Identification of Surface EMG Signals Using Wavelet Packet Entropy

Xiao Hu[1], Xiaomei Ren[2]
1, Information, Machinery and Electronics College, Guangzhou University, Guangzhou 510006, China
2, Department of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

Abstract: - This paper introduces a novel and simple algorithm to extract the feature from Surface EMG signals recorded from the skin surface over forearm muscles. Surface EMG signal is decomposed into 16 frequency bands (FB) by wavelet packet transform (WPT), and then wavelet packet entropy (WPE) of every surface EMG signal is calculated by its relative wavelet energy in every FB. WPE is regarded as the feature to distinguish forearm supination (FS) surface EMG signals from forearm pronation (FP) surface EMG signals. The results show that WPE is an effective method to extract the feature from surface EMG signal.

Key-Words: -Surface EMG signal; Wavelet packet transform; Wavelet packet entropy; Bayes decision

1 Introduction

The surface EMG signal recorded from the skin surface over limb muscles in the process of limb movement is called action surface EMG (ASEMG) signal. Due to its non-invasive and easy measurement, ASEMG signal has more and more been used as the raw signal from which some features are extracted to control the limb prostheses. Englehart et al [1,2] have taken such methods as wavelet transform and wavelet packet transform to extract the features from ASEMG signal. Using hidden markov models, Chan et al [3] identified ASEMG signals, and Hu et al [4] used fractal dimension to represent different ASEMG signal patterns. Although the methods can perform well, there is still an obvious motivation to explore some more effective algorithms which extract the features from shorter ASEMG signal and to reduce the error identification rate. Surface EMG signal is constituted by many individual motor unit (MU) action potential (MUAP) trains, generated by the active MUs in muscle, and background noise. If a limb takes different limb actions, the spectral energy distribution will take on different properties, because the active MUs which are stimulated during different limb actions are not identical. The spectral analysis is therefore an effective means to explore the clinical diagnostic information [5,6] or the control information for limb prostheses [1,2] from surface EMG signal. Based on the spectral energy distribution, this paper introduces a novel and effective algorithm to extract the features from ASEMG signal.

2 Methods

2.1 ASEMG Signal Acquisition

All ASEMG signals are recorded from right forearm flexor of 30 healthy subjects in the EMG room at Hua Shan Hospital in Shanghai, China. The sampling frequency f_s is 1000Hz. It is about 2cm between two measuring electrodes which are put on the skin surface over the pronator teres in the right forearm along the flexor, and the ground electrode is on the flexor carpi radialis and the palmaris longus.

During the acquiring process, every subject is instructed to do two different kinds of limb actions: FS and FP. Two sets of ASEMG signals are recorded from every subject’s forearm flexor: one is for FS and the other is for FP. Then, from the start time of forearm actions, one segment of ASEMG signal lasting for 0.5s is segmented from every ASEMG signal. Thus among the 60 segments of ASEMG signals obtained in all, there are two ASEMG signal patterns: FS ASEMG signal and FP ASEMG signal, 30 sets for each pattern.

2.2 Wavelet Packet Entropy

When a signal $s(t)$ is decomposed to the fourth resolution level (j=-4) with wavelet packet transform, the whole signal with frequencies in the interval $[0,2^{-1}f_s]$ is divided into 16 FBs with frequencies correspondingly in the interval $[(n-1)2^{j-1}f_s, n2^{j-1}f_s]$, n=1,2,…,16. The sub-signal at the nth FB on the jth level can be reconstructed by

$$s_j^n(t) = \sum_k D_k^{j,n} \psi_{j,k}(t) \qquad k \in Z \tag{1}$$

Here, $D_k^{j,n}$ is the wavelet packet coefficients at the nth FB on the jth level and $\psi_{j,k}(t)$ is the wavelet function. Since the wavelet $\{\psi_{j,k}(t)\}$ is an orthogonal basis at $L^2(\Re)$, the energy of the sub-signal $s_j^n(t)$ is calculated by

$$E_n = \sum_k |D_k^{j,n}|^2 \tag{2}$$

The total energy of $s(t)$is

$$E = \sum_n E_n \tag{3}$$

In consequence, the relative energy (RE) at the nth FB is

$$RE_n = E_n / E \tag{4}$$

RE quantifies the probability distribution of the spectral energy of signal $s(t)$. Obviously, the RE set $\{RE_1, RE_2, \ldots, RE_{16}\}$, covering the whole frequency band [0-500Hz], describes the spectral energy distribution of ASEMG signal and includes some feature information. However, it could not quantitatively describe the feature information. In information theory, the Shannon entropy can provide a measure of the information of any probability distribution. The wavelet packet entropy (WPE) [7] is defined to measure the spectral energy distribution of signal $s(t)$.

$$WPE(RE) = -\sum_n RE_n \cdot \ln[RE_n] \tag{5}$$

It is well known that the information carried by the coefficients of wavelet packet transform depends on the joint characterristics of the analyzed signal and the selected wavelet function; the more similar are the two functions, the less spread the significant coefficients in the time-scale plane. Because Daubechies family of wavelet packets seems to resemble MUAPs most [8] and the simplest of these wavelets is Daubechies 2 (db2), db2 is adopted as the mother wavelet.

3 Results

3.1 Spectral Energy Distribution

Estimating REs in 16 FBs from 0.5s-long ASEMG signal and then analyzing the spectral energy distribution which is instituted by the REs, we encouragingly find some striking results. No matter that ASEMG signal is FS ASEMG signal or FP ASEMG signal, its spectral energy in 0-250Hz is above 95% of its total spectral energy in 0-500Hz. Fig.1 (a) and (b) show respectively the average REs of FS ASEMG signals and the average REs of FP ASEMG signals estimated by the following formula

$$P_n = \tfrac{1}{30} \sum_{i=1}^{30} RE_n^i \tag{6}$$

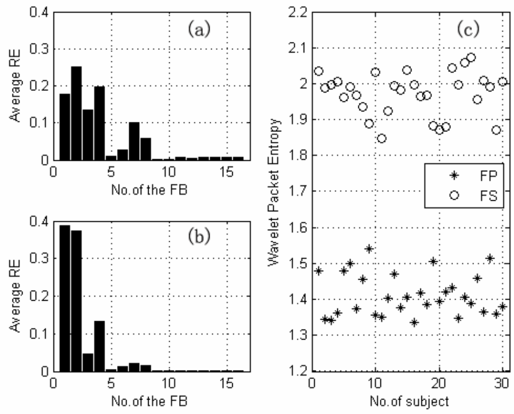

Fig. 1. Average RE and WPE's distribution
(a) Average RE of FS ASEMG signal;
(b) Average RE of FP ASEMG signal;
(c) WPE's distribution. Asterisk (*) and circle (○), respectively, symbolize FP and FS signals.

The average REs set characterizes the general characteristics of the spectral energy distribution about different patterns of ASEMG signals. From Fig.1, it is easily found that FS ASEMG signals and FP ASEMG signal have different spectral energy distribution. In the 1th and 2th FB, the spectral energy of FS ASEMG signal is smaller than that of FP ASEMG signal. On the contrary, in higher FBs, FS ASEMG signal has more spectral energy than FP ASEMG signal. In the other word, the spectral energy of FS ASEMG signal is more dispersive than FP ASEMG signal. Fig.1 (c) shows the WPE distribution of FS and FP ASEMG signals (each for 30). Obviously, WPEs of all FS ASEMG signals are bigger than those of all FP ASEMG signal. Visually, all FS ASEMG signals can be distinguished from all FP ASEMG signals by their WPEs. Therefore, WPE is undoubtedly an alternative and appropriate measurement to capture the distinction between the spectral energy distribution of FP ASEMG signal pattern and that of FS ASEMG signal pattern.

3.2 The Error Identification Rate Based on Bayes Decision

Let ω_1 and ω_2 be the two classes (FS and FP ASEMG signal patterns) to which our patterns belong. Feature vector x represents an unknown pattern. The Bayes rule is

$$P(\omega_i / x) = \frac{p(x/\omega_i)P(\omega_i)}{\sum_{i=1}^{2} p(x/\omega_i)P(\omega_i)} \qquad (7)$$

Here, $p(\omega_i/x)$ is the ith conditional probability and $P(\omega_i)$ is priori probability. In this paper, $P(\omega_1)= P(\omega_2)=0.5$. $p(x/\omega_i)$ is the class-conditional probability density function. One of the most commonly encountered probability density functions in practice is the normal density function. The major reasons for its popularity are its computational tractability and the fact that it models adequately a large number of cases.

The Bayes decision can be stated as: if $p(\omega_1/x)> p(\omega_2/x)$, x is classified to ω_1; otherwise, x is classified to ω_2. In practice, the decision errors are unavoidable. If the curve of $p(x/\omega_1)$ is assumed to be left to the curve of $p(x/\omega_2)$ on x-axis, the total probability, P_e , of committing a decision error is given by

$$P_e = \tfrac{1}{2}\int_{-\infty}^{x_0} p(x/\omega_2)dx + \tfrac{1}{2}\int_{x_0}^{\infty} p(x/\omega_1)dx \qquad (8)$$

x_0 is the crossing point of $p(x/\omega_1)$ and $p(x/\omega_2)$. The error decision rate is calculated by

$$R_e = P_e \times 100\% \qquad (9)$$

When the sampling point of signal changes from 200 points (0.2s) to 500 points (0.5s), the WPEs are calculated from signals with different sampling points. And then the WPEs are regarded as the features to calculate the error decision rate according to (7) and (8).

According to Englehart's work [1], the features which are obtained by the combination of wavelet packet transform and principal components analysis can get the smallest error identification rate under a linear discriminant analysis. The features are called as WPT features in this paper.

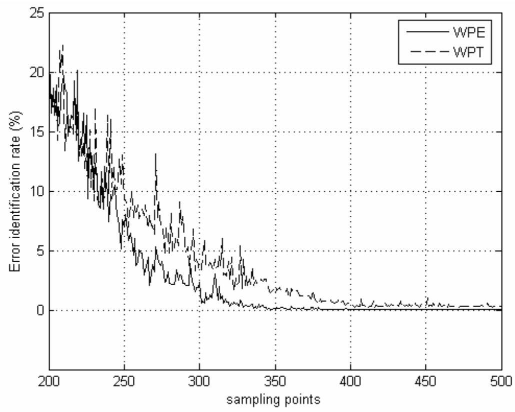

Fig. 2. The Error Decision Rate Based on Bayes Decision VS. the sampling points of signal.

Like the WPE features, the error identification rate of the WPT features using a linear discriminant analysis is computed with the increasing of signal's sampling points.

Fig.2 depicts the error identification rate vs. signal sampling points. The curves on Fig.2 are the real error decision rate to the sampling points. With the increasing of the sampling points, the signal obviously includes more and more feature information, so it is undoubted that the error decision decreases with the increasing of the sampling points. However, from Fig.2, we find another result that the WPE features perform better than the WPT features. No matter how long the signal is, the error decision by the WPE features is lower than the WPT features. Furthermore, when the sampling points is above 350, the error decision by the WPE features is almost 0.

4 Discussion and conclusions

In the flexor of forearm, there are many muscles (e.g. Brachioradialis, pronator teres, flexor carpi radialis, flexor digitorum superficialis, etc) taking charge of or assisting one kind of forearm actions. Major MUAPs of ASEMG signal, which is analyzed in this study, obviously come from the muscles. However, the muscles responsible for FP and FS are only pronators (pronator teres and pronator quadratus) and supinators. When a subject wants to take FP or FS, the muscles will actively alter their contraction conditions and other muscles will not change or change slightly. As a result, in all MUAPs in FB (0-500Hz), the percentage of MUAPs in some FBs must change greatly [9]. The change can be successfully captured by RE_n. Spectral energy distributions represent different patterns of ASEMG signals. From the results in this paper, WPE can precisely measure the spectral energy distribution.

In pattern recognition, the feature set insisted of some features is known as feature vector. The number of the features in the feature vector is called as dimensionality. In general cases, whether one pattern of signal can be identified effectively and accurately depends much upon two important factors. One is a set of optimal features. A set of desired features should contain general information which characterizes one pattern of signals and ignore special information which only exists in some particular signals or sometime might be the result of noisy measurements. From the results in this paper, both WPE feature and WPT feature can capture the general information of ASEMG signal. The other is dimensionality reduction. There is more than one reason for the necessity to reduce the number of features to a sufficient minimum. Computational complexity is the obvious and important one. The computational complexity for pattern recognition is reduced by dealing with the features in a lower dimensional space. In this paper, WPE feature vector only need one feature, but in WPT feature vector, we used 8 features.

In summary, wavelet packet entropy is an more effective method to extract feature from ASEMG signal than WPT feature. FS ASEMG signals can be successfully distinguished from FP ASEMG signals by the WPE features.

ACKNOWLEGMENTS

This work was supported in part by the National Basic Research Program (973) of China (No.2005CB724303)

References:

[1] K. Englehart, B. Hudgins, and P.A. Parker, A wavelet-based continuous classification scheme for multifunction myoelectric control, *IEEE Trans. Biomed. Eng.*, Vol. 48, No.3, 2001. pp.302-311.

[2] K. Englehart and B. Hudgins, A robust, real-time control scheme for multi-function myoelectric control, *IEEE Trans. Biomed. Eng.*, Vol.50, No.7, 2003, pp.848-854.

[3] A. Chan and K. Englehart, Continuous myoelectric control for powered prostheses using hidden markov models, *IEEE Trans. Biomed. Eng.*, Vol.52,No.1, 2005,pp.121-124.

[4] Hu Xiao,Wang Zhizhong and Ren Xiaomei, Classification of surface EMG signal with fractal dimension, *J Zhejiang Univ SCI.* Vol.6B,No.8, 2005, pp.844-848.

[5] S. Karlsson, J. Yu, and M. Akay, Enhancement of spectral analysis of myoelectric signals during static contractions using wavelet methods, *IEEE Trans. Biomed. Eng.*, Vol.46,No.6, 1999, pp.670-684.

[6] S. Karlsson, J. Yu, and M. Akay, Time-frequency analysis of myoelectric signals during dynamic contractions: a comparative study, *IEEE Trans. Biomed. Eng.*, Vol.47, No.2, 2000,pp.228-238.

[7] S.Blanco, A. Figliola, R.Quiroga, O.A. Rosso and E.Serrano, Time-frequency analysis of electroencephalogram series. III. Wavelet packets and information cost function, Vol.57, 1998,pp.932 -940.

[8] M. Flanders, Choosing a wavelet for single-trial EMG, *Journal of Neuroscience Methods*, Vol.116, 2002, pp.165-177.

[9] Xiao Hu and Zhizhong Wang, Detecting the motor unit action potential from surface EMG signals based on wavelet transform, *2004 IEEE International workshop on Biomedical circuits & system*, 2004, S2.6-15.