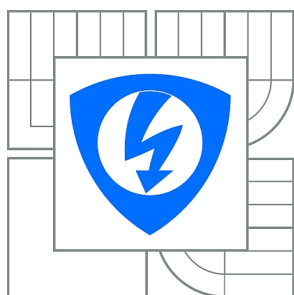


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ**
ÚSTAV TELEKOMUNIKACÍ

**FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS**

ZJIŠTĚNÍ PARKINSONOVY NEMOCI NA ZÁKLADĚ ANALÝZY ŘEČOVÉHO ZÁZNAMU

DIAGNOSING PARKINSON'S DISEASE FROM ANALYSIS OF SPEECH RECORDING

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

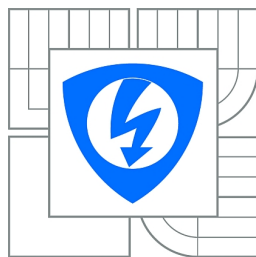
AUTOR PRÁCE
AUTHOR

PETR VYMLÁTIL

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. JAKUB LNĚNIČKA

BRNO 2013



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Bakalářská práce

bakalářský studijní obor
Teleinformatika

Student: Petr Vymlátil

ID: 125342

Ročník: 3

Akademický rok: 2012/2013

NÁZEV TÉMATU:

Zjištění Parkinsonovy nemoci na základě analýzy řečového záznamu

POKYNY PRO VYPRACOVÁNÍ:

Na základě vyhledání vhodných parametrů, které efektivně a komplexně popisují vlastnosti řečového signálu, je možné rozpoznávat Parkinsonovu nemoc, která se projeví ve změně základního tónu řeči, v prozodii, v mikrointonaci, monotónnosti hlasu apod. Cílem projektu je nalézt vhodnou metodu, která je schopna provést klasifikaci s dostatečnou přesností.

DOPORUČENÁ LITERATURA:

- [1] PSUTKA, J., MULLER, L., MATOUŠEK, J., RADOVÁ, V.: Mluvíme s počítačem česky. ACADEMIA, Praha 2006. ISBN 80-2100-1309-1
- [2] KRČMOVÁ, M.: Fonetika. Elektronické texty. MU Brno 2003.
<http://is.muni.cz/do/1499/el/estud/ff/js07/fonetika/materialy/index.html>
- [3] ZAPLATÍLEK, K., DOŇAR, B.: MATLAB – začínáme se signály. BEN, Praha 2010. ISBN 80-7300-200-0

Termín zadání: 11.2.2013

Termín odevzdání: 5.6.2013

Vedoucí práce: Ing. Jakub Lněnička

Konzultanti bakalářské práce:

prof. Ing. Kamil Vrba, CSc.
Předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské 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

Tato práce se zabývá zjištěním Parkinsonovi choroby na základě analýzy řečového záznamu. V úvodních kapitolách je popsán mechanismus vzniku lidského hlasu, jeho základní vlastnosti a vliv hypokinetické dysartrie na řeč. V další kapitole jsou popsány vlastnosti řečového signálu a některé metody jeho předzpracování. Následuje popis a způsob extrakce vybraných jednotlivých příznaků potřebných pro diagnózu nemoci a stručný popis metod redukce a klasifikátorů. Praktická část práce srovnává úspěšnost klasifikace naivního bayesovského klasifikátoru v závislosti na použité redukci.

KLÍČOVÁ SLOVA

Parkinsonova nemoc, hypokinetická dysartrie, řečový signál, klasifikátory, gaussovy smíšené modely, naivní bayesovský klasifikátor, analýza signálu, formant, předzpracování signálu, jitter, schimmer, mRMR, genetické algoritmy.

ABSTRACT

This thesis is focused on diagnosing Parkinson's disease from analysis of speech recording. Introduction of this work deals with description of voice production mechanism, it's basic qualities and influence of hypokinetic dysarthria on speech. In next chapter, there is described voice signal and some methods of it's preprocessing. Next part continues dealing with description of chosen individual symptoms, which are needed for PD diagnosing, followed by definition of chosen reduction methods and classifiers. There is a comparison of classify succes of naive bayes classifier, depending on chosen reduction method in last chapter of this work.

KEYWORDS

Parkinson's disease, hypokinetic dysarthria, voice signal, classifiers, gaussian mixture models, naive bayes classifier, signal analysis, formant, signal preprocessing, jitter, schimmer, mRMR, genetic algorithms.

VYMLÁTIL, Petr *Zjištění Parkinsonovy nemoci na základě analýzy řečového záznamu*: bakalářská práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2012. 53 s. Vedoucí práce byl Ing. Jakub Lněnička

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Zjištění Parkinsonovy nemoci na základě analýzy řečového záznamu“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

Brno

.....

(podpis autora)

PODĚKOVÁNÍ

Tímto bych rád poděkoval vedoucímu práce Ing. Jakubovi Lněničkovi za odborné vedení, konzultace, podnětné návrhy a podklady, díky kterým bylo možné tuto práci realizovat.

Brno

.....

(podpis autora)

OBSAH

Úvod	10
1 Vznik a vlastnosti lidské řeči	11
1.1 Prozodické vlastnosti řeči	11
1.1.1 Tempo	11
1.1.2 Intenzita řeči	12
1.1.3 Pauzy	13
1.1.4 Barva hlasu	13
2 Hypokinetická dysartrie:	14
2.1 Poruchy respirace	14
2.2 Poruchy fonace	14
2.3 Poruchy artikulace	14
2.4 Poruchy prozodie řeči	15
2.5 Plynulost řeči	15
3 Analýza řečového signálu	16
3.1 Znělost a neznělost řeči	16
3.2 Předzpracování signálu	17
3.2.1 Analogové předzpracování signálu	17
3.2.2 Digitální předzpracování signálu	17
3.2.3 Preemfáze	17
3.2.4 Ustředění	18
3.2.5 Segmentace pomocí oken	18
3.3 Určování hraničních bodů promluvy	19
3.4 Krátkodobá energie signálu	19
4 Extrakce vybraných příznaků	21
4.1 Měření základní frekvence hlasu	21
4.2 Detekce formantových frekvencí	22
4.3 Jitter	22
4.4 Schimmer	23
4.5 Statistické funkce použité pro výpočet některých příznaků	23
4.5.1 Medián	23
4.5.2 Průměr	24
4.5.3 Směrodatná odchylka	24

5	Diagnóza PN z hlasu	25
5.1	Selekce vhodných parametrů	25
5.2	Redukce parametrů	26
5.2.1	Míra geometrické oddělitelnosti	26
5.2.2	Metoda minimální redundance a maximální relevance (mRMR)	27
5.2.3	Genetické algoritmy	28
5.2.4	Křížová validace	30
5.3	Klasifikace	30
5.3.1	Gaussovy smíšené modely	30
5.3.2	Naivní bayesovský klasifikátor	31
6	Metoda řešení	32
6.1	Klasifikace na základě analýzy samohlásek	32
6.2	Zpracování nahrávek	33
6.3	Extrakce parametrů ze signálu a předzpracování signálu	33
6.4	Redukce příznaků	36
6.4.1	Metoda minimální redundance a maximální relevance (mRMR)	36
6.4.2	Míra geometrické oddělitelnosti	37
6.4.3	Genetické algoritmy	38
6.5	Klasifikace	40
7	Závěr	42
	Literatura	44
	Seznam symbolů, veličin a zkratk	47
	Seznam příloh	48

SEZNAM OBRÁZKŮ

1.1	Věta „Pes opravdu nechce být sám.“, se znázorněnou intenzitou řeči.	12
3.1	Slovo „sám“ a detail jeho znělého a neznělého úseku, ve kterém jsou vyznačeny jednotlivé základní hlasivkové periody.	16
3.2	Jednotlivé fáze předzpracování signálu	17
4.1	Reálné kepstrum stacionárního segmentu znělé hlásky „a“. Přejato z [24].	22
6.1	Blokové schéma navržené metody.	32
6.2	Extrakce parametrů pomocí programu Praat. Pod zobrazením zvukové nahrávky samohlásky „á“ je vykreslen spektrogram, jsou zde znázorněny jednotlivé formanty (červeně) a křivka ukazující základní hlasivkovou frekvenci (modře).	34

SEZNAM TABULEK

6.1	Příznaky vybrané pomocí metody mRMR - muži	37
6.2	Příznaky vybrané pomocí metody mRMR - ženy	37
6.3	Příznaky vybrané pomocí metody geometrické oddělitelnosti - muži .	38
6.4	Příznaky vybrané pomocí metody geometrické oddělitelnosti - ženy .	38
6.5	Příznaky vybrané metodou genetických algoritmů pro muže	39
6.6	Příznaky vybrané metodou genetických algoritmů pro ženy	39
6.7	Tabulka shrnující klasifikační úspěšnost u všech použitých metod. Q značí metodu geometrické oddělitelnosti, mRMR minimální redun- danci a maximální relevanci a GA genetické algoritmy	40
6.8	Příznaky vybrané metodou genetických algoritmů pro data bez spe- cifikace pohlaví	41

ÚVOD

Parkinsonova choroba je chronické, neurodegenerativní a pomalu se rozvíjející onemocnění, na které v současné době není známá žádná léčba a lze pouze zmírňovat jeho příznaky. Nejčastější výskyt počátku nemoci se pohybuje okolo 45.–60. rokem života. [26] Výskyt v populaci se odhaduje na 100-250 postižených na 100 000 obyvatel [6]. Postihuje jak muže, tak i ženy, roli nehraje společenská vrstva, způsob stravování, vzdělání, množství požitého alkoholu ani sociální stav. Problematika této nemoci spočívá především v její včasné diagnostice, která je velmi obtížná, neboť příznaky nemoci mohou vést k chybné diferenciální diagnostice. PN bývá často popisována jako komplex symptomů, mezi které patří akineze (zpomalenost pohybů), rigidita (ztuhlost) a tremor (třes). [17, 26] Další z projevů je tzv. hypokinetická dysartrie, která má zásadní vliv na řeč postiženého, proto jí bude v této práci věnována zvláštní pozornost.

Porucha řeči může postiženému značně znepříjemnit, mnohdy až znemožnit komunikaci se svým okolím, pro takovou osobu se stanou problematické i zcela běžné úkony každodenního života. To může vést ke špatnému psychickému stavu, objevují se deprese a celkové zhoršení kvality života.

Na základě analýzy řečového signálu, lze odhalit projev hypokinetické dysartrie na mluvené řeči osoby postižené PN, a to už v jejím poměrně ranném stádiu. Tato práce by proto měla vést k nalezení vhodných metod analýzy mluvené řeči, na základě kterých bude možné diagnostikovat tuto nemoc.

1 VZNIK A VLASTNOSTI LIDSKÉ ŘEČI

Mechanismus vytváření hlasu (fonace) spočívá v proudění vzduchu z plic přes překážku vytvořenou hmotou hlasivek, které mu uzavírají cestu. Vzhledem k pružnosti této hmoty proudící vzduch způsobí rozkmitání hlasivek a ty se přitom střídavě postupně otevírají a prudce uzavírají. Procházející vzduch se tímto mechanismem změnil ve střídající se kvanta hustšího a řidšího vzduchu, neboli ve zvukové vlny. Jako základní (hlasivkový) tón se označuje právě toto vlnění a představuje nosný zvuk řeči. Frekvence základního hlasivkového tónu, což je vlastně frekvence kmitání hlasivek, se potom označuje jako F_0 . Běžně se označuje také jako fundamentální frekvence. Tato frekvence představuje fyzikální charakteristiku řečového signálu a posluchač ji vnímá jako výšku hlasu. Její obrácená hodnota $T_0 = 1/F_0$ se nazývá perioda základního hlasivkového tónu. Rozsah základního hlasivkového tónu se udává asi 60-400 Hz. [20] Na frekvenci základního hlasivkového tónu se projevuje vliv intonace promluvy a to především v delším řečovém úseku, není proto konstantní.

1.1 Prozodické vlastnosti řeči

U řeči nás kromě obsahu promluvy mluvčího zajímá také tzv. prozodická informace, neboli to, jak to mluvčí říká. Tato informace se projevuje především u delších mluvených úseků a bývá přenášena několika parametry řeči, konkrétně to jsou již zmíněná 1 frekvence základního hlasivkového tónu (jeho změny tvoří melodii řeči, resp. intonaci), intenzita a časování řeči. Mezi parametry specifikující časování řeči se řadí rytmus, (který má vliv i na přízvuk) a rychlost řeči. Všechny tyto parametry jsou mezi sebou poměrně silně vázané, a je proto těžké popisovat jednu bez závislosti na druhé. Význam prozodické složky řeči je především proto, aby posluchač snáze pochopil obsah a význam sdělované informace, jde-li o tázací či oznamovací větu, jaká je její důležitost apod. Lze z ní také odvodit emocionální naladění mluvčího. [20] [11]

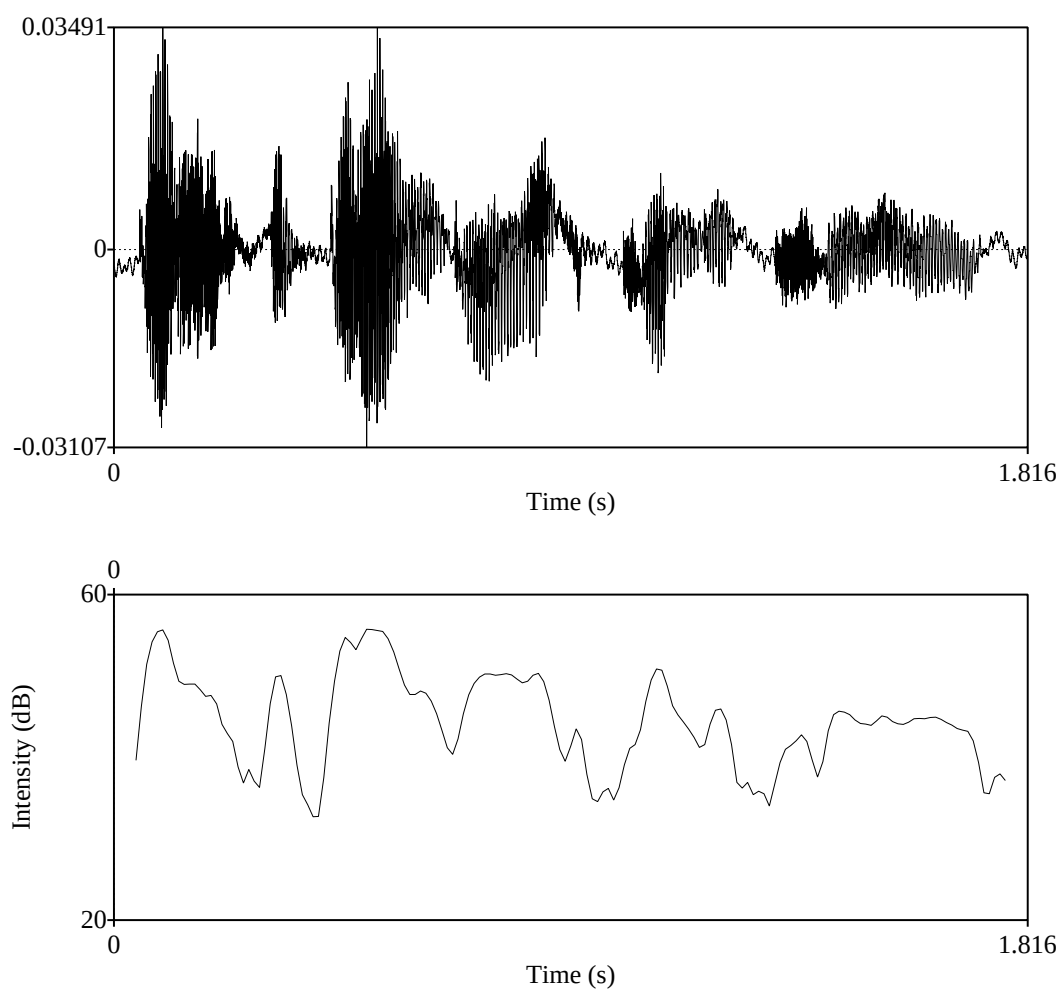
1.1.1 Tempo

Tempo, neboli rychlost řeči, popisuje jak rychle mluvčí hovoří a obvykle bývá popsána počtem vyslovených slabik za minutu, přičemž průměr se uvádí kolem 120 slov za minutu. [20] Jednotlivé jazyky se liší svým průměrným tempem řeči, stejné však lze říci i o jednotlivých osobách, na určité tempo řeči si lze navyknout. V mimořádně rychlém či pomalém tempu lze vysledovat některé emoce a postoj mluvčího k posluchačům. (Zpomalení tempa řeči z důvodu očekávání menší chápavosti posluchačů

pod.) Zvolená rychlost řeči určuje, do jaké míry se zkrátí či protáhnou jednotlivé hlásky.

1.1.2 Intenzita řeči

Intenzita řeči, neboli síla hlasu, je založena na amplitudě rozkmitu hlasivek. V závislosti na tom, jak bude mluvčí měnit sílu proudu vydechovaného vzduchu přes artikulační svalstvo, jehož napětí se bude taktéž měnit, bude ovlivněna i síla hlasu. Tyto vlastnosti jsou dány fyziologicky a jsou individuální u každé osoby. Regulace intenzity hlasu se taktéž projevují na vyjádření postoje k posluchači, např. zesílení intenzity při rozčilení mluvčího apod. [11]



Obr. 1.1: Věta „Pes opravdu nechce být sám.“, se znázorněnou intenzitou řeči.

1.1.3 Pauzy

Pauzy jsou další z parametrů, ovlivňující časování i srozumitelnost řeči. Jedná se o přerušení řeči, především z důvodu potřeby doplnit dech. Na vlastní srozumitelnost má především vliv umístování a délky pauz. Pokud bude pauza umístěna uvnitř větného úseku - tj. např. mezi dvě slova se silnou syntaktickou vazbou, tak velmi výrazně ovlivní celkovou srozumitelnost projevu. Správně by nádech neměl oddělovat výrazy, které k sobě významově patří. (např. lyžování v Alpách, oslava narozenin apod.) U spontánní konverzační řeči se umístování a délka pauz projeví více než při čtené, protože jejich zastoupení v prvním případě je až polovina celkového času, kdežto u čtené řeči je to pouze kolem 20 procent. Na délku pauz má přímý vliv i tempo řeči, které jejich délku prodlužuje při pomalejší řeči, a naopak zrychluje při rychlejší. [20]

1.1.4 Barva hlasu

Jedná o další složku řeči, která dokáže vlivem interference tónů a šumů vznikajících v artikulačním ústrojí mluvcího ovlivnit znění všech artikulačních souhlásek. Porovnáním s běžnou barvou hlasu na základě zkušeností dokážeme rozpoznávat emoční naladění mluvcího, barvu hlasu lze využít i při identifikaci mluvcího. [11]

2 HYPOKINETICKÁ DYSARTRIE:

Jedná se o poruchu motorické stránky tvorby řeči u PN a vyskytuje se až u 70-80 procent osob s touto nemocí. Z klinického hlediska ji lze rozdělit na HD lehkého, středního až těžkého stupně. Jak se HD projevuje lze sledovat na jednotlivých faktorech řeči, které budou postupně rozebrány. Jedná se především o poruchy respirace, fonace, a z hlediska fonetiky jsou to poruchy artikulace, prozodie a plynulosti řeči. Může se projevovat i jako chybné nasazování hlasové polohy z hlediska zvýšení či snížení základního posazení hlasu.[26, 17, 25]

2.1 Poruchy respirace

Poruchy respirace (dýchání) jsou způsobeny zkrácením fází nádechu a výdechu, což vede k ovlivňování tvorby hlasu, má vliv na jeho kvalitu, sílu a délku promluvy. U osob s PN se potom dá pozorovat zrychlování tempa řeči, mluvení v krátkých větách apod.

2.2 Poruchy fonace

Další z příznaků dysartrie je porucha fonace (tvorby hlasu), která se projevuje jako hlasový tremor (nedokonalé sevření hlasivek a redukované, asymetrické nebo pomalé chvění hlasivek) . U osob s PN klesá síla hlasu při delším projevu daleko rychleji, než u osob zdravých, nemoc také snižuje rozsah výšky hlasu. Hlas se potom jeví jako monotónní. Objevuje se také snížení celkové úrovně síly hlasu (hypofonie) a narušení kvality hlasu (dysfonie), hlas bývá chraptivý, zastřený nebo drsný, pacient s PN má problém správně hlas zesilovat či zeslabovat. [26]

2.3 Poruchy artikulace

Porucha artikulace je způsobena zejména vlivem PN na artikulační svalstvo, u něhož bývá snížený jeho rozsah a rychlost pohybu. To se projevuje zejména na souhláskách, správné vyslovování samohlásek bývá nedotčené. Z logopedického hlediska lze rozdělit souhlásky do několika skupin, rozlišených podle způsobu jejich utváření logoped potom vyšetřuje postižení výslovnosti jednotlivých skupin. V případě lehké až těžké poruchy artikulace lze potom na úrovni slov, vět a delších promluv sledovat nepřesnou a nedbalou výslovnost s vlivem na srozumitelnost řeči. [17]

2.4 Poruchy prozodie řeči

Jak již bylo zmíněno výše 1.1, prozodické faktory řeči mají vliv zejména na rozpoznání emočního naladění u mluvčího a zdůrazňují obsah sdělení či jeho význam. PN má vliv na všechny tyto faktory. Vlivem absence melodie a přízvuku se projev jeví jako monotónní, narušením tempa řeči vzniká zpomalení a zrychlování projevu a chybné umístování pauz v řeči ovlivňuje celkovou srozumitelnost.

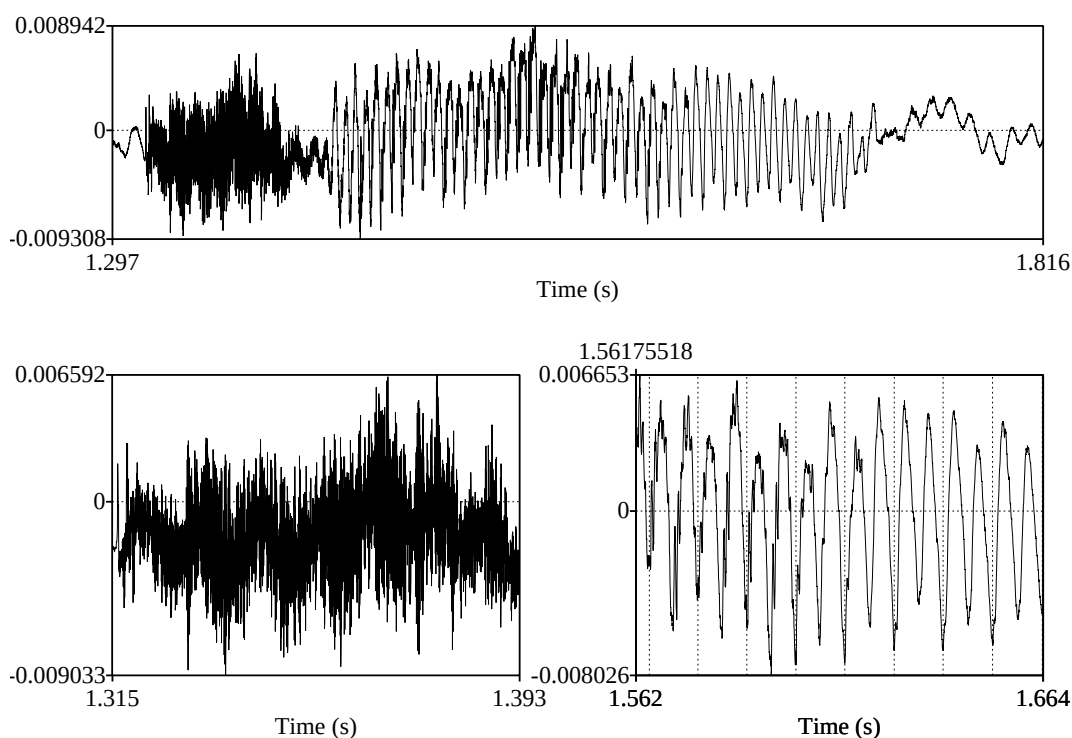
2.5 Plynulost řeči

Plynulost řeči bývá narušena dvěma formami poruch - palilalie a hezitace. [25] V případě palilalie se jedná o zrychlené opakování slabik na začátku, uprostřed či na konci slov, či opakování celých slov. Hezitace jsou pauzy, které se v řeči vyskytují neplánovaně a to kdekoliv v průběhu projevu. Jedná se například o ulpívání na hláskách či slabikách.

3 ANALÝZA ŘEČOVÉHO SIGNÁLU

3.1 Znělost a neznělost řeči

Z akustického hlediska lze vlastnosti řeči sledovat především na výše 1 zmíněných parametrech, konkrétně na základní hlasivkové frekvenci, intenzitě a rozdělení spektrální energie. V případě, že budeme pozorovat vlastnosti časového průběhu vlny u lidské řeči, lze zjistit, že bývá tvořena oblastmi, z nichž jedna má periodický charakter a druhá má v podstatě stochastický průběh. V prvním případě se jedná o znělou část řeči, a periodicitu této oblasti je daná základní frekvencí hlasu. Ve druhém případě se jedná o neznělou řeč, a její charakter se dá nejlépe přirovnat k šumu.[20, 12]

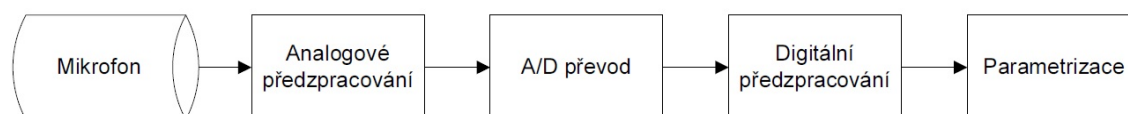


Obr. 3.1: Slovo „sám“ a detail jeho znělého a neznělého úseku, ve kterém jsou vyznačeny jednotlivé základní hlasivkové periody.

Na charakteru samohlásek a znělých souhlásek se také podílejí tzv. formanty, což jsou oblasti koncentrace akustické energie, soustředěné kolem určitého rezonančního kmitočtu. Formanty vznikají vlivem rezonance hlasivkového tónu uvnitř hrdelní, ústní a případně i nosní dutiny. Jejich přítomnost může značně zkomplikovat detekci základní hlasivkové frekvence. Formanty se označují jako $f_0, f_1, \dots, f_n$ a z hlediska obsahu informace o charakteru řeči jsou nejvíce podstatné první tři z nich.

3.2 Předzpracování signálu

Vzhledem k poměrně vysoké variabilitě lidské řeči a přítomnosti mnoha rušivých vlivů, které vznikají již při snímání zvuku (zkreslení při přenosu signálu vlivem vlastností jednotlivých prvků akustického řetězce, hluky na pozadí, rušení) vzniká potřeba signál nejprve vhodně předzpracovat, a tím minimalizovat tyto rušivé elementy, které by jinak mohly mít vliv na kvalitu extrakce jednotlivých parametrů signálu. Předzpracování lze rozdělit podle fáze ve které probíhá na analogové a digitální.



Obr. 3.2: Jednotlivé fáze předzpracování signálu

3.2.1 Analogové předzpracování signálu

Analogovým předzpracováním lze nazývat sled operací, které upravují vlastnosti signálu ještě před jeho digitalizací. Tyto operace se signálem zahrnují například zařazení antialiasingového filtru pro potlačení vysokých frekvencí signálu, jejichž přítomnost nemá pro extrakci parametrů potřebných pro tuto práci význam, mohou však vést k aliasingu a tím podstatnému zkreslení signálu. Podle potřeby může také zahrnovat zařazení dalších nástrojů, které mohou pomoci zlepšit celkovou kvalitu signálu, v tomto případě to však nebývá příliš běžné. Jako příklad lze uvést zařazení vhodného mikrofonního předzesilovače či použití kvalitnější kabeláže pro vedení zvukového signálu, jež by minimalizovala vliv okolního rušení.

3.2.2 Digitální předzpracování signálu

Zpracování v číslicové oblasti nastává po převodu signálu do diskrétní podoby pomocí A/D převodníku s dostatečnou vzorkovací frekvencí. Nejvýznamnější operace lze přehledně shrnout v níže popsanych bodech.

3.2.3 Preemfáze

Pro zpracování řeči je vhodné, aby spektrum signálu bylo rovnoměrné. Při dlouhodobé analýze spektra řeči je však patrné, že jeho střední část klesá se sklonem 6dB/oktávu. Vzhledem k vysokému množství informací o řeči, obsažených v pásmu

nad 300 Hz a rovnoměrnému spektru kvantizačního šumu, který se proto projevuje o to více u energeticky slabších složek, je vhodné tento efekt potlačit. To lze provést pomocí číslicového filtru s touto charakteristikou:

$$H[z] = 1 - \lambda z^{-1}. \quad (3.1)$$

Preemfáze se potom nazývá úprava signálu takovým filtrem, který zvýrazňuje jeho vyšší kmitočty a lze jej popsat vztahem:

$$s[n] = s[n] - \lambda s[n-1]. \quad (3.2)$$

Koeficient λ se volí obvykle v intervalu 0.9–1. Lze také použít tzv. adaptivní preemfázi:

$$\lambda = \frac{R[1]}{R[0]}, \quad (3.3)$$

kde se koeficient λ mění v závislosti na podílu prvních dvou autokorelačních koeficientů.[12]

3.2.4 Ustředění

V případě, že by při snímání signálu došlo k jeho nechtěnému stejnosměrnému posunutí v kladném či záporném směru, což by mělo negativní dopad např. na výpočet energie, je vhodné signál ustředít a tím tuto stejnosměrnou složku potlačit. Pro ustředění jednoho rámce lze použít vztah:

$$s'[n] = s[n] - \frac{1}{N} \sum_{i=1}^N s[i], \quad (3.4)$$

platící pro $1 \leq n \leq N$, kde $s'[n]$ je ustředěná hodnota, $s[n]$ je n -tá hodnota rámce a N je počet vzorků obsažených v jednom rámci. [16]

3.2.5 Segmentace pomocí oken

Vzhledem k podstatě vzniku lidského hlasu, je možné přijmout předpoklad, že řečový signál má kvazistacionární vlastnosti, neboli že změny v hlase jsou v čase dostatečně pomalé na to, aby se v řečovém signálu daly vytvořit ekvidistantní úseky, neboli segmenty, o délce N vzorků, a každý takový segment mohl být zvlášť popsán vektorem příznaků.

Zvolením délky segmentu 10–25 ms zajistí dostatečně malou velikost k tomu, aby bylo možné popsat jeho vlastnosti konstantními hodnotami, a zároveň byl tento segment dostatečně dlouhý pro změření těchto parametrů. Tento interval úzce souvisí

s fyziologickými vlastnostmi hlasového ústrojí. V případě, že se budou segmenty částečně překrývat, lze docílit částečného vyhlazení časových průběhů parametrů signálu, což však sebou nese nevýhody, jako zvýšení výpočetních nároků a zpomalení časového posunu.

Přidělování určité váhy jednotlivým segmentům signálu se provádí pomocí váhovací funkce, jejíž typ určuje typ použitého okna. Pro analýzu řečového signálu se nejčastěji používá pravoúhlé nebo Hammingovo okno, z nichž vhodnější je právě Hammingovo. To je například z důvodu jeho potlačování vzorků na okrajích segmentů, což zajišťuje lepší stabilitu výpočtů. Funkce tohoto okna je

$$w[n] = \begin{cases} 0,54 - 0,46\cos\left(\frac{2\pi n}{N}\right) & \text{pro } n = 1, 2, \dots, N \\ 1 & \text{pro ostatní } n \end{cases}. \quad (3.5)$$

N zde představuje délku okna. Hammingovo okno je zde výhodné také proto, že při použití pravoúhlého okna vzniká rozmazání a rozptyl spektra, obojí související s charakterem hlavního a vedlejších laloků spektra tohoto okna. [20, 12]

3.3 Určování hraničních bodů promluvy

Aby bylo možné zjistit, kde vlastně v analyzovaném signálu začíná vlastní řeč, je potřeba nalézt počáteční a koncové body promluvy. Toho lze dosáhnout pomocí krátkodobé funkce středního počtu průchodu signálu nulou, která je definovaná jako

$$Z_c = \frac{1}{2M} \sum_{i=0}^{M-1} |\text{sgn}[x(i)] - \text{sgn}[x(i-1)]| \quad (3.6)$$

kde M je počet analyzovaných vzorků a $x(i)$ je předzpracovaný vstupní signál. [20]

3.4 Krátkodobá energie signálu

K detekci a oddělení segmentů ticha od segmentů řeči, nebo též k oddělení znělé a neznělé řeči, lze využít detekci krátkodobé energie signálu, jejíž funkce je

$$E_n = \sum_{k=-\infty}^{\infty} [s(k)w(n-k)]^2, \quad (3.7)$$

a délku mikrosegmentu lze doporučit shodnou s délkou okna, a to 10–20 ms při frekvenci vzorkování odpovídající kvalitě zpracovávaného signálu [20]. Tato charakteristika je však velmi citlivá na velké změny úrovně signálu, proto se spíše využívá krátkodobá intenzita:

$$E_n = \sum_{k=-\infty}^{\infty} |s(k)|w(n-k), \quad (3.8)$$

která již není na tyto změny natolik citlivá. Takto vypočtený segment energie však jako samotný neudává informaci o tom, zda je analyzovaná řeč nebo pauza, musí se proto srovnat s odpovídající prahovou hodnotou, definovanou jako

$$E_p = 1,5 \cdot E_d, \quad (3.9)$$

kde E_d je úroveň energie hluku pozadí[1]aktualizována podle vztahu

$$E_d = (1 - p)E_d^o + pE, \quad (3.10)$$

kde E_d^o je poslední známá hodnota energie. Tato aktualizace se provádí v řečových pauzách, E je energie signálu v segmentu a p se volí kolem 0,5. [1]. Pokud je E_n větší, než E_d , lze analyzovaný signál ohodnotit jako řeč, v opačném případě je energie šumu aktualizována. [20]

4 EXTRAKCE VYBRANÝCH PŘÍZNAKŮ

V této kapitole bude objasněn způsob výpočtu některých příznaků, které byly vybrány pro realizování této práce. Ty jsou uvedené v A.1. Větší část těchto příznaků vychází ze základních parametrů analyzované části signálu, mezi které patří především základní hlasivková frekvence, frekvence jednotlivých formantů a příznaky odvozené z parametrů jitter a schimmer.

4.1 Měření základní frekvence hlasu

Měření základního hlasivkového tónu bývá problematické, zejména vlivem jednotlivých formantů, velkým rozdílem ve frekvenci základního tónu mezi jednotlivými řečníky a variabilitou u jedné osoby, kdy se tato frekvence může měnit při vyslovování různých hlásek až v poměru 2:1. Problém s formantami se dá řešit vhodným předzpracováním signálu. Nejvíce výhodné pro zjištění základní frekvence hlasu je aplikování autokorelační funkce, která je definována pro rámec $x[n]$ jako

$$R[m] = \sum_{k=0}^{L-1-m} x[k]x[k+m], \quad (4.1)$$

kde L je velikost segmentu, která se musí pohybovat vzhledem k běžnému rozsahu frekvencí v základním tónu, viz 1, mezi 20–40 ms. Hodnotu maxima označíme jako R_{max} . Základní hlasivkový tón potom nalezneme pomocí vztahu

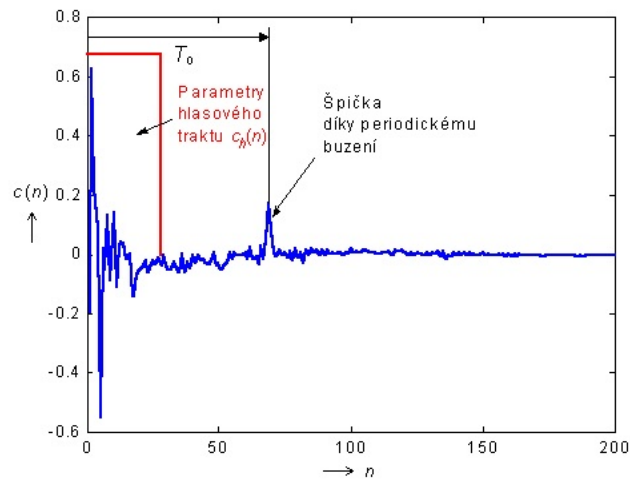
$$\frac{R_{max}}{R[0]} > \alpha, \quad (4.2)$$

kde se α volí přibližně 0,3. [12] V případě vyhovění této rovnici se dá analyzovaná část signálu označit jako znělá část promluvy a výsledná hodnota určuje její základní periodu, v opačném případě byla analyzována neznělá část řeči.

Jako další metodu pro určení fundamentální frekvence lze uvést keprální analýzu signálu, [15] kde po výpočtu reálného kepru signálu pomocí vztahu

$$C_r[n] = IDFT\{\ln|DFT\{x[n]\}|\}, \quad (4.3)$$

lze získat charakteristiku, ve které jsou parametry hlasivkového traktu promítnuty blízko počátku, kdežto rychlé periodické změny průběhu spektra vlivem hlasivkového buzení lze pozorovat jako špičku v čase odpovídající základní hlasivkové periodě. Rozsah kvefrencce, ve kterém lze tutu špičku hledat lze nastavit na 32–320 [-], což odpovídá intervalu 50–500 Hz ve vzorkovací frekvenci 16 kHz.[15] V případě, že bude analyzována neznělá řeč, nebudou rychlé změny ve spektru již periodické a tudíž se v kepru nepromítnou ve formě špičky v odpovídajícím čase. Při této analýze je doporučeno [15] používat Hammingovo okénko, které má lepší spektrální vlastnosti než okénko obdélníkové. [15, 12]



Obr. 4.1: Reálné kepstrum stacionárního segmentu znělé hlásky „a“. Přejato z [24].

4.2 Detekce formantových frekvencí

Při odhadu formantových frekvencí dochází k několika problémům. Mezi hlavní patří především výskyt nepravých vrcholů ve spektrální obálce a splývání formantů. V prvním případě se jedná o problém, kdy namísto toho, aby se ve spektrální obálce vyskytovaly maxima způsobená pouze formanty, tak zde vznikají vrcholy, které mohou být vyvolány různými poruchami. Druhý případ nastává tehdy, když dojde ke splývání špiček dvou formantů do takové míry, že je už nelze od sebe rozlišit. Existují algoritmy, založené například na aplikaci skrytých Markovových modelů při modelování hlásek, nebo na sinusovém rozkladu řečových kmitů, které dokáží efektivně odhadnout formantové frekvence i přes různé nastávající komplikace. Avšak vzhledem k tomu, že ve spektrální obálce bývají nalezeny přesně tři vrcholy prvních formantů ve většině případů (85–90 %), [20] nebude jejich implementace pro potřeby této práce důležitá.

Pro odhad těchto frekvencí bude tedy nejvhodnější použití metody vyhledávání vrcholů spektrální obálky. Po výpočtu obálky podle vzorce lze tedy jednotlivé formanty nalézt jako hodnoty frekvencí v místech lokálních maxim. Je také vhodné zohlednit šířku pásma jednotlivých vrcholů, a to z toho důvodu, že vrcholy formantů nepřesahují šířku 500 Hz. [20]

4.3 Jitter

Jitter se často využívá jako parametr pro měření kvality hlasu. Lze jej stručně definovat jako krátkodobou nepravidelnost délek jednotlivých period řečového signálu,

existuje však více variant a modifikací tohoto parametru. Pro výpočet lokálního kolísání periody lze postupovat tak, že se nejprve definuje jitter absolutní jako průměrná absolutní hodnota rozdílu mezi sousedními periodami, a to pomocí vztahu

$$jitter(absolute) = \frac{1}{n-1} \sum_{i=1}^{n-1} |T_i - T_{i+1}|, \quad (4.4)$$

kde T_i jsou vypočítané délky period F_0 a n je celkový počet těchto period. Takto získaná hodnota se vydělí průměrnou délkou periody. Výsledný vztah lze potom zapsat jako

$$jitter = \frac{\frac{1}{n-1} \sum_{i=1}^{n-1} |T_i - T_{i+1}|}{\frac{1}{n} \sum_{i=1}^n T_i}. \quad (4.5)$$

4.4 Schimmer

Schimmer popisuje amplitudovou nestabilitu analyzovaného signálu. Stejně jako jitter, i schimmer má několik různých variant, lišící se jak v jednotkách, kterými tento jev popisujeme, tak počtem period zahrnutých do analýzy. Pro relativní hodnotu schimmeru vyjádřenou v procentech platí vztah

$$schimmer = \frac{\frac{1}{n-1} \sum_{i=1}^{n-1} |A_i - A_{i+1}|}{\frac{1}{n} \sum_{i=1}^n A_i}, \quad (4.6)$$

kde A_i je špičková hodnota signálu a n je počet extrahovaných základních hlasivkových period.

4.5 Statistické funkce použité pro výpočet některých příznaků

Pro výpočet některých příznaků, viz A.1, je použito několik statistických funkcí. Jsou to především průměr, medián a směrodatná odchylka. Ty budou v této kapitole stručně popsány.

4.5.1 Medián

Pomocí této statistické funkce bude vypočten především příznak F_0 median. Způsob určení F_0 je popsán v 4.1. Jako medián se označuje hodnota, která dělí podle velikosti

seřazený statistický soubor na dvě poloviny o stejném počtu prvků. V případě, že by soubor obsahoval sudý počet těchto prvků, a nebylo zde tedy možné určit jednoznačně hodnotu uprostřed, využívá se průměru prostředních dvou hodnot.

4.5.2 Průměr

Mezi nejtypičtější míru polohy patří aritmetický průměr, který je definován jako součet všech hodnot dělený jejich počtem: [18]

$$x = \frac{1}{n} \sum_{i=1}^n x_i. \quad (4.7)$$

4.5.3 Směrodatná odchylka

Míru proměnlivosti rozdělení statistického souboru kolem jeho střední hodnoty se nejčastěji určuje pomocí směrodatné odchylky. Tu lze definovat jako druhou odmocninu rozptylu: [18]

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (4.8)$$

5 DIAGNÓZA PN Z HLASU

Postup diagnózy PN z řečového signálu se dá rozdělit do několika kroků. Nejprve je třeba vybrat správné parametry z hlasového signálu, u kterých se očekává projev nemoci. Poté je potřeba tyto příznaky vhodnou metodou redukovat a tím zajistit, že se nadále bude pracovat jen s takovými parametry, které zajistí úspěšné rozpoznání mezi osobou zdravou a osobou s PN. Nakonec se provede klasifikace, která zajistí zařazení testovaného signálu do správné množiny, tedy do skupiny označené jako řečový signál zdravých či nemocných lidí. Jednotlivými kroky se bude zabývat tato kapitola.

5.1 Selekce vhodných parametrů

V současnosti není přesně určeno, kterým parametrům a metodám u akustické analýzy řečového signálu pro diagnostikování PN vlastně dát přednost a které sebou přinášejí nejvyšší úspěšnost rozpoznání nemoci. V těchto názorech se literatura často rozchází. Například [21] se věnuje analýze založené na intenzitě a rychlosti řeči, kde se jako podstatná výhoda této metody uvádí její univerzálnost vzhledem k pohlaví pacientů, které se převážně projevuje v metodách založených na základní hlasivkové frekvenci. Ve studii [23] bylo z tohoto důvodu testování založené na parametru F_0 provedeno na pacientech mužského i ženského pohlaví zvláště a u parkinsoniků zde byl zjištěn velký pokles variability F_0 vzhledem k první a poslední čtené větě, bez ohledu na pohlaví. Ohledně změn v parametrech řeči prezentuje podobné výsledky studie [8], která se mimo jiné zabývá rozdílným projevem mezi konverzační řečí a řečí, kdy je mluvčí požádán, aby mluvil srozumitelně. Jako zajímavý fakt je zde uvedeno, že skupina s PN používá jiný způsob z hlediska vkládání pauz a regulování artikulační rychlosti pro produkci srozumitelné/konverzační řeči než kontrolní skupina. Uvádí také, že osoby s PN mají o něco vyšší střední hodnotu F_0 než kontrolní skupina a také pomalejší artikulační rychlost.

Další zajímavou metodu zkoumá [2]. V prvním experimentu zkoumá vliv hlasitosti konverzace a čtení textu testované skupiny v závislosti na hlasitosti jejich vlastního hlasu přivedeného do sluchátek konverzujícího. Zatímco kontrolní skupina snižovala hlasitost konverzace v závislosti na zvyšující se hlasitosti jejich vlastního hlasu ve sluchátkách, osoby s PN na něj reagovaly jen minimálně. Ve druhém experimentu byla použita podobná metoda, kde do sluchátek byl přiveden růžový šum. Ukázalo, že se zvyšující se hlasitostí rušení ve sluchátkách kontrolní skupina zvyšuje hlasitost vlastního hlasu, kdežto osoby s PN projeví jen minimální reakci na tento rušivý vliv. V posledním experimentu je pomocí vnějších instrukcí testován rozsah

hlasitostí řeči u kontrolní skupiny a u osob s PN. V tomto experimentu se ukázalo, že osoby s PN mají tento rozsah menší než kontrolní skupina. Ve všech experimentech této studie se také ukázalo, že osoby s PN mluví celkově více potichu.

Další metodou se zabývá studie [14], ve které se autoři rozhodli pro široký výběr parametrů, popisujících řečové vlastnosti signálu a z nahrávek pořízených ve Fakultní nemocnici u sv. Anny v Brně testovali, kterými parametry lze nejlépe rozlišit kontrolní skupinu od skupiny s PN. Jako nejvíce podstatné jsou zde (avšak pouze pro mužské pohlaví) uvedeny koeficienty odvozené z prvního formantu, koeficienty odvozené z harmonicity (odstup harmonických složek od šumu) a parametr jitter, popisující frekvenční nestabilitu v rámci periody. V této práci měření založená na energetických vlastnostech signálu jsou vynechána z důvodu nekalibrovaného mikrofónu a v době psaní této studie bylo ještě pořízeno příliš malé množství nahrávek.

5.2 Redukce parametrů

Pro úspěšnou klasifikaci je potřeba nejprve vybrat takové parametry, které jsou pro tento účel nejvhodnější. Existuje více metod, z nichž každá vykazuje rozdílnou efektivitu v závislosti na účelu jejího použití.

5.2.1 Míra geometrické oddělitelnosti

Jedna z možností je využití výpočtu míry geometrické oddělitelnosti pro každý příznak. Tato metoda spočívá v určení rozptylu a vzdálenosti mezi třídami a na základě těchto parametrů ohodnotí příznak jako kvalitní či nikoliv. Tedy v případě, že se prvky jedné třídy vyskytují v okolí střední hodnoty a zároveň se jednotlivé střední hodnoty tříd dostatečně liší, lze příznak ohodnotit pozitivně. Postup celého výpočtu podle [9] je následující:

Nejprve je potřeba určit kvadrát vzdálenosti mezi středními hodnotami jednotlivých tříd:

$$D_{v,w}^2 = (\mu_v - \mu_w)^2. \quad (5.1)$$

Zde v, w jsou jednotlivé třídy a μ označuje střední hodnotu. Dále je potřeba určit aritmetickou střední hodnotu vzdáleností mezi třídami:

$$D^2 = \frac{1}{V(V-1)} \sum_{v=1}^V \sum_{w=1}^V D_{v,w}^2. \quad (5.2)$$

V zde označuje celkový počet tříd. Dále se určí kvadrát rozptylu třídy v okolo střední hodnoty:

$$S_v^2 = (x - \mu_v)^2, \quad (5.3)$$

kde x označuje vektor příznaků. Aritmetický průměr hodnot S^2 jednotlivých tříd se poté spočítá podle vzorce 4.7 a nakonec se určí celková geometrická oddělitelnost tříd:

$$Q = \frac{S^2}{S^2 - D^2}, \quad (5.4)$$

kde se kvalita příznaku Q pohybuje mezi hodnotou 1 a 0, přičemž platí, že čím je hodnota bližší 0, tím je příznak kvalitnější, protože v rámci své třídy vykazuje malé rozdíly, kdežto mezi třídami velké.

5.2.2 Metoda minimální redundance a maximální relevance (mRMR)

Další z možností jak určit vhodnost příznaku pro klasifikaci, je metoda mRMR. Algoritmus mRMR vychází z myšlenky, že obecně je minimální klasifikační chyba definovatelná jako maximální statistická závislost třídy c na rozložení dat v prostoru příznaků. [7] Využívá přitom maximální relevanci a minimální redundanci. Metoda maximální relevance podle [9] spočívá v hledání funkce odpovídající vzorci

$$\max D(S, c), \quad D = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c), \quad (5.5)$$

která odpovídá $D(S, c)$ ze vzorce

$$\max D(S, c), \quad D = I(x_i, i = 1, \dots, m; c). \quad (5.6)$$

$I(x_i; c)$ ve vzorci 5.5 je největší vzájemná výměna informací v třídě c , což odráží největší závislost na cílové třídě. Z hlediska vzájemné výměny informací, je účelem výběru příznaků najít sadu funkcí S , se střední hodnotou všech vzájemně vyměněných informací mezi funkcí x_i a třídou c . [9]

Takto vybrané funkce budou s velkou pravděpodobností na sobě závislé, budou tedy redundantní. Při vysoké závislosti dvou prvků je možné jeden z nich odstranit, aniž by se výrazněji změnila odlišnost tříd. Využívá se proto dále vzorec pro minimální redundanci [9] :

$$\min R(S), \quad R = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i; x_j), \quad (5.7)$$

který v kombinaci se vztahem 5.5 dává dohromady kritérium mRMR, které lze vyjádřit pomocí vztahu

$$\max \phi(D, R), \quad \phi = D - R. \quad (5.8)$$

5.2.3 Genetické algoritmy

Genetické algoritmy patří do podmnožiny tzv. evolučních algoritmů, které vycházejí z principu inspirovaným evolucí. Podobně jako v přírodě, kde vlivem evoluce přežívají jen nejsilnější organismy, jsou jednotlivé objekty ohodnocovány pomocí fitness funkce, která určuje kvalitu daného řešení a na základě výsledku tohoto ohodnocení se rozhodne o jejich výběru. [4] Genetické algoritmy využívají ke své práci několika základních operátorů. Jsou to selekce, křížení, mutace a rekombinace. [4]

Selekce

Selekce označuje výběr jedinců z populace, kteří se budou moci stát rodiči a vznikne z nich tak nová populace. Vybírání bývají většinou jedinci s vysokým ohodnocením fitness funkcí, pro rozmanitost populace se však do výběru zahrnují i jedinci s nižší hodnotou této funkce. [4] Existuje více metod selekce jedinců:

1. Proporciální selekce, kde výběr závisí na fitness funkci a k zamezení postupného nahrazení populace jedincem s vysokou fitness hodnotou se využívá úpravy této funkce buď komprimací, kterou lze popsat pomocí vztahu

$$f'(i) = f(i) + \beta^t, \quad (5.9)$$

kde β označuje nejhorší fitness funkci v t -generaci, nebo pomocí sigma škálování:

$$f'(i) = \max(f'(i)(-E(f) - c \cdot \sigma_f), 0), \quad (5.10)$$

kde $c=2$. [4]

2. Lineární uspořádání, které vyžaduje setříděnou populaci, ve které se ohodnocují jedinci tak, že nejhorší jedinec má index 1 a nejlepší N . Selektce se pak provádí na základě následujícího vztahu: [4]

$$p(i) = \frac{2-s}{N} + \frac{2i(s-1)}{N(N-1)} \quad (5.11)$$

kde $i \in \{1, 2, \dots, N\}$ a s je hodnota selekční intenzity, která se volí obvykle 1,5 nebo 2.

3. Exponenciální uspořádání, které se podobá lineárnímu, s tím rozdílem, že rozložení pravděpodobnosti má exponenciální závislost. Lze ji vyjádřit vztahem:

$$p(i) = \frac{1 - e^{-i}}{c}, \quad (5.12)$$

kde $i \in \{1, 2, \dots, N\}$ a c je normalizační konstanta, vyjadřující velikost populace, volí se v rozsahu $0 < c < 1$. [4]

4. Turnajová selekce, jejíž výhodou je nepotřebnost setříděnosti populace. Princip této selekce je založen tak, že z N jedinců se vybere t jedinců, a z těchto t jedinců se vybere právě ten jedinec, jehož fitness funkce je nejvyšší. [4]

Křížení

Pod pojmem křížení se v genetických algoritmech rozumí vytváření nového jedince na základě genetických informací, předávaných jeho rodiči. Tak vzniká nová rozmanitá populace. Používají se dva druhy křížení:

1. *N*-bodové, které je založené na výběru jednoho bodu v genetických informacích rodičů, který ji rozdělí na dvě části a ty se potom vymění mezi potomky. [4]
2. Uniformní křížení, u kterého se s pravděpodobností 50% zdědí jeden gen po rodiči a druhý po druhém. Takto se zajistí obrana před brzkou konvergencí procesu. [4]

Mutace

Mutace spočívá v změně hodnot v genech v hodnoty opačné, tedy v bitové negaci. Probíhá však s velmi malou pravděpodobností, aby se zabránilo nestabilitě vývoje dalších populací. Výhoda tohoto procesu spočívá v tom, že tak může v dané generaci vzniknout dosud nepředaná vlastnost. [4] [12]

Rekombinace

Rekombinace je proces, při kterém se obnoví populace a existují dva způsoby, jakými probíhá:

1. S úplnou obnovou - rodiče jsou nahrazeni potomky v celém rozsahu
2. S částečnou obnovou - jeden potomek nahradí jedince, který je nejslabší

V praxi se ovšem často používají metody obě, kdy je například polovina staré populace nahrazena potomky [4]

Průběh procesu genetických algoritmů

Způsob, jakým probíhá celý proces lze shrnout takto: [4]

1. Počáteční inicializace a vytvoření náhodné populace
2. Ohodnocení jedinců fitness funkcí
3. Vytvoření cyklu, který bude probíhat dokud nedojde ke splnění stanovené podmínky
4. Inkrementace čísla populace
5. Selektce rodičů na základě stanovené metody výběru
6. Rekombinace vybraných rodičů
7. Mutace potomků
8. Ohodnocení potomků fitness funkcí
9. Obnovení populace

5.2.4 Křížová validace

Křížová validace, někdy nazývaná jako cross-validace je metoda, díky které lze hodnotit kvalitu naučeného algoritmu na malém počtu prvků, a využívá k práci jak množinu dat určenou k trénování tak i k testování klasifikátoru. Je to proto metoda vhodná i pro případy, kde je použito malé množství prvků. [4]

V praxi se používají dvě metody křížové validace, a to:

1. k-fold validace, kdy se data rozdělí například na deset skupin tak, že se vždy vyjme jedna desetina dat a použije se pro testování a zbytek dat bude sloužit k učení klasifikátoru. Tento postup se zopakuje desetkrát a výsledek se zprůměruje. V takovém případě se jedná o desetinásobnou křížovou validaci.
2. leave-one-out validace, kdy je ze skupiny dat vyjmut jeden prvek, který se použije pro testování a zbytek dat se použije k učení klasifikátoru. Tento postup se zopakuje tolikrát, kolik existuje v množině prvků. Nevýhodou této metody je její vysoká výpočetní náročnost. [4]

5.3 Klasifikace

Po extrakci parametrů je potřeba provést klasifikaci. Tato operace v podstatě rozhoduje, zda zkoumaný objekt nebo vlastnost patří do určité skupiny. V tomto případě se na základě příznaků extrahovaných z řečového signálu rozhodne, zda mluvčí patří do skupiny zdravých či lidí s PN. Existuje několik metod vhodných pro klasifikaci v rámci této práce. Konkrétně to mohou být například Gaussovy smíšené modely (GMM) nebo bayesovský klasifikátor. [12] [9]

5.3.1 Gaussovy smíšené modely

Tato metoda využívá statistické rozpoznávání vzorů. Vychází přitom z předpokladu, že některé statistické vlastnosti mohou být podobné u modelů stejných tříd. U tohoto algoritmu se využívá jedné nebo několika Gaussových funkcí rozložení pravděpodobnosti. Jednorozměrná funkce rozložení pravděpodobnosti je definována jako

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (5.13)$$

kde μ je střední hodnota a σ^2 rozptyl. Pro vícerozměrnou Gaussovu funkci platí vztah

$$f(x) = \frac{1}{\sigma\sqrt{2 \cdot \det(\Sigma)}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad (5.14)$$

kde Σ je kovarianční matice a μ je vektor středních hodnot.

Gaussovy smíšené modely potom pracují s lineární kombinací více Gaussových funkcí:

$$f(x) = \sum_{i=1}^M \alpha_i N_i(x). \quad (5.15)$$

Zde M představuje počet Gaussových funkcí rozložení pravděpodobnosti, α_i je váhový parametr a $N_i(x)$ je funkce rozložení pravděpodobnosti. Samotná GMM klasifikace proběhne pomocí vzorce

$$f(x) = \sum_{i=1}^M \alpha_i \cdot \frac{1}{\sigma \sqrt{2 \cdot \det(\Sigma)}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \quad (5.16)$$

ve kterém bude potřeba definovat kovarianční matici Σ , vektor středních hodnot μ a váhový parametr α . [12] [9]

5.3.2 Naivní bayesovský klasifikátor

Tento klasifikátor pracuje na principu podmíněné pravděpodobnosti [7]

$$P(y|x), \quad (5.17)$$

kde x představuje všechny příznaky analyzovaného vzorku a y třídu, přičemž přiřazuje vzorek k takové třídě, k jaké je jeho hodnota 5.17 nejvyšší. Po aplikaci Bayesova teorému lze obdržet vztah

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)}, \quad (5.18)$$

ve kterém $P(x)$ představuje apriorní (předem danou) pravděpodobnost daného vzorku a $P(y)$ apriorní pravděpodobnost třídy, přičemž obojí určíme z trénovací množiny. $P(x|y)$ - podmíněná pravděpodobnost daného vzorku ve třídě y , jež nelze z trénovací množiny určit, se získá pomocí vztahu 5.19 za předpokladu nezávislosti příznaků (odtud slovo „naivní“ v názvu):

$$P(x|y) = \prod_{i=1}^n P(x_i|y). \quad (5.19)$$

Zde x_1 až x_n představují jednotlivé příznaky vzorku x , jejichž pravděpodobnosti již lze z trénovací množiny určit. Jelikož zde není potřeba znát přesné hodnoty pravděpodobnosti z 5.18, lze vypustit jmenovatele zlomku, jež je konstantou nezávisle na třídě y . Na základě tohoto poznatku a za použití vztahů 5.18 a 5.19 lze definovat vztah

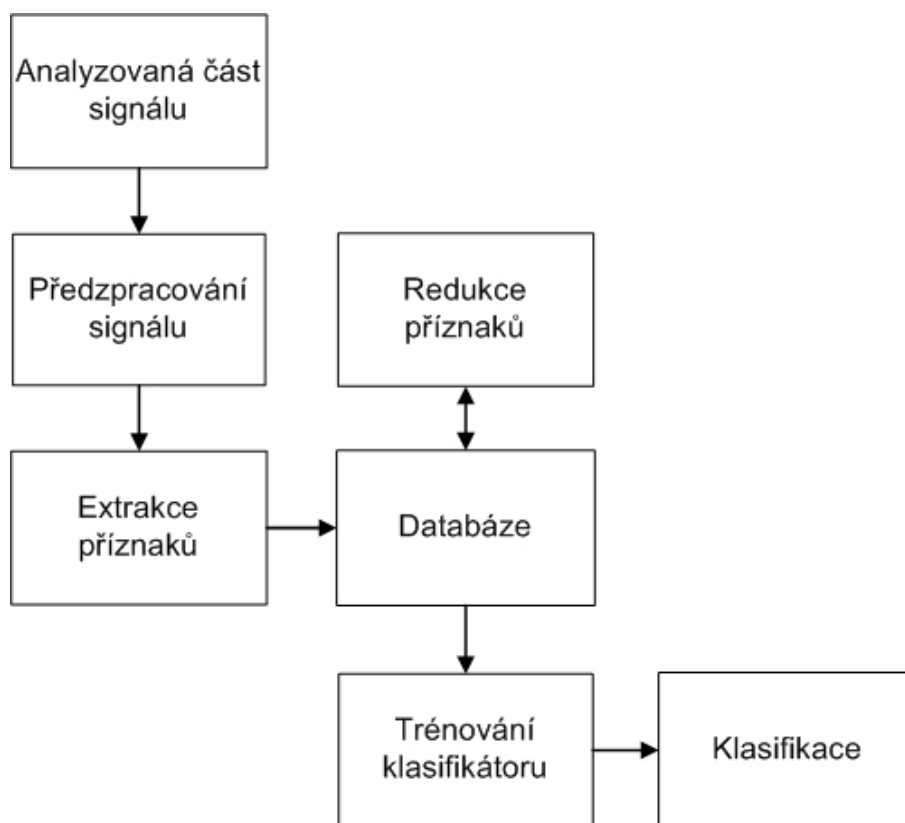
$$y = q(x) = \operatorname{argmax}_y \prod_{i=1}^n P(x_i|y)P(y). \quad (5.20)$$

6 METODA ŘEŠENÍ

Pro tuto práci jsou k dispozici zvukové nahrávky popisované na konci kapitoly 5.1. Od doby vzniku práce [14] již byly pořízeny další nahrávky, dá se však předpokládat, že aparatura a přístup pro pořizování jednotlivých záznamů zůstaly stejné, proto ani v tomto případě nebudou zvolené metody klást důraz na parametry odvozené z energie signálu, jako například srovnávání maximální síly hlasu u jednotlivých řečníků.

6.1 Klasifikace na základě analýzy samohlásek

Metoda této práce je založena na klasifikaci, vycházející z příznaků extrahovaných z jednotlivých samohlásek vyslovovaných dlouze. Celkové schéma metody lze rozdělit do dvou fází a znázornit pomocí tohoto blokového diagramu:



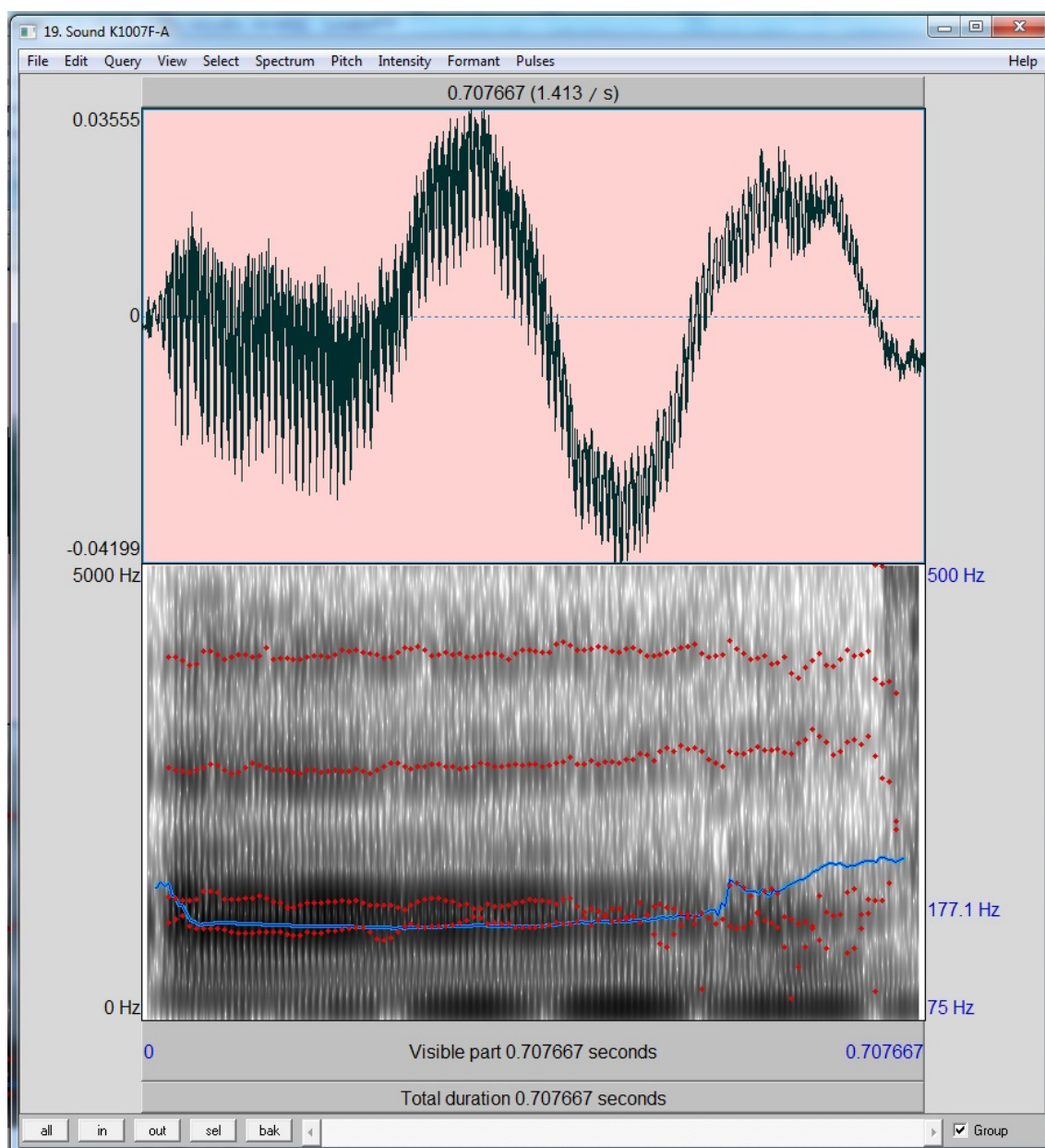
Obr. 6.1: Blokové schéma navržené metody.

6.2 Zpracování nahrávek

První tři bloky vedou k extrakci parametrů řečového signálu, na základě kterých bude poté provedena klasifikace. Protože však nahrávky obsahují velké množství různých cvičení, je nejprve nutné lokalizovat hledané samohlásky a uložit je do zvláštního souboru. Jednotlivé nahrávky lze časově označkovat a vybrat tak, u které jejich části předpokládáme výskyt signálu s projevem nemoci. Toho lze využít při automatizovaném výběru části nahrávky kterou jsme označili, případně lze využít jednodušší, avšak pracnější možnosti vystříhnutí požadované sekvence pomocí softwaru na editaci zvuku. V případě této práce byla zvolena druhá možnost. K tomu byl použit software s bezplatnou licencí Audacity, který má lepší podporu pro zpracování dlouhých nahrávek než program Praat, jehož jediná výhoda pro provedení této práce tkví v tom, že lze pomocí něj přesněji lokalizovat začátek a konec požadované části promluvy, a to za pomoci zobrazení křivek intenzity hlasu či základní hlasivkové frekvence. Proto byl použit spíše vyjimečně a to u nahrávek, kde nebylo možné zcela přesně nalézt tyto části promluvy jen za pomoci klasického zobrazení zvuku v časové oblasti. Celkově pro tyto účely bylo zpracováno 555 nahrávek, obsahující jednotlivě vyslovené samohlásky jak z kontrolní skupiny, tak od lidí trpících PN.

6.3 Extrakce parametrů ze signálu a předzpracování signálu

Samotná extrakce příznaků byla provedena kompletně ve volně stažitelném programu Praat. Autory tohoto software specializovaného na analýzu řeči ze zvukového signálu jsou Paul Boersma a David Weenink z Amsterodamské univerzity. Je několik způsobů, jak získat základní hlasové parametry pomocí tohoto Praatu, buď je můžeme získat jednotlivě za pomoci využívání nástrojů, které jsou k dispozici v menu po otevření nahrávky která se má zpracovat, nebo lze celý proces automatizovat za pomoci skriptu, napsaného přímo v Praatu. Při použití kteréhokoliv z obou způsobů je výstupem vždy textový soubor, ve kterém se nachází výsledek extrakce. Praat lze také využít pro tvorbu grafů a jiných grafických zobrazení parametrů, se kterými se pracuje.



Obr. 6.2: Extrakce parametrů pomocí programu Praat. Pod zobrazením zvukové nahrávky samohlásky „á“ je vykreslen spektrogram, jsou zde znázorněny jednotlivé formanty (červeně) a křivka ukazující základní hlasivkovou frekvenci (modře).

Vzhledem k velkému množství nastříhaných nahrávek a relativně velkému počtu příznaků, které se z nich v rámci této práce mají vypočítat, je daleko rychlejší použít skript, napsaný přímo pro potřeby této práce. Tento skript byl napsán na základě nápovědy přímo na stránkách tvůrců Praatu [5], jsou optimalizovány pro extrakci parametrů ze samohlásek a jeho zdrojové soubory byly zahrnuty do přílohy. Existují dvě verze tohoto skriptu, jedna verze je určena pro práci v klasické verzi Praatu pro Windows, kde lze po spuštění skriptu sledovat jeho práci s jednotlivými nahrávkami,

druhá verze je určena pro práci v jeho konzolové verzi. Oba tyto skripty jsou identické až na několik detailů, zahrnující například práci s analyzovanými položkami v menu programu, což ze zřejmých důvodů v konzolové verzi chybí. Spuštění skriptu v konzolové verzi programu Praat, například za pomoci příkazové řádky, se dá provést následujícím příkazem:

```
praatcon.exe -a scrcon.praat d:\cesta_k_nahravkam >> lib_jmeno.txt
```

kde parametr `-a` musí být přítomen v případě, že je potřeba výstup z Praatu přesměrovat do souboru, `scrcon.praat` je jméno souboru se spouštěným skriptem, dále následuje cesta k souborům které se mají zpracovat a jméno souboru za operátorem `>>` označuje výstupní soubor, do kterého budou extrahované příznaky zapsány. Příkaz musí být zadán ve složce, ve které se nachází konzolová verze programu, `Praatcon.exe` a zdrojový skript. Extrahovány budou parametry ze všech souborů, nacházející se v adresáři, zadaném jako vstupní parametr. Konzolová verze skriptu narozdíl od verze pro klasický Praat generuje výstupní soubor s extrahovanými parametry, definovaný v posledním parametru spouštěcího příkazu. Druhý skript při spuštění automaticky generuje jednotlivé soubory s příponou `.txt`, pojmenované stejně jako nahrávka, ze které jsou parametry extrahované, navíc však ještě výstupní soubor, který je typově totožný se souborem, generovaným konzolovou verzí skriptu, jehož název je definovaný v kódu skriptu jako „komplet.txt“. Všechny tyto výsledné soubory se automaticky ukládají na místo, odkud je spouštěný skript, nikoliv program Praat. Ukázka skriptu pro extrakci příznaků:

```
jmeno_souboru$ = selected$ ("Sound")
select Sound 'jmeno_souboru$'
To Formant (burg)... 0.0 5 5500 0.025 50
select Sound 'jmeno_souboru$'
To Pitch... 0.0 75 180
select Sound 'jmeno_souboru$'
To PointProcess (periodic, cc)... 75 180

select Formant 'jmeno_souboru$'
a = Get mean... 1 0.0 0.0 Hertz
b = Get minimum... 1 0.0 0.0 Hertz Parabolic
c = Get maximum... 1 0.0 0.0 Hertz Parabolic
```

Na základě literatury uvedené v kapitole 5.1 a dalších prací [9] [18] [4] bylo v rámci této práce vybráno 45 příznaků, uvedených v tabulce a umístěné v příloze A.1.

Po extrakci z jednotlivých nastříhaných nahrávek byly tyto příznaky importovány do databáze v Matlabu, a v této podobě jsou zároveň se skripty pro jejich získání umístěné na DVD, přiloženému k této práci. Databáze byla rozdělena na několik skupin a to podle pohlaví řečníků na analyzovaných nahrávkách, podle vyslovované samohlásky a třídy, do které nahrávka patří. Formát databází jsou matice, kde jednotlivé řádky představují příznaky, jejichž pořadí odpovídá pořadí uvedenému v tabulce v příloze, a sloupce každou zpracovanou nahrávku. Jednotlivé databáze s příznaky jsou značeny následujícím způsobem:

- pd_ma - Databáze obsahuje příznaky mužských řečníků s PN a extrahovány byly ze samohlásky [á]
- kntrl_fe - Databáze obsahuje příznaky ženských řečníků a extrahovány byly ze samohlásky [é]

6.4 Redukce příznaků

Jak již bylo naznačeno v kapitole 5.2, pro navýšení úspěšnosti klasifikace je zapotřebí vybrat takové příznaky, které přispějí ke kvalitnější klasifikaci a nevnášejí do ní naopak chaos. V této práci byly s různou úspěšností vyzkoušeny tři metody redukce.

6.4.1 Metoda minimální redundance a maximální relevance (mRMR)

Princip, na kterém tato metoda pracuje, byl podrobně popsán v kapitole 5.2.2. Pro realizaci byl využit skript do Matlabu autora [19], jenž jej uveřejnil na stránkách The MathWorks, Inc. - Matlab central a lze jej zde bezplatně stáhnout. Pro výpočet mRMR byly zadány následující vstupní parametry:

- Počet vybraných příznaků - 20
- Použitá metoda - MID (Mutual Information Difference)
- Typ dat - Categorical

Prvních pět nejlépe ohodnocených příznaků metodou mRMR shrnují následující tabulky:

Příznak	Ohodnocení
F2b max-min	0.976
jitter (localabs)	-0.000
hnr nh	-0.000
hnr ut	-0.000
jitter (local)	-0.166

Tab. 6.1: Příznaky vybrané pomocí metody mRMR - muži

Příznak	Ohodnocení
F2b max	0.976
jitter (localabs)	-0.000
hnr nh	-0.000
hnr ut	-0.000
jitter (ppq5)	-0.056

Tab. 6.2: Příznaky vybrané pomocí metody mRMR - ženy

Detailnější výpis výsledků lze nalézt v příloze A.5.

6.4.2 Míra geometrické oddělitelnosti

Tato metoda je podrobněji rozebrána v kapitole 5.2.1. Vzhledem k tomu, že pro realizaci těchto výpočtů je potřeba jen několika jednoduchých statistických funkcí, byla pro realizaci vybrána aplikace Microsoft Excel. Výpočet pro každý příznak probíhal následujícím způsobem:

1. Vypočtení střední hodnoty určeného příznaku
2. Vypočtení kvadrátu vzdálenosti mezi oběma třídami
3. Vypočtení kvadrátu rozptylu
4. Vypočtení aritmetického průměru kvadrátů rozptylu obou tříd
5. Vypočtení kvality příznaku Q podle vzorce 5.4

Výpočty byly provedeny jednotlivě pro každou samohlásku a dodatečně i pro skupinu obsahující všechny samohlásky dohromady. Výsledky byly podrobně zapsány v posledním listu přiloženého sešitu geometrie.xlsx. Z této metody byly nejlépe hodnoceny příznaky, shrnuté v následujících tabulkách:

Příznak	Hodnota Q
jitter (localabs)	1,07E-07
hnr ut	0,091195466
hnr nh	0,351320818
shimmer (localabs)	0,727241431
jitter (ppq5)	0,780546308

Tab. 6.3: Příznaky vybrané pomocí metody geometrické oddělitelnosti - muži

Příznak	Hodnota Q
hnr nh	0,10874093
hnr ut	0,116206725
shimmer (localabs)	0,575123007
hnr hn	0,948339174
shimmer (apq3)	0,94848844

Tab. 6.4: Příznaky vybrané pomocí metody geometrické oddělitelnosti - ženy

Celá tabulka vypočtených kvalit jednotlivých příznaků A.3 A.4 je vytištěna a přidána k přílohám a nachází se také v souboru *geometrie.xlsx* přiloženém na DVD, kde se také nachází podrobnější výpis celkových výsledků. Z těchto výsledků lze jednoznačně vyzvednout příznak jitter (localabs) u mužů a příznaky odvozené z harmonicity signálu u mužů i žen. Ostatní příznaky se již velmi blížily k 1, a nemělo tedy smysl hodnotit jejich kvalitu v porovnání s výše uvedenými.

Největší rozdíl mezi muži a ženami v této metodě je patrný u nejlépe hodnoceného příznaku jitter (localabs), kde u mužů vykazoval jen velmi malou závislost na vyslovované samohlásce v rámci intervalu $0 < x < 1$, kdežto u žen byl tento příznak kladně ohodnocen pouze u samohlásek [á], [ó] a [ú]. U ostatních kladně hodnocených příznaků (ostatní příznaky odvozené z jitteru, mimo jitter ddp, hnr nh, hnr ut a shimmer localabs) se závislost na jednotlivých samohláskách projevuje na obou pohlavích, i když s různým výsledkem. Podrobnější výsledky jsou dobře patrné z již zmiňovaných příloh A.3 A.4

6.4.3 Genetické algoritmy

Jako poslední metoda redukce byly vyzkoušeny genetické algoritmy, které byly podrobněji popsány v kapitole 5.2.3. Realizace potřebných výpočtů byla provedena v programu Matlab, za pomoci funkce *ga.m*, [13] jež je součástí tohoto programu. Podrobný návod k využití této funkce je zakomentován přímo v jejím kódu a lze ji

tak pro tyto účely využít pomocí nové funkce, která specifikuje požadované vstupní parametry. Tyto funkce ke své práci využívají i křížovou validaci a klasifikátor, které byly teoreticky rozebrány v kapitolách 5.2.4 5.3 a jejich praktické realizace jsou dále popsány v 6.5. Výstup této metody se ustálil na těchto příznacích:

Příznaky	
F0 mean	F0 min
F0 max-min	F1 mean
F2b mean	F2b std
F2b max-min	jitter (ppq5)
F1 max	F2 max-min
F1b min	hnr nh
hnr hn	hnr ut

Tab. 6.5: Příznaky vybrané metodou genetických algoritmů pro muže

Příznaky	
F1 max	F1 max-min
F2 mean	F2 median
F2 std	F1b mean
F1b max	F1b max-min
F2b max	jitter (local)
jitter (ppq5)	jitter (ddp)
hnr nh	

Tab. 6.6: Příznaky vybrané metodou genetických algoritmů pro ženy

6.5 Klasifikace

Výběr metody klasifikace byl proveden na základě kapitoly 5.3. Ze zmiňovaných klasifikátorů byl vybrán naivní bayesovský klasifikátor, jehož implementace v prostředí Matlab je přiložena v přílohách této práce. Tento skript využívá ke své práci i křížovou validaci, která byla podrobně rozebrána v kapitole 5.2.4. Skripty jsou nastaveny tak, aby mohly pracovat s maticí dat, využívané v této práci. Matice pro obě pohlaví mají stejný formát jako databáze pro jednotlivé samohlásky zmiňované v 6.3, pro účely klasifikace však byly obě třídy i se všemi samohláskami importovány do jedné databáze. Pro muže je tedy výsledná matice složena z 24 nahrávek pro každou samohlásku z obou tříd, jdoucí po sobě v obvyklém pořadí [a] [e] [i] [o] [u] a to nejprve z nahrávek z kontrolní skupiny a poté ze skupiny lidí s PN. Důvod, proč nebyly použity všechny analyzované nahrávky byl ten, že klasifikátor ke své činnosti vyžaduje stejný počet nahrávek z obou tříd, v tomto případě vzniklo omezení počtem nahrávek kontrolní skupiny. U žen vznikla situace opačná, kdy použitelné nahrávky ze skupiny lidí s PN dosáhly počtu 23 nahrávek na samohlásku, z toho vždy jedna pro každou samohlásku byla vyřazena z důvodu lepšího rozdělení celé databáze pro křížovou validaci. Vznikla tak matice s celkovým počtem 220 nahrávek pro ženy a 240 pro muže. Klasifikátor tedy bylo nutné nastavovat pro každé pohlaví zvlášť, především pak parametry pro křížovou validaci. Ta byla u mužů nastavena tak, aby bylo vždy vybráno 12 nahrávek z obou tříd pro testování a zbytek se použil pro trénování klasifikátoru. Celý proces se opakoval desetkrát, jednalo se tedy o desetinásobnou křížovou validaci. U žen proces probíhal stejným způsobem s jediným rozdílem, což byl počet nahrávek pro testování, který byl nastaven na 11.

Nejprve byla provedena klasifikace u obou pohlaví bez jakékoliv redukce. U mužů dosahovala klasifikační úspěšnost 55,83% a u žen 45%. Poté byla klasifikace provedena znovu, a to s redukovánými příznaky na základě všech tří metod, provedených v rámci této práce. Celkovou klasifikační úspěšnost u všech tří metod shrnuje následující tabulka:

Pohlaví	Bez redukce	Q	mRMR	GA
Muži	55,83%	58,33%	61,25%	73,33%
Ženy	45%	47,73%	49,09%	66,36%

Tab. 6.7: Tabulka shrnující klasifikační úspěšnost u všech použitých metod. Q značí metodu geometrické oddělitelnosti, mRMR minimální redundanci a maximální relevanci a GA genetické algoritmy

Z předchozí tabulky lze vyčíst, že nejúspěšnější metoda redukce při použití

bayesovského klasifikátoru jsou jednoznačně genetické algoritmy. Následovány jsou metodou mRMR a nejhůře zde dopadla geometrická oddělitelnost, která přinesla jen malé zlepšení výsledků. Také je zde patrný podstatný rozdíl v úspěšnosti klasifikace u žen oproti mužům.

Na závěr byla otestována úspěšnost klasifikace na matici dat, která nebyla rozdělena podle pohlaví. Úspěšnost klasifikace za využití stejného klasifikátoru jako dosud dosahovala 52,39%. Dále byly použité příznaky redukovány pomocí metody, která dosud jevila nejlepší výsledky, genetických algoritmů. Vybrané příznaky shrnuje následující tabulka:

Příznaky			
F0 min	F0 median	F0 sd	F1 mean
F1 std	F2 median	F1b max-min	F2b min
F2b max	jitter (localabs)	jitter (ddp)	hnr nh

Tab. 6.8: Příznaky vybrané metodou genetických algoritmů pro data bez specifikace pohlaví

Po redukování parametrů podle tabulky 6.8 dosáhla výsledná úspěšnost klasifikace 60,68%, což je méně, než při klasifikaci provedené zvlášť na každém pohlaví po redukcii pomocí GA.

7 ZÁVĚR

Tato bakalářská práce se zabývá diagnózou Parkinsonovy choroby na základě analýzy řečového záznamu, přičemž využívá faktu, že hypokinetická dysartrie, popsaná v kapitole 2, bývá jedním z průvodních jevů této nemoci. Metoda této práce tedy spočívá v odhalení HD pomocí příznaků, extrahovaných z řečového signálu.

V první kapitole této práce jsou popsány základní vlastnosti vzniku lidské řeči a několik jejích důležitých parametrů.

Následuje charakteristika hypokinetické dysartrie, kde je v několika stručných kapitolách shrnuto, jak se tato porucha tvorby řeči projevuje na jednotlivých aspektech tvorby hlasu.

Dále jsou v této práci popsány základní metody pro zpracování řečového signálu. Tato kapitola nejprve popisuje rozdíl mezi znělou a neznělou řečí a jakým způsobem je lze od sebe odlišit, dále popisuje několik základních kroků v digitalizaci signálu a nakonec definuje metodu, díky které lze rozpoznat počátek a konec promluvy, stejně jako znělou řeč od neznělé.

Ve čtvrté kapitole jsou vysvětleny postupy, jakým jsou extrahovány jednotlivé základní příznaky, používané ke klasifikaci a také je zde stručně popsáno několik statistických funkcí, sloužících k výpočtu příznaků odvozených ze základních.

Pátá kapitola se zabývá popisem metody diagnózy PN z hlasu. V první části je popsáno několik experimentů, na jejichž základě byly nakonec vybrány příznaky použité ke klasifikaci. Následuje charakteristika jednotlivých metod, které slouží k redukci příznaků a tím k docílení úspěšnější klasifikace. Konec této kapitoly se zaměřuje na popis dvou klasifikátorů, které je možné pro řešení této práce použít.

V praktické části této práce je nejprve popsána a pomocí blokového diagramu znázorněna zvolená metoda řešení, jejíž princip je založen na extrakci zvolených příznaků ze samohlásek. Následuje popis způsobu zpracování nahrávek, které byly nastříhány zvláště pro každou samohlásku, pohlaví i třídu pomocí volně přístupného programu Audacity. Dále je popsán způsob extrakce zvolených příznaků pomocí programu Praat, k čemuž byly vytvořeny skripty, přiložené k této práci. Způsob práce s těmito skripty je také objasněn v této kapitole. Následuje popis realizace metod mRMR, geometrické oddělitelnosti a genetických algoritmů. Jsou zde také přiloženy tabulky s příznaky, označenými jednotlivými metodami jako nejvhodnější. Poslední část této kapitoly se zabývá samotnou klasifikací. K tomu byl zvolen naivní bayesovský klasifikátor, který bez jakékoliv redukce dosahoval úspěšnosti 55,83% u mužských a 45% u ženských řečníků. Nejlepších výsledků potom dosahoval při použití genetických algoritmů, kdy u mužských řečníků dosahoval úspěšnosti 73,33% a 66,36% u ženských. Redukce pomocí mRMR přinesla zpřesnění 5,42% u mužů a 4,09% u žen. Nejhorší výsledky projevuje metoda geometrické oddělitelnosti, jejíž

přínos byl jen 2,5% u mužů a 2,73% u žen. Celkově lze pozorovat horší přesnost klasifikace u ženského pohlaví, což může být způsobeno použitím příznaků, jejichž vlastnosti jsou na pohlaví závislé. Na závěr byla provedena klasifikace matice všech dat, bez specifikace pohlaví, kdy úspěšnost bez využití redukce dosahovala 52,39%, avšak ani po aplikování genetických algoritmů nepřesáhla 60,68%, což je méně, než v případě kdy byla obě pohlaví klasifikována zvlášť.

Na základě poslední kapitoly lze jako příznaky vhodné ke klasifikaci doporučit jitter (localabs), příznaky odvozené z harmonicity signálu (hnr ut, hnr nh) a příznaky popisující maximální a minimální šířky pásem prvních dvou formantů. Jako nejvíce závislé na pohlaví se projevily příznaky odvozené ze základní hlasivkové frekvence. Závislost na jednotlivých samohláskách se projevovala především u příznaků hnr nh, hnr ut, shimmer (localabs) a jitter (localabs), poslední zmiňovaný pouze u žen. Určování těchto závislostí však vychází z metody geometrické oddělitelnosti, která nepřinesla příliš velký přínos pro přesnost. Taktéž lze na základě výsledků v části 6.8 doporučit, v případě zachování stejné metody, provádět klasifikaci o každého pohlaví zvlášť.

Jako další rozšíření této práce by bylo možné doporučit především otestování dalších klasifikátorů, jako GMM nebo kNN a závislost přesnosti zmiňovaných redukcí na těchto klasifikátorech. Taktéž by bylo možné aplikovat ostatní metody redukce, včetně těch, které byly zmiňované v této práci, na klasifikaci dat obsahující extrahované příznaky pro obě pohlaví, a to z důvodu, že klasifikace bez využití jakékoliv redukce byla v tomto případě přibližně střední hodnotou mezi oběma úspěšnostmi u dat s oddělenými pohlavími a až po aplikování GA byl celkový výsledek nižší než u ženských řečníků. Dále by bylo vhodné vyzkoušet klasifikaci i na základě jiných metod, než extrakce uvedených příznaků ze samohlásek, a to například metodu založenou na artikulační rychlosti. Taktéž by bylo možné vyzkoušet srovnání postupného klesání hlasitosti u vyslovování hlásek [pataka], zmiňovanou v [21] a to i navzdory nekalibrovanému mikrofону, protože princip je založen na srovnání relativní energie signálu na začátku a na konci promluvy u každého řečníka zvlášť.

LITERATURA

- [1] ADAMEC M. *Moderní rozpoznávače řečové aktivity*. Brno: Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2008. 75 s. Diplomová práce. Vedoucí práce byl Mgr. Pavel Rajmic, Ph.D.
- [2] AILEEN, K. Ho, John L. BRADSHAW, Robert IANSEK a Robin ALFREDSON. *Speech volume regulation in Parkinson's disease: effects of implicit cues and explicit instructions*. Neuropsychologia [online]. 1999, č. 37, s. 1453–1460 [cit. 2012-11-19]. Dostupné z: <<http://www.sciencedirect.com/science/article/pii/S0028393299000676>>.
- [3] AILEEN, K. Ho, John L. BRADSHAW, Robert IANSEK a Robin ALFREDSON. *Predicting Severity of Parkinson's Disease from Speech*. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) [online]. 2010, č. 32, s. 5201–5204 [cit. 2012-11-19]. Dostupné z: <http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=1&SID=S1f1AhmbB0gEHc27F0i&page=1&doc=2>.
- [4] BÍLÝ, Ondřej. *Moderní řečové příznaky používané při diagnóze chorob*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2010. 76 s. Vedoucí práce byl Ing. Jiří Mekyska
- [5] BOERSMA, Paul a David WEENINK. *Praat: doing phonetics by computer*. Scripting [online]. 2013 [cit. 2013-05-13]. Dostupné z: <<http://www.fon.hum.uva.nl/praat/manual/Scripting.html>>.
- [6] BOLKOVÁ, Zuzana. *Léčebně-rehabilitační plán a postup u Parkinsonovy choroby* [online]. 2012 [cit. 2012-11-03]. Bakalářská práce. Masarykova univerzita, Lékařská fakulta. Vedoucí práce Eva Pospíšilová.
- [7] DOHNALOVÁ, Ludmila. *Systém pro selekci příznaků z EEG signálu* [online]. 2012 [cit. 2012-11-03]. Bakalářská práce. České vysoké učení technické v Praze, Fakulta elektrotechnická, Katedra kybernetiky. Vedoucí práce Ing. Václav Gerla. Dostupné z: <<http://cyber.felk.cvut.cz/research/theses/papers/179.pdf>>.

- [8] GOBERMAN, Alexander M. a Lawrence W. ELMER. *Acoustic analysis of clear versus conversational speech in individuals with Parkinson disease*. *Journal of Communication Disorders* [online]. 2005, č. 38, s. 215–230 [cit. 2012-11-19]. Dostupné z: <<http://www.sciencedirect.com/science/article/pii/S0021992404000991>>.
- [9] KARÁSEK, Michal *Diagnóza Parkinsonovy choroby z řečového signálu*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav telekomunikací, 2011. 50 s. Vedoucí práce byl Ing. Jiří Mekyska
- [10] KOVAŘÍK, J. *Detektor řečové aktivity v signálovém procesoru*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2012. 64 s. Vedoucí diplomové práce Ing. Petr Sysel, Ph.D.
- [11] KRČMOVÁ, M.: *Fonetika*. Elektronické texty. MU Brno 2003. Dostupné z: <<http://is.muni.cz/do/1499/el/estud/ff/js07/fonetika/materialy/index.html>>.
- [12] LNĚNIČKA, J. *Rozpoznávání emočních stavů na základě řečového záznamu*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2011. 45 s. Vedoucí semestrální práce prof. Ing. Zdeněk Smékal, CSc.
- [13] Copyright 2003-2012 The MathWorks, Inc.
- [14] MEKYSKA, Jiří, Irena REKTOROVÁ a Zdeněk SMÉKAL. *Selection of Optimal Parameters for Automatic Analysis of Speech Disorders in Parkinson's Disease*. 34th International Conference on Telecommunications and Signal Processing (TSP) [online]. 2011, č. 34, s. 408–412 [cit. 2012-11-19]. Dostupné z: <http://apps.webofknowledge.com/full_record.do?product=WOS&search_mode=GeneralSearch&qid=2&SID=S1f1AhmbB0gEHc27F0i&page=1&doc=1>.
- [15] MIZERA, P. *Pitch-synchronní segmentace řečového signálu*. Praha: České vysoké učení technické v Praze, Fakulta elektrotechnická, 2012. 58 s. Vedoucí diplomové práce Doc. Ing. Petr Pollák, CSc.
- [16] Navrátil, M. *Rozpoznávání emočních stavů pomocí analýzy řečového signálu*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2008. 66 s. Vedoucí diplomové práce prof. Ing. Zdeněk Smékal, CSc.
- [17] HILLEROVÁ, Kateřina. *Využití psychomotorických cvičení u seniorů s Parkinsonovou chorobou* [online]. 2012 [cit. 2012-10-21]. Diplomová práce. Masarykova univerzita, Fakulta sportovních studií. Vedoucí práce Dagmar Trávníková.

- [18] PAVLÍČEK, T. *Zpracování řečových signálů ve spektrální oblasti*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2012. 37 s., 1 s. příloh. Vedoucí semestrální práce prof. Ing. Milan Sigmund, CSc.
- [19] PENG, Hanchuan. *MRMR Feature Selection (using mutual information computation)*. In: THE MATHWORKS, INC. Matlab Central [online]. 2007 [cit. 2013-05-25]. Dostupné z: <<http://www.mathworks.com/matlabcentral/fileexchange/14608>>
- [20] PSUTKA, Josef, Luděk MÜLLER, Jindřich MATOUŠEK a Vlasta RADOVÁ. *Mluvíme s počítačem česky*. 1. vyd. Praha: Academia, 2006. ISBN 80-200-1309-1.
- [21] RUSZ, Jan, Roman ČMEJLA, Hana BACHUROVÁ a Jan JANDA. *Akustická analýza intenzity a rychlosti řeči u Parkinsonovy nemoci*. [online]. s. 5 [cit. 2012-11-19]. Dostupné z: <http://dsp.vscht.cz/konference_matlab/MATLAB08/prispevky/091_rusz.pdf>.
- [22] RUSZ, Jan, Roman ČMEJLA a Hana RŮŽIČKOVÁ. *Analýza základní frekvence, amplitudového a frekvenčního kolísání hlasivek u Parkinsonovy nemoci*. Akustické listy [online]. 2009, roč. 15, č. 1, s. 13-18 [cit. 2013-05-06]. Dostupné z: <<http://www.czakustika.cz/listy/casopis/15-1.pdf>>.
- [23] SKODDA, Sabine, Wenke GRÖNHEIT a Uwe SCHLEGEL. *Intonation and Speech Rate in Parkinson's Disease: General and Dynamic Aspects and Responsiveness to Levedopa Admission*. Journal of Voice. 2011, roč. 25, č. 4, s. 199–205.
- [24] VONDRA, Martin. *Kepstrální analýza řečového signálu*. Elektrorevue [online]. 2001, č. 48 [cit. 2012-10-27]. ISSN 1213-1539. Dostupné z: <<http://www.elektrorevue.cz/clanky/01048/index.html>>.
- [25] ZAMIŠKOVÁ, Gabriela, Pavel RESSNER, Jana DLOUHÁ a Dana ŠIGUTOVÁ. *Poruchy řeči u Parkinsonovy nemoci* [online]. 2010 [cit. 2012-10-22]. Dostupné z: <http://www.solen.sk/index.php?page=pdf_view&pdf_id=4371&magazine_id=3>.
- [26] ZWETTLEROVÁ, H. *Speciálně andragogická intervence v oblasti běžných denních činností u osob s Parkinsonovou nemocí* [online]. 2012 [cit. 2012-10-21]. Diplomová práce. Masarykova univerzita, Pedagogická fakulta. Vedoucí práce Kateřina Sayoud Solárová.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

PN Parkinsonova nemoc

HD Hypokinetická dysartrie

DFT Diskrétní Fourierova transformace

GA genetické algoritmy

GMM Gaussovy smíšené modely

IDFT Inverzní diskrétní Fourierova transformace

mRMR minimum Redundancy Maximum Relevance Feature Selection

f_{vz} vzorkovací kmitočet

kNN metoda nejbližších sousedů

SEZNAM PŘÍLOH

Jednotlivé přílohy této práce:

- Tabulka příznaků vybraných k extrakci z nahrávek rozdělená na dvě části, A.1 a A.2
- Tabulku s výslednými hodnotami geometrické oddělitelnosti pro muže A.3
- Tabulku s výslednými hodnotami geometrické oddělitelnosti pro ženy A.4
- Tabulku s výstupními daty metody mRMR A.5
- Přiložené CD, které obsahuje skripty použité pro klasifikaci, soubor geometrie.xlsx použitý pro výpočet Q geometrické oddělitelnosti i s výsledky, skripty použité pro extrakci příznaků pro konzolovou i klasickou verzi Praatu a složku Data, obsahující matice vypočtených příznaků pro jednotlivé skupiny.

A PŘÍLOHA 1

A.1 Tabulka vybraných příznaků 1

Označení	Popis
F_0 mean	střední hodnota F_0
F_0 min	minimální hodnota F_0
F_0 max	maximální hodnota F_0
F_0 median	medián F_0
F_0 std	směrodatná odchylka F_0
F_0 max-min	rozdíl maximální a minimální hodnoty F_0
F_0 var	relativní kolísání F_0
F_0 sd	relativní směrodatná odchylka F_0
F_1 mean	střední hodnota F_1
F_1 min	minimální hodnota F_1
F_1 max	maximální hodnota F_1
F_1 median	medián F_1
F_1 std	směrodatná odchylka F_1
F_1 max-min	rozdíl maximální a minimální hodnoty F_1
F_2 mean	střední hodnota F_2
F_2 max	maximální hodnota F_2
F_2 median	medián F_2
F_2 std	směrodatná odchylka F_2
F_2 max-min	rozdíl maximální a minimální hodnoty F_2
F_{1b} mean	střední hodnota šířky pásma F_{1b}
F_{1b} min	minimální hodnota šířky pásma F_{1b}
F_{1b} max	maximální hodnota šířky pásma F_{1b}
F_{1b} median	medián šířky pásma F_{1b}
F_{1b} std	směrodatná odchylka šířky pásma F_{1b}
F_{1b} max-min	rozdíl maximální a minimální hodnoty šířky pásma F_{1b}
F_{2b} mean	střední hodnota šířky pásma F_{2b}
F_{2b} min	minimální hodnota šířky pásma F_{2b}
F_{2b} max	maximální hodnota šířky pásma F_{2b}

A.2 Tabulka vybraných příznaků 2

Označení	Popis
F_2b median	medián šířky pásma F_2b
F_2b std	směrodatná odchylka šířky pásma F_2b
F_2b max-min	rozdíl maximální a minimální hodnoty šířky pásma F_2b
jitter local	lokální kolísání periody
jitter localabs	absolutní hodnota kolísání periody
jitter rap	průměrný rozdíl mezi periodou a průměrem dvou sousedních period
jitter ppq5	průměrný rozdíl mezi periodou a průměrem čtyř sousedních period
jitter ddp	průměrný rozdíl mezi za sebou jdoucích rozdílů period
schimmer local	lokální kolísání amplitudy
schimmer local abs	lokální kolísání amplitudy v dB
schimmer apq3	průměrný rozdíl mezi amplitudou a průměrem dvou sousedních amplitud
schimmer apq5	průměrný rozdíl mezi amplitudou a průměrem čtyř sousedních amplitud
schimmer apq11	průměrný rozdíl mezi amplitudou a průměrem deseti sousedních amplitud
schimmer ddp	průměrný rozdíl mezi za sebou jdoucích rozdílů amplitud
hnr hn	střední hodnota poměru signál-šum v dB
hnr nh	střední hodnota poměru signál-šum
hnr ut	střední hodnota autokolerace

A.3 Výsledky geometrické oddělitelnosti - muži

Muži	a	e	i	o	u	Celek
F0 mean	0,99992209	0,99993979	0,99999999	0,99999791	0,99904685	0,99999976
F0 min	0,99998636	0,99998159	0,99940322	0,99998422	0,99926633	0,99996868
F0 max	0,99994254	0,99994898	0,99998211	0,99986504	0,99992894	0,99999808
F0 median	0,99999491	0,99991833	0,99997625	0,99999827	0,9997239	0,99998611
F0 std	0,99739738	0,99992492	0,99999774	0,99826489	0,99580873	0,99989396
F0 max-min	0,99983725	0,99996827	0,99999785	0,99995072	0,9999209	0,99999419
F0 var	0,99970402	0,99994173	0,99999813	0,99994769	0,99996803	0,9999991
F0 sd	0,99999491	0,99991833	0,99997625	0,99999827	0,9997239	0,999745
F1 mean	0,99992111	0,9999593	0,99999551	0,99993368	0,99999826	0,99999842
F1 min	0,99995462	0,9999982	0,9999637	0,99999974	0,9999632	0,99999826
F1 max	0,9999997	0,99999985	1	0,99999934	0,99999999	0,99999969
F1 median	0,99997226	0,99999181	0,99996075	0,99996605	0,9999976	0,99999879
F1 std	0,99999735	0,99999808	1	0,9999966	0,99999959	0,9999946
F1 max-min	0,9999999	0,99999997	0,99999991	1	1	0,99999943
F2 mean	0,99999482	0,99999979	0,9999985	0,99999997	0,99999909	0,99999554
F2 max	0,99999979	0,99999998	0,99999992	0,99999991	0,99999996	0,99999997
F2 median	0,99999985	0,99999746	0,99999999	1	0,99999978	0,99999608
F2 std	0,99983725	0,99996827	0,99999785	0,99995072	0,9999209	0,99999419
F2 max-min	0,99999998	0,99999997	0,99999993	1	0,99999993	0,99999815
F1b mean	0,99999101	0,99998116	0,99999098	0,99999131	0,9999889	0,99999562
F1b min	0,99985282	0,99958624	0,99991233	0,99999908	0,99989663	0,99994539
F1b max	1	0,99999998	0,99999998	0,99999981	0,99999989	0,99999999
F1b median	0,999992	0,99987719	0,99998196	0,99998274	0,99996599	0,99999572
F1b std	0,99999946	0,99999829	0,99999303	0,99998614	1	0,99999515
F1b max-min	1	0,99999998	0,99999998	0,99999986	0,99999995	0,99999999
F2b mean	0,99999934	0,99997778	1	0,99999814	1	0,99999711
F2b min	0,99998667	0,99989754	0,99993628	0,99997734	0,99999354	0,99999847
F2b max	0,99999998	0,99999999	1	0,99999999	0,99999986	0,99999997
F2b median	0,99999868	0,99997888	0,99999882	0,99999892	0,99999745	0,99999724
F2b std	0,9999996	0,9999965	0,9999992	0,99999997	0,99999717	0,99999848
F2b max-min	0,99999998	0,99999999	1	0,99999999	0,99999999	0,99999998
jitter (local)	0,45827055	0,95374541	0,48254059	0,87306394	0,99850723	0,93819126
jitter (localabs)	3,1224E-08	4,4668E-07	2,9538E-06	1,0682E-07	3,7569E-07	1,2929E-06
jitter (rap)	0,1748091	0,83317788	0,38761588	0,69132239	0,987936	0,82863754
jitter (ppq5)	0,32810783	0,44350436	0,82217575	0,95024791	0,93459205	0,78054631
jitter (ddp)	0,65595153	0,97823706	0,85067178	0,9527335	0,99864503	0,97753831
shimmer (local)	0,99450477	0,99957902	0,99975451	0,99989209	0,99999887	0,99955856
shimmer (localabs)	0,53008176	0,74713038	0,8908075	0,85179127	0,91791244	0,72724143
shimmer (apq3)	0,99560769	0,99992158	0,99783576	0,99938067	0,99783325	0,99625964
shimmer (apq5)	0,98451371	0,9993725	0,9944761	0,99513625	0,99671436	0,99887523
shimmer (apq11)	0,99426684	0,99994258	0,9960252	0,99958008	0,99987775	0,9993221
shimmer (dda)	0,99951005	0,99999129	0,99975907	0,99993115	0,99975879	0,99958302
hnr hn	0,99646094	0,9954184	0,98713481	0,99917191	0,98046429	0,99829244
hnr nh	0,14614791	0,54034496	0,51560148	0,07830099	0,02268157	0,35132082
hnr ut	0,84061605	0,18222554	0,09305334	0,10591864	0,00542612	0,09119547

A.4 Výsledky geometrické oddělitelnosti - ženy

Ženy	a	e	i	o	u	Celek
F0 mean	0,99959378	0,99993686	0,99989758	0,99999982	0,99990283	0,99984605
F0 min	0,99999902	0,99998961	0,99987486	0,99980343	0,99999956	0,99999982
F0 max	0,99999433	0,99968519	0,99991607	0,99995373	0,99999407	0,99995749
F0 median	0,99996305	0,99999297	0,99998161	0,99999869	0,99906497	0,99998892
F0 std	0,99944033	0,99999809	0,9996019	0,99555092	0,99663448	0,99967158
F0 max-min	0,99994209	0,99994363	0,99985123	0,99966934	0,99892555	0,99998764
F0 var	0,99993221	0,99973704	0,9998702	0,99904865	0,99860902	0,99993414
F0 sd	0,99981346	0,99999979	0,99999984	0,99235659	0,99376985	0,99999996
F1 mean	0,9999916	0,99998011	0,9999991	0,99999247	0,99997803	0,99999959
F1 min	0,99999995	0,99999593	0,99999698	0,99999212	0,99999473	0,99999994
F1 max	0,99999933	0,99999493	0,99999997	0,9999992	0,99999752	0,99999887
F1 median	0,99999951	0,99999968	0,99999095	0,99998107	0,99998489	0,99999998
F1 std	0,99990657	0,99995849	0,99999666	0,99997239	0,99999727	0,9999777
F1 max-min	0,99999774	0,99999922	0,99999988	0,99999982	0,999999	0,99999943
F2 mean	0,99999943	0,99999966	0,99999924	0,99999866	0,99999981	0,99999998
F2 max	0,99999929	0,99999432	0,99999948	0,9999998	0,99999989	0,99999999
F2 median	0,99999837	0,99999954	0,99999994	0,99999976	0,99999992	1
F2 std	0,99994209	0,99995659	0,99988909	0,99966934	0,99892555	0,99999417
F2 max-min	0,99999956	0,9999986	0,99999928	0,99999988	0,99999943	0,99999983
F1b mean	0,99991729	0,9999798	0,99999984	0,99996183	0,99995633	0,99998659
F1b min	0,99991282	0,99999997	0,99999999	0,99975768	0,99991237	0,99999994
F1b max	0,99999999	0,99999991	1	0,99999983	0,99999957	0,99999993
F1b median	0,99991277	0,99999761	0,99999941	0,99999576	0,99999999	0,99999671
F1b std	0,99999109	0,9999942	0,9999982	0,99998198	0,99994906	0,99999331
F1b max-min	0,99999999	0,9999999	1	0,99999985	0,99999957	0,99999994
F2b mean	0,99999706	0,99999768	0,99999915	0,99999609	0,99999586	0,99999971
F2b min	0,9999193	0,99998543	0,99999872	0,99992056	0,99999318	0,999978
F2b max	0,99999998	0,9999998	1	0,99999998	0,99999999	1
F2b median	0,9999989	0,99999853	0,99999984	0,99999854	0,99999301	0,99999993
F2b std	0,99999836	0,99999389	0,99999984	0,9999803	0,99999836	0,99999977
F2b max-min	0,99999998	0,9999998	1	0,99999998	1	1
jitter (local)	0,99989368	0,87998477	0,8574105	0,94275052	0,63073896	0,99935421
jitter (localabs)	1,2736E-07	0,99924871	0,99955048	1,4004E-06	1,57E-08	0,99378158
jitter (rap)	0,93189853	0,16625165	0,96102254	0,96243764	0,94600681	0,99794874
jitter (ppq5)	0,9726172	0,98597714	0,51045472	0,34827755	0,99625286	0,98479822
jitter (ddp)	0,9919456	0,95712333	0,99857332	0,99568224	0,99369831	0,99995508
shimmer (local)	0,99916084	0,96196868	0,96539991	0,99818777	0,92569698	0,9865917
shimmer (localabs)	0,91028528	0,56321457	0,53185885	0,88617252	0,23798664	0,57512301
shimmer (apq3)	0,9934582	0,99321628	0,82011827	0,98839013	0,75321927	0,94848844
shimmer (apq5)	0,99900273	0,93707142	0,96947876	0,99783136	0,90027169	0,97573876
shimmer (apq11)	0,99309851	0,94906779	0,99769312	0,99383058	0,9988165	0,9986961
shimmer (dda)	0,99926888	0,99924168	0,97620937	0,99869656	0,96487484	0,99400186
hnr hn	0,94367708	0,95213508	0,98028367	0,98019535	0,87856583	0,94833917
hnr nh	0,00743808	0,34817204	0,38650507	0,0031175	0,01663046	0,10874093
hnr ut	0,00395376	0,40072229	0,6800719	0,00196356	0,00643077	0,11620672

A.5 Výstupní data metody mRMR

Ženy				Muži			
Příznak	Relevance	Příznak	mRMR	Příznak	Relevance	Příznak	mRMR
F2b max	0.976	F2b max	0.976	F2b max-min	0.976	F2b max-min	0.976
F2 max	0.976	jitter (localabs)	-0.000	F2b min	0.976	jitter (localabs)	-0.000
F1 max	0.967	hnr nh	-0.000	F1b max	0.967	hnr nh	-0.000
F2 max-min	0.967	hnr ut	-0.000	F2 median	0.967	hnr ut	-0.000
F1b max-min	0.967	jitter (ppq5)	-0.056	F1b max-min	0.963	jitter (rap)	-0.166
F1b max	0.967	jitter (rap)	-0.070	F2 max-min	0.958	shimmer (localabs)	-0.208
F2b max-min	0.967	shimmer (localabs)	-0.151	F1 max	0.958	jitter (ppq5)	-0.198
F2b std	0.936	jitter (local)	-0.207	F2 max	0.930	jitter (local)	-0.338
F2 mean	0.918	F1 max	-0.304	F2 mean	0.927	F1 max-min	-0.407
F2b mean	0.905	jitter (ddp)	-0.483	F1 max-min	0.888	jitter (ddp)	-0.636
F2 median	0.900	shimmer (apq3)	-0.626	F1 mean	0.869	shimmer (apq3)	-0.660
F1 max-min	0.873	hnr hn	-0.639	F2b median	0.856	F0 std	-0.731
F1 median	0.861	shimmer (apq5)	-0.754	F2b mean	0.851	hnr hn	-0.781
F1 mean	0.847	shimmer (apq11)	-0.853	F2b std	0.838	shimmer (apq5)	-0.810
F1b std	0.844	F0 sd	-0.965	F1 median	0.816	F0 sd	-0.892
F2b median	0.777	shimmer (local)	-1.016	F1b std	0.771	shimmer (apq11)	-0.938
F1 min	0.751	F0 max	-1.160	F1b mean	0.754	shimmer (local)	-1.019
F1 std	0.748	shimmer (dda)	-1.282	F1 std	0.730	shimmer (dda)	-1.211
F1b mean	0.742	F0 std	-1.321	F1b median	0.727	F0 min	-1.448
F1b median	0.728	F0 min	-1.416	F1 min	0.688	F0 var	-1.563