Australas Phys Eng Sci Med (2016) 39:59–69
DOI 10.1007/s13246-015-0395-9

SCIENTIFIC PAPER

A sparse Bayesian learning based scheme for multi-movement recognition using sEMG

Shuai Ding[1] · **Liang Wang**[1]

Received: 9 May 2015 / Accepted: 1 November 2015 / Published online: 17 November 2015

Abstract This paper proposed a feature extraction scheme based on sparse representation considering the non-stationary property of surface electromyography (sEMG). Sparse Bayesian learning was introduced to extract the feature with optimal class separability to improve recognition accuracy of multi-movement patterns. The extracted feature, sparse representation coefficients (SRC), represented time-varying characteristics of sEMG effectively because of the compressibility (or weak sparsity) of the signal in some transformed domains. We investigated the effect of the proposed feature by comparing with other fourteen individual features in offline recognition. The results demonstrated the proposed feature revealed important dynamic information in the sEMG signals. The multi-feature sets formed by the SRC and other single feature yielded more superior performance on recognition accuracy, compared with the single features. The best average recognition accuracy of 94.33 % was gained by using SVM classifier with the multi-feature set combining the feature SRC, Williston amplitude (WAMP), wavelength (WL) and the coefficients of the fourth order autoregressive model (ARC4) via multiple kernel learning framework. The proposed feature extraction scheme (known as SRC + WAMP + WL + ARC4) is a promising method for multi-movement recognition with high accuracy.

✉ Shuai Ding
aztlaztl@163.com

[1] School of Automation Science and Electrical Engineering, Beihang University, NO. 37, Xueyuan Road, Haidian District, Beijing, China

Keywords Surface electromyography (sEMG) · Feature extraction · Non-stationarity · Sparse representation · Temporal MMV sparse Bayesian learning (T-MSBL)

Introduction

Surface electromyography (sEMG) signal is a kind of the electrical activity generated during muscular contractions [1]. The signal can be recorded from the surface of skeletal muscle by using surface electrodes. Studies have shown that the sEMG signal is a highly useful electrophysiological signal in the fields of medicine and engineering. Efficient recognition of human multi-movement pattern using sEMG has attracted research interests in human machine interaction and rehabilitation engineering fields during the past three decades [2, 3].

Researchers have identified some sEMG-based movements including head, gross hand, finger, wrist, arm, lower limb and facial movements. Most studies have mainly focused on pattern recognition of arm, wrist and gross hand movements [4, 5]. Despite high recognition accuracy (RA), these movements cannot completely satisfy the needs of most myoelectric systems. Some researchers have studied the recognition of finger movements. However, it is still a challenging task to recognize subtle hand movements, such as individual finger movements and multi-finger movements, with the high accuracies. It is chiefly because most of human forearm muscles are small and long skeletal muscles. The muscles cover the forearm layer by layer and they are intertwined. The same muscle contracts when we perform two different movements. Under this circumstance, we think similarity exists between the two types of movements. The similar movements are generally identified with low recognition accuracies. The pattern-

recognition-based strategy has been proposed to recognize multi-movement patterns, and to enhance the classification accuracy. Hargrove et al. had found that the features extracted from sEMG are more crucial than the classifiers for the performance of movement recognition [6]. The effective and distinguishable features make contribution to improve RA radically. This paper focused on extracting the features with optimal class separability from sEMG to improve the RA of multi-movement patterns, especially subtle hand movements.

Along with the development of signal processing theory, a large number of different methods have been used for sEMG feature extraction. The time, frequency and time–frequency processing methods are widely used to extract sEMG features for movement classification. The features in time domain group include mean absolute value (MAV), root mean square (RMS), zero crossing (ZC), slope sign change (SSC), Williston amplitude (WAMP), integral of sEMG (IEMG), and wavelength (WL) and so on [7, 8]. The features in frequency domain group include autoregressive model (AR) coefficients and power spectral density (PSD) and more [9, 10]. The time–frequency features can be extracted by using time–frequency analysis methods, such as wavelet transform and wavelet packet transform [10, 11]. These analysis methods of feature extraction have their own advantages and shortcomings. The time and frequency processing methods are quite simple. The sEMG signal is assumed as a stationary or short-time stationary signal in both time and frequency domains. The neurophysiological system is a complicated non-linear dynamic system, which leads to the non-stationarity in sEMG signal. Published reports have indicated that the potential non-stationarity in sEMG signal affects the performance of movement recognition [12]. Therefore, the time and frequency analysis methods have their limitations to represent the sEMG signal. The wavelet transform and wavelet packet transform are powerful tools in analyzing the non-stationary signals. They can be used in multi-resolution analysis of time domain and frequency domain. However, they generally do not lead to sparse decompositions (i.e., with very few non-zero decomposition coefficients), so the time–frequency resolution is not very high. Thus we cannot obtain the features with good class separability for subtle movement recognition.

Considering the limitation of conventional sEMG analysis methods, we introduced sparse representation to feature extraction of non-stationary sEMG signal in this study [13]. Studies have shown that sparse representation of signal over a dictionary (i.e., the set of time–frequency atoms) is applicable for analyzing and processing the non-stationary signals [14]. Sparse representation model holds that the signal can be decomposed into a linear combination of a few columns chosen from a dictionary. The signal sparse representation aims to represent the original signal using as few atoms in the dictionary matrix as possible. Sparse representation provides adaptive time–frequency analysis for the non-stationary sEMG with high time–frequency resolution. The atoms with nonzero representation coefficients capture the key and significant properties of original sEMG signal and reflect important information of movement patterns. Different movements produce different sEMG signal waveforms. Hence the corresponding nonzero representation coefficients provide hints for multi-movement recognition. Many natural signals tend to be compressible rather than sparse [15]. The sEMG signal is not sparse in the time domain. But the signal is compressible in some transformed domains [16]. The sEMG signal can be sparsely represented by the nonzero representation coefficients with maximum efficiency. These nonzero coefficients contain the significant information of sEMG signal, and they can be used as the sEMG features for multi-movement recognition.

The idea of sparseness has been used in many fields of sEMG investigation, such as signal decomposition [17], blind source separation (BSS) [18–20], signal classification [21], signal recovery [22], and classifier design [23], etc. However, in addition to the wavelet analysis, some other methods for sparse representation are not exploited widely for the sEMG-based multi-movement recognition, especially for feature extraction of sEMG signal. In most recent study, signal processing methods, such as independent component analysis (ICA) [20, 24] and non-negative matrix factorization (NMF) [25], which are closely related to sparseness methods, have been used for the sEMG-based hand movement recognition. ICA and NMF are matrix factorization techniques. ICA converts a multidimensional vector into statistically independent components (i.e., ICs). In the BSS and ICA, the sources can be extracted or recovered to find the non-gaussianity maximization points. The non-gaussianity makes ICA estimation (i.e., estimation of signal sources) sparse to some extent [26]. NMF uses multiplicative updates to ensure the non-negativity of the factorisation. NMF provides sparseness of the solutions as a result of a side effect caused by non-negativity constraints [27]. Studies have shown that sparse Bayesian learning (SBL) algorithm can improve the performance of sparse decomposition and be well-suited for finding maximally sparse representation [28, 29]. This property would make the SBL algorithm valuable to extract the features with good separability from sEMG signals. Through consulting a great deal of literature, we found that there has not been any work done on the analysis of SBL for the sEMG-based application of movement recognition. Temporal multiple measurement vector sparse Bayesian learning (T-MSBL) algorithm is an improved algorithm based on the SBL algorithm [30]. This paper employed T-MSBL

algorithm to extract effective features from the sEMG signals for multi-movement recognition.

We preliminarily focused on the multi-movement recognition for healthy non-amputees. The rest of this paper is organized as follows: "Methods" section presents the proposed scheme for sEMG feature extraction. "Experiments and discussion" section provides experiments to demonstrate the performance of the proposed scheme. Conclusions are given in the last section.

Methods

Subjects and data acquisition

The sEMG signals were amplified by the BIOPAC EMG100C amplifier (gain: 2000). The raw sEMG signals were sampled at a rate of 1000 Hz and acquired using the BIOPAC MP150 data monitoring system (BIOPAC Systems Inc., Goleta, CA, USA). The signals were filtered between 20 and 500 Hz. The signals were collected by several pairs of standard disposable Ag/AgCl ECG disc electrodes with conductive paste. All electrodes were pasted on the skin surface of forearm of each subject. The skin areas where the electrodes were attached were degreased beforehand with 95 % medicinal alcohol to ensure a better interface between the skin and the electrodes. The pair of electrodes had a center-to-center spacing of 15 mm.

Six human hand/finger movement patterns were planned to classify: (1) thumb extension (EXTF), (2) index extension (EXIF), (3) thumb flexion (FLTF), (4) palm extension (EXPM), (5) HOOK and (6) OKAY in our work, as shown in Fig. 1. The movements selected in this study are more subtle, compared with the movements of arm and wrist [4, 5]. And they are common and easy to perform in our daily lives. Our preliminary work focused on the movements with no external load. We considered not only mutually opposite movements (e.g., EXTF and FLTF), but also similar movements (e.g., EXIF and HOOK, EXPM and OKAY) to further complicate the movement recognition. We collected four channels of sEMG to classify these hand/finger movements. According to human anatomy, four pairs of electrodes were pasted over extensor pollicis brevis (EPB), flexor pollicis longus (FPL), extensor indicis proprius (EIP) and extensor digitorum (EDM), respectively [31]. The muscle positions for the electrode locations are generally used to acquire sEMG signal for recognizing the six classes of movements [32].

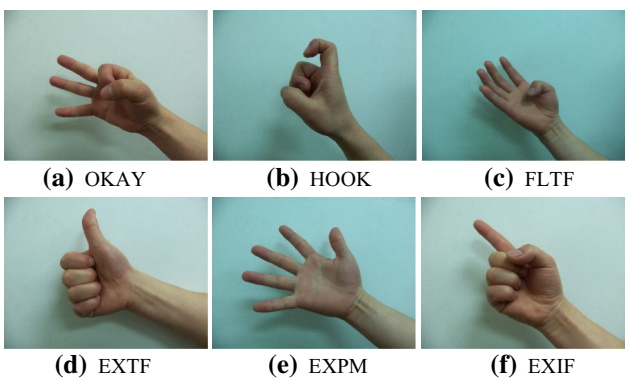

Fig. 1 Six classes of hand and finger movements: **a** OKAY, **b** HOOK, **c** FLTF, **d** EXTF, **e** EXPM, and **f** EXIF

Twelve able-bodied subjects (six females and six males, aged 20–28 years) participated in our experiment. All subjects had signed their written informed consent. Beihang University ethics committee approved the study protocol. Each subject was conducted to perform twenty groups of individual finger and multi-finger movements. The subject completed six classes of movements as described above and each movement was repeated ten times in each group. Each movement was held for 2–3 s. The interval of adjacent movements in each group was 3 s. The subjects relaxed their hands without any muscle contraction during the interval. The subjects should wait at least 2 h before the next group to avoid muscle fatigue. Steady-state sEMG (i.e., the continuous sEMG samples recorded after 300 ms once the contraction was established) were analyzed to discriminate the different movements.

The steady-state sEMG data were segmented by the analysis window to extract the features. The window increment (i.e., the overlap between two adjacent analysis windows) would determine the response of myoelectric control systems. The response of system should be less than 300 ms for reducing users' perceived lag [8]. Therefore, the analysis window was set to 256 ms sliding (one segment) with a 128 ms window increment. For the twenty groups of collected data, training and testing of one classifier were implemented using leave-one-out cross-validation.

Sparse reconstruction modeling of sEMG

Sparse representation of sEMG is to find a linear combination of a minimum number of atoms from an over-complete dictionary to approximate the original signal.

$$\hat{\rho} = \arg\min_{\rho}\left(\|z - G\rho\|_2^2 + \zeta\upsilon(\rho)\right) s.t. \quad z = G\rho + \varepsilon \qquad (1)$$

where $z \in \mathbf{R}^{d\times 1}$ denotes the sEMG signal vector, $G \in \mathbf{R}^{d\times K}(K > d)$ denotes the over-complete dictionary, $\rho \in \mathbf{R}^{K\times 1}$ is the vector of sparse coefficients (i.e., the representation of signal), $\upsilon(\rho)$ is sparsity constraint item, usually

it can be expressed as $\upsilon(\rho) = \|\rho\|_1$. $\zeta(\zeta > 0)$ is the balance parameter between the data fidelity term $\|z - G\rho\|_2^2$ and $\upsilon(\rho)$, ε denotes the modeling errors.

As indicated above, z in the time domain and ρ in the G domain are equivalent representations of sEMG signal. Once $\hat{\rho}$ has been estimated, nonzero entries in $\hat{\rho}$ express the distinguishable features of the original sEMG, which provide the recognition hints of individual finger and multi-finger movements.

One of the major challenges for the common sparse representation algorithms is the high computational cost and large memory capacity. For such a reason, we also introduce compressed sensing (CS) theory [33] into the feature extraction scheme of sEMG to alleviate the problem. The original sEMG signals are first mapped to lower dimensional measurement domain using the CS theory in the proposed scheme. Then representation coefficients of the signals are estimated by solving convex optimization problems. The coefficients of sparse representation are used as characteristic parameters for the recognition task of multi-movement patterns. Assume the sEMG signal $s \in \mathbf{R}^{K\times 1}$ is compressible in some transform domain D, i.e., $s = D\rho$. The sparse reconstruction modeling of the sEMG signal can be described mathematically as follows.

$$y = \Theta s + e = \Theta D\rho + e = A\rho + e \tag{2}$$

where $\Theta \in \mathbf{R}^{c\times K}$ is a $c \times K$ random measurement matrix ($c < K$), $D \in \mathbf{R}^{K\times K}$ is the dictionary matrix for sparse representation and assume that the basis is orthonormal for simplicity, $A \in \mathbf{R}^{c\times K}$ is written as $A = \Theta D$, $\rho \in \mathbf{R}^{K\times 1}$ denotes sparse coefficients of the original signal s over the dictionary D, $y \in \mathbf{R}^{c\times 1}$ is the vector of the compressed measurements in the low dimension space, e is the modeling errors.

T-MSBL algorithm

Sparse Bayesian learning (SBL) algorithm is a powerful tool for sparse representation of signals. This algorithm uses the prior to promote the sparsity of signal to the largest extent. The SBL algorithm is equivalent to an algorithm of iterative reweighted L1-norm minimization. Many studies have shown that iterative reweighted algorithms obtain more sparse solutions and better performance than conventional algorithms do [29]. The introduction of SBL algorithm to feature extraction of sEMG would help to obtain the features with good class separability.

SBL algorithm was developed as a method of machine learning originally and the basic SBL algorithm was derived by Tipping [28]. Then it has been applied to sparse representation, regression and CS [16]. T-MSBL algorithm, a member in the family of SBL, is mainly adopted to solve multiple measurement vectors (MMV) problem. The T-MSBL has superior performance when the dictionary matrix is highly coherent. Single measurement vector (SMV) model is a special case of the MMV model. The T-MSBL algorithm can also be applicable to solving the SMV problem. Under this circumstance, the TMSBL algorithm is similar to the expectation–maximization-based SBL (EM-SBL) [34]. The key difference between the two is the learning rule of error variance [35].

The MMV model is given by

$$Z = \Psi P + E \tag{3}$$

where $Z \in R^{d\times L}$ is a measurement matrix consisting of L measurement vectors, E is error matrix, and $P \in R^{K\times L}$ is an unknown coefficient matrix which has only a few nonzero rows. $\Psi \in R^{d\times K}$ denotes mapping matrix. The key assumption in the T-MSBL algorithm is that each row P_i satisfies a parameterized Gaussian distribution:

$$p(P_i; \gamma_i, B_i) \sim N(0, \gamma_i B_i) \quad i = 1, 2, \ldots, N \tag{4}$$

where γ_i and B_i are hyperparameters. γ_i is a nonnegative scalars controlling the row sparsity of P. When $\gamma_i = 0$, the corresponding i-th row element P_i becomes 0. Most γ_i equal 0 in noiseless cases or tend to very small values in noisy cases because of the mechanism of automatic relevance determination. Generally those γ_i which tend to very small values are set to 0 using T-MSBL with a threshold. B_i is a positive definite matrix capturing the temporal correlation structure of P_i, which is adaptively learned from the sEMG data.

To conveniently derive T-MSBL, by letting $z = \mathrm{vec}(Z^{\mathrm{T}}) \in R^{dL\times 1}$, $G = \Psi \otimes \mathrm{I}_L$, $\rho = \mathrm{vec}(P^{\mathrm{T}}) \in R^{KL\times 1}$, $\varepsilon = \mathrm{vec}(E^{\mathrm{T}})$, the MMV model (see Eq. 3) is equivalently transformed to the single vector model (see Eq. 1), i.e., $z = G\rho + \varepsilon$. Assume each element in the noise vector ε has a Gaussian distribution $p(\varepsilon; \lambda) \sim N(0, \lambda \mathrm{I})$, where λ is variance. According to the Bayes rule, the posterior probability $p(\rho|z; \Gamma)$ is written as $p(\rho|z; \Gamma) = N(\mu, \Sigma)$ with the mean μ and the covariance matrix Σ given by

$$\mu = \frac{1}{\lambda}\Sigma G z \tag{5}$$

$$\Sigma = \left(\Sigma_0^{-1} + \frac{1}{\lambda}G^T G\right)^{-1} = \Sigma_0 - \Sigma_0 G^T\left(\lambda \mathbf{I} + G\Sigma_0 G^T\right)^{-1} G\Sigma_0 \tag{6}$$

where Γ is the set of all the hyperparameters $\{\lambda, \gamma_i, B_i\}$, Σ_0 is given by $\Sigma_0 = diag\{\gamma_i B_i, \ldots, \gamma_N B_N\}$. After the hyperparameters are estimated, the estimate of ρ, i.e., $\hat{\rho}$, is given by the mean of the posterior probability.

The T-MSBL algorithm can be summarized as follows.

Step 1 Initialize γ, let $\gamma = 1$, or a non-negative random vector.

Step 2 Calculate μ and Σ using Eqs. (5) and (6).

Step 3 Update hyperparameters using the expectation maximization (EM) rule.

$$\gamma_i \leftarrow \frac{1}{L}\rho_i B^{-1}\rho_i^T + (\Xi_\rho)_{ii}$$
$$\tilde{B} \leftarrow \sum_{i=1}^{N}\frac{\rho_i^T\rho_i}{\gamma_i} + \eta\mathbf{I}$$
$$\lambda \leftarrow \frac{1}{NL}\|z - G\rho\|_F^2 + \frac{\lambda}{N}Trace\left[GWG^T\left(\lambda\mathbf{I} + GWG^T\right)^{-1}\right] \quad (7)$$

where $\Xi_\rho = \left(W^{-1} + \frac{1}{\lambda}G^TG\right)^{-1}$, $\rho = WG^T(\lambda\mathbf{I} + GWG^T)^{-1}z$, $W = diag\{\gamma_i, \ldots, \gamma_N\}$, η is a positive scalar.

Step 4 Check for convergence. If convergence criterion is not satisfied, iterate Step 2 and Step 3. If it has converged, go to Step 5.

Step 5 Calculate μ using Eq. (5). Let $\hat{\rho} = \mu$.

Feature extraction based on the T-MSBL algorithm

According to "Sparse reconstruction modeling of sEMG" and "T-MSBL algorithm" sections, s denoted the raw sEMG signals generated by a certain movement, K denoted the length of each segment of signal and $K = 256$. To compress the sEMG signals segment by segment, a sparse binary matrix of the size 128×256 was used as the measurement matrix Θ. Note that the sparse binary sensing matrix has been widely used in CS-based compression because of its efficiency in storage and matrix–vector multiplication [36]. Each column of Θ contained only a few entries equal to 1 with random locations, while other entries were 0. Sparse representation provides a good approximation to the initial sEMG signal. The optimal approximation depends on the selection of the best dictionary. However, heavy computation for selecting the best dictionary becomes an obstacle in application. The sEMG signal is compressible in some wavelet bases. Therefore wavelet bases are chosen as dictionary matrixes to balance computation and practical application. The selection of wavelet bases depends on the specific signal. The shapes of some wavelets in Daubechies family are similar to single motor action unit potential (MUAP) [37]. We chose Daubechies (order of 6) wavelet as the dictionary matrix D to sparsely represent the sEMG signals in this research. The dictionary matrix D had the size of 256×256. The compressed data of each segment y was obtained using $y = \Theta s$. The T-MSBL algorithm calculated the estimate of ρ (i.e., $\hat{\rho}$ in Eq. (2) using y and A ($A = \Theta D$).

In order to further reduce feature dimensions and enhance classification speed, the statistical features of the decomposition coefficients were taken as feature vectors in this study. RMS of the decomposition coefficients was adopted as the input vector of classifier. The RMS value ω_{rc} value was calculated as follows.

$$\omega_{rc} = \sqrt{\frac{1}{M}\sum_{l=1}^{M}\rho^2(l)} \quad (8)$$

where ρ denoted the vector of sparse coefficients in each analysis window, M denoted the number of nonzero entries of ρ.

Experiments and discussion

The proposed feature of sparse representation coefficients (SRC) was used to recognize the six classes of movements. The recognition performance of SRC feature was compared with those of some common features. The common features in our work were: MAV, RMS, WL, ZC, SSC, IEMG, histogram of EMG (HEMG), WAMP, AR coefficients (ARC), median frequency (MDF), frequency-band energy ratio after wavelet decomposition (ERA), RMS extracted from the NMF separated signals (EMGNMF) [25] and RMS extracted from the ICA separated signals (EMGICA) [24]. All the extracted features were normalized. The recognition performance was quantified by RA which was defined as follows.

$$RA_i = \frac{R_i}{T_i} \times 100\ \% \quad i = 1, 2, \ldots, C \quad (9)$$

$$ARA_i = \frac{1}{C}\sum_{L=1}^{C} RA_i \quad (10)$$

where R_i denoted the number of correctly classified samples of the ith movement. T_i denoted total number of testing samples of the ith movement. C denoted the number of movements in our work. ARA_i denoted the average RA result of the ith movement.

Two experiments were performed to validate the feasible of the proposed feature extraction scheme. The first experiment related to the recognition performance comparison of different single features. The second involved the recognition performance comparison of different multi-feature sets.

Performance comparison of single EMG features

We realized movement recognition by SVM classifier in this experiment. The average RA results with different single features are shown in Fig. 2.

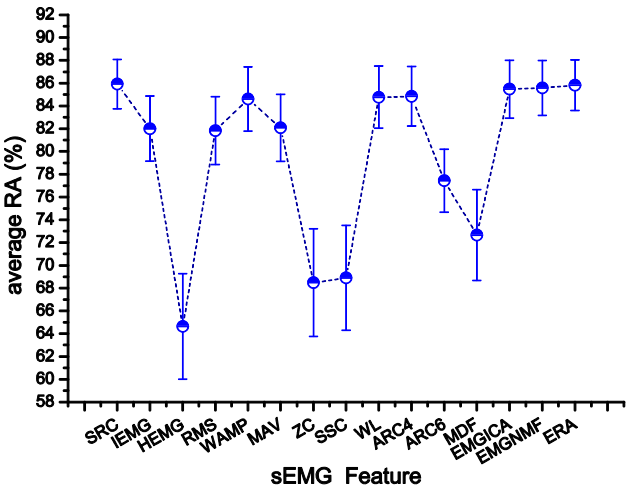

Fig. 2 The mean/std recognition accuracies of different single features

As shown in Fig. 2, the SRC has the highest average RA result of 85.92 % among all the single features, while the HEMG has the lowest one (64.65 %). The features of ERA, EMGNMF and EMGICA have better accuracy, which are 85.83, 85.58 and 85.48 %, respectively. The features of ARC4 (the fourth order of autoregressive model), WL and WAMP have comparable recognition performance, whose accuracies are over 84.50 %. The features of IEMG, ZC and SSC have relatively lower RA results. Most features in the time domain and frequency domain cannot represent the sEMG signals exactly because of the non-stationarity of sEMG signals. Hence their average RA results are unsatisfactory. The algorithms of wavelet, NMF and ICA can achieve sparseness of the solutions to some certain extent. However, the sparseness provided by the three analysis methods is rather weak. Hence the features of ERA, EMGNMF and EMGICA just offer the mean recognition accuracies of 85.83, 85.58 and 85.48 %, respectively. The SBL algorithm is equivalent to an iterative reweighted L1 minimization algorithm. Hence the global minima of SBL are always the sparsest solution [30]. Hence in contrast to wavelet, NMF and ICA, SBL can learn much sparser features. Compared to the other single features, the feature SRC provides more inherent information of sEMG. So the feature SRC gets the best RA result. It should be noted that the orthogonal wavelet base was adopted to analyze the sEMG signals because of the complexity of signals. The optimal sparse solutions over predefined orthogonal wavelet base might not be obtained. Hence the feature SRC has no huge advantage over other single features. We can consider the multi-features combining different single features to classify different movements. These will be discussed in the next section.

Some criteria, such as Davies–Bouldin index (DBI), entropy, Euclid-distance and Fisher's linear discriminate index (FLDI), can quantitatively assess the separability of sEMG features [38]. The FLDI represents clusters' dispersion comparing to their scatter [39]. Note that a larger FLDI value means a higher possibility of linearly discriminating the clusters in the feature space. Figure 3 shows the FLDI values of different single features. The trend of the FLDI values demonstrates that the separability of SRC is superior to those of other single features.

Figure 4 displays the confusion matrix for six movements using the feature SRC and SVM classifier. According to the results in Fig. 4, the movements of gross hand and individual finger, such as EXPM, FLTF, EXTF and EXIF, are easily recognizable. Their average RA results are over 85 %. The movements with more fingers may cause confusion and be misidentified, especially the multi-finger movements which are similar to the other ones. For example, HOOK and OKAY have lower RA results among all movements. The movement HOOK is misclassified as the movement EXIF. The movement OKAY is misclassified as the movement EXPM. The possible reasons for misidentification include measurement conditions, physiological structures of forearm muscles, etc. Regarding the results in Fig. 4, it can be due to the similarity between two movements. For example, extensor indicis proprius (EIP) served a similar contractile function when we performed the movement HOOK and EXIF, respectively.

Comparison of recognition performance among multi-feature sets

The multi-motion recognition with a single feature makes it difficult to obtain a satisfactory RA result. Studies have

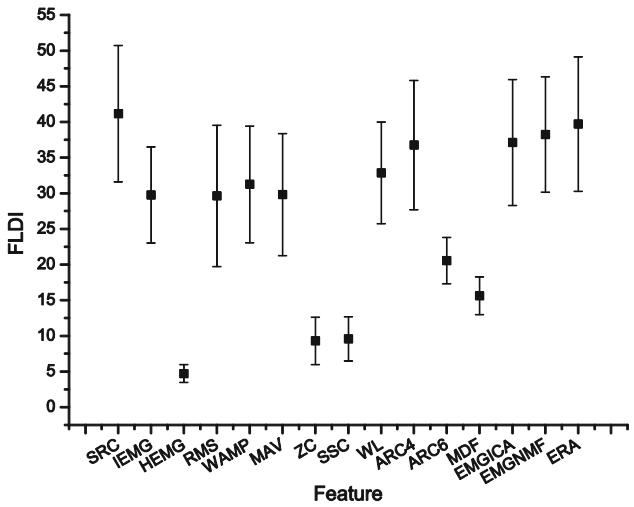

Fig. 3 The FLDI values of different single features

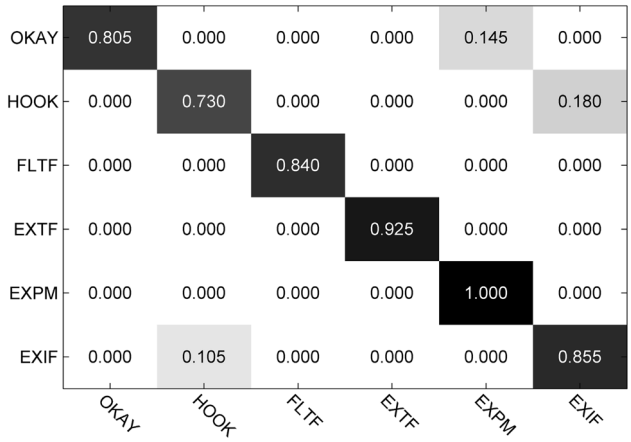

Fig. 4 Confusion matrix for six movements using the feature SRC and SVM classifier

proved that the recognition performances of multi-feature sets constituted by single features are much better than those of the single features [40, 41] for the multi-motion recognition using sEMG. If the single feature constituting a multi-feature set has no correlation to the others, i.e., they are all independent of each other and they have no effect on each other, the RA result with the multi-feature set can be improved greatly. The multi-feature sets constituted by the single features in "Performance comparison of single EMG features" section were employed in the second experiment. We classified the six movements with these multi-feature sets and SVM classifier. The RA results are shown in Fig. 5. The possible number of multi-feature sets was so large that we could not show all combinations here. The multi-feature sets whose RA results were good were shown in Fig. 5. The multi-feature sets with poor RA results were not shown.

According to the results in Fig. 5, the multi-feature sets obtain more satisfying and more stable performance than the single features do. The average RA results of multi-feature sets, such as MAV + WL + SRC, ARC4 + SRC, WAMP + SRC and SRC + WL, are more than 88 %. The multi-feature set SRC + WL has the best RA result, 91.67 %, among all multi-feature sets.

Figure 6 displays the FLDI values of different multi-feature sets. The trend of the FLDI values demonstrates that the separability of SRC + WL is superior to those of the other multi-feature sets.

Figure 7 presents the confusion matrix for the six movements using the multi-feature set SRC + WL and SVM classifier. The results indicate that the multi-feature set SRC + WL with better class separability achieves higher recognition performance in movement classification. The average RA results of six movements using the multi-feature set SRC + WL are improved greatly when compared with the results shown in Fig. 4. The movements of EXTF and EXPM have the average RA results of greater than 95 %. The average RA results of EXIF and FLTF are over 92 %. The average RA result of OKAY increases from 80.50 to 85.50 % and the average RA result of HOOK increases from 73.00 to 81.00 %. The results demonstrate that the combined features have more superior performance when compared with the single features.

The combinations of multiple features in [40, 41] and the above experiment in "Comparison of recognition performance among multi-feature sets" section are some simple concatenations of features, which assume each feature type contributes equally to the classification and may not give the best performance. To access this problem, we should consider about more flexible combinations, such as multiple kernel learning (MKL) framework [42] that

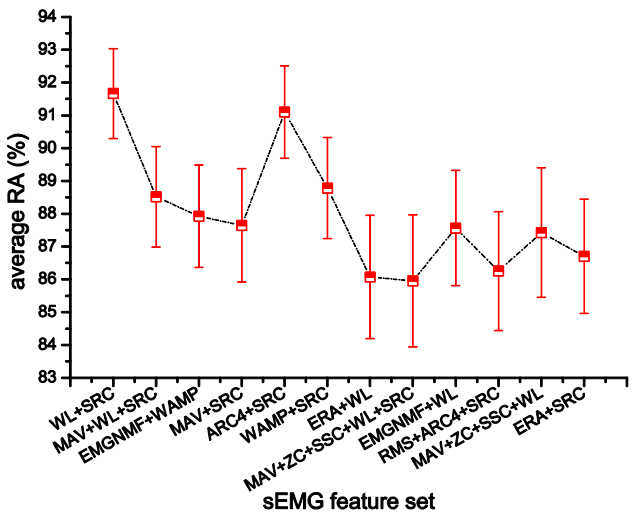

Fig. 5 The mean/std recognition accuracies of different feature sets

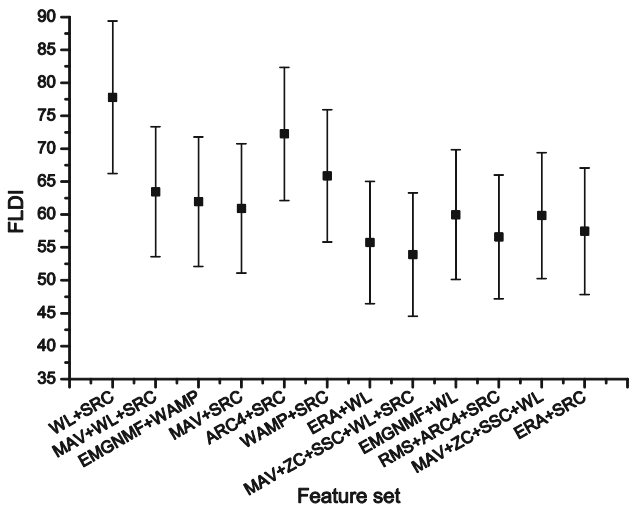

Fig. 6 The FLDI values of different multi-feature sets

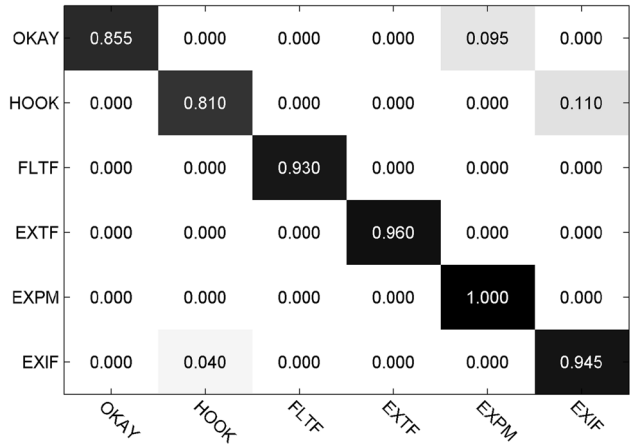

Fig. 7 Confusion matrix for six movements using the multi-feature set SRC + WL and SVM classifier

adaptively combines features with the optimal combination coefficients automatically learned. In this study, we employed simpleMKL algorithm [43] to combine different types of sEMG features for more effective classification. SimpleMKL is an efficient optimization algorithm to solve the MKL problem. The algorithm estimates the optimal coefficients for kernel combination using an iterative gradient descent algorithm and learns a weighted linear combination of kernels. For the detailed implementation of SimpleMKL, refers to Algorithm 1 in [43].

Linear, polynomial, Gaussian radial basis function (RBF) and sigmoid kernels are the common and major kernel functions in the MKL framework. To choose the appropriate kernel function for each feature type mentioned above, the linear, polynomial, RBF and sigmoid kernel functions were used to evaluate the classification performance, respectively. Table 1 lists the mean classification accuracies of different single features based on each single kernel function across the six movements.

From Table 1, it shows that in the classification of six movements, SRC, WAMP, ARC4, ARC6, MDF, ERA, EMGNMF, and EMGICA are well-suited with a RBF kernel; IEMG, HEMG, RMS, MAV, ZC, SSC, and WL are well-suited with a polynomial kernel. So the Gaussian RBF kernel was adopted for the features of SRC, WAMP, ARC4, ARC6, MDF, ERA, EMGNMF, and EMGICA. The polynomial kernel was adopted for the other seven features. We integrated the 15 features mentioned in "Performance comparison of single EMG features" section with the polynomial kernel and the RBF kernel using the SimpleMKL algorithm. We realized movement recognition with SVM classifier and the combination of multiple features. The average RA reached 94.38 % when the kernel weights of SRC, IEMG, HEMG, RMS, WAMP, MAV, ZC, SSC, WL, ARC4, ARC6, MDF, EMGICA, EMGNMF, and

Table 1 Comparison of performance for different single features based on different single kernel functions

Feature	Mean classification (%)			
	Linear	Polynomial	RBF	Sigmoid
SRC	83.79	84.75	**85.92**	83.84
IEMG	80.56	**82.18**	82.01	80.61
HEMG	63.25	**64.86**	64.65	63.73
RMS	81.47	**82.05**	81.84	81.52
WAMP	83.47	84.11	**84.62**	83.25
MAV	81.73	**82.22**	82.08	81.79
ZC	67.75	**68.63**	68.49	67.59
SSC	66.81	**68.97**	68.91	66.96
WL	84.69	**84.85**	84.77	84.54
ARC4	80.69	81.82	**84.85**	81.99
ARC6	73.52	75.66	**77.45**	75.59
MDF	69.95	70.29	**72.67**	70.31
EMGICA	83.42	83.59	**85.48**	83.27
EMGNMF	83.33	83.64	**85.58**	84.71
ERA	83.53	84.91	**85.83**	83.57

The best RA result for each single feature under one of four kernel functions is highlighted in bold

ERA were 0.7068, 0.0000, 0.0000, 0.0004, 0.0110, 0.0000, 0.0000, 0.0000, 0.2700, 0.0099, 0.0003, 0.0000, 0.0004, 0.0005, and 0.0007, respectively.

Figure 8 shows the weight of each kind of features. From Fig. 8, the weights of RMS, ARC6, EMGICA, EMGNMF and ERA were very low (close to 0) in the final combined kernel. In order to save the time of feature extraction, the features with the very low weights were dropped from our consideration. So we used the SimpleMKL algorithm to combine the features of SRC, WAMP, WL and ARC4. The multi-feature set SRC + WAMP + WL + ARC4 had a good and satisfying

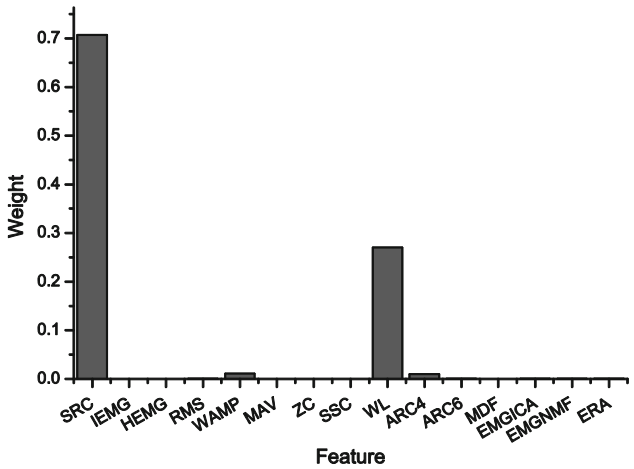

Fig. 8 Contributions (weights) of the sEMG features for movement classification

average RA, which was 94.33 %. The kernel weights of SRC, WAMP, WL and ARC4 were 0.7079, 0.0112, 0.2705, and 0.0104, respectively. The average RA of the multi-feature set SRC + WAMP + WL + ARC4 is much higher than the concatenations of features which had been studied in our previous experimental work. Compared with the multi-feature set integrated by the 15 features, the average RA of the multi-feature set SRC + WAMP + WL + ARC4 fell very rarely.

Figure 9 presents the confusion matrix for the six movements using the multi-feature set SRC + WAMP + WL + ARC4 and SVM classifier. From Fig. 9, we can see that the average RA results of six movements using the multi-feature set SRC + WAMP + WL + ARC4 are further improved when compared with the results shown in Fig. 7. The movements of EXTF and EXIF have the average RA results of greater than 97 %. The average RA of FLTF is over 95 %. The average RA result of OKAY increases from 85.50 to 89.00 % and the average RA result of HOOK increases from 81.00 to 85.00 %. The results demonstrate that the combined features via the MKL framework have more superior performance when compared with the concatenation of features. In the experiment, the FLDI value of the multi-feature set SRC + WAMP + WL + ARC4 was 89.54, which was greater than those of the other multi-feature sets. It demonstrates that this multi-feature set has the optimal class separability.

We had assessed the performance of the features including multi-feature sets and single features using SVM classifier. Now all of the sEMG features mentioned above were evaluated using different classifiers. According to the research results in [6], the effect of different classifiers on recognition performance is not significant when the extracted features are more distinguishable than other features. The following experimental results will demonstrate their conclusions in [6]. We focused more on the features with better performance. Here we chose the five multi-feature sets with the better RA to assess the recognition performance. They were WL + SRC, ARC4 + SRC, WAMP + SRC, MAV + WL + SRC, and SRC + WAMP + WL + ARC4. The multi-feature sets of WL + SRC, ARC4 + SRC, WAMP + SRC and MAV + WL + SRC were formed by simply concatenating different features together. The multi-feature set SRC + WAMP + WL + ARC4 was obtained by the SimpleMKL algorithm. Four common classifiers, k-nearest neighbor (k-NN), linear discriminant analysis (LDA), multi-layer perceptron (MLP) and SVM were employed in this research. The five multi-feature sets were formed first and then were applied as input to the classifiers when the classifiers of k-NN, LDA and MLP were employed to recognize the movements. The recognition results were shown in Fig. 10.

The results in Fig. 10 show that the multi-feature set SRC + WAMP + WL + ARC4 has the better accuracy for all four classifiers when compared with the other feature sets. The best RA result is gained by using the feature set SRC + WAMP + WL + ARC4 and SVM among all the features and classifiers. But the distinctions among the accuracies of four classifiers with the same multi-feature set are less sharp. Our result is consistent with the previous result in [6]. Hence it is testified that different classifiers have no significant effect on the RA results when more distinguishable features are extracted.

We used the analysis of variance (ANOVA) method to determine the effect of sEMG features and classifiers on the RA results. The Bonferroni correction was employed for multiple comparisons. The results are declared statistically significant if the significance level α is less than 0.05. The α value of the feature factor was 0 ($\alpha < 0.05$) in

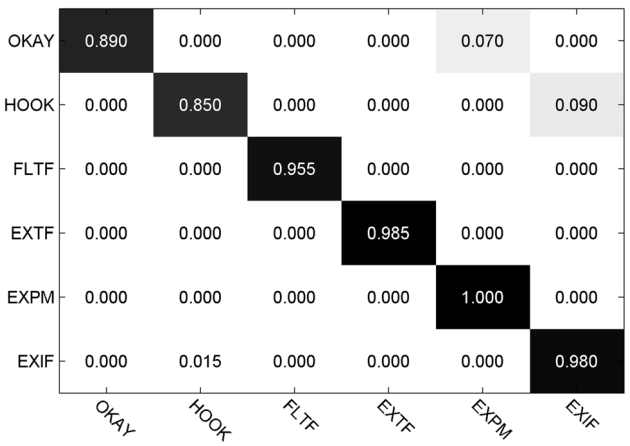

Fig. 9 Confusion matrix for six movements using the multi-feature set SRC + WAMP + WL + ARC4 and SVM classifier

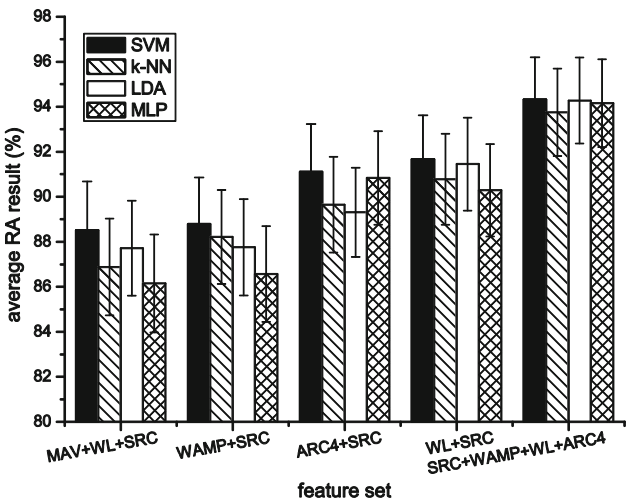

Fig. 10 The recognition accuracies of five multi-feature sets using different classifiers

this study. The α value of the classifier factor was 0.23 ($\alpha > 0.05$). The results indicate the features take significant effect on the RA results, while the effect of classifiers is not significant. These results are consistent with the previous reports in [40, 41], which have indicated the effect of features plays a more crucial role than that of classifiers in recognition performance. The multiple comparison analysis further indicates that the RA result of the multi-feature set SRC + WAMP + WL + ARC4 differs from those of the other features. It shows that the performance of SRC + WAMP + WL + ARC4 is superior to those of the other features.

We evaluated all the methods of sEMG feature extraction using MATLAB 2010a on a computer with Intel Core i5, 2.3 GHz and 2 GB memory. We reported the run time for each single feature from the multi-feature sets showed in Fig. 8. The running time for each single feature was 35 ms for SRC, 11 ms for ARC4 and less than 1 ms for MAV, WL and WAMP, respectively. The running time for the combination of multiple features via MKL was 5–6 ms. The classifiers ran within 2–4 ms for each feature.

Conclusions

This paper introduced and evaluated SBL approach to extract the features of sEMG signal and to discriminate different types of subtle movements including individual finger and multi-finger movements. The feature SRC was extracted by the T-MSBL algorithm. This feature effectively represented time-varying characteristics of sEMG during movements because of the compressibility of signal in the dictionary. The performance of SRC was compared with a variety of conventional sEMG features in terms of RA and FLDI value. The results demonstrated that the multi-feature set SRC + WAMP + WL + ARC4 fused using the SimpleMKL algorithm outperformed the multi-feature set formed by simply concatenating different features together and all single features. The multi-feature set SRC + WAMP + WL + ARC4 with optimal separability led to the improvement of recognition accuracy and the reduction of false recognition rate for multi-movement recognition. In particular when the similar movements were performed, the method was a powerful tool of feature extraction for movement recognition. We also found the standard deviation of SRC + WAMP + WL + ARC4 was relatively small. This suggested that SRC + WAMP + WL + ARC4 had more robust performance than other features in the movement recognition. Moreover, the statistical analysis confirmed that the performance of SRC + WAMP + WL + ARC4 was superior to those of other features, and feature factor had more significant effect on recognition performance than the classifier factor do. In conclusion, the proposed scheme provided satisfactory performance in terms of recognition results. In the future work, evaluating the proposed scheme on more complex hand movements in daily life will be the focus of our study. Besides, the proposed scheme of feature extraction will also be considered to test on prosthetic hands worn by the amputees.

Acknowledgments We thank S. L. Fang, C. Y. Zhang, C. C. Yu, Y. C. Liu, S. X. Cheng and T. Z. Xia for discussions and assistance with this work.

Compliance with ethical standards

Conflict of interest No conflict of interest.

References

1. Wheeler KA, Kumar DK, Shimada H (2010) An accurate bicep muscle model with sEMG and muscle force outputs. J Med Biol Eng 30:392–398
2. Khushaba RN, Kodagoda S, Liu D, Dissanayake G (2013) Muscle computer interfaces for driver distraction reduction. Comput Methods Programs Biomed 110:137–149
3. Geethanjali P, Ray KK (2011) Identification of motion from multi-channel EMG signals for control of prosthetic hand. Australas Phys Eng Sci Med 34:419–427
4. Pang M, Guo S, Song Z (2012) A surface EMG signals-based realtime continuous recognition for the upper limb multi-motion. In: IEEE international conference on mechatronics and automation (ICMA), pp 1984–1989
5. Kim KS, Choi HH, Moon CS, Mun CW (2011) Comparison of k-nearest neighbor, quadratic discriminant and linear discriminant analysis in classification of electromyogram signals based on the wrist-motion directions. Curr Appl Phys 11:740–745
6. Hargrove LJ, Englehart K, Hudgins B (2007) A comparison of surface and intramuscular myoelectric signal classification. IEEE Trans Biomed Eng 54:847–853
7. Nadzri AABA, Ahmad SA, Marhaban MH, Jaafar H (2014) Characterization of surface electromyography using time domain features for determining hand motion and stages of contraction. Australas Phys Eng Sci Med 37:133–137
8. Hudgins B, Parker P, Scott RN (1993) A new strategy for multifunction myoelectric control. IEEE Trans Biomed Eng 40:82–94
9. Chen X, Zhu X, Zhang D (2009) Use of the discriminant Fourier-derived cepstrum with feature-level post-processing for surface electromyographic signal classification. Physiol Meas 30:1399–1413
10. Xie HB, Huang H, Wu J, Liu L (2015) A comparative study of surface EMG classification by fuzzy relevance vector machine and fuzzy support vector machine. Physiol Meas 36:191
11. Rafiee J, Rafiee MA, Yavari F, Schoen MP (2011) Feature extraction of forearm EMG signals for prosthetics. Expert Syst Appl 38:4058–4067
12. Lorrain T, Jiang N, Farina D (2011) Influence of the training set on the accuracy of surface EMG classification in dynamic contractions for the control of multifunction prostheses. J NeuroEng Rehabil 8:25
13. Rubinstein R, Bruckstein AM, Elad M (2010) Dictionaries for sparse representation modeling. Proc IEEE 98:1045–1057
14. Wang J, Gao XZ, Guo P (2013) Feature extraction based on sparse representation with application to epileptic EEG classification. Int J Imaging Syst Technol 23:104–113

15. Tropp JA, Wright SJ (2010) Computational methods for sparse solution of linear inverse problems. Proc IEEE 98(6):948–958
16. Balouchestani M, Raahemifar K, Krishnan S (2013) New channel model for wireless body area network with compressed sensing theory. IET Wirel Sens Syst 3:85–92
17. Maoqi C, Ping Z (2015) A novel framework based on fastICA for high density surface EMG decomposition. IEEE Trans Neural Syst Rehabil, Eng
18. Naik GR, Dinesh KK, Marimuthu P (2014) Signal processing evaluation of myoelectric sensor placement in low-level gestures: sensitivity analysis using independent component analysis. Expert Syst Appl 31:91–99
19. Boudjellal A, Abed-Meraim K, Aissa-El-Bey A, Belouchrani A (2014) Sparsity-based algorithms for blind separation of convolutive mixtures with application to EMG signals. In: 2014 IEEE workshop on statistical signal processing (SSP)
20. Naik GR, Dinesh KK (2012) Subtle electromyographic pattern recognition for finger movements: a pilot study using BSS techniques. J Mech Med Biol 12:1250078
21. Shokrollahi M, Sridhar K (2012) Sleep EMG analysis using sparse signal representation and classification. In: Engineering in Medicine and Biology Society (EMBC), 2012 annual international conference of the IEEE, pp 3480–3483
22. Salman A, Allstot EG, Chen AY, Dixon AM, Gangopadhyay D, Allstot DJ (2011) Compressive sampling of EMG bio-signals. In: Circuits and systems (ISCAS), 2011 IEEE international symposium on, pp 2095–2098
23. Xie HB, Huang H, Wu J, Liu L (2015) A comparative study of surface EMG classification by fuzzy relevance vector machine and fuzzy support vector machine. Physiol Meas 36:191
24. Naik GR, Acharyya A, Nguyen HT (2014) Classification of finger extension and flexion of EMG and Cyberglove data with modified ICA weight matrix. In: Engineering in Medicine and Biology Society (EMBC), 2014 36th annual international conference of the IEEE, pp 3829–3832
25. Naik GR, Nguyen HT (2015) Nonnegative matrix factorization for the identification of EMG finger movements: evaluation using matrix analysis. IEEE J Biomed Health Inf 19:478–485
26. Vigário R, Oja E (2008) BSS and ICA in neuroinformatics: from current practices to open challenges. IEEE Rev Biomed Eng 1:50–61
27. Kameoka H, Ono N, Kashino K, Sagayama S (2009) Complex NMF: a new sparse representation for acoustic signals. In: Acoustics, speech and signal processing, ICASSP 2009. IEEE international conference on pp 3437–3440
28. Tipping ME (2001) Sparse Bayesian learning and the relevance vector machine. J Mach Learn Res 1:211–244
29. Wipf D, Nagarajan S (2010) Iterative reweighted and methods for finding sparse solutions. IEEE J Sel Top Signal Process 4:317–329
30. Zhang Z, Rao B (2011) Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning. IEEE J Sel Top Signal Process 5:912–926
31. Lu ZN, Zeng QX, Li CY, Yu SZ (2002) Practical electromyography. People's Medical Publishing House, Beijing
32. Yang D, Zhao J, Jiang L, Liu H (2012) Dynamic hand motion recognition based on transient and steady-state EMG signals. Int J Humanoid Rob 9:1250007
33. Eldar YC, Kutyniok G (eds) (2012) Compressed sensing: theory and applications. Cambridge University Press, Oxford city
34. Wipf DP, Rao BD (2004) Sparse Bayesian learning for basis selection. IEEE Trans Signal Process 52:2153–2164
35. Zhang Z, Rao BD (2011) Clarify some issues on the sparse Bayesian learning for sparse signal recovery University of California, San Diego, Tech. Rep
36. Liu B, Zhang Z, Xu G, Fan H, Fu Q (2014) Energy efficient telemonitoring of physiological signals via compressed sensing: a fast algorithm and power consumption evaluation. Biomed Signal Process Control 11:80–88
37. Sobahi NM (2011) Denoising of EMG signals based on Wavelet transform. Asian Trans Eng 1:17–23
38. Phinyomark A, Hirunviriya S, Nuidod A, Phukpattaranont P, Limsakul C (2011) Evaluation of EMG feature extraction for movement control of upper limb prostheses based on class separation index. In: International conference on biomedical engineering, pp 750–754
39. Oskoei MA, Hu H, Gan JQ (2013) Feature-channel subset selection for optimising myoelectric human-machine interface design. Int J Biomech Biomed Robot 2:195–208
40. Phinyomark A, Quaine F, Charbonnier S, Serviere C, Tarpin-Bernard F, Laurillau Y (2013) EMG feature evaluation for improving myoelectric pattern recognition robustness. Expert Syst Appl 40:4832–4840
41. Ibrahimy MI, Ahsan MR, Khalifa OO (2013) Design and performance analysis of artificial neural network for hand motion detection from EMG signals. World Appl Sci J 23:751–758
42. Althloothi S, Mahoor MH, Zhang X, Voyles RM (2014) Human activity recognition using multi-features and multiple kernel learning. Pattern Recognit 47:1800–1812
43. Rakotomamonjy A, Bach FR, Canu S, Grandvalet Y (2008) SimpleMKL. J Mach Learn Res 9:2491–2521