



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

ANALÝZA VLIVU PŘÍTOMNOSTI ŠUMU V NAHRÁVKÁCH NA AUTOMATIZOVANOU DETEKCI HYPOKINETICKÉ DYSARTRIE

DOPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Nikola Havelková

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Daniel Kováč

BRNO 2022

Diplomová práce

magisterský navazující studijní program **Audio inženýrství** specializace

Zvuková produkce a nahrávání

Ústav telekomunikací

Studentka: Bc. Nikola Havelková

ID: 203727

Ročník: 2

Akademický rok: 2021/22

NÁZEV TÉMATU:

Analýza vlivu přítomnosti šumu v nahrávkách na automatizovanou detekci hypokinetické dysartrie

POKYNY PRO VYPRACOVÁNÍ:

Hypokinetická dysartrie je porucha motorické realizace řeči vyskytující se v oblasti respirace, fonace, artikulace a prozódie. Její automatizovaná detekce spočívá ve vhodné parametrizaci řečového signálu s následným užitím strojového učení. Řečové nahrávky však mohou být z různých důvodů znehodnocené šumem, což ovlivňuje jednotlivé parametry řeči a s tím i výslednou úspěšnost klasifikace. V rámci diplomové práce bude aplikován vhodný algoritmus pro zvýraznění řeči na nahrávky s aditivním šumem. Bude provedena parametrizace signálů a následné porovnání těchto parametrů extrahovaných ze signálu čistého, zašuměného a signálu zpracovaného metodou zvýraznění řeči. Také ve všech scénářích proběhne klasifikace nahrávek na řeč zdravou a řeč ovlivněnou dysartrií.

DOPORUČENÁ LITERATURA:

[1] GE, Fei. Brief Review of Recent Researches in Speech Enhancement from Filters to Neural Networks. In: IEEE. Stanford, CA, USA: IEEE, 2020, s. 260-264. ISBN 978-1-7281-7107-4. Dostupné z: doi:10.1109/CDS49703.2020.00059

[2] KOVÁČ, D.; MEKYSKA, J.; GALÁŽ, Z.; BRABENEC, L.; KOŠŤÁLOVÁ, M.; RAPCSAK, S; REKTOROVÁ, I. Multilingual Analysis of Speech and Voice Disorders in Patients with Parkinson's Disease. In 2021 44th International Conference on Telecommunications and Signal Processing. 2021. s. 273-277. ISBN: 978-1-6654-2933-7.

Termín zadání: 7.2.2022

Termín odevzdání: 24.5.2022

Vedoucí práce: Ing. Daniel Kováč

doc. Ing. Jiří Schimmel, Ph.D.

předseda rady studijního programu

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

Práce se zabývá analýzou vlivu přítomnosti šumu v nahrávkách na automatizovanou detekci hypokinetické dysartrie. Pro zvýraznění řeči jsou zvoleny a v prostředí MATLAB R2022a implementovány vhodné jednonábové metody, a to konkrétně spektrální odečítání a Kalmanův filtr. Metody jsou rovněž použity i u nahrávek bez zatížení šumem, ke kterým byl přidán aditivní bílý šum. Účinnost těchto metod je pak hodnocena objektivně pomocí hodnot odstuou signálu od šumu. Po zvýraznění řeči jsou z nahrávek extrahovány řečové příznaky. Vliv přítomnosti šumu, a i jeho následné potlačení jednotlivými metodami, je následně vyhodnoceno statistickou analýzou, konkrétně za pomoci Kruskal-Wallisova testu a post-hoc Dunnova testu. Distribuce pravděpodobnosti parametrů nahrávek čistých, zašuměných i zvýrazněných, u nichž je dle statistických testů vliv šumu signifikantní, jsou vykresleny za pomoci houslových a krabicových grafů. Na závěr byla provedena klasifikace nahrávek logistickou regresí za pomoci strojového učení, kdy byl dle hodnot plochy pod ROC křivkou popsán vliv přítomnosti šumu a následné zvýraznění řeči na automatizovanou detekci hypokinetické dysartrie.

Klíčová slova

Hypokinetická dysartrie, řečové příznaky, zvýraznění řeči, spektrální odečítání, Kalmanův filtr, odstup signálu od šumu, logistická regrese

Abstract

This thesis deals with the automated detection of hypokinetic dysarthria by analysing the influence of noise present in recordings. Appropriate single-channel methods, specifically the spectral subtraction and Kalman filter, are selected and implemented in the MATLAB R2022a to enhance speech. These methods are also used for noise-free recordings, to which additive white noise was added. Afterwards, the effectiveness of these methods is objectively evaluated by using signal-to-noise ratio values. After enhancing of speech, interferences are extracted from the recordings. The effect of the presence of noise, as well as its subsequent suppression by individual methods, is then evaluated by statistical analysis, specifically using the Kruskal-Wallis test and the post-hoc Dunn's test. The probability of distributing parameters of clean, noisy and enhanced recordings, for which the effect of noise is significant, according to statistical tests, are plotted using violin and box graphs. Finally, the classification was done by logistic regression with the help of machine learning, where the effect of the presence of noise and subsequent speech enhancement on automated detection of hypokinetic dysarthria was described according to the area values under the ROC curve.

Keywords

Hypokinetic dysarthria, speech features, speech enhancement, spectral subtraction, Kalman filter, signal-to-noise ratio, logistic regression

Bibliografická citace

HAVELKOVÁ, Nikola. *Analýza vlivu přítomnosti šumu v nahrávkách na automatizovanou detekci hypokinetické dysartrie* [online]. Brno, 2022 [cit. 2022-05-24]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/141420>. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce Daniel Kováč.

Prohlášení autora o původnosti díla

Jméno a příjmení studenta:	<i>Bc. Nikola Havelková</i>
VUT ID studenta:	<i>203727</i>
Typ práce:	<i>Diplomová práce</i>
Akademický rok:	<i>2021/22</i>
Téma závěrečné práce:	<i>Analýza vlivu přítomnosti šumu v nahrávkách na automatizovanou detekci hypokinetické dysartrie</i>

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné 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 závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si 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.

V Brně dne: 24. května 2022

podpis autora

Poděkování

Děkuji vedoucímu diplomové práce Ing. Danielu Kováčovi za účinnou metodickou, pedagogickou a odbornou pomoc, za ochotu, vstřícnost a cenné rady při zpracování mé diplomové práce.

V Brně dne: 24. května 2022

podpis autora

Obsah

SEZNAM OBRÁZKŮ	9
SEZNAM TABULEK	11
1. ÚVOD	12
2. HYPOKINETICKÁ DYSARTRIE	13
2.1 PROJEVY HYPOKINETICKÉ DYSARTRIE.....	13
2.1.1 <i>Poruchy fonorespirace</i>	14
2.1.2 <i>Poruchy faciokineze</i>	14
2.1.3 <i>Poruchy fonetiky</i>	14
2.2 PARAMETRIZACE ŘEČI.....	15
2.2.1 <i>Fonace</i>	16
2.2.2 <i>Artikulace</i>	19
2.2.3 <i>Prozódie</i>	20
3. METODY ZVÝRAZNĚNÍ ŘEČI	22
3.1 SPEKTRÁLNÍ ODEČÍTÁNÍ.....	22
3.2 KALMANŮV FILTR.....	23
4. IMPLEMENTACE METOD ZVÝRAZNĚNÍ ŘEČI	25
4.1 DATABÁZE.....	25
4.2 HLAVNÍ SKRIPT.....	25
4.2.1 <i>Výpočet odstupů signálu od šumu (SNR)</i>	26
4.3 SPEKTRÁLNÍ ODEČÍTÁNÍ.....	29
4.4 KALMANŮV FILTR.....	30
5. PARAMETRIZACE	31
6. STATISTICKÁ ANALÝZA	33
7. KLASIFIKACE	34
8. VÝSLEDKY	35
8.1 VÝSLEDKY IMPLEMENTACÍ METOD PRO ZVÝRAZNĚNÍ ŘEČI.....	35
8.2 VÝSLEDKY STATISTICKÉ ANALÝZY.....	42
8.2.1 <i>Kruskal-Wallisův test</i>	42
8.2.2 <i>Dunnův test</i>	43
8.3 VÝSLEDKY KLASIFIKACE.....	48
9. DISKUZE	54
9.1 METODY ZVÝRAZNĚNÍ ŘEČI.....	54
9.2 STATISTICKÁ ANALÝZA.....	54
9.3 KLASIFIKACE.....	55
10. ZÁVĚR	58
11. LITERATURA	60
SEZNAM SYMBOLŮ A ZKRATEK	63

SEZNAM PŘÍLOH.....	64
PŘÍLOHA A: ČASOVÝ PRŮBĚH A SPEKTROGRAM ÚLOHY ČTENÍ TEXTU (HU)	65
PŘÍLOHA B: ČASOVÝ PRŮBĚH A SPEKTROGRAM ÚLOHY PRODLOUŽENÉ FONACE (HU)	67
PŘÍLOHA C: HODNOTY SNR (CZ)	69
PŘÍLOHA D: HODNOTY SNR (HU).....	73
PŘÍLOHA E: STRUKTURA KÓDU	74

SEZNAM OBRÁZKŮ

Obr. 2.1: Diagram odhadu šumového a harmonického spektra pro HNR [7]	17
Obr. 2.2: Signál zdravé řeči [8]	17
Obr. 2.3: Signál řeči lehce dysfonické [8].....	18
Obr. 2.4: Signál řeči silně dysfonické [8].....	18
Obr. 2.5: a) hlasivky zdravého řečnicka b) běžná míra addukčního napětí c) addukční napětí dyšné fonace [10].....	19
Obr. 4.1: Porovnání časového zpoždění signálů.....	27
Obr. 4.2: Porovnání kompenzace časového zpoždění signálů	28
Obr. 4.3: Diagram implementace spektrálního odečítání.....	29
Obr. 4.4: Diagram implementace Kalmanova filtru	30
Obr. 8.1: Řečový signál originální české nahrávky čtení textu	35
Obr. 8.2: Řečový signál české nahrávky čtení textu zatížené šumem	36
Obr. 8.3: Řečový signál české nahrávky čtení textu zvýrazněné spektrálním odečítáním	36
Obr. 8.4: Řečový signál české nahrávky čtení textu zvýraznění Kalmanovým filtrem.....	37
Obr. 8.5: Řečový signál originální české nahrávky prodloužené fonace	37
Obr. 8.6: Řečový signál české nahrávky prodloužené fonace zatížené šumem	38
Obr. 8.7: Řečový signál české nahrávky prodloužené fonace zvýrazněné spektrálním odečítáním	38
Obr. 8.8: Řečový signál české nahrávky prodloužené fonace zvýraznění Kalmanovým filtrem.....	39
Obr. 8.9: Hodnoty SNR nahrávek čtení textu – potlačení aditivního šumu s odstupem 10 dB od originálního signálu	39
Obr. 8.10: Hodnoty SNR nahrávek prodloužené fonace – potlačení aditivního šumu s odstupem 10 dB od originálního signálu	40
Obr. 8.11: Hodnoty SNR nahrávek čtení textu – potlačení aditivního šumu s odstupem 20 dB od originálního signálu	40
Obr. 8.12: Hodnoty SNR nahrávek prodloužené fonace – potlačení aditivního šumu s odstupem 20 dB od originálního signálu	40
Obr. 8.13: Hodnoty SNR nahrávek čtení textu – potlačení aditivního šumu s odstupem 30 dB od originálního signálu	41
Obr. 8.14: Hodnoty SNR nahrávek prodloužené fonace – potlačení aditivního šumu s odstupem 30 dB od originálního signálu	41
Obr. 8.15: Porovnání skupin pro RFA1 úlohy TSK2	44
Obr. 8.16: Porovnání skupin pro RFA2 úlohy TSK2	44
Obr. 8.17: Porovnání skupin pro relF1SD úlohy TSK2.....	45
Obr. 8.18: Porovnání skupin pro relF2SD úlohy TSK2.....	45
Obr. 8.19: Porovnání skupin pro Indmark úlohy TSK2	45
Obr. 8.20: Porovnání skupin pro NST úlohy TSK2	46
Obr. 8.21: Porovnání skupin pro TPT50 úlohy TSK2	46
Obr. 8.22: Porovnání skupin pro PPR úlohy TSK2.....	46
Obr. 8.23: Porovnání skupin pro GNE úlohy TSK6.....	47
Obr. 8.24: Porovnání skupin pro HRF úlohy TSK6	47
Obr. 8.25: Porovnání skupin pro QOQ úlohy TSK6	47
Obr. 8.26: Porovnání skupin pro NAQ úlohy TSK6	48
Obr. 8.27: Porovnání skupin pro relF2SD úlohy TSK6.....	48
Obr. 8.28: ROC křivka klasifikace čistých nahrávek v českém jazyce	50
Obr. 8.29: ROC křivka klasifikace nahrávek v českém jazyce zatížených aditivním šumem	50

Obr. 8.30: ROC křivka klasifikace nahrávek v českém jazyce zvýrazněných spektrálním odečítáním.....	51
Obr. 8.31: ROC křivka klasifikace nahrávek v českém jazyce zvýrazněných Kalmanovým filtrem.....	51
Obr. 8.32: ROC křivka klasifikace nahrávek v maďarském jazyce zatížených šumem	52
Obr. 8.33: ROC křivka klasifikace nahrávek v maďarském jazyce zvýrazněných spektrálním odečítáním	52
Obr. 8.34: ROC křivka klasifikace nahrávek v maďarském jazyce zvýrazněných Kalmanovým filtrem..	53

SEZNAM TABULEK

Tab. 4.1: Struktura databáze nahrávek.....	25
Tab. 5.1: Souhrn užitých parametrů.....	31
Tab. 8.1: P-hodnoty Kruskal-Wallisova testu pro jednotlivé parametry	42
Tab. 8.2: Q-hodnoty Dunnova testu pro jednotlivé parametry	43
Tab. 8.3: AUC hodnoty klasifikace logistickou regresí.....	49

1. ÚVOD

Tato práce se zabývá problematikou výskytu šumu v nahrávkách při stanovení parametrů pro následnou automatizovanou detekci hypokinetické dysartrie.

Začátkem práce je v prve řadě popsána samotná hypokinetická dysartrie, tedy jak vzniká a jakým způsobem se projevuje. Tyto projevy jsou zde rozebrány ze tří základních pohledů – fonorespirace, faciokineze a fonetiky. Druhým úsekem této teoretické části je popsání a rozebrání definic řečových příznaků, jež bylo následně využito pro klasifikaci nahrávek.

Další část práce se již zabývá stanovením dvou metod pro zvýraznění řeči, respektive pro potlačení nežádoucího šumu vyskytujícího se v nahrávkách, který je zde buďto přirozeně, jako je tomu u maďarských nahrávek, nebo se jedná o úmyslné přičtení aditivního bílé šumu, což je případ nahrávek česky mluvících řečníků.

Metody byly voleny tak, aby zde byla zastoupena jak metoda implementačně a výpočetně nenáročná, tedy metoda spektrálního odečítání, tak i metoda pokročilejší, kterou je Kalmanův filtr. Po teoretickém rozebrání těchto možností úpravy řečového signálu následuje samotná implementace v prostředí MATLAB R2022a. Po zvýraznění řeči v nahrávkách byla provedena parametrizace, jejíž výstupní hodnoty řečových příznaků byly podrobeny statistické analýze pomocí Kruskal-Wallisova a post-hoc Dunnova testu pro zjištění, zdali je vliv šumu v nahrávkách čistých, zašuměných či zvýrazněných signifikantní. Distribuce pravděpodobnosti u takovýchto parametrů byla graficky vyjádřena houslovým a krabicovým grafem.

Na závěr byla provedena klasifikace nahrávek logistickou regresí za pomoci strojového učení, jejímiž výstupními hodnotami plochy pod ROC křivkou byl zhodnocen vliv přítomnosti šumu v nahrávkách pro tuto automatizovanou detekci hypokinetické dysartrie.

Cílem této práce tak bylo nejen zhodnotit, zdali přítomnost šumu v nahrávkách může ovlivňovat řečové příznaky a tím i automatizovanou detekci, jako je tomu například ve studii [1], kde je dokonce využito více typů aditivních šumů. Součástí této práce je ovšem i analýza vlivu následného zvýraznění řeči zatížené šumem, a to prostřednictvím výše uvedených metod, tedy pomocí spektrálního odečítání a Kalmanova filtru.

2. HYPOKINETICKÁ DYSARTRIE

Hypokinetická dysartrie (HD) je porucha motorické realizace řeči, jejíž příčina sahá do oblasti mozku, která je centrem motoriky, kognitivních funkcí a funkcí limbického systému. Tato část centrální nervové soustavy se nazývá bazální ganglia, která hrají významnou roli při kontrole komplexnějších motorických procesů, mezi které se řadí například i řeč.

Vzniklé léze v oblasti bazálních ganglií, konkrétně tedy podél párové struktury středního mozku zvané černá substance (*substantia nigra*), negativně ovlivňují funkci dopaminergních neuronů tvořících jádro (*pars compacta*) této substance, přičemž právě dopamin jakožto neurotransmitter zajišťuje rovnováhu mezi okruhy zapojení bazálních ganglií, tedy dráhou přímou a nepřímou, a to prostřednictvím receptorů D1 (pro zvýšení aktivity přímé dráhy) a D2 (pro snížení aktivity nepřímé dráhy).

Účelem přímé dráhy je aktivace žádoucích pohybů, kdy se tímto procesem sníží inhibice *thalamu*, což mimo jiné excitačně ovlivňuje segmentální motorickou oblast mozkové kůry.

Funkcí nepřímé dráhy je redukce nechtěné motorické aktivity. Během tohoto procesu dochází ke snížení inhibice *subthalamu*, a naopak ke zvýšení inhibice *thalamu*, následkem čehož je snížen vliv excitačních neuronů na *cortex cerebri*, tedy kůru mozkovou.

Snížení aktivity přímé dráhy a nedostatečný útlumem cesty nepřímé tak patří mezi jednu z hlavních příčin Parkinsonovy nemoci (PN) projevující se tímto jako hypokinetická dysartrie (HD). [2]

2.1 Projevy hypokinetické dysartrie

Mezi patologické projevy, které se u HD mohou vyskytnout, patří poruchy fonorespirace, faciokineze či poruchy fonetiky. V klinické praxi jsou tyto oblasti posuzovány v tzv. 3F testu [3], jehož cílem je stanovení dysartrického profilu. Pro posouzení dysartrie jsou využívány i mnohé další testy, přičemž mezi ty významnější, především tedy v zahraničí, lze zařadit například test FDA-2 (*Frenchay Dysarthria Assessment*).

2.1.1 Poruchy fonorespirace

Oblast fonorespirace se zabývá nejen funkčností samotného hlasivkového ústrojí, ale také respirací, která je nedílnou součástí tvorby fonace, která bývá zpravidla první oblastí, kde je možné HD z hlediska tvorby hlasu detekovat.

Vlivem rigidity fonačního ústrojí dochází k neúplnému uzavírání a asymetrickému chvění hlasivek. Za pomoci elektromyografu byla rovněž u pacientů s HD zjištěna zvýšená aktivita hrtanového svalstva, a také jejich nedostatečná koordinace se svaly respiračního ústrojí. Pacienti tak nedokážou efektivně pracovat s dechem, proud vzduchu kolísá, což se projevuje jako hlasový třes. Rovněž nejsou schopni tyto svalové oblasti udržet ve stabilní pozici po delší časový úsek. U prodloužených vokálů tak dochází k rozptylu základního tónu. V případě náhlého uvolnění proudu dechu dutinou nosní se v modulovém frekvenčním spektru projeví hypernazalita zdůrazněním některých formantových frekvencí [3] [4] [5].

2.1.2 Poruchy faciokineze

Z hlediska faciokineze je zkoumáno narušení pohyblivosti svalových skupin zajišťujících motoriku mluvidel. Mezi tato svalstva spadá například kruhový sval ústní, sval bradový, trubačský atd. Omezená pohyblivost těchto svalů se klasifikuje jako hypokineze, bradykineze, akineze či rigidita. Poruchy inervace orofaciálního svalstva tak mají za následek nejen omezení mimiky (hypomimie), tedy neverbální komunikace, ale především také špatnou artikulaci [3] [4] [5].

2.1.3 Poruchy fonetiky

V této oblasti se jedná o poruchy akustických projevů při tvorbě řeči, spadá sem tedy narušení artikulace, prozódie či plynulosti řeči.

- **Artikulace**

Narušení artikulace je nejvíce zřejmé při vyslovování souhlásek (neboli konsonant), především pak u souhlásek plozivních, které vznikají jako exploze či imploze při uvolnění tlaku vzniklého v dutinách řečového ústrojí. Objem těchto dutin, ve kterých dochází k rezonanci při tvorbě vokálů, bývá spojován se správnou artikulací. Tento parametr je možné měřit za pomoci formantových oblastí.

Špatná výslovnost je obecně dána nedostatečnou schopností řízení a koordinace aktivních artikulačních orgánů, mezi které se řadí rty, jazyk, hrtanová přiklopka a další. Výsledkem je tak nesrozumitelná řeč, jenž působí nedbale až nepřírozně.

Pro hodnocení míry narušení artikulace se využívá diadochokinetického testu, který spočívá ve vyslovování řady slabik, u kterých je za potřeby rychle měnit postavení artikulačních orgánů. Pacienti se mohou snažit o udržení kadence, ovšem může se tak snížit pohyblivost mluvidel, což má za následek zvýšenou nesrozumitelnost [3] [4] [5].

- **Prozódie**

Prozódie je označení pro souhrn zvukových vlastností řeči, které dodávají sdělovanému obsahu význam a zdůraznění. Mezi hlavní prozodické znaky se řadí kmitočet základního tónu, intenzita a tempo řeči. Jako další znaky lze také označit rytmus, intonaci, přízvuk atd.

Narušení těchto vlastností, tzv. dysprozódie, tak pro pacienty znamená špatnou schopnost intonace, čímž může projev působit monotónně, dále dochází k náhlým změnám v tempu řeči (*tachyfemie/bradyfemie*) či ke ztrátě přirozené melodie věty [3] [4] [5].

- **Plynulost řeči**

Při HD může být také narušená přirozená fluence řeči. Pacient může neúmyslně tvořit pauzy ve větě či ve slovech, tzv. hezitace, nebo také může zrychleně opakovat slabiky i celá slova (tzv. palilalie) [3] [4] [5].

2.2 Parametrizace řeči

Řečový signál obsahuje vysoké množství redundantních dat. Je tak zaváděna parametrizace, již jsou extrahovány informace přínosné pro danou aplikaci. Výsledkem tak jsou tzv. řečové příznaky, které se člení na příznaky segmentální a suprasegmentální, dle toho, zdali jsou extrahovány pouze z krátkého úseku řeči, tedy segmentu, či jsou získány z úseku delšího, kupříkladu celé řečové nahrávky. Parametry je možné členit rovněž dle příslušných oblastí fonetiky na příznaky fonační, artikulační a prozodické.

2.2.1 Fonace

- **Fundamentální frekvence**

Frekvence základního tónu F_0 je jedním z hlavních suprasegmentálních rysů, kdy se při diagnóze využívá především jeho odchylky a střední hodnoty. Míru variability fundamentální frekvence lze získat za pomoci směrodatné odchylky a tzv. *jitteru* neboli kolísání frekvence F_0 . Kolísání amplitudy je naopak definováno parametrem *shimmer*.

Pro tuto analýzu je *jitter* a *shimmer* využit v podobě percentuelního vyjádření jakožto kvocienty PPQ (Period perturbation quotient) a APQ (Amplitude perturbation quotient).

Normalizací F_0 lze získat výpočty pro další parametry, jako například relativní směrodatnou odchylku $relF_0SD$ popisující nepravidelné kolísání hlasu či monotónnost řeči z hlediska výšky hlasu [4] [5]

$$relF_0SD = \frac{std(F_0)}{mean(F_0)} \cdot 100 \text{ [%]}. \quad (1)$$

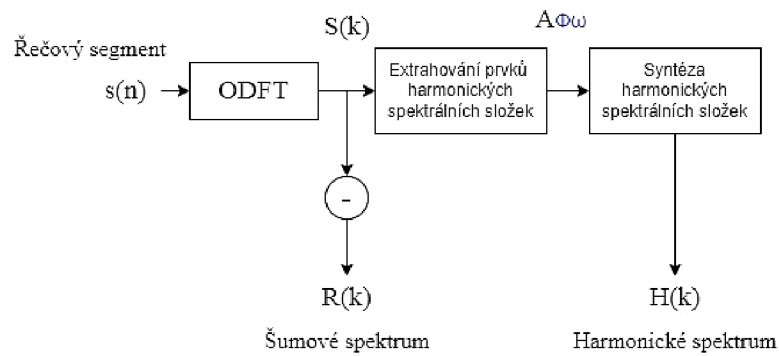
Obdobně jako se počítá hodnota parametru $relF_0SD$, lze vypočítat relativní kolísání pro první a druhý formant, tedy parametry $relF1SD$ a $relF2SD$. Tyto formanty jsou využity z toho důvodu, že souvisí s rezonancí budícího signálu v oblasti oronazofaryngeálního traktu, které mohou být modifikovány polohou jazyku a čelisti, díky čemuž je možné stanovit např. míru chvění či naopak ztuhlost čelisti a jazyku, za předpokladu, že jsou tyto parametry extrahovány z úloh prodloužené fonace [6].

- **HNR (Harmonic to noise ratio)**

Parametr HNR je logaritmické vyjádření poměru mezi energií periodické (harmonické) a aperiodické (tedy šumové) složky. Při HD totiž může docházet k nedovírání hlasivek, což zapříčiňuje vyšší množství neznělých složek v řeči.

$$HNR = 10 \log \frac{\sum_{k=1}^{N/2} |H(k)|^2}{\sum_{k=1}^{N/2} |N(k)|^2} \quad (2)$$

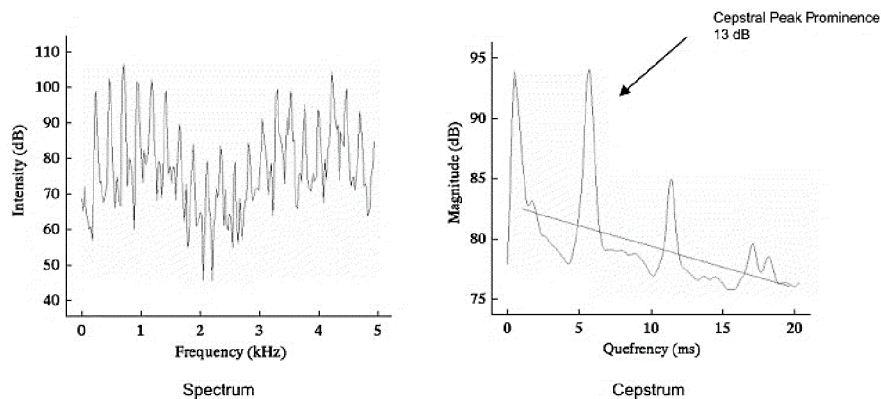
Harmonickou $H(k)$ a šumovou složku $N(k)$ lze z řeči získat autokorelační funkcí, viz diagram obr. 2.1 [7].



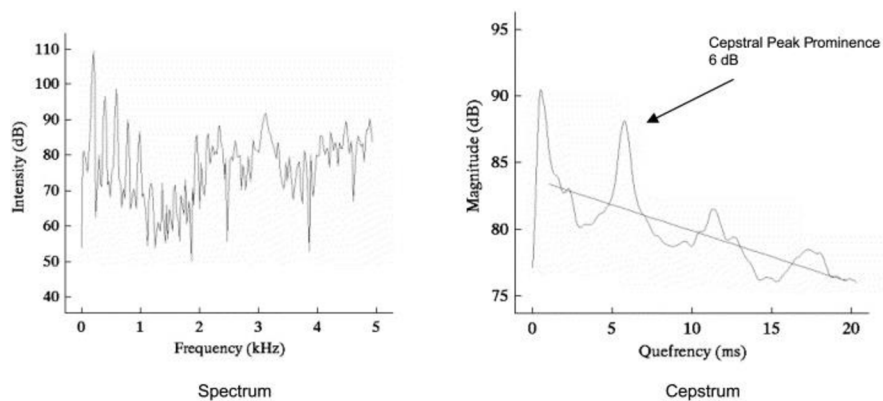
Obr. 2.1: Diagram odhadu šumového a harmonického spektra pro HNR [7]

- **CPP (Cepstral peak prominence)**

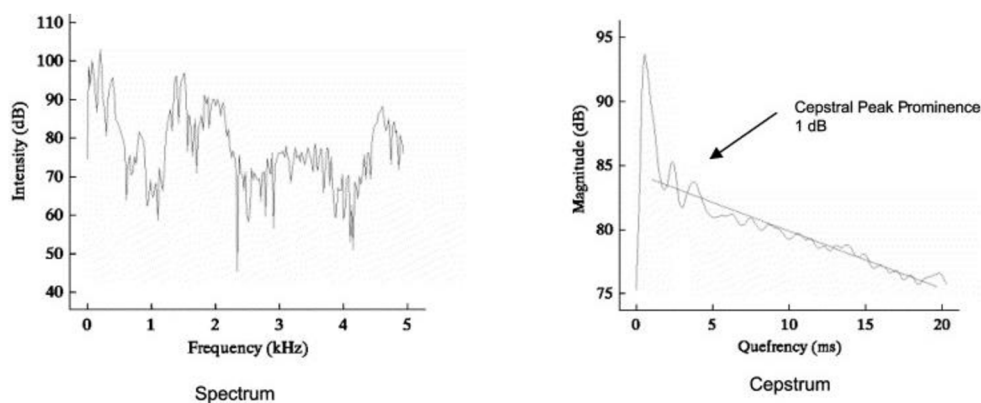
CPP vyjadřuje míru s jakou jsou špičkové hodnoty v keprstru výrazné. CPP je tedy definováno jako rozdíl mezi špičkovou hodnotou základního tónu a přímkou lineární regrese. U HD lze očekávat nižší hodnoty tohoto parametru, jelikož vlivem nedovírání hlasivek dochází ke snížení míry periodického buzení, což se v tomto případě projeví jako méně výrazná špička rezonančního maxima. Tento jev lze vidět na obr. 2.2 až 2.4, kde je vyobrazeno spektrum, a především pak keprstrum nahrávek zdravé, lehce dysfonické a silně dysfonické řeči [8].



Obr. 2.2: Signál zdravé řeči [8]



Obr. 2.3: Signál řeči lehce dysfonické [8]



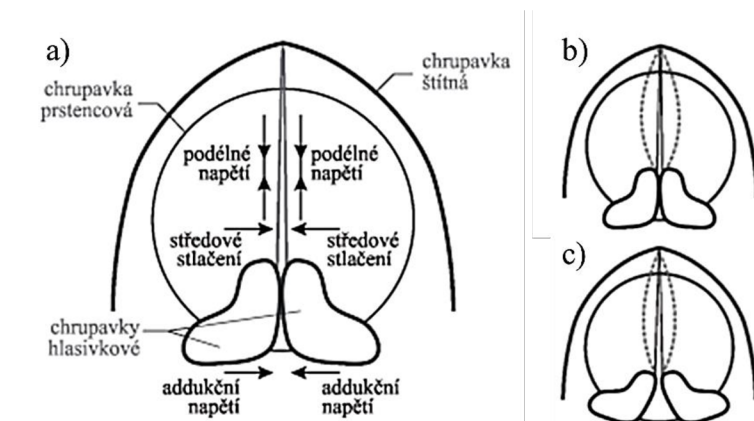
Obr. 2.4: Signál řeči silně dysfonické [8]

- **HRF (Harmonic richness factor)**

Tento parametr popisuje míru znělosti produkovaného signálu, a to jako poměr mezi energií první harmonické složky, tedy základním tónem a celkovou energií vyšších harmonických složek [9].

- **NAQ (Normalized amplitude quotient)**

NAQ je kvocient, který na základě fáze uzavírání hlasivek popisuje míru intenzity hlasu, a také jeho fonační typ, jako například fonace dyšná, jejíž příčinou je nižší addukční napětí. Hlasivkové chrupavky se tak nemohou plně sevřít, čímž dochází k nedokonalému přerušení vzduchového proudu, kdy takto vzniklou šterbinou část vzduchu prochází [10].



Obr. 2.5: a) hlasivky zdravého řečníka b) běžná míra addukčního napětí c) addukční napětí dyšné fonace [10]

Matematicky lze k hodnotě tohoto kvocientu dojít podílem amplitudy glotálního pulzu A vůči derivaci záporné amplitudy pulzu d_{peak} získané inverzní filtrací. Tento vztah je pak normalizován celkovou dobou trvání hlasivkového pulzu, tedy $NAQ = \frac{A}{d_{peak} \cdot T}$ [11].

Odpovídající ztuhlost hlasivek je pak možné vyjádřit pomocí směrodatné odchylky značené jako $relNAQSD$ [11].

- **QQQ (Quasi-open quotient)**

QQQ udává energii harmonických složek a celkovou intenzitu hlasu vztahem mezi dobou otevírací fáze hlasivkového pulzu a celkovou periodou pulzu. Obdobně jako u kvocientu NAQ je možné výpočtem směrodatné odchylky tohoto parametru popsat ztuhlost hlasivek [9].

- **DUV (Degree of voiceless)**

Parametr DUV procentuálně udává stupeň neznělých úseků v řeči, kdy se za patologický jev považuje jakákoliv hodnota vyšší nežli 0 %, kdy příčinou zvýšení této hodnoty může být např. už i pouhé nachlazení [12].

2.2.2 Artikulace

Při tvorbě slova prochází turbulentní proud vzduchu respiračním a artikulačním ústrojím, čímž dochází ke vzniku rezonancí na určitých frekvencích, které jsou nazývány jako kmitočty formantové. Nejvíce vypovídajícími formanty skrze produkci vokálů jsou

formanty F_1 a F_2 . Spolu s jejich charakteristikami je zle využít pro popis míry postižení u pacientů s PN.

Vzhledem k tomu, že změna polohy mluvidel má za následek změnu objemu dutin, kde k rezonancím dochází, mění se s tím i hodnoty formantů. Na základě toho je tak možné sledovat hybnost těchto mluvidel, např. jazyku, s čímž je spjata i výsledná artikulace [4] [5] [13].

- **RFA (Resonant frequency attenuation)**

Parametr RFA určuje na základě útlumu rezonanční frekvence poruchu artikulace.

Hodnota RFA1 odpovídá ve spektru lineární predikční analýzy vzdálenosti mezi frekvencí druhého formantu a jemu předcházejícímu antiformantu (první lokální minimum) [14].

V případě parametru RFA2 odpovídá hodnota vzdálenosti opět od frekvence druhého formantu, ovšem tentokrát k hodnotě antiformantu nadcházejícímu (druhé lokální minimum) [14].

Celkový počet lokálních maxim je definován jako parametr *#loc_max*, který se rovněž užívá k analýze poruch artikulace [14].

Při práci s parametry jako je například RFA1 a RFA2 je třeba si uvědomit, že formanty jsou silně závislé na věku a pohlaví jedince. Dle studie [15] bylo ale zjištěno, že samohlásky *i* a *u* jsou nejméně závislé na těchto faktorech [4] [5] [13].

- **Lndmrk (Landmark)**

Parametr *lndmrk* poukazuje především na náhlé změny v řečovém signálu za využití hranic akustických příznaků, přičemž je množství těchto oblastí vztaženo k celkovému času signálu.

2.2.3 Prozódie

- **Tvorba pauz**

Pro sledování přerušování řeči a prodlužování pauz lze využít několika parametrů, mezi které patří:

PPR, tedy procentuální poměr pauz, který lze vypočítat jako poměr mezi celkovým trváním pauz (*TPT*) a absolutním trváním řečového signálu (*TST*).

SPIR je tzv. index rytmičnosti, jehož hodnota odpovídá celkové délce mezislovních pauz, jež jsou delší než 50 ms, vůči celé délce trvání řeči [4] [5] [16].

S tvorbou pauz, především pak s délkou trvání těchto tichých úseků řeči jsou spjaty následující parametry popisující prozódie, kterých bylo při analýze využito.

DurMED (Median duration) je průměrná doba trvání tichých úseků, jejichž délka je přesahující 50 ms [16].

DurMAD (Median absolute deviation) popisuje proměnlivost délky trvání tichých úseků v řeči, a to jako střední absolutní odchylka tichých úseků, které jsou opět delší než 50 ms [16].

NST (Net speech time) odpovídá relativní době trvání řeči vztažené k celkové době trvání [16].

- **Intenzita řeči**

Pacienti s PN mívají spíše nižší úroveň intenzity řeči, kterou sice dokážou úmyslně zvýšit, ale pouze v krátkém časovém úseku, po kterém se jejich hlasitost vrací zpět do jim přirozené úrovně [4] [5].

Proměnlivost hlasitosti je analyzována za pomoci parametru *relSE0SD*, který je definován jako směrodatná odchylka intenzity buzeného signálu v podílu ku střední hodnotě intenzity buzeného signálu bez tichých úseků, jež přesahovaly délku trvání 50 ms [16].

EEVOL (Energy evolution) definuje strmostí intenzity, jak se vyvíjí energie řeči a zdali dochází ke kolísání intenzity.

- **relF0SD**

Řečový parametr *relF0SD* již byl popsán jakožto příznak fonační, a to za předpokladu jeli extrahován z řečové úlohy prodloužené fonace. Pakliže je získán z úlohy čtení textu, spadá již mezi příznaky prozodické, kdy je jím kvantifikována monotónnost výšky hlasu [16].

3. METODY ZVÝRAZNĚNÍ ŘEČI

Pro tuto práci byly zvoleny dvě různé metody odstraňující nežádoucí šum z nahrávek určených k detekci HD. Metody byly zvoleny tak, aby zde byly zastoupeny různé úrovně obtížnosti implementace i výpočetní náročnosti. První metoda, spektrální odečítání, tedy reprezentuje způsob implementačně a výpočetně přívětivý a druhou metodou je Kalmanův filtr, který byl zvolen jako metoda pokročilejší s předpokladem dosažení lepších výsledků nežli pomocí metody první.

3.1 Spektrální odečítání

Metoda spektrálního odečítání je jednou z prvních metod pro zvýraznění řečového signálu. Patří tak k samotnému základu této problematiky a za dlouhou dobu od jejího vzniku bylo vytvořeno mnoho modifikací kompenzujících nedostatky základní verze. Ačkoliv je tato metoda poněkud triviální, má mnohé výhody, mezi než patří snadná implementace, nízká výpočetní náročnost, a i přes to poměrně vysoká účinnost. Nevýhodou této metody je však fakt, že její princip spočívá i v určité intuici a heuristickém přístupu. Druhou nevýhodou je vznik hudebního šumu, který se při odečítání modulů generuje jako náhodné frekvence harmonických složek [17] [18].

Podstata této metody tkví v tom, že vstupním signálem je signál řečový $x(n)$, který je znehodnocený šumem $d(n)$. Tento vstupní signál lze vyjádřit za pomoci čistého signálu $s(n)$ a šumu následovně

$$x(n) = s(n) + d(n). \quad (3)$$

Aby tedy bylo možné na výstupu získat řeč čistou, je touto metodou šum $d(n)$ odečítán, respektive jeho přesné spektrum není možné určit, proto dochází k odečítání pouze jeho odhadu získaného z oblastí, kde se nevyskytuje řeč. Odhad šumu je možné získat za předpokladu, že je tento signál považován za stacionární, čehož lze dosáhnout segmentací, kdy se v krátkých úsecích signál, v tomto případě tedy šum, nemění, či dochází pouze k pomalým změnám [19] [20] [21].

Je nezbytné mít na paměti, že vzhledem k tomu, že od vstupního signálu je odečítán pouhý odhad spektra nežádoucího signálu, je i výsledná spektrální charakteristika výstupního signálu opět jen odhadem [19] [20] [21].

Po aplikaci Fourierovy transformace na rovnici 4 a vyjádření proměnných v polárním tvaru je spektrum odhadu čistého řečového signálu popsáno následovně:

$$\hat{S}(\omega) = [|X(\omega)| - |\hat{D}(\omega)|]e^{j\Phi_x(\omega)}. \quad (4)$$

Při využívání této metody je nutné sledovat míru odečítání redundantní informace, jinak by mohlo dojít k odečtení, a tedy ke ztrátě části užitečného řečového signálu. Naopak nebude-li šum odečítán dostatečně, bude v řečovém signálu setrvávat i po zpracování řeči zvolenou metodou [17] [19] [20] [21].

3.2 Kalmanův filtr

Kalmanův filtr byl navržen matematikem Robertem Emilem Kalmanem v 60. letech 20. století. Té doby byl ovšem vytvořen se záměrem jej využít především v oblastech navigačních systémů, robotiky apod. První zmínkou o využití této metody pro zvýraznění řeči byla roku 1987 práce K. Paliwala a A. Basu [22], která je stěžejním dokumentem pro mnohé další výzkumy či práce zabývající se touto problematikou.

Tato metoda je matematickým postupem, jehož princip stojí na predikčních a korekčních mechanismů. Je zde využíváno jak známých vlastností daného systému, tak i změřených či odhadovaných požadovaných parametrů.

Jednou z výhod této metody je to, že počítá i s nestacionárními signály, jako je například tedy signál řečový, který je zde popsán v rámci segmentu pomocí autoregresivního modelu. Filtr je nejvhodnější využít pro odstranění šumu bílého charakteru, jelikož tento algoritmus předpokládá, že šum narušující řečový signál je nekorelovaný a s běžnou distribucí [17] [23].

Čistý řečový signál $s(k)$ lze pro aktuální vzorek popsat rovnicí

$$\begin{aligned} s(k) &= a_1s(k-1) + \dots + a_ps(k-p) + \mathbf{q}(k) \\ &= \sum a_i s(k-i) + \mathbf{q}(k) \quad i = 1 \dots p. \end{aligned} \quad (5)$$

Parametr p představuje hodnotu řádu autoregresivního filtru, a_i odpovídá lineárně predikčním koeficientům filtru a vektor $\mathbf{q}(k)$ je chybou predikce, pro kterou platí vztah $\mathbf{q}(k) = \mathbf{G}\mathbf{u}(k)$, kde \mathbf{G} značí vektor zesílení a $\mathbf{u}(k)$ je vektorem budícího aditivního šumu [17] [18] [22] [23].

Výše uvedenou rovnici 5 je možné vyjádřit také stavovou rovnicí systému:

$$\mathbf{Z}(k) = \mathbf{A}\mathbf{Z}(k-1) + \mathbf{G}\mathbf{u}(k), \quad (6)$$

jejíž obecné řešení sestává z homogenní a partikulární části:

$$z(t) = e^{A(t-t_0)}z(t_0) + \int_{t_0}^t e^{A(t-\tau)}Gu(\tau)d\tau. \quad (7)$$

Homogenní část $e^{A(t-t_0)}z(t_0)$ obecného tvaru je stavovou odezvou systému na počáteční podmínky při nulovém vstupním signálu $u(t)$ a část partikulární, tedy $\int_{t_0}^t e^{A(t-\tau)}Gu(\tau)d\tau$, je stavovou odezvou na vstup $u(t)$ při nulových počátečních podmínkách [17] [18] [22] [23].

Matrice e^{At} , která je v rovnici 6 označena jako \mathbf{A} , je stavovou maticí přechodu mezi stavy $z(k)$ a $z(k-1)$ a \mathbf{G} odpovídající vektoru zesílení nabývající délky řádu filtru lze získat z hodnoty výkonu chyby predikce α a Tyto prvky jsou definovány následovně:

$$\mathbf{A} = \begin{bmatrix} 0 & \dots & 1 \\ -a_p & \dots & -a_1 \end{bmatrix} \quad (8)$$

$$\mathbf{G} = \sqrt{\alpha} = \sqrt{R[0] + \sum_{k=1}^p a_k R[k]}. \quad (9)$$

Za předpokladu, že se znehodnocený signál $x(k)$ rovná vztahu $x(k) = s(k) + d(k)$, je možné jej v maticovém tvaru vyjádřit jako

$$x(k) = \mathbf{H}X(k) + d(k), \quad (10)$$

kde \mathbf{H} značí observační matici, pro kterou platí výpočet $\mathbf{H} = \mathbf{G}T$, kde T odpovídá délce signálu [17] [18] [22] [23].

4. IMPLEMENTACE METOD ZVÝRAZNĚNÍ ŘEČI

4.1 Databáze

Pro implementaci jednotlivých jednokanálových metod zvýraznění řeči, které byly po teoretické stránce rozebrány v kapitole 3, byla využita databáze nahrávek 214 českých a 23 maďarských řečníků. Struktura databáze z hlediska zastoupení zdravé kontroly (ZK), Parkinsonovy nemoci (PN), pohlaví či věku je uvedeno pro oba jazyky v tab. 4.1.

Tab. 4.1: Struktura databáze nahrávek

Jazyk	Typ nahrávky		Pohlaví		Věk			
	ZK	PN	Muž	Žena	Min	Max	Průměr	Sm. odchylka
CZ	81	133	119	95	45	87	67	8,23
HU	12	11	14	9	45	81	65	7,08

Pro rozřazení těchto nahrávek byla napsána funkce *Dataprep.m*, jejíž cílem je prohledat jednotlivé složky řečníků v databázi, z níž jsou vybrány pouze nahrávky úloh čtení textu (dále značeno jako TSK2) a prodloužené fonace (TSK6). Data takto vyhledaných nahrávek (včetně hodnoty vzorkovací frekvence) jsou přečtena interní funkcí *audioread* a uložena do vytvořené struktury *database*, respektive jejich podstruktur dle typu úlohy. Zároveň je do další podstruktury uložena zdrojová cesta jednotlivých wav souborů, která je potřebná z hlediska identifikace dat.

4.2 Hlavní skript

Prvním krokem hlavního skriptu *Main.m* je volání funkce *Dataprep.m*, jejímž výstupem je výše popsáná struktura *database*, čímž jsou tedy přečtena všechna potřebná data nahrávek. Následně proběhne pro jednotlivé řečníky cyklus *for*, jehož součástí je rozlišení typu úlohy, detekce a následné přeskočení chybějících nahrávek, určení, o jaký jazyk se jedná, na základě čehož je pak definována příslušná proměnná. Tato proměnná je nezbytná především z hlediska přičtení bílého Gaussovského šumu (s odstupem 10, 20 a 30 dB od čistého signálu), což je provedeno pouze u záznamů v českém jazyce. Přičtení šumu je provedeno interní funkcí *awgn*.

```
database.(tsk){pat,3}=awgn(database.(tsk){pat,2},20,"measured");
```

Před voláním této funkce je úroveň signálu ještě normalizována, a to na hodnotu amplitudy 0,891, což odpovídá hodnotě -1 dB.

```
database.(tsk){pat,2} =  
(0.891/max(abs(database.(tsk){pat,2}))) * database.(tsk){pat,2};
```

Takto předpřipravená data vstupují do funkcí pro jednokanálová zvýraznění řeči. Jako první v pořadí proběhne výpočetně snazší funkce pro spektrální odečítání *Spectral_sub.m*, v níž probíhá předzpracování signálu, tedy segmentace a váhování Hammingovým oknem, výpočet odhadu šumu, VAD a samotný proces této metody, podrobněji popsáno v kapitole 4.3, stejně tak funkci pro aplikaci Kalmanova filtru je věnována samostatná kapitola 4.4.

Po zvýraznění řeči jsou wav soubory zapsány do nově vytvořené databáze, strukturou shodné databázi vstupní, která je prvně rozčleněna na skupinu českého a maďarského jazyka, ve kterých se nacházejí složky jednotlivých řečníků pojmenované jejich identifikačními čísly. V takto vytvořených složkách jsou pak již samotné nahrávky, opět popsány ID řečníka a typem úlohy. Aby byla databáze kompletní, jsou zde zapsány i nahrávky originální a s přidaným šumem. Soubory jsou odlišeny předponou C (čisté), N (zatížené šumem), S (po aplikaci spektrálního odečítání) a K (po aplikaci Kalmanova filtru).

Vzhledem k poněkud vyšší výpočetní náročnosti zapříčiněné především Kalmanovým filtrem a velikostí databáze, jsou v tomto hlavním skriptu přidána vyskakovací okna informující o průběhu celého procesu, aby byl zajištěn bližší přehled o fázi výpočtu.

4.2.1 Výpočet odstupů signálu od šumu (SNR)

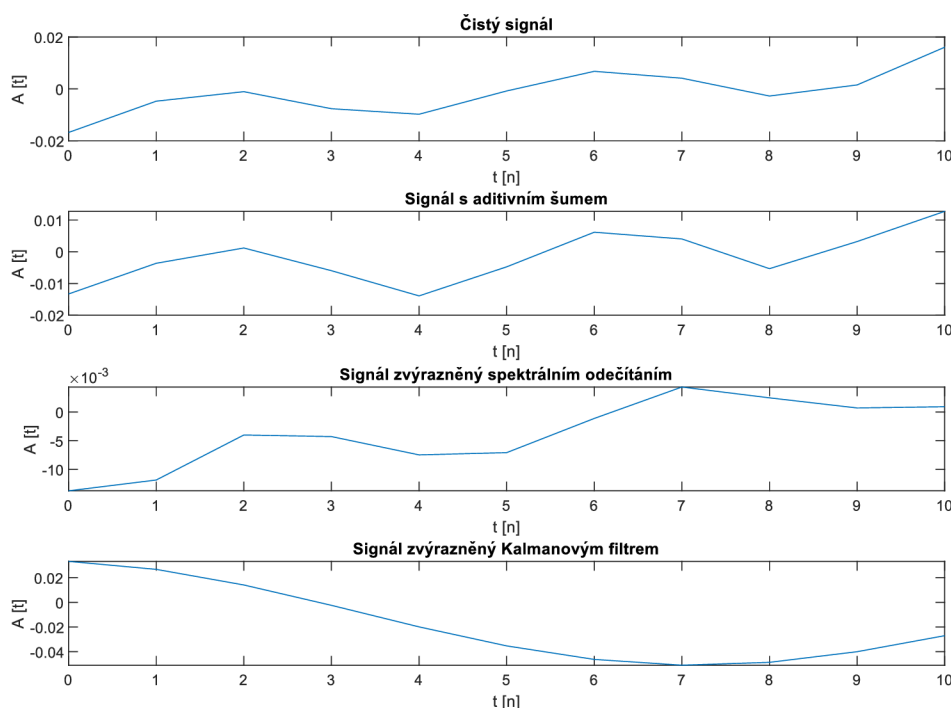
Před ukončením cyklu *for* proběhne ještě výpočet odstupů signálu od šumu, aby bylo možné vyhodnotit úspěšnost implementovaných metod. Je tedy volána funkce *Mysnr.m*.

```
database.(tsk){pat,4} = Mysnr(Y_S,  
    database.(tsk){pat,lan}(1:length(Y_S)));  
database.(tsk){pat,5} = Mysnr(Y_K,  
    database.(tsk){pat,lan}(1:length(Y_K)));
```

Vstupními hodnotami je buďto signál originálně čistý se signálem zatíženým aditivním šumem (pro ověření, že byl opravdu přidán šum o hodnotě SNR 10, 20 či 30 dB) nebo zde vstupuje signál upravený metodou pro zvýraznění řeči se signálem

před aplikací této metody, tedy opět řeč zatížená aditivním šumem. Cílenou hodnotou je v obou případech 10, 20 či 30 dB (dle konkrétního nastavení SNR aditivního šumu), jelikož je snahou se co nejvíce přiblížit originální nahrávce.

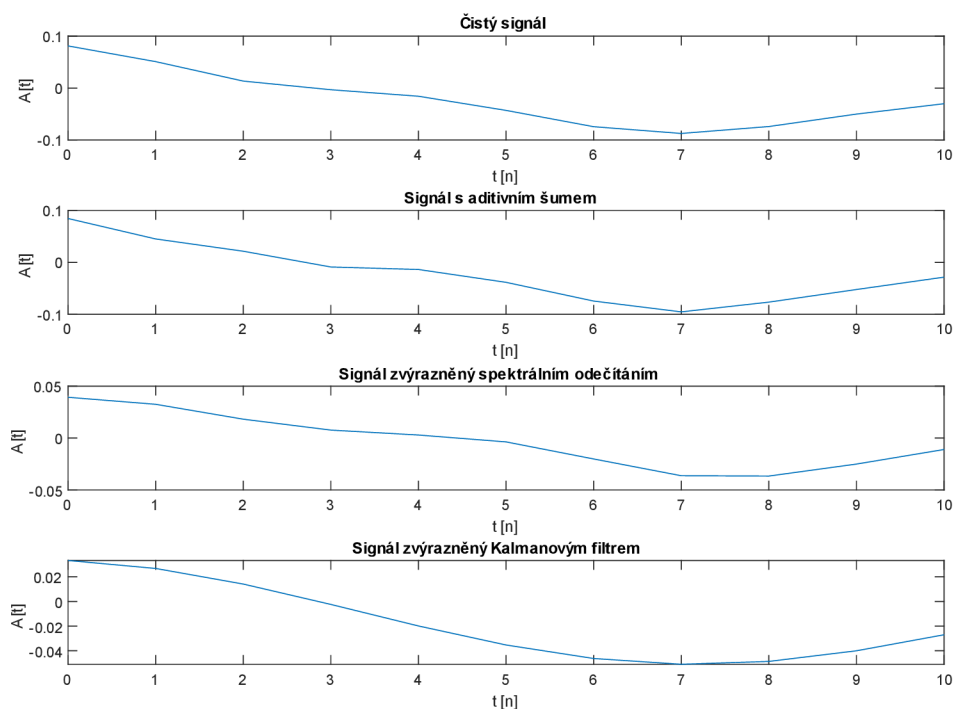
Před samotným výpočtem SNR bylo zapotřebí zvýrazněné nahrávky upravit z hlediska časového posunu vzorků zapříčiněného výše popsaným zpracováním těchto signálů, jelikož jsou běžné metody pro výpočet SNR náchylné na sebemenší asynchronii mezi vzorky porovnávaných signálů. Porovnání jednotlivých zpoždění lze vidět na obr. 4.1, kde je vykresleno 10 po sobě jdoucích vzorků signálů úlohy TSK2 pro všechny čtyři scénáře. Z průběhů je patrné, že k posunu dochází opravdu pouze až po implementacích jednokanálových metod zajišťujících zvýraznění řeči v nahrávkách.



Obr. 4.1: Porovnání časového zpoždění signálů

Aby tedy bylo možné získat hodnoty SNR, a tím i objektivně zhodnotit míru úspěšnosti spektrálního odečítání a Kalmanova filtru, bylo nezbytné vyrovnat tato zpoždění, což bylo provedeno pomocí interní funkce *alignsignals*, která byla zahrnuta do funkce *Mysnr.m*.

```
[enh, noisy] = alignsignals(enh, noisy);
```



Obr. 4.2: Porovnání kompenzace časového zpoždění signálů

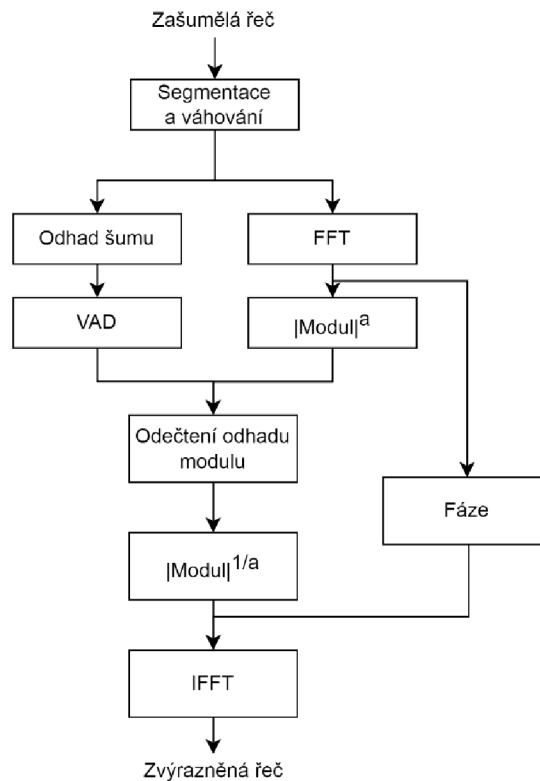
Funkce *Mysnr.m* dále vyrovná vstupující signály z hlediska jejich délek, zároveň odstraní stejnosměrnou složku a normalizuje signály totožným způsobem, jako je to provedeno před vstupem signálů do funkcí pro zvýraznění řeči. Po tomto předzpracování signálů je vypočtena hodnota energie signálu bez zatížení šumem a energie odhadovaného šumu, získaného jako rozdíl vstupních signálů funkce. Z těchto hodnot se již vypočítá samotný odstup užitečného signálu od šumu dle vzorce 11.

$$SNR = 10 \log_{10} \frac{E_y}{E_d} = 10 \log_{10} \frac{\sum_k Y^2[k]}{\sum_k [Y[k] - x[k]]^2} \text{ [dB]}. \quad (11)$$

Hodnota E_y je energií signálu Y (enh – konkrétně tedy Y_S či Y_K , dle příslušné metody) a E_d je energie šumu získaného rozdílem upraveného signálu Y a signálu zatíženého šumem (noisy).

Takto byly vypočteny hodnoty SNR pro odstup aditivního šumu 10 dB, 20 dB a 30 dB, na základě čehož byla pro další zpracování využita pouze jedna z těchto variant, a to konkrétně 20 dB.

4.3 Spektrální odečítání



Obr. 4.3: Diagram implementace spektrálního odečítání

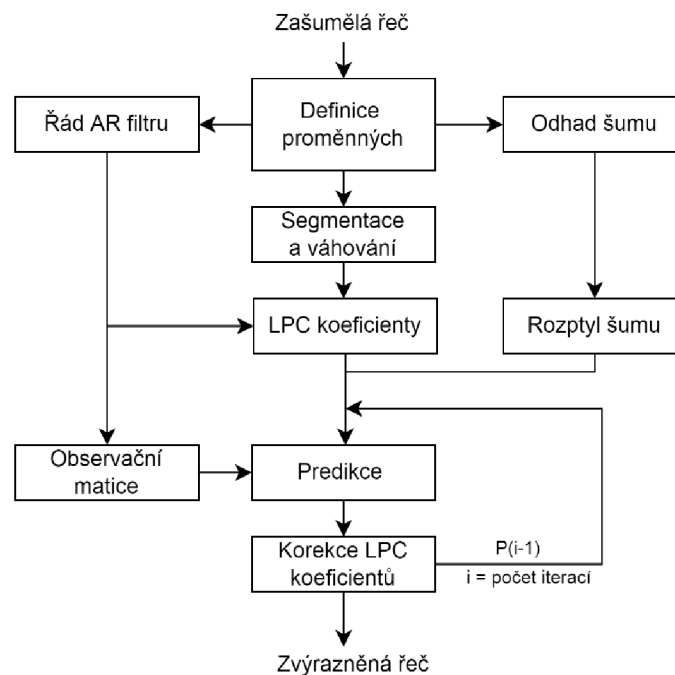
Implementována byla základní varianta spektrálního odečítání, a to funkcí *Spectral_sub.m* následujícím způsobem, jenž je rovněž vyobrazen diagramem na obr. 4.3. V prve řadě je znehodnocený vstupní signál této funkce rozdělen na jednotlivé segmenty, aby tak bylo možné signál v rámci jednoho tohoto segmentu více méně považovat za stacionární. Překryv segmentů je nastaven na běžně používaných 50 % a následně je provedeno váhování Hammingovým oknem.

Odhad šumu je v tomto případě získán z prvních pěti segmentů, což je sice implementačně snadná metoda, avšak ukázala se u tohoto způsobu zvýraznění řeči jako účinná. Tento způsob ovšem není možné úspěšně aplikovat na nahrávky s prodlouženou fonací. Navození fonace je zde totiž ve většině případů okamžité, proto nelze počáteční segmenty považovat za šum, a zároveň se zde nevyskytují jakékoliv pauzy, není zde tedy aplikovatelný ani detektor řečové aktivity (VAD). Vzhledem k tomu, že nebyl nalezen alternativní výpočet odhadu šumu, který by zde byl pro nahrávku přínosnější nežli degradující, ale zároveň spadal do kategorie výpočetně přívětivějších přístupů, bylo toto

úskalí vyřešeno tak, že výstupem každého spektrálního odečítání pro úlohy TSK2 je i odhadovaný šum, který je pak aplikován na nadcházející úlohu TSK6. Aby nedošlo ke změně hodnot odhadovaného šumu aktuální nahrávkou, je u TSK6 přeskočen výpočet, respektive aktualizace odhadu šumu pomocí VAD.

Jelikož je již signál rozčleněn na jednotlivé segmenty, je tak možné aplikovat rychlou Fourierovu transformaci, kterou lze použít pouze pro signál stacionární. Následně je tak zjištěn modul a fáze spektra (která není jakkoliv pozměněna) a dle principu uvedeného v kapitole 3.1 je vypočten odhad čistého signálu. Výsledný signál s potlačeným šumem je získán inverzní Fourierovou transformací odhadu čistého signálu.

4.4 Kalmanův filtr



Obr. 4.4: Diagram implementace Kalmanova filtru

Tato komplexnější metoda byla implementována funkcí *Kalman_filter.m* dle [24]. Před samotnou filtrací do funkce vstupuje odhad šumu získaný funkcí *Connfreq_estimation.m*, a to úspěšně pro obě úlohy. Dále následuje typické předzpracování signálu funkcí *Chopper.m*, po čemž je dle definovaného řádu autoregresivního filtru inicializována matice popsaná blíže v kapitole 3.3 a zároveň jsou pomocí interní funkce *lpc* nalezeny lineárně predikční koeficienty filtru. Následně proběhne cyklus *for*, jehož počet cyklů odpovídá stanovenému počtu iterací. V rámci tohoto cyklu jsou LPC koeficienty AR filtru aktualizovány za využití dříve inicializovaných matic.

5. PARAMETRIZACE

V tab. 5.1 je uveden souhrn všech použitých parametrů, včetně typu úlohy, ze které byly extrahovány, a charakteristické poruchy, pro jejichž detekci a stanovení míry těchto poruch jsou využívány. Blíže popsány, především pak z hlediska teorie, jsou tyto parametry v kapitole 2.2.

Tab. 5.1: Souhrn užitých parametrů

ÚLOHA	AKUSTICKÝ PARAMETR	CHARAKTERISTICKÉ PORUCHY
ARTIKULACE		
TSK2	RFA1	Porucha artikulace
TSK2	RFA2	Porucha artikulace
TSK2	#loc_max	Porucha artikulace
TSK2	relF1SD	Ztuhlost jazyka a čelisti
TSK2	relF2SD	Ztuhlost jazyka a čelisti
TSK2	#lndmrk	Nepřesná artikulace
PROZODIE		
TSK2	relF0SD	Monotónnost hlasu
TSK2	relSE0SD	Snížená variace hlasitosti
TSK2	EEVOL	Nestabilní hlasitost
TSK2	SPIR	Tvorba pauz
TSK2	PPR	Vyšší poměr pauz
TSK2	DurMED	Delší trvání pauz
TSK2	DurMAD	Vyšší proměnlivost délky trvání pauz
TSK2	NST	Vyšší poměr pauz
FONACE		
TSK6	HNR	Nárůst šumu
TSK6	CPP	Zvýšená dyšnost
TSK6	HRF	Zvýšená dyšnost
TSK6	NAQ	Zvýšená ostrost hlasu
TSK6	relNAQSD	Ztuhlost hlasivek
TSK6	QQQ	Zvýšená ostrost hlasu
TSK6	relQQSD	Ztuhlost hlasivek
TSK6	relF0SD	Nepřavidelné kolísání výšky hlasu
TSK6	Jitter (PPQ)	Mikroperturbace frekvence
TSK6	Shimmer (APQ)	Mikroperturbace amplitudy
TSK6	DUV	Aperiodicita
TSK6	relF1SD	Chvění brady
TSK6	relF2SD	Chvění brady

Samotná parametrizace byla provedena opět v prostředí MATLAB R2022a, a to pomocí knihovny funkcí BDALabu (*Brain Diseases Analysis Laboratory*), konkrétně skriptu *getfeatures.m* a jemu přidruženým funkcím pro samotné výpočty jednotlivých parametrů. Skript je napsán jako *for* cyklus pro postupné čtení a následné zpracování jednotlivých nahrávek, při čemž je prvně zpracována úloha TSK2, a poté úloha TSK6. V tomto *for* cyklu jsou volány příslušné funkce parametrů pro odpovídající úlohu, kdy výstupy těchto funkcí jsou ukládány do matice, která je poté zapsána jako tabulka. Výstupem parametrizace je tedy xlsx soubor, kde se ve sloupcích vyskytují výsledné hodnoty parametrů odpovídající danému mluvčímu, čemuž odpovídají řádky popsané jednotlivými identifikačními čísly. Tímto způsobem jsou zpracovány všechny čtyři scénáře, tedy nahrávky originálně čisté, zatížené šumem, zvýrazněné spektrálním odečítáním a zvýrazněné Kalmanovou filtrací.

Takto získané tabulky parametrizace jsou následně využity pro statistickou analýzu a klasifikaci za pomoci metod strojového učení.

6. STATISTICKÁ ANALÝZA

Pro statistickou analýzu bylo využito Kruskal-Wallisova testu. Jedná se o neparametrickou statistickou metodu vycházející z analýzy rozptylu (ANOVA), avšak s tím, že je Kruskal-Wallisův test přizpůsoben analýze parametrů, u nichž není předpokládáno vykazování normálního rozdělení. Tento test je tak modifikací Mann-Whitneyho testu, s tím rozdílem, že právě Kruskal-Wallisova metoda je vhodná pro porovnání více skupiny, kdežto Mann-Whitneyho lze aplikovat pouze na dva datové sady. Vzhledem k tomu, že bylo cílem porovnat hodnoty parametrů pro celkem čtyři scénáře, byl zvolen test Kruskal-Wallisův.

Výsledný vztah, respektive rozdíl, mezi všemi skupinami je vyhodnocen pomocí p -hodnoty. Pakliže je tato hodnota vyšší nežli koeficient *alfa* ($\alpha = 0,05$), nulová hypotéza není zamítnuta, což znamená, že rozdíl mediánů porovnávaných skupin není signifikantní a dané parametry pochází ze stejného rozdělení pravděpodobnosti.

Touto metodou ovšem není možné zjistit rozdíly mezi jednotlivými kombinacemi dvojic analyzovaných skupin, čímž by bylo možné posoudit vliv šumu na jednotlivé parametry a rovněž i dopad implementovaných jednonálových metod pro zvýraznění řeči. Je proto aplikována další statistická metoda, tzv. Dunnův neparametrický post-hoc test, který při porovnávání zároveň vyrovnává růst chyby prvního řádu. Hlavním výstupem Dunnova testu je Q -hodnota, která buďto zamítá, či vyvrací zamítnutí nulové hypotézy, a to na základě porovnání Q -hodnoty vůči její kritické hodnotě, která byla získána výpočtem zahrnující Šidákův parametr pro korekci chyby prvního řádu dle počtu porovnávaných skupin.

Pro tento postup statistické analýzy byl v prostředí MATLAB R2022a napsán skript, jenž čte data z tabulek získaných parametrizací a následně je analyzuje dle výše uvedeného, a to interní funkcí *kruskalwallis* a funkcí *dunn*. Veškeré výstupní hodnoty těchto funkcí jsou zapsány do vytvořené struktury *Stats*, která je dále rozčleněna dle příslušného jazyku nahrávek, akustických parametrů až na jednotlivé statistické metody.

Pro požadované parametry se zadáním hodnoty 1 až 30 (odpovídající pořadí voleného parametru) v příslušné podmínce *if* při volání funkce *kruskalwallis* vykreslí graficky znázorněné distribuce a vzájemné porovnání parametru mezi skupinami, a to grafem houslovým a krabicovým.

7. KLASIFIKACE

Parametry byly klasifikovány binárně dle toho, zdali byly extrahovány z nahrávky zdravé kontroly (ZK) či mluvčího s Parkinsonovou nemocí (PN). Byla proto volena statistická metoda logistická regrese využívající strojové učení, která byla implementována v MATLAB R2022a pomocí toolboxu *Classification Learner*.

Do tohoto nástroje byly jako vstupní prvky přidány tabulky získané parametrizací rozšířené o určení, zdali se jedná o ZK či PN. Chybějící či nulové hodnoty zapříčiněné chybějící úlohou daného řečníka byly nahrazeny nečíslnou hodnotou NaN, aby jimi nedošlo k ovlivnění modelu.

V případě českých nahrávek bylo pro vstupní data využito 10násobné křížové validace. U nahrávek maďarsky mluvících řečníků byl zvolen speciální případ k -násobné křížové validace zvané *leave-one-out*, kdy počet podmnožin k odpovídá počtu vstupních dat, v tomto případě tedy počtu nahrávek. Tato metoda je vhodná pro menší databáze (jako je tomu právě u maďarských nahrávek, kterých je pouze 23), kdy je tak zajištěna vyšší úspěšnost, která navíc není závislá na nahodilém namíchání podmnožin. U velkých datových souborů by ovšem touto validací mohlo dojít k nadhodnocení úspěšnosti.

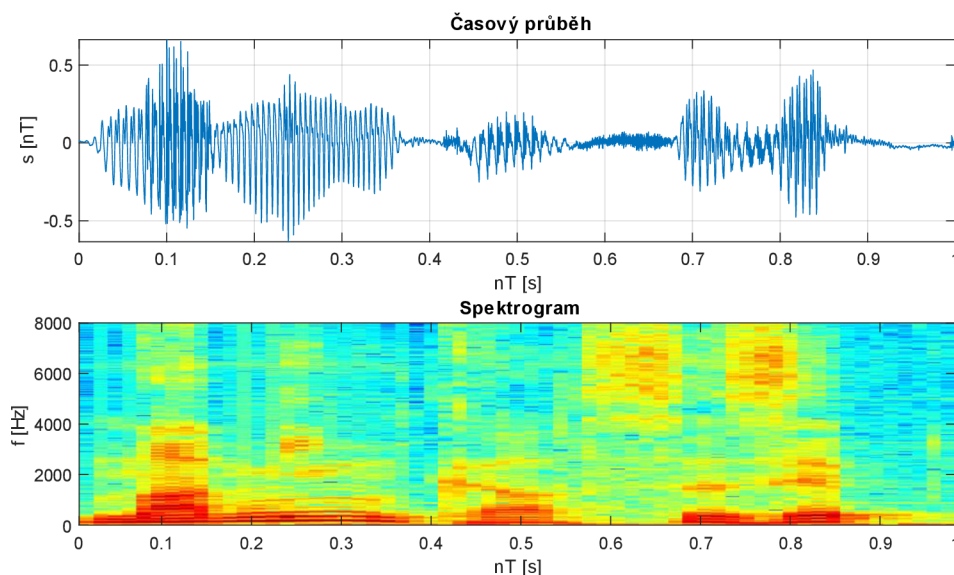
Po načtení dat byly pro samostatně zadané regrese zvoleny jednotlivé parametry pro daný scénář a jako poslední byla navolena regrese na jejíž vstup byly přivedeny všechny parametry současně. Pro jeden scénář jednoho jazyku nahrávek tak vzniklo celkem 31 logistických regresí, pro které byly následně vykresleny ROC křivky (*Receiver Operating Characteristic*, tedy křivka prahové operační charakteristika), jejíž prostřednictvím lze posoudit schopnosti dané statistické metody dle její senzitivity a specificity. Touto křivkou byly rovněž získány, jakožto průměr všech validací, hodnoty AUC (*Area Under Curve*), tedy hodnoty plochy pod ROC křivkou, která se může pohybovat v rozmezí 0 až 1, kdy platí, že čím vyšší hodnoty AUC nabývá, tím je daná klasifikace efektivnější.

8. VÝSLEDKY

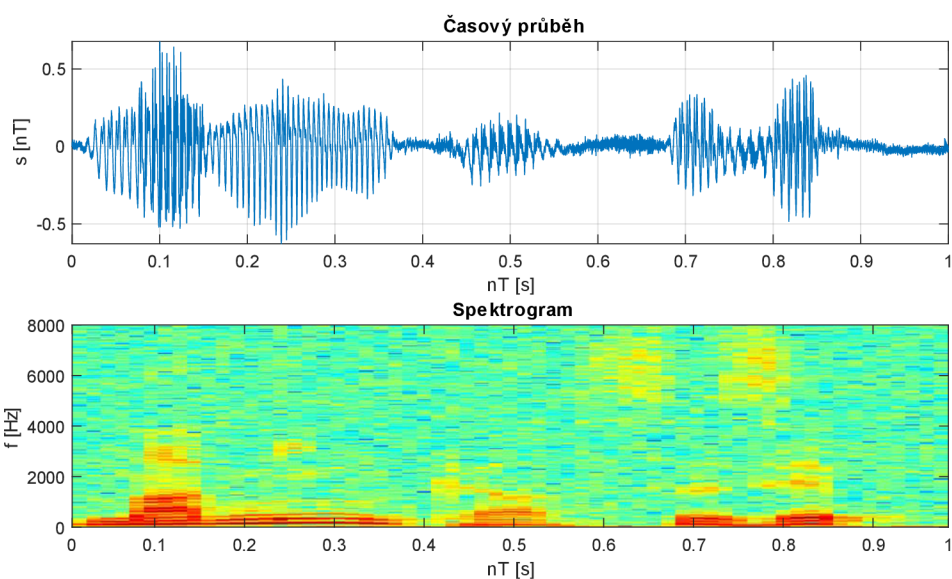
8.1 Výsledky implementací metod pro zvýraznění řeči

Výsledky implementací jednotlivých metod pro zvýraznění řeči jsou níže graficky vyobrazeny časovým průběhem a spektrogramem úzkopásmovým (1 sekundy při vzorkovací frekvenci $f_{vz} = 16$ kHz), aby byla zajištěna lepší frekvenční nežli časová čitelnost. Pro porovnání jsou zde tato grafická zobrazení uvedena i pro nahrávku čistou, tedy bez zatížení šumem, a i pro nahrávku rušenou aditivním šumem.

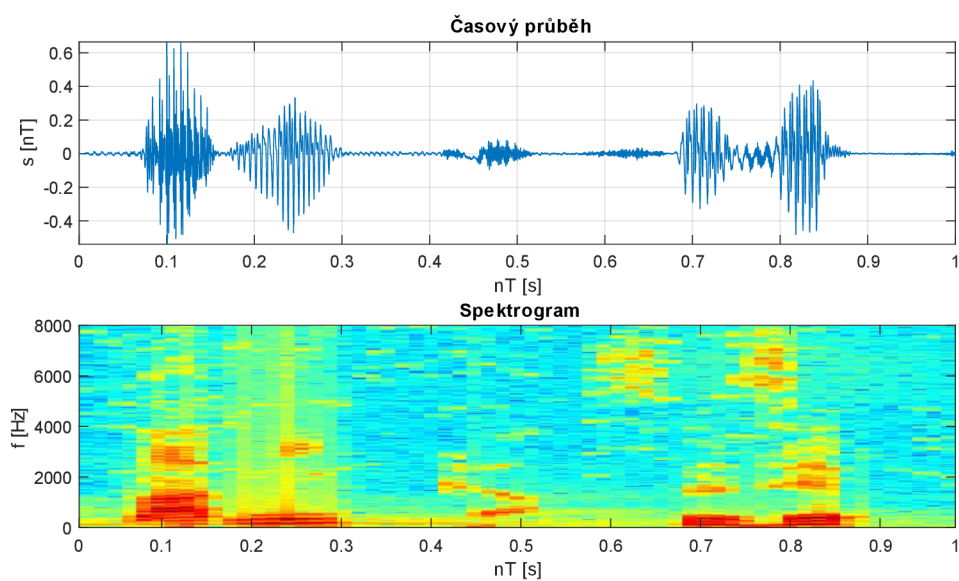
Z hlediska objektivního posouzení celé implementace byly vypočteny hodnoty odstupů signálu od šumu (SNR), které jsou vzhledem k rozsahu databáze uvedeny v přílohách C a D. Níže je však možné vidět jejich grafickou interpretaci za využití spojnicového grafu, aby tak bylo umožněno snadné vyčtení úspěšnosti těchto metod, přičemž jsou zde vyobrazeny hodnoty SNR pouze pro nahrávky české, jelikož je k nim známa referenční hodnota na základě nastaveného SNR aditivního šumu vůči čisté nahrávce, tedy 10 dB, 20 dB či 30 dB. Na ose x je tak vyneseno n nahrávek a na ose y jsou vyneseny nahrávkám odpovídající hodnoty odstupů signálu od šumu.



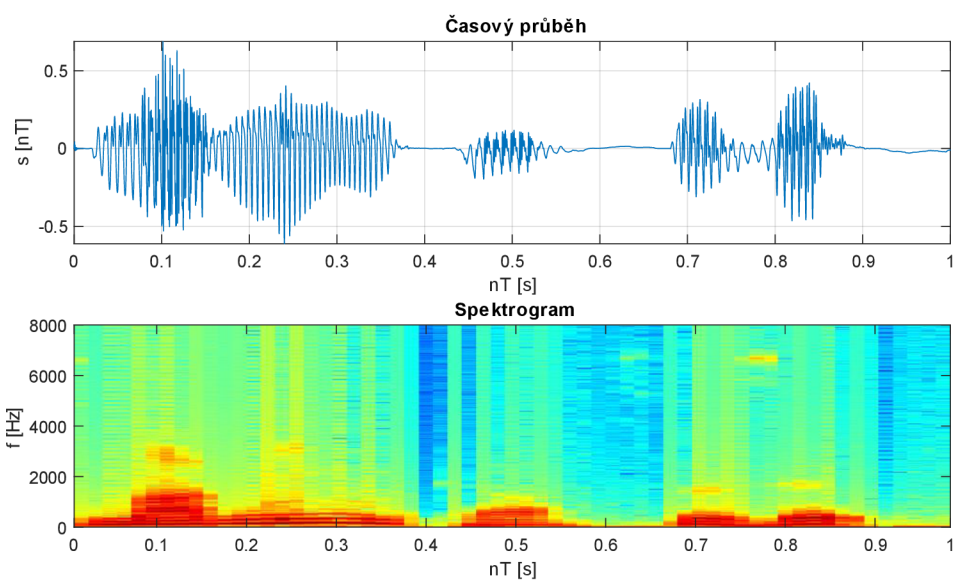
Obr. 8.1: Řečový signál originální české nahrávky čtení textu



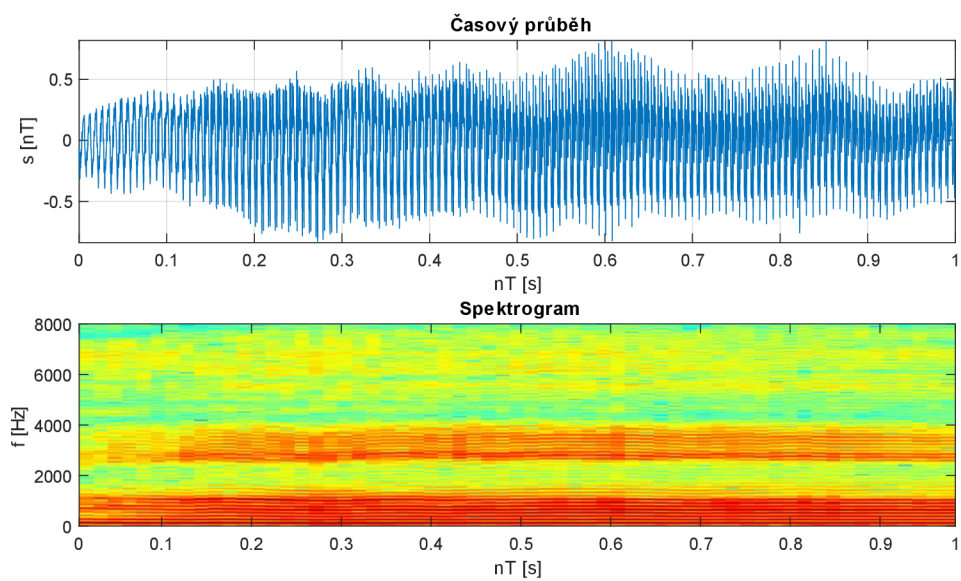
Obr. 8.2: Řečový signál české nahrávky čtení textu zatížené šumem



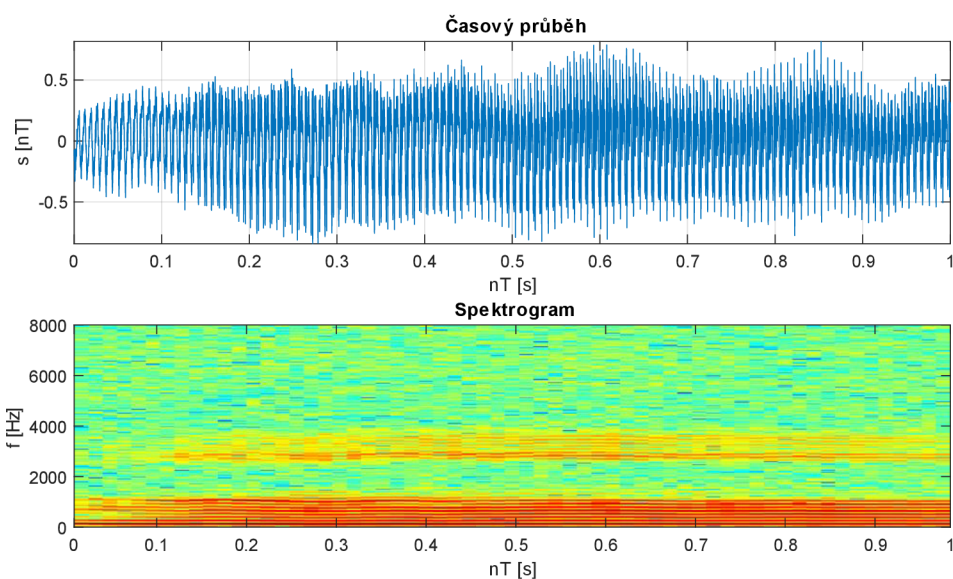
Obr. 8.3: Řečový signál české nahrávky čtení textu zvýrazněné spektrálním odečítáním



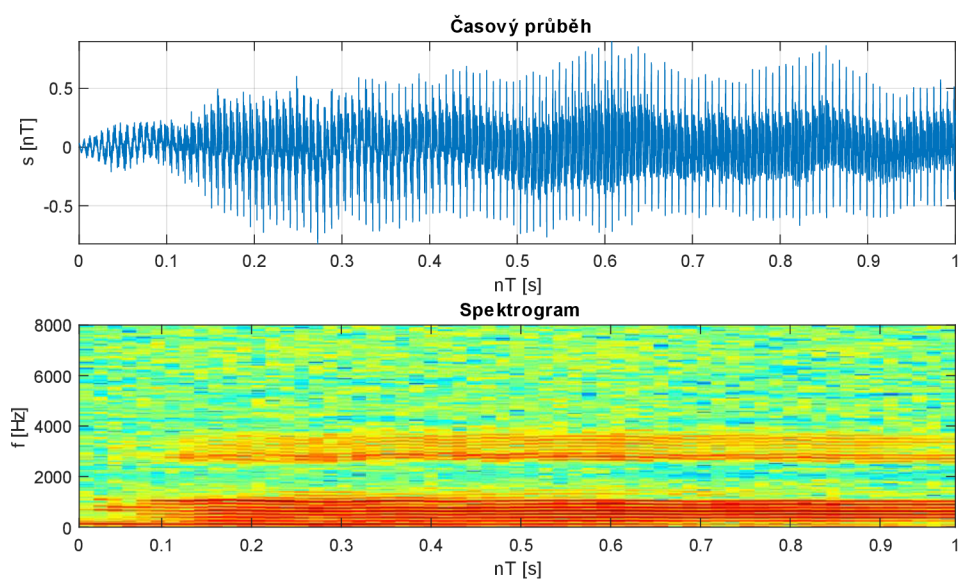
Obr. 8.4: Řečový signál české nahrávky čtení textu zvýraznění Kalmanovým filtrem



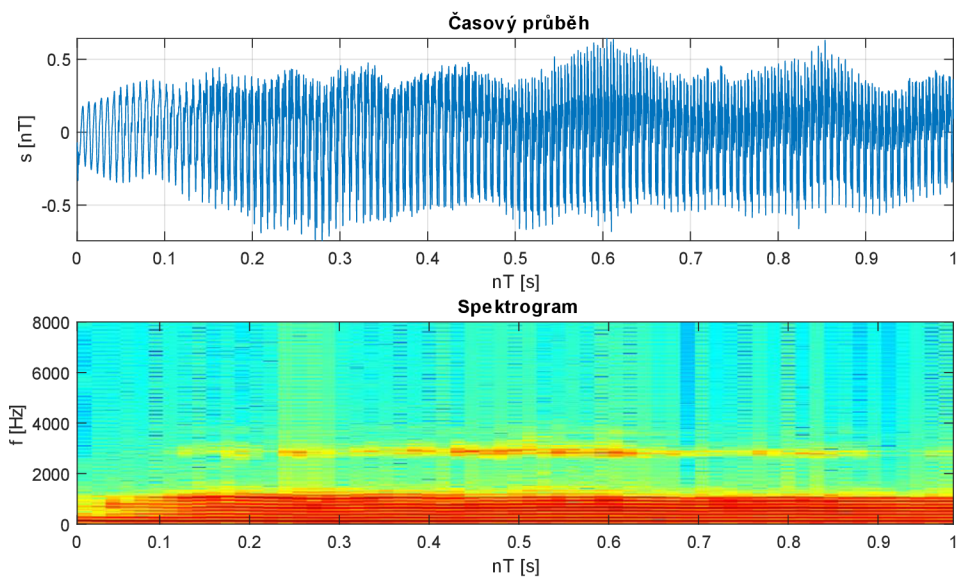
Obr. 8.5: Řečový signál originální české nahrávky prodloužené fonace



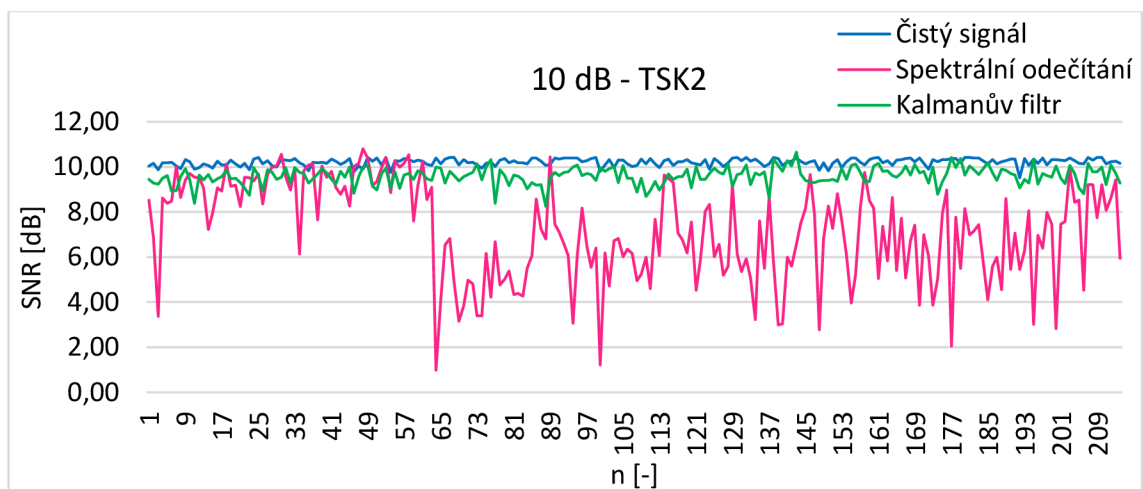
Obr. 8.6: Řečový signál české nahrávky prodloužené fonace zatížené šumem



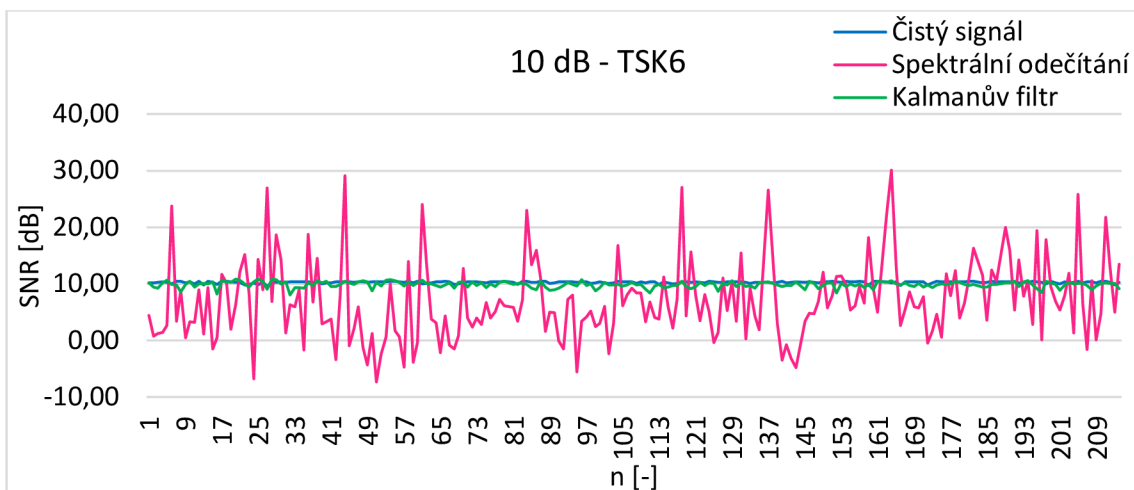
Obr. 8.7: Řečový signál české nahrávky prodloužené fonace zvýrazněné spektrálním odečítáním



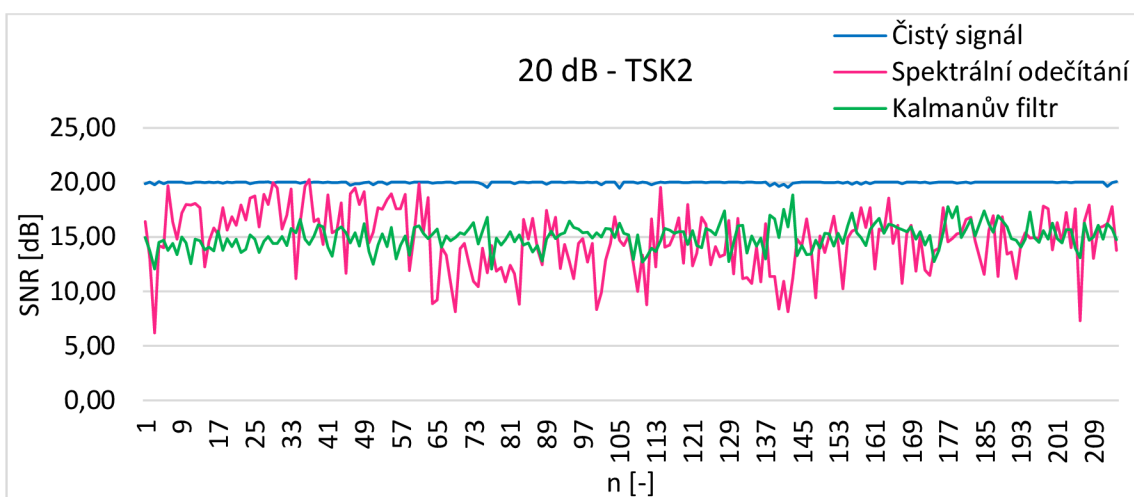
Obr. 8.8: Řečový signál české nahrávky prodloužené fonace zvýraznění Kalmanovým filtrem



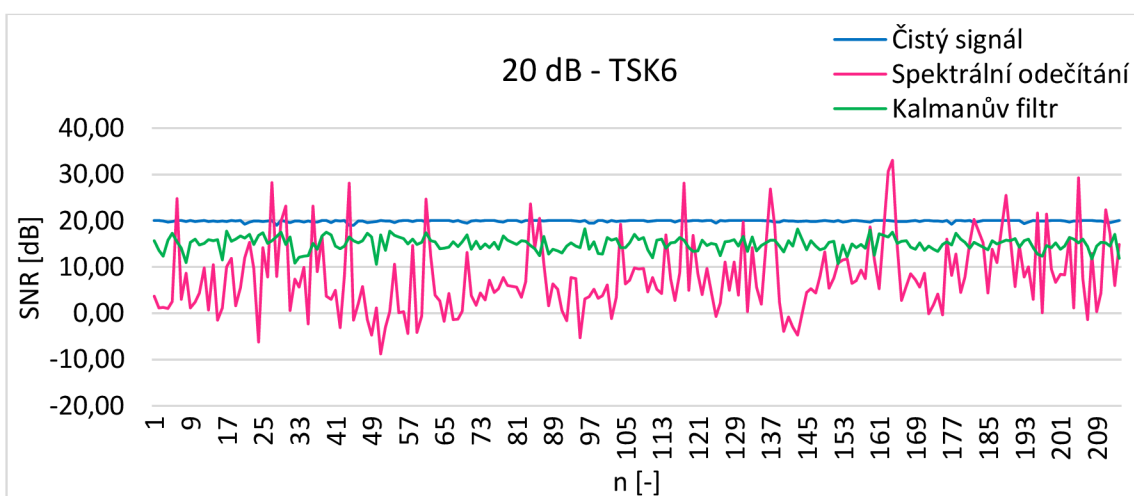
Obr. 8.9: Hodnoty SNR nahrávek čtení textu – potlačení aditivního šumu s odstupem 10 dB od originálního signálu



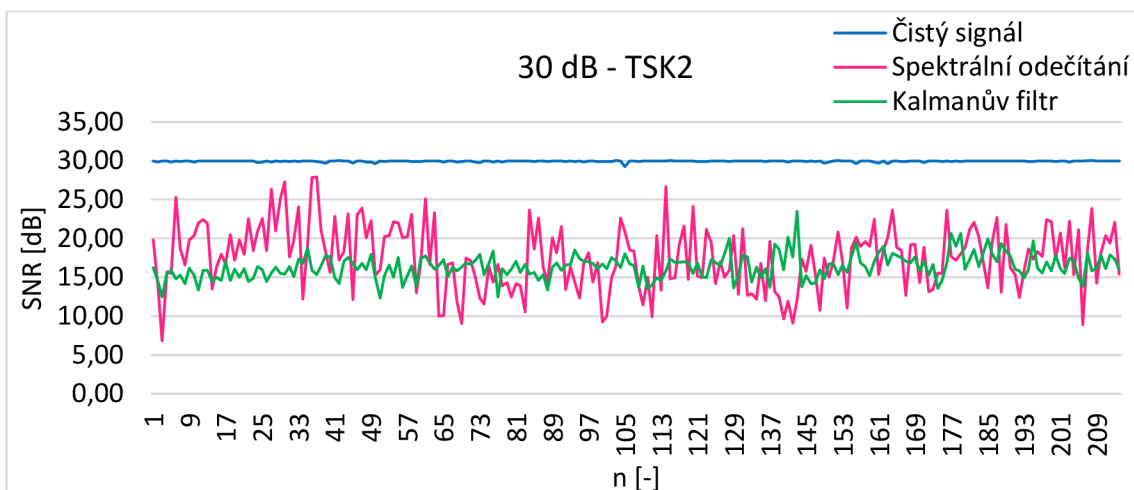
Obr. 8.10: Hodnoty SNR nahrávek prodloužené fonace – potlačení aditivního šumu s odstupem 10 dB od originálního signálu



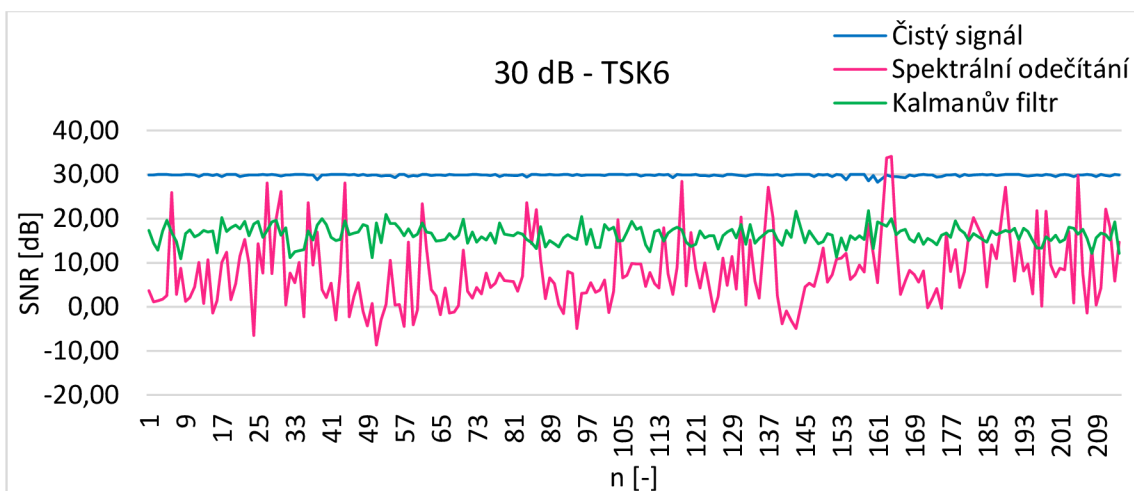
Obr. 8.11: Hodnoty SNR nahrávek čtení textu – potlačení aditivního šumu s odstupem 20 dB od originálního signálu



Obr. 8.12: Hodnoty SNR nahrávek prodloužené fonace – potlačení aditivního šumu s odstupem 20 dB od originálního signálu



Obr. 8.13: Hodnoty SNR nahrávek čtení textu – potlačení aditivního šumu s odstupem 30 dB od originálního signálu



Obr. 8.14: Hodnoty SNR nahrávek prodloužené fonace – potlačení aditivního šumu s odstupem 30 dB od originálního signálu

8.2 Výsledky statistické analýzy

8.2.1 Kruskal-Wallisův test

V tab. 8.1 jsou uvedeny výstupní p-hodnoty Kruskal-Wallisovy statistické metody pro porovnání jednotlivých parametrů napříč všemi skupinami pro nahrávky v českém jazyce, přičemž hodnoty ve druhém sloupci jsou pro přehlednost zaokrouhlením hodnot sloupce prvního.

U parametrů s p-hodnotou vyšší než $\alpha = 0,05$ není nulová hypotéza zamítnuta, což znamená, že rozdíl mediánů mezi jednotlivými skupinami není signifikantní. Hodnoty nižší než α (označeno *) odpovídají zamítnutí nulové hypotézy. V případě, že je hodnota výrazně nižší nežli α (přesněji $<0,01$), je značena **.

Tab. 8.1: P-hodnoty Kruskal-Wallisova testu pro jednotlivé parametry

Parametry	p_hodnota	Parametry	p_hodnota		
TSK2_RFA1	3,84E-08**	0,00	TSK6_GNE	1,15E-42**	0,00
TSK2_RFA2	7,68E-21**	0,00	TSK6_CPP	1,00E+00**	1,00
TSK2_loc_max	9,50E-18**	0,00	TSK6_HRF	1,01E-05**	0,00
TSK2_relF1SD	9,90E-58**	0,00	TSK6_QOQ	1,41E-09**	0,00
TSK2_relF2SD	1,48E-51**	0,00	TSK6_relQOQSD	4,00E-13**	0,00
TSK2_lndmrk	6,43E-24**	0,00	TSK6_NAQ	8,49E-14**	0,00
TSK2_relF0SD	3,85E-05**	0,00	TSK6_relNAQSD	1,05E-15**	0,00
TSK2_relSE0SD	3,32E-01	0,33	TSK6_relF0SD	1,28E-02*	0,01
TSK2_EEVOL	6,00E-01	0,60	TSK6_Jitter_PPQ	1,65E-01	0,17
TSK2_TST	9,98E-01	1,00	TSK6_Shimmer_APQ	8,15E-01	0,82
TSK2_NST	1,30E-05**	0,00	TSK6_HNR	2,30E-02*	0,02
TSK2_TPT_50	2,47E-10**	0,00	TSK6_DUV	2,34E-02*	0,02
TSK2_PPR	7,29E-07**	0,00	TSK6_relF1SD	3,19E-09**	0,00
TSK2_DurMED	6,73E-08**	0,00	TSK6_relF2SD	2,46E-23**	0,00
TSK2_DurMAD	1,04E-03**	0,00		<0,01 **	
TSK2_SPIR	7,60E-02	0,08		<0,05 *	

8.2.2 Dunnův test

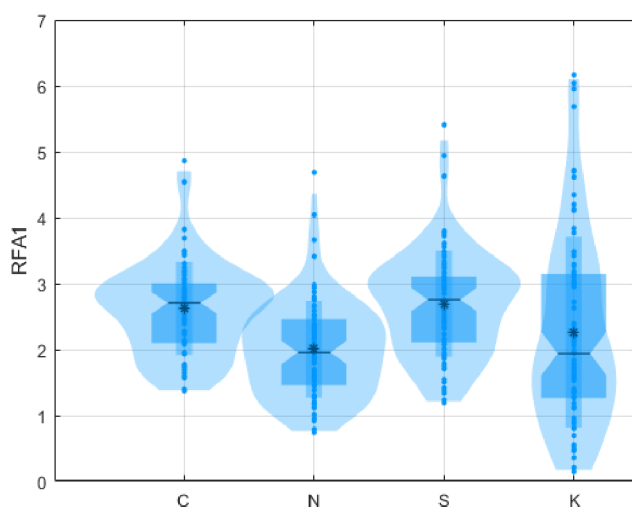
Výsledné Q-hodnoty Dunnova testu jsou vypsány v tab. 8.2, která je ve sloupcích rozčleněna dle jednotlivých kombinací porovnávaných skupin. První sloupec udává, zdali má šum zatěžující nahrávky vliv na analyzované parametry a sloupce S-C a K-C jsou vyhodnocením efektivity implementací metod pro zvýraznění řeči.

Tab. 8.2: Q-hodnoty Dunnova testu pro jednotlivé parametry

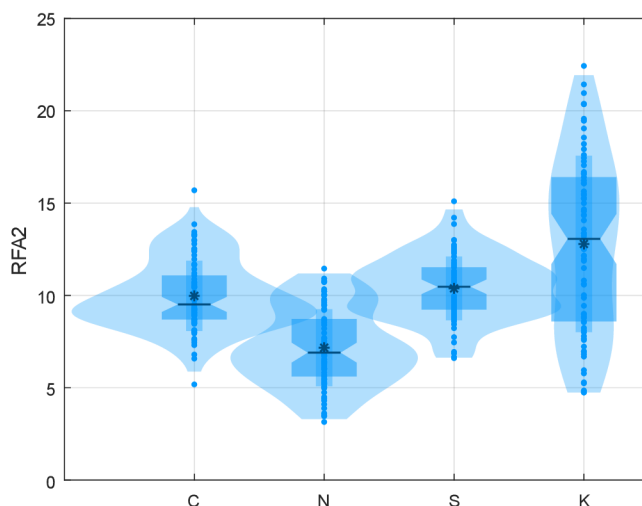
Parametry	Vliv šumu			Zvýraznění řeči		
	C-N	N-S	N-K	S-C	K-C	S-K
TSK2_RFA1	4,687	5,157	1,484	0,471	3,203	3,674
TSK2_RFA2	7,345	9,314	6,062	1,969	1,282	3,252
TSK2_loc_max	5,005	8,630	2,343	3,625	2,661	6,286
TSK2_relF1SD	8,370	16,343	7,659	7,973	0,711	8,684
TSK2_relF2SD	8,656	15,417	8,425	6,761	0,231	6,992
TSK2_lndmrk	6,802	9,425	2,099	2,623	4,703	7,325
TSK2_relF0SD	2,419	4,696	1,539	2,277		3,157
TSK2_relSE0SD	1,542	1,636	0,897			
TSK2_EEVOL	0,591	1,350	0,492			
TSK2_TST	0,146	0,146	0,146			
TSK2_NST	3,804	4,688	2,226		1,578	2,462
TSK2_TPT_50	5,535	6,344	4,034		1,501	2,310
TSK2_PPR	2,687	5,181	0,898	2,494	1,789	4,283
TSK2_DurMED	2,317	5,464	0,571	3,147		4,893
TSK2_DurMAD	1,905	3,732	0,650	1,826		3,081
TSK2_SPIR	1,847	1,973	0,120			
TSK6_GNE	8,849	13,806	8,797	4,957	0,053	5,010
TSK6_CPP	0,000	0,000	0,000			
TSK6_HRF	3,566	4,901	2,455		1,111	2,446
TSK6_QOQ	4,124	6,571	3,680	2,447	0,444	2,892
TSK6_relQOQSD	0,660	6,739	0,538	6,079		6,201
TSK6_NAQ	5,783	7,594	5,278		0,505	2,316
TSK6_relNAQSD	1,929	7,491	0,275	5,562		7,216
TSK6_relF0SD	0,539	2,962	0,410			2,552
TSK6_Jitter_PPQ	1,061	2,219	0,762			
TSK6_Shimmer_APQ	0,385	0,882	0,096			
TSK6_HNR	2,596	2,741	1,665			1,076
TSK6_DUV	1,148	2,861	0,443			2,418
TSK6_relF1SD	0,975	5,807	0,672	4,832		5,136
TSK6_relF2SD	5,445	10,364	4,511	4,919	0,934	5,853

C – originální nahrávka, N – nahrávka zatížena šumem, S – nahrávka zvýrazněna spektrálním odečítáním, K – nahrávka zvýrazněna Kalmanovým filtrem

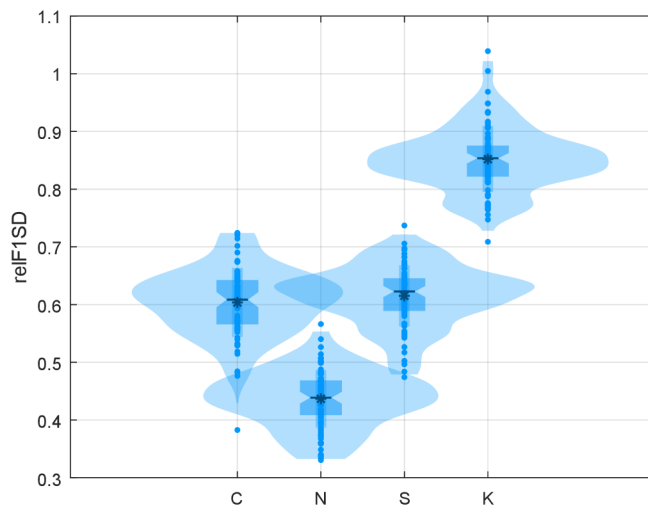
Kritickou hodnotou rozdělující výše uvedená data na zamítnutí a vyvrácení zamítnutí nulové hypotézy je $Q_{crit} = 2,631$. Q-hodnoty nižší (v tabulce nepodbarveny) tedy odpovídají vyvrácení zamítnutí nulové hypotézy, stejně tak tomu je u prázdných buněk, kdy pro tato data byl rozdíl při porovnání natolik nízký, že nebyla hodnota Q vypočítána. Hodnoty podbarvené modře ve sloupcích zvýraznění řeči ukazují na parametry, u kterých byla implementace spektrálního odečítání či Kalmanova filtru efektivní. Pro tyto parametry je proto níže graficky znázorněna distribuce hodnot daného parametru všech skupin.



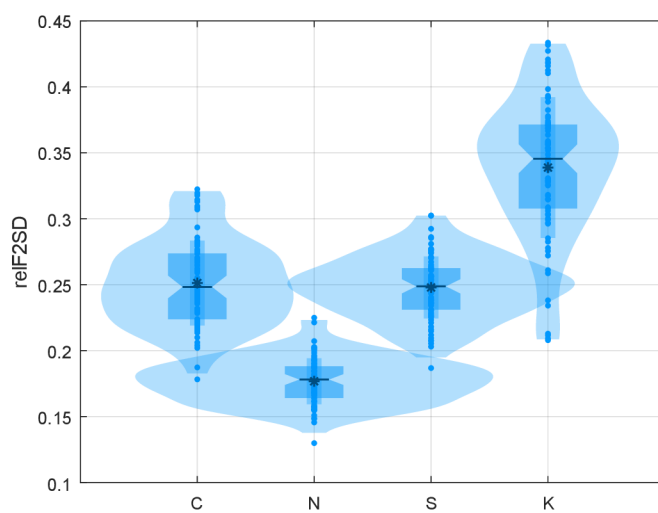
Obr. 8.15: Porovnání skupin pro RFA1 úlohy TSK2



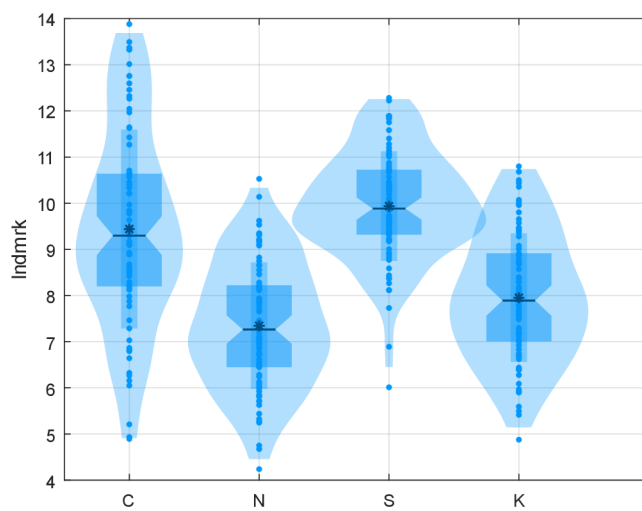
Obr. 8.16: Porovnání skupin pro RFA2 úlohy TSK2



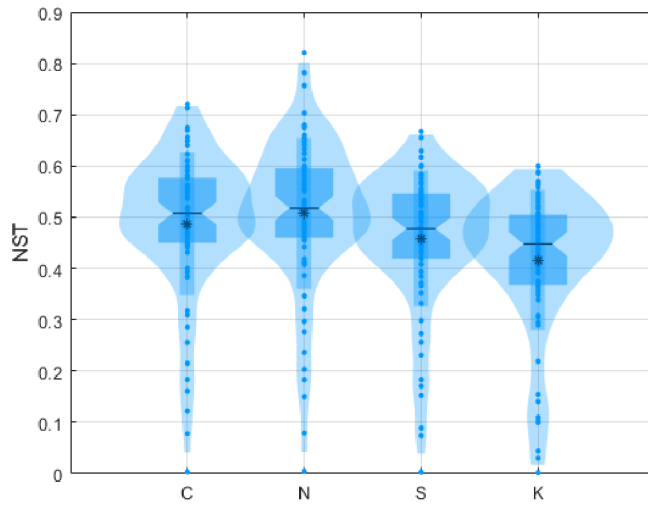
Obr. 8.17: Porovnání skupin pro relF1SD úlohy TSK2



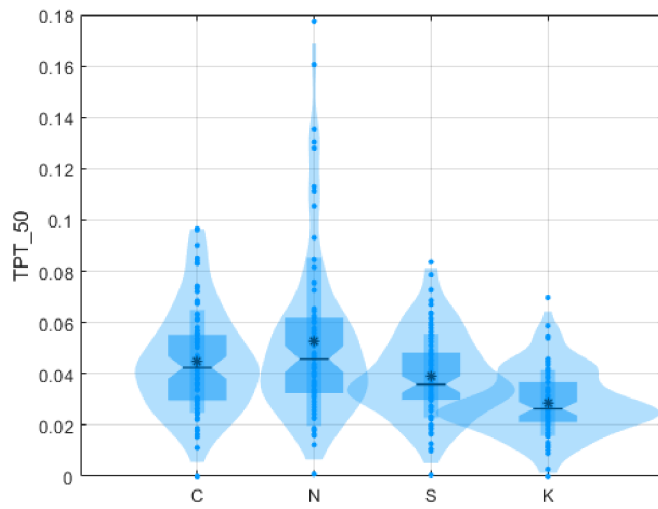
Obr. 8.18: Porovnání skupin pro relF2SD úlohy TSK2



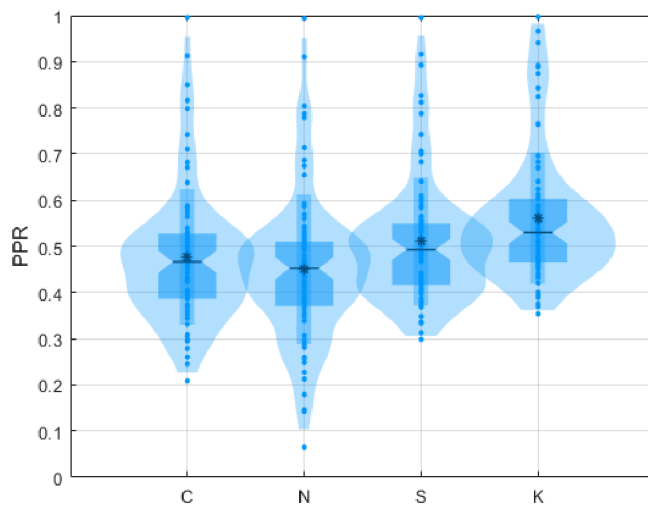
Obr. 8.19: Porovnání skupin pro Indmark úlohy TSK2



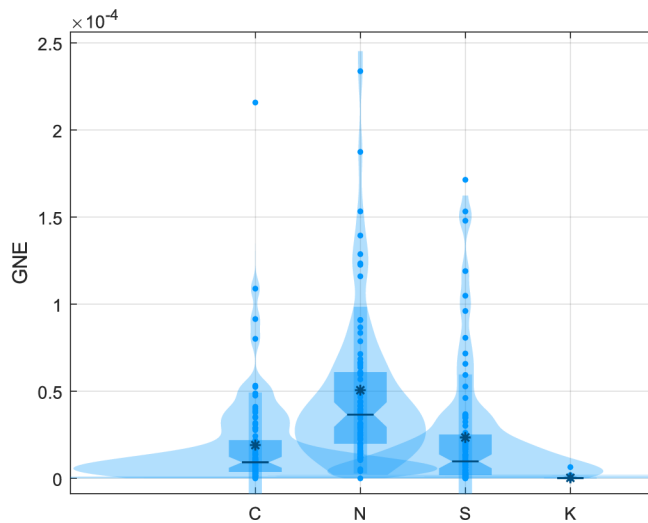
Obr. 8.20: Porovnání skupin pro NST úlohy TSK2



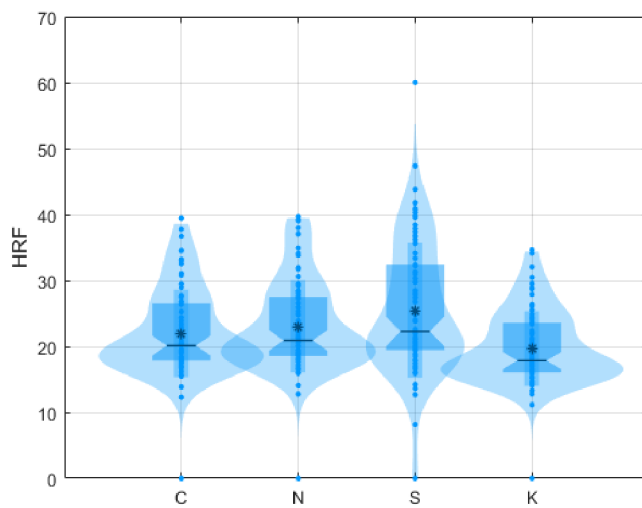
Obr. 8.21: Porovnání skupin pro TPT50 úlohy TSK2



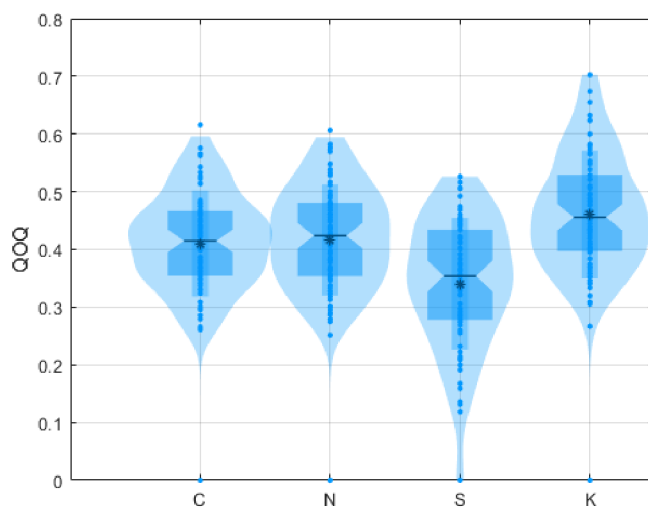
Obr. 8.22: Porovnání skupin pro PPR úlohy TSK2



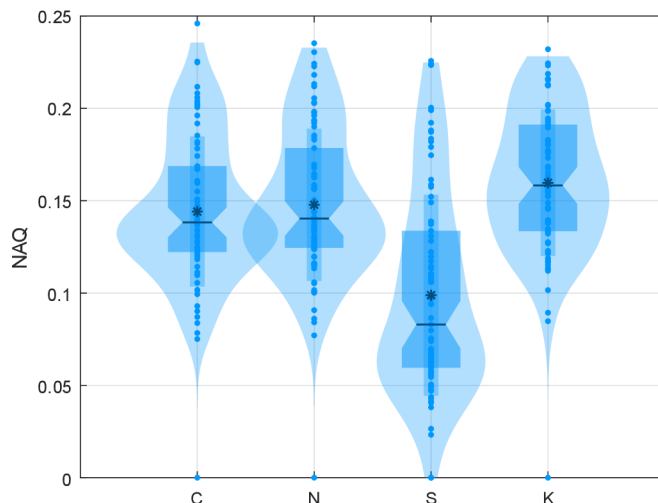
Obr. 8.23: Porovnání skupin pro GNE úlohy TSK6



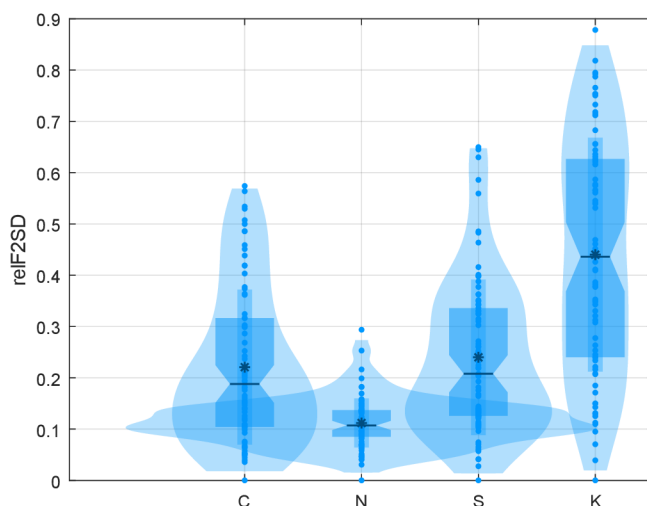
Obr. 8.24: Porovnání skupin pro HRF úlohy TSK6



Obr. 8.25: Porovnání skupin pro QOQ úlohy TSK6



Obr. 8.26: Porovnání skupin pro NAQ úlohy TSK6



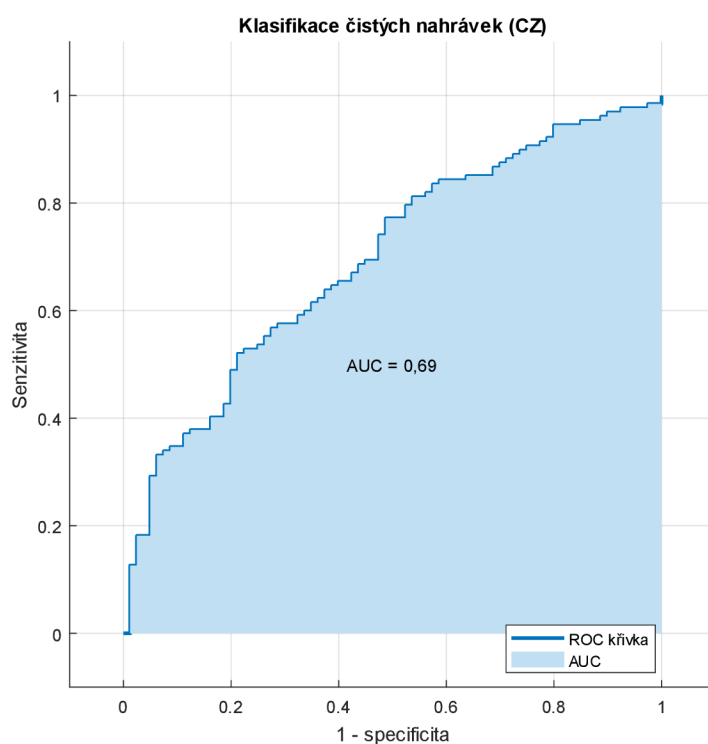
Obr. 8.27: Porovnání skupin pro relF2SD úlohy TSK6

8.3 Výsledky klasifikace

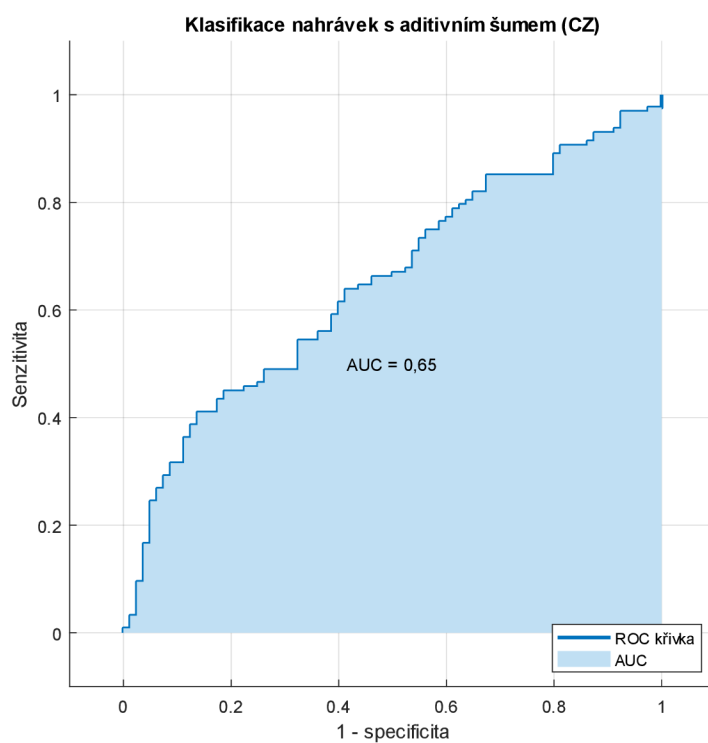
Výsledky klasifikace logistickou regresí jsou níže prezentovány jako ROC křivky (včetně uvedení AUC hodnot) pro případy, kdy jsou vstupními daty všechny parametry najednou. Jsou zde uvedeny výsledky pro klasifikaci jak českých nahrávek, tak i nahrávek maďarských. AUC hodnoty pro klasifikaci jednotlivých parametrů jsou uvedeny v tab. 8.3, ve které parametry psané tučným písmem jsou z hlediska klasifikace ovlivněny aditivním šumem, hodnoty podbarvené modře značí zlepšení klasifikace zvýrazněním nahrávek příslušnou metodou a parametry podbarvené šedě by dle statistické analýzy měly být, po stránce ovlivnění šumem, robustní, jinými slovy tedy rozdíl hodnot napříč skupinami není signifikantní.

Tab. 8.3: AUC hodnoty klasifikace logistickou regresí

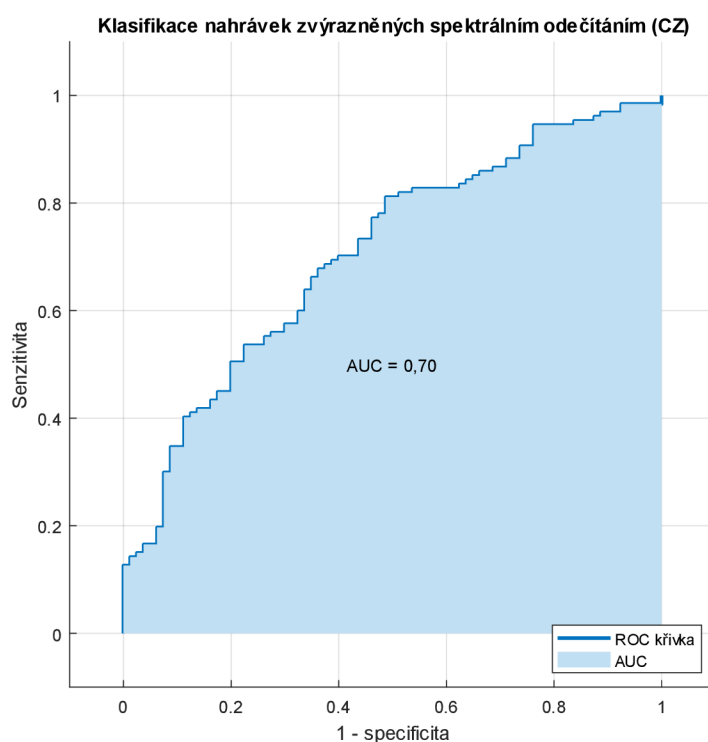
Parametry	AUC_CZ				AUC_HU		
	C	N	S	K	N	S	K
TSK2_RFA1	0,67	0,65	0,69	0,63	0,62	0,51	0,47
TSK2_RFA2	0,59	0,61	0,66	0,72	0,79	0,66	0
TSK2_loc_max	0,37	0,52	0,50	0,65	0,69	0,68	0,58
TSK2_relF1SD	0,56	0,58	0,54	0,51	0	0,45	0
TSK2_relF2SD	0,62	0,54	0,60	0,72	0	0,25	0,69
TSK2_indmrk	0,51	0,47	0,50	0,57	0,6	0,61	0,59
TSK2_relF0SD	0,64	0,55	0,61	0,61	0	0	0,43
TSK2_relSE0SD	0,44	0,42	0,48	0,60	0,36	0,3	0,06
TSK2_EEVOL	0,42	0,45	0,41	0,45	0,50	0,43	0,43
TSK2_TST	0,48	0,46	0,51	0,71	0,48	0,48	0,48
TSK2_NST	0,46	0,44	0,50	0,54	0,59	0,55	0,54
TSK2_TPT_50	0,44	0,59	0,51	0,56	0	0,17	0,69
TSK2_PPR	0,48	0,38	0,42	0,53	0,63	0,59	0,56
TSK2_DurMED	0,48	0,45	0,50	0,56	0	0	0,45
TSK2_DurMAD	0,49	0,45	0,51	0,53	0	0	0
TSK2_SPIR	0,63	0,60	0,62	0,43	0	0	0
TSK6_GNE	0,48	0,53	0,55	0,54	0,10	0	0
TSK6_CPP	0,48	0,48	0,49	0,49	0	0	0
TSK6_HRF	0,55	0,57	0,45	0,59	0,20	0,73	0
TSK6_QOQ	0,40	0,44	0,53	0,52	0,03	0,62	0,48
TSK6_relQOQSD	0,56	0,44	0,57	0,63	0,60	0,25	0,61
TSK6_NAQ	0,51	0,53	0,58	0,52	0	0,43	0
TSK6_relNAQSD	0,50	0,42	0,57	0,65	0,73	0	0,64
TSK6_relF0SD	0,51	0,51	0,52	0,38	0	0,33	0
TSK6_Jitter_PPQ	0,56	0,50	0,49	0,60	0	0	0,04
TSK6_Shimmer_APQ	0,52	0,52	0,44	0,80	0,45	0,09	0,41
TSK6_HNR	0,48	0,49	0,51	0,75	0	0,41	0
TSK6_DUV	0,53	0,53	0,56	0,60	0	0,05	0
TSK6_relF1SD	0,52	0,56	0,53	0,62	0,48	0	0
TSK6_relF2SD	0,46	0,39	0,57	0,47	0,61	0	0,47
Všechny parametry	0,69	0,65	0,70	0,78	0,58	0,60	0,56



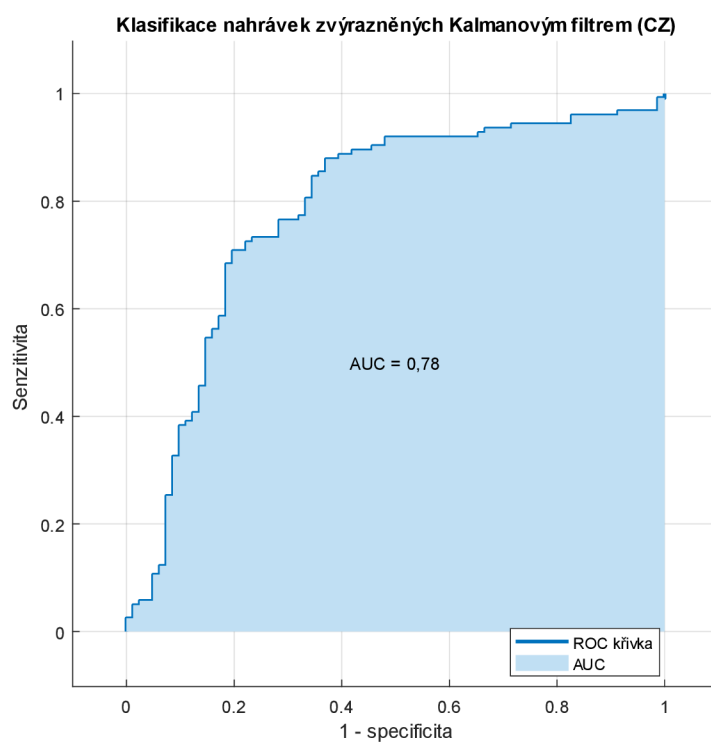
Obr. 8.28: ROC křivka klasifikace čistých nahrávek v českém jazyce



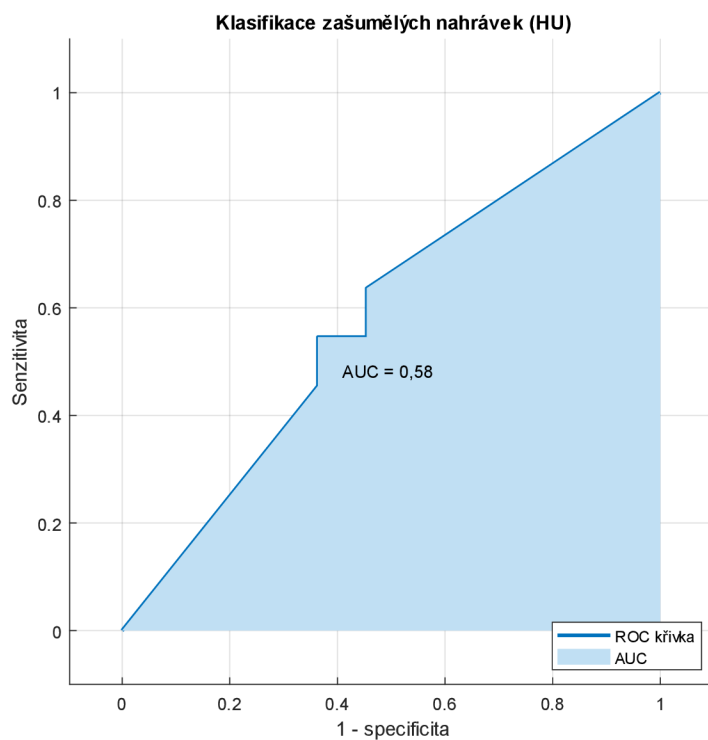
Obr. 8.29: ROC křivka klasifikace nahrávek v českém jazyce zatížených aditivním šumem



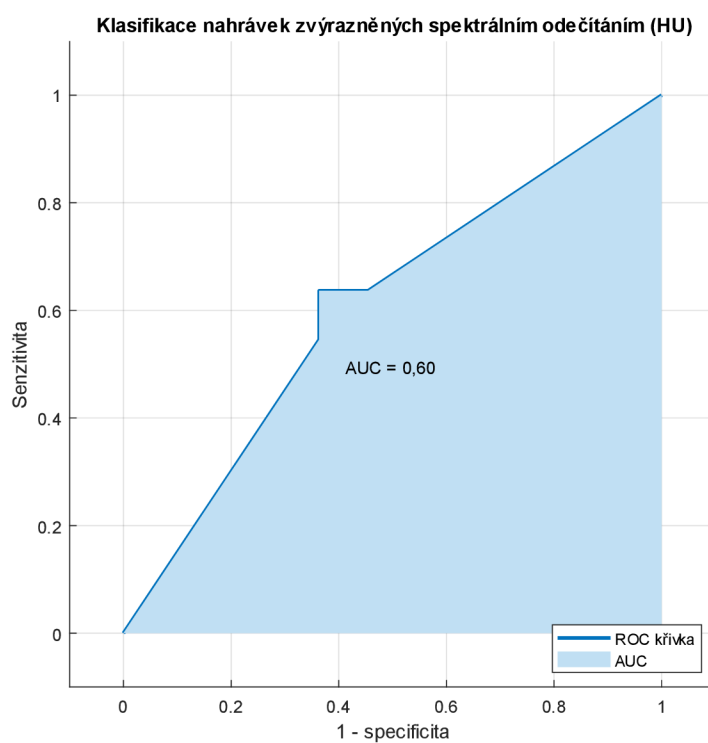
Obr. 8.30: ROC křivka klasifikace nahrávek v českém jazyce zvýrazněných spektrálním odečítáním



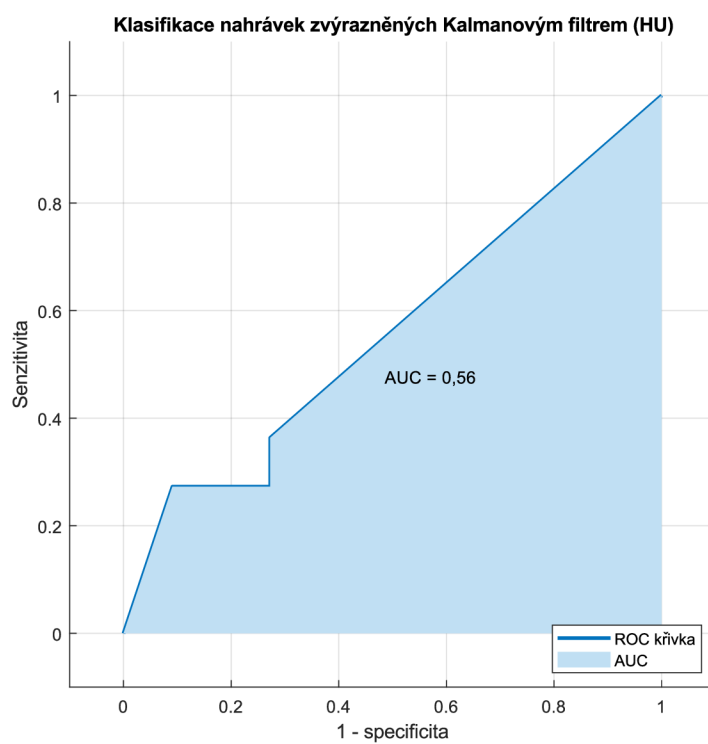
Obr. 8.31: ROC křivka klasifikace nahrávek v českém jazyce zvýrazněných Kalmanovým filtrem



Obr. 8.32: ROC křivka klasifikace nahrávek v maďarském jazyce zatížených šumem



Obr. 8.33: ROC křivka klasifikace nahrávek v maďarském jazyce zvýrazněných spektrálním odečítáním



Obr. 8.34: ROC křivka klasifikace nahrávek v maďarském jazyce zvýrazněných Kalmanovým filtrem

9. DISKUZE

9.1 Metody zvýraznění řeči

Dle výsledných hodnot SNR, a také podle poslechu jednotlivých nahrávek lze považovat implementaci metod pro zvýraznění řeči v nahrávkách určených k automatizované detekci HD za více méně úspěšnou. Jsou-li porovnány tyto metody mezi sebou, obecně lze říci, že metoda využívající Kalmanovu filtraci dopadla lépe, především pak, co se úloh TSK6 týče, kde u spektrálního odečítání i přes využití odhadu šumu z úloh TSK2 nebylo zvýraznění řeči v některých nahrávkách příliš markantní.

Jak již bylo zmíněno v kapitole 3.1, jednou z hlavních nevýhod výpočetně nenáročného spektrálního odečítání je pravděpodobný vznik hudebního šumu, který je jak ve spektrogramech na obr. 8.3 a obr. 8.7, tak i při poslechu samotných nahrávek lehce znatelný, avšak během samotného poslechu nepůsobí příliš rušivě.

Při využití Kalmanovy filtrace jsou spolu se šumem do určité míry odstraněny i vyšší frekvence užitečného signálu, což se při poslechu takto zvýrazněných nahrávek může mírně projevit jako aplikace low-pass filtru. Dle výsledků statistické analýzy a klasifikace se ovšem nejeví, že by tato skutečnost měla na parametry značný vliv, a to ani v případě např. příznaku HRF, který je definován jako poměr mezi energií první harmonické složky vůči celkové energii vyšších harmonických složek, které právě Kalmanovým filtrem mohou být dle spektrogramů v kapitole 8.1 lehce potlačeny.

9.2 Statistická analýza

Statistickými metodami pro analýzu vlivu šumu na parametry užívané k automatizované detekci HD bylo dle p-hodnot získaných Kruskal-Wallisovým testem (viz tab. 8.1) zjištěno, že k vyvrácení zamítnutí nulové hypotézy, tedy že ke změně mezi všemi skupinami nedochází u následujících parametrů:

- TSK2_relSE0SD
- TSK2_EEVOL
- TSK2_TST
- TSK2_SPIR

- TSK6_Jitter_PPQ
- TSK6_Shimmer_APQ

Lze tak tedy říci, že vliv šumu na tyto parametry není signifikantní.

Následně aplikovaným Dunnovým testem bylo zjištěno, pro které parametry bylo dané zvýraznění řeči efektivní.

V obou případech, tedy ať už bylo využito spektrálního odečítání či Kalmanova filtru, byly tyto metody efektivní u parametrů:

- TSK2_RFA2
- TSK2_PPR
- TSK6_QOQ

Samostatně pak bylo zvýraznění spektrálním odečítáním účinné pro parametry:

- TSK2_RFA1
- TSK2_Indrk

A zvýraznění řeči Kalmanovým filtrem se projevilo jako efektivní u parametrů:

- TSK2_relF1SD
- TSK2_relF2SD
- TSK2_NST
- TSK2_TPT50
- TSK6_GNE
- TSK6_HRF
- TSK6_NAQ
- TSK6_relF2SD

9.3 Klasifikace

Ačkoliv výsledné hodnoty AUC získané klasifikací nahrávek jednotlivými akustickými příznaky nejsou vyšší než 0,8, což znamená že úspěšnost klasifikace není 100%, je i přes to možné analyzovat vliv aditivního šumu i efektivitu použitých metod pro zvýraznění řeči.

Dle výsledků klasifikace logistickou regresí za pomoci strojového učení uvedených v tab. 8.3 bylo potvrzeno, že u téměř všech parametrů, které by měly být dle statistické analýzy robustní a vliv šumu by u nich neměl být příliš markantní, nedochází k ovlivnění

klasifikace využívající tyto příznaky. Výjimkou je pouze řečový příznak *jitter*, kde dochází k poklesu hodnoty AUC vlivem zatížení šumem, konkrétně z 0,56 na 0,50. K výraznějšímu ovlivnění dochází u parametrů:

- relF2SD
- lndmrk
- relF0SD
- PPR
- DurMAD
- relQOQSD
- relNAQSD
- relF2SD

Dle výsledků klasifikace je dále zjevný dopad zvýraznění řeči, především metodou využívající Kalmanovu filtraci, která v případě českých nahrávek zvýšila účinnost klasifikace celkem u 24 řečových příznaků, přičemž u 22 příznaků byl progres vyšší alespoň o 5 %. Mezi tyto parametry spadají i ty, u nichž se neprojevil značný vliv aditivního šumu. Tento aspekt je pravděpodobně zapříčiněn tím, že jsou spolu s velkým podílem šumu odfiltrovány i pro daný parametr redundantní složky, což může mít na výslednou efektivitu pozitivní dopad. Maximálním zlepšením účinnosti je zvýšení hodnoty AUC o 28 %, konkrétně u parametru *shimmer*, kdy je tak výsledná hodnota plochy pod křivkou rovna 0,8.

Co se spektrálního odečítání týče, zde u českých nahrávek došlo k úspěšnému zvýraznění mající vliv na klasifikaci celkem 22 příznaků, z čehož u 13 byla zvýšena plocha pod křivkou alespoň o 5 %, přičemž nejvýraznějším bylo ovlivnění parametru relF2SD, u kterého byla AUC hodnota navýšena o 18 %.

V případě modelu logistické regrese, kdy je klasifikace českých nahrávek trénována všemi parametry současně je určité zhoršení aditivním šumem opět zřetelné, stejně tak jako zlepšení klasifikace po zvýraznění řeči, především tedy Kalmanovým filtrem, kterým došlo k navýšení plochy pod ROC křivkou nejen vůči nahrávkám zatížených šumem, ale dokonce i v porovnání s nahrávkami původními.

Klasifikace prováděna z nahrávek řečníků hovořících maďarským jazykem v porovnání s nahrávkami českými dopadla výrazně hůře, což je pravděpodobně dáno malým množstvím dat této databáze, konkrétně se jedná o pouhých 23 nahrávek. I přes

to je však zjevné, že parametry, které by měly být dle předešle uvedeného odolnější vůči šumu, vykazují robustnost i v těchto nahrávkách. Respektive, v tomto případě neexistují referenční nahrávky řečových úloh, ale nedochází u těchto parametrů k výraznému či dokonce žádnému zlepšení klasifikace implementací metod pro zvýraznění řeči, jako je tomu například u parametru relF2SD, relF0SD, HNR a další.

Nižší efektivita implementovaných metod u této maďarské databáze vůči české může být dána rovněž faktem, že nahrávky maďarských mluvčích nejsou zatíženy aditivním bílým šumem, ale šumem neznámého charakteru spolu s dalšími ruchy.

10. ZÁVĚR

V první části práce byla popsána teorie nezbytná pro pochopení rozebírané problematiky zvýraznění řečového signálu. Mezi tuto teorii spadá jednak oblast vzniku a projevu hypokinetické dysartrie a parametry využívané pro popis řeči v nahrávkách. Dále byly rozebrány zvolené metody pro potlačení šumu, tedy implementačně a výpočetně přívětivé spektrální odečítání, a naopak o něco složitější a výpočetně náročný Kalmanův filtr. Tyto metody byly následně implementovány v prostředí MATLAB R2022a. Zvýraznění řeči proběhlo na databázi 214 originálně čistých nahrávek česky mluvících řečníků, ke kterým byl dále přičten aditivní bílý Gaussovský šum, a na 23 nahrávkách maďarsky hovořících mluvčích, kdy již původní verze nahrávek je znehodnocena šumem a dalšími ruchy.

Výsledky implementací byly z objektivního hlediska posouzeny vykreslením časových průběhů a spektrogramů a vypočtením hodnot odstupe užitečného signálu od šumu, tedy hodnot SNR. Dle těchto výsledků byla pro následné zpracování a analýzu zvolena pouze jedna úroveň odstupe signálu českých nahrávek od aditivního šumu. Touto hodnotou bylo SNR 20 dB, a to z toho důvodu, že v tomto případě je řeč stále srozumitelná, šum je zde dostatečně znatelný a zároveň je následné zvýraznění řeči efektivní.

Bylo rovněž zjištěno, že implementace jednokanálových metod pro zvýraznění řeči je úspěšnější, pakliže je úroveň rušivého šumu v nahrávce vyšší. U šumu slabšího, například s odstupem 30 dB SNR, nejsou metody tolik efektivní.

Dalším krokem byla parametrizace nahrávek čistých, zatížených šumem i zvýrazněných oběma metodami pro potlačení šumu. Vznikly tak celkem čtyři skupiny hodnot parametrů pro nahrávky české a tři skupiny pro maďarské. Takto rozčleněná data byla následně podrobena statistické analýze za využití Kruskal-Wallisova testu a Dunnova post-hoc testu. Výsledné p-hodnoty a Q-hodnoty určily pravděpodobnost distribuce daných dat, přičemž pro parametry, u kterých se jevil pozitivní dopad spektrálního odečítání či Kalmanova filtru byly krabicovým a houslovým grafem znázorněny distribuce dat ve všech čtyřech porovnávaných skupinách. Mezi tyto parametry patří, v případě úlohy čtení textu, RFA1, RFA2, relF1SDm relF2SD, lndmrk, NST, TPT50, PPR. Pro prodlouženou fonace jimi jsou GNE, QOQ, NAQ a relF2SD.

Na závěr byla provedena klasifikace nahrávek pomocí parametrů za využití logistické regrese aplikované za pomoci strojového učení. Pro klasifikaci, rovněž prováděnou v prostředí MATLAB R2022a, bylo využito aplikace *Classification Learner*, ve které byla logistická regrese za využití k-násobné křížové validace provedena klasifikace pro všechny parametry zvlášť a následně i dohromady. Ačkoliv přesnost logistické regrese nebyla excelentní, nebyl tento fakt brán příliš v potaz, jelikož stěžejní oblastí této klasifikace bylo porovnání výsledných AUC hodnot a zhodnocení jejich změn mezi skupinami. Tímto se prokázala robustnost parametrů, u nichž byla tato vlastnost předpokládána na základě statistické analýzy. Rovněž byl touto metodou potvrzen pozitivní dopad metod zvýrazňující řečový signál v nahrávkách, přičemž lze říci, že obecně byla úspěšnější metoda Kalmanovy filtrace, ovšem na úkor výpočetní a implementační náročnosti.

11. LITERATURA

- [1] MADRUGA, Mario, Yolanda CAMPOS-ROCA a Carlos PÉREZ. *Impact of noise on the performance of automatic systems for vocal fold lesions detection* [online]. 2021 [cit. 2022-05-22]. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S0208521621000887>
- [2] ČIHÁK, Radomír. *Anatomie 3*. Třetí, upravené a doplněné vydání. Grada, 2016. ISBN 978-80-247-5636-3.
- [3] KOŠTÁLOVÁ, Milena. *Test 3F Dysartrický profil – normativní hodnoty řeči v češtině* [online]. 2013 [cit. 2022-05-23]. Dostupné z: <https://www.csnn.eu/casopisy/ceska-slovenska-neurologie/2013-5/test-3f-dysartricky-profil-normativni-hodnoty-rci-v-cestine-41400>
- [4] ZAMIŠKOVÁ, Gabriela, Pavel RESSNER a Jana DLOUHÁ. *Poruchy řeči u Parkinsonovy nemoci* [online]. In: . 2010 [cit. 2022-05-23]. Dostupné z: <https://www.neurologiepropraxi.cz/pdfs/neu/2010/02/10.pdf>
- [5] REKTOROVÁ, Irena, Jiří MEKYSKA a Zdeněk SMÉKAL. *Motorické aspekty poruch řeči u Parkinsonovy nemoci a jejich hodnocení* [online]. In: . 2011 [cit. 2022-05-23].
- [6] TEIXEIRA, João, Carla OLIVEIRA a Carla LOPES. *Vocal Acoustic Analysis – Jitter, Shimmer and HNR Parameters* [online]. In: . 2013 [cit. 2022-05-23]. ISSN 22120173. Dostupné z: <https://doi.org/10.1016/j.protcy.2013.12.124>
- [7] J. T. DE SOUSA, Ricardo. *A new accurate method of harmonic-to-noise ratio extraction* [online]. In: . SciTePress, 2009 [cit. 2022-05-23]. Dostupné z: [doi:10.5220/0001552903510356](https://doi.org/10.5220/0001552903510356)
- [8] D. HEMAN-ACKAH, Yoland. *The Relationship Between Cepstral Peak Prominence and Selected Parameters of Dysphonia* [online]. In: . The Voice Foundation, 2002 [cit. 2022-05-23]. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S089219970200067X>
- [9] DRUGMAN, Thomas, Baris BOZKURT a Thierry DUTOIT. *A Comparative Study of Glottal Source Estimation Techniques* [online]. In: . 2019 [cit. 2022-05-23]. Dostupné z: <https://arxiv.org/pdf/2001.00840.pdf>
- [10] SKARNITZL, Radek. *Co dokáže náš hlas? Fonetický pohled na variabilitu řečové produkce* [online]. In: . 2017 [cit. 2022-05-23]. Dostupné z: https://www.researchgate.net/publication/315449555_Co_dokaze_nas_hlas_Foneticky_pohled_na_variabilitu_recove_produkce

- [11] ALKU, Paavo, Tom BÄCKSTRÖM a Erkki VILKMAN. *Normalized amplitude quotient for parametrization of the glottal flow* [online]. In: . 2002 [cit. 2022-05-23]. Dostupné z: <https://www.csc.kth.se/utbildning/kth/kurser/DT2212/Singing%20Alku%20et%20al%20JASA.pdf>
- [12] V DI, Nicola, Fiorella ML, Spinelli DA a Fiorella R. *Acoustic analysis of voice in patients treated by reconstructive subtotal laryngectomy. Evaluation and critical review* [online]. In: . Acta Otorhinolaryngol Ital. [cit. 2022-05-23]. Dostupné z: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2639982/>
- [13] TYKALOVÁ, T., J. RUSZ a R. ČMEJLA. *Formantové charakteristiky během vyjádření důrazů u Parkinsonovy nemoci* [online]. In: . [cit. 2022-05-23]. Dostupné z: https://www2.humusoft.cz/www/papers/tcp13/060_tykalova.pdf
- [14] KOVÁČ, Daniel. *Hodnocení poruch artikulace pacientů s Parkinsonovou nemocí založené na rezonancích v hlasovém traktu* [online]. In: . [cit. 2022-05-23]. Dostupné z: <http://www.elektrorevue.cz/cz/clanky/zpracovani-signalu/0/hodnoceni-poruch-artikulace-pacientu-s-parkinsonovou-nemoci-zalozene-na-rezonancich-v-hlasovem-traktu--assessment-of-articulation-impairment-in-patients-with-parkinson-disease-based-on-resonances-in-the-vocal-tract/>
- [15] SAPIR, Shimon, Jennifer SPIELMAN, Lorraine O RAMIG a Brad H STORY. *Effects of Intensive Voice Treatment (the Lee Silverman Voice Treatment [LSVT]) on Vowel Articulation in Dysarthric Individuals With Idiopathic Parkinson Disease: Acoustic and Perceptual Findings*. In: *Journal of Speech Language and Hearing Research* [online]. 2007 [cit. 2022-05-23]. Dostupné z: doi:10.1044/1092-4388(2007/064)
- [16] CoBeN speech and voice parameterization toolbox. In: *CoBeN* [online]. [cit. 2022-05-23]. Dostupné z: <https://coben.ceitec.cz/software/>
- [17] BENESTY, Jacob, M. SONDHI a Yiteng HUANG. *Springer Handbook of Speech Processing*. Springer, 2008. ISBN 978-3-540-49125-5.
- [18] QUATIERI, Thomas F. *Discrete-Time Speech Signal Processing: Principles and Practice*. Prentice Hall, 2002. ISBN 0-13-242942-X.
- [19] UPADHYAY, Navneet a Abhijit KARMAKAR. *Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study* [online]. In: . 2015 [cit. 2022-05-23]. Dostupné z: <https://doi.org/10.1016/j.procs.2015.06.066>
- [20] PHILIPOS, Loizou C. *Speech Enhancement: Theory and Practice*. Second Edition. CRC Press, 2013. ISBN 9781466504219.

- [21] BENESTY, Jacob, Shoji MAKINO a Jingdong CHEN. *Speech Enhancement*. Springer, 2005. ISBN 3-540-24039-X.
- [22] PALIWAL, K.K. a Anjan BASU. *A speech enhancement method based on Kalman filtering* [online]. In: . 1987 [cit. 2022-05-23]. Dostupné z: doi:DOI:10.1109/ICASSP.1987.1169756
- [23] MATHE, Mariyadasu, S.P. NANDYALA a T. KUMAR. *Speech enhancement using Kalman Filter for white, random and color noise* [online]. In: . 2012 [cit. 2022-05-23]. Dostupné z: doi:10.1109/ICDCSyst.2012.6188703
- [24] *Speech Enhancement Kalman Filter* [online]. In: . 2016 [cit. 2022-05-23]. Dostupné z: <https://github.com/Vishnu-Dhanabalan/Speech-Enhancement-Kalman-Filter>

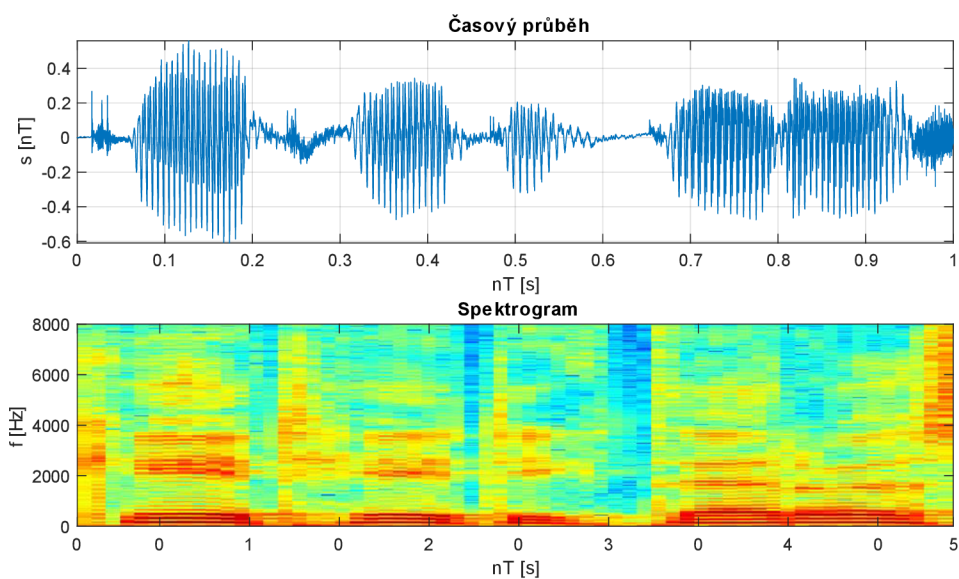
SEZNAM SYMBOLŮ A ZKRATEK

HD	hypokinetická dysartrie
PN	Parkinsonova nemoc
CNS	centrální nervová soustava
ZK	zdravá kontrola
$x(n)$	znehodnocený signál
$d(n)$	šum
$e(n)$	střední kvadratická chyba
$s(k)$	čistý signál
$y(k)$	výstupní vyčištěný signál
$u(k)$	budící aditivní šum
$z(t)$	signál obecné stavové rovnice
E_y	energie výstupního vyčištěného signálu y
E_d	energie šumu d
a_i	koeficienty AR filtru
H	observační matice
G	zesílení
T	délka signálu

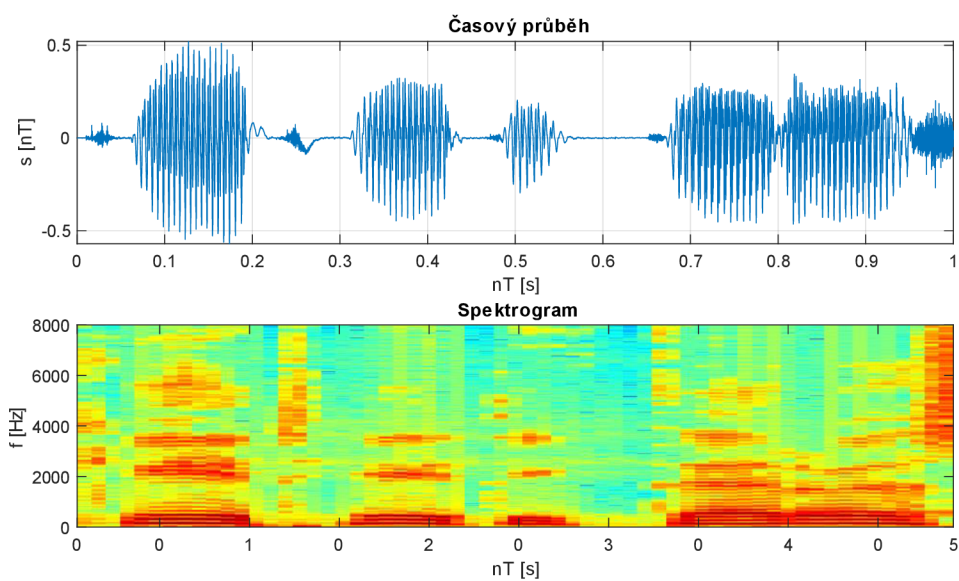
SEZNAM PŘÍLOH

Příloha A: Časový průběh a spektrogram úlohy čtení textu (HU).....	65
Příloha B: Časový průběh a spektrogram úlohy prodloužené fonace (HU).....	67
Příloha C: Hodnoty SNR (CZ)	69
Příloha D: Hodnoty SNR (HU).....	73
Příloha E: Struktura kódu	74

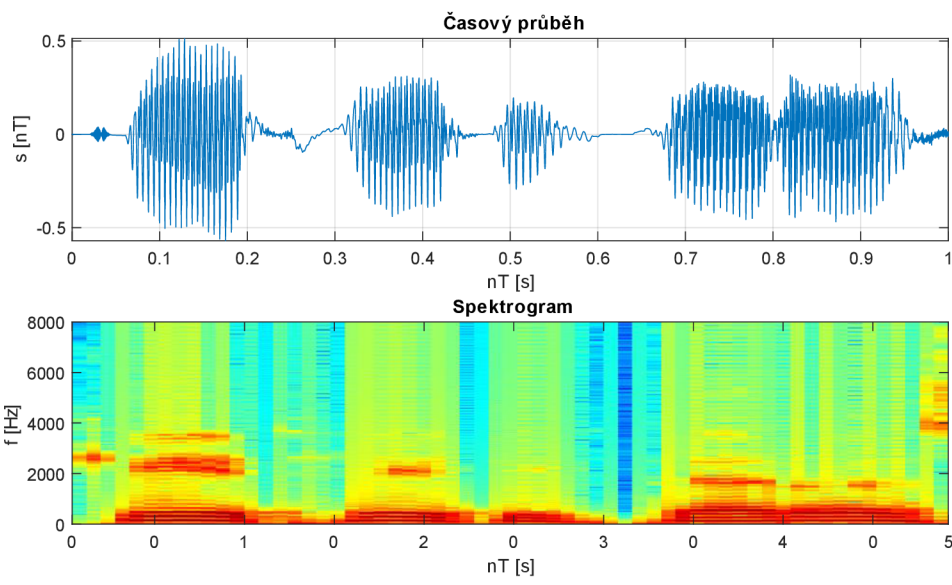
PŘÍLOHA A: ČASOVÝ PRŮBĚH A SPEKTROGRAM ÚLOHY ČTENÍ TEXTU (HU)



Řečový signál maďarské zašuměné nahrávky čtení textu

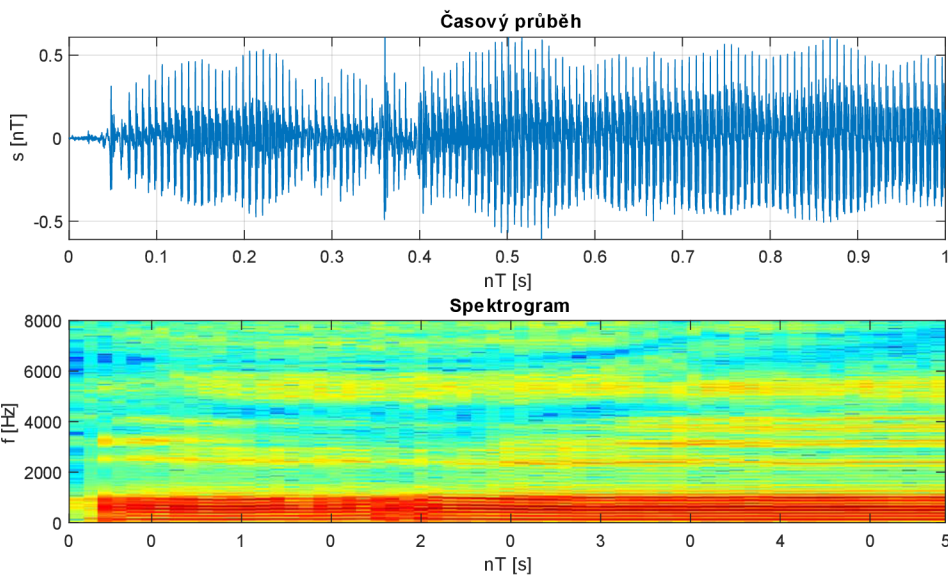


Řečový signál maďarské nahrávky čtení textu zvýrazněné spektrálním odečítáním

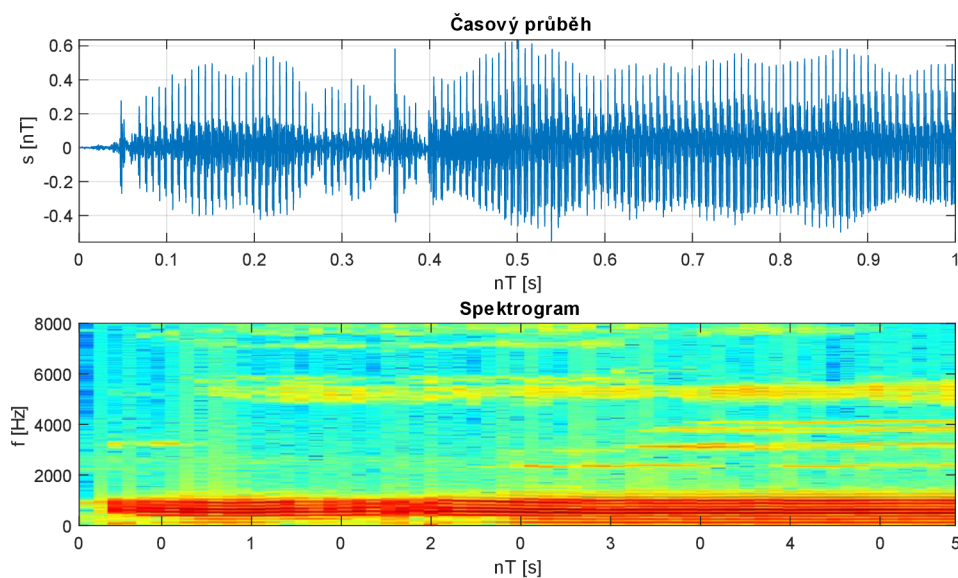


Řečový signál maďarské nahrávky čtení textu zvýrazněné Kalmanovým filtrem

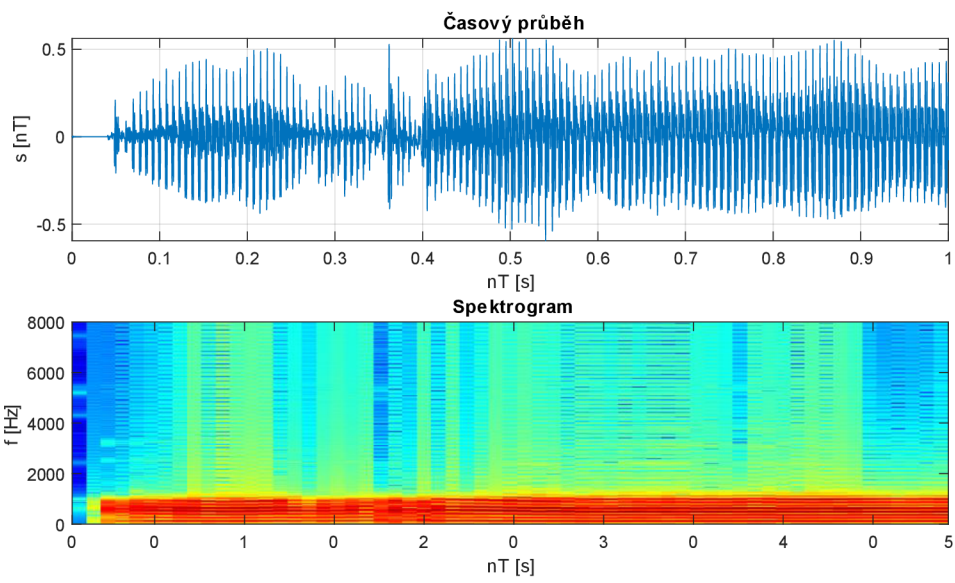
PŘÍLOHA B: ČASOVÝ PRŮBĚH A SPEKTROGRAM ÚLOHY PRODLOUŽENÉ FONACE (HU)



Řečový signál maďarské zašuměné nahrávky prodloužené fonace



Řečový signál maďarské nahrávky prodloužené fonace zvýrazněné spektrálním odečítáním



Řečový signál maďarské nahrávky prodloužené fonace zvýrazněné Kalmanovým filtrem

PŘÍLOHA C: HODNOTY SNR (CZ)

Tabulka hodnot SNR českých nahrávek s aditivním šumem o odstupu od řečového signálu 20 dB

TSK2			TSK6		
C	S	K	C	S	K
19,89	16,42	14,92	20,00	3,65	15,62
20,01	12,99	13,70	20,02	1,17	13,62
19,79	6,19	12,06	19,93	1,25	12,32
20,05	14,22	14,49	19,75	0,99	15,83
19,88	13,99	14,69	19,84	2,50	17,24
20,03	19,69	13,78	19,99	24,72	15,50
20,05	16,37	14,38	20,04	2,97	14,09
20,04	14,80	13,38	19,87	8,69	10,92
20,04	17,17	14,99	20,04	1,17	15,36
19,95	17,98	14,44	19,78	2,47	15,98
19,92	17,91	12,55	19,97	4,42	14,70
20,01	18,07	14,78	20,05	9,82	15,10
20,04	17,69	14,67	19,87	0,70	15,92
19,97	12,25	13,81	19,91	10,52	15,66
20,05	14,59	14,07	19,84	-1,48	15,93
20,00	15,81	13,71	19,91	1,18	11,49
20,03	15,25	15,52	19,87	9,97	17,71
19,94	17,70	13,74	20,01	11,89	15,60
20,03	15,62	14,83	19,94	1,63	16,07
19,99	16,85	14,12	20,05	5,59	16,70
20,02	16,08	14,81	19,24	11,98	16,21
20,03	17,91	13,59	19,74	15,27	17,11
20,04	16,55	13,88	19,96	9,68	14,88
19,86	18,56	15,20	19,91	-6,26	16,80
20,00	18,73	14,81	19,84	14,12	17,42
20,05	15,90	13,57	19,88	7,86	15,05
20,01	18,88	14,59	20,01	28,20	15,72
20,05	17,95	15,06	19,06	7,94	16,59
19,92	19,98	14,42	20,08	19,77	17,49
20,04	19,51	14,38	19,83	23,18	14,86
20,04	15,70	15,02	19,54	0,53	16,42
20,02	16,99	14,20	19,95	7,42	10,84
20,03	19,37	15,79	19,95	5,63	12,14
20,03	11,18	15,40	19,70	9,90	12,29
19,95	16,32	16,61	19,97	-2,31	12,48
20,04	19,65	14,78	19,71	23,16	15,11
19,99	20,27	14,29	19,71	8,99	13,85
20,03	16,41	15,06	20,04	16,54	16,66
20,01	16,63	16,11	20,04	3,74	17,49
19,99	14,30	15,93	19,54	3,05	16,88
20,02	18,85	14,06	20,02	4,93	14,56
19,98	15,39	13,24	19,93	-3,12	13,99
19,97	15,63	15,60	20,00	8,07	14,48
20,04	18,11	15,90	19,02	28,10	16,42
20,04	11,67	15,38	19,00	-1,48	15,66
19,74	18,96	14,46	19,98	2,09	15,23
19,88	19,47	15,39	19,91	5,77	15,61
19,87	18,00	14,18	19,59	-1,43	17,25
19,99	19,15	16,19	19,66	-4,69	16,49

20,04	14,45	13,66	19,78	1,14	10,59
19,79	15,45	12,47	20,03	-8,78	16,97
20,04	17,66	14,25	19,90	-3,04	13,59
20,05	17,54	15,27	19,93	0,30	17,71
19,82	18,41	14,10	19,53	10,55	16,88
20,03	18,97	15,88	19,99	0,15	16,46
20,03	17,58	13,00	20,02	0,39	16,14
20,02	17,60	14,19	20,00	-4,42	15,01
20,01	18,92	15,10	19,86	14,59	16,07
19,94	11,90	13,31	20,06	-4,11	14,82
20,01	15,28	15,88	20,05	-0,52	15,35
20,03	19,89	16,00	19,74	24,60	17,35
20,01	15,45	15,35	20,04	12,18	15,62
20,01	18,59	14,83	20,00	3,89	15,43
19,94	8,86	15,21	20,04	2,67	13,92
19,98	9,19	15,74	20,05	-1,77	14,11
19,99	14,06	14,06	20,03	4,27	14,28
20,04	13,30	15,08	19,80	-1,41	15,48
20,02	10,83	14,65	20,00	-1,31	14,38
19,93	8,13	14,94	19,71	0,44	15,42
20,04	13,89	15,37	19,49	13,11	16,97
20,03	14,39	15,24	19,98	3,86	13,79
20,04	12,50	15,78	20,05	1,72	15,59
20,02	10,91	16,31	19,89	4,34	13,90
19,99	10,43	14,37	20,06	2,91	14,97
19,84	13,97	15,54	20,01	7,18	14,19
19,53	11,73	16,78	20,04	4,45	15,32
20,03	14,19	12,05	19,86	5,34	13,78
20,05	11,86	14,91	19,73	7,67	16,69
20,03	12,21	14,28	20,02	6,02	15,84
20,05	10,89	14,74			
20,05	12,38	15,48	20,02	5,61	14,86
19,88	11,62	14,53	19,69	3,44	15,64
20,02	8,80	15,18	20,00	6,84	15,57
20,01	16,59	14,26	20,06	23,62	14,58
19,99	14,80	14,47	20,00	14,11	13,88
20,04	16,71	13,63	20,05	20,56	12,49
20,04	13,93	14,23	19,93	10,74	16,63
20,04	12,45	12,80	20,00	1,63	12,84
19,84	17,41	14,79	20,03	6,34	13,86
20,02	15,17	15,50	20,02	5,19	13,45
20,04	16,80	14,84	20,04	0,58	13,00
20,00	12,09	15,24	20,04	-1,64	14,53
19,97	14,29	15,38	20,03	7,76	15,24
20,04	12,71	16,44	19,98	7,46	14,48
20,02	11,18	15,89	19,85	-5,29	14,14
19,98	14,40	15,73	20,05	3,13	18,25
19,97	14,82	15,39	19,52	3,54	13,88
20,04	12,68	15,43	19,49	5,14	15,49
19,98	14,42	14,90	20,04	3,24	12,88
20,04	8,33	15,38	20,03	3,75	12,75
19,80	9,90	14,97	19,65	6,13	16,33
20,02	12,88	15,78	20,02	-1,15	15,76
20,03	14,42	15,74	19,77	3,42	16,07
20,03	16,82	14,93	19,98	19,25	14,20

19,47	14,71	16,26	19,95	6,39	14,18
20,05	14,22	15,34	20,00	7,13	15,24
20,02	15,07	15,16	20,01	9,74	17,07
20,02	12,50	12,91	20,03	9,60	15,92
19,94	10,01	15,16	20,06	9,66	16,36
20,02	13,31	12,68	19,82	4,63	13,67
19,97	8,75	13,16	19,96	7,75	12,00
19,78	16,64	13,96	20,02	5,21	15,78
19,94	12,26	13,67	20,03	4,23	16,00
20,02	19,56	14,65	20,04	16,90	14,01
19,98	14,03	15,76	20,02	7,38	15,25
20,03	14,27	15,61	19,69	2,73	15,36
20,03	15,26	15,31	20,03	8,75	16,35
20,05	16,76	15,46	19,98	28,15	15,73
20,00	12,59	15,47	19,93	4,94	14,17
20,00	17,99	14,29	20,05	16,81	13,36
20,03	12,32	15,59	20,06	8,52	13,54
20,04	13,53	14,19	19,94	4,05	15,80
20,00	16,82	14,00	20,05	9,68	14,62
19,99	16,18	15,75	20,05	4,68	15,10
20,03	12,42	15,57	19,46	-0,75	14,92
20,03	14,11	15,19	20,04	2,13	12,49
20,02	13,17	16,22	19,91	11,07	15,41
19,97	13,37	17,38	20,04	4,93	15,58
20,01	16,50	12,72	19,99	11,10	15,86
20,04	11,62	14,42	20,03	3,87	14,46
20,03	16,70	15,99	20,06	19,67	16,63
19,97	11,16	16,06	20,07	0,36	13,34
20,03	11,28	13,50	19,99	14,15	16,48
20,01	10,74	15,11	20,04	5,62	13,49
19,99	14,07	14,15	20,02	2,01	14,69
19,97	10,87	14,90			
20,05	16,19	12,99	19,94	26,89	15,82
19,66	11,37	16,99	19,69	19,04	15,73
19,92	11,39	16,67	19,68	2,44	14,47
19,65	8,37	14,96	20,03	-3,86	13,30
19,84	10,92	17,51	19,99	-0,76	15,62
19,56	8,13	15,92	19,96	-3,02	14,52
19,92	11,23	18,85	19,85	-4,67	18,22
19,98	14,82	13,28			
20,01	14,20	14,19	19,98	4,50	13,72
20,04	16,63	13,37	19,80	5,34	15,64
20,05	14,20	13,44	19,79	4,39	14,54
20,02	9,40	14,70	19,97	8,00	13,68
20,04	15,07	13,95	20,00	13,22	14,05
19,98	13,56	15,34	19,89	5,45	15,35
20,00	15,05	15,27	19,82	7,59	15,52
20,00	16,91	14,13	20,03	10,85	10,71
20,02	14,44	15,32	19,71	11,51	14,73
19,92	10,26	14,42	19,77	11,75	12,25
20,05	14,88	16,03	20,02	6,45	14,96
19,83	15,53	17,20	20,01	7,08	14,21
20,01	15,70	15,37	19,98	9,37	14,91
19,85	17,68	14,96	19,80	7,54	14,06
20,03	15,88	14,19	19,70	18,64	18,00

19,86	17,67	15,67	20,04	11,60	12,60
20,02	12,07	16,26	20,02	5,26	17,13
20,01	15,71	16,72			
20,04	15,46	15,31	20,04	30,62	16,48
20,03	18,58	16,14	19,97	33,04	17,53
20,05	14,42	16,05	19,84	14,42	15,12
20,04	16,05	15,84	19,80	2,73	15,51
19,89	10,72	15,66	19,81	5,85	15,62
20,04	15,32	15,45	19,97	8,47	14,33
20,00	16,08	15,93	20,01	7,25	13,83
20,01	11,88	14,78	19,77	5,64	15,15
19,96	15,09	15,46	20,00	8,61	13,59
20,03	11,93	14,28	20,06	-0,06	14,49
19,93	11,45	15,14	19,95	1,85	13,86
19,96	13,74	12,72	19,96	4,14	13,38
20,03	13,94	13,77	19,82	-0,38	14,81
20,02	17,68	15,40	19,99	16,02	15,44
20,04	14,53	17,77	19,25	8,14	14,78
20,05	14,82	16,77	20,04	12,79	17,27
19,91	15,28	17,76	20,06	4,51	16,07
19,97	15,39	14,95	19,97	7,81	15,30
20,03	16,63	15,72	20,06	15,29	14,02
19,94	16,81	16,58	19,57	20,30	15,31
20,04	14,65	15,08			
			20,03	15,22	14,33
20,04	11,58	17,40	20,04	4,39	13,76
20,04	15,61	16,16	20,04	13,80	15,69
20,00	16,96	15,44	20,02	10,97	15,02
20,01	11,37	16,95			
20,03	16,82	16,50	19,99	25,52	15,74
20,02	13,42	15,93	20,01	16,23	15,72
20,02	13,61	14,85	20,02	5,82	16,12
20,04	11,19	14,68	20,04	14,62	14,28
20,04	14,40	14,04	19,39	7,79	15,63
20,01	15,40	14,77	19,75	9,97	15,98
20,01	14,87	17,29	20,01	2,99	14,15
20,04	14,91	14,94	19,97	21,63	12,76
20,04	14,83	14,50	19,89	0,09	12,28
20,05	17,84	15,57	19,98	21,47	14,64
20,04	17,58	14,79	20,08	9,51	14,18
20,05	13,81	16,26	20,04	6,64	15,23
19,99	16,31	14,86	20,04	8,39	13,80
20,04	14,51	14,45	19,91	8,34	14,48
20,04	17,25	15,69	19,76	16,35	16,33
19,99	14,00	15,65	19,92	1,13	16,00
20,04	17,59	13,95	19,98	29,23	15,21
20,03	7,30	13,05	20,04	7,44	16,02
20,04	16,45	16,24	20,03	-1,38	14,42
20,04	17,91	14,68	20,00	12,94	11,78
20,02	13,05	14,99	19,90	0,33	14,53
20,04	15,80	16,08	19,99	4,34	15,35
20,04	15,97	14,81	19,69	22,30	15,20
19,62	16,20	16,21	19,61	17,24	14,52
19,98	17,78	15,74	19,78	6,00	17,07
20,05	13,77	14,75	19,99	14,88	11,84

PŘÍLOHA D: HODNOTY SNR (HU)

Tabulka hodnot SNR českých nahrávek s aditivním šumem o odstupe od řečového signálu 20 dB

TSK2		TSK6	
S	K	S	K
19,92	14,75	-3,07	17,49
28,33	15,56	10,87	18,15
19,28	16,66	9,15	19,32
		14,17	20,31
14,55	16,06	15,51	18,63
21,28	16,26	16,48	20,14
26,08	16,29	11,60	17,95
21,19	14,23	5,65	21,14
18,59	14,31	1,26	15,55
27,47	15,59	18,21	16,89
29,54	16,09	19,61	16,76
29,28	15,16	13,84	21,14
23,32	14,17	12,68	17,39
25,49	16,26	13,51	21,36
17,84	14,68	8,58	19,75
17,87	15,22	10,79	20,58
19,86	13,44	12,92	18,80
28,64	14,87	19,02	21,10
21,27	15,33	7,91	13,87
31,27	16,46	20,98	18,04
25,59	17,02	23,35	19,31
31,31	16,31	30,51	18,78
25,14	15,68	16,77	18,15

PŘÍLOHA E: STRUKTURA KÓDU

