Czech technical univeristy in Prague

Faculty of electrical engineering

Department of cybernetics

DIPLOMA THESIS

Tremor Detection Algorithm for Parkinsonian Patients

Author: Eduard Bakštein

Supervisor: Kevin Warwick

Prague, 2010

Czech Technical University in Prague
Faculty of Electrical Engineering

Department of Cybernetics

DIPLOMA THESIS ASSIGNMENT

Student: Bc. Eduard B a k š t e i n

Study programme: Electrical Engineering and Information Technology

Specialisation: Biomedical Engineering

Title of Diploma Thesis: Tremor Detection Algorithm for Parkinsonian Patients

Guidelines:

The aim is to develop a tremor detection algorithm for Parkinsonian patients.

Input to the detection algorithm is provided by electrical signals from the patients´ brainstem, while a corresponding EMG signal is also available for training purposes.

The goal is to be achieved by preprocessing, feature description of the signals provided and classifier design for the features selected.

Work will be implemented in MATLAB environment.

Bibliography/Sources: Will be provided by the supervisor.

Diploma Thesis Supervisor: prof. Kevin Warwick

Valid until: the end of the winter semester of academic year 2010/2011

prof. Ing. Vladimír Mařík, DrSc.
Head of Department

doc. Ing. Boris Šimák, CSc.
Head

Prague, May 25, 2009

České vysoké učení technické v Praze
Fakulta elektrotechnická

Katedra kybernetiky

ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: Bc. Eduard Bakštein

Studijní program: Elektrotechnika a informatika (magisterský), strukturovaný

Obor: Biomedicínské inženýrství

Název tématu: Detekce třesu u pacientů trpících Parkinsonovou chorobou

Pokyny pro vypracování:

Cílem práce je návrh detekčního algoritmu pro pacienty trpící Parkinsonovou chorobou.

Vstupem algoritmu jsou poskytnuté elektrické signály mozkového kmene, korespondující EMG signály jsou k dispozici pro účely trénování.

Cíle bude dosaženo předzpracováním, výběrem příznaků a natrénováním klasifikátoru pro vybrané příznaky.

Práce bude provedena v prostředí MATLAB.

Seznam odborné literatury: Dodá vedoucí práce.

Vedoucí diplomové práce: prof. Kevin Warwick

Platnost zadání: do konce zimního semestru 2010/2011

prof. Ing. Vladimír Mařík, DrSc.
vedoucí katedry

doc. Ing. Boris Šimák, CSc.
děkan

V Praze dne 25. 5. 2009

Prohlášení

Prohlašuji, že jsem svou diplomovou práci vypracoval samostatně a použil jsem pouze podklady (literaturu, projekty, SW atd.) uvedené v přiloženém seznamu.

V Praze dne 4.1.2010

podpis

Acknowledgement

I would like to express my appreciation to Mr. Kevin Warwick, my supervisor, who provided me with great support during the work on this thesis, as well as to Mr. Øyvind Stavdahl, my supervisor at NTNU Trondheim, for his countless suggestions and ideas that kept the progress of my work on the right way.

Abstrakt

Tato práce se zabývá vývojem metody vhodné pro detekci tremoru u pacientů s Parkinsonovou chorobou. Detekce je založena na signálech subtalamického jádra, zaznamenaných prostřednictvím elektrod používaných pro hlubokou mozkovou stimulaci.

Hlavní část této práce je věnována zkoumání rozličných vlastností zaznamenaných signálů subtalamického jádra a srovnávání jejich vlastností na nahrávkách s tremorem a bez něj. Na základě provedených pozorování je navržena řada příznaků, vycházejících z časových, spektrálních autokorelačních a dalších vlastností zkoumaného signálu. Tyto příznaky jsou následně testovány na reálných datech a jejich význam vzhledem k detekci tremoru je ověřen pomocí klasifikačního procesu.

Výsledky hodnocení příznaků i klasifikace ukazují na velký význam spektrálních vlastností zkoumaných signálů, přičemž frekvenční pásma 0–1 Hz a 3.5–5.5 Hz se ukázala jako zvlášť významná. Na mnoha datech byl pozorován vysoký podíl energie v oblasti posledně jmenovaného pásma, jehož frekvence velmi dobře koresponduje s frekvencí tremoru, pozorovanou v souběžných EMG signálech. V klasifikačním procesu bylo dosaženo velice dobrých výsledků pro 3 z 5 pacientů, kdy senzitivita dosahovala až 94% při specificitě rovné jedné. Pro zbylé dva pacienty podala metoda naopak velice špatné výsledky, což může být zdůvodněno velice nízkým množstvím dat, které bylo na hranici proveditelnosti použitých metod.

Abstract

This diploma thesis deals with development of a suitable method for detection of tremor in patients with Parkinson's disease. The detection is based on recorded local field potentials of the Subthalamic nucleus (STN), captured through an electrode of a deep brain stimulation device.

The main part of this work is dedicated to research of various properties of the STN signal, compared on tremor and non-tremor recordings in different patients. Features based on temporal, spectral, statistical, fractal and autocorrelation-related discovered properties of the signals are developed and implemented. Evaluation and comparison of the features is done on the available data during a classification process, using short sections of preprocessed signals.

The results show great importance of spectral properties of the signal, which corresponds with previous research in this field. In particular, the frequency bands 0–1 and 3.5–5.5 Hz proved great significance to the problem, the latter corresponding with the tremor frequency found in the simultaneous EMG recordings. Classification based on multiple features showed very good results for 3 out of 5 patients, reaching up to 94% sensitivity with specificity equalling one. On the contrary, the method did not work with the two remaining patients, possible reason being substantial lack of training data.

Contents

Abbreviations

PD	Parkinson's disease	Chronic neurological disorder, affecting mainly motor function of the diseased. Tremor, akinesia and rigidity are among the most common symptoms.
DBS	Deep brain stimulation	Therapeutic method based on application of electrical impulses to the structures in the human brain.
BG	Basal ganglia	Brain structure positioned mainly in the midbrain, involved also in modulation of movement. A dysfunction of this structure causes the Parkinson's disease.
STN	Subthalamic nucleus	Structure in the brain, functional unit of the Basal ganglia, involved in the modulation of motor functions. A target structure for DBS in Parkinson's disease.
GPi	Globulus Pallidus	Structure of the Basal ganglia, alternative target of DBS in Parkinson's disease patients.
LFP	Local field potentials	Summarized activity of the neuronal tissue in a specific region as captured by an electrode pair.
L-dopa	Levodopa	Dopamine precursor, medication used for conventional treatment of the symptoms of the Parkinson's disease.
ROC	Reciever operating characteristic	Graphical plot of binary classifier system properties, used for system evaluation.
AUC	Area under curve	Area under the ROC curve, used for simple evaluation of classifier properties.
FDR	Fisher discriminant ratio	Measure used to characterize difference between two probability distribu tions, taking both mean and variance into account.

Chapter 1

Introduction

Over the last decades, technology has become an essential part of modern medicine, being present in nearly all branches of this rapidly developing field. Besides the common technical areas of health care such as diagnosis, surgical and therapeutic tools, communication and other, technical devices can also assist in applications where traditional procedures fail or are completely unsuitable. Dating back to the 60's when the clinical application of implantable heart pacemakers was introduced, the chronic electrical stimulation of human tissue became an assistive method of that kind, supporting life of thousands of patients. Several decades later, when the electrical stimulation of the heart had already become a common procedure, the benign effects of a similar stimulation of the human brain were discovered for motor, psychiatric and other neurodegenerative disorders. It was shown that the application of electrical stimuli to specific structures in the brain can reduce or even eliminate symptoms of these severe diseases and help the patients lead normal lives. To accomplish this, special implantable devices have been developed, delivering electrical impulses directly into the brain. These devices, commonly referred to as the *Deep brain stimulators* (DBS), have found their way to clinical application over the years. However, the mechanisms of their positive effects still remain mostly unclear.

Just as any other implantable device, the DBS suffer from the great limitation of battery life, implicating the need of re-implantation in a period of circa 2 years. Even though the device itself is placed in the chest cavity and no neurosurgery is involved during the re-implantation process, the repeated surgery of the unchanging region of the body brings many complications and discomfort. And this is true especially in elderly patients—the most common recipients of this therapeutic method. In case of the DBS for Parkinson's disease, which is the main topic of this work, the uninterrupted stimulation is unnecessary for a great group of the patients, as the symptoms occur only at times. Contrary to the

possibility of switching the device state manually, available in the current stimulation devices, an automatic detection of the onset of the symptoms could be developed, switching the device on just when necessary—on demand. Although the occurrence of the disease onset in the activity of the stimulated brain structures is yet to be proved, it is believed that the signals of these structures, recorded through the stimulation electrodes, could be used for the detection. Such a solution could significantly improve the battery lifespan of the device. This way, the beneficial effects of the method could be preserved without need of extra inputs or modifications to the surgical procedures.

This work tries to offer a procedure to accomplish this task for one of the severe symptoms of the Parkinson's disease: tremor. The work has been conducted in cooperation with Kevin Warwick and Jon Burgess from the University of Reading, who provided their current knowledge on the topic and the recorded patient data, used in the study.

First of all, the basic information on Parkinson's disease and Deep brain stimulation is provided in this introductory chapter. Then, the process of sorting and preprocessing the data is described in Chapter 2.

The process of looking for suitable features of the brain signals is described in Chapter 3. The examination of the possible approaches to the detection is based partly on current scientific findings about the brain signals, partly on my own research of the data. The result is formed by a feature set, covering several different domains, including temporal, spectral, statistical and other. Both successful and unsuccessful research is presented in this chapter.

The evaluation of the developed features is performed in Chapter 4, where the features are subjected to experiments with the patient data, evaluating their classification power with regard to the tremor. The patient data is divided into small sections, for which the features are calculated and which then form training and testing samples for a subsequent classification process.

First, the features are tested individually on the whole dataset, providing an overall information about them. Later on, a subset of features is selected and tested using training data for single patients. In the end, the results of both steps are combined, presented and discussed together with the classification results.

In Chapter 5, the results and achievements are summarized, weaknesses are discussed and further research together with possible improvements to the outlined methods are proposed.

References to the implemented methods, performing the respective tasks, are given throughout the text. All the methods have been implemented for the purpose of this

work, unless stated otherwise.

1.1 Parkinson's disease and the Deep brain stimulation

The Parkinson's disease (PD) is a chronic neurological disorder, affecting mainly elderly patients over the age of 60, however can be found in younger people too. The most common symptoms are connected with motor functions and include:

- Tremor; shaking incontrollable movement of a body part, usually a limb (hand, arm or leg), occurring when the person is awake and in a still position
- Muscle rigidity (stiffness)
- Bradykinesia; Slow, difficult movement, especially when the patient moves from resting position
- Gait disorders; problems with stability

From the physiological point of view, the PD is caused by lack of dopamine, a neurotransmitter in a brain structure called the basal ganglia (BG). In PD, the dopamine producing cells in substantia nigra (a part of the basal ganglia) deteriorate, making the dopamine production insufficient. This causes an impaired function of other basal ganglia parts, such as the subthalamic nucleus, an important part of the motor system.

In spite of the ongoing research of PD causes, the actual reasons of this disorder are unknown. Late reports show connection between PD and exposition to specific chemicals (toxins) as well as genetic predispositions. This is true especially in patients with early developed symptoms [15].

The standard treatment of PD is based on supplying patients with dopamine precursor replacements such as Levodopa (l-dopa), which is effectual in preventing some of the PD symptoms. Nevertheless, after a few years of l-dopa application, the positive effect goes off and side effects such as hypotensia, arrythmia or nausea become more significant, which grounded the research for new and progressive methods.

One of the possible ways to treat the symptoms of the disease is the Deep Brain Stimulation. In this treatment, electrical pulses are delivered to the brain using an

electrode implanted in a part of the Basal Ganglia, usually the Subthalamic nucleus or the Globulus pallidus. This requires an advanced neurosurgery, due to the deep position of the target structures in the midbrain. However, the method gives another chance to the patients who have lost the response to the medication or for whom the standard therapeutic procedures are unavailable due to various contraindications. Moreover, fewer side-effects have been identified with the DBS than with medication.

The DBS system consists of the simulator itself, implanted in the chest cavity, and the implanted electrode, connected to the device. The connecting wires are hidden under the patient's skin and are attached to the device using special connectror, allowing exchange of the device if needed. The electrode consists of four pins in a row, spaced 5 mm apart. Once the electrode is in place, the medical experts search for appropriate electrode combination and stimulation pulse setting. The settings with the best symptom-reducing effect is chosen.

The signals, recorded from the electrodes of the DBS device, capture the summarized electrical activity of the neuron population between them, called the *Local field potentials* (LFP). These signals are different from the *microelectrode recordings*—a type of recording used during the operation—due to the different electrode sizes and spacing.

Chapter 2

Data

Description of the dataset is provided in this chapter, including the process of selecting and preprocessing the input files.

The original data files for this study come from the Theatre of Radcliff infirmary, University of Oxford, UK. The recordings were done during the peri-operative period when the deep-brain electrodes are already in place and the patient is overseen at hospital. It is the period prior to the implantation of the DBS device. The recordings were taken during several sessions in the years 1999 to 2006.

2.1 Data content

Each of the recordings comes in a text file and contains several signals, including recorded:

- LFP - Local Field Potentials of the Basal ganglia
- surface EMG from flexor and-or extensor forearm muscle
- accelerometer signal from the forearm

and is accompanied by a document, describing recording conditions.

As stated in Chapter 1, the deep-brain electrode is implanted in the basal ganglia (Subthalamic nucleus or Globulus Pallidus) and consists of 4 electrode pins, spaced 5 mm apart. This provides 3 electrode pairs, when only the neighboring electrodes are considered. In some patients, provided with a bilateral implant, 6 channels are hypothetically available, 3 for each hemisphere. For some patients all the possible signals are available in the recordings, while only one or one per file is at disposal for others.

As for the supportive signals, at least one corresponding EMG signal is present in each file, providing information about the movement of patient's forearm. Two EMG signals are usually available, one for flexor and one for extensor forearm muscle. Signal from an accelerometer, attached to the patient's forearm is also available in some recordings, providing more information about patient's movement. The supplementary signals were recorded from the opposite side of the body than the LFP signals, due to the crosswise connection of the motor paths in the human body.

The document attached to each file contains description of the recording conditions, patient information and information about the channels in the data files.

2.2 File selection

At the time of recording, not all the data were intended for tremor detection, which made most of the obtained files unusable for the purposes of this work. There were files recorded during different modes of patient movement (no movement, active movement of the wrist, wrist moved passively by the physician), different settings of the stimulation device (off or stimulating into different electrodes) and also recordings taken from different parts of the Basal Ganglia (STN or GPi). This provided a very wide range of conditions, varying from patient to patient. The original idea at the beginning of this work was to include as much data as possible, covering various recording conditons. Nevertheless, the selection was reduced to specific conditions later on. The decision was based partly on insufficiency of the data for such a general task, partly on the huge variability of the data across different conditions, which would make the original task almost unfeasible. The required recording conditions were restrained to:

- Signals exclusively from the STN.
- No movement present during the recording (patient at rest)
- Stimulation off during and prior to recording.

Another factor that resulted in exclusion of more files was the lack of description, which made the determination of recording conditions impossible in some cases. Some of the files also lacked some of the crucial signals (LFP or at least one corresponding signal) and had to be excluded as well.

The original file set contained 231 mostly unsorted files with description. Applying the restrictions specified above, only 33 files met the conditions to be included in the dataset. The main reasons for file exclusion were[1]:

- Movement or DBS present (around 50%)
- Crucial signals missing (STN or EMG, around 20%)
- Lack of information (description missing, unclear etc, less than 20%)
- Unclear data (unable to specify tremor and non-tremor periods, around 10%)

The loss of data the restrictions caused is significant. However, the impact was considered carefully and the requirements were set in order to retain the maximum number of patients with suitable data. From the 14 patients in the original file set, 10 were included in the file set, as some of the patient had a GPi implant, and some lacked documentation for all the files.

2.3 Data sectioning

Once the dataset was selected, the individual files could be described for further use. The conditions changed during the recording in some cases; some recordings contained tremor onset or changes to the movement and stimulation mode, so that only a part of these files could be used. In order to use the maximal possible amount of data, even single file sections were used where possible. To achieve this, markers have been attached to the files at positions specified either by the documentation (stimulation on or movement started at certain moment) or the corresponding EMG signal[2].

The signal artefacts observed at the beginning and end of some recordings, as well as some other factors (e.g. changes of the recording conditions) were also removed from the useful sections in order to achieve higher consistency of the data. Figure 2.1 shows an example of time series from an original file, prior to any preprocessing.

[1] Reasons presented in order of the percentage of files excluded.

[2] To carry out the tremor detection from the EMG signal, it had to be preprocessed first. The preprocessing was done according to thestandard method used in prosthesis control [12], including 50 Hz filtering, rectification and smoothing using the moving average filter. This way an envelope EMG signal was produced, showing the tremor activity.

An important note, implying further design of the classifier algorithm, is on the nature of the tremor onset: As can also be seen in the Figure 2.1, the onset of a tremor is rather a dynamic process, than a discrete event at a discrete time. A section of a special type was therefore placed around the tremor onset in order not to mix the tremor and non-tremor properties in the sections. Nevertheless, the period shortly before the onset was the only non-tremor section available in some cases and the section could not be made long enough to cover the whole onset-related development. This might have had negative impact to the classification process, as described in Chapter 4.

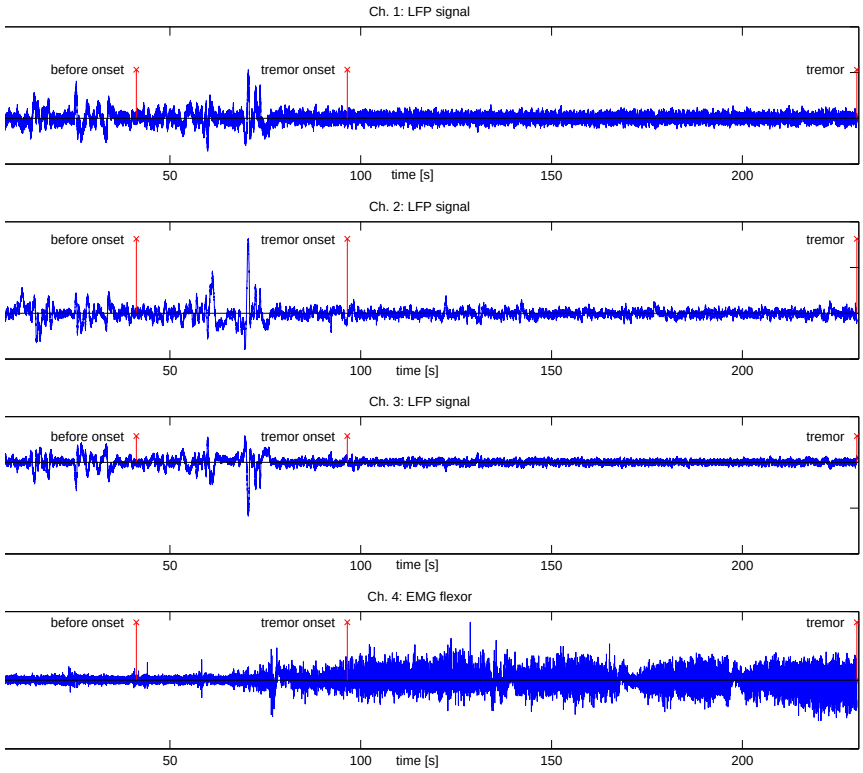

Figure 2.1: An example of an unprocessed file (Patient 9, file 28), with added section descriptions as plotted by the function parkinsonian_plot_data.

The Table 2.1 shows the numbers of files and sections for each of the patients after the selection. Unfortunately, the data is missing in either tremor or non-tremor class for 5 out of 10 patients. This means these patients can not be fully included in the classification process. Nevertheless, the data of these patients can be used for general analysis and

estimation of qualities of the feature set.

Loading of the data and the sectioning process are carried out by the following implemented functions and scripts:

`parkinsonian_file_load.m` - loading a text file

`load_files.m` - instructions for the file loading function, containing information about channel and other settings for each of the selected files

`parkinsonian_plot_data.m` - time, spectral and spectrogram plots of the signals of one file

Table 2.1: Overview of the final dataset, showing number of sections and file of specific content for each patient.

patient			**files**				**sections**			
id	initials	LFP ch.	total	T	NT	onset	total	T	NT	onset
1	be	1	6	1	5		8	2	6	
2	ep	1	3	3			6	6		
3	ma	1	5	3	2		5	3	2	
4	dc	1	3	3			4	4		
5	CG	3	4	4	2	4	10	4	2	4
6	DS	6	1		1		1		1	
7	GA	6	1		1		4		4	
8	rb	1	4	4			8	8		
9	RB	3	2	2	2	2	6	2	2	2
10	sw	1	4	2	2		4	2	2	
total		24	33	22	15	6	56	19	31	6

2.4 Preprocessing

Investigation of the dataset showed that several preprocessing and normalization steps were necessary before the data could be evaluated.

2.4.1 Unification of the sampling frequency

First, the sampling frequency varied from 250–1000 Hz, which made further preprocessing and feature calculation uneasy. Therefore all the data was downsampled to 250 Hz, using the Matlab function `resample`. This function applies appropriate low-pass filtering prior to downsampling [14]. The downsampling routine for the whole parkinsonian files, including also recalculation of section markers etc. is implemented in function `parkinsonian_data_resample.m`

2.4.2 50 Hz noise filtering

Second issue, faced during the data processing, was a very high level of 50 Hz noise in most recordings. The high energy carried by this frequency overrode any useful signal in most recordings. For the purpose of the first analyses, the data was preprocessed with a narrow notch filter (`remove50.m`). Once it was clear that all the designed features work only with frequencies up to 20 Hz, he preprocessing step has been changed to a 50 Hz Type II Chebyshev low-pass filter. The filtering was done using zero-lag filtering Matlab function `filtfilt.m`.

2.4.3 Amplitude normalization

Another property, varying highly across the dataset was the amplitude scale of the signals; The recording level was not specified for any of the input files. This was probably also caused by the high amount of noise, which changed the overall signal amplitude rapidly in many files. Once the noise was removed, the amplitude normalization turned out to be necessary, as the signal levels varied in several orders across the set. To perform this task, the `parkinsonian_data_normalize.m` function was implemented, normalizing the mean value of the rectified signal to one[3]. All the sections containing useful signal were used together for calculation of the normalization coefficient. The whole signal was then normalized. The normalization was done for each channel separately

[3]The mean value of the signal is subtracted prior to rectification.

2.5 Data containing structures

The data has been loaded to a MATLAB structure array, carrying all available information about the original file, numbers of channels in the original text file and section content. Each item in the dataset can contain any number of channels and section. The original text files can be loaded using the already-mentioned function `parkinsonian_file_load.m`.

When the preprocessing process is done, the sections of a desired type, can be extracted using the function `parkinsonian_data_getsections.m`. The sections produced contain only one LFP channel each, bound together with all corresponding signals and information about the original file. For example: 6 section files are produced from a preprocessed data file of 3 LFP channels and 2 sections. These section files are further divided into small portions, used for feature calculation and classification, as described in Chapter 4

2.6 Chapter summary

The process of data selection and preprocessing has been described in this chapter. Description of the consequent analyses of the data can be found in the next chapter, dedicated to feature calculation and implementation. However, it can be derived from the first rough views of the data that some changes obviously occur during the tremor onset in the STN signals. It should therefore be possible to track the onset using the methods of artificial intelligence.

Chapter 3

Features

Correct feature selection is fundamental for achieving good results in a detection or classification task. Nature of the features can range from very simple to very complex ones, using advanced calculations or transforms. In processing of well known signals, basic features or properties that lead to successful detection results are usually known. The detection task then consists of finding correct combination of these features, improving them or searching for new ones, that would substitute the current methods or utilize other aspects of the signal of interest. Many proved features or feature sets exist in biomedical engineering for signals such as ECG or EEG or EMG that have been known and researched for many years.

In case of this study, deep brain stimulation for Parkinsonian disease is a relatively young topic, with the main outbreak dating to 1990's [2]. To record electrical signals of the basal ganglia, an electrode implantation inlcuding complicated surgery is necessary. Therefore, very little research of these signals had been done before the positive effects of electrical stimulation of this brain structure were discovered. Anyway, even after the deep brain stimulation found its way among respected treatment techniques for the PD, it took several years to start research of the signals. Particularly, there is a growing knowledge about electrical activity of the STN signal, even though no generally applicable method has been found for tremor detection.

The research done so far, consisted of tracking changes in the signal spectrum during tremor onset [16]. This method is successful for some data, but not robust enough for being generally applicable. As a result of this, further research is carried out, investigating different aspects of the STN signal, part of which is also this thesis.

The original aim was to cover the widest possible area of different approaches to create varied feature set and avoid important signal properties being missed. Basic domains, in

which the features were searched included:

- Time domain: Common signal descriptors such as amplitude, energy etc.
- Statistics: Statistical moments of different orders.
- Frequency: Energy in different bands, shape of the frequency spectrum.
- Signal theory: Shannon entropy.
- Autocorrelation function-based features
- Other: fractal dimension, sample average etc.

After obtaining the whole dataset it was found that a great number of features will have to be left out due to necessary amplitude normalization (see Section 2.4.3). This included mainly d.c. component and absolute amplitude based features, which are usually the first-choice parameters in signal processing tasks. This caused a significant loss of information, carried by the data and should therefore be avoided during the signal-recording sessions to come.

The following section focuses on individual features, describing motives for including them in the feature set, as well as the development process with attempted methods used. The name of respective function implemented is given at each section. Some of the simple features are implemented directly in the main function `parkinsonian_get_feature.m`, which is used for obtaining any (or all) of the features from an input sequence.

3.1 Time domain

Examination of the actual values of the signal as they come in time is the first-choice method for most detection or signal processing tasks. This approach does not usually require any transformation of the input data and is therefore easy to accomplish with simple means. Not any different in this application, the simple properties of the signals should be researched first, prior to introduction of more sophisticated and complex methods.

The main problem of the time–domain methods arises from the amplitude normalization, described in Chapter 2. The absolute values cannot be compared across different files nor patients. Despite these facts, several time-domain features were implemented.

Only those features, less dependent on the exact amplitude value were implemented, while several others (peak-to-peak value and mean value) were excluded from the feature set.

3.1.1 Short-time energy

Signal power, or short time energy is first of the features implemented. It calculates time average of energy, according to the formula (for discrete series of N samples and sampling frequency f_s):

$$P_f = \frac{f_s}{N} \sum_{n=1}^{N} |f(n)|^2 \tag{3.1}$$

and is implemented directly in the `parkinsonian_get_feature.m` method

3.1.2 Zero crossing rate

The zero-crossing rate calculates the time-weighted number of times the signal value crosses zero. This means the sampling frequency is taken into account. In simple cases, this feature can give a rough measure of the prevailing frequency in the signal without the need of more sophisticated calculations. Another advantage is that it can be calculated in real time. The feature is implemented in the function `zerocross.m` and is weighted by the length of the sequence and the sampling frequency.

3.1.3 The first differential

Another possibility is the first differential that can be roughly and simply calculated as the difference between values of the consecutive samples in the series. Two features were implemented, based on the first differential:

- Maximum of the first differential on the segment: *maxdiff*
- Average of the first differential on the segment: *avgdiff*

Both features are length-invariant and implemented direcly in the `parkinsonian_get_feature.m` method.

3.2 Statistics

Another possibility is to ignore the time order of the values and view the sequence of interest as a set of numbers. Based on this idea, the statistical moments can be calculated, taking only the signal values and number of their occurrences into account. Due to the normalization, which removes the mean value from the series, this basic property cannot be used. However, the moments about the mean (or standardized moments) of different orders can be calculated. *Variance*, the second order moment, representing the expected square deviation from the mean is calculated according to:

$$\sigma^2 = E[(X - \mu)^2], \tag{3.2}$$

where X is the series, μ is the mean value $\mu = E[X]$, and E is the expected value operator.

Skewness, the third order moment about the mean, representing the skew of the distribution is calculated according to Matlab manual [14]:

$$y = \frac{E[(X - \mu)^3]}{\sigma^3}, \tag{3.3}$$

and *kurtosis*, representing further properties of the distribution (envelope shape), calculated as follows [14]:

$$k = \frac{E[(X - \mu)^4]}{\sigma^4}. \tag{3.4}$$

The Matlab native functions were used to calculated the values of these features. A well specified distribution is assumed during the calculation of these features, which requires a relatively high number of samples. Apparently, the higher order moment is used, describing finer properties of the distribution, the higher number of samples is neeeded. It is therefore very probable that only the variance will give reasonable results for short signal segments, while the result of the higher moments will be closer to random, with no actual information about the distribution.

3.3 Frequency spectrum

The frequency spectrum is a common means used for examination of time series, widely used in biomedical signal processing. Several frequency bands are specified in the human

brain, which can be used for determination of current functional state of this organ. It is therefore very sensible to research the spectral properties with regard to Parkinsonian tremor. Moreover, the spectral properties have already been used for tremor detection and they were proved successful for specific conditions [16]. This study can compare the detection power of the spectral properties to other possible approaches and provide more information about their actual significance.

3.3.1 Spectral properties of the STN LFP signals

Before the frequency based detection a note about the observations in spectral domain on the data, together with remarks of other researchers on this topic will be presented.

Previous research on the spectral properties of the LFP signals

An ongoing research of the signals of Parkinsonian patients' basal ganglia is aimed to different symptoms of the Parkinson disease. Among the numerous publications on the BG activity, only little literature is available for Parkinsonian tremor.

The spectrum of the activity of the Basal Ganglia can be classified into 3 groups: <8 Hz, 8–30 Hz and >60 Hz, according to recent findings [1]. The best characterised group of oscillations is in the range 8–30 Hz, reported to be of an *akinetic* character, i.e. inversely related to motor activity. Some studies[23] discuss a synchronous activity in this frequency range, commonly referred to as *beta activity* as a cause of bradykinesia and akinesia—other symptoms of the Parkinson disease. Also, this range is reportedly strongly modulated by voluntary movement activity and medication[11]. According to other studies [17] this frequency range appears not to be related to tremor.

As for the high frequency range, some researchers report dependency of activity in this range and response to levodopa [1] and the connection with the motor symptoms remain unclear.

The activity at low frequencies around 4 Hz have been reported to be correlated with the occurrence of tremor in the intraoperative microelectrode recordings. However it has been found not to be a strong feature of the LFP signals [1]. The authors suggest that capturing of these low frequencies may be affected by the asynchrounous character of the BG neurons' activity. The correlation between the tremor frequency in the globulus pallidus—another DBS target in PD patients—has been also reported [10], although highly position-dependent.

As for the tremor detection, the frequency spectrum was used for this purpose by K. Warwick et al. [16]

Summarized, the oscillations in the activity of the BG neurons seem to be little understood and although the scientfic evidence suggests connection between specific effects found in the signal spectrum and specific symptoms of the Parkinsonian disease, no clear decision can be made about the occurrence of tremor related features in the BG LFP signals.

Spectral properties observed on the dataset

An analysis was done on the dataset, comparing frequency spectra of tremor and non-tremor data. Regarding the frequency ranges specified in the previous section, no significant activity was observed in the high frequency range >60 Hz, as well as in the top of the beta range, approximately above 20 Hz. Therefore, the further examination was focused on the range below 20 Hz. The examination was undertaken on the spectra of the individual LFP channels and sections, which can also be plotted by the function `parkinsonian_plot_data.m`.

Normalized spectra, averaged over the sections of both tremor states are shown in Figure 3.1. First, the spectra were calculated for each section, the sum was normalized for the frequencies under 20 Hz. Then, the result was avaraged with its respective class. The spectrum was produced by the function `parkinsonian_plot_avgspectra.m`, which also implements plotting of multiple sections' spectra in a 3D plot.

The most remarkable notice was that a significant peak around 4.5 Hz was present in most of the tremor recordings. The central frequency of this narrow (less than 1 Hz wide) peak varied from about 3.5 to 5.5 Hz among the different patients and correlated well with the frequency found in the corresponding EMG recording. Moreover, peaks at multiples of this frequency (*tremor frequency*) were observed in some of the files. This could be interpreted as harmonics of the tremor frequency. The multiple sharp peaks in this frequency range can also be seen in Figure 3.1. Not all the patients' LFP signals exhibited this peak, but it can be seen as a predominant factor in the tremor recordings. Moreover, this band was not present in any of the non-tremor signals. In patients with more LFP channels avaiable, the peak was usually dominant in one of the channels, while lower or missing in the other.

Further, some activity has been observed around 15 Hz in both tremor and non-tremor files, which makes it a weak feature with regard to the tremor occurrence. Generally

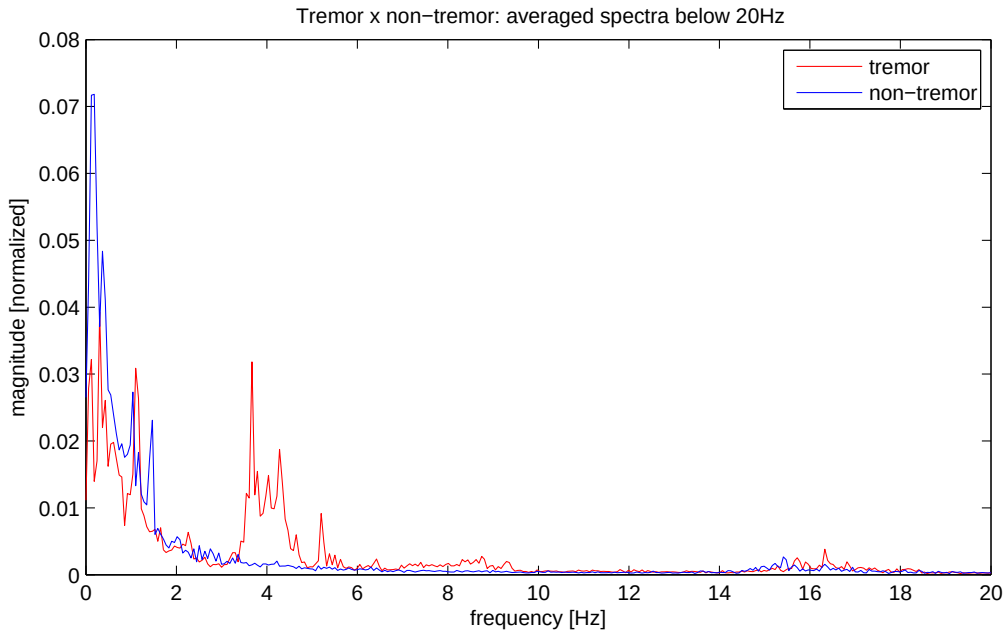

Figure 3.1: The averaged normalized spectra for tremor and non-tremor data.

speaking, the non-tremor recordings had rather flat spectrum with a high peak below 1 Hz, which was probably caused by the normalization, combined with the lack of activity at other frequencies. The 0–1 Hz and primarily the tremor frequency were observed as the strongest features with respect to tremor.

3.3.2 Energy in spectral bands

The first way used to utilize the spectral properties of the signal, was calculating the discrete Fourier transform of the signal and computation of the energy in selected spectral bands. It was important to decide, how wide should range of the frequencies be and how narrow bands should be used as single features. As discussed in the previous paragraphs, the main information content was observed in the frequency range 0–20 Hz. However low this may seem, the traditional EEG does not usually work with frequencies over the beta range (12–30 Hz). In accordance with the first observations of the spectral properties of the data, only the range 0–20 Hz was included to the feature set.

Based on the information above, the Fourier spectrum was divided into the following bands and respective features:

Table 3.1: Frequency bounds of the spectral features

feature name	**frequency [Hz]**
specbands1	0–1
specbands2	1–2
specbands3	2–3,5
specbands4	3,5–5,5
specbands5	5,5–7
specbands6	7–10
specbands7	10–12
specbands8	12–16
specbands9	16–20

Special attention was paid to the tremor frequency range, as well as to the low frequency ranges. This was the reason for the uneven widths and precision of different bands.

In order to make the comparison of energy in different bands possible, the spectrum was normalized to sum to 1 on the 0–20 Hz range. Moreover, the sum of the energy in the bands was divided by the band width to provide a width-independent measure.

3.3.3 Other spectral properties

Apart from the comparison of energy in different frequency regions, the spectrum can also be viewed as a whole and a specific tendency can be searched.

Repetition in spectrum

During the analysis of the data, a significant repetitive character was observed, appearing as harmonics of the tremor frequency (e.g. peaks at around 4, 8, 12 Hz). On the other hand, no such property was observed on the non-tremor data. This lead to an idea of creating a feature that would extract the repetitive property of the spectrum. The following was conducted:

Fourier transform can be used when searching for repetitive components in a time

series. An analogous approach could perhaps be used to extract the repetitive information from the spectrum. This concept is used in the speech processing and is called *cepstrum*, calculated according to the following formula:

$$C = |\mathscr{F}\{log_{10}\mathscr{F}\{x\}\}|\,, \tag{3.5}$$

where x denotes the original signal and $\mathscr{F}$ the Fourier transform.

To provide an output that would be easy to evaluate, Shannon entropy was calculated from the cepstrum in the next step. Presumably, a signal with no repetitive character in the spectrum will have flat cepstrum, which will then lead to a lower entropy.

This method is relatively complex and would require high computational power to be done in a real-time application. Inclusion of this feature to the feature set should be seen as an attempt to unveil other possible properties of the tremor problem. These could be later developed in another implementation of a simpler method, based on the same signal properties.

The feature is implemented under the name *specrep* in the `specrep.m` function. The implemented function `sigentropy`, used for the calculation for Shannon entropy is described in 3.4.

Character of the frequency spectrum

As seen in Figure 3.1, the non-tremor recordings exhibit rather flat spectra in average. On the contrary, more peaks are visible in most tremor recordings. Based on a similar idea as the previously described specrep feature, this characteristic can be extracted by calculation of the entropy of the spectrum. The calculation is done only for spectrum under 20 Hz, and is implemented in the function `specentropy.m`.

3.4 Information theory

Another possible way of gaining information from the time series is by the use of information theory. One feature, the Shannon Entropy, was implemented, calculating the expected information contents of the message, according to [18]:

$$H = -\sum_{i=1:N}^{N} p(x_i) log_b p(x_i) \tag{3.6}$$

The problem, faced when applying this theory to waveforms is that the idea is suited for a sequence of symbols, rather than a discrete waveform. Applied to short portions of a LFP signal, counting ca 500 samples (2 s chunk), each value is with high probability observed only once. This is based on the high number precision used and leads to the same value of entropy for any chunk. To prevent this the series was quantized to evenly spaced intervals prior to calculation of entropy.

The number of quantization steps was experimentally set to 40, which provided the best distribution of the output values on the 2 s chunks. The base of the algorithm in (Equation 3.6) was selected as 2, p_i was calculated from the number of occurences within the chunk.

Both steps are implemented in the function `sigentropy.m`.

3.5 Sample average

According to findings, presented in Section 3.3.1, a significant peak in frequency band 3.5–5.5 Hz can be seen in most tremor recordings' spectra. At the same point in this text, correspondence between this energy in STN and EMG signals is also described. The idea of sample average presumed, that the STN signal exhibited specific features within single tremor cycles. In other words: a specific tremor waveform or pattern could be discovered. This would be achieved by averaging the signal over single tremor cycles, which—given a sufficient number of cycles—might result in removal of all the tremor non-related activity and production of a representative pattern. This could then be used for tremor detection. The idea of sample average was successfully applied in EMG onset detection [8]. It is also a basic of brain evoked potentials analysis of the EEG [3]. In both applications the synchronizing signal is external, given at a moment specified by the examiner, which makes the starting point better specified.

The following part of the text describes the research carried out on this matter that was later considered unsuccessful and was discontinued. This parameter is *not included* in the final parameter set.

3.5.1 Overview of the Sample average detection method

To create the average waveform, identification of the individual tremor cycles had to be done first. Originally, the corresponding EMG signal was used for this purpose, owing to the easily detectible tremor cycles it contained. Once the high occurrence of tremor frequency was found also in the STN signals, the detection was modified and based directly on the STN signal. This solution was thought to be more reliable and straightforward.

3.5.2 Sectioning process

The detection is carried out by searching for a peak in the Fourier spectrum within the tremor frequency range specified above. This detected peak or more precisely its position and complex value were used to create a new spectrum with the length of the original one, filled with zeros. The detected peak, together with its complex conjugate on a symmetrical position in the vector, formed the only non-zero elements of the created spectrum. The sequence was transferred back to time domain, using inverse Fourier transform. This way a sine signal was created, specifying tremor cycles in the original STN time-series (red line in Figure 3.2). Taking a closer look at the signal in the upper plot of the figure, we can see the tremor period varies over time slightly in the STN signal, which might cause inaccuracies in the subsequent averaging process. To avoid this, the signal was smoothened with an equally-weighted moving average filter of order 10 (dark blue line). Then, the peaks were found in this smoothed signal around the maxima of the sine wave. The resulting positions were used as reference points to create signal sections. Even though the sections produced contained in most cases useful signal, transient effects or other non-standard behavior can be found in most signals. These would cause undesired distortion, if added to the average. Outlier sections can be detected by its length, which significantly differs from average one. This is caused by a failure of the peak detection on a non-standard signal. The cycle-searching method is implemented in the function `Parkinsonian findtemcycle.m`.

3.5.3 Averaging process

The averaging process starts with the first selection, determined by the first peak in the calculated sine function and detected period length, which is taken as the first average.

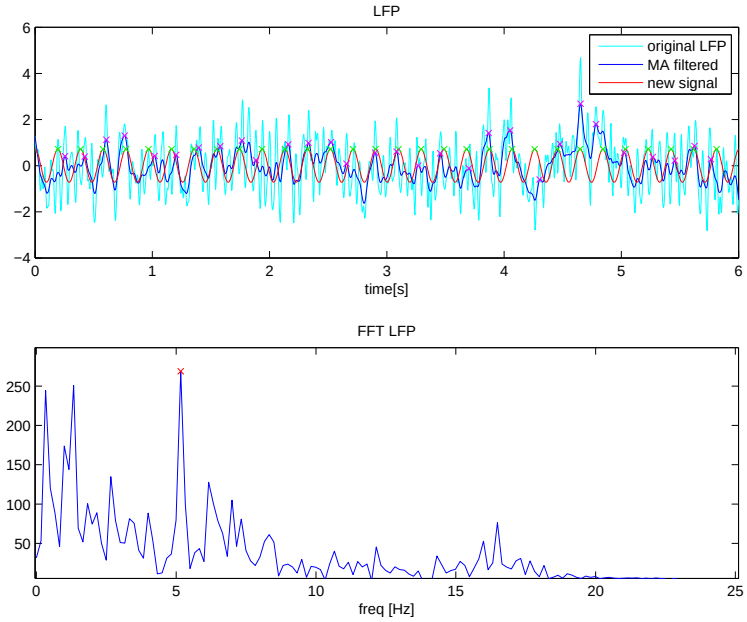

Figure 3.2: Example of cycle detection in LFP signal.

In each step, the start of next section is searched around the position of next expected period start. The signal is compared to current sample average using vector distance (Equation 3.7) and position is found with maximum value of this measure. If the maximum exceeds a taken threshold, section is added to the average, starting at the position where the maximum match was recorded, otherwise is thrown as an outlier. The process continues until the end of the signal. Inlier, outlier and average sections, calculated for 2 different patient's LFP signals are shown in Figure 3.3.

$$d = \frac{a \cdot b}{|a||b|} \tag{3.7}$$

Weakness of this method is selecting the first section as a temporary average. This may lead to undesired results in cases, where the first section is formed by an outlier. For the testing purposes, the signal was always chosen to start with a regular tremor cycle.

The averaging process is implemented in the function `parkinsonian_averagecycle.m`.

3.5.4 Detection

The detection was carried out by sliding the sample average (or pattern) over the signal and calculating the vector distance (Equation 3.7). For successful use of this similarity

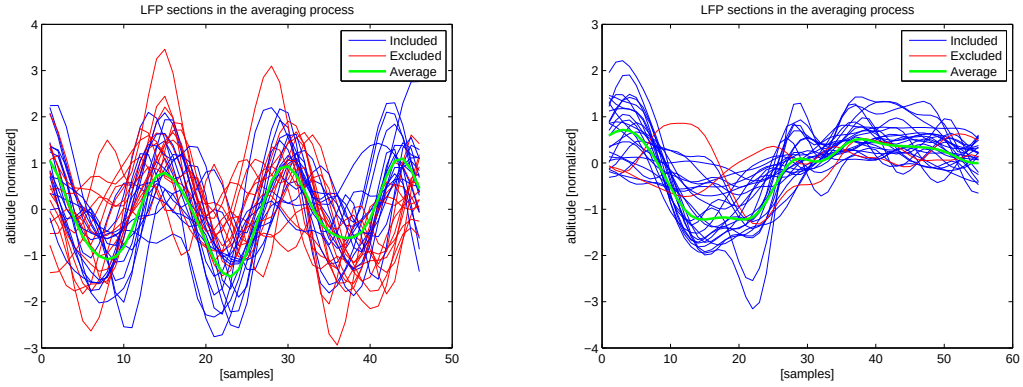

Figure 3.3: Averaging carried out on 2 different LFP sequences.

measure, it is necessary to have additive value removed from both signal sections, as this affects the result significantly. The mean value was therefore removed prior to calculation of the vector distance. It is the same method, as the one was used in the averaging process for measuring section distance. The result of this sliding process is in ideal case an alternating waveform, having maxima at every start of a tremor period, symbolizing a maximal match. Values of the detection output between these peaks are lower, which comes from the nature of the method. For detection, only the peaks should be taken into account, as they symbolize the real match with the pattern. A low peak amplitude or no visible peaks symbolize a low match with the pattern, while high values and great peak-to-peak difference symbolize a good match.

3.5.5 Experimental results, discussion

The detection was done on tremor onset data with special attention to peak values of the detection function output. First of all, the pattern was calculated for the tremor part of the signal to be used for detection on the whole file. The testing file contained non-tremor, tremor onset and tremor part. It was found that the value of the detection function exhibits unrecognizably similar behavior in both resting and onset part of the data (Figure 3.4). This was most probably caused by strong higher frequency content in the signal, represented basically by a signal around 16 Hz. To improve the result, this frequency could be removed, but removing all frequencies above, say, 12 Hz would cause a crucial loss of shape information and the most of the information would be lost.

A decision was done not to continue the sample average research, as no promising

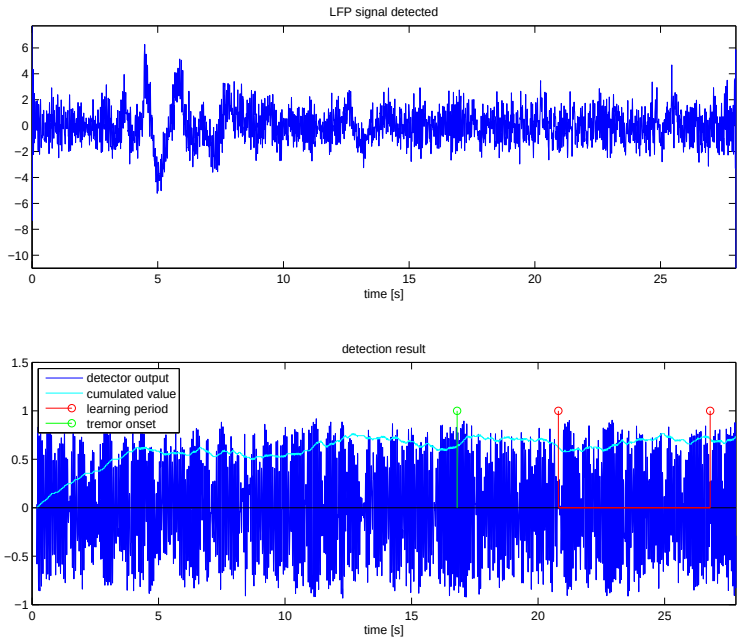

Figure 3.4: Detection with sample average on tremor onset signal.

properties can be seen in the data. Moreover, further development of the averaging function would require great effort, while no great improvement in the results was to be expected. Anyway, the research done can not be considered thorough enough to reject the existence of a—no matter if patient-dependent—tremor pattern. Considering the results, better outcome would probably be achieved using wavelet transform which provides means for this type of analysis. Neverheless, the wavelet transform was not included in the feature set, as it is being researched thoroughly by others, working with the same data.

3.6 Autocorrelation function based features

Examining the signal of interest, the autocorrelation function can provide new point of view and unveil specific properties of the series. Despite its relatively high computational complexity for longer signals and calculation in the full range, the use of autocorrelation in real-time application can be considered thanks to the constraints described further.

As seen in the frequency spectra of the original signals across the dataset, there are

specific components with high energy, that also affect the shape of the autocorrelation function. Several courses of the autocorrelation function, selected as representative for the dataset are shown in Figure 3.5.

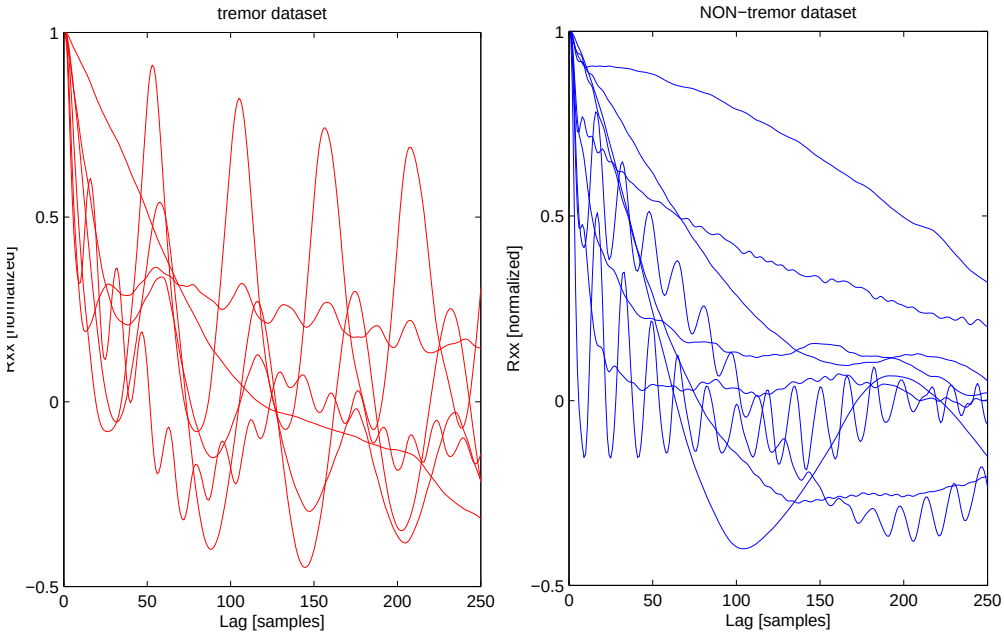

Figure 3.5: Selected examples of autocorrelation function courses for tremor and non-tremor data.

Apart from the shapes common for both groups, which can be considered outliers (the number of these courses formed a reasonably small part of the dataset), a specific tendency can be noted on the tremor data: For most cases the autocorrelation function exhibits a peak at the lag in the region between 45–75 samples, while the values are lower in the region 10–40 Hz. The former range points to a significant periodic character of the input sequence at the equivalent frequencies ca 3.3–5.5 Hz, the latter ca 6–25 Hz, calculated for 250 Hz sampling frequency. It seems that the tremor frequency based periodicity is dominant in tremor recordings, while not present in any sample from the non-tremor set. Based on these facts, a feature can be implemented, utilizing the observed character of the autocorrelation function.

The short-period cycles, apparent in both datasets in the figure, correspond to the frequency of ca 16 Hz. Higher energy around this frequency was observed in both datasets, as described in the Section 3.3.1. Although these oscillations may cause difficulties in the detection process, filtering out this frequency band would cause great harm to the

original signal and was therefore not implemented. Low pas filtering at less than 15 Hz would cause a high loss of the autocorrelation function precision by removing the higher frequency information. On the other hand, the band-pass filtering would require a narrow filter, which would lead either to a long lag or a phase shift, distorting the autocorrelation function as well. These facts were taken into account during the implementation of the two following autocorrelation-function based features.

Two different implementations were done, producing two features: *acpeaks* and *acratio*. Both of them utilize the same properties of the signal and both were included in the final feature set, as their evaluation is better done by the feature selection algorithms.

3.6.1 Implementation

The first implementation, *acpeaks*, is based on finding peaks in the autocorrelation function in the lag region 45–73 samples (the values were taken from the observation of the autocorrelation function on the data). The peak value is divided by the average peak value in the lower lag region. This way, the output value of the feature for a given signal segment is produced. Summarized, a lower number indicates rather non-tremor behaviour, while a higher value means the opposite.

The problem of this first implementation is its dependency on the peak detection, which might not be a robust and flexible solution. Thus, the autocorrelation function can probably provide detection information even in cases where no peaks can be found, but a specific character is present. In this function, the output value is set to one when some peak is found in the *tremor* region and no peaks can be detected for lower lags, indicating a tremor behaviour. On the contrary, zero is the output for the cases when no peak can be detected in the tremor region. Apparently, this is not an optimal solution and the method should be generalized to provide more robust results. This was attempted in the next implementation: *acratio*.

In the second implementation, no peaks were searched in the autocorrelation function. The detection is based on the maximum of the function in the region of interest, divided by the average value in a lower, neighboring region, as represented by the following equation:

$$r_{ACR} = \frac{max(R_t)}{mean(R_l)}\,, \tag{3.8}$$

where R_t is a vector of the values of the Autocorrelation function in the determined

tremor region (45–73 samples) and R_l contains values of the autocorrelation function in a neighboring lower region (28–44 samples). The r_{ACR} element is directly the output of the method.

This second implementation should provide more robust detection, however, the solution presented here serves as a proposal for a future implementation, based on a wider dataset.

An important note can be done, regarding the target application, which works in real time; the maximum lag, that has to be calculated in the autocorrelation function is 73 for both feature implementations, which makes the calculation reasonably faster.

Both implementations were tested on the whole original sections and a simple classification was done, setting a threshold for the function output to determine the tremor state. Both methods were very successful, correctly classifying up to 80% of the sections. The more specific acpeaks implementation reached better results and both were included in the feature set, to be evaluated using unbiased methods. The result of this simple test is very general and the actual success of these features can be evaluated only from the classification results of the experiment. However, the high success means some significance to the detection task exists.

The features are implemented in the functions:

`parkinsonian_feature_acpeaks.m`

`parkinsonian_feature_acratio.m`

3.7 Fractal dimension

Another approach to the STN signals is considering them to be a chaotic output of a nonlinear dynamic system. In other words, we do not treat the signal of interest as completely random but we expect an underlying structure, carrying information about current brain function. Research projects done on his topic [7] show, that for many biological processes, such as neural tissues or vascular system, the chaos theory can provide new level of understanding or assist at specifying properties of the system.

Among other complex variables, describing nature of a chaotic system, the *fractal dimension* can provide a basic measure of the system chaoticness by calculating a non-integral dimension of the assumed fractal. The deviation of the calculated dimension and that of a space-filling (conventional) object can then provide basic idea of the system

nature in terms of chaos theory. However, usage of fractal dimension estimates is relatively common in processing of other biomedical signals, such as ECG (Hearth Rate Variability [13]). The iterative algorithms described below calculate estimates of fractal dimension.

Different fractal character was found for example in patients during epilepsy seizure [7], which can be—with is temporary and brain disorder related character—considered similar to Parkinsonian tremor.

3.7.1 Box-counting dimension

The box-counting dimension (also called Minkowski–Bouligand dimension or Kolmogorov capacity) is the first of the two fractal methods implemented. It is based on iterative splitting the object (signal in this case) to boxes of given size. In each iteration , the size of the box is changed exponentially and the input object is all split into boxes with side the current length. Then, the number of boxes containing the object and total number of boxes in the particular step are counted and stored. The algorithm goes from one box, covering the whole data, to boxes of unit size, containing one data point each. When boxes are counted for the whole range of selected side lengths, local dimensions can be calculated according to:

$$D_c \approx \frac{\triangle\ log_2(N(r))}{\triangle\ log_2(r)}\,, \tag{3.9}$$

where r is box size and $N(r)$ number of boxes for the respective box size. When calculating the overall fractal dimensuion estimate, additional conditions have to be taken into account to utilize the part of boxcount series, for which the result is senzible [9]. The choice of logartithm base as 2 is due to the selected step of box sizes [9].

In one-dimensional case, the algorithm searches for boxes covering all non-zero elements of the input sequence, while for two-dimensional case boxes covering the area of the object are counted. The researched signal is a sequence of numbers and has to be treated as 2D object. The algorithm was adapted for waveforms, assuming the function to be continuous and therefore crossing every value between maximum and minimum on the specified interval. To make the algorithm work properly for waveforms, the input sequence is first normalized into unit square. The overall dimension is estimated in an interval, meeting the conditions mentioned above.

The method is implemented as `boxcountwf.m`.

3.7.2 Sevcik fractal dimension estimate

Another estimate for fractal dimension was implemented, based on work by Carlos Sevcik [19]. It is intended specially for waveforms and therefore no adaptations are necessary. The algorithm consists of mapping the input sequence into a unit square, calculating the length of the curve and calculating the dimension as follows:

$$D = 1 + \frac{ln(L)}{ln(2 \cdot (N-1))}, \tag{3.10}$$

where L is the calculated length of the curve in normalized coordinates (unit square) and N is number of samples of the input discrete waveform.

The curve length is calculated as Euclidean distance between consecutive points in the input sequence.

The method is implemented as `fractdim.m`

Both implemented methods were tested for different lengths of LFP signals with stable results, which differ between both function within reasonable range. First tests of distribution of fractal dimension done on the whole dataset show slightly higher fractal dimension number for tremor signal.

3.8 Chapter summary

The implementation of all the features in the feature set was presented in this chapter. These features will be further tested in the next chapter, using the experimental data, and the results will be discussed. This way, one or more suitable approaches to the tremor detection can be found, based on classification a test on the data. A complete list of the implemented features follows.

3.8.1 List of the implemented features

Temporal

- *power*: Signal power
- *zerocross*: Zero-crossing rate
- *avgdiff*: Average of the first differential

- *maxdiff*: Maximum of the first differential

Statistics

- *variance*
- *skewness*
- *kurtosis*

Spectral

- *specbands1–9*: Energy in the spectral bands
- *specrep*: Entropy in spectrum
- *specentropy*: Entropy in cepstrum

Information Theory

- *entropy*: Information entropy of the signal

Auto-correlation

- *acpeaks*: Comparison of peaks in the autocorrelation function
- *acratio*: Comparison of values of the autocorrelation function

Fractal dimension

- *fractdim*: Waveform fractal dimension estimate
- *boxcount*: Box-counting fractal dimension estimate

Chapter 4

Classification

In terms of this study, the classification process serves as an assessment tool for evaluation of the designed features. As described in Chapter 3, the tremor-indicating of the STN signals are unknown. Thus, the discriminative properties of the implemented features are unknown either, unless a classification process is carried out.

This chapter describes the methods used for individual feature evaluation, feature set selection, classification and related tasks, needed for the classification process. Discussion of important aspects of the detection problem can be found at the beginning of the chapter.

4.1 General considerations

When searching for a suitable classification method for the onset problem, it is necessary to keep respect to the goal application: real-time detection of parkinsonian tremor in the patient's brain. Addressing the facts about the DBS process, there are several important issues that should be considered. This includes the following subsequently discussed facts:

- Tremor onset should be detected with the shortest lag possible
- Detection has to be done in real time and should be reasonably fast to provide sufficient time for the starting stimulation to take effect.
- STN signal can have specific properties for each patient. Classifier can therefore be trained as patient-independent or specifically for a given patient.

- Multiple signals are present for some of the patients: 3 pairs produced by 4 neighboring electrodes, 6 signals for bilateral implants. The optimal electrode selection is a priori unknown.

4.1.1 Onset detection

Going through the dataset characteristics, shown in Table 2.1, a significant lack of tremor onset recordings is apparent in the dataset. It seems reasonable to use short sections of the input signal, similar to those, that would be used in a real application. These sections would be used for feature evaluation and determination of the differences between tremor and non-tremor periods in the signal. Due to the impossibility of determining an exact moment of tremor onset (see Section 2.3), the onset sections can not be used in this approach. Instead, determination of suitable features can be done, that can later be developed to an onset detection algorithm on suitable data.

4.1.2 Real-time

To provide a fast system response, the feature set has to be calculated only from a specified number of last samples. This number should be reasonably low to capture the possible tremor onset as soon as possible. These classified sections can overlap, which would probably be the method of choice for practical application. For the training phase, described in this text, the use of non-overlapping sections was chosen in order to maintain independency of the input samples. As there is no need for fast response in the parameter assessment and connected classification, this approach should make no harm to the results.

The following method was chosen for the classification: the input data, formed by the marked sections of files with recordings of simultaneous channels that count usually tens of seconds, were cut into short series of just several seconds, providing data closer to those in real application. These sections are referred to as chunks in the following text, as well as in the Matlab source codes.

Using shorter time sections of the signal should capture fast changes in the signals that might be overseen if longer signals were used.

The difference between file, section and chunk is clarified in Figure 4.1.

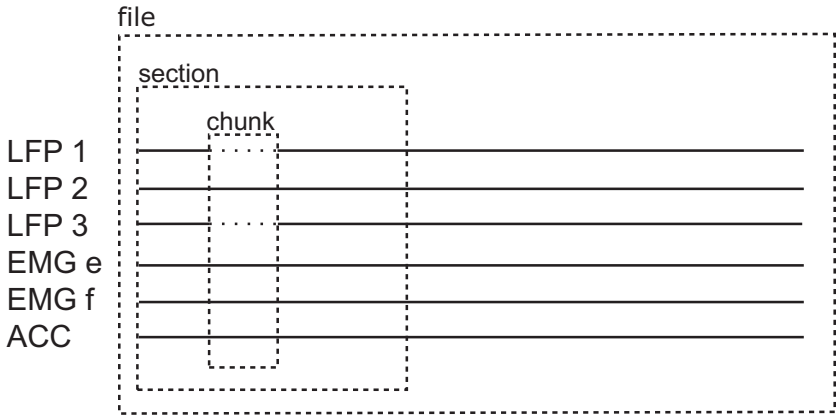

Figure 4.1: Clarification of the terms *section* and *chunk* as used in this text

4.1.3 In-section signal dependency

In spite of the advantages the assessment of shorter signal chunks brings, considerations have to be done to keep the classification results consistent. Dividing the data into subsets is often required during the classification process, e.g. in creating training and testing subsets. A higher level of consistency (or similarity) can be expected across chunks coming from one signal section (sibling chunks), compared to those coming from other sections. Furthermore, all the sibling chunks form just one input file in the dataset. Based on this, a decision was done not to separate the sibling chunks and keep them together when dividing the data.

This policy was kept especially during the classifier training, while the implementation would be very complicated for the feature selection process and was therefore not undertaken. Anyway, this should not make any great harm to the results, as the classification output is assessed in the end.

The automatic chunk creation is implemented in `parkinsonian_chunks_cut.m`

4.1.4 Patient-specific detection

One of the important general issues to consider is whether to attempt design of a universal classifier, suitable for any possible patient, or whether to try creating a patient-specific solution. The original idea at the beginning of this project was to create a classifier that would be as universal as possible, being trained on various patients in different situations. However, once the dataset sorting process was finished it turned out, that the quality of

the dataset is insufficient for such a general task. Moreover, this task would probably be too complex and complicated, given the low level of general preliminary knowledge on this topic and thus, just a smaller task of classification of signals from patients in resting state was chosen. Anyway, handling data coming from different patients is still an important issue in this task and deserves more study.

Recalling the aim of the work, the task is to detect starting tremor in the signals of patient's brain, which is clearly a patient specific task. On the other hand, the designed solution should be versatile, giving reasonable results for the widest possible range of patients. Small adjustments or personalisation would be no problem. Based on this, a feature set should be selected with respect to the highest possible versatility, while the classifier parameters should be trained specifically, for each patient separately. Alternatively, groups of patients could be identified together with specific feature sets. The classification would then consist of assigning the examined patient a previously specified group and therefore a feature subset. This approach would probably be optimal in case more different types of tremor STN activity were found and several patient groups could be identified. In this case the data is insufficient for such task (number of patients too low) and the patients involved are therefore viewed as one group. A general feature set is selected for all the patients and a subset of is is chosen further specifically for each patient.

Summing up the considerations done in the previous paragraph, the method selected consists of first selecting a set of significant features on the whol dataset. Then, a subset of features is selected for each patient specifically.

4.1.5 Presence of multiple channels

As already mentioned in Chapter 2.3, 3 or 6 input channels in some recordings, produced by the pairs of neighbouring electrodes. Making a comparison of the size of the STN and the difference between the electrodes' tips, it is clear, that only some of the electrode pairs contain information about the target brain structure. The other electrodes may be placed outside of the target structure in the brain and therefore be of no significance to the problem. In this study, being given a number of input signals, the selection of a signal with highest tremor significance is a priori unknown. However, the situation is not too different in the practical application of DBS. In real application, the electrode pair for stimulation is searched experimentally [21], even though more information about the exact electrode placement is available, provided by medical imaging techiques used

during the implantation.

To solve the problem of multiple available signals, it was assumed that one optimal signal exists for each patient, having the highest significance with respect to tremor detection. To find this channel, separability measures are calculated for each available channel and all features. Subsequently, the signal with the best rating is selected.

This problem becomes more complex for cases when the channel has to be selected for more patients with multiple available channels at the same time. The solution is described in the Section 4.2.4.

4.2 Description of the classification process

This section describes the classification process step by step, with references to the implemented functions.

To carry out the per-patient feature selection and classification, a PRTools pattern recognition toolbox for Matlab, developed by the researchers from TU Delft [6] was used. The methods implemented in this toolbox were combined with newly implemented functions for classification result assessment and other related tasks.

The chart in Figure 4.2 shows the whole classification process, as implemented in `classification.m`. The individual steps of the algorithm are described in the following sections.

4.2.1 Input to the classification algorithm

The input to the algorithm, as shown in Figure 4.2, is formed by marked sections of preprocessed data. All the files, carrying also information about the content in different sections, are all first downsampled, LFP signals low-pass filtered, normalized in amplitude and cut into sections, according to the file description. The whole preprocessing process is described in Chapter 2.

4.2.2 Cutting the sections into chunks

According to Section 4.1.2 earlier in this chapter, the classification is done on short sections of the input signals, called chunks. These are produced automatically from the pre-

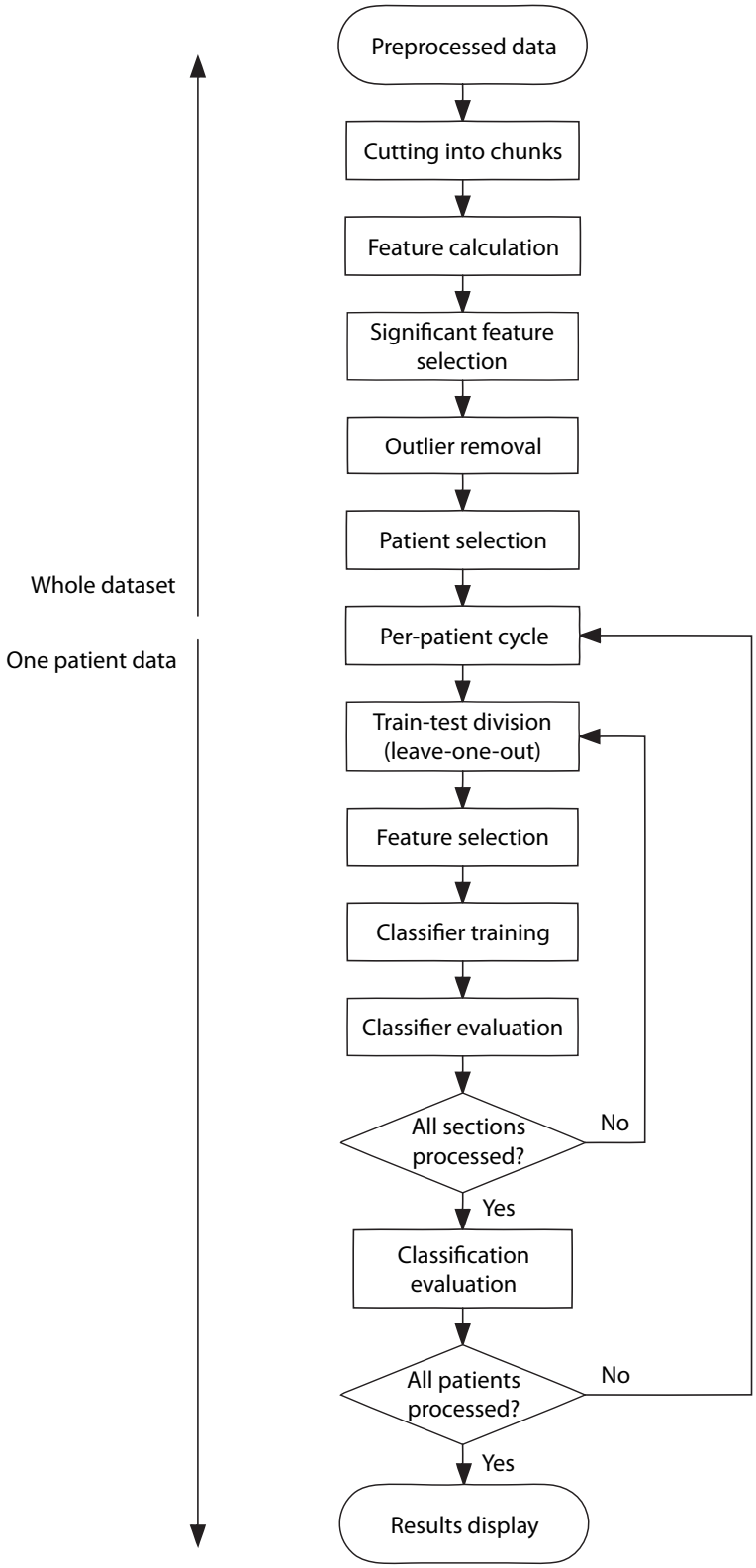

Figure 4.2: A flowchart showing the classification process.

processed signal sections using the implemented function `parkinsonian_chunks_cut.m`. The chunks are stored in a data structure, carrying information about the original patient, file, channel and position along with the signals. Every chunk contains only one LFP channel, which means more chunks are produced for one specific moment in patients, where more simultaneous channels are available. This is done due to an easier handling and assessment of the individual chunks and is taken into account in the consequent processing. Apart from the LFP channel, every chunk also carries corresponding EMG and accelerometer signals, if applicable. Even though this structure contains a lot of redundant information (e.g. the section information could be derived from the original file and position, as well as the patient can be derived from the file of origin), it turned out to be useful, avoiding references to multiple data structures.

As for the chunk length, the number of samples necessary for the calculation of the features has to be taken into account. This is true especially for the features based on spectral bands, where the number of samples affects the precision of the calculated spectrum. The number of samples used affects the accuracy of the result and should therefore not be too low. Apart from spectral based features, the statistical features—the higher moments especially—are reasonable only for higher numbers of samples that makes the approximated distribution function more precise. On the other hand, a problem of overseeing of some short-time effects arises for longer signal chunks.

Considering these aspects, multiple chunks lengths were tested and results compared. The main decision was then done between 1 s and 2 s chunks, as shorter chunks can not provide sufficient number of samples for feature calculation and, on the other hand, using longer chunks causes the dataset too small for some patients, making the classifier design impossible. Better classification results were achieved for chunks 2 s long, numbering 500 samples per chunk. This number seems reasonable to maintain sufficient accuracy of feature calculation, as well as fast response and was therefore chosen for the final evaluation. No chunk overlapping was used, as it would bring dependencies to the training set and make the assessment of the results more complicated.

4.2.3 Feature calculation

Having all the data cut into short segments—chunks—the set of observations is ready to be transferred from the time domain to the feature space. This is done by calculating the whole set of 23 features for each individual chunk. Calculated features are then transferred to a matrix, suitable for PRTools toolbox, as well as subsequent steps of the

classification process. In this matrix form, every observation is represented by one row of all the calculated features, while the corresponding class (tremor or non-tremor) is carried in a special vector. It is necessary to keep in mind, that there are important dependencies in the training set produced, formed by the parallel corresponding LFP channels and multiple chunks, produced from one section.

Normalization of all the calculated feature values was done as a part of the classification process, recalculating the feature distributions to zero mean and unit variance, which is desirable for an easier cross-feature evaluation and comparison.

The feature calculation and normalization process is done by the implemented functions:

`Parkinsonian_get_feature.m`

`Parkinsonian_feat_tomatrix.m`

`normfeatures.m`

4.2.4 Significant feature selection

The set of designed features,described in the previous chapter, was based on the knowledge of the dataset, as well as on existing solutions for other biomedical signals. However, the actual relevance of the features is not known, unless tested on real data. A number of tests has to be done on the individual features, that will unveil their relevance to tremor detection. This can be achieved by characterization of their separability capabilities by different measures. That is, how well can the respective feature distinguish between the two target classes. Not only should the feature value be significantly different for the tremor and non-tremor samples, but it should be consistent on each of the classes as well. The output of this process will be a subset of features, having a proven significance to the task. At later stage, this set can serve as a set of general descriptors for the tremor task, from which a small number of features can be selected for each individual patient.

In order to make the resulting feature set versatile, all the available files were included in the evaluation process. This condition should reduce the number of features necessary to calculate, as well as ease the consequent per-patient feature selection.

The feature selection method consists of two main steps:

1. Selection of one LFP channel per patient
2. Selection of the significant features on the resulting set

The implementation of these two steps is described in the following paragraphs.

Multiple-channel issue

One of the important issues, which had to be solved during the feature selection, was the already-mentioned presence of multiple channels in some patients (Section 4.1.5). It was necessary to select only one channel per patient to achieve a realistic evaluation of the individual features. Recalling the characteristics of the It is presumed, that the relevalce of some of the channels is significantly low, while higher for others. The solution to this problem was to base the channel selection upon separability measures, determined by suitable features. Obviously, the optimal channel selection should be indicated by the highest class determination capability of the features. At this point, another problem comes into account: The suitable features are yet to be selected. Moreover, there are more patients with multiple channels available and the selection becomes even less specified.

The proposed solution consists of calculation of feature separability measures for all possible channel selections on the whole set and then choosing the selection with the highest possible ranking. As the feature selection is still to be done, the evaluation is based on four features with the highest rank for each channel combination. The channel selection with the highest average value of the 4 best features is then selected. The selected comparison criterion was the AUC, as it provided easy-to-interpret evaluation measure. For a better understanding of the channel selection process, the Figure 4.3 shows channel assessment values for all possible channel selections. Maximum value for each series is marked with a red cross. Except for the worst feature measure, that plays mostly just illustrative role in this figure, the other three measures give relatively stable results. The average of the best 4 features was selected as the representative value for the channel set and the set number 110 was selected. The apparent periodicity of the results is caused by the channel-set generation method[1]. The total number of channel selections is given by the multiple of all the possible channels for each patient (see Table 2.1), which gives a total of 324 possibilities.

The channel selection process is implemented in the following functions:

`parkinsonian_select_patchans.m`

`parkinsonian_get_patlfpchannums.m`

`generateposvectors.m`

`featauc.m`.

[1] Only some of the neighboring samples differ in the selection of one channel.

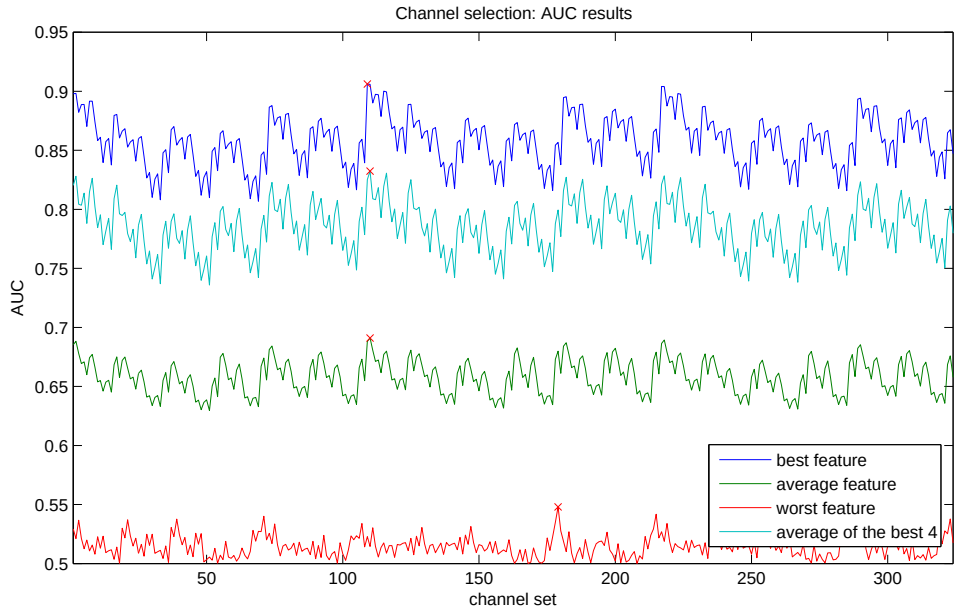

Figure 4.3: The results of the AUC measure over all possible channel selections. The connecting lines serve shown just for improved readability of the figure.

Statistical comparison methods

The task of selection of suitable feature features is complex and has no ideal solution. Statistical assessment of the feature values is first of the possible approaches. It is based on a presumption that probability density function (PDF) of a good feature can be viewed as a combination of multiple PDFs, one per each target class. In case these distributions overlap completely, the feature has no separating power. Assuming the pdfs to be a normal (Gaussian) distribution, basic parameters—mean and variance—can be tested for difference on the two classes. Several statistical methods exist to verify this property with a specific confidence level. These methods are based on statistical hypothesis testing and should provide relevant answer to the question of the parameter difference.

First, the paired t-Test was intended to be used. It is based on testing the hypothesis that the feature values come from two normal distributions with different mean, one for each target class. This can also be viewed as testing a difference distribution, produced by subtraction of the two hypothetical distributions, for zero mean. Based on this test, the hypothesis can be either verified or rejected at a selected significance level.

The main problem of the paired t-Test is that the two compared classes must be of

the same cardinality, i.e. the same number of instances. This problem is very serious and blocks use of this method for other cases. In this case, the tremor dataset contains unequal number of sections per class and the t-Test can therefore not be used. Moreover, this approach to statistical testing in classification application was reported to be unsuitable [4]

The solution was found in another method—A Mann-Whitney-Wilcoxon test—which is a non-parametric alternative to the paired t-Test and enables calculation for different set cardinalities [20]. Generally speaking, the nonparametric methods are not based on the assumption of normal distributions and are therefore more suitable for smaller datasets, where this assumption is problematic. The Wilcoxon test compares ranks of the data (order of values in this case) and tests the hypothesis that the data from the two classes have equal means against the hypothesis that the means are different. The output of this test is a p-value, providing a confidence measure of the comparison. The Matlab implementation of this method, `ranksum.m` was used for the value calculation.

The last statistical method implemented was Fisher discriminant ratio, comparing mean values and variances over the two classess, according to [22]:

$$FDR = \frac{\mu_1 - \mu_2}{\sigma_1^2 + \sigma_2^2} \tag{4.1}$$

This feature was implemented in function `fdr.m`, but was excluded from the feature assessment due to disputable assessment of the output values for small sets. Eventually, the Wilcoxon rank-sum test remained as the only statistical method involved.

Other comparison methods

Apart from the discussed statistical comparison methods, which provide information about the relevance of the particular features, the AUC measure was implemented to assess the actual classification strength. It is based on the calculation of the Reciever operating characteristic (ROC curve), which compares the true positive rate (sensitivity) to the false positive rate (1 - specificity). The ROC curve is a common measure of the class discrimination properties[22], which can be evaluated from the area below the curve – AUC. This value will converge to 0.5 for a feature with low discrimination capability, while on the contrary, the value will be closer to one for a good feature. The parameter is calculated using the implemented function featauc.m.

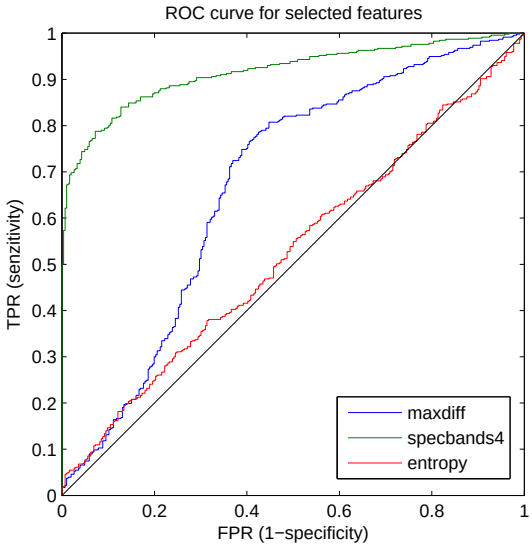

Figure 4.4: Calculated ROC curves for 3 selected features with different discrimination abilities.

Combined feature evaluation

To gain an overall view of the individual feature performance, a combination of both approaches was chosen. The adopted significance criterion requires specific level of confidence at the Wilcoxon test (upper bound of the p-value) and minimum value of the AUC at the zame time. This ensures the feature is both significant and has exhibits practical separability properties as well. The actual values, together with the feature selection results are presented in the next section.

The function `parkinsonian_feature_assess.m` was implemented to assist feature assessment measures calculation for multiple features.

4.2.5 Outlier removal

The position of outlier removal in the algorithm may be considered disputable, as the position before feature normalization would be expected. Nevertheless, to classify a sample as an outlier, it is necessary to compare its value (i.e. the value of the respective feature vector) to the rest of the dataset. At this point, a needles loss of samples could be caused in case some of the features, used for outlier identification, had no significance to the problem.

On the other hand, the position before the patient reduction (next subsection) serves

to gain a higher level of generalization, so that the individual observation can be compared with a more general set than just the reduced set of the patients, suitable for classification.

The outlier removal is carried out by a comparison of the feature values on individual features from the significant feature subset. Then, the value of each sample is compared to the distribution of the rest of the data, while single distribution per each class is presumed. The difference from mean is expressed as a multiple of standard deviation of the particular class. To maintain sufficient consistency of the set, a criterion has to be adopted to exclude unusual observations with extreme values. These properties, indicating extraordinary conditions during the recording, which could negatively affect the classification process. It seems reasonable to compare not only a value of one feature, but to employ multiple features. The actual threshold and group size chosen is presented along with the outlier numbers in Section 4.3.1.

The outlier removal process is done by the function `parkinsonian_filter_outliers.m`.

4.2.6 Patient selection

This step serves to set aside data of the patients that lack data in one of the two target group and are therefore unsuitable for the classification process. This is done with the help of the function `parkinsonian_get_patnumsections.m`.

The need of at least two sections in each class arises from the classification process, described further: when one of the sections is left out for classifier testing, at least one section of the same class must remain in the training dataset.

4.2.7 Patient cycle

This section of the chart in Figure 4.2 refers to a cycle, where classifier is trained and evaluated for each individual patient. Due to the lack of data (chunks coming from the same section are considered dependent and are kept together – see Section 4.1.3), the leave-one-out method was selected for the classifier evaluation.

4.2.8 Leave-one-out cycle

The method consists of a main cycle, in which the data is divided into a testing set of one section, and the rest, providing the training data. The classifier is trained on the training set and the classification of the testing sample is recorded. All possible selections of the sets are tested iteratively and the classifier error is calculated according to the gathered information. This method should give a good error estimate and is suitable especially for small datasets, due to the high computational complexity (high number of folds) [22]

In terms of this study, one sample is represented by all observations coming from the same section. As there are very few sections for each patient, this approach seems to be very suitable with a maximum actual number of 8 iterations per patient (Patient 1). A special implementation of the classifier evaluation is needed, due to the presence of multiple samples in each section. In this case, the number of erroneously classified samples was calculated for each algorithm fold and the overall error was weighted by the total chunk number. Together with the class information about the testing sample, this information can serve to calculation of all the main error measures.

4.2.9 Feature selection

The selection of a general set of features, suitable for the tremor detection, was selected as described in Section 4.2.4. This selection was based on the whole dataset, which should provide the output set with higher versatility and applicability for other possible patients, not covered by the data. This selection does not reflect the dependencies within the set, as only individual features were assessed. And there are obvious dependencies in the test, for example the features *acpeaks* and *acratio* describe the same property of the signal, only with little different means. To carry out the classification successfully, a selection of features with little dependencies is desired. Moreover, the number of features in the significant feature set is relatively high (14 features), while only few of them should be sufficient to carry out the detection for the particular patient. Summarized, the described aspects lead to an effort to select a low number of highly significant features with low level of a reciprocal dependency, suited for the given patient specifically.

Many methods exist for this purpose, based on different approaches and assessment methods. In this work, the algorithm of sequential backward feature selection was adopted, which is implemented in the PRTools toolbox under the function featselb.m [6].

The backward algorithm starts with the full feature set. Firstly, value of a selected criterion is calculated for the entire set. In the next step, all possible subsets with one removed feature are produced and the criterion calculated. A subset with the best ranking is selected and the process continues until a desired number of features is achieved [22].

To proceed with this algorithm, two decisions have to be made and the following has to be selected:

- Assessment criterion
- Feature number

A great majority of feature set assessment methods the PRTools toolbox offers was tried and the sum of estimated Mahalanobis distances was selected, according to the classification results achieved. The Mahalanobis distance takes into account correspondences in the vector elements through the use of covariance matrix and gives therefore more realistic results for multiple feature estimation than would the Euclidean distance in the same place.

The optimum number of features should be possible to find according to the recorded results of the feature-set assessment, done during the backward selection algorithm. The value of the calculated criterion for a group of specified size is shown in Figure 4.5. The optimal feature number should be rather small (Occam's razor), while its rank reasonably high. The number should therefore be determined at the end of the steep part of the graph in the figure. Apparently, the optimum point varies among different patients. Based on the rules described, the optimum should be about 3 for patients 1 and 9, 5 for patient 3 and about 10 for patients 5 and 10. Nevertheless, after inspection of the classification results for different settings of the feature selection algorithm, the number of 5 parameters together with Mahalanobis criterion were evaluated as the best combination. This settings produced the best overall results.

To achieve an unbiased method evaluation, the feature selection is included in the leave-one-out cycle, being done only on the training data in each step.

4.2.10 Classifier training

Once the set of features is selected, classifier training begins on the training samples. Multiple classifier types were included into the process, listed as follows:

- Parzen classifier—a classifier based on Parzen window estimation [5]

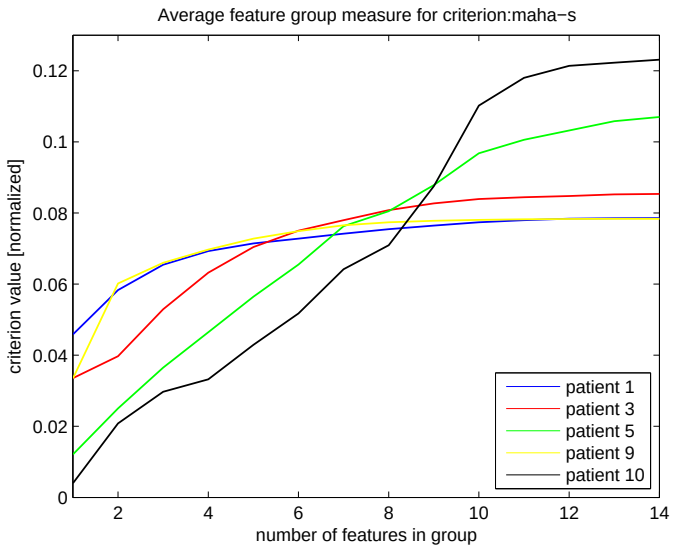

Figure 4.5: The intra-class Mahalanobis distance for the best groups of the specified size, as calculated by the backward algorithm. The connecting lines are shown to improve readability of the graph (measured values at the integer points).

- K-Nearest-neighbors classifier with automatic selection of k
- Support vector classifier with linear kernel
- Neural network with one hidden layer of 3 units
- Binary decision tree with pruning

For a successful design of some of the classifier types, setting of prior probabilities is necessary. In real case, these should be based on the probability of tremor occurrence, i.e. the ratio of time with registered tremor occurrence to the time without. As this probabilities are unknown and can not be derived from the dataset (the set consists of selected periods from a longer-term signal), the probabilities for tremor and non-tremor were assumed to be even. This should be reconsidered and corrected in further work on this problem using wider data.

The training process, as well as the classifier evaluation was realized by the PRTools toolbox. A brief description of the included classifiers can be found in the following paragraph.

Parzen classifier

This type of classifier is based on the probability density estimation by Parzen windows. Each instance of the training set is fitted with a windowing function of specified parameters over the feature vectors. This way an N-dimensional probability-density-function is estimated, given a feature vector of length N. When this is done, a threshold is found, dividing the target class (high densities) from the rest of the data. The properties of the classifier depend highly on the selection of the windowing function used: a smaller, more precisely specified windowing function will lead to better fitting of the classifier to the training set, while better generalization properties will be found for wider functions [5]. In this case, the selection of the windowing function was done automatically by the PRTools toolbox, based on the training set.

PRTools: parzenc

K-nearest-neighbors

Nearest neighbor classifier divides the feature space into sections, based on the class of the closest example of the training set. When a new instance is presented to the classifier, it is assigned a class according to the section it falls in, based on its feature vector (i.e. it is assigned the class of the nearest training sample). This is further generalized in the k-NN form, where the classification is based not on the proximity of one, but multiple training example, providing the algorithm with more stable results. Again, the successful classification depends on the selection of k. This decision is done by the PRTools toolbox, based on the leave-one-out error on the training set [6].

PRTools: knnc

SVC classifier

Support vector classifier (or "Support vector machine") is based on optimal fitting of a specified kernel function to the training data. The optimality is given by the maximum distance of the training samples to the dividing border, also taking penalties for possible classification errors into account (when the dataset is not separable by the selected kernel function). The function was trained with linear kernel, giving subjectively best results.

PRTools: svc

Neural network classifier

The neural network classifier included in the set is based on a three-layer architecture with one hidden layer of 3 elements. During the training, fitting of the output of the classifier to the training examples is done by tuning weights of the individual network elements with back propagation.

PRTools:bpxnc

Decision tree

In the decision tree classifier a specific parameter is chosen in each step of the algorithm, creating several branches that are further developed. A binary decision tree was included in the set, meaning there are two branches in each node. To maintain better generalization, pruning was introduced into the process, removing nodes with insufficient significance.

PRTools: treec

4.2.11 Classifier testing

When all the classifiers from the test are trained on the training data, testing on the testing set starts. This way, the classifier is presented with unseen data (data it was not trained on), which should provide estimation of the classifier performance in a real application. As mentioned before (Section 4.1.2), one sample of the dataset is represented by one section of the original data, divided further into shorter sections – chunks. Therefore, despite the use of the leave-one-out method, the training set is formed by not one, but several training feature vectors (instances). To calculate the errors, the testing set is classified by each of the classifiers and the number of correctly and erroneously classified samples is counted and stored. This is done by the implemented function `classerror.m`, which uses PRTools internal classification function labeld.

4.2.12 Classifier evaluation

The evaluation of the classifier performance comes after all the cycles of the leave-one-out method are finished and the error for each of the possible training-testing set selection is known. As the instance numbers differ section to section (the original signal sections have different lengths), it is necessary to take this fact into account. Therefore, the total

number of incorrectly classified chunks is calculated for each class, divided by the total relevant chunks count. This way, several common classifier performance measures are calculated, including:

- Accuracy: $\frac{TP+TN}{P+N}$
- Sensitivity: $\frac{TP}{P}$
- Specificity: $\frac{T}{N}$

N and P refer to the total number of samples from the negative (non-tremor) and positive (tremor) class, respectively, while TN and TP stand for the correctly classified samples for each class.

Apparently, apart from the accuracy measure, this gives an overall measure for the classifier performance, sensitivity will be of a great interest in the results evaluation, as it describes the ratio of overseen tremor instances to their total number. It is more likely to respect erroneously classified non-tremor samples (the DBS device runs without need) to a reasonable extent, than overseen tremor, which would mean the stimulation device does not fulfil its purpose when needed.

The calculation of the classifier evaluation metrics is carried out by the implemented function `classerrm`.

4.3 Experimental results

Having described the dataset, feature calculation methods and aspects of the classification process, the experimental results of these steps will be presented in this section.

4.3.1 Feature selection

The results of two corresponding steps are presented in this section: the general, individual feature evaluation, based on the whole data, and results of the feature group selection, done during the iterations of the leave-one-out method on one patient training set.

When combined, the results from both feature selection steps provide valuable information about the strength of individual features.

Individual feature evaluation

The individual feature evaluation was done on the whole dataset, including data from all the 10 patients. The aim of this step was to determine a general test of features, suitable for the tremor detection task.

The feature evaluation was based on calculation of two feature performance measures: Wilcoxon test and AUC (Section 4.2.4). Only the features passing the following criteria passed the process: Wilcoxon test result (P-value) lower than 0.001 (statistical significance above 99.9%) and AUC above 0.6.

These criteria were considered as sensible with regard to the feature set and a sufficient number of 14 out of 23 features passed the test. The evaluation results for the whole feature set are presented in Table 4.1. The results of Wilcoxon test, Fisher discriminant ratio and AUC are shown. The P-value represents the result from Wilcoxon test and can be interpreted as a significance level. It is therefore important to evaluate mainly the exponent of this parameter, which gives the basic information about the feature performance. The division between features selected as significant and the rest is indicated by the horizontal line. The f_{max} column informs about the upper frequency bounds where applicable. A visualisation of the results is provided in Figure 4.6.

Making a short investigation of the result, a correspondence can be noted between the different feature measures. While high value of AUC and FDR is required, a low P-value signifies a high confidence of statistical relevance. It can be noted, that features with high AUC usually have low P-value (high negative exponent) and also a high FDR. Nevertheless, P-value together with FDR were considered sufficient for the feature selection.

Note that the best features were coherently evaluated as very good by all the three implemented metrics, reaching the confidence level of the order of $\min 85$ and AUC 0.92 in the best case. Unlike the AUC and Wilcoxon test result, the Fisher discriminant ratio was not as coherent with the other two ratings which was one of the reasons for its exclusion from the assessment criterion.

It is worth to be mentioned that the selection of significant features was very stable over the tested range of chunk lengths between 0.5–4 s, differing in maximum of one feature. This fact points to a good level of time-invariability of the features. The rating of the best feature was very high on the whole dataset and a high rate of selection of these features during the feature-group selection is to be expected.

With the exception of the spectral band 3 (2–3.5 Hz), all the spectral bands passed

the test, as well as the specrep feature, characterizing the repetitivity in the spectrum. Both acratio and acpeaks features, based on the autocorrelation function of the signal, passed the test with a very high ranking. From the other feature types, only the maxdiff feature together with variance passed the test.

Average of the first differential together with fractal entropy estimates and higher order statistical moments ended up at the end of the ranking table and the values achieved show, that their excluxion from further testing is well reasoned.

Looking at the results once again, it is clear that no feature failed the adopted significance criterion only due to a low P-value. Anyway, except for the zerocross feature, all the excluded features failed at both parts of the criterion.

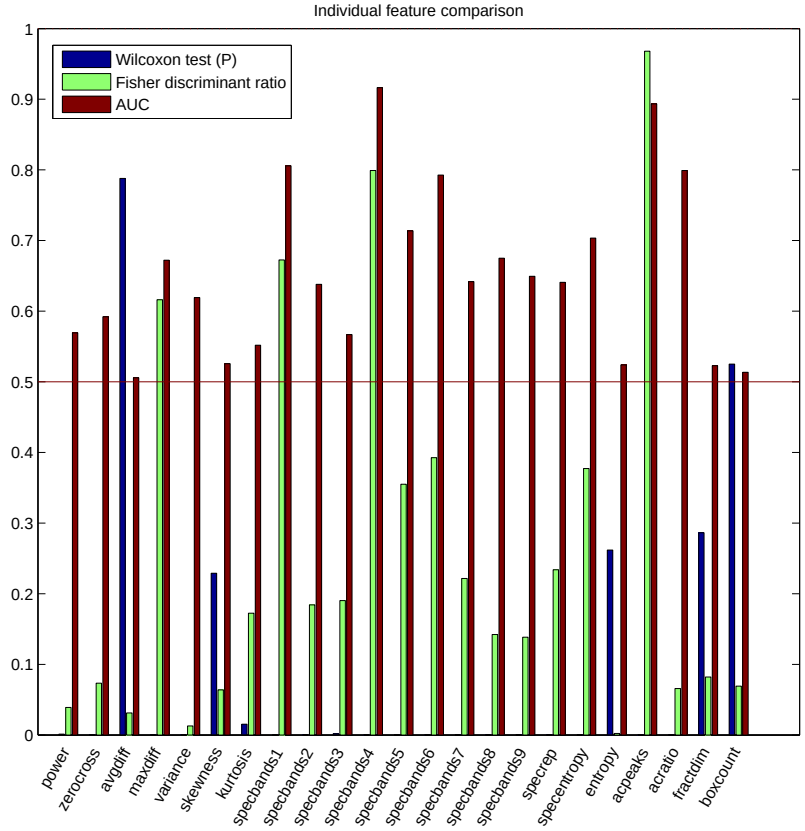

Figure 4.6: Individual feature evaluation results. Note the correspondence between high P-values and low AUC (close to 0.5), as well high FDR and high AUC (P-value very low in these cases).

Table 4.1: Individual feature evaluation results (whole dataset)

id	**featurename**	f_{max}	**P-value**	**FDR**	**AUC**
11	specbands4	5,5	8,52E-85	0,80	0,92
20	acpeaks		2,57E-72	0,97	0,89
8	specbands1	1	1,51E-46	0,67	0,81
21	acratio		1,14E-44	0,07	0,80
13	specbands6	10	8,36E-43	0,39	0,79
12	specbands5	7	1,28E-23	0,35	0,71
18	specentropy		1,47E-21	0,38	0,70
15	specbands8	16	2,37E-16	0,14	0,68
4	maxdiff		7,45E-16	0,62	0,67
16	specbands9	20	2,61E-12	0,14	0,65
14	specbands7	12	2,96E-11	0,22	0,64
17	specrep		2,21E-11	0,23	0,64
9	specbands2	2	1,02E-10	0,18	0,64
5	variance		2,31E-08	0,01	0,62
2	zerocross		2,11E-05	0,07	0,59
1	power		1,16E-03	0,04	0,57
10	specbands3	3,5	1,74E-03	0,19	0,57
7	kurtosis		1,52E-02	0,17	0,55
6	skewness		2,29E-01	0,06	0,53
19	entropy		2,62E-01	0,00	0,52
22	fractdim		2,86E-01	0,08	0,52
23	boxcount		5,25E-01	0,07	0,51
3	avgdiff		7,88E-01	0,03	0,51

Feature group selection

Seeing that the individual feature selection ignored dependencies between the features, it is interesting to mention the results of the feature group selection as done by the backward algorithm during the cycles of the leave-one out method. A set of 5 features was selected on the training set in each of the iterations, producing a total number of 27 selections.

Being done on rather small subsets of the training data (the selection is done on single patient's training sets), the table provides an important measure of the classification power. As a product of a backward algorithm, the selections take also the in-set dependencies into account. Different aspects of the generated feature-selection set will be discussed further in this text.

The Table 4.2 shows the number of occurrences of the features in the total of 27 feature selections done by the backward algorithm during the classifier training. The values of the measures P-value and AUC done during the initial significant feature selection are shown along for comparison. It can be noted that the rating of the best (mostly selected) features in the original significance tests is very high with a high value of AUC and high negative exponent of P-value. The full list of selected features can be found in Appendix A.

Seen from the results in the table, the best features were *acpeaks*—a feature based on autocorrelation function of the input signal, selected in more than 60% cases. It was followed by *specbands4*, capturing the energy level in the parkinsonian tremor band 3.5–5.5 Hz. Considering the third most selected feature—spectral band 0–1 Hz—it is apparent that the spectral properties of the signals played the most important role.

4.3.2 Outlier removal

Based on the observation of outlier counts for different features, the filtering criterion was set to exceeding of the per-class standard deviation at least 3 times in at least 2 features at the same time. This lead to removal of 61 outliers, out of which 25 belonged to tremor and 36 to non-tremor class. The total number of chunks was 1671, making the removed outliers 3.65% of the dataset. Outlier counts for different features and different criteria are shown in Table 4.3. The distance of the sample value from the mean value of the respective class, measured as a multiple of standard deviation was the measure calculated for each feature separately. The number of outliers, for which the criterion of 3σ distance was exceeded for a specific number of features is shown in the *group* column. No samples

Table 4.2: The number of occurences of the specific feature in the feature sets, selected in the leave-one-out-cycles, together with the results of individual feature evaluation from Table 4.1.

	feature	**freq. [Hz]**	**occurrences**	**P-value**	**AUC**
20	acpeaks		17	2,57E-72	0,89
11	specbands4	3,5–5,5	16	8,52E-85	0,92
8	specbands1	0–1	13	1,51E-46	0,81
12	specbands5	5,5–7	13	1,28E-23	0,71
16	specbands9	16–20	13	2,61E-12	0,65
5	variance		12	2,31E-08	0,62
4	maxdiff		10	7,45E-16	0,67
15	specbands8	12–16	7	2,37E-16	0,68
18	specentropy		7	1,47E-21	0,7
21	acratio		7	1,14E-44	0,8
9	specbands2	1–2	6	1,02E-10	0,64
13	specbands6	7–10	6	8,36E-43	0,79
14	specbands7	10–12	5	2,96E-11	0,64
17	specrep		3	2,21E-11	0,64

that would exceed the criterion at more than 3 features at the same time were found.

Table 4.3: The number of outliers, found for a certain criterion setting and feature count, shown for different features.

		single f.		**groups** (3σ)		
	feature	3σ	4σ	1	2	3
4	maxdiff	8	7	0	7	1
5	variance	12	11	7	4	1
8	specbands1	0	0	0	0	0
9	specbands2	11	2	11	0	0
11	specbands4	14	5	11	3	0
12	specbands5	33	18	21	10	2
13	specbands6	35	14	26	9	0
14	specbands7	27	8	18	9	0
15	specbands8	41	29	13	23	5
16	specbands9	50	29	23	23	4
17	specrep	25	13	6	16	3
18	specentropy	2	0	0	2	0
20	acpeaks	0	0	0	0	0
21	acratio	13	9	7	4	2
number of outliers		271	145	143	55	6

4.3.3 Classification

Eventually, this section presents the classification done. The evaluation for each classifier was calculated from all the leave-one-out cycles, carried out for each patient as described in Section 4.2.12. The results of the test for each of the 5 patients with sufficient data are shown in Table 4.4. The target (positive) class was tremor.

Evaluating the classification results from Table 4.4, accuracy, carrying information about the total ration of correctly classified training samples, is the first parameter to inspect. It can be noted, that the overall classification result, measured by this property

Table 4.4: The overall classification results for different patients and classifiers

Patient	**classifier**	**Accuracy**	**specificity**	**sensitivity**
	Parzen	0,982	1,000	0,923
	k-NN	0,982	1,000	0,923
1	SVC	0,973	0,988	0,923
	MLP	0,964	0,988	0,885
	Tree	0,938	0,942	0,923
	Parzen	0,875	1,000	0,769
	k-NN	0,861	0,939	0,795
3	SVC	0,736	0,667	0,795
	MLP	0,903	1,000	0,821
	Tree	0,875	0,939	0,821
	Parzen	0,657	0,200	0,733
	k-NN	0,714	0,000	0,833
5	SVC	0,800	0,000	0,933
	MLP	0,743	0,200	0,833
	Tree	0,743	0,200	0,833
	Parzen	0,210	0,358	0,118
	k-NN	0,080	0,113	0,059
9	SVC	0,022	0,019	0,024
	MLP	0,123	0,038	0,176
	Tree	0,043	0,000	0,071
	Parzen	0,969	1,000	0,941
	k-NN	0,969	1,000	0,941
10	SVC	0,969	1,000	0,941
	MLP	0,844	1,000	0,706
	Tree	0,969	1,000	0,941

was very high for patients 1, 3 and 10, moderate for patient 5 and very poor for patient 9.

However, considering the target application, sensitivity is the main parameter of interest, showing the ratio of correctly classified tremor samples. A sensitivity value very close or equal to one is necessary for a proper, applicable tremor detector. According to the results, most of the tremor samples were correctly classified for patients 1,3,5,10, while almost no tremor was detected for patient 9.

At last, specificity, the last parameter shown in the table, shows the ratio of correctly classified examples from the non-target class—non-tremor. This value shows how well the classifier performs on the non-tremor parts of the dataset and can be viewed as a measure of the stimulation time that could be saved. The values of specificity were outstanding for the patients 1, 3 and 10, while almost all non-tremor samples were misclassified for the patients 5 and 9.

The best possible value of sensitivity was aimed during the classifier design process, while keeping the specificity reasonably low. The classification results are discussed in the following section.

4.4 Experimental results discussion

After the listing of the experimental results, a discussion of these is undertaken in this section. Descriptions of links and dependencies, observed in the classification results, are given together with discussion of possible impact of the low amount of data, contained in the dataset.

4.4.1 Aspects implied by the properties of the dataset

In agreement with the numerous notices throughout the whole work, the main problem of the classification issue was the rest of the data. As seen in the table Table 2.1, the number of sections was at the minimum limit—two sections for each class—for most of the patients included in the set. The lack of the data even caused the exclusion of half of the present patients from the classifier testing and their data was used only for the feature selection.

The lack of the data affects the quality and precision of the error estimates too, as a different classification of a single chunk changes the classifier error to the extent of several percent, making the results unstable and decreasing their credibility considerably.

Another aspect—the unfixed amplitude scale of the data—induced loss of a number of simple features, such as the peak-to-peak value, that belong to the first-choice parameters during the processing of any signal. These parameters could be revealed to carry important information of tremor and non-tremor states and help to achieve better classification results. Although the absolute value of the signal can be highly patient-dependent, being affected by the properties and exact position of the electrode tips used, the value could be compared for each patient separately. Ideally, this could have been done in this work as well, had the very high differences in amplitude scale across the different patients not made the normalization necessary in order to work with the data as a consistent dataset.

4.4.2 Feature selection

There are several debatable aspects in both steps of the feature selection: The general significant feature selection and the feature group selection done prior to the classifier training. Firstly, the former will be discussed.

Feature significance

As described in the process description in Section 4.2.4 of this text, the individual feature selection consists of two steps: selection of one channel per patient and the feature evaluation itself. In both processes, the selection of a suitable criterion has very great impact of the calculated results. The choice of a suitable criterion has to be done prior to either of these methods can be performed. There is no clear solution to the selection and the only relevant measure can be the estimated error of the classification process. As the only way to finding the optimal settings of the criterion is trying all the possible methods and values, the optimal setting is impossible to find, due to the very high computational complexity of such task. Thus, the consideration of these aspects led to selection of AUC criterion for both tasks, based mainly upon its simplicity and low demands on the set size (no statistical assumptions necessary).

The selection of significant features is then supported by the Wilcoxon test, which evaluates other aspects of the set. Nevertheless, the Wilcoxon test, as well as the calculation of AUC is based on the ordering of the instances according to their value, which makes the results correlate very well. On the other hand, it also causes a lower credibility of the selection. The the significance of the selected method can be affirmed by the comparison of Fisher discriminat ratio, based on a different assumption of a normally distributed probability density for each class. The values of this measure correlate well with the results of the other measures, especially for the best rated features. The selection of methods for feature significance evaluation can therefore be considered consistent.

Alongside with the aspects already mentioned, credibility of the feature selection depends highly on the data, on which the selection is performed. To achieve the highest possible versatility of the feature set, it was chosen to include all the available data and not to perform the selection on training and testing set. This was reasoned mainly by the lack of data in both classes for some of the patients, which would cause their complete exclusion from the set, while they can provide important information that would be lost. The step of significant features should be seen as an attempt to determine the set of generally applicable features for the tremor task, rather than determination of exact feature ranking. Even though some of the rankings presented seem impressively high (e.g. the AUC and P-value for the three best ranked features), the method chosen cannot satisfy criteria for a highly consistent test and the values should be considered relative. However, a high significance of the best features in a real test is very expectable.

Generally speaking, the exclusion of the features with very low or almost no observed

significance provides better performance to the classification and further feature-selection algorithms, in which the number of available features commonly increases the number of necessary calculations vitally. As such, this step provides an important link in the process and also enables almost half of the dataset to be taken into account.

As for the channel selection, a big drawback of this method is not taking the actual position of the electrode in the brain into account. This is caused by the lack of this information in the input data and it is highly probable, that a channel selection, done by an experienced physician could lead to better results. However, the rules for optimal channel selection for tremor detection are not yet available, and thus the proposed approach, based on estimation of the tremor-detection power of the individual channel offers a consistent solution to the problem, being suitable for real application as well. Again, the selection of a suitable measure comes into play, which can hardly be answered without a test on wider dataset.

Similarly to the evaluation criteria, the adjustment of the individual features affects the results as well. Due to the number of possible decisions in the whole process, the tuning of the parameters of individual features becomes complicated, as the result is influenced by many factors and a contribution of the single feature is hard to track. Thus, the failure of some of the features in the significance test can be caused by their improper setting.

Inspecting the actual values of the selected features, several notes can be done on the ranking of some of the features. Firstly, both fractal features failed the test. This can be caused by an insufficient length of the input sequence (2s chunk length creates input sequences of 500 samples), as well as the simple insignificance of these features to the task. The low sample level can also be reason for the failure of the higher statistical moments (kurtosis, variance), as these fine properties of the probability density function require a high number of samples to provide sensible results. On the contrary, the most basic of the statistical features—variance—ended up as the last of the included features and is therefore very likely, that the statistical features have simply no significance to the problem. Further, entropy and power ended up at the end of the list as well. Unfortunately, no general conclusions can be done about them, being the only features of their kind. As for the highly ranked features, it is reasonable to evaluate these according to the results in the feature group selection, which is done on more specific data and carries more information about their actual importance.

Features selected in groups

Several conclusions can be done about the features, based on the number of their occurrences in the classification sets. Unlike the feature significance measures, the group selection was calculated only on training data for each patient and the consistency of the selected feature groups over the iteration of the classification algorithm can provide a lot of information about the stability of the method. The respect the adopted backward algorithm pays to the feature dependencies is also important to be noted.

The first conclusion about the features, based on the occurrences in the Table 4.2 is that the most selected features—*acpeaks* and *specbands4*—as well as the following ones, reached a very high rating in the significant feature selection, done on different and wider data, as well. Moreover, there were no completely unselected features and it can be therefore concluded, that the significant feature selection is with these results in a great correspondence and therefore seems to be done suitably.

The tremor frequency seems to play an important role in the dataset, according to the features at the top positions in the occurrence table—-acpeaks and specbands 4 (3.5–5.5 Hz)—which both utilize the tremor frequency. On the on the other hand, at least one of the tremor-frequency based features (specbands4, acpeaks or acratio) was selected in 22 out of 27 cases, which indicates the tremor frequency is not a measure suitable for all the recordings.

The question on presence of the tremor frequency in the Parkinsonian LFP signals remains the greatest question of this work and should definitely be further researched on wider data.

Generally, based on the feature evaluation data, the frequency spectrum has obviously a key meaning for the tremor detection, with main points of interest in the tremor (3.5–5.5 Hz) and lowest (0–1 Hz) bands. Other most important bands were identified in the 5.5–7 Hz and 16–20 Hz. The former could be caused also by the variance of the tremor frequency from patient to patient. The main band of this apparent spectral peak was registered mainly in the range 3,5-5,5 (a sharp peak at a frequency somewhere in this range), but can probably differ even more for other patient. The main frequency of most tremor recordings is carried by the tremor frequency and its harmonics. On the contrary, the registered spectrum shape of the non-tremor recordings seemed to be very flat with a high peak close to zero, which is probably caused by the lack of other spectral components and the performed spectral normalization, recalculating the sum of the whole spectrum to one.

Inspecting the individual selected feature groups, the selection varies from one training-testing set selection to another, which is probably caused by the lack of the data. Nevertheless, there are high overlaps between the selected sets and just very few outlier sets can be found, differing highly from the others for the same patient. The differences are caused by the high differences between the training sets, composed of just very few samples. The same presumption can be derived about the classification error, which varied highly between the different selections of the training set.

4.4.3 Classification

Before starting the discussion of the classification results, let us recall the aim of this work. The final applied tremor detection device should reduce the time of stimulation applied to the human brain stem by switching the stimulation on demand, only in periods where necessary. A very important requirement is that the reduction of stimulation time should cause no noticeable harm to the efficiency of the therapeutic process, meaning that any starting tremor should be captured early enough for the stimulation to take effect. This way, the patient experience should remain unchanged, while the battery life of the device could be improved significantly.

In terms of the evaluation of the classification results, the reliability of the detection can be estimated from the sensitivity parameter, which should be very close or equal to one in order to meet the requirement of an unchanged patient experience. On the other hand, the energy-conserving ability of the detection algorithm can be read from the specificity measure. A high specificity would mean no unnecessary stimulation, while high sensitivity would mean a great reliability of the system. Unfortunately, these two requirements are antagonistic in classifier design: moving the decision hyperplane towards higher sensitivity will cause lower specificity (more false positives) and vice versa. A limit exists for each specific classification method and problem, causing a certain achievable classifier performance that can not be exceeded, assuming a reasonable level of the important generalization properties of the classifier are kept.

In order to satisfy the target application, the classifiers should be set for a very high sensitivity, keeping just a reasonable level of specificity. This was attempted during the classifier design process, even though not fully achieved. Comparing the results in 4.4, a considerably good performance can be observed for patients 1,3 and 10, where the sensitivity achieved was in the range of 0.82–0.94 while specificity was as high as one. According to what was discussed above, the classifier threshold could be shifted in order to

increase sensitivity at the expense of lower specificity. To achieve this shift, a specialized classifier design would be necessary, altering the evaluation criteria. This step is rather complicated and was therefore not carried out in this classification, which was designed to give a rough overall idea about the possible results.

As for the other patients, the results were inferior for patient 5, for which all the training samples were classified as tremor, causing the specificity goes to zero. This means there would be no benefit in application of the method for this patient. Inspecting the original LFP files, an important fact was discovered about the data of this patient: all the sections used come from just two files – the ones available. Both these files record a tremor onset, starting with a low tremor activity and ending with a full tremor. The non-tremor section was taken from the signals before the assumed tremor onset, while the tremor signal was taken from the recording later in time. As the tremor onset is a process of certain duration and progress, determination of the exact onset point is hardly feasible. Therefore an unused section was marked around the assumed onset to prevent eventual contamination of the sections by the properties of the opposite class. However, the determination of the tremor onset was done subjectively, based on the corresponding EMG signals and it is possible that the differences between the two final sections produced were to low for the general feature set to capture. This is a very optimistic explanation, suggesting the tremor could be captured even before the available recording started. On the contrary, the results can be simply caused by the lack or non-homogeneity of the data.

Another patient with very poor results was patient 9, for whom the classifier performed much worse than would a random guess, classifying almost all the samples incorrectly. This reflects probable high inconsistency of the data for this patient, causing the properties of the training set change extremely from iteration to iteration. Inspecting the original LFP time signals for this patient, no significant spectral components were found in them and all the spectra exhibited approximately the same form, no matter which class they originated from. At this point it is reasonable to question the proper recording and description of the data was done, as the signals differ very highly from the rest of the dataset. Unfortunately, no satisfactory answer can be offered, according to the information available. The last step of the classification process, which was intended to be done, was the classification of the whole onset recordings for patients where the data was available. In this process, original file would be divided into highly overlapping chunks, indifferent of the tremor marks. Then, a previously trained classifier for the patient would be used to classify the chunks and the output classification would be shown together with

the time series. This way an easy-to-evaluate idea of the actual performance of the system could be gained. Unfortunately, the classification results were very bad for the only two patients, where the onset recordings were available (5, 9) and the intention could not be carried out.

4.5 Chapter summary

In this chapter, classification has been carried out on the dataset, using the features described in Chapter 3. The feature set was subjected to a two-fold evaluation: In the first step, the features were evaluated individually on the whole 10-patient dataset and their overall significance with respect to the tremor detection was estimated. In the second step, the power of the features was evaluated according to the selection done on single patient data in multiple runs of the classification cycle. The results of both classification and feature evaluation have been discussed: the spectral bands 1 and 4, together with the acpeaks feature were the most significant ones. The classification achieved good results for patients 1,3 and 10, while for the patient 5 and 9 the results were poor and the method failed.

Chapter 5

Conclusion

A novel approach to the detection of tremor in Parkinsonian patients' recorded local field potentials of the basal ganglia has been presented in this work. Features of different kinds, including temporal spectral, based on the autocorrelation function, information theory or fractal dimension estimation have been evaluated on a 10-patient dataset, selected out at the beginning of the work.

According to the classification results, the classifier performance was very good for 3 out of 5 patients, with sensitivity up to 94% and specificity equalling one. The whole suggested method seems to be a suitable solution for tremor detection. However, some steps could be markedly simplified by additional knowledge about the origin of the files (e.g. the channel selection). Further knowledge could also lead to the exclusion of generally unnecessary steps in the process, providing possibly more information about the problem. The amplitude normalization of the time series can be seen as a good example of such step.

Major role of spectral properties of the LFP signals has been confirmed by evaluation of the single features as well as feature selection during the classification process. This affirms the previous research on the topic [16]. The importance of the spectral properties was found to be superior to other possible approaches, leaving the importance of the signal variance and maximum of the first differential for discussion. Other approaches, such as entropy or fractal dimension were found very improper for the task in this test.

A way of simple detection of tremor presence, based on the autocorrelation function, has been developed and successfully acknowledged. The *acpeaks* feature, implementing this approach was, the most frequently selected from all the available features. Its separability measure ranking reached high values as well. This approach should bring a simple and reliable detection, given the tremor frequency is present in the LFP signals of the

given patient. However, despite the very high occurrence of the tremor frequency in the dataset, wide applicability of this theory is yet to be approved.

The lack of training data can be seen as a major drawback of this work, affecting the accuracy of the classification results radically. Use of a wider dataset should provide more relevant results, especially in the classifier and feature-group selection phase, which suffered from the lack of data most. The results of individual feature evaluation, done on a wide 10-patient dataset, and further evaluation training subsets correlated very well. This indicates a probable good general applicability of the method to wider population.

5.1 Further work

Many of the problems discussed in this thesis were connected with lack of suitable data or description. This affected the accuracy and made the evaluation impossible for some of the patients. Furher work should therefore be done on wider data, including more patients and more patient data as well. Another possibility, that could simplify the process, would be obtaining more information about the recording conditions. More exact information about the electrode position could provide better understanding and easier handling of the signals.

As the most successful features were based on the occurrence of tremor frequency in the STN signals, further work should be aimed to prove presence of this frequency in the signals of more patients. Proving the presence of this frequency for at least a group of patients would mean a relatively simple way to detect tremor exists.

Apart from further research on wider data, the reasons why the method did not work for some of the patients should be further researched. Possible causes such as absence of the tremor frequency or erroneous recording should be either proved or rejected.

Bibliography

[1] Brown, P., Williams, D. Basal ganglia local field potential activity: Character and functional significance in the human. In *Clinical Neurophysiology*, Official journal of the IFCN, Vol. 116, No. 11, November 2005. ISSN 1388-2457

[2] Butler,M. A., Rosenow, J. M., Okun M. S. History of the Therapeutic Use of Electricity on the Brain and the Development of Deep Brain Stimulation in *Deep Brain Stimulation in Neurological and Psychiatric Disorders*, ed. Tarsy, D. et al (editor),NJ USA:Humana Press, 2008. ISBN: 978-1-58829-952-9

[3] Chiappa, K. A. Principles of evoked potentials in *Evoked Potentials in Clinical Medicine*. 3rd ed., Philadelphia: Lippincott - Raven Publishers. ISBN 0-397-51659-2

[4] Dietterich, T. G. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. In *Neural Computation*,October 1998, Vol. 10, No. 7, Pages 1895-1923. ISSN 0899-7667

[5] Duda, O., Hart, P. E., Stork, D. G. *Pattern Classification*. 2nd ed., Wiley, 2000. ISBN 978-0-471-05669-0

[6] Duin, R.P.W., et al. *PRTools4 - A Matlab Toolbox for Pattern Recognition*. Version 4.1, Delft University of Technology, 2007. ¡http://www.prtools.org¿

[7] Elbert, T., et al. Chaos and Physiology: Deterministic Chaos in Excitable Cell Assemblies in *Physiological Reviews*, The American Physiological Society, 1994, Vol. 74, No. 1. ISSN 1522-1210

[8] Hudgins, B., Parker, P. A New Strategy For Myoelectric Control, *IEEE Transactions on Biomedical Engineering*. Vol. 40, p. 82-94, 1993. ISSN 0018-9294

[9] Kruger, A. Implementation of a fast box-counting algorithm. In *Computer Physics Communications*. Elsevier B. V., 1996, Vol. 98, p. 224-234. ISSN 0010-4655

[10] Lemstra, A. W, et al. Tremor-frequency (3–6 Hz) activity in the sensorimotor arm representation of the internal segment of the globus pallidus in patients with Parkinson's disease. In *Neuroscience Letters*, Vol. 267, Issue 2, May 1999, pages 129–132, ISSN 0304-3940

[11] Levy, R., et al. Dependence of Subthalamic Nucleus Oscillations on Movement and Dopamine in Parkinson's Disease. In *Brain*, Vol. 125, No. 6, June 2002. ISSN 0006-8950

[12] Lovely, D. F.: Signals and Signal Processing for Myoelectric Control. In *Powered upper limb prostheses: control, implementation and clinical application*, Muzumdar, A, Ed., Springer, 2004. ISBN 3540404066

[13] Mašek, O. *Heart Rate Variability Analysis*, Praha, 2009, Diplomová práce na katedře kybernetiky FEL ČVUT

[14] *Matlab and Simulink* [PC software], version 7.8 (R2009a), MA USA: The MathWorks, 2009, <http://www.mathworks.com>

[15] Olanow, C. W., Tatton, W. G. Etiology and Pathogenesis of Parkinson's Disease. In *Annual Review of Neuroscience*, Vol. 22, pages 123–144, March 1999, ISSN 1545-4126

[16] Pan, S., Warwick, K., et al. Prediction of Parkinson's Disease Tremor Onset using Artificial Neural Networks. In *BioMed 2007 - Biomedical Engineering*, ed. Gardner, J.W., Innsbruck: ACTA, 2007. ISBN 0-88986-648-5

[17] Ray, N. J., et al. Local field potential beta activity in the subthalamic nucleus of patients with Parkinson's disease is associated with improvements in bradykinesia after dopamine and deep brain stimulation. In *Experimental Neurology*, Issue 212, 2008. ISSN 0014-4886

[18] Shannon, C. E. A Mathematical Theory of Communication. Reprinted in *Mobile Computing and Communications Review*, vol. 5, number 1, 2001, Association for Computing Machinery. ISSN 1559-1662

[19] Sevcik, C. A procedure to Estimate the Fractal Dimension of Waveforms. In *Complexity International*, Monash University, 1998, Vol. 5. ISSN 1320-0682

[20] Sprent, P., Smeeton, N. C. *Applied nonparametric statistical methods*. 3rd ed., Boca Ranton (US): CRC Press, 2001. ISBN 1-58488-145-3

[21] Steigerwald, F., Volkmann, J. Intra-operative Microrecording and Stimulation in *Deep Brain Stimulation in Neurological and Psychiatric Disorders*. Tarsy, D. et al., Editor. NJ USA:Humana Press, 2008. ISBN 978-1-58829-952-9

[22] Theodoridis, S., Koutroumbas, K. *Pattern Recognition*. 3rd ed., CA (USA): Academic Press, 2006. ISBN 0-12-369531-7

[23] Weinberger, M., Hutchinson, W. D., Dostrovsky, J. O. Pathological subthalamic nucleus oscillations in PD: Can they be the cause of bradykinesia and akinesia? in In *Experimental Neurology*, Issue 212, 2008. ISSN 0014-4886

Appendix A

Selected feature groups

Table A.1: Feature groups as selected by the backward algorithm in the individual iterations of the leave-one-out method

Patient	**Selected features**				
1	4	11	14	20	21
	11	12	18	20	21
	11	12	18	20	21
	4	11	12	20	21
	11	14	18	20	21
	4	11	12	17	20
	11	12	16	20	21
	8	11	12	18	20
3	8	9	13	15	16
	5	11	13	14	20
	5	8	12	16	20
	5	8	16	18	20
	5	8	16	18	20
5	8	9	11	12	16
	4	9	11	12	13
	8	12	15	16	17
	4	5	8	12	16
	4	5	11	12	15
	4	9	11	12	15
9	4	5	8	11	15
	4	8	9	15	16
	4	5	8	14	18
	8	11	14	16	20
10	5	13	16	20	21
	5	8	9	11	20
	5	13	15	16	20
	5	13	16	17	20