Computers in Biology and Medicine 112 (2019) 103372

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Improve computational efficiency and estimation accuracy of multi-channel surface EMG decomposition via dimensionality reduction

Yong Ning[a], Nicholas Dias[b], Xuhong Li[c], Jing Jie[a], Jinrong Li[a], Yingchun Zhang[b,*]

[a] School of Automation and Electrical Engineering, Zhejiang University of Science & Technology, Hangzhou, 310023, China
[b] Department of Biomedical Engineering, University of Houston, Houston, TX, USA 77204
[c] The Third Xiangya Hospital, Central South University, Changsha, China

ARTICLE INFO

Keywords:
LMMSE
SVD
Multi-channel
Clinical diagnosis
Rehabilitation engineering

ABSTRACT

Electromyography (EMG) decomposition serves as a powerful tool to provide valuable information regarding motor control strategy and muscle pathology. Various approaches have recently been developed to decompose surface electromyography (sEMG) signals into motor unit action potential (MUAP) trains, but there is still room to improve decomposition accuracy and efficiency. In this article, the singular value decomposition (SVD) method is employed to decompose extended multi-channel sEMG signals based on the linear minimum mean square error (LMMSE) estimation and convolution kernel compensation (CKC). The column dimension of the right unitary matrix obtained via SVD is compressed to reduce the run-time of the decomposition process. As such, an innervation pulse train (IPT) can be extracted only using the compressed right unitary matrix. We studied the effect of reduced dimensionality on the computational efficiency and accuracy of the proposed method. Results from both simulated and experimental data demonstrate that this algorithm can decompose multi-channel sEMG efficiently and accurately. In most cases, 20%–60% of run-time can be reduced. In addition, the number of extracted IPTs can also be increased, especially when the signal-to-noise ratio of the EMG data is low. The method developed in this study has potential applications in clinical diagnosis, rehabilitation engineering and research on human motion control systems.

1. Introduction

Surface electromyography (sEMG) is composed of superimposed motor unit action potential (MUAP) trains generated by different motor units (MUs). A wealth of information regarding motor control strategy and muscle pathology can be extracted by analyzing the recruitment, de-recruitment, and firing rate of MUs. These motor unit properties are usually obtained through EMG decomposition, which is the process of classifying and identifying individual MUAPs in the interference pattern detected with either intramuscular or surface electrodes. Consequently, the constituent innervation pulse trains (IPTs) that compose the signal can be determined.

Efforts have been devoted to intramuscular (iEMG) decomposition [1–8] and some gratifying results have been achieved. However, iEMG is painful to subjects, and requires medical expertise to operate. Surface electrodes can be easily applied in a noninvasive manner with a high degree of repeatability, and require minimal medical knowledge or supervision [9]. More importantly, the spatial coverage offered by sEMG can describe global muscle activity. However, sEMG decomposition is very challenging due to the low pass filter effect of skin and fat tissue [10]. Furthermore, the low-pass filtering effect of the volume conductor further reduces differences in MUAP shapes, greatly increasing the difficulty of correctly classifying MUs. As a result, the properties of MUAP shapes in sEMG are not identifiable. Together, these aspects make sEMG decomposition using traditional spike-sorting techniques employed in the decomposition of iEMG a challenging task. Regardless, great strides have been made towards the efficient and accurate decomposition of sEMG signals. So far, sEMG decomposition techniques can be divided mainly into two categories. The first is based on template recognition, De Luca et al. [11,12] and others have achieved good results with this method. Template recognition methods require signal acquisition at a high sampling frequency, and then the templates of the MUAP waveform, which correspond to different MUs, are identified. These templates are then used to identify and classify the corresponding MUAP waveform to decompose sEMG signals. The second is the blind source separation (BSS) methods, of which Holobar et al. first proposed the CKC method [13]. By directly employing the multi-channel measurements, the CKC method does not involve any

* Corresponding author. Department of Biomedical Engineering, University of Houston, 3605 Cullen Blvd, Room 2027, Houston, TX, 77204, USA.
E-mail address: yzhang94@uh.edu (Y. Zhang).

https://doi.org/10.1016/j.compbiomed.2019.103372
Received 5 November 2018; Received in revised form 26 July 2019; Accepted 26 July 2019

correlation calculations of the demixing matrix that are required in other BSS methods. CKC requires simple calculations, produces highly accurate decomposition results and extracts a large number of IPTs. Next, the gradient (gCKC) method was proposed [14], which uses the natural gradient algorithm to optimize the IPT sequence, improving upon the original CKC method. Finally, KmCKC [15] was proposed based upon the original CKC method. This method first classifies the multi-channel signal vectors corresponding to the MU firing times using k-means clustering, and then reconstructs the initial IPT using the classified vectors. On this basis, the correlation vector between the IPT and the measurements is reconstructed, and finally, the corresponding IPT is estimated. KmCKC can significantly improve the decomposition results when compared to the previously discussed methods. Later, Ning et al. [16] proposed an approach based on the construction of the measurement correlation matrix to decompose the sEMG signals with LMMSE. This method has two main steps. First, the measurement correlation matrix is calculated, and then an iterative approach is adopted to reconstruct the correlation matrix between IPTs and the measurements. This method can obtain reliable decomposition performance, especially in cases where the MUAP superposition is severe. Although various approaches have been developed [11–23] and a high number of promising results have been achieved, sEMG decomposition accuracy and efficiency can still be markedly improved, and the importance of such undertakings cannot be understated. Many sources of noise can obfuscate sEMG signals recorded under experimental conditions. In this case, it is vital to design a decomposition algorithm that is more robust to noise. Additionally, it is imperative that sEMG decomposed for clinical and physiological research data is accomplished in a manner that produces the most information from a single recording session.

Several sEMG decomposition methods have been developed previously [11–17,23], however, the decomposition performance with dimensionality reduction using eigenvalue truncation has not been extensively investigated. In this study, the singular value decomposition (SVD) and compressed unitary matrix were employed to process multi-channel sEMG signals to improve the performance of the sEMG decomposition achieved using linear minimum mean square error estimation (LMMSE) and convolution kernel compensation (CKC) [13]. The obtained results were compared with the KmCKC method, which shares several features with previously proposed methods [13,17,23]. Our simulation and experimental results demonstrate that the run-time of the decomposition process is significantly reduced by compressing the column dimension of the right unitary matrix derived from SVD.

2. Materials and methods

2.1. Data model

Multi-channel sEMG can be described with a time-invariant linear Multi-Input-Multi-Output system. The data model can be given using the following equation [13]:

$$\boldsymbol{X}(n) = A\bar{\theta}(n) + b(n) \tag{1}$$

Where $X(n) = [x_1(n),...,x_M(n)]^T$ is the sEMG signals measured by M channels, $x_j(n)$ is the n-th sample in the j-th channel, $b(n) = [b_1(n),...,b_M(n)]^T$ is a spatially and temporally zero-mean white noise vector.

$$\bar{\theta}(n) = [\theta_1(n), \theta_1(n-1),...,\theta_1(N-P+1),..., \theta_N(n), \theta_N(n-1),...,\theta_N(n-P+1)]^T$$

is the expanded form of N motor unit pulse trains (MUPTs) $\theta(n) = [\theta_1(n),...,\theta_N(n)]^T$. A is the mixing matrix which is made up of all the channel responses $a_{ij} = [a_{ij}(0),...,a_{ij}(P-1)]$ with length P. Where a_{ij} denotes the j-th response appeared in the i-th channel.

It is necessary to extend $X(n)$ in Equation (1) with L–1 delayed repetitions of each original measurement [13] to improve the knowns-to-unknowns ratio:

$$\overline{X}(n) = [x_1(n), x_1(n-1),...,x_1(n-L+1),...,x_M(n), x_M(n-1),...,x_M(n-K+1)]^T \tag{2}$$

Consequently, there are a total of LM rows in the matrix$\overline{X}(n)$.

2.2. Decomposition methods using SVD and LMMSE

Per the expression of LMMSE in Ref. [24], a linear estimation of θ which is statistically correlated with $\overline{X}$ can be written as

$$\hat{\theta} = E(\theta) + C_{\overline{X}\theta}C_{\overline{X}\overline{X}}^{-1}(\overline{X} - E(\overline{X})) \tag{3}$$

where $\hat{\theta}$ is an estimation of θ, E(θ) is an expectation of θ, and $\overline{X}$is a known parameter which is statistically correlated with θ.$C_{\overline{X}\theta}$ is the cross-correlation of $\overline{X}$ and θ, and $C_{\overline{X}\overline{X}}^{-1}$ is the inverse of self-correlation of $\overline{X}$. Let E(θ) = E($\overline{X}$) = 0, then it can be written

$$\hat{\theta} = C_{\overline{X}\theta}C_{XX}^{-1}\overline{X} \tag{4}$$

where $C_{\overline{X}\overline{X}} = E[\overline{X}\overline{X}^T]$.

For multi-channel sEMG, $\hat{\theta}$ denotes an estimation of the IPT. $\overline{X}$ denotes the extended multi-channel sEMG signals and $C_{X\theta}$ is an unknown parameter that needs to be estimated [13]. According to Ref. [13], if all firing times of an IPT are known in advance, the estimation equation of $C_{\overline{X}\theta}$ can be written as

$$\mathbf{C}_{X\theta}{}^T = \frac{1}{card(\psi_j)}\sum \overline{X}(\psi_j) \tag{5}$$

where $\psi_j$stands for a set of firing times of the j-th MU, and card(ψ_j) stands for the cardinal number of ψ_j.

To decompose $\overline{X}$ using economy-size singular value decomposition (SVD), shown by

$$\overline{X} = VDU^T \tag{6}$$

Substitute Equations (5) and (6) into (4), to yield

$$\begin{aligned}\hat{\theta}(n) &= \bar{X}^T(n_j)\{E[\bar{X}(n)\bar{X}^T(n)]\}^{-1}\bar{X}(n)\\ &= U(n_j)E[D^TV^TV(DD^T)^{-1}V^TVD]U^T(n)\\ &= U(n_j)U^T(n)\end{aligned} \tag{7}$$

Where n_j denotes the n_j-th time instant. It can be seen from Equation (7) that an estimation of IPTs can be obtained only using $U(n)$ which is the right unitary matrix on the right-hand side in Equation (6). This is because the time-domain characteristic information of the sEMG signals is included in the column vectors of $U(n)$ [25].

2.3. Dimension reduction and decomposition

The right unitary matrix $U(n)$ with LM columns is obtained after decomposing $\overline{X}$ via economy-size SVD. To simplify the matrix calculation and shorten the run-time of the decomposition algorithm, the number of columns in $U(n)$ is reduced by removing columns corresponding to LM-r ($r < LM$) smallest singular values from the rearward part of the matrix. As a result, only the first r columns in $U(n)$ are retained, significantly decreasing the size of $U(n)$. Consequently, the computational efficiency of the decomposition is considerably improved. Furthermore, the time domain characteristic information of sEMG signals is preserved. The detailed decomposition procedure is shown below:

1. Decompose the matrix $X^T(n)$ into $X^T(n) = U_{Ns\times LM}(n)D_{LM\times LM}(n)V_{LM\times LM}(n)^T$ using economy-size SVD, where ^{T} denotes the transpose, and Ns is the number of samples in each measurement.
2. Reduce the number of columns from LM to r in matrix $U(n)$.
3. Compute $\gamma(n) = X^T(n)X(n)$,(n = 1,2, …) and $n_0 = \underset{\text{arg}}{median}(\gamma(n))$[13]. Then compute $IPT_{n_0} = U(n_0)_{1\times r}U^T(n)_{r\times Ns}$.

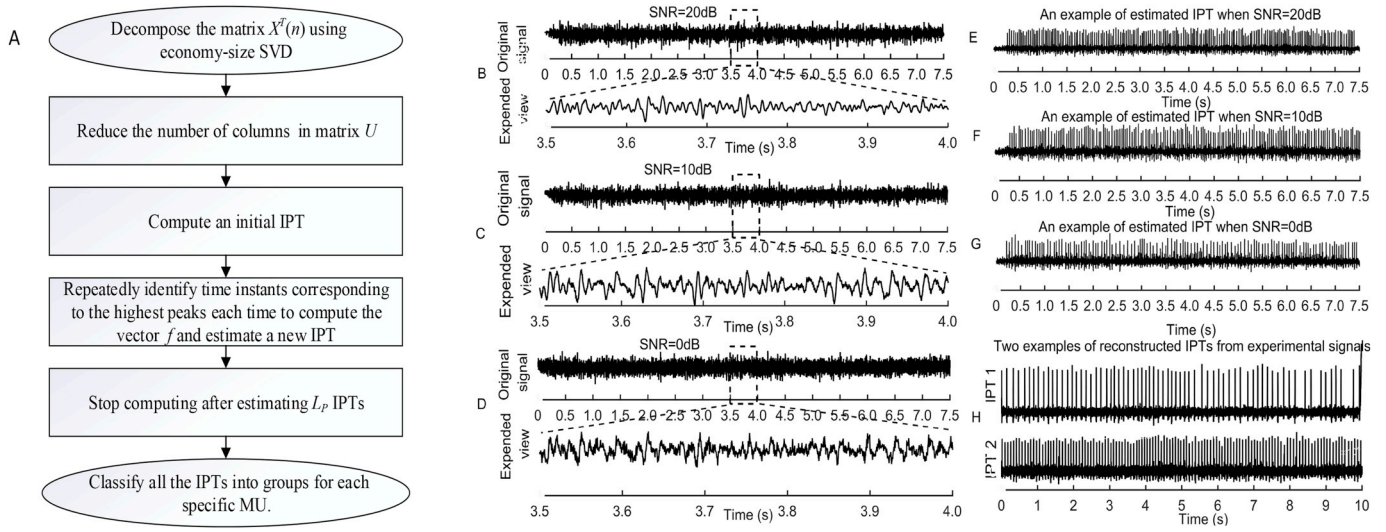

Fig. 1. The algorithm step diagram, simulated sEMG signals and examples of reconstructed IPTs. (A) algorithm step diagram, (B) simulated sEMG signals with SNR of 20 dB, (C) simulated sEMG signals with SNR of 10 dB, (D) simulated sEMG signals with SNR of 0 dB, (E) an example of estimated IPT when SNR was set to 20 dB, (F) an example of estimated IPT when SNR was set to 10 dB, (G) an example of estimated IPT when SNR was set to 0 dB, (H) two examples of reconstructed IPTs from experimental signals.

4. Identify d_k ($d_k = C \times k$, where k is the k-th iteration, and C is an integer greater than zero and usually between 3 and 6) time instants$\phi_{nx} = \{n_{x1}, \ n_{x2}, \ \ldots, \ n_{xd_k}\}$, corresponding to the highest peaks in IPT_{n_0}, then replace $U(n_0)_{1\times r}$in step 3 with $f = \frac{1}{d_k}\sum U(\phi_{nx})_{1\times r}$. Then a new IPT $IPT(n) = f_{1\times r}U^T(n)_{r\times Ns}$will be obtained. The vector f will be gradually improved by repeating this iterative process until$d_k > N_p$, where N_P is a rough estimate number of firing times in each IPT, at which point the final IPT will be obtained.
5. Set $\gamma(n_0) = 0$.
6. Repeat steps 3 to 5 for Lp times (Lp can be set according to actual condition)
7. Classify all the IPTs for each specific MU.

Note that every d_k time instant needs to be identified in each estimated IPT in step 4. N_p is theoretically equal to the number of firing times of MUs. However, it is difficult to estimate the number of firing times of MUs in a practical case. Therefore, N_p can be estimated according to the length of signal acquisition and firing rate, the latter of which can be set to a normative value for the target muscle, which was chosen to be 30 Hz in this case. The algorithm step diagram is shown in Fig. 1A.

2.4. Simulated surface EMG signals

According to Refs. [26,27], the extracellular single fiber action potential (SFAP) can be represented via the sum of three basic Gaussian functions as shown below:

$$\zeta(t) = \sum_{i=1}^{3} W_i e^{-\left(\frac{t-G_i}{Z_i}\right)^2} \tag{8}$$

where t is time, Z_i is the bandwidth, W_i is the amplitude factor, and G_i is the center of the peak. Using these three basic Gaussian functions, any triphasic action potential waveform can be approximated accurately by adjusting W_i and Z_i. Suppose each fiber is parallel to the skin surface, then the SFAP shape detected by the electrodes can be considered as a function of muscle fiber conduction velocity and the fiber position in a 3-dimensional Cartesian coordinate system. W_i and Z_i in Equation (8) were represented as defined in Ref. [27]:

$$W_i = f_1(d, e, f, cv) \tag{9}$$

$$Z_i = f_2(d, e, f, cv) \tag{10}$$

where e denotes the fiber depth in the vertical direction below the skin surface, f is the center position along the fiber in the f-d plane, d is the fiber center position in the f-d plane perpendicular to the f direction, and cv is the muscle fiber conduction velocity. The MUAP shapes were depicted as the summation of the SFAPs included in a single MU. Then the MUAP trains were produced by convolving their corresponding firing times with the MUAP shapes. Finally, the synthetic sEMG signals were modeled as linear summations of the MUAP trains. The shape, duration and amplitude distribution of the MUAP depended on the morphological properties of active muscle fibers innervated by the corresponding MUs.

In this test, a series of sEMG signals were simulated. Additive zero-mean Gaussian noise of varying amplitude was added into the simulated sEMG signals. The center depths of all MUs were uniformly distributed from 1 mm to 10 mm. A random number of muscle fibers uniformly distributed between 50 and 200 was assumed in active MUs. All semi-fiber lengths were set to 50 mm, and the endplate and tendon positions of the fibers were uniformly distributed in the range of ± 5 mm. MU conduction velocities were set to 4.0 m/s. The motor unit recruitment and discharge patterns in sEMG simulation primarily depended on [28]. The mean and standard deviation of firing rates of the MUs were normally distributed and set from 20 ± 5 to 30 ± 5 Hz. A total of 130 MUs were simulated for each respective group. The starting times of each MU was set from 10 ms to 300 ms. A 16 × 8 electrode-array grid with a 4.5-mm inter-electrode distance in both directions was generated and centered above the simulated MUs. A sampling rate of 2,000 Hz was assumed. The numbers of fibers, location of active MUs, and MU discharge patterns were all randomly generated. Fig. 1B, C, D shows the simulated sEMG signals with an SNR of 20 dB, 10 dB and 0 dB, respectively. The extension factor L was set to 16. Hence, a total of 2048 channels were produced after original measurements $X(n)$ were extended.

2.5. Experimental surface EMG signals

The vastus lateralis (VL) of one subject (Subject 1) (the contraction level was not measured), the first dorsal interosseous (FDI) of three adult subjects (Subjects 2, 3 and 4) (with contraction force of 2 N), the biceps brachii muscle of one subject (Subject 5) (the contraction level was not measured) and the FDI of three adult subjects (Subject 6, 7 and

8) at different force levels were studied. The protocol was approved by the Institutional Review Board of University of Houston, Houston, TX, USA. Written informed consent was obtained before conducting each experiment. Two 8 × 8 arrays of silver-plated electrodes with an 8.5-mm inter-electrode distance were used to acquire sEMG from the VL of Subject 1. An 8 × 8 array of silver-plated electrodes with a 4-mm inter-electrode distance was used to acquire sEMG from the FDI of Subjects 2, 3, 4, 6, 7 and 8. An 8 × 8 array of silver-plated electrodes with an 8.5-mm inter-electrode distance was used to acquire sEMG from the biceps brachii of Subject 5. Isopropyl alcohol was used to sanitize the skin over the target muscle, and the skin was lightly abraded with an abrasive paste to reduce the skin-electrode impedances. The electrode was affixed over the target muscle. The ground electrode was attached to the wrist of the subject via a Velcro strap. The subjects were configured in a manner to ensure an isometric contraction of the target muscle. Signals were recorded with a 136 channel Refa amplifier (TMSi, the Netherlands) with a sampling rate of 2,048 Hz. The sEMG signals were band-pass filtered with a range of 10 Hz–500 Hz. The extension factor L for all subjects was set to 16. Hence, the total number of channels after extension was 2048, 1024, 1008, 944, 1024, 1024, 1024 and 1024, for each subject respectively, in numerical order. Fig. 2A shows one channel of signal measured from subject 1. Fig. 2B shows one channel of EMG from subject 7 at 15% maximal voluntary contraction (MVC).

3. Results

The proposed decomposition algorithm, abbreviated here to “SL”, was tested with simulated sEMG signals, and experimental sEMG signals recorded from 8 subjects, including 5 healthy subjects and 3 stroke patients. The main purpose of this study is to investigate the effect of column dimension of U(n) on the computational efficiency, amount of decomposed IPTs, and decomposition accuracy of the algorithm. The performance of the proposed approach was evaluated by comparing against the developed KmCKC method [15]. The number of main decomposition loops was set to 500. All data results were averaged over 5 trials.

3.1. Simulation results

Fig. 3 shows the influence of reduced dimensions of matrix $U(n)$ on run-time corresponding to SNRs of 20 dB (Figure 3A), 10 dB (Fig. 3B), and 0 dB (Fig. 3C). As U_d is decreased, the run time of the algorithm also decreases linearly.

Fig. 4 shows the influence of the reduction of the column dimension of matrix $U(n)$ on the number of extracted IPTs and the accuracy corresponding to SNRs of 20 dB (Figure 4A), 10 dB (Fig. 4B), and 0 dB (Fig. 4C). At a high noise level (SNR = 0 dB), only an average of 2.8 IPTs could be extracted when U_d was set to 2048. Compressing U_d improves the decomposition performance and results in IPTs from over 10 MUs being extracted. Further compression of U_d to 256, however, resulted in a reduced number of extracted IPTs.

The accuracy was calculated according to Ref. [15]. The firing time tolerance was set to ± 1 sample. Thus, each detected firing time was considered true if it was identified within ± 0.5 ms (sampling frequency of 2000 Hz) from its actual position along the signal. In addition, pulse-to-noise-ratio (PNR) [29] was also used to verify the performance of the proposed method. The PNR was varied from 23.98 dB to infinity. The average PNR was 36.89 dB after removing the maximum of infinity.

Table 1 summarizes the simulation results with varying signal SNR, extension factor L, and dimensionality of $U(n)$. R_C denotes the percentage of the first r columns reserved (corresponding to r largest singular values) in matrix $U(n)$, which were varied from 100% to 25%. It can be seen from the table that when L was 6 and the SNR was 20 dB and 0 dB, the average number of reconstructed IPTs reached the maximum of 28.8 and 8.2 respectively, when R_C was set at 75%. While the average number of reconstructed IPTs reached the maximum of 18.4 when the SNR was 10 dB and R_C was set at 50%, respectively. When SNR, L, and Rc were set to 20 dB, 16, and 25% respectively, the largest number of IPTs could be reconstructed, at an average of 35.2 IPTs. When SNR and L were decreased to 10 dB and 10, the resultant largest number of reconstructed IPTs decreased to an average of 30.6 IPTs when Rc were set to 50%. Similarly, when SNR was further reduced to 0 dB, L was set to 16, IPT yield was significantly increased by reducing Rc, yielding a maximum of 13.4 IPTs. Fig. 1E, F and G shows an example of an estimated IPT when SNR was set to 20 dB, 10 dB and 0 dB, respectively.

3.2. Experimental results

Figs. 5 and 6 present an analysis of decomposition results obtained from experimental sEMG signals recorded from 5 healthy subjects. Fig. 5A presents the results obtained from Subject 1. The number of reconstructed IPTs improved when U_d was decreased to 1500, but further reductions in dimensionality deteriorated the decomposition results. Fig. 5B presents the results obtained from Subject 2. Decomposition results were slightly affected by small reductions in U_d (less than 50%). After this point, the number of extracted IPTs was drastically reduced. SL outperformed KmCKC in both run-time and accuracy. Fig. 5C presents the results obtained from subject 3.

The number of extracted IPTs increased when U_d was decreased to 490, but any further reductions in dimensionality resulted in reduced IPT yield. Fig. 6A presents the results obtained from subject 4. The extracted IPT number increased as U_d when was decreased to 472. Further reductions in U_d yielded less IPTs. Fig. 6B presents the results obtained from subject 5. The number of extracted IPTs increased to 10.6 when U_d was decreased to 768 but yielded less IPTs when U_d was further reduced. Dimensionality reduction of $U(n)$ drastically decreased the run time of the algorithm in all healthy subjects. Fig. 1H presents the two reconstructed IPTs from subject 1.

Table 2 shows results from subjects 6, 7 and 8 at different contraction levels. Where the forces generated by subjects 6 and 8 were 20%, 30% and 40% of MVC, and those by subject 7 were 10%, 15% and 20%

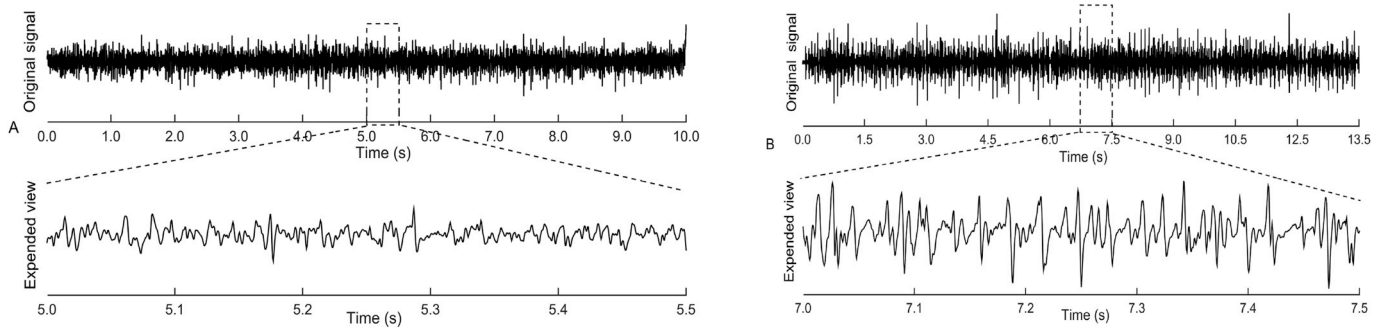

Fig. 2. The experimental signals. (A) one channel of EMG from subject 1, (B) one channel of EMG from subject 7 at 15% MVC.

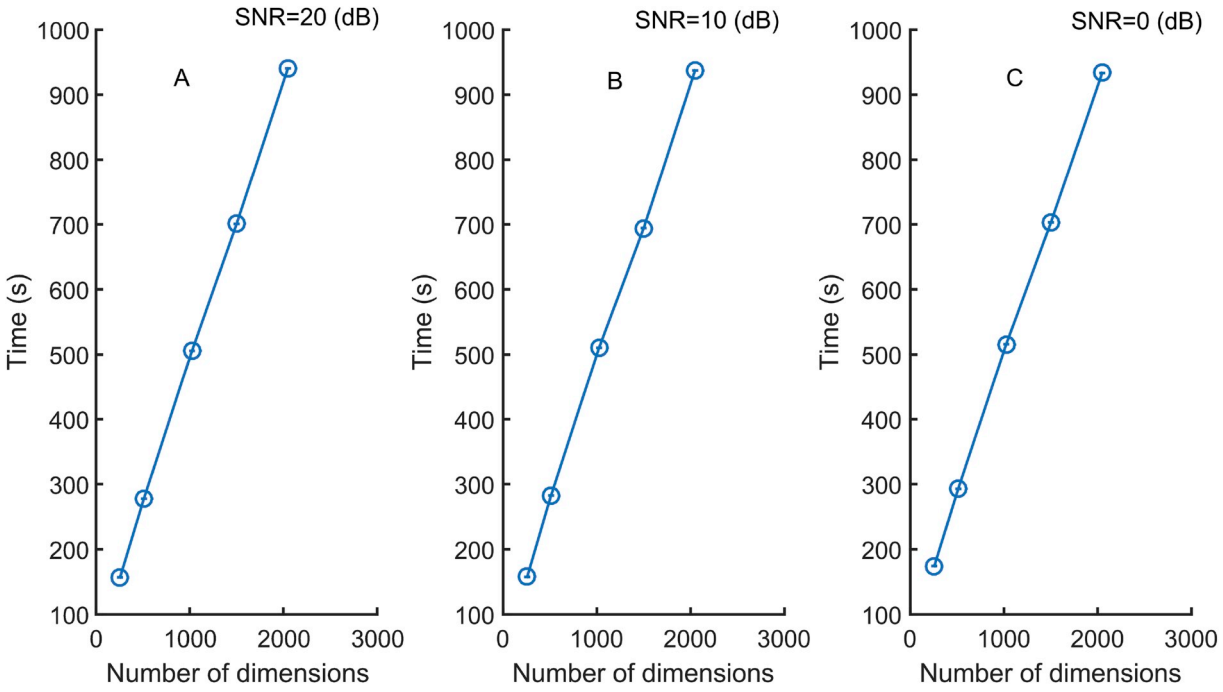

Fig. 3. The results of simulated signals generated by a Gaussian function with 50 active MUs. The influence of the number of column dimensions in matrix U(n) on running time with SNR of (A) 20 dB, (B)10 dB and (C) 0 dB, respectively.

of MVC. As seen from the table, in most cases, the average number of extracted IPTs is increased by reducing $U(d)$. Moreover, the average calculation time is reduced significantly. Fig. 7A and Fig. 7B shows the 25 IPTs extracted from subject 1 and the instantaneous firing rates of 14 MUs from subject 7 at 15% MVC. In Fig. 7A, each vertical bar represents a firing time. In Fig. 7B, each dot indicates a MU instantaneous firing rate at a given time instant, where each vertical axis has a range of 0–30 pulses per second (pps).

4. Discussion

The method presented in this article aims to enhance LMMSE by employing SVD to reduce the number of column dimensions of right unitary matrix $U(n)$, thus, significantly shortening the computing time, and, in some cases, achieving a more representative decomposition result. To the best of our knowledge, the presented approach has not been previously reported in the literature. In this article, the proposed method and the advanced KmCKC were compared in terms of comprehensive performance. Our results show that IPT yield and calculation time are improved.

After the extended measurement matrix $\overline{X}(n)$ is decomposed by SVD, the original information of the time domain characteristic is reserved in matrix $U(n)$, where the most important information is reserved in the columns corresponding to r largest singular values at the

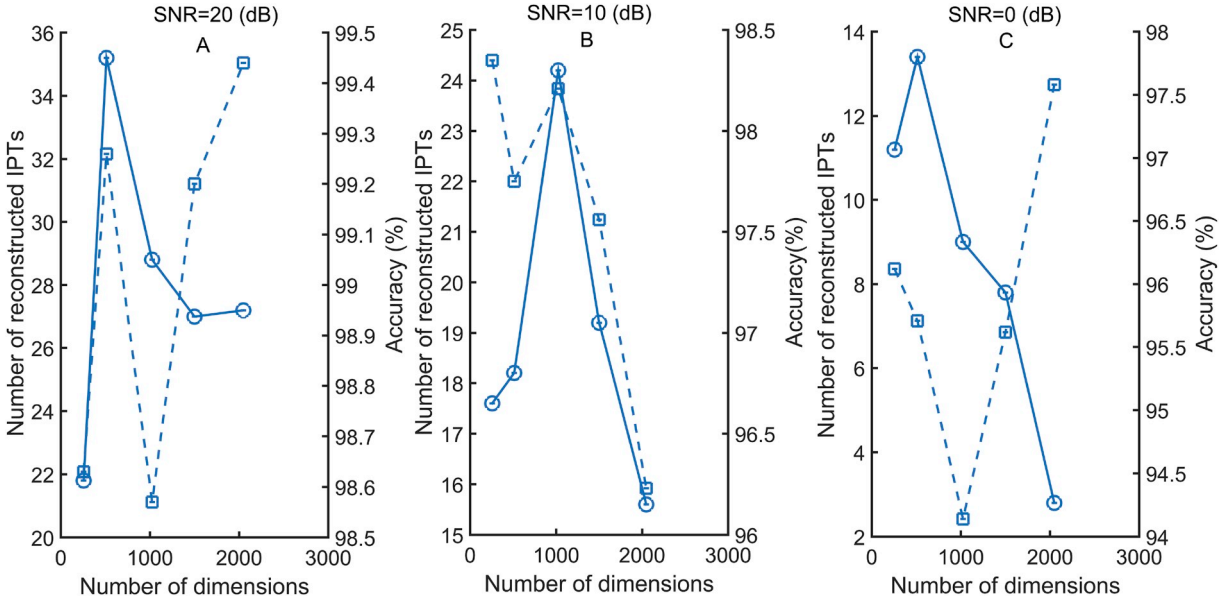

Fig. 4. The results of simulated signals generated by a Gaussian function with 50 active MUs. The influence of the number of column dimensions in matrix U(n) on the number of extracted IPTs and decomposition precision with SNR of (A) 20 dB, (B) 10 dB and (C) 0 dB, respectively. The solid line indicates the number of reconstructed IPTs, the dashed line indicates the precision.

Table 1
Simulation results achieved using different extension factors and different numbers of columns of $U(n)$.

R_C (%)	Extension factor L						SNR (dB)
	6		10		16		
	Number of IPTs	Accuracy (%)	Number of IPTs	Accuracy (%)	Number of IPTs	Accuracy (%)	
100	21.4	97.1	25.6	98.8	27.2	99.4	20
75	28.8	97.4	27.4	98.2	27.0	99.2	
50	26.6	97.8	30.6	98.6	28.8	98.5	
25	24.4	97.4	27.2	98.3	35.2	99.2	
100	14.4	97.5	16.2	97.5	19.2	96.2	10
75	15.8	96.7	23.0	97.0	18.2	97.2	
50	18.4	96.5	19.6	96.8	24.2	98.2	
25	15.6	96.6	19.2	97.7	18.2	97.7	
100	6.6	91.0	7.2	94.7	2.8	97.5	0
75	8.2	91.0	9.0	91.8	7.8	95.6	
50	6.2	92.6	10.2	93.7	9.0	94.1	
25	6.0	94.7	8.2	95.0	13.4	95.7	

front of $U(n)$. Therefore, the column dimensions of $U(n)$ can be reduced by removing columns in the rearward position. Finally, satisfactory decomposition results can be achieved utilizing only the reserved information. Our results show that the optimal dimensionality of $U(n)$ varies. The diverse nature of EMG signals recorded from different experimental protocols explains this variance. Factors including the muscle properties of various subjects, muscle anatomy (small vs large), number of electrodes, size of electrodes, and SNR, are different for all experimental cases. These factors all influence the optimal dimensionality of $U(n)$. Therefore, for each experimental case, ideal $U(n)$ dimensionality should first be estimated if necessary. In practice, several matrix dimensionalities can be selected at large intervals. For example, 75%, 50% and 25% of the original matrix dimensionality can be selected, then sEMG can be decomposed to obtain the desired results. Both the simulated and experimental results show that the greater compression of $U(n)$, the more the run-time performance of the algorithm is increased. However, if too many dimensions in $U(n)$ are removed, useful information may not be retained. Consequently, the decomposition accuracy will be negatively affected. If $U(n)$ is not compressed, not only is the run-time of the algorithm increased, the IPT yield may not be optimal. Therefore, run-time performance and decomposition quality should be balanced when determining the optimal degree of dimensionality reduction of $U(n)$. According to the results achieved in this study, an appropriate balance between the computing time and the number of decomposed IPTs was found with a 25%–50% reduction in column dimensions of $U(n)$, depending on signal characteristics. There is no single standard for the choice of r columns, it depends on the goals of the operator.

In most cases, r is set to half the number of $U(n)$ columns to obtain the best balance of efficiency and IPT yield. It is also possible to set multiple r values according to the characteristics of the specific signal, and then select an r value with the best decomposition results. An important difference between KmCKC and SL is that the sEMG signal matrix cannot be compressed during KmCKC. Therefore, the dimensionality of the inverse correlation matrix of extended observations in KmCKC is unique, that is, it can only be assessed in the full-dimensional case, which leads to a time-consuming calculation. The ability to reduce the dimensionality of the unitary matrix $U(n)$ is an advantage of

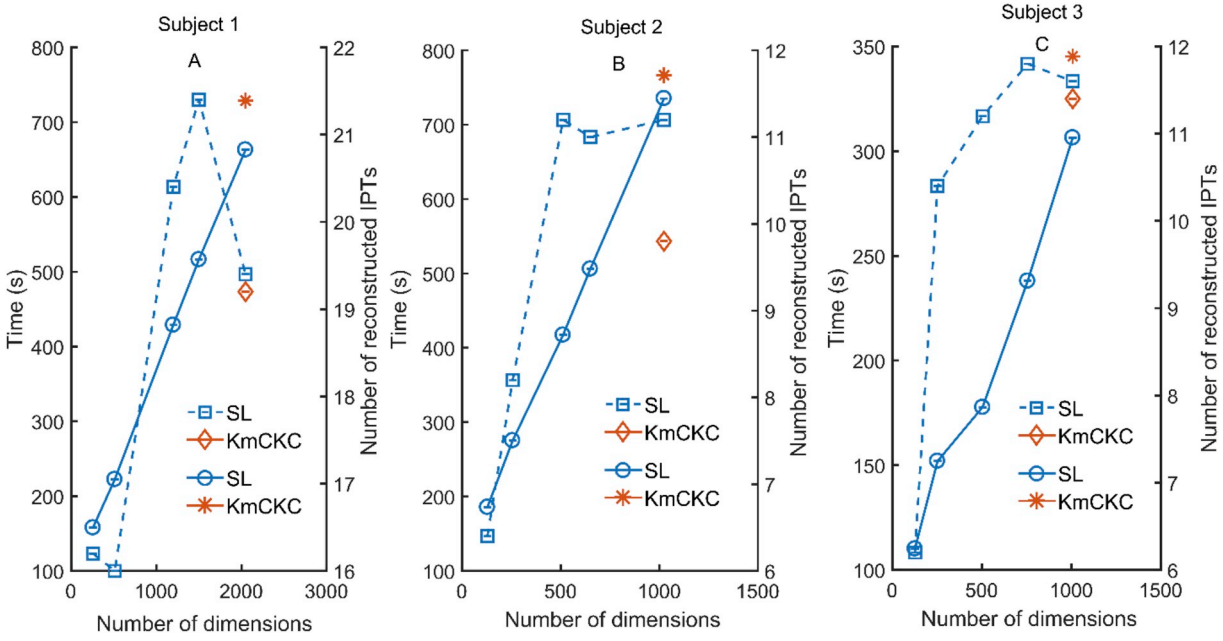

Fig. 5. The results of experimental signals. The influence of number of column dimensions in matrix U(n) on the decomposition time and the number of extracted IPTs for (A) subject 1, (B) subject 2 and (C) subject 3. The solid line indicates the running time, the dashed line indicates the number of extracted IPTs. For KmCKC, the diamond and asterisk symbol represents number of reconstructed IPTs and time respectively.

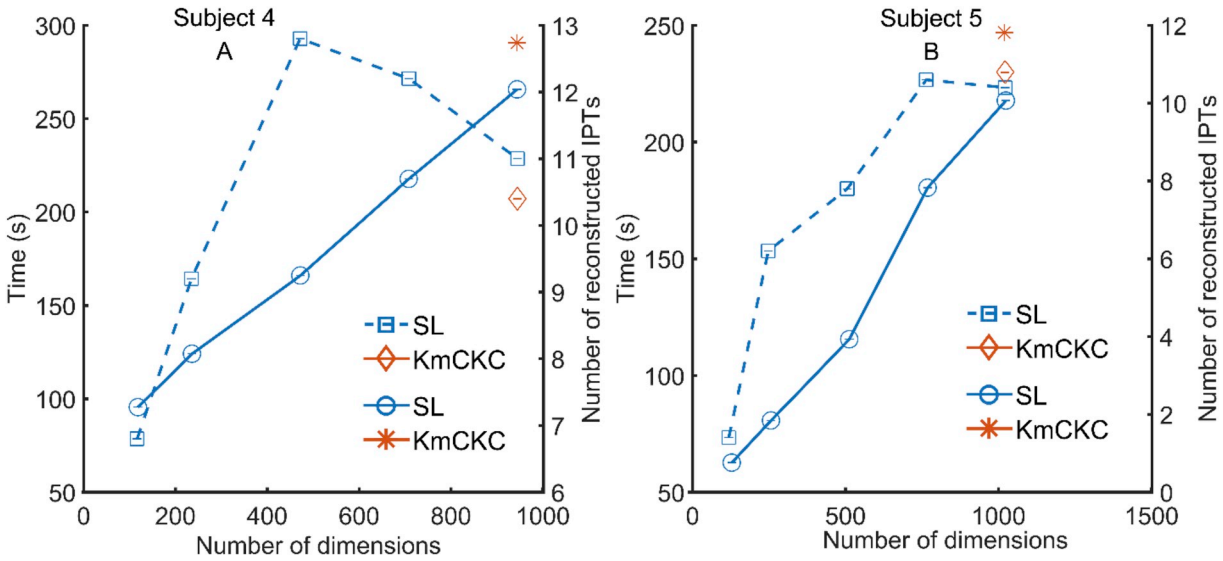

Fig. 6. The results of experimental signals. The influence of the number of column dimensions in matrix $U(n)$ on the decomposition time and the number of extracted IPTs for (A) subject 4 and (B) subject 5. The solid line indicates the running time, the dashed line indicates the number of extracted IPTs. For KmCKC, the diamond and asterisk symbol represents number of reconstructed IPTs and time respectively.

Table 2
Experimental results achieved at different contraction levels (%MVC) and different numbers of columns of $U(n)$.

MVC(%)	Rc(%)	Number of IPTs	RunningTime(s)	Subject
20	100	11.2	715.2	Subject 6
	75	13.0	407.1	
	50	10.8	287.2	
	25	8.4	185.1	
30	100	9.4	717.9	
	75	9.2	543.5	
	50	7.8	383.3	
	25	4.2	240.6	
40	100	7.8	666.8	
	75	4.4	394.1	
	50	8.4	276.9	
	25	4.2	179.3	
10	100	8.4	644.4	Subject 7
	75	9.8	495.0	
	50	9.8	350.0	
	25	8.2	232.1	
15	100	12.2	544.2	
	75	12.6	433.6	
	50	12.2	307.1	
	25	11.4	204.3	
20	100	11.0	583.2	
	75	11.0	465.9	
	50	9.2	331.0	
	25	7.6	218.8	
20	100	9.8	515.0	Subject 8
	75	11.4	410.4	
	50	8.4	297.9	
	25	8.0	187.9	
30	100	6.0	491.4	
	75	6.2	383.3	
	50	3.4	276.9	
	25	2.2	178.7	
40	100	9.2	573.4	
	75	10.4	456.8	
	50	8.2	332.3	
	25	5.6	210.9	

SL, significantly improving run-time performance. Since SL directly estimates the IPTs from the measurement matrix, channel extension is necessary to obtain suitable results [13]. Without channel extension, IPT yield and accuracy will be diminished, and in some cases, no IPTs will be extracted. The proposed dimensionality reduction can improve the efficiency and the number and accuracy of the extracted IPTs when compared with previously published related methods. When the unitary matrix $U(n)$ can be compressed, and necessary information is retained to the greatest extent, the influence of interference information is diminished or discarded. On the other hand, by applying an advanced iterative method, the greatest amount of firing MU firing information can be obtained, thus the correlation vector f between the measurements and the IPTs can be better constructed. Note that dimensionality reduction can be performed in several ways, which may also improve computational efficiency and decomposition accuracy. The improvement is particularly apparent when the noise level in the signals is relatively high, as seen from the simulated results.

Although channel extension greatly improves results, it leads to increased computational burden. Therefore, the proposed method of compressing $U(n)$ can solve the problems posed by increased computational complexity after channel extension. Moreover, in some cases compressing $U(n)$ can significantly improve the number and accuracy of the reconstructed IPTs. According to the previous literature [13,15,23], the extension factor L can be set from 10 to 16, which directed our decision to use L values from 10 to 16 in our tests. We tested the relationship between L and dimension of $U(n)$, as presented in Table 1. Sometimes choosing a larger L value (such as 16) may improve results (see Table 1), at the expense of computational speed. So far, published literature [13,15] most commonly chose an L of 10 to balance trade-offs between accuracy and computation time. The economy-size SVD method was first employed which can achieve decomposition accuracy at the level very close to non-economy-size SVD decomposition methods, but with higher efficiency. Whether or not step 4 converges depends largely on the initial condition. If the initial condition is representative of the initial firing times of only one IPT, it very likely will converge. However, if the initial firing times set for iteration are inaccurate, it may not converge.

5. Conclusions

Multi-channel sEMG signal decomposition has become more accessible since the introduction of LMMSE and CKC into the field. In this paper, SVD is utilized to decompose sEMG signals based on LMMSE and CKC. The number of extracted IPTs, computational efficiency and accuracy were improved by compressing the right unitary matrix $U(n)$. Dimensionality reduction simplifies the decomposition process and

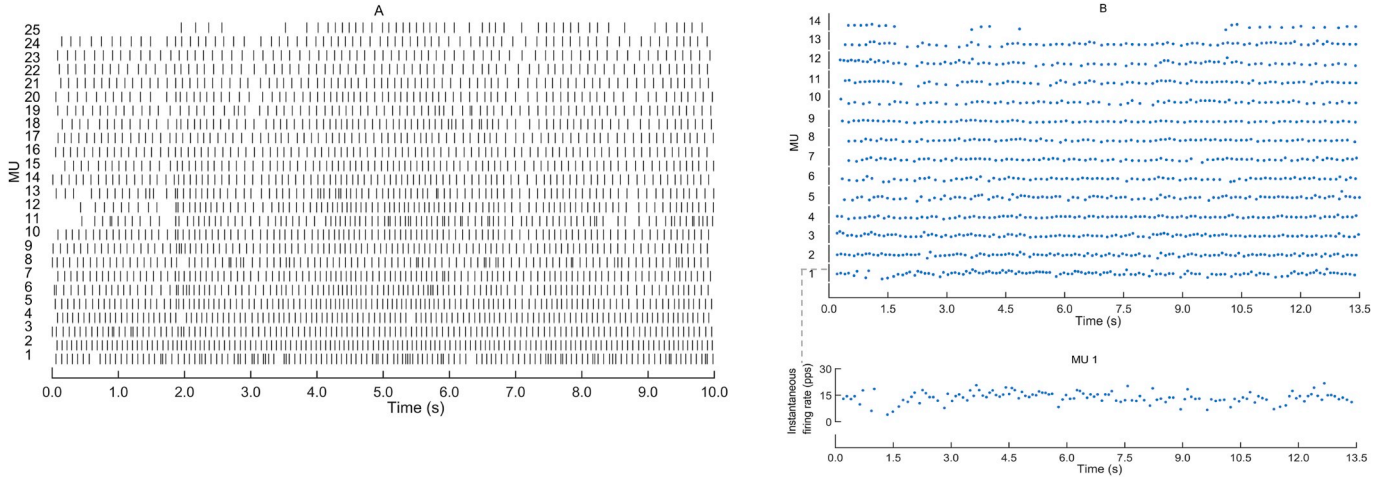

Fig. 7. The extracted IPTs and instantaneous firing rate (A) IPTs extracted from subject 1, (B) instantaneous firing rate of MUs from subject 7 with 15% MVC.

facilitates sEMG decomposition. As mentioned in 2.2, the important time-domain characteristic information of the sEMG signals is included in $U(n)$. Thus, an IPT can be estimated from only the right unitary matrix from the SVD of the measurement matrix. Our results indicate that dimensionality reduction of $U(n)$ can improve the efficiency and accuracy of sEMG decomposition in simulated and experimental signals, while greatly reducing computation time, even in cases of poor SNR.

Authors contributions

YN contributed to the experimental design, data analysis and paper writing; ND, JJ, and JL contributed to data analysis and paper writing; XL contributed to human subject recruitment and experimental data collection; YZ contributed to the study design, experimental design, result interpretation and paper writing.

Funding

This work was supported in part by National Natural Science Foundation of China 51677171, Zhejiang Provincial Natural Science Foundation of China LY17C100001, China Scholarship Council and the University of Houston.

Conflicts of interest

NY, ND, XL, JJ, JL and YZ have no conflicts of interest.

References

[1] D. Stashuk, Emg signal decomposition: how can it be accomplished and used? J. Electromyogr. Kinesiol. 11 (3) (2001) 151–173, https://doi.org/10.1016/S1050-6411(00)00050-X.
[2] K.C. McGill, Z.C. Lateva, H.R. Marateb, Emglab: an interactive EMG decomposition program, J. Neurosci. Methods 149 (2) (2005) 121–133, https://doi.org/10.1016/j.jneumeth.2005.05.015.
[3] J.R. Florestal, P.A. Mathieu, K.C. McGill, Automatic decomposition of multichannel intramuscular EMG signals, J. Electromyogr. Kinesiol. 19 (1) (2009) 1–9, https://doi.org/10.1016/j.jelekin.2007.04.001.
[4] S.H. Nawab, R.P. Wotiz, C.J. De Luca, Decomposition of indwelling EMG signals, J. Appl. Physiol. 105 (2) (2008) 700–710, https://doi.org/10.1152/japplphysiol.00170.2007.
[5] R.S. Lefever, C.J. De Luca, A procedure for decomposing the myoelectric signal into its constituent action-potentials-I: technique, theory, and implementation, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. (1982) 149–157, https://doi.org/10.1109/TBME.1982.324881 BME-29(3).
[6] X. Ren, C. Zhang, X. Li, G. Yang, Y.Z. Zhang, Intramuscular EMG decomposition basing on MUAPs detection and superposition resolution, Front. Neurol. 9 (2) (2018), https://doi.org/10.3389/fneur.2018.00002.
[7] J. Roussel, P. Ravier, M. Haritopoulos, D. Farina, O. Buttelli, Decomposition of multi-channel intramuscular EMG signals by cyclostationary-based blind source separation, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (11) (2017) 2035–2045, https://doi.org/10.1109/TNSRE.2017.2700890.
[8] M.G. Jahromi, H. Parsaei, A. Zamani, M. Dehbozorgi, Comparative analysis of wavelet-based feature extraction for intramuscular EMG signal decomposition, Journal of Biomedical Physics and Engineering 7 (4) (2017), https://doi.org/10.22086/jbpe.v0i0.538.
[9] P. Zhou, W.Z. Rymer, Muap number estimates in surface emg: template-matching methods and their performance boundaries, Ann. Biomed. Eng. 32 (7) (2004) 1007–1015, https://doi.org/10.1023/B:ABME.0000032463.26331.b3.
[10] M.R. Parker, R. Merletti, Electromyography: Physiology, Engineering, and Non-invasive Applications, Wiley Interscience, 2004 0-47.
[11] C.J. De Luca, A. Adam, R. Wotiz, L.D. Gilmore, S.H. Nawab, Decomposition of surface EMG signals, J. Neurophysiol. 96 (3) (2006) 1646–1657, https://doi.org/10.1152/jn.00009.2006.
[12] S.H. Nawab, S.S. Chang, C.J. De Luca, High-yield decomposition of surface EMG signals, Clin. Neurophysiol. 121 (10) (2010) 1602–1615, https://doi.org/10.1016/j.clinph.2009.11.092.
[13] A. Holobar, D. Zazula, Multichannel blind source separation using convolution kernel compensation, IEEE Trans. Signal Process. 55 (9) (2007) 4487–4496, https://doi.org/10.1109/TSP.2007.896108.
[14] A. Holobar, D. Zazula, Gradient convolution kernel compensation applied to surface electromyograms, International Conference on Independent Component Analysis and Signal Separation, Springer, Berlin, Heidelberg, 2007, pp. 617–624, , https://doi.org/10.1007/978-3-540-74494-8_77.
[15] Y. Ning, X.J. Zhu, S.A. Zhu, Y.C. Zhang, Surface EMG decomposition based on K-means clustering and convolution kernel compensation, IEEE J Biomed Health Inform 19 (2) (2015) 471–477, https://doi.org/10.1109/JBHI.2014.2328497.
[16] Y. Ning, Y. Zhao, A. Juraboev, P. Tan, J. Ding, J. He, Multichannel surface EMG decomposition based on measurement correlation and LMMSE, Journal of healthcare engineering (2018), https://doi.org/10.1155/2018/2347589.
[17] M. Chen, P. Zhou, A novel framework based on fastica for high density surface EMG decomposition, IEEE Trans. Neural Syst. Rehabil. Eng. 24 (1) (2016) 117–127, https://doi.org/10.1109/TNSRE.2015.2412038.
[18] D. Zazula, A. Holobar, An approach to surface EMG decomposition based on higher-order cumulants, Comput. Methods Progr. Biomed. 80 (2005) S51–S60, https://doi.org/10.1016/S0169-2607(05)80006-9.
[19] B.U. Kleine, J.P. van Dijk, B.G. Lapatki, M.J. Zwartsa, D.F. Stegemana, Using two-dimensional spatial information in decomposition of surface EMG signals, J. Electromyogr. Kinesiol. 17 (5) (2007) 535–548, https://doi.org/10.1016/j.jelekin.2006.05.003.
[20] I. Gligorijevic, J.P. van Dijk, B. Mijovic, S.V. Huffel, J.H. Blok, M.D. Vos, A new and fast approach towards sEMG decomposition, Med. Biol. Eng. Comput. 51 (5) (2013) 593–605, https://doi.org/10.1007/s11517-012-1029-y.
[21] M. Chen, X. Zhang, X. Chen, P. Zhou, Automatic implementation of progressive FastICA peel-off for high density surface EMG decomposition, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (1) (2018) 144–152, https://doi.org/10.1109/TNSRE.2017.2759664.
[22] Chen M, Zhang X, Lu Z, Li X, Zhou P. Two-source validation of progressive FastICA peel-off for automatic surface EMG decomposition in human first dorsal interosseous muscle. Int. J. Neural Syst.. doi:10.1142/S0129065718500193.
[23] F. Negro, S. Muceli, A.M. Castronovo, A. Holobar, D. Farina, Multi-channel intramuscular and surface EMG decomposition by convolutive blind source separation, J. Neural Eng. 13 (2) (2016) 026027, , https://doi.org/10.1088/1741-2560/13/2/026027.
[24] S.M. Key, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice-Hall Int., Englewood Cliffs, NJ, 1993.
[25] W. Ou, M.S. Hämäläinen, P. Golland, A distributed spatio-temporal EEG/MEG inverse solver, Neuroimage 44 (3) (2009) 932–946, https://doi.org/10.1016/j.neuroimage.2008.05.063.

[26] J. Clark, R. Plonsey, A mathematical evaluation of the core conductor model, Biophys. J. 6 (1) (1966) 95–112, https://doi.org/10.1016/S0006-3495(66)86642-0.
[27] Y. Ning, Y.C. Zhang, A new approach for multi-channel surface emg signal simulation, Biomed. Eng. Lett. (2017), https://doi.org/10.1007/s13534-017-0009-4.
[28] A.J. Fuglevand, D.A. Winter, A.E. Patla, Models of recruitment and rate coding organization in motor-unit pools, J. Neurophysiol. 70 (6) (1993) 2470–2488, https://doi.org/10.1152/jn.1993.70.6.2470.
[29] A. Holobar, M.A. Minetto, D. Farina, Accurate identification of motor unit discharge patterns from high-density surface EMG and validation with a novel signal-based performance metric, J. Neural Eng. 11 (1) (2014) 016008, , https://doi.org/10.1088/1741-2560/11/1/016008.