A New Measure of Movement Symmetry in Early Parkinson's Disease Patients Using Symbolic Processing of Inertial Sensor Data

Anita Sant'Anna*, Arash Salarian, *Member, IEEE*, and Nicholas Wickström, *Member, IEEE*

***Abstract*—Movement asymmetry is one of the motor symptoms associated with Parkinson's disease (PD). Therefore, being able to detect and measure movement symmetry is important for monitoring the patient's condition. The present paper introduces a novel symbol based symmetry index calculated from inertial sensor data. The method is explained, evaluated, and compared to six other symmetry measures. These measures were used to determine the symmetry of both upper and lower limbs during walking of 11 early-to-mid-stage PD patients and 15 control subjects. The patients included in the study showed minimal motor abnormalities according to the unified Parkinson's disease rating scale (UPDRS). The symmetry indices were used to classify subjects into two different groups corresponding to PD or control. The proposed method presented high sensitivity and specificity with an area under the receiver operating characteristic (ROC) curve of 0.872, 9% greater than the second best method. The proposed method also showed an excellent intraclass correlation coefficient (ICC) of 0.949, 55% greater than the second best method. Results suggest that the proposed symmetry index is appropriate for this particular group of patients.**

***Index Terms*—Gyroscope, Parkinson's disease, symbolization, symmetry.**

I. Introduction

MOTOR dysfunctions have long been associated with neurological conditions. There is evidence that movement asymmetry is commonly observed in conjunction with a decline in health status [1]. Parkinson's disease (PD) patients, in particular, may exhibit very asymmetrical gait [2], and asymmetrical hand movements [3].

Although many symmetry measures have been suggested in the literature, there is still no accepted standard. Each symmetry measure may convey different information, and there is no correspondence between most measures. That is, one symmetry measure cannot be inferred from another. The lack of a standard measure and the fact that results are not transferable from one symmetry index to another suggest that each application may require a different optimal measure. Early PD patients, for example, exhibit very subtle movement characteristics that can only be evidenced by very sensitive tools. The objective of the present paper is to introduce, evaluate, and benchmark a new measure of movement symmetry, appropriate for early-to-mid-stage PD patients. The present paper will not dwell on clinical aspects of the proposed method; such analysis will be addressed in future works.

The sample of PD patients chosen for this study showed minimal motor abnormality and practically no postural instability. Characterizing the differences in movement symmetry between these patients and controls is a challenging task. The symmetry measure proposed here aims at detecting subtle movement asymmetries, which are characteristic of early PD symptoms. The proposed method is based on the analysis of symbolized inertial sensor data. The data were collected using gyroscopes attached to upper and lower limbs of the trial subjects during walking. These data were used to evaluate the performance of the proposed symmetry index in comparison with six other methods: symmetry index ($\mathrm{SI}_{\mathrm{index}}$), symmetry angle ($\mathrm{SI}_{\mathrm{angle}}$), gait asymmetry ($\mathrm{SI}_{\mathrm{GA}}$), trend symmetry ($\mathrm{SI}_{\mathrm{trend}}$), maximum angular velocity ratio ($\mathrm{SI}_{\mathrm{ratio}}$), and latency-corrected ensemble average symmetry magnitude ($\mathrm{SI}_{\mathrm{LCEA}}$). All methods were evaluated for sensitivity and test-retest reliability.

II. Related Work

Symmetry measures typically compute the similarity between the movements of the right and left sides of the body. Once the movement information has been obtained (via sensors, cameras or others), there are several different ways to compute the similarity between both sides. Generally speaking, gait symmetry measures can be computed: 1) from certain discrete values, e.g., double support time, stride length, or 2) from continuous signals, e.g., EMG signals, accelerometer signals. Through this document, the former will be referred to as *discrete* methods, and the latter as *continuous* methods.

Discrete methods are the most common. They require simple movement information, and they can be easily computed. For this reason, the earlier symmetry indices belong to this category. In the late 1980s, Robinson *et al.* [4] proposed a simple way to quantify symmetry between right and left lower limbs. Since then, several other authors have proposed small variations to Robinson's method, e.g., [5]–[8]. Robinson's formula has also been used to quantify asymmetry of upper limbs [9]. The major limitation of this method is that a reference value must be

Manuscript received November 9, 2010; revised February 14, 2011 and April 4, 2011; accepted April 20, 2011. Date of publication May 2, 2011; date of current version June 17, 2011. *Asterisk indicates corresponding author.*

*A. Sant'Anna is with the Intelligent Systems Lab at Halmstad University, PO Box 823, Halmstad, Halland, Sweden (e-mail: anita.santanna@gmail.com).

A. Salarian is with the Balance Disorders Laboratory at Oregon Health and Science University, Portland, OR 97239-3098 USA (e-mail: arash.salarian@gmail.com).

N. Wickström is with the Intelligent Systems Lab at Halmstad University, PO Box 823, Halmstad, Halland, Sweden (e-mail: nicholas.wickstrom@hh.se).

Digital Object Identifier 10.1109/TBME.2011.2149521

chosen, and inappropriate choices of reference may cause artificial inflation [7]. Zifchock *et al.* [7] proposed a symmetry angle measure that did not require a reference value. Zifchock's symmetry angle has also been applied to upper limb data [10]. Other examples of *discrete* symmetry measures are presented in [1], [11], [12], [13]. All these measures are limited to discrete values. If the motion data are obtained from a continuous source, preprocessing is needed to extract values which can later be used to estimate symmetry. *Discrete* methods normally make use of spatiotemporal gait measurements, or features extracted from kinematic data. Most gait symmetry studies have focused on lower limb measurements. However, some recent studies have also observed the role of arm swing symmetry during walking, e.g., [9], [10], [13].

Continuous methods compare similarities between two continuous signals; they, therefore, require continuous motion information such as EMG, acceleration, angular displacement or position. These methods are more easily applied to whole-body data. Moe-Nilssen and Helbostad [14], for example, introduced an unbiased autocorrelation method using trunk acceleration data. Crenshaw and Richards [15] calculated trend symmetry based on the variance around the 1st principal component of a right-side versus left-side plot. Miller *et al.* [16] introduced a latency corrected ensemble average (LCEA) method to evaluate the symmetry and variability of EMG signals. Sant'Anna and Wickström [17] recently suggested a symbolic method for estimating gait symmetry.

Although *discrete* methods have been proved very useful, *continuous* methods are potentially more informative. While the former may only consider isolated spatiotemporal values, the latter can compare whole movement sequences. To exemplify this difference, consider the following thought experiment. Imagine that one evaluates a gait pattern only by looking at the footprints a subject creates over time. These footprints convey spatiotemporal information that is typically used in *discrete* symmetry measures. It is possible, however, that two people with different walking patterns create the exact same sequence of footprints. The only way to differentiate between two such walking patterns would be to observe what happens above ground, in between footprints. This is the kind of information one would like to use in *continuous* symmetry measures.

The method presented here is a *continuous* method based on previous work by Sant'Anna and Wickström [17]. The six other methods chosen for comparison are:

1) symmetry index ($\mathrm{SI}_{\mathrm{index}}$) [4];
2) symmetry angle ($\mathrm{SI}_{\mathrm{angle}}$) [7];
3) gait asymmetry ($\mathrm{SI}_{\mathrm{GA}}$) [11];
4) maximum angular velocity ratio ($\mathrm{SI}_{\mathrm{ratio}}$) [18];
5) trend symmetry ($\mathrm{SI}_{\mathrm{trend}}$) [15];
6) LCEA symmetry magnitude ($\mathrm{SI}_{\mathrm{LCEA}}$) [16].

Method 1 is the most commonly used, e.g., [4], [7], [15], [19]. Methods 2 and 3 were among the *discrete* methods investigated by Patterson *et al.* [20]. These measures, calculated for swing time and stance time, were able to detect statistically significant differences between stroke patients and the healthy control group. Method 4 is a variant of method 1 that may be used to measure the symmetry of both lower and upper limbs. Methods 5 and 6 were chosen among a small number of *continuous* methods in the literature. They were the only *continuous* methods that could be implemented using the data available in the present study.

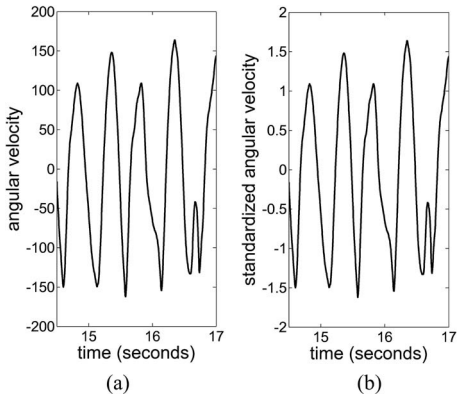

Fig. 1. Standardization. Consider an original signal, S_0, with mean μ_0, and standard deviation σ_0. The original signal is standardized to zero mean and unitary standard deviation according to the formula: $S = (S_0 - \mu_0)/\sigma_0$. (a) shows an example of original signal, and (b) illustrates the same signal after standardization.

III. Method

A. Data Analysis

The proposed symbolic symmetry index, $\mathrm{SI}_{\mathrm{symb}}$, is based on the concept of motion language [17], [21]. The present method, however, differs from [17] in symbolization method and type of sensors used. The main characteristic of this method is that the continuous signal is segmented and transformed into a string composed of a finite number of symbols. One of the advantages of using symbols instead of the raw signal is that it makes the analysis more robust to measurement noise and to small data variations which are not useful to the intended analysis. Another interesting property of this method is that it incorporates information about the shape and temporal characteristics of the signal. In addition, no gait events such as heel strike or toe-off need to be calculated, eliminating a potential source of error. On the other hand, symbolic methods are greatly influenced by the choice of symbols. Inadequate symbols may overlook important information or exacerbate the influence of irrelevant factors.

1) Symbolization: In order to calculate the novel symmetry index, the data must be symbolized. In order to achieve this, the original signal, Fig. 1(a), is first standardized to zero mean, $\mu = 0$, and standard deviation one, $\sigma = 1$, Fig. 1(b). This standardization is needed to ensure that all signals are within the same range, which facilitates segmentation. This standardization does not compromise the analysis of the data because the symbolic symmetry index aims to compare the shape of two signals, not their absolute values. It is unlikely that two signals will have the same shape and different absolute values. When the subject is walking in a straight line, the average speed of both hands must coincide with the average walking speed. Therefore, a faster swing on one side needs to be compensated by a different pattern of motion.

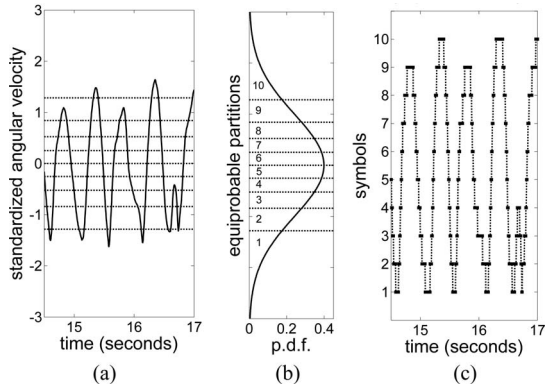

Fig. 2. Symbolization. The continuous signal is segmented into 10 equiprobable quantization levels (deciles). (a) shows an example of standardized signal. The horizontal lines correspond to the quantization levels based on a normal probability density function (p.d.f) with zero mean and unit standard deviation, as shown in (b). (c) illustrates the same signal after symbolization. Each quantization level corresponds to a symbol.

The standardized signal, Fig. 2(a), is then partitioned into ten different quantization values. The quantization values are based on deciles, i.e., ten equiprobable partitions of a normal distribution, Fig. 2(b). This segmentation technique is known as symbolic aggregate approximation (SAX) [22]. This segmentation method is not limited to ten partitions, different numbers of partitions could be used. For this analysis, alternative partitions were investigated (see Section VI). The results obtained with ten partitions were most satisfactory. All values between two consecutive partitions are then joined into a segment and represented by a symbol, Fig. 2(c). The segment's mid-point is used to represent when that symbol segment occurs in time. For simplicity, the symbols are numbers from 1 to 10.

2) Symbolic Symmetry Index: Once the signal has been symbolized, Fig. 3(a), the symmetry index is calculated as follows. For each symbol, i, construct a *period histogram* by finding the period between consecutive segments of the same symbol, Fig. 3(b). These periods are sorted into intervals of 0.1 s each and organized as a histogram, Fig. 3(c). Intervals of 0.1 s were chosen based on previously observed frequency characteristics of human movements [23]. It has been shown that 99% of the acceleration power during walking is concentrated below 15 Hz, and that the lower frequencies are the ones that directly result from voluntary muscular work. Considering that rotational data typically contains lower frequencies than acceleration data, a period of 0.1 s (10 Hz) was judged sufficient to capture relevant movement information.

The period histogram is then normalized by its sum, Fig. 3(d), so as to consider the relative distribution of symbol periods, disregarding the number of steps recorded. Period histograms are calculated for both the right (h_{Ri}) and left (h_{Li}) limbs, for each symbol.

The symbolic symmetry index ($\mathrm{SI}_{\mathrm{symb}}$) is a comparison between the histograms of the right and left signals:

$$\mathrm{SI}_{\mathrm{symb}} = \frac{\sum_{i=1}^{Z} \frac{1}{n_i} \sum_{k=1}^{K} |h_{Ri}(k) - h_{Li}(k)|}{\sum_{i=1}^{Z} \frac{1}{n_i} \sum_{k=1}^{K} |h_{Ri}(k) + h_{Li}(k)|} 100 \qquad (1)$$

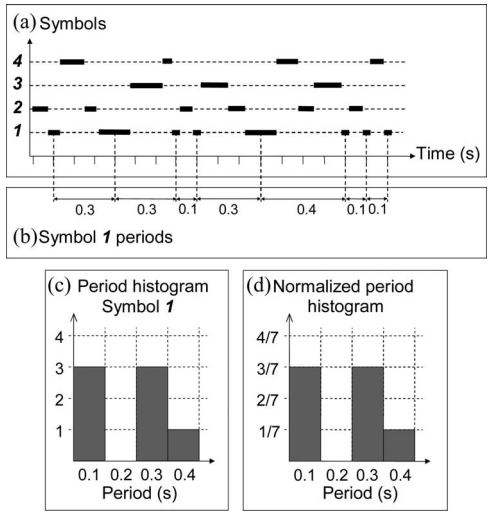

Fig. 3. Period Histograms. Period histograms are formed by measuring the time between two consecutive segments of the same symbol. (a) is a simple example of how symbols may occur over time. (b) illustrates how the periods of symbol *1* are calculated based on their occurrence over time. (c) shows how the histogram is constructed from the calculated periods. (d) shows the period histogram normalized by its sum.

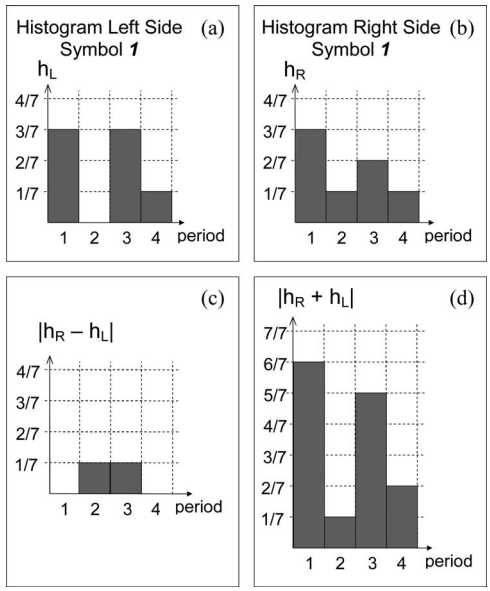

Fig. 4. Example. Comparing period histograms for one symbol. (a) and (b) show simple examples of possible period histograms. (c) is the absolute difference between the histograms in (a) and (b). (c) shows the sum of the histograms in (a) and (b).

where $Z = 10$ is the number of symbols, $K \leq 10(MAX_{\mathrm{period}})$ is the number of bins in the histograms and MAX_{period} is the largest symbol period; n_i is the number of non-empty histogram bins (for either foot) for symbol i, $h_{Ri}(k)$ is the normalized value for bin k in the period histogram i for the *right* foot, and $h_{Li}(k)$ is the normalized value for bin k in the period histogram i for the *left* foot.

The histograms' difference (numerator) and the histograms' sum (denominator) for each symbol are multiplied by a weight inversely proportional to the number of bins in the histogram.

This weight penalizes the symbols with very variable periods. The $\mathrm{SI}_{\mathrm{symb}}$ is then the sum of weighted histograms' differences divided by the sum of weighted histograms' sums. If the left and right histograms are the same, $\mathrm{SI}_{\mathrm{symb}} = 0$ and the signals are considered symmetric. If the left and right histograms have nothing in common, the absolute difference equals the sum and $\mathrm{SI}_{\mathrm{symb}} = 100$.

Fig. 4 illustrates the calculus of $\mathrm{SI}_{\mathrm{symb}}$ for one particular symbol, where $K = 4$, $i = 1$, $n_i = 4$. Fig. 4(c) shows the absolute difference of the histograms in 4(a) and 4(b), and corresponds to the numerator of (1), e.g., $\sum_{k=1}^{4} |h_{R1}(k) - h_{L1}(k)| = 2/7$. Fig. 4(d) is the sum of the histograms in 4(a) and 4(b) and corresponds to the denominator of (1), e.g., $\sum_{k=1}^{4} |h_{R1}(k) + h_{L1}(k)| = 14/7 = 2$. This process is repeated for all ten symbols. The results are weighted and summed, and the quotient is taken.

B. Reference Symmetry Indices

The reference methods were implemented as follows:

1) Symmetry index [4]:

$$\mathrm{SI}_{\mathrm{index}} = 100 \cdot \frac{|X_L - X_R|}{\max(X_L, X_R)}$$

2) Gait asymmetry [11]:

$$\mathrm{SI}_{\mathrm{GA}} = 100 \cdot \left| \ln \left\{ \frac{\min(X_R, X_L)}{\max(X_R, X_L)} \right\} \right|$$

3) Symmetry angle [7]:

$$\mathrm{SI}_{\mathrm{angle}} = 100 \cdot \frac{45 - \arctan\{X_L / X_R\}}{90}$$

where X_R is the variable from the right side, averaged over all cycles, and X_L the respective variable from the left side.

4) Maximum angular velocity ratio [18]:

$$\mathrm{SI}_{\mathrm{ratio}} = 100 \cdot \frac{|\mathbf{MA}_L - \mathbf{MA}_R|}{\max(\mathbf{MA}_L, \mathbf{MA}_R)}$$

where $\mathbf{MA}_{R(L)}$ is the maximum angular velocity during a cycle, averaged over all cycles.

5) Trend symmetry [15]:

$$\mathrm{SI}_{\mathrm{trend}} = 100 \cdot \mathrm{mean}_n \left\{ \frac{\mathrm{var}(\boldsymbol{X}\mathbf{rot}_L^n)}{\mathrm{var}(\boldsymbol{X}\mathbf{rot}_R^n)} \right\}$$

where

$$\boldsymbol{X}\mathbf{rot}_R^n = \boldsymbol{X}\boldsymbol{m}_R^n \cos(\theta) + \boldsymbol{X}\boldsymbol{m}_L^n \sin(\theta),$$
$$\boldsymbol{X}\mathbf{rot}_L^n = -\boldsymbol{X}\boldsymbol{m}_R^n \sin(\theta) + \boldsymbol{X}\boldsymbol{m}_L^n \cos(\theta)$$

θ is the angle between the first eigenvector of $\boldsymbol{M} = [\boldsymbol{X}\boldsymbol{m}_R^n, \boldsymbol{X}\boldsymbol{m}_L^n]$ and the horizontal axis, $\boldsymbol{X}\boldsymbol{m}_{R(L)}^n = \boldsymbol{X}_{R(L)}^n - \mathrm{mean}(\boldsymbol{X}_{R(L)}^n)$, and $\boldsymbol{X}_{R(L)}^n$ is the signal from the right (left) side for cycle n.

6) LCEA symmetry magnitude [16]:

$$\mathrm{SI}_{\mathrm{LCEA}} = 100 \cdot \max\left\{ \frac{\rho_{RL}}{\sqrt{\rho_{\mathrm{RR}} \cdot \rho_{\mathrm{LL}}}} \right\}$$

where ρ_{RL} is the cross-correlation between LCEA_R and LCEA_L, $\rho_{\mathrm{RR(LL)}}$ is the autocorrelation of $\mathrm{LCEA}_{R(L)}$, and $\mathrm{LCEA}_{R(L)}$ is the column-wise average of an $L \times N$ matrix $S_{R(L)}$. Each row of $S_{R(L)}$ contains the signal for one cycle of the data, normalized to L samples. N is the total number of cycles in the data set for the right (left) side. Each row of $S_{R(L)}$ is also time shifted in order to synchronize the cycles as best as possible.

C. Statistical Analysis

In the present study, the clinical evaluation did not provide movement asymmetry information. As a result, no reference data are available. In order to investigate the sensitivity and reliability of each index, a classification task was performed. The motivation for this classification based on symmetry indices is the assumption that the initial symptoms of PD set asymmetrically, affecting one side more than the other [24]. Therefore, we expect PD patients to exhibit greater asymmetry than control subjects. The following statistical analyses were performed on all symmetry indices. All hypotheses were nondirectional and the critical significance level was 0.05.

Two-Sample t-Test: A two-sample *t*-test was used to determine if the symmetry measures distinguished controls' and patients' results as originating from two different distributions.

Receiver Operating Characteristic (ROC) Curve: The area under the ROC curve (AUC) was used to evaluate the discriminatory power of each index. The ROC curves were constructed based on tests performed on the same individuals. Therefore, any statistically significant comparison between different AUC must take into account the correlated nature of the data. A nonparametric approach based on generalized U-statistics was used to estimate the covariance matrix of the different curves [25].

Intraclass Correlation Coefficient (ICC): Eight controls and eight PD patients were measured twice. This data was used to assess the test-retest reliability of each index using ICC type A-1 as a measure of absolute agreement [26].

IV. Experimental Setup

A. Subjects

The data used in the present study were derived from the study by Zampieri *et al.* [18]. The control group consisted of 15 healthy subjects, aged 61.4 ± 7.8 years, and weighing 81.4 ± 20.5 kg. The patient group was composed of 11 subjects with idiopathic PD, aged 60.1 ± 8.6 years, and weighing 79.7 ± 13.5 kg. The differences in age and weight of the groups were not statistically significant. The selected PD patients had been diagnosed 13.7 ± 12.9 months earlier and had never been treated with antiParkinson medications. They were evaluated by a movement-disorder neurologist according to the unified

TABLE I
DATA SUMMARY—CONTROL SUBJECTS

Hand sensor data – control subjects		
Measurement	**Left side**	**Right side**
Cycle [ms]	986 ± 82	982 ± 129

Shank sensor data – control subjects		
Measurement	**Left side**	**Right side**
Stance[1] [%]	58.0 ± 6.7	58.6 ± 6.6
Swing[1] [%]	40.0 ± 6.8	40.8 ± 2.6
Double support[1] [%]	17.7 ± 5.1	
Stride length[2] [%]	70.6 ± 10.2	77.6 ± 10.2
Stride time [ms]	980 ± 53	980 ± 53

Summary of spatiotemporal measurements for all control subjects. The shank sensor data were processed, as described in [8]. The hand sensor cycle time was determined, as described in Section IV-D. The values are reported as: mean value ± standard deviation. ms: milliseconds.

[1]Normalized to stride time.

[2]Normalized to subject's height.

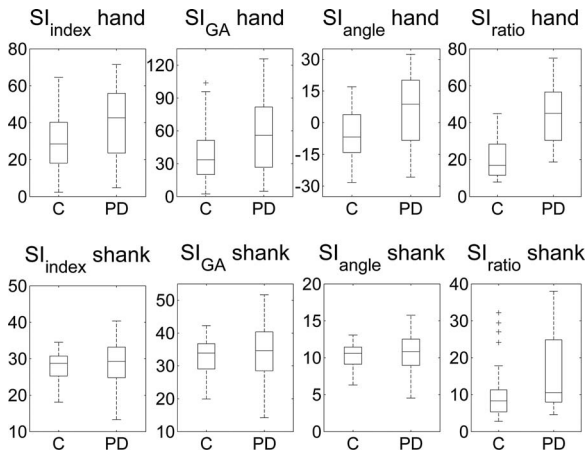

Fig. 5. *Discrete* symmetry indices—distributions. Results for all *discrete* methods are shown as box-plots. The whiskers represent the smallest and largest observations, the edges of the box correspond to the lower and upper quartiles, the horizontal line indicates the median, the plus sign marks probable outliers. C: controls, PD: patients, hand: hand sensor data, shank: shank sensor data.

TABLE II
DISCRETE SYMMETRY INDICES—SUMMARY

Hand sensor data – PD patients		
Measurement	**Left side**	**Right side**
Cycle [ms]	1034 ± 148	1053 ± 113

Shank sensor data – PD patients		
Measurement	**Left side**	**Right side**
Stance[1] [%]	57.4 ± 6.6	57.4 ± 6.6
Swing[1] [%]	40.6 ± 6.9	41.9 ± 2.2
Double support[1] [%]	16.0 ± 4.1	
Stride length[2] [%]	75.4 ± 11.2	77.4 ± 11.2
Stride time [ms]	1051 ± 73	1051 ± 73

[1]Normalized to stride time.

[2]Normalized to subject's height.

Summary of spatiotemporal measurements for all PD The shank sensor data was processed as described in [8]. The hand sensor cycle time was determined as described in Section IV-D. The values are reported as: mean value ± standard deviation. ms: milliseconds.

TABLE III
DATA SUMMARY—PD PATIENTS

Hand sensor data		
Index	**Control**	**PD**
SI_{index}	32.38 ± 17.98	45.29 ± 22.00
SI_{GA}	43.24 ± 31.59	67.85 ± 41.02
SI_{angle}	-1.68 ± 15.45	5.70 ± 21.71
SI_{ratio}	23.62 ± 16.27	40.55 ± 19.78

Shank sensor data		
Index	**Control**	**PD**
SI_{index}	26.58 ± 4.74	28.95 ± 5.41
SI_{GA}	31.09 ± 6.44	34.43 ± 7.55
SI_{angle}	9.72 ± 1.95	10.72 ± 2.27
SI_{ratio}	12.77 ± 9.36	13.29 ± 7.70

Results for all *discrete* methods are shown as: mean value ± standard deviation. The top part of the table reports the results calculated from hand sensor data whereas the bottom part reports the results from shank sensor data.

Parkinson's disease rating scale (UPDRS), motor section [27], and Hoehn and the Yahr (H&Y) scale [28].

Clinical evaluation of the patients' gait according to the motor UPDRS showed minimal gait deficiency, the postural instability and gait disorder (PIGD) subscores was 0.7 ± 1.1 on a scale of 0 to 16 where 0 means no impairment. Though they showed minimal gait abnormalities, they might have had other motor dysfunctions such as slight hand tremors. The H&Y stages of the patient group were 1.6 ± 0.6.

Subjects were excluded if they presented any neurological disorder other than PD, or other conditions which could interfere with gait. All participants signed informed consent forms approved by the Oregon Health & Science University Internal Review Board.

B. Apparatus

During the measurements, subjects wore a portable datalogger (Physilog from BioAGM, Switzerland) with four inertial sensors attached to their body. A 1-D gyroscope (range 600°/s) was attached to the anterior shank of each limb, 4 cm above the ankle joint. A 2-D gyroscope (range ±1200°/s) along pitch and roll axes was attached to the dorsum of each wrist. Data were recorded at a sampling rate of 200 Hz with 16 bits/sample and stored in a flash memory card.

C. Protocol

The subjects were equipped with the sensors and walked at their preferred speed along a straight 30-meter hallway for two minutes. At the end of the hallway the subject turned around and walked in the opposite direction. Typically, each subject turned four times. The data were processed in order to remove the turns and the five initial and final cycles, so as to consider only steady-state, straight walking patterns. This is necessary because the turn data are asymmetric in nature and would cause a bigger effect on the results than the occasional boundary irregularities that arise from creating gaps in the data. To assess test-retest reliability, eight subjects in each group repeated the same protocol a second time after 1 h.

D. Data Analysis

$\mathrm{SI}_{\mathrm{index}}$, $\mathrm{SI}_{\mathrm{GA}}$, and $\mathrm{SI}_{\mathrm{angle}}$ were calculated using temporal variables derived from both hand and shank sensor data. The variable chosen for the hand was mean cycle time, i.e., the time between maximum swing velocity front-to-back from one cycle

TABLE IV
CONTINUOUS SYMMETRY INDICES—SUMMARY

Hand sensor data		
Index	**Control**	**PD**
SI_{trend}	4.68 ± 3.56	9.87 ± 11.49
100 - SI_{LCEA}	48.88 ± 8.15	45.91 ± 10.03
SI_{symb}	23.61 ± 14.12	38.74 ± 11.61

Shank sensor data		
Index	**Control**	**PD**
SI_{trend}	1.51 ± 0.90	1.73 ± 1.17
100 - SI_{LCEA}	0.76 ± 0.51	1.48 ± 1.51
SI_{symb}	16.60 ± 5.26	17.85 ± 5.78

Results for all *continuous* methods are shown as: mean value ± standard deviation. The top part of the table reports the results calculated from hand sensor data, whereas the bottom part reports the results from shank sensor data.

TABLE V
TWO-SAMPLE t-TEST

	Hand sensor data		Shank sensor data	
Index	**Reject H_0**	**p-value**	**Reject H_0**	**p-value**
SI_{index}	YES	0.0431	NO	0.2946
SI_{GA}	YES	0.0260	NO	0.2819
SI_{angle}	NO	0.1690	NO	0.2860
SI_{ratio}	YES	0.0017	NO	0.8151
SI_{trend}	NO	0.1225	NO	0.6391
SI_{LCEA}	NO	0.4313	NO	0.1028
SI_{symb}	YES	<0.0001	NO	0.4060

T-test results for the indices calculated on hand and shank sensor data. H_0: results from PD and control subjects are consistent with one single distribution.

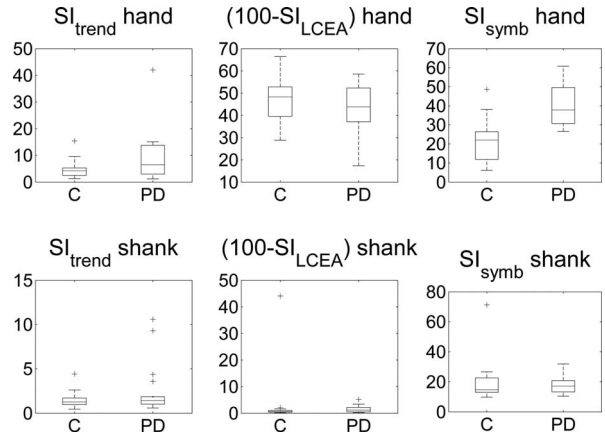

Fig. 6. *Continuous* symmetry indices—distributions. Results for all *continuous* methods are shown as box-plots. C: controls, PD: patients, hand: hand sensor data, shank: shank sensor data.

TABLE VI
AREA UNDER ROC CURVE (AUC)

Hand sensor data			
Index	**AUC**	**95% C.I.**	**p-value**
SI_{index}	0.673	[0.445, 0.902]	0.1368
SI_{GA}	0.673	[0.445, 0.902]	0.1368
SI_{ratio}	0.798	[0.629, 0.968]	<0.0001
SI_{symb}	0.872	[0.720, 1.000]	<0.0001

AUC results for the indices that rejected the two-sample H_0. The C.I. lower limits for SI_{index} and SI_{GA} are smaller than 0.5; these indices are not statistically significantly greater than 0.5. C.I.: confidence interval.

to the next. The variable chosen for the lower limb data was swing time, one of the measures used by Patterson *et al.* [20] that was available in our dataset. $\mathrm{SI_{ratio}}$, $\mathrm{SI_{trend,}}$ and $\mathrm{SI_{LCEA}}$ were calculated from both hand and shank sensor data. The cycles were determined according to strides of the corresponding foot [8]. $\mathrm{SI_{symb}}$ was calculated from hand and shank data as explained in Section III. Because the shank sensors are 1-D and the hand sensors are 2-D, the two axes of the hand data were combined into a resultant signal, $\mathrm{res} = \sqrt{\mathrm{axis}_1^2 + \mathrm{axis}_2^2}$. The number of strides used in this analysis varied across subjects, the average number of steps was 89.7 ± 9.4. Overall, eight *discrete* and six *continuous* measures were computed for each subject. All the acquired data were considered for the sensitivity and specificity analysis (t-test and ROC) of the symmetry indices. Only the data from the eight subjects who repeated the test were used for the test-retest reliability analysis (ICC). All data analysis was undertaken in MATLAB (MathWorks, Natick, MA).

V. RESULTS

A summary of the variables extracted from the data is shown in Tables I and III. There are no clear differences between control subjects and PD patients besides hand cycle and stride time. For both groups hand cycle duration is consistent with stride time, and stance and swing are approximately 60% and 40% of stride time, respectively.

A summary of the *discrete* symmetry indices is shown in Table II, alongside their distributions in Fig. 5. Regarding the shank sensor data, the spread of the distributions is so large compared to the differences in mean values that it is difficult to differentiate between control and PD. The indices calculated from the hand sensor data present a similarly large spread, but the greater differences in mean values make these better candidates, with the exception of $\mathrm{SI_{angle}}$. These observations are corroborated by the results of the two-sample t-test shown in Table V.

Similarly, a summary of the *continuous* symmetry indices is shown in Table IV. The corresponding distributions are illustrated in Fig. 6. Although the distributions of the shank data are much narrower than before, the differences between mean values is also much smaller. The only *continous* index that appears as a good candidate is $\mathrm{SI_{symb}}$, as evidenced by the two-sample t-test reported in Table V.

In order to further investigate the four indices that successfully rejected the t-test H_0, the area under ROC curve (AUC) was calculated. Table VI shows the the resulting AUC along with the corresponding 95% confidence intervals (CI). An AUC of 0.5 indicates an inability to differentiate between control and PD. Both $\mathrm{SI_{index}}$ and $\mathrm{SI_{GA}}$ present CI lower limit smaller than 0.5. These areas are, therefore, not statistically significantly greater than 0.5, as evidenced by the corresponding p-values. On the other hand, $\mathrm{SI_{ratio}}$ and $\mathrm{SI_{symb}}$ present favourable AUC.

The ROC curves for these two indices are illustrated in Fig. 7. $\mathrm{SI_{symb}}$ presents an AUC greater than $\mathrm{SI_{ratio}}$. A χ^2-test was performed to judge if this difference is statistically relevant. The test resulted in a χ^2 coefficient of 1.086 and a p-value of 0.297, which indicates that there is a 30% chance that the difference between the areas is due to chance alone.

In addition to AUC, an ICC analysis was used to judge the test-retest reliability of $\mathrm{SI_{ratio}}$ and $\mathrm{SI_{symb}}$. The ICC results reported

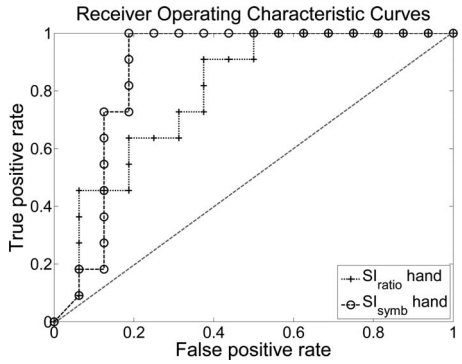

Fig. 7. ROC curves. ROC curves for methods showing AUC statistically greater than 0.5. An ROC curve which coincides with the diagonal line indicates a test that cannot distinguish between control and patients.

TABLE VII
INTRACLASS CORRELATION COEFFICIENTS (ICC).

Hand sensor data			
Index	**ICC**	**95% C.I.**	**p-value**
SI_{ratio}	0.611	[0.178, 0.849]	0.0049
SI_{symb}	0.949	[0.859, 0.982]	<0.0001

ICC results for the three indices with AUC statistically greater than 0.5. H_0: intra-subject variance is similar to inter-subject variance. C.I.: confidence interval.

in Table VII show that $\mathrm{SI}_{\mathrm{symb}}$ is considerably more reliable than $\mathrm{SI}_{\mathrm{ratio}}$. In fact, the upper CI limit for $\mathrm{SI}_{\mathrm{ratio}}$ is smaller than the lower CI limit for $\mathrm{SI}_{\mathrm{symb}}$.

VI. DISCUSSION

The objective of the present analysis was to evaluate which symmetry indices were able to differentiate between control subjects and early-to-mid-stage PD patients. Out of all seven symmetry indices calculated on upper and lower limb data, four were able to pick up on differences between PD and control groups according to a two-sample *t*-test. An ROC analysis indicated that, out of these four, only $\mathrm{SI}_{\mathrm{ratio}}$ and $\mathrm{SI}_{\mathrm{symb}}$ were truly able to classify patients and controls. Between these two methods, $\mathrm{SI}_{\mathrm{symb}}$ showed better performance according to AUC and ICC results.

To better understand how the $\mathrm{SI}_{\mathrm{symb}}$ works, consider some of the differences between the control and PD data illustrated in Fig. 8. In this example, the most striking dissimilarities present in the PD data are the higher frequency content and the lower periodicity of the left hand signal. To understand how these characteristics are transformed by the symbolic analysis, consider the normalized symbol histograms shown in Fig. 9. These histograms were computed from symbol 1, which is the symbolic representation of the data points that fall below the dotted line shown in Fig. 8. The higher frequency with which the left-hand PD signal crosses the line causes the skewness of the symbol histogram to the left. In addition, the lower periodicity of the signal flattens the distribution causing the main frequency of the signal to be less apparent. These characteristics of the symbolic histogram transformation are the basis for the symbolic symmetry index.

Results show that the hand data are, in general, more asymmetric than lower limb data, even among healthy subjects. This

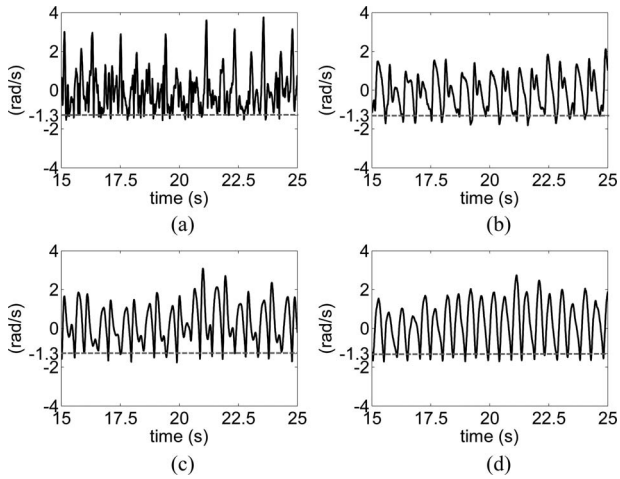

Fig. 8. Signal comparison. (a) and (b) show the resultant hand signals for a patient. (c) and (d) show the resultant hand signals for a control subject. The dotted gray line indicates the limit of the first symbol level (symbol 1).

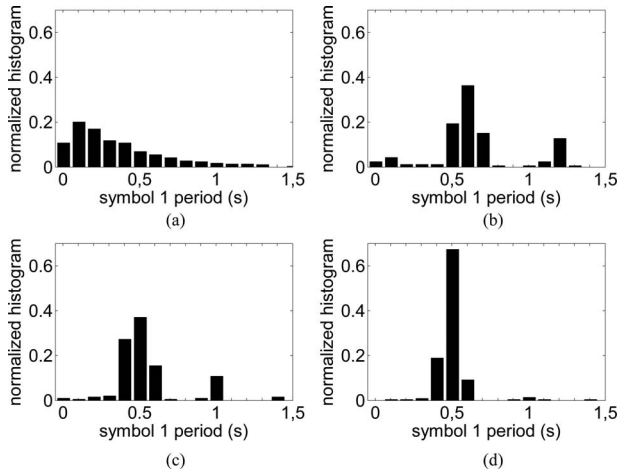

Fig. 9. Histogram comparison. (a) and (b) show the period histograms of symbol 1 for the patient signals shown in Figs. 8(a) and 8(b) respectively. (c) and (d) show the period histograms of the same symbol for the control signals shown in Figs. 8(c) and 8(d) respectively.

decoupling between upper and lower limb symmetry is not exclusive to PD patients. A study by Kuhtz-Buschbeck *et al.* [9] determined that arm swing asymmetry is common in healthy subjects, and cannot be explained by lower limb kinematics. This study extracted arm swing amplitude and step length from 3-D motion analysis data, among other variables, and utilized $\mathrm{SI}_{\mathrm{index}}$ to determine upper and lower limb symmetry. The study determined that the asymmetry of upper limb variables is considerably greater than that of lower limb variables.

The analysis presented here stems from the assumption that PD patients present more asymmetric arm movements than the control group. Indeed, other studies have determined that reduced arm swing is often one of the first motor symptoms of the disease [29], [30]. More recently a quantitative study by Lewek *et al.* investigated the symmetry of lower and upper limbs during walking in PD [10]. This work used $\mathrm{SI}_{\mathrm{angle}}$ on variables extracted from 3D motion analysis data. Lewek *et al.* also found that lower limb symmetry was not significantly different in PD

TABLE VIII
COMPARING $\mathrm{SI}_{\mathrm{symb}}$ EQUATIONS

Hand sensor data		
Equation	**ICC**	**95% C.I.**
Eq. 1	0.949	[0.859, 0.982]
Eq. 2	0.921	[0.788, 0.972]
Eq. 3	0.928	[0.807, 0.975]

ICC coefficients for each of the three alternative equations.

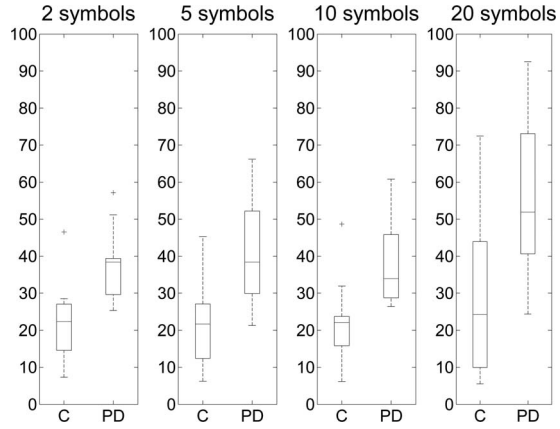

Fig. 10. Dependency on number of symbols. Results of the $\mathrm{SI}_{\mathrm{symb}}$ calculated on hand sensor data considering different numbers of symbols.

patients, but arm swing symmetry was. The lower limb $\mathrm{SI}_{\mathrm{angle}}$ results reported in [10] are slightly smaller than the ones presented here. This can be due to the fact that their choice of variable was stride time, instead of swing time. Another difference is that Lewek *et al.* found the arm swing $\mathrm{SI}_{\mathrm{angle}}$ significantly different for the PD group. This can be attributed to the fact that they used arm swing distance as the observed variable instead of arm swing cycle time. Another influencing factor is that features extracted from 3-D motion data are slightly more accurate than features extracted from inertial sensor data. On the other hand, inertial sensors can be used continuously and in uncontrolled environments, e.g. at home.

As mentioned in Section III-A, the proposed symbolic analysis of gyroscope data presents a few issues. The first issue relates to 1. This equation compares the left and right period histograms, however, there are many alternatives which could serve the same purpose. The following equations were also investigated:

$$\mathrm{SI} = \frac{\sum_{i=1}^{Z}\sum_{k=1}^{K}|h_{Ri}(k) - h_{Li}(k)|}{\sum_{i=1}^{Z}\sum_{k=1}^{K}|h_{Ri}(k) + h_{Li}(k)|}100 \tag{2}$$

$$\mathrm{SI} = \frac{1}{Z}\sum_{i=1}^{Z}\frac{\sum_{k=1}^{K}|h_{Ri}(k) - h_{Li}(k)|}{\sum_{k=1}^{K}|h_{Ri}(k) + h_{Li}(k)|}100. \tag{3}$$

All three equations were compared in terms of ICC and AUC. All methods result in the same ROC curve and ICC results are very similar, Table VIII.

Another issue is that the number of symbols and the symbolization technique used can have a considerable effect on the results. Too few symbols would result in a too coarse measurement. Too many would cause the measures to be too sensitive to noise or unimportant details. Therefore, the number of symbols must be optimized to the intended application. Fig. 10 shows the $\mathrm{SI}_{\mathrm{symb}}$ results considering different numbers of symbols. In all cases, the quantization levels were chosen to match equiprobable partitions of a normal distribution. The distributions created by using ten partitions was one of the most satisfactory based on ROC analysis and ICC results.

Fig. 5 and Fig. 6 show some data points as outliers. No subjects were considered outliers based on all methods. Therefore, no data points were excluded from the analysis.

VII. CONCLUSION

The present study introduced a new symbol based movement symmetry index and compared it to six other symmetry measures. The methods were evaluated on data from early-to-mid-stage PD patients who showed minimal motor abnormalities and practically no postural instability according to the UPDRS. Sensitivity and test-retest reliability were computed for each method, on both upper limb and lower limb data.

Results illustrate how important it is to chose the most appropriate method given a particular data set. The method proposed here is sensitive enough to detect movement asymmetry in early-to-mid-stage PD patients. Two of the evaluated methods performed significantly better than the other five. Out of these two, the proposed method showed superior sensitivity and test-retest reliability. Results also show that for this group of patients, movement asymmetry is more easily observed in upper limb data. It is therefore important to employ a symmetry index which can be applied not only to gait spatiotemporal parameters but also to upper limb data.

ACKNOWLEDGMENT

The authors thank Prof. Fay B. Horak and Prof. Kamiar Aminian who kindly shared with us the dataset used in the present study.

REFERENCES

[1] G. Yogev, M. Plotnik, C. Peretz, N. Giladi, and J. Hausdorff, "Gait asymmetry in patients with Parkinson's disease and elderly fallers: When does the bilateral coordination of gait require attention?," *Exp. Brain Res.*, vol. 177, pp. 336–346, 2007.

[2] R. Baltadjieva, N. Giladi, L. Gruendlinger, C. Peretz, and J. M. Hausdorff, "Marked alterations in the gait timing and rhythmicity of patients with *de novo* Parkinson's disease," *Eur. J. Neurosci.*, vol. 24, no. 6, pp. 1815–1820, 2006.

[3] C. Isenberg and B. Conrad, "Kinematic properties of slow arm movements in Parkinson's disease," *J. Neurol.*, vol. 241, pp. 323–330, 1994.

[4] R. Robinson, W. Herzog, and B. Nigg, "Use of force platform variables to quantify the effects of chiropractic manipulation on gait symmetry," *J. Manipulative Physiol. Therapeut.*, vol. 10, pp. 172–176, 1987.

[5] T. Karaharju-Huisnan, S. Taylor, R. Begg, J. Cai, and R. Best, "Gait symmetry quantification during treadmill walking," in *Proc. 7th Australian New Zealand Intell. Inf. Syst. Conf.*, 2001, pp. 203–206.

[6] G. Vagenas and B. Hoshizaki, "A multivariable analysis of lower extremity kinematic asymmetry in running," *J. Appl. Biomech.*, vol. 8, no. 1, pp. 11–29, 1992.

[7] R. Zifchock, I. Davis, J. Higginson, and T. Royer, "The symmetry angle: A novel, robust method of quantifying asymmetry," *Gait Posture*, vol. 27, no. 4, pp. 622–627, 2008.

[8] A. Salarian, H. Russmann, F. Vingerhoets, C. Dehollain, Y. Blanc, P. Burkhard, and K. Aminian, "Gait assessment in Parkinson's disease: Toward an ambulatory system for long-term monitoring," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 8, pp. 1434–1443, Aug. 2004.

[9] J. P. Kuhtz-Buschbeck, K. Brockmann, R. Gilster, A. Koch, and H. Stolze, "Asymmetry of arm-swing not related to handedness," *Gait Posture*, vol. 27, no. 3, pp. 447–454, 2008.
[10] M. D. Lewek, R. Poole, J. Johnson, O. Halawa, and X. Huang, "Arm swing magnitude and asymmetry during gait in the early stages of Parkinson's disease," *Gait Posture*, vol. 31, no. 2, pp. 256–260, 2010.
[11] M. Plotnik, N. Giladi, Y. Balash, C. Peretz, and J. M. Hausdorff, "Is freezing of gait in Parkinson's disease related to asymmetric motor function?," *Ann. Neurol.*, vol. 57, no. 5, pp. 656–663, 2005.
[12] M. Plotnik, N. Giladi, and J. M. Hausdorff, "A new measure for quantifying the bilateral coordination of human gait: Effects of aging and Parkinsons disease," *Exp. Brain Res.*, vol. 181, pp. 561–570, 2007.
[13] J. Riad, S. Coleman, D. Lundh, and E. Broström, "Arm posture score and arm movement during walking: A comprehensive assessment in spastic hemiplegic cerebral palsy," *Gait Posture*, vol. 33, no. 1, pp. 48–53, 2011.
[14] R. Moe-Nilssen and J. L. Helbostad, "Estimation of gait cycle characteristics by trunk accelerometry," *J. Biomech.*, vol. 37, no. 1, pp. 121–126, 2004.
[15] S. J. Crenshaw and J. G. Richards, "A method for analyzing joint symmetry and normalcy, with an application to analyzing gait," *Gait Posture*, vol. 24, no. 4, pp. 515–521, 2006.
[16] R. A. Miller, M. H. Thaut, G. C. McIntosh, and R. R. Rice, "Components of EMG symmetry and variability in parkinsonian and healthy elderly gait," *Electroen. Clin. Neuro./Electromyogr. Motor C.*, vol. 101, no. 1, pp. 1–7, 1996.
[17] A. Sant'Anna and N. Wickström, "A symbol-based approach to gait analysis from acceleration signals: Identification and detection of gait events and a new measure of gait symmetry," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 5, pp. 1180–1187, Sep. 2010.
[18] C. Zampieri, A. Salarian, P. Carlson-Kuhta, K. Aminian, J. G. Nutt, and F. B. Horak, "The instrumented timed up and go test: Potential outcome measure for disease modifying therapies in Parkinson's disease," *J. Neurol., Neurosurg. Psychiatry*, vol. 81, no. 2, pp. 171–176, 2010.
[19] H. Sadeghi, P. Allard, F. Prince, and H. Labelle, "Symmetry and limb dominance in able-bodied gait: A review," *Gait Posture*, vol. 12, no. 1, pp. 34–45, 2000.
[20] K. K. Patterson, W. H. Gage, D. Brooks, S. E. Black, and W. E. McIlroy, "Evaluation of gait symmetry after stroke: A comparison of current methods and recommendations for standardization," *Gait Posture*, vol. 31, no. 2, pp. 241–246, 2010.
[21] G. Guerra-Filho and Y. Aloimonos, "A language for human action," *Computer*, vol. 40, no. 5, pp. 42–51, 2007.
[22] J. Lin, E. Keogh, L. Wei, and S. Lonardi, "Experiencing SAX: A novel symbolic representation of time series," *Data Mining Knowl. Discov.*, vol. 15, pp. 107–144, 2007.
[23] C. Bouten, K. Koekkoek, M. Verduin, R. Kodde, and J. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 3, pp. 136–147, Mar. 1997.
[24] D. J. Gelb, E. Oliver, and S. Gilman, "Diagnostic criteria for parkinson's disease," *Arch. Neurol.*, vol. 56, no. 1, pp. 33–39, 1999.
[25] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, 1988.
[26] K. O. McGraw and S. P. Wong, "Forming inferences about some intraclass correlation coefficients," *Psychol. Methods*, vol. 1, pp. 30–46, 1996.
[27] S. Fahn and R. Elton, *Unified Parkinson's Disease Rating Scale*. Florham Park, NJ: Macmillan Healthcare Information, 1987.
[28] M. Hoehn and M. Yahr, "Parkinsonism: Onset, prognosis, and mortality," *Neurology*, vol. 17, pp. 427–442, 1967.
[29] F. Buchthal and M. Fernandez-Ballesteros, "Electromyographic study of the muscles of the upper arm and shoulder during walking in patients with Parkinson's disease," *Brain*, vol. 88, pp. 875–896, 1965.
[30] A. J. Lees, "When did Ray Kennedy's Parkinson's disease begin?," *Movement Disord.*, vol. 7, no. 2, pp. 110–116, 1992.

Anita Sant'Anna received the B.Sc. degree in electrical engineering, in 2007, from Universidade Federal de Santa Catarina, Florianopolis, Brazil (UFSC). She is currently a doctoral candidate at Halmstad University, Halmstad, Sweden (HH).

She has focused her research interests on biomedical signal processing, wearable sensors, and human movement analysis.

Arash Salarian (M'04) received the B.Sc. degree in computer engineering, in 1993, from Isfahan University of Technology, Isfahan, Iran (IUT), the M.Sc. degree in computer architecture, in 1997, from Shairf University of Technology, Tehran, Iran (SUT), and the Ph.D. degree in biomedical engineering, in 2006, from Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. He finished a post-doctoral training at EPFL and a second post-doctoral training at Oregon Health & Science University (OHSU), Portland, OR.

He is currently a Senior Research Associate in the Balance Disorders Laboratory at Oregon Health & Science University (OHSU), Portland, OR. His research interests include biomedical signal processing and wearable sensors and their application in human movement analysis.

Nicholas Wickström (M'11) was born in Uppsala, Sweden,in 1970. He received the Ph.D. degree in computer engineering from Chalmers University of-Technology, Gothenburg, Sweden, in 2004.

He is currently an Associate Professor of computer systems engineering at the School of Information Science, Computer and Electrical Engineering, Halmstad University, Sweden. His current research interests include biomedical sensing and signal processing.