Surface EMG Signals Determinism Analysis Based on Recurrence Plot for Hand Grasps

Gaoxiang Ouyang, Zhaojie Ju and Honghai Liu, Senior Member, IEEE
Intelligent Systems & Biomedical Robotics Group
School of Creative Technologies, University of Portsmouth
England, PO1 2DJ, UK
honghai.liu@port.ac.uk

Abstract—This paper proposes determinism measure (DET) based on recurrence plot, which is capable of showing the recurrence property of a deterministic dynamical system, to evaluate the dynamical characteristics of the surface electromyogram (sEMG) during three different hand movements. In addition, the linear discriminant analysis (LDA) is applied to evaluate the performance of the above measures to identify these three hand grasp movements. The experimental result shows that the recognition rate, 96.7%, based on the combination of the linear and non-linear measures is much higher than that with only linear measures, and DET might be a potential tool to reveal the sEMG hidden characteristics of hand grasp movements.

Keywords—Determinism; hand grasp; recurrence plot (RP); surface electromyogram (sEMG)

I. INTRODUCTION

The surface electromyogram (sEMG) is a measure of the summed activity of a number of motor unit action potentials (MUAP) lying in the vicinity of the recording electrode. It may provide insight into the neural activation and dynamics of muscles [1], and thus the analysis of sEMG signals is particularly attractive in that it provides relatively easy access to the physiological processes that allow the muscles to generate force and movements [2]. One potentially important application using sEMG signals lies in controlling prosthetic devices [3] for instance hand prostheses [4], due to the fact that sEMG signals contain the electrical activities of arm muscle contraction that undergo many complex transitions in different hand movements [5]. However, there are inherent difficulties in deriving a general model on the relationship between the recorded sEMG and hand gestures when humans perform contractions [6]. In order to control a prosthetic device, the fundamental challenge is to efficiently handle sEMG signals and identify the intention of users [7]. Meanwhile, the exploration of inherent features of sEMG signals could be of benefit in understanding the mechanisms of neuromuscular systems [8], and therefore there is a growing interest in tackling this challenge [9].

Various methods have been proposed to extract valuable information from sEMG signals, in which a series of traditional methods such as time and frequency analysis has been widely utilized [10][11]. In the time domain, EMG amplitude analysis mainly consists of the root mean squared (RMS) value, absolute mean value (AMV) and difference absolute mean value [10]. On the other hand, the dominant methods in the frequency domain are comprised of the mean and median frequency [11][12]. To some extent, these traditional methods of EMG amplitude and spectral analysis have the capability to track muscular changes. However, they are not effective in analyzing complex transient signals [13]. In particular, the spectral analysis is based on an assumption that observed variations of electrical field of muscle activities are time-invariant (stationary) processes [12]. Therefore, these traditional methods fail to detect the critical feature of sEMG data during transient human movements since this assumption is inconsistent with sEMG dynamics.

Nonlinear methods derived from the theory of chaos theory recently have been proposed to analyse the sEMG signals [12], which can extract informative features from sEMG data to detect the changes in different muscle status [14]. However, chaos-based approaches can only be employed to analyse sEMG data in constrained conditions since they assume that a signal is stationary and originates from a low-dimensional nonlinear system [15][16]. Nevertheless, the sEMG signal – a complex signal embedded in noise, is non-stationary and stems from a high-dimensional nonlinear system. Therefore, it is promising to develop new nonlinear algorithms to characterize sEMG changes in different neural activation of muscles. Moreover, the introduction of new sEMG analysis techniques and tools has made it possible to extract substantial meaningful information from the signals of the neuromuscular system [12].

It is evident that nonlinear methods have inspired the research in sEMG signal analysis in depth. For instance, a recurrence quantification analysis (RQA) [17] was employed to analyse the nonlinear dynamical characteristics of sEMG data, where one of its key features is the ability to describe nonlinear nature of short and non-stationary signal corrupted with noise [18]. RQA methods have been broadly applied to the analysis of physiological data, such as EEG data [19], fMRI data [20], heart signals [21], and EMG data [22]. A number of experimental studies showed that the RQA methods have a high potential for detecting changes in sEMG due to fatigue [23] or clinical pathology [24]. In this paper, we aim to investigate whether or not the determinism measures of sEMG data can effectively extract hidden information in
II. MATERIALS AND METHODS

Eight (2 female, 6 male) healthy right-handed subjects volunteered for this study. Their ages range from 23 to 40 and average is 32.5 years. All participants gave informed consent prior to the experiments according to the University of Portsmouth CCI Faculty Ethics Committee.

A. Experimental Procedure

The experiment consisted of both freely and different grasp gestures. Each type of grasp was repeated 10 times. Every motion lasted about 2 seconds. Between every two repetitions, participants had to relax their hand for 2 seconds in the intermediate state which is opening hand naturally without any muscle contraction. Once one motion with ten repetitions was finished, participants had to relax hand for 2 minutes before the next motion started. This was designed to overcome the effects of muscle fatigue.

B. Data Collection

The four EMG electrodes were applied to the subject’s right forearm muscles, i.e. flexor carpi radialis (channel 1), flexor carpi ulnaris (channel 2), flexor pollicis longus (channel 3) and flexor digitorum profundus (channel 4). The sEMG data were recorded using DataLINK system from Biometrics LTD with a gel-skin contact area of about 4 cm²/2 for each bipolar electrode and a centre to centre recording distance of 20 mm. The sampling frequency of DataLINK system in our experiment was set to be 1000Hz and sEMG signals were amplified 1000 times and bandwidth is 20 to 460 Hz using a sEMG amplifier (SX230FW sEMG Amplifier, Biometrics LTD). To obtain good-quality signals, subjects were scrubbed with alcohol and shaved if necessary and then electrodes were applied over the body using the die cut medical grade double sided adhesive tape. Electrodes locations were selected according to the Musculoskeletal of these four muscles and confirmed by muscle specific contractions, which include manually resisted finger flexion, extension and abduction. The captured sEMG signals were visualized on a computer screen giving participants feedback to choose the positions of electrodes with stronger sEMG signals.

To investigate the dynamical characteristics of sEMG data during different grasp gestures, EMG signals were selected and dissected from no action (dataset I), grasp and lift a can full of rice using five fingers with the thumb abduction (dataset II) and hold and lift a dumbbell (dataset III) intervals. In this study, 80 4-channel 1-sec EMG epochs were selected for each dataset. Short (1-sec) EMG signals were used since the duration of the grasp gesture is only about a few seconds.

C. Phase Space Reconstruction

EMG data is non-stationary time series. The first step in the analysis of a signal using non-linear dynamics theory is the reconstruction of the phase space trajectory of the signal [25]. The methods of delay [25] are usually used to embed a scalar time series into a m-dimensional space, it is

\[ x_k = (u_k, u_{k+\tau}, \ldots, u_{k+(m-1)\tau}) \]  

where \( k = 1, 2, \ldots, L - (m - 1)\tau \), \( \tau \) is the delay time and \( m \) is the embedding dimension, \( m \geq 2 \). The difficulty of this method is the choice of the delay time, \( \tau \), and the minimum embedding dimension, \( m \).

The most common method for choosing a proper time delay is based on detection of the first local minimum of the mutual information (MI) function [26], since the first minimum of the \( MI(\tau) \) portrays the time delay where the signal \( u_{i+\tau} \) adds maximal information to the knowledge we have from \( u_i \) [26]. As for the choice of the embedding dimension, Cao proposed a robust and efficient method that determines the minimum embedding dimension; it can overcome some shortcomings of false nearest neighbors [27]. In this study, the MI method [26] and Cao’s method [27] are employed to estimate the delay time and the embedding dimension, respectively.

D. Recurrence Plot and Determinism Measure

Recurrence plots (RP), proposed by Eckmann et al. [28] can describe the recurrence property of a deterministic dynamical system, i.e. visualizing the time dependent behavior of orbits \( x_i \) in a phase space. The key step of RP is to calculate the following \( N \times N \) matrix

\[ R_{i,j}(\epsilon) = \Theta(\epsilon - \|x_i - x_j\|), \quad i, j = 1, \ldots, N \]  

where \( N = L - (m - 1)\tau \), \( \epsilon \) is a predefined cutoff distance, \( \| \cdot \| \) is the norm (e.g. the Euclidean norm) and \( \Theta(x) \) is a Heaviside function. The phase space vector \( x_i \) can be reconstructed using Takens’ time delay method, \( x_i = (u_i, u_{i+\tau}, \ldots, u_{i+(m-1)\tau}) \) [52], based on the observations \( u_i \).

The cutoff distance \( \epsilon \) defines a sphere centered at \( x_j \), if \( x_i \) falls within this sphere, i.e. the state is close to \( x_j \), then \( R_{i,j} = 1 \); otherwise \( R_{i,j} = 0 \). The binary values of \( R_{i,j} \) can be simply visualized with the colors black (1) and white (0). Thereby, RP can be considered as a visual inspection of a high dimensional phase space trajectory; in other words, RP indicates the time evolution of a trajectory [28].

In order to further investigate the properties of RP, several measures of complexity that quantify the small-scale structures in RP called recurrence quantification analysis (RQA), have been proposed [29]. These measures are based on the recurrence point density and the diagonal line structures of the RP.

The classical measure of RQA is the recurrence rate (RR)

\[ RR(\epsilon) = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}(\epsilon) \]
which simply counts the black dots in the RP. \( RR \) is a measure of the density of recurrence points, and the value

\[
N_n(\varepsilon) = \frac{1}{N} \sum_{i,j=1}^{N} R_{i,j}(\varepsilon)
\]

is the simple average number of neighbours that each point on the trajectory has in its \( \varepsilon \) – neighbourhood.

The determinism measure is based on the diagonal lines. The frequency distribution of the lengths \( l \) of the diagonal structures in the RP is \( P^\varepsilon(l) = \{ l_i; i = 1, 2, \ldots, N \} \). Processes with stochastic behavior cause none or very short diagonals, whereas deterministic processes cause longer diagonals and less single, isolated recurrence points. Therefore, the ratio of recurrence points on the diagonal structures (of at least length \( l_{\text{min}} \)) to all recurrence points is called the \( DET \), and is introduced as a determinism (or predictability) measure of the system. Its formulation is given as below:

\[
DET = \frac{\sum_{l_{\text{min}}}^{N} IP^\varepsilon(l)}{\sum_{l=1}^{N} R_{i,j}(\varepsilon)}
\]  

where \( l_{\text{min}} \) is the threshold, which excludes the diagonal lines formed by the tangential motion of a phase space trajectory, which in this study we fixed at \( l_{\text{min}} = 2 \).

A parameter specific to the RP is the cutoff distance \( \varepsilon \). If it is too large, almost every point is a neighbour of every other point, which produces a saturation of the RP including irrelevant points; on the contrary, if it is too small, there may be almost no recurrence points, which loses information of the underlying system [19]. Several criteria for the choice of the cutoff distance \( \varepsilon \) have been advocated in the literature [30]. One approach uses a fixed number of neighbours, \( N_n \), for every point of the trajectory, called fixed amount of nearest neighbours (FAN) [29]). In this approach, the cutoff distance \( \varepsilon_i \) changes for each state \( x_i \) to ensure all columns of the RP have the same recurrence density. Using this neighbourhood criterion, \( \varepsilon_{i} \) can be adjusted in such a way that the recurrence rate has a fixed predetermined value (i.e. \( RR = N_n/N \)) [19]; and the RP is invariant under enlarging or reducing of the amplitude of time series. In this paper, the \( DET \) is computed for all 1-sec EMG recordings (the number of data \( L = 1000 \)) with a fixed number of neighbours \( N_n = 50 \). Using this \( N_n \) value, the recurrence rate \( RR \) is about 0.05.

\[ E. \quad \text{Linear Analysis Methods} \]

In order to compare the extraction information of EMG between linear and nonlinear methods, two well known linear methods, root mean square (RMS) and median frequency (MF), are also used to analyse the sEMG data. Given a signal \( u_1, u_2, \ldots, u_L \), the RMS value of the signal is calculated as the root of the mean of the square of all samples in a signal:

\[ \text{RMS} = \sqrt{\frac{1}{L} \sum_{i=1}^{L} u_i^2} \]

The power spectrum of signal \( u(t) \) is calculated using the fast discrete Fourier transform. Then, the MF is defined as the frequency that divides the power density spectrum in two regions having the same amount of power [14].

\[ \text{III. \quad RESULTS} \]

\[ \text{A. \quad Choice of Time Delay and Embedding Dimension} \]

The first step in RQA is the reconstruction of the \( m \)-dimensional phase space trajectory. The delay time \( \tau \) and the embedding dimension \( m \) of the EMG epochs are determined by using the MI method and the Cao’s method. In order to determine proper time delay, the MIs as a function of time delay \( \tau \) are calculated on each EMG segment. The optimum values of \( \tau \), based on the detection of the first local minimum of the MI function, ranged from 2 to 8 samples (mean and standard deviation is 3.36 \( \pm \) 1.66) for different EMG epochs. So the optimum delay time, \( \tau = 4 \), is selected for the phase space reconstruction of the three EMG data sets. In order to determine the proper embedding dimension, Cao’s method is applied to all EMG epochs. The optimum embedding dimension \( m \) ranged from 6 to 12 (mean and standard deviation are 8.81 \( \pm \) 0.98) for different EMG segments. Therefore, \( m = 9 \) is suitable for the topologically proper reconstruction of the EMG data.

\[ \text{B. \quad Determinism Measures of sEMG Data} \]

The determinism measure \( DET \) is applied to analyse all 960 1-sec sEMG epochs in this study (80 x 4-channel from each dataset I, II and III). For example, the \( DET \) values for the sEMG epochs from flexor carpi radialis averaged \( 0.372 \pm 0.107 \), \( 0.412 \pm 0.093 \) and \( 0.588 \pm 0.066 \) (mean \( \pm \) SD) in dataset I, II and III, respectively. The \( DET \) values of the no action are lower than those in the grasp and lift a can and hold and lift a dumbbell, and hold and lift a dumbbell had the highest \( DET \) values. Statistical analysis of the \( DET \) values in each dataset from different forearm muscles is carried out to determine whether their distributions over the three groups were significantly different. A one-way ANOVA was performed using standard tool of numerical analysis (Matlab’s anova1 function, statistics toolbox) [31]. The results are presented in Fig. 1. The lower and upper lines of the “box” are the 25th and 75th percentiles of the sample, the distance between the top and bottom of the box is the inter quartile range and the line in the middle of the box is the sample median. Outliers (plus sign) are cases with values that are more than 1.5 times the interquartile range. For all four forearm muscles, it can be seen that the \( DET \) values of sEMG data from no action are lower than those of grasp and lift a can and hold and lift a dumbbell.
To test these observed mean differences statistically, the one-way ANOVA test is performed for DET values of three different sets. In order to further investigate the EMG activities of different forearm muscles during hand grasp, statistical analysis of the DET values in each dataset are carried out for flexor carpi radialis, flexor carpi ulnaris, flexor pollicis longus and flexor digitorum profundus, respectively. The F value is 117.8, 115.5, 72.0 and 121.5 for flexor carpi radialis, flexor carpi ulnaris, flexor pollicis longus and flexor digitorum profundus, respectively. The results suggest that the determinism of sEMG data in no action, grasp and lift a can and hold and lift a dumbbell are significantly different.

Thus, a Scheffe's post-hoc test (Matlab’s multcompare function, statistics toolbox) [31] for all pairwise comparisons between the means is shown in Fig. 2. We can use it for calculating a critical difference, to allow us to determine whether a difference can be considered to be statistically significantly different, or not. For example, Fig. 2(A) shows the results of multiple comparisons analysis of sEMG data from flexor carpi radialis; the critical difference is 0.037. Thus, in this case, the mean difference must exceed 0.037 to be considered statistically significant, at the 5% level. The DET values for the sEMG epochs from flexor carpi radialis averaged 0.372, 0.412 and 0.588 in dataset I, II and III, respectively. Therefore, the results suggest that the DET values for the sEMG data from flexor carpi radialis during grasp a can and hold a dumbbell each have significantly higher values than the no action, and grasp a can significantly differs from hold a dumbbell. Similarly, the critical differences for multiple comparisons analysis for flexor carpi ulnaris, flexor pollicis longus and flexor digitorum profundus are 0.038, 0.040 and 0.038 as shown in Fig. 2(B), (C) and (D), respectively. It can be seen that the difference between different groups is statistically significant, except between group II and III for flexor pollicis longus.

C. Linear Analysis of sEMG Data

The linear techniques RMS and MF are applied to analyse all 960 1-sec sEMG epochs in this study (80 x 4-channel from each dataset I, II and III). The results are shown in Fig. 3 and Fig. 4, respectively. It is found that: the RMS values of the no action are lower than those in the grasp and lift a can and hold and lift a dumbbell, and sEMG data in dumbbell group has the highest RMS values; and almost of MF values of the no action are higher than those in the grasp and lift a can and hold and lift a dumbbell, and sEMG data in dumbbell group has the lowest MF values.
Figure 4. ANOVA results of MF values of sEMG from flexor carpi radialis (A), flexor carpi ulnaris (B), flexor pollicis longus (C) and flexor digitorum profundus (D). Boxplots for the DET values of the all sEMG epochs, grouped by no action (I), grasp and lift a can (II) and hold and lift a dumbbell (III).

D. Classification

The performance of the above measures to discriminate among groups is also evaluated by means of a linear discriminant analysis (LDA) [32]. First, the ability of the linear techniques in classifying different hand grasp is evaluated using a LDA algorithm. The calculated RMS and MF measures are used as input data with 8 features (dimension of the extracted feature vectors – each 4 RMS and MF values for muscles) in the LDA classifier. These features are projected down to a two dimensional space as shown in Fig. 5A. It is shown that there is considerable overlap between the features in the no action and grasp a can, so the traditional linear techniques can not distinguish well among different hand grasp movements. The classification results is illustrated in Table I. Of 240 hand movements in three groups, 215 are classified correctly. The total classification accuracy is 89.6%.

Table I. Classification Results with RMS and MF

<table>
<thead>
<tr>
<th>Desired result</th>
<th>Output result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free</td>
<td>Cup</td>
</tr>
<tr>
<td>Free</td>
<td>72</td>
</tr>
<tr>
<td>Cup</td>
<td>12</td>
</tr>
<tr>
<td>Dumbbell</td>
<td>3</td>
</tr>
</tbody>
</table>

Next, in order to investigate whether or not the DET measures can effectively extract additional hidden information from sEMG data, the calculated DET, RMS and MF measures are used as input data with 12 features (dimension of the extracted feature vectors – each 4 DET, RMS and MF values for muscles) in the LDA classifier. These features are projected down to a two dimensional space as shown in Fig. 5B. It can be seen that the data separate into well-defined clusters, which is able to correctly classify 232 out of 240 subjects, as illustrated in Table II, giving approximately 96.7% separability.

Table II. Classification Results with DET, RMS and MF

<table>
<thead>
<tr>
<th>Desired result</th>
<th>Output result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free</td>
<td>Cup</td>
</tr>
<tr>
<td>Free</td>
<td>78</td>
</tr>
<tr>
<td>Cup</td>
<td>3</td>
</tr>
<tr>
<td>Dumbbell</td>
<td>0</td>
</tr>
</tbody>
</table>

IV. Conclusion

In this study, we have analysed non-linear deterministic structure in sEMG data during different hand grasp states using the RQA measure, DET. The novelty of this method is that it can deal with a linear and non-linear time series to quantify the activity of a system irrespective of the number or dynamical nature of the individual sources. It is found that there is a significant increase of the determinism of the sEMG data from no action state to hand grasp states.

Furthermore, our results showed that there is a significant difference in the determinism of the sEMG data from forearm
muscles between grasp a can and hold a dumbbell. Our results demonstrated that the DET values for the sEMG data from flexor carpi radialis and flexor carpi ulnaris have significantly higher values during hold a dumbbell than during grasp a can. This finding is consistent with the fact that the contraction of flexor carpi radialis and flexor carpi ulnaris during hold a dumbbell are stronger than those during grasp a can.

Finally, the LDA algorithm was applied to evaluate the performance of linear and nonlinear measures to discriminate among three hand grasp movements. A total classification accuracy of 89.6% is achieved based on the traditional linear techniques: RMS and MF. However, a higher classification accuracy of 96.7% is obtained when using the DET, RMS and MF measures. It can be seen that the determinism measure, DET, of sEMG data can effectively extract additional hidden information; and this information can not be extracted by traditional linear techniques.

REFERENCES


