THE PERFORMANCE OF ONSET DETECTION METHODS FOR SURFACE ELECTROMYOGRAPHIC DATA

A. Andrews, L. McLean
Motor Performance Laboratory, Queen's University, Kingston, Canada

ABSTRACT

The performance of five different algorithm-based onset detection methods for surface electromyography (SEMG) data was analyzed relative to visually determined onset times from three expert volunteers. It was hypothesized that at least one algorithm would out-perform the others in terms of degree of correctness.

Three-hundred data plots from three previous studies on motor control were selected as source data. The automated algorithms tested included: i) a forward-moving sliding window, considering window mean value and number of points above a threshold [12], ii) a backward-moving sliding window starting at the data peak, considering window mean value compared to a threshold (developed by authors), iii) a backward-moving sliding window starting at the highly smoothed data peak, considering window mean value and number of points below a threshold (developed by authors), iv) a forward-moving sliding window, considering window mean value and number of points above a threshold lasting for a minimum number of consecutive windows [13], and v) a system based on log likelihood ratios and maximum likelihood estimation [11]. Each method went through parameter optimization as part of the testing.

The three expert volunteers determined onset times by visual inspection for the 300 data plots on two occasions for each of two filter methods: a 2nd order Butterworth filter with 6 Hz cut-off frequency, and a 20 ms sliding window root-mean-square (RMS) filter.

The algorithms were ranked based on i) a Z value, expressing the average of each data plot's deviation from the visually determined onset distribution and ii) having no more than 10% erroneous results. The forward moving sliding window, considering window mean value and number of points above a threshold [12], resulted in the most accurate determination of onset time, with an RMS averaged Z value of 0.7738. However, the algorithms were comparable and several had unique advantages.

INTRODUCTION

The determination of electromyographic (EMG) onset time is an important step in electrophysiological data analysis [1-10]. The EMG onset marks the time at which a muscle is first considered to be active and is a fundamental characteristic used in biomechanics and motor control research. However, despite its clinical importance, there is not yet a widely accepted standard with which the onset time is determined.

A significant amount of literature reports the use of visually determined EMG onsets [1,2,5,8,11,12], but this method can lead to lengthy data processing times and rater bias. The advantages provided by automated onset detection processes include the speed which onset times can be found and the inherent repeatability. However, a fully automated system must have high accuracy in order to be trusted. Without high accuracy, the advantages of speed and precision are lost because the data will have to be reviewed manually to ensure correctness. This is commonly the case and thus results of automated processes must often be re-examined by visual inspection [3].

The purpose of this study was to determine which of the five different algorithm-based onset detection methods provided the most accurate performance when tested against a distribution of visually determined onset times.

METHOD

Source Data

Three-hundred data plots were selected evenly from three studies performed in the Motor Performance Laboratory at Queen's University. The muscles from which SEMG data were used included anterior deltoid, infraspinatus, pectoralis major, posterior deltoid, serratus anterior, subscapularis, supraspinatus, teres major, levator ani, rectus abdominis, internal obliques, external obliques, and multifidus.

Abdominal muscle data were acquired using surface Ag-AgCl electrodes (Kendall Meditrace™) and a Bortec AMT-8 EMG amplifier system (two stage gain with preamplifier gain fixed at x200 and overall gain set to x1000, CMRR -120dB at 60Hz, R_{IN}>1 GOhm, band-pass 10Hz-1kHz) and shoulder muscle data were acquired using DelSys™ DE 2.1 differential electrodes and an AMT-8 amplifier system (two stage gain with pre-amplifier gain fixed at x500 and overall gain set to x1000, CMRR -80dB at 60Hz, R_{IN}>100

MOhm and band-pass 20Hz-450Hz). All data were acquired using a 16bit National Instruments AD Converter at a sampling rate of 1000Hz.

Visually Determined Onsets

Visually determined onset (VDO) times for all data plots were determined during four separate sessions by each of the three experts. The experts each had at least five years of experience working with kinesiological EMG data in a research setting. The time between detection sessions varied between one and three weeks.

Each session consisted of selecting the onset location on all 300 plots, filtered with a 20 ms sliding window RMS filter for the first and third sessions, and a 2nd order, 6 Hz cut-off Butterworth filter with zero lag for the second and fourth sessions. A computer program developed with MATLAB V7.0 R14 facilitated efficient onset detection for the experts. The program allowed the experts to zoom-in to any level and to select the desired onset point with the mouse cursor.

The following plot exclusion criteria were put in place: (1) Plots were excluded from analysis if at least one expert considered a particular plot unusable or the onset location uncertain. (2) A VDO value was excluded if it was deemed an outlier when compared to other experts' estimates for a given plot, using Chauvenet's criterion [14].

In order to determine if any onset value was an outlier, the data set was first assumed to be Gaussian. Chauvenet's criterion states that for a Gaussian distribution, a measurement should be discarded if the probability of its observance lies below 1/(2*n), where n is the size of the distribution. Therefore, for the 12 VDO values determined for a given plot, any point with a probability of occurrence less than 1/(2*12)=4.17% is considered an outlier. This probability corresponds to a Z-value threshold of ±2.0368.

After the exclusions, the remaining onsets for each plot were used as a reference for the judgement of algorithm determined onsets. The set of remaining VDOs is referred to as the reference distribution from this point forward.

Algorithm Determined Onsets

Algorithm determined onset (ADO) times were determined by five separate algorithms, implementing a variety of parameter combinations. The initial ranges of parameters used for each algorithm were chosen according to the results of an unpublished pilot study. Each parameter had several possible values and every combination of different parameter values was used. All necessary parameter descriptions are given for each algorithm below and possible values are shown in Tables 1 and 2.

Parameter descriptions

Sliding window: All algorithms implement a sliding window with a width defined by this parameter, which is used to detect a significant rise in activity.

Threshold: The amplitude threshold is defined as the mean of baseline activity plus a multiple of its standard deviation. The multiple is defined by this parameter for algorithms 1 to 4. For algorithm 5, this parameter is used in the maximum likelihood estimation.

2nd Threshold (%): Some algorithms require a certain number of data points within the sliding window to be above an amplitude threshold in order for muscle activity to be detected. The number of points is defined by this parameter.

Peak window (ms): The 3rd algorithm finds the peak signal value within the window with the highest mean value, in order to ignore non-physiological spikes. The window size is defined by this parameter.

Baseline window (ms): This is the size of window used to calculate baseline mean and standard deviation.

Baseline method: This parameter describes the location in the data series from which the baseline activity is taken.

Table 1: Parameters for Algorithms 1-4

Parameter	Alg.	Possible Values			
Sliding Window (ms)	1,2,3,4	25	100	175	250
Threshold	1,2,3,4	1.5	3	4.5	6
2nd threshold (%)	1,3,4	20	40	60	80
Peak Window (ms)	3	150	300	450	600
Baseline Window (ms)	1,2,3,4	150	300	450	600
Baseline Method	1,2,3,4	(1) Earliest window in plot. (2) Latest window in plot. (3) Use baseline window location that yields lowest possible threshold value.			

Table 2: Parameters for Algorithm 5

Parameter	Possible Values			
Sliding Window (ms)	10	40	70	100
Threshold	20	45	70	95
Whitening Filter Order	6	7	8	9

The algorithm descriptions can be found below. However, note that no numerical values are provided for amplitude thresholds, window sizes or other parameters as all possible values can be found in Tables 1 and 2.

1. *Developed by Studenski et al.* [12]

This algorithm used a double threshold approach to find the onset. First, the earliest window with some percentage of data points exceeding an amplitude threshold was found. Next, the leftmost point in the window was chosen to be the onset.

2. Developed by Authors

The latest window with a mean amplitude below an amplitude threshold and before the peak value in the series was determined. The rightmost point in the window with amplitude below the amplitude threshold was chosen to be the onset.

3. Developed by Authors

Similar to (2), but with two differences. First, the plot peak was chosen from the window with the highest mean value in the signal. Second, in addition to the mean of the window being below some threshold, it was required that a percentage of data points were also below the amplitude threshold.

4. *Developed by Bonato et al.* [13]

Similar to (1), but with two additional conditions. First, the window mean must also exceed an amplitude threshold. Second, the window was only considered to hold the onset if the subsequent 29 windows satisfied the same amplitude-related criteria.

5. Developed by Staude [11]

The range from which the onset time was selected was the earliest pair of adjacent windows with a log likelihood ratio exceeding a threshold. The onset value was then selected from this range using maximum likelihood estimation.

Algorithms 1 to 4 were implemented on data smoothed with a 20 ms sliding window with 19 ms overlap, whereas algorithm 5 was applied to raw data. The determination of ADOs for all algorithms and parameter combinations was done using MATLAB V7.0 R14. The program would return an error value if a given method could not find an onset.

Data Processing

For a given plot, the value that an algorithm returned was categorized into one of three values. It was either an error value as determined by the algorithm when a onset value should have been found (Type I error), a onset value as determined by the algorithm when an error value should have been found (Type II error), or a correct value to some degree of accuracy.

All outliers and rejected data plots were removed during VDO determination, and so the remaining data plots were all considered to have a distinct onset value. Therefore, when an algorithm could not find an onset value and returned a corresponding error value, that error was classified as a Type I.

A Type II error occurred when an algorithm chose an onset value that was considered incorrect when judged against the reference onset distribution. Essentially, the algorithm did not correctly reject a plot. The corresponding calculations were done similarly to the outlier calculations during the VDO determination, using Chauvenet's Criterion. First, the ADO was temporarily added to the reference distribution and then the sample mean and standard deviation were recalculated. The ADO was then classified as a Type II error if it had a Z value, relative to the temporary reference distribution, that exceeded the threshold calculated from Chauvenet's Criterion.

Algorithm performance was gauged by two factors: (1) Type I & II error counts and (2) the accuracy of correct onset values relative to the reference distribution. A threshold of 10% was set as an acceptable limit of Type I & II errors, and the accuracy was calculated as the root mean square of the Z values of all non-erroneous onset values, as shown in Equation 1.

$$A_a = \sqrt{\frac{1}{N}\sum_{i=1}^{N} Z_{a,i}^2} \qquad (1)$$

Where A_a represents algorithm accuracy, N is the number of total plots used, and $Z_{a,i}$ is the Z value for algorithm a, data series i.

RESULTS

The most accurate results from all possible parameter combinations were calculated for each algorithm as presented in Table 3. Note that after exclusions, 237 plots were used in the analysis and so the 10% error threshold allowed no more than 24 error values. Algorithms 4 and 5 were not able to meet the error value threshold, and so the next best cases were entered into Table 3.

Table 3: Parameter-optimized algorithm results

Algorithm	1	2	3	4	5
RMS Z Value	0.7738	0.9039	0.7992	0.7155	1.33 10
Type I error	2	0	3	3	3
Type II error	21	19	21	22	61
Total error	23	19	24	25	64
Searching Window (ms)	175	100	25	100	100
Threshold	1.5	1.5	1.5	1.5	45
Second Threshold (%)	80	N/A	60	80	N/A
BL Window (ms)	300	600	600	600	N/A
BL Method	1	1	3	1	N/A

Peak Window (ms)	N/A	N/A	150	N/A	N/A
Whitening Filter Order	N/A	N/A	N/A	N/A	7

DISCUSSION

Given the criteria for successful algorithms, the best results were found for algorithm 1. Though algorithm 4 did not meet the specified error value criteria, it did provide significantly more accurate onset estimates and only exceeded the accepted error value threshold by 1 value. The number of error values from the 237 plots were comparable for algorithms 1 to 4: 23, 19, 24, and 25, respectively. The refinements that separated Algorithm 3 from algorithm 2 did significantly improve performance, though at the cost of computational time and more error values. However, for off-line analysis the computational time for all algorithms was considered negligible for the size of data set used. Algorithm 5 showed no advantages over the other four algorithms.

Algorithm 3 presented a successful parameter combination between the third baseline method, which chooses the baseline window location that yields the smallest possible baseline value, and the minimum possible window length. The algorithm's good performance suggests that this combination can be very successful. The optimal threshold values for algorithms 1 to 4 were all found to be the minimum possible threshold value. For algorithms 2 to 4, the small threshold value was matched with the largest possible window length used to detect baseline noise. This suggests that algorithms are often successful when considering a small deviation from a relatively large baseline noise value.

The number of parameter combinations tested was limited by computational time. Though a greater resolution in parameter values would produce more effective parameter combinations, it was not believed that the improvement would yield significantly better algorithm results. The number of algorithms tested was a further limitation of the paper, though a good representation of algorithm accuracy was still provided. Further investigation will take place into the strengths and weaknesses of different algorithms and also the suitability of each algorithm and parameter combination to data with specific characteristics. Some examples are low signal-to-noise ratio, ECG interference, environmental noise, and EMG sensor movement artefacts.

Overall, algorithms 1, 3 and 4 had comparable performance when parameters were optimized. The successful parameter combinations of all algorithms were presented and future developments were discussed. As the source data used was real and not simulated, the results of this study have direct application in clinical research.

ACKNOWLEDGEMENTS

Thanks are extended to Kristina Calder, Dr. Joan Stevenson and Dr. Arthur MacDonald for their time and valuable assistance.

REFERENCES

[1] Dong-Qing Xu, Jing-Xian Li, Youlian Hong, "Effect of regular Tai Chi and jogging exercise on neuromuscular reaction in older people," *Age and Ageing ,* vol. 34, pp. 439-444, 2005.

[2] C.W. Linford et al., "Effects of neuromuscular training on the reaction time and electromechanical delay of the peroneus longus muscle," *Arch Phys Med Rehabil*, vol. 87, pp. 395-401, 2006.

[3] B.K. Barry, G.E. Warman, R.G. Carson, "Age-related differences in rapid muscle activation after of force development training of the elbow flexors," *Exp Brain Res*, vol 162, pp. 122-132, 2005.

[4] J.M. Castellote, J. Valls-Solé, M.T. Sanegre, "Ballistic reactions under different motor sets," *Exp Brain Res*, vol. 158, pp. 35-42, 2004.

[5] A. Buccolieri, L. Avanzino, C. Trompetto, G. Abbruzzese, "Relaxation in distal and proximal arm muscles: a reaction time study," *Clinical Neurophysiology*, vol. 114, pp. 313-318, 2003.

[6] R.W. Simmons, C. Richardson, "The effects of muscle activation on postural stability in diabetes mellitus patients with cutaneous sensory deficit in the foot," *Diabetes Research and Clinical Practice*, vol. 53, pp. 25-32, 2001.

[7] D.J. Oeffinger, R. Shapiro, J. Nyland, D. Pienkowski, D.N.M. Caborn, "Delayed gastrocnemius muscle response to sudden perturbation in rehabilitated patients with anterior cruciate ligament reconstruction," *Knee Surg, Sports Traumatol, Arthrosc*, vol. 9, pp. 19-27, 2001.

[8] S. Ashford, L. De Souza, "A comparison of the timing of muscle activity during sitting down compared to standing up," *Physiotherapy Research International*, vo. 5, pp. 111-128, 2000.

[9] S.J. Shultz et al., "Assessment of neuromuscular response characteristics at the knee following a functional perturbation," *Journal of Electromyography and Kinesiology*, vol. 10, pp. 159-170, 2000.

[10] H.A. Latimer, J.E. Tibone, M.M. Pink, K.J. Mohr, J. Perry, "Shoulder reaction time and muscle-firing patterns in response to an anterior translation force," *J Shoulder Elbow Surg*, vol. 7, pp. 610-615, 1998.

[11] G. Staude, "Precise onset detection of human motor responses using a whitening filter and the log-likelihood-ratio test," *IEEE Transactions on Biomedical Engineering*, vol. 48, pp. 1292-1305, 2001.

[12] P. Hodges, B. Bui, "A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography," *Electroencephalography and Clinical Neurophysiology*, vol. 101, pp. 511-516, 1996.

[13] P. Bonato, T. D'Alessio, M. Knaflitz, "A statistical method for the measurement of muscle activation intervals from surface myoelectric signal during gait," *IEEE Transactions on Biomedical Engineering*, vol. 45-3, pp. 287-209, 1998.

[14] P.R. Bevington, D.K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 2nd edition, McGraw-Hill, USA, 1992.