1. INTRODUCTION

A motor unit (MU) is the smallest physiological unit of the human motor system, comprising a motor neuron and its axon, the neuromuscular junction, and the group of muscle fibers innervated by this neuron. Neuromuscular disorders can be roughly divided into neurogenic disorders, which are marked by a loss of functioning neurons, and myogenic diseases, which affect the muscle fibers. Both types have in common that they influence the shape of the action potentials that are generated in a muscle when its MUs are activated (Denny-Brown (1949), Roeleveld et al. (1997), Drost et al. (2006)). These pathophysiological changes tend to be fairly specific for the underlying class of disorders. Therefore, estimating the shapes of the so-called motor unit action potentials (MUAPs) , as obtained by needle or surface electromyography (EMG), is important for the diagnostic investigation and follow-up of such diseases (Daube and Rubin (2009)). Data on the firing patterns of the motor units provide information on the motor control strategies of the central nervous system and, hence, a valuable additional perspective (Dorfman et al. (1989)).

Until recently, the investigation of MU activity has been the sole privilege of intramuscular EMG (iEMG), which uses a needle electrode to pick up MUAPs in the muscle. Over the years, several dedicated algorithms have been developed that allow decomposition of trains of MUAPs from multiple MUs that are recorded with such electrodes. The result of these procedures is a more or less accurate set of MUAP shapes and a list of firing instances. A review of these methods and their limitations has been provided by Stashuk (2001). A major disadvantage of needle EMG is that it is invasive. It can be uncomfortable and sometimes unsuitable for the patient. Moreover, iEMG measurements are difficult to reproduce as the same motor units cannot be followed longitudinally. For follow-up investigations, this implies that MUAP shapes can vary significantly between recordings, even when made from the same, healthy MU.

For these reasons, surface EMG and, in particular, high-density surface EMG (HD-sEMG) recordings have been introduced as noninvasive alternatives. HD-sEMG measures the propagated electrical activity from the muscle on the skin overlying this muscle using a grid of multiple densely spaced electrodes (Blok et al. (2002), Drost et al. (2006)). MUAPs observed on such a grid vary not only between MUs but also from electrode to electrode. These spatial variations within a MUAP are generally sufficiently characteristic for the underlying MUs to allow discrimination between them. Therefore, the MUAP recorded with HD-sEMG is sometimes referred to as “MUAP signature”. The spatial differences between and within HD-sEMG MUAPs are due to anatomical and physiological differences between the MUs (e.g., in their size) as well as to differences induced by spatial filtering from the source (muscle fibers) to the recording site (electrode) (Blok et al. (2005), van Dijk et al. (2008)).

Although the signal-to-noise ratio (SNR) is generally lower for sEMG than for iEMG because of the larger distance between source and electrode, the noninvasiveness of sEMG, its larger field of view and better reproducibility outweigh this disadvantage for

specific purposes (such as recordings in children and monitoring). sEMG has been shown to be able to offer important insights into muscle physiology and anatomy (Wood et al. (2001), Drost et al. (2001), Drost et al. (2004), Lapatki et al. (2010)). However, the decomposition of sEMG signals tends to be much more difficult than that of needle EMG. The low-pass filtering of the tissue between the source and electrode reduces the MUAPs' distinctive features, rendering them more similar and, hence, more difficult for algorithms to differentiate. So far, a limited number of methods are available that can extract the contributions of individual MUs from sEMG recordings (for a review see Merletti et al. (2008)). One of the major openstanding challenges is to disentangle the measured activity when the MUAPs of several simultaneously active units overlap, creating composite signals or mixtures.

When a MU is active, it fires at fairly regular intervals, resulting in a train of repetitive MUAPs. This property supports the assignment of MUAPs to a specific MU. The term full decomposition is used to describe an algorithm that is able to properly detect and classify each firing of all MUs contributing to a recording. If achieved, such a technique provides a complete set of MU firing patterns. To accomplish this, the overlapping mixtures of various MUs have to be "demixed".

Over the past decade, several studies have proposed solutions to this problem, many in the context of needle EMG. In case of sEMG, some of these are able to extract several mean MUAP shapes per recording (e.g. Gazzoni et al. (2004), Xu et al. (2001)) but were not able to deal with mixtures. Others reported full decomposition approaches (e.g. Zazula et al. (2005), De Luca et al. (2006), Kleine et al. (2007), Florestal et al. (2006), Florestal et al. (2009)). However, various problems concerning either automation, processing time, or the number of different MUAP signatures that could be extracted (usually not more than six per recording) remain and ask for new solutions.

A more recent, promising approach for full decomposition is the so-called Convolution Kernel Compensation (CKC), reported by Holobar et al. (2007) and later verified in Holobar el al. (2010). It represents a so-called generalized blind source separation approach for MIMO systems with a finite symbol dictionary. Instead of dealing with the MUAP shapes, the method compensates for these shapes and reconstructs the input pulse trains carrying the information about the symbol (MUAP) rising times directly. Another solution has recently been reported by Nawab et al. (2010), which was later tested by De Luca and Hostage (2010). It builds upon the previously used Precision Decomposition technique (De Luca et al. (2006)). However, the solution is currently only offered as an integrated system of Delsys company.

In this study we introduce a novel approach towards the ultimate aim of full decomposition. Our automated method first extracts MUAP signatures of contributing MUs and then uses these shapes as prior knowledge for subsequent full decomposition. In addition, we designed our method to be computationally inexpensive, robust, easy to use for possible diagnostic applications, and compatible with multiple surface electrode configurations and equipment. The paper is organized as follows. In Section 2, we describe how the mean shapes or signatures for individual MUs are obtained by means of

a clustering approach. We then explain in detail how we resolve superimposed action potentials using a complete dictionary of the obtained mean MU signatures in the optimization problem that we refer to as demixing. We verified our method on both simulated signals and real-life data, as described in Section 3.

2. METHODS

2.1 Algorithm

2.1.1 Outline of the algorithm

The algorithm comprises three major steps (see outline in Fig. 1): detection of action potentials (spikes) (which can either be individual MUAPs or mixtures), clustering of similar spikes, and demixing. The goal of clustering is to group similar MUAPs belonging to active MUs and obtain their mean shapes. The demixing stage provides the time instances at which each MU was active – timestamps. The output of the algorithm is a set of MUAP signatures and their corresponding firing patterns. The second step exploits the fact that the MUAPs come in trains and have a constant, all-or-none signature on the recording grid. These properties allow us to cluster similar signatures generated by active MUs. However, this step leaves many spikes unclustered due to overlapping of MUAPs from different, simultaneously active MUs or because of clustering imperfections. Therefore, in the third step, all of the detected shapes are passed through a demixing process, which assumes that each observation is a superposition of one or more previously obtained shapes and noise. The demixing is performed by minimizing the residue that remains after subtraction of the estimated shapes at proper time lag.

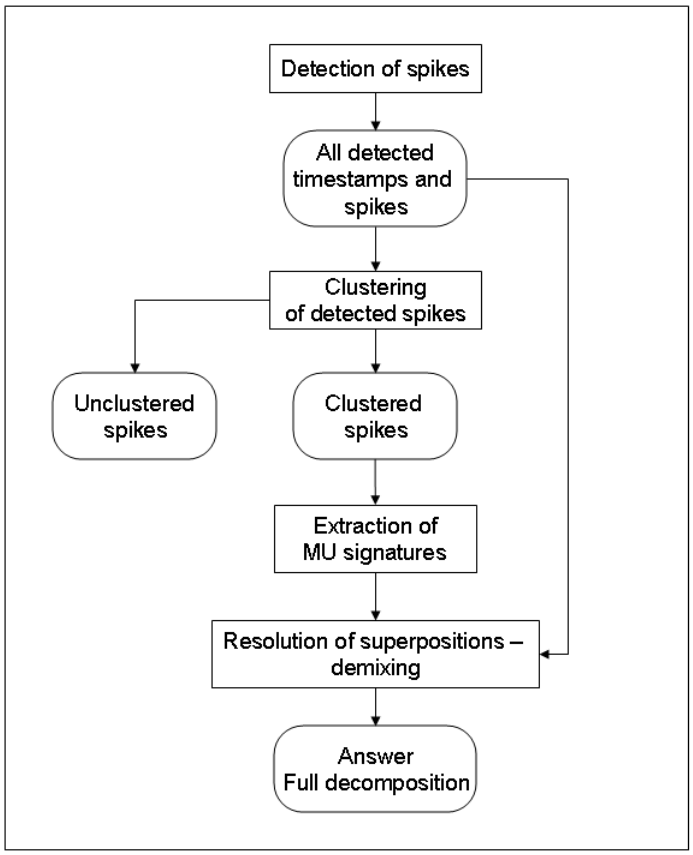


Figure 1: Schematic outline of the algorithm. Three major steps can be recognized: 1) spike detection, 2) clustering of spikes and extraction of MU signatures, 4) demixing the interference pattern by using the MU signatures obtained in step #2.

2.1.2 Detection and clustering of spikes

Spike detection

The HD-sEMG array that we employ contains 126 electrodes, which record the signals relative to a far-away reference (i.e., we have monopolar recordings on 126 channels) (Blok et al. (2002)). Because the electrodes are densely spaced, the information content of the signals from adjacent electrodes overlaps to a large extent. Therefore, using all of the available channels for detection and clustering is time consuming and redundant. Especially during clustering, channels with low-amplitude signals do not contribute to the differentiation between MUAPs and can even deteriorate clustering performance. As the optimal electrodes for detection of individual MUAPs is unknown, we proposed to use scattered electrodes that together optimally “cover” the array. In this way, even small MUAPs should be picked up by at least one nearby electrode. We used a limited set of eight scattered electrodes and their corresponding channels for detection and clustering, both in the simulation study and for the real-life data in this paper. We refer to these channels as “clustering channels”.

For each of the eight clustering channels, the noise level was estimated using the median-based measure introduced by Donoho and Johnston (1994). This measure was chosen because of its robustness in the presence of spikes. Segments of the signal exceeding the noise threshold were defined as spikes. Within the recorded samples of each detected spike, the one with the maximum amplitude was found and the corresponding time instant of this peak (timestamp) was stored. Because a MUAP could be detected on more than one channel simultaneously, timestamps of different channels were compared and multiples removed. To prevent spikes with multiple peaks being wrongly detected as two or more different spikes, detector dead time was introduced as the smallest interspike interval allowed (8 ms in our case).

Input vectors for clustering

For each detected spike, the data from all the clustering channels are used as input for clustering. Initially, 5ms equaling to 22 samples (11 before and after the detected peak time) are stored for each clustering channel. These samples were empirically found to represent most of the discriminatory information of the spike and are least influenced by MUAP summation if multiple MUs are simultaneously active. The eight 22-sample segments from the clustering channels are then concatenated to form a pre-input vector for clustering as shown in Fig. 2a. Our clustering was performed on a wavelet-based model for representing this raw data. The use of wavelets localizes distinctive features of spikes and, hence, increases the discriminatory power of the clustering. We used the Haar wavelet, which is orthogonal and has a form of square function, avoiding prior assumptions on spike shape. The discrete wavelet transform was performed on all the pre-input vectors. Subsequently, Kolmogorov-Smirnov criteria were used as in Quian Quiroga et al. (2004) to find the 44 (25%) most discriminating coefficients to represent the full pre-input vectors. These coefficients were then used as input vectors for clustering.

Clustering

Clustering of the input vectors corresponding to detected spikes is performed with a multichannel extension of the Wave_clus algorithm of Quian Quiroga et al. (2004) as reported in Gligorijević et al. (2011). The set of vectors of wavelet coefficients for all detected spikes is used as input. The algorithm returns both clusters of vectors with very similar coefficients and a set of nonclustered vectors that appear to be outliers. Similar elements that are successfully clustered are assumed to belong to individual MUs while the outliers are mainly mixtures that are not repetitive. The number of clusters is automatically estimated and the clustering is unsupervised. The concept underlying this approach is taken from the statistical mechanics of ferromagnetism; the algorithm employs the hierarchical superparamagnetic clustering method introduced by Blatt et al. (1996). In this concept, data vectors are treated as points in a multidimensional feature space. A measure, called Potts spin, is assigned to each data point and the interaction of neighboring points is subsequently observed. This new “magnetic” system exhibits three phases out of which individual clusters can be identified as ordered groups through their

spin-spin correlation function. Fig. 2 shows an example of the pre-input (raw data) vectors before (Fig. 2a) and after (Fig. 2b, c) clustering.

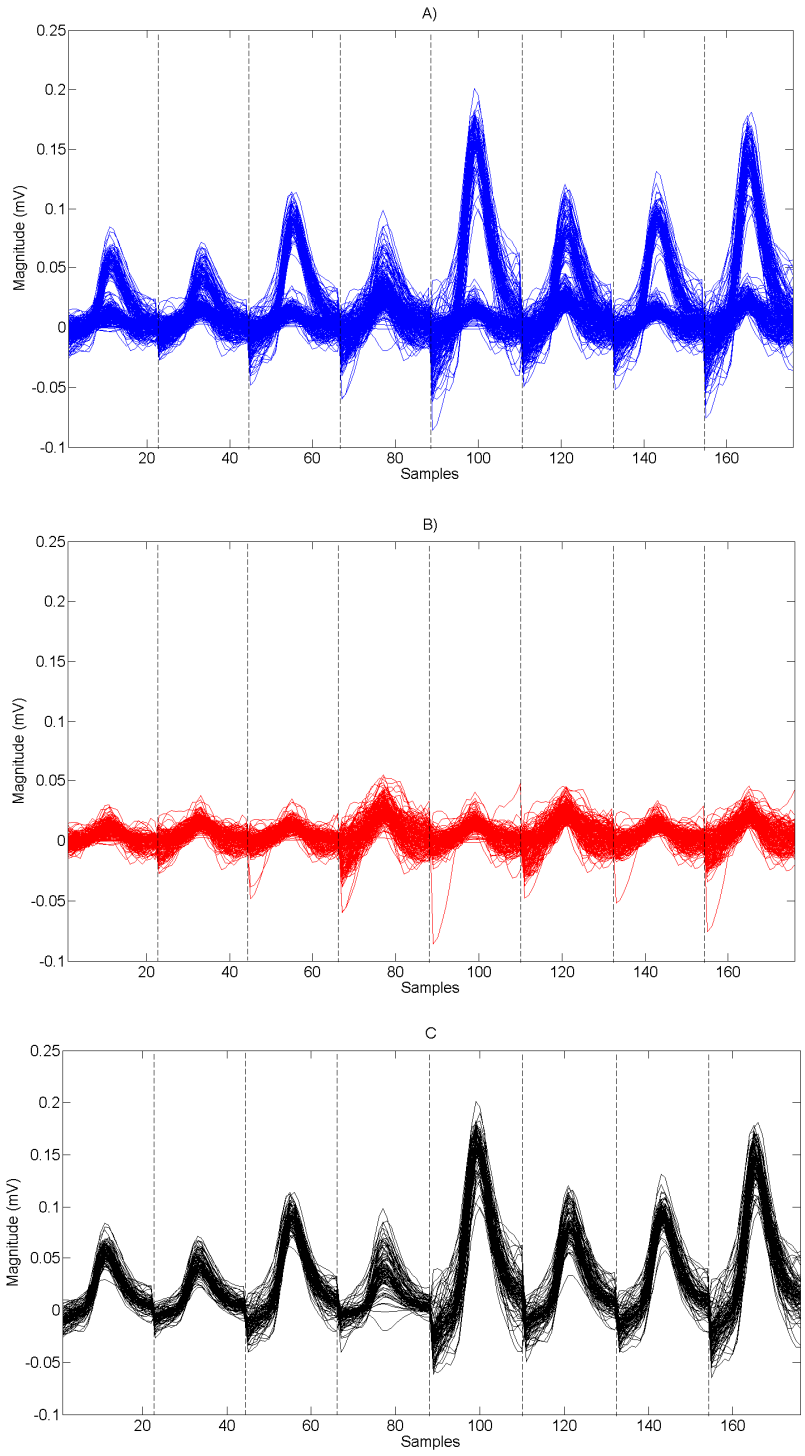


Figure 2: Example of clustering. (A) Overlapped and aligned pre-input vectors based on concatenated spikes from the eight clustering electrodes. (B) and (C) Clustering places the detected spikes accurately into two groups with small intragroup variability and

significantly larger intergroup difference. The vector data in (B) can therefore be assigned reliably to a different MU than those in (C).

Once clusters have been identified, a large number of timestamps corresponding to individual MUs is identified. MUAP signatures are estimated as the median of the recorded temporal channel observations for each MU (Fig. 3). The median was chosen because it is more robust to clustering imperfections than the mean. Next, each signature was stored, with duration of the stored segment equal to a predefined length. Additional de-noising of these template MUAP signatures was obtained by computing the best rank-5 approximation from a singular value decomposition (SVD). These signatures will serve for decomposition of all previously detected spikes. An example of a signature and the associated MUAPs is shown in Fig. 3.

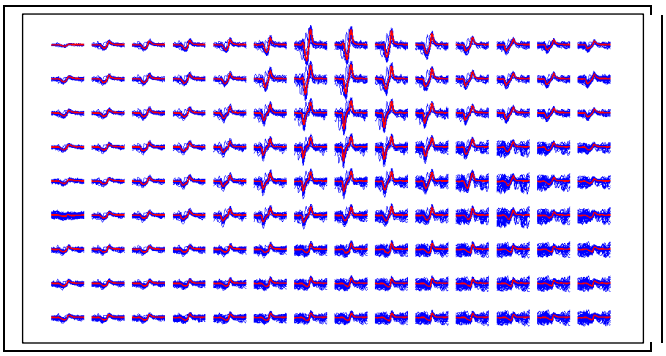


Figure 3: An example of overlaid MUAPs belonging to the same MU (blue) together with the median shape in red (signature).

2.1.3 Demixing of mixtures of MUAPs

After clustering, a number of vectors remained unclustered. Many of those belonged to mixtures of MUAPs generated by more or less simultaneously active MUs. MUAPs that are heavily corrupted by noise could also appear as outliers. With the prior knowledge of the MUAP signatures that were present in the signal, demixing could now be performed by optimizing the fit of estimated shapes to a detected spike. For this purpose, we first need a model how the activity of MUs is reflected on the recording electrodes.

Multiple-input-multiple-output (MIMO) model

The motor neuron firing sequence can be represented as a pulse train of Dirac functions, as described by Holobar and Zazula (2007). Surface EMG recordings can then be seen as the output of a multichannel linear, time-invariant system. The system is defined by the transfer function that occurs between the source of the activity (muscle fiber) through the tissue to the recording electrodes. The output is the result of the convolution between the stream of Dirac impulses and this system. If a single motor neuron is active and the resulting MUAP is recorded with multiple electrodes, the system can be mathematically described with a single-input-multiple-output (SIMO) model. Accordingly, in the case of activity of multiple MUs we have a MIMO system. Since finding the formal solution of this system is computationally very demanding, we propose to reduce this model to a

computationally less expensive and yet accurate, multiple-input-single-output (MISO) problem. The result of this optimization is a new set of timestamps corresponding to each MU that reveals its firing pattern. We will first introduce the convolutive single-input-multiple-output (SIMO) system, which is similar to that of Abed-Meraim et al. (1997), and then show how it extends to MIMO and is finally reduced to MISO.

We represent the identified MUAP shapes by a system (**H**), a set of spatial filters that is linear and shift invariant. The sources (**s**) are Dirac pulses, on-off events of amplitude equal to one, that mark the instances at which the central nervous system triggers the respective MUs. The convolution between the sources and the filters yields the output (**y**), which is corrupted by additive noise (**w**). Given that in our setup **H** is completely determined by the estimated shapes as obtained from the previous clustering steps and that the output is observable, the only problem is to determine the time instances of the Dirac pulses.

Mathematically, we have the following relation between a single input discrete time series $s(k)$ and M discrete output time series $y(k)$:

$$\begin{aligned} y_1(k) &= s(k) * h_1(k) + w_1(k) \\ y_2(k) &= s(k) * h_2(k) + w_2(k) \\ &... \\ y_M(k) &= s(k) * h_M(k) + w_3(k) \end{aligned} \qquad \dots (1)$$

Here, M denotes the number of outputs or recording channels, and $w_i(k)$ represents the noise present in the i-th channel. In matrix form this set of equations becomes:

$$\mathbf{y}=\mathbf{H_M}\mathbf{s}+\mathbf{w}; \qquad \mathbf{H_M}=\begin{bmatrix} \mathbf{H_{(1)}} \\ \mathbf{H_{(2)}} \\ \mathbf{...} \\ \mathbf{H_{(M)}} \end{bmatrix}; \qquad H_{(i)}=\begin{bmatrix} h_i(L) & \dots & h_i(0) & \dots & 0 \\ \vdots & \ddots & & \ddots & \vdots \\ 0 & \dots & h_i(L) & \dots & h_i(0) \end{bmatrix}; \quad \dots (2)$$

Every channel i on the grid has the corresponding filter $H_{(i)}$. If the length of the observed shape is L (number of the filter coefficients), dimensions of the matrix $H_{(i)}$ are $N \times (N+L)$. Here, N is the number corresponding to any possible "shift" that can occur with a shape, so that any single observation can be described. In other words, one shape can "corrupt" another only if they overlap by at least one sample. In the extremes, this may be the last sample of one shape and the first of the other, or the other way around. Therefore, we have to observe a number of samples that is three times that of the spike length. Accordingly, number N has to be chosen up to the value of 2L. Here, $h_{(i)}$ are filter coefficients representing the MUAP shape on the corresponding channel. The optimization problem aims at finding the moment in time at which the output of the system maximally correlates with the observed measurement on this particular channel.

Fig. 4 illustrates this procedure for the recording of a single MUAP with one electrode (single input-single output (SISO) model). The output **s** (proper time lag) is derived by using the filter matrix **H** and minimizing the residue between a fitted model and the observed signal.

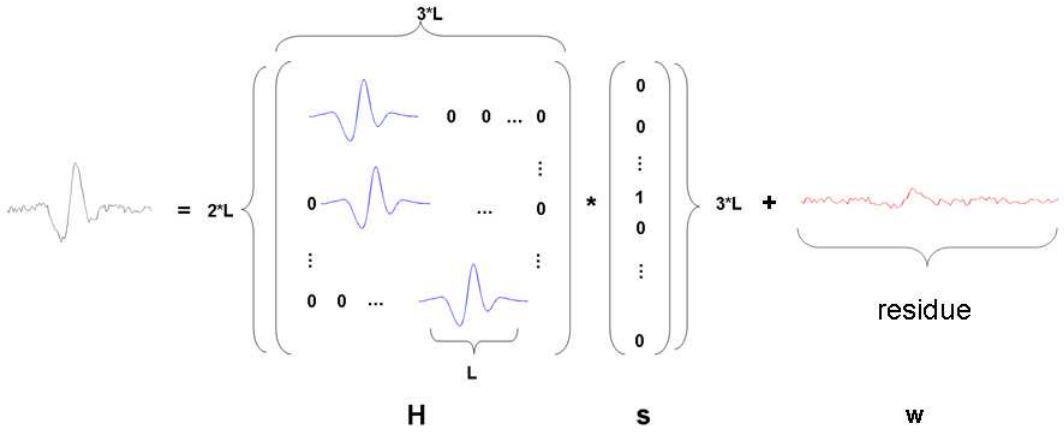


Figure 4: A recorded spike (left) is described as the sum of the previously identified shape at a proper time lag and noise

Mathematically, the SIMO model can easily be extended to a MIMO system, in which the output represents the observations from multiple electrodes, by having many input vectors and corresponding system filters. This means that for each MU *T* and for each electrode *i*, we have a different $H_{(T)i}$ filter matrix. To reduce the huge computational complexity associated with solving this new, large system of equations, we propose to select one electrode to represent each MU. This optimization strategy reduces the system and results in a computationally affordable solution.

Optimization of the demixing algorithm

The single-channel spikes coming from individual MUs that together constitute a HD-sEMG MUAP signature can look almost the same on certain electrodes, while being substantially different on others. Channels with similar outputs do not add much information. Therefore, instead of using information from all the channels, we limit ourselves to only one per MU – the one with the highest signal to noise ratio (SNR). In case more MUs have their maximal SNR on the same channel, its information is used as many times, increasing the weight of this channel. Thus, we use a number of electrodes that equals the number of detected MUs that we will refer to as optimization electrodes. For the selected electrodes, we concatenate the observations. For instance, if we have four units, this step in the analysis yields 4 concatenated template spikes from the selected channels.

As action potentials are conducted over the muscle fiber from the neuromuscular junction towards the tendons, the MUAP may appear with different latencies at different electrodes. Under non fatiguing conditions and constant temperature, these latency differences within the MUAP signature remain fairly constant from firing to firing for one particular MU. This implies that, when "shifting" the signals to determine the minimal residue, the relative peak latency between the used channels remains constant for every MU. Therefore, the concatenated shape corresponding to one MU can only be "shifted" left or right as a whole. This brings us to a single filter $\mathbf{H}_{\mathbf{TM}}$ for each MU T, with size $N \times (N + L_{large})$ where L_{large} is (number of electrodes)*(spike length) samples.

Observations that are subject to the optimization process

Each detected timestamp is treated in the same way: for each of the selected electrodes, one spike length is taken centered on this timestamp and these segments are subsequently concatenated. This concatenated vector constitutes a single observation in the demixing process. Thus, the system can be efficiently reduced to a multiple input-single output (MISO) system. The observed shape may be the result of activity of a single unit (Fig. 5a) or a combination of units (Fig. 5b).

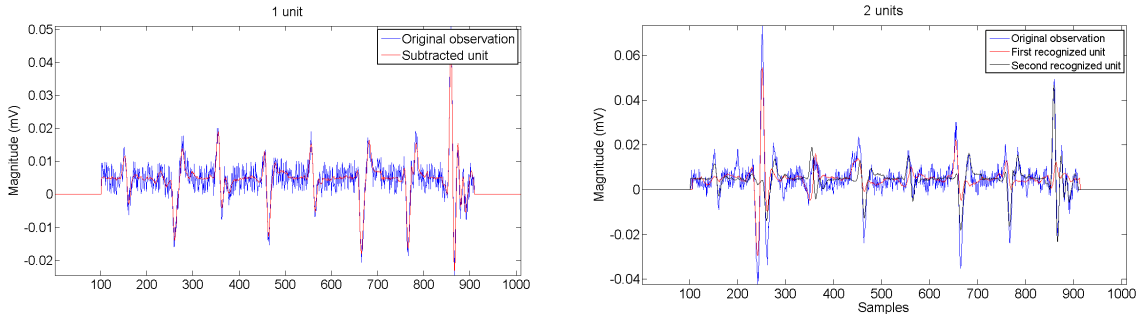


Figure 5: (a) observation of a single unit; (b) superposition of two units

For each detected timestamp, an observation vector, ***Ov*** is first created. We repeat subtracting the MU shapes until the residue drops below a given threshold. The optimization process will try to fit up to 3 simultaneously active units. For each unit, the possibly optimal position is found by finding the maximum of the scalar product between each **H** matrix and the ***Ov***. The corresponding **H** filter defines the proper MU while the column number of global maximum provides the proper time lag for this unit – modified timestamp. Our goal is to minimize the residue vector (***Rv***) obtained by subtraction of a particular observation from the ***Ov*** at the indicated lag. We additionally define the grid residue (***Rg***) as the sum of residues from all electrodes on the grid on the time lag indicated by the smallest ***Rv***. Its purpose is to check whether the overall residue is in the noise level. If this residue is too large after optimization, no successful decomposition is achieved, and we reject the corresponding spike and its timestamp from our analysis. If the residue is lower than 1.5 times the noise estimate on the whole electrode grid, we assume that successful demixing is achieved.

The algorithm scheme for the demixing process is outlined in Algorithm 1:

Algorithm 1

1. Create matrices **H**
2. Calculate residue constant
3. For all detected timestamps:
   A. Make a vector of observation (***Ov***)
      Repeat for combination of up to 3 units:
      i. Multiply each **H** matrix with the vector ***Ov***
      ii. Find the maximum of the product
      iii. Calculate vector residue ***Rv*** when indicated unit is subtracted and corresponding timestamp
      iv. Calculate grid residue ***Rg***
      v. Assign ***Rv*** as the new observation and create new vector ***Ov***
      end
   B. Find the smallest ***Rv*** – this will indicate number of units allowed to participate in the mixture (from 1 to 3)
   C. Check if the corresponding ***Rg*** residue is below the residue constant:
      yes → assign all indicated timestamps to according clusters
      no → leave the initial timestamp (observation) as unclassified one
4. Final decomposition obtained

2.2 Data

2.2.1 Simulated data

The proposed algorithm was validated using MUAPs obtained from HDsEMG recordings from the thenar muscles. MUAP template shapes obtained from Maathuis et al. (2008) were used and combined with simulated firing sequences to generate an artificial sEMG interference pattern. MUAP signatures with mutual spatial correlation value up to 0.95 were randomly chosen from the recorded set for creating the simulation signals.

The template MUAP signatures were recorded as described in Maathuis et al. (2008). A flexible surface array of 126 Ag-AgCl electrodes (9 rows and 14 columns) was used with a electrode diameter of 1.5 mm and the center-to-center electrode distance was 4 mm in both directions (Lapatki 2004).

The simulated spike trains had interspike intervals that followed a Poisson distribution and that ranged from 8 to 24 Hz. Interference patterns based on contributions from 3 to 10 active MUs (in increments of 1 MU) were created. For each set of generated spike trains (and template MUAP signatures) this resulted in 8 different signals – the first containing three and the eighth containing ten different MUAP signatures. The duration of the simulated signals was 20 seconds and the sampling frequency was set to 4096 Hz. Five of these sets were created on the basis of different sets of spike trains and mean MUAP signatures.

To increase the resemblance with real signals, Gaussian noise was added. The noise level was set to one three different values (3, 10, and 15 μV root-mean-square (rms) power) to assess its influence. The resulting signals were low-pass filtered at 1000 Hz with a fourth order Butterworth filter. Each data set was coupled with each of the noise levels so that overall 120 signals were used for evaluation of the algorithm (five repetitions for each of three to ten active MUs and each of the three noise levels).

First, for each MU signature, the electrode with a peak spike value was chosen. Then, the SNR was estimated as the ratio between the power of the spike divided by the noise power on this electrode in dB:

$$SNR = 10 * \log_{10} \frac{P_{spike}}{P_{noise}} \qquad \ldots \qquad (3)$$

Power of spike and noise were estimated by summation of their squared values.

2.2.2 Real-life data

We also validated the algorithm presented in Section 2.1 against the HDsEMG data set that was previously reported by Kleine et al. (2008). This data set consisted of 20-second recordings from the lateral vastus leg muscle of nine healthy subjects. The contraction level was 1% of the maximum voluntary contraction (MVC) in 7 subjects, 2% MVC in

one and 4% MVC for the remaining subject. A major advantage of using this data set was that the 5-s long middle section of all recordings had been manually analyzed by two independent, trained operators in the context of the original study. The inter-operator agreement was very high (in 20 out of 25 identified units it was above 95%, the lowest being 63% on a single unit), implying that the manually detected firing patterns could be used as gold standard for our evaluation.

The main features of the acquisition system have been described in detail elsewhere (Blok et al. (2002); Kleine et al. (2008)). In short, the electrode array consisted of 10×13 electrodes with center-to-center electrode distance of 5 mm in both directions. Sampling frequency was 2000 Hz and the data were band-pass filtered (3-400Hz). After A/D conversion, a high-pass Butterworth filter of order 6 with cutoff frequency of 10 Hz was used to further reduce the baseline drift.

Because SNR values for lateral vastus recordings are lower than those of thenar recordings, both monopolar and bipolarly filtered data were used. Bipolar signals were obtained by subtracting the signals from adjacent electrodes in the muscle fiber direction. Monopolar and bipolar signals were examined visually to identify corrupted channels and channels with high noise. These channels were excluded prior to the analysis. Next, the monopolar recordings were used to have a first, rough go at detecting clusters. On the basis of these initial clusters, the channels were identified in which the corresponding bipolar signatures had their maximum peak value. The motivation for this was the more localized nature – presence on fewer channels – of the bipolar signals. These channels were then used as clustering channels in bipolar representation to detect MUAPs and to perform the clustering (as data-driven alternative to the scattered electrode approach described in 2.1.2). After the clusters were identified in the bipolar signals, the “peak” channel was selected for each cluster as described above, and these sets of channels were used for the demixing process.

2.3 Validation

To verify the quality of our analysis, temporal and spatial matching parameters were introduced. While temporal parameters quantify the accuracy of the clustering and level of matching between exact and extracted values of timestamps, spatial parameters grade the quality of extracted MUAP signatures. Both simulated and real-life signals were treated with the same set of quality measures.

Temporal parameters used were sensitivity and precision. Sensitivity is defined as the ratio of correctly classified compared to all the present timestamps (for the particular MU). The number of correctly classified timestamps divided by the number of timestamps for each MU is defined as precision.

Spatial quality was estimated using 2 parameters: normalized root mean square error (NRMSE) and correlation coefficient. Each of the signatures obtained by clustering is aligned and compared with the original shape using these two measures. NRMSE is calculated using formula (4), where $\hat{x}$ represents the estimated shapes obtained by

clustering, and x the template (exact) forms. Summation is performed over all samples and channels.

$$NRMSE = \sqrt{\frac{\sum(\hat{x} - x)^2}{\sum x^2}} \quad \ldots (4)$$

In case of the simulated signals, exact values for timestamps and template shapes were available. For the real datasets that were used, template MUAP signatures were calculated as the median forms obtained from timestamps grouped in clusters by operators. Similarly, timestamp values marked by the operators were taken as a gold standard for the comparison. In cases when the operators did not achieve total agreement, both values (comparison with both operators individually) were included in the analysis.

3. RESULTS

3.1 Simulated data

The simulated data sets were decomposed in a fully automatic way. Execution time per each 20 s data set was up to 5 minutes on an office desktop computer (Intel Core 2 Quad processor on 2.66 GHz with 4 GB of RAM). Figure 6 shows the average percentage of detected and properly classified MUAP signatures for different noise levels.

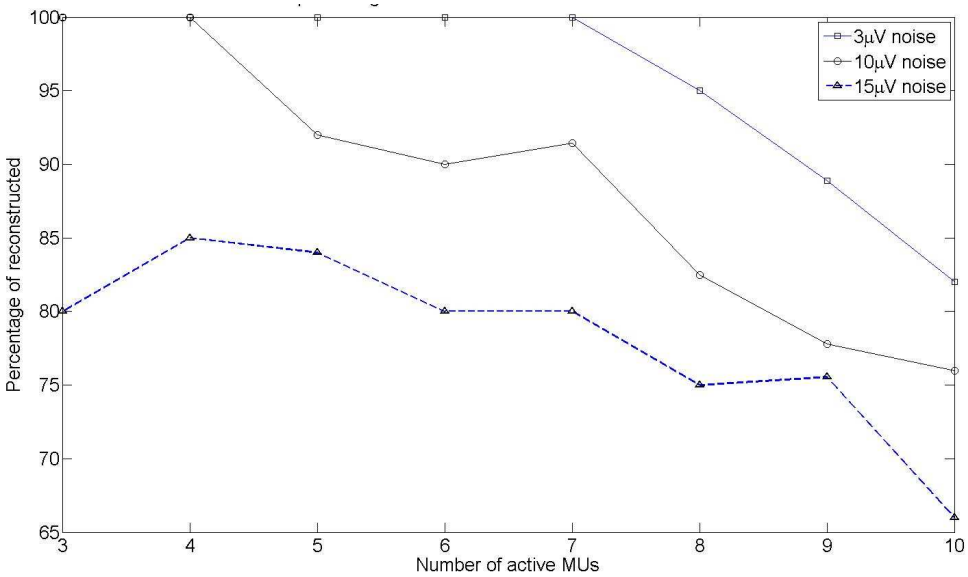


Figure 6: Mean percentage of reconstructed MUAP signatures as a function of noise level and the number of present active units

The maximal number of correctly detected units in a single run was 9. Units that went undetected were often (but not always) the ones with the lowest SNR. For instance, all the units that were not detected in signals containing three active MUs had a SNR value lower than -5 dB on the detection/clustering electrodes.

The sensitivity and precision of the analysis were evaluated after both the clustering and demixing steps. The sensitivity values as a function of noise and number of active MUs are given in Fig. 7. This Figure illustrates the significant gain in performance that results from the demixing procedure. Sensitivity levels sharply increase through demixing and reach average values of 85-90%. The precision before and after demixing was in the range of 0.97 to 1. Despite these consistently high values, there was still a significant difference between the precision before and after demixing for the simulations at the 3 and 10 µV noise levels (p=0.03 for both levels). This means that the demixing process significantly decreased the number of wrongly classified units. For the noise level of 15 µV, results were comparable.

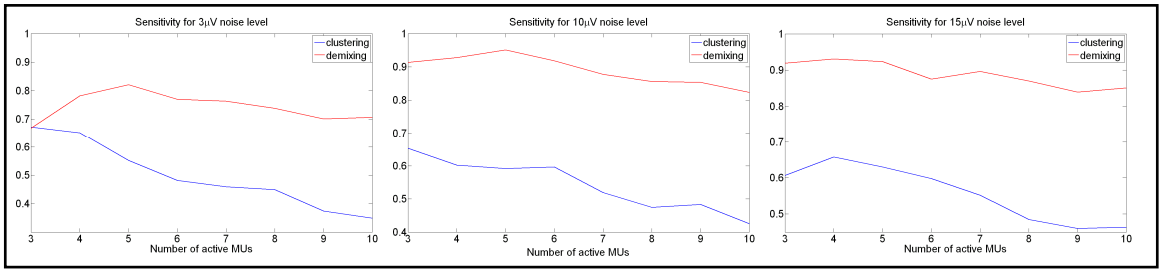


Figure 7: Effect of (additional) demixing on the sensitivity of identified MUs' trains for 5, 10 and 15 µV noise (A, B and C respectively) as a function of number of active MUs

All of the detected MUAPs appearing in large clusters (containing at least 30 elements) were successfully reconstructed. The values for the spatial reconstruction parameters (correlation and NRMSE) are shown in Table 1. Increasing the number of units had only a marginal effect on these values. The reconstructed MUAPs almost exactly matched the template shapes.

Table 1: Quality of spatial reconstruction parameters for different noise levels (mean ± standard deviation) expressed as Pearson correlation coefficient and the normalized root mean square error (NRMSE).

	Noise level		
	3µV	10 µV	15 µV
correlation	0.982 ± 0.005	0.975 ± 0.006	0.972 ± 0.007
NRMSE	0.293 ± 0.013	0.313 ± 0.012	0.319 ± 0.014

Fig. 8 shows an example of an interspike interval (ISI) histogram of a single spike train before and after the demixing process. Note the higher orders (multiples) of the basic interspike interval (Fig. 8A), indicating that one, two,…, etc. firings had been missed. The demixing process accurately decomposes the resulting MUAP mixtures, thereby correctly identifying the missing firings. This results in a regular, complete firing pattern

(peak in the histogram Fig. 8B) with very few missed firings (scattered ISIs around 180 ms).

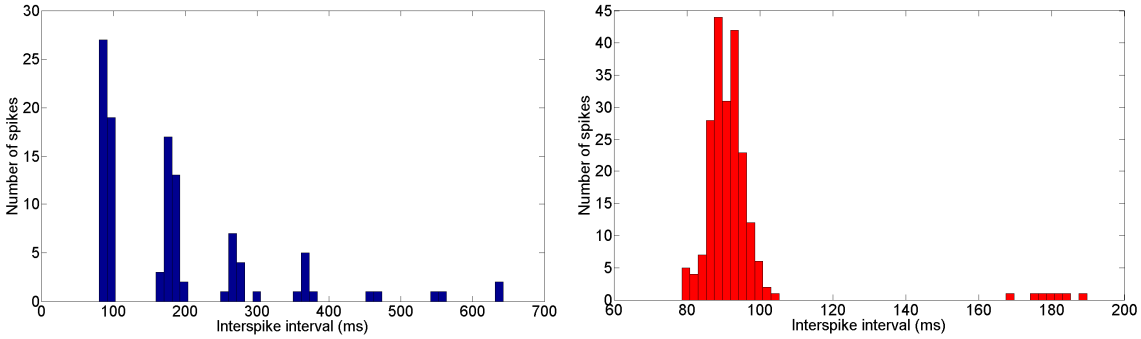


Figure 8: ISI histogram of the firings of a single simulated spike train after the clustering (A) and demixing (B) stages. The demixing (B) almost fully reconstructs the firing pattern revealing the smooth distribution of ISIs compared to a lot of undetected MUAPs resulting in multiples of the mean firing frequency (A).

3.2 Real data

In the nine investigated subjects, two trained investigators found 25 reliable, active MUs (Kleine et al.; 2008). Our automated method extracted 23 of these; 2 units were not detected. Spatial matching between the operator and program extracted shapes was very good, testifying to a successful reconstruction of all identified MUAP signatures. Measures of the temporal aspects of the spike trains were mostly good (sensitivity and precision > 0.8). As in the case of the simulated data, precision was high but sensitivity dropped in the presence of more units (see, e.g., Subject 1) as well as for signals with smaller amplitude (lower SNR). The results of our algorithm were compared against the classifications of each of the investigators separately as well as against their combined detections. In the latter case, the subset of timestamps where both operators agreed on detection was taken as a reference. The results are summarized in Table 2.

Table 2: Comparison of automatic decomposition program and the operators; each of the MUs indicated by the operators was compared in terms of the quality of reconstructed MUAP signatures and the associated firing patterns.

MU	Subject	Force	Firings				Sensitivity			Precision			Spatial correlation	NRMSE
		(%MVC)	program	op. 1	op. 2	inter-op. agreement	with op. 1	with op. 2	with both op-s.	with op. 1	with op. 2	with both op-s.		
1	1	1	21	47	49	47	0.47	0.43	0.47	1.00	1.00	1.00	0.90	0.60
2	1	1	20	45	45	45	0.44	0.44	0.44	1.00	1.00	1.00	0.94	0.43
3	1	1	13	25	25	25	0.48	0.48	0.48	0.92	0.92	0.92	0.92	0.53
4	1	1	7	30	30	30	0.23	0.23	0.23	1.00	1.00	1.00	0.90	0.50
5	1	1	42	49	49	49	0.84	0.84	0.84	0.98	0.98	0.98	0.99	0.28
6	1	1	23	26	25	25	0.81	0.84	0.84	0.91	0.91	0.91	0.96	0.35
7	2	1	42	42	42	42	0.95	0.95	0.95	0.95	0.95	0.95	0.98	0.26
8	2	1	26	38	38	38	0.68	0.68	0.68	1.00	1.00	1.00	0.91	0.44
9	2	1	45	50	50	42	0.60	0.68	0.69	0.67	0.76	0.64	0.83	0.57
10	3	1	31	35	35	35	0.86	0.86	0.86	0.97	0.97	0.97	0.97	0.28

11	4	2	*	*	*	*	*	*	*	*	*	*	*	*
12	4	2	28	27	27	27	1.00	1.00	1.00	0.96	0.96	0.96	1.00	0.24
13	5	1	32	31	31	31	0.94	0.94	0.94	0.91	0.91	0.91	0.97	0.30
14	5	1	39	39	39	39	0.95	0.95	0.95	0.95	0.95	0.95	0.98	0.25
15	5	1	51	51	51	51	0.98	0.98	0.98	0.98	0.98	0.98	0.99	0.15
16	6	1	43	35	35	24	0.66	0.60	0.67	0.53	0.49	0.37	0.93	0.45
17	6	1	24	36	33	23	0.39	0.36	0.39	0.58	0.50	0.38	0.90	0.51
18	6	1	41	43	43	40	0.84	0.84	0.85	0.88	0.88	0.83	0.98	0.32
19	6	1	36	35	35	35	1.00	1.00	1.00	0.97	0.97	0.97	0.96	0.32
20	7	1	33	36	36	35	0.78	0.78	0.77	0.85	0.85	0.82	0.94	0.39
21	7	1	*	*	*	*	*	*	*	*	*	*	*	*
22	7	1	36	41	41	41	0.78	0.78	0.78	0.89	0.89	0.89	0.97	0.25
23	8	4	49	45	45	45	0.96	0.96	0.96	0.88	0.88	0.88	0.98	0.24
24	9	1	40	40	40	40	0.98	0.98	0.98	0.98	0.98	0.98	0.96	0.34
25	9	1	39	50	50	49	0.66	0.68	0.67	0.85	0.87	0.85	0.96	0.32

* - unit was not found using the program

To illustrate how well the signatures of the MUAPs after decomposition matched those detected by the operators (our gold standard), Fig. 9 presents both signatures for MU #9 from Table 2, the unit with the lowest correlation value (0.83). The relatively low value is mainly due to noise corruption in channels where the signature is insignificant.

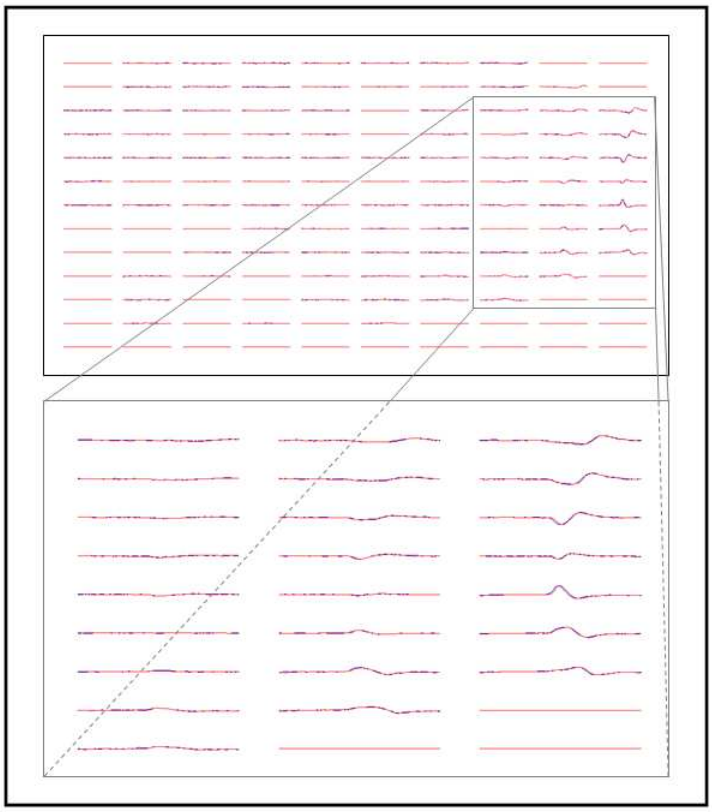


Figure 9: The MUAP with the lowest spatial correlation (0.83; MU #9 in Table 2) between the original (blue) and recovered (red) signatures. The expanded, lower figure zooms in on the part of the array that is closest to the MU (has the largest signals). This example illustrates that even in the cases with a relatively low correlation, the reconstructed and original shapes match very well. The low correlation value is primarily

due to the very small signal size in most channels in combination with the influence of noise.

4. DISCUSSION

In this paper, we presented a new approach aiming at the decomposition of voluntary muscle activity recorded with HD-sEMG. Our algorithm presents a solution following a general framework consisting of two steps. The first step obtains the mean signatures of MUs while the second, demixing step recovers their firing patterns using the previously extracted information. The method was successfully validated on both simulation and real life data. It offers a fast, robust and automated solution for the HD-sEMG decomposition problem. The main advantage of it is that its computational effectiveness allows the fast assessment of both the MUAP signatures of active MUs and their firing patterns. Although envisioned as a fully automated and robust solution, it also allows for a user to manually replace the electrodes used for detection and clustering of MUAPs in search for possibly more optimal result. An important conclusion of this study is that the reconstructed MUAP shapes matched almost perfectly with the used templates (in case of the simulation signals) and the signatures extracted by the two trained operators (for the real life data). In addition, the firing patterns of the MUs generating the MUAP trains were almost fully reconstructed.

We have demonstrated that the multichannel extension of the Wave_clus clustering algorithm (Quan Quiroga et al. (2004)) is able to reliably extract the signatures of MUs. We found that our decision to use only a small number of scattered electrodes for the clustering step enhances its performance, both in the sense of speed as well as quality. It allowed us to obtain the spatial distribution information while at the same time avoiding the bias that would be introduced by channels with high redundancy. This bias is due to the fact that many corresponding features for clustering do not differentiate between the data and thus diminish the relative significance of those which do. While implementing and fine tuning the approach, we found that for proper performance of the algorithm it is essential to use only a very small number of samples around the detected spike peak as input for clustering. Otherwise, the number of MUAPs that can be reliably identified decreases considerably, as previously shown in the work of Zhou and Rymer (2004). Another aspect contributing to the reliable extraction of signatures was that only clusters with a larger number of elements were retained, avoiding smaller ones corrupted by noise or similar mixtures grouped together. The automated criteria for estimating the number of clusters sometimes placed similar shapes that belonged to different MUs in a single cluster. In cases of excessive noise and/or few similar MUAPs, some signatures could not be extracted. It is important to notice, however, that even when some signatures are missed during the clustering stage, the extraction of identified MU shapes was not impaired. However, the quality of the clustering is critically important for the final outcome of the analysis, because the number and accuracy of the MU signatures extracted at this stage determine the performance limits of the next, demixing step for which the signatures serve as necessary input and where no new shapes can be extracted.

Solving the computationally demanding MIMO system of equations needed for the demixing of the superpositioned MUAPs was successfully reduced to a MISO problem. The devised optimization strategy proved to be a very fast and effective alternative. Prior knowledge about the identified MU signatures was exploited through the use of one electrode for representation of each MU. Somewhat surprisingly, using several channels per MU did not necessarily result in better performance (data not shown). Especially in the case of bipolar filtered signals, additional channels would sometimes diminish the significance of the useful information rather than enhance the quality of the decomposition. This bias tended to increase with increasing number of active MUs. For instance, if for each of 8 active MUs 3 electrodes are assigned, this results in a set of 24 representing channels. Particularly when some MUs have similar signatures, this can result in misclassification, since the majority of the electrodes used would not contribute to the differentiation. The criterion for choosing the optimal channel(s) for representation is still subject of discussion. For our initial implementation, we used the one with highest SNR, but using the channel that maximizes the difference with all other signatures (according to some difference criterion to be defined) may be an alternative. The algorithm allows up to three MUAPs to participate in a mixture. Because of the sequential fitting, the error accumulates after each subtraction step and can become larger than the noise in cases if more MUAPs are present. Possibly, more complex mixtures could be treated by means of subsample fitting or through a stricter approach that allows arbitrary mixtures (e.g., similar to McGill's work (2002) on iEMG decomposition) at higher computational cost.

To achieve full decomposition, however, our strategy requires improvement in two respects. First, it should be modified to allow resolution of cases of phase cancellation between nearly simultaneous MUAPs (which may currently render the resulting mixture irresolvable). Second, and more importantly, we need to be able to increase the maximum number of identified signatures beyond 10. One way to achieve this would be through the use of the various spatial filters as reported by Farina et al. (2003), which might increase the differences between MUAPs and/or using the several combinations of scattered electrodes.

Simulated data

The simulated signals were designed to encompass all relevant temporal and spatial aspects of the thenar muscles voluntary activity recorded with HD-sEMG. At realistic noise level of 10 μV rms, our algorithm was able to retrieve nine out of ten units present in the simulated signal. As might be anticipated, we found that higher noise levels reduced the probability of recovering all MUAP signatures (Fig. 6); this effect was stronger for low-amplitude signals. As a result, the MUAPs of a single MU were occasionally assigned to multiple small clusters, each of which was, by itself, not large enough to be associated with an individual MU. Less rigorous conditions on the size of a cluster and/or a way to merge small clusters could potentially remedy this problem.

The precision of the results after both the clustering and the demixing process was generally very high. This implies that the identified spike timestamps are reliable. The

current sensitivity seems adequate for most clinical and research purposes that can be addressed with low-intensity, steady contractions. Modifications or improvements are required in other conditions, e.g., when the firings of different MUs are correlated. Fig. 7 however, indicates one seemingly unexpected phenomenon: sensitivity for the smallest noise level is lower than that of the higher ones. This discrepancy comes from the fact that the automatic estimation of the noise level is not perfect. When a unit is identified and assigned in the demixing process, the remaining residue often stays above the estimated noise value, resulting in characterizing the event as unrecognized. This should be taken into account when little noise is present. However, in this study we preferred to only present results of a fully automated method. Perhaps better ways to assess the noise level might improve this. Sensitivity is only slowly decaying for the noise levels of 10 and 15μV while preserving the precision.

Real-life data

In the second part of the study, we used data from the lateral vastus leg muscle. These data sets have lower SNR values than those of the thenar muscle as used in the simulations. Part of this "noise" stems from physiological sources such as activity from deep MUs or distant muscles. To reduce the effect of this physiological noise, bipolar spatial filtering was applied prior to the processing. Although bipolar signals somewhat differ from monopolar observations, none of these significantly influenced the algorithm's performance.

Out of the 25 different MUAP signatures detected by the two operators in the recorded data, 23 were identified with our automated approach. Possibly, application of a different spatial filter (Farina et al., 2003) or an optimized choice of clustering electrodes would enable detection of the remaining two units.

The sensitivity achieved by our algorithm is similar to the 80-90% level reported in other HD-sEMG decomposition studies (De Luca et. al (2006), Holobar et. al (2010)). It decreased significantly when many (similar or small) unidentifiable MUAPs were present (as in data set 1), because these resulted in many unrecognized shapes and an inability of the algorithm to decompose these mixtures. In the case of dataset 6, first 2 units' clusters were mixed, due to the very similar shapes (correlation = 0.88, NRMSE = 0.26). This lowered the precision considerably, although MUAP signatures were successfully reconstructed. Nevertheless, even in these cases the firing patterns could still be reconstructed and the dominant mean firing frequency derived.

It is worth mentioning that our algorithm identified several additional units. These had not been detected by the operators and were, therefore, excluded from our analysis. Some additional firings, associated with the units observed by the operators, were also found. Both these findings suggest that our algorithm outperforms the human observer, but unfortunately such a claim cannot be substantiated without reference to another, better standard for comparison. Therefore, the next best step appears to be further validation of our algorithm by means of co-registration and correlation of intramuscular and surface

EMG recordings as performed in other studies (De Luca et al. (2006), Holobar et al. (2010)).

REFERENCES

[1] Drost G., D.F. Stegeman, B. G. M. Van Engelen and M. J. Zwarts. Clinical applications of high-density surface EMG: A systematic review. *J. Electromyography Kinesiol.* 16:586-602, 2006.

[2] Drost G., J. H. Blok, D. F. Stegeman, J. P. van Dijk, B. G. M. Van Engelen and M. J. Zwarts. Propagation disturbance of motor unit action potentials during transient paresis in generalized myotonia. *Brain* 124:352-360, 2001.

[3] Stashuk, D. EMG signal decomposition: how can it be accomplished and used? *J. Electromyography Kinesiol.* 11:151-173, 2001.

[4] Wood S. M., J. A. Jarratt, A. T. Barker and B. H. Brown. Surface electromyography using electrode arrays: a study of motor neuron disease. *Muscle Nerve* 24:223-230, 2001.

[5] Blok J. H., J. P. van Dijk, G. Drost, M. J. Zwarts and D. F. Stegeman. A high-density multichannel surface electromyography system for the characterization of single motor units. *Rev. Sci. Instrum.* 73(4):1887-1898, 2002.

[6] Maathuis E. M., J. Drenthen, J. P. van Dijk, G. H. Visser and J. H. Blok. Motor unit tracking with high-density surface EMG. *J. Electromyography Kinesiol.* 18:920-930, 2008.

[7] Kleine B. U., J. P. van Dijk, B. G. Lapatki, M. J. Zwarts and D. F. Stegeman. Using two-dimensional spatial information in decomposition of surface EMG signals. *J. Electromyography Kinesiol.* 17:535–548, 2007.

[8] Roeveld K., D. F. Stegeman, B. Falck and E. V. Stålberg. Motor unit size estimation: confrontation of surface EMG with macro EMG. *J. Electroen Clin Neuro.* 105:181-188, 1997.

[9] Merletti R., A. Holobar and D. Farina. Analysis of motor units with high-density surface electromyography. *J. Electromyography Kinesiol.* 18:879-890, 2008.

[10] Gazzoni M., D. Farina and R. Merletti. A new method for the extraction and classification of single motor unit action potentials from surface EMG signals. *J. Neurosci. Methods.* 136(2):165-177, 2004.

[11] Florestal J. R., P. A. Mathieu and A. Malanda. Automated decomposition of intramuscular electromyographic signals. *IEEE T. BIO-MED ENG.* 53(5):832-839, 2006.

[12] Florestal J.R., P.A. Mathieu and K.C. McGill. Automatic decomposition of multichannel intramuscular EMG signals. *J. Electromyography Kinesiol.* 19:1-9, 2009.

[13] Blok J. H, J. P. van Dijk, M. J. Zwarts and D. F. Stegeman. Motor unit action potential topography and its use in motor unit number estimation. *Muscle Nerve* 32: 280–291, 2005.

[14] McGill K. C. Optimal resolution of superimposed action potentials. *IEEE T. BIO-MED ENG.* 49(7):640-650, 2002.

[15] Zazula D, A. Holobar. An approach to surface EMG decomposition based on higher-order cumulants. *Comput. Meth. Prog. Bio.* 80(1):51-60, 2005.

[17] De Luca C. J., A. Adam, R. Wotiz, L. D. Gilmore and S. Hamid Nawab. Decomposition of Surface EMG Signals. *J Neurophysiol* 96:1646-1657, 2006.

[18] Holobar A. and D. Zazula. Multichannel blind source separation using convolution kernel compensation. *IEEE Trans. Signal Process.* 55(9):4487-4496, 2007.
[19] Holobar A., M. A. Minetto, A. Botter, F. Negro and D. Farina. Experimental analysis of accuracy in the identification of motor unit spike trains from high-density surface EMG. *IEEE Trans. Neural Syst. Rehabil. Eng.* 18(3):221-229, 2010.
[20] Drost G., D. F. Stegeman, M. L. Schillings, H. L. D. Horemans, H. M. H. A. Janssen, M. Massa, F. Nollet and M. J. Zwarts. Motor unit characteristics in healthy subjects and those with Postpoliomyelitis syndrome: a high-density surface EMG study. *Muscle Nerve* 30: 269–276, 2004.
[21] Lapatki B. G. , R. Oostenveld, J. P. van Dijk, I. E. Jonas, M. J. Zwarts and D. F. Stegeman. Optimal placement of bipolar surface EMG electrodes in the face based on single motor unit analysis**.** *Int. J. Psychophysiol.* 47(2):299-314, 2010.
[22] Farina D., E. Schulte, R. Merletti, G. Rau and C. Disselhorst-Klug. Single motor unit analysis from spatially filtered surface electromyogram signals. Part I: spatial selectivity. *Med Biol Eng Comput.* 41(3):330-7, May 2003.
[23] Donoho D. L., and J. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika.* 81(3):425-455, 1994.
[24] Quian Quiroga R., Z. Nadasdy and Y. Ben-Shaul. Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Computation* 16:1661-1687, 2004.
[25] Gligorijević I., M. De Vos, J. H. Blok, B. Mijović, J. P. van Dijk and S. Van Huffel. Automated way to obtain motor units' signatures and estimate their firing patterns during voluntary contractions using HD-sEMG. *IEEE Boston EMBC 2011 conference*, 2011.
[26] Blatt, M., S. Wiseman and E. Domany, Superparamagnetic clustering of data. *Phys. Rev. Lett.*, 76: 3251–3254. 1996.
[27] Abed-Meraim K., Q. Wanzhi and Hua Y.. Blind system identification. *P. IEEE*, 85(8):1310-1322, 1997.
[28] van Dijk J. P. , J. H. Blok, B. G. Lapatki, I. N. van Schaik, M. J. Zwarts, D. F. Stegeman. Motor unit number estimation using high-density surface electromyography. *Arch. Clin. Neuropsych.* 119(1):33-42, 2008.
[29] Kleine B. U., J. P. van Dijk, M. J. Zwarts, D. F. Stegeman. Inter-operator agreement in decomposition of motor unit firings from high-density surface EMG. *J. Electromyography Kinesiol.* 18:652–661, 2008.
[30] Zhou P. and W. Z. Rymer. MUAP Number Estimates in Surface EMG: Template-Matching Methods and Their Performance Boundaries. *Ann Biomed Eng.* 32(7): 1007-1015, 2004.
[31] Denny-Brown D., Interpretation of the electromyogram. Arch. Neurol. Psychiatry, 61:99–128, 1949
[32] Lapatki B. G., J. P. van Dijk, I. E. Jonas, M. J. Zwarts and D. F. Stegeman. A thin, flexible multielectrode grid for high-density surface EMG. J. Appl. Physiol. 96(1):1327-336, 2004