Biomedical Signal Processing and Control 75 (2022) 103615

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Evaluation of decomposition parameters for high-density surface electromyogram using fast independent component analysis algorithm

Long Meng [a], Qiong Chen [a], Xinyu Jiang [a], Xiangyu Liu [b], Jiahao Fan [a], Chenyun Dai [a,*], Wei Chen [a,*]

[a] Center of Biomedical Engineering, School of Information Science and Technology, Fudan University, Shanghai 200438, China
[b] College of Communication and Art Design, University of Shanghai for Science and Technology, Shanghai 200093, China

ARTICLE INFO

Keyword:
Biosignal processing
Surface electromyogram (sEMG)
sEMG decomposition
Blind source separation
Independent component analysis

ABSTRACT

The surface electromyogram (sEMG) signal reveals the electrical neuromuscular activities and offers theoretical and clinical information. Recently, exploring motoneuron discharge events at the microscopic level through decomposition has been proposed as a promising analysis approach, which can surpass traditional global sEMG-based analysis method in some aspects. However, computational efficiency is an essential issue when discharge events of individual motor unit are decomposed using blind source separation algorithms. Therefore, choosing proper parameters of decomposition algorithms for different research purposes is important. Accordingly, we have systematically investigated the influences between computation time and decomposition performance for fast independent component analysis (FastICA)-based sEMG decomposition algorithm under different value selections of five decomposition parameters, namely the percentage of eliminated channels, extension factor, the number of decomposition iteration loops, the number of maximum inner loops in each iteration and sampling frequency. We employed high-density sEMG signals from 14 intact subjects during muscle contractions of four-digit extension and flexion at different force levels (20% and 50% maximum voluntary contraction) and a public dataset for sEMG decomposition. According to obtained results, we offer four preference suggestions (less computation time, more motor units, higher accuracy and trade-off). Results show that the trade-off values with consideration of decomposition performance and computation time are recommended as 25% of channels with minimal root mean square, extension factor of 4, 200 iteration loop numbers of decomposition, 20 maximum inner loop numbers and 2048 sampling frequency. Overall, this paper provides a guide for researchers to determine proper decomposition parameters for sEMG decomposition-related works.

1. Introduction

Surface electromyogram (sEMG) signal is the summation of hundreds of motor unit action potential (MUAP) trains originating from motor unit (MU) discharges during voluntary contractions. The global sEMG signal, which can reflect the motor control information from neural activities, has been widely applied in a variety of fields, such as the control of human–machine systems [1,2], ergonomics [3,4], sports science [5,6], rehabilitation [7,8], neuromuscular user authentication [9], etc. However, the continuous development of the use of global sEMG signal is limited due to its intrinsic defects, such as signal cross-talk generated from adjacent muscles or muscle compartments [10], signal amplitude cancellation caused by the superimposition of MU action potential waveforms [11]. These issues lead to the inaccurate interpretation of neural information in the convolutive sEMG signals. In contrast, with the development of flexible high-density sEMG (HD-sEMG) electrode array techniques, Farina et al. have made a breakthrough of sEMG-based analysis through shifting the neuromuscular information mining approach from the macroscopic level (via global sEMG) to the microscopic level (via motoneuron discharge events decomposed from HD-sEMG) [12]. The motoneuron-based analysis uses binary MU spike trains to represent motoneuron discharges, thereby reflecting the motor control information [13]. This novel approach can overcome the inherent disadvantages of the macroscopic counterpart by

This work is supported by National Natural Science Foundation of China (Grant Number: 62173094), Shanghai Municipal Science and Technology International R&D Collaboration Project (Grant No. 20510710500) and Natural Science Foundation of Shanghai (Grant Number: 20ZR1403400).

* Corresponding authors.
E-mail addresses: chenyundai@fudan.edu.cn (C. Dai), w_chen@fudan.edu.cn (W. Chen).

https://doi.org/10.1016/j.bspc.2022.103615
Received 29 September 2021; Received in revised form 23 February 2022; Accepted 27 February 2022
Available online 7 March 2022

eliminating the influences generated from MU potential waveforms [14]. Therefore, it has been considered to be a promising approach for the development of sEMG-based applications in the future. Previous studies have shown that motoneuron-based analysis can surpass global sEMG-based analysis in different application scenarios. For engineering studies, numerous recent studies have established the findings that the accuracy of using neuromuscular information from motor units is higher than that of directly using global sEMG information for force control [15] or joint angle estimation [16]. For physiological studies, the motoneuron discharge events can represent the activities of neural control information from a higher level (e.g. spinal cord) for related researches, such as exploring the MU short-term synchronization in diverse contraction tasks [17], the changes of MU discharge properties after strength training [18], etc.

Although the microscopic analysis has shown great advantages, an obvious shortcoming cannot be ignored. To analyze myoelectric signals at the microscopic level, discharge events of individual MU need to be decomposed from multi-channel sEMG signals using blind source separation algorithms. Although accurately decomposing global sEMG into MUAP is crucial, the computational efficiency (balance of computational cost and decomposition performance) of the signal decomposition algorithms is an essential issue for practical applications. To support applications of sEMG decomposition techniques in as many scenarios as possible, we mainly aim to investigate the trade-off between decomposition performance and computation time in this work. Most recent studies employed the fast independent component analysis (FastICA) method due to its relatively high computational efficiency, compared with other approaches [19,20][21][22]. The high computational efficiency of FastICA algorithm can be further validated by the comparisons with the decomposition performance of two other variants of independent component analysis (ICA), i.e., Information Maximization (InfoMax) [22] and second-order blind identification (SOBI) [23] in this study.

The FastICA algorithm involves six main steps: (1) channel selection, (2) signal extension, (3) whitening, (4) FastICA-based deconvolution, (5) peak detection through clustering, and (6) duplicated MU removal. The decomposition performance (both time consumption and accuracy) highly depends on the selection of parameters in six aforementioned steps [20]. Choosing a different parameter value that improves the decomposition accuracy, may result in huge computational costs. For example, increasing the signal extension factor can slightly improve both the accuracy of decomposition and the number of decomposed MUs [19,20]. However, the corresponding computation time grows substantially when the signal extension factor ascends. Therefore, we need to find a trade-off between the decomposition accuracy and the computational costs according to the specific tasks.

To systematically investigate the influences of selecting decomposition parameter values for the decomposition accuracy and computation time, we used the following approaches:

1. Five main parameters (the percentage of channels eliminated, extension factor, the number of decomposition iteration loops, the number of maximum inner loops in a single MU deconvolution iteration, and sampling rate) were selected for evaluation.
2. The decomposition results were tested with different parameter values on experimental HD-sEMG data from 14 intact subjects with different muscles and force levels. Four metrics from different perspectives were used to evaluate the decomposition performance. The corresponding computation time was also recorded.
3. We gave suggestions on the selection of proper parameters for different research aims.

By studying the effect of tuning different parameters for different muscles and tasks, we can provide a deep insight into the mechanism on how different parameters impact the final decomposition accuracy and computation time, which in turn can help readers customize their own parameter selection according to the unique requirements in their own application scenarios, such as physiological analysis at MU level [18], prosthetic control [23], gesture recognition [24], etc.

2. Materials and Method

2.1. Private Dataset

2.1.1. Subject

Fourteen healthy subjects (ten males, four females; aged from 21 to 31) without known neurological disorders or injuries were recruited for the sEMG–force data acquisition. All subjects provided written informed consent. The study was approved by the Institutional Review Board of Fudan University (BE2036).

2.1.2. Experimental Setup

The subject was seated in an experimental chair with his of her elbow supported on a soft pad and the dominant forearm comfortably put on the experimental table. The entire forearm and wrist were fixed between a couple of padded boards in a zero degree flexion–extension and neutral pronation-supination position to minimize wrist movements. The four fingers including index, middle, ring and pinky were abducted comfortably, with each finger secured to one load cell (Interface, SM-200 N, with resolution of 0.01 N) using a padded Velcro strap (fixed at the middle and proximal phalanges). In the experiment, the participants were divided into two groups. Six males and two females, who participated in group one, were asked to isometrically extend four fingers in each trial. In contrast, the other four males and two females, who participated in group two, were asked to isometrically flex four fingers in each trial. At the same time, the finger force data were recorded with a sampling rate of 1 kHz. The force output was also displayed through a custom-built MATLAB program (MathWorks Inc.) as a real-time feedback to the subject. We began by measuring the maximum voluntary contraction (MVC) of four-finger extension (or flexion). The subject was asked to extend (or flex) the four fingers with maximum force and maintain approximately 3 s. The force values during the contraction plateau of 2 s were averaged as the MVC.

During the data collection session, a 20-s sin-wave target force trajectory was provided on a screen. Subjects in group one were asked to control the force of four-digit extension to track the target with two different force level contractions (20% and 50% MVC). Subjects in group two were asked to control the flexion force of four-digit with one force contraction (50% MVC). For the sine-wave extension or flexion force contraction, subjects took 3 s to ramp up to 20% or 50% effort, maintained for 5 s, then performed three 3-s repeated sin-wave force within 10% to 20% MVC or 25% to 50% MVC, and finally ramped down to relaxed condition with 3 s. To avoid the issues from muscle fatigue, a 2-min rest was provided between two successive trials. Five repetitions were performed for each subject. Therefore, a total of 110 trials (group one: five repetitions $\times$ two force levels $\times$ eight subjects $= 80$ trials; group two: five repetitions $\times$ one force level $\times$ six subjects $= 30$ trials) were recorded for further analysis.

2.1.3. sEMG Recordings

sEMG signals were obtained simultaneously via an 8×20 (for extension tasks) or 8×16 (for flexion tasks) monopolar HD-sEMG electrode grid with 10 mm inter-electrode distance (OT Bioelettronica, Torino, Italy), shown in Fig. 1. The reference electrode was placed on the

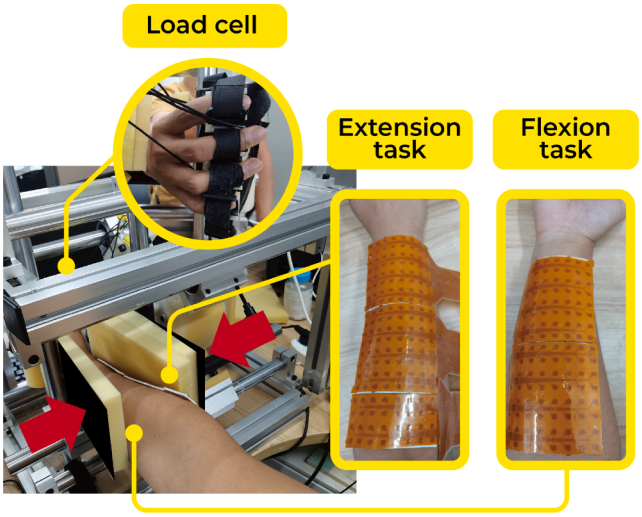

Fig. 1. Electrode placement and exemplar experimental setup.

elbow. The long side of the electrode grid was aligned to the muscle fibres. Since two different tasks (four-finger extension and flexion) were implemented in this work, the electrode array was gelled and placed on the extensor digitorum communis (EDC) muscle or flexor digitorum superficialis (FDS) muscle for four-finger extension or flexion tasks. Before the electrode placement, the skin surface was cleaned with alcohol pad to minimize the skin impedance. The sEMG signals were recorded from Quattrocento system (OT Bioelettronica, Torino, Italy) with a gain of 1000, a sampling rate of 2048 Hz, and a bandpass filter with bandwidth at 10–900 Hz.

2.2. Public Dataset

To validate the generalization capacity of selected parameters, besides the designed data acquisition in this study, we also evaluated the decomposition parameters on a public dataset: SurfacE Electromyographic with hanD kinematics (SEEDS) [25]. For the SEEDS dataset, 25 non-disabled subjects (11 males, 14 females; aged from 21 to 55) participated in the experiment to perform 13 different hand movements. A total of 450 trials (25 subjects $\times$ 3 sessions/subject $\times$ 6 repetitions/session) were collected for each movement. Half of the repetitions were performed at a controlled slow speed and the other half at a normal pace. For the data collection, HD-sEMG data were acquired with a 126 (9 $\times$ 14) high-density electrode array on the flexor muscle, and finger-joint angles were recorded with a data-glove (CyberGlove III, Cyber Glove Systems LLC, San Jose, CA, USA). The sampling rates of HD-sEMG and kinematic data available in SEEDS dataset were 2048 Hz and 256 Hz, respectively. For the preprocessing, HD-sEMG data were filtered with a zero-phase 3^{rd} order band-pass Butterworth filter (cut-off frequency: 10–500 Hz), and kinematic data were linearly normalized in [0,1] using Min–Max scaling approach for each participant. In this study, the experiment was conducted with the acquired HD-sEMG data to estimate the metacarpophalangeal joints of two representative movements (1. fist, 2. thumb, index and middle fingers pinch).

2.3. sEMG Decomposition

Algorithm 1 shows the steps of the FastICA-based sEMG decomposition method with the parameters to be explored shown in red. Given an observed HD-sEMG signals $\boldsymbol{X}_{\mathrm{O}} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_i, \ldots, \boldsymbol{x}_{N_{\mathrm{O}}}) \in \mathbb{R}^{N_{\mathrm{O}} \times N_{\mathrm{S}}}$, where $\boldsymbol{x}_i \in \mathbb{R}^{1 \times N_{\mathrm{S}}}$ denotes the sEMG data acquired with the i-th channel; N_{O} denotes the number of channels; and N_{S} denotes the number of samples for each channel. The details of the proposed algorithm are described in the following steps:

Algorithm 1: Steps of sEMG decomposition algorithm. The parameters investigated are marked in red.

Algorithm 1: Steps of sEMG decomposition algorithm. The parameters investigated are marked in red.

Input: HD-sEMG data: $\boldsymbol{X}_{\mathrm{O}}$
Result: Spike trains with good quality
Step 1: Eliminate N_e of original channels with minimal RMS, and get $\boldsymbol{X}_{\mathrm{R}}$
Step 2: Extend channels of $\boldsymbol{X}_{\mathrm{R}}$, and get $\boldsymbol{X}_{\mathrm{E}}$
Step 3: Whiten the extended signals $\boldsymbol{X}_{\mathrm{E}}$, and get $\boldsymbol{X}_{\mathrm{W}}$
Step 4: FastICA-based deconvolution:
for m=1:M **do**
  Initialize $\boldsymbol{w}_m^{(1)}$ and $\boldsymbol{w}_m^{(2)}$ with random numbers
  for n=2:NN **do**
    if $\left|\boldsymbol{w}_m^{(n-1)\top} \cdot \boldsymbol{w}_m^{(n)} - 1\right| > Threshold$ **then**
      $\boldsymbol{w}_m^{(n)} = \boldsymbol{X}_{\mathrm{W}} \cdot g\left(\boldsymbol{w}_m^{(n-1)} \cdot \boldsymbol{X}_{\mathrm{W}}\right) - A \cdot \boldsymbol{w}_m^{(n-1)}$
      with $A = E\left[g'\left(\boldsymbol{w}_m^{(n-1)\top} \cdot \boldsymbol{X}_{\mathrm{W}}\right)\right]$
      Orthogonal projection:
      $\boldsymbol{w}_m^{(n)} = \boldsymbol{w}_m^{(n)} - \boldsymbol{B}\boldsymbol{B}^{\top}\boldsymbol{w}_m^{(n)}$
      Normalization: $\boldsymbol{w}_m^{(n)} = \frac{\boldsymbol{w}_m^{(n)}}{\left\|\boldsymbol{w}_m^{(n)}\right\|}$
    else
      break;
    end
  end
  $\boldsymbol{B} = [\boldsymbol{B}, \boldsymbol{w}_m]$
end
$\boldsymbol{S} = \boldsymbol{X}_{\mathrm{W}}^{\top} \cdot \boldsymbol{B}$
Step 5: Cluster peaks: spike trains = kmeans($\boldsymbol{S}$)
Step 6: Regulate time delay and remove duplicates or MU spike trains with low silhouette measure (SIL) values

1. Channel selection. Exclude N_e of channels with minimal root mean square (RMS) values from total channels of $\boldsymbol{X}_{\mathrm{O}}$, and get the channel-removed sEMG signals $\boldsymbol{X}_{\mathrm{R}} \in \mathbb{R}^{N_{\mathrm{R}} \times N_{\mathrm{S}}}$, where N_e denotes the percentage of original channels, N_{R} denotes the number of channels after the channel selection step, and $N_{\mathrm{R}} = N_{\mathrm{O}} \times (1 - N_e)$. The reason for conducting this step is that the channels with a low signal-to-noise ratio may have few contributions to decomposition accuracy but increase the computation costs. An example of the RMS heat map during muscle contraction is shown in Fig. 2.
2. Channel extension. Duplicate the remaining channels by adding R delays of the original channels, and get the extended sEMG signals $\boldsymbol{X}_{\mathrm{E}} \in \mathbb{R}^{N_{\mathrm{E}} \times N_{\mathrm{S}}}$, where N_{E} denotes the number of channels after the channel extension step, and $N_{\mathrm{E}} = N_{\mathrm{R}} \times (1 + R)$. To be specific, samples of each channel are delayed by 1 sample, 2 samples,…, and R samples to construct R new channels. The extended sEMG signals $\boldsymbol{X}_{\mathrm{E}}$ are then formed by combining data of $\boldsymbol{X}_{\mathrm{R}}$ and R new channels.
3. Signal whitening. Whiten the extended sEMG signals $\boldsymbol{X}_{\mathrm{E}}$ to remove the correlation of observations, and get the whitened sEMG signals $\boldsymbol{X}_{\mathrm{W}} \in \mathbb{R}^{N_{\mathrm{E}} \times N_{\mathrm{S}}}$.
4. FastICA-based decomposition. Perform a total of M iterations. One source signal can be decomposed in each iteration. The FastICA method based on negative entropy maximization executes the following steps in each loop:

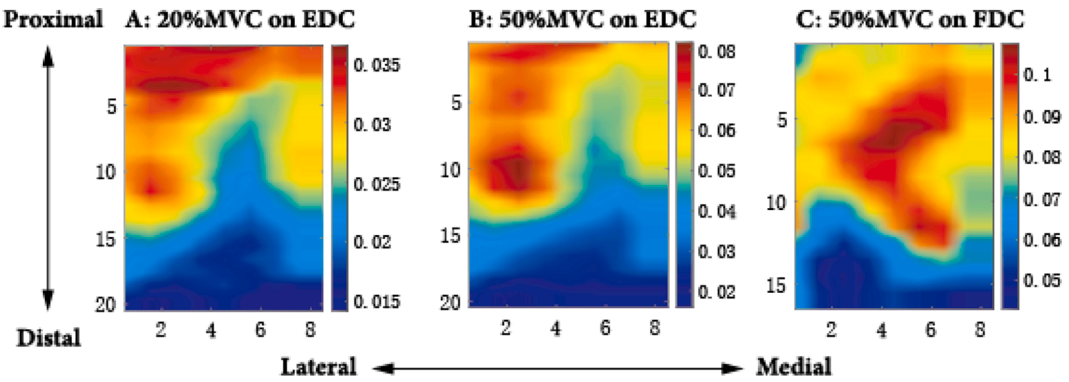

Fig. 2. Examples of the energy distribution of the high-density sEMG during muscle contractions.

- For m-th loop, the inner loop continues if either the inner product of separation vector$\boldsymbol{w}_m^{(n-1)}$ and $\boldsymbol{w}_m^{(n)}$ minus one (representing the convergence condition to maximize non-gaussianity of source) exceeds a very small threshold (the threshold is set as 10^{-4} referred to [19]), or before the end of the loop NN, where $\boldsymbol{w}_m \in \mathbb{R}^{N_E \times 1}$ is the m-th separation vector; n is the inner iteration number; and NN is the maximum number of inner iteration.
- While maximizing the non-guassianity, the contrast function $G(x)$is used to accelerate convergence speed. In this study, the $G(x)$was set as $\frac{1}{3}x^3$ referred to [19]:

$$g(x) = \frac{dG(x)}{dx} \tag{1}$$

where $x$in $G(x)$represents the independent variable of contrast function. Then $\boldsymbol{w}_m$ is updated by orthogonal projection and normalization to avoid converging to the same MU source.

5. Peak detection. After M iterations, M source signals S can be obtained through multiplying the optimal separation matrix ($B = [w_1, w_2, \ldots, w_M]$) by whitened sEMG signals X_W, where $B \in \mathbb{R}^{N_E \times M}$ and $S \in \mathbb{R}^{N_S \times M}$. For each source signal, two obvious clusters can be found in each source signal after decomposition. One class with larger amplitude represents the MU discharge events, and another class with smaller amplitude represents baseline noise in the decomposed source. Thus, the class with higher peak values can be regarded as discharge events of one MU. The discharge events of each decomposed MU are represented as a binary spike train.
6. Duplicated MU removal. In the above 5 steps, the decomposed MUs can be identical due to the extension step. Hence, the duplicated MUs are removed if any pair of MU spike trains have more than 50% of synchronized spike events within ±1ms after regulating the time delay. Besides, silhouette measure (SIL), which has a strong correlation with decomposition accuracy [20], is used to remove the decomposed MU with low accuracy on the estimated sources. The SIL is calculated as the mean of distances of each point between within-cluster distance summation and between-cluster distance summation. If SIL is greater than a threshold (set to 0.7 based on our previous studies [20,22]), the decomposed MU is automatically reserved.

2.4. Validation Methodology

The five parameters have been investigated in this study as mentioned above, namely N_e (percentage of eliminated channel numbers with minimal RMS), R (extension factor, indicating the number of delayed replicas of original channels), M (the number of MU deconvolution iteration), NN (the number of maximum inner loops in a single MU deconvolution iteration), and Fs (sampling rate). Based on our previous experience [20,22], the parameter values were selected as follows: (1) the percentage of excluded channels N_e was selected as 10%, 25% and 50%, (2) the extension factor R ranged from 0 to 8 with an interval of 2, (3) the number of MU deconvolution iteration M ranged from 100 to 300 with an interval of 50, (4) the number of maximum inner loops NN ranged from 20 to 80 with an interval of 30, and (5) two different sampling rate Fs values—1024 and 2048 Hz. For each trial, all possible parameter combinations were tested. All parameters were performed without interruptions from other programs under MATLAB R2019b (MathWorks Inc.), and Intel Core i9-9960@3.10 GHz with 64 GB of memory environment.

To evaluate the decomposition performance, five metrics were used including computation time, the number of detected MUs (NumMU), SIL, coefficient of determination (R^2) and root-mean-square error (RMSE). The computation time and number of detected MUs are two objective metrics and can be directly measured through decomposition results. The computation time was defined as the time consumption of step 1 - step 6 of FastICA decomposition algorithm on the 20-s signal. Besides, since the decomposition results of experimental data had no ground truth, the decomposition accuracy was measured by three indirect measures-SIL, R^2 and RMSE. For the SIL measurement [26], it has been proved to have a positively linear relationship with decomposition accuracy and widely used as the accuracy metric for experimental data. Hence, the mean SIL of decomposed MUs was adopted. For the R^2 and RMSE measures, they represented the degree of correlation and regression error between MU-based estimated finger force (sEMG–force data)/joint angle (SEEDS dataset) and measured ground-truth force/angle using linear regression. Previous study [12] have proved that the discharge frequency of pooled MU spike trains after HD-sEMG decomposition could be used for force estimation. Based on our previous study [15], higher R^2 and lower RMSE values were strong (although indirect) indicators of better decomposition performance, and the operation procedures of the two measures were performed the same as [15].

For a clear visualization, we plotted five metrics for each parameter by averaging corresponding metrics obtained from all combinations of the other four parameters. Since the scales of five metrics were different and the trends of metrics were with the top priority, we rescaled all metrics in [0,1] using Min–Max scaling approach by $Metric_{norm} = \frac{Metric - Metric_{min}}{Metric_{max} - Metric_{min}}$, where $Metric$is the original metric values we obtained, $Metric_{max}$ represents the maximum value of all metric values obtained from all corresponding trials, and $Metric_{min}$ denotes the minimum value of all metric values obtained from all corresponding trials.

Table 1
Normalization factors of results.

	NumMU	MeanSIL	RMSE	R^2	Time (s)
Minimum value	1	0.7002	0.0051	0.00059	1.6546
Maximum value	42	0.9761	0.4029	0.9858	832.37

2.5. Statistical Analysis

Since the obtained results did not follow Gaussian distribution, we chose nonparametric Friedman's test to detect whether there was an overall significant difference [27]. If the overall significant difference was founded, post hoc analysis using Wilcoxon signed-rank test was applied for pair-wise comparison. To avoid multiple comparison error, Bonferroni correction was adopted. The significance value was set at $p < 0.05$.

3. Result

For the self-collected dataset, the overall results of different muscles and force levels are shown in Fig. 3., and that with SEEDS dataset are presented in Fig. 4. The corresponding normalization factors are presented in Table 1.

3.1. Results under 20% MVC on EDC Muscle with sEMG–force Data

The overall decomposition results of 20% MVC on EDC muscle are shown in Fig. 3 (a).

From the performance of eliminated channel numbers N_e in the first column of Fig. 3 (a), we can find that 1) the number of detected MUs slightly decreases from $N_e = 10\%$ to 25%, and then sharply decreases from $N_e = 25\%$ to 50%; 2) mean SIL and computation time tend to consistently decrease; 3) RMSE tends to keep unchanged from $N_e = 10\%$ to 25%, then increases from $N_e = 25\%$ to 50%; 4) R^2 slightly increases from $N_e = 10\%$ to 25%, then diminishes from $N_e = 25\%$ to 50%; The results indicate that removing channels with low RMS can greatly shorten computation time without sacrificing decomposition accuracy. Once the channels with high RMS are eliminated, the decomposition accuracy greatly deteriorates.

From the performance of extension factor R was investigated in the second column of Fig. 3 (a), we can find that 1) the number of decomposed MUs, R^2 and computation time increase with R^2 in a logarithmic shape. Accordingly, there are no obvious changes for the number of decomposed MUs and R^2 after $R = 4$; 2) mean SIL tends to linearly increase; 3) RMSE decreases with R in a logarithmic shape, and tends to be stable after $R = 4$; The results indicate that extending the number of observation channels at a small number can improve decomposition performance. When the observation channel numbers are large enough, the improvement of decomposition performance tends to slow down but the time consumption continues to increase.

From the performance of the number of decomposition iteration loops M in the third column of Fig. 3 (a), we can find that: 1) the number of decomposed MUs increases with M in a logarithmic shape, and tends to be stable after $M = 200$; 2) mean SIL slightly goes down from $M = 100$ to 150, then tends to be stable from $M = 150$ to 300; 3) RMSE and R^2 tend to be unchanged with M; 4) the computation time tends to linearly increase with M. The results indicate that the growing number of decomposition iteration loops facilitates the improvement of decomposed MU numbers, but substantially escalates time consumption. However, the overall decomposition performance (R^2 and RMSE) does not significantly improve with the mounting numbers of decomposition iteration loops.

From the performance of the number of maximum inner loops in a single MU deconvolution iteration *NN* in the fourth column of Fig. 3 (a), we can find that: 1) the computation time monotonically increases with *NN*; 2) the other four metrics remain no obvious changes as *NN* grows. The results indicate that the decomposition performances do not change with the rising number of maximum inner loops in a single MU deconvolution iteration, except for the growing tendency of computation time.

From the performance of sampling rate *Fs* in the fifth column of Fig. 3 (a), we can find that: with the *Fs* shifting from 1024 to 2048 Hz, RMSE significantly decreases. In contrast, the other four metrics (the number of decomposed MUs, mean SIL, R^2 and consumption time) greatly increase. The results indicate that increasing the sampling rate can improve the decomposition performance, except for the growing tendency of computation time.

3.2. Results under 50% MVC on EDC Muscle with sEMG–force Data

The overall decomposition results of 50% MVC on EDC muscle are shown in Fig. 3 (b). Since Fig. 3 (a) and Fig. 3 (b) are the results under the same part of muscle but with different forces, we can find that the tendencies of five metrics with the change of five parameter values in Fig. 3 (b) can be found similar to that in Fig. 3 (a). However, RMSE and mean SIL in Fig. 3 (b) have higher values than that in Fig. 3 (a). On the contrary, the number of detected MUs in Fig. 3 (b) is lower than that in Fig. 3 (a).

3.3. Results under 50% MVC on FDS Muscle with sEMG–force Data

The overall decomposition results of 50% MVC on FDS muscle are shown in Fig. 3 (c).

In Fig. 3 (c), the tendencies of five metrics with the change of three parameters (*M*, *NN* and *Fs*) can be found similar to the corresponding tendencies in Fig. 3 (a).

From the performance of eliminated channel numbers N_e in the first column of Fig. 3 (c), we can find that 1) RMSE monotonously increases with N_e; 2) R^2 tends to keep unchanged; 3) the tendencies of the other three metrics (the number of detected MUs, mean SIL, computation time) with N_e can be found similar to the corresponding three metric tendencies with N_e in Fig. 3 (a).

From the performance of extension factor R on FDS muscle shown in the second column of Fig. 3 (c), we can find that 1) there are no obvious changes of R^2 after $R = 2$ rather than $R = 4$ on the EDC muscle; 2) the tendencies of other four metrics (the number of decomposed MUs, mean SIL, RMSE and computation time) are similar to the corresponding metric tendencies with R in Fig. 3 (a).

3.4. Results of Finger Joint Movements on FDS muscle with SEEDS Dataset

Besides the decomposition parameters evaluation with the sEMG–force data, we conducted the same experiment procedure with the SEEDS public dataset to explore the validity of selected parameters in a different scenario (i.e., estimations of finger-joint movements). The sEMG decomposition results are presented in Fig. 4.

Regarding the sEMG decomposition performances on FDS muscle, we can observe that the tendencies of five metrics in Fig. 4 are highly consistent with that in Fig. 3 (c) under the varying values of N_e, R, *NN*

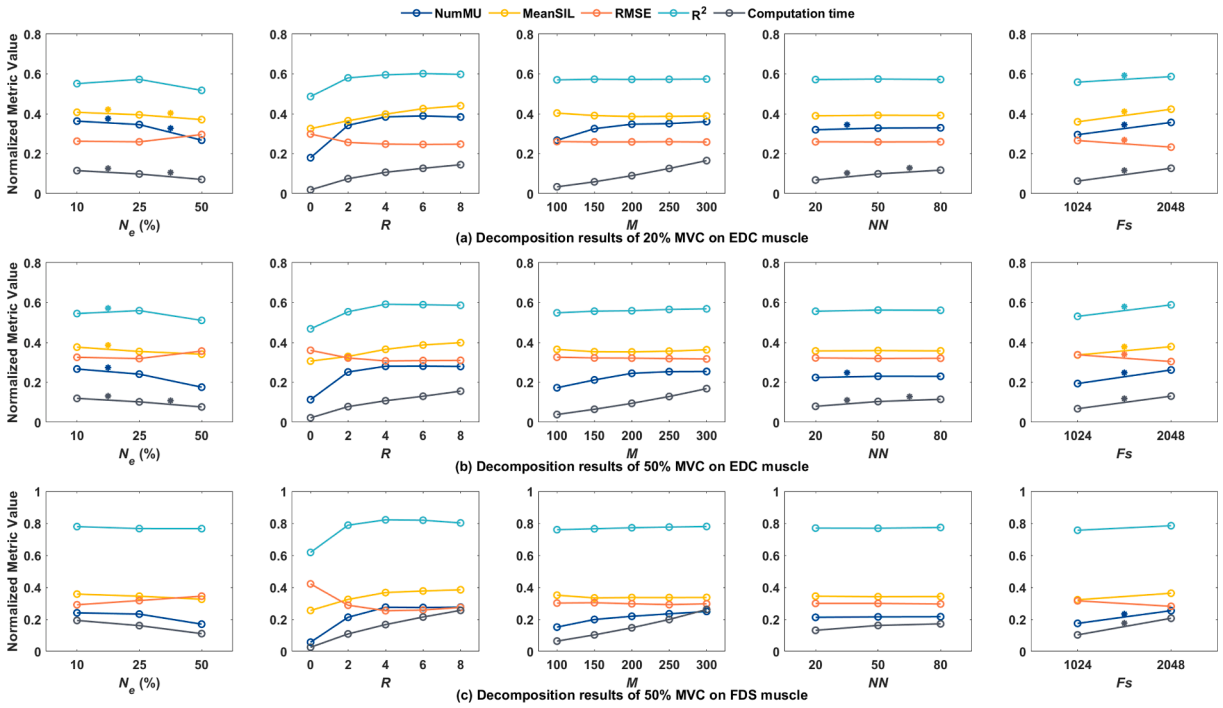

Fig. 3. Performances of five metrics for FastICA decomposition algorithm under five different parameters on different muscles and force. (a) decomposition results of 20% MVC on extensor muscle; (b) decomposition results of 50% MVC on extensor muscle; (c) decomposition results of 50% MVC on flexor muscle. Each marker represents the mean values of trials under specific parameter combinations. A star is marked between two adjacent metric values if a significant different ($p < 0.05$) is detected between the two values.

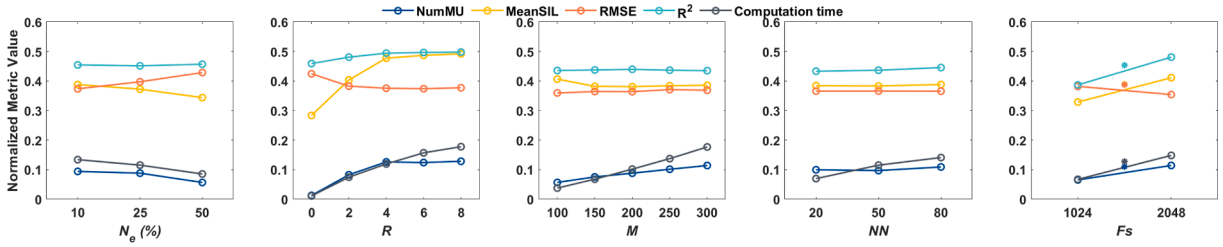

Fig. 4. Performances of five metrics for FastICA decomposition algorithm under five different parameters on FDS muscle. Each marker represents the mean values of trials under specific parameter values. A star is marked between two adjacent metric values if a significant difference ($p < 0.05$) is detected between the two values.

and *Fs*. From the performance of the number of decomposition iteration loops *M* in Fig. 4, tendencies of mean SIL, RMSE, R^2 and computation time can be found similar to that in Fig. 3 (c). In contrast, the number of detected MUs shows a logarithmic increase in Fig. 3 (c), but tends to linearly increases in Fig. 4.

4. Discussion

This work systematically evaluates the performance of FastICA-based HD-sEMG decomposition algorithm under the changes of five parameters on different muscles and tasks. We can find similar overall tendencies of five metrics on different muscles and tasks. The results on two muscles manifest that 1) excluding channels with low RMS makes a minor difference in decomposition accuracy, but can shrink the computation time. However, the decomposition accuracy can be seriously affected if channels with high RMS are removed; 2) increasing the extension factor *R* at a small number boosts the decomposition accuracy, but leads to escalating computation time. When the extension factor is greater than 4, the improvement of decomposition performance in most metrics is not obvious, but with computation time continuously increasing; 3) elevating the number of deconvolution iteration loops *M* has a minor influence on decomposition performance, which only leads to an expanding number of decomposed MUs and causes a growth of computation time; 4) the change of inner loop number in a single MU deconvolution iteration has little impact on decomposition accuracy, except for inducing a monotonous growth of time consumption with rising *NN*; 5) doubling the sampling rate can rise the decomposition accuracy conspicuously, without incurring additional computation time.

Table 2
Suggested parameter sets.

	Less time	More MU[1]	Higher accuracy[2]	Trade-off
N_e(%)	50	10	10	25
R	0	4	8	4
M	100	200	200	200
NN	20	20	20	20
Fs (Hz)	1024	2048	2048	2048

[1] Represents the goal that the decomposed MU number reaches the highest level even though huge time costs or low accuracy exist.

[2] Represents the goal that the balance of R2, mean SIL and RMSE reach the highest level without the consideration of time costs and decomposed MU numbers.

4.1. Preference Suggestions and Explanations of Overall Conclusions for Each Parameter

Based on results, we have the following preference suggestions for different project requirements, such as less computation time, more number of decomposed MUs, high decomposition accuracy or trade-off selection between decomposition performance and computation time. The overall recommendation is shown in Table 2.

When the percentage of removed channels N_e increases from 10% to 50%, the computation time monotonously decreases. Besides, an obvious inflection point of the number of MUs, RMSE and R^2 can be noticed at $N_e = 25\%$ for the SEEDS dataset and the two forces levels on EDC muscle in self-collected dataset. In contrast, the decomposition performances on FDS muscle in self-collected dataset tend to linearly deteriorate, except that there is a not obvious inflection point for the number of decomposed MUs at $N_e = 25\%$. Therefore, 25% of channels with minimal RMS are suggested for the N_e trade-off selection. The hidden reason is that channels with low RMS values cover very few muscle activations, but record baseline noise. The channels with only baseline noise can even decrease the decomposition accuracy, but require excessive computational resources. However, with the increasing of the eliminated channels number, channels containing useful information are removed, which makes the accuracy decline and RMSE grow. For the linear decrease of consumed time, the reason is that removing signal channels leads to a linear decrease of processed data size.

When the extension factor R isometrically increases from 0 to 8 with 2 interval, three metrics (the number of detected MUs, RMSE and R^2) tend to increase at start, all plateau after $R = 4$ across muscles and datasets. Therefore, $R = 4$ is preferred as the trade-off suggestion. Similar tendencies of MU quantity have been observed in previous studies [19,20]. The potential reason is that the decomposition performance can improve with the increasing of number of observations. When the number of observations is large enough to decompose all observed MUs, further increasing the extension factor contributes little to the decomposition results. In addition, the increase of R leads to the increase of the matrix size of decomposed data, resulting in high computation loads.

When the number of deconvolution iteration loops M isometrically increases from 100 to 300 with 50 interval, the number of detected MUs shows a logarithmic increase, and tends to be stable after $M = 200$ for the self-collected dataset, but tends to linearly increase for the SEEDS dataset. However, mean SIL slightly goes down from $M = 100$ to 150, then tends to be stable in both datasets. Besides, the RMSE and R^2 do not improve, and the time consumption soars up. Therefore, $M = 200$ is preferred for the trade-off suggestion. The reason is that the orthogonal projection steps, which avoid converging to the same MUs during continuous iteration, use the information of the separation matrix from all previous iteration loops for calculation. When a new MU is decomposed in an iteration loop, the separation matrix needs to be updated by adding a new column (a new separation vector of the decomposed MU). The size of separation matrix equals to the number of iteration loops that have been performed. A larger M substantially consumes more computation resources. However, the algorithm tends to separate good MUs with high SIL values at the beginning of loops, which leads to the slight decrease in mean SIL. With the increment of M, very few good MUs can be composed. Therefore, the two indicators (R^2 and RMSE) of decomposition performance tend to remain unchanged.

When the number of inner loop NN in a single MU deconvolution iteration isometrically increases from 20 to 80 with 30 interval, there is a small increase in computing time and no obvious change for the other four metrics in both datasets. Therefore, $NN = 20$ is recommended for preference suggestions. The reason is that if the algorithm can find the local optimal separation vector for a MU (the number of inner loop is typically less than 20) without waiting for the end of inner loop [17]. Two possible reasons for the small increase of time consumption are that 1) a small amount of MUs may require more iterations in inner loops for convergence, and 2) if one certain loop cannot converge, the algorithm has to carry out all NN iterations.

When the sampling rate Fs increases from 1024 to 2048 Hz, decomposition performances of four metrics (the number of decomposed MUs, mean SIL, RMSE and R^2) improve, but computation time increases in both datasets. Therefore, $Fs = 2048$ is preferred for the trade-off suggestion. This situation can arise from the fact that increasing the sampling rate provides more high-frequency information, which makes the MU spikes sharper with a higher amplitude. Sharper spikes can make a better clustering performance during ICA and improve the decomposition performance. However, the increase of Fs leads to higher dimensional data and more computational complexity, thus enlarging the computation time.

4.2. Time Consumption on the Embedded System

Considering the importance of time consumption in practical use, we implemented FastICA-based sEMG decomposition with the four suggested parameter sets on a commonly-used embedded system (Raspberry Pi 4B), which is portable and has been widely used for hardware implementation of real-time algorithms in biomedical engineering filed [28,29]. All suggested parameter sets were performed under Python 3.9.2, and ARM Cortex-A72@1.5 GHz with 4 GB of memory environment. Table 3 presents the average computation time with the server and Raspberry.

The results reveal that the computation time on the Raspberry is significantly longer than that on the server in each situation (all $p < 0.0001$), mainly caused by lowing the computational configuration. The shortest time spent on the Raspberry is 16.3 s with the parameter set of 'less computation time' in SEEDS dataset, which does not meet the requirement of portable and real-time needs. Considering the limited computing power of embedded systems, we recommend conducting the decomposition work on cloud servers for real-time engineering applications, instead of current embedded systems.

4.3. Comparisons of Different ICA-based Algorithms

Besides the FastICA algorithm, there are many ICA-based algorithms for blind source separation. Therefore, to have a comprehensive discussion, we investigated two other well-known ICA algorithms, i.e., InfoMax and SOBI. Since parameter sets can substantially influence the decomposition results and the trade-off parameter set can take both accuracy and computation time into consideration, we only applied the trade-off parameters to all three algorithms for comparison.

Fig. 5 presents the normalized results of the three ICA-based algorithms. We can conclude that 1) SOBI has the worst performance, except for the computation time. 2) FastICA can identify the largest number of MUs and has a faster convergence speed than InfoMax; 3) FastICA has similar performances as InfoMax in terms of mean SIL, RMSE and R^2 metrics. Regarding the computation time, SOBI performs ICA with the

Table 3
Average computation time (s) on the server and Raspberry.

		Less time	More MU	Higher accuracy	Trade-off
20% MVC on EDC muscle	Server	4.5	91.4	157.6	78.3
	Raspberry	34.4	359.0	478.6	323.1
50% MVC on EDC muscle	Server	4.3	85.2	147.2	75.6
	Raspberry	33.6	356.2	480.4	321.2
50% MVC on EDC muscle	Server	4.3	85.2	147.2	75.6
	Raspberry	33.6	356.2	480.4	321.2
SEEDS dataset	Server	2.9	55.7	118.6	50.7
	Raspberry	16.3	276.8	396.2	256.7

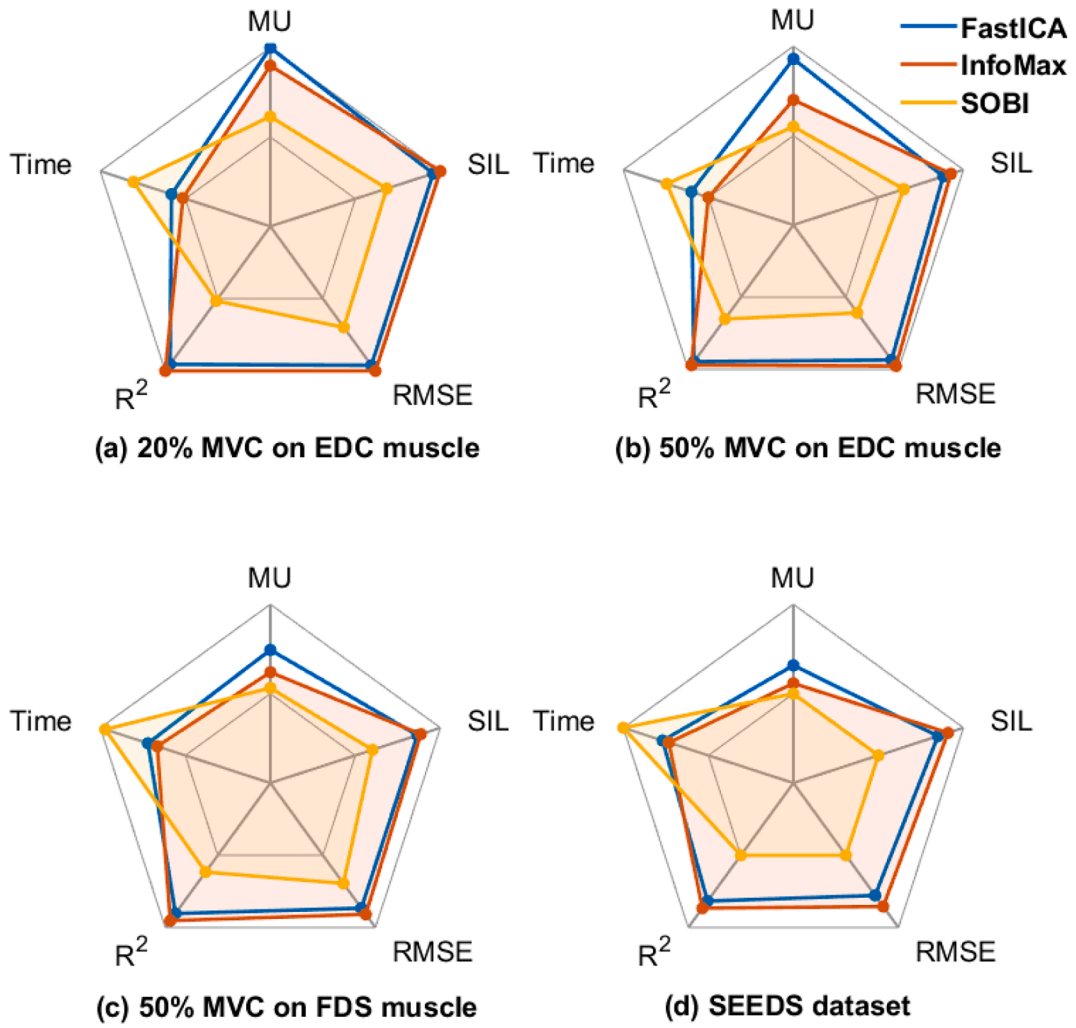

Fig. 5. Comparisons of different ICA-based algorithms. For a clear visual comparison, each metric is linearly normalized in [0, 1] with the minimum and maximum averaged value of this metric under various conditions. The edges of inner and outer pentagons represent the rescaled values of 0 and 1, respectively. To be consistent that a larger value represents a better result, the RMSE and computation time metrics are converted to negative values for the normalization in this figure.

second-order cumulants of the data. Correspondingly, the sources are estimated based on the simultaneous diagonalization of covariance matrices, contributing to the shortest computation time. In contrast, the InfoMax algorithm conducts the source separation with the aim of minimizing the mutual information among all sources. It is worth noting that InfoMax has to calculate the global mutual information of different sources in each iteration. On the other hand, FastICA employs the deflationary approach where only one source is extracted in each iteration, which can greatly lower the computation time, resulting in a lower computation time compared with Infomax. The worst performances in the other four metrics with SOBI may arise from the fact that SOBI is very sensitive to the time-lags parameter, the selection of which can be subject-dependent [30]. For the number of identified MUs, all ICA-based algorithms assume that the sources are independent. Actually, previous studies [31,32] have proven that firings of MUs have a certain degree of synchronization. Since the procedure of InfoMax decomposition is based on minimum mutual information of all sources, it potentially treats multiple MUs with strong firing synchronization as one single MU, which decreases the number of detected MUs.

4.4. Limitations and Further Improvement

This study explored the decomposition performances with different parameter values on different muscles and force levels, and provided preference suggestions for several research purposes. As the first study of parameter evaluation and selection from the perspective of microscopic analysis, we only investigated the decomposition performances with different parameter values on different muscles (EDC and FDS muscles) and force levels (20% MVC and 50% MVC), the performances under other experimental factors (e.g., different detection systems, contraction types, etc.) were not under consideration due to the limited experimental conditions.

5. Conclusion

In this work, we systematically studied the relationship between computation time and decomposition performance (the number of decomposed MUs, mean SIL, R^2 and RMSE) for FastICA-based sEMG decomposition algorithm under different selections of five decomposition parameters: percentage of channels eliminated, the number of extended channels, the number of decomposition iteration loops, the number of maximum inner loops in each iteration and sampling frequency. Based on the analysis of obtained results, we offered preference suggestions for different needs, in which, the trade-off suggestion of each parameter was 25%, 4, 200, 20, and 2048, respectively. In general, the results can help researchers find suitable parameter values while implementing HD-sEMG decomposition-based studies.

CRediT authorship contribution statement

Long Meng: Conceptualization, Methodology, Software, Formal analysis, Writing - original draft, Visualization, Validation, Data curation. **Qiong Chen:** Conceptualization, Data curation, Software, Resources, Writing - original draft, Validation. **Xinyu Jiang:** Formal analysis, Writing - review & editing. **Xiangyu Liu:** Visualization. **Jiahao Fan:** Writing - review & editing. **Chenyun Dai:** Funding acquisition, Investigation, Project administration, Writing - review & editing. **Wei Chen:** Supervision, Funding acquisition, Project administration, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] D.Y. Barsakcioglu, D. Farina, A real-time surface emg decomposition system for non-invasive human-machine interfaces, in: 2018 IEEE Biomedical Circuits and Systems Conference (BioCAS), IEEE, 2018, pp. 1–4.

[2] X. Jiang, et al., Open Access Dataset, Toolbox and Benchmark Processing Results of High-Density Surface Electromyogram Recordings, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 1035–1046.

[3] W.S. Marras, Overview of electromyography in ergonomics, in Proc. Hum. Factors Ergon. Soc. Annu. Meet., vol. 44, no. 30. SAGE Publications Sage CA: Los Angeles, CA, 2000, pp. 5–534.

[4] G. Hagg et al., Applications in ergonomics, Electromyography: Physiology, engineering, and noninvasive applications, pp. 343–363, 2004.

[5] A.M. Ndermann, et al., Orthotic comfort is related to kinematics, kinetics, and emg in recreational runners, Med. Sci. Sports Exerc. 195 (9131/03) (2003), 3510–1710.

[6] C.A. Hautier, et al., Influence of fatigue on emg/force ratio and cocontraction in cycling, Med. Sci. Sports Exerc. 32 (4) (2000) 839–843.

[7] D. Leonardis, et al., An emg-controlled robotic hand exoskeleton for bilateral rehabilitation, IEEE Trans. Haptics 8 (2) (2015) 140–151.

[8] R. Sun, et al., Complexity analysis of emg signals for patients after stroke during robot-aided rehabilitation training using fuzzy approximate entropy, IEEE Trans. Neural Syst. Rehabil. Eng. 22 (5) (2013) 1013–1019.

[9] X. Jiang et al., Neuromuscular Password-based User Authentication, IEEE Trans. Ind. Inform., vol. X, no. X, pp. 1–1, 2020.

[10] C.J. De Luca, R. Merletti, Surface myoelectric signal cross-talk among muscles of the leg, Electroencephalogr. Clin. Neurophysiol. 69 (6) (1988) 568–575.

[11] K.G. Keenan, et al., Influence of amplitude cancellation on the simulated surface electromyogram, J. Appl. Physiol. 98 (1) (2005) 120–131.

[12] D. Farina et al., Man/machine interface based on the discharge timings of spinal motor neurons after targeted muscle reinnervation, Nat. Biomed. Eng., vol. 1, no. 2, 2017.

[13] X. Liu, et al., Changes in synchronization of the motor unit in muscle fatigue condition during the dynamic and isometric contraction in the Biceps Brachii muscle, Neurosci. Lett. 761 (2021), 136101.

[14] X. Liu et al., Generalized Finger Motion Classification Model Based on Motor Unit Voting, Motor Control, vol. 25, no. 1, 2020.

[15] C. Dai, et al., Prediction of Individual Finger Forces Based on Decoded Motoneuron Activities, Ann. Biomed. Eng. 47 (6) (2019) 1357–1368.

[16] C. Dai, X. Hu, Finger Joint Angle Estimation Based on Motoneuron Discharge Activities, IEEE J. Biomed. Health Inform. 24 (3) (2020) 760–767.

[17] C. Dai, et al., Altered motor unit discharge coherence in paretic muscles of stroke survivors, Front. Neurol. 8 (MAY) (2017) 1–9.

[18] A. Del Vecchio, et al., The increase in muscle force after 4 weeks of strength training is mediated by adaptations in motor unit recruitment and rate coding, J. Physiol. 597 (7) (2019) 1873–1887.

[19] F. Negro et al., Multi-channel intramuscular and surface EMG decomposition by convolutive blind source separation, J. Neural Eng., vol. 13, no. 2, 2016.

[20] C. Dai, X. Hu, Independent component analysis based algorithms for high-density electromyogram decomposition: Systematic evaluation through simulation, Comput. Biol. Med. 109 (2019) 171–181.

[21] M. Chen, P. Zhou, A novel framework based on fastica for high density surface emg decomposition, IEEE Trans. Neural Syst. Rehabil. Eng. 24 (1) (2015) 117–127.

[22] C. Dai, X. Hu, Independent component analysis based algorithms for high-density electromyogram decomposition: Experimental evaluation of upper extremity muscles, Comput. Biol. Med. 108 (2019) 42–48.

[23] D. Farina, et al., Man/machine interface based on the discharge timings of spinal motor neurons after targeted muscle reinnervation, Nat. Biomed. Eng. 1 (2) (2017) 1–12.

[24] G.R. Naik, D.K. Kumar, Estimation of independent and dependent components of non-invasive emg using fast ica: validation in recognising complex gestures, Comput. Methods Biomech. Biomed. Engin. 14 (12) (2011) 1105–1111.

[25] A. Matran-Fernandez, et al., SEEDS, simultaneous recordings of high-density EMG and finger joint angles during multiple hand movements, Sci. Data 6 (1) (2019) 1–10.

[26] T. Kapelner, et al., Decoding motor unit activity from forearm muscles: perspectives for myoelectric control, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (1) (2017) 244–251.

[27] L. Meng, et al., Exploration of Human Activity Recognition Using a Single Sensor for Stroke Survivors and Able-Bodied People, Sensors 21 (3) (Jan. 2021) 799.

[28] S.A. Khadar, et al., A simple realtime algorithm for automatic external defibrillator, Biomed. Signal Process. Control 51 (2019) 277–284.

[29] F. Luna-Perejón, et al., IoT garment for remote elderly care network, Biomed. Signal Process. Control 69 (2021), 102848.

[30] C. Brunner, et al., Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis, Pattern Recognit. Lett. 28 (8) (2007) 957–964.

[31] J.G. Semmler, Motor unit synchronization and neuromuscular performance, Exerc. Sport Sci. Rev. 30 (1) (2002) 8–14.

[32] A.K. Datta, J.A. Stephens, Synchronization of motor unit activity during voluntary contraction in man, J. Physiol. 422 (1) (1990) 397–419.