On Artefact Elimination in High Density Electromyograms by Independent Component Analysis

Aljaž Frančič
University of Maribor,
Faculty of Electrical
Engineering and Computer Science,
Koroška cesta 46, 2000 Maribor, Slovenia
aljaz.francic@um.si

Aleš Holobar
University of Maribor,
Faculty of Electrical
Engineering and Computer Science,
Koroška cesta 46, 2000 Maribor, Slovenia
ales.holobar@um.si

Milan Zorman
University of Maribor,
Faculty of Electrical
Engineering and Computer Science,
Koroška cesta 46, 2000 Maribor, Slovenia
milan.zorman@um.si

Abstract

We propose a novel approach to artefact detection and elimination in high density electromyograms by using previously introduced Activity Index and Independent Component Analysis (ICA). 28 electromyographic recordings of the biceps brachii muscle were analysed for the presence of artefacts. Using the technique presented in this study, we eliminated an average of 1.07 ± 1.18 artefacts per HDEMG recording. The mean number of eliminated artefacs per recording with at least one detected artefact was 1.88 ± 0.96. In the HDEMGs used in our study, each artefact was found in a separate ICA component.

Keywords biomedical signal processing, electromyography, blind source separation, independent component analysis, activity index, artefact

1 Introduction

In humans, movement and locomotion is regulated by muscles. An electrical signal travels from the central nervous system towards muscles, where it is electrically amplified. By using non-invasive surface electromyography, it is possible to detect the subtle changes in the voltage on the surface of the skin that originate from the muscles, even through the skin and the subcutaneous fat layers. Such recordings are called surface electromyograms (SEMG). When an array of tens of electrodes is used, we call the resulting recordings high density electromyograms (HDEMGs). Due to many factors involved in recording of HDEMGs that are often impossible to control for, the resulting recordings are prone to contain artefacts from various sources, such as power line interference, inadequate electrode-skin contact, electrode drift, subpar quality of the equipment, movement artefacts, etc.

A motor unit (MU) is made up of a motor neuron and the skeletal muscle fibers innervated by that motor neuron's axonal terminals. Groups of motor units often work together to coordinate the contractions of a single muscle. We can think of the signal that travels from the motor neuron to the skeletal muscle fiber as a time series of zeroes, which represent no firing, and ones, which represent the firing of a muscle fiber and, thus, the MU. Because of the refractory period, the spikes are few and far between, making the resulting signal sparse in time. This signal is called the MU spike train.

MUs fire asynchronously and their contributions are superimposed into HDEMG, forming a highly complex signals that are difficult to interpret. By using computer-aided methods, such as Convolution Kernel Compensation (CKC) [2], it is possible to decompose the HDEMGs into contributions of individual MUs and, therefore, identify the firing times of MUs. This gives us insight into the status of the motor system in humans.

However, due to the reasons mentioned in the initial paragraph of this section, HDEMGs often contain various artifacts. These artefacts hinder the ability of CKC to decompose the HDEMG into contributions of individual motor units. Also worth noting: a single artefact might sometimes be present in multiple EMG channels at the same time, so eliminating a single HDEMG channel (for a period of time) is not always effective.

In this study, we exemplified our technique for artefact detection and elimination in HDEMG by using previously introduced Activity Index [1] and Independent Component Analysis (ICA) techniques [5].

2 Materials and Methods

2.1 HDEMG model and Activity Index

In isometric contractions of skeletal muscles, HDEMG signals can be modeled by the following convolutive model [3]:

$$\boldsymbol{y}(n) = \mathbf{H}t(n) + \boldsymbol{\omega}(n) \qquad (1)$$

where

$$\boldsymbol{y}(n) = [y_1(n) \,...\, y_1(n-F+1) \,...\, y_M(n-F+1)]^T \quad (2)$$

stands for time-wise extended vector of HDEMG signals, with extension factor F set between 10 and 20 [3],

$$\boldsymbol{\omega}(n) = [\omega_1(n) \,...\, \omega_M(n-F+1)]^T \quad (3)$$

is the time-wise extended noise vector, and

$$\mathbf{t}(n) = [t_1(n) \,...\, t_1(n-L-F+1) \,...\, t_J(n-L-F+1)]^T \quad (4)$$

is similarly extended vector of MU spike trains. Here, the spike train of the j-th MU is defined as

$$t_j(n) = \sum_k \delta(n - \tau_j(k)), \quad j = 1, \,...\, ,J \quad (5)$$

where $\delta()$ denotes the Delta function and the k-th spike of the j-th MU appears at time $\tau_j(k)$.

The mixing matrix $\mathbf{H}$ comprises L samples long MUAPs of J active MUs, as detected by M uptake electrodes [3].

The Convolution Kernel Compensation (CKC) method [3] estimates the MU filter iteratively as

$$\hat{\mathbf{f}}_j = \hat{\mathbf{f}}_j + \alpha E(g(\hat{t}_j(n))\boldsymbol{y}(n))^T \mathbf{C}_{\boldsymbol{y}}^{-1} \quad (6)$$

$$\hat{\mathbf{f}}_j = \frac{\hat{\mathbf{f}}_j}{\left\| \hat{\mathbf{f}}_j \right\|} \quad (7)$$

where α determines the speed of convergence, $E()$ stands for mathematical expectation, $g(t)$ is a nonlinear weighting function, e. g. $g(t) = log(1 + t^2)$ and $\mathbf{C}_{\boldsymbol{y}} = E(\boldsymbol{y}(n)\boldsymbol{y}^T(n))$ represents the correlation matrix of extended HDEMG measurements. After each iteration of Eqs. 6 and 7, the estimate of MU spike train gets updated

$$\hat{t}_j(n) = \hat{\mathbf{f}}_j^T \boldsymbol{y}(n) \quad (8)$$

The Activity Index I_A at a given time n is defined as [3]

$$I_A(n) = \boldsymbol{y}^T(n)\mathbf{C}_{\boldsymbol{y}}^{-1}\boldsymbol{y}(n) \quad (9)$$

Calculating the Activity Index is the first step of the CKC decomposition approach [3] and is susceptible to HDEMG artefacts. Thus, it can also be used to detect them before further processing. If we do not eliminate the artefacts in HDEMG, they can cause problems for decomposition, preventing the convergence of MU filter in Eq. 6 or segmentation of MU firing moments from identified MU spike train $\hat{t}_j(n)$.

2.2 Independent Component Analysis

In signal processing, ICA is a method for separating a multivariate signal into its additive subcomponents. This is done by assuming that the subcomponents are non-Gaussian signals and that they are statistically independent from each other [4]. ICA is a special case of blind source separation.

In our case, the fastICA [5] decomposition was applied to HDEMG, yielding the individual independent components. Deflation was chosen as the decorrelation approach in fastICA decomposition, and the nonlinearity $g(u) = u^3$ was selected in the fixed-point algorithm. The stopping criterion ϵ was set to 0.0001 and the maximum number of iterations was set to 1000 (the default values).
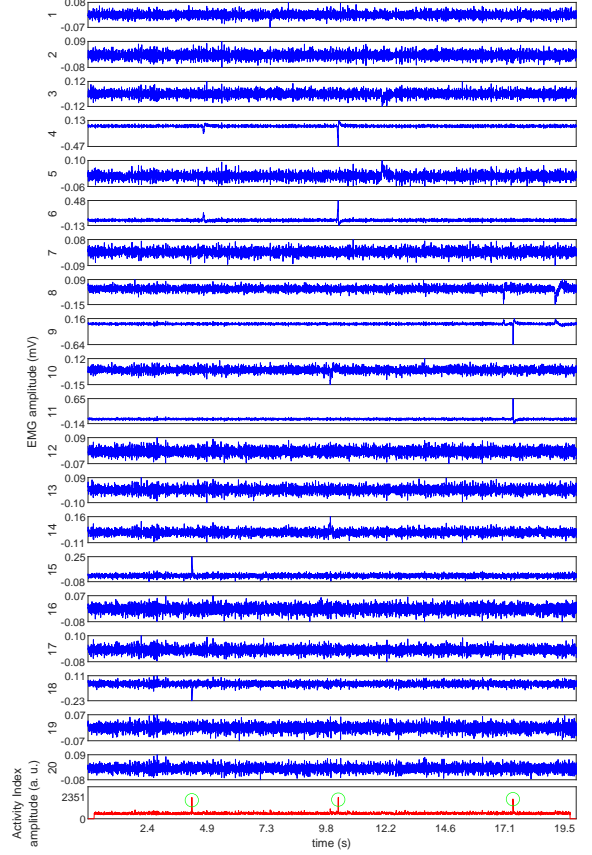

Figure 1: Representative HDEMG channels recorded during the isometric contraction of biceps brachii at 10 % MVC contraction (in blue) and the corresponding Activity Index (in red), with the detected outliers encircled in green. The two most prominent artefacts are noticeable in the 4th and 6th EMG channel at approximately 10th second, in the 9th and 11th EMG channel at approximately 17th second and in the 15th and 18th EMG channel at approximately 4th second. A single artefact is often present in several EMG channels. Also evident, not all the HDEMG artefacts were detrimental for Activity index calculation and, therefore, for MU identification.

2.3 Artefact detection and elimination

An example of a few representative HDEMG channels recorded during isometric contraction of biceps brachii muscle at 10 % of maximum voluntary contraction (MVC) is shown in Fig. 1.

The artefact detection process began by calculating the Activity Index (red line in Fig 1) from HDEMG. In our study, the extension factor F was set to 1. This yielded one time series from several HDEMG channels.

Next, the outliers (denoted by green circles on the red line in Fig 1) were found in the Activity Index. In our case, an outlier was defined as an element that was more than 15 scaled median absolute deviations (MAD) away from the local median. Noteworthy, this could be parameterized and the outliers could be determined in some other fashion. The local median was defined

as the median inside the window, where the window size was equal to 2048 samples. To the best of our knowledge there is no established method for artefact elimination using the Activity Index, hence the MAD threshold for determining the outliers in the Activity Index was selected empirically using visual inspection of the results by an expert. As an additional step, we also ignored any outlier that was fewer than 200 samples away from an already detected outlier. In this way, we prevented the multiple identifications of the same artefact.

Afterwards, fastICA decomposition was applied to the HDEMG, yielding the individual independent components. All the identified components were taken into consideration, as we would like to preserve as much of the information as possible after artefact elimination. Each independent component was then excluded and a new Activity Index I_{EX} was calculated without the excluded independent component. Noteworthy, calculating the Activity Index from either HDEMG or all it's ICA components will always yield the same result. Then, for each detected outlier in the Activity Index I_A, we tried to identify the ICA component that contributed the most to the outlier in the Activity Index (to the HDEMG artefact) by observing the difference in the Activity Index at the time of the outlier before and after the exclusion of the individual ICA component. We did this by using the interest metric defined as:

$$IntMet_{EX}(n) = 1 - I_{EX}(n)/I_A(n) \qquad (10)$$

at any given time n. In partucular, for each detected outlier at time x in the old Activity Index $I_A(x)$, we looked at all the new Activity Indices $I_{EX}(x)$ (with individual ICA components excluded) and calculated the interest metrics $IntMet_{EX}(x)$ at the time x of the outlier in the Activity Index $I_A(x)$.

The interest metric $IntMet_{EX}$ helped to expose the relationship between the old Activity Index I_A and the new Activity Indices I_{EX}. Higher value of the interest metric corresponded to a greater chance that there was an artefact in the excluded ICA component. Using this metric it was possible to determine which ICA component contains the artefact. It was assumed that only a single ICA component will contain a single artefact, as ICA works under the assumption that the sources are statistically independent from each other. If we were to see artefacts in multiple ICA components at the same time, this would imply that they came from statistically independent sources. This would imply artefact co-occurrence, which is unlikely given our findings about the number of outliers and artefacts found per recording (Section 3). Interest metric threshold, above which we considered the artefacts to be successfully eliminated was set to 0.5.

By using the ICA algorithm it would also possible to reconstruct the original recordings from ICA components. This would allow us to first transform the HDEMGs to ICA space, eliminate the artefacts to the best of our ability and then transform the ICA components back to HDEMG space using simple matrix multiplication. In our current study, we eliminated the whole ICA component, but it would be worth considering "repairing" that component (e.g. locally setting the component elements to 0).

2.4 Dataset and evaluation

To evaluate our method for artefact detection and elimination, we used 20 second long HDEMG recordings from 7 neurologically intact young subjects performing isometric contractions of the biceps brachii muscle, at 5, 10, 15 and 20 % of MVC for a total of 28 HDEMG recordings. 13×5 electrode array was used. Visual feedback on force was provided to the participants. All the experiments were conducted in accordance with the Declaration of Helsinki, and were approved by the local Ethical Committee.

In Section 3 we reported the mean ± the Standard Deviation (SD) of the number of outliers found in the Activity Index per HDEMG recording, the mean number of eliminated artefacts per HDEMG recording, the mean number of outliers per HDEMG, where we identified at least one outlier, the mean number of eliminated artefacts per HDEMG where we eliminated at least one artefact as well as the mean interest metric $IntMet_{EX}$ of the eliminated artefacts. We also provided a visual example of an elimination of an artefact.

3 Results

The mean number of outliers found in the Activity Index per HDEMG recording was equal to 1.25 ± 1.29. The mean number of eliminated artefacts per HDEMG recording was equal to 1.07 ± 1.18. The mean number of outliers in Activity Index per HDEMG with at least one outlier was 2.06 ± 1.03 and the mean number of eliminated artefacs per HDEMG with at least one eliminated artefact present was 1.88 ± 0.96. The mean interest metric $IntMet_{EX}$ of the eliminated artefacts was 0.85 ± 0.11. In the HDEMGs used in our study, each artefact was found in a separate ICA component. A representative example of our results is provided in Fig. 2.

4 Discussion

Our results indicated that it was possible to eliminate artefacts in HDEMG using Activity Index and ICA. However, it was difficult to accurately quantify the efficiency of this approach, as we did not know the ground truth about the artefacts' locations in time, nor the HDEMG channels where they were present. We could simulate certain kinds of artefacts at known times and in known HDEMG channels. However this would only account for certain types of artefacts. For example, we could induce an artefact by touching certain electrodes during recording at a predefined time, or during the whole recording by incorrectly applying the contact gel. But we would still be left with other artefacts that we have little control over. Moreover, not all the artefacts have a significant impact on the Activity Index and on MU identification. By observing the Activity Index and comparing it to the HDEMG channels, it is quite clear, that certain artefacts are more detrimental for the Activity Index than others. Therefore, without using the Activity Index or a similar metric, the assessment of artefact impact on EMG decomposition is often difficult.
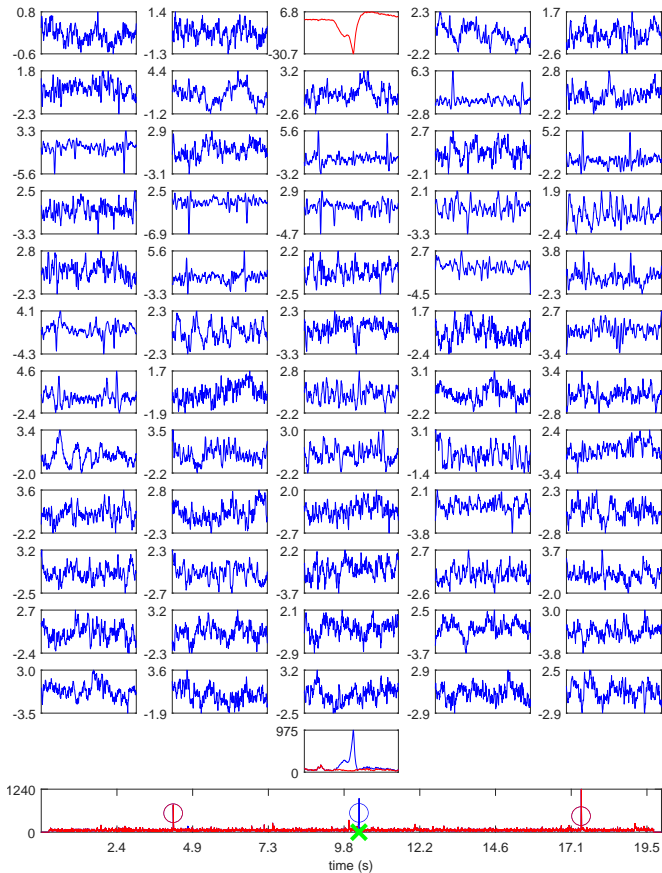
Figure 2: Artefact detection and elimination in ICA components of recorded HDEMG. The bottom panel depicts the I_A of all the ICA components (in blue) and the I_{EX} without the excluded component (in red). The green cross denotes the detected artefact time. The blue and red circles represent the detected outliers in I_A and I_{EX}, respectively. The second row from the bottom shows the 200 samples of Activity Index before and after the detected artefact time (green cross from the bottom panel). The blue line shows the I_A and the red line shows I_{EX}. The first dozen rows show the ICA components 200 samples before and after the detected artefact time (the green cross). The excluded component is depicted in red. It contains an artefact with an interest metric value of 0.96. All y-axis units are arbitrary.

In our method, determining the outliers in the Activity Index depend upon several parameters. The first one was the extension factor F, which was set to 1 for the purposes of this study. Also important was the scaled median absolute deviation (MAD) threshold, above which we considered an element of the Activity Index to be an outlier. In our case, we set it to 15. Lowering this threshold would yield more outliers in the Activty Index. Another parameter was the sliding window size in the Activity Index outlier detection, which in our case was set to 2048 samples.

The presented outlier location detection could also be refined, as using the currently described technique does not guarantee the identified outlier location to be at the location of the actual outlier peak. Instead, we aimed at identifying the first of the 200 samples that is at least 15 MAD away from the local median. This selection could have significant implications in the case of longer artefacts in HDEMG signals.

In our dataset, we found the mean interest metric $IntMet_{EX}$ value of the eliminated artefacts to be 0.85 $\pm$ 0.11, while the threshold, above which we considered the artefacts to be eliminated was set to 0.5. We also found slightly more outliers than actual artefacts as identified by visual inspection of HDEMG signals. This indicated that the current parameters (especially the extension factor for Activity Index calculation F of 1, the MAD threshold for outlier detection of 15 and the interest metric threshold of 0.5) were suitable to identify artefacts in the HDEMGs. However, further tests are required to confirm these findings in other muscles and contraction levels.

Acknowledgment

References

[1] HOLOBAR, A., AND ZAZULA, D. Correlation-based decomposition of surface electromyograms at low contraction forces. *Medical and Biological Engineering and Computing 42*, 4 (2004), 487–495.

[2] HOLOBAR, A., AND ZAZULA, D. Gradient convolution kernel compensation applied to surface electromyograms. In *International Conference on Independent Component Analysis and Signal Separation* (2007), Springer, pp. 617–624.

[3] HOLOBAR, A., AND ZAZULA, D. Multichannel blind source separation using convolution kernel compensation. *IEEE Transactions on Signal Processing 55*, 9 (2007), 4487–4496.

[4] HYVÄRINEN, A. Independent component analysis: recent advances. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 371*, 1984 (2013), 20110534.

[5] HYVÄRINEN, A., AND OJA, E. Independent component analysis: algorithms and applications. *Neural networks 13*, 4-5 (2000), 411–430.