Adaptive Electrode Calibration Method Based on Muscle Core Activation Regions and Its Application in Myoelectric Pattern Recognition

Ruochen Hu\textsuperscript{c}, Graduate Student Member, IEEE, Xiang Chen\textsuperscript{e}, Member, IEEE, Xu Zhang\textsuperscript{c}, Member, IEEE, and Xun Chen\textsuperscript{\textcopyright}, Senior Member, IEEE

Abstract—To reduce the bad effect of electrode shifts on myoelectric pattern recognition, this paper presents an adaptive electrode calibration method based on core activation regions of muscles. In the proposed method, the high-density surface electromyography (HD-sEMG) matrix collected during hand gesture execution is decomposed into source signal matrix and mixed coefficient matrix by fast independent component analysis algorithm firstly. The mixed coefficient vector whose source signal has the largest two-norm energy is selected as the major pattern, and core activation region of muscles is extracted by traversing the major pattern periodically using a sliding window. The electrode calibration is realized by aligning the core activation regions in unsupervised way. Gestural HD-sEMG data collection experiments with known and unknown electrode shifts are carried out on 9 gestures and 11 participants. A CNN+LSTM-based network is constructed and two network training strategies are adopted for the recognition task. The experimental results demonstrate the effectiveness of the proposed method in mitigating the bad effect of electrode shifts on gesture recognition accuracy and the potentials in reducing user training burden of myoelectric control systems. With the proposed electrode calibration method, the overall gesture recognition accuracies increase about (5.72\textpm7.69)\%. In specific, the average recognition accuracy increases (13.32\textpm7.30)\% when using only one batch of data in data diversity strategy, and increases (12.01\textpm13.75)\% when using only one repetition of each gesture in model update strategy. The proposed electrode calibration algorithm can be extended and applied to improve the robustness of myoelectric control system.

Index Terms—Electrode calibration, core activation region of muscles, HD-sEMG, FastICA algorithm, myoelectric pattern recognition.

I. INTRODUCTION

MYOELECTRIC control is a technique that interprets physical movements or intention as machine command by surface Electromyography (sEMG) signals, which has potential medical application prospects such as prosthesis, rehabilitation training, etc. [1], [2], and commercial application prospects such as human-computer interaction, electronic consumption, etc. [3], [4]. The main steps of myoelectric control include sEMG signals collection, myoelectric pattern recognition and the implementation of hardware equipment. Among them, the accuracy of myoelectric pattern recognition largely determines the performance of myoelectric control systems. Although most myoelectric pattern recognition methods can achieve satisfactory accuracy in ideal laboratory conditions, they are difficult to be applied to practical myoelectric control systems because of the sensitivity to electrode shifts, the change of contraction level, muscle fatigue and so on. In particular, electrode shifts caused by the replacement of electrodes or human movement can bring large distribution differences between training data and test data, thus reduce the performance of classifiers.

Since the replacement of electrodes or human movement are inevitable in practical application, effective measures must be taken to overcome the influence of electrode shifts. Young et al. [5], [6] attempted to improve the robustness of myoelectric pattern recognition by changing inter-electrode distance of bipolar channels and trying various electrode configurations. They investigated how the size of the electrode detection surface and electrode orientation affect system robustness in the situations of with or without electrode shifts. Although it was found that optimizing inter-electrode intervals, space structures, and the size or orientation can improve certainly the system performance, this kind of method is relatively inconvenient and difficult to adopt in the practical application.

Aiming to adapt to the change of data distribution brought by electrode shifts, data diversity strategy, in which training data are collected from different positions, has often been adopted to enhance the generalization ability of classifier. In the works of [7] and [8], in order to train classifier with strong generalization, various electrode shifts were made manually during the gesture data collection experiment to ensure that diverse muscle activation states can be captured and learned. The drawback of data diversity strategy lies that classifier is poor when the distributions of training data and test data have significant discrepancy. Model update strategy, in which the incoming data is used to assist updating or...
fine-tuning the parameters of the prior well-trained classifier, is another training strategy usually adopted to reduce the effect of electrode shifts and improve the robustness of classifier. Vidovic et al. proposed a method called mixed-LDA (linear discriminant analysis) to compensate for the effect of electrode shifts by updating a pre-trained LDA model with distribution parameters of post-shift EMG data [9]. Zhang et al. proposed a semi-supervised adaptation strategy to learn the new distribution information of post-shift samples, thus improve classification performance [10]. Ameri et al. proposed a supervised adaptation approach based on transfer learning to mitigate the influence of electrode shift, which only requires few post-shift samples to recalibrate the system [11]. The disadvantage of the model updating strategy is that a certain quantity of data is needed to collect for the fine-tuning of the model parameters, which is unfriendly to the practical application of myoelectric control system.

From the perspective of classification algorithm, the hand-crafted features based methods such as support vector machine (SVM) [12], linear discriminant analysis (LDA) [13], k-nearest neighbor (KNN) [14] etc. have been widely applied to myoelectric pattern recognition. The change of data distribution characteristics caused by electrode shifts has a great influence on the performance of the system taking handcrafted features based methods. In recent years, deep learning nets, which have high generalization ability and can automatically extract the optimal features with strong specificity, have also been introduced into myoelectric pattern recognition [15], [16]. Compared with the handcrafted features based methods, end-to-end deep learning has huge potential for improving the accuracy and robustness of the myoelectric pattern recognition.

From the way of data collection, HD-sEMG technology has inherent advantages in overcoming the influence of electrode shifts. HD-sEMG electrode array of sufficient size can cover fully the target muscle, so as to capture substantial amount of muscle activation spatial information. HD-sEMG has demonstrated remarkable performance in many applications, particularly in compensating for the effect of electrode shifts. Stango et al. used HD-sEMG grid to collect gesture data and found that a measure of the degree of spatial correlation - the variogram - had lower sensitivity to electrode shifts (up to 1 cm) [17]. He et al. used the gray-level co-occurrence matrix to describe the spatial distribution among HD-sEMG grid, and found it is insensitive to electrode shifts [18]. Daley et al. adopted HD-sEMG grid as a tool to identify optimal electrode sites for resisting electrode shifts [19]. Using a 10 × 10 electrode array, Zhang et al. proposed an unsupervised manner to detect and calibrate electrode shifts [20]. Overall, most researches based on HD-sEMG mainly make full use of spatial information to overcome the bad effect of electrode shifts.

In a previous work on HD-sEMG-based muscle force estimation, we reported that the location of the core activation region of muscles can be achieved by applying matrix decomposition algorithms such as nonnegative matrix factorization (NMF) [21] and independent component analysis (ICA) [22] etc. on HD-sEMG signals. For a force-related task, the NMF algorithm was used to factorize the HD-sEMG matrix into activation patterns and corresponding time-varying coefficient vectors, and the activation pattern with the highest activation intensity was adopted to extract the core region of muscle activation. The results of force estimation demonstrated that signals from the core activation region could track force change effectively.

Inspired by the previous work, we speculate that the core activation region of muscles accompanied by other type of human movement can also be extracted. With the assumption that the core activation region of muscle is relatively stationary and consistent for the same gesture, the bad effect of electrode shifts on the performance of myoelectric pattern recognition is expected to be solved by aligning the core activation regions extracted from different repetitions. Therefore, taking advantages of HD-sEMG technology, we propose an adaptive electrode calibration method based on core activation region of muscle for the realization of robust gesture recognition in this paper. Unlike related works, the novelty of this study is trying to reduce the influence of the electrode shifts by locating and aligning core activation regions of muscle. The feasibility and effectiveness of the proposed electrode calibration method in mitigating the bad effect of electrode shifts on gesture recognition accuracy and the potentials in reducing the user training burden of myoelectric control systems are verified through two types of hand gesture recognition experiments.

II. MATERIALS AND METHODOLOGY

The block diagram in Fig. 1 illustrates the overall framework of hand gesture recognition based on adaptive electrode calibration method proposed in this study. When a gesture task is performed, HD-sEMG signals are collected from target muscles, FastICA algorithm is applied on HD-sEMG matrix to extract the core activation region of muscles, and the signals from the core region are fed into a CNN+LSTM-based network for gesture recognition.

A. Hand Gesture Set, Participants and Data Collection

In this study, the target gesture set consists of 9 hand gestures (shown in Fig. 2(a)) involving various state combinations of wrist joint, metacarpophalangeal joint, and interphalangeal joint. The duration of each gesture is about 5s, which can be divided into three stages, namely: the start stage (about 1s), the steady stage (about 3s), and the end stage (about 1s). In the start stage, the gesture is about to perform and the related muscles begin to contract. In the steady stage, the gesture is carried out, all joints and the muscle contraction modes are
In order to verify the feasibility and superiority of the proposed electrode calibration method, two types of hand gesture data collection experiments are carried out as follows:

1) Supervised Electrode Shift (S-ES) Experiment: The main purpose of S-ES experiment is to simply and intuitively explore whether the calibration algorithm can capture the electrode shifts. One representative subject (Participant 2) participates this experiment. As shown in Fig. 2(c), a 48-channel HD-sEMG grid is placed on the forearm extensors of the right hand for the data collection. Five batches of gesture data with known electrode shifts are collected. In the first trial, data is collected at initial position determined via the knowledge of human anatomy. The target muscles are covered adequately and located in the center of the grid as much as possible. In the following four trials, taking the initial position as the reference, the grid shifts towards the medial-distal, lateral-distal, medial-proximal, and lateral-proximal directions respectively with the spacing of an inter-electrode interval, namely 14 mm. In each trial, the participant executes all 9 gestures at a comfortably medium force level, and a 5-minute rest is asked to avoid muscle fatigue after each gestural task. Eight repetitions of each gesture are performed with the interval of 3s.

2) Unsupervised Non-Rotational Electrode Shift (UN-ES) Experiment: The main purpose of UN-ES experiment is to further explore the validity of the proposed calibration method from the actual application scenarios. Two electrodes are used to capture more activation information of muscles for gesture recognition. All 11 participants participate this experiment. For each participant, five batches of gesture data with unknown electrode shifts are collected at five different trials during 2-3 days. As shown in Fig. 2(d), two 48-channel HD-sEMG grids are placed on the forearm extensors and flexors respectively. Because the five trials are carried out at different days, the replacement of the grids inevitably led to electrode shifts. It should be noted that the electrodes can shift towards medial/lateral or distal/proximal directions, but cannot rotate. In each trial, participants execute all gestures at a comfortably medium force level, and a 5-minute rest is asked to avoid muscle fatigue after each gestural task. Eight repetitions of each gesture are performed with the interval of 3s.

Raw HD-sEMG signals are preprocessed by the following procedures. First, few channels (about 2~3 channels) whose signal amplitude out of the reasonable range are discarded and replaced by the mean value of neighboring channels. Then, the signals are band-pass filtered digitally with 20-500 Hz bandwidth (finite impulse response filter, Hanning window, 20th order) to remove the low-frequency noise and high-frequency interference.

B. ICA-Based Adaptive Electrode Calibration Algorithm

The matrix decomposition algorithms such as principle component analysis (PCA), independent component analysis (ICA) and non-negative matrix analysis (NMF) can be used to extract muscle activation modes or contribution sources from HD-sEMG signals [21]–[23]. The core activation region of muscles, which is determined via the modes or sources that contribute significantly to gesture task, can be further used for electrode self-adaptive calibration. Due to the stability and
physiological interpretability [22], [24], the FastICA algorithm is adopted in this study.

With FastICA algorithm, the mean-removed HD-sEMG signal matrix \( X \in \mathbb{R}^{m \times T} \) of one grid (\( m \) channels and \( T \) samples) can be regarded as the combination of weighted source signals \( S \in \mathbb{R}^{n \times T} \) (\( n \) represents the number of source signals) according to mixed coefficient matrix \( A \in \mathbb{R}^{m \times n} \), as shown in Formula (1). Each row of \( S \) is one source signal and each column of \( A \) is its corresponding weighting coefficient vector [22], [24].

\[
X = A \times S \tag{1}
\]

Before applying FastICA algorithm, the HD-sEMG signals need to be processed by principal component analysis (PCA) algorithm and whitened. The processing by PCA is to transform the mean-removed HD-sEMG signals matrix \( X \) into a matrix \( Y \in \mathbb{R}^{k \times T} \) consisting of \( k \) uncorrelated principal components through orthogonal transformation technique, which is shown in Formula (2). Where \( U \in \mathbb{R}^{m \times k} \) is the orthogonal eigenmatrix of the covariance matrix of \( X \). \( U^T \) is the transposition form. Each row of \( U \) is an eigenvector corresponding to the eigenvalue \( \lambda_i \), \( i = 1, 2, \ldots, k \). The several principal components whose variability account for (VAF, shown in Formula (3)) attains 0.95 are generally selected to describe the characteristics of original data [22], [24]. So the number of principal components \( k \) can be determined by the threshold.

\[
Y = U^T \times X \tag{2}
\]

\[
VAF = \sum_{i=1}^{k} \frac{\lambda_i}{\sum_{i=1}^{m} \lambda_i} \tag{3}
\]

In the FastICA framework, the number of sources \( n \) generally equals to the number of principal modes \( k \). The whitened matrix \( Y_1 \in \mathbb{R}^{k \times T} \) can be calculated by Formula (4) and (5) [22], [24].

\[
Y_1 = C \times Y \tag{4}
\]

\[
C = \text{inv}(\text{sqrt}(D)) \tag{5}
\]

where \( C \) is a whitened operator with the information of eigenvalues so as to normalize all the principal components, \( D \) is an diagonal matrix with diagonal elements equaling to all eigenvalues, and \( \text{sqrt}(\cdot) \) and \( \text{inv}(\cdot) \) are square root transformation and inversion operators respectively.

Based on central limit theorem, FastICA theory assumes that the source signals, which are non-Gaussian compared to the mixed signals, could be obtained by maximizing the non-Gaussianity. In this study, the FastICA algorithm adopting a fixed point iteration based on maximizing the negentropy as well as the non-Gaussianity is selected. The main iterative process can be expressed by Formula (6) and (7) [22], [24].

\[
w \leftarrow E[Y_1 \cdot g(w'Y_1)] - E[g'(w'Y_1)]w \tag{6}
\]

\[
w = w/||w|| \tag{7}
\]

where \( w \) is the de-mixed vector corresponding to one source signal, \( E(\cdot) \) is the expectation operator, and \( g(\cdot) \) is a particular function (e.g. \( g(x) = x^3 \)). The iteration process would not stop until \( w \) converges. All the obtained \( w \) vector (the number of vectors is \( n \)) constitutes the de-mixed matrix \( W \) after Schmidt orthogonalization, the source signal matrix could be obtained using Formula (8), and then the mixed coefficient matrix \( A \) can be calculated by Formulas (9) [22], [24].

\[
S = W \times Y_1 = W \times \text{inv}(\text{sqrt}(D)) \times U^T \times X \tag{8}
\]

\[
A = U \times \text{sqrt}(D) \times \text{inv}(W) \tag{9}
\]

Assume that the electrode grid is in the size of \( P \times Q \) (the product of \( P \) and \( Q \) is \( m \)), the electrode calibration algorithm based on core activation region of muscle is applied in each electrode grid respectively according to the specific steps as follows:

- **Step 1:** To decompose the HD-sEMG signal \( X \) of one electrode into source signals \( S = [S_1, S_2, \ldots, S_n] \) \((n \leq m)\) and the corresponding mixed coefficient vectors \( A = [A_1, A_2, \ldots, A_n] \) according to the criterion of \( \text{VAF} \geq 95\% \).
- **Step 2:** To normalize each mixed coefficient vector \( A_j (j = 1, 2, \ldots, n) \) by dividing the sum of squares of all \( m \) elements, and multiply the corresponding source signal \( S_j \) by the same value.
- **Step 3:** To select the mixed coefficient vector with the largest two-norm energy of the source signal as the major pattern, calculate the absolute of major pattern and re-label each element with the sequence number \( 1 \sim m \) from small to large.
- **Step 4:** To map the core activation region to the corresponding position of the grid. To traverse the processed major pattern periodically using a sliding window with size of \( p \)-channel \( \times q \)-channel \((p \leq P, q \leq Q)\). The region with the largest sum of sequence number in the sliding window is selected as the core activation region. If there are several cases in which the sum is same, the one with the smallest variance is chosen.
- **Step 5:** To normalize the signals in the core activation region using the maximum value across all channels for subsequent gesture recognition.

### C. Hand Gesture EMG Sample Establishment

The entire EMG data stream of one gesture repetition is segmented into \( M \) overlapping analysis windows (length: 256ms, increment: 128ms). The window length is determined by experimental results. Because sEMG signals are relatively stable in the steady stage, the analysis windows are extracted from that and the number of \( M \) was set to 20. For each analysis window, four time domain features, namely: mean amplitude value (MAV), variance (VAR), waveform length (WL), and zero-cross (ZC), are calculated on the selected channels and then reshaped into sEMG images. In other words, the data sample of each analysis window is in the form of \( 4 \) (dimensions of features) \( \times \) height \( \times \) width, where the product of height and width is the number of channels. In the following section, when training CNN model, the gesture dataset contains \( 9 \) (gestures) \( \times \) \( 8 \) (repetitions) \( \times \) \( 5 \) (batches) \( \times \) \( 20 \) (analysis windows) = 7200 samples for each participant. The one-hot labels are matched to analysis windows in accordance with the gesture categories. When training LSTM model, the number of samples is \( 9 \) (gestures) \( \times \) \( 8 \) (repetitions) \( \times \) \( 5 \) (batches) = 360 for each participant. The one-hot labels are matched to repetitions in accordance with gesture categories.
The architecture of the CNN: each hidden layer, and the Softmax layer. The layer is flattened into one-dimensional vector through flatten problem behind each max pooling layer. The output of middle neurons lost at a probability is used to prevent the over-fitting in convolution layers. Dropout [29] that make covariate shift [28]. The Normalization is used in each layer for reducing the internal features and reduce the dimensions of features [25]. Batch the pooling layer is mainly used to further extract effective some convolution kernels for local features extraction, and connected layers. The convolution layer is accompanied with convolution layers, pooling layers, flatten layer, and fully-connected layers. The complete CNN model consists of several layers including convolution layers, pooling layers, flatten layer, and fully-connected layers. The convolution layer is accompanied with some convolution kernels for local features extraction, and the pooling layer is mainly used to further extract effective features and reduce the dimensions of features [25]. Batch Normalization is used in each layer for reducing the internal covariate shift [28]. The $L_2$ regularizer is applied to prevent over-fitting in convolution layers. Dropout [29] that make neurons lost at a probability is used to prevent the over-fitting problem behind each max pooling layer. The output of middle layer is flattened into one-dimensional vector through flatten layer. The $Relu$ non-linearity function [30] is adopted after each hidden layer, and the $Softmax$ function is adopted in the last fully-connected layer. The Adaptive moment estimation (Adam) [31] is selected as the network optimizer. The mini-batch training strategy [32] are adopted to avoid the problem that it is difficult to converge in one single batch training, and the batch size is usually set to between one and the sample size.

The CNN block is more likely to be a feature extractor, and the feature vectors outputted via the flatten layer of CNN are fed into the LSTM. Each input of LSTM layer is M-dimensional feature vector corresponding to a gesture repetition. The LSTM layers integrate the time information of M analysis windows to make decision. The $Relu$ non-linearity function is adopted after each hidden layer. In the training phase, each analysis window is treated as a sample to train a generic CNN model first. The well-trained CNN is adopted to be a feature extractor, whose flatten layer is taken as the input of the LSTM model. Then the LSTM network is trained. In the test phase, the samples can be directly input into the CNN block, and the classification results are given by the back-end LSTM block.

2) The Determination of the Structure and Parameter of the Network: In the neural network, the more layers and the number of neurons can bring the larger expressed feature space, but the more time and space complexity will be consumed, and the phenomenon of over-fitting will appear. In order to optimize the structure and hyper parameters of the CNN+LSTM-based network, a lot of parameter adjustment should be carried out, mainly including: the number of convolution layers and convolution kernels of CNN, the number of LSTM layers, the dropout value, and the initial learning rate etc.. The structure and hyper parameters are adjusted according to the loss function value and recognition accuracy of validation set. When the training error is convergent and generalization error decreases to a relatively lower level, the structure and hyper parameters are considered as the optimal selection.

3) Two Training Strategies of the Network: To demonstrate the effectiveness and superiority of the proposed electrode calibration algorithm, gesture recognition experiments adopting and not adopting electrode calibration algorithm are carried out respectively under two network training strategies.

a) Data diversity strategy: The CNN+LSTM-based network is trained with multiple batches of data with different electrode shifts for enhancing the generalization ability. Among all batches of data, some of them are selected as training set, and the rest as test set. The cross-validation method is adopted when dividing the training set and test set. A small part of the training set (about 10%) is removed and used as the validation set when training the network.

b) Model update strategy: The CNN+LSTM-based network is trained with a batch of data, and then the parameters of LSTM block is updated or fine-tuned via post-shift data, so as to adapt to the change of data distribution brought by electrode shifts. Among all batches of data, the cross-validation method is adopted when determining the training set and test set. A small part of the training set (about 10%) is removed and used as the validation set when training the network. After the network is well-trained, the parameters of CNN block is frozen, different number of repetitions of each class in test set are used to fine-tune only the parameters of sequence model LSTM, and the remaining repetitions are used to evaluate the performance of the network. The Full Batch Learning [33] is adopted due to the small scale of the incoming data.

E. Performance Evaluation Index and Statistical Analysis

The Pearson correlation coefficient serves to quantify the similarity of extracted major patterns. The recognition accuracy of the gesture samples is used to evaluate the performance...
of hand gesture pattern recognition. The analysis of variance (ANOVA; SPSS 22, Chicago, IL, USA) is used for the statistical analysis of experimental results, and the significance level is set to 5%.

III. EXPERIMENTAL RESULTS

A. The Results of Major Pattern Extraction Based on ICA Analysis

The mean VAF values of the FastICA analysis for data from all available analysis windows of 9 gestures and 11 participants in both two experiments are analyzed. The VAF values of all participants reaches 95% when the number of source signals is set to 4 for both two grids. In a few cases, the VAF values exceeds 95% when the number is set to 3, but the standard deviations are greater than 1%. Therefore, 4 source signals are thought to be sufficient to describe the HD-sEMG signals for both two grids in the 9-gesture task. Consequently, all the ICA-based decompositions are conducted with 4 source signals in the following section.

The representative ICA decomposition results of No. 1 gesture of Participant 2 in S-ES experiment are shown in Fig. 4. These results are from three analysis windows belonging to different repetitions. For each analysis window, the mixed coefficient vector maps and the corresponding source signals are given. The source signal with the highest two norm energy is considered as the main contribution source, and the corresponding mixed coefficient vector is considered as the major pattern. It can be seen that although the energy differences between source signals are obvious, the major patterns are relatively fixed and consistent in the three analysis windows. The Pearson correlation coefficient between the three extracted major patterns are 0.9341 ((a), (b)), 0.9815 ((a), (c)), and 0.9107 ((b), (c)) respectively. In UN-ES experiment, the major patterns extracted from FE and FF respectively, and the correlation coefficients between the extracted major patterns are calculated in pairs among all analysis windows of the same gestures in one batch. The overall results of 11 gestures, two muscle groups and all data batches are shown in Fig. 5. It can be found that the mean correlation coefficients are about 0.8. Above results verify the consistency and rationality of the major pattern extraction based on ICA.

B. The Results of Electrode Calibration Based on Core Activation Regions of Muscle

In this study, it is found that 4-channel × 4-channel could cover core activation regions in most cases, so this size of sliding window is used in the implement of the proposed the calibration algorithm. Taking No.1, No.3, and No.7 gestures of Participant 2 as examples, Fig. 6 illustrates the extracted major patterns and core activation regions from one representative analysis window under each known electrode shift in S-ES experiment. In the figure, each row represents the results of one gesture. The first column represents the electrode Baseline (BL) position, and the last four columns represent the four electrode shift conditions, namely towards medial-distal (MD), lateral-distal (LD), medial-proximal (MP), and lateral-proximal (LP) directions with the distance of an inter-electrode interval respectively. The core activation regions are encircled with 4 × 4 red rectangle boxes, and the relative shift (Δ) to Baseline is marked below the subfigure in the form of the direction and stride. Form Fig. 6, it can be found that: 1) For the same gesture, the shifts of the core activation regions are consistent with the known electrode shifts; 2) Under one known electrode shift condition, the core activation regions of three gestures have the same shift direction. These experiment
results verify the effectiveness of the core activation region extraction algorithm.

Similarly, Fig. 7 shows the extracted major patterns and core activation regions of No.1, No.4, and No.6 gestures of Participant 2 in UN-ES experiment. For one gesture, the results of two electrode grids (FE and FF) are given respectively, and the shifts of the two electrode grids are independent of each other. For each data batch, the extracted major pattern of one representative analysis window is given. In the major pattern, the core activation regions are encircled with 4 red rectangle boxes, and the relative shift (Δ) to Batch 1 is marked below the subfigure in the form of the direction and stride. From Fig. 7, it can be observed obviously that, for the same gesture, the core activation regions are different among different batches, which reflects that shifts brought by the replacement of the electrode arrays. For different gestures, taking Batch 1 as reference, the shift direction and stride of the core activation regions extracted form data of the same batch are consistent with each other, which demonstrates that electrode shifts in FE or FF can be tracked clearly according to the change of core activation regions. This research further verifies the validity of the major pattern extraction and core activation region location method proposed in this study.

C. The Results of Gesture Recognition

1) The Determination of the Structure of CNN+LSTM-Based Network: The specific network structures are designed for the two experiments respectively. The structure of two types of networks are both determined by the training set and validation set from Participant 2. Table I shows the structure and parameters of CNN+LSTM-based network obtained in two types of experiments. We have tried one, two, and three convolution layers, and found that double convolution layers can well meet the needs of high recognition rate and low computation cost. For the number of filters in each convolution layer of CNN blocks, we have tried the combinations of 64-32, 32-32, 32-16, 16-16, and 16-8 for the double convolution layers, and each filter is in the size of 2 × 2. In the absence of the electrode calibration algorithm, the scale of the sEMG map is 4 × 8 × 6 for S-ES experiment or 4 × 4 × 8 for UN-ES experiment. It is found that the combinations of 16-16 is usually the optional selection for S-ES experiment, and 32-16 is the optional selection for UN-ES experiment. In the two types of networks, the 1st and 2nd convolution layers consist of the filters with stride of 1, no padding, and the 1st and 2nd max pooling layers consist of one filter with stride of 1, no padding. The network is followed by a fully connected (FC) layer with 64 units and a Batch Normalization layer. The number of layers of LSTM is explored as well, and it is found that the one-layer LSTM is the optional selection. When the electrode calibration algorithm is operated, the scale of the sEMG map is 4 × 4 × 4 for S-ES experiment or 4 × 4 × 8 for UN-ES experiment. In this case, double convolution layers with the combination of filters of 16-8 are usually appropriate selection for both two types of networks. The details of the combination layer and max pooling layer are similar to that adopted in the absence of the electrode calibration algorithm except for the padding operation was needed. The one-layer LSTM is the optional selection. As for the dropout, 0.4 is usually the good choice in most cases. Additionally, the batch size of training set is usually set to 1/8 ~ 1/4 of the quantity of samples, and it is better to be power of two. The initial learning rate (Lr) is set to 0.001.

2) The Results of Gesture Recognition: The gesture recognition results in the two experiments under two network training strategies are presented in Fig. 8 and Fig. 9 respectively. For the data diversity strategy, a two-way ANOVA is performed to explore the influence of calibration algorithm and gesture on recognition accuracy; for the model update strategy, a two-way ANOVA is performed to explore the influence of calibration algorithm and gesture repetition on recognition accuracy. The results of the statistical analysis under two strategies are both shown in Table II. From the results shown in Fig. 8, Fig. 9, and Table II, the following phenomena can be observed:

First, the electrode calibration algorithm can improve significantly the recognition accuracy (p < 0.01). On the whole,
the recognition accuracies of using the calibration algorithm are significantly higher than those of not using the calibration algorithm. When the calibration algorithm is adopted, the average recognition accuracies are 93.47 ± 6.15% for S-ES experiment under data diversity strategy, 94.51 ± 4.56% for S-ES experiment under model update strategy, 90.64 ± 6.67% for UN-ES experiment under data diversity strategy, and 93.80% ± 4.21% for UN-ES experiment under model update strategy. When the calibration algorithm is not adopted, the average recognition accuracies are 84.97% ± 7.76%, 92.87% ± 5.94%, 83.86% ± 5.64%, and 87.64% ± 8.19% in four different cases.

Second, under data diversity strategy, the batch of data used to train the network has significant influence on the recognition accuracies in both two experiments ($p < 0.001$). In the situations of using or not using the calibration algorithm, the networks have different requirements on the number of data batches. With the calibration algorithm, the recognition accuracies are less affected by data batches. Even ONLY one batch is adopted, the average recognition accuracy can reach almost 90%. Without the calibration algorithm, recognition accuracies increase with data batches, even the network trained with four batches of data cannot approach to that trained with a batch of data using the calibration algorithm.

Third, under model update strategy, the number of gesture repetitions used to fine-tune the classifier has significant effect on the recognition accuracies in both two experiments ($p < 0.001$). In the cases of using or not using the calibration algorithm, the networks have different requirements on the number of gesture repetitions. With the calibration algorithm, the recognition accuracies are less affected by gesture repetitions. Even ONLY one repetition of each class is adopted, the average recognition accuracy of more than 90% can be obtained. When the number of gesture repetitions is greater than 2, the stable recognition accuracies can be reached. Without the calibration algorithm, recognition accuracies increase with gesture repetitions. Only when the number reaches 7, the recognition accuracies approach those obtained using the calibration algorithm.

IV. DISCUSSION

HD-sEMG has been adopted widely for gesture pattern recognition due to its innate advantages in overcoming electrode shifts. In this study, we innovatively propose a HD-sEMG electrode array calibration algorithm by aligning the core activation regions of muscle for high accuracy recognition of hand gestures. As far as we know, this is the first work to combine the characteristics of muscle activation with sEMG images for overcoming the bad effect of electrode shifts on myoelectric pattern recognition. Combined with relevant researches, the feasibility, superiority and limitations of the proposed calibration algorithm are discussed as follows.

A. The Effectiveness of the Proposed Calibration Algorithm in Mitigating the Bad Effect of Electrode Shifts

Based on the experimental results of supervised electrode shift (S-ES) experiment and unsupervised non-rotational electrode shift (UN-ES) experiment, the effectiveness of the proposed calibration algorithm in mitigating the bad effect of electrode shifts can be understood from the two aspects: 1) the extracted core activation regions of muscle among repetitions of the same gesture are similar regardless of electrode shifts, thus the electrode calibration can be achieved by aligning the core activation regions; 2) the electrode calibration can improve significantly the accuracy of gesture recognition.

The experimental results in Fig.4 ~ Fig.7 verify that, regardless of the electrode shifts, the ICA-based adaptive electrode calibration algorithm can extract relatively stationary and consistent core activation regions of muscles from repetitions of the same gesture. In specific, Fig. 4 and Fig.5 verified the consistency of the extracted major patterns of the same gesture with about 0.8 of mean correlation coefficients. Fig. 6 shows that the changes of direction and stride of the extracted core activation regions are consistent with those of the known electrode shifts. Fig. 7 demonstrates that the spatial distributions of the core activation regions of the same gesture from different data batches have certain similarity, and the shifts for different gestures are basically consistent with each other in the same batch. The experimental results in Fig. 8 and Fig.9 demonstrate that the proposed calibration algorithm can improve significantly the performance of gesture recognition in both S-ES experiment and UN-ES experiment. With
the electrode calibration, the overall recognition accuracies have an increase of (5.72~7.69) %. In specific, the average recognition accuracy increases (13.32~17.30) % when using only one batch of data in data diversity training strategy, and increases (12.01~13.75) % when using only one repetition of each gesture in model update training strategy.

In related researches, the spatial information of the HD-sEMG images is usually adopted to calibrate electrode shifts by proposing new features not sensitive to the shifts [17], [18], or using the specific template extracted from training images to detect or match the target in testing images [20]. For instance, Stango et al. [17] proposed a variogram feature to weaken the effects of electrode shifts. For nine-gestural recognition task, the average classification accuracy of 90% was obtained, with an increase of about 10% compared with other classical time domain features. He et al. [18] used gray-level co-occurrence matrix to represent the spatial distribution. For a ten-class task, the average recognition accuracy was about 95%, which increased about 15% compared with other time and frequency domain features. Zhang et al. [20] adopted specific template extracted from training images to overcome electrode shifts. For the six-gesture task, around 95% classification accuracies were obtained.

Only from the improvement of recognition accuracy, the proposed method does not seem to be better than the related researches. However, compared to the target gestures related to the movement of wrist joint and elbow joint in the literatures [17], [18], [20], the 9 gestures in this study involve fine movements of finger joints, which make the recognition task more challenging. The satisfactory performance (92.17% recognition accuracy) certainly indicates the effectiveness of the proposed method. In addition, relevant studies often used adjacent channels of HD-sEMG grid to simulate electrode shifts [17], [18]. In other words, the experimental data were recorded simultaneously in the same trials. In this study, the electrode array was completely removed and then replaced between different trials in both S-ES experiment and UN-ES experiment, which may incur much data variations related to the activation states and the mode of power generation and bring more challenges to the recognition task. Obviously, the schemes conducted in this study is closer to the real application scenarios of myoelectric control systems.

B. The Potentials of the Proposed Calibration Algorithm in Reducing User Training Burden of Myoelectric Control Systems

The two types of training strategies including data diversity strategy and model update strategy are often adopted in related researches to reduce the effect of electrode shifts and improve the robustness of classifier. In summary, enough training data should be collected for both the two training strategies, which can bring certain training burden for the user and is unfriendly to the practical application of myoelectric control system. In this study, we also adopted these two types of training strategies and found the potentials of the proposed calibration method in reducing user training burden.

The characteristic of the data diversity strategy is that the training dataset should contain data from multiple batches to cover various electrode shifts. In the works of Hargrove et al. [7], [8], for a ten-gesture recognition task, more than 90% recognition accuracy was obtained when the training set consists of data from four plausible shift locations. Compared to that obtained using training data from one position, a relative increase of over 20% was obtained. In this study, the experimental results also verify the importance of the diversity of training data when not using the calibration algorithm. As shown in Fig. 8, the recognition accuracy increases significantly with the number of data batches in training set. However, with the proposed calibration algorithm, the recognition accuracies are less affected by data batches. Even only one batch is adopted, the average recognition accuracy reaches almost 90%. This result demonstrates that the proposed calibration method can reduce significantly the requirement of data batches in the training phase.

For the model update strategy, a certain quantity of gesture repetitions are needed for the fine-tuning of the model parameters. For a recognition task of eight movements, Vidovic et al. [9] proposed a mixed-LDA to compensate for the effect of electrode shifts. The distribution parameters of pre-trained LDA model were updated by some post-shift samples. The classification accuracy attained above 92% for incoming data with calibration. Zhang et al. [10] proposed a semi-supervised adaptation strategy to improve the classification performance for ten arm- and hand-movements. After the LDA was well-constructed by the training set, some testing data and the corresponding prediction class were collected regularly in the testing phase, and then replaced a part of samples of the training set. The classifier was retrained periodically so that it can learn the new distribution information of post-shift data. It was found that a substitution proportion over 40% could provide a good classification performance up to 95%. In this study, as shown in Fig. 9, the proposed calibration method can reduce significantly the requirement of gesture repetitions when the model update strategy was adopted. With the proposed calibration algorithm, the recognition accuracies are less affected by gesture repetitions. Even only one repetition of each class is adopted, the average recognition accuracy of almost 92% with the improvement of 10% is obtained for two experiments.

From the perspective of user acceptance, it is better to use none or only a small number of post-shift samples to train or fine-tune the classifier. The proposed algorithm can be realized in an unsupervised and self-adaptive way, and is characterized by low requirements for multiple batches of training data or post-shift samples, which has good practical application value in myoelectric control systems.

C. Limitations and Future Work

We would like to point out the limitations of current study. First, the electrode calibration algorithm proposed in this paper cannot be applied to the case of electrode shifts in rotation. The algorithm suitable for rotational electrode shifts should be explored in future work to really realize the calibration of any unsupervised shift. Second, the feasibility and superiority of the proposed electrode calibration algorithm only were verified
on a nine-gesture task. The follow-up work should continue to extend the application of the algorithm to more gestures. Third, the proposed method was only evaluated in the offline analysis. The calibration algorithm needs further optimization to ensure the feasibility of its application in online analysis. Finally, only FastICA is adopted to extract muscle activation modes or contribution sources from HD-sEMG signals in this study. We will continue to explore if the electrode calibration can be performed with other decomposition methods such as NMF and PCA.

V. CONCLUSION

The main contribution of this paper is presenting an adaptive electrode calibration method based on core activation region of muscles for HD-sEMG based myoelectric pattern recognition. The experimental results of supervised electrode shift experiment and unsupervised non-rotational electrode shift experiment demonstrate that, the proposed calibration algorithm has the advantages of mitigating the bad effect of electrode shifts on gesture recognition accuracy, and the potentials to reduce user training burden of myoelectric control systems. The proposed electrode calibration algorithm can be applied to improve the robustness of myoelectric control system, and further employed in the fields of myoelectric prosthetic, exoskeleton devices, and remote control devices.

REFERENCES
