# On the Robustness of EMG Features for Pattern Recognition Based Myoelectric Control; A Multi-Dataset Comparison

E. Scheme, *Member, IEEE*, and K. Englehart, *Senior Member, IEEE*

*Abstract* – The selection of optimal features has long been a subject of debate for pattern recognition based myoelectric control. Studies have compared many features, but have typically used small or constrained data sets. Herein, the performance of various features is evaluated using data from *six* previously reported data sets. The number of channels, the contraction dynamics (dynamic vs static), and classifier type all yielded significant interactions ($p<0.05$) with the feature set. When using 8 channels, the addition of the tested features produced no improvement over a standard time domain (TD) feature set alone ($p>0.05$). When using fewer channels, however, autoregressive, Cepstral coefficients, Willison amplitude and sample entropy features all provided significant improvement during dynamic contractions ($p<0.05$). The simple Willison amplitude is highlighted, showing that it can provide significant improvement when used as a *replacement* for any one of the standard TD features.

## I. Introduction

Pattern recognition based myoelectric control has been a greatly researched subject, particularly over the last decade. Substantial effort has been invested to advance the state of the art and produce a control system that is sufficiently robust for clinical and commercial use. Recent findings have suggested that the traditional design of these systems using constrained experiments translates poorly to clinical performance [1]. Several groups have identified deficiencies that arise when testing statically trained systems with perturbations such electrode shift [2], muscle fatigue [3], residual limb position [4] [5], and varying contraction intensity [6] [7]. While in each case dynamic training approaches were able to mitigate these effects, it is clear that isolated testing of algorithmic advances does not provide a complete picture.

The literature has shown that the proper selection of representative features is of particular importance [8] [9]. Several groups have published detailed comparisons of the various feature extraction methods routinely employed in myoelectric pattern recognition. Oskei *et al.* [9] reviewed several time domain (TD), frequency domain (FD) and time-frequency features. In 2009, Phinyomark *et al.* [10] compared the noise robustness of several EMG features, although only 3 subjects were tested and the noise was simulated using additive white Gaussian noise. More recently, Phinyomark *et al.* [11] evaluated 50 different features of the electromyogram (EMG). They concluded that, over a period of 21 days, sample entropy was the best single EMG feature. These results, however, were obtained from a single able-bodied subject (previously reported by Kaufmann *et al.* [12]), and included only static, steady-state contractions.

In 2010, Tkach *et al.* [13] examined the stability of various features in the presence of electrode shift, induced muscle fatigue, and change in contraction effort/intensity. They used two grids of twelve monopolar electrodes placed over the biceps and triceps to classify four static contractions; *elbow flex/extension* and *wrist pro/supination*. Their results suggested that a pairing of time (*mean absolute value* and *wavelength*) and frequency (*autoregressive* and *Cepstral coefficients*) domain features produced the most robust performance. No consistent favorite was determined, however, and it is unclear how these results might translate to the musculature of the forearm for use by transradial amputees.

Despite their having been a number of studies comparing the suitability of features for EMG pattern recognition [1], no clear consensus appears to have been reached. Furthermore, little emphasis has been placed on the complexity of feature calculations or the robustness of features in varying measurement conditions. The most commonly referenced set of features remains the set of time-domain (TD) features first introduced by Hudgins *et al.* [14]. Herein, we use data from *six* previously reported data sets to evaluate the performance of a selection of promising features with varying numbers of channels and contraction dynamics.

## II. Methodology

Data from *six* previously reported pattern recognition based myoelectric control data sets were used. All data used were collected from able bodied subjects to ease cross study comparisons, and represent over 60 separate data collection sessions (some subjects may have participated in more than one study). From each dataset, 7 classes of motion were extracted: *wrist flexion & extension, wrist pronation & supination, hand open & close* and *no motion*. All data were

Manuscript submitted Feb 12, 2014.
E. Scheme is the Research Engineer at the Institute of Biomedical Engineering at the University of New Brunswick, Fredericton, NB Canada. (phone: 506-458-7029, fax: 506-453-4827)
(email: escheme@unb.ca)
K. Englehart is the Director of the Institute of Biomedical Engineering at the University of New Brunswick, Fredericton, NB Canada.
(e-mail: kengleha@unb.ca).
This work was supported in part by NSERC Discovery Grants 217354.

**Table 1 - A description of the datasets used in this work. The "static or dynamic" column refers to the use of constant or variable intensity contractions during training.**

| DATA SET | DESCRIPTION | # OF SUBJECTS | MAX # OF CHANNELS | STATIC OR DYNAMIC |
|---|---|---|---|---|
| $D_1$ | Confidence Based Rejection for Improved Pattern Recognition [19] | 10 | 6 | Dynamic |
| $D_2$ | Selective Ensemble-Based Classification [20] | 10 | 6 | Dynamic |
| $D_3$ | Principal Components Analysis for Pattern Recognition [21] | 10 | 8 | Static |
| $D_4$ | Effect of Limb Position on Pattern Recognition [4] | 12 | 8 | Static |
| $D_5$ | Effect of Proportional Control on Pattern Recognition [7] | 10 | 8 | Static |
| $D_6$ | Effect of Proportional Control on Pattern Recognition [7] | 10 | 8 | Dynamic |

recorded using bipolar pairs of electrodes, placed equidistantly around the circumference of the dominant forearm at the area of largest muscle bulk (approximately one third the length of the forearm, proximal to the elbow).

Table I briefly summarizes the datasets. The term *static* or *dynamic* in the final column refers to the use of constant or variable intensity contractions during training and testing. The static data sets used subjectively governed constant, medium intensity contractions for training and testing. The dynamic data sets employed a *ramp* approach, where users subjectively increased from rest to a moderate intensity contraction during both training and testing. For further details, please refer to the original studies, as referenced in Table I.

Results were calculated for each trial of each study using an $n$-fold cross validation process, training with $n$-1 repetitions and testing with the $n^{th}$ one iteratively until each repetition is tested. Results were calculated for the each of the Hudgins' time domain features [14]; *slope sign changes*, *zero crossings, mean absolute value* and *wavelength*, as well as *autoregressive* (AR, $6^{th}$ order) [15], *cepstral coefficients* (CC) [16], *Willison amplitude* ($W_{AMP}$) [17] and *sample entropy* (ENT) [18]. Results were computed using both a linear discriminant classifier (LDA) and a multiclass support vector machine with linear kernel (SVM), and repeated for combinations of 4, 6 and 8 channels.

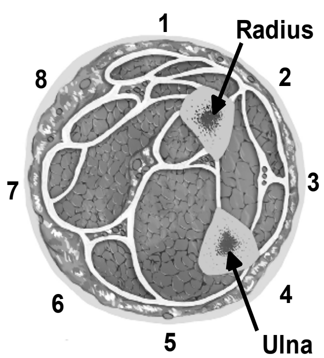


**Figure 1 - Cross-sectional view of the forearm with approximate electrode locations.**

Fig. 1 shows the approximate electrode locations on a cross-sectional image of the forearm. The six channel set comprised the set of {2 3 4 6 7 8}, while the four channel set used channels {2 4 6 8}.

Statistical analysis of the results was conducted using a within factors repeated measures analysis of variance with a significance level of 95%. Post-hoc comparisons were completed using the Tukey-Kramer test.

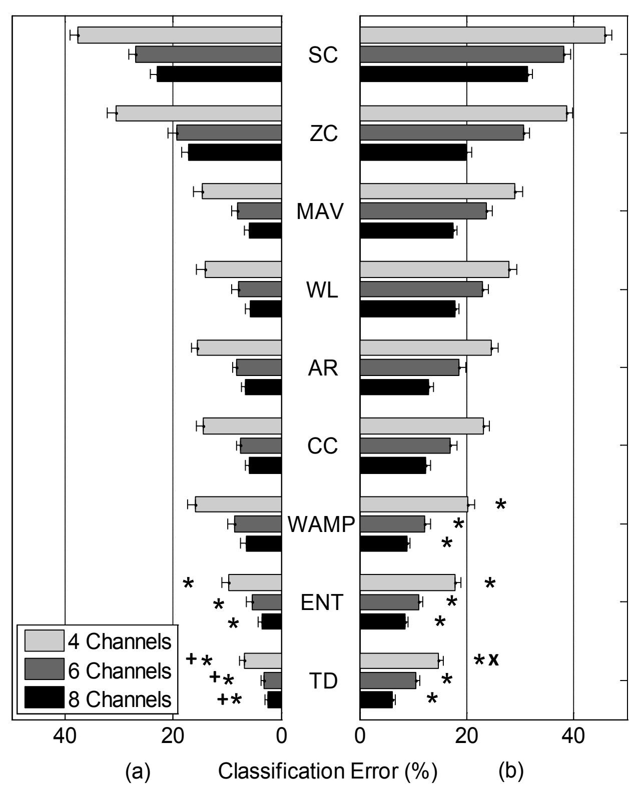


**Figure 2 - Mean LDA classification error (%) of the individual features along with the TD feature set for the static (a) and dynamic (b) data sets. Note that the * symbol denotes a significant improvement ($p$<0.05) over each of the SC, ZC, MAV, WL, AR & CC features. The + and x symbols denote a significant improvement ($p$<0.05) over the $W_{AMP}$, and ENT features, respectively.**

## III. Results

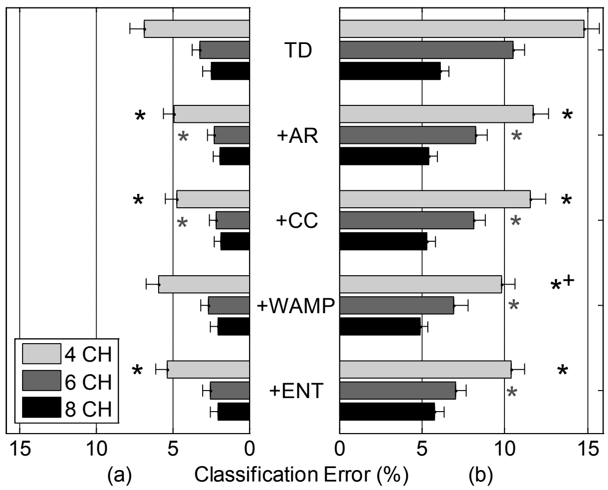


**Figure 3 - Mean LDA error (%) of the TD and TD+ feature sets for the static (a) and dynamic (b) data sets. Note that * and + symbols denote a significant improvement ($p$<0.05) over the TD, and the +AR and +CC feature sets, respectively.**

The performance of each feature is shown in Fig. 2, along with the TD feature set for comparison. The left (a) and right (b) sides of the figure show the results for the static and dynamic data sets, respectively. Note that while the TD features produced the lowest average error in both cases, only the 8CH dynamic data sets showed a significant improvement ($p$<0.05) over the ENT and $W_{AMP}$ features alone. Fig. 3 shows the performance of the full TD feature set compared to the combination of the TD set plus each of: the AR features (+AR), the CC features (+CC), the $W_{AMP}$ features (+$W_{AMP}$), and the ENT features (+ENT). Note that when using 8 channels, none of the features added significant discriminatory ability to the TD feature set. Conversely, when using only 4 channels, nearly all of the features added significant information.

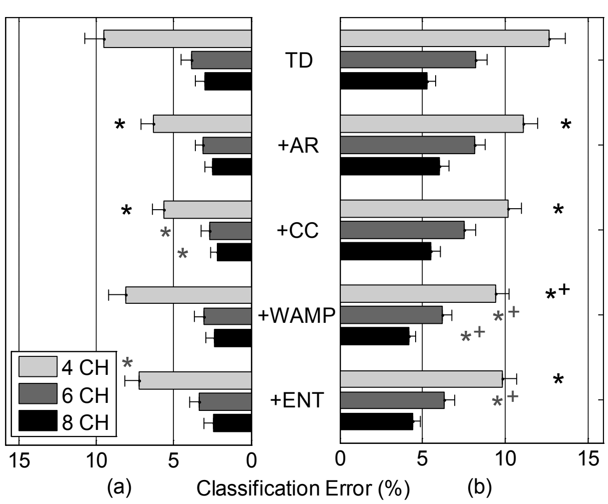


**Figure 4 - Mean SVM error (%) of the TD and TD+ feature sets for the static (a) and dynamic (b) data sets. Note that * and + symbols denote a significant improvement ($p$<0.05) over the TD, and the +AR and +CC feature sets, respectively.**

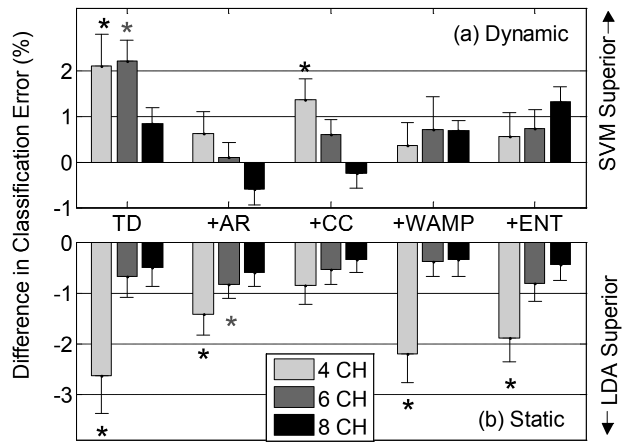


**Figure 5 - Mean difference in classification error (%) between the LDA and the SVM classifiers for the dynamic (a) and static (b) data sets. Results were computed using LDA – SVM, so positive numbers indicate superior SVM performance. Note that the * symbols denote a significant difference ($p$<0.05) between classifiers.**

The computational ease of the $W_{AMP}$ feature and its superior performance during dynamic tasks makes it a desirable addition to Hudgins' TD feature set [14]. Fig. 6 shows that the performance benefit of the +$W_{AMP}$ is maintained ($p$>0.05) even when *replacing* any one of the SC, ZC, MAV or WL features with the $W_{AMP}$.

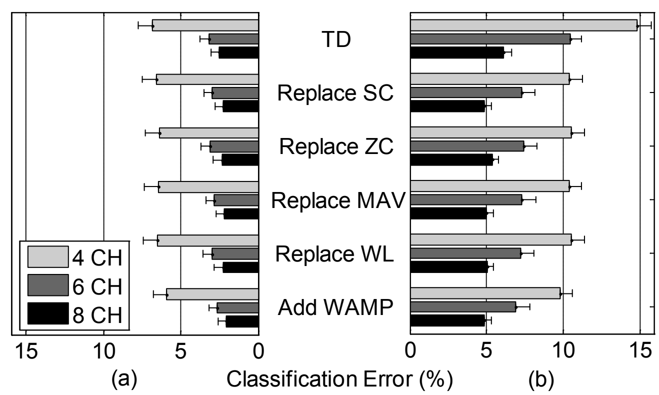


**Figure 6 - Mean classification error (%) when adding the $W_{AMP}$ feature to the TD set, or replacing any one of the TD features with the $W_{AMP}$, for the static (a) and dynamic (b) data sets.**

## IV. Discussion

Data compiled from six different data sets spanning five years were analyzed in this work. In all, these results represent over 60 user sessions, totaling more than 2500 separate contractions. These data were used to examine the performance of different feature sets, classifiers, number of channels and contraction dynamics.

As commonly reported [1], the TD feature set performed well when using 8 channels and static contractions. In fact, no significant improvement was seen when adding additional features, or between the LDA and SVM. As the complexity of the problem was increased, however (by reducing the number of channels or increasing the data variability using dynamic contractions), improvement was achieved by including additional features and differences between the classifiers emerged.

When using static data, the classes are more highly separable resulting in sparsely populated boundaries, rendering the choice of classifier less important. As dynamic data are added, the interclass space becomes more populated placing higher importance on the method of discrimination. The superiority of the SVM in these cases implies that these boundaries may become non-linear. The superiority of the LDA in static experiments is less clear; it may be that the training exemplars used by the SVM in these cases do no readily depict the true decision boundaries.

Although several features offered improvement over Hudgins' TD feature set [14], the performance of the $W_{AMP}$ feature was of particular interest; combined with the TD features, this simple-to-calculate time domain feature significantly ($p$<0.05) outperformed the more complex frequency based +AR and +CC sets in the dynamic data sets. This improvement was consistent for even the 8 channel case when using the SVM classifier. Furthermore, it was found that a similar improvement ($p$>0.05) could be had when *replacing* any one of the TD features with the $W_{AMP}$.

Ultimately, given the desire to improve the clinical reliability and robustness of pattern recognition systems, the more challenging dynamic data sets may provide more meaningful and telling test cases for algorithmic comparisons. This works shows that they reveal differences not readily visible when using highly constrained and separable data.

## References


[1] E. Scheme and K. Englehart, "EMG Pattern Recognition for the 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, 2011.

[2] A. Young, L. Hargrove and T. Kuiken, "Improving myoelectric pattern recognition robustness to electrode shift by changing interelectrode distance and electrode configuration," *IEEE Transactions on Biomedical Engineering,* vol. 59, no. 3, pp. 645-652, Mar 2012.

[3] J. Song, J. Jung, S. Lee and Z. Bien, "Robust EMG pattern recognition to muscular fatigue effect for powered wheelchair control," *Journal of Intelligent and Fuzzy Systems,* vol. 20, no. 1, pp. 3-12, 2009.

[4] E. Scheme, A. Fougner, O. Stavdahl, A. Chan and K. Englehart, "Examining the Adverse Effects of Limb Position on Pattern Recognition Based Myoelectric Control," in *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology*, Buenos Aires, Aug. 2010.

[5] N. Jiang, S. Muceli, B. Graimann and D. Farina, "Effect of arm position on the prediction of kinematics from EMG in amputees," *Medical and Biological Engineering and Computing,* p. Epub, 2012.

[6] T. Lorrain, N. Jiang and D. Farina, "Influence of the training set on the accuracy of surface EMG classification in dynamic contractions for the control of multifunction prostheses," *Journal of Neuroengineering and Rehabilitation,* vol. 8, no. 25, 2011.

[7] E. Scheme and K. Englehart, "Training Strategies for Mitigating the Effect of Proportional on Classification in Pattern Recognition Based Myoelectric Control," *Journal of Prosthetics and Orthotics,* vol. 25, no. 2, pp. 76-83, Apr. 2013.

[8] L. Hargrove, K. Englehart and B. Hudgins, "A Comparison of Surface and Intramuscular Myoeletric Signal Classification," *IEEE Transactions on Biomedical Engineering,* vol. 54, no. 5, pp. 847-853, 2007.

[9] M. Oskoei, "Myoelectric control systems - A survey," *Biomed. Signal Process. Control,* vol. 2, pp. 275-294, 2007.

[10] A. Phinyomark, C. Limsakul and P. Phukpattaranont, " "A novel feature extraction for robust EMG pattern recognition"," *Journal of Computing,* vol. 1, no. 1, pp. 71-80, Dec 2009.

[11] A. Phinyomark, F. Quaine, S. Charbonnier, 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, Sept 2013.

[12] P. Kaufmann, K. Englehart and M. Platzner, "Fluctuating EMG signals: Investigating long-term effects of pattern matching algorithms," in *Conference Proceedings of IEEE EMBS*, Buenos Aires, Argentina, 2010.

[13] D. Tkach, H. Huang and T. Kuiken, "Study of stability of time-domain features for electromyographic pattern recognition," *Journal of Neuroengineering and Rehabilitation,* vol. 7, no. 21, May 2010.

[14] B. Hudgins, P. Parker and R. N. Scott, "A New Strategy for Multifunction Myoelectric Control," *IEEE Transaction on Biomedical Engineering,* vol. 40, pp. 82-94, Jan. 1993.

[15] Y. H. Huang, K. Englehart, B. Hudgins and A. Chan, "A Gaussian Mixture Model Based Classification Scheme for Myoelectric Control of Powered Upper Limb Prostheses," *IEEE Transactions on Biomedical Engineering,* vol. 52, no. 11, pp. 1801-1811, Nov. 2005.

[16] W. J. Kang, J. R. Shiu, C. K. Cheng, J. S. Lai, H. W. Tsao and T. S. Kuo, "The application of cepstral coefficients and maximum likelihood method in EMG pattern recognition," *IEEE Trans. Biomed. Eng.,* vol. 42, no. 8, pp. 777-785, 1995.

[17] M. Zardoshti-Kermani, B. Wheeler, K. Badie and R. Hashemi, "EMG feature evaluation for movement control of upper extremity prostheses," *IEEE Transaction on Rehabiliation Engineering* , vol. 3, no. 4, pp. 324-33, 1995.

[18] X. Zhang and P. Zhou, "Sample Entropy Analysis of Surface EMG for Improved Muscle Activity Onset Detection against Spurious Background Spikes," *Journal of Electromyography and Kinesiology,* vol. 22, no. 6, pp. 901-907, Dec 2012.

[19] E. Scheme, B. Hudguns and K. Englehart, "Confidence Based Rejection for Improved Pattern Recognition Myoelectric Control," *IEEE Trans. Biomed. Eng,* vol. 60, no. 6, pp. 1563-1570, Jun. 2013.

[20] E. Scheme and K. Englehart, "Validation of a Selective Ensemble-Based Classification Scheme for Myoelectric Control Using a Three Dimensional Fitts' Law Test," *Transactions on Neural Systems and Rehabilitation Engineering,* vol. 21, no. 4, pp. 616-623, July 2013.

[21] L. Hargrove, G. Li, K. Englehart and B. Hudgins, "Principal Components Analysis Preprocessing for Improved Classification Accuracies in Pattern-Recognition-Based Myoelectric Control," *IEEE Transactions on Biomedical Engineering,* vol. 56, no. 5, pp. 1407-1414, May 2009.