CLASSIFICATION OF EMG SIGNALS BASED ON CURVELET TRANSFORM AND RANDOM FOREST TREE METHOD

SUBHANI SHAIK¹, DR. UPPU RAVIBABU²

¹Research Scholar, Department of CSE, Acharya Nagarjuna University, India.
²Professor, Department Of CSE, India.

ABSTRACT

Electromyography (EMG) signal is most powerful signal processing tools for electrical activity of neuromuscular associated with a corresponding muscle. In this paper, the analysis of EMG signals using curvelet transform and Random forest tree is presented. The EMG signal including noise though dissimilar media. The curvelet transform is used for clear away noise from the surface electromyography and superior order of statistics is used for analyzing the signal. The first level is to evaluate the surface of EMG signal and extract features using curvelet transform. The second level is best EMG quality segment was chosen and the rebuilding of the useful data signal was finished using random forest classifier. The intention of this work is introducing a novel approach for discover, analyzing and classifying of EMG signals. The proposed method is applied using clinical dataset and the parameters like mean root mean square, correlation coefficient and absolute value are calculated and to get better quality of class separability. A comparison is made with other traditional methods and the EMG characteristics extracted from rebuilding of EMG signals provide the enhancement of class separability in feature space than. Statistical results shows maximum classification accuracy of 99% and higher information transfer rate is achieved.

Keywords: Electromyography (EMG) Signal, Curvelet Transform, Random Forest Tree, Clinical Dataset.

1. INTRODUCTION

The electromyography is a bio-signal provides muscular activity communication for the generation of movement related signals to drive an assistive device. EMG may be recorded invasively by using pointer electrodes added directly into the muscle over the skin or may be recorded from the surface of the skin without any aggression of the body. Electromyography recordings during muscular activity tasks are frequently used as input signals. Single-trial identification of EMG is one of the key methods in the muscular activity. It can change the skeletal muscular function in the prime sensory domains, so as a result it can serve to generate self-induced variations of the EMG. Classification of electromyography signal is an open area of research in muscular activity and it detects the different states produced by a subject to control an external prosthesis. Muscular techniques are used to assist disabled people to translate these signals to control commands imitating peculiar human thoughts based on electromyography signal processing. In this process still drawbacks for identification and features of existing nonlinearities in the surface electromyography signal, assessment of the level, acquiring correct data due to derivation from normality (M. B. I. Reaz et al. 2006, Abdulhamit Subasi).

In this section, we provide a brief review of articles in different domains of discovery and classification of EMG signals. They must allow the amputee’s volitional muscle manage to be in a way that give accurate estimation of the condition of muscle activity. Graupe et. al(1975) stated that a fourth-order time-series design of the EMG signals can be classified by a linear discrimination function, but this approach enters a huge complexity in calculation. The outcomes of Kelly and Parker’s work stated that a Hopfield neural network could generate AR coefficients from the EMG signals in a
short time. Zardoshti et al (1995) extracted some features such as the number of zero crossings, variance, integral of absolute value and auto-regressive model parameters from upper limb EMG signals and then calculate them with K-nearest neighbor. Kang et al (1995) compared AR and capstrum coefficients and presented that these coefficients are quite helpful to enhance classification rate. The time–frequency transforms technique also introduced as a novel mathematical method in the time–frequency domain. Biomedical signals, especially EMG signals, have been processed by time–frequency transforms in order to extract more characteristic features to enhance the rate of classification of actions. Jung et al (1994) implemented the Wigner–Ville transform on the upper limb EMG signals to classify six dissimilar movements. So the literature includes so many articles which explore the extraction of characteristics from the EMG for managing prosthetic limbs, make quantitative comparison of their quality. Overall, an elevated quality of EMG feature space should have the following properties such as maximum class separability, Robustness and Complexity.

We presented a fresh method depends on curvelet transform for classification of EMG signals. This paper is ordered as follows: The next section, EMG signal mode is presented. Section 3 describes Methodology related to detection and classification of EMG signals. Section 4 describes the results and discussions and lastly finishes with conclusion in Section 5.

2. MODEL FOR THE EMG SIGNAL

The foundation of the EMG signal is the electrical activity of a solo skeletal muscle process. The muscle processes belonging to sole motor unit (MU) are warmed at the neuromuscular junction through their matching motoneuron. This commences an action potential (AP) proliferate along each process from the neuromuscular connection towards the top. Our virtual EMG representation fundamentally contain of two segments. The primary part is the computation of the AP, i.e. alter of the transmembrane potential in time, along a solo skeletal muscle process. For this activity, we make employ of the chemoelectromechanical skeletal muscle design, which uses a total biophysical design and the mono-domain equation to determine for the AP proliferate along a single process. In a second level, the sharing and proliferate of the electric potential within the muscle and in the fat/skin cover has to be calculated, in order to achieve the SEMG signal.

A synopsis of the major equations for level two can be seen in Fig. 1. The equation that has to be resolve for the potential sharing within the muscle field, $\Omega^M$, is the additional cellular bi-domain equation (c.f. Eq. (1)). The potential proliferate through the fat/skin cover and also body part, $\Omega^B$, is characterized by the generalized Laplace equation (c.f. Eq. (2)).

$$\nabla \cdot \left[ \left( \sigma_i(x) + \sigma_g(x) \right) \nabla \phi_e(x, t_k) \right] = - \nabla \cdot \left[ \sigma_i(x) \nabla V_m(x, t_k) \right] \text{in } \Omega^M$$  \hspace{1cm} (1)

$$V_m(x, t_k) \quad \nabla \cdot \left[ \sigma_o(x) \nabla \phi_e(x, t_k) \right] = 0 \text{ in } \Omega^B$$  \hspace{1cm} (2)

$$\phi_e(x) = \phi_o(x) \quad \text{on } T_D^M$$  \hspace{1cm} (3)

$$\left[ \left( \sigma_i(x) + \sigma_g(x) \right) \nabla \phi_e(x, t_k) \right] \cdot n^M = - \left[ \sigma_o(x) \nabla \phi_o(x, t_k) \right] \cdot n^B \quad \text{on } T_N^M$$  \hspace{1cm} (4)

$$\int_{\Omega} \phi(x) \ dx = 0 \quad \text{on } T_N^B$$  \hspace{1cm} (5)

The following equation presents a simple design of the EMG signal: [1]

$$Y(m) = \sum_{s=0}^{M-1} f(s)d(m-s) + w(m)$$  \hspace{1cm} (7)

Where, $y(m)$, new EMG signal, $d(n)$, point calculated, presents the observe impulses, $f(s)$, presents the MUAP, $w(m)$, zero mean obsessive white Gaussian noise and $M$ is the number of motor unit firings. The fundamental model for EMG signal model shown below:

Figure 1: Summary Of The Central Equations For The Calculation Of A Virtual EMG Signal
3. METHODOLOGY

The curvelet Transform an extremely flexible approach to signal decomposition. Whatever an EMG signals is being recorded from a muscle activity; dissimilar types of noise will distort recorded signals. So, processing and classifying EMG signals is tremendously hard for the complex patterns of EMG signals. These signals affected by the anatomical and physiological operations of muscles. The electrical noise which would influence EMG signals can be classified into the following types of noise. They are Motion Artifact, Inherent Instability of Signal, and Inherent Noise in Electronics Equipment, Ambient Noise, Electrocardiographic (ECG) Artifacts and cross Talk etc. These types of noises are eliminated before decomposition of EMG signals.

EMG signals are the superior of muscular activities in signal processing. It is essential to break down the EMG signal to expose the structure connected to muscle and nerve control. Break down of EMG signal has been done by curvelet coefficients and random forest tree. Raw EMG signals offer us precious data in proper form different signal-processing techniques are implemented on raw EMG to attain the accurate EMG signal. The curvelet transform is processed by successive high-pass filtering and low-pass filtering in the discrete time domain. The CT of a signal x(n) is measured by transmit it through a series of different filters. Initially the samples x(n) are passed through a low-pass filter with impulse reaction g(n), resulting in a difficulty of y(n). The signal also goes at the same time through a high-pass filter with impulse response h(n). The outcomes give the details of coefficients and the estimation coefficients. A threshold is fixed for the raw EMG signal which is practiced on the curvelet coefficients after the CT. The curvelet transform coefficients are used to guess the noise and compute the threshold. Curvelet transform of a time domain signal x(n) is described as

\[ C(j, l, k) = \int f(x) \varphi_{j, l, k}(x) dx \]  

(2)

Since digital curvelet transforms activate in the frequency domain, it will be show valuable to employ Plancherel’s theorem. The entire product of internal as the integral over the frequency plane.

\[ C(j, l, k) = \int \int f(\omega) \tilde{\phi}_{j, l, k}(\omega) d\omega = \int \int f(\omega) V_j(R_{\theta}(\omega)) e^{i \theta} \tilde{\phi}_{j, l, k}(\omega) d\omega \]  

(3)

In wavelet concept, we use coarse scale elements and launch the low-pass window \( \delta_0 \) accepting

\[ |\delta_0(\omega)|^2 + \sum_{k \in \mathbb{Z}} |\delta(2^{-j} \omega)|^2 = 1 \]  

(4)

for \( k_1, k_2 \in \mathbb{Z} \), describe coarse scale curvelets as

\[ \Phi_{j_0, k}(y) = \varphi_{j_0}(y - 2^{-j_0} k), \tilde{\varphi}_{j_0}(\omega) = 2^{-j_0} \delta_0(2^{-j_0} \omega) \]  

(5)

So, coarse scale curvelets are non-directional. The complete curvelet transform contain fine scale directional elements \( (\varphi_l, l, k) \geq j_{\text{max}}, l, k \) of the coarse-scale isotropic family of wavelets \( (\Phi_{j_0, k}) \) k.

Random forest (as shown in fig 4) uses more independent decision trees which are formed by randomly chosen variables and the
independent trees are built by an algorithm. Random forest composes number of tree predictors. Random forest composes number of tree predictors.

In this forest tree, every tree is presented by a random vector which is autonomously picked from the same distribution in the forest. The number of the trees enlarges in the forest. The limit assembles to a generalization error. The power of the singular tree association amid the trees affects of generalization error. Every tree in the forest generates a outcome, they are designated for the majority passable class.

The pseudo procedure for analysis of EMG signal analysis based on curvelet transforms is shown below:

Figure 3: Basic Model For EMG Feature Extraction

The pseudo procedure for analysis of EMG signal analysis based on curvelet transforms is shown below:
Algorithm: EMG Signal analysis based on Curvelet Transform

1. Task: Classification of EMG signals using Curvelet transform.
2. Parameter: Cross validation folds, Random forest classifier.

Methodology:

3. For each EMG signals do the following steps:
4. Before decomposition, apply pre-processing techniques for removing the noise.
5. Analyze the surface EMG signal and extract features using Curvelet transform:

\[
C(j, l, k) = (f, \varphi_{j,l,k}(Y)) = \int_{\mathbb{R}^2} f(X) \varphi_{j,l,k}(Y) \, d(Y)
\]

6. The rebuild of the data signal was done using random forest classifier.
7. END

Output: EMG features extracted from rebuilded EMG signals of the primary level and the secondary level coefficients give up the enhancement feature space of class separability.

The following parameters are calculated for each dataset in order to improve the class severability property:

The correlation coefficient of two variables in a data set equivalent to their covariance segmented by the product of their corresponding standard deviations. It is normalized computation of two linearly connected. Generally, the correlation coefficient is described by the following formula,

\[
r = \frac{\text{Covariance}(x,y)}{SD(x) \cdot SD(y)} = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{N} (y_i - \bar{y})^2}}
\]

(6)

Where: \( N \) = number of score, \( \Sigma xy \) = sum of the product of the scores, \( \Sigma x = \) sum of x scores, \( \Sigma y = \) sum of y scores \( \Sigma x^2 = \) sum of squared x scores \( \Sigma y^2 = \) sum of squared y scores. Similarly, the correlation coefficient is described as follows, where \( \sigma_x \) and \( \sigma_y \) are the standard deviations, and \( \sigma_{ab} \) is the covariance. If statistics is defined by standard scores:

\[
r = \frac{1}{n-1} \sum_{i=1}^{n} \left( \frac{x_i - \bar{x}}{s_x} \right) \left( \frac{y_i - \bar{y}}{s_y} \right)
\]

(8)

Where \( \bar{x}, \bar{y}, \bar{x}, \) and \( s_x \) - standard score, sample mean and sample standard deviations are measured using \( (N-1) \) in denominator. If the data comes from the population, then

\[
\rho = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{x_i - \mu_x}{\sigma_x} \right) \left( \frac{y_i - \mu_y}{\sigma_y} \right) \frac{X_j - \mu_x}{\sigma_x}
\]

(7)

\( \mu_x, \sigma_x \) - standard score, sample mean and sample standard deviations are computed using \( (N) \) in denominator.

If the correlation coefficient is near to one, it would point out that the variables are absolutely linear associated and the scatter plot falls nearly a straight line with positive slope. For -1, it describe that the variables are negatively linear associated and the scatter plot nearly falls along a straight line with negative slope. And for zero, it would point out a weak linear association between the variables. The mean absolute error is one of the important correlate forecasts with their ultimate outcomes. The mean absolute error formula is given by

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |f_i - y_i|
\]

(8)

The mean absolute error is a modest of the absolute errors \( |e_i| = |f_i - y_i| \), where \( f_i \) is the retrieve and \( y_i \) the right value. RMSE is a quadratic count rule that also compute the standard magnitude of the error. It’s the square root of the moderate of squared dissimilarities between prediction and actual observations.
The employ of RMSE is very common and it makes an outstanding general purpose error metric for arithmetical predictions. Moderate to the similar Mean Absolute Error, RMSE amplifies and severely abuse big errors.

\[
\text{MSE} = \frac{1}{m \times n} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [f(i,j) - g(i,j)]^2
\]

\[
\text{RMSE} = \sqrt{\frac{1}{m \times n} \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} [f(i,j) - g(i,j)]^2}
\]

\[
\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{MSE}} \quad (9)
\]

The relative absolute error is tremendously alike to the relative squared error in the sense relation to a plain predictor, which is just the modest of the real values. In this study of PSNR, the fault is currently the complete absolute error as an option of the whole squared error. Thus, the relative absolute error catches the whole absolute error and normalizes it and divide by the entire absolute error of the plain predictor.

Relative error \( \frac{\Delta A}{A} = [\frac{\Delta a}{a} + \frac{\Delta b}{b}] \)

Absolute error: \( \Delta A = [\frac{\Delta a}{a} + \frac{\Delta b}{b}] a = [\frac{\Delta a}{a} + \frac{\Delta b}{b}] ab = (\Delta a)b + (\Delta b)a \quad (11) \)

The experimental results are carried out using clinical dataset that is shown in the next session.

4. RESULTS AND DISCUSSION

In this paper, we presented the classification EMG signals Curvelet transform and neural networks. Here, two-dimensional EMG signals are decomposes a set of coefficients associated with different scales and directions. The combination of curvelet transform with Curvelet coefficients was selected as inputs of artificial neural networks. Random forest tree analysis was utilized to classify features into different classes that represent the muscular activities. The performance was tested by the clinical dataset. The experiment is carried out using clinical SEMG dataset presents in table 1 and the parameter values are considered correlational coefficient: 0.6494, Mean absolute error (MAE): 0.0595, Root mean squared error (RMS): 0.1203, Relative absolute error (RAE): 78.0216 %, Root relative squared error (RSE): 76.3248 %. These parameters are calculated; enhance the quality of class separability. A comparative study is made with aggressive and normal datasets as presented in table 2 and 3. Statistical results shows maximum classification accuracy of 99% and higher information transfer rate is achieved.

<table>
<thead>
<tr>
<th>Table1: Sample Clinical SEMG Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
</tr>
<tr>
<td>----</td>
</tr>
<tr>
<td>-0.0181</td>
</tr>
<tr>
<td>-0.0278</td>
</tr>
<tr>
<td>-0.0218</td>
</tr>
<tr>
<td>-0.0158</td>
</tr>
<tr>
<td>-0.0105</td>
</tr>
<tr>
<td>-0.0068</td>
</tr>
<tr>
<td>-0.0015</td>
</tr>
<tr>
<td>0.0007</td>
</tr>
<tr>
<td>0.0007</td>
</tr>
<tr>
<td>0.003</td>
</tr>
<tr>
<td>0.0007</td>
</tr>
<tr>
<td>-0.003</td>
</tr>
<tr>
<td>-0.0053</td>
</tr>
</tbody>
</table>
Table 2: Aggressive Datasets

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Dataset</th>
<th>Correlation coefficient</th>
<th>Mean absolute error</th>
<th>Root mean squared error</th>
<th>Relative absolute error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Elbowing</td>
<td>0.422</td>
<td>279.23</td>
<td>458.30</td>
<td>93.92 %</td>
</tr>
<tr>
<td>2</td>
<td>Front kicking</td>
<td>0.4106</td>
<td>315.21</td>
<td>498.87</td>
<td>92.14 %</td>
</tr>
<tr>
<td>3</td>
<td>Hamering</td>
<td>0.4488</td>
<td>744.66</td>
<td>1128.45</td>
<td>93.95 %</td>
</tr>
<tr>
<td>4</td>
<td>Headering</td>
<td>0.3245</td>
<td>538.37</td>
<td>872.69</td>
<td>96.67 %</td>
</tr>
<tr>
<td>5</td>
<td>Kneeing</td>
<td>0.5123</td>
<td>2745.05</td>
<td>3048.81</td>
<td>84.44 %</td>
</tr>
<tr>
<td>6</td>
<td>Pulling</td>
<td>0.4866</td>
<td>2443.17</td>
<td>2823.19</td>
<td>86.49 %</td>
</tr>
<tr>
<td>7</td>
<td>Punching</td>
<td>0.507</td>
<td>2715.8</td>
<td>3026.57</td>
<td>84.58 %</td>
</tr>
<tr>
<td>8</td>
<td>Pushing</td>
<td>0.4777</td>
<td>2333.74</td>
<td>2715.45</td>
<td>86.46 %</td>
</tr>
</tbody>
</table>

Table 3: Normal Datasets

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Dataset</th>
<th>Correlation coefficient</th>
<th>Mean absolute error</th>
<th>Root mean squared error</th>
<th>Relative absolute error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bowing</td>
<td>0.3319</td>
<td>24.99</td>
<td>33.65</td>
<td>95.28 %</td>
</tr>
<tr>
<td>2</td>
<td>Clapping</td>
<td>0.3043</td>
<td>55.70</td>
<td>89.01</td>
<td>92.11 %</td>
</tr>
<tr>
<td>3</td>
<td>Handshaking</td>
<td>0.1808</td>
<td>12.26</td>
<td>15.56</td>
<td>98.67 %</td>
</tr>
<tr>
<td>4</td>
<td>Hugging</td>
<td>0.2283</td>
<td>70.41</td>
<td>94.58</td>
<td>97.51 %</td>
</tr>
<tr>
<td>5</td>
<td>Jumping</td>
<td>0.4029</td>
<td>2994.50</td>
<td>3286.58</td>
<td>88.27 %</td>
</tr>
<tr>
<td>6</td>
<td>Running</td>
<td>0.5087</td>
<td>2283.92</td>
<td>2653.24</td>
<td>85.50 %</td>
</tr>
<tr>
<td>7</td>
<td>Seating</td>
<td>0.1842</td>
<td>11.93</td>
<td>14.94</td>
<td>99.72 %</td>
</tr>
<tr>
<td>8</td>
<td>Standing</td>
<td>0.0637</td>
<td>16.25</td>
<td>23.85</td>
<td>91.32 %</td>
</tr>
</tbody>
</table>
EMG signals are commonly recorded by inserting concentric pointer electrodes bottomless (0.45 mm diameter with a recording surface area 0.07 mm²; impedance at 20 Hz below 200 kΩ) within the muscle. Pointer EMG provides better selectivity with strong attenuates cross talk and can be used when goal of small muscles. Approximately 20 dissimilar MUPs were achieved from every muscle through five to seven muscle insertions. Among ant two muscles, the pointers were reserved for at least 5 mm. The position of the pointer close to active muscle processes was cultured by auditory and visual control of the EMG signal. Under isometric conditions, the EMG signal was stored at force levels approximately 30% of maximum voluntary contraction (MVC). The signal was acquired for 5 s, band pass filtered at 5–10 kHz, and sampled at 20 kHz with 12-bit A/D resolution. The EMG signal was low-pass filtered at 2 kHz. The following table 4 shows comparative study of clinical dataset, aggressive dataset and normal dataset. The classification accuracy rates are 99% for clinical, 98% for aggressive and 98.9 for normal data sets. It is clearly shows that proposed method achieved higher information transfer rate than traditional methods.

**Table 4: Comparative study of clinical dataset, aggressive dataset and normal dataset**

<table>
<thead>
<tr>
<th>Dataset / Method</th>
<th>Clinical dataset</th>
<th>Aggressive dataset</th>
<th>Normal Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet Transform</td>
<td>97%</td>
<td>97.2%</td>
<td>96.5%</td>
</tr>
<tr>
<td>PCA</td>
<td>96.5%</td>
<td>95.6%</td>
<td>97.5%</td>
</tr>
<tr>
<td>Cosine Transform</td>
<td>95%</td>
<td>98.5%</td>
<td>96.8%</td>
</tr>
<tr>
<td>Fourier Transform</td>
<td>98%</td>
<td>96.7%</td>
<td>98.7%</td>
</tr>
<tr>
<td>Curvelet Transform</td>
<td>99%</td>
<td>98%</td>
<td>98.9%</td>
</tr>
</tbody>
</table>
5. CONCLUSION

Electromyography (EMG) signal is most powerful signal processing tools for electrical activity of neuromuscular associated with a corresponding muscle. In this paper, the analysis of EMG signals using curvelet transform and Random forest tree is presented. The EMG signal including noise though dissimilar media. The curvelet transform is used for clear away noise from the surface electromyography and superior order of statistics is used for analyzing the signal. The first level is to evaluate the surface of EMG signal and extract features using curvelet transform. The second level is best EMG quality segment was chosen and the rebuilding of the useful data signal was finished using random forest classifier. The intention of this work is introducing a novel approach for discover, analyzing and classifying of EMG signals. The proposed method is applied using clinical dataset and the parameters like mean root mean square, correlation coefficient and absolute value are calculated and to get better quality of class separability. A comparison is made with other traditional methods and the EMG characteristics extracted from rebuilding of EMG signals provide the enhancement of class separability in feature space than. Statistical results shows maximum classification accuracy of 99% and higher information transfer rate is achieved. It is possible to reduce complexity analysis by using other classification techniques.

REFERENCES


