A Heterogeneous Database for Movement Knowledge Extraction in Parkinson's Disease

Albert Samà[1], Carlos Pérez[1], Daniel Rodríguez-Martin[1], Joan Cabestany[1],
Juan Manuel Moreno Aróstegui[1] and Alejandro Rodríguez-Molinero [2] [*]

1- Technical Research Centre for Dependency Care and Autonomous Living
(CETpD) Rambla de l'Exposició, 59-69 08800 Vilanova i la Geltrú - Spain

2- Electrical & Electronic Engineering Department, NUI Galway (NUIG), Ireland

Abstract. This paper presents the design and methodology used to create a heterogeneous database for knowledge movement extraction in Parkinson's Disease. This database is being constructed as part of REMPARK project and is composed of movement measurements acquired from inertial sensors, standard medical scales as Unified Parkinson's Disease Rating Scale, and other information obtained from 90 Parkinson's Disease patients. The signals obtained will be used to create movement disorder detection algorithms using supervised learning techniques. The different sources of information and the need of labelled data pose many challenges which the methodology described in this paper addresses. Some preliminary data obtained are presented.

1   Introduction

Parkinson's disease (PD) is a neurodegenerative disease which symptoms are caused by a decrease in the levels of dopamine, due to the death of the dopamine-producing nerve cells in the brain. PD medications help increase dopamine production and enable patients to reduce tremor and to improve motor control. However, disease progression and long-term levodopa therapy produce complications in PD patients, being motor fluctuations the most common one. PD patients alternate between 'almost normal' periods from the motor symptoms point of view, known as *ON periods* and normally associated with the dopaminergic medication intake, and periods where motor symptoms are more evident, known as *OFF periods* usually produced when the medication blood level decreases. Although disease experience is variable, patients might cycle between motor states from three to four times every day. During ON periods and end-of-dose wearing OFF phenomena (the recurrence of PD motor symptoms before next medication intake), dyskinesia may appear as involuntary movements or dystonic movements, respectively. During OFF periods, patients can experience the full range of classic PD symptoms: tremor, stiffness, dystonia or bradykinesia (slowness of movement), among others.

An accurate reporting of PD motor states and symptoms during patients' daily life will enable doctors to personalize medication intakes and, therefore,

improve the response to treatment. Thus, many research efforts have been made on the ambulatory detection of PD symptoms. Inertial sensors (accelerometers, gyroscopes and magnetometers) are the most common sensors used. For instance, in [1] the severity of tremor, bradykinesia and hypokinesia was detected in 6 PD patients through four inertial sensors, at wrist, foot, thigh and sternum. Tremor and bradykinesia were detected and quantified in 20 PD patients by using two tri-axial gyroscopes located on each of the forearms [2]. In [3], gait parameters of 22 patients were examined in laboratory and gait of only one was analyzed at home through an accelerometer placed at the lower back. In [4], the severity of tremor, bradykinesia and dyskinesia in 12 patients was estimated while patients performed specific Unified Parkinson's Disease Rating Scale (UP-DRS) motor exercises in laboratory conditions. 8 uniaxial accelerometers were located in upper and lower limbs. In [5], the detection of postures and gait in 32 people at laboratory and 5 PD patients at home while performing basic and domestic activities was examined. In [6], dyskinesia and tremor were detected by means of 5 accelerometers in 2 patients in uncontrolled home environment.

Most of the previous studies on PD symptoms detection were located in laboratory and patients performed specific activities [1] [2] [3] [4]. Only few studies took place at patient's home: 1 patient in [3], 5 patients in [5] and 2 patients in [6]. Thus, to the best of our knowledge, there is not any study that collects movement signals from a large number of PD patients in an uncontrolled home or outdoors environment. However, previous studies on PD symptom's detection rely on supervised learning techniques such as Decision Trees [1], linear classifiers [2] or Support Vector Machines [4] [7] [6]. Considering that the conditions in which the training database is constructed limit the applicability of the learning algorithms, we believe that a database of signals collected in an uncontrolled environment is necessary to obtain valid algorithms for the daily life of patients.

This paper describes the design and the technical solutions used to construct a database that addresses the need of movement signals acquired in uncontrolled environments. This database is constructed as part of the Personal Health Device for the Remote and Autonomous Management of Parkinson's Disease (REM-PARK) project. The database will not only consist of movement sensor signals that will allow the training of learning algorithms in uncontrolled environments but also standard medical scales and information about the activities performed by patients. This database is being constructed from data gathered in Spain, Italy, Ireland and Israel. The database will be finished by June of 2013.

The paper is divided as follows. In Section 2 we describe the design and requirements for the database. In Section 3 the methodology and the technical solution employed to implement the database design are described and real data already gathered are given. In Section 4, we present the conclusions.

2   Movement Signals Database Design

REMPARK database aims to create algorithms for identifying motor status and motor symptoms of PD patients in an uncontrolled environment. As previously

stated, it is required that the database construction takes place in an uncontrolled environment, both at patient's home and in its surroundings, in order to obtain algorithms suitable to be used during the daily life of PD patients. The database protocol used to collect the data was designed by medical experts and complied with ethical approval. The data collection protocol was designed to be homogeneous to facilitate reproducibility of results.

The main objectives are the obtaining of a database of properly identified inertial signals in order to I) train learning algorithms for motor states detection; II) train learning algorithms for detecting the following symptoms: tremor, Freezing of Gait (FoG), bradykinesia of lower and upper limbs and dyskinesia; III) obtain signals corresponding to movements and activities that can be mistaken for PD motor symptoms, potential false positives (FP); and IV) develop algorithms for extracting the gait speed and step/stride length.

In order to accomplish these objectives, before the training phase of the learning algorithms for symptoms detection, inertial signals must be labelled as the presence or absence of each symptom. For instance, the training of algorithms for dyskinesia detection would need several minutes, or hours, of sensor signals with dyskinesia and several more without dyskinesia. For non-symptoms detection, the value to estimate is needed as a gold-standard, e.g. for stride length estimation the actual stride length of several patients and their corresponding sensor signals are needed. These requirements have been taken into account in the database design in such a way that the collection and labelling of presence/absence data is facilitated by asking patients to perform specific activities. For instance, patients are asked to walk through a door, make turns and walk through narrow spaces with the aim of provoking and recording FoG episodes [8]. However, at the same time free activity is allowed so that daily life activities are recorded and labelled as well. Thus, the database includes two types of experiments:

- Short controlled tests where the patient is asked to perform certain activities so that specific motor symptoms, FP, UPDRS scale values, gait speed and step/stride length are obtained. These tests must be closely controlled so that an accurate gold-standard is used.

- Free activity monitoring of the patient so that natural symptoms and activities are recorded. This monitoring may last hours and the gold-standard used is required to be not as accurate as in the previous case.

The obtention of signals that could be mistaken with PD symptoms is a difficult task since there are, potentially, infinite FP. Researchers thought many FP's and they were included in the experiments as short controlled tests. However, assuming that there are more FP, the free activity monitoring is expected to provide the habitual FP that are produced in the daily life of PD patients and that were not thought by the researchers.

The two types of experiments take place alternatively during the data gathering of a patient, which is specified to last from 6 to 8 hours to have at least two

motor states. Thus, when the patient is in the first OFF phase, the specific controlled tests for the OFF symptoms will be conducted and the remaining OFF state time will be used for monitoring their free and natural activity. Similarly, when the patient enters in the first ON phase, some short tests for capturing ON symptoms will be performed and the rest of the time is devoted to monitor the free natural activity.

Since supervised learning techniques will be used, it is crucial that signals are correctly labelled, i.e. a reliable gold standard is used. Thus, on the one hand, video recording will be used as a gold-standard of the first kind of experiments and, on the other hand, during the second type of experiments only annotations in an electronic support are required. REMPARK database is composed of the video, annotations and signals acquired during the experimentation. UPDRS values and other sociodemographic and PD information are the rest of it.

3   Methodology and technical implementation

This section describes the different issues that arise during the implementation of the design presented in the previous section. The technical solutions taken are presented and the process applied to the collected data until they are ready for the creation of learning algorithms is described.

The database design specifies two gold-standards: video recordings and annotations. A mobile phone camera (Nexus S Google) was chosen since mobile phones are less intrusive than standard video cameras, which are bigger, and due to its HD quality and small weight. A specific JAVA-based application that is executed in a tablet (DELL Latitude ST) was developed to allow annotations of symptoms, daily life activities, postures and medication intakes.

Additionally to the gold-standards, movement sensor signals are the third source of information. Two inertial sensors are used to collect data: one in the wrist for assessing tremor and bradykinesia and one in the waist near the Iliac Crest for detecting the rest of symptoms, as well as bradykinesia. Wrist device contains a triaxial accelerometer (LIS344ALH, $\pm 6g$ range where $g \approx 9.81\text{m/s}^2$), a microcontroller (PIC24F) and a communications unit. The microcontroller acquires samples at 80 Hz and sends them to the waist sensor unit using a Bluetooth (BT) 2.1 link. Waist device is composed of a triaxial accelerometer (LIS3LV02DQ, $\pm 6g$ range), a triaxial gyroscope (IDG650+ISZ650, $\pm 2000°$/s range), a triaxial magnetometer (HMC6042+HMC1041Z, $\pm 6G$ range), a microcontroller (dsPIC33F), a BT unit and a memory unit. The microcontroller is in charge of sampling data at 200 Hz and saving them, together with the wrist data received by BT, in a μSD memory card. Finally, waist sensor includes a Real-Time Clock/Calendar (RTCC, ABS06 at 32.768 Khz) that was calibrated against both Network Time Protocol (NTP) servers and frequency meters. A temperature correction by using an internal sensor is applied, obtaining a deviation lower than 40 ms. in a 48-h period. The sensor RTCC synchronization is given from the tablet to the waist sensor by BT previously to collect data.

The three sources of information enable the training of learning algorithms

only if they share the same timeline. Moreover, this is even more crucial since the labels of the signals corresponding to the videotaped parts will be obtained from the video. Thus, clocks of the three sources must be synchronized. There are two common ways of synchronizing clocks against the same timeline: obtaining current time from NTP servers or GPS signals. In our case, GPS signal might not be available at patient's home, neither Internet connection for NTP. Therefore, both mobile phone and tablet clocks, and consequently movement sensor clock as well since it gets the time from the tablet, are updated in the investigator's laboratory by NTP before going to patient's home. However, the clocks can deviate from each other during the experimentation given the minimum duration of 6 hours. The solution employed consists of measuring the time difference between tablet and mobile phone clocks several times during the experiment (before and after a video recording) so that they can be later corrected through a linear transformation, similarly to [10]. Time differences are obtained in the tablet by cable after connecting the mobile phone.

The described synchronization process provides a synchronization error between signals and video of until 1 s. This error is due to the delay that the mobile phone Operative System adds, since the video's initial and final time might be saved with delay. In order to correct this delay, an event which can be identified in both video recordings and movement signals is used to synchronize them. A fall of the sensor has been chosen, since a peak is obtained which clearly identifies the signals, as shown in Fig. 1. This event is performed at the beginning and at the end of each video recording. This way, both video and signals can be synchronized with less than 100 ms. of error.

The synchronization process described is applied by means of specifically developed applications once the experimentation has finished. These applications enable the labelling of videos as well. Fig. 2 shows the signals obtained from the OFF videotaped part of one of the first patients and some of the corresponding labels obtained with the applications.

Fig. 1: Synchronization event

4   Conclusions

In this paper, a methodology that is being used to create a database of PD symptoms measured by means of a waist and a wrist movement sensor has been shown. The database will enable the creation of detection algorithms in real daily life conditions through supervised learning techniques that will allow clinics to
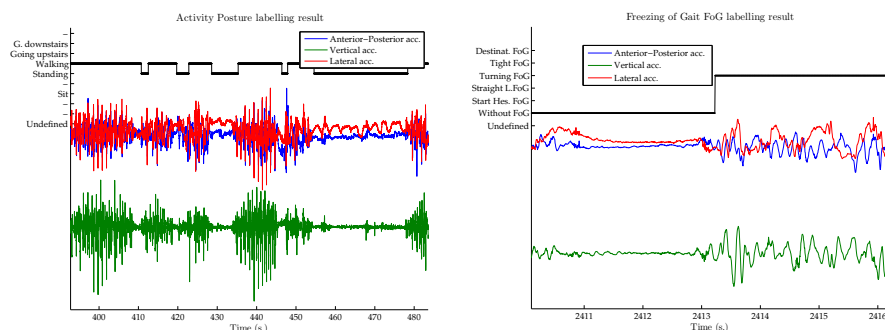
Fig. 2: Waist data and its synchronized Activity/Posture and FoG label

manage PD in a more accurate way. The total number of recruited patients will
be 90, distributed among 4 countries. The access to the database is restricted
to part of REMPARK consortium.

References

[1] Zwartjes, D.G.M.; Heida, T.; van Vugt, J.P.P.; Geelen, J.A.G. and Veltink, P.H., Ambu-
latory Monitoring of Activities and Motor Symptoms in PD, *Biom. Eng., IEEE Trans-
actions on*, 57(11):2778-86, 2010.

[2] Salarian, A.; Russmann, H.; Wider, C.; Burkhard, P.R.; Vingerhoets, F.J.G. and
Aminian, K., Quantification of Tremor and Bradykinesia in PD Using a Novel Ambu-
latory Monitoring System, *Biom. Eng., IEEE Trans. on*, 54(2):313-22, 2007.

[3] Weiss, Aner; Sharifi, Sarvi; Plotnik, Meir; van Vugt, Jeroen P P; Giladi, Nir and Haus-
dorff, Jeffrey M, Toward automated, at-home assessment of mobility among patients with
PD, using a body-worn accelerometer, *Neurorehab. and Neural Repair*, 25:810-8, 2011.

[4] Patel, S.; Lorincz, K.; Hughes, R.; Huggins, N.; Growdon, J.; Standaert, D.; Akay, M.;
Dy, J.; Welsh, M. and Bonato, P., Monitoring Motor Fluctuations in Patients With PD
Using Wearable Sensors, *Inf. Tech. in Biomedicine, IEEE Trans. on*, 13(6):864-73, 2009.

[5] Baukje Dijkstra; Ype P. Kamsma and Wiebren Zijlstra, Detection of Gait and Postures
Using a Miniaturized Triaxial Accelerometer-Based System: Accuracy in Patients With
Mild to Moderate PD, *Archives of Phys. Med. and Rehab.*, 91(8):1272-77, 2010.

[6] Das, Samarjit; Amoedo, Breogan; Torre, Fernando De la and Hodgins, Jessica, Detecting
PD symptoms in uncontrolled home environments: A multiple instance learning approach.
In *Proceedings of the Engineering in Medicine and Biology Society (EMBC), 2012 I. C.*,
p. 3688-91, August 28-September 1, San Diego (USA), 2012.

[7] Sama, A.; Perez-Lopez, C.; Romagosa, J.; Rodriguez-Martin, D.; Catala, A.; Cabestany,
J.; Perez-Martinez, D.A. and Rodriguez-Molinero, A., Dyskinesia and motor state detec-
tion in PD patients with a single movement sensor. In *Proceedings of the EMBS, 2012
Int. Conf.*, p. 1194-97, August 28-September 1, San Diego (USA), 2012.

[8] Morris, M.E., Locomotor training in people with PD, *Arch.P.Med.&Reh.*,86:1426-35,2006.

[9] N. Giladi; H. Shabtai; E.S. Simon; S. Biran; J. Tal and A.D. Korczyn, Construction of
FoG questionnaire for patients with Parkinsonism, *Park. & Rel. Dis.*, 6(3):165-170, 2000.

[10] Jones, T.; Koenig, G.A., A Clock Synchronization Strategy for Minimizing Clock Vari-
ance at Runtime in High-End Computing Environments. In *Comp. Arch. and High Perf.
Comp., 22nd Int. Symp. on* (SBAC-PAD 2010), p. 207-14, Oak Ridge (USA), 2012.