Surface Electromyographic Onset Detection Based On Statistics and Information Content

Natalia M. López¹, Eugenio Orosco², Fernando di Sciascio².

1- Gabinete de Tecnología Médica, Facultad de Ingeniería, Universidad Nacional de San Juan.

2- Instuto de Automática, Facultad de Ingeniería, Universidad Nacional de San Juan.

nlopez@gateme.unsj.edu.ar

Abstract. The correct detection of the onset of muscular contraction is a diagnostic tool to neuromuscular diseases and an action trigger to control myoelectric devices. In this work, entropy and information content concepts were applied in algorithmic methods to automatic detection in surface electromyographic signals.

1. Introduction

The surface electromyographic signal (sEMG) is a common diagnostic tool to detect neuromuscular pathologies while it also finds applications in rehabilitation engineering, sports, ergonomy and research. Temporal and energy characteristics supply parameters to estimate force of contraction, so the accuracy in the detection of beginning and end allows the evaluation of posture and movement.

Differences of 20 ms in conduction velocity studies, or in alignment of movement-related potentials in the electroencephalogram (EEG) are significant to the diagnosis of several pathologies [1- 4]. In others applications, as prosthetic or assistive devices control, the correct detection of sEMG beginning is critical in applications where that event triggers an action, or in signal classification where the operator’s intent is extracted in the first 300 ms of the muscular contraction [5].

Subjective visual determination by a trained operator is the most common way to evaluate contraction detection algorithms and used as reference [6-8]. However, precision and repeatability are low requiring a large number of well characterized EMG’s if any meaningful statistical analysis is wanted. Also, this method is restricted to offline applications. For this reason, an algorithmic method for efficient onset detection is required to reduce the operator’s induced errors and allow online processing. Several research groups propose algorithms based in lowpass filters [9], statistical considerations [10], whitening filters [11], energy operators [12] and others algorithms. A systematic comparison of the methods present in bibliography can be found in the work of Staude [13] or by Hodges and Bui [14].

In this paper we introduce an approach based in probabilistic considerations and information content. The concepts of information content, entropy and algorithmic complexity theory have found application in biological signal analysis [15-17], speech and audio recognition, and other fields.

Onset detection based in probabilistic models is a usual technique in the field of music analysis, where the onsets of significant events are perceived as surprising moments, based on the notion of observer’s familiarity with the evolution of the signal. This observer generates a probability trace of the baseline and the occurrence of an event can be seen as an unexpected change in probability.
building a temporal structure [18]. Thus, a dynamically evolving measure of surprise can be used as a
detection function in the analysis of sEMG, since muscular contraction is a surprising event.

There is a close relation between information content and entropy. Complexity and information
measures have been used in several applications in the field of biological signal processing, as speech
recognition [19] and sEMG characterization using approximate entropy [20], so the Shannon Entropy
of the signal was analyzed in this context. Shannon’s entropy describes the uncertainty or the degree
of disorder of the data [21, 22], that is, the information content of the signal.

In order to perform a statistical test of the method here proposed, a large amount of sEMG signals
is required, with parameters such as start point, amplitude and signal to noise ratio (SNR) perfectly
known. In this paper we use the same approach as the proposed by Staude et al [23], who use
simulated EMG data with the possibility of changing the parameters that affect the sensitivity and
improve the reproducibility. The surprise and entropy algorithms were compared with other detection
methods, with simulated and real EMG signals and the performance was evaluated under different
levels of SNR.

In what follows, we will present the database used to validate the algorithms, the two methods
proposed and the results obtained.

2. Materials

2.1. sEMG database

2.1.1. Synthetic sEMG database.

Automatic event detection from sEMG signals requires a model containing abrupt changes in signal
properties and variable parameters in the baseline.
A zero mean steady stochastic process, of Gaussian distribution, band limited (10-500Hz) and
contaminated with additive white noise appears as a good modeling choice [24, 25].
According to the model of Staude [23] an all-pole representation by an autoregressive (AR) filter
provides an adequate bandwidth while the excitation \( x_k \) results from two statistically independent
noise sources, \( m_k \) and \( w_k \), dynamical mean signal and system noise, respectively (Fig. 1). The
complete EMG digitized signal \( y_k \) can be model as:

\[
(1) \quad y_k = \tilde{y}_k + v_k \\
(2) \quad \tilde{y}_k = -\sum_{i=1}^{p} a_i \tilde{y}_{k-i} + b_0 x_k \\
(3) \quad x_k = m_k + w_k
\]

where \( \tilde{y}_k \) denotes the output of the AR filter of order \( p \) with coefficients \( a_i \) and static gain \( b_0 \), and \( v_k \)
is the measurement and other biological signals acting as interference .
The onset of the contraction is flagged by an abrupt change in the profile of \( m_k \) at the arbitrary time
\( t_1 \), modeled by a ramp between \( t_2 - t_1 \).

In the following, the onset time \( t_1 \) will be setting randomly among the interval of 3-5s. The final ramp
time \( t_2 \), noise baseline and maximal contraction amplitude were setting in the same way, in a bounded
span. To preserve the Nyquist considerations the sampling frequency (fs) is 1KHz. A typical simulated
sEMG signal is shown in Fig. 2.
Figure 1: sEMG generation model. A gaussian white noise model and a linear filter $H(z)$ model the $emg_s$ signal. Internal and measurement noise are modeled by additional independent processes. Abrupt changes are driven by deterministic coefficients included in variance variable.

2.1.2. Real sEMG database.

To optimize onset detection algorithms, a sEMG real database was constructed with bipolar recordings made with Ag-AgCl electrodes (3M RedDot) fixed 2 cm apart in the biceps of 6 volunteers during right arm flexion. The signal was amplified, filtered (10 Hz/500 Hz for low/high frequencies cut-off, respectively), and digitized ($fs=1KHz$) on-line with BrainNet BNT36 technology (LYNX TECNOLOGIA ELETRÔNICA LTDA, SP, Brasil). Several acquisition conditions were tested using different noise sources and the analysis was performed off line.

3. Methods

Di Fabio’s algorithm is most commonly used to detect onset time and it is also a common reference for a new technique [9], [14]. The method uses a fixed threshold ($Th$), calculated by adding a number of standard deviations above the mean value of the baseline of the EMG, in resting state. Besides, the signal must surpass the threshold during a whole period, to reduce false positive detections. Both conditions are selected empirically to obtain the best output.

\[ Th = \overline{emg_{(rest)}} + 3SD \]

We use this technique to compare with our results.

3.1. Surprise Function

In order to measure the degree of unexpectedness $s(E)$ of an event $E$, it is reasonable to assume the following conditions:

- $s(E) = f(p(E))$, $s(E)$ only depend on the probability of the event $E$.
- $s(E)$ is a monotone decreasing function of $p(E)$
- $s(E) = 0$ for $p(E) = 1$, and $s(E) = \infty$ for $p(E) = 0$

The logarithm satisfies these conditions, thus the surprise can be defined as in [18]:

\[ s(E) = -\log p(E) \]

Using the concept of a negative log-probability as a measure of the surprise or unexpectedness of the contraction in the sEMG signal, it is possible to determine the onset in terms of an implicit
probability model. This model can be unconditional or conditional, according to the behavior of the signal in relation with the past samples. In this work, the unconditional assumption was applied to reduce the computational cost.

Calculation of \( s(E) \) function requires segmentation of the signal in disjointed windows \( (\nu(k)) \), independent and identically distributed. Now, if we assume that the elements of \( \nu(k) \) are Gaussian distributed with zero mean and variance \( \sigma^2 \), the probability density function (pdf) is:

\[
(6) \quad P(\nu) = \prod_{i=1}^{n} \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{\nu_i^2}{2\sigma^2}}
\]

where \( \nu_i \) represents each of the \( n \) elements of the window elements. From (5) and (6) we obtain the measure of surprise, being \( c \) a constant. As can be seen, the surprise is closely related to the energy concept.

\[
(7) \quad s(E) = \frac{1}{2\sigma^2} \sum_{i=1}^{n} \nu_i^2 + c
\]

In the work of Clancy and Hogan [25], the probability density function of the sEMG was revised theoretically and experimentally. They propose a model using a Laplacian density. Hence, a Laplacian distribution is assumed to compute the \( s(E) \) function, that is,

\[
(8) \quad P(\nu) = \frac{1}{2} e^{-\frac{\nu_i}{\lambda}}
\]

\[
(9) \quad s(E) = \lambda \sum_{i=1}^{n} |\nu_i| + c
\]

In this case, the algorithm gives the same results that rectification and smoothing, a procedure very usual in sEMG processing, related to the energy contained in the signal.

For both distributions, the success of the onset detection implies the correct fitting of the statistical model.

3.2. Entropy and information content

Entropy as a thermodynamic state variable was introduced into physics as a measure of system disorganization. Afterward, entropy was promoted to a generic measure of information content and Shannon [21, 22] formalized the concept applied to signal processing:

\[
(10) \quad S(m) = -\sum_{i=1}^{n} p_i \log p_i
\]

where \( p_i \) is an estimate of the pdf for sample \( i \), \( n \) represents the window size and \( m \) is the window number. A proposed relation between entropy and sEMG processing is based on the hypothesis that the measurement and environmental noise is supposed to have the highest entropy value while the contraction signal has significantly lower entropy, as a large amount of information (i.e. muscular force) is present in the signal. To detect the contraction onset a modified version of Shannon equation was applied, computing the empirical histogram of a signal segment and counting the number of repetitions of each value in this section.
The EMG activity beginning was setting when at least 3 consecutive windows met the following criterion,

\[ S(m) > \bar{S}(\text{rest}) \]

That is, the window’s entropy must be larger than the mean baseline entropy.

For both algorithms, surprise function and Shannon’s entropy, the window size was empirically set as a function of the sampling frequency recalling that the smaller the number of samples, the better the detection precision (meaning repeatability); however, if the window becomes too small, false detections are more likely to occur due to noise spikes. Keeping the aforementioned criteria led us to choose the window size \( n = 50 \) samples.

4. Results

Figure 2 shows onset detection results by the surprise algorithms and Shannon’s entropy in a synthetic signal.

![Figure 2: Results obtained with surprise algorithms, Shannon entropy and di Fabio method. Onset estimates are indicated by arrows.](image)

A set of 1000 signals randomly created according to Section 2.1.1 was tested along a SNR range of 16 dB. To verify the algorithms, visual decision of an expert and Di Fabio’s method were used as reference. Signal to noise ratio (SNR) in dB was computed as:

\[ SNR = 10 \cdot \log_{10} \frac{\sigma_{\text{signal}}^2}{\sigma_{\text{baseline}}^2} \]

The detection error was calculated as: \( error = abs(t_i - t_d) \), where \( t_d \) is the time of onset detected by the algorithms here proposed.
Table 1 depicts the statistical performance of the onset algorithms herein proposed, over the synthetic database. The Gaussian surprise function shows an unacceptable error and was discarded, while mean, standard deviation and mode in Shannon’s entropy and Laplacian surprise are in the same order that di Fabio’s method.

|                  | Mean error | Error STD | Mode error | Range error |
|------------------|------------|-----------|------------|-------------|
| Di Fabio         | 1.219 ms   | 0.9995 ms | 0.003756 ms| 4.433 ms    |
| Shannon Entropy  | 1.604 ms   | 1.478 ms  | 0.5507 ms  | 11.11ms     |
| Gaussian Surprise| 3.328 ms   | 0.5186 ms | 3.34 ms    | 2.654 ms    |
| Laplacian Surprise| 1.392 ms  | 0.8479 ms | 0.00766 ms | 4.248 ms    |

Table 1: Statistics for onset algorithms, range of SNR 16dB.

Figure 3 illustrates the detection error at different levels of SNR. The Laplacian surprise function shows a slightly better performance at low SNR values and can be considered as an option to suitable onset detection. Shannon Entropy algorithm’s presents a stepped behavior and needs to be improved, to be considered as a suitable choice.

Figure 3: Dependence of error on SNR, with a 50 samples window’s size.

In real EMG signals the performance of the algorithms is similar; Laplacian surprise is closer to the onset detected by the expert than di Fabio’s algorithm. A statistical analysis is not suitable because the expert decision is not an exact parameter. A typical signal is shown in Fig. 4.
5. Discussion and Conclusions

This paper proposes the use of information theory based techniques to automated detection of the onset sEMG contraction, based on statistical considerations and Laplacian distribution. Other techniques as Gaussian distribution and even Shannon Entropy were discarded as noisier, less reliable and more costly from a computational point of view.

The Gaussian surprise function is unable to detect the onset of contraction, because the statistical fitting supposes a change in distribution and both, baseline noise and sEMG, were simulated with a white Gaussian noise source. All detectors herein presented are related to signal energy estimation, which is consistent with an optimal detector under the assumption of Gaussian distribution for both signal to noise base.

On other hand, the Laplacian surprise detection error is in the order of the window’s size, while a smaller window (i.e. 50 samples) yields to false detections. This algorithm shows better performance when high noise levels are expected, even better than Di Fabio threshold method.

References

[1] Adamovich SV, Levin MF, Feldman AG. “Merging different motor patterns: coordination between rhythmical and discrete single- joint movements”. Exp Brain Res.99:325–37; 1994
[2] Takada K, Yashiro K. “Automatic measurement of on/off periods of EMG activity”. Medinfo.8:751–4; 1995
[3] Rafał RD, Winhoff A, Friedman JH, Bernstein E. “Programming and execution of sequential movements in Parkinson’s disease.” J. Neurol Neurosurg Psychiatry. 50:1267–73. 1987.
[4] Haagh S, Spijkers W, Van den Boogaart B., Van Boxtel A., “Fractioned reaction time as a function of response force.” Acta Psychol., vol. 66, pp. 21–35, 1987.
[5] Hudgins B., Parker F., Scott R. “A new strategy for multifunction myoelectric control”. IEEE Trans. Biomed. Eng., vol. 40 pp. 82-94. 1993.
[6] Barrett G, Shibasaki H, Neshige R. “A computer-assisted method for averaging movement-related cortical potentials with respect to EMG onset”. Electroenceph Clin Neurophysiol 1985;60:276–81.
[7] Badke MB, Duncan PW. “Patterns of rapid motor responses during postural adjustments when standing in healthy subjects and hemiplegic patients”. Phys Ther 1983;63:13–20.
[8] Walter CB. “Temporal quantification of electromyography with reference to motor control research”. Human Movement Science 1984;3:155–62.
[9] Di Fabio RP. “Reliability of computerized surface electromyography for determining of onset of muscle activity”. Physical Therapy 1987;67(1):43–8.
[10] P. Bonato, T. D’Alessio, and M. Knaflitz, “A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait,” IEEE Trans.Biomed.Eng., vol. 45, pp. 287–298, 1998.
[11] G. Staude, W. Wolf, U. Appel, “Automatic event detection in surface EMG of rhythmically activated muscles,” Int. Conf.IEEE BME Soc., vol. 17, pp. 1351–1352, 1995.
[12] Solnik S, DeVita P, Rider P, Long B, Hortobagyi T. “Teager-Kaiser Operator improves the accuracy of EMG onset detection independent of signal to noise ratio” Acta Bioeng Biomech 10(2), 65-68. 2008.
[13] Staude G, Flachenecker C, Daumer M, Wolf W. “Onset detection in surface electromyographic signals: a systematic comparison of methods”. EURASIP Journal on Applied Signal Processing. 2001;2:67–81.
[14] Hodges PW, Bui BH. “A comparison of computer-based methods for determination of onset of muscle contraction using electromyography”. Electroenceph Clinical Neurophysiology 1996;101:511–9.
[15] Abásolo D, Hornero R, Gómez C, García M, López M, “Analysis of EEG background activity in Alzheimer’s disease patients with Lempel–Ziv complexity and central tendency measure”. Medical Engineering & Physics 28, 315–322. 2006
[16] G. Barrett, H. Shibasaki, and R. Neshige, “A computer-assisted method for averaging movement-related cortical potentials with respect to EMG onset,” Electroenceph.Clin.N europhysiol. vol. 60, pp. 276–281, 1985.
[17] P. Brodin, T. S. Miles, and K. S. Turk, “Simple reaction time responses to mechanical and electrical stimuli in human masseter muscle,” Arch. Oral Biol., vol. 38, pp. 221–226, 1993.
[18] Abdallah S., Plumbley M, “Probability as metadata: event detection in music using ICA as a conditional density model”, Proc. 4th Int. Symp. Independent Component Analysis and Signal Separation (ICA2003), Nara, Japón. pp. 233–238, 2003.
[19] Torres M. E., Gamero L. G., “Relative Complexity Changes in Time Series using Information Measures”, Physica A (ISSN: 0378-4371), Vol. 286, Iss. 3-4, pp.457-473, Oct. 2000.
[20] Chen W, Zhuang J, Yu W, Wang Z., “Measuring complexity using FuzzyEn, ApEn, and SampEn”, Med Eng Phys. 2009 Jan; 31(1):61-8.
[21] Shannon, C. E., “A mathematical theory of communication”. Bell System Technical Journal, 27. p. 379-423.
[22] Shannon C., Weaver W. “The Mathematical Theory of Communication”. Univ of Illinois Press, 1963.
[23] Staude G, Flachenecker C, Wolf W. “A dynamic generator model for motor related biosignals”. Ann Int Conf IEEE BME Soc 1998;20:2374–7.
[24] De Luca CJ. “Physiology and mathematics of myoelectric signals”. IEEE Trans Biomed Eng 1979;26(6):313–25.
[25] Clancy E, Hogan N. “Probability density of the surface electromyogram and its relation to amplitude detectors,” IEEE Trans.Biomed.Eng., vol. 46, pp. 730–739, 1999.