2012 IEEE/SICE International Symposium on System Integration (SII)
Kyushu University, Fukuoka, Japan
December 16-18, 2012

Classification of Parkinson's Disease Patients' Gait Variability

Leo Ota[1], Hirotaka Uchitomi[1], Satoshi Orimo[2] and Yoshihiro Miyake[1]

Abstract— **Parkinson's disease (PD) is a neurodegenerative disorder. With progression of PD, movement disorder such as gait disturbance and balance impairment is frequently observed. Hoehn and Yahr scale (HY) is an indicator to evaluate the severity of motor signs of PD. Recently, objective measurement comes to be widely spread. Previous studies pointed out that human walking comes from complex interaction, and it comes to be seen as nonlinear dynamics. Amplitude of PD patients' stride time variability was reported to be larger than that of healthy people. Coefficient of variation (CV) is commonly used to see amplitude of variability. The fractal property of PD patients' stride time is lower than that of healthy people. The fractal property was measured by scaling exponent α calculated by detrended fluctuation analysis (DFA). However, the relationship between the stride time variability and the severity of motor signs of PD remains to be clarified. In this study, we tried to investigate the relationship between these indicators. Clarifying the relationship between practical severity index and objective data provides us information to make a diagnosis of PD. Forty-five PD patients walked 200 meter corridor at their preferred pace. As control group, 35 healthy people, which include young and elderly people, are participated. In order to separate between the PD patients and healthy people or to classify HY scale, linear discriminant analysis on both CV and DFA was applied. When we separated into all PD patients group and all healthy people group, the accuracy was 0.76. We tried to separate 2 groups. One is group of PD patients with HY2.5 or higher and the other is the group of healthy elderly people or mild PD patients. The specificity was 0.86. When we tried to separate between PD patients with HY3 and PD patients with HY2.5, the sensitivity was 0.93. We conclude that gait variability is one of the indicators of motor severity seen in PD.**

I. INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disorder. Due to strong depression of motor control and dysfunction of rhythm generation in basal ganglia, movement disorders such as tremor, akinesia, rigidity, and postural instability are typical motor symptoms of PD [1]. When PD progresses, postural instability or gait disturbances appears in many cases. When a patient has characteristic gait disturbance such as festination, short gait step and freezing gait, physicians can make a diagnosis of PD much easier. There are some severity index of PD [2], [3]. Particularly, Hoehn and Yahr (HY) scale is commonly used in order to describe the natural progression of PD [4]. The range of HY is from 1 to 5.

[1]L. Ota, H. Uchitomi and Y. Miyake are with the Department of Computational Intelligence and Systems Science, Tokyo Institute of Technology, Midori, Yokohama 226-8502, Japan ohta at myk.dis.titech.ac.jp, uchitomi at myk.dis.titech.ac.jp, miyake at dis.titech.ac.jp

[2]S. Orimo is with the Department of Neurology, Kanto Central Hospital, Setagaya, Tokyo 158-8531, Japan orimo at kanto-ctr-hsp.com

In recent years, objective measurement of human walking comes to be easy. And biological or physiological time series including gait data is thought to be come from nonlinear dynamics, which is composed from complex interactions. From perspective of nonlinear dynamics, variability of gait cycle is attracted. Magnitude of the stride time variability of PD patients was bigger than that of healthy people [5]. These evidences were provided by the magnitude of variability by means of standard deviation or coefficient of variation (CV).

In addition, the fractal property of the gait cycle fluctuation has been reduced by specific disease such as PD or Huntington's disease [6]–[8]. Detrended Fluctuation Analysis (DFA) is used to see fractal property. By DFA, we can analyze the fractal properties of non-stationary time series data [9], [10]. Based on these backgrounds, there is possibility of diagnosis of the severity of PD using CV or DFA.

Correlation with the severity index of PD and CV has been suggested [5]. However, there are some exceptions [8]. Therefore, to classify the severity of PD with only CV is difficult. On the other hand, correlation between the fractal property of the gait cycle fluctuation and the severity index of PD is also reported [8]. However fractal property of healthy elderly people is likely to be reduced compared to young people [6]. Therefore, classifying the gait variability into different age or severity of PD groups using only CV or using only DFA has problems.

In this study, we tried to analyze stride time variability comprehensively using both the DFA and the CV. The clarification of the relationship between practical severity index and objective gait data has a possibility to contribute to diagnosis of PD or remedy for PD patients. Our hypothesis is that we can clarify the difference in age or the severity of PD by combining magnitude of stride time variability and fractal property of stride time. In order to verify this hypothesis, we divided participants into healthy young group, healthy elderly people, mild PD patients groups, moderate PD patients and relatively severe PD patients. Then we compared each group in terms of both CV and DFA.

II. MATERIAL AND METHOD

A. Data Description

Forty-five patients (24 women, 21 men) with idiopathic PD participated in the experiment. Among them, 20 PD patients gait data were measured in a previous study [11]. We recruited the patients whose HY is 1, 1.5, 2, 2.5, 3 or 3.5, and they did not exhibit freezing gait. Patients were divided into 3 groups, HY1-2, HY2.5 and HY3-3.5. Table I shows description for these groups. Table II shows the explanation for the modified HY score [3], [12]. Common point of

TABLE I

DESCRIPTION OF 5 GROUPS (M:*Male*, F:*Female*)

Group	number(M:F)	Age (y/o)	Disease duration (yr)
HY1-2	19 (6:13)	66.2 ± 8.7	4.4 ± 2.9
HY2.5	11 (7:4)	72.4 ± 7.5	4.0 ± 2.8
HY3-3.5	15 (8:7)	74.2 ± 4.2	6.0 ± 5.6
Young	18 (16:2)	24.7 ± 2.7	(healthy control)
Elderly	17 (10:7)	70.2 ± 2.8	(healthy control)

TABLE II

MODIFIED HOEHN AND YAHR SCALE (HY) [3] [12]

score	Definition
HY1	Unilateral involvement only
HY1.5	Unilateral and axial involvement
HY2	Bilateral disease without balance impairment
HY2.5	Mild bilateral disease with recovery on pull test
HY3	Mild to moderate bilateral disorder Some postural instability but physically independent
HY3.5	Bilateral disorder with balance impairment Require an attendant participants or wheelchair in the street but physically independent inside the house [12]

HY1, HY1.5, HY2 is without balance impairment. HY2.5 is the intermediate level between HY2, and HY3. HY2.5 is defined as bilateral disease with recovery on pull test. The same point between HY3 and HY3.5 is bilateral disorder with balance impairment, but physically independent inside home at least [3]. HY3.5 is defined in the study by Araki et al. [12]. Group HY1-2 consisted of 19 patients whose HY stage was 1, 1.5 or 2 (mean age = 66.2 years; s.d. = 8.7 years), Group HY2.5 (n = 11, mean age = 74.2 years; s.d. = 7.5 years) and group HY3-3.5 consisted of 15 patients whose HY stage was 3 or 3.5 (mean age = 74.2 years; s.d. = 4.2 years; see Table I). Mean duration of disease of HY1-2 was 4.4 years (s.d. = 2.9 years), that of HY2.5 was 4.0 years (s.d. = 2.8 years) and that of HY3 was 6.0 years (s.d.=5.6years). All were tested while on anti-parkinsonian medication. Eighteen younger healthy controls (16 men, 2 woman, mean age = 24.7 years; s.d. = 2.7 years; measured in a previous study [11].) and elderly healthy controls (10 men, 7 woman, mean age = 70.2 years; s.d. = 2.8 years) also participated. Informed consent was provided and participants were paid for participating. Experimental procedures were approved by the Kanto Central Hospital Ethics Committee.

B. Task and Experimental setup

Participants were instructed to walk at a natural and comfortable pace around a long corridor. The length of the course was 200m. On average, each trial lasted about 3 minutes and contained approximately 320 footsteps. Foot step timing was collected via foot switches (OT-21BP-G, Ojiden, Japan) attached to participants' shoes, was relayed to a laptop (CF-W5, Panasonic, Japan) via radio frequency every 10 ms, and was processed in real time. Two transceivers (S-1019M1F, Smart Sensor Technology, Japan) and a receiver (WM-1019M1F, Smart Sensor Technology, Japan) were used.

Fig. 1. The picture of experimental space. During gait measurement, each participants walked with a few staffs. Each participants walked 200 meter corridor.

Fig. 2. Devices of experiment. Foot switches were in these devices, and all participants wore these devices while gait measurement.

Fig. 1 and fig. 2 shows a scene of the walking experiments and foot switches. The computer algorithm controlling the above experimental system was run on the laptop.

C. Data Analysis

The time series stride time were analyzed. The time series data are represented by $u(i)$ in the following (1),

$$u(i) = T(i+1) - T(i), \tag{1}$$

where $u(i)$ is the i-th stride interval, and $T(i)$ represents the i-th step timing.

1) Coefficient of Variation(CV): Fluctuation magnitude is evaluated by coefficient of variation (CV). This is standard deviation (s.d.) normalized by the mean value, as in (2).

$$CV = \frac{u_{s.d.}}{u_{ave}} \times 100 \qquad [\%], \tag{2}$$

where u_{ave} is average of stride time and $u_{s.d.}$ is standard deviation of stride time.

2) Detrended Fluctuation Analysis(DFA): We quantified the long-range correlations using Detrended Fluctuation Analysis (DFA). This technique offers certain advantages over other methods (e.g., spectral or Hurst analyses) when dealing with non-stationary time series, because it "avoids

spurious detection of apparent long-range correlations that are an artifact of non-stationarity" [9], [10].

At first, stride time $u(i)$ is integrated as in (3),

$$y(k) = \sum_{i=1}^{k} (u(i) - u_{ave}), \tag{3}$$

where u_{ave} is the average of stride time. Then, this integrated time series $y(k)$ is divided into non-overlapped equal boxes of length, n. In each box of length n, a least-squares line is fit to the data, which represents the trend in each box. The fluctuation $F(n)$ for each box is then calculated as the root-mean-square deviation between the integrated time-series $y(k)$ and its local trend $y_n(k)$ as in (4).

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^{N} [y(k) - y_n(k)]^2} \tag{4}$$

This calculation is repeated while box sizes n is between 10 and a half of max length of stride time data number to provide a relationship between $F(n)$ and the box size n.

Typically, the fluctuation, $F(n)$, will increase with larger box sizes. A linear relationship between n and $F(n)$ on a log-log plot indicates self-similar scaling property, in that fluctuations in the smaller boxes are related to the fluctuations in the larger boxes in a power-law relation. The slope of the line $\log_{10} F(n)$ over $\log_{10} n$ is the DFA scaling exponent α, and gives a measure of long time correlation of the original stride time. Using DFA, a DFA scaling exponent $\alpha = 0.5$ corresponds to unpredictable white noise; $\alpha = 1.0$ corresponds to 1/f-like noise and long-range correlations [9]. Analyses were performed on the stride time of the left leg. In the case of problems of left foot switches, stride time of right side was used. analyzed data of stride time, which excludes the first and last 10 steps each.

3) Linear Discriminant Analysis: We operated the linear discriminant analysis (LDA) [13] on both CV and DFA of stride time at a time. LDA is one of the supervised classification methods. In this method, likelihood probability of each class is estimated as a normal distribution with equal variation and covariance matrices for all classes are calculated. We evaluate the classification performance with leave-one-out cross validation method [13]. Then we calculate accuracy, sensitivity, specificity, by (5), (6) and (7).

$$accuracy = \frac{True}{True + False} \tag{5}$$

$$sensitivity = \frac{TruePD}{TruePD + FalsePD} \tag{6}$$

$$specificity = \frac{TrueHealthy}{TrueHealthy + FalseHealthy} \tag{7}$$

In which, "True" means the numbers of all true data. "False" means the number of false data. "TruePD" are the numbers of true data estimated to PD. "FalsePD" represents the numbers of false data estimated to PD. "TrueHealthy" and "FalseHealthy" are as well.

III. Result and Discussion

Fig. 3 shows samples of the time series data of stride time (upper panels) and the DFA plot (lower panels). From left to right in fig. 3, each figures shows samples of healthy young, healthy elderly, HY2, HY2.5 and HY3, respectively. Comparing these five data, the temporal variation of stride time of healthy young and healthy elderly were shown to be smaller than that of HY2, HY2.5 or HY3. Using Kruskal-Wallis rank sum test, the significant difference among healthy young, elderly, HY1-2, HY2.5 and HY3 group was shown ($\chi^2(3) = 27.4$, $p < 0.001$). The fractal property of these time series data suggests that the fractal exponent becomes lower in older age. DFA scaling exponent α of PD patients distributed in wide range, and it overlapped to healthy people. However, DFA scaling exponent α of PD patient of HY3 is lower than that of healthy young participants. Using Kruskal-Wallis rank sum test, the significant difference among 5 groups was shown ($\chi^2(3) = 9.52$, $p < 0.05$). Difference between participants of HY3 and that of HY2.5 is whether balance impairment is seen, or not. Therefore, It's indicated that motor control mechanism of participants whose HY is 3 and higher is critically altered from healthy state.

Fig. 4 shows the gait patterns associated with the healthy young people, healthy elderly people, PD patients in HY1-2 group, PD patients in HY2.5 group and PD patients in HY3-3.5 group, in the feature space that is configured with CV and DFA scaling exponent α. It can be observed that healthy people's patterns mainly congregated in the area where $1.3\% < CV < 2.5\%$, whereas most of PD patient's patterns were in the area where $1.5\% < CV < 3.0\%$. Most of CV of HY3-3.5 is between 2.5% and 4.0%. CV of HY2.5 is between 1.8% and 4.4%. And most of CV of HY1-2 is between 1.8% and 3.3%. Almost all healthy young people's patterns were in the area where $0.60 < \alpha < 1.3$, in contrast to healthy elderly people's patterns, which mostly were in the area where $0.40 < \alpha < 0.90$. DFA scaling exponent α of PD patients were almost same level as elderly people. In PD patient's groups, HY1-2 patterns were in the area where $0.63 < \alpha < 1.0$, but most of HY3-3.5 patterns were in the area where $0.50 < \alpha < 0.85$. These results are consistent with previous findings [5], [6]. Fig. 4 indicates that the distribution of DFA scaling exponent α of PD patients were similar to healthy elderly people. And the distribution of HY1-2, HY2.5, HY3-3.5 were wide and overlapped.

As a first step, linear discriminant analysis(LDA) is applied to separate healthy people and PD patients. The independent variables were CV and DFA scaling exponent α of stride time. The straight line in fig. 4 shows boundary which separate healthy people and PD patients. Table III and is the cross-tabulation tables of LDA. Twenty-eight healthy people are truly predicted to healthy people. Total number of healthy people is 35. So specificity is 0.80 as in (8).

$$specificity = \frac{TrueHealthy}{TrueHealthy + FalseHealthy} = \frac{28}{28 + 7} = 0.80 \tag{8}$$

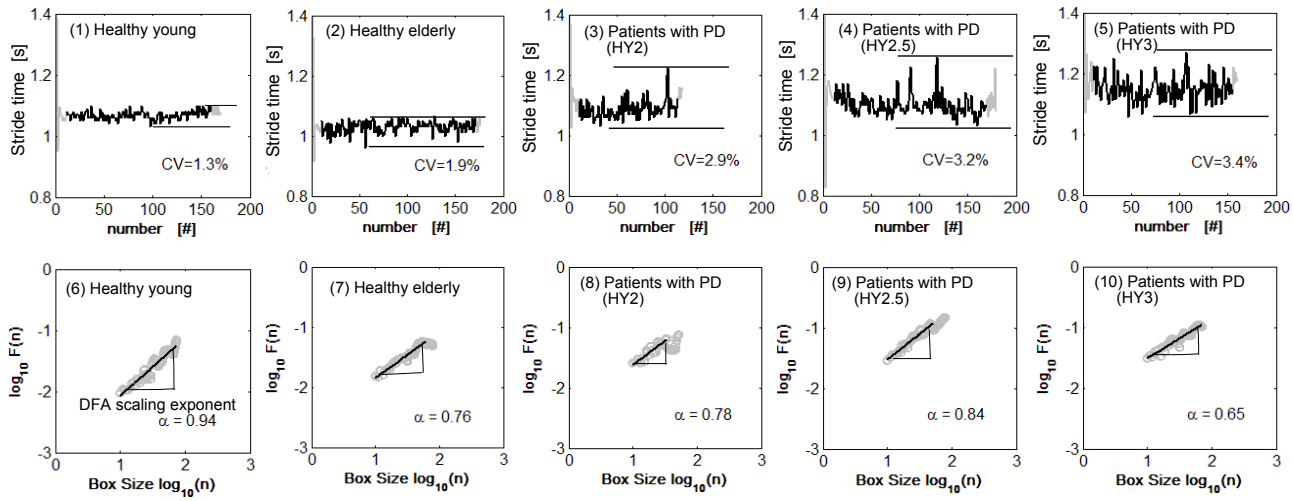

Fig. 3. Samples of time series data of stride interval (figures above) and DFA scaling exponents α (figures below). Stride time variability of healthy young people tends to be smaller than that of elderly people and PD patients. And the DFA scaling exponent α represents the property of stride interval time series structure. Fractal exponent of healthy young people is likely to be higher than that of elderly people and PD patients.

TABLE III
CROSS-TABULATION TABLE (HEALTHY PEOPLE - PD PATIENTS: SEPARATED BY STRAIGHT LINE IN FIG. 4)

Group	Healthy people	PD patients
Healthy people	28	7
PD patients	12	33

Thirty-three PD patients are truly estimated to PD patients, and total number of PD patients is 45. Therefore sensitivity is 0.73 as in (9).

$$sensitivity = \frac{TruePD}{TruePD + FalsePD} = \frac{33}{33+22} = 0.73 \quad (9)$$

The sensitivity is lower than specificity. This result is considered to be come from the overlap between healthy people and milder PD patients. This result speculates the possibility of overlooking early PD patients, the difference between healthy people and PD patients comes to be clearer. The number of true data is 61 and the total number of data is 80, so accuracy is 0.76 as in (10).

$$accuracy = \frac{True}{True + False} = \frac{28+33}{28+7+33+22} = 0.76 \quad (10)$$

Considering the risk of missing detection, analysis using both CV and DFA fractal exponent α of stride time has a potential to be an objective severity index.

Then we divided elderly people into 2 groups using LDA. One is healthy elderly people or mild PD patients group(HE&HY1_2), and the other is severer PD patients group(HY2.5_3.5). The straight line in fig. 5 shows boundary which separate the 2 groups. Table IV is the result of LDA.

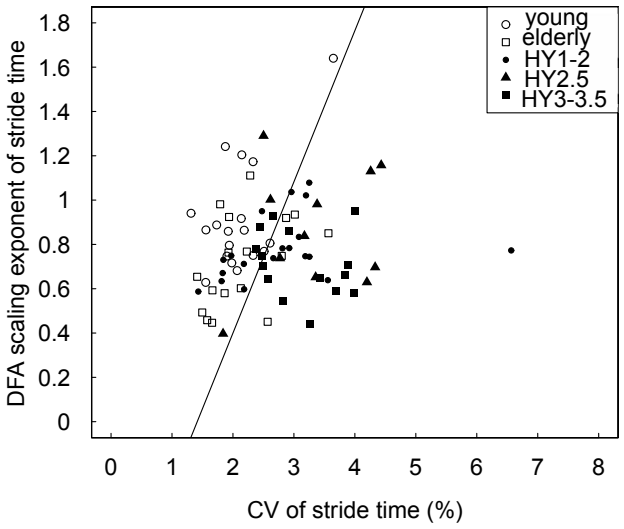

Fig. 4. Distribution of healthy young(white circle), healthy elderly(white square), HY1-2(black circle), HY2.5(black triangle) and HY3-3.5(black square) groups in feature space configured with CV and DFA. The straight line shows the boundary between healthy people(Healthy) and PD patients(PD).

The accuracy, the sensitivity, the specificity are calculated as (11), (12), (13).

$$accuracy = \frac{True}{True + False} = \frac{31+13}{31+5+13+13} = 0.71 \quad (11)$$

$$sensitivity = \frac{TrueHY2.5\_3.5}{TrueHY2.5\_3.5 + FalseHY2.5\_3.5} = \frac{13}{13+13} = 0.50 \quad (12)$$

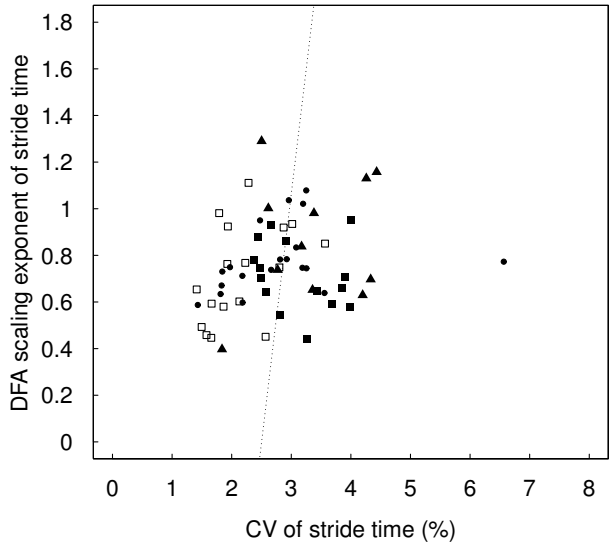

Fig. 5. Distribution of healthy elderly(white square), HY1-2(black circle), HY2.5(black triangle) and HY3-3.5(black square) groups in feature space configured with CV and DFA. The dotted line shows the boundary between severer PD patients and the other group. Severer PD patients are defined as PD patients with HY2.5 or higher(HY2.5_3.5), and the other group are defined as healthy elderly people and PD patients whose HY is 2 or less(HE&HY1_2).

TABLE IV
CROSS-TABULATION TABLE (HEALTHY ELDERLY OR MILD PD PATIENTS: HE&HY1_2 - PD PATIENTS WITH HY2.5 OR HIGHER: HY2.5_3.5 ; SEPARATED BY DOTTED LINE IN FIG. 5)

Group	Healthy elderly or mild PD	Severer PD patients
Healthy elderly or mild PD	31	5
Severer PD patients	13	13

$$specificity = \frac{TrueHE\&HY1\_2}{TrueHE\&HY1\_2 + FalseHE\&HY1\_2} = \frac{31}{31+5} = 0.86 \quad (13)$$

The specificity is much higher than sensitivity. The mild PD patients were very similar to healthy elderly people, however severer PD patients seems to altered from healthy elderly people.

Finally, we separate PD patients with HY2.5 of PD patients with HY3 using LDA. The dashed line in fig. 6 shows boundary which separate between HY3-3.5 group and HY2.5 group. This result indicate that in the group whose member with large stride time variation, the lower the DFA scaling exponent α was, the severer the progression of PD was. Table V is the result of LDA. The accuracy, the sensitivity, the specificity are calculated as (14), (15), (16).

$$accuracy = \frac{True}{True + False} = \frac{14+5}{14+1+5+6} = 0.73 \quad (14)$$

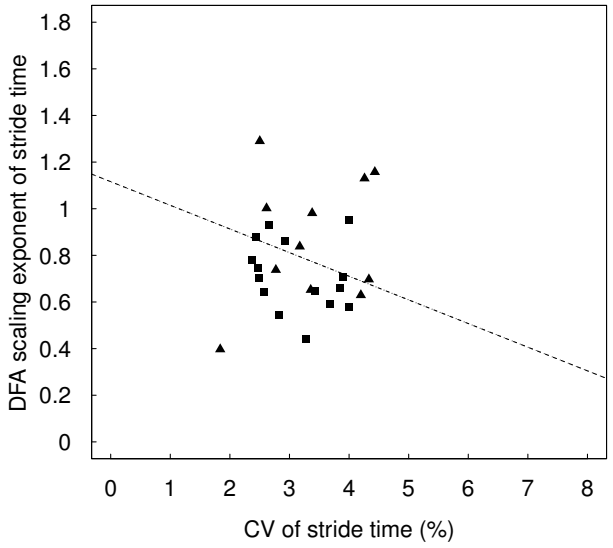

Fig. 6. Distribution of HY2.5(black triangle) and HY3-3.5(black square) groups in feature space configured with CV and DFA. The dashed line shows the boundary between PD patients whose HY is 2.5(HY2.5) and PD patients whose HY is 3(HY3_3.5).

TABLE V
CROSS-TABULATION TABLE (PD PATIENTS WITH HY2.5(HY2.5) - PD PATIENTS WITH HY3-3.5(HY3_3.5): SEPARATED BY DASHED LINE IN FIG. 6)

Group	HY2.5	HY3-3.5
HY2.5	5	6
HY3-3.5	1	14

$$sensitivity = \frac{TrueHY3\_3.5}{TrueHY3\_3.5 + FalseHY3\_3.5} = \frac{14}{14+1} = 0.93 \quad (15)$$

$$specificity = \frac{TrueHY2.5}{TrueHY2.5 + FalseHY2.5} = \frac{5}{5+6} = 0.38 \quad (16)$$

The sensitivity is much higher than specificity. The prediction is not so reliable. However it is suggested that DFA scaling exponent α is associated to postural instability.

From these results, the possibility of classifying between PD patients and healthy people from stride time variability were speculated. Furthermore, the severer PD patients' gait property tends to be altered from the gait property of healthy people or mild PD patients. As a next step, we plan to analyze kinematic variability of PD patients' locomotion in order to find better features which represent PD patients' severity.

ACKNOWLEDGMENT

The authors express our deepest gratitude to participants. We appreciate K. Suzuki and M.J. Hove's contribution on

data acquisition. The authors gratefully acknowledge the contribution of Kanto Central Hospital. We are grateful for anonimous reviewers.

REFERENCES

[1] D. A. Bennett, L. A. Beckett, A. M. Murray, K. M. Shannon, C. G. Goetz, D. M. Pilgrim, and D. A. Evans, "Prevalence of parkinsonian signs and associated mortality in a community population of older people," *New England Journal of Medicine*, vol. 334, no. 2, pp. 71–76, 1996.

[2] C. Ramaker, J. Marinus, A. M. Stiggelbout, and B. J. van Hilten, "Systematic Evaluation of Rating Scales for Impairment and Disability in Parkinsonfs Disease," *Movement Disorders*, vol. 17, no. 5, pp. 867–876, 2002.

[3] C. G. Goetz, W. Poewe, O. Rascol, C. Sampaio, G. T. Stebbins, C. Counsell, N. Giladi, R. G. Holloway, C. G. Moore, G. K. Wenning, M. D. Yahr, and L. Seidl, "Movement Disorder Society Task Force Report on the Hoehn and Yahr Staging Scale : Status and Recommendations," *Society*, vol. 19, no. 9, pp. 1020–1028, 2004.

[4] M. M. Hoehn and M. D. Yahr, "Parkinsonism: onset, progression, and mortality." *Neurology*, vol. 17, no. 5, pp. 427–442, Nov. 1967. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/11775596

[5] J. M. Hausdorff, M. E. Cudkowicz, R. Firtion, J. Y. Wei, and A. L. Goldberger, "Gait Variability and Basal Ganglia Disorders : Stride-to-Stride Variations of Gait Cycle Timing in Parkinson's Disease and Huntington 's Disease," *Movement Disorders*, vol. 13, no. 3, pp. 428–437, 1998.

[6] J. Hausdorff, S. L. Mitchell, R. Firtion, C.-K. Peng, M. E. Cudkowicz, J. Y. Wei, and A. L. Goldberger, "Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease," *Journal of Applied Physiology*, vol. 82, no. 2, pp. 262–269, January 1997

[7] J. M. Hausdorff, "Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking." *Human movement science*, vol. 26, no. 4, pp. 555–89, Aug. 2007.

[8] O. Henmi, Y. Shiba, T. Saito, H. Tsuruta, A. Takeuchi, M. Shirataka, S. Obuchi, M. Kojima, and N. Ikeda, "Spectral Analysis of Gait Variability of Stride Interval Time Series : Comparison of Young , Elderly and Parkinson's Disease Patients," *Journal of Physical Therapy and Science*, vol. 21, pp. 105–111, 2009.

[9] C. K. Peng, S. V. Buldyrev, S. Havtin, M. Simons, H. E. Stanley, and A. L. Goldberger, "Mosaic organization of DNA nucleotides," *Physical Review E*, vol. 49, no. 2, pp. 1685–1689, 1994.

[10] C. K. Peng, S. Havlin, H. E. Stanley, and A. L. Goldberger, "Quantification of scaling exponent and crossover phenomena in nonstationary heartbeat time series," *Chaos (Woodbury, N.Y.)*, vol. 5, no. 1, pp. 82–87, 1995.

[11] M. Hove, K. Suzuki, H. Uchitomi, S. Orimo, and Y. Miyake, "Interactive rhythmic auditory stimulation reinstates natural 1/f timing in gait of parkinson's patients," *PLoS ONE*, vol. 7, no. 3, p. e32600, 2012.

[12] I. Araki and S. Kuno, "Assessment of voiding dysfunction in parkinson's disease by the international prostate symptom score," *J Neurol Neurosurg Psychiatry*, vol. 68, pp. 429–433, 2000.

[13] R. Duda, P. Hart, and D. Stork, *Pattern classification*. New York: Wiley-Interscience, 2001.