

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

DIPLOMOVÁ PRÁCE

Brno, 2016

Bc. Ján Mucha



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY

A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

**AKUSTICKÁ ANALÝZA PŘEDNESU BÁSNĚ U PACIENTŮ
S PARKINSONOVOU NEMOCÍ**

ACOUSTIC ANALYSIS OF POEM RECITATION IN PATIENTS WITH PARKINSON'S DISEASE

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Ján Mucha

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Zoltán Galáž

BRNO 2016



Diplomová práce

magisterský navazující studijní obor **Telekomunikační a informační technika**

Ústav telekomunikací

Student: Bc. Ján Mucha

ID: 147417

Ročník: 2

Akademický rok: 2015/16

NÁZEV TÉMATU:

Akustická analýza přednesu básně u pacientů s Parkinsonovou nemocí

POKYNY PRO VYPRACOVÁNÍ:

Parkinsonova nemoc je druhé nejčastější neurodegenerativní onemocnění na světě. Odhaduje se, že až u 60 - 90% pacientů trpících Parkinsonovou nemocí se vyskytuje dysfunkce motorické realizace řeči, nazývaná hypokinetická dysartrie. Vyznačuje se především v oblasti fonace, artikulace a prozodie řečového projevu. V rámci této práce bude provedena analýza přednesu básně jako prozodicky a artikulačně náročného úkonu. Toto řečové cvičení bude parametrizováno moderními metodami zpracování řeči. Následně budou tyto parametry analyzovány pomocí metod binární klasifikace. Výsledkem bude systém automatické diagnostiky Parkinsonovy nemoci. Uvedené postupy budou naprogramovány v jazyku MATLAB.

DOPORUČENÁ LITERATURA:

[1] PSUTKA, J.; et al. Mluvíme s počítačem česky. Praha: Academia, 2006. 752 s. ISBN 80-200-1309-1.

[2] SMÉKAL, Z. Číslicové zpracování řeči (MZPR). Elektronická skripta pro magisterská studia, Ústav telekomunikací, FEEC, VUT Brno, 2009.

[3] RUSZ, J.; CMEJLA, R.; RUZICKOVA, H.; RUZICKA, E.: Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. J Acoust Soc Am, roč. 129, č. 1, 2011: s. 350–367.

Termín zadání: 1.2.2016

Termín odevzdání: 25.5.2016

Vedoucí práce: Ing. Zoltán Galáž

Konzultant diplomové práce:

doc. Ing. Jiří Mišurec, CSc., předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Diplomová práca je zameraná na akustickú analýzu prednesu básne u pacientov trpiacich Parkinsonovou chorobou. Táto choroba je sprevádzaná poruchou reči nazývanou hypokinetická dysartria. Jedným z cieľov práce je oboznámenie sa s priebehom, príznakmi a liečbou týchto chorôb. V práci je uvedené predspracovanie a parametrizácia rečového signálu a metódy binárnej klasifikácie. Následne je uvedený návrh modulárneho systému automatickej diagnostiky Parkinsonovej choroby na základe akustickej analýzy reči. Navrhnutý systém je implementovaný v prostredí Matlab. Klasifikácia vypočítaných parametrov je realizovaná metódami náhodných lesov a podporných vektorov. Výsledky týchto metód sú porovnané a uvedené v diplomovej práci. Hlavným cieľom a výsledkom diplomovej práce je systém automatickej diagnostiky Parkinsonovej choroby na základe akustickej analýzy prednesu básne.

KĽÚČOVÉ SLOVÁ

Parkinsonova nemoc, Hypokinetická dysartria, akustická analýza prednesu básne, akustická analýza reči, systém automatickej diagnostiky Parkinsonovej choroby

ABSTRACT

Diploma thesis is focused on the acoustic analysis of poetry recitation in patients with Parkinson's disease. This disease is associated with speech disorder called hypokinetic dysarthria. One objective of this thesis was familiarization with process, symptoms and treatment of these diseases. In thesis is described preprocessing and parametrization of the speech signal and the binary classification methods. Subsequently, it is the above proposal modular system of auto-diagnosis of Parkinson's disease based on acoustic analysis of the speech. The proposed system is implemented in MATLAB. Classification of calculated parameters is realized using the method of Random forest and Support vector machine. The results of these methods are compared and listed in the thesis. The main objective and the result of this thesis is a system of automatic diagnosis of Parkinson's disease based on acoustic analysis of the poem recitation.

KEYWORDS

Parkinson's disease, Hyperkinetic dysarthria, acoustic analysis of poems, poem analysis, speech analysis, auto-diagnosis system of Parkinson's disease

MUCHA, Ján *Analýza přednesu básně u pacientu s parkinsonovou nemocí: diplomová práca*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2016. 69 s. Vedúci práce bol Ing. Zoltán Galáž,

PREHLÁSENIE

Prehlasujem, že som svoju diplomovú prácu na tému „Analýza přednesu básně u pacientu s parkinsonovou nemocí“ vypracoval(a) samostatne pod vedením vedúceho diplomovej práce, využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autor(ka) uvedenej diplomovej práce ďalej prehlasujem, že v súvislosti s vytvorením tejto diplomovej práce som neporušil(a) autorské práva tretích osôb, najmä som nezasiahol(-la) nedovoleným spôsobom do cudzích autorských práv osobnostných a/nebo majetkových a som si plne vedomý(-á) následkov porušenia ustanovenia § 11 a nasledujúcich autorského zákona č. 121/2000 Sb., o právu autorskom, o právach súvisejúcich s právom autorským a o zmeně niektorých zákonov (autorský zákon), vo znení neskorších predpisov, vrátane možných trestnoprávných dôsledkov vyplývajúcich z ustanovenia časti druhej, hlavy VI. diel 4 Trestného zákoníka č. 40/2009 Sb.

Brno

.....

podpis autora(-ky)

POĎAKOVANIE

Rád by som poďakoval vedúcemu mojej diplomovej práce, pánovi inžinierovi Zoltánovi Galážovi, za veľmi kvalitné vedenie práce, potrebnú pomoc a poskytnutie dôležitých materiálov potrebných pre realizáciu tejto diplomovej práce.

Brno

.....

podpis autora(-ky)



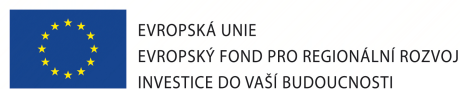
Faculty of Electrical Engineering
and Communication
Brno University of Technology
Purkynova 118, CZ-61200 Brno
Czech Republic
<http://www.six.feec.vutbr.cz>

POĎAKOVANIE

Výzkum popsaný v tejto diplomovej práci bol realizovaný v laboratóriách podporených projektom SIX; registračné číslo CZ.1.05/2.1.00/03.0072, operačný program Výzkum a vývoj pro inovace.

Brno

.....
podpis autora(-ky)



OBSAH

Úvod	12
1 Parkinsonova choroba a Hypokinetická dyzartria	14
1.1 Parkinsonova choroba	14
1.2 Hypokinetická dyzartria	16
1.2.1 Fonácia	16
1.2.2 Artikulácia	17
1.2.3 Prozódia	19
1.2.4 Plynulosť reči	19
1.2.5 Faciokinézia	19
1.3 Liečba Hypokinetickej dyzartrie	20
2 Predspracovanie rečového signálu a Parametrizácia	22
2.1 Predspracovanie rečového signálu	22
2.1.1 Ustredenie	23
2.1.2 Preemfázova filtrácia	23
2.1.3 Segmentácia	24
2.2 Parametrizácia	26
2.2.1 Prozodické parametre	26
2.2.2 Artikulačné parametre	28
2.2.3 Fonačné parametre	28
2.2.4 Parametre hodnotiace kvalitu signálu	29
3 Štatistické a klasifikačné metódy	32
3.1 Štatistika	32
3.2 Klasifikácia	33
3.2.1 Random forest	33
3.2.2 Support vector machine SVM	34
3.2.3 Validácia	35
3.2.4 Confusion matrix	36
3.2.5 Korelačný koeficient	37
4 Návrh systému	38
4.1 Popis modulárneho systému	38
4.2 Blokový diagram	39

5	Realizácia systému	40
5.1	Načítanie databázy rečových záznamov	40
5.1.1	Databáza	40
5.1.2	Načítanie databázy	40
5.2	Predspracovanie (Segmentácia)	41
5.3	Parametrizácia	41
5.3.1	Dekódovanie názvu	41
5.3.2	Výpočet parametrov	42
5.3.3	Výpočet štatistických hodnôt	43
5.3.4	Individuálna klasifikácia	46
5.3.5	Klasifikácia viacerých parametrov	47
6	Vyhodnotenie	49
6.1	Individuálna analýza	49
6.2	Analýza viacerých parametrov	53
	Záver	57
	Literatúra	60
	Zoznam symbolov, veličín a skratiek	65
	ZOZNAM PRÍLOH	68
	A Obsah priloženého DVD	69

ZOZNAM OBRÁZKOV

1.1	Zobrazenie častí a ciest v mozgu obsahujúcich doapmín[7]	14
1.2	Hlasový trakt človeka[17]	18
2.1	Časový priebeh Hammingovho a pravouhlého okna a ich modul spektra	25
3.1	Confusion matrix	36
4.1	Blokový diagram systému	39
5.1	Ukážka štruktúry <i>matrix_of_param.mat</i> , obsahujúca všetky počítané parametre pre každého rečníka.	45
6.1	Zobrazenie odhadu strednej hodnoty prvých 10 najlepšie hodnotených parametrov podľa mRMR. Zobrazenie je po dvojiciach kde každá dvo- jica znázorňuje hodnotu pre chorých a zdravých rečníkov.	56

ZOZNAM TABULIEK

6.1	Najlepšie hodnotené parametre podľa ACC	49
6.2	Najlepšie hodnotené parametre podľa SEN	50
6.3	Najlepšie hodnotené parametre podľa SPC.	51
6.4	Prehľad najlepších parametrov individuálnej analýzy klasifikovaných metódou Random forest. V tabulke je uvedený názov parametra, ACC, SEN, SPC, Pearmanov korelačný koeficient(PK) a jeho p hodnota(pPK) a Spearmanov korelačný koeficient(SK) a jeho p hodnota(pSK).	52
6.5	Prehľad najlepších parametrov individuálnej analýzy klasifikovaných metódou SVM. V tabulke je uvedený názov parametra, ACC, SEN, SPC, Pearmanov korelačný koeficient(PK) a jeho p hodnota(pPK) a Spearmanov korelačný koeficient(SK) a jeho p hodnota(pSK).	53
6.6	Skupiny parametrov klasifikované metódou Rnadam forest zoradené podľa ACC, v ktorých poradie parametrov je hodnotené podľa algoritmu mRMR.	53
6.7	Skupiny parametrov zoradené podľa ACC, v ktorých poradie parametrov je hodnotené podľa algoritmu mRMR. Skupiny parametrov sú klasifikované metódou SVM.	54
6.8	Výpis parametrov z najlepšej skupiny podľa ACC. Parametre sú zoradené podľa hodnotenia mRMR.	55

ÚVOD

Ludská reč patrí medzi najpoužívanejší a najdôležitejší komunikačný nástroj na svete. Prešla dlhodobým vývinom od pravekých ľudí až po dnešnú formu mnohých sofistikovaných jazykov spolu s nespočtým množstvom nárečí. Stala sa tak neodmysliteľnou súčasťou života. Poškodenie reči má na kvalitu života mimoriadne negatívny dopad. Pre človeka trpiaceho akoukoľvek poruchou reči sa stávajú každodenné rutinne úkony náročnými úlohami. Stráca možnosť jasne a zreteľne komunikovať so svojím okolím a mnohokrát sa začínajú prejavovať aj psychické poruchy. Vzhľadom na svoju poruchu reči sa snaží vyhýbať priamej komunikácii a tým pádom sa stáva menej spoločenským. Stráca svoje postavenie v spoločnosti, uzatvára sa pred ostatnými ľuďmi, čím sa jeho rečová porucha a zároveň aj psychický stav len zhoršuje.

Parkinsonova choroba je druhým najčastejším neurodegeneratívnym ochorením s chronicky progredujúcim priebehom na svete. Postihuje približne 1,5 % populácie vo veku nad 65 rokov, pričom riziko vzniku rastie s vekom. Odhaduje sa, že až u 60 – 90 % pacientov trpiacich Parkinsonovou chorobou sa vyskytuje dysfunkcia motorickej realizácie reči, nazývaná Hypokinetická dyzartria, postihujúca najmä oblasti fonácie, artikulácie a prozódie.

Diplomová práca je zameraná na akustickú analýzu prednesu básne u pacientov s Parkinsonovou chorobou. Je rozdelená na šesť častí. Prvá časť práce venuje pozornosť problematike Parkinsonovej choroby a hypokinetickej dyzartrie. Popisuje jednotlivé ochorenia, ich príznaky, priebeh, momentálne liečebné procesy, ako aj ich spojitost.

Pri analýze rečového signálu je dôležitým krokom výpočet jeho parametrov, avšak rečový signál je nutné pred parametrizáciou predspracovať. V druhej časti diplomovej práce je uvedený popis metód predspracovania signálu a to ustredenie, preemfáza a segmentácia, ich výpočet a použitie. Následne sú popísané parametre počítané pri analýze rečového signálu. Jedná sa o parametre prozodické, artikuláčnne, fonačné a parametre hodnotiace kvalitu signálu, pričom najväčší význam pre túto prácu majú parametre prozodické a artikuláčnne.

Aplikáciou parametrizácie na rečový signál môže byť výsledkom skalárna hodnota, alebo vektor hodnôt. V diplomovej práci sú použité klasifikačné metódy pre rozpoznanie Parkinsonovej choroby. Pre tieto metódy je nežiaduce, aby bol výsledok akéhokoľvek parametru vo forme vektora, a preto je potrebné použiť štatistické

metódy pre dosiahnutie skalárnych hodnôt. Popis použitých štatistických a klasifikačných metód je uvedený v tretej časti diplomovej práce.

Štvrtá časť diplomovej práce obsahuje návrh systému automatickej diagnostiky Parkinsonovej choroby, pozostávajúci z blokového diagramu a popisu jednotlivých blokov. Následná realizácia celého systému a jeho vyhodnotenie sa nachádza v piatej a šiestej časti diplomovej práce. Systém je realizovaný v prostredí Matlab.

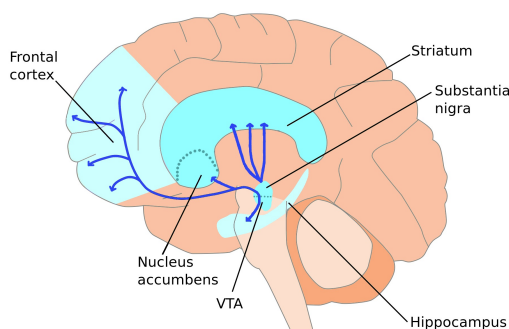
Ciele diplomovej práce sú naštudovanie Parkinsonovej choroby a Hypokinetickéj dyzartrie na znalostnú úroveň umožňujúcu následné riešenie automatického systému jej detekcie. Oboznámenie sa s problematikou predspracovania rečového signálu a jeho následnej parametrizácie a taktiež naštudovanie aplikovania klasifikačných metód pre danú problematiku. Hlavným cieľom a výsledkom diplomovej práce je systém automatickej diagnostiky Parkinsonovej choroby na základe akustickej analýzy prednesu básne.

1 PARKINSONOVA CHOROBA A HYPOKINETICKÁ DYZARTRIA

1.1 Parkinsonova choroba

Presná príčina vzniku Parkinsonovej choroby nie je doposiaľ presne známa[1]. Prvýkrát ju popísal vo svojej práci *An Essay on the Shaking Palsy* James Parkinson v roku 1817[2], avšak najstarší opis symptómov sprevádzajúcich Parkinsonovu chorobu je spísaný v literárnych pamiatkach *The Four Vedas*, ktorých existencia siaha až 5000 rokov p.n.l. do Indie[3].

Jedná sa o idiopatickú neurodegeneratívnu chorobu. V priemere postihuje asi 1,5 % populácie vo veku nad 65 rokov, pričom riziko vzniku rastie s vekom. Vo veku 85 rokov narastá prevalencia na 4-5 %. Po Alzheimerovej chorobe predstavuje druhé najčastejšie neurodegeneratívne ochorenie s chronicky progredujúcim priebehom. Mužov postihuje 1,5-krát častejšie ako ženy. Predpokladá sa, že ochorenie vzniká kombináciou vonkajších faktorov a individuálnej genetickej predispozície[4]. Parkinsonský syndróm je prevažne zapríčinený poruchou dopaminergnej inervácie striata, vedúceho k poruche funkcie spätne-väzobných okruhov nazývaných extrapyramídový systém. Konštantným nálezom u Parkinsonovej choroby je zánik neurónov a makroskopicky zrejmá depigmentácia v *substantia nigra pars compacta* (tzv. "čierne jadro"[5]). Najvýznamnejším biologickým nálezom je úbytok obsahu dopamínu v bazálnych gangliách, najmä v striate, odkiaľ je následne uvoľňovaný do synapsií, ktoré majú zásadný význam pri komunikácii jednotlivých buniek nervového systému. Dopamín(4-(2) aminoethyl benzén-1,2-diol) slúži v ľudskom tele ako transmitter nigrostriánej dráhy[6].



Obr. 1.1: Zobrazenie častí a ciest v mozgu obsahujúcich doapmín[7]

Mozog túto látku vylučuje prevažne vo chvíľach pocitovo vyjadrujúcej túžby, vedie mozog k činnosti, je nepostrádateľný pri pohybe svalstva. K poklesu dopamínu dochádza v priebehu života aj u zdravých ľudí, najmä v starobe, nie však až v takom merítku aby došlo k prejavom Parkinsonovej choroby. Tie sa objavujú až v štádiu, keď je zničená minimálne polovica buniek slúžiacich na jeho tvorbu v substantia nigra a jeho množstvo v striate poklesne o viac než 70 - 80 % oproti pôvodnému stavu. Úbytok dopamínu je spojený s poruchou motorických funkcií.[5] Nie je však pozorovaný dopad na dĺžku života, ale výrazne pôsobí na jeho kvalitu. Pacient s Parkinsonovou chorobou sa môže javiť ako menej inteligentný, bez emócií, pri komunikácii sa môže zdať že nepočúva, nereaguje, má problém s vyjadrovaním a podobné sprievodné príznaky poruchy pohybového aparátu. To vyplýva z troch charakteristických kardinálnych príznakov pri onemocnení na Parkinsonovu chorobu, a to tremor, rigidita, pohybové poruchy a s tým spojená porucha reči, ktorá bude analyzovaná v tejto práci[3][4][6].

Tremor/Tras Ide o najznámejší prejav Parkinsonovej choroby. Je to pravidelný rytmický, oscilačný pohyb. Jeho frekvencia je okolo 5 Hz. Vzniká na základe nedostatku dopamínu v bazálnych gangliách. Dopamín za normálnych okolností pôsobí tlmivo. Pri jeho nedostatku prevažuje vplyv iného neurotransmiteru, ktorým je acetylcholín s opačným účinkom, pôsobí povzbudivo. Chýbajúce tlmivé pôsobenie dopamínu, spôsobuje výkyvy pohybových funkcií, ktoré sa prejavujú práve tremorom. Tremor sa samozrejme vyskytuje aj u iných ochorení, preto nemôže byť považovaný za jednoznačný ukazovateľ diagnózy. Parkinsonský tremor má niekoľko základných charakteristík. Najčastejšie je pozorovaný na končatinách, prípadne na brade, vzácné na hlave. V počiatku sa objavuje na prstoch na hornej končatine, pričom na jednej strane tela býva výraznejší. Neskôr sa presunie na dolnú končatinu rovnakej strany tela a následne aj na druhú stranu tela. Spravidla sa tremor objavuje keď sa pacient nehýbe a ruky mu spočívajú voľne pozdĺž tela v tzv. kludovej fáze. Pri aktívnom pohybe sa tremor znižuje alebo mizne. Výraznejšie sa môže prejaviť v závislosti na zmene emočných stavov [6][5][8].

Rigidita/Svalová stuhlosť Ako väčšina hlavných príznakov Parkinsonovej choroby súvisí priamo s nedostatkom dopamínu v mozgu, respektíve v bazálnych gangliách. Aj tento príznak sa spravidla zo začiatku prejavuje iba na jednej strane ľudského tela. Prejavuje sa zvýšením svalového napätia oproti normálnemu stavu. Postihnutý sval vykazuje pri aktivite zvýšený odpor, ktorý je zo strany pacienta definovaný ako pohyb uskutočňovaný pod vodou. Už v počiatkoch choroby rigiditu sprevádzajú bolestivé stavy, pocit nepríjemného napätia alebo tlaku v oblasti ramien a chrbta. Navonok sa môže rigidita prejaviť ako vysunutie hlavy dopredu, odborne

nazývané protrúzia, poprípade sa vyskytuje aj trup v predklone. Tendencia k predklonu býva patrná pri chôdzi pacienta. Postihnutú skupinu svalstva sa pacient snaží prirodzene šetriť, čím dochádza naopak k preťažovaniu iných svalov a tým k vyššie popisovaným stavom[6][5].

Pohybové poruchy Spomalenosť, ochudobnenie a obmedzenie rozsahu pohybu, v odbornej terminológii nazývané bradykinéza, akinéza a hypokinéza. Sú nezávislé na miere rigidity. Prejavujú sa najmä v počiatočnej fáze pohybu, čiže pri snahe o začiatok chôdze z kludového stavu, pri prechode zo sedu do stoja a podobne. Taktiež sú ovplyvnené aj mimovoľné pohyby ako napríklad pohyb paží pri chôdzi, mimika tváre, reč, ale aj písmo. V niektorých prípadoch sa prejavuje náhle prerušenie pohybu a pacient nie je schopný na určitý časový úsek pohyb vykonať[6][9].

1.2 Hypokinetická dyzartria

Narušenie pohybovej zložky tvorby reči nazývame dyzartria. Porucha reči vyskytujúca sa vo väčšine prípadov ochorenia Parkinsonovou chorobou sa nazýva Hypokinetická dyzartria. Je to špecifický typ dyzartrie, ktorý sa v takej miere pri inom neurodegeneratívnom ochorení nevyskytuje. Vedecké publikácie sa líšia v percentuálnej hodnote, ktorá udáva počet Parkinsonistov postihnutých Hypokinetickou dyzartriou a to 70 až 90 %. Hypokinetická dyzartria v rôznej miere a variabilite ovplyvňuje rôzne subsystemy podieľajúce sa na tvorbe reči[11][12]. V počiatočných štádiách onemocnenia Parkinsonovou chorobou sú deficity reči nepatrné. Bola zistená aj závislosť na pohlaví pacienta. Vznik Hypokinetickej dyzartrie je samozrejme ovplyvnený rigiditou bradykinézou a tremorom, čiže sprievodnými znakmi Parkinsonovej choroby, spojené s úbytkom dopamínu. Hypokinetická dyzartria sa prejavuje predovšetkým v týchto oblastiach:

- Fonácia
- Artikulácia
- Prozódia
- Plynulosť reči
- Faciokinéza

Z hľadiska spracovania rečového signálu sa diplomová práca sústreďuje predovšetkým na prvé 4 spomenuté oblasti[13].

1.2.1 Fonácia

Ako prvá oblasť v ktorej sa Hypokinetická dyzartria prejavuje je fonácia. Nedokonalá práca hlasiviek spôsobená stuhnutím svalstva má za následok napríklad ich ne-

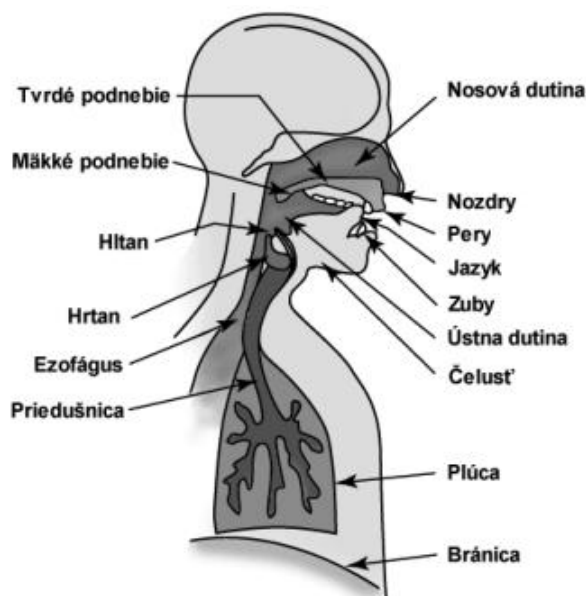
úplné uzavretie. Spolu s turbulentným prúdením z pľúc charakteristickým výrazným kolísaním, sa vo výsledku prejavuje ako hlasový tremor [14]. U Pacientov trpiacich Parkinsonovou chorobou je pozorovateľná zvýšená stredná hodnota frekvencie základného tónu F0 pri rozprávaní dlhších vetných celkov, ale aj u rozprávaní predĺžených vokálov. Taktiež bolo zistené, že s predlžujúcou sa dobou postihnutia stredná hodnota frekvencie základného tónu F0 rastie, čo znamená, že tento parameter môže byť využitý k posudzovaniu pokročilosti štádia Parkinsonovej choroby[15]. Kolísanie a rozptyl frekvencie základného tónu F0 popisovaný bežne pomocou smerodatnej odchýlky alebo jitteru nadobúda vyšších hodnôt pri predĺžených vokáloch, zapríčinených neschopnosťou udržať hrtanové svaly v stabilnej pozícii po dlhšiu dobu. Avšak z globálneho hľadiska. pri sledovaní dlhších vetných úsekov, hovoria pacienti veľmi monotónne[13].

Hypofónia a dysfónia sú ďalšie prítomné poruchy reči u pacientov. Pri hypofónii hovoríme o znížení sily hlasu, kde sa hlas javí ako tichý a nevýrazný. Dysfónia je narušenie kvality hlasu spôsobujúce drsný, zastrený a chraplavý hlas[14]. Úroveň intenzity rečového signálu je nízka a v dôsledku bradykinézy a rigidity zároveň príliš nekolíše, odpovedajúci shimmer nadobúda nízkych hodnôt[9]. Po vyzvaní pacientov k zvýšeniu hlasitosti reči, sú schopný na krátku dobu hlas zosilniť, avšak po chvíli intenzita hlasu opäť klesá.

Proces tvorby hlasu úzko súvisí s respiráciou, tvorenou dvoma fázami, a to inspirium (nádych) a expirium (výdych). Obe tieto fázy bývajú narušené a poruchy mávajú obštrukčný charakter. Sila inspiračných svalov zostáva v úvode choroby zachovaná, ale v jej priebehu sa znižuje. Pri expiračných svaloch je priebeh podobný. Klinicky je sledovaná skrátená doba oboch fáz dýchania, čo ovplyvňuje tvorbu hlasu, jeho kvalitu a silu, dĺžku prehovorenia a podobne. Hospodárenie s dychom je neefektívne, pacient hovorí v krátkych vetách, mení tempo reči, obtiažne koordinuje dýchanie a reč, nevhodne sa pri reči nadychuje, niekoľko krát opakuje začiatky slov alebo vytvára pauzy tam kde nemá a podobne. Pri prudkom uvoľnení dychového prúdu cez nosnú dutinu, môže dôjsť k hypernazalite (nezrozumiteľnosti). Hypernazalita sa v modulovom spektre reči prejavuje zdôraznením niektorých formantových frekvencií[9][14][15].

1.2.2 Artikulácia

Druhou najčastejšie narušenou oblasťou je artikulácia, prejavujúca sa najmä pri vyslovovaní spoluhlások, a to až pri 45 % pacientov[16]. Najväčšie problémy nastávali predovšetkým pri vyslovovaní plozív /p/, /t/, /b/, /d/, /g/. Z hľadiska správ-



Obr. 1.2: Hlasový trakt človeka[17]

nej výslovnosti je pre pacientov najťažšie vysloviť plosivum /k/. Je to opäť dané nesprávnou prácou rečového svalstva, v tomto prípade aktívnych artikulačných orgánov, čo sú pery, špička jazyka, stred jazyka, koreň jazyka, hrtanová príklopka a hrtan[13]. Zistené artikulačné chyby boli spôsobené predovšetkým nedostatočnou eleváciou jazyka a nesprávnym uzatváraním dýchacích ciest u niektorých hlásiek. Ľahké až ťažké poruchy artikulácie sa vyskytujú na úrovni slov, viet alebo dlhších prehovorení a prejavujú sa ako nepríjemná, nedbalá, nepresná výslovnosť s vplyvom na jej zrozumiteľnosť[9].

K hodnoteniu precíznosti artikulácie sa často využívajú diadochokinetické úlohy DDK (Diadochokinetic Tasks). Počas týchto úloh vyslovujú rečníci opakujúce sa slabiky obsahujúce kombinácie plosíva-samohláska, napríklad "pa-ta-ka-pa-ta-ka-pa-ta-ka ...". Pri vyslovení týchto kombinácií dochádza k rýchlym zmenám v postavení artikulačných orgánov. V niektorých prípadoch pacienti nie sú schopní tieto rýchle zmeny uskutočňovať, v iných prípadoch udržujú kadenciu na úkor zníženia pohyblivosti rečových ústrojenstiev, tým pádom dochádza ku zvýšenej nezrozumiteľnosti. Správne vyslovenie vokálov závisí od vhodne zvoleného objemu dutiny hrdelnej a ústnej, v ktorých dochádza k rezonanciám. Tie je možné merať pomocou formantových frekvencií, ktoré s dutinovým priestorom úzko súvisia[13].

1.2.3 Prozódia

Medzi prozodické vlastnosti reči patria prízvuk, melódia, tempo, intonácia, frázovanie, rytmus, pauzy a iné. Sú významnými nositeľmi emočného ladenia hovoriacej osoby, dodávajú vyjadreniu význam, ale aj zdôrazňujú jeho obsah. Pacienti trpiaci Parkinsonovou chorobou prejavujú deficity vo všetkých prozodických faktoroch. Reč sa stáva monotónna napríklad pri absencii prízvuku alebo melódie. Narušenie tempa reči je prejavované bradyfémioiou, čiže spomalením, ale aj tachyfémioiu - zrýchlením. Tachyfémia zhoršuje artikuláciu, má vplyv na umiestňovanie páuz v hovorenom prejave a významne ovplyvňuje zrozumiteľnosť reči[9]. Keďže prozódia človeka sa odvíja od fonácie, medzi tri základné suprasegmentálne rysy patria práve frekvencia základného tónu udávajúca výšku hlasu, intenzita a tempo reči[13]. Poruchu prozódie nazývame dysprozódia.

Niekoľko výskumov sa zaoberalo aj expresívnym aspektom ľudskej prozódie. Ich zistením je, že pacienti nedostatočne vyjadrujú aktívne emócie. Všetky aspekty dysprozódie spojené so všeobecnými poruchami reči významne prispievajú k celkovému zníženiu zrozumiteľnosti reči a dokonca aj k strate komunikačných schopností a častej sociálnej izolácií pacientov s Parkinsonovou chorobou[18].

1.2.4 Plynulosť reči

Poruchy plynulosti reči prejavujúce sa pri hypokinetickej dyzartrii, sú palilália a hezitácia. Obe choroby sa vyskytujú izolovane aj súbežne. V prípade palilálie sa jedná o zrýchlené opakovanie slabík na začiatku, uprostred alebo na konci slov, respektíve opakovanie celých slov. Hezitácia je charakterizovaná ako porušenie súvislej reči nezámernými či neplánovanými pauzami v úvode alebo kedykoľvek behom rečového prejavu[9]. Rôzne výskumy sa nezhodujú na tom, či je tempo u pacientov zvýšené alebo znížené, prípadne sa oproti tempu reči kontrolných rečníkov nemení[13].

1.2.5 Faciokinézia

V oblasti faciokinézie sú pozorovateľné poruchy orofaciálneho svalstva, ktoré má vplyv na správnu artikuláciu, prežúvanie, prehĺtanie a mimiku. Tremor v orofaciálnej oblasti sa podľa[19] vyskytuje vzácne a pokiaľ je prítomný, tak menovite je to u kruhového svalu ústneho alebo bradového svalu. Obecne je možné povedať, že tremor v orofaciálnej oblasti býva skôr súčasťou symptomatiky parkinsonských symptómov, než priamo Parkinsonovej choroby. Zjavným symptómom narušenej faciokinézie je hypomimia, negatívne ovplyvňujúca a obmedzujúca celkový výraz tváre, vyjadrovanie emócií a neverbálnu komunikáciu.[9] Ako však už bolo spomenuté, faciokinézia

nieje objektom skúmania pri analýze rečového prejavu skúmaného v diplomovej práci.

1.3 Liečba Hypokinetickej dyzartrie

Väčšina metód liečby reči u pacientov s Parkinsonovou chorobou je zameraná predovšetkým na tlmenie hlavných motorických príznakov. Najrozšírenejšia farmakologická liečba je dopamínerná liečba L-DOPA, ktorá má pozitívny vplyv približne na 80 % pacientov po dobu prvých 4 až 5 rokov užívania. Po tejto dobe sa u niektorých pacientov objavujú komplikácie ako napríklad ospalosť, zmätenosť, halucinácie, rozmazané videnie, zápcha, arytmia a iné[20][21]. Vedecké publikácie väčšinou tvrdia, že L-DOPA nemá na reč vplyv, avšak v poslednej dobe sa vedecká komunita začína zhodovať na individuálnosti vplyvu dopamínu na reč[13]. Liečba pomocou L-DOPA prestáva byť efektívna v momente, keď počet dopaminerných neurónov v substantia nigra prekročí kritickú hranicu a u pacientov sa začnú vyskytovať hybné komplikácie. Behom posledných rokov, bolo zistené, že u týchto pacientov je možné použiť neuromodulačný druh liečby. Jedná sa najčastejšie o STN-DBS (Subthalamic Nucleus Deep Brain Stimulation) - hlbokú mozgovú stimuláciu subthalamických jadier pomocou vhodne zavedených elektród napájaných zo stimulátoru umiestneného pod kľúčnou kosťou. Okrem stimulácie subthalamických jadier sa stimulujú aj iné oblasti mozgu. Po operácii je nastavená vhodná frekvencia a úroveň stimulačných impulzov tak, aby boli čo najviac potlačené príznaky tremoru, rigidity, bradykinézie a dyskinézie[22]. Aj na túto liečbu je nutné pozeráť skôr individuálne, vzhľadom na rozdielnosť vplyvov uvedených v publikáciách. Obecne sa však vedci zhodujú, že stimulácia ľavej hemisféry mozgu má na reč horší vplyv ako stimulácia pravej hemisféry[13]. Akustická analýza kvality hlasu je rýchla a neinvazívna metóda vhodná k nastaveniu priebehu stimulačných impulzov tak, aby STN-DBS mala pozitívny vplyv aj na reč a preto má do budúcnosti táto liečba veľký potenciál[23].

Duodopová pumpa a apomorfinová pumpa sú ďalšie možnosti liečby. Rovnako ako pri STN-DBS sa jedná o invazívnu liečbu. Pri duodopovej pumpke sa používa gél Duodopa, ktorý sa kontinuálne zavádza do tenkého čreva a následne je rýchlo vstrebávaný do krvného obehu. Gél obsahuje L-DOPA, takže ide aj o dopamínernú liečbu. K prenosu tejto látky sa používa pumpa umiestnená na tele pacienta. Výhodou tejto metódy je, že pacientov sa nevyskytujú neskoré hybné komplikácie. U druhého typu pumpy, čiže apomorfinovej sa kontinuálne podkožne zavádzajú dopamínerné agonisty[13]. U oboch typov púmp v dobe písania práce chýbajú publikácie, ktoré by detailnejšie opisovali vplyv liečby na hypokinetickú dyzartriú.

Ako neinvazívna metóda liečby Parkinsonovej choroby s vysokým potenciálom sa ukazuje rTMS (repetitive Transcranial Magnetic Stimulation), čiže repetitívna trasnkraniálna magnetická stimulácia. Ide o liečbu krátkymi impulzami, ktoré vytvárajú silné magnetické pole o indukcii 1 až 2 T indukujúce v mozgovom tkanive elektrický potenciál. Magnetické pole vytvára cievka, ktorá sa prikladá pacientovi na požadovanú kortikálnu oblasť. Behom jedného sedenia je mozog stimulovaný približne 2000 impulzami o frekvencii 5 až 10 Hz. Momentálne existuje iba zopár publikácií sledujúcich vplyv rTMS na hypokinetickú dyzartriu. V publikácii [25] neboli pozorované žiadne vplyvy 10 Hz rTMS nad motorickým kortexom na reč pacientov, avšak išlo o pacientov s miernou formou dyzartrie. Pozitívny vplyv 5 Hz rTMS nad primárnym motorickým kortexom, predovšetkým zlepšenie fonácie a zvýšenie rozsahu intenzity hlasu, bol pozorovaný v publikácii[26]. 10 Hz rTMS nad ľavým primárnym motorickým kortexom, opakované pri 10 sedení, má podľa [27] pozitívny vplyv na zrozumiteľnosť reči a hybnosť artikulačných orgánov, predovšetkým jazyka[13].

Ďalšia, zatiaľ veľmi okrajová metóda liečby hypokinetickej dyzartrie je transorálna injekcia kolagénu do hlasiviek. Kolagén je prirodzená súčasť hlasiviek, čiže ho tkanivá dobre vstrebávajú, čo je výhodou tejto metódy. Nevýhodou je že metóda je invazívna a nepríjemná. Pri aplikácii tejto metódy bol pozorovaný výrazný nárast rozsahu základného tónu, intenzity reči a maximálnej doby fonácie[13][28].

Doteraz najúspešnejšia metóda, zameraná priamo na liečbu hypokinetickej dyzartrie u pacientov s parkinsonovou chorobou je behaviorálna terapia LSVT (Lee Silverman Voice Treatment). Metóda bola vyvinutá logopedičkami Lorraine Ramig a Cynthia Fox v USA. Ide o individuálnu mesačnú terapiu, kde pacient 4-krát týždenne uskutočňuje sadu rečových cvičení, pri ktorých sa učí hovoriť nahlas. Učí sa kontrolovať intenzitu hlasu a správne dýchať najskôr na slovách, potom na vetách a na konci pri bežnej konverzácii. Na konci terapie pozorovali zlepšenie fonácie, prehĺtania, mimiky, prozódie a artikulácie. Pre český a slovenský jazyk zatiaľ tento druh liečby nebol upravený[13][29].

2 PREDSPRACOVANIE REČOVÉHO SIGNÁLU A PARAMETRIZÁCIA

Pri analýze rečového signálu ze účelom diagnostiky poruchy je potrebné pre jeho spracovanie rečový signál parametrizovať. Zo záznamu rečového prejavu je možné vypočítať isté hodnoty, čiže parametre signálu. Požiadavka na tieto parametre je čo najpresnejšie popisovať vlastnosti rečového signálu, ktoré sú objektom skúmania. Pri diagnostike reči sa pracuje s takými parametrami, ktoré hodnotia poškodené, respektíve postihnuté časti rečového ústrojenstva. Najčastejšie sa jedná o parametre sledujúce prácu fonačného ústrojenstva, anomálie artikulačného ústrojenstva, parametre sledujúce intonáciu reči a podobne. Čiže úlohou parametrizácie je vyjadriť rečový signál obmedzeným množstvom hodnôt. Klasická literatúra uvádza dve metódy popisu rečového signálu a to[30]:

- *neparametrický popis*, založený iba na poznatkoch o spracovaní signálu (banky filtrov, FFT a pod.)
- *parametrický popis*, založený na poznatkoch o tvorbe reči, avšak táto technika používa mnoho techník neparametrického popisu, čiže nieje jednoduché a niekedy ani žiadúce tieto dve techniky oddeliť

Vypočítané hodnoty použitím oboch týchto techník nazývame rovnako a to rečovými parametrami.

2.1 Predspracovanie rečového signálu

Rečový signál, respektíve ľudská reč je značne časovo premenná. Analýza ľudskej reči je proces s vysokou mierou zložitosti, vzhľadom nato, že rečník nikdy nevyšloví jedno slovo dvakrát úplne rovnako tak, aby dodržal rovnaký prízvuk, hlasitosť, výšku základného tónu, rýchlosť prehovorenia a podobne, nehovoriac o rozdieloch medzi jednotlivými rečníkmi. Pri zázname reči a prenose reči je signál zanesený skreslením, hlukom okolitého prostredia, inak povedané ako šum a rušenie čo má taktiež negatívny vplyv na spracovanie reči. Na tomto skreslení sa podieľajú napríklad frekvenčné charakteristiky mikrofónov, ekvalizérov, zosilňovačov, vlastnosti prenosových médií, a rôzne ďalšie technické aspekty. Vzhľadom nato že rečový signál sa spracováva výhradne v číslicovej podobe je vhodné ho pred jeho prevodom do číslicovej podoby upraviť, čiže predspracovať.[31]

Predspracovanie rečového signálu prebieha väčšinou pomocou filtrácie vstupného signálu, čo je operácia upravujúca signál tak, aby zdôraznila vybrané časti frekvenčného spektra signálu. Taktiež je jej cieľom zlepenie odstupu signálu od šumu

a podobne. Takto sa signál stáva efektívnejším a ľahšie spracovateľným. Metódy používané na predspracovanie rečového signálu sú:

- **Ustredenie** (odstránenie jednosmernej zložky)
- **Preemfáza** (zdôraznenie vyšších frekvencií)
- **Segmentácia** (rozdelenie signálu na menšie časti, tzv. segmenty)

2.1.1 Ustredenie

Jedná sa o odčítanie jednosmernej zložky, čiže strednej hodnoty signálu. Jednosmerná zložka nenesie žiadnu užitočnú informáciu, ba práve naopak, môže pôsobiť pre ďalšie spracovanie signálu rušivým vplyvom, ako napríklad pri výpočte energie signálu alebo pri výpočte počtu prechodov signálu nulovou úrovňou. Je teda vhodné ju odstrániť, ako už bolo spomenuté, jej odčítaním od signálu[30]

$$s'[n] = s[n] - \mu_s \quad (2.1)$$

kde $s'[n]$ je výsledný signál po odčítaní strednej hodnoty, $s[n]$ je pôvodný signál, a μ_s je potrebné odhadnúť.

Strednú hodnotu môžeme počítať dvoma nasledujúcimi spôsobmi:

1. **Off-line** odhad strednej hodnoty, čo je aritmetický priemer hodnôt signálu.

$$s'[n] = \frac{1}{N} \sum_{n=0}^{N-1} s[n] , \quad (2.2)$$

kde $s'[n]$ je výsledný signál po odčítaní strednej hodnoty, N je počet vzorkou pôvodného filtrovaného signálu a $s[n]$ je pôvodný signál.

2. **On-line** odhad strednej hodnoty. Používa sa vtedy, keď nemáme k dispozícii celý signál alebo ak je príliš dlhý a neustále s časom pribúda. Stredná hodnota sa dá potom odhadnúť rekurzívne

$$s'[n] = \gamma s'[n-1] + (1 - \gamma)s[n] , \quad (2.3)$$

kde γ sa blíži k 1.

2.1.2 Preemfázova filtrácia

Frekvenčná (modulová) charakteristika prostredia vplýva na rečový signál potlačením vyšších kmitočtových zložiek rečového signálu približne od 2 kHz a vyššie, významných pre zrozumiteľnosť daného signálu. Na kompenzáciu tohto javu sa v číslicovom spracovaní signálov používa filtrácia preemfázovým číslicovým filtrom

typu horná prepust' (FIR) s konečnou impulznou odozvou, ktorého prenosová funkcia je rovná[31]:

$$H_{pre}(z) = 1 - a_{pre}z^{-1} , \quad (2.4)$$

kde a_{pre} je koeficient, ktorý býva v rozsahu od 0,9 až 1.

Ide v podstate o filter prvého rádu, ktorý zdôrazňuje vyššie frekvenčné zložky rečového signálu približne o 20 dB na dekádu za účelom vyrovnania modulového frekvenčného spektra signálu. Diferenčná rovnica, ktorá odpovedá prenosovej funkcii $H(z)$ má tvar[31]:

$$y[n] = x[n] - a_{pre}x[n - 1] . \quad (2.5)$$

2.1.3 Segmentácia

Rýchlosť fyziologického pohybu artikulačných orgánov je obmedzená a preto je možné pozorovať v rečovom signále krátke úseky, v ktorých sa vlastnosti reči menia dostatočne pomaly. Rečový signál považujeme za nestacionárny, avšak na výpočet parametrov potrebujeme, aby bol signál stacionárny. Práve preto sa rečový signál spracováva metódami krátkodobej analýzy. Rečový signál sa zvyčajne delí na segmenty s dĺžkou cca 20 ms, kde sa javí ako stacionárny, avšak stále je dĺžka dostatočne veľká na presný odhad požadovaných parametrov. Táto hodnota bola stanovená podľa maximálnej rýchlosti zmeny artikulačných orgánov čo je približne od 10 po 30 ms. Segmentáciu delíme na [31]:

- **Segmentácia bez prekrývania rámcov** zaisťuje rýchly časový posun v signále, má malé nároky na pamäť a procesor, avšak hodnoty parametrov sa môžu od jedného rámca k druhému výrazne líšiť.
- **Segmentácia s prekrývaním rámcov** zaisťuje pomalý časový posuv, dostávame "vyhladené-priebehy parametrov, avšak má veľké nároky na pamäť a procesor. Výsledné parametre môžu byť rámec od rámca príliš podobné, čo je proti požiadavkám štatistickej nezávislosti.

Kompromisom je typická dĺžka prekrytia 10 ms, čiže 50 % rámca. Počet rámcov je možné vypočítať pomocou vzťahu:

$$N_{ram} = 1 + \left\lceil \frac{N - I_{ram}}{S_{ram}} \right\rceil , \quad (2.6)$$

kde N_{ram} je počet rámcov, N je dĺžka segmentovaného signálu, I_{ram} je veľkosť rámca vo vzorkách a S_{ram} je veľkosť neprekrytej časti rámca vo vzorkách.

Realizácia segmentácie, resp, tvorby rámcov sa uskutočňuje tzv. okienkovou funkciou. Pri spracovaní rečového signálu sa najčastejšie používa pravouhlé a Hammingovo okno. Pravouhlé okno definujeme nasledovne[31]:

$$w[n] = 1, \text{ pre } n = 0, 1, \dots, N - 1, \quad (2.7)$$

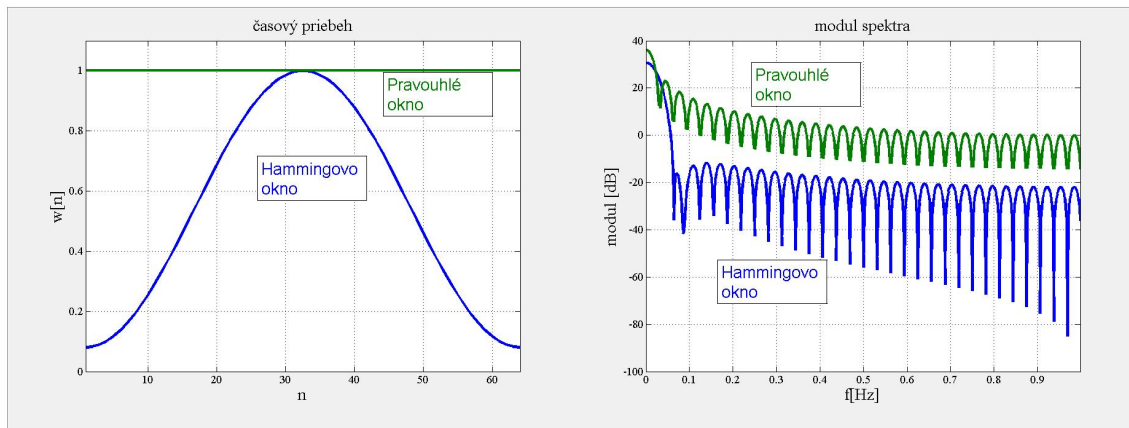
kde $w[n]$ je hodnota vzorku signálu po násobení oknom. Pre ostatné hodnoty n je $w[n] = 0$.

Hammingovo okno je rovne:

$$w[n] = 0,54 - 0,46 \cos \left[n \frac{2\pi}{N} \right], \text{ pre } n = 0, 1, \dots, N - 1, \quad (2.8)$$

kde opäť platí že pre ostatné hodnoty n je $w[n] = 0$.

Rámec vznikne vynásobením okienkovej funkcie a časovej oblasti. Zmení sa aj spektrum signálu vzhľadom na to, že násobenie v časovej oblasti znamená konvolúciu v oblasti frekvenčnej. Na nasledujúcom obrázku je znázornené pravouhlé a Hammingovo okno. U Hammingovho okna je viditeľný širší hlavný lalok. To má za následok horšie frekvenčné rozlíšenie a širšie prechodové pásmo ako u pravouhlého okna. Oproti pravouhlému oknu má však Hammingovo okno väčšie potlačenie postranných lalokov, takže má lepšie vlastnosti v nepriepustnom pásme a do spektra signálu daného okna sa nedostanú takmer žiadne spektrálne zložky z vedľajších okien/rámcov. Preto sa používa skôr Hammingovo okno ako okno pravouhlé.



Obr. 2.1: Časový priebeh Hammingovho a pravouhlého okna a ich modul spektra

2.2 Parametrizácia

Výstup parametrizácie v oblasti analýzy rečových signálov je označovaný ako rečový príznak, respektíve parameter, čo môže byť buď skalárna hodnota, vektor alebo matica. Počas niekoľkých dekád výskumu, bolo vyvinutých mnoho parametrizačných metód rečových signálov[13]. V nasledujúcej časti práce budú uvedené vybrané parametre, používané na analýzu reči a vhodné na analýzu prednesu básne. U niektorých parametrov je uvedený len anglický názov, vzhľadom na nedefinovaný oficiálny preklad. Preklad je ponechaný na čitateľovi. Pre lepšiu orientáciu sú parametre rozdelené do niekoľkých skupín, avšak niektoré sa vyskytujú vo viacerých skupinách, preto je rozdelenie iba orientačné. Skupiny sú nasledovné:

- Prozodické parametre
- Artikulačné parametre
- Fonačné parametre
- Parametre hodnotiace kvalitu signálu

Najväčší význam pre túto prácu budú mať Prozodické a Artikulačné parametre.

2.2.1 Prozodické parametre

Príznačky Hypokinetickej dyzartrie z hľadiska prozódie budú popísané nasledujúcimi parametrami, a to pre:

- **monopitch** (monotónnosť)
Frekvencia základného hlasivkového tónu F_0
- **monoloudness** (monohlasitosť)
Krátkodobá energia (STE)
Teager-Kaiser energetický operátor (TKEO)
- **tempo reči**
Total Speech Time (TST)
Total Pause Time (TPT)
Net Speech Time (NST)

Frekvencia základného hlasivkového tónu

Frekvencia základného hlasivkového tónu F_0 sa často používa pri spracovaní rečového signálu a sledovaní intonácie reči. Ide o základnú frekvenciu hlasu, tzv. *pitch*. Z hľadiska poslucháča sa jedná o výšku hlasu. V spektrálnej oblasti odpovedá frekvenciou prvej harmonickej zložky signálu. F_0 je obsiahnutá u hlások budených kvázi-periodickým kmitaním hlasiviek. Jeho diferenciálne formy výborne sledujú melódiu

rečového prejavu. Vzťah pre výpočet základného hlasivkového tónu F_0 je:

$$F_0 = \frac{f_{vz}}{k_{min}} , \quad (2.9)$$

kde f_{vz} je vzorkovací kmitočtet, a k_{min} je pozícia (vo vzorkách) prvého maxima autokorelačnej funkcie. Tento vzťah popisuje iba jednu možnosť detekcie F_0 , avšak existuje viacero postupov detekcie v časovej a spektrálnej oblasti. Frekvencia základného hlasivkového tónu F_0 je veľmi vhodný parameter na rozpoznanie zdravej a poškodenej reči. Je rôzna u detí a dospelých taktiež je rozdiel aj medzi pohlavím[31].

Krátkodobá energia

Krátkodobá energia vyjadruje energiu rečového signálu v jednotlivých segmentoch. Slúži napríklad k rozdeleniu znelých a neznelých hlások, a to tak, že u znelých hlások je energia vyššia ako u neznelých. Tento parameter sa využíva aj na rozpoznanie prechodu medzi znelým a neznelým úsekom rečového signálu. To sa dá využiť napríklad pri tzv. detektoroch rečovej aktivity, ktoré používajú jednoduchý rozpoznávač znelosti segmentu. Krátkodobá energia pre celý rečový signál teda predstavuje vektor energií pre jednotlivé segmenty a je počítaný pomocou nasledujúceho vzorca:

$$STE = \sum_{k=0}^{N-1} s[k]^2 , \quad (2.10)$$

kde $s[k]$ je vzorka diskretného rečového signálu a N je počet vzorkou segmentu.

Teager-Kaiser energetický operátor

Ide o ďalší parameter popisujúci intenzitu reči. TKEO je nelineárny operátor, poskytujúci odhad okamžitej frekvencie a amplitúdy z frekvenčnej a amplitúdovej modulácie. Energia je počítaná z celého rečového signálu a nie zo segmentov ako to je u STE. TKEO je pre diskretný čas definovaný ako:

$$\psi[s(n)] = s(n)^2 - s(n+1)s(n-1) , \quad (2.11)$$

kde $s(n)$ je vstupný rečový signál[13].

Parametre hodnotiace tempo reči

Tempo reči je možné vyjadriť rôznymi parametrami, založených na predpoklade pôsobenia choroby reči na rôzne rečové ústrojenstvá.

- **Total Speech Time - TST** je absolútne trvanie rečového signálu, obsahujúce aj tiché, napríklad medzislovné pauzy.

- **Total Pause Time - TPT** je trvanie páuz v rečovom signále, dlhších ako 10 ms.
- **Net Speech Time - NST** je trvanie čisto rečového signálu, odpovedá rozdielu medzi TST a TPT čiže[13]:

$$\text{NST} = \text{TST} - \text{TPT} . \quad (2.12)$$

2.2.2 Artikulačné parametre

Vybrané artikulačné parametre pre analýzu rečového signálu sú:

- Formantové frekvencie
- Šírky pásma formantových frekvencií

Formantové frekvencie

Počas prechodu turbulentného prúdenia hlasovým traktom dochádza v dutine hrdelnej, nosnej a ústnej k rezonanciám. Frekvencie, na ktorých dochádza k týmto rezonanciám sa nazývajú formanty (F_1, F_2, F_3). Pomocou formantových frekvencií je možné sledovať napr. hybnosť jazyka. Bolo zistené, že prvé dva formanty F_1 a F_2 sú parametre, ktoré najviac vypovedajú o produkcii vokálov[32]. Pri parametrizácii rečového signálu sa často formanty dopĺňujú o odpovedajúce šírky pásiem B_1, B_2, B_3 . Formantové frekvencie sú silno závislé na veku a pohlaví rečníka, čo je ich veľkou nevýhodou. Práve kvôli tomu niektorí výskumníci používali pomer druhých formantov samohlások **i** a **u**, F_{2i}/F_{2u} , u ktorého bolo dokázané, že je schopný nezávisle na pohlaví robustne odlíšiť zdravú reč od reči postihnutej hypokineticou dyzartriou[33][13].

2.2.3 Fonačné parametre

Parametre hodnotiace fonáciu sú napr.:

- Jitter
- Shimmer
- Pitch period entropy (PPE)

Jitter

Frekvenčná nestabilita - kolísanie dĺžky základnej periódy (prevrátenej hodnoty F_0). Perióda jednotlivých pulzov základného hlasivkového tónu F_0 sa môže líšiť aj v rámci veľmi krátkeho úseku rečového signálu. Jedná sa o parameter popisujúci tras, resp. chvenie hlasiviek (mikroprozodia). Počíta sa ako rozdiel dvoch susedných

periód hlasivkových pulzov delený priemernou dĺžkou periódy[31]:

$$J_g = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_g[i] - T_g[i-1]|}{\frac{1}{N} \sum_{i=0}^{N-1} T_g[i]}, \quad (2.13)$$

kde N je počet periód hlasivkových pulzov (segmentov), kde pre každý segment je vypočítaná základná perióda a $T_g[i]$ je i -tá hodnota základnej periódy.

Shimmer

Amplitúdová nestabilita - kolísanie amplitúdy hlasivkových impulzov. Podobne ako u Jitteru aj amplitúda hlasivkových pulzov sa môže líšiť aj v rámci veľmi krátkeho úseku rečového signálu. Jedná sa o parameter popisujúci vibrácie alebo kolísanie pulzov. Počíta sa ako rozdiel veľkostí dvoch susedných pulzov delený ich priemernou veľkosťou[31]:

$$S_g = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_g[i] - A_g[i-1]|}{\frac{1}{N} \sum_{i=0}^{N-1} A_g[i]}, \quad (2.14)$$

kde N je počet amplitúd hlasivkových pulzov a $A_g[i]$ je i -tá hodnota amplitúdy.

Pitch Period Entropy

Väčšina spomenutých parametrov má nevýhodu v závislosti na pohlaví rečníka. Práve preto Little et al.[34] navrhli parameter potlačujúci túto závislosť. Pri výpočte tohto parametra je uvažovaná logaritmická poltónová škála, inverzná filtrácia a odhad entropie[13].

2.2.4 Parametre hodnotiace kvalitu signálu

Ako parametre na hodnotenie kvality signálu budú použité:

- Harmonic-to-Noise ratio (HNR)
- Noise-to-Harmonic ratio (NHR)
- Zero-Crossing rate (ZCR)
- Spektrálny tok (SF)
- Medián výkonovej spektrálnej hustoty (MPSD)

Harmonic-to-Noise ratio / Noise-to-Harmonic ratio

Harmonic-to-Noise ratio (HNR) je odstup harmonickej zložky signálu od šumu. Pri analýze rečových signálov je HNR používaný na detekciu rečovej aktivity. Parameter porovnáva úroveň energie harmonickeho signálu s úrovňou energie šumu,

vyjadruje stupeň periodicity skúmaného priebehu. HNR počítame pomocou nasledujúceho vzorca[30]:

$$\text{HNR} = 10 \log_{10} \left(\frac{\sum_{i=1}^{\frac{N}{2}} |S_i|^2}{\sum_{i=1}^{\frac{N}{2}} |N_i|^2} \right), \quad (2.15)$$

kde $|S_i|$ je odhad energie harmonického signálu a $|N_i|$ je odhad energie šumu.

Noise-to-Harmonic ratio (NHR) je odstup šumu od harmonickej zložky signálu. Je to prevrátená hodnota HNR, čiže ju môžeme počítat ako:

$$\text{NHR} = \frac{1}{\text{HNR}}. \quad (2.16)$$

Zero-Crossing rate

Parameter Zero-Crossing rate (ZCR) popisuje počet prechodov rečového signálu nulovou úrovňou. Ide pravdepodobne o najjednoduchší parameter hodnotiaci kvalitu reči, avšak je veľmi citlivý na šum a na posun jednosmernej zložky. Na základe metód ZCR boli vypracované postupy pre určenie napr. základného hlasivkového tónu alebo približné určenie najsilnejšieho formantu[35]. Počítame ho nasledovne[30]:

$$\text{ZCR} = \frac{1}{2} \sum_{n=1}^{N-1} |\text{sign } s[n] - \text{sign } s[n-1]|, \quad (2.17)$$

kde $s[n]$ sú vzorky segmentovaného signálu, n predstavuje poradie vzorkou v segmente, N je počet vzorkou v jednom segmente a $\text{sign } s[n]$ je zjednodušená znamienková funkcia:

$$\text{sign } s[n] = \begin{cases} +1 & \text{pre } s[n] \geq 0 \\ -1 & \text{pre } s[n] < 0 \end{cases}. \quad (2.18)$$

Funkcia $|\text{sign } s[n] - \text{sign } s[n-1]|$ dáva hodnotu 2 pre každý prípad, že sa od vzorku $s[n-1]$ ku vzorku $s[n]$ zmení znamienko.

Spektrálny tok

Parameter kvantitatívne vyjadruje zmenu výkonového spektra signálu v čase. Spektrálny tok (SF) sa počíta ako normovaný rozdiel výkonového spektra aktuálneho segmentu signálu a výkonového spektra predchádzajúceho segmentu signálu. Normovaný rozdiel znamená euklidovskú vzdialenosť dvoch normovaných výkonových spektier. Parameter počítame nasledovne[36]:

$$S_{\text{flux}}(k) = \sum_{i=0}^{N-1} \sqrt{(S_k(i) - S_0)^2 (S_{(k-1)}(i) - S_0)^2}, \quad (2.19)$$

kde S_0 je normalizačný faktor, ktorý je počítaný ako aritmetický priemer odhadu strednej hodnoty výkonových spektier pre všetky segmenty daného signálu.

Medián výkonovej spektrálnej hustoty

Výkonová spektrálna hustota je parameter definovaný ako výkon signálu v určitom frekvenčnom pásme. Odhad výkonovej spektrálnej hustoty je možné vypočítať viacerými metódami, napr. použitím tzv. periodogramu. Na spracovanie rečových signálov v tejto práci bude použitá neparametrická metóda odhadu PSD, vzhľadom na rozdelenie rečového signálu na viacero vzájomne sa prekrývajúcich segmentov (max 50 %), čo odpovedá Welchovej metóde priemerovania modifikovaných periodogramov. Analyzované dáta sú rozdelené podľa[37]:

$$s_i[n] = s[n + iD] , \quad (2.20)$$

kde iD je počiatočným bodom i -tej postupnosti, $i = 0,1,\dots,L-1$. L je počet po sebe idúcich segmentov.

Modifikovaný periodogram podľa Welchovej metódy je definovaný nasledovne[37]:

$$\tilde{P}_{xx}^{(i)}(e^{j2\pi f}) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_{(i)}[n]w[n]e^{-2j\pi fn} \right|^2 , \quad (2.21)$$

kde $i = 0,1,\dots,L-1$, M je dĺžka segmentu, $w[n]$ je dĺžka okna a U je normalizačný koeficient rovný[37]:

$$U = \frac{1}{M} \sum_{n=0}^{M-1} w^2[n] . \quad (2.22)$$

Welchov odhad výkonovej spektrálnej hustoty sa následne počíta ako priemer modifikovaných periodogramov[37]:

$$P_{xx}^w(e^{j2\pi f}) = \frac{1}{L} \sum_{i=0}^{L-1} \tilde{P}_x^{(i)}(e^{j2\pi f}) . \quad (2.23)$$

Výkonová spektrálna hustota bude použitá na výpočet jej mediánu (MPSD), čo slúži na zistenie stredného výkonu v určitom pásme skúmaného priebehu.

Výpočet mediánu MPSD[36]:

- pre nepárny počet vzoriek PSD

$$\text{med} \left(P_{xx}^w(e^{j2\pi f}) \right) = P_{xx}^w(e^{j2\pi f}) \left[\frac{n+1}{2} \right] , \quad (2.24)$$

- pre párny počet vzoriek PSD

$$\text{med} \left(P_{xx}^w(e^{j2\pi f}) \right) = \frac{1}{2} \left(P_{xx}^w(e^{j2\pi f}) \left[\frac{n}{2} \right] + P_{xx}^w(e^{j2\pi f}) \left[\frac{n}{2} + 1 \right] \right) . \quad (2.25)$$

3 ŠTATISTICKÉ A KLASIFIKAČNÉ METÓDY

3.1 Štatistika

Výstupom parametrizácie, ako bolo spomenuté v predchádzajúcej kapitole, môže byť skalárna hodnota, vektor hodnôt poprípade matica. Záleží to na konkrétnom rečovom parametre. Niektoré parametre popisujú isté vlastnosti rečového signálu pre jednotlivé segmenty signálu, čo má za následok, že parameter je vektor, v niektorých prípadoch dokonca matica. Pre následnú klasifikáciu je potrebné vyjadriť štatistické vlastnosti týchto parametrov, vzhľadom na použité klasifikačné metódy. V diplomovej práci sú použité nasledujúce štatistiky, respektíve ich **odhady** (pre lepšiu neskoršiu orientáciu sú uvedené aj anglické názvy):

- **Stredná hodnota** - Mean
- **Medián** (Prostredná hodnota) - Median
- **Smerodajná odchýlka** - Standard deviation (std)
- **Rozsah** - Range
- **Relatívna smerodajná odchýlka** - Relative standard deviation
- **Relatívny rozsah** - Relative range
- **Priemerná absolútna odchýlka** - Mean absolute deviation (mad)
- **Prvý kvartil** - First quartile
- **Tretí kvartil** - Third quartile
- **Medzikvartilový rozsah** - Interquartile range (IQR)

3.2 Klasifikácia

Klasifikácia je základným problémom strojového učenia. Jedná sa o problém rozpoznania jednotlivých objektov a ich priradenie do tried. Aby bolo možné objekty správne rozpoznať je potrebné ich popísať, buď pomocou parametrov, alebo popisom štruktúry objektu. Aby mohol stroj riešiť problémy musí sa to najprv naučiť. Používajú sa dva základné prístupy učenia[42]:

- **s učiteľom** – predkladajú sa známe objekty zo známou klasifikáciou, ktoré si stroj zapamätá a je schopný podľa nich zaradiť neznámy objekt do definovanej triedy. Pri tejto metóde je riziko, že učiteľ vnesie objekt do zlej triedy a metóda je pomerne časovo náročná.
- **bez učiteľa** - stroj musí najprv odhaliť triedy objektov v parametrickom priestore, napríklad pomocou vyhľadávania zhlukov v priestore – zhluková analýza, hierarchické zhukovanie.

Existuje mnoho klasifikátorov založených na rôznych algoritmoch s rôznorodou spoľahlivosťou a úspešnosťou. V diplomovej práci sa jedná o klasifikáciu rečového signálu prednesu básne na základe vypočítaných rečových parametrov na zdravých ľuďoch a pacientov trpiacich Parkinsonovou chorobou.

3.2.1 Random forest

Metóda Random forest, čiže náhodný les, je kombinovaná učiacia sa metóda, ktorá náhodne vytvára viacero rozhodovacích stromov používaná pre klasifikáciu a predikciu. Rozhodovací strom je vývojový diagram v ktorej každý vnútorný uzol reprezentuje test na atribút, každá vetva predstavuje výsledok testu a každý koncový uzol reprezentuje označenie triedy. Cesta od koreňa po koncové lístky reprezentuje klasifikačné pravidlá. Každý uzol v rozhodovacom strome je rozdelený pomocou najlepšieho rozkolu medzi všetkými premennými, zatiaľ čo pri náhodnom lese je každý uzol rozdelený pomocou najlepšej podmnožiny prediktorov náhodne vybraných v danom uzle. Táto trochu rozporuplná stratégia sa ukázala byť veľmi výkonne porovnateľná z mnohými ostatnými klasifikátormi. Taktiež je dobre odolná voči preučeniu. Výhodou je jej užívateľská prívetivosť v tom, že obsahuje iba dva parametre a to počet premenných a počet stromov a zvyčajne nie je veľmi citlivá na ich hodnoty[43]. Každý strom rastie nasledujúcim spôsobom[44]:

1. Ak je počet prípadov v tréningovom sete rovný N , vzorka N je v náhodnom poradí od originálneho setu. Táto vzorka bude tréningovým setom pre rast stromu.
2. Ak je na vstupe M premenných, číslo $m \ll M$ je stanovené tak, že v každom uzle sú m premenné vybrané náhodne z M a najlepší rozkol týchto m je použitý

- k rozdeleniu uzla. Hodnota m je udržiavaná ako konštanta počas rastu stromu.
3. Každý strom sa pestuje v čo najväčšej možnej miere. Neexistuje žiadne prerezanie.

Chybovosť lesov závisí na dvoch veciach[44]:

- Korelácia medzi akýmkoľvek dvoma stromami v lese – zvyšovaním korelácia sa zvyšuje chybovosť lesa.
- Sila každého jednotlivého stromu v lese – strom s nízkou chybovosťou je silným klasifikátorom. Zvyšovanie sily stromu znižuje chybovosť lesa.

Znížením m sa znižuje ako korelácia tak aj sila stromu. Zvýšením sa oba parametre zvyšujú. Preto je potrebné nájsť najlepšiu hodnotu m v jej rozsahu, ktorý býva obvykle pomerne široký. Je to jediný nastaviteľný parameter, na ktorý je náhodný les citlivý.

3.2.2 Support vector machine SVM

Support vektor machine, alebo algoritmus podporných vektorov patrí k metódam strojového učenia tvoriacich kategóriu tzv. jadrových algoritmov - "kernel machines". Jadrové algoritmy vychádzajú z matematickej teórie pre nájdenie lineárnej hranice a zároveň sú schopné reprezentovať zložité nelineárne funkcie. Základný princíp je prevod daného pôvodného priestoru do iného, viacdimeziálneho priestoru, kde je možné jednotlivé triedy od seba oddeliť lineárne[45].

Problémom SVM je ako najlepšie zvoliť deliacu hranicu týchto priestorov tak, aby boli vedené čo najefektívnejšie z hľadiska kategorizácie budúcich dát, ktoré pri tréningu neboli použité. Tento problém rieši SVM nájdením polroviny, snažiacu sa o maximalizáciu okrajov voči podporným vektorom. Klasifikácia pomocou SVM je vykonávaná v dvoch fázach[36]

1. **Učenie** - každý bod o N príznakoch je umiestnený do odpovedajúceho miesta v N -rozmernom priestore, pričom má každý bod pevne definovanú triedu. Po vytvorení zhlukov dát je hľadaná deliaca hranica pomocou podporných vektorov.
2. **Klasifikácia** - každý bod vstupnej množiny je znova umiestnený do priestoru a pomocou hranice je priradený do odpovedajúcej triedy.

Okrem lineárneho jadra SVM existujú rôzne ďalšie varianty (bodové, radiálne, gausovské,...), ktoré majú rôzne parametre pre optimalizáciu klasifikácie. Vo všetkých systémoch sa však stretne s parametrom C a ξ . Výsledná zložitosť modelu závisí na oboch parametroch súčasne spolu s použitým jadrom SVM, kde[45]:

- Parameter C - určuje kompromis zložitostou modelu a mierov odchýliek väčších než ξ , ktoré sú tolerované optimalizačnou rovnicou. To znamená, ak je C príliš veľké je snaha minimalizovať riziko na základe skúsenosti, bez ohľadu na zložitost modelu.
- Parameter ξ - ovplyvňuje šírku ξ - necitlivej oblasti, ktorá je používaná pre nastavenie vzoriek tréovacích dát. To znamená že čím väčšie je ξ , tým menej je podporných vektorov, a tým pádom viac vyhladená predpoveď.

3.2.3 Validácia

Aplikovaním klasifikačných metód na dáta je vytvorený natrénovaný model pripravený na klasifikáciu nových nezávislých vzoriek. Pred tým je však potrebné vytvorený model zvalidovať, čím zistíme jeho presnosť klasifikácie. Existuje viacero validačných metód ako napríklad krížová validácia (Cross-validácia), alebo validácia zvyškov. V diplomovej práci je použitá metóda krížovej validácie.

Základnou myšlienkou krížovej validácie je rozdeliť vstupné dáta na viacero podmnožín a následne použiť na tréovanie modelu iba určitý počet podmnožín. Zvyšné podmnožiny, ktoré neboli použité pri tréovaní sa použijú na testovanie natrénovaného modelu. Existuje viacero metód krížovej validácie[46]:

- **Holdout** metóda je najjednoduchšia metóda krížovej validácie, kde sa vstupné dáta rozdelia na dve podmnožiny, tréovacia a testovacia. Na vytvorenie modelu sa používa iba tréovacia podmnožina a na testovanie iba testovacia podmnožina. Výhodou tejto metódy je vyššia presnosť ako u validácií zvyškov a krátka doba výpočtu. Nevýhodou je vysoký rozptyl vyhodnotenia, ktorý silno závisí na rozdelení tréovacej a testovacej podmnožiny.
- **K-fold** metóda je jeden zo spôsobov ako vylepšiť metódu Holdout. Vstupné dáta sú rozdelené do k podmnožín a metóda Holdout sa vykonáva k -krát, pričom sa vždy používa iba jedna podmnožina z k na testovanie a $k-1$ podmnožín je spojených do jednej podmnožiny určenej na tréovanie modelu. Výhodou K-fold metódy je menšia závislosť na rozdelení vstupných dát, keďže každý bod podmnožiny je presne jeden-krát v testovacej podmnožine a $k-1$ -krát tréovacej. Rozptyl výsledného odhadu sa znižuje s narastajúcim k . Nevýhodou je väčšia časová náročnosť výpočtu, vzhľadom na opakované spúšťanie algoritmu podľa veľkosti k . Variantou tejto metódy je náhodne rozdeliť dáta na tréovacia a testovacia podmnožina k -krát, čím je umožnené nezávisle voliť veľkosť každej testovacej podmnožiny.
- **Leave-one-out** je logický extrém K-fold metódy, kde veľkosť k je rovné celkovému počtu bodov N vstupnej množiny dát. To znamená, že model je tré-

novaný N -krát všetkými dátami okrem jedného bodu a práve týmto bodom je model testovaný. Rozptyl výsledného odhadu je nízky, avšak časová náročnosť je omnoho vyššia oproti predchádzajúcej metóde.

3.2.4 Confusion matrix

Confusion matrix (preklad je ponechaný na čitateľovi) je špeciálny typ kontingenčnej tabuľky o dvoch rozmeroch – skutočný a predpovedaný, s identickými triedami v oboch rozmeroch. Táto matica umožňuje vizualizáciu výkonu klasifikačného algoritmu, kde každý stĺpec matice reprezentuje inštanciu v predpovedanej triede, pričom každý riadok reprezentuje inštanciu v skutočnej triede (alebo naopak) [47].

		Predpoveď	
		Pozitívna	Negatívna
Skutočnosť	Pozitívna	TP	FN
	Negatívna	FP	TN

Obr. 3.1: Confusion matrix

Na obrázku 3.1 je znázornená Confusion matrix kde:

- TP je počet správnych predpovedí ktorých trieda je pozitívna.
- FP je počet nesprávnych predpovedí ktorých trieda je pozitívna.
- FN je počet nesprávnych predpovedí ktorých trieda je negatívna.
- TN je počet správnych predpovedí ktorých trieda je negatívna.

Po zostavení Confusion matrix je možné počítať rôzne hodnoty určujúce kvalitu natrénovaného modelu. V diplomovej práci budú počítané tieto hodnoty [47]:

- **Presnosť** (ACC - Accuracy) meria podiel celkového počtu predpovedí, ktoré boli správne označené:

$$ACC = \frac{TP + TN}{L}, \quad (3.1)$$

kde L je celkový počet prvkov vstupných dát.

- **Citlivosť** (SEN - Sensitivity) meria podiel pozitívnych výsledkov, ktoré sú správne označené:

$$SEN = \frac{TP}{TP + FN}. \quad (3.2)$$

- **Špecifickosť** (SPC - Specificity) meria podiel negatívnych výsledkov, ktoré sú správne označené:

$$SPC = \frac{TN}{FP + TN}. \quad (3.3)$$

3.2.5 Korelačný koeficient

Korelačný koeficient vyjadruje mieru závislosti medzi dvoma, alebo viacerými premennými, ktorého hodnota môže dosahovať hodnoty od -1 po +1, kde -1 predstavuje najvyššiu negatívnu koreláciu a +1 najvyššiu pozitívnu koreláciu. Hodnota korelácie rovná 0 znamená nezávislosť premenných. Existuje viacero typov korelačných koeficientov. V diplomovej práci budú počítané dva typy a to[48]:

- **Pearsonov** korelačný koeficient

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (3.4)$$

kde x_i, y_i sú konkrétne hodnoty dvoch vstupných vektorov, \bar{x}, \bar{y} sú priemerné hodnoty týchto vektorov a n je celkový počet prvkov vektora.

- **Spearmanov** korelačný koeficient

$$\rho = 1 - \frac{6 \sum (Rx_i - Ry_i)^2}{n(n^2 - 1)}, \quad (3.5)$$

kde Rx_i, Ry_i je poradie hodnôt x_i, y_i .

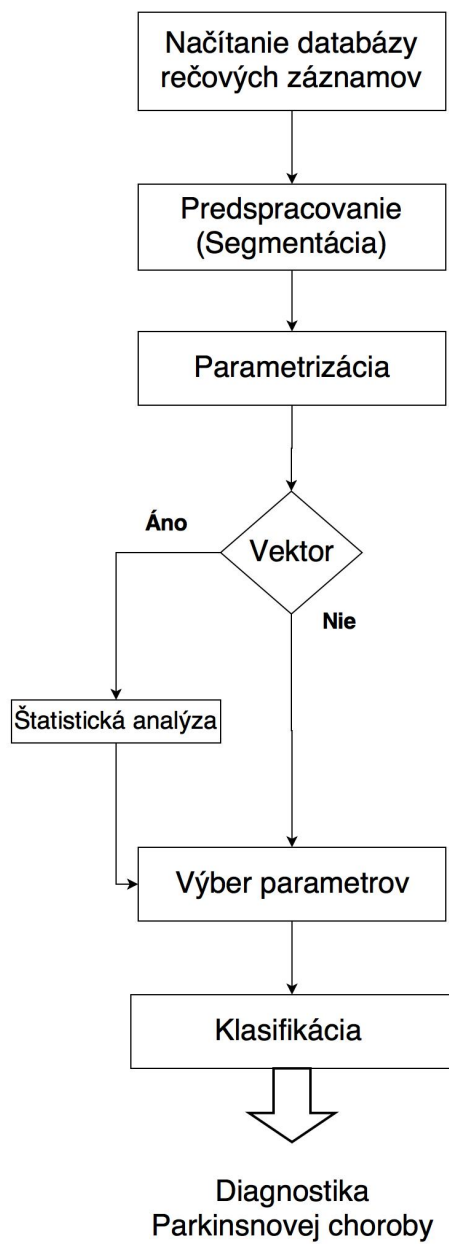
4 NÁVRH SYSTÉMU

4.1 Popis modulárneho systému

Navrhnutý systém diagnostiky Parkinsonovej choroby na základe analýzy rečového signálu prednesu básne je zložený z nasledujúcich blokov:

- **Načítanie databázy rečových záznamov** - databáza rečových záznamov pozostáva z nahrávok rečového signálu prednesu básne zdravých ľudí a ľudí trpiacich Parkinsonovou chorobou. Databáza sa nahrá do prostredia Matlab vo vhodnom formáte pre následné spracovanie.
- **Predspracovanie (Segmentácia)** - jednotlivé nahrávky načítané z databázy sa segmentujú na menšie časti. Segmentácia je na signál aplikovaná z dôvodu nestacionárnej vlastnosti rečového signálu. Signál sa delí na bloky s dĺžkou 20 ms, kde sa rečový signál javí ako kvázi-stacionárny, čo je potrebné pre jeho ďalšie spracovanie.
- **Parametrizácia** - po predspracovaní signálu sú počítané požadované rečové parametre zo vstupného signálu. Výsledkom parametrizácie môže byť buď skalárna hodnota alebo vektor hodnôt. Súčasťou bloku je rozhodovací blok, ktorý v prípade ak je výsledkom parametrizácie vektor hodnôt aplikuje štatistickú analýzu, ktorú predstavujú popisné štatistické metódy ako napr. stredná hodnota, smerodajná odchýlka, medián, medzikvartilový rozsah a iné.
- **Výber parametrov** - pre každý parameter je počítaný korelačný koeficient, na základe ktorého sú vyberané najvhodnejšie parametre pre tréovanie a validáciu systému. Výber parametrov vylučuje parametre, ktoré do systému neprinášajú informáciu vhodnú pre popis Hypokinetickej dyzartrie.
- **Klasifikácia (tréovanie a validácia)** - v konečnom bloku prebieha analýza vybraných najlepších parametrov metódou binárnej klasifikácie. Na validáciu systému je použitá metóda krížovej validácie.

4.2 Blokový diagram



Obr. 4.1: Blokový diagram systému

5 REALIZÁCIA SYSTÉMU

System automatickej detekcie Parkinsonovej choroby je implementovaný v prostredí Matlab. V nasledujúcej kapitole sa nachádza popis implementácie všetkých blokov systému a ich výstupy. Okrem vlastných zdrojových súborov sú použité v diplomovej práci na implementáciu niektorých algoritmov zdrojové súbory vytvorené inými osobami zaoberajúcich sa rovnakou respektíve podobnou problematikou. Obsah všetkých týchto použitých zdrojových súborov podlieha licencií dovoľujúcej použitie pre nekomerčné účely (tzv. open-source). Všetky vytvorené a použité zdrojové súbory sú priložené na DVD.

5.1 Načítanie databázy rečových záznamov

5.1.1 Databáza

Pre účely diplomovej práce je použitá databáza nahraných rečových signálov prednesu básne ako u pacientov trpiacich na Parkinsonovu chorobu tak aj u zdravých ľudí. Databáza bola nahrávaná na 1. neurologickej klinike Fakultnej nemocnice u Sv. Anny v Brne v Českej Republike. Rodná reč všetkých rečníkov je Český jazyk. Báseň obsahuje dva rýmy. Rečníci prečítali báseň z papiera a následne sa snažili recitovať z pamäti, avšak stále sa mohli pozeráť do papiera. Báseň v originálnom jazyku znie nasledovne:

*Chcete vidět velký lov?
Budu lovit v džungli slov.
Osledlám si pegasa,
chytím báseň do lasa.*

Databáza obsahuje celkovo 152 nahrávok, z toho je 53 zdravých a 99 pacientov. U zdravých rečníkov je 27 žien a 26 mužov, u pacientov je 40 žien a 59 mužov. Nahrávka každého rečníka obsahuje záznam prednesu básne, ktorá je súčasťou testu patologickej reči tzv. 3F test. Vzorkovacia frekvencia nahrávok $f_{vz} = 16$ kHz.

5.1.2 Načítanie databázy

Načítanie databázy rečových záznamov do prostredia Matlab realizuje prvý blok navrhnutého systému. Aby bolo možné databázu načítať, je potrebné vygenerovať súbor obsahujúci cesty s názvom `__file_list.txt` k jednotlivým nahrávkam v súborovom systéme. To je uskutočnené pomocou predpripraveného dávkového súboru v databáze `__gen_file_list.bat` obsahujúci príkaz volaný interpretom príkazového riadku

na vytvorenie požadovaného súboru. Následne je vytvorený súbor načítaný do premennej v matlabe a pomocou priloženého skriptu *load_database.m* je načítaná celá databáza. Výstupom tohto skriptu je bunkové pole obsahujúce cesty k jednotlivým záznamom a štruktúra *pat.mat* obsahujúca dva atribúty a to rečový signál y ako vektor a vzorkovaciu frekvenciu fs rečového signálu v Hz.

5.2 Predspracovanie (Segmentácia)

Segmentácia rečového signálu je aplikovaná z dôvodu nestacionárnej vlastnosti rečového signálu. Ako bolo uvedené v kapitole 2.1.3, rečový signál sa delí na segmenty o dĺžke 20 ms, kde sa signál javí ako stacionárny a je možné počítať vybrané parametre. Na rečový signál sa aplikuje segmentácia s prekrývaním rámcov o dĺžke prekrytia 10 ms, čiže 50 % zvolenej dĺžky jedného segmentu. Pomocou priloženého skriptu *segmentation.m*, ktorý obsahuje funkciu *segmentation* je v prostredí matlab aplikovaná segmentácia na rečový signál. Vstupom funkcie je požadovaný signál, vzorkovacia frekvencia, dĺžka okna a veľkosť prekrytia. Dĺžka okna má východziu hodnotu nastavenú na 20 ms a veľkosť prekrytia na 10 ms. Výstupom funkcie je matica s veľkosťou $[počet\ vzorkou\ v\ segmente] \times [počet\ segmentov]$. Táto funkcia je volaná v skripte *load_database.m*, po načítaní databázy a jej výstup je pripísaný do štruktúry *pat.mat* pod novým atribútom *seg*.

5.3 Parametrizácia

V predchádzajúcom kroku bola vytvorená štruktúra, obsahujúca rečový signál, jeho vzorkovaciu frekvenciu a maticu rozsegmentovaného signálu, pripravená pre výpočet parametrov. K tomu slúži skript *parametrization.m*, ktorého výstupom je štruktúra obsahujúca vypočítané všetky parametre pre každý rečový signál.

5.3.1 Dekódovanie názvu

V prvom rade je dekodovaný klinický stav a pohlavie rečníka z názvov jednotlivých nahrávok. Dekódovanie je uskutočnené postupným čítaním každého znaku názvu. Logika názvu je nasledovná:

- Prvým znakom v názve je písmeno reprezentujúce klinický stav. Zdravého rečníka reprezentuje písmeno **K**, rečníka trpiaceho Parkinsonovou chorobou reprezentuje písmeno **P**.
- Druhým znakom je číslo reprezentujúce pohlavie rečníka. Číslo 1 reprezentuje ženské pohlavie, číslo 2 mužské.

- Zvyšné znaky v názve sú číslice označujúce poradie rečníka a verziu cvičenia. Tieto znaky nie sú v štruktúre dekodované.

Vytvorí sa nová štruktúra *matrix_of_param.mat*, ktorej prvé tri atribúty sú označenie rečníka *participant_label*, klinický stav rečníka *clinical_state* a jeho pohlavie *gender*.

5.3.2 Výpočet parametrov

V predchádzajúcom kroku, bola vytvorená štruktúra, pripravená na zapisovanie výsledkov jednotlivých parametrov. Každá výsledná hodnota pre daný parameter je zapísaná ako nový atribút štruktúry *matrix_of_param.mat*.

STE, je počítaný priamo v skripte *parametrization.m*. Jeho výpočet je popísaný v kapitole 2.2.1. Výsledkom je vektor hodnôt, z čoho vyplýva, že je potrebné aplikovať tzv. post-processing parametra pomocou popisných štatistík popísaných v kapitole 3.1.

TKEO, pre jeho výpočet, ktorý je uvedený v kapitole 2.2.1, je vytvorený skript *tkeo.m*, ktorý obsahuje funkciu počítajúcu tento parameter. Požadovaným vstupom funkcie je rečový signál, výstupom funkcie je vektor, na ktorý je aplikovaný post-processing.

ZCR, na jeho výpočet, popísaný v kapitole 2.2.4, je vytvorený skript *ZCR.m* obsahujúci funkciu, ktorej vstupom je rozsegmentovaný rečový signál a výstupom je počet prechodu signálu nulovou úrovňou.

F₀, výpočet je vykonávaný pomocou programu **praatcon.exe** volne dostupného na http://www.fon.hum.uva.nl/praat/download_win.html. Parameter je počítaný pomocou skriptu *praat_F0.m*, ktorého požadovaný vstup je rečový signál a vzorkovacia frekvencia. Tento skript volá program **praatcon.exe** do konzoly operačného systému spolu z rôznymi parametrami potrebnými pre výpočet F_0 . Výstupom je vektor, na ktorý je následne aplikovaný post-processing.

PPE, popis parametru na nachádza v kapitole 2.2.3. Jeho výpočet je implementovaný pomocou skriptu *pitch_period_entropy.m* [34], obsahujúci funkciu ktorej vstupom je F_0 a rád lineárneho odhadu, ktorý je predvolene nastavený na hodnotu 12. Výstupom je skalárna hodnota PPE.

Jitter, Shimmer, HNR/NHR, tieto parametre sú počítané pomocou skriptu *praat_info.m*, ktorý opäť volá program **praatcon.exe**. Funkcia obsiahnutá v skripte má požadovaný vstup rečový signál a vzorkovaciu frekvenciu. Výstupom funkcie je matica viacerých parametrov popisujúcich vstupný signál. Z tejto matice parametrov sú vybrané požadované parametre a zapísané do štruktúry *matrix_of_param.mat*. Popis týchto parametrov sa nachádza v kapitolách 2.2.3 a 2.2.4.

MPSD, popis jeho výpočtu je uvedený v kapitole 2.2.4. Pre jeho výpočet je použitý skript *MPSD.m*, obsahujúci funkciu, ktorej vstupom je rečový signál a voliteľná hodnota dĺžky FFT, predvolene nastavenej na dĺžku vstupného rečového signálu. Výstupom funkcie je hodnota MPSD.

Formantové frekvencie a šírky ich pásiem, sú počítané pomocou skriptu *praat_formants.m*, ktorý volá program **praatcon.exe**. Ich popis sa nachádza v kapitole 2.2.2. Vstupom pre výpočet je rečový signál a vzorkovacia frekvencia. Výstupom je matica obsahujúca prvú až tretiu formantovú frekvenciu a šírky ich pásiem, pre každý segment vstupného signálu. Na tento výstup je aplikovaný post-processing.

SF, pre výpočet SF je použitý skript *spectrum_flux.m*, obsahujúci funkciu, ktorej požadovaným vstupom je rečový signál a jeho vzorkovacia frekvencia. Výstupom je vektor vypočítaných hodnôt, na ktorý je aplikovaný post-processing. Popis výpočtu SF sa nachádza v kapitole 2.2.4.

TST, TPT, NST, ich popis je uvedený v kapitole 2.2.1. TST je počítaný v sekundách ako celkový počet vzorkou rečového signálu, podelený vzorkovacou frekvenciou. Pre výpočet NST bol použitý skript *detectVoiced.m*[49], ktorý slúži na odstránenie tichých miest z rečového signálu. Jeho vstupom je rečový signál a výstupom je vektor obsahujúci rečový signál bez tichých miest. Dĺžka signálu v sekundách je počítaná rovnako ako pri TST. TPT je rozdiel TST a NST.

5.3.3 Výpočet štatistických hodnôt

Ako už bolo spomenuté výsledkom parametrizácie môže byť buď skalárna hodnota, alebo vektor. Ako je to vektor aplikujú sa naň štatistické metódy popísané v kapitole 3.1. V matlabe sú pre všetky použité štatistické metódy, okrem relatívnej hodnoty smerodajnej odchýlky a rozsahu, už obsiahnuté funkcie, takže nie je potrebné implementovať ich výpočet. Stačí ich jednoducho zavolať:

- **mean (parameter)** vráti odhad strednej hodnoty,
- **median (parameter)** vráti odhad mediánu,

- **std** (**parameter**) vráti odhad smerodajnej odchýlky,
- **range** (**parameter**) vráti odhad rozsahu ,
- **mad** (**parameter**) vráti odhad priemernej absolútnej odchýlky,
- **prctile** (**parameter**, **25**) vráti odhad prvého kvartilu,
- **prctile** (**parameter**, **75**) vráti odhad tretieho kvartilu,
- **iqr** (**parameter**) vráti odhad medzikvartilového rozsahu.

Každá štatistická hodnota vypočítaná pre jednotlivý parameter je zapísaná do štruktúry *matrix_of_param.mat* ako nový parameter. Štatistická analýza je aplikovaná priamo po výpočte jednotlivých parametrov v skripte *parametrization.m*. Parametre na ktoré sa aplikuje štatistická analýza sú nasledovné:

- STE
- TKEO
- F_0
- SF
- Prvé tri formantové frekvencie a ich šírky pásiem

Výsledkom parametrizácie a aplikovaním štatistickej analýzy na vektorové parametre je kompletná štruktúra *matrix_of_param.mat* s celkovým počtom parametrov 110. Ukážka štruktúry je na obrázku 5.1:

matrix_of_param

1x152 struct with 113 fields

Fields	parti	clir	gen	mean_STE	median_STE	std_STE	std_rel_STE
1	'K1003'	'OK'	'female'	0.0041	0.0028	0.0045	109.7938
2	'K1004'	'OK'	'female'	0.0027	0.0014	0.0039	143.7919
3	'K1005'	'OK'	'female'	0.0056	0.0020	0.0086	153.2034
4	'K1006'	'OK'	'female'	0.0062	0.0026	0.0086	138.6168
5	'K1007'	'OK'	'female'	0.0030	0.0014	0.0044	146.1541
6	'K1012'	'OK'	'female'	0.0036	0.0019	0.0048	132.2267
7	'K1017'	'OK'	'female'	0.0335	0.0170	0.0434	129.7161
8	'K1018'	'OK'	'female'	0.0250	0.0145	0.0338	135.0870
9	'K1019'	'OK'	'female'	0.0412	0.0155	0.0827	200.7959
10	'K1020'	'OK'	'female'	0.0194	0.0113	0.0247	127.1085
11	'K1021'	'OK'	'female'	0.0203	0.0121	0.0283	139.3245
12	'K1022'	'OK'	'female'	0.0464	0.0208	0.0610	131.5965
13	'K1023'	'OK'	'female'	0.0151	0.0098	0.0164	108.9107
14	'K1024'	'OK'	'female'	0.0541	0.0193	0.1025	189.4945
15	'K1025'	'OK'	'female'	0.0271	0.0150	0.0373	137.8392
16	'K1026'	'OK'	'female'	0.0321	0.0156	0.0536	166.8942
17	'K1029'	'OK'	'female'	0.0244	0.0167	0.0267	109.6007
18	'K1030'	'OK'	'female'	0.0282	0.0169	0.0308	109.2315

Obr. 5.1: Ukážka štruktúry *matrix_of_param.mat*, obsahujúca všetky počítané parametre pre každého rečníka.

5.3.4 Individuálna klasifikácia

Po vytvorení štruktúry, pristupujeme ku samotnej klasifikácii nami vypočítaných parametrov a následnej validácii. K tomu je vytvorený skript `v`, v ktorom sa okrem klasifikácie a validácie vykonáva výpočet *confusion matrix* a korelačného koeficientu. Výstupom je štruktúra, ktorá obsahuje názov parametru a k nemu vypočítané ACC, SEN, SPC a korelačné koeficienty. Na klasifikáciu sú použité dve klasifikačné metódy, a to Random forest a SVM, ktorých popis je ubedený v kapitole 3.2.1. Oba modely sú tréňované s Cross-validáciou.

Random forest V Matlabe je táto metóda už implementovaná pod názvom *fitctree*. Vo vytvorenom skripte *clasification_RF.m* sa pre každý parameter volá funkcia *fitctree* nasledovným spôsobom:

```
fitctree(X, Y, 'CrossVal', 'on'),
```

kde `X` je vektor obsahujúci vypočítané hodnoty jedného parametra pre všetkých rečníkov, `Y` je vektor obsahujúci klinický stav rečníkov a `'CrossVal','on'` je príznak aplikujúci Cross-validáciu, s predvolenou hodnotou rozdelenia vstupných dát na 10 podmnožín(10-fold validation).

Po natrénovaní modelu pre každý parameter je model testovaný pomocou funkcie *kfoldPredict*, ktorá je taktiež už implementovaná v Matlabe. Jej vstupom je natrénovaný model daného parametra, a výstupom je vektor odhadu klinického stavu rečníka na základe hodnoty parametra. Funkcia *kfoldPredict* si sama automaticky prečíta z modelu, akou metódou bol natrénovaný a patrične k danej metóde model otestuje.

Po každom testovaní je počítaná confusion matrix pomocou funkcie *confusionmat*, ktorá je taktiež už implementovaná v Matlabe. Vstupom tejto funkcie je vektor odhadnutého klinického stavu pre daný parameter a vektor obsahujúci skutočný klinický stav rečníkov. Výstupom funkcie je matica obsahujúca jednotlivé hodnoty confusion matrix, z ktorých sú následne počítané hodnoty ACC, SEN a SPC, výpočet ktorých je uvedený v kapitole 3.2.4. ACC je počítaná pomocou skriptu *classification_accuracy.m*, ktorého vstupom sú TP, FP, FN a TN, výstupom je hodnota ACC pre daný model parametru. SEN je počítaná pomocou skriptu *classification_sensitivity.m*, ktorého vstupom sú TP a FN a výstupom je hodnota SEN pre daný model parametru. A SPC je počítaná pomocou skriptu *classification_specificity.m*, ktorého vstupom sú FP a TN a výstupom je hodnota SPC pre daný model parametru.

SVM Klasifikácia pomocou metódy SVM je implementovaná obdobne ako Random forest. Implementácia SVM sa nachádza v skripte *clasification_SVM.m*. Jediný rozdiel oproti Random forest je vo volaní klasifikačnej metódy. Pre SVM je v matlabe taktiež už vytvorená funkcia *fitcsvm*, ktorá sa pre každý parameter volá nasledovným spôsobom:

```
fitcsvm(X,Y,'CrossVal','on','KernelFunction','RBF'),
```

kde 'KernelFunction','RBF' je znak pre použitie radiálneho jadra SVM. Všetok ostatný proces je identický ako pri metóde Random forest.

Celý proces klasifikácie pri oboch metódach je opakovaný 5-krát, z dôvodu zníženia pravdepodobnosti omylu vzhľadom na rozdielnosť hodnôt pri každom tréningu. Vzhľadom na použitú metódu klasifikácie s rozdelením vstupných dát na 10 podmnožín je klasifikácia vykonávaná 50-krát pre jeden parameter. 5-krát je vykonávané testovanie a počítanie ACC, SEN a SPC. Výsledok týchto hodnôt sa sprisporuje a zapíše sa do výslednej štruktúry *clasifi_score.mat*, ktorej prvým atribútom je názov parametra nasledovaný ACC, SEN a SPC pre daný parameter.

Následne je počítaný korelačný koeficient pre každý parameter. Na výpočet korelačného koeficientu je použitá funkcia Matlabu *corr*, a to nasledovne:

```
corr(X,Y,'type','Pearson'),
```

kde X je vektor obsahujúci vypočítané hodnoty jedného parametra pre všetkých rečníkov, Y je vektor obsahujúci klinický stav rečníkov pretransformovaný do binárnej podoby (0=zdravý, 1=chorý) a 'type','Pearson' je príznak určujúci typ korelačného koeficientu. V diplomovej práci sú počítané dva typy, a to Pearsnov a Spearmanov korelačný koeficient, ktorých výpočet je uvedený v 3.2.5. Absolútne hodnoty vypočítaných korelačných koeficientov spolu s hodnotou *p* sú zapísané do výslednej štruktúry *clasifi_score.mat*. Vyhodnotenie individuálnej analýzy sa nachádza v kapitole 6.1.

5.3.5 Klasifikácia viacerých parametrov

Po individuálnej klasifikácii je použitá klasifikácia viacerých parametrov zároveň. K výberu najlepších parametrov je použitý algoritmus mRMR[50]. Tento algoritmus vyvinul popredný výskumník v oblasti analýzy patologickej reči Athanasios Tsanas. Algoritmus je použitý v diplomovej práci kvôli jeho výberu parametrov na základe Spearmanovho korelačného koeficientu. V matlabe je tento algoritmus implementovaný ako funkcia *mRMR_Spearman.m*, ktorej vstupom je matica všetkých

parametrov a vektor klinického stavu rečníkov. Výstupom je vektor v ktorom sa nachádzajú ohodnotené parametre od 1 po N, kde 1 znamená najlepšie hodnotený parameter.

Pre klasifikáciu viacerých parametrov zároveň sú vytvorené dva skoro identické skripty *feature_selection_mRMR_RF.m* a *feature_selection_mRMR_SVM.m*, pre obe klasifikačné metódy. Rozdiel je vo volaní inej klasifikačnej metódy. V skripte je vytvorená štruktúra s názvom a ohodnotením daného parametru pomocou algoritmu mRMR. Táto štruktúra je následne zoradená od najlepšie hodnoteného parametru až po najhorší. Pre klasifikáciu skupiny parametrov zároveň sú vytvorené dve funkcie *clasification_mRMR_RF.m* a *clasification_mRMR_SVM.m*, ktorých vstupom je štruktúra zoradených a ohodnotených parametrov *best_features.mat* a výstupom je ACC, SEN a SPC pre danú skupinu parametrov. Funkcia volá na klasifikáciu jednu z klasifikačných metód, buď *fitctree*, alebo *fitcsvm*, rovnakým spôsobom ako pri individuálnej analýze, akurát vstupom je matica obsahujúca skupinu parametrov. Natrénovanie a testovanie modelu sa vykonáva 5-krát a pri každom testovaní sa počítajú hodnoty ACC, SEN a SPC, ktoré sa potom spriemerujú. V skripte *feature_selection_mRMR_RF.m(..._SVM.m)* je ďalej vytvorený algoritmus pre nájdenie najlepšieho počtu parametrov. Tento algoritmus volá funkciu *clasification_mRMR_RF.m(..._SVM.m)*, jej výsledok zapíše do štruktúry *best_mRMR.mat*, v ktorej je zapísaný počet parametrov v skupine a ich ACC, SEN a SPC. Potom vymaže najhoršie hodnotený parameter a znova prebieha proces klasifikácie. To sa opakuje až kým nezostane v skupine jeden parameter. Výstupom tohto skriptu je teda štruktúra *best_mRMR.mat*, ktorá obsahuje počet parametrov v skupine (110 až 1) a hodnoty ACC, SEN a SPC pre danú skupinu. Vyhodnotenie analýzy viacerých parametrov je uvedené v kapitole 6.2.

6 VYHODNOTENIE

6.1 Individuálna analýza

Parametre sú vyhodnocované na základe výsledných hodnôt ACC, SEN a SPC pre jednotlivé parametre. Ako je vidno v nasledujúcich tabuľkách, hodnoty individuálnej analýzy sú pomerne nízke a v praxi nepoužiteľné. Vybraných je 10 najlepších parametrov pre každú hodnotu, podľa danej klasifikačnej metódy. z výsledného hodnotenia je odstránený parameter 1. kvartil F_0 z dôvodu nulových hodnôt pre všetkých rečníkov.

ACC - v tabuľke 6.1 je 10 najlepších parametrov zoradených podľa ACC, z toho sú:

Random forest	SVM
5x artikulačné - formanty	4x artikulačné - formanty
2x prozodické - STE, NST	3x prozodické - STE, F_0
2x fonačné - jitter, shimmer	1x fonačné - jitter
1x hodnotiace kvalitu signálu - SF	2x hodnotiace kvalitu signálu - SF, HNR

Tab. 6.1: Najlepšie hodnotené parametre podľa ACC

Random forest		SVM	
Názov	ACC[%]	Názov	ACC[%]
relatívna std 1. formantu	66,58±0,9	relatívna std 1. formantu	72,63±1,9
1. kvartil STE	64,87±0,7	std SF	72,37±0,9
std SF	64,74±1,8	jitter	71,18±2,1
shimmer	64,34±1,8	relatívna std 3. formantu	68,95±2,2
relatívna std 3. formantu	64,08±2,6	1. kvartil STE	68,42±1,5
NST	63,55±0,9	1. kvartil 2. formantu	67,50±0,8
mad šír. p. 2. formantu	63,55±1,4	relatívna std STE	67,50±1,8
jitter	63,03±1,4	HNR	66,58±2,3
3.kvartil šír. p. 1. formantu	61,58±1,4	IQR F_0	66,18±1,3
medián 2. formantu	61,44±1,9	mad šír. p. 1. formantu	66,05±1,4

Pri obidvoch použitých klasifikačných metódach, je najlepšie hodnotený parameter, ktorý popisuje variabilitu 1. formantu. Tento parameter vyjadruje obmedzenosť pohybu rečových orgánov, spôsobenú svalovou rigiditou, čo je bežným príznakom Parkinsonovej choroby. SF je taktiež jedným z najlepšie hodnotených parametrov. SF popisuje kvalitu hlasu, čo len potvrdzuje obmedzenosť pohybu rečových orgánov,

u pacientov s Parkinsonovou chorobou. Hodnota ACC je pri klasifikačnej metóde SVM vyššia v priemere cca o 5 % oproti metóde Random forest.

SEN - najlepších 10 parametrov podľa SEN je zobrazených v tabuľke 6.2, z toho sú:

Random forest	SVM
5x artikulačné - formanty	5x artikulačné - formanty
3x prozodické - STE, NST	3x prozodické - STE, F_0
1x fonačné - shimmer	1x fonačné - jitter
1x hodnotiace kvalitu signálu - SF	1x hodnotiace kvalitu signálu - SF

Tab. 6.2: Najlepšie hodnotené parametre podľa SEN

Random forest		SVM	
Názov	SEN[%]	Názov	SEN[%]
1. kvartil STE	51,32±2,8	relatívna std 1. formantu	52,08±5,4
mad šír. p. 2. formantu	50,19±3,9	std SF	37,74±2,3
relatívna std 1. formantu	47,17±2,8	1. kvartil STE	33,21±4,5
std SF	46,42±2,9	relatívna std 3. formantu	32,08±3,5
NST	44,91±2,1	1. kvartil 2. formantu	31,70±1,6
shimmer	43,40±1,9	jitter	31,70±3,4
mad šír. p. 1. formantu	42,64±2,5	mad šír. p. 1. formantu	30,94±2,9
medián 2. formantu	40,38±4,1	1. kvartil 1. formantu	29,81±3,1
iqr STE	40,00±4,5	relatívna std STE	27,17±3,2
relatívna std 3. formantu	40,00±6,3	IQR F_0	26,04±3,1

Hodnota SEN je pri SVM nižšia v priemere cca o 13 % oproti metóde Random forest. Táto hodnota je obzvlášť kritická, vzhľadom nato, že vypovedá o počte chorých označených správne za chorých. Hodnota SEN pre praktické použitie musí byť čo najvyššia. V diplomovej práci je však brané do úvahy radenie podľa ACC a to z dôvodu porovnateľnosti výsledkov s inými prácami, kde sa väčšinou hodnoty SEN a SPC ani neuvádzajú.

SPC - podľa SPC je 10 najlepších parametrov zobrazených v tabuľke 6.3, z toho sú:

Random forest	SVM
3x artikulačné - formanty	3x artikulačné - formanty
3x prozodické - STE, TST, NST	3x prozodické - STE, F_0
2x fonačné - jitter, shimmer	1x fonačné - jitter
2x hodnotiace kvalitu signálu - SF	3x hodnotiace kvalitu signálu - SF, HNR

Pri hodnotení podľa SPC je možné pozorovať vyrovnanie počtu parametrov podľa jednotlivých kategórií, čo naznačuje vhodnosť kombinovania jednotlivých kategórií parametrov počas klasifikácie. Hodnota SPC je pri SVM vyššia v priemere o cca 15 % oproti metóde Random forest.

Tab. 6.3: Najlepšie hodnotené parametre podľa SPC.

Random forest		SVM	
Názov	SPC[%]	Názov	SPC[%]
relatívna std 1. formantu	76,97±0,6	jitter	92,32±1,5
relatívna std 3. formantu	76,97±2,7	rozsah SF	92,12±2,2
shimmer	75,55±3,3	rozsah F_0	91,11±1,8
jitter	75,55±0,7	std SF	90,91±1,6
iqr SF	74,55±1,7	HNR	90,07±2,6
rozsah STE	74,55±2,3	rel. roz. šír. p. 2. formantu	90,30±1,5
3. kvartil šír. p. 1. formantu	74,55±2,5	relatívna std STE	89,09±1,3
std SF	74,55±1,8	mad šír. p. 3. formantu	88,89±2,6
TST	73,74±1,1	relatívna std 3. formantu	88,69±1,8
NST	73,53±0,9	rozsah STE	88,69±1,8

V tabľke 6.4 je uvedených 10 najlepších parametrov klasifikovaných metódou Random forest z hodnotami ACC, SEN, SPC, Pearmanov korelačný koeficient a jeho p hodnota a Spearmanov korelačný koeficient a jeho p hodnota. Zoradené sú podľa ACC. V tabľke 6.5 je uvedených 10 najlepších parametrov klasifikovaných metódou SVM.

Z výsledkov individuálnej analýzy parametrov je zrejmé, že pre zvýšenie úspešnosti analýzy Parkinsonovej choroby je potrebné kombinovať jednotlivé parametre, resp. nájsť najvhodnejšiu kombináciu týchto parametrov. Dominujú artikulačné parametre, čo mení počiatočný predpoklad dôležitosti prozodických parametrov pri akustickej analýze prednesu básne. Nevýhodou artikulačných parametrov je ich závislosť na pohlaví, čo značí o vhodnosti rozdelenia rečníkov na mužov a ženy. Ďalším výsledkom individuálnej analýzy je porovnanie klasifikačných metód Random forest a SVM. Metóda SVM sa javí ako vhodnejšia na základe vyšších hodnôt ACC(cca o 5 %) a SPC(cca o 15 %), avšak hodnota SEN je pri SVM nižšia(cca o 13 %). Vzhľadom na výpovedný charakter hodnoty SEN je táto hodnota pre praktické účely najsignifikantnejšia. Z praktického hľadiska nie až taký veľký problém, ak je zdravý rečník označený za chorého, lebo následné medicínske testy

ukázu jeho zdravotný stav. Avšak je veľmi nežiadúce, aby bol chorý rečník označený za zdravého. Z toho hľadiska je teda pri individuálnej analýze úspešnejšia klasifikačná metóda Random forest. Nízke hodnoty korelačných koeficientov pre oba počítané druhy korelácie, tj. Spearmanov koeficient a Pearsonov koeficient poukazujú na len nepatrnú monotónnu, respektíve lineárnu závislosť medzi hodnotami jednotlivých rečových parametrov a klinickým stavom rečníkov. Navyše, korelačné koeficienty uvádzané v tabuľkách nedosahujú dostatočnú štatistickú signifikantnosť ($p < 0.05$), okrem parametra *shimmer*.

Tab. 6.4: Prehľad najlepších parametrov individuálnej analýzy klasifikovaných metódou Random forest. V tabuľke je uvedený názov parametra, ACC, SEN, SPC, Pearmanov korelačný koeficient(PK) a jeho p hodnota(pPK) a Spearmanov korelačný koeficient(SK) a jeho p hodnota(pSK).

Názov	ACC [%]	SEN [%]	SPC [%]	PK	p PK	SK	p SK
relatívna std 1. formantu	66,58±0,9	47,17±2,8	76,97±0,6	0,05	0,51	0,05	0,54
1. kvartil STE	64,87±0,7	51,32±2,8	72,12±1,4	0,15	0,07	0,11	0,17
std SF	64,74±1,8	46,42±2,9	74,55±1,8	0,06	0,44	0,08	0,35
shimmer	64,34±1,8	43,40±1,9	75,56±3,3	0,18	0,03	0,14	0,08
relatívna std 3. formantu	64,08±2,6	40,00±6,3	76,97±2,7	0,01	0,94	0,01	0,9
NST	63,55±0,9	44,91±2,1	73,54±0,9	0,13	0,12	0,15	0,06
mad šírky pásma 2. formantu	63,55±1,4	50,19±3,9	70,71±2,7	0,09	0,29	0,1	0,23
jitter	63,03±1,4	39,62±4,5	75,56±0,7	0,09	0,26	0,05	0,57
3. kvartil šírky pásma 1. formantu	61,58±1,4	37,36±3,9	74,55±2,5	0,07	0,38	0,06	0,43
medián 2. formantu	61,45±1,9	40,38±4,1	72,73±1,9	0,04	0,63	0,02	0,85

Tab. 6.5: Prehľad najlepších parametrov individuálnej analýzy klasifikovaných metódou SVM. V tabuľke je uvedený názov parametra, ACC, SEN, SPC, Pearmanov korelačný koeficient(PK) a jeho p hodnota(pPK) a Spearmanov korelačný koeficient(SK) a jeho p hodnota(pSK).

Názov	ACC [%]	SEN [%]	SPC [%]	PK	p PK	SK	p SK
relatívna std 1. formantu	72,63±1,9	52,08±5,4	83,64±2,5	0,05	0,51	0,05	0,54
std SF	72,37±0,9	37,74±2,3	90,91±1,6	0,06	0,44	0,08	0,35
jitter	71,18±2,1	31,70±3,4	92,32±1,5	0,09	0,26	0,05	0,57
relatívna std 3. formantu	68,95±2,2	32,08±3,5	88,69±1,8	0,01	0,94	0,01	0,9
1. kvartil STE	68,42±1,5	33,21±4,5	87,23±1,8	0,15	0,07	0,11	0,17
1. kvartil 2. formantu	67,50±0,8	31,70±1,6	86,67±1,1	0,06	0,45	0,04	0,65
relatívna std STE	67,50±1,8	27,17±3,2	89,09±1,3	0,04	0,65	0,03	0,73
HNR	66,58±2,3	21,51±2,5	90,07±2,6	0,22	0,01	0,24	0,01
IQR F0	66,18±1,3	26,04±3,1	87,68±1,7	0,12	0,15	0,1	0,21
mad šírky pásma 1. formantu	66,05±1,4	30,94±2,9	84,85±2,0	0,03	0,68	0,03	0,72

6.2 Analýza viacerých parametrov

Výsledkom analýzy viacerých parametrov zároveň je skupina parametrov s najlepšou hodnotou ACC, kde parametre boli hodnotené pomocou algoritmu mRMR. Tabuľka 6.6 ukazuje prvých 5 najlepších skupín klasifikovaných metódou Random forest a ich hodnoty ACC, SEN a SPC.

Tab. 6.6: Skupiny parametrov klasifikované metódou Random forest zoradené podľa ACC, v ktorých poradie parametrov je hodnotené podľa algoritmu mRMR.

Random forest			
Počet parametrov v skupine	ACC [%]	SEN [%]	SPC [%]
36	70,66±2,6	59,25±6,3	76,77±2,5
30	70,00±3,5	59,25±2,9	75,76±4,7
64	69,61±3,6	60,75±8,9	74,35±4,6
32	69,08±2,6	57,74±6,9	75,15±2,9
43	68,95±2,2	56,60±6,1	75,56±4,4

Ako môžeme vidieť, hodnoty ACC a SEN sa oproti individuálnej analýze pri metóde Random forest zvýšili. ACC sa zvýšila zhruba o 5 %, hodnota SEN zhruba o 10 %. Hodnota SPC sa pohybuje v rovnakých medziach ako pri individuálnej analýze. V tabuľke 6.7 je zobrazených prvých 5 najlepších skupín klasifikovaných metódou SVM a ich hodnoty ACC, SEN a SPC.

Tab. 6.7: Skupiny parametrov zoradené podľa ACC, v ktorých poradie parametrov je hodnotené podľa algoritmu mRMR. Skupiny parametrov sú klasifikované metódou SVM.

SVM			
Počet parametrov v skupine	ACC [%]	SEN [%]	SPC [%]
33	69,34±2,1	27,92±4,5	91,52±1,5
31	68,95±2,3	29,06±2,9	90,30±2,7
34	68,95±2,3	27,92±4,1	90,91±1,6
36	68,95±1,8	27,92±4,7	90,91±1,0
46	68,55±1,6	23,77±3,9	92,53±1,5

Pri metóde SVM sú hodnoty ACC a SEN nižšie oproti individuálnej analýze o cca 2 %. Hodnota SPC sa pohybuje v rovnakých medziach. Avšak hodnota SEN pri metóde SVM je v priemere približne až o 30 % nižšia, ako pri metóde Random forest. Vzhľadom na tento fakt je najlepšia skupina parametrov vybraná spomedzi skupín klasifikovaných metódou Random forest. Najvyššia hodnota ACC je pri skupine, v ktorej sa nachádzajú parametre hodnotené podľa mRMR od 1 po 36, čiže ide o skupinu s počtom parametrov 36. Tieto parametre sú vypísané v tabuľke 6.8, z toho je:

- 20 artikulačných parametrov (formanty),
- 13 prozodických parametrov (STE, TST, F_0 , TKEO),
- 2 fonačné parametre (jitter, PPE),
- 1 parameter hodnotiaci kvalitu signálu (SF).

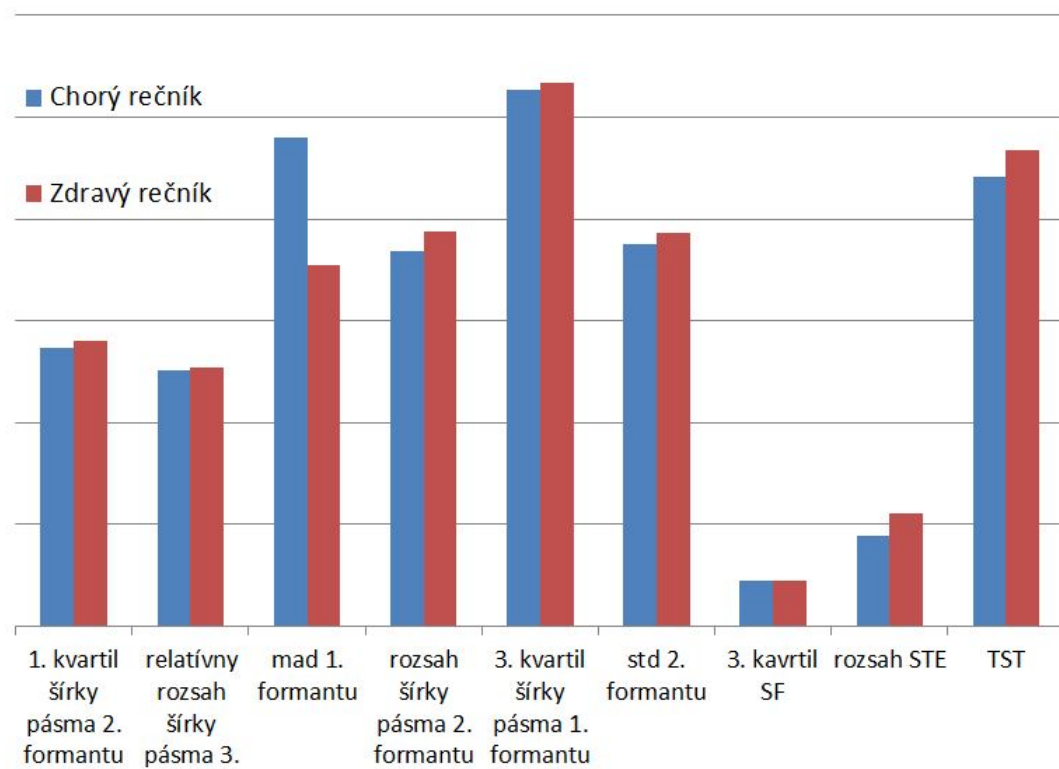
Artikulačné parametre obsadzujú približne 56 % všetkých parametrov zo skupiny, čo poukazuje o ich významnosti pri analýze reči. Taktiež sa potvrdila dôležitosť prozodických parametrov, ktorých je z celkového počtu parametrov v skupine cca 36 %. Avšak, je dobré si povšimnúť aj fakt, že fonačné parametre boli počítané iba 3 a z toho 2 sa taktiež nachádzajú vo vybranej skupine parametrov. Z výsledkov analýzy viacerých parametrov zároveň je možné konštatovať, že pri analýze reči je vhodnejšie klasifikovať skupinu parametrov, ako klasifikovať každý parameter individuálne. Taktiež je možné konštatovať, že metóda Random forest je vhodnejšia

na akustickú analýzu prednesu básne ako metóda SVM. Pomocou vybranej skupiny parametrov je dosiahnutý automatický systém detekcie Parkinsonovej choroby na základe akustickej analýzy prednesu básne s celkovou úspešnosťou $70,66 \pm 2,6$ %. Úspešnosť správneho označenia chorého pacienta za chorého je $59,25 \pm 6,3$ %. Oproti individuálnej analýze je zvýšenie úspešnosti minimálne, čo je pravdepodobne spôsobené výberom algoritmu mRMR, ktorý nepatrí medzi najrobustnejšie algoritmy pre výber parametrov. Na obrázku 6.1 je znázornené porovnanie strednej hodnoty 10 najlepších parametrov podľa mRMR pre zdravých a chorých rečníkov. Z tohto porovnania je možné sledovať nižšie hodnoty variability formantových frekvencií u chorých rečníkov. To len potvrdzuje, obmedzenú pohyblivosť artikulačných orgánov u pacientov s Parkinsonovou chorobou. Taktiež je možné pozorovať napríklad nižšiu hodnotu celkovej dĺžky rečového záznamu u chorých rečníkov, čo je spôsobené poruchou prozodických vlastností.

Tab. 6.8: Výpis parametrov z najlepšej skupiny podľa ACC. Parametre sú zoradené podľa hodnotenia mRMR.

1	1. kvartil šírky pásma 2. formantu	19	mad 3.formantu
2	relatívny rozsah šír. p. 3. formantu	20	PPE
3	mad 1. formantu	21	IQR šírky pásma 3. formantu
4	rozsah šírky pásma 2. formantu	22	std šírky pásma 2. formantu
5	3. kvartil šírky pásma 1. formantu	23	stredná hodnota TKEO
6	std 2. formantu	24	relatívna std TKEO
7	3. kvartil SF	25	1. kvartil STE
8	rozsah STE	26	stredná hodnota 1. formantu
9	TST	27	stredná hodnota F0
10	relatívna std šír. p. 2. formantu	28	1. kvartil šírky pásma 1. formantu
11	relatívny rozsah 2. formantu	29	medián STE
12	jitter	30	1. kvartil 1. formantu
13	medián F0	31	medián šírky pásma 2. formantu
14	rozsah F0	32	rozsah šírky pásma 3. formantu
15	relatívny rozsah 1. formantu	33	IQR F0
16	medián 1. formantu	34	stredná hodnota STE
17	std šírky pásma 2. formantu	35	IQR STE
18	3. kvartil 2. formantu	36	std TKEO

Zobrazenie rozdielov vybraných parametrov



Obr. 6.1: Zobrazenie odhadu strednej hodnoty prvých 10 najlepšie hodnotených parametrov podľa mRMR. Zobrazenie je po dvojiciach kde každá dvojica znázorňuje hodnotu pre chorých a zdravých rečníkov.

ZÁVER

Jedným z cieľov diplomovej práce bolo oboznámenie sa s Parkinsonovou chorobou a s ňou spojenou poruchou reči nazývanou hypokinetická dyzartria. Problematika je opísaná v prvej časti práce, kde je uvedený popis jednotlivých chorôb, ich priebeh, príznaky a momentálne liečebné postupy. Táto časť diplomovej práce bola časovo veľmi náročná, vzhľadom na početnosť odborných medicínskych výrazov a celkovú náročnosť danej problematiky. Ďalej sa práca venuje predspracovaniu rečového signálu. V tejto časti sú uvedené jednotlivé metódy používané pri predspracovaní a to ustredenie, preemfázová filtrácia a segmentácia. Z týchto metód bola použitá a nainplementovaná metóda segmentácie rečového signálu. Rečový signál je potrebné segmentovať na menšie úseky, vzhľadom na jeho nestacionárny charakter. Vďaka segmentácií sa rečový signál stáva kvázi-stacionárnym a je možné aplikovať výpočty požadovaných parametrov. Dĺžka segmentov je 20 ms s 50% prekrytím jednotlivých segmentov. Následne na to je popísaná parametrizácia rečového signálu. Parametre sú rozdelené podľa ich charakteru a to na artikulačné, prozodické, fonačné a parametre hodnotiace kvalitu signálu. V nasledujúcej časti sú uvedené klasifikačné metódy použité v diplomovej práci.

V štvrtej časti diplomovej práce je uvedený návrh systému automatickej diagnostiky Parkinsonovej choroby na základe akustickej analýzy prednesu básne a popis jednotlivých blokov. Návrh sa skladá zo šiestich blokov. V prvom bloku je načítanie databázy rečových záznamov do prostredia matlab. V druhom bloku je na jednotlivé rečové záznamy aplikovaná segmentácia signálu. Po rozsegmentovaní signálu je možné počítat parametre. Parametrizácia rečového signálu je v treťom bloku, ktorého súčasťou je rozhodovací blok. Ten v prípade, ak je výstupom parametrizácie vektor hodnôt, aplikuje na výstup štatistiku. Nasleduje blok s výberom parametrov. Parametre sú vyberané pomocou algoritmu mRMR. V poslednom bloku je aplikovaná klasifikácia a validácia. V diplomovej práci sú použité dve klasifikačné metódy, a to Random forest a SVM. Na validáciu je použitá metóda Cross-validácie.

V piatej časti diplomovej práce je uvedený postup implementácie navrhnutého systému. Celý systém je implementovaný v prostredí matlab. V diplomovej práci sú použité aj skripty vytvorené tretími stranami. Tvorcami týchto skriptov sú poprední odborníci v spracovaní rečových signálov, ako je napríklad Athanasios Tsanas. Všetky mnou vytvorené skripty a funkcie ako aj skripty tretích strán sa nachádzajú na priloženom DVD.

V poslednej časti, sú uvedené výsledky použitých klasifikačných metód. Ako prvá je uvedená individuálna analýza parametrov. Parametre sú vyhodnocované na základe výsledných hodnôt ACC, SEN a SPC pre jednotlivé parametre. Z výsledkov individuálnej analýzy parametrov je zrejmé, že pre zvýšenie úspešnosti analýzy Parkinsonovej choroby je potrebné kombinovať jednotlivé parametre, resp. nájsť najvhodnejšiu kombináciu týchto parametrov. Dominujú artikulačné parametre, čo mení počiatočný predpoklad dôležitosti prozodických parametrov pri akustickej analýze prednesu básne. Nevýhodou artikulačných parametrov je ich závislosť na pohlaví, čo značí o vhodnosti rozdelenia rečníkov na mužov a ženy. Ďalším výsledkom individuálnej analýzy je porovnanie klasifikačných metód Random forest a SVM. Metóda SVM sa javí ako vhodnejšia na základe vyšších hodnôt ACC(cca o 5 %) a SPC(cca o 15 %), avšak hodnota SEN je pri SVM nižšia(cca o 13 %). Vzhľadom na výpovedný charakter hodnoty SEN je táto hodnota pre praktické účely najsignifikantnejšia. z praktického hľadiska nie až taký veľký problém, ak je zdravý rečník označený za chorého, lebo následné medicínske testy ukážu jeho zdravotný stav. Avšak je veľmi nežiadúce, aby bol chorý rečník označený za zdravého. Z toho hľadiska je teda pri individuálnej analýze úspešnejšia klasifikačná metóda Random forest.

Po individuálnej analýze je vyhodnotená analýza viacerých parametrov zároveň. Výsledkom analýzy viacerých parametrov je skupina parametrov s najlepšou hodnotou ACC, kde parametre boli hodnotené pomocou algoritmu mRMR. Z výsledkov analýzy viacerých parametrov zároveň je možné konštatovať, že pri analýze reči je vhodnejšie klasifikovať skupinu parametrov, ako klasifikovať každý parameter individuálne. Taktiež je možné konštatovať, že metóda Random forest je vhodnejšia na akustickú analýzu prednesu básne ako metóda SVM. Pomocou vybranej skupiny parametrov je dosiahnutý hlavný cieľ diplomovej práce, a to automatický systém detekcie Parkinsonovej choroby na základe akustickej analýzy prednesu básne s celkovou úspešnosťou $70,66 \pm 2,6$ %. Úspešnosť správneho označenia chorého pacienta za chorého je $59,25 \pm 6,3$ %. Tento systém je v praxi nepoužiteľný, avšak poukazuje na obrovský potenciál výskumu v tejto oblasti.

V oblasti akustickej analýzy dyzartrickej reči sa poprední vedci prikláňajú k analýze reči pomocou nelineárnych dynamických parametrov, ktoré narozdiel od klasických konvenčných parametrov dokážu popísať vlastnosti reči, ktorá je značne poškodená. Parametre použité v diplomovej práci sú klinicky interpretovateľné aj napriek nižšej úspešnosti klasifikácie. V spolupráci s 1. neurologickou klinikou Fakultnej nemocnice u Sv. Anny v Brne prebieha na fakulte výskum Hypokinetickej dyzartrie, ktorý nie je zameraný na úspešnosť detekcie, ale skôr na klinický popis

patologickej reči a analýzu vzťahu medzi Parkinsonovou chorobou a Hypokinetickou dyzartriou. Pre vytvorenie použiteľného systému je potreba čo najlepšia voľba robustnejších rečových parametrov. K tomu je potrebné použiť komplexnejší algoritmus na výber parametrov, ktorý by mal uvažovať aj koreláciu medzi jednotlivými parametrami. Taktiež je potrebné hlbšie a pokročilejšie použitie klasifikačných metód, ktoré sú kľúčovým bodom pri automatickej diagnostike. Zvýšenie úspešnosti klasifikácie je navyše závislé na voľbe rečovej úlohy. Pre komplexné popísanie Hypokinetickej dyzartrie je potrebné použiť rečový protokol, ktorý by obsahoval viacero typov cvičení, ako je napríklad fonácia samohlások, čítanie slov, viet, diadochokinézia, monológ a podobne. V diplomovej práci sa jednalo čisto o akustickú analýzu prednesu básne, pričom je dokázané, že táto prozodicky a artikulačne náročná úloha dokáže aj pri použití jednoduchých lineárnych parametrov poskytnúť obstojné výsledky. Analýza rečových záznamov by mala byť vykonávaná pre každé pohlavie samostatne, vzhľadom na rozdielnu anatómiu hlasového traktu jednotlivých pohlaví. Rozdelenie rečníkov podľa pohlavia by s veľkou pravdepodobnosťou taktiež viedlo k zvýšeniu účinnosti systému. V neposlednom rade je dôležitá aj rozsiahla a kvalitná databáza rečových záznamov chorých a zdravých ľudí, pre natrénovanie čo najpresnejšieho modelu pre automatickú diagnostiku Parkinsonovej choroby.

LITERATÚRA

- [1] Parkinson.sk.[online]. 2015 [cit. 2015-10-17]. Dostupné z: <http://www.parkinson.sk/O-chorobe/Priciny-vzniku-choroby/b-f.article.aspx>
- [2] Parkinson J. An essay on the shaking palsy. *J Neuropsychiatry Clin Neurosci* 2002; 14(2): 223–236.
- [3] TSANAS, Athanasios. Accurate telemonitoring of Parkinson's disease symptom severity using nonlinear speech signal processing and statistical machine learning. University of Oxford, St. Cross College, 2012. University of Oxford, St. Cross College.
- [4] KOLLÁROVÁ, Katarína, Pavel RESSNER a Petr KANOVSKÝ. Genetika parkinsonovej choroby. *Neurologie pro praxi* [online]. 2007, (8(6): 357-359)[cit. 2015-10-17]. Dostupné z: <http://www.neurologiepropraxi.cz/pdfs/neu/2007/06/08.pdf>
- [5] Pharma news [online]. 2006. [cit. 2015-10-19]. Dostupné z: http://pharmanews.cz/2006_05/parkinson.html
- [6] Rektorová I. Současné možnosti diagnostiky a terapie Parkinsonovy nemoci. *Neurol. prax* 2009; 10 (Suppl 2): 5–36.
- [7] OIST: Okinawa institute of science and technology. 2013. Dopamine Pathways [online]. 1919-1 Tancha, Onna-son, Kunigami-gun Okinawa, Japan [cit. 2015-10-19]. Dostupné z: <https://www.oist.jp/news-center/photos/dopamine-pathways>
- [8] Rektorová I, Rektor I. Parkinsonský syndrom. In: Rektor I,Rektorová I et al. Centrální poruchy hybnosti v praxi. *Movement disorders*. Praha Triton 2003, p. 21–70.
- [9] ZAMIŠKOVÁ, G, P RESSNER, J DLOUHÁ a D ŠIGUTOVÁ. Poruchy řeči u Parkinsonovy nemoci. *Neurologie pro praxi*. 2010, 11(2): 112/116.
- [10] Neubauer K, et al. Neurogení poruchy komunikace u dospělých. Praha: Portál 2007: 227.
- [11] Hartelius, L.; Svensson, P.: Speech and Swallowing Symptoms Associated with Parkinson's Disease and Multiple Sclerosis: a Survey. *Folia Phoniatr Logop*, ročník 46, c. 1, 1994: s. 9–17.
- [12] Ramig, L. O.; Fox, C.; Sapir, S.: Speech treatment for Parkinson's disease. *Expert Rev Neurother*, ročník 8, c. 2, 2008: s. 297–309.

- [13] MEKYSKA, Jiří Analýza řečových promluv pro IT diagnostiku neurologických onemocnění: dizertační práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2014. 160 s. Vedoucí práce byl prof. Ing. Zdeněk Smékal, CSc.
- [14] MEKYSKA, J, Z SMÉKAL, M KOŠŤÁLOVÁ, M MRAČKOVÁ, S SKUTILOVÁ a I REKTOROVÁ. Motorické aspekty poruch řeči u Parkinsonovy nemoci a jejich hodnocení. *Cesk Slov Neurol.* 2011, (6): 74/107.
- [15] Metter, E. J.; Hanson, W. R.: Clinical and Acoustical Variability in Hypokinetic Dysarthria. *J Commun Disord*, ročník 19, c. 5, 1986: s. 347–366.
- [16] Logemann JA, Fisher HB. Vocal tract control in Parkinson's Disease. *Journal of Speech and Hearing Disorders* 1981;46:348–352.
- [17] Základný opis ľudskej reči [online]. [cit. 2015-12-09]. Dostupné z: ftp://kemt-old.fe.i.tuke.sk/SpracovanieAprenosRecovychAudiosignalov_261117_SPA/_web/voice/kurz/rr/rr2.htm
- [18] J. F. Caekebeke, A. Jennekens-Schinkel, M. E. van der Linden, O. J. Buruma, R. A. Roos, The Interpretation of Dysprosody in Patients with Parkinson's Disease, *J Neurol Neurosurg Psychiatry* 54 (1991) 145–148.
- [19] Ružicková H. Poruchy hlasu, reci a komunikace u Parkinsonovy nemoci. In Roth J, Sekyrová M, Ružicka E, et al. *Parkinsonova nemoc*. 4. vyd. Praha: Maxdorf 2009: 107–121.
- [20] Aarsland, D.; Bronnick, K.; Ehrt, U.; De Deyn, P. P.; Tekin, S.; aj.: Neuropsychiatric symptoms in patients with Parkinson's disease and dementia: frequency, profile and associated care giver stress. *J Neurol Neurosurg Psychiatry*, ročník 78, c. 1, 2007: s. 36–42.
- [21] Jankovic, J.: Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry*, ročník 79, c. 4, 2008: s. 368–376.
- [22] McConnell, G. C.; So, R. Q.; Hilliard, J. D.; Lopomo, P.; Grill, W. M.: Effective Deep Brain Stimulation Suppresses Low-Frequency Network Oscillations in the Basal Ganglia by Regularizing Neural Firing Patterns. *J Neurosci*, ročník 32, c. 45, 2012: s. 15657–15668.
- [23] Chenausky, K.; MacAuslan, J.; Goldhor, R.: Acoustic Analysis of PD Speech. *Parkinson's Disease*, ročník 2011, 2011: s. 1–13.

- [24] Bispebjerg Hospital. Duodopabehandling via sonde [online]. 2011 [cit.2015-11-02]. Dostupné z: <https://www.bispebjerghospital.dk/afdelinger-ogklinikker/neurologisk-afdeling/sygdom-og-behandling/sygdomme-og-behandlingsforloeb/Sider/duodopabehandlingvia-sonde.aspx>
- [25] Hartelius, L.; Svantesson, P.; Hedlund, A.; Holmberg, B.; Revesz, D.; aj.: Short-Term Effects of Repetitive Transcranial Magnetic Stimulation on Speech and Voice in Individuals with Parkinson's Disease. *Folia Phoniatr Logop*, ročník 62, c. 3, 2010: s. 104–109.
- [26] Dias, A. E.; Barbosa, E. R.; Coracini, K.; Maia, F.; Marcolin, M. A.; aj.: Effects of repetitive transcranial magnetic stimulation on voice and speech in Parkinson's disease. *Acta Neurol Scand*, ročník 113, c. 2, 2006: s. 92–99.
- [27] Murdoch, B. E.; Ng, M. L.; Barwood, C. H. S.: Retracted: Treatment of articulatory dysfunction in Parkinson's disease using repetitive transcranial magnetic stimulation. *Eur J Neurol*, ročník 19, c. 2, 2012: s. 340–347.
- [28] Sewall, G. K.; Jiang, J.; Ford, C. N.: Clinical Evaluation of Parkinson's-Related Dysphonia. *Laryngoscope*, ročník 116, c. 10, 2006: s. 1740–1744.
- [29] Sapir, S.; Ramig, L. O.; Fox, C. M.: Intensive voice treatment in Parkinson's disease: Lee Silverman Voice Treatment. *Expert Rev Neurother*, ročník 11, c. 6, 2011: s. 815–830.
- [30] CERNOCKÝ, J. Zpracování recových signálů. Ústav počítačové grafiky a multimédií, FIT, VUT Brno, 2006.
- [31] SMÉKAL, Z. Číslicové zpracování reci (MZPR). Elektronická skripta pro magisterská studia, Ústav telekomunikací, FEEC, VUT Brno, 2009.
- [32] Hillenbrand, J.; Getty, L. A.; Clark, M. J.; Wheeler, K.: Acoustic Characteristics of American English Vowels. *J Acoust Soc Am*, ročník 97, č. 5, 1995: s. 3099–111.
- [33] Sapir, S.; Ramig, L. O.; Spielman, J. L.; Fox, C.: Formant Centralization Ratio (FCR): A proposal for a new acoustic measure of dysarthric speech. *J Speech Lang Hear Res*, ročník 53, č. 1, 2010: str. 1–20.
- [34] Little, M.; McSharry, P.; Hunter, E.; Spielman, J.; Ramig, L.: Suitability of Dysphonia Measurements for Telemonitoring of Parkinson's Disease. *IEEE T Bio-Med Eng*, ročník 56, č. 4, 2009: s. 1015–1022.
- [35] PSUTKA, Josef, et al. Mluvíme s počítačem česky. Praha: ACADEMIA, 2006. 752 s. ISBN 80-200-1309-1.

- [36] GALÁŽ, Z. Analýza ručně psaného projevu u pacientů postižených neurologickými onemocněními. Brno: Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií, 2013. 98 s. Vedoucí diplomové práce Ing. Jiří Mekyska.
- [37] SMĚKAL, Z. Číslicové zpracování signálů (MCSI). Elektronická skripta pro magisterská studia, Ústav telekomunikací, FEEC, VUT Brno, 2009.
- [38] Medián. Matematika.cz [online]. Brno-město: Nová média, s. r. o., 2010 [cit. 2016-04-20]. Dostupné z: <http://www.matematika.cz/median>
- [39] Gavora, Peter a kol. 2010. Elektronická učebnica pedagogického výskumu. [online]. Bratislava : Univerzita Komenského, 2010. Dostupné na: <http://www.e-metodologia.fedu.uniba.sk/> ISBN 978-80-223-2951-4.
- [40] SKODDA, Sabine, Heiko RINSCHKE a Uwe SCHLEGEL. Progression of dysprosody in Parkinson's disease over time-A longitudinal study. Movement Disorders [online]. 2009, 24(5), 716-722 [cit. 2016-04-20]. DOI: 10.1002/mds.22430. ISSN 08853185. Dostupné z: <http://doi.wiley.com/10.1002/mds.22430>
- [41] Mean Absolute Deviation. Khan Academy [online]. 2016 [cit. 2016-04-23]. Dostupné z: <https://www.khanacademy.org/math/cc-sixth-grade-math/cc-6th-data-statistics/cc-6-mad/v/mean-absolute-deviation>
- [42] BASTL, P., J. KUČERA a D. LÉWOVÁ. Metody strojového učení [online]. 2003, 4 [cit. 2016-04-23]. Dostupné z: <http://tydenvedy.fjfi.cvut.cz/2003/cd/prispevky/sbpdf/strojuc.pdf>
- [43] COOPER, J. P. (ed.). The new Cambridge modern history: 1609-48/59. Cambridge: Cambridge University Press, c1970. ISBN 0-521-29713-3.
- [44] Random Forests Leo Breiman and Adele Cutler. Random Forests Leo Breiman and Adele Cutler [online]. [cit.2016-04-27]. Dostupné z: https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm
- [45] BURGET, R. Teoretická informatika. Brno: Vysoké učení technické v Brně, 2012. s. 1-198. ISBN: 978-80-214-4897- 1. (cs)
- [46] Cross Validation. Carnegie Mellon University [online]. Pittsburgh, 1997 [cit. 2016-04-28]. Dostupné z: <https://www.cs.cmu.edu/~schneide/tut5/node42.html>
- [47] Confusion Matrix. Confusion Matrix [online]. 1995 [cit. 2016-04-28]. Dostupné z: http://www2.cs.uregina.ca/dbd/cs831/notes/confusion_matrix/confusion_matrix.html

- [48] Alena Škaloudová, Ph.D. Alena Škaloudová, Ph.D. [online]. [cit. 2016-04-28]. Dostupné z: <http://kps.pedf.cuni.cz/skalouda/korelace.doc>
- [49] GIANNAKOPOULOS, Theodoros. A method for silence removal and segmentation of speech signals, implemented in Matlab [online]. 2010, , 3 [cit. 2016-05-04]. Dostupné z: www.di.uoa.gr/~tyiannak
- [50] A. Tsanas, M.A. Little, P.E. McSharry: "A methodology for the analysis of medical data", Chapter 7 in Handbook of Systems and Complexity in Health, pp. 113-125, Eds. J.P. Sturmburg, and C.M. Martin, Springer, 2013

ZOZNAM SYMBOLOV, VELIČÍN A SKRATIEK

ACC	Accuracy
dB	decibel
DDK	Diadochokinetic Tasks
DVD	Digital Versatile Disc
e	Eulerovo číslo(2,71828...)
F_0	Frekvencia základného hlasivkového tónu
FFT	Fast Furier Transformation
FIR	Finite impulse response
FN	false negative
FP	false positive
f_{vz}	vzorkovací kmitočet
HNR	Harmonic to Noise ratio
Hz	Hertz
IQR	Inter-quartil range
J_g	Jitter
kHz	kiloherzt
L-DOPA	dopamínerná liečba Hypokinetickej dyzartrie
log	logaritmus
LSVT	Lee Silverman Voice Treatment
mad	mean absolute deviation
MPSD	medián výkonovej spektrálnej hustoty
mRMR	minimum Redundancy Maximum Relevance
ms	milisekunda
NHR	Noise to Harmonic ratio

NST	Net Speech Time
π	Ludolfovo číslo (3.14159...)
PK	Pearmanov korelačný koeficient
PPE	Pitch Period Entropy
pPK	p hodnota Pearmanovho korelačného koeficientu
PSD	výkonová spektrálna hustota
pSK	p hodnota Spearmanovho korelačného koeficientu
$Q_{0,25}$	1. kvartil
$Q_{0,75}$	3. kvartil
rTMS	repetitive Transcranial Magnetic Stimulation
s	sekunda
SEN	Sensitivity
SF	Spectral flux
S_g	Shimmer
SK	Spearmanov korelačný koeficient
SPC	Specificity
std	smerodajná odchýlka
STE	Short Time Energy
STN-DBS	(Subthalamic Nucleus Deep Brain Stimulation)
Σ	suma
SVM	Support Vector Machine
T	Tesla
TKEO	Teager-Kaiser Energy operator
TN	true negative
TP	true positive

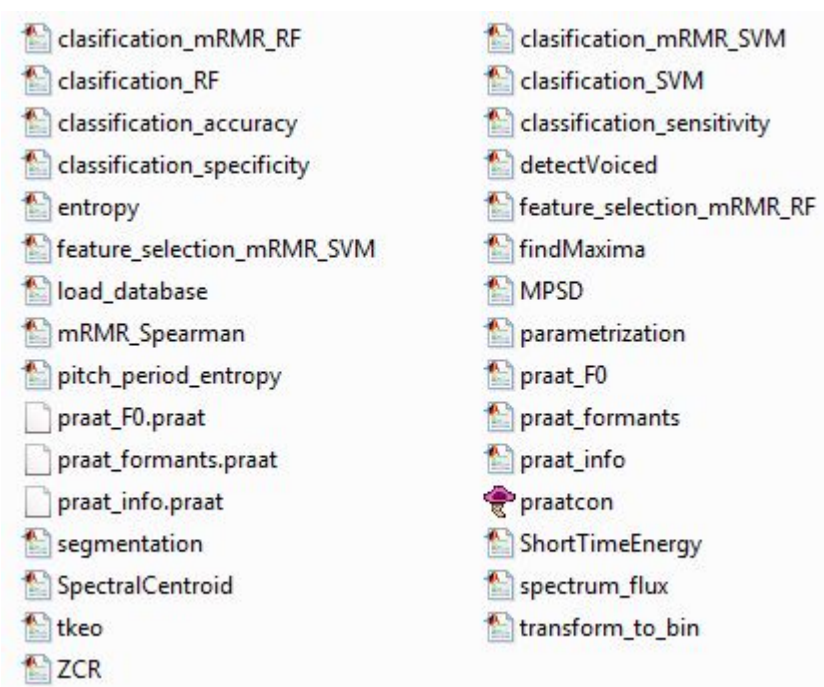
TPT	Total Pause Time
TST	Total Speech Time
\bar{x}	priemerná hodnota
ZCR	Zero-Crossing rate

ZOZNAM PRÍLOH

A OBSAH PRILOŽENÉHO DVD	69
-------------------------------	----

A OBSAH PRILOŽENÉHO DVD

Priložené DVD obsahuje súbory vytvorené a použité v rámci diplomovej práce. Jedná sa o súbory *.mat obsahujúce jednotlivé skripty a funkcie potrebné pre funkčnosť navrhnutého systému automatickej detekcie Parkinsonovej choroby. Tieto súbory sa nachádzajú v priečinku **zdrojové súbory**. Databáza rečových záznamov na priloženom DVD nie je, pretože sa jedná o dôverné dáta ku ktorým nie je povolený voľný prístup. Obsah priečinka **zdrojové súbory** je teda nasledovný:



Na DVD sa okrem zdrojových súborov nachádza aj elektronická verzia diplomovej práce pod názvom **diplomova_praca.pdf**.