

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Programa de doctorat

AUTOMÀTICA, ROBÒTICA I VISIÓ

Tesi doctoral

**Human movement analysis by means of accelerometers:
Application to human gait and motor symptoms of
Parkinson's Disease**

Albert Samà Monsonís

Director: Francisco Javier Ruiz Vegas

Co-directora: Núria Agell Jané

Juliol 2013

Als meus pares, en Manel i la Maria Mercè i, en especial, als meus germans Jordi i Mercè, els millors companys de viatge possibles, i a l'Alba, la meva motivació.

Abstract

This thesis presents the original contributions of the author on the field of human movement analysis from signals captured by accelerometers. These sensors are capable of converting acceleration from some body parts into electric signals for further analysis. The progressive refinement and miniaturization of accelerometers has allowed the development of minimally invasive devices that can be used to ambulatory monitor human movements during daily live activities.

The study's contributions mainly fall under two heads: first, the analysis of movement in Parkinson's disease (PD); and, second, the relationship between accelerometer signals and characteristics of gait. To this end, new methods for obtaining speed and length of strides and, also, for identifying people have been developed. In all these studies, a single sensor fixed to the patient's waist has been used.

PD is a neurodegenerative disease characterized by movement alterations. The main motor symptoms of PD are 1) tremor, 2) bradykinesia or slowness of movements, 3) freezing of gait and 4) dyskinesia or abnormal involuntary movements. The first three symptoms primarily occur when the medication has not yet reached an effective therapeutic effect. These periods are commonly known as OFF periods or OFF motor state. On the other hand, periods when the patient is suitably responding to the medication are known as ON periods or ON motor state. Dyskinesias mainly appear when the medication blood level is excessive.

Both dyskinesias and OFF motor states are caused by a defect in the medication administration. In this sense, a wearable device capable of detecting and recording dyskinesias and OFF periods represents an important tool that enables clinics to more accurately prescribe the medication regimen of a patient. The work done in the field

of PD consisted in developing algorithms able to detect dyskinesias and both ON and OFF periods. These algorithms have been adapted to provide real-time detection, which enabled their employment in a pilot study. This clinical study has tested, for the first time, the automatic adjustment of medication performed by means of a subcutaneous infusion pump according to the dyskinesias appearance and motor state of PD patients.

The experience gained in the treatment of accelerometric signals from PD has led to contribute in the field of gait analysis. First, new methods for obtaining speed and length of strides from a single sensor fixed to the patient's waist have been obtained. Not only the PD can benefit from this study, but other diseases such as diabetes or some orthopedic-traumatological diseases can also benefit from its results. Finally, using some of the techniques of the previous studies, another important contribution has been made in the field of biometric person identification. The work presented shows how the signal obtained from a single accelerometer located at the waist not only enables the extraction of some gait characteristics but also permits the identification of a person through its gait pattern.

The main theoretical contribution of this thesis has been the development of techniques based on the reconstruction of attractors. It has been shown that the usage of only a small number of features that characterize the reconstructed attractor obtained from a time series of acceleration measurements makes possible the extraction of important parameters of gait and the person identification.

Resumen

La tesis que este documento recoge es una aportación al análisis del movimiento humano a partir de las señales capturadas por acelerómetros. Estos sensores permiten convertir la aceleración producida por algunas partes del cuerpo en señales eléctricas susceptibles de un análisis posterior. El progresivo perfeccionamiento y miniaturización de los acelerómetros ha permitido construir sensores poco invasivos y que pueden ser usados de forma ambulatoria para registrar los movimientos realizados en las actividades de la vida diaria del sujeto.

La tesis se focaliza en dos ámbitos. El primero, fundamentalmente clínico, se ha centrado en el análisis del movimiento en enfermos de la Enfermedad de Parkinson (EP). El segundo ámbito, más general, ha consistido en relacionar las señales acelerométricas con características de la marcha. Con este fin, se han desarrollado métodos para la obtención de la velocidad de la marcha y la longitud del paso, así como para la identificación de personas. En ambos ámbitos se ha empleado un único acelerómetro situado en un lado de la cintura.

La EP es una enfermedad neurodegenerativa que produce primordialmente trastornos del movimiento en los pacientes que la sufren. Los principales síntomas motores de la enfermedad son 1) los temblores, 2) la lentitud de movimientos o bradicinesia, 3) la congelación de la marcha (freezing of gait FoG) y 4) los movimientos involuntarios o discinesias. Los tres primeros síntomas aparecen cuando la medicación no ha alcanzado un efecto terapéutico efectivo. Estos periodos se conocen habitualmente como período o estado motor OFF. Los periodos en los cuales la medicación hace efecto y los pacientes presentan una movilidad normal (o casi normal) son los periodos o estados motores ON. Las discinesias aparecen principalmente cuando el nivel

de medicación en sangre es excesivo.

Tanto las discinesias como los estados OFF son consecuencia de un defecto en la administración de la medicación. Un dispositivo no invasivo que detecte y registre las discinesias y ambos periodos ON y OFF supone una importante herramienta que permite al médico prescribir con mayor precisión la dosis de medicamento adecuada a su paciente. El trabajo realizado en esta tesis en el ámbito de la EP ha consistido en el desarrollo de algoritmos de detección de discinesias y periodos OFF. Estos algoritmos han sido adaptados para proporcionar una detección en tiempo real, de forma que se han empleado ya en un estudio piloto en los que el ajuste de medicación suministrada por una bomba de infusión subcutánea se realiza de forma automática en función de la presencia de discinesias y el estado motor del paciente.

La experiencia ganada en el tratamiento de la señal acelerométrica proveniente de enfermos de Parkinson ha permitido contribuir en el campo del análisis de la marcha y realizar una aportación que relaciona varios parámetros de la misma con la señal que suministra un único acelerómetro situado en la cintura. No solo la EP puede beneficiarse de este estudio, sino también otras enfermedades como la Diabetes o algunas enfermedades ortopédicas y traumatológicas pueden aprovecharse de sus resultados.

Por último, usando algunas de las técnicas de los estudios anteriores, se ha realizado una importante contribución en el ámbito de la identificación biométrica de personas. Se ha puesto de manifiesto que la señal proveniente de un único acelerómetro situado en la cintura no solo permite obtener algunas de las características de la marcha sino también permite identificar a la persona a través del patrón de su marcha.

La principal contribución teórica de esta tesis ha sido el desarrollo de técnicas basadas en la reconstrucción de atractores. Se ha evidenciado que un número muy reducido de características procedentes del atractor reconstruido a partir de una serie temporal de medidas de aceleración permite la extracción de los parámetros de la marcha y la identificación de personas.

Acknowledgements

It is always difficult to acknowledge all the support received in these 4-years long period. I would like to thank the support of my doctoral advisors, Prof. Francisco Ruiz and Prof. Núria Agell, from who I learned so much, both in the research and in the personal fields. I would also like to express my gratitude to the thesis reviewers Prof. John Nelson and Prof. Davide Anguita, and to the members of my jury.

Most of the clinical ideas that shaped this thesis came from the researchers and professors from the Technical Research Center for Dependency Care and Autonomous Living. I would like to specially thank Alejandro Rodríguez Molinero, Joan Cabestany and Andreu Català, their ideas made this PhD possible. I would like to also acknowledge the support and the ideas of Carlos Pérez, his brilliance and constancy really helped me during these years. Finally, Diego Pardo, Daniel Rodríguez, Xavi Pérez and Jaume Romagosa also supported me with nice ideas and technical support.

Glossary

Bradykinesia. It is a motor complication that usually appears during OFF periods and consists in slowness of movements.

Dyskinesia. It is a motor complication that consists in involuntary movements with a variable occurrence following the intake of dopaminergic drugs.

Freezing of gait, FoG. It consists in a blockade of the motor activity, resulting in a sudden inability to start or continue walking.

Hoehn and Yahr scale. It is a commonly used system for describing how the symptoms of Parkinson's disease progress. This scale allocates stages from 0 to 5 to indicate the relative level of disability.

Levodopa. Levodopa is a medication used to treat Parkinson's disease. Parkinson's disease is associated with low levels of a chemical component called dopamine in the brain. Levodopa is turned into dopamine in the body and therefore increases levels of this chemical component.

ON/OFF periods. When Parkinson's disease patients start to take medication, usually levodopa, they notice that their symptoms go away for hours at a time (ON period), then the symptoms return (OFF period).

Stance phase. Gait phase in which the foot is in contact with the ground.

Swing phase. Gait phase in which the foot is suspended and moves forward.

PD. Parkinson's Disease

UPDRS. It is a rating scale used to follow the longitudinal course of Parkinson's disease. It is made of different sections, including Hoehn and Yahr scale, that evaluate current motor and non-motor symptoms. In 2007, the Movement Disorder Society published a revision of the UPDRS, that is currently used.

Contents

Abstract	v
Resumen	vii
Acknowledgements	ix
Glossary	xi
1 Introduction	1
1.1 Purpose and framework	2
1.2 Objectives	5
1.3 Motor Symptoms of Parkinson’s Disease	6
1.3.1 The disease’s impact	6
1.3.2 Parkinson’s Disease: causes, symptoms and present treatment	7
1.4 Human gait	9
1.4.1 Human gait analysis	10
1.4.2 Gait cycle	10
I Dyskinesia and ON/OFF motor states detection in Parkinson’s Disease patients	13
2 State of the art in dyskinesia and motor states detection in PD patients	17
2.1 The importance of detecting and monitoring motor states and dyskinesia	17

2.2	Inertial sensors	18
2.3	Accelerometer signal analysis for ON/OFF motor states detection . .	20
2.4	Accelerometer signal analysis for dyskinesia detection	22
2.5	Conclusions	23
3	Data collection with PD patients	25
3.1	Sensor device: position and components	26
3.2	Data framework for offline dyskinesia and motor states detection algorithms	27
3.3	Data framework for online dyskinesia and motor states detection algorithms	30
3.4	Conclusions	33
4	Dyskinesia detection	35
4.1	Offline algorithm for dyskinesia detection	36
4.1.1	Parameter tuning for the dyskinesia detection offline algorithm	40
4.1.2	Offline dyskinesia detection algorithm in signals obtained from patients' home	43
4.2	Dyskinesia detection online algorithm in healthy people: false positive identification	44
4.3	Online dyskinesia detection in PD patients	45
4.3.1	Available data for the online dyskinesia detection validation .	46
4.3.2	Results and discussion	47
4.4	Conclusions	49
5	ON/OFF motor states detection	53
5.1	Offline algorithm for the ON/OFF motor states detection	54
5.1.1	Algorithm results in laboratory tests on PD patients	56
5.1.2	Algorithm results in daily life environment on PD patients . .	58
5.2	Online detection of ON/OFF motor states	61
5.2.1	Adapting the offline motor states detection to the real-time processing	64

5.2.2	Threshold Adjustment Results	66
5.2.3	Online ON/OFF motor states detection results	67
5.2.4	Discussion	77
5.3	Conclusions	81
 II Human movement analysis for the extraction of gait parameters and gait recognition tasks		83
6	Signal treatment methods for gait analysis in the temporal and frequency domains	87
6.1	Accelerometry signal analysis for the extraction of gait parameters . .	87
6.2	Accelerometry signal analysis for gait identification	93
6.3	Conclusions	95
7	Signal analysis methods based on attractor reconstruction	97
7.1	Optimal values for the state space reconstruction	99
7.1.1	Estimation of time-lag τ optimal value	100
7.1.2	Estimation of embedding dimension k optimal value	101
7.2	Singular Spectrum Analysis	102
7.3	Cell-to-cell Mapping	106
7.4	Measurements to characterize dynamical systems in the reconstructed state space	108
7.5	Gait analysis through non-linear signal methods	109
7.6	Conclusions	111
8	Gait parameters extraction in PD patients by means of accelerometers	113
8.1	Gait parameters extraction in the temporal domain	114
8.1.1	Methodology	114
8.1.2	Experiments and results	117
8.2	Gait parameters extraction through reconstructed attractors	118
8.2.1	Methodology	119
8.2.2	Experiments and results	122

8.2.3	Discussion	124
8.3	Conclusions	125
9	Gait recognition by means of a single waist-worn accelerometer	127
9.1	Gait recognition in the temporal and frequency signal domains	128
9.1.1	Segmentation process	128
9.1.2	Representative cycle	129
9.1.3	Gait recognition experiments	129
9.1.4	Results	130
9.2	Granular approach for gait recognition	130
9.2.1	Granular approach	131
9.2.2	Experiments	133
9.2.3	Results and discussion	133
9.3	Box Approximation Geometry for gait recognition	137
9.3.1	Box Approximation Geometry	138
9.3.2	Gait identification by means of Box Approximation Geometry	139
9.3.3	Experiments	141
9.3.4	Results	142
9.3.5	Discussion	146
9.4	Conclusions	148
III	Final remarks	151
10	Conclusions and future work	153
10.1	Dyskinesia detection and motor state identification in Parkinson’s disease	153
10.2	Accelerometer-based gait analysis	155
10.3	Future Work	156
10.4	Publications	158
	Bibliography	161

List of Tables

4.1	Validity of the offline dyskinesia detection algorithm in the daily life of 15 PD patients	43
4.2	Duration of the dyskinesia online detection tests for its validation in healthy users	45
4.3	Frequency of the activities that provoked false positives (FP) in the online dyskinesia detection	45
4.4	Online detection of dyskinesia results from HELP study	49
5.1	ON/OFF Motor States Detection Results in 10 last patients from laboratory tests (only 6 had motor fluctuations)	58
5.2	ON/OFF Motor States Detection Results in 10 last patients from laboratory tests averaging 5 strides (only 6 had motor fluctuations)	58
5.3	Validity of the OFF detection algorithm for each patient (algorithm applied to walking segments of any length)	62
5.4	Validity of the OFF detection algorithm for each patient (algorithm applied to walking segments longer than 9 strides)	63
5.5	Threshold values for all patients.	69
5.6	Real-time ON/OFF motor states detection results	71
8.1	Summary of results for step velocity and step length.	118
8.2	Summary of results for step velocity and step length.	123
8.3	Velocity Error Case Analysis (Average Velocity = 68 cm/s)	124
8.4	Length Error Case Analysis (Average Length = 83.24 cm)	124

List of Figures

1.1	Gait cycle phases and events [32].	11
3.1	Movement sensor used and its position in a neoprene belt.	27
3.2	Apomorphine pump used in HELP project.	31
4.1	On the left, signal from a patient without dyskinesia. Patient sat during 40-47 s., stood (47-52 s.) and walked (55-55 s.). On the right, signal from another patient with dyskinesia. Low frequency and amplitude harmonics are observed for dyskinetic patient. Patient remained sat (until 50 s.), was stood (50-56 s.) and then walked. Higher amplitudes are shown with warm colours and lower amplitudes with cold colours.	37
4.2	ON/OFF diary filled by a PD patient during HELP study. Data missing in the diary are shown.	48
4.3	Dyskinesia online detection for patient 5 th . Online dyskinesia probability (see Eq. 4.2) and the dyskinesia predicted label are shown. Dyskinesia probability was computed based on the analysis of 1 minute periods. The algorithm successfully monitored this patient during this day.	50
5.1	Three stages of the ON and OFF motor states characterization and classification	54
5.2	Three strides detected by the algorithm described. Strides are divided into two steps. Gyroscope signals from shanks that validate the detection are also shown.	57

5.3	Frequency content of a stride during an ON state and a stride during an OFF state. Lower amplitudes are observed in the OFF state stride.	59
5.4	Outcomes of the dyskinesia and ON-OFF detection algorithms. Dyskinesia ends with the motor state change.	62
5.5	Graphical representation of the strides collected during the first week for (a) patient 1 (Madrid-1), (b) patient 2 (Madrid-2), (c) patient 3 (Barcelona-1), (d) patient 4 (Barcelona-2) and (e) patient 6 (Barcelona-4). Data showed correspond to \mathcal{X}_1 , i.e. minutes in which the sensor detected at least 1 stride are depicted.	68
5.6	Graphical representation of the strides collected during the second week and the two thresholds used to determine the motor state for (a) patient 1 (Madrid-1), (b) patient 2 (Madrid-2), (c) patient 3 (Barcelona-1), (d) patient 4 (Barcelona-2) and (e) patient 6 (Barcelona-4). Data showed correspond to \mathcal{X}_1 , i.e. minutes in which the sensor detected at least 1 stride are depicted.	70
5.7	ON/OFF online detection during the second validation day of patient 3. Data showed correspond to $k=3$. Crosses are the measurements that correspond to the characterized strides that were obtained in real-time by the ON/OFF detection algorithm.	72
5.8	ON/OFF online detection during the fifth validation day of patient 3. Data showed also correspond to $k=3$. Misclassified elements do not allow a sensitivity greater than 0.72.	73
5.9	ON/OFF online detection during the fourth day of patient 3 first week. Data showed correspond to $k=5$. Motor fluctuations are very clear in the sensor measurements. However, the second OFF state is detected at 14:15 by the sensor, which is 45 min. before the patient described this state in the diary (15:00). The first and last intermediate states are perfectly determined by the algorithm.	74

5.10	ON/OFF online detection during the second day of patient 3 second week. Data showed correspond to $k=5$. Fluctuations in the intermediate and the ON state are shown in the ON/OFF detection algorithm's output. There is not any data available during the OFF period in this day.	75
5.11	ON/OFF online detection during the last day of patient 3 validation week. Data showed correspond to $k=5$. Algorithm's output during the first OFF period shows a clear difference compared to the previous state, although measurements are not correctly classified. The last OFF period trend is clearly obtained in advance.	76
5.12	ON/OFF real-time detection for patient 6 during 1st training day. Data showed correspond to $k=1$. First fluctuation from OFF to ON is detected with a delay of 1 h. Second fluctuation is not detected although a decrease in the measurements is observed.	77
5.13	ON/OFF real-time detection for patient 6 during 2nd training day. Data showed correspond to $k=1$. First, third and fourth OFF periods are correctly detected, although there was not any data available for the second one. Second, third and fourth ON periods are correctly identified. First ON period lacks data and in the last one measurements were increased although an it was not correctly detected.	78
5.14	ON/OFF real-time detection for patient 6 during 1st validation day. Data showed correspond to $k=1$. All OFF periods are detected and, with the exception of the third period, all of them were correctly identified.	79
5.15	ON/OFF real-time detection for patient 6 during 3rd validation day. Data showed correspond to $k=1$. All fluctuations are shown by the algorithm measurement trends.	79
6.1	Gyroscope and accelerometer used without biomechanical model [90]. Angular rotation provides the current slope θ enabling to separate the forward acceleration (stride length) from the rest of accelerations. . .	89

6.2	The upper figure shows a gait cycle division according to swing and stance phases. The lower figure presents the CoM trajectory during these phases according to the inverted pendulum model.	91
6.3	On the top, sensor locations in a double inverted pendulum model. On the bottom, model used in [10]. Thigh and shank lengths, l_1 and l_2 , are used to establish α and β angles, for which gyroscope signal is integrated so displacement d is obtained	92
6.4	On the left, accelerometry signal ($F_s = 40$ Hz) segmented into steps through the autocorrelation function similarly to [40]. On the right, autocorrelation function values for two seconds of signal. A local maxima after 25 samples appears (0.5 s.) with the new step. Another local maxima appears after 50 samples, which corresponds to the new stride.	95
7.1	Time series, scree plots and 2-d reconstruction for white noise, sinusoidal and Lorenz series.	105
8.1	Forward acceleration obtained from L3 region and forward acceleration obtained from the lateral side of waist. For the L3 region, acceleration is similar between left and right steps. Similarity between steps is altered when the sensor is located in the lateral side of waist.	115
8.2	Original and latent variables during both slow and fast gait. Amplitude of the latent variables during fast gait are higher than in the slow gait.	120
8.3	Eigenvalues of each latent variable. The first two latent variables comprise almost half of the variance of the trajectory matrix.	121
8.4	Latent space for fast and slow gait. The filtered reconstructed attractor is shown for both fast and slow gait. Different amplitudes are observed.	122
8.5	Recurrence plot obtained from the accelerometer signal registered during patient 1 gait. Gait cycles are reflected by long and uninterrupted diagonals	123

9.1	Accelerometer signal obtained during gait for three users and its reconstructed trajectories. Abstraction levels 3 and 10 are shown. Parameters used are $w=5$, $\tau=5$ and $m=10$	132
9.2	Gait recognition results for different values of w , a and τ . $w = 5, m = 5$ and $a = 10$ provide the best results.	134
9.3	Abstraction level effect for different values of w and τ . Increasing the abstraction level increases the accuracy until a certain level, after which the accuracy does not increase.	135
9.4	Two consecutive reconstructed trajectories for $a = 10$ and $a = 30$. No differences appear for $a = 10$ while for $a = 30$ some distinctions can be appreciated	136
9.5	Acceleration magnitude measurements and their reconstruction into 2-latent space by using the first and third dimensions, $m=20$ and a sampling frequency of 200 Hz.	140
9.6	Latent variables 1 to 12 in approximately phase quadrature for parameters $\tau = 5, m = 30$ and $N = 200$	141
9.7	Scree plots obtained from the accelerometer signals of 2 different volunteers using 3 different time lag values.	142
9.8	Reconstruction of the trajectory in the first and third dimensions of the latent space from the accelerometer signals of 2 different volunteers ($m = 20$; $N = 2000$, $N = 500$ and $N = 222$ for $\tau = 1, \tau = 4$ and $\tau = 9$, respectively).	143
9.9	Effect of τ and w in the gait recognition accuracy in 20 volunteers . .	143
9.10	Effect of τ and m in the gait recognition accuracy in 20 volunteers . .	144
9.11	Effect of w and m in the gait recognition accuracy in 20 volunteers . .	144
9.12	Effect of k and m in the gait recognition accuracy in 20 volunteers . .	145
9.13	Effect of k in the gait recognition accuracy for 4 combination of parameters	145
9.14	False Nearest Neighbors results for different embedding dimensions. Y axis represents the proportion of False Nearest Neighbours.	146

9.15 Average Mutual Information results for four different training signals. Best recognition rates are provided by $\tau = 2, 3$ and 4.	147
---	-----

Chapter 1

Introduction

This thesis seeks to further the analysis of human movement. The analysis is based on the use of accelerometers which - as the name indicates - measure acceleration. Thanks to the giant strides made in miniaturisation, these tiny devices are ideal for measuring human movements without inconveniencing the people to whom they are fitted. The study's contributions mainly fall under two heads: analysis of movement in Parkinson's Disease (PD) and analysis of gait.

The first head is focused on the use of accelerometers to analyse human movement in PD. PD is a neuro-degenerative disease that affects patients' movements. Sufferers display falling levels of dopamine, a hormone and neuro-transmitter needed to ensure proper control over one's body movements. As a result, PD patients suffer from a set of motor symptoms such as tremor, slowness of movements or bradykinesia, freezing of gait (FoG) and involuntary movements or dyskinesia. The first three symptoms mainly occur when medication proves therapeutically ineffective and is termed the 'OFF motor state' or 'OFF period'. The spells during which patients exhibit normal mobility are termed 'ON motor states' or 'ON periods'. Dyskinesia usually appears when patients reach peak-dose, which is when the level of medication in the blood hits a peak, and when patients flip from OFF to ON or from ON to OFF.

The main contributions of this thesis in the PD field are the development of algorithms to analyse movements in PD in order to detect dyskinesia and ON/OFF motor states. This contribution is of special relevance given that these symptoms

allow one to establish: (1) when the patient lacks medication (OFF state); (2) when the patient is in a normal state (ON state); (3) when the patient is over-medicated (dyskinesia). Knowing the cycle of medication surfeit and insufficiency through the application of the developed algorithms would help the doctor to adjust the medication doses throughout the day to minimise OFF periods and dyskinesia. The latest contributions in the treatment of PD are in this scope. They involve the modification of algorithms for detecting ON/OFF motor states and dyskinesia in real time. This opens a new way to treat the disease: by administering medication to patients using a sub-cutaneous infusion pump to ensure precisely-controlled dosing based upon the subject's current motor state. This approach has been tested in a pilot study, as have the real-time detection algorithms discussed in this paper.

Gait is the second field covered by this thesis. Gait analysis has many applications, including clinical ones, for diagnosing diseases affecting walking and optimising their treatment, and sporting ones, improving athletes' gait and performance.

This thesis has contributed to the accelerometer-based analysis of gait in two ways. First, it has examined gait from a biometric standpoint since walking can identify an individual in a similar way to fingerprints. Gait has also been analysed by means of accelerometers to identify speed and length of strides for clinical ends. The information provided by the developed algorithms could be used in the treatment of Parkinson's Disease and Diabetes, which could be improved through the use of movement sensors capable of measuring these parameters. The main theoretical contribution made by this study to gait analysis was the development of techniques based on the reconstruction of attractors for recognising individuals and for identifying gait parameters.

1.1 Purpose and framework

The research was carried out at the Technical Research Center for Dependency Care and Autonomous Living (CETpD). The CETpD is an applied research and technology transfer center in Vilanova i la Geltrú (Catalunya) and was set up by the Universitat Politècnica de Catalunya (UPC) and the Fundació Hospital Comarcal Sant Antoni

Abat (FHCSAA) to meet needs in the fields of: social robotics; user experience; smart settings; technology for the elderly and analysis of human movement.

Human movement is a complex phenomenon that is affected by many factors, including physiological, mechanical and psychological ones. Evaluating the quality of movement, monitoring its evolution and quantifying it yields fundamental knowledge in sundry medical disciplines that can be used in the treatment of various diseases and in rehabilitation [43]. Technological advances now make it possible to produce cheap miniaturised sensors, accelerometers, gyroscopes and magnetometers, that greatly facilitate such tasks. Given that these devices produce reliable information on movement, they are of great interest in facilitating the diagnosis of: osteo-arthritis [64]; obesity [30]; chronic pulmonary diseases [67]; back pain [115]; falls risk [21] and Parkinson's Disease [119]; among others.

Gait analysis have greatly benefited from accelerometers and gyroscopes. This kind of sensor offers a lower cost than classic laboratory gait analysis equipment for measuring posture and makes it possible to measure a patient's normal daily movements [59]. In the clinical field, gait has been studied for years in the evaluation and treatment of individuals with diseases that affect walking [122]. Treatment of many diseases and medical conditions could benefit from better clinical evaluation, including: cerebral palsy; Parkinson's Disease; muscular dystrophy; osteo-arthritis; rheumatoid arthritis; lower limb amputation; stroke; head injury; spinal cord injury; myelodysplasia and multiple sclerosis [122]. The sporting community has also shown a great deal of interest in studying gait to help athletes run faster and to identify posture and movements linked to injuries [121]. Last but not least, biometrics (the recognition of individuals through their physical traits or behaviour) has recently led to efforts to establish gait patterns that are unique to each person, in the same way as fingerprints are [40].

As noted earlier, the gait analysis carried out by CETpD also covers PD - which is the second most common neuro-degenerative disorder after Alzheimer's Disease. The symptoms produced by PD stem from falling levels of dopamine due to the death of nerve cells in the brain. Pharmaceutical treatment of the disease seeks to boost dopamine production through drugs such as levodopa. The main problem for PD

patients continually taking this or similar drugs over the years is flipping between ON and OFF periods. Each patient experiences the disease differently, given that the duration and severity of the symptoms in each motor state is unique to a given individual. A patient with moderate to severe PD may flip between ON/OFF motor states three or four times a day. Thus patients need to learn to organise their daily activities depending on these flips and the medication intakes.

The techniques developed in this study will facilitate:

- Tailoring drug dosing to individual needs and thus improve treatment, avoiding or cutting the length of time the patient is in an OFF motor state.
- Treatment of PD using a sub-cutaneous infusion pump that administers a drug depending on the patient's motor state and symptoms at any given moment.

The CETpD has been researching on human movement since 2009 to achieve its objectives. To these ends, it is carrying out various research projects:

- The first research project related with this thesis was 'Mobility Monitoring of Parkinson's Patients' (MoMoPa, PI08/90756), covering Spain. It began in 2009 and ended in mid-2011. This project marked the beginning of CETpD's research into PD and provided the first results in the detection of ON/OFF motor states, dyskinesia, walking parameters, freezing of gait and falls in 35 patients with PD by means of inertial sensors taped to patients' bodies.
- Another project covered the second objective and was titled 'Home-Based Empowered Living for PD patients' (HELP, AAL-2008-1-022). The project was funded as part of the EU's Seventh Framework Programme through Ambient Assisted Living (AAL) and was scored as the best project submitted for the call. It began in mid-2009 and ended in March 2013. The project tried a new form of treating PD based on subcutaneous pump injection of a drug. The dose administered depended on patient's motor state and the presence of dyskinesia, which was detected by the algorithms included in an inertial sensor fixed to the patient's waist. The real-time detection algorithms were developed as part of this work and are shown in the part covering contributions.

- The third project related with this thesis is ‘Personal Health Device for the Remote and Autonomous Management of PD’ (REMPARK, FP7-ICT-2011-7-287677), begun at the end of 2011 and planned for completion in 2015, continues CETpD’s research on PD. This project will further develop the algorithms given in this thesis and seek new ways of detecting other symptoms such as FoG and dyskinesia. It will thus advance PD treatment by giving two more levels of control. REMPARK plans to develop a closed loop system that acts on the patient in real time and in a similar way to the HELP system. Second, REMPARK plans to control the long-term development of the disease by identifying trends in the measurements yielded by the detection algorithms.
- Finally, the gait analysis for the identification of gait parameters that has been done in this thesis is part of the SENSORIAL project (TIN2010-20966-C02-02) funded funded by the Spanish Ministry of Education and Science.

1.2 Objectives

The main objective of this thesis has consisted in developing algorithms for treating accelerometer signals with the aim of extracting gait parameters, identifying individuals through gait recognition and detect dyskinesia and motor states in PD patients. The following sub-objectives were set to achieve these ends:

- To study and develop algorithms for analysing the accelerometer signal to detect ON/OFF motor states.
- To study and develop algorithms for analysing the accelerometer signal to detect dyskinesia.
- To adapt the algorithms developed to detect dyskinesia and motor states in order to obtain a real-time version of them and to implement the algorithms in a small memory computer processor.
- To study and develop algorithms for analysing the accelerometer signal to identify gait parameters.

- To study and develop algorithms for analysing the accelerometer signal to identify individuals from their gait.

The first, second and third objectives, bearing on PD, are covered in the first part of this thesis. The fourth and fifth objectives, bearing on gait analysis, are covered in its second part. The last chapter of this thesis is devoted to the conclusions. The rest of this section gives an introduction to Parkinson's Disease and details gait characteristics.

1.3 Motor Symptoms of Parkinson's Disease

This section presents Parkinson's disease impact on the society and its main causes, symptoms and current treatments.

1.3.1 The disease's impact

Parkinson's is the second most common neurological disease after Alzheimer's. PD is the result of degeneration of the neurons producing dopamine, a hormone and neuro-transmitter required for proper control over human movements. Although it may occur at any age, it is extremely unusual in people under thirty and only in ten per cent of cases it does begin before 40. In Europe, its annual incidence is 10-20 cases per 100.000 inhabitants per year and it has a prevalence of around 1% in the population aged over 65. The World Health Organization states that 5.2 million people worldwide suffer from PD of which over 2 million are in Europe [85].

PD is fast becoming a major health problem given that it mainly affects the elderly, which make up an ever-larger part of the age pyramid [2]. The costs of health care for those with PD are considerable. It is conservatively estimated (after taking into account productivity losses) that the total cost comes to some €20,000 million a year. Around four fifths of this stems from indirect costs. Hospital costs and prescription drugs are the main items of expenditure and average 30% and 21% of direct costs, respectively [77] [36] [65].

1.3.2 Parkinson's Disease: causes, symptoms and present treatment

Our movements are controlled by nerve cells sited in the basal ganglia region of the brain. When we want to move, these cells send messages via chemical neurotransmitters in the form of electrical impulses through the nervous system. Dopamine is the main neuro-transmitter in the basal ganglia. These messages are efficiently sent and received in healthy people. However, in PD patients, lack of dopamine causes interruption of these messages, with the consequent loss of control over muscle movements. In PD patients, between 70 and 80% of dopamine-producing cells, found in a small region of the midbrain called the substantia nigra, are damaged.

Unfortunately, oral administration of dopamine has no effect given that the substance does not pass to those areas of the brain where it is needed. The current treatment is based on drugs that can be absorbed by the organism and then converted into dopamine. Levodopa is one of these drugs and is the commonest treatment for PD. Dopamine antagonists such as bromocryptine and apomorphine are also used. The last of these is usually administered subcutaneously and is quickly absorbed. There are also other treatments, such as surgical ones in which a nerve stimulator is implanted to electrically stimulate those regions of the brain where PD symptoms are found [18].

Over the longer term, especially from the fifth year of the disease onwards, motor complications arise from both the course of Parkinson's and the continuous use of drugs stimulating the use of dopamine [88]. These motor complications may give rise to:

- On the one hand, medication has an ever shorter-lived effect, giving rise to the so-called 'wearing-off' effect [23]. This forces patients to raise the number of times they take medication and to split doses.
- On the other hand, other phenomena such as dyskinesia, freezing of gait (FoG), tremor and bradykinesia appear:
 - Dyskinesia is characterised by excessive movement such that the patient

cannot avoid moving his limbs repetitively in a way that makes walking, or indeed any other activity, virtually impossible.

- A further symptom is freezing of gait (FoG) which, as the name suggests, involves the patient freezing up and being unable to continue walking or make any other movement. FoG can last minutes, cause falls and lead to high anxiety and greatly worsen patients' quality of life [81].
- Bradykinesia is revealed in slowness of movements. Walking is the activity where this is most evident but it can also be seen in other exercises such as finger-tapping.

These problems seem to be linked to either a rise of pro-dopamine drug levels in the blood, in the case of dyskinesia, or a fall in those levels in all other cases.

The main problem patients suffer from after taking medication for several years is flipping between ON/OFF motor states [88]:

- During ON periods, patients are able to control their movements relatively easily. During ON states, symptoms such as bradykinesia, rigidity and dyskinesia are very slight and they are only noticed by clinicians. Dyskinesia is the only evident symptom of the disease during ON states.
- During OFF periods, patients may suffer from dyskinesia, bradykinesia, FoG, shortened stride, rigidity, lack of muscular co-ordination and pain. The range and severity of symptoms during OFF states differ greatly among patients.

In addition to motor symptoms, patients could also suffer non-motor symptoms such as dementia, depression, anxiety and sleeplessness [88].

For years now, it has been argued that the best way to treat PD is to either keep drug levels in the blood constant or to match them to the patient's activity. Various treatments have been developed to one of these two ends [4]:

- The first treatment was the intake of pills every few hours in order to keep blood medication levels constant.

- An innovative step was the use of rotigotine, a dopamine antagonist, skin patches to ensure continuous drug administration over a 24-hour period. However, patients' mobility needs vary during the day, which means the dose is too small at certain times and too great at others.
- Another innovation is occasional injections of dopamine antagonists, which have a quick effect on patients with sudden OFF symptoms. The most commonly-used antagonist is apomorphine, given that it is rapidly absorbed. However, it requires either the patient or his carer to inject the drug.
- A recent new advance in this field has been the introduction of programmable infusion pumps which patients can carry with them. This kind of pump is programmed to ensure steady dosing of apomorphine throughout the day. The main advantage of this system is that the patient may administer a booster dose called *bolus* over a short period of time to obviate the need for carer-administered injections, as is the case with dopamine antagonists. The drawback is that the doses given by the pump are set when the neurologist sees the patient, which is usually once every three or six months [1]. Accordingly, the pump cannot be regulated to meet patients' needs at a given time, save in the case of the occasionally administered bolus shots.

Hence the appearance of self-programming infusion pumps, which would cater to the patient's needs at any given time, would be greeted from both clinicians and patients as a great advance in controlling PD symptoms. This idea was tried out for the very first time in November 2012 within the HELP project. The algorithms for detecting symptoms and motor states used in the test with patients were developed as part of this thesis.

1.4 Human gait

Gait, as noted earlier, has been studied from three perspectives: clinical practice, sports and biometrics. This section describes the key concepts of gait that will be used throughout the study.

1.4.1 Human gait analysis

The simplest and most commonly used definition of 'gait' is 'the way someone walks'. This simple definition is the most widely used because it is neither unduly restrictive, for example, it includes walking with the aid of crutches, nor too broad, it excludes running.

Man's gait is currently analysed from the standpoint of its phases. The first studies, which date back to the Renaissance, focused on establishing the body's center of mass during walking. The next remarkable study was carried out by the Weber brothers in 1836 and analysed gait with precise descriptions of each of its phases. At the end of the 19th Century, Braune and Fischer produced the first study to precisely define the legs' trajectory, speed and acceleration during walking. Furthermore, by including the masses in them, calculations were made of the forces involved during each phase of walking. Once the forces had been ascertained, walking was analysed in muscular terms, the most important work in this field being carried out by Verne Inman of the University of California during the 1940s and 50s [55] [56]. In the 1960s, researchers focused their efforts on the forces acting on joints during walking [104]. Over the last few years, measurement methods have taken a giant leap forward, with Motion Capture (MoCap) systems [80] and the use of MEMS technology [76] playing a major part.

1.4.2 Gait cycle

The gait cycle - also called stride - is defined as the time elapsing between successive repetitions of the same event [122]. Although any event may be chosen to define the walking cycle, *initial contact*, that is, when one foot first touches the ground, is the most commonly used. For example, if the initial contact begins with the right foot, the cycle will end when the right foot touches the ground again. The left foot goes through the sequence but remains out of phase with that of the right foot by half a cycle.

Walking comprises two phases: the stance phase, which is also termed the support phase or contact phase; and the swing phase, during which the suspended foot moves

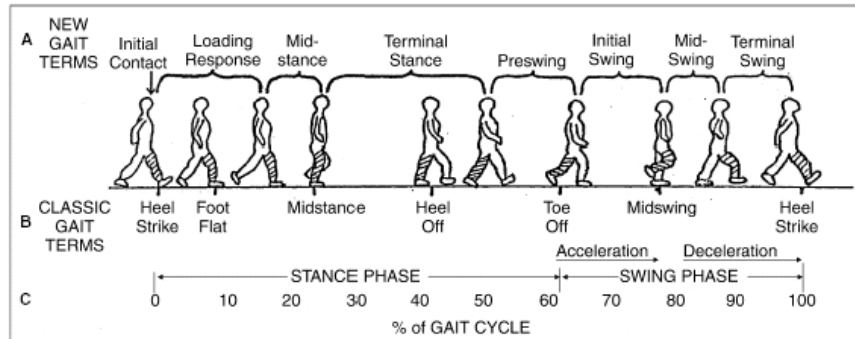


Figure 1.1: Gait cycle phases and events [32].

forward [32]. Figure 1.1 shows the different phases and events of gait [32].

The stance phase may be sub-divided into the following stages:

- Initial contact - the instant the foot touches the ground. This instant is also called Heel Strike, although this term is somewhat inaccurate given that in many cases initial contact is not made with the heel.
- Loading response - the period following initial contact in which the impact is cushioned and ends with the alignment of the ankles.
- Mid-stance - the period during which the ankles are aligned.
- Terminal stance - the period during which the ankles remain aligned just before the initial contact of the other foot.
- Pre-swing - the period that runs from when the other foot touches the ground until the foot under consideration lifts off the ground.

The swing phase comprises the following sub-phases:

- Initial swing, which begins when the foot lifts off the ground and ends when the knee reaches maximum flexing.
- Mid-swing, which follow the initial swing and ends with the tibia vertical.

- Terminal swing, which begins with the tibia in vertical position and ends the instant before there is initial contact.

The following terms, in order of appearance, are used to identify the phases of the walking cycle:

- Initial contact,
- Opposite toe off, the moment when the other foot lifts off the ground.,
- Heel rise, the instant in which the reference foot lifts off the ground,
- Opposite initial contact, which is the initial contact of the other foot,
- Toe off

The nomenclature used to describe walking cycles varies from one publication to another. This study will mainly use the terms presented in the beginning of this section: phase swing, phase stance and the initial contact.

Part I

Dyskinesia and ON/OFF motor states detection in Parkinson's Disease patients

In this part of the thesis different chapters are presented, which are devoted to describe the state of the art and the contributions done in the field of accelerometry signal analysis for the dyskinesia and ON/OFF motor states detection in PD patients. The contributions presented are, mainly:

- New signal analysis methods for the detection of dyskinesia and motor states developed through the movement signals obtained from 20 patients in a controlled environment.
- Offline validation of the previously developed methods by means of the movement signals obtained from other 15 PD patients in a non-controlled environment consisting in the patients' home and its surroundings.
- Online validation of the algorithms that were adapted to work in a real-time fashion. The online algorithms were used in the first pilot test in which a subcutaneous apomorphine pump was used to automatically administer personalized apomorphine doses according to the patient's motor state.

This part of the thesis is divided into four chapters:

- Chapter 2 presents the state of the art in inertial sensors and their use in motor state and dyskinesia detection.
- Chapter 3 describes the protocol and data gathering with PD patients for the development and validation of the algorithms.
- Chapters 4 and 5 detail the offline and online methods developed for dyskinesia and motor state detection, respectively. Results and discussions are also provided.

Chapter 2

State of the art in dyskinesia and motor states detection in PD patients

This chapter firstly justifies the need of ambulatory dyskinesia and motor states detection in PD. Then, it provides a review of the sensors currently used for the ambulatory monitoring of human movement: inertial sensors. In addition, recent studies related to dyskinesia and motor states detection and the signal analysis performed in these studies are presented. Finally, a short discussion is given.

2.1 The importance of detecting and monitoring motor states and dyskinesia

An ambulatory and continuous monitoring of PD in terms of motor states and dyskinesia is the main aim of the first part of this thesis. Such a monitoring, based upon objective measurements, would provide a detailed report on patient's motor states and would constitute the following three advances:

- It would allow the doctor to tailor drug dosing to each patient in an objective way. Current methods of monitoring PD over the long term are based on

the patient noting down his motor state in a diary. However, self-evaluation by patients poses various problems and is notoriously unreliable [118]. A system capable of providing reliable, continuous information to the neurologist on the evolution of the disease will help him/her change therapeutic protocols to achieve better treatment results.

- If detection of the sufferer's motor state and symptoms is carried out in real time, it would be possible to treat the patient using an infusion pump with doses finely adjusted to his needs at any given moment.
- It would give specialists an objective measurement tool that would be invaluable in evaluating new pharmaceuticals. Currently, new pharmaceuticals used for PD are evaluated by measuring the extent to which they reduce OFF time, which is logged by patients themselves in their diaries [13]. An objective measurement tool for this purpose would provide much more reliable information on the efficacy of such drugs.

In current clinical practice, there are several methods for objectively evaluating the motor state of a patient with Parkinson's Disease. For example, the UPDRS measures the patient's behaviour, mood and movement and, among other things, yields an index of the patient's disability [73]. Other more specific scales are the *Unified Dyskinesia Rating Scale* (UDRS), which evaluates the patient's involuntary movements. However, such methods can only be used under the supervision of a trained clinical observer [44]. Furthermore, the information gathered only gives a snapshot of the disease yet practical treatment requires continuous observation of motor states and the flips between them.

2.2 Inertial sensors

Inertial sensors such as accelerometers, gyroscopes and magnetic sensors have been gradually introduced for ambulatory monitoring of human movement [43]. These sensors use *Micro Electro-Mechanical Systems* (MEMS), a technology that can be defined as miniaturized mechanical and electro-mechanical devices and structures that

are made using microfabrication techniques. The physical dimensions of MEMS devices can vary from one micron to several millimeters. Not only the size of MEMS is an interesting feature, but their production methodology leverages the same fabrication techniques used in the integrated circuit industry. Consequently, it is possible to achieve high performance, small size, light weight and low consumption at relatively low cost. Currently, there are MEMS sensors which measure acceleration, rotation, magnetic field, pressure, temperature, proximity, radiation and acoustic emission, among others. They are a crucial component of many consumer electronics, such as mobile phones or energy expenditure monitoring devices. MEMS devices are widely extended in the automotive industries, for instance, in suspension systems or in the crash detection performed in the airbag systems. They are also used in medical applications, for instance to monitor blood pressure, to measure respiratory capacity and as a part of hemodialysis equipment. In the military and aerospace industries, they are used in guidance and navigation systems.

Accelerometers are commonly used sensors in human movement studies [43] [52] [120] [9]. There are different types of accelerometers according to their sensing principle. Capacitive accelerometers are the most common type. They have two planar surfaces, one of them charged with an electrical current. Changing the distance between the plates changes the electrical capacity of the system, which can be measured as a voltage output. This method of sensing is known for its high accuracy and stability. Other types of accelerometers are piezoelectrics, piezoresistive, magnetoresistive, based on Hall effect, heat transfer, etc. However, capacitive accelerometers typically have higher performances, especially in the low frequency range. They are also less prone to noise and variation with temperature, typically dissipate less power, and can have larger bandwidths due to internal feedback circuitry.

Gyroscopes are devices for measuring orientation. They use vibrating mechanical elements to sense rotation and are based on the transfer of energy between two vibration modes of a structure caused by Coriolis acceleration (the drive and the sense modes) [12]. The angular velocity of the sensor is determined by measuring the amplitude in the second mode (sense mode). The resonant frequencies of these two modes must be very close to each other for the gyroscope to work at resonance for maximum

sensitivity. However, achieving its maximum is very difficult due to tolerances of the manufacturing process. This is why gyroscopes are very sensitive to the manufacture conditions and packagings. For this reason, and taking into account that the Coriolis force is extremely small, gyroscopes need a more complex signal conditioning stage than accelerometers.

Magnetometers are sensors used to measure the strength (scalar magnetometer) or the direction of the magnetic field. This type of sensors relies on the mechanical motion of the MEMS structure due to the Lorentz force acting on the current-carrying conductor in the magnetic field. The mechanical motion of the micro-structure is sensed either electronically or optically [63]. When applied to human movement monitoring, usually three magnetometers in mutual perpendicular directions measure the orientation with respect to the magnetic North. The main problem is that magnetometer signals are strongly altered by ferromagnetic materials as metal beams, which limits its usefulness in indoor spaces given the noise added in the measurements obtained.

In this thesis, we are focusing on accelerometers owing to the considerable advantages referred to above. Concretely, accelerometers do not have problems in indoor spaces and they do not need excessive signal conditioning. These advantages justify why most of studies devoted to human movement analysis using inertial sensors were performed using accelerometers.

2.3 Accelerometer signal analysis for ON/OFF motor states detection

The research conducted so far on the ON/OFF motor states detection by means of inertial sensors is mainly based on the patient symptoms characterization. This line is followed by the first work conducted in 2004 by scientists of the Medical Center at Leiden University [52]. This study focused on determining the state motor of 50 patients through the characterization of hypokinesia, bradykinesia and tremor, which are closely related to the motor state. To this end, two triaxial accelerometers located

on the wrists are used and three symptoms are characterized through three signal features: time the patient spent immobile, mean acceleration and percentage of time with tremor respectively, each calculated in intervals of 30 minutes. A linear classifier was used to determine the patient's motor state, so that a threshold is applied to each feature and, thus, both motor states can be separated. Results demonstrate a relatively good performance since they achieved a specificity value (accuracy rate in ON detection) equal to 0.7 and a sensitivity value (accuracy rate in OFF detection) close to 0.7 too.

Researchers at the *École Polytechnique Fédérale de Lausanne* evaluated the use of 7 gyroscopes and 2 accelerometers located on the forearms, shins and trunk detecting motor states [15]. They evaluated the use of 15 parameters or subsets of them representing the presence or absence of tremor, bradykinesia, postural transitions, body posture, and gait parameters to determine the motor state of 13 patients using logistic regression. They used 10-minute windows in tests lasting from 3 to 6 hours. Results showed that, for a subset of 9 features, a specificity of 0.9 and sensitivity of 0.76 were achieved.

A distinct tendency for the evaluation of patient symptoms involves scoring patient response to some exercises similar to those performed in the evaluation of the *Unified Parkinson's Disease Rating Scale* (UPDRS) scale from the sensor signal. The UPDRS score provides a value that measures the severity in which PD is affecting a patient at the time of evaluation. This way, motor fluctuations can be monitored because there is a close relationship between the UPDRS value and ON/OFF motor state [109]. For instance, Kinesia device, developed by the American company Great Lakes Technology, consists of a triaxial accelerometer and a gyroscope which measure patient's finger movement. This device is capable of providing measures correlated with UPDRS [78]. Another example is the work of researchers from Harvard Medical School, who used 9 accelerometers to estimate the value of UPDRS in different hand, feet and heels exercises [86]. The latter two methods have the disadvantage of requiring the patient to perform specific exercises other than those usually performing during daily life. This way, the capabilities of capturing the rapid fluctuations of the disease are limited. Furthermore, these experiments force patients to stop their

normal activity.

2.4 Accelerometer signal analysis for dyskinesia detection

Most sensors used in the detection of dyskinesia are also accelerometers, although some works also used other sensors such as gyroscopes and magnetometers [41] [46]. Since this symptom can occur both in the upper and lower extremities and in the trunk and the head [19], in order to detect dyskinesia, accelerometers have been placed in those parts of the body, particularly on the trunk and extremities in works [60] [51] [113] and shoulder in the study [69].

One of the most commonly used signal analysis methods in the dyskinesia detection is the frequency analysis. Several studies have shown that the power spectrum corresponding to the band from 1 to 3 Hz of the accelerometer signal increases if dyskinesia appear [69] [60]. In another study, conducted at the Leiden University Medical Center (Netherlands), however, it is considered the power of the harmonics of less than 8 Hz to be corresponded to dyskinesia [51]. As detailed in the work from Manson et al. [69], which performed at the Hospital of Queen Square (London), one of the problems of analyzing the power spectrum of a particular band is that it can allow the occurrence of false positives since other activities also increase the power of that frequency band. For example, the activities of walking and stairs climbing and descending, measured by a sensor on the waist provide a signal that increases its power at frequencies up to 20 Hz. This will produce false positives if the dyskinesia detection was just associated to the bands described above. That is why, for example, in the study performed by Manson, signals obtained during walking were excluded from the dyskinesia detection analysis [69], preventing the occurrence of false positives.

In the study conducted at the University of Nijmegen (Netherlands) [60], power spectrum of the dyskinesia band was combined with other features not related to the signal frequency and that have proven useful in detecting dyskinesia. These consider the amplitude of the movements of the limbs [41] [46] [29] [113], or consist of statistic

values calculated from the derivative of the acceleration respect to the derivative of the acceleration with respect of time or jerk, as the mean value, the deviation, etc. [60].

Classification in dyskinesia presence or absence of the values extracted by previous methods has been carried out mainly through supervised learning methods: Artificial Neural Networks [60] [29] [113] and linear classification [69] [51]. The remaining studies used statistical tests to assess the ability of dyskinesia detection [41] [46].

2.5 Conclusions

This chapter has presented current approaches followed to ambulatory monitor motor states and dyskinesia in PD patients. As it has been shown, the signal processing applied to the inertial sensor signals mainly consist in characterizing symptoms through some specific features, both in temporal and frequency domains, that aim to represent patient's movement and, consequently, enable the identification of symptoms.

As it has been described, an objective system able to report patient motor fluctuations and the appearance of dyskinesia would improve the disease management. In chapters 4 and 5, the development of specific algorithms that aim to provide this objective system are presented. The next chapter presents the data collection with PD patients.

Chapter 3

Data collection with PD patients for the development of dyskinesia and motor states detection algorithms

This chapter describes the sensor devices and the tests that PD patients performed that enabled the data gathering for the development and validation of the dyskinesia and motor states detection algorithm. First, the location used by the patients to wear the sensor device and its components are described. The sensor and location of the sensor were fixed in a study previous to the thesis. Second, a movement protocol is described, which was followed by patients and enabled the development of the dyskinesia and motor state detection algorithms. This protocol is part of the first stage of MoMoPa project and consists of a set of activities performed in a laboratory by 20 patients. Next, the validation framework for the developed algorithms is presented which comprises, on the one hand, the second stage of MoMoPa project as the offline validation framework and, on the other hand, the HELP clinical study as the online validation framework. The offline validation consisted on the accelerometer signal acquisition from 15 new patients in their home during several hours. Finally, the online validation framework enabled the testing of the online adapted algorithms

and, at the same time, enabled the validation of a new treatment for PD patients, as it is detailed in the following sections.

The algorithms developed through the clinical protocols presented in this chapter were created by the author of this thesis as his main role in both MoMoPa and HELP research projects. Data used to develop and validate the algorithms was gathered by a team of engineers, doctors and nurses from UPC and Hospital Sant Antoni Abat.

3.1 Sensor device: position and components

According to the literature, in previous studies on symptoms and motor state detection in PD patients, as well as gait analysis studies, human movement has been analyzed by means of inertial sensor from many different body locations. Concretely, sensors have been worn in the chest [60] [51] [113], arms [60] [51] [113], shoulders [69], ankle [66], foot [17] [90], shanks and thighs [10] [60] [51] [113] and on the waist or close to the CoM [126] [9] [72]. This last location is commonly the recommended to monitor human movement [106] [75] since it provides reliable information [124] with the exception of specific movements from arms and legs [74].

Previously to this thesis, a study in the CETpD was carried out to establish the most suitable position to locate the sensor in order to measure PD patients movement¹. This position should enable the detection of the different symptoms and motor states of PD patients but, at the same time, should enable wearing the sensor in a comfortable way in order to not disturb the daily life activities. In the study, ten PD patients wore the movement sensor in 5 different positions: on the lower part of the illiac crest, on both shanks and on both foot insteps. Patients chose the location they felt more comfortable and less disturbing for the daily life. All patients chose the lower part of the illiac crest, which is the upper border of the pelvis major bone, in a completely lateral position to the trunk.

In this manner, given the result of the study and the literature reviewed, previously to this thesis the sensor position was established as the lateral side of the waist. This position has been used to locate the movement sensor in all the tests with PD patients.

¹Study carried out as part of MoMoPa project



Figure 3.1: Movement sensor used and its position in a neoprene belt.

This position enables each accelerometer axis to match, at least approximately, with some of the following directions during a straight walk: frontal acceleration (X axis), vertical acceleration (Y axis) and lateral acceleration (Z axis).

Regarding the sensor used to record inertial signals, it is a prototype of inertial measurement unit developed at the Universitat Politècnica de Catalunya, which includes triaxial sensors to record acceleration, angular velocity and magnetic field. Figure 3.1 shows the sensor position and location, which is in a neoprene belt. The triaxial sensor of lineal acceleration (LIS3LV02DQ, ST Microelectronics) can be used to measure acceleration magnitudes up to 6 g ($1 \text{ g} = 9.81 \text{ m/s}^2$), with 2.94 mg sensitivity and 2% maximum nonlinearity on the full scale range. The inertial device is "wearable" and equipped with data processing and wireless information transfer capabilities and rechargeable batteries.

3.2 Data framework for offline dyskinesia and motor states detection algorithms

This experimentation was performed in two different stages. The first stage was used to develop the detection algorithms and the second stage to validate the offline algorithms. During the first stage, a database comprising inertial signals obtained in a laboratory was obtained from 20 PD patients wearing the previously described sensor in the lateral side of the waist. On this basis, algorithms were developed to detect dyskinesia and motor fluctuations (ON/OFF states) in patients with Parkinson's

disease. During the second stage, the developed algorithms were validated at patient's home by using a new sample of 14 Parkinson patients wearing the sensor device. This second group of patients did not participate in the first phase.

The study was carried out during 2009 and 2010 with patients diagnosed idiopathic Parkinson's disease according to the criteria of the Brain Bank, London, aged between 55 and 75 years and living in the province of Barcelona, Spain. During the second stage, only patients with moderate-severe disease and motor fluctuations were included, while during the first stage, patients with milder forms of disease were allowed to participate.

The first stage of the study was conducted in the facilities of the Universitat Politècnica de Catalunya and two regional hospitals: Consorci Sanitari del Garraf and Consorci Sanitari del Maresme, all of them located in Barcelona (Spain). The second stage was conducted at the patients' homes. All data were collected by the same researcher team (a doctor, two nursing assistants and three engineers) who had been specifically trained in the study procedures and administration of the corresponding questionnaires on the basis of common instructions and clinical guidelines.

Data acquired during the first stage was obtained by means of two different movement protocols:

- The first 10 recruited patients participated in the experiment under controlled conditions; they were asked to walk 5 meters straight (3 times) while wearing the inertial sensor device attached to the waist. The gait speed and stride length of these patients was obtained. Patients who presented motor fluctuations repeated this experiment both while in the ON and OFF states (OFF state was induced by withdrawal or reduction of the patient's usual dopaminergic morning medication). These 10 patients wore 4 inertial sensors, in addition to the waist sensor, that were located in each shank and each instep. These extra inertial sensors were synchronized by Bluetooth with the waist sensor and were used to validate the step detection, as it will be shown in the following subsections.
- The remaining 10 patients participated in a mixed experiment under controlled and uncontrolled conditions. They were asked to complete a movement circuit

in the laboratory, which included walking straight, walking up and down stairs, walking up and down inclined planes, turning, taking different positions such as sitting, standing up or lying down, walking while carrying a glass of water, walking while carrying a heavy object and other more complex activities, such as setting the table for a meal. After that, they took a 15-minute walk outside the involved facilities.

During the second stage of the study, patients were asked to wear the sensor device during 3 to 5 hours, while engaged in their usual activities, in an environment that was familiar to them and therefore uncontrolled (their home, neighbourhood or habitually visited places). All patients participating in this stage had moderate-severe disease with freezing of gait episodes and/or motor fluctuations. During this stage, the OFF state was not induced by modifying the medication schedule; instead, OFF data were collected during naturally occurring OFF-state episodes. Data collected during the second stage of the study were used to validate the processing algorithms created during the first stage.

Video-recording was the gold standard during the initial period of stage 1 (laboratory) to identify patients' movements and positions, as well as ON or OFF state episodes. Additionally, patients were asked to confirm the occurrence of their ON and OFF episodes. Recorded inertial data were then labelled as corresponding to ON state, OFF state and with/without dyskinesia.

Gait speed and stride length of the 10 first patients from the first stage were obtained by means of two methods. First, subject's shoe soles were painted to leave footprints over a 6-meter-length paper placed over the ground, so that step length was provided. The process was recorded by a video-camera in order to obtain step duration time and, consequently, step velocity. The video-camera was also used on the second method in order to obtain the velocity and length of steps since, in this case, visual markers were placed on the ground along the walking space every 15 cm. The signal processing algorithm developed to obtain the gait properties is presented in the next part where gait analysis methods are presented.

The gold standard during the experimental stages conducted under uncontrolled conditions (on the street, at home, . . .), was a trained observer who accompanied the

patient all the time and recorded the presence or absence of dyskinesia as well as the ON/OFF motor state. Data were recorded in real time with a JAVA-based software developed ad-hoc by the researchers of the Universitat Politècnica de Catalunya and installed on a tablet where the following patients' symptoms and events were recorded: ON or OFF state, dyskinesia, FoG, festination, medication intakes, daily activities, falls and loss of balance.

3.3 Data framework for online dyskinesia and motor states detection algorithms

Algorithms developed through the signals gathered in laboratory conditions, as described in Chapter 2, were validated using the signals obtained in the patients' home. This validation was done, consequently, in an offline way since data was registered in a digital support. The offline algorithms, as presented in the next chapter, were posteriorly adapted in order to be executed in an online manner in the sensor device. The online algorithms were validated in the clinical study HELP. In this study 7 PD patients participated: 2 patients were from Madrid, 4 from Catalunya and 1 from Tel Aviv, Israel. This section describes this clinical study that enabled the validation of the real-time algorithms.

The clinical study for the online validation of the algorithms has different objectives of efficacy, security and validity. The most important objective consists in measuring the validity of a new treatment for PD based on automatically regulate the dose administered by a subcutaneous apomorphine pump according to the patient's motor state and the presence of dyskinesia. Another important objective, for this thesis purposes, is the validity measurement of the real-time dyskinesia and ON/OFF motor states detection algorithms. The rest of the objectives consist, first, in evaluating whether the treatments provokes side effects or no and, second, in evaluating the system's usability.

A set of devices and web applications were developed for the HELP study. This set of tools will be denoted as *HELP system*, which is composed of:



Figure 3.2: Apomorphine pump used in HELP project.

- A sensor device, which implements the real-time algorithms developed in this thesis and which is in charge of evaluating patient's state.
- A subcutaneous apomorphine pump whose dose is regulated in real-time by a clinician through a web application. Figure 3.2 shows the pump used in HELP project.
- A server that provides the web application for clinicians. Clinicians are able to follow patient's evolution and their current state in this web application according to the data sent by the sensor. Moreover, clinicians may remotely change the apomorphine dose that the pump administers to the patient.
- A mobile phone which communicates the previous three elements. The mobile phone connects the sensor device and the pump through ZigBee and it sends and receive data from the server through 3G connection. The mobile phone has a router functionality: on the one hand, it sends sensor's data to the server while, on the other hand, it receives the apomorphine pump doses decided by clinicians and resends them to the pump.

According to the study inclusion criteria, patients willing to participate must already be under a subcutaneous apomorphine pump treatment or must be about to

start the treatment. The reason is that a previous apomorphine's dose fixed by a neurologist must be available. In the current treatments, patients wear the apomorphine pump during several hours a day and the dose always remains constant. In the new treatment, the apomorphine's dose is changed between a *high* dose, which should allow the patient go from an OFF period to an ON period, and a *low* dose, which should allow the patient remain in an ON state and, at the same time, avoid dyskinesia. The dose that the patient receives previously to enter the clinical study becomes the low dose. The high dose is established as the maximum dose that the patient tolerates without suffering adverse effects up to a 50% increase of the low dose.

The HELP clinical study is divided into three phases:

- During the first phase, which lasts a week, the patient follows the treatment that has before entering the clinical study. During this week, the patient wears: 1) the apomorphine pump that was using before the study with the same dose (low dose), 2) the sensor device and 3) the mobile phone. Thus, the sensor's online algorithms evaluate patient's movement and the outputs are sent in real-time to the server. Moreover, during this week the patient fills a diary with the actual motor state and the presence of dyskinesia in each hour of the day. This diary will enable the performance validation of the algorithms.
- In the second phase, patient is hospitalized during two days. The goal is twofold: first, the high dose is found, and second, the detection algorithms are tuned. As it will be shown in the next chapters, the ON/OFF motor state algorithm must be personalized, which requires the patient to walk short distances in both ON/OFF motor states. Regarding the high dose, it is found by progressively increasing the low dose a 10, 20 and a 50%. The high dose is the highest dose which does not provoke adverse effects (nausea, vomits, etc.).
- The third phase has, again, 1 week length. During this phase, the patient follows the new treatment, so the remotely regulable pump administers a dose which alternates according to the patient's motor state and dyskinesia. Furthermore, during this week the patient fills a diary with the actual motor state and the

presence of dyskinesia in each hour of the day, which will enable the performance validation of the algorithms.

The study design makes clear whether the new treatment improves patient's quality of life or not. This is the fundamental variable to measure and its improvement requires the treatment to reduce the time patients spend in OFF state and not to provoke adverse effects.

3.4 Conclusions

This chapter has presented the sensor and the position in which it is worn by the patients. The sensor, which is able to sample acceleration, angular rotation and magnetic field at 200 Hz, is worn in the lateral side of the waist inside a neoprene belt.

This chapter has also described the data collection with PD patients. The data collection is divided in two parts: the protocol followed for the creation and validation of offline dyskinesia and motor state detection algorithms, which was both in controlled and uncontrolled environments; and the protocol for validating the online algorithms. The data collection was performed by a team of doctors, nurses and engineers from UPC and Hospital Sant Antoni Abat.

Chapter 4

Dyskinesia detection

In this chapter, the contributions made in the dyskinesia detection field are presented. The methods developed for the accelerometer signal treatment are described, as well as their results in the data collected, according to the previous chapter.

As it has been described previously, different methods may be found in the literature for detecting dyskinesia in PD patients. These methods, however, have been only tested inside controlled environments, that is, in a laboratory. The thesis' main contribution in this field consists in extending the dyskinesia analysis to the patients environment. This contribution has been possible because of the data gathered under MoMoPa project. Thus, the main contributions in the dyskinesia detection field are:

- The development and evaluation of a dyskinesia detection algorithm in the daily life environment of patients.
- The development of an algorithm able to detect dyskinesia in real-time.
- The identification of the daily life activities able to provoke false positive in the dyskinesia detection.

The work presented in this chapter has been published and presented in the *Engineering in Medicine and Biology Society Annual Conference of the IEEE* (San Diego) in 2012 [99]. Moreover, the results of the dyskinesia detection, as well as the motor state ones, have been recently submitted to the 'Movement Disorders' journal.

The contributions made in the dyskinesia field are relevant since the detection method developed has been tested against acceleration signals obtained from PD patients in non-controlled environments, which was possible due to, essentially, the sensor device that was designed to be comfortably wearable. In contrast, the dyskinesia detection algorithms available in the research literature have been tested in controlled environments in which the activities performed by the patients were limited to those listed in the study protocol. The algorithm presented in this thesis, instead, was tested while patients performed their daily life activities since the data collection took place in the patients' home. This way, signals were collected while patients were driving, shopping, cooking, cleaning, reading, playing chess, etc. These activities may be also common in PD patients and, therefore, must be included in, at least, the validation of an algorithm that is intended to be used in the daily life of patients. This way, the possible false positives that may arise in the home environment are able to be identified, as it will be shown.

4.1 Offline algorithm for dyskinesia detection

The offline algorithm for dyskinesia detection has been developed by means of the data gathered during the first stage of MoMoPa project. The algorithm is based on, similarly to Manson algorithm [69] and, to a lesser extent, to Keijsers work [60], in calculating the power spectra in the dyskinesia band upon X, Y and Z axis. The dyskinesia band is considered in the [1, 4] Hz range since, although other studies consider it until the 8 Hz, frequencies higher than 4 Hz correspond to tremor according to a consensus of the *Movement Disorders Society* [34], one of the most relevant societies in the PD research field. In this consensus, it is detailed that the minimum Parkinsonian tremor is 4 Hz and the maximum, instead, is not clearly identified since it is able to go from 7 to 9 Hz. This way, the upper limit of the dyskinesia band is set to the lowest frequencies of tremor.

Power spectra in [1, 4] Hz band is not only increased during dyskinesia since other activities, such as gait or go down or up stairs, also increase it [69]. This is the main reason why dyskinesia detection is considered to be extended to the whole

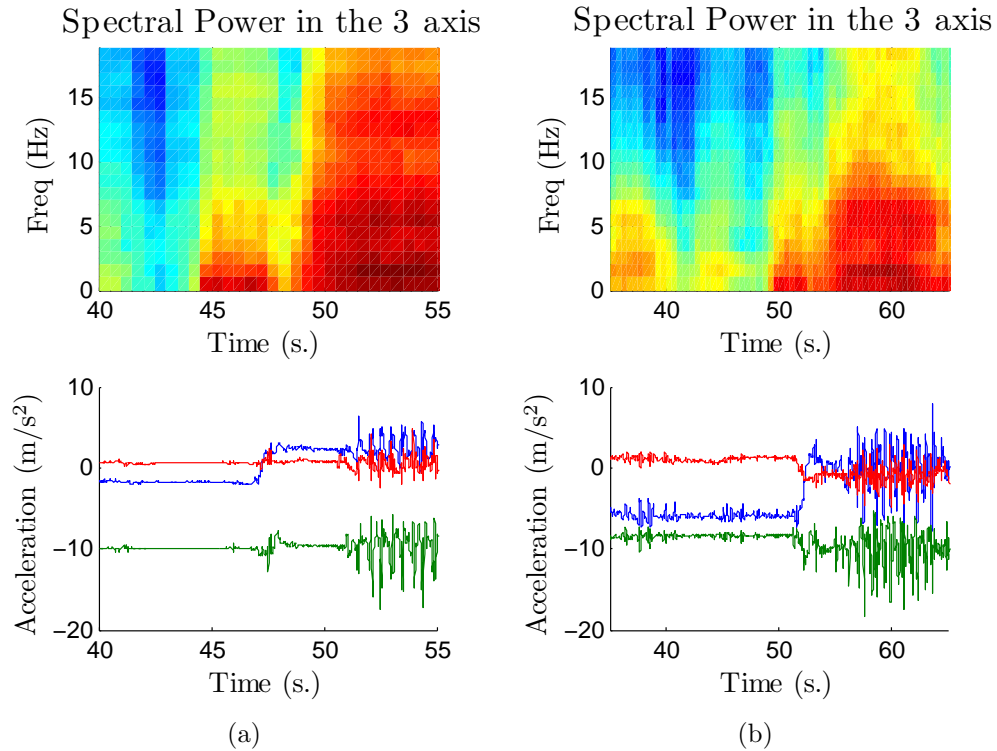


Figure 4.1: On the left, signal from a patient without dyskinesia. Patient sat during 40-47 s., stood (47-52 s.) and walked (55-55 s.). On the right, signal from another patient with dyskinesia. Low frequency and amplitude harmonics are observed for dyskinetic patient. Patient remained sat (until 50 s.), was stood (50-56 s.) and then walked. Higher amplitudes are shown with warm colours and lower amplitudes with cold colours.

spectra. Figure 4.1 shows the signal obtained from a patient with dyskinesia and the signal obtained from another patient without dyskinesia. It is observed that, when the patient without dyskinesia is sitting, there are not any harmonics in the signal while, when the patient suffers dyskinesia, harmonics among $[1, 4]$ Hz appear. Harmonics in this band are also present during gait in the signal gathered from the patient without dyskinesia. According to these observations, the strategy developed to detect dyskinesia is based on splitting the spectra into two different bands:

- Dykinetic band $[1, 4]$ Hz. A high power spectra in this band is an indicator that either the patient has dykinesia or the patient is walking or going up/down

stairs. In order to distinguish between both cases, the following band is analyzed.

- Non-dyskinetic band is considered to be the [8, 20] Hz. This band will enable us to know if an increase in the dyskinetic band is due to dyskinesia appearance or due to gait (or similar activities). Given that gait has the 99% of its power spectra below 20 Hz [14], 20 Hz are considered as the upper limit of the non-dyskinetic band. The lower limit of the band is provided by the maximum frequencies obtained in some dyskinesia research works, which have been reported to reach the 8 Hz [51]. This way, when power spectra in [8, 20] Hz is high, it is considered that the patient is walking or doing a similar activity. Since these activities produce harmonics in the whole spectra below 20 Hz [14], the power in the dyskinesia band will be also high. Consequently, harmonics with high amplitude in both dyskinetic and non-dyskinetic bands must not be considered as a positive dyskinesia detection.

Given the signal contained in a window time i , its spectral power in the dyskinesia band is denoted as P_i^d and its spectral power in the non-dyskinetic band as $P_i^{\bar{d}}$. Then, the dyskinesia algorithm's output is defined on them according to:

$$\text{Window Decision} \begin{cases} s_i = 1 & (\text{Dyskinesia}) & \text{if } P_i^d > d \wedge P_i^{\bar{d}} < \bar{d} \\ s_i = 0 & (\text{Non - dyskinesia}) & \text{otherwise} \end{cases} \quad (4.1)$$

where s_i represents the presence or absence of dyskinesia in window i , d and \bar{d} are the thresholds for dyskinetic and non-dyskinetic bands, respectively.

This way, activities such as walking or going up/down stairs and for which $P_i^{\bar{d}} \geq \bar{d}$ is hold, the output's algorithm will be negative despite of the patient has dyskinesia. That is, if a patient has dyskinesia the algorithm will not identify it during gait, even if it lasts minuted. This behaviour is actually associated with visual appreciation: when a patient has dyskinesia and walks, they cannot be distinguished in the waist since gait hides their movement.

The algorithm described in equation (4.1) enables us, ideally, to know whether in a given time window the patient has dyskinesia. However, dyskinesia are a symptom which may last several minutes. The most suitable window to examine the appearance of dyskinesia is, then, of minutes length. Nevertheless, the calculus of power spectra in current microcontrollers that are susceptible to be used during hours for real-time dyskinesia detection, that is, in those microcontrollers whose energy consumption is manageable through small batteries, requires us to use windows that do not exceed few seconds length, considering their limited amount of memory and a frequency sampling around 50 Hz. Therefore, using windows whose length is lower than the duration of the symptoms makes necessary to aggregate the algorithm's output during a certain time period. It is proposed to aggregate the output obtained in windows of w -seconds length overlapped at 50%, that is, starting a new window every $w/2$ seconds, along a period of T seconds. The algorithm provides, then, an output for each one of the $2T/w$ analyzed windows, which are aggregated by:

$$\text{Aggregated Decision} = \begin{cases} i. & \text{Dyskinesia} & \text{if } \sum_{i=0}^{2T/w} \frac{s_i}{n_d} > t_p \wedge \frac{w \cdot n_d}{2T} > t_c \\ ii. & \text{Non - dyskinesia} & \text{otherwise} \end{cases} \quad (4.2)$$

where $s_1, \dots, s_{2T/w}$ are the algorithm outputs represented in Equation (4.1), n_d is the number of time windows in which the condition for the non-dyskinetic band $P_i^{\bar{d}} \geq \bar{d}$ was not held, t_p is the threshold which sets the dyskinesia detection percentage needed to consider the T -seconds period as dyskinetic and t_c sets the minimum amount of analyzed windows in the period of T seconds.

From now on, the dyskinesia probability is set as the summation in Equation (4.2), which is always a number between 0 and 1. In order to consider a T -seconds period as dyskinetic, this probability must be greater than a certain threshold t_p . However, the probability value might not be enough by itself since this value could not be reliable if the patient walked during most of the analyzed period. This way, a confidence index is defined as $w \cdot n_d/2T$, which represents, from the total of $2T/w$ analyzed windows, the number of windows which were not rejected because of the condition $P_i^{\bar{d}} \geq \bar{d}$. In this manner, a low confidence index is considered to provide a not reliable dyskinesia

probability, since few windows could be analyzed. Consequently, the confidence index is required to be greater than a certain threshold t_c .

4.1.1 Parameter tuning for the dyskinesia detection offline algorithm

The parameters on which the algorithm depends are, according to equations (4.1) and (4.2), d , \bar{d} , w , T , t_p , and t_c . This section describes the methodology used to set their values.

The power spectra of a frequency band is computed, in this thesis, through the *Discrete Fourier Transform* (DFT) obtained by means of the *Fast Fourier Transform* (FFT). Given the signal of the accelerometer contained in a window of w seconds x_0, \dots, x_{N-1} , its DFT is composed of the complex values X_0, \dots, X_{N-1} , that are calculated according to the following expression:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k \frac{n}{N}} \quad (4.3)$$

where $k = 0, \dots, N - 1$.

The FFT algorithm is a method developed by Cooley and Turkey and presented in 1965 [31] that enables the DFT computation in a time proportional to $N \cdot \log(N)$. The FFT result is a set of complex values, where each one describes the amplitude and the phase of a frequency component. This way, each complex value represents a sinusoidal wave of a certain frequency, amplitude and phase. The frequency associated to value X_k is k/F_s Hz, where F_s is the signal sampling frequency. The amplitude of the frequency component X_k is $|X_k|$, where the absolute value of a complex number is the magnitude of its real and imaginary parts. Finally, the frequency component phase, or initial angle, is the angle between the real and the imaginary parts.

The spectral power of a frequency band $[f_1, f_2]$ from a signal x_1, \dots, x_{N-1} is obtained, in this thesis, as:

$$\sum_{i=k_1}^{k_2} |X_i|^2 \quad (4.4)$$

where k_1 corresponds to the index of the frequency component closer to f_1 and k_2 to the index closer to f_2 .

Finally, it is considered that the spectral power in a frequency band of the three signals provided by a triaxial accelerometer consists in the summation of the power spectra in the frequency band of each signal.

FFT influences the election of the window size w , since its value has been determined according to the capabilities of the DSPic33F microcontroller that the sensor device employs. This microcontroller has a limited amount of memory and its capabilities for the FFT computation limits the w value. As it is derived from the previous paragraphs, the FFT calculation depends on, first, the frequency sampling F_s since the number of harmonics in which the signal is decomposed is $F_s \cdot w$. Minimizing F_s enables saving memory and calculations and, given that in the computation of $P_i^{\bar{d}}$ frequencies up to 20 Hz are used, its minimum value is 40 Hz. Value w is set, then, considering that the number of samples in the window w is restricted by the microprocessor's Digital Signal Processing (DSP) to a power of 2 among: 32 (0.8 s.), 64 (1.6 s.), 128 (3.2 s.) and 256 (6.4 s.) samples. Given the microcontroller energy consumption, 128 samples are chosen and, consequently, w value is set to 3.2 s.

Thresholds d and \bar{d} are tuned according to a linear-kernel Support Vector Machine (SVM) [101].

- First, threshold \bar{d} , which enables the detection of gait and similar activities periods through the analysis of the [8, 20] Hz frequency band, is set. To this end, a training set was constructed composed of two classes whose patterns are one-dimensional real values. On the one hand, the first class consists of $P_i^{\bar{d}}$ values obtained from the signals corresponding to the first MoMoPa stage that were labelled as walking or going up/down stairs. On the other hand, the second class is composed of the $P_i^{\bar{d}}$ values computed in the rest of the signals. The dataset is, then, made of the elements from both classes, which are denoted as $\{P_1, \dots, P_N\}$, and were employed to train a linear kernel SVM:

$$\begin{aligned}
\min_{\mathbf{w}, b, \xi} \quad & \frac{1}{2} \|v\|_2^2 + C \sum_{i=1}^N \xi_i \\
s.t. \quad & \xi_i \geq 0 \\
& y_i [w \cdot P_i + b] \geq 1 - \xi_i
\end{aligned} \tag{4.5}$$

where y_1, \dots, y_N are the label associated to each P_i with $y_i = 1$ in case an activity similar to gait was performed and $y_i = -1$ otherwise; and v represents the hyperplane that separates the activities such as gait from the rest of activities.

Given that P_i is a scalar value, v is a threshold, v value obtained from Equation (4.5) is used as \bar{d} . Parameter C is determined as the value that maximizes the accuracy among the values $10^{-2}, \dots, 10^2$ in a 10-fold Cross-Validation [62].

- Second, d was set through a similar method applied to a different dataset. In this case, the dataset was composed of P_i^d values computed from the window signals gathered from the first 20 patients that satisfied $P_i^{\bar{d}} < \bar{d}$, that is, the patient neither walked nor went up/down stairs. The label consisted in the presence or absence of dyskinesia. The separation found through the minimization problem in Equation (4.5) was used to set d .

Values obtained for d and \bar{d} are 1.67 and 5.28, respectively. The SVM tool employed is LibSVM [28].

Threshold t_p establishes the proportion of windows that must be detected as dyskinetic in the period of $\frac{w \cdot n_d}{2T}$ seconds in order to consider the whole T -seconds period as dyskinetic. Threshold t_c , instead, establishes the minimum proportion of windows that must have been analyzed in the T -seconds period in order to enable considering it as dyskinetic. An empirical analysis of the signals from 5 patients obtained during the first MoMoPa stage was used to set their values, which are 0.7 for t_p and 0.3 for t_c .

The period of T seconds in which a dyskinetic decision is given depends on the data analyzed. In the laboratory data, a period of $T = 6$ seconds was established

Patient	Specificity	Sensitivity	Positive Predictive Value	Negative Predictive Value
1	1.00	-	-	1.00
2	1.00	0.77	1.00	0.79
3	1.00	0.81	1.00	0.67
4	1.00	0.18	1.00	0.69
5	1.00	0.00	-	0.98
6	1.00	-	-	1.00
7	1.00	0.00	-	0.67
8	1.00	-	-	1.00
9	1.00	-	-	1.00
10	1.00	-	-	1.00
11	1.00	-	-	1.00
12	1.00	0.13	1.00	0.45
13	1.00	-	-	1.00
14	1.00	-	-	1.00
15	1.00	0.92	1.00	0.73
Total	1.00	0.40	1.00	0.86

Table 4.1: Validity of the offline dyskinesia detection algorithm in the daily life of 15 PD patients

given that data gatherings were short; for instance, some tests lasted less than a minute. The signals collected from patient' homes were treated through periods of $T = 60$ seconds, since these tests had a duration of several hours.

4.1.2 Offline dyskinesia detection algorithm in signals obtained from patients' home

Table 4.1 shows the results of the offline algorithm for dyskinesia detection applied to the 15 signals obtained from 15 patients that participated in the second MoMoPa stage. The signals were obtained in the patients' home and its surroundings. The algorithm, which produces an output every 60 seconds, analyzed 1937 signal minutes (33 h.), from which 460 corresponded to dyskinetic periods and 1477 to non-dyskinetic periods. Results show an excellent specificity and a lower sensitivity. The high specificity shows that no false positive were found in the patients activities.

4.2 Online algorithm for dyskinesia detection in healthy people: false positive identification

The dyskinesia online algorithm was obtained from the adaptation of the offline algorithm to the resources that the microcontroller dsPIC33F offers. The resulting algorithm computes the probability and its confidence in periods of $T = 60$ seconds.

As it has been presented in the previous section, the offline algorithm employs the FFT to analyze 3.2-seconds windows. In the online algorithm, the DSP that the microcontroller includes has been used to perform the DFT since it enables its computation in a single clock cycle. The correct implementation of the real-time algorithm was evaluated by comparing the online values obtained from $P_i^{\bar{d}}$ and P_i^d against the offline values. To this end, several tests with a duration greater than 5 hours were performed, during which the original accelerometer signal together with the $P_i^{\bar{d}}$ and P_i^d values calculated in real-time were saved in the microSD card. The absolute error between both measurements was measured, which was observed to be around few thousandths.

Once the online algorithm implementation was validated, it was tested in the daily life of 8 healthy people previously to its usage in patients with PD. The goal of this test was evaluating its behaviour and identifying in advance the possible false positives that the algorithm could provide. To this end, 8 healthy people wore the sensor device with the real-time algorithm implemented along their daily life during several days. The algorithm analyzed 113 hours of signal distributed in 24 different tests, as Table 4.2 shows. The daily life activities that were performed by the users included, among others, the following ones: driving, using elevators, travel by car, travel by train, travel by underground, ride a bicycle, take a walk, walk a dog and use a laptop. Among them, only three activities provoked false positives in the dyskinesia detection algorithm, as Table 4.3 presents. Driving a car produces a false positive with a high probability, around 70%. Travelling by train and metro also provokes them, with a lower probability of 30%. These false positive are due to the waist movement that appears in the means of transport, which is similar to mild dyskinesia.

The real-time results show that the online dyskinesia detection algorithm is able

User initials (num. tests)	Recordings duration
A1 (5 tests)	52 h.
L1 (1 tests)	1 h.
C (2 tests)	9 h.
A2 (1 tests)	3.5 h.
G (3 tests)	15 h.
B (1 tests)	5 h.
D (8 tests)	13.5 h.
L2 (3 tests)	14 h.
Total (24 tests)	113 h.

Table 4.2: Duration of the dyskinesia online detection tests for its validation in healthy users

Activity	Number of FP	Number of times that the activity was performed	Frequency
Driving	12	17	70.6 %
Traveling by metro	1	4	25.0 %
Traveling by train	2	11	18.2 %

Table 4.3: Frequency of the activities that provoked false positives (FP) in the online dyskinesia detection

to provoke false positives in means of transport (car, train and metro). Thus, false positives might be obtained as well during the evaluation of the algorithm in PD patients during similar activities.

4.3 Online dyskinesia detection in PD patients

The online implementation of the dyskinesia algorithm was validated in a total of 7 PD patients. During the clinical study in which the validation took place, a doctor called the patient approximately every hour in order to ask for the patient's current motor state and the presence of dyskinesia. Patients could respond with 3 options in each case. Patients could say that their motor state was ON, OFF or something between which was called *intermediate state*. Intermediate state is considered a motor state that can appear during a smooth fluctuation from ON to OFF and from OFF to ON. Regarding dyskinesia, patients reported if they were not suffering the symptom

or they had mild dyskinesia or severe dyskinesia.

According to this information, the doctor could remotely change the dose administered by the apomorphine infusion pump that the patient wore by means of the HELP web portal. The motor state and the dyskinesia presence was annotated in a diary by the same doctor who called the patient. Concurrently, patients were asked to annotate every half an hour their own motor state and whether they were suffering dyskinesia or not. As it has been described in the introduction, patients are usually not able to distinguish their motor state and they are commonly not aware of dyskinesia presence. Thus, it is important to note that the gold-standard of the algorithms in this clinical study is not very reliable and the algorithm validation conditions are, consequently, difficult for achieving very high specificities and sensitivities such as those obtained in the controlled conditions. Moreover, this difficulty is even more increased considering that some diaries written by the clinicians do not match with those written by the patients.

4.3.1 Available data for the online dyskinesia detection validation

Among the 7 patients that participated in the clinical study, only data from 5 of them have been usable for the validation of the online dyskinesia detection. The gold-standard of the 7th patient turned out to be not valid given that its diary dates were wrongly gathered and did not match the dates of the movement sensor values available in the HELP server. The other patient whose data could not be analyzed corresponds to the 5th patient, from which the server did not register any sensor value since the patient did not have mobile coverage at home. Thus, the online dyskinesia detection and motor states validation was performed in 5 PD patients.

Data considered for the online dyskinesia detection are:

- Dyskinesia's probability and confidence computed by the sensor in real-time in periods of 1 minute according to Equation (4.2). Both values were sent by the sensor to the HELP server by means of the mobile phone that patients worn.

- **Dyskinesia label.** The gold-standard used in the dyskinesia detection has been the intersection of the information contained in the diaries gathered by clinicians and patients. More concretely, when both diaries coincide in the presence of dyskinesia in a certain time, the gold-standard took that value. Instead, when the diaries values did not match, the gold-standard was set as no available for those specific times.

Figure 4.2 shows one of the diaries gathered by a patient during the clinical study, in this case the information corresponding to the motor states is presented. These diaries had to be filled by the patient every half an hour. This way, it was possible to know the presence of dyskinesia in the time that the information was written down. The value obtained from the diary was considered to be valid since 15 minutes before the time in which it was written until 15 minutes after that moment.

4.3.2 Results and discussion

Table 4.4 shows the results for the online dyskinesia detection. The following considerations arise from the table:

- **Positive Predictive Value (PPV)** is low, which means that many False Positives appear. Dyskinesia algorithm must be improved since some activities have been shown to be confused with dyskinesia. Unfortunately, it has not been possible to maintain a complete list of activities that produce dyskinesia false positive. The partial list of identified activities are: cooking, sweeping and ironing.
- Many dyskinesias are not detected, i.e. results show a low sensitivity. This low sensitivity is presented together with a low PPV since there are more false negatives than true positives. This is due to the fact that very mild dyskinesia were annotated in the diary as dyskinesia, which are not detected by the sensor.
- A good **Negative Predictive Value (NPV)** is obtained, which means that the algorithm correctly detects a lack of dyskinesia. This is compatible with a low

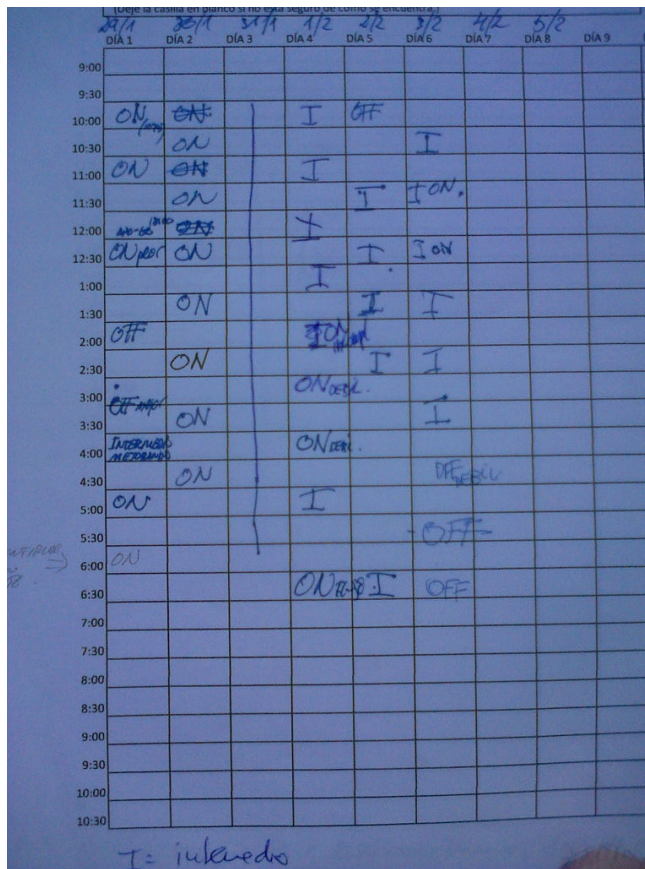


Figure 4.2: ON/OFF diary filled by a PD patient during HELP study. Data missing in the diary are shown.

sensitivity since, although some false negatives are obtained, much more true negatives are produced while, at the same time, more false negatives than true positives are provided.

Figure 4.3 shows a successful detection of the algorithm for the fifth patient. It is shown how the algorithm correctly detected the dyskinesic and non-dyskinesic periods. In this case, no activities were confused as dyskinesia and mild dyskinesia could be detected. It is shown how the dyskinesia probability fluctuates among time, which could be improved with a low-pass filter. However, a simple probability average would correctly reflect the presence or absence of dyskinesia in a quantitative way.

In summary, the dyskinesia algorithm has good specificity and negative predictive value, thus detects well the absence of dyskinesia. In some cases, the algorithm has successfully monitored dyskinesias. However, the sensitivity and positive predictive values must be improved, as there are many daily life activities which the algorithm takes by dyskinesia. In order to achieve this, the activities that may be confused as dyskinesia must be involved in the development of the algorithm. Moreover, dyskinesia severity could help on the correct detection of the symptom since a different value of l could be used to detect both of them.

Patient	Patterns	Dysk. patt.	Non-dysk. patt.	Sens.	Spec.	PPV	NPV
1 - Madrid-1	2009	321	1688	0.18	0.99	0.71	0.86
2 - Madrid-2	1585	90	1495	0	0.97	0	0.94
3 - Barcelona-1	558	167	391	0.11	0.96	0.54	0.72
4 - Barcelona-2	3161	578	2583	0.59	0.51	0.21	0.85
6 - Barcelona-4	1413	463	950	0.51	0.91	0.74	0.79

Table 4.4: Online detection of dyskinesia results from HELP study

4.4 Conclusions

A dyskinesia detection algorithm based on a single accelerometer placed on the patient's waist has been presented. It analyzes the signal spectrum and allows detecting the presence of dyskinesias except when the patient walks. The reason behind is that gait movements overlap the involuntary movements associated to dyskinesias.

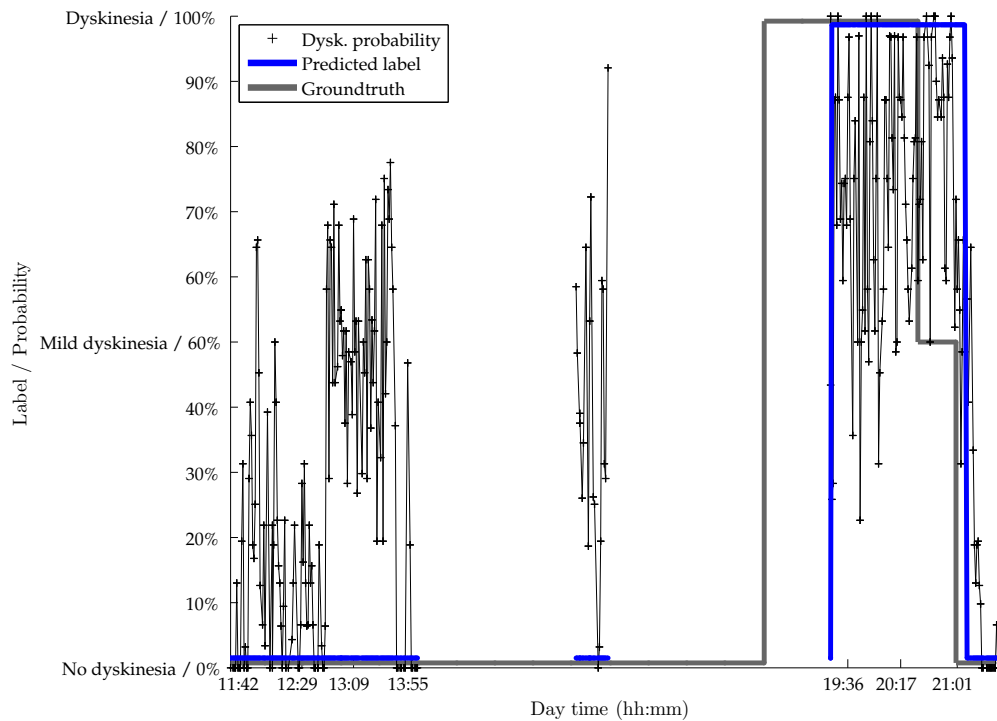


Figure 4.3: Dyskinesia online detection for patient 5th. Online dyskinesia probability (see Eq. 4.2) and the dyskinesia predicted label are shown. Dyskinesia probability was computed based on the analysis of 1 minute periods. The algorithm successfully monitored this patient during this day.

The validation of the dyskinesia algorithms in its offline versions was made by means of the signals recorded in 15 patients performing daily life activities in their own environment and for several hours. An online implementation of the algorithm in a low consuming microcontroller has been tested in 7 PD patients and in 8 healthy patients. Results show that severe dyskinesia is effectively detected and that some false positive arise, specially in means of transport.

The algorithm has been employed in the first pilot in which an apomorphine pump automatically controlled the medication of the patient. The results obtained show that, with a false positive rejection, it may be the tool for neurologists that would help them to effectively manage PD.

Chapter 5

ON/OFF motor states detection

In this chapter, contributions in ON/OFF motor states detection are presented. The development of the detection methods, as well as the results obtained in the clinical study previously presented, are detailed.

The main goal of the ON/OFF detection algorithm consists of obtaining a measuring instrument correlated to ON and OFF motor states. Thus, an objective measurement of the patient's motor state is expected: the greater the measurement, the more probable the patient is, for instance, in the ON state and, at the same time, the lower the measurement the more probable the patient is, according to the previous example, in the OFF state. Intermediate values, consequently, would correspond to an intermediate state.

The ON/OFF motor state detection algorithm has been developed by using the data from 20 patients of the MoMoPa's first phase, which corresponds to data obtained from the laboratory. The data from the first 10 patients have been used to train the developed algorithm and the other 10 patients to validate it. Secondly, the algorithm has been tested on the data belonging to the MoMoPa's 2nd phase patients, which was obtained in patient's daily life environment. Finally, a real-time version of the algorithm has been tested during two weeks in 5 PD patients.

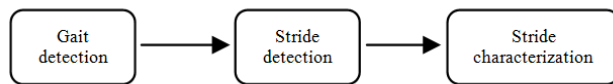


Figure 5.1: Three stages of the ON and OFF motor states characterization and classification

5.1 Offline algorithm for the ON/OFF motor states detection

ON/OFF motor states, as described in Chapter 1, are commonly revealed in three main motor symptoms: as the presence of bradykinesia or FoG during the OFF states or the existence of dyskinesia in the ON periods or in the wearing off periods. Bradykinesia and FoG are two symptoms that fundamentally alter patient's gait, in such a way that gait is slowed down or even frozen during, mainly, the OFF states [109] [119] [120]. Consequently, gait is one of the activities in which patient's motor state is more clearly distinguished. Therefore, gait is considered to be analyzed to detect the motor state of a patient.

This way, the ON and OFF motor states analysis is based on characterizing gait cycles, i.e. strides. A prerequisite is that patient should be walking; thus, a gait detector should enable gait cycles analysis. The latter analysis should perform, on the one hand, an identification of gait cycles from the accelerometer signal, i.e. stride detection, and, on the other hand, it should characterize gait cycles by some features that correlates motor states. Then, a three-step ON and OFF characterization method is obtained, which is represented in Figure 5.1:

1. Gait detection consists of a pattern recognition process applied to the accelerometer signal in order to identify whether the patient is walking or not. This way, a bi-class problem arises. SVM are employed to solve the bi-classification problem since they are able to find an optimal classifier without suffering from local minima, as gradient methods for neural networks do, while providing state-of-the-art performances and good generalization abilities [101]. SVM input consists of several features obtained from 3.2 s. windows of the accelerometer signal.

A training dataset was created by means of the data belonging to the first 10 patients that performed the laboratory experiments. The dataset is composed of the following features obtained from each axis of the accelerometer and their magnitude:

- maximum and minimum value, range, median, average and standard deviation ,
- spectrum power and maximum amplitude in the spectrum or spectra band between $[b_1, b_2]$ Hz s.t. $b_1 < b_2, b_1$ and $b_2 \in \{0, 0.5, \dots, 20\}$ Hz .

The two most significant features in gait detection, considered as those that maximize inter-class distance and minimize intra-class distance, were selected according to a Relief algorithm [89]. The obtained SVM after the learning process was validated against the data belonging to the last 10 patients that performed the laboratory experiments.

2. The stride detection process, which is performed when the SVM gives a positive output, takes advantage of how acceleration signals from the lower part of the trunk behave due to biomechanical characteristics of gait, as described in the literature. The onset of gait stance phase, when the heel makes contact with the ground, can be determined by a local minimum in the forward acceleration observed from the lower trunk [126]. This event of the gait cycle is also known as ‘Initial Foot Contact’ and is considered to establish the starting of a step. However, we are interested in strides, which are composed of two consecutive steps. Discrimination between right and left steps can be performed by analyzing relative extrema on lateral acceleration of the lower trunk, since it approximately describes a sinus period during a gait cycle [126]. Consequently, forward acceleration provides step identification and lateral acceleration allows the determination of strides.

Figure 5.2 shows the result of the described stride detection algorithm based on step detection, which was previously used in [93]. It also shows the stride detection validation against the gyroscope signal obtained from the shanks sensors.

This previously has been shown to determine swing phase [10], during which the foot is off the ground and which is preceded by stance phase as a minimum peak of the signal.

3. The stride characterization process aims to extract a feature representing the smoothness of movement of the patient, i.e. bradykinesia and rigidity, from previously detected strides. This way, several statistics are applied and evaluated. The best statistic is considered to be one that maximizes the separation between ON and OFF motor states. Furthermore, we are interested in a representation that linearly separates both motor states and that intuitively represents the smoothness of the movement. Thus, the best statistic should maximize the Area Under ROC Curve (AUC). Furthermore, it is considered that motor states are user-dependent; ideally, the border between ON and OFF depends on the stage of the disease and the patient. Thus, the threshold that best distinguishes both motor states in a certain patient is expected to have a different value than the best threshold for another patient. This way, AUC is evaluated separately in each patient

5.1.1 Algorithm results in laboratory tests on PD patients

The features selected by Relief algorithm as the most relevant for gait detection are the tri-axial power spectra between [0.1, 3] Hz and [0.1, 10] Hz. The training set, obtained from phase-one patients' data, was used to train a SVM by a 10-fold CV with a Radial Basis Function (RBF) kernel. The validation of SVM on the testing set obtained from phase-two patient's data provides the following results: specificity 0.84; sensitivity 0.90, and accuracy 0.94. These results validate the resulting SVM as a gait detector for PD patients.

The stride characterization process evaluates how several features representing the strides separate ON and OFF motor states by measuring the AUC. Features maximizing the distance between motor states are expected to represent the smoothness of movement. Tested features are the same as used for gait detection listed in the previous section. During the experiments, six PD patients had ON-OFF fluctuations,

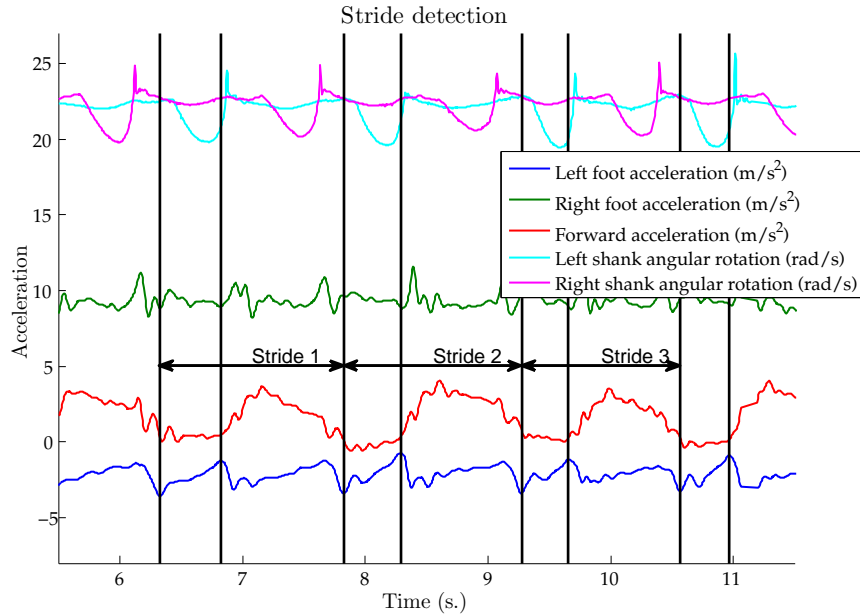


Figure 5.2: Three strides detected by the algorithm described. Strides are divided into two steps. Gyroscope signals from shanks that validate the detection are also shown.

and the results of the best 5 features are shown in Table 5.1. The best five features relates to tri-axial power spectra values, more specifically features F_1 to F_5 are power spectra in bands $[0, 10]$, $[0, 8]$, $[0, 7]$, $[0, 6]$ and $[0, 5]$ Hz. How the feature values averaged between 5 consecutive strides affect the results was also tested, in order to add robustness to the method. The obtained AUC values are shown in Table 5.2.

Results show that an excellent AUC value is provided by the tri-axial power spectra in the band of $[0, 10]$ Hz averaged between 5 strides and used as a linear classifier. This result can be explained according to PD symptoms. The reduction of the step frequency and the shortening of stride length and speed are common PD gait alterations. Figure 5.3 shows two typical strides obtained during ON and OFF states and it can be observed that OFF motor state provides less amplitude in both temporal and frequency domains.

It is important to note that data from ON and OFF motor states are quite pure in the sense that patients performed the experiments in OFF state by avoiding the first medication intake in the morning. After some recover time, which may took up

to 3 hours, the experiments in ON were performed. Hence, the excellent separation between motor states had been facilitated by the bi-modal distribution obtained.

Patient	AUC F_1	AUC F_2	AUC F_3	AUC F_4	AUC F_5
P_1	0.820	0.875	0.851	0.812	0.812
P_2	0.904	0.895	0.889	0.858	0.848
P_3	0.822	0.770	0.808	0.767	0.802
P_4	0.855	0.825	0.815	0.883	0.848
P_5	0.889	0.922	0.879	0.898	0.858
P_6	0.805	0.804	0.816	0.781	0.828
Avg.	0.849	0.848	0.843	0.833	0.833

Table 5.1: ON/OFF Motor States Detection Results in 10 last patients from laboratory tests (only 6 had motor fluctuations)

Patient	AUC F_1	AUC F_2	AUC F_3	AUC F_4	AUC F_5
P_1	0.979	0.974	0.981	0.981	0.980
P_2	0.934	0.905	0.894	0.861	0.856
P_3	0.827	0.811	0.816	0.818	0.816
P_4	0.997	0.977	0.960	0.940	0.920
P_5	0.936	0.936	0.921	0.905	0.890
P_6	0.997	0.997	0.997	0.997	0.997
Avg.	0.945	0.933	0.928	0.917	0.910

Table 5.2: ON/OFF Motor States Detection Results in 10 last patients from laboratory tests averaging 5 strides (only 6 had motor fluctuations)

5.1.2 Algorithm results in daily life environment on PD patients

This section describes the validity of the offline motor states detection applied to the 15 patients who performed the tests in their own house and the surroundings during several hours. Thus, a continuous change between motor states was obtained in some patients, which is very different to the laboratory test data where pure motor states were obtained.

To assess the validity of the ON/OFF detection algorithm, its results, which consist of a continuous numerical variable, were classified into "ON" or "OFF" categories

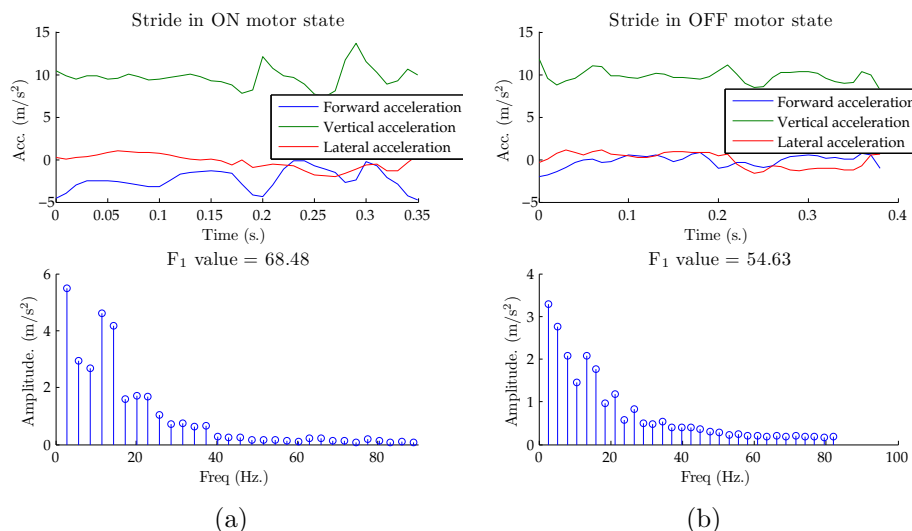


Figure 5.3: Frequency content of a stride during an ON state and a stride during an OFF state. Lower amplitudes are observed in the OFF state stride.

after establishing a splitting ON/OFF threshold. Measurements describing patient's strides were split into two datasets. One of these datasets was used to fix the ON/OFF splitting threshold; thus, strides over the threshold were considered as walking in ON-state while those below the threshold were considered OFF-state. The other dataset was used to assess validity of the algorithm and threshold. The first dataset included 20% of the measurements consecutively recorded both in ON and OFF. The remaining data were included into the second dataset. All specificity, sensitivity, PPV and NPV values reported in this thesis correspond to the analysis of the second dataset. To minimize the effects of an arbitrary selection of the elements used to fix a threshold, the splitting process was repeated 30 times; mean and standard deviation are provided.

The ON/OFF splitting threshold was established by using a linear kernel Support Vector Machine (SVM), similar to the process followed for dyskinesia threshold adjustment. Thus, since data to be classified were scalar, the splitting hyperplane found during the training process is a real-valued threshold. Finally, when enough data were available, 10-fold Cross Validation (CV) was used; otherwise, 2-fold (minimum) CV was used.

Thus, the sensitivity, specificity, positive and negative predictive values corresponding to the ON/OFF detection algorithm were studied by applying it to the detected walking segments containing more than 10 strides or to any walking segment of a patient. The first and last two strides of each walk segment were rejected since gait initiation and termination were considered not to be stable. Signal segments corresponding to a state that was described as "indefinite" or "intermediate" between ON and OFF by the patient were disregarded. Signal segments lacking comparison standard because of technical errors or artifacts, as well as the 5-minute segments immediately prior and following any change of state (ON/OFF; ON/intermediate; intermediate/OFF; OFF/ON) were disregarded with the aim of reducing potential errors derived from time gaps between actual changes and the moment an observer recorded such changes. The validity of the algorithm was studied for individual patients and the results were averaged.

Tables 5.3 and 5.4 show sensitivity and specificity values corresponding to the ON/OFF detection algorithm. A total of 46.9 hours of inertial sensor signals were recorded corresponding to the movement records of 15 subjects who participated in the validation phase (average 3.1 hours per patient; range 1.4-5.5 hours). The ON/OFF detection algorithm applied to these data yielded 2812 results: 1185 of them corresponded to ON state, 341 corresponded to OFF state and 156 to an "undetermined" state between ON and OFF. The remaining 1130 results were not included into the analysis because they fell within the range of potential synchronization inaccuracy between the sensor and the gold standard (fixed by the researchers in 5 minutes immediately before and after any state change). Out of the 1185 sensor results corresponding to the ON state, 850 derived from walking segments equal to or longer than 10 strides and the rest of them from shorter walking segments. Out of the 341 sensor results corresponding to the OFF state, 253 derived from walking segments equal to or longer than 10 strides.

For four patients participating in stage 2, not enough motor state (ON/OFF) data could be recorded as to apply the validation method. For the remaining 11 patients, average ON/OFF detector validity values were: sensitivity 0.91 (median 1, interquartile range (IQR): 0.85 - 1), specificity 0.90 (median 0.92, IQR: 0.81 -

0.92), positive predictive value 0.80 (median 0.80, IQR: 0.70 - 0.95) and negative predictive value 0.94 (median 1, IQR: 0.89 - 1). When only walking segments equal to or longer than 10 strides were considered (10 patients with a complete dataset), average ON/OFF detector validity values were: sensitivity 0.96 (median 1, IQR: 0.93 - 1), specificity 0.94 (median 0.96, IQR: 0.90 - 1), positive predictive value 0.90 (median 0.92, IQR: 0.80 - 1) and negative predictive value 0.98 (median 1, IQR: 0.97 - 1).

Figure 5.4 shows as an example the results produced by the ON/OFF and the dyskinesia detection algorithms, together with the corresponding motor state (ON/OFF) gold standard for patient number 3:

- *Dyskinesia algorithm*: A black line represents the output of the dyskinesia algorithm. Dyskinesia is discontinuously detected during the ON state and ends with the motor state change.
- *ON-OFF detection algorithm*: Blue crosses represent the values obtained by the ON/OFF algorithm. The higher the value is, the more ON is the patient. The lower the value is, a deeper OFF period the patient has.
- *Gold standard*: Grey and blue lines depict the ON/OFF and dyskinesia groundtruth, respectively.
- *Interpretation*: The ON-OFF detection algorithm showed higher values during the ON phase, lower values during the OFF phase and intermediate values during the transition phase. Dyskinesia detections were more frequent during the ON phase, and were not present during the OFF state.

5.2 Online detection of ON/OFF motor states

The online algorithm for the ON/OFF motor state detection has been tested in the HELP clinical study under the same difficulties, described in the previous chapter, hold for the online dyskinesia detection.

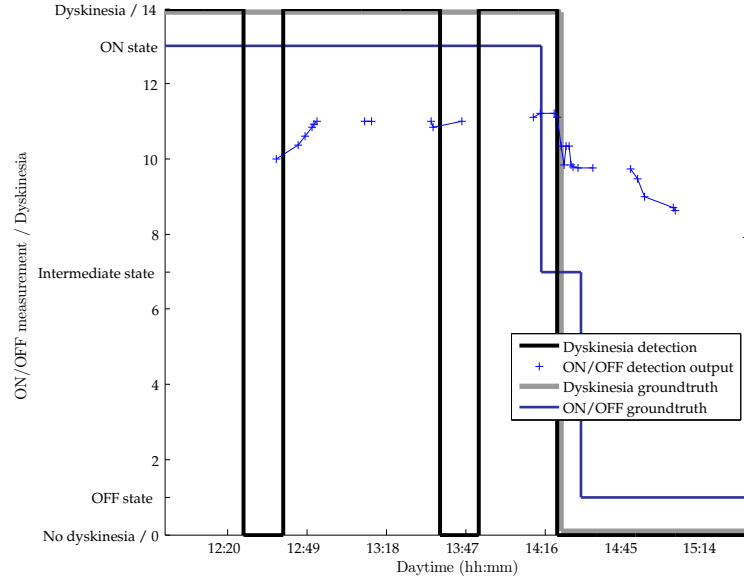


Figure 5.4: Outcomes of the dyskinesia and ON-OFF detection algorithms. Dyskinesia ends with the motor state change.

Patient	Specificity	Sensitivity	PPV	NPV
1	0,97 ± 0,04	1,00 ± 0,00	0,99 ± 0,01	1,00 ± 0,00
2	0,92 ± 0,12	1,00 ± 0,00	0,81 ± 0,27	1,00 ± 0,00
3	-	-	-	-
4	0,83 ± 0,13	0,89 ± 0,23	0,80 ± 0,15	0,95 ± 0,09
5	0,99 ± 0,01	1,00 ± 0,00	0,95 ± 0,03	1,00 ± 0,00
6	0,94 ± 0,04	1,00 ± 0,00	0,70 ± 0,20	1,00 ± 0,00
7	0,96 ± 0,03	1,00 ± 0,00	0,94 ± 0,05	1,00 ± 0,00
8	0,90 ± 0,04	1,00 ± 0,00	0,59 ± 0,17	1,00 ± 0,00
9	-	-	-	-
10	0,80 ± 0,27	0,56 ± 0,23	0,76 ± 0,31	0,70 ± 0,14
11	-	-	-	-
12	0,81 ± 0,16	0,85 ± 0,20	0,48 ± 0,30	0,99 ± 0,02
13	1,00 ± 0,00	0,90 ± 0,05	1,00 ± 0,00	0,77 ± 0,11
14	-	-	-	-
15	0,77 ± 0,11	0,80 ± 0,32	0,71 ± 0,10	0,89 ± 0,15

Table 5.3: Validity of the OFF detection algorithm for each patient (algorithm applied to walking segments of any length)

Patient	Specificity	Sensitivity	PPV	NPV
1	$0,95 \pm 0,05$	$1,00 \pm 0,00$	$0,99 \pm 0,01$	$1,00 \pm 0,00$
2	$0,94 \pm 0,03$	$1,00 \pm 0,00$	$0,68 \pm 0,15$	$1,00 \pm 0,00$
3	-	-	-	-
4	$0,88 \pm 0,09$	$0,96 \pm 0,07$	$0,78 \pm 0,15$	$0,99 \pm 0,03$
5	$1,00 \pm 0,00$	$1,00 \pm 0,00$	$1,00 \pm 0,00$	$1,00 \pm 0,00$
6	$1,00 \pm 0,00$	$1,00 \pm 0,00$	$1,00 \pm 0,00$	$1,00 \pm 0,00$
7	$0,96 \pm 0,03$	$1,00 \pm 0,00$	$0,94 \pm 0,04$	$1,00 \pm 0,00$
8	$0,97 \pm 0,04$	$1,00 \pm 0,00$	$0,81 \pm 0,21$	$1,00 \pm 0,00$
9	-	-	-	-
10	$0,83 \pm 0,26$	$0,82 \pm 0,11$	$0,87 \pm 0,17$	$0,86 \pm 0,08$
11	-	-	-	-
12	-	-	-	-
13	$1,00 \pm 0,00$	$1,00 \pm 0,00$	$1,00 \pm 0,00$	$1,00 \pm 0,00$
14	-	-	-	-
15	$0,90 \pm 0,08$	$0,83 \pm 0,24$	$0,89 \pm 0,07$	$0,91 \pm 0,13$

Table 5.4: Validity of the OFF detection algorithm for each patient (algorithm applied to walking segments longer than 9 strides)

The ON/OFF motor states detection algorithm requires a threshold against which the measurements obtained from strides are compared. The comparison result establishes whether a patient is in the ON or OFF state. This threshold has been set through the data from the first stage, the first week of the clinical study, instead of the two hospitalization days. This way, data from the second week has been used to obtain the performance measurements of the algorithm. The measurements obtained by the algorithm during the hospitalization, in the controlled environment, were observed to show different distribution than those obtained in the patient's home. OFF data distribution obtained in controlled environments showed a tendency to provide higher values than the OFF data obtained at patient' homes. Therefore, fixing the threshold by using the hospitalization data decreases the performance measurements. For instance, the threshold found for patient 3 during the hospitalization days was ≈ 12 , which would provide an almost zero sensitivity in the OFF motor state detection. This fact reasonably suggests that a motor state detection suitable for daily life

conditions must be trained in the same conditions and not in a controlled environment.

5.2.1 Adapting the offline motor states detection to the real-time processing

The motor state detection algorithm produces a value for each set of 5 strides detected. In the HELP study, this algorithm had to be adapted since the sensor could only send a measurement to the HELP server every minute according to the system specifications. Thus, instead of obtaining a feature for each set of 5 strides, the online algorithm produced different values that aggregated the information from all the strides detected in a minute. More concretely, the measurements that were sent by the sensor and stored in the HELP server were:

$$\mathcal{X} = \{x_1, \dots, x_{r+s}\} \quad (5.1)$$

$$\mathcal{N} = \{n_1, \dots, n_{r+s}\} \quad (5.2)$$

$$\mathcal{Y} = \{y_1, \dots, y_{r+s}\} \quad (5.3)$$

where

- x_i is a real value corresponding to the average value of the ON/OFF characteristic that aggregates the strides performed in minute i . Additionally, it is considered that there were obtained r values in ON state and s values during the OFF state.
- n_i is the number of steps at minute i .
- y_i is a label that follows the Support Vector Machine notation, i.e. $y_i = 1$ if the patient annotated that the motor state during that period was OFF and $y_i = -1$ otherwise.

\mathcal{X} set is filtered in the following way, so that a new vector $\mathcal{X}' = \{x'_1, \dots, x'_{r+s}\}$ is obtained by using a weighted mean according to the number of steps:

$$x'_i = \sum_{j=i-15}^{i+15} \frac{x_j \cdot n_j}{\sum_{k=i-15}^{i+15} n_k} \quad (5.4)$$

Thus, the values considered to weight a measurement are those that were obtained 15 min. after and 15 min. before the considered measurement.

Note that Equation (5.4) implies that all data provided by the online algorithms are used, even short walks with a single stride. The online algorithm rejects the first two strides and the last two as well, so that a minute with a single stride implies that the patient performed actually 5 of them. However, these short walks may not contain useful information on the motor state of the patient since, during them, the maximum speed is not reached and, consequently, the information contained in the strides may be confused with bradykinesia. To avoid this, a subset of \mathcal{X} , \mathcal{S} and \mathcal{Y} are used:

$$\mathcal{X}_k = \{x_i\} \quad \forall i \in 1, \dots, r+s \quad s.t. \quad n_i \geq k \quad (5.5)$$

$$\mathcal{N}_k = \{n_i\} \quad \forall i \in 1, \dots, r+s \quad s.t. \quad n_i \geq k \quad (5.6)$$

$$\mathcal{Y}_k = \{y_i\} \quad \forall i \in 1, \dots, r+s \quad s.t. \quad n_i \geq k \quad (5.7)$$

In this way, the data contained in these subsets are those corresponding to the minutes in which the patient walked at least $k \in \mathbb{N}$ strides. Note that these sets are a generalization of the previous ones, since $\mathcal{X}_1 = \mathcal{X}$. The filtering showed in Equation (5.4) is applicable to \mathcal{X}_k , which provides \mathcal{X}'_k .

Given the resulting \mathcal{X}'_k set, a Support Vector Machine with linear kernel is trained by using the data from the first week, in the same way that the dyskinesia thresholds were obtained. Thus, the following primal problem with linear kernel is solved (\mathcal{X}' notation is used for simplification):

$$\min_{w,b,\xi} \frac{1}{2} \|w\|_2^2 + C_{-1} \sum_{i=1}^r \xi_i + C_1 \sum_{i=r+1}^{r+s} \xi_i \quad (5.8)$$

$$\begin{aligned}
s.t. \quad & \xi_i \geq 0 \\
& y_i[w \cdot x'_i + b] \geq 1 - \xi_i
\end{aligned}$$

where w represents the hyperplane that separates the ON/OFF classes, $C_{-1} = C \cdot \frac{s}{r+s}$ and $C_1 = C \cdot \frac{r}{r+s}$ in order to weight the classes.

Since x'_i are scalar, w becomes a threshold. Thus, w is set as the value that distinguishes among motor states. In order to fix parameter C , a standard stratified 10-fold Cross Validation is used. C possible values are $10^{-2}, \dots, 10^2$.

Once the ON/OFF threshold w is found, it is applied to the data gathered in the second week of the study, i.e. the home monitoring that took place after the two hospitalization days. Since patients describe their motor state in three different possibilities (ON / intermediate state / OFF) and only ON and OFF motor states are clinically relevant, only these two states are needed to be identified. Measurements obtained during an intermediate state are measurements between those obtained in the ON and OFF motor state, as Figure 5.4 shows. The intermediate measurements zones were estimated to correspond to $\pm 10\%$ in the patients from the third experimentation, as Figure 5.4 describes: ON values correspond to those greater than 10 and OFF measures to those lower than 8. Thus, the ON/OFF motor state detection was designed to provide the decision by means of the following equation:

$$Motor \ state = \begin{cases} i. \ ON & \text{if } x'_i > 1.1 \cdot w \\ ii. \ OFF & \text{if } x'_i < 0.9 \cdot w \end{cases} \quad (5.9)$$

5.2.2 Threshold Adjustment Results

Figure 5.5 graphically depict the ON/OFF motor state characterization of each patient during the first week. Each cross corresponds to the characterization of one stride according to the procedure previously described in this chapter, i.e. each cross represents the walking behavior of a patient during a minute of the first week. Horizontal bars are the threshold found according to the linear-kernel SVM for each patient. Every patient has a different threshold in the same sense that every patient has a different ON/OFF motor state. The threshold values showed as horizontal bars

in the figures are shown numerically in the Table 5.2.2.

The following observations can be made from Figure 5.5 on the data distribution:

- ON and OFF data from patient 3 are clearly separable by a simple threshold.
- ON and OFF data from patient 6 are not linearly separable distributions although they can be distinguished as two different density functions.
- Motor states data from patients 1 and 4 are not only separable but they are completely overlapped, since seem to have similar values for both ON and OFF states. Patients 1 and 2 barely had bradykinesia fluctuations, which agrees with the data.

Thus, ON/OFF algorithm shows some differences for the motor states of patients 3 and 6, according to the first week data analysis. On the other hand, data from the other trials cannot be separated into ON and OFF states. As it will be shown, the motor state of the latter three patients will not be detected during the second week. Therefore, one-week analysis is enough to conclude if the sensor and its algorithm are able to detect the motor state of a certain patient.

It is important to note that, for instance, third patient's threshold value found according the sensor signals obtained in the hospital is 12.87. Thus, the OFF measurements obtained in the hospitalization days were higher than those obtained at home. Consequently, hospitalization's OFF was less severe than the OFF obtained at home. In this way, it is concluded that the adaptation of the ON/OFF state detection algorithm must be performed through data obtained at home and not in a controlled environment since measurements could be altered.

5.2.3 Online ON/OFF motor states detection results

This section describes the results of the online motor state detection during the second week of the study. The distributions of the data collected during the second week for each patient are shown in Figure 5.6. The values obtained follow a similar distribution to those showed in Figure 5.5.

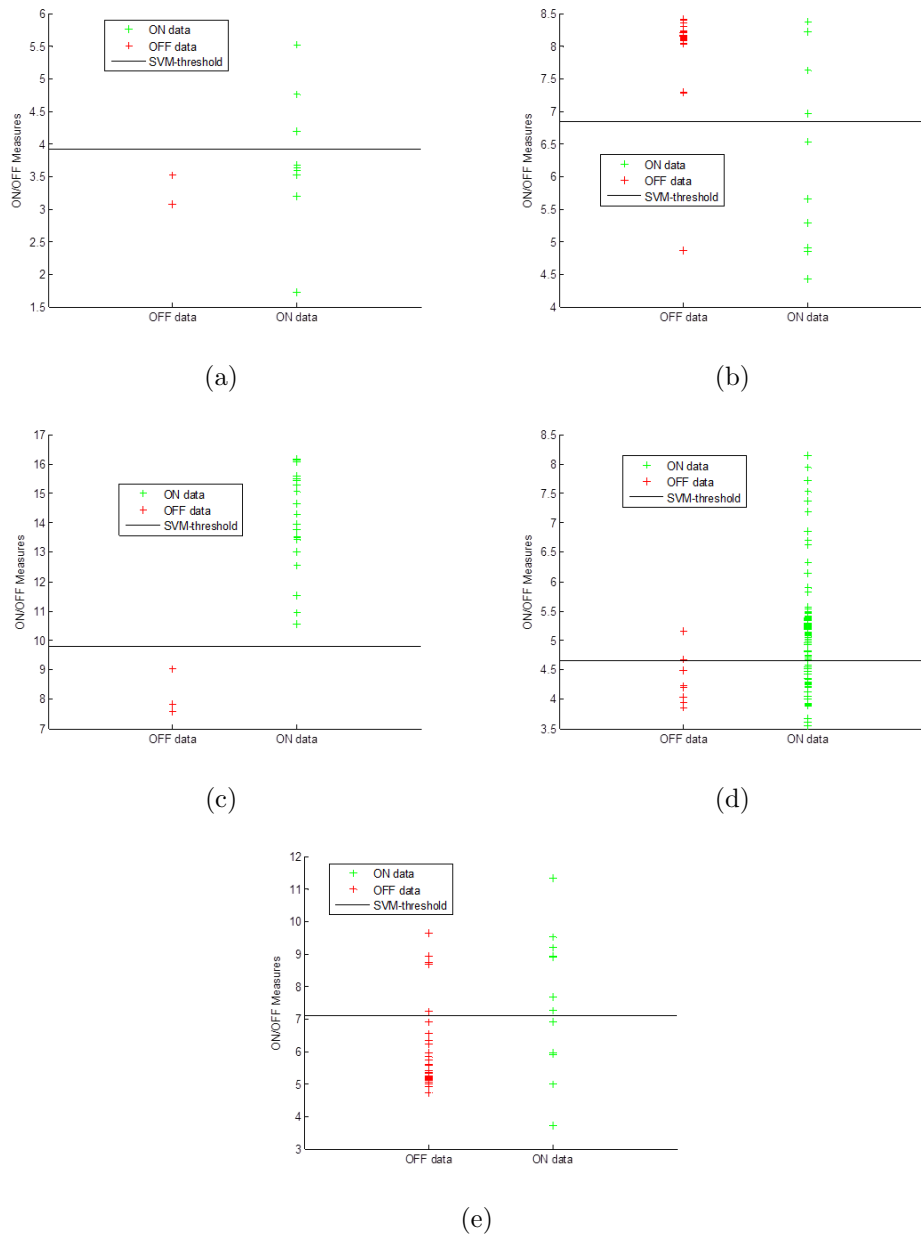


Figure 5.5: Graphical representation of the strides collected during the first week for (a) patient 1 (Madrid-1), (b) patient 2 (Madrid-2), (c) patient 3 (Barcelona-1), (d) patient 4 (Barcelona-2) and (e) patient 6 (Barcelona-4). Data showed correspond to \mathcal{X}_1 , i.e. minutes in which the sensor detected at least 1 stride are depicted.

Patient	SVM Threshold
1 - Madrid-1 *	3.91
2 - Madrid-2	6.83
3 - Barcelona-1	9.78
4 - Barcelona-2	4.64
6 - Barcelona-4 **	7.10

Table 5.5: Threshold values for all patients.

* Patient 1 - Madrid-1 threshold was fixed by using only the last two days of the first week since the previous days the sensor was worn together with the pump in the same belt, which resulted in an incorrect location of the sensor. It was solved by putting the pump in the neck through a collar.

** Patient 6 - Barcelona-4 threshold was fixed by using only the first three days of the first week since only 5 days of data were available.

The ON/OFF detection depend on value k in Equation (5.7), the minimum number of strides per minute. Table 5.2.3 shows the results for each patient according to k values from 1 up to 10.

Regarding patient 1 and patient 2, the corresponding data distributions obtained during second week are very similar to that obtained in the first week, according to Figure 5.6. Thus, there is a clear overlapping between the ON data and the OFF data in both distributions, which does not allow the algorithm to distinguish the patient state. In this sense, the overlapped measurements make the algorithm provide poor accuracies in terms of sensitivity and specificity. Main causes of the overlapped distributions and, consequently, the poor accuracies are:

- Patient had bradykinesia in both OFF and ON states. The patient did not clearly fluctuate in terms of movement slowness during gait, which is the main feature that the ON/OFF algorithm evaluates.
- Patient had much more freezing of gait in OFF than in ON state. Then, only short periods of gait with few strides are obtained, which does not allow a good accuracy of the sensor. In the Table 5.2.3, it is shown than the number of OFF patterns is very low compared to the number of ON patterns.

Patient 3 (Barcelona-1) has clear fluctuations between ON/OFF motor states in

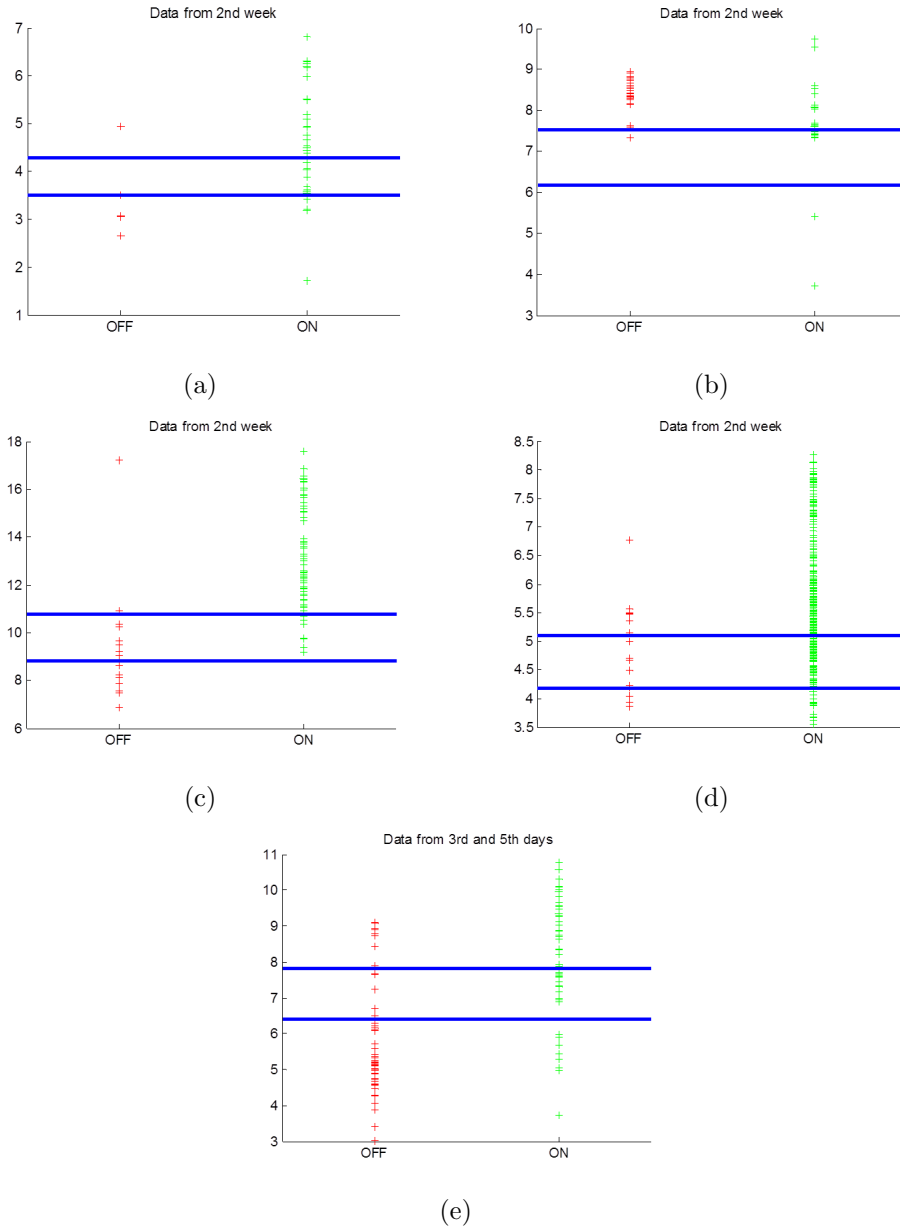


Figure 5.6: Graphical representation of the strides collected during the second week and the two thresholds used to determine the motor state for (a) patient 1 (Madrid-1), (b) patient 2 (Madrid-2), (c) patient 3 (Barcelona-1), (d) patient 4 (Barcelona-2) and (e) patient 6 (Barcelona-4). Data showed correspond to \mathcal{X}_1 , i.e. minutes in which the sensor detected at least 1 stride are depicted.

	k	Patterns	ON patterns	OFF patterns	Sensitivity	Specificity	PPV	NPV
Patient 1 (Madrid-1)	1	77	61	16	0,44	0,08	0,11	0,36
	2	74	59	15	0,47	0,14	0,12	0,5
	3	66	53	13	0,38	0,09	0,09	0,38
	4	64	53	11	0,45	0,09	0,09	0,45
	5	60	50	10	0,5	0,1	0,1	0,5
	6	56	46	10	0,5	0	0,1	0
	7	49	44	5	1	0	0,1	NaN
	8	49	44	5	1	0	0,1	NaN
	9	49	44	5	1	0	0,1	NaN
	10	44	39	5	1	0	0,11	NaN
Patient 2 (Madrid-2)	1	159	94	65	0,4	0,65	0,44	0,61
	2	155	91	64	0,41	0,64	0,44	0,6
	3	157	93	64	0,41	0,61	0,42	0,6
	4	150	95	55	0,47	0,61	0,41	0,67
	5	150	95	55	0,47	0,6	0,41	0,66
	6	150	95	55	0,47	0,6	0,41	0,66
	7	130	75	55	0,47	0,52	0,42	0,57
	8	129	74	55	0,47	0,5	0,41	0,56
	9	122	71	51	0,51	0,48	0,41	0,58
	10	116	65	51	0,51	0,43	0,41	0,53
Patient 3 (Barcelona-1)	1	129	94	35	0,63	0,99	0,96	0,88
	2	126	91	35	0,63	1	1	0,88
	3	114	89	25	0,72	1	1	0,93
	4	120	89	31	0,61	0,99	0,95	0,88
	5	105	85	20	0,4	1	1	0,88
	6	101	77	24	0,5	1	1	0,87
Patient 4 (Barcelona-2)	1	459	444	15	1	0,01	0,03	1
	2	451	437	14	1	0,01	0,03	1
	3	447	431	16	1	0,01	0,04	1
	4	449	433	16	1	0,02	0,04	1
	5	434	418	16	1	0,03	0,04	1
	6	424	408	16	1	0,03	0,04	1
	7	424	408	16	1	0,04	0,04	1
	8	428	416	12	1	0,05	0,03	1
	9	430	414	16	0,75	0,06	0,03	0,86
	10	427	409	18	0,78	0,07	0,04	0,87
Patient 6 (Barcelona-4)	1	89	39	50	0,9	0,72	0,8	0,85
	2	76	26	50	0,9	0,58	0,8	0,75
	3	76	26	50	0,9	0,62	0,82	0,76
	4	76	28	48	0,9	0,64	0,81	0,78
	5	68	21	47	0,89	0,52	0,81	0,69
	6	64	18	46	0,91	0,44	0,81	0,67
	7	60	14	46	0,91	0,29	0,81	0,5
	8	56	14	42	0,9	0,29	0,79	0,5
	9	56	14	42	0,9	0,29	0,79	0,5
	10	48	10	38	0,89	0,4	0,85	0,5

Table 5.6: Real-time ON/OFF motor states detection results

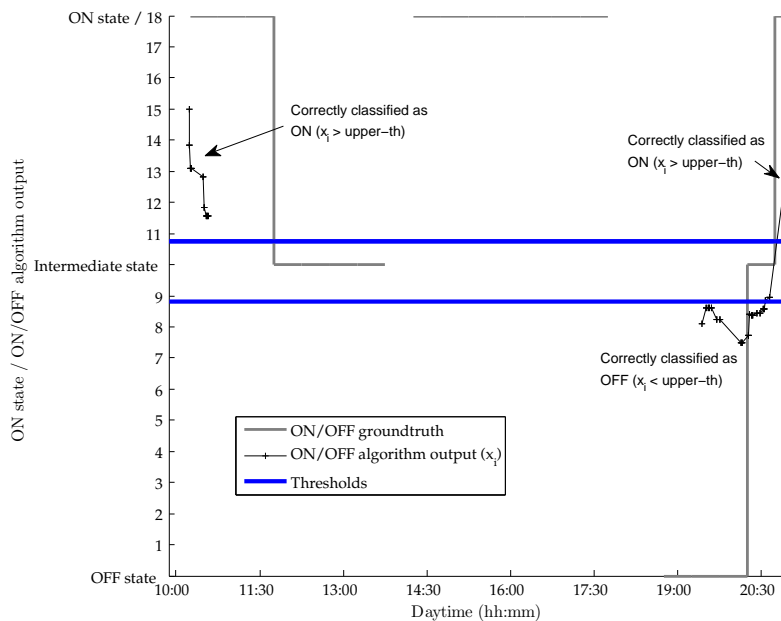


Figure 5.7: ON/OFF online detection during the second validation day of patient 3. Data showed correspond to $k=3$. Crosses are the measurements that correspond to the characterized strides that were obtained in real-time by the ON/OFF detection algorithm.

terms of bradykinesia. During the trial, the mobile phone had a very bad connection which made that few measurements were received by the server. Figure 5.7 shows the ON/OFF detection algorithm results for the second day of the second week. The relatively low sensitivity of the algorithm is mainly produced because of the data labeled as 'OFF' that the algorithm detects as ON during the 5th validation day, as showed in Figure 5.8. The algorithm performs well with this patient since, for instance, for $k = 3$ a sensitivity and specificity of 0.72 and 1, respectively, are obtained with very high (> 0.9) positive and negative predictive values. These results are exceptionally good given the conditions in which the validation process took place.

Patient 4 (Barcelona-2) had motor fluctuations in terms of bradykinesia. During the two hospitalization days, the sensor algorithm was able to capture the fluctuation by showing different measurements for each state. However, data obtained in the daily life environment does not correlate good with the ON/OFF diary since both first week and second week data distributions are, although distinguishable, overlapped. Figure

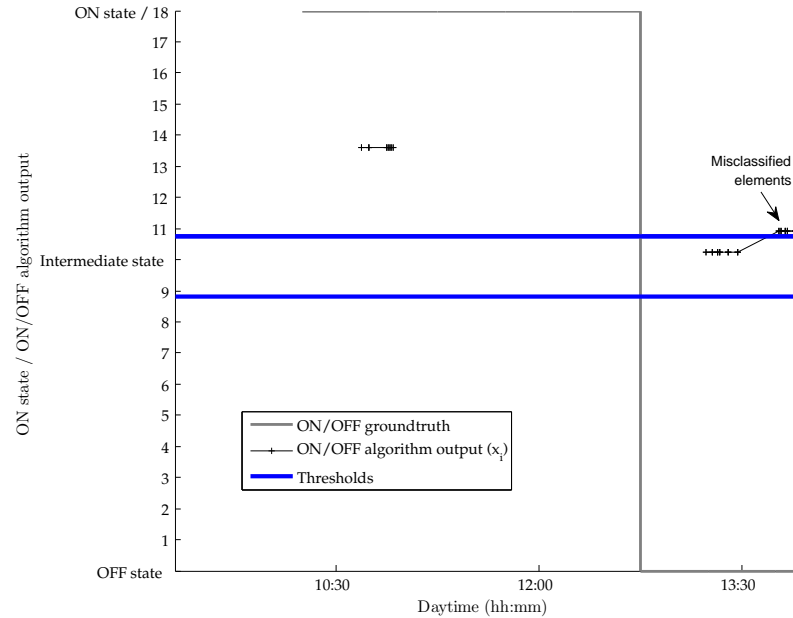


Figure 5.8: ON/OFF online detection during the fifth validation day of patient 3. Data showed also correspond to $k=3$. Misclassified elements do not allow a sensitivity greater than 0.72.

5.9 shows that motor fluctuations were very clear in the algorithm output values for the selected day. Note that the second OFF state was detected in advance by the sensor, which is an outstanding result. However, it decreases the sensitivity values. A similar result is obtained in Figure 5.11, where the last OFF period trend is observable before the OFF state was provided by the patient. Another interesting result is observable in Figure 5.10, where measurements obtained during the first OFF period show a clear difference in relation to the previous state. However, this OFF period is not correctly classified by the algorithm since its measurements are not located below the lower threshold. This result suggests that an additional classification could be performed in terms of relative changes: since motor states actually depend on patient's feelings and may appear as sudden motor changes, i.e. wearing OFF periods, sudden decreasements in the algorithm output could be classified as a motor state change into OFF period.

Regarding patient 6, only 5 days of data sent by the movement sensors are available. Thus, data from all days, except the 4th one during which few measurements

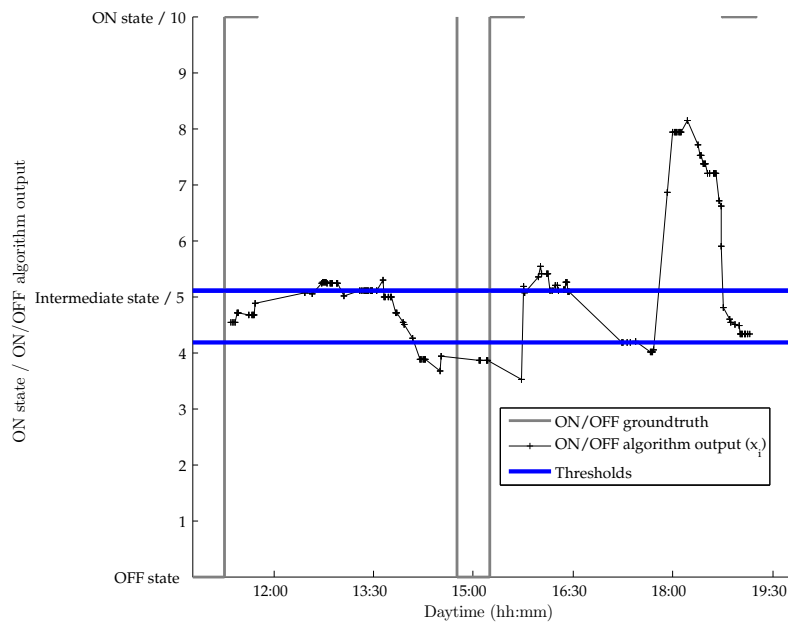


Figure 5.9: ON/OFF online detection during the fourth day of patient 3 first week. Data showed correspond to $k=5$. Motor fluctuations are very clear in the sensor measurements. However, the second OFF state is detected at 14:15 by the sensor, which is 45 min. before the patient described this state in the diary (15:00). The first and last intermediate states are perfectly determined by the algorithm.

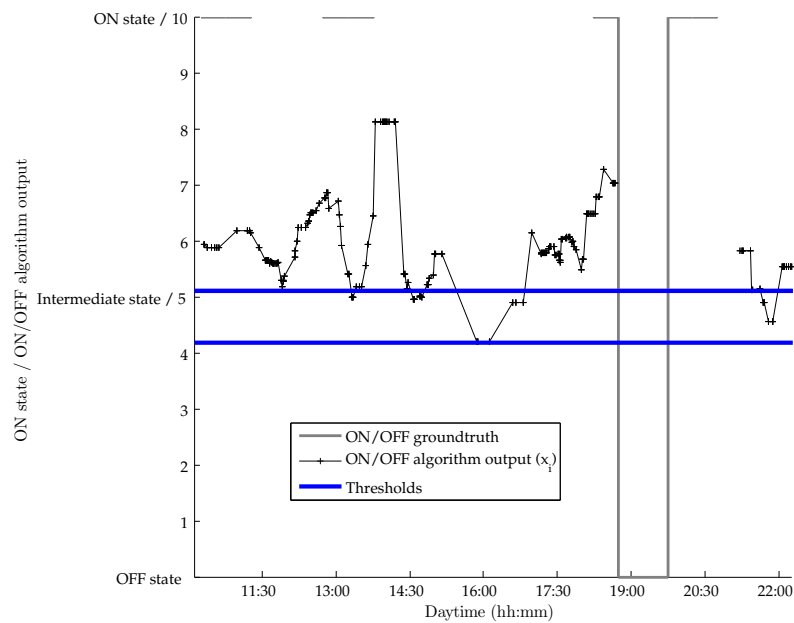


Figure 5.10: ON/OFF online detection during the second day of patient 3 second week. Data showed correspond to $k=5$. Fluctuations in the intermediate and the ON state are shown in the ON/OFF detection algorithm's output. There is not any data available during the OFF period in this day.

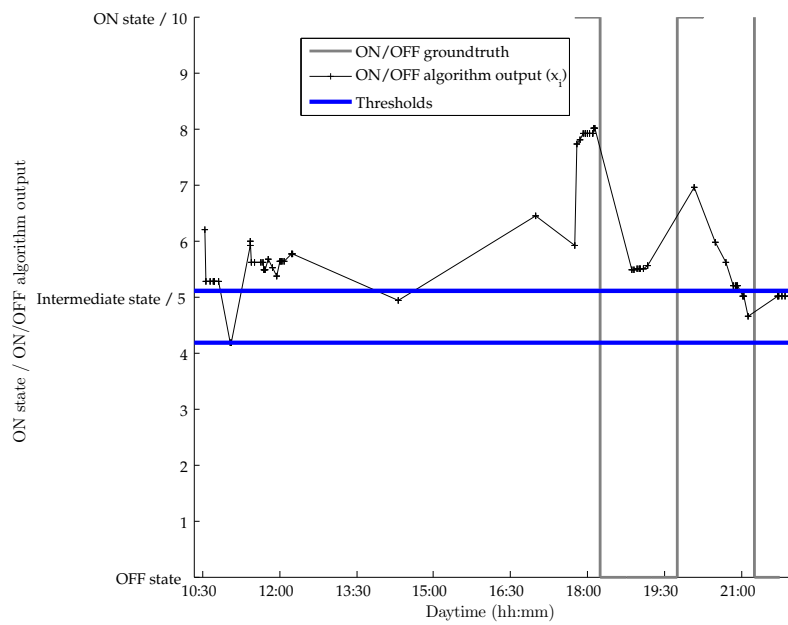


Figure 5.11: ON/OFF online detection during the last day of patient 3 validation week. Data showed correspond to $k=5$. Algorithm's output during the first OFF period shows a clear difference compared to the previous state, although measurements are not correctly classified. The last OFF period trend is clearly obtained in advance.

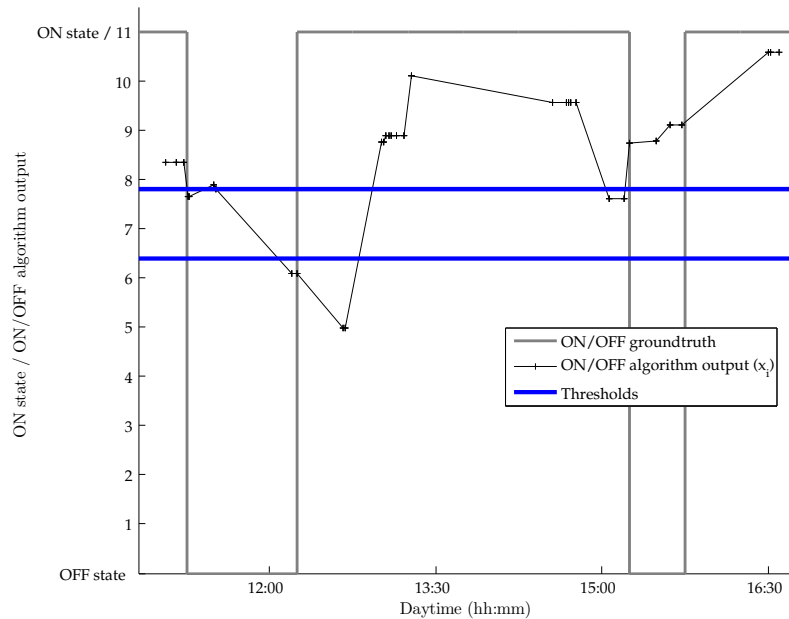


Figure 5.12: ON/OFF real-time detection for patient 6 during 1st training day. Data showed correspond to $k=1$. First fluctuation from OFF to ON is detected with a delay of 1 h. Second fluctuation is not detected although a decrease in the measurements is observed.

were received by the server, are showed in Figures 5.12, 5.13, 5.14 and 5.15. These figures show that the algorithm's output is a very good indicator of the patient's motor state although some delays can be found. More concretely, data from 19th of February in Figure 5.12 show a clear correlation between ground-truth and sensor measurements, although delays of approximately 30 min. in the OFF detection appear. Data from 20th and 21st of February in Figures 5.13 and 5.14 show the same correlation with no delays. Finally, data from 23rd of February in Figure 5.15 have a wrong detection at 16:30 and a delay in the last OFF phase. The delays make the specificity decrease until 0.71 although a positive and a negative predictive values over 0.8 are obtained. These results are very relevant given the harsh study conditions.

5.2.4 Discussion

The following issues arise from the previous sections:

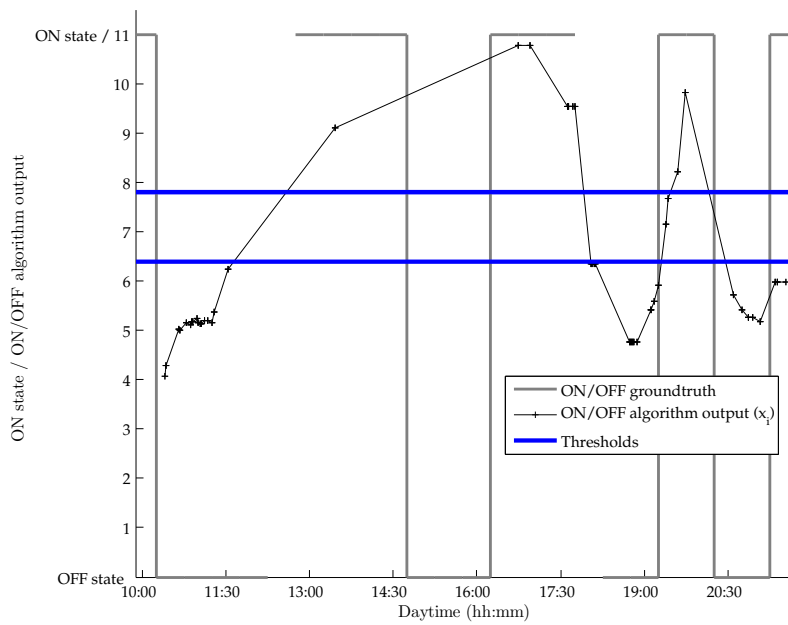


Figure 5.13: ON/OFF real-time detection for patient 6 during 2nd training day. Data showed correspond to $k=1$. First, third and fourth OFF periods are correctly detected, although there was not any data available for the second one. Second, third and fourth ON periods are correctly identified. First ON period lacks data and in the last one measurements were increased although an it was not correctly detected.

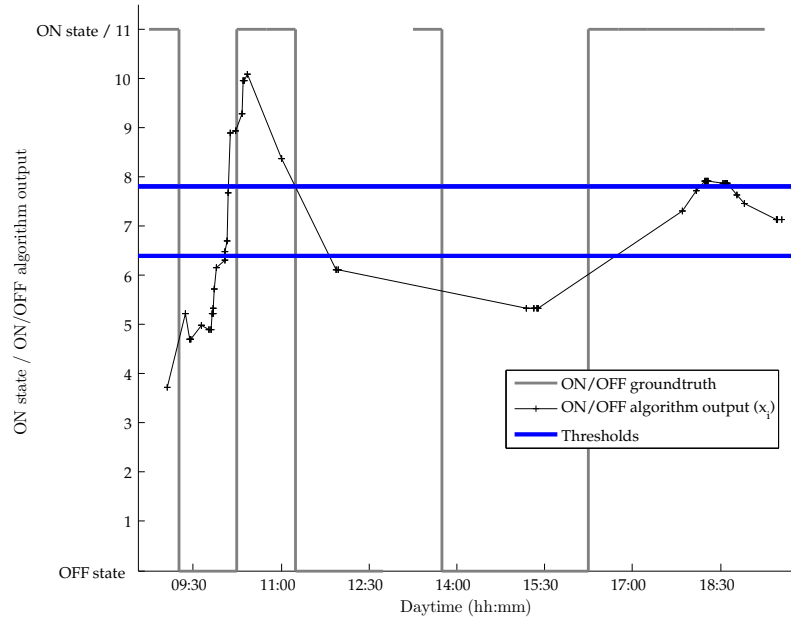


Figure 5.14: ON/OFF real-time detection for patient 6 during 1st validation day. Data showed correspond to $k=1$. All OFF periods are detected and, with the exception of the third period, all of them were correctly identified.

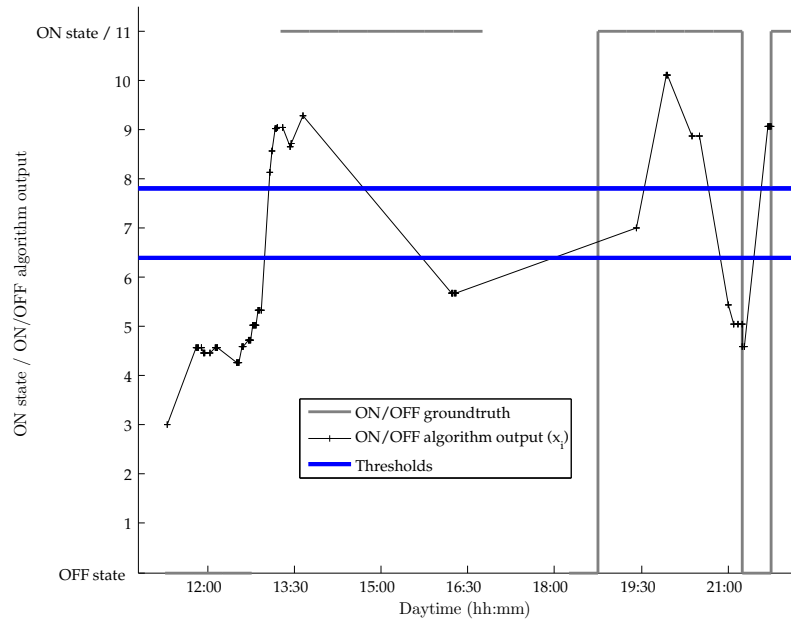


Figure 5.15: ON/OFF real-time detection for patient 6 during 3rd validation day. Data showed correspond to $k=1$. All fluctuations are shown by the algorithm measurement trends.

- The algorithm worked for those patients whose OFF state is characterized with bradykinesia, while it did not work for the rest of them. The main reason is that all patients who participated in the development of the algorithm had bradykinesia. Thus, bradykinesia is the main symptom that the algorithm detects, which agrees with the fact that the algorithm is based on the signal amplitude, as depicted in Figure 5.3. Consequently, patients who did not have bradykinesia could not be monitorized by the online algorithm. Patients 1 and 4 OFF state was mainly characterized by FoG episodes, rather than bradykinesia. Patient 2 barely had motor symptoms. Thus, the developed algorithm used together with a FoG detection algorithm, such as [16], could monitor patients with different motor fluctuations.
- The online algorithm has provided high specificity and sensitivity values in patients with bradykinesia. More concretely, a specificity of 1 and a sensitivity of 0.72 for patient 3 were obtained, respectively; and 0.71 and 0.90 for patient 6. The PPV and NPV values are higher: 1.00 and 0.92 for patient 3, respectively; and 0.80 and 0.84 for patient 6. k values used are 3 for patient 3 and 1 for patient 6.
- Graphical results show a clear correlation between the gold-standard (ON/OFF diaries filled by patients) and the sensor output. However, the gold-standard is a binary value while sensor provides continuous measurements that can objectively detect improvement in patient's treatment. Moreover, it is widely known that ON/OFF diaries are very inexact since patients commonly do not recognize their motor state and they can forget to fill it. Finally, some delays between the gold-standard and the sensor outputs, sometimes in advance and sometimes late, decrease the performance measurements. However, even under these restrictions, the sensor has produced outstanding correlation between them, as Figures 5.8 and 5.13 show.
- OFF measurements obtained in the hospitalization days were, in general, different than those obtained at home. More concretely, hospitalization's OFF

was, usually, less severe than the OFF obtained at home. In this way, it is concluded that the adaptation of the ON/OFF state detection algorithm must be performed through data obtained at home and not in a controlled environment since measurements may be altered.

- One-week analysis is enough to conclude if the sensor and its algorithm are able to detect the motor state of a certain patient. Indeed, it was concluded that the sensor works for patients 3 and 6 and it does not for the rest of patients.

5.3 Conclusions

An ON/OFF detection algorithm based on a single accelerometer placed on the patient's waist has been presented. While dyskinesia algorithm analyzes all patient movements with the exception of walking, the ON/OFF detection algorithm evaluates patient's gait to establish the motor state. The algorithm analyzes the spectral power of strides, which has been shown to reflect the motor fluctuations.

The validation of the ON/OFF algorithm in its offline versions, similarly to the dyskinesia algorithm, was made by means of the signals recorded in 15 patients performing daily life activities in their own environment and for several hours. An online implementation of the algorithm in a low consuming microcontroller has been tested in 7 PD patients in the first pilot in which an apomorphine pump automatically controlled the medication of the patient. Results show that the algorithm effectively detects ON/OFF motor states in patients who suffer bradykinesia in the OFF periods.

Part II

Human movement analysis for the
extraction of gait parameters and
gait recognition tasks

In this part of the thesis, the state of the art and the contributions on the gait recognition and gait parameters extraction are presented. The contributions are based on the accelerometry signal treatment through non-linear signal analysis techniques that reconstruct the state space. More concretely, contributions consist of:

- Extraction of the most important gait parameters: step length and gait speed. These parameters are useful in the evaluation of PD since its variation enable the identification of the motor states [20]. In the chapter devoted to this contribution, a new method for the extraction of the gait parameters based on the state space reconstruction techniques is presented, as well as its application in 10 patients with PD.
- A new method for the recognition of subjects by means of their gait is presented. Its validation by means of 20 users with a 95% accuracy is also detailed.

This second part of the thesis is composed of four chapters:

- Chapter 6 introduces gait analysis based on inertial sensors.
- Chapter 7 presents non-linear signal analysis. This chapter includes a a short literature review of the use of these techniques in gait evaluation.
- Chapter 8 describes the contributions on gait parameter extraction using both time-domain and non-linear signal analysis methods.
- Chapter 9 details three different methods developed for gait recognition.

Chapter 6

Signal treatment methods for gait analysis in the temporal and frequency domains

This chapter describes current accelerometry signal treatment methods in the temporal and frequency domains that enable the extraction of gait parameters and the identification of a person by means of gait.

6.1 Accelerometry signal analysis for the extraction of gait parameters

Currently most widely extended sensors employed in the extraction of gait parameters are accelerometers. These sensors have been located in multiple body positions: waist or the center of the back (at the L3 level) [126] [9] [106] [72]; in the foot, either the ankle [66] or the foot instep [17]; and in the thigh and shank [10].

Current methods for the extraction of gait parameters need to establish the beginning and end of gait cycles. Two different kinds of techniques are distinguished. On the one hand, there exist some methods which rely on biomechanical models in

which the sensor measurements are integrated. These biomechanical models simulate the lower limbs movement by means of a simple inverted pendulum model or a double inverted pendulum model. The first biomechanical model considers the lower extremities as two unique rigid bodies and it expresses the trajectories of the Center of Mass (CoM) during gait by means of a simple inverted pendulum model. The second biomechanical model separates each lower limb into two different rigid bodies and considers the angular rotations performed by each part. These model-based techniques have the advantage of being able to personalize the biomechanical model by including the extremities length of the user.

The second kind of techniques for the extraction of gait parameters make use of direct integration of the sensor values without any biomechanical models. These techniques take advantage of the fact that accelerometers and gyroscopes provide measurements directly related to the parameters aimed to be obtained, which are actually spatial or velocity measurements.

Methods which do not use biomechanical models locate the sensors in one or both feet, as Figure 6.1 shows. These methods use two different sensors [17] [90]:

- First, a gyroscope is employed in order to know the foot slope. The angular rotation signal is integrated in the toe-off event, sometimes assuming that part of the initial slope is zero degrees, and ends with the initial contact. Only one angular rotation signal is needed to establish the foot slope, which is the axis perpendicular to the plane in which the leg movement is contained.
- Once the foot slope is known, then, acceleration measurements are used. The acceleration measurements are reoriented through the instant slope obtained from the gyroscope. Then, the resulting measurements are integrated one or twice in order to obtain the gait velocity and the stride length, respectively.

The simple inverted pendulum model uses a unique sensor, an accelerometer in this case, in the posterior part of the back at the L3 vertebra level. This model assumes a fixed relationship between the stride length and the gait pattern given by:

$$SL = 2\sqrt{2lh - h^2} \quad (6.1)$$

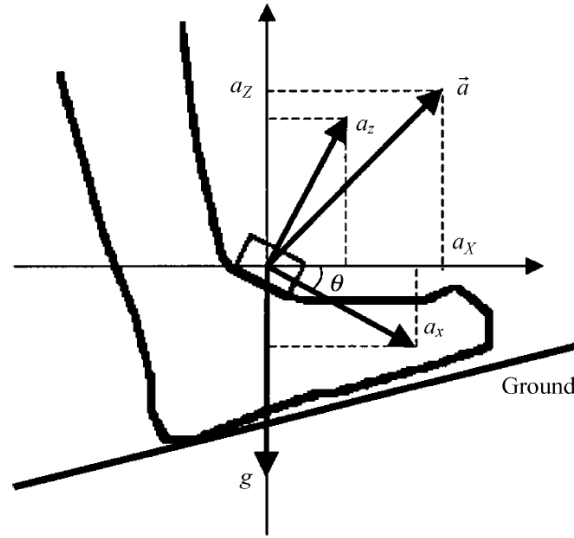


Figure 6.1: Gyroscope and accelerometer used without biomechanical model [90]. Angular rotation provides the current slope θ enabling to separate the forward acceleration (stride length) from the rest of accelerations.

where l is the leg length, h is the vertical displacement of CoM during a gait cycle and SL is the stride length. Thus, after h is computed through the integration of the vertical acceleration, the stride length SL is obtained. Figure 6.2 shows the evolution of h during the gait cycle. Stride velocity is obtained by considering the stride duration.

Different versions of the simple inverted pendulum model have been developed. One of the most extended consists in including an experimental constant in order to adapt the model to the user and reduce the estimation error, so that the relation is transformed into [9]:

$$SL = 2K\sqrt{2lh - h^2} \quad (6.2)$$

The double pendulum model is based on measuring the angular rotation from thighs and shanks to obtain feet's speed and displacement. Figure 6.3 shows the sensors location and the angles measured through the integration of the angular rotation. This technique considers a double pendulum model during the swing phase

and a double inverted pendulum model during the stance phase. Stride velocity is obtained through measuring the time duration of the gait cycle.

The methods presented need to integrate the measurements obtained from either the accelerometer or the gyroscope and, in case biomechanical models are not used, it is necessary, then, to integrate both measurements. The integration introduces drifts in the resulting values, which makes that the values obtained are only valid for short time periods [111]. Thus, the integration periods must be limited. To this end, integrations are performed along a gait cycle and some initial conditions are assumed to facilitate its computation. For instance, when no biomechanical models are used, it is assumed that the initial acceleration value is zero since the foot starts (toe-off event) in a static position. This assumption is not valid if the sensor is located in the waist since it is never in a static position. However, in the inverted pendulum models, it has been proposed to restart the integration when the initial velocity is approximately zero, which coincides with the moment in which both foot are on the ground.

A different method to the previously described ones was developed by Martin et al [72]. This method employs an accelerometer in the waist and enables obtaining the gait speed and the stride length without a biomechanical model by means of wavelets. In this work, the wavelet decomposition permits the authors to establish the gait speed from among 3 possibilities through the energy of weighted details, considered until the fifth level. The energy value obtained, according to this work, shows that the stride length may be estimated with an error below 5%.

The most accurate methods to measure gait parameters are the double pendulum model and those which do not use biomechanical model, since they provide the lowest error. The squared mean squared error (RMSE) corresponding to the simple inverted pendulum model [126] is in the range from 0.08 to 0.25 m/s and from 0.05 to 0.14 m/s depending on the velocity in which the user walked. The RMSE of the double pendulum model [10] is 0.07 m. (7.2 %) for stride length and 0.06 m/s (6.7%) for gait speed. Finally, the method without a biomechanical model [17] provided an error of 0.08 m.

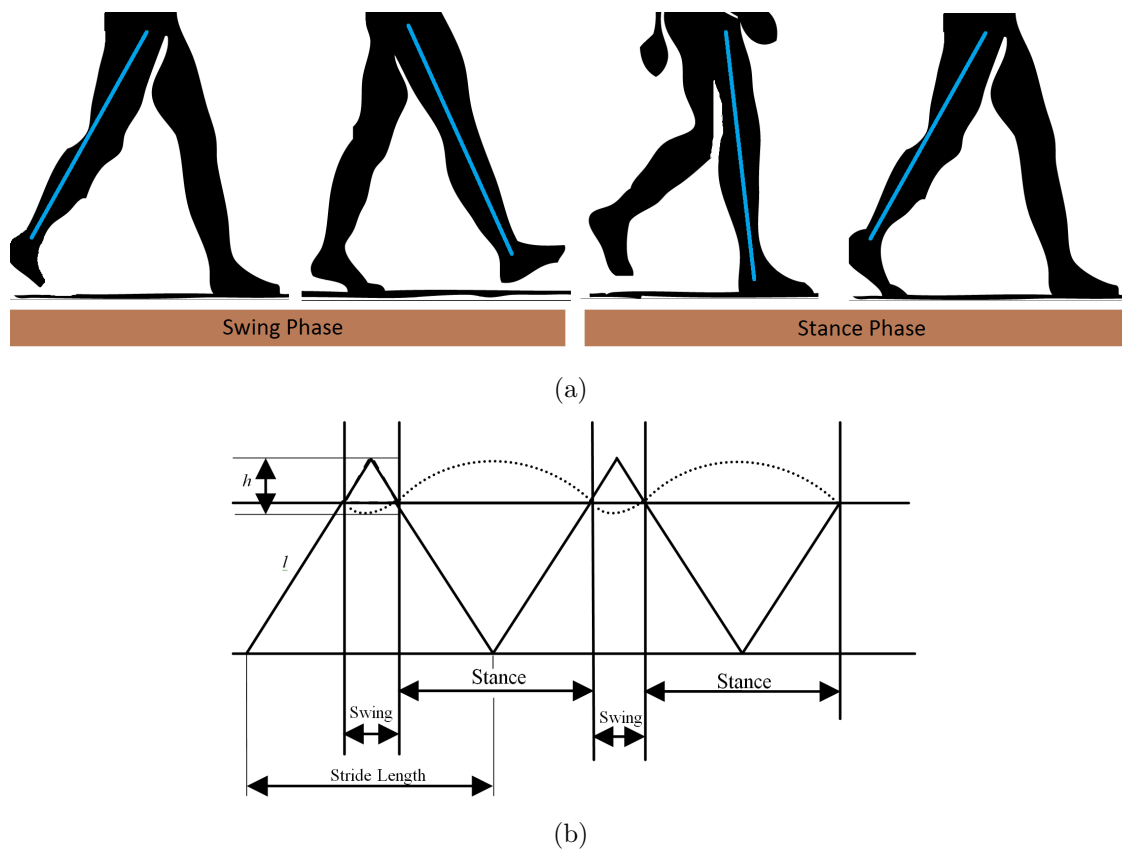


Figure 6.2: The upper figure shows a gait cycle division according to swing and stance phases. The lower figure presents the CoM trajectory during these phases according to the inverted pendulum model.

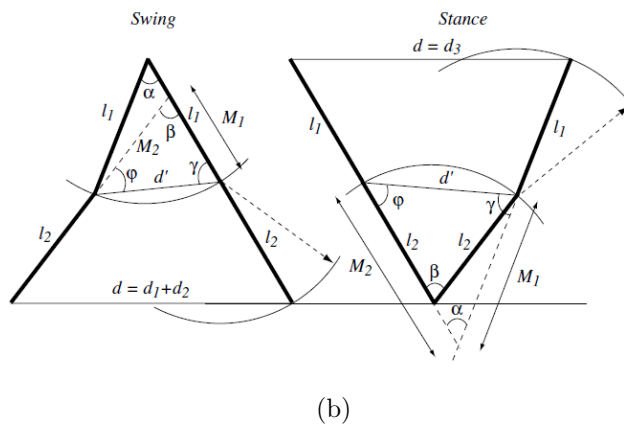
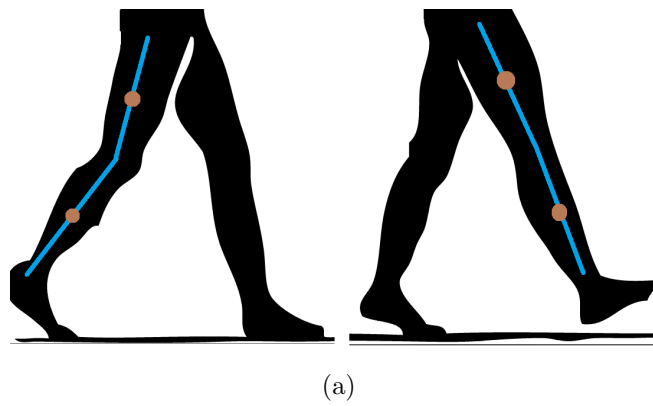


Figure 6.3: On the top, sensor locations in a double inverted pendulum model. On the bottom, model used in [10]. Thigh and shank lengths, l_1 and l_2 , are used to establish α and β angles, for which gyroscope signal is integrated so displacement d is obtained

6.2 Accelerometry signal analysis for gait identification

Gait recognition or gait identification is the term used for the task of recognizing people by means of gait. This is a field of great interest whose research covers security and access control applications. Typical identification systems analyze fingerprints, face, hand and ear geometry, as well as speech or iris. In this section, the most recent methods for gait recognition are presented.

Existing gait recognition methods can be categorised into three groups: vision-based; floor-sensor based; and inertial-sensor based. In the first category, a sequence of images or silhouettes is obtained using a camera. After that, image processing techniques are used to extract gait features for recognition [84]. The main advantage of vision methods is that identification can be performed from afar and without personal interaction. In floor sensor-based gait recognition, a set of force sensors are installed on the floor and recognition is performed using floor pressure data. These methods offer unobtrusive data collection and also provide location information, for instance within a building by using sensors that are unaffected by visual obstructions and light changes. The third category is based on measurements obtained from one or more inertial sensors, normally accelerometers although gyroscopes and magnetometers might be used as well. Gait identification based on accelerometers can be used to protect personal portable devices such as mobile phones, wearable computers, intelligent clothing and other smart devices from unauthorised usage [70]. In the present thesis, we focus on the third category.

Most of the previous studies on inertial-based biometric gait recognition are based on characterizing gait cycles [39] [70] [40] [107]. In these studies, the signals obtained from different users who were walking are used to build a personalized stride model for each user. Then, given another set of signals, the user who provided them is determined by assigning the strides detected in the signals to that model the strides are most similar to. Therefore, the identification is divided into two steps: training and testing. In the training phase, first, signals are segmented into gait cycles. Then, once signals have been divided, each cycle is represented by a feature vector. Finally,

the last training phase consists in learn through a supervised learning method to classify the strides into a class which represents the corresponding user. On the other hand, the testing phase employs the trained supervised learning method to classify the new strides represented by their feature vector.

The accelerometry signal segmentation might be performed through mainly two different techniques. First, the cyclical behaviour of gait enables the usage of relative extrema in the accelerometer signal to segment it into cycles. These techniques have the disadvantage of requiring a previous analysis of the signal behaviour in the specific place in which the sensor is located. For instance, different works have placed the sensor in the waist and have used minima and maxima extrema in the lateral acceleration in order to identify the steps [107] [39] [70]. Second, gait cycles are able to be detected through the identification of those signal parts that are repeated during gait. For instance, the autocorrelation function has been used to treat the vertical acceleration to determine the gait cycles [40]. This method finds local maxima in the autocorrelation function and identifies them as the start of the next cycle. Using autocorrelation function has the advantage to be applicable to any position in which the sensor is located. Figure 6.4 shows the signal segmentation into cycles performed by an autocorrelation function in the same conditions that are described in [40].

The characterization of gait cycles can be performed in several ways. The most simple one consists in normalizing the signals obtained during different gait cycles in length and amplitude and, then, average them to obtain a representative stride of a person through mean [70] or median [39]. A different method consist in computing the probability of obtaining a gait cycle with a specific characteristic, so that histograms with a fixed number of bins are used as a feature vector [40] [39] [70]. Correlations between axis and cumulants [39] have been used, as well, for the characterization of strides.

Most extended current gait recognition methods significantly depend on detecting strides. However, the most common methods for gait cycle detection have been demonstrated to provide a poor performance in ambulatory conditions [71]. Thus, a gait identification method which does not require step detection would improve the existing weakness in current methods.

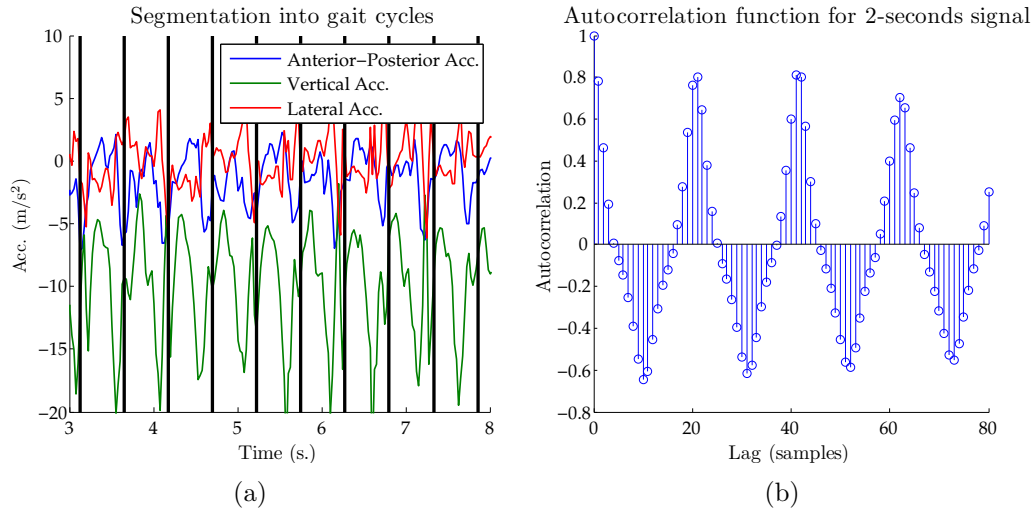


Figure 6.4: On the left, accelerometry signal ($F_s = 40$ Hz) segmented into steps through the autocorrelation function similarly to [40]. On the right, autocorrelation function values for two seconds of signal. A local maxima after 25 samples appears (0.5 s.) with the new step. Another local maxima appears after 50 samples, which corresponds to the new stride.

6.3 Conclusions

This chapter has presented current signal processing methods for extracting gait parameters and performing gait recognition tasks from inertial sensors. Methods for extracting gait parameters depends on sensor location. On the one hand, direct integration is performed when sensors are located on the foot. On the other hand, biomechanical models are employed when sensors are located in the lower limbs or the hip.

Most current gait recognition methods employ step detection methods and characterize steps through features that model gait. Step detection methods also vary according to sensor position. Although the autocorrelation function might be used for this purpose in any position, relative extrema detection depends on the concrete sensor location. The main disadvantage of most current gait recognition methods consists in relying on methods for gait cycle detection, which have been demonstrated to provide a poor performance in ambulatory conditions [71].

Chapter 7

Signal analysis methods based on attractor reconstruction

In this chapter, a special set of non-linear signal analysis techniques are described, which come from the dynamical systems field. These techniques are based on reconstructing the trajectory that a dynamical system describes in its state space by means of a scalar observation of the system. From now on, the state of a dynamical system in a certain time t is denoted as x_t . This state belongs to the state space \mathcal{X} , i.e. $x_t \in \mathcal{X}$. The state of a system is described through a minimum set of variables that identify and unambiguously characterize the system. Thus, the system evolution along time is seen as a trajectory in the state space. This trajectory, after sufficient time, remains limited in specific zone of the state space called *attractor*. Geometrically, an attractor may be a point, a surface, a differentiable manifold or even a complex structure such as a fractal, which gives rise to the *strange attractors*.

The techniques described in this chapter are based on reconstructing the attractor of a dynamical system by means of periodic observations of it. These observations are scalar values and are denoted as s_t . The s_t measurements are the result of an observation function f , i.e. $s_t = f(x_t)$. The s_t value does not offer a complete description of the dynamical system in the same sense that the state x_t does. However, Takens' theorem [110] establishes the conditions for a k -tuple of consecutive observations to enable a topologically equivalent reconstruction of the original dynamical

system attractor.

Takens' theorem considers, first, a function $\varphi : M \rightarrow M$ where M is a m -dimensional compact manifold that defines the evolution of the dynamical system according to $x_{t+1} = \varphi(x_t)$. Let us consider φ and function $f : M \rightarrow \mathbb{R}$ which provides smooth scalar values of the visited states. The evolution of the dynamical system is a sequence $\{x_n = \varphi^n(x_0)\}$, where n is a natural number and x_0 the initial state of the dynamical system. Assuming that at each time step n the system state x_n is not observable but a discrete value $s_n = f(x_n)$ is, during the dynamical system evolution a time series of consecutive measurements $\{s_n = f(\varphi^n(x_0)) = f(x_n)\}$ is obtained. Takens' theorem establishes the conditions for which the so-called reconstruction vectors $\mathbf{Rec}_n^k \in \mathbb{R}^k$ [25] comprised of k consecutive observations $\mathbf{Rec}_n^k = (f(x_n), f(\varphi(x_n)), \dots, f(\varphi^{k-1}(x_n)))$ are a reconstruction of the state x_n , that is, the conditions for which the following function:

$$M \ni x \mapsto (f(x), f(\varphi(x)), \dots, f(\varphi^{k-1}(x))) \in \mathbb{R}^k \quad (7.1)$$

is a diffeomorphism between M and \mathbb{R}^k .

The conditions that Takens' theorem establishes are:

- The number k of employed measurements in the reconstruction vectors must be greater than twice the attractor dimension m .
- (φ, f) belong to an open and dense subset \mathcal{U} included in the product of the diffeomorphisms space Diff^l of M and the space of functions C^l , i.e. $\mathcal{U} \subset \text{Diff}^l(M) \times C^l(M)$, where $l \geq 1$.

For practical purposes, the conditions established in the Takens's theorem may not be strictly satisfied. For instance, the signal sampling and the data discretization due to finite representation prevent the observation function from being differentiable. On the other hand, the presence of noise in both the system and the measurements also prevents the evolution function from being differentiable. Despite of this, Takens' theorem has been widely employed in many different works and the resulting

reconstruction is considered to be a good approximation of the original attractor topology.

In the rest of the chapter, different non-linear signal analysis techniques which employ or are related with Taken's theorem are presented. These techniques have been recently used in the treatment of physiological signals, such as electrocardiograms (ECG) [112] [27] or *Local Field Potential*, a sort of signal similar to electroencephalograms (EEG) [87]. Furthermore, these techniques have been employed to analyze gait and race through different sensors such as movement capture systems based on cameras and electrophysiological signals. The greatest exponent of their use in the gait analysis field is the publication of a *Focus Issue* in Chaos journal in 2009 [79].

7.1 Optimal values for the state space reconstruction

Takens' theorem demonstrates the feasibility of reconstructing the dynamics of a system by means of time series composed of sufficient scalar observations. For practical purposes, the reconstruction process requires the usage and tuning of the time-lag parameter τ and the embedding dimension k . This is a crucial task since the quality of the resulting attractor considerably depends on these parameters [38].

Many different methods have been developed to estimate the optimal value for these parameters. However, different works point that the optimal value greatly depends on the concrete application [58]. In this section, the most extended methods for the estimation of the optimal parameters are presented: *Average Mutual Information* (AMI) for τ and *False Nearest Neighbourhs* (FNN) for k .

Although theory implies that any time-lag enables the reconstruction of the attractor, the usage of a limited set of data and the presence of non-controlled noise provoke that its value considerably influences the correct reconstruction of the attractor. The estimation of time-lag τ is usually done previously to the embedding dimension k since the latter requires the usage of reconstruction vectors, so that τ is necessarily fixed. Instead, the time-lag estimation does not require to know the

embedding dimension, so that the tuning of the value of τ is done previously to k .

7.1.1 Estimation of time-lag τ optimal value

The most extended method to determine τ is AMI, which was developed by Frase and Swinney in 1986. Since the first attempts to set the optimal value of τ , it was observed that using a low value of τ would make the consecutive measurements employed in the reconstruction to be indistinguishable. In such a case, the reconstruction obtained, for instance, in a plane ($k=2$) would be contained in a straight line. In order to avoid this problem, it is assumed that τ value must provide independent consecutive measurements.

According to this principle, a naive method to find the optimal value for τ and that provides independent measurements would consist in requiring a linear independence among them [38]. Consequently, the optimal value for τ would be fixed as the zero-crossing value of the autocorrelation function.

Average Mutual Information employs, instead, the concept of mutual information, introduced by C. Shannon in 1948 [105]. Mutual information is related with the uncertainty reduction in a random variable due to knowing the value of another random variable. In this manner, the general dependency between both variables is measured and not only the linear dependency as the autocorrelation function does.

Mutual Information comes from the entropy definition from Shannon. The entropy in a random variable S measures the *quantity of surprise you should feel upon reading a sample* of S [38] and it is defined in its discrete form as:

$$H(S) = - \sum P(s_i) \log P(s_i) \quad (7.2)$$

where the random variable S may take values s_1, \dots, s_N and $P(s_i)$ is the probability of obtaining the value s_i .

If we assume that S provides a scalar value s_t of a dynamical system and Q provides the measurement $s_{t+\tau}$, $H(Q)$ turns out to be the uncertainty of obtaining $s_{t+\tau}$ and $H(Q|S)$ is the uncertainty of obtaining $s_{t+\tau}$ when s_t is given. Thus, the uncertainty reduction of obtaining $s_{t+\tau}$ given s_t consists in:

$$\begin{aligned}
I(Q, S) &= H(Q) - H(Q|S) \\
&= H(Q) + H(S) - H(S, Q) = I(S, Q)
\end{aligned}
\tag{7.3}$$

The expression in equation (7.3) is the Mutual Information. Optimal value for τ is found observing the evolution of the Mutual Information for growing values of τ . According to the principle detailed at the beginning of this section, small values of τ provoke consecutive measurements to be dependent and, then, the value provided by equation (7.3) would be high. For higher values of τ , the independence among measurements would be higher and the value provided by equation (7.3) would be decreased. Thus, the optimal value of τ was proposed to be the one that provides the first local minima in the sequence of mutual information values [38].

7.1.2 Estimation of embedding dimension k optimal value

The FNN method was developed by Kennel et al. [61] in 1992 and considers the estimation of the optimal value of k through the observation of the reconstructed states in growing embedding dimensions.

Let d_a be the dimension of the attractor to be reconstructed. Kennel et al. observed that, given a reconstruction with $k > 2d_a$, two close states in the state space \mathcal{X} must be also close in the reconstructed space of dimension k , since the value of k is enough to unfold the attractor. On the other hand, if k is too small, states that are far among them in \mathcal{X} could be neighbours in \mathbb{R}^k since the attractor is projected into a low dimensional space. These states are called *false neighbours*. FNN method determines the optimal embedding dimension, the minimum value of k that allows a correct reconstruction, through measuring the percentage of false neighbours for growing values of k .

Let $\mathbf{Rec}_i^k = (s_i, s_{i+\tau}, s_{i+(k-1)\tau})$ a k -dimensional reconstruction vector and $\mathbf{Rec}_{n_i}^k$ its closest reconstruction vector. FNN method assumes that both vectors are false neighbours if including one more dimension makes them not to be close anymore. This way, when most of the states are false neighbours the dimension k is considered to not enable the correct unfolding of the attractor.

Formally, Kennel et al defined that the closest neighbour of a certain state i is considered as a false neighbour either the following condition is satisfied:

$$\sqrt{\frac{\|\mathbf{Rec}_{n_i}^{k+1} - \mathbf{Rec}_i^{k+1}\|^2 - \|\mathbf{Rec}_{n_i}^k - \mathbf{Rec}_i^k\|^2}{\|\mathbf{Rec}_{n_i}^k - \mathbf{Rec}_i^k\|^2}} \quad (7.4)$$

$$= \frac{|s_{n_i+k\tau} - s_{i+k\tau}|}{\|\mathbf{Rec}_{n_i}^k - \mathbf{Rec}_i^k\|} > R_{tol} \quad (7.5)$$

where R_{tol} is a threshold that establishes the minimum distance to recognize a state as a false neighbour, or, on the other hand, if this condition is satisfied:

$$\frac{\|\mathbf{Rec}_{n_i}^k - \mathbf{Rec}_i^k\|}{R_A} > A_{tol} \quad (7.6)$$

where $R_A = \frac{1}{N} \sum_{i=1}^N |s_i - \bar{s}|$ and A_{tol} is a second threshold for short time series.

The second criterium was proposed with the aim of enable the detection of signals with noise or short length. This criteria allowed the authors to correctly identify white noise as a non-deterministic series. Regarding the thresholds R_{tol} and A_{tol} defined in equations (7.5) and (7.6), the same authors found robust results for $R_{tol} \approx 10$ and $A_{tol} \approx 2$.

For practical purposes, the percentage of false neighbours defined by equations (7.5) and (7.6) is observed for growing values of k . It is usually observed that after a certain value of k the false neighbour percentage does not provide a significative decrease. This value is considered the optimal value for k [3].

7.2 Singular Spectrum Analysis

Singular spectrum analysis (SSA) [117] is a nonparametric estimation method that enables the filtering of reconstructed spaces obtained through scalar observations of dynamical systems. SSA is essentially a Principal Component Analysis (PCA),

although other spectra decomposition methods may be considered, that extracts information from short and noisy time series without prior knowledge of its dynamics. SSA unravels the information embedded in the delay coordinate phase space by decomposing the time series into statistically independent components. The idea is to introduce a new set of orthonormal basis vectors in the embedding space so that the projections onto a given number of these directions preserve the maximal fraction of the variance of the original vectors. Solving this problem leads to an eigenvalue problem. The orthonormal eigenvectors determine the *latent variables*. By considering only a few of this directions, those with largest eigenvalues, they are considered to be sufficient to represent most part of the embedded attractor.

SSA first step consists in obtaining the reconstruction vectors in the same way that was described in the previous section. These reconstruction vectors are placed as column vectors in order to fill a matrix. This matrix is denoted as trajectory matrix and has the following form:

$$\begin{aligned}
 M &= (\mathbf{Rec}_1^k \mathbf{Rec}_2^k \dots \mathbf{Rec}_{N-k-1}^k)^T = \\
 &= \begin{pmatrix} s_1 & s_2 & \cdots & s_k \\ s_2 & s_3 & \cdots & s_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{N-(k-1)} & s_{N-(k-2)} & \cdots & s_N \end{pmatrix} \quad (7.7)
 \end{aligned}$$

In general, this trajectory matrix is referred as a Hankel matrix, which means that all the elements along the diagonal $i + j = \text{constant}$ are equal. However, it is true only when $\tau = 1$.

The second stage of SSA consists of determining the eigenvalues and eigenvectors of the positive semidefinite matrix $M^T \cdot M$. Denote the eigenvalues by $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k \geq 0$. For each eigenvalue λ_i there is a corresponding eigenvector v_i such as $(M^T \cdot M) \cdot v_i = \lambda_i \cdot v_i$. The eigenvalues of $M^T \cdot M$ are most often calculated by undertaking a Singular Value Decomposition of M . The right singular vectors of M are identical with the eigenvectors of $M^T \cdot M$ and the eigenvalues are the squares of the corresponding singular values of M .

From the eigenvectors v_i , new vectors $u_i = M \cdot v_i$ of length $N - (k - 1)$ are

defined. These vectors are referred as *latent variables* and are the projections of the embedded time series onto the eigenvectors. If the eigenvectors v_i are chosen orthonormal, i.e. $v_i^T \cdot v_j = 0$, ($i \neq j$) and $v_i^T \cdot v_i = 1$, the latent variables u_i are orthogonal, $u_i^T \cdot u_j = 0$, ($i \neq j$), and $u_i^T \cdot u_i = \lambda_i$. The space formed by the union of consecutive k latent variables, u_1, \dots, u_k is denoted as *k-latent space*.

The most important behavior of a time series is found by taking into account only the first few latent variables. The *scree plot*, which is a plot of the eigenvalues versus the order, can be used to select the relevant latent variables.

In Figure 7.1, the scree plot, the representation of the 2-latent space and the first and last dimension of the 3-latent space for white noise, sinusoidal series and the second variable of the Lorenz series are shown. In all cases, a value of $k = 20$ is considered. In the case of white noise, the scree plot is flat; all eigenvalues are approximately distributed around value 1, which means that the reconstruction needs a large number of dimensions (theoretically infinite). In the second case, the sinus series, only two eigenvalues are significantly different than zero. Thus, the reconstruction can be made by using two dimensions. This reconstruction corresponds to the ellipse in the figure. Finally, one of the variables of Lorenz series is considered. In this case, only three eigenvalues are significantly different than zero although the third eigenvalue is small compared to the first two eigenvalues.

SSA has been widely employed in the treatment of physiological signals. A recent work has demonstrated that Local Field Potentials signals obtained from the brain can be treated with SSA in order to detect if a subject is being exposed to visual stimulus [87]. In this case, components obtained by SSA are more significative than those obtained by *Empirical Mode Decomposition*, a decomposition method that separates a signal into the so-called *Intrinsic Mode Functions* that belong to the temporal domain. On the other hand, SSA has been used to characterize brain's connectivity, which enables, for instance, detecting cerebral dysfunctions [103]. In this work, magnetic resonance imaging have been employed. Finally, SSA has been employed in the treatment of ultrasound images to detect braquitherapy seed implants, showing an effective detection for a reliable clinical use [6].

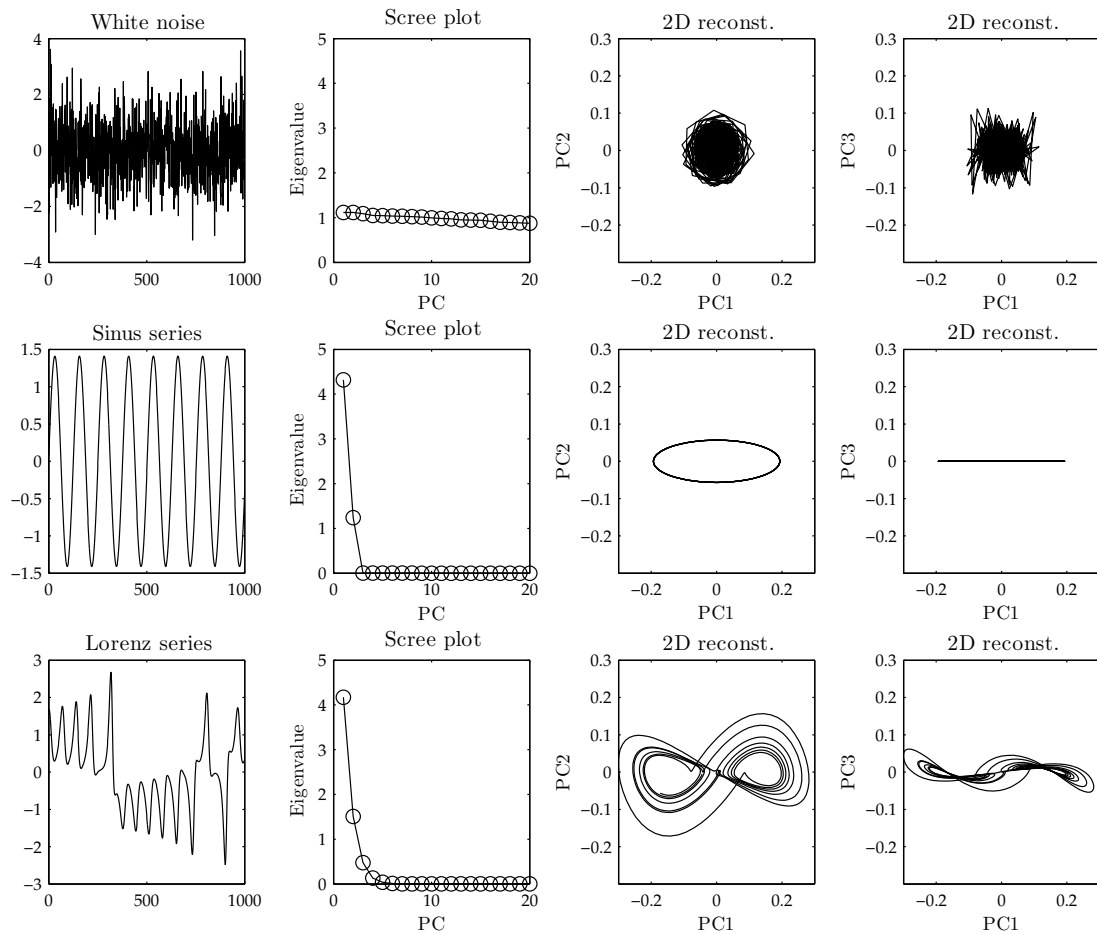


Figure 7.1: Time series, scree plots and 2-d reconstruction for white noise, sinusoidal and Lorenz series.

7.3 Cell-to-cell Mapping

Cell-to-cell Mapping mapping techniques are a sort of techniques which emerged with the aim of solving some issues in the numerical methods that are applied in the dynamical systems analysis. More concretely, it is argued that numerical methods introduce some errors in calculations due to the limited accuracy that computers offer. Moreover, since any experimental measurement is obtained with noise, it is possible to conclude that both in experimental and numerical methods physical quantities are not possible to be obtained in an exact form. For these reasons, it was argued that a state variable should not be treated as a continuous variable but as a discrete one. This consideration guided the development of discrete techniques for the analysis of dynamical systems [53]. The same reasoning is valid for the analysis of dynamical systems through reconstructed attractors, since measurements s_n obtained from the dynamical system inevitably contain noise and are not able to be represented in an exact form.

Discretization performed by Cell-to-cell Mapping methods occurs in a state space or a reconstructed state space \mathbb{R}^k , depending if the dynamical system analysis is considered through its system's state equation or through the embedding given by (φ, f) , respectively. The space \mathbb{R}^k is divided, then, in *cells* that are called *cell state space*. In principle, cells may take any arbitrary form. For practical purposes, k -dimensional rectangular cells are preferably chosen. In this manner, discretization first considers that state space variables or, equivalently, the dynamical system observations s_n have an upper and lower thresholds:

$$l \leq s_n \leq u \quad (7.8)$$

Division of \mathbb{R}^k in rectangular cells may be performed by dividing each one of the k dimensions in C_i intervals of h_i length. Thus,

$$h_i = \frac{u - l}{M_i}, \quad i = 1..k \quad (7.9)$$

This way, the space \mathbb{R}^k is divided in M rectangular cells, with

$$M = \prod_{i=1}^k M_i \quad (7.10)$$

where cells are distinguished by means of an index $j \in \{1, \dots, M\}$.

This way, the dynamical system is not described by a reconstruction vector \mathbf{Rec}_i^k but by the index $\epsilon(t) \in \{0, \dots, M\}$ corresponding to the cell in which the point \mathbf{Rec}_i^k falls. Therefore, different reconstruction vectors may be represented with the same cell and denoted with the same index.

The evolution of a dynamical system may be seen in the cell state space as the indexes succession of the visited cells, i.e. $\epsilon(0), \epsilon(1), \epsilon(2), \dots$, so that the periodic behaviour of the attractor may be represented through the mapping $C : \mathbb{N} \rightarrow \mathbb{N}$, named *Simple Cell Mapping* (SCM), in the following manner:

$$\epsilon(n+1) = C(\epsilon(n)) \quad (7.11)$$

Two sort of cells are distinguished: periodical and transitional cells. A periodical cell satisfies $\epsilon(n) = C^m(\epsilon(n))$, where $m \in \mathbb{N}$. If a cell ϵ is periodical, all cells $C(\epsilon(n)), \dots, C^{m-1}(\epsilon(n))$ are also periodical and make up a group of m periodical cells. A cell which is not periodical is a transitional cell. It is possible to determine the chaotic behaviour of a dynamical system through this cell classification since the chaotic movements would be observed in periodical groups that contain most of the available cells. In this sense, a chaotic attractor would be observable in a set of cells that cover most part of the space \mathbb{R}^k .

SCM and other extensions, such as *Improved Generalized Cell Mapping* [127] that considers the probability of visiting one cell from another one, have demonstrated to be a robust and efficient way method for the analysis of dynamical systems, for instance, as a nonlinear gear system [125] or in a four-wheeled robot similar to a car [45].

7.4 Measurements to characterize dynamical systems in the reconstructed state space

Dynamical systems characterization through measurements obtained from the reconstructed attractor is one of the fields in which there have been a wide research effort since this field has many practical applications. For instance, trending or complexity measurements enable the identification of the current state of a dynamical system and have demonstrated to, for instance, characterize cardiac diseases by means of ECG signals [48] [114], enable the observation of aging effects by measuring the force in the contraction of certain muscles [108] or characterizing fluctuations in electroencephalograms during cardiac failures [68].

Takens' theorem provides the conditions under which a valid reconstruction of a dynamical system attractor may be obtained. Although the resulting attractor is topologically equivalent to the original attractor, the dynamical system evolution observed in the reconstructed vectors is, in general, different to the observed in the original state space. However, it is possible to extract measurements which are invariant to the reconstruction.

Attractor's complexity has been one of the measurements in which the research community has paid more attention. It is characterized as the dimensionality that the reconstructed attractor occupies in \mathbb{R}^k . The reason to measure the dimensionality is that fractional dimensions are an indicator of the presence of chaos in attractors. There are many different methods to obtain the dimensionality, such as *Box-Counting Dimension*, *Information Dimension* or *Hausdorff Dimension*, but *Correlation Dimension* [47] is one of the most employed measurements given its facility to be computed. Correlation dimension computes the correlation integral $C(\epsilon)$ which represents the proportion of each possible pair of reconstructed vectors whose distance is lower than ϵ :

$$C(\epsilon) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i,j=1; i \neq j}^{\infty} H(\epsilon - \|x_i - x_j\|) \quad (7.12)$$

where H is the Heaviside step function.

As the number of reconstruction vectors tends to infinite and the distance among them tends to zero, the correlation integral provides, for small values of ϵ , the form $C(\epsilon) \sim \epsilon^\nu$, where ν is the correlation dimension. Thus, observing the growing of $\log \sum_{i,j=1;\ i \neq j}^{\infty} H(\epsilon - \|x_i - x_j\|)$ on the basis of $\log \epsilon$, a line is obtained whose slope is the dimension ν that data occupy. The correlation dimension has demonstrated to distinguish a chaotic attractor from deterministic behaviour and noise.

Hausdorff dimension is defined through the Minkowski-Bouligand dimension [102], which aims to represent the number of balls $N(r)$, where r is the radius, needed to cover the attractor embedded in \mathbb{R}^k . Hausdorff dimension is the value of d such that $N(r)$ grows as $\frac{1}{r^d}$ when r tends to zero. Formally, let $F(r)$ be the area covered by a circle of radius r located over a fractal curve. Then, if the following limit exists, Hausdorff dimension is defined by:

$$D_M = \lim_{r \rightarrow 0} \frac{\ln F(r)}{-\ln(r)} \quad (7.13)$$

Correlation dimension, in the same way than the Information, Hausdorff or box-counting dimensions, require very large time series to obtain accurate estimations of the dimension. *Statistical Dimension*, instead, is a more accessible measurement of an attractor dimensionality [116] since it can be calculated in short time series. This dimension provides an upper limit to the minimum number of degrees of freedom required to describe the attractor with the data precision used. This dimension can be estimated through SSA: the statistical dimension is considered as the number of eigenvalues above the noise level.

7.5 Gait analysis through non-linear signal methods

In this section, the studies that have analyzed gait through non-linear signal analysis techniques and their most interesting conclusions are presented.

One of the gait parameters that have been widely analyzed has been the time interval between strides. This parameter is similar to the *RR interval*, which consists

in the time interval spent between two heart beats. A beat measured through ECG provides a signal whose peak is named *R wave*, so RR interval is the time interval between two consecutive R peaks. Regarding the gait analysis, currently, it is known that time series composed of the time intervals among strides are selfsimilar, so they satisfy some fractal properties [100]. Moreover, the fractal behaviour of these time series is similar to the behaviour found in the RR intervals [57] since, concretely, both time series from young healthy people showed a higher fractal behaviour than in healthy older people.

Gait in PD patients has been also evaluated through non-linear signal analysis techniques [50]. Concretely, the relation between step length and gait variability has been analyzed, at the same time than the effect of some experimental conditions. The most common way to determine gait variability is to analyze the first and second order statistics of time series composed of the stride length or its duration. Instead, in Hausdorff's work [50] it is demonstrated that, although in healthy people a stride is related, at least statistically, with hundreds of subsequent strides, the relation between strides is lost in PD patients. From a medical point of view, this result is interesting since it means that the locomotor system in healthy people has a sort of 'memory' and a long-term correlation, while it does not occur in PD patients. Causes attributed to this change is the lost of gait automation due to the disease, as can be observed during FoG episodes.

In another paper from the same Chaos special issue [79], Nessler et al. analyzed the movement synchronization among humans during gait, that is, walking *side by side* [83]. Concretely, they analyzed the time variability between consecutive strides while subjects 1) walked alone, 2) walked freely and side by side with another person and 3) walked with forced synchronization between steps and side by side with another person. Variability among strides was incremented during side by side walk in respect of walking alone. Moreover, gait was synchronised unintentionally during free walks side-by-side. Finally, the variability found in healthy people compared to people with some motor disabilities was higher. Thus, the authors suggested that variability in certain aspects of the performance may be an indicator of a healthy system.

Finally, gait has been previously analyzed with biometric purposes in the recent

work of J. Frank [37]. The objective consisted in performing gait recognition tasks through the triaxial accelerometers included in the mobile phones. In this work, gait's attractor was reconstructed and was analyzed through SSA. The reconstruction parameters were arbitrarily fixed without using FNN or AMI. The user identification was performed in the latent variables space. The method keeps the trajectory followed by the attractor in the training data and compares these trajectories to those obtained from the testing data by comparing all the reconstructed states by means of the so-called *Geometric Template Matching* algorithm. Results show an excellent accuracy in gait recognition tasks. However, the algorithm requires to store trajectories to perform the recognition, which may be computationally expensive depending on the value of m .

7.6 Conclusions

This chapter has presented attractor reconstruction techniques that come from dynamical systems theory. These techniques enables us to deal with dynamical systems whose state space cannot be precisely known. The reconstruction of their attractors through scalar time series enable their characterization, which has been useful to treat physiological signals, as it has been shown in the literature reviewed. In the last part of this chapter, the application of these non-linear signal analysis techniques into gait evaluation has been presented.

Attractor reconstruction depends on parameters τ and k , whose optimal values are estimated by two mathematical approaches, AMI and FNN respectively, that have been detailed in this chapter. Moreover, two different methodologies that employ the attractor reconstruction to analyze a dynamical system have been described: SSA and Cell-to-cell mapping. On the one hand, SSA filters the trajectory matrix that contains dynamic information through a PCA analysis. On the other hand, Cell-to-cell mapping employs a qualitative approach to describe the dynamics of a given system. Both approaches will be exploited in the following chapters in order to extract gait parameters and perform gait recognition tasks.

Chapter 8

Gait parameters extraction in PD patients by means of accelerometers

Many motor complications affect spatiotemporal properties of the PD patients' gait. Length and speed of the steps change due to the increment of the impact of some diseases. In this section, contributions on gait parameters extraction in PD patients are presented. Contributions are twofold: first, an approach that works in the temporal domain is given and, second, a method which employs non-linear signal analysis techniques is described. The contributions presented have been published in two conferences: Neural Information Processing Systems in 2009 [92] and International Joint Conference on Neural Networks in 2010 [93]; and in Neurocomputing journal in 2011 [91].

The sensor device used is sewed to a belt in order to obtain two main properties: step length and step velocity. The signals database used are those gathered in MoMoPa's first phase, which has been described in the third chapter. It is composed of movement signals from 10 PD patients obtained in laboratory and the length and speed of their steps.

8.1 Gait parameters extraction in the temporal domain

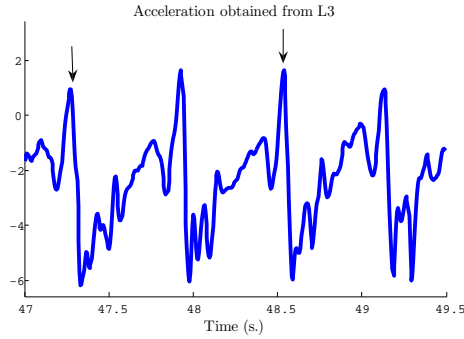
In this section, the extraction of gait properties is performed in the temporal domain. The sensor, as in the previous experiments, was located in the lateral side of the waist. When standing in anatomical position, the orientation of the accelerometer conforms ISB recommendations [123], i.e. positive x - values correspond to anterior acceleration, positive y - values to upward acceleration and positive z - values to the acceleration to the right.

Accelerations signals obtained from the lateral side of the waist differ from usual signals obtained from the region near L3. Figure 8.1 shows anterior-posterior acceleration on normal gait obtained by the sensor from both locations. The signal obtained from the L3 region follows the behaviour reported in the literature [126], where negative peaks of the anterior-posterior acceleration are due to the end of the single support phase and the beginning of double support phase. These peaks are preceded by a positive peak, produced in the feet-floor contact. On the other hand, when the accelerometer is located in the lateral side of the waist, the negative peaks are also observed, but not all the positive peaks appear. Only those contacts generated by the foot of the side where the sensor is located produce positive peaks. This way, gait parameters extraction in the lateral side of the waist must take into account the changes provided by the new position of the sensor.

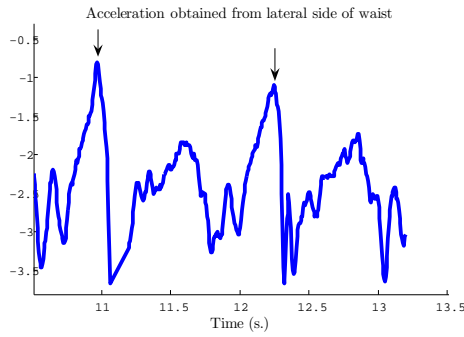
8.1.1 Methodology

A two-phases methodology for the extraction of gait parameters is presented. The first phase consists in segmenting the signal into steps and characterizing them through features. The second phase establishes the relationship between these features and the gait properties through a non-linear regression.

Biomechanical characteristics of gait allow to automatically identifying steps from tri-axial acceleration signals [126]. Figure 5.2 shows how segments of acceleration signals related with a step are automatically detected in the step detection method



(a) L3 region.



(b) Lateral side.

Figure 8.1: Forward acceleration obtained from L3 region and forward acceleration obtained from the lateral side of waist. For the L3 region, acceleration is similar between left and right steps. Similarity between steps is altered when the sensor is located in the lateral side of waist.

previously described in Section 5.1. This method takes into account that the steps performed by the leg corresponding to the same side of the sensor are measured with a greater amplitude than the opposite ones and, moreover, exploits the same principles that were described in [126] while adapts them to the new position. Acceleration signals are low-pass filtered previously to the step detection analysis using a second-order zero-lag Butterworth filter with a cut-off frequency of 15Hz, which is enough given that 99% of energy is contained below 15Hz [14].

Once the acceleration signal is segmented into steps, these steps are characterized through different features. To this end, let us define as:

$$\mathbf{a}(k) = (a_x(k), a_y(k), a_z(k)) \quad (8.1)$$

to the acceleration vector provided by the sensor at time k and its magnitude as $r(k)$:

$$r(k) = \sqrt{a_x^2(k) + a_y^2(k) + a_z^2(k)} \quad (8.2)$$

Based on the acceleration vector $\mathbf{a}(k)$ and its magnitude $r(k)$ along each time-variant segmented acceleration signals, i.e. each step, three features are empirically defined to characterize it:

1. The mean of acceleration magnitudes,

$$\rho_{1,s} = \frac{1}{T_s} \sum_{k=k_{s,0}}^{k_{s,0}+T_s-1} r(k)$$

where T_s is the number of samples during the s^{th} step, and $k_{s,0}$ the starting sample for the s^{th} segmented step.

2. The sum of the absolute value of the components,

$$\rho_{2,s} = \sum_{k=k_{s,0}}^{k_{s,0}+T_s-1} (|a_x(k)| + |a_y(k)| + |a_z(k)|)$$

which is related to well-known energy expenditure indicators [22].

3. The mean of the absolute values of the time-step increments,

$$\rho_{3,s} = \frac{1}{T_s - 1} \sum_{k=k_{s,0}+1}^{k_{s,0}+T_s-1} |r(k) - r(k-1)|$$

i.e the mean of the jerk absolute value.

These features are affected by the stride dynamics since, for instance, the energy expended on a stride is related to its length and velocity. The faster the step is the higher the acceleration signal values and its increments are, so the mean norm and the mean jerk reflect it.

Additionally, the length and velocity of a given step, measured directly from the experiments, are denoted as l_s and v_s , respectively. The problem at hands can be formulated as finding out mappings $f(\cdot)$ and $g(\cdot)$ such that,

$$l_s = f(\rho_{1,s}, \rho_{2,s}, \rho_{3,s}) \quad , \quad v_s = g(\rho_{1,s}, \rho_{2,s}, \rho_{3,s})$$

The ϵ -SVR [101] was the kernel method selected in order to extract this relationship from experimental data as it is able to establish non-linear relations between input and output.

8.1.2 Experiments and results

Ten PD patients were asked to walk several times over a plain surface of 6 m. length while a tri-axial accelerometer recorded the measures as part of MoMoPa 1st phase. As mentioned before, the accelerometer was attached to a belt and it was approximately located at the lateral side of the waist. Its exact orientation depended on how the belt was worn by the patient, thus quite different positions and orientations were used between patients. Experiments were recorded by a video camera. Two methods were used to obtain actual step length: footprints and visual markers. Firstly, patients's shoe soles were painted and footprints left were measured. Secondly, visual markers distributed every 30 cm. were used to determine by video recordings step lengths. Actual step velocity was obtained by dividing step length by its duration obtained from recordings. These actual values are used as ground truth required to perform regressions against the sensor information. Local ethics committee approved the study, and subjects participation was informed consent.

A ϵ -SVR with a cubic polynomial and a Radial Basis Function (RBF) kernel is designed based on the defined features. In order to evaluate its prediction ability, a randomly selected set composed by 80% of the steps is used to train the ϵ -SVR, and the remaining data are used to establish mean squared error (MSE) rate. Finally, ϵ -SVR is also compared against linear regression on the same training and evaluation sets for each repetition.

Results for step velocity are summarized in Table 8.1.2. It can be observed that

the mean MSE of ϵ -SVR is significantly lower than the error obtained with linear regression. Moreover, the cubic kernel provides the lower error, a mean RMSE error of 14.64 cm/s (MSE=214.37) for predicting new step velocities, which outperforms the RBF kernel.

Features used on predicting step length are the same used on step velocity combined with duration time of the step. Results are also summarized in Table 8.1.2. MSE value for ϵ -SVR is much lower than obtained on linear regression. However, an MSE value of 340.9 cm² is obtained, which is greater than the obtained for the case of velocity. It means that prediction of the step length is more complex than the estimation of the step velocity using the same set of features.

This way, since the cubic kernel provides the best results to estimate the step velocity and the step length in the temporal domain, it is chosen as the optimal one. Its results will be compared in the following section against the results obtained through the reconstructed attractors methods.

	Step velocity	Step length
Regression	Mean MSE (cm/s) ²	Mean MSE (cm ²)
Linear	301.5	605.5
ϵ -SVR (RBF)	233.08	381.42
ϵ -SVR (cubic)	214.37	340.9

Table 8.1: Summary of results for step velocity and step length.

8.2 Gait parameters extraction through reconstructed attractors

As a second approach in gait analysis, human gait is analyzed as a dynamical system where internal states behave differently according to the dynamic state of the system. Gait parameters are considered as unknown intrinsic properties of that dynamical system. Acceleration signals are processed through a Singular Spectrum Analysis approximation, in which PCA is applied to a conveniently organized set of time series provided by sensor measurements, according to previous sections. As it will be shown,

SSA unravels the internal dynamics of the system and correlate it with the actual spatiotemporal properties obtained during gait.

8.2.1 Methodology

The reconstruction of the attractor of human gait for the extraction of gait parameters follows a methodology that is designed from the following observations:

- The size of the reconstruction space is determined by the chosen dimension m , which should be large enough for leading to a space with capacity to capture the system dynamics. At the same time, using a large number of sensor data, i.e., a large time-series, leads to a large representation space with the valuable information spread along the matrix columns.
- The use of PCA transforms the trajectory matrix M , a huge database of time-series provided by sensor signals, into a new matrix in which the relevant information of the dynamical system behavior is concentrated in first latent variables. Thus, only these latent variables are employed, and the discarded latent variables are considered to contain only noise. Distinguishing between relevant or noisily variables is performed by observing its contribution: noisily variables barely contribute in comparison to the relevant ones. Figure 8.2 shows an example of both, the original magnitude for the acceleration and the first and second more informative latent variables during a gait episode, which includes fast and slow walking dynamics. In this figure, $\tau = 4$ and $m = 20$ values were employed. Vertical line discriminates walking velocity during a gait episode, the first cycle corresponding to a gait velocity of 35 cm/s, while the second one represents a magnitude of 189 cm/s.
- It can be observed how latent variables behave similarly to a sinusoidal signal. A direct relationship exists between the new signal amplitude and human gait velocity. Therefore, the problem of estimating the stride length and velocity could be solved by using regression on latent variables. As depicted in Figure 8.2(b), both latent variables would discriminate a slow gait episode from a fast

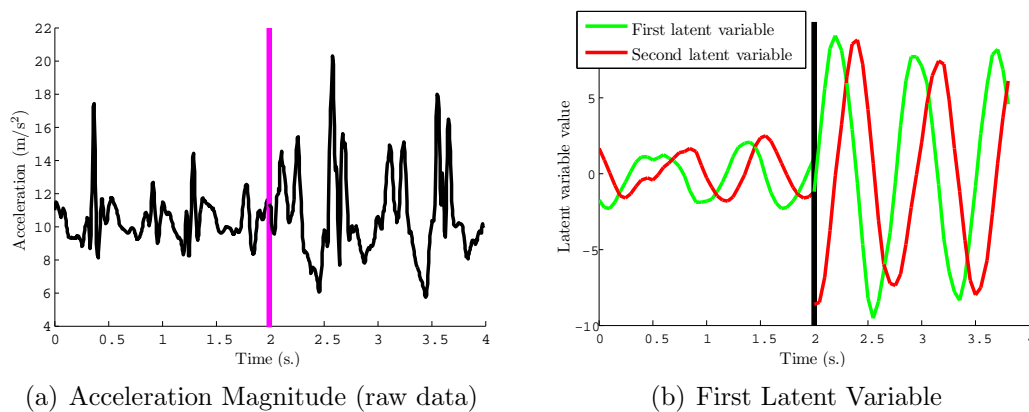


Figure 8.2: Original and latent variables during both slow and fast gait. Amplitude of the latent variables during fast gait are higher than in the slow gait.

one, therefore, for the estimation of step length and velocity both variables ($\mathbf{y}_1, \mathbf{y}_2$) are chosen as input space for the regressor. First two latent variables contain main information described as covariance, as Figure 8.3 presents. Only these two latent variables are out of the noise level.

Considering these observations, the methodology used to estimate gait parameters in the reconstructed space consists of:

1. Segmenting the accelerometer signal into steps by considering the lateral waist position of the sensor. To this end, the same step detection method described in the previous section, and that was also used in Section 5.1, is employed.
2. Once steps are detected, the signal comprised in each step is used to reconstruct the dynamics of gait by using SSA on the acceleration magnitude as in the preceding examples. Therefore, only two of the resulting latent variables are used, those which are above the noise level.
3. Since the amplitude of the resulting latent variables increases together with gait speed and stride length, the latent variables obtained from each step segmented are characterized through ρ_4 :

$$\rho_4 = \max_j(w_j \cdot \tau) \quad (8.3)$$

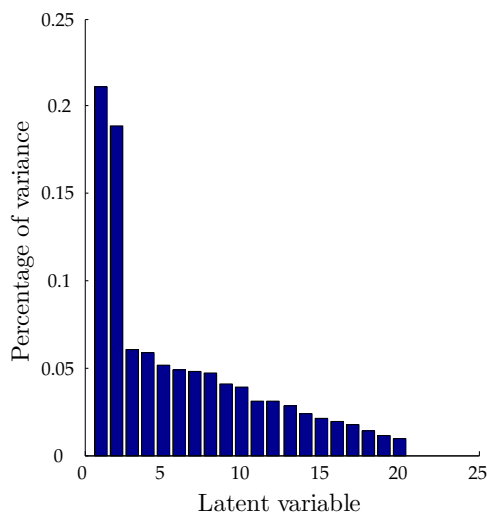


Figure 8.3: Eigenvalues of each latent variable. The first two latent variables comprise almost half of the variance of the trajectory matrix.

where

$$w_j = \sqrt{(y_{j,1})^2 + (y_{j,2})^2}. \quad (8.4)$$

The sum of the square components in (8.4) represents the instantaneous radius of the trajectory in the space formed by the first two latent variables. A graphical representation is presented in Figure 8.4, where two different gait velocities are described. The values of the first latent values are shown against the values of the second one. Thus, the trajectories in the reconstructed and filtered attractor are observed.

In this work, since we are measuring gait properties, recurrence plots may clarify whether our reconstruction is valid. Recurrence plots are a common technique helpful to visualize the recurrences of dynamical systems [35]. Given a sequence of (reconstructed) states x_1, \dots, x_n of a system, a matrix $M_{n \times n}$ is considered where each element m_{ij} may have two values: 1 when $x_i \approx x_j$ and 0 when not. Note that similarity is defined by ϵ -insensitivity. This matrix is plotted, and periodic motions are reflected by long and uninterrupted diagonals. The vertical distance between these lines, which is measured in states, corresponds to the period of the oscillation. Figure 8.5 shows the recurrence plot for a patient when using embedding dimension 30 and $\epsilon = 10.5$. Similar results were obtained by the rest of patients. Periodic orbits

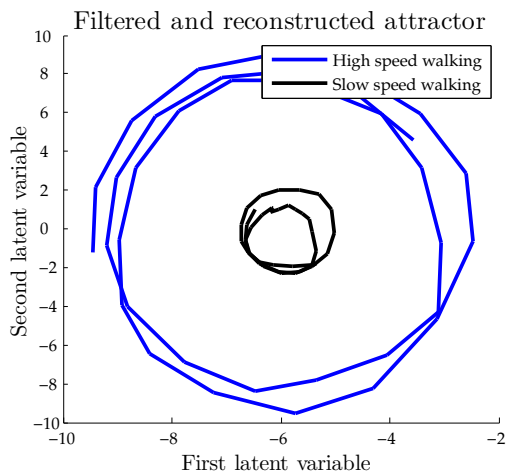


Figure 8.4: Latent space for fast and slow gait. The filtered reconstructed attractor is shown for both fast and slow gait. Different amplitudes are observed.

are clearly seen which completely agrees with the periodic nature of gait, therefore suggesting that a good reconstruction is obtained.

8.2.2 Experiments and results

The method presented was applied on the same data used in the previous section for comparison purposes. This way, this section compares the estimation results of spatiotemporal gait parameters using the reconstructed attractors feature ρ_4 against the temporal features ρ_1, ρ_2, ρ_3 from the precedent experiment.

Table 8.2 shows the best results obtained when predicting length and velocity for each regression model ($m=30$, sampling at 50Hz). In order to evaluate the prediction capability of the regression using raw data, a randomly generated set (80% of the samples) was used to train ϵ -SVR model, the remaining samples were used to establish the MSE error rate. This process was repeated one hundred times to obtain significant values. A 10-fold cross validation was performed in the training data in order to set the ϵ -SVR parameters. This process was repeated 30 times to establish significant values. For the embedding approach, the best result was obtained when using ϵ -SVR

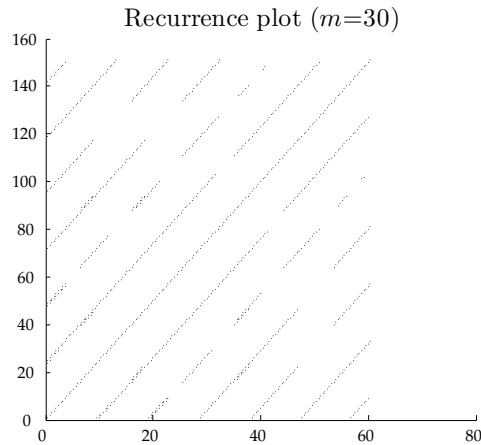


Figure 8.5: Recurrence plot obtained from the accelerometer signal registered during patient 1 gait. Gait cycles are reflected by long and uninterrupted diagonals

with a RBF kernel for both spatiotemporal parameters, which improves the results provided by other type of regressions, (e.g., polynomial or linear kernels). When using those features extracted from raw data, the best result was that provided by the ϵ -SVR, using a cubic polynomial kernel. From the medical point of view these results may be accurately enough to value the subjects' gait performance.

	Step velocity	Step length
Regression	Mean MSE (cm/s) ²	Mean MSE (cm ²)
Direct (ϵ -SVR cubic kernel)	214.37	340.9
Embedding(ϵ -SVR RBF kernel)	110.2	240.2

Table 8.2: Summary of results for step velocity and step length.

A more understandable prediction error measures of the results are listed in Table 8.3 and 8.4, where a specific estimation and its errors are presented. The relative error between the RMSE and the average value of each spatio-temporal parameters is presented.

Approach	RMSE (cm/s) - (%)
Embedding (ϵ -SVR RBF kernel)	10.50 – (15.3%)
Direct (ϵ -SVR 3-degree polynom. kernel)	14.64 – (21.3%)

Table 8.3: Velocity Error Case Analysis (Average Velocity = 68 cm/s)

Approach	RMSE (cm) - (%)
Embedding (ϵ -SVR RBF kernel)	15.49 – (18.6%)
Direct (ϵ -SVR 3-degree polynom. kernel)	18.46 – (22.18%)

Table 8.4: Length Error Case Analysis (Average Length = 83.24 cm)

8.2.3 Discussion

Table 8.2 compares results from embedding approach against those obtained by direct model, which are also presented in Table 8.1.2. Time embedding approach outperforms the direct model results. A single value calculated from the reduced reconstruction space outperforms predictions based on different features of the original signal. Figure 8.2 suggest that it would be easier to extract information from latent variables, and estimation results evidence it.

Estimation errors are presented in Table 8.3 and 8.4. The embedding approach shows better performance on estimating both spatiotemporal parameters. Relative errors between RMSE and the average value are also shown. Embedding approach provides an error of 15.3% for step velocity and 18.6% for step length. This apparently high error may not be a problem in most medical applications, since step length and velocity assessment does not require an extremely precise measurements. A long ambulatory assessment of the step length and velocity of patients can be performed through this estimation method since significative changes will be reflected on the estimations.

A previous study which used 4 gyroscopes reported a RMSE of 7 cm/s for step velocity and 8 cm for step length [11]. Sensors should be located carefully in thigh

and shanks, and they were sensitive to orientation. Such results are more accurate than those obtained by the embedding approach used in this study. However, it must be taken into account that this lower precision is a consequence of usability restrictions. An unique wearable sensor attached to waist, which must be insensible to orientation provides quite different and irregular signals between patients. Thus, embedding approach has been able to adapt to such signals but resulting in higher errors.

Reconstructing the state space of human gait is shown to be more precise than using signal features to estimate parameters of gait. It is possible that other features could obtain more accurate estimations. In order to find them, a large process comprising feature definition, feature selection and testing should be done. By contrast, state space reconstruction combined with PCA analysis has provided a representation of gait easily interpretable that made feature definition obvious providing more accurate results.

Step velocity estimations are more precise than those obtained for length. The reason may be caused by the nature of the sensor measures, which are accelerations. A measure error induced by noise may affect length regression stronger than velocity since there is a double cumulative relation which is likely to disturb.

8.3 Conclusions

In this chapter, contributions on gait parameters extraction in PD patients have been presented. Two different methods have been developed: the first method employs the temporal domain and the second one exploits non-linear signal analysis techniques based on reconstructing human gait attractors.

The temporal-domain method uses a two-phases methodology for the extraction of gait parameters. It first segments the signal into steps and characterizes them through features. Then, it establishes the relationship between these features and the gait properties through a non-linear regression. On the other hand, the non-linear signal analysis method includes another phase in which, after detecting steps, signals are used to reconstruct the dynamics of gait by using SSA on the acceleration

magnitude. Then, the amplitude of the resulting latent variables are used as an input to a non-linear regression.

Reconstructing the state space of human gait is shown to be more precise than using the temporal-domain features employed to estimate parameters of gait in the specific sensor position used. State space reconstruction combined with PCA analysis has provided a representation of gait easily interpretable that made feature definition obvious and providing more accurate results.

Chapter 9

Gait recognition by means of a single waist-worn accelerometer

In this chapter, the contributions made in this thesis in the field of gait recognition by means of gait signal accelerometry signals are presented. The gait recognition methods developed have been tested in a different signal database than the used in the previous chapters. In this case, signals from healthy users have been obtained using the same sensor location that was used in the signal gathering from PD patients, which was the lateral side of the waist. Signals were obtained while healthy volunteers walked 20 m. at normal speed twice. Before walking the second time, users removed the belt and located it again. Thus, signals from second walk were obtained in, at least, a slightly different position. Signals from first walk were used to train the methods and signals from second walk were used to test them.

The first contribution presents an approach that employs the temporal and frequency domain to recognize people from gait. The second contribution presents a qualitative approach to granularly represent the gait reconstructed attractors is presented, which employs Cell-to-cell mapping technique. Finally, a new gait recognition method called ‘Box Approximation Geometry’ that uses Singular Spectrum Analysis to filter the gait reconstructed attractor is detailed. The three contributions have been presented as two paper conferences in the International Workshop on Qualitative Reasoning (2012) [95], in the International Work-Conference on Artificial Neural

Networks (2011) [97] and in the Neurocomputing Journal (2013) [96].

9.1 Gait recognition in the temporal and frequency signal domains

The gait recognition method described in this section is denoted as a time and frequency domain method in order to differentiate them from those presented in the next sections that performs a reconstruction of the state space.

According to the state of the art presented, most common gait recognition methods perform the identification of a person by means of gait cycles. More concretely, a representative gait cycle for each user is constructed in a training phase. Then, the recognition process consists in comparing the new stride to be recognized to the representative cycles obtained in the training phase. Finally, the new stride is decided to belong to the user whose representative cycle is most similar to the new stride. Thus, a segmentation process is followed in both training and validation phases in order to obtain the strides from the acceleration signal.

9.1.1 Segmentation process

The objective of segmenting is to divide accelerometer signal into gait cycles, i.e. each stride is isolated. This process can be performed by two different main methods, as described in the state of the art. On the one hand, a simple minimum peak detection based on lateral acceleration has been used in [107] [39] [70]. This kind of analysis is valid only when the accelerometer is located at the waist. A different analysis would be necessary if the sensor was attached to a different position. On the other hand, autocorrelation function of vertical acceleration may be used for step detection [40]. This method would be valid for almost any position that the sensor can be attached to.

The minimum-peak-based detection has been observed to fail in many users since some gaits are irregular and produce occasional minimum peaks that result in a bad segmentation. Thus, in this work, strides are detected by using the autocorrelation

function in the vertical acceleration. Figure 6.4 shows an example of a segmentation process result by using autocorrelation function.

9.1.2 Representative cycle

Once cycles from training data have been extracted, it is necessary to construct a representative stride in order to compare it to the new cycles to be classified. One of the simplest approaches consists of normalizing the cycles by length and amplitude and, then, average them to obtain the representative stride by means of the mean [70] or median [39]. Another very common approach is to use a histogram with a fixed number of bins [40] [39] [70], where features represent the number of values relying in each bin. Note that normalization in time and amplitude is also needed. Correlation between axis [40] [39] [70] and cumulant coefficients of order 1 to 4 [39] has been also used for characterizing gait cycles. In this work, the combination of all these features that have been shown to enable the gait recognition accuracy have been employed.

9.1.3 Gait recognition experiments

The sensor used in the experiments gathered measurements from the triaxial accelerometer at 200 Hz. In order to use the minimum number of samples required, since frequency content is known to be below 20 Hz [14], signal is resampled from 200 Hz to 40 Hz.

During the signal segmentation into steps, it was observed that left and right steps of a user may be quite different. A classification based on steps should take into account those differences, but it has been also observed that the stride composed by both of them remain constant in a user. This observation was also remarked in [70]. Thus, the elements considered to be identified are the strides.

Once a stride has been identified, features are extracted from it. The histogram technique is applied after normalizing in time and amplitude. As in [40] [70], each stride is interpolated into 100 values and the quantity remained in the 10 bins of the histogram are used to characterize the step. Correlation between axes, kurtosis and skewness are also used, similarly to [70].

9.1.4 Results

Time-domain gait recognition training process is performed by Classification and Regression Trees (CART) [24] using the strides obtained in the first 20 m. walked by 5 volunteers. Accuracies are obtained by classifying the second 20 m. walk. CART methodology used is the standard cross-validation prune, where the optimal tree is the one with least nodes whose accuracy is within 1 standard error of the minimum cost tree.

Time-domain gait identification achieves 82% of accuracy for the 5 users. This percentage is similar to the result described in [39] (86%), and quite below than the percentage in [40], which was 95% although this paper located the accelerometer at the leg. Accelerations from leg show clearer the behavior of gait than those obtained when the sensor is at the waist, as can be observed in Figure 5.1, which may explain the higher performance.

9.2 Granular approach for gait recognition

Granular Computing is a general computation paradigm that performs information processing tasks by considering granularity levels of information such as classes, subsets, groups and intervals to build an efficient computational model for complex applications with huge amounts of data. This section describes the granular approach followed, which proposes different abstract levels to granularly represent reconstructed attractors in a suitable way to perform gait recognition.

Let the reconstructed trajectory described by the k first latent variables above the noise level be:

$$\mathbf{U}_k = [\mathbf{u}_1 \ \mathbf{u}_2 \ \dots \ \mathbf{u}_k] = [\mathbf{w}_1 \ \mathbf{w}_2 \ \dots \ \mathbf{w}_{N-(m-1)\cdot\tau}]^T \quad (9.1)$$

where the rows of \mathbf{U}_k , named \mathbf{w}_j , are the reconstructed trajectory points in the latent space \mathbb{R}^k .

The shape of the *filtered* trajectory (filtered because noisy contribution are rejected) depends on the characteristics of the system and is expected to allow the

identification of the dynamical system and hence the person whose gait has been measured. A previous work on Gait Recognition [37] used representative model trajectories to characterize each possible person or person activity. In that work, the new trajectory to be classified is compared to each of them and considered to belong to the person which representative trajectory is most similar.

However, a granularity problem may appear. The comparison between trajectories can be easily done between filtered time-delays, i.e. comparing rows in \mathbf{U}_k . However, higher gait speeds provide wider orbits, as showed in the previous section, and fewer states included in a gait cycle [91], so that state-by-state comparison may fail when the trajectories to compare were obtained at different gait speeds.

Abstraction levels are proposed to perform gait recognition tasks by avoiding this issue. Since the trajectory shape is similar but wider at different gait speeds and the number of states included in a gait cycle is different, a density estimation measurement that depends on an abstraction level is proposed. This density estimation may be seen as a density measurement of the cell-to-cell mapping, in which the cell sizes are controlled through the abstraction levels.

9.2.1 Granular approach

Given an integer abstraction level a , a filtered trajectory matrix \mathbf{U}_k is represented by its pattern $\mathbf{d} = \{d_{i_1, \dots, i_k}\}$, $i_j \in \{1, \dots, a\}$ that contains the density estimations for the corresponding abstraction level.

Let us assume that \mathbf{U}'_k is the matrix \mathbf{U}_k which each column is normalized between 0 and 1:

$$\mathbf{U}'_k = [\mathbf{u}'_1 \ \cdots \ \mathbf{u}'_k] = [\mathbf{w}'_1 \ \cdots \ \mathbf{w}'_{N-(m-1)\cdot\tau}]^T \quad (9.2)$$

where

$$u_j^i = \frac{1}{\max(\mathbf{u}_j) - \min(\mathbf{u}_j)} \cdot (u_j^i - \min(\mathbf{u}_j)) \quad (9.3)$$

then, d_{i_1, \dots, i_k} is the proportion of the filtered reconstructed attractor in the corresponding region of the latent space:

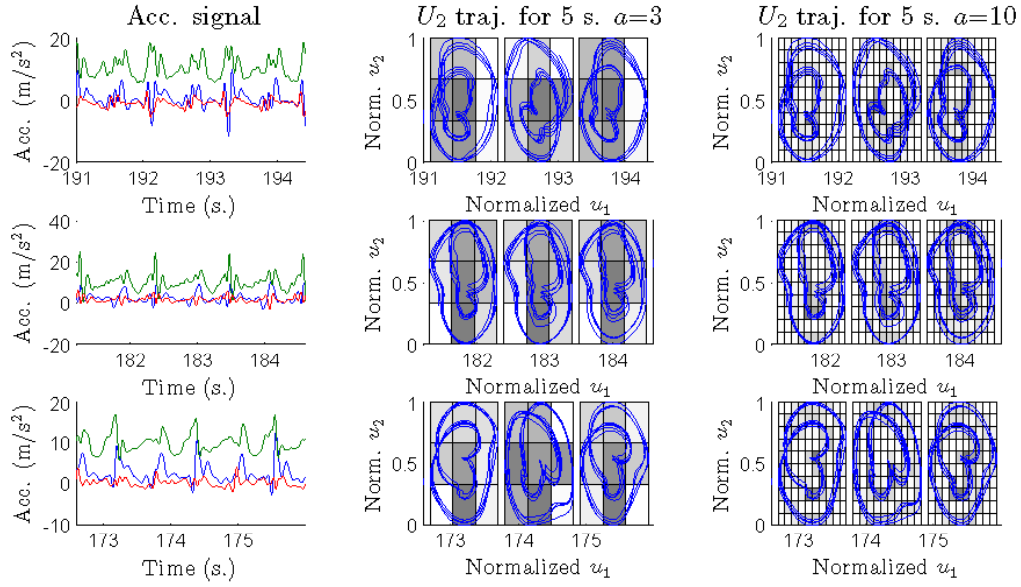


Figure 9.1: Accelerometer signal obtained during gait for three users and its reconstructed trajectories. Abstraction levels 3 and 10 are shown. Parameters used are $w=5$, $\tau=5$ and $m=10$

$$d_{i_1, \dots, i_k} = \frac{|l(\mathbf{U}'_{\mathbf{k}}, i_1, \dots, i_k)|}{N - (m-1) \cdot \tau} \quad (9.4)$$

where:

$$l(\mathbf{U}'_{\mathbf{k}}, i_1, \dots, i_k) = \left\{ \mathbf{w}' = (w^1, \dots, w^k) \text{ row of } \mathbf{U}'_{\mathbf{k}} / w^j \in \left[\frac{i_j-1}{a}, \frac{i_j}{a} \right] \right\} \quad (9.5)$$

and $|\cdot|$ is the cardinality of a set.

Figure 9.1 shows abstraction levels 5 and 10 of a 2- d filtered trajectory ($k=2$) applied to three accelerometer signals obtained during walking. Trajectories from the same user have similar density values at different abstraction levels although changing a value may maximize differences between users. Note that a point-by-point comparison would be obtained for $a = \infty$.

9.2.2 Experiments

The acceleration magnitude is again used as the scalar measure to reconstruct the state space. Experiments will test different values for embedding dimension m , time-lag τ and window size w . Signals used correspond to twelve healthy volunteers walking 20 m at normal speed twice, as in the previous section.

Experiments will test different values for embedding dimension m , time-lag τ and window size w . Patterns obtained through the first walk are used to train a SVM with a RBF-kernel through a 10-fold Cross Validation. Second-walk patterns are used to obtain the accuracies reported. Filtered trajectories are classified by using a normalized version of patterns defined in equation 9.4. Patterns used are $\mathbf{d}' = \frac{d_{i_1, \dots, i_k}}{\max(\mathbf{d})}$ to allow the different density probabilities maximize its differences.

Window sizes tested are $w = \{1, 3, 5, 10\}$, embedding dimensions are $m = \{3, 5, 10, 20, 30\}$, time-lag values $\tau = \{1, 3, 5, 7, 9\}$ and abstraction levels $a = \{1, \dots, 30\}$. Dimension k has been fixed to 2 since it allows us to visually evaluate the results.

9.2.3 Results and discussion

Figure 9.2 shows the gait recognition results. It can be seen that a window size greater than 1 second is necessary to have an accuracy higher than 70%. Since a gait cycle, i.e. two steps, takes approximately 1 second, it suggests that several cycles are necessary to have a good accuracy. Regarding the abstract level, small values do not allow good gait recognition accuracies, but higher values from 5 to 10 provide the best accuracies.

Figure 9.3 shows the effect of the abstract level in the test accuracy for few parameter values. It also includes the best accuracy obtained (87%), which corresponds to $w = 5$, $m = 5$ and $a = 10$. Abstraction levels above a certain value, that according to figure 9.3 results could be fixed in $a=10$, make the gait recognition accuracy decrease. Figure 9.4 illustrates the main problem behind high values of a . It can be observed two consecutive filtered trajectories for $a = 10$ and $a = 30$. For $a = 10$ both trajectories have the same density values. However, for $a = 30$ some differences appear which make the accuracy on gait recognition decrease. This fact suggests that

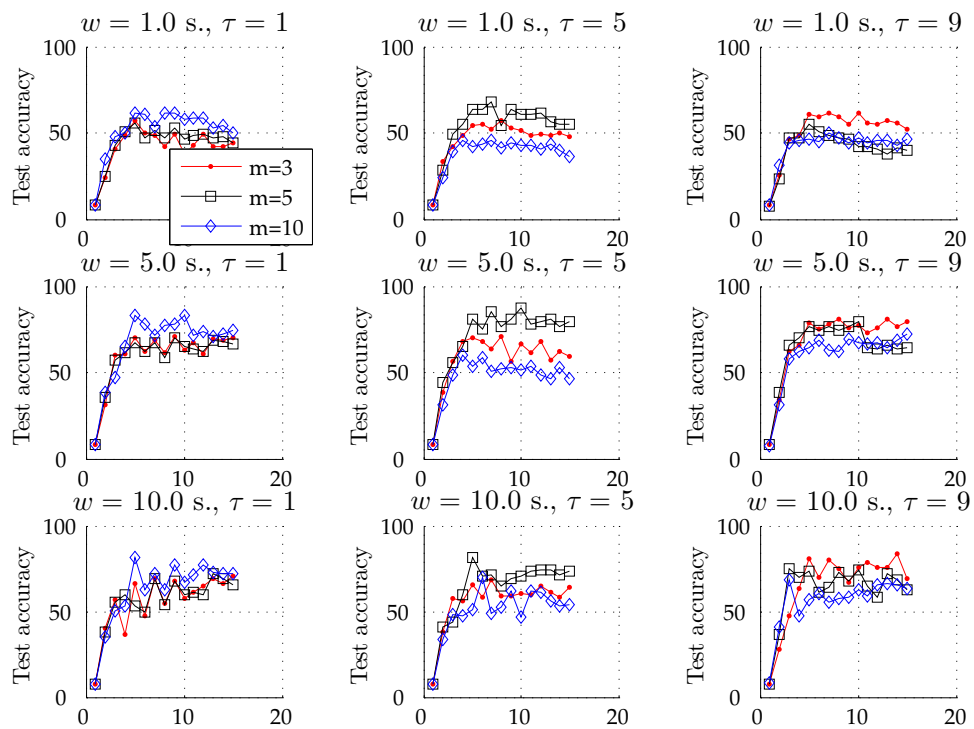


Figure 9.2: Gait recognition results for different values of w , a and τ . $w = 5$, $m = 5$ and $a = 10$ provide the best results.

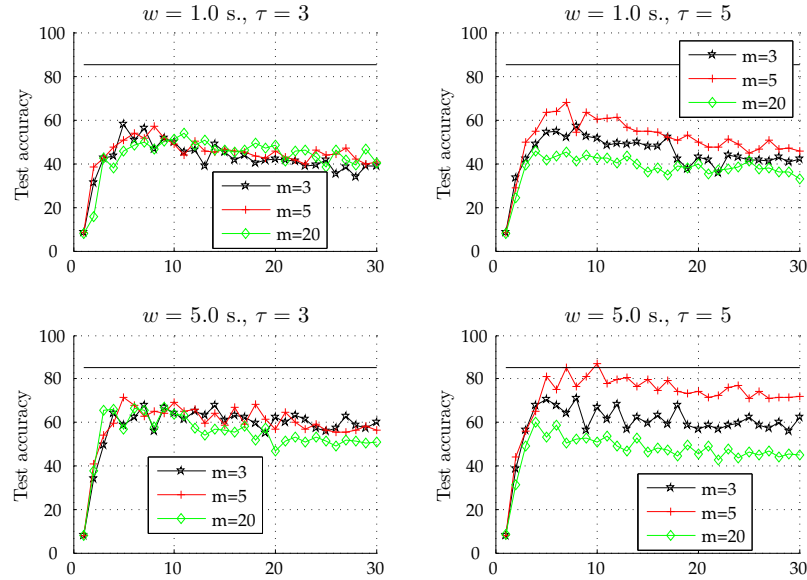


Figure 9.3: Abstraction level effect for different values of w and τ . Increasing the abstraction level increases the accuracy until a certain level, after which the accuracy does not increase.

in the extreme, for $a = +\infty$, which corresponds to a element-by-element comparison of the trajectories, the accuracy would extremely decrease. Thus, the granular approach proposed considers reconstructed trajectories in such an abstraction levels that provide good recognition capabilities for low a values. It is remarkable that the results obtained have only considered $k=2$ dimensions.

Abstract level largely impact the computational time. Patterns that are classified have a dimensionality of a^k . Thus, an abstract level of $a = 30$ provides a highly dimensional dataset. In order to save computational time and provide real-time detection, low values for a are desired. According to the results, $a = 5$ could be used instead of $a = 10$ in order to save computational costs by slightly decreasing the accuracy obtained.

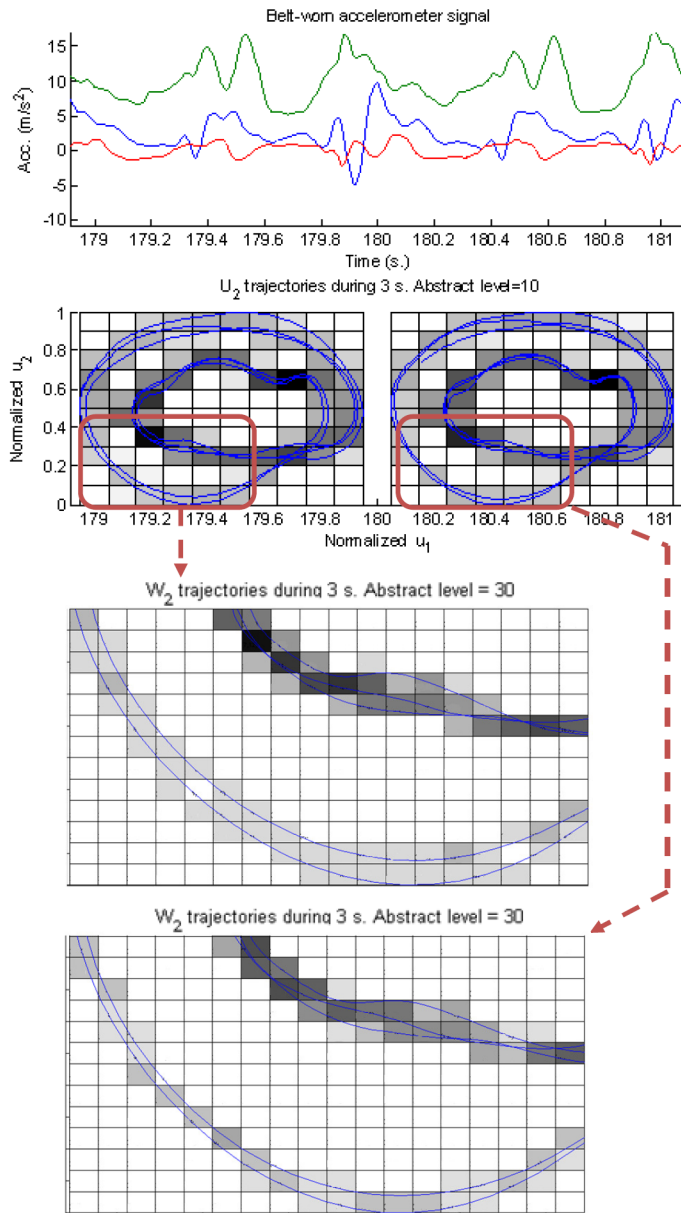


Figure 9.4: Two consecutive reconstructed trajectories for $a = 10$ and $a = 30$. No differences appear for $a = 10$ while for $a = 30$ some distinctions can be appreciated

9.3 Box Approximation Geometry for gait recognition

This section presents a novel gait recognition that employs a few singular values obtained by means of Singular Spectrum Analysis applied to scalar measurements from the inertial sensor. Singular values can be interpreted as the approximate edge length of the bounding box wrapping the attractor in the latent space. Effects of different parameters on the gait recognition performance using patterns from 20 different subjects are analysed.

Most of the previous studies on inertial-based gait recognition use a gait cycle model to represent the identification characteristic of every subject [40] [70] [33] [107] [39]. A disadvantage of these methods is that they significantly depend on existing gait cycle detection methods, which have been proved to show poor performance in ambulatory conditions [71]. An interesting alternative which uses techniques from nonlinear time series analysis [58] is proposed by Frank et al. [37]. This method consists of considering the human gait as the result of a complex dynamical system influenced by muscles, cardiovascular system, skeleton and even by the frame of mind. The sensor measurements are treated as a time series that enable an attractor reconstruction of the dynamical system. In [37] the recognition is performed using an algorithm named *Geometric Template Matching* that assesses how well the reconstructed time series obtained from a segment of an acceleration signal fit a particular model. Thus, this method requires to save the observed trajectory in the reconstructed state space in order to be able to be compared to new trajectories.

The method presented in this section also considers the dynamical system approach. However, it will be shown that identification tasks can be performed by taking into account only a reduced number of features obtained from the reconstructed attractor of a short segment of gait signals. The features considered are associated with the geometry of the *box* that contains a filtered version of the reconstructed attractor, i.e. the sizes of the hyperrectangle edges where the filtered attractor is embedded, which are the so-called *Box Approximation Geometry*. These features represent the amplitude of the reconstructed trajectory in each of the dimensions so,

consequently, trajectories are not required to be saved as in [37]. The hyperrectangle edges measurements are related to the time series spectrum obtained through Singular Spectrum Analysis (SSA) methodology. SSA has been shown to be a very useful tool for noise-reduction and oscillatory components identification [49]. This reduction is achieved by selecting the eigenvalues and the associated principal components from the lagged-covariance matrix of the time series data [54]. In this work, it is shown that not only the components associated to small eigenvalues can be rejected, but also one component of each consecutive pair can be discarded due to the redundancy caused by strong correlation between components in phase-quadrature.

9.3.1 Box Approximation Geometry

According to [116], the statistical dimension can be estimated by counting the eigenvalues above the noise floor after the trajectory matrix decomposition into orthogonal components. However, eigenvalues contain more information since they give an idea of the reconstructed attractor shape. Specifically, they represent a measure of the bounding box enclosing the attractor. The approach proposed in this work consists of characterising the dynamical system, i.e. identifying the personal gait, through what is termed as *Box Approximation Geometry* (BAG) of the attractor in the latent space obtained through the same trajectory matrix decomposition that SSA performs. The term Box Approximation Geometry is used to refer to the k -dimensional hyperrectangle containing most of the filtered trajectory in latent space. The dimension sizes of this hyperrectangle are closely related to the k -larger eigenvalues. This leads to define $BAG(k)$ as the vector:

$$BAG(k) = (\lambda_1, \dots, \lambda_k) \tag{9.6}$$

$BAG(k)$ represents the quantity of information among the attractor dimensions. We will show later that taking into account only a subset of $BAG(k)$ is enough to accurately identify a person.

Another algorithm that uses the filtered reconstructed attractor for activity recognition and gait identification is *Geometric Template Matching* (GTM) [37]. This algorithm assesses how well a given time series projected onto the eigenvectors, obtained with other training examples, fit a particular model. Short segments of the projected data are then compared geometrically against their nearest neighbours and they are assigned to the model that obtains the highest average score. Thus, several elements of the latent space are compared against others, which is avoided when the geometry of the attractor is employed since only the amplitude of the trajectories are needed.

It has been shown that the projection of a quasi-periodic time series into the eigenvectors of the time-delayed matrix provides pairs of principal components which are in an approximated phase quadrature [42]. Vautard and Ghil [116] argued that, subject to certain statistical significance tests, such pairs correspond to the nonlinear counterpart of a sine-cosine pair in the standard Fourier analysis of linear problems. However, the advantage over sines and cosines is that the components obtained from SSA are not necessarily harmonic functions and they can capture highly anharmonic oscillation shapes. Considering this, BAG method is designed to use only the eigenvalues corresponding to odd principal components so that redundant information is avoided.

9.3.2 Gait identification by means of Box Approximation Geometry

Accelerometers are the most common inertial sensors used in movement analysis. They can be used to measure vibration on cars, machines, building, process control systems and safety installations. Nowadays, they are integrated on laptops, tablets, digital cameras and mobile phones providing a new opportunity to protect these devices from unauthorised usage. Triaxial accelerometers sense linear acceleration along three orthogonal directions.

The magnitude of the acceleration is used again as the scalar measure to reconstruct the state space. Figure 9.5 shows the typical walking acceleration measures obtained by an accelerometer located on the waist during stance and swing phases.

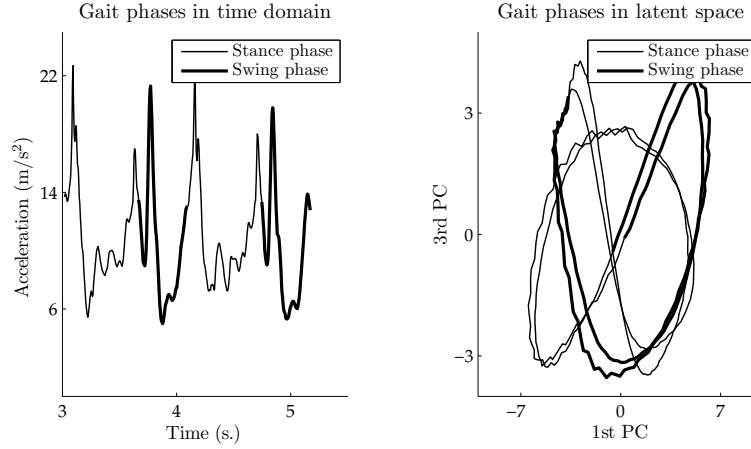


Figure 9.5: Acceleration magnitude measurements and their reconstruction into 2-latent space by using the first and third dimensions, $m=20$ and a sampling frequency of 200 Hz.

Gait phases have been estimated by the lateral acceleration obtained from the waist according to [126]. Figure 9.5 also shows the reconstruction obtained by using $\tau = 5$, $m = 10$ and a sampling frequency of 200 Hz. Stance and swing phases are shown to be different parts of the trajectory and they clearly form a complete gait cycle. Since a correctly reconstructed trajectory in the phase space cannot intersect itself along an orbit, it turns out that the reconstruction observed needs more than just two latent variables in order to be valid, so the statistical dimension is greater than two.

As mentioned before, components associated to consecutive eigenvalues in the spectrum are approximately in phase quadrature. This behaviour is illustrated in Figure 9.6, where the first twelve principal components are represented in pairs (1-2, 3-4, ..., 11-12). This behaviour confirms that considering only one component from each consecutive pair enables us to avoid redundant information. Consequently, only odd eigenvalues are taken into account, improving the efficiency of the method. Thus, a new *BAG* denoted as $BAG^*(k)$ is defined in the context of gait identification as follows:

$$BAG^*(k) = (\lambda_1, \lambda_3, \dots, \lambda_{\lceil \frac{k}{2} \rceil}) \quad (9.7)$$

The characterisation of the human gait is dependent on the time lag τ , the number

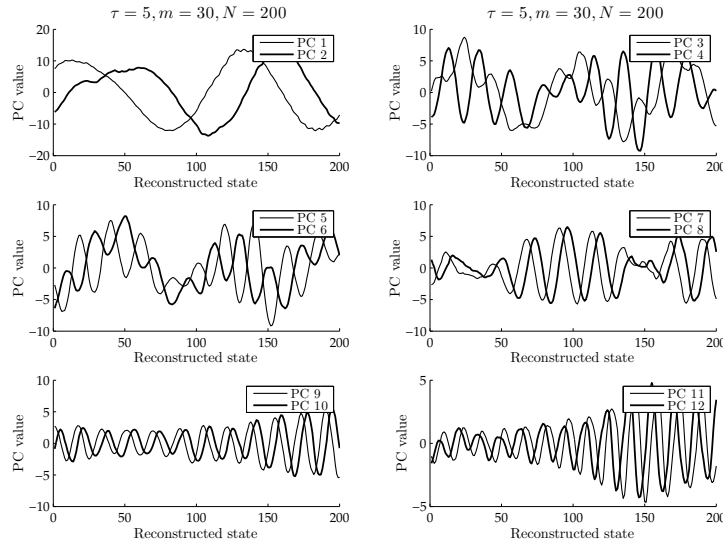


Figure 9.6: Latent variables 1 to 12 in approximately phase quadrature for parameters $\tau = 5$, $m = 30$ and $N = 200$.

of measures used in the delay coordinates m , and the number of latent variables employed, which should be lower than its statistical dimension. For instance, the time lag effect is shown in Figures 9.7 and 9.8 by observing the eigenvalues for three different time lags in two individuals. These figures have been obtained by using $m = 10$ and a sampling frequency of 200 Hz.

9.3.3 Experiments

CETpD sensor device accelerometer was sampled at 200Hz. Twenty volunteers participated in this study. Each volunteer walked at a normal speed twice and the first walk was used for training the gait recognition algorithm and the second walk for testing.

Different values for the parameters that affect the characterisation method have been tested: the time-lag; the embedding dimension; and the feature dimension vector k , i.e. the number of the hyperrectangle edges considered in $BAG^*(k)$ defined by Eq. (9.7). Another important tested parameter that affects gait identification is the window length measured in seconds into which the reconstruction is performed,

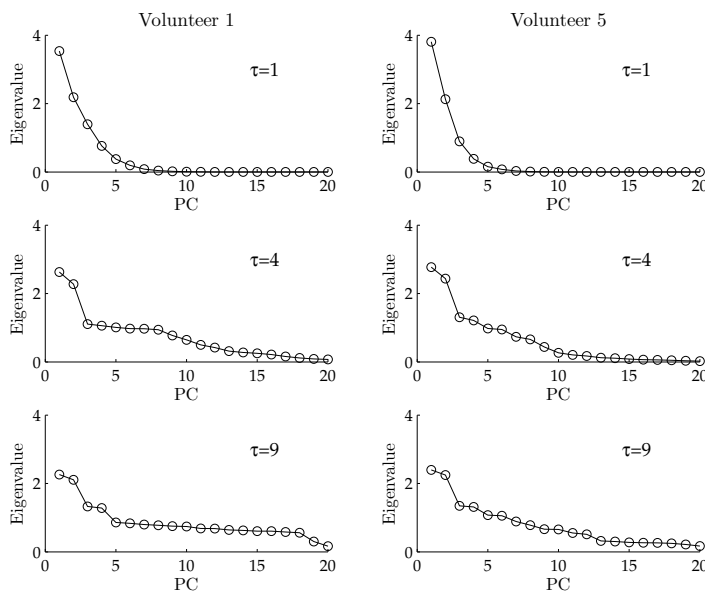


Figure 9.7: Scree plots obtained from the accelerometer signals of 2 different volunteers using 3 different time lag values.

henceforth denoted by w , where $N = w \cdot F_s$ and F_s is the sampling frequency. It is expected that larger windows provide more accurate gait recognitions since a larger state space trajectory enables better reconstruction. Optimal values of τ and m are commonly estimated according to Average Mutual Information and False Nearest Neighbours. In this work, they have been calculated to establish whether the optimal values provide the best recognition rate or not.

9.3.4 Results

BAG patterns, which are composed of eigenvalues according to Eq. (9.7), were extracted by means of overlapped windows of w seconds obtained from the accelerometer. Every $w/2$ seconds a new window was started. The number of patterns obtained depends on w , τ and m (from 25 to 52 patterns per each individual). Thus, a classification problem of 20 classes and a maximum of 1025 patterns has been obtained. BAG values from the first walk have been used to train a Support Vector Machine (SVM) with a Gaussian kernel. The resulting SVM was tested using values obtained

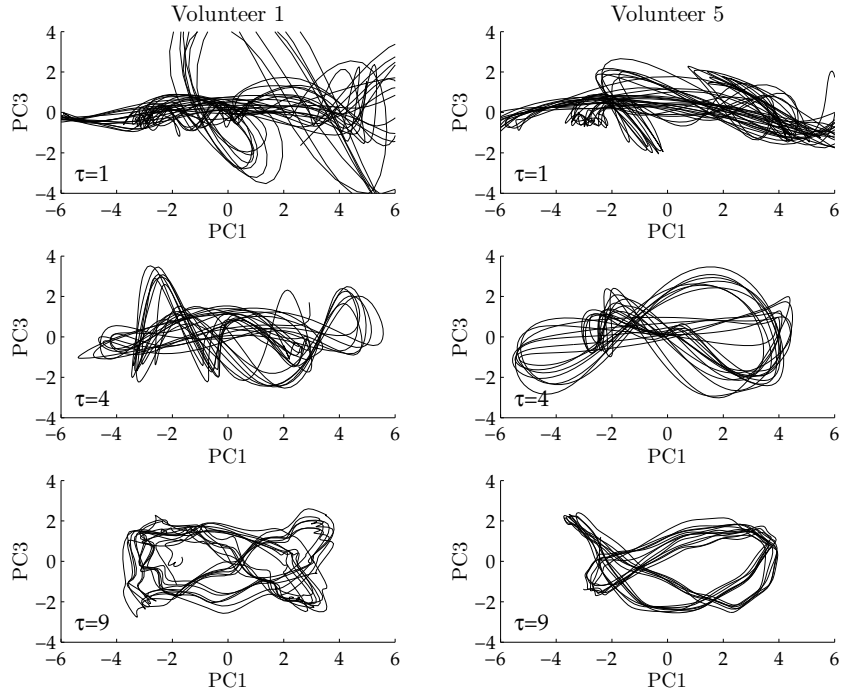


Figure 9.8: Reconstruction of the trajectory in the first and third dimensions of the latent space from the accelerometer signals of 2 different volunteers ($m = 20$; $N = 2000$, $N = 500$ and $N = 222$ for $\tau = 1$, $\tau = 4$ and $\tau = 9$, respectively).

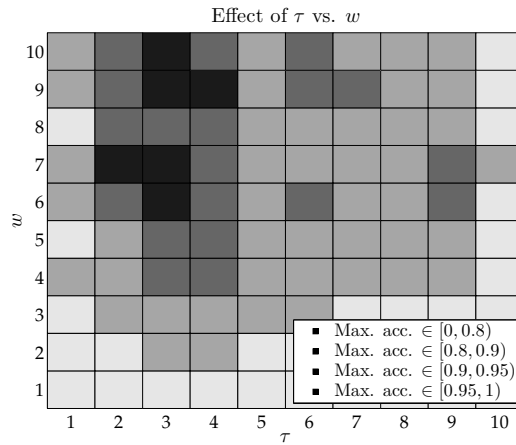
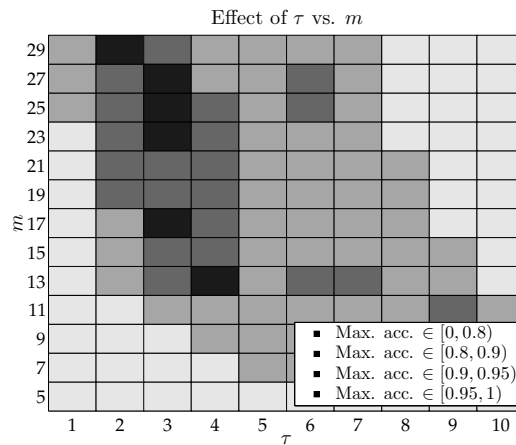
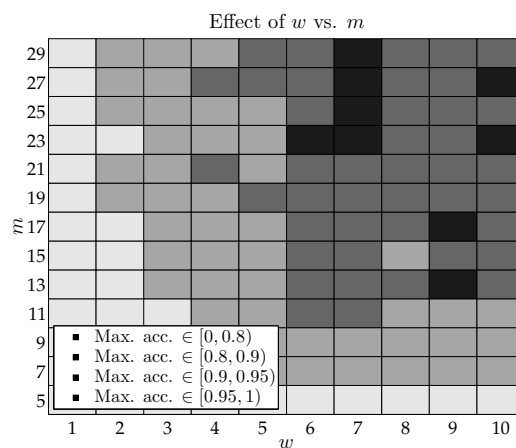


Figure 9.9: Effect of τ and w in the gait recognition accuracy in 20 volunteers

Figure 9.10: Effect of τ and m in the gait recognition accuracy in 20 volunteersFigure 9.11: Effect of w and m in the gait recognition accuracy in 20 volunteers

from the second 20 meters walk.

The parameters that affect the method have been tested with several values. More concretely, w and τ tested values are $1, \dots, 10$. Regarding the other parameters, $m = 5, 7, 9, \dots, 29$ and $k = 1, 3, 5, \dots, 29$ were used, since only odd latent variables are considered. The number of features used is $(k + 1)/2$.

The maximum accuracy obtained fixing the value of a pair of parameters among $(\tau, m, w$ and $k)$ and testing the previously mentioned values for the rest of them are shown in Figure 9.9, 9.10, 9.11 and 9.12. The highest accuracy is 96.4% with $\tau=3$, $w=7$, $m=23$ and $k=23$. Figure 9.13 shows the impact of increasing k in the

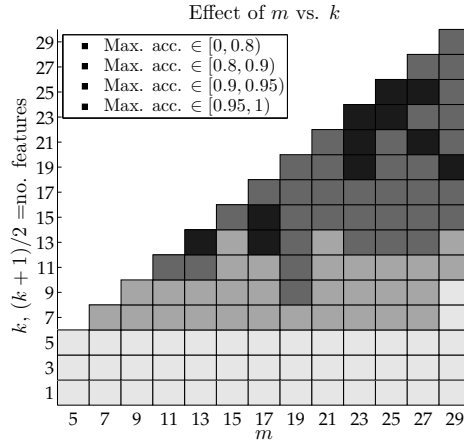


Figure 9.12: Effect of k and m in the gait recognition accuracy in 20 volunteers

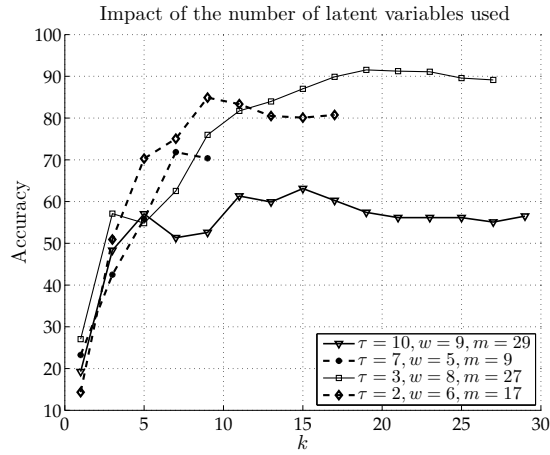


Figure 9.13: Effect of k in the gait recognition accuracy for 4 combination of parameters

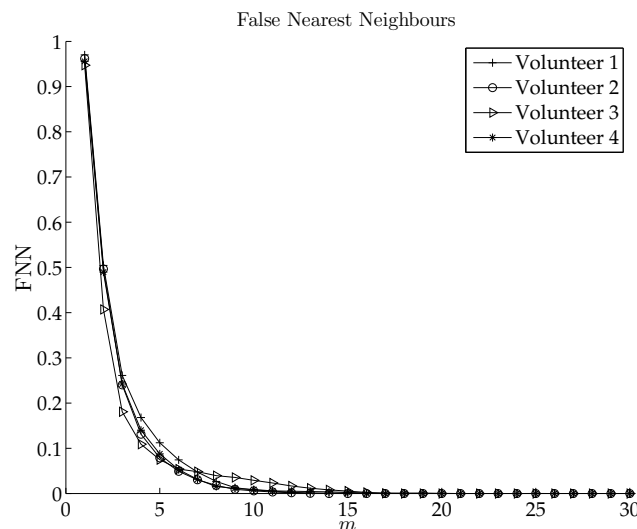


Figure 9.14: False Nearest Neighbours results for different embedding dimensions. Y axis represents the proportion of False Nearest Neighbours.

accuracy of four given examples. The False Nearest Neighbours and the Average Mutual Information results are shown in Figure 9.14 and 9.15, respectively.

9.3.5 Discussion

A gait cycle, i.e. two steps, usually takes more than one second at normal speed, so a window length of 1 second does not enable the whole trajectory to be unfolded. This could explain that accuracies obtained for $w = 1$ s. do not exceed 80%, as Figure 9.9 presents. Larger window lengths reach accuracies over 90%. On the other hand, gait recognition performance is also affected by the embedding dimension m , as Figure 9.10 shows. For instance, $m = 5$ leads to accuracies below 80%. This is an expected result according to the False Nearest Neighbours scheme since, as Figure 9.14 shows, the attractor is not yet unfolded when an embedded dimension below 10 is used. Regarding time lag τ , good results can be obtained by any value as long as it is lower than 10, i.e. samples must be taken at a frequency sampling greater than 20 Hz.

The number of the BAG hyperrectangle edges used greatly influences the gait recognition performance. Although the maximum accuracy is obtained by means of 12 features, accuracies over 90% can be obtained with only 5 features ($k=9$) and over

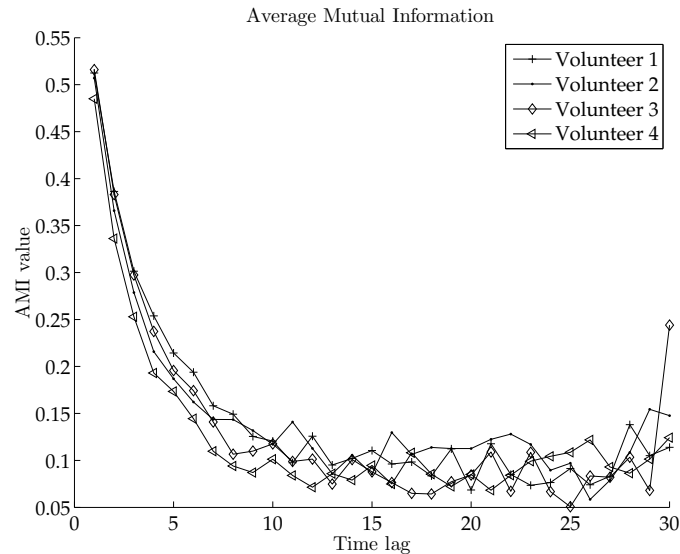


Figure 9.15: Average Mutual Information results for four different training signals. Best recognition rates are provided by $\tau = 2, 3$ and 4.

95% with 7 features ($k = 13$). In Figure 9.12 and 9.13 it is clearly shown that if the number of edges used is increased, the performance also augments. This is obviously explained because the more dimensions used, the more information from the attractor is considered and higher accuracies are obtained. However, according to Figure 9.13, there is a value of k after which the performance does not significantly increase since the last variables only correspond to noise [42] [49].

The result of the Average Mutual Information applied to different time lags is shown in Figure 9.15. This function measures the dependence between a signal and its corresponding τ -lagged delayed signal. Thus, $\tau=1$, provides the highest dependence of a sample and the following one, which is obtained after 0.005 ms. Time lag values greater than 10 provide similar results. This value is the optimal one according to the Average Mutual Information criterion. However, the best gait recognition accuracies are obtained for lower time lag values ($\tau=2, 3$ and 4) according to Figure 9.9. Thus, the optimal τ estimated through AMI does not provide in this case the best accuracy, confirming that there is no any universal approach to select the optimal τ and that the optimal choice depends on the application [58].

In terms of accuracy, the results obtained are similar to those provided by other

methods. For instance, in [39] an accuracy of 86% was reported and 95% was achieved in [40]. However, these methods use more features than the 5 used by BAG and, in addition, they depend on existing gait cycle detection methods, which have been proved to show poor performance in ambulatory conditions [71]. In [37] a similar recognition rate is achieved without detecting gait cycles. However, in that work, successive elements in the latent space that compose a trajectory were used as features, instead of just the rough shape of the attractor as the method introduced in this paper does. Using BAG, less and easier interpretable features are employed in the recognition task.

In this work, a specifically developed device was used while in [107] [39] [37] a mobile phone was employed instead. Using a specifically developed device may reduce the difficulty of our task since a stable frequency sampling is obtained. Regarding the sensor position, the belt-worn location corresponding to this work is the same than in [37] but different to the one used in other works. For instance, in [39] the sensor was in a trouser pocket and in [40] it was located on the leg.

In summary, results show that identifying a person through the dynamical system characterisation approach proposed is possible by using a standard SVM-classifier with an overall accuracy of 95%. The most suitable time lag values are from 2 to 4 when a sampling frequency of 200 Hz is employed, which are below the optimal value considered by Average Mutual Information. The necessary reconstruction dimension to achieve these results is at least 10, although higher values may be used, and three or more seconds are needed as window length. The quantity of eigenvalues necessary to achieve 90% accuracy, and therefore the number of features, is just 5.

9.4 Conclusions

In this section, three different approaches for gait recognition tasks have been presented. The first method employs the temporal and frequency domains to recognize people from gait. The second one consists in a qualitative approach to granularly represent the gait reconstructed attractors, which employs Cell-to-cell mapping techniques. Finally, the third approach is a new gait recognition method called ‘Box

Approximation Geometry' that uses Singular Spectrum Analysis to filter the gait reconstructed attractor.

The first method was tested against 5 volunteers with an accuracy of 82%. The granular approach was validated in 12 healthy adults with an accuracy up to 87%. Finally, the Box Approximation Geometry approach was tested in 20 volunteers showing an accuracy greater than 95%. Thus, reconstructing gait's attractor has provided the best results on gait recognition and, moreover, Box Approximation Geometry approach outperforms the other methods using, at the same time, the minimum number of features.

Part III

Final remarks

Chapter 10

Conclusions and future work

The aim of this thesis is to contribute to the development of new methods for analyzing accelerometer signals in the field of human motion analysis. As it may be deduced from the structure of the memory, contributions are twofold. First, new methods have been introduced and new results have been obtained for dyskinesia detection and motor state (ON/OFF) identification in PD. Second, new techniques have been proposed for gait parameters identification based on state space reconstruction. These new techniques have also been applied as a biometric measure to recognize users from accelerometer signals.

10.1 Dyskinesia detection and motor state identification in Parkinson's disease

Parkinson's disease is a chronic neurodegenerative disease that affects motor control and movement. Presently, there is not any cure for this disease, so current treatments mainly consist in counteracting motor symptoms. Patients typically experience a good response to the medication in the early stages of the disease. However, as the disease advances, the effect of medication begins to wear off in a few hours after each dose, leaving patients anticipating the need for their next dose. In a moderate stage of the disease, patients fluctuate between ON periods, during which they enjoy a good

response to medication, and OFF periods, during which they experience symptoms of the underlying parkinsonism such as bradykinesia, freezing of gait or dyskinesias. Motor state is a denomination that reflects both ON and OFF periods. Dyskinesias are not a symptom related to state OFF since this symptom usually appears during ON periods. This symptom is mainly related to peak doses of the blood medication level and to transitions from ON to OFF and from OFF to ON periods.

The need for medication in PD patients greatly varies from patient to patient and depends on the progression of the disease, the severity of symptoms and the activities undertaken by each patient during the day. Until now, the only way a doctor may adjust the exact dose of medication is the observation of the disease evolution during the periodic consultations and the use of questionnaires that the patient or some relative fills. Automatic ambulatory detection of motor fluctuations is an important tool that enable doctors to accurately adjust the medication to the circumstances of each patient and, thus, minimize the OFF periods and periods of over-medication, in which usually dyskinesias appear.

To this end, in this thesis a detection algorithm of dyskinesias has been developed as well as a motor state detection algorithm based on a single accelerometer placed on the patient's waist. Dyskinesia detection algorithm analyzes the signal spectrum and allows us to detect the presence of dyskinesias except when the patient walks. The reason behind is that gait movements overlap the involuntary movements associated to dyskinesias. Instead, motor states detection is performed while the patient walks. The algorithm analyzes the spectral power of strides, which has been shown to reflect the motor fluctuations in patients who suffer bradykinesia during OFF states.

The validation of the algorithms in their offline versions was made by means of the signals recorded in 15 patients performing daily life activities in their own environment and for several hours. The results of this validation have amply shown the ability of these algorithms to be used as the tool that will allow neuroscientists to improve the therapeutic regimen of patients.

Detecting dyskinesias and motor states in real time, as discussed in the introduction, would allow us to use an apomorphine pump to automatically control the medication of the patient. This idea has already been assayed in the clinical study

of the HELP project, in which the automatic administration of apomorphine in PD patients has been tested for the first time. Consequently, online versions of the algorithms have been developed and implemented using the same sensor. Results show that:

- The algorithm of dyskinesia detection has shown a good performance detecting the involuntary movements during the study. However, a few daily life activities resulted in false positives in the detection.
- The motor states works correctly in patients suffering bradykinesia only during OFF states.

10.2 Accelerometer-based gait analysis

The second focus of this thesis is the description of gait by means of techniques from the field of dynamical systems. Specifically, the approach used exploits the attractor reconstruction for characterizing gait. In this thesis, the orbits described by the dynamical system have been characterized for two different purposes. First, two gait parameters, velocity and stride length, were extracted from the reconstructed attractors. Secondly, the reconstructed attractors have been considered as gait signature and it was used for person identification.

Gait parameters extraction has been performed in PD patients since these parameters are indicative of the motor state of the patients. To this end, two methods have been developed. The first performs stride detection from the event: initial contact of gait cycle and then, each stride is characterized by means some statistics that have been used to estimate gait parameters using a nonlinear regression. The second method employs the latent variables obtained from Singular Spectrum Analysis to estimate gait parameters also using a nonlinear regression. Results have shown that the use of techniques from the field of dynamic systems provides more accuracy than characterizing gait by means of statistics obtained from the raw times series.

People identification through gait analysis was performed using three different methods:

- First, a method based on the existing literature on gait recognition has been implemented. This method, which uses the time domain, has the disadvantage of requiring the detection of gait cycles, which has shown to provide poor precision in ambulatory environments.
- Second, a new method that does not require detecting gait cycles and uses techniques from the field of dynamical systems has been developed. Specifically, this method performs a granular approach based on cell-to-cell mapping to characterize the subjects' gait.
- Finally, we have developed an alternative method that also uses the reconstruction of the state space but, in this case, employs the values of the latent variables to identify each person while walking. This method has shown excellent results in the task of identifying 20 different people.

10.3 Future Work

Within the scope of the PD, it is noteworthy that the experience gained in the development of the algorithms presented is used as a basis for further research on how to generate a map of the symptoms experienced by PD patients through an accelerometer placed on the waist. The work on this thesis project continues in the REMPARK European project in which a larger database formed by motion signals obtained from 90 PD patients from four different countries (Ireland, Italy, Spain and Israel) is collected [94]. Thus, it is expected to evolve and enhance the methods presented in this thesis in order to improve the quality of life of PD patients:

- For the motor states detection algorithm it is important to emphasize that the algorithm is still not effective during the OFF state for those patients who did not show bradykinesia, or showed it simultaneously to FoG. In the latter case, it was observed that freezing episodes limited the amount of strides that the patient performed, so that the detection algorithm could not identify the motor state of the patient. Consequently, a possible improvement of the motor

state detector developed in this thesis is to combine it with a FoG detector, as outlined in the REMPARK project.

- The dyskinesia detection algorithm, however, showed that it could produce false positives in the detection while performing certain activities of daily life. These activities should be incorporated in the dyskinesia detection analysis to avoid confusion with the symptom. This line of work is covered in REMPARK project.

Regarding the gait analysis methods developed in the state space reconstruction, there are several paths that can be explored in the future:

- First, it could be convenient to explore new reconstructed attractor filters different from SSA, since this method provides a linear subspace of the reconstructed space. For instance, decomposition methods such as Diffusion Maps or Local Linear Embedding could bring a dimensionality reduction of the attractor maintaining the attractor locality while taking into account the nonlinearity.
- Secondly, techniques based on state space reconstruction can be extended to other tasks such as characterizing motor patterns in PD patients. This approach can also be applied to the detection and characterization of other symptoms such as freezing of gait, which has been proven to be a very difficult task from a single sensor located at the waist [16].
- Finally, the methods obtained for the extraction of gait parameters are not only applicable to PD but they could be also useful in other diseases as, for instance, diabetes. Diabetic patients may suffer disfunctions in the vestibular system because of the high levels of blood glucose and may develop diabetic neuropathies, which is a damage in the nerves [8] [26]. The consequences of these dysfunctions are that patients suffer gait alterations, which produce a higher risk of suffering injuries such as fractures, falls and sprained ankles [82]. The monitorization of gait parameters in diabetic patients, such as gait speed and stride length, could allow clinics to prevent neuropathies and reduce the risk of injuries for their patients [8].

10.4 Publications

First, publications in the PD field are:

- Dyskinesia and motor state detection in PD patients with a single movement sensor. **Samà, A.**; Perez-Lopez, C.; Romagosa, J.; Rodriguez-Martin, D.; Catala, A.; Cabestany, J.; Perez-Martinez, D.A. and Rodriguez-Molinero, A. Engineering in Medicine and Biology Society, EMBC, 2012 Annual International Conference of the IEEE [99]
- HELP: Optimizing Treatment of Parkinson's Disease Patients. Ahlrichs, C.; Herrlich, S.; **Samà, A.**; Rodríguez-Molinero, A.; and Rovira, J. 3rd International Conference on Elderly and New Technologies (2012) [5]
- Detection of Gait Parameters, Bradykinesia and Falls in Patients with Parkinson's Disease by Using a Unique Triaxial Accelerometer. Rodríguez-Molinero, A.; Pérez-Martínez, D.; **Samà, A.**; Sanz, P.; Gálvez, M.; Pérez-López, C.; Romagosa, J.; Català, A. World Parkinson Congress, 28th September to 2nd de October, Glasgow (2010) [7].

Second, publications in the field of gait analysis by means of accelerometers are provided:

- Gait identification by means of box approximation geometry of reconstructed attractors in latent space **Samà, A.**; Ruiz, F.J.; Agell, N.; Pérez-López, C.; Català, A.; Cabestany, J. Neurocomputing (2013) [96].
- Analyzing human gait and posture by combining feature selection and kernel methods. **Samà, A.**; Angulo, C.; Pardo-Ayala, D.E.; Català A.; and Cabestany, J. Neurocomputing (2011) [91]
- Time Series Analysis of inertial-body signals for the extraction of dynamic properties from human gait. **Samà, A.**; Pardo-Ayala, D.E.; Cabestany, J.; and Rodríguez-Molinero, A. Neural Networks (IJCNN), The 2010 International Joint Conference on. [93]

- Extracting Gait Spatiotemporal Properties from Parkinson's Disease Patients. **Samà, A.**; Catala, A.; Rodríguez-Molinero, A.; and Angulo, C. Neural Information Processing Systems, Mini Symposia on Assistive Machine Learning for People with Disabilities (2009) [92]
- Gait recognition by using Spectrum Analysis on state space reconstruction. **Samà, A.**; Ruiz, F.J.; Pérez-López, C.; and Català, Andreu. Artificial Intelligence Research and Development: Proceedings of the 14th International Conference of the Catalan Association for Artificial Intelligence [98]
- Gait identification by using spectrum analysis on state space reconstruction. **Samà, A.**; Ruiz, F.J.; Pérez-López, C.; and Català, A. Lecture Notes in Computer Science, International Work-Conference on Artificial Neural Networks (2011) [97]
- Granular Singular Spectrum Analysis for Gait Recognition. **Samà, A.**; Ruiz, F.J.; and Agell, N. 26th International Workshop on Qualitative Reasoning (2012), Los Angeles, California (USA) [95]

Bibliography

- [1] *La prise en charge de votre maladie de Parkinson*. Haute Autorité de Santé, France, 2007.
- [2] *A focus on Parkinson's disease and the European Society*. 1st European Brain Policy Forum. Brussels, Belgium, 2008.
- [3] Henry D. I. Abarbanel and Matthew B. Kennel. Local false nearest neighbors and dynamical dimensions from observed chaotic data. *Phys. Rev. E*, 47:3057–3068, May 1993.
- [4] Alison Abbott. Levodopa: the story so far. *Nature*, 466(7310):S6–S7, August 2010.
- [5] Claas Ahlrichs, Albert Sama, Jordi Rovira Simon, Simon Herrlich, and Alejandro Rodraguez-Molinero. Help: Optimizing treatment of parkinson's disease patients. In *3rd International Conference on the Elderly and New Technologies*, Castellan, Spain, April 2012.
- [6] S.K. Alam, J. Mamou, E.J. Feleppa, A. Kalisz, and S. Ramachandran. Comparison of template-matching and singular-spectrum-analysis methods for imaging implanted brachytherapy seeds. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 58(11):2484 –2491, 2011.
- [7] Albert Samà Pilar Sanz Matilde Gálvez Carlos Pérez-López-Jaume Romagosa Andreu Català Alejandro Rodríguez-Molinero, David Pérez-Martínez. Detection of gait parameters, bradykinesia and falls in patients with parkinson's

- disease by using a unique triaxial accelerometer. In *World Parkinson Congress, 28th September to 2nd de October, Glasgow*, September 2010.
- [8] Lara Allet, Stphane Armand, Rob A. de Bie, Zoltan Pataky, Kamiar Aminian, Francois R. Herrmann, and Eling D. de Bruin. Gait alterations of diabetic patients while walking on different surfaces. *Gait and Posture*, 29(3):488 – 493, 2009.
- [9] D. Alvarez, R.C. Gonzalez, A. Lopez, and J.C. Alvarez. Comparison of step length estimators from wearable accelerometer devices. In *Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE*, pages 5964 –5967, 30 2006-sept. 3 2006.
- [10] K. Aminian, B. Najafi, C. Bla, P.-F. Leyvraz, and Ph. Robert. Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *Journal of Biomechanics*, 35(5):689 – 699, 2002.
- [11] Kamiar Aminian and Bijan Najafi. Capturing human motion using body-fixed sensors: outdoor measurement and clinical applications. *Computer Animation and Virtual Worlds*, 15(2):79–94, 2004.
- [12] R. Antonello, R. Oboe, L. Prandi, and F. Biganzoli. Automatic mode matching in mems vibrating gyroscopes using extremum-seeking control. *Industrial Electronics, IEEE Transactions on*, 56(10):3880 –3891, oct. 2009.
- [13] Angelo Antonini, Pablo Martinez-Martin, Ray K. Chaudhuri, Marcelo Merello, Robert Hauser, Regina Katzenschlager, Per Odin, Mark Stacy, Fabrizio Stocchi, Werner Poewe, Oliver Rascol, Cristina Sampaio, Anette Schrag, Glenn T. Stebbins, and Christopher G. Goetz. Wearing-off scales in parkinson’s disease: Critique and recommendations. *Movement Disorders*, 26(12):2169–2175, 2011.
- [14] Erik K Antonsson and Robert W Mann. The frequency content of gait. *Journal of biomechanics*, 18(1):39–47, 1985.

- [15] Salarian Arash. *Ambulatory monitoring of motor functions in patients with Parkinson's disease using kinematic sensors*. PhD thesis, École polytechnique fédérale de Lausanne, 2006.
- [16] M Bachlin, Meir Plotnik, Daniel Roggen, Inbal Maidan, Jeffrey M Hausdorff, Nir Giladi, and G Troster. Wearable assistant for parkinsons disease patients with the freezing of gait symptom. *Information Technology in Biomedicine, IEEE Transactions on*, 14(2):436–446, 2010.
- [17] S. Bamberg, A.Y. Benbasat, D.M. Scarborough, D.E. Krebs, and J.A. Paradiso. Gait analysis using a shoe-integrated wireless sensor system. *Information Technology in Biomedicine, IEEE Transactions on*, 12(4):413–423, july 2008.
- [18] Alim Louis Benabid, Stephan Chabardes, John Mitrofanis, and Pierre Pollak. Deep brain stimulation of the subthalamic nucleus for the treatment of parkinson's disease. *The Lancet Neurology*, 8(1):67–81, 2009.
- [19] Roongroj Bhidayasiri, Daniel Tarsy, Roongroj Bhidayasiri, and Daniel Tarsy. Parkinson's disease: Levodopa-induced dyskinesia. In *Movement Disorders: A Video Atlas*, pages 10–11. 2012.
- [20] O. Blin, A.M. Ferrandez, and G. Serratrice. Quantitative analysis of gait in parkinson patients: increased variability of stride length. *Journal of the Neurological Sciences*, 98(1):91–97, 1990.
- [21] A.K. Bourke, J.V. OBrien, and G.M. Lyons. Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait and Posture*, 26(2):194–199, 2007.
- [22] Carlijn V Bouten, Klaas R Westerterp, Maarten Verduin, Jan D Janssen, et al. Assessment of energy expenditure for physical activity using a triaxial accelerometer. *Medicine and science in sports and exercise*, 26(12):1516–1523, 1994.

- [23] D. Bravi, M. M. Mouradian, J. W. Roberts, T. L. Davis, Y. H. Sohn, and T. N. Chase. Wearing-off fluctuations in parkinson's disease: Contribution of postsynaptic mechanisms. *Annals of Neurology*, 36(1):27–31, 1994.
- [24] Leo Breiman, Jerome H Friedman, Richard A Olshen, and Charles J Stone. Classification and regression trees. wadsworth & brooks. *Monterey, CA*, 1984.
- [25] Henk Broer, Floris Takens, Henk Broer, and Floris Takens. Reconstruction and time series analysis. In *Dynamical Systems and Chaos*, volume 172 of *Applied Mathematical Sciences*, pages 205–242. Springer New York, 2011.
- [26] P.R. Cavanagh, J.A. Derr, J.S. Ulbrecht, R.E. Maser, and T.J. Orchard. Problems with gait and posture in neuropathic patients with insulin-dependent diabetes mellitus. *Diabetic Medicine*, 9(5):469–474, 1992.
- [27] Fang SC Chao PK Wei JD. Chan HL, Wang CL. Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *Ann Biomed Eng*, 38(3):813 – 823, 2010.
- [28] Chih-Chung Chang and Chih-Jen Lin. Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011.
- [29] Mircea I. Chelaru, Christian Duval, and Mandar Jog. Levodopa-induced dyskinesias detection based on the complexity of involuntary movements. *Journal of Neuroscience Methods*, 186(1):81 – 89, 2010.
- [30] Rachel C. Colley, Andrew P. Hills, Neil A. King, and Nuala M. Byrne. Exercise-induced energy expenditure: Implications for exercise prescription and obesity. *Patient Education and Counseling*, 79(3):327 – 332, 2010.
- [31] James W Cooley and John W Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.
- [32] S. Cuccurullo and S.J. Cuccurullo. *Physical Medicine and Rehabilitation Board Review*. Demos Medical Publishing, 2004.

- [33] M.O. Derawi, P. Bours, and K. Holien. Improved cycle detection for accelerometer based gait authentication. In *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on*, pages 312–317, oct. 2010.
- [34] G Deuschl, P Bain, M Brin, and Adhoc Scientific Committee. Consensus statement of the movement disorder society on tremor. *Movement Disorders*, 13 Suppl 3:2–23, 1998.
- [35] J-P Eckmann, S Oliffson Kamphorst, and David Ruelle. Recurrence plots of dynamical systems. *EPL (Europhysics Letters)*, 4(9):973, 2007.
- [36] Leslie J. Findley. The economic impact of parkinson’s disease. *Parkinsonism and Related Disorders*, 13, Supplement(0):S8 – S12, 2007.
- [37] Jordan Frank, Shie Mannor, and Doina Precup. Activity and gait recognition with time-delay embeddings. In *AAAI*, 2010.
- [38] Andrew M. Fraser and Harry L. Swinney. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A*, 33:1134–1140, Feb 1986.
- [39] D. Gafurov, E. Snekkenes, and P. Bours. Gait authentication and identification using wearable accelerometer sensor. In *Automatic Identification Advanced Technologies, 2007 IEEE Workshop on*, pages 220–225, june 2007.
- [40] Davrondzhon Gafurov, Kirsi Helkala, and Torkjel Sndrol. Biometric gait authentication using accelerometer sensor. *Journal of Computers*, 1(7), 2006.
- [41] Mehrdad Ghassemi, Sarah Lemieux, Mandar Jog, Roderick Edwards, and Christian Duval. Bradykinesia in patients with parkinson’s disease having levodopa-induced dyskinesias. *Brain Res Bull*, 69(5):512–8, 2006.
- [42] M. Ghil, M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. Advanced spectral methods for climatic time series. *Reviews of Geophysics*, 40(1):1003–1, September 2002.

- [43] A. Godfrey, R. Conway, D. Meagher, and G. Laighin. Direct measurement of human movement by accelerometry. *Medical Engineering and Physics*, 30(10):1364 – 1386, 2008.
- [44] C. G. Goetz, G. T. Stebbins, H. M. Shale, A. E. Lang, D. A. Chernik, T. A. Chmura, J. E. Ahlskog, and E. E. Dorflinger. Utility of an objective dyskinesia rating scale for parkinson’s disease: Inter- and intrarater reliability assessment. *Movement Disorders*, 9(4):390–394, 1994.
- [45] M. Gomez Plaza, T. Martinez-Marin, S.S. Prieto, and D. Meziat Luna. Integration of cell-mapping and reinforcement-learning techniques for motion planning of car-like robots. *Instrumentation and Measurement, IEEE Transactions on*, 58(9):3094 –3103, sept. 2009.
- [46] Jackie Gour, Roderick Edwards, Sarah Lemieux, Mehrdad Ghassemi, Mandar Jog, and Christian Duval. Movement patterns of peak-dose levodopa-induced dyskinesias in patients with Parkinson’s disease. *Brain research bulletin*, 74(1-3):66–74, September 2007.
- [47] Peter Grassberger and Itamar Procaccia. Measuring the strangeness of strange attractors. *Physica D: Nonlinear Phenomena*, 9(12):189 – 208, 1983.
- [48] T.J. Harris and Hui Yuan. Heart rate variability analysis using correlation dimension and detrended fluctuation analysis. *ITBM-RBM*, 23(6):333–339, 2002.
- [49] T.J. Harris and Hui Yuan. Filtering and frequency interpretations of singular spectrum analysis. *Physica D: Nonlinear Phenomena*, 239(20-22):1958 – 1967, 2010.
- [50] Jeffrey M Hausdorff. Gait dynamics in parkinson’s disease: common and distinct behavior among stride length, gait variability, and fractal-like scaling. *Chaos*, 19(2):026113, 2009.
- [51] J. I. Hoff, A. A. V/d Plas, E. A. H. Wagemans, and J. J. van Hilten. Accelerometric assessment of levodopa-induced dyskinesias in Parkinson’s disease. *Movement Disorders*, 16(1):58–61, 2001.

- [52] J I Hoff, V van der Meer, and J J van Hilten. Accuracy of objective ambulatory accelerometry in detecting motor complications in patients with parkinson disease. *Clin Neuropharmacol*, 27(2):53–7, 2004.
- [53] C.S. Hsu. *Cell-to-Cell Mapping: A Method of Global Analysis for Nonlinear Systems*. Number v. 64 in Applied Mathematical Sciences. Springer, 1987.
- [54] Baoxin Hu, Qingmou Li, and A. Smith. Noise reduction of hyperspectral data using singular spectral analysis. *International Journal of Remote Sensing*, 30:2277–2296, January 2009.
- [55] Verne T Inman. Functional aspects of the abductor muscles of the hip. *The Journal of Bone and Joint Surgery (American)*, 29(3):607–619, 1947.
- [56] Verne T Inman, Howard D Eberhart, et al. The major determinants in normal and pathological gait. *The Journal of Bone and Joint Surgery (American)*, 35(3):543–558, 1953.
- [57] N. Iyengar, C. K. Peng, R. Morin, A. L. Goldberger, and L. A. Lipsitz. Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics. *American Journal of Physiology*, 271(4):1078–84, 1996.
- [58] H Kantz and T Schreiber. *Nonlinear Time Series Analysis*, volume 7. Cambridge University Press, 1997.
- [59] Justin J. Kavanagh and Hylton B. Menz. Accelerometry: A technique for quantifying movement patterns during walking. *Gait and Posture*, 28(1):1 – 15, 2008.
- [60] N. L. Keijsers, M. W. Horstink, and S. C. Gielen. Automatic assessment of levodopa-induced dyskinesias in daily life by neural networks. *Movement disorders : official journal of the Movement Disorder Society*, 18(1):70–80, January 2003.

- [61] Matthew B. Kennel, Reggie Brown, and Henry D. I. Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A*, 45:3403–3411, Mar 1992.
- [62] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint Conference on artificial intelligence*, volume 14, pages 1137–1145. Lawrence Erlbaum Associates Ltd, 1995.
- [63] J. Lenz and S. Edelstein. Magnetic sensors and their applications. *Sensors Journal, IEEE*, 6(3):631–649, june 2006.
- [64] Tuomas Liikavainio, Timo Bragge, Marko Hakkarainen, Pasi A. Karjalainen, and Jari P. Arokoski. Gait and muscle activation changes in men with knee osteoarthritis. *The Knee*, 17(1):69–76, 2010.
- [65] Peter Lindgren. Economic evidence in parkinson’s disease: a review. *The European Journal of Health Economics*, 5:s63–s66, 2004.
- [66] Irvin Hussein López-Nava and Angélica Muñoz Meléndez. Towards ubiquitous acquisition and processing of gait parameters. In *Proceedings of the 9th Mexican international conference on Advances in artificial intelligence: Part I, MICAI’10*, pages 410–421, 2010.
- [67] Vanesa Lores, Francisco García-Río, Blas Rojo, Sergio Alcolea, and Olga Mediano. Recording the daily physical activity of copd patients with an accelerometer: An analysis of agreement and repeatability. *Archivos de Bronconeumología*, 42(12):627–32, 2006.
- [68] Roberto Maestri, Maria Teresa La Rovere, Elena Robbi, and Gian Domenico Pinna. Fluctuations of the fractal dimension of the electroencephalogram during periodic breathing in heart failure patients. *Journal of Computational Neuroscience*, 28(3):557–565, 2010.

- [69] A J Manson, P Brown, J D O’Sullivan, P Asselman, D Buckwell, and A J Lees. An ambulatory dyskinesia monitor. *J Neurol Neurosurg Psychiatry*, 68(2):196–201, 2000.
- [70] Jani Mäntyjärvi, Mikko Lindholm, Elena Vildjiounaite, Satu marja Mäkelä, and Heikki Ailisto. Identifying users of portable devices from gait pattern with accelerometers. In *in IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005.
- [71] Michael Marschollek, Mehmet Goevercin, Klaus-Hendrik Wolf, Bianying Song, Matthias Gietzelt, Reinhold Haux, and Elisabeth Steinhagen-Thiessen. A performance comparison of accelerometry-based step detection algorithms on a large, non-laboratory sample of healthy and mobility-impaired persons. *Conf Proc IEEE Eng Med Biol Soc*, pages 1319–22, 2008.
- [72] E. Martin, V. Shia, and R. Bajcsy. Determination of a patient’s speed and stride length minimizing hardware requirements. In *Body Sensor Networks (BSN), 2011 International Conference on*, pages 144 –149, may 2011.
- [73] P. Martnez-Martn, A. Gil-Nagel, L. Morln Gracia, J. Balseiro Gmez, J. Martnez-Sarris, and F. Bermejo. Unified parkinson’s disease rating scale characteristics and structure. *Movement Disorders*, 9(1):76–83, 1994.
- [74] Merryn J Mathie, Adelle C F Coster, Nigel H Lovell, and Branko G Celler. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiological Measurement*, 25(2):R1, 2004.
- [75] M.J. Mathie, J. Basilakis, and B.G. Celler. A system for monitoring posture and physical activity using accelerometers. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 4, pages 3654–3657 vol.4, 2001.

- [76] Ruth E. Mayagoitia, Anand V. Nene, and Peter H. Veltink. Accelerometer and rate gyroscope measurement of kinematics: an inexpensive alternative to optical motion analysis systems. *Journal of Biomechanics*, 35(4):537 – 542, 2002.
- [77] Paul McCrone, Liesl M. Allcock, and David J. Burn. Predicting the cost of parkinson’s disease. *Movement Disorders*, 22(6):804–812, 2007.
- [78] Thomas O. Mera, Dustin A. Heldman, Alberto J. Espay, Megan Payne, and Joseph P. Giuffrida. Feasibility of home-based automated parkinson’s disease motor assessment. *Journal of Neuroscience Methods*, 203(1):152 – 156, 2012.
- [79] John G. Milton. Introduction to focus issue: Bipedal locomotion—from robots to humans. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 19(2):026101, 2009.
- [80] Thomas B Moeslund and Erik Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, 81(3):231–268, 2001.
- [81] Orna Moore, Chava Peretz, and Nir Giladi. Freezing of gait affects quality of life of peoples with parkinson’s disease beyond its relationships with mobility and gait. *Movement Disorders*, 22(15):2192–2195, 2007.
- [82] Michael J Mueller, Scott D Minor, Shirley A Sahrman, James A Schaaf, and Michael J Strube. Differences in the gait characteristics of patients with diabetes and peripheral neuropathy compared with age-matched controls. *Physical therapy*, 74(4):299–308, 1994.
- [83] Jeff A Nessler, Charles J De Leone, and Sara Gilliland. Nonlinear time series analysis of knee and ankle kinematics during side by side treadmill walking. *Chaos*, 19(2):026104, 2009.
- [84] Mark S. Nixon, Tieniu N. Tan, and Rama Chellappa. *Human Identification based on Gait*. International Series on Biometrics. Springer, 2005.

- [85] World Health Organisation. *Global Burden of Disease 2004*. World Health Organization, 2004.
- [86] Shyamal Patel, Konrad Lorincz, Richard Hughes, Nancy Huggins, John Growdon, David Standaert, Metin Akay, Jennifer Dy, Matt Welsh, and Paolo Bonato. Monitoring motor fluctuations in patients with parkinson’s disease using wearable sensors. *Trans. Info. Tech. Biomed.*, 13(6):864–873, November 2009.
- [87] Peter De Weerd Eric Lowet Mark Roberts Ronald Westra Olivier Meste Pietro Bonizzi, Jol Karel and Ralf Peeters. Singular spectrum analysis improves analysis of local field potentials from macaque v1 in active fixation task. In *Proceedings in EMBC12*, 2012.
- [88] Werner Poewe. Treatments for parkinson disease—past achievements and current clinical needs. *Neurology*, 72(7 Suppl):S65–73, 2009.
- [89] Marko Robnik-Šikonja and Igor Kononenko. Theoretical and empirical analysis of relieff and rrelieff. *Mach. Learn.*, 53(1-2):23–69, October 2003.
- [90] A. M. Sabatini, C. Martelloni, S. Scapellato, and F. Cavallo. Assessment of walking features from foot inertial sensing. *IEEE Transactions on Biomedical Engineering*, January 2005.
- [91] Albert Samà, Cecilio Angulo, Diego E. Pardo, Andreu Català, and Joan Cabestany. Analyzing human gait and posture by combining feature selection and kernel methods. *Neurocomputing*, 74(16):2665–2674, 2011.
- [92] Albert Samà, Andreu Català, A Rodríguez-Molinero, and C Angulo. Extracting gait spatiotemporal properties from parkinson’s disease patients. In *Neural Information Processing Systems, Mini Symposia on Assistive Machine Learning for People with Disabilities*, 2009.
- [93] Albert Samà, Diego E Pardo-Ayala, Joan Cabestany, and A Rodríguez-Molinero. Time series analysis of inertial-body signals for the extraction of

- dynamic properties from human gait. In *Neural Networks (IJCNN), The 2010 International Joint Conference on*, pages 1–5. IEEE, 2010.
- [94] Albert Samà, Carlos Pérez-López, Daniel Rodríguez-Martín, Cabestany Joan, Juan Manuel Moreno-Arostegui, and Alejandro Rodríguez-Molinero. A heterogeneous database for movement knowledge extraction in parkinsons disease. In *21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, April 2013.
- [95] Albert Samà, Francisco J Ruiz, and Núria Agell. Granular singular spectrum analysis for gait recognition. In *26th International Workshop on Qualitative Reasoning (QR) 2012, Los Angeles, California (USA)*, 2012.
- [96] Albert Samà, Francisco J Ruiz, Núria Agell, Carlos Pérez-López, Andreu Català, and Joan Cabestany. Gait identification by means of box approximation geometry of reconstructed attractors in latent space. *Neurocomputing*, 2013.
- [97] Albert Samà, Francisco J Ruiz, Carlos Pérez, and Andreu Català. Gait identification by using spectrum analysis on state space reconstruction. In *Lecture Notes in Computer Science*, pages 597–604. Springer Berlin Heidelberg, 2011.
- [98] Albert Samà, Francisco J Ruiz, Carlos Pérez-López, and Andreu Català. Gait recognition by using spectrum analysis on state space reconstruction. *Artificial Intelligence Research and Development: Proceedings of the 14th International Conference of the Catalan Association for Artificial Intelligence*, 232:228, 2011.
- [99] C.; Romagosa J.; Rodríguez-Martín D.; Català A.; Cabestany J.; Pérez-Martínez D.A. Samà, A.; Pérez-López and A. Rodríguez-Molinero. Dyskinesia and motor state detection in pd patients with a single movement sensor. In *Engineering in Medicine and Biology Society, EMBC, 2012 Annual International Conference of the IEEE*, pages 1194–97, Aug 28-Sept 1 2011 2012.
- [100] Nicola Scafetta, Damiano Marchi, and Bruce J West. Understanding the complexity of human gait dynamics. *Chaos*, 19(2):026108, 2009.

- [101] B. Schölkopf and A.J. Smola. *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. the MIT Press, 2002.
- [102] Manfred Schroeder. *Fractals, Chaos, Power Laws: Minutes From an Infinite Paradise*. W. H. Freeman, 1991.
- [103] Abd-Krim Seghouane and Adnan Shah. Functional brain connectivity as revealed by singular spectrum analysis. In *Proceedings in EMBC12*, 2012.
- [104] A. Seireg and R.J. Arvikar. The prediction of muscular load sharing and joint forces in the lower extremities during walking. *Journal of Biomechanics*, 8(2):89 – 102, 1975.
- [105] Claude Elwood Shannon and Warren Weaver. A mathematical theory of communication, 1948.
- [106] T. Shany, S.J. Redmond, M.R. Narayanan, and N.H. Lovell. Sensors-based wearable systems for monitoring of human movement and falls. *Sensors Journal, IEEE*, 12(3):658–670, March.
- [107] Sebastijan Sprager and Damjan Zazula. Gait identification using cumulants of accelerometer data. In *Proceedings of the 2nd WSEAS International Conference on Sensors, and Signals and Visualization, Imaging and Simulation and Materials Science*, pages 94–99, 2009.
- [108] Dinesh K. Kumar Sridhar P. Arjunan and Teodiano Bastos. Fractal based complexity measure and variation in force during sustained isometric muscle contraction: Effect of aging. In *Proceedings in EMBC12*, 2012.
- [109] Glenn T. Stebbins, Christopher G. Goetz, Anthony E. Lang, and Esther Cubo. Factor analysis of the motor section of the unified parkinson’s disease rating scale during the off-state. *Movement Disorders*, 14(4):585–589, 1999.
- [110] F. Takens. Detecting strange attractors in turbulence. In D. A. Rand and L.S. Young, editors, *Dynamical Systems and Turbulence, Warwick, 1980, Lecture notes in mathematics vol. 898*, pages 366–381, Berlin, 1981. Springer.

- [111] D. Titterton, J. Weston, D.H. Titterton, and J.L. Weston. *Strapdown Inertial Navigation Technology, 2nd Edition*. Iee Radar Series. The Institution of Engineering and Technology, 2004.
- [112] Ho Huu Minh Tam Truong Quang Dang Khoa and Vo Van Toi. Extracting fetal electrocardiogram from being pregnancy based on nonlinear projection. *Mathematical Problems in Engineering*, 2012:1–13, 2012.
- [113] M.G. Tsipouras, A.T. Tzallas, D.I. Fotiadis, and S. Konitsiotis. On automated assessment of levodopa-induced dyskinesia in parkinson’s disease. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 2679 –2682, 30 2011-sept. 3 2011.
- [114] Rajendra Acharya U, N Kannathal, and S M Krishnan. Comprehensive analysis of cardiac health using heart rate signals. *Physiological Measurement*, 25(5):1139, 2004.
- [115] M.G.H. van Weering, M.M.R. Vollenbroek-Hutten, T.M. Tnis, and H.J. Hermens. Daily physical activities in chronic lower back pain patients assessed with accelerometry. *European Journal of Pain*, 13(6):649–654, 2009.
- [116] R. Vautard and M. Ghil. Singular spectrum analysis in nonlinear dynamics with applications to paleoclimatic time series. *Physica D*, 35:395, 1989.
- [117] R Vautard, P Yiou, and M Ghil. Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D: Nonlinear Phenomena*, 58(1-4):95–126, 1992.
- [118] C. Vitale, M. T. Pellecchia, D. Grossi, N. Fragassi, T. Cuomo, L. Di Maio, and P. Barone. Unawareness of dyskinesias in parkinson’s and huntington’s diseases. *Neurological Sciences*, 22:105–106, 2001.
- [119] Aner Weiss, Talia Herman, Meir Plotnik, Marina Brozgol, Inbal Maidan, Nir Giladi, Tanya Gurevich, and Jeffrey M. Hausdorff. Can an accelerometer enhance the utility of the timed up and go test when evaluating patients with parkinson’s disease? *Medical Engineering and Physics*, 32(2):119 – 125, 2010.

- [120] Aner Weiss, Sarvi Sharifi, Meir Plotnik, Jeroen P P van Vugt, Nir Giladi, and Jeffrey M Hausdorff. Toward automated, at-home assessment of mobility among patients with parkinson disease, using a body-worn accelerometer. *Neurorehabil Neural Repair*, 25(9):810–8, 2011.
- [121] P. G. Weyand and M. W. Bundle. Last word on point:counterpoint: Artificial limbs do make artificially fast running speeds possible. *Journal of Applied Physiology*, 108(4):1019, April 2010.
- [122] Michael Whittle. *Gait analysis: an introduction*. Elsevier, 3 edition, 2002.
- [123] Ge Wu, Sorin Siegler, Paul Allard, Chris Kirtley, Alberto Leardini, Dieter Rosenbaum, Mike Whittle, Darryl D DLima, Luca Cristofolini, Hartmut Witte, et al. Isb recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motionpart i: ankle, hip, and spine. *Journal of biomechanics*, 35(4):543–548, 2002.
- [124] Che-Chang Yang and Yeh-Liang Hsu. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, 10(8):7772–7788, 2010.
- [125] Wang Yangang, Zheng Haiqi, Guan Zhenzhen, and Gao Feng. Application of cell mapping method in global analysis of fault gear system. In *Network Computing and Information Security (NCIS), 2011 International Conference on*, volume 2, pages 352 –354, may 2011.
- [126] Wiebren Zijlstra and At L Hof. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait and Posture*, 18(2):1 – 10, 2003.
- [127] HaiLin Zou and JianXue Xu. Improved generalized cell mapping for global analysis of dynamical systems. *Science in China Series E: Technological Sciences*, 52:787–800, 2009.

