Gaussian Filtering of EMG Signals for Improved Hand Gesture Classification

I. F. Ghalyan, Z. M. Abouelenin, and V. Kapila

Mechanical and Aerospace Engineering Department, NYU Tandon School of Engineering, Brooklyn, NY, USA
{ibrahim.ghalyan, zma244, vkapila}@nyu.edu

***Abstract*—This paper considers the problem of classifying human hand gestures by using electromyography (EMG) signals that are usually corrupted with noise. Noisy EMG signals result in significant degradation of classification performance and to enhance the performance, a Gaussian Smoothing Filter (GSF) is employed to remove the noise in the sensed EMG signals. The filtered signals, along with various classification schemes, are used to classify several hand gestures. The features of the GSF include: high filtering efficiency, simple implementation, and equal support in frequency and time domains, endowing the GSF with the ability to filter out the noise while partially retaining high frequency components of the original signal. The use of GSF produces smoothed EMG signals that not only enhances the classification accuracy but also reduces the computational time required to develop and test the classifiers. Experiments are conducted on EMG signals, captured from a MYO band, using multiple classification techniques and a significant improvement is observed in the classification performance when using the GSF to filter out the noise in the EMG signals. The classification performance for the EMG signals, for both unfiltered and filtered cases, is compared and the use of GSF is shown to yield significant performance enhancement. Moreover, a significant reduction in the computational time is reported when employing the GSF-based classification, demonstrating the advantages of the GSF for classifying EMG signals. Finally, a comparison is performed for classifying the EMG signals smoothed using a Median Filter (MF) *versus* the GSF and the superiority of the GSF is shown.**

I. INTRODUCTION

The Electromyography (EMG) is an approach for recording the electric response of muscles by measuring the electric potential produced by the muscle cells when they are activated and engaged in a certain action [1]. The EMG signals are widely used in diverse applications, e.g., neuromuscular monitoring in the myasthenia gravis patients [2], computer interface for limb disabled [3], nerve function assessment using needle EMG [4], robot movement control [5], to name a few. Many current applications of EMG study the statistical features of the sensed signals and develop data-driven models that are suited to the application needs. One of these data-driven models is the EMG signals classification, which is considered one of the attractive elements in multiple EMG signals-based applications [6].

One of the earliest efforts to classify EMG signals was reported by Graupe and Cline who employed the autoregressive-moving-average (ARMA) in building a parametric classification approach for interpreting the EMG time series signal [7]. Using various time-frequency representations, promising classification performance was reported by employing Fourier and wavelet transforms [8]. In [9], a real-time EMG signals-based approach was developed, which produced good results, for simultaneously detecting multiple hand motions and learning to adapt to individual human operator with an application to prosthetic hand. Hidden Markov and autoregressive models were combined for developing efficient models of human hand gesture using the sensed EMG signals and promising results were reported [10]. In [11], Support Vector Machine (SVM) was efficiently employed in real-time classification of EMG signals for distinguishing human hand gestures while the arm joint angles were estimated by developing simple linear models relating the EMG signals to the joint angles. Other techniques were suggested for efficiently classifying EMG signals, e.g., k-Nearest Neighbor (k-NN) [12], Linear Discriminant Analysis (LDA) [13], Deep Neural Network (DNN) [14], among others.

This paper suggests improving the classification performance of the EMG signals when using some of the aforementioned classification schemes. The performance of the classification process may be degraded when using the EMG signals corrupted with a significant amount of noise. To overcome this limitation, the EMG signals classification is initaited by employing a Gaussian Smoothing Filter (GSF) to filter out the noise encountered in EMG signals. The main features of the GSF are its simplicity in implementation, its equal support in both frequency and time domains, and excellent noise suppression performance. These features provide a significant impetus for the applicability of the GSF to the EMG classification since the noisy EMG signals have a degrading effect on the classification performance. Moreover, employing the GSF is shown to reduce the computational cost required for training and testing the classification models.

The rest of the paper is organized as follows. Section II describes the classification problem for EMG signals and the noise encountered in EMG signals. Section III explains the GSF and several classification techniques, namely SVM, k-NN, Naive Bayes Classifier (NBC), and LDA used for EMG signals classification. Section IV details the experimental validation and the enhancement in classification when using the GSF in smoothing the EMG signals. Finally, Section V contains concluding remarks and recommendations for future work.

II. PROBLEM DESCRIPTION

Consider the sensed signal of a human hand for two gestures as shown in Figure 1: (a) with the hand closed and (b) with the hand opened. The corresponding sensed

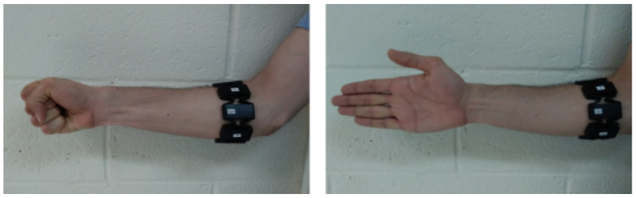

Figure 1. MYO band mounted on an arm with: (a) hand closed and (b) hand opened.

signal from one of the MYO band sensors corresponding to phases in Figure 1 (a) and (b) are shown in Figure 2 (a) and (b), respectively. One can formulate the problem of classifying the two phases (phase_ℓ, $\ell = 1, \ldots, L$, where L denotes the total number of phases and $L = 2$ for the the case of two phases) shown in Figure 1 (a) and (b) as below

$$y_{\text{phase}_\ell}(t) = \begin{cases} 1, & \text{if } x(t) \in \text{phase}_\ell \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

where $y_{\text{phase}_\ell}(t) \in \mathbb{B} \triangleq \{0, 1\}$ is the binary desired classifier output of the ℓ^{th} phase at time instance t and $x \in \mathbb{R}^8$ is the corresponding EMG signal vector. [1] The main objective of the classification process is to develop models that can realize the nonlinear mapping given in (1) as accurately as possible. However, it is obvious from Figure 2 (a) and (b) that the sensed EMG signals are corrupted with noise that can degrade the performance of the classification process. The noise can arise due to sensors, communication, human body, or their combinations. Thus, the features of the noise are unknown and the main objective of this paper is to enhance the classification process by filtering out noise from the sensed EMG signals by employing the GSF before performing the classification step.

III. GSF-Based Enhanced Classification Process

To outline the suggested GSF-based enhanced classification approach, we begin with the review of several related topics, including the concept of GSF and various existing classification techniques. Then, the GSF and classification techniques are used to enhance the process of distinguishing hand gestures using the EMG signals.

A. Gaussian Smoothing Filter (GSF)

The GSF can be characterized by the following impulse response [15]

$$g(x(t)) = \frac{e^{\frac{-x(t)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}, \tag{2}$$

where $x(t)$ is the signal to be smoothed and σ is the standard deviation of the GSF. Since (2) represents the impulse response of the filter under consideration, one can

[1]Note that we have $x \in \mathbb{R}^8$ since the MYO band considered throughout the article is composed of 8 sensors. Alternatively, for cases where we consider q EMG sensors in the device, then x would be a q-dimensional vector, i.e. $x \in \mathbb{R}^q$.

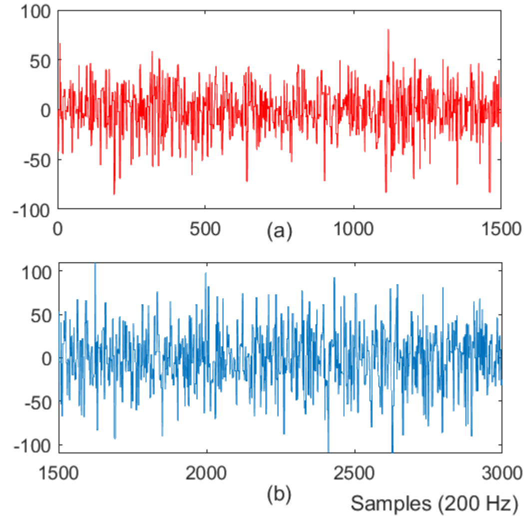

Figure 2. The EMG signals of Figure 1: (a) hand closed and (b) hand opened.

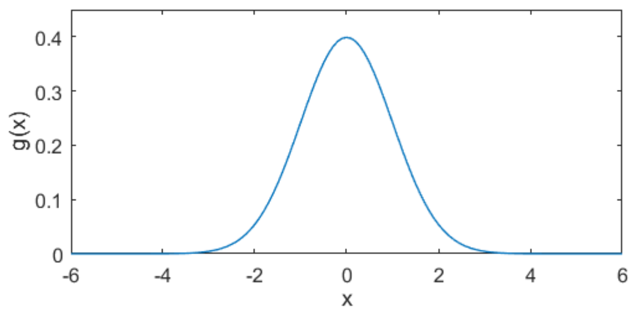

Figure 3. An example of the impulse response of a Gaussian smoothing filter with $\sigma = 1$.

use convolution to obtain the output $\hat{x}(t)$ of the GSF as follows

$$\hat{x}(t) = x(t) \circledast g(x(t)), \tag{3}$$

where $\circledast$ denotes the convolution operator. Next, using the convolution integral, (3) can be rewritten as

$$\hat{x}(t) = \int_{-\infty}^{\infty} x(\tau) g(x(t-\tau)) d\tau. \tag{4}$$

Let x denote the vector of signals captured from noisy sensors and suppose that x is required to develop a classifier with the corresponding target output y. Figure 3 shows the impulse response of the GSF and taking the Fourier transform for (2), we obtain

$$G(f) = e^{-2\pi^2 f^2 \sigma^2}, \tag{5}$$

where f denotes the frequency. Eq. (5) is also a Gaussian function, thus revealing that both the time and frequency domain responses of GSF have a similar support that is a Gaussian function. Even though a GSF behaves like a low-pass filter, its Gaussian function provides a good compromise of, partially retaining high frequency components of the original signal and reducing possible distortions of the signal while smoothing noisy signals.

B. Filtered EMG Signals Classification

Using (4), one can rewrite (1) in terms of the filtered signal as

$$y_{\text{phase}_\ell}(t) = \begin{cases} 1, & \text{if } \hat{x}(t) \in \text{phase}_\ell \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

To realize (6), the existing classification techniques can be employed for developing models that accurately approximate (6). The pattern classification literature offers many techniques that can be employed for such a classification task. Below, we summarize four well-known classification techniques for which further details can be found in the classification and statistical modeling literature (see [16], [17]).

1) Support Vector Machine (SVM): Given the training data set $D = \{(x_1, y_1), ..., (x_N, y_N)\}$, the SVM classification technique can be formulated as a solution to the following quadratic optimization problem [18]

$$\min_{\boldsymbol{w}, \boldsymbol{\xi}_i} J(w, \xi_i) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \xi_i, \tag{7}$$

subject to

$$y_i \psi(w^{\mathrm{T}} x_i) \geq 1 - \xi_i, \tag{8}$$

where (w, ξ_i) are the SVM parameters with $\xi_i \geq 0$, C is a constant vector, and $\psi(\cdot)$ characterizes the classifier. Thus, estimation of w and ξ_i for a given training set produces the information on the boundaries for class separation, which yields good classification performance. The optimization problem of (7), (8) can be solved using various techniques, e.g., the Lagrange multiplier optimization method, which is frequently employed and yields excellent classification performance. See [16], [17] for further details about the SVM technique.

2) k-Nearest Neighbor (k-NN): Given the set of data D, the k-NN technique can be employed in estimating the output of a classifier using

$$\hat{y}(x) = \frac{1}{k} \sum_{x_i \in N_k(x)} y_i, \tag{9}$$

where $\hat{y}(x)$ is the classifier output, $N_k(x)$ is the neighborhood containing k closest points of x, namely, $x_i, i = 1, \dots, k$, and y_i is the class label corresponding to the point x_i. The neighborhood can be specified by using any of the following distance measures, among others,

$$d(x_i, x_j) = \sqrt{\sum_{i=1, i\neq j}^{k} (x_i - x_j)^2}, \text{ (Euclidean)} \tag{10}$$

$$d(x_i, x_j) = \sum_{i=1, i\neq j}^{k} |x_i - x_j|, \text{ (Manhattan)} \tag{11}$$

$$d(x_i, x_j) = (\sum_{i=1, i\neq j}^{k} (x_i - x_j)^p)^{\frac{1}{p}}. \text{ (Minkowski)} \tag{12}$$

Now if $\hat{y}(x)$ is greater than a certain threshold, e.g., 0.5 for binary classification with 0 or 1 outputs, then x is deemed to belong to one class otherwise it belongs to the other class (see [17] for further details about the k-NN technique).

3) Naive Bayes Classification (NBC): Suppose that K is the total number of possible categories, or labels, that y_i can take and let π_k be the prior probability that x_i corresponds to the k^{th} class with $k = 1, 2, \dots, K$. Let $\mathbb{P}_k(x) = p(x = x_i|y = c_k)$ denote the density function of x obtained from the k^{th} class where c_k is the label of the k^{th} class of y. Using Bayes' rule, one can show that

$$p(y = c_k|x = x_i) = \frac{\pi_k \mathbb{P}_k(x)}{\sum_{j=1}^{K} \pi_j \mathbb{P}_j(x)}. \tag{13}$$

In (13), $p(y = c_k|x = x_i)$ is called the posterior probability that suggests $y = c_k$ given the predictor x_i. Thus, the class with the largest posterior probability for a predictor x_i is judged to be the class to which the predictor x_i corresponds and this is called the Naive Bayes Classification (NBC) [16], [17]. Note that $\mathbb{P}_k(x)$ is a key factor in specifying the accuracy of classification and one of the simplest, yet efficient, scheme is to employ the Gaussian function for approximating $\mathbb{P}_k(x)$.

4) Linear Discriminant Analysis (LDA): The LDA classification ([16], [17]) is developed based on the NBC. If one assumes that the data x is drawn from a Gaussian distribution, then by taking the log of (13), one obtains

$$\hat{\delta}_k(x) = x\frac{\hat{\mu}_k}{\hat{\sigma}^2} - \frac{\hat{\mu}_k^2}{2\hat{\sigma}^2} + \log(\hat{\pi}_k), \tag{14}$$

where

$$\hat{\mu}_k = \frac{1}{n_k} \sum_{i:y_i = c_k} x_i, \tag{15}$$

$$\hat{\sigma}^2 = \frac{1}{N - K} \sum_{k=1}^{K} \sum_{i:y_i = c_k} (x_i - \hat{\mu}_k)^2, \tag{16}$$

$$\hat{\pi}_k = \frac{n_k}{N}, \tag{17}$$

n_k is the number of training samples of the k^{th} class. A sample is deemed to belong to a certain class if (14) results in the largest value for that class. This is why (14) is called a discriminant and its linearity with respect to the predictor gives the linear attribute of the LDA classification process.

IV. Experimental Validation

To evaluate the performance of the GSF and its impact on the EMG classification problem for hand gesture recognition, we consider the scenario shown in Figure 4 with six distinct hand gestures: extension, flexion, wrist flexion, wrist extension, pinching, and index extension that are named in this paper as Phases 1, 2, 3, 4, 5, and 6, respectively, (i.e., $\text{phase}_\ell, \ell = 1, \dots, 6$). A Thalmic Labs MYO band, containing eight EMG sensors, is used for capturing the EMG signals of the hand during the aforementioned hand gestures performed by a single subject. The sampling rate of the EMG signals is 200 Hz and the EMG data from the MYO band is communicated using its built-in Bluetooth capability. A laptop computer uses a compatible

Bluetooth 4.0 low energy USB adapter provided with the MYO band to receive the data and stores it for further processing. All computations for signal processing and classification are performed using MATLAB running on the aforementioned 64-bit computer with Microsoft Windows 7 Operating System, an AMD 2 GHz CPU, and 16 GB RAM. Figure 5 shows the unfiltered EMG signals for all six phases of the hand gestures considered in this paper. For the experiments of this section, number of samples used in six phases considered were: Phase 1: 2570, Phase 2: 2121, Phase 3: 2147, Phase 4: 1483, Phase 5: 1161, and Phase 6: 1229. To evaluate the performance of the EMG signals classification considered in this paper, a ten-fold cross validation was employed, as detailed in [17], without any random redistribution of the sensed time-series signals since the time-series is already a random variable and random redistribution of the signals might impose another distribution, affecting the prior distributions of the given samples. Furthermore, it was shown in [19] that the regularized empirical risk functional, for modeling a time-series, is related to the relative positions of the samples of the given time-series signal and changing the sequence might change the modeling problem settings.

Employing the SVM, LDA, NBC, and k-NN to classify the EMG signals of Figure 5 resulted in classification accuracy of 79.82%, 83.33%, 86.31%, and 93.56%, respectively. Note that amplitude values received from the eight channels of the MYO band served as the input features to the four classification algorithms. Using the GSF with $\sigma = 2$, which

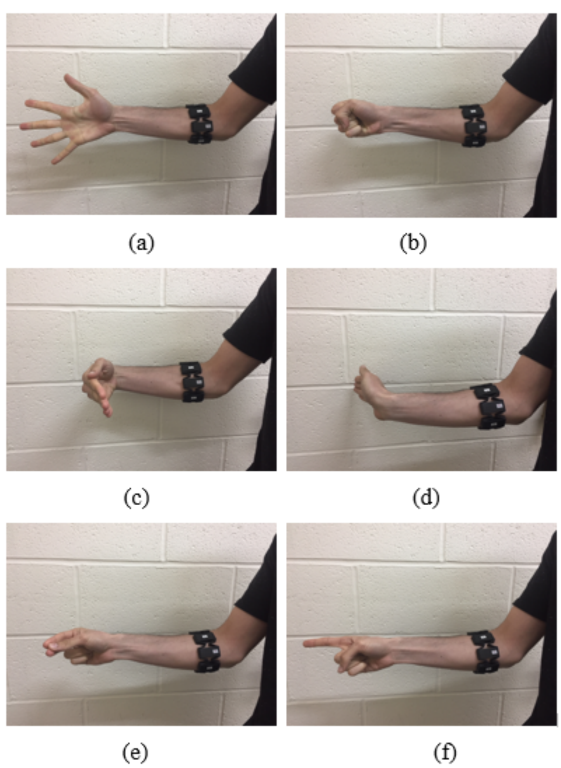

Figure 4. MYO band with multiple situations: (a) extension, (b) flexion, (c) wrist flexion, (d) wrist extension, (e) pinching, and (f) index extension.

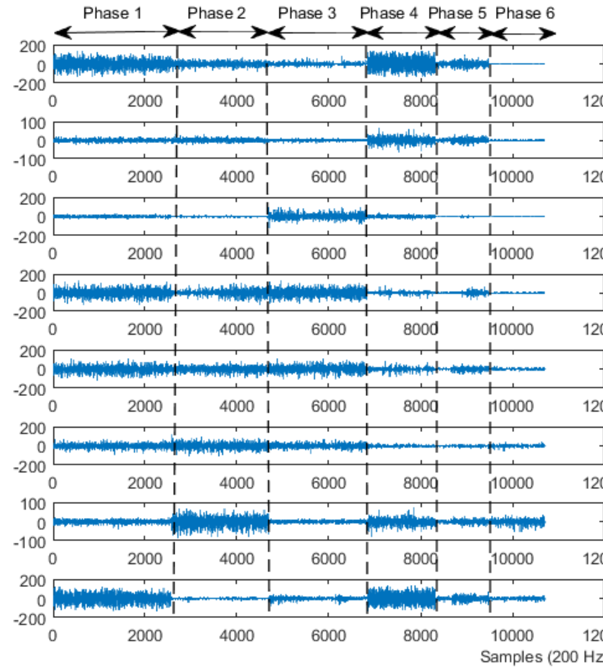

Figure 5. The unfiltered MYO band EMG signals, with signals from eight sensors on separate subplots.

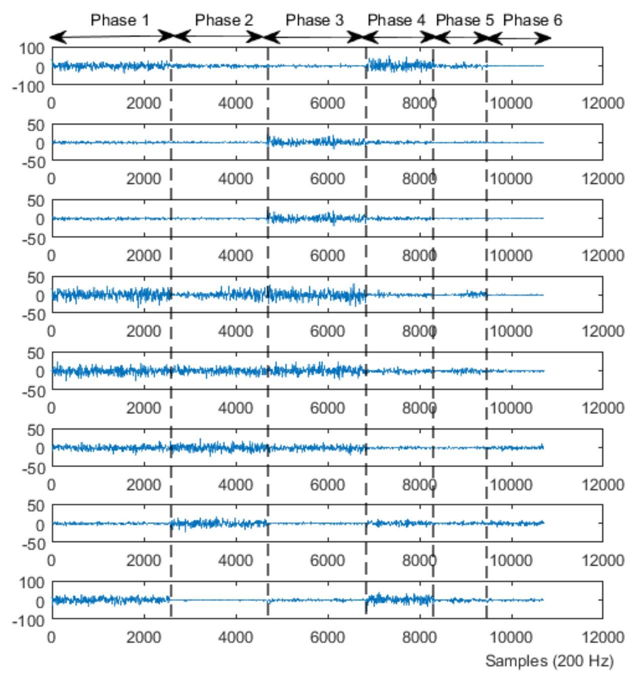

Figure 6. The filtered MYO band EMG signals, with signals from eight sensors on separate subplots.

was manually tuned and selected, the filtered EMG signals shown in Figure 6 are obtained. With the filtered EMG signals of Figure 6 used for classifying the hand gestures under consideration, the SVM, LDA, NBC, and k-NN classification techniques yielded classification accuracy of 94.52%, 93.19%, 93.79%, and 98.09%, respectively. Thus, it is seen that the use of GSF in the EMG classification process results in significant improvement in performance for all four classification techniques considered in this

Table I
ENHANCEMENT OF EMG SIGNALS CLASSIFICATION.

Scheme	Performance (%) without GSF	Performance (%) with GSF
SVM	79.82	94.52
LDA	83.33	93.19
NBC	86.31	93.79
k-NN	93.56	98.09

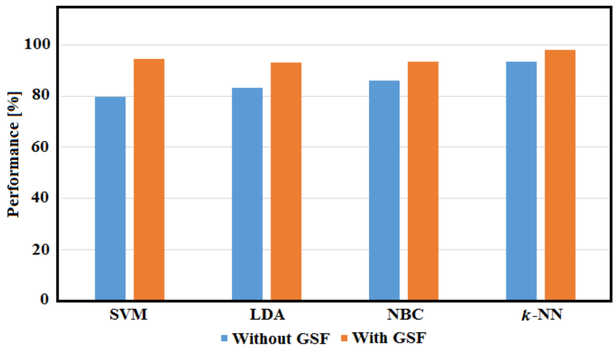

Figure 7. The enhancement in classification performance.

paper. Such a classification performance enhancement is a consequence of filtering out the noise from the EMG signals. Table I and Figure 7 summarize the classification accuracy of the aforementioned techniques with and without the use of GSF in filtering out noise. The GSF filter has only one parameter, the standard deviation σ, and its implementation is simple. The main reason behind the efficient performance of GSF in filtering out the noise stems from the nature of the similarity of supports in both time and frequency domains, i.e. both of them are Gaussian functions, since the frequency transform of a Gaussian function is Gaussian as well. Thus, high frequency noise is eliminated without suppressing and deteriorating the corresponding EMG signal quality rendering efficient noise filtering process that is reflected on the EMG classification task. Figure 7 provides a visual representation of results of Table I and explicitly shows the classification enhancement in the four classification techniques considered in this paper.

Measuring the time required for developing and testing the models of the unfiltered EMG signals resulted a total computational time of 402.96 sec when using the SVM, 9.33 sec for the LDA, 10.62 sec when employing the NBC, and 10.18 sec in the case of k-NN. The corresponding time required for developing and testing the models for the case of filtered EMG signals was found to be 381.11 sec when using the SVM for classifying the EMG signals, 4.14 sec for the case of LDA classifier, 6.54 sec when employing the NBC while the k-NN required 5.14 sec. Table II summarizes this time measurement data. It is obvious from Table II that the measured computational time is reduced significantly, for all classification techniques when using the GSF in filtering out the noise from the sensed EMG signals. The rate of approximation r_N in a learning process is related to the smoothness of a signal by the relation [20]

$$r_N = N^{-\frac{s}{N_d}}, \tag{18}$$

where s is a smoothness measure of the signal and N_d is the dimensionality of the input training space. Thus, for a fixed N and N_d, according to (18) one can deduce that increment in the smoothness s of the EMG signals results in decrement of the rate of approximation r_N leading to a faster approximation of the risk functional of the EMG classification process.[2] When employing the GSF, the value of s increases resulting in reduced values of r_N which speeds up the process of approximating and minimizing the risk functional of the classification process. Figure 8 provides a visual representation of results of Table II and it is obvious that employing the GSF, in filtering out the noise of the EMG signals, significantly reduces the computational cost.

Table II
ACCUMULATIVE COMPUTATIONAL TIME.

Scheme	Time (sec) without GSF	Time (sec) with GSF
SVM	402.96	381.11
LDA	9.33	4.14
NBC	10.62	6.54
k-NN	10.18	5.14

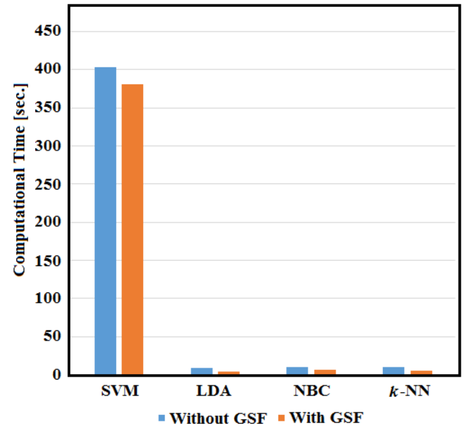

Figure 8. The reduction in computational time.

Employing a 10^{th}-order Median Filter (MF) in smoothing the EMG signals of Figure 5 results in 90.91%, 89.50%, 89.53%, and 95.43% classification performance with SVM, LDA, NBC, and k-NN, respectively. Clearly, the MF provides a significant enhancement in classifying the considered EMG signals. With the previously obtained classification performance when using the raw EMG signals and GSF-based EMG signals, (see Table I, and Figure 7), we compute the improvement in the classification performance of the four techniques considered in this paper with GSF *versus* raw signals and MF *versus* raw signals. These

[2] See [18] for more details about the relation between the risk functional approximation and the smoothness of a function.

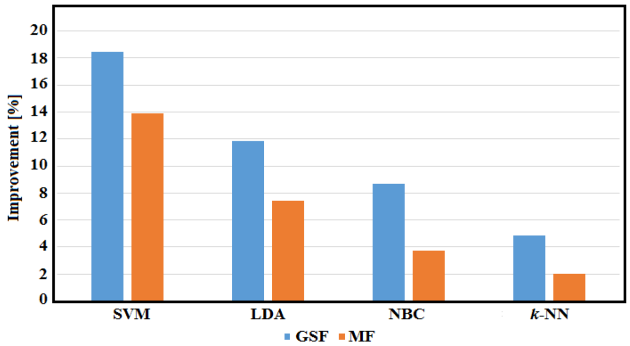

Figure 9. Improvement in classification performance when using the Gaussian Smoothing Filter (GSF) and Median Filter (MF).

improvements in classification performance with the use of smoothing filters are shown in Figure 9, which shows that the GSF-based EMG signals yield superior performance with all four classification techniques. Figure 9 illustrates that the SVM classification, for both cases of the GSF and MF, has the highest value of improvement, indicating that the SVM technique has higher sensitivity to noise. Clearly, from Figure 9, LDA, NBC, and k-NN techniques also witnessed a significant enhancement, reflecting their sensitivity to the noise as well. Noise in the EMG signals results in a change in the decision boundary of the classifiers degrading their performance.

V. Conclusion and Future Work

In this paper, a Gaussian Smoothing Filter (GSF) is employed to enhance the classification process for hand gestures sensed using the Electromyography (EMG) sensors. The sensed EMG signals are shown to be contaminated with a significant amount of noise that results in a degraded classification performance. Employing the GSF, the noise of the EMG signals is filtered out to eliminate the undesirable and unpredictable effect of noise, enhancing the classification performance. Experiments are conducted using a MYO band for a classification scenario consisting of six distinct hand gestures. To examine the performance of the GSF, four EMG signals classification techniques are considered: Support Vector Machine (SVM), k-Nearest Neighbor (k-NN), Naive Bayes Classifier (NBC), and Linear Discriminant Analysis (LDA). Each of the aforementioned classification techniques is shown to yield enhanced performance when presented with signals smoothed using GSF versus the noisy, raw signals. Furthermore, employing the GSF is shown to reduce the computational time of the learning process. Finally, a comparison is performed for classifying the EMG signals smoothed using the Median Filter (MF) versus the GSF, and the GSF is shown to yield superior performance.

Despite the excellent performance reported in this paper with the use of GSF, its standard deviation is not optimized and this may affect the filtering process, producing non-optimal classification results. Thus, future work will focus on developing an enhanced GSF algorithm where the optimal value of the standard deviation is estimated and integrated in the classification process.

Acknowledgment

This work is partially supported by the National Science Foundation grants DRK-12 DRL: 1417769, ITEST DRL: 1614085, and RET Site EEC: 1542286, and NY Space Grant Consortium grant 76156-10488.

References

[1] G. Kamen and D. Gabriel, *Essentials of Electromyography*. Human Kinetics, 2010.
[2] S. Y. Botelho, "Comparison of simultaneously recorded electrical and mechanical activity in myasthenia gravis patients and in partially curarized normal humans," *The American J. Medicine*, vol. 19, no. 5, pp. 693 – 696, 1955.
[3] C. Choi and J. Kim, "A real-time EMG-based assistive computer interface for the upper limb disabled," in *IEEE 10th Int. Conf. on Rehabilitation Robotics*, June 2007, pp. 459–462.
[4] A. Sandoval, "Electrodiagnostics for low back pain," *Phys. Med. Rehabil. Clin. N. Am.*, vol. 21, pp. 767–776, 2010.
[5] S. Sharma and A. K. Dubey, "Movement control of robot in real time using EMG signal," in *2nd Int. Conf. on Power, Control and Embedded Systems*, Dec 2012, pp. 1–4.
[6] N. Nazmi *et al.*, "A review of classification techniques of EMG signals during isotonic and isometric contractions," *Sensors*, vol. 16, no. 8, 2016.
[7] D. Graupe and W. K. Cline, "Functional separation of EMG signals via ARMA identification methods for prosthesis control purposes," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-5, no. 2, pp. 252–259, 1975.
[8] K. Englehart, B. Hudgins, P. Parker, and M. Stevenson, "Classification of the myoelectric signal using time-frequency based representations," *Med. Eng. Phys.*, vol. 21, no. 6-7, pp. 431–438, 1999.
[9] D. Nishikawa, W. Yu, H. Yokoi, and Y. Kakazu, "EMG prosthetic hand controller using real-time learning method," in *IEEE Int. Conf. on Systems, Man, and Cybernetics*, Oct 1999, pp. 153–158.
[10] P. Ju, L. P. Kaelbling, and Y. Singer, "State-based classification of finger gestures from electromyographic signals," in *Proc.* 7^{th} *Int. Conf. Mach. Learn.*, 2000, pp. 439–446.
[11] M. Yoshikawa, M. Mikawa, and K. Tanaka, "Real-time hand motion estimation using EMG signals with support vector machines," in *SICE-ICASE Int. Joint Conf.*, Oct 2006, pp. 593–598.
[12] M. Murugappan, "Electromyogram signal based human emotion classification using KNN and LDA," in *IEEE Int. Conf. on System Engineering and Technology*, June 2011, pp. 106–110.
[13] S. Negi, Y. Kumar, and V. M. Mishra, "Feature extraction and classification for EMG signals using linear discriminant analysis," in *2nd Int. Conf. on Advances in Computing, Communication, Automation (ICACCA)*, Sept 2016, pp. 1–6.
[14] A. D. Orjuela-Cañón, A. F. Ruíz-Olaya, and L. Forero, "Deep neural network for EMG signal classification of wrist position: Preliminary results," in *IEEE Latin American Conf. on Computational Intelligence (LA-CCI)*, Nov 2017, pp. 1–5.
[15] L. G. Shapiro and G. Stockman, *Computer Vision*, 1st ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.
[16] C. M. Bishop, *Pattern Recognition and Machine Learning*. Berlin, Heidelberg: Springer-Verlag, 2006.
[17] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, ser. Springer Series in Statistics. Springer, 2009.
[18] V. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. New York, USA: Springer-Verlag, 2000.
[19] V. Vapnik and R. Izmailov, "V-matrix method of solving statistical inference problems," *Journal of Machine Learning Research*, vol. 16, pp. 1683 – 1730, 2015.
[20] G. Lorentz, *Approximation of Functions*. New York, USA: Holt-Rinehart-Winston, 1966.