An algorithm for the estimation of the Signal-to-Noise Ratio in surface myoelectric signals generated during cyclic movements

Original

Availability:
This version is available at: 11583/2440716 since:

Publisher:
IEEE

Published
DOI:10.1109/TBME.2011.2170687

Terms of use:
openAccess
This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)
An Algorithm for the Estimation of the Signal-To-Noise Ratio in Surface Myoelectric Signals Generated During Cyclic Movements

Valentina Agostini and Marco Knaflitz, Member, IEEE

Abstract—In many applications requiring the study of the surface myoelectric signal (SMES) acquired in dynamic conditions, it is essential to have a quantitative evaluation of the quality of the collected signals. When the activation pattern of a muscle has to be obtained by means of single- or double-threshold statistical detectors, the background noise level \( e_{\text{noise}} \) of the signal is a necessary input parameter. Moreover, the detection strategy of double-threshold detectors may be properly tuned when the signal-to-noise ratio (SNR) and the duty cycle (DC) of the signal are known. The aim of this work is to present an algorithm for the estimation of \( e_{\text{noise}} \), SNR and DC of a SMES collected during cyclic movements. The algorithm is validated on synthetic signals with statistical properties similar to those of SMES, as well as on more than a hundred real signals.

Index Terms—Cyclic movements, signal-to-noise ratio (SNR), surface electromyography (sEMG), surface myoelectric signal (SMES).

I. INTRODUCTION

The analysis of the myoelectric signal allows to determine if a specific muscle is active or silent [1]. Surface myoelectric probes – applied on the skin above the muscle – permit a non-invasive evaluation of the muscle activity.

The traditional use of surface myoelectric signal (SMES) in physiological and biomechanical research [2] was extended to applied research: SMES achieved a well established value as an evaluation tool in surgery planning [3], rehabilitation [4], biofeedback [5]–[6], sports medicine and training [7], and ergonomics research [8].

Among other innovative applications, it is gaining interest demonstrating its usefulness in orthopedics, treatment of cerebral palsy, and a number of other clinical applications [13].

Well established techniques to determine the activation pattern of a muscle are based on single- or double-threshold statistical detectors [17]–[18]. These detectors require, as a necessary input parameter to set the (first) threshold, the background noise level. Furthermore, double-threshold detectors require additional parameters in order to fine-tune the second threshold. In particular, it is important to estimate the signal-to-noise ratio (SNR) and, to a minor extent, the duty cycle (DC) of the detected signal.

The vast majority of the above mentioned applications evaluates muscle activity in dynamic conditions and, in particular, during cyclic or repetitive motor tasks [10]–[12], e.g. walking (gait analysis) [13]–[14] or biking [15]. The movements of the corresponding body segments imply cyclic contractions of muscles. As a consequence, SMES detected from a specific muscle results in a sequence of pseudo-periodical activation bursts [16]. In the time elapsed between the end of a muscle burst and the beginning of the successive one, the muscle under study is silent. However, the probe detects a background noise that is unavoidable in any dynamic test. This background noise is mainly due to crosstalk [2] from neighboring muscles: it is a ‘physiological’ noise, more relevant than that due to instrumentation.

In order to evaluate the quality of the recorded signals, the background noise and the signal-to-noise ratio are usually estimated by means of manual or automatic segmentation of the signal in the time-domain. To this purpose, myoelectric bursts are separated from muscle baseline activity and the corresponding powers are calculated and compared.

In the analysis of dynamic SMES it is of paramount importance to obtain the onset and offset time instants of a muscle burst, distinguishing the muscle activity from the background noise. The relevance of considering the muscle activation timing is supported by several studies demonstrating its usefulness in orthopedics, treatment of cerebral palsy, and a number of other clinical applications [13].

The method we present in this contribution allows to estimate the root-mean-square value of the background noise \( e_{\text{noise}} \), the SNR and the DC of an SMES generated during...
cyclic movements. This is done following a statistical approach that does not require any a priori knowledge on the signal. In this way, we obtain the input parameters necessary to run double-threshold algorithms without the need of pre-processing the signal in the time domain.

II. MATERIALS AND METHODS

A. The surface myoelectric signal in cyclic movements

The SMES recorded during cyclic contractions may be considered as the superposition of the signal generated by the observed muscle during its contraction and the background noise. This noise is mainly due to the activity of neighboring muscles, collected because of the limited spatial selectivity of the detection probe. Moreover, due to the cyclic nature of the movement, this process may be defined as cyclostationary [19]. Its periodicity depends on the cyclic movement under investigation.

In its period of cyclostationarity, the signal can be modeled as the superposition of two stationary processes: noise only, when the muscle is not active, and a second process corresponding to the muscle activity.

When the muscle is non-active (OFF-state) only background noise is present. This noise can be modeled as a Gaussian process with zero-mean and variance $\sigma_n^2$:

$$n(t) \in N(0, \sigma_n^2).$$

During muscle activity (ON-state), the SMES can be modeled as a zero-mean Gaussian process given by the superimposition of two Gaussian processes corresponding to signal and background noise, respectively:

$$x(t) = s(t) + n(t) \in N(0, \sigma_s^2 + \sigma_n^2),$$

being $\sigma_s^2$ the variance of $s(t)$.

The percentage of time in which the muscle is active with respect to the total cycle duration is referred to as duty cycle.

In the following, it is assumed that the signal $x(t)$ is sampled with a sampling period $T_s$ that satisfies the Nyquist criterion. In particular, since SMES collected by means of usual surface probes typically has more than 99% of the signal power below 500 Hz, we consider a sampling frequency $f_s = 2$ kHz, that results in a two-time oversampling.

B. Separating signal from noise

The probability density function of the cyclostationary signal considered above is given by the superposition of the probability density functions corresponding to $x(t)$ and $n(t)$ and depends also on the DC. As an example, Fig. 1 reports the histograms of three cyclostationary processes corresponding to $e_{\text{noise}}$ equal to 1 $\mu V$, SNR equal to 20 dB and DC equal to 20%, 50%, and 80%, respectively.

The separation of the ON-state from the OFF-state could be obtained by applying a proper detector, but such detectors would require, as an input parameter, the root-mean-square value of the background noise $e_{\text{noise}}$ [18], that is unknown. A possible solution for estimating $e_{\text{noise}}$, without actually separating the ON and OFF states in the time domain, consists of considering an auxiliary time series with a $\chi^2$ distribution. The amplitude histogram of the auxiliary time series has two separated modes, one corresponding to the noise variance and the other to the signal variance.

The auxiliary time series $C_r(k)$ is obtained subdividing the time series $x(t)$ in $M$ epochs constituted by $r$ consecutive samples and then considering the normalized sum of squares of each epoch:

$$C_r(k) = \frac{1}{r} \sum_{j=1}^{r} X^2_{kj}, \quad k = 1, \ldots, M.$$
If the signal follows the hypothesized model, the time series has a bimodal distribution with two separated modes. The larger is the difference among $\sigma_n^2$ and $\sigma_s^2$, the greater is the distance between the modes.

As an example, Fig. 2 reports the bimodal distributions obtained for $r = 2$, $r = 10$ and $r = 50$, respectively. It is evident that increasing $r$ allows for a better separation of the two modes, but contemporarily, due to the reduced size of the series $C_r(k)$, we have a reduced number of samples. Moreover, increasing the $r$ value causes a loss of time resolution from 1 ms ($r = 2$) up to 25 ms ($r = 50$). A satisfactory tradeoff may be obtained by choosing $r = 10$ (time resolution equal to 5 ms). This guarantees an acceptable time resolution, a good separation between the noise and the signal modes, and a sufficient number of samples to build the histogram.

The assumption of gaussianity of the SMES allows to treat the auxiliary time series $C_r(k)$ as $\chi^2$-distributed. This is useful to determine the number of consecutive samples ($r$) that constitute $C_r(k)$ in an optimal way. However, even if the SMES is not exactly Gaussian, the auxiliary time series will still be two-bell shaped and the algorithm described below will give reliable results.

### C. Description of the algorithm

In order to estimate $e_{\text{noise}}$, SNR and DC of an SMES generated during cyclic movements we use the following algorithm:

1) Consider the time series \{x_i\}, $i = 1, \ldots, N$, being $N$ the number of samples. In the following, we refer to a time series with a duration equal to 30 s sampled at sampling frequency equal to 2 kHz. It follows that the number of samples $N$ is equal to 60000.

2) Divide \{x_i\} into $M = N/r$ epochs. Considering $r = 10$ we have:

\[{X_{ij}} = \{X_{i1}, X_{i2}, \ldots, X_{ij}\}, \quad j = 1, \ldots, 10\]

3) Obtain the auxiliary time series of the normalized sum of squares:

\[C(k) = \sum_{j=1}^{10} X_{ij}^2 \quad k = 1, \ldots, M. \quad (4)\]

4) Obtain the histogram of the series $\log_{10} C$. The bins of the histogram are defined as:

\[\text{bins}(m) = m \cdot \frac{\max(\log_{10} C) - \min(\log_{10} C)}{2 \cdot \text{Nbins}} + \min(\log_{10} C), \quad m = 1, 3, \ldots, 2 \cdot \text{Nbins} - 1. \quad (5)\]

where Nbins is the number of bins. Since in our case $M = 6000$, to have a sufficient sample numerosity for each bin we choose Nbins = 60. In general, a number of bins in the range 50-100 is an acceptable choice.

5) Search for local maxima of the curve that interpolates the frequencies of the histogram. Locate the absolute maximum and the highest relative maximum. The leftmost point of maximum is associated to noise ($I_{\text{noise}}$), the rightmost is associated to signal ($I_{\text{signal}}$).

6) Estimate the mean power of the noise, averaging five bins around $I_{\text{noise}}$:

\[P_{\text{noise}} = \frac{\sum_{i=1}^{I_{\text{noise}}-2} \text{bins}(i) \cdot \text{Freq}(i)}{\sum_{i=1}^{I_{\text{noise}}-2} \text{Freq}(i)}. \quad (6)\]

7) Estimate the mean power of the signal, averaging five bins around $I_{\text{signal}}$:...
TABLE I

<table>
<thead>
<tr>
<th></th>
<th>Rectangular Window</th>
<th>Gaussian Window with $\sigma = 1$</th>
<th>Gaussian Window with $\sigma = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Noise ($e_{\text{noise}}$) (µV)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SNR (dB)</strong></td>
<td>6 dB</td>
<td>12 dB</td>
<td>18 dB</td>
</tr>
<tr>
<td>20%</td>
<td>1.00 ± 0.01</td>
<td>1.00 ± 0.02</td>
<td>1.00 ± 0.01</td>
</tr>
<tr>
<td>40%</td>
<td>1.00 ± 0.04</td>
<td>1.00 ± 0.02</td>
<td>1.00 ± 0.02</td>
</tr>
<tr>
<td>60%</td>
<td>1.02 ± 0.03</td>
<td>1.01 ± 0.02</td>
<td>1.00 ± 0.02</td>
</tr>
<tr>
<td>80%</td>
<td>1.52 ± 0.63</td>
<td>1.01 ± 0.04</td>
<td>1.00 ± 0.02</td>
</tr>
</tbody>
</table>

9) Estimate the SNR (in dB):

$$SNR = 10 \cdot \log_{10} \frac{P_{\text{signal}} - P_{\text{noise}}}{P_{\text{noise}}}. \quad (9)$$

10) Estimate the duty cycle (%):

$$DC = 100 \cdot \frac{\sum_{i=1}^{\text{bins(i)}} \text{Freq}(i)}{\sum_{i=1}^{\text{bins(i)}} \text{Freq}(i) + \sum_{i=1}^{\text{bins(i)}} \text{Freq}(i)}. \quad (10)$$

In order to clarify how the algorithm works, Fig. 3 represents a 30s-cyclostationary signal with DC = 50% and SNR = 20 dB. Fig. 4 reports a similar example with SNR = 6 dB. As one could have expected, in the 6dB-case there is a partial superposition between the bell-shaped curve relative to the noise and that relative to the signal. However, their modes are still clearly distinguishable.

### III. RESULTS AND DISCUSSION

#### A. Validation of the algorithm

Synthetic signals are generated with a time duration of 30 s. They are cyclic with a cycle duration of 1 s. They are obtained adding a random process simulating the background noise to a process simulating the SMES generated during a muscle contraction. Fig. 5 represents 1 cycle of the synthetic SMES obtained with (a) a rectangular window, (b) a Gaussian window with $\sigma = 1$ and (c) a Gaussian window with $\sigma = 2$. In this example SNR = 20 dB and DC = 50%.
TABLE III
GAUSSIAN WINDOW WITH $\sigma = 2$

<table>
<thead>
<tr>
<th>e_noise (µV)</th>
<th>20%</th>
<th>40%</th>
<th>60%</th>
<th>80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 dB</td>
<td>1.02 ± 0.03</td>
<td>1.05 ± 0.03</td>
<td>1.08 ± 0.04</td>
<td>1.14 ± 0.12</td>
</tr>
<tr>
<td>12 dB</td>
<td>1.01 ± 0.02</td>
<td>1.02 ± 0.01</td>
<td>1.05 ± 0.03</td>
<td>1.20 ± 0.33</td>
</tr>
<tr>
<td>18 dB</td>
<td>0.99 ± 0.02</td>
<td>1.00 ± 0.02</td>
<td>1.02 ± 0.02</td>
<td>1.49 ± 1.04</td>
</tr>
<tr>
<td>24 dB</td>
<td>1.00 ± 0.02</td>
<td>1.00 ± 0.02</td>
<td>1.00 ± 0.03</td>
<td>1.36 ± 1.36</td>
</tr>
<tr>
<td>30 dB</td>
<td>1.00 ± 0.02</td>
<td>0.99 ± 0.02</td>
<td>0.99 ± 0.03</td>
<td>1.44 ± 2.19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>20%</th>
<th>40%</th>
<th>60%</th>
<th>80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 dB</td>
<td>4.5 ± 2.5</td>
<td>3.9 ± 2.0</td>
<td>1.2 ± 2.1</td>
<td>0.3 ± 2.1</td>
</tr>
<tr>
<td>12 dB</td>
<td>9.2 ± 1.0</td>
<td>9.6 ± 0.6</td>
<td>9.6 ± 0.8</td>
<td>8.2 ± 2.5</td>
</tr>
<tr>
<td>18 dB</td>
<td>16.6 ± 0.5</td>
<td>16.3 ± 0.9</td>
<td>16.2 ± 0.5</td>
<td>13.7 ± 4.9</td>
</tr>
<tr>
<td>24 dB</td>
<td>21.7 ± 1.0</td>
<td>22.1 ± 0.3</td>
<td>22.4 ± 0.4</td>
<td>21.1 ± 4.5</td>
</tr>
<tr>
<td>30 dB</td>
<td>28.2 ± 0.5</td>
<td>28.3 ± 0.5</td>
<td>28.4 ± 0.3</td>
<td>27.4 ± 4.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DC (%)</th>
<th>20%</th>
<th>40%</th>
<th>60%</th>
<th>80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 dB</td>
<td>10.5 ± 1.6</td>
<td>20.9 ± 1.3</td>
<td>36.4 ± 8.1</td>
<td>46.7 ± 9.7</td>
</tr>
<tr>
<td>12 dB</td>
<td>10.0 ± 0.4</td>
<td>22.0 ± 0.8</td>
<td>35.6 ± 0.9</td>
<td>50.6 ± 1.2</td>
</tr>
<tr>
<td>18 dB</td>
<td>10.0 ± 0.4</td>
<td>22.2 ± 0.8</td>
<td>39.2 ± 0.8</td>
<td>61.3 ± 2.0</td>
</tr>
<tr>
<td>24 dB</td>
<td>10.1 ± 0.8</td>
<td>22.9 ± 0.6</td>
<td>39.9 ± 0.9</td>
<td>64.1 ± 1.1</td>
</tr>
<tr>
<td>30 dB</td>
<td>10.4 ± 0.4</td>
<td>23.3 ± 0.5</td>
<td>40.8 ± 0.7</td>
<td>64.8 ± 0.8</td>
</tr>
</tbody>
</table>

Estimation of background noise ($e_{\text{noise}}$), signal-to-noise ratio (SNR) and duty cycle (DC) for synthetic signals obtained with a Gaussian window with $\sigma = 2$. It is reported the mean value estimated over ten realizations ± its standard deviation.

We considered 10 realizations of the described synthetic signals and estimated $e_{\text{noise}}$, SNR and DC for each of them. Then, we calculated the mean value and standard deviation of these parameters over the 10 realizations. Table 1 reports the results relative to the rectangular window, Table 2 those obtained using the Gaussian window with $\sigma = 1$, and Table 3 those relative to the Gaussian window with $\sigma = 2$.

The estimated background noise ($e_{\text{noise}}$) shows values that are almost always very close to the expected value of 1µV. In the large majority of cases, the corresponding error is lower than 1%. In the rectangular window case (Table 1) the estimated SNR and DC show values that are very close to their expected values. Again, in the large majority of simulation conditions, the error is lower than 1%. In the case of the Gaussian window with $\sigma = 1$, there is a slight underestimation of SNR and DC (see Table 2). This behavior is reinforced in the case of the Gaussian window with $\sigma = 2$. In fact, the estimated values are systematically and appreciably lower than the expected ones. This is not surprising and it is due to the shape of the applied window. When dealing with real signals similar to this typology, the estimated values of SNR and DC are always underestimated.

On the contrary, when the Gaussian windows are considered and DC is equal to 80%, the estimated value of the background noise is overestimated up to 60% relative to its true value. The effect of this overestimation is a slight reduction of the sensitivity of the single- or double-threshold detectors, that generally does not compromise their performances.

Although the method herein presented was developed to allow the tuning of statistical detectors, other applications are possible. As an example, the availability of the estimates of $e_{\text{noise}}$ and SNR makes it possible to evaluate the quality of the acquired signal in a user independent way. This is of paramount importance to allow for a prompt detection of signals whose quality is not high enough to guarantee reliable results of the processing techniques applied to them.

B. Test of the algorithm on real SMES

Although the validation of the proposed algorithm was carried out working with synthetic SMES, we tested the algorithm also on real signals.

In order to test the applicability of the algorithm to real SMES, recorded during cyclic human movements, we consider - as an example - signals acquired during gait. Our database consists of a total of 142 SMES collected from:
tibialis anterior (48), gastrocnemius lateralis (32), lateral hamstrings (24), vastus lateralis (24), rectus femoris (14). The SMES acquisition s were carried out positioning s urface myoelectric probes (STEP32, DemItalia, Italy) over the muscle’s belly [1]. As for the synthetic signals, we considered a time support of 30 s. The real signals are “pseudo-cyclic” with a cycle duration equal to the subject gait cycle.

Fig. 6 shows an example of SMES collected from (a) tibialis anterior and (b) gastrocnemius lateralis. The corresponding two-bell shaped histograms are shown aside.

We estimated the parameters both with the algorithm described above and by means of an alternative method. The parameters estimated through the proposed algorithm will be indicated as $\text{enoise}_1$, $\text{SNR}_1$, $\text{DC}_1$, while those estimated with the alternative procedure will be indicated as $\text{enoise}_2$, $\text{SNR}_2$, $\text{DC}_2$. The alternative procedure consists of segmenting the 30-s signals in the time domain, obtaining the ON and OFF states by means of the double-threshold detector [18]. An example of signal segmentation is shown in Fig. 7.

Indicating with $N$ the number of segmented gait cycles we define:

$$\text{enoise}_2 = \frac{1}{N} \sum_{i=1}^{N} \sigma_{B_1}^i,$$

where $\sigma_{B_1}^i$ is the standard deviation of the i-th background noise segment (see Fig. 7).

Furthermore, we define:

$$\text{SNR}_2 = 10 \cdot \log_{10} \left( \frac{\sigma_A^2}{\sigma_B^2} - 1 \right),$$

where $\sigma_A^2$ and $\sigma_B^2$ are the mean variances of the ON and OFF states, respectively.

Finally, we define:

$$\text{DC}_2 = \frac{1}{N} \sum_{i=1}^{N} \frac{A_i}{A_i + B_i},$$

where $A_i$ and $B_i$ are the ON and OFF states of the signal.

Fig. 7. SMES segmentation in the time domain. The segments labeled as $A_1$, $A_2$, $B_1$, $B_2$, … correspond to muscle activations (= ON states), while $B_1$, $B_2$, … indicate background noise (= OFF states).

Fig. 8 shows the comparison of the two methods on the set of 142 real signals. This figure displays, for each parameter, the scatter plot of the values obtained with the proposed algorithm (x-axis) and with the alternative method (y-axis). To determine how well the proposed algorithm retrieves the values obtained with the “direct” time-domain method we calculated the regression line among the points. The line among $\text{enoise}_1$-$\text{enoise}_2$ points has a slope close to 1 (0.97) and a very small y-intercept (0.41 $\mu$V rms), demonstrating the accuracy of the proposed algorithm in the estimation of background noise. Results are still acceptable when considering the SNR1-SNR2 scatter plot, also if it can be noticed a slight underestimation of SNR1 with respect to SNR2. For what concerns the duty cycle the values obtained with the algorithm are systematically lower than those obtained with the alternative method of about 20% of the gait cycle. These findings are similar to those already commented for the synthetic signals with Gaussian window with $\sigma=1$ and, more markedly, in the case $\sigma=2$.

IV. CONCLUSION

This work presents an algorithm for the estimation of background noise, signal-to-noise ratio and duty cycle of SMES generated during cyclic movements. The algorithm was tested on synthetic and real SMES and the obtained results show that, in most practical situations, it provides accurate and stable measures of the aforementioned parameters.

We adopted this method to choose the parameters of a double-threshold statistical detector we previously developed [18] and that we have been using in the past years to carry-out a user independent analysis of the SMES detected during walk. Results we obtained in that field are fully satisfactory and we believe that the approach herein presented could be beneficial also in other applications, when dealing with operator independent processing of SMES or other biomedical

Fig. 8. Scatter plots illustrating the comparison between the two methods. Each point in the plots represents the values of the parameters estimated with: (1) the proposed algorithm (x-axis) and (2) the time-domain method (y-axis), respectively. In each plot the regression line is shown superimposed as well as its corresponding equation.
signals with similar characteristics.

REFERENCES


Valentina Agostini was born in Rome, Italy, on November 2, 1970. She received the Italian Laurea in physics summa cum laude from the University of Torino, Torino, Italy, in 1995 and the Ph.D. degree in physics from Politecnico di Torino, Torino, Italy, in 2002. Since 2004 she is with the Dipartimento di Elettronica of Politecnico di Torino, Torino, Italy, as Research Assistant. She focused her research interests in medical imaging and biomedical signal processing since 2004. She worked at the evaluation of the signal-to-noise ratio of infrared image sequences and at the motion artifact reduction in breast dynamic infrared imaging. Her current research interests are in statistical gait analysis and static posturography.

Marco Knaflitz (M’92) was born in Turin, Italy, on January 20, 1955. He received the Italian Laurea in electrical engineering and the Ph.D. in electrical engineering from Politecnico di Torino, Torino, Italy. From 1986 to 1990 he was with the Neuromuscular Research Center of Boston University, Boston, MA, as Research Assistant Professor. From 1995 to 2000 he was appointed Adjunct Professor at the Union Institute of Cincinnati (OH). Since 1990 he has been with the Dipartimento di Elettronica di Politecnico di Torino, Torino, Italy, where he is presently Full Professor of Biomedical Engineering and he teaches Biomedical Instrumentation and Implantable Active Devices. He has been active since 1985 in the fields of design of biomedical instrumentation and biomedical signal analysis. Specifically, his research interests are mainly focused on the detection and analysis of the myoelectric signal for research and clinical applications.

Prof. Knaflitz is a member of GNB, the Italian national group of bioengineering.