

Classification of Hand Gestures using Surface Electromyography Signals For Upper-Limb Amputees

Gregory Luppescu
Stanford University

Michael Lowney
Stanford University

Raj Shah
Stanford University

I. INTRODUCTION

For our project we classified hand movements based on surface electromyography (EMG) signals from upper-limb amputees. Approximately 38 to 50 percent of patients with upper-limb amputations discontinue use of their prosthetic because the cost of carrying it outweighs its (limited) usage [1]. Machine learning has the potential to greatly improve the functionality of myoelectric prosthetic limbs. After a patient loses a limb, they still contain all the necessary nerves to control their non-existing limb. By using EMG to measure the electrical signals sent through these nerves, amputees can potentially control a robotic prosthetic in the same way that they once controlled their original limb. Much research still needs to be done in order for the latest generation of prosthetics to understand the electrical signals coming from the user. In our project, we utilize Linear Discriminant Analysis, Naïve Bayes, and Support Vector Machines to classify a set of gestures based on EMG signals provided by several upper-limb amputees. We also propose our own collection of features to use for each classifier.

II. DATASET AND FEATURES

A. Dataset

The data set was provided by [2], and consists of raw EMG signals recorded from nine different transradial amputees (seven traumatic and two congenital) with unilateral amputation, where electrodes were placed on and around the stump where a hand (or prosthetic) would be located. Ten pairs of electrodes were placed on each subject, providing 10 channels of EMG data. Each amputee imagined to perform six different hand gestures using their amputated limb. The six gestures in the data set are the spherical grip, index flexion, hook grip, thumb flexion, fine pinch, and tripod grip. For each of the six gestures, three levels of force - low, medium, and high - were imagined as well. Each trial (signal) consists of an 8-12 second holding phase

of a gesture, and 5-8 trials were recorded for each force level, giving a total of over 810 distinct gesture signals. An example of a raw EMG signal is shown in Fig. 1. For our project, we used data from four of the nine subjects. These four subjects all had traumatic amputations on their left forearm.

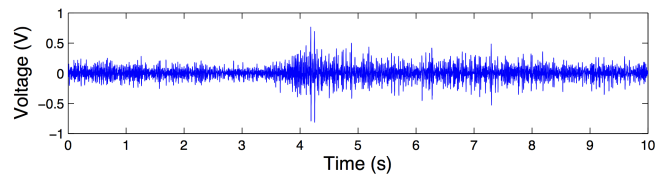


Fig. 1. Raw EMG signal from a single channel



Fig. 2. Experimental setup used in [2] to acquire EMG data

B. Features

A total of 18 features were examined in our project. These features are popular in modern research for EMG pattern recognition [4]. Fifteen of the features were extracted in the time domain, and three were extracted from the frequency domain. Six of the time domain features were time-dependent Power Spectrum density features suggested in [2]. All the features except the time-dependent Power Spectrum density features are listed in Table I. To compute the other features, we first calculate the zero, second, and fourth order root squared moments by

$$\bar{m}_0 = \sqrt{\sum_{i=1}^N x_i^2} \quad \bar{m}_2 = \sqrt{\sum_{i=1}^N \Delta x_i^2} \quad \bar{m}_4 = \sqrt{\sum_{i=1}^N \Delta^2 x_i^2}$$

where N is the total number of samples, and x_i is the sample at index i . The moments are then scaled and normalized by a factor of γ to make them more robust to noise

$$m_k = \frac{\bar{m}_k^\gamma}{\gamma}$$

where a value of $\gamma = 0.1$ was used. The six features f_1, f_2, \dots, f_6 are extracted as follows:

$$f_1 = \log(m_0)$$

$$f_2 = \log(m_0 - m_2)$$

$$f_3 = \log(m_0 - m_4)$$

$$f_4 = \log\left(\frac{m_0}{\sqrt{m_0 - m_2} \sqrt{m_0 - m_4}}\right)$$

$$f_5 = \frac{m_2}{\sqrt{m_0 m_4}}$$

$$f_6 = \log\left(\frac{\sum_{j=0}^{N-1} |\Delta x|}{\sum_{j=0}^{N-1} |\Delta^2 x|}\right)$$

f_4 is a measure of sparseness and f_5 is a measure of the ratio of the zero crossings to the number of peaks as given in [2]. f_6 is simply the ratio of the waveform length of the first derivative to that of the second derivative. In order to negate the effects of force, the final six features are found by a cosine similarity measure between the six features mentioned above and their logarithmically scaled versions. This helps in negating the force effects and only takes into account the orientation of the features [2]. The final 6 time-dependent Power Spectrum density features are defined as

$$F_i = \frac{-2 * f_i * \log(f_i^2)}{f_i^2 + \log(f_i^2)^2}$$

Thus, we have defined all of our features taken into consideration for each channel. Each feature describes a particular aspect of the EMG signal. For example, Waveform Length gives a measure of the EMG signal's complexity, and both Wilson Amplitude and Log Detector give a metric for the level of muscle contraction observed.

Features	Definition
Time domain features	
Mean absolute value (MAV)	$\frac{1}{N} \sum_{i=1}^N x_i $
Integrated EMG (IEMG)	$\sum_{i=1}^N x_i $
Variance (VAR)	$\frac{1}{N-1} \sum_{i=1}^N x_i^2$
Root Mean Square (RMS)	$\frac{1}{N-1} \sum_{i=1}^N x_i^2$
Waveform Length (WL)	$\sum_{i=1}^{N-1} x_{i+1} - x_i $
Log detector (LD)	$\exp(\frac{1}{N} \sum_{i=1}^N \log(x_i))$
Wilson Amplitude (WA)	$\sum_i 1(x_i - x_{i+1} > \epsilon)$
Slope Sign Change (SSC)	The number of times the sign of the slope changes (with a threshold ϵ)
Zero Crossing (ZC)	The number of times the signal crosses a threshold ϵ ($\epsilon=0.05$)
Frequency domain features	
Mean Frequency (MF)	$\sum_{i=1}^M f_j P_j / \sum_{i=1}^M P_j$
Median Frequency (MEF)	Median Frequency of the Power Spectrum
Modified-mean Frequency (MMF)	$\sum_{i=1}^M f_j A_j / \sum_{i=1}^M A_j$
Time domain Power Spectrum Descriptors	
$F_1, F_2, F_3, F_4, F_5, F_6$ as defined in II B.	

x_i - i^{th} sample of the EMG signal, N - Length of the EMG signal, f_j - Frequency at the j^{th} sample of the spectrum, P_j - Power spectrum at f_j , A_j - Amplitude spectrum at f_j , M - Length of the spectrum

TABLE I

THE DEFINITIONS FOR THE FEATURES USED FOR EACH CHANNEL.

III. METHODOLOGY

A. Preprocessing

Before extracting features, it is important to process the raw EMG signals to ensure the features of interest are not obfuscated by various sources of noise. To do this, we apply a cascade of five filters: a high pass fifth order Butterworth filter with a cutoff frequency of 20 Hz, a low pass third order Butterworth filter with a cutoff frequency of 450 Hz, and three notch filters with stop bands centered at 50 Hz, 150 Hz, and 250 Hz [3]. The low-pass and high-pass filters smooth the data and restrict the frequency components of our signal to the frequency range of normal EMG signals, and the notch filters suppress interference caused by power lines and electrical wires (50 Hz in most countries, 60 Hz in the United States).

B. Feature Extraction

The preprocessed signals are segmented into windows of length 300 ms with 50% overlap. This is done to leverage the fact that the signals are "pseudo-stationary" in small regions, and leads to better classification results than if one used non-overlapping windows [4]. For each window, the 18 features described in the previous section are extracted for each of the ten channels, giving a grand total of 180 features per window. An example window is shown in Fig. 3.

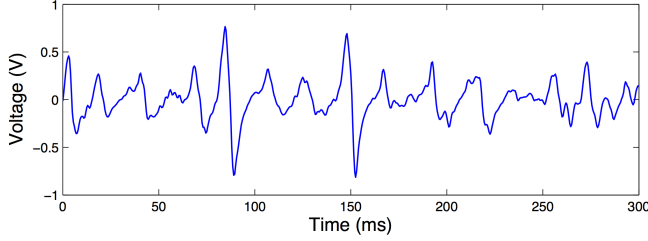


Fig. 3. Windowed EMG signal for which we extracted features

C. Feature Selection

Real time classification is needed for modern myoelectric prosthetics. For this reason, it would be advantageous to reduce the size of the feature set from 180 to a smaller number to reduce computation time. For each classifier, we performed forward feature selection to determine what the optimal features were, and how few features could be used while maintaining high test accuracy.

There are two potential ways to perform feature selection: one in which we consider all 180 features across the ten channels, and one in which we find a subset of the 18 features which is applied to each of the ten channels. Due to changes in electrode placement from subject to subject, we chose to find which of the 18 features per channel would give the best results. Using this method, the optimal feature set can be applied irrespective of the channel placement or number of channels.

Feature selection was run on each individual for each classifier. The test errors were calculated using 5-fold cross-validation. The optimal feature set for each classifier was decided by using a voting scheme to average the results across each of the four subjects. A weighting was applied to each feature depending on its rank in the output of feature selection. We chose the optimal feature set for each classifier based on the features that cumulatively ranked the highest across all individuals.

D. Training Procedure

We trained three different classifiers to predict hand gestures: Linear Discriminate Analysis (LDA), Naïve Bayes, and multiclass SVM. For the Naïve Bayes classifier the feature vectors are discretized into 128 different values. The classifier is then trained using a multivariate model. For the multiclass SVM, we took the "one-vs-all" approach, where we trained a single classifier per class, with the samples of a specific class as positives and all other samples as negatives. For testing, the classifier yielding the highest confidence level determined the predicted gesture. The classifiers were trained on each subject individually due to the fact that EMG signals can vary from person to person due to biological and environmental factors. Two different training and testing procedures were performed, one in which only the gestures were classified (6 classes), and one in which both the gesture and force level were classified (18 classes).

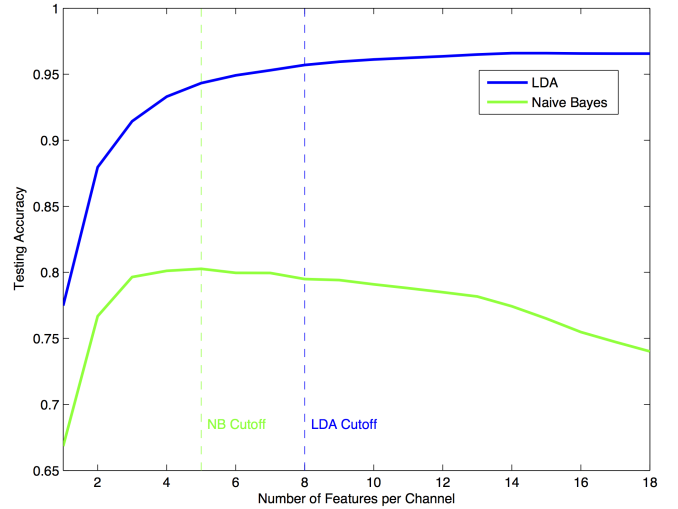


Fig. 4. Test accuracy vs. number of features per channel for LDA and NB classification

IV. RESULTS AND DISCUSSIONS

The feature selection algorithm was run for the classification of hand gestures without force labels. The selected feature sets for each classifier were then used to classify hand gestures with and without force labels. The results of the feature selection algorithm on LDA and Naïve Bayes are shown in Fig 4. The Naïve Bayes classifier quickly experiences over-fitting when there are six or more features. Also, the training and testing accuracy is the lowest for Naïve Bayes. This was expected as Naïve Bayes makes some strong

assumptions about the independence between different features, which is certainly not the case in this application. The optimal feature set for Naïve Bayes was of size five and can be found in Table II. The features are listed in the order they were selected from the feature selection algorithm. The LDA classifier did not experience any over-fitting, but we still decided to pick the top eight features in order to reduce the feature set, which would be needed in real time applications. The optimal feature set for LDA can also be found in Table II.

Classifier	Features used for each channel (by ranking)
LDA	F1, F6, RMS, F4, WL, F5, F3, IE
Naïve Bayes	F6, F5, WA, F1, F3
SVM	WL, F3, F1, F4, RMS, MAV, F5

TABLE II
THE FEATURE SETS CHOSEN FOR EACH CLASSIFIER

The confusion matrices for LDA and NB lead to an interesting result. Among both methods, the two gestures that are mistakenly classified for one another are the hook grip gesture and the spherical grip gesture. These rates of misclassification (when compared to all other rates of misclassification) are in accordance with the fact that the two gestures are similar to each other. In general, the rate of misclassification between two gestures seems to be correlated with the level of similarity between gestures.

Gesture Classification

Classifier	Training Accuracy	Testing Accuracy
LDA	96.53%	96.18%
Naïve Bayes	79.09%	78.65%
SVM	100%	88.76%

TABLE III
TRAINING AND TESTING ACCURACY FOR GESTURE CLASSIFICATION (6 CLASSES)

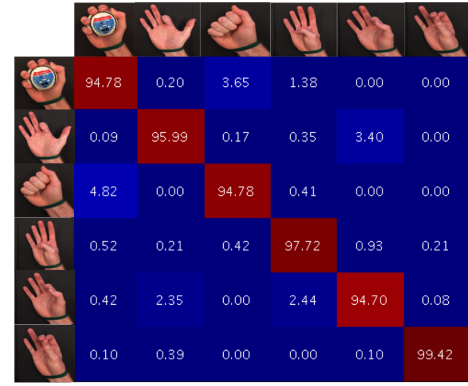
For SVM, the optimal feature set can be found in Table II. Due to time and computational constraints, we only ran feature selection for 14 out of the 18 features, but noted that features after this point were most likely either causing over-fitting, or were not contributing much more to the test accuracy. Thus, we felt it was valid to halt the process early. We initially

Gesture and Force Classification

Classifier	Training Accuracy	Testing Accuracy
LDA	93.99%	93.11%
Naïve Bayes	77.59%	76.86%
SVM	99.94%	86.53%

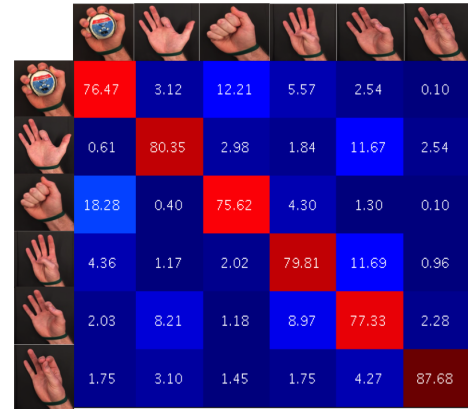
TABLE IV
TRAINING AND TESTING ACCURACY FOR GESTURE AND FORCE CLASSIFICATION (18 CLASSES)

implemented the SVMs with linear kernels, which gave poor testing accuracy. Consequently, we implemented multiclass SVM with Gaussian kernels which achieved high training accuracy, most likely due to the fact that the Gaussian kernel can theoretically fit an infinite number of points. However, the testing accuracy was also considerably high and hence, we decided to go with a Gaussian kernel.



	94.78	0.20	3.65	1.38	0.00	0.00
	0.09	95.99	0.17	0.35	3.40	0.00
	4.82	0.00	94.78	0.41	0.00	0.00
	0.52	0.21	0.42	97.72	0.93	0.21
	0.42	2.35	0.00	2.44	94.70	0.08
	0.10	0.39	0.00	0.00	0.10	99.42

Fig. 5. Confusion matrix for LDA gesture only classification



	76.47	3.12	12.21	5.57	2.54	0.10
	0.61	80.35	2.98	1.84	11.67	2.54
	18.28	0.40	75.62	4.30	1.30	0.10
	4.36	1.17	2.02	79.81	11.69	0.96
	2.03	8.21	1.18	8.97	77.33	2.28
	1.75	3.10	1.45	1.75	4.27	87.68

Fig. 6. Confusion matrix for NB gesture only classification

The feature selection also presented some important observations. Firstly, all the features selected for the

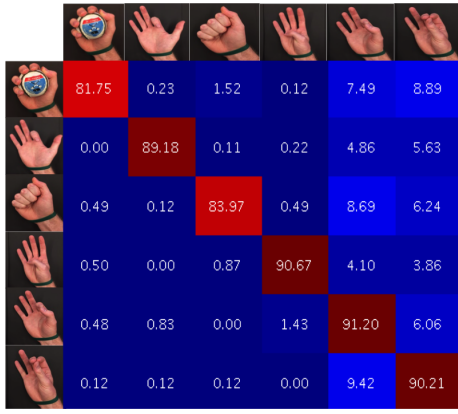


Fig. 7. Confusion matrix for SVM gesture only classification

classifiers were time domain features. This suggests that there is little variance in the frequency domain features to help classify the EMG signals. Also, a number of time-dependent Power Spectrum density features from [2] are a part of the optimal feature set. This implies that these features coupled with others could prove to be a strong feature set for classifying EMG signals.

Finally, LDA proved to be best classifier for this application with high training and test accuracy. Also, it was observed that increasing the number of features did not lead to overfitting for the LDA classifier, unlike Naïve Bayes and SVM. SVM gave high training accuracies, but testing accuracies lesser than LDA hinting that regularization may have helped to improve the testing error.

V. CONCLUSIONS AND FUTURE WORK

In this project, we provided an effective feature set for real time classification of EMG signals using LDA, Naïve Bayes, and multiclass SVM, and showed that LDA provided the best overall performance. We were also able to predict with fairly high accuracy common gestures that would be useful for modern prosthetic limbs. We also demonstrated that the same set of features resulted in fairly high performance for the classification of gesture and force simultaneously. For future work, it is worth exploring ways to make the multiclass SVM classifier more efficient and effective by regularization and by using different kernels. Feature selection can be implemented for all 180 features across all channels, which can help to understand the dependence of features on different channels (different locations on the arm). Also, it would be useful to implement a method to identify important channels

out of the 10 channels. This information could be used to make prosthetic arms less cumbersome by including only the necessary channels for classification. Lastly, it would be worthwhile to explore if the feature sets chosen could generalize to all EMG signals by testing on data acquired from other sources of the body. For instance, one could acquire data on a lower limb amputee to characterize foot movements and use our methodologies for prediction.

ACKNOWLEDGMENT

This project was a part of the CS229 Machine Learning course at Stanford conducted by Prof. John Duchi. We would like to thank Prof. Duchi and the TAs for this insightful course and for guiding us through this project.

REFERENCES

- [1] Dromerick, Alexander W., et al. "Feedforward control strategies of subjects with transradial amputation in planar reaching." *Journal of rehabilitation research and development* 47.3 (2010): 201.
- [2] Al-Timemy, Ali, et al. "Improving the performance against force variation of emg controlled multifunctional upper-limb prostheses for transradial amputees." (2015).
- [3] E. Scheme and K. Englehart, "A comparison of classification based confidence metrics for use in the design of myoelectric control systems," 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, 2015, pp. 7278-7283. doi: 10.1109/EMBC.2015.7320072
- [4] Hakonen M, Piitulainen H, Visala A. Current state of digital signal processing in myoelectric interfaces and related applications. *Biomed Signal Process Control* Elsevier Ltd. 2015;18:33459.