Biomedical Signal Processing and Control 31 (2017) 265–271

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Measuring signal fluctuations in gait rhythm time series of patients with Parkinson's disease using entropy parameters

Yunfeng Wu[a,b,*], Pinnan Chen[a], Xin Luo[a], Meihong Wu[a], Lifang Liao[a], Shanshan Yang[a], Rangaraj M. Rangayyan[c]

[a] School of Information Science and Technology, Xiamen University, Xiamen, Fujian, China
[b] Fujian Key Laboratory of Sensing and Computing for Smart City Xiamen University, 422 Si Ming South Road, Xiamen, Fujian 361005, China
[c] Department of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada

ARTICLE INFO

Article history:
Received 5 August 2015
Received in revised form 21 August 2016
Accepted 22 August 2016

Keywords:
Approximate entropy
Gait analysis
Generalized linear regression analysis
Parkinson's disease
Signal turns count
Stride time
Symbolic entropy

ABSTRACT

Gait rhythm disturbances due to abnormal strides indicate the degenerative mobility regulation of motor neurons affected by Parkinson's disease (PD). The aim of this work is to compute the approximate entropy (ApEn), normalized symbolic entropy (NSE), and signal turns count (STC) parameters for the measurements of stride fluctuations in PD. Generalized linear regression analysis (GLRA) and support vector machine (SVM) techniques were employed to implement nonlinear gait pattern classifications. The classification performance was evaluated in terms of overall accuracy, sensitivity, specificity, precision, Matthews correlation coefficient (MCC), and area under the receiver operating characteristic (ROC) curve. Our experimental results indicated that the ApEn, NSE, and STC parameters computed from the stride series of PD patients were all significantly larger (Wilcoxon rank-sum test: $p < 0.01$) than those of healthy control subjects. Based on the distinct features of ApEn, NSE, and STC, the SVM provided an accuracy rate of 84.48% and MCC of 0.7107, which are better than those of the GLRA (accuracy: 82.76%, MCC: 0.6552). The SVM and GLRA methods were able to distinguish PD gait patterns from healthy control cases with area of 0.9049 (SVM sensitivity: 0.7241, specificity: 0.9655) and 0.9037 (GLRA sensitivity: 0.8276, specificity: 0.8276) under the ROC curve, respectively, which are better or comparable with the classification results achieved by the other popular pattern classification methods.

1. Introduction

Parkinson's disease (PD) is a hypokinetic neurological disease due to apoptosis of dopaminergic cells in the substantia nigra [1]. The chronic deterioration of dopaminergic cells in the cerebrum decreases neural interactions such that neuronal signals are not properly transmitted from one neuron to another. The effects of PD may involve cognitive disorders, such as dementia, depression, disturbances in rapid eye movement sleep, visual difficulty, and dysphonia [2]. Degeneration of the central nervous system also leads to motor dysfunction, with the manifest symptoms in terms of noticeable tremor at 4–6 Hz, bradykinesia, postural instability, and rigidity [1]. Freezing of gait and impaired balance increase risks of falling for PD patients [3]. Lower limb stiffness, slow movement, small shuffling strides, and other apparent gait disturbances can be observed in PD patients with mild and moderate motor impairments [4].

Wearable sensors and portable mobility measurement systems are useful in ambulatory posture monitoring and gait assessment for PD patients [5–9]. Rigas et al. [10] and Salarian et al. [11] measured body movement activities by using a group of accelerometers to detect and quantify tremor and bradykinesia in PD. Patel et al. [6] set up a body sensor network with wearable sensors to fuse accelerometer data to estimate the severity of motor symptoms and other PD complications. Moore et al. [12] and Bachlin et al. [13] computed postural and kinematic features associated with freezing of gait from the acceleration signals recorded by inertial measurement systems. Su et al. [14] measured the asymmetry of frequency sub-band components of the ground reaction force time series to detect pathological gait patterns in Parkinson's disease. Corbier et al. [15] used autoregressive moving average models with reduced order using the Huberian approach to characterize the stochastic process of gait rhythm signals of patients with PD and Huntington's disease. Predominant gait features may provide meaningful information to assess the pathological condition [16–18] and evaluate

* Corresponding author at: School of Information Science and Technology, Xiamen University, Xiamen, Fujian, China.
E-mail address: y.wu@ieee.org (Y. Wu).

http://dx.doi.org/10.1016/j.bspc.2016.08.022
1746-8094/

the effects of anti-PD medicine intervention or physical therapy [19]. Hausdorff [20] emphasized the importance of fluctuations in stride process for movement disorder analysis.

In recent years, fractal analysis and statistical methods have been effectively applied to study the gait variability in neurological diseases [21–23]. Xia et al. [24] calculated the Lempel-Ziv complexity, fuzzy entropy, and Teager-Kaiser energy features to characterize the different gait patterns of patients with PD, amyotrophic lateral sclerosis, and Huntington's disease. Daliri [25] used the short-time Fourier transform to analyze the spectrum of gait signals, and computed the chi-square distances between the histograms of frequency variances for gait pattern classification. Ertugrul et al. [26] proposed the shifted 1D local binary patterns to characterize local disturbances in gait signals. Wahid et al. [27] quantified several spatial-temporal gait features for 23 PD patients and 26 age-matched healthy controls, and then compared the classification results of random forest, support vector machine (SVM), and kernel Fisher discriminant analysis methods. Lee and Lim [28] computed the wavelet transforms, and used a neural network with weighted fuzzy membership functions to distinguish pathological gait patterns in PD. Classification of gait patterns may assist neurologists to effectively identify and analyze the abnormal biomarkers related to movement disorders [29]. The motivation of this work is to compute the approximate entropy (ApEn), normalized symbolic entropy (NSE), and signal turns count (STC) features to measure the intrinsic irregularities as gait disturbance indicators in PD. The primary hypothesis is to test whether the gait rhythm irregularities in PD represented by these three features are significantly greater than those of age-matched healthy subjects. Based on the entropy and signal variability features, the Parkinsonian gait patterns can be effectively distinguished by different nonlinear classifiers.

2. Gait data set

The gait data used in the present study were provided by Yogev et al. [30], and can be accessed via the PhysioNet website [31]. 29 patients with idiopathic PD (20 males and nine females, age mean ± standard deviation (SD): 71.1 ± 8.1 years, body mass: 73.8 ± 15.7 kg, height: 169 ± 11 cm) and 29 age-matched healthy control (CO) subjects (16 males and 13 females, age: 71.9 ± 6.5 years, body mass: 73 ± 12.3 kg, height: 167 ± 8 cm) were recruited from the Tel-Aviv Sourasky Medical Center, Israel [30]. The neurological impairment stages of the PD patients ranged from 2 to 3 on the Hoehn and Yahr (H&Y) scale [32], which were confirmed by neurological examinations. The severity of PD was also quantified by the Unified Parkinson's Disease Rating Scale (UPDRS) [33]. The mean ± SD of the H&Y scale and UPDRS were 2.34 ± 0.4 and 32.9 ± 12.3 for the PD patients, respectively. The healthy control subjects were recruited from the local community in Tel-Aviv, Israel. All of the participants were able to ambulate independently, without a mobility-assistive device, and did not suffer from any other pathological condition, such as cardiovascular disease, respiratory disease, musculoskeletal disease, or other neurological disease. The subjects were requested to provide written informed consent. The experimental protocol and subject consent documents were approved by the Human Studies Committee of Tel-Aviv Sourasky Medical Center [30]. Data analysis methodology documents were reviewed and approved by the Ethics Committee of Xiamen University.

According to the experimental protocol of Yogev et al. [30], the subjects were asked to walk at their comfortable pace along a straight path on level ground for 2 min. The raw gait data were quantified by a force-sensitive system that contains a pair of shoes and a portable data acquisition module [34]. Each shoe contained eight ultrathin load sensors which measured vertical forces underneath the foot, with a sampling rate of 100 Hz and 12 bits per sample of quantizing resolution. The data acquisition module (dimensions: 19 × 14 × 4.5 cm; weight: 1.5 kg) was worn on the waist [34]. The gait cycle time series were processed using the algorithm proposed by Hausdorff et al. [35], which determines gait parameters such as stride time, stance time, and swing time. Although the stride time of both feet were recorded, we only considered the right-foot stride time for statistical analysis in the present study. To eliminate the start-up and ending effects of walking postures, which were somewhat different from the normal walking patterns, we excluded the first four strides (start-up from standing to initial walking) and the last four strides (from normal walking to ending the last stride) in the raw time series. A median filter [36,4] was applied to detect and remove the stride outliers, the amplitudes of which were 3 SD larger than the median value of the entire stride time series.

3. Gait rhythm analysis

3.1. Approximate entropy

The ApEn parameter, proposed by Pincus [37], is a statistical approach that measures the irregularity and subtle fluctuations in a physiological process [38]. For a stride-time series of N samples, $\{s(i)\}$, the ApEn model is expressed as ApEn(m, r, N), where the positive integer $m \in \mathbb{Z}^+$ denotes a window length for similarity comparison and the positive real number $r \in \Re^+$ is the tolerance parameter for accepting similarity matches [39]. Let us define a sequence of vectors, $\{\mathbf{x}^1(i), \mathbf{x}^2(i), \ldots, \mathbf{x}^{(N-m+1)}(i)\}$, in which each vector $\mathbf{x}^m(i) = [s(i), s(i+1), \ldots, s(i+m-1)]$ is composed of m consecutive data samples of the time series. The distance $d[\mathbf{x}^m(i), \mathbf{x}^m(j)]$ is defined as the maximum absolute difference between the corresponding elements from the vectors $\mathbf{x}^m(i)$ and $\mathbf{x}^m(j)$, respectively [40], i.e.,

$$d[\mathbf{x}^m(i), \mathbf{x}^m(j)] = \max_{k=1,2,\ldots,m} |s(i+k-1) - s(j+k-1)|. \quad (1)$$

For each i, $1 \leqslant i \leqslant N-m+1$, let $C_i^m(r)$ quantify the probability of j satisfying the condition that the distance between $\mathbf{x}^m(i)$ and the template $\mathbf{x}^m(j)$ is smaller than r, i.e.,

$$C_i^m(r) = \frac{\text{number of } j \text{ such that } d[\mathbf{x}^m(i), \mathbf{x}^m(j)] < r}{N-m+1}. \quad (2)$$

The value of $C_i^m(r)$ represents the frequency of similarity matches within a tolerance r between the vector $\mathbf{x}^m(i)$ of window length m and a given template $\mathbf{x}^m(j)$. Define the function $\phi^m(r)$ as the average of the natural logarithms of $C_i^m(r)$ [41]:

$$\phi^m(r) = \frac{\sum_{i=1}^{N-m+1} \ln C_i^m(r)}{N-m+1}. \quad (3)$$

Then, the approximate entropy ApEn(m, r) is given by [39]

$$\text{ApEn}(m, r) = \lim_{N\to\infty} [\phi^m(r) - \phi^{m+1}(r)]. \quad (4)$$

For a finite-length time series, the approximate entropy ApEn(m, r, N) is expressed as

$$\begin{aligned}\text{ApEn}(m, r, N) &= [\phi^m(r) - \phi^{m+1}(r)] \\ &= \frac{\sum_{i=1}^{N-m+1} \ln C_i^m(r)}{N-m+1} - \frac{\sum_{i=1}^{N-m} \ln C_i^{m+1}(r)}{N-m}. \end{aligned} \quad (5)$$

When N is much larger than m, that is $N \gg m$, the value of ApEn(m, r, N) can be approximately estimated as

$$\text{ApEn}(m, r, N) \approx \frac{\sum_{i=1}^{N-m} \ln \left[C_i^m(r)/C_i^{m+1}(r)\right]}{N-m}. \quad (6)$$