Tensor Decomposition of Gait Dynamics in Parkinson's Disease

Tuan D. Pham, *Senior Member, IEEE*, and Hong Yan, *Fellow, IEEE*

***Abstract*—*Objective:* The study of gait in Parkinson's disease is important because it can provide insights into the complex neural system and physiological behaviors of the disease, of which understanding can help improve treatment and lead to effective developments of alternative neural rehabilitation programs. This paper aims to introduce an effective computational method for multichannel or multisensor data analysis of gait dynamics in Parkinson's disease. *Method:* A model of tensor decomposition, which is a generalization of matrix-based analysis for higher dimensional analysis, is designed for differentiating multisensor time series of gait force between Parkinson's disease and healthy control cohorts. *Results:* Experimental results obtained from the tensor decomposition model using a PhysioNet database show several discriminating characteristics of the two cohorts, and the achievement of 100% sensitivity and 100% specificity under various cross validations. *Conclusion:* Tensor decomposition is a useful method for the modeling and analysis of multisensor time series in patients with Parkinson's disease. Significance: Tensor-decomposition factors can be potentially used as physiological markers for Parkinson's disease, and effective features for machine learning that can provide early prediction of the disease progression.**

***Index Terms*—Parkinson's disease, gait dynamics, time series, multi-sensors, tensor decomposition, pattern classification.**

I. Introduction

PARKINSON'S disease (PD) is one of the most common progressive disorders of the nervous system that causes tremor of hands, arms, legs, jaw and face, and stiffness or slowing of movement. Statistics on PD have reported it affects approximately 10 million people worldwide, and about 4% of them before the age of 50 [1]. A disturbed gait is a common symptom of patients with PD, while those with a severely disturbed gait are likely to suffer from falls and may lose their postural balance and coordination. It is therefore important to study the dynamics of gait patterns in PD to diagnose its variety of severity, which has implications for the prediction of falls in PD, improvement of treatment and rehabilitation strategies [2], [3]. The implications also include socio-economic impacts because the association of mobility with falls and freezing of gait is a trauma in daily lives of the patients, and the fear of falling can result in social withdrawal that leads to depression in patients with PD.

Computerized quantifications of stride-interval (the time between consecutive heel strikes of the same foot) and vertical ground reaction force changes can be of economic benefit and potential for better characterizing gait disturbance patterns in patients with PD to accurately predict stages of the disease, while it is difficult to identify gait in the clinical setting [4]. Several studies on statistical analysis and automated classification of gait patterns exhibited by patients with PD and HC subjects have been reported in literature. An early work by Hausdorff [5] described typical gait changes in patients with PD in terms of the fractal feature of gait in PD, and the association between the fractal feature of gait, stride length, and gait variability. Fractal methods were also carried out to analyze long time series of strides to differentiate between healthy control (HC) and PD groups [6].

Wu and Krishnan [7] applied the non-parametric Parzen window method to estimate the probability density functions of stride interval, swing interval, stance interval, and the signal turn counts of the gait time series to characterize and classify HC and PD groups using the least squares support vector machines. Khorasani and Daliri [8] classified the time series of the right stride interval of HC and PD subjects with the double-stochastic method of hidden Markov models. Su *et al.* [9] introduced gait asymmetry measures by using the wavelet transform to decompose vertical ground reaction force time series of the left and right feet to evaluate the difference in gait asymmetry between the control and PD subjects obtained from the same PhysioNet database used in this study. Zeng and Wang [10] applied the theory of dynamic learning to classify HC from PD, HD (Huntington's disease), and ALS (amyotrophic lateral sclerosis) subjects using features extracted from the left and right swing intervals and the left and right stance intervals obtained from the gait dynamics in neuro-degenerative disease database of the PhysioNet. Zeng and Wang [11] also applied deterministic learning for the classification of HC and PD subjects using another PhysioNet database, which was used in this study. Kamath [12] studied gait dynamics with aging and PD using complexity analysis developed in information theory. Ren *et al.*

Manuscript received August 27, 2017; revised October 23, 2017; accepted November 30, 2017. Date of publication December 4, 2017; date of current version July 17, 2018. This work was carried out while the first author visited the Signal Processing and BioComputing Lab, Department of Electronic Engineering, City University of Hong Kong, Hong Kong. This work was partially supported by the Hong Kong Research Grants Council (Project C1007-15G). *(Corresponding author: Tuan D. Pham.)*

T. D. Pham is with the Department of Biomedical Engineering, Linkoping University, Linkoping 581 83, Sweden (e-mail: tuan.pham@liu.se).

H. Yan is with the Department of Electronic Engineering, City University of Hong Kong.

Digital Object Identifier 10.1109/TBME.2017.2779884

[13] studied the gait fluctuations of HC, PD, HD, and ALS using the method of empirical mode decomposition to extract features for assessing the dissimilarity between the four cohorts, and trained several classifiers with these features to differentiate the four groups. Pham [14] most recently applied fuzzy recurrence plots to extract texture features from the time series of stride intervals to classify PD, HD, and ALS from HC subjects.

Tensor decompositions can effectively deal with the complexity of multi-sensor (multi-channel) data to allow effective analysis of the interactions between multiple modes or dimensions of a multi-index numerical array. In this paper, we present a tensor decomposition model for dynamic gait analysis in PD, where the tensor-decomposition coefficients can be used as useful features for differentiating patients with PD from HC subjects.

The rest of the paper is organized as follows. Section II describes the gait database of the PhysioNet to be analyzed in the present study. Section III presents a tensor decomposition of gait dynamics in PD and HC cohorts. Section IV reports results of gait analysis. Section V discusses findings from the experimental results and comparisons with other methods. Finally, Section VI is the conclusion of the research.

II. Gait Data in Parkinson's Disease

The database of Gait in Parkinson's Disease [15]–[18], which is publicly available at the PhysioNet website [19], was used in this study.

The database contains records of gait from 93 patients with idiopathic Parkinson's disease (PD), where the PD mean age is 66.3 years, and 63% of the PD cohort are men, and 72 healthy control (HC) subjects, where the HC mean age is 66.3 years; and 55% of the HC cohort are men. The database stores the vertical ground reaction force (VGRF) measurements of the subjects as they walked at their usual and self-selected pace for approximately 2 minutes on the ground level. The computerized force-sensitive system consists of a pair of shoes and a recording unit. Each shoe is equipped with 8 load sensors (Ultraflex Computer Dyno Graphy, Infotronic Inc) that cover the surface of the sole and measure the vertical forces (in newton) under each foot as a function of time. The recording unit, whose size and weight respectively are $19 \times 14 \times 4.5$ cm and 1.5 kg, was carried on the waist [15]. The output of each of these 16 sensors under the left and right feet were digitized and recorded at 100 samples per second.

According to the description obtained from the Gait in Parkinson's Disease database [19], when a subject is standing with both legs being parallel to each other, the locations of 8 sensors inside the insole in each foot can be depicted as lying approximately at the (X, Y) coordinates as shown in Table I, where the coordinates of origin (0,0) is between the legs and the subject is facing toward the direction of the positive side of the Y axis. The (X, Y) coordinates are in an arbitrary coordinate system, which reflects the relative positions of the sensors within each insole. During walking, the sensors inside each insole remain at the same relative position, but the two feet are no longer parallel to each other. This coordinate system therefore enables a calculation of a proxy for the location of the center of pressure under each foot. Furthermore, the demographic information of the PD and HC subjects and relevant measures of disease severity are also available from the database.

TABLE I
Locations of 8 Sensors on Each Foot [19], Where "L" and "R" Stand for Left and Right, Respectively

Sensor	X	Y
L1	−500	−800
L2	−700	−400
L3	−300	−400
L4	−700	0
L5	−300	0
L6	−700	400
L7	−300	400
L8	−500	800
R1	500	−800
R2	700	−400
R3	300	−400
R4	700	0
R5	300	0
R6	700	400
R7	300	400
R8	500	800

Figs. 1 and 2 show short segments of the VGRF time series recorded with 8 sensors on the right foot of an HC subject and a PD patient, respectively (the database consists of only records of time and forces without annotations of phases of the gait cycle such as heel stride, foot flat, mid-stance, heel-off, toe-off, and mid-swing). In a normal gait, the VGRF signal has two peaks: the first peak is when the foot strikes the ground and the second peak is caused by the force pushing off from the ground. The shape of the VGRF signal is abnormal in PD [20], [21], as can be seen in Fig. 2(h), showing the VGRF measured with sensor #8 located under the right foot of a PD subject, where narrow peaks in the VGRF signal can be observed. The magnitudes of the VGRF signals of the PD subject are generally lower than those of the HC subject due to reduced forces for the heel contact and the pushing-off phase.

III. Tensor Decomposition of Gait Dynamics

With the use of multi-linear algebra and the advanced power of computer computation, tensor analysis, which is known as the multi-way array method, has recently been gaining increasing interest in the analysis of multi-sensor data [22]–[26]. Tensor computing has much better capacity and flexibility in modeling multi-dimensional data for extracting useful features than vector-based (one-way) or matrix-based (two-way) methods [25].

Two well-known models for tensor decomposition are the canonical polyadic decomposition (CPD) [27], and the Tucker decomposition [28]. The CPD is also known as the PARAFAC [29], which stands for parallel factor analysis, and the CANDECOMP [30], which stands for canonical decomposition.

In this study, the PARAFAC, which is a generalization of the principal component analysis (PCA) to arrays of higher orders, was applied for decomposing tensors of gait data with multi-sensors. In tensor analysis, the terms *mode*, *way*, and

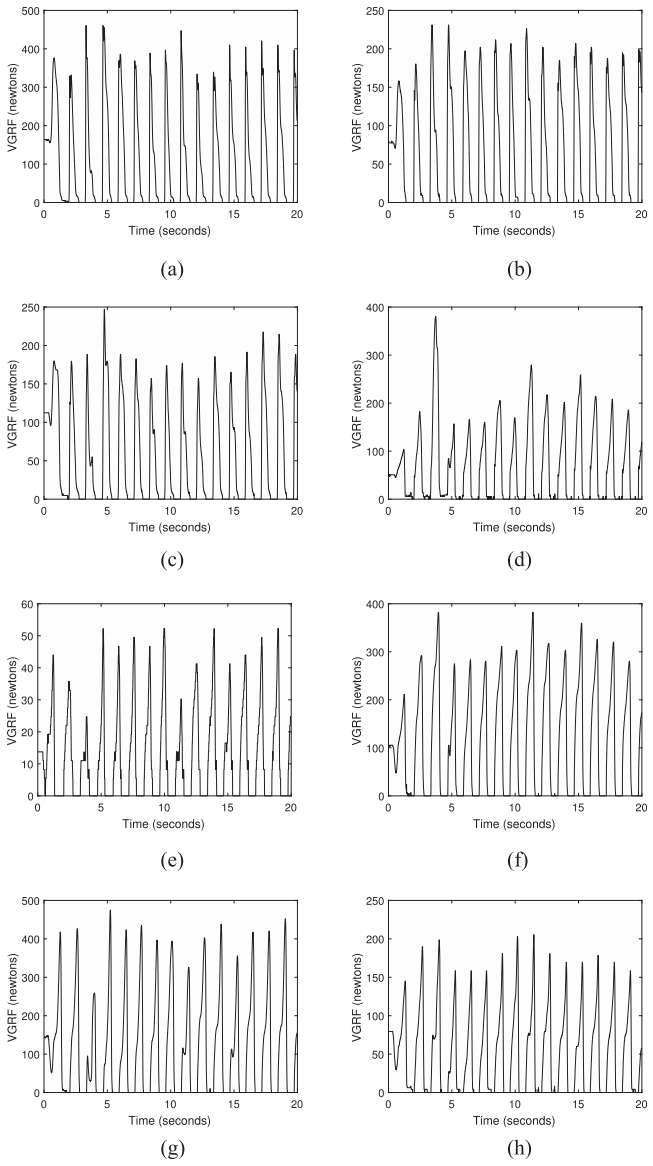

Fig. 1. Vertical ground reaction force (VGRF) of first 2000 time points recorded with 8 sensors placed under the right foot of a healthy control subject. (a) Sensor #1. (b) Sensor #2. (c) Sensor #3. (d) Sensor #4. (e) Sensor #5. (f) Sensor #6. (g) Sensor #7. (h) Sensor #8.

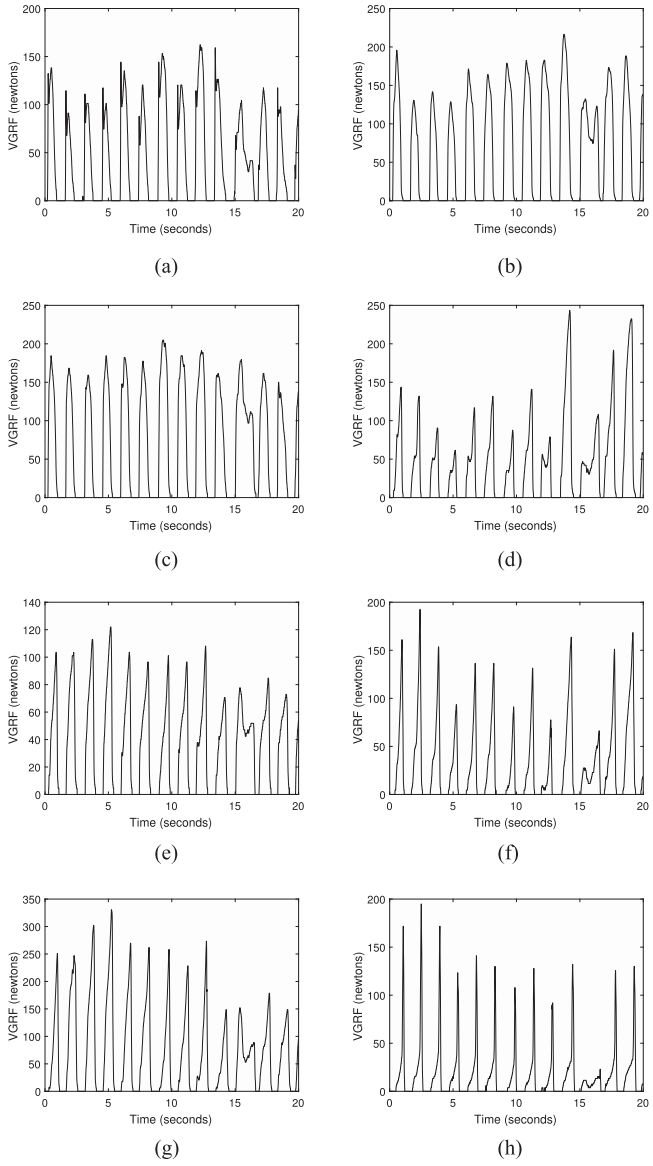

Fig. 2. Vertical ground reaction force (VGRF) of first 2000 time points recorded with 8 sensors placed under the right foot of a patient with Parkinson's disease. (a) Sensor #1. (b) Sensor #2. (c) Sensor #3. (d) Sensor #4. (e) Sensor #5. (f) Sensor #6. (g) Sensor #7. (h) Sensor #8.

order generally have the same meaning; and also, the terms *component* and *factor* are used interchangeably [22]. In general, the PARAFAC model of an N-mode tensor of dimensions $L_j, j = 1, \ldots, N$, can be mathematically expressed as

$$\underline{\mathbf{X}} = \sum_{f=1}^{F} \mathbf{a}_f^{(1)} \otimes \mathbf{a}_f^{(2)} \otimes \cdots \otimes \mathbf{a}_f^{(N)} + \underline{\mathbf{E}}, \tag{1}$$

where $\underline{\mathbf{X}} \in \mathcal{R}^{L_1 \times L_2 \times \cdots \times L_N}$ is the tensor, $\underline{\mathbf{E}}$ is the model error tensor, $\otimes$ is the outer product, F is the number of factors, and $\mathbf{a}_f^{(j)}, j = 1, \ldots, N$, are the f-th columns of the loading matrices $\mathbf{A}^{(j)} = [\mathbf{a}_1^{(j)}, \ldots, \mathbf{a}_F^{(j)}]$.

The tensor model can also be defined as

$$\underline{\mathbf{X}} = \underline{\hat{\mathbf{X}}} + \underline{\mathbf{E}} \approx \underline{\hat{\mathbf{X}}}, \tag{2}$$

where $\underline{\hat{\mathbf{X}}}$ is the estimate of the tensor, which is the linear combination of f-term tensors expressed in (1).

The solution to the PARAFAC model can be obtained using the alternating least squares method (ALS), which was described in [22], [31]. The PARAFAC ALS algorithm works by successively initializing the loadings in two modes, then iteratively estimating the last mode by the least squares regression until the estimates of the loading matrices converge.

The fourth-order tensor for gait dynamics in HC and PD cohorts can be modeled with 4 modes that include *subjects*,

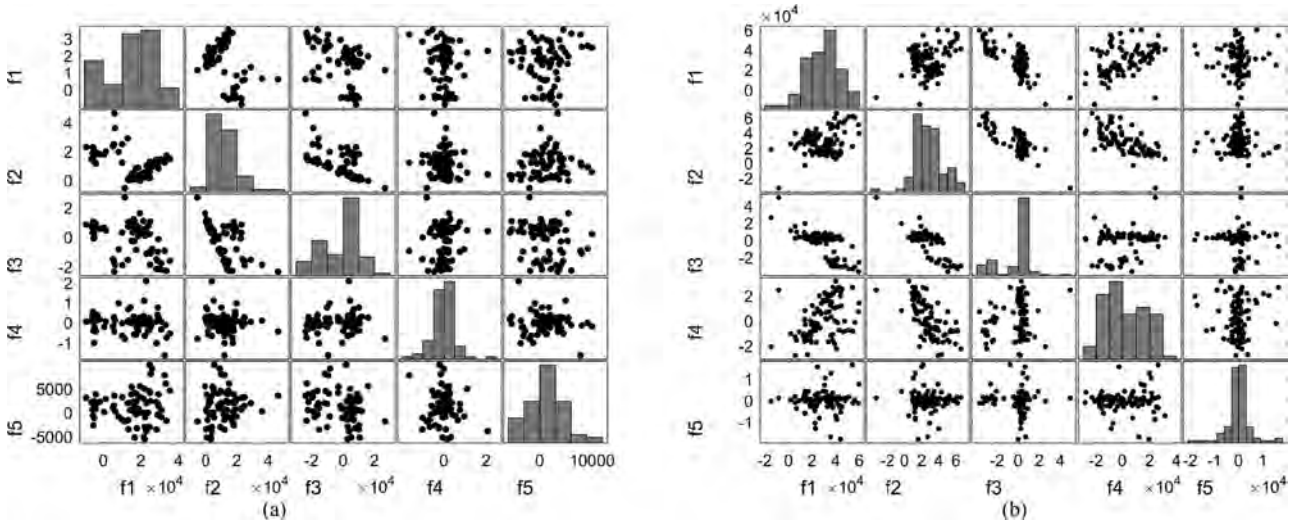

Fig. 3. Bivariate scatter plot matrices between 5 factors (f1 ... f5), with a univariate histogram for each factor (diagonal subplots), obtained from 4-way tensor decomposition of VGRF time series of healthy control (HC) and Parkinson's disease (PD) cohorts, exhibiting different distribution patterns and magnitudes between the two cohorts. (a) HC. (b) PD.

feet, *sensors*, and *VGRF* as follows:

$$\underline{\mathbf{X}}_{HC} \approx \sum_{f=1}^{F} \mathbf{a}_{f,HC}^{(1)} \otimes \mathbf{a}_{f}^{(2)} \otimes \mathbf{a}_{f}^{(3)} \otimes \mathbf{a}_{f,HC}^{(4)}, \tag{3}$$

where $\underline{\mathbf{X}}_{HC}$ stands for the tensor of the HC cohort, $\mathbf{a}_{f,HC}^{(1)}$ is the first-mode vector whose length is equal to the number of HC subjects, $\mathbf{a}_{f}^{(2)}$ is the second-mode vector whose length is equal to 2, which stands for the left and right feet, $\mathbf{a}_{f}^{(3)}$ is the third-mode vector whose length is equal to the number of sensors, and $\mathbf{a}_{f,HC}^{(4)}$ is the fourth-mode vector whose length is equal to the length of the VGRF time series of the HC; and

$$\underline{\mathbf{X}}_{PD} \approx \sum_{f=1}^{F} \mathbf{a}_{f,PD}^{(1)} \otimes \mathbf{a}_{f}^{(2)} \otimes \mathbf{a}_{f}^{(3)} \otimes \mathbf{a}_{f,PD}^{(4)}, \tag{4}$$

where $\underline{\mathbf{X}}_{PD}$ stands for the tensor of the PD cohort, and other variables are defined likewise as in (3).

IV. Results

The PARAFAC of the freely available N-way Toolbox for Matlab [31] was used for the tensor decomposition of the gait-dynamics time series of the HC and PD subjects in this study, where no constraints were imposed on the loadings of the 4 different modes of the tensor model. The PARAFAC algorithm arranges the factors as in the PCA, where the most important factor is made the first factor, and so on. The PARAFAC model of the 4-way array with 5 factors was used in this study, which was based on the formulation proved in [32] to ensure the uniqueness of the PARAFAC solution [33]: $\sum_{j=1}^{N} k^{(j)} \geq 2F + (N-1)$, where $k^{(j)}$ is the k-rank of $\mathbf{A}^{(j)}$.

Fig. 3 shows the bivariate scatter plots of the 5 factors, with a univariate histogram for each factor, obtained from the 4-way tensor decomposition of the gait time series of the HC and PD cohorts. Fig. 4(a)–(d) respectively show the scatter plots of the first and second factors, second and third factors, third and fourth factors, and fourth and fifth factors of the first model loading matrix (subjects) of the 4-way tensor decomposition of the HC and PD cohorts. Table II shows the means and standard deviations of the 5 factors of the two cohorts. Fig. 5 shows the plot of the first three factors of the second loading matrix (feet) of the 4-way tensor decomposition of the left feet against the right feet of the HC and PD subjects. Table III shows values of all 5 factors for the left and right feet of the HC and PD cohorts.

To validate the effectiveness of the tensor decomposition of the gait dynamics, all 5 factors of the first loading matrices of the HC subjects and PD patients were used for the binary classification using the least-squares support vector machines (LS-SVM) [34], where the Matlab software is publicly available [35]. To show the effectiveness of the tensor decomposition of multi-sensor data for gait analysis in comparison with the PCA, the first 5 principal component coefficients extracted from the gait data of the HC and PD groups were used for the classification by the LS-SVM. Table IV shows the leave-one-out, 2-fold, 5-fold, and 10-fold cross-validation results of the tensor decomposition (TD), principal component analysis (PCA), and the deterministic learning (DL) [11].

V. Discussion

The matrices of bivariate scatter plots (Fig. 3) shows different patterns in the relationships between pairs of the factors between the HC and PD cohorts. All sub-figures in Fig. 4 suggest that the factor distributions of the PD cohort are more disperse than those of the HC cohort. The first two TD factors of the HC and PD cohorts are all positive, where those of the HC are about double greater than those of the PD, and the standard deviations of the PD cohort are higher than those of the HC cohort (Table II), suggesting a higher degree of variability in gait pattern among the PD group than the HC. A clear separation of the plot of

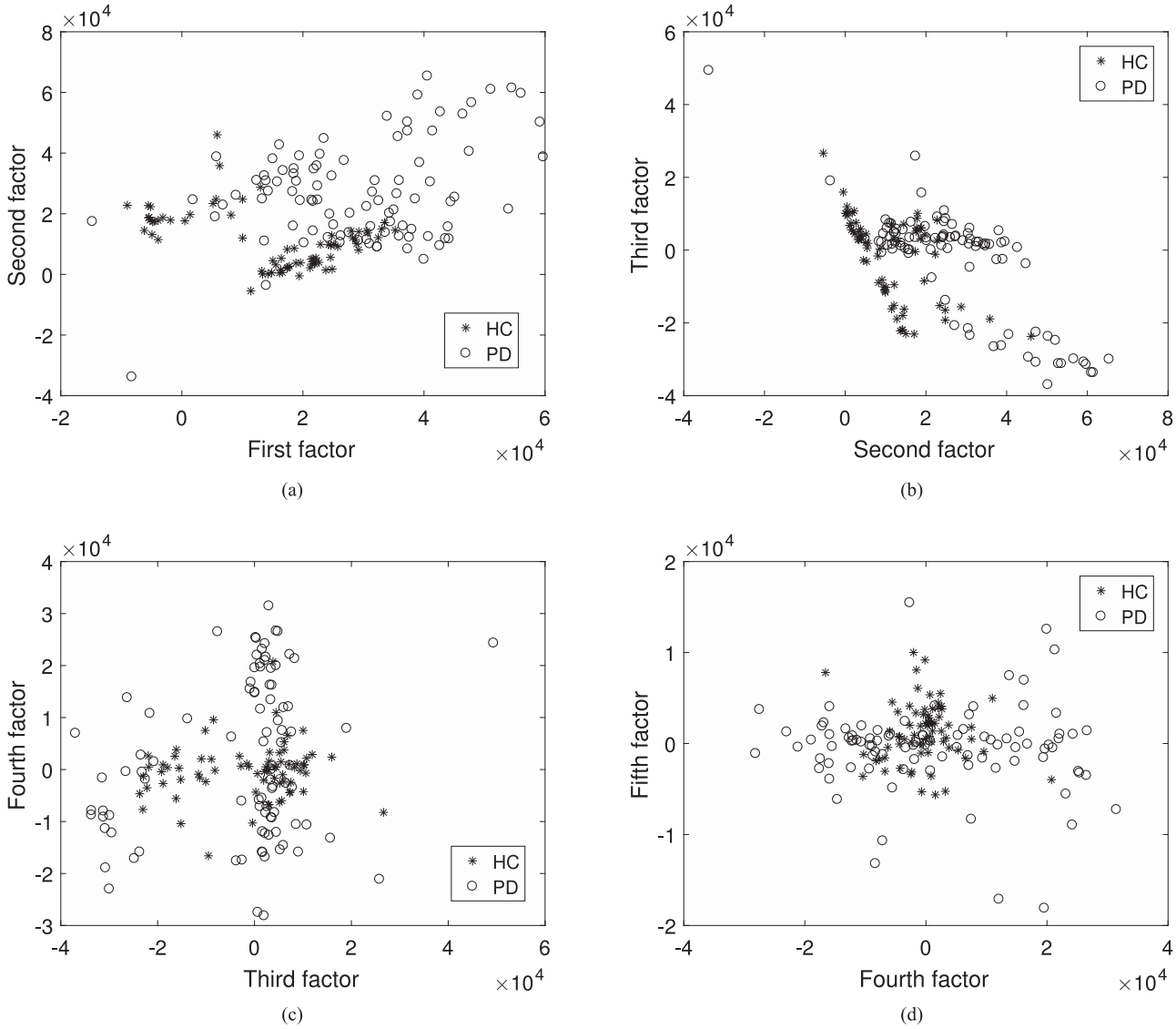

Fig. 4. Scatter plots of pairs of factors obtained from 4-way tensor decomposition of VGRF time series of healthy control (HC) and Parkinson's disease (PD) cohorts. The plots of the PD cohort show more dispersion than those in the HC cohort.

TABLE II
MEANS ($\times 10^4$) AND STANDARD DEVIATIONS ($\times 10^4$) OF 5 FACTORS (F1 ... F5) OF THE FIRST LOADING MATRIX (SUBJECTS) FOR HEALTHY CONTROL (HC) AND PARKINSON'S DISEASE (PD) COHORTS OBTAINED FROM 4-WAY TENSOR DECOMPOSITION

Cohort	f1	f2	f3	f4	f5
HC	1.4842 ± 1.2308	1.1203 ± 0.9245	-0.1781 ± 1.1474	-0.0294 ± 0.5179	0.1269 ± 0.3262
PD	2.9087 ± 1.4132	2.7041 ± 1.6323	-0.2882 ± 1.4963	0.1823 ± 1.5031	-0.0316 ± 0.5051

the first three factors of the second loading matrix (feet) of the 4-way tensor decomposition of the left feet against the right feet of the HC and PD subjects can be observed in Fig. 5, where the factors of the right foot of the HC cohort are low and the factors of the left foot of the HC cohort are high, and vice versa for the PD cohort. The first 4 factors for the left and right feet of the HC and PD shown in Table III agree with the observation from Fig. 5. These findings may have an important implication as tensor-based physiological markers for differentiating gait dynamics between HC and PD subjects.

As shown in Table IV, except for the case of 10-fold cross-validation, where there is no comparison with the DL, all the cross-validation results obtained from the TD are higher than those obtained from the DL and PCA. The average area under the receiver operating characteristic curve (AUC), where an AUC of 1 means a perfect test and an AUC of 0.5 represents

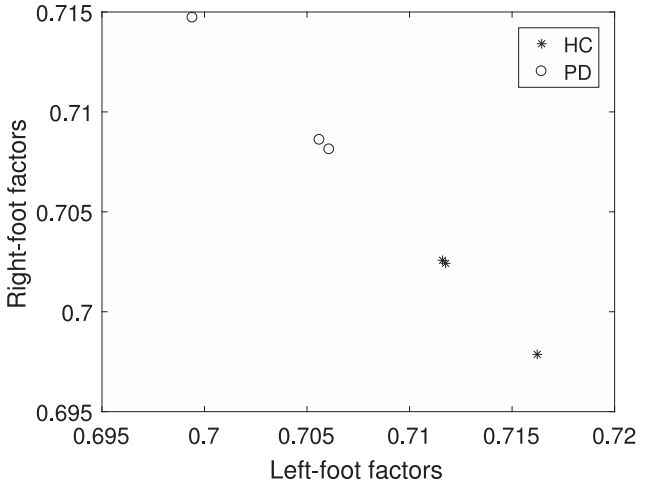

Fig. 5. First 3 factors of 4-way tensor decomposition of right-foot vs. left-foot VGRF time series of healthy control (HC) and Parkinson's disease (PD) cohorts.

TABLE III
VALUES OF 5 FACTORS (F1 ... F5) OF THE SECOND LOADING MATRIX (LEFT AND RIGHT FEET) FOR HEALTHY CONTROL (HC) AND PARKINSON'S DISEASE (PD) COHORTS OBTAINED FROM 4-WAY TENSOR DECOMPOSITION

Cohort	f1	f2	f3	f4	f5
HC left foot	0.7116	0.7118	0.7162	0.7101	0.8092
HC right foot	0.7026	0.7024	0.6979	−0.7041	0.5876
PD left foot	0.7061	0.7056	0.6994	0.7098	0.7198
PD right foot	0.7081	0.7086	0.7147	0.7044	−0.6941

a random result, was computed for each cross-validation. The average AUC values of the four cross-validations obtained from LS-SVM using TD features are all 1, which indicate the high performance of the proposed method for studying gait dynamics. The DL method extracted features as the differences between the VGRF values of L3 and R3 sensors, L6 and R6 sensors, and the sum of the eight sensor outputs from each foot, resulting in a 4-dimensional feature vector for each HC subject and each PD patient. These features were used to model and learn the gait dynamics with the radial basis function based neural network. While the cross-validation results of the DL methods were fairly high, this method mainly concerned with the classification of HC subjects and PD patients, and was not able to discover novel physiological characteristics of the two cohorts hidden in the multi-sensor data. The proposed TD method systematically and fully utilized the data provided by all the sensors to model gait dynamics, and produced the TD factors as properties that show distinct attributes between the two groups as well as effective features for pattern classification. Based on the comparative results obtained from the PCA and 4-way TD (Table IV), where the TD performed much higher than the PCA, it is obvious that the TD is a more adequate, robust and interpretable model than the PCA for modeling gait dynamics obtained from multi-sensor data. Another advantage of the TD over the PCA is that higher-order tensor decompositions are often unique, whereas matrix decompositions are not [23].

TABLE IV
CROSS-VALIDATION RESULTS FOR CLASSIFICATION OF HEALTHY CONTROL AND PARKINSON'S DISEASE SUBJECTS OBTAINED FROM DETERMINISTIC LEARNING (DL), PRINCIPAL COMPONENT ANALYSIS (PCA), AND TENSOR DECOMPOSITION (TD) METHODS IN TERMS OF SENSITIVITY (SEN), SPECIFICITY (SPE), AREA UNDER ROC CURVE (AUC), AND ACCURACY (ACC), WHERE 'X' INDICATES NOT AVAILABLE

Method	SEN (%)	SPE (%)	AUC	ACC (%)
		Leave-one-out		
DL	92.47	94.52	x	93.37
PCA	76.34	79.17	0.84	77.58
TD	100	100	1	100
		2-fold		
DL	91.49	91.67	x	91.57
PCA	86.02	87.50	0.92	86.67
TD	100	100	1	100
		5-fold		
DL	96.77	95.89	x	96.39
PCA	79.57	80.56	0.88	80.00
TD	100	100	1	100
		10-fold		
DL	x	x	x	x
PCA	86.02	87.50	0.92	86.67
TD	100	100	1	100

In this study, the number of factors F was chosen to be 5 in the 4-way FARAFAC model, because the sum squared prediction error given by $F = 5$ was lowest among $F = 1, \ldots, 5$ in both HC and PD tensors. When F was set to 6, some factors in both HC and PD tensors were highly correlated as the Tucker's congruence coefficients [37], which are indices of the similarity between factors, were greater than 0.85, suggesting to decrease the number of factors [31], [37]. The core consistency diagnostic (CORCONDIA) [38] was introduced for the heuristic determination of the appropriate number of factors in PARAFAC models. The CORCONDIA takes a tensor and its F-component PARAFAC decomposition as the input and produces a measure of consistency expressed in percentage, denoted as c, where the range of c is from some negative value to 100%. The PARAFAC model is considered valid if c is close to 100%. Given several inputs of the decompositions, the appropriate number of factors can be practically selected as the last high consistency value before the sharp decrease in c. However, it is pointed out that CORCONDIA can be misleading [39], and cannot be applied to large arrays, such as those used in this study ($93 \times 2 \times 8 \times 26366$ for the PD and $72 \times 2 \times 8 \times 12119$ for the HC), due to its lack of scalability and its modeling assumptions [40]. An extension of the core consistency diagnostic was described in [40], which requires the specification of a maximum rank to find a good number of factors for the tensor decomposition. The distinct observations of the plots of the factors of HC subjects and PD patients shown in Figs. 3 and 4, and the highest classification performance of the PARAFAC model of the 4-way array with 5 factors as shown in Table IV provide evidence that the use of this number of factors is appropriate.

Instead of using the 5 factors of all loading matrices of the PARAFAC model, which constitute a very high dimensional space for pattern classification, only 5 factors of the first loading matrix that represents the number of subjects was used in this experiment. It is because such a number of factors can be

sufficient enough for training the LS-SVM. In fact, the performance of the LS-SVM classification with the use of the 5 factors of the first loading matrix resulted in maximum accuracy (Table IV), illustrating the effectiveness of the proposed design of the tensor decomposition of the gait dynamics.

In particular, the classification accuracy rates provided by the tensor decomposition of the PD and HC gait dynamics are much higher than those obtained from the PCA in all cross-validations (Table IV), showing an advantage of the tensor decomposition over the PCA for multi-sensor data analysis.

In comparison with the best results obtained using combinations of the wavelet features and multilayer perceptron neural networks for classification (80% training, 10% validation, and 10% testing) reported in [9], the values for sensitivity = 80.1% (obtained from combining symlet and Daubechies wavelets), accuracy = 86.2% (obtained from combining Symlet and Daubechies wavelets), and AUC = 0.78 (obtained from Coiflet wavelet), all of which are lower than those obtained from the tensor decomposition.

In another comparison of the present results with those obtained from a similar study [36] using the PCA of VGRF tested on a different non-public database consisting of smaller cohorts (38 normal subjects, and 5 patients with lower limb fractures), the scatter plots of the pairs of principal component coefficients (PCC) in all cases show similar scattering patterns of the factor pairs for the two groups, which are less informative than those found in the 4-way tensor decomposition of the VGRF time series of the HC and PD cohorts. Furthermore, the best PCA-based leave-one-out cross-validation results (using 6 PCCs) for sensitivity = 84.60%, specificity = 100%, and accuracy = 96.10%, which are less than the sensitivity (100%), specificity (100%), and accuracy (100%) obtained from the tensor decomposition, testing with much larger cohorts of controls and patients.

Highlights of technical advantages of the proposed tensor decomposition of gait dynamics over other methods discussed above are that this tensor decomposition is a better match for exploratory signal analysis of complex multi-sensor data, where similar successful applications of tensor decompositions have been found in multi-sensor and multi-modal signal processing [25], it offers a great flexibility in the design for finding hidden components in multi-sensor data in the context of data fusion to discover discriminating information between patients with PD and HC subjects, and its factors are effective and robust (in the presence of noise and missing data [25]) to be used as compact features for training classifiers for automated recognition of gait disorders.

VI. Conclusion

The discovery of the distinct relationships of the tensor-decomposition factors of the controls and PD patients as well as the high performance of machine learning in the differentiation between the two groups under various cross-validations show the potential of the tensor model and its decomposition for gait analysis. The findings suggest the utilization of the proposed method as an objective means for quantifying locomotion associated with the disease, monitoring responses to therapeutic interventions and rehabilitation, as well as early prediction of falls, where reliable predictors of future falls in patients with Parkinson's disease are still inadequate at present [41].

The tensor model developed in this study can be readily applied for the analysis of multi-sensor time series in pathophysiological patterns from other disorders. In this study, the PARAFAC for tensor decomposition was applied to the gait analysis, application of the Tucker decomposition will also be explored in future study.

Unlike the case of matrices, the rank of a tensor is still not well understood. It was shown that the problem of computing the rank of a tensor is NP-hard [42]. The selection of the number of PARAFAC components in this study was somehow made ad hoc. Therefore, automated determination of optimal numbers of PARAFAC components for the tensor decomposition of gait would be expected to gain further insights into the pathophysiological mechanism of the disease.

The proposed method used the full length of the VGRF time series without considering the minimum number of strides required for effective tensor decomposition analysis of the gait dynamics. Tensor decompositions can effectively operate on signals of very short durations [25], therefore study on the minimum stride number is worth investigating. The application can be useful for short gait trials, because it is usually difficult to measure a large number of strides in clinical settings [6].

References

[1] Statistics on Parkinson's, *Parkinson's Disease Foundation*, New York, NY, USA, 2017. [Online]. Available: http://www.pdf.org/parkinson_statistics

[2] B. Lindholm *et al.*, "External validation of a 3-step falls prediction model in mild Parkinson's disease," *J. Neurol.*, vol. 263, pp. 2462–2469, 2016.

[3] A. Delval *et al.*, "Why we should study gait initiation in Parkinson's disease," *Neurophysiol. Clinique/Clinical Neurophysiol.*, vol. 44, pp. 69–76, 2014.

[4] P. H. Chen *et al.*, "Gait Disorders in Parkinson's Disease: Assessment and Management," *Int. J. Gerontol.*, vol. 7, pp. 189–193, 2013.

[5] J. M. Hausdorff, "Gait dynamics in Parkinson's disease: common and distinct behavior among stride length, gait variability, and fractal-like scaling," *Chaos*, vol. 19, 2009, Art. no. 026113.

[6] M. Kirchner *et al.*, "Detrended fluctuation analysis and adaptive fractal analysis of stride time data in Parkinson's disease: Stitching together short gait trials," *PLoS ONE*, vol. 9, 2014, Art. no. e85787.

[7] Y. Wu and S. Krishnan, "Statistical analysis of gait rhythm in patients with Parkinson's disease," *IEEE Trans Neural Syst. Rehabil. Eng.*, vol. 18, no. 2, pp. 150–158, Apr. 2010.

[8] A. Khorasani and M. R. Daliri, "HMM for classification of Parkinson's disease based on the raw gait data," *J. Med. Syst.*, vol. 38, no. 147, 2014, http://doi.org/10.1007/s10916-014-0147-5.

[9] B. L. Su *et al.*, "Characterizing gait asymmetry via frequency sub-band components of the ground reaction force," *Biomed. Signal Process. Control*, vol. 18, 2015, pp. 56–60, 2015.

[10] W. Zeng and C. Wang, "Classification of neurodegenerative diseases using gait dynamics via deterministic learning," *Inf. Sci.*, vol. 317, pp. 246–258, 2015.

[11] W. Zeng and C. Wang, "Parkinson's disease classification using gait analysis via deterministic learning," *Neurosci. Lett.*, vol. 633, pp. 268–278, 2016.

[12] C. Kamath, "Analysis of altered complexity of gait dynamics with aging and Parkinson's disease using ternary Lempel-Ziv complexity," *Cogent Eng.*, vol. 3, 2016, Art. no. 1177924.

[13] P. Ren *et al.*, "Gait rhythm fluctuation analysis for neurodegenerative diseases by empirical mode decomposition," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 52–60, Jan. 2017.

[14] T. D. Pham, "Texture classification and visualization of time series of gait dynamics in patients with neuro-degenerative diseases," *IEEE Trans. Neural Syst. Rehabil. Eng.*, 27 Jul. 2017, doi: 10.1109/TNSRE.2017.2732448 .

[15] S. Frenkel-Toledo *et al.*, "Effect of gait speed on gait rhythmicity in Parkinson's disease: variability of stride time and swing time respond differently," *J. NeuroEng. Rehabil.*, vol. 2, no. 23, 2005, doi: 10.1186/1743-0003-2-23.
[16] S. Frenkel-Toledo *et al.*, "Treadmill walking as a pacemaker to improve gait rhythm and stability in Parkinson's disease," *Movement Disorders*, vol. 20, pp. 1109–1114, 2005.
[17] J. M. Hausdorff *et al.*, "Rhythmic auditory stimulation modulates gait variability in Parkinson's disease," *Eur. J. Neurosci.*, vol. 26, pp. 2369–2375, 2007.
[18] G. Yogev *et al.*, "Dual tasking, gait rhythmicity, and Parkinson's disease: Which aspects of gait are attention demanding?," *Eur. J. Neurosci.*, vol. 22, pp. 1248–1256, 2005.
[19] Gait in Parkinson's Disease, *PhysioNet*. [Online]. Available on: https://physionet.org/physiobank/database/gaitpdb/, Accessed on: Jun. 2017.
[20] S. H. Koozekanani *et al.*, "Ground reaction forces during ambulation in Parkinsonism: pilot study," *Arch Phys. Med. Rehabil.*, vol. 68, pp. 28–30, 1987.
[21] E. Ueno *et al.*, "Gait disorders in Parkinsonism. A study with floor reaction forces and EMG," *Adv. Neurol.* vol. 60, pp. 414–418, 1993.
[22] R. Bro, "PARAFAC. Tutorial and applications," *Chemometrics Intell. Lab. Syst.*, vol. 38, pp. 149–171, 1997.
[23] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, pp. 455–500, 2009.
[24] H. Becker *et al.*, "EEG extended source localization: Tensor-based vs. conventional methods," *NeuroImage*, vol. 96, pp. 143–157, 2014.
[25] A. Cichocki *et al.*, "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.
[26] F. Cong *et al.*, "Tensor decomposition of EEG signals: A brief review," *J. Neurosci. Methods*, vol. 248, pp. 59–69, 2015.
[27] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Phys.*, vol. 6, pp. 164–189, 1927.
[28] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, pp. 279–311, 1966.
[29] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis," *UCLA Working Papers Phonetics*, vol. 16, pp. 1–84, 1970.
[30] J. D. Carroll and J. J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, pp. 283–319, 1970.
[31] C. A. Andersson and R. Bro, "The N-way Toolbox for MATLAB," *Chemometrics Intell. Lab. Syst.*, vol. 52, pp. 1–4, 2000.
[32] A. Stegeman, "On uniqueness conditions for Candecomp/Parafac and Indscal with full column rank in one mode," *Linear Algebra Appl.*, vol. 431, pp. 211–227, 2009.
[33] J. B. Kruskal, "Rank, decomposition, and uniqueness for 3-way and n-way arrays," in *Multiway Data Analysis*, R. Coppi and S. Bolasco, Eds., Amsterdam, The Netherlands: North Holland, 1989, pp. 7–18.
[34] J. A. K. Suykens *et al.*, *Least Squares Support Vector Machines*. Singapore: World Scientific, 2002.
[35] Least squares—Support vector machines MATLAB/C toolbox. [Online]. Available on: http://www.esat.kuleuven.be/sista/lssvmlab/, Accessed on: Apr. 2017.
[36] A. M. S. Muniz and J. Nadal, "Application of principal component analysis in vertical ground reaction force to discriminate normal and abnormal gait," *Gait Posture*, vol. 29, pp. 31–35, 2009.
[37] U. Lorenzo-Seva and J. M. F. ten Berge, "Tucker's congruence coefficient as a meaningful index of factor similarity," *Methodology*, vol. 2, pp. 57–64, 2006.
[38] R. Bro and H. A. L. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *J. Chemometrics*, vol. 17, pp. 274–286, 2003.
[39] [Online]. Available: http://www.models.life.ku.dk/čourses/parafac/chap2parafac.htm, Accessed on: Oct. 2017.
[40] E. E. Papalexakis and C. Faloutsos, "Unsupervised tensor mining for big data practitioners," *Big Data*, vol. 4, pp. 179–191, 2016.
[41] M. Hoskovcova *et al.*, "Predicting falls in Parkinson disease: What is the value of instrumented testing in OFF medication state?," *PLoS ONE*, vol. 10, 2015, Art. no. e0139849. [Online]. Available on: https://doi.org/10.1371/journal.pone.0139849
[42] C. J. Hillar and L. H. Lim, "Most tensor problems are NP-hard," *J. ACM*, vol. 60, 2013, Art. no. 45, doi:10.1145/2512329.