Biomedical Signal Processing and Control 18 (2015) 254–262

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Classification of gait rhythm signals between patients with neuro-degenerative diseases and normal subjects: Experiments with statistical features and different classification models

Yi Xia [a], Qingwei Gao [a,*], Qiang Ye [b]

[a] School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China
[b] Information Technology Research Centre, Nanjing Sport Institute, Nanjing 210014, China

ARTICLE INFO

Article history:
Received 15 August 2014
Received in revised form 14 January 2015
Accepted 2 February 2015

Keywords:
Gait classification
Neuro-degenerative diseases
Statistical features
Feature selection

ABSTRACT

For the purpose of realizing an intelligent and highly accurate diagnosis system for neuro-degenerative diseases (NDD), such as amyotrophic lateral sclerosis (ALS), Parkinson's disease (PD) and Huntington's disease (HD), the present study investigated the classification capability of different gait statistical features extracted from gait rhythm signals. Nine statistical measures, including several seldom-used variability measures for these signals, were calculated for each time series. Next, after an evaluation of four popular machine learning methods, the optimal feature subset was generated with a hill-climbing feature selection method. Experiments were performed on a data set with 16 healthy control (CO) subjects, 13 ALS, 15 PD and 20 HD patients. When evaluated with the leave-one-out cross-validation (LOOCV) method, the highest accuracy rate for discriminating between groups of NDD patients and healthy control subjects was 96.83%. The best classification accuracy (100%) was obtained with two subtype binary classifiers (PD vs. CO and HD vs. CO).

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

From a neurophysiological control viewpoint, neuro-degenerative diseases are characterized as the progressive loss of structure or function of neurons, including the death of neurons. Typical neurodegenerative diseases include: amyotrophic lateral sclerosis (ALS), Parkinson's disease (PD) and Huntington's disease (HD). Different neurodegenerative diseases may have different clinical symptoms. PD manifests itself as bradykinesia, rigidity, a resting tremor and posture instability. Primary clinical features for HD are chorea and cognitive and personality changes, while the most evident symptom for ALS is the degeneration and atrophy of the muscles. Meanwhile, as a result of an impairment of the neural control network, the common external characteristic for NDD is the disorder of movement, which has been investigated and confirmed by several studies.

Hausdorff et al. [1] studied the magnitude of the stride-to-stride fluctuations and perturbations of the gait rhythm in subjects with ALS disease compared to healthy control subjects. Their study revealed longer stride intervals for ALS patients vs. normal gait. In another similar study, by using the detrended fluctuation analysis technique, Hausdorff et al. [2] found that stride-interval fluctuations for patients with Huntington's disease are more random than healthy control subjects. Liao et al. [3] investigated gait symmetry in ALS patients using the multi-resolution entropy analysis of stance time fluctuations. Their study verified that gait rhythm in ALS patients exhibits a prominent asymmetry. As for PD patients' gait dynamics, Morris et al. [4] demonstrated that PD had greatly influenced the subjects' gait by reducing the speed, stride length, and total range of movement during walking. These studies demonstrated that for NDD patients, stride interval or stride-to-stride variability gait dynamics manifested important differences when compared with control subjects.

Currently, the most commonly used method to diagnose and evaluate an improvement in NDD patients is still dependent on all types of questionnaires. It is well known that the use of questionnaires may lead to subjective results. Therefore, an objective assessment of patients' physical functional performance is of great importance to current clinical practice, and with the help of it, clinical doctors tend to develop a more reasonable treatment plan and a more scientific assessment of the outcome of therapy. Quantitative gait analysis as a non-invasive method for the detection of movement disorders has captured considerable attention in previous studies. In recent years with the rapid development of the machine

* Corresponding author. Tel.: +86 13856922457.
E-mail address: qingweigao@ahu.edu.cn (Q. Gao).

http://dx.doi.org/10.1016/j.bspc.2015.02.002
1746-8094/© 2015 Elsevier Ltd. All rights reserved.

learning technique, various feature extraction methods and classification models have been employed to realize an automated and accurate diagnosis system for clinical assistant purposes.

Duta et al. [5] employed Elman's recurrent neural network for the automatic identification of healthy and pathological gait based on stance, swing and double support intervals of gait. The relevant features in their work were from cross-correlograms of those gait signals with corresponding signals of a reference subject. Daliri [6] extracted features such as the minimum, maximum, average and standard deviation (SD) of each gait stride sequence and then fed this information to a support vector classifier for the purpose of distinguishing the healthy group from the group of subjects with NDD. Wu and Krishnan [7] quantified the gait variability in patients with PD using the signal turns count method. They found that the swing interval turns count parameter presents a significant difference between the healthy control subjects and PD patients. In another study [8], the nonparametric Parzen-Window method was utilized to estimate the probability density functions (PDFs) of gait in terms of the stride interval and then by computing the mean of the left-foot stride interval and the modified Kullback–Leibler divergence from the estimated PDFs, they realized the classification of gait in subjects with ALS and normal subjects.

In addition to considerable work toward an automatic gait analysis, recent decades also saw the rapid development of gait recognition, which has similar procedures as gait classification. Moustakidis et al. [9] applied wavelet packet (WP) decomposition to ground reaction force measurements during walking, and then a compact set of powerful and complementary features from WP coefficients was obtained with a fuzzy complementary criterion. By using a single camera, Zeng and Wang [10] extracted features from a series of subject silhouettes and then realized gait recognition under the framework of the deterministic learning theory. To realize multi-view gait recognition, Hu [11] first applied a set of dual-tree complex wavelet transforms (DTCWT) with different scales and orientations to gait energy images and then proposed a two-stage Gaussian mixture model to represent these DTCWT coefficients. The view angle of a probe gait sequence was then estimated by using a simple nearest-neighbor classifier.

The present study focuses on an automatic gait classification for NDD patients by applying gait rhythm signals because it is easy to collect with different sensors [12,13]. Previous studies have demonstrated the use of statistical and nonlinear computational methods for feature extraction from gait rhythm signals and some nonlinear classifiers for automatic diagnosis of pathological gait. However, for a practical expert system to be widely used in clinical practice, accuracy is a dominant factor. When the public gait dataset used in this study is evaluated with the LOOCV method, the best accuracy reported is only 90.32% for PD vs. CO and 92.3% for ALS vs. CO, which is obtained by using an SVM classifier [7,8]. To obtain more performance, a better feature set and a more suitable classifier are needed. Difficulty comes from the following two basic facts: (1) gait cadence exhibits complex and nonlinear behavior in both NDD and CO subjects because of the nonlinear dynamics of the human system [14,15]; (2) due to the impaired movement ability of NDD patients, it is hard to design an experimental protocol to sample long time series gait rhythm signals, for example, only 5 min of walking for each time series is available in the present study. The purpose here is to investigate the classification performance of different statistical features of gait rhythm signals under different nonlinear classifiers. In addition to the previously calculated statistical features for gait rhythm signals, several other features, which were seldom applied before such as complexity, fuzzy entropy and Teager energy, were also extracted for classification purposes. Then, for different gait features, their classification performance was compared with the most suitable classifier, and an optimal feature set was generated by a hill-climbing feature selection algorithm. With this optimized feature set, several different diagnosis expert systems for NDD patients were implemented with high accuracy. In addition to the classification tasks between abnormal and normal gaits as investigated in most previous studies, the present study also introduced a multi-class solver to differentiate different subtypes of abnormal gaits, which could be helpful in monitoring altered gait dynamics for different NDD patients. This paper has been organized as follows: an overview of different feature extraction algorithms and classification models is given in Section 2. The proposed method is evaluated by actual gait data in Section 3. Finally, the summary and conclusions are presented in Section 4.

2. Methodologies

2.1. Gait dataset

The gait dataset used in this study was contributed by Hausdorff et al. [16]. The database contains 64 recordings of gait from four different groups: 20 subjects with HD, 13 subjects with ALS, 15 subjects with PD and 16 healthy control subjects. After some checking, it was found that the data with the 20th HD subject has some problems. Therefore, it was removed in this study. According to the experimental protocol, the subjects were asked to walk at their normal pace along a straight hallway that was 77 m in length for 300 s. The gait signals were measured with ultrathin force-sensitive switches placed inside each subject's shoes. The signals were recorded by an on-board analog-to-digital converter at a 300-Hz sampling rate and with a 12-bit resolution per sample. Different gait cadence parameters were derived with the stride detection algorithm proposed by Hausdorff et al. [12]. Currently, the database provides the following seven gait parameters: left and right stride interval (time from initial contact of one foot to the immediate subsequent contact), left and right stance interval (amount of time when one foot is on the ground), left and right swing interval (amount of time when one foot is in the air), and double support interval (time of bilateral foot contact). The first 20 s of each record were removed to eliminate the start-up effects, and a median filter was applied to remove data points that were three SDs greater than or less than the median value. These outliers were most likely caused by the turns at the end of the hallway. All of the pre-processing steps were suggested in the previous study by Hausdorff et al. [16].

Heights and weights of the control and NDD subjects were not significantly different. The subjects with NDD were not using a wheelchair for mobility and were free of other ailments that might affect the gait. The presence or absence of symptoms that might affect lower extremity weakness was determined by a qualified physician. The severity of ALS, PD and HD is represented by the corresponding questionnaire score, which is often used in clinical practice. Some basic information for the different categories of subjects has been summarized in Table 1. Additionally, some stride-to-stride plots of gait cadence parameters for a healthy subject and three NDD patients are shown in Fig. 1.

2.2. Feature extraction

Because the stride interval is the sum of the swing and stance interval among the seven gait parameters provided in the gait dataset, five independent gait time series are chosen for feature extraction. These selected gait time series are the left and right stride interval, left and right stance interval and double support interval. For each gait time series, features describing the major statistical characteristics of its sample distribution were extracted as: (1) MEAN (mean value), (2) STD (standard deviation), (3) MAX

Download English Version:

https://daneshyari.com/en/article/6951405

Download Persian Version:

https://daneshyari.com/article/6951405

Daneshyari.com