Biomedical Signal Processing and Control 46 (2018) 182–191

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Using the sEMG signal representativity improvement towards upper-limb movement classification reliability

Vinicius Horn Cene*, Alexandre Balbinot

Department of Electrical Engineering, Laboratory of Electro-Electronic Instrumentation (IEE), Federal University of Rio Grande do Sul (UFRGS), Avenue Osvaldo Aranha 103, 206-D Porto Alegre, RS, Brazil

ARTICLE INFO

Article history:
Received 21 December 2017
Received in revised form 31 May 2018
Accepted 31 July 2018

Keywords:
Upper-limb
Extreme learning machines
EMG
Reliability
IEE
Differential Enhanced Signal

ABSTRACT

Several Machine Learning techniques have been employed to process sEMG signals in order to provide a reliable control biosignal. Although some papers report accuracy rates superior to 90%, there is a lack of more detailed reasoning for reliable systems capable of providing control signals to users that may, for instance, control a prosthetic device. In this paper, we combined two strategies in order to increase the representativity of the sEMG signals: (a) the use of a stochastic filter based on Antonyan Vardan Transform (AVT) prior the extraction of the signal features that reduces the stochastic behavior of the sEMG signal; and (b) a novel sEMG feature called Differential Enhanced Signal (DES), designed to increase the signal representativity in the sEMG transition sections where features based on time-domain are usually inefficient. Thus, using only RMS and DES features, we were able to mitigate the class overlap in the transition sections and consequentially increase the overall classification accuracy for training and testing of the system. Since a reliable output signal is desired to perform ultimate prosthetic control, a reliability metric was defined and evaluated, and once a non-reliable classification is detected, the system autonomously activates auxiliary methods based on post-processing and data discard to maintain the classification consistency. Three preliminary scenarios involving the AVT filter, a Wavelet filter and the unfiltered signal were compared in terms of accuracy rate to define the most efficient filtering technique. The signal representation using the combination of RMS and DES features was also compared to a set of Time Domain (TD) features to test its enhancement capabilities. The AVT-based filter and the DES feature were able to present higher accuracy rates in both accuracy scenarios tested. Three different databases including 60 subjects among amputees and non-amputees were used to appraise the system, which was able to reach a mean accuracy rate of 99.1% in the best-case scenario.

1. Introduction

The study of myoelectric signals to control assistive devices has continuously increased since 40s [1]. Since then, the developments were focused on the manipulation of the biosignals to enable the natural control of devices, granting more autonomy to amputee and disabled people [2–7]. The sEMG signal has known ranges of frequency and amplitude but also is characterized by a significant stochastic behavior which varies even for the same person. This signal behavior precludes the obtainment of an ideal Machine Learning model for the classification of the signals. Seeking to mitigate this issue, there were proposed several approaches centered on signal filtering, feature extraction, and different classifiers to obtain the lowest classification error at desirable application time.

1.1. sEMG signal filtering and feature extraction

To remove artifacts within the bandwidth of the sEMG signal, adaptive filters are commonly used as in [8–13]. Among the various solutions proposed for the filtering of sEMG signals, it is possible to find the use of Wavelet [14,15], Bayesian [8] and energy-based filters through the Teager-Kaiser Energy operator (TKE) [12]. Other filters are also used to remove specific artifacts such as heartbeats [9,13], ambient noise [10], and line noise [11].

Regardless the filtering and classification methods used, feature extraction is a vital stage in the classification process. In the context of the classification of sEMG signals, the features extracted must

* Corresponding author.
E-mail addresses: vinicius.cene@gmail.com (V.H. Cene), alexandre.balbinot@ufrgs.br (A. Balbinot).

https://doi.org/10.1016/j.bspc.2018.07.014
1746-8094/

provide enough repetitive and descriptive statistical parameters to enable the construction of a classifier with the lowest classification error possible. The choice of appropriate features is often more important than the control scheme itself concerning multifunctional control performance [16]. The extraction of features takes place in domains of time (TD) and frequency (FD) or using hybrid features in Time-Frequency Domain (TFD), or Time-Series Domain (TSD) approaches [1,17].

1.2. sEMG signal pre-processing and classification

Along with the extraction of features, some transformations such as the Principal Component Analysis (PCA) became very popular in the classifier model construction. The PCA is commonly applied in feature reduction [18,19], when usually only the most significant principal components are identified and used. Another typical application of PCA relies on the use of PCA for feature projection to obtain low-variance data values provided by the main component coefficients as input for the classifier. This process is widely used between feature extraction blocks and the classifier, presenting a low-pass filter effect on the system. This PCA implementation is used, for instance, in [20] for online automatic sEMG signal segmentation and in [21], to remove ECG noise from a single sEMG channel. Moreover, applications of PCA and different combinations of features have been used to identify hand movements [22], and the method is cited in several reviews in the area such as [1].

Several Machine Learning techniques such as Support Vector Machines (SVM) [1,23], Linear Discriminant Analysis (LDA) [24,22,25] and Artificial Neural Networks (ANN) [26,25] have been used to classify the sEMG signals. Besides, most recent methods such as Neuro-Fuzzy [27] and unsupervised methods combined with clustering strategies and classifiers [18,28] are also described in the literature.

Recently, some classical ANN architectures have been updated with newer strategies, forming interesting solutions such as Deep Learning, Convolutional Networks and Single Layer Feed-Forward Networks (SLFN) structures [29]. In the SLFN family, the ELM method has achieved compelling results on a broad variety of applications, reaching optimal solutions with higher computational efficiency when compared to iterative methods based on gradients and more traditional Feed-Forward structures [30–35].

1.3. Extreme learning machines

The ELM is a supervised learning method based on a Single Hidden-Layer Feedforward Neural Networks (SLFNs), which analytically reaches an optimal solution using a pseudo-inverse matrix to generate the output weights of the model. In addition to the computational speed offered by the ELM since there is no iteration or backpropagation to solve the system, another clear advantage is the lack of hyper-parameters tuning which are mandatory in other classifiers [31–35].

The most important (unique in the non-regularized form) parameter of ELM is the number of hidden neurons used in the network. Generally, the regularized version of the classifier tends to be more robust to the variation of this parameter. In a non-regularized classifier, those hidden neurons, when used on a small number, lead the method to underfit, and when used in excess, cause the overfitting of the model, these both situations tend to preclude the classifier accuracy. To find an adequate number of hidden neurons, the state of the art approach is focused on “pruning algorithms” often based on methods such as Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) [34,36–39], despite the empirical determination of this parameter been quite common.

ELM classifier is typically designed to process an **X** input matrix formed by **N** training samples and **d** features. As in the most classical models of ANN, the input data is delivered to a vector of input weights **W** formed by **d** features and **L** hidden neurons. The weights and biases values of these neurons are attributed randomly within an interval of [−1;1] and [−1,5;1,5], respectively. A small range of these values is essential to mitigate noise in the signal, so the classifier itself presents a low-pass filter behavior [34]. Once the inputs advance through the **W** input neurons, a ø kernel function is used to project the features and form the matrix **H**, as presented in Eq. (1). The kernel maps the features between the linear inputs $\mathbf{X}\cdot\mathbf{W}$ and linear output $\mathbf{H}\cdot\beta$, where β is the output formed by **L** hidden neurons for each class. The labeling occurs on the matrix **T**, which aggregates all output values of the sequential β vectors wherein the higher output class value labels each tested sample on an *argmax* scenario.

$$\mathbf{H}=\begin{bmatrix} ø(w_1x_1+b_{1,1}) & \cdots & ø(w_Lx_1+b_{1,L}) \\ \vdots & \ddots & \vdots \\ ø(w_1x_N+b_{N,1}) & \cdots & ø(w_Lx_N+b_{N,L}) \end{bmatrix} \quad (1)$$

Once the value of matrix **H** is obtained, the goal is to find a matrix β to solve Eq. (2) with the minimum error concerning the labels **T**.

$$\mathbf{H}\beta=\mathbf{T} \quad (2)$$

The inverse **H** matrix provides a simple solution of Eq. (2). However, since Eq. (2) usually describes an under or over-determined system, the direct inverse of **H** does not exist. Thus, a pseudo-inverse matrix represented by $\mathbf{H}^{\dagger}$ is used in Eq. (3) to obtain the β values of the model and solve the system to determine **H**.

$$\beta=\mathbf{H}^{\dagger}\mathbf{T} \quad (3)$$

The $\mathbf{H}^{\dagger}$ can be obtained efficiently using Singular Value Decomposition (SVD) to approximate the pseudo-inverse of an exact inverse of the matrix **H** giving a pre-defined tolerance. Alternatively, a regularized form of the ELM method could use strategies such as Ridge-Regression to define an optimal pseudo-inverse aiming to minimize the error of the ℓ_2 norm of the model, as presented on Eq. (4).

$$(\mathbf{T}-\mathrm{H}\beta)^{\mathbf{2}}+\lambda||\beta||_{\mathbf{2}} \quad (4)$$

From where the parameter β is calculated according to Eq. (5), using λ as regularization parameter that is generally expressed as $\frac{1}{C}$ (where I is an identity used to avoid singularities and C is the regularization parameter) [31,40,30]. Thus, the pseudo-inverse to solve the system usually is obtained by Eq. (6).

$$\beta=\mathrm{H}^{\mathrm{T}}(\mathrm{H}^{\mathrm{T}}\mathrm{H}+\lambda\mathrm{I})^{-\mathbf{1}}\mathrm{T} \quad (5)$$

$$\mathrm{H}^{\dagger}=\mathrm{H}^{\mathrm{T}}\left(\mathrm{H}^{\mathrm{T}}\mathrm{H}+\frac{\mathrm{I}}{\mathrm{C}}\right)^{-\mathbf{1}}\mathrm{T} \quad (6)$$

For the sEMG signal classification, as in other classifiers, the ELM algorithm has been used in combination with different pre-processing methods such as Wavelets [41], PCA [42–44] and Regression [41,45]. Late developments such as [41] are also mixing some post-processing methods such Major-Voting and Analysis Of Variance (ANOVA) to evaluate the classification of fingers movements of the amputee and non-amputee subjects. Moreover, the ELM method could also be used in its regularized form, as presented on [34,46].

1.4. Reliability of classification

According to [47], the reliability of a system could be defined as “*the probability of a product or service to operate properly for a*