Classification of Gesture based on sEMG Decomposition: A Preliminary Study

Anbin Xiong*†, Daohui Zhang*†, Xingang Zhao*†, Jianda Han*, Guangjun Liu**

* State Key Laboratory of Robotics, Shenyang Institute of Automation (SIA) Chinese Academy of Sciences (CAS), Shenyang, Liaoning, 110016, China (e-mail: xiongab, zhangdaohui, zhaoxingang, jdhan@lia.cn)
† University of Chinese Academy of Sciences (CAS), Beijing, 100049, China
** Department of Aerospace Engineering, Ryerson University, Toronto, Canada (e-mail: gjliu@ryerson.ca)

Abstract: Multi-channel surface electromyography (sEMG) recognition has been investigated extensively by researchers over the past several decades. However, due to the nature of sEMG sensors, the more sensors are used, the greater chance for the sEMG to be influenced by environment noise. Furthermore, it is not feasible to use multi-sensors in some cases because of the bulky size of the sensors and the limited area of muscles. This paper proposes a novel sEMG recognition method based on the decomposition of single-channel sEMG. At first, sEMG is acquired while the participant does 5 predetermined hand gestures. Then, this signal is decomposed into its component motor unit potential trains (MUAPTs), which includes 4 steps: 2-order differential filtering, spikes detection, dimension reduction and clustering with Gaussian Mixture Model (GMM). Finally, 5 MUAPTs are obtained and used for hand gestures classification: four features, integral of absolute value (IV), maximum value (MAX), median value of non-zero value (NonZeroMed) and index of NonZeroMed (Ind) are extracted to form feature matrix, which is then classified with the algorithm of Linear Discriminate Analysis (LDA). The classification results indicate this method can achieve an accuracy of 74.7% while the accuracy of traditional classification method for single-channel sEMG is about 52.6%.

Keywords: sEMG, Pattern recognition, sEMG decomposition, Gaussian Mixture Model, Linear Discriminate Analysis.

1. INTRODUCTION

Surface electromyography (sEMG) is a technique to record the electrical activity produced by skeletal muscles and can be readily measured on the skin to provide an assessment to human neuromuscular system (Tassinary et al., 2007). It is safe, noninvasive, real-time and convenient to be implemented. The signals can be analyzed to detect medical abnormalities, activation level, recruitments of motor units and to analyze the biomechanics of human or animal movements (Bilodeau et al, 2003; Chan et al, 2003; Frigo et al, 2000). sEMG is weak, low Signal Noise Ratio (SNR), non-stationary and easy to be influenced by environment noise. Hence, its process and recognition methods have drawn more and more attention.

The conventional sEMG recognition approach is implemented as follows. Multi-channel sEMG are acquired on specific muscles while the researchers are doing some gestures or actions; a certain length of time window and sliding window are used to extract the corresponding time domain, frequency domain and time-frequency domain features, including Integral Absolute Value (IV), Zero Crossing (ZC) (Xiong et al, 2011), Median Frequency (MedFre) (Chang et al, 2013), Power Spectral Density (PSD) (Singh et al, 2007) and Wavelet Transfer Coefficients(WTF) (Li, et al, 2008), etc.; after that, some classification or clustering algorithm such as Neural network (Xiong et al, 2012), Support Vector Machine (Naik et al, 2010) and Linear Discriminate Analysis (Al-Timemy et al, 2013) are employed to recognize different action.

However, it is known to us that sEMG is vulnerable to environmental noise because the sEMG sensors are adhered to the skin surface and the sweat, sebum or large range of motion will inevitably cause the sensor to loosen. Therefore, the more sensors are used, the greater chance for the sEMG signal to be interfered (Scheme et al, 2011). Furthermore, it is not feasible to use multi-sensors in some cases because of the bulky size of the sensors and the limited area of muscles. In this paper, we will present a novel sEMG recognition method based on the decomposition of sEMG, aiming to achieve higher sEMG recognition accuracy with fewer EMG sensors.

sEMG signal is composed of the action potentials (APs) from groups of muscle fibers organized into functional units called motor units (MUs). It is a non-linear summation of the motor unit action potentials (MUAP) in a muscle (De Luca et al 2006; Nawab et al, 2010). Each action potential from each
active unit contributes a small and constant increase in the total electrical signal recorded at the skin. The sEMG decomposition is to decompose the original signal into its component MUAP and hence analyze the recruitment and firing of motor units (Tracy et al, 2005).

When researchers pose different actions, the recruited MUs are different although the amplitudes of sEMG might be almost the same. Consequently, to decompose sEMG into MUAPs; that is, to analyze sEMG from a micro scale, may help to improve the accuracy of recognition of sEMG. Although various decomposition (Stashuk 1999; Florestal et al, 2009; McGill et al, 1985) and corresponding validation (McGill et al, 2011; Parsaei et al, 2012) methods have been proposed, to the best of our knowledge, this paper presents the first sEMG recognition approach based on sEMG decomposition.

The rest of the paper is organized as follows. Section 2 presents the sEMG processing algorithms. The experimental results are analyzed and reported in Section 3. Section 4 draws the conclusions.

2. METHOD

In this study, single-channel sEMG is decomposed into several trains of MUAPs. Then, the MUAPs trains are used for hand gestures recognition. The flowchart of the signal process procedure is illustrated in Fig.1.

![Flowchart of sEMG process](image)

2.1 sEMG Decomposition Method

2.1.1 Signal pre-process and MUAP detection

According to (McGill et al, 1985), sEMG is filtered with a 2-order differential filter, that is,

$$x_i = y_{i+2} - 2y_{i+1} + y_i$$  \hspace{1cm} (1)

where $y_i$ is the sampled raw signal and $x_i$ is the sampled filtered signal, $i$ is the sample time. Then, the threshold $\alpha$ is calculated according to (2)

$$\alpha = c_i \left( \frac{\sum_{i=1}^{N} x_i^2 I(a,t)}{\sum_{i=1}^{N} I(a,t)} \right)$$  \hspace{1cm} (2)

where $c_i$ is a constant which is equal to 3.5, $x_i$ is the sampled sEMG signal and

$$I(a,t) = \begin{cases} 1, & \text{if } |x| < \alpha \\ 0, & \text{otherwise} \end{cases}$$

The filtered signal is scanned for locations where the detection threshold is exceeded. Once a MUAP spike is detected, its firing time is defined by the location of the maximum value found within the next 1 ms (Stashuk 1999).

2.1.2 Dimension Reduction with Principal Component Analysis (PCA)

According to (McGill et al, 1985), the spike is made up of 16 data due to the 2 order differential filter; that is, the dimension of samples for the following cluster algorithm is 16. In order to reduce the required size of samples and computation load, improve the numerical stability and accelerate the convergence, Principal Component Analysis (PCA) is employed (Duda et al, 2012). Principal Component Analysis (PCA) in general transforms multivariate data into a set of linearly independent components that contain eigenvectors or principal axes onto which the original data are projected. To clarify this notation, recall that the calculation of modes is realized by diagonalization of the data’s covariance matrix, to which eigenvectors and eigenvalues correspond (Staudenmann et al, 2006). Representative data are decomposed and ranked in terms of the eigenvalues that, for every mode, reflect its contribution to the data in terms of variance. The exact use of PCA in our study will be introduced in Section 3.2 in detail.

2.1.3 Clustering with Gaussian Mixture Model (GMM)

A Gaussian mixture model is a weighted sum of $M$ component Gaussian densities as given by the equation:

$$p(x | \lambda) = \sum_{i=1}^{M} \omega_i g(x | \mu_i, \Sigma_i)$$  \hspace{1cm} (3)

Where $M$ is determined by Akaike information criterion (AIC) (Akaike 1974), $x$ is a D-dimensional continuous-valued data vector (i.e. the dimension-reduced sEMG data with PCA method), $\omega_i$, $i = 1,...,M$, are the mixture weights, and $g(x | \mu_i, \Sigma_i)$, $i = 1,...,M$, are the component Gaussian densities. Each component density is a D-variate Gaussian function of the form:

$$g(x | \mu, \Sigma) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} \exp \left\{ -\frac{1}{2} (x - \mu) \Sigma^{-1} (x - \mu) \right\}$$  \hspace{1cm} (4)
with mean vector $\mathbf{\mu}_i$ and covariance matrix $\Sigma_i$. The mixture weights satisfy the constraint that $\sum_{i=1}^N \omega_i = 1$.

The complete Gaussian mixture model is parameterized by the mean vectors, covariance matrices and mixture weights from all component densities.

According to the GMM, we can cluster the spikes samples into $M$ clusters, namely $M$ trains of MUAPs.

### 2.2 Motion Recognition based on MUAPs

#### 2.2.1 Feature extraction

On account of the complexity and the short-time stationary random characteristic of the sEMG, it is important to choose the proper features to represent the raw sEMG (Ferrarin et al., 2007). In this study, the EMG segments are divided into overlapping epochs of length 800 ms and 25% overlap; namely, the EMG is segmented with 800ms time windows and 200ms sliding windows. The integral of absolute value (IAV), maximum value (MAX), mean value of non-zero value (NonZeroMed) and index of NonZeroMed (Ind) are then extracted. Their description and calculation are given below.

a) Integral of absolute value (IAV)

$$IAV = \frac{1}{N} \sum_{i=1}^N |x_i|$$  \hspace{1cm} (5)

where $x_i$ is the $i$th sampling point and $N$ is the length of time windows.

b) Maximum value (MAX)

$$MAX = \max_{i=1,\ldots,N} |x_i|$$  \hspace{1cm} (6)

where $x_i$ is the $i$th sampling point and $N$ is the length of time windows.

c) Median value of non-zero value (NonZeroMed)

$$NonZeroMed = \text{median value of} \{x_1, x_2, \ldots, x_N\}$$  \hspace{1cm} (7)

where $x_i$ is the $i$th sampling point and $N$ is the length of time windows.

d) Index of NonZeroMed (Ind)

$$Ind = \text{original location of the NonZeroMed in the time windows}$$  \hspace{1cm} (8)

The selection and the number of features as well as the length of time and sliding windows have been tested several times; and the aforementioned features and windows provide best performance.

#### 2.2.2 Classification with Linear Discriminate Analysis (LDA)

After features are extracted, LDA is utilized to classify the features samples into 5 classes corresponding to 5 predetermined hand gestures. Since the purpose of the classification is not to introduce complex classification method but to show the improvement of the proposed approach, we want to use a well-known and simple classification algorithm.

In this experiment, we define a linear transfer matrix $A$, which projects features extracted from $M$ trains of MUAPs, into a reduced subspace, namely

$$y = A^T D$$  \hspace{1cm} (9)

The training dataset is

$$D = \begin{bmatrix} IAV_{i1} & IAV_{i2} & \cdots & IAV_{ip} \\ MAX_{i1} & MAX_{i2} & \cdots & MAX_{ip} \\ NonZeroMed_{i1} & NonZeroMed_{i2} & \cdots & NonZeroMed_{ip} \\ Ind_{i1} & Ind_{i2} & \cdots & Ind_{ip} \\ \vdots & \vdots & \ddots & \vdots \\ IAV_{p1} & IAV_{p2} & \cdots & IAV_{pp} \\ MAX_{p1} & MAX_{p2} & \cdots & MAX_{pp} \\ NonZeroMed_{p1} & NonZeroMed_{p2} & \cdots & NonZeroMed_{pp} \\ Ind_{p1} & Ind_{p2} & \cdots & Ind_{pp} \end{bmatrix}$$  \hspace{1cm} (10)

where $IAV_{ij}, MAX_{ij}, NonZeroMed_{ij}, Ind_{ij}$ are defined in Section 2.2.1 and $i = 1, 2, \ldots, p$; $j = 1, 2, \ldots, M$. For the subscript $4M \times p$, $4M$ denotes the number of features of $M$ trains of MUAPs and $p$ is the number of training samples.

$$p = \frac{\text{total number of sEMG data}}{\text{Time Window length}} \times \frac{\text{sliding window length}}{\text{sliding window length}}$$  \hspace{1cm} (11)

The reduced subspace is represented as $y = \{y_1, y_2, \ldots, y_p\}$, where $y_j \in \{1, 2, \ldots, k, \ldots, c\}$, $k$ denotes the 4th cluster, and $c = 5$ in this study corresponding to the 5 hand gestures defined earlier in Section 2.2.2.

LDA is mainly based on a set of functions of scatter matrices. The within cluster scatter matrix is defined as

$$S_w = \sum_{i=1}^c \sum_i (D - m_i)(D - m_i)^T$$  \hspace{1cm} (12)

The between clusters scatter matrix is defined as

$$S_b = \sum_{i=1}^c N_i (m_i - m)(m_i - m)^T$$  \hspace{1cm} (13)

where $m = \frac{1}{N} \sum_i D$ and $m_i = \frac{1}{N_i} \sum_i D_i$.

LDA optimization is defined as

$$\max_A \frac{\det (A^T S_b A)}{\det (A^T S_w A)}$$  \hspace{1cm} (14)
Accordingly, the matrix $\mathbf{A}$ can be obtained. The matrix $\mathbf{A} \in \mathbb{R}^{d \times r}$ is made up of $r$ eigenvectors corresponding to the $r$ largest eigenvalues of $S_w' S_w$. The dimension of the matrix $\mathbf{A}$ is determined by the number of clusters $c$.

As mentioned in (9), the subspace $\mathbf{y}$ can be calculated accordingly.

3 EXPERIMENT AND RESULTS

3.1 Equipment and Procedure

In this study, sEMG is acquired while the participant poses predetermined 5 gestures, including rest, fist clenching (FC), palm stretching (PS), index finger pinching (IFP), and middle finger pinching (MFP), and the gesture of rest is used constantly as the transition from one movement to another.

Fig. 2 The 5 gestures: pinching the index finger, pinching the middle finger, rest, clenching the fist, and outstretching the palm.

The sensor used for acquiring sEMG is MyoScan-Flex (Thought Technology Co. Ltd., Canada) with a gain of 1000 and a common rejection mode ratio >100 dB. The disposable surface electrodes (Ag/AgCl) with a diameter of 8 mm and a center to center distance of 20 mm, which adhered to the skin that was cleaned previously with alcohol. The electrode is affixed on flexor digitorum superficialis (see Fig. 3), whose primary function is flexion of the middle phalanges of the fingers at the proximal interphalangeal joints, however under continued action it also flexes the metacarpophalangeal joints and wrist joint (Tubiana et al, 1988). All data are sampled with a frequency of 1kHz and then digitally filtered by a bandpass filter of 2-500Hz and a notch filter of 50Hz for the further process with MATLAB.

3.2 Experiment

In our experiment, the sEMG is recorded while the participant completes the 5 gestures described in Section 3.1. Totally 960000 sEMG data points, are used for training.

The raw sEMG is filtered by the 2-order differential filter mentioned in (1), which can suppress the low-frequency background activity as well as removing the outliers.

Then, the threshold to detect MUAP spikes is calculated to be 39.5 and the spikes are marked with red circles at their peaks locations (Fig. 4). The sEMG data which are less than the threshold are replaced by 0.

![Fig. 4 The filtered sEMG signal and detected spikes which are marked with circles located at the peaks positions.](image-url)

After that, PCA algorithm is used to reduce the dimension of spikes into a lower dimension subspace. The selection of the subspace dimension has been tested several times and we finally choose 4 because it is the compromise between the computation load and clustering accuracy. Then, we use GMM to cluster the dimension-reduced spikes and 5 clusters are obtained due to the AIC.

Furthermore, the MUAP spikes within the same clusters form a MUAP train according to their original location; and the other locations without MUAP are set to be 0. The 5 trains of MUAPs are plotted in Fig. 5(a). We will use the 5 trains of MUAPs to recognize hand gestures later.

A 800ms time window and a 200ms sliding window are employed to extract features, including $I_{AV}$, $MAX$, $NonZeroMed$ and $Ind$. To describe it clearly, we select 1 of the sEMG files to plot it and its corresponding features (see Fig. 5). The red lines denote the onset and end time of hand gestures.

These features are combined to form a matrix, where $p$ can be calculated according to (11):

$$p = \frac{960000 \times 800}{200 = 2397}$$

LDA algorithm is employed to classify the matrix into 5 classes. The classification training result is plotted in Fig. 14. Different colors denote different gestures samples and they scatter in the 3 dimension space. Moreover, the projection matrix $\mathbf{A}$ can be obtained according to (12)-(14), and the matrix $\mathbf{Center}$ is calculated by averaging all the samples in each clusters.
method can achieve a mean accuracy of 74.7%, as shown in Table 1.

<table>
<thead>
<tr>
<th>Gestures</th>
<th>Rest</th>
<th>FC</th>
<th>PS</th>
<th>IFC</th>
<th>MFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.63</td>
<td>0.83</td>
<td>0.79</td>
<td>0.98</td>
<td>0.50</td>
</tr>
<tr>
<td>Mean accuracy</td>
<td>74.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. The classification result for each gesture**

Furthermore, we compare this result with that of a traditional method, which is to classify the single channel sEMG with LDA but without sEMG decomposition. The latter result is shown in Table 2. The mean accuracy is about 52.6%.

<table>
<thead>
<tr>
<th>Gestures</th>
<th>Rest</th>
<th>FC</th>
<th>PS</th>
<th>IFC</th>
<th>MFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>1.00</td>
<td>0.37</td>
<td>0.27</td>
<td>0.48</td>
<td>0.49</td>
</tr>
<tr>
<td>Mean accuracy</td>
<td>52.6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. The classification result for each gesture with LDA only**

4. CONCLUSIONS

In this paper, we propose a novel gesture recognition method based on sEMG decomposition. sEMG on flexor digitorum superficialis is acquired while the participant does 5 predetermined gestures. After that, the single channel sEMG is resolved into 5 trains of MUAPs utilizing 2-order differential filter, spike detection, PCA dimension reduction and GMM clustering. Then, 4 features, such as IAV, MAX, NonZeroMed and Ind features extracted from MUAPs trains, respectively.

Additionally, we implement an online sEMG recognition experiment. 384 seconds of sEMG data are recorded in 12 trials when the participant completes the aforementioned actions. Results are shown in Fig. 6 and Table 1. The x-axis in Fig. 6 is the number of samples (the time windows). The y-axis denotes the gestures. The red line is the real label that we marked during the experiment and the blue dotted line is the estimated result. We can see the estimate errors appear during the transition moment from one action to another.

Fig. 6 The sEMG recognition result. The x-axis is the number of samples (time windows). The y-axis denotes the gestures. The red line is the real label that we marked during the experiment and the blue dotted line is the estimated result. We can see the estimate errors appear during the transition moment from one action to another.

In this paper, we propose a novel gesture recognition method based on sEMG decomposition. sEMG on flexor digitorum superficialis is acquired while the participant does 5 predetermined gestures. After that, the single channel sEMG is resolved into 5 trains of MUAPs utilizing 2-order differential filter, spike detection, PCA dimension reduction and GMM clustering. Then, 4 features, such as IAV, MAX, NonZeroMed and Ind are extracted to form features matrix and finally LDA algorithm is employed to classify it into 5 classes. The accuracy of classification is about 74.7%, which is much higher than that of conventional recognition method for single channel sEMG.
REFERENCES


