



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

ZVÝŠENÍ KVALITY AUDIO SIGNÁLŮ POMOCÍ KOMBINACE VÍCE LEVNÝCH A JEDNOHO DRAŽŠÍHO A/D PŘEVODNÍKU

AUDIO ENHANCEMENT USING DIFFERENT-QUALITY A/D CONVERTERS

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Vojtěch Kovanda

VEDOUCÍ PRÁCE

SUPERVISOR

prof. Mgr. Pavel Rajmic, Ph.D.

BRNO 2024

Diplomová práce

magisterský navazující studijní program **Audio inženýrství**
specializace Zvuková produkce a nahrávání
Ústav telekomunikací

Student: Bc. Vojtěch Kovanda

ID: 209409

Ročník: 2

Akademický rok: 2023/24

NÁZEV TÉMATU:

Zvýšení kvality audio signálů pomocí kombinace více levných a jednoho dražšího A/D převodníku

POKYNY PRO VYPRACOVÁNÍ:

Student se bude věnovat rekonstrukci signálů. Práce by měla odpovědět na následující otázku: Jak vysoké kvality audio signálů je možné dosáhnout pomocí řady A/D převodníků s nízkou bitovou hloubkou (např. 16 bitů) a nízkým vzorkovacím kmitočtem (48 kHz) v kombinaci s jedním převodníkem s vysokou bitovou hloubkou (24, 32 bitů) a vysokým vzorkovacím kmitočtem (96 kHz či více)? Student se nejprve seznámí s vlastnostmi relevantních A/D převodníků. Poté si nastuduje typy kvantování a sestaví jejich matematický model. Na základě literatury navrhne optimalizační postup, jak z množiny různě kvantovaných signálů rekonstruovat signál s vysokou bitovou hloubkou a vysokým vzorkovacím kmitočtem. Postup otestuje na syntetických i reálných signálech a vyhodnotí získanou kvalitu pomocí objektivních evaluačních kritérií jako jsou SDR nebo PEMO-Q.

DOPORUČENÁ LITERATURA:

- [1] Mokrý, O., Rajmic, P., Záviška, P. Flexible framework for audio restoration. In Proceedings of the 23rd International Conference on Digital Audio Effects (DAFx) 2020.
- [2] Záviška, P., Rajmic, P., Mokrý, O. Audio Dequantization Using (Co)Sparse (Non)Convex Methods. In Proceedings IEEE ICASSP, Canada, June 2021.

Termín zadání: 5.2.2024

Termín odevzdání: 21.5.2024

Vedoucí práce: prof. Mgr. Pavel Rajmic, Ph.D.

doc. Ing. Jiří Schimmel, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

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

ABSTRAKT

Práce řeší dekvantizaci několika audio signálů na základě řídkosti, při čemž využívá kombinace řady levnějších a jednoho dražšího A/D převodníku. Problém řeší jako konvexní optimalizaci za použití určitého proximálního algoritmu, který je implementovaný v prostředí MATLAB. Za použití algoritmu byly provedeny rekonstrukce reálných a syntetických signálů. Srovnání úspěšnosti rekonstrukcí je vyhodnocováno pomocí objektivních kritérií, jako je odstup signálu od zkreslení (SDR) a PEMO-Q, které zohledňuje posluchové vlastnosti signálů.

KLÍČOVÁ SLOVA

Analogově digitální převod, analogově digitální převodníky, dekvantizace audia, proximální algoritmy, konvexní optimalizace, řídká reprezentace signálů

ABSTRACT

This paper deals with sparsity based audio dequantization using combination of few cheaper and one more expensive A/D converters. Paper finds solution of this problem as convex optimization using generic proximal algorithm implemented in MATLAB. Real and synthetic signal reconstructions were performed using this algorithm. The reconstruction performance is evaluated using objective criteria such as Signal-to-Distortion Ratio (SDR) and PEMO-Q, which consider the perceptual properties of the signals.

KEYWORDS

Analog-to-digital conversion, analog-to-digital converters, audio dequantization, proximal algorithms, convex optimization, sparsity

KOVANDA, Vojtěch. *Zvýšení kvality audio signálů pomocí kombinace více levných a jednoho dražšího A/D převodníku*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2024, 39 s. Diplomová práce. Vedoucí práce: prof. Mgr. Pavel Rajmic, Ph.D.

Prohlášení autora o původnosti díla

Jméno a příjmení autora:	Bc. Vojtěch Kovanda
VUT ID autora:	209409
Typ práce:	Diplomová práce
Akademický rok:	2023/24
Téma závěrečné práce:	Zvýšení kvality audio signálů pomocí kombinace více levných a jednoho dražšího A/D převodníku

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom 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*

*Autor podepisuje pouze v tištěné verzi.

Obsah

Úvod	8
1 Teoretická část studentské práce	9
1.1 Analogově digitální převod	9
1.1.1 Vzorkování	9
1.1.2 Kvantizace	11
1.1.3 Analogově digitální převodníky	15
1.2 Dekvantizace	16
1.2.1 Řídká reprezentace signálů	17
1.3 Konvexní optimalizace	18
1.3.1 Proximální algoritmy	18
1.3.2 Condatův algoritmus	20
2 Výsledky studentské práce	22
2.1 Formulace problematiky	22
2.2 Navrhnuté řešení	23
2.3 Řešení a výsledky	30
2.3.1 Implementace v prostředí MATLAB	30
2.3.2 Výsledky	31
Závěr	38
Literatura	39

Seznam obrázků

1.1	Původní signál před vzorkováním	9
1.2	Periodický sled impulzů	10
1.3	Navzorkovaný signál	10
1.4	Aliasing v časové doméně signálu	11
1.5	Kvantizace signálu	12
1.6	Kvantizace mid-tread	13
1.7	Kvantizace mid-riser	13
1.8	Kvantizační chyba	14
1.9	Spektrogram nahrávky klavíru	17
2.1	Rozhodovací úrovně kvantizace	29
2.2	Projekce signálu na množinu přípustných hodnot	29
2.3	Úsek syntetického signálu	30
2.4	SDR v závislosti na počtu iterací pro nahrávky smyčců	32
2.5	SDR pro bitovou hloubku 8 bitů pro nahrávky smyčců	33
2.6	Průměrné SDR a ODG pro různou bitovou hloubku pro nahrávky smyčců	33
2.7	SDR v závislosti na počtu iterací pro syntetické signály	34
2.8	SDR pro bitovou hloubku 8 bitů pro syntetické signály	34
2.9	průměrné SDR a ODG pro různou bitovou hloubku pro syntetické signály	35
2.10	SDR v závislosti na počtu iterací pro klavír	36
2.11	SDR pro bitovou hloubku 8 bitů pro různé nástroje	37
2.12	SDR a ODG pro různou bitovou hloubku pro různé nástroje	37

Úvod

Zvyšování bitové hloubky audia je problematika, kterou se snaží řešit celá řada různých typů algoritmů, při které se uplatňují klasické i novější metody. Stejnou problematiku řeší například zdroj [2], který různé metody pro zvyšování bitové hloubky audia srovnává. Stejně metody jsou základem také pro tuto práci, která je ale jedinečná svojí myšlenkou použití v praxi, kde se při zvyšování bitové hloubky více signálů uvažuje i jejich součet, který teoreticky poskytuje více informací pro rekonstrukci jednotlivých signálů.

Práce se zabývá zvyšováním kvality audio signálů na základě kombinace několika levnějších a jednoho dražšího A/D převodníku. Předpokládejme, že známe několik signálů po A/D převodu v horší kvalitě a jejich součet v lepší kvalitě. Toho lze využít při dekvantizaci jednotlivých signálů. Kritérium pro dekvantizaci zde bude řídká reprezentace signálu.

Práce je členěna do několika částí. V teoretické části, která se dále člení do tří podkapitol je prvně popsán analogově digitální převod a jsou zde uvedeny vybrané typy A/D převodníků. Další podkapitola se zabývá dekvantizací audia a poslední podkapitola teoretické části se věnuje konvexní optimalizaci. V kapitole Výsledky studenstské práce je nejprve formulována problematika vycházející ze zadání práce. Dále je nastíněna metoda pro její řešení. V poslední části je popsána implementace algoritmu v prostředí MATLAB a jsou zde zobrazeny výsledky rekonstrukcí hodnoceny objektivními kritérii jako je SDR nebo PEMO-Q.

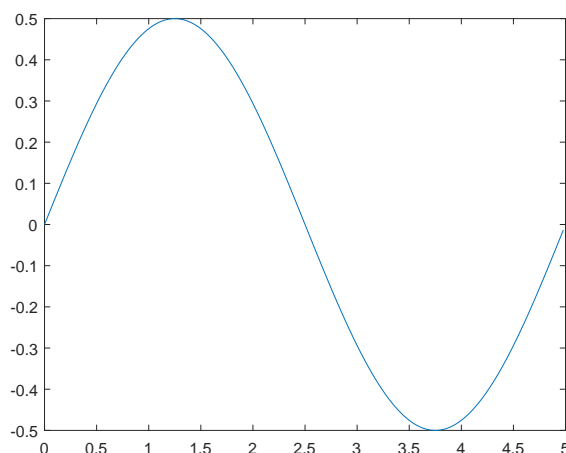
1 Teoretická část studentské práce

1.1 Analogově digitální převod

Reálné signály z praxe, ať už audio nebo jiné, kterými se zabýváme, jsou signály spojité a ve spojitém čase a abychom s nimi mohli pracovat ve výpočetní technice je nutné je reprezentovat digitálně, to je diskretizovat signály v čase a jednotlivým časovým vzorkům přiřadit konečnou hodnotu veličiny, kterou signál reprezentuje. K tomu slouží A/D převod a nástroje provádějící tento úkol nazýváme A/D převodníky. Základem digitalizace signálů je tedy vzorkování a kvantizace, těmto procesům se tato kapitola budeme dále věnovat.

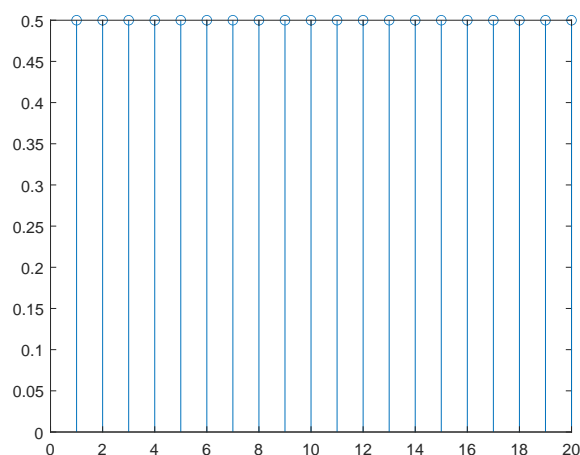
1.1.1 Vzorkování

Jak už bylo zmíněno, při převodu signálu z analogového na digitální je třeba jej diskretizovat v časové oblasti. To znamená periodicky odečíst vzorek původního signálu vždy s určitou vzorkovací frekvencí f_{vz} . Ideální vzorkování se provádí pomocí sledu periodických impulzů o konstantní periodě a amplitudě. Původní signál je amplitudově modulovaný sledem impulzů [1]. Důležitý je pohled na zmíněné signály v kmi-

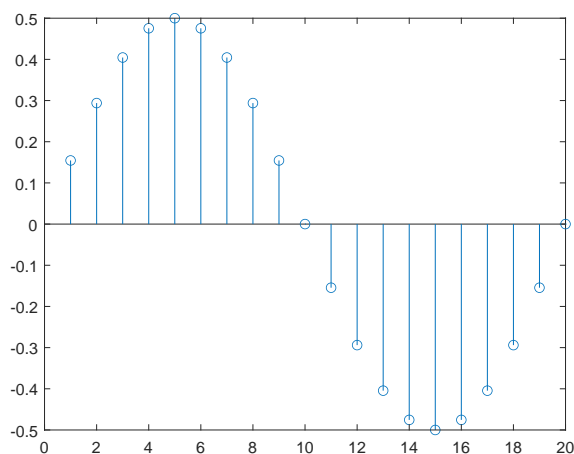


Obr. 1.1: Původní harmonický signál. Na ose y je amplituda signálu, na ose x je čas

točtové oblasti. Spektrum sledu impulzů je teoreticky diskrétní a periodické s celočíselnými násobky vzorkovacího kmitočtu. Výsledné spektrum navzorkovaného signálu představuje spektrum původního signálu, které se periodicky opakuje se vzorkovací frekvencí f_{vz} [1].



Obr. 1.2: Periodický sled impulzů s určitou vzorkovací frekvencí. Na ose x je čas vyjádřený pořadím impulzů.



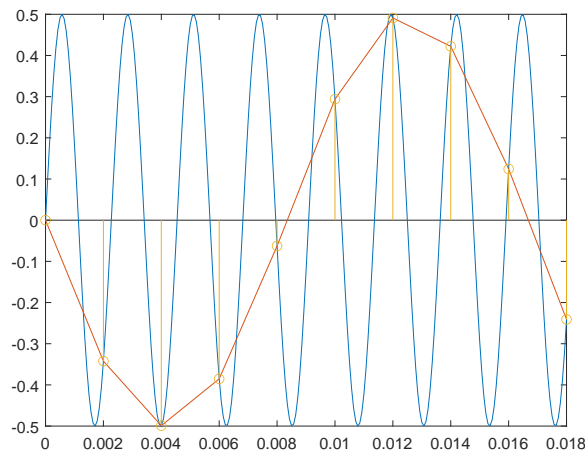
Obr. 1.3: Navzorkovaný harmonický signál vzniklý násobením signálů na obrázcích 1.1 a 1.2. Na ose x je pořadí vzorků signálu s určitou vzorkovací frekvencí v tomto případě 20 vzorků na periodu.

Aliasing

Při nevhodném zvolení vzorkovací frekvence může dojít k překrývání jednotlivých period spektra a tím ke zkreslení původního signálu. Takovému zkreslení se říká aliasing, který lze snadno reprezentovat i v časové doméně signálu. Aby k takovému zkreslení nedošlo, musí se dodržet vzorkovací poučka známá například jako Nyquistův–Shannonův teorém, ze kterého vyplývá, že vzorkovací frekvence musí být nejméně dvojnásobná než maximální frekvence vzorkovaného signálu.

$$f_{vz} > 2f_{max} \quad (1.1)$$

Jelikož většina analogových signálů není pásmově omezená, používají se v praxi



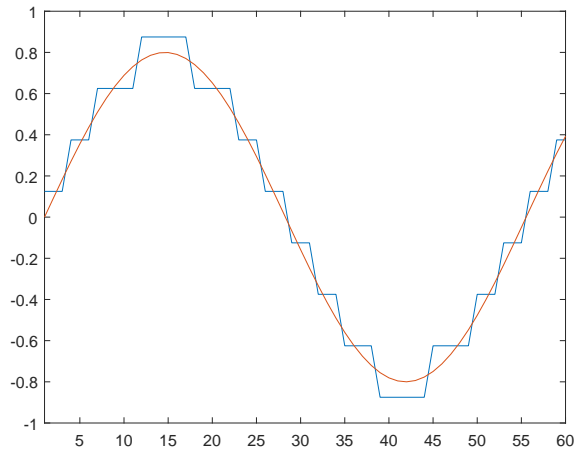
Obr. 1.4: Aliasing v časové doméně signálu. Modře původní signál o frekvenci 440 Hz, výsledek vzorkování o nevhodné vzorkovací frekvenci je propojen červeně.

filtry typu dolní propust tak, aby před vzorkováním odstranily vysoké kmitočty, které nevyhovují vzorkovací poučce. Chybná reprezentace signálu způsobená nevhodným vzorkovacím kmitočtem je na obrázku 1.4.

1.1.2 Kvantizace

Hodnoty analogového signálu odpovídají množině reálných čísel. To znamená, že nejmenší rozdíl přiřazený nejbližším hodnotám analogového signálu je teoreticky nekonečně malý. Kvantizace vyjadřuje hodnoty analogového signálu hodnotami, které jsou vyjádřeny čísly s konečnou délkou. Ty si můžeme představit jako kvantizační hladiny, hodnoty původního signálu jsou nahrazeny nejbližšími kvantizačními hladinami. Podle rozložení kvantizačních hladin se kvantizace rozděluje na uniformní, to je když jsou kvantizační hladiny rovnoměrně rozloženy a neuniformní, když rovnoměrně rozloženy nejsou. Tato práce se bude dále zabývat jenom kvantizací uniformní. Hladiny jsou reprezentovány binárními čísly. Kvantizace používá následující parametry jako je počet bitů, kterými se hodnoty signálu kódují B , kvantizační krok Q určuje interval, při kterém jsou nejbližší hodnoty signálu přiřazeny určité kvantizační hladině, nebo rozsah R . Například kvantizace signálu napětí s rozsahem od -1 do 1 voltu, který se kóduje 3 bity, je rozdělen rovnoměrně na 2^3 různých kvantizačních hladin vzdálených od sebe stejnou hodnotou jako je kvantizační krok $2/2^3$, obrázek 1.5. Pro zmíněné parametry tedy platí rovnice

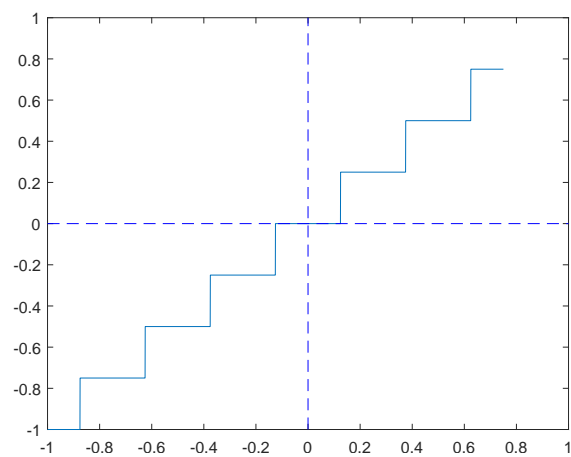
$$Q = \frac{R}{2^B}. \quad (1.2)$$



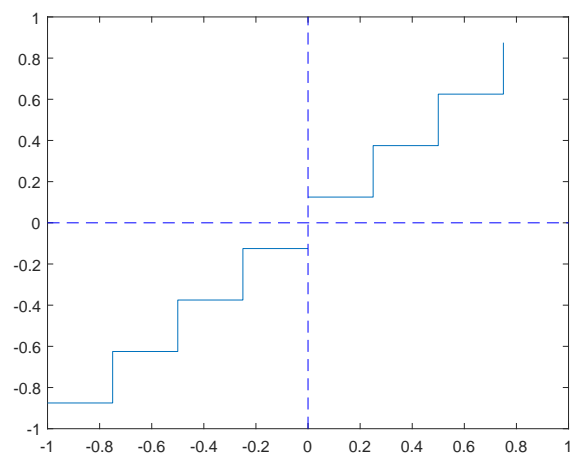
Obr. 1.5: Kvantizace signálu, v příslušném rozsahu má signál pouze osm kvantizačních hladin.

Mid-tread a Mid-riser

V praxi se používají dva typy kvantizace mid-tread a mid-riser podle toho, kde se nacházejí kvantizační hladiny [1]. Při kvantizaci mid-tread se kvantizační hladiny nacházejí uprostřed kvantizačního kroku. To znamená, že nula je jednou z kvantizačních hladin. Tento typ kvantizace se častěji používá u audia právě proto, že dokáže úplné ticho vyjádřit nulami. Naopak u kvantizace mid-riser není nula kvantizační hladina, ale rozhodovací úroveň mezi dvěma nejbližšími kvantizačními hladinami. Převodové funkce u těchto typů kvantizace jsou na obrázcích 1.6 a 1.7.



Obr. 1.6: Převodová funkce u kvantizace mid-tread.



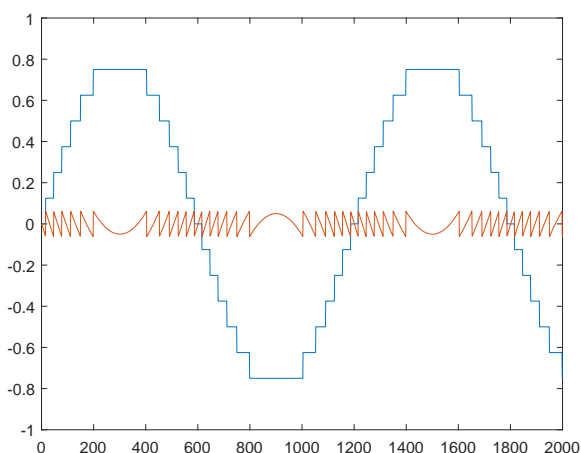
Obr. 1.7: Převodová funkce u kvantizace mid-riser.

Kvantizační chyba

Při kvantizaci se dopouštíme nepřesnosti v reprezentaci původního signálu. Toho si lze všimnout i v převodové funkci kvantizace, převodová funkce přesné reprezentace původního signálu by měla tvar přímky. Rozdíl původního signálu a kvantovaného signálu se říká kvantizační chyba nebo kvantizační šum. V následující rovnici je kvantizační chyba e vyjádřena, T značí vzorkovací periodu a n pořadí vzorku, jde tedy o vzorkované signály, horní index Q říká, že jde o nakvantovaný signál.

$$e(nT) = x^Q(nT) - x(nT) \quad (1.3)$$

Kvantizační chyba je zobrazená spolu s kvantizovaným signálem na obrázku 1.8.



Obr. 1.8: Kvantizační chyba zobrazená spolu s nakvantovaným signálem.

Snížení amplitudy kvantizační chyby lze docílit zvýšením počtu bitů převodu, které způsobí zmenšení kvantizačního kroku Q . Je totiž patrné, že kvantizační chyba nebude nabývat větší hodnoty než je polovina kvantizačního kroku [5].

$$-\frac{Q}{2} \leq e \leq \frac{Q}{2} \quad (1.4)$$

V audio signálech je kvantizační chyba nežádoucí, jelikož zní nepříjemně lidskému uchu. Její potlačení je motivací také pro dekvantizaci, která z kvantovaného signálu odhaduje původní signál před kvantizací.

Dithering

Pro signály s vysokou amplitudou se amplituda kvantizační chyby zdá být náhodná a jeví se podobně jako šum, dokonce je v praxi často maskována v rámci kritického pásma. Naopak pro signály s nižší amplitudou se kvantizační chyba více podobá kvantovanému signálu a jeví se spíše jako zkreslení. Na místo navýšením bitového

čísla se tento problém může řešit přidáním nízkourovňového šumu před kvantováním, který způsobí, že kvantizační chyba více odpovídá přidanému šumu než původnímu signálu a jeví se opět jako šum [1]. Této technice se říká dithering. Ač dithering přidá signálu slyšitelný šum, výrazně potlačí zkreslení slyšitelné u kvantizační chyby, což je pro poslech audia přijatelnější. Proces výrazně závisí na úrovni přidaného šumu a jeho rozdělení pravděpodobnosti. V audiu se nejčastěji používají tři typy rozdělení pravděpodobnosti, které jsou gaussovské, obdélníkové a trojúhelníkové.

1.1.3 Analogově digitální převodníky

V této kapitole se budeme zabývat samotnými A/D převodníky, které A/D převod provádějí. Bude zmíněno několik základních typů převodníků, které jsou zmíněny ve zdroji [1] a jejich důležité parametry.

Paralelní A/D převodníky

Jsou nejjednodušším typem A/D převodníků. Každý kvantizační interval má svůj komparátor. Hranice napětí pro každý kvantizační interval je dána řetězem rezistorů připojených k referenčnímu napětí. Referenční napětí je jedním vstupem pro každý napěťový komparátor, druhým vstupem je analogový signál. Komparátor můžeme vnímat jako jednobitový A/D převodník. Jelikož je pro každý kvantizační interval potřeba jeden komparátor, pro 8bitový převod je potřeba 255 komparátorů. Pro 16bitový převod by bylo potřeba 65 535 komparátorů, z toho důvodu se nehodí pro přímý převod audia.

Převodníky s postupnou aproximací

Další metoda A/D převodu vyžadující menší komplexnosti obvodu. Každý bit je testovaný zvlášť. Začíná se od nejvýznamnějšího bitu. Je-li úroveň vstupního signálu vyšší než polovina rozsahu, je nejvýznamnější bit zachován a použije se jako základ pro testování dalšího bitu. Ten je zachován v případě, že vstupní napětí převyší tři čtvrtiny rozsahu a tak dále.

Převodníky $\Sigma\Delta$

Převodníky, které mají v amplitudě dosahovat většího rozlišení používají techniky jako je oversampling a noise-shaping. Oversampling znamená, že je použita daleko větší vzorkovací frekvence než je Nyquistův kmitočet, což má za následek rozprostření kvantizačního šumu ve spektru, které má mnohonásobně větší šířku pásma, pásmo přenášené. To znamená, že odstup signálu od šumu, který je dvakrát nadvzorkovaný klesne o 6 dB, pro signál, který je nadvzorkovaný čtyřikrát o 12 dB a tak

dále. Což vytváří možnost použití převodníků s menšíbitovou hloubkou, jelikož šum nebude slyšitelný. Výstupní nadvzorkovaný signál se následně filtruje digitálním filtrem, který nepotřebné frekvence, vyšší než je Nyquistův kmitočet, spolu s částí šumu odstraní. Spolu s nadvzorkováním se používá také tvarování šumu (noise shaping), které díky zpětné vazbě potlačuje šum na nižších frekvencích. Šum ve spektru potom není rovnoměrně rozložený a filtrací se odstraní daleko větší část šumu [1].

Převodníky $\Sigma\Delta$ těchto technik využívají. Jejich realizace bývá rozdělena do dvou hlavních částí, které jsou modulátor $\Sigma\Delta$, ve kterém dochází k jednobitovému převodu vstupního signálu za mnohonásobně vyšší vzorkovací frekvence než je Nyquistův kmitočet, dále obsahuje zpětnou vazbu. Tato část slouží k tvarování šumu. Další část převodníku je digitální decimační filtr, který slouží k převzorkování signálu, kde se odstraní vysoké kmitočty spolu se šumem. Tyto převodníky dosahují vyšší kvality převodu a jsou vhodné pro převod audia.

1.2 Dekvantizace

Dekvantizace je jedním z procesů, při kterém se restauruje digitální signál. Máme-li digitální signál, úkolem dekvantizace je co nej přesněji odhadnout vzorky signálu před kvantizací, tedy neomezené délkou slova, které vzorky signálu vyjadřuje. Motivací pro dekvantizaci je potlačení nežádoucích vlastností kvantovaného signálu, jako je kvantizační šum. Při uniformní kvantizaci mid-riser můžeme napsat podle [2] rovnici pro kvantovaný signál.

$$(\mathbf{x}^Q)_n = \text{sgn}^+(x_n) \cdot Q \cdot \left(\left\lfloor \frac{|x_n|}{Q} \right\rfloor - \frac{1}{2} \right) \quad (1.5)$$

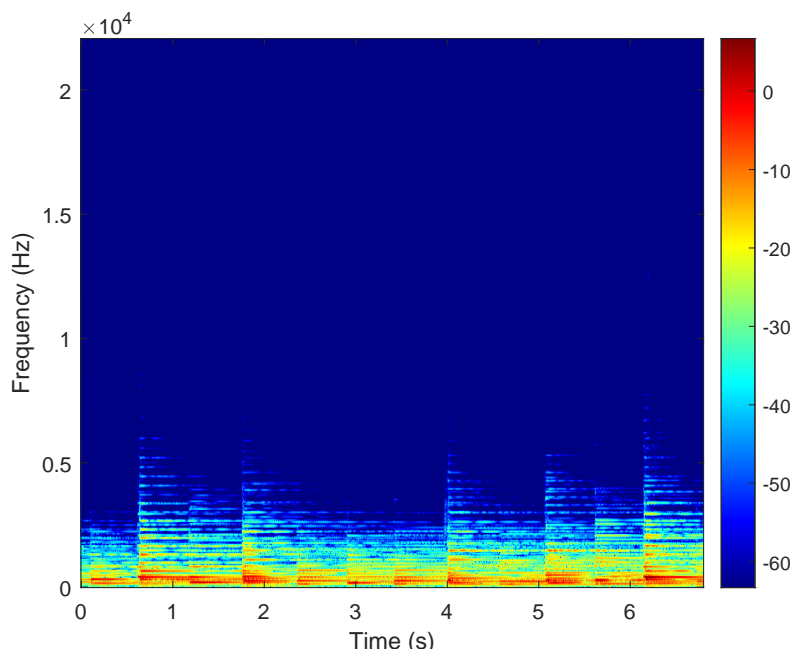
V této rovnici \mathbf{x}^Q značí kvantizovaný signál, \mathbf{x} je původní signál, n je pořadí vzorku a Q je kvantizační krok daný rozlišením kvantizace. Operace $\lfloor \cdot \rfloor$ znamená *floor*, což znamená zaokrouhlení dolů. Zpětně když známe kvantizovaný signál můžeme sestavit množinu přípustných hodnot, které lze přiřadit původnímu signálu. Takovou množinu budeme značit Γ a bude pro ni platit následující rovnice [2].

$$\Gamma = \{\mathbf{x} \in \mathbb{R}^P \mid \|\mathbf{x} - \mathbf{x}^Q\|_\infty < Q/2\} \quad (1.6)$$

P zde značí počet vzorků signálu. Γ je potom taková množina, kde každý vzorek původního signálu náleží množině reálných čísel a rozdíl všech vzorků původního signálu od příslušných vzorků kvantizovaného signálu je maximálně polovina kvantizačního kroku Q , $\|\cdot\|_\infty$ značí maximovou normu vracející absolutní hodnotu z vektoru rozdílů $\mathbf{x} - \mathbf{x}^Q$.

1.2.1 Řídká reprezentace signálů

Byla definována množina, na které se budou nacházet odhady původního signálu. Dále je také potřeba stanovit podle jakého kritéria je vhodný odhad původního signálu hledat. Podle [3] většina hudebních signálů má převážnou část koeficientů ve spektrogramu, který představuje časově-frekvenční reprezentaci signálu, zanedbatelnou a je možné je považovat za řídká řešení, proto je vhodné hledat takové signály, které jsou ve spektrogramech řídké. Hledáme tedy takové signály \mathbf{x} , které



Obr. 1.9: Spektrogram nahrávky klavíru.

náleží množině Γ a které mají co nejřidší koeficienty v časově-frekvenční doméně [2].

$$\arg \min_{\mathbf{x} \in \mathbb{R}^P} \|\mathbf{A}\mathbf{x}\|_0 \quad \text{vzhledem k } \mathbf{x} \in \Gamma \quad (1.7)$$

V této minimalizační úloze \mathbf{A} představuje časově-frekvenční analýzu signálu \mathbf{x} a operace $\|\cdot\|_0$ je ℓ_0 -norma vracející počet nenulových hodnot. Použití ℓ_0 -normy dělá z úlohy NP problém [2] a proto se místo ní používá ℓ_1 -norma, relaxovaná řídkost, vracející součet absolutních hodnot všech prvků množiny, která z celého problému dělá problém konvexní a umožňuje použití konvexní optimalizace. Minimalizační problém můžeme potom přepsat na následující tvar

$$\arg \min_{\mathbf{x} \in \mathbb{R}^P} \|\mathbf{A}\mathbf{x}\|_1 \quad \text{vzhledem k } \mathbf{x} \in \Gamma. \quad (1.8)$$

Takto zapsaný problém řeší například Douglas–Rachfordův nebo Chambolle–Pockův algoritmus.

1.3 Konvexní optimalizace

Konvexní optimalizace se obecně zabývá minimalizací konvexní funkce na konvexní množině. Konvexní funkce u spojitých funkcí je taková funkce, jejíž druhá derivace je nezáporná. Konvexnost funkce je definována v rovnici (1.9). Konvexní množina je taková množina, kde spojnice dvou libovolných bodů uvnitř množiny je opět uvnitř množiny. Funkce $f : \mathbb{R}^n \rightarrow \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$ tehdy když

$$\forall \mathbf{x}_1, \mathbf{x}_2 \in X, \forall \lambda \in (0, 1) : f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2). \quad (1.9)$$

Funkce f je ryze konvexní tehdy když

$$\forall \mathbf{x}_1, \mathbf{x}_2 \in X, \forall \lambda \in (0, 1) : f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) < \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_2). \quad (1.10)$$

Obecná úloha konvexní optimalizace může vypadat takto:

$$\arg \min_{\mathbf{x}} f(\mathbf{x}) \quad \text{vzhledem k } \mathbf{x} \in X. \quad (1.11)$$

Pro konvexní úlohy platí:

- Každé lokální minimum f na X je globální.
- Množina všech optimálních argumentů funkce f je konvexní.
- Je-li f ryze konvexní, existuje maximálně jedno optimální řešení.

Optimalizační úlohy můžeme dále rozdělit na ty ve tvaru omezeném a neomezeném. Obecná úloha (1.11) i úlohy z předchozí kapitoly jsou úlohy v omezeném tvaru. Takovou úlohu lze převést na neomezený tvar pomocí indikátorové funkce [3]. Indikátorová funkce pro úlohu (1.11) lze zapsat jako:

$$i_X : \mathbf{x} \rightarrow \begin{cases} 0 & \text{pro } \mathbf{x} \in X \\ \infty & \text{jinak.} \end{cases} \quad (1.12)$$

S využitím indikátorové funkce se původní úloha přepíše do neomezeného tvaru:

$$\arg \min_{\mathbf{x}} f(\mathbf{x}) + i_X(\mathbf{x}). \quad (1.13)$$

1.3.1 Proximální algoritmy

Proximální algoritmy jsou postupy pro nalezení řešení optimalizačních úloh. Jsou to iterativní algoritmy, které slouží k minimalizaci součtu funkcí pomocí postupného vyhodnocování jejich gradientů nebo proximálních operátorů [6, 3]. Proximální algoritmy řeší celou škálu konvexních úloh, v práci bude dále rozebrán obecný proximální algoritmus pro konvexní optimalizaci.

Proximální operátor

Mějme funkci $f : \mathbb{R}^n \rightarrow \mathbb{R}$, která je konvexní, zdola polospojité s neprázdným definičním oborem. Jednoznačné řešení následného minimalizačního problému budeme nazývat proximální operátor funkce f a bude se značit prox_f [6].

$$\arg \min_{\mathbf{x}' \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{x} - \mathbf{x}'\|_2^2 + f(\mathbf{x}') \quad (1.14)$$

Řešení problému jsou odhady \mathbf{x}' , které nejsou příliš daleko od známého signálu \mathbf{x} a které jsou dále regularizovány s ohledem na funkci f . Ta je konvexní funkcí vynucující některé vlastnosti jako například relaxovanou řídkost $f(\mathbf{y}) = \|\mathbf{y}\|_1$. Je vhodné poznamenat, že proximální operátor indikátorové funkce je projekcí \mathbf{x}' na konvexní množinu například X , vůči které je indikátorová funkce definována [6] a bude se značit proj_X .

$$\text{prox}_{i_X}(\mathbf{x}) = \text{proj}_X(\mathbf{x}') \quad (1.15)$$

Další případ, který poznamenáme je proximální operátor pro nulovou funkci. Dosadíme-li do rovnice (1.14) dostaneme

$$\arg \min_{\mathbf{x}' \in \mathbb{R}^n} \frac{1}{2} \|\mathbf{x} - \mathbf{x}'\|_2^2 + 0. \quad (1.16)$$

Z takto zapsaného problému je patrné, že řešení minimalizačního problému je $\mathbf{x}' = \mathbf{x}$. Proximální operátor nějaké funkce f , která je nulovou funkcí je tedy

$$\text{prox}_0 = \text{Id}. \quad (1.17)$$

Proximální operátor ℓ_1 -normy

S ohledem na hledané řídké reprezentace signálu je vhodné uvést proximální operátor pro relaxovanou řídkost $f(\mathbf{y}) = \lambda \|\mathbf{y}\|_1$, λ je regularizační skalár ovlivňující váhu funkce v rámci minimalizačního problému. Hledaný proximální operátor můžeme zapsat jako

$$\text{prox}_{\lambda \|\cdot\|_1}(\mathbf{x}) = \arg \min_{\mathbf{x}'} \frac{1}{2} \|\mathbf{x} - \mathbf{x}'\|_2^2 + \lambda \|\mathbf{x}'\|_1. \quad (1.18)$$

Výsledkem je

$$x'_i = \frac{x_i}{|x_i|} \max(|x_i| - \lambda, 0). \quad (1.19)$$

Výsledná funkce se nazývá měkké prahování a provádí se pro každou složku zvlášť. Bude se značit $\mathbf{x}' = \text{soft}_\lambda(\mathbf{x})$ [3].

1.3.2 Condatův algoritmus

V této části bude popsán Condatův obecný proximální algoritmus pro konvexní optimalizaci, viz [6]. Tento algoritmus je zobecněním některých klasických proximálních algoritmů řešící dekvantizaci na základě řídké reprezentace signálů, jako je například Chambolle–Pock nebo Douglas–Rachford algoritmus uvedený v [2]. Tento algoritmus předpokládá možnost, že se v minimalizačním problému objevuje součet více konvexních funkcí, které mohou a nemusí být hladké a které mohou a nemusí být složeny s lineárním operátorem. Algoritmus pracuje na základě *splitting* rozdělení operací, což znamená, že proximální operátory nebo subgradienty jednotlivých funkcí vyhodnocuje v každé iteraci zvlášť. Obecně minimalizační problém definuje jako:

$$\arg \min_{\mathbf{x} \in X} f(\mathbf{x}) + g(\mathbf{x}) + \sum_{m=1}^M h_m(L_m \mathbf{x}). \quad (1.20)$$

Kde $M \in \mathbb{N}$, $X = \mathbb{R}^n$ a $U_m = \mathbb{R}^{p_m}$, přičemž p_m odpovídá dimenzi, tak že $L_m \mathbf{x} \in U_m$. Dále:

- $f, g : X \rightarrow \mathbb{R}, h_m : U_m \rightarrow \mathbb{R}$.
- Operátory $L_m : X \rightarrow U_m$ jsou lineární a ohraničené.
- f je diferencovatelná na X a její gradient je β -Lipschitz spojitý pro reálnou konstantu $\beta > 0$.

Řešení minimalizačního problému (1.20) řeší algoritmus 1. V algoritmu se objevuje

Algoritmus 1: Condatův algoritmus

Zvolit parametry $\tau > 0$, $\sigma > 0$, $\rho > 0$

a počáteční odhady $\mathbf{x}^{(0)} \in X$, $\mathbf{u}_1^{(0)} \in U_1, \dots, \mathbf{u}_M^{(0)} \in U_M$.

for $i = 0, 1, \dots$ **do**

$\tilde{\mathbf{x}}^{(i+1)} := \text{prox}_{\tau g}(\mathbf{x}^{(i)} - \tau \nabla f(\mathbf{x}^{(i)}) - \tau \sum_{m=1}^M L_m^* \mathbf{u}_m^{(i)}),$

$\mathbf{x}^{(i+1)} := \rho \tilde{\mathbf{x}}^{(i+1)} + (1 - \rho) \mathbf{x}^{(i)},$

for $m = 1, \dots, M$, **do**

$\tilde{\mathbf{u}}_m^{(i+1)} := \text{prox}_{\sigma h_m^*}(\mathbf{u}_m^{(i)} + \sigma L_m(2\tilde{\mathbf{x}}^{(i+1)} - \mathbf{x}^{(i)})),$

$\mathbf{u}_m^{(i+1)} := \rho \tilde{\mathbf{u}}_m^{(i+1)} + (1 - \rho) \mathbf{u}_m^{(i)}.$

h_m^* , což je konvexní konjugát funkce h_m , který se v algoritmu objevuje spolu s *Moreau* identitou:

$$\text{prox}_h(\mathbf{x}) + \text{prox}_{h^*}(\mathbf{x}) = \mathbf{x} \quad (1.21)$$

V algoritmu spolu s konstantou σ vypadá výsledný $\text{prox}_{\sigma h_m^*}$ takto:

$$\text{prox}_{\sigma h_m^*}(\mathbf{u}) = \mathbf{u} - \sigma \text{prox}_{h_m/\sigma}(\mathbf{u}/\sigma). \quad (1.22)$$

Dále v algoritmu vystupují L_m^* , které můžeme chápat jako transponované matice k příslušným lineárním operátorům L_m .

Podmínky konvergence

Předpokládejme, že parametry v algoritmu splňují:

- $\tau \left(\frac{\beta}{2} + \sigma \left\| \sum_{m=1}^M L_m^* L_m \right\| \right) < 1$, kde $\|\cdot\|$ je operátorová norma a β je Lipschitz konstanta.
- $\rho \in]0, 1]$.

Potom obě sekvence $(\tilde{\mathbf{x}}^{(i)})$ a $(\mathbf{x}^{(i)})$ konvergují k řešení minimalizačního problému $\hat{\mathbf{x}} \in X$.

Je-li funkce f nulová, množiny X a U_m mají konečnou dimenzi a parametry algoritmu splňují

- $\tau \sigma \left\| \sum_{m=1}^M L_m^* L_m \right\| \leq 1$,
- $\rho \in]0, 2[$,

potom sekvence $(\tilde{\mathbf{x}}^{(i)})$ a $(\mathbf{x}^{(i)})$ opět konvergují k řešení $\hat{\mathbf{x}} \in X$.

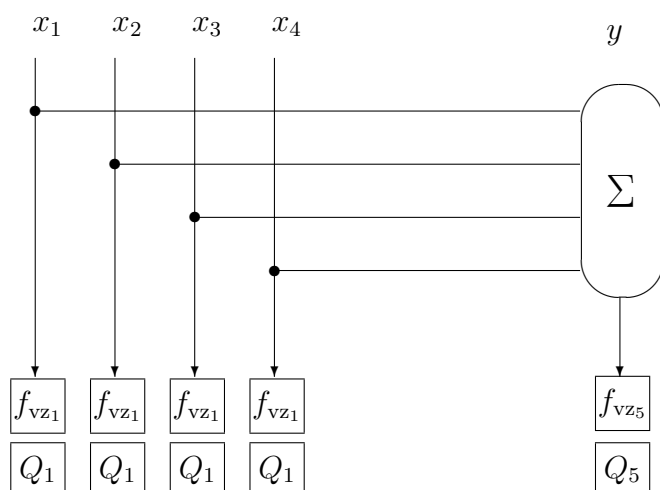
2 Výsledky studentské práce

V této části se budeme zabývat konkrétním problémem, který odpovídá zadání této práce. Dále budou nastíněny metody pro jeho řešení.

2.1 Formulace problematiky

Otázka shrnující zadání této práce je následující: Jak vysoké kvality audio signálu je možné dosáhnout pomocí řady A/D převodníků s nízkou bitovou hloubkou a nízkým vzorkovacím kmitočtem v kombinaci s jedním převodníkem s vysokou bitovou hloubkou a vysokým vzorkovacím kmitočtem?

Mějme tedy několik analogových audio signálů, které budou podrobeny A/D převodu v řadě levnějších A/D převodníků a potom jejich součet, který bude podroben převodu s kvalitnější A/D převodníkem podobně jako na následujícím schématu.



Zde $x_1 \dots x_4$, představují původní analogové signály, y je součet původních signálů, sekce f_{vz} je navzorkování původních signálů s příslušnou vzorkovací frekvencí. Sekce f_{vz_5} by mohlo být vzorkování s vyšším vzorkovacím kmitočtem. Tato práce ale dále pracuje se stejným vzorkovacím kmitočtem pro jednotlivé signály i jejich součet. Q je kvantizace s indexem 1 v horší kvalitě a 5 v lepší kvalitě, který převádí součet vstupních signálů y .

Nastíněné řešení se podrobně zabývá dekvantizací a předpokládá, že počty vzorků jednotlivých signálů a jejich součtu si odpovídají.

Pro následnou rekonstrukci původních čtyř signálů po naznačené kvantizaci zde vznikají jasné podmínky zmíněné v části 1.2. V rovnicích budeme značit x jako původní signály, \mathbf{x}^Q jako původní signály po kvantizaci a \mathbf{x}' jako odhady signálů,

tedy výsledek rekonstrukce.

$$\begin{aligned}
Q_1(\mathbf{x}'_1) &= \mathbf{x}_1^Q & \text{to je } \mathbf{x}'_1 &\in \Gamma_1, \\
Q_1(\mathbf{x}'_2) &= \mathbf{x}_2^Q & \text{to je } \mathbf{x}'_2 &\in \Gamma_2, \\
Q_1(\mathbf{x}'_3) &= \mathbf{x}_3^Q & \text{to je } \mathbf{x}'_3 &\in \Gamma_3, \\
Q_1(\mathbf{x}'_4) &= \mathbf{x}_4^Q & \text{to je } \mathbf{x}'_4 &\in \Gamma_4.
\end{aligned} \tag{2.1}$$

Pro jednotlivé Γ platí rovnice (1.6) podle vlastností kvantizace u A/D převodu.

Dále pro součet jednotlivých odhadů původních signálů musí platit:

$$Q_5(\mathbf{x}'_1 + \mathbf{x}'_2 + \mathbf{x}'_3 + \mathbf{x}'_4) = y^Q \quad \text{to je } \mathbf{x}'_1 + \mathbf{x}'_2 + \mathbf{x}'_3 + \mathbf{x}'_4 \in \Gamma_5. \tag{2.2}$$

Dalším krokem je vybrat kritérium, podle kterého chceme rekonstrukci signálů provádět, což v tomto případě může být řídká reprezentace signálů. To znamená, že budeme hledat takové signály, jejichž spektrální koeficienty jsou co nejvíce řídké a které vyhovují podmínkám zapsaných v rovnicích (2.1) a (2.2). To lze zapsat jako následující minimalizační problém.

$$\arg \min_{\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3, \mathbf{x}'_4 \in \mathbb{R}^P} \lambda_1 \|\mathbf{A}\mathbf{x}'_1\|_1 + \lambda_2 \|\mathbf{A}\mathbf{x}'_2\|_1 + \lambda_3 \|\mathbf{A}\mathbf{x}'_3\|_1 + \lambda_4 \|\mathbf{A}\mathbf{x}'_4\|_1 \tag{2.3}$$

vzhledem k $\mathbf{x}'_1 \in \Gamma_2, \mathbf{x}'_2 \in \Gamma_1, \mathbf{x}'_3 \in \Gamma_3, \mathbf{x}'_4 \in \Gamma_4, \mathbf{x}'_1 + \mathbf{x}'_2 + \mathbf{x}'_3 + \mathbf{x}'_4 \in \Gamma_5$

Parametry $\lambda_1 \dots, \lambda_4$ jsou váhovací koeficienty ℓ_1 -norm. Předpokládáme tedy, že vstupní signály rozlišujeme podle řídkosti. Výběr váhovacích koeficientů by měl probíhat tak, aby měly ℓ_1 -normy spektrálních koeficientů více řídkých signálů větší váhu. Dále je možné podmínky přepsat pomocí indikátorových funkcí jako v rovnici (1.13).

$$\begin{aligned}
&\arg \min_{\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3, \mathbf{x}'_4 \in \mathbb{R}^P} \lambda_1 \|\mathbf{A}\mathbf{x}'_1\|_1 + \lambda_2 \|\mathbf{A}\mathbf{x}'_2\|_1 + \lambda_3 \|\mathbf{A}\mathbf{x}'_3\|_1 + \lambda_4 \|\mathbf{A}\mathbf{x}'_4\|_1 + \\
&\quad + \iota_{\Gamma_1}(\mathbf{x}'_1) + \iota_{\Gamma_2}(\mathbf{x}'_2) + \iota_{\Gamma_3}(\mathbf{x}'_3) + \iota_{\Gamma_4}(\mathbf{x}'_4) + \iota_{\Gamma_5}(\mathbf{x}'_1 + \mathbf{x}'_2 + \mathbf{x}'_3 + \mathbf{x}'_4)
\end{aligned} \tag{2.4}$$

Argumenty minimalizačních problémů odpovídají dimenzi P , představující délku jednotlivých signálů před rekonstrukcí.

2.2 Navrhnuté řešení

Pro nalezení řešení minimalizačního problému zmíněného v sekci 2.1 je vhodné použít Condatův algoritmus popsany v sekci 1.3.2. Aby úloha odpovídala tvaru, ve kterém algoritmus hledá řešení, je potřeba hledané argumenty přepsat do jednoho vektoru, který budeme značit \mathbf{x}' a bude pro něj platit:

$$\mathbf{x}' = (\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3, \mathbf{x}'_4)^T. \tag{2.5}$$

Dále zavedeme maticové operátory, které budou z vektoru \mathbf{x}' vybírat jednotlivé úseky $\mathbf{x}'_1 \dots \mathbf{x}'_4$, které budeme značit $B_1 \dots B_4$, pro které platí:

$$\begin{aligned} B_1 \mathbf{x}' &= \mathbf{x}'_1 \\ B_2 \mathbf{x}' &= \mathbf{x}'_2 \\ B_3 \mathbf{x}' &= \mathbf{x}'_3 \\ B_4 \mathbf{x}' &= \mathbf{x}'_4. \end{aligned} \tag{2.6}$$

Z rovnice (2.6) vyplývá, že matice B_n budou mít rozměr $P \times 4P$ a budou vypadat tak, že na sloupcích $(n-1)P + 1$ až nP budou mít na diagonále samé 1 a zbytek prvků v matici bude nulových. Zjednodušený příklad pro jednotlivé signály s délkou vzorků $P = 2$ bude vypadat takto

$$\begin{aligned} B_1 &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ B_2 &= \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\ B_3 &= \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \\ B_4 &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.7}$$

Nyní lze minimalizační problém (2.4) přepsat do tvaru

$$\begin{aligned} \arg \min_{\mathbf{x}' \in \mathbb{R}^{4P}} \lambda_1 \|AB_1 \mathbf{x}'\|_1 + \lambda_2 \|AB_2 \mathbf{x}'\|_1 + \lambda_3 \|AB_3 \mathbf{x}'\|_1 + \lambda_4 \|AB_4 \mathbf{x}'\|_1 + \\ + \iota_{\Gamma_1}(B_1 \mathbf{x}') + \iota_{\Gamma_2}(B_2 \mathbf{x}') + \iota_{\Gamma_3}(B_3 \mathbf{x}') + \iota_{\Gamma_4}(B_4 \mathbf{x}') + \iota_{\Gamma_5}((B_1 + B_2 + B_3 + B_4) \mathbf{x}'). \end{aligned} \tag{2.8}$$

Nyní přiřadíme jednotlivé funkce z rovnice (2.8) funkcím vystupujícím v rovnici (1.20). V tomto případě $f(\mathbf{x}')$ a $g(\mathbf{x}')$ bude 0. Zbývající funkce tedy budou $h_m(L_m \mathbf{x}')$, které následně rozepíšeme.

$$\begin{aligned} h_1 &= \lambda_1 \|\cdot\|_1, h_2 = \lambda_2 \|\cdot\|_1, \\ h_3 &= \lambda_3 \|\cdot\|_1, h_4 = \lambda_4 \|\cdot\|_1, \\ h_5 &= \iota_{\Gamma_1}, \\ h_6 &= \iota_{\Gamma_2}, h_7 = \iota_{\Gamma_3}, \\ h_8 &= \iota_{\Gamma_4}, h_9 = \iota_{\Gamma_5}. \end{aligned} \tag{2.9}$$

Lineární operátory L_m definujeme jako

$$\begin{aligned}
L_1 &= AB_1, & L_2 &= AB_2, \\
L_3 &= AB_3, & L_4 &= AB_4, \\
L_5 &= B_1, & L_6 &= B_2, \\
L_7 &= B_3, & L_8 &= B_4, \\
L_9 &= B_1 + B_2 + B_3 + B_4.
\end{aligned} \tag{2.10}$$

Indexy m funkcí L nabývají hodnot $1 \dots 9$. A je lineární operátor představující časově frekvenční analýzu signálu.

Dále v algoritmu vystupují proximální operátory funkcí h_m , které jsou v případě indikátorových funkcí projekce do příslušných množin a v případě ℓ_1 -normy měkké prahování jako v části 1.3.1. A transponované matice L_m^* , které vypadají takto:

$$\begin{aligned}
L_1^* &= B_1^* A^*, & L_2^* &= B_2^* A^*, \\
L_3^* &= B_3^* A^*, & L_4^* &= B_4^* A^*, \\
L_5^* &= B_1^*, & L_6^* &= B_2^*, \\
L_7^* &= B_3^*, & L_8^* &= B_4^*, \\
L_9^* &= (B_1 + B_2 + B_3 + B_4)^*.
\end{aligned} \tag{2.11}$$

A^* představuje syntézu signálu z příslušných koeficientů spektrogramu. Matice B^* jsou potom transponované matice původních B_m , které doplňují signál nulami do příslušné délky $4P$.

V algoritmu dále vystupují konstanty, které podléhají podmínkám konvergence, a počáteční odhady, které pro $\mathbf{u}_1 \dots \mathbf{u}_4$ rozměrově odpovídají koeficientům spektrogramu, zbytek odhadů \mathbf{u} odpovídají délce segmentu například \mathbf{x}'_1 a počáteční odhad $\mathbf{x}^{(0)}$ odpovídají velikosti hledanému odhadu signálů \mathbf{x}' .

Podmínky konvergence

Jak už bylo zmíněno, aby odhady v jednotlivých iteracích konvergovaly k hledanému řešení minimalizačního problému \hat{x} , musí parametry τ , σ a ρ vyhovovat podmínkám konvergence, které byly změněny v sekci 1.3.2, pro případ kdy je funkce f nulová.

$$\begin{aligned}
\tau \sigma \left\| \sum_{m=1}^M L_m^* L_m \right\| &\leq 1, \\
\rho &\in]0, 2[.
\end{aligned} \tag{2.12}$$

První podmínku rozepíšeme s ohledem na operátory L_m definované v rovnici (2.10).

$$\tau \sigma \left\| \sum_{m=1}^4 B_m^* A^* A B_m + \sum_{m=1}^4 B_m^* B_m + (B_1 + B_2 + B_3 + B_4)^* (B_1 + B_2 + B_3 + B_4) \right\| \leq 1. \tag{2.13}$$

Pro vypsane členy platí určité vlastnosti. Předpokládáme, že výsledek složeného operátoru A^*A , což představuje analýzu a následnou syntézu signálu je matice identity Id , za použití *tight frames* [3]. Maticové součiny $B_m^*B_m$ pro představu vypíšeme pro $P = 2$.

$$B_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$B_1^* = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (2.14)$$

Součin matic $B_1^*B_1$ vypadá takto.

$$B_1^*B_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.15)$$

Součin dalších matic $B_2^*B_2$ bude vypadat takto:

$$B_2^*B_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.16)$$

a analogicky budou vypadat součiny zbývajících matic.

Z matic je patrné, že výsledek součtu $B_m^*B_m$ dá pouze jedničky na hlavní diagonále, tedy identitu matice Id . Rovnici (2.13) nyní přepíšeme do tvaru:

$$\tau\sigma\|\text{Id} + \text{Id} + (B_1 + B_2 + B_3 + B_4)^*(B_1 + B_2 + B_3 + B_4)\| \leq 1. \quad (2.17)$$

Součin posledních dvou závorek dá matici identity stejně jako ve dvou prvních případech. Původní podmínka se tedy přepíše do tvarů:

$$\begin{aligned}\tau\sigma\|\text{Id} + \text{Id} + \text{Id}\| &\leq 1 \\ \tau\sigma 3\|\text{Id}\| &\leq 1.\end{aligned}\tag{2.18}$$

Maticová norma $\|\cdot\|$ z matice identity vrací hodnotu 1. Podmínka pro parametry τ a σ bude tedy vypadat takto:

$$\tau\sigma \leq \frac{1}{3}.\tag{2.19}$$

Nyní známe všechny členy objevující se v algoritmu 1 a můžeme jej použít pro konstrukci signálů $x_1 \dots x_4$. Algoritmus tedy přepíšeme do konkrétnější podoby. Rozměry počátečních odhadů $\mathbf{x}^{(0)}$ budou $4P$, $\mathbf{u}_1^{(0)} \dots \mathbf{u}_4^{(0)}$ budou rozměrově odpovídat koeficientům spektrogramu signálu o délce P . Zbývající počáteční odhady $\mathbf{u}_5^{(0)} \dots \mathbf{u}_9^{(0)}$ mají rozměr P . V prvním řádku algoritmu 1 se objevuje proximální operátor funkce g , která je v tomto případě nulová. Proximální operátor vyřešíme stejně jako v rovnici (1.17). Dále se v prvním řádku objevuje gradient funkce f , která je taky nulová, proto i gradient bude nabývat nulové hodnoty. Argumenty proximálních operátorů konvexních konjugátů funkcí h , definované v rovnici (1.22), které vystupují na konci algoritmu budeme substituovat jako \mathbf{p}_m následovně:

$$\mathbf{p}_m = \mathbf{u}_m^{(i)} + \sigma L_m(2\tilde{\mathbf{x}}^{(i+1)} - \mathbf{x}^{(i)}).\tag{2.20}$$

Tedy pro $m = 1, \dots, 4$ L_m jsou AB_m , pro $m = 5, \dots, 8$ jsou B_{m-4} , a L_9 je $B_1 + B_2 + B_3 + B_4$ jako v rovnici (2.10). Konkrétní podoba Condatova algoritmu je v algoritmu 2. Zde se objevují odhady \mathbf{x} a $\tilde{\mathbf{x}}$, které odpovídají výsledkům rekonstrukce dříve značeny jako \mathbf{x}' .

Kvůli zjednodušení algoritmu jsou odhady $\mathbf{u}_5, \dots, \mathbf{u}_8$ zapsané v jedné proměnné \mathbf{u}_s

$$\mathbf{u}_s = (\mathbf{u}_5, \mathbf{u}_6, \mathbf{u}_7, \mathbf{u}_8)^T.\tag{2.21}$$

A rovnice (2.8) se přepíše takto.

$$\begin{aligned}\arg \min_{\mathbf{x}' \in \mathbb{R}^{4P}} \lambda_1 \|AB_1 \mathbf{x}'\|_1 + \lambda_2 \|AB_2 \mathbf{x}'\|_1 + \lambda_3 \|AB_3 \mathbf{x}'\|_1 + \lambda_4 \|AB_4 \mathbf{x}'\|_1 + \\ + \iota_{\Gamma_s}(\mathbf{x}') + \iota_{\Gamma_5}((B_1 + B_2 + B_3 + B_4)\mathbf{x}').\end{aligned}\tag{2.22}$$

Kde Γ_s je množina přípustných hodnot pro signál o délce $4P$, tvořený po sobě jdoucími množinami $\Gamma_1, \dots, \Gamma_4$.

$$\Gamma_s = (\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4)^T.\tag{2.23}$$

Stejným způsobem jsou v algoritmu zapsané i argumenty proximálních operátorů $\mathbf{p}_5, \dots, \mathbf{p}_8$.

$$\mathbf{p}_s = (\mathbf{p}_5, \mathbf{p}_6, \mathbf{p}_7, \mathbf{p}_8)^T.\tag{2.24}$$

Algoritmus 2: Condatův algoritmus v konkrétní podobě pro řešení optimalizace (2.22)

Zvolit parametry $\tau > 0$, $\sigma > 0$, $\rho > 0$

a počáteční odhady $\mathbf{x}^{(0)} \in X$, $\mathbf{u}_1^{(0)} \in U_1, \dots, \mathbf{u}_m^{(0)} \in U_m$.

for $i = 0, 1, \dots$ **do**

$\tilde{\mathbf{x}}^{(i+1)} :=$

$\mathbf{x}^{(i)} - \tau \sum_{m=1}^4 B_m^* A^* \mathbf{u}_m^{(i)} - \tau \sum_{m=1}^4 B_m^* \mathbf{u}_{m+4}^{(i)} - \tau (B_1 + B_2 + B_3 + B_4)^* \mathbf{u}_9^{(i)},$

$\mathbf{x}^{(i+1)} := \rho \tilde{\mathbf{x}}^{(i+1)} + (1 - \rho) \mathbf{x}^{(i)},$

for $m = 1, \dots, 4$, **do**

$\tilde{\mathbf{u}}_m^{(i+1)} := \mathbf{p}_m - \text{soft}_{\lambda_m}(\mathbf{p}_m),$

$\mathbf{u}_m^{(i+1)} := \rho \tilde{\mathbf{u}}_m^{(i+1)} + (1 - \rho) \mathbf{u}_m^{(i)},$

$\tilde{\mathbf{u}}_s^{(i+1)} := \mathbf{p}_s - \sigma \text{proj}_{\Gamma_s}(\mathbf{p}_s / \sigma),$

$\mathbf{u}_s^{(i+1)} := \rho \tilde{\mathbf{u}}_s^{(i+1)} + (1 - \rho) \mathbf{u}_s^{(i)}.$

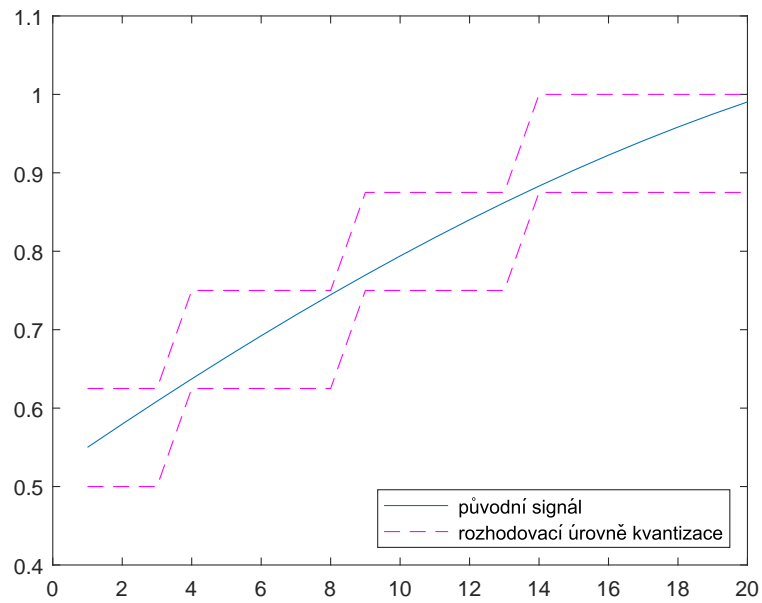
$\tilde{\mathbf{u}}_9^{(i+1)} := \mathbf{p}_9 - \sigma \text{proj}_{\Gamma_5}(\mathbf{p}_9 / \sigma),$

$\mathbf{u}_9^{(i+1)} := \rho \tilde{\mathbf{u}}_9^{(i+1)} + (1 - \rho) \mathbf{u}_9^{(i)}.$

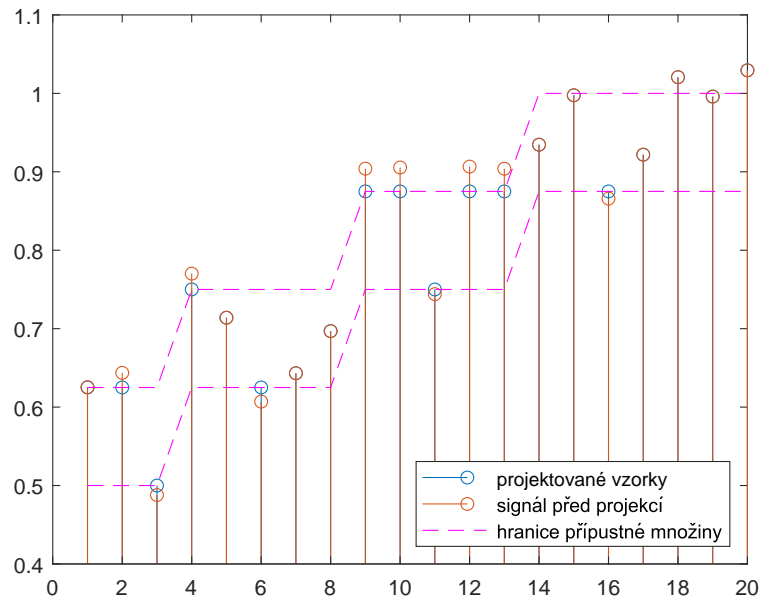
Projekce

Stěžejní částí algoritmu je zmíněná projekce. Ta jako proximální operátor indikátorových funkcí upravuje vstupy \mathbf{p}_s a \mathbf{p}_9 do příslušných přípustných množin Γ_s a Γ_9 . Projekce pracuje s každým vzorkem zvlášť tak, že hodnoty, které se nachází v přípustné množině ponechá stejně. Hodnoty nespadaající do přípustné množiny posune vždy na její hranici, která je blíže původní hodnotě vstupu. Hraniční případ projekce je, když je vzorek nad nebo pod úrovní příslušné Γ a množina přípustných hodnot se zároveň nachází mezi nejvyšší nebo nejnižší kvantovací hladinou. Takové vstupní hodnoty projekce ponechává stejné, protože původní signál před kvantizací mohl takových hodnot nabývat. Příklady kvantizace původního signálů a následné projekce jsou vidět na obrázku 2.1 a 2.2. Rovnice (2.25) popisuje fungování projekce, kde se k popisu množiny Γ využívá příslušný kvantizační krok Q jako v rovnici (1.6) a kvantovaného signálu \mathbf{x}^Q .

$$\text{proj}_{\Gamma} : \mathbf{x} \rightarrow \begin{cases} x_i^Q + \frac{Q}{2} & \text{pro } \left(x_i - x_i^Q > \frac{Q}{2}\right) \wedge \left(x_i^Q < 1 - \frac{Q}{2}\right) \\ x_i^Q - \frac{Q}{2} & \text{pro } \left(x_i^Q - x_i > \frac{Q}{2}\right) \wedge \left(x_i^Q > \frac{Q}{2} - 1\right) \\ x_i & \text{jinak} \end{cases} \quad (2.25)$$



Obr. 2.1: Rozhodovací úrovně kvantizace



Obr. 2.2: Projekce signálu na množinu přípustných hodnot. Vzorky, u kterých je jenom jedno kolečko, je projekce identická se vstupním signálem.

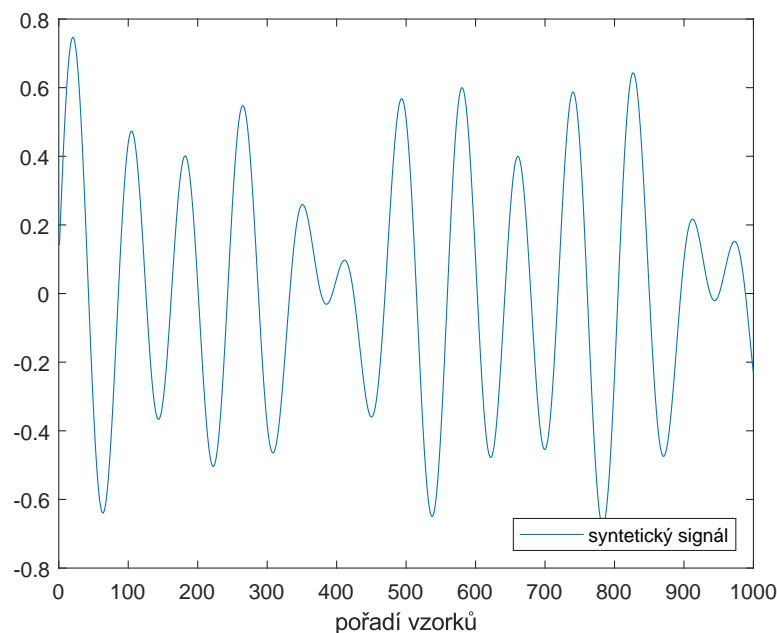
2.3 Řešení a výsledky

2.3.1 Implementace v prostředí MATLAB

Nastíněný algoritmus pro řešení formulovaného problému byl implementovaný v prostředí MATLAB R2023a. Byl vytvořen jeden hlavní script, ve kterém se nastavují vstupy a další parametry algoritmu a několik dalších scriptů, které se z hlavního scriptu volají. V této části je hlavní script po úsecích popsán.

Načtení vstupních signálů

Jako vstupní signály se načítají nahrávky z vytvořené databáze, která obsahuje nahrávky hudebních nástrojů, které jsou různě řídké. Nahrávky mají bitovou hloubku 24 bitů a vzorkovací kmitočet 48 kHz. Zaznamenané hudební nástroje jsou housle, cello, klavír a baskytara. Další možností vstupních signálů jsou syntetické signály, které se generují v připraveném scriptu. Generované signály jsou směsy harmonických signálů o různých fázích a kmitočtech. Příklad části syntetického signálu je na obrázku 2.3.



Obr. 2.3: Úsek syntetického signálu

Kvantizace

V další části kódu se načtené vstupní signály zkvantizují. Nejprve každý signál zvlášť s nižší bitovou hloubkou a potom jejich součet s vyšší bitovou hloubkou. Kvantizace

typu mid-riser probíhá v samostatném scriptu, který předpokládá, že vstupní signál je v rozsahu -1 až 1 . Script provádějící kvantizaci je převzatý ze zdroje [2].

Parametry časově-frekvenční analýzy a syntézy signálů

V této části kódu se vybírají parametry uplatňující se při diskrétní Gaborově transformaci, která představuje časově-frekvenční spektrum. Časově-frekvenční operace jsou prováděny v toolboxu LTFAT. Bylo použito Hannovo okno s délkou 2048 vzorků s 25% překryvem a se 4096 frekvenčními kanály, což je případ, který odpovídá těsnému framu viz [3].

Parametry dekvantizace

Před zavoláním scriptu, který provádí samotnou dekvantizaci je ještě nutné zvolit příslušné parametry popsané v algoritmu 2. Hodnoty parametrů souhlasí s podmínkami konvergence (2.18). Z pohledu rychlosti konvergence v závislosti na počtu iterací algoritmu se ukázaly optimální hodnoty $\rho = 0,8$, $\tau = \frac{1}{9}$ a $\sigma = 3$. Před samotnou dekvantizací se ještě nastavuje počet iterací, který má algoritmus provést a váhy $\lambda_1, \dots, \lambda_4$ pro jednotlivé signály. Použité počáteční odhady $\mathbf{x}^{(0)}$ a $\mathbf{u}_m^{(0)}$ jsou nulové hodnoty o příslušných rozměrech X nebo U_m .

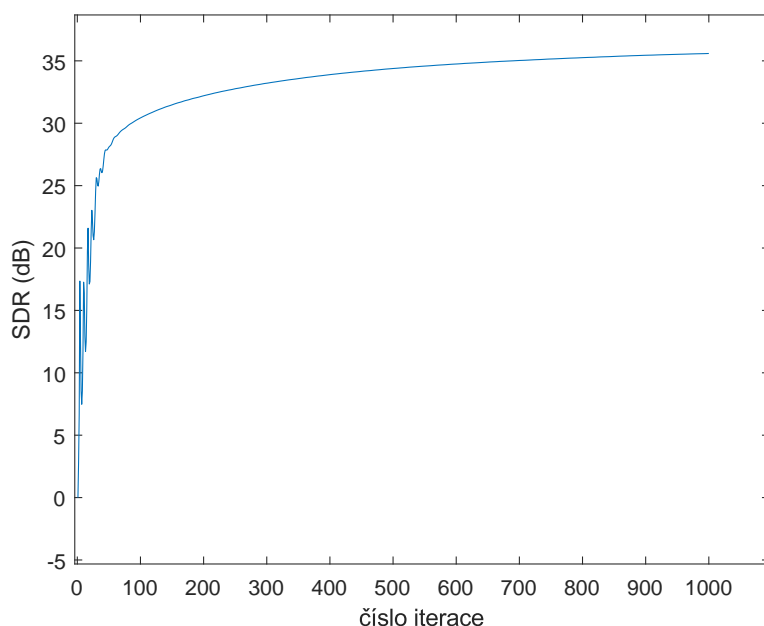
Vyhodnocení rekonstrukce

Po získání rekonstrukcí jednotlivých signálů se vyhodnocuje jejich úspěšnost na základě jejich podobnosti s původními signály ve 24bitové kvalitě pro reálné signály a ve vyšší kvalitě pro signály syntetické. Podobnost se měří pomocí SDR (signal to distortion ratio), vyjadřující míru zkreslení signálu, které je popsané rovnicí (2.26), kde \mathbf{x} vyjadřuje původní signál před kvantizací, tedy ve vyšší bitové hloubce a \mathbf{x}' je příslušná rekonstrukce. Dále se měřila podobnost na základě PEMO-Q ODG (objective difference grade), která při hodnocení podobnosti signálů uvažují poslechové vlastnosti signálů [7].

$$\text{SDR}(\mathbf{x}, \mathbf{x}') = 10 \log_{10} \frac{\|\mathbf{x}\|^2}{\|\mathbf{x} - \mathbf{x}'\|^2} \quad (2.26)$$

2.3.2 Výsledky

Dekvantizace tedy probíhala pro různé signály kvantované na bitovou hloubku 8, 10 nebo 12 bitů. Jako první byly použity čtyři signály s nahrávkami cello a houslí. Průběh SDR se zvyšujícím se počtem opakování algoritmu byl obdobný pro všechny použité bitové hloubky. A pro bitovou hloubku 8 bitů je vyobrazený na obrázku 2.4.

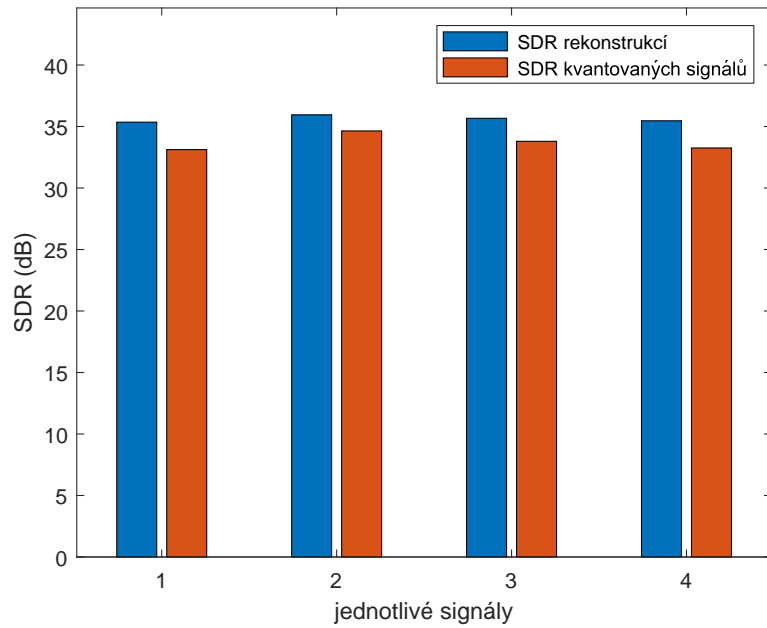


Obr. 2.4: SDR v závislosti na počtu iterací pro nahrávky smyčců

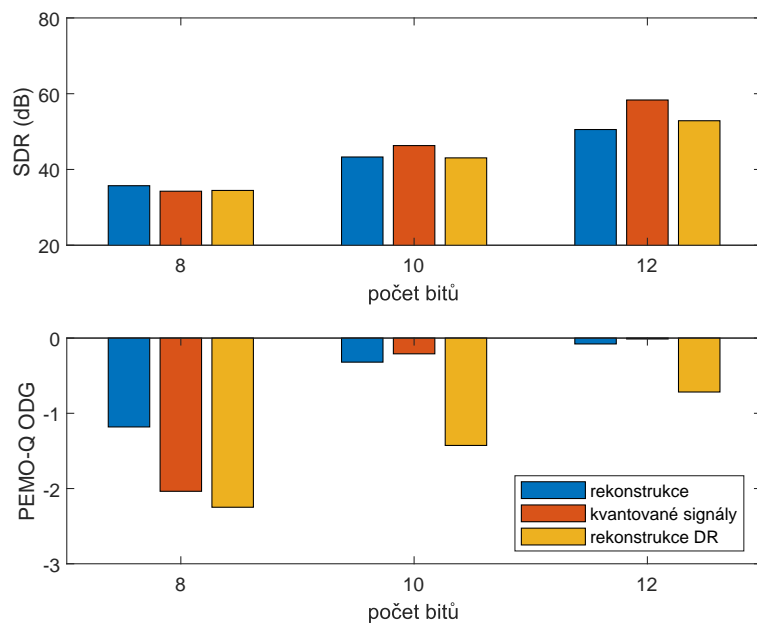
Signály byly zkráceny na 70 tisíc vzorků, což odpovídá asi sekundě a půl. První dva signály představovali dvoje různé housle, další dva signály potom dvě různá cello. Nahrávky cello byly spektrálně méně řídké. Proto měly menší váhu, $\lambda_3 = \lambda_4 = 0,8$, zatímco nahrávky houslí měly váhu $\lambda_1 = \lambda_2 = 1$. Na obrázku 2.5 je vidět SDR rekonstruovaných a kvantovaných signálů s bitovou hloubkou 8 bitů po 1000 iteracích. Pro všechny signály je SDR rekonstrukcí o jednotky dB vyšší než SDR kvantovaných signálů.

Srovnání průměru SDR a ODG pro čtyři různé nahrávky nástrojů s měnící se bitovou hloubkou kvantovaného signálu je na obrázku 2.6. Se zvyšující se bitovou hloubkou kvantovaného signálu se zvyšuje i SDR rekonstruovaného signálu. Při bitové hloubce 10 bitů už ale není vyšší než SDR příslušného kvantovaného signálu. Stejně tak signály dosahují lepšího ODG. Od bitové hloubky 10 bitů dosahují kvantované signály lepších hodnot než rekonstrukce. Podobných výsledků dosahovaly i rekonstrukce v práci [2] za použití algoritmů neuvažující součet signálů před A/D převodem s vyšší bitovou hloubkou. Na obrázku je i srovnání stejné rekonstrukce za použití Douglas–Rachfordova algoritmu převzatého právě z [2], který dosahuje podobných hodnot SDR jako algoritmus navržený v této práci, v ODG má však horší výsledky. Z tohoto zjištění lze usoudit, že rekonstrukce navrženým algoritmem dosahují lepších výsledků než rekonstrukce prováděné bez znalosti součtu signálů ve větší kvalitě.

Další serie testování proběhla na syntetických signálech, které byly všechny stejně

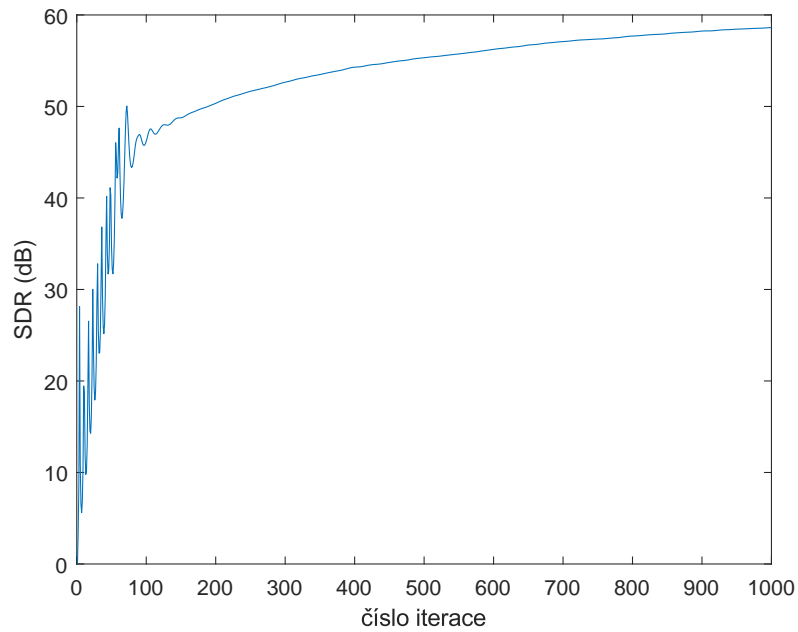


Obr. 2.5: SDR pro bitovou hloubku 8 bitů pro nahrávky smyčců



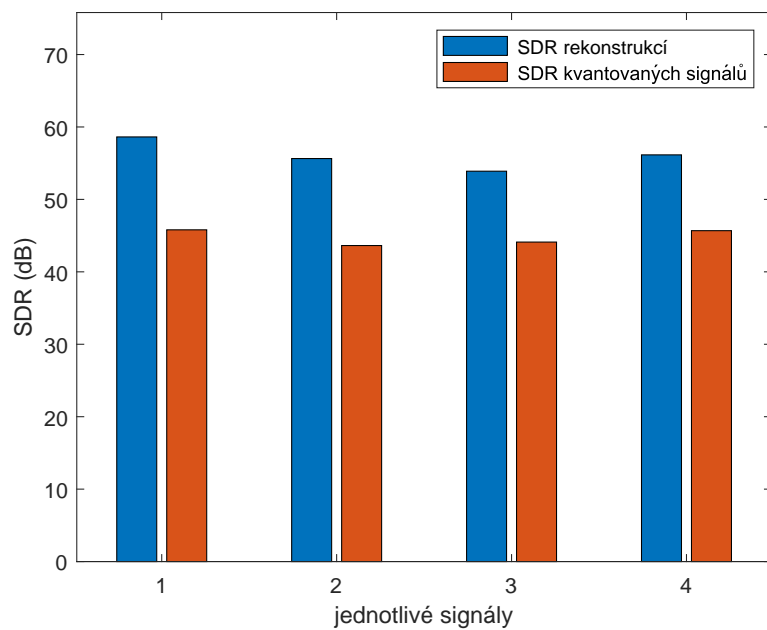
Obr. 2.6: průměrné SDR a ODG pro různou bitovou hloubku pro nahrávky smyčců

řidké. Váhovací koeficienty λ tedy měly všechny hodnotu 1. Vyobrazené SDR v závislosti na počtu iterací je na obrázku 2.7. Zde jde vidět, že rekonstrukce dosahují daleko lepších výsledků. To si vysvětluji především větší řídkostí použitých signálů.



Obr. 2.7: SDR v závislosti na počtu iterací pro syntetické signály

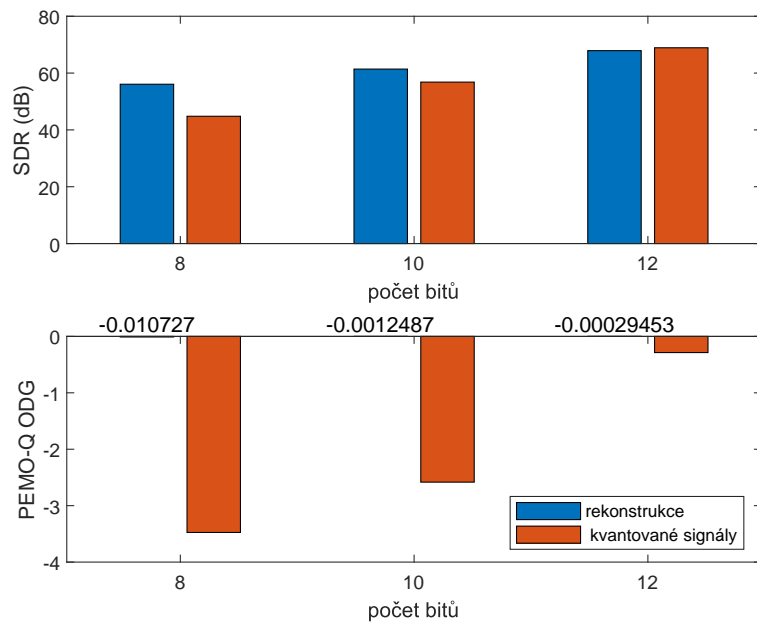
Na obrázku 2.8 je potom znázorněno SDR pro čtyři různé signály po 1000 iteracích algoritmu se signály kvantovanými 8 bity. Zde dosahuje SDR asi o 10 dB lep-



Obr. 2.8: SDR pro bitovou hloubku 8 bitů pro syntetické signály

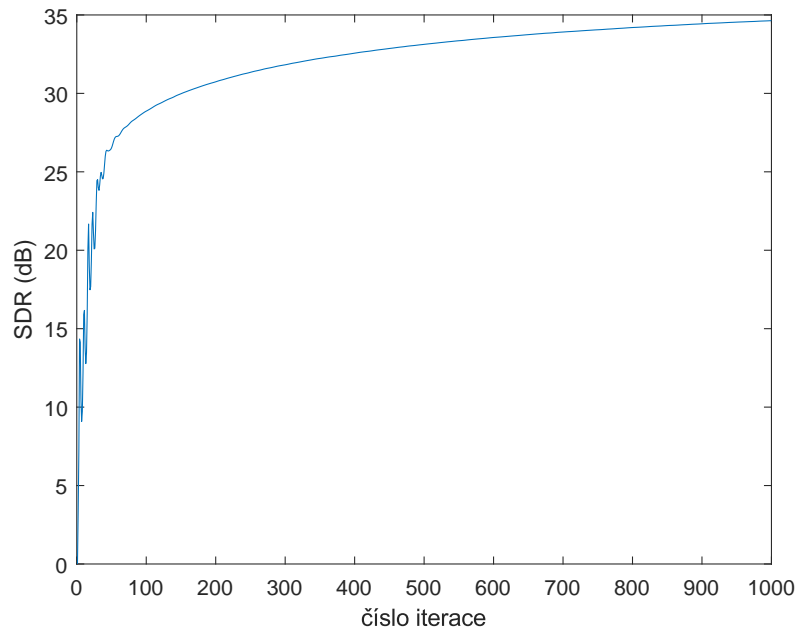
ších hodnot než kvantované signály. Obrázek 2.9 srovnává SDR a ODG syntetických

signálů pro různou bitovou hloubku. SDR nedosahuje horších výsledků než u kvan-



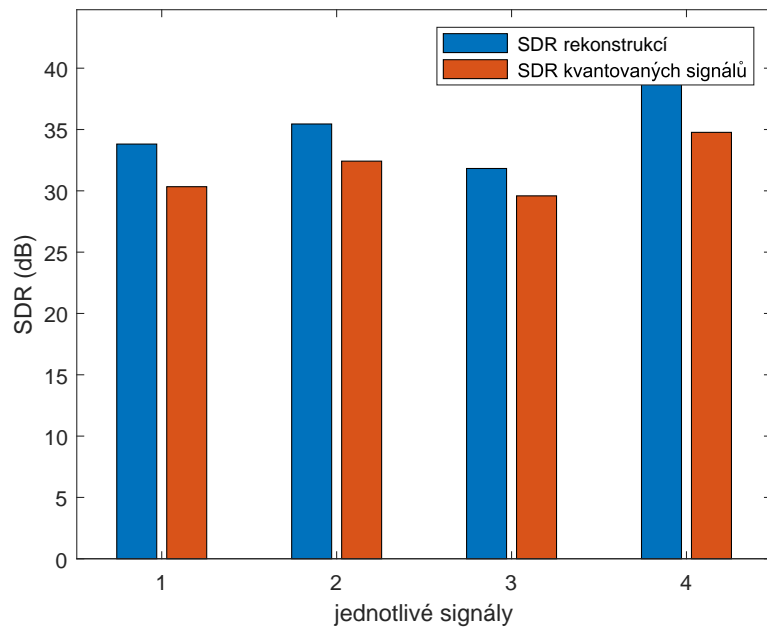
Obr. 2.9: Průměrné SDR a ODG pro různou bitovou hloubku pro syntetické signály, ve spodním grafu je ODG rekonstrukcí tak blízké nule, že je pod hranicí grafického rozlišení.

tovaného signálu. ODG při různé bitové hloubce je ale vždy výrazně lepší.

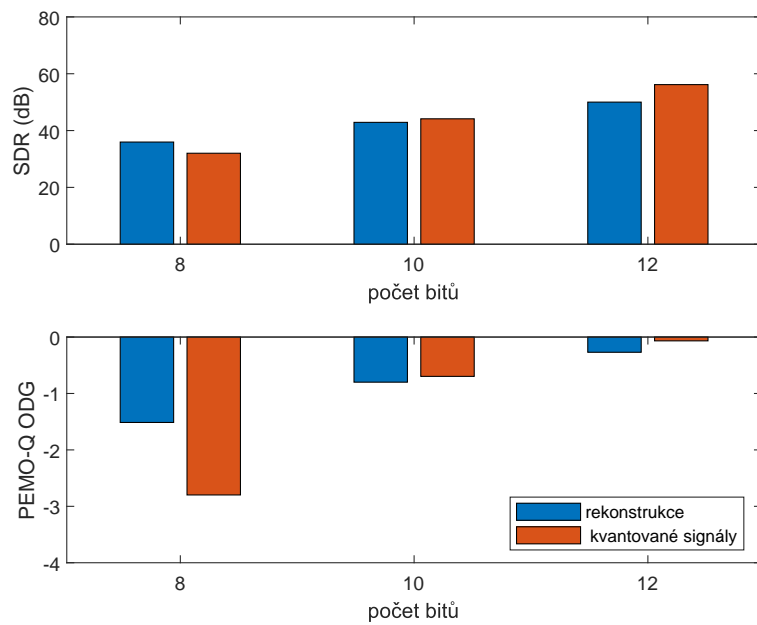


Obr. 2.10: SDR v závislosti na počtu iterací pro klavír

Poslední testovací serie obsahovala housle, cello, klavír a baskytaru jako ve spektru rozdílnější nástroje. Druhá nahrávka v pořadí je klavír, který dostal váhovací koeficient $\lambda_2 = 0,6$ jakožto nejméně řídký signál. Zbylé váhovací koeficienty λ_1 , λ_3 a λ_4 mají hodnotu 1. A na následujících obrázcích 2.10 2.11 a 2.12 je jako v předchozích případech analogicky vidět průběh SDR v průběhu iterací, tentokrát pro klavír. SDR při bitové hloubce 8 bitů a SDR a ODG pro různou bitovou hloubku. Pro všechny předchozí případy trvá 1000 iterací algoritmu pro 70 tisíc vřezků asi 220 sekund. Zajímavý je i subjektivní dojem při poslechu rekonstruovaných a kvantovaných nahrávek především u reálných nahrávek. Při poslechu signálu kvantovaného na bitovou hloubku 8 bitů je šum slšitelný, zatímco u rekonstrukcí není žádný šum postřehnutelný. Lepší ODG u rekonstrukcí tedy odpovídá i subjektivnímu poslechu nahrávek. Při poslechu signálu kvantovaného na bitovou hloubku 10 bitů je šum také ještě postřehnutelný, zatímco u rekonstrukcí nikoliv. Při bitové hloubce 10 bitů ale ODG u kvantovaných signálů dosahuje mírně lepších hodnot, které nesouhlasí se subjektivním hodnocením kvality nahrávek. Pro kvantované signály s bitovou hloubkou 12 bitů hodnotím rozdíl rekonstrukcí a kvantovaných signálů jako nepostřehnutelný.



Obr. 2.11: SDR pro bitovou hloubku 8 bitů pro různé nástroje



Obr. 2.12: SDR a ODG pro různou bitovou hloubku pro různé nástroje

Závěr

Pro kombinaci několika A/D převodníků s horší kvalitou a jedním převodníkem s lepší kvalitou byl formulován optimalizační problém, na základě něhož je možné rekonstruovat signály převedené v horší kvalitě. Navrhnuté řešení pro takto formulovanou konvexní úlohu je použití Condatova obecného proximálního algoritmu, který je v práci představen a jsou v ní rozepsány všechny kroky potřebné k jeho použití.

Tímto způsobem bylo v práci otestováno několik různých reálných i syntetických nahrávek. Výsledky dosažené popsáním algoritmem se podobají výsledkům navyšování bitové hloubky jako ve zdroji [2], signály dokonce dosahují lepších výsledků, které se vyhodnocovaly na základě SDR a PEMO-Q, v závislosti na bitové hloubce použitých kvantovaných signálů.

Výsledný algoritmus pracuje se stejnými vzorkovacími kmitočty pro všechny A/D převodníky, i když zadání práce nastiňovalo možnost použití různých vzorkovacích kmitočtů pro A/D převodníky různé kvality. To by mohlo vést ke zlepšení výsledků rekonstrukcí a práce by tímto směrem mohla pokračovat.

Literatura

- [1] Watkinson, John. The art of digital audio. 3rd ed. Amsterdam: Elsevier, c2001. ISBN 0-240-51587-0.
- [2] Závíška, P., Rajmic, P., Mokrý, O. Audio Dequantization Using (Co)Sparse (Non)Convex Methods. In Proceedings IEEE ICASSP, Canada, June 2021.
- [3] Rajmic, Pavel. Řídké a nízkohodnostní reprezentace signálů s aplikacemi: Sparse and low-rank representations of signals with applications : zkrácená verze habilitační práce. Brno: VUTIUM, 2015. ISBN 978-80-214-5195-7.
- [4] Mokrý, O., Rajmic, P., Závíška, P. Flexible framework for audio restoration. In Proceedings of the 23rd International Conference on Digital Audio Effects (DAFx) 2020.
- [5] Orfanidis, Sophocles J. Introduction to signal processing. Englewood Cliffs, N.J.: Prentice Hall, c1996. ISBN 0132091720.
- [6] L. Condat, A generic proximal algorithm for convex optimization—application to total variation minimization, IEEE Signal Processing Letters, 2014.
- [7] R. Huber and B. Kollmeier, "PEMO-Q—A New Method for Objective Audio Quality Assessment Using a Model of Auditory Perception," in IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, no. 6, pp. 1902-1911, Nov. 2006, doi: 10.1109/TASL.2006.883259.