COMBINED CLASSIFICATION AND REGRESSION FOR SIMULTANEOUS AND PROPORTIONAL EMG CONTROL OF WRIST FORCES

Mohammad Hossein Shahmoradi, Mohammad Ali Akhaee, Maryam S. Mirian

School of Electrical and Computer Engineering
College of Engineering, University of Tehran, Tehran, Iran
{mh.shahmoradi, akhaee, mmirian}@ut.ac.ir

ABSTRACT

In this study, a novel method for estimating wrist forces from surface electromyogram (EMG) measured from the upper limb is proposed, which can be applied for unilateral transradial amputees. Three degrees of freedom (DoFs) of wrist including flexion-extension, abduction-adduction, and pronation-supination were used. We first classify feature vectors extracted from the EMG signals into three classes namely positive output, negative output and dead zone output, using a multiple kernel learning (MKL) algorithm. Then for each DoF and each class, a neural network was trained to associate EMG features to their corresponding force outputs. We will show that this classification prior to regression plays an important role in increasing the performance of force estimation. The accuracy of estimation ranges from 90% to 94% (R^2 index) in 8 able-bodied subjects, which is proved to be significantly higher (p<0.05) than that of the previous works.

Index Terms— Electromyography (EMG), multiple kernel learning (MKL), artificial neural network (ANN), prosthetic control

1. INTRODUCTION

Myoelectric control has been the subject of a significant body of research in the last three decades. Researchers have been trying to develop prostheses which are intuitively controlled, i.e. the will of the amputee closely resembles the produced movement. This is achievable by exploiting electromyographic (EMG) signals measured at the end of the amputated limb. These signals are in fact electrical messages from the brain which pass through the neurons to reach the muscles and produce contractions [1]. Therefore, one may estimate the contractions based on the corresponding EMG signals [2].

This approach can be used to control prostheses and is referred to as myoelectric control.

There exists a significant literature based on EMG classification techniques e.g. [3-6]. These methods classify EMG features into a set of predefined corresponding motions (classes).

However classification based methods, provide a sequential on-off control which will make the prosthetic limb's movement less intuitive. To address this problem, recent work has proposed methods to incorporate simultaneous control of multiple degrees of freedom (DoFs) [7] instead of sequential control or proportional control instead of on-off control [8]. However a viable estimation method which would be both simultaneous and proportional is yet to be developed.

One of the major recent challenges is to develop a simultaneous proportional estimation method which provides intuitive control. Jiang et al. [9] have proposed an estimation method based on nonnegative matrix factorization (NMF) that doesn't need force data for training and can be used for bilateral amputees. But this method gives poor performance in presence of pronation-supination. Nielsen et al. [10] suggest mirrored bilateral training method based on the association between the surface EMG from one limb and force from the contralateral limb. This training strategy can be practically applied in unilateral amputees. This method again suffers from a decrease in performance as more DoFs especially supination-pronation are used simultaneously. Mucelli et al. [11] try to tackle this problem by using high density EMG channels. The results of this study shows that the performance for the supination- pronation DoF is influenced by the reduction in number of EMG channels to a greater extent than the other DoFs. Kamavuako et al. [12] test bilateral training on intramuscular EMG.

In the current study, we propose a new method for simultaneous and proportional control of the three DoFs of the wrist including the most challenging supination-pronation. This method improved the force estimation from EMG by means of using a novel structure for the algorithm. The organization of this paper is as follows: we briefly summarize the method introduced in [10] for mirrored bilateral training in section 2. Section 3 introduces the reader to the setup used in the experiments. The proposed method is described in section 4. Section 5 represents the simulation results. Section 6 states our conclusion.

2. BACKGROUND

In this section we briefly describe the main protocol for mirrored bilateral training employed by Nielsen et al. [10]. The block diagram of this method is depicted in Fig. 1. In the training phase, a multilayer perceptron (MLP) neural

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

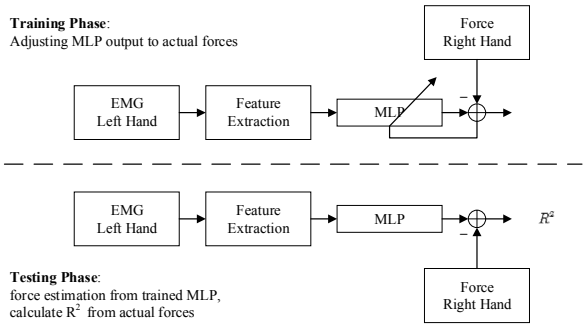

Fig. 1. Mirrored bilateral training algorithm for each DoF (Nielsen et al. [10])

network (used as regression tools) for each DoF learns the association between EMG signals measured on the left hand and the force signal of the contralateral (right) hand. The input to all 3 networks is a vector of features extracted from EMG channels. The output of MLP for each DoF is the force in that DoF (each MLP is trained for one DoF). the window length is typically chosen in the range of *100-250ms*. To reduce the processing time we can use non-overlapping windows or windows with a small overlap. In this way we also have to downsample the force data. This is depicted in Fig. 2.

During the testing phase the trained neural networks are used to estimate force in each DoF based on the EMG feature set of the opposite limb. In order to measure the performance of estimation method we use the R^2 metric, which has been previously shown to be an effective performance measure. The global R^2 is defined as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{D}\sum_{n=1}^{N}\left(\hat{f}_i[n] - f_i[n]\right)^2}{\sum_{i=1}^{D}\sum_{n=0}^{N}\left(f_i[n] - \overline{f_i[n]}\right)^2} \quad (1)$$

Where D is the number of force sequences (same as the number of DoFs), N is the number of force samples, $f_i[n]$ is the ith force sequence, and $\hat{f}_i[n]$ is the estimated value of the ith force sequence. $\overline{f_i[n]}$ is the mean of ith force sequence over n. Similarly R_i^2 for individual DoFs can be defined as:

$$R_i^2 = 1 - \frac{\sum_{n=0}^{N}\left(\hat{f}_i[n] - f_i[n]\right)^2}{\sum_{n=0}^{N}\left(f_i[n] - \overline{f_i[n]}\right)^2} \quad (2)$$

Closer values to 1 indicate better performances. Mean performance could be calculated using a cross validation procedure.

Selection of a good feature set is vital in achieving good performance. Many recent works have addressed this issue e.g. [13].

Training and testing sets are comprised of various recorded articulations, including contractions of single and combined DoFs. Estimating the force sequences when all of the DoFs are used simultaneously, is the most challenging part of this problem.

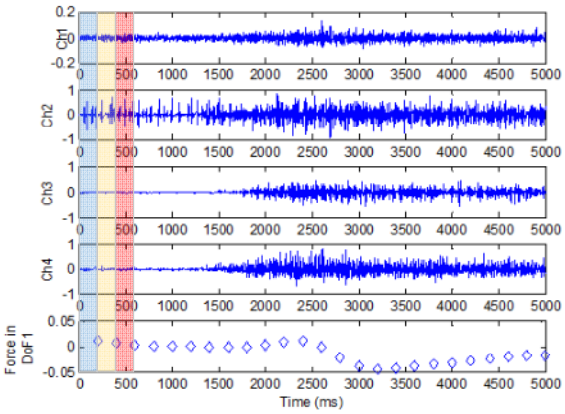

Fig. 2. Windowing the EMG signals. Each L (here 200) milliseconds of EMG signals in different channels is used to estimate one sample of each DoF's force at the end of the window. Three non-overlapping windows are shown in different colors.

3. EXPERIMENTAL SETUP

The three DoFs of the wrist; flexion-extension, abduction-adduction and pronation-supination, were investigated. Eight bipolar surface electrodes were attached to each arm, equally spaced around the circumference of the forearm. The electrodes were placed at approximately one-third of the forearm length, measured from the olecranon of ulna. The EMG data were sampled at 1 KHz using a 12 bit A/D converter.

Subjects sat in a chair, with arms secured to armrests. Their hands were fixed in a neutral position with palms facing inward using two vertical handles attached to a steel frame mounted in front of the chair. A commercially available 3-axis force/torque transducer (Gamma FT-130-10, ATI Industries) was mounted between the left handle and the steel frame, so that the x-axis corresponded to flexion-extension, y-axis to abduction-adduction and z-axis to pronation-supination. The force data were recorded at 1 KHz using a 12 bit A/D converter.

4. PROPOSED METHOD

4.1. Motivation

If we take a closer look at EMG and force signals, we'll see that when only one DoF is stimulated, some of EMG channels undergo a visible change and some of them remain almost unchanged. As can be seen in figure 3 the direction of DoF movement determines which channels have changed. This is quite natural, because for performing two opposite direction movements of each DoF (e.g. flexion and extension), different muscles are contracted, and these contractions won't occur simultaneously.

The described patterns shown in Fig. 3, brings this idea to the mind that we can first classify training data, and then

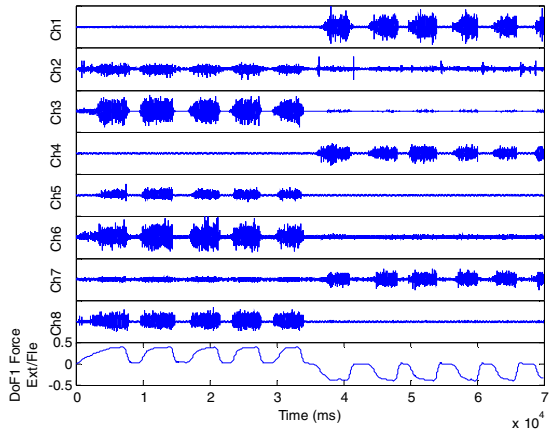

Fig. 3. EMG channels and the corresponding force in DoF 1 when only this DoF is stimulated.

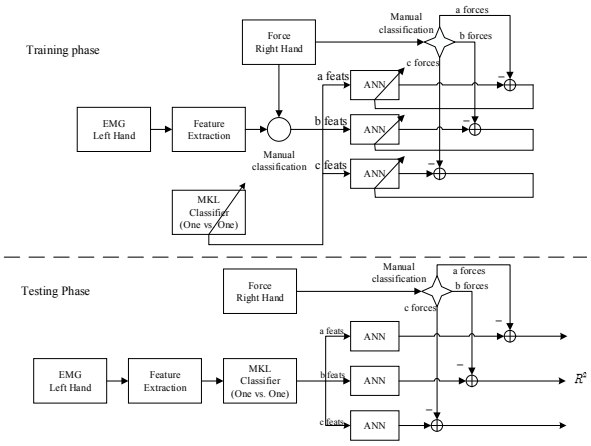

Fig. 4. Block diagram of the proposed algorithm

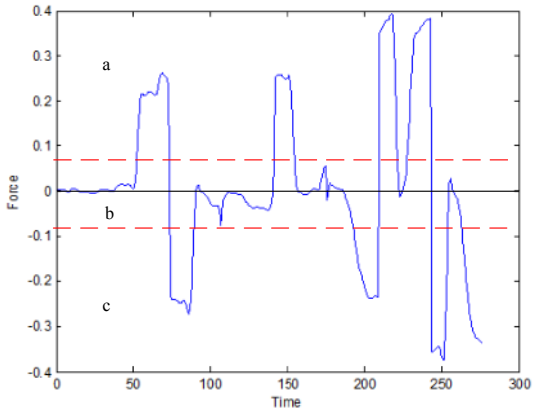

Fig. 5. Plane division

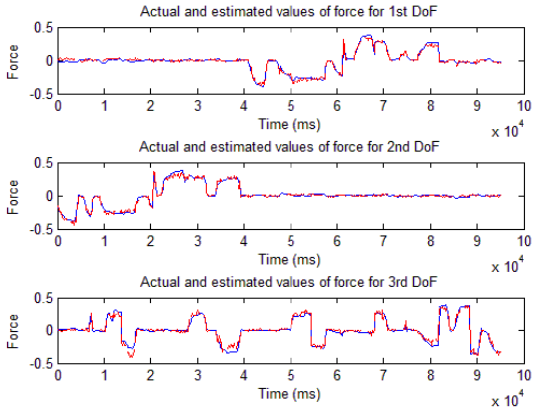

Fig. 6. Actual (blue) and estimated (red) values of force for each DoF in ideal mode.

use a neural network as regression tool for each class separately.
We will show that this simple idea can improve the estimation performance.

4.2. Algorithm

The block diagram of the algorithm is shown in Fig. 4. The blocks are explained in the next sub-sections.

4.2.1. Plane division

We divide the force planes (for each DoF) into three regions (Fig. 5); (a) positive region, (b) dead zone region, and (c) negative region. For the feature vectors on the training set we exactly know that which region each vector should be mapped to. Therefore we divide the EMG feature vector space, based on the vectors' output regions into three classes. The dead zone region is included to account for noise and unwanted signals and its bounds must be chosen small. Our experiments show that best performance is achieved when the bounds are set to some value in the range [0.03,0.1]. For each DoF, plane division must be done separately.

For each class and its corresponding output in any DoF, an MLP neural network was trained. These operations should be done for all of the DoFs in parallel.
Assume that we know the labels of the feature vectors in testing set also. As Fig. 6 represents, in case of this ideality we obtain a very good estimation performance. Especially in pronation-supination (the DoF with the lowest performance metric in our experiments) we get as much performance as the other DoFs, and the performance for each of the DoFs (R^2 metric) is more than 95%. But in reality we don't have such information.
In order to predict the label of the feature vectors we need to "classify" them using a proper pattern recognition method.

4.2.2. Multiple kernel learning (MKL) method

After the first use of the "kernel trick" in machine learning by Cortes and Vapnik [14] in 1995, several methods have been proposed to use a combination of kernels instead of a single one. These different kernels may correspond to using different notions of similarity or may be exploiting information coming from multiple sources [15]. We briefly explain this method here.
Support vector machines (SVMs) are binary classifiers based on the theory of structural risk minimization. As-

sume that we are given a set of N training samples $\{(\mathbf{x}_i, y_i)\}$ where $\mathbf{x}$ is the D-dimensional input feature vector and $y_i \in \{-1, +1\}$ is its corresponding label. SVM finds the linear discriminant function with the maximum margin in the feature space induced by the mapping function $\Phi : \mathbb{R}^D \rightarrow \mathbb{R}^S$. A discriminant function of the form

$$f(\mathbf{x}) = \left\langle \mathbf{w}, \Phi(\mathbf{x}) \right\rangle + b \tag{3}$$

is assumed, where $\mathbf{w}$ is the weight vector and b is a constant. SVM finds the optimal discriminant function by solving an optimization problem [15].
The resulting discriminant function is of the form:

$$f(\mathbf{x}) = \sum_{i=1}^{N} \alpha_i y_i k(\mathbf{x}_i, \mathbf{x}) + b \tag{4}$$

where $\boldsymbol{\alpha}$ is a vector of dual variables calculated during optimization procedure and $k : \mathbb{R}^D \times \mathbb{R}^D \rightarrow \mathbb{R}$ is a kernel function. Instead of using a single kernel function, MKL algorithms use a weighted sum of kernels, in which the weight vector is optimized during training. The mixed kernel function can be written as:

$$k_\eta(\mathbf{x}_i, \mathbf{x}_j) = f_\eta\left(\left\{k_m(\mathbf{x}_i^m, \mathbf{x}_j^m)\right\}_{m=1}^{P}\right) \tag{5}$$

where the combination function f_η can be a linear or a nonlinear function. Kernel functions take P feature representations (not necessarily different) of data instances: $x_i = \{\mathbf{x}_i^m\}_{m=1}^{P}$ where $x_i \in \mathbb{R}^{D_m}$ and D_m is the dimensionality of the corresponding feature representation.
Several learning methods for determining the kernel combination function. For a comprehensive comparison of these methods the reader can refer to [15].

4.2.3. Multiclass classification

There are three classes at the output of each classifier. So it's important to choose a proper binary classifier extension to match to the problem. We've tested one vs. one and one vs. all methods. One vs. one produces better results.

4.2.4. Feature selection

We have investigated the best choice for the feature set using forward selection. The following features have been chosen (the dimensionality of each feature subset is indicated in curly brackets); 1)The mean absolute value (MAV) feature used in [3] {8}, 2) Slope sign changes used in [3] {8}, 3) Waveform length used in [3] {8}, 4) Mean frequency used in [13] {8}, 5) Willison amplitude with threshold set to 0.025 used by [13] {8}, and 6) AR model parameters used by [3,13] {16}. Using AR model of an order greater than 2 decreased the performance in our experiments.

4.2.5 Feature grouping:

	Type of contraction
1	Abduction
2	Abduction-Pronation
3	Abduction-Supination
4	Adduction
5	Adduction-Pronation
6	Adduction-Supination
7	Extension
8	Extension-Pronation
9	Extension-Supination
10	Flexion
11	Flexion-Pronation
12	Flexion-Supination
13	Pronation
14	Supination

Table 1. Types of contractions

The main advantage of MKL techniques is that these methods can account for different sources features are coming from. The type of kernel, and its parameters for each arbitrary subset of features is chosen by the user. We have investigated the best grouping of the features and it has been found to be grouping of each of the aforementioned feature subsets separately. Putting all time domain features and all frequency domain features in two separate subsets gave lower performances.

4.2.6. MKL Parameters selection:

For each of 6 feature subsets, we've used an RBF kernel function. The typical value for RBF gamma parameters is set to be 4. The typical value for cost parameter C is 6. These parameters have also been optimized for each subject. The performance using the typical values and the optimized values has been measured.

4.2.7. MKL Classifier type selection:

As mentioned before, many MKL Classifiers have been proposed in recent years. We have investigated the use of different classifiers mentioned in [15] on the performance metric. The best results achieved by using SimpleMKL method [16].

5. SIMULATION RESULTS

We have simulated the method in [10] and our proposed method using MATLAB. There were 14 trials listed in Table 1. 80% of the data from each trial have been used for training and the remaining 20% for testing in each fold of a five fold cross validation procedure. In each case ANNs were trained 15 times with different initial weights, and the ones with the best performance were chosen.
Results from the simulation are plotted in figure 7. As can be seen, our proposed method outperforms the method in [10], for both typical and optimized parameters. In addition, a one way ANOVA test followed by a post-hoc Tukey-Kramer test was performed to ensure the significance of the improvement. The test indicated that performance

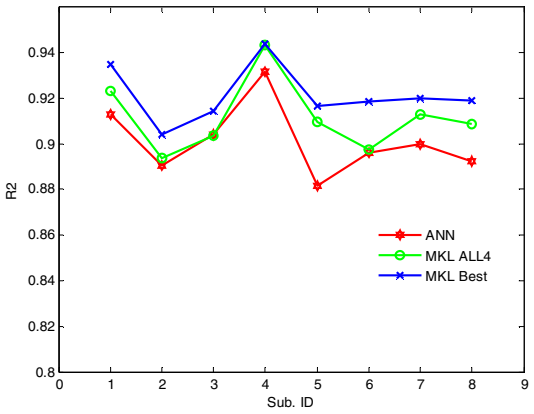

Fig. 7. Performance metric (R^2) for different methods (tested on 8 subjects)

of the proposed method with optimized parameters is significantly higher ($p < 0.05$) than that of [10]. Furthermore this improvement is mainly due to a significant increase in performance metric of the weakest DoF (pronation-supination), up to 4.5% for some of the subjects.

6. CONCLUSION

This study proposed a method based on both classification and regression to improve the performance of force estimation in multiple degrees of freedom. While most of the efforts in the literature are focused on improving the classification or regression performance by adding new features or making physical changes to the setup, our results show that it is possible to further improve the performance just by means of using a new configuration for processing. There are some patterns in the EMG signals which might not be fully exploited just by extracting more features. Instead, we have first classified these useful patterns into classes and then used a separate neural network for each class as a regression tool. This configuration produced significantly better results.

REFERENCES

[1] J. E. Hall (2010). Guyton and Hall textbook of medical physiology. Elsevier Health Sciences.

[2] P. Parker, K. Englehart, & B. Hudgins (2006). Myoelectric signal processing for control of powered limb prostheses. Journal of electromyography and kinesiology, 16(6), 541-548.

[3] B. Hudgins, P. Parker, & R. N. Scott (1993). A new strategy for multifunction myoelectric control. Biomedical Engineering, IEEE Transactions on, 40(1), 82-94.

[4] K. Englehart, B. Hudgins, P. A. Parker, & Stevenson M. Stevenson (1999). Classification of the myoelectric signal using time-frequency based representations. Medical engineering & physics, 21, 431-438.

[5] M. F. Lucas, A. Gaufriau, S. Pascual, C. Doncarli, & D. Farina (2008). Multi-channel surface EMG classification using support vector machines and signal-based wavelet optimization. Biomedical Signal Processing and Control,3(2), 169-174.

[6] L. J. Hargrove, E. J. Scheme, K. B. Englehart, & B. S. Hudgins (2010). Multiple binary classifications via linear discriminant analysis for improved controllability of a powered prosthesis. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 18(1), 49-57.

[7] A. J. Young, L. H. Smith, E. J. Rouse, & L. J. Hargrove (2012). A new hierarchical approach for simultaneous control of multi-joint powered prostheses. In Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on (pp. 514-520). IEEE.

[8] A. Fougner, Ø. Stavdahl, P. J. Kyberd, Losier Y. G. Losier, & P Parker (2012). Control of upper limb prostheses: terminology and proportional myoelectric control—a review. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 20(5), 663-677.

[9] N. Jiang, K. B. Englehart, & P. Parker (2009). Extracting simultaneous and proportional neural control information for multiple-DOF prostheses from the surface electromyographic signal. Biomedical Engineering, IEEE Transactions on, 56(4), 1070-1080.

[10] J. L. Nielsen, S Holmgaard., N. Jiang, K. B. Englehart, D. Farina, & P. Parker (2011). Simultaneous and proportional force estimation for multifunction myoelectric prostheses using mirrored bilateral training. Biomedical Engineering, IEEE Transactions on, 58(3), 681-688.

[11] S. Muceli, & D. Farina (2012). Simultaneous and proportional estimation of hand kinematics from EMG during mirrored movements at multiple degrees-of-freedom. Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 20(3), 371-378.

[12] E. N. Kamavuako, D. Farina, K. Yoshida, & W. Jensen (2012). Estimation of grasping force from features of intramuscular EMG signals with mirrored bilateral training. Annals of biomedical engineering, 40(3), 648-656.

[13] A. Phinyomark, C. Limsakul, & P. Phukpattaranont (2009). A novel feature extraction for robust EMG pattern recognition. arXiv preprint arXiv:0912.3973.

[14] C. Cortes, & V. Vapnik (1995). Support-vector networks. Machine learning,20(3), 273-297.

[15] M. Gönen, & E. Alpaydın (2011). Multiple kernel learning algorithms. The Journal of Machine Learning Research, 12, 2211-2268.

[16] A. Rakotomamonjy, F. Bach, S. Canu, & Y. Grandvalet (2008). SimpleMKL.Journal of Machine Learning Research, 9, 2491-2521.