33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

On Automated Assessment of Levodopa-Induced Dyskinesia in Parkinson's Disease

Markos G. Tsipouras, *Member, IEEE*, Alexandros T. Tzallas, *Member, IEEE*, Dimitrios I. Fotiadis, *Senior Member, IEEE*, and Spyridon Konitsiotis

***Abstract*—A method for the analysis of accelerometer and gyroscope signals in order to automatically assess the Levodopa-induced dyskinesia (LID) in patients with Parkinson's disease is presented in this paper. Several accelerometers and gyroscopes are placed on certain positions on the subject's body and the obtained signals are analyzed in order to extract several features that depict the energy distribution over the frequency spectrum and the non-linear properties of the signal. These features are fed into an artificial neural network which is used for LID detection and severity classification. The method has been evaluated using a group of 29 subjects. Results are presented related to the body locations that the accelerometers and the gyroscopes are placed. The obtained results indicate high classification ability (84.3% average classification accuracy).**

I. Introduction

Parkinson's Disease (PD) is a complex, progressive neurodegenerative disease of the central nervous system that is manifested clinically with various movement disorder such as tremor, muscle rigidity, bradykinesia and postural instability [1]. The symptoms of the disease are controlled with suitable medication. The use of levodopa (L-Dopa) is highly effective in reducing the symptoms and after many decades of universal usage L-Dopa therapy remains a gold standard for the treatment of PD [2]. However this therapy also has limitations in various facets. The long-term use of L-Dopa is often complicated by the appearance of motor fluctuations and dyskinesias, referred as levodopa-induced dyskinesia (LID) [1]. The frequency of LID in parkinsonian patients has been reported in several studies, with figures ranging from 30 to 80% [2]. Differences in the methods used to assess the presence of LID (self-assessment diary, objective measurement, use of activation procedures), in the clinical setting and in different patient populations (community-based or clinic-based) may account for this variability [3].

Development of a scale for assessment of LID represents one of the most challenging issues to the study of this disabling phenomenon [4]. One critique concluded that an ideal scale for LID would define clinical phenomenology, assess anatomical distribution, rate severity of movements, determine their impact on activities of daily living, and be short and easy to use. However, current rating scales do not fulfill these criteria. In addition, patients' self-evaluation diaries are increasingly used on the effect of drugs potentially useful in the treatment of LID. Data derived from diary cards are accepted by the Food and Drug Administration (FDA) as an end point for regulatory purposes, but the accuracy and compliance to diary completion has been seriously challenged. The quality of data derived from use of diaries may be increased by careful training of patients.

Several computer-based methods have been used to quantify LID including surface electromyography [5], Doppler ultrasound [6], gyroscopes [7], accelerometers [8-10], magnetic motion trackers [11-13] and digital drawings [14] which may be particularly useful for research purposes but are not suitable for widespread use in clinical studies since the datasets used are composed from recordings which include only LID symptoms, while recordings with other PD symptoms (such as tremor and bradykinesia) are not included. Also, the recorded signals include a small number of involved activities since the patients are instructed to refrain from making voluntary movements. A major challenge for an automated method for LID assessment is to be able to distinguish LID symptoms from any other clinical symptoms that a PD patient may present, during any kind of movements (voluntary or not).

The above (LID assessment in ubiquitous environment) is the main goal of this study. For this reason special care have been taken in order to include a large variety of recordings in the study, including PD patients with different LID severities, PD patients presenting tremor and bradykinesia, which are the most common PD motor disabilities, PD patients not presenting any kind of symptom and healthy subjects. Also, voluntary movements and random events are very common in the recordings used for this study.

Manuscript received March 26, 2010. This work was funded in part by the ICT project of the European Commission PERFORM (Project: FP7-ICT-2007-1-215952).

M. G. Tsipouras, is with the Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering, University of Ioannina, GR45110 Ioannina, Greece (e-mail: markos@cs.uoi.gr).

A.T. Tzallas is with the Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering, University of Ioannina, GR45110 Ioannina, Greece (e-mail: atzallas@cc.uoi.gr).

D.I. Fotiadis is with the Unit of Medical Technology and Intelligent Information Systems, Dept. of Materials Science and Engineering, University of Ioannina, GR45110 Ioannina, Greece (e-mail: fotiadis@cs.uoi.gr).

S. Konitsiotis is with the Dept. of Neurology, Medical School, University of Ioannina, GR45110 Ioannina, Greece (e-mail: skonitso@cc.uoi.gr).

II. Materials and Methods

A. Dataset

The dataset used for this study included 29 subjects belonging into three categories: (i) 5 healthy subjects, (ii) 14 PD patients not presenting LID symptoms but other PD common motor symptoms (such as tremor and bradykinesia), (iii) ten PD patients with LID symptoms of various severities. PD patients suffering from LID showed LID severity varying between no dyskinesia to moderate dyskinesia, rating between 0 and 3 on the Unified Parkinson's Disease Rating Scales (UPDRS) [15]. The experiments were approved by the Medical Ethical Committee of the Hospital of the University of Ioannina in Greece.

The movements and postures are measured using accelerometers and gyroscopes and a portable data recorder. Six sets of three orthogonal accelerometers and two sets of three orthogonal gyroscopes (ANCO Devices [16]) are used. These are placed at six different positions of the body: right and left wrist (LW and RW), right and left leg (LL and RL), chest (CH) and waist (WS). Each accelerometer/gyroscope records three signals, one for each axis (x, y and z axis). The above placement of the sensors on the patient's body is illustrated in Fig. 1. All sensors transmit data using Bluetooth to a portable PC, with sampling rate 62.5 Hz.

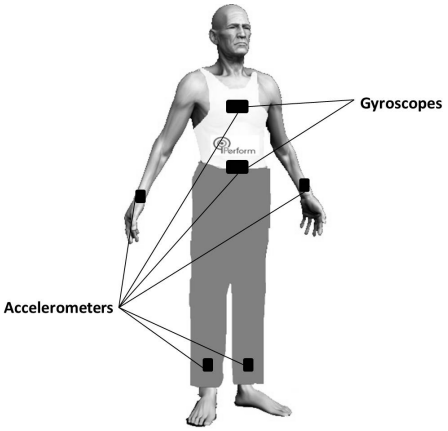

Fig. 1. Placement of the sensors (accelerometers and gyroscopes) on the subject's body.

The recording protocol consists of three major tasks: (i) lying on the bed, (ii) sitting on a chair and (iii) standing up from the chair and performing a series of activities. In the current study only the second part is included (sitting on the chair). During the recording the patients were instructed to act freely, speak and make voluntary movements. Also, several random evens that occurred are included in the study (speaking to cellular phone, writing, eating/drinking, reading from a magazine). All the procedure was videotaped and the medical annotation related to the LID severity was made from expert neurologists using the video footage. Some video snapshots from the recordings are presented in Fig. 2.

B. Signal Analysis

A 2-second moving window with 1 second overlapping is used over each single lead and, for each window several features were extracted. These features are presented in Table I. Features 1 to 3 are mean values of the three axes (x, y and z). For features 4-10, initially the power spectrum density of each axis is calculated and then they are superimposed to provide the signal's power spectrum density, based on which the features are calculated.

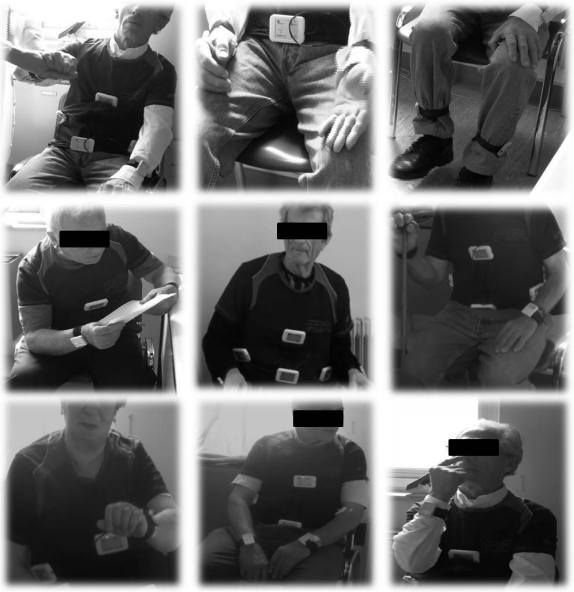

Fig. 2. Snapshots from the video recordings, from various patients, while acting freely and with various body and hand postures.

C. LID Assessment

For each window a feature vector is created, consisting from the above features. LID severity assessment is made for four different body positions: wrists, legs, chest and waist. All signals recorded from the left and the right wrists are considered as a single dataset since mirror-image movements are not distinguished based on the set of the

TABLE I
Extracted Features

1	Mean value (mean for three axis)
2	Standard deviation (mean for three axis)
3	Mean Entropy (mean for three axis)
4	Portion of Signal Energy in 0.5-2 Hz
5	Portion of Signal Energy in 2-4 Hz
6	Portion of Signal Energy in 4-10 Hz
7	Signal Energy in 2-4 Hz
8	Signal Energy in 4-10 Hz
9	Spectral Entropy
10	Spectral standard deviation

above described features. The above also applies for the left and right legs. In the case of the chest and the waist, the above features are extracted from both accelerometers and gyroscopes, placed in these positions. Thus, the dimension of the feature vector is 10 for wrists and legs LID assessment while it is 20 for the chest and the waist. LID assessment have been performed using an artificial neural network (ANN), with 10 or 20 inputs (for wrists/legs and waist/chest, respectively), one hidden layer with 10 neurons and an output layer with 4 neurons.

III. Results

Based on the signal analysis described above four different classification problems have been formulated, i.e. the LID assessment in wrists, legs, chest and waist body positions. Tables II-V present the confusion matrix for the LID assessment in wrists, legs, chest and waist, respectively. Data from wrists and legs are formulated as a single classification problem, since the extracted features are not affected from the axis system orientation (that is mirror-image between left and right hand/leg). The results are extracted using the leave-one-patient-out cross validation technique.

TABLE II
WRISTS CONFUSION MATRIX

		Assessment			
		S0	S1	S2	S3
Annotation	S0	2124	102	45	0
	S1	177	187	81	0
	S2	86	44	432	2
	S3	0	3	10	7

TABLE III
LEGS CONFUSION MATRIX

		Assessment			
		S0	S1	S2	S3
Annotation	S0	2217	93	16	0
	S1	163	260	53	1
	S2	23	77	264	36
	S3	0	7	16	74

TABLE IV
CHEST CONFUSION MATRIX

		Assessment			
		S0	S1	S2	S3
Annotation	S0	1108	34	3	1
	S1	70	113	19	5
	S2	6	16	102	12
	S3	10	9	72	60

TABLE V
WAIST CONFUSION MATRIX

		Assessment			
		S0	S1	S2	S3
Annotation	S0	1097	38	11	0
	S1	70	94	31	0
	S2	57	30	175	2
	S3	0	4	18	23

Additionally, evaluation metrics such as sensitivity and specificity for each LID severity, average sensitivity and sensitivity for all LID severities, and classification accuracy are also calculated and they are presented in Table VI.

TABLE VI
EVALUATION METRICS

		LID assessment				
		S0	S1	S2	S3	average
Wrists	Sensitivity (%)	93.56	42.02	76.6	35	61.76
	Specificity (%)	88.98	55.65	76.06	77.78	74.62
	Accuracy (%)	83.33				
Legs	Sensitivity (%)	95.31	54.51	66	76.29	73.03
	Specificity (%)	92.26	59.50	75.64	66.67	73.52
	Accuracy (%)	85.3				
Chest	Sensitivity (%)	96.68	54.59	75	39.74	66.5
	Specificity (%)	92.80	65.70	52.04	76.92	71.86
	Accuracy (%)	84.33				
Waist	Sensitivity (%)	95.72	48.21	66.29	51.11	65.33
	Specificity (%)	89.62	56.63	74.47	92.00	78.18
	Accuracy (%)	84.18				

IV. Discussion

In this paper, a method for the automated analysis of accelerometer and gyroscope signals is presented. The method is based on the analysis of signals that are obtained from six accelerometers (placed on the wrists, legs, chest and waist of the patient) and two gyroscopes (placed on the chest and the waist of the patient), using a moving window. Several frequency domain and non-linear features are extracted from each window, while the LID severity classification is made using an ANN. The method has been evaluated using recordings from 29 subjects including healthy people, PD patients not suffering from LID and patients that presented LID with severity 0 to 3 at the UPDRS. The obtained results indicate that the proposed methodology is highly efficient for automated LID severity detection and classification.

The medical annotation of the LID severity is based on the determination of time intervals on the video footage that the patient presents LID. These time intervals were then annotated with a LID severity value. However, small variations in the LID severity during a time interval were not recorded, i.e. during a time interval were the patient presents LID with severity 1, small parts were the patient stands still (i.e. presents LID with severity 0) or his symptoms get worst (i.e. presents LID with severity 2) are not annotated from the medical experts since they are not important to the overall clinical image of the patient. However, these intervals are detected from the method and are considered as misclassifications. This can be clearly observed from the obtained results: from the total misclassifications 75.64%, 90.31%, 86.77% and 72.41% occur in neighbour classes for the wrists, legs, chest and waist classification problems, respectively.

Another, important issue is that there are several LID episodes that are annotated as "0-1" severity. All the above cases were considered as LID episodes of severity 1. Also, during a period were the patient doesn't present LID, small parts (such as a single hand or leg movement) that are affected from LID of low severity (1) are not considered of clinical importance and thus are not annotated. The above issues can be clearly observed from the obtained results: from the total misclassifications 50.73%, 52.78%, 40.47% and 41.38% are between the LID severity 0 and 1.

Misclassification on the normal subjects was very low (less than 5%) and mainly towards the LID severity 1; 97.83% of misclassifications were LID of severity 0 that misclassified as LID of severity 1 and only 2.17% misclassified as LID of severity 2 or 3. The same results occurred also to the PD patients that do not present LID; total misclassification was 8.76% with 93.12% being LID of severity 0 that misclassified as LID of severity 1 and only 6.88% misclassified as LID of severity 2 or 3.

In Table IV several similar works for LID assessment that have been proposed in the literature are presented. A direct comparison is not feasible, since different datasets have been employed. However, this study compares well since it involves the second largest dataset (in recorded signals total duration) which includes a large number of activities, and also presents high classification accuracy.

TABLE VI
COMPARISON WITH OTHER WORKS PRESENTED IN THE LITERATURE

Author	Experiment Setting	Dataset	Analysis	Results
Burkhard et al. [6]	2 Gyroscopes *(upper limps)*	24 patients 1 min rec. (150)	Statistical analysis	
Keijsers et al. [7]	8 Accelerometers	16 patients 7 min rec.	MLP	Correlation: Trunk: 0.7, Arm: 0.72, Leg: 0.64
Hoff et al. [8]	8 Accelerometers	16 patients 7 min rec.	Statistical Analysis	
Keijsers et al. [9]	6 Accelerometers	13 patients 2.5 hrs rec.	MLP	Accuracy: Trunk: 83%, Arm: 77%, Leg: 76.9%
Ghassemi et al. [10]	15 magnetic motion trackers	30 subjects 60 sec rec. (90)	Statistical Analysis	
Gourb et al. [11]	15 magnetic motion trackers	30 subjects 60 sec rec. (90)	Statistical Analysis	
Chelaru et al. [12]	15 magnetic motion trackers	29 subjects 60 sec rec. (90)	Statistical Analysis - MLP	Accuracy: 100%
this work	6 accelerometers 2 gyroscopes	29 subjects 6 min rec. (47)	MLP	Accuracy: Wrists: 83%,Legs: 85% Chest: 84%,Waist:84%

REFERENCES

[1] J. Jankovic, "Parkinson's disease: clinical features and diagnosis," J Neurol Neurosurg Psychiatry, vol. 79, no. 4, pp. 368–376, 2008.

[2] H. Yuan, Z.W. Zhang, L.W. Liang, Q. Shen, X.D. Wang, S.M. Ren, H.J. Ma, S.J. Jiao, and P. Liu, "Treatment strategies for parkinson's disease," Neurosci Bull, vol. 26, no. 1, pp. 66–76, 2010.

[3] N.L. Keijsers, M.W. Horstink, and S.C. Gielen, "Online Monitoring of Dyskinesia in Patients with Parkinson's disease," IEEE Eng Med Biol Mag, vol. 22, pp. 96–103, May 2003.

[4] J.I. Hoff, J.J. van Hilten, and R.A. Roos, "A review of the assessment of dyskinesias," Mov Disord, vol. 14, no. 5, pp. 737–743, 2001.

[5] N.Yanagisawa, "EMG characteristics of involuntary movements," in *Dyskinesias*, G. W. Bruyn, Ed. Sandoz BV, Uden, 1984, pp. 142–159.

[6] J. Haines, and P. Sainsbury, "Ultrasound system for measuring patients' activity and disorders of movement," Lancet, vol. 14, no. 2(7781), pp. 802–803, 1972.

[7] P.R. Burkhard, H. Shale, L.W. Langston, and J.W. Tetrud, "Quantification of dyskinesia in Parkinson's disease: validation of a novel instrumental method," Mov Disord, vol. 14, no. 5, pp. 754–63, 1999.

[8] N.L. Keijsers, M. W. Horstink, J.J. van Hilten, J.I. Hoff, and S.C. Gielen, "Detection and assessment of the severity of Levodopa-induced dyskinesia in patients with Parkinson's Disease by neural networks," Mov Disord, vol. 15, no. 6, pp. 1104–1111, 2000.

[9] J.I. Hoff, A.A. van den Plas, E.A. Wagemans, and J.J. van Hilten, "Accelerometric assessment of levodopa-induced dyskinesias in Parkinson's disease," Mov Disord, vol. 16, pp. 58–61, 2001.

[10] N.L. Keijsers, M.W. Horstink, and S.C. Gielen, "Automatic assessment of Levodopa-induced dyskinesias in daily life by neural networks," Mov Disord, vol. 18, no. 1, pp.70–80, 2003.

[11] M. Ghassemi, S. Lemieux, M. Jog, R. Edwards, and C. Duval, "Bradykinesia in patients with Parkinson's disease having levodopa-induced dyskinesias," Brain Res Bul, vol. 69, pp. 512–518, 2006.

[12] J. Gourb, R. Edwards, S. Lemieux, M. Ghassemi, M. Jog, and C. Duval, "Movement patterns of peak-dose levodopa-induced dyskinesias in patients with Parkinson's disease," Brain Res Bul, vol. 74, pp. 66–74, 2007.

[13] M.I. Chelaru, C. Duval, and M. Jog, "Levodopa-induced dyskinesias detection based on the complexity of involuntary movements," J Neurosc Meth, vol. 186, no. 1, pp. 81–89, 2010.

[14] X. Liu, C.B. Carroll, S.Y. Wang, J. Zajicek, and P.G. Bain, "Quantifying drug-induced dyskinesias in the arms using digitized spiral-drawing tasks," J Neurosc Meth, vol. 144, pp. 47–52, 2005.

[15] S. Fahn, and R. Elton, "Unified Parkinson's Disease Rating Scale," in Recent Developments in Parkinson's Disease, Edited by: S. Fahn, C. Marsden, and D. Calne, Eds. Florham Park, NJ: Macmillan Health Care Information, 1986, pp. 153–164.

[16] ANCO. Available: http://www.anco.gr/.