Neural muscle activation detection: A deep learning approach using surface electromyography

Iman Akef Khowailed *, Ahmed Abotabl

Doctor of Physical Therapy Program, University of St. Augustine for Health Sciences, San Marcos, CA, USA

A R T I C L E   I N F O

A B S T R A C T

The timing of muscles activation which is a key parameter in determining plenty of medical conditions can be greatly assessed by the surface EMG signal which inherently carries an immense amount of information. Many techniques for measuring muscle activity detection exist in the literature. However, due to the complex nature of the EMG signal as well as the interference from other muscles that is observed during the measurement of the EMG signal, the accuracy of these techniques is compromised. In this paper, we introduce the neural muscle activation detection (NMAD) framework that detects the muscle activation based on deep learning. The main motivation behind using deep learning is to allow the neural network to detect based on the appropriate signal features instead of depending on certain assumptions. Not only the presented approach significantly improves the accuracy of timing detection, but because of the training nature, it can adapt to operate under different levels of interference and signal-to-noise ratio.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Surface Electromyography (sEMG) has been widely used to evaluate the neuromuscular function and muscle recruitment patterns with acceptable reliability (Farina et al., 2004). Temporal analysis of muscle activation in terms of burst onset, duration and offset (Solnik et al., 2008) is one important example that can explain unique muscle activation patterns pertaining to stroke (Babyar et al., 2007), Parkinson's disease (Kumru et al., 2004) and cerebral palsy (Tedroff et al., 2008). For example, the temporal analysis of muscle activation patterns recorded during gait provide insights into the changes in neuromuscular strategies that occur with age (Hortobágyi and DeVita, 2000). Onset detection is also a signification factor in the determination of the electromechanical delay under different conditions (Howatson et al., 2009).

The onset timing of muscular activity is a particularly interesting variable to describe control strategies during different activities. Appropriate muscle activation prior to ground contact during gait increases the sensitivity of the muscle spindles and could contribute to proper joint stiffness and stabilization during different sport activities (Klyne et al., 2012).

Muscle activity timing can be detected via sEMG signal by detecting the first motor units action potentials that build up constantly before touchdown. Detecting the muscle activation via sEMG signal can be done either via visual inspection or threshold-based methods (Solnik et al., 2008).

The sEMG signal is a result of the superposition of signals that stems from the motor units (MU). The timing of the sEMG signal depends on the timing of the MUs however, several factors play a role in the detected sEMG signal formation. The most obvious factor is the noise that is added to the sEMG which might lead to a large margin of error in case the detected sEMG has a significantly low signal-to-noise ratio (SNR). Another main factor is the crosstalk which is the interference from neighboring muscles (De Luca and Merletti, 1988). Neither the noise nor the crosstalk can be avoided or completely eliminated which makes it even harder to understand the sEMG signal of interest. Other factors that affect the nature of sEMG signal such as measurement inaccuracies also exist, and all these factors contribute to the accuracy of the detected activation timing.

The existing techniques in the literature suffer from two main drawbacks. First, the decision of the muscle activity timing is based on some metric which is not necessarily a sufficient metric. For example, threshold-based methods can rely on the signal energy or amplitude which is only one feature of the sEMG signal and ignoring the many other features can lead to accuracy deterioration. Second, interference from neighboring muscles is usually ignored (treated as noise) even though it has special properties that might lead to more accurate muscle activity timing detection if taken into account.

In this paper, we approach this problem via a deep learning technique. The main idea is that instead of relying on some heuristic features of the sEMG signal to solely determine the timing activity, an artificial neural network can be trained to find the activation

* Corresponding author.

timing. Due to the time-series nature of the sEMG signal, a recurrent neural network (RNN) (Merlo et al., 2003) is a reasonable choice for identifying the timing of the muscle activity. Our results show that a properly trained RNN can achieve significantly higher accuracy than the state-of-art algorithms in the literature. Moreover, the trained neural network can work accurately under a wide range of SNR and/or cross-talk levels.

1.1. Related work

One of the most effective ways in improving the timing detection is the Teager-Kaiser (TK) Energy operator that gives an indication of the energy of the signal, not only based on the signal amplitude but the frequency components as well. The TK operator at a time instant n is given by

$$y(n) = x^2(n) - x(n-1)x(n+1) \tag{1}$$

where $x(n)$ is the raw signal and $y(n)$ is the resulting signal after applying the TK operator. One of the main drawbacks in this approach is its vulnerability to the noise level. At high noise levels, the TK methods can be easily confused between noise and real EMG signal resulting from a muscle activation.

Other methods in the literature rely on transferring the signal into other domains using the wavelet transform (Torrence and Compo, 1998) of the Fourier transform (Bracewell and Bracewell, 1986) and then process the signal in these domains. For example, the continuous wavelet transform was used in (Olmo et al., 2000) to detect the muscle activation using the sEMG signal. However, for these methods to work well, certain assumptions such as an identical signal generated from each motor unit have to be met. Also, the accuracy of these methods is greatly compromised under high crosstalk.

A neural network can naturally include plenty of these techniques a special case in their structure and weights and hence, the performance of a neural network can be at least as good as the best of these methods if it is trained properly. For example, a very simple neural network can perform the same operation of the TK measure. In order to show this, assume the neural network in Fig. 1 with $x(n), x(n+1)$ and $x(n-1)$ as inputs. Assuming a quadratic activation at the nodes of the hidden layer according to $f(x) = x^2$ and a linear activation at the output node, the output of the neural network is then given by

$$y(n) = k_1(w_{11}x(n+1))^2$$

$$+ k_2(w_{22}x(n))^2$$

$$+ k_3(w_{33}x(n-1))^2$$

$$+ k_4(w_{14}x(n+1) + w_{34}x(n-1))^2$$
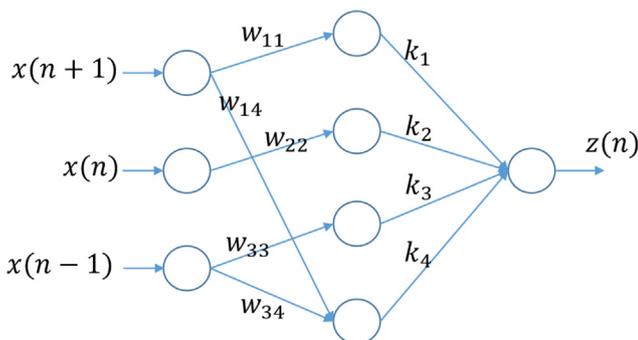
and with the appropriate choice of the weights of the neural network as for example, $k_1 = k_2 = k_3 = 1, k_4 = -1, w_{11} = w_{14} = 2, w_{22} = 1$, and $w_{33} = w_{34} = \frac{1}{4}$, the output will be identical to (1).

A key point in the previous analogy is the use of a quadratic activation unlike the activation function that are usually used in a neural network. However, as shown in (Cheng et al., 2018), an activation function such as the sigmoid function in (4) or the ReLu function in (3) can be approximated with a Taylor series that will eventually perform the same function of a quadratic activation. Therefore, under the usual activation function, the TK measure is still a special case of a neural network.

2. Neural muscle activation detection (NMAD)

In order to be able to train a neural network, an EMG signal and its accurate and correct muscle activity timing should be known. For this purpose, we use a simulated EMG signal whose activation is perfectly known and even designed. The following section gives details about the simulations of the sEMG signal.

2.1. sEMG signal simulation

Several models for the sEMG signal exist in the literature. One of the earliest models that simulate the sEMG signal uses a white Gaussian noise and filter it according to the following transfer function (Stulen and De Luca, 1981)

$$H_{sf}(s) = \frac{ks(s\pi f_h)^2}{(s + 2\pi f_l)(s + 2\pi f_h)^2} \tag{2}$$

where s is the Laplace variable, k is a scaling factor and f_l and f_h are the cutoff frequencies set to 80 Hz and 120 Hz respectively. Even though the frequency parameters of the EMG signal according to this filtering captures many of the signal properties, it still misses many of the sEMG signal features and cannot capture crosstalk properly.

In this paper, we use the more accurate simulated EMG signal that is generated according to the model in (Farina and Merletti, 2001) which simulates synthetic motor units action potentials generated by finite length fibers and then detected by surface electrodes. The model in (Farina and Merletti, 2001) controls the location of the detection probe, the number of fibers of each motor unit, the fiber density in the muscle and more.

We follow the same model parameters described in (Merlo et al., 2003); the detection of the surface EMG signal is simulated by a single differential detection system, with inter-electrode distance of 20 mm and circular electrodes with radius 5 mm. The minimum firing rate is 8 pulses per second and the maximum firing rate is 35 pulses per second. For more details about the parameters of the simulated model, the reader is referred to (Merlo et al., 2003). An example of the simulated sEMG signal with 5 dB SNR is shown in Fig. 2.

2.2. Recurrent neural networks

As mentioned earlier, our approach is based on training an artificial neural network. The choice of a recurrent neural network is because of the following reasons:

- In order to increase the accuracy and avoid the confusion with noise or cross-talk, the muscle activation should be judged by looking at the entire signal and not based on a short snapshot.
- Recurrent neural networks as will be explained next has feedback links that allow learning based on current and previous inputs to the network and hence, allowing the network to decide based on a longer observation. For the same reasons RNNs are used in speech processing (Graves et al., 2013).
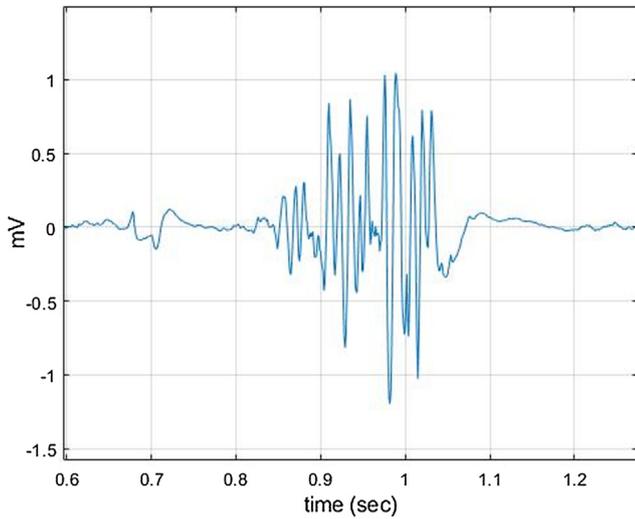

Fig. 1. An example of a neural network that can perform the TK measure.

Fig. 2. Simulated sEMG signal at 8 dB SNR.

A recurrent neural network is defined as a class of artificial neural networks. Unlike the feed-forward neural networks, recurrent neural networks have feed-back links, as shown in Fig. 3, that gives the network a memory which makes it able to process time sequences. The example shown in Fig. 3 shows a recurrent neural network with one hidden layer. The figure shows one step feedback which means that the output at a given time depends on the input at two consecutive time instants. Longer feedback would allow the output to depend on a longer snapshot of the input.

Similar to any neural networks, the nodes of the RNN perform the so-called activation function. In this paper, we used ReLu activation function (LeCun et al., 2015) given in (3) for all the hidden layers and a sigmoid activation function given in (4) for the output layer.

$$f(x) = \max(x, 0) \tag{3}$$
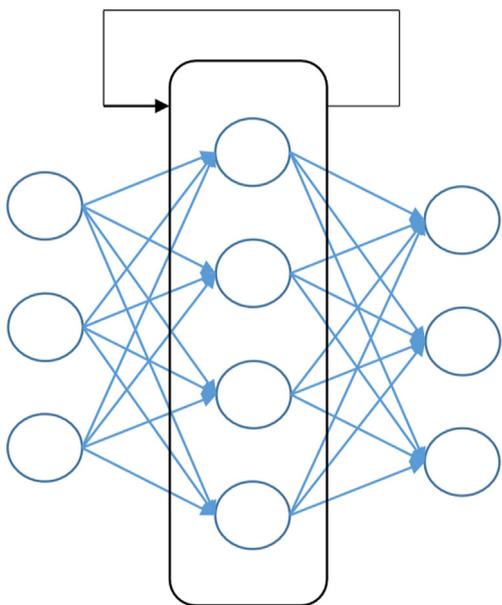
$$f(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

The range of the sigmoid function is [0,1] therefore, we apply a hard limiter that rounds the output to the limits of the range for a final decision on the activation where 1 means that the muscle is active and 0 means that the muscle is not active.

Remark 1. The hard limiter has a zero derivative for the entire domain except at the comparison threshold where the derivative is undefined. Therefore, the error that is used to calculate the gradient is calculated before the hard limiter.

2.3. Activity detection timing

Now that the simulated sEMG signal and the exact activation timing is known, the input to the neural network would play an important role in the performance of the network. In our approach, we use the spectrogram of the sEMG signal since it gives the power spectrum density at each given interval. For example, the spectrogram of the sEMG signal in Fig. 2 is shown in Fig. 4. Clearly, the spectrogram shows higher power especially in the frequency range [10,150] Hz during the time interval [0.85, 1.05]. However, we would like to emphasize that it is not only a matter of power since you can observe high power around the second 0.7 which is clearly due to noise or crosstalk. The main idea as mentioned before is to let the network learn the necessary features that determine the exact activation timing.

The training data is collected by simulating several EMG signals with different activation durations with an average of 0.2 s and a standard deviation of 0.1 s as observed in experimental measurements of the Vastus Lateralis muscle during running. This duration can be different for different muscles therefore, training can be performed via certain characteristics of a certain muscle and in this case, the neural network would be more suitable for the muscle of interest. However, in this paper, we examine training with different activation durations to accommodate different muscles. The neural network is updated with a batch of 100 simulated sEMG signals. Each signal is in the batch represent 2 s and the start of the sEMG signal is chosen randomly between second 0 and second 0.3.

As discussed earlier, the noise level would play a role in detection accuracy. Therefore, we examined the training under different SNR values. Interestingly, our findings show that the best training results are observed under a moderate SNR value rather than a large SNR value. The intuition behind this observation is that the network needs to learn the features of the sEMG signal, but it also needs to learn how the noise structure looks like when it is added to the sEMG signal. A very low SNR or a very high SNR value would not let the neural network learn either the sEMG signal structure or
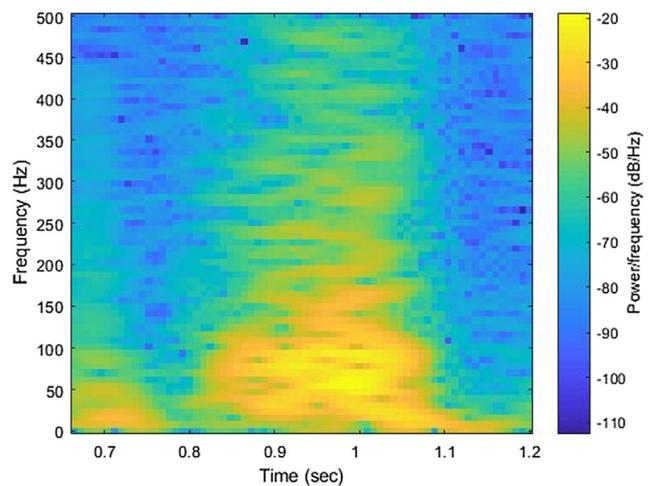

Fig. 3. An example of a recurrent neural network with one hidden layer.

Fig. 4. Spectrogram of sEMG signal of Fig. 2.

the noise effect respectively. Our training results shows that training under an SNR of 3 dB gives the best performance even under different SNR values.

Remark 2. The output of the neural network is 1 when the muscle is active and is 0 when the muscle is not active. This can be used to determine the onset which is the moment the output changes from 0 to 1. The activation duration and the offset are detected similarly.

3. Results and comparison with other techniques

Using the same simulated sEMG signals, different techniques are compared with our proposed method. Based on the TK measure, the single and double threshold methods are considered in the comparison as well as the more statistically involved approach, the generalized likelihood ratio (GLR) (Staude, 2001).

The comparison results are presented in terms of the error in the onset timing detection. The detected onset time can be either earlier, same or after the true onset timing. Therefore, the absolute difference between the detected onset time and the true onset time is a reasonable error metric to assess the performance of different detection methods. As shown in Fig. 5, the accuracy of the neural network muscle activity detection is much higher than the other approaches. The figures also show that the performance of the neural network is not vulnerable to the noise level.

Fig. 6 on the other hand shows a simulated sEMG signal at 5 dB SNR along with the true activation indication and the detected activation indication. The figure shows that the proposed NMAD method follows, almost perfectly, the true activation timing.

3.1. Activation detection of real sEMG signals

In this section, we use the trained RNN in the detection of the muscle activity based on real measured sEMG signal from the Vastus Lateralis as well as the Lateral Hamstring of a runner. The purpose of this comparison is to validate the accuracy of our method with real measure sEMG signals even though the training was performed over simulated signals. All procedures and protocols were approved by the Institutional Review Board of Loma Linda University. Before generating the spectrogram of the measured sEMG signal, the signal is first filtered via a bandpass filter with passing band [10.500] Hz to filter out any noise power outside of the frequency range of the sEMG signal. Consequently, the spectrogram
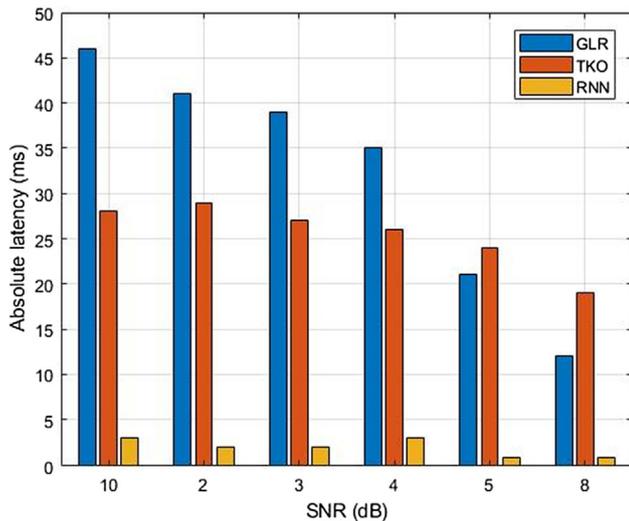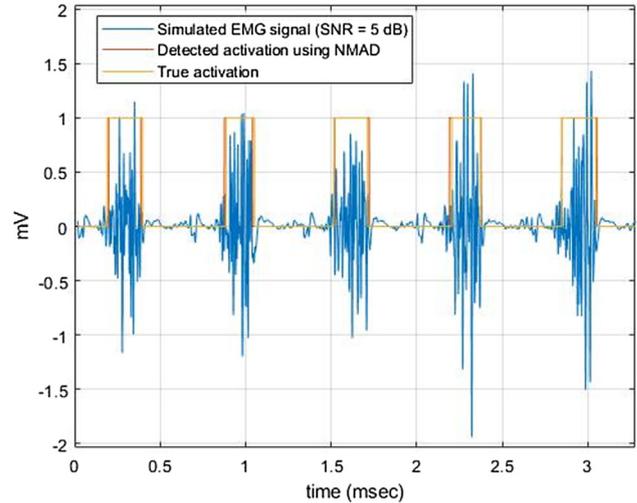

Fig. 6. Simulated sEMG signal, detected activation via NMAD and the true activation.

of the spectrogram of the filtered signal is generated and passed to the RNN as an input. Fig. 7 shows the measured sEMG signal after filtering along with the detected activation according to the RNN output. As shown in the figure, the network is able to accurately detect the activation even for real sEMG signals.

The timing of muscle activity is a commonly applied analytic method to understand how the nervous system controls movement. Detection of the moment when a muscle begins to activate on the basis of EMG signal is an important task for a number of clinical biomechanical studies. In biomechanics, the off-line analysis of EMG is used to add a physiologic context to observed patterns of movement or specific events during movement, such as heel-strike in walking. During a defined movement, the EMG activation from different muscles may be compared if theory dictates that differential activation may cause a predisposing factor towards injury. Thus, this paper introduced a high accuracy onset neural muscle activation detection (NMAD) approach based on deep neural networks. The main idea behind using neural networks is to allow the network to learn the appropriate features that can be used in the detection process instead of relying on some metric that can be statistically insufficient. The presented approach provides a significantly higher detection accuracy under different
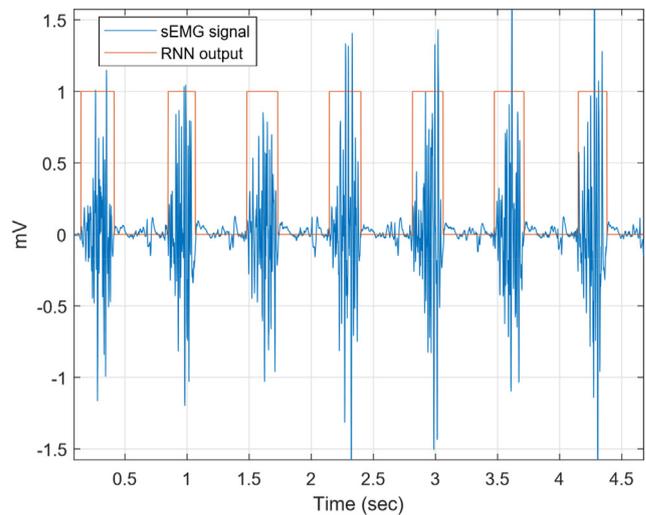

Fig. 5. Performance comparison between NMAD and other methods in the literature.

Fig. 7. Vastus Lateralis sEMG signal and the RNN output.

ranges of SNR and interference levels. Compared to other techniques in the literature such as threshold detection of the TK measure of the signal, the proposed approach requires more computations however, these computations can be easily parallelized. Enabling real time muscle activity detection. The higher accuracy that the presented approach provides not only enables better clinical assessment but also provides researchers with an accurate activation timing that can significantly affect future clinical interventions and preventive strategies.

Declaration of Competing Interest

The authors declare that there is no conflict of interest.

References

Babyar, S.R., McCloskey, K.H., Reding, M., 2007. Surface electromyography of lumbar paraspinal muscles during seated passive tilting of patients with lateropulsion following stroke. Neurorehabilitation and neural repair 21 (2), 127–136.

Bracewell, R.N., Bracewell, R.N., 1986. The Fourier transform and its applications, vol. 31999. McGraw-Hill, New York.

Cheng, X., Khomtchouk, B., Matloff, N., Mohanty, P., 2018. Polynomial regression as an alternative to neural nets. arXiv preprint arXiv:1806.06850.

De Luca, C.J., Merletti, R., 1988. Surface myoelectric signal cross-talk among muscles of the leg. Electroencephalogr. Clin. Neurophysiol. 69 (6), 568–575.

Farina, D., Merletti, R., 2001. A novel approach for precise simulation of the EMG signal detected by surface electrodes. IEEE Trans. Biomed. Eng. 48 (6), 637 646.

Farina, D., Merletti, R., Enoka, R.M., 2004. The extraction of neural strategies from the surface EMG. J. Appl. Physiol. 96 (4), 1486–1495.

Graves, A., Mohamed, A.-R., Hinton, G., 2013. Speech recognition with deep recurrent neural networks. In: In IEEE international conference on acoustics, speech and signal processing, pp. 6645–6649.

Hortobagyi, T., DeVita, P., 2000. Muscle pre-and coactivity during downward stepping are associated with leg stiffness in aging. J. Electromyogr. Kinesiol. 10 (2), 117–126.

Howatson, G., Glaister, M., Brouner, J., Van Someren, K.A., 2009. The reliability of electromechanical delay and torque during isometric and concentric isokinetic contractions. J. Electromyogr. Kinesiol. 19 (5), 975–979.

Klyne, D.M., Keays, S.L., Bullock-Saxton, J.E., Newcombe, P.A., 2012. The effect of anterior cruciate ligament rupture on the timing and amplitude of gastrocnemius muscle activation: a study of alterations in EMG measures and their relationship to knee jointstability. J. Electromyogr. Kinesiol. 22 (3), 446–455.

Kumru, H., Summerfield, C., Valldeoriola, F., Valls-Sole, J., 2004. Effects of subthalamic nucleus stimulation on characteristics of EMG activity underlying reaction time in Parkinson's disease. Mov. Disord. 19 (1), 94–100.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436.

Merlo, A., Farina, D., Merletti, R., 2003. A fast and reliable technique for muscle activity detection from surface EMG signals. IEEE Trans. Biomed. Eng. 50 (3), 316–323.

Olmo, G., Laterza, F., Presti, L.L., 2000. Matched wavelet approach in stretching analysis of electrically evoked surface EMG signal. Signal Process. 80 (4), 671–684.

Solnik, S., DeVita, P., Rider, P., Long, B., Hortobagyi, T., 2008. Teager–kaiser operator improves the accuracy of EMG onset detection independent of signal-to-noise ratio. Acta Bioeng. Biomech./Wroclaw Univ. Technol. 10 (2), 65.

Staude, G.H., 2001. Precise onset detection of human motor responses using a whitening filter and the log-likelihood-ratio test. IEEE Trans. Biomed. Eng. 48 (11), 1292–1305.

Stulen, F.B., De Luca, C.J., 1981. Frequency parameters of the myoelectric signal as a measure of muscle conduction velocity. IEEE Trans. Biomed. Eng. 7, 515–523.

Tedroff, K., Knutson, L.M., Soderberg, G.L., 2008. Co-activity during maximum voluntary contraction: a study of four lower-extremity muscles in children with and without cerebral palsy. Dev. Med. Child Neurol. 50 (5), 377–381.

Torrence, C., Compo, G.P., 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc. 79 (1), 61–78.