

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ**

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

**NÁSTROJ PRO SUBJEKTIVNÍ HODNOCENÍ KVALITY
RESTAURACE AUDIOSIGNÁLŮ**

TOOL FOR SUBJECTIVE QUALITY EVALUATION OF AUDIOSIGNALS RESTORATION

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Viktor Černý

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Pavel Záviška

BRNO 2018

Bakalářská práce

bakalářský studijní obor **Audio inženýrství**

Ústav telekomunikací

Student: Viktor Černý

ID: 186602

Ročník: 3

Akademický rok: 2017/18

NÁZEV TÉMATU:

Nástroj pro subjektivní hodnocení kvality restaurace audiosignálů

POKYNY PRO VYPRACOVÁNÍ:

Prostudujte možné modely znehodnocení audiosignálů a seznamte se s různými typy poslechových testů sloužících k subjektivnímu zhodnocení kvality zvukových signálů. Seznamte se rovněž s programovacím jazykem Java, tvorbou grafického uživatelského rozhraní a knihovnamy pro práci se zvukovými soubory. V Javě následně implementujte program s GUI, ve kterém bude možné vytvořit a následně provést několik možných subjektivních psychoakustických testů a vhodným způsobem vyexportovat výsledky testování a provést základní statistickou analýzu. Funkčnost programu otestujte vytvořením a vyhodnocením několika subjektivních testů.

DOPORUČENÁ LITERATURA:

[1] Eckel, B. Thinking in Java 4th edition, 2006. ISBN 978-0131872486.

[2] Madisetti, V. K. The Digital Signal Processing Handbook, 2nd Edition. 2009. ISBN 978-1420045635.

Termín zadání: 5.2.2018

Termín odevzdání: 29.5.2018

Vedoucí práce: Ing. Pavel Záviška

Konzultant:

prof. Ing. Jiří Mišurec, CSc.
předseda oborové rady

UPOZORNĚNÍ:

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

ABSTRAKT

Bakalářská práce se zabývá teorií znehodnocení a rekonstrukce audio signálů, objektivními testy a subjektivními metodami pro vyhodnocení kvality rekonstruovaných signálů. Nejprve je popsáno několik druhů poškození audio signálů a následně rekonstrukční typy pro jejich rekonstrukci. Po rekonstrukčních typech je zmíněno šest objektivních a šest subjektivních testů pro vyhodnocení kvality audio signálů. Další část bakalářské práce je věnována popisu programu implementovaného v programovacím jazyce Java. Program slouží k navrhování poslechových testů typu MOS, MUSHRA a ABX, a také k provádění samotného testování. Na konci práce jsou výsledky poslechových testů, které byly použity pro otestování programu, a uživatelské hodnocení aplikace.

KLÍČOVÁ SLOVA

audio, znehodnocení signálů, rekonstrukce signálů, objektivní testy, subjektivní hodnocení, grafické uživatelské rozhraní, aplikace, poslechové testy

ABSTRACT

Bachelor thesis deals with the theory of audio signal depreciation and reconstruction, with objective tests and subjective methods of reconstructed signal evaluation. At first, there are described several types of depreciation of audio signals and then reconstruction types for their reconstruction. After the reconstruction, types six objective and six subjective tests are mentioned for the evaluation of audio signal quality. The next part of the Bachelor thesis is the description of the program that was implemented in the programming language Java. The program is designed to design MOS, MUSHRA and ABX listening tests as well as to use for the actual testing. At the end of the thesis are the results of the listening tests that were used to test the program and the user ratings of the application.

KEYWORDS

audio, depreciation of signals, reconstruction of signals, objective tests, subjective evaluation, graphical user interface, application, listening tests

ČERNÝ, Viktor. *Nástroj pro subjektivní hodnocení kvality restaurace audio signálů*. Brno, 2018, 50 s. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Pavel Závíška

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Nástroj pro subjektivní hodnocení kvality restaurace audio signálů“ jsem vypracoval(a) samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

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

Brno

.....

podpis autora(-ky)

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu bakalářské práce panu Ing. Pavlovi Záviškovi za odborné vedení, konzultace, trpělivost a podnětné návrhy k práci. Chtěl bych se také poděkovat panu doc. Mgr. Pavlovi Rajmicovi, Ph.D. a panu doc. Ing. Jířimu Schimmelovi, Ph.D. pro umožnění použití laboratoře SC1.23 pro poslechové testy. Děkuji každému, kdo poslechové testy trpělivě udělal a tím mi pomohl nasbírat data pro bakalářskou práci.

Brno

.....

podpis autora(-ky)

OBSAH

Úvod	10
1 Poškození audio signálů	11
1.1 Chybějící vzorky	11
1.2 Clipping	13
1.2.1 Hard clipping	13
1.2.2 Soft clipping	15
1.3 Zašumění	16
1.4 Clicking	16
1.5 Crackling	16
1.6 Hissing	17
1.7 Rumbling	17
2 Rekonstrukce audio signálů	18
2.1 Declipping	18
2.2 Inpainting	18
2.3 Denoising	19
2.4 Declicking	19
2.5 Decrackling	21
2.6 Dehissing	21
2.7 Derumbling	21
3 Vyhodnocení kvality	22
3.1 Objektivní testy	22
3.1.1 SNR	22
3.1.2 PEMO-Q	23
3.1.3 PEAQ	23
3.1.4 PESQ	24
3.1.5 POLQA	24
3.1.6 ViSQOL	25
3.2 Subjektivní metody	25
3.2.1 MOS	26
3.2.2 MUSHRA	27
3.2.3 ABX	27
3.2.4 Duo-trio test	28
3.2.5 Triangle test	28
3.2.6 Paired comparison test	28

4	Popis programu	30
4.1	JavaFX	30
4.2	Hlavní nabídka	30
4.3	Seznam testů	31
4.4	Správce	32
4.5	Přidání nového testu	32
4.6	Testy	34
5	Testování programu	37
5.1	Rozdíly mezi formáty	37
5.2	Výsledky testů ABX	37
5.3	Výsledky testu MUSHRA	39
5.4	Vyhodnocení testů	41
5.5	Hodnocení programu	42
6	Závěr	43
	Literatura	44
	Seznam symbolů, veličin a zkratk	48
	Seznam příloh	49
A	Obsah přiloženého CD	50

SEZNAM OBRÁZKŮ

1.1	Nepoškozený sinusový signál o kmitočtu 440 Hz.	12
1.2	Sinusový signál o kmitočtu 440 Hz se 4 chybějícími vzorky.	12
1.3	Sinusový signál o kmitočtu 440 Hz s 20 chybějícími vzorky.	12
1.4	Jednostranné spektrum sinusového signálu o kmitočtu 440 Hz.	14
1.5	Sinusový signál o kmitočtu 440 Hz po hard clippingu.	14
1.6	Jednostranné spektrum sinusového signálu o kmitočtu 440 Hz po hard clippingu.	14
1.7	Sinusový signál o kmitočtu 440 Hz po soft clippingu.	15
1.8	Jednostranné spektrum sinusového signálu o kmitočtu 440 Hz po soft clippingu.	16
2.1	Sinusový signál o kmitočtu 440 Hz s clickingem.	20
2.2	Sinusový signál o kmitočtu 440 Hz po manuálním declickingu.	20
4.1	Hlavní nabídka programu.	30
4.2	Seznam testů.	31
4.3	Přihlašovací okno správce.	32
4.4	Přidání nového testu MOS.	33
4.5	Přidání nového testu MUSHRA.	33
4.6	Přidání nového testu ABX.	34
4.7	Testovací okno MOS.	35
4.8	Testovací okno MUSHRA.	35
4.9	Testovací okno ABX.	36
5.1	Četnost správných výběrů testu ABX – WAV vs MP3 320kbps	37
5.2	Četnost správných výběrů testu ABX – WAV vs MP3 128kbps	38
5.3	Četnost správných výběrů testu ABX – MP3 320kbps vs MP3 128kbps	38
5.4	Střední hodnoty testu MUSHRA pro klasickou hudbu	39
5.5	Střední hodnoty testu MUSHRA pro jazz	39
5.6	Střední hodnoty testu MUSHRA pro akustickou hudbu	40
5.7	Střední hodnoty testu MUSHRA pro rockovou hudbu	40
5.8	Střední hodnoty testu MUSHRA pro popovou hudbu	41

SEZNAM TABULEK

3.1	Pětibodová škála MOS pro hodnocení poškození zvukových souborů .	26
3.2	Hodnotící škála testu MUSHRA	27
5.1	Hodnocení programu podle uživatelů	42

ÚVOD

Tato bakalářská práce se věnuje problematice vyhodnocování kvality rekonstruovaných audio signálů. Před vyhodnocovacími testy jsou však zmíněny některé typy poškození audio signálů. Nejznámější z nich jsou clipping, zašumění signálů a chybějící vzorky v audio signálech. Nicméně degradace audio signálů může být způsobena i jinými způsoby, ze kterých jsou zmíněny následující: clicking, crackling, hissing a rumbling.

Po typech poškození audio signálů jsou probírány rekonstrukční typy pro jejich opravy. První je declipping, který je chápán jako inverzní proces clippingu, dále je zmíněn inpainting, které se používá při výpočtu hodnot chybějících vzorků v audio signálech a následuje denoising, který se používá pro odstranění šumu ze zašumělých signálů. Declicking, decrackling, dehissing a derumbling jsou operace pro zmírnění nepříjemných vlivů při poslechu clickingu, cracklingu, hissingu a rumblingu.

Po rekonstrukčních typech audio signálu následuje stručný úvod pro objektivní testy a následně je vyjmenováno šest z nich. Tyto objektivní testy se používají kvůli jejich rychlosti a jednoduchosti. Některé z těchto metod nám ale neudávají relevantní informace o vnímané kvalitě audio signálů, právě proto je velká snaha o vytvoření testů, které do výsledku co nejlépe zahrnou vlastnosti lidského sluchu. Tím dosáhneme výsledky, které se blíží subjektivní kvalitě zvukových ukázek. Nejznámější objektivní testy jsou SNR, PEMO-Q a PEAQ.

Po popisu jednotlivých objektivních testů následuje stručný úvod do problematiky subjektivních testů, jejichž největší nevýhoda je časová náročnost. Nicméně když jsou testy prováděny správnými způsoby, podávají nám reálné informace o subjektivní kvalitě audio signálů. Ze subjektivních vyhodnocovacích testů je šest vyjmenováno a taktéž stručně vysvětleno; jsou to MOS, MUSHRA, ABX, Duo-trio test, Triangle test a Paired comparison test.

Další část bakalářské práce popisuje program pro subjektivní hodnocení kvality restaurace audio signálů, který je přizpůsoben testům MOS, MUSHRA a ABX. Je zde popsán princip fungování programu s obrázky grafického uživatelského rozhraní. Program byl napsán v programovacím jazyce Java, právě proto je kompatibilní s více platformami. Hlavní funkcí programu je provádění poslechových testů, ale v režimu správce lze také vytvářet nové testy. Výsledky testů se ukládají do CSV souborů, které lze snadno zpracovávat.

Na konci bakalářské práce jsou výsledky poslechových testů, které byly použity pro testování programu a také hodnocení programu podle uživatelů.

1 POŠKOZENÍ AUDIO SIGNÁLŮ

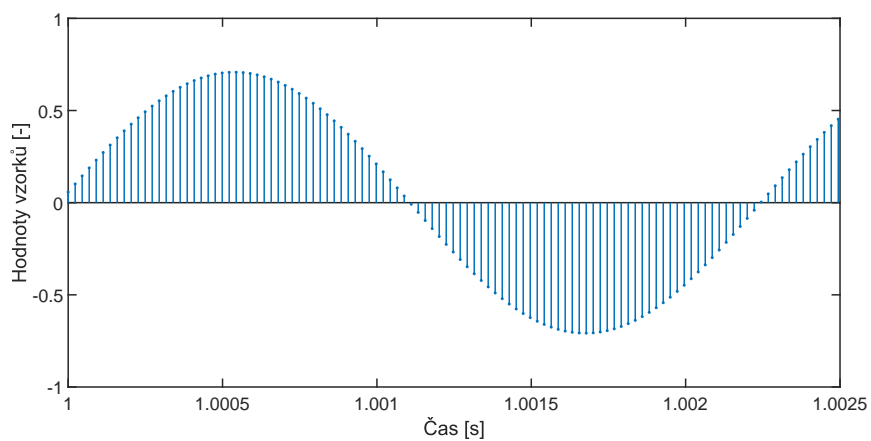
Digitální zvukové signály jsou při nahrávání, zpracovávání nebo přenosu velice náchylné na různé vlivy prostředí, které pak mohou způsobit nežádoucí jevy v signálech např. šum, zkreslení, praskání nebo úplný výpadek signálu (resp. výpadek vzorků signálu). Tyto problémy mohou nastat při přehrávání poškrábaného CD, při nekvalitním záznamu zvuku, při přenosu signálu mezi přijímači nebo v mnoha dalších případech. Rekonstrukcí poškozených signálů se zabývá několik studií. Existuje mnoho algoritmů a řešení odlišných metod, které dosahují různé výsledky v různých aplikacích [1].

1.1 Chybějící vzorky

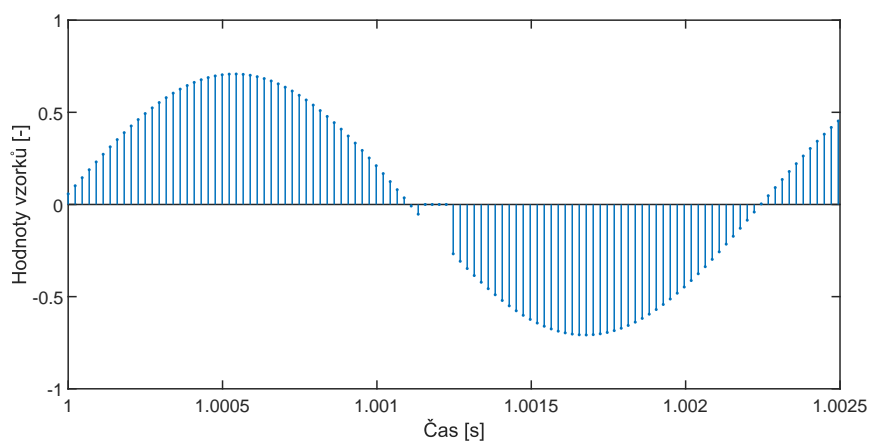
V digitálním světě se často stane, že vzorky se poškodí, jak již bylo zmíněno, např. při poškrábaném CD. Tyto vzorky (anglicky sample) se dají různými metodami znovu dopočítat. Některé typy takových rekonstrukcí audio signálů budou podrobněji rozebrány v kapitole 2. Teď ale nás zajímají hlavně příčiny ztráty vzorků ze signálů. U hardwarových nosičů to funguje celkem logicky: poškodí se nosič, signál se také poškodí, což může způsobit ztrátu celého signálu nebo částečnou ztrátu, tj. ztrátu některých vzorků v signálu.

Při současném nahrávání a digitálním zpracování zvuku s více (digitálními) zařízeními je třeba tato zařízení synchronizovat, aby nedošlo k poškození signálu při převodu z analogové platformy do digitální. Synchronizace těchto zařízení není složitá, můžeme jí docílit např. přes S/PDIF (Sony/Philips Digital Interface Format), ADAT (Alesis Digital Audio Tape) nebo pomocí Word Clock za použití koaxiálních kabelů. (Anglické názvy rozhraní byly převzaty z literatury [2].) Když tato zařízení nejsou synchronizována, tak nepracují se stejnými vzorky, nebo nepoužívají stejné vzorkovací kmitočty pro převod analogového signálu na digitální (ve výsledku se stane totéž, nepoužívají stejné vzorky signálu), to může vést i k neopravitelnému poškození nahrávky. V tomto případě některé vzorky nebudou přítomné v digitální podobě signálu, může se to stát jen s několika po sobě jdoucími vzorky, což se ještě dá dopočítat. Když se to ale stane s větším množstvím vzorků (stačí i několik tuctů), nahrávka bude zcela znehodnocena. Při poslechu takové nahrávky můžeme slyšet lupnutí (silnější nebo slabší) tam, kde vzniklo digitální ticho kvůli chybějícím vzorkům.

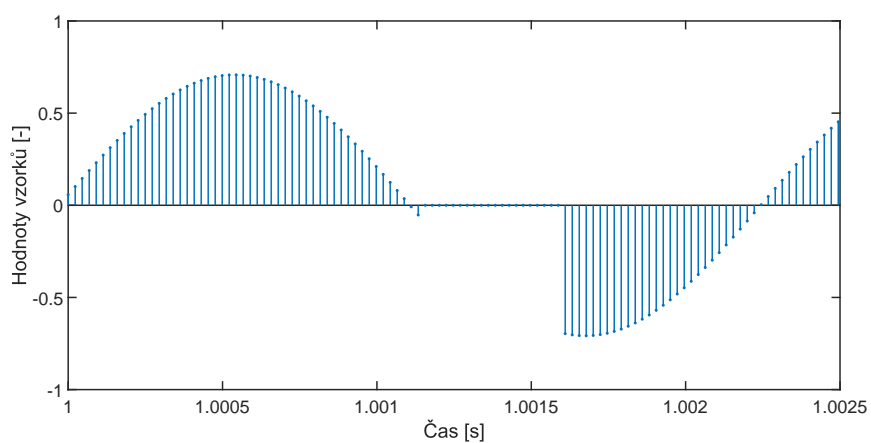
Na následujících obrázcích jsou znázorněny chybějící vzorky sinusového signálu o kmitočtu 440 Hz. Na přiloženém CD jsou i zvukové ukázky těchto průběhů, kde si můžeme vyslechnout, že čím víc vzorků chybí, tím průraznější je lupnutí při přehrávání.



Obr. 1.1: Nepoškozený sinusový signál o kmitočtu 440 Hz.



Obr. 1.2: Sinusový signál o kmitočtu 440 Hz se 4 chybějícími vzorky.



Obr. 1.3: Sinusový signál o kmitočtu 440 Hz s 20 chybějícími vzorky.

1.2 Clipping

„Clipping je forma nelineárního zkreslení signálu, která nastane, pokud zpracovávaný signál disponuje větším dynamickým rozsahem, než je systém schopný zpracovat. Výstupní signál systému pak bude přichycen na prahové úrovni systému a ztratíme tak část informace původního signálu“ [3].

Clipping může nastat při nahrávání, když audio signál překročí maximální hladinu, kterou převodník dokáže z analogového spojitého průběhu navzorkovat do digitální domény. Signál, který překročí daný práh, bude vždy vzorkován na maximální úrovni, kterou je zařízení schopno uložit. Proto průběh převedeného signálu v digitální podobě nebude stejný jako originální analogový průběh, tj. zdeformuje/zkreslí se. Tenhle případ může nastat i v případě převodu digitálního signálu na analogový, např. při přehrávání.

Clipping můžeme, ale nemusíme vždy slyšet. Když slyšíme zkreslení, nebo občasné lupnutí v záznamu, mohou to být znaky clippingu. Jsou známé dva typy clippingu. Hard clipping (část 1.2.1) a soft clipping (část 1.2.2).

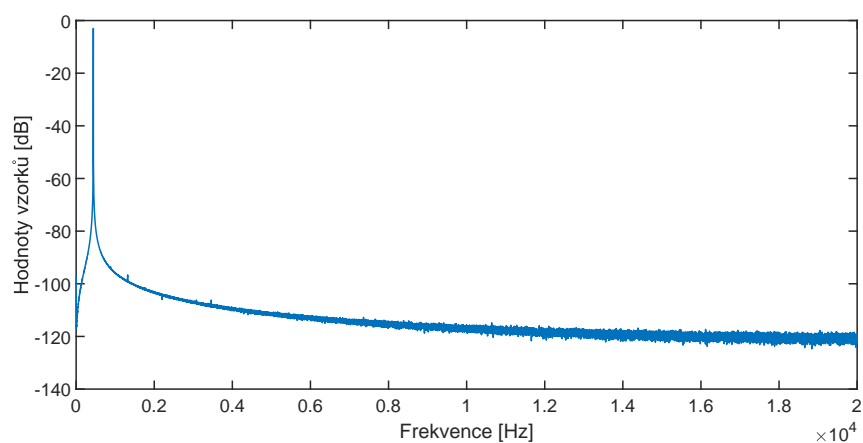
1.2.1 Hard clipping

Při takzvaném hard clippingu je signál $y_c[n]$, překračující danou úroveň θ_c , přísně omezen. Tuto operaci lze matematicky vyjádřit následovně [3]:

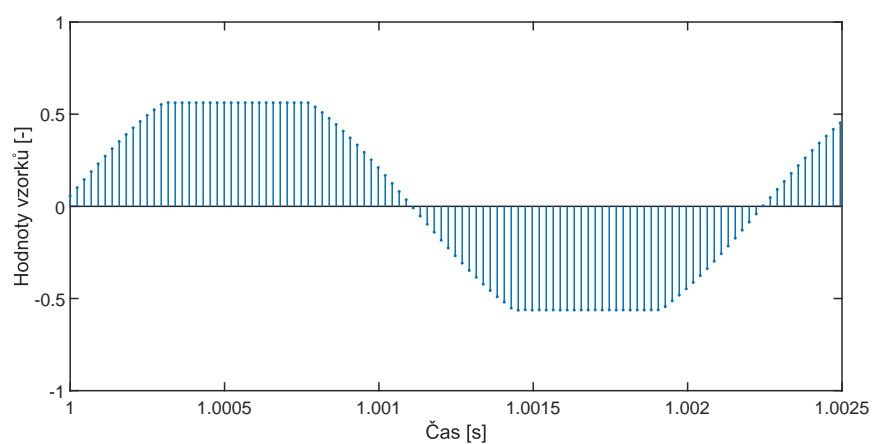
$$y_c[n] = \begin{cases} y[n] & \text{pro } |y[n]| < \theta_c, \\ \theta_c \cdot \text{sgn}(y[n]) & \text{pro } |y[n]| \geq \theta_c. \end{cases} \quad (1.1)$$

Mezi prahy $-\theta_c$ a θ_c je přenos ze vstupu na výstup lineární, tj. vstupní signál je přenášen na výstup systému beze změny. Když ale vstupní signál překročí prahovou úroveň, bude omezen na hodnoty $\pm\theta_c$ [3]. Způsobuje to ploché odříznutí maximálního průběhu. Při této operaci ve spektru signálu vzniká nekonečně mnoho vyšších harmonických frekvencí.

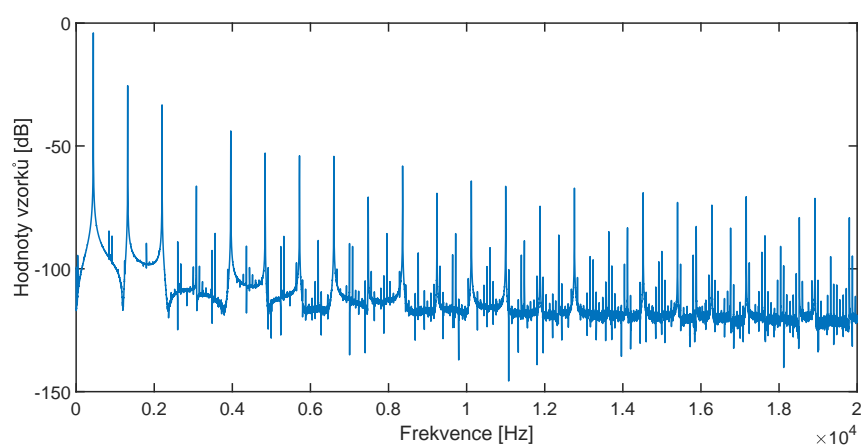
Pro znázornění hard clippingu (obr. 1.5 a obr. 1.6) použijeme stejný sinusový signál o kmitočtu 440 Hz, který byl použit při demonstraci chybějících vzorků na obr. 1.1.



Obr. 1.4: Jednostranné spektrum sinusového signálu o kmitočtu 440 Hz.



Obr. 1.5: Sinusový signál o kmitočtu 440 Hz po hard clippingu.



Obr. 1.6: Jednostranné spektrum sinusového signálu o kmitočtu 440 Hz po hard clippingu.

1.2.2 Soft clipping

Při takzvaném soft clippingu signál $y_c[n]$, který překročí danou úroveň θ_c , je sice omezen, ale ne tak přísně jako u hard clippingu. Hodnoty, které přesahují daný práh, jsou redukovány, ale sledují průběh originálního signálu. Při této operaci také vznikají vyšší harmonické frekvence ve spektru signálu.

Převodní charakteristika soft clippingu není pevně daná, obecně ale lze říci, že soft clipping se dá modelovat pomocí libovolné sigmoidy (např. $\arctan(x)$, $\tanh(x)$ nebo $\frac{x}{1+|x|}$). Jedna z možností implementace soft clippingu je následovná [3]:

$$y_c[n] = \begin{cases} y[n] & \text{pro } |y[n]| \leq r, \\ \text{sgn}(y[n]) \left[(\theta_c - r) \tanh\left(\frac{|y[n]| - r}{\theta_c - r}\right) + r \right] & \text{pro } |y[n]| > r, \end{cases} \quad (1.2)$$

kde θ_c je nastavitelný práh a r je práh, od kterého bude výstupní signál komprimován. Mezi úrovněmi $-r$ a r je přenosová charakteristika lineární, tj. výstupní a vstupní signál je totožný.

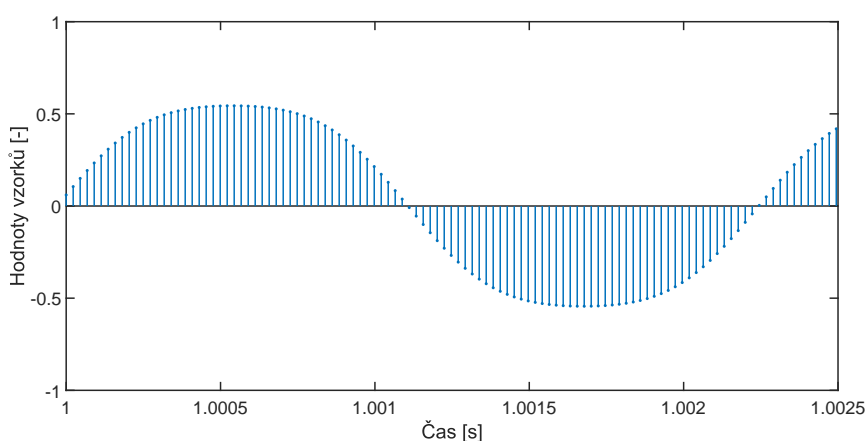
Proměnná k , tzv. „tvrdost“ charakteristiky nabývá hodnoty od 0 do 100. Pokud známe přesnou hodnotu k můžeme r vypočítat pomocí vztahu:

$$r = 0,01 \cdot k \cdot \theta_c \quad (1.3)$$

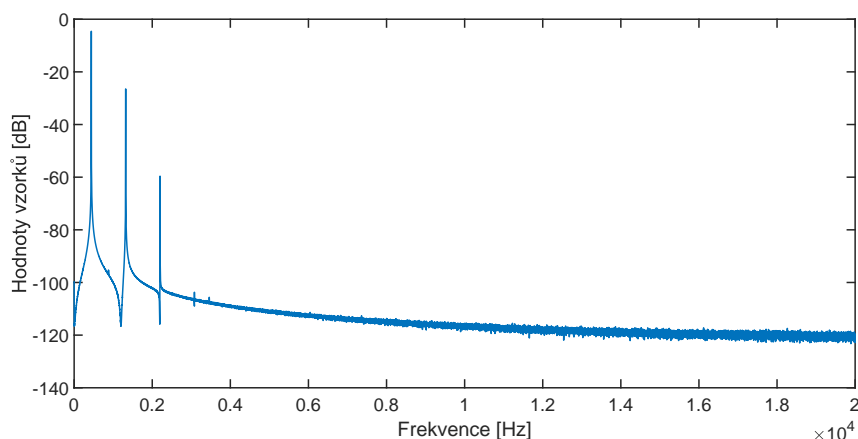
Když k nabývá hodnotu 0, převodní charakteristika bude co možná nejměkčí. V druhém mezním případě, když $k = 100$, pak $r = \theta_c$, což znamená, že převodní charakteristika je stejná jako u hard clippingu [3].

Pro znázornění soft clippingu (obr. 1.7 a obr. 1.8) použijeme sinusový signál o kmitočtu 440 Hz (obr. 1.1).

Na příloženém CD jsou i zvukové ukázky těchto průběhů.



Obr. 1.7: Sinusový signál o kmitočtu 440 Hz po soft clippingu.



Obr. 1.8: Jednostranné spektrum sinusového signálu o kmitočtu 440 Hz po soft clippingu.

1.3 Zašumění

Analogové audio signály jsou často ovlivněny šumem. Šumy/hluky jsou nejlépe popsány jako nežádoucí poruchy překrývající užitečný signál. Hladina šumu je obvykle vyjádřeno v dB vzhledem k referenční hodnotě a je označován jako SNR (Signal-to-noise ratio – odstup signálu od šumu) [4].

1.4 Clicking

Anglické slovo *clicking* je používáno v audio světě k popsání slyšitelného cvaknutí nebo lupnutí při poslechu audio signálu. Lupnutí však v tomto případě není způsobeno chybějícími vzorky v signálu, ale krátkými vyčnívajícími náhodnými maximy (obr. 2.1) [5].

1.5 Crackling

Anglické slovo *crackling* česky znamená praskání, šustění, což je také nežádoucí jev v audio signálech. Praskání a šustění může být způsobeno několika faktory, typickou ukázkou je například praskání gramofonových desek. „Crackles“ (čili praskoty) jsou tišší clicky, které se dějí pravidelněji a předvídatelněji [5].

1.6 Hissing

Hissing je odvozen z anglického slova *hiss*, což česky znamená syčení. Říká se tomu i „high end hiss“, kde „high end“ poukazuje na vysoké frekvence, které ještě jsou přítomné v rozsahu slyšitelnosti lidského sluchu. Syčení je tedy nepříjemný přebytek energií vysokých frekvencí ve spektru audio signálů.

1.7 Rumbling

Anglické slovo *rumble* česky znamená dunění, hučení. V oblasti zvuku se tomu říká i „low end rumble“, kde „low end“ označuje nízké frekvence, které ještě jsou přítomné v rozsahu slyšitelnosti lidského sluchu. Dunění, hučení je tedy přebytek energií nízkých frekvencí ve spektru audio signálů.

2 REKONSTRUKCE AUDIO SIGNÁLŮ

Když už audio signál byl poškozen, buď jedním ze způsobů zmíněných v kapitole 1, nebo z nějakého jiného důvodu, zbývá nám pouze možnost rekonstrukce. Avšak v některých případech ani rekonstrukce nepřipadá v úvahu, audio signál může být tak poškozen, že už není způsob, jak ho zachránit.

Rekonstrukce signálu je proces, při němž signál je vhodnou technikou opraven, nebo je docílena co největší podobnost rekonstruovaného signálu s původním nepoškozeným signálem. Dosáhnutí nejbližší podobnosti ale nemusí být vždy nejdůležitějším cílem tohoto procesu. Důležitější je, aby výsledek při poslechu nepůsobil rušivým dojmem v místech rekonstrukce, a aby tato místa nebyla pro posluchače rozeznatelná [6].

Tato kapitola bude zaměřena na některé typy rekonstrukcí signálů, zejména na declipping (podkapitola 2.1), inpainting (podkapitola 2.2), denoising (podkapitola 2.3), declicking (podkapitola 2.4), decrackling (podkapitola 2.5), dehissing (podkapitola 2.6) a na derumbling (podkapitola 2.7).

2.1 Declipping

Declipping je typ rekonstrukce, při které se snažíme ze saturovaného (tedy poškozeného) signálu získat zpět originální signál nebo alespoň signál co nejpodobnější původnímu, ze kterého jsou co nejlépe odstraněny rušivé jevy saturace (zkreslení). Declipping je chápán jako inverzní proces clippingu.

Declipping se dá aplikovat na oba typy clippingu, hard clipping (část 1.2.1) i soft clipping (část 1.2.2), ale většinou se používá pro rekonstrukci signálů po zkreslení hard clippingem. Jedním důvodem, proč je k tomu declipping v praxi využíván, je, že vliv hard clippingu na signál je mnohem nepříjemnější při poslechu než v případě soft clippingu. Kromě toho když známe převodní charakteristiku u soft clippingu, pak pomocí inverze této charakteristiky je možné původní signál rekonstruovat [3].

V literatuře [3] je přehled mnoha metod pro declipování signálů, které využívají např. princip interpolace, lineární predikci, bayesovskou statistiku, Monte Carlo filtrování, řídkou reprezentaci signálů nebo rekursivní projekci vektorů.

2.2 Inpainting

Audio inpainting je definován jako proces výpočtu, který je založen na konceptu image inpainting (používán při rekonstrukci obrázků – při doplňování chybějících pixelů). Je to všeobecně problematická situace, která se vyskytuje v mnoha případech.

Danou část audio signálu tvoří „spolehlivá“ data zvuku a druhou část „nespolehlivá“ data, která buď ze signálu zcela chybějí nebo jsou vysoce degradovaná (i v tomto případě jsou považována za chybějící). Ze „spolehlivých“ dat jsou následně odhadnuta „nespolehlivá“ data tohoto audio signálu. V některých případech můžeme mít i dodatečné informace o chybějících datech.

Ve všeobecném zpracovávání zvukových signálů mohou data, zmíněná v předchozím odstavci, být například vzorky z průběhu audio signálu. Máme tedy dvě skupiny vzorků. V první skupině jsou vzorky známých hodnot a v druhé skupině jsou vzorky, u kterých hodnoty neznáme (buď chybějí nebo jsou degradovány). Problematika tohoto typu rekonstrukce je popsána jako obnovení hodnot vzorků v druhé skupině na základě znalosti hodnot vzorků v první skupině.

Audio inpainting využívá např. autoregresní model (AR), bayesovský odhad, sinusové modelování nebo umělé neuronové sítě pro výpočet chybějících dat [7].

2.3 Denoising

Takzvaný audio noise reduction je proces, při kterém šum/hluk z audio signálu je zeslaben, potlačen nebo zcela utlumen. Existují různé aplikace (např. u mluveného slova nebo u rekonstrukce hudby), u nichž je technika audio denoisingu nezbytně důležitá. Je to zejména důležité v aplikacích, pomocí nichž můžeme telefonovat přes internet. V tomto případě je audio signál převeden z digitální podoby do analogové a teprve poté zesílen. Během amplifikace je tento signál velice náchylný na šumy/ruchy prostředí. Většina nahrávacích zařízení (buď digitální nebo analogové) může být také citlivá na šumy/hluky prostředí [8].

Cílem audio denoisingu je zeslabení šumu/hluku v audio signálu a zároveň obnovení originálního signálu. To ale neznamená, že signál po denoisingu musí být zcela čistý (tzn. bez jakéhokoli šumu nebo ruchu). Byla navržena řada metod pro odstranění šumu z audio signálů, ale v mnoha případech metody nedosahují příliš vysoké účinnosti [9].

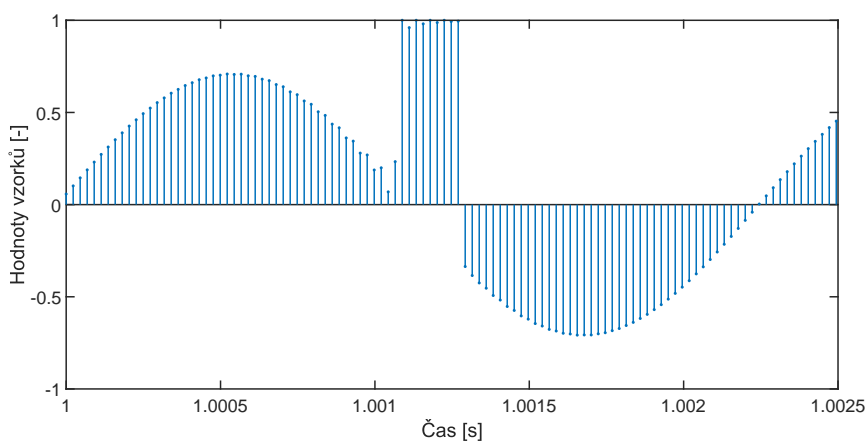
Audio denoising metody využívají např. řídkou reprezentaci signálů, Gaussův filtr, anisotropickou difúzi, vlnkovou transformaci, Fisherovu informaci nebo umělé neuronové sítě pro rekonstrukci audio signálů.

2.4 Declicking

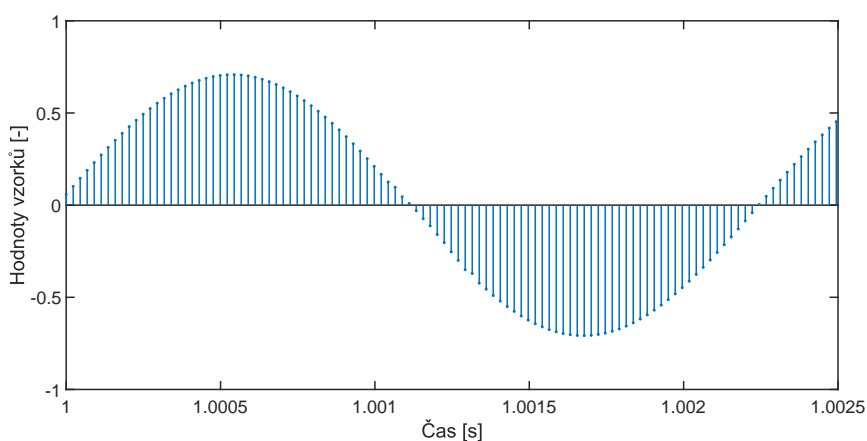
Declicking je proces, přičemž krátká vyčnívající náhodná maxima (clicky) jsou zredukována na patřičnou úroveň a předchozí průběh signálu je následně překreslen na nový průběh (již bez náhodných maxim).

Declicking je možné provést manuálně nebo automatizovaně. Manuální declicking se provádí v programech, kde průběh signálu je znázorněn a dá se manuálně překreslit. Proces je velice jednoduchý: najdeme místa clickingu a manuálně překreslíme maxima, aby nevyčnívala z průběhu. Můžeme také použít pluginy/programy, které průběh v daném místě clickingu překreslí. Při automatizovaném declickingu se používají pluginy/programy, které lupnutí a cvaknutí odstraňují v reálném čase při přehrávání signálu. Výhodou automatizovaného declickingu je, že uživatel nemusí hledat místa náhodných maxim v průběhu signálu [5].

Pro znázornění clickingu (obr. 2.1) a manuálního declickingu (obr. 2.2) byl použit sinusový signál o kmitočtu 440 Hz (obr. 1.1). Na přiloženém CD jsou i zvukové ukázky těchto průběhů. Na obr. 2.2 je vidět, že manuální declicking není zcela přesný, ale při poslechu zvukových ukázek tyto malé nepřesnosti jsou vlastně nerozeznatelné.



Obr. 2.1: Sinusový signál o kmitočtu 440 Hz s clickingem.



Obr. 2.2: Sinusový signál o kmitočtu 440 Hz po manuálním declickingu.

2.5 Decrackling

Při operaci decrackling je praskání, šustění co nejlépe potlačen, v některých případech úplně odstraněn. Crackling není jednoznačně viditelný z průběhu audio signálu, právě proto manuální opravu není možné provést. Při decracklingu se používají programy/pluginy, které praskání a šustění detekují a následně potlačí v uživatelem zadané míře. Tyto programy fungují taktéž v reálném čase, tj. při přehrávání již slyšíme (buď částečně nebo úplně) rekonstruovaný audio signál [5].

2.6 Dehissing

Při procesu dehissing je přebytek energií vysokých frekvencí ve spektru audio signálu různými nástroji potlačen na akceptovatelnou úroveň. Tzv. de-esser plugin je nástroj, který provádí dynamické úpravy vysokých frekvencí ve spektru audio signálu. Sice tyto pluginy nebyly vyvinuty na úlohu dehissingu, ale z principu fungování jsou vhodnými nástroji k provádění této operace.

2.7 Derumbling

Derumbling je podobný proces jako dehissing, rozdíl je v tom, že dehissing je proces pro potlačení přebytku energií vysokých frekvencí ve spektru audio signálu, naopak derumbling je proces, přičemž je potlačen přebytek energií nízkých frekvencí ve spektru audio signálu. Derumbling je prováděn pomocí pluginů.

3 VYHODNOCENÍ KVALITY

O kvalitě vybraného typu rekonstrukce se můžeme více dozvědět na základě objektivních testů (podkapitola 3.1) nebo pomocí subjektivních metod vyhodnocování (podkapitola 3.2). Tato kapitola bude zaměřena na tyto testy a metody vyhodnocování kvality rekonstrukce audio signálů. Z objektivních testů budou zmíněny: SNR (část 3.1.1), PEMO-Q (část 3.1.2), PEAQ (část 3.1.3), PESQ (část 3.1.4), POLQA (část 3.1.5) a ViSOQL (část 3.1.6). Ze subjektivních metod budou probírány: MOS (část 3.2.1), MUSHRA (část 3.2.2), ABX (část 3.2.3), Duo-trio test (část 3.2.4), Triangle test (část 3.2.5) a Paired comparison test (část 3.2.6).

3.1 Objektivní testy

Objektivní testy vyhodnocující kvalitu audio signálů jsou především používány kvůli jejich rychlosti a jednoduchosti. Výsledky dostaneme okamžitě po testování. Jsou nazývány objektivními kvůli faktu, že výsledky testů nejsou ovlivněny vlivy prostředí ani jedinců. Testované signály budou vyhodnocovány vždy stejně bez ohledu na místo a čas testování. Nicméně některé z těchto metod nám neudávají relevantní informace o vnímané, subjektivní kvalitě zvukových nahrávek. Právě proto byly také navrženy testy, které se do testování snaží zahrnout vlastnosti lidského sluchu, tím pádem výsledky objektivních testů se přibližují skutečné vnímané kvalitě.

3.1.1 SNR

Signal-to-noise ratio (dale jen SNR) je nejpoužívanější objektivní test pro vyhodnocování kvality rekonstruovaných audio signálů, nicméně korektnost této metody je oprávněně zpochybňována [10]. Signal-to-noise ratio – odstup signálu od šumu – obecně znamená bezrozměrný poměr výkonu signálu k výkonu šumu obsaženého v záznamu [11].

V případě hodnocení kvality rekonstrukce, SNR znamená něco mírně odlišného: „Odstup rekonstruovaného signálu od originálního signálu“. Výsledkem je číslo v jednotkách decibelů. Čím větší toto číslo je, tím podobnější signály jsou. Pokud jsou signály totožné, výsledek se bude blížit nekonečnu [6].

Matematický vztah pro výpočet hodnoty SNR je následující [6]:

$$\text{SNR}(\mathbf{y}, \hat{\mathbf{y}}) = 10 \log \frac{\|\mathbf{y}(\mathbf{I})\|_2^2}{\|\mathbf{y}(\mathbf{I}) - \hat{\mathbf{y}}(\mathbf{I})\|_2^2}, \quad (3.1)$$

kde $\mathbf{y}(\mathbf{I})$ je vektor původního signálu a $\hat{\mathbf{y}}(\mathbf{I})$ je vektor rekonstruovaného signálu.

Největší problém metody SNR je, že výsledek nám udává informace jenom o podobnosti dvou signálů. Signály neprochází žádnými filtry a nevyužívá se žádný akustický model, proto nelze jednoznačně tvrdit, že čím větší je hodnota SNR, tím příjemnější bude rekonstruovaný signál na poslech [6].

3.1.2 PEMO-Q

Sofistikovanější objektivní metodou ke srovnání a vyhodnocení kvality je PEMO-Q, která se snaží do výsledku zahrnout psychoakustický model. Vychází se z původního a poškozeného signálu, které jsou na základě lidského sluchu dále zpracovávány. Po průchodu sadou filtrů jsou tyto signály křížově korelovány a v některých případech i váhovány a kvantovány. Proces zpracování signálů této metody je poměrně složitý.

Výstupem jsou dvě hodnoty: první Perceptual Similarity Measure – míra vjemové podobnosti (PSM) a druhá Temporary Perceptual Similarity Measure – míra dočasné vjemové podobnosti (PSMt). U PSMt jsou data signálu zpracovávána v 10 milisekundových segmentech a v poslední části zpracování jsou váhována a kvantována. Díky tomu PSMt je méně ovlivněn typem a charakterem signálu než PSM. Obě tato čísla vyjadřují kvalitu rekonstrukce a nabývají hodnot od -1 a 1 . Pokud se hodnoty blíží -1 , tak mluvíme o špatné kvalitě rekonstrukce a naopak když se tyto hodnoty blíží 1 , pak kvalita rekonstrukce je velice dobrá [6].

Z literatury [12] byl převzat matematický vztah pro přepočtení hodnoty PSMt na hodnoty škály MOS ODG (podrobněji probírána v části 3.2.1), která nám prozradí víc o vnímané kvalitě rekonstrukci audio signálu než hodnoty PSM a PSMt. Kromě toho pak tuto hodnotu je možné použít při porovnávání testu PEMO-Q se subjektivními testy vyhodnocující kvalitu audio signálů.

$$f(x) = \begin{cases} \max \left\{ -4, \frac{a}{x-b} + c \right\} & \text{pro } x < x_0, \\ d \cdot x - d & \text{pro } x \geq x_0, \end{cases} \quad (3.2)$$

kde x je hodnota PSMt, a , b , c , d a x_0 jsou předem dané konstanty o hodnotách $a = -0,22$, $b = 0,98$, $c = -4,13$, $d = 16,4$ a $x_0 = 0,864$. Tímto způsobem vypočítané hodnoty funkce $f(x)$ jsou pak hodnoty škály ODG.

3.1.3 PEAQ

PEAQ (Perceptual Evaluation of Audio Quality – vjemové hodnocení kvality zvuku), též známé jako ITU-R BS.1387, je normou doporučené objektivní měření vnímané kvality zvuku zavedené v roce 1998 agenturou ITU (International Telecommunication Union – Mezinárodní telekomunikační unie). Využívá se v něm software pro simulace vjemových vlastností lidského ucha, následně jsou výsledky integrovány, čím

dostaneme subjektivní kvalitu testovaného zvuku. PEAQ je považována za jednu z nejefektivnějších objektivních metod pro vyhodnocování vnímané zvukové kvality. Velké množství softwaru měřící objektivní kvalitu zvuku bylo navrženo v souladu s touto normou. Mohou být využity dva typy PEAQ-u: základní verze a pokročilá verze. Jádrem této metody je psychoakustický model lidského sluchu [13].

3.1.4 PESQ

PESQ (Perceptual Evaluation of Speech Quality – vjemové hodnocení kvality řeči) je normou ITU-T P.862 doporučené měření objektivní kvality řeči pro úzkopásmová použití zavedené v roce 2001 agenturou ITU. Tyto úzkopásmové systémy mohou původní signál degradovat filtrováním, zpožděním, zkreslením, nebo použitím kodeků s nízkou bytovou rychlostí. PESQ řeší tyto nežádoucí efekty ekvalizací, časovým vyrovnaním a s algoritmem pro průměrování zkreslení. Test porovnává původní signál $X(t)$ s degradovaným signálem $Y(t)$, který je výsledkem průchodu signálu $X(t)$ přes komunikační systém. Výstupem testu je předpověď vnímané kvality signálu $Y(t)$.

Je třeba také poznamenat, že algoritmus PESQ neposkytuje komplexní hodnocení kvality přenosu. Měří pouze účinky jednosměrného zkreslení řeči a šumu. Ztráta hlasitosti, zpoždění, ozvěny a další poruchy vyskytující se v obousměrné komunikaci ve výsledku testu PESQ nejsou zahrnuty. Proto je možné mít vysoké hodnocení v testu PESQ, ale špatnou vnímanou kvalitu signálu [14].

Výsledky testu PESQ nabývají hodnoty mezi -0,5 a 4,5. Pro přepočítání těchto hodnot na škálu MOS SDG (podrobněji probírána v části 3.2.1) můžeme použít vztah z literatury [15]:

$$y = 0,999 + \frac{4,999 - 0,999}{1 + e^{-1,4945 \cdot x + 4,6607}}, \quad (3.3)$$

kde x je výsledná hodnota testu PESQ a y je hodnota na škále MOS.

Přepočítání hodnot škály MOS SDG na výslednou hodnotu testu PESQ je také možný a to podle vzorce [15]:

$$x = \frac{4,6607 - \ln\left(\frac{4,999 - y}{y - 0,999}\right)}{1,4945}, \quad (3.4)$$

kde y je známá hodnota hodnotící škály MOS a x je vypočtená hodnota testu PESQ.

3.1.5 POLQA

POLQA (Perceptual Objective Listening Quality Assessment – objektivní hodnocení vnímané kvality poslechu) je normou ITU-T P.863 doporučené měření objektivní

kvality řeči současných telefonních systémů (i telefonních systémů blízké budoucnosti), které využívají širokou škálu kódovacích a transportních technologií, zavedené v roce 2014 agenturou ITU. Test POLQA byl navržen pro současné průmyslové požadavky, právě proto lze test využít i pro širokopásmové systémy a pro sítě a kodeky s časovou deformací. Algoritmus testu POLQA je vylepšeným algoritmem testu PESQ. Z tohoto důvodu test POLQA také porovnává původní signál $X(t)$ s degradovaným signálem $Y(t)$, který je výsledkem průchodu signálu $X(t)$ přes komunikační systém. Výstupem testu je předpověď vnímané kvality signálu $Y(t)$, které je již přizpůsoben škále MOS.

Je třeba také poznamenat, že obdobně jako test PESQ, ani algoritmus POLQA neposkytuje komplexní hodnocení kvality přenosu. Měří pouze účinky jednosměrného zkreslení řeči a šumu. Poruchy vyskytující se v obousměrné komunikaci ve výsledku nejsou zahrnuty, proto je možné mít vysoké hodnocení v testu POLQA, ale špatnou vnímanou kvalitu signálu [16].

3.1.6 ViSQOL

ViSQOL (Virtual Speech Quality Objective Listener – objektivní posluchač virtuální řečové kvality) je psychoakustický model citlivosti lidského sluchu na degradaci v kvalitě řeči, který byl vyvinut výzkumnou skupinou Sigmedia (Signal Processing and Media Applications). Test porovnává referenční signál s degradovaným signálem. Výstup testu je hodnota kvality řeči vnímaná průměrnou osobou na škále MOS. Model má pět hlavních fází zpracování: předběžné zpracování, časové vyrovnaní, předpovídání deformace, porovnání podobnosti a postprocesní přepoččet podobnosti signálů na škálu MOS. Test ViSQOL implementovaný v MATLABu lze stáhnout z webové stránky skupiny Sigmedia [17].

Algoritmus testu ViSOQLAudio také od výzkumné skupiny Sigmedia vychází z algoritmu testu ViSQOL, ale byl rozvinut a přizpůsoben k vyhodnocování kvality hudby a zvukových nahrávek. Výstupem testu již není hodnota škály MOS, ale hodnoty mezi 0 a 1, vyjadřující podobnost dvou signálů (0 – odlišné, 1 – totožné). Test ViSOQLAudio implementovaný v MATLABu je k dispozici ke stažení na webové stránce autorů [18].

3.2 Subjektivní metody

Subjektivní metody vyhodnocování kvality zvukových nahrávek jsou pojmenovány subjektivními z faktu, že bez subjektů, které se zúčastní na testováních, by nebylo možné tyto testy realizovat. Nevýhodou je časová náročnost: testy musí každý účastník dokončit, teprve pak je možné začít zpracovávat nasbíraná data a vypočítat

výsledek testování. Další nevýhodou je, že místo, čas testování a fyzický či psychický stav jedinců podílejících se v testu mohou mít nežádoucí vlivy na výsledek. Avšak výsledkem těchto testů je skutečná subjektivní kvalita dané zvukové ukázky za předpokladu, že testy byly provedeny správnými způsoby (tzn. dostatečný počet testovaných subjektů, stejné podmínky pro všechny subjekty při testování, správně navržený test).

3.2.1 MOS

ITU (International Telecommunication Union – Mezinárodní telekomunikační unie) definuje bodové hodnocení jako „hodnoty v předem definované stupnici, které subjekt vybírá podle jeho názoru na výkonnost systému“. MOS (Mean Opinion Score – průměrná známka hodnocení) je průměr těchto subjektů vybraných hodnocení. Stal se nejpobulárnějším deskriptorem vnímané kvality médií. V oblasti kvality řeči dosáhl velký úspěch a následně byl použit i pro jiné oblasti jako například audio, obraz, video a audiovizuální obsah v řadě aplikací, od laboratorních testů až po monitorování v provozu.

MOS je nyní vlastně nejvíce používaná forma pro kvantifikaci vnímané kvality médií. Pětibodová stupnice MOS (5 – vynikající, 4 – dobrý, 3 – přijatelný, 2 – špatný, 1 – nepřijatelný) je mimořádně populární. Tato forma hodnocení kvality byla přínosem v rámci zvýšení povědomí o důležitosti vyhodnocování vnímané kvality. Jasnou výhodou měřítka je, že tyto hodnoty jsou známé a používané široce ve světě, nicméně nepoužívá se vždy ze správných důvodů nebo ve správném kontextu [19].

Při vyhodnocování kvality rekonstruovaných nahrávek tato pětibodová škála má dva druhy: SDG (Subjective Difference Grade – subjektivní rozdílový stupeň) a ODG (Objective Difference Grade – objektivní rozdílový stupeň). Tyto stupnice jsou vlastně totožné, rozdíl je pouze v hodnotách škály, což můžeme vidět v tab. 3.1 (Česká pojmenování jednotlivých stupňů byla převzata z literatury [20]).

Tab. 3.1: Pětibodová škála MOS pro hodnocení poškození zvukových souborů

Poškození (<i>Impairment</i>)	SDG	ODG
Neznatelné (<i>Imperceptible</i>)	5,0	0,0
Znatelné, ale ne nepříjemné (<i>Perceptible, but not annoying</i>)	4,0	−1,0
Trochu nepříjemné (<i>Slightly annoying</i>)	3,0	−2,0
Nepříjemné (<i>Annoying</i>)	2,0	−3,0
Velmi nepříjemné (<i>Very annoying</i>)	1,0	−4,0

3.2.2 MUSHRA

Test MUSHRA (Multi Stimulus test with Hidden Reference and Anchor – Vícestimulový test se skrytou referencí a kotvou) je standardizovaná metoda pro subjektivní testování kvality zvuku. Při testování je posluchači předloženo několik zvukových stop včetně známé reference. Mezi stopami je i skrytá reference a tzv. kotva, což je nahrávka s nejvíce zhoršenou kvalitou.

Jedna zkouška testu MUSHRA může obsahovat maximálně patnáct zvukových stop, ze kterých jedna je známá reference, druhá je náhodně skrytá reference, třetí je náhodně skrytá kotva a ostatní nahrávky jsou testované systémy. Kotva se používá pro lepší vymezení hodnotící škály (mezi kotvou a referencí). Nejčastěji používané kotvy jsou dolním propustem (např. s mezním kmitočtem 3,5 kHz) filtrované referenční signály. Z tohoto faktu vyplývá, že kotva má nejhorší kvalitu testovaných stop, a právě proto by ji měl posluchač hodnotit nejhůř ze všech ukázek. (Při testování rekonstrukčních metod audio signálů jako kotva se používá přímo poškozený signál.) Testované stopy nesmí přesáhnout délku dvaceti sekund [20].

Tab. 3.2: Hodnotící škála testu MUSHRA

Kvalita (<i>Quality</i>)	Hodnocení
Vynikající (<i>Excellent</i>)	100 – 80
Dobrá (<i>Good</i>)	80 – 60
Příjemná (<i>Fair</i>)	60 – 40
Špatná (<i>Poor</i>)	40 – 20
Nepříjemná (<i>Bad</i>)	20 – 0

3.2.3 ABX

Test ABX je rozlišovací test, který může svědčit buď o shodě vnímané kvality dvou ukázek nebo naopak o odlišnosti. V testech ABX jsou vždy tři ukázky. První dvě jsou známé referenční vzorky a jsou označovány písmeny A a B. Jedna z těchto vzorků je původní nahrávka a druhá je modifikovaná (v našem případě rekonstruovaná). Třetí ukázka (X) je jedna ze vzorků A a B, ale bez označení. Subjekt testu má za povinnost vybrat jednu z prvních dvou ukázek, který je podle něho totožná s třetí ukázkou [21].

Při testování ABX můžeme dostat jen 2 výsledky (buď správný nebo špatný výběr ze vzorků A a B), právě proto nám může pomoci při rozhodování, jestli daná rekonstrukční metoda je vhodná k rekonstrukci audio signálů, či nikoliv.

3.2.4 Duo-trio test

Duo-trio test (česky „Zkouška duo-trio“) je daná mezinárodní normou ISO 10399 zavedené organizací ISO (International Organization for Standardization – Mezinárodní organizace pro normalizaci). Tato norma popisuje postup jak určit, zda mezi dvěma vzorky existuje vnímatelný rozdíl nebo podobnost. Metoda využívá postup nucené volby. Test lze použít když mezi vzorky existuje jen jeden rozdíl v sensorických vlastnostech, ale i v případě více rozdílů. Test je pro posuzovatele snadnější provést než Triangle test (část 3.2.5), ale je statisticky méně účinný.

Subjekty testu obdrží tři vzorky, z nichž jeden vzorek je označen jako referenční a další dva vzorky mají různá označení. Subjekty jsou informovány, že jeden z těchto dvou vzorků je totožný s referenčním vzorkem a druhý nikoliv. Posuzovatele mají předem dané, jestli mají vybírat nahrávku, která je podle jejich posudku totožná s referenční nahrávkou nebo naopak, mají vybírat nahrávku, která se podle nich liší od referenční nahrávky. Kterou nahrávku vybírat záleží na cíli testu. Cílem testu může být buď hledání rozdílu nebo hledání podobnosti. Počet správných odpovědí je počítán a výsledek testu je určen podle statistické tabulky [22].

3.2.5 Triangle test

Triangle test (česky „Trojúhelníková zkouška“) je daná mezinárodní normou ISO 4120 zavedené organizací ISO. Tato norma popisuje postup jak určit, zda mezi dvěma vzorky existuje vnímatelný rozdíl nebo podobnost, stejně jako u testu Duo-trio. Tato metoda taktéž využívá postup nucené volby. Test lze použít když mezi vzorky existuje jen jeden rozdíl v sensorických vlastnostech, ale i v případě více rozdílů. Test má limitovanější použití než Duo-trio test (část 3.2.4), ale je statisticky účinnější.

Subjekty testu obdrží tři vzorky a jsou informováni o tom, že dva ze vzorků jsou shodné a třetí je odlišný. Posuzovatele musí vybrat z těchto tří vzorků jeden, který se podle nich liší od zbývajících dvou, i tehdy, když jejich výběr je pouze odhadem. Počet správných odpovědí je počítán a výsledek testu je určen podle statistické tabulky [23].

3.2.6 Paired comparison test

Paired comparison test (česky „Párová porovnávací zkouška“) je daná mezinárodní normou ISO 5495 zavedené organizací ISO. Tato norma popisuje postup jak určit, zda existuje znatelný rozdíl nebo podobnost mezi dvěma vzorky týkající se jejich kvality. Jedná se vlastně o porovnávací test s metodou nucené volby mezi dvěma alternativami. Test je nejjednodušší existující klasifikační test, neboť se v něm používají pouze dva vzorky. Metoda je použitelná pro určení, jestli existuje nějaký rozdíl

mezi těmito vzorky v dané vlastnosti, ale neposkytuje žádný údaj o rozsahu tohoto rozdílu. Absence rozdílu ve výsledku testu nemusí nutně znamenat, že mezi vzorky není žádný rozdíl. Tato metoda je použitelná pouze tehdy, jsou-li testované ukázky relativně homogenní.

Subjekty testu obdrží soubor dvou vzorků (tj. dvojici vzorků). Posuzovatele vybírají vzorek, který považují za kvalitnější, i tehdy, když jejich výběr je pouze odhadem. Počítá se kolikrát jsou dané vzorky vybrané a výsledek testu je následně určen podle statistické tabulky. Test lze také použít k porovnání dvou vzorků z hlediska preference subjektů [24].

4 POPIS PROGRAMU

V této kapitole je popsán princip fungování programu pro vyhodnocování kvality restaurace audio signálů, který byl vytvořen v programovacím jazyce JavaFX (verze Java SE 8) ve vývojovém prostředí Eclipse Neon (verze 4.6). Při popisu programu jsou jednotlivé části GUI (graphical user interface – grafické uživatelské rozhraní) znázorněny i na obrázcích.

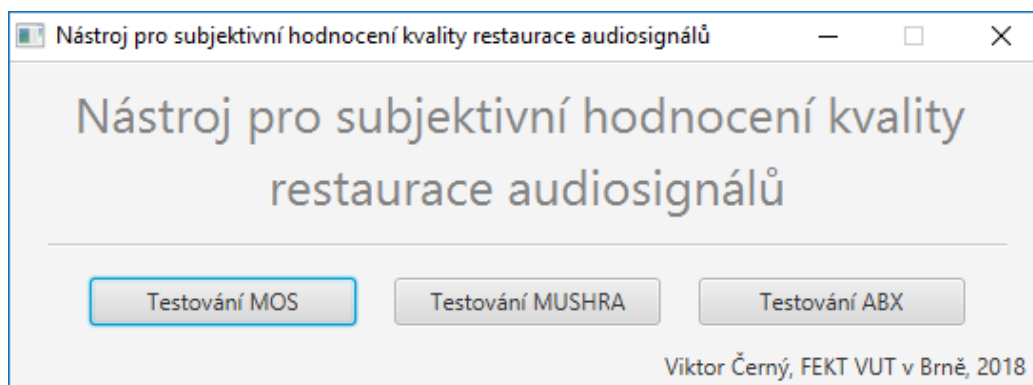
4.1 JavaFX

JavaFX je sada grafických a média balíčků, které umožňují vývojářům navrhovat, vytvářet, testovat, ladit a nasazovat aplikace, které fungují na různých platformách.

JavaFX je vlastně knihovna programovacího jazyku Java, skládající se ze tříd a rozhraní, které jsou napsány v Java kódu. Právě proto jsou při používání JavaFX dostupné a lze použít i ostatní knihovny programovacího jazyku Java. Používá se hlavně pro implementování aplikací, které využívají grafické uživatelské rozhraní, audio, video nebo animace a musí být spustitelné na více platformách [25].

4.2 Hlavní nabídka

Po spuštění programu se načte databáze testů, která je uložena ve složce „data“, v souboru „database.dat“. Program nás varuje, pokud soubor není nalezen nebo pokud databázi nelze načíst. Po úspěšném načtení databáze se dostaneme na hlavní nabídku programu, která je velice jednoduchá (obr. 4.1). V programu jsou implementovány tři testovací metody, podle kterých jsou pojmenována tlačítka: „Testování MOS“, „Testování MUSHRA“ a „Testování ABX“. Kliknutím na dané tlačítko si vybereme testovací metodu a dostaneme se na seznam testů vybrané metody.



Obr. 4.1: Hlavní nabídka programu.

4.3 Seznam testů

Okna pro výběr testů jsou stejná pro všechny testovací metody (např. seznam testů MOS – obr. 4.2). Testy jsou shrnuty do tabulky a je možné je seřadit podle ID, názvu, počtu stran nebo počtu vzorek. S dvojitým kliknutím na daný test nebo po výběru testu a kliknutí na tlačítko „Začít testování“ zahájíme vybraný test (podrobněji v podkapitole 4.6).

Tlačítka „Přidat test“ a „Odebrat test“ nemohou být v uživatelském režimu použita. Abychom mohli přidat nebo smazat test, musíme se přihlásit pomocí tlačítka „Přihlášení správce“. Po kliknutí na tlačítko se zobrazí přihlašovací okno (obr. 4.3).

Seznam testů MOS

— □ ×

Hlavní nabídka testování MOS

Seznam testů:

ID	Název testu	Počet stran	Počet vzorků	
1	WAV vs MP3s	6	30	

Přihlášení správce

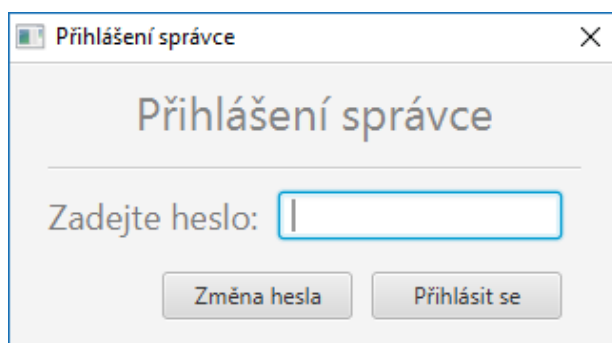
Přidat test

Odebrat test

Začít testování

Obr. 4.2: Seznam testů.

4.4 Správce



Obr. 4.3: Přihlašovací okno správce.

Přihlašovací okno správce neslouží jen k přihlášení, ale je zde umístěno i tlačítko „Změna hesla“, které otevře nové okno pro změnu hesla správce. V tomto okně musíme zadat staré heslo jednou a následně nové heslo dvakrát, aby došlo ke změně hesla. MD5 hash hesla je uložen ve složce „data“, v souboru „password.dat“. Po úspěšném přihlášení se tlačítka „Přidat test“ a „Odebrat test“ aktivují a tlačítko pro přihlášení správce se změní na tlačítko pro odhlášení. Nyní lze v režimu správce smazat a přidat testy.

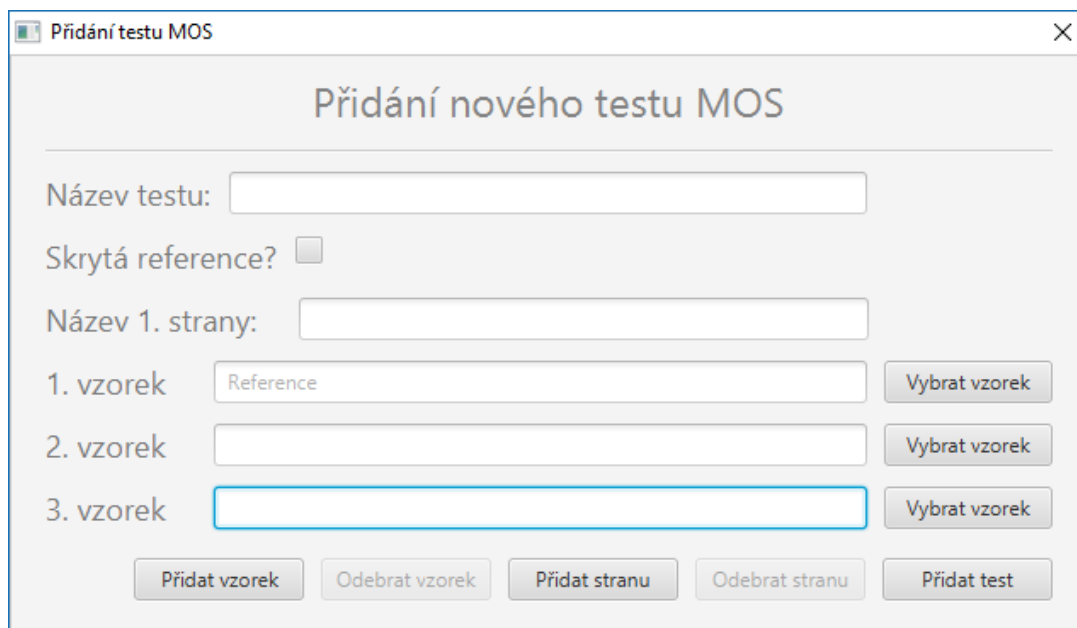
4.5 Přidání nového testu

Okna pro vytvoření nových testů jsou v zásadě stejná, liší se jen v detailech. U všech testovacích metod musíme zadat název nového testu, pojmenovat všechny vzorky a strany testu a vybrat audio soubory pro všechny vzorky.

Tlačítka „Přidat vzorek“ a „Odebrat vzorek“ slouží k přidávání dalších vzorků na danou stranu testu nebo k odebrání vzorků z ní. Jedna strana testu musí obsahovat minimálně 3 vzorky (kromě metody ABX), maximální počet vzorků je omezen na 10. Tlačítka „Přidat stranu“ a „Odebrat stranu“ slouží k přidávání dalších stran nebo k vymazání už existující strany. Minimální počet stran jednoho testu je 1, maximální počet stran testu není omezen. Když jsou všechny pole platně vyplněny, kliknutím na tlačítko „Přidat test“ program přidá test do databáze a zkopíruje audio soubory do nové složky s ID nového testu do složky s názvem testovací metody ve složce „data“ (např. do složky ../data/mos/1).

V následující části jsou popsány rozdíly mezi jednotlivými metodami a obrázky oken pro vytvoření nových testů.

Při vytvoření nového testu MOS máme možnost skryté reference při testování (obr. 4.4). První vzorek je označen jako referenční vzorek.

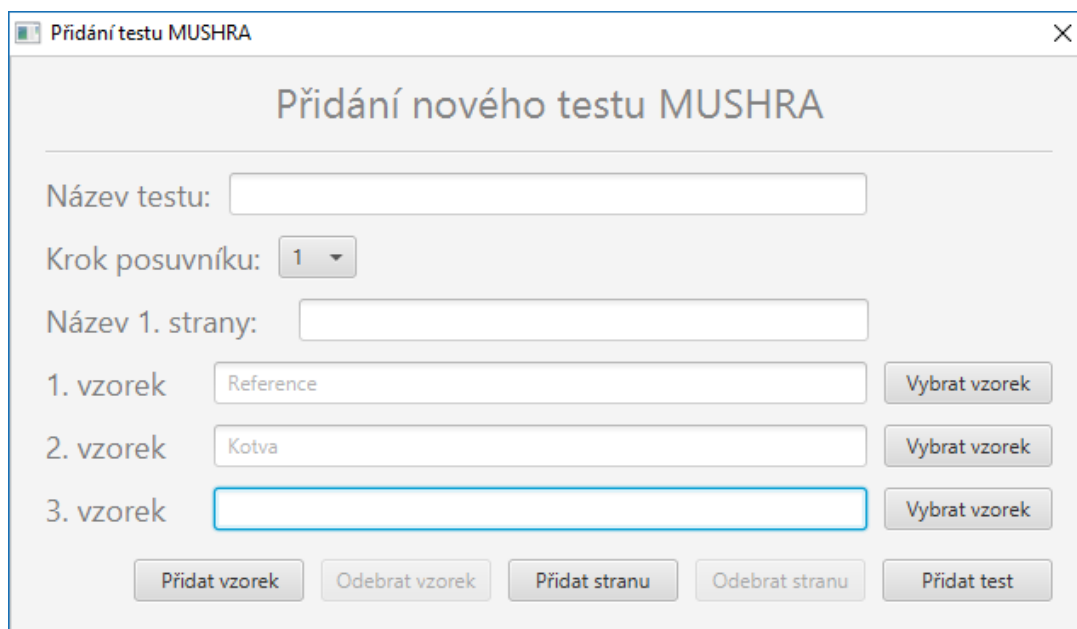


The screenshot shows a dialog box titled "Přidání nového testu MOS". It contains the following fields and controls:

- Název testu:** A text input field.
- Skrytá reference?** A checkbox.
- Název 1. strany:** A text input field.
- 1. vzorek:** A text input field containing the word "Reference", with a "Vybrat vzorek" button to its right.
- 2. vzorek:** An empty text input field, with a "Vybrat vzorek" button to its right.
- 3. vzorek:** An empty text input field, with a "Vybrat vzorek" button to its right.
- At the bottom, there are five buttons: "Přidat vzorek", "Odebrat vzorek", "Přidat stranu", "Odebrat stranu", and "Přidat test".

Obr. 4.4: Přidání nového testu MOS.

Při vytvoření nového testu MUSHRA máme možnost nastavit krok posuvníku při testování (obr. 4.5). První vzorek je referenční vzorek a druhý vzorek je kotva.

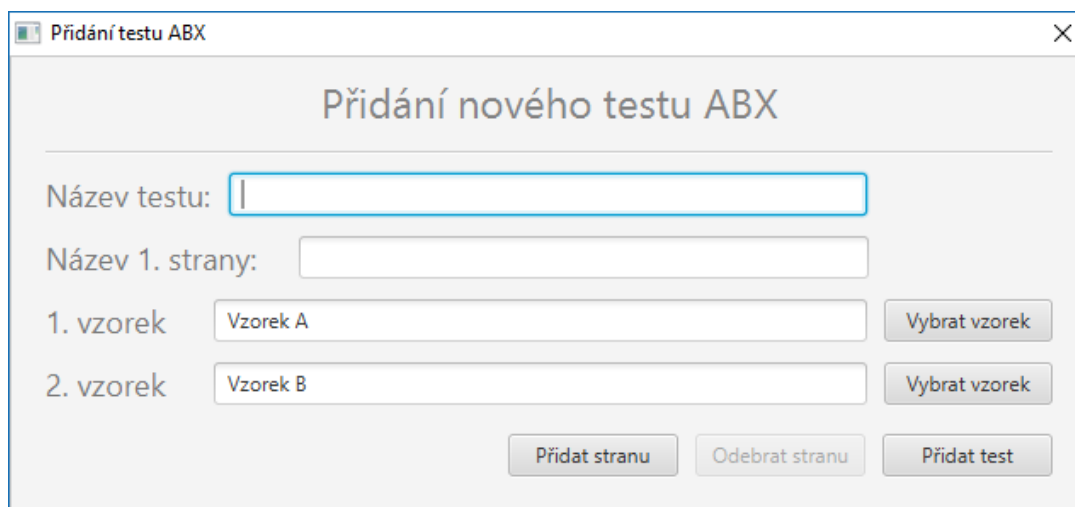


The screenshot shows a dialog box titled "Přidání nového testu MUSHRA". It contains the following fields and controls:

- Název testu:** A text input field.
- Krok posuvníku:** A dropdown menu currently showing the value "1".
- Název 1. strany:** A text input field.
- 1. vzorek:** A text input field containing the word "Reference", with a "Vybrat vzorek" button to its right.
- 2. vzorek:** A text input field containing the word "Kotva", with a "Vybrat vzorek" button to its right.
- 3. vzorek:** An empty text input field, with a "Vybrat vzorek" button to its right.
- At the bottom, there are five buttons: "Přidat vzorek", "Odebrat vzorek", "Přidat stranu", "Odebrat stranu", and "Přidat test".

Obr. 4.5: Přidání nového testu MUSHRA.

Rozdílné je, oproti předchozím metodám, vytvoření nového testu ABX, kde počet vzorků pro jednu stranu testu je pevně nastaven na 2 a vzorky jsou pojmenovány „Vzorek A“ a „Vzorek B“ (obr. 4.6).



Obr. 4.6: Přidání nového testu ABX.

4.6 Testy

Po zahájení testu se objeví nové okno s vybraným testem. Každé okno, bez ohledu na testovací metodu, má 2 tlačítka: „Předešlá strana“ a „Další strana“. Tlačítka se používají pro procházení stran. Na poslední straně testu je místo tlačítka „Další strana“ tlačítko „Uložit výsledky“. Po správném vyplnění testu toto tlačítko uloží výsledky do CSV (Comma-separated values – hodnoty oddělené čárkami) souboru s názvem testu do složky s názvem testovací metody ve složce „results“.

V testech MOS a MUSHRA jsou vzorky na jednotlivých stranách testu vypsány v náhodném pořadí. Jedinou výjimkou je referenční vzorek, který je první ze vzorků a je pojmenován. V testech MOS referenční vzorek může být i skrytý, to znamená, že není značen a je zamíchán mezi ostatní vzorky (obr. 4.7). V testech MUSHRA kromě známé reference je reference spolu s kotvou také zamíchána mezi ostatní vzorky (obr. 4.8).

Vedle posuvníků jsou textová pole, která jsou přednastavena prázdná. Po ohodnocení vzorků posuvníkem se v nich objeví vybraná hodnota. Pokračovat na další stranu testu a uložit výsledky lze, jen když jsou všechny vzorky ohodnoceny, tj. žádné textové pole není prázdné.

Test MOS - 1. strana

WAV vs MP3s - Atmospheric pop

1. vzorek	<input type="range" value="3"/>	<input type="checkbox"/>	Přehrát vzorek
2. vzorek	<input type="range" value="3"/>	<input type="checkbox"/>	Přehrát vzorek
3. vzorek	<input type="range" value="3"/>	<input type="checkbox"/>	Přehrát vzorek
4. vzorek	<input type="range" value="3"/>	<input type="checkbox"/>	Přehrát vzorek
5. vzorek	<input type="range" value="3"/>	<input type="checkbox"/>	Přehrát vzorek

Předešlá strana Další strana

Obr. 4.7: Testovací okno MOS.

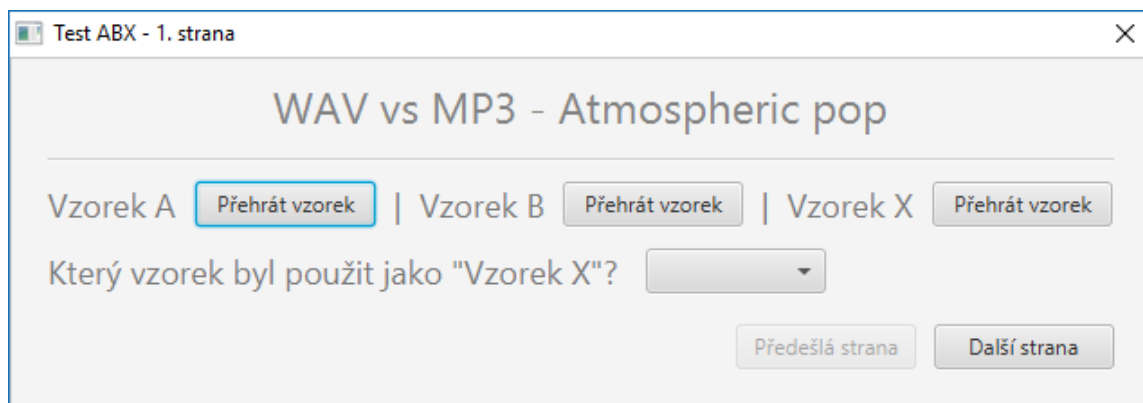
Test MUSHRA - 1. strana

WAV vs MP3s - Atmospheric pop

Reference		<input type="checkbox"/>	Přehrát vzorek
1. vzorek	<input type="range" value="50"/>	<input type="checkbox"/>	Přehrát vzorek
2. vzorek	<input type="range" value="50"/>	<input type="checkbox"/>	Přehrát vzorek
3. vzorek	<input type="range" value="50"/>	<input type="checkbox"/>	Přehrát vzorek
4. vzorek	<input type="range" value="50"/>	<input type="checkbox"/>	Přehrát vzorek
5. vzorek	<input type="range" value="50"/>	<input type="checkbox"/>	Přehrát vzorek

Předešlá strana Další strana

Obr. 4.8: Testovací okno MUSHRA.



Obr. 4.9: Testovací okno ABX.

V testech ABX vzorek X je náhodně vybrán ze vzorků A a B. Pro pokračování na další stranu a pro uložení výsledků je nutné vybrat jednu z možností v tzv. Choice-Boxu na všech stranách testu (obr. 4.9). Export výsledků je jednodušší než u předchozích metod, neboť není třeba uložit číselné hodnoty, jen to, jestli uživatelův výběr byl správný či nikoliv.

5 TESTOVÁNÍ PROGRAMU

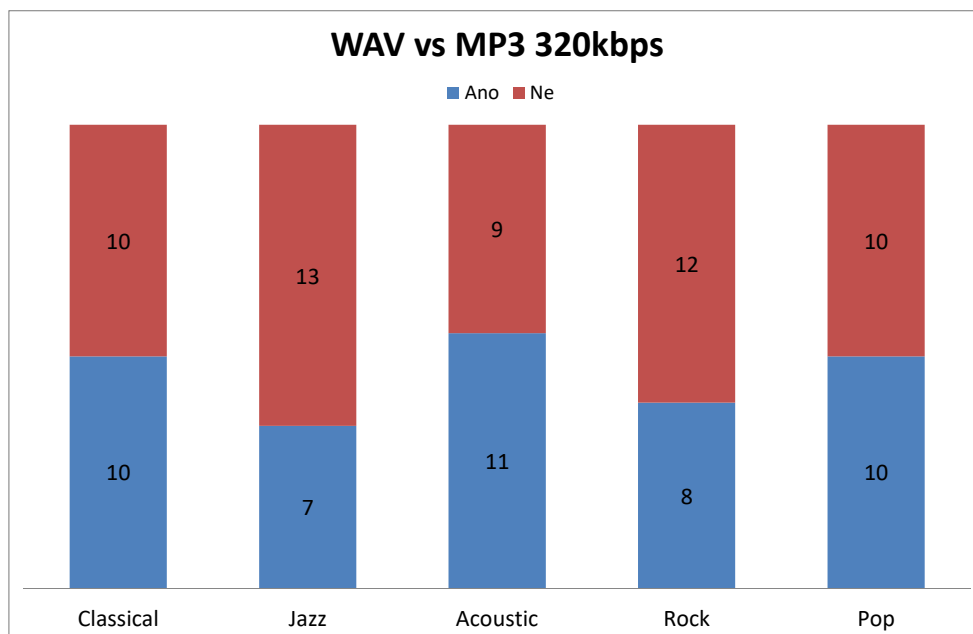
Pro testování programu byly vytvořeny dohromady čtyři poslechové testy: jeden test MUSHRA a tři testy ABX. Téma testování bylo porovnání subjektivní kvality formátů WAV (Waveform Audio File Format), kvalitní a méně kvalitní MP3 (MPEG-1 Audio Layer III nebo MPEG-2 Audio Layer III). Testování probíhalo v laboratoři SC1.23 na FEKT VUT v Brně. Program byl otestován dvaceti lidmi.

5.1 Rozdíly mezi formáty

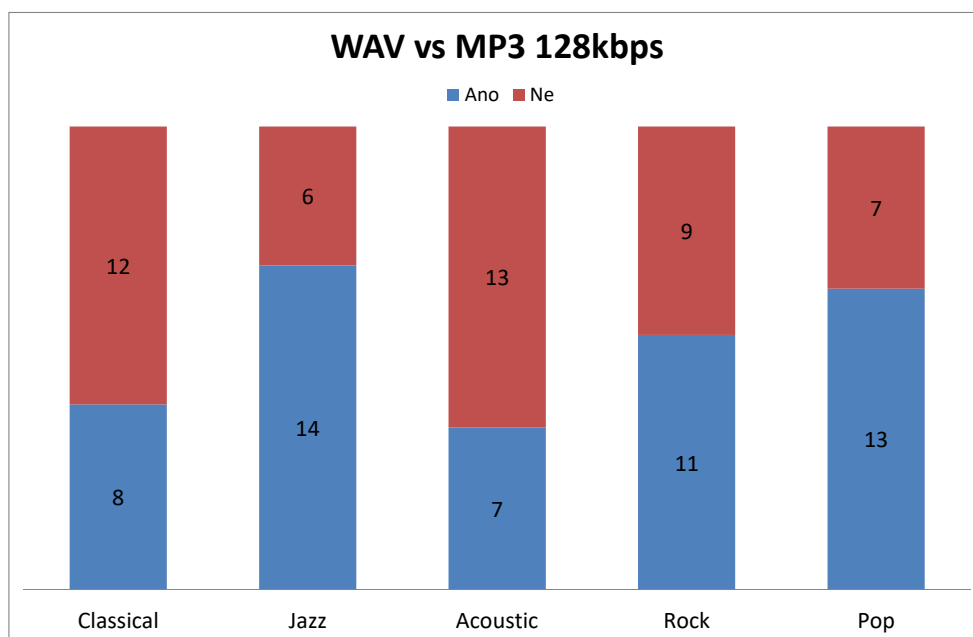
Formát WAV je bezztrátový formát, který je používán pro ukládání audio signálů do souborů. Kvůli vlastnosti bezztrátového ukládání soubory formátu WAV zabírají relativně velké místo na hard disku [26].

Formát MP3 je ztrátový formát a byl navržen tak, aby audio soubory zabíraly méně místa, bez významných rozdílů pro posluchače. Formát MP3 se snaží odstranit podle psychoakustického modelu nadbytečné informace z původního signálu. MP3 tedy odstraňuje informace ze signálu, které vlastně nejsou pro lidské ucho slyšitelné nebo významné a tím zmenšuje velikost souborů (zhruba na desetinu formátu WAV) [27].

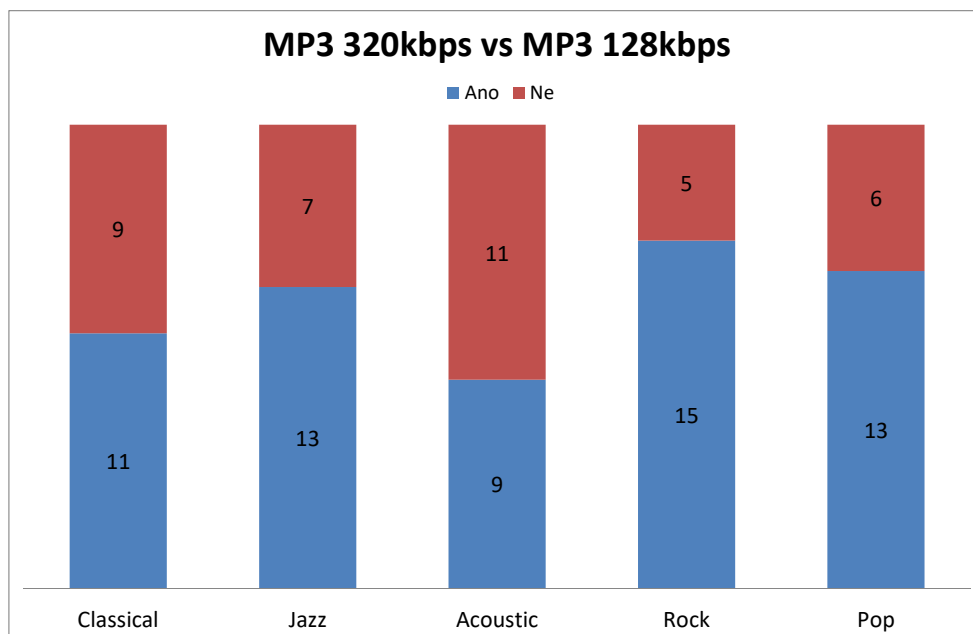
5.2 Výsledky testů ABX



Obr. 5.1: Četnost správných výběrů testu ABX – WAV vs MP3 320kbps

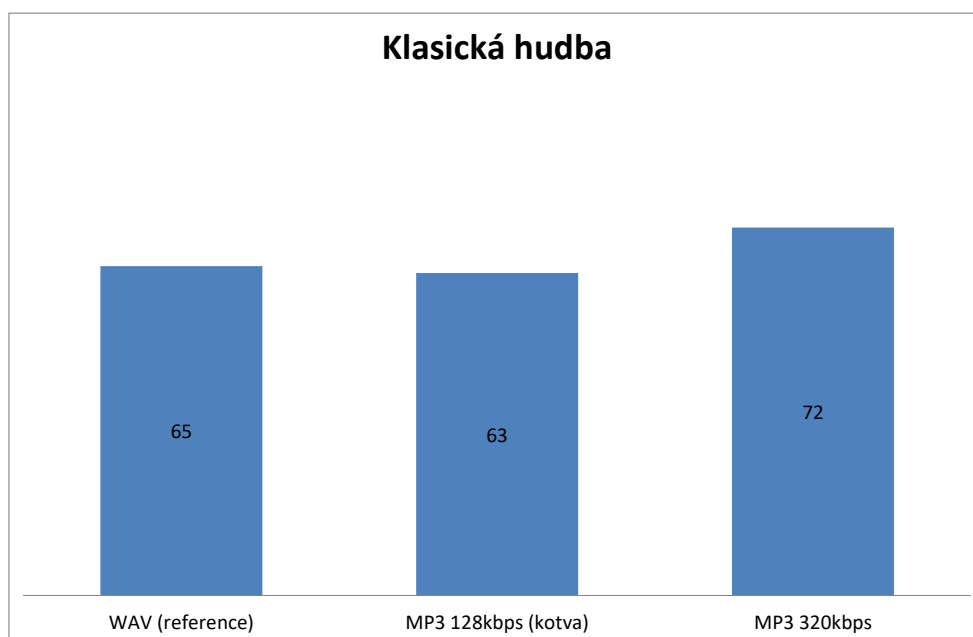


Obr. 5.2: Četnost správných výběrů testu ABX – WAV vs MP3 128kbps

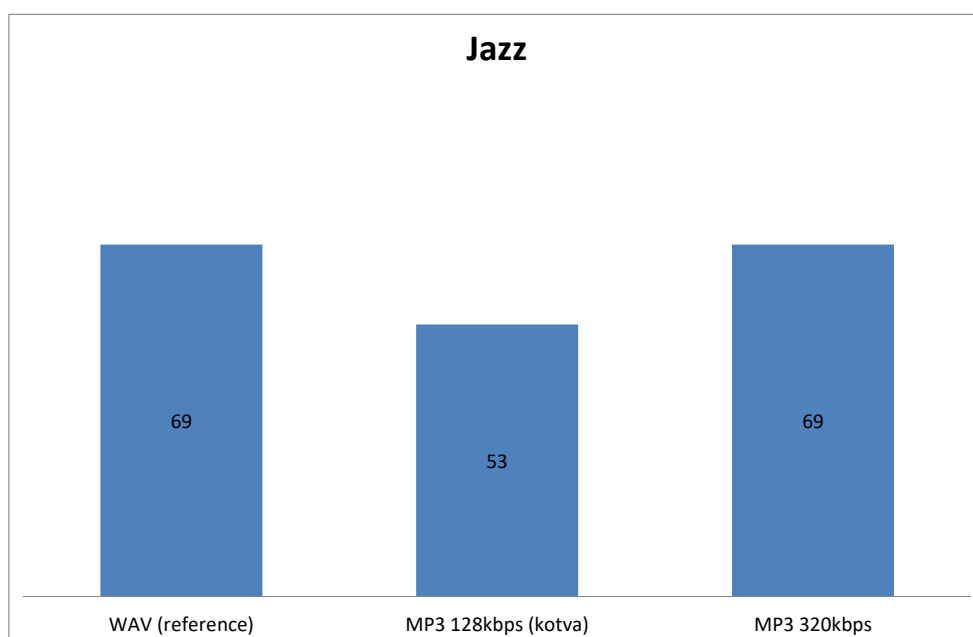


Obr. 5.3: Četnost správných výběrů testu ABX – MP3 320kbps vs MP3 128kbps

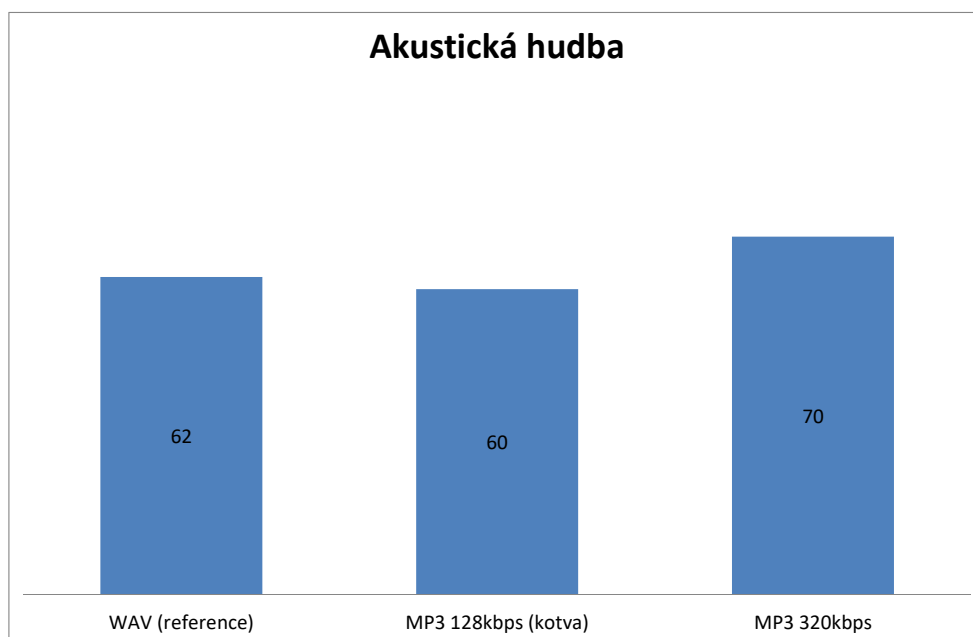
5.3 Výsledky testu MUSHRA



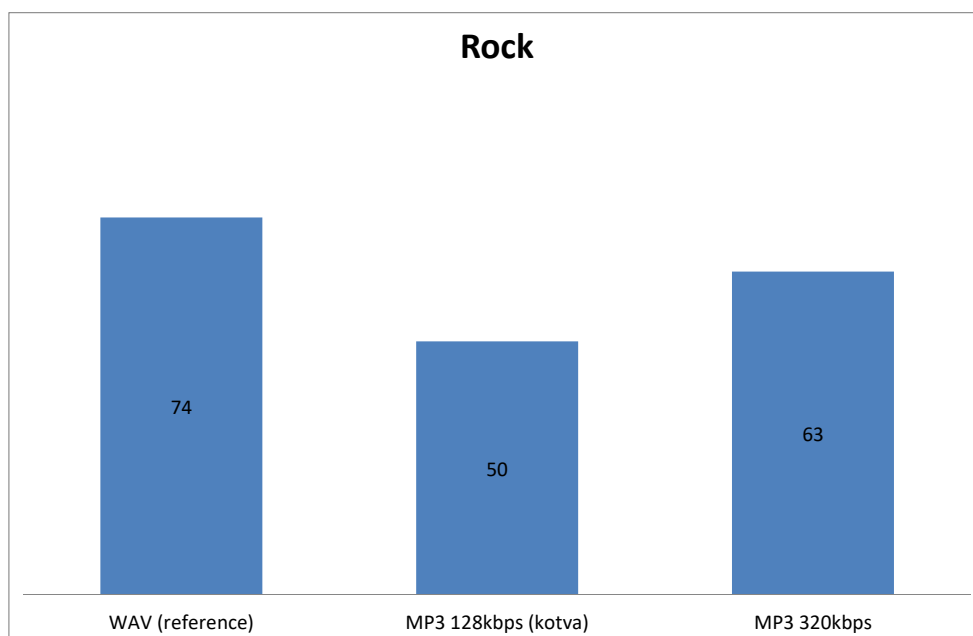
Obr. 5.4: Střední hodnoty testu MUSHRA pro klasickou hudbu



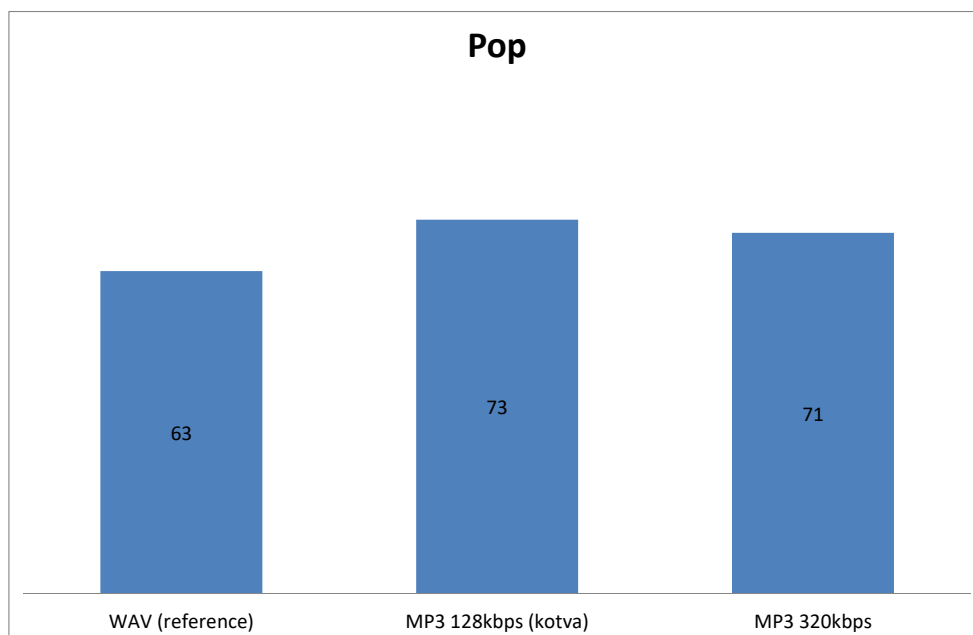
Obr. 5.5: Střední hodnoty testu MUSHRA pro jazz



Obr. 5.6: Střední hodnoty testu MUSHRA pro akustickou hudbu



Obr. 5.7: Střední hodnoty testu MUSHRA pro rockovou hudbu



Obr. 5.8: Střední hodnoty testu MUSHRA pro popovou hudbu

5.4 Vyhodnocení testů

Podle grafů 5.1, 5.2 a 5.3 lze konstatovat, že ve většině případů u testů ABX výsledky nejsou jednoznačné. Nejméně správných výběrů (46%) bylo mezi formáty WAV a kvalitní MP3, což znamená, že je nejmenší rozdíl mezi subjektivní kvalitou těchto formátů. Mezi méně kvalitní MP3 a formátem WAV můžeme říci, že výběr správného vzorku záviselo hodně na žánru, nicméně správných odpovědí u tohoto testu bylo více než u předchozího (53%). Nejvíce správných výběrů (61%) bylo mezi porovnáváním kvalitní a méně kvalitní MP3, což znamená, že pro posluchači největší subjektivní rozdíl je mezi těmito formáty.

U testu MUSHRA jsou výsledky o něco jednodušší. Lze konstatovat, že kromě žánru „Pop“ jsou vždy nejhůř hodnoceny méně kvalitní MP3. Hodnoty u žánru „Rock“ vyšly zcela podle očekávání, bezztrátový WAV formát měl nejvyšší hodnocení, naopak 128kbps MP3 nejnižší. Zajímavé ale je, že podle výsledku testu můžeme říci, že ve většině případů je lépe hodnocen kvalitní MP3 formát než WAV, i když podle teorie obsahuje méně informací o původním signálu. Dle mého názoru je to způsobeno rozšířením formátu MP3, na který se lidské ucho zvyká velice rychle.

Celkově můžeme říci, že i když jsou mezi formáty WAV a MP3 rozdíly, podle výsledků testů nejsou tyto rozdíly veliké, mimoto je zajímavé, že většina uživatelů preferovala kvalitní (320kbps) ztrátový formát MP3.

5.5 Hodnocení programu

Při provádění poslechových testů program fungoval přesně podle očekávání, všechno fungovalo bez problémů. Každého, kdo se zúčastnil na poslechovém testu, jsem se ptal na celkové hodnocení programu. V tabulce 5.1 jsou všechny tyto hodnoty podle škály MOS. Průměrné hodnocení programu je tedy 4,3. Celkově se dá říci, že uživatelé byly spokojeni s programem a hodnotili ho velmi kladně.

Tab. 5.1: Hodnocení programu podle uživatelů

Pořadí	Hodnocení
1	4
2	4
3	4
4	4
5	4
6	5
7	4
8	5
9	5
10	5
11	3
12	4
13	5
14	4
15	5
16	4
17	5
18	4
19	3
20	5

6 ZÁVĚR

První kapitola bakalářské práce se zabývá především problematikou poškození signálů. Je zde popsáno celkem sedm typů poškození. Nejznámějšími z nich jsou problematika chybějících vzorků signálů, zašumění signálů a clipping. Chybějící vzorky v signálech a oba typy clippingu (soft clipping a hard clipping) jsou zde znázorněny i na grafech. Dále lze zde nalézt i matematické vyjádření hard clippingu a soft clippingu. Ostatní typy poškození signálů jsou jenom zmíněny a stručně shrnuty, jsou to clicking, crackling, hissing a rumbling.

Druhá kapitola je zaměřena na rekonstrukci již v prvním kapitole zmíněných poškození signálů. Na začátku kapitoly je stručný popis rekonstrukce audio signálů, pak jsou rozebírány typy rekonstrukcí, z nichž největší pozornost je věnována úlohám declipping, inpainting a denoising. U těchto třech typů jsou zmíněny i principy, kterých rekonstrukční metody využívají. Následující typ rekonstrukce audio signálů je declipping, který je stručně popsán a znázorněn i na grafech. Další tři rekonstrukční typy jsou zmíněny a krátce vysvětleny, jsou to decrackling, dehissing a derumbling. Poslední čtyři rekonstrukční metody se využívají hlavně při provádění audio masteringu.

Ve třetí kapitole byly popsány objektivní i subjektivní testy vyhodnocující kvalitu audio signálů. Z objektivních testů je nejprve zmíněn asi nejznámější ze všech, SNR. Následně jsou popsány sofistikovanější testy, které se snaží do výsledku zahrnout i vlastnosti lidského sluchu, jsou to PEMO-Q, PEAQ, PESQ, POLQA a ViSQOL. Ze subjektivních metod vyhodnocení kvality audio signálů je nejprve rozebírán MOS, kde je uvedena pětibodová škála, která je známá a používaná v celém světě. Následně je popsáno dalších pět testů, jsou to test MUSHRA, ABX, Duo-trio test, Triangle test a Paired comparison test.

Cílem bakalářské práce bylo implementování programu pro subjektivní hodnocení kvality restaurace audio signálů v programovacím jazyce Java. Tento cíl byl splněn, program je stručně popsán ve čtvrté kapitole. Dalším cílem bylo testování programu, které bylo provedeno prostřednictvím poslechových testů. Výsledky poslechových testů jsou zobrazeny v páté kapitole. Na konci páté kapitoly je stručné shrnutí výsledků těchto testů a hodnocení programu podle uživatelů.

LITERATURA

- [1] POSPÍŠIL, Jiří. *Doplňování chybějících úseků v audio signálu*. Brno, 2012, 49 s. Bakalářská práce. VUT v Brně. Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Václav Mach.
Dostupné z URL: <<https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/9382/xpospi65.pdf>>.
- [2] SCHIMMEL, Jiří. *Studiová a hudební elektronika*. Brno: VUT v Brně. Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací 2015. s. 1-197. 978-80-214-4452-2. (cs)
Dostupné z URL: <https://www.vutbr.cz/www_base/priloha.php?dpid=61387>.
- [3] ZÁVIŠKA, Pavel. *Restaurace audiosignálů založená na řídkých reprezentacích*. Brno, 2017, 94 s. Diplomová práce. VUT v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: doc. Mgr. Pavel Rajmic, Ph.D.
Dostupné z URL: <<https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/65668/Restaurace%20audiosignalu%20zalozena%20na%20ridkych%20reprezentacich.pdf>>.
- [4] ROBIN, Michael. Analog audio noise. *Broadcast Engineering* [online]. Overland Park: New Bay Media, 2005, 47(8), 20 [cit. 2017-11-28]. ISSN 00071994.
Dostupné z URL: <<http://search.ebscohost.com.ezproxy.lib.vutbr.cz/login.aspx?direct=true&db=bth&AN=18098388&lang=cs&site=ehost-live>>.
- [5] MACHAT, Karl. Declicking The Clicks! From the website: *Mister's Mastering House* [cit. 2017-11-20].
Dostupné z URL: <<http://www.about-audio-mastering-software.com/declicking.html>>.
- [6] OZDOBINSKI, Roman. *Aplikace metod učení slovníku pro Audio Inpainting*. Brno, 2014, 56 s. Diplomová práce. VUT v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce: Ing. Václav Mach.
Dostupné z URL: <https://dspace.vutbr.cz/bitstream/handle/11012/32908/dp_ozdobinski.pdf>.

- [7] ADLER, Amir, EMIYA Valentin, JAFARI Maria G., ELAD Michael, GRIBONVAL Rémi a PLUMBLEY Mark D. Audio Inpainting. *Audio, Speech, and Language Processing, IEEE Transactions on* [online]. USA: IEEE, 1203, 20(3), 922-932 [cit. 2017-10-25]. DOI: 10.1109/TASL.2011.2168211. ISSN 1558-7916. Dostupné z URL: <<http://ieeexplore.ieee.org.ezproxy.lib.vutbr.cz/stamp/stamp.jsp?arnumber=6020748>>.
- [8] LEVADA, Alexandre L. M. a CORREA Debora C. An adaptive approach for contextual audio denoising using local Fisher information. In: *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on* [online]. IEEE Publishing, 1105, s. 125-128 [cit. 2017-11-01]. DOI: 10.1109/ISCAS.2011.5937517. ISBN 978-1-4244-9473-6. ISSN 0271-4302. Dostupné z URL: <<http://ieeexplore.ieee.org.ezproxy.lib.vutbr.cz/stamp/stamp.jsp?arnumber=5937517>>.
- [9] ALI, M. A. a SHEMI P. M. An improved method of audio denoising based on wavelet transform. In: *Power, Instrumentation, Control and Computing (PICC), 2015 International Conference on* [online]. IEEE, 1512, s. 1-6 [cit. 2017-11-01]. DOI: 10.1109/PICC.2015.7455802. Dostupné z URL: <<http://ieeexplore.ieee.org.ezproxy.lib.vutbr.cz/stamp/stamp.jsp?arnumber=7455802>>.
- [10] MACH, Václav. *Inpainting of Missing Audio Signal Samples*: doctoral thesis. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Telecommunications, 2016. 125 p. Supervised by doc. Mgr. Pavel Rajmic, Ph.D. Dostupné z URL: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/58311/mach_disertace.pdf>.
- [11] JOHNSON, Don H. Signal-to-noise ratio. *Scholarpedia*, 2006, 1.12: 2088. Dostupné z URL: <www.scholarpedia.org/article/Signal-to-noise_ratio>.
- [12] RAO, Sreenivasa K. Voice conversion by mapping the speaker-specific features using pitch synchronous approach. *Computer Speech & Language* [online]. Elsevier, 2010, 24(3), 474-494 [cit. 2017-11-22]. DOI: 10.1016/j.csl.2009.03.003. ISSN 0885-2308. Dostupné z URL: <<http://ieeexplore.ieee.org.ezproxy.lib.vutbr.cz/stamp/stamp.jsp?arnumber=1709880>>.

- [13] HU, Xiaopeng, HE Guiming a ZHOU Xiaoping. PEAQ-based psychoacoustic model for perceptual audio coder. In: *Advanced Communication Technology, 2006. ICACT 2006. The 8th International Conference* [online]. IEEE, 2006, 3, 5 pp.-1823 [cit. 2017-11-02]. DOI: 10.1109/ICACT.2006.206344. ISBN 89-5519-129-4.
Dostupné z URL: <<http://ieeexplore.ieee.org.ezproxy.lib.vutbr.cz/stamp/stamp.jsp?arnumber=1625949>>.
- [14] ITU-T Recommendation P.862. *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*. Methods for objective and subjective assessment of quality, 2001.
- [15] ITU-T Recommendation P.862.1. *Mapping function for transforming P.862 raw result scores to MOS-LQO*. Methods for objective and subjective assessment of quality, 2003.
- [16] ITU-T Recommendation P.863. *Perceptual objective listening quality assessment*. Methods for objective and subjective assessment of speech quality, 2014.
- [17] HINES, Andrew, SKOGLUND Jan, KOKARAM Anil C, HARTE Naomi. *VISQOL: an objective speech quality model*. EURASIP Journal on Audio, Speech, and Music Processing, 2015, 2015:13.
Dostupné z URL: <<https://asmp-urasipjournals.springeropen.com/track/pdf/10.1186/s13636-015-0054-9>>.
- [18] HINES, Adnrew, GILLEN Eioin, KELLY Damien, SKOGLUND Jan, KOKARAM Anil, HARTE Naomi. *ViSQOLAudio: An Objective Audio Quality Metric for Low Bitrate Codecs*. J. Acoust. Soc. Am., EL 137 (6), June 2015.
Dostupné z URL: <<http://asa.scitation.org/doi/pdf/10.1121/1.4921674>>.
- [19] STREIJL, Robert, WINKLER Stefan a HANDS David. Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives. *Multimedia Systems* [online]. Berlin/Heidelberg: Springer Berlin Heidelberg, 1603, 22(2), 213-227 [cit. 2017-11-02]. DOI: 10.1007/s00530-014-0446-1. ISSN 0942-4962.
Dostupné z URL: <<https://link-springer-com.ezproxy.lib.vutbr.cz/content/pdf/10.1007%2Fs00530-014-0446-1.pdf>>.

- [20] ŠPETA, Marek. *Programová implementace subjektivních testů zvukové kvality*. Brno: VUT v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav Telekomunikací, 2011. 69 s. Vedoucí diplomové práce Ing. Ondřej Rášo. Dostupné z URL: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/6139/xspeta00_dip_2011.pdf>.
- [21] HUANG, Yu-Ting a LAWLESS Harry T. SENSITIVITY OF THE ABX DISCRIMINATION TEST. *Journal of Sensory Studies* [online]. Oxford, UK: Blackwell Publishing, 9807, 13(2), 229-239 [cit. 2017-11-02]. DOI: 10.1111/j.1745-459X.1998.tb00085.x. ISSN 0887-8250. Dostupné z URL: <<http://onlinelibrary.wiley.com.ezproxy.lib.vutbr.cz/doi/10.1111/j.1745-459X.1998.tb00085.x/epdf>>.
- [22] ISO Standard 10399:2004. *Sensory analysis – Methodology – Duo-trio test*. Geneva, Switzerland: International Organization for Standardization.
- [23] ISO Standard 4120:2004. *Sensory analysis – Methodology – Triangle test*. Geneva, Switzerland: International Organization for Standardization.
- [24] ISO Standard 5495:2005. *Sensory analysis – Methodology – Paired comparison test*. Geneva, Switzerland: International Organization for Standardization.
- [25] POTTS, Jasper, HILDEBRADNT Nancy, GORDON Joni a CASTILLO Cindy. JavaFX. *Getting Started with JavaFX*. Release 8. E50607-02. Copyright © 2008, 2014, Oracle [cit. 2018-04-18]. Dostupné z URL: <<https://docs.oracle.com/javase/8/javafx/JFXST.pdf>>.
- [26] WAV, Wikipedia [online]. 2018 [cit. 2018-05-17]. Dostupné z URL: <<https://cs.wikipedia.org/wiki/WAV>>
- [27] MP3, Wikipedia [online]. 2018 [cit. 2018-05-17]. Dostupné z URL: <<https://cs.wikipedia.org/wiki/MP3>>.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

ADAT	Alesis Digital Audio Tape
CSV	Comma-separated values – hodnoty oddělené čárkami
GUI	graphical user interface – grafické uživatelské rozhraní
ISO	International Organization for Standardization – Mezinárodní organizace pro normalizaci
ITU	International Telecommunication Union – Mezinárodní telekomunikační unie
MOS	Mean Opinion Score – průměrná známka hodnocení
MP3	MPEG-1 Audio Layer III nebo MPEG-2 Audio Layer III
MUSHRA	MUlti Stimulus test with Hidden Reference and Anchor – Více-stimulový test se skrytou referencí a kotvou
ODG	Objective Difference Grade – objektivní rozdílový stupeň
PEAQ	Perceptual Evaluation of Audio Quality – vjemové hodnocení kvality zvuku
PESQ	Perceptual Evaluation of Speech Quality – vjemové hodnocení kvality řeči
POLQA	Perceptual Objective Listening Quality Assessment – objektivní hodnocení vnímané kvality poslechu
PSM	Perceptual Similarity Measure – míra vjemové podobnosti
PSMt	Temporary Perceptual Similarity Measure – míra dočasné vjemové podobnosti
SDG	Subjective Difference Grade – subjektivní rozdílový stupeň
Sigmedia	Signal Processing and Media Applications
SNR	Signal-to-noise ratio – odstup signálu od šumu
S/PDIF	Sony/Philips Digital Interface Format
ViSQOL	Virtual Speech Quality Objective Listener – objektivní posluchač virtuální řečové kvality
WAV	Waveform Audio File Format

SEZNAM PŘÍLOH

A Obsah přiloženého CD

50

A OBSAH PŘÍLOŽENÉHO CD

/	kořenový adresář přiloženého CD
└─ audio	zvukové ukázky průběhů
└─ sin440 4 nulove vzorky.wav	
└─ sin440 20 nulovych vzorku.wav	
└─ sin440 click.wav	
└─ sin440 declick.wav	
└─ sin440 hardclip.wav	
└─ sin440 neposkozen.wav	
└─ sin440 softclip.wav	
└─ figures	grafy z Matlabu
└─ sin440_4nul_vzor.fig	
└─ sin440_20nul_vzor.fig	
└─ sin440_click.fig	
└─ sin440_declick.fig	
└─ sin440_hardclip.fig	
└─ sin440_hardclip_spektrum.fig	
└─ sin440_neposkozen.fig	
└─ sin440_softclip.fig	
└─ sin440_softclip_spektrum.fig	
└─ sin440_spektrum.fig	
└─ program	složka s aplikací
└─ project	složka s Eclipse projektem aplikace
└─ results	výsledky testů
└─ abx.xlsx	
└─ mushra.xlsx	
└─ user.xlsx	hodnocení uživatelů