and 96.85%, respectively, under three different data divisions for a nine-class problem, which are significantly higher than the cases of two-modal fusion or only a single modality.

The rest of this paper is organized as follows: Section II gives an insight into the collection and preprocessing of multimodal data. The multimodal fusion structure and implementation are depicted in Section III. Section IV presents the experimental results and discussions. Finally, Section V concludes the paper.

II. Data Collection and Preprocessing

The procedures of data collection and preprocessing are described in this section. Firstly, the subjects' basic information and the deployment of different sensors are depicted. Next, the experimental protocol and data preprocessing method are reported, respectively.

A. Subjects and Sensor Deployment

Twelve healthy subjects (six male and six female, age = 25.69 ±2.39 years, height = 166.83 ± 8.16 cm, weight = 57.13 ± 7.61 kg) were recruited in data recording. The dominant-legs of these subjects are right legs, which can be determined by the mobilizing functions of the legs when kicking a ball or hitting a target [31], [32]. This study was approved by the Ethics Committee of China Rehabilitation Research Center and an informed consent form was signed by every subject.

The brain activities were recorded by a Mindo-32 system (Mindo, Taiwan) which integrates a band-pass filter between 0.1 to 30 Hz. The sampling rate was set as 256 Hz. During the acquisition, each subject needed to wear an electrode cap with the reference electrode and ground electrode placed on the left and right mastoid processes, respectively. Although this system can simultaneously acquire 32 channels of EEG signals, only 9 channels over the motor cortex (FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4) were active and other channels were idle. The reason for this is twofold: Firstly, the EEG channels in this area have stronger relevance with human movements. Secondly, using fewer channels can be faster and more convenient for data processing and analysis. In order to guarantee the quality of the acquired EEG signals, all the subjects were asked to try to avoid head shaking when doing exercises.

The muscular activities were monitored by collecting EMG and MMG signals, respectively. Seven channels of surface EMG signals from right leg muscles were collected by the MyoScan sensors (Thought Technology Ltd., Canada) with a 50 Hz notch filter and a 10–100 Hz band-pass filter. The sampling frequency was also set as 256 Hz. The recorded muscles included the iliopsoas, gluteus maximus, rectus femoris, vastus medialis, long head of biceps femoris, tibialis anterior, medial head of gastrocnemius. The surface electrodes were placed based on the human anatomical locations [33]. Seven-channel MMG signals were captured from the same muscles mentioned above by using seven high sensitivity (1000 mV/g) accelerometers (CA-DR-1001, Sinocera Piezotronics Inc., China) fixed with medical tapes. The MMG signals were band-pass filtered (5–55 Hz), notch filtered (50 Hz) using the designed filter circuits and sampled at 256 Hz by a 16-bit acquisition module. To ensure the synchronization between different acquisition devices, digital clock signals for synchronous beginning and ending were generated by a computer that was used to store data.

Unluckily, there were very large noises in a subject's EEG signals, whose data should be ruled out before next step. Besides, another subject's data was removed because a sensor measuring MMG lost contact with the skin surface during the movement. Consequently, the further analysis is based on the remaining ten subjects' data.

B. Experimental Protocol

Different lower limb exercises are considered in this study. They are cycling, walking, and going up and down a step, which are the most common activities of daily living. All the exercises are carried out on the fitness equipments by the Technogym (Shanghai) International Trading Ltd., China. Each kind of exercise is subdivided into three different forms. As a result, there are a total of nine motion patterns which are more than that in the literature reviewed in Section I. In the present paper, the cycling exercise is conducted on a stationary exercise bicycle which can provide different exercise loads. Three load conditions, i.e., 20 W (lower load), 70 W (medium load) and 120 W (higher load), are considered, respectively. The pedaling rate is kept at 25 revolutions per minute. The walking exercise is performed on a treadmill which can offer a way to walk at an adjustable incline. The three subcategories of the walking exercise are discriminated by different slopes that are set as 0.5° (slight slope), 5.5° (moderate slope) and 10.5° (slightly steeper slope). The walking speed is maintained at 1.2 kilometres per hour. For going up and down a step, the three subcategories are separated by different movement frequencies. This exercise is conducted on a stepper which is a fitness equipment for up and down step training. The movement frequencies can be controlled by the equipment. The specific frequency settings are 15 times per minute (lower frequency), 20 times per minute (medium frequency) and 25 times per minute (higher frequency), respectively. The category names of different exercises are shown in Table I, which are labeled using the Arabic numerals from 1 to 9.

Before data collection, all the subjects were required to adapt to the different types of exercises and surrounding environment. For each subject, each motion mode was conducted for 60 seconds. After a 20-second adaptation period, multimodal data were simultaneously recorded during the following 40 seconds. In order to prevent muscle fatigue, the subjects were allowed to have a short rest for 2-3 minutes between different motion modes. The data collection scene during the cycling exercise from one subject is shown in Fig. 1.

TABLE I
CATEGORY NAMES OF DIFFERENT EXERCISES

Label	Category name
1	Cycling under a lower load
2	Cycling under a medium load
3	Cycling under a higher load
4	Walking on a slight slope
5	Walking on a moderate slope
6	Walking on a slightly steeper slope
7	Going up and down a step at lower frequency
8	Going up and down a step at medium frequency
9	Going up and down a step at higher frequency

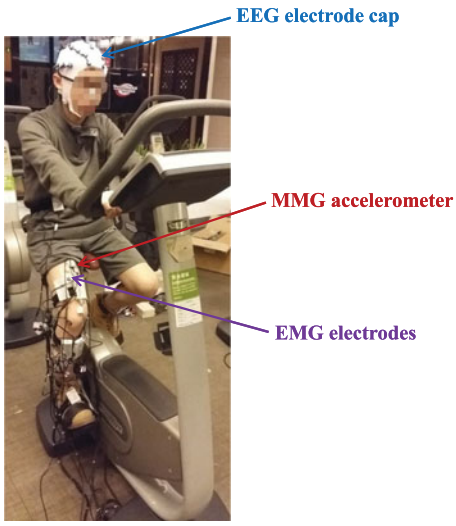

Fig. 1. Data collection scene during the cycling exercise.

In general, all the exercises were carried out at a relatively low speed, which were designed after considering several practical factors. These factors are mainly from the following three aspects. 1) The first purpose is to ensure the quality of the collected biosignals. Fast motion will cause large motion artifacts in the three biosignals, which will lead to bad effects on recognition results. Some sensors may shift or even fall off from the subject's leg during fast movements. 2) The second is the feasibility of performing multiple different motion forms. As these experiments involve a total of nine kinds of locomotion patterns, this design helps the subjects to accomplish these movement modes successfully without fatigue. 3) The third one is for the consideration of applications. The ultimate purpose of our study is to help and assist the patients with lower limb motor dysfunction, whose movements are usually slow.

It is noted that most parts of the data collection procedure were conducted under the guidance of two rehabilitation physicians.

C. Preprocessing

In order to continuously classify different intended movements, the data segmentation of biosignals is based on the windowing technique [34]. Normally, biosignals are divided into a series of fixed-length analysis windows whose size determines the amount of data for generating one class prediction [35]. In order to reduce the delay between adjacent analysis windows, the overlapping technique is commonly used [35], which means that the current window overlaps with the previous one in a certain proportion. An experimental sample in this study is defined as the data in an analysis window. The length of every analysis window is set as 62.5 milliseconds (ms) and the overlapping ratio is set as 50%. Therefore, for each modal biosignal of one subject, the number of experimental samples is 1,279 in every motion pattern. The total number of samples is 11,511 for each modal biosignal of one subject in a nine-class problem, which is a large amount of samples for this specific task.

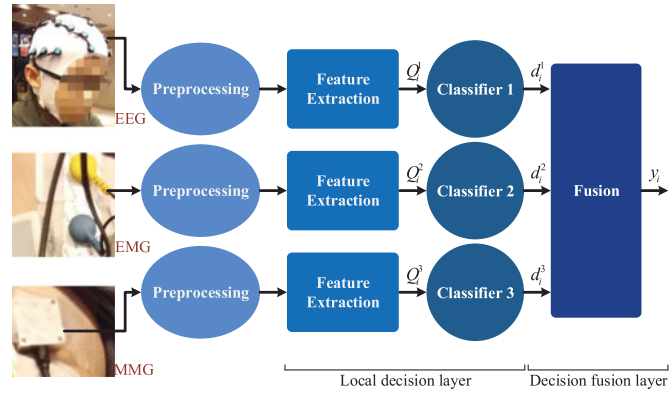

Fig. 2. Multimodal fusion framework for combining EEG, EMG and MMG signals.

III. DATA ANALYSIS

In this section, a series of data analysis methods are conducted on the collected biosignals to recognize different motion intentions of subjects' lower extremity. As a first step, an overall view of the multimodal fusion framework based on EEG, EMG and MMG is provided and discussed. Then, the specific implementation approaches are described in detail. Finally, the performance metric of the data analysis methods is given.

A. Multimodal Fusion Framework

The block diagram of the proposed multimodal fusion framework for combining EEG, EMG and MMG signals is shown in Fig. 2. After data collection and segmentation pre-processing, the biosignals in each analysis window are fed to this fusion framework which can be represented as a two-layer structure. The first one is the local decision layer, where three individual classifiers produce local decisions by using their respective features extracted from each analysis window. Next, the fusion algorithm integrates local decisions to achieve the final class prediction at the decision fusion layer. The above procedure can be formulized as follows. The input feature vector Q_i^s extracted from the ith analysis window of the sth modality can be defined as

$$Q_i^s = \left[\eta_1^s, \ldots, \eta_n^s, \ldots, \eta_{N_s}^s\right]^T, s = 1, 2, \ldots, S, i = 1, 2, \ldots, I \quad (1)$$

where η_n^s is the nth feature of Q_i^s, S is the number of modalities and I is the number of analysis windows. The output of the sth individual classifier is the local decision of the sth modality at

the ith analysis window, which can be expressed as

$$d_i^s = f^s(Q_i^s) \tag{2}$$

where f^s denotes the individual classifier of the sth modality and d_i^s is the local decision. Then the outputs of these individual classifiers are concatenated together to form a set of local decisions. At last, the fusion model predicts the final result for the ith analysis window by combining the local decisions. This procedure can be expressed as

$$y_i = g(d_i^1, \ldots, d_i^s, \ldots, d_i^S) \tag{3}$$

where g represents the fusion algorithm and y_i is the final predicted class for the ith analysis window. It is worth to note that this framework can be easily extended to the parallel usage of more than three modal signals.

By considering the influences of individual differences between subjects, the classification and fusion models are subject-specific, all of which need to be independently trained based on each subject's data. In this work, the training procedure is performed in the following steps. To start with, the training data of each subject are divided into two parts, i.e., individual-training data and fusion-training data. Next, different candidate individual classifiers are established by using the individual-training data. After that, these candidate classifiers are used to process the fusion-training data and three best individual classifiers are selected by comparing recognition accuracies. Finally, the outputs of three selected classifiers and the ground truth labels are fed to the fusion model to optimize tunable parameters. During the training procedure, 5-fold cross validation is applied to determine the hyper-parameters in some classification or fusion models. If there are no special instructions, the percentages of the individual-training data and fusion-training data for one subject are set to about 50% and 30%, respectively. The remaining approximately 20% of one subject's data are used as testing data to evaluate the recognition performance of this framework.

B. Implementation of the Two-Layer Structure

In order to implement the proposed two-layer structure, different feature extraction methods and individual classifiers are adopted at the local decision layer. Furthermore, various decision fusion algorithms are considered and applied in the decision fusion layer.

1) Local Decision Layer: The feature extraction approaches for EEG, EMG and MMG signals are described as follows.

For EEG signals, by considering the complex transient characteristics such as abrupt changes, spikes and drifts, the wavelet packet transform (WPT) is applied to extract features on the analysis window of EEG data. The Daubechies wavelet of order four is chosen to decompose EEG signals into three levels, and the features are extracted from the transformation coefficients of the third level. The logarithmic energy feature of the WPT coefficients can be denoted by [36]

$$\Gamma_{\Omega_{j,k}} = \log\left(\frac{\sum_{n=1}^{N_l}\left(w_{j,k,n}^T x\right)^2}{N_l}\right),\ j = 3, k = 0, 1, \ldots, 7 \tag{4}$$

where $\Gamma_{\Omega_{j,k}}$ is calculated by summing the squares of the transform coefficients for each entry in the node $\Omega_{j,k}$ from a binary wavelet packet tree and normalized by the number of coefficients in $\Omega_{j,k}$. Then the features on different nodes are concatenated to form the feature vector of EEG signals.

For EMG signals, the filter bank technique is applied to divide the data into different frequency bands. The filter bank is made up of three second-order Butterworth band-pass filters which can decompose the EMG signals into three frequency bands, namely 10–40 Hz, 40–70 Hz and 70–100 Hz. From different bands, different signal characteristics can be obtained. In this study, the mean absolute value (MAV) feature is extracted from the analysis window of each frequency band signal. The MAV feature has been widely investigated in EMG signal analysis [37], which is used here for its low computational complexity. The definition of MAV feature can be written as

$$\text{MAV} = \frac{1}{L}\sum_{m=1}^{L}|u_m| \tag{5}$$

where u_m is the mth data point of an EMG analysis window and L is the number of data points in an analysis window. Then the MAV features of different frequency bands are concatenated together to construct the EMG feature vector.

Similar to the processing of EMG, the MMG signals are decomposed into several frequency bands, and simple features are extracted from each frequency band. Specifically, three frequency bands of 5–20 Hz, 20–35 Hz, 35–55 Hz are separated by a filter bank composed of three second-order Butterworth band-pass filters. After that, the root mean square (RMS) feature on the analysis window from each frequency band is calculated, which is commonly used in the analysis of MMG signals [38]. The definition of RMS feature can be expressed as

$$\text{RMS} = \sqrt{\frac{1}{L}\sum_{m=1}^{L} v_m^2} \tag{6}$$

where v_m is the mth value of one MMG analysis window and L has been defined in (5). Next, the RMS features extracted from the three frequency bands are concatenated to build the MMG feature vector.

After feature extraction, the features from different modalities are fed to their respective individual classifiers. For each modality, eight commonly used linear and non-linear classifiers are regarded as candidate individual classification models in this study. Specifically, the eight classification models [39] are the support vector machine (SVM), neural network (NN), linear discriminant analysis (LDA) classifier, quadratic discriminant analysis (QDA) classifier, decision trees (DT), random forest (RF), naive Bayes (NB) classifier and k-nearest neighbor (KNN) classifier. In these experiments, the SVM model uses a radial basis kernel function, and the classification scores are transformed

to posterior probabilities by using Platt's method [40]. The NN model consists of a feedforward network structure with three layers. The activation functions of the hidden layer and output layer are sigmoid function and softmax function, respectively.

The output of the sth individual classifier can be a predicted probability vector or a predicted class, which depends on the subsequent fusion algorithm. Consequently, the local decision can be further represented as

$$p_i^s = f^s(Q_i^s) \text{ or } c_i^s = f^s(Q_i^s) \tag{7}$$

where Q_i^s and f^s have been defined in (1) and (2), respectively, p_i^s is the predicted probability vector and c_i^s is the predicted class. As it is mentioned in Section III-A, the individual-training data are fed to these candidate individual models for training (except KNN). After that, the classifier with the highest accuracy for each modality is selected by comparing the recognition performance on the fusion-training data. They are SVM for EEG, RF for EMG and NN for MMG. Then the local decisions produced by the three selected best classifiers on the fusion-training data are further employed to train the fusion model.

2) Decision Fusion Layer: At the decision fusion layer, various decision fusion algorithms are utilized to combine local decisions from three selected individual classifiers (SVM, RF and NN). According to literature [41], the involved decision fusion approaches can be divided into two types: rule-based approaches and classification-based approaches. The former combines multimodal data by pre-established rules while the latter is based on appropriate pattern classification techniques. In order to facilitate the description of different methods, some new variables are defined as follows. B_i is defined as the intended motion pattern of the ith analysis window, which need to be assigned to one of the J possible classes $(w_1, w_2, \ldots, w_J)$. Let $P(w_k|Q_i^s)$ denote the kth element of the predicted probability vector p_i^s which has been defined in (7).

For the rule-based fusion methods, three static rules are considered, including the max rule (MR), average rule (AR) and majority voting rule (MVR) [42]. The mathematical definitions of these rules can be denoted as follows. Max rule (MR): Assign $B_i \rightarrow w_k$ if

$$\max_{s=1}^{S} P(w_k|Q_i^s) = \max_{j=1}^{J} \max_{s=1}^{S} P(w_j|Q_i^s).$$

Average rule (AR): Assign $B_i \rightarrow w_k$ if

$$\frac{1}{S}\sum_{s=1}^{S} P(w_k|Q_i^s) = \max_{j=1}^{J}\left(\frac{1}{S}\sum_{s=1}^{S} P(w_j|Q_i^s)\right).$$

Majority voting rule (MVR): Assign $B_i \rightarrow w_k$ if

$$\sum_{s=1}^{S} \Phi_i^{ks} = \max_{j=1}^{J} \sum_{s=1}^{S} \Phi_i^{js}$$

where

$$\Phi_i^{zs} = \begin{cases} 1 & \text{if } P(w_z|Q_i^s) = \max_{j=1}^{J} P(w_j|Q_i^s) \\ 0 & \text{otherwise} \end{cases}.$$

It is noted that the MVR combines the predicted classes rather than the predicted probabilities, which is different from other fusion methods used in this paper. Moreover, the MVR can only be used under the constraint of at least three modalities, i.e., $S \geq 3$.

There are a total of nine classification-based fusion approaches involved in this study, which include the eight algorithms described in Section III-B1 and the Bayesian fusion (BF) algorithm [43]. However, these eight algorithms here are used to establish the fusion models which are different from the previous individual models. The inputs of the fusion models are the predicted probabilities from the three selected individual classifiers mentioned before. For convenience, the classification-based fusion methods can be uniformly expressed as

$$y_i = G_{cf}(p_i^1, \ldots, p_i^s, \ldots, p_i^S) \tag{8}$$

where G_{cf} represents the classification-based fusion model, y_i and p_i^s have been defined in (3) and (7), respectively.

To the best of our knowledge, in addition to the AR and BF methods whose performance have been reported in the reviewed literature, other decision fusion methods applied in this paper have not yet been studied in the recognition of human motion intentions.

C. Performance Metric

In this study, decoding different motion intentions of human lower extremity can be formulated as a multiclass classification problem. Consequently, the accuracy, one of the most commonly used performance metrics, is selected to evaluate the different data analysis methods. Mathematically, the accuracy is calculated as [39]

$$r = \frac{1}{N}\sum_{i=1}^{N} I(y_i = l_i) \tag{9}$$

where r is the accuracy, N is the number of samples, l_i is the ground truth label of the ith sample, y_i has been defined in (3) and $I(\cdot)$ is an indicator function. In this paper, the ground truth labels are marked according to the specific lower limb exercises performed during data recording and the specific experimental sample has been defined in Section II-C.

IV. Experimental Results and Discussions

To evaluate the practicability of the above multimodal framework, extensive experiments and relevant statistical analysis are conducted on the testing data. The classification performance by using only a single modality is reported at first. Then different combinations of multimodal signals are discussed and evaluated. Furthermore, recognition results of the above schemes are sufficiently compared in different aspects. Finally, the impact of the recognition performance improvement is discussed.

A. Recognition Results of Single-Modal Schemes

The classification accuracy based on a single modality is examined on the testing data. As the three best individual classifiers (SVM, RF and NN) are selected on the fusion-training data, the performance on the testing data can further validate the rationality of this selection. Moreover, the evaluation results on the single modality can be regarded as the comparison reference for that of multimodal fusion. The average classification

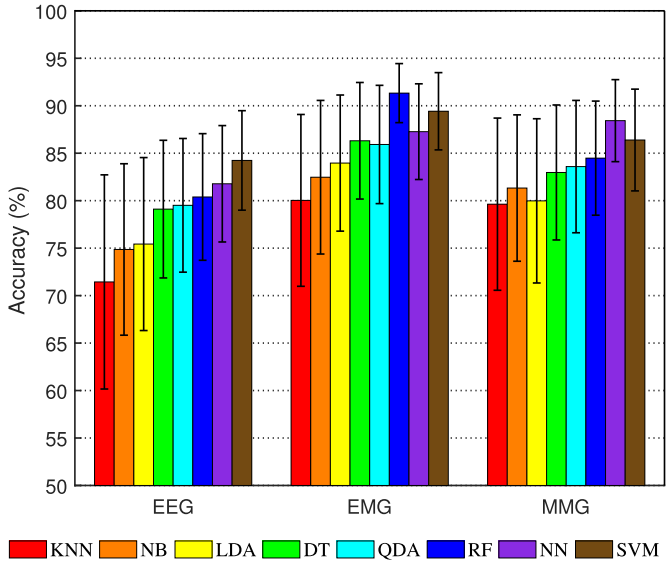

Fig. 3. Average classification accuracy and its standard deviation of single-modal schemes.

TABLE II
AVERAGE CLASSIFICATION ACCURACY (%) AND ITS STANDARD DEVIATION OF SINGLE-MODAL SCHEMES

Individual classifier	EEG only	EMG only	MMG only
KNN	71.44±11.28	80.03±9.05	79.63±9.07
NB	74.86±9.03	82.47±8.09	81.33±7.71
LDA	75.43±9.11	83.96±7.17	79.98±8.65
DT	79.11±7.25	86.31±6.14	82.97±7.11
QDA	79.51±7.04	85.92±6.23	83.59±6.97
RF	80.39±6.67	**91.33±3.11**	84.48±6.01
NN	81.78±6.13	87.27±5.04	**88.43±4.32**
SVM	**84.24±5.24**	89.42±4.07	86.39±5.36

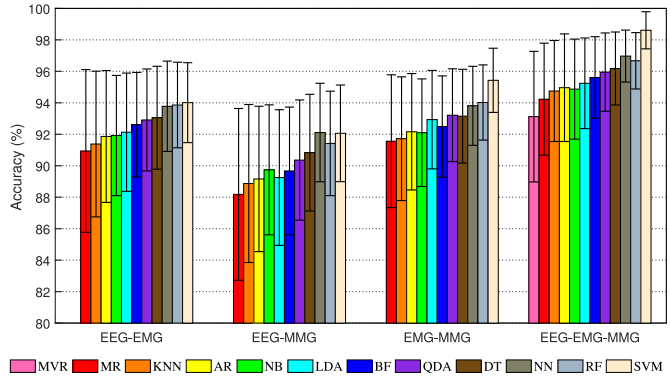

Fig. 4. Average classification accuracy and its standard deviation of multimodal fusion schemes.

TABLE III
AVERAGE CLASSIFICATION ACCURACY (%) AND ITS STANDARD DEVIATION OF MULTIMODAL FUSION SCHEMES

Fusion method	EEG-EMG	EEG-MMG	EMG-MMG	EEG-EMG-MMG
MVR	–	–	–	93.12±4.15
MR	90.94±5.17	88.18±5.46	91.56±4.22	94.23±3.56
KNN	91.38±4.63	88.87±5.02	91.72±3.93	94.75±3.21
AR	91.86±4.19	89.16±4.62	92.16±3.70	94.96±3.42
NB	91.92±3.82	89.74±4.13	92.10±3.42	94.87±3.18
LDA	92.13±3.76	89.25±4.31	92.93±3.13	95.24±2.88
BF	92.61±3.32	89.67±4.06	92.49±3.22	95.61±2.59
QDA	92.91±3.24	90.36±3.82	93.21±2.95	95.95±2.49
DT	93.05±3.27	90.83±3.71	93.15±2.98	96.18±2.32
NN	93.78±2.87	**92.11±3.13**	93.81±2.51	96.97±1.65
RF	93.86±2.72	91.42±3.32	94.02±2.39	96.67±1.79
SVM	**94.01±2.54**	92.06±3.07	**95.43±2.04**	**98.61±1.18**

accuracies on ten subjects' data obtained from different individual classifiers of each modality are shown in Fig. 3 and Table II where the best result for each modality is indicated in bold. For EEG, EMG and MMG, the highest accuracies on testing data are achieved by SVM, RF and NN, respectively. The selected classifiers are consistent with the previous, which means that the selection for the best individual classifier of each modality is reasonable. By comparing the average accuracies on different single modalities, the best scheme for single modality (BSSM) obtains an accuracy of 91.33%, which is implemented by RF classification model using EMG signals. The recognition result of BSSM is shown in the grey highlighted cell in Table II.

B. Recognition Results of Multimodal Fusion Schemes

Different combinations of multimodal signals (EEG, EMG and MMG) are explored to decode human motion intentions. The combining forms can be either two or three modal fusion. In order to verify the practical usefulness of the proposed multimodal fusion framework, all kinds of combinations are evaluated, i.e., EEG-EMG, EEG-MMG, EMG-MMG and EEG-EMG-MMG. Twelve fusion algorithms outlined in Section III-B2 are tested to find the most advantageous method for each combination form. The recognition results for all the fusion algorithms of every combination are shown in Fig. 4 and Table III where the best fusion result for each combination is displayed in bold. According to the constraint condition of MVR in Section III-B2, the MVR can only be utilized under the condition of at least three modalities. Therefore, the evaluation results of the MVR for two-modal fusion have not been calculated, which are represented by short horizontal lines in Table III. On the whole, most classification-based fusion methods outperform rule-based fusion methods. The reason is that the pre-established fusion rules are fixed, which may not be appropriate for different subjects' data. By comparing the results from different combinations of two modalities, the best scheme for two-modal fusion (BSTWF) achieves an accuracy of 95.43%, which is implemented by SVM fusion model using both EMG and MMG. For the combination of three modalities, the best scheme for three-modal fusion (BSTHF) yields an accuracy of 98.61%, which is implemented by SVM fusion model using the three modal signals. The identification results of BSTWF and BSTHF are indicated in the grey highlighted cells in Table III.

C. Comparison on Recognition Results of Different Schemes

The comparison and analysis of experimental results are carried out to verify the effectiveness of the proposed multimodal

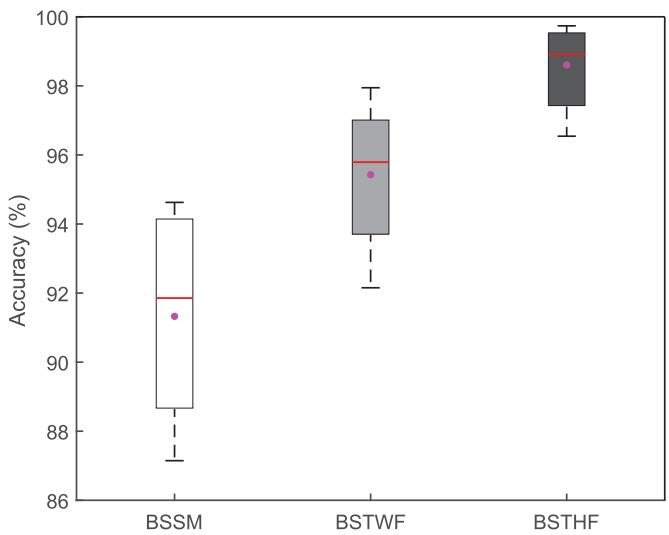

Fig. 5. Classification accuracy distributions of the three best schemes. The magenta dot and the red line are the average value and median value of the accuracy, respectively.

TABLE IV
COMPARISON ON THE AVERAGE CLASSIFICATION ACCURACY (%) OF THE THREE BEST SCHEMES AND SIGNIFICANCE TEST RESULTS ($\alpha = 0.05$)

Scheme	Performance	Pairs	p-value	p_c-value
BSSM	91.33 ± 3.11	BSSM vs. BSTWF	0.0371	0.1113
BSTWF	95.43 ± 2.04	BSSM vs. BSTHF	0.0020	0.0060
BSTHF	98.61 ± 1.18	BSTWF vs. BSTHF	0.0059	0.0177

fusion framework. The specific procedure is conducted from the following aspects.

1) Comparison on the Average Classification Accuracy: The box-and-whisker diagrams of recognition accuracies of the BSSM, BSTWF and BSTHF are shown in Fig. 5, which provide an overall view of accuracy distributions on ten subjects' data. For direct comparison, the average accuracies of the three best schemes mentioned above are duplicated on the left half of Table IV. As shown in this table, in comparison with the results of BSSM, the average accuracy increases 4.10% by using BSTWF and 7.28% by performing BSTHF, respectively. Furthermore, the BSTHF outruns BSTWF by 3.18%. Then, the Wilcoxon signed-rank test [44] is conducted to investigate the significance of difference in this study. The significance level is set at $\alpha = 0.05$. The p-values corresponding to the different pairs are presented on the right half of Table IV. By considering the multiple comparisons problem, the Bonferroni correction [45] are used to correct the p-value. The p_c-value in Table IV denotes the corrected p-value after Bonferroni correction. It can be clearly seen that all the p-values are less than 0.05, which means that BSTWF is significantly superior to BSSM, and BSTHF is significantly better than the other two schemes in terms of the average classification accuracy. After Bonferroni correction, except for the difference between BSSM and BSTWF, other differences in the multiple comparisons are still significant. The above results demonstrate that the recognition accuracy can be significantly enhanced by the proper fusion of more modal signals in this specific task.

TABLE V
DIFFERENT DATA DIVISIONS

Individual-training data	Fusion-training data	Testing data
50%	30%	20%
45%	25%	30%
40%	20%	40%

TABLE VI
COMPARISON ON AVERAGE CLASSIFICATION ACCURACY (%) OF THE THREE BEST SCHEMES UNDER DIFFERENT DATA DIVISIONS

Scheme	50%/30%/20%	45%/25%/30%	40%/20%/40%
BSSM	91.33 ± 3.11	90.10 ± 3.25	88.69 ± 3.42
BSTWF	95.43 ± 2.04	94.42 ± 2.15	93.23 ± 2.28
BSTHF	**98.61 ± 1.18**	**97.78 ± 1.25**	**96.85 ± 1.34**

2) Comparison on the Effects of Different Data Divisions: All the previous results are based on the same data division described in Section III-A, i.e., approximately 50%, 30%, 20% for individual-training data, fusion-training data and testing data. Since all the methods used in this study are essentially data-driven approaches, the different data divisions probably have influences on the recognition results. Besides, it is usually very difficult to get enough multimodal data for training recognition models in practical application. We doubt whether the proposed multimodal fusion framework can still maintain good performance as the training data become less and the test data become more. To this end, the recognition experiments under different data divisions are carried out. The specific forms of data division are shown in Table V. It is worth to note that in these experiments the training and testing procedures are consistent with that mentioned in Section III-A. During the training of the local decision layer, the best individual classifier for each modality is reselected under the latter two new data divisions. By comparing the recognition performance on the fusion-training data, the selection outcomes are SVM for EEG, RF for EMG and NN for MMG, which are the same as before. During the testing procedure, the best schemes by using single modality, two-modal fusion and three-modal fusion are separately evaluated again under the latter two data divisions. By comparing the recognition results, the BSSM, BSTWF and BSTHF under the latter two data divisions are kept in line with that of the previous, i.e., BSSM implemented by RF classification model using EMG, BSTWF implemented by SVM fusion model using both EMG and MMG, BSTHF implemented by SVM fusion model using the three modalities. Due to space limitation, we only report the results of the above three best schemes here.

Fig. 6 shows the overall distributions of the recognition results for the three best schemes under different data divisions. Table VI shows the average classification accuracies obtained by BSSM, BSTWF and BSTHF under different data divisions, where the best recognition result under each data division is displayed in bold. With the increase in the proportion of testing data, the recognition results for all the three schemes present a continuous downward trend. Specifically, when the testing data

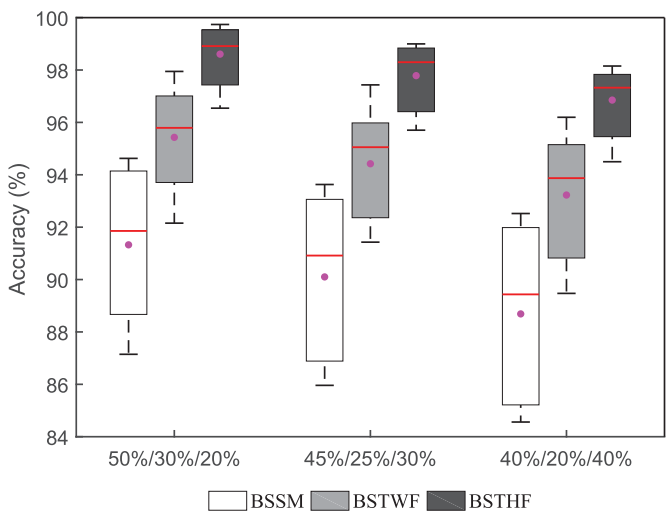

Fig. 6. Classification accuracy distributions of the three best schemes under different data divisions. The magenta dot and the red line are the average value and median value of the accuracy, respectively.

TABLE VII
SIGNIFICANCE TEST RESULTS UNDER DIFFERENT DATA DIVISIONS ($\alpha = 0.05$)

Pairs	45%/25%/30%		40%/20%/40%	
	p-value	p_c-value	p-value	p_c-value
BSSM vs. BSTWF	0.0098	0.0294	0.0098	0.0294
BSSM vs. BSTHF	0.0020	0.0060	0.0020	0.0060
BSTWF vs. BSTHF	0.0098	0.0294	0.0020	0.0060

increase by 10%, the average accuracies for BSSM, BSTWF and BSTHF decrease by 1.23%, 1.01% and 0.83%, respectively. When the testing data rise by 20%, the accuracies reduce by 2.64%, 2.20% and 1.76%, respectively. Compared with BSSM and BSTWF, BSTHF has obvious advantages in terms of the average accuracy and robustness against changes in data division. In addition, the outcomes of the Wilcoxon signed-rank test ($\alpha = 0.05$) performed on different pairs are given in Table VII, which reveal that BSTWF significantly outruns BSSM, and BSTHF significantly outperforms the other two schemes under different data divisions. After Bonferroni correction, all the differences in the multiple comparisons remain significant. These results further indicate that the classification accuracy can be significantly improved by appropriately combining more modal signals even if the data division is changed. Therefore, the proposed fusion framework has some potential to keep good classification performance in practical application with insufficient training data.

3) Comparison on the Processing Delay: Since continuous classification is based on the windowing technique, the processing delay for an analysis window on testing data should be evaluated and discussed. For the single-modal scheme, the processing delay only contains the time for generating a decision. For the multimodal fusion scheme, the processing delay consists of the time for producing local decisions and the time for fusing local decisions. For a 62.5 ms analysis window which contains 16 data points for each modality, the processing delays by using BSSM, BSTWF and BSTHF are roughly 7 ms, 12 ms and 15 ms, respectively. Since the overlapping ratio of the window is set as 50%, all the above processing delays are less than 31.25 ms that is the interval time between two adjacent analysis windows. Consequently, the total delay between two consecutive predictions is equal to 31.25 ms (much less than 100 ms) for the above three schemes, which is an acceptable delay for a real-time HRI [46].

D. Impact of the Recognition Performance Improvement

The application background of this study is to design a multimodal HRI to control the biorobots for assisting the patients with lower extremity motor dysfunction. Therefore, the better the recognition performance of human motion intentions is, the better the practical usefulness will be. There are many factors that affect the performance improvement, such as types of signal (i.e., modalities), signal quality, preprocessing methods, feature extraction methods, classifiers, fusion methods, etc. We focus on the selection of modalities, classifiers and fusion methods. The selection of classifiers for single modality is discussed at first. From the experimental results in Table II, for each single modality, some classifiers indeed achieve higher accuracies than others. The highest accuracy yielded by single-modal schemes is up to 91.33%, but the performance still cannot meet the demand for precise control of biorobots. This means that only selecting appropriate classifiers cannot satisfy the actual demand. In order to further improve the recognition performance, different combination forms of multimodal signals and different decision fusion methods have been explored and evaluated. From the experimental results in Table III, for each combination form of multimodalities, some fusion methods indeed obtain better results than others. For each fusion method, the general trend is that the accuracy is improved as the number of modalities increases. The highest accuracy can be improved to 98.61%, which is much better than the cases of only using a single modality. It is worth pointing out that the decision fusion methods used in this study are implemented based on the output of the individual classifier for each single modality. Therefore, all the three factors (i.e. modalities, classifiers and fusion methods) considered in this paper have certain impacts on the recognition performance of human motion intentions. The relative importance and correlation of the three factors affecting the performance will be further explored in our future work.

V. CONCLUSION

This study provides a multimodal fusion framework based on three different modal biosignals for lower limb motion intention recognition. Under this framework, the cortical and muscular activities can be effectively integrated by synchronously processing EEG, EMG and MMG. Perhaps the most exciting results are that with our collected data, the proposed best three-modal fusion scheme not only brings significant improvement in the average recognition accuracy compared with the cases of two-modal fusion or only a single modality, but also has

advantages in the robustness against changes in data division. In addition, the processing delay produced by the best three-modal fusion scheme can meet the requirement of continuous recognition during the human-robot interaction. The subsequent work will continue to explore more advanced multimodal fusion applications, including the classification of other activities of daily life and the popularization of feasibility experiments on the patients.

Acknowledgment

The authors would like to thank all the subjects who participated in data collection.

References

[1] S. Wang *et al.*, "Design and control of the MINDWALKER exoskeleton," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 2, pp. 277–286, Mar. 2015.

[2] K. H. Ha, S. A. Murray, and M. Goldfarb, "An approach for the cooperative control of FES with a powered exoskeleton during level walking for persons with paraplegia," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 4, pp. 455–466, Apr. 2016.

[3] H. Fuketa *et al.*, "1 μm-thickness ultra-flexible and high electrode-density surface electromyogram measurement sheet with 2 V organic transistors for prosthetic hand control," *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 6, pp. 824–833, Dec. 2014.

[4] S. Huang, J. P. Wensman, and D. P. Ferris, "Locomotor adaptation by transtibial amputees walking with an experimental powered prosthesis under continuous myoelectric control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 5, pp. 573–581, May 2016.

[5] H. Toreyin and P. T. Bhatti, "A low-power ASIC signal processor for a vestibular prosthesis," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 3, pp. 768–778, Jun. 2016.

[6] D. Novak and R. Riener, "A survey of sensor fusion methods in wearable robotics," *Rob. Auton. Syst.*, vol. 73, pp. 155–170, 2015.

[7] N. Rodrìguez, J. Weissberg, and G. E. Loeb, "Flexible communication and control protocol for injectable neuromuscular interfaces," *IEEE Trans. Biomed. Circuits Syst.*, vol. 1, no. 1, pp. 19–27, Mar. 2007.

[8] C.-T. Lin *et al.*, "Wireless and wearable EEG system for evaluating driver vigilance," *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 2, pp. 165–176, Apr. 2014.

[9] T. Roh, K. Song, H. Cho, D. Shin, and H.-J. Yoo, "A wearable neurofeedback system with EEG-based mental status monitoring and transcranial electrical stimulation," *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 6, pp. 755–764, Dec. 2014.

[10] M. Guermandi, R. Cardu, E. F. Scarselli, and R. Guerrieri, "Active electrode IC for EEG and electrical impedance tomography with continuous monitoring of contact impedance," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 1, pp. 21–33, Feb. 2015.

[11] X. Liu, M. Zhang, B. Subei, A. G. Richardson, T. H. Lucas, and J. V. der Spiegel, "The PennBMBI: Design of a general purpose wireless brain-machine-brain interface system," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 2, pp. 248–258, Apr. 2015.

[12] S. Benatti *et al.*, "A versatile embedded platform for EMG acquisition and gesture recognition," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 5, pp. 620–630, Oct. 2015.

[13] V. Nathan and R. Jafari, "Design principles and dynamic front end reconfiguration for low noise EEG acquisition with finger based dry electrodes," *IEEE Trans. Biomed. Circuits Syst.*, vol. 9, no. 5, pp. 631–640, Oct. 2015.

[14] F. N. Guerrero, E. M. Spinelli, and M. A. Haberman, "Analysis and simple circuit design of double differential EMG active electrode," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 3, pp. 787–795, Jun. 2016.

[15] B. G. D. Valle, S. S. Cash, and C. G. Sodini, "Low-power, 8-channel EEG recorder and seizure detector ASIC for a subdermal implantable system," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 6, pp. 1058–1067, Dec. 2016.

[16] H. Ding, Q. He, L. Zeng, Y. Zhou, M. Shen, and G. Dan, "Motion intent recognition of individual fingers based on mechanomyogram," *Pattern Recognit. Lett.*, vol. 88, pp. 41–48, 2017.

[17] J. D. R. Millán *et al.*, "Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges," *Front. Neurosci.*, vol. 4, no. 161, pp. 1–15, 2010.

[18] D. G. E. Robertson, G. E. Caldwell, J. Hamill, G. Kamen, and S. N. Whittlesey, *Research Methods in Biomechanics*. Champaign, IL, USA: Human Kinetics, 2004.

[19] C. Orizio, "Muscle sound: Bases for the introduction of a mechanomyographic signal in muscle studies." *Crit. Rev. Biomed. Eng.*, vol. 21, no. 3, pp. 201–243, 1993.

[20] R. Leeb, H. Sagha, R. Chavarriaga, and J. d. R. Millán, "A hybrid brain-computer interface based on the fusion of electroencephalographic and electromyographic activities." *J. Neural Eng.*, vol. 8, no. 2, pp. 587–589, 2011.

[21] K. Kiguchi and Y. Hayashi, "A study of EMG and EEG during perception-assist with an upper-limb power-assist robot," in *Proc. IEEE Int. Conf. Robot. Autom.*, Saint Paul, Minnesota, 2012, pp. 2711–2716.

[22] E. A. Kirchner and M. Tabie, "Closing the gap: Combined EEG and EMG analysis for early movement prediction in exoskeleton based rehabilitation," in *Proc. 4th Eur. Conf. Technol. Assisted Rehabil.*, Berliny, 2013, pp. 1–4.

[23] P. Xie, X. Chen, P. Ma, X. Li, and Y. Su, "Identification method of human movement intention based on the fusion feature of EEG and EMG," in *Proc. World Congress Eng.*, London, 2013, pp. 1–5.

[24] Y. Song, Y. Du, X. Wu, X. Chen, and P. Xie, "A synchronous and multi-domain feature extraction method of EEG and sEMG in power-assist rehabilitation robot," in *Proc. IEEE Int. Conf. Robot. Autom.*, Hong Kong, 2014, pp. 4940–4945.

[25] N. N. Shusharina, E. A. Bogdanov, V. A. Petrov, E. V. Silina, and M. V. Patrushev, "Multifunctional neurodevice for recognition of electrophysiological signals and data transmission in an exoskeleton construction," *Biol. Med.*, vol. 8, no. 6, pp. 1–7, 2016.

[26] P. Prociow, A. Wolczowski, T. G. B. Amaral, O. P. Dias, and J. Filipe, "Identification of hand movements based on MMG and EMG signals," in *Proc. 1st Int. Conf. Biomed. Electron. Devices*, Funchal, Madeira, 2008, pp. 534–539.

[27] W. Guo, X. Sheng, D. Zhang, and X. Zhu, "Development of a hybrid surface EMG and MMG acquisition system for human hand motion analysis," in *Proc. 8th Int. Conf. Intell. Robot. Appl.*, Portsmouth, 2015, pp. 329–337.

[28] M. Kurzynski, "Multiclassifier system with dynamic model of classifier competence applied to the control of bioprosthetic hand," in *Proc. Global Conf. Artif. Intell.*, Tbilisi, Georgia, 2015, pp. 163–175.

[29] M. Kurzynski and M. Majak, "Meta-Bayes classifier with Markov model applied to the control of bioprosthetic hand," in *Intelligent Decision Technologies 2016*. Cham, ZG, Switzerland: Springer, 2016, pp. 107–117.

[30] M. Kurzynski and A. Wolczowski, "Control system of bioprosthetic hand based on advanced analysis of biosignals and feedback from the prosthesis sensors," in *Proc. 3rd Int. Conf. Inf. Technol. Biomed.*, Gliwice, 2012, pp. 199–208.

[31] J. P. Chapman, L. J. Chapman, and J. J. Allen, "The measurement of foot preference." *Neuropsychologia*, vol. 25, no. 3, pp. 579–84, 1987.

[32] S. Coren, "The lateral preference inventory for measurement of handedness, footedness, eyedness, and earedness: Norms for young adults," *Bull. Psychon. Soc.*, vol. 31, no. 1, pp. 1–3, 1992.

[33] A. Perotto, *Anatomical Guide for the Electromyographer for Limbs and Trunk*, 4th ed. Springfield, IL, USA: Charles C Thomas, 2005.

[34] K. B. Englehart, B. Hudgins, and P. A. Parker, "A wavelet-based continuous classification scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 3, pp. 302–311, Mar. 2001.

[35] K. B. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 7, pp. 848–854, Jul. 2003.

[36] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 121–131, Jan. 2011.

[37] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7420–7431, 2012.

[38] W. Youn and J. Kim, "Feasibility of using an artificial neural network model to estimate the elbow flexion force from mechanomyography." *J. Neurosci. Methods*, vol. 194, no. 2, pp. 386–93, 2011.

[39] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.

[40] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in Large Margin Classifiers*. Cambridge, MA, USA: MIT Press, 2000, pp. 61–74.
[41] P. K. Atrey, M. A. Hossain, A. El-Saddik, and M. S. Kankanhalli, "Multimodal fusion for multimedia analysis: A survey," *Multimedia Syst.*, vol. 16, no. 6, pp. 345–379, 2010.
[42] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, Mar. 1998.
[43] G. L. Rogova and V. Nimier, "Reliability in information fusion: Literature survey," in *Proc. 7th Int. Conf. Inf. Fusion*, Stockholm, 2004, pp. 1158–1165.
[44] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bull.*, vol. 1, no. 6, pp. 80–83, 1945.
[45] J. M. Bland and D. G. Altman, "Multiple significance tests: The Bonferroni method," *Br. Med. J.*, vol. 310, no. 6973, p. 170, 1995.
[46] T. R. Farrell and R. F. Weir, "The optimal controller delay for myoelectric prostheses," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 1, pp. 111–118, Mar. 2007.

Chengkun Cui received the B.E. degree in automation from Tianjin University of Technology and Education, Tianjin, China, in 2012, and the M.E. degree in control engineering from Huazhong University of Science and Technology, Wuhan, China, in 2014. He is currently working toward the Ph.D. degree at the Institute of Automation, Chinese Academy of Sciences, Beijing, China.

His research interests include human motion analysis, information fusion, and signal processing.

Gui-Bin Bian (M'13) received the B.E. degree in mechanical engineering from North China University of Technology, Beijing, China, in 2004, and the M.E. and Ph.D. degrees in mechanical engineering from Beijing Institute of Technology, Beijing, China, in 2007 and 2010, respectively.

He is currently an Associate Professor in the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China, where his research interests include medical robotics and human robot interaction.

Zeng-Guang Hou (SM'09) received the B.E. and M.E. degrees in electrical engineering from Yanshan University (formerly North-East Heavy Machinery Institute), Qinhuangdao, China, in 1991 and 1993, respectively, and the Ph.D. degree in electrical engineering from Beijing Institute of Technology, Beijing, China, in 1997. From May 1997 to June 1999, he was a Postdoctoral Research Fellow in the Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing. He was a Research Assistant at the Hong Kong Polytechnic University, Hong Kong, China, from May 2000 to January 2001. From July 1999 to May 2004, he was an Associate Professor in the Institute of Automation, Chinese Academy of Sciences, and has been a Full Professor since June 2004. From September 2003 to October 2004, he was a Visiting Professor at the Intelligent Systems Research Laboratory, College of Engineering, University of Saskatchewan, Saskatoon, SK, Canada. He is a Professor and Deputy Director of the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include neural networks, robotics, and intelligent systems.

He was an Associate Editor of the IEEE COMPUTATIONAL INTELLIGENCE MAGAZINE and the IEEE TRANSACTIONS ON NEURAL NETWORKS, and the Chair of Neural Network Technical Committee of Computational Intelligence Society (CIS). He is currently the Chair of Adaptive Dynamic Programming and Reinforcement Learning Technical Committee of CIS. He is an Associate Editor of the IEEE TRANSACTIONS ON CYBERNETICS and *ACTA Automatica Sinica*, etc. and an editorial board member of *Neural Networks*.

Jun Zhao received the B.Med. degree in clinical medicine from Qingdao University, Shandong, China, in 1991, and the M.Med. and Ph.D. degrees in rehabilitation medicine and physical therapy from Capital Medical University, Beijing, China, in 2005 and 2012, respectively.

He is currently the Chief Physician in the Department of Neurology, Beijing Bo'ai Hospital, China Rehabilitation Research Center, Beijing, China. His research interests include diagnosis and treatment of cerebrovascular disease and brain damage dysfunction, and neurological rehabilitation.

Hao Zhou received the B.Med. degree in clinical medicine from Jining Medical University, Shandong, China, in 2010, and the M.Med. degree in neurology from Jinzhou Medical University, Liaoning, China, in 2013.

He is currently a Resident Physician in the Department of Neurology, Beijing Bo'ai Hospital, China Rehabilitation Research Center, Beijing, China. His research interests include electromyogram and neurological rehabilitation.