

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

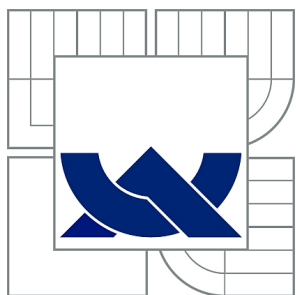
VARIABILNÍ SEGMENTACE PRO ZPRACOVÁNÍ ZVUKOVÝCH
SIGNÁLŮ

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

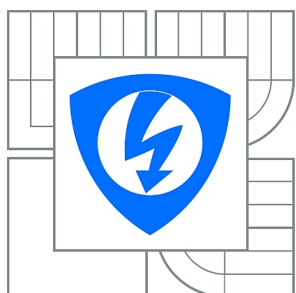
Bc. SZABOLCS GARAI

BRNO 2010



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ

ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

VARIABILNÍ SEGMENTACE PRO ZPRACOVÁNÍ ZVUKOVÝCH SIGNÁLŮ

VARIABLE SEGMENTATION FOR SOUND SIGNAL PROCESSING

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. SZABOLCS GARAI

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. MIROSLAV BALÍK, Ph.D.

BRNO 2010



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Diplomová práce

magisterský navazující studijní obor
Telekomunikační a informační technika

Student: Bc. Szabolcs Garai

ID: 78486

Ročník: 2

Akademický rok: 2009/2010

NÁZEV TÉMATU:

Variabilní segmentace pro zpracování zvukových signálů

POKYNY PRO VYPRACOVÁNÍ:

Proveďte srovnání jednotlivých běžných metod segmentace v závislosti na délce segmentu a použitém váhovacím okně při prahování koeficientů. Pro tento účel využijte databázi zvukových signálů. Realizujte variabilní segmentaci s proměnnou délkou segmentu, konstantním překryvem a s proměnnou délkou segmentu i překryvu.

DOPORUČENÁ LITERATURA:

[1] O. Julius, III. Smith Spectral Audio Signal Processing. Center for Computer Research in Music and Acoustics (CCRMA) Department of Music, Stanford University, Stanford, California 94305 USA, 2009
[online] <https://ccrma.stanford.edu/~jos/sasp/index.html>

[2] D. Rudoy, P. Basu, T. F. Quatieri, B. Dunn, and P. J. Wolfe, Adaptive Short-Time Analysis-Synthesis for Speech Enhancement. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 4905–4908, 2008

Termín zadání: 29.1.2010

Termín odevzdání: 26.5.2010

Vedoucí práce: Ing. Miroslav Balík, Ph.D.

prof. Ing. Kamil Vrba, CSc.

Předseda oborové rady

UPOZORNĚNÍ:

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

ABSTRAKT

Táto práca je zameraná na metódy používané hlavne pri filtrácii hudobných signálov – potlačenia šumu. Realizuje variabilnú segmentáciu s premennou dĺžkou segmentu, konštantným prekryvom a segmentáciu s premennou dĺžkou segmentu aj prekryvu. Tieto metódy potom porovnáva s bežnými metódami segmentácie v závislosti na dĺžke segmentu a použitom váhovacím okne pri prahovaní koeficientov. Pre tento účel využíva databázu zvukových nahrávok.

V prvej časti vysvetľuje spôsob spracovania zvukového signálu bežnou metódou segmentácie. Na základe tejto metódy pokračuje v návrhu spracovania signálu variabilnou metódou segmentácie, pre ktorú je potreba upraviť aj obálku váhovacích okien, ktorých vplyvy a vlastnosti sú vysvetlené v nasledujúcej kapitole. V praktickej časti potom podrobne popisuje implementované metódy v programovom prostredí MATLAB spolu s jednotlivými krokmi testovania. Pokračuje zoznamom priložených súborov a zhodnotením výsledkov posluchových testov.

KLÍČOVÁ SLOVA

príčítanie prekryvu, variabilná segmentácia, váhovacie okno, spracovanie zvuku, COLA, VOLA, MATLAB

ABSTRACT

This paper describes the methods used mainly in the filtration of audio signals – noise reduction. It realizes segmentation with variable length of segment, constant overlap add and segmentation with variable length of segment and overlap add. These methods are then compared with common methods of segmentation in dependence of the length of segment and used window function by the thresholding method. For this purpose it uses the database of audio records.

In the first part it describes the technique of audio signal processing with common method of segmentation. According to this method it continues in design of signal processing by variable segmentation method, in which it is needed to modify the shape of window function, which influences and attributes are explained in next chapter. In practical part it describes the implemented methods in MATLAB programming language with each steps of testing. It continues with chart of enclosed files and the evaluation of the results of hearing tests.

KEYWORDS

overlap add, variable segmentation, window function, sound signal processing, COLA, VOLA, MATLAB

GARAI, Szabolcs *Variabilní segmentace pro zpracování zvukových signálů*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2009. 58 s. Vedoucí práce byl Ing. Miroslav Balík, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Variabilní segmentace pro zpracování zvukových signálů“ jsem vypracoval samostatně pod vedením vedoucího diplomové 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 diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové 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 § 152 trestního zákona č. 140/1961 Sb.

Brno

.....

(podpis autora)

POĎAKOVANIE

Moja vďaka patrí všetkým, ktorý mi určitou formou pomohli pri riešení danej problematiky. Obzvlášť ďakujem vedúcemu práce Ing. Miroslavovi Balíkovi, Ph.D. za motiváciu, cenné rady a metodickú pomoc pri spracovaní diplomovej práce.

Brno

.....

(podpis autora)

OBSAH

Úvod	11
1 Spracovanie zvukového signálu	12
1.1 Spracovanie signálu s pričítaním prekryvu	12
1.2 Spracovanie dlhého signálu	14
1.2.1 COLA	15
1.2.2 Nestacionarita hudobného signálu	18
1.2.3 VOLA 1	18
1.2.4 VOLA 2	20
1.2.5 Tvrdé prahovanie	21
1.3 Váhovacie okná pre spektrálnu analýzu	22
1.3.1 Obdĺžnikové okno	23
1.3.2 Rodina Hammingových okien	26
1.3.3 Hannovo okno	27
1.3.4 Hammingovo okno	28
1.3.5 Rodina Blackmanových okien	29
1.3.6 Blackmanovo okno	30
1.3.7 Bartlettovo okno	30
1.3.8 Vplyv váhovacích okien	31
1.3.9 Prekryv váhovacích okien	32
2 Výsledky študentskej práce	33
2.1 Implementované metódy segmentácie	33
2.1.1 Algoritmus COLA	33
2.1.2 Algoritmus VOLA1	36
2.1.3 Algoritmus VOLA2	38
2.1.4 Tvorba váhovacích okien	40
2.1.5 Zašumenie signálu	42
2.1.6 Odhad prahov	43
2.1.7 Spôsob prahovania	43
2.2 Testovanie implementovaných metód	43
2.2.1 Zvolené zvukové nahrávky	44
2.2.2 Spôsob testovania	44
2.2.3 Výpočtová náročnosť	47
2.3 Popis priložených súborov	47
2.3.1 Spôsob značenia spracovaných nahrávok	48
2.3.2 Zoznam priložených súborov	49

3	Testovanie posluchovej kvality	51
3.1	Subjektívne metódy merania kvality	51
3.2	Metóda absolútneho hodnotenia	51
3.3	Zhodnotenie výsledkov	52
4	Záver	56
	Literatura	57
	Seznam symbolů, veličin a zkratek	58

SEZNAM OBRÁZKŮ

1.1	Bloková schéma spracovania signálu.	12
1.2	Princíp metódy COLA s Bartlettovým oknom.	15
1.3	Symetrické a periodické Bartlettovo okno, $N = 8$	16
1.4	Princíp metódy VOLA 1 s modifikovaným Bartlettovým oknom.	19
1.5	Princíp metódy VOLA 2 s modifikovaným Bartlettovým oknom.	21
1.6	Váhovanie diskretného signálu Hannovým oknom, $N = 1024$	23
1.7	Priebeh váhovacích okien, $N = 256$	24
1.8	Modulové spektrum váhovacích okien	26
2.1	Obálka Blackmanovho okna po rekonštrukcii signálu s prekryvom $O = 75\%$	35
2.2	Segmentácia vstupného signálu metódou COLA.	36
2.3	Rekonštrukcia signálu metódou COLA pre A, 50 % a B, 75 % prekryv.	37
2.4	Bloková schéma testovania implementovaných metód	45
2.5	Bloková schéma odhadu prahov	46
2.6	Bloková schéma spracovania segmentov	46
3.1	Výsledky hodnotenia zvukových nahrávok pre SNR = -25 dB	54
3.2	Výsledky hodnotenia zvukových nahrávok pre SNR = -35 dB	55

SEZNAM TABULEK

1.1	Vlastnosti váhovacích okien	32
1.2	Teoretické prekryvy váhovacích okien	32
2.1	Vstupné parametre algoritmu cola	33
2.2	Korekčné konštanty váhovacích okien pre určité prekryvy	36
2.3	Vstupné parametre algoritmu vola1	37
2.4	Vstupné parametre algoritmu vola2	39
2.5	Vstupné parametre algoritmu cw	41
2.6	Parametre zvukových nahrávok	44
2.7	Výpočtová náročnosť implementovaných metód	47
3.1	Hodnotiaca stupnica pre metódu MOS podľa normy ITU	51

ÚVOD

V oblasti spracovania hudobných a rečových signálov sú známe viaceré techniky. Jednou z metód je krátkodobá Fourierova analýza, v ktorej sa tvarom váhovacej okennej funkcie volí kompromis medzi časovým a frekvenčným rozlíšením pevne určeného časového intervalu – segmentu. Metóda vychádza z predpokladu, že signál je v tomto intervale stacionárny a jeho dĺžka sa volí v rozmedzí 15 až 30 ms.

V skutočnosti je takýto predpoklad nesprávny, pretože audio signál je obecné nestacionárny. Vhodnejším riešením je prispôbenie dĺžky segmentu podľa aktuálnej potreby. Pasáže, v ktorých sú evidentné pomalé zmeny signálu možno pokladať za generované kvázi stacionárnymi ustálenými procesmi a spracovávať ich v dlhšom časovom intervale. Naopak pasáže, v ktorých dochádza k rýchlym zmenám a prechodom, je vhodnejšie spracovávať na kratšom časovom intervale.

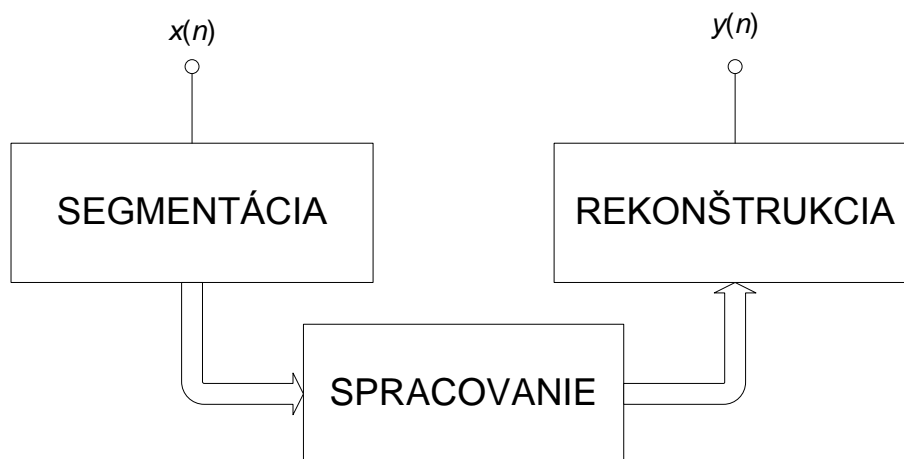
Spracovanie signálov pomocou Fourierovej transformácie je popísané v práci [9], ktorá sa zameriava hlavne na analýzu signálu s konštantnou dĺžkou segmentu. Nepresne a pevne stanovená dĺžka spôsobuje skreslenie, ktorá má podľa [1] za následok výskyt nežiaducich kmitočtov vo frekvenčnej oblasti a prejavuje sa ako hudobný šum alebo takzvané klikanie v časovej oblasti. Spomínaná práca sa zaoberá práve metódami pre potlačenie takýchto artefaktov.

Táto práca je zameraná na metódy používané hlavne pri filtrácii hudobných signálov – potlačenia šumu. Popisuje implementáciu a testovanie nových metód navrhnutých vedúcim diplomovej práce Ing. Miroslavom Balíkom, Ph.D.. Jedná sa o variabilnú segmentáciu s premennou dĺžkou segmentu, konštantným prekryvom a segmentáciu s premennou dĺžkou segmentu aj prekryvu. Tieto metódy potom porovnáva s bežnými metódami segmentácie v závislosti na dĺžke segmentu a použitom váhovacom okne pri prahovaní koeficientov. Pre tento účel využíva databázu zvukových nahrávok.

V prvej časti vysvetľuje spôsob spracovania zvukového signálu bežnou metódou segmentácie. Na základe tejto metódy pokračuje v návrhu spracovania signálu variabilnou metódou segmentácie, pre ktorú je potreba upraviť aj obálku váhovacích okien, ktorých vplyvy a vlastnosti sú vysvetlené v nasledujúcej kapitole. V praktickej časti potom podrobne popisuje implementované metódy v programovom prostredí MATLAB spolu s jednotlivými krokmi testovania. Pokračuje zoznamom priložených súborov a zhodnotením výsledkov posluchového testu.

1 SPRACOVANIE ZVUKOVÉHO SIGNÁLU

Jedným z princípov spracovania signálu digitálnou technikou je rozdelenie vstupného toku vzorkov $x(n)$ na krátke časové úseky – segmenty, ktoré sa následne spracúvajú. Spracované úseky sa znova rekonštruujú do pôvodného poradia, čím vzniká výstupný tok vzorkov $y(n)$ (Obr. 1.1). Jednotlivé operácie majú za následok určité časové oneskorenie.



Obr. 1.1: Bloková schéma spracovania signálu.

1.1 Spracovanie signálu s pričítaním prekryvu

Kapitola pojednáva o použití DFT (diskrétna Fourierova transformácia – Discrete Fourier Transform) pre implementáciu lineárneho filtrovania v kmitočtovom spektre. Ako uvádza SMITH III. [9], vzhľadom k rýchlosti FFT (rýchla Fourierova transformácia – fast Fourier transform), DFT poskytuje najefektívnejšiu implementáciu u väčšiny FIR (konečná impulzová charakteristika – finite impulse response) filtrov v oblasti spracovania audio signálov.

$$\mathbf{W}_i(\omega) = \sum_{n=-\infty}^{\infty} x(n)w(n - iR)e^{-j\omega n} \quad (1.1)$$

$$= \text{DFT}_{\omega}(x \cdot \text{SHIFT}_{iR}(w)) \quad (1.2)$$

$x(n)$... vstupný signál v čase n

$w(n)$... váhové okno

$\mathbf{W}_i(\omega)$... DFT váhovaných vzorkov času iR
 R ... dĺžka skoku, vo vzorkoch, medzi jednotlivými oknami

Pokiaľ váhujeme oknom $w(n)$ s konštantným pričítaním presahu so skokom R ,

$$\sum_{i=-\infty}^{\infty} w(n - iR) = 1, \quad \forall n \in \mathbf{Z} \quad (w \in \text{COLA}(R)), \quad (1.3)$$

potom súčet po sebe idúcich DFT v priebehu času je rovná DFT celého signálu $\mathbf{X}(\omega)$:

$$\sum_{i=-\infty}^{\infty} \mathbf{X}_i(\omega) = \mathbf{X}(\omega) = \text{DFT}_{\omega}(x) \quad (1.4)$$

Dôsledkom toho, inverzné DFT je jednoducho súčtom:

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=-\infty}^{\infty} \mathbf{X}_i(\omega) e^{-j\omega n} d\omega = \sum_{i=-\infty}^{\infty} x_i(n)$$

Modifikácia spektra je založená na násobení každého koeficientu $\mathbf{X}_i(\omega)$ frekvenčnou odozvou filtra $\mathbf{H}_i(\omega)$:

$$\mathbf{Y}_i(\omega) = \mathbf{H}_i(\omega) \mathbf{X}_i(\omega). \quad (1.5)$$

Hodnoty \mathbf{H}_i môžu byť rôzne pre každý segment rôzne, čo umožňuje realizáciu filtrov premenných v čase. Spektrum výstupného filtrovaného signálu je

$$\mathbf{Y}(\omega) = \sum_{i=-\infty}^{\infty} \mathbf{Y}_i(\omega) \quad (1.6)$$

pričom platí

$$y(n) = \sum_{i=-\infty}^{\infty} y_i(n) \quad (1.7)$$

kde aplikáciou inverznej DFT

$$y_i(n) = \text{DFT}_n^{-1}(\mathbf{Y}_i) = x_i h_i. \quad (1.8)$$

Vyššie uvedené vzťahy možno použiť aj pre efektívne FFT, čo možno tiež chápať ako vzorkované DFT. Pre kvalitné spektrálne modifikácie by hustota vzorkovania mala byť čo najvyššia. V konečnom dôsledku vyzerá FFT konvolúcia nasledovne:

$$y = \sum_{i=-\infty}^{\infty} \text{SHIFT}_{iR} \left(\text{FFT}_N^{-1} \{ \mathbf{H}_i \text{FFT}_N [\text{SHIFT}_{-iR}(x) \cdot w_N] \} \right) \quad (1.9)$$

Súčet cez i možno interpretovať ako pridanie separovaných filtrovaných segmentov $y_i = x_i h_i$. Vzhľadom k tejto filtrácii sa segmenty musia vždy prekrývať, aj pri použití obdĺžnikového okna. V dôsledku toho sa tento systém nazýva ako OLA (pričítanie presahu – OverLap Add) procesor z čoho vyplýva aj COLA (konštantné pričítanie presahu – constant OverLap Add) – OLA procesor s konštantnou dĺžkou váhovacích okien. V princípe je to sekvencia FFT transformácií, ktoré možno modifikovať, potom inverzne transformovať a sčítať.

1.2 Spracovanie dlhého signálu

Vykonávanie konvolúcie dvoch segmentov rovnakej dĺžky pomocou FFT je efektívne, v reálnom živote sa však vyskytujú situácie, ako napr. jeden zo signálov je veľmi dlhý, alebo je potrebné vykonávať operácie v reálnom čase (vstupom je signál nekonečnej dĺžky), uvádza SMITH III. [9]. Pri priamej konvolúcii tieto situácie nie je potrebné riešiť. Napríklad pri kauzálnom systéme s konečnou impulznou odozvou (FIR) h dĺžky L je potrebné zaznamenať pozledných $L - 1$ vzorkov vstupného signálu x na výpočet nasledujúceho výstupného vzorku, pretože

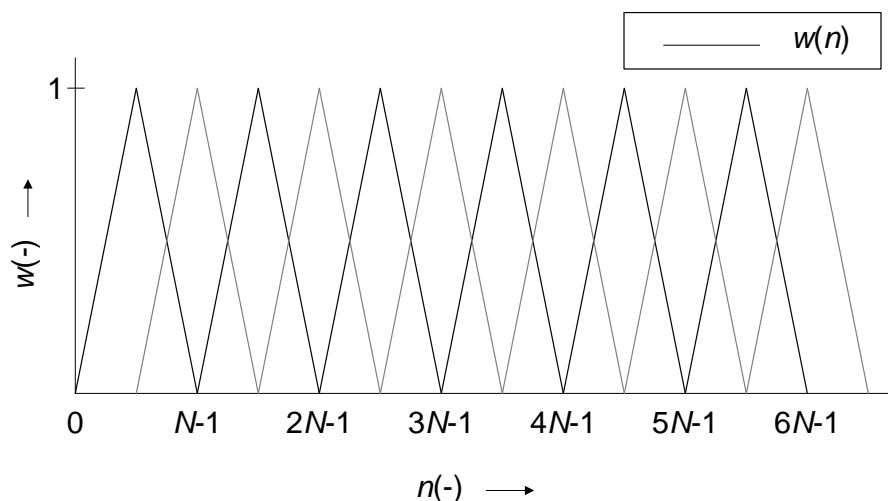
$$\begin{aligned} y(n) &= (h * x)(n) = \sum_{i=0}^n h(i)x(n-i) \\ &= h(0)x(n) + h(1)x(n-1) + \dots + h(L-1)x(n-L+1) \end{aligned}$$

V každom okamihu n sa výstup $y(n)$ počíta ako lineárna kombinácia aktuálneho vstupného vzorku $x(n)$ a aktuálnym stavom filtra $\{x(n-1), \dots, x(n-L+1)\}$. Pokiaľ je teda vstupný signál príliš dlhý a požaduje sa výhoda rýchlej FFT konvolúcie, vstupný signál x sa rozdelí do segmentov a konvolúcia sa vykoná na každom segmente zvlášť. Výstupom je potom súčet oddelených filtrovaných segmentov. Prekrýv segmentov sa volí kôli zvlenueniu filtra. Pre filter s nulovou fázou sa každý segment prekrýva s oboma susediacimi segmentami. Pre kauzálne filtre sa každý segment prekrýva iba so susedným segmentom vpravo (nasledujúci segment v čase). Fakt, že sa jednotlivé segmenty prekrývajú a musia byť sčítané je základom pre metódu COLA pracujúcu s dlhými signálmi.

Ak sa jednoduchá FFT konvolúcia vykonáva medzi signálom x a FIR filtrom h , nie je dôvod použiť váhové okno na každý vstupný segment. V takomto prípade je možné uvažovať, že vstupný signál je spracovaný obdĺžnikovým oknom. Ak je dĺžka okna N potom postupuje N vzorkov pre nasledujúci rámec (veľkosť skoku = N vzorkov). Vstupné segmenty sa neprekrývajú. Na druhú stranu, ak sa vykonávajú spektrálne modifikácie, alebo sa filter mení každým segmentom ($h \leftarrow h_i$ kde i je číslo segmentu) je potreba použiť niektoré z váhovacích okien a menšiu hodnotu skoku ako bude vysvetlené ďalej.

1.2.1 COLA

Metódou COLA (Obr. 1.2) sa rozumie delenie vstupného signálu x na segmenty použitím váhového okna w s konečnou dĺžkou N .



Obr. 1.2: Princíp metódy COLA s Bartlettovým oknom.

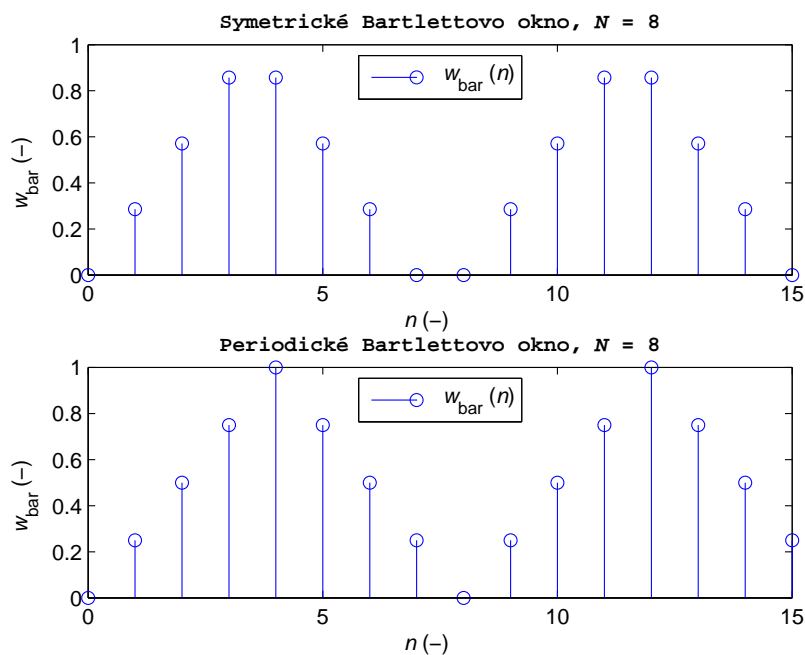
Pre metódu COLA sa aktuálne váhové okno definuje zo symetrického váhového okna $w_T(n)$

$$w_T(n) \quad n = 0, 1, \dots, N \quad (1.10)$$

ako periodické váhové okno (Obr. 1.3) s postupnosťou čísiel

$$w(n) = w_T(n) \quad n = 0, 1, \dots, N - 1 \quad (1.11)$$

kde N je aktuálna dĺžka váhového okna. Symetrické váhové okno sa získa zrkadlovým otočením jednostranného okna, pre metódy segmentácie sa ale vždy



Obr. 1.3: Symetrické a periodické Bartlettovo okno, $N = 8$

používa periodické váhové okno, ktoré je nesymetrické. Metóda COLA pracuje vo väčšine prípadov s $\frac{1}{2}$ a $\frac{3}{4}$ prekryvom, pri Blackmanovom okne s $\frac{2}{3}$ prekryvom. Aktuálny segment sa získa pomocou

$$x_i(n) = x(n)w(n - iR), \quad n \in (-\infty, +\infty) \quad (1.12)$$

$x_i(n)$... aktuálny segment

i ... poradie segmentu

R ... hodnota skoku (skok segmentu),

alebo

$$x_i = x \cdot \text{SHIFT}_{iR}(w) \quad (1.13)$$

pričom

$$R = N(1 - O), \quad (1.14)$$

O ... hodnota prekryvu
 N ... dĺžka segmentu.

Veľkosť skoku je počet vzorkov medzi počiatočnými vzorkami susedných segmentov, čo je vlastne vzdialenosť medzi jednotlivými segmentami. Pri takomto krokování spektrálneho spracovania je potrebné spätne rekonštruovať z každého jednotlivého prekryvajúceho sa segmentu pôvodný signál, ideálne sčítaním v ich pôvodných časových pozíciách. Preto platí

$$\begin{aligned} x(n) &= \sum_{i=-\infty}^{\infty} x_i(n) \\ &= \sum_{i=-\infty}^{\infty} x(n)w(n - iR) \\ &= x(n) \sum_{i=-\infty}^{\infty} w(n - iR), \end{aligned}$$

avšak výraz $x = \sum_i x_i$ platí iba vtedy, ak

$$\sum_{i \in \mathbf{Z}} w(n - iR) = 1, \forall n \in \mathbf{Z}. \quad (1.15)$$

Pri sčítavaní presahov je potrebné dosiahnuť čo najmenšie zvlnenie spracovaného signálu. Toto zvlnenie je priamo závislé na zvolenej okennej funkcii pre váhovanie segmentu signálu pred aplikáciou FFT. Všetky váhovacie okná s konštantným prekryvom dokonale rekonštruujú vstupný signál x zo segmentov $x_i = x \cdot \text{SHIFT}_{iR}(w)$ pomocou metódy COLA. Typ zvoleného váhovacieho okna nie je obmedzený, jedine tým, že hodnota skoku je konštantná.

Vo svojej práci SMITH III. [9] poukazuje na skutočnosť, že pokiaľ sa implementuje jednoduchá FIR filtrácia (nemenná v čase) a nie je potrebné pracovať s DTFT, potom je viac efektívne použiť obdĺžnikové okno s hodnotou skoku $R = N$ a nastaviť $N = F - L + 1$, kde L je dĺžka filtru h a F je vhodná veľkosť FFT.

COLA príklady

- obdĺžnikové okno s 0 % prekryvom (hodnota skoku $R \approx N$ dĺžka okna)
- obdĺžnikové okno s 50 % prekryvom ($R \approx \frac{N}{2}$)
- obdĺžnikové okno so 75 % prekryvom ($R \approx \frac{N}{4}$)
- Bartlettovo okno s 50 % prekryvom ($R \approx \frac{N}{2}$)
- Hammingovo okno s 50 % prekryvom ($R \approx \frac{N}{2}$)

- Hammingovo okno so 75 % prekryvom ($R \approx \frac{N}{4} \approx 25$ % skok)
- Hannovo okno s 50 % prekryvom ($R \approx \frac{N}{2}$)
- Blackmanovo okno s $\frac{2}{3}$ prekryvom ($R \approx \frac{N}{3}$ skok)

1.2.2 Nestacionarita hudobného signálu

Prvé metódy odšumovania zvukových signálov vychádzajú z predpokladu, že signál je stacionárny v špecifickom časovom intervale. Signál, ktorý má byť spracovaný s takýmto predpokladom sa v časovej doméne rozdelí na sériu časových intervalov – segmentov s konštantnou dĺžkou. DSP (číslkové spracovanie signálov – Digital Signal Processing) algoritmus potom spracováva každý segment oddelene.

Faktom však je, že audio signál (rozumie sa rečový aj hudobný) je obecné nestacionárny. V skutočnosti nie je možné predpokladať stacionaritu v každom z týchto segmentov. Práve preto spracovávať signál s pevne danou dĺžkou segmentu nie je vhodné, pretože transformované koeficienty sú potom obohatené o nechcené kmitočty, ktorými sa rekonštruovaný signál skreslí (napr. hudobný šum).

Jedným z riešení je voliť pre každý úsek signálu vhodnú dĺžku segmentu podľa potreby. Intervaly, v ktorých sú evidentné pomalé zmeny signálu možno pokladať za kvázi stacionárne a spracovávať ich s dlhými segmentami (napr. 4096 alebo 8192 vzorkov). Naopak, intervaly, v ktorých dochádza k rýchlym zmenám, je vhodné spracovávať s kratšími segmentami (napr. 256 alebo 512 vzorkov). Nasledujúce kapitoly popisujú dva spôsoby riešenia takouto metódou.

1.2.3 VOLA 1

Metóda VOLA 1 (Obr. 1.4) spočíva vo variabilnej zmene dĺžky segmentu, ktorého hodnoty sú nadefinované ako postupnosť dĺžok N_i . Následkom toho sú dĺžky segmentov rozdielne v každom iteračnom kroku. Prekryv jednotlivých segmentov je konštantný a určuje sa pred začiatkom segmentácie.

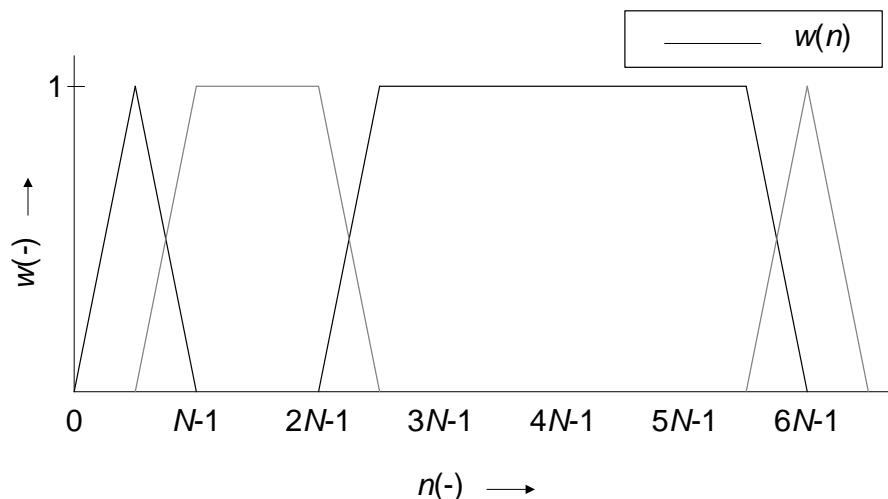
Váhovacie okno vychádza z váhovacieho okna metódy COLA (1.11) a je nutné ho rozdeliť na tri časti, na ľavú stranu, strednú časť a pravú stranu

$$w_{SL}(n) \quad n = 0, 1, \dots, N_{SL} - 1 \quad (1.16)$$

$$w_C(n) \quad n = 0, 1, \dots, N_C - 1 \quad (1.17)$$

$$w_{SR}(n) \quad n = 0, 1, \dots, N_{SR} - 1 \quad (1.18)$$

kde N_{SL} je dĺžka ľavej strany okna, N_C je dĺžka strednej časti a N_{SR} je dĺžka pravej strany okna. Ďalej pre VOLA 1 platí



Obr. 1.4: Princíp metódy VOLA 1 s modifikovaným Bartlettovým oknom.

$$N_{\text{SL}} = N_{\text{SR}} = N_{\text{s}} = \text{konst.}$$

Potom je váhové okno $w(n)$ definované pomocou svojich jednotlivých častí ako

$$w(n) = w_{\text{SL}}(n) + w_{\text{C}}(n - N_{\text{s}}) + w_{\text{SR}}(n - N_{\text{s}} - N_{\text{C}}) \quad (1.19)$$

$$n = 0, 1, \dots, N - 1$$

kde pre dĺžku aktuálneho váhového okna platí vzťah

$$N_i = N_{\text{Ci}} + 2N_{\text{s}} \quad i = 0, 1, \dots, I - 1$$

kde i je index aktuálneho segmentu a I je celkový počet segmentov vzniknutých behom segmentácie. Pri metóde VOLA 1 možno použiť váhové okná iba pre $\frac{1}{2}$ prekryv, vyššie hodnoty prekryvu by zasahovali do ostatných častí váhových okien, tým pádom nemožno použiť Blackmanovo okno, ktoré pracuje s $\frac{2}{3}$ prekryvom. Proces segmentácie VOLA 1 je potom definovaný ako

$$x_i(n) = w_i(n)x \left(n + \sum_{j=1}^i N_{\text{C}(j-1)} + iN_{\text{s}} \right) \quad (1.20)$$

$$\begin{aligned}
x_i(n) &= w_{SLi}(n)x \left(n + \sum_{j=1}^i N_{C(j-1)} + iN_s \right) + \\
&+ w_{Ci}(n - N_s)x \left(n + \sum_{j=1}^i N_{C(j-1)} + iN_s \right) + \\
&+ w_{SRi}(n - N_s - N_C)x \left(n + \sum_{j=1}^i N_{C(j-1)} + iN_s \right)
\end{aligned}$$

$$n = 0, 1, \dots, N_i - 1, i = 0, 1, \dots, I - 1$$

kde N_i je dĺžka aktuálneho segmentu. Rekonštrukcia signálu je definovaná ako

$$y(n) = \sum_{i \in \mathbf{Z}} x_i \left(n - \sum_{j=1}^i N_{C(j-1)} - iN_s \right), \forall n \in \mathbf{Z}, \quad (1.21)$$

1.2.4 VOLA 2

Metóda VOLA 2 (Obr. 1.5) je založená na variabilnej zmene veľkosti prekryvu, ktorého hodnoty sú definované ako postupnosť prekryvov. Každou iteráciou sa mení aj dĺžka segmentu (v krokoch, kde bola vykonaná zmena veľkosti prekryvu). Tým pádom pri zmene veľkosti prekryvu vznikajú aj nesymetrické okná (vplývajúce na fázové spektrum). Takéto okná sú nežiaduce, avšak potrebné pre zmenu veľkosti segmentu.

Aktuálne váhové okno $w(n)$ pre metódu VOLA 2 je definované rovnako ako pri metóde COLA (1.11) a je nutné ho rozdeliť na dve polovice, na ľavú a pravú polovicu

$$w_L(n) \quad n = 0, 1, \dots, N_L - 1 \quad (1.22)$$

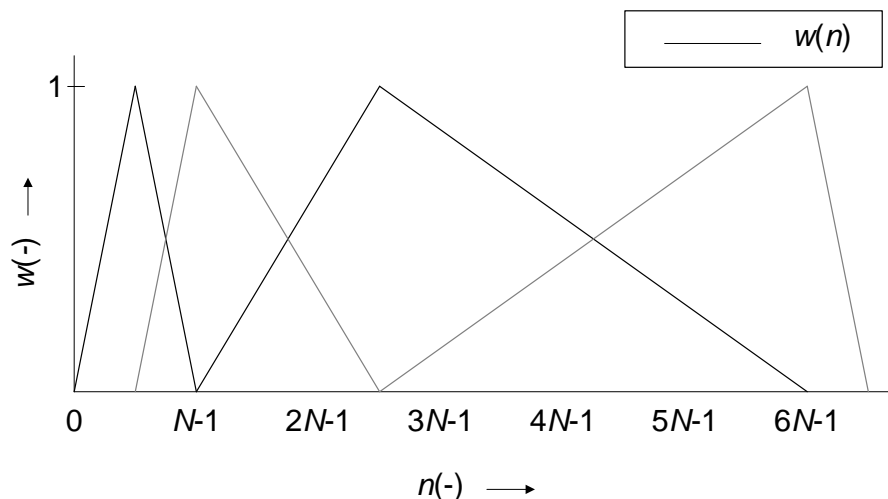
$$w_R(n) \quad n = 0, 1, \dots, N_R - 1 \quad (1.23)$$

kde N_L je dĺžka ľavej polovice okna a N_R je dĺžka pravej polovice okna. Potom je váhové okno definované pomocou svojich dvoch polovic ako

$$w(n) = w_L(n) + w_R(n - N_L) \quad n = 0, 1, \dots, N - 1 \quad (1.24)$$

keď pre dĺžku aktuálneho okna a dĺžky polovic susediacich okien platia vzťahy

$$N_i = N_{Li} + N_{Ri}, N_{Li} = N_{R(i-1)}, N_{Ri} = N_{L(i+1)} \quad i = 0, 1, \dots, I - 1 \quad (1.25)$$



Obr. 1.5: Princíp metódy VOLA 2 s modifikovaným Bartlettovým oknom.

kde i je index aktuálneho segmentu a I je celkový počet segmentov vzniknutých behom segmentácie. Rovnako, ako pri metóde VOLA 1, aj pri metóde VOLA 2 možno použiť váhové okná iba pre $\frac{1}{2}$ prekryv, vyššie hodnoty prekryvu by zasahovali do ostatných častí váhovacích okien, tým pádom nemožno použiť Blackmanovo okno, ktoré pracuje s $\frac{2}{3}$ prekryvom.

Proces segmentácie VOLA2 je potom definovaný ako

$$x_i(n) = w_i(n)x \left(n + \sum_{j=1}^i N_{L(j-1)} \right) \quad (1.26)$$

$$= w_{Li}(n)x \left(n + \sum_{j=1}^i N_{L(j-1)} \right) + w_{Ri}(n - N_{Li})x \left(n + \sum_{j=1}^i N_{L(j-1)} \right) \quad (1.27)$$

$$n = 0, 1, \dots, N_i - 1, i = 0, 1, \dots, I - 1$$

kde N_i je dĺžka aktuálneho segmentu. Rekonštrukcia signálu je definovaná ako

$$y(n) = \sum_{i \in \mathbf{Z}} x_i \left(n - \sum_{j=1}^i N_{L(j-1)} \right), \forall n \in \mathbf{Z}, \quad (1.28)$$

1.2.5 Tvrde prahovanie

Ako uvádza BALÍK [1], metódy pre potlačenie šumu zvukových signálov majú za následok nežiadúce zvukové artefakty. Jedna z metód, spôsobujúca artefakty, je

tvrdé prahovanie DFT koeficientov. Tvrdé prahovanie nuluje koeficienty, ktoré sa nachádzajú pod úrovňou zvoleného prahu. Po spätnom transformovaní do časovej oblasti spôsobuje medzi susediacimi segmentami diskontinuitu, čo sa prejaví jednak ako hudobný šum alebo takzvané klikanie. Pri určovaní prahu sa volí kompromis medzi mierou potlačenia šumu a mierou skreslenia výsledného signálu.

1.3 Váhovacie okná pre spektrálnu analýzu

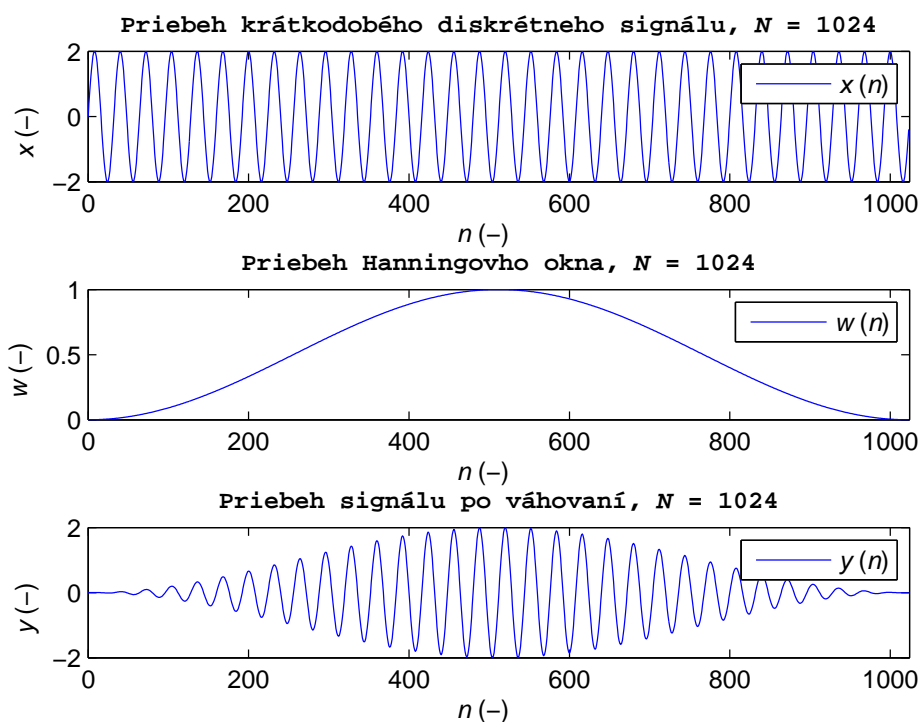
Vymedzenie signálu na časovej ose sa nazýva váhovaním. Časová funkcia, ktorou sa analyzovaný signál násobí (váhuje), sa nazýva váhovacím oknom. Toto okno sa posúva po časovej ose a pre každú polohu sa aplikuje výpočet spektra. Tento proces sa dá jednoducho predstaviť ako nekonečné opakovanie okna na časovej ose. Ideálny prípad nastáva, ak vstupný signál je periodický a dĺžka časového okna je rovná jednej perióde signálu. Potom vypočítané spektrálne čiary po prevode do frekvenčnej oblasti súhlasia s jeho skutočnými harmonickými zložkami.

V spektrálnej analýze prirodzene sa vyskytujúcich audio signálov sa takmer vždy analyzuje krátky úsek signálu, než celý signál. Je tomu tak z mnohých dôvodov. Snáď najviac podstatné je, že ucho podobne ako Fourierova analýza, analyzuje naraz len krátky segment audio signálu v čase (rádovo 10 – 20 ms). Preto je pri vykonávaní spektrálnej analýzy majúč časovo-frekvenčné rozlíšenie porovnateľné s ľudským sluchom potrebné obmedziť segment podľa času. Spôsob extrahovania krátkeho časového segmentu dĺžky N z dlhšieho signálu je násobenie pomocou váhovacieho okna (Obr. 1.6).

Digitálna analýza z dôvodu rýchleho výpočtu spektra používa algoritmus FFT. Vo svojej najjednoduchšej podobe pracuje s dĺžkou časového okna vyjadrenou počtom vzorkov v mocninách čísla 2. Diskontinuity vzniknuté pevným vymedzením dĺžky okna (COLA) sa eliminujú jeho vhodným tvarom, ktorý k okrajom okna postupne znižuje veľkosť vzorkov, najčastejšie na nulovú hodnotu. Použitím vhodného okna sa znižuje úroveň skresľujúcich frekvenčných zložiek výkonového spektra, ale mení sa aj fázové spektrum.

Hlavným prínosom pri výbere dobrého váhovacieho okna pre spektrálnu analýzu je minimalizácia bočných lalokov (filtračné vlastnosti), ktoré spôsobujú v odhadovanom spektre výskyt nežiadúcich kmitočtov a tiež zachovanie fázového spektra. Existuje mnoho typov váhovacích okien a každé z nich je vhodné pre rôzne situácie a analytické techniky.

Predovšetkým sa jedná o šírku hlavného laloku a potlačenie takzvaných postranných lalokov. Čím širší je hlavný lalok, tým väčšie je prelievanie energie spektra a tým pádom aj nižšia ostrosť spektra. Ak je lalok užší, má za následok viac postranných



Obr. 1.6: Váhovanie diskřetného signálu Hannovým oknom, $N = 1024$

lalokov a tým pádom aj viac nežiaducích kmitočtov v spektre.

V tejto kapitole je s prehľadom popísané obdĺžnikové okno nasledované niekoľkými ďalšími druhmi váhovacích okien, ako Bartlettovo, Hammingovo, Hannovo a Blackmanovo okno (Obr. 1.7).

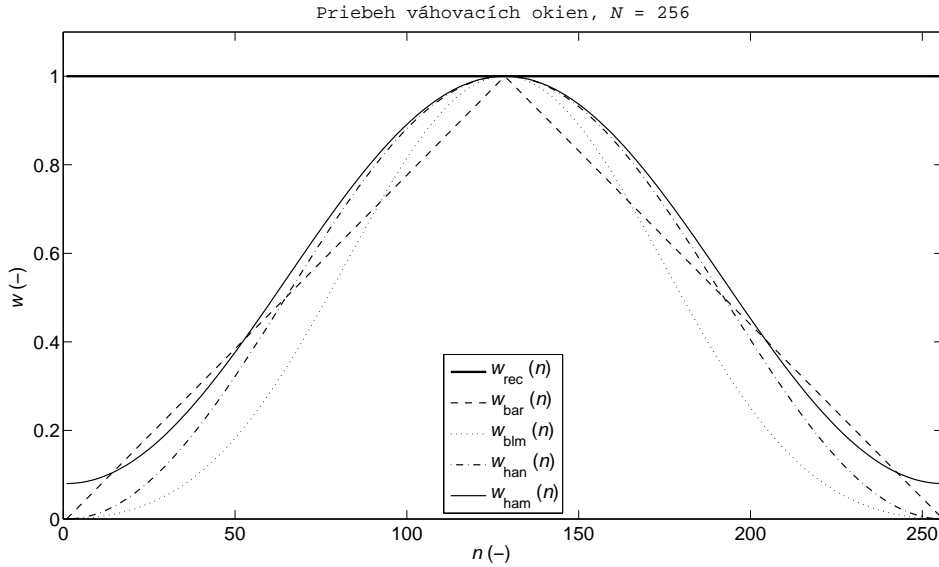
1.3.1 Obdĺžnikové okno

Obdĺžnikové okno (Obr. 1.7) patrí medzi najjednoduchšie a vyjadruje všeobecnú okennú funkciu, ktorá nemení svoje funkčné hodnoty. Pre transformovanie je vhodné iba v prípade, že je zaručená rovnaká dĺžka segmentu ako je celočíselný násobok periód analyzovaných kmitočtov.

Obdĺžnikové okno s nulovým centrováním je v práci Julius O. Smith III. [9] definované ako

$$w_{\text{REC}}(n) = \begin{cases} 1, & -\frac{N-1}{2} \leq n \leq \frac{N-1}{2} \\ 0, & \text{pre ostatné } n \end{cases}, \quad (1.29)$$

kde N je dĺžka okna vo vzorkoch (pre nepárny predpoklad). Niekedy je vhodné



Obr. 1.7: Pribeh váhovacích okien, $N = 256$

definovať okná tak, aby zisk ich jednosmernej zložky bol rovný 1, v takom prípade sa vyššie uvedená definícia násobí s $\frac{1}{N}$. Pribeh okna pre dĺžku $N = 256$ je znázornený na grafe (Obr. 1.7).

Pribeh vo frekvenčnej doméne sa zistí aplikáciou DFT (diskrétna Fourierova transformácia – Discrete Fourier Transform) na okno:

$$\mathbf{W}(\omega) = \text{DFT}_\omega(w) = \sum_{n=-\infty}^{\infty} w(n)e^{-j\omega n}, \omega \in [-\pi, \pi] \quad (1.30)$$

$$= \sum_{n=-\frac{N-1}{2}}^{\frac{N-1}{2}} e^{-j\omega n} = \frac{e^{j\omega \frac{N-1}{2}} - e^{-j\omega \frac{N+1}{2}}}{1 - e^{-j\omega}} \quad (1.31)$$

kde posledný riadok je odvodený pomocou uzavretého tvaru geometrickej rady:

$$\sum_{n=L}^U r^n = \frac{r^L - r^{U+1}}{1 - r}. \quad (1.32)$$

Z lineárnej fázy $e^{-j\omega_k \Delta}$ (fáza je lineárnou funkciou frekvencie) vyplýva

$$\mathbf{W}(\omega) = \frac{e^{-j\omega \frac{1}{2}}}{e^{-j\omega \frac{1}{2}}} \left[\frac{e^{j\omega \frac{N}{2}} - e^{-j\omega \frac{N}{2}}}{e^{j\omega \frac{1}{2}} - e^{-j\omega \frac{1}{2}}} \right] \quad (1.33)$$

$$= \frac{\sin\left(N\frac{\omega}{2}\right)}{\sin\left(\frac{\omega}{2}\right)} = N \cdot \text{asinc}_N(\omega) \quad (1.34)$$

kde $\text{asinc}_N(\omega)$ označuje alias sinc funkciu:

$$\text{asinc}_N(\omega) = \frac{\sin\left(\frac{N\omega}{2}\right)}{N \cdot \sin\left(\frac{\omega}{2}\right)} \quad (1.35)$$

(tiež nazývanú Dirichletova funkcia alebo periodická sinc funkcia). Tento výsledok platí pre obdĺžnikové okno centrované na nulu. Pre kauzálny prípad sa lineárna fáza javí:

$$\mathbf{W}^c(\omega) = e^{-j\frac{N-1}{2}\omega} \cdot N \cdot \text{asinc}_N(\omega) \quad (1.36)$$

Pojem alias sinc funkcie poukazuje na skutočnosť, že sa môže získať zo sinc funkcie $\text{sinc}(\omega) = \frac{\sin(\omega)}{\omega}$ aliasingom sinc $\frac{\omega N}{2}$ na bloku N vzoriek. (V časovej oblasti je amplitúda $\frac{1}{N}$ kontinuálneho obdĺžnikového pulzu od času $t = -\frac{N}{2}$ po $t = \frac{N}{2}$ jednoducho vzorkovaná cez celé čísla.)

Vzhľadom k tomu, že vzorkovacia frekvencia ide do nekonečna, aliasovaná sinc funkcia sa priblíži k sinc funkcii

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \quad (1.37)$$

Konkrétne,

$$\lim_{\substack{T \rightarrow 0 \\ NT = \tau}} \text{asinc}_N(\omega T) = \text{sinc}(\tau f), \quad (1.38)$$

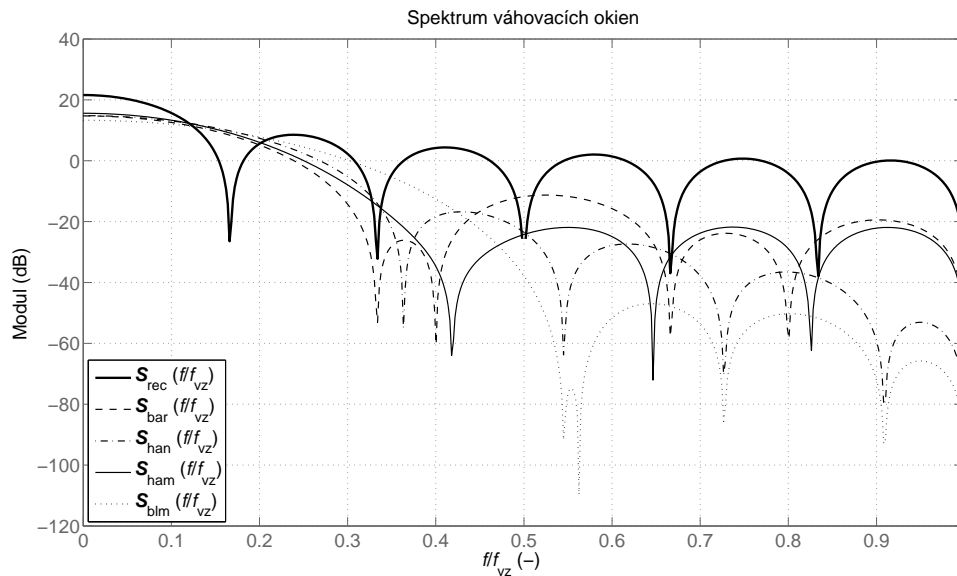
kde $\omega = 2\pi f$.

Fáza pravouhlého okna transformácie $\mathbf{W}(\omega)$ je nulová pre $|\omega| < \frac{2\pi}{N}$, čo je šírka hlavného laloku. To je aj dôvod, prečo sú váhové okná centrované na nulu nazývané oknami s nulovou fázou. Kým sa fáza strieda medzi 0 a π radiánov, π hodnoty sa vyskytujú len v rámci bočných lalokov, ktoré sú bežne zanedbávané (váhové okno je zvyčajne konštruované tak, aby zabezpečilo, že všetky postranné laloky možno zanedbať).

Pri práci so zvukom sa modul transformovaného okna typicky zobrazuje v decibelovej škále (dB), ako vidno na grafe (Obr. 1.8). Obyčajne sa veľkosť vrcholu normalizuje na 0 dB.

Vlastnosti Obdĺžnikového okna

- nulové prechody na celočíselných násobkoch $\Omega_N = \frac{2\pi}{N}$,



Obr. 1.8: Modulové spektrum váhovacích okien

- šírka hlavného laloku $2\Omega_N = \frac{4\pi}{N}$,
- zväčšovaním N sa hlavný lalok zužuje (lepšie frekvenčné rozlíšenie),
- hodnota N nemá vplyv na výšku postranných lalokov (Gibbsov jav),
- prvý postranný lalok má pokles iba -13 dB voči hlavnému laloku,
- bočné laloky klesajú približne -6 dB na oktávu,
- skoková zmena funkčných hodnôt medzi 0 a 1,
- prekryv $O = 0,75$ alebo $O = 0,5$

1.3.2 Rodina Hammingových okien

Hannovo a Hammingovo okno vychádzajú z univerzálnej Hammingovej rodiny okien. Okná sú konštruované pridaním jednej periódy funkcie cosínus k obdĺžnikovému oknu. Výhodou prídania cosínového segmentu sú nižšie postranné laloky. Cenou za túto výhodu je, že hlavný lalok sa zdvojnásobí v šírke. Najznámejšími členmi Hammingovej rodiny sú Hannovo a Hammingovo okno.

Základná myšlienka spočíva v sčítaní troch funkčných priebehov, uvádza SMITH III. [9]. Prvá vychádza zo zmenšenej transformácie obdĺžnikového okna

$$0,5 \cdot \mathbf{W}(\omega) = 0,5 \cdot N \cdot \text{asinc}_N(\omega) \quad (1.39)$$

a zvyšné dva sú posuny rovnakej funkcie

$$0,25 \cdot \mathbf{W}(\omega \pm \Omega_N). \quad (1.40)$$

Ich súčtom sa dosiahne to, že sa zrušia niektoré postranné laloky, avšak šírka hlavného laloku sa zdvojnásobí. Transformáciu pravouhlého okna (nulová fáza, funkčná hodnota rovná jednej) je možné popísať ako

$$\mathbf{W}(\omega) = \alpha \mathbf{W}(\omega) + \beta \mathbf{W}(\omega - \Omega_N) + \beta \mathbf{W}(\omega + \Omega_N) \quad (1.41)$$

kde $\alpha = \frac{1}{2}$ a $\beta = \frac{1}{4}$.

Pre každé $x \in \mathbf{C}^N$ a integrál Δ platí

$$\text{DFT}_k[\text{SHIFT}_\Delta(x)] = e^{-j\omega_k \Delta} \mathbf{X}(k). \quad (1.42)$$

Použitím tohto teorému je možné inverznou transformáciou rovnice získať

$$w_H(n) = \alpha w_{\text{REC}}(n) + \beta e^{-j\Omega_N n} w_{\text{REC}}(n) + \beta e^{j\Omega_N n} w_{\text{REC}}(n), \quad (1.43)$$

alebo

$$w_H(n) = w_{\text{REC}}(n) \left[\alpha + 2\beta \cos\left(\frac{2\pi n}{N}\right) \right], \quad (1.44)$$

kde $w_H(n)$ charakterizuje rodinu Hammingových okien.

Výberom rôznych parametrov pre α a β sa získajú mnohé váhové okná, z ktorých majú niektoré svoje vlastné pomenovania.

Vlastnosti rodiny Hammingových okien

- okná sú zložené z obdĺžnikového okna plus pozdvihutej funkcie cosínus (tá má jednu periódu cez okno),
- šírka hlavného laloku $4\Omega_N$ radiánov na úsek, kde $\Omega_N = \frac{2\pi}{N}$
- symetrické (\Rightarrow nulová alebo lineárna fáza),
- pozitívne (na základe α a β),
- nulové prechody v transformácii okna na intervaloch Ω_N po stranách hlavného laloku.

1.3.3 Hannovo okno

Hannovo okno (Obr. 1.7) je podľa SMITH III. [9] definované na základe nastavení $\alpha = \frac{1}{2}$ a $\beta = \frac{1}{4}$:

$$w_{\text{HAN}}(n) = w_{\text{REC}}(n) \left[\frac{1}{2} + \frac{1}{2} \cos(\Omega_N n) \right] = w_{\text{REC}}(n) \cos^2\left(\frac{\Omega_N}{2} n\right) \quad (1.45)$$

Okno je špeciálny prípad sínusového okna a je často používaným oknom s vyhovujúcim kompromisom medzi ostrosťou spektra a potlačenia nežiadúcich. Postranné laloky sú malé a hlavný lalok je úzky. Veľmi často používané v oblasti krátkodobej analýzy. Pri analýze sa používa s prekryvom $O = 0,5$.

Postranné laloky váhového okna klesajú približne -18 dB na oktávu. Okrem značne zrýchlenému poklesu má prvý postranný lalok pokles z -13 dB (obdĺžnikové okno) dole na $-31,5$ dB. Šírka hlavného laloku je samozrejme dvojnásobná oproti obdĺžnikovému oknu (Obr. 1.8).

Vlastnosti Hannovho okna

- pomalý prechod na nulu pri koncových hodnotách,
- šírka hlavného laloku $4\Omega_N$,
- prvý postranný lalok má pokles -31 dB voči hlavnému laloku,
- bočné laloky klesajú približne -18 dB na oktávu (ako $T \rightarrow 0$),
- prekryv $O = 0,5$

1.3.4 Hammingovo okno

Hammingovo okno (Obr. 1.7) je podľa SMITH III. [9] určené voľbou α a $\beta = 1 - \alpha$ s účelom vyrušenia najväčšieho postranného laloku. Takýto výsledok sa dosiahne voľbou

$$\begin{aligned}\alpha &= \frac{25}{46} \approx 0,54 \\ \beta &= \frac{1 - \alpha}{2} = \frac{21}{46} \approx 0,46.\end{aligned}$$

Definícia okna po dosadení (1.44) znie

$$w_{\text{HAM}}(n) = w_{\text{REC}}(n) \left[0,54 + 2(0,46) \cos\left(\frac{2\pi n}{N}\right) \right]. \quad (1.46)$$

Pri analýze sa volí polovičný prekryv. Postranný lalok má potom približne $-42,76$ dB pokles. Táto voľba je veľmi blízko k tej, ktorá minimalizuje vrchol bočného laloku až na $-43,19$ dB – najnižšia možná hodnota v rámci generalizovanej rodine Hammingových okien ($\alpha = 0,53836\dots$).

Transformovaním okna pomocou DFT je aj toto okno, podobne ako Hannovo okno, jednou periódou pozdvihutej funkcie cosínus. Je však tak vysoká, že jej negatívne vrcholy sú nad nulou okno má diskontinuitu (nespojitosť) v amplitúde opúšťajúc okno. To má za následok veľmi nízky pokles postranného laloku. Na druhú stranu, v najhoršom prípade klesne postranný lalok na -41 dB, čo je práve sieľom tohto typu

okennej funkcie. Je to o 10 dB lepšie ako v prípade Hannovho okna a o 28 dB lepšie v porovnaní s obdĺžnikovým oknom. Šírka hlavného laloku je približne $4\Omega_N$ široká, tak ako pri všetkých oknách spadajúcich do tejto rodiny.

Bočné laloky v blízkosti hlavného laloku sú silne tvarované optimalizáciou. Pokles takmer -6 dB na oktávu sa vyskytuje iba na vnútornom intervale spektra, medzi hlavným lalokom a polovicou vzorkovacej frekvencie. Prejaví sa to hlavne pri väčšej hodnote dĺžky váhovaciego okna N , keď postranné laloky zaberajú menší kmitočtový interval.

Vzhľadom k tomu, že postranné laloky sú položené tak nízko, je toto okno vhodnou voľbou pre menej presné systémy, ako napríklad 8-bitové systémy, pretože pri nich je zriedka dôvod požadovať postranné laloky ležiace hlboko pod kvantizačným šumom. Tento druh okna bol často používaný v telekomunikačnej technike, keďže 8-bitové kodeky boli štandardom mnoho rokov. Pre vyššiu kvalitu spracovania zvukového signálu sú požadované kvalitnejšie typy okien, najmä ak pôsobia ako lowpass filtre.

Vlastnosti Hammingovho okna

- diskontinuálny pokles na nulu v koncových bodoch,
- bočné laloky sú bližšie k rovnakému zvlenniu,
- prvý postranný lalok má pokles -41 dB voči hlavnému laloku = o 10db lepšie ako Hannovo okno,
- bočné laloky klesajú približne -6 dB na oktávu,
- prekryv $O = 0, 5$.

1.3.5 Rodina Blackmanových okien

Všeobecne sa rodina Blackman-Harrisových okien¹ získa pridaním viacerých pozdvihnutých sinc funkcií. Ako uvádza SMITH III. [9], definícia znie:

$$w_B(n) = w_{\text{REC}}(n) \sum_{l=0}^{L-1} \alpha_l \cos(l\Omega_N n), \quad (1.47)$$

kde $\Omega_N = \frac{2\pi}{N}$ a $w(n)$ je dĺžka N obdĺžnikového okna s nulovou fázou (pre nenulovú $n \in [-\frac{N-1}{2}, \frac{N-1}{2}]$).

Voľbou $L = 3$ sa získa rodina Blackmanových okien:

$$w_B(n) = w_{\text{REC}}(n) [\alpha_0 + \alpha_1 \cos(\Omega_N n) + \alpha_2 \cos(2\Omega_N n)] \quad (1.48)$$

¹rodina Blackman-Harrisových okien – vychádza z priamočiareho zovšeobecnenia rodiny Hammingových okien

Vzhľadom k všeobecnej Hammingovej rodiny sa pridala ešte jedna cosínus funkcia vážená s α_2 . Hammingova rodina sa zakladá na normalizácii amplitúdy okna a maximalizácii spádu (Hannovo okno) alebo vypúšťaním postranných lalokov (Hammingovo okno). Blackmanova rodina okien optimalizuje buď rovnaké ciele, alebo je možné ich skombinovať, čo vedie k trom subtypom okien v tejto rodine. V nasledujúcej kapitole bude popísané klasické Blackmanovo okno.

1.3.6 Blackmanovo okno

Blackmanovo okno (Obr. 1.7) je špecifický prípad pre ktorý sa volí $\alpha_0 = 0,42$, $\alpha_1 = 0,5$ a $\alpha_2 = 0,08$:

$$w_{\text{BLA}}(n) = w_{\text{REC}}(n) [0,42 + 0,5 \cos(\Omega_N n) + 0,08 \cos(2\Omega_N n)]. \quad (1.49)$$

Okno je charakteristické širším hlavným lalokom a lepším potlačením nežiadúcich kmitočtov, uvádza vo svojej práci SMITH III. [9]. Následkom je však menej ostré spektrum. V krátkodobej analýze sa používa pri sčítaní presahov s $\frac{2}{3}$ prekryvom.

Vlastnosti Blackmanovho okna

- postranné laloky majú pokles približne -18 dB na oktávu (ako Hannovo okno),
- úroveň postranného laloku je približne -58 dB (najhorší prípad),
- tri typy Blackmanovho okna,
- prvý typ okna sa používa pre zvýšenie spádu z -6 dB na oktávu (pravouhlé okno) až na -18 dB na oktávu nastavením amplitúdy a sklonu na nulu pri koncových bodoch,
- druhý typ sa používa na minimalizáciu postranných lalokov (ako Hammingovo okno)
- tretí typ sa používa pre škálovanie stupnice okna,
- prekryv $O = \frac{2}{3}$.

1.3.7 Bartlettovo okno

Bartlettovo okno (Obr. 1.7) sa rozlišuje od ostatných okien hlavne svojim ostrým tvarom, ktorý pripomína trojuholník. Pri analýze sa používa s polovičným prekryvom. Okno je definované v práci Julius O. Smith III. [9] nasledovne:

$$w_{\text{BAR}}(n) = w(n) \left[1 - \frac{|n|}{\frac{N-1}{2}} \right] \quad (1.50)$$

po transformovaní

$$\mathbf{W}(\omega) = \left(\frac{N-1}{2}\right) \text{asinc}^2_{\frac{N-1}{2}}(\omega) \quad (1.51)$$

Vlastnosti Bartlettovho okna

- konvolúcia dvoch obdĺžnikových okien s polovičnou dĺžkou,
- hlavný lalok dvakrát tak široký ako pri rovnakej dĺžke obdĺžnikového okna N (nie polovičnej dĺžky, ktorá generuje Bartlettovo okno cez seba-konvolúciu),
- prvý postranný lalok má dvojnásobný spád v porovnaní s obdĺžnikovým oknom (-26 dB),
- často používané implicitne na vzorkovanie korelácie konečnej dĺžky údajov,
- $N-1$ často nahradzované za N alebo $N+1$ pre vynechanie koncových nulových bodov,
- prekryv $O = 0, 5$.

1.3.8 Vplyv váhovacích okien

Váhovacie okná vplývajú ako na modulové spektrum signálu tak aj na fázové spektrum. Najlepšie to vidno na váhovaní obdĺžnikovým oknom po transformovaní, pri ktorom vznikajú nežiadúce kmitočty (koeficienty) v okolí hlavného kmitočtu v modulovom spektre signálu. Pri Hammingovom okne k tomuto javu dochádza takisto ale v menšej miere. Najväčšie potiaže badať pri meraní slabších signálov v blízkosti silnejších. Spomínaný jav spôsobujú postranné laloky váhovacieho okna (Obr. 1.8). Preto sa pre vlastnosti váhovacieho okna uvádza veľkosť hlavného laloku, veľkosť postranných lalokov a ich pokles voči hlavnému laloku. Platí, že čím menšie sú postranné laloky, tým menej nežiaducích kmitočtov sa objaví v spektre a čím širší je hlavný lalok, tým viac dochádza k zmenšeniu rozlišovacej schopnosti spektrálnej analýzy.

Dĺžka váhovacieho okna súvisí s požiadavkom na frekvenčnú rozlíšiteľnosť alebo presnosť v spektre pri analýze. Dlhému oknu zodpovedá vysoká frekvenčná rozlíšiteľnosť, ale stráca sa rozlíšiteľnosť časových zmien frekvenčného spektra v dôsledku ich spriemerovania, uvádza SYROVÝ². Pokiaľ sa zvolí mala dĺžka N , získa sa presnejšia informácia o časovom umiestnení signálu, ale horšie sa rozoznávajú kmitočtové špičky. Váhovacie okná obsahujú vlastný širokopásmový šum, ktorý sa vďaka úprave dostane aj do analyzovaného signálu. Šírka hlavného laloku a ekvivalentná šírka šumu (Tab. 1.1) je vztiahnutá k f_{vz} (vzorkovací kmitočet)

²SYROVÝ, Václav. *Hudební akustika*. 2. vydání, Praha : Akademie múzických umění, 2008. 440 s. ISBN 978-80-7331-127-8, s. 133

$$\Delta f = \frac{f_{vz}}{N}. \quad (1.52)$$

Tab. 1.1: Vlastnosti váhovacích okien

Typ váhovacieho okna	Šírka hlavného laloku $k_c \Delta f$	Najvyšší postranný lalok [dB]	Ekvivalentná šírka šumu
Obdĺžnikové	Δf	-13	Δf
Bartlettovo	$2\Delta f$	-26	$1,33\Delta f$
Blackmanovo	$3,5\Delta f$	-58	$1,73\Delta f$
Hannovo	$2\Delta f$	-31	$1,36\Delta f$
Hammingovo	$2\Delta f$	-41	$1,5\Delta f$

1.3.9 Prekryv váhovacích okien

Po výbere váhovacieho okna $w(n)$ a s dĺžkou N je potrebné stanoviť prekryv okna. Teoretický prekryv sa počíta podľa vzťahu

$$\frac{100(N-n)}{N} = 100 \left(1 - \frac{1}{2k_c}\right) \quad [\%], \quad (1.53)$$

kde n je krok segmentácie, $\frac{N-n}{N}$ je vlastný prekryv a k_c reprezentuje šírku hlavného laloku. Dosadením sa vypočítajú teoretické prekryvy (Tab. 1.2).

Tab. 1.2: Teoretické prekryvy váhovacích okien

Typ váhovacieho okna	k_c	Teoretický prekryv [%]
Obdĺžnikové	1	50
Bartlettovo	2	75
Blackmanovo	3,5	85
Hannovo	2	75
Hammingovo	2	75

2 VÝSLEDKY ŠTUDENTSKEJ PRÁCE

2.1 Implementované metódy segmentácie

Kapitola popisuje jednotlivé metódy, ktoré som implementoval v programovacom prostredí MATLAB verzie 7.9.0.529 (R2009b) 32-bit. MATLAB (MATrix LABoratory) je určený pre mnohé vedeckotechnické numerické výpočty, teda aj pre meranie a spracovanie signálov. Kľúčovou dátovou štruktúrou pri výpočtoch v MATLABe sú matice. Pre implementovanie algoritmov segmentácie mi ale bohaté postačili vektory.

2.1.1 Algoritmus COLA

Zvolil som tvar pre funkciu algoritmu COLA.

```
function y = cola(x,Fvz,typ_okna,N,Op)
```

Jeho vstupné parametre som popísal v tabulke (Tab. 2.1). Výstupom algoritmu je spracovaný signál (s potlačeným šumom).

Tab. 2.1: Vstupné parametre algoritmu cola

Názov parametru	Popis parametru
x	vstupný signál x (zašumený)
Fvz	vzorkovací kmitočet f_{vz} vstupného signálu x
typ_okna	typ váhovacieho okna $w(n)$ (Kap. Tvorba váhovacích okien)
N	dĺžka váhovacieho okna N
Op	hodnota prekryvu O v pomere (napr. $O = \frac{1}{2}$)

Algoritmus som navrhol podľa nasledujúcich krokov:

1. vytvorenie váhovacieho okna,

```
[w meno_okna] = cw(typ_okna,N);
```

2. určenie počtu segmentov podľa dĺžky vstupného signálu,

```
% L - dlzka vstupneho signalu  
L = length(x);  
% O - dlzka prekryvu  
O = fix(N*Op);  
% Sn - vysledny pocet segmentov  
Sn = ceil((L-O)/(N-O));
```

3. alokovanie pamäte pre výstup aj zásobník,

```
% alokacia vystupu
t = ((Sn-1)*(N-0)) + N;
y = zeros(1,t);
% alokacia zasobnik a
acc = zeros(1,N);
```

4. úprava vstupného signálu pridaním núl (pokiaľ je potrebné) na základe vypočítaných segmentov,

```
% uprava vstupu
x = [x zeros(1,t-L)];
```

5. spracovanie signálu – segmentácia a rekonštrukcia,

```
for i=1:Sn
    % vypocet indexov pre vyber vzorkov zo vstupneho signalu
    indxL = (i-1)*(N-0)+1;
    indxR = indxL + N - 1;
    xx = x(indxL:indxR);
    % vahovanie oknom
    xx = xx.*w';

    % koli rovnakemu pomeru koeficientov
    if (typ_okna==1)
        if (O_p==0.5) xx = xx * 1/2;
        elseif (O_p==0.75) xx = xx * 1/4;
        end
    end

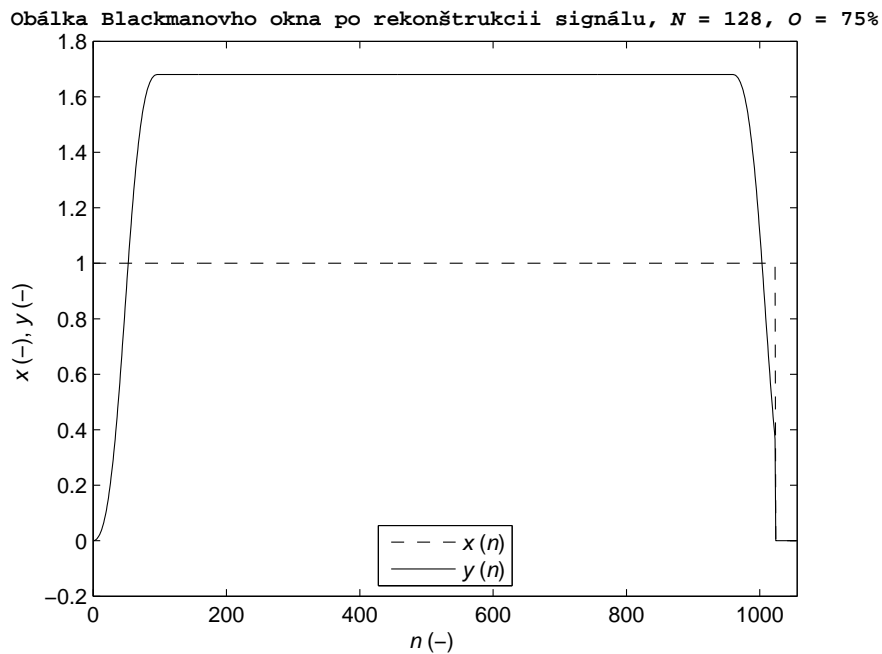
    % "spracovanie" vo frekvencnej oblasti -> yy
    yy = xx;

    % scitanie so zasobnikom
    yy = acc + yy;
    % ulozenie vzorkov do zasobnika
    acc = [yy(N-0+1:N) zeros(1,N-0)];
    % spracovane vzorky na vystup
    y(indxL:indxL+(N-0-1)) = yy(1:N-0);
end;
```

6. úprava rekonštruovaného signálu korekčnou konštantou.

```
switch typ_okna
  case 5 % blackman
    if 0_p == 0.75
      y = y * (1/1.68);
    elseif 0_p == 2/3
      y = y * (1/1.26);
    end
  case 4 % hamming
    if 0_p == 0.75
      y = y * 1/2;
    end
end
```

Pri istých hodnotách prekryvu som po rekonštrukcii segmentovaného vstupného signálu $x(n) = 1$ dostal výslednú obálku váhovacieho okna $y(n)$ s maximom viac ako 1 (Obr. 2.1), preto som do algoritmu vložil úpravu výstupného signálu korekčnou konštantou, ktorej hodnoty som vypočítal ako prevrátenú hodnotu maxima (Tab. 2.2).

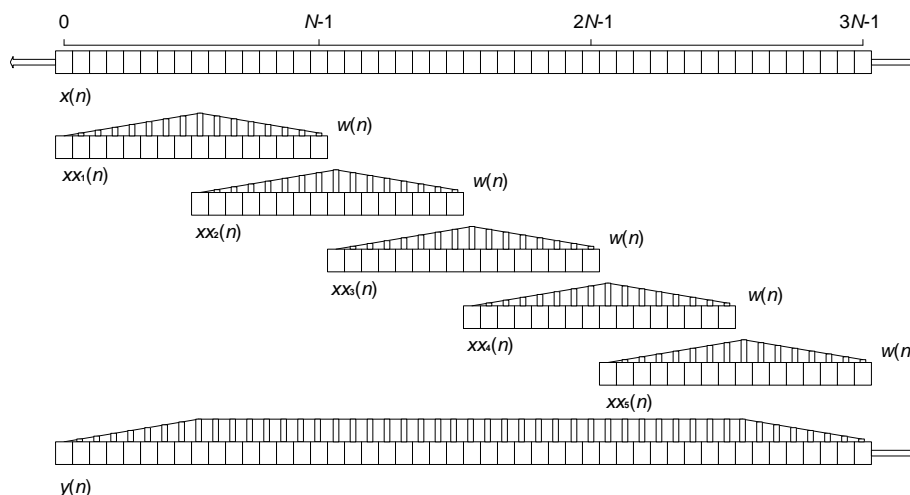


Obr. 2.1: Obálka Blackmanovho okna po rekonštrukcii signálu s prekryvom $O = 75\%$

Tab. 2.2: Korekčné konštanty váhovacích okien pre určité prekryvy

Typ váhovacieho okna	Prekryv	Hodnota korekčnej konštanty
Obdĺžnikové	$\frac{1}{2}$	$\frac{1}{2}$
Obdĺžnikové	$\frac{3}{4}$	$\frac{1}{4}$

Segmentácia vyberá hodnoty vzorkov vstupného signálu $x(n)$ do segmentu $xx_k(n)$ a váhuje ich váhovacím oknom $w(n)$ (Obr. 2.2).



Obr. 2.2: Segmentácia vstupného signálu metódou COLA.

Rekonštrukcia signálu $y(n)$ sa vykonáva sčítaním segmentov $xx_k(n)$ pomocou zásobníka $acc_k(n)$ (Obr. 2.3).

2.1.2 Algoritmus VOLA1

Zvolil som tvar pre funkciu algoritmu VOLA1.

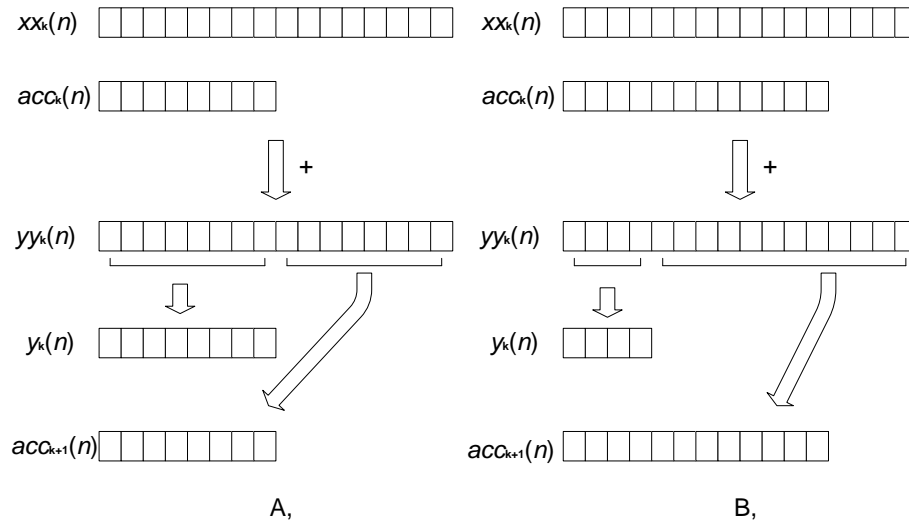
```
function y = vola1(x,Fvz,typ_okna,N)
```

Jeho vstupné parametre som popísal v tabulke (Tab. 2.3). Výstupom algoritmu je spracovaný signál (s potlačeným šumom).

Algoritmus som navrhol podľa nasledujúcich krokov:

1. definovanie pevnej veľkosti prekryvu, minimálnej a maximálnej dĺžky segmentu (kôli zásobníku),

```
% prekryv
O = 64;
```



Obr. 2.3: Rekonštrukcia signálu metódou COLA pre A, 50 % a B, 75 % prekryv.

Tab. 2.3: Vstupné parametre algoritmu volal

Názov parametru	Popis parametru
x	vstupný signál x (zašumený)
Fvz	vzorkovací kmitočet f_{vz} vstupného signálu x
typ_okna	typ váhovacieho okna $w(n)$ (Kap. Tvorba váhovacích okien)
N	postupnosť dĺžok segmentov N

```

% najmensia dlzka segmentu
N_min = 0*2;
% najvacsia dlzka segmentu
N_max = 8192;

2. vytvorenie váhovacieho okna,

[w meno_okna] = cw(typ_okna,N_min);

3. rozdelenie váhovacieho okna na ľavú a pravú časť,

% rozdelenie okna na lavu a pravu cast
wL = w(1:N_min/2)';
wR = w((N_min/2)+1:N_min)';

4. alokovanie pamäte,

% pocet okien (z postupnosti dlzok)
pocet_okien = length(N);
% alokovanie miesta pre vystup

```

```

dlzka_vystupu = sum(N) - pocet_okien*O + 0;
y = zeros(1,dlzka_vystupu);
% zasobnik pre "prekryv"
acc = zeros(1,N_max);
% zasobnik pre spracovany signal
yy = zeros(1,N_max);

```

5. spracovanie signálu – segmentácia, úprava váhovacieho okna, váhovanie, rekonštrukcia,

```

indxT = 1;
for i=1:pocet_okien
    % uprava vahovacieho okna
    wN = [wL ones(1,N(i)-N_min) wR];
    % vypocet indexov pre vyber vzorkov zo vstupneho signalu
    indxL = indxT;
    indxR = indxL + N(i) - 1;
    indxT = indxR - O + 1;
    xx = x(indxL:indxR);
    % vahovanie oknom
    xx = xx.*wN;

    Nd = length(xx);
    if typ_okna==1
        konstanta = [1/2*ones(1,N_min/2) ...
                     ones(1,Nd-N_min) ...
                     1/2*ones(1,N_min/2)];
        xx = xx .* konstanta;
    end

    % desegmentacia
    yy = [xx zeros(1,N_max-N(i))];
    yy = acc + yy;
    acc = [yy(N(i)-O+1:N(i)) zeros(1,N_max-O)];
    y(indxL:indxR-O) = yy(1:N(i)-O);
end

```

2.1.3 Algoritmus VOLA2

Zvolil som tvar pre funkciu algoritmu VOLA2.

```
function y = vola2(x,Fvz,typ_okna,0)
```

Jeho vstupné parametre som popísal v tabulke (Tab. 2.4). Výstupom algoritmu je spracovaný signál (s potlačeným šumom).

Tab. 2.4: Vstupné parametre algoritmu vola2

Názov parametru	Popis parametru
x	vstupný signál x (zašumený)
Fvz	vzorkovací kmitočet f_{vz} vstupného signálu x
typ_okna	typ váhovacieho okna $w(n)$ (Kap. Tvorba váhovacích okien)
O	postupnosť prekryvov segmentov O

Algoritmus som navrhol podľa nasledujúcich krokov:

1. definovanie maximálnej dĺžky segmentu (kôli zásobníku),

```
N_max = 8192;
```

2. vytvorenie okna maximálnej dĺžky segmentu (podvzorkovaním sa získa lavá a pravá časť oblálky modifikovaného váhovacieho okna),

```
[w meno_okna] = cw(typ_okna,N_max);
```

3. alokovanie pamäte,

```
% pocet okien (z postupnosti prekryvov)
pocet_okien = length(O);
% alokovanie miesta pre vystup
dlzka_vystupu = sum(O) + O(1);
y = zeros(1,dlzka_vystupu);
% zasobnik pre "prekryv"
acc = zeros(1,N_max);
% zasobnik pre spracovany signal
yy = zeros(1,N_max);
```

4. spracovanie signálu – segmentácia, úprava váhovacieho okna, váhovanie, rekonštrukcia.

```
O_t = O(1);
indxT = 1;
```

```
for i=1:pocet_okien
    % vytvorenie okna (cutw podvzorkovava okno s N=8192)
```

```

wN = [cutw(w,0_t*2,'l'); cutw(w,0(i)*2,'r')]';
% vypocet indexov pre vyber vzorkov zo vstupneho signalu
indxL = indxT;
indxR = indxL + 0_t + 0(i) - 1;
indxT = indxR - 0(i) + 1;
xx = x(indxL:indxR);
% vahovanie oknom
xx = xx.*wN;

if typ_okna==1
    konstanta = 1/2;
    xx = xx .* konstanta;
end

% rekonstrukcia
yy = [xx zeros(1,N_max-length(xx))];
yy = acc + yy;
acc = [yy(0_t+1:0_t+0(i)) zeros(1,N_max-0(i))];
y(indxL:indxR-0(i)) = yy(1:length(xx)-0(i));
0_t = 0(i);
end

```

2.1.4 Tvorba váhovacích okien

Zvolil som tvar funkcie vytvárajúcej váhovacie okno.

```
function [w meno_okna] = cw(typ_okna,N)
```

Vstupné parametre som popísal v tabulke (Tab. 2.5). Výstupom algoritmu je periodické váhovacie okno dĺžky N spolu s jeho názvom.

```

switch typ_okna
case 1
    w = rectwin(N);
    meno_okna = 'Obdlznikove okno';
case 2
    w = bartlett(N+1); w = w(1:N);
    meno_okna = 'Bartlettovo okno';
case 5
    w = blackman(N,'periodic');

```

Tab. 2.5: Vstupné parametre algoritmu cw

Názov parametru	Popis parametru
typ_okna	typ váhovacieho okna: 1 – Obdĺžnikové okno 2 – Bartlettovo okno 3 – Hannovo okno 4 – Hammingovo okno 5 – Blackmanovo okno
N	dĺžka váhovacieho okna N

```

        meno_okna = 'Blackmanove okno';
    case 3
        w = hann(N,'periodic');
        meno_okna = 'Hannovo okno';
    case 4
        w = 1-0.84*cos(2*pi*(0:N-1)/N)';
        w = w/2;
        meno_okna = 'Hammingovo okno';
    otherwise
        disp('Neznamy typ vahovacieho okna.')
end

```

Okenné funkcie programového vybavenia MATLAB majú voliteľný druhý parameter (okrem Bartlettovho okna) nastavený implicitne na 'symmetric'. Pre správny chod metód popísaných vyššie som potreboval vytvoriť periodické váhovacie okná (Obr. 1.3), preto som nastavil druhý parameter na 'periodic', ktorý je ekvivalentom pre vytvorenie okna s dĺžkou $N + 1$ a následným orezaním posledného vzorku:

```

w = bartlett(N+1);
w = w(1:N);

```

Metóda VOLA2 vytvára obálku váhovacieho okna z dvoch hodnôt prekryvov, z ľavej a pravej časti dvoch váhovacích okien, ktoré sú vytvorené podvzorkovaním z váhovacieho okna maximálnej dĺžky. Implementoval som preto nasledovnú funkciu, ktorá na základe vstupných parametrov podvzorkuje vstupné váhovacie okno na dĺžku `size` a vracia pravú (`side = 1`), alebo ľavú (`side = r`) časť obálky váhovacieho okna:

```

function wN = cutw(w, size, side)

```

```

incr = length(w) / size;
wT = zeros(1,size);

for i = 1:incr:length(w)
    ii = size+1 - ((length(w)-i+1)/incr);
    wT(ii) = w(i);
end

wT = wT';
wN = zeros(1, size/2);
switch lower(side)
    case 'l'
        wN = wT(1:size/2);
    case 'r'
        wN = wT((size/2)+1:size);
end
end

```

2.1.5 Zašumenie signálu

Zašumenie signálu $x(n)$ som implementoval na základe poznatkov z literatúry [1] podľa nasledujúcich krokov:

1. určenie SNR,

```
snrdb=-35;
```

2. načítanie wav súboru,

```
[x Fvz] = wavread('nahravka.wav');
```

3. výpočet výkonu,

```
sig_c = 10*log10(sum(x.*x)/length(x));
```

4. vygenerovanie bieleho šumu,

```
sumsum = wgn(1,length(x),sig_c+snrdb);
```

5. sčítanie šumu a nahrávky,

```
x = x + sumsum';
```

6. zápis do súboru

```
wavwrite(x,Fvz,['nahravka_SNR_',num2str(snrdb),'.wav']);
```

2.1.6 Odhad prahov

Pre odhad prahov šumu som použil nasledujúci spôsob, podľa [1]:

1. vygenerovanie bieleho šumu dĺžky L s výkonom `sumdb`,

```
sum = wgn(1,L,sumdb);
```

2. použitie jednej z metód segmentácie (COLA, VOLA1, VOLA2),
3. výpočet koeficientov pre každý segment dĺžky N ,

```
koef(i,:) = fft(xx)/sqrt(N);
```

4. spriemerovanie smerodatnej odchyľky koeficientov.

```
STDX=mean(std(koef));
```

2.1.7 Spôsob prahovania

Potlačenie šumu som implementoval do funkcie `odsum`, ktorej vstupnou hodnotou je segment a odhadnutý prahovací koeficient. Funkcia potlačuje šum tvrdým prahovaním. Algoritmus bol navrhnutý na základe poznatkov z literatúry [1].

```
function yy = odsum(xx,STDX)
    N = length(xx);
    XX = fft(xx)/sqrt(N);
    vek = abs(XX)<STDX;
    for j = 1:length(XX)
        if(vek(j))
            XX(j) = XX(j) * 0;
        end
    end
    yy = ifft(XX)*sqrt(N);
end
```

2.2 Testovanie implementovaných metód

Kapitola pojednáva o zvukových predlohách, ktoré som zvolil pre testovanie implementovaných algoritmov, ďalej o vstupných podmienkach a jednotlivých krokoch testovania.

2.2.1 Zvolené zvukové nahrávky

Implementované metódy segmentácie som testoval na troch zvukových ukážkach od rôznych interpretov s dĺžkou trvania približne 6 s. Parametre jednotlivých nahrávok sú prehľadne napísané v tabuľke (Tab.2.6). Nahrávky boli zvolené vzhľadom na kvázi stacionaritu signálu v pomalých pasážach a vzhľadom na nestacionaritu signálu v rýchlych pasážach (rytmika). Pri variabilných metódach sa určením dĺžky segmentu v intervale nestacionárneho signálu stáva tento signál stacionárnym vzhľadom k dĺžke okna.

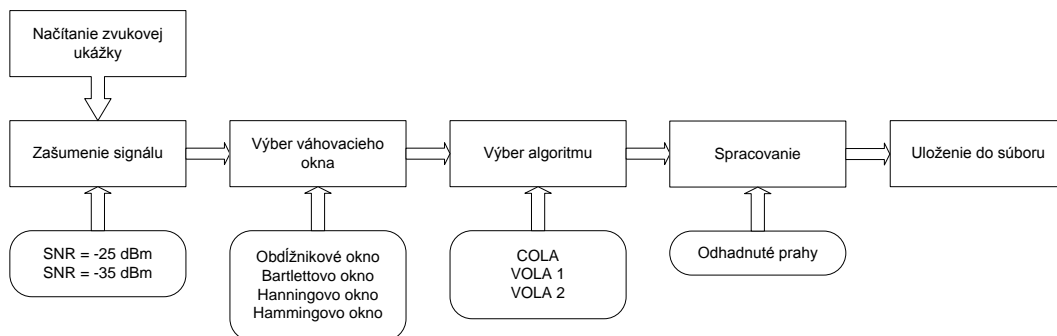
Tab. 2.6: Parametre zvukových nahrávok

Spoločné parametre zvukových nahrávok	
Vzorkovací kmitočet	44100 Hz
Počet kanálov	1 (MONO)
Dĺžka trvania	512000 vzorkov
Parametre nahrávky č.1	
Interpret	Black
Názov skladby	Wonderful Life
Časový interval	00:06 až 00:11
Označenie	black.wav
Parametre nahrávky č.2	
Interpret	blink-182
Názov skladby	Stay Together For The Kids
Časový interval	00:37 až 00:42
Označenie	blink.wav
Parametre nahrávky č.3	
Interpret	Steely Dan
Názov skladby	Sign In A Stranger
Časový interval	00:06 až 00:11
Označenie	dan.wav

2.2.2 Spôsob testovania

Pre testovanie zvukových nahrávok (Obr. 2.4) som zvolil nasledujúci postup:

1. Načítanie jednej zo zvukových ukážok do vektoru, **vstupný parameter**:
 - black.wav,
 - blink.wav,
 - dan.wav.



Obr. 2.4: Bloková schéma testovania implementovaných metód

2. Voľba SNR, **vstupný parameter:**

- SNR = -25 dB,
- SNR = -35 dB.

3. Zašumenie signálu podľa hodnoty SNR

4. Výber jedného z váhovacích okien, **vstupný parameter:**

- Obdĺžnikové okno,
- Bartlettovo okno,
- Hannovo okno,
- Hammingovo okno.

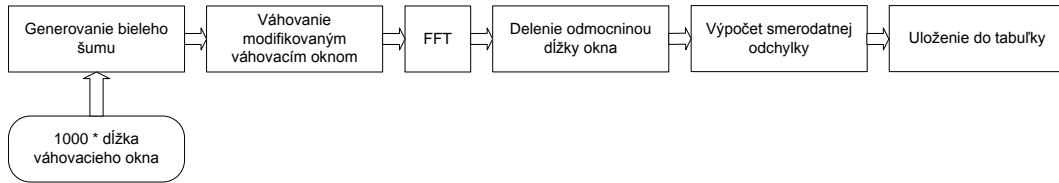
5. Voľba metódy segmentácie, **vstupný parameter:**

- COLA, prekryv 1024 vzorkov, konštantná dĺžka segmentu 2048 vzorkov,
- COLA, prekryv 2048 vzorkov, konštantná dĺžka segmentu 4096 vzorkov,
- VOLA 2, variabilný prekryv, variabilná dĺžka segmentu,
- VOLA 1, prekryv 64 vzorkov, konštantná dĺžka segmentu 2048 vzorkov,
- VOLA 1, prekryv 64 vzorkov, konštantná dĺžka segmentu 4096 vzorkov,
- VOLA 1, prekryv 64 vzorkov, variabilná dĺžka segmentu.

6. Načítanie postupnosti dĺžok a prekryvov segmentov z externého súboru (len pre metódy VOLA 1, VOLA 2). Konkrétne dĺžky som určil empiricky na základe zastúpenia jednotlivých kmitočtov na spektrograme. Dĺžky okien a prekryvov som volil tak, aby približne odzrkadlovali rýchle a pomalé prechody.

7. Načítanie odhadnutých prahov, ktoré som vypočítal nasledovne (Obr. 2.5):

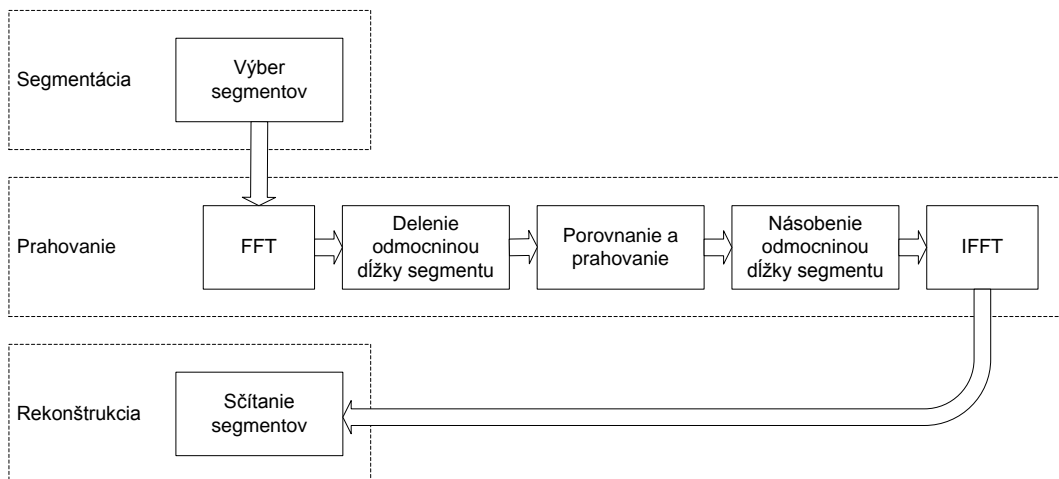
- (a) generovanie bieleho šumu s 1000-násobnou dĺžkou váhovacieho okna (modifikovaného pre metódy s variabilnými dĺžkami segmentov),
- (b) váhovanie zvoleným váhovacím oknom,
- (c) transformácia do kmitočtovej oblasti (FFT, alebo DFT pri určitých dĺžkach váhovacích okien metódy VOLA 2),
- (d) delenie koeficientov odmocninou dĺžky váhovacieho okna,



Obr. 2.5: Blokovaná schéma odhadu prahov

- (e) výpočet smerodatnej odchylky,
- (f) spriemerovanie vypočítaných smerodatných odchýliek,
- (g) uloženie do tabuľky (externého súboru).

8. Spracovanie signálu na základe zvolených vstupných parametrov (Obr. 2.6) :



Obr. 2.6: Blokovaná schéma spracovania segmentov

- (a) výber jednotlivých váhovaných segmentov,
 - (b) transformácia vzorkov do kmitočtovej oblasti,
 - (c) delenie koeficientov odmocninou dĺžky aktuálneho segmentu,
 - (d) porovnanie koeficientov s odhadnutými prahmi, ak je modul koeficientu menší ako prah, aplikuje sa tvrdé prahovanie – vynulovanie
 - (e) násobenie koeficientov odmocninou dĺžky aktuálneho segmentu,
 - (f) spätná transformácia koeficientov do časovej oblasti,
 - (g) rekonštrukcia signálu – sčítanie jednotlivých segmentov.
9. Výstupom je 144 testovacích súborov (3 nahrávky, 2 hodnoty SNR, 4 váhovacie okná, 6 spôsobov segmentácie)

2.2.3 Výpočtová náročnosť

K jednoduchému výpočtu náročnosti implementovaných metód segmentácie som vygeneroval MONO signál s trvaním 1 minuty ($f_{vz} = 44100$ Hz) s pseudonáhodnými hodnotami, na ktorý som aplikoval použité metódy. K variabilným metódam segmentácie boli postupnosti dĺžok a prekryvov segmentov generované taktiež náhodne, s dodržaním pravidiel. Výsledné časy (Tab. 2.7) sú spriemerované z desiatich meraní, pričom konfigurácia počítača, na ktorom prebiehal test bola nasledovná:

- Procesor: AMD Turion(mt) 64 Mobile, Technology MT-32, 1,80 GHz,
- Pamäť: 1,50 GB RAM,
- Operačný systém: Microsoft Windows XP Professional SP3,
- MATLAB ver. 7.9.0.529 (R2009b) 32-bit

Tab. 2.7: Výpočtová náročnosť implementovaných metód

Metóda	Trvanie [s]
COLA, prekryv 1024 vzorkov, dĺžka segmentu 2048 vzorkov	6,034526
COLA, prekryv 2048 vzorkov, dĺžka segmentu 4096 vzorkov	5,899029
VOLA 2, variabilný prekryv, variabilná dĺžka segmentu	7,850573
VOLA 1, prekryv 64 vzorkov, dĺžka segmentu 2048 vzorkov	3,167125
VOLA 1, prekryv 64 vzorkov, dĺžka segmentu 4096 vzorkov	3,234752
VOLA 1, prekryv 64 vzorkov, variabilná dĺžka segmentu	3,543947

Ako vidno, metóda VOLA 2 potrebuje na výpočet viac času v porovnaní s metódou COLA, čo je spôsobené hlavne tým, že pri zmene veľkosti dĺžky segmentu sa generujú dĺžky, ktoré nie sú násobkami dvojkovej sústavy a tým pádom sa namiesto FFT počíta DFT. Metóda VOLA 1 v porovnaní s metódou COLA potrebuje približne polovičný čas, pretože má značne menšiu hodnotu prekryvu pre jednotlivé segmenty, takže sa dostane skôr na koniec signálu.

2.3 Popis priložených súborov

Na priložené CD som do zložiek prehľadne uložil súbory, s ktorými som pracoval:

- zdrojové kódy implementovaných metód,
- tabulky odhadnutých prahov,
- postupnosti dĺžok a prekryvov segmentov,
- pasáže vybraných zvukových nahrávok,
- zašumené zvukové nahrávky,
- spracované zvukové nahrávky s potlačeným šumom.

V koreňovom adresári sa nachádzajú dve hlavné zložky:

- **SPRACOVANE** – obsahuje spracované zvukové nahrávky s potlačeným šumom so zvoleným značením, ktoré je presnejšie popísané v nasledujúcej kapitole
- **MATLAB** – obsahuje ostatné, vyššie spomínané súbory potrebné pre spracovanie zvolenými metódami

2.3.1 Spôsob značenia spracovaných nahrávok

Pre označenie spracovaných zvukových nahrávok v zložke **SPRACOVANE** som zvolil formát tvaru

`nahravka_SNR_xx_okno_metoda_info.wav`

kde význam jednotlivých častí znamená

- **nahravka** – označenie zvukovej nahrávky,
- **xx** – hodnota SNR v dB,
- **okno** – typ váhového okna:
 1. **RECT** – Obdĺžnikové okno,
 2. **BART** – Bartlettovo okno,
 3. **HANN** – Hannovo okno,
 4. **HAMM** – Hammingovo okno,
- **metoda** – zvolená metóda segmentácie:
 1. **COLA** – metóda COLA,
 2. **VOLA1** – metóda VOLA 1,
 3. **VOLA2** – metóda VOLA 2,
- **info** – informácia o veľkosti segmentov:
 1. **var** – variabilná veľkosť segmentov (načítaná z vektoru),
 2. **yyyy** – konštantná veľkosť segmentov (yyyy značí číselnú hodnotu, napr. 2048).

Pre nahrávku pomenovanú ako `black` s odstupom signálu od šumu -25 dB spracovanou Hammingovým oknom metódou COLA s dĺžkou segmentu 4096 vzorkov bude formát označenia

`black_SNR_-25_HAMM_COLA_4096.wav`

a pre spracovanie metódou VOLA 1 s variabilnou dĺžkou segmentu

`black_SNR_-25_HAMM_VOLA1_var.wav`

2.3.2 Zoznam priložených súborov

Priložené súbory v zložke **MATLAB** majú nasledujúci význam:

- **tab** – zložka s tabuľkami odhadnutých prahov (hodnoty uložené v súboroch sa načítavajú pred spracovaním)
- **vec** – zložka s postupnosťami dĺžok a prekryvov segmentov (hodnoty uložené v súboroch sa načítavajú pred spracovaním) s označením:

nahravka_Z.txt

kde Z odlišuje či sa jedná o dĺžky alebo prekryvy segmentov:

1. pre dĺžky segmentov $Z = N$ (napr. **black_N.txt**)
 2. pre prekryvy segmentov $Z = O$ (napr. **black_O.txt**)
- **wav** – zložka s originálnymi a zašumenými pasážami zvolených zvukových nahrávok s označením:

nahravka_SNR_xx.txt

kde xx označuje hodnotu SNR v dB:

1. pre SNR -25 dB $xx = -25$ (napr. **black_SNR_-25.wav**)
2. pre SNR -35 dB $xx = -35$ (napr. **black_SNR_-35.wav**)
3. pre nezašumenú originálnu nahrávku sa xx nedefinuje (napr. **black.wav**)

a nahravka označuje jednu zo zvolených nahrávok:

1. nahravka = **black.wav** pre Black – Wonderful Life (6. s až 11. s z originálnej nahrávky)
 2. nahravka = **blink.wav** pre blink-182 – Stay Together For The Kids (37. s až 42. s z originálnej nahrávky)
 3. nahravka = **dan.wav** pre Steely Dan – Sign In A Stranger (6. s až 11. s z originálnej nahrávky)
- **aktiv.m** – funkcia spustí všetky metódy segmentácie pre zvolenú nahrávku pri zvolenom zašumení a uloží ich ako wav súbory
 - **cola.m** – funkcia pre metódu COLA, bez prahovania (popis v kap. Algoritmus COLA)
 - **cutw.m** – funkcia pre podvzorkovanie váhovacieho okna
 - **cw.m** – funkcia pre vytvorenie váhovacieho okna (popis v kap. Tvorba váhovacích okien)
 - **odsum.m** – funkcia potlačenie šumu (popis v kap. Spôsob prahovania)
 - **sig.m** – funkcia pre výpočet energie
 - **spustenie_prahovanie_cola.m** – funkcia potlačí šum pre zvolenú nahrávku metódou COLA
 - **spustenie_prahovanie_vola1.m** – funkcia potlačí šum pre zvolenú nahrávku metódou VOLA 1

- **spustenie_prahovanie_vola2.m** – funkcia potlačí šum pre zvolenú nahrávku metódou VOLA 2
- **spustenie_tabulka_prahov.m** – funkcia vypočíta odhadnuté prahy (popis v kap. Odhad prahov)
- **testovaci_signal.m** – funkcia vygeneruje testovací signál so zadanou 1. a 2. harmonickou
- **txt2N.m** – funkcia prevedie hodnoty z externého súboru do vektoru
- **vola1.m** – funkcia pre metódu VOLA 1, bez prahovania (popis v kap. Algoritmus VOLA 1)
- **vola2.m** – funkcia pre metódu VOLA 2, bez prahovania (popis v kap. Algoritmus VOLA 2)
- **zasumenie.m** – funkcia pre zašumenie nahrávky so zvoleným SNR (popis v kap. Zašumenie signálu)

3 TESTOVANIE POSLUCHOVEJ KVALITY

Kapitola popisuje metódu, ktorou som subjektívne meral kvalitu zvolených zvukových nahrávok.

3.1 Subjektívne metódy merania kvality

Ako uvádza [7], zrozumiteľnosť je jednou z vlastností, ktoré ukazujú na úspešnosť algoritmov pre spracovanie audio signálov. Ďalším kritériom hodnotenia je kvalita spracovaného audio signálu. Subjektívne metódy merania kvality reči môžeme rozdeliť do dvoch skupín:

- metóda relatívnej preferencie,
- metóda číselného hodnotenia kvality.

K testovaniu implementovaných algoritmov som zvolil druhú skupinu, metódu absolútneho hodnotenia.

3.2 Metóda absolútneho hodnotenia

Zdroj [7] ďalej uvádza, že pri metóde MOS (Mean Opinion Scores) poslucháči hodnotia kvalitu spracovaného audio signálu podľa priebežnej stupnice v rozsahu [0 – 5] (Tab. 3.1). Jednotlivé rozsahy som označil písmenami pre lepšiu prehľadnosť v tabulke výsledkov.

Tab. 3.1: Hodnotiaca stupnica pre metódu MOS podľa normy ITU

	rozsah na škále	kvalita spracovania (potlačenie šumu)	úroveň skreslenia
A	[4 – 5]	vynikajúca (excellent)	nerozpoznateľné
B	[3 – 4]	dobrá (good)	rozpoznateľné, ale nie nepríjemné
C	[2 – 3]	prijateľná (fair)	rozpoznateľné a trochu nepríjemné
D	[1 – 2]	zlá (poor)	nepríjemné, ale nie protivné
E	[0 – 1]	neprijateľná (unsatisfactory)	velmi nepríjemné a protivné

Výhodou metódy MOS je, že poslucháči môžu priradiť svoje hodnotenie absolútne, bez porovnávania. Nevýhodou zároveň je skutočnosť, že každý poslucháč môže mať iný názor na pojem dobrá kvalita. Meraná kvalita je určená ako priemer výsledkov, získaných od všetkých poslucháčov. Aby bol rozdiel medzi názormi poslucháčov na pojem dobrej kvality zjednotený, tak sa testy metódy MOS uskutočňujú v dvoch

fázach: tréovanie a hodnotenie. V tréovacej fázi poslúchajú poslucháči skupinu referenčných signálov s výbornou, strednou a zlou kvalitou. Metóda merania MOS sa používa hlavne pre hodnotenie kodérov reči. Ďalšia podobná metóda je metóda DAM (Diagnostic Acceptability Measure). Využíva sa pre hodnotenie rečových signálov so strednou a vysokou kvalitou. Tu je rečový signál vyhodnocovaný na 16 samostatných 100 bodových škálach pre tri kategórie: kvalita signálu, kvalita pozadia a celková kvalita. Poslucháči sa tak môžu viac sústrediť na dielcie vlastnosti a nemusia pri rozhodovaní robiť kompromisy. V porovnaní s metódou MOS je metóda DAS viac časovo náročná a vyžaduje tréovaných poslucháčov. Pred každým posluhom sú poslucháči požiadaný o vyhodnotenie dvoch hraničných referenčných signálov a štyroch degradovaných referenčných signálov. Prvé dva referenčné signály sú nízkej a vysokej kvality a majú pomôcť poslucháčom nastaviť svoje medze kvality, v ktorých sa môžu pohybovať. Ďalšie štyri degradované referenčné signály ukazujú poslucháčom, na aké typy rušenia sa majú sústrediť. Vyhodnotenie referenčnej fáze potom slúži ku korekcii výsledkov subjektívneho merania kvality jednotlivými poslucháčmi.

3.3 Zhodnotenie výsledkov

Moje subjektívne hodnotenie posluchovej kvality som zapísal do dvoch tabuliek. Tabuľka obsahuje výsledky pre tri zvolené zvukové nahrávky s hodnotami $\text{SNR} = -25 \text{ dB}$ (Obr. 3.1) a $\text{SNR} = -35 \text{ dB}$ (Obr. 3.2). Každá nahrávka spracovaná vybraným váhovacím oknom a metódou segmentácie má štyri hodnoty podľa značenia:

1. **SNRE** – rozdiel odstupú signálu od šumu výstupnej spracovanej nahrávky a vstupnej zašumenej nahrávky (objektívna hodnota),
2. **POTL** – hodnotí subjektívnu kvalitu potlačenia šumu,
3. **SKRES** – hodnotí subjektívnu mieru skreslenia,
4. **DYCH** – značí výskyt dýchania hudobného šumu.

Subjektívne hodnoty majú význam podľa stupnice metódy MOS.

Zvolené metódy mali veľmi dobré výsledky pre potlačenie šumu, čo vysvetľuje fakt, že pri zašumovaní zvukových nahrávok bol použitý biely šum, ktorého prahovacie koeficienty sa ľahko odhadujú. Objektívne hodnotenie vykazuje aj záporné hodnoty. Výpočet totiž nerozoznáva biely šum zašumenej nahrávky od hudobného šumu a skreslenia, ktoré spôsobujú váhovacie okná a tvrdé prahovanie koeficientov po spracovaní nahrávky.

Hlavný rozdiel medzi používanou metódou COLA a implementovanými metódami VOLA je ten, že pri metódach VOLA sa podstatne minimalizuje výskyt dýchania

hudobného šumu (výskyt záleží od vhodne zvolených postupností dĺžok segmentov a prekryvov). Dôkazom toho je nahrávka `black.wav` v častiach s rýchlejšou rytmikou. Metóda VOLA 2 vykazuje lepšie výsledky ako VOLA 1, avšak je výpočtovo náročnejšia.

V pomalých častiach tejto nahrávky a nahrávky `blink.wav` sa pri všetkých metódach prejavovalo počuteľné pískanie. Lepšie potlačenie by vyriešila voľba dlhšieho segmentu v týchto častiach (použitá bola maximálna dĺžka 8192 vzorkov). Pri hodnotení použitých váhovacích okien som dospel k názoru, že Obdĺžnikové okno je nevhodné, pretože v ňom zreteľne počut praskanie medzi susednými segmentami. Hammingovo a Bartletovo okno dávali dobré výsledky, voľba Hannovho okna dávala najlepšie výsledky.

Celkovo bolo hodnotenie spracovaných zvukových nahrávok testované iba mnou, teda jedným poslucháčom. Hodnotil som za použitia slúchadiel KOSS Porta Pro a integrovanej zvukovej karty Realtek AC97 Audio.

Váhovacie okno		Metóda		SNR = -25 dB											
				black.wav				blink.wav				dan.wav			
				SNRE	POTL	SKRES	DYCH	SNRE	POTL	SKRES	DYCH	SNRE	POTL	SKRES	DYCH
Obdĺžnikové	COLA_2048	1,932	B	D	áno	5,674	B	D	áno	1,279	B	D	áno		
	COLA_4096	4,762	B	C	áno	7,952	B	C	áno	2,944	B	D	áno		
	VOLA2_var	4,335	B	C	nie	4,713	B	C	nie	1,092	B	D	nie		
	VOLA1_2048	-2,189	B	D	áno	0,533	B	D	áno	0,284	B	D	áno		
	VOLA1_4096	-1,837	B	D	áno	1,298	B	D	áno	0,606	B	D	áno		
	VOLA1_var	-2,623	B	C	nie	0,425	B	D	nie	0,286	B	D	nie		
Bartlettovo	COLA_2048	1,744	B	D	áno	5,196	B	D	áno	1,396	B	C	áno		
	COLA_4096	4,629	B	B	áno	7,958	B	C	áno	2,988	B	B	áno		
	VOLA2_var	4,322	B	B	nie	4,771	B	B	nie	1,056	B	B	nie		
	VOLA1_2048	-2,438	B	D	áno	0,860	B	D	áno	0,191	B	D	áno		
	VOLA1_4096	-1,926	B	D	áno	1,477	B	D	áno	0,586	B	D	áno		
	VOLA1_var	-2,690	B	C	nie	0,439	B	C	nie	0,229	B	C	nie		
Hannovo	COLA_2048	1,810	B	D	áno	5,277	B	D	áno	1,442	B	C	áno		
	COLA_4096	4,655	B	B	áno	8,061	B	B	áno	3,021	B	B	áno		
	VOLA2_var	4,344	B	B	nie	4,720	B	B	nie	1,167	B	B	nie		
	VOLA1_2048	-2,429	B	D	áno	0,902	B	D	áno	0,138	B	D	áno		
	VOLA1_4096	-1,919	B	D	áno	1,547	B	D	áno	0,604	B	D	áno		
	VOLA1_var	-2,679	B	C	nie	0,422	B	C	nie	0,221	B	C	nie		
Hammingovo	COLA_2048	1,754	B	D	áno	5,170	B	D	áno	1,407	B	C	áno		
	COLA_4096	4,626	B	B	áno	7,952	B	B	áno	3,001	B	B	áno		
	VOLA2_var	4,328	B	B	nie	4,647	B	B	nie	1,079	B	B	nie		
	VOLA1_2048	-2,423	B	D	áno	0,825	B	D	áno	0,181	B	D	áno		
	VOLA1_4096	-1,920	B	D	áno	1,520	B	D	áno	0,610	B	D	áno		
	VOLA1_var	-2,682	B	C	nie	0,524	B	C	nie	0,254	B	C	nie		

Obr. 3.1: Výsledky hodnotenia zvukových nahrávok pre SNR = -25 dB

Váhovacie okno		Metóda		SNR = -35 dB																					
				black.wav						blink.wav						dan.wav									
				SNRE	POTL	SKRES	DYCH	SNRE	POTL	SKRES	DYCH	SNRE	POTL	SKRES	DYCH	SNRE	POTL	SKRES	DYCH						
Obdĺžnikové	COLA_2048	10,632	A	C	áno	13,945	A	B	áno	8,052	A	C	áno	14,144	A	C	áno	17,373	A	B	áno	10,807	A	B	áno
	COLA_4096	13,672	A	B	nie	13,201	A	B	nie	7,572	A	B	nie	13,672	A	C	áno	4,994	A	B	áno	1,800	A	C	áno
	VOLA2_var	0,902	A	C	áno	5,663	A	B	nie	3,774	A	B	nie	2,196	A	C	áno	3,774	A	B	nie	1,831	A	C	áno
	VOLA1_2048	2,196	A	B	áno	13,988	A	B	áno	8,096	A	B	áno	-0,795	A	B	nie	17,456	A	A	áno	2,075	A	B	nie
	VOLA1_4096	13,992	A	A	áno	17,456	A	A	áno	10,903	A	A	áno	10,469	A	B	áno	13,988	A	B	áno	8,096	A	B	áno
	VOLA1_var	13,659	A	A	nie	13,181	A	A	nie	7,588	A	A	nie	13,992	A	A	áno	17,456	A	A	áno	10,903	A	A	áno
Bartlettovo	VOLA2_var	0,228	A	B	áno	4,880	A	B	áno	1,613	A	B	áno	13,659	A	A	nie	13,181	A	A	nie	7,588	A	A	nie
	VOLA1_2048	2,033	A	B	áno	5,576	A	B	áno	1,772	A	B	áno	0,228	A	B	áno	4,880	A	B	áno	1,613	A	B	áno
	VOLA1_4096	-1,055	A	A	nie	3,743	A	A	nie	1,940	A	A	nie	2,033	A	B	áno	5,576	A	B	áno	1,772	A	B	áno
	VOLA1_var	10,499	A	B	áno	14,012	A	B	áno	8,141	A	B	áno	-1,055	A	A	nie	3,743	A	A	nie	1,940	A	A	nie
	COLA_2048	14,000	A	A	áno	17,475	A	A	áno	10,954	A	A	áno	10,499	A	B	áno	14,012	A	B	áno	8,141	A	B	áno
	COLA_4096	13,663	A	A	nie	13,199	A	A	nie	7,625	A	A	nie	14,000	A	A	áno	17,475	A	A	áno	10,954	A	A	áno
Hannovo	VOLA2_var	0,251	A	B	áno	4,884	A	B	áno	1,634	A	B	áno	13,663	A	A	nie	13,199	A	A	nie	7,625	A	A	nie
	VOLA1_2048	2,048	A	B	áno	5,572	A	B	áno	1,819	A	B	áno	0,251	A	B	áno	4,884	A	B	áno	1,634	A	B	áno
	VOLA1_4096	-1,017	A	A	nie	3,740	A	B	nie	1,957	A	A	nie	2,048	A	B	áno	5,572	A	B	áno	1,819	A	B	áno
	VOLA1_var	10,471	A	B	áno	13,986	A	B	áno	8,094	A	B	áno	-1,017	A	A	nie	3,740	A	B	nie	1,957	A	A	nie
	COLA_2048	13,983	A	A	áno	17,448	A	A	áno	10,914	A	A	áno	10,471	A	B	áno	13,986	A	B	áno	8,094	A	B	áno
	COLA_4096	13,660	A	A	nie	13,199	A	A	nie	7,601	A	A	nie	13,660	A	A	nie	13,199	A	A	nie	7,601	A	A	nie
Hammingovo	VOLA2_var	0,252	A	B	áno	4,892	A	B	áno	1,622	A	B	áno	0,252	A	B	áno	4,892	A	B	áno	1,622	A	B	áno
	VOLA1_2048	2,038	A	B	áno	5,712	A	B	áno	1,791	A	B	áno	2,038	A	B	áno	5,712	A	B	áno	1,791	A	B	áno
	VOLA1_4096	-1,037	A	A	nie	3,614	A	B	nie	1,968	A	A	nie	-1,037	A	A	nie	3,614	A	B	nie	1,968	A	A	nie
	VOLA1_var	10,471	A	B	áno	13,986	A	B	áno	8,094	A	B	áno	10,471	A	B	áno	13,986	A	B	áno	8,094	A	B	áno
	COLA_2048	13,983	A	A	áno	17,448	A	A	áno	10,914	A	A	áno	13,983	A	A	áno	17,448	A	A	áno	10,914	A	A	áno
	COLA_4096	13,660	A	A	nie	13,199	A	A	nie	7,601	A	A	nie	13,660	A	A	nie	13,199	A	A	nie	7,601	A	A	nie

Obr. 3.2: Výsledky hodnotenia zvukových nahrávok pre SNR = -35 dB

4 ZÁVER

V práci som popísal metódy pre segmentáciu zvukového signálu s variabilnou dĺžkou segmentu s konštantným prekryvom a s variabilnou dĺžkou segmentu s variabilným prekryvom pri použití rôznych druhov váhovacích okien a prahovaní koeficientov. Metódy som potom implementoval v programovom prostredí MATLAB a porovnal s bežnými metódami segmentácie. Porovnával som spracované zvukové ukážky a hodnotil ich podľa stupnice MOS.

Najlepšie vychádza použitie metódy VOLA 2 so zvoleným Hannovým oknom. Algoritmus VOLA 2 je však v porovnaní s algoritmom COLA výpočtovo náročnejší približne o tretinu na signále trvajúcim 1 minúta pri vzorkovacej frekvencii 44100 Hz. Pri zmene veľkosti segmentu vznikajú dĺžky, ktoré nemajú hodnoty mocniny čísla 2 a tým pádom sa v programovom prostredí MATLAB namiesto algoritmu FFT počíta Fourierova transformácia algoritmu DFT.

Algoritmus VOLA 1 je najmenej náročný na výpočet, pretože má pevne stanovenú dĺžku prekryvu (v tejto práci na 64 vzorkov) a tým pádom spracuje signál použitím variabilných dĺžok segmentov skôr ako algoritmus COLA, ktorý má pevne stanovenú dĺžku segmentu vždy s polovičným prekryvom. Rýchlosť je v porovnaní s COLA dvojnásobná na signále trvajúcim 1 minúta pri vzorkovacej frekvencii 44100 Hz.

Hlavným rozdielom metód s variabilnou dĺžkou segmentu oproti bežnej metóde COLA je minimalizácia výskytu dýchania hudobného šumu. Minimalizuje sa korektným určením postupnosti dĺžok a prekryvov segmentov, ktoré rozdeľujú nestacionárny hudobný signál na stacionárne intervaly v rámci dĺžky segmentu. V tejto práci som určoval postupnosti empiricky na základe zastúpenia jednotlivých kmitočtov v spektrograme. Pre adaptívne určovanie postupností dĺžok a prekryvov segmentov by sa mohli použiť autokorelačné metódy.

Potlačovanie šumu implementovanými metódami dávalo dobré výsledky pre nahrávky zašumené bielym šumom. Posluchové testy pre iné druhy šumov neboli uskutočnené. Tvrdé prahovanie koeficientov spôsobovalo badateľné skreslenie, v odšumených nahrávkach sa vyskytovalo mierne pískanie a praskanie medzi susediacimi segmentami. Pre výsledky s menším skreslením by sa v práci mohlo pokračovať implementovaním iných spôsobov prahovania a určovania prahov, pridaním filtrov pre hudobný šum a použitím ďalších metód transformácie z časovej do kmitočtovej oblasti.

LITERATURA

- [1] BALÍK, Miroslav, RÁŠO, Ondrej. *Sound Artifacts Caused By The Thresholding Method*. The LASTED International Conference on Modelling, Simulation, and Identification (MSI 2009), Beijing, China, pp. 1-8. ISBN: 978-0-88986-810-6, 2009.
- [2] Natick, Massachusetts: The MathWorks Inc. *MATLAB - Documentation*. [online]. 2010, [cit. 10. 5. 2010]. Dostupné z URL: <<http://www.mathworks.co.uk/access/helpdesk/help/techdoc/>>.
- [3] OPPENHEIM, A.V., SCHAFER, R.W. *Discrete-Time Signal Processing*. Prentice-Hall, 1989. 879 s. ISBN 978-0132162920.
- [4] RÁŠO, Ondrej, BALÍK, Miroslav. *Half-soft parabolic thresholding of DFT coefficients*. Proc. 32nd International Conference Telecommunication and Signal Processing 2009, Dunakiliti, Hungary, pp. 1-6. ISBN: 978-963-06-7716-5, 2009.
- [5] RÁŠO, Ondrej, BALÍK, Miroslav. *Polynomial thresholding based on probability distribution of noise signal*. Department of Telecommunications, The Faculty of Electrical Engineering and Communication, Brno University of Technology, 2009.
- [6] RUDOY, Daniel, BASU, Prabahan, QUATIERI, Thomas F., DUNN, Bob, WOLFE, Patrick J.. *Adaptive Short-Time Analysis-Synthesis for Speech Enhancement*. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 4905–4908, 2008.
- [7] SMÉKAL, Zdeněk. *Číslíkové zpracování řeči*. Ústav telekomunikací FEKT VUT v Brně, 2009.
- [8] SMÉKAL, Zdeněk. *Číslíkové zpracování signálu*. Ústav telekomunikací FEKT VUT v Brně, 2009.
- [9] SMITH III., O. JULIUS. *Spectral Audio Signal Processing*. [online]. 2009, poslední aktualizace 1. 3. 2009 [cit. 10. 5. 2010]. Dostupné z URL: <<https://ccrma.stanford.edu/jos/sasp/index.html>>.
- [10] SYROVÝ, Václav. *Hudební akustika*. 2. vydání, Praha : Akademie múzických umění, 2008. 440 s. ISBN 978-80-7331-127-8.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

FFT rýchlá Fourierova transformácia – fast Fourier transform

IFFT Spätná forma rýchlej Fourierovej transformácie

SNR Odstup signálu od šumu – Signal to Noise Rate

DFT diskrétna Fourierova transformácia – Discrete Fourier Transform

OLA pričítanie presahu – OverLap Add

FIR konečná impulzová charakteristika – finite impulse response

COLA konštantné pričítanie presahu – constant OverLap Add

VOLA variabilné pričítanie presahu – variable OverLap Add

DSP číslicové zpracování signálů – Digital Signal Processing

f_{vz} vzorkovací kmitočet