Available online at www.sciencedirect.com
ScienceDirect
Biomedical Signal Processing and Control 3 (2008) 169–174
Biomedical Signal Processing and Control
www.elsevier.com/locate/bspc

Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization

Marie-Françoise Lucas [a], Adrien Gaufriau [a], Sylvain Pascual [a], Christian Doncarli [a], Dario Farina [b,*]

[a] Institut de Recherche en Communication et Cybernétique de Nantes (IRCCyN), Nantes Cedex, France
[b] Center for Sensory-Motor Interaction (SMI), Department of Health Science and Technology, Aalborg University, Fredrik Bajers Vej 7 D-3, DK-9220 Aalborg, Denmark

Received 22 March 2007; received in revised form 19 September 2007; accepted 26 September 2007
Available online 7 November 2007

Abstract

The study proposes a method for supervised classification of multi-channel surface electromyographic signals with the aim of controlling myoelectric prostheses. The representation space is based on the discrete wavelet transform (DWT) of each recorded EMG signal using unconstrained parameterization of the mother wavelet. The classification is performed with a support vector machine (SVM) approach in a multi-channel representation space. The mother wavelet is optimized with the criterion of minimum classification error, as estimated from the learning signal set. The method was applied to the classification of six hand movements with recording of the surface EMG from eight locations over the forearm. Misclassification rate in six subjects using the eight channels was (mean $\pm$ S.D.) 4.7 $\pm$ 3.7% with the proposed approach while it was 11.1 $\pm$ 10.0% without wavelet optimization (Daubechies wavelet). The DWT and SVM can be implemented with fast algorithms, thus, the method is suitable for real-time implementation.

Keywords: Multi-resolution analysis; Wavelet design; Multi-channel signal classification; Support vector machine; Electromyography; Myoelectric prostheses

1. Introduction

Surface electromyographic (EMG) signals are detected over the skin surface and are generated by the electrical activity of the muscle fibers during contraction. Since each movement corresponds to a specific pattern of activation of several muscles, multi-channel EMG recordings, performed with electrodes placed on the involved muscles, can be used to identify the movement. This concept has been applied for the development of myoelectric prostheses [1], where the control of the prostheses is obtained by classification of EMG signals. For this purpose, various pattern recognition schemes, consisting of feature extraction and classification, have been applied [2].

Surface EMG signals are the summation of motor unit action potential trains [3]. Motor unit action potentials are compact support waveforms which repeat with similar shape over time. Thus, classical representation techniques based on autoregressive modeling or time-frequency features (decomposition on time-windowed sine bases) may not be suited to this type of signals. For this reason, wavelets have been previously proposed as basis functions to project EMG signals for feature extraction and subsequent classification [4]. In the present study, the discrete dyadic wavelet transformation (DWT) will be applied with optimization of the corresponding representation space by signal-based selection of the mother wavelet.

The DWT projects a signal into a set of basis functions that are scaled and delayed versions of a prototype function, called mother wavelet. The mother wavelet determines the projection space, thus, infinite feature spaces can be obtained from the DWT with varying mother wavelet. The feature space may be chosen a priori using a particular standard mother wavelet (e.g., Daubechies), or selected based on optimal discrimination. It is expected that different mother wavelets would lead to different discrimination abilities between classes, thus, the need for signal-based wavelet selection. With the purpose of classification, a natural optimization criterion is the classification error

* Corresponding author. Tel.: +45 96358821; fax: +45 98154008.
E-mail address: df@hst.aau.dk (D. Farina).

1746-8094/$ – see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bspc.2007.09.002

estimated from a training set of signals and an efficient method for exploring a large family of wavelet functions is wavelet parameterization by imposing conditions on the scaling filter in a multi-resolution analysis (MRA) framework [6]. Signal-based optimization of DWT for pattern recognition has been previously proposed by our group for simulated EMG signals [7], where the focus was mainly on the feature space while a simple classification method (nearest representative) was applied in a single channel and bi-class context. In this study we propose a method based on the optimization of the representation space in combination with a robust classifier (support vector machine, SVM), in the case of multi-channel signals, in a multi-class context. Moreover, the method will be applied to experimental signals recorded during six movements of the hand. The main aims are (1) to prove that wavelet signal representation is a suitable space for control of myoelectric prostheses and (2) that selection of the mother wavelet has an impact in the discrimination ability of wavelet features, thus, the need for signal-based optimization.

2. Methods

In this section, we present the parameterization of the mother wavelet, the feature space used for the multi-channel signal, the multi-class decision rule, and the criterion that allows optimization of the mother wavelet.

2.1. Wavelet parameterization

The parameterization chosen for the mother wavelet, defined within the MRA framework [8–10], was previously used for signal classification [6], for biomedical signal compression [11,12], and for blind source separation [13]. It corresponds to a parameterization of the filter defining the wavelet, and is briefly recalled here.

In the MRA, the filters h and g are associated to the scaling function φ and wavelet ψ through the two-scale relations: $\varphi(t/2) = \sqrt{2}\sum_n h[n]\varphi(t-n)$ and $\psi(t/2) = \sqrt{2}\sum_n g[n]\varphi(t-n)$. In the case of orthogonal wavelets, the wavelet filter g can be deduced from h by $g[k] = (-1)^{1-k}h[1-k]$ and then the mother wavelet is completely defined by the scaling filter h coefficients. However, to generate an orthogonal MRA wavelet, h must satisfy some conditions. For a Finite Impulse Response (FIR) filter of length L_h, there are $L_h/2+1$ sufficient conditions to ensure the existence and orthogonality of the scaling function and wavelet [14,15]. Thus, $L_h/2-1$ degrees of freedom remain to design the filter h. The lattice parameterization described by Vaidyanathan and Hoang [16] offers the opportunity to design h with $L_h/2-1$ new free parameters, called parameter vector θ. If $L_h = 4$, the design parameter vector $\theta = [\alpha]$ is reduced to a scalar parameter:

$$
\begin{aligned}
i = 0,3 : h[i] &= \frac{1-\cos(\alpha) + (-1)^i \sin(\alpha)}{2\sqrt{2}} \\
i = 1,2 : h[i] &= \frac{1+\cos(\alpha) + (-1)^{i-1} \sin(\alpha)}{2\sqrt{2}}
\end{aligned}
\tag{1}
$$

If $L_h = 6$, we need a two-component design vector $\theta = [\alpha, \beta]$, and h is given by:

$$
\begin{aligned}
i = 0,1 : h[i] &= \frac{[(1+(-1)^i\cos\alpha + \sin\alpha)(1-(-1)^i\cos\beta - \sin\beta) + (-1)^i 2\sin\beta\cos\alpha)]}{(4\sqrt{2})} \\
i = 2,3 : h[i] &= \frac{[1+\cos(\alpha-\beta) + (-1)^i\sin(\alpha-\beta)]}{(2\sqrt{2})} \\
i = 4,5 : h[i] &= \frac{1}{\sqrt{2} - h(i-4) - h(i-2)}
\end{aligned}
\tag{2}
$$

For other values of L_h, expressions of h are given in Refs. [10,17].

2.2. Multi-channel feature space

We consider a multi-channel signal x composed of K channels x_k: $x = \{x_1, \ldots, x_k, \ldots, x_K\}$. Given a mother wavelet ψ, the DWT decomposes the channel x_k on the corresponding discrete wavelet basis, where all the wavelets are dyadic dilated and translated versions of ψ. It provides a set of coefficients $d_{x_k}(s,u) = \langle x_k(t), \psi_{s,u}(t)\rangle$ where $\psi_{s,u}(t) = 2^{-s/2}\psi(2^{-s}t-u)$. The N coefficients $d_{x_k}(s,u)$ of the decomposition of the discrete channel x_k of length N are computed by Mallat's algorithm using h and g [8].

The waveforms are supposed to occur at unknown instants, randomly distributed inside a class. In order to be insensitive to occurrence instants, we use the marginals of each level of the decomposition as the channel features. We define the marginals of the DWT as:

$$
m_{x_k}(s) = \sum_{u=0}^{N/2^s-1} |d_{x_k}(s,u)|, \quad s = 1, \ldots, S
$$

where S is the deepest level of the decomposition ($S = \lfloor \log_2 N \rfloor$). The features representing the channel x_k are the components of the vector $M_{x_k} = [m_{x_k}(1), \ldots, m_{x_k}(S)]$. The vector M_{x_k} allows the representation of the channel by the contributions of each dyadic scale.

The multi-channel representation space is defined by the set of the mono-channel feature spaces. In this space, the signal x is represented by:

$$
M_x = [M_{x_1}, \ldots, M_{x_k}, \ldots, M_{x_K}]
$$

2.3. Multi-class decision rule using SVM

The classification is based on the SVM approach, using a two-class SVM discrimination for each of the n_c classes [5]: each class ω_{+i} against the rest of the population ω_{-i}, $i = 1, \ldots, n_c$. After training, a discriminant function f_i is obtained for each two-class SVM separation problem. For a given signal x, the final decision is taken from the n_c discriminant functions f_i:

$$
\text{``}x \text{ belongs to } \omega_j\text{''} \text{ with } j = \underset{i=1,\ldots,n_c}{Arg\max}\ f_i(x)
\tag{3}
$$

We recall briefly in the following some principles of the SVM approach. The description follows Burges [18].

2.4. Linear SVM

Consider a two-class problem with a learning set of labeled data in the representation space $\{x_p, y_p\}$, $p = 1, \ldots, P$, with $y_p = +1$ if $x_p \in \omega_i$, $y_p = -1$ if $x_p \in \omega_{-i}$. Assume that a separating hyperplane separates the positive from the negative examples. If this hyperplane is parameterized by w and b, it satisfies $\langle x, w\rangle + b > 0$ if $y = +1$ and $\langle x, w\rangle + b < 0$ if $y = -1$. Let d_+ (d_-) be the shortest distance from a given separating hyperplane to the closest positive (negative) example, the margin is defined as $d_+ + d_-$.

For this separable case (no points on the bad side of the border or inside the margin), the support vector algorithm simply provides the separating hyperplane with largest margin, under the constraint $d_+ = d_-$. This corresponds to the hyperplane (w_0, b_0) which is solution of the following constrained quadratic optimization problem:

$$\begin{aligned}
&\max_{w,b} 2\frac{\langle w, x\rangle + b}{||w||} && (\text{with } x \text{ being on a margin hyperplane})\\
&\text{under the } P \text{ constraints}: && \frac{\langle x_p, w\rangle + b}{||w||} \geq d, \quad \text{for } y_p = +1\\
& && \frac{\langle x_p, w\rangle + b}{||w||} \leq -d, \quad \text{for } y_p = -1
\end{aligned}$$

Using the normalization $\langle w, x\rangle + b = 1$ when x is on a margin hyperplane, the problem can be written as:

$$\begin{aligned}
& && \min_{w,b} ||w||^2\\
&\text{under the } P \text{ constraints}: && \langle x_p, w\rangle + b \geq 1, \quad \text{for } y_p = +1\\
& && \langle x_p, w\rangle + b \leq -1, \quad \text{for } y_p = -1
\end{aligned}$$

The solution is obtained using a Lagrangian formulation of the problem. Under the Kuhn–Tucker conditions, the Lagrange multipliers are non-zero only when the constraints are saturated, i.e., for the points located on the margin hyperplanes when $y_p(\langle x_p, w\rangle + b) - 1 = 0$. The corresponding points x_p are called support vectors. The results of optimization are the parameters w_0, b_0, which only depend on scalar products of the training data $\langle x_m, x_p\rangle$. Then, for a given vector x, the discriminant function that is used in the multi-class rule of Eq. (5) is written as:

$$f_i(x) = \langle w_0, x\rangle + b_0 \tag{4}$$

In the non-separable case, the previous constrained problem has no solution and it is necessary to relax the constraints for allowing points either within the margin or in the wrong class. It is done by introducing positive slack variables ξ_p, $p = 1, \ldots, P$ in the constraints, so that the problem becomes:

$$\begin{aligned}
& && \min_{w,b} ||w||^2 + C\sum_p \xi_p\\
&\text{under the constraints}: && \langle x_p, w\rangle + b \geq 1 - \xi_p, \quad \text{for } y_p = +1\\
& && \langle x_p, w\rangle + b \leq -1 + \xi_p, \quad \text{for } y_p = -1\\
& && \xi_p \geq 0, \quad p = 1, \ldots, P
\end{aligned}$$

where C controls the penalty to errors. The optimal parameters still only depends on the scalar products $\langle x_m, x_p\rangle$.

2.5. Non-linear SVM

The above method is generalized by mapping the data from the original representation space to another (possibly of infinite dimension) Euclidian space H, in which the mapped data are linearly separable. This is possible through the "kernel trick". The data only appear through scalar products in the above formulations. Given a mapping function Φ, the previous results can be used replacing the scalar products $\langle x_m, x_p\rangle$ by $\langle \Phi(x_m), \Phi(x_p)\rangle$. If there exists a kernel function K such that $K\langle x_m, x_p\rangle = \langle \Phi(x_m), \Phi(x_p)\rangle$, it is even not necessary to use or know Φ. A Gaussian kernel $K\langle x_m, x_p\rangle = \exp(-||x_m - x_p||^2/2\sigma^2)$ verifies these properties and corresponds to an infinite H. In this paper, the discriminant functions are searched in the class of linear functions.

2.6. Wavelet optimization criterion

The criterion used to optimize the mother wavelet is an estimation of the probability of classification error. A classification error occurs each time a signal x is assigned to ω_i, while belonging to ω_j with $j \neq i$. Thus, for a given wavelet parameter θ, the corresponding probabilities of classification error are:

$$P_e^\theta(\omega_i) = \text{Prob}(x \quad \text{assigned} \quad \text{to} \quad \omega_i | x \in \omega_j, j \neq i), 1, \ldots, n_c \tag{5}$$

The overall probability of classification error is the average value of the n_c probabilities $P_e^\theta(\omega_i)$:

$$P_e^\theta = \frac{1}{n_c}\sum_{i=1}^{n_c} P_e^\theta(\omega_i) \tag{6}$$

A regularized estimation of this classification error probability is computed from the available data (i.e., the learning set), by using the following cross-validation procedure [19]: P_e^θ is estimated from the average of the misclassification rates computed on several subsets of the learning set which did not participate to the decision rule elaboration (see step 5 of the summary of wavelet optimization). We optimize θ by minimizing this criterion. In this paper, this optimization is done by an exhaustive search on a grid (sampling of the parameters and computation of the criterion on each node). In the following, the optimal parameter will be denoted $\hat{\theta}$. The definitive classification rule corresponding to the optimal wavelet is then derived from all the learning sets (see the summary of wavelet optimization).

2.7. Summary of wavelet optimization (parameter θ)

The optimization of the parameter θ is performed with the following steps:

(1) Select a value of the parameter θ;

(2) Decompose the multi-channel signals of the training set Ω via the DWT derived from $h(\theta)$;
(3) Compute the marginal vectors of these decompositions

$\{M_x = [M_{x_1}, \ldots, M_{x_k}, \ldots M_{x_K}]\}_{x \in \Omega}$;

(4) Divide Ω in L subsets: $\Omega^1, \ldots, \Omega^L$;
(5) For $l = 1$ to L (learning cross-validation; in the present study $L = 3$):
   - Compute the discriminant functions f_i, $i = 1, \ldots, n_c$ (Eq. (4)) on $\Omega - \Omega^l$
   - Classify all the elements of Ω^l using the decision rule of Eq. (3)
   - For each class ω_i, calculate $P_e^\theta(\omega_i)(l)$ (Eq. (5))
   - Calculate $P_e^\theta(l)$ (Eq. (6))
(6) The criterion is Criterion$(\theta) = \frac{1}{L}\sum_{l=1}^{L} P_e^\theta(l)$;
(7) Repeat steps 1–6 for each value of the parameter vector θ and choose $\hat{\theta}$ as the θ value (which defines the optimal wavelet) that minimizes *Criterion*(θ).

The learning of the decision rule with the optimal wavelet is performed with the following steps:

(1) Decompose the multi-channel signals of all the training set Ω via the DWT derived from $h(\hat{\theta})$;
(2) Compute the marginal vectors of these decompositions $\{M_x = [M_{x_1}, \ldots, M_{x_k}, \ldots, M_{x_K}]\}_{x \in \Omega}$;
(3) Compute the discriminant functions f_i, $i = 1, \ldots, n_c$ (Eq. (4)) on Ω. The decision rule is then defined by Eq. (3).

2.8. Performances evaluation

To assess the performance of the method, a new (external to the learning step) cross-validation is applied: the set of available labeled signals is divided in M subsets, $M - 1$ used for training (input of the previous procedure) and one for testing. The sets are permuted and the average error rate from the M test set classifications is used as a measure of performance. In the present study, $M = 3$.

2.9. Experimental EMG signals

Surface EMG signals were detected from the forearm muscles of six healthy, male subjects (age range 25–34 years). Eight bipolar electrode systems (Neuroline 72001-k, Medicotest, Denmark; 2-cm inter-electrode distance) were placed equi-spaced around the circumference of the forearm at a distance from the elbow of one third of the elbow–wrist distance. Six tasks were considered: (1) wrist flexion, (2) wrist extension, (3) hand supination, (4) hand pronation, (5) hand opening, and (6) hand closing. Each task had a duration of 4 s, with 1 s of absence of activity (pre-movement phase), approximately 0.5 s of movement from the initial to the final position, and maintenance of the final position for the remaining time interval. Each movement was repeated 40 times, with randomization of the task order, with 5-min rest every 60 trials.

The EMG signals were amplified (16-channel surface EMG amplifier, EMG-16, LISiN-OT Bioelettronica, Torino, Italy), filtered (3 dB bandwidth, 10–500 Hz), sampled at 2048 Hz, and converted to a numerical format using a 12 bit A/D converter.

For each subject, the available data consisted of 180 recordings per class, divided in 120 for learning, and the remaining 60 for performance evaluation within a three-fold cross-validation procedure (see performances evaluation section). The classification results, obtained using the optimal wavelet filter of length 4 were compared with those obtained with the worst wavelet (maximum criterion), with a standard wavelet of same length (Daubechies) and a spectral representation in which the same features were extracted with uniform subband decomposition (using the power spectrum estimate and same number of subbands as in the wavelet analysis).

3. Results

Fig. 1 shows examples of recorded signals. The misclassification rates computed from the test set with the described cross-validation procedure for each subject are reported in Table 1. The accuracy of classification with optimal wavelets is always equal or better than with the catalog wavelet. When performance is poor with the catalog wavelet (subjects 2, 4, and 6), it is substantially improved by optimization while the improvement is modest when classification accuracy is already good with catalog wavelet. The averaged misclassification rate was (mean ± S.D.) 4.7 ± 3.7% with the proposed approach (wavelet optimization by minimization of the criterion), 11.1 ± 10% using Daubechies wavelet defined by a filter of

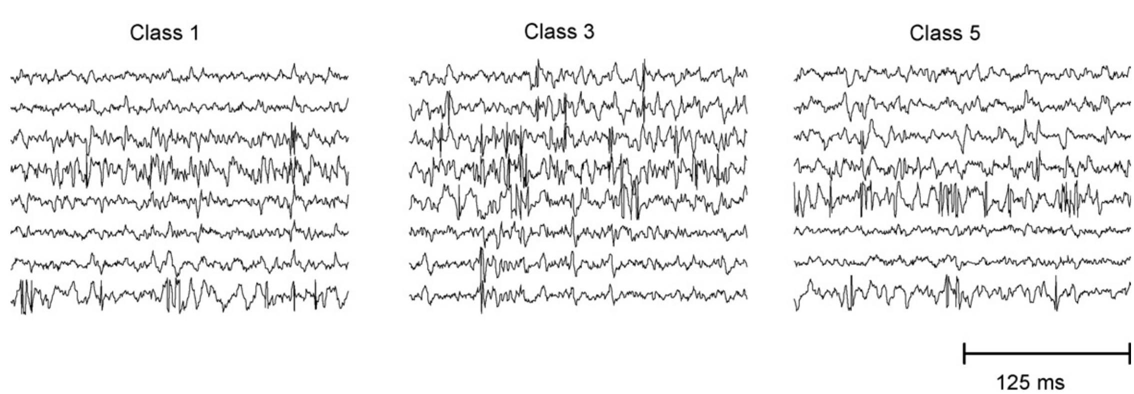

Fig. 1. Representative recordings (eight channels) from one subject in the last 250 ms of the movement (steady part) for three of the six classes considered. The vertical scale is in arbitrary units. Class 1 = wrist flexion; Class 3 = hand supination; Class 5 = hand opening.

Table 1
Misclassification rate (%) computed from the test set for each subject

	Daubechies	Optimal	Worst	Spectrum
Subject 1	1.0	0.46	1.3	7.2
Subject 2	13.2	8.9	16.6	20.4
Subject 3	5.2	4.1	6.0	44.0
Subject 4	16.9	10.5	20.0	25.0
Subject 5	1.0	0.9	13.2	7.1
Subject 6	29.0	3.4	41.0	6.8
Mean $\pm$ S.D.	11.1 $\pm$ 10.0	4.7 $\pm$ 3.7	16.4 $\pm$ 12.7	18.4 $\pm$ 13.5

Results are reported for Daubechies wavelet, optimal wavelet, worst wavelet and with uniform subband analysis (spectrum).

length 4, 16.4 $\pm$ 12.7% with the worst wavelet (selection of wavelet by maximization of the criterion), and 18.4 $\pm$ 14.8% with the spectral representation.

Fig. 2 reports the criterion for one subject. The criterion showed a large variability across wavelet functions for all subjects. The optimal wavelet was different for each subject (Fig. 3), thus, it was not possible to perform an a priori selection of the mother wavelet.

4. Discussion

Classification of multi-channel surface EMG signals was performed with an approach based on SVM applied on wavelet coefficients. The results obtained on five subjects have shown that the optimization of the feature space (mother wavelet) led to substantial improvement of performance with respect to the a priori selection of the mother wavelet.

4.1. Wavelet optimization

The influence of wavelet optimization was shown by comparison of the classification results obtained with a commonly used mother wavelet (Daubechies) and with the

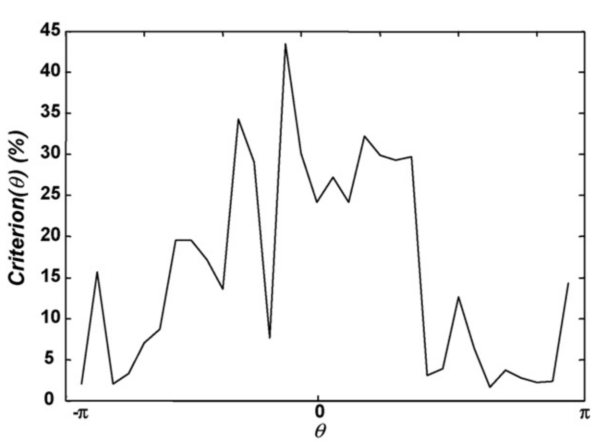

Fig. 2. Representative example of the criterion for selection of the optimal wavelet computed on the learning signal set for each tested wavelet. The length of the filter h is $L = 4$ in this example, thus, the design parameter vector θ is reduced to a scalar parameter. Each value of the scalar parameter θ corresponds to a filter h, according to Eqs. (1), and to a mother wavelet. The optimal wavelet is selected as that leading to the minimum criterion and is then applied to the test set that has not been used for the selection of the mother wavelet (see text for details).

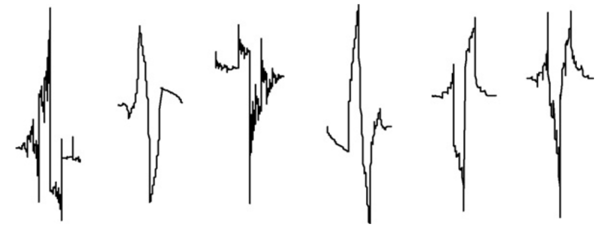

Fig. 3. The wavelets that minimized the criterion (probability of error estimated on the training set) for the six subjects.

wavelet leading to the worst criterion among those tested. The classification error was on average approximately 50% lower with optimized wavelet with respect to Daubechies wavelet. This approach does not need any a priori information and leads to simple implementation.

The improvement in performance by optimal wavelet selection confirms previous work on a simpler classification approach [6]. Moreover, the optimal wavelet was different among subjects (Fig. 3). It is not possible to predict the mother wavelet which allows best discrimination among the infinite possible, thus, the need for a supervised selection.

The proposed approach requires an exhaustive test of a pre-defined set of mother wavelets, obtained by parameterization, thus, the computational time is increased with respect to a priori selection of the transformation. However, only the learning step is affected by the increased computational time. It is expected that performance may be further improved by optimization of a different mother wavelet for each channel independently or by global optimization of the best combination of mother wavelets for the set of available channels. These approaches may, however, require iterative procedures that avoid the test of each combination of parameter values since this would not be feasible when the number of parameters to joint optimize increases. Best basis selection (wavelet packet) may also be a further improvement of the classification method. Finally, methods for reduction of the feature space dimension can be applied, such as principal component analysis. Despite these potential improvements, the method proposed in this paper allowed a misclassification rate for six-class discrimination of 5% using only 512 samples (250 ms) from the steady part of the task. The achieved performance makes the method suitable for practical applications.

4.2. Control of myoelectric prostheses

Various methods for extraction of features and classification have been proposed for discriminating movements from surface EMG signals with the purpose of controlling myoelectric prostheses (for review see Ref. [2]). The misclassification errors obtained with the proposed approach are in the same range as those recently reported by Huang et al. [20] and show high accuracy in the determination of the movement. The feature extraction and the SVM classification rule can be effectively implemented with fast algorithms (after training) for real-time applications. Moreover, the feature extraction is adapted to the subject through the learning step, thus, allowing optimization of performance for each subject.

5. Conclusion

A novel scheme for multi-channel supervised classification of surface EMG signals has been proposed. Application of this method to experimental EMG signals allowed discrimination of six hand movements with average misclassification rate of 5%.

References

[1] P.A. Parker, R.N. Scott, Myoelectric control of prostheses, Crit. Rev. Biomed. Eng. 13 (1986) 283–310.

[2] P.A. Parker, K. Englehart, B. Hudgins, The control of upper limb prostheses, in: R. Merletti, P.A. Parker (Eds.), Electromyography: Physiology, Engineering, and Non-invasive Applications, Wiley-IEEE Press, 2004.

[3] R. Merletti, P.A. Parker, Electromyography: Physiology, Engineering and Non-Invasive Applications, Wiley-IEEE Press, 2004.

[4] K. Englehart, B. Hudgins, P.A. Parker, M. Stevenson, Classification of the myoelectric signal using time-frequency based representations, Med. Eng. Phys. 21 (1999) 431–438.

[5] E.L. Allwein, R.E. Schapire, Y. Singer, Reducing multi-class to binary: a unifying approach for margin classifiers, in: Proceedings of the 2000 International Conference on Machine Learning, Stanford University, (2000), pp. 9–16.

[6] A. Maitrot, M.F. Lucas, C. Doncarli, Design of wavelets adapted to signals and application, in: IEEE International Conference an Acoustics, Speech and Signal Processing (ICASSP05), Philadelphia, USA, 2005.

[7] A. Maitrot, M.F. Lucas, C. Doncarli, D. Farina, Signal-dependent wavelets for electromyogram classification, Med. Biol. Eng. Comput. 43 (2005) 487–492.

[8] S. Mallat, A theory for multi-resolution signal decomposition: the wavelet representation, IEEE Trans. Pattern Anal. Machine Intell. 11 (1989) 674–693.

[9] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1998.

[10] C.S. Burrus, R.A. Gopinath, H. Guo, Introduction to Wavelets and Wavelet Transforms—A Primer, Prentice Hall, 1997, pp. 53–66.

[11] M. Nielsen, E.N. Kamavuako, M.M. Andersen, M.F. Lucas, D. Farina, Optimal wavelets for biomedical signal compression, Med. Biol. Eng. Comput. 44 (2006) 561–568.

[12] L. Brechet, M.F. Lucas, C. Doncarli, D. Farina, Compression of biomedical signals with mother wavelet optimization and best-basis wavelet packet selection, IEEE Trans. Biomed. Eng., in press.

[13] D. Farina, M.F. Lucas, C. Doncarli, Optimized wavelets for blind separation of non-stationary surface myoelectric signal, IEEE Trans. Biomed. Eng., in press.

[14] W. Lawton, Tight frames of compactly supported affine wavelets, J. Math. Phys. 31 (1990) 1898–1901.

[15] W. Lawton, Necessary and sufficient conditions for constructing orthonormal wavelet base, J. Math. Phys. 32 (1991) 1440–1443.

[16] P.P. Vaidyanathan, P.Q. Hoang, Lattice structures for optimal design and robust implementation of two-band perfect reconstruction QMF banks., IEEE Trans. Acoustic, Speech, Signal Proc. 36 (1988) 81–94.

[17] I.W. Selesnick, Maple and the parameterization of orthogonal wavelet bases. Online, October 1997, available: http://taco.poly.edu/selesi/theta2h/.

[18] C. Burges, A tutorial on support vector machines for pattern recognition, Data Mining Knowl. Dis. 2 (1998) 121–167.

[19] R.O. Duda, P.E. Hart, D.G. Stork, Pattern classification, second ed., John Wiley & Sons, Inc., New York, 2001.

[20] Y. Huang, K.B. Englehart, B. Hudgins, A.D. Chan, A Gaussian mixture model based classification scheme for myoelectric control of powered upper limb prostheses, IEEE Trans. Biomed. Eng. 52 (2005) 1801–1811.