Biomedical Signal Processing and Control 27 (2016) 24–31

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Distance and mutual information methods for EMG feature and channel subset selection for classification of hand movements

Haitham M. Al-Angari [a,*], Gunter Kanitz [b], Sergio Tarantino [b], Christian Cipriani [b]

[a] Department of Biomedical Engineering, Khalifa University, PO Box 127788, Abu Dhabi, United Arab Emirates
[b] The BioRobotics Institute, Scuola Superiore Sant'Anna, 56025 Pontedera, PI, Italy

ARTICLE INFO

Article history:
Received 3 June 2015
Received in revised form 8 December 2015
Accepted 25 January 2016
Available online 12 February 2016

Keywords:
Myoelectric control
Pattern recognition
Features selection
Mahalanobis distance
Mutual information

ABSTRACT

Different approaches have been proposed to select features and channels for pattern recognition classification of myoelectric upper-limb prostheses. The goal of this work is to use deterministic methods to select the feature-channels pairs that best classify the hand postures at different limb positions. Two selection methods were tried. One is a distance-based feature selection (DFSS) that determines a separability index using the Mahalanobis distance between classes. The second method is a correlation-based feature selection (CFSS) that measures the amount of mutual information between features and classes. To evaluate the performance of these selection methods, EMG data from 10 able-bodied subjects were acquired when performing 5 hand postures at 9 different arm positions and 10 time-domain and frequency-domain features were extracted. Classification accuracy using both methods was always higher than including all the features and channels and showed slight improvement over classification using the state-of-art TD features when evaluated against limb variation. The CFSS method always used less feature-channel pairs compared to the DFSS method. Using both methods, selection of channels placed on the posterior side of the forearm was significantly higher than anterior side. Such methods could be used as fast screening filters to select features and channels that best classify different hand postures at different arm positions.

1. Introduction

The search for physiologically appropriate, yet robust and reliable strategies for controlling motorized upper limb prostheses has been and is still today one of the main challenges in the field of rehabilitation engineering. Pattern recognition (PR) algorithms applied on the surface electromyogram (EMG) of the residual muscles have been investigated for predicting the individual's intent and controlling myoelectric prostheses since the late 1960s [1]. This technique is based on the premise that humans can voluntarily activate repeatable and distinct EMG signal patterns for different motor tasks and that these patterns can be recognized by appropriate algorithm and used to send appropriate commands to the prosthesis.

PR algorithms have demonstrated remarkable performance within ideal laboratory conditions [2–4]. However, their clinical viability is still unclear due to several factors that curtail their online performance in real-world scenarios. In particular, PR cannot adapt with the changes of EMG signals with time, like sweat, electrode displacement, contact impedance and fatigue [4] or with the changes of the physical setup like the variation of the limb position, the effects of weight and inertia [5–10]. The only example of clinically available PR myoelectric controller is represented by CoApt [11]. CoApt is a PR socket system and software program that serves the transhumeral and transradial amputation levels. The system consists of multiple sensors that are placed throughout a prosthetic socket. These sensors measure the strength and pattern of the muscle contractions when the user attempts the movements of the phantom arm. The patterns are stored in the PR system. The prosthesis is programed through a series of movements when the user first puts the socket on and can be reprogramed when the user feels like control is waning. To date CoApt assessed in a limited number of clinical cases in the US only.

In a conventional PR system a set of features describing the signal is extracted from each of an array of EMG channels, using time windows and is used to train and test a machine learning algorithm, associated with the relative motion class [12–14]. Better classification results and lower computational costs are achieved by reducing the size of the processed features, i.e., the dimensionality of the problem.

* Corresponding author. Tel.: +971 2 401 8186.
E-mail addresses: haitham.alangari@kustar.ac.ae, haitham.angari@gmail.com (H.M. Al-Angari).

http://dx.doi.org/10.1016/j.bspc.2016.01.011
1746-8094/

There is a general agreement that different machine learning algorithms like Artificial Neural Network (ANN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Multi-Layer Perception (MLP), perform similarly in terms of classification accuracy [15–18], however it is still an open question whether there are optimal or suboptimal features or feature sets, that decode the neural ensemble. Since Hudgins and colleagues proposed what became a widely used set of features, in 1993 [3], several new time-domain and frequency-domain features have been proposed (such as auto-regression coefficients, EMG histogram, Wilson Amplitude, Cepstrum coefficients, Wavelet decomposition coefficients, etc.). The performance of these features and of combination of features extracted from all the channels was assessed and compared [19–25] but very little work was done for searching the best features contained in each channel.

To our knowledge Khushaba and colleagues [26] and Oskoei et al. [27], were the only who investigated the importance of selecting individual features from different channels. They defined this search as combined feature/channel subset selection (FSS). In both studies evolutionary computational techniques were used as search strategies (particle swarm optimization by Khushaba et al.; multi-objective genetic algorithm by Oskoei et al.,) and classification rate and/or separability indexes as the objective functions. The selected subsets of features resulted in higher classification accuracies compared to the original sets of features, without the need of preliminary user specific adjustments. In other words FSS was proven as a significant aid to pattern recognition systems. For best channel selection, both works fixed the features subset first- i.e., the selection of best features and channels were not done simultaneously. Recently, Liu et al. have proposed two similar methods for feature and channel selection (using minimum Redundancy Maximum Relevance and the Markov random field). However, the work neither compared the features selected in best subset using the two methods nor the importance of electrode locations to improve the classification accuracy [28].

Based on these premises and on the promise hold, in this work we investigated whether FSS could represent a possible solution to address the so called position effect. The latter is the known degradation of the pattern recognition classification accuracy when the system is trained in one arm position and tested in another position. This degradation particularly affects intrinsic hand movements when the EMGs are collected from the residual forearm. The position effect, together with the above mentioned drawbacks severely curtail the clinical viability of physiologically appropriate pattern recognition systems (which use the muscles that would normally contribute to the function to be restored by the prosthesis) for multi-digit transradial prostheses commercially available. In this study we proposed two FSS deterministic methods, as opposed to evolutionary algorithms used in previous studies [26,27]. In particular we used a correlation-based method (CFSS) and a distance-based method (DFSS).

In the CFSS method, the selection of feature/channel (hereafter F/C) is based on the amount of mutual information they represent about the motion classes. This method was used in different fields including computer networking and gene expression studies [29–31]; recently we have proposed it for classification of EMG patterns during hand movements, due to its effective selection performance and computational efficiency, especially with large number of F/Cs [32,33]. In the DFSS method the quality of a F/C is characterized based on its ability to discriminate samples of one motion class from samples of other motion classes using some measures of inter-class distance. Different measures have been used to quantify the inter-class distance such as the Fisher ratio [34,35], the Mahalanobis distance [36] and Bhattacharyya distance [37]. In this work we further explored the work by Bunderson and Kuiken [38] by using the Mahalanobis distance.

Fig. 1. From left to right, the following hand postures: relax, pinch grasp, power grasp, lateral grasp and open hand. (1-column fitting image.)

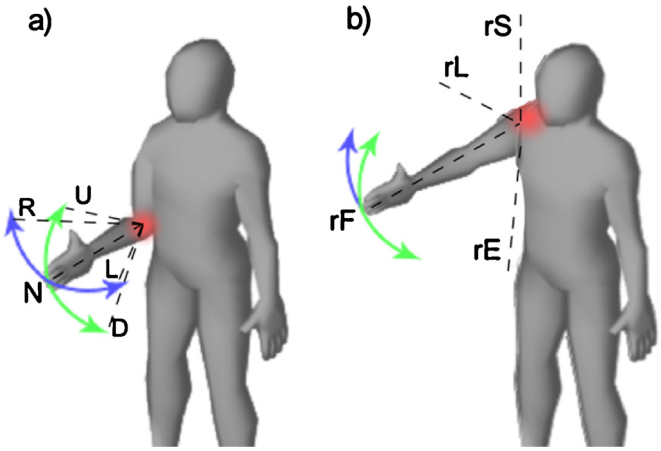

Fig. 2. Arm Positions: (a) "Normal" position and 45° variations (Right, Up, Left, and Down). (b) Reaching Sky, reach Lateral, reach Front and reach Earth. (1-column fitting image).

The goal of this work was to compare the classification performance, the importance of features and electrode locations of the CFSS and the DFSS methods in selecting the optimum set of features and channels to classify different hand postures at different arm positions. We enrolled 10 able-bodied subjects and we recorded EMG signals from their forearms during the execution of five hand grips/postures in nine arm positions for a total of 225 repetitions. Offline, we computed 10 state of art features including time-domain features, the Wavelet Transform, and the Sample entropy. The method evaluates the

F/Cs at different arm positions and selects those which are most robust against position variation which was not addressed in the previous works. Our results indeed show that FSS could be used for improving the performance of PR. Such proposed fast deterministic methods can be implemented as a subject-based approach for preselection of features and electrode placement to optimize the control myoelectric prostheses.

2. Methods

2.1. Experimental protocol

Ten able-bodied subjects were enrolled in this study. The subjects were asked to keep their right hand in one of the following five postures: *Relax*, *Pinch*, *Power*, *Lateral*, and *Open* for two seconds (Fig. 1). The hand postures were performed in nine different arm positions. Each posture was repeated 5 times following a randomized order of arm positions and hand postures. The nine arm positions were: Normal, Up, Right, Down, Left, Sky, Earth, Front, and Lateral (Fig. 2). The first five positions (Fig. 2a) were performed with the shoulder adducted while in the last four (Fig. 2b) the elbow was fully extended towards one of the four directions. In the Normal position, the elbow was bent for about 90° while the forearm was kept parallel to the horizontal plane. Instead, positions Up, Right, Down, and Left were reached by moving the forearm about 45° from the Normal position towards the proper direction. This range of

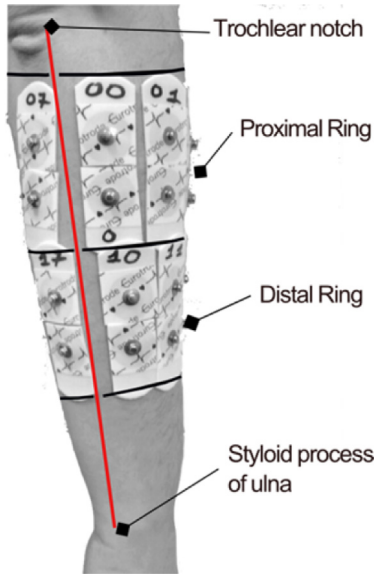

Fig. 3. Electrode placement around the forearm: (a) First electrode of each ring was placed to the right of the Ulna groove (red line). (b) Fifteen surface electrodes were placed in two rings (proximal and distal). (1-column fitting image). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

positions was chosen in order to cover much of the arm workspace. 15 pairs of electrodes (Ag–AgCl) were used to record the myoelectric activity from the forearm muscles. The electrodes were placed around the forearm to form two rings, one proximal to the elbow with the centres of electrode-pairs placed approximately 6 cm from the olecranon. The other ring was placed more distally with the centres of the electrode-pairs placed approximately 10 cm from the proximal ring (Fig. 3b). The proximal ring contained between 8 and 10 electrodes depending on the circumference of the subject's forearm and the remaining electrodes were placed on the distal ring. The first electrode-pair in each ring was placed directly right of the ulna groove (Fig. 3a). The raw EMG data were acquired using the EMG-USB2 system, (OT Bioelettronica, Turin, Italy) at a sampling frequency of 2048 Hz, band passed filtered (10–500 Hz), amplified with a gain of 1000, and stored for offline analysis in MatLab environment. A computer screen in front of the subjects provided instructions about the movements to perform by showing the arm position and the target hand posture, before the beginning of each repetition.

2.2. Data processing

The EMG signals were filtered using a 6th order BPF Butterworth filter (5–500 Hz) and were segmented using a 250 ms overlapping sliding window (225 ms overlap). Ten time-domain and frequency-domain features were extracted from each segment (Table 1). Details in computing these features are reported in Section 5. The FSS search, was made by evaluating and ranking the features and channels as pairs (F/C) (i.e., the performance of each feature from each channel was evaluated and the top ranked F/C were selected as the best subsets). This was done to have simultaneous selection of best features and channels and to speed up the selection process. We investigated a correlation-based method (CFSS) and a distance-based method (DFSS), described hereafter.

2.3. CFSS method

In the CFSS method, a correlation measure is applied to evaluate the effectiveness of feature subset on the hypothesis that a good feature subset is the one that contains features highly correlated with the class, yet uncorrelated with each other. The Pearson correlation coefficient is sensitive to linear relationship between two variables however it does not capture correlations that are not linear in nature [39]. To overcome this shortcoming, a correlation measure based on the information-theoretical concept of entropy (a measure of uncertainty of a random variable) is used.

Table 1
Time-domain and frequency-domain features with their number of coefficients per channel.

Feature	Abbreviation	No. of coefficients per channel
(1) Mean absolute value	MAV	1
(2) Standard deviation	SD	1
(3) Waveform length	WL	1
(4) Energy	EN	1
(5) Zero crossings	ZC	1
(6) Slope sign change	SSC	1
(7) Auto-regression	AR	6
(8) Wavelet decomposition	WDC	4
(9) Wavelet decomposition difference	WDCDIF	4
(10) Sample entropy	SamEn	1

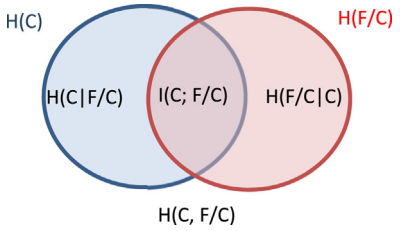

Fig. 4. A Venn diagram representing the relationship between entropy and mutual information of a F/C and the class C. Individual [$H(C)$, $H(F/C)$], join [$H(C, F/C)$] and conditional entropies for a pair of correlated subsystems C and F/C, with mutual information $I(C; F/C)$. (1-column fitting image).

For our application, the entropy of the class vector C was defined as:

$$H(C) = \sum_i p(c_i) \log(p(c_i)) \tag{1}$$

where $p(c_i)$ is the probability for an EMG segment to belong to class c_i and was estimated using the simple histogram approach: $p(c_i)$ = (number of EMG segments belongs to class i)/(Total number of segments). Similarly, a histogram was generated for F/C(n) and the entropy of C given the nth F/C was defined as:

$$H(C|\mathrm{F/C}(n)) = \sum_j p\left(\mathrm{F/C}(n)_j\right) \sum_i p\left(c_i|\mathrm{F/C}(n)_j\right) \log\left(p\left(c_i|\mathrm{F/C}(n)_j\right)\right) \tag{2}$$

where $p(c_i|\mathrm{F/C}(n)_j)$ is the probability for an EMG segment to belong to class c_i given that it belongs to the jth bin of F/C(n) histogram.

The mutual information between C and F/C(n), i.e., the amount of information gained about C after observing F/C(n), was computed as:

$$I(C; \mathrm{F/C}(n)) = H(C) - H(C|\mathrm{F/C}(n)) \tag{3}$$

and this relation can be presented with the Venn diagram shown in Fig. 4.

The symmetrical uncertainty (SU: the mutual information normalized over the sum of entropies) was then computed as

$$\mathrm{SU}(C; \mathrm{F/C}(n)) = \frac{2 \cdot I(C; \mathrm{F/C}(n))}{H(C) + H(\mathrm{F/C}(n))} \tag{4}$$

The SUs were sorted in descending order and since no prior information was available on the best subset size, the pairs with

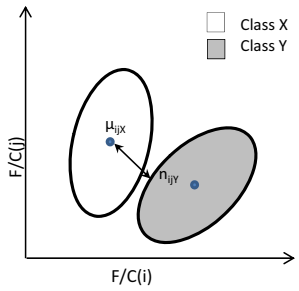

Fig. 5. The mahalanobis distance between the centre of the ellipse of class X (μ_{ijX}) and the nearset point from all classes (n_{ijY}) between the ith and jth pairs (F/C(i), F/C(j)) is used to compute the separability index (SI). (1-column fitting image).

top one-third SU were preselected in the best subset. Then, within the preselected set, the sum of the SU between any mth F/C and the other pairs: $\sum_k \mathrm{SU}(\mathrm{F/C}(m);\mathrm{F/C}(k))$ was computed (where $k=1$: size of the preselected set, and $k \neq m$). The mth F/C that does not satisfy the following inequality:

$$\frac{\sum_k \mathrm{SU}(\mathrm{F/C}(m);\mathrm{F/C}(k))}{\mathrm{SU}(C;\mathrm{F/C}(m))} > \frac{\mu_{\mathrm{F/C}}(m)}{\mu_{C;\mathrm{F/C}}(m)} \quad (5)$$

was considered redundant and removed from the final subset. In Eq. (5), $\mu_{C;\mathrm{F/C}(m)}$ and $\mu_{\mathrm{FC(m)}}$ were the means of $\mathrm{SU}(C;\mathrm{F/C}(m))$ and $\sum_k \mathrm{SU}(\mathrm{F/C}(m);\mathrm{F/C}(k))$, respectively. In order to find the optimum subset, the selected F/C pairs were added one-by-one as inputs to a SVM classifier (starting from the top one) until the increase of the overall accuracy was less than 0.1%. The corresponding F/C subset was then considered the optimum.

2.4. DFSS method

The Mahalanobis Distance is a measure of the distance between a point P and a distribution D. It is a multi-dimensional generalization of the idea of measuring how many standard deviations is P away from the mean of D. It is unit-less and scale-invariant, and takes into account the correlations of the data set.

To implement this method, the centroid (μ) and covariance (S) of the F/C space for each class were used to construct a hyper ellipsoid as shown in Fig. 5. We defined a separability index (SI) for the ith pair: F/C(i), as one-half the Mahalanobis distance between the centre of the ellipse of F/C(i) and the jth pair: F/C(j) for class X: μ_{ijX} and the nearest point of all the other classes n_{ijY}, averaged across all pairs and across all classes:

$$\mathrm{SI}(\mathrm{F/C}(i)) = \frac{1}{\mathrm{Nc}(\mathrm{Np}-1)} \sum_{X=1}^{\mathrm{Nc}} \sum_{j=1, j\neq i}^{\mathrm{Np}} \min_{Y=1:\mathrm{Nc}, Y\neq X} \frac{1}{2} \sqrt{\left(\mu_{ijX} - n_{ijY}\right)^T S_{ijX} \left(\mu_{ijX} - n_{ijY}\right)} \quad (6)$$

where $\mathbf{S_{ijX}}$ is the covariance of F/C(i) and F/C(j) at class X, Nc is the number of classes and Np is the number of F/C pairs. The SI indices were sorted in descending order. Also here, the F/C with top one-third SI were preselected and the best subset was determined as described for the CFSS method.

2.5. Repeatability of selection

For both methods, the optimum subset of the F/C pairs was searched from all of the five repetitions of all movement in all arm position. A repeatability index (RI) was introduced, in order to evaluate the generalization properties of the found subset. The RI was computed as the percentage of the F/C pairs in one repetition that was selected also in the other repetitions, averaged across all repetitions. In other words, the RI assessed the reliability of an optimum F/C subset across repetitions in all arm positions.

2.6. Performance of the FSS methods versus limb variation

The robustness of the two FSS methods against limb-position changes was assessed by evaluating the performance of the selected subset when training the classifier in one arm position and testing it in all the other positions. The classification accuracy using the individual features was computed for the sake of comparison with the optimized subsets. The classification performance was compared to the best single feature, all features (with all channels), and the state-of-art TD features (MAV, WL, ZC, SSC) suggested by Hudgins et al. [3].

2.7. Channel selection analysis

The goal of this analysis was to determine the muscle sites (electrode locations) that mostly contributed in encoding the hand grasps used in the experiment. There are three complex muscles in the forearm involved in hand motions through the wrist joint [40]. Since the exact location of the forearm muscles vary depending on the anatomy of the forearm, we divided the forearm muscles into four regions: proximal posterior, distal posterior, proximal anterior and distal anterior regions and the percentages of the selected channels from these regions were determined. This analysis was done using data from all arm positions and repetitions.

2.8. Classification and statistical analysis

All classifications were done using a state-of-art SVM classifier with a radial basis function (RBF) kernel [41]. For the purpose of comparison between the performances of the different selection methods, the penalty parameter C of the RBF kernel was set to equal 1. The Kolmogorov–Smirnov test was performed to check whether the results were normally distributed. Statistical significant differences were evaluated using the Kurskal–Wallis analysis of variance followed by a multi-comparison test with Bonferroni adjustment.

3. Results

The total number of F/C was 15×21 (as some features have more than one coefficient) = 315. The 2 s EMG signal with 250 ms window and 225 ms overlap made: (2000–225)/(250–225) = 71 segments. The number of segments for all 5 hand postures in all 9 arm positions and 5 repetitions was: $71 \times 5 \times 9 \times 5 = 15{,}975$ segments per F/C (per subject). Ranking of the F/Cs was done by evaluating these 15,975 segments according to the equations of the CFSS and DFSS methods and the performance of the best F/C subset was determined by computing its accuracy of classifying these 15,975 segments per subject.

3.1. Classification accuracy of the individual features

The performance achieved by the individual features provides the baseline for comparison with the performance achieved with the F/C subsets. For this analysis, the SVM classifier was trained in all arm positions with leave-one-repetition-out for testing. The MAV, SD, and WL allowed the highest accuracy; the accuracy of SamEn and WDCDIF were significantly lower than those of MAV and SD ($p < 0.05$) (Fig. 6).

3.2. Features subset selection

The histograms of the selected features using the CFSS and DFSS methods averaged across subjects are presented in Fig. 7. The CFSS

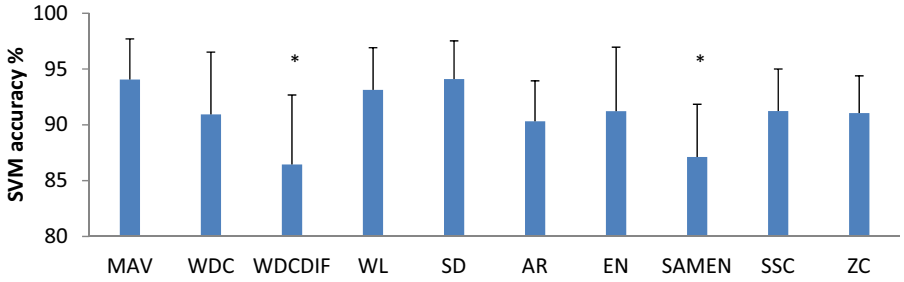

Fig. 6. SVM accuracy of the ten features trained and tested in all arm positions represented as mean ± SD across all subjects. (*: significantly lower than MAV and SD, $p < 0.05$). (2-column fitting image).

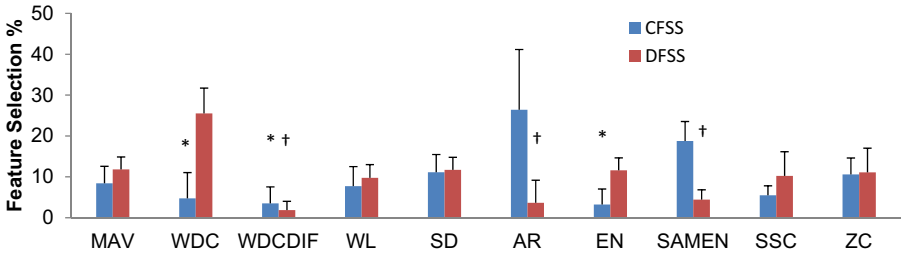

Fig. 7. Selection percentage of the ten features in the best F/C subset using the CFSS and SI methods (*: significantly lower than AR and SamEn, †: significantly lower than WDC, $p < 0.05$). (2-column fitting image).

Table 2
Number of F/C and RI using the two FSS methods for each subject.

Subject	CFSS		DFSS	
	# F/C	RI (%)	# F/C	RI (%)
SUB1	35	68	65	87
SUB2	40	72	60	86
SUB3	50	75	60	80
SUB4	40	79	50	86
SUB5	35	64	60	71
SUB6	40	68	65	71
SUB7	40	71	40	70
SUB8	30	71	60	73
SUB9	35	64	60	76
SUB10	35	85	45	76

approach turned out selecting the AR coefficients and SamEn. AR was the top selected feature for seven subjects and third highest selected feature for two subjects (not shown). SamEn was the top selected feature for 3 subjects and the second highest selected feature for 7 subjects (not shown). On the contrary, the WDC, WDCDIF, and EN were much less selected (significantly lower than AR and SamEn, $p < 0.05$). The WDC for example, was not selected in five subjects and with a percentage below 5% for two other subjects.

The DFSS method mostly selected the WDC (highest selected feature in all subjects—not shown). The AR, SamEn and WDCDIF were rarely or not selected at all (significantly lower than WDC, $p < 0.05$).

When the features as selected by the two methods were compared, AR and SamEn resulted selected by the CFSS significantly more than by the DFSS ($p < 0.01$, $p < 0.001$, respectively). The selection of WDC, EN, and MAV was significantly higher using DFSS ($p < 0.001$, $p < 0.001$, and $p < 0.05$, respectively).

3.3. Repeatability of selection

Table 2 shows the RI and the no. of *F, Ch* selected by the two selection methods. The RI of the DFSS (mean ± SD: 78 ± 6.8%) was slightly higher than the CFSS (mean ± SD: 72 ± 6.6%), across the subjects. The number of selected F/C pairs was lower for the CFSS (1st quartile–3rd quartile: 35–37.5) than for the DFSS (1st quartile–3rd quartile: 50–60).

3.4. Performance of the FSS methods versus limb variation

Fig. 8(a) compares the classification performance of the two selection methods, the MAV (the best individual feature), the state-of-art TD features: (MAV, WL, SSC, ZC), and all features with all channels (315 pairs in total) for each subject when training the classifier in one arm position and testing in all other positions. The classification accuracy using the subset achieved with the CFSS method was significantly higher than using the MAV for five subjects (SUB5,6,10: $p < 0.001$, SUB3,8: $p < 0.01$). It was significantly higher than using all features for four subjects (SUB4,5,8: $p < 0.05$, SUB10: $p < 0.01$) and significantly higher than using the TD features for two subjects (SUB5,10: $p < 0.05$).

The classification accuracy using the DFSS pairs was higher than the MAV for five subjects (SUB2,6,8: $p < 0.05$, SUB3,10: $p < 0.01$). It was significantly higher than using all features for three subjects (SUB2,5: $p < 0.05$, SUB10: $p < 0.01$) and significantly higher than using the TD features for two subjects (SUB3,5: $p < 0.05$). In one case i.e., one subject, the accuracy achieved using the CFSS, MAV, TD, and all features was significantly higher than when using the DFSS pairs (SUB7: $p < 0.01$). The performance of the classifier at each training position across all subjects is shown in Fig. 8(b). Both CFSS and DFSS methods showed higher accuracy compared to the MAV, all, and TD features at every training position. There was a noticeable drop in the performance at reach Sky and reach Earth training positions for all selection methods. At reach Sky position, the performance of DFSS was significantly higher than MAV ($p < 0.05$). At reach Earth position, the performance of CFSS, DFSS was significantly higher than MAV ($p < 0.05$) and all features ($p < 0.01$). At both reach Lateral and reach Front positions, all CFSS, DFSS, and TD was significantly higher than MAV ($p < 0.05$).

3.5. Channel selection analysis

The percentage of channels selected from the anterior and posterior regions and the two electrode rings using the CFSS and DFSS

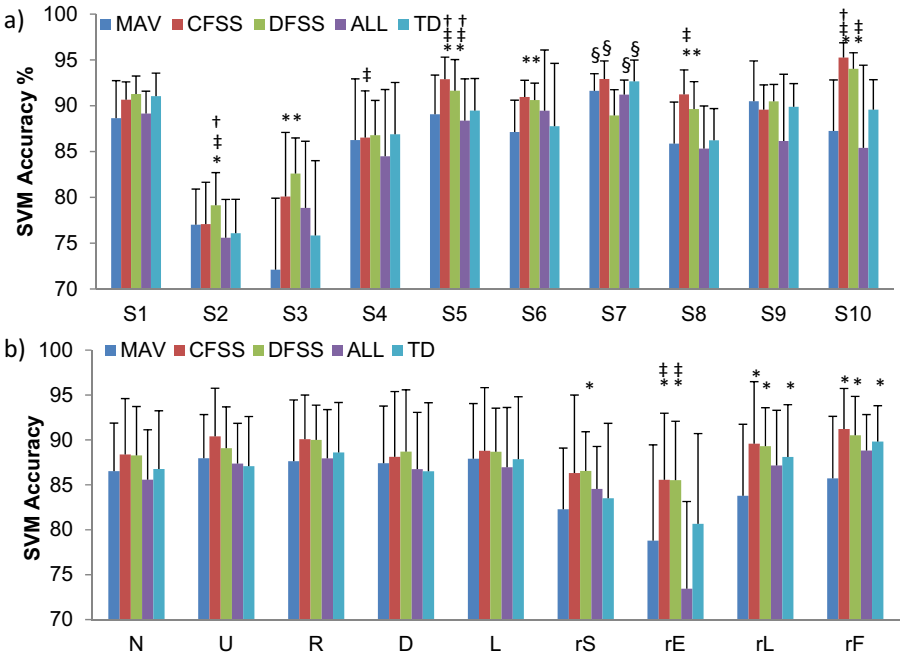

Fig. 8. Classification Performance of the MAV, CFSS, DFSS, ALL, and TD features with classifier trained in one position and tested in all other positions: (a) Performance per subject across all nine training (arm) positions, (b) Performance per training position across all subjects (N: Normal, U: Up, R: Right, D: Down, L: Left, rS: reach Sky, rE: reach Earch, rL: reach Lateral, rF: reach Front position). *: significantly higher than MAV, ‡: significantly higher than ALL pairs, †: significantly higher than TD features, §: significantly higher than DFSS pairs. (2-column fitting image).

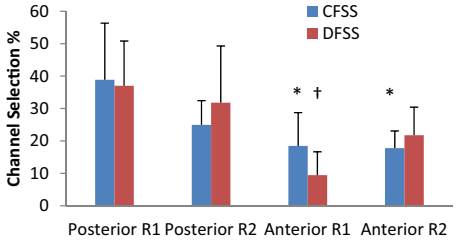

Fig. 9. Percentage of selected channel (mean ± SD) from Posterior and Anterior forearm electrode rings (R1: proximal ring, R2: distal ring) using the CFSS and DFSS methods (*: significantly lower than Posterior R1, $p < 0.05$, †: significantly lower than Posterior R1. (1-column fitting image).

filters is shown in Fig. 9. For most subjects and using both methods, the channels that contained most of the information needed for the investigated classification problem were those placed on the posterior side (8 subjects with CFSS, $p < 0.01$ and 9 subjects with DFSS, $p < 0.001$). Using CFSS, the selected anterior channels from both rings were significantly lower than the selected posterior channels from the proximal ring ($p < 0.05$). Using DFSS, the selected anterior channels from the proximal ring was significantly lower than the selected posterior channels from both rings ($p < 0.05$).

4. Discussion

This work presented two methods for selection of best pairs of features and channels one based on a measure of distance and the other one based on the amount of mutual information. The selection of the best subset was done by evaluating the features and channels as pairs (F/C). This was performed to have a simultaneous pick of best features and channels (i.e., which feature and from which channel) and to speed up the selection process. The quality of selection was tested against limb-effect and compared with other combinations of features: the MAV, the state-of-art TD features, and all features.

Regarding CFSS feature selection, the high percentage of the AR and SamEn, represents the high amount of information they contain about the class. The AR coefficients have been reported to be widely selected in the best feature subset when classifying hand postures in one arm position [27] and when evaluating the performance stability against electrode shift, muscle contraction effort, and fatigue [42]. SamEn, although poorly classified alone (Fig. 6), it showed importance when combined with the other features, thus indicating that it adds extra information about the class that is not provided by the other features. This is consistent with recent results that showed improvements in the classification accuracy off-line and virtual hand application when SamEn was added to MAV, variance, and AR features [22]. For on-line application however, the use of SamEn might introduce some delay due to the computational load and this needs to be examined experimentally [43]. Interestingly, these two features were rarely selected using the DFSS method. The high selection of the WDC features using the DFSS method is also consistent with its wide selection in best feature subset regardless of the type of mother wavelet function [27].

Both methods resulted with reasonably good selection of the TD features (i.e., MAV, WL, SD, ZC, and SSC); this confirms their effectiveness in classifying hand postures. The WDCDIF (although represented by 4 coefficients) was poorly selected in the best subsets in either method which suggests that it is not a good feature to be used for hand posture classification. In fact, we introduced this feature to evaluate whether a noticeable variation of the EMG energy while holding a certain grasp could be useful for classification of different hand postures, however the performance of this feature was the lowest among the others (Fig. 6).

One explanation of the significant variation in the selection percentage of the AR, SamEn, and WCD between the two methods could be due to the type of measurements. The DFSS is based on a linear measure of distance between different classes evaluated for every F/C, whereas the CFSS is a nonlinear measure of mutual information between each F/C and the classes. This suggests that the importance of (some) features among a combination of features depends on the FSS method and not on the features themselves. In

other words there might be multiple sub-optimal feature subsets that will differ based on the FSS used, providing basically the same overall performance. Further analysis on selection methods using different types of class separation measurement is needed to verify this observation.

The considerably high RI for both selection methods (Table 2) indicates that the majority of the selected F/C consistently had high ranks across repetitions. The number of pairs selected by the CFSS method was always lower than DFSS for all subjects, due to the removal of the redundant pairs from the preselected subset in the CFSS method. The number of repetition was selected to be 5. This was due the noticed fatigue on the subjects when trying higher number of repetitions especially with having the required arm position and hand grasp appearing randomly. Even with 5 repetitions and continuous recording of EMG, it took more than an hour to finish the whole experiment.

The improved performance of the selection methods over the state-of-art TD features, and all features when training in one position and testing in all other positions indicates that selected features and channels were the best to count for position effect. As more new EMG features are proposed, such methods could be useful for proper preselection of features and electrode placement for classification of different hand postures at different arm positions.

With regards to channel selection, the high selection of the electrodes placed on the posterior side of the forearm is probably due to a higher role those muscles play in performing the grasps investigated in this study. It could be interpreted as the posterior muscles are more discriminative during co-contraction compared to the anterior muscles. This opposes the findings of Oskoei et al. [27] where both sides of the forearm showed even importance and is most likely due to type of the hand posture used in this study that involved more finger flexion (pinch, power, and lateral) compared to the even number of flexion and extension movement in Oskoei's study. However, this is a general evaluation since the electrodes did not target specific muscles and further analysis would be needed to determine the important muscle associated with different hand postures to improve classification accuracy.

5. Appendix A

The ten features used in this work were:

(1) Mean absolute value (MAV)

$$\mathrm{MAV} = \frac{1}{N}\sum_{k=1}^{N}\left|x_k\right| \tag{A.1}$$

where x_k is the kth sample of the signal x and N is the number of samples in the time window.

(2) Standard deviation (SD)

$$\mathrm{SD} = \sqrt{\frac{1}{N}\sum_{k}^{N}(x_k-\mu)^2} \tag{A.2}$$

where μ is the mean of the time window

(3) Waveform length (WL)

$$\mathrm{WL} = \sum_{k=1}^{N}\left|\Delta x_k\right| \tag{A.3}$$

where $\Delta x_k = x_k - x_{k-1}$

(4) Signal energy (EN)

$$\mathrm{EN} = \sum_{k=1}^{N}\left|x_k\right|^2 \tag{A.4}$$

(5) Zero crossing

It is defined as the number of times the waveform crosses zero and is incremented if:

$$\mathrm{sgn}\left(-x_k \times x_{k+1}\right) \quad \text{and} \quad \left|x_k - x_{k+1}\right| \geq \text{threshold} \tag{A.5}$$

where sgn(x) = 1 if $x > 0$ and 0 otherwise

(6) Slope sign changesDefined as the number times the slope of the waveform changes sign and is incremented if:

$$\left(x_k - x_{k-1}\right) \times \left(x_k - x_{k+1}\right) \geq \text{threshold} \tag{A.6}$$

(7) Autoregressive coefficients (AR)This feature model the waveform as a linear autoregressive time series defined as:

$$x_k = \sum_{i=1}^{p} a_i x_{k-i} + e_k \tag{A.7}$$

where a_i represents the AR coefficients, p is the AR model order, and e_k is the residual white noise.

(8) Energy of a 4-level Wavelet Decomposition Transform (WDC) evaluated at each level of decomposition (with db-1 as the mother wavelet)

(9) Energy difference of a 4-level Wavelet Decomposition Transform (WDCDIF)

This feature was computed by splitting each segment into two halves and computing the difference in WDC between the two halves.

(10) Sample entropy (SamEn)

With a template size $m = 2$, and tolerance $r = 0.25 \times \mathrm{SD}$.

Acknowledgement

This work was supported by the European Commission under the WAY project (FP7-ICT-288551).

References

[1] F.R. Finley, R.W. Wirta, Myocoder studies of multiple myopotential response, Arch. Phys. Med. Rehabil. 11 (48) (1967) 598–601.
[2] D. Graupe, J. Salahi, K.H. Kohn, Multifunctional prosthesis and orthosis control via microcomputer identification of temporal pattern differences in single-site myoelectric signals, J. Biomed. Eng. 4 (1) (1982) 17–22.
[3] B. Hudgins, P. Parker, R.N. Scott, A new strategy for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 40 (1) (1993) 82–94.
[4] E. Scheme, K. Englehart, Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use, J. Rehabil. Res. Dev. 48 (6) (2011) 643–660.
[5] L. Hargrove, K. Englehart, B. Hudgins, The effect of electrode displacements on pattern recognition based myoelectric control, Conf. Proc. IEEE Eng. Med. Biol. Soc. (2006) 2203–2206.
[6] L. Hargrove, K. Englehart, B. Hudgins, A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control, Biomed. Signal Process. Control 3 (2008) 175–180.
[7] E. Scheme, K. Englehart, Training strategies for mitigating the effect of proportional control on classification in pattern recognition based myoelectric control, J. Prosthet. Orthot. 25 (2013) 76–83.
[8] A. Fougner, E. Scheme, A.D. Chan, K. Englehart, Ø. Stavdahl, Resolving the limb position effect in myoelectric pattern recognition, IEEE Trans. Neural Syst. Rehabil. Eng. 19 (2011) 644–651.
[9] C. Cipriani, M. Controzzi, G. Kanitz, R. Sassu, The effects of weight and inertia of the prosthesis on the sensitivity of electromyographic pattern recognition in relax state, J. Prosthet. Orthot. 24 (2012) 86–92.
[10] Y. Geng, P. Zhou, G. Li, Toward attenuating the impact of arm positions on electromyography pattern-recognition based motion classification in transradial amputees, J. Neuroeng. Rehabil. 9 (2012) 1–11 (74-0003-9-74).
[11] CoApt, 2014, Retrieved June 30, 2014, Available: ⟨http://www.coaptengineering.com/pattern-recognition.html⟩.
[12] M.A. Oskoei, H. Hu, Myoelectric control systems—a survey, Biomed. Signal Process. Control 2 (4) (2007) 275–294.
[13] K. Englehart, B. Hudgin, P.A. Parker, A wavelet-based continuous classification scheme for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 48 (3) (2001) 302–311.
[14] K. Englehart, B. Hudgins, A robust, real-time control scheme for multifunction myoelectric control, IEEE Trans. Biomed. Eng. 50 (7) (2003) 848–854.

[15] G. Kanitz, C. Antfolk, C. Cipriani, F. Sebelius, M.C. Carrozza, Decoding of individuated finger movements using surface EMG and input optimization applying a genetic algorithm, Conf. Proc. IEEE Eng. Med. Biol. Soc. (2011) 1608–1611.
[16] I. Kuzborskij, A. Gijsberts, B. Caputo, On the challenge of classifying 52 hand movements from surface electromyography, Conf. Proc. IEEE Eng. Med. Biol. Soc. (2012) 4931–4937.
[17] M. Ortiz-Catalan, R. Brånemark, B. Håkansson, BioPatRec: a modular research platform for the control of artificial limbs based on pattern recognition algorithms, Source Code Biol. Med. 8 (11) (2013), http://dx.doi.org/10.1186/1751-0473-8-11.
[18] S. Guo, M. Pang, B. Gao, H. Hirata, H. Ishihara, Comparison of sEMG-based feature extraction and motion classification methods for upper-limb movement, Sensors 15 (2015) 9022–9038.
[19] M. Zardoshti-Kermani, B.C. Wheeler, K. Badie, R.M. Hashemi, EMG feature evaluation for movement control of upper extremity prostheses, IEEE Trans. Neural Syst. Rehabil. Eng. 3 (1995) 324–333.
[20] S.H. Park, S.P. Lee, EMG pattern recognition based on artificial intelligence techniques, IEEE Trans. Rehabil. Eng. 6 (1998) 400–405.
[21] F. Chan, Y. Yang, F. Lam, Y. Zhang, P. Parker, Fuzzy EMG classification for prosthesis control, IEEE Trans. Rehabil. Eng. 8 (3) (2000) 305–311.
[22] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, Y. Laurillau, Emg feature evaluation for improving myoelectric pattern recognition robustness, Expert Syst. Appl. 40 (12) (2013) 4832–4840.
[23] Z. Ju, G. Ouyang, M. Wilamowska-Korsak, H. Liu, Surface emg based hand manipulation identification via nonlinear feature extraction and classification, IEEE Sens. J. 13 (9) (2013) 3302–3311.
[24] R.N. Khushaba, M. Takruri, J.V. Miro, S. Kodagoda, Toward limb position invariant myoelectric pattern recognition using time-dependent spectral features, Neural Netw. 55 (2014) 42–58.
[25] Y. Zhang, G. Wang, C. Teng, Z. Sun, J. Wang, The analysis of hand movement distinction based on relative frequency band energy method, BioMed Res. Int. 2014 (2014) 1–8.
[26] R.N. Khushaba, A. Al-Jumaily, Channel and feature selection in multifunction myoelectric control, Conf. Proc. IEEE Eng. Med. Biol. Soc. (2007) 5182–5185.
[27] M.A. Oskoei, H. Hu, J. Gan, Feature-channel subset selection for optimizing myoelectric human-machine interface design, Int. J. Biomechatron. Biomed. Rob. 2 (2013) 195–208.
[28] J. Liu, X. Li, G. Li, P. Zhou, EMG feature assessment for myoelectric pattern recognition and channel selection: A study with incomplete spinal cord injury, Med. Eng. Phys. 36 (2014) 975–980.
[29] X. Lu, X. Peng, P. Liu, Y. Deng, B. Feng, B. Liao, A novel feature selection method based on CFS in cancer recognition, in: Conf. Rec. IEEE Int. Conf. Systems Biology, 2012, pp. 226–231.
[30] T. Chou, K. Yen, J. Luo, N. Pissinou, K. Makki, Correlation-based feature selection for intrusion detection design, in: Conf. Rec. IEEE Conf. Military Communications, 2007, pp. 1–7.
[31] H.C. Peng, F. Long, C. Ding, Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy, IEEE Trans. Pattern Anal. Mach. Intell. 27 (2005) 1226–1238.
[32] M.A. Hall, L.A. Smith, Feature selection for machine learning: comparing a correlation-based filter approach to the wrapper, in: Proc. 12th Int. Florida AI Research Society Conf, Orlando, FL, 1999.
[33] H.M. Al-Angari, G. Kanitz, S. Tarantino, J. Rigosa, C. Cipriani, Feature and channel selection using correlation based method for hand posture classification in multiple arm positions, in: Proc. 2nd Int. Conf. Neuro Rehab, 2014, pp. 227–236.
[34] R.O. Duda, P.E. Hart, D.G. stork, Pattern Classification, second ed., John Wiley & sons, New York, 2001.
[35] H. Tang, H. Maitre, No. Boujemaa, W. Jiang, On the relevance of linear discriminative features, Inf. Sci. 180 (2010) 3422–3433.
[36] P.C. Mahalanobis, On the generalised distance in statistics, Proc. Natl. Inst. Sci. India 2 (1) (1936) 49–55.
[37] K. Fukunaga, Introduction to Statistical Pattern Recognition, second ed., Academic Press, San Diego, 1990.
[38] N.E. Bunderson, T.A. Kuiken, Quantification of feature space changes with experience during electromyogram pattern recognition control, IEEE Trans. Neural Syst. Rehabil. Eng. 20 (2012) 239–246.
[39] F.E. Croxton, E. Frederick, D.J. Cowden, S. Klein, Applied General Statistics, Pitman, London, UK, 1968.
[40] ExRx.net LLC, ACSM Resource Manual for Guidelines for Exercise Testing and Prescription, 2006, pp. 224–349, Available: ⟨http://www.exrx.net/Lists/ExList/ForeArmWt.html⟩ [online].
[41] S.S. Keerthi, C.J. Lin, Asymptotic behaviors of support vector machines with Gaussian kernel, Neural Comput. 15 (2003) 1667–1689.
[42] D. Tkach, H. Huang, T.A. Kuiken, Study of stability of time-domain features for electromyographic pattern recognition, J. NeuroEng. Rehabil. 7 (21) (2010), http://dx.doi.org/10.1186/1743-0003-7-21.
[43] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, Am. J. Physiol. Heart Circ. Physiol. 278 (2000) h2039–h2049.