



Published in final edited form as:

Med Eng Phys. 2014 July ; 36(7): 975–980. doi:10.1016/j.medengphy.2014.04.003.

EMG Feature Assessment for Myoelectric Pattern Recognition and Channel Selection: A Study with Incomplete Spinal Cord Injury

Jie Liu¹, Xiaoyan Li², Guanglin Li³, and Ping Zhou^{2,4}

¹Sensory Motor Performance Program, Rehabilitation Institute of Chicago, IL, USA

²Department of Physical Medicine and Rehabilitation, University of Texas Health Science Center at Houston, and TIRR Memorial Hermann Research Center, Houston, TX, US

³Key Lab of Health Informatics of Chinese Academy of Sciences (CAS), Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong, China

⁴Biomedical Engineering Program, University of Science and Technology of China, Hefei, China

Abstract

Myoelectric pattern recognition with a large number of electromyogram (EMG) channels provides an approach to assessing motor control information available from the recorded muscles. In order to develop a practical myoelectric control system, a feature dependent channel reduction method was developed in this study to determine a small number of EMG channels for myoelectric pattern recognition analysis. The method selects appropriate raw EMG features for classification of different movements, using the minimum Redundancy Maximum Relevance (mRMR) and the Markov random field (MRF) methods to rank a large number of EMG features, respectively. A k-nearest neighbor (KNN) classifier was used to evaluate the performance of the selected features in terms of classification accuracy. The method was tested using 57 channels' surface EMG signals recorded from forearm and hand muscles of individuals with incomplete spinal cord injury (SCI). Our results demonstrate that appropriate selection of a small number of raw EMG features from different recording channels resulted in similar high classification accuracies as achieved by using all the EMG channels or features. Compared with the conventional sequential forward selection (SFS) method, the feature dependent method does not require repeated classifier implementation. It can effectively reduce redundant information not only cross different channels, but also cross different features in the same channel. Such hybrid feature-channel selection from a large number of EMG recording channels can reduce computational cost for implementation of a myoelectric pattern recognition based control system.

Keywords

EMG; myoelectric control; feature selection; channel reduction; spinal cord injury

Address Correspondence To: Ping Zhou, Ph.D. 1333B Moursund Street, Room 230, Houston, TX, 77030, dr.ping.zhou@ieec.org.

Competing interests: None declared.

Ethical approval: The study was approved by the Institutional Review Board of Northwestern University, Chicago, IL, USA (Reference number: STU00023682).

I. INTRODUCTION

Surface electromyogram (EMG) signals from residual muscles of amputee patients have been used for myoelectric prosthesis control for many years [1] [2]. Surface EMG can also be used as control signals for assistive or rehabilitative devices in robot-aided therapy for individuals with neurological injuries [3–6]. To increase dexterity of myoelectric control, EMG pattern recognition has been developed to overcome the limitations of conventional proportional control by extracting multiple features from EMG signals rather than solely relying on EMG amplitude [7–9]. Such features describing EMG signals can be used to identify different intended movements. To date, with advances in both feature extraction and classifier design techniques, high accuracies can be achieved in classification of different movements using surface EMG signals [10–12].

A recent emerging approach to assessing myoelectric control information is to use a large number of EMG recording channels (or high-density surface EMG recording) in combination with pattern recognition analysis. Such an approach has been applied in residual muscles of amputee subjects after the surgery of targeted muscle reinnervation [13], and in paretic muscles of stroke or spinal cord injury (SCI) subjects [14] [15]. It has been demonstrated that substantial neural control information can be extracted from the recorded muscles. In spite of encouraging findings from high-density surface EMG recording and analysis, using a large number of EMG electrodes hinders the practical application. Therefore, it is crucial to select a small number of appropriate channels which can yield the desired classification accuracy. Previous studies have used a sequential forward selection (SFS) method for this purpose [16] [17]. The method selects the best single channel for classification and then adds one channel at a time that can maximize classification performance in combination with the selected channels. With this method, the classifier has to be repeatedly implemented in order to select a single channel each time. Moreover, the redundancy of features in the same channel is not considered, since features extracted from each channel are fixed during the process. Thus, the SFS algorithm can not remove the already selected features that might become obsolete after the addition of other candidate features (or channels).

To overcome the limitations of the SFS method, this study introduces a different channel reduction strategy based on evaluation of EMG features extracted from high-density surface EMG recordings. The proposed strategy does not require repeatable implementation of the classification. Instead, it minimizes the number of channels by ranking the most discriminative features derived from all the EMG channels. The performance of such a strategy was confirmed by high classification accuracies using a small number of selected features, comparable to those derived from all the features or channels. The advantage of the proposed strategy was also demonstrated by comparing with the SFS channel selection method.

II. METHODS

A. Dataset description

The datasets used for this study were collected from 9 subjects with incomplete cervical spinal injury (6 males, 3 females; age range 31–62 year; Neurological injury level C4–C8; ASIA class: C or D; Upper extremity motor score: 30–45). The experimental protocols were detailed in our previous report [15]. In brief, subjects were comfortably seated and instructed to rest their forearm on a table. Six grasp patterns were performed including cylindrical grip, tool grip, power grip, lateral pinch, open pinch, and fine pinch. These patterns were selected because of their high usage frequency in daily life [24] [25]. For each pattern, subjects were instructed to perform the task with a comfortable and consistent level of effort for approximately 5 s, and then relax. This was repeated for 8 times, recorded in one trial. Thus, the surface EMG signals for six grasp patterns were recorded in six trials, respectively. Additionally, spontaneous EMG signal during hand relaxation was recorded as a different class. Sufficient relaxation time between trials and between repetitions of the same task were allowed.

A high-density EMG system (Refa 128 model, TMS International BV, Netherlands) was used to record surface EMG signals from the forearm and hand muscles in the weaker side of each subject [15]. A total of 57 surface electrodes (5 mm in diameter) were used for recording, among which 48 were placed over the forearm (Figure 1). To facilitate placement of the electrodes, 8 equally spaced electrodes were first incorporated into a custom-made stretchable fabric belt. In total, six electrode belts were made and equally placed around the forearm at locations from approximately 12.5% to 75.0% (with 12.5% increments) of the entire distance from the medial epicondyle of the humerus to the styloid process of the ulna. In addition to the forearm, 9 electrodes, 3 in group, were placed on three hand muscles (the first dorsal interosseous, thenar and hypothenar muscles), respectively. A reference electrode was located on the olecranon with a system feedback of the mean of all the recording channels provided to each individual channel. The surface EMG signals were sampled at 2000 Hz per channel, with a band pass filter setting at 20–500 Hz.

B. Feature extraction

A manual data segmentation scheme based on examination of the EMG amplitude was used to determine the onset and offset of the active segments, corresponding to each repetition of a hand task. For each active segment, 57-channel EMG data were further divided into a series of overlapping analysis windows (window length: 256 ms, window overlap: 128 ms). For each analysis window, a set of features was extracted to characterize the EMG signals for classification of the different tasks. Two feature sets were investigated in this study: the time domain (TD) feature set which includes four time domain statistics, i.e. the mean absolute value (MAV), the number of zero crossings (ZC), the waveform length (WL), and the number of slope sign changes (SSC); and the AR+RMS feature set which includes a sixth order autoregressive (AR) model coefficients plus the root mean square (RMS) amplitude of the signal [8]. For each analysis window, the features extracted from all the channels were concatenated to form a feature vector.

C. Raw EMG feature selection

For each analysis window a large number of EMG channels resulted in a very high order of feature vector dimensions. These raw EMG features were then ranked using the minimum Redundancy Maximum Relevance (mRMR) criteria [18] and the Markov random field (MRF) based Fisher-Markov selector [19], respectively.

The mathematical details of the mRMR and the MRF feature selection methods and their implementation are described in [18] [19]. In brief, the mRMR criteria choose features that are mutually dissimilar to each other and marginally similar to the classification variable, by ranking candidate component features based on compromise between relevance and redundancy which can be measured in different forms such as mutual information, correlation, distances, etc [18]. If a ranked feature has a smaller index, it achieves a better trade-off between the maximum relevance and minimum redundancy and will be more important for classification. Mutual information difference criteria were used in this study to search for the features in mRMR optimization conditions.

The MRF based selector aims to identify the best subset of candidate features that maximizes the between-class distance while minimizing the within-class distance in a higher dimensional kernel space [19]. The MRF feature selection was designed toward efficiently selecting the globally optimal subset of features, which are the most useful in characterizing differences among the possible classes. By using special kernel functions, MRF optimization techniques can be employed to efficiently solve the formulated objective functions and achieve global optimum for feature selection [19]. The MRF method assigns a coefficient indicating its importance to each feature. The larger coefficients indicate higher importance of the selected features. Thus, the distribution of such coefficients relative to the raw EMG features can be used to rank the features.

D. Performance evaluation of raw feature selection

The performance of the raw EMG feature selection method was assessed by the resultant classification accuracy. In this study, the k-nearest neighbor (KNN) classifier (k=5) [20] was used. An eightfold cross-validation scheme was used to evaluate the classification performance. The EMG data within any random seven active segments were assigned as training dataset, and sequentially the EMG data of the remaining active segment were used as testing dataset. The performance accuracy for each intended hand pattern was the percentage of correctly classified windows over all the analysis windows in its testing dataset. An overall performance was then calculated as the percentage of correctly classified windows over all the analysis windows in the testing datasets across all hand grasp patterns. The performance of the raw EMG feature selection was further compared with the SFS based method previously used for channel reduction of high-density surface EMG recordings [17]. The SFS methods using individual raw EMG features and individual channels were respectively implemented for comparison.

III. RESULTS

A. Classification performance with raw EMG feature selection

The performance of the raw EMG feature selection was evaluated by the resultant classification accuracies, as demonstrated in Figure 2, where the classification accuracies achieved from different feature sizes were compared, for TD and AR+RMS feature sets respectively. A maximum of 20 raw EMG features were respectively selected using the mRMR, MRF and SFS (based on individual raw features) methods. We observed that for the TD feature set, the MRF method achieved the best performance among the three methods. For the AR+RMS feature set, the MRF and mRMR methods achieved similar performance, slightly better than the SFS method. With sequentially adding the selected raw EMG features, the classification accuracy dramatically increased at the beginning and then gradually approached to a maximum value. The results indicate that it is feasible to greatly reduce the number of raw EMG features or channels) while maintaining high classification accuracies achieved using high-density surface EMG. For example, across all subjects, the selected 4 raw EMG features using the MRF method achieved an average overall classification accuracy higher than 90%.

B. Comparing with individual channel based SFS method

The comparison of the raw EMG feature selection using the MRF and SFS (based on individual channels) methods is demonstrated in Figure 3. It was observed that the MRF method approached the maximum classification accuracies more quickly than the SFS method. The average overall classification accuracy of approximately 95% was achieved by 7 TD features (Figure 3a) or 6 AR+RMS features (Figure 3b) selected using the MRF method, which was similar to that achieved by 20 TD features or 42 AR+RMS features selected using the individual channel based SFS method. This can also be observed from the channel point of view. With increase of the channel number, the MRF method tended to approach similar or higher classification accuracies compared with the SFS method for both TD (Figure 3c) and AR+RMS (Figure 3d) feature sets. However, the MRF method required extraction of significantly fewer EMG features per channel than the SFS method.

C. Analysis of selected raw EMG features and channels

For each subject, we further investigated the selected raw EMG features using the MRF method. Figure 4 shows an example of the distribution of the coefficients used to determine the importance of each raw EMG feature [19]. From the coefficient distribution a small fraction of highly discriminative features in the presence of a large number of irrelevant or redundant features can be selected. From Figure 4 it was observed that for the TD feature set the waveform length (WL) was the most important feature for the datasets, while the RMS amplitude was the most important feature among the AR+RMS feature set. We examined the selected features using the MRF method and the corresponding EMG channels for each subject. It was found that for either TD or AR+RMS feature set, no consistent patterns of selected channels were observed across different subjects. Furthermore, for the same subjects, the selected channels might be partly or completely different for different feature sets (for example, Subject 5: first 4 channel indexes: 10, 26, 2, 37 for TD, and 10, 2, 37, 14 for AR+RMS; Subject 1: 4, 47, 12,7 for TD, and 6, 44, 14, 10 for AR+RMS).

IV. DISCUSSIONS

This study presents a different strategy from the SFS method to reduce the number of channels used in high-density surface EMG recordings, toward developing a practical myoelectric control system. Recent development in EMG instrumentation, particularly in EMG amplifier design, has allowed simultaneous recording of more than 100 channels from single or multiple muscles. Commercial high-density surface EMG data acquisition systems have become available (e.g. TMS International BV, Netherlands; OT Bioelettronica, Italy). High-density surface EMG recording has achieved increasing applications. For example, using an electrode array recording from a single muscle, it is possible to extract motor unit activity and the spatial information which may be impractical to obtain with a needle electrode [21] [22].

An important application of high-density surface EMG is to perform pattern recognition analyses toward developing a practical myoelectric control system for helping restoration of limb function (of amputee, stroke and SCI subjects) [13–15] [23]. The results of the current study using the raw feature selection methods are in consistent to the previous findings with different subject populations that there is much redundant information contained in high-density surface EMG recordings, which should be removed for implementation of a practical myoelectric control system [13–15]. It is noted that the redundancy of information does not compromise the importance of high-density EMG recording and analysis. An important finding from the previous analyses was that myoelectric pattern recognition performance was subject-specific, for both amputee subjects who received nerve transfer surgery [13] and subjects with neurological disorders [14] [15]. There was lack of common patterns across each group of tested subjects in regard to class-to-class misclassification performance, the best location or number of the EMG electrodes, etc. This may be due to the difference between subjects (such as impairment level, injury location, or different nerve transfer surgeries, etc). Thus, the most appropriate design (such as number and locations of EMG channels for myoelectric pattern recognition) is not available for each specific subject. High-density surface EMG analysis is useful for guiding implementation of a practical myoelectric control system because it offers an approach to appropriate selection of a small number of EMG channels and determination of their locations.

To select optimal channels from high-density surface EMG for myoelectric pattern recognition, the performance of all the combinations of the desired number of channels should be assessed to select the one with the highest classification accuracy. Such an approach is impractical considering a large number of possible combinations (for example, selection of 5 from 57 channels would require testing of 4,187,106 different combinations). Therefore, in the previous studies, a sub-optimal method of sequential feed-forward selection (SFS) was used, which iteratively added one channel producing the highest classification accuracy combined with the previously selected channels. However, the SFS algorithm can not remove selected features that become obsolete after the addition of other candidate channels. Instead of focusing on individual channels, the current study used a feature dependent method to select the most informative EMG features in order to reduce the number of EMG channels without significantly degrading classification performance. The feature dependent channel selection has two advantages over the SFS method. First,

each iteration of the SFS method requires repeatedly training and testing of the classifier for $N-m+1$ times (N is the total number of electrodes, m is the iteration index, $m=1, 2, \dots, M$, M is the number of selected channels), while the feature dependent method does not require classifier implementation. Instead, it ranks and selects discriminative features applicable to different classifiers in general. Second, the feature dependent channel selection treats features from the same or different channels in the same manner. Thus it can effectively reduce the redundant information not only cross different channels, but also cross different features in the same channel. Compared to the SFS based channel selection, the hybrid feature-channel selection method can reduce computational cost for implementation of a practical myoelectric control system.

It is noted that even for the same subject, the selected EMG channels might not be the same for TD and AR+RMS feature sets, suggesting that the channel selection is not only subject specific, but also related to different feature sets. Therefore, determination of appropriate number and location of EMG channels during implementation of a practical myoelectric control system (for neurological injury rehabilitation) should consider user difference (such as user need, remaining motor control capacity, etc). Additionally, the effects of different features on classification performance and channel selection should also be considered.

In this study, the performance of the EMG feature selection methods was evaluated using a combination of TD or AR+RMS feature set and the KNN classifier. The hybrid feature-channel selection method can be applied to other features. The selected features are also suitable for pattern recognition analysis with other classifiers in general. The hybrid feature-channel selection was tested using EMG data from incomplete SCI subjects. The method can also be used for high-density EMG recordings from different populations. As an evaluation criterion, calculation of average overall classification accuracies was used to confirm the performance of the selected features or channels in this study. We acknowledge that the robustness of the individual features (such as with respect to electrode shift, electrode size, orientation, etc) should also be considered for designing or implementing a practical myoelectric control system.

Acknowledgments

We would like to acknowledge the availability of the code at <http://penglab.janelia.org/proj/mRMR/> for the mRMR method implementation, and at <http://www.cs.siu.edu/~qcheng/featureselection/> for the MRF method implementation.

Funding: This work was supported in part by the National Institutes of Health of the U.S. Department of Health and Human Services under Grant 1R21NS075463, in part by the Memorial Hermann Foundation, and in part by the National Natural Science Foundation of China under Grant 61135004.

REFERENCES

1. Parker P, Englehart K, Hudgins B. Myoelectric signal processing for control of powered limb prostheses. *Journal of Electromyography and Kinesiology*. 2006; 16(6):541–548.
2. Zecca M, Micera S, Carrozza MC, Dario P. Control of multifunctional prosthetic hands by processing the electromyographic signal. *Critical Review in Biomedical Engineering*. 2002; 30(4–6):459–485.

3. Dipietro L, Ferraro M, Palazzolo JJ, Krebs HI, Volpe BT, Hogan N. Customized interactive robotic treatment for stroke: EMG-triggered therapy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2005; 13(3):325–334. [PubMed: 16200756]
4. Song R, Tong K, Hu X, Li L. Assistive control system using continuous myoelectric signal in robot-aided arm training for patients after stroke. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2008; 16(4):371–379. [PubMed: 18701384]
5. Zariffa J, Kapadia N, Kramer JL, Taylor P, Alizadeh-Meghbrazi M, Zivanovic V, Willms R, Townson A, Curt A, Popovic MR, Steeves JD. Feasibility and efficacy of upper limb robotic rehabilitation in a subacute cervical spinal cord injury population. *Spinal Cord*. 2012; 50(3):220–226. [PubMed: 21912402]
6. Jiang N, Falla D, d'Avella A, Graitmann B, Farina D. Myoelectric control in neurorehabilitation. *Critical Review in Biomedical Engineering*. 2010; 38(4):381–391.
7. Englehart KB, Hudgins B. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*. 2003; 50(7):848–854. [PubMed: 12848352]
8. Huang Y, Englehart KB, Hudgins B, Chan AD. A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses. *IEEE Transactions on Biomedical Engineering*. 2005; 52(11):1801–1811. [PubMed: 16285383]
9. Scheme E, Englehart K. Electromyogram pattern recognition for control of powered upper-limb prostheses: state of the art and challenges for clinical use. *Journal of Rehabilitation Research & Development*. 2011; 48(6):643–659.
10. Hargrove LJ, Li G, Englehart KB, Hudgins BS. Principal components analysis preprocessing for improved classification accuracies in pattern-recognition-based myoelectric control. *IEEE Transactions on Biomedical Engineering*. 2009; 56(5):1407–1414. [PubMed: 19473932]
11. Huang H, Zhou P, Li G, Kuiken T. Spatial filtering improves EMG classification accuracy following targeted muscle reinnervation. *Annals of Biomedical Engineering*. 2009; 37(9):1849–1857. [PubMed: 19526342]
12. Chen X, Zhang D, Zhu X. Application of a self-enhancing classification method to electromyography pattern recognition for multifunctional prosthesis control. *Journal of Neuroengineering and Rehabilitation*. 2013; 10(1):44. [PubMed: 23634939]
13. Zhou P, Lowery M, Englehart KB, Huang H, Li G, Hargrove L, Dewald JPA, Kuiken TA. Decoding a new neural-machine interface for control of artificial limbs. *Journal of Neurophysiology*. 2007; 98:2974–2982. [PubMed: 17728391]
14. Zhang X, Zhou P. High density myoelectric pattern recognition towards improved stroke rehabilitation. *IEEE Transactions on Biomedical Engineering*. 2012; 59(6):1649–1657. [PubMed: 22453603]
15. Liu J, Zhou P. A novel myoelectric pattern recognition strategy for hand function restoration after incomplete cervical spinal cord injury. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2013; 21:96–103. [PubMed: 23033334]
16. Nazarpour K, Sharafat AR, Firoozabadi SM. Application of higher order statistics to surface electromyogram signal classification. *IEEE Transactions on Biomedical Engineering*. 2007; 54(10):1762–1769. [PubMed: 17926674]
17. Huang H, Zhou P, Li G, Kuiken TA. An analysis of EMG electrode configuration for targeted muscle reinnervation based neural machine interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2008; 16(1):37–45. [PubMed: 18303804]
18. Peng H, Long F, Ding C. Feature selection based on mutual information: Criteria of max-dependency, max-relevance and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2005; 27(8):1226–1238. [PubMed: 16119262]
19. Cheng Q, Zhou H, Cheng J. The Fisher-Markov selector: fast selecting maximally separable feature subset for multiclass classification with applications to high-dimensional data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2011; 33:1217–1233. [PubMed: 21493968]
20. Cover TM, Hart PE. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*. 1967; 13:21–27.

21. Merletti R, Holobar A, Farina D. Analysis of motor units with high-density surface electromyography. *Journal of Electromyography and Kinesiology*. 2008; 18(6):879–890.
22. Drost G, Stegeman DF, van Engelen BG, Zwarts MJ. Clinical applications of high-density surface EMG: a systematic review. *Journal of Electromyography and Kinesiology*. 2006; 16(6):586–602.
23. Daley H, Englehart K, Hargrove L, Kuruganti U. High density electromyography data of normally limbed and transradial amputee subjects for multifunction prosthetic control. *Journal of Electromyography and Kinesiology*. 2012; 22(3):478–484.
24. Cutkosky MR. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *IEEE Transactions on Robotics and Automation*. 1989; 5(3):269–279.
25. Zheng JZ, Rosa SDL, Dollar AM. An investigation of grasp type and frequency in daily household and machine shop tasks. *Proceedings of IEEE International Conference on Robotics and Automation*. 2011:4169–4175.

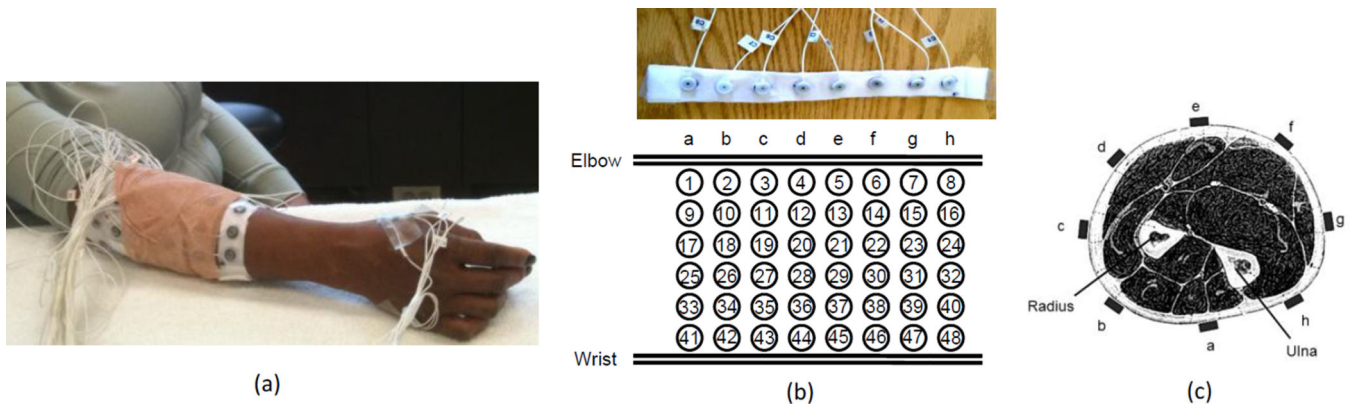


Figure 1. (a) An overall review of the electrode placement for 57-channel EMG signal recordings; (b) The channel index for the 48 electrodes placed on the forearm; (c) Cross section of the forearm and the electrode positions.

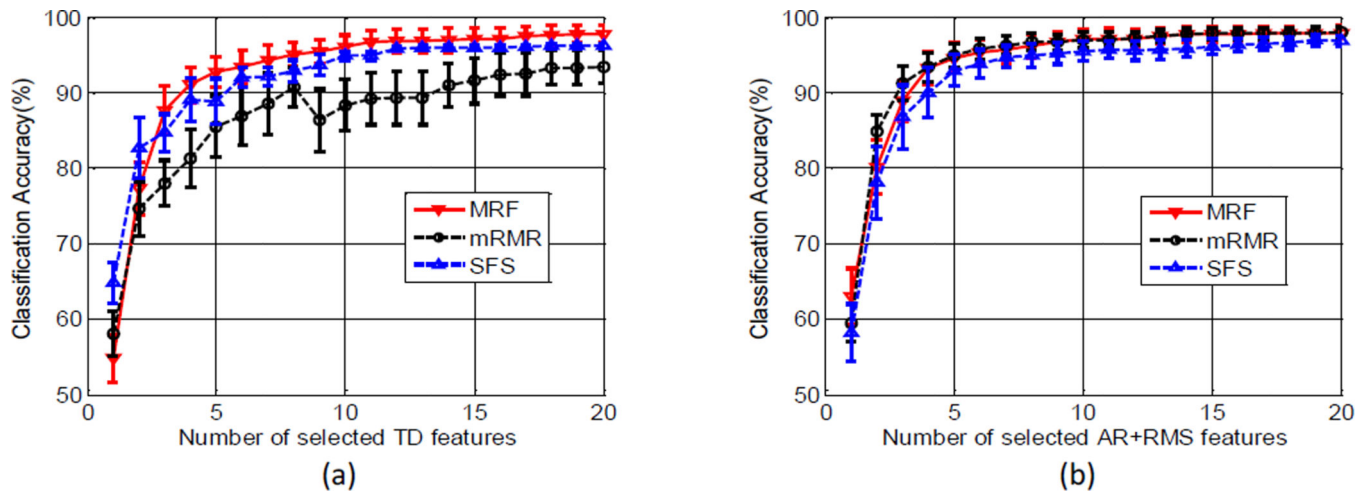


Figure 2.

Comparison of the mRMR, MRF and SFS methods for raw EMG feature selection. The SFS method was based on individual raw EMG features. The average overall classification accuracy across all subjects was presented as a function of the number of selected TD or AR +RMS features using the three methods, respectively.

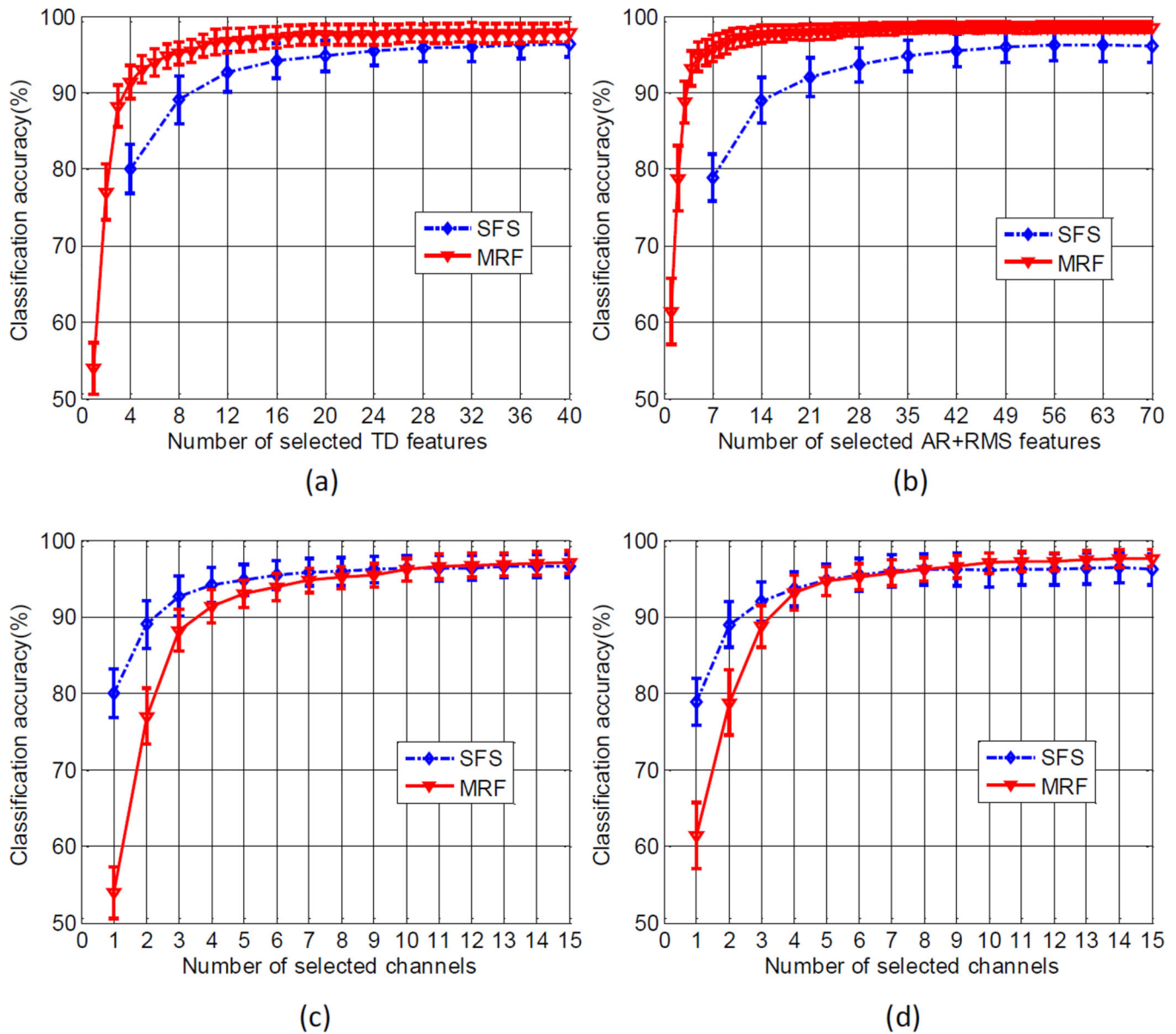


Figure 3.

(a) and (b): comparison of the classification performance with increase of the number of raw EMG features selected by the MRF and SFS methods for TD and AR+RMS features, respectively. The SFS method was based on individual EMG recording channels. (c) and (d): comparison of the classification performance with increase of the number of EMG channels selected by the MRF and SFS methods for (c) TD and (d) AR+RMS features, respectively.

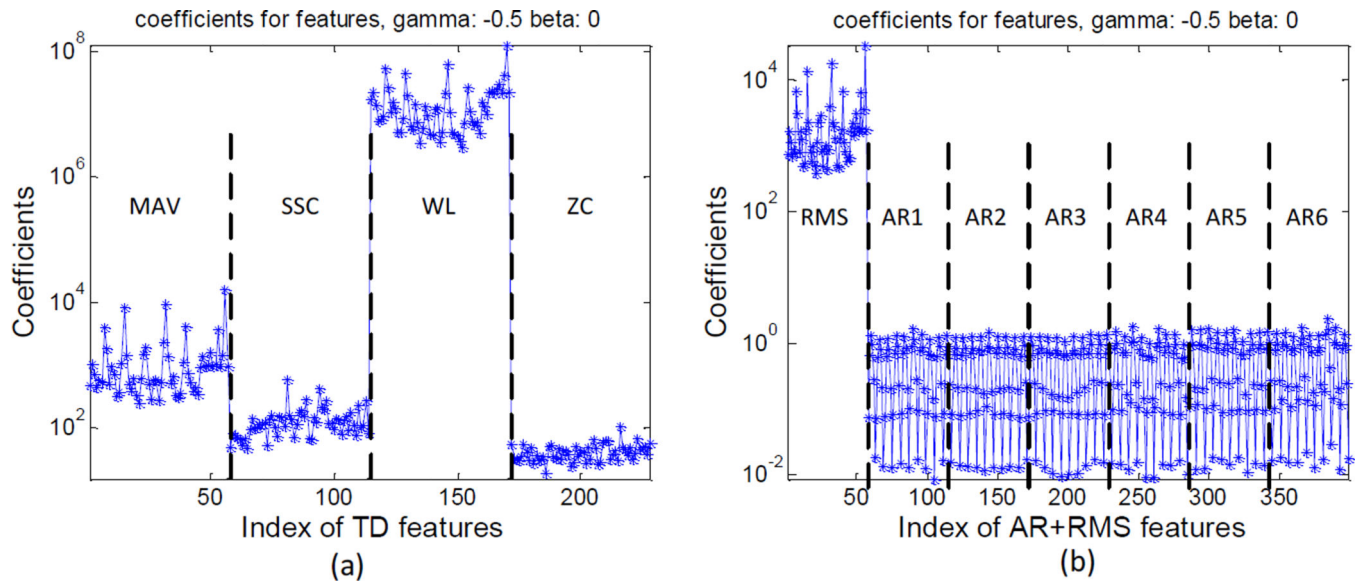


Figure 4.

An example of the importance coefficient distribution of the raw EMG features using the MRF method. β and γ are two parameters used in the MRF feature selection criterion [19]. β is a constant factor related to the number of selected features from the high dimensional feature vectors. γ is a free parameter. Both β and γ were set the same values as suggested in [19].