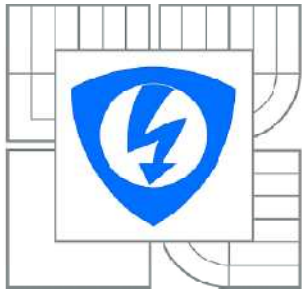


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ

ÚSTAV RADIOELEKTRONIKY

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF RADIO ELECTRONICS

METODY PRO ODSTRANĚNÍ RUŠIVÉHO POZADÍ Z AUDIO SIGNÁLU

METHODS FOR REMOVING NOISE BACKGROUND FROM AUDIO SIGNAL

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

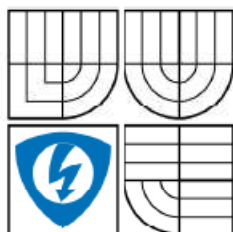
DAVID RUMAN

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. KAREL POVALAČ

BRNO 2010



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav radioelektroniky

Bakalářská práce

bakalářský studijní obor
Elektronika a sdělovací technika

Student: David Ruman
Ročník: 3

ID: 98427
Akademický rok: 2009/2010

NÁZEV TĚMATU:

Metody pro odstranění rušivého pozadí z audio signálu

POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s problematikou odstranění rušivého šumu z audio signálu. Nalezněte metody, které by se daly za tímto účelem použít. Vypracujte studii s teoretickým popisem metod a srovnajte je simulacemi v prostředí MATLAB.

Praktické provedení odstranění rušivého pozadí proveďte na reálném audio signálu, který je rušen hlukem automobilu nebo větru. Získané výsledky jednotlivých metod srovnajte.

DOPORUČENÁ LITERATURA:

[1] JAN, J. Číslíková filtrace, analýza a restaurace signálů. Brno: VUTIUM Press Brno, 2002. 427 s. ISBN: 80-214-1558-4.

Termín zadání: 8.2.2010

Termín odevzdání: 28.5.2010

Vedoucí práce: Ing. Karel Povalač

prof. Dr. Ing. Zbyněk Raida
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.

LICENČNÍ SMLOUVA POSKYTOVANÁ K VÝKONU PRÁVA UŽÍT ŠKOLNÍ DÍLO

uzavřená mezi smluvními stranami:

1. Pan/paní

Jméno a příjmení: David Ruman
Bytem: Lačnov 263, 756 12 Horní Lideč
Narozen/a (datum a místo): 30. března 1987 ve Vsetíně
(dále jen „autor“)

a

2. Vysoké učení technické v Brně

Fakulta elektrotechniky a komunikačních technologií
se sídlem Údolní 53, Brno, 602 00
jejímž jménem jedná na základě písemného pověření děkanem fakulty:
prof. Dr. Ing. Zbyněk Raida, předseda rady oboru Elektronika a sdělovací technika
(dále jen „nabyvatel“)

Čl. 1

Specifikace školního díla

1. Předmětem této smlouvy je vysokoškolská kvalifikační práce (VŠKP):

- disertační práce
 - diplomová práce
 - bakalářská práce
 - jiná práce, jejíž druh je specifikován jako
- (dále jen VŠKP nebo dílo)

Název VŠKP: Metody pro odstranění rušivého pozadí z audio signálu

Vedoucí/ školitel VŠKP: Ing. Karel Povalač

Ústav: Ústav radioelektroniky

Datum obhajoby VŠKP: _____

VŠKP odevzdal autor nabyvateli*:

- v tištěné formě – počet exemplářů: 2
- v elektronické formě – počet exemplářů: 2

2. Autor prohlašuje, že vytvořil samostatnou vlastní tvůrčí činností dílo shora popsané a specifikované. Autor dále prohlašuje, že při zpracovávání díla se sám nedostal do rozporu s autorským zákonem a předpisy souvisejícími a že je dílo dílem původním.
3. Dílo je chráněno jako dílo dle autorského zákona v platném znění.
4. Autor potvrzuje, že listinná a elektronická verze díla je identická.

* hodící se zaškrtněte

Článek 2

Udělení licenčního oprávnění

1. Autor touto smlouvou poskytuje nabyvateli oprávnění (licenci) k výkonu práva uvedené dílo nevýdělečně užít, archivovat a zpřístupnit ke studijním, výukovým a výzkumným účelům včetně pořizování výpisů, opisů a rozmnoženin.
2. Licence je poskytována celosvětově, pro celou dobu trvání autorských a majetkových práv k dílu.
3. Autor souhlasí se zveřejněním díla v databázi přístupné v mezinárodní síti
 - ihned po uzavření této smlouvy
 - 1 rok po uzavření této smlouvy
 - 3 roky po uzavření této smlouvy
 - 5 let po uzavření této smlouvy
 - 10 let po uzavření této smlouvy
(z důvodu utajení v něm obsažených informací)
4. Nevýdělečné zveřejňování díla nabyvatelem v souladu s ustanovením § 47b zákona č. 111/ 1998 Sb., v platném znění, nevyžaduje licenci a nabyvatel je k němu povinen a oprávněn ze zákona.

Článek 3

Závěrečná ustanovení

1. Smlouva je sepsána ve třech vyhotoveních s platností originálu, přičemž po jednom vyhotovení obdrží autor a nabyvatel, další vyhotovení je vloženo do VŠKP.
2. Vztahy mezi smluvními stranami vzniklé a neupravené touto smlouvou se řídí autorským zákonem, občanským zákoníkem, vysokoškolským zákonem, zákonem o archivnictví, v platném znění a popř. dalšími právními předpisy.
3. Licenční smlouva byla uzavřena na základě svobodné a pravé vůle smluvních stran, s plným porozuměním jejímu textu i důsledkům, nikoliv v tísní a za nápadně nevýhodných podmínek.
4. Licenční smlouva nabývá platnosti a účinnosti dnem jejího podpisu oběma smluvními stranami.

V Brně dne: 28. května 2010

.....
Nabyvatel

.....
Autor

ABSTRAKT

Cílem této bakalářské práce na téma “Metody pro odstranění rušivého pozadí z audio signálu” je seznámit se s jednotlivými metodami používanými pro odrušení řečového signálu. V první části bude postupně teoreticky rozebrána lineární filtrace a spektrální odečítání, v druhé části budou tyto metody aplikovány na reálný signál, detailně rozebrány a následně vyhodnoceny.

ABSTRACT

The aim of the bachelor work “Methods for removing interference noise from an audio signal” is to introduce individual methods used for dissolution a speech signal. The first part deals with linear filtration and spectral subtraction in a theoretical way. The second part applies these methods to a real signal, analyzes and consequently evaluates them.

RUMAN, D. - Metody pro odstranění rušivého pozadí z audio signálu. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav radioelektroniky, 2010. 45 s., 0 s. příloh. Bakalářská práce. Vedoucí práce: Ing. Karel Povalač.

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma *Metody pro odstranění rušivého pozadí z audio signálu* 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 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.

V Brně dne

.....

(podpis autora)

PODĚKOVÁNÍ

Děkuji vedoucímu bakalářské práce Ing. Karlovi Povalačovi za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé bakalářské práce.

V Brně dne

.....

(podpis autora)

OBSAH

Seznam obrázků	9
Seznam tabulek	10
Úvod	11
1 Zpracování signálu	12
1.1 Řečový signál.....	12
1.1.1 Charakter řečového signálu	12
1.1.2 Model vytváření řeči.....	14
1.2 Kepstrální analýza.....	15
1.3 Časově - frekvenční analýza	16
2 Metody odrušení řečového signálu	18
2.1 Detekce řečové aktivity	18
2.2 Lineární filtrace.....	18
2.2.1 Filtry FIR	19
2.3 Koncept adaptivní filtrace.....	20
2.4 Spektrální odečítání	21
2.4.1 Základní algoritmus	21
2.4.2 Zlepšení základního algoritmu.....	22
3 Použití jednotlivých metod v praxi	24
3.1 Lineární filtrace.....	24
3.1.1 Návrh filtru metodou oken pomocí Matlabu	24
3.1.2 Ruční návrh filtru metodou oken.....	30
3.2 Spektrální odečítání	35
4 Závěr	43
Literatura	45

SEZNAM OBRÁZKŮ

Obr. 1 : Časový průběh slova AFÉRA	13
Obr. 2 : Znělá část slova AFÉRA	13
Obr. 3 : Neznělá část slova AFÉRA	14
Obr. 4 : Model vytváření řeči.....	14
Obr. 5 : Blokové schéma detektoru řeči	18
Obr. 6 : Amplitudová frekvenční charakteristika filtrů	19
Obr. 7 : Obecná struktura adaptivního filtru.....	20
Obr. 8 : Blokové schéma spektrálního odečítání	21
Obr. 9 : Modulová frekvenční charakteristika pásmové propusti s Pravoúhlým oknem o různých délkách.	24
Obr. 10 : Modulová frekvenční charakteristika pásmové propusti s Hammingovým oknem o různých délkách.	25
Obr. 11 : Modulová frekvenční charakteristika pásmové propusti s Hannovým oknem o různých délkách.	25
Obr. 12 : Modulová frekvenční charakteristika pásmové propusti s Bartlettovým oknem o různých délkách.	25
Obr. 13 : Modulová frekvenční charakteristika pásmové propusti s Pravoúhlým oknem o různých délkách.	26
Obr. 14 : Fázová frekvenční charakteristika pásmové propusti s Hammingovým oknem o různých délkách.	26
Obr. 15 : Fázová frekvenční charakteristika pásmové propusti s Hannovým oknem o různých délkách.	26
Obr. 16 : Fázová frekvenční charakteristika pásmové propusti s Bartlettovým oknem o různých délkách.	27
Obr. 17 : Modulová charakteristika jednotlivých oken v časové oblasti.....	31
Obr. 18 : Modulová charakteristika jednotlivých oken ve frekvenční oblasti.....	31
Obr. 19 : Impulzní odezva filtru	32
Obr. 20 : Modulová frekvenční charakteristika filtru s ukázkou všech oken.....	32
Obr. 21 : Fázová frekvenční charakteristika filtru s ukázkou všech oken.....	33
Obr. 22 : Impulzní odezvy při různých metodách návrhu filtru	34
Obr. 23 : Princip plnění segmentů	35
Obr. 24 : Výkonová charakteristika segmentů při použití Pravoúhlého okna	36
Obr. 25 : Výkonová charakteristika segmentů při použití Hammingova okna	36

Obr. 26 : Výkonová charakteristika segmentů při použití Hannova okna.....	37
Obr. 27 : Výkonová charakteristika segmentů při použití Bartlettova okna	37
Obr. 28 : Spektrum hluku	38
Obr. 29 : Skládání výsledného signálu bez překrývání segmentů	38
Obr. 30 : Skládání výsledného signálu při 25% překrývání segmentů	39
Obr. 31 : Skládání výsledného signálu při 50% překrývání segmentů	39
Obr. 32 : Skládání výsledného signálu při 75% překrývání segmentů	40

SEZNAM TABULEK

Tab. 1 : Vyhodnocení odrušení pomocí filtru navrženého metodou oken.....	29
Tab. 2 : Vyhodnocení odrušení pomocí ručně spočítaného filtru.....	33
Tab. 3 : Vyhodnocení spektrálního odečítání s Pravoúhlým násobícím oknem.....	41
Tab. 4 : Vyhodnocení spektrálního odečítání s Hammingovým násobícím oknem	41
Tab. 5 : Vyhodnocení spektrálního odečítání s Hannovým násobícím oknem	42
Tab. 6 : Vyhodnocení spektrálního odečítání s Bartlettovým násobícím oknem	42

ÚVOD

Řeč je nejdůležitější a nejpřirozenější způsob komunikace mezi lidmi. Často však obsahuje nežádoucí složky, které můžou řeč částečně, či úplně zarušit. Existuje spousta rušivých elementů a každý může mít jiný charakter rušení. Proto se používají jiné metody pro odstranění stacionárního rušivého signálu a jiné pro rychle se měnící signály. Některé z metod se dají sice použít na více typů rušení, ovšem účinnost odrušení se často výrazně liší.

Řečové signály se mohou zpracovávat analogově nebo digitálně. Číslicové zpracování signálů může dále probíhat přímo v časové oblasti, úpravou posloupnosti vzorků, nebo ve frekvenční oblasti, úpravou spektra v technikách krátkodobé analýzy.

Úkolem této bakalářské práce bude se zabývat metodami pro odrušení audio signálu, který bude zarušen dvojím typem hluku – zvukem motoru automobilu a hlukem, jež způsobuje vítr. Detailně bude rozebrán samotný řečový signál a jeho zpracování v časové a frekvenční oblasti. Dále budou teoreticky zpracovány metody pro odrušení řeči založené na lineární filtraci a spektrálním odečítání a následně budou obě metody aplikovány na reálnou nahrávku. Závěrem bude popsáno nastavení parametrů v jednotlivých metodách a zhodnotí se možné použití v praxi.

1 ZPRACOVÁNÍ SIGNÁLU

Aby bylo možné pracovat s audio signálem, je třeba si ujednotit jednotlivé analýzy, kterými bude audio signál zkoumán. Je třeba si také uvědomit, že signál většinou obsahuje nejen spektrum řečového signálu ale i další spektra, jenž jsou často nežádoucí a aby mohlo dojít k jejich odstranění, je nutné znát charakter jednotlivých spekter, který dokážeme získat pomocí níže uvedených metod.

1.1 Řečový signál

1.1.1 Charakter řečového signálu

Lidskou řečí se rozumí časově proměnný souvislý proces, který je nejvíce ovlivňován:

- akustickou strukturou (velikostí signálu, spektrem)
- lingvistickou strukturou (jazyk, dialekt, akcent)
- osobností mluvčího (barva hlasu, intonace, rytmus)
- vliv pozadí a prostředí jimž se zvuk šíří

Takto vzniká u každého člověka různorodá řeč, která je dále přenášena prostředím pomocí akustického vlnění.

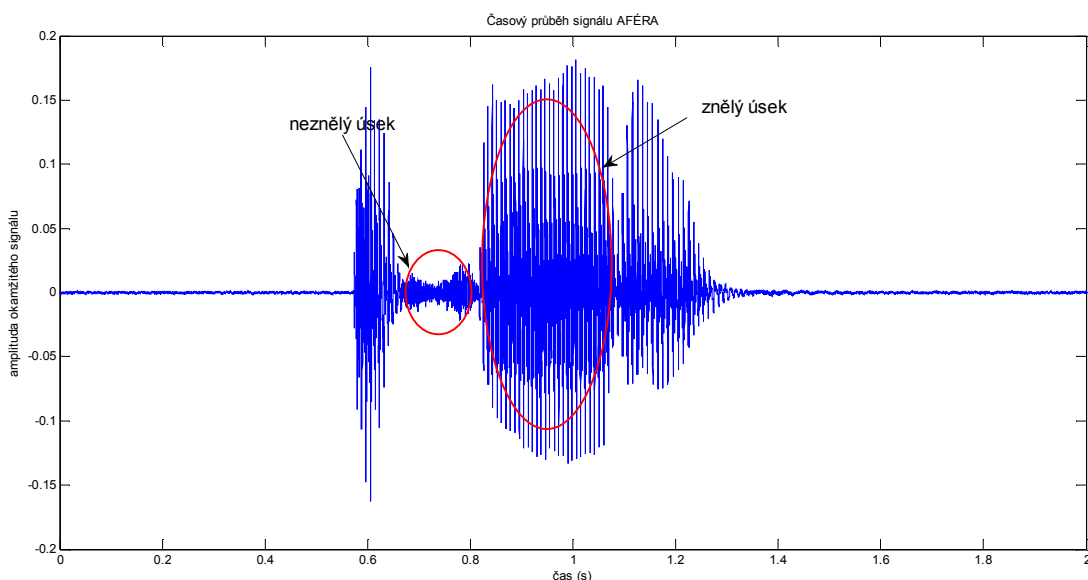
Řeč tvoří segmenty, které na sebe navzájem navazují a vznikají změnami překážek v hlasovém traktu. Tyto segmenty jakožto nejmenší foneticky odlišné jednotky, se nazývají *fonémy*. Z fonémů pak skládáme jednotlivé slabiky a z těch následně celá slova. Hlasovým projevem fonému je hláska. V češtině rozeznáváme dva druhy hlásek – samohlásky a souhlásky. Na základě rozpoznávání jednotlivých fonémů, může teprve dojít k rozpoznání řeči.

Pokud vezmeme v úvahu průběh vzniku řeči, je nutné si uvědomit, že jednotlivá slova vznikají změnou hlasového traktu, přičemž tato změna není skoková, kvůli síle svalů a setrvačnosti při přechodu traktu z jednoho stavu do druhého. Toto způsobuje tzv. *koartikulaci*, což je vzájemné ovlivňování předchozí na následující hlásce. Hláska tak zní díky okolním hláskám pokaždé jinak. Koartikulace dala vzniknout tedy dalším jednotkám, které toto můžou popsat. Známé jsou především *difóny* (posloupnost samohláska - souhláska) a *trifóny* (závislost na levém a pravém sousedovi) . Jejich výhodou je právě postih koartikulace, ovšem nevýhodou je pak jejich velký počet [1].

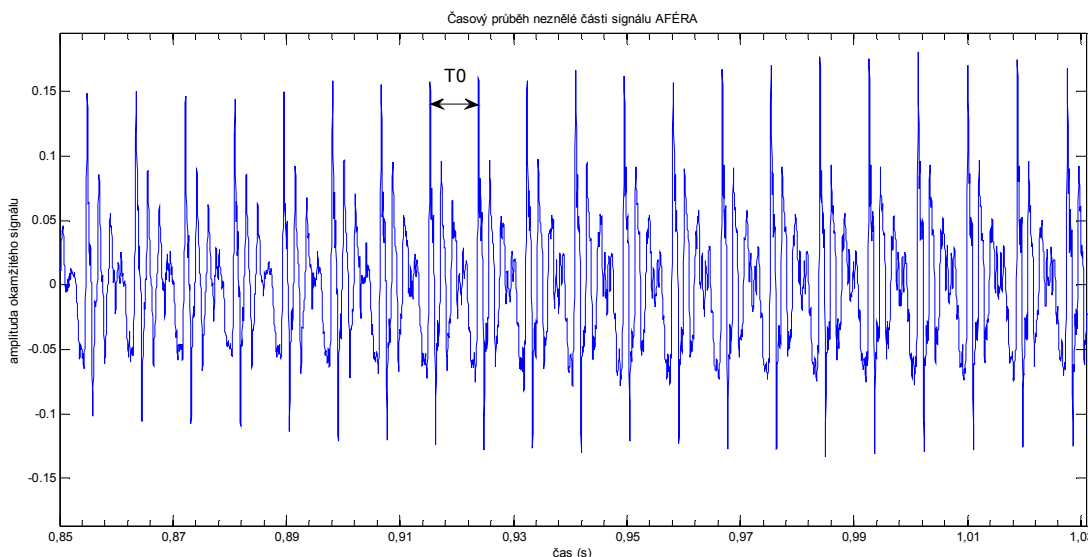
Řečový signál $s(t)$ je posloupnost diskrétních vzorků a jak můžeme vidět na obr.1 je složen ze dvou typů signálů:

- a) *Periodický signál* – představuje znělé části textu a jeho hlavním parametrem je základní perioda řeči (pitch period), která je značena T_0 .
- b) *Zašumělý signál* – oblasti s charakterem šumu představují neznělé části řeči.

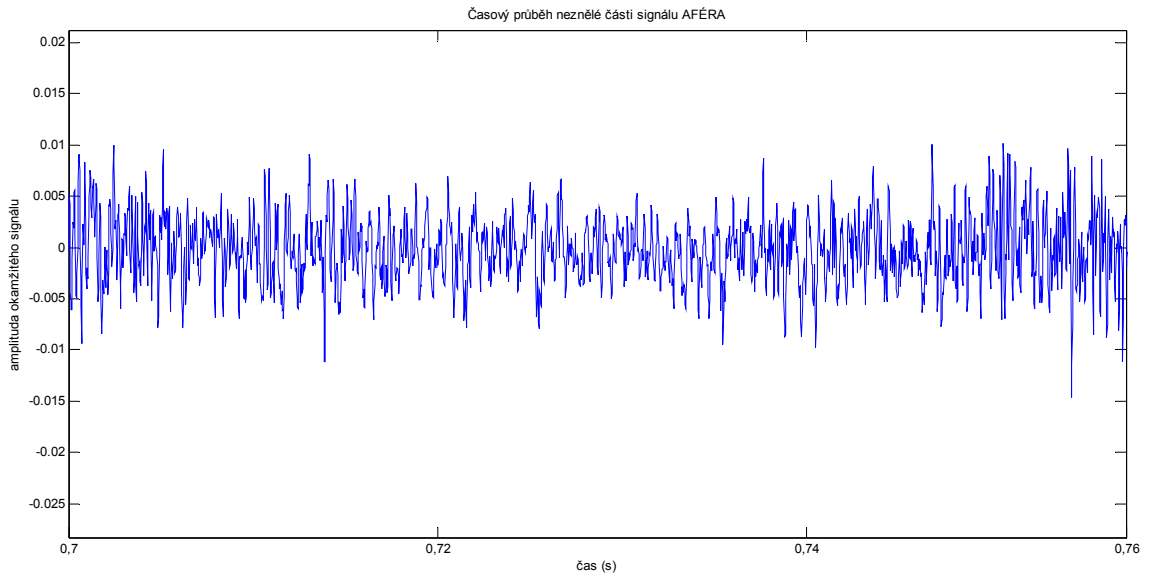
Kmitající hlasivky se dají pokládat za tvůrce všech znělých zvuků. Jestliže člověk mlčí, pak hlasivkovou štěrbinou proudí vzduch bez odporu. Při promluvě se hlasivky svírají a roztahují, čímž vzniká základní tón lidského hlasu F_0 – odpovídá kmitočtu v rozmezí 500 – 2000 Hz a je základem pro všechny znělé zvuky. Znělost, popř. neznělost, je způsobena tím, jestli dojde, popř. nedojde, k rozkmitání hlasivkové štěrbinou výdechovým proudem vzduchu vytlačovaného z plic. [2]



Obr. 1 : Časový průběh slova AFÉRA



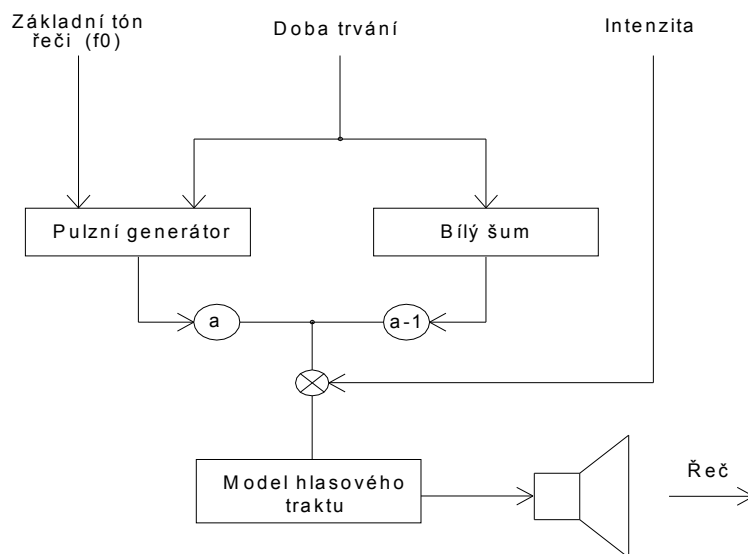
Obr. 2 : Znělá část slova AFÉRA



Obr. 3 : Neznělá část slova AFÉRA

1.1.2 Model vytváření řeči

Podle předchozích úvah lze sestavit jednoduchý model vytváření řeči, na který můžeme nahlížet jako na zjednodušený hlasový trakt. Základem je co nejjednodušší funkčnost, přičemž model funguje na základě několika mála parametrů – koeficientů, které popisují hlasové ústrojí člověka. Jelikož se jednotlivé části hlasového traktu v závislosti s časem mění, musí být model nestacionární. Díky pomalé změně jednotlivých ústrojí, můžeme říct, že se tato změna během 10 – 30ms neprojeví, a proto může mít model v tomto časovém intervalu konstantní parametry. Ovšem každý segment již má parametry jiné. [3]



Obr. 4 : Model vytváření řeči

Existují základní dva principy vytváření řeči. Jestliže model bude vytvářet znělý segment řeči, je využíváno pulzního generátoru, jehož parametrem je základní hlasivková frekvence f_0 .

$$T_0 = \frac{1}{f_0}, \quad (1.1.1) [3]$$

Neznělé segmenty jsou vytvářeny vybuzením generátoru bílého šumu s plochým výkonovým spektrem. Ve skutečnosti se vždycky nejedná jen o úplně znělé, či neznělé segmenty. Toho je možno docílit např. pomocí přimíchávání barevného šumu s definovaným výkonovým spektrem ke generátoru impulzů. [4]

1.2 Kepsrální analýza

Pomocí diskretní Fourierovy transformace DFT lze určit diskretní spektrum zkoumaného signálu. Budeme brát v úvahu, že konečný úsek N vzorků diskretního signálu

$$\{f_n = f(nT)\} \quad n = 0, 1, \dots, N-1, \quad (1.2.1) [5]$$

může být vyjádřen v kvazispojité reprezentaci jako součet vážených posunutých Diracových distribucí

$$f_s = \sum_{n=0}^{N-1} f_n \delta(t - nT), \quad (1.2.2) [5]$$

jeho spektrum ve smyslu integrální Fourierovy transformace tedy je

$$\begin{aligned} F_s(\omega) &= F\left\{\sum_{n=0}^{N-1} f_n \delta(t - nT)\right\} = \sum_{n=0}^{N-1} f_n F\{\delta(t - nT)\} = \\ &= \sum_{n=0}^{N-1} f_n e^{-j\omega nT} = DTFT\{f_n\}. \end{aligned} \quad (1.2.3) [5]$$

Navzorkujeme-li toto spojité spektrum, které je ovšem periodické s periodou $2\pi/T$, N vzorky v jedné periodě dostaneme

$$\{F_k\} = \{F_s(k\Omega)\} = \left\{\sum_{n=0}^{N-1} f_n e^{-jk\Omega nT}\right\} = DFT\{f_n\}. \quad (1.2.4) [5]$$

Můžeme tedy konstatovat, že diskretní Fourierova transformace poskytuje přesné vzorky konečného spektra a navzorkovaného úseku analyzovaného signálu. [6]

Máme-li časově neomezený spojité signál, je potřeba jej v první řadě zkrátit na konečnou délku a to vynásobením obdélníkovým oknem o délce NT (okno však může mít i jiný tvar). Ve spektrální oblasti tím dojde ke konvoluci původního spektra se spektrem okna. Vzorkování, které lze také interpretovat jako násobení, tentokrát Diracových distribucí, má za následek periodizace spektra s periodou rovnou vzorkovacímu kmitočtu. Shrňme-li důsledky diskretizace analyzovaného signálu, můžeme konstatovat následující :

- dochází ke snižování rozlišovací schopnosti ve spektru v důsledku omezení délky signálu, způsobujícího tzv. prosakování ve spektru. Můžeme si pomoci prodloužením okna, či vhodným zvolením tvaru okna.
- vzniká periodizace spektra v důsledku vzorkování signálu. Může dojít také ke zkreslení spektra tzv. aliasingem, pokud spektrum po předchozím kroku má významné složky nad Nyquistovým kmitočtem vzorkování. Náprava je možná především zvýšením vzorkovacího kmitočtu a dodržením vzorkovacího teorému nebo předzpracováním signálu antialiasingovým filtrem.

Při číslicovém výpočtu spektra je nutno navzorkovat také spektrum, což je vyjádřeno na straně spektra součinem s diskretizační posloupností impulzů. Při diskretizaci spektra podle definice DFT dochází ovšem k periodizaci signálu s periodou rovnou délce vstupního okna, takže původně nulovým úsekům vně okna jsou formálně vnuceny opakované hodnoty signálu uvnitř okna. Zpracovávaný úsek to však nepostihuje. Vzorkování spektra však může vézt k tomu, že zobrazený diskretní výsledek nedává dobrou informaci o spektru původního signálu. [7]

1.3 Časově - frekvenční analýza

Při analýze signálů, jejichž charakter se v čase rychle mění, tedy signálů přechodového charakteru, je často účelné uvažovat o frekvenčním obsahu krátkých signálových úseků, což znamená rozvinout koncept tzv. krátkodobých spekter. Tento koncept umožňuje formulovat obecněji spektrum jako dvojrozměrnou funkci, závislou nejen na frekvenci, ale i na pozici v čase. Ačkoliv integrální Fourierova transformace v čisté podobě pracuje se signály nekonečné délky, praktická analýza vždy vychází jen z konečných úseků signálu, vymezených použitým oknem. Pokud tedy okno má správnou délku, a je formulováno jako klouzavé na časové ose, může tento přístup být použit pro časově-frekvenční analýzu. Uvážíme-li, že postup, uvedený v předchozím paragrafu, také vycházel z konečných úseků signálu, můžeme zřejmě v zásadě téhož postupu použít i zde, i když v klasické frekvenční analýze jde o určení velikosti a fáze harmonických složek různých frekvencí a neomezeného trvání, nyní jde o nejpřesnější lokalizaci výskytu složek signálu jak pokud jde o frekvenční, tak i časový údaj. Z tohoto rozdílného přístupu také vyplývá odlišná metodyka návrhu pozorovacích intervalů. V klasické analýze bylo účelné tento interval v zájmu minimalizace zkreslení co nejvíce prodloužit, zatímco v časově-frekvenční analýze je pozorovací interval určen kompromisem mezi požadavkem na dostatečnou rozlišovací schopnost ve frekvenční oblasti (rozlišitelná diference frekvencí je nepřímo úměrná délce okna) a současnou snahou o velké rozlišení také v čase (minimální rozšiřitelný časový rozdíl je délce okna úměrný). [7],[8]

Jeden z těchto požadavků bývá v konkrétní praktické aplikaci určující a délku okna je pak třeba mu přizpůsobit. Použitím DFT na takto stanovené úseky signálu, tak vznikají zmíněná krátkodobá spektra, charakterizující frekvenční obsah i fázové poměry signálu v příslušném časovém úseku. Hodnoty krátkodobého spektra odpovídají celému časovému úseku okna a nelze již rozlišit, zda uvnitř tohoto intervalu došlo k nějakému vývoji – tím je v zásadě určena časová rozlišovací schopnost. Na druhé straně, délka

signálového okna určuje také frekvenční rozlišovací schopnost, která je zhruba N -tým dílem v rozsahu $\langle -\omega_s/2, \omega_s/2 \rangle$, a je tudíž konstantní v celém rozsahu zpracovávaných frekvencí.[7]

Krátkodobá spektra se zpravidla pořizují v celých sériích na základě signálových dat z delšího úseku signálu, v jehož rámci si přejeme sledovat vývoj frekvenčního obsahu. Přitom lze v nejjednodušším případě použít dělení posloupnosti signálových vzorků na úseky o délce N vzorků a z celkového úseku signálu o M vzorcích, který je pro analýzu k dispozici.[7] Přitom dosahujeme frekvenční rozlišovací schopnosti, odpovídající délce okna NT , a tento časový úsek znamená současně také nejmenší rozložitelný rozdíl v čase. Pokud jde časovou rozlišovací schopnost zvýšit tím, že dílčí okna mají zvolený přesah, např. o polovinu své délky. Pak dostaneme podél časové osy přiměřeně více spekter a lze lépe sledovat případný rychlý vývoj zejména na straně vysokých kmitočtů. Takový soubor spekter, tzv. spektrogram, může být názorně zobrazen jako dvojrozměrný obraz, v němž jedna souřadnice odpovídá frekvenci, druhá času a barva nebo úroveň jasu resp. šedi odpovídá amplitudě odpovídajících koeficientů spekter.

V prostém spektrogramu se ukazuje nevýhoda DFT, jejíž absolutní frekvenční rozlišovací schopnost je konstantní. Délka okna musí být tedy určena potřebnou frekvenční rozlišovací schopností na straně nízkých kmitočtů a omezuje tak časové rozlišení. Přitom by zpravidla pro praktické aplikace bylo užitečnější, kdyby byla konstantní relativní frekvenční rozlišovací schopnost, tj. kdyby rozlišitelný frekvenční interval byl jistým zlomkem z výchozí frekvence. Pak by ovšem na straně vysokých kmitočtů absolutní frekvenční rozlišení stačilo menší, k tomu by postačilo kratší okno a časové rozlišení by tedy u vyšších kmitočtů mohlo být detailnější. To vyplývá z toho, že složky vyšších kmitočtů lze zachytit a s potřebnou přesností identifikovat v kratší době.[9]

2 METODY ODRUŠENÍ ŘEČOVÉHO SIGNÁLU

2.1 Detekce řečové aktivity

Tento proces probíhá ve třech krocích, tak jako to můžeme vidět na obr.3 . Nejdříve je nalezena kepstrální reprezentace signálu, která umožňuje nejen zjednodušit náročnost algoritmu snížením objemu dat, ale také blíže specifikovat vstupní signál pomocí parametrů, což se může hodit k dalšímu zpracování. V další fázi dochází k rozpoznání vzdálenosti pro popis odlišnosti daného segmentu signálu od odhadu charakteristik pozadí řeči. Pro toto určení se vychází přímo z kepstrálních koeficientů nebo diferenčních kepstrálních koeficientů, které rovněž charakterizují řečový signál. Poslední fází je porovnání získané vzdálenosti s prahovou hodnotou pro konečné rozhodnutí o přítomnosti řeči a následné vyhlazení pro potlačení krátkých falešných skoků. [10],[11]



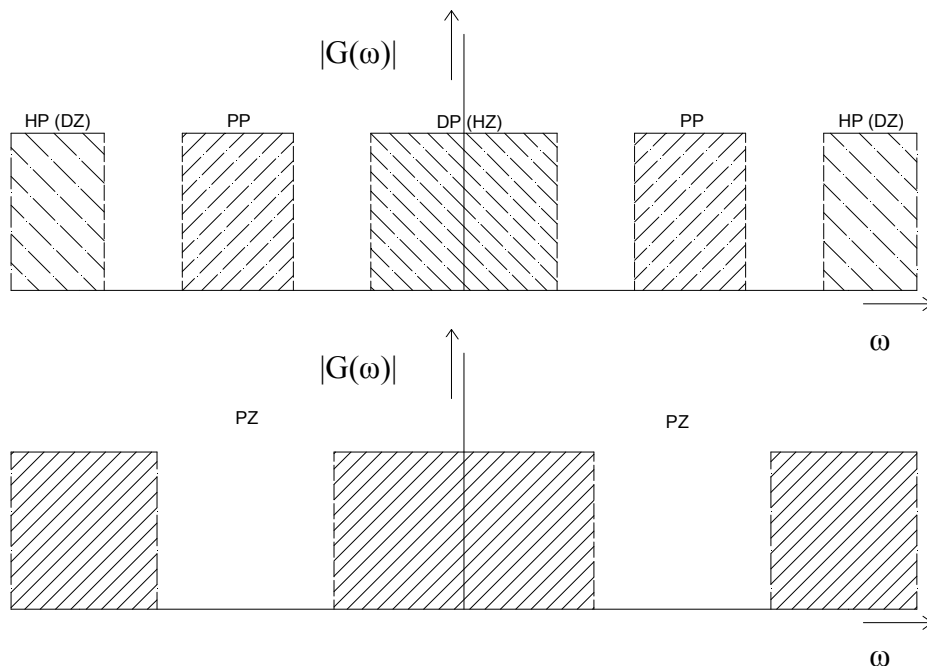
Obr. 5 : Blokové schéma detektoru řeči

2.2 Lineární filtrace

Proces, kdy dochází ke zpracování signálu, lze označit jako filtrace. Dochází zde k výběru patřičných složek ze směsi více signálů. Filtrace nám umožňuje měnit vlastnosti jednotlivých složek (např. poměrné zastoupení užitečné složky ve výsledném signále).

Složky signálu jsou většinou chápány ve frekvenční oblasti. Úpravu signálu vystihuje amplitudová a fázová frekvenční charakteristika. Z nich je zřejmá úprava amplitud a časových vztahů při filtračním procesu.[12] Jestliže budeme brát v potaz diskrétní signál, jsou tyto charakteristiky periodické, a proto se udávají jejich hodnoty pouze v rozsahu kmitočtů $< 0, \omega_s / 2 >$, kde ω_s je úhlový vzorkovací kmitočet.

Vyjímkou jsou ovšem pásmové filtry, které mají v ideálním případě v některých frekvenčních pásmech amplitudový přenos roven jedné a mimo tyto pásma je přenos nulový. Vlastnosti realizovatelných filtrů se od těch ideálních často liší, především ve strmosti náběžných a sestupných hran mezi propustnými a nepropustnými pásmy. Jestliže požadujeme jednotné zpoždění výstupu vůči vstupu, je žádoucí dodržet lineární fázovou charakteristiku. Tato podmínka ovšem nemusí být dodržena pro nepropustná pásma. [7],[13]



Obr. 6 : Amplitudová frekvenční charakteristika filtrů typu pásmová propust (PP) a pásmová zadrž (PZ), dolní propust (DP) – respektive horní zadrž (HZ) a horní propust (HP) – respektive dolní zadrž (DZ)

Filtry jsou tedy založeny na konvoluci mezi jejich impulzovou charakteristikou a vstupním signálem a lze je rovnocenně charakterizovat kteroukoliv z charakteristik .

2.2.1 Filtry FIR

FIR filtr je filtr s konečnou impulzní charakteristikou a je definován N hodnotami této charakteristiky, které tvoří vektor systémových konstant (koeficientů filtru)

$$\mathbf{h}=[h_n], \quad n \in \langle 0, N-1 \rangle, \quad (2.2.1), [7]$$

N odpovídá počtu vzorků. Filtr je založen na konečné diskrétní konvoluci

$$y_n = \sum_{k=0}^{N-1} x_{n-k} h_k . \quad (2.2.2), [7]$$

Pokud sumu rozepíšeme, získáváme diferenční rovnici filtru

$$y_n = h_0 x(n) + h_1 x(n-1) + \dots + h_n x(n-N) . \quad (2.2.3), [7]$$

Při hledání impulzní charakteristiky filtru je důležité přivést na vstup filtru diskrétní jednotkový impulz ζ_n , který bude mít tvar

$$\zeta_n = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}, \quad (2.2.4), [7]$$

což je i realizační algoritmus, ze kterého je zřejmé, že systémovými realizačními konstantami jsou hodnoty impulzní charakteristiky. Obrazový přenos je tedy

$$H(z) = \sum_{n=0}^{N-1} h_n z^{-n}, \quad (2.2.5), [14]$$

a reprezentace v rovině z je dána nulovými body.

Frekvenční charakteristika je periodická funkce s periodou $2\pi/T$, vyjádřená Furierovou řadou s koeficienty h_n . Návaznost systémových koeficientů s frekvenční charakteristikou umožňuje poměrně snadné navržení tohoto filtru.

$$G(\omega) = H(e^{j\omega T}) = \sum_{n=0}^{N-1} h_n e^{-j\omega n T}. \quad (2.2.6), [14]$$

Z konečného počtu složek uvedené řady vyplývá, že frekvenční charakteristika je spojitá funkce.

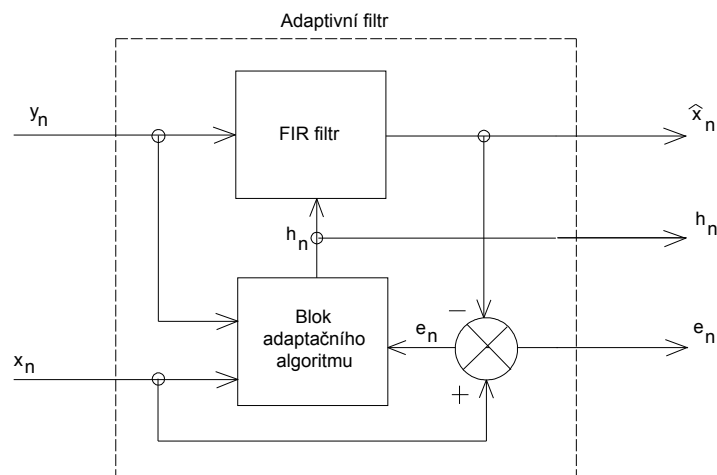
U filtrů typu FIR můžeme dosáhnout lineární fázové charakteristiky, jestliže bude platit

$$h_n = h_{(N-1-n)} \quad \text{nebo} \quad h_n = -h_{(N-1-n)}, \quad (2.2.7), [7]$$

bude se jednat o tzv. symetrickou, popř. nesymetrickou charakteristiku.

2.3 Koncept adaptivní filtrace

Jestliže však filtr pracuje v neznámém prostředí, kdy předem není jasná identifikace šumu, není možno využívat klasický Wienerův filtr, ale je třeba vytvořit adaptivní filtr, který je schopen získávat odhady nezbytných veličin a na základě vyhodnocení těchto informací se v průběhu práce přizpůsobit. Očekává se tedy, že filtr s určitou rychlostí bude reagovat i na změny prostředí a zpracovávat tedy i nestacionární procesy. Cenou za to, že adaptivní filtry nevyžadují znalost zdroje signálu, je nutné poskytovat tzv. trénovací signál, jenž je přiváděn na další vstup. Tento trénovací signál musí úzce souviset s požadovaným výstupem filtru. [7], [15]



Obr. 7 : Obecná struktura adaptivního filtru

Obecné blokové schéma můžeme vidět na Obr. 7. Jak můžeme vidět, filtr má dva vstupy, měřený signál y_n a trénovací signál x_n , který více či méně těsně popisuje žádoucí výstupní signál. Výstupy jsou pak filtrovaný signál x_n' , dále chybový signál $e_n = x_n - x_n'$ a identifikační signál h_n , což je posloupnost okamžitých hodnot vektoru parametru filtru. Pozorovaný signál je zpracováván FIR filtrem je v každém časovém okamžiku popsán vektorem koeficientů, tj. impulzní charakteristikou h_n . Tato časově proměnná charakteristika je vytvářena blokem adaptačního algoritmu, který vektor h_n počítá pro každý takt n znovu na základě dvojice vstupujících signálů – vstupujícího (pozorovaného) signálu y_n a rozdílového signálu e_n . Filtr poskytuje také více či méně přesný odhad vzorků x_n' žádoucího signálu.

2.4 Spektrální odečítání

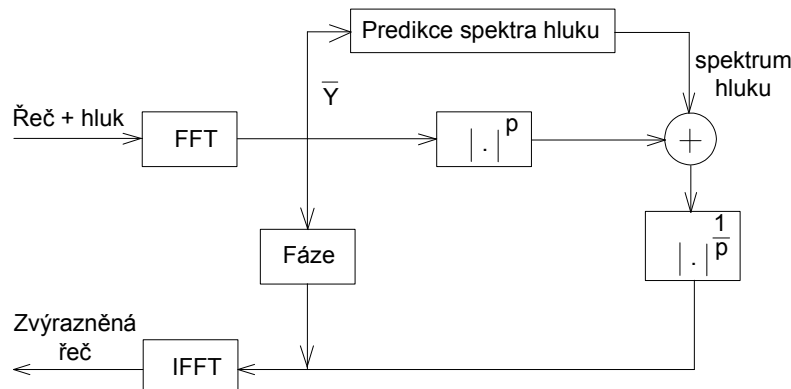
Princip spektrálního odečítání vychází z předpokladu, že řečový signál je zarušený hlukem okolí. Když známe charakter rušivého signálu, můžeme pak určit tzv. odhad spektra řečového signálu pomocí rozdílu spekter zkresleného řečového signálu a spektra hluku. Jestliže však spektrum hluku není známo, lze získat jeho odhad odečtením ze zarušeného signálu v době, kdy se nemluví. Z vypočteného spektra se pak určí pomocí IFFT časový průběh signálu. Je nutné dodat, že tato metoda je určena pro odstranění časově neměnných (popřípadě minimálně se měnících) rušivých elementů.

2.4.1 Základní algoritmus

Algoritmus spektrálního odečítání je obecně vyjádřen vztahem

$$|\hat{X}(\omega)|^p = |Y(\omega)|^p - |\hat{D}(\omega)|^p \quad \text{K} \quad 2.4.1, [16]$$

Kde $|\hat{X}(\omega)|$ je amplitudové spektrum zvýrazněného signálu, $|Y(\omega)|$ je amplitudové spektrum zarušeného signálu a $|\hat{D}(\omega)|$ je odhad amplitudového spektra hluku. Mocnina p určuje charakter rovnice. Jestliže $p = 1$, jedná se odečítání amplitudových spekter, pro $p = 2$ se budou odečítat výkonová spektra.



Obr. 8 : Blokové schéma spektrálního odečítání

Při nesprávném odhadu spektra hluku $|\widehat{D}(\omega)|$ se může stát, že výsledné amplitudové spektrum $|\widehat{X}(\omega)|$ bude nabývat záporných hodnot ($|\widehat{Y}(\omega)| < |\widehat{D}(\omega)|$). Pak je třeba spektrum usměrnit. Používají se dva druhy – jednocestné a dvoucestné. Při jednocestném usměrnění dochází k vynulování všech záporných hodnot. U dvoucestného usměrnění se záporné hodnoty berou za kladné.

Při usměrnění vznikají v kmitočtovém spektru výsledného signálu malé osamocené vrcholky, které se projevují po IFFT jako tzv. hudební tóny, jejichž kmitočty se nahodile mění.

2.4.2 Zlepšení základního algoritmu

Abychom eliminovali výskyt hudebních tónů, je důležité nahradit záporné hodnoty hodnotami kladnými, nenulovými. Můžeme využít Beroutiho algoritmus

$$\begin{aligned} |\widehat{X}(\omega)|^2 &= |Y(\omega)|^2 - \alpha |\widehat{D}(\omega)|^2 \quad \text{pro} \quad |Y(\omega)|^2 > (\alpha + \beta) |\widehat{D}(\omega)|^2, \\ |\widehat{X}(\omega)|^2 &= \beta |\widehat{D}(\omega)|^2 \quad \text{pro ostatní případy,} \end{aligned} \quad (2.4.2), [16]$$

kde $\alpha \geq 1$ je substrakční faktor a $0 < \beta \ll 1$ je parametr spektrálního pozadí. Takhle zajistíme, že hodnoty neklesnou pod hodnotu $\beta |\widehat{D}(\omega)|^2$. Při odečtu odhadu spektra hluku od spektra vstupního signálu s hlukem, zůstávají ve výsledném spektru širokopásmové a úzkopásmové vrcholky. Při nastavení parametru $\alpha > 1$ je dosaženo poklesu amplitudy širokopásmových vrcholků ve spektru. Parametr α tedy ovlivňuje zkreslení výsledného signálu. Je-li α příliš velké, dochází ke značnému zkreslení, což výrazně sníží srozumitelnost. Další zlepšení nám zajišťuje parametr β , který ovlivňuje množství hluku ve zpracovávaném signálu. Jestliže je hodnota β příliš vysoká, pak nebudou hudební tóny tak patrné, ovšem zbytkový hluk bude příliš slyšitelný. Parametr α se určuje pro každý zpracovávaný segment a platí pro něj [16]

$$\alpha = \alpha_0 - \frac{3}{20} SNR \quad \text{pro} \quad -5 \text{ dB} \leq SNR \leq 20 \text{ dB} \quad (2.4.3), [16]$$

kde α_0 je hodnota parametru α při $SNR = 0$ dB a SNR je odhad poměru signál/šum pro každý zpracovávaný segment. Dobrých výsledků algoritmu je dosahováno pro α_0 v rozsahu od 3 do 6. Beroutiho algoritmus předpokládá, že hluk v okolí ovlivňuje stejnou měrou všechny kmitočtové složky spektra řečového signálu [16],[17].

Další možností odstranění hudebních tónů a především zkvalitnění metody je nelineární spektrální odečítání, které již využívá frekvenčně závislé parametry a lépe tak odečítá rušivý hluk. Je založena na myšlence, aby se na kmitočtech s malým SNR odečítaly větší hodnoty hluku a na kmitočtech s vysokým SNR hodnoty menší. Platí zde vztahy

$$\begin{aligned} |\widehat{X}(\omega)| &= |\overline{Y}(\omega)| - \alpha(\omega)N(\omega) \quad \text{pro} \quad |\overline{Y}(\omega)| > \alpha(\omega)N(\omega) + \beta |\overline{D}(\omega)|, \\ |\widehat{X}(\omega)| &= \beta |\overline{Y}(\omega)| \quad \text{pro ostatní případy,} \end{aligned} \quad (2.4.4), [17]$$

kde $|\bar{Y}(\omega)$ a $|\bar{D}(\omega)$ je hladký odhad zkreslené řeči a hluku, $\alpha(\omega)$ je frekvenčně závislý parametr a $N(\omega)$ je nelineární funkce spektra hluku. Pro hladký odhad zkreslené řeči a hluku platí

$$\begin{aligned} |\bar{Y}_i(\omega)| &= \mu_y |\bar{Y}_{i-1}(\omega)| + (1 - \mu_y) |Y_i(\omega)| \\ |\bar{D}_i(\omega)| &= \mu_d |\bar{D}_{i-1}(\omega)| + (1 - \mu_d) |\hat{D}_i(\omega)| \end{aligned} \quad (2.4.5), [17]$$

kde $|\bar{Y}_i(\omega)$ je amplitudové spektrum řeči v i -tém rámci, $|\bar{D}_i(\omega)$ je odhad amplitudového spektra hluku v i -tém rámci, μ_y je konstanta $0.1 \leq \mu_y \leq 0.5$, μ_d je konstanta $0.5 \leq \mu_d \leq 0.9$. $N(\omega)$ představuje maximální hodnotu amplitudového spektra odhadu hluku $|\bar{D}_i(\omega)|$, která se určuje z posledních 40 rámců. [18]

Smyslem je, aby parametr $\alpha(\omega)$ nabýval velkých hodnot pro kmitočty s malými hodnotami SNR a naopak. Příkladem vhodné funkce je

$$\alpha(\omega) = \frac{1}{1 + \gamma \frac{|\bar{Y}(\omega)|}{|\bar{D}(\omega)|}} \quad (2.4.6), [18]$$

Kde γ je normovací činitel.

3 POUŽITÍ JEDNOTLIVÝCH METOD V PRAXI

Praktická část je založena na teoretických znalostech uvedených v předchozích kapitolách. Při práci bylo použito programů MATLAB a BENTLEY SYSTEM – Microstation.

Načtením signálu do prostředí Matlab, získáváme matici o rozměrech $[X,1]$, či $[X,2]$. Počet řádků X je dán počtem vzorků signálu x , přičemž se vychází ze vztahu

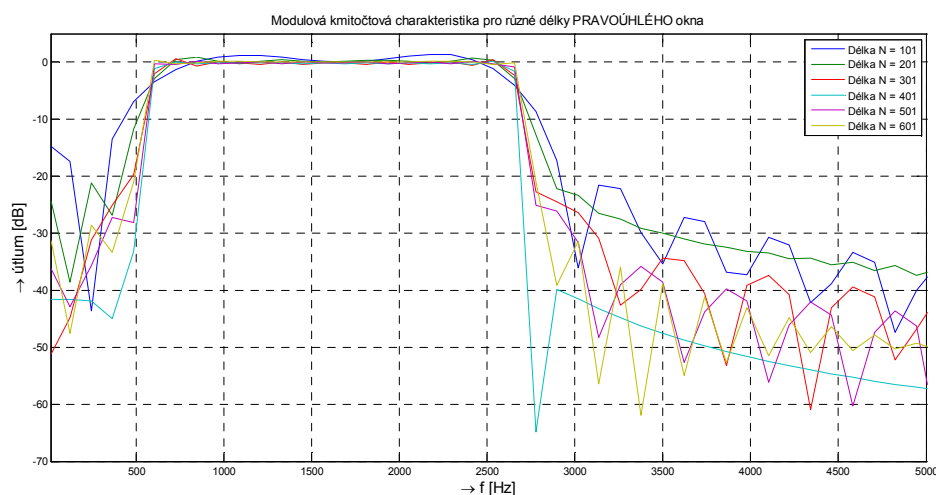
$$x = \frac{t}{f_{vz}}, \quad (3.1)$$

kde t představuje délku nahrávky a f_{vz} její vzorkovací frekvenci. Druhý rozměr má rozměr 1 - jedná-li se o mono nahrávku, 2 – jde-li o stereo. Všechny nahrávky použité v této práci jsou navzorkovány frekvencí 48 000 Hz.

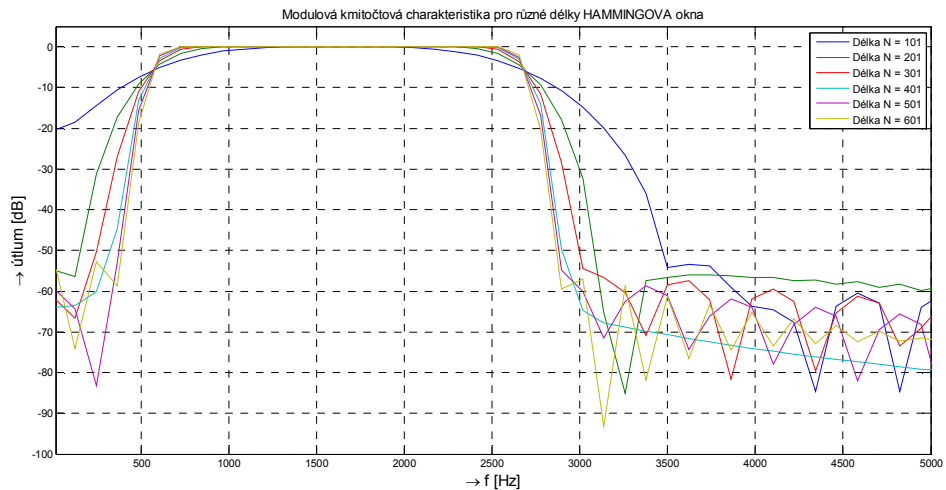
3.1 Lineární filtrace

3.1.1 Návrh filtru metodou oken pomocí Matlabu

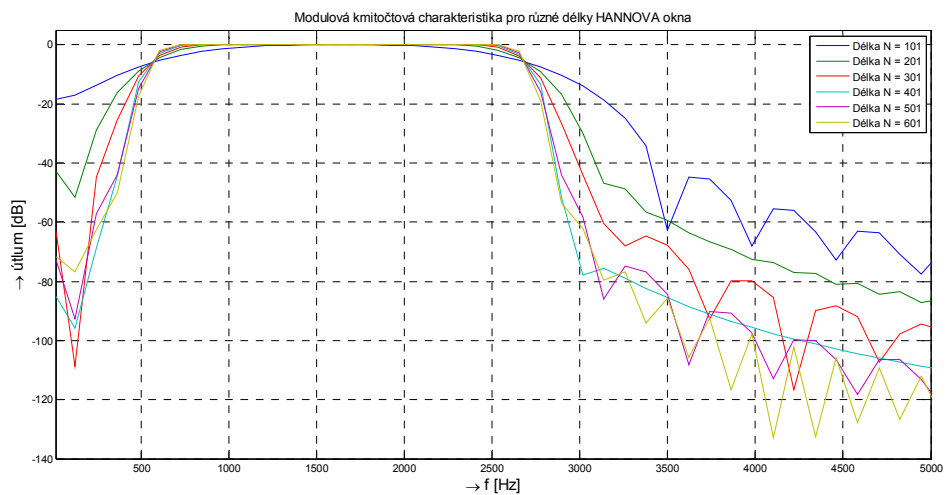
Filtr je navrhován jako pásmová propust (bandpass), a využívá se při návrhu v prostředí MATLAB funkce $b=fir1(n, Wn, 'ftype', window)$, která vrací vektor koeficientů filtru b o velikosti $n+1$, přičemž n představuje řád filtru, Wn vektor normované mezní frekvence, $'ftype'$ typ filtru (v našem případě je použit typ $'high'$) a $window$ typ násobícího okna o délce $n+1$. Takto vytvořené koeficienty jsou následně použity ve funkci $y=filter(b, a, X)$, která využívá námi navržený filtr, respektive jeho koeficienty b , a aplikuje jej na na vektor X , představující zarušený signál.



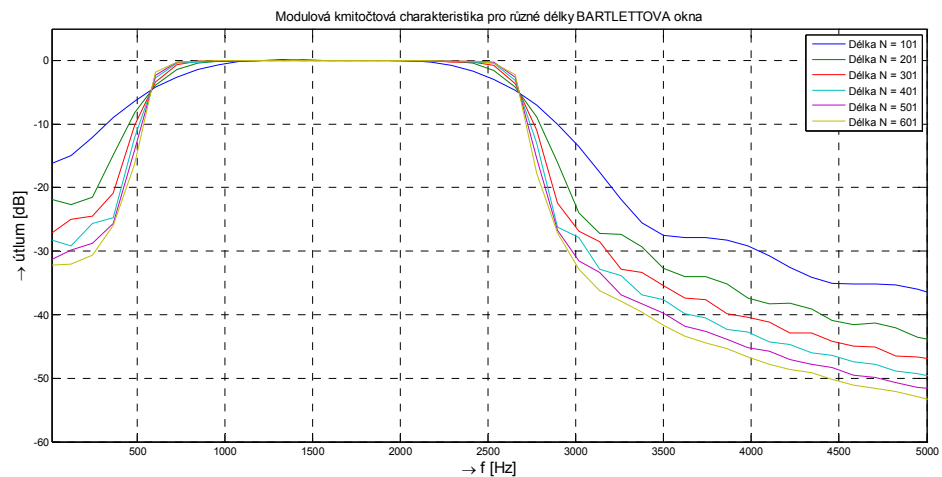
Obr. 9 : Modulová frekvenční charakteristika pásmové propusti s Pravoúhlým oknem o různých délkách.



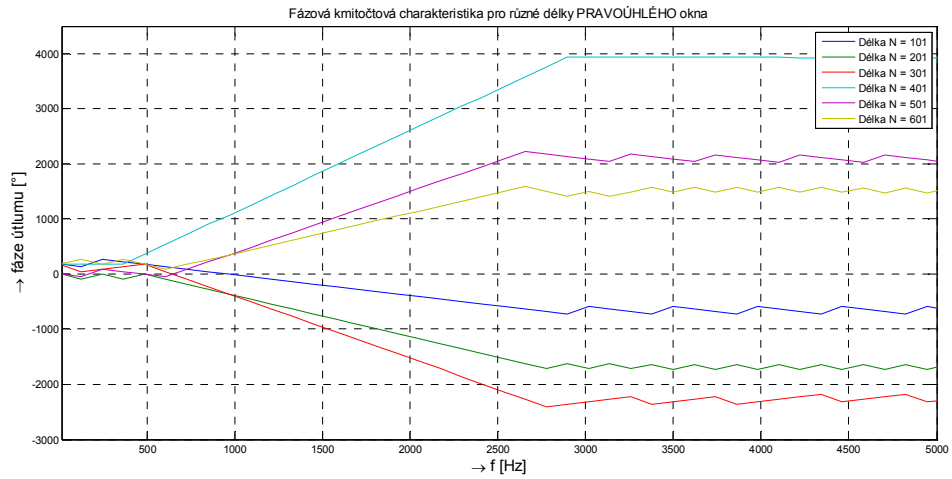
Obr. 10: Modulová frekvenční charakteristika pásmové propusti s Hammingovým oknem o různých délkách.



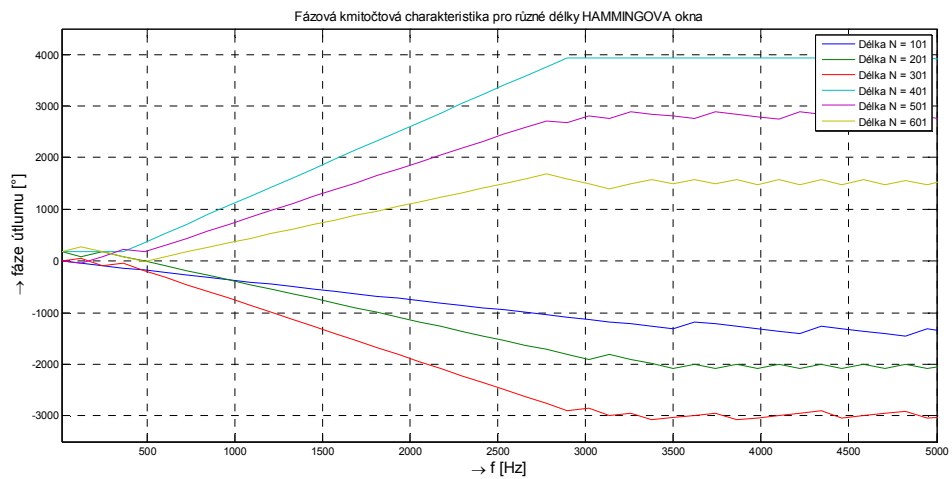
Obr. 11: Modulová frekvenční charakteristika pásmové propusti s Hannovým oknem o různých délkách.



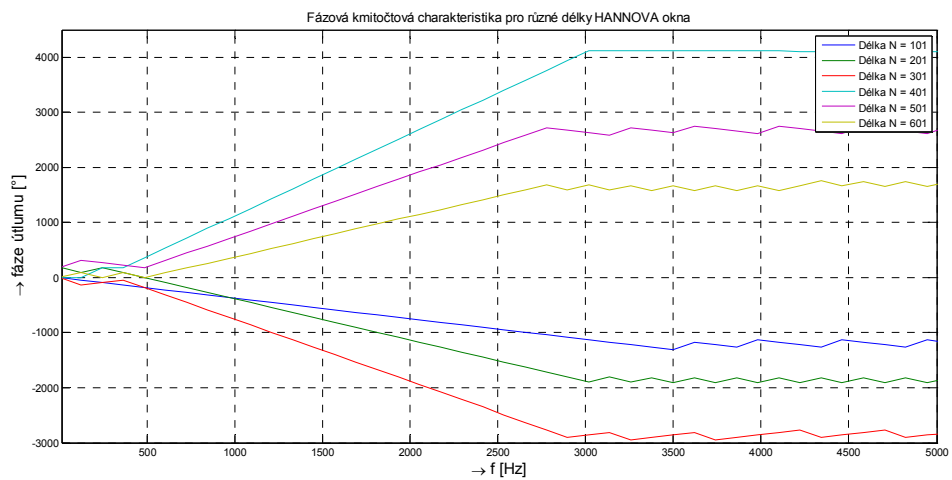
Obr. 12: Modulová frekvenční charakteristika pásmové propusti s Bartlettovým oknem o různých délkách.



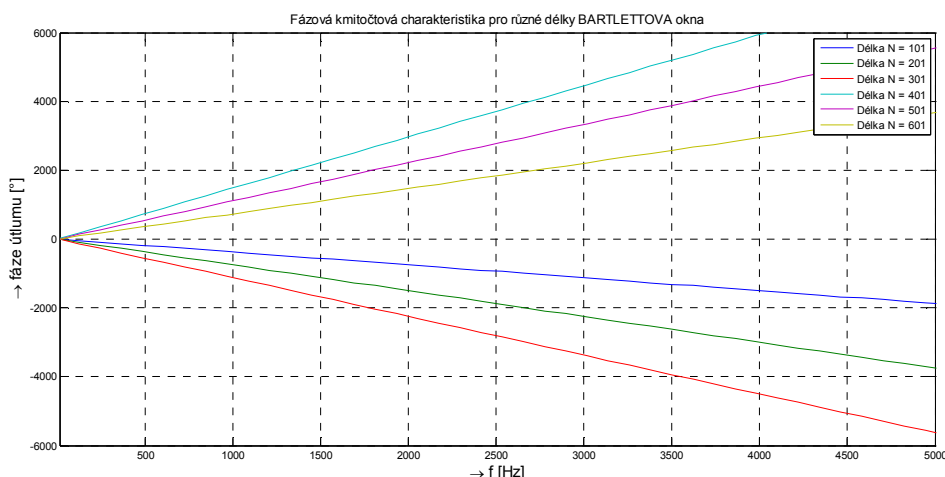
Obr. 13: Modulová frekvenční charakteristika pásmové propusti s Pravoúhlým oknem o různých délkách.



Obr. 14: Fázová frekvenční charakteristika pásmové propusti s Hammingovým oknem o různých délkách.



Obr. 15 : Fázová frekvenční charakteristika pásmové propusti s Hannovým oknem o různých délkách.



Obr. 16 : Fázová frekvenční charakteristika pásmové propusti s Bartlettovým oknem o různých délkách.

V tab.1 máme souhrn všech výsledků dosažených v této kapitole. Výstupní soubory, jenž byly vytvořeny při měnící se dolní mezní frekvenci f_d postupně pro všechny zmíněná okna, byly přehrávány 4 lidem ve věku od 18 do 25 let, jejich úkolem bylo hodnotit úroveň srozumitelnosti a četnost nežádoucích frekvenčních složek body od 1 do 10, přičemž hodnocení 1 je nejkvalitnější a 10 nejhorší.

Jak můžeme vidět, srozumitelnost klesala se zvyšující se dolní mezní frekvencí F_d a právě samotné odrušení má opačný charakter. Vše je způsobeno frekvenčním rozsahem řečového signálu [4], který se překrývá se spektrem rušivého zvuku motorů. Rušivý zvuk větru se podařilo odfiltrovat úplně, jelikož jeho rozsah začíná na hranici 3500 Hz a zde již spektrum řečového signálu nezasahuje.

Právě odfiltrováním spodní části spektra dochází k výrazné změně barvy hlasu a hodnocení kvality samotného řečového signálu v tomto případě sleduje, do jaké míry se hlas mluvčího mění. Nejlepších výsledků bylo dosaženo při nižších dolních mezních frekvencích propustného pásma a nižších řádech filtru. Je třeba však dodat, že důležitým faktorem je volba násobícího okna. Tvar modulových charakteristik jednotlivých oken se více, či méně odlišuje a to nejen v časové oblasti, ale i ve frekvenční. Tím vznikají rozdíly u hodnocení kvality při stejné hodnotě dolní mezní frekvence a různých typů oken. Jak můžeme vidět níže, nejlepších výsledků kvality (ovšem na úkor odrušení nežádoucích frekvenčních složek) bylo dosaženo při nejnižší zvolené dolní mezní frekvenci a u nejnižšího řádu filtru s použitým Hammingovým a Hannovým oknem.

Hlavní parametr ovlivňující četnost nežádoucích frekvenčních složek ve výsledném signálu je řád filtru, jenž výrazně ovlivňuje charakter přechodové charakteristiky filtru. Především tvar přechodových pásem, strmost a zároveň délku přechodového pásma, jenž klesá s rostoucím řádem filtru. Silně také ovlivňuje útlum v nepropustných pásmech. Například u pravoúhlého okna je vidět jak s rostoucím řádem klesá zvlnění v propustném pásmu. U ostatních oken je již zvlnění propustného pásma zanedbatelné, ovšem zase mnohem výrazněji je vidět, do jaké míry se mění délka propustného pásma. Rozdíl mezi řády 101 a 601 je u všech typů oken více než 900 Hz. Dalším parametrem je útlum nepropustných pásem, který je nejvíce ovlivňován typem zvoleného okna.

Největšího útlumu dosahuje Hannovo okno s Hammingovým, téměř dvojnásobného než je u Pravoúhlého a Bartlettova. Podle hodnocení odrušení z tab.1 je nejlepších výsledků dosaženo při volbách vyšších řádů, přičemž výrazně klesá podobnost původního mluvčího hlasu.

Kompromisem při nastavení filtru použitelného pro naše účely se rozumí, aby filtr nijak výrazně neměnil barvu hlasu mluvčího a při tom bylo odrušeno co možná největší počet nežádoucích složek. Podle tab.1 by tímto kompromisem mohl být filtr s řádem 601, s dolní mezní frekvencí 450Hz a použitým Pravoúhlým a Hammingovým oknem. Jelikož jde o filtr s velkým počtem řádu, což by bylo komplikované při realizaci, je možné volit filtr s počtem řádů 301, dolní mezní frekvence se již změnila na 550Hz a typ použitého okna na Hammingovo a Hannovo.

Tab. 1 : Vyhodnocení odrušení pomocí filtru navrženého metodou oken

FIR filr navržený fcí FIR1				Hodnotící subjekty									
Typ okna	Řád filtru	Toleranční pásmo		Osoba č.1		Osoba č.2		Osoba č.3		Osoba č.4		Průměr	
		N	F _d (Hz)	F _h (Hz)	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.
Pravouhíle okno	101	450	2700	3	5	4	5	3	5	5	7	3,75	5,5
		550	2700	3	4	5	5	4	4	5	6	4,25	4,75
		650	2700	4	3	6	4	5	3	6	5	5,25	3,75
	301	450	2700	4	4	5	6	4	5	6	6	4,75	5,25
		550	2700	5	3	6	5	5	4	7	5	5,75	4,25
		650	2700	5	2	5	3	5	3	7	4	5,5	3
	601	450	2700	3	2	4	3	4	3	5	4	4	3
		550	2700	5	4	5	3	5	4	7	6	5,5	4,25
		650	2700	6	3	6	3	5	3	7	5	6	3,5
Hammingovo okno	101	450	2700	2	6	4	6	3	6	4	7	3,25	6,25
		550	2700	3	4	4	5	3	4	5	6	3,75	4,75
		650	2700	3	4	4	4	4	4	5	6	4	4,5
	301	450	2700	3	4	4	5	3	4	4	5	3,5	4,5
		550	2700	4	3	5	4	4	3	6	5	4,75	3,75
		650	2700	4	3	5	4	4	4	6	7	4,75	4,5
	601	450	2700	3	2	4	4	4	3	5	4	4	3,25
		550	2700	4	2	4	3	4	3	6	4	4,5	3
		650	2700	5	3	5	3	5	3	7	5	5,5	3,5
Hannovo okno	101	450	2700	2	6	4	7	3	6	4	7	3,25	6,5
		550	2700	3	5	5	6	4	6	5	6	4,25	5,75
		650	2700	3	5	5	4	4	5	5	6	4,25	5
	301	450	2700	3	3	5	4	4	4	5	5	4,25	4
		550	2700	3	3	5	3	4	3	5	5	4,25	3,5
		650	2700	4	2	6	3	5	3	5	4	5	3
	601	450	2700	4	4	5	4	5	5	6	5	5	4,5
		550	2700	5	4	6	4	5	4	6	5	5,5	4,25
		650	2700	6	3	6	3	5	4	7	4	6	3,5
Bartlettovo okno	101	450	2700	3	7	4	7	4	7	5	6	4	6,75
		550	2700	3	6	5	8	4	6	5	6	4,25	6,5
		650	2700	3	5	5	6	5	6	5	5	4,5	5,5
	301	450	2700	4	5	5	6	5	6	6	7	5	6
		550	2700	4	5	5	6	5	5	6	7	5	5,75
		650	2700	5	4	5	5	5	5	6	6	5,25	5
	601	450	2700	4	5	5	5	4	5	6	7	4,75	5,5
		550	2700	5	4	6	5	5	4	6	6	5,5	4,75
		650	2700	5	4	6	4	5	4	6	6	5,5	4,5

 * F_d ... dolní mezní frekvence; F_h ... horní mezní frekvence

3.1.2 Ruční návrh filtru metodou oken.

Pro naše účely byla navržena pásmová propust (bandpass), která je konstruována pomocí metody oken. Mezní frekvence byly určeny z předchozích pokusů jako :

$$f_{-p} = 650 \text{ Hz}, f_p = 2600 \text{ Hz}, f_{-s} = 450 \text{ Hz}, f_s = 2800 \text{ Hz},$$

přičemž dolní a horní mezní frekvence propustného pásma jsou označeny jako f_{-p} a f_p , dolní a horní mezní frekvence nepropustného pásma f_{-s} a f_s .

Filtr budeme normovat na polovinu vzorkovacího kmitočtu, jenž odpovídá hodnotě $fvz = 48 \text{ kHz}$. Jednotlivé kmitočty normujeme podle vzorce 3.2 a dostáváme :

$$\omega_{-p} = 0,0851, \omega_p = 0,3403, \omega_{-s} = 0,0589, \omega_s = 0,3665.$$

$$\omega = \frac{\pi f}{0.5 fvz} = \frac{2\pi f}{fvz}. \quad (3.2)$$

Dále si určíme šířku horního a dolního přechodového pásma (pro jednoduchost byly voleny počáteční frekvence tak, aby měla obě pásma stejnou velikost). Platí tedy

$$\Delta\omega_d = \omega_{-p} - \omega_{-s} \quad \text{a} \quad \Delta\omega_h = \omega_s - \omega_p, \quad (3.3)$$

$$\Delta\omega = \Delta\omega_d = \Delta\omega_h = 0,0262.$$

Vypočítáme si středy jednotlivých pásem podle 3.4 :

$$\omega_{ds} = \omega_{-p} - \Delta\omega \quad \omega_{hs} = \omega_p - \Delta\omega \quad (3.4)$$

$$\omega_{ds} = 0,0720 \quad \omega_{hs} = 0,3534$$

Následuje výpočet koeficientů, který je rozdílný pro každý typ okna. Pro Pravoúhlé okno platí 3.5, pro Hammingovo okno 3.6, pro Hannovo okno 3.7, pro Bartlettovo okno 3.8.

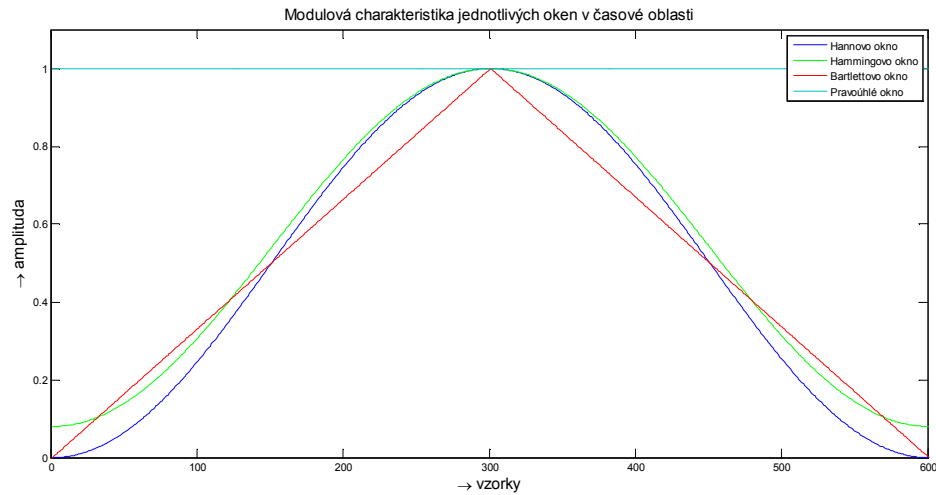
$$\omega[n] = \begin{cases} 1 & 0 \leq n \leq N-1 \\ 0 & \text{jinak} \end{cases} \quad (3.5)$$

$$\omega[n] = \begin{cases} 0.54 - 0.46 \cos\left(2\pi \frac{n}{N}\right) & 0 \leq n \leq N-1 \\ 0 & \text{jinak} \end{cases} \quad (3.6)$$

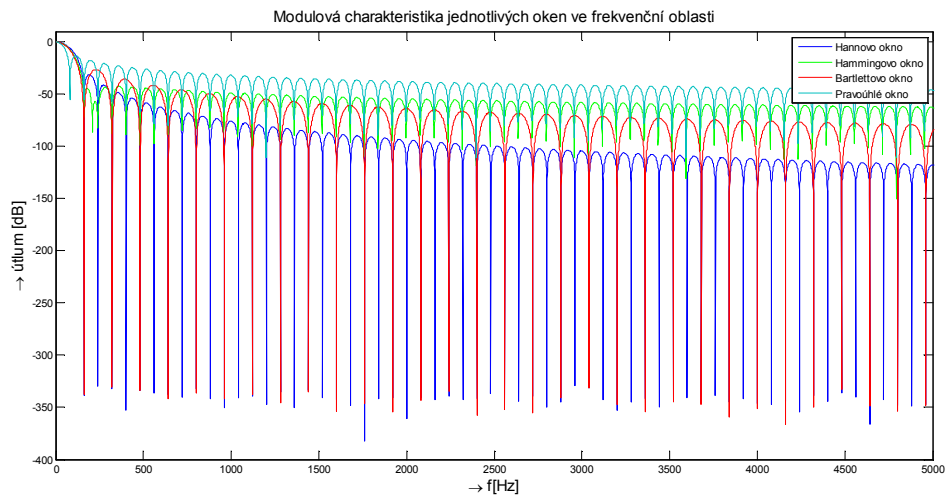
$$\omega[n] = \begin{cases} 0.5 \left[1 - \cos\left(\frac{2\pi n}{N-1}\right) \right] & 0 \leq n \leq N-1 \\ 0 & \text{jinak} \end{cases} \quad (3.7)$$

$$\omega[n] = \begin{cases} \frac{2n}{N} & 0 \leq n \leq \frac{N+1}{2} - 1 \\ 2 - \frac{2(N-n)}{N} & \frac{N+1}{2} \leq n \leq N \end{cases} \quad (3.8)$$

Na Obr.17 vidíme, koeficienty jednotlivých oken v časové oblasti a na Obr.18 ve frekvenční oblasti.



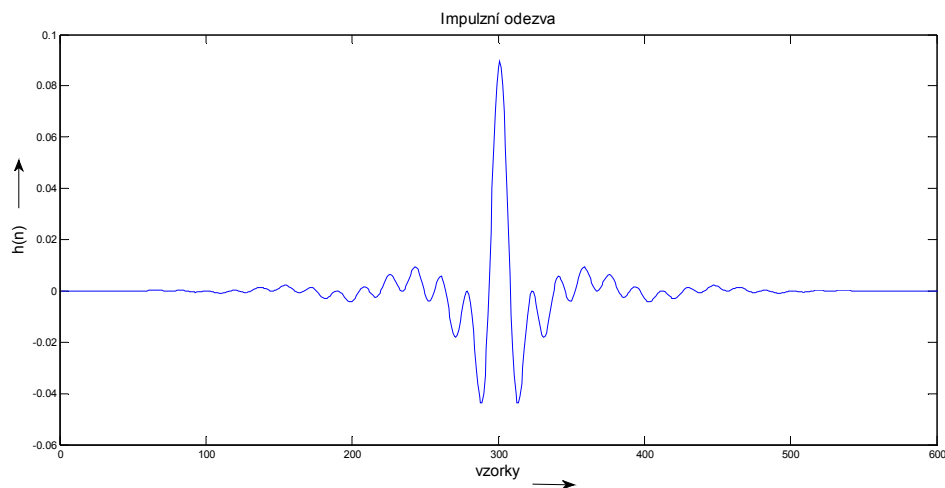
Obr. 17 : Modulová charakteristika jednotlivých oken v časové oblasti



Obr. 18 : Modulová charakteristika jednotlivých oken ve frekvenční oblasti

Impulzovou odezvu filtrů počítáme podle 3.9 a můžeme ji vidět na Obr.19

$$h[n] = \begin{cases} \frac{\omega_{hS} - \omega_{dS}}{\pi} & m = 0 \\ \frac{1}{m\pi} [\sin(\omega_{hS} * m) - \sin(\omega_{dS} * m)] & m \neq 0 \end{cases} \quad (3.9)$$

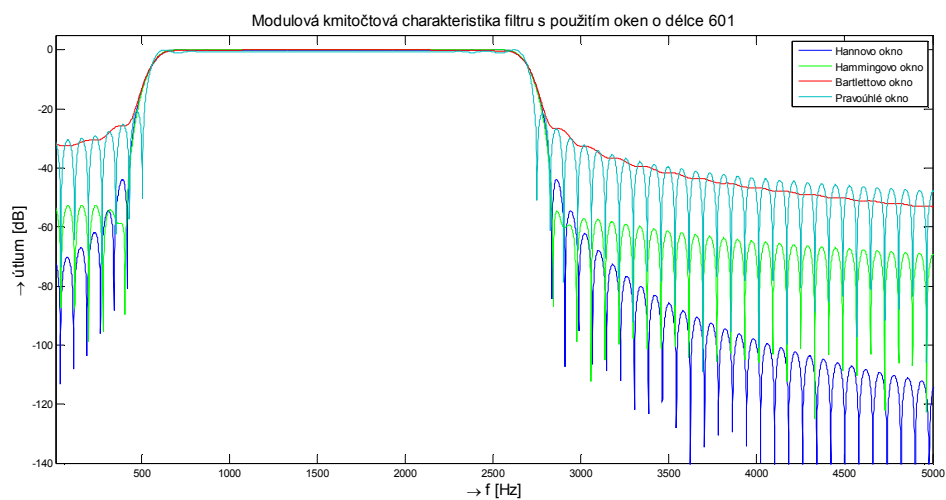


Obr. 19 : Impulzní odezva filtru

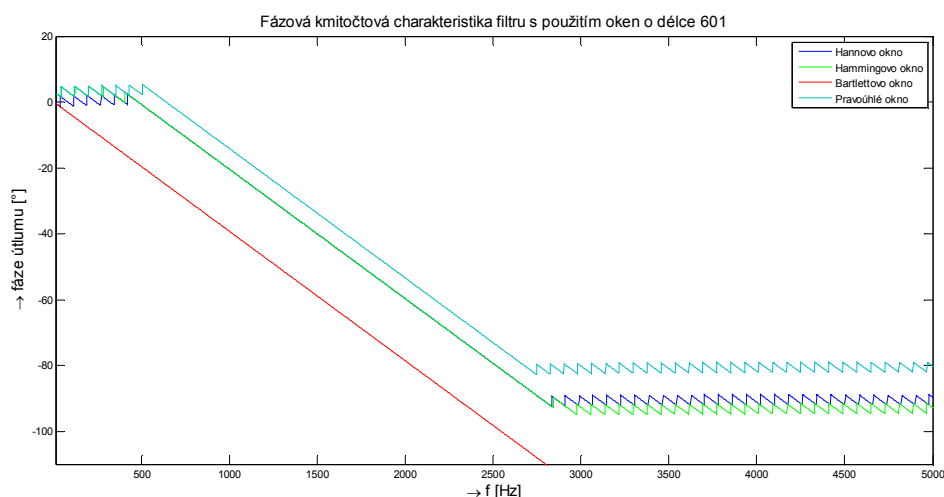
Koeficienty naší pásmové propusti získáme přenásobením koeficientů ideální impulzní odezvy $h[n]$ s koeficienty pro námi vybrané okno $w[n]$.

$$h_{pp}[n] = h[n] * w[n] \quad (3.10)$$

Níže můžeme vidět do jaké míry ovlivňuje pásmovou propust typ zvoleného okna. Největší útlumu nepropustných pásem jsme schopni docílit pomocí Hannova okna.



Obr. 20 : Modulová frekvenční charakteristika filtru s ukázkou všech oken



Obr. 21 : Fázová frekvenční charakteristika filtru s ukázkou všech oken

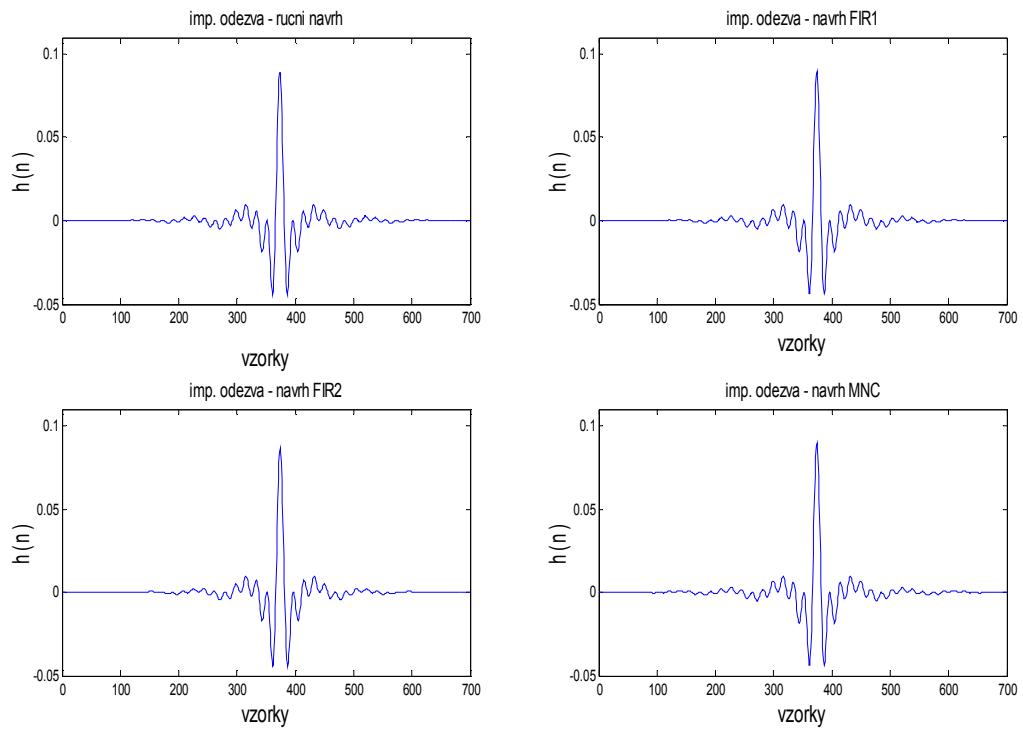
Tab. 2 : Vyhodnocení odrušení pomocí ručně spočítaného filtru

FIR filr navržený ručně				Hodnotící subjekty									
Typ okna	Řád filtru	Toleranční pásmo		Osoba č.1		Osoba č.2		Osoba č.3		Osoba č.4		Průměr	
	N	Fd(Hz)	Fh(Hz)	Kval.	Oduš.	Kval.	Oduš.	Kval.	Oduš.	Kval.	Oduš.	Kval.	Oduš.
Pravoúhl.	601	550	2700	6	5	5	3	5	4	7	6	5,5	4,25
Hammig.				6	4	4	3	4	3	6	4	4,5	3
Hann.				6	5	6	4	5	4	6	5	5,5	4,25
Bartlett.				6	5	6	5	5	4	6	6	5,5	4,75

Ruční návrh filtru byl spíše ukázkou, do jaké míry nám zjednoduší funkce Matlabu práci, přičemž docílíme stejného výsledku. Podle Tab.2 jsou výsledné signály na poslech totožné se signály, jenž byly vytvořeny pomocí integrovaných funkcí Matlabu. Matlab nabízí několik možností návrhu filtru :

- *fir1* – návrh filtru pomocí metody oken (použita v kapitole 3.1.1.)
- *fir2* – návrh filtru pomocí metody frekvenčního vzorkování
- *firls* – návrh filtru pomocí metody nejmenších čtverců (MNC)

Na Obr.22 vidíme nepatrné rozdíly impulzních charakteristik získaných na základě jednotlivých metod.

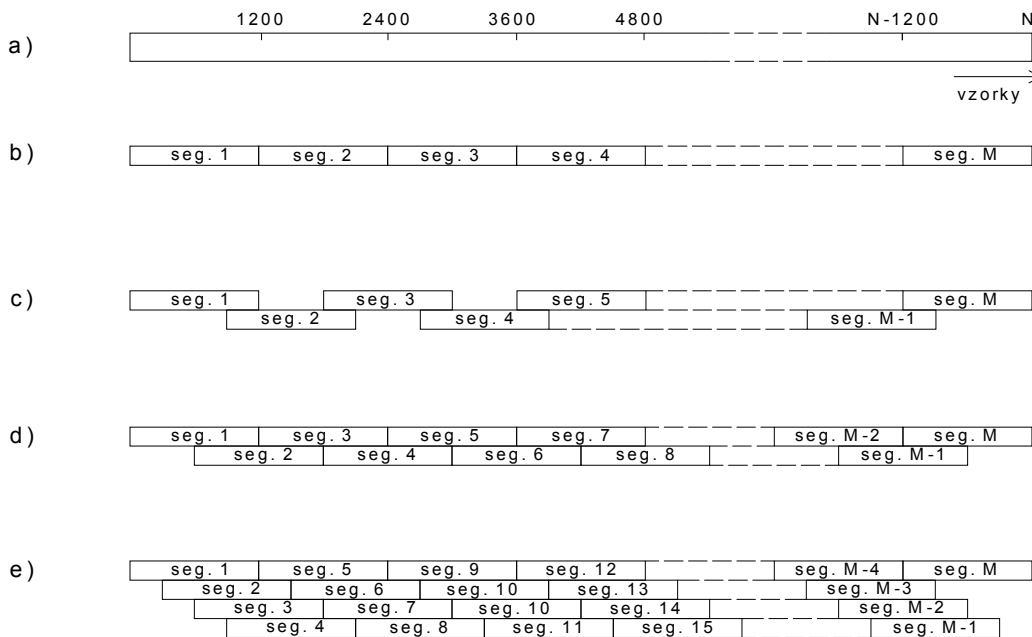


Obr. 22: Impulzní odezvy při různých metodách návrhu filtru

3.2 Spektrální odečítání

Po načtení signálu do *Workspace* musí dojít k rozčlenění signálu do jednotlivých úseků, tzv. segmentů. Velikost segmentu je zvolena podle [19] na 25 ms, což odpovídá 1200 vzorkům. Jak můžeme vidět na obr.16, je několik možností, jak můžeme jednotlivé segmenty naplnit:

- s překrytím 0% (bez překrytí) – segment č.1 bude obsahovat 1. až 1200. vzorek z původního signálu, segment č.2 1201. až 1400., atd. .
- segmenty se budou překrývat o 25% - segment č.1 bude obsahovat 1. až 1200. vzorek z původního signálu, segment č.2 901. až 2100., segment č.3 1801. až 3000., atd. .
- segmenty se budou překrývat o 50% - segment č.1 bude obsahovat 1. až 1200. vzorek z původního signálu, segment č.2 601. až 1800., segment č.3 901. až 2100., atd. .
- segmenty se budou překrývat o 75% - segment č.1 bude obsahovat 1. až 1200. vzorek z původního signálu, segment č.2 301. až 1501., segment č.3 601. až 1800., atd. .



Obr. 23 : Princip plnění segmentů:

- Původní signál
- Plnění segmentů bez překrytí
- Plnění segmentů s překrytím 25%
- Plnění segmentů s překrytím 50%
- Plnění segmentů s překrytím 75%

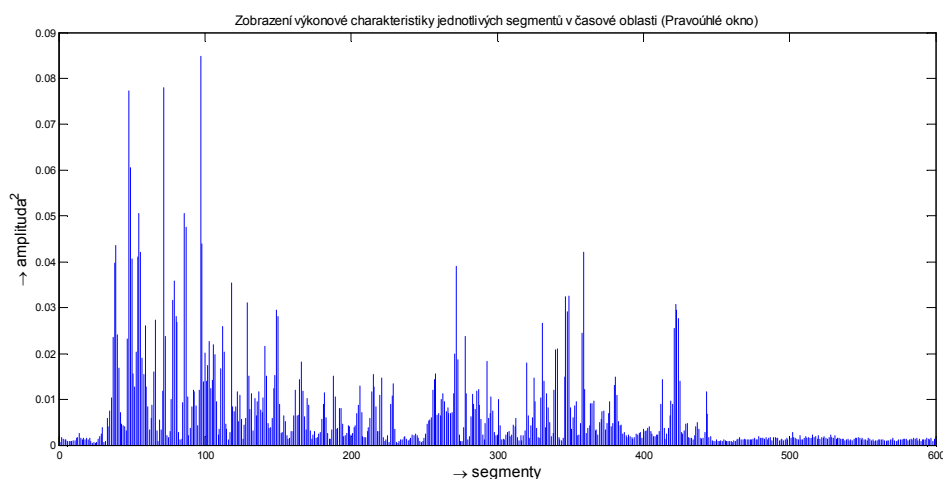
Jednotlivé segmenty pak jsou postupně roznásobeny (kvůli testování kvality) Pravoúhlým, Hammingovým, Hannovým a Bartlettovým oknem.

Provedením diskrétní Fourierovy transformace FFT získáme tvar jednotlivých

vzorků v komplexním tvaru $a+bi$, kde a je reálná složka a bi je složka imaginární. Jelikož je metoda spektrálního odečítání založena na odečítání amplitudového spektra šumu od amplitudového spektra původního signálu a s fází se nepracuje, je nutno hodnoty signálu v jednotlivých segmentech upravit tak, abychom je měli v goniometrickém tvaru $|Y|(\cos\varphi+i\sin\varphi)$, což nám umožňují integrované funkce Matlabu jako jsou $abs(.)$ a $angle(.)$. Tím dostáváme nám vyhovující tvar pro další práci.

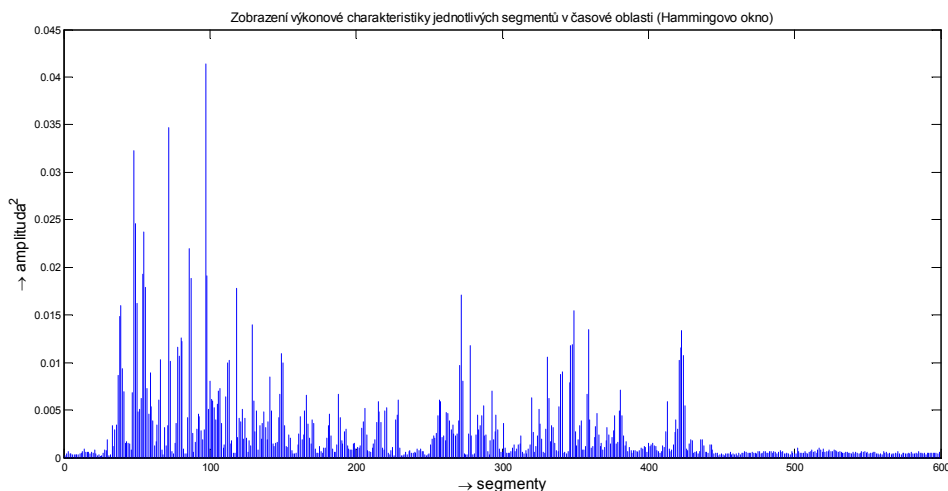
Dalším úkolem je zjistit amplitudové spektrum hluku (šumu), které budeme odečítat od amplitudového spektra signálu. Můžeme jej odečíst z mezer mezi jednotlivými slovy a nebo v části signálu, kde se nemluví. Jelikož budeme pracovat ve výkonovém spektru, je třeba umocnit amplitudové spektrum na druhou a z něj odečíst hodnotu, která odpovídají prahové hladině šumu. Ty se u jednotlivých oken liší a jsou nastavovány následovně:

- Pravoúhlé okno – 0.005, 0.01, 0.02



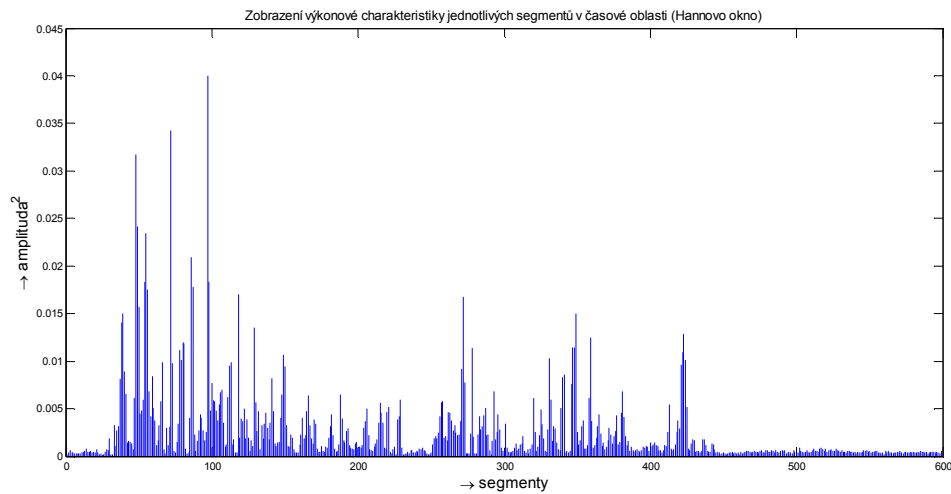
Obr. 24 : Výkonová charakteristika segmentů při použití Pravoúhlého okna

- Hammingovo okno – 0.0025, 0.005, 0.01



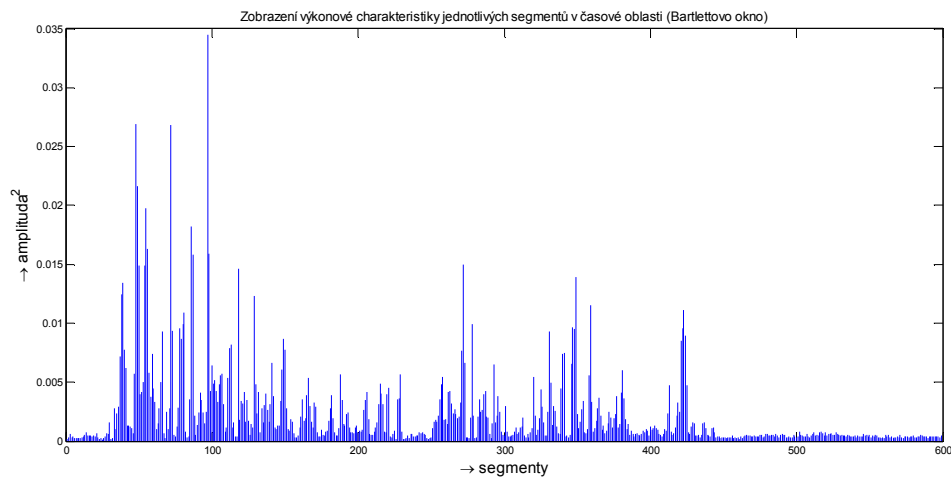
Obr. 25 : Výkonová charakteristika segmentů při použití Hammingova okna

- Hannovo okno – 0.001, 0.005, 0.01



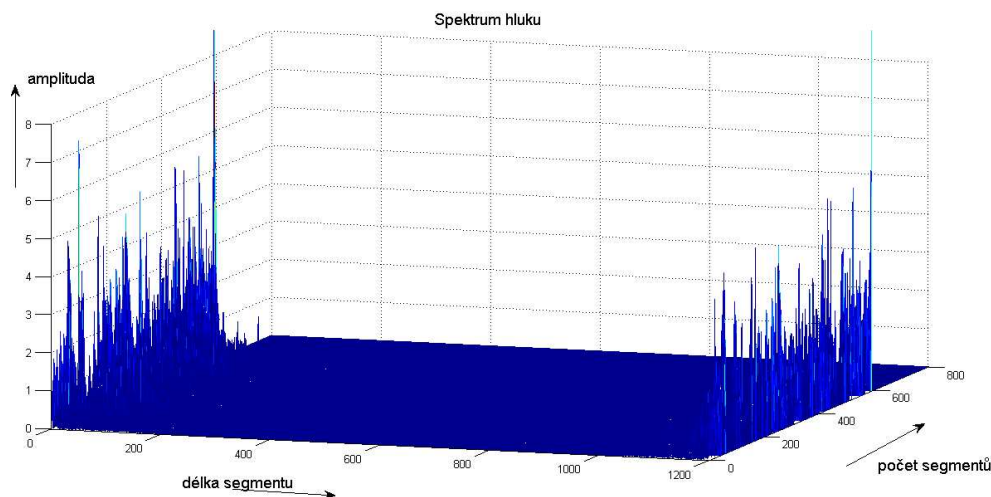
Obr. 26 : Výkonová charakteristika segmentů při použití Hannova okna

- Bartlettovo okno – 0.0025, 0.005, 0.01



Obr. 27 : Výkonová charakteristika segmentů při použití Bartlettova okna

Druhým parametrem, jež bylo možné nastavit, byl časový úsek, kdy program rozpoznává pauzu v řeči, ovšem tento parametr nijak výrazně neovlivňoval výsledný signál a byl teda nastaven na jednu hodnotu 0.1s. Na základě porovnávání mnou nastavených prahových hodnot s jednotlivými segmenty se vyberou všechny, jež nemají výkonovou střední hodnotu (pomocí funkce *mean(..)*) větší než nastavené prahy a vloží se do matice, jež bude mít stejný charakter jako matice, z níž se segmenty vybíraly. Nově vzniklá matice představuje spektrum šumu v jednotlivých segmentech.



Obr. 28 : Spektrum hluku

Jelikož se nám podařilo získat oba operandy, může dojít k samotnému odečtení podle vztahu 2.4.1. Abychom mohli provést zpětnou Fourierovu transformaci IFFT, je potřeba mít jednotlivé prvky matice v komplexním tvaru. Z vlastností pravoúhlého trojúhelníku získáváme podle rovnice 3.11 a 3.12 vztahy pro výpočet reálné a imaginární složky [19].

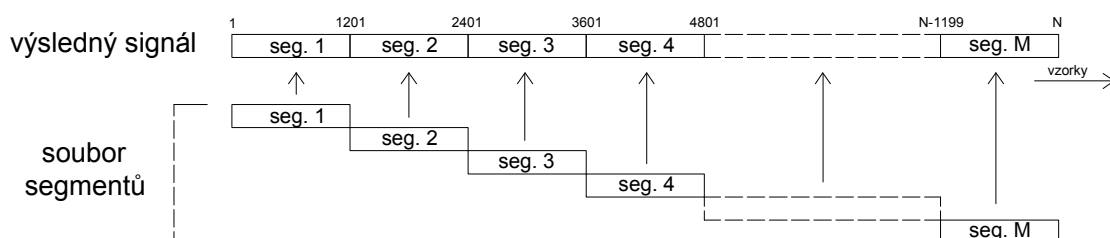
$$\text{Re} = |Y| \cos \alpha \quad (3.11)$$

$$\text{Im} = |Y| \sin \alpha \quad (3.12)$$

kde $|Y|$ je amplituda opraveného signálu a α je fáze původního signálu. Tento výpočet je aplikován pro všechny prvky matice a následně je provedena IFFT.

Poslední věc, kterou je potřeba provést, je vytvořit z jednotlivých segmentů celistvý signál. Tato operace probíhá jinak u segmentů vytvořených bez překrytí a jinak u překrývajících se segmentů.

V prvním případě se jednotlivé segmenty řadí postupně za sebou.



Obr. 29 : Skládání výsledného signálu bez překrývání segmentů

Jestliže kompletujeme původně překrývané segmenty, je důležité dodržet podmínku, aby se přednostně využívalo středových hodnot segmentu a tím se eliminovalo obsahu ořezaných segmentů ve výsledném signále. Obecně platí

```
for A=1:Nseg-3
signal_finish(1, ((1+(Nwnd/2)+(Nwnd/(2*sh)))+(Nwnd/2)*A) : ((Nwnd/2)+(Nwnd/(2*sh)))+(A+1)*(Nwnd/2)))=signal_komplet(A+2, (1+(Nwnd/(2*sh))):(Nwnd/2)+(Nwnd/(2*sh))))
end
```

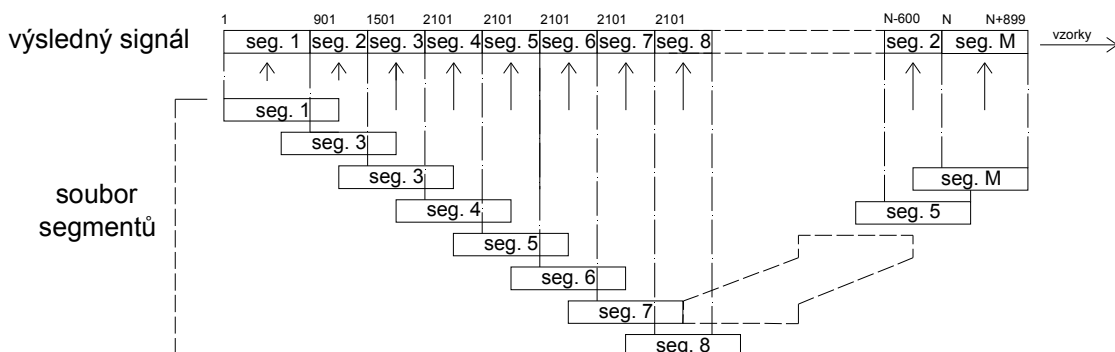
kde *signal_finish* odpovídá výslednému signálu, komplexní signál je označen jako *signal_komplet*, *Nwnd* představuje délku okna, *sh* posunutí segmentů a *Nseg* počet vzniklých segmentů z původního signálu.

Skládání signálu při překrývání segmentů 25% je ukázáno na Obr.30. Ze segmentu č.1 se použije prvních 1051 vzorků. Z každého dalšího segmentu se pak postupně zapisují vzorky z pozic 151 až 900. Tímto způsobem se zapisuje do vstupního signálu až po segment č.M, z něhož se zapíší vzorky 151 až 1200.



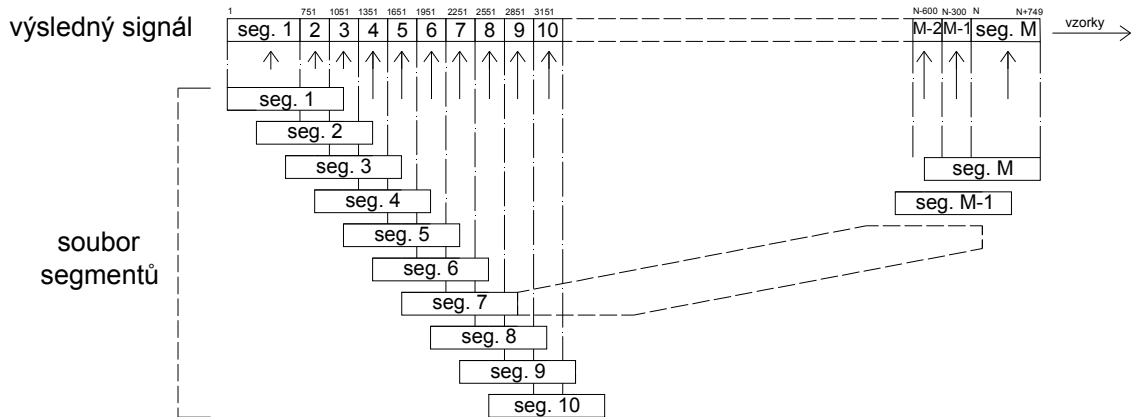
Obr. 30 : Skládání výsledného signálu při 25% překrývání segmentů

Skládání signálu při překrývání segmentů 50% je zobrazeno na Obr.31. Ze segmentu č.1 se použije prvních 900 vzorků. Z každého dalšího segmentu se pak postupně zapisuje z pozic 301 až 900. Tímto způsobem se zapisuje do vstupního signálu až po segment č.M, z něhož se zapíší vzorky 301 až 1200.



Obr. 31 : Skládání výsledného signálu při 50% překrývání segmentů

Skládání signálu při překrývání segmentů 75% je zobrazeno na Obr.32. Ze segmentu č.1 se použije prvních 750 vzorků. Z každého dalšího segmentu se pak postupně zapisuje z pozic 451 až 750. Tímto způsobem se zapisuje do vstupního signálu až po segment č.M, z něhož se zapíší vzorky 451 až 1200.



Obr. 32 : Skládání výsledného signálu při 75% překrývání segmentů

Vytvoření souboru pro zpětnou kontrolu ve formátu *.wav zajišťuje funkce *wavwrite(.)*.

Vyhodnocování výsledných souborů bylo provedeno podobně jako v předchozí kapitole, ovšem při hodnocení kvality řečové části signálu se tentokrát hodnotí srozumitelnost (především souvislost řeči a podobnost s původním signálem). Přičemž tato metoda byla zkoumána pro rozdílné nastavení prahu šumu a rozdílné zpracování signálu při použití různých oken při roznásobování jednotlivých segmentů.

Kvalitu řeči ve výsledném signálu silně ovlivňuje průběžné zpracování signálu. Tab.3, Tab.4, Tab.5 a Tab.6 ukazují, jak významnou roli hraje typ násobícího okna u rozdílného překrývání segmentů signálu. Jestliže se segmenty nepřekrývají vůbec a je použito Pravoúhlé okno, je dosaženo kvality odpovídající původnímu signálu. Když ale násobíme jednotlivé segmenty jiným oknem než Pravoúhlým, dochází k ořezávání okrajů segmentů a následně při kompletaci vzniká výrazně deformovaný signál (signál má charakter uměle vytvořené řeči), který je u všech zbylých oken podobný. Tuto deformaci částečně odstraňuje zpracování, kdy je signál rozčleněn do segmentů, jež se překrývají o 25%. Při překrytí segmentů o 50% dochází ke zlepšení kvality na dostatečnou úroveň, ale bohužel ani při překrytí segmentů o 75% není dosaženo kvality, kterou zajišťuje Pravoúhlé okno.

Hodnocení odrušení bylo v tomto případě velmi komplikované, protože signál byl odrušen buď jen částečně a nebo se ve výsledku objevovaly nové rušivé složky, tzv. hodební tóny. Jejich četnost rostla se zvyšující výškou prahu, určující úroveň hluku. Tyto tóny vznikají především při špatně určeném odhadu spektra hluku a nebo při proměnlivém charakteru hluku.

Za kompromis této metody je možné považovat spektrálního odečítání s použitím Pravoúhlého okna a jakýmkoliv překrýváním segmentů, čímž bylo dosaženo nejlepší kvality a podobného odrušení jako u ostatních oken při překrýváním segmentů o 75 %.

Jedná se o metodu velice náročnou na výpočty, především při zvyšujícím se

překrývání segmentů.

Tab. 3 : Vyhodnocení spektrálního odečítání s Pravoúhlým násobícím oknem

Spektrální odečítání			Hodnotící subjekty										
Typ okna	Nastavené prahy		Překrytí segmentů	Osoba č.1		Osoba č.2		Osoba č.3		Osoba č.4		Průměr	
	Ta (s)	Tb (s)		Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.
Pravoúhlé okno	0,005	0,1	0%	2	7	3	7	2	8	3	7	2,5	7,25
			25%	2	6	3	6	2	8	3	8	2,5	7
			50%	2	6	3	7	2	8	3	7	2,5	7
			75%	2	7	3	7	2	7	3	6	2,5	6,75
	0,01	0,1	0%	2	6	3	6	2	8	3	6	2,5	6,5
			25%	2	6	3	6	2	8	3	7	2,5	6,75
			50%	2	6	3	7	2	8	3	7	2,5	7
			75%	2	6	3	7	2	7	3	6	2,5	6,5
	0,02	0,1	0%	2	5	3	6	2	7	3	6	2,5	6
			25%	3	5	3	5	3	7	4	7	3,25	6
			50%	3	6	3	5	3	8	4	7	3,25	6,5
			75%	2	6	3	7	3	7	3	7	2,75	6,75

Tab. 4 : Vyhodnocení spektrálního odečítání s Hammingovým násobícím oknem

Spektrální odečítání			Hodnotící subjekty										
Typ okna	Nastavené prahy		Překrytí segmentů	Osoba č.1		Osoba č.2		Osoba č.3		Osoba č.4		Průměr	
	Ta (s)	Tb (s)		Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.
Hammingovo okno	0,0025	0,1	0%	8	8	8	7	7	8	8	7	7,75	7,5
			25%	7	7	7	7	6	7	7	7	6,75	7
			50%	5	7	6	7	5	6	6	7	5,5	6,75
			75%	4	7	5	7	4	6	5	7	4,5	6,75
	0,005	0,1	0%	8	7	8	7	7	7	8	8	7,75	7,25
			25%	7	7	7	7	7	8	8	8	7,25	7,5
			50%	5	6	6	7	6	7	6	8	5,75	7
			75%	4	6	5	7	6	7	5	8	5	7
	0,01	0,1	0%	7	6	8	7	8	8	8	9	7,75	7,5
			25%	6	8	7	7	7	8	6	9	6,5	8
			50%	6	7	6	7	6	8	6	8	6	7,5
			75%	5	7	6	8	6	8	5	8	5,5	7,75

Tab. 5 : Vyhodnocení spektrálního odečítání s Hannovým násobícím oknem

Spektrální odečítání				Hodnotící subjekty									
Typ okna	Nastavené prahy		Překrytí segmentů	Osoba č.1		Osoba č.2		Osoba č.3		Osoba č.4		Průměr	
	Ta (s)	Tb (s)		Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.
Hannovo okno	0,001	0,1	0%	7	6	8	7	8	7	9	6	8	6,5
			25%	6	6	6	7	7	6	8	5	6,75	6
			50%	5	6	5	7	5	6	6	6	5,25	6,25
			75%	4	6	3	6	4	7	4	6	3,75	6,25
	0,005	0,1	0%	7	6	8	6	8	7	8	7	7,75	6,5
			25%	6	5	6	6	7	6	7	6	6,5	5,75
			50%	5	5	5	6	6	7	6	6	5,5	6
			75%	5	6	5	6	6	7	6	6	5,5	6,25
	0,01	0,1	0%	6	5	7	6	8	8	8	6	7,25	6,25
			25%	6	5	6	5	7	5	7	7	6,5	5,5
			50%	5	6	5	5	6	5	6	7	5,5	5,75
			75%	5	5	5	5	6	5	6	7	5,5	5,5

Tab. 6 : Vyhodnocení spektrálního odečítání s Bartlettovým násobícím oknem

Spektrální odečítání				Hodnotící subjekty									
Typ okna	Nastavené prahy		Překrytí segmentů	Osoba č.1		Osoba č.2		Osoba č.3		Osoba č.4		Průměr	
	Ta (s)	Tb (s)		Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.	Kval.	Odruš.
Bartlettovo okno	0,0025	0,1	0%	7	9	8	8	8	9	7	8	7,5	8,5
			25%	6	8	7	8	6	9	6	7	6,25	8
			50%	5	8	5	7	5	8	5	7	5	7,5
			75%	4	8	5	7	6	8	4	7	4,75	7,5
	0,005	0,1	0%	7	8	8	7	7	8	6	8	7	7,75
			25%	6	7	7	7	7	8	6	7	6,5	7,25
			50%	5	8	6	7	6	8	5	7	5,5	7,5
			75%	5	7	5	7	6	8	6	6	5,5	7
	0,01	0,1	0%	7	7	8	7	7	8	6	8	7	7,5
			25%	6	7	7	7	7	8	6	8	6,5	7,5
			50%	6	7	7	6	7	7	7	8	6,75	7
			75%	5	6	7	6	6	7	5	7	5,75	6,5

4 ZÁVĚR

V bakalářské práci byly zkoumány metody pro odrušení řečového signálu z nahrávky, kde byla řeč rušena zvukem motoru a zároveň větrem vnikajícím do auta otevřeným oknem. Teoreticky byla rozebrána metoda odrušení pomocí lineárního filtru, metoda nazvaná spektrální odečítání, a metoda využívající adaptivního filtru. Prakticky však byly zpracovány pouze první dvě jmenované. Vyhodnocení výsledků probíhalo na základě poslechu odrušených nahrávek 4 lidmi, kteří hodnotili jak srozumitelnost, tak úspěšnost odstranění rušivých složek.

V první části praktické části byly zkoumány rozdíly různých metod návrhu filtru FIR, jenž byl navrhován jako pásmová propust nejdříve ručně a následně v prostředí Matlab metodou oken, metodou frekvenčního vzorkování a metodou nejmenších čtverců. Výsledky jednotlivých metod se od sebe téměř nelišily, proto byla pro další zkoumání použita již pouze metoda oken, jež je implementovaná v Matlabu jako funkce *fir1*. Testován byl filtr, u kterého se postupně měnil jeho řád (101, 301, 601), použitý typ násobícího okna (Pravoúhlé okno, Hammingovo okno, Hannovo okno, Bartlettovo okno) a dolní mezní frekvence propustného pásma (450Hz, 550Hz, 650Hz). Horní mezní frekvenci propustného pásma nebylo potřeba měnit, jelikož spektrum hluku způsobeného větrem bylo dostatečně vzdáleno od spektra řečového signálu a celé se nacházelo v nepropustném pásmu filtru. Z výsledku jež jsou zobrazeny v tab.1 je vidět, jak roste odrušení nežádoucích signálů se zvyšující se dolní mezní frekvencí a s řádem filtru. Bohužel s tím jak roste odrušení, klesá podobnost výsledného řečového signálu k původnímu řečovému signálu (mění se barva hlasu). Je to způsobeno překrýváním spektra zvuku motoru se spektrem řečového signálu. Jestliže je potřeba odfiltrovat ze signálu rušivou složku, která zasahuje do spektra žádoucího signálu, není tato metoda nejvhodnější, pokud se nespokojíme s určitým kompromisem. V našem případě byl kompromisem filtr s nastaveným mezním kmitočtem 450Hz, řádem 601 a Pravoúhlým, či Hammingovým násobícím oknem. Obecně platí – při volbách nižších řádů se volí dolní mezní frekvence vyšší a horní mezní frekvence zase nižší. Ale ani takovým to způsobem není možno dosáhnout stejných výsledků jako u filtru s vysokým řádem. Snížením řádu filtru totiž klesá i strmost přechodových pásem. Například pro řád 101 je již vhodnější dolní mezní kmitočet propustného pásma nastavit na hodnotu 650 Hz, ovšem při tomto nastavení již vykazuje nejlepších výsledků Hannovo okno.

V další části byla zkoumána metoda spektrálního odečítání, jež je založena na nalezení spektra šumu a jeho následném odečtením od původního signálu. Základem celé metody je správné odhadnutí spektra hluku. Jestliže je odhad nesprávný, dochází při odečítání odhadnutého spektra hluku od původního spektra signálu ke vzniku chyb, jenž mají za následek vznik tzv. hudebních tónů, jenž mohou mít ve výsledku rušivější charakter, než původní hluk. Dalším důležitým faktorem u této metody je způsob zpracování signálu tzv. segmentováním a následně roznásobením jednotlivých segmentů typem zvoleného okna. Jestliže se segmenty nepřekrývají alespoň o 50 % je důležité, aby bylo voleno Pravoúhlé násobící okno, protože jinak dochází k výrazné deformaci samotného řečového signálu (čím je menší překrytí segmentů, tím dochází k větší deformaci). Při větším překrytí segmentů (testováno na 75%) je již kvalita odrušení u všech oken obdobná, ovšem srozumitelnost zůstává pořád nejlepší u Pravoúhlého okna.

Je třeba podotknout, že spektrální odečítání je metoda vhodná především na odstranění nežádoucích složek, jež v čase svůj charakter nemění vůbec, či pouze minimálně. Jelikož byla tato metoda testována na stejné nahrávce jako byl testován filtr - s relativně rychle se měnícím rušivým zvukem, dosažené výsledky byly pouze uspokojivé. Nejlepších výsledků však bylo dosaženo při spektrálním odečítání s použitým Pravoúhlým oknem a jakýmkoliv překrýváním segmentů, čímž bylo dosaženo nejlepší kvality a podobného odrušení jako u ostatních oken při překrývání segmentů o 75 %. Ještě je nutné dodat, že tato metoda nepatří mezi nenáročná metody co se týká výpočtů (náročnost roste s překrýváním jednotlivých segmentů).

LITERATURA

- [1] BARTÁK, J. , BECHYNĚ, M., *Encyklopedický slovník*. Odeon. Praha, 1993
- [2] GOLD, B. ,MORGAN, N., *Speech and audio signal processing.*, John Wiley & Song, 2002, p.6-18.
- [3] PSUTKA, J., *Komunikace s počítačem mluvenou řečí*. Academica, Praha, 1995
- [4] HESS, W.J., *Pitch Determination of Speech Signals*. Springer, New York, 1993
- [5] ČÍŽEK, V., *Diskrétní Fourierova transformace a její použití*. SNTL, Praha, 1981
- [6] GRIFFIN, D.W., LIM, J.S., Signal estimation from modified short time Fourier transform., *IEEE Transaction on Acoustics, Speech and Signal Processing*, 1984
- [7] JAN, J., *Číslicová filtrace, analýza a restaurace signálů*. 2. vydání, VUTIUM, Brno, 2002.
- [8] SMUTNÝ, J., PAZDERA, L., *Časově frekvenční analýza reálných signálů*, ECON, Brno, 2003
- [9] HAWKSFORD, M., Time-Quantized Frequency Modulation, Time-Domain Dither, Dispersive Codes, and Parametrically Controlled Noise Shaping in SDM. *The Journal Audio Engineering Society*, Vol. 52, No.6, 2004 June
- [10] RAJNOHA, J., POLLÁK, P., *Detektory řečové aktivity na bázi keprstrální analýzy*. ČVUT Praha. 2008, p. 2-4.
- [11] HAIGH, J.H., MASON, J.S. . A voice activity detektor based on cepstral analysis. *Eurospeech '93-Proceedings of the 3rd European Conference on Speech, Comunication and Technology*. Berlin 1993, vol. 52, no. 1, p. 103-1106.
- [12] HLAVÁČEK, V., SEDLÁČEK, M., *Zpracování signálů a obrazů*. Dotisk 2. vydání, Praha, ČVUT, p. 15-35.
- [13] CARTER,B., An audio circuit collection, Part 1. *Analog Applications Journal*, Texas, 2000, p.39-43.
- [14] UHLÍŘ, J., SOVKA, P., *Číslicové zpracování signálů*, Vydavatelství ČVUT, Praha, 2002
- [15] KUECH, V., KELLERMANN, W., Partioned Block Frequency – Domain Adaptive Filter, *IEEE Transactions on Signal processing*, Vol.53, No.2, Februar 2005
- [16] LOIZOU, C.P. . *Speech enhancement Theory and practice*. New York, CRC Press, Taylor&Francis Group, 2007.
- [17] HOVORKA,J. , Moderní přístupy pro zabezpečení srozumitelnosti komunikace uvnitř bojových vojenských vozidel. *Sborník příspěvků 8. mezinárodní vědecké konference – Měření, diagnostika, spolehlivost palubních soustav letadel*. Brno 2008, p. 8-14.
- [18] VASEGHI, S.V., *Advanced Digital Signal Processing and Noise Reduction*, Chichester 2005
- [19] [81] PALIWAL, K. . Modulation Domain Spectral Subtraction for Speech Enhancement. *Interspeech 2009*. Brighton UK 2009, p. 1327-1330.
- [20] HÁJEK, K., SEDLÁČEK, J., *Kmitočtové filtry*, BEN, Praha, 2003