



TECHNICKÁ UNIVERZITA V LIBERCI
Fakulta mechatroniky, informatiky
a mezioborových studií ■

METODY POSTFILTROVÁNÍ PRO REDUKCI ŠUMU

Bakalářská práce

Studijní program: B2646 – Informační technologie
Studijní obor: 1802R007 – Informační technologie

Autor práce: **Petr Bartoš**
Vedoucí práce: doc. Ing. Zbyněk Koldovský, Ph.D.





TECHNICAL UNIVERSITY OF LIBEREC
Faculty of Mechatronics, Informatics
and Interdisciplinary Studies ■

POSTFILTERING METHODS FOR NOISE REDUCTION

Bachelor thesis

Study programme: B2646 – Information Technology
Study branch: 1802R007 – Information Technology

Author: **Petr Bartoš**
Supervisor: doc. Ing. Zbyněk Koldovský, Ph.D.



ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Petr Bartoš**
Osobní číslo: **M12000105**
Studijní program: **B2646 Informační technologie**
Studijní obor: **Informační technologie**
Název tématu: **Metody postfiltrování pro redukci šumu**
Zadávající katedra: **Ústav informačních technologií a elektroniky**


Z á s a d y p r o v y p r a c o v á n í :

1. Nastudujte a implementujte metody pro jednokanálovou redukci šumu s využitím referenčního signálu.
2. Metody porovnávejte na simulovaných zarušených nahrávkách. K vyhodnocování použijte kritéria PEASS.
3. Výsledky interpretujte podle významu kritérií a porovnejte své závěry s výsledky v knize I. Tasheva.
4. Simulujte situace, kdy referenční signál je zbarven dvoukanálovým filtrem, který referenci získává potlačením cílového signálu (ze stereo záznamu).
5. Opakujte totožné experimenty jako v bodu 2 s takto zbarveným referenčním signálem a proveďte vyhodnocení.
6. Navrhněte způsoby opravy zbarvení referenčního signálu a opět je experimentálně vyhodnoťte.

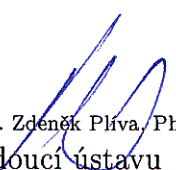
Rozsah grafických prací: **Dle potřeby dokumentace**
Rozsah pracovní zprávy: **cca 30 stran**
Forma zpracování bakalářské práce: **tištěná/elektronická**
Seznam odborné literatury:

- [1] **I. Tashev, "Sound Capture and Processing: Practical Approaches", Wiley, 2008.**
- [2] **Valentin Emiya, Emmanuel Vincent, Niklas Harlander and Volker Hohmann, Subjective and objective quality assessment of audio source separation, IEEE Transactions on Audio, Speech and Language Processing, 2011, 19 (7), pp. 2046-2057.**

Vedoucí bakalářské práce: **doc. Ing. Zbyněk Koldovský, Ph.D.**
Ústav informačních technologií a elektroniky
Konzultant bakalářské práce: **Ing. Jiří Málek, Ph.D.**
Ústav informačních technologií a elektroniky
Datum zadání bakalářské práce: **12. září 2014**
Termín odevzdání bakalářské práce: **15. května 2015**


prof. Ing. Václav Kopecký, CSc.
děkan




prof. Ing. Zdeněk Pliva, Ph.D.
vedoucí ústavu

V Liberci dne 12. září 2014

Prohlášení

Byl jsem seznámen s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu TUL.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím mé bakalářské práce a konzultantem.

Současně čestně prohlašuji, že tištěná verze práce se shoduje s elektronickou verzí, vloženou do IS STAG.

Datum: 14.5.2015

Podpis: 

Poděkování

Chtěl bych poděkovat doc. Ing. Zbyňku Koldovskému PhD. za vedení mé bakalářské práce, cenné rady, odborný dohled a vstřícnost při konzultacích.

Abstrakt

Tato práce se věnuje problému odstranění šumu ze zvukového signálu. Jsou zde zkoumány jednokanálové metody pro redukci šumu, zejména Wienerův filtr. Tyto metody tu porovnávám za použití referenčního signálu a vyhodnocuji podle kritérií PEASS. Dále se zabývám použitím těchto metod na signál zabarvený dvoukanálovým filtrem.

Klíčová slova: Postfiltrování, redukce šumu, PEASS, Wiener

Abstract

This thesis focuses on the problem of noise reduction in audio signals. A number of methods designed for the purpose of noise reduction are tested in this thesis, mainly the Wiener suppression rule. The methods were being evaluated subjectively and objectively by using the PEASS toolkit. Each method was tested on a reference signal and on a signal modified using a dual-channel filter.

Keywords: Postfiltering, noise reduction, PEASS, Wiener

Obsah

Úvod	10
1 Předpoklady	12
1.1 Zvuk	12
1.2 Šum	12
1.3 SNR	13
1.4 Digitální filtr	13
1.5 Diskrétní Fourierova transformace	14
1.6 Použití filtrů k redukci šumu	15
2 Jednokanálové metody	16
2.1 Definice A-priori SNR a A-posteriori SNR	16
2.2 Wienerova metoda redukce šumu	16
2.3 Upravená Wienerova metoda	18
2.4 Ephraim a Malah	19
2.5 JMAP SAE	20
3 Sada nástrojů PEASS	21
4 Implementace	23
4.1 Nahrávání	23
4.2 Zašumění	23
4.2.1 Bílý šum	24
4.2.2 Růžový šum	25
4.2.3 Šum typu babble	26
4.2.4 Zašumění jiným řečníkem	26
4.3 Metody redukce šumu	27
5 Vyhodnocení	28
5.1 Ideální situace	28
5.1.1 Wienerův filtr	28
5.1.2 MMSE	30
5.1.3 Log-MMSE	31
5.1.4 JMAP SAE	31
5.2 Situace s odhadnutým šumem	32
5.2.1 Wienerova metoda	33
5.2.2 MMSE	34
5.2.3 Log-MMSE	36
5.2.4 JMAP SAE	37

5.3	Oprava zabarveného signálu	37
5.4	Porovnání opraveného signálu	41
	Závěr	43
	Použitá literatura	44
	Obsah příloženého CD	46
	Přílohy	47
	Příloha A – Tabulky pro Wienerovu metodu při ideální situaci	47
	Příloha B – Tabulky pro metodu MMSE při ideální situaci	48
	Příloha C – Tabulky pro metodu Log-MMSE při ideální situaci	49
	Příloha D – Tabulky pro metodu JMAP SAE při ideální situaci	50
	Příloha E – Tabulky pro Wienerovu metodu při odhadu šumu	51
	Příloha F – Tabulky pro metodu MMSE při odhadu šumu.....	52
	Příloha G – Tabulky pro metodu Log-MMSE při odhadu šumu	53
	Příloha H – Tabulky pro metodu JMAP SAE při odhadu šumu	54

Seznam obrázků

Obrázek 1: Spektrální hustota bílého šumu	25
Obrázek 2: Spektrální hustota růžového šumu s logaritmickou osou x	26
Obrázek 3: Impulzní odezva	39
Obrázek 4: Spektrální hustota originálního a odhadnutého šumu	40
Obrázek 5: Spektrální hustota upraveného šumu a originálního šumu	41

Seznam tabulek

Tabulka 1: Výsledky PEASS pro Wienerovu metodu a bílý šum	28
Tabulka 2: Výsledky PEASS pro Wienerovu metodu a růžový šum	29
Tabulka 3: Výsledky PEASS pro Wienerovu metodu a šum typu babble.....	29
Tabulka 4: Výsledky PEASS pro metodu MMSE a ženský hlas zašuměný mužským hlasem	30
Tabulka 5: Výsledky PEASS pro metodu Log-MMSE a ženský hlas zašumění mužským hlasem ...	31
Tabulka 6: Výsledky PEASS pro metodu Wienerova filtru pro reálný odhad bílého šumu	33
Tabulka 7: Výsledky PEASS pro metodu Wienerova filtru pro reálný odhad růžového šumu.....	33
Tabulka 8: Výsledky PEASS pro metodu MMSE pro reálný odhad bílého šumu	34
Tabulka 9: Výsledky PEASS pro metodu MMSE pro reálný odhad růžového šumu	35
Tabulka 10: Výsledky PEASS pro metodu MMSE a mužský hlas zašumění ženským hlasem.....	35
Tabulka 11: Výsledky PEASS pro metodu Log-MMSE pro reálný odhad bílého šumu.....	36
Tabulka 12: Výsledky PEASS pro metodu Log-MMSE pro reálný odhad růžového šumu	36

Tabulka 14: Výsledky PEASS pro metodu JMAP SAE pro reálný odhad růžového šumu.....	37
Tabulka 13: Výsledky PEASS pro metodu JMAP SAE pro reálný odhad bílého šumu.....	37
Tabulka 15: Porovnání výsledků PEASS pro zabarvený signál a signál s opraveným spektrem ...	42

Seznam zkratek

MMSE (Minimal Mean Square Error) – minimální střední kvadratická hodnota chyby

JMAP SAE (Joint maximum a-posteriori spectral amplitude and phase estimator) – společný odhad amplitudového a fázového spektra s důrazem na a-posteriori hustotu

SNR (Signal to Noise Ratio) – poměr signálu k šumu

SDR (Signal to Distortion Ratio) – kritérium PEASS, vyjadřující poměr signálu ke zkreslení

ISR (source Image to Spatial distortion Ratio) – kritérium PEASS, vyjadřující poměr původního signálu k prostorovému zkreslení

SIR (Signal to Interference Ratio) – kritérium PEASS, vyjadřující poměr užitečného signálu k rušivému signálu

SAR (Signal to Artifacts Ratio) – kritérium PEASS, vyjadřující poměr signálu vůči artefaktům

OPS (Overall Perceptual Score) – kritérium PEASS, vyjadřující celkové vnímání nahrávky

TPS (Target-related Perceptual Score) – kritérium PEASS, vyjadřující vnímání zkreslení signálu

IPS (Interference-related Perceptual Score) – kritérium PEASS, vyjadřující vnímání rušivé složky signálu

APS (Artifacts-related Perceptual Score) – kritérium PEASS, vyjadřující vnímání vzniklých artefaktů.

Úvod

Problém šumu existuje v oblasti zpracování zvuku od samého počátku. Ať už se jednalo o nekvalitní nahrávací zařízení, elektromagnetické rušení či prostý hluk z okolních jevů, šum byl vždy součástí zvukových nahrávek. Proto bylo jenom otázkou času, než se přijde na způsob omezení či kompletního odstranění šumu. Vylepšením hardwaru se potlačil šum způsobený nekvalitními součástkami či rušením, nicméně šum ve formě nežádoucích okolních zvuků se hardwarem nedá zcela potlačit. Na řadu přišla softwarová řešení, zejména ve formě digitálních filtrů. Tyto filtry pracují nad diskrétními hodnotami a potlačují, případně odstraňují některou část signálu.

Hlavním důvodem výběru tohoto tématu byla moje záliba ve zvuku, zejména v hudbě, a zároveň potřeba komunikace přes Internet prostřednictvím počítače, jež je v dnešní době, zejména v oblasti IT, nevyhnutelná. Způsob, jakým se šum odstraňuje, mě vždy zajímal, takže téma této práce mě ihned oslovilo.

Cílem práce je seznámit se s problematikou jednokanálové redukce šumu. V současné době se používají metody, buď zastaralé a postavené na ideálním případě, kde známe jak originální signál, tak rušivý signál, příkladem může být Weinerova metoda [1] nebo metody náročné na výpočet jako například metody Short-term MMSE [2] a Log-MMSE od Ephraima a Malaha [3]. Objevují se zde však i nové metody řešení redukce šumu ve formě Joint maximum a-posteriori spectral amplitude and phase estimator (JMAP SAE) [4], která byla navržena s ohledem na rychlost výpočtu v reálném čase.

S jednotlivými metodami se v práci seznámím a následně je budu implementovat pomocí vývojového prostředí MATLAB. Dále pomocí tohoto prostředí nahraji signál, v mém případě řeč, a tento signál zaruším. Na zašuměný signál budu postupně aplikovat jednotlivé metody pro redukci šumu a výsledný odrušený signál porovnam a vyhodnotím míru odrušení.

Vyhodnocení bude probíhat dvěma způsoby. Prvním způsobem je vyhodnocení pomocí nástrojové sady PEASS [5]. Tato nástrojová sada vyhodnotí opravenou nahrávku oproti původní a vypíše zlepšení podle daných kritérií, jako jsou např. SIR, SDR a další. Dále budu také hodnotit, jak selepší nahrávka při poslechu.

Zásadním problémem při redukci šumu u jednokanálových nahrávek je problém neznalosti původního rušivého signálu. Tento problém se pokusím vyřešit aplikováním dvoukanálového filtru na nahrávku. Takto nasimuluji reálnou situaci, kdy získávám signál pomocí 2 mikrofonů, u kterých znám jejich vzdálenost, a tudíž i časové zpoždění. To je potřebné pro správné odečtení signálů, v mém případě řeči, získaných na těchto mikrofonech, kde zbylým signálem bude odhadnutý šum.

Problém u takto získaného šumu je však zabarvení jeho spektra. Pokusím se tyto nahrávky vyhodnotit a zjistit vliv zabarvení na jednotlivé metody.

Poté se pokusím opravit toto zabarvení ve formě opravy amplitudového spektra odhadnutého signálu a následně zjistím, jak moc se změnilo vyhodnocení.

1 Předpoklady

1.1 Zvuk

Zvuk je mechanické vlnění, které je v látkovém prostředí schopno vyvolat sluchový vjem přenosem na sluchové orgány. Toto vlnění může nabývat různých frekvencí, z nichž je pouze část slyšitelná lidským sluchovým aparátem. Slyšitelný frekvenční rozsah se liší člověk od člověka, nicméně je známo, že se nachází v intervalu od 16 Hz do 20 kHz a s rostoucím věkem horní hranice klesá [6].

Oblast informačních technologií se zabývá mimo jiné zpracováním digitálního signálu, tedy signálu zaznamenaného pomocí nahrávacího zařízení, většinou mikrofону. Mikrofon převede akustický tlak vyvolaný vlněním na elektrický signál, který je posléze pomocí digitálně analogového převodníku převeden na digitální data.

Takto získaná data, lze dále zpracovávat pomocí různých softwarových pomůcek k tomu určených. V mém případě jsem použil interaktivní programové prostředí MATLAB od společnosti MathWorks. Tento program umožňuje maticové počítání, implementaci algoritmů, vykreslování 2D a 3D grafů apod.

1.2 Šum

Ideální signál se skládá pouze z užitečných informací, podle svého účelu. Reálně se však každý signál skládá jak z užitečné složky, která je potřebná ke zpracování, tak ze složek nepotřebných, většinou rušivých. Těmto složkám se říká šum a snahou je podstatně zmenšit, popřípadě úplně odstranit tyto složky. Pozorovaný signál lze popsat pomocí následující rovnice, převzaté z [1].

$$y(t) = x(t) + d(t)$$

(Vzorec 1)

Zde proměnná x představuje cílový signál, proměnná d představuje šum a proměnná y představuje jejich kombinaci, tedy zašuměný signál, všechny proměnné jsou vyjádřeny v závislosti na čase t .

Ačkoli by se mohlo zdát při pohledu na vzorec 1, že odstranění šumu je jednoduchou záležitostí, opak je pravdou. Hlavním důvodem je především neznalost původního rušivého signálu d , šumu. Jelikož šum není znám, nelze jej ani odstranit ze získaného signálu.

1.3 SNR

Signal-to-noise ratio neboli poměr užitečného signálu k šumu je jeden z hlavních atributů určujících kvalitu signálu. Tato hodnota je vyjádřena pomocí jednotky decibel a vypočítá se podle následující rovnice.

$$\text{SNR} = 10 \log_{10} \left(\frac{\sum_t x(t)^2}{\sum_t d(t)^2} \right)$$

(Vzorec 2)

1.4 Digitální filtr

Digitální filtr je systém, který pracuje s daty z diskretního signálu. Hlavním účelem filtrů je kompletní či částečné potlačení některých složek signálu.

Každý digitální filtr je popsán impulzní odezvou, odezvou na jednotkový impuls na vstupu, v následující podobě.

$$h[n] = \sum_{i=0}^N b_i \cdot \delta[n-i] = \begin{cases} b_n & 0 \leq n \leq N \\ 0 & \text{pro ostatní} \end{cases}$$

(Vzorec 3)

Filtry mohou být také popsány pomocí přenosové funkce.

$$H(z) = \sum_{n=-\infty}^{\infty} h[n] \cdot z^{-n}$$

(Vzorec 4)

$$H(z) = \frac{Y(z)}{X(z)}$$

(Vzorec 5)

Zde je filtr vyjádřen jako poměr výstupů systému v čitateli a vstupů systému ve jmenovateli.

Tyto filtry se dělí do dvou podkategorií:

- FIR (Finite impulse response)
- IIR (Infinite impulse response)

1.5 Diskrétní Fourierova transformace

Je druh integrální transformace sloužící k převedení signálu z časové oblasti do oblasti frekvenční pomocí funkcí komplexní exponenciály. Je popsána následujícími rovnicemi.

$$D(n) = \sum_{k=0}^{N-1} d(k)e^{-ink2\pi/N}, n = 0, \dots, N-1$$

(Vzorec 6)

$$d(k) = \frac{1}{N} \sum_{n=0}^{N-1} D(n)e^{ink2\pi/N}, k = 0, \dots, N-1$$

(Vzorec 7)

Vzorec 6 je přímá diskrétní Fourierova transformace a vzorec 7 je inverzní diskrétní Fourierova transformace, kde D a d jsou posloupnosti ve frekvenční a v časové oblasti a N je počet prvků těchto posloupností.

Diskrétní Fourierova transformace je důležitá v oblasti zpracování signálů zejména díky převedení operace konvoluce v časové oblasti na operaci násobení v oblasti frekvenční. Toto umožňuje pouhým násobením aplikovat filtr na signál převedený do frekvenční oblasti a posléze pomocí zpětné Fourierovy transformace se převede zpět do časové oblasti, popsáno rovnicí níže,

$$h(t) * y(t) = F^{-1}\{F\{h(t)\} \cdot F\{y(t)\}\} \quad (\text{Vzorec 8})$$

kde F znázorňuje Fourierovu transformaci, F^{-1} znázorňuje zpětnou Fourierovu transformaci, $h(t)$ je filtr a $y(t)$ je zašuměný signál v časové oblasti.

Kvůli složitosti výpočtu Fourierovy transformace na velké úseky signálu se většinou signál dělí na kratší úseky stejné délky. Na tyto úseky se aplikuje rychlá Fourierova transformace, druh Fourierovy transformace používaný v počítačích, provede se filtrace, zpětná rychlá Fourierova transformace a jednotlivé úseky se opět složí dohromady.

1.6 Použití filtrů k redukci šumu

Jakmile je signál ve frekvenční oblasti, lze začít s redukcí šumu. Redukce šumu spočívá v aplikování zvolené jednokanálové funkce na zašuměný signál. Výsledkem je odhadnutý užitečný signál, tedy částečné odstranění šumu. Tento proces je popsán následující rovnicí,

$$\hat{X}_k^{(n)} = H_k^{(n)} \cdot Y_k^{(n)} \quad (\text{Vzorec 9})$$

kde k označuje index frekvenční rámeček, n je index v současném frekvenčním rámci, X je odhadnutý užitečný signál, H je filtr vytvořený podle použité funkce a Y je původní zašuměný signál [1]. V bakalářské práci jsem se omezil na filtry H , které nabývají pouze reálných hodnot.

Ze vzorce 9 je jasné, že původní signál nelze získat úplně a vždy se bude jednat o odhad, přiblížení původnímu signálu. Tedy nikdy nebude dosaženo takových hodnot SNR, jaké byly v původním nezarušeném signálu. Efektivita odstranění šumu je závislá na zvolené metodě pro redukci a na kvalitě zašuměného signálu.

2 Jednokanálové metody

2.1 Definice A-priori SNR a A-posteriori SNR

Většina metod, se kterými pracuji v této bakalářské práci, používaných pro potlačení šumu, odhadnutí a následné odstranění, používá a-priori SNR nebo a-posteriori SNR [1]. A-priori SNR ξ_k a a-posteriori SNR γ_k se definuje následovně.

$$\xi_k \triangleq \frac{\lambda_x(k)}{\lambda_d(k)} \quad (\text{Vzorec 10})$$

$$\gamma_k \triangleq \frac{|Y_k|^2}{\lambda_d(k)} \quad (\text{Vzorec 11})$$

Tedy a-priori SNR ξ_k je poměr odhadu užitečné složky signálu λ_x vůči odhadu šumu λ_d a a-posteriori SNR γ_k je poměr zašuměného signálu Y vůči odhadu šumu λ_d . K výše uvedeným rovnicím mohu poznamenat, že k jejich výpočtu je potřeba znát odhad šumu a v případě a-priori SNR dokonce odhad původního signálu, což nejsou triviální operace.

2.2 Wienerova metoda redukce šumu

Wienerova metoda pro redukci šumu je jedna z prvních metod a je považována za výchozí pro další metody. Původně byla tato metoda odvozená v časové doméně jako FIR filtru N-tého řádu, popsána následovně podle [1].

$$\hat{x}(n) = \sum_{i=0}^N h_i \cdot y(n-i) \quad (\text{Vzorec 12})$$

Chyba odhadu pro tuto metodu je následující

$$e(n) = x(n) - \hat{x}(n) \quad (\text{Vzorec 13})$$

a účelem je najít filtr h_{opt} s koeficienty h_i , který minimalizuje střední kvadratickou hodnotu této chyby.

$$E\{|e(n)|^2\}$$

(Vzorec 14)

Výsledný filtr vypadá následovně,

$$h_{opt} = R_{yy}^{-1} \cdot r_{yy}(0)$$

(Vzorec 15)

kde $r_{yy}(n)$ je korelační vektor původního signálu s původním signálem zpožděným o n prvků a R_{yy} je autokorelační matice.

Ve frekvenční doméně vypadá filtr následovně,

$$S_{ss} = H_{opt} \cdot S_{yy}$$

(Vzorec 16)

kde S_{ss} je výkonová spektrální hustota původního užitečného signálu a S_{yy} je výkonová spektrální hustota pozorovaného signálu, tj. signálu zarušeného. Díky statistické nezávislosti užitečného a rušivého signálu platí,

$$S_{yy} = S_{ss} + S_{dd}$$

(Vzorec 17)

kde S_{dd} je výkonová spektrální hustota rušivého signálu a optimální filtr je pak popsán takto.

$$H_{opt} = \frac{S_{ss}}{S_{dd} + S_{ss}} = \frac{\lambda_s}{\lambda_d + \lambda_s}$$

(Vzorec 18)

Tento vzorec se dá jednoduše přepsat z hlediska a-priori SNR pro každý frekvenční bin následovně.

$$H_k = \frac{\xi_k}{1 + \xi_k}$$

(Vzorec 19)

Vzorec 19 je Wienerova metoda redukce šumu, která minimalizuje střední kvadratickou hodnotu chyby odhadnutého signálu [1].

Tato metoda naráží na několik nedostatků při použití na reálné signály v praxi. Zejména kvůli použití a-priori SNR, kde je třeba znát jak původní signál, tak rušivý signál. Výše uvedené vzorce jsou převzaty z [1]

2.3 Upravená Wienerova metoda

Logickým krokem proto bylo použití a-posteriori SNR γ_k , k odhadu ξ_k , jelikož a-posteriori SNR potřebuje pouze odhad šumu, a tím vznikne nový filtr v následujícím tvaru.

$$H_k^{(n)} = \frac{\xi_k}{1 + \xi_k} \approx \frac{\gamma_k - 1}{\gamma_k} = \frac{|Y_k^{(n)}|^2 - \lambda_d(k)}{|Y_k^{(n)}|^2}$$

(Vzorec 20)

Na vzorci 20 si lze všimnout, že k jeho výpočtu je třeba zašuměný signál Y a odhad šumu λ_d . Což je výrazné zlepšení oproti předešlému výpočtu, nicméně zůstává zde problém odhadu šumu. Toto je problém sám o sobě a v praxi se řeší mnoha způsoby. Jedním z nich je například použití detekce aktivity řeči, kdy v úsecích ticha, když řečník nemluví, změříme signál, který nám představuje odhad šumu.

Tento vzorec je však potřeba omezit, aby zůstaly hodnoty filtru kladné. Po úpravách vypadá finální filtr následovně.

$$H_k^{(n)} = \frac{\max[0, |Y_k^{(n)}|^2 - \lambda_d(k)]}{|Y_k^{(n)}|^2 + \varepsilon}$$

(Vzorec 21)

Tyto úpravy se skládaly z přidání malého čísla ε pro zaručení nenulového jmenovatele, a omezení čitatele pomocí funkce max, aby filtr nedosahoval záporných hodnot.

Tento finální filtr už přináší dobrou redukci šumu a zlepšuje SNR, nicméně na výstupu mohou být slyšet artefakty a zkreslení [1].

2.4 Ephraim a Malah

Ephraim a Malah vytvořili metodu Short-term MMSE estimator [2] neboli krátkodobý odhad minimálních středních kvadratických hodnot chyby. Tato metoda je vyjádřena pomocí následujícího filtru.

$$H_k = \frac{\sqrt{\pi v_k}}{2\gamma_k} \left[(1 + v_k) I_0\left(\frac{v_k}{2}\right) + v_k I_1\left(\frac{v_k}{2}\right) \right] \exp\left(\frac{v_k}{2}\right)$$

(Vzorec 22)

Kde $I_0()$ a $I_1()$ označují modifikované Besselovské funkce nultého a prvního řádu a v_k je následující.

$$v_k \triangleq \frac{\xi_k}{1 + \xi_k} \gamma_k$$

(Vzorec 23)

Tato metoda je efektivní v potlačení šumu a oproti Wienerově filtru má menší zkreslení a počet artefaktů. Oproti Wienerově filtru je ale definovaná jako funkce a-priori SNR ξ_k a a-posteriori SNR γ_k zároveň [1].

Ephraim a Malah vzali v potaz, že lidský sluchový vjem je v logaritmickeém měřítku a vymysleli novou funkci Short-term Log-MMSE estimator [3], která je optimální v logaritmickeém amplitudovém spektru. Tato metoda je vyjádřena pomocí následujícího filtru.

$$H_k = \frac{\xi_k}{1 + \xi_k} \exp\left\{ \frac{1}{2} \int_{v_k}^{\infty} \frac{\exp(-t)}{t} dt \right\}$$

(Vzorec 24)

Problém této metody je, že obsahuje integrál, který se musí počítat v reálném čase. Jelikož je toto jedna z nejlepších metod pro redukci šumu, bylo navrženo mnoho metod pro aproximaci vnořeného integrálu.

2.5 JMAP SAE

Kvůli složitosti výpočtů metod od Ephraima a Malaha, zejména v závislosti na integrálech, exponentech a Besselovských funkcích, a kvůli nedokonalosti Wienerova filtru byla pány Wolfem a Godsillem vyvinuta metoda, která je rychlá a efektivní. Touto metodou je Joint maximum a-posteriori spectral amplitude and phase estimator (JMAP SAE) [4], která je založena na principu společného odhadu amplitudového a fázového spektra, popsána níže.

$$H_k = \frac{\xi_k + \sqrt{\xi_k^2 + 2(1 + \xi_k) \frac{\tilde{\xi}_k}{\gamma_k}}}{2(1 + \xi_k)}$$

(Vzorec 25)

Tato metoda je rychlejší než předchozí metody a v rozmezí od -30 do 30 dB by se měla lišit od metod MMSE a Log-MMSE v řádu 1 dB [1].

3 Sada nástrojů PEASS

Jelikož ve své práci porovnávám jednotlivé filtry, bylo potřeba pro zachování objektivitu použít nějaký nástroj na vyhodnocení jednotlivých filtrů v závislosti na výstupním signálu. Pro tyto účely jsem ve své práci použil sadu nástrojů PEASS [5]. Tato sada nástrojů slouží k vyhodnocení signálu podle 8 stanovených kritérií, kde 4 kritéria jsou objektivní a 4 simulující subjektivní vjem.

Při vyhodnocení signálu předpokládá PEASS, že je signál zkreslený, vzniknou artefakty a je zde přítomné nějaké rušení. Kvůli tomuto předpokladu se signál rozdělí na jednotlivé části, které představují původní signál, zkreslení tohoto signálu, artefakty a rušivou složku, šum, podle následující rovnice.

$$\hat{s}_{ij}(t) - s_{ij}(t) = e_{ij}^{target}(t) + e_{ij}^{interf}(t) + e_{ij}^{artif}(t) \quad (\text{Vzorec 26})$$

Zde $\hat{s}_{ij}(t)$ je odhadnutý signál, $s_{ij}(t)$ je původní užitečný signál, v mém případě řeč, $e_{ij}^{target}(t)$ je odhadnuté zkreslení signálu, $e_{ij}^{interf}(t)$ je odhadnutý šum působící na signál a $e_{ij}^{artif}(t)$ jsou odhadnuté artefakty. K tomuto rozdělení by samozřejmě nemohlo dojít, pokud by nebyl znám původní nezarušený signál, tudíž lze říci, že tuto sadu nástrojů, lze využít jenom v laboratorních podmínkách.

Po rozdělení signálu, následuje porovnání rušivých prvků oproti původnímu signálu a z toho vycházející 4 objektivní kritéria. Toto porovnání je ve formě poměrů energií jednotlivých složek. Jednotlivé poměry jsou popsány rovnicemi níže.

$$SDR_j = 10 \log_{10} \frac{\sum_i \sum_t |s_{ij}(t)|^2}{\sum_i \sum_t |\hat{s}_{ij}(t) - s_{ij}(t)|^2} \quad (\text{Vzorec 27})$$

$$ISR_j = 10 \log_{10} \frac{\sum_i \sum_t |s_{ij}(t)|^2}{\sum_i \sum_t |e_{ij}^{target}(t)|^2} \quad (\text{Vzorec 28})$$

$$SIR_j = 10 \log_{10} \frac{\sum_i \sum_t |s_{ij}(t) + e_{ij}^{target}(t)|^2}{\sum_i \sum_t |e_{ij}^{interf}(t)|^2} \quad (\text{Vzorec 29})$$

$$SAR_j = 10 \log_{10} \frac{\sum_i \sum_t |s_{ij}(t) + e_{ij}^{target}(t) + e_{ij}^{interf}(t)|^2}{\sum_i \sum_t |e_{ij}^{artif}(t)|^2} \quad (\text{Vzorec 30})$$

Vzorec 27 popisuje SDR (Signal to Distortion Ratio), neboli poměr signálu ke zkreslení. Vzorec 28 popisuje ISR (source Image to Spatial distortion Ratio), poměr původního signálu k prostorovému zkreslení. Vzorec 29 popisuje složku SIR (Signal to Interference Ratio), poměr užitečného signálu k rušivému signálu. Vzorec 30 popisuje SAR (Signal to Artifacts Ratio), poměr signálu k vzniklým artefaktům [5].

Dále se zde vyhodnocují 4 kritéria simulující subjektivní vjem, založená na lidském vjemu zvuku pomocí sluchového aparátu. Vyhodnocuje se OPS (Overall Perceptual Score), celkové vnímání nahrávky, TPS (Target-related Perceptual Score), vnímání zkreslení signálu, IPS (Interference-related Perceptual Score), vnímání rušivé složky signálu a APS (Artifacts-related Perceptual Score), vnímání vzniklých artefaktů. Vyhodnocení těchto složek probíhá pomocí PSM (Perceptual Similiarity Measure) v modelu PEMO-Q [7]. Postup výpočtu je popsán následovně,

$$\begin{aligned} q_j^{overall} &= PSM(\hat{s}_j, s_j) \\ q_j^{target} &= PSM(\hat{s}_j, \hat{s}_j - e_j^{target}) \\ q_j^{interf} &= PSM(\hat{s}_j, \hat{s}_j - e_j^{interf}) \\ q_j^{artif} &= PSM(\hat{s}_j, \hat{s}_j - e_j^{artif}) \end{aligned} \quad (\text{Vzorec 31})$$

kde jednotlivé koeficienty q , jsou sdruženy do vektorů a namapovány na jednotlivé koeficienty OPS, TPS, IPS a APS.

4 Implementace

4.1 Nahrávání

Před samotnou implementací jednotlivých metod pro redukci šumu musím nejdříve samostatný signál nahrát. Pro nahrání užitečného signálu, v mém případě lidské řeči, jsem použil objekt *audiorecorder*, který slouží pro nahrávání audio nahrávek v prostředí MATLAB. Použil jsem metodu *recordblocking*, která nahraje audio signál na vstupu, v mém případě mikrofon připojený k počítači, po určenou dobu. Objektu *audiorecorder* jsem nastavil následující kritéria:

- Vzorkovací frekvence: 44,1 kHz
- Počet bitů pro kvantování: 16

Z objektu *audiorecorder* jsem získal audio data pomocí příkazu *getaudiodata*. Posléze jsem nahrávku uložil jako soubor typu WAV pomocí příkazu *audiowrite*. Takto nahrané audiosignály byly dále zašuměny a byly na ně aplikovány metody pro redukci šumu.

4.2 Zašumění

Nahrávky získané způsobem popsáným výše byly zašuměny několika způsoby. Prvním úkolem bylo zašumět nahrávku tak, aby měla požadovanou úroveň SNR. Tohoto jsem dosáhl aplikací níže popsané odvozené rovnice.

Výpočet výkonu signálu P .

$$P = \frac{1}{N} \sum_i^N x_i^2$$

(Vzorec 32)

Poměr výkonu užitečného signálu k výkonu rušivého signálu k .

$$k = \frac{\frac{1}{N} \sum_i^N s_i^2}{\frac{1}{N} \sum_i^N n_i^2}$$

(Vzorec 33)

Výpočet SNR v decibelech.

$$\text{SNR} = 10 \cdot \log_{10} \left(\frac{1}{\alpha^2} \cdot k \right)$$

(Vzorec 34)

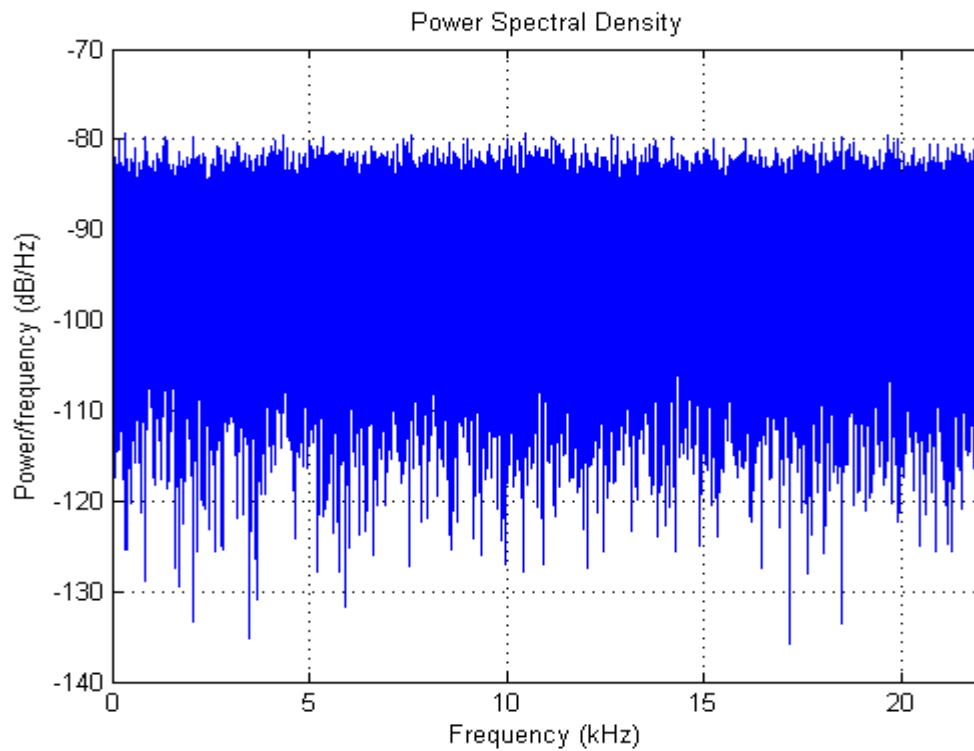
Ve vzorci 34 znám požadovanou hodnotu SNR v decibelech, ale chci zjistit hodnotu α , koeficient, kterým budu násobit šum, abych dosáhl požadované hodnoty SNR. Koeficient k je poměr výkonů užitečného signálu k šumu. Tento poměr můžu modifikovat právě koeficientem α . Po úpravě získám takovýto vztah.

$$\text{Koeficient } \alpha: \alpha = \sqrt{\frac{1}{(10^{(\text{SNR}_{dB}/10)}/k)}}$$

(Vzorec 35)

4.2.1 Bílý šum

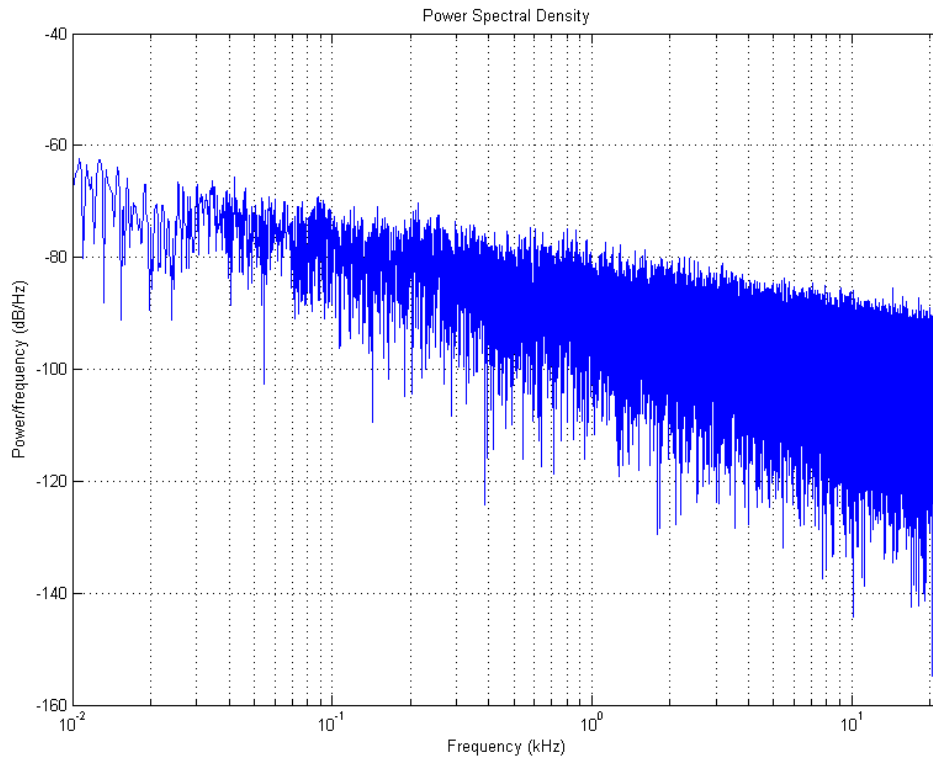
První druh zašumění jsem zvolil tzv. bílý šum. Toto je náhodný signál s výkonovou spektrální hustotou rovnoměrnou v celé frekvenční oblasti. To znamená, že všechny stejně široká frekvenční pásma mají stejný výkon. Tento signál se vytvoří velice jednoduše vygenerováním náhodných čísel požadované délky. V mém případě je tato délka shodná s délkou nahrávky pro zašumění. Ukázka spektra bílého šumu se nachází na obrázku 1.



Obrázek 1: Spektrální hustota bílého šumu

4.2.2 Růžový šum

Růžový šum je druh signálu, kde výkonová spektrální hustota je nepřímo úměrná frekvenci signálu. Hlavní vlastností takového šumu je stejný výkon na každou oktávu, tj. frekvence dvakrát vyšší nebo poloviční. Ukázka takového signálu je na obrázku 2.



Obrázek 2: Spektrální hustota růžového šumu s logaritmickou osou x

4.2.3 Šum typu babble

Tento druh šumu představuje nahrávání v jedné místnosti jako je např. jídelna, kancelář apod., kde hovoří zároveň mnoho lidí. Takový druh zašumění může v praxi nastat velmi často, a proto jsem jej testoval.

4.2.4 Zašumění jiným řečníkem

Posledním druhem zašumění, které v této práci uvažuji, je zašumění jiným řečníkem, tedy simulace situace, kdy mluví více lidí najednou. Pro tento druh zašumění jsem si stanovil 4 situace. Tyto jednotlivé situace představovali zašumění mužského hlasu hlasem ženským, zašumění mužského hlasu jiným mužským hlasem, zašumění ženského hlasu mužským hlasem a zašumění ženského hlasu jiným ženským hlasem.

4.3 Metody redukce šumu

Díky prostředí MATLAB byla implementace jednotlivých metod pro redukci šumu zjednodušená, zejména díky funkcím už obsažených ve vývojovém prostředí, které zprostředkovávají složitější matematické funkce např. Besselovské funkce. Jednotlivé metody jsem aplikoval na signál, který byl rozdělen na rámce o délce 1024 prvků. Po rozdělení vstupního signálu jsem na jednotlivé rámce aplikoval FFT, abych získal signál ve frekvenční oblasti a mohl na něj aplikovat filtr H představující způsob redukce šumu podle dané rovnice. Po aplikování filtru jsem provedl IFFT, zpětnou rychlou Fourierovu transformaci, a nakonec jsem spojil jednotlivé rámce do jednoho celku. Vzniklá nahrávka je pak připravená k vyhodnocení.

5 Vyhodnocení

5.1 Ideální situace

V první části vyhodnocení se zaměřím na ideální situaci, kdy je znám jak původní signál, tak signál který je zdrojem šumu, viz. vzorec 2. Tento případ nastane pouze v laboratorních podmínkách, nicméně je důležitý pro zhodnocení jednotlivých metod. V pokusech jsem se soustředil na 5 úrovní SNR a to -10, -5, 0, 5, 10 dB. Jednotlivé nahrávky byly dlouhé 5 sekund, vzorkované pomocí vzorkovací frekvence 44,1 kHz a úroveň kvantování byla stanovená 16 bity. Ve většině nahrávek je cílovým signálem mužská řeč, pokud není uvedeno jinak.

5.1.1 Wienerův filtr

Jako první zhodnotím výsledky Wienerova filtru. Tento filtr dokázal zvýšit SNR nahrávky a potlačit šum adekvátně, nicméně přinesl nežádoucí vedlejší efekty ve formě artefaktů, zejména v úsecích nahrávek, kde docházelo k přechodu z řeči na ticho. I přesto bylo dosaženo velmi kvalitního zlepšení signálu oproti původní zašuměné nahrávce. V příloze A se nachází kompletní tabulka výsledků pro Wienerovu metodu, dále budou následovat tabulky, které jsou pouze jejími částmi. V tabulce 1 jsou uvedeny výsledky kritérií vypočítané pomocí sady nástrojů PEASS pro signál zašuměný bílým šumem.

Tabulka 1: Výsledky PEASS pro Wienerovu metodu a bílý šum

	SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	23,4	18,3	13,9	11,5	14,9
	ISR [dB]	31,6	26,8	21,4	18	22,4
	SIR [dB]	25,8	20,7	16,8	14,9	17,8
	SAR [dB]	31,3	26,5	21,3	18,6	22,7
	OPS [%]	54	57	49	42	53
	TPS [%]	69	68	77	83	75
	IPS [%]	69	73	73	72	74
	APS [%]	70	52	27	18	27

Z tabulky je vidět zlepšení v hodnotách SNR, reprezentované v tabulce zkratkou SDR, zejména u signálů s nižším počátečním SNR. Zajímavé je zlepšení hodnot u signálu s hodnotou SNR -10 oproti signálu s hodnotou SNR -5. Lze si také všimnout zvyšující se hodnoty IPS, tedy zvýšení míry vnímání zarušení. Celkově však došlo k velmi dobré redukci šumu na úroveň skoro neslyšitelnou. V tabulce 2 jsou uvedeny výsledky při zašumění signálu růžovým šumem.

Tabulka 2: Výsledky PEASS pro Wienerovu metodu a růžový šum

	SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	19,8	16,6	13,1	10,4	7,8
	ISR [dB]	28	24,8	19,7	16,3	12,6
	SIR [dB]	22,3	19,4	16,6	14,2	11,8
	SAR [dB]	27,2	23,8	20,1	16,9	14
	OPS [%]	63	59	42	31	22
	TPS [%]	64	67	80	86	86
	IPS [%]	76	74	75	74	74
	APS [%]	62	54	21	7	3

U růžového šumu jsou hodnoty objektivních kritérií celkově nižší, ale nenastává zde problém u signálu se SNR -10. Opět zde bylo slyšet jemné zkreslení, ale pouze při delším poslechu.

V tabulce 3 jsou znázorněny výsledky u šumu typu babble. Opět ve všech případech byl signál krásně odrušen.

Tabulka 3: Výsledky PEASS pro Wienerovu metodu a šum typu babble

	SNR [dB]	10	5	0	-5	-10
Babble	SDR [dB]	20,6	17,5	14,6	12,2	10
	ISR [dB]	28,4	24,8	21,1	17,9	15
	SIR [dB]	23,4	20,7	18,3	16,4	14,5
	SAR [dB]	27	23,8	20,8	18,2	15,9
	OPS [%]	79	74	61	43	32
	TPS [%]	65	65	72	78	84
	IPS [%]	83	83	81	79	77
	APS [%]	67	52	36	21	10

Wienerova metoda fungovala dobře i pro zarušení jednoho řečníka jiným řečníkem.

5.1.2 MMSE

Dále zhodnotím výsledky u metody Ephraima a Malaha short-term MMSE. Tato metoda vykazovala o něco horší úroveň odstranění šumu oproti Wienerově metodě, ale na rozdíl od ní vznikal menší počet artefaktů. Zajímavostí byl však slyšitelný šum v nahrávkách s nízkou hodnotou SNR, kde jak cílovým signálem, tak šumem byla lidská řeč.

Tabulka 4: Výsledky PEASS pro metodu MMSE a ženský hlas zašuměný mužským hlasem

	SNR [dB]	10	5	0	-5	-10
Ž x M	SDR [dB]	17,9	14,9	12,3	10,1	8,2
	ISR [dB]	25,2	22,7	19,6	16,5	13,6
	SIR [dB]	19,5	16,7	14,5	12,7	11,2
	SAR [dB]	28,4	24,6	21,2	18,3	15,9
	OPS [%]	97	89	88	73	60
	TPS [%]	93	81	63	66	71
	IPS [%]	95	93	88	85	81
	APS [%]	87	87	64	37	34

V tabulce 4 je vidět ukázka takového signálu, v tomto případě zašumění ženské řeči mužským řečníkem. Efektivita metody stoupala v závislosti na snižujícím se SNR, je zde vidět, že signál se SNR -10 se zlepšil na hodnotu 8,2 SDR. Zbylé tabulky jsou v příloze B.

5.1.3 Log-MMSE

Tato metoda od Ephraima a Malaha dopadla podle kritérií PEASS lépe než předešlé 2 metody. U nahrávek s nízkými hodnotami SNR už nebylo možno rozeznat šum, nicméně zůstalo zde zkreslení signálu zejména u nahrávek se SNR -10. Jednotlivé tabulky se nacházejí v příloze C. V tabulce 5 jsou znázorněny výsledky PEASS pro nahrávku s cílovou ženskou řečí zašuměnou mužskou řečí.

Tabulka 5: Výsledky PEASS pro metodu Log-MMSE a ženský hlas zašumění mužským hlasem

	SNR [dB]	10	5	0	-5	-10
Ž x M	SDR [dB]	19,1	16	13,3	10,9	8,7
	ISR [dB]	27,5	22,9	18,9	15,3	12,3
	SIR [dB]	21,2	18,7	16,5	14,7	13,1
	SAR [dB]	28,5	25,1	22	19,1	16,6
	OPS [%]	98	94	92	77	62
	TPS [%]	93	88	67	65	70
	IPS [%]	96	92	89	85	83
	APS [%]	87	84	66	44	39

Je zde patrné zlepšení oproti výsledkům v tabulce 4.

5.1.4 JMAP SAE

Tato metoda vykazovala podobné výsledky jako předchozí metoda Log-MMSE. Vzhledem k menší výpočetní náročnosti metody JMAP SAE jsem byl výsledky velmi překvapen. Jak objektivní výsledky PEASS, tak subjektivní vjem byl skoro k nerozeznání oproti předešlé metodě Log-MMSE, která má daleko složitější výpočet. Výsledky PEASS najdete v příloze D.

5.2 Situace s odhadnutým šumem

Jedná se o situaci, kdy je neznámý původní zdroj šumu, ale získá se jako rozdíl signálů na 2 mikrofonech. Tuto situaci jsem nasimuloval pomocí databáze impulzních odezev [8], kde jsou uloženy měření impulzní odezvy v testovací místnosti za použití pole mikrofonů. Pro simulaci jsem zvolil situaci, kde byly použity 2 mikrofony. Tento výpočet vychází ze znalosti filtru h , což je relativní impulzní odezva, tj. vyjadřuje rozdíl, jak cílový signál doputuje na jednotlivé mikrofony, v závislosti na umístění a vzdálenosti těchto mikrofonů. Postup je popsán následujícími rovnicemi,

$$\begin{aligned}x_L(n) &= s(n) + y_L(n) \\x_R(n) &= h * s(n) + y_R(n)\end{aligned}$$

(Vzorec 36)

kde x_L a x_R jsou signály přicházející na levý a pravý mikrofon, s je cílový signál, v našem případě řeč, a y_L a y_R představují šum. Kdy v tomto případě je levý mikrofon blíže ke zdroji zvuku. Odečtením zpožděného signálu na pravém mikrofону od signálu získaného na mikrofону levém, získám odhad šumu, který mohu použít pro jednotlivé metody redukce šumu.

5.2.1 Wienerova metoda

Z tabulky 6 lze vyčíst, že zlepšení při použití odhadnutého šumu není tak velké jako v ideálním případě. Nicméně stále se podařilo výrazně zlepšit signál, i když u signálů s nízkou úrovní SNR je stále slyšet šum a vznikají zde artefakty. Nastal zde zajímavý případ, kdy u signálu zašuměného bílým šumem se SNR 0 vyšly lepší hodnoty než u signálu se SNR 5, tento efekt se objevuje i u ostatních metod.

Tabulka 6: Výsledky PEASS pro metodu Wienerova filtru pro reálný odhad bílého šumu

	SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	17,5	11	12,5	4,7	3,1
	ISR [dB]	24,5	25	21,5	16	14,7
	SIR [dB]	19,5	11,3	13,6	5,1	3,4
	SAR [dB]	27	24,9	23,7	17,1	16
	OPS [%]	44	37	37	25	17
	TPS [%]	72	74	78	70	73
	IPS [%]	58	48	47	17	16
	APS [%]	66	53	47	16	7

Při zašumění signálu růžovým šumem se začali více projevovat efekty odhadnutého šumu a vzdálení výsledků PEASS od ideální situace. Výsledky jsou zobrazeny v tabulce 7. Dosažené zisky byly minimální oproti signálu zašuměného šumem bílým, avšak poslechově bylo slyšet zlepšení oproti původní nahrávce.

Tabulka 7: Výsledky PEASS pro metodu Wienerova filtru pro reálný odhad růžového šumu

	SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	11,8	7	2,3	-2,8	-7,5
	ISR [dB]	27	23,9	20,8	17,5	14,5
	SIR [dB]	12,1	7	2,2	-3,1	-8,1
	SAR [dB]	28,2	25,5	22,7	21,1	20
	OPS [%]	28	32	24	16	9
	TPS [%]	58	68	81	64	51
	IPS [%]	43	27	22	9	6
	APS [%]	68	57	30	29	22

U nahrávek s řečníky, tj. kdy jak užitečný signál, tak rušivý signál byly lidská řeč, dosahovala Wienerova metoda podobných výsledků jako u zašumění růžovým šumem. U zašumění šumem typu babble, vycházeli výsledky naznačující dokonce zhoršení signálu, nicméně poslechově zde žádný rozdíl nebyl. Všechny tabulky pro tuto metodu jsou v příloze E.

5.2.2 MMSE

Tato metoda už vykazovala větší zlepšení signálu. Šum byl potlačen velmi dobře ve všech úrovních SNR. Zkreslení nebylo tak výrazné jako u předchozí Wienerovy metody, avšak v době ticha byly slyšet artefakty a zbytkový šum, zejména při nižších úrovních SNR. V následující tabulce 8 jsou výsledky získané pomocí metody PEASS.

Tabulka 8: Výsledky PEASS pro metodu MMSE pro reálný odhad bílého šumu

	SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	18,9	14,1	14,9	9,1	7,5
	ISR [dB]	26	24,9	24,2	18,1	16,6
	SIR [dB]	20,2	14,8	15,9	10,7	8,9
	SAR [dB]	31,2	25,3	26	18,6	17,4
	OPS [%]	51	52	55	24	20
	TPS [%]	61	81	72	96	96
	IPS [%]	73	71	71	65	63
	APS [%]	76	31	45	0	0

Lze si všimnout, že u úrovní SNR -5 a -10 se APS neboli vnímání artefaktů dostalo k nule, zatímco vnímání zkreslení signálu TPS se blíží maximu, tj. při vyhodnocení artefakty splynuli se zkreslením. V tabulce 9 jsou uvedeny výsledky pro odhad růžového šumu. Oproti předchozí Wienerově metodě bylo dosaženo značných zlepšení zejména u signálů s nižší hodnotou SNR, kde byl šum potlačen velmi dobře, nicméně se zde projevovало výrazné zkreslení signálu řeči. Podobně fungovala tato metoda na signál zašumění šumem typu babble.

Tabulka 9: Výsledky PEASS pro metodu MMSE pro reálný odhad růžového šumu

	SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	14,7	10,8	7,1	3,4	0,5
	ISR [dB]	25,4	22,9	19,8	14,7	9,8
	SIR [dB]	15,3	11,5	7,9	4,7	2,2
	SAR [dB]	25,9	20,6	16,7	12,6	9,3
	OPS [%]	51	48	27	14	11
	TPS [%]	63	85	95	94	86
	IPS [%]	69	70	68	60	52
	APS [%]	64	22	1	0	0

U signálů zašuměných lidskou řečí fungovala metoda o něco lépe, jak je znázorněno v tabulce 10. Zbylé tabulky jsou v příloze F.

Tabulka 10: Výsledky PEASS pro metodu MMSE a mužský hlas zašumění ženským hlasem

	SNR [dB]	10	5	0	-5	-10
M x Ž	SDR [dB]	18,2	14,9	11,6	8,5	5,6
	ISR [dB]	26,2	24,8	22,4	19,1	15,5
	SIR [dB]	19,5	16,1	13,1	10,3	7,9
	SAR [dB]	30,4	26	21,8	17,8	14,2
	OPS [%]	27	30	49	42	29
	TPS [%]	55	55	69	88	87
	IPS [%]	75	68	65	69	73
	APS [%]	87	83	57	12	5

5.2.3 Log-MMSE

Tato metoda fungovala velmi podobně jako předchozí metoda MMSE, co se výsledků PEASS, např. pro bílý šum zobrazený v tabulce 10, a zkruslení signálu týče. Pro signál s bílým šumem, zde došlo oproti předešlé metodě MMSE k daleko lepšímu potlačení původního šumu, dalo by se říci na úroveň těžko slyšitelnou běžným posluchačem a to pro všechny úrovně vstupního SNR.

Tabulka 11: Výsledky PEASS pro metodu Log-MMSE pro reálný odhad bílého šumu

	SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	19,8	14,4	15,5	9,7	8,2
	ISR [dB]	27,7	25,7	23,7	16,6	15,2
	SIR [dB]	21,7	15,6	17,2	11,9	10,3
	SAR [dB]	30,1	24,9	25,8	19	17,8
	OPS [%]	50	48	46	29	25
	TPS [%]	60	72	71	86	83
	IPS [%]	68	64	63	67	67
	APS [%]	70	47	47	6	4

V tabulce 12 jsou zobrazeny výsledky při odhadu růžového šumu. Výsledky jsou opět velmi podobné předešlé metodě MMSE.

Tabulka 12: Výsledky PEASS pro metodu Log-MMSE pro reálný odhad růžového šumu

	SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	14,8	10,8	7,1	3,4	0,4
	ISR [dB]	28	23,2	19,1	13,9	9,2
	SIR [dB]	15,6	11,8	8,2	4,8	2,2
	SAR [dB]	25,4	20,4	16,6	12,4	9,1
	OPS [%]	31	46	34	19	16
	TPS [%]	57	83	91	87	82
	IPS [%]	59	68	65	63	65
	APS [%]	78	23	5	1	0

Pro nahrávky zašuměné lidskou řečí fungovala skoro identicky s předchozí metodou MMSE. Zbylé tabulky jsou v příloze G

5.2.4 JMAP SAE

Tato metoda dokázala potlačit šum stejně dobře jako metoda Log-MMSE, ale měla horší zkreslení signálu a větší výskyt artefaktů. Vezmu-li v potaz rozdíl ve výpočetní složitosti těchto 2 metod, metoda JMAP SAE byla čtyřikrát rychlejší na výpočet, tak bych mohl říci, že v praxi by se nejvíce osvědčila právě tato metoda. Na následujících tabulkách 13 a 14 je vidět, že metoda má podobný výstup z PEASS jako Log-MMSE pro bílý a růžový šum. Tabulky pro ostatní průběhy jsou v příloze H.

Tabulka 14: Výsledky PEASS pro metodu JMAP SAE pro reálný odhad bílého šumu

	SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	19,8	14,3	15,4	9,5	8,1
	ISR [dB]	30	27,4	25,5	17,7	16,1
	SIR [dB]	21,2	15,3	16,8	11,5	10
	SAR [dB]	31	25,2	26,1	18,9	17,8
	OPS [%]	48	49	49	31	27
	TPS [%]	61	78	73	91	91
	IPS [%]	68	70	66	66	65
	APS [%]	73	28	40	3	3

Tabulka 13: Výsledky PEASS pro metodu JMAP SAE pro reálný odhad růžového šumu

	SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	14,8	10,7	7	3,3	0,3
	ISR [dB]	29,5	24,2	19,9	14,4	9,5
	SIR [dB]	15,5	11,7	8,1	4,8	2,2
	SAR [dB]	25,7	20,5	16,7	12,5	9,2
	OPS [%]	50	46	32	23	14
	TPS [%]	63	85	94	90	85
	IPS [%]	67	68	66	60	61
	APS [%]	62	21	3	1	0

5.3 Oprava zabarveného signálu

U předchozí úlohy nastal problém s odlišností spektrální hustoty odhadnutého rušivého signálu oproti původnímu rušivému signálu. Toto bylo způsobené výpočtem odhadu šumu, kde byly odečteny signály na mikrofonech a posléze zde zbyl rozdíl signálů šumu, popsany vzorcem 37.

$$h * x_L - x_R = h * y_L - y_R$$

(Vzorec 37)

Zde h představuje impulzní odezvu filtru, x_L a x_R představují zarušený signál nahraný na levém a pravém mikrofону a y_L a y_R představují šum nahraný na těchto mikrofonech.

Za předpokladu, že šum na pravém mikrofону je pouze původní šum zpožděný a zesílený filtrem impulzní odezvy pak,

$$y_R(n) = \alpha y(n - D)$$

(Vzorec 38)

kde hodnota n je současný index prvku, D je zpoždění signálu a α je zesílení získané z impulzní odezvy, y je původní rušivý signál, maximální hodnota z filtru h , která je zobrazena na obrázku 3. Dále zavedu předpoklad, že impulzní odezva filtru je pouze zpoždění signálu, v tomto případě pak,

$$h * y_L(n) \approx y(n - Q)$$

(Vzorec 39)

kde Q je zpoždění signálu v závislosti na impulzní odezvě. Za těchto předpokladů lze napsat, že odhadnutý šum y_{est} se rovná,

$$y_{est}(n) \approx y(n - Q) - \alpha y(n - D)$$

(Vzorec 40)

což se dá přepsat do následujícího tvaru,

$$y_{est}(n) = y(n) - \alpha y(n - R)$$

(Vzorec 41)

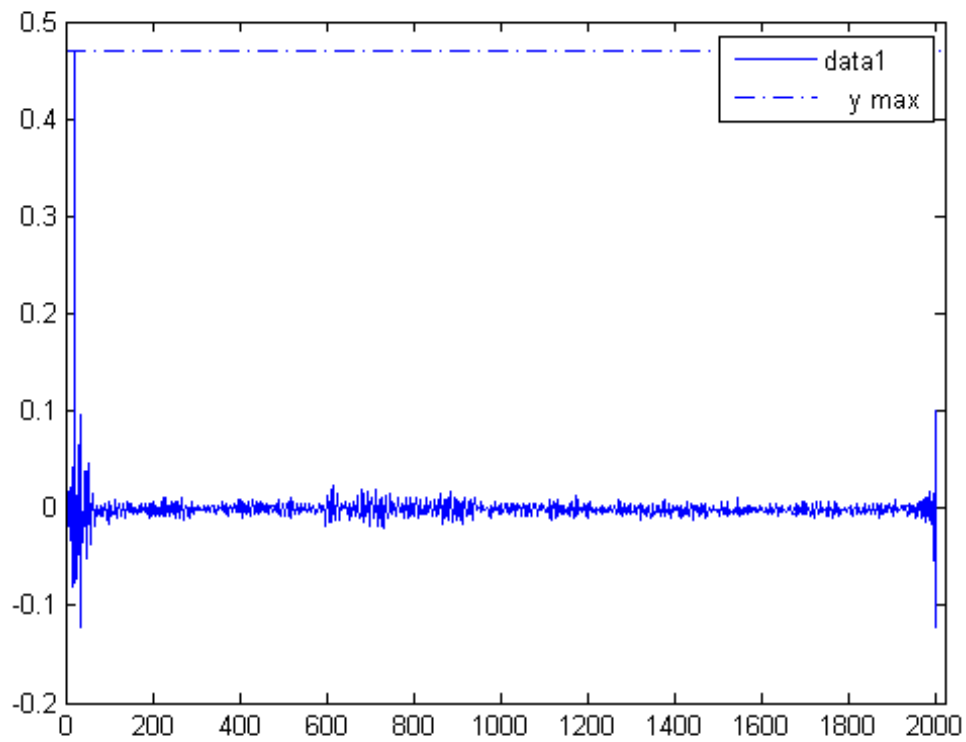
, kde R je zpoždění signálu.

Vzorec pro opravu tohoto spektra je pak následující

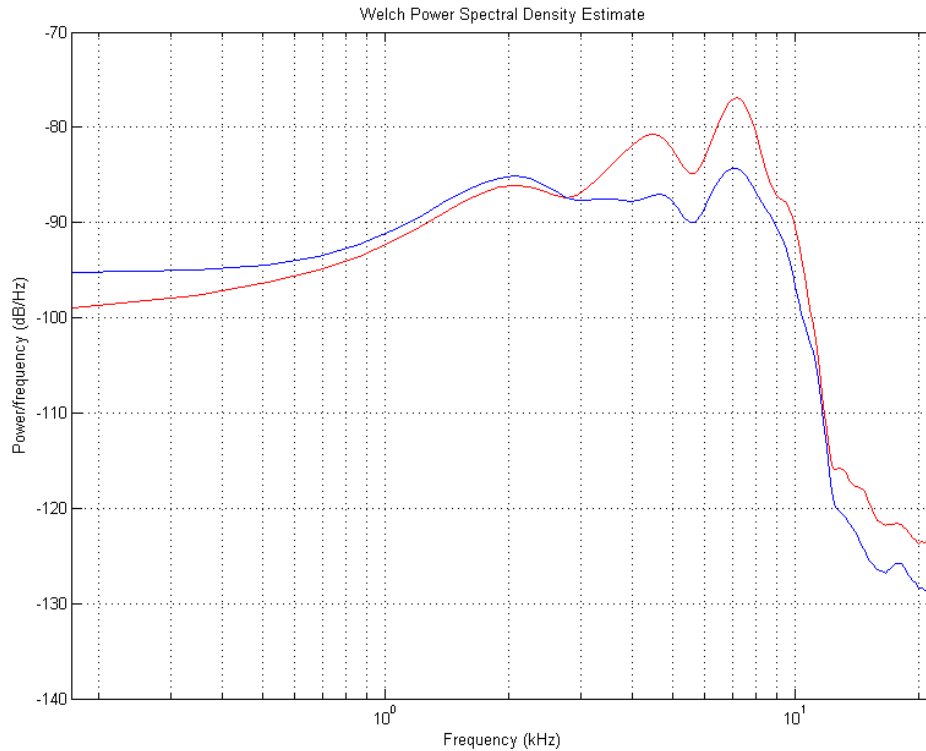
$$\hat{y}(n) = y_{est}(n) + \alpha y_{est}(n - R)$$

(Vzorec 42)

Jelikož znám hodnoty y_{est} a α , stačilo zjistit hodnotu R . Tuto hodnotu získám empiricky postupným zvětšováním a kontrolou výstupního spektra, dokud se spektrum nepodobá původnímu spektru signálu.

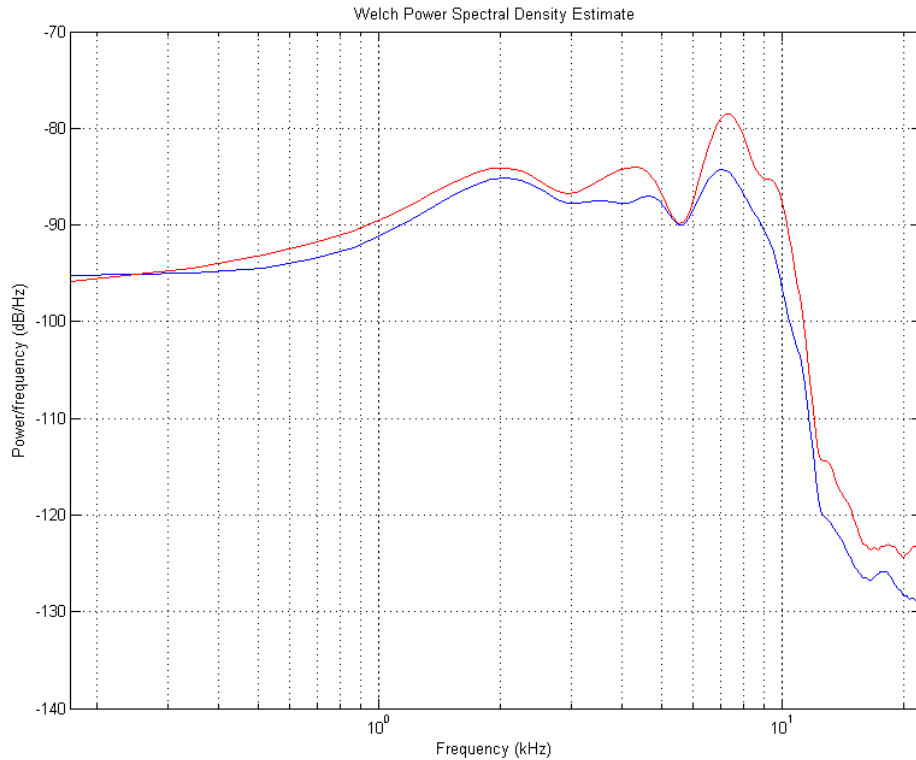


Obrázek 3: Impulzní odezva



Obrázek 4: Spektrální hustota originálního a odhadnutého šumu

Na obrázku 4 je zobrazený graf spektrální hustoty originálního signálu šumu modře a odhadnutého šumu červeně. Lze si všimnout, že grafy se podobají, nicméně nejsou příliš stejné, zejména v oblasti do 6 kHz, která je nejdůležitější pro odšumění frekvencí, kde se vyskytuje lidská řeč. Po aplikování metody na opravu spektra s hodnotou $R = 4$, tj. odstup signálu byl 4 prvky, se spektrum v této oblasti zlepšilo. Upravené spektrum je zobrazené na obrázku 5.



Obrázek 5: Spektrální hustota upraveného šumu a originálního šumu

Na obrázku 5 je vidět zlepšení a přiblížení se původnímu signálu v oblasti do 6 kHz. To má za následek zlepšení redukce šumu v této oblasti. Avšak zhoršil se průběh spektra ve zbylém frekvenčním rozsahu. Toto by však nemělo mít velký vliv na výsledný subjektivní vjem, zejména pokud je původní užitečný signál lidská řeč, která takto vysokých frekvencí nedosahuje.

5.4 Porovnání opraveného signálu

Všechny metody po opravení spektra odhadnutého rušivého signálu, vykazovaly zlepšení u všech objektivních kritérií PEASS.

V tabulce 15 jsou zobrazeny výsledky průměrné změny objektivních kritérií PEASS pro jednotlivé metody v závislosti na typech šumu.

Tabulka 15: Porovnání výsledků PEASS pro zabarvený signál a signál s opraveným spektrem

		Wiener	MMSE	Log-MMSE	JMAP SAE
Bílý šum	SDR [dB]	0,52	0,20	0,26	0,28
	ISR [dB]	0,48	0,00	-0,12	-0,02
	SIR [dB]	0,74	0,26	0,38	0,36
	SAR [dB]	-1,24	-0,08	-0,18	-0,10
Růžový šum	SDR [dB]	0,36	0,66	0,76	0,72
	ISR [dB]	0,24	0,08	-0,18	-0,02
	SIR [dB]	0,36	0,86	0,96	0,92
	SAR [dB]	-1,34	0,00	-0,10	-0,04
Babble	SDR [dB]	0,18	0,90	1,02	0,96
	ISR [dB]	-0,42	0,22	0,16	0,26
	SIR [dB]	0,14	1,10	1,14	1,18
	SAR [dB]	-2,00	0,20	0,16	0,14
M x M	SDR [dB]	1,04	0,70	0,84	0,82
	ISR [dB]	-1,12	-0,08	-0,56	-0,36
	SIR [dB]	1,06	0,86	1,04	0,94
	SAR [dB]	-2,36	0,16	0,06	0,18
Ž x M	SDR [dB]	1,04	0,70	0,84	0,82
	ISR [dB]	-1,12	-0,08	-0,56	-0,36
	SIR [dB]	1,06	0,86	1,04	0,94
	SAR [dB]	-2,36	0,16	0,06	0,18
Ž x Ž	SDR [dB]	1,24	0,86	1,02	1,00
	ISR [dB]	-1,36	-0,14	-0,62	-0,46
	SIR [dB]	1,30	1,10	1,42	1,28
	SAR [dB]	-2,58	0,00	-0,10	-0,02
M x Ž	SDR [dB]	1,00	0,64	0,80	0,64
	ISR [dB]	-0,42	0,10	-0,18	0,76
	SIR [dB]	1,06	0,86	1,08	0,76
	SAR [dB]	-2,04	0,10	0,04	0,30

Z tabulky 15 lze vyčíst, že opravou spektra bylo dosaženo zisků u kritéria SDR okolo 1 dB pro signály zašuměné lidskou řečí. Pro signály zašuměné šumem typu babble je vidět, že Wienerova metoda nezvládla signál vylepšit a to ani s opraveným spektrem. U signálů zašuměných bílým šumem nebyl zisk tak veliký, jelikož bylo dosaženo vysokého zlepšení už se zabarveným šumem. Celkově se zisk pohyboval od 0,2 do 1,2 dB.

Závěr

V průběhu této bakalářské práce jsem se seznámil s problematikou redukce šumu a různými jednokanálovými metodami k tomu určenými. Při vyhodnocování nahrávek jsem si vyzkoušel práci se sadou nástrojů PEASS, díky čemuž jsem zjistil, jaká kritéria se používají k vyhodnocení signálu. S tímto nástrojem se mi pracovalo velmi dobře, zejména kvůli dobré dokumentaci a podrobným návodům. Využil jsem svých dosud získaných znalostí k implementování těchto metod ve vývojovém prostředí MATLAB. Toto prostředí jsem si vybral zejména kvůli předchozím zkušenostem nabytých v mém bakalářském studiu a vzhledem k užívání tohoto programu v běžné praxi. Při použití metod pro redukci šumu došlo k několika zajímavým situacím jako je například větší zlepšení u signálu se SNR 0 oproti signálu se SNR 5 při zašumění bílým šumem.

Při použití dvoukanálového filtru k získání šumu jsem zjistil efektivitu jednotlivých metod v reálné situaci, kde je zdroj šumu neznámý, a seznámil se s problematikou odhadu šumu. Tyto výsledky byly v souladu s mým původním očekáváním. Nejúčinnějšími metodami na redukci šumu byly Log-MMSE a JMAP SAE, a naopak Wienerova metoda se prokázala jako neefektivní při úlohách s odhadnutým šumem, zejména u signálu s nízkými hodnotami SNR. Vzhledem k výpočetní náročnosti metody Log-MMSE bych pro využití v praxi doporučil metodu JMAP SAE. Nakonec jsem navrhl způsob jak opravit zabarvení odhadnutého rušivého signálu neboli jeho výkonového spektra tak, aby se přiblížil výkonovému spektru původního rušivého signálu. Díky čemuž došlo ke zlepšení redukce šumu ve většině případů.

Díky zhotovení mé bakalářské práce jsem měl možnost proniknout do problematiky redukce šumu, která mě velmi zaujala. Průběh zpracování audio nahrávek byl zajímavý, zejména zjištění efektivitu jednotlivých metod na nahrávku pomocí poslechu.

Použitá literatura

- 1 TASHEV, Ivan J. *Sound capture and processing: practical approaches*. Chichester: John Wiley, 2009, xx, 365 s. ISBN 978-0-470-31983-3.
- 2 EPHRAIM, Y. a D. MALAH. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing* [online]. 1984, vol. 32, issue 6, s. 1109-1121 [cit. 2015-05-14]. DOI: 10.1109/tassp.1984.1164453.
- 3 EPHRAIM, Y. a D. MALAH. Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing* [online]. 1985, vol. 33, issue 2, s. 443-445 [cit. 2015-05-14]. DOI: 10.1109/tassp.1985.1164550.
- 4 WOLFE, Patrick J. a Simon J. GODSILL. Efficient Alternatives to the Ephraim and Malah Suppression Rule for Audio Signal Enhancement. *EURASIP Journal on Advances in Signal Processing* [online]. 2003, vol. 2003, issue 10, s. 1043-1051 [cit. 2015-05-14]. DOI: 10.1155/s1110865703304111.
- 5 EMIYA, Valentin, Emmanuel VINCENT, Niklas HARLANDER a Volker HOHMANN. Subjective and Objective Quality Assessment of Audio Source Separation. *IEEE Transactions on Audio, Speech, and Language Processing* [online]. 2011, vol. 19, issue 7, s. 2046-2057 [cit. 2015-05-14]. DOI: 10.1109/tasl.2011.2109381.
- 6 BERNAT, Petr. *Akustika, vznik a šíření zvuku...* [online]. [cit. 2015-05-14].
Dostupné z: http://homen.vsb.cz/~ber30/texty/varhany/anatomie/pi_staly_akustika.htm

- 7 HUBER, R. a B. KOLLMEIER. PEMO-Q; A New Method for Objective Audio Quality Assessment Using a Model of Auditory Perception. *IEEE Transactions on Audio, Speech and Language Processing* [online]. 2006, vol. 14, issue 6, s. 1902-1911 [cit. 2015-05-14]. DOI: 10.1109/tasl.2006.883259.
- 8 HADAD Elior, Florian HEESE, Peter VARY a Sharon GANNOT. Multichannel audio database in various acoustic environments. In: *2014 14th International Workshop on Acoustic Signal Enhancement (IWAENC)* [online]. 2014 [cit. 2015-05-14]. DOI: 10.1109/iwaenc.2014.6954309.

Obsah přiloženého CD

- Text bakalářské práce ve formátu PDF.
- Zdrojové kódy programu.
- Použitá verze sady nástrojů PEASS.
- Nahrávky použité při vyhodnocení.
- Tabulky s výsledky vyhodnocení nahrávek.

Příloha B – Tabulky pro metodu MMSE při ideální situaci

	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	21,9	17,7	13,4	11	14,3	Babble	SDR [dB]	20	17,1	14,2	11,4	8,8
	ISR [dB]	26,5	25,5	22,3	19,2	22,7		ISR [dB]	26	24,5	21,8	18,3	14,8
	SIR [dB]	24,1	19,2	15,4	13,4	16,1		SIR [dB]	22,3	19,3	16,7	14,5	12,5
	SAR [dB]	34,4	28	21,8	18,9	23,2		SAR [dB]	27,7	23,9	20,5	17,4	14,8
	OPS [%]	19	57	50	44	63		OPS [%]	97	89	60	35	30
	TPS [%]	69	61	81	90	76		TPS [%]	70	62	77	86	88
	IPS [%]	59	71	77	75	81		IPS [%]	94	91	85	81	77
APS [%]	87	61	18	11	33	APS [%]	76	57	20	8	7		
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	19,1	16,2	12,8	9,8	7,1	Ž x M	SDR [dB]	17,9	14,9	12,3	10,1	8,2
	ISR [dB]	25,8	24,6	21,1	17,7	13,6		ISR [dB]	25,2	22,7	19,6	16,5	13,6
	SIR [dB]	20,8	18	15	12,5	10,2		SIR [dB]	19,5	16,7	14,5	12,7	11,2
	SAR [dB]	28,3	24,3	20,2	16,7	13,8		SAR [dB]	28,4	24,6	21,2	18,3	15,9
	OPS [%]	54	58	50	27	15		OPS [%]	97	89	88	73	60
	TPS [%]	60	69	81	92	92		TPS [%]	93	81	63	66	71
	IPS [%]	71	74	75	77	73		IPS [%]	95	93	88	85	81
APS [%]	69	42	21	2	1	APS [%]	87	87	64	37	34		
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x M	SDR [dB]	19,2	16,6	13,9	11,2	9,2	Ž x Ž	SDR [dB]	17,2	14,5	11,9	9,5	7,4
	ISR [dB]	26,3	25	22,8	19,9	16,9		ISR [dB]	25,3	23,2	20	16,3	12,9
	SIR [dB]	20,8	18,1	15,5	13	11,3		SIR [dB]	18,4	15,8	13,6	11,7	10,2
	SAR [dB]	30,4	26,8	23,4	20,2	17,5		SAR [dB]	28,8	25,1	21,7	18,5	16,1
	OPS [%]	78	69	63	55	54		OPS [%]	98	88	81	73	63
	TPS [%]	82	64	62	72	75		TPS [%]	93	78	62	62	68
	IPS [%]	90	80	76	76	72		IPS [%]	96	92	87	82	82
APS [%]	87	75	54	29	39	APS [%]	87	85	74	46	36		
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x Ž	SDR [dB]	20	17,3	14,5	11,8	9,5		SDR [dB]	20	17,3	14,5	11,8	9,5
	ISR [dB]	26	24,8	22,6	19,6	16,5		ISR [dB]	26	24,8	22,6	19,6	16,5
	SIR [dB]	22,3	19,5	17	14,7	12,9		SIR [dB]	22,3	19,5	17	14,7	12,9
	SAR [dB]	30,7	26,6	22,8	19,4	16,3		SAR [dB]	30,7	26,6	22,8	19,4	16,3
	OPS [%]	44	63	62	53	47		OPS [%]	44	63	62	53	47
	TPS [%]	72	62	63	76	82		TPS [%]	72	62	63	76	82
	IPS [%]	80	78	75	73	72		IPS [%]	80	78	75	73	72
APS [%]	87	74	51	32	21	APS [%]	87	74	51	32	21		

Příloha C – Tabulky pro metodu Log-MMSE při ideální situaci

Bílý šum	SNR [dB]	10	5	0	-5	-10	Babble	SNR [dB]	10	5	0	-5	-10
	SDR [dB]	23,9	18,6	14,2	11,7	15,1		SDR [dB]	21,4	18,1	15	12,1	9,5
	ISR [dB]	34,1	27,3	21,2	17,4	21,9		ISR [dB]	29,9	25,1	20,7	16,7	13,3
	SIR [dB]	25,4	20,8	17,1	15,2	18,1		SIR [dB]	23,9	21,2	18,7	16,5	14,5
	SAR [dB]	33,4	27,4	22	19,3	23,3		SAR [dB]	28	24,6	21,5	18,6	16,1
	OPS [%]	23	59	53	47	52		OPS [%]	97	88	66	44	29
	TPS [%]	72	62	77	80	71		TPS [%]	71	62	73	83	84
	APS [%]	64	72	75	72	72		APS [%]	95	90	86	81	78
Růžový šum	SNR [dB]	10	5	0	-5	-10	Ž x M	SNR [dB]	10	5	0	-5	-10
	SDR [dB]	20,2	17	13,4	10,5	7,7		SDR [dB]	19,1	16	13,3	10,9	8,7
	ISR [dB]	29,2	25,1	19,6	15,6	11,7		ISR [dB]	27,5	22,9	18,9	15,3	12,3
	SIR [dB]	22,3	19,6	16,8	14,3	11,8		SIR [dB]	21,2	18,7	16,5	14,7	13,1
	SAR [dB]	28,2	24,4	20,7	17,6	14,4		SAR [dB]	28,5	25,1	22	19,1	16,6
	OPS [%]	64	61	53	32	20		OPS [%]	98	94	92	77	62
	TPS [%]	60	66	78	87	87		TPS [%]	93	88	67	65	70
	APS [%]	76	76	75	76	72		APS [%]	96	92	89	85	83
M x M	SNR [dB]	10	5	0	-5	-10	Ž x Ž	SNR [dB]	10	5	0	-5	-10
	SDR [dB]	20,8	17,7	14,9	12,1	9,8		SDR [dB]	18,4	15,4	12,7	10,2	8
	ISR [dB]	31,9	27,2	23	19,1	15,8		ISR [dB]	27,4	22,9	18,7	14,8	11,6
	SIR [dB]	22,5	19,8	17,5	15	13,1		SIR [dB]	20,1	17,7	15,6	13,8	12,2
	SAR [dB]	29,9	26,4	23,2	20,3	17,7		SAR [dB]	28,6	25,2	22,1	19,2	16,8
	OPS [%]	84	79	71	60	47		OPS [%]	98	91	84	74	65
	TPS [%]	88	73	62	65	72		TPS [%]	93	85	65	65	68
	APS [%]	90	83	78	78	79		APS [%]	96	91	85	81	81
M x Ž	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
	SDR [dB]	21,4	18,3	15,3	12,5	10,1		SDR [dB]	18,4	15,4	12,7	10,2	8
	ISR [dB]	31,3	27	22,9	19	15,6		ISR [dB]	27,4	22,9	18,7	14,8	11,6
	SIR [dB]	23,9	21,2	18,8	16,6	14,8		SIR [dB]	20,1	17,7	15,6	13,8	12,2
	SAR [dB]	30,1	26,4	22,8	19,7	16,7		SAR [dB]	28,6	25,2	22,1	19,2	16,8
	OPS [%]	72	69	68	59	50		OPS [%]	98	91	84	74	65
	TPS [%]	89	69	62	66	73		TPS [%]	93	85	65	65	68
	APS [%]	85	77	76	76	76		APS [%]	96	91	85	81	81

Příloha D – Tabulky pro metodu JMAP SAE při ideální situaci

	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
	SDR [dB]	23,7	18,5	14	11,6	15		SDR [dB]	21,3	17,9	14,8	12	9,4
	ISR [dB]	35,6	29,1	22,7	18,8	23,4		ISR [dB]	31,5	26,7	22,1	17,9	14,3
	SIR [dB]	25	20,2	16,5	14,6	17,4		SIR [dB]	23,3	20,5	18	15,9	13,9
	SAR [dB]	34,1	27,8	22	19,3	23,4		SAR [dB]	28	24,4	21,2	18,2	15,6
	OPS [%]	20	57	51	46	55	Babble	OPS [%]	97	88	64	38	27
	TPS [%]	69	62	79	86	71		TPS [%]	71	63	76	86	86
	IPS [%]	61	71	75	75	74		IPS [%]	94	91	86	81	77
	APS [%]	87	57	22	14	32		APS [%]	77	53	23	9	8
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
	SDR [dB]	20,1	16,9	13,3	10,4	7,6		SDR [dB]	18,9	15,8	13,1	10,8	8,7
	ISR [dB]	30,8	26,9	21,1	16,9	12,8		ISR [dB]	28,9	24,2	20	16,3	13,1
	SIR [dB]	21,8	19	16,1	13,7	11,3		SIR [dB]	20,6	18	15,8	14,1	12,6
	SAR [dB]	28,4	24,5	20,6	17,3	14,2		SAR [dB]	28,6	25	21,8	18,9	16,4
	OPS [%]	57	59	51	30	16	Ž x M	OPS [%]	98	94	91	75	61
	TPS [%]	60	68	80	90	89		TPS [%]	93	82	65	65	70
	IPS [%]	72	75	75	76	73		IPS [%]	96	94	89	86	82
	APS [%]	67	42	22	4	2		APS [%]	87	85	63	38	37
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
	SDR [dB]	20,5	17,5	14,7	11,9	9,7		SDR [dB]	18,1	15,2	12,6	10,1	8
	ISR [dB]	33,5	28,8	24,3	20,3	16,8		ISR [dB]	29	24,3	20	15,9	12,4
	SIR [dB]	21,9	19,3	16,8	14,3	12,6		SIR [dB]	19,5	17	14,9	13,1	11,7
	SAR [dB]	30,4	26,8	23,5	20,4	17,7		SAR [dB]	28,9	25,3	22,1	19	16,6
	OPS [%]	82	72	66	55	46		OPS [%]	99	92	85	75	65
	TPS [%]	88	68	62	66	74		TPS [%]	93	78	66	62	67
	IPS [%]	90	81	75	76	78		IPS [%]	97	93	87	82	81
	APS [%]	87	75	50	25	19		APS [%]	87	83	71	50	44
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
	SDR [dB]	21,3	18,2	15,2	12,4	10,1		SDR [dB]	21,3	17,9	14,8	12	9,4
	ISR [dB]	32,6	28,2	24	20	16,5		ISR [dB]	31,5	26,7	22,1	17,9	14,3
	SIR [dB]	23,3	20,7	18,2	16	14,3		SIR [dB]	23,3	20,5	18	15,9	13,9
	SAR [dB]	30,5	26,6	23	19,7	16,6		SAR [dB]	28	24,4	21,2	18,2	15,6
	OPS [%]	54	68	66	58	50		OPS [%]	97	88	64	38	27
	TPS [%]	85	67	62	67	75		TPS [%]	71	63	76	86	86
	IPS [%]	80	77	76	76	77		IPS [%]	94	91	86	81	77
	APS [%]	87	74	51	33	24		APS [%]	77	53	23	9	8
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
	SDR [dB]	21,3	18,2	15,2	12,4	10,1		SDR [dB]	21,3	17,9	14,8	12	9,4
	ISR [dB]	32,6	28,2	24	20	16,5		ISR [dB]	31,5	26,7	22,1	17,9	14,3
	SIR [dB]	23,3	20,7	18,2	16	14,3		SIR [dB]	23,3	20,5	18	15,9	13,9
	SAR [dB]	30,5	26,6	23	19,7	16,6		SAR [dB]	28	24,4	21,2	18,2	15,6
	OPS [%]	54	68	66	58	50		OPS [%]	97	88	64	38	27
	TPS [%]	85	67	62	67	75		TPS [%]	71	63	76	86	86
	IPS [%]	80	77	76	76	77		IPS [%]	94	91	86	81	77
	APS [%]	87	74	51	33	24		APS [%]	77	53	23	9	8
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
	SDR [dB]	21,3	18,2	15,2	12,4	10,1		SDR [dB]	21,3	17,9	14,8	12	9,4
	ISR [dB]	32,6	28,2	24	20	16,5		ISR [dB]	31,5	26,7	22,1	17,9	14,3
	SIR [dB]	23,3	20,7	18,2	16	14,3		SIR [dB]	23,3	20,5	18	15,9	13,9
	SAR [dB]	30,5	26,6	23	19,7	16,6		SAR [dB]	28	24,4	21,2	18,2	15,6
	OPS [%]	54	68	66	58	50		OPS [%]	97	88	64	38	27
	TPS [%]	85	67	62	67	75		TPS [%]	71	63	76	86	86
	IPS [%]	80	77	76	76	77		IPS [%]	94	91	86	81	77
	APS [%]	87	74	51	33	24		APS [%]	77	53	23	9	8

Příloha E – Tabulky pro Wienerovu metodu při odhadu šumu

	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	17,5	11	12,5	4,7	3,1		SDR [dB]	9,6	4,6	-0,4	-5,4	-10,4
	ISR [dB]	24,5	25	21,5	16	14,7		ISR [dB]	30,7	29,6	27,6	24,8	21,8
	SIR [dB]	19,5	11,3	13,6	5,1	3,4		SIR [dB]	9,6	4,6	-0,5	-5,5	-10,6
	SAR [dB]	27	24,9	23,7	17,1	16		SAR [dB]	31,6	30,9	30	29,5	29,5
	OPS [%]	44	37	37	25	17		OPS [%]	9	13	18	18	16
	TPS [%]	72	74	78	70	73		TPS [%]	52	49	59	81	90
	IPS [%]	58	48	47	17	16		IPS [%]	33	20	12	7	5
APS [%]	66	53	47	16	7		APS [%]	79	76	72	63	43	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	11,8	7	2,3	-2,8	-7,5		SDR [dB]	12	7,2	2,3	-2,6	-7,6
	ISR [dB]	27	23,9	20,8	17,5	14,5		ISR [dB]	27,5	25,3	22,6	19,5	16,2
	SIR [dB]	12,1	7	2,2	-3,1	-8,1		SIR [dB]	12,2	7,2	2,2	-2,8	-7,9
	SAR [dB]	28,2	25,5	22,7	21,1	20		SAR [dB]	28,5	25,9	23,2	21,2	20,3
	OPS [%]	28	32	24	16	9		OPS [%]	23	12	9	11	12
	TPS [%]	58	68	81	64	51		TPS [%]	63	52	47	43	38
	IPS [%]	43	27	22	9	6		IPS [%]	70	47	28	15	9
APS [%]	68	57	30	29	22		APS [%]	84	84	81	76	71	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x M	SDR [dB]	12,1	7,2	2,3	-2,6	-7,6		SDR [dB]	12,1	7,4	2,6	-2,3	-7,2
	ISR [dB]	30,1	28,2	25,4	21,9	18		ISR [dB]	25,1	22,8	20,5	18	15,1
	SIR [dB]	12,2	7,3	2,3	-2,7	-7,7		SIR [dB]	12,4	7,5	2,6	-2,5	-7,5
	SAR [dB]	29	26,4	23,5	21,4	20,4		SAR [dB]	28,1	25,6	23	20,9	19,9
	OPS [%]	18	16	16	16	11		OPS [%]	24	12	13	14	17
	TPS [%]	53	51	53	62	79		TPS [%]	69	51	47	44	48
	IPS [%]	46	31	17	10	5		IPS [%]	70	47	28	17	12
APS [%]	75	75	74	75	71		APS [%]	84	81	78	73	69	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x Ž	SDR [dB]	12,5	7,7	2,8	-2,2	-7,2		SDR [dB]	12,5	7,7	2,8	-2,2	-7,2
	ISR [dB]	28,7	26,8	24,2	21,2	17,8		ISR [dB]	28,7	26,8	24,2	21,2	17,8
	SIR [dB]	12,7	7,7	2,7	-2,3	-7,3		SIR [dB]	12,7	7,7	2,7	-2,3	-7,3
	SAR [dB]	28,7	26	22,9	20,7	19,6		SAR [dB]	28,7	26	22,9	20,7	19,6
	OPS [%]	21	17	17	15	12		OPS [%]	21	17	17	15	12
	TPS [%]	55	50	51	62	75		TPS [%]	55	50	51	62	75
	IPS [%]	37	19	12	7	5		IPS [%]	37	19	12	7	5
APS [%]	71	71	71	70	69		APS [%]	71	71	71	70	69	

Příloha F – Tabulky pro metodu MMSE při odhadu šumu

	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	18,9	14,1	14,9	9,1	7,5		SDR [dB]	14	9,9	6,1	2,7	-0,3
	ISR [dB]	26	24,9	24,2	18,1	16,6		ISR [dB]	25,5	23,3	19,9	16	12,1
	SIR [dB]	20,2	14,8	15,9	10,7	8,9		SIR [dB]	14,5	10,6	7	4	1,6
	SAR [dB]	31,2	25,3	26	18,6	17,4		SAR [dB]	23,4	18,8	14,5	11	8,3
	OPS [%]	51	52	55	24	20		OPS [%]	22	53	42	24	12
	TPS [%]	61	81	72	96	96		TPS [%]	54	69	88	94	91
	IPS [%]	73	71	71	65	63		IPS [%]	70	69	73	70	68
APS [%]	76	31	45	0	0		APS [%]	87	55	10	1	0	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	14,7	10,8	7,1	3,4	0,5		SDR [dB]	16,4	13	9,5	6,3	3,3
	ISR [dB]	25,4	22,9	19,8	14,7	9,8		ISR [dB]	26,2	24,5	21,5	17,7	13,7
	SIR [dB]	15,3	11,5	7,9	4,7	2,2		SIR [dB]	17,1	13,6	10,4	7,4	4,8
	SAR [dB]	25,9	20,6	16,7	12,6	9,3		SAR [dB]	29,4	24,7	20,2	16	12,5
	OPS [%]	51	48	27	14	11		OPS [%]	83	42	50	55	47
	TPS [%]	63	85	95	94	86		TPS [%]	79	63	58	70	78
	IPS [%]	69	70	68	60	52		IPS [%]	92	82	72	71	73
APS [%]	64	22	1	0	0		APS [%]	87	87	74	44	24	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x M	SDR [dB]	17,2	13,7	10,4	7,2	4,4		SDR [dB]	15,7	12,3	9	6,1	3,3
	ISR [dB]	26,4	25,4	23,3	20,3	16,7		ISR [dB]	25,9	24	20,6	16,6	12,6
	SIR [dB]	17,9	14,4	11,1	8,1	5,4		SIR [dB]	16,2	12,8	9,8	7,1	4,9
	SAR [dB]	29,6	25,3	21,3	17,8	14,7		SAR [dB]	29,2	24,7	20,6	16,9	13,6
	OPS [%]	38	22	49	37	32		OPS [%]	86	39	40	59	49
	TPS [%]	54	53	73	90	93		TPS [%]	78	61	56	67	77
	IPS [%]	82	62	65	72	71		IPS [%]	93	81	72	74	75
APS [%]	87	84	58	5	2		APS [%]	87	87	80	53	25	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x Ž	SDR [dB]	18,2	14,9	11,6	8,5	5,6		SDR [dB]	18,2	14,9	11,6	8,5	5,6
	ISR [dB]	26,2	24,8	22,4	19,1	15,5		ISR [dB]	26,2	24,8	22,4	19,1	15,5
	SIR [dB]	19,5	16,1	13,1	10,3	7,9		SIR [dB]	19,5	16,1	13,1	10,3	7,9
	SAR [dB]	30,4	26	21,8	17,8	14,2		SAR [dB]	30,4	26	21,8	17,8	14,2
	OPS [%]	27	30	49	42	29		OPS [%]	27	30	49	42	29
	TPS [%]	55	55	69	88	87		TPS [%]	55	55	69	88	87
	IPS [%]	75	68	65	69	73		IPS [%]	75	68	65	69	73
APS [%]	87	83	57	12	5		APS [%]	87	83	57	12	5	

Příloha G – Tabulky pro metodu Log-MMSE při odhadu šumu

	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	19,8	14,4	15,5	9,7	8,2		SDR [dB]	13,8	9,7	5,9	2,4	-0,5
	ISR [dB]	27,7	25,7	23,7	16,6	15,2		ISR [dB]	29,4	24,8	20,2	15,9	11,8
	SIR [dB]	21,7	15,6	17,2	11,9	10,3		SIR [dB]	14,6	10,7	7,1	4,1	1,7
	SAR [dB]	30,1	24,9	25,8	19	17,8		SAR [dB]	23,2	18,6	14,4	10,9	8,2
	OPS [%]	50	48	46	29	25		OPS [%]	19	52	44	26	15
	TPS [%]	60	72	71	86	83		TPS [%]	55	68	85	89	82
	IPS [%]	68	64	63	67	67		IPS [%]	65	68	71	71	69
APS [%]	70	47	47	6	4		APS [%]	87	53	17	3	1	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	14,8	10,8	7,1	3,4	0,4		SDR [dB]	16,8	13,2	9,6	6,3	3,3
	ISR [dB]	28	23,2	19,1	13,9	9,2		ISR [dB]	30,5	26,2	21,6	17,3	13,3
	SIR [dB]	15,6	11,8	8,2	4,8	2,2		SIR [dB]	17,6	14,1	10,8	7,8	5
	SAR [dB]	25,4	20,4	16,6	12,4	9,1		SAR [dB]	28,8	24,5	20,1	16	12,4
	OPS [%]	31	46	34	19	16		OPS [%]	87	41	52	55	47
	TPS [%]	57	83	91	87	82		TPS [%]	84	69	59	68	75
	IPS [%]	59	68	65	63	65		IPS [%]	92	80	71	71	74
APS [%]	78	23	5	1	0		APS [%]	87	87	71	43	26	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x M	SDR [dB]	17,6	14	10,6	7,4	4,5		SDR [dB]	16	12,5	9,3	6,3	3,4
	ISR [dB]	32,3	28,6	24,4	20,3	16,5		ISR [dB]	28,3	24,3	20	15,7	11,9
	SIR [dB]	18,5	14,9	11,6	8,5	5,7		SIR [dB]	16,8	13,5	10,4	7,7	5,3
	SAR [dB]	29	25,2	21,4	17,9	14,8		SAR [dB]	28,5	24,5	20,6	16,9	13,5
	OPS [%]	67	44	55	48	34		OPS [%]	81	37	42	58	49
	TPS [%]	61	57	64	79	83		TPS [%]	83	66	58	65	74
	IPS [%]	90	84	70	73	76		IPS [%]	91	78	71	73	75
APS [%]	87	87	54	23	7		APS [%]	87	87	79	54	28	
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x Ž	SDR [dB]	18,8	15,2	11,9	8,7	5,8		SDR [dB]	16	12,5	9,3	6,3	3,4
	ISR [dB]	31	27,2	23,1	19,1	15,3		ISR [dB]	28,3	24,3	20	15,7	11,9
	SIR [dB]	20,1	16,8	13,8	11	8,5		SIR [dB]	16,8	13,5	10,4	7,7	5,3
	SAR [dB]	29,7	25,7	21,7	17,8	14,2		SAR [dB]	28,5	24,5	20,6	16,9	13,5
	OPS [%]	38	31	54	45	28		OPS [%]	81	37	42	58	49
	TPS [%]	58	57	66	80	82		TPS [%]	83	66	58	65	74
	IPS [%]	81	68	70	71	75		IPS [%]	91	78	71	73	75
APS [%]	87	83	54	21	7		APS [%]	87	87	79	54	28	

Příloha H – Tabulky pro metodu JMAP SAE při odhadu šumu

	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Bílý šum	SDR [dB]	19,8	14,3	15,4	9,5	8,1		SDR [dB]	13,8	9,7	5,8	2,4	-0,6
	ISR [dB]	30	27,4	25,5	17,7	16,1		ISR [dB]	30,1	25,1	20,4	16	11,8
	SIR [dB]	21,2	15,3	16,8	11,5	10		SIR [dB]	14,5	10,6	7,1	4	1,7
	SAR [dB]	31	25,2	26,1	18,9	17,8		SAR [dB]	23,4	18,7	14,5	11	8,3
	OPS [%]	48	49	49	31	27		OPS [%]	20	51	44	25	13
	TPS [%]	61	78	73	91	91		TPS [%]	55	67	87	92	90
	APS [%]	68	70	66	66	65		IPS [%]	66	67	71	71	69
	APS [%]	73	28	40	3	3		APS [%]	87	56	13	2	0
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
Růžový šum	SDR [dB]	14,8	10,7	7	3,3	0,3		SDR [dB]	16,7	13,1	9,5	6,2	3,2
	ISR [dB]	29,5	24,2	19,9	14,4	9,5		ISR [dB]	32,3	27,4	22,4	17,8	13,6
	SIR [dB]	15,5	11,7	8,1	4,8	2,2		SIR [dB]	17,4	13,9	10,6	7,6	5
	SAR [dB]	25,7	20,5	16,7	12,5	9,2		SAR [dB]	29,3	24,7	20,2	16	12,4
	OPS [%]	50	46	32	23	14		OPS [%]	88	45	51	55	47
	TPS [%]	63	85	94	90	85		TPS [%]	82	67	59	68	76
	IPS [%]	67	68	66	60	61		IPS [%]	93	82	71	72	73
	APS [%]	62	21	3	1	0		APS [%]	87	87	73	45	26
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x M	SDR [dB]	17,5	13,9	10,4	7,2	4,3		SDR [dB]	15,9	12,4	9,1	6,1	3,3
	ISR [dB]	34,1	29,8	25,3	20,9	16,8		ISR [dB]	30,2	25,5	20,8	16,3	12,2
	SIR [dB]	18,3	14,7	11,4	8,4	5,6		SIR [dB]	16,6	13,2	10,2	7,5	5,2
	SAR [dB]	29,4	25,3	21,4	17,9	14,8		SAR [dB]	29,1	24,7	20,7	16,9	13,6
	OPS [%]	64	18	53	47	32		OPS [%]	84	38	41	58	49
	TPS [%]	59	55	64	80	86		TPS [%]	82	65	57	65	75
	IPS [%]	90	62	69	73	76		IPS [%]	91	79	72	73	75
	APS [%]	87	87	56	19	4		APS [%]	87	87	80	55	26
	SNR [dB]	10	5	0	-5	-10		SNR [dB]	10	5	0	-5	-10
M x Ž	SDR [dB]	18,8	15,2	11,9	8,7	5,8		SDR [dB]	15,9	12,4	9,1	6,1	3,3
	ISR [dB]	31	27,2	23,1	19,1	15,3		ISR [dB]	30,2	25,5	20,8	16,3	12,2
	SIR [dB]	20,1	16,8	13,8	11	8,5		SIR [dB]	16,6	13,2	10,2	7,5	5,2
	SAR [dB]	29,7	25,7	21,7	17,8	14,2		SAR [dB]	29,1	24,7	20,7	16,9	13,6
	OPS [%]	38	31	54	45	28		OPS [%]	84	38	41	58	49
	TPS [%]	58	57	66	80	82		TPS [%]	82	65	57	65	75
	IPS [%]	81	68	70	71	75		IPS [%]	91	79	72	73	75
	APS [%]	87	83	54	21	7		APS [%]	87	87	80	55	26