I D

IMPROVED DECOMPOSITION OF INTRAMUSCULAR EMG SIGNALS

S. Hamid Nawab[1], Robert Wotiz[1], Lorin Hochstein[1] and Carlo J. De Luca[2]

College of Engineering,
Boston University
Boston, MA 02215

ABSTRACT

We present a framework, initial implementation, and experimental results for the decomposition of intramuscular electromyographic (EMG) signals. A multi-channel EMG signal from an electrode inserted into a muscle may be modeled as a sum of N quasi-periodic pulse trains. Typically, $N \in [1,20]$. Treating the occurrence of pulses within a train as an interval process, we use MAP receiver techniques to initiate and track the trains. Employing an object-oriented program architecture enables us to experiment with signal processing strategies to deal with pulse overlap, track merging, track extension, and re-classification of misclassified detections. The initial implementation is already performing more accurate decomposition than the best system predating it.

1. INTRODUCTION

The process by which the brain causes muscles to produce force is not yet fully understood by researchers. This process is mediated by the motor control system, a complex network of interconnected neurons. There are a large number of research areas that depend on the study of the motor control system, ranging from the study of muscular disorders and diseases such as (potentially fatal) spasmodic dysphonia to the effects of zero-gravity on the human body.

Muscle fibers are stimulated by neurons whose cell bodies are located in the spinal cord. These muscle fibers together with the motor neuron are referred to as a motor unit. To study the firing patterns of motor units, a multi-channel electrode is inserted into the muscle, which records an intramuscular EMG signal. When the muscle fibers near the electrode contract, the electrode records a pulse, known as a motor unit action potential.

Supported by:
NICHD Biomed. Res. Partnership Grant # 1R24HD38585-01

[1] S. Hamid Nawab, Robert Wotiz and Lorin Hochstein are with the ECE Department (*email*: hamid@bu.edu).
[2] Carlo J. De Luca is with the BME Department and is the Director of the NeuroMuscular Research Center at Boston University (*email*: cjd@bu.edu).

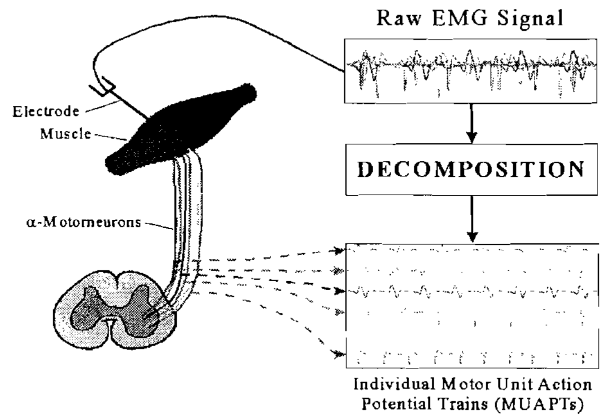

Fig. 1 Schematic of EMG pulse tracking (Courtesy of NMRC, Boston University)

Throughout a muscle contraction, the motor units fire repetitively, generating motor unit action potential trains. An electrode will detect pulses from several nearby fibers. Some fibers will belong to the same motor unit, and some will belong to different motor units. Therefore the recorded EMG signal is a summation of trains from all of the motor units which have fibers near the electrode.

To study the firing patterns of the motor units, the pulses corresponding to each motor unit must be tracked throughout the contraction. However, since the detected pulse shapes can interfere with each other when they overlap in time (see Fig. 2), and since the pulse shapes can change shape and amplitude over time, tracking is a very challenging task.

In the late 1970's and through the 1980's, the NeuroMuscular Research Center (NMRC) at Boston University [1-4] developed a system that decomposes an intramuscular EMG signal into the constituent motor unit action potential trains. The application of this "precision decomposition" system to experimental EMG data has led to significant physiological findings. However, to address many of the remaining physiological issues it has become necessary to overcome this system's limitations:

- Processing is very slow. The automated component requires several hours of execution time.

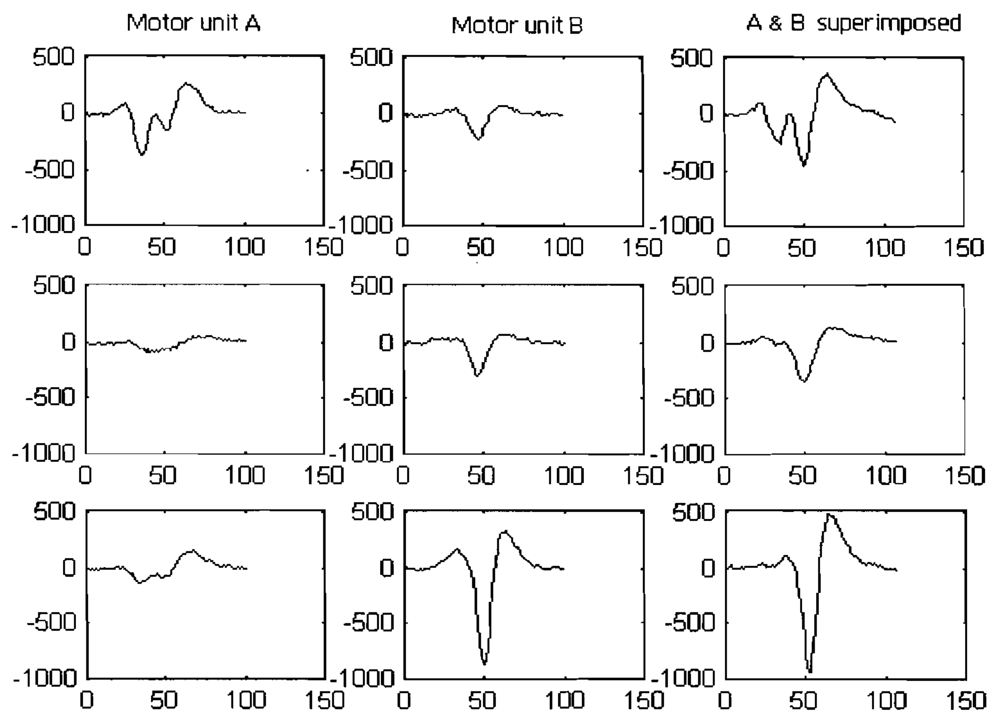

Fig. 2 Typical pulse shapes

- Automated results are not accurate enough, requiring manual editing by a human operator. Manual editing takes several days to complete.
- For a given experiment, the system only tracks up to 20 motor units.
- For a given experiment, the system only handles signals containing at most 3000 pulses.
- The system only handles data sampled at 50 kHz, despite the fact that the signal bandwidth is only 20 kHz.
- The system can only handle 3 channels of EMG data.
- The system uses outdated software which only runs on an obsolete platform (VAX/VMS).

The system presented in this paper is an entirely new implementation developed to overcome the limitations of the previous system. Preliminary results indicate that the new system outperforms the original when applied to typical intramuscular EMG signals that are obtained through experiments.

2. FRAMEWORK

The EMG signal consists of a sum of multi-channel pulse trains. Our problem is to assign each pulse to its train. This is done by passing each multi-channel pulse through a maximum a posteriori (MAP) receiver to determine the train to which it belongs or to initiate a new train. To implement the receiver, we can acquire the a priori probability of occurrence of each train's pulse at a given time through the use of the hazard (or failure rate) function associated with the corresponding interval process. The prototype pulse used by the MAP receiver for any particular train also has to be regularly updated since pulse-shape characteristics are generally time-varying.

2.1. The pulse detector

An EMG signal is characterized by pulses of short duration (typically 2 ms.) separated by longer intervals (typically 80 ms.). Pulses may be located by examining regions where the energy of the signal exceeds a given threshold. The exact location of the pulse within an identified region may be determined by identifying the largest amplitude peak across all channels.

The pulse detection process is more involved when there is a degree of overlap between pulses from different trains. The basic strategy from [1,2] is to first detect and classify the pulse with the highest amplitude peak and then to subtract the corresponding prototype pulse from the data before trying to detect any overlapping pulses. Since the subtraction process can introduce error energy, there is the potential for the system to erroneously detect this energy as another pulse. To avoid such errors, we have introduced the following innovation with respect to [1,2]: After subtraction of a prototype pulse, data values whose magnitudes are below the corresponding prototype pulse magnitude are ignored when attempting to detect another pulse. In other words, we model the subtraction process as introducing errors proportional to the values being subtracted.

2.2. The MAP receiver

The task of the MAP receiver is to associate a particular pulse train with a detected multi-channel pulse. Denoting the i'th pulse train by the symbol u_i, the MAP receiver selects a particular symbol, say u_k, such that:

$$P(u_k|\bar{\rho}) > P(u_i|\bar{\rho}) \quad 1 \le i \le M, \quad i \ne k \tag{1}$$

where $\bar{\rho}$ is the vector consisting of the concatenated data channels of the detected pulse. In the presence of white Gaussian noise with variance σ^2, the MAP decision criterion may be restated as:

$$|\bar{\rho}-\bar{s}_k|^2 - 2\sigma^2 \ln P(u_k) < |\bar{\rho}-\bar{s}_i|^2 - 2\sigma^2 \ln P(u_i),$$
$$1 \le i \le M, i \ne k \tag{2}$$

where $\bar{s}_i$ is the vector consisting of the concatenated data channels of the prototype pulse for the i'th train, and $P(u_i)$ is the a priori probability that the detected pulse belongs to the pulse train u_i.

In the EMG application, the lengths of vector waveforms for each prototype pulse differ, and may vary over time for the same train. As a result, the MAP decision criterion in (2) may be modified [1,2] by normalizing the error signals as follows:

$$\frac{|\bar{\rho}-\bar{s}_k|^2}{|\bar{s}_k|^2} - 2\sigma^2 \ln P(u_k) < \frac{|\bar{\rho}-\bar{s}_i|^2}{|\bar{s}_i|^2} - 2\sigma^2 \ln P(u_i),$$
$$1 \le i \le M, \quad i \ne k. \tag{3}$$

An innovation with respect to [1,2] is the manner in which $\bar{\rho}$ is aligned with $\bar{s}_i$ before computing their difference. Rather than simply basing the alignment on the locations of the highest peaks in each of the vectors, our alignment process also considers the second highest peak in $\bar{\rho}$ as well as the polarities (negative or positive) of the peaks whose locations are being aligned. We have found this change to be a major factor in the improved results from our system with respect to previous systems.

2.3. The a priori probabilities

When a pulse is detected at a particular time, we are interested in estimating the a priori probability that the pulse belongs to a particular pulse train. To accomplish this, we utilize the hazard function associated with each of the pulse trains. Experimental evidence [1,2] indicates that the corresponding interval processes are non-Poisson and thus their arrival rates are time-varying. It has been suggested [1,2] that the first order inter-arrival times (T) have a Gaussian density function and thus the hazard function for the i'th train may be obtained as:

$$R_i(t) = \frac{f_{T_i}(t)}{1 - F_{T_i}(t)}, \tag{4}$$

where $f_{T_i}(t)$ is the density function for the inter-arrival time and $F_{T_i}(t)$ is the corresponding cumulative distribution function. There is always the possibility that some of the pulses may not be detected (e.g. due to overlap with pulses from other trains). Such missed detections may be accounted for in the hazard function calculation by modeling the missed detections as a Bernoulli process with probability of miss p. This results in the following expression for $f_{T_i}(t)$ in (4):

$$f_{T_i}(t) = \sum_{n=1}^{\infty} \frac{1}{\sqrt{2\pi n \sigma_i^2}} e^{-\left[\frac{(t-n\mu_i)^2}{2n\sigma_i^2}\right]} (1-p)p^{n-1}, \tag{5}$$

where μ_i is the mean and σ_i^2 is the variance of the inter-arrival time for the i'th train.

Following [1,2], we estimate the a priori probability of occurrence of a pulse from the i'th train at time t as:

$$P_i(t) = \frac{R_i(t)}{\sum_{j=1}^{M} R_j(t)}, \tag{6}$$

where M is the total number of pulse trains at time t.

2.4. Prototype pulses: initiation and updating

The prototype pulse representing each pulse train needs to be initiated and regularly updated as the characteristics of the trains' pulse shapes evolve over time. Our basic strategy to initiate a prototype pulse for a new pulse train is as follows. When pulse train u_k is selected by the MAP receiver, we find the scalar A which minimizes the error $|\bar{\rho} - A\bar{s}_k|^2$. We then determine the degree of shape mismatch as $XN = |\bar{\rho} - A\bar{s}_k|$. If $|A-1|$ and/or XN are unacceptably large, we initiate a track representing a new pulse train. When neither $|A-1|$ nor XN are unacceptably large, the detected pulse is assigned to pulse train u_k selected by the MAP receiver. In this instance, the prototype pulse $\bar{s}_k$ is updated by a weighted average of $\bar{s}_k$ and the corresponding portion of $\bar{p}$.

2.5. Resolution of pulse overlap

As illustrated in Fig. 2, it is possible in some instances for the received data $\bar{p}$ to correspond to significantly overlapped pulses of two or more pulse trains. Such a situation typically gives rise to a false pulse train initiation. To avoid such unnecessary initiations, we may perform a search for combinations of prototype pulses that can account for the received data. This is part of our ongoing research.

So far, we have found the pulse overlap resolution strategies in [1,2] to have unacceptably large error rates for the type of signal data reported in section 4. Furthermore, these strategies are computationally expensive because they search through all possible combinations of prototype pulses to account for the signal data. We have focused instead on ruling out the possibility of pulse-overlap by using heuristics involving pulse duration and pulse energy. By minimizing the number of times pulse-overlap resolution is performed, our system gains in accuracy and speed of execution.

3. IMPLEMENTATION

The implementation of our system is within an object-oriented framework, in C++. We have utilized the IPUS C++ Platform (ICP) [5] for flexibly coordinating the actions of the initial pulse-train tracking program with that of other programs (under current development) for extending/repairing/merging tracks. The initial pulse-tracking procedure, whose prototype implementation is now complete, is partitioned into independent stages – pulse detection, pulse overlap resolution, prototype update, etc. Here we have employed a state machine architecture, with the different stages of pulse tracking implemented by different classes. The separate stages are thus decoupled and can be modified in isolation. Unlike typical state machine architectures, the state transitions are triggered by intermediate data processing results rather than external events. For example, the identity of the next state in one case depends upon whether or not a significant peak is detected within a signal region.

Our system is able to handle multi-channel signal data with up to 9 channels and it allows the user to specify the sampling rate and various parameters associated with the MAP receiver, prototype initiation, prototype update etc. It also saves intermediate results (such as the evolution of pulse prototypes) for use in later processing. For example, we have developed a program to fill gaps (due to misclassifications) in the initial tracks by carrying out <u>anti-causal</u> tracking that uses selected pulse prototypes from the initial run.

4. RESULTS

The initial implementation of our EMG decomposition program (NEW-DECOMP) has already out-performed the best decomposition program (ORIG-DECOMP) predating it. To illustrate this, we display the NEW-DECOMP output in Fig. 3 and the ORIG-DECOMP output in Fig. 4 when the input to each program is a three-channel intramuscular EMG signal, approximately 65 seconds long and sampled at 50KHz. Each of these figures uses dots to represent inter-pulse intervals (IPI's) in milliseconds for each detected track. Also shown is the level of force applied by the subject, which initially ramps up to 50% maximum voluntary contraction (MVC) for approximately one second, then decreases to 20% MVC for approximately 50 seconds. For comparison purposes, we have displayed in Fig. 5 the result of modifying the ORIG-DECOMP output after extensive manual inspection of the multi-channel input data by an experienced operator. In essence, Fig. 5 may be viewed as representing the desired output for the given input signal. Even a cursory visual inspection indicates that the NEW-DECOMP output in Fig. 4 is closer to the desired output than the ORIG-DECOMP output in Fig. 3. At a 0.6% incorrect classification rate, ORIG-DECOMP detects and correctly classifies approximately 70% of the total number of pulses, while NEW-DECOMP detects and correctly classifies approximately 90% of the total number of pulses. While these percentages are specific to one particular input signal, we have consistently obtained better performance from NEW-DECOMP for other input signals as well.

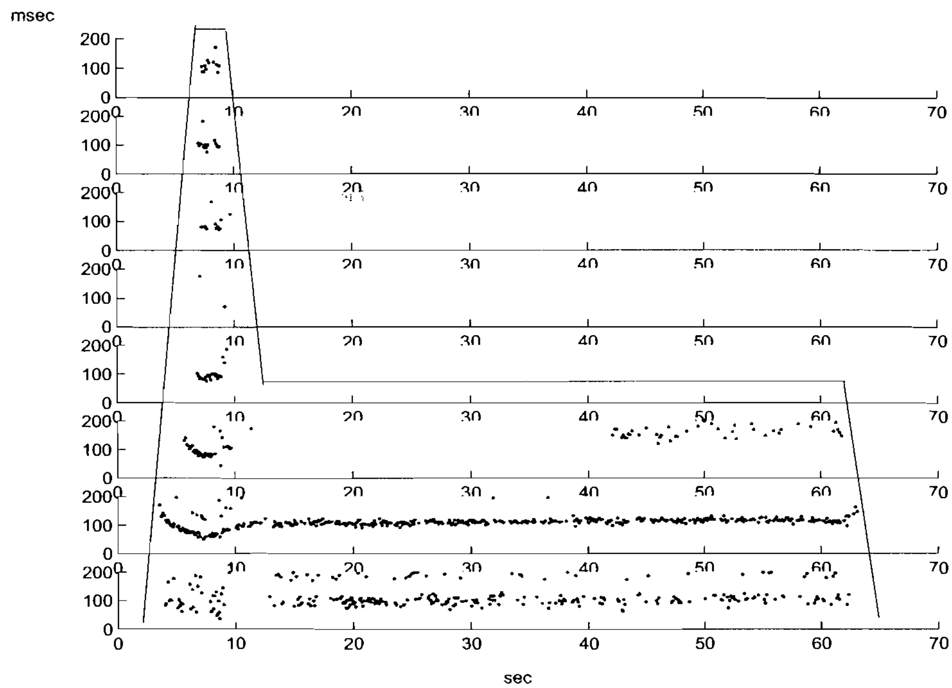

Fig. 3 Results from ORIG-DECOMP

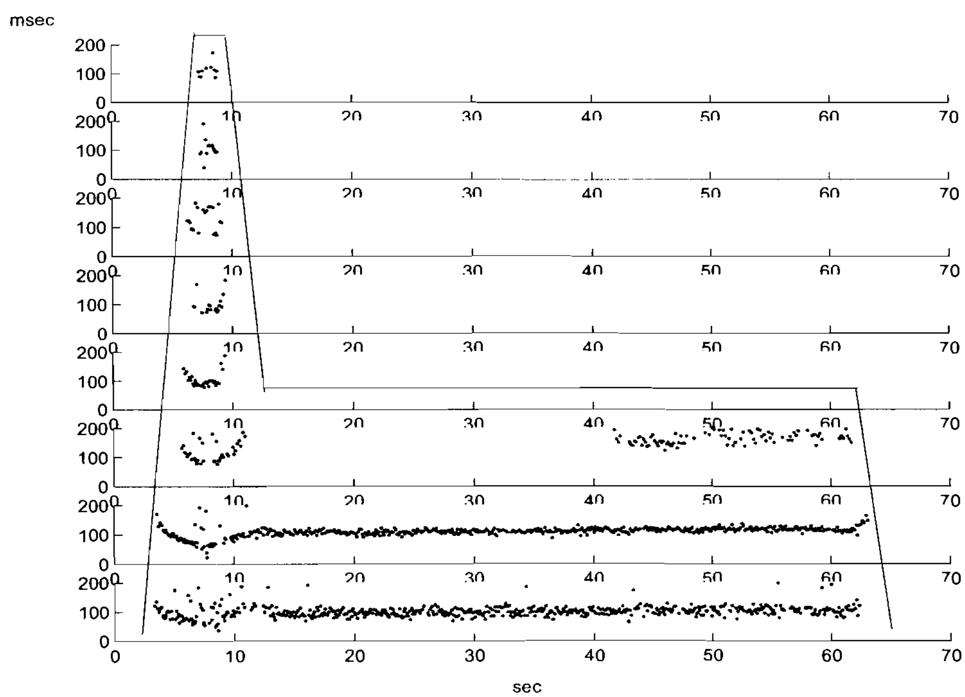

Fig. 4 Results from NEW-DECOMP

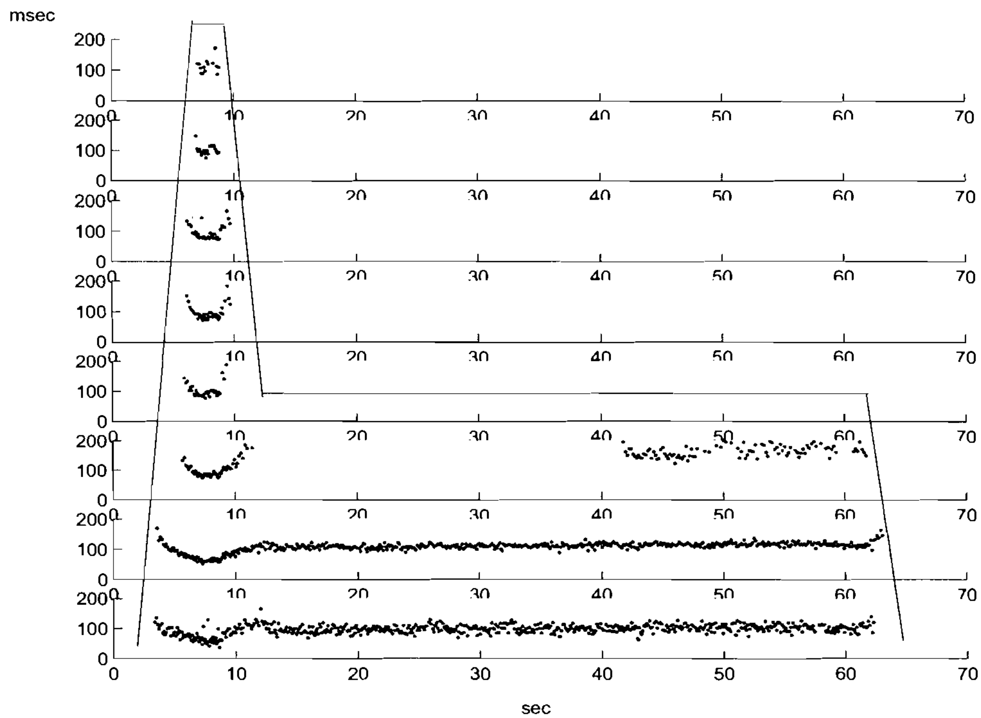

Fig. 5 Manually edited results

5. CONCLUSION

We have described a signal processing framework and its initial software implementation to overcome the limitations of previous systems for the decomposition of intramuscular EMG signals. Major signal processing changes already incorporated with respect to the previously most successful approach are in the following areas: (1) pulse detection, (2) alignment of data pulse and MAP prototype pulse, and (3) the resolution of pulse overlap. Although we are continuing research on improving various signal processing aspects, our system is already providing superior performance with respect to previous systems for real EMG signals such as the one considered in this paper.

6. ACKNOWLEDGEMENTS

We would like to thank Alexander Adam, Prof. Zeynep Erim, and Shane Migliore of the NeuroMuscular Research Center, Boston University, for their advice and contributions to this research.

7. REFERENCES

[1] R. S. LeFever, "Statistical Analysis of Concurrently Active Human Motor Units," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, February 1980.

[2] R. S. LeFever, and C. J. De Luca, "A Procedure for Decomposing the Myoelectric Signal Into Its Constituent Action Potentials – Part I: Technique, Theory, and Implementation," *IEEE Trans. Biomed. Eng.*, BME-29:149-157, 1982.

[3] R. S. LeFever, A. P. Xenakis, and C. J. De Luca, "A Procedure for Decomposing the Myoelectric Signal Into Its Constituent Action Potentials – Part II: Execution and test for accuracy," *IEEE Trans. Biomed. Eng.*, BME-29:158-164, 1982.

[4] H. Broman, "Knowledge-Based Signal Processing in the Decomposition of Myoelectric Signals," *IEEE Engineering in Medicine and Biology Magazine*, June 1988.

[5] J. M. Winograd, and S. H. Nawab, "A C++ Software Environment for the Development of Embedded Signal Processing Systems," *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, 4:2715-2718, 1995.