

**Cross-Comparison of Three Electromyogram
Decomposition Algorithms Assessed with Simulated and
Experimental Data**

by

Chenyun Dai

A Thesis

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Master of Science

in

Electrical and Computer Engineering

by

March 2013

APPROVED:

Dr. Edward A. Clancy, Major Advisor

Abstract

High quality automated electromyogram (EMG) decomposition algorithms are necessary to insure the reliability of clinical and scientific information derived from them. In this work, we used experimental and simulated data to analyze the decomposition performance of three publicly available algorithms—EMGLAB [McGill et al., 2005] (single-channel data only), Fuzzy Expert [Erim and Lin, 2008] and Montreal [Florestal et al., 2009]. Comparison data consisted of quadrifilar needle EMG from the tibialis anterior of 12 subjects (young and elderly) at three contraction levels (10, 20 and 50% MVC), single-channel clinical EMG from the biceps brachii of 10 subjects, and matched simulation data for both electrode types. Performance was assessed via *agreement* between pairs of algorithms for experimental data and *accuracy* with respect to the known decomposition for simulated data. For the quadrifilar data, median agreements between the Montreal and Fuzzy Expert algorithms at 10, 20 and 50% MVC were 95.7, 86.4 and 64.8%, respectively. For the single-channel data, median agreements between pairs of algorithms were 94.9% (Montreal vs. Fuzzy Expert) and 100% (EMGLAB vs. either Montreal or Fuzzy Expert). Accuracy on the simulated data exceeded this performance. Agreement/accuracy was strongly related to trial Complexity, as was motor unit signal to noise ratio, Dissimilarity and Decomposability Index. When agreement was high between algorithm pairs applied to the simulated data, so was the individual accuracy of each algorithm.

Acknowledgements

First and foremost, I shall greatly thank my academic advisor, Dr. Edward A. Clancy. He is not only a respectable and responsible person, but also provides valuable guidance and support to my research. Besides, I am grateful to him for recommending me as his PhD student. I hope I can have a very pleasant collaboration with him on my PhD life.

Thanks to three coauthors of two papers, Anita Christie, Paolo Bonato, Kevin C. McGill for providing data and many good suggestions.

Thanks to my senior alumnus Lukai Liu, my senior alumna Pu Liu and my partner Yejin Li. They give me a lot of help on my research.

Thanks to my family. My mother Zhiyuan, the most important person for me, gives me life, love and whatever I want unconditionally. Also, thanks go to my 'wife' Tachibana Kanade, a person who has changed my life.

TABLE OF CONTENTS

Chapter 1: Introduction	5
Chapter 2: Conference Paper Draft	14
Chapter 3: Journal Paper Draft	16
Appendix: A Report Regarding EMG Decomposition.....	24

CHAPTER 1—INTRODUCTION

Contribution: This whole project is a team project. Coauthors Christie, Clancy and McGill of the two enclosed papers provided experimental data and the three EMG decomposition algorithms (via the MATLAB toolbox “EMGlab”). I mainly focused on running different algorithms on experimental data and simulated data and also generated the simulator EMG signal. Lejin Li, my research partner, mainly focused on result fitting and analysis. Most of the work, in fact, was finished together under the direction of Dr. Clancy. Yejin and I wrote the entire Appendix, which details the research methods and results. Dr. Clancy drafted the journal paper based on this report. Yejin and I drafted the conference paper based on our report and an early draft of the journal paper. Drs. McGill, Christie and Bonato advised the project remotely and edited the conference and journal manuscripts. Dr. McGill is the primary author of one of the decomposition algorithms and Dr. Christie had previously collected one of the data sets. Dr. Bonato provided an additional data set that was not included in the final project.

Main contents of thesis: During my Master’s study, my major is electrical and computer engineering and the field I focus on is biomedical signal processing. All study and research are under the instruction from my advisor Dr. Edward Clancy. Human tissues can generate very weak voltages; a goal of biomedical signal processing is to collect and analyze such weak signals. Most of my work is related with electromyogram (EMG) signal processing. The electromyogram is the electrical activity of human skeletal muscles and has several important functions for diagnosing and treating muscle diseases. This thesis mainly focuses on the performance and reliability of EMG signal decomposition results.

EMG signal generation: Our muscle consists of many small units called motor units (MU). The motor unit includes two parts—one is the motor nerve and another is innervated muscle fibers (see Figure 1). When a muscle contracts, individual motor units in our muscles electrically discharge (see Figure 2). An electrical “motor unit action potential (MUAP)” can be recorded. The average frequency of discharging is called the firing rate. If one motor unit is activated, its initial firing rate is about 4–10

pulses/sec. When force increases, the firing rates can increase up to 20 pulses/sec, or higher. EMG signal recordings are the sum of voltages due to each active motor unit. Typically, many motor units can be active at the same time. Firing times are generally uncorrelated in time within one motor unit and uncorrelated across motor units. Since different motor units generate signals with different shapes and each (healthy) motor unit generates similar shapes each excitation, EMG signal decomposition becomes possible and useful. The purpose of EMG signal decomposition is to separate the composite interference pattern into its constituent motor unit (MU) firing times, permitting the evaluation and study of individual MU firing patterns and action potential shapes (see Figure 3).

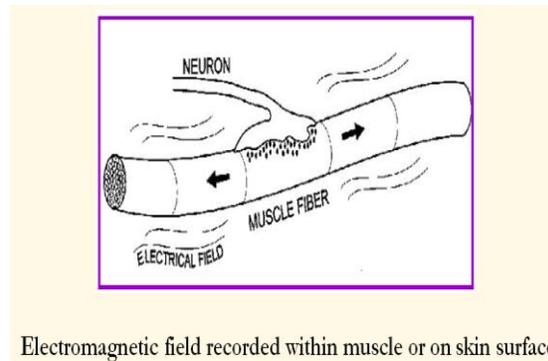
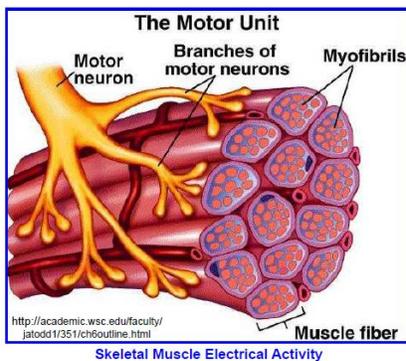


Figure (1): the structure of the motor unit

Figure (2): record EMG signal generated by motor unit

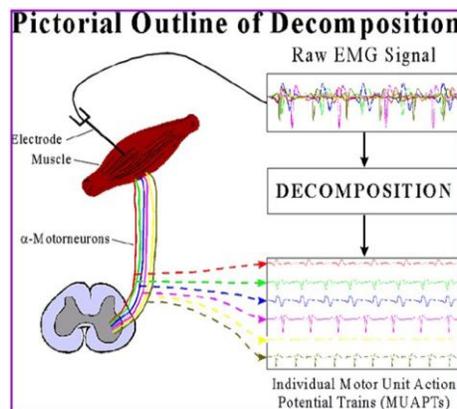


Figure (3): schema of EMG signal decomposition

EMG data collection: Since the signal from our body is always very weak and with high background noise, we need to collect and process raw EMG signals carefully. The data used in this paper come from two parts: experimental data and simulated data:

1. For experimental data, we used both multi-channel data previously acquired at the University of Massachusetts (UMass) and single channel data available from EMGLab.net.

For the UMass data, recordings were acquired from a total of 16 subjects covering a variety of ages (classified as young and old), genders and contraction levels—10%, 20% and 50% of maximum voluntary contraction (MVC). Three channels data of EMG data were simultaneously acquired using a quadrifilar needle electrode and multi-channel decomposition was utilized. Exclusion criteria were first applied for all data, based on the level of noise and duration of stable activity. Finally 12 subjects—7 young (3 males and 4 females) and 5 old (2 males and 3 females) were retained as the usable recordings for further processing.

For EMGLab.net data, we had available the N2001 database of clinical signals which consisted of various subject groups such as one normal control group, one group of patients with myopathy and one group of ALS patients. We only used data from the normal control group for our single channel decomposition. Our sample group consisted of 10 normal subjects aged from 21 to 37 years old, 4 females and 6 males. Each subject had 15 recordings at low-level contraction and another 15 at moderate level. Since the low level recordings were at a very low contraction level and too easy to analyze, we only chose one moderate level data recording from each subject according to background noise and complexity.

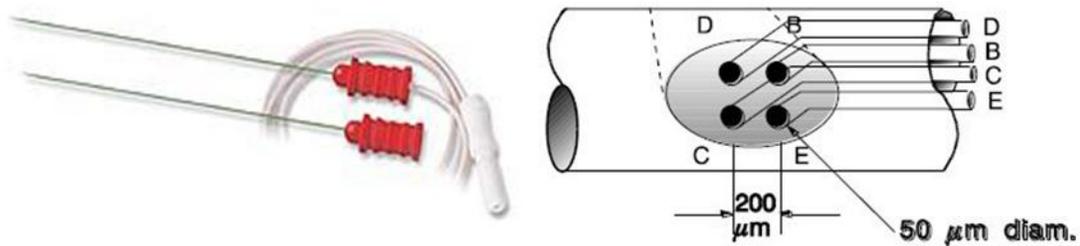


Figure (4): two different types of needles -concentric (single-channel) needle (left) and quadrifilar needle (right).

2. A physiologically based simulation of clinical EMG signals had been developed and was publically available. We designed the simulated data to be similar to the UMass data for multi-channel analysis and similar to the EMGlab.net data for single channel analysis.

The application of decomposition: Decomposition is useful in a wide range of clinical and scientific studies of the neuromuscular system. The number of motor units for a normal muscle in general does not change. And the shape of a healthy motor unit action potential of a muscle also does not change (always similar, but not exact), excluding changes due to fatigue. (Fatigue was avoided when data for this study were collected.) However, if a muscle is diseased (example.g., myopathic diseases), the number of motor units can be reduced and the average diameter of the fibers in motor units can decrease. Disease will change the structure of motor units so that the shapes of motor unit action potentials detected and recorded will change. Not only muscle diseases, but also other effects, such as age and fatigue can cause a similar change in action potential shape. Thus, if we can decompose the original EMG signals into different motor units and characterize the changes of these motor units, corresponding diagnosis and treatment can be implemented. For most decomposition-based studies, an automated algorithm is utilized to perform most of the decomposition, with expert manual editing often completed thereafter. Methods for automated decomposition were pioneered by DeLuca and colleagues [11]. Since that time, a number of other significant approaches and variations have been developed and refined.

General algorithm for EMG decomposition: Since signals generated from our human body are weak and acquired in the presence of high background noise, we need to process raw signals carefully before we decompose. When an EMG signal is acquired, amplification and filtering are common and efficient ways to eliminate background noise and to enlarge the power of the weak signal. All such work is called preprocessing. However all parameters of the preprocessing should be chosen in a very careful way since too much filtering may not lead to an expected high SNR (Signal-to-Noise Ratio). Another challenge is that different motor units are usually active at the same time. This will cause superimposition in the recorded signal. How to deal with superimposition and how to separate the interference signal into several small motor units is the most challenging issue of decomposition. The core concept of most decomposition algorithms is to classify different motor units into clusters based on templates. Like all other algorithms, there is always a trade-off between the performance and the computing time. In this project, different settings such as firing rates statistics and possible combinations (possible number of concurrently active motor units considered when resolving a superimposition) had a great impact on the computing time especially for the Fuzzy Expert algorithm of Erim and Lin.

The primary steps to a classical decomposition algorithm are: pre-processing, detection, clustering and superimposition resolution [1–3, 11]. As noted above, the primary preprocessing step is the application of a highpass filter. The goal of this filter is to accentuate the differences between motor unit spikes, which primarily are found in the higher frequencies. For detection, a simple threshold detector is most common. If the threshold is set too high, motor unit spikes can be missed; if the threshold is set too low, noise spikes can be detected. Clustering is then used to associate the various spikes with motor units. Generally, spikes are only clustered if their shape sufficiently matches the template shape, so as to limit clustering of noise spikes. In addition, every superimposed spike tends to have a different shape. Thus, the general clustering stage tends to purposely not desire to classify superimposed shapes. Many algorithms perform clustering in multiple passes. During the first pass, the most similar spikes are clustered, after which robust templates are formed. During subsequent passes, the templates are improved as more units are added. Finally, superimposition resolution is performed on the unclustered spikes. Several techniques are possible. The simplest technique is an exhaustive search

method of trying all combinations of two or more templates at all possible relative time displacements. Unfortunately, this technique tends to be prohibitively time intensive, particularly when testing superimpositions of three or more templates. Note that most algorithms will not classify all detected spikes. Many variants to this classical algorithm have been developed.

In this project, three of the major decomposition algorithms are now publically available within the MATLAB software environment—the McGill algorithm [1], the Fuzzy Expert algorithm [2] and the Montreal (MTL) algorithm [3]. In addition, a detailed indwelling EMG simulator is also publically available [4]. Hence, we cross-compared the performance of these three algorithms utilizing a variety of experimental and simulated EMG needle data.

Before decomposition, each signal was high-pass filtered in order to improve the accuracy of results. The reason to use a high-pass filter is that the signal information at frequencies less than 500 Hz to 1000 Hz tends to look rather similar for all motor units. But, the higher frequency content is more discriminable. When the shapes of different motor units are different, decomposition algorithms can distinguish them much more easily. Therefore, we eliminate the lower frequency portion of the signal. However, after we filter this low frequency part, the spikes of the motor units become smaller, which decreases the SNR. Thus, we need to choose a suitable cut-off frequency between 500 Hz and 1000 Hz carefully to keep both good SNR and distinguishability. For the UMass multi-channel data, an analog high-pass filter (1000 Hz) had been applied before digitizing/ Since some residual low frequency background noise existed, a 1st-order zero-phase Butterworth digital high-pass filter with 100 Hz cut off frequency was used. For the EMGlab.net database, the single channel signal was processed in analog hardware (prior to sampling) by a first-order high-pass filter with 2 Hz cut off and low-pass filter with 10 kHz cut off. We then used a 500 Hz first-order zero-phase Butterworth high-pass digital filter.

Each single channel signal was decomposed separately by three algorithms and each multi-channel signal by MTL (Montreal) and Fuzzy Expert. All three algorithms automatically detected spikes of motor units and established their discharge times in the signal. In order to compare different algorithms under the same circumstance, the results of these algorithms after decomposition were saved in a

uniform format, including discharging time, motor unit ID number and channel for each spike in the signal in an annotation file (.eaf file).

The parameters of Fuzzy Expert should be set carefully or the computing time would be long and inefficient. Most of the parameters can be chosen as default settings. We modified some key parameters as below: a) passes =10; b) Min Template to Fill = 0.2; c) Max MU Combo for super-position: 3 for 1st and 2nd pass, 5 for 3rd pass and 6 for the rest.

The performance of automated decomposition algorithms (emphasis of thesis): The performance of automated decomposition algorithms has primarily been evaluated in a few manners [5]. First, “reference” or “truth” annotations have been achieved via manual expert editing of an experimental data set [3], [4], [6], [7]. This technique can be extremely time consuming (e.g., one hour per second of data for Fuzzy Expert) and its own accuracy is difficult to assess. Nonetheless, assessment on experimental data guarantees signal conditions representative of actual use. Second, some experimental data sets have been evaluated manually, but the evaluation has been limited to clinical classification of each MU as normal vs. abnormal [8], [9]. This manual evaluation is much more time efficient, but does not quantitatively assess the intermediate algorithm steps of spike detection and spike classification. Third, EMG signals have been simulated [3], [7], [8], [10]–[12]. In this case, the truth annotations are known to be correct. However, even highly detailed simulations cannot guarantee all of the complexities of an actual signal. Fourth, a few studies have recorded EMG from multiple indwelling needles, each of which is decomposed [7], [13]–[15]. Some of the MUs recorded from the distinct electrodes are common. Agreement in their firing times is strong evidence of correct detection and classification of those firings. Recent studies have also compared indwelling decomposition to that accomplished by surface EMG arrays [16]–[18]. Most commonly, a combination of evidence—experimental and simulation—is used to evaluate an algorithm, as each evaluation technique has its own strengths and weaknesses.

To date, very little direct comparison has been made between the performances of various automated algorithms [19]. Such comparison is important, since the reported performance of an

algorithm is a strong function of the data used for evaluation. Recordings are known to be more difficult to decompose, for example, when: more spikes occur per second, distinct MUs exhibit similar shape, the signal-to-noise ratio (SNR) is low, MU shapes change over time and firing times are irregular [7]. Hence, direct comparison between reported algorithm accuracies is confounded. In addition, further support is given to the efficacy of decomposition, in general, if multiple algorithms are able to arrive at common solutions.

For experimental data and simulated data, we developed comparison based on *agreement* among different algorithms and *accuracy* based on truth annotations, respectively. We also computed four measures of decomposition difficulty. High agreement and accuracy versus these measures would reflect the reliability of the automated decomposition algorithms. Details of the difficulty measures—SNR, DI, Dissimilarity and Complexity—are provided in the journal paper draft (below).

The remainder of this thesis is structured as follows. Chapter 2 is the conference paper draft accepted to the 2013 IEEE 39th Annual Northeast Bioengineering Conference [20], which only presents the cross-comparison between the two multi-channel EMG decomposition algorithms based on DI, due to the page limitation. Chapter 3 is the draft of the journal paper, which includes a broader range of the work including both multi-channel and single channel comparison using all three decomposition algorithms. Appendix is the report regarding EMG decomposition, which presents all the detailed information, such as intermediate steps of the processing results and unmodified figures.

REFERENCES IN INTRODUCTION

- [1] McGill KC, Lateva ZC, Marateb HR, “EMGLAB: An interactive EMG decomposition program,” J Neurosci Meth, vol. 149, pp. 121–133, 2005.
- [2] Erim Z, Lin W, “Decomposition of intramuscular EMG signals using a heuristic fuzzy expert system,” IEEE Trans Biomed Eng, vol. 55, pp. 2180–2189, 2008.
- [3] Florestal JR, Mathieu PA, McGill KC, “Automatic decomposition of multichannel intramuscular EMG signals,” J Electromyogr Kinesiol, vol. 19, pp. 1–9, 2009.
- [4] Hamilton-Wright A, Stashuk DW, “Physiologically based simulation of clinical EMG signals,” IEEE Trans Biomed Eng, vol. 52, 171–183, 2005.
- [5] Farina D, Colombo R, Merletti R, Olsen HB, “Evaluation of intra-muscular EMG signal decomposition algorithms,” J Electromyogr Kinesiol, vol. 11, pp. 175–187, 2001.
- [6] Ge D, Le Carpentier E, Farina D, “Unsupervised bayesian decomposition of multiunit EMG recordings using tabu search,” IEEE Trans Biomed Eng, vol. 57, pp. 561–571, 2010.

- [7] Nawab SH, Wotiz RP, DeLuca CJ, “Decomposition of indwelling EMG signals,” *J Appl Physiol*, vol. 105, pp. 700–710, 2008.
- [8] Chauvet E, Fokapu O, Hogrel JY, Gamet D, Duchene J, “Automatic identification of motor unit action potential trains from electromyographic signals using fuzzy techniques,” *Med Biol Eng Comput.*, vol. 41, pp. 646–653, 2003.
- [9] Christodoulou CI, Pattichis CS, “Unsupervised pattern recognition for the classification of EMG signals,” *IEEE Trans Biomed Eng*, vol. 46, pp. 169–178, 1999.
- [10] Fang J, Agarwal GC, Shahani BT, “Decomposition of multiunit electromyographic signals,” *IEEE Trans Biomed Eng*, vol. 46, pp. 685–697, 1999.
- [11] LeFever RS, DeLuca CJ, “A procedure for decomposing the myoelectric signal into its constituent action potentials—Part I: Technique, theory, and implementation,” *IEEE Trans Biomed Eng*, vol. 29, pp. 149–157, 1982.
- [12] Zennaro D, Wellig P, Koch VM, Moschytz GS, Laubli T, “A software package for the decomposition of long-term multichannel EMG signals using wavelet coefficients,” *IEEE Trans Biomed Eng*, vol. 50, pp. 58–69, 2003.
- [13] DeLuca CJ, “Reflections on EMG signal decomposition,” in *Computer-Aided Electromyography and Expert Systems*, Ed. JE Desmedt, Elsevier, 1989, pp. 33–37.
- [14] Mambrito B, DeLuca CJ, “A technique for the detection, decomposition and analysis of the EMG signal,” *EEG Clin Neurophysiol*, vol. 58, pp. 175–188, 1984.
- [15] McGill KC, Lateva ZC, Johanson ME, “Validation of a computer-aided EMG decomposition method,” in *IEEE Eng Med Biol Conf*, 2004, pp. 4744–4747.
- [16] DeLuca CJ, Adam A, Wotiz R, Gilmore LD, Nawab SH, “Decomposition of surface EMG signals,” *J Neurophysiol*, vol. 96, pp. 1646–1657, 2006.
- [17] Holobar A, Minetto MA, Botter A, Negro F, Farina D, “Experimental analysis of accuracy in the identification of motor unit spike trains from high-density surface EMG,” *IEEE Trans Neural Sys Rehab Eng*, vol. 18, pp. 221–229, 2010.
- [18] Marateb HR, McGill KC, Holobar A, Lateva ZC, Mansourian M, Merletti R, “Accuracy assessment of CKC high-density surface EMG decomposition in biceps femoris muscle,” *J Neural Eng*, vol. 8, 066002, 2011.
- [19] McGill KC, “A comparison of three quantitative motor unit analysis algorithms,” *Suppl Clin Neurophysiol*, vol. 60, 273–278, 2009.
- [20] Li Y, Dai C, Clancy EA, Bonato P, Christie A, McGill KC, “Cross-comparison between two multi-channel EMG decomposition algorithms assessed with experimental and simulated data,” in *IEEE Northeast Bioeng Conf*, 2013, in press.

CHAPTER 2 Conference paper (in press)

Cross-Comparison Between Two Multi-Channel EMG Decomposition Algorithms Assessed with Experimental and Simulated data

Yejin Li, Chenyun Dai, Edward A. Clancy
ECE Dept., Worcester Polytechnic Institute, USA
yli1@wpi.edu, cdai@wpi.edu, ted@wpi.edu

Paolo Bonato, Dept. Phys. Med. Rehab.,
Harvard Med. School, USA, pbonato@partners.org

Anita Christie, Dept. Human Physiology,
U. Oregon, USA, adc1@uoregon.edu

Kevin C. McGill, Rehab. R&D Center
VA Palo Alto, USA, kemcgill43@gmail.com

Abstract—The reliability of automated electromyogram (EMG) decomposition algorithms is important in clinical and scientific studies. In this paper, we analyzed the performance of two multi-channel decomposition algorithms—Montreal [1] and Fuzzy Expert [2] using both experimental and simulated data. Comparison data consisted of quadrifiler needle EMG from the tibialis anterior muscle of 12 subjects (young and elderly) at three contraction levels (10, 20 and 50% MVC), and matched simulation data. Performance was assessed via agreement between the two algorithms for experimental data and accuracy with respect to the known decomposition for simulated data. For the experimental data, median agreement between the Montreal and Fuzzy Expert algorithms at 10, 20 and 50% MVC was 95.7, 86.4 and 64.8%, respectively. For the simulation data, median accuracy was 99.8%, 100% and 95.9% for Montreal, and 100%, 98% and 93.5% for Fuzzy Expert at the different contraction levels.

Keywords—EMG; Motor unit potential; Decomposition; SNR; Composite Decomposability Index (CDI); Cross-comparison.

I. INTRODUCTION

Decomposition of EMG signals separates the composite signal into its constituent motor unit (MU) firing times and action potential shapes. Decomposition is used in many clinical and research studies of the neuromuscular system [3]. In most cases, an automatic algorithm and expert manual editing are combined to produce a more reliable result [4].

Evaluating the performance between multiple automated algorithms is crucial, since EMG recordings can be extremely difficult to decompose under conditions of low signal to noise ratio (SNR) and/or high similarity of MU shapes. Hence, if high agreement is achieved between algorithms, more confidence is gained in each decomposition.

Two major multi-channel decomposition algorithms are now publically available within the MATLAB software environment [1], [2]. In addition, a detailed indwelling EMG simulator is also publically available [5]. Hence, we cross-compared the performance of these two algorithms, utilizing a variety of experimental and simulated data at different contraction levels.

II. METHODS

A. Experimental and Simulated Data

The data used in this paper consisted of two parts: the experimental data were recorded at the University of Massachusetts and simulated data were generated. The data reanalysis was approved and supervised by the WPI Institutional Review Board.

Three-channel quadrifiler needle EMG signals were acquired from the tibialis anterior muscle of the dominant leg of seven young (three male, four female) and five elderly (two male, three female), healthy subjects during isometric contractions of 10%, 20% and 50% MVC. Four 50- μ m diameter wires with 200 μ m distance between electrode pairs comprised the recording surfaces. The signals were bandpass filtered from 1,000–10,000 Hz and sampled at 51,200 Hz at 16-bit resolution. One 5s segment of each signal was analyzed. Thus, 36 recordings of 5 s duration each were used (12 subjects \times 3 levels of contraction).

Quadrifiler data were simulated using the EMG needle simulator of Hamilton-Wright [5]. The resulting signals resembled those acquired experimentally as closely as possible including electrode configuration and shape, recording duration, contraction levels and background noise.

B. Automated Decomposition Algorithms

Before decomposition, each recording was highpass filtered. For the experimental data, the acquisition hardware already provided some filtering. Thus, a 1st-order zero-phase Butterworth high-pass filter with 100 Hz cut off frequency was applied digitally. For the simulated data, the same process was implemented, except that the cut off frequency was 500 Hz.

Both algorithms detected EMG spikes, classified spikes with similar shapes and patterns, and resolved superimpositions. The first algorithm was the “Montreal” algorithm [1]. This algorithm has no adjustable parameters. The second algorithm was the “Fuzzy Expert” algorithm [2]. We utilized ten passes and limited resolution of superimpositions to three MUs on the first two passes, five MUs on the third pass and six MUs thereafter.

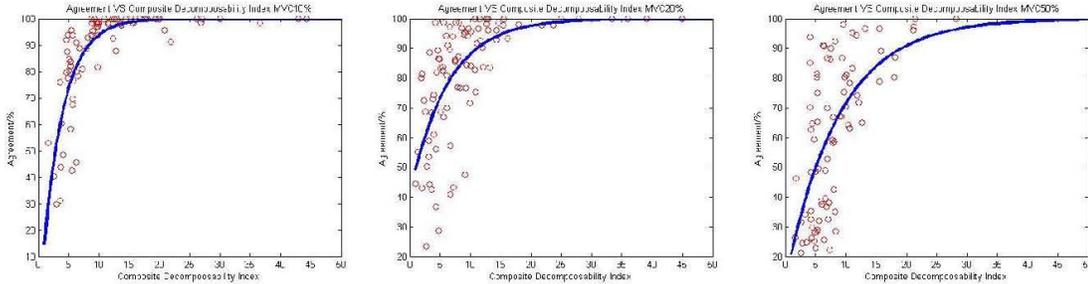


Figure 1. The relationship between Decomposability Index and agreement at MVC 10%, 20% and 50% (Montreal vs. Fuzzy Expert).

C. Methods of Analysis

The experimental data were analyzed separately from the simulated data. After highpass filtering, all experimental and simulated quadrifiler data were automatically decomposed by the Fuzzy Expert and Montreal algorithms.

For each MU in the experimental signals that was identified by both algorithms, the agreement rate was calculated as the percentage of the MU firings on which the two algorithms agreed to within ± 1 ms [6]. For each MU in the simulated signals, the accuracy of each algorithm was calculated as the percentage of firings which the algorithm correctly identified to within ± 1 ms. The "decomposability index" [1] of each MU was calculated as the minimum RMS difference between that MU and any other MU or the baseline, divided by the RMS value of the entire signal. The index was computed from each channel and the norm of the individual indexes reported.

III. RESULTS

Average performance results vs. contraction level are listed in the Abstract. Fig. 1 shows the relationship between agreement rate and decomposability index for the experimental data at each contraction level. For 10% MVC, the number of matched MUs ranged from 7–10 per subject (total of 103 MUs). For 20% MVC, the number ranged from 7–13 (total 110). For 50% MVC, the number ranged from 7–12 (total 120). Fig. 2 shows the accuracy vs. decomposability index for the simulated data at 20% MVC. The number of matched MUs identified by Fuzzy Expert for each subject ranged from 7–13 (total of 117 MUs) and from 7–15 by Montreal (total of 114 MUs).

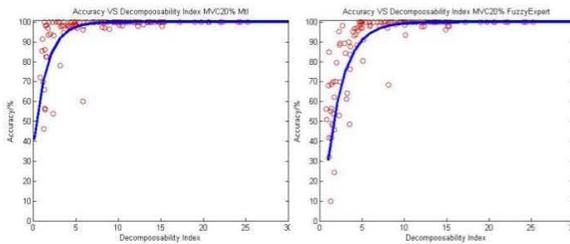


Figure 2. The relationship between Decomposability Index and accuracy at MVC 20% (Montreal vs. Fuzzy Expert).

The data from each plot were then fit to the exponential model: $Agreement = 100 - a \cdot e^{-b \cdot x}$, where x is the Decomposability index of each motor unit.

IV. DISCUSSION

This study evaluated two automated decomposition algorithms when applied to experimental and simulated data. The results show good accuracy and substantial agreement between the algorithms, especially for MUs with a larger decomposability index at lower levels of contraction. These results provide a measure of confidence that the algorithms perform reliably on real EMG signals.

Performance was poorer for MUs with lower decomposability indices and in signals from higher levels of contraction (Figs. 1 and 2). This result is expected, as MU recruitment and firing rates rise with contraction level, producing more overlapping discharges and increasing the likelihood of MU shape similarity.

REFERENCES

- [1] Florestal JR, Mathieu PA, Malanda A. Automated decomposition of intramuscular electromyographic signals. *IEEE Trans Biomed Eng.* 2006;53:832–839.
- [2] Erim Z, Lin W. Decomposition of intramuscular EMG signals using a heuristic fuzzy expert system. *IEEE Trans Biomed Eng.* 2008;55:2180–2189.
- [3] Stashuk D. EMG signal decomposition: How can it be accomplished and used? *J Electromyogr Kinesiol.* 2001;11:151–173.
- [4] Nawab SH, Wotiz RP, DeLuca CJ. Decomposition of indwelling EMG signals. *J Appl Physiol* 2008;105:700–710.
- [5] Hamilton-Wright A, Stashuk DW. Physiologically based simulation of clinical EMG signals. *IEEE Trans Biomed Eng* 2005;52:171–183. [Software available at <http://emglab.net>.]
- [6] McGill KC, Lateva ZC, Johanson ME. Validation of a computer-aided EMG decomposition method. *IEEE Eng Med Biol Conf* 2004;4744–4747.

Cross-Comparison of Three Electromyogram Decomposition Algorithms Assessed with Experimental and Simulated Data

Chenyun Dai, Yejin Li, Anita Christie, Paolo Bonato, *Senior Member, IEEE*, Kevin C. McGill, *Member, IEEE*, and Edward A. Clancy, *Senior Member, IEEE*

Abstract— High quality automated electromyogram (EMG) decomposition algorithms are necessary to insure the reliability of clinical and scientific information derived from them. We used experimental and simulated data to analyze the performance of three publicly available decomposition algorithms—EMGLAB [1] (single channel data only), Fuzzy Expert [2] and Montreal [3]. Comparison data consisted of quadrifilar needle EMG from the tibialis anterior of 12 subjects at 10%, 20% and 50% maximum voluntary contraction (MVC), single channel needle EMG from the biceps brachii of 10 subjects, and matched simulation data for both electrode types. Performance was assessed via agreement between pairs of algorithms for experimental data and accuracy with respect to the known decomposition for simulated data. For the quadrifilar experimental data, median agreements between the Montreal and Fuzzy Expert algorithms at 10%, 20% and 50% MVC were 95.7%, 86.4% and 64.8%, respectively. For the single channel experimental data, median agreements were 94.9% (Montreal vs. Fuzzy Expert) and 100% (EMGLAB vs. Montreal or Fuzzy Expert). Accuracy on the simulated data exceeded this performance. Agreement/accuracy was strongly related to motor unit signal to noise ratio. When agreement was high between algorithm pairs applied to the simulated data, so was the accuracy of each algorithm.

Index Terms—Electromyography (EMG), motor units, decomposition, intramuscular EMG, biomedical signal analysis.

I. INTRODUCTION

DECOMPOSITION of indwelling electromyogram (EMG) recordings is the process of separating the composite interference pattern into its constituent motor unit (MU) firing times, permitting the evaluation and study of individual MU firing patterns and action potential shapes. Decomposition is useful in a wide range of clinical and scientific studies of the neuromuscular system (for reviews, see [4]–[6]. For most

decomposition-based studies, an automated algorithm is utilized to perform most of the decomposition, with expert manual editing often completed thereafter. Methods for automated decomposition were pioneered by DeLuca and colleagues [7], [8]. Since that time, a number of other significant approaches and variations have been developed and refined [2], [4], [9]–[17].

The performance of automated decomposition algorithms has primarily been evaluated in a few manners [18]. First, “reference” or “truth” annotations have been achieved via manual expert editing of an experimental data set [2], [4], [12], [13], [17]. This technique can be extremely time consuming (e.g., one hour per second of data [2]) and the accuracy of the reference annotations can be difficult to assess. Nonetheless, assessment on experimental data guarantees signal conditions representative of actual use. Second, EMG signals have been simulated [8], [9], [11]–[13], [15], [17]. In this case, the truth annotations are known to be correct. However, even highly detailed simulations cannot guarantee all of the complexities of an actual signal. Third, a few studies have recorded EMG from multiple indwelling needles, each of which is decomposed [4], [19]–[21]. Some of the MUs recorded from the distinct electrodes are common. Agreement in their firing times is strong evidence of correct detection and classification of those firings. Recent studies have also compared indwelling decomposition to that accomplished by surface EMG arrays [22]–[24]. Most commonly, a combination of evidence—experimental and simulation—is used to evaluate an algorithm, as each evaluation technique has its own strengths and weaknesses.

To date, very little direct comparison has been made between the performance of various automated algorithms [25]. Since algorithm performance depends on the characteristics of the signal being analyzed, the same set of signals should be used when comparing different algorithms. Signals are known to be more difficult to decompose, for example, when: more spikes occur per second, distinct MUs exhibit similar shape, the signal to noise ratio (SNR) is low, MU shapes change over time and firing times are irregular [4]. Moreover, when multiple algorithms are able to agree on the annotation of a particular signal, it increases confidence that the annotation is objectively correct. Three of the major decomposition algorithms are now publicly available within the MATLAB software

This paragraph of the first footnote will contain the date on which you submitted your paper for review.

C. Dai, Y. Li and E. A. Clancy (corresponding author) are with Worcester Polytechnic Institute, Worcester, MA 01609 USA (e-mail: cdai@wpi.edu, yli1@wpi.edu, ted@wpi.edu).

A. Christie is with the University of Oregon, Eugene, OR 97403 USA (e-mail: adc1@uoregon.edu).

P. Bonato is with Harvard Medical School, Boston, MA 02114 (e-mail: pbonato@partners.org).

K. C. McGill is with the VA Palo Alto Health Care System, Palo Alto, CA 94304 USA (e-mail: kcmcgil143@gmail.com).

environment [1]–[3]. In addition, a detailed indwelling EMG simulator is also publically available [26]. Hence, we cross-compared the performance of these three algorithms utilizing a variety of experimental and simulated EMG needle data.

II. METHODS

A. Experimental Data

Portions of experimental data from two prior studies were reanalyzed, and simulated data were generated. No new subject data were collected. The data reanalysis was approved and supervised by the WPI Institutional Review Committee.

Three-channel quadrifilar needle EMG had been acquired from the dominant leg of seven young (three male, four female; aged 18–30 years) and five elderly (two male, three female; aged 65 years or older), healthy subjects at the University of Massachusetts. Subjects were seated, the upper leg of their preferred limb restrained and the ipsilateral foot secured to a stiff transducer that measured ankle dorsiflexion force. The skin over the tibialis anterior (TA) muscle was cleaned with rubbing alcohol and a 27-gauge four-wire quadrifilar needle electrode was inserted into the belly of the TA muscle, avoiding the innervation zone. The needle was maneuvered into a position from which activity from several MUs could be obtained. Four 50- μm diameter platinum-iridium wires terminating at a side port 7.5 mm from the tip of the electrode comprised the recording surfaces [27]. The four wires in this electrode were arranged in a square array with approximately 200 μm on each side. The signals detected with this needle were connected to three differential amplifiers (10^{12} Ω input resistance; 25 pA bias current), bandpass filtered from 1,000–10,000 Hz and sampled at 25,600 Hz at 16-bit resolution. These data were upsampled off-line by a factor of two to a sampling rate of 51,200 Hz. Lower leg skin temperature was maintained at approximately 34.0° C. Prior to electrode insertion, maximum voluntary contraction (MVC) dorsiflexion force was measured as the average of 3–5 maximum contractions of 5 s duration each. Following electrode insertion, subjects performed contractions at 10%, 20% and 50% MVC, with target force levels displayed on a video monitor. Subjects slowly increased their force to the target level, and then maintained the force while a 30 s recording was made. A rest period of three minutes was provided between each contraction to prevent fatigue. One, 5-s segment of each recording was analyzed. Thus, 36 recordings of 5 s duration each were used (12 subjects \times 3 levels of contraction).

Single channel needle EMG were reanalyzed from ten control subjects (6 males, 4 females; aged 21–37 years) in the publically-available “N2001” database of Nikolic [28]. Recordings were acquired from the biceps brachii muscles during “moderate” level contractions using a concentric needle electrode in accordance with standard clinical recording procedures. The signals were bandpass filtered between 2–10,000 Hz and sampled at 23,437.5 Hz with 16 bit resolution. Ten 5-s recordings (ten subjects \times one recording/subject) were

used for analysis.

For each experimental signal, a “Complexity” measure was computed, expressing the number of MU firings per second. Within the analyzed 5 s segment of each recording, the number of pulses exceeding the background noise was manually counted. Spikes of duration greater than 3 ms, representing superimpositions, were counted as two pulses. Those with duration greater than 6 ms were counted as three pulses, etc. For multiple-channel data, all three channels were simultaneously viewed. The Complexity measure was expressed in pulses per second (pps). Complexity measures from the experimental data were used to guide generation of the simulation data.

B. Simulation Data

Quadrifilar and single channel data were simulated using the publically-available EMG needle simulator of Hamilton-Wright and Stashuk [26]. The resulting signals closely resembled those acquired experimentally. The simulation parameters modeled the physical layout of the muscle, MU firing patterns, action potential propagation and type of EMG electrode. To emulate quadrifilar recordings, four noise-free monopolar tip electrodes (50 μm diameter) were simultaneously simulated in a square array configuration at 200 μm distances. This configuration mimics a quadrifilar needle. The three differential voltages were then computed offline in MATLAB, and 20 dB of white Gaussian noise was added. For each experimental contraction level to be simulated, trial and error was used to determine the contraction level parameter input value of the simulator software such that the average Complexity of the simulated data matched the average Complexity of the corresponding experimental data. Five-second recording segments were created at force levels representing 10%, 20% and 50% MVC. Simulation was iterated 12 times, providing 12 realizations, to give the same number of subjects as with the quadrifilar experimental data. The truth time instances and identities of each MU firing (i.e., MU annotations), which are fully known in a simulation, were recorded along with the simulated signals (sampled at 31,250 Hz, 16-bit resolution). To emulate single channel recordings, one 10 mm concentric electrode was simulated with 20 dB of white Gaussian noise added. The Complexity of these simulations was matched to the average Complexity of the single channel needle (N2001) data set, again via selection of the contraction level parameter input value of the simulator software. Ten recordings, each of 5 s duration, were created at a sampling rate of 31,250 Hz with 16-bit resolution, along with the truth MU annotations.

C. Automated Decomposition Algorithms

Three publically-available decomposition algorithms were compared. Each is implemented in MATLAB, which was used for all computation. Each algorithm was used *without* manual editing, although such editing is the norm in scientific studies. Prior to automated decomposition, the quadrifilar experimental data were highpass filtered at 100 Hz and the single channel

experimental data at 500 Hz. Different filters were used because the hardware filtering differed between the data sets. Both simulated data sets were highpass filtered at 1,000 Hz. In all cases, a first-order Butterworth filter was designed, and then applied in the forward and reverse time directions to achieve zero phase shift.

All three algorithms detected EMG spikes, classified spikes with similar shapes and resolved superimpositions. The first automated decomposition algorithm was the default algorithm implemented in the publically-available “EMGLAB” software [1]. This algorithm can only analyze single channel EMG data and thus was only used for our single channel data. The second algorithm was the “Montreal” algorithm [3]. This algorithm has no adjustable parameters. The third algorithm was the “Fuzzy Expert” algorithm [2]. With Fuzzy Expert, we utilized ten algorithm passes and limited resolution of superimpositions to three MUs on the first two passes, five MUs on the third pass and six MUs thereafter.

D. Methods of Analysis

After highpass filtering (described above), all experimental and simulated quadrifilar data were automatically decomposed by the Fuzzy Expert and Montreal algorithms. The single channel experimental and simulated data were decomposed by all three automated algorithms. Decompositions of *experimental* signals were compared pair-wise between algorithms for each signal. Each MU annotation was said to match if both algorithms found the same MU within a ± 1 ms match window, after adjusting for the difference in MU registration locations between the different algorithms [18], [24]. “Agreement” was measured as the number of matched annotations, divided by the sum of: (1) matched annotations and (2) unmatched annotations from either algorithm. Agreement results were expressed in percent. Results are only presented for those MUs that exhibited a minimum of 20 matches (average of 4 matches per second over a 5 s recording duration). Decompositions of *simulated* signals were also compared directly to the truth annotations in a similar manner, this result being denoted “Accuracy,” since the true annotations were known.

For each identified MU, a signal to noise ratio (SNR_{MU}) was computed from one EMG channel as the peak-to-peak height of the MU divided by the RMS value of the entire signal [21]. A ten-bin histogram of all (negative/positive) peak values from all firings of a MU was computed. A peak value was estimated as the average height of all values contributing to the histogram mode bin. This selection helps to reject peak values that might be unrepresentative due to MU superimpositions. For multiple-channel data, the SNR_{MU} was computed separately for each channel and then averaged. The SNR_{MU} is non-dimensional. For experimental signals, SNR_{MU} was computed multiple times, using the annotations from each respective decomposition algorithm. For simulated signals, the measures were also computed using the truth annotations. Cross-plots of SNR_{MU} vs. agreement (or accuracy) were created for each contraction level for each data set. The data from each plot were then least squares fit to the exponential model:

$Agreement = 100 - a \cdot e^{-b \cdot SNR_{MU}}$. Performance differences were tested statistically using one-way ANOVA, with ANOVAs also used for pair-wise *post hoc* testing.

III. RESULTS

Table I lists the number of MUs analyzed in the four data sets. General statistical comparisons of agreement and accuracy are shown in Fig. 1 for each of the four data sets. Fig. 1 also indicates statistically significant differences in results from one-way ANOVA comparisons. Agreement/accuracy generally decreased with MVC level for the multiple-channel data.

Complexity values for the quadrifilar experimental data at 10%, 20% and 50% MVC were 100.1 ± 49.8 , 119.3 ± 46.4 and 211.8 ± 54.6 pps, respectively. An ANOVA showed a significant difference in Complexity between MVC levels [F(2,33)=16.9, $p < 10^{-5}$], with pair-wise *post hoc* ANOVAs showing that Complexity was significantly higher at 50% MVC ($p < 2 \times 10^{-4}$). Complexity values for the corresponding quadrifilar simulation signals at 10%, 20% and 50% MVC were 99.2 ± 21.3 , 120.7 ± 25.5 and 215.9 ± 59.4 pps, respectively. Complexity values for the single channel experimental data were 61.8 ± 19.8 pps, while the corresponding values for the single channel simulations were 62.6 ± 12.3 pps. Hence, the average experimental and simulation Complexity values were quite well matched, as designed.

Fig. 2 shows *agreement* results (Montreal vs. Fuzzy Expert) vs. SNR_{MU} for the experimental quadrifilar data. Figs. 3 and 4 show *accuracy* vs. SNR_{MU} for the simulated quadrifilar data, both as a function of MVC level and combined across levels. Similarly, agreement and accuracy results for experimental and simulated single channel data are shown vs. SNR_{MU} in Fig. 5. Each plot also shows the best-fit exponential model. Quantitatively, it is desired that SNR_{MU} provide an association with agreement/accuracy. Here, that relation is expressed by the goodness-of-fit of the exponential model, also listed in the figures. In general, agreement/accuracy increased with SNR_{MU} , as shown by the decaying exponent in the exponential model.

Finally, Fig. 6 shows agreement between algorithm pairs and their individual accuracies for the quadrifilar simulations and, separately, for the single channel simulations. Each paired agreement value (paired between decomposition algorithms) is plotted twice, once corresponding to each individual accuracy value. The vast majority of accuracy values are higher than their corresponding agreement values. Note that most of the plotted values are clustered in the upper right corner of each plot, with multiple values over-plotted. The “tails” extending towards the origin primarily depict the limited number of low accuracy/agreement values.

IV. DISCUSSION

This study evaluated the *agreement* between pairs of automated decomposition algorithms when applied to experimental data, as well as the *accuracy* of these algorithms when applied to simulated data. A large subject pool and wide range of contraction levels (10%, 20% and 50% MVC) was considered, as well as quadrifilar and single channel electrode

recordings. Only the automated portion of the algorithms was evaluated—algorithm parameters, when available, were not varied. In research practice, algorithm parameters may be tuned as a function of the data set and decomposition results are manually edited. As such, it is likely that the agreements/accuracies listed herein represent a lower bound on those that might be found in practice [23]. In particular, cursory examination of the results showed that one common source of errors occurred when one algorithm created two or more distinct (often non-overlapping) MU trains that really corresponded to the same MU. In our analyses, the partial MU train with the larger number of firings would tend to be correctly paired, but the unit with the smaller number of firings would produce an error at each firing. Manual editing tends to find and correct this issue, with a single “merging” of MUs correcting a large number of agreement/accuracy errors. Hence, improved automated merging of distinct MUs might substantially improve the automated performance of these algorithms.

For the multiple channel data, there was a clear performance decrement as contraction level increased (Fig. 1). This result is expected, as MU recruitment and firing rates rise with contraction level, producing more overlapping discharges and increasing the likelihood of MU shape similarity. Both of these issues are known challenges to decomposition algorithms [3], [4], [8]. Additionally, median simulation *accuracy* exceeded experimental *agreement*, particularly at higher MVC levels. Of course, *agreement* between two algorithms counts an error whenever either algorithm errs, whereas *accuracy* with respect to the known firings of simulated data only counts an error when the test algorithm produces a mistaken firing. Additionally, the decomposition difficulty may have differed between the experimental and simulation data. Even though their Complexity values were matched, this one measure may not adequately capture all factors that affect decomposition difficulty. Note that most experimental validation studies are based on data from relatively low contraction levels. Thus, the 95.7% median agreement for the 10% MVC quadrifilar experimental data, as well as the near-perfect accuracies for the quadrifilar simulation results are consistent with prior performance assessments of these and other mature decomposition algorithms [1]–[4]. Anecdotally, it was noted that the Montreal algorithm more frequently produced MU decompositions with very high accuracy (e.g., 25th percentile accuracy of 98%), but occasionally failed to detect MUs with high SNR_{MU} .

The experimental single channel data led to perfect and near-perfect median agreement results between all three pairs of algorithms. Agreement was not statistically different between any algorithm pair. These data were recorded at much lower contraction levels (described as clinically “moderate”), thus were likely a less challenging data set. For the single channel simulation, the Montreal and EMGLAB algorithms out-performed Fuzzy Expert.

Fig. 6 compared agreement vs. accuracy for the two simulations. Accuracy nearly always exceeded agreement, perhaps again reflecting that accuracy performance is only

influenced by errors in the test algorithm while agreement is influenced by errors in both algorithms. Most important, whenever agreement was high (e.g., above 90–95%), accuracy was similarly high—suggesting that each algorithm was correctly detecting the same MU discharges and properly classifying them. Fig. 6 also seems to indicate that agreement is largely determined by the algorithm with the weaker performance. The weaker algorithm has many values along the diagonal (line of agreement).

Figs. 2–5 show a rather clear relationship between SNR_{MU} and agreement/accuracy. By combining the data from the three MVC levels (as shown in the figures), it appeared that SNR_{MU} captured most of the Complexity information that varies with contraction level. These figures also show that the RMS errors from the best fit exponential model between agreement/accuracy and SNR_{MU} were in the range of 6–20%, depending on MVC level and electrode recording type. This RMS error seems moderately high, suggesting that SNR_{MU} is useful but not necessarily definitive. Although not shown here (see [29] for additional results), we similarly analyzed the relationship between agreement/accuracy and Complexity, and also the Decomposability Index of Florestal et al. [3]. Result trends were not distinguishable from those found using SNR_{MU} . These measures are population-based statistical models of performance vs. quality measure. McGill and Marateb [30] developed a quality measure based on the properties of each individual recording.

The evidence found herein did not support universal selection of one algorithm over the other. In fact, the high agreement/accuracy at lower contraction levels suggests that each of the three algorithms is largely detecting and similarly classifying the same MU discharges. Further, Fig. 6 strongly suggests that when two algorithms are in strong *agreement*, then they are each likely to also be highly *accurate*. Additionally, Figs. 2–5 suggest that high accuracy is most probably achieved whenever the SNR_{MU} is high.

Reliable assessment of EMG decomposition algorithms is an important, yet difficult, challenge with many inherent limitations. In simulation studies, known true decompositions exist, but the simulator cannot fully capture the character of real EMG. In experimental studies, the true decomposition is not known. Thus, it is common to study decomposition algorithm performance with both real and simulated data. Some algorithms focus on decomposition of only those MUs exhibiting a large SNR. Algorithms that decompose all detected spikes—even those corresponding to low SNR units—might be penalized with a lower *average* accuracy/agreement performance, even though they perform as well on those units with high SNR. Anecdotally, this type of detection error was observed from the Fuzzy Expert algorithm, whose default settings tended to detect more low-SNR MUs compared to the other algorithms. False positive noise spike detections seemed more prevalent in these low-SNR MUs. And, of course, distinct algorithms should be compared using the same data, since algorithm performance is influenced by the characteristics of the signal recordings. Accordingly, some algorithms may perform better on certain signals as compared to other

algorithms, and vice versa.

To our knowledge, this study is the first systematic comparison of performance across algorithms commonly used for MU decomposition. Each of the algorithms that we compared is publicly available. Our comparison information is important to the interpretation of results across studies using different analysis techniques. For the quadrifilar experimental data, median agreement between the Montreal and Fuzzy Expert algorithms at 10%, 20% and 50% MVC were 95.7, 86.4 and 64.8%, respectively. For the single channel data, median agreement between pairs of the three algorithms was 94.9% (Montreal vs. Fuzzy Expert) and 100% (EMGLAB vs. either Montreal or Fuzzy Expert). Agreement across algorithms and accuracy within algorithms were strongly related to SNR_{MU} . When agreement was high between algorithm pairs applied to the simulated data, so was the individual accuracy of each algorithm. These results, therefore, provide confidence that the algorithms perform reliably on experimental quadrifilar and single channel indwelling EMG signals.

REFERENCES

- [1] K. C. McGill, Z. C. Lateva, and H. R. Marateb, "EMGLAB: An interactive EMG decomposition program," *J. Neurosci. Meth.*, vol. 149, pp. 121–133, 2005.
- [2] Z. Erim and W. Lin, "Decomposition of intramuscular EMG signals using a heuristic fuzzy expert system," *IEEE Trans. Biomed. Eng.*, vol. 55, pp. 2180–2189, 2008.
- [3] J. R. Florestal, P. A. Mathieu, and K. C. McGill, "Automatic decomposition of multichannel intramuscular EMG signals," *J. Electromyogr. Kinesiol.*, vol. 19, pp. 1–9, 2009.
- [4] S. H. Nawab, R. P. Wotiz, and C. J. DeLuca, "Decomposition of indwelling EMG signals," *J. Appl. Physiol.*, vol. 105, pp. 700–710, 2008.
- [5] D. Stashuk, "EMG signal decomposition: How can it be accomplished and used?," *J. Electromyogr. Kinesiol.*, vol. 11, pp. 151–173, 2001.
- [6] R. C. Thornton and A. W. Michell, "Techniques and applications of EMG: Measuring motor units from structure to function," *J. Neurol.*, vol. 259, pp. 585–594, 2012.
- [7] R. S. LeFever and C. J. DeLuca, "A procedure for decomposing the myoelectric signal into its constituent action potentials—Part I: Technique, theory, and implementation," *IEEE Trans. Biomed. Eng.*, vol. 29, pp. 149–157, 1982.
- [8] R. S. LeFever, A. P. Xenakis, and C. J. DeLuca, "A procedure for decomposing the myoelectric signal into its constituent action potentials—Part II: Execution and test for accuracy," *IEEE Trans. Biomed. Eng.*, vol. 29, pp. 158–164, 1982.
- [9] E. Chauvet, O. Fokapu, J. Y. Hogrel, D. Gamet, and J. Duchene, "Automatic identification of motor unit action potential trains from electromyographic signals using fuzzy techniques," *Med. Biol. Eng. Comput.*, vol. 41, pp. 646–653, 2003.
- [10] C. I. Christodoulou and C. S. Pattichis, "Unsupervised pattern recognition for the classification of EMG signals," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 169–178, 1999.
- [11] J. Fang, G. C. Agarwal, and B. T. Shahani, "Decomposition of multiunit electromyographic signals," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 685–697, 1999.
- [12] D. Ge, E. Le Carpentier, Farina D, "Unsupervised Bayesian decomposition of multiunit EMG recordings using tabu search," *IEEE Trans. Biomed. Eng.*, vol. 57, pp. 561–571, 2010.
- [13] R. Gut and G. S. Moschytz, "High-precision EMG signal decomposition using communications techniques," *IEEE Trans. Sig. Proc.*, vol. 48, pp. 2487–2494, 2000.
- [14] H. Parsaei, D. W. Stashuk, S. Rasheed, C. Farkas, and A. Hamilton-Wright, "Intramuscular EMG signal decomposition," *Crit. Rev. Biomed. Eng.*, vol. 38, pp. 435–465, 2010.
- [15] D. Zennaro, P. Wellig, V. M. Koch, G. S. Moschytz, and T. Laubli, "A software package for the decomposition of long-term multichannel EMG signals using wavelet coefficients," *IEEE Trans. Biomed. Eng.*, vol. 50, pp. 58–69, 2003.
- [16] K. C. McGill, K. L. Cummins, and L. J. Dorfman, "Automatic decomposition of the clinical electromyogram," *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 470–477, 1985.
- [17] J. R. Florestal, P. A. Mathieu, and A. Malanda, "Automated decomposition of intramuscular electromyographic signals," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 832–839, 2006.
- [18] D. Farina, R. Colombo, R. Merletti, and H. B. Olsen, "Evaluation of intra-muscular EMG signal decomposition algorithms," *J. Electromyogr. Kinesiol.*, vol. 11, pp. 175–187, 2001.
- [19] C. J. DeLuca, "Reflections on EMG signal decomposition," in *Computer-Aided Electromyography and Expert Systems*, J. E. Desmedt, Ed. Elsevier, 1989, pp. 33–37.
- [20] B. Mambrito and C. J. DeLuca, "A technique for the detection, decomposition and analysis of the EMG signal," *EEG Clin. Neurophysiol.*, vol. 58, pp. 175–188, 1984.
- [21] K. C. McGill, Z. C. Lateva, and M. E. Johanson, "Validation of a computer-aided EMG decomposition method," in *IEEE Eng. Med. Biol. Conf.*, 2004, pp. 4744–4747.
- [22] C. J. DeLuca, A. Adam, R. Wotiz, L. D. Gilmore and S. H. Nawab, "Decomposition of surface EMG signals," *J. Neurophysiol.*, vol. 96, pp. 1646–1657, 2006.
- [23] A. Holobar, 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. Sys. Rehab. Eng.*, vol. 18, pp. 221–229, 2010.
- [24] H. R. Marateb, K. C. McGill, A. Holobar, Z. C. Lateva, M. Mansourian, and R. Merletti, "Accuracy assessment of CKC high-density surface EMG decomposition in biceps femoris muscle," *J. Neural. Eng.*, vol. 8, 066002, 2011.
- [25] K. C. McGill, "A comparison of three quantitative motor unit analysis algorithms," *Suppl. Clin. Neurophysiol.*, vol. 60, 273–278, 2009.
- [26] A. Hamilton-Wright and D. W. Stashuk, "Physiologically based simulation of clinical EMG signals," *IEEE Trans. Biomed. Eng.*, vol. 52, 171–183, 2005.
- [27] A. Christie and G. Kamen, "Doublet discharges in motoneurons of young and older adults," *J. Neurophysiol.*, vol. 95, 2787–2795, 2006.
- [28] M. Nikolic, "Detailed Analysis of Clinical Electromyography Signals EMG Decomposition, Findings and Firing Pattern Analysis in Controls and Patients with Myopathy and Amyotrophic Lateral Sclerosis," Ph.D. thesis, Faculty of Health Science, University of Copenhagen, 2001.
- [29] Y. Li, C. Dai, E. A. Clancy, P. Bonato, A. Christie, and K. C. McGill, "Cross-comparison between two multi-channel EMG decomposition algorithms assessed with experimental and simulated data," in *IEEE Northeast. Bioeng. Conf.*, 2013, in press.
- [30] K. C. McGill and H. R. Marateb, "Rigorous *a posteriori* assessment of accuracy in EMG decomposition," *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol. 19, pp. 54–63, 2011.

Author biographies provided if/when requested.

TABLE I
 NUMBER OF MOTOR UNITS (MUS) ANALYZED. NUMBER OF MATCHES
 SHOWN FOR EXPERIMENTAL DATA SETS. NUMBER OF TRUTH MUS SHOWN
 FOR SIMULATION DATA SETS.

Data Set	Number of MUs per Signal	Total Number of MUs		
		10% MVC	20% MVC	50% MVC
Quadrifilar Experiment	3-11	78	90	81
Quadrifilar Simulation	7-13	103	110	120
		Montreal vs. Fuzzy	Montreal vs. EMGLAB	EMGLAB vs. Fuzzy
Single Channel Experiment	3-10	51	63	52
Single Channel Simulation	6-9	51	73	68

Data Set	10% MVC	20% MVC	50% MVC										
Quadrifilar Experiment (Agreement)	95.7 [80.9, 100]	86.1 [71.6, 97.3]	81.8 [35.3, 83.8]										
Quadrifilar Simulation:													
Montreal (Accuracy)	99.8 [98, 100]	100 [96.2, 100]	95.9 [73.9, 100]										
Fuzzy Expert (Accuracy)	100 [87.9, 100]	98 [69.8, 100]	93.5 [62.1, 97.5]										
<table border="1"> <thead> <tr> <th>Montreal vs.</th> <th>Montreal vs.</th> <th>EMGLAB vs.</th> </tr> <tr> <th>Fuzzy</th> <th>EMGLAB</th> <th>Fuzzy</th> </tr> </thead> <tbody> <tr> <td>Single Channel Experiment (Agreement)</td> <td>94.9 [78.2, 100]</td> <td>100 [98, 100]</td> <td>100 [97.3, 100]</td> </tr> </tbody> </table>				Montreal vs.	Montreal vs.	EMGLAB vs.	Fuzzy	EMGLAB	Fuzzy	Single Channel Experiment (Agreement)	94.9 [78.2, 100]	100 [98, 100]	100 [97.3, 100]
Montreal vs.	Montreal vs.	EMGLAB vs.											
Fuzzy	EMGLAB	Fuzzy											
Single Channel Experiment (Agreement)	94.9 [78.2, 100]	100 [98, 100]	100 [97.3, 100]										
<table border="1"> <thead> <tr> <th>Montreal</th> <th>Fuzzy Expert</th> <th>EMGLAB</th> </tr> </thead> <tbody> <tr> <td>Single Channel Simulation (Accuracy)</td> <td>100 [96, 100]</td> <td>84.9 [78.2, 100]</td> <td>100 [97.3, 100]</td> </tr> </tbody> </table>				Montreal	Fuzzy Expert	EMGLAB	Single Channel Simulation (Accuracy)	100 [96, 100]	84.9 [78.2, 100]	100 [97.3, 100]			
Montreal	Fuzzy Expert	EMGLAB											
Single Channel Simulation (Accuracy)	100 [96, 100]	84.9 [78.2, 100]	100 [97.3, 100]										

Fig. 1. Summary statistics of decomposition agreement and accuracy. Each cell lists “x [y, z]”, where x is the median agreement (or accuracy), y is the 25th percentile agreement/accuracy and z is the 75th percentile agreement/accuracy. All values exclude outliers. Sample sizes are provided by Table I. Symbol “*” denotes marginally significant difference ($0.01 \leq p < 0.05$), symbol “**” denotes significant difference ($p < 0.01$).

Quadrifilar Experiment: Montreal - Fuzzy Expert

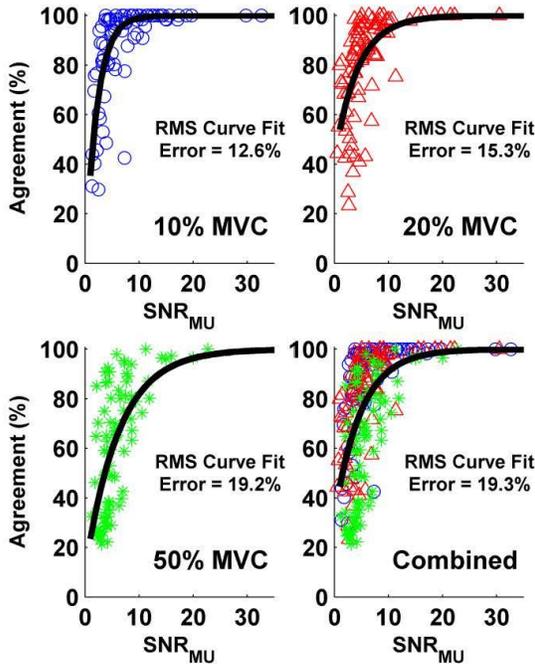


Fig. 2. Agreement between the Fuzzy Expert and Montreal algorithms as a function of SNR_{MU} for the quadrifilar experimental data. Results shown separately for each MVC level, and for all levels combined. Best fit exponential model shown in each plot, along with the RMS fit error.

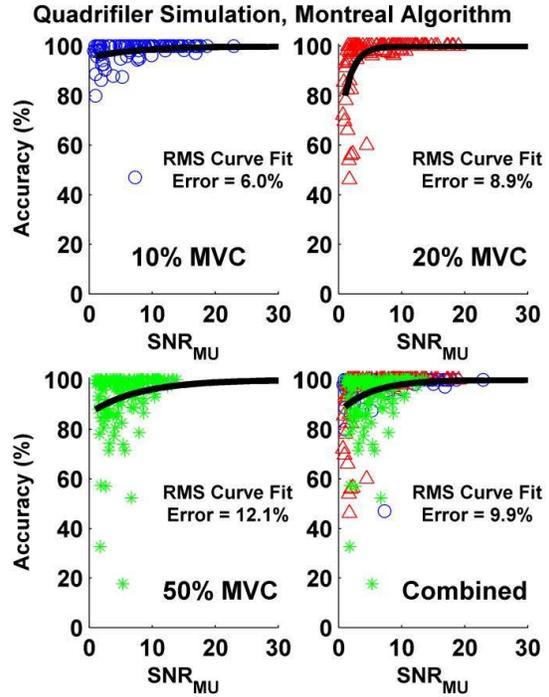


Fig. 3. Accuracy with respect to the true decomposition for the quadrifilar simulations as a function of SNR_{MU} for the Montreal algorithm. Results shown separately for each MVC level, and for all levels combined. Best fit exponential model shown in each plot, along with the RMS fit error.

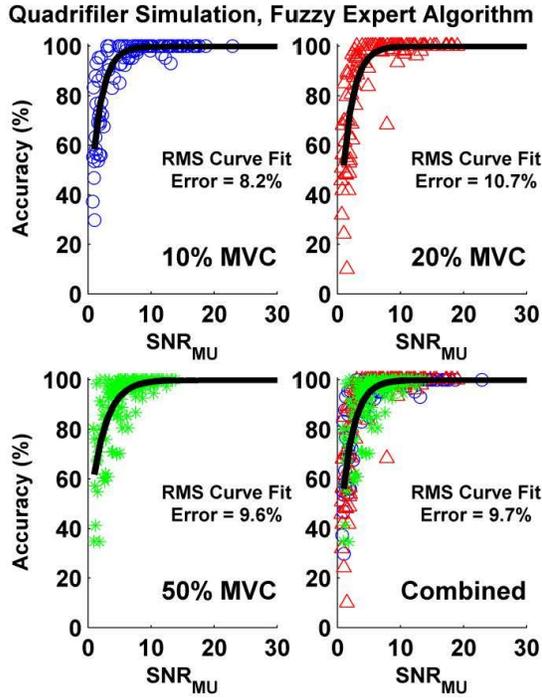


Fig. 4. Accuracy with respect to the true decomposition for the quadrifilar simulations as a function of SNR_{MU} for the Fuzzy Expert algorithm. Results shown separately for each MVC level, and for all levels combined. Best fit exponential model shown in each plot, along with the RMS fit error.

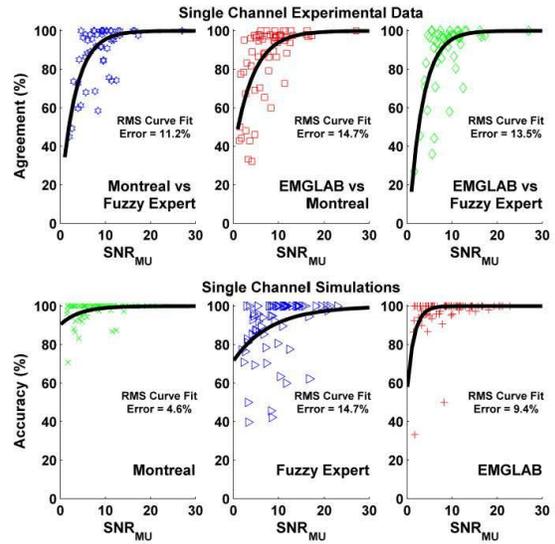


Fig. 5. Agreement (top) between algorithm pairs (as labeled) as a function of SNR_{MU} for the single channel experimental data. Accuracy (bottom) with respect to the true decomposition for the single channel simulated data as a function of SNR_{MU} . Best fit exponential model shown in each plot, along with the RMS fit error.

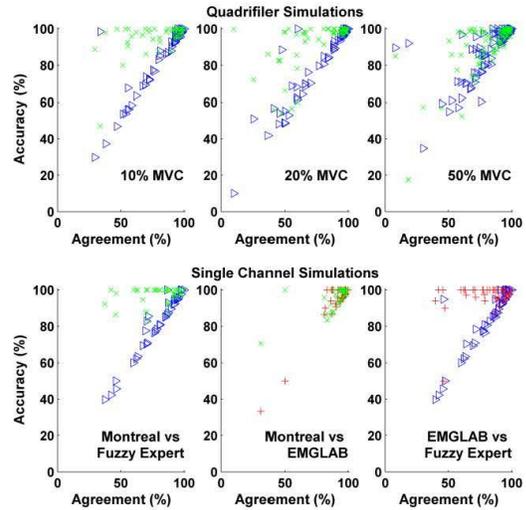


Fig. 6. (Top) Agreement between Montreal and Fuzzy Expert algorithms vs. their individual accuracies for the quadrifilar simulations, separated by MVC level. Each x -axis agreement value corresponds to two y -axis accuracy values. (Bottom) Agreement between algorithm pairs vs. their individual accuracies for the single channel simulations. For all plots, red plus=EMGLAB, green "X"=Montreal, blue triangle=Fuzzy Expert.

Appendix A: Informal Report Regarding EMG Decomposition

1. Introduction

The most important thing we need to consider about EMG decomposition is to evaluate the performance of decomposition algorithms. The purpose of this paper is to compare the accuracy of different EMG signal decomposition algorithms——MTL, FuzzyExpert and Emglab. Three algorithms were tested on single channel. Only MTL and FuzzyExpert were used on multi-channel decomposition, since Emglab only can decompose one channel signal. Two main approaches to evaluate the performance have been proposed in this paper:

1. For the real data which have no accurate results, we built agreement to reflect the performance between the two instead.
2. Simulated data can be used as a reference. The huge advantage of simulated data is that it has truth annotation and accuracy can be computed. Besides, agreement was also computed in order to set up a relationship between agreement and accuracy. This can help evaluate accuracy from the agreement of real data.

2. Method

2.1 EMG signal recording

The data used in this paper come from two parts: real data recorded in the hospital and simulated data generated by the simulator:

1. For real data, we used both Multi-channel data from UMass and single channel data from EMGLab.net.

For UMass data, a total number of 16 subjects covered a variety of age (classified as young and old), gender and MVC contraction level (including 10%, 20% and 50%) were three channel data and used as multi-channel decomposition. An excluding criteria was first applied for all data which based on the level of noise and duration of stable activity. Finally 12 subjects——7 young including 3 males and 4 females and 5 old including 2 males and 3 females were judged as the usable recording for further processing. The data were recorded simultaneously using three bipolar electrodes called quadrifilar (? based on Anita's Email) and the ADC resolution is 16-bit.

For EMGlab.net data, we used N2001 database of clinical signals which consisted of a various groups such as one normal control group, one group of patients with myopathy and one group of ALS patients. Here we only used the normal control group for our single channel decomposition. The group consisted of 10 normal subjects aged from 21 to 37 years old, 4 females and 6 males. Each subject had 15 recordings at low-level contraction and another 15 at moderate level. Since the low level recordings were just above the threshold and too easy to analyze, we only chose one moderate level data from each subject according to background noise and complexity. The data was sampled at 23437.5 Hz and 16-bit ADC. The electrode type was concentric needle and the muscle type was biceps brachii.

* Please cite this data as:
Nikolic M. Detailed Analysis of Clinical Electromyography Signals EMG Decomposition, Findings and Firing Pattern Analysis in Controls and Patients with Myopathy and Amytrophic Lateral Sclerosis. PhD Thesis, Faculty of Health Science, University of Copenhagen, 2001.
[The data are available as dataset N2001 at <http://www.emglab.net>]

2. A physiologically based simulation of clinical EMG signals had been developed. We want the simulated data we generate are similar to UMass data for multi-channel and EMGlab.net data for single channel. The sampling rate of both data was 31250 Hz and ADC resolution is 16-bit. There are different key parameters for simulator setting such as:
 1. Jitter
The shape of a same motor unit from different firing times often has some difference. Jitter is used to measure this kind of diversity.
 2. Muscle setting
Muscle setting includes number of motor units in muscle, muscle fiber density, muscle fiber area and motor unit diameter. Most of the parameters here were used default setting which was calculated by the average human main muscle activity.
 3. Electrode type and position
There are different electrode types can be selected: concentric, monopolar and bipolar.

For multi-channel UMass data, the electrode type is bipolar which basically consists of two adjacent monopolar electrodes and records the differential voltage between the two but here we used one pair of monopolar electrodes to mimic each bipolar electrode because in this way the distance between the electrodes can be better controlled. The length of the tip of each electrode is 10 mm. The diameter of each monopolar electrode is $50\mu\text{m}$ and the distance between two electrodes is $200\mu\text{m}$. The positions of monopolars are placed as figure (0). There are six possible combinations of the differential voltage between monopolar electrodes which can be regards as six bipolar electrodes and only three of them which are independent were picked.

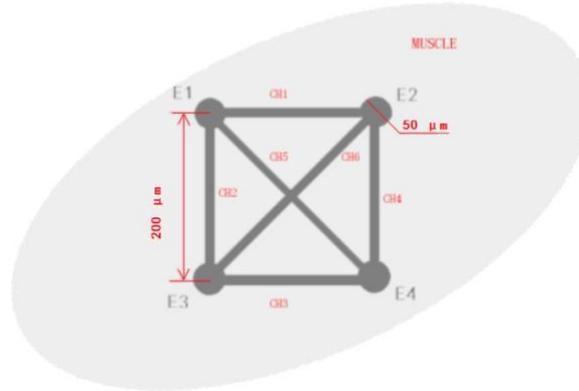


Figure (0)

For single channel EMGlab.net data, the electrode type was set as concentric and the length of the tip of the electrode was 10 mm.

4. MVC contraction level and pulse per second (PPS).

In order to use simulator to best mimic the contraction level of the real data, complexity is used to measure the contraction level. The complexity of data mainly measured by pulses per second (PPS). PPS of each subject was pre-computed manually which was a more reliable way to get exact number. The standard is that for each normal pulse with amplitude more than the max peak amplitude of background noise will be counted one pulse. For the case of superimposition, when the duration of one pattern which contains two consecutive pulses is

more than 3ms, it will be counted as two pulses. And if the duration of one pattern which contains three consecutive pulses is more than 6ms, it will be counted as three pulses and so on so forth. For the case of multichannel, we will take a comprehensive consideration of all three channels. If one spike is identified in each channel, it will be counted as one pulse else won't. For the stable period selected for decomposition was 5s, which was not long enough and varied from subject to subject, we only counted the spikes of the stable period to get PPS.

The PPS of each MVC level in UMass data were: 10%-100.1, 20%-119.3 and 50%-211.8. The standard deviations of each level MVC data were: 10%-49.8, 20%-46.4 and 50%-54.6.

The PPS of multichannel simulation data were: 10%-99.2, 20%-120.7, and 50%-215.9.

The standard deviations of multichannel simulation data were: 10%-21.3, 20%-25.5 and 50%-59.4.

The average PPS of EMGlab.net database was 61.8 with standard deviation of 19.8.

The average PPS of single channel simulator was 62.56 with standard deviation of 12.3.

5. Noise

In order to get a best simulated effect, a white Gaussian noise was added on differential signals. The value of noise was measured on Signal-to-Noise Ratio SNR based on signal power:

$$SNR = 20 * \log_{10} \frac{Power_{signal}}{Power_{noise}}$$

2.2 Decomposition

Before decomposition, each signal was high-pass filtered in order to improve the accuracy of results. For UMass multi-channel data, the process of high-pass filter had been done at machine level and since some low frequency background noise existed, a 1-order zero-phase Butterworth high-pass filter with 100Hz cut off frequency was used. For EMGlab.net database the single channel signal was processed by a first high-pass filter with 2 Hz cut off and low-pass with 10 kHz cut off. We then used a 500 Hz 1-order zero-phase Butterworth high-pass filter to keep both good SNR and distinguishability.

Each signal channel signal was decomposed separately by three algorithms and each multi-channel signal by MTL and FuzzyExpert. All three algorithms automatically detected spikes of motor units and found out their discharges in the signal. Then, some sophisticated methods were used to align spikes and resolve superimpositions. In order to compare different algorithms under the same circumstance, the results of three algorithms after decomposition were saved as a uniform format, including discharging time, motor unit ID number and channel for each spike in the signal saved as annotation file (.eaf file).

The parameters of FuzzyExpert should be set carefully or the computing time would be long and inefficient. Most of the parameters can be chosen as default settings. We modified some key ones as below: a) passes =10; b) Min Template to Fill = 0.2; c) Max MU Combo for super-position: 3 for 1st and 2nd pass, 5 for 3rd pass and 6 for the rest. The parameters setting of one pass was show as Figure (1).

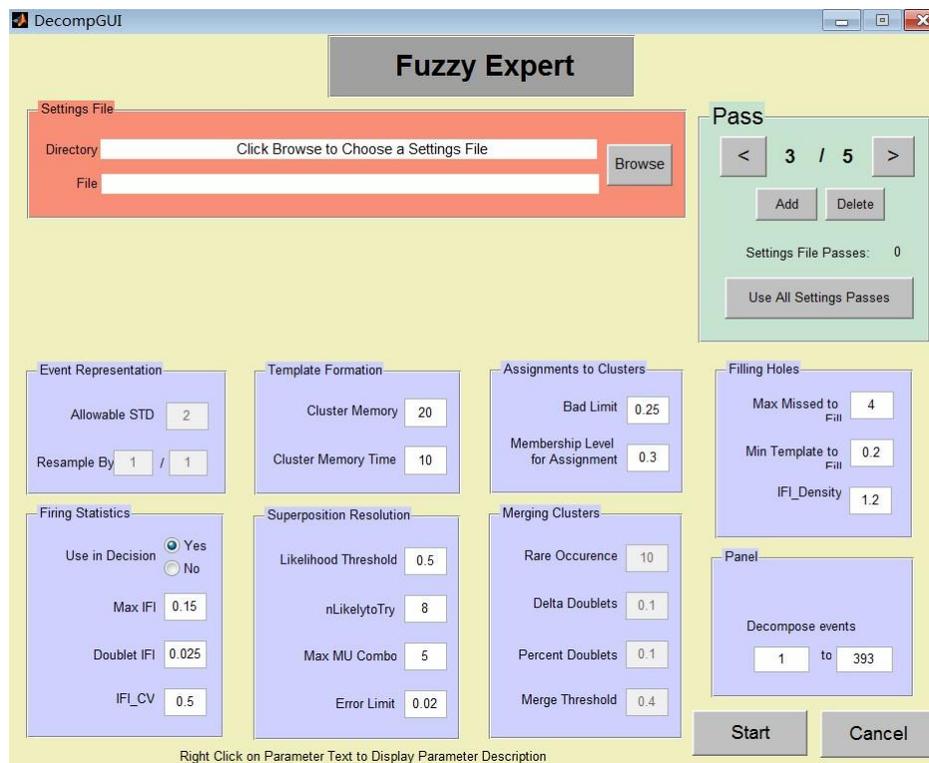


Figure (1): parameters of the FuzzyExpert

2.3 Decomposition comparison

The results of three different algorithms for each signal were not compared in one group assessment. Instead, a series of comparisons were made between pairs of algorithms in order to make the results more sensible and easier to follow. If only one algorithm detects a particular MU while another not, then it is basically a smaller unit and it can be ignored safely. Two comparison scenarios were built on both accuracy comparison for simulated data and agreement comparison for real data. First, between a known correct truth annotation from simulated data and a test annotation from the result of one of three algorithms, the truth annotation was taken as a standard in this case. Our goal is to portray how well the test annotation replicates the truth annotation. Second, between two annotations of which neither is considered as a standard. In this case, we wish to determine how well the two algorithms agree. The information in neither file should hold neither more nor less weight in determining the comparison outcome. Truth-test terminology will be used in the first scenario. For agreement comparison, the annotation of first algorithm will be taken as the truth and the annotation of second algorithm will be taken as the test. The main steps for truth-test comparison will include: (For agreement comparison, repeat the below computation and just pick first annotation as reference instead of truth annotation.)

1. Associate annotations with time offsets.

For truth-test comparison, loop over the test annotations. For each truth discharging time, find a closest test discharging time within time offset 1ms. After associating discharging times, some discharging times in truth annotation may have no test discharging time associated with them within 1ms offset. Record these discharging times without associated as not found (NF). Similarity, there are some discharging times without associated existing in test annotation. Record these discharging times as not included (NI).

2. Combine discharging times with motor unit ID number

After truth-test pairs were found, it is time to judge whether those pairs are correct with motor unit ID number. For each spike, we have both discharging time and motor unit ID number. If a truth-test pair has a same discharging time within 1ms offset but has different motor unit ID numbers, it recorded as false positives (FP). Only if both discharging time and motor unit ID number were matched, a pair recorded as true positive (TP).

3. Form Confusion Matrix

Now that truth and test motor unit IDs have been matched, a matrix was built to show the final results intuitively. The matrix was shown as figure (2):

TRUTH	TEST File Motor Unit Numbers										NF	
	1	2	3	4	5	6	7	8	9	10		
1	79*	0	0	0	0	0	0	0	0	0	0	0
2	0	0	87*	0	0	0	0	0	0	0	0	2
3	0	0	0	0	0	0	54*	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	71*	0	0
5	0	0	0	0	0	0	0	99*	0	0	0	0
6	0	0	0	0	0	0	0	0	71*	0	0	0
7	0	71*	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	63*	0	0	0	0	0	2
9	0	0	0	54*	4	1	0	0	0	0	0	26
10	0	0	0	0	45*	0	0	0	0	0	0	8
NI	0	0	1	21	10	1	0	0	0	0	0	0

Figure (2): 1. The first column is the truth motor unit ID number and the first row is the test motor unit ID number. 2. The last column is the quantity of spikes not found by test annotation (NF) for each truth motor unit and the last row is the quantity of spikes not included by truth motor unit (NI) for each test motor unit. 3. The rest of the matrix shows the matched conditions of each motor unit. The digits with asteroid are the quantity of true positives (TP) for all truth-test motor unit pairs. The digits without asteroid show the quantity of false positives (FP) for all truth-test motor unit pairs. It is quite easy to find out all NF, NI, TP and FP for each truth motor unit. For example, Number 9 motor unit of true annotation has 26 NFs 21NIs, 54 TPs which match with Number 4 motor unit of test annotation and 5 FPs which include 4 FPs classified as Number 5 motor unit and 1 FP classified as Number 6 motor unit from test annotation.

4. Evaluation of the final accuracy and agreement.

After confusion matrix was built, several important parameters can be computed in order to evaluate the performance of algorithms for each motor unit.

- a. The overall accuracy unit can be computed as:

$$Accuracy = \frac{TP}{TP + FP + NF + NI}$$

- b. The overall sensitivity can be computed as:

$$Sensitivity = \frac{TP + FP}{TP + FP + TF}$$

- c. The overall positive predictivity can be computed as:

$$Predictivity = \frac{TP}{TP + FP}$$

For agreement evaluation, the overall agreement for each motor unit is similar to overall accuracy, and other two parameters are no longer usable.

2.4 Analysis of comparison results

In order to get a better description of decomposition results, PPS, signal-to-noise ratio (SNR) and similarity unit were used in this paper.

2.4.1 PPS

PPS is defined in 2.1.2.4. Since PPS is in term of the whole signal, it is associated with total accuracy or agreement.

2.4.2 SNR

SNR is defined as the peak to peak amplitude (maximum subtract minimum for each spike) of each motor unit divided by the rms of the whole signal amplitude. Since SNR is evaluated separately for each motor unit, it is associated with accuracy or agreement of single motor unit.

Considering about the condition of superimposition, we cannot simply average the SNR of all spikes for each motor unit. Therefore, some strategies were needed to compute the SNR of each motor unit carefully, especially for real data which has no truth result. The main steps for calculating the SNR of a certain motor unit without truth includes :

1. Find out the matched IDs of annotations from two annotations for each motor unit.
2. Calculate peak to peak amplitude of all spikes. Since real data has no truth result, we need to calculate peak to peak amplitude according to two different annotations separately in terms of matched IDs. After getting all values of peak-to-peak amplitude, we will plot them in a histogram for each annotation shown as figure (3). We can get a statistic distribution of peak-to-peak amplitude.

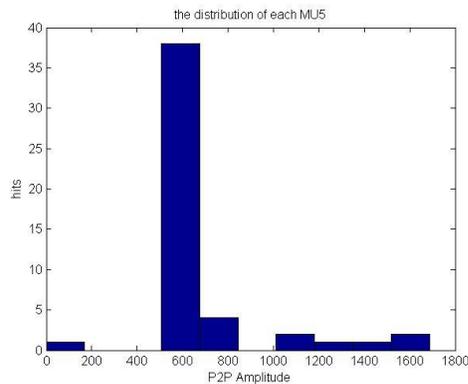


Figure (3): histogram of distribution of SNR for a certain motor unit

3. Since spikes without superimposition should have a dominant quantity and superimposition should always have different peak to peak amplitudes, the highest bar which means more amplitude distributed will be considered as the peak to peak amplitude of this certain motor unit. By averaging the peak-to-peak amplitude of the dominant bar for each annotation, then the mean of peak-to-peak amplitude of two annotations is computed. If it is multichannel, the average of the three is computed to get an overall SNR.
4. Calculate the RMS of the whole signal, then get the SNR of this motor unit by divided by RMS.

For simulated data with truth annotation, test annotation was no longer considered. Therefore, calculating the mean value of the peak-to-peak amplitude separately according to two annotations in terms of matched IDs was not needed. We only need to do the similar steps according to truth annotation.

2.4.3 Dissimilarity

If two kinds of motor units in the same signal are quite similar to each other, this will definitely influence the final result even if they have a relatively large SNR. Therefore, dissimilarity was

introduced to study its influence on agreement or accuracy. Dissimilarity is defined as:

$$Dissimilarity = \frac{\|m_{ki} - m_{ki}^*\|}{RMS}$$

, where denominator is the *RMS* of the whole signal, m_{ki} is k th motor unit in channel i , m_{ki}^* is the most similar motor unit to m_{ki} and the nominator is the norm of the difference of m_{ki} and m_{ki}^* . For dissimilarity measurement, it is still based on each motor unit.

2.4.4 CDI

A revised measurement called composite decomposability index (CDI) was also introduced by Kevin McGill and Florestal to quantify the difficulty of decomposition. CDI is defined as:

$$CDI = \frac{\min \{\|m_{ki}\|, \|m_{ki} - m_{ki}^*\|\}}{RMS}$$

, where denominator is the *RMS* of the whole signal, and the nominator contains norms of two parts that only smaller one will be selected. m_{ki} is k th motor unit in channel i , m_{ki}^* is the most similar motor unit to m_{ki} .

3. Result

3.1 Multi-channel UMass result

3.1.1 Results of SNR, dissimilarity and CDI versus agreement for each motor unit

Since the UMass data were recorded by hospital, the true annotations were unknown. So, we only built agreement to reflect the results. Figure (4) shows the results of agreement versus SNR for each motor unit. Each point indicates a pair of trains for each matched motor unit. In this paper, Matlab curve fitting toolbox was used to try to fit all points. The mathematical expression of the blue curve can be expressed as:

$$Agreement = -A * e^{-B*SNR} + C$$

A is used to adjust the range of Y axis. A larger A reflects the agreement may reach a quite low level with the same B.

B indicates the relationship between SNR and relationship. A larger B means the agreement can reach a high level with a smaller SNR.

C is an offset to make the range of Y axis from 0 to 100%. In general, C is always equal to 100.

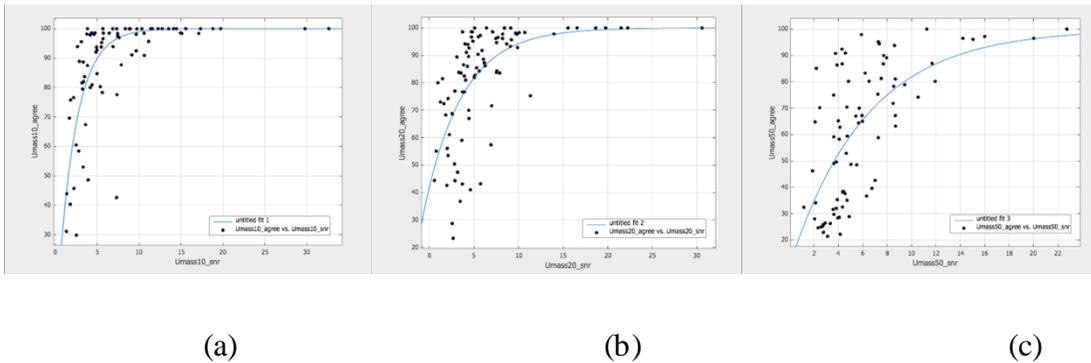


Figure (4): The figure shows the relationship between agreement and SNR as the MVC level rises. (a) The number of matched motor units identified for each subject ranged from 3 to 10. The total matched number was 78 pairs. The SNR for MVC 10% data was mostly from 1 to 15. The agreement mostly ranged from 40% to 90% for a small motor unit with SNR under 5. For motor unit with SNR from 5 to 10, the agreement can almost reach over 80%. When SNR is larger than 10, the agreement in general can be up to 90% or even 100%. The exponential expression for MVC 10% is $Agreement = -102.9 * e^{-0.465 * SNR} + 100$, which means an estimating agreement can be evaluated by a given SNR. The RMSE of the fitting model is 12.56. (b) The number of matched motor units identified for each subject ranged from 3 to 11. The total matched number was 90 pairs. The agreement becomes lower as MVC contraction level increases. In this case, the agreement is from 30% to 90% for some smaller motor units with SNR less than 5. The agreement can reach 80% to 100% with SNR between 5 and 15, and more points in the domain 90% to 100% when SNR larger than 10. The agreement will reach 95 or even 100% as SNR increases to 15 or more. The exponential expression for MVC 20% is $Agreement = -58.02 * e^{-0.2314 * SNR} + 100$. The RMSE is 15.26. (c) The number of matched motor

units identified for each subject ranged from 4 to 9. The total matched number was 81 pairs. The agreement drops sharply and the number of matched motor units decreases instead as MVC contraction level increases to a high level. This is because almost cases may be the case of superimposition although the number of motor units should increase theoretically. In this case, the agreement is from 20% to 90% for motor units with SNR less than 10. The agreement can reach 60% to 90% with SNR between 10 and 15. The agreement will reach 90% or more as SNR increases to 15 or more. The exponential expression for MVC 50% is $Agreement = -89.74 * e^{-0.1598 * SNR} + 100$ with a RMSE of 19.15. We can see clearly the slope of the function goes down as the MVC level goes up, i.e. the more complicated the data is the lower agreement two algorithms will achieve.

The detailed analyses of other figures are shown in table (see Appendix part).

The result of dissimilarity and CDI is similar to SNR. An exponential function can be also used to try to all points (shown in Appendix). The general expression of similarity and SNR fitting expression is the same as SNR. The figures of dissimilarity are shown as below (more specific analyses are shown in appendix part):

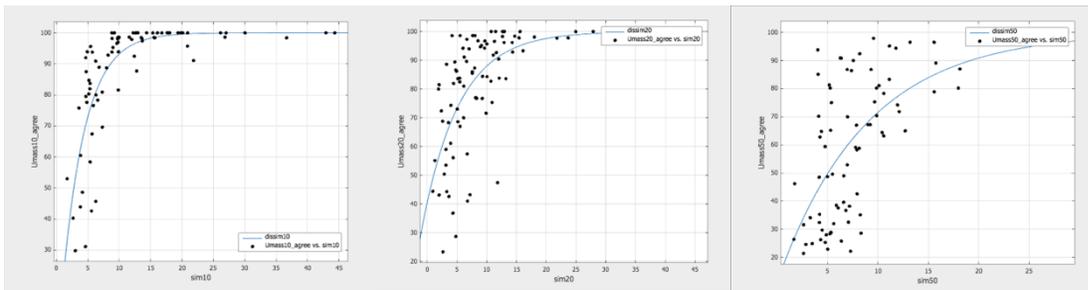


Figure (5): dissimilarity VS agreement with different MVC levels

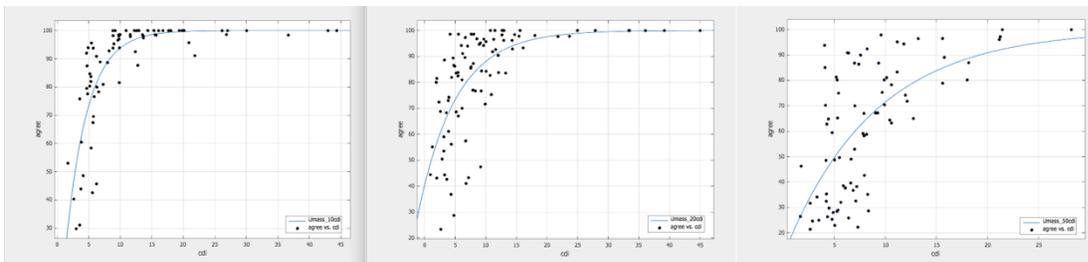


Figure (6): CDI VS agreement with different MVC levels

A one-way Anova was introduced to analyze the statistic result of agreement for each motor unit with different MVC level. The link of this method introduction can be found at:

<http://vault.hanover.edu/~altermattw/methods/stats/anova/one-way2b.html>.

The Anova result is shown as table 1.

Group name	Anova result
MVC10%, MVC20%, MVC50%	$F(2, 246) = 33.04777, p = 1.94 \cdot 10^{-13}$
MVC10%, MVC20%	$F(1, 166) = 4.13840, p = 0.043512$
MVC20%, MVC50%	$F(1, 169) = 33.76966, p = 3.02 \cdot 10^{-8}$
MVC10%, MVC50%	$F(1, 157) = 54.8418, p = 7.51 \cdot 10^{-12}$

Table 1: one-way Anova result of agreement for each motor unit with different MVC levels

3.1.2 Results of complexity versus agreement for each trial

SNR and similarity measurement is for single motor unit of each trial. Then, we developed complexity (mainly based on the PPS and the number of motor units identified of each trial) versus agreement to measure the overall agreement for each trial. Each point indicates the average of agreement of all motor units for each trial. Since we have 3 contraction levels and 12 subjects for each level, 36 points were shown in figure (7).

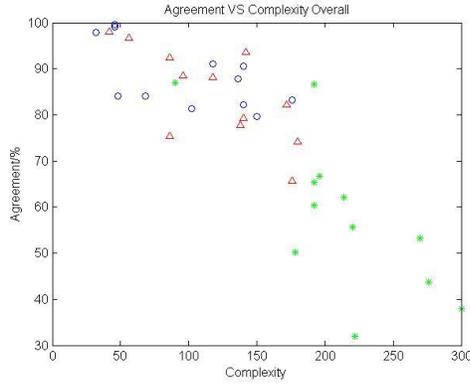


Figure (7): The overall agreement versus complexity measurement for each trial. The blue circles are the 10% MVC, the red triangles are the 20% MVC and the green asteroids are the 50% MVC. It also indicates that the higher complexity will arrive at a lower agreement.

The Anova result of complexity versus agreement for each trial with different MVC level is shown as table 2.

Group name	Anova result
MVC10%, MVC20%, MVC50%	$F(2, 33) = 21.52, p = 1.04 \times 10^{-6}$
MVC10%, MVC20%	$F(1, 22) = 1.2488, p = 0.27584$
MVC20%, MVC50%	$F(1, 22) = 20.67718, p = 1.59 \times 10^{-4}$
MVC10%, MVC50%	$F(1, 22) = 31.64667, p = 1.18 \times 10^{-5}$

Table 2: one-way Anova result of complexity versus agreement for each trial with different MVC levels

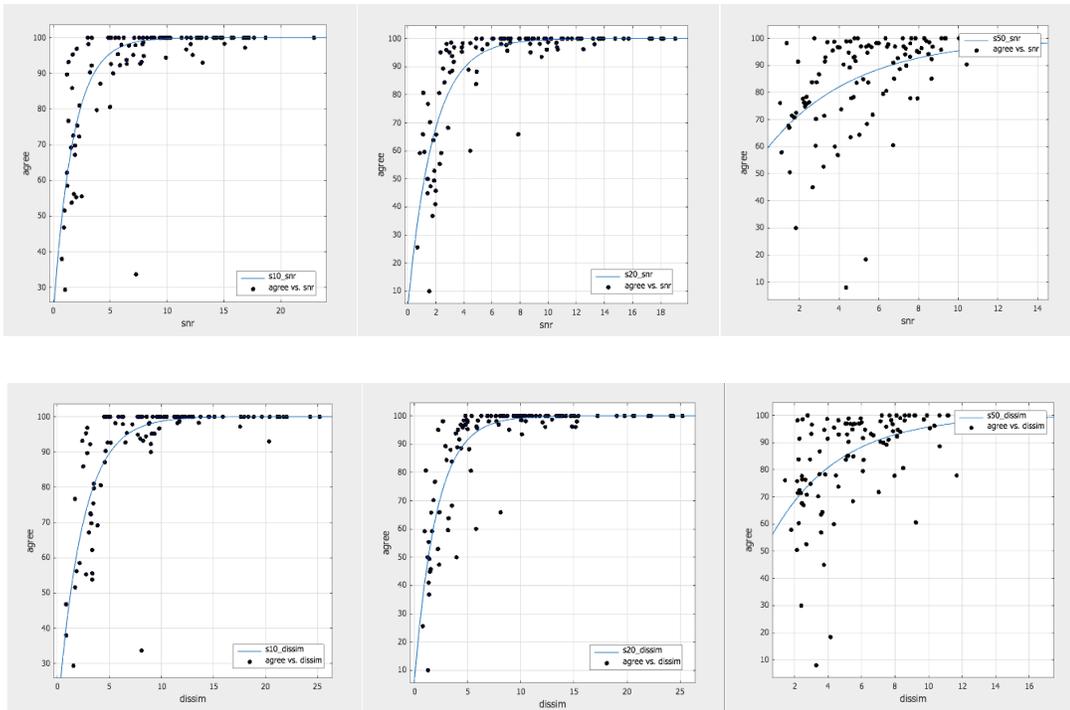
3.2 multi-channel simulated data result

For simulated data, we can not only compute the agreement between the two algorithms in a similar way as the real data but also we can calculate each algorithm’s accuracy based on the true annotation of

the simulator. In addition, we plot the relationship between accuracy and agreement for simulated data. It can be used as a reference to evaluate the accuracy by agreement when decomposing real data.

3.2.1 Results of SNR, dissimilarity and CDI versus agreement for each motor unit

First, the results of SNR, dissimilarity and CDI versus agreement are shown as figure (8). For MVC10%, the number of matched motor units identified for each subject ranged from 7 to 10 and the total matched number was 103 pairs. For MVC20%, the number of matched motor units identified for each subject ranged from 7 to 13 and the total matched number was 110 pairs. For MVC50%, the number of matched motor units identified for each subject ranged from 7 to 12 and the total matched number was 120 pairs.



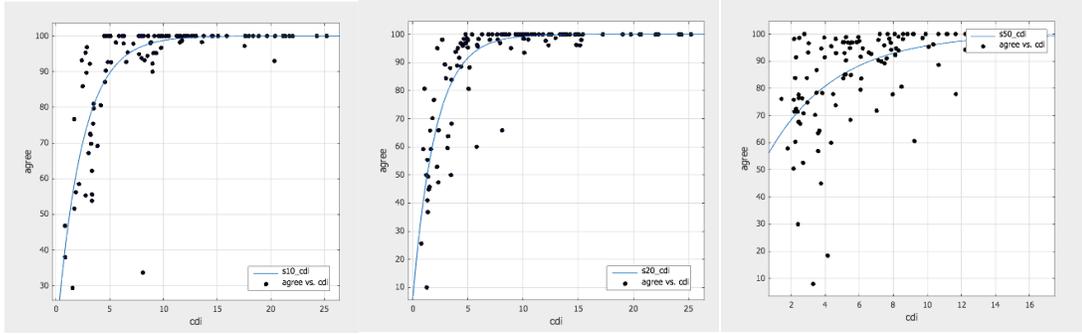


Figure (8): The relationship between SNR (dissimilarity, CDI) and agreement of the simulated data with different contraction level was computed the same way as the UMass data.

The Anova result of agreement for each trial with different MVC level is shown as table 3.

Group name	Anova result
MVC10%, MVC20%, MVC50%	$F(2, 400) = 13.97501, p = 1.36 \cdot 10^{-6}$
MVC10%, MVC20%	$F(1, 235) = 3.994058, p = 0.046813$
MVC20%, MVC50%	$F(1, 291) = 9.842108, p = 0.00188$
MVC10%, MVC50%	$F(1, 274) = 25.41835, p = 8.4 \cdot 10^{-7}$

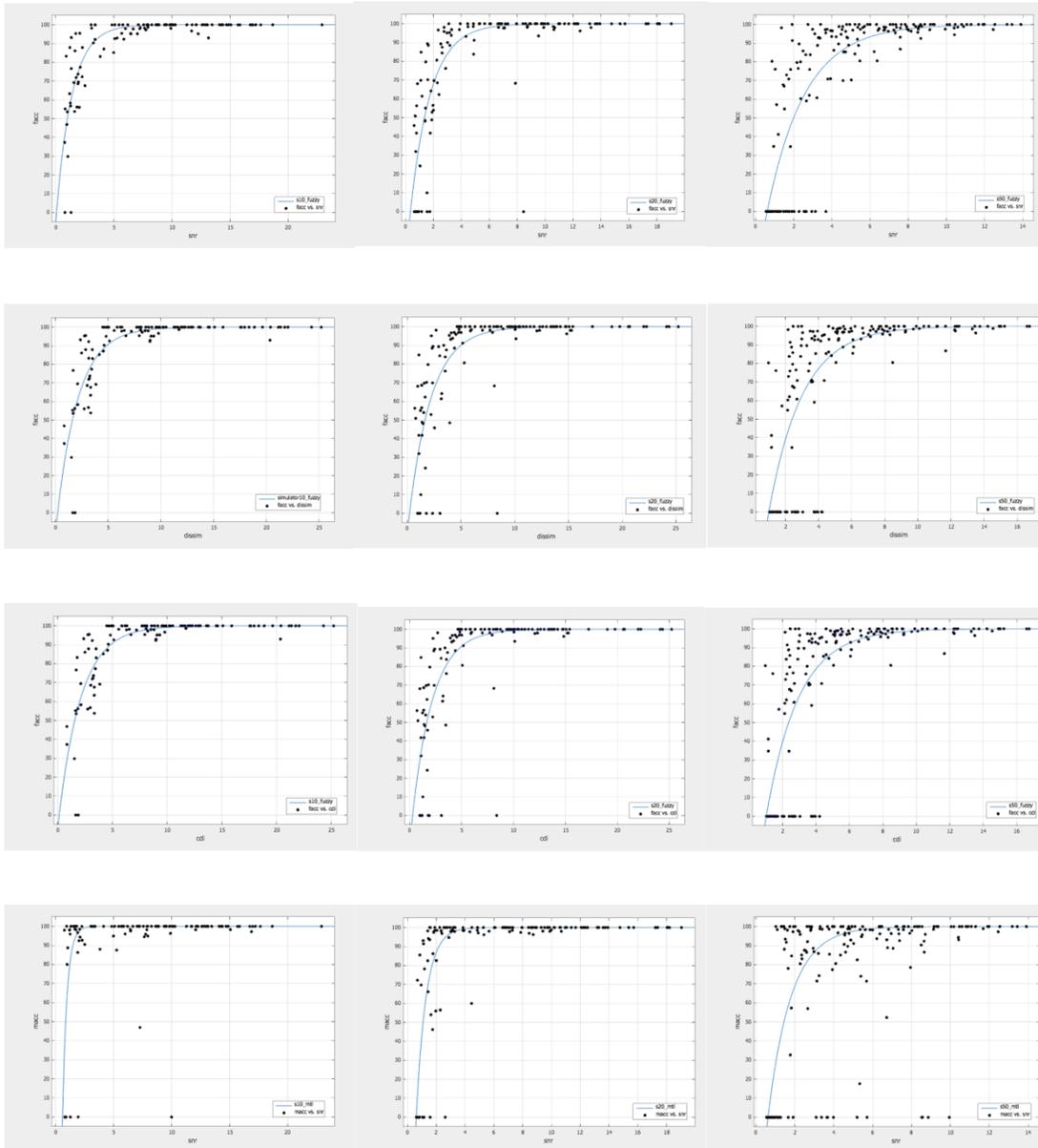
Table 3: one-way Anova result of agreement for each trial with different MVC levels

3.2.2 Results of SNR, dissimilarity and CDI versus accuracy of two algorithms for each motor unit

Second, since the simulated data has true annotation, accuracy of simulated data for each algorithm can be calculated.

For MVC10%, the number of motor units matching with truth identified by Fuzzy Expert or Mtl for each subject both ranged from 7 to 10, and the total matched number was 108 pairs for Fuzzy Expert and

104 pairs for Mtl. For MVC20%, the number of matched motor units identified by Fuzzy Expert for each subject ranged from 7 to 13 and from 7 to 15 by Mtl, and the total matched number was 117 pairs for Fuzzy Expert and 114 pairs for Mtl. For MVC20%, the number of matched motor units identified by Fuzzy Expert or Mtl for each subject ranged from 7 to 14 and from 4 to 15 by Mtl, and the total matched number was 133 pairs for Fuzzy Expert and 132 pairs for Mtl.



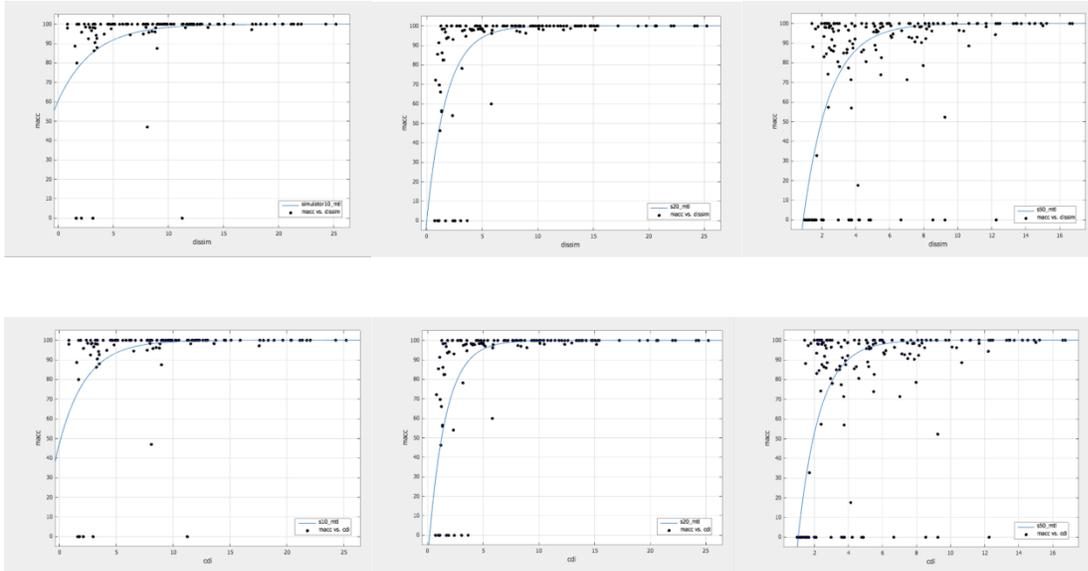


Figure (9): shows the accuracy versus SNR (dissimilarity, CDI) for two different algorithms. The decomposition result was compared with truth annotation and if the motor units is not found it will be judged as a miss and mark as zero in the plot. For SNR, in the low contraction level, MTL has a better performance—accuracy above 80% for 10%MVC and above 70% for 20%MVC, while FuzzyExpert is above 50% for 10%MVC and above 40% for 20%MVC. Both algorithms have similar accuracy range of 40% to 100% for the case of high contraction level of 50%MVC. However MTL is more likely to miss some templates with larger SNR. For dissimilarity and CDI, the result is similar to SNR and can be easily gotten from figures.

The Anova results of accuracy of two algorithms for each motor unit with different MVC level are shown as table 4 and table 5.

Group name	Anova result
MVC10%, MVC20%, MVC50%	$F(2, 400) = 9.26936, p = 1.16 \cdot 10^{-4}$
MVC10%, MVC20%	$F(1, 235) = 5.52475, p = 0.01958$
MVC20%, MVC50%	$F(1, 291) = 4.34973, p = 0.03787$
MVC10%, MVC50%	$F(1, 274) = 17.56522, p =$

	$3.75 \cdot 10^{-5}$
--	----------------------

Table 4: one-way Anova result of accuracy of Fuzzy Expert for each motor unit with different MVC levels

Group name	Anova result
MVC10%, MVC20%, MVC50%	$F(2, 400) = 11.51291, p = 1.38 \cdot 10^{-5}$
MVC10%, MVC20%	$F(1, 235) = 3.543954, p = 0.060998$
MVC20%, MVC50%	$F(1, 291) = 7.989511, p = 0.005031$
MVC10%, MVC50%	$F(1, 274) = 20.28175, p = 9.90 \cdot 10^{-6}$

Table 5: one-way Anova result of accuracy of Mtl for each motor unit with different MVC levels

3.2.3 Relationship between agreement and accuracy of two algorithms for each motor unit

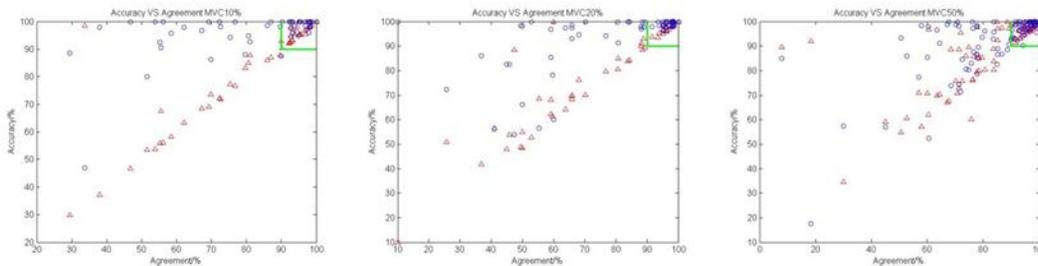
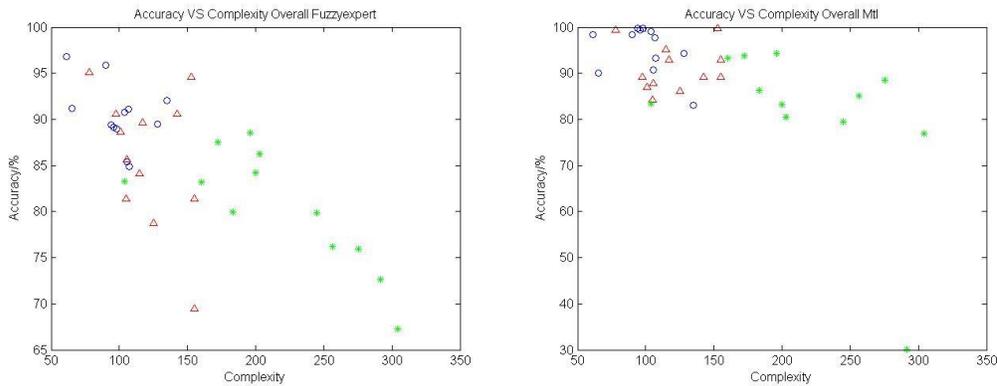


Figure (10): shows the cross relationship of two algorithms' agreement against accuracy. Triangles present for the accuracy of MTL, circles present for the accuracy of FuzzyExpert. Each agreement on x-axis will map to two different accuracy values on y-axis. The zone in the top right corner indicates that the more two algorithms agree with each other the higher accuracy they will achieve of the decomposition.

3.2.4 Results of complexity versus agreement for each trial



Figure(11): shows the accuracy versus complexity. MTL has a more concentrated range of high accuracy of above 80% over all MVC level while FuzzyExpert ranges from 70% to 95%. The blue circles are the 10%MVC, the red triangles are the 20%MVC and the green asteroids are the 50% MVC.

The Anova result of complexity versus agreement of two algorithms for each trial with different MVC levels is shown as table 6 and 7.

Group name	Anova result
MVC10%, MVC20%, MVC50%	$F(2, 33) = 8.485675, p = 0.001063$
MVC10%, MVC20%	$F(1, 22) = 3.852877, p = 0.062428$
MVC20%, MVC50%	$F(1, 22) = 3.737978, p = 0.066159$
MVC10%, MVC50%	$F(1, 22) = 22.63882, p = 9.48 \cdot 10^{-5}$

Table 6: one-way Anova result of complexity versus accuracy of Fuzzy Expert for each trial with different MVC levels

Group name	Anova result
MVC10%, MVC20%, MVC50%	$F(2, 33) = 5.444128, p = 0.009052$

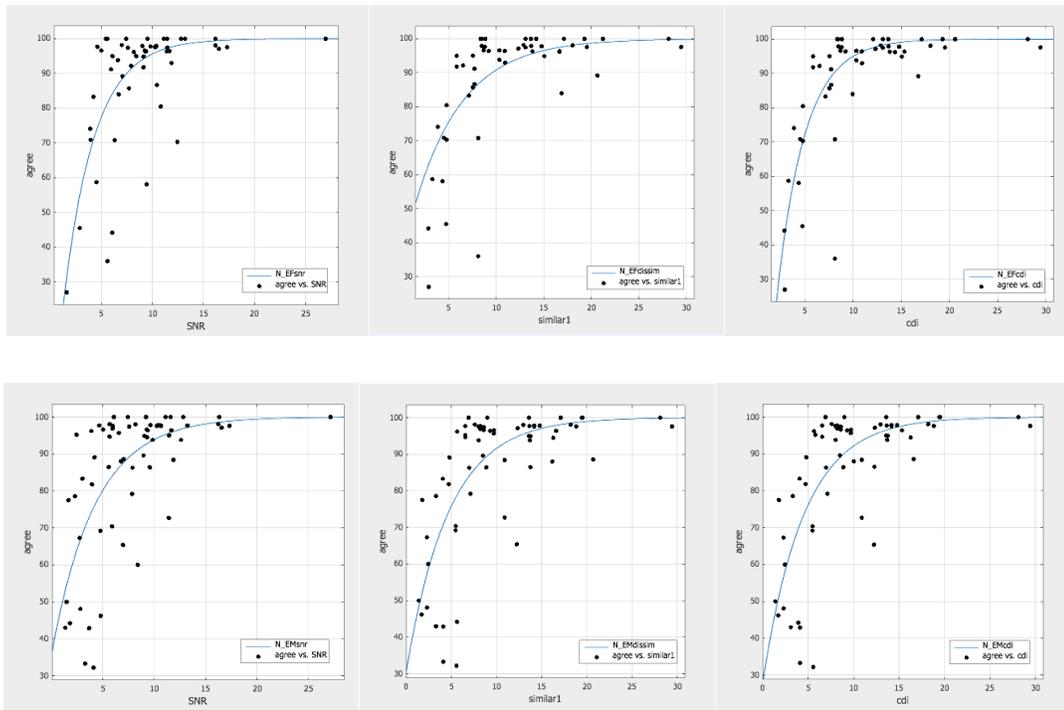
MVC10%, MVC20%	$F(1,22) = 4.166531, p = 0.053405$
MVC20%, MVC50%	$F(1, 22) = 3.63428, p = 0.069422$
MVC10%, MVC50%	$F(1, 22) = 7.443184, p = 1.23 \cdot 10^{-2}$

Table 7: one-way Anova result of complexity versus accuracy of Mtl for each trial with different MVC levels

3.3 single channel Emglab data result

For single channel data, comparisons were made between pairs of algorithms (Emglab VS Fuzzy Expert, Emglab VS Mtl and Mtl VS Fuzzy Expert). Only one contraction level (moderate level) was used for testing.

3.3.1 Results of SNR, dissimilarity and CDI versus agreement for each motor unit



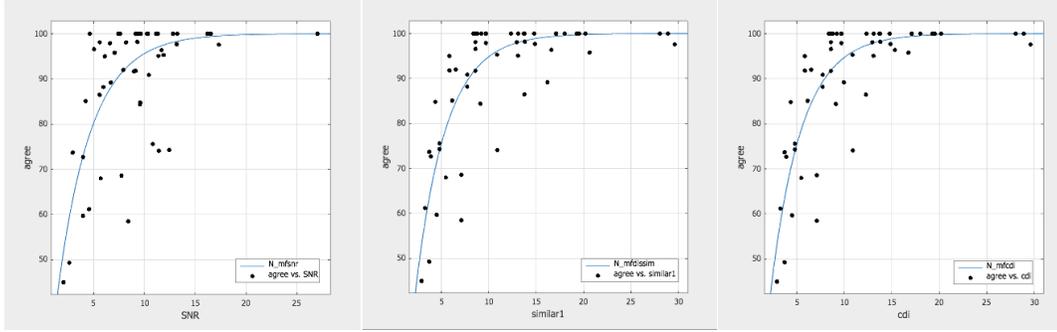


Figure (12): The figures show the results of different pairs of two algorithms. First row is the result of SNR, dissimilarity and CDI versus agreement about Emglab-FuzzyExpert (EF) pair. Second row is the result about Emglab-Mtl (EM) pair. Third row is the result about Mtl-Fuzzy (MF) pair. Similarly, the relationship between agreement and SNR (dissimilarity, CDI) can be analyzed for different pairs from figures. For EF pair, the number of matched motor units identified for each subject ranged from 4 to 8 and the total matched number was 52 pairs. For EM pair, the number of matched motor units identified for each subject ranged from 3 to 10 and the total matched number was 63 pairs. For MF pair, the number of matched motor units identified for each subject ranged from 4 to 9 and the total matched number was 51 pairs.

The Anova result of agreement for each algorithm pair is shown as table 8.

Group name	Anova result
EF pair, EM pair, MF pair	$F(2, 163) = 0.732179, p = 0.482435$

Table 8: one-way Anova result of agreement for each algorithm pair

3.3.2 Results of complexity versus agreement for each trial (Since the trials of single channel of Nikolic data are limited, complexity measurement can be omitted).

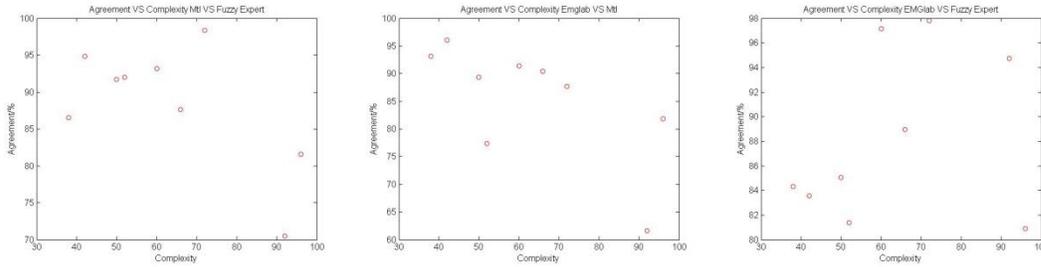


Figure (13) shows the result of total 10 trials of overall agreement versus complexity. For each trial, we have three results of pairs, so total 10 points were shown in each figure. A tendency that agreement decreases as complexity increases was shown in the figure. Since only one contraction level was tested and the range of complexity is relatively small comparing with multi-channel data, this kind of tendency is not obvious enough.

The Anova result of complexity versus agreement of three algorithm pairs for each trial is shown as table 9.

Group name	Anova result
EF pair, EM pair, MF pair	$F(2, 27) = 0.338457, p = 0.715851$

Table 9: The Anova result of complexity versus agreement of three algorithm pairs

3.4 single channel simulated data result

A group of simulated data was generated to evaluate the accuracy of three algorithms. Similar to what we have done in 3.2. Since we have three algorithms for single channel data, the results of three pairs of comparison were shown as below. And a relationship between agreement and accuracy was also shown to reflect the reliability of agreement for real data without true annotations.

3.4.1 Results of SNR, dissimilarity and CDI versus agreement for each motor unit with different algorithm pairs.

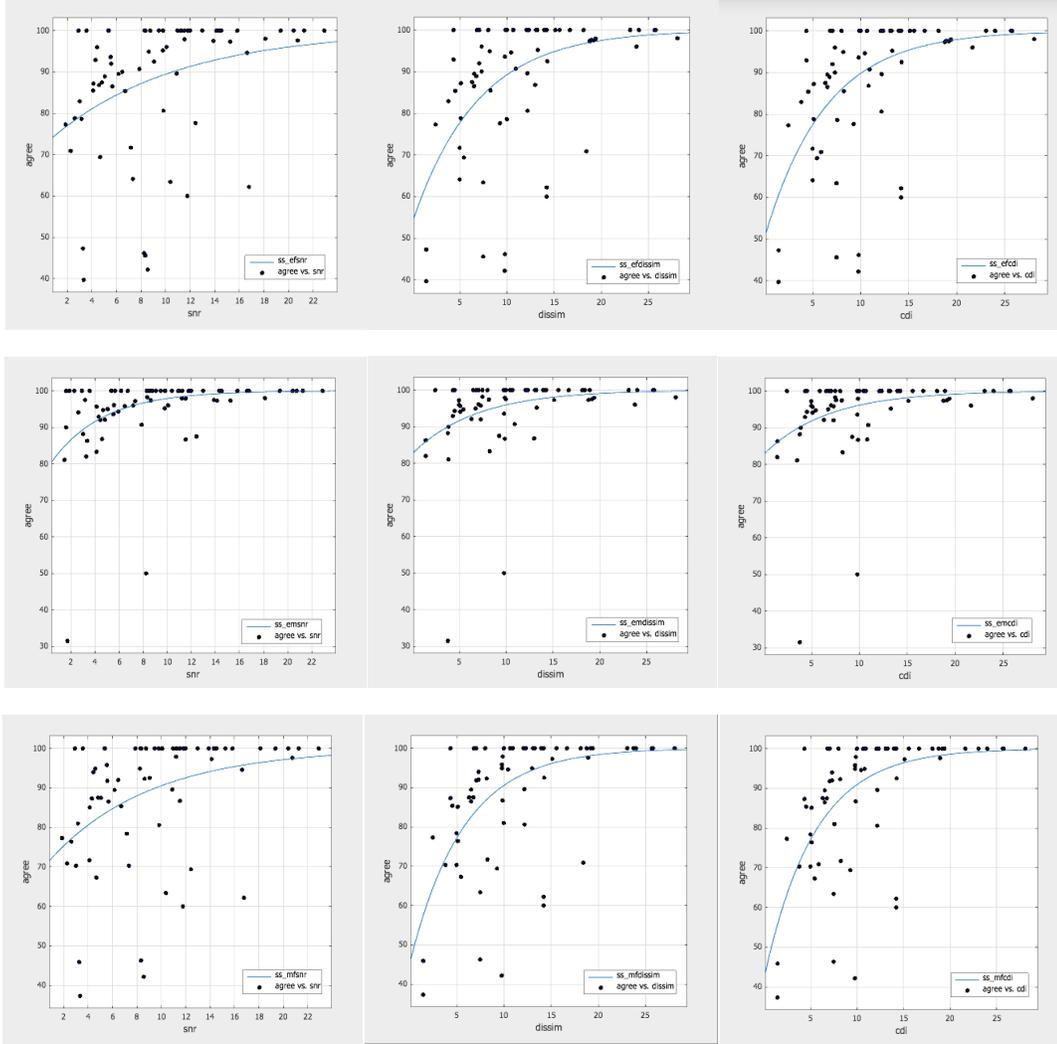


Figure (14): The relationship between SNR (dissimilarity, CDI) and agreement of the simulated data with different algorithm pairs. First row is the result of SNR, dissimilarity and CDI versus agreement about EF pair. Second row is the result about EM pair. Third row is the result about MF pair. Similarly, the relationship between agreement and SNR (dissimilarity, CDI) can be analyzed for different pairs from figures. For EF pair, the number of matched motor units identified for each subject ranged from 6 to 8 and the total matched number was 68 pairs. For EM pair, the number of matched motor units identified for each subject ranged from 6 to 9 and the total matched number was 73 pairs. For MF pair, the number of matched motor units identified for each subject ranged from 6 to 8 and the total matched number was 68 pairs.

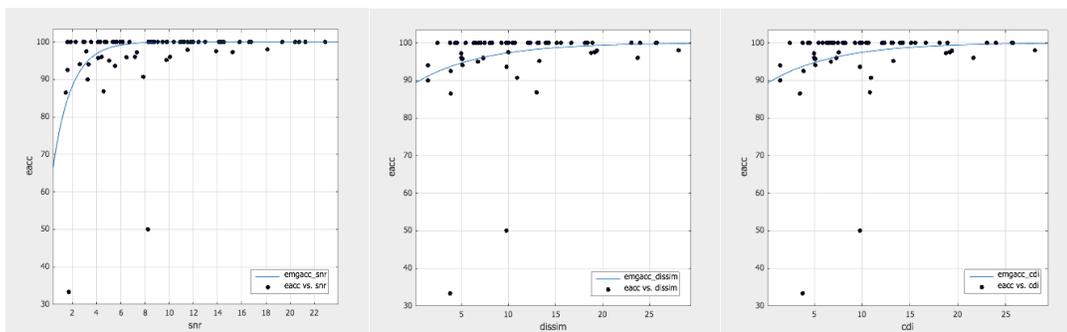
The Anova result of agreement for each algorithm pair is shown as table 10.

Group name	Anova result
EF pair, EM pair, MF pair	$F(2, 216) = 7.995563, p = 0.000447$
EM pair, MF pair	$F(1, 144) = 14.53662, p = 0.000203$
EF pair, EM pair	$F(1, 144) = 15.37151, p = 0.000136$
EF pair, MF pair	$F(1, 144) = 0.008002, p = 0.928844$

Table 10: one-way Anova result of agreement for each algorithm pair

3.4.2 Results of SNR, dissimilarity and CDI versus accuracy for each motor unit with different algorithms.

Second, since the simulated data has true annotation, accuracy of simulated data for each algorithm can be calculated.



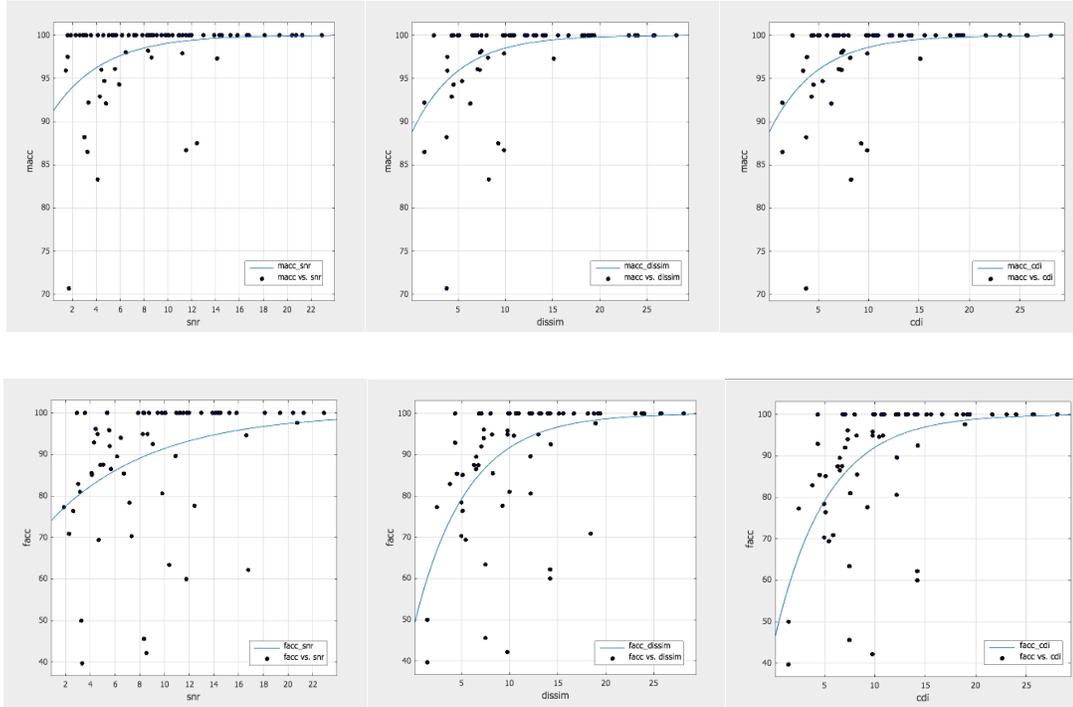


Figure (15): The relationship between accuracy and SNR (dissimilarity, CDI) can be analyzed for each algorithm. First row is the result of SNR, dissimilarity and CDI versus accuracy about Emglab algorithm. Second row is the result about Mtl algorithm. Third row is the result about Fuzzy Expert algorithm. Similarly, the relationship between accuracy and SNR (dissimilarity, CDI) can be analyzed for different algorithms from figures. For Emglab, the number of matched motor units identified for each subject ranged from 6 to 9 and the total matched number was 73 pairs. For Mtl, the number of matched motor units identified for each subject ranged from 6 to 9 and the total matched number was 73 pairs. For Fuzzy Expert, the number of matched motor units identified for each subject ranged from 6 to 8 and the total matched number was 68 pairs.

The Anova result of accuracy is shown as table 11.

Group name	Anova result
Emglab, Fuzzy Expert, Mtl	$F(2, 216) = 17.23247, p = 1.14 \times 10^{-7}$
Emglab, Fuzzy Expert	$F(1, 144) = 20.82178, p =$

	$1.07 \cdot 10^{-5}$
Emglab, Mtl	$F(1, 144) = 0.655483, p = 0.419495$
Fuzzy Expert, Mtl	$F(1, 144) = 16.38691, p = 8.4 \cdot 10^{-5}$

Table 11: one-way Anova result of accuracy for each algorithm

3.2.3 Relationship between agreement and accuracy of three algorithm pairs for each motor unit

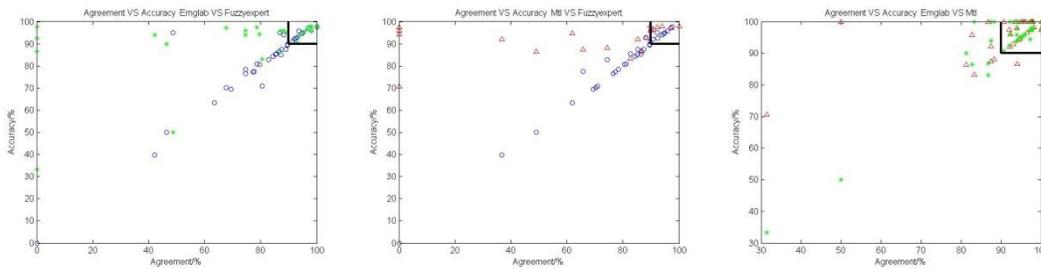


Figure (16) shows the cross relationship of two algorithms' agreement against accuracy. Red triangles present for the accuracy of MTL, circles present for the accuracy of FuzzyExpert and asterisks present for the accuracy of Emglab. When the agreement reaches 90%, the accuracy of two algorithms also can be 90% or more. Especially for MTL-Emglab pair (figure b), two algorithms often have a high agreement with high accuracy.

3.2.4 Results of complexity versus accuracy for each trial (Since the trials of single channel of simulated data are limited, complexity measurement can be omitted).

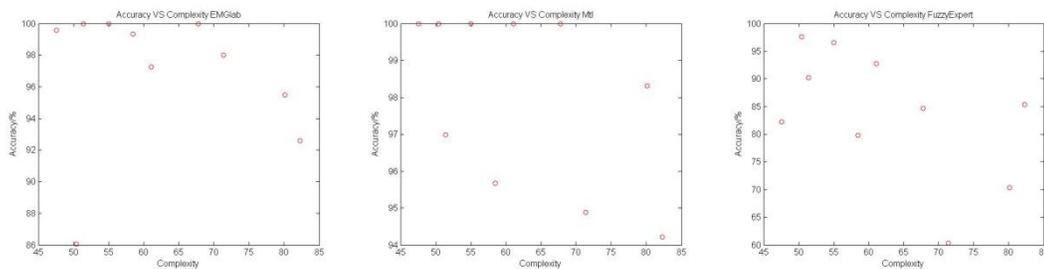


Figure (17) shows the result of total 10 trials of overall accuracy versus complexity for each algorithm.

The Anova result of complexity versus accuracy of three algorithms for each trial is shown as table 11.

Group name	Anova result
Emglab, Mtl, Fuzzy Expert	$F(2, 27) = 11.22016, p = 0.000284$
Emglab, Mtl	$F(1, 18) = 0.535726, p = 0.473632$
Mtl, Fuzzy Expert	$F(1, 18) = 13.87775, p = 0.001549$
Emglab, Fuzzy Expert	$F(1, 18) = 10.56945, p = 0.004435$

Table 11: The Anova result of complexity versus accuracy of three algorithms

4. Discussion

1、 High-pass filter

Various cut-off frequency settings of the high-pass filter may influence the result of the decomposition a lot. For the recording with less background noise, the high frequency (500Hz or 1000Hz) will provide a better performance. Since the multi-channel UMass data have already done the filtering on the hardware level, only a 100Hz high-pass filter is applied to eliminate the low frequency off-set.

2、 comment on different algorithms

Basically, the MTL has a better performance on multi-channel. Especially for the simulated data, the majority of the low level complexity (10% MVC and 20%MVC) decomposition results can achieve

100% accuracy. However, at the high contraction level it tends to miss some templates with high SNR. In the other hand, the FuzzyExpert has more sensitivity but sometime will detect too many small motor units. Different passes and parameters can be changed may have a big influence on the computing time, which careful setting should be considered.

3、 single channel data performance

EMGLab Auto Decomp and MTL agree with each other well which have an agreement over 80%.

4、 Confusion issue (merging)

One motor unit in one algorithm may map to several in the other algorithm due to the background noise and various key parameters like jitter and difference tolerance. One-to-two is a more general case and we currently don't have an efficient solution but only merge the separated motor units manually, i.e. find each motor units in the compare matrix, combine them into one and re-compute the agreement between the two new motor units.

5. Appendix

All tables are a detailed analysis of figures in result part.

Table 1 and 2 are for UMass data. 3 to 8 are for multi-channel simulated data. Table 9 is for EMGlab.net data. Table 10 and 11 are for single simulated data.

data	fitting	a	b	RMSE
UMass MVC10%	SNR vs. Agreement	-102.9	0.465	12.56
UMass MVC20%	SNR vs. Agreement	-58.02	0.2314	15.26
UMass MVC50%	SNR vs. Agreement	-89.74	0.1598	19.15

Table 1 (corresponding to figure 4): The result of SNR VS Agreement of Multi-channel UMass data

data	fitting	a	b	RMSE
UMass MVC10%	Dissimilarity vs. Agreement	-106.9	0.275	11.3
UMass MVC20%	Dissimilarity vs. Agreement	-59.52	0.1598	14.93
UMass MVC50%	Dissimilarity vs. Agreement	-88.85	0.1139	19.4

Table 2 (corresponding to figure 5): The result of Dissimilarity VS Agreement of Multi-channel UMass data

data	fitting	a	b	RMSE
UMass MVC10%	cdi vs. Agreement	-114.2	0.2916	10.77
UMass MVC20%	cdi vs. Agreement	-59.75	0.1609	14.79
UMass MVC50%	cdi vs. Agreement	-88.85	0.1139	19.4

Table 3 (corresponding to figure 6): The result of cdi VS Agreement of Multi-channel UMass data

data	fitting	a	b	RMSE
Simulator MVC10%	SNR vs. Accuracy (Fuzzy)	-104.5	0.7428	11.06
Simulator MVC10%	SNR vs. Accuracy (MTL)	-459.1	2.57	20.22
Simulator MVC10%	SNR vs. Agreement	-77.69	0.5741	10.66
Simulator MVC10%	Dissimilarity vs. Accuracy (Fuzzy)	-111.7	0.4971	10.89
Simulator MVC10%	Dissimilarity vs. Accuracy (MTL)	-39.57	0.3171	22.01
Simulator MVC10%	Dissimilarity vs. Agreement	-85.16	0.4218	10.33
Simulator MVC10%	cdi vs. Accuracy	-107.4	0.4914	11.19

	(Fuzzy)			
Simulator MVC10%	cdi vs. Accuracy (MTL)	-52.83	0.4219	21.65
Simulator MVC10%	cdi vs. Agreement	-85.06	0.4217	10.34
Simulator MVC20%	SNR vs. Accuracy (Fuzzy)	-127.7	0.6927	17.88
Simulator MVC20%	SNR vs. Accuracy (MTL)	-241.7	1.458	18.72
Simulator MVC20%	SNR vs. Agreement	-96.44	0.5534	11.09
Simulator MVC20%	Dissimilarity vs. Accuracy (Fuzzy)	-111	0.4816	19.59
Simulator MVC20%	Dissimilarity vs. Accuracy (MTL)	-99.65	0.5658	24.18
Simulator MVC20%	Dissimilarity vs. Agreement	-94.03	0.481	10.46
Simulator MVC20%	cdi vs. Accuracy (Fuzzy)	-115.1	0.5075	19.2
Simulator MVC20%	cdi vs. Accuracy (MTL)	-112.9	0.6489	23.65
Simulator MVC20%	cdi vs. Agreement	-94.73	0.4845	10.33

Simulator MVC50%	SNR vs. Accuracy (Fuzzy)	-127.7	0.6927	17.88
Simulator MVC50%	SNR vs. Accuracy (MTL)	-241.7	1.458	18.72
Simulator MVC50%	SNR vs. Agreement	-44.29	0.226	14.79
Simulator MVC50%	Dissimilarity vs. Accuracy (Fuzzy)	-163.5	0.4925	23.36
Simulator MVC50%	Dissimilarity vs. Accuracy (MTL)	-180.4	0.6459	33.38
Simulator MVC50%	Dissimilarity vs. Agreement	-51.83	0.2488	14.38
Simulator MVC50%	cdi vs. Accuracy (Fuzzy)	-170.8	0.5187	24.45
Simulator MVC50%	cdi vs. Accuracy (MTL)	-228.4	0.7903	32.5
Simulator MVC50%	cdi vs. Agreement	-51.83	0.2488	14.38

Table 4 (corresponding to figure 8): The result of Multi-channel Simulated data

data			fitting	a	b	RMSE
Fuzzy	Nikolic	MTL vs.	SNR vs. Agreement	-88.61	0.3008	11.2
	Nikolic	MTL vs.	Dissimilarity vs. Agreement	-116.8	0.3122	8.636
	Nikolic	MTL vs.	cdi vs. Agreement	-114.8	0.3077	8.521
EMGlab	Nikolic	MTL vs.	SNR vs. Agreement	-64.57	0.236	14.69
	Nikolic	MTL vs.	Dissimilarity vs. Agreement	-69.56	0.2114	13.74
	Nikolic	MTL vs.	cdi vs. Agreement	-72.04	0.2216	13.33
Fuzzy	Nikolic	EMGlab vs.	SNR vs. Agreement	-116	0.3282	13.52
	Nikolic	EMGlab vs.	Dissimilarity vs. Agreement	-157.8	0.3485	10.94
	Nikolic	EMGlab vs.	cdi vs. Agreement	-153.8	0.3417	10.82

Table 5(corresponding to figure 11): The result of EMGlab.net single-channel data

data	fitting	a	b	RMSE
Simulator Single	SNR vs. Accuracy (Fuzzy)	-28.5	0.1197	14.68
Simulator Single	Dissimilarity vs. Accuracy (Fuzzy)	-51.88	0.1834	13.33
Simulator Single	cdi vs. Accuracy (Fuzzy)	-54.83	0.1929	12.92
Simulator Single	SNR vs. Accuracy (MTL)	-9.611	0.2334	4.638
Simulator Single	Dissimilarity vs. Accuracy (MTL)	-11.51	0.2051	4.519
Simulator Single	cdi vs. Accuracy (MTL)	-11.5	0.2118	4.545
Simulator Single	SNR vs. Accuracy (EMGlab)	-44.79	0.6674	9.424
Simulator Single	Dissimilarity vs. Accuracy (EMGlab)	-10.87	0.1414	9.762
Simulator Single	cdi vs. Accuracy (EMGlab)	-10.92	0.1459	9.765

Simulator MTL vs. Fuzzy	Single	SNR vs. Agreement	-31.59	0.121	15.02
Simulator MTL vs. Fuzzy	Single	Dissimilarity vs. Agreement	-54.85	0.1758	13.46
Simulator MTL vs. Fuzzy	Single	cdi vs. Agreement	-57.77	0.1853	13.05
Simulator MTL vs. EMGlab	Single	SNR vs. Agreement	-21.34	0.2324	9.79
Simulator MTL vs. EMGlab	Single	Dissimilarity vs. Agreement	-17.37	0.1457	9.988
Simulator MTL vs. EMGlab	Single	cdi vs. Agreement	-17.32	0.1495	10.01
Simulator EMGlab vs. Fuzzy	Single	SNR vs. Agreement	-27.99	0.09753	15.56
Simulator EMGlab vs. Fuzzy	Single	Dissimilarity vs. Agreement	-46.14	0.1457	14.48
Simulator EMGlab vs. Fuzzy	Single	cdi vs. Agreement	-49.48	0.1577	14.11

Table 6 (corresponding to figure 13): The result of Emblab.net single-channel data

data	fitting	a	b	RMSE
Simulator Single	SNR vs. Accuracy (Fuzzy)	-28.5	0.1197	14.68
Simulator Single	Dissimilarity vs. Accuracy (Fuzzy)	-51.88	0.1834	13.33
Simulator Single	cdi vs. Accuracy (Fuzzy)	-54.83	0.1929	12.92
Simulator Single	SNR vs. Accuracy (MTL)	-9.611	0.2334	4.638
Simulator Single	Dissimilarity vs. Accuracy (MTL)	-11.51	0.2051	4.519
Simulator Single	cdi vs. Accuracy (MTL)	-11.5	0.2118	4.545
Simulator Single	SNR vs. Accuracy (EMGlab)	-44.79	0.6674	9.424
Simulator Single	Dissimilarity vs. Accuracy (EMGlab)	-10.87	0.1414	9.762
Simulator Single	cdi vs. Accuracy (EMGlab)	-10.92	0.1459	9.765

Table 7 (corresponding to figure 14): The result of single-channel simulated data