Medical Engineering & Physics 33 (2011) 347–355

Contents lists available at ScienceDirect

Medical Engineering & Physics

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

Analysis of altered gait cycle duration in amyotrophic lateral sclerosis based on nonparametric probability density function estimation☆

Yunfeng Wu[a,*], Lei Shi[b]

[a] Department of Communication Engineering, School of Information Science and Technology, Xiamen University, 422 Si Ming South Road, Xiamen, Fujian 361005, China
[b] Department of Orthopedics, Zhongshan Hospital Xiamen University, 201 Hubin South Road, Xiamen, Fujian 361004, China

ARTICLE INFO

Article history:
Received 17 December 2009
Received in revised form 31 August 2010
Accepted 29 October 2010

Keywords:
Gait analysis
Amyotrophic lateral sclerosis
Movement disorders
Probability density function
Parzen window
Support vector machine

ABSTRACT

Human locomotion is regulated by the central nervous system (CNS). The neurophysiological changes in the CNS due to amyotrophic lateral sclerosis (ALS) may cause altered gait cycle duration (stride interval) or other gait rhythm. This article used a statistical method to analyze the altered stride interval in patients with ALS. We first estimated the probability density functions (PDFs) of stride interval from the outlier-processed gait rhythm time series, by using the nonparametric Parzen-window approach. Based on the PDFs estimated, the mean of the left-foot stride interval and the modified Kullback–Leibler divergence (MKLD) can be computed to serve as dominant features. In the classification experiments, the least squares support vector machine (LS-SVM) with Gaussian kernels was applied to distinguish the stride patterns in ALS patients. According to the results obtained with the stride interval time series recorded from 16 healthy control subjects and 13 patients with ALS, the key findings of the present study are summarized as follows. (1) It is observed that the mean of stride interval computed based on the PDF for the left foot is correlated with that for the right foot in patients with ALS. (2) The MKLD parameter of the gait in ALS is significantly different from that in healthy controls. (3) The diagnostic performance of the nonlinear LS-SVM, evaluated by the leave-one-out cross-validation method, is superior to that obtained by the linear discriminant analysis. The LS-SVM can effectively separate the stride patterns between the groups of healthy controls and ALS patients with an overall accurate rate of 82.8% and an area of 0.869 under the receiver operating characteristic curve.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a type of neurodegenerative disorder that affects the motor neurons in the central nervous system, which directly or indirectly control muscular contractions during ambulation [1]. Deterioration of motor neurons or their myelin sheath will result in neurological dysfunction, with which the impulses initiated by the cerebrum cannot be transmitted to the target muscle fibers [2]. During the degenerative process, those weak motor neurons will be progressively replaced with fibrous astrocytes, and the muscle tissue begins to atrophy, which would lead to lower extremity weakness, and ultimately motor paralysis [3,4].

Due to the interruption of the pathway from the cerebrum to the muscle, the lower limbs may not properly perform voluntary movements so that the gait in a patient with ALS could be altered [5]. Analysis of gait parameters is very useful for a better understanding of the mechanisms of movement disorders, and also has high potential in monitoring of the progression of a particular neurodegenerative disease [6,7].

In recent related studies [8–15], computer-assisted tools have been utilized to measure the gait parameters in healthy adults, and also to describe the distinct characteristics of the gait in neurodegenerative diseases. The research based on the fractal analysis indicated that the temporal dynamics of stride interval in healthy adults possess a fractal (self-similar) structure in small time scales [8,9], and the fluctuation magnitude is relatively small in healthy subjects [14]. Hausdorff et al. [10] quantified the locomotor instability in ALS by measuring the fluctuation magnitude (in terms of the coefficient of variation) in the gait rhythm time series. Their results suggested that the gait speed is apparently slower in patients with ALS, and the fluctuation magnitude of the strides in ALS is larger than that in healthy adults. By analyzing the gait rhythm time series with the detrend fluctuation analysis method, Hausdorff et al. [10] also observed that the stride-to-stride fluctuation dynamics are perturbed in ALS, but the changes of the gait variability (characterized by the fractal scaling index and nonstationary index) in ALS do not reach the level of significant difference.

☆ A preliminary version of this paper was presented at the 32nd International Annual Conference of IEEE EMBS, Buenos Aires, Argentina, September, 2010.

* Corresponding author. Tel.: +86 13959276699; fax: +86 592 5095058.
E-mail address: y.wu@ieee.org (Y. Wu).

1350-4533/$ – see front matter © 2010 IPEM. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.medengphy.2010.10.023

Liao et al. [12] used the multi-resolution entropy approach to analyze the fluctuations of gait rhythm in ALS, Parkinson's disease (PD), and Huntington's disease (HD). Their study demonstrated that the gait rhythm in ALS exhibits a higher degree of asymmetry than in PD or HD. Scafetta et al. [13] used the supercentral pattern generator model to simulate human gait dynamics, and also discussed the stochastic and fractal properties of the gait in ALS, PD, and HD. Aziz and Arif [15] converted the stride time series into a kind of symbol sequence, and then applied a threshold dependent symbolic entropy method in the analysis of gait complexity. They observed that the normalized corrected Shannon entropy of the symbolic stride sequences is much lower in ALS at different short thresholds.

Although a few nonlinear analysis methods have been used in previous studies to investigate the complexity of human locomotion process, further quantitative studies call for more computational tools that are suited for gait analysis [16]. The present study used the statistical methods and pattern classifiers to analyze the altered stride interval in patients with ALS [17]. In order to characterize the dominant statistical parameters of the gait, we first applied the nonparametric Parzen-window method to estimate the probability density functions (PDFs) of stride interval for healthy subjects and ALS patients, respectively. Based upon the PDFs obtained, we computed two statistical parameters of stride interval, i.e., the mean of the left-foot stride interval and the modified Kullback–Leibler divergence. It is hypothesized that the statistical parameters of stride interval in ALS could be statistically different from those in healthy adults. The implementation of classification tools is also useful in gait analysis. With the decision boundaries provided by the classifiers, the different gait patterns may be automatically categorized in the feature space, which helps monitor the altered gait rhythm parameters in ALS at an early stage. We expected that the kernel-based nonlinear classifier employed in the present study could properly distinguish the ALS stride patterns with a higher degree of accuracy.

2. Methods

2.1. Gait assessment

2.1.1. Subjects

The gait database used in the present study was contributed by Hausdorff et al. [10], and can also be downloaded via the web page of PhysioNet[1] [18]. Hausdorff et al. [10] reported that 13 patients with ALS (age mean ± standard deviation, SD: 55.6 ± 12.8 years; 10 males and 3 females) were recruited from the Neurology Outpatient Clinic at Massachusetts General Hospital, Boston, MA, USA. These ALS patients were free of visual, respiratory, cardiovascular, or other neurological diseases, which might lead to lower extremity weakness. The pathological duration since the diagnosis of ALS ranged from 1 to 54 months (mean ± SD: 18.3 ± 17.8 months). The ALS patients did not usually use a wheelchair for mobility, and the medication usage was not changed. Sixteen healthy subjects (age mean ± SD: 39.3 ± 18.5 years; 2 males and 14 females) without gait disturbances were recruited to form the control (CO) group. These healthy subjects were free of neurological, respiratory, or cardiovascular disorders. Either height or weight of the CO subjects was not significantly different from that of the ALS patients. All the CO and ALS subjects provided informed consent as approved by the Institutional Review Board of the Massachusetts General Hospital.

[1] Online available at http://www.physionet.org/physiobank/database/gaitndd/.

2.1.2. Experimental protocol

According to the experimental protocol established by Hausdorff et al. [10], the subjects were instructed to walk at their normal pace along a straight 77-m long hallway for 5 min without stopping (unless they had to turn at the end of the hallway) on level ground. Ultrathin pressure-sensitive switches were placed in each subject's shoes to record the force underneath the foot during ambulation [19]. The temporal signals were digitized by an on-board analog-to-digital converter at the sampling rate of 300 Hz with 12-bit resolution per sample, and then stored in a recorder (dimensions: 5.5 cm × 2 cm × 9 cm; weight: 0.1 kg). The recorder was worn on the ankle cuff of each foot and held in place with a wallet on the ankle [19]. The present study focuses on the gait rhythm in terms of stride interval (time from initial contact of one foot to the subsequent contact of the same foot), which can be obtained with the algorithm proposed by Hausdorff et al. [19].

2.1.3. Preprocessing of stride interval time series

Before the statistical analysis, we removed the samples of stride interval recorded in the first 20 s to minimize the start-up effects, which was the same as implemented in the previous studies [10,12,13,16]. During the monitoring period of 5 min, each time the subjects reached the end of the hallway, they had to turn around and then continue walking [10]. The strides associated with these turning events (regarded as outliers) should be excluded for the present statistical analysis, because they were different from those recorded during walking in a straight line along the hallway. According to the well-known "three-sigma rule" [20], about 99.7% of the normally distributed probability values lies within 3-SD distance from the mean. For each subject, we removed the stride outliers that were 3 SDs in amplitude greater or less than the median value of stride interval over the entire time series for both feet [16]. In the stride interval preprocessing procedure, the median value was used instead of the mean of the corresponding time series, because some outliers possessed large values and could affect the computing of the mean over the time series. With the annotations of the outliers detected, the time series of right-foot stride interval were calibrated with the left-foot time series, for the purpose of probability-density-function-based comparative study. Illustrations of the detection of stride outliers for a healthy control and an ALS patient are provided in Fig. 1, and the results are presented in Section 3.

2.2. Gait analysis

2.2.1. Nonparametric probability density function estimation

With the outlier-processed time series of stride interval, we first computed the histogram, as a reference to the PDF to be estimated. For each subject, the histogram of stride interval was established with B bins, which helped calculate the probability of occurrence with B containers of equal length in the amplitude range of stride interval. According to Scott's choice [21], the optimal bin number B can be derived by minimizing the mean squared error between the estimated histogram and the Gaussian density function, i.e.,

$$B = \frac{(g^h - g^l)}{3.49\, s\, n^{-1/3}}, \tag{1}$$

where s and n denote the SD and the number of samples in the stride interval time series, respectively, and the highest and lowest values of the range of stride interval g are presented as g^h and g^l, respectively.

For each subject, the PDF of stride interval was estimated from each outlier-processed time series, by using the nonparametric Parzen-window method [22]. If a M-length stride interval time series, the samples of which are sorted in amplitude as $\{x_1, x_2, \cdots,$

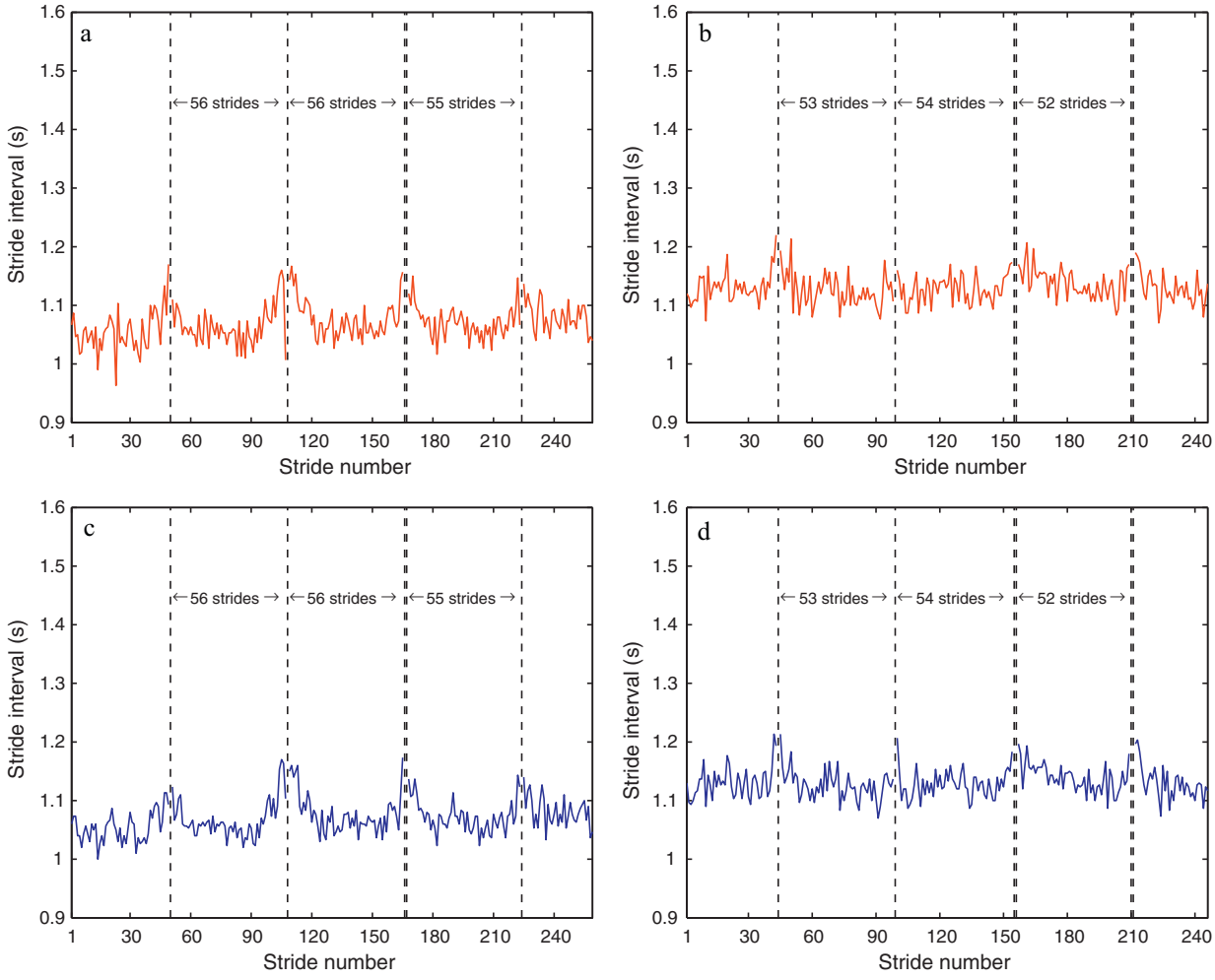

Fig. 1. Illustrations of outlier-processed stride interval time series recorded from (a) left foot and (c) right foot of a healthy control subject (age: 57, female); (b) left foot and (d) right foot of a patient with amyotrophic lateral sclerosis (age: 40, female, pathological duration: 14.5 months). *Dashed lines* locate the outlier strides which were 3 standard deviations greater or less than the median value of stride interval over the entire time series. The first strides in (a)–(d) each were counted from the corresponding time series after the start-up 20 s. During the monitoring period of 5 min (without the start-up 20 s), the numbers of strides produced by the healthy control subject and the patient with amyotrophic lateral sclerosis were 259 (5 outliers) and 246 (6 outliers), respectively. The height is 1.94 m for the healthy control subject and 1.7 m for the patient with amyotrophic lateral sclerosis, respectively.

$x_M\}$, then the underlying PDF $p(x)$ can be estimated as [23]

$$\hat{p}(x) = \frac{1}{M}\sum_{m=1}^{M} w(x - x_m), \tag{2}$$

where $w(\cdot)$ is a window function that integrates to unity. In the present study, we considered the Gaussian window, which may be expressed as

$$w(x - x_m) = \frac{1}{\sigma_P\sqrt{2\pi}} \exp\left[-\frac{(x - x_m)^2}{2\sigma_P^2}\right], \tag{3}$$

where σ_P represents the spread parameter that determines the width of a Gaussian window, the center of which is located at x_m. In order to determine the optimal spread parameter, the estimated PDF was allocated into the same B containers as computed for the histogram, so that the Parzen-window PDF can be presented as $\hat{p}(g_b)$, in which g_b, $b = 1, 2, \cdots, B$, indicates the range of stride interval g represented with B bins. The value of spread parameter was varied over the range from 0.01 to 0.1, with an increment of 0.01, and then the optimal spread parameter that minimizes the mean squared error between the Parzen-window PDF and the histogram can be selected in the range. In our experiment, the optimal values of σ_P were 0.02 and 0.03 for the CO and ALS subject groups, respectively. The results of the histogram and the nonparametric PDF estimation are presented in Section 3.

2.2.2. Extraction of statistical parameters

Based on the Parzen-window PDF of stride interval, $\hat{p}(g_b)$, we may compute the mean of stride interval as

$$\mu = \sum_{b=1}^{B} g_b \hat{p}(g_b), \tag{4}$$

where the mean values of stride interval computed for the left foot and right foot are presented as μ_L and μ_R, respectively. We did not compute the variance of stride interval or the coefficient of variation, because these two fluctuation magnitude parameters have been discussed in the study of Hausdorff et al. [10]. In probability theory, the Kullback–Leibler divergence (KLD) is a statistical measure of the difference between two PDFs [24]. However, such a relative-entropy-based parameter does not fit the metric criterion, due to its asymmetry. In order to describe the difference between the PDFs of stride interval estimated from the left foot and right

foot, we considered the modified KLD (MKLD) defined as

$$\begin{aligned}\mathrm{MKLD} &= D\left[\hat{p}_{\mathrm{L}}(g_b)|\hat{p}_{\mathrm{R}}(g_b)\right] + D\left[\hat{p}_{\mathrm{R}}(g_b)|\hat{p}_{\mathrm{L}}(g_b)\right] \\ &= \sum_{b=1}^{B}\hat{p}_{\mathrm{L}}(g_b)\log_e\left[\frac{\hat{p}_{\mathrm{L}}(g_b)}{\hat{p}_{\mathrm{R}}(g_b)}\right] + \sum_{b=1}^{B}\hat{p}_{\mathrm{R}}(g_b)\log_e\left[\frac{\hat{p}_{\mathrm{R}}(g_b)}{\hat{p}_{\mathrm{L}}(g_b)}\right], \end{aligned} \quad (5)$$

where $D\left[\hat{p}_{\mathrm{L}}(g_b)\,\middle|\,\hat{p}_{\mathrm{R}}(g_b)\right]$ denotes the KLD of the Parzen-window PDF of right-foot stride interval, $\hat{p}_{\mathrm{R}}(g_b)$, from the PDF of stride interval estimated for the left foot, $\hat{p}_{\mathrm{L}}(g_b)$. The logarithms in the definition of the MKLD are taken to base e (with respect to natural logarithms) so that the information is measured in nats (one nat corresponds to about 1.44 bits). The MKLD is nonnegative, i.e., $\mathrm{MKLD} \geq 0$. The minimum $\mathrm{MKLD} = 0$ occurs if and only if $\hat{p}_{\mathrm{L}}(g_b) = \hat{p}_{\mathrm{R}}(g_b)$, for all $b = 1, 2, \cdots, B$. The details of the feature analysis are presented in Section 3, and then discussed in Section 4.

2.2.3. Classification of stride patterns

In the experiment, we applied the least squares support vector machine (LS-SVM) with nonlinear kernels for classification of the stride patterns in ALS. The LS-SVM was proposed by Suykens et al. [25], with some modifications to the Vapnik SVM formulation [26]. Similar to the Vapnik SVM, the LS-SVM is also a type of kernel-based artificial neural network, the learning of which follows the structural risk minimization rule [27]. By choosing the nonlinear inner-product kernels, the LS-SVM is able to perform the same function as the polynomial learning machine (with polynomial kernels), the radial basis function network (with Gaussian kernels), or the multilayer perceptron with a single hidden layer (with sigmoid kernels) [28]. During the model parameter optimization, a subset of the training data is selected as the support vectors, which are considered to be the most informative for the classification task, because these support vectors are geometrically located close to the decision boundary.

Compared with the Vapnik SVM, the LS-SVM has an improvement of moderate complexity and works more efficiently, which is the reason that we considered the LS-SVM instead of the Vapnik SVM in the present study. The major differences between the Vapnik SVM and the LS-SVM are summarized as follows [25]:

- The optimization of the Vapnik SVM is obtained by solving a constrained quadratic programming (CQP) problem related to the convex cost function, whereas the learning of the LS-SVM is implemented by minimizing a regularized least squares cost function with equality constraints under the Karush–Kuhn–Tucker (KKT) condition.
- The solution vector of Lagrange multipliers obtained by the Vapnik SVM is sparse, i.e., only the Lagrange multipliers assigned to the representative training data points (support vectors) are with non-zero values, and the values of the rest Lagrange multipliers (for the remaining subset of training data) are equal to zero. For the LS-SVM, on the other hand, all the training data contribute to the model, as no Lagrange multiplier will be exactly equal to zero. Nevertheless, Suykens et al. [25] pointed out that, after optimizing the LS-SVM, the Lagrange multipliers assigned to the representative data points in the training set are with large absolute values, and the others for those less important training data have small absolute values (or close to zero). In other words, although the solution of Lagrange multipliers provided by the LS-SVM is not sparse, the support vectors can still be picked up from the optimum solution by defining a threshold level.
- The LS-SVM reformulates the convex cost function to be a regularized function in the squared sense, and the optimization of the LS-SVM is achieved by solving a set of linear KKT equations, instead of solving a CQP problem. Therefore, compared with the Vapnik SVM, the LS-SVM has a modest computation complexity which leads to higher performance.

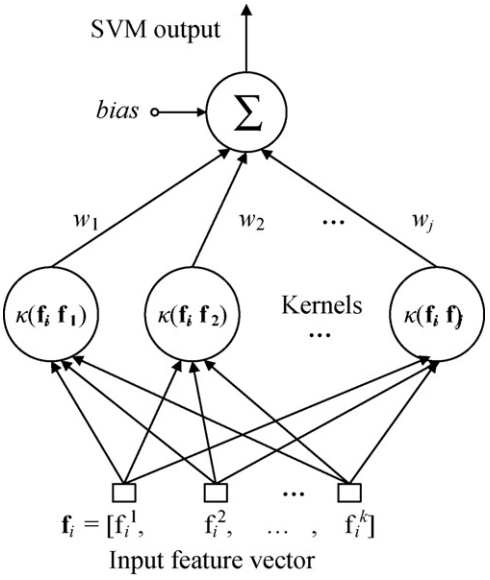

Fig. 2. Schematic representation of the least-squares support vector machine used for the gait pattern analysis. The input feature vector was constructed with the modified Kullback–Leibler divergence (MKLD) and the mean of the left-foot stride interval parameters. The inner kernels were Gaussian kernels with the optimal variance parameter equal to 0.25.

The schematic representation of the LS-SVM employed in our experiment is shown in Fig. 2, and the optimization process of the LS-SVM is summarized in Table 1. For more details of the theoretical framework of the LS-SVM, readers are referred to the book authored by Suykens et al. [25].

In our experiment, the classification results were characterized by the overall accuracy (i.e., the correct classification rate in percentage) and the diagnostic receiver operating characteristic (ROC) curve. The ROC curve and the area (A_z) under the curve were obtained by using the software ROCKIT[2] provided by the University of Chicago, Chicago, IL, USA [29]. We compared the classification accuracy obtained by the LS-SVM based on three types of nonlinear kernels (i.e., the polynomial, sigmoid, and Gaussian kernels), and chose the Gaussian kernels, together with the optimal variance parameter $\sigma_\kappa^2 = 0.25$ and the regularization parameter $\gamma = 2$ that supported the highest classification accuracy. The Gaussian kernel function $\kappa(\cdot)$ used in the experiment can be written as

$$\kappa\left(\mathbf{f}_i, \mathbf{f}_j\right) = \exp\left[\frac{-\left\|\mathbf{f}_i - \mathbf{f}_j\right\|^2}{\sigma_\kappa^2}\right], \quad (6)$$

where $\mathbf{f}_i$ denotes the input feature vector associated with the ith subject, and σ_κ^2 represents the variance parameter of the Gaussian kernel. In addition to the all-training–all-testing (ATAT) style, i.e., all available subjects for both the training and testing of the classifier, we also considered the leave-one-out (LOO) cross-validation method [28] to evaluate the generalization ability of the LS-SVM. For performance comparison purpose, the linear discriminant analysis (LDA) [28], a standard linear classifier used in a variety of clinical applications, was also used to implement a linear classification.

[2] Online available at http://www-radiology.uchicago.edu/krl/KRL_ROC/software_index6.htm.

Table 1
A optimization paradigm of the least squares support vector machine.

Training set: $\{\mathbf{f}_i, y_i\}_{i=1}^N$; features of the i th input pattern: $\mathbf{f}_i$; class label: $y_i \in \{-1, 1\}$ (negative or positive);
Classification: $y_i = y(\mathbf{f}_i) = \text{sign}[\mathbf{w}^T\varphi(\mathbf{f}_i) + c]$; high-dimensional feature space mapping: $\varphi(\cdot)$; bias: c;
Optimization problem:

$$\begin{cases} \min\limits_{\mathbf{w},c,e} J(\mathbf{w}, \mathbf{e}) = \dfrac{1}{2}\mathbf{w}^T\mathbf{w} + \gamma\sum\limits_{i=1}^{N} e_i^2 \\ \text{s.t.} \quad y_i[\mathbf{w}^T\varphi(\mathbf{f}_i) + c] = 1 - e_i, \ \forall_i, \end{cases}$$

where γ is the regularization parameter that controls the tradeoff between the complexity of the machine and the number of nonseparable data points, and e_i represent slack variables that measure the deviation of a data point from the ideal condition of pattern separability;
Lagrangian:

$L(\mathbf{w}, c, \mathbf{e}; \boldsymbol{\alpha}) = J(\mathbf{w}, \mathbf{e}) - \sum\limits_{i=1}^{N} \alpha_i \left\{ [\mathbf{w}^T\varphi(\mathbf{f}_i) + c] - 1 + e_i \right\}$; Lagrange multiples: $\alpha = [\alpha_1, \cdots, \alpha_N]^T$;

Conditions for optimality:

$$\begin{cases} \dfrac{\partial L}{\partial \mathbf{w}} = 0 \to \mathbf{w} = \sum\limits_{i=1}^{N} \alpha_i y_i \varphi(\mathbf{f}_i) \\ \dfrac{\partial L}{\partial c} = 0 \to \sum\limits_{i=1}^{N} \alpha_i y_i = 0 \\ \dfrac{\partial L}{\partial e_i} = 0 \to \alpha_i = \gamma e_i, \quad i = 1, \cdots, N \\ \dfrac{\partial L}{\partial \alpha_i} = 0 \to [\mathbf{w}^T\varphi(\mathbf{f}_i) + c] - 1 + e_i = 0, \quad i = 1, \cdots, N \end{cases}$$

The equivalent linear system:

$$\left[\begin{array}{ccc|c} \mathbf{I} & 0 & 0 & -\mathbf{Z}^T \\ 0 & 0 & 0 & -\mathbf{Y}^T \\ 0 & 0 & \gamma\mathbf{I} & -\mathbf{I} \\ \hline \mathbf{Z} & \mathbf{Y} & \mathbf{I} & 0 \end{array}\right] \left[\begin{array}{c} \mathbf{w} \\ c \\ \mathbf{e} \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ \hline \mathbf{I}_v \end{array}\right],$$

where $\mathbf{Z} = [\varphi(\mathbf{f}_1)^T y_1, \cdots, \varphi(\mathbf{f}_N)^T y_N]^T$, $\mathbf{Y} = [y_1, \cdots, y_N]^T$, $\mathbf{I}$ is a unit matrix, $\mathbf{I}_v = [1, \cdots, 1]^T$, $\mathbf{e} = [e_1, \cdots, e_N]^T$, $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_N]^T$;
Elimination of $\mathbf{w}$ and $\mathbf{e}$ gives:

$$\left[\begin{array}{c|c} 0 & \mathbf{Y}^T \\ \hline \mathbf{Y} & \mathbf{Z}^T\mathbf{Z} + \mathbf{I}/\gamma \end{array}\right] \left[\begin{array}{c} c \\ \hline \alpha \end{array}\right] = \left[\begin{array}{c} 0 \\ \hline \mathbf{I}_v \end{array}\right];$$

The kernel trick is applied to the matrix $\boldsymbol{\Omega} = \mathbf{Z}^T\mathbf{Z}$ with $\boldsymbol{\Omega}_{ij} = y_i y_j \varphi(\mathbf{f}_i)^T\varphi(\mathbf{f}_j) = y_i y_j \kappa(\mathbf{f}_i, \mathbf{f}_j)$, $\quad i, j = 1, \cdots, N$;
The corresponding LS-SVM classification in the dual space:
$y_i = y(\mathbf{f}_i) = \text{sign}\left[\sum_{j=1}^{N} \alpha_j y_j \kappa(\mathbf{f}_i, \mathbf{f}_j) + c\right]$; α_j are also used to determine the support vectors;

3. Results

Fig. 1 illustrates the detections of stride outliers present in the left-foot and right-foot time series of a 57-year-old healthy CO subject and of a 40-year-old ALS patient, respectively. Based on the outlier annotations, three unabridged 77-m walking distances can be segmented (see the *dashed lines* in Fig. 1). According to the stride numbers for these walking distances shown in Fig. 1, the averaged stride length of these two subjects can be estimated and normalized by their height, i.e., $(3 \times 77/(56 + 56 + 55) \times 1.94) = 0.71$ m for the CO subject (height: 1.94 m) and $(3 \times 77/(53 + 54 + 52) \times 1.7) = 0.85$ m for the ALS patient (height: 1.7 m), respectively. The ALS patient did not modulate as many strides as the CO subject practiced in the same walking distance (3×77 m), which indicated that the averaged stride length was larger in ALS, as confirmed in the study of Hausdorff et al. [10]. From Fig. 1, for the same subject (either the CO subject or the ALS patient), the number of strides taken within a segmented walking distance (equal to the length of the hallway) was similar to those taken within the other two segments. It can be inferred that the averaged stride length taken by a subject from one segmented walking distance to another was almost identical.

Fig. 3 plots the histogram graphics and the Parzen-window PDFs of stride interval for a 57-year-old CO subject and a 65-year-old ALS patient, respectively. The stride interval PDFs of the ALS patient exhibited a number of distinct characteristics, compared with those of the CO subject in the same scale: (1) the mean of stride interval was significantly increased in either left-foot or right-foot time series of the ALS patient; (2) the spread of either the left-foot or right-foot PDFs of the ALS patient was much wider than that of the CO subject; (3) the difference between the left-foot and right-foot PDFs of stride interval was apparent in the ALS patient.

As described in Section 2.2.2, the mean values of stride interval were computed for the left-foot and right-foot time series (labeled as μ_L and μ_R), respectively. We found that μ_L and μ_R were highly

Table 2
Mean and standard deviation (SD) values of two statistical parameters computed for 16 healthy control (CO) subjects and 13 patients with amyotrophic lateral sclerosis (ALS). Student's t-test was used to evaluate the significance of separability (p-value) for each statistical parameter.

Statistical parameters	Mean ± SD		p-Values [a]
	CO subjects	ALS patients	CO vs. ALS
Mean of left-foot stride interval (s)	1.093 ± 0.091	1.354 ± 0.196	<0.001
Modified Kullback–Leibler divergence (nats)	0.106 ± 0.160	0.445 ± 0.499	0.016

[a] $p < 0:05$, significant.

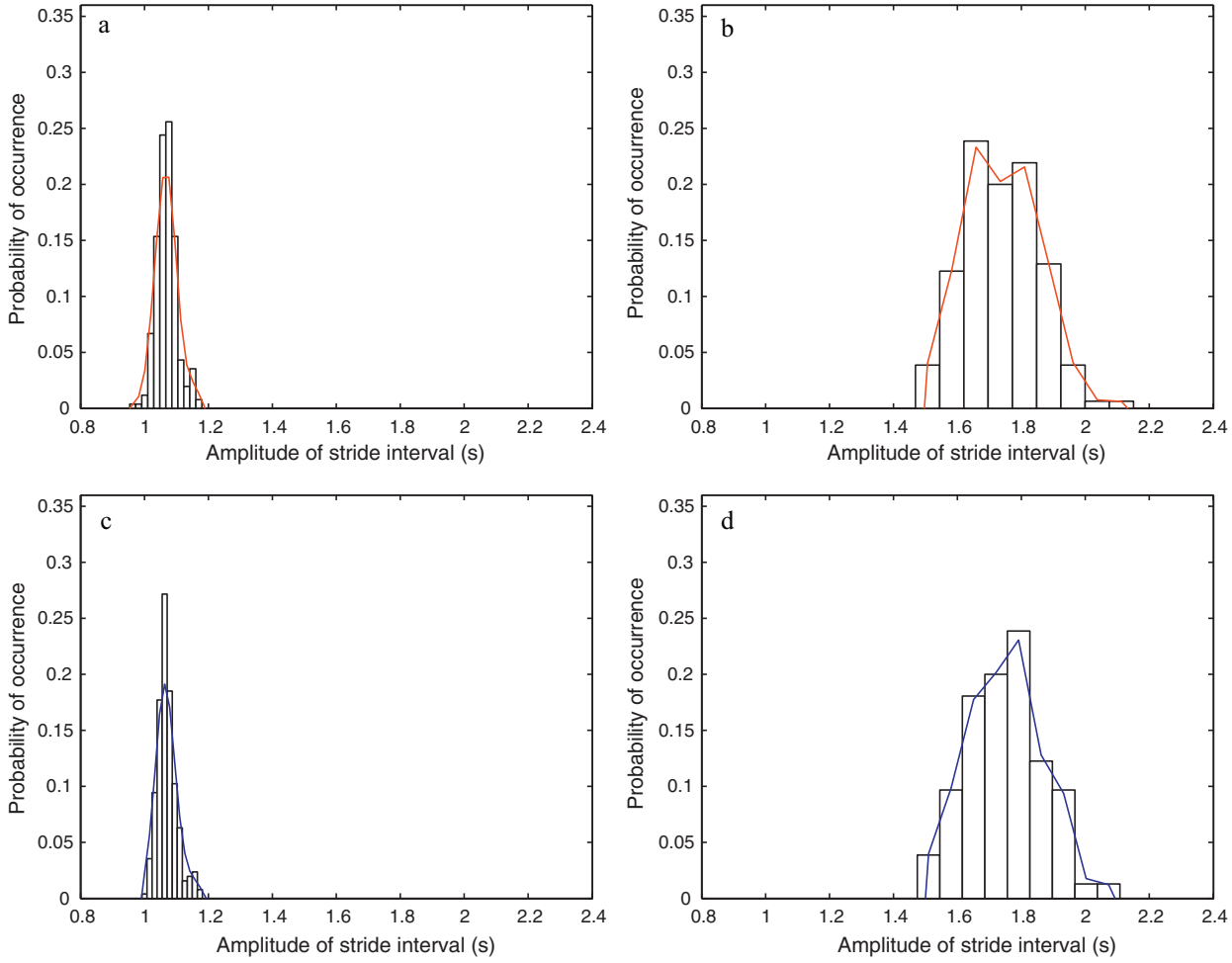

Fig. 3. Histograms and Parzen-window probability density functions (PDFs) of stride interval estimated from (a) left foot and (c) right foot of a healthy control subject (age: 57, female); (b) left foot and (d) right foot of a patient with amyotrophic lateral sclerosis (age: 65, male). The number of bins computed for the histograms of the healthy control and the patient with amyotrophic lateral sclerosis (ALS) were 12 and 9, respectively. The Parzen-window parameters σ_P computed for the PDFs of the healthy control and the ALS patient were 0.02 and 0.03, respectively.

correlated (over 0.99 in terms of Pearson's correlation coefficient), for either the CO or ALS subject group. Thus we only considered the mean of the left-foot stride interval, μ_L, as one of the dominant features. The statistics of the mean of the left-foot stride interval for the CO and ALS subject groups are listed in Table 2, from which we may observe that μ_L was increased by 23.9% (an increment of 0.261 s) in the group of ALS patients, compared with the CO subject group. The corresponding p-value obtained with the Student's t-test [30] was lower than 0.001, indicating that the μ_L parameter of the ALS patients was significantly different from that of the CO subjects. In addition, it is clear from Table 2 that the mean value of MKLD for the group of ALS patients was four times larger than that of the CO subject group, which implied that the stride interval time series became asymmetric in the ALS patients, as confirmed by Liao et al. [12]. The p-value of the MKLD ($p < 0.02$) also indicated a significantly difference between the two subject groups, so that we utilized the μ_L and MKLD parameters to construct the set of dominant features for the pattern classification task. Fig. 4 plots in the two-dimensional feature space the stride patterns of the CO and ALS subject groups. Note that most of the CO patterns congregated in a restricted region where $\mu_L < 1.2$ s and MKLD < 0.6 nat, whereas the ALS patterns were widely dispersed in the upper district where $\mu_L > 1.1$ s.

The classification results of the Gaussian-kernel-based LS-SVM in the ATAT and LOO training–testing styles are presented in Table 3. The LS-SVM trained in the ATAT style, labeled as LS-SVM/ATAT, can distinguish the stride patterns of two subject groups with an accurate rate of 93.1% (sensitivity: 0.923; specificity: 0.938). The LS-SVM evaluated with the LOO method, labeled as LS-SVM/LOO, provided an accurate rate of 82.8% (sensitivity: 0.769; specificity:

Table 3
Confusion matrix for the performance of the least squares support vector machine (LS-SVM) evaluated by the leave-one-out (LOO) cross-validation method. Subjects were categorized into the healthy control (CO) and amyotrophic lateral sclerosis (ALS) groups.

Evaluation methods	Predicted groups	Actual groups		Sensitivity	Specificity	Overall accuracy
		ALS	CO			
All-training–all-testing	ALS	12	1	0.923	0.938	93.1%
	CO	1	15			
Leave-one-out	ALS	10	2	0.769	0.875	82.8%
	CO	3	14			

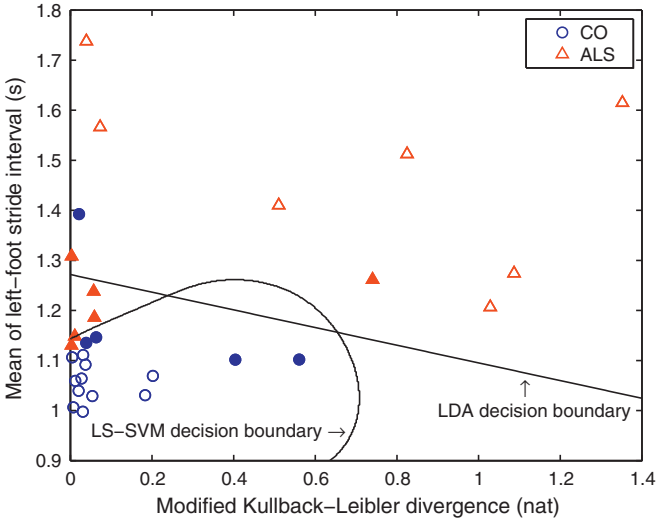

Fig. 4. Scatter plot of the stride patterns of the healthy control (CO) subjects, marked as *circles*, and of the patients with amyotrophic lateral sclerosis (ALS), marked as *triangles*. The stride patterns were displayed with the features of the modified Kullback–Leibler divergence (MKLD) and the mean of the left-foot stride interval (μ_L). The linear and nonlinear decision boundaries were provided by the linear discriminant analysis (LDA) and the least squares support vector machine (LS-SVM) with Gaussian kernels, respectively. Both of two classifiers were trained with the statistical features computed from all 16 healthy CO subjects and 13 ALS patients. Solid data points represent the support vectors, for which the absolute values of Lagrange multipliers were larger than 1.

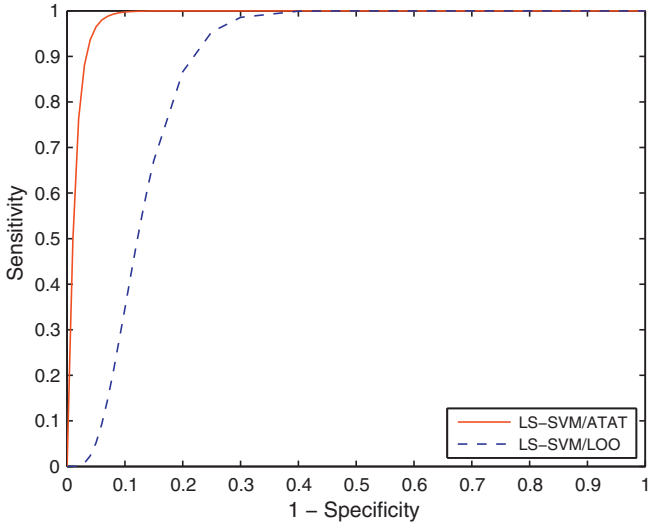

Fig. 5. Receiver operating characteristic (ROC) curves obtained by the Gaussian-kernel-based least squares support vector machine (LS-SVM), which was evaluated by the all-training–all-testing style (labeled as LS-SVM/ATAT) and the leave-one-out cross-validation method (labeled as LS-SVM/LOO), respectively. The input features were the modified Kullback–Leibler divergence (MKLD) and the mean of the left-foot stride interval (μ_L). The corresponding areas (A_z) under the ROC curves for the LS-SVM/ATAT (*solid line*) and LS-SVM/LOO (*dashed line*) were 0.985 (standard error, SE: 0.025) and 0.869 (SE: 0.072), respectively.

0.875). By comparing the decision boundaries generated by the LS-SVM and the LDA (both implemented in the ATAT style), we may observe from Fig. 4 that the decision boundary of the LS-SVM was much better than that of the LDA, which demonstrated the advantages of the LS-SVM based on nonlinear kernels and the support vectors (solid *circles* and *triangles* in Fig. 4), for analysis of the gait in ALS. It was not able to provide the decision boundary of the LS-SVM/LOO, as both of the training and testing sets were varied in the cross-validation test, such that the support vectors and the discriminant function of the LS-SVM/LOO could be altered from one training set to another. From Fig. 5, the A_z value under the ROC curve produced by the LS-SVM/ATAT remarkably reached 0.985, with a very low standard error (SE) equal to 0.025. Although the overall accurate rate obtained by the LS-SVM/LOO was not as high as that of the LS-SVM/ATAT, the LS-SVM/LOO was still able to generate an excellent ROC curve (A_z: 0.869; SE: 0.072), as shown in Fig. 5.

4. Discussion

In the present study, the *p*-values of μ_L and MKLD parameters indicated that the statistical characteristics were significantly perturbed in the gait of ALS patients. Based on these dominant features, the LS-SVM with Gaussian kernels was able to provide excellent diagnostic ROC performance on the 29 subjects studied. It is worth noting that either μ_L or μ_R can be used as an alternative feature for the pattern analysis, as they were highly correlated with each other. However, such a high correlation does not indicate the symmetry in the left-foot and right-foot gait rhythm time series of ALS patients, because the mean of stride interval computed based on the PDF does not contain any information about gait fluctuations.

Although the MKLD parameter can be used to identify the stride patterns in ALS, it is worth noting that the MKLD is not suitable to serve as a measure of gait symmetry, as such a statistical parameter does not contain any temporal information. For example, two time series with completely different temporal structures could also cause a near-zero-value MKLD, as long as their samples produce similar probability distributions of amplitude, whereas the MKLD with a large value definitely indicates some degree of asymmetry in the gait. Therefore, the MKLD with a zero or near-zero value is a necessary but not sufficient condition for gait symmetry. In the present study, we only used the MKLD parameter as a discriminant feature, rather than a gait symmetry indicator, for the analysis of the stride patterns in two subject groups.

The same gait database was also investigated in a few recent studies. Carletti et al. [31] proposed a linear model to interpret the stride interval time series in healthy controls, ALS, PD, and HD. Unfortunately, such a linear model was not very suited for analysis of the gait in ALS, as only 57% of the stride interval time series in ALS can be correctly predicted using this linear model, which was the worst among the stride-to-stride prediction results. The nonlinear model used in the study of Scafetta et al. [13] is useful for analysis of the fractal properties of human gait, but their model is established on a few prior assumptions (e.g., the standard deviation of the random walk process remains constant), which definitely limit the applications. Dutta et al. [32] extracted three types of cross-correlation-based features (i.e., centroid, mean-square abscissa, and variance of the abscissa) from the left-foot and right-foot time series of double support time, stance interval, and swing interval, and then applied the Elman's recurrent neural network to distinguish the stride patterns in patients with ALS, PD, and HD, from those of healthy subjects. Although the gait rhythm features illustrated in their paper showed notable correlations between each other, however, they did not analyze correlation relationships between these features. In our opinion, analysis of feature correlations is necessary before classifications, because we observed in the present study that the mean value of stride interval for the left-foot time series was highly correlated with that for the right-foot time series. The PDF-based analysis presented in this article improves the signal turns count method used in the study of Wu and Krishnan [16]. Since the signal turns count method can only be applied to continuous time series, the removal of stride outliers would not be feasible. Instead, Wu and Krishnan [16] had to use the median value of stride interval to replace the outliers. In the present study, the mean of stride interval and the MKLD would not be affected

by the removal of stride outliers, because the Parzen-window PDF estimate does not require the time indices of the stride interval samples.

The present study also contributed the accuracy improvements to the classification of ALS strides. By using the same ATAT evaluation method, the averaged correct classification rate obtained by the Elman's recurrent neural network method was reported to be 91.7% for ALS patients in the study of Dutta et al. [32], which is a bit lower than the 92.3% accuracy in the present study. Besides the lack of feature correlation study, the work of Dutta et al. [32] did not either remove the stride outliers before the feature extraction, which caused a higher error rate of the classification. Aziz and Arif [15] creatively introduced the sequence coding into the stride analysis, and obtained an A_z value of 0.82 under the ROC curve using only one feature (the normalized corrected Shannon entropy). In contrast, the present study showed that a combination of the dominant features could improve the performance of the classifier (A_z of LS-SVM/LOO: 0.869), in categorizing the ALS strides.

5. Limitations and conclusion

Since the age and gender of the subjects in two groups were not matched very well, the present study has some limitations. However, it was reported in the previous study that the effect of gender on usual walking patterns is very small [33]. The effect of age on the gait is complex [34], however, studies suggested that neurological disorders seem to contribute much more to altered gait rhythm than physiological aging [6]. In addition, the latest study of Hernandez et al. [35] suggested that the gait speed and step length of healthy elderly subjects (age >60 years) are not significantly different from those of young subjects (age <40 years), although the center of mass is performed in a different style in elderly adults. In the experiment, a 20-year-old healthy subject possessed a μ_L value of 1.393 s, which was much larger than the values of the elderly subjects in the healthy CO group, or even comparable to the averaged level (1.354 s) in the ALS subject group. In addition, the database available cannot support the study of the gait in post-stroke hemiplegia or cerebellar ataxia, which may also cause a larger averaged stride interval. The size of the current database is small, which limits the test of the generalization ability of the LS-SVM classifier. Future studies need to recruit more subjects to construct a training data set and a separate testing set for a better classification performance evaluation.

The turning-related stride outliers is usually regarded as a disturbance in the long-term indoor gait monitoring, because the subjects would reach the end of limited-length walking path very soon and then have to turn around. To overcome such a drawback, the treadmill could be introduced in the study, although some patients with movement disorders may not totally adapt to the speed control configuration. The present study, on the other hand, made use of the unpleasant stride outliers, that is to utilize the outlier annotations to estimate the averaged stride length, as described in Section 3.

The proposed MKLD parameter also contributed to the stride pattern classifications, and the computing of this feature was not affected by the removal of the outliers in the stride interval time series. However, it is worth noting that the advantage of the MKLD parameter also brings the disadvantage discussed in Section 4, because the nonparametric PDF estimate discarded the temporal structures of the stride time series.

The excellent decision boundary of the LS-SVM illustrated in Fig. 4 demonstrated the merit of nonlinear classifiers in automatic gait analysis. Although the statistical parameters and the nonlinear classifiers showed encouraging results and even the potential in helping analyze early indicators of ALS, we have to admit that the effectiveness of the proposed MKLD parameter and the classification methods was limited in gait monitoring due to the small size of the data set. To explore a better metric for monitoring the progression of ALS and the effects of intervention therapies in the future studies requires to recruit a larger number of ALS patients at different disease stages and the matched healthy subjects, and also demands to compare the gait parameters in ALS before and after the intervention therapies. Hausdorff et al. [36] recently studied the effects of external cueing (e.g., by means of a metronome) on gait variability, and the results demonstrated that rhythmic auditory stimulation set to 110% of the step rate may reduce fall risk in PD patients. It would be interesting to investigate whether the rhythmic auditory stimulation could also improve the mobility in ALS, although ALS is not a disorder caused by basal ganglia dysfunction.

Acknowledgments

This work was supported by the 2009 Fundamental Research Funds for the Central Universities of China under Grant No. 2010121061. The author would like to thank the anonymous reviewers whose valuable comments and suggestions help improve the quality of the paper.

Conflict of interest

There is no conflict of interest.

References

[1] Hirano A. Neuropathology of ALS: an overview. Neurology 1996;47(4 Suppl 2):63–6.
[2] Brown RH, Meininger V, Swash M. Amyotrophic lateral sclerosis. London, UK: Martin Dunitz Publishers; 1999.
[3] Brooks BR. Natural history of ALS: symptoms, strength, pulmonary function, and disability. Neurology 1996;47(4 Suppl 2):71–81.
[4] Sharma KR, Kent-Braun JA, Majumdar S, Huang Y, Mynhier M, Weiner MW, Miller RG. Physiology of fatigue in amyotrophic lateral sclerosis. Neurology 1995;45(4):733–40.
[5] Goldfarb BJ, Simon SR. Gait patterns in patients with amyotrophic lateral sclerosis. Archives of Physical Medicine and Rehabilitation 1984;65(2):61–5.
[6] Hausdorff JM, Mitchell SL, Firtion R, Peng CK, Cudkowicz ME, Wei JY, Goldberger AL. Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. Journal of Applied Physiology 1997;82(1):262–9.
[7] Hausdorff JM, Alexander NB. Gait disorders: evaluation and management. New York, NY: Informa Healthcare; 2005.
[8] Hausdorff JM, Peng CK, Ladin Z, Wei JY, Goldberger AL. Is walking a random walk? Evidence for long-range correlations in stride interval of human gait. Journal of Applied Physiology 1995;78(1):349–58.
[9] Hausdorff JM, Purdon PL, Peng CK, Ladin Z, Wei JY, Goldberger AL. Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations. Journal of Applied Physiology 1996;80(5):1448–57.
[10] Hausdorff JM, Lertratanakul A, Cudkowicz ME, Peterson AL, Kaliton D, Goldberger AL. Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis. Journal of Applied Physiology 2000;88(6):2045–53.
[11] Hausdorff JM. Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking. Human Movement Science 2007;26(4):555–89.
[12] Liao FY, Wang J, He P. Multi-resolution entropy analysis of gait symmetry in neurological degenerative diseases and amyotrophic lateral sclerosis. Medical Engineering and Physics 2008;30(3):299–310.
[13] Scafetta N, Marchi D, West BJ. Understanding the complexity of human gait dynamics. Chaos 2009;19(2):026108.
[14] West BJ, Griffin L. Allometric control, inverse power laws and human gait, Chaos. Solutions and Fractals 1999;10(9):1519–27.
[15] Aziz W, Arif M. Complexity analysis of stride interval time series by threshold dependent symbolic entropy. European Journal of Applied Physiology 2006;98(1):30–40.
[16] Wu YF, Krishnan S. Computer-aided analysis of gait rhythm fluctuations in amyotrophic lateral sclerosis. Medical and Biological Engineering and Computing 2009;47(11):1165–71.
[17] Wu YF, Ng SC. A PDF-based classification of gait cadence patterns in patients with amyotrophic lateral sclerosis. In: Proceedings of the 32nd annual international conference of IEEE engineering in medicine and biology society (EMBC'10). 2010. p. 1304–7.
[18] Moody GB, Mark RG, Goldberger AL. PhysioNet: a web-based resource for the study of physiologic signals. IEEE Engineering in Medicine and Biology Magazine 2001;20(3):70–5.

[19] Hausdorff JM, Ladin Z, Wei JY. Footswitch system for measurement of the temporal parameters of gait. Journal of Biomechanics 1995;28(3):347–51.
[20] Hahn GJ, Shapiro SS. Statistical models in engineering. Hoboken, NJ: Wiley; 1994.
[21] Scott DW. On optimal and data-based histograms. Biometrika 1979;66(3):605–10.
[22] Parzen E. On estimation of a probability density function and mode. Annals of Mathematical Statistics 1962;33(3):1065–76.
[23] Rangayyan RM, Wu YF. Screening of knee-joint vibroarthrographic signals using probability density functions estimated with Parzen windows. Biomedical Signal Processing and Control 2010;5(1):53–8.
[24] Kullback S, Leibler RA. On information and sufficiency. Annals of Mathematical Statistics 1951;22(1):79–86.
[25] Suykens JAK, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. Least squares support vector machines. Singapore: World Scientific Publishing; 2002.
[26] Vapnik VN. Statistical learning theory. New York, NY: Wiley; 1998.
[27] Cortes C, Vapnik VN. Support-vector networks. Machine Learning 1995;20(3):273–97.
[28] Duda RO, Hart PE, Stork DG. Pattern classification. 2nd ed. New York, NY: Wiley; 2001.
[29] Metz C, Herman B, Shen J. Maximum-likelihood estimation of ROC curves from continuously distributed data. Statistics in Medicine 1998;17(9):1033–53.
[30] Box JF. Gosset, fisher, and the t distribution. American Statistician 1981;35(2):61–6.
[31] Carletti T, Fanelli D, Guarino A. A new route to non invasive diagnosis in neurodegenerative diseases? Neuroscience Letters 2006;394(3):252–5.
[32] Dutta S, Chatterjee A, Munshi S. An automated hierarchical gait pattern identification tool employing cross-correlation-based feature extraction and recurrent neural network based classification. Expert Systems 2009;26(2):202–17.
[33] Gabell A, Nayak USL. The effect of age on variability in gait. Journal of Gerontology 1984;39(6):662–6.
[34] Kerrigan DC, Todd MK, Della Croce U, Lipsitz LA, Collins JJ. Biomechanical gait alterations independent of speed in the healthy elderly: evidence for specific limiting impairments. Archives of Physical Medicine and Rehabilitation 1998;79(3):317–22.
[35] Hernandez A, Silder A, Heiderscheit BC, Thelen DG. Effect of age on center of mass motion during human walking. Gait and Posture 2009;30(2):217–22.
[36] Hausdorff JM, Lowenthal J, Herman T, Gruendlinger L, Peretz C, Giladi N. Rhythmic auditory stimulation modulates gait variability in Parkinson's disease. European Journal of Neuroscience 2007;26(8):2369–75.