Adjacent Features for High-Density EMG Pattern Recognition

Ian M. Donovan, *Student Member, IEEE*, Kazunori Okada, *Member, IEEE*, and Xiaorong Zhang, *Member, IEEE*

***Abstract*—In the infancy of electromyography (EMG) based pattern recognition (PR) limited numbers of electrode channels were anatomically placed over muscles of interest. Modern methods have shown that regularly spaced electrodes around the circumference of a limb are equally effective and have been demonstrated in consumer-ready myoelectric control systems such as Thalmic Labs' Myo armband. In addition to linear arrays, grid arrays have also been applied in this field of research. Although electrode arrays have mainly been adopted to simplify placement, other benefits will be exploited in this work.**

Presented in this paper is a novel spatial-temporal feature set that separately analyzes the intensity and structure of the measured electrical signals (MES) and evaluates the similarities between adjacent electrodes, hence the name Adjacent Features (AF). Results in this paper show that AF produced classification accuracies about 4%-6% greater than autoregression (AR) coefficients and Hudgins' time-domain (TD) features for classifying 47 hand and wrist gestures, while having a computational simplicity similar to the TD features.

I. Introduction

The application of surface electromyogram (sEMG) pattern recognition (PR) in fields such as powered prostheses[1-2], rehabilitation robotics [3], and gesture control interfaces [4] have been widely studied in the past decades. The general process for sEMG PR starts with the segmentation of the raw measured signals into analysis windows, then characteristics or features of the signal are extracted from the windows. Once these values have been concatenated with the features from the other channels, the resulting feature vector is used by a classification method to discriminate which gesture the subject is intending.

Common features used for sEMG characterization include those extracted from time domain (TD) (e.g. Hudgins' TD features [5]), frequency domain (e.g. autoregression (AR)-based features), and time-frequency domain (e.g. wavelet-derived features) [6]. Due to their relatively low computational requirements, TD features have been the most popular in real-time sEMG PR-controlled applications. Most of these feature extraction methods were developed based on a few anatomical placed electrodes and focus on the extraction of sEMG characteristics from individual sEMG channels without exploiting the relations between the channels, or information within the space-domain (SD).

This work is partly supported by the San Francisco State University (SFSU) Ken Fong Translational Research Fund, the SFSU Center for Computing for Life Sciences (CCLS) Mini Grant, and the California State University Program for Education & Research in Biotechnology.

The authors are with SFSU, San Francisco, CA, 94132, USA.

With the introduction of high-density (HD) sEMG arrays in the field of EMG research, information in the SD can be harnessed. These HD sEMG electrodes generally consist 16 of to 256 regularly spaced electrodes and were originally developed for clinical diagnosis of neuromuscular diseases [7] and offline analysis of motor unit activities [8] due to the high computational burden of resolving these finer details. Nevertheless, with the increase in the number of channels and knowledge of the spatial relations of the channels, real-time methods of PR are now being explored [9]. Some of the methods utilize the information from the SD; however, most of these methods still extract the standard single-channel TD features such as intensity, and evaluate the spatial relations of these resulting features to tackle issues such as electrode shift [10]. This neglects the spatial-temporal relations between the raw data.

This paper aims to prove the strength of spatial-temporal features for grid and linear electrode arrays. In particular, a set of novel, computational efficient features based on adjacent channels have been developed to exploit the spatial relationships of sEMG signals. The resulting method is not only computationally efficient for real-time EMG PR; it also results in higher classification accuracies than Hudgins' TD features and AR coefficients for classifying various hand and wrist gestures.

II. Methods

A. Data Collection

This study is conducted with Institutional Review Board (IRB) approval at San Francisco State University (SFSU) and informed consent of subjects. Our dataset includes sEMG data of 47 hand and wrist gestures recorded from 5 subjects. The first 30 gestures are isolated finger movements, subjects are instructed to press, lift, push, pull, deviate right, and deviate left each of their five fingers. The next four gestures are also isolated motions, but of the wrist being lifted, pressed, right (ulnar) deviation, and left (radial) deviation. Four gestures of each finger gripping against the thumb and nine complex hand gestures, including Fist Close, Hand

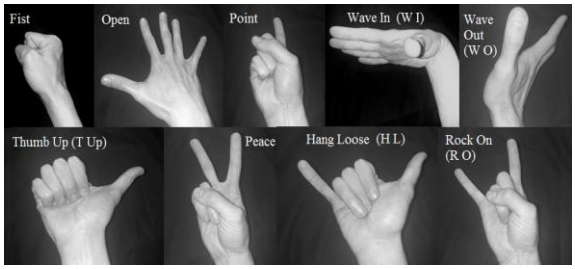

Figure 1. Nine complex gestures of the 47 gesture set.

Open, Index Point, Wave In, Wave Out, Thumb up, Peace, Hang Loose, and Rock On (Figure 1). Data acquisition was conducted with the OT Bioelettronica's EMG-USB2 amplifier at 2048 samples per second with two surface electrode grids (placed on subjects dominate forearm) with 10mm spacing in an 8 by 8 arrangement, resulting in 128 channels.

Each subject preformed the 47 gestures in sequence three times. Each gesture was performed for 6 seconds with 5 seconds rest time in-between. The first and last half second of each gesture was removed to avoid transitions, resulting in 5 seconds of analyzed data.

B. Feature Extraction

Spatial and temporal filters are applied to the raw recorded data. A forth order Butterworth bandpass filter of 10Hz to 500Hz is used and then longitudinal single differential (LSD) (along arm) is applied reducing the data from 8 rows to 7.

The filtered sEMG data is segmented by overlapped analysis windows for feature extraction and pattern classification. The length and increment of the analysis window were set to 200ms (i.e. 410 samples) and 50ms (i.e. 102 samples) respectively as suggested in [11] for real-time control applications.

Standard Features Used for Comparison

The Hudgins' TD features [5] have been used extensively in real-time EMG PR for decades due to the low computational requirements and accuracy. AR features are also commonly used because of their effectiveness in EMG PR and they typically require more computations than TD features. Both of these common feature sets were used in this study for comparison. The TD features included mean absolute value (MAV), wavelength (W), zero crossings (Z), and sign slope changes (T). AR features model the signal within the analysis window as each sample being a polynomial function of the previous values with the addition of a white Gaussian noise. Generally AR features sets are used in combination with an intensity feature like MAV or root mean square (RMS).

Scaled Intensity

An additional feature presented in our previous work [12] is also used in this study, scaled intensity. The intensity of the signal is the most effective and telling indicator of the muscle activity and is calculated with MAV or RMS. Scaled intensity can be applied to any multi-channel system and simply scales out the overall gesture intensity from the individual channels. For example, the scaled MAV feature extracted from input channel i is expressed as

$$SMAV_i = \frac{MAV_i}{\frac{1}{ch}\sum_{j=1}^{ch} MAV_j}$$

where MAV_i is the non-scaled intensity of channel i, ch is the total number of channels, and the denominator is the average of all channels' intensities, or the gesture intensity.

This results in feature vectors from the same gestures being more consistent regardless of the intensity by which the gesture is being preformed. The gesture intensity is relevant to proportional controls, but is not necessary for gesture discrimination, as seen from our previous work [12].

Adjacent Features (AF)

With the use of electrode grid and linear arrays the spatial relations between channels can be exploited to provide features based on the relations of signal shapes from adjacent channels. Whereas standard single channel feature extraction methods attempt to quantize the absolute structure of the individual measured electrical signals (MES), the AF methods use the neighboring MES's structure as a template to evaluate the relative structural similarities.

The sEMG signals being analyzed are the result of multiple pulses or motor unit action potentials (MUAP) propagating through the muscle tissues. Figure 2 shows sEMG signals from seven LSD filtered channels aligned with the muscle fibers while the subject performed the closed fist gesture. Between two neighboring channels, similarities in pulses can be observed, some being synchronized while others show a slight temporal shift (marked in red line in Figure 2) due to the conduction velocity of the MUAPs. The proposed features aim to approximate and quantify these characteristics.

As the focus of these features are on the similarities in the signal shapes or structure, and the signal intensity for each channel is represented by other features, the analysis windows of all channels are normalized to have an intensity (MAV or RMS) of 1.

Consider a single MUAP wave propagating at a slight angle to the longitudinal alignment of the electrode array. Figure 3 demonstrates the process of extracting the proposed AF for one channel (labeled as reference (R)). As shown in the left diagram of Figure 3, a slanted line represents the peak of an idealized MUAP wave traveling in the direction of the arrow. The MES of the three electrodes (transverse (T), reference (R), and longitudinal (L)) in the middle of the diagram show this wave influencing them at the different times. To calculate the AF, the transverse and longitudinal

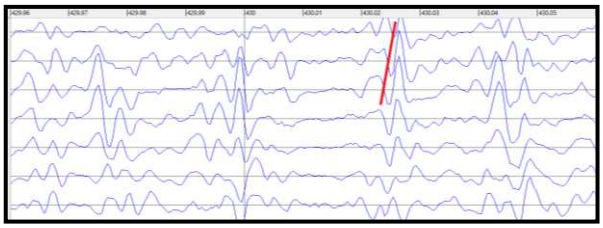

Figure 2. Seven longitudinally aligned channels of EMG for a period of 0.1s. Red line illustrates the effect of conduction velocity.

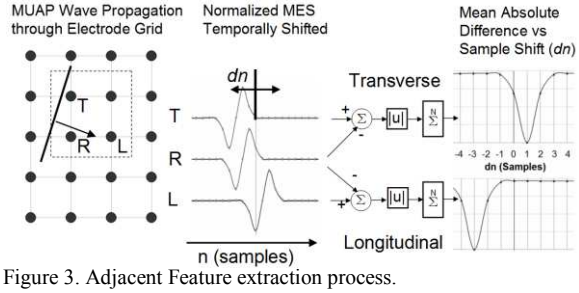

Figure 3. Adjacent Feature extraction process.

MES are shifted forward and backward by *dn* samples, and then compared to the reference MES.

Multiple methods were explored to compare MES in our previous work including correlation coefficients and linear regression; however, a computationally simpler method was found which produced better accuracies [12]. The Mean Absolute Difference (MAD) is similar to MAV of the resulting time series from subtracting one normalized MES from another. Assuming an *i* by *j* grid of MESs, the MAD in the longitudinal and transverse direction are calculated as

$$MAD_L[i,j,dn] = \frac{\sum_{n=1}^{wl} \left| X_{i,j}[n] - X_{i+1,j}[n+dn] \right|}{wl},$$

$$MAD_T[i,j,dn] = \frac{\sum_{n=1}^{wl} \left| X_{i,j}[n] - X_{i,j+1}[n+dn] \right|}{wl}$$

where *n* is the sample within the window, *wl* is the number of samples, *dn* is the number of samples to shift by, and $X_{i,j}$ is the normalized MES at the *i*, *j* location.

Note that the proposed AF cannot be extracted from the last row or column as they require a neighboring MES to compare against. A continuous ring of channels around a limb would avoid this in one dimension; however, in this setup both a row and column of AF are omitted. This results in 6 rows by 7 columns of AF from each of the 8 by 8 grid electrodes.

C. Validation and Testing

Evaluation of the effectiveness of these feature extraction methods was based on the resulting accuracies from classification. A simple linear discriminant analysis (LDA)-based classifier was adopted in this study for pattern classification because of its previous success in EMG PR and its computational efficiency for real-time processing. For each subject, a classification accuracy was calculated using three-fold cross validation. The resulting accuracies were then averaged across all five subjects and the results were reported as the average classification accuracies (ACA).

Along with evaluating these features extracted from the 128 channels discriminating 47 gestures (labeled as *full mesh 47* in the Results section), three additional configurations where also investigated. A subset of input channels from the full mesh was extracted which uses 3 of the 8 rows of raw data to calculate two rows of TD features and a single row of the AF, labeled as the row mesh. The second configuration was to classify 47 gestures using the row-mesh data (*row mesh 47*). The gesture set was also reduced to 13 gestures which still included the 9 complex gestures shown in Figure 1 and the four fingers individually gripping against the thumb. The third and fourth configurations were designed to classify these 13 gestures using both full-mesh (*full mesh 13*) and row-mesh (*row mesh 13*) data.

III. Results

Table 1 shows a selection of feature sets tested on the recorded data and the resulting ACA for the four configurations. The adjacent features are denoted by AF followed by the *dn* used in the longitudinal (L, along muscle fibers) and transverse (T) directions. In our experiments either only a *dn* of zero or a *dn* of zero and a positive and negative integer pair were used for each direction. For example, the feature set 'AF L0 T0,1' refers to 4 values per channel, with *dn* = 0 in the longitudinal and *dn* = -1, 0, and 1 in the transverse.

A. Classification Results with Scaled Intensity

Comparing the first four rows of Table 1, the positive effects of scaled intensity (SMAV) are evident, increasing the ACA by at least 2% in all cases. These results are similar to those from our previous work with a lower resolution acquisition system [12].

B. Classification Results with Adjacent Features

The last eight rows of Table 1 show the resulting ACA for AF sets both with and without the SMAV feature. The four sets of *dn* presented demonstrate three different sizes of feature vectors and the resulting increases in ACA. 'L0 T0' uses no sample shift and represents only two values. The last two sets of *dn* both represent 6 values per channel and show how the selected values of these shifts only slightly alters the accuracies.

As expected, AR features out-perform the standard TD feature sets, but only by as much as the use of SMAV. The use of scaled intensity with AR features does increase AR results (not shown), but not as much as AF sets.

Overall the AF sets performed the best. Increasing the number of *dn* of the AF sets increases the ACA, but with diminishing returns; however, this still shows the benefit of comparing temporally shifted MES of adjacent channels.

One anomaly seen in the results is the slightly lower ACA when including SMAV with the AF features for the full mesh discriminating 13 gestures. Considering this does not happen with the other configurations, it is believed this is due to the large feature vector of the full mesh and the reduced gesture set not fully training the models, or 'the curse of dimensionality '[13].

TABLE I. CLASSIFICATION ACCURURACY RESULTS OF FEATURE SETS

Data Points Used	Full Mesh		Row Mesh	
# of Gestures	47	13	47	13
MAV	72.33%	83.58%	66.47%	79.85%
SMAV	75.93%	86.72%	70.92%	84.72%
MAV, W, ZC, TC	75.71%	86.71%	71.52%	84.55%
SMAV, W, ZC, TC	77.95%	89.54%	73.82%	87.85%
MAV, 4th Order AR	78.19%	88.78%	73.05%	86.69%
MAV, 6th Order AR	78.04%	89.30%	73.50%	86.63%
MAV, 8th Order AR	77.76%	88.94%	73.53%	86.24%
AF L0 T0	78.86%	88.94%	68.35%	83.74%
AF L0 T0,1	81.40%	**90.67%**	74.78%	88.58%
AF L0,2 T0,1	81.71%	90.66%	76.53%	89.73%
AF L0,4 T0,1	81.84%	90.58%	76.35%	88.99%
SMAV,AF L0 T0	79.76%	88.78%	76.03%	89.26%
SMAV,AF L0 T0,1	81.71%	90.05%	78.16%	91.07%
SMAV,AF L0,2 T0,1	81.93%	90.44%	79.06%	**91.78%**
SMAV,AF L0,4 T0,1	**82.11%**	90.37%	**79.13%**	91.67%

C. *Confusion Matrices*

Figures 4 and 5 show the between-gesture confusion matrices of the ACAs derived from the standard TD features (2nd row of Table 1) and from the scaled intensity and AF set (last row of Table 1) respectively; for the reduced mesh and gesture set configurations (last column). Figure 4 shows significant confusion between the thumb grip gestures, as well as a small confusion of wave in and wave out. Figure 5 shows that the AF set noticeably reduces the thumb grip confusions and nearly eliminates the confusion between wave in and wave out.

IV. Discussion and Conclusion

This paper has investigated the effectiveness of spatial-temporal adjacent feature sets including scaled intensity. These methods have shown potential improvements over standard TD feature sets as well as AR features. Though this is based off of data taken from equally space electrode arrays, applications of other electrode configurations may be possible.

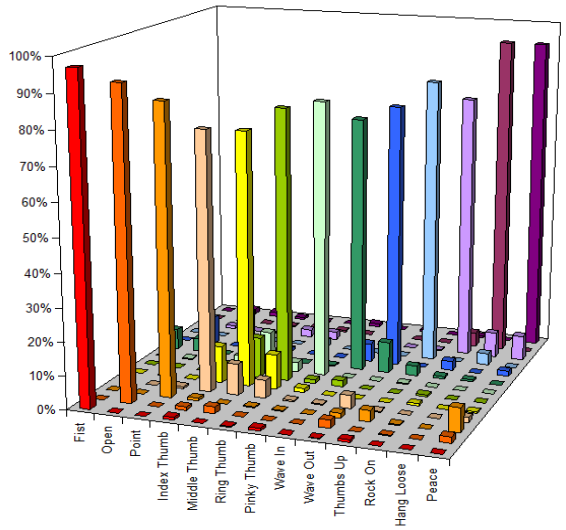

Figure 4. Confusion Matrix from TD features of row mesh discriminating 13 gestures.

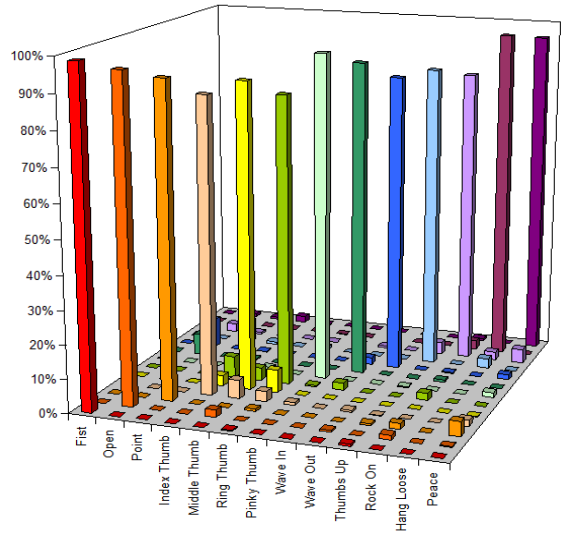

Figure 5. Confusion Matrix from adjacent features of row mesh discriminating 13 gestures.

The AF set computational requirements are comparable with those of TD features and are estimated to be a seventh of the required computational requirements of a 6th order auto-regression. AF requires a single multiplication per sample, per analysis window to normalize the data, then for each series compared (number of values of *dn*) a calculation similar to the one required to calculate wavelength is executed (a single subtraction and absolute summation) per sample per window. This poses significantly less time-complexity compared to the at least 42 multiplications required per sample per window of a 6th order AR. Due to the mathematical simplicity of the AF set, these should be applicable for real-time computing on embedded systems.

The proposed adjacent features have the potential of improving accuracies in other EMG PR-based applications. Future work includes evaluating the AF with more advanced classifiers as well as real-time experiments.

References

[1] Kuiken TA, Li G, Lock BA, Lipschutz RD, Miller LA, Stubblefield KA, Englehart KB: Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms. Jama 2009, 301:619-628.

[2] Zhang X, Liu Y, Zhang F, Ren J, Sun YL, Yang Q, Huang H: On Design and Implementation of Neural-Machine Interface for Artificial Legs. Industrial Informatics, IEEE Transactions on 2012, 8:418-429.

[3] Cesqui B, Tropea P, Micera S, Krebs HI: EMG-based pattern recognition approach in post stroke robot-aided rehabilitation: a feasibility study. Journal of neuroengineering and rehabilitation 2013, 10:75.

[4] Donovan I, Valenzuela K, Ortiz A, Dusheyko S, Jiang H, Okada K, Zhang X: MyoHMI: A Low-Cost and Flexible Platform for Developing Real-Time Human Machine Interface for Myoelectric Controlled Applications.

[5] Hudgins B, Parker P, Scott RN: A new strategy for multifunction myoelectric control. Biomedical Engineering, IEEE Transactions on 1993, 40:82-94.

[6] Oskoei MA, Hu H: Myoelectric control systems—A survey. Biomedical Signal Processing and Control 2007, 2:275-294.

[7] Drost G, Stegeman DF, van Engelen BG, Zwarts MJ: Clinical applications of high-density surface EMG: a systematic review. Journal of Electromyography and Kinesiology 2006, 16:586-602.

[8] Merletti R, Holobar A, Farina D: Analysis of motor units with high-density surface electromyography. Journal of Electromyography and Kinesiology 2008, 18:879-890.

[9] Rahimi A, Benatti S, Kanerva P, Benini L, Rabaey JM: Hyperdimensional biosignal processing: A case study for EMG-based hand gesture recognition. In Rebooting Computing (ICRC), IEEE International Conference on. IEEE; 2016: 1-8.

[10] Stango A, Negro F, Farina D: Spatial correlation of high density EMG signals provides features robust to electrode number and shift in pattern recognition for myocontrol. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2015, 23:189-198.

[11] Smith LH, Hargrove LJ, Lock BA, Kuiken TA: Determining the optimal window length for pattern recognition-based myoelectric control: balancing the competing effects of classification error and controller delay. Neural Systems and Rehabilitation Engineering, IEEE Transactions on 2011, 19:186-192.

[12] Donovan IM, Puchin J, Okada K, Zhang X: Simple Space-Domain Features for Low-Resolution sEMG Pattern Recognition.

[13] Hargrove LJ, Scheme EJ, Englehart KB, Hudgins BS: Multiple binary classifications via linear discriminant analysis for improved controllability of a powered prosthesis. IEEE Transactions on Neural Systems and Rehabilitation Engineering 2010, 18:49-57.