A Power Wheelchair Controlled using Hand Gestures, a Single sEMG Sensor, and Guided Under-determined Source Signal Separation

Luis A. Rivera, Student Member, IEEE, and Guilherme N. DeSouza, Senior Member, IEEE
Department of Electrical and Computer Engineering, University of Missouri

Abstract—Surface Electromyographic signals (sEMG) find applications in many areas such as rehabilitation, prosthetics and human-machine interaction. Systems reliant on these muscle-generated electrical signals require some form of machine learning algorithm for recognition of specific patterns of muscle activity. Those systems vary in terms of the signal detection methods, the feature selection and the classification algorithm used, however, in all those cases, the use of multiple sensors is a constant requirement. In this paper, we present a power wheelchair control system that relies on a single sEMG sensor and a new technique for signature recognition called Guided Under-determined Source Signal Separation (GUSSS). Compared to other approaches in the literature, the proposed technique achieves comparable results even when using a simple distance classifier and a very small number of features.

I. INTRODUCTION

The ability to recognize Motor Unit Action Potential Trains (MUAPT) using electromyographic signals collected at the surface of the skin (sEMG) have been used in many applications, including rehabilitation, prosthesis, computer interfacing, exoskeleton robotics, etc. [1], [2], [3], [4], [5], [6]. When it comes to assistive technology, more specifically for power wheelchair control, sEMG signals have often been used as on/off switches. In those cases, menu driven approaches [7], finite state machines [8], and a combination of multiple muscles and sensors [9] are common techniques employed to expand these simple on/off patterns of activation. In general, sEMG-based systems require more sophisticated pattern recognition techniques and they vary widely in terms of the classification approach employed, the feature selection criteria, and the number of sensors used [10], [8], [9].

In terms of the classification algorithm, the most common methods used to classify muscle activity are Artificial Neural Networks (ANN) [4], [11], [5], Fuzzy Logic and Fuzzy Control systems [4], [12]. For example, in [4] an ANN was compared to a Fuzzy Inference System (FIS) for classification and control of a hand prosthesis. In this work, the authors concluded that for their application the best performance was using the FIS classifier which achieved 83% accuracy. In another work [5], several techniques for classification were employed in order to identify hand gestures using sEMG signals extracted from the forearm of human subjects. The authors compared the performance of ANN, Random Forest (RF), 1-Nearest-Neighbor (1NN), Support Vector Machine (SVM), Decision Tree (DT) and Decision Tree with Boosting (DT/B) as possible classification techniques. They reported the ANN as the approach with best performance among those methods.

In terms of feature selection, the features can be extracted from time or time-frequency domains [4], [11], [3]. These features typically include: number of Zero Crossings (ZC), Mean Absolute Value (MAV), Slope Sign Changes (SSC), coefficients of Auto-regressive models (AR) [4], [11]; Absolute Maximum/Minimum, Maximum minus Minimum, Median Value (Med), Variance, Waveform Length (WL) [3]; coefficients of the Short Time Fourier Transform (STFT) [3]; Wavelets Transform (WT) [3], [2], etc.

Given the wide range of features and their large dimensionality, many systems also employ dimensionality reduction techniques. In those cases, Class Separability (CS), Principal Component Analysis (PCA), Analysis of Variance (ANOVA) or Multivariate ANOVA (MANOVA) are the techniques frequently used. In [4], for example, the authors developed a feature selection employing CS and PCA for dimensionality reduction. In that system, as well as in [5] where ANOVA was the technique of choice, the main concern was, as usual, to reduce dimensionality without affecting classification.

Finally, in terms of number of sensors used, as far as we know all systems developed to date have relied on multiple sEMG signals and a large number of features. For example, in [4], the authors reported using two differential sEMG electrodes, multiple features, and PCA to reduce dimensionality of those features. In [5], the system relied on even more sensors – 5 to be more specific – and an ANN as the classification algorithm.

As it can be inferred from the literature, one constant in most systems is the use of a large number of sensors and the use of sophisticated classification algorithms to help coping with a major disadvantage of surface EMG – i.e. the occurrence of cross-talk from adjacent muscles [1]. Our goal in this work is to present a much simpler and yet effective technique using a single EMG sensor, freeing other muscles to be used in other interfaces or to add modalities of operation to the interface.

In this paper, we propose a system for operating a wheelchair that recognizes muscle movements derived from hand gestures. In our framework, we propose a new technique to separate the “cross-talked” MUAPT signals from a single sEMG sensor called “Guided Under-determined Source
Signal Separation” (GUSSS). This technique was inspired on Independent Component Analysis (ICA), but unlike other methods based on ICA, e.g. [12], our method relies on a single sEMG source. Our proposed method combined with a simple distance classifier was applied to the control of a power wheelchair using three hand gestures or eyebrow movements.

II. BACKGROUND AND RELATED WORK

In our method, only two features extracted from a single sEMG signal are used for classification. Since one of these features is based on ICA, in this section we present a quick overview of traditional Blind Source Signal Separation using ICA [13], [14]. In Section III, we explain the proposed technique to eliminate cross-talk, which we named “Guided Under-determined Source Signal Separation” (GUSSS), and the derived GUSSS ratio.

Traditional Blind Source Signal Separation using ICA (BSSS-ICA) is a powerful technique for sEMG signal separation [12]. In those scenarios, it is assumed that a sEMG sensor captures a combination of statistically independent MUAPTs due to cross talk [1], [12]. It is important to notice that each MUAPT is itself a sum of many activations within a muscle. However, here we are interested in separating the entire MUAPT originating from a single muscle. In order to apply BSSS-ICA, each sEMG signal must be captured by a specific sensor placed close to the muscle responsible for that MUAPT.

Mathematically, the goal of BSSS-ICA would be to recover \( N \) source MUAPTs, \( S = [s_1(t), \ldots, s_N(t)]^T \) which are linearly combined, producing the observed signals \( X = [x_1(t), \ldots, x_M(t)]^T \). An analogous example would be that of \( N \) independent sounds emanating from different sources and being detected as mixed signals by \( M \) microphones spread over the space [14]. Figure 1 depicts this idea for three sources and one microphone.

Traditional ICA methods are able to separate the signals whenever \( M \geq N \), that is, the number of observed signals is at least equal to the number of independent sources. In those cases, the sources and the signals can be related in a matrix form such as \( X = AS \) where \( A \) is called the mixing matrix and contains the coefficients of the linear combination of the observed sources. The fact that \( M \) is greater or equal to \( N \) allows for BSSS-ICA to solve an overdetermined system of equations through the expression \( S = A^{-1}X = WX \). The solution is found using a constrained optimization algorithm that maximizes the independence of the signals in \( S \).

For the under-determined cases, that is, when the number of sensors is smaller than the number of independent sources \( (M < N) \), methods for signal separation have been proposed [15], [16] and referred to as Under-determined BSs. However, these methods produce losses in the recovered (separated) sources, which increase with the reduction of the number of sensors.

A. Guided Under-determined Source Signal Separation

As in other systems, here we also assume that an sEMG signal is a mixture of MUAPT originating from different muscles [1]. In other words, the sensed signals are linear combinations of independent MUAPTs due to cross-talk inside the subject’s arm.

In the proposed Guided Under-determined Source Signal Separation, we let \( x_1 \) be such linear combination of \( N \) independent MUAPTs. That is, \( x_1 \) represents a sensed signal from the single sensor and \( s_p \) is a particular known MUAPT, or signature, that the system is trying to identify within the
observed signal \( x_1 \). Since the sensor captures not only \( s_p \), but also various other MUAPTs \( s_i \), we can write:

\[
\begin{align*}
    x_1 &= c_1 s_1 + c_2 s_2 + \cdots + c_p s_p + \cdots + c_N s_N \\
    &= c_p s_p + \sum_{i \neq p} c_i s_i \\
    &= c_p s_p + \hat{s}
\end{align*}
\]  

(1)

where \( c_i, i = 1, \ldots, N \) are unknown mixing coefficients. We assume \( c_i \geq 0 \). The final expression above is simply to stress the fact that \( x_1 \) can be considered a linear combination of the desired signature plus a mixture of other MUAPTs \( \hat{s} \). Since we are interested in separating or identifying only \( s_p \), we will assume that \( \hat{s} \) is independent from \( s_p \). This is an obvious consequence of the assumption that all \( N \) MUAPTs are independent – i.e. if \( N \) MUAPTs are independent, any one of the MUAPTs must also be independent of the remaining \( N - 1 \) MUAPTs. Therefore, an algorithm can successfully identify \( s_p \) within \( x_1 \) iff \( c_p \neq 0 \). So, the question remaining becomes how to determine \( c_p \).

In fact, two situations may arise: the desired signature is indeed present in the mixed signal \( x_1 \), or it is not. In order to distinguish between those two situations, the algorithm creates a second, synthesized signal \( x_p \) by injecting a weighted copy of the particular signature \( s_p \) into the sensed signal \( x_1 \). That is:

\[
x_p = w_1 x_1 + w_p s_p
\]  

(2)

where \( w_1 \) and \( w_p \) are arbitrarily chosen constants. Substituting eq. (1) in eq. (2), we obtain:

\[
x_p = w_1 (c_p s_p + \hat{s}) + w_p s_p = w_1 \hat{s} + (w_1 c_p + w_p) s_p
\]  

(3)

which leads to

\[
x_1 = \hat{s} + c_p s_p \\
x_p = w_1 \hat{s} + k_p s_p
\]

where \( k_p = w_1 c_p + w_p \). Finally, we can express these equations in matrix form as

\[
X_p = AS
\]

where

\[
X_p = \begin{bmatrix} x_1^T \\ x_p^T \end{bmatrix}, \quad A = \begin{bmatrix} 1 & c_p \\ w_1 & k_p \end{bmatrix}, \quad S = \begin{bmatrix} \hat{s}^T \\ s_p^T \end{bmatrix}
\]

The last step of the algorithm is to solve for \( S \). Since we now have two independent components and two linear equations on \( s_p \) and \( \hat{s} \), we can apply a traditional ICA algorithm to separate the \( s_p \) and \( \hat{s} \) components. Moreover, since a sub product of the ICA algorithm is the mixing matrix \( A \), the coefficients of such matrix can be used to infer whether or not a particular signature was present in the sensed signal \( x_1 \). For example, if we consider the case where the particular signature \( s_p \) is not present in the mixture signal \( x_1 \), the mixing coefficient \( c_p \) should be zero. On the other hand, if \( s_p \) is indeed present in the mixture \( x_1 \), that coefficient should be different from zero.

In practice, the coefficient \( c_p \) is never exactly zero. However, it should be small whenever the particular signature is not present in \( x_1 \) and it should be large otherwise. In the proposed framework, we define the GUSSS ratio as:

\[
r_p = \frac{1}{c_p}
\]

Finally, while what constitutes a “large” or a “small” value for the coefficient \( c_p \), may not be obvious, it is clear that the derived GUSSS ratio can be used as a feature for determining if a particular signature is present or not in the sensed signal.

In the proposed framework, the GUSSS ratio is used as such a feature for the distance classifier.

**Identifying multiple signatures in \( x_1 \):** In the previous discussion, we explained how a particular signature can be identified or separated from \( x_1 \). In order to identify the presence or not of all possible signatures, the framework employs an iterative method. That is, first, we assume that the system needs to identify \( N \) MUAPT signatures, each one predominant in one of the \( N \) possible hand gestures. Next, from the test signal \( x_1 \), we obtained \( N \) ratios by injecting iteratively the desired signature into \( x_1 \) – equations (2)-(4). That is, we find

\[
x_p = x_1 + s_p \quad \text{for } p = 1 \text{ to } N
\]

and once again, we apply the ICA algorithm to each

\[
X_p = \begin{bmatrix} x_1^T \\ x_p^T \end{bmatrix} \quad \text{for } p = 1 \text{ to } N
\]

to obtain the ratios \( r_1, r_2, \ldots, r_N \). Finally, it should go without saying that if \( r_1 \) is the smallest of the \( N \) ratios found by the GUSSS, it is likely that the sensed signal \( x_1 \) contains the signature \( s_1 \), and thus, the hand gesture \( i \) is the one being sought.

**B. Mean Absolute Value as a Classification Feature**

We considered a second feature for the classifier: the Mean Absolute Value (MAV) of the signals. The MAV of a signal \( x(t) \) is obtained by calculating the average of the absolute values of \( x \) at all instants \( t \). If the signal is discrete, then

\[
\text{MAV} = \frac{1}{N} \sum_{n=1}^{N} |x(n)|
\]

\[
\text{Noise, inter-dependence between MUAPTs, similarity of gestures, etc. can cause the coefficient not to be zero.}
\]

\[
\text{We will explain how to obtain the signatures in Section IV}
\]
\[
\text{MAV} = \frac{1}{K} \sum_{k=1}^{K} |x(k)|
\]

where \( K \) is the number of samples that constitute \( x(k) \).

C. Classification Module

As we pointed out earlier, the goal of the GUSSS is to identify which signature is present in the observed sEMG signal \( x_1 \). In order to do so, the same signature must be injected to synthesize a secondary signal \( x_p \). Our framework uses a training set of sEMG signals to learn those signatures and in the results section we will explain the approach used for this purpose. Here, we will simply assume that the signatures are available.

Furthermore, from the training signals we can also learn the average MAVs for the different signatures, i.e. for the different gestures. In mathematical terms, let \( \mu_1, \mu_2, \ldots, \mu_N \) be the average MAVs obtained from the training set and corresponding to \( N \) different gestures to be recognized. Let \( \sigma_1, \sigma_2, \ldots, \sigma_N \) be the corresponding standard deviations. Given the input signal \( x_1 \), the algorithm calculates its MAV, \( m_1 \), and based on this value, it computes the Mahalanobis distances to the average MAVs of the gestures. That is:

\[
d_p = \frac{|m_1 - \mu_p|}{\sigma_p} \quad \text{for } p = 1 \text{ to } N
\]

It should be noted that if \( x_1 \) contains a MUAPT that is predominant in gesture \( i \), it is likely that \( m_i \) is similar to \( \mu_i \). In that case, \( d_i \) would also be the smallest of the \( N \) distances above.

Using both the GUSSS ratios and the MAV distances above, we can construct the distance classifier based on the normalization of both features. That is, we define the normalized GUSSS ratio and the normalized MAV distance as, respectively:

\[
\bar{r}_p = \frac{r_p}{\sum_{j=1}^{N} r_j} \quad \bar{d}_p = \frac{d_p}{\sum_{j=1}^{N} d_j}
\]

The features are grouped in the feature vector: \( \bar{v}_p = \left[ \bar{r}_p \bar{d}_p \right] \) for \( p = 1, \ldots, N \), corresponding to each of the \( N \) gestures to be identified. The classification is obtained by assigning \( x_1 \) to that gesture (i.e. the class) for which the corresponding feature vector \( \bar{v}_j \) is smallest. The reason for the normalization of the ratios and distances is, of course, to allow both features to have the same weight in the classification process.

IV. RESULTS

In this section, we explain how we applied the proposed classification framework. For this experiment we used three hand gestures and one resting gesture, which are illustrated in Figure 3. The sEMG signals of interest, i.e., the ones to be associated with each gesture, are those generated during the transition from the rest position to the actual hand gesture and back to the rest position. The signals were acquired using a Tinkertron EMG switch. This device consists of circuitry for detection and amplification of the raw sEMG signals. The signals from the Tinkertron were then sampled by a TS-7250 embedded device from Technologic Systems running the first software module in Figure 2 – Signal Detection and Acquisition. This module monitors the EMG signals, waiting for their levels to cross a certain threshold. Once the threshold is detected, the program stores the signal and then transmits the signal to the second module of the system for the purpose of feature extraction and classification. The latter then sends the commands to the wheelchair on-board processor, which then sends the control signals to the motors.

As we explained before, our framework relies on a single sEMG sensor. So, for this experiment we placed a pair of differential electrodes on the Extensor Carpi Radialis muscle along the forearms of three human subjects. A reference (ground) electrode was also placed on the wrist of the opposite arm of the subjects. We choose this muscle for convenience and ease of use. Later, we also tested the system for eyebrow movement.

As we mentioned earlier, three different hand gestures were considered in our experiments (Figure 3). Before the subjects can use the wheelchair they need to go through a training process. The system allows the user to repeat each gesture a certain number of times and it associates that gesture to one of the possible motions of the chair. The training signals obtained are analyzed and processed to create the signatures, which are then stored in the server.

Once the signatures are learned, the user can start using the wheelchair. But before we present the actual results from our experiment, we must explain the method used to learn the signatures.

A. Obtaining the Signature Signals

Given a training set with \( 3 \times T \) samples – i.e. \( T \) samples from each hand gesture in Figure 3 (b, c and d), we did an averaging of the training signals grouped per hand gesture. That is, each of the \( T \) samples belonging to the same gesture \( p \) were averaged creating a single time signal signature.
A simplified framework for the method consists of the following parts: 1) signal acquisition and pre-processing; 2) feature extraction; and 3) classification. The features used and the classification module implemented here are as described in Section III-C.


d_{p}(t) = \frac{1}{T} \sum_{i=1}^{T} y_{ip}(t), \text{ where } p \text{ is the index of the gesture, } y_{ip} \text{ is the } i^{th} \text{ training signal of gesture } p.

B. Offline Test using National Instruments Digitizer

The main goal here was to validate the proposed GUSSS method for classification of sEMG signals. Therefore, we collected sEMG signals offline and under well-controlled conditions. We processed and analyzed those signals to get classification rates, leaving out the last part of the framework presented in Section III (i.e. command transmission and execution by the wheelchair). A simplified framework for the method consists of the following parts: 1) signal acquisition and pre-processing; 2) feature extraction; and 3) classification. The features used and the classification module implemented here are as described in Section III-C.

\[
\text{Table I: Confusion matrices for offline classification of hand gestures.}
\]

<table>
<thead>
<tr>
<th>Subject 1</th>
<th>Assigned gesture</th>
<th>Actual hand gesture</th>
<th>Executed command</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>clench</td>
<td>stop</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>up</td>
<td>forward</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>tapping</td>
<td>turn</td>
<td>0</td>
</tr>
<tr>
<td>Correct classification: 74%</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{Table II: Confusion matrices for online classification and control of the power wheelchair using hand gestures.}
\]

<table>
<thead>
<tr>
<th>Subject 1</th>
<th>Executed command</th>
<th>Actual hand gesture</th>
<th>Stop</th>
<th>Forward</th>
<th>Turn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>clench</td>
<td>43</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up</td>
<td>7</td>
<td>37</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>tapping</td>
<td>0</td>
<td>19</td>
<td>31</td>
</tr>
<tr>
<td>Correct classification: 74%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\text{Table III: Confusion matrices for classification using eyebrow movements.}
\]

<table>
<thead>
<tr>
<th>Subject 1</th>
<th>Assigned gesture</th>
<th>Real gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Up</td>
<td>Down</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Correct classification: 78%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Three test subjects were asked to perform a total of 120 test movements: 40 movements for each gesture.

For these offline experiments, the signals from the Tektron were sampled using a National Instruments (NI) digitizer. In order to reduce undesired noise coming from the power lines, we implemented a digital filter to remove the 60 Hz component. Table I shows the confusion matrices for the classification results obtained for each one of the human subjects.

C. Online Test with the Wheelchair

A video showing the operation of the wheelchair is included in the video proceedings and it can also be found at http://vigir.missouri.edu/EMGWheelchair.htm

As before, three test subjects were asked to go through the training process using 10 signals per gesture. This time, each subject performed a total of 150 test movements: 50 movements for each gesture. Also, for these experiments, the “clench” movement was assigned to the stop command, the “up” movement was assigned to the forward command and the “finger tapping” was assigned to the turn command. Table II shows the confusion matrices for the classification results of these experiments.

It is important to mention that the results presented here encompass all steps required for the real time control of the wheelchair. In other words, the system had to capture the signal, localize the time window of muscle activity, extract the features and classify the gesture correctly in order to be considered a successful classification. As a result, while still reasonable, the accuracy of the classification presented on these tables is worse than the accuracy in the previous section.

D. Eyebrow Movement Recognition

In a more practical scenario, where severe disabilities lead to limited to no muscle control below the neck, a system relying on eyebrow movements or other facial muscle
activity should be more useful. Here, we present the results from applying the proposed method to the classification of different eyebrow movements. Two human subjects were used for this test: one healthy and the other with a severe muscular dystrophy. Both subjects were allowed to select the eyebrow movements that were easier for them to perform. The first subject selected: eyebrows up, eyebrows down, and crossed eyebrows (one up and one down). The second subject preferred: eyebrow up, eyebrow down, a quick up movement followed by a quick down movement, and a quick down followed by a quick up movement.

For these tests, a pair of differential electrodes was placed on the forehead of the test subjects, right between the eyebrows. The training was done in the same way as for the hand gestures: 10 signals per eyebrow movement were collected to extract the signatures and the MAV parameters. We collected over 40 additional signals per eyebrow movement, and we applied the classification to those signals. These experiments were also done offline, using the NI digitizer. Confusion matrices are shown in Table III.

V. DISCUSSION AND CONCLUSIONS

This paper introduced a novel technique for the extreme case of under-determined source signal separation – i.e. one single observation and multiple sources. We showed how this technique can be employed in robotic assistive technologies, more specifically, control of a power wheelchair. The proposed framework was tested both offline and online. In all tests, a single sEMG sensor was used to recognize three and four different signatures. The classification accuracy obtained in the offline experiments was slightly higher than the accuracy obtained online due to the quality of the AD converter used in each case. Also, the algorithm used online to automatically locate the window containing the EMG signature for classification had an impact on the accuracy of the actual control of the wheelchair.

Despite the problems with the embedded device, the results achieved here were still quite reasonable, especially given the highly reduced number of features and the use of a simple distance classification algorithm. In contrast, other systems found in the literature achieve slightly better results using multiple sEMG sensors and elaborated algorithms for classification. The training process used was also very simple and was carried out practically on-the-fly – which is yet another advantage of our method over other methods in the literature.

VI. FUTURE WORK

Many practical applications, such as prosthetic hands, wheelchairs, etc. would require higher classification rates than the ones obtained here. Also, a larger number of recognizable gestures would be useful for a “real-world” application. In that sense, our framework must be enhanced by the addition of other techniques also found in the literature: as for example, the use of a better classifier – i.e. instead of a distance classifier we could use ANN, FIS, etc. Another improvement for the proposed framework should be easily achievable by the simple addition of extra features. Also, new ways of obtaining signatures during training could be explored. In this paper, we used a small set of features and other simplifications in the classifier in order to emphasize the discriminant power of the GUSS ratio. The main goal was to support the claim that the introduction of GUSS presents a new and very powerful method for separation and identification of patterns in signals, not only in the context of sEMG signals, but also in other areas, as we demonstrated in [17] for Terahertz signals for explosive detection.

ACKNOWLEDGEMENTS

The authors would like to thank Dinal Andreasen from Tinkertron LLC for providing the sEMG switch and for his technical support on using the device.

REFERENCES