9th International IEEE EMBS Conference on Neural Engineering
San Francisco, CA, USA, March 20 - 23 , 2019

Differences in EMG Feature Space between Able-Bodied and Amputee Subjects for Myoelectric Control

Evan Campbell[1], Angkoon Phinyomark[1], Ali H. Al-Timemy[2], Rami N. Khushaba[3], Giovanni Petri[4], and Erik Scheme[1]

Abstract— **Difficulties accessing amputee populations has resulted in the widespread adoption of able-bodied subjects in virtual environments for the development of myoelectric prostheses. Factors such as scar tissue, different physiologies or surgical outcomes, and reduced visual and proprioceptive feedback, however, may contribute to differences in electromyogram (EMG) patterns between these groups. As such, studies have consistently found worse results when comparing the performance of amputee subjects to that of their able-bodied counterparts under the same conditions. To identify the source of this performance degradation, a topology-based data analysis method, called Mapper, was employed to visualize the "shape" of EMG feature spaces derived from amputee and able-bodied subjects. The information content of amputee EMG features was found to differ from those of non-amputee subject in three ways: 1) the loss of nonlinear complexity and frequency information, 2) the loss of time-series modeling information, and 3) the segmentation of unique information. The empirical effects of these differences were visualized by classifying motion classes using consistent and migratory features from functional feature groups. In summary, this work characterized inconsistencies in EMG features between amputee and able-bodied populations by theoretical means, highlighted the empirical effects when these are ignored, and proposed a solution for future studies with able-bodied subjects.**

I. INTRODUCTION

Electromyography (EMG) has been used by researchers and clinicians to extract information related to volitional movement for use in the control of prostheses [1], electric power wheelchairs [2], and diagnostic tools [3]. In particular, pattern recognition based myoelectric control has been an area of research with clinical influence as early as the 1960's. The performance of these systems are highly dependent on the availability of high quality and robust features that ensure class-separability with minimal complexity and redundancy with other selected features. The state-of-the-art pattern recognition architecture consists of pre-processing, data segmentation, feature extraction, further dimensionality reduction, and classification stages for the recognition of multiple gestures. Each component of this pipeline has been the subject of much research, resulting in numerous modifications and introductions of new techniques that have each incrementally improved in-laboratory and clinical performance.

The prohibitive cost and lack of availability of research prostheses, coupled with a lack of access to amputee populations, have resulted in many studies resorting to the testing of able-bodied subjects in virtual environments to validate controller strategies. Studies that have evaluated both populations, however, have commonly reported lower system performance when controlled by amputees than their able-bodied counterparts [4]. For example, Scheme and Englehart [1] showed consistent relative trends between the groups when evaluating different pattern recognition configurations, but consistently lower amputee performance. Five transradial amputee subjects experienced higher classification errors than ten able-bodied subjects despite training with fewer motion classes (i.e., 7 and 11 classes for amputee and able-bodied subjects, respectively). This performance margin was consistent across the eleven state-of-the-art classifiers employed in the study, with differences between the subject groups exceeding 10% in some cases.

Nevertheless, myoelectric control systems intended for amputee populations have regularly been designed and optimized using able-bodied subjects, naively translating the results to amputee subjects assuming the solution applies. While this has commonly been understood and, largely, accepted in the field, an explicit investigation of the credibility of this approach from a design perspective has not been conducted. In this work, we explore the differences between EMG features derived from amputee and able-bodied subjects using a topological data analysis tool called Mapper that simplifies feature space for better visualization and understanding. Using this tool, the relationship between feature types and their categorization into functional groups can help identify the causes of performance differences between user groups.

II. METHODOLOGY

A. EMG Data

This analysis used able-bodied and amputee data collected as part of previous experiments according to the Declaration of Helsinki. The able-bodied subject group was comprised

*This work was supported in part by the New Brunswick Health Research Foundation, the New Brunswick Innovation Foundation, and the Natural Sciences and Engineering Research Council of Canada grant number DG 2014-04920.

[1]Evan Campbell, Angkoon Phinyomark, and Erik Scheme are with the Institute of Biomedical Engineering, University of New Brunswick, Fredericton, NB E3B 5A3, Canada evan.campbell1@unb.ca, aphinyom@unb.ca, escheme@unb.ca

[2]Ali H. Al-Timemy is with the Biomedical Engineering Department, Al-Khwarzmi College of Engineering, University of Baghdad, Baghdad 10071, Iraq ali.altimemy@kecbu.uobaghdad.edu.iq

[3]Rami N. Khushaba is with the Faculty of Engineering and Information Technology, University of Technology, Sydney, New South Wales 2007, Australia rami.khushaba@uts.edu.au

[4]Giovanni Petri is with the ISI Foundation, Turin 10126, Italy and the ISI Global Science Foundation, New York City 10036, USA giovanni.petri@isi.it

978-1-5386-7921-0/19/$31.00 ©2019 IEEE

of twenty subjects (10 male, age 21.5±0.97; 10 female, age 21.2±0.79). Subjects' right arms were sanitized by an alcohol solution, then four pairs of Ag/AgCl electrodes were used to measure the EMG activity of four forearm muscles during eight hand motions. EMG signals were captured using a sampling rate of 1024 Hz. For more detail about the able-bodied data, readers are encouraged to consult [5]. The amputee subject group consisted of nine transradial amputees, seven of which were traumatic amputees (male, age 33.7±10.8), and two of which were congenital amputees (female, age 19, 31). Eight pairs of Ag/AgCl electrodes were placed around the residual limb of the subjects after the site had been prepared by alcohol and abrasive skin preparation gel. EMG signals were captured during six finger and hand motions using a sampling rate of 2000 Hz. For more detail about the amputee data collection process, readers are encouraged to consult [6].

Although different in origin, both datasets were pre-processed identically to minimize the potential effects of external factors such as electrical interference or motion artifacts. Pre-processing consisted of resampling both datasets to 1000 Hz, bandpass filtering at 20-500 Hz, and removing power-line interference by notch filtering at 50 Hz. Data were then segmented into 250 ms frames with 50% overlap (125 ms) for feature extraction.

B. Feature Extraction

Feature extraction is a technique that emphasizes the discriminating information of an input signal. In this work, 58 state-of-the-art EMG feature extraction methods in time domain and frequency domain, as described in Table 1 of [7], were employed to define 81 features that captured most known characteristics of the surface EMG signal [6], [8], [9]. It is important to note that some feature extraction methods provided more than one feature values. Afterwards, feature scaling was applied to ensure a zero-mean unit-variance distribution of features across all subjects and muscles.

C. Topological Data Analysis

Mapper, a topological simplification technique rooted in topological data analysis, can be used to extract key insights from complex, nonlinear, low signal-to-noise, highly variability data [7]. By producing a controlled simplification of high-dimensional data, a topological network robust to perturbations can be created. Here, Mapper was computed via a four stage pipeline:

(1) The raw EMG data were transformed into a point cloud. For each feature, feature values from all windows were combined to represent one entity through a principal component analysis (PCA) method where the number of principal components defined 95% of the variance of the feature vector. For instance, the able-bodied dataset contained 38,400 elements for each feature and was transformed into 28 PCA-dimensional space.

(2) The PCA-dimensional EMG feature point cloud was transformed by a filter function that preserves a global characteristic. In Mapper, regions that span the transformed space define a resolution specified by the number of regions and the overlap between regions; within this study, these were chosen to be 3, and 50%, respectively. The Euclidean distance of the kth-nearest-neighbour (k-NN) filter function, $k = 2$, was used as an indicator for feature similarity.

(3) Ward's hierarchical clustering method was applied to identify local relationships between features within each region by deriving mutually exclusive nodes that minimized the error sum of squares.

(4) Connections were formed between regions indicating the level of mutual information between nodes using edge thickness. The resulting nodes and connections formed the topological simplification of the original data cloud.

Analyses of the topological simplification illuminates the presence of functional feature groups, where features that characterize similar information are clustered within the same node or have high edge thickness with connected nodes. Specifically, for the k-NN filter function, features within groups that occupy low k-NN distance have strong correlation; whereas high k-NN distance indicates weak correlation. Repetition of this protocol on amputee and able-bodied datasets provides an environment to profile feature grouping differences between the populations and identify areas where the information content of amputee data is diminished.

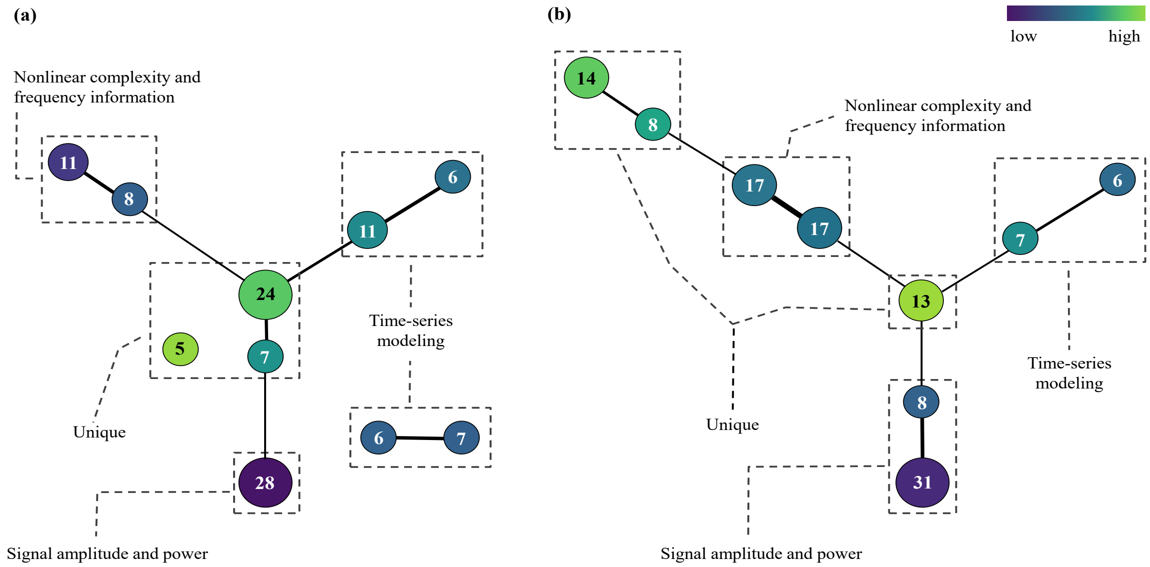
D. Classification

To provide empirical support for conclusions determined during topological data analysis, features from the various functional groups were used to build classification models. Additionally, pairs of key features were used to illustrate a scenario where features have low correlation to one another in able-bodied subjects (providing unique information); but are redundant in amputee subjects. For all classification tasks, support vector machine (SVM) classifiers with a linear kernel were trained using a 10-fold cross-validation to determine feature accuracy.

III. RESULTS

The topological networks constructed using Mapper that represent the information composition of able-bodied and amputee EMG feature space are shown in Fig 1. The able-bodied and amputee topological networks, both have a main structure shaped like the letter Y composed of three arms connected to a central core, and were comprised of 10 and 9 nodes, respectively. Analysis of the composition of able-bodied nodes revealed clusters of four functional feature groups based on information content: (1) signal amplitude and power, (2) nonlinear complexity and frequency information, (3) time-series modeling, and (4) unique features [7]. Examples of features within these functional groups for the able-bodied network were the root mean square (RMS) and mean absolute value (MAV) in the signal amplitude and power feature group; zero-crossings (ZC), slope sign change (SSC) and approximate entropy (ApEn) in the nonlinear complexity and frequency information feature group; autoregressive coefficients (AR) and correlation coefficients (CC)

Fig. 1. Topological feature networks. Node size is proportional to the number of features contained within the node, which is also explicitly given by the labeled number. Node color is representative of the average filter values, with blue indicative of low distance and green of high distance. (*a*) able-bodied subjects and (*b*) amputee subjects.



in the time-series modelling feature group; and histogram (HIST) and time domain power spectral density (TDPSD) in the unique feature group.

The amputee nodes differed from the able-bodied nodes in three notable ways; (1) loss of nonlinear complexity and frequency information, (2) loss of time-series modeling information, and (3) spatial segmentation of unique information. First, the loss of nonlinear complexity and frequency information was identified by the migration of maximal fractal length (MFL), myopulse percentage rate (MYOP), and Willison amplitude (WAMP) features from the nonlinear complexity and frequency information feature group to signal amplitude and power feature group. These features rely on the extraction of frequency characteristics through the time domain, thus indicating the inability to capture this information within amputee EMG signatures. Secondly, the loss of time series modelling information was identified by the migration of seven features including the first- and third-order AR and the second- and fourth-order CC, to the nonlinear complexity and frequency information feature group. Thirdly, the unique functional feature node location in the topological network experienced partial translation along the network, moving a portion of its features distal to the nonlinear complexity and frequency information region. In addition to this migration, the clustering of unique functional group features changed profoundly. The central unique amputee node was composed of features like HIST, kurtosis, and signal-to-motion-artifact ratio. The distal unique amputee nodes were composed of features like TDPSD, and skewness.

The SVM classification accuracies derived from both datasets were included for a subset of select features common to the four functional feature groups in Table I. Within the table, features are distinguished as migrated (M) when they changed functional feature groups across populations and constant (C) when they remain within the same functional feature group.

TABLE I

SVM CLASSIFICATION ACCURACY OF KEY EMG FEATURES ON ABLE-BODIED AND AMPUTEE DATASETS

Feature	Able-Bodied (%)	Amputee (%)	Migrate (M/C)
RMS	85.00	84.86	C
MAV	85.61	85.30	C
ZC	78.16	65.29	C
SSC	71.76	63.70	C
MFL	89.32	85.92	M
MYOP	80.86	75.10	M
WAMP	85.20	75.17	M
AR	79.85	84.84	M
CC	76.97	84.58	M
HIST	75.09	67.27	M
TDPSD	90.48	89.41	C
AR+ZC	86.11	86.17	M+C

IV. DISCUSSION

Inspection of the topological networks shows that the majority of features remain within the same functional groups between both populations. Without exception, signal amplitude and power features remained within their functional group, signifying that amplitude information remains consistently important between able-bodied and amputee subjects.

An implication the migration of other features, however, is that the optimal feature sets determined using able-bodied data through feature selection techniques (like sequential forward selection) may be sub-optimal for amputee motion recognition if the selected features provide fundamentally different information than originally intended. Specifically, when feature sets are translated across populations, performance degradation may be at least partially explained by information redundancy caused by feature migration.

From Table I, everything being equal, amputee performance *should* have exceeded able-bodied performance due to the inclusion of more measurement channels and fewer motion classes. However, nearly all features show consistent or worse performance on the amputee dataset. The effect of information redundancy can be seen in AR+ZC, where able-bodied features capture time series modelling, and nonlinear complexity and frequency information, whereas amputee features capture only nonlinear complexity and frequency information, resulting in similar performance despite a more complex classification task (more motion classes, fewer EMG channels) in the able-bodied case.

While feature migration provides some insight into areas of performance degradation, these migrations have yet to be tied to specific physiological factors. The sources of performance degradation for amputees originate from numerous factors, such as the loss of visual and proprioceptive feedback beyond the amputation site. Additionally, signals from traumatic amputees may be hindered by scar tissue, and post-amputation plasticity of the cortical somatotopic map [10] introducing inter-subject nerve signal propagation variability. For congenital amputees neither scar tissue nor change in cortical somatotopic map are generally a concern; however, some inter-subject variability in muscle geometry is still present. In spite of the general knowledge of these factors, able-bodied subjects have remained the standard candidate for amputee research for practical reasons.

To develop an appropriate environment for amputee research, we propose two potential solutions. The first is to overcome the scarcity of amputee research subjects through widespread collaboration of multiple research groups. Previously recommended by Resnik et al. [11], a multi-site collection harnessing the geographic dispersion of the amputee population would be a costly endeavor. However, with substantial coordination in standardizing aspects like sampling frequency and electrode sites across the numerous case studies in future work, a dataset suitable for big data techniques could be continuously expanded. The second solution is to develop an adaptive system that could be applied to able-bodied subject EMG signals to mimic the deviations caused by morphological changes present in amputee populations. Signal characteristics have been observed in studies like Waris et al. [12], where the effect of long-term EMG pattern recognition performance was compared between able-bodied and amputee subjects. Specifically, able-bodied subjects experienced similar performance across 7 days of study; however, amputee subjects experienced performance degradation as time between training and testing increased. Additionally, amputee performance was found to be proportional to residual limb size, indicating an anthropomorphic model could be beneficial. These findings motivate accompanying research into the variance of outcomes between amputee and able-bodied populations under dynamic factor perturbations like limb position, forearm orientation, contraction intensity, electrode shift, muscle fatigue, and noise [13].

In conclusion, this study explored the sources of differences between control strategies when used by able-bodied and amputee populations. Specifically, the migration of EMG features to different functional groupings highlights meaningful changes in information content between subject groups, and should be considered during experiments that use able-bodied subjects exclusively. Taking these differences into consideration may help inform the design of feature sets or experimental conditions that better allow for translation from able-bodied subjects to amputee end users.

References

[1] E. Scheme and K. Englehart, "Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use," *Journal of Rehabilitation Research & Development*, vol. 48, no. 6, pp. 643–660, July 2011.

[2] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "A review of control methods for electric power wheelchairs based on electromyography signals with special emphasis on pattern recognition," *IETE Technical Review*, vol. 28, no. 4, pp. 316–326, 2011.

[3] X. Zhang, P. E. Barkhaus, W. Z. Rymer, and P. Zhou, "Machine learning for supporting diagnosis of amyotrophic lateral sclerosis using surface electromyogram," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 96–103, Jan 2014.

[4] S. M. Wurth and L. J. Hargrove, "A real-time comparison between direct control, sequential pattern recognition control and simultaneous pattern recognition control using a fitts' law style assessment procedure," *Journal of NeuroEngineering and Rehabilitation*, vol. 11, no. 1, p. 91, May 2014.

[5] A. Phinyomark, P. Phukpattaranont, C. Limsakul, and M. Phothisonothai, "Electromyography (EMG) signal classification based on detrended fluctuation analysis," *Fluctuation and Noise Letters*, vol. 10, no. 03, pp. 281–301, 2011.

[6] A. H. Al-Timemy, R. N. Khushaba, G. Bugmann, and J. Escudero, "Improving the performance against force variation of EMG controlled multifunctional upper-limb prostheses for transradial amputees," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 6, pp. 650–661, June 2016.

[7] A. Phinyomark, R. N. Khushaba, E. Ibáñez-Marcelo, A. Patania, E. Scheme, and G. Petri, "Navigating features: a topologically informed chart of electromyographic features space," *Journal of The Royal Society Interface*, vol. 14, no. 137, 2017.

[8] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, and Y. Laurillau, "EMG feature evaluation for improving myoelectric pattern recognition robustness," *Expert Systems with Applications*, vol. 40, no. 12, pp. 4832 – 4840, 2013.

[9] A. Phinyomark, F. Quaine, S. Charbonnier, S. Serviere, S. Tarpin-Bernard, and Y. Laurillau, "Feature extraction of the first difference of EMG time series for EMG pattern recognition," *Computer Methods and Programs in Biomedicine*, vol. 117, no. 2, pp. 247 – 256, 2014.

[10] P. Montoya, K. Ritter, E. Huse, W. Larbig, C. Braun, S. Tpfner, W. Lutzenberger, W. Grodd, H. Flor, and N. Birbaumer, "The cortical somatotopic map and phantom phenomena in subjects with congenital limb atrophy and traumatic amputees with phantom limb pain," *European Journal of Neuroscience*, vol. 10, no. 3, pp. 1095–1102, 1998.

[11] L. Resnik, H. H. Huang, A. Winslow, D. L. Crouch, F. Zhang, and N. Wolk, "Evaluation of EMG pattern recognition for upper limb prosthesis control: a case study in comparison with direct myoelectric control," *Journal of NeuroEngineering and Rehabilitation*, vol. 15, no. 1, p. 23, Mar 2018.

[12] A. Waris, I. K. Niazi, M. Jamil, O. Gilani, K. Englehart, W. Jensen, M. Shafique, and E. N. Kamavuako, "The effect of time on EMG classification of hand motions in able-bodied and transradial amputees," *Journal of Electromyography and Kinesiology*, vol. 40, pp. 72 – 80, 2018.

[13] E. Scheme and K. Englehart, "On the robustness of EMG features for pattern recognition based myoelectric control; a multi-dataset comparison," *36th Annual International Conference of the IEEE Engineering in Medicine and Biology*, pp. 650–653, August 2014.