Evaluating a motor unit potential train using cluster validation methods

Hossein Parsaei[1] and Daniel W. Stashuk

Systems Design Engineering, University of Waterloo, Waterloo, Canada;

ABSTRACT

Assessing the validity of motor unit potential trains (MUPTs) obtained by decomposing a needle-detected electromyographic (EMG) signal is a crucial step in using these trains for quantitative EMG analysis. In general, for MUPT validation a train is assessed using the shapes of its motor unit potentials (MUPs) and the motor unit firing pattern it represents. Here, two methods to assess the validity of a given MUPT using its MUP shape information are presented. These methods are based on the gap statistic and jump algorithms presented for estimating the number of clusters in a dataset. They evaluate the shapes of the MUPs of a MUPT to determine whether it represents the activity of a single MU (i.e. it is a valid MUPT) or not. Evaluation results using both simulated and real data show the gap statistic method is more accurate than the jump method in correctly categorizing a train. The accuracy of the gap static method was 92.3% for simulated data and 93.8% for real data while accuracy for the jump method was 88.3% and 91.0%, respectively. The results are encouraging and suggest that using these methods can improve EMG signal decomposition results, and facilitate automatic MUPT validation.

Keywords: EMG signal decomposition, motor unit potential train, motor unit potential wave shape, motor unit potential train validation, cluster validation.

1 INTRODUCTION

An electromyographic (EMG) signal is simply the superposition of the electrical activity detected by an electrode of the motor units (MUs) that are active during muscle contraction. A motor unit is a single α-motor neuron, its axon and all the connected muscle fibers. MUs are repetitively active during sustained voluntary contraction and generate trains of motor unit potentials (MUPs), each of which is called a motor unit potential train (MUPT). Therefore, an EMG signal is the summation of MUPTs and when detected using suitable electrodes reflect the characteristics of the muscle from which it was detected.

Recent development in computer technology, signal processing and pattern recognition techniques have provided researchers and engineers with opportunities to develop new techniques for extracting valuable information regarding a contracting muscle from an EMG signal detected from this muscle. One of these techniques is EMG signal decomposition which is the process of resolving a composite EMG signal into is constituent MUPTs. This is implemented by employing digital signal processing and pattern recognition techniques in four steps: signal preprocessing, signal segmentation and MUP detection, feature extraction, clustering of detected MUPs, and supervised classification of detected MUPs [1].The first step is to remove background noise and low-frequency information from the detected EMG signal, to shorten the duration of MUPs and decrease MUP temporal overlap, and to sharpen the MUPs and increase discrimination between them. The second step is to section the signal into segments containing possible MUPs that were generated by active motor units and contribute significantly to the detected EMG signal. The detected MUPs are represented by a feature vector in the third steps and finally are sorted into MUPTs using clustering and/or supervised classification techniques. The obtained MUPTs provide information regarding the temporal behavior and morphological layout of the generating MUs. This information can assist with the diagnosis of various neuromuscular diseases and the study of motor unit control, and lead to a better understanding of healthy, pathological, ageing or fatiguing neuromuscular systems [1-5]. However, this is achieved only when this information is valid. In fact, before using decomposition results and the MUP shape and MU firing pattern information for either clinical or research purposes the validity of the extracted MUPTs needs to be confirmed. Although many EMG signal decomposition methods have been developed, automatic validation of the extracted MUPTs has not been investigated in detail.

---

[1] Corresponding author. E-mail: hparsaei@engmail.uwaterloo.ca, Telephone: +1(519) 888- 4567 Ext.33746.

To date, a decomposition-created MUPT is evaluated qualitatively by an expert operator using the shapes and occurrence times of the MUPs assigned to the train [1,5,6]. MUP shape-based validation of a MUPT is conducted by assessing the raster/shimmer plots of the MUPs assigned into this train to determine if their wave shapes are consistent or not. In fact, if the shapes of the MUPs assigned to a given MUPT are consistent, one can conclude that this MUPT represents the MUPs of a single MU and is valid; otherwise it is an invalid train. Firing pattern-based validation of a MUPT is made by evaluating its inter-discharge interval (IDI) histogram [1] and the instantaneous firing rate of the corresponding MU versus time. The MU discharges corresponding to a valid MUPT occur at regular intervals, while an invalid MUPT that does not represent the activity of a single MU and contains many false positive errors will have large variations in its firing rate plot and will not have a Gaussian like IDI distribution. A train is considered valid if it satisfies both temporal and shape criteria.

Although qualitative evaluation of MUPTs does not depend on the decomposition algorithm used or the signal decomposed, it does depend on operator experience and skill. Moreover, it is time consuming and hence cannot be practically completed in a busy clinical environment. To overcome these issues, methods need to be developed to automatically assess MUPT validity. Parsaei at al. [7] developed a supervised method for validating a MUPT using its firing pattern information. In these methods, ten features of the MUPT firing pattern are extracted and then fed to two supervised classifiers and to a linear model to determine whether they represent the firings of a single MU or the merged activity of more than one MU, and if it is a single train whether the estimated levels of false positive and false negative errors in it are acceptable or not. Here, automatic MUP-shape based validation of a MUPT is explored. Two methods based on cluster validation concepts are proposed to automatically validate MUPTs using their MUP shape information. They evaluate the shapes of the MUPs of a MUPT to determine whether they are consistent or not. If a train passes this test, it can be concluded that it represents the activity of a single motor unit and hence is valid. The composition of these methods, their objectives and how they were evaluated using both simulated and real data are presented below.

2 VALIDATING A MUPT USING MUP SHAPE INFORMATION

To convey the concept of assessing validity of a MUPT using its MUP shape information, two examples are provided in Figure 1. The left column shows a valid MUPT and the right column shows an invalid train. The valid train was obtained from decomposing a simulated EMG signal. The invalid train was created by merging two valid MUPTs. As shown, the shapes of MUPs assigned to the valid MUPT are consistent while that of the invalid train are inconsistent. The shapes of MUPs in the invalid train are different for samples 11 to 25. The goal of developing a MUPT validation method is to perform this assessment automatically during or after decomposition. The advantages of using such an algorithm during EMG signal decomposition is that detecting invalid trains during decomposition can improve decomposition accuracy by improving estimation of the number of MUPTs, their MU firing pattern statistics and MUP templates.

On the whole, the process of EMG signal decomposition can be considered as a clustering problem because neither the number of MUPTs (i.e. clusters) nor the labels of the MUPs are known in advance. During EMG signal decomposition, detected MUPs are clustered into groups called MUPTs. Therefore, shape-based validation of a MUPT can be considered a cluster validation problem and the decision to be made is whether a decomposition-created MUPT represents one cluster in terms of the shapes of the assigned MUPs or not. If the MUPs of a given MUPT are homogeneous in terms of their shapes, they will represent one cluster and hence it can be concluded

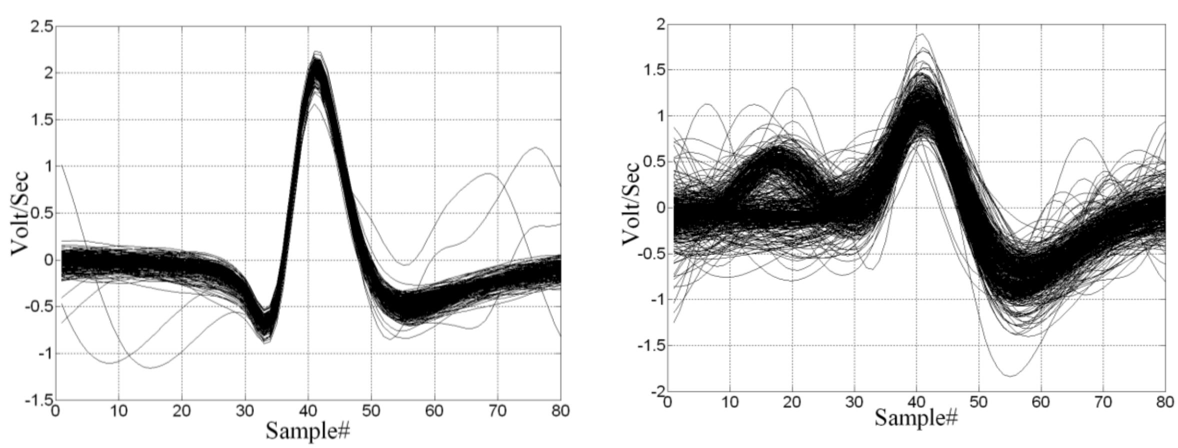

Figure 1. A valid MUPT (left) and a simulated invalid MUPT (right).

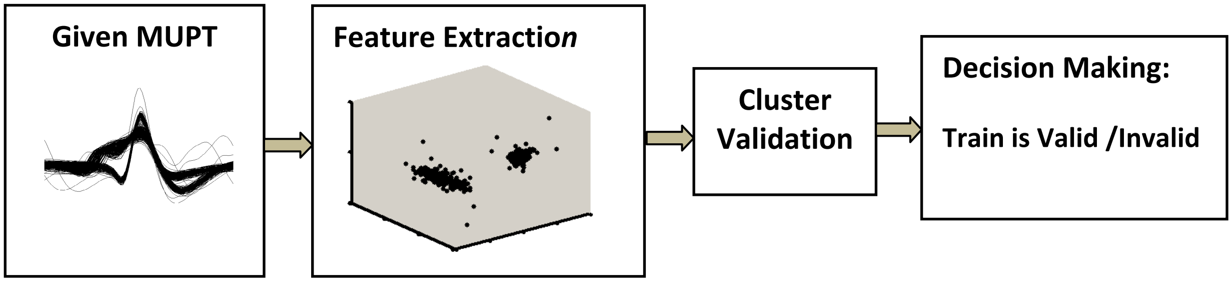

Figure 2. Overview of an automatic MUPT validation algorithm using its MUP shape information.

that this MUPT represents the MUPs of a single motor unit and is valid; otherwise this is an invalid train. Figure 2 summarizes the principal steps of a MUP shape–based MUPT validation algorithm.

In this work, two algorithms based on the gap statistic [8] and the jump methods [9] have been developed for this purpose. Both algorithms include two steps: feature extraction and cluster validation. For features extraction, two techniques were employed. For the first feature space, 80 time samples of the MUPs filtered by a low-pass differentiator (LPD) filter and then centered on the position of their peaks were used. The LPD filtered samples were used instead of unfiltered samples because they discriminate between the MUPs generated by different motor units better than the raw samples. For the second feature space, a reduced number of uncorrelated features selected via principle components analysis (PCA) were used. The MUPs of a given train were represented by the first T principal components that represent β% of the data variance. Given the features extracted, the validity of the given train was assessed using either the gap statistic or jump method.

Both the gap statistic and jump methods were mainly developed for estimating the optimum number of groups ($\widehat{\mathrm{K}}$) in a data set, but they can also be used for evaluating a single cluster such that they are classified as global cluster validation methods. In general, global cluster validation methods compare two/more clustering results for deciding which one fits the data best. For finding $\widehat{\mathrm{K}}$ in a given data set via these methods, the quality of clustering results is measured by a criterion and then is optimized as a function of the number of clusters when the entire data set is clustered into K groups for K=1, 2, 3,…, K*. Where and K* is the maximum possible number of clusters in the given data set. It is clear that $\widehat{\mathrm{K}}$ will be the value of K for which this criterion is optimal. In general, the criteria are defined based on the assumption that a cluster is compact and well separated from other clusters. Hence, they are often defined based on the within-cluster scatter dispersion ($\mathrm{W_K}$) and between-cluster scatter variability ($\mathrm{B_K}$). $\mathrm{W_K}$ is given by [8]:

$$W_K = \sum_{i=1}^{K} \sum_{\underline{x} \in c_i} (\underline{x} - \underline{m}_i)(\underline{x} - \underline{m}_i)^T \tag{1}$$

where $\underline{\mathrm{m}}_{\mathrm{i}}$ is the sample mean for the $\mathrm{N_i}$ members of the ith cluster.

The gap statistic [8] method estimates the number of clusters by comparing $\log(\mathrm{W_K})$ to its expected value (i.e., $\mathrm{E}\{\log(\mathrm{W_K})\}$) estimated using an appropriate null reference distribution of the given data set. Defining the gap statistic as $\mathrm{Gap_K} = \mathrm{E}\{\log(\mathrm{W_K})\} - \log(\mathrm{W_K})$, the best value for $\widehat{\mathrm{K}}$ corresponds to the first local maximum of $\mathrm{Gap_K}$. Tibshirani et al. [8] proposed two methods to generate a reference data set for the gap statistic method. In the first method, it is sampled uniformly from the range of the observed values for each feature. In the second method, the reference data set is also sampled uniformly but here over a box aligned with the principal components of the data. The gap statistic method based on the first method is known as the gap/uni method and that based on the second method is known as the gap/pc method. The former is simpler than the latter but may not be as accurate because the gap/pc considers the shape of the cluster in generating the reference data set and hence has a better estimation of $\mathrm{E}\{\log(\mathrm{W_K})\}$.

The jump method [9] applies an appropriate transformation to the curve of $\mathrm{W_K}$ and then determines the largest jump in the transformed curve. The value of K associated with this jump is considered as the best value for $\widehat{\mathrm{K}}$. Sugar and James [9] proposed the following transformation for $\mathrm{W_K}$

$$\mathrm{W_K^*} = \mathrm{W_K^{-Y}} \tag{2}$$

where Y is the transformation power. A typical value for Y is $^{\mathrm{d}}/_{2}$, where d is the dimension of the feature space. Defining $\mathrm{W_0^*} = 0$, the jump index ($\mathrm{J_K}$) is given by $\mathrm{J_K} = \mathrm{W_K^*} - \mathrm{W_{K-1}^*} = \mathrm{W_K^{-Y}} - \mathrm{W_{K-1}^{-Y}}$; K=1,2,…,K*. Given $\mathrm{J_K}$, $\widehat{\mathrm{K}} = argmax_{\mathrm{K}}(\mathrm{J_K})$. The theoretical results provided by Sugar and James [9] show that $^{\mathrm{d}}/_{2}$ is the best value for Y only when the data has a multivariate independent Gaussian distribution. For the cases that this assumption is not valid, they suggest trying several values of Y to find the best value for this parameter.

Table 1. The parameters and accuracy of the four methods studied for validating a MUPT using its MUP shape information applied to simulated and real data. **VasV** stands for the valid MUPTs classified as valid, ***Iv as Iv*** represents the invalid MUPTs recognized as invalid, and **Acc** represents the total accuracy.

		Simulated Data			Real Data		
Method	Parameters	V as V %	Iv as Iv %	Acc %	Vas V %	Iv as Iv %	Acc %
Gap statistic	-	83.0±0.4	98.2±0.3*	90.6±0.3	89.2	98.3*	93.8
PGS	$\beta = 90$	92.0±0.6*	93.3±0.3	92.7±0.3*	93.1	97.9	95.5*
Jump	$Y = 3$	81.5±0.8	90.5±0.6	86.0±0.5	87.9	94.0	91.0
PJ	$\beta = 50; Y = 2$	86.2±0.6	90.4±0.6	88.3±0.4	89.1	93.7	91.4

In order to assess the validity of a given MUPT using the jump method, the train is split into K=1 to K* sub trains using a K-means algorithm and then if $max\ (\mathrm{J_K}) = \mathrm{J_1}$, the given train is labeled a valid train otherwise it is labeled an invalid train. If the gap statistic method is to be used for validating this train, the same procedure will be repeated but the gap criterion will be used for cluster validation proposes. However, it is shown that for a given data set including compact and well separated clusters $\mathrm{W_K}$ decreases monotonically as the value of K increases, but when K reaches the true number of clusters (i.e., $\hat{\mathrm{K}}$) this decrease becomes smoother. Thus, $\mathrm{Gap_K} < \mathrm{Gap_{\hat{K}}}$ when $K < \hat{K}$ and $\mathrm{Gap_K} > \mathrm{Gap_{\hat{K}}}$ when $K > \hat{K}$. In other words, if $\mathrm{Gap_1} > \mathrm{Gap_2}$, $max\ (\mathrm{Gap_K})$ will be equivalent to the gap value for K=1. Consequently, since the goal is to determine whether a MUPT represents one cluster in terms of the assigned MUP shapes or not, only $\mathrm{Gap_K}$ for K=1 and 2 are sufficient. Therefore, for the gap statistic-based MUPT validation, the algorithm is only run for K=1 and 2 and if $\mathrm{Gap_1} > \mathrm{Gap_2}$ the MUPT under question is flagged a valid train otherwise it is classified an invalid train. This decreases the algorithms processing time and hence makes it practical for clinical applications. For the jump-based MUPT validation, however, the algorithm must be run for K=1,2,3,…, K* because even if $\mathrm{J_1} > \mathrm{J_2}$ the is no guarantee that $max\ (\mathrm{J_K}) = \mathrm{J_1}$ (i.e., $\hat{\mathrm{K}}$ =1).

3 RESULTS AND DISCUSSION

The effectiveness of the developed method was studied using both simulated and real data. In total four methods, as listed in Table 1 were evaluated. In this Table, PGS and PJ stand for the gap statistic and jump methods when the features used were selected using PCA, respectively.

For simulated data, 261 EMG signals each of 30s length with different levels of intensity (24-93 pps), MUP shape stability (with jitter values from 50-150µs) and IDI variability (CV from 0.10-0.45) were generated using an EMG signal simulator [10]. These data allowed the performance relative to signal intensity, number of trains and MUP shape variability to be studied. The simulated signals were decomposed (using the DQEMG software [11]) and the resulting MUPTs visually assessed to determine valid and invalid MUPTs. Additional invalid trains were generated by merging valid MUPTs (up to 4) randomly selected from each signal. Additional valid trains were generated by selecting valid MUPTs with greater than 100 MUPs and randomly splitting them into sub trains of at least 50 MUPs. In total 36000 MUPTs (18000 valid and 18000 invalid trains) were tested. The number of merged trains generated was 234778, but only 18000 were randomly selected to have equal class sizes. Out of the 18000 selected merged trains, 60% include MUPs of two valid MUPTs, 30% include MUPs of three valid MUPTs, and 10% include MUPs of four valid MUPTs. This data set was divided into 30 subsets each containing 600 valid and 600 invalid trains.

For real data, EMG signals provided by M. Nikolic of Rigshospitalet, Copenhagen, Denmark [12] were used. These signals were detected from normal, myopathic and neuropathic muscles using a standard concentric needle electrode during constant low level voluntary contractions. The same analysis as with the simulated data was done on these signals. This dataset includes 3130 MUPTs (1565 valid and 1565 invalid trains).

Before evaluating the four considered methods using the provided simulated and real data, the best values for their user defined parameters were determined empirically using one of the thirty subsets of the simulated data described above. In using the gap statistic method, the gap/pc was used because it outperformed the gap/uni method. For the other methods, the values used for their parameters are listed in the second column of Table 1.

The average accuracy of the developed method in determining valid MUPTs and invalid MUPTs for both the simulated and real data sets are summarized in Table 1. The accuracy here is defined as percentage of the number of correctly classified MUPTs. For example, in the column presenting the results for valid trains, the accuracy represents the percentage of the number of examined valid train labeled as valid by the studied algorithm. The numbers given for simulated data are obtained by testing each method using the thirty different data sets described above. For each column, in the simulated data category the best method as determined by a t-test using a 5% significance level are indicated by '*'. Based on these results, the PGS method is the best algorithm because it is the most accurate in terms of overall accuracy and labeling valid trains correctly. In classifying invalid MUPTs the gap statistic is the best performer, but the probability of error of this method for valid trains is high. It is 0.17 on average which causes duplication of MUPTs during EMG signal decomposition. The PGS method is the second most accurate with an accuracy of 93% which is acceptable in a practical sense. The results obtained using the real data support the conclusion drawn from the simulated data results that the PGS method is the best algorithm. All four methods studied performed better using the real dataset than the simulated dataset because the variably of the MUPs in the real dataset are lower than those of the simulated EMG signals and also the MUPs created by different MUs in the real dataset are less similar than those in the simulated signals. Most of the valid trains recognized as invalid are trains with highly variable MUP shapes caused by either high numbers of superimpositions (for the signals with high intensity) or very high jitter (around 150μs). Therefore, the accuracy of this method in determining valid MUPTs will be higher for trains provided by EMG decomposition algorithms that resolve superimposed MUPs. Most of the invalid trains that labeled incorrectly are trains with very similar MUP shapes. Such trains are hard to assess using only shape information, but firing pattern information can assist with label them correctly [7]. Nevertheless, the obtained results are encouraging and suggest that using these methods can facilitate automatic validation of a MUPT extracted from a decomposed EMG signal. It can also improve EMG signal decomposition results, by obtaining more accurate estimates of the number of MUPTs, and the MUP template and MU firing pattern statistics of each MUPT.

Acknowledgements

The authors gratefully acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] D.W. Stashuk, "EMG signal decomposition: how can it be accomplished and used? " *J. of Electromyography and Kinesiology*, **1**, pp. 151-173, 2001.

[2] A. Fuglsangfrederiksen, "The role of different EMG methods in evaluating myopathy," *Clinical Neurophysiology*, **117**, pp. 1173-1189, 2006.

[3] E. Sålberg and B. Falck, "The role of electromyography in neurology," *Electroencephalography and Clinical Neurophysiology*, **103**, pp. 579-598, 1997.

[4] C. J. de Luca, R.S. LeFever, M.P. McCue, and A.P. Xenakis, "Control scheme governing concurrently active human motor units during voluntary contractions," *J. of Physiology*, 329, pp. 129-142, 1982.

[5] K.M. Calder, D.W. Stashuk, and L. McLean, "Physiological characteristics of motor units in the brachioradialis muscle across fatiguing low-level isometric contractions," *J. of Electromyography and Kinesiology*, **18**, pp. 2-15, 2008.

[6] S.G. Boe, D.W. Stashuk, W.F. Brown, and T.J. Doherty, "Decomposition-based quantitative electromyography: Effect of force on motor unit potentials and motor unit number estimates," *Muscle & Nerve*, **31**, pp. 365-373, 2005.

[7] H. Parsaei, F. Jahanmiri Nezhad, D.W. Stashuk, and A. Hamilton-Wright, "Validation of motor unit potential trains using motor unit firing pattern information," in 31st Annu. Int. Conf. of IEEE EMBS Minnesota, USA, pp. 974-977, 2009.

[8] R. Tibshirani R, G. Walther G, and T. Hastie, "Estimating the number of clusters in a data set via the gap statistic," *J. of the Royal Statistical Society*, **63**, pp. 411–423, 2000.

[9] A.C. Sugar, and G. M. James, "Finding the number of clusters in a data set: an information theoretic approach," *J. of the American Statistical Association*, **98**, pp.750–763, 2003.

[10] A. Hamilton-Wright and D.W. Stashuk, "Physiologically based simulation of clinical EMG signals," *IEEE Trans. on Biomed. Eng.*, **52**, pp.171-183, 2005.

[11] D. W. Stashuk, "Decomposition and quantitative analysis of clinical electromyographic signals," *Med. Eng.& Phys.*, **21**, pp.389-404,1999

[12] M. Nikolic, "Detailed Analysis of Clinical Electromyography Signals EMG Decomposition, Findings and Firing Pattern Analysis in Controls and Patients with Myopathy and Amytrophic Lateral Sclerosis," Ph.D. dissertation, Faculty of Health Science, University of Copenhagen, 2001.