Biomedical Signal Processing and Control 45 (2018) 267–273

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Effect of threshold values on the combination of EMG time domain features: Surface versus intramuscular EMG

Asim Waris [a,b], Ernest Nlandu Kamavuako [a,c,*]

[a] Department of Health Science and Technology, Aalborg University, Fredrik bajers vej 7 D3, 9220 Aalborg, Denmark
[b] Department of Robotics and Intelligent Machine Engineering, School of Mechanical and Manufacturing Engineering, National University of Sciences and Technology, H-12, Islamabad, Pakistan
[c] Center for Robotics Research, Department of Informatics, King's College London, 30 Aldwych, WC2B 4BG, London, United Kingdom

ARTICLE INFO

Article history:
Received 20 September 2017
Received in revised form 1 April 2018
Accepted 28 May 2018

Keywords:
Electromyography
Pattern recognition
Time domain features
Subject based optimum threshold
Population based optimum threshold

ABSTRACT

In myoelectric control, the calculation of a number of time domain features uses a threshold. However there is no consensus on the choice of the optimal threshold values. In this study, we investigate the effect of threshold selection on the classification for prosthetic use. Surface and intramuscular EMG were recorded concurrently from four forearm muscles on nine able-bodied subjects. Subjects were prompted to elicit comfortable and sustainable contractions corresponding to eight classes of motion. Four repetitions of three seconds were collected for each motion during medium level steady state contractions. The threshold for each feature was computed as a factor (R = 0:0.02:6) times the average root mean square of the baseline. For each threshold value, classification error was quantified using linear discriminant analysis (LDA) and k-nearest neighbor (KNN, k = 4) first for each individual feature and when combined. Three-way ANOVA revealed no significant difference between surface and intramuscular EMG (P = 0.997). However there was a significant difference between the features (P = 0.006) and between the classifiers (P < 0.001). The most dominant feature combination depended on the EMG (surface and intramuscular) and classifier. Results have demonstrated that using appropriate threshold value is very important to assure acceptable performance. For surface EMG, zero crossings (ZC) and slope sign changes (SSC) require no threshold, while a low threshold (R = 0.1:1) different from zero must be applied for willison amplitude (WAMP), myopulse percentage rate (MYOP) and cardinality (CARD). For intramuscular EMG, there is similar observation when using LDA as classifier. When using KNN, ZC SSC showed tendency to benefit from a low value threshold as well. Furthermore we propose the inclusion of a threshold that makes CARD robust to data precision.

1. Introduction

The electromyography (EMG) signal is one of the electrophysiological signals representing the neuromuscular activity during movements. It provides useful information about muscle condition. EMG based pattern recognition system have been extensively used in applications such as multifunctional upper limb prostheses [1,2], powered exoskeletons [3,4], rehabilitation robots [5,6], assistive computers [6,7] and wearable devices. Additionally, researchers and clinicians use profiles created by EMG features to investigate and diagnose neuromuscular conditions. Various pattern recognition algorithms have been proposed in the literature for the detection of motion intent [8–10]. Pattern recognition based control schemes extracts a set of features (time domain, frequency domain and time-frequency domain) that characterize the EMG signals in order to classify the user intended motions.

Multiple studies have evaluated the ability of various EMG features and classifiers to recognize different motions [10–12]. Classification accuracies > 90 percent have been reported in the literature [13,14]. Comparison of accuracies in these studies demonstrated that choice of features has more significant impact on the classification performance than the choice of classifiers [15,16,9]. Time domain (TD) features have been widely used in myoelectric control due to their computational simplicity and because they are easy to implement and do not require any signal transformation. Hudgins et al. introduced four TD features and now mostly referred to as Hudgins TD feature set [17]. Combining more robust and stable time domain features can significantly improve classification performance without increasing computation complexity [18].

* Corresponding author.
E-mail address: ernest.kamavuako@kcl.ac.uk (E.N. Kamavuako).

https://doi.org/10.1016/j.bspc.2018.05.036
1746-8094/

Table 1
Description of all features used in this study. N represents the total number of samples in a signal window; n is the sample index and $\boldsymbol{\varepsilon}$ is the threshold values defined in Eq. (1).

Feature	Description	Formula
MAV	Mean Absolute Value (MAV) is the average of the absolute value of the EMG signal. It is an indication of muscle contraction levels.	$MAV = \frac{1}{N}\sum_{n=1}^{N}\lvert x_n\rvert$
WL	Waveform length (WL) is related to the fluctuations of a signal when the muscle is active. Thus, the feature provides combined information about the frequency, duration, and waveform amplitude of the EMG signal.	$WL = \sum_{n=1}^{N-1}\lvert x_n - x_{n+1}\rvert$
ZC	Zero Crossing (ZC) measures the number of crosses by zero of the signal and is related to the frequency content of the signal. This feature provides an approximate estimation of frequency domain properties	$ZC = \sum_{k=1}^{N-1}[(x_n \cdot x_{n+1} < 0) \cap (\lvert x_n - x_{n+1}\rvert > \varepsilon)]$
SSC	Slope Sign Changes (SSC) measures the number of times the sign changes in the slope of the signal. It is another method to represent the frequency information of sEMG signal.	$SSC = \sum_{n=2}^{N-1}[(x_n - x_{n-1}) \cdot (x_n - x_{n+1})] > \varepsilon$
WAMP	Willison Amplitude (WAMP) estimates the number of active motor units, which is an indicator of the level of muscle contraction.	$WAMP = \sum_{n=1}^{N-1}\lvert x_n - x_{n+1}\rvert > \varepsilon$
MYOP	Myopulse Percentage Rate (MYOP) is defined to be the average value of the myopulse output. The myopulse output is defined as one when the absolute value of a signal is above a threshold and Zero otherwise.	$MYOP = \frac{1}{N}\sum_{n=1}^{N}\lvert x_n\rvert > \varepsilon$
CARD	Cardinality of a set is a measure of the number of distinct values. This can be computed in two steps. Data needs to be sorted and one sample is distinct from the next if the difference is above a predefined threshold.	Step 1: $y_n = sort(x_n),\ n = 1 : N$Step 2: $CARD = \sum_{n=1}^{N-1}\lvert y_n - y_{n+1}\rvert > \varepsilon$

Hudgins' set comprises the mean absolute value (MAV), waveform length (WL), slope sign change (SSC) and zero crossings (ZC). Among many other features that have been proposed, willison amplitude (WAMP) and myopulse percentage rate (MYOP) [19] have shown to contribute significantly to classification [10,20]. Recently Cardinality (CARD) has been proposed as a suitable feature [21] with improved performance. Features such as ZC, SSC, WAMP and MYOP are typically computed with a threshold value to attenuate the effect of background noise [18]. CARD feature does not require a threshold value, however, according to [21], attention must be paid to the unit length (precision) used for signal processing prior to the computation of cardinality. Altering the dimension of the unit used for sampling (ADC resolution) to a high precision unit (for example a double) would alienate the discrimination power of cardinality. This is because every sample value would be unique so that cardinality will always be equal to the number of samples in the time window [21]. We acknowledge the value of this comment and we introduce a threshold value that should eliminate dependency of this feature on unit length. Meaning a sample is considered unique if and only if its distance to the previous sample (after sorting) is greater than the defined threshold.

Although several studies have investigated proper selection of representative features, very few have investigated the effect of optimum thresholds on classification accuracies. Variable threshold values have been reported in the literature and in most of the studies threshold values were ignored or arbitrarily fixed. Hudgins et al. used threshold value of 2 μV for computing ZC and SSC for decoding hand motions [13]. Phinyomark et al. used threshold values between 0.5 and 50 mV for WAMP [15], and showed that 5 mV threshold performed best. Furthermore they extended the work by quantifying optimum threshold values for ZC and WAMP and concluded that thresholds are gain and instrument specific [22].

Recently [23] extensively studied the effect of threshold selection for ZC and SSC on the feature space and classification accuracy. Threshold for each feature was computed as a factor (R = 0:0.01:4) times the average root mean square of the data during rest period. Results demonstrated that threshold value has strong impact on features space and an optimum threshold value for each feature could be quantified, though with limited generalization ability. Nevertheless, the investigation was limited to surface EMG with two features and using only one classifier. Thus it is not known whether intramuscular EMG, which is gaining importance in myoelectric control, require different threshold levels compared to surface EMG.

In this study seven features MAV, WL, ZC, SSC, WAMP, CARD, MYOP features were investigated individually and in combinations to quantify the effect of each feature on classification error when the threshold is optimized for either surface or intramuscular EMG using Linear discriminant analysis (LDA) and *k*-nearest neighbor (KNN).

2. Materials and methods

2.1. Features

For this investigation, the focus is on the TD features which computation may require selection of a threshold value such WAMP, CARD, MYOP and we consider ZC and SSC for completeness. Table 1 summarizes all the features.

As proposed by [23], threshold is defined as a factor (R) times the average (across channels) root mean square value of the EMG signal at rest (during no contraction) (Eq. (1)). In this study R ranges from 0 to 6 with a step of 0.02.

$$\varepsilon = R * \sqrt{\frac{1}{N}\sum_{j=1}^{N}(x_{NM}[j])^2} \tag{1}$$

where $x_{NM}[j]$ are the samples of the signal at rest, and N is the total number of samples and NM stands for no motion which the signal measured during rest.

2.2. Experimental procedures

For this investigation, we used data for both surface and intramuscular EMG. Experiments were conducted on nine able-bodied subjects (age range: 19–26 yrs). The procedures were in accordance with the Declaration of Helsinki and approved by the local ethic committee of Northern Jutland (approval no.: N-20080045).