

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA STROJNÍHO INŽENÝRSTVÍ
ÚSTAV MATEMATIKY
FACULTY OF MECHANICAL ENGINEERING
INSTITUTE OF MATHEMATICS

DOPLŇOVÁNÍ CHYBĚJÍCÍCH DAT VE ZVUKOVÝCH SIGNÁLECH

AUDIO INPAINTING ALGORITHMS

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. HANA BARTLOVÁ

VEDOUCÍ PRÁCE
SUPERVISOR

doc. Mgr. PAVEL RAJMIC, Ph.D.

BRNO 2015

Vysoké učení technické v Brně, Fakulta strojního inženýrství

Ústav matematiky

Akademický rok: 2014/2015

ZADÁNÍ DIPLOMOVÉ PRÁCE

student(ka): Bc. Hana Bartlová

který/která studuje v **magisterském navazujícím studijním programu**

obor: **Matematické inženýrství (3901T021)**

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a kúšebním řádem VUT v Brně určuje následující téma diplomové práce:

Doplňování chybějících dat ve zvukových signálech

v anglickém jazyce:

Audio inpainting algorithms

Stručná charakteristika problematiky úkolu:

Signály z reálného světa se často skládají z relativně malého počtu stavebních bloků (ve vztahu vůči celkové délce signálu). Toto zjištění je základem pro moderní tzv. řídké reprezentace signálů. Metodami lineární algebry a optimalizace lze nejen takové řídké vyjádření pro audiosignály získat, ale také efektivně využít při jejich zpracování.

Práce je zaměřena na konkrétní aplikaci doplňování chybějících dat do audiosignálu (restaurace chybějících či silně poškozených úseků). Dnes již obvyklé metody (viz literaturu) by práce měla vylepšit pomocí konceptu tzv. strukturované řídkosti a nahrazení syntezujícího modelu analyzujícím, a rovněž zefektivnit pomocí zohlednění vlastností diskrétní Gaborovy transformace.

Cíle diplomové práce:

Nastudovat principy a algoritmy hledání tzv. řídkých reprezentací a komprimovaného snímání.

Ujasnit si principy stávajících metod doplňování informací do audiosignálů.

Definovat jejich přednosti a slabiny.

Navrhnout modifikace metody na základě strukturované řídkosti.

Navrhnout modifikace metody na základě analyzujícího modelu.

Implementovat v Matlabu a testovat.

Zrychlit výpočty použitím efektivního výběru Gaborových atomů.

Shrnout a interpretovat výsledky.

Seznam odborné literatury:

ELAD, M., Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, Springer, 2010.

HRBÁČEK, R.; RAJMIC, P.; VESELÝ, V.; ŠPIŘÍK, J. Řídké reprezentace signálů: Úvod do problematiky. Elektrorevue - Internetový časopis, 2011, roč. 2011, č. 50, s. 1-10. ISSN: 1213-1539.

BALAZS, P., DOERFLER, M. et al., Theory, Implementation and Application of Nonstationary Gabor Frames, J. Comput. Appl. Math., In Press. (2011)

ADLER, A. et al., A Constrained Matching Pursuit Approach to Audio Declipping. Acoustics, Speech and Signal Processing, IEEE International Conference on (ICASSP 2011), IEEE, Prague, Czech Republic, 2011.

SIEDENBURG, K., DÖRFLER, M.: Structured Sparsity for Audio Signals. Proceedings of the 14th International Conference on Digital Audio Effects, DAFX-11 September 19-23, 2011 Paris, France.

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

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2014/2015.

V Brně, dne 20.11.2014

L.S.

prof. RNDr. Josef Šlapal, CSc.
Ředitel ústavu

prof. RNDr. Miroslav Doupovec, CSc., dr. h. c.
Děkan fakulty

Abstrakt

Tato práce se zabývá doplňováním chybějících dat do zvukových signálů. Na úvod jsou shrnuty základní poznatky využívané dále v textu. Před samotnou aplikační částí je představena řídká reprezentace signálů a některé algoritmy jejího hledání. V kapitole věnované doplňování dat je formulován problém a krátce popsány a srovnány dosavadní metody řešení. Poté je uveden nejnovější přístup řešení využívající harmonickou strukturu signálů a provedeny experimenty. Na závěr je odvozen algoritmus pro zajištění maximální efektivity výpočtu.

Summary

This thesis deals with audio inpainting problem. Firstly, basic concepts are summarized. Then, sparse representation of signals is introduced along with several algorithms. In the main part dedicated to the audio inpainting, the problem is defined and actual methods are presented and compared. The newest approach using the harmonic structure of sound signals is then introduced, followed by several experiments and evaluation. Lastly, an algorithm ensuring the maximal computational efficiency is derived.

Klíčová slova

frames, řídká reprezentace signálů, proximální algoritmy, doplňování chybějících dat do zvukových signálů, strukturovaná řídkost

Keywords

frames, sparse representation of signals, proximal algorithms, audio inpainting, structured sparsity

BARTLOVÁ, H. *Doplňování chybějících dat ve zvukových signálech*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2015. 71 s. Vedoucí diplomové práce doc. Mgr. Pavel Rajmic, Ph.D.

Prohlašuji, že jsem svou diplomovou práci na téma *Doplňování chybějících dat ve zvukových signálech* zpracovala samostatně pod vedením doc. Mgr. Pavla Rajmice, Ph.D. a s použitím zdrojů uvedených v seznamu, jenž je součástí této práce.

Bc. Hana Bartlová

Děkuji všem, kteří mi pomohli při přípravě této práce, zejména svému vedoucímu doc. Mgr. Pavlu Rajmicovi, Ph.D. a Ing. Václavu Machovi za velmi přínosné připomínky.

Bc. Hana Bartlová

Obsah

1	Úvod	13
2	Základní pojmy	14
3	Báze a framy	20
3.1	Systém generátorů vektorového prostoru	20
3.2	Báze	20
3.3	Framy	21
3.4	Prostory nekonečné dimenze	24
3.5	Gaborovy framy	25
4	Řídké reprezentace signálů	27
4.1	Postačující podmínky pro jednoznačnost řešení	27
4.2	Algoritmy hledání řídkých reprezentací	28
4.2.1	ℓ_1 relaxace	29
4.2.2	Hladové (greedy) algoritmy	32
4.3	Proximální algoritmy	34
4.4	Strukturovaná řídkost	39
4.5	Komprimované snímání	40
5	Doplňování chybějících dat do zvukového signálu	43
5.1	Formulace problému	44
5.2	Dosavadní metody	44
5.2.1	AR modely	44
5.2.2	Neuronové sítě	46
5.2.3	Metody využívající řídkou reprezentaci	47
5.2.4	Výhody a nevýhody dosavadních metod	49
5.3	Využití strukturované řídkosti	49
6	Experiment	52
6.1	Maximalizace efektivity výpočtu	59
6.1.1	Značení	60
6.1.2	Odvození algoritmu	61
6.1.3	Modifikace pro strukturovanou řídkost	64
7	Závěr	65
	Literatura	66
	Seznam použitých zkratk a symbolů	69
A	Obsah CD	71

1. Úvod

Zpracování signálů zahrnuje mnoho oblastí matematiky – lineární algebru, funkcionální analýzu, optimalizaci, pravděpodobnost a řadu dalších. Moderní přístupy zpracování signálů využívají tzv. řídkou reprezentaci. To znamená, že signál je možné přesně vyjádřit (popř. dobře aproximovat) lineární kombinací pouze několika málo vektorů reprezentačního systému. Aplikací řídkých reprezentací je celá řada, např. odstraňování šumu, komprese signálu, doplňování chybějících úseků, komprimované snímání. . .

Počátky řídkých reprezentací signálů spadají do roku 1993, kdy vědci Mallat a Zhang prezentovali algoritmus Matching Pursuit [27]. V roce 1996 publikoval R. Tibshirani přístup využívající ℓ_1 -relaxaci (metoda LASSO) [35], dále článek [9] uvedl ekvivalentní metodu basis pursuit.

Tato práce se zabývá doplňováním chybějících dat ve zvukových signálech. Při přenosu signálu a jeho zpracování často dochází k jeho narušení. Jsou uvedeny dosavadní metody doplňování dat a testujeme nejnovější přístup užití řídkých reprezentací. Také je prezentován algoritmus pro zrychlení výpočtu.

V kapitole 2 je uveden přehled základních pojmů užívaných dále v textu. Kapitola 3 pojednává o teorii reprezentačních systémů. Poté je kapitola 4 věnována řídkým reprezentacím a některým algoritmům jejich hledání. Na konci této kapitoly jsou krátce shrnuty principy komprimovaného snímání. Následně se kapitola 5 zaměřuje na doplňování chybějících dat. Je zformulován problém a popsány různé metody jeho řešení. V kapitole 6 jsem provedla příslušné experimenty a zabývala se zrychlením výpočtů.

2. Základní pojmy

Nejdříve uvedeme základní pojmy používané dále v textu. Vektory a matice jsou značeny tučně (\mathbf{x} , \mathbf{A}). Číslo komplexně sdružené k číslu a označujeme a^* . Tzv. Hermitovskou transpozici matice značíme také symbolem $*$. \mathbf{A}^* je matice, která vznikla složením transpozice matice a komplexním sdružením každého jejího prvku. Platí $(\mathbf{A}^*\mathbf{B}^*) = \mathbf{B}^*\mathbf{A}^*$. Jádro lineárního zobrazení určeného maticí \mathbf{A} značíme $\ker(\mathbf{A})$.

Vektorový prostor, normy

Definice 1 *Nechť je \mathbb{V} neprázdná množina. Dále nechť:*

1. *Ke každým dvěma prvům $\mathbf{x}, \mathbf{y} \in \mathbb{V}$ je jednoznačně přiřazen prvek $\mathbf{z} \in \mathbb{V}$ nazývaný jejich součet a označovaný $\mathbf{x} + \mathbf{y}$, přičemž platí:*

(a) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.

(b) $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$.

(c) *Ve \mathbb{V} existuje takový prvek označený \mathbf{o} , že $\forall \mathbf{x} \in \mathbb{V}$*
 $\mathbf{x} + \mathbf{o} = \mathbf{x}$.

(d) *Ke $\forall \mathbf{x} \in \mathbb{V}$ existuje takový prvek označený $(-\mathbf{x})$, že*
 $\mathbf{x} + (-\mathbf{x}) = \mathbf{o}$.

2. *Ke $\forall \alpha$ ($\alpha \in \mathbb{R}$ nebo $\alpha \in \mathbb{C}$) a ke $\forall \mathbf{x} \in \mathbb{V}$ je jednoznačně přiřazen prvek $\alpha\mathbf{x} \in \mathbb{V}$, přičemž platí:*

(a) $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$.

(b) $1 \cdot \mathbf{x} = \mathbf{x}$.

(c) $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$.

(d) $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.

Pak množinu \mathbb{V} nazýváme vektorovým prostorem. Podle toho, zda pod čísla α, β, \dots rozumíme reálná či komplexní čísla, hovoříme o reálném či komplexním vektorovém prostoru. Prvky množiny \mathbb{V} nazýváme vektory, čísla α, β, \dots skaláry.

Definice 2 *Nechť \mathbb{V} je vektorový prostor a $f : \mathbb{V} \rightarrow \mathbb{R}^+$ je funkce splňující $\forall \mathbf{x} \in \mathbb{V}$:*

1. $f(\mathbf{x}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{o}$,

2. $f(\mathbf{x} + \mathbf{y}) \leq f(\mathbf{x}) + f(\mathbf{y})$,

3. $f(\alpha\mathbf{x}) = |\alpha| f(\mathbf{x})$.

Pak se funkce f nazývá normou ve vektorovém prostoru \mathbb{V} a tento prostor je normovaný vektorový prostor. Normu prvku $\mathbf{x} \in \mathbb{V}$ značíme $\|\mathbf{x}\|$.

Ve vektorovém prostoru se skalárním součinem $\langle \cdot, \cdot \rangle$ lze zavést normu indukovaně: $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Prostor \mathbb{C}^n nejčastěji doplňujeme o skalární součin

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i^* = \mathbf{y}^* \mathbf{x}$$

a indukovanou normu

$$\|\mathbf{x}\|_2 := \|\mathbf{x}\| = \sqrt{\sum_{i=1}^n |x_i|^2} = \mathbf{x}^* \mathbf{x}.$$

Vektory uvažujeme jako sloupcové a s konečným počtem prvků. Indexování předpokládáme počínaje jedničkou, tedy $\mathbf{x} = [x_1, \dots, x_n]^T$. Kardinalitu množiny (počet prvků) značíme jako absolutní hodnotu, např. $|\{1, 3, 5, 7, 9\}| = 5$.

Definice 3 [19] *Nosičem vektoru \mathbf{x} rozumíme množinu jeho indexů, v nichž má vektor nenulové hodnoty. Tuto množinu pak značíme $\text{supp}(\mathbf{x}) = \{i | x_i \neq 0\}$.*

Pro signál $\mathbf{x} = [x_1, \dots, x_n]^T = [1, 0, 5, 3, 2, 7, 0, 0, 9]^T$ platí $\text{supp}(\mathbf{x}) = \{1, 5, 3, 2, 7, 9\}$ a $|\text{supp}(\mathbf{x})| = 6$.

Definice 4 [19] ℓ_p -norma vektoru $\mathbf{x} \in \mathbb{C}^N$ je definována jako

$$\|\mathbf{x}\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p} \quad \text{pro } 1 \leq p < \infty.$$

Dále speciální případy pro $p \in (0, 1)$, kdy se ve skutečnosti nejedná o normu (neplatí pozitivní homogenita), a pro $p = \infty$. Kvůli zjednodušení však užíváme jednotně ℓ_p -norma pro všechna p .

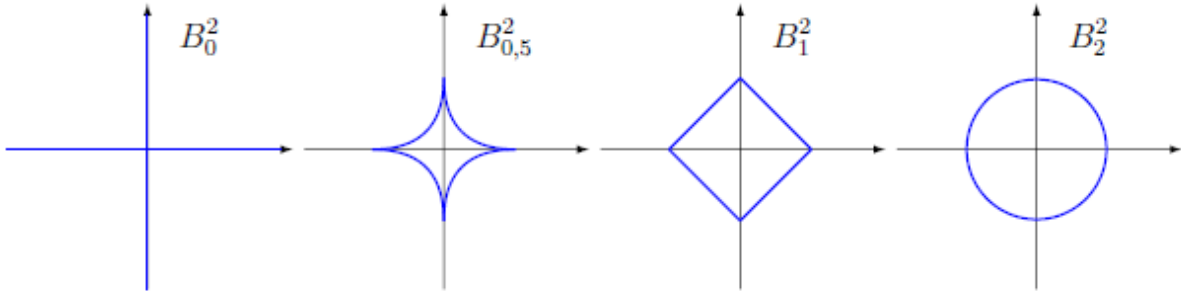
$$\begin{aligned} \|\mathbf{x}\|_\infty &:= \max_i |x_i|, \\ \|\mathbf{x}\|_p &:= \sum_{i=1}^n |x_i|^p \quad \text{pro } 0 < p < 1, \\ \|\mathbf{x}\|_0 &:= |\text{supp}(\mathbf{x})|. \end{aligned}$$

Norma $\|\cdot\|_1$ vyjadřuje součet absolutních hodnot prvků vektoru, $\|\cdot\|_0$ pak počet jeho nenulových složek.

Definice 5 [19] *Jednotkovou kouli B_p^N v ℓ_p -normě definujeme jako*

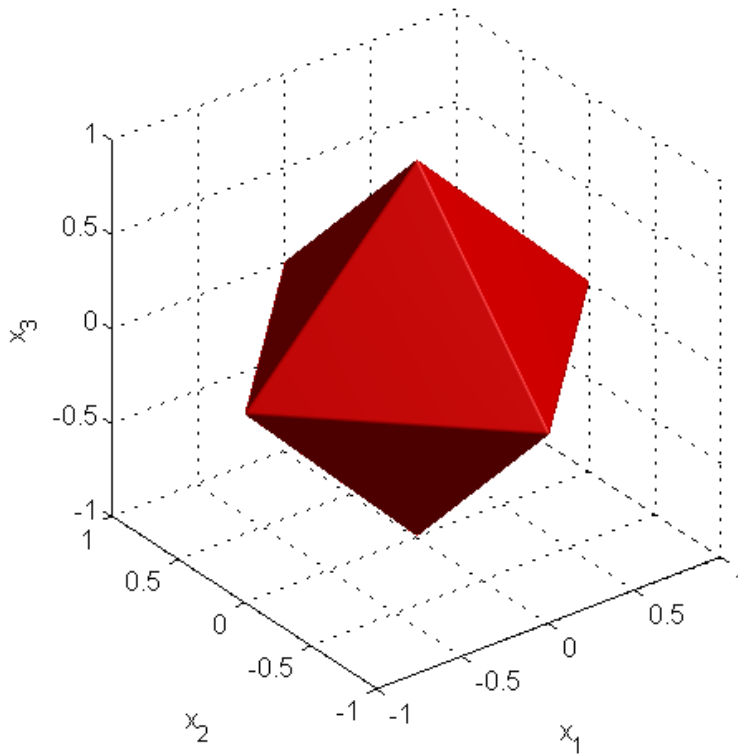
$$B_p^N := \left\{ \mathbf{x} \in \mathbb{C}^N \mid \|\mathbf{x}\|_p \leq 1 \right\}.$$

Ilustrace na obr. 2.1:



Obrázek 2.1: Jednotkové koule v normách $\ell_0, \ell_{0,5}, \ell_1, \ell_2$ [19].

Jednotková koule B_1^3 je znázorněna na obr. 2.2, koule B_2^3 pak na obr. 2.3.



Obrázek 2.2: Jednotková koule v normě $\|\cdot\|_1$.

Definice 6 [19] Vektor $\mathbf{x} \in \mathbb{C}^N$ nazýváme k -řídský (k -sparse), pokud splňuje

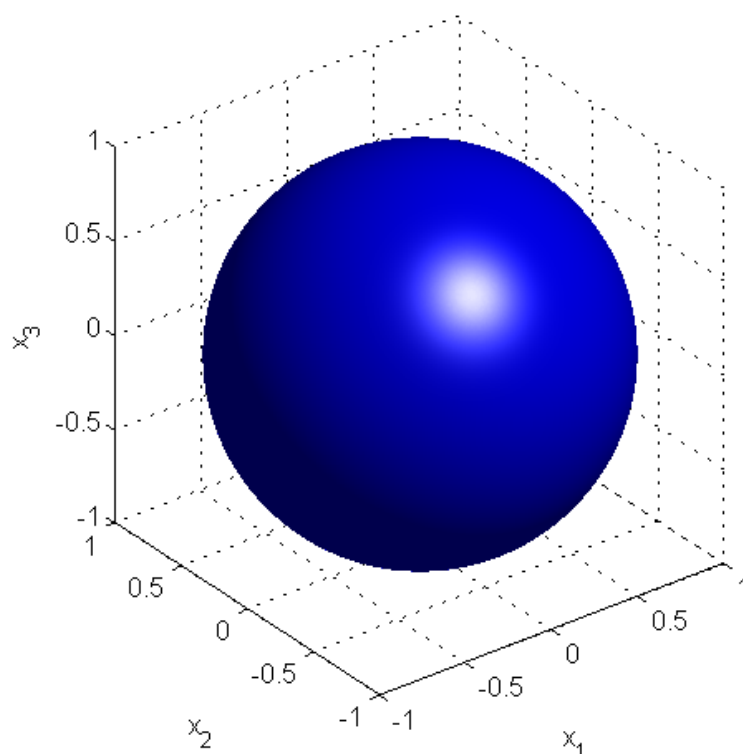
$$\|\mathbf{x}\|_0 \leq k.$$

k -řídský vektor má tedy nejvýše k nenulových složek. Poměrem $\frac{k}{N}$ rozumíme relativní řídkost vektoru \mathbf{x} o délce N .

Definice 7 [33] (p, q) -smíšená norma matice $\mathbf{A} \in \mathbb{R}^{m \times n}$ je obecně definována

$$\|\mathbf{A}\|_{p,q} = \left(\sum \|\mathbf{a}_i\|_p^q \right)^{1/q},$$

kde \mathbf{a}_i je i -tý sloupec matice \mathbf{A} a $\|\mathbf{a}_i\|_p$ klasická ℓ_p -norma.

Obrázek 2.3: Jednotková koule v normě $\|\cdot\|_2$.

Tedy pro příklad $p = 1, q = 2$

$$\|\mathbf{A}\|_{1,2} = \|\|\mathbf{a}_1\|_1, \|\mathbf{a}_2\|_1, \dots, \|\mathbf{a}_n\|_1\|_2 = (\|\mathbf{a}_1\|_1^2 + \|\mathbf{a}_2\|_1^2 + \dots + \|\mathbf{a}_n\|_1^2)^{1/2}.$$

Podobně pro $p = 2, q = 1$. Jednotková koule v normě $\|\cdot\|_{1,2}$, kdy prvky x_1, x_2 tvoří skupinu je zobrazena na obr. 2.4, podobně v normě $\|\cdot\|_{2,1}$ na obr. 2.5.

Pseudoinverze (Moore-Penrose)

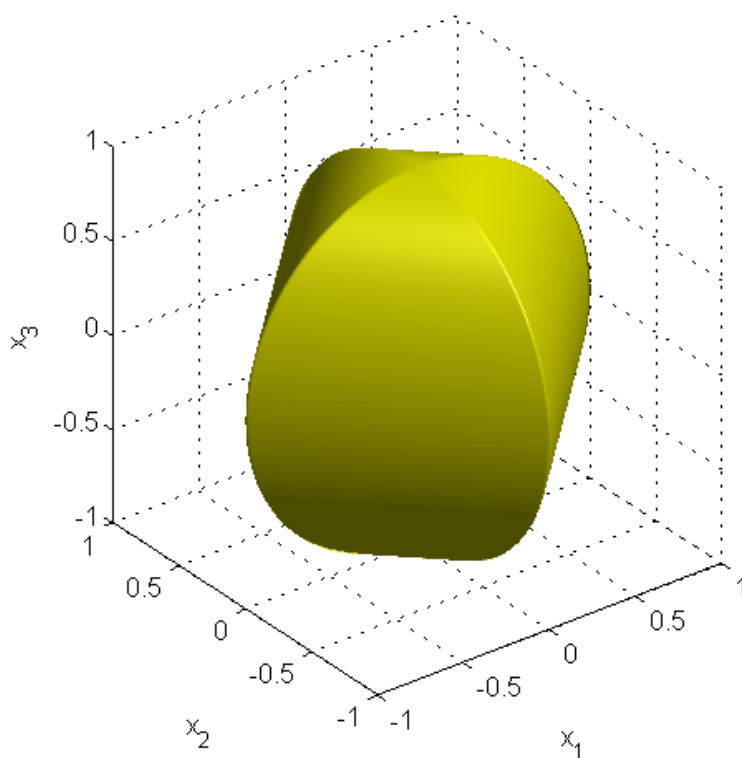
Definice 8 Matice pseudoinverzní k matici \mathbf{A} je \mathbf{A}^+ , pro kterou platí:

$$\begin{aligned} \mathbf{A}\mathbf{A}^+\mathbf{A} &= \mathbf{A}, \\ \mathbf{A}^+\mathbf{A}\mathbf{A}^+ &= \mathbf{A}^+, \\ (\mathbf{A}\mathbf{A}^+)^* &= \mathbf{A}\mathbf{A}^+, \\ (\mathbf{A}^+\mathbf{A})^* &= \mathbf{A}^+\mathbf{A}. \end{aligned}$$

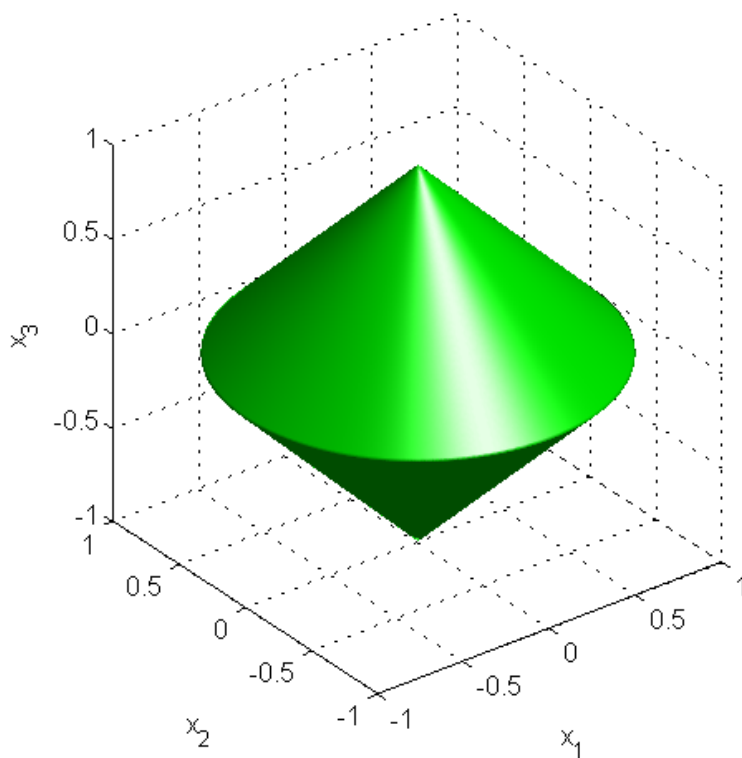
Jde tedy o zobecnění inverze pro neregulární matice: $\mathbf{A}\mathbf{A}^+\mathbf{x} = \mathbf{x}$ [19].

Jestliže řešíme nedourčenou soustavu lineárních rovnic $\mathbf{A}\mathbf{x} = \mathbf{y}$, pak $\mathbf{A}^+\mathbf{y}$ je řešení s nejnižší energií. Lze ukázat [19], že $\mathbf{A}^+ = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^+$, v případě plné řádkové hodnosti matice \mathbf{A} pak

$$\mathbf{A}^+ = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}. \quad (2.1)$$



Obrázek 2.4: Jednotková koule v normě $\|\cdot\|_{1,2}$.



Obrázek 2.5: Jednotková koule v normě $\|\cdot\|_{2,1}$.

Konvexní množina a funkce

Definice 9 [4] Řekneme, že množina $M \subset \mathbb{V}$ je konvexní, jestliže pro $\forall \mathbf{x}_1, \mathbf{x}_2 \in M$ platí

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

Konvexní množina je taková množina, která obsahuje konvexní kombinace všech svých bodů.

Definice 10 [4] Funkce $f(\mathbf{x})$ na konvexní množině se nazývá konvexní, jestliže pro $\forall \mathbf{x}_1, \mathbf{x}_2 \in \text{dom} f$ platí

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

Funkce je tedy konvexní, pokud graf této funkce na každém intervalu $[\mathbf{x}_1, \mathbf{x}_2]$ leží pod úsečkou spojující funkční hodnoty $f(\mathbf{x}_1), f(\mathbf{x}_2)$.

3. Báze a framy

Pokud signál uvažujeme jako vektor délky n , pohybujeme se v teorii vektorových prostorů. Základním pojmem je v této oblasti *báze* vektorového prostoru. Vektory lze jednoznačně vyjádřit jako lineární kombinaci prvků báze. Místo práce s obecnými vektory lze tedy obdobně pracovat s prvky báze. Nevýhodou tohoto konceptu je ovšem požadavek na lineární nezávislost prvků báze a často i jejich ortogonalitu vzhledem ke skalárnímu součinu. Pokud jsou kladeny další požadavky, často bývá těžké nalézt bázi, která jim vyhovuje.

Framy jsou v tomto ohledu všestrannější. Frame ve vektorovém prostoru s vnitřním součinem také umožňuje vyjádřit každý vektor jako lineární kombinaci svých prvků, odpadá však požadavek na lineární nezávislost mezi prvky framu. Pokud tedy uvažujeme reprezentaci signálu pomocí systému lineárních rovnic, v případě framu je tento systém nedourčený. Vyskytuje se v něm více neznámých než lineárních vztahů mezi nimi a tím pádem není reprezentace signálu určena jednoznačně. Hledáme pak takovou reprezentaci, která je v jistém smyslu optimální. Například tzv. *řádkou* reprezentaci signálu (signál je v systému určen jen několika jeho prvky) popisujeme v kapitole 4. Nejdříve připomeneme základní pojmy ohledně reprezentačních systémů. Uvažujeme konečnědimenzionální vektorový prostor $\mathbb{V} \neq \{ \}$ s vnitřním součinem $\langle \cdot, \cdot \rangle$.

3.1. Systém generátorů vektorového prostoru

Jestliže lze každý vektor $\mathbf{x} \in \mathbb{V}$ vyjádřit jako konečnou lineární kombinaci prvků z určité podmnožiny vektorů $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ prostoru \mathbb{V} , nazveme tuto množinu *systém generátorů* vektorového prostoru \mathbb{V} . Pokud je generátorů více než je dimenze n , daný vektor může mít více reprezentací. Toto je již výše zmíněná *nedourčenost*. Lze-li zapsat \mathbf{x} v systému generátorů jako

$$\mathbf{x} = \sum_{k=1}^n c_k \mathbf{e}_k,$$

skaláry c_k se nazývají *souřadnice* \mathbf{x} v \mathbf{E} .

3.2. Báze

Posloupnost vektorů $\{\mathbf{b}_k\}_{k=1}^n$ ve vektorovém prostoru \mathbb{V} dimenze n se nazývá *báze*, jestliže platí [10]:

- $\mathbb{V} = \text{span}\{\mathbf{b}_k\}_{k=1}^n$
- $\{\mathbf{b}_k\}_{k=1}^n$ je *lineárně nezávislá*, tj. pokud $\sum_{k=1}^n c_k \mathbf{b}_k = \mathbf{0}$ pro nějaké skalární koeficienty $\{c_k\}_{k=1}^n$, pak $c_k = 0 \quad \forall k = 1, \dots, n$.

Báze je tedy minimální systém generátorů vektorového prostoru. Každý prvek $\mathbf{x} \in \mathbb{V}$ lze v dané bázi pomocí *jednoznačných* c_k vyjádřit jako

$$\mathbf{x} = \sum_{k=1}^n c_k \mathbf{b}_k$$

Definice 11 [10]. Báze $\{\mathbf{b}_k\}_{k=1}^n$ se nazývá ortogonální, když pro každé dva vektory z $\{\mathbf{b}_k\}_{k=1}^n$ platí

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0, \quad \langle \mathbf{b}_i, \mathbf{b}_i \rangle \neq 0.$$

Báze se nazývá ortonormální, jestliže navíc $\|\mathbf{b}_k\| = 1, \forall k = 1, \dots, n$.

Někdy, zvláště ve vícedimenzionálních prostorech, je náročné najít koeficienty $\{c_k\}_{k=1}^n$. Ovšem pokud je báze *ortonormální*, lze koeficienty nalézt snadno: Vnitřní součin vektoru \mathbf{x} s libovolným \mathbf{b}_j dává

$$\langle \mathbf{x}, \mathbf{b}_j \rangle = \left\langle \sum_{k=1}^n c_k \mathbf{b}_k, \mathbf{b}_j \right\rangle = \sum_{k=1}^n c_k \langle \mathbf{b}_k, \mathbf{b}_j \rangle = c_j,$$

takže

$$\mathbf{x} = \sum_{k=1}^n \langle \mathbf{x}, \mathbf{b}_k \rangle \mathbf{b}_k.$$

3.3. Framy

Jestliže je počet generátorů vektorového prostoru \mathbb{V} vyšší než jeho dimenze, stále je možné reprezentovat každý vektor konečnou lineární kombinací, generátory jsou však v tomto případě již lineárně závislé. Takový systém generátorů se nazývá (*konečný*) *frame* ve \mathbb{V} . Framy jsou tedy zobecněním pojmu báze, používají se kvůli své flexibilitě (jsou méně omezené). Jejich nevýhoda spočívá v riziku numerické nestability a náročnějších výpočetních metodách.

Definice 12 [10]. Nejvýše spočetná množina vektorů $\{\phi_k\}_{k \in I}$ ve \mathbb{V} v něm tvoří *frame*, pokud existují konstanty $0 < A \leq B < \infty$ tak, že

$$A \|\mathbf{x}\|^2 \leq \sum_{k \in I} |\langle \mathbf{x}, \phi_k \rangle|^2 \leq B \|\mathbf{x}\|^2, \quad \forall \mathbf{x} \in \mathbb{V}.$$

Konstanty A, B se nazývají *meze framy*, nejsou jednoznačné. *Optimální dolní mez framy* je supremum přes všechny dolní meze, podobně *optimální horní mez framy* infimum přes všechny horní meze. Prvky framy ϕ_k se nazývají *atomy*.

V konečnědimenzionálním prostoru lze uvažovat framy $\{\phi_k\}_{k \in I}$ sestávající z nekonečně mnoha prvků. My se omezíme na konečné množiny $\{\phi_k\}_{k=1}^m, m \in \mathbb{N}$. V tomto případě Cauchy-Schwarzova nerovnost dává [10]

$$\sum_{k=1}^m |\langle \mathbf{x}, \phi_k \rangle|^2 \leq \sum_{k=1}^m \|\phi_k\|^2 \|\mathbf{x}\|^2, \quad \forall \mathbf{x} \in \mathbb{V}.$$

3.3. FRAMY

Tedy podmínka horní meze je automaticky splněna. Dolní mez $A > 0$ existuje, pokud $\text{span}\{\phi_k\}_{k=1}^m = \mathbb{V}$.

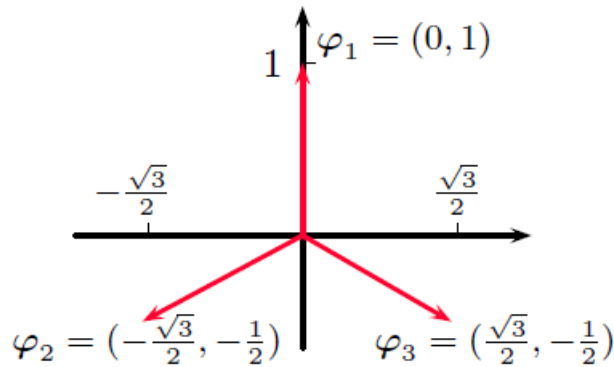
Frame nazveme *těsný* [10], pokud je v definici možné vybrat konstanty $A = B$, tedy platí

$$\sum_{k=1}^m |\langle \mathbf{x}, \phi_k \rangle|^2 = A \|\mathbf{x}\|^2, \quad \forall \mathbf{x} \in \mathbb{V}.$$

Příkladem těsného framu je tzv. Mercedes-Benz frame v \mathbb{R}^2 [10]:

$$\Phi = \begin{pmatrix} 0 & -\sqrt{3}/2 & \sqrt{3}/2 \\ 1 & -1/2 & -1/2 \end{pmatrix}$$

Znázorněný je na obr. 3.1.



Obrázek 3.1: Mercedes-Benz frame.

Výraz $\sum_k |\langle \mathbf{x}, \phi_k \rangle|^2$ můžeme přepsat: skalární součin $\langle \mathbf{x}, \phi_k \rangle$ jako $\phi_k^* \mathbf{x}$, a tedy posloupnost $\langle \mathbf{x}, \phi_k \rangle_k$ jako $\Phi^* \mathbf{x}$. ϕ_k jsou obsaženy v Φ jako sloupce a Φ^* je určena jednoznačně. Potom

$$\sum_k |\langle \mathbf{x}, \phi_k \rangle|^2 = (\Phi^* \mathbf{x})^* \Phi^* \mathbf{x} = \mathbf{x}^* \Phi \Phi^* \mathbf{x},$$

což je kvadratická forma určená $\Phi \Phi^*$ [34]. Definici framu pak přepíšeme na

$$A \leq \mathbf{x}^* \Phi \Phi^* \mathbf{x} \leq B, \quad \forall \mathbf{x} \in \mathbb{V} : \|\mathbf{x}\| = 1.$$

Optimální horní mez framu A je tedy rovna nejmenšímu vlastnímu číslu $\Phi \Phi^*$ a optimální horní mez B tomu největšímu [34].

Uvažujme frame $\{\phi_k\}_{k=1}^m$ ve vektorovém prostoru \mathbb{V} . Definujme lineární zobrazení [10]

$$\mathcal{T} : \mathbb{C}^m \rightarrow \mathbb{V}, \quad \mathcal{T}\{c_k\}_{k=1}^m = \sum_{k=1}^m c_k \phi_k.$$

\mathcal{T} nazýváme *syntezující operátor*. Adjungovaný operátor je dán

$$\mathcal{T}^* : \mathbb{V} \rightarrow \mathbb{C}^m, \quad \mathcal{T}^* \mathbf{x} = \{\langle \mathbf{x}, \phi_k \rangle\}_{k=1}^m,$$

nazývá se *analyzující operátor*. Složením \mathcal{T} a \mathcal{T}^* získáme *framový operátor*

$$\mathcal{S} : \mathbb{V} \rightarrow \mathbb{V}, \quad \mathcal{S}\mathbf{x} = \mathcal{T}\mathcal{T}^*\mathbf{x} = \sum_{k=1}^m \langle \mathbf{x}, \phi_k \rangle \phi_k.$$

Věta 1 [10] *Nechť $\{\phi_k\}_{k=1}^m$ je frame ve \mathbb{V} s framovým operátorem \mathcal{S} . Pak platí následující:*

1. \mathcal{S} je invertibilní a samoadjungovaný
2. $\forall \mathbf{x} \in \mathbb{V}$ lze vyjádřit jako

$$\mathbf{x} = \sum_{k=1}^m \langle \mathbf{x}, \mathcal{S}^{-1}\phi_k \rangle \phi_k = \sum_{k=1}^m \langle \mathbf{x}, \phi_k \rangle \mathcal{S}^{-1}\phi_k. \quad (3.1)$$

3. Pokud $\mathbf{x} \in \mathbb{V}$ má reprezentaci $\mathbf{x} = \sum_{k=1}^m c_k \phi_k$ pro nějaké skalární koeficienty $\{c_k\}_{k=1}^m$, pak

$$\sum_{k=1}^m |c_k|^2 = \sum_{k=1}^m |\langle \mathbf{x}, \mathcal{S}^{-1}\phi_k \rangle|^2 + \sum_{k=1}^m |c_k - \langle \mathbf{x}, \mathcal{S}^{-1}\phi_k \rangle|^2.$$

Vztah (3.1) ve větě 1 se nazývá *framový rozklad*. Věta říká, že pro obecné framy lze každý vektor $\mathbf{x} \in \mathbb{V}$ reprezentovat ve tvaru $\mathbf{x} = \sum_{k=1}^m \langle \mathbf{x}, \mathbf{f}_k \rangle \phi_k$ pro vhodnou volbu $\{\mathbf{f}_k\}_{k=1}^m$. Navíc, mezi všemi skalárními posloupnostmi $\{c_k\}_{k=1}^m$, pro které platí $\mathbf{x} = \sum_{k=1}^m c_k \phi_k$, mají koeficienty $\{\langle \mathbf{x}, \mathcal{S}^{-1}\phi_k \rangle\}_{k=1}^m$ minimální energii (ℓ_2 -normu).

Definice 13 [10] *Nechť je $\mathbf{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ frame vektorového prostoru \mathbb{V} , $m > n = \dim(\mathbb{V})$. Vektory framu jsou lineárně závislé, každý vektor z \mathbb{V} lze vyjádřit nekonečně mnoha kombinacemi $\mathbf{x} = \sum_k c_k \mathbf{e}_k$. Frame $\mathbf{F} = \{\mathbf{f}_1, \dots, \mathbf{f}_m\}$, který generuje tentýž prostor se nazývá duální k \mathbf{E} , pokud $\forall \mathbf{x} \in \mathbb{V}$*

$$\mathbf{x} = \sum_k \langle \mathbf{x}, \mathbf{f}_k \rangle \mathbf{e}_k = \sum_k \langle \mathbf{x}, \mathbf{e}_k \rangle \mathbf{f}_k,$$

tedy $\mathbf{x} = \mathbf{E}\mathbf{F}^*\mathbf{x} = \mathbf{F}\mathbf{E}^*\mathbf{x}$.

Duálních framů může být více. Jen pro jeden duální frame \mathbf{F} k framu \mathbf{E} , definovaný $\{\mathbf{f}_k\}_{k=1}^m = \{\mathcal{S}^{-1}\mathbf{e}_k\}_{k=1}^m$ (neboli $\mathbf{F} = \mathcal{S}^{-1}\mathbf{E}$), má posloupnost jeho souřadnic (tzv. framových koeficientů) dle věty 1 minimální energii. Tento frame se nazývá *kanonický duální frame* [10] k framu \mathbf{E} . Z definice tedy plyne, že souřadnice c_k vektoru v původním framu hledáme pomocí duálního: $\langle \mathbf{x}, \mathbf{f}_k \rangle = c_k$ a naopak: $\langle \mathbf{x}, \mathbf{e}_k \rangle = d_k$, tj. $\mathbf{F}^*\mathbf{x} = \mathbf{c}$ a $\mathbf{E}^*\mathbf{x} = \mathbf{d}$.

Kanonický duální frame lze vytvořit také použitím Moore-Penroseovy pseudoinverze [34]. Vycházíme z (2.1):

$$\mathbf{E}^+ = \mathbf{E}^*(\mathbf{E}\mathbf{E}^*)^{-1},$$

kde \mathbf{E} je původní frame. Protože framový operátor je definovaný $\mathcal{S}\mathbf{x} = \sum_k \langle \mathbf{x}, \mathbf{e}_k \rangle \mathbf{e}_k$, lze jej v konečné dimenzi ztotožnit s maticí $\mathbf{S} = \mathbf{E}\mathbf{E}^*$, a můžeme pak rovnici pro pseudoinverzi zapsat

$$\mathbf{E}^+ = \mathbf{E}^*\mathbf{S}^{-1}.$$

3.4. PROSTORY NEKONEČNÉ DIMENZE

Následně upravíme hermitovskou transpozicí

$$(\mathbf{E}^+)^* = (\mathbf{E}^+ \mathbf{S}^{-1})^* = \mathbf{S}^{-1} \mathbf{E} = \mathbf{F}.$$

Hledání duálního framu je často náročné (velké rozměry matice či špatná podmíněnost), v případě *těsných* framů ovšem platí $\mathcal{S} = \mathbf{A}\mathbf{I}$ [10] a duální frame vytvoříme z původního jednoduše $\mathbf{f}_k = \mathbf{e}_k/A$. Tedy platí

$$\mathbf{x} = \frac{1}{A} \sum_k \langle \mathbf{x}, \mathbf{e}_k \rangle \mathbf{e}_k.$$

Toto vede ke hledání právě těsných framů v různých prostorech.

3.4. Prostory nekonečné dimenze

V prostorech nekonečné dimenze některá samozřejmá fakta z konečněrozměrných prostorů neplatí. Prvky jsou vyjadřovány pomocí nekonečných řad [10]

$$\mathbf{x} = \sum_{k=-\infty}^{\infty} c_k \mathbf{f}_k$$

a je potřeba zde vyžadovat bezpodmínečnou konvergenci, kde součet nezávisí na pořadí sčítanců. Pojem lineární nezávislosti je zobecněný: nekonečná množina je považována za lineárně nezávislou, jestliže každá její konečná podmnožina je lineárně nezávislá. Pracujeme s tzv. *separabilním Hilbertovým prostorem* [10]. Hilbertův prostor je definován jako úplný vektorový prostor se skalárním součinem. *Úplnost* znamená, že každá cauchyovská posloupnost vektoru je zároveň v tomto prostoru konvergentní. *Separabilnost* znamená, že obsahuje nejvýše spočetnou hustou podmnožinu (tedy tato podmnožina generuje podprostor hustý v původním prostoru).

Nejčastěji využívané báze v tomto prostoru jsou tzv. *Schauderova* a *Riezsova*. Frame je definovaný stejným způsobem jako v konečné dimenzi s rozdílem, že množina I je tentokrát spočetná [10].

Příkladem tohoto prostoru nekonečné dimenze je prostor funkcí s konečnou energií $L^2(\mathbb{R})$:

$$L^2(\mathbb{R}) = \left\{ f : \sqrt{\int_{-\infty}^{\infty} |f(t)|^2 dt} < \infty \right\},$$

f je komplexní funkce reálné proměnné. Skalární součin definujeme jako

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(t)g^*(t),$$

indukovanou normu $\|f\|_2 = \sqrt{\int_{-\infty}^{\infty} |f(t)|^2 dt}$. V následující části se věnujeme systému generátorů pro $L^2(\mathbb{R})$, kde jsou prvky framu konstruovány pomocí translace a modulace.

3.5. Gaborovy framy

Teorie Gaborovy analýzy v prostoru $L^2(\mathbb{R})$ je založena na dvou třídách operátorů, *translace* T_a o $a \in \mathbb{R}$ a *modulace* E_b číslem $b \in \mathbb{R}$:

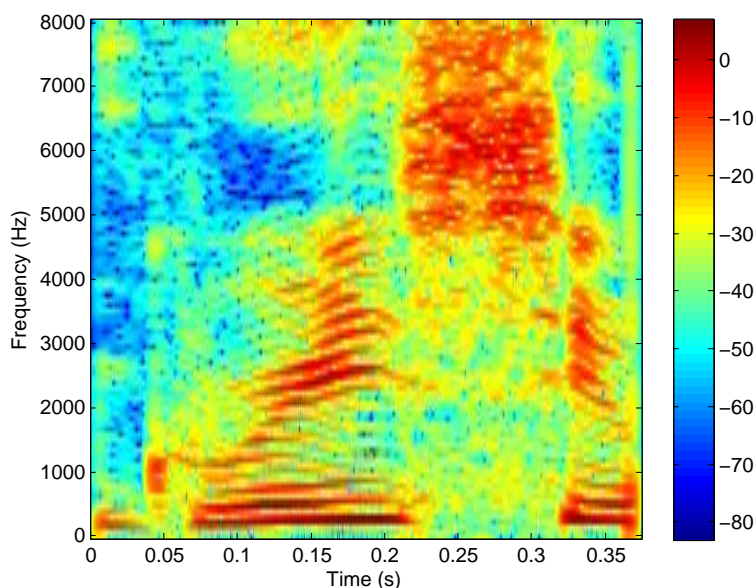
$$\begin{aligned} T_a : L^2(\mathbb{R}) &\rightarrow L^2(\mathbb{R}), & (T_a f)(x) &= f(x - a), \\ E_b : L^2(\mathbb{R}) &\rightarrow L^2(\mathbb{R}), & (E_b f)(x) &= e^{2\pi i b x} f(x). \end{aligned}$$

Funkce $f \in L^2(\mathbb{R})$ jsou reprezentovány jako superpozice posunutých a modulovaných verzí pevně dané funkce $g \in L^2(\mathbb{R})$. Získaný soubor funkcí

$$\{E_{mb}T_{na}g\}_{m,n \in \mathbb{Z}} = \{e^{2\pi i m b x} g(x - na)\}_{m,n \in \mathbb{Z}}$$

se nazývá Gaborův systém [3],[10].

Snaha zavést tento systém plyne z potřeby časově-kmitočtové reprezentace: Fourierova transformace využívá harmonických funkcí, které ovlivňují podobu celého signálu. Lidské ucho ale dokáže umístit kmitočtovou strukturu zvuků v čase, proto potřebujeme pro analýzu časově lokální spektra signálů. Pomocí Gaborových framů můžeme signál analyzovat či syntetizovat za pomoci stavebních bloků lokalizovaných v čase i frekvenci. Tato reprezentace vede na tzv. *spektrogram* (obr. 3.2). Časově-kmitočtové rozlišení má bohužel svoje omezení. Žádný signál totiž nemůže být najednou koncentrován jak v čase, tak v kmitočtu. Toto popisuje tzv. Heisenbergův princip neurčitosti [13].



Obrázek 3.2: Spektrogram signálu - vyslovení 'greasy'. Čím teplejší barva, tím významnější je daná frekvence v daném čase.

Definice 14 [10] *Gaborův frame je frame pro $L^2(\mathbb{R})$ tvaru $\{E_{mb}T_{na}g\}_{m,n \in \mathbb{Z}}$, kde $a, b > 0$ jsou dané a $g \in L^2(\mathbb{R})$ je fixní funkce.*

Funkce g se nazývá *okénková funkce* nebo *generátor*.

Otázkou je výběr parametrů $a, b > 0$ a funkce $g \in L^2(\mathbb{R})$ tak, abychom získali frame v $L^2(\mathbb{R})$.

3.5. GABOROVY FRAMY

Věta 2 [10] *Nechť je dána funkce $g \in L^2(\mathbb{R})$ a konstanty $a, b > 0$. Pak platí:*

1. *Pokud $ab > 1$, pak $\{E_{mb}T_{na}g\}_{m,n \in \mathbb{Z}}$ nemůže být frame pro $L^2(\mathbb{R})$.*
2. *Pokud je $\{E_{mb}T_{na}g\}_{m,n \in \mathbb{Z}}$ frame, pak $ab = 1 \Leftrightarrow \{E_{mb}T_{na}g\}_{m,n \in \mathbb{Z}}$ je Rieszova báze.*

Tedy systém $\{E_{mb}T_{na}g\}_{m,n \in \mathbb{Z}}$ tvoří v $L^2(\mathbb{R})$ frame jen za podmínky $ab \leq 1$, což znamená, že mřížka $\{(na, mb)\}_{m,n \in \mathbb{Z}}$ musí být dostatečně hustá (toto je pouze nutná podmínka).

Pro rekonstrukci signálu je potřeba, aby měl duální frame ke Gaborovu framu opět stejnou strukturu. Platí, že každý kanonický duální frame toto splňuje [10]. Pro získání duálního okna je potřeba provést inverzi framového operátoru, což bývá (s výjimkou těsných framů) nelehký úkol.

Pro praktické účely se přesouváme do konečných posloupností v prostoru $\ell^2(\mathbb{Z})$. Jisté podmínky, kladené na Gaborův frame $\{E_{mb}T_{na}g\}_{m,n \in \mathbb{Z}}$ pro $L^2(\mathbb{R})$ implikují, že můžeme sestavit frame pro $\ell^2(\mathbb{Z})$ s podobnou strukturou [10]. Diskrétní Gaborův systém generovaný posloupností $g \in \ell^2(\mathbb{Z})$ s modulačním parametrem $1/M$ a translačním parametrem a , $(M, a \in \mathbb{N})$ je definován jako soubor posloupností

$$\{E_{m/M}T_{na}g\}_{n \in \mathbb{Z}, m=0, \dots, M-1}.$$

Tedy j -tý člen posloupnosti je

$$E_{m/M}T_{na}g(j) = e^{2\pi i \frac{m}{M}j} g(j - na).$$

Nutnou podmínkou, aby $\{E_{m/M}T_{na}g\}_{n \in \mathbb{Z}, m=0, \dots, M-1}$ byl frame v $\ell^2(\mathbb{Z})$ je $\frac{a}{M} \leq 1$. Navíc, pokud je tento systém frame, je Rieszovou bází právě tehdy, když $a = M$ [10].

Bylo ukázáno, že je možné zkonstruovat soubor posloupností v $\ell^2(\mathbb{Z})$ na základě Gaborových systémů v $L^2(\mathbb{R})$ diskretizací funkce g za splnění určitých podmínek na tuto funkci [10]. Okénkové funkce používané v aplikacích jsou např. Hannova, Hammingova, Blackmanova, Nuttaloa a další [31].

Koeficienty $x_{m,n}$ vektoru $\mathbf{y} \in \mathbb{C}^L$ v Gaborově framu tedy získáme

$$x_{m,n} = \langle \mathbf{y}, \mathbf{g}_{m,n} \rangle,$$

kde $\mathbf{g}_{m,n}$ jsou posunutá a modulovaná verze okénkové funkce \mathbf{g} tvořící atomy Gaborova framu, $n = 0, \dots, N-1, m = 0, \dots, M-1$. N -násobek parametru translace a M -násobek parametru modulace navíc dává délku vektoru L . Pokud uvažujeme těsný frame s mezí $A = 1$ (tzv. Parsevalův), syntézu lze provádět analyzujícím oknem, tedy celkově signál $\mathbf{y} \in \mathbb{C}^L$ reprezentujeme

$$\mathbf{y} = \sum_{m,n} \langle \mathbf{y}, \mathbf{g}_{m,n} \rangle \mathbf{g}_{m,n}.$$

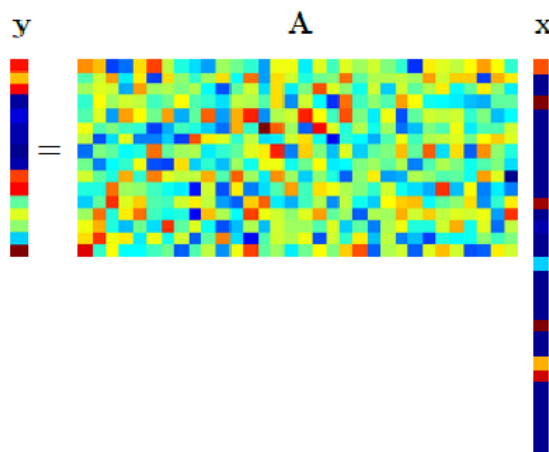
4. Řídké reprezentace signálů

Signál můžeme reprezentovat pomocí systému lineárních rovnic $\mathbf{Ax} = \mathbf{y}$, který je nedourčený. Tímto má soustava nekonečně mnoho řešení a v rámci této reprezentace nás zajímají tzv. řídká řešení (mají co nejvíce neznámých současně nulových). Výhodou této reprezentace je snadná a silná komprese, numerická stabilita, jednoduchá interpretace a další.

Tedy chceme řešit soustavu $\mathbf{Ax} = \mathbf{y}$, přičemž hledané řešení má být co nejřidší:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{vzhledem k} \quad \mathbf{Ax} = \mathbf{y}, \quad (\text{P0})$$

známý je pro nás vektor $\mathbf{y} \in \mathbb{C}^m$, který vyjadřuje signál (měření, pozorování), a matice $\mathbf{A} \in \mathbb{C}^{m \times N}$ (tzv. *slovník*, její sloupce jsou tzv. *atomy*). Předpokládáme, že \mathbf{A} je plné hodnosti a $m < N$. Všechna \mathbf{x} splňující $\mathbf{Ax} = \mathbf{y}$ nazýváme přípustné reprezentace vektoru \mathbf{y} (obr. 4.1).



Obrázek 4.1: Nedourčený systém lineárních rovnic $\mathbf{Ax} = \mathbf{y}$. Vektor \mathbf{y} je lineární kombinace pouze několika sloupců matice \mathbf{A} . Modrá barva reprezentuje nulu a s teplotou barvy roste hodnota [19].

Často se stává, že je signál zašuměný, a tím pádem přesné řešení neexistuje. V tomto případě se povoluje malá odchylka:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{vzhledem k} \quad \|\mathbf{Ax} - \mathbf{y}\|_2 \leq \delta. \quad (\text{P0}\delta)$$

4.1. Postačující podmínky pro jednoznačnost řešení

Důležitým pojmem pro rozhodnutí o existenci a vlastnostech řešení je *spark* matice.

Definice 15 [19] Číslo $\text{spark}(\mathbf{A})$ je definováno jako nejmenší počet sloupců matice \mathbf{A} , které jsou lineárně závislé.

$$\text{spark}(\mathbf{A}) = \min_{\mathbf{z} \in \ker(\mathbf{A}), \mathbf{z} \neq \mathbf{0}} \|\mathbf{z}\|_0.$$

4.2. ALGORITMY HLEDÁNÍ ŘÍDKÝCH REPREZENTACÍ

Pro nenulovou matici $\mathbf{A} \in \mathbb{C}^{m \times N}$, $m < N$, platí $\text{spark}(\mathbf{A}) \in \{2, \dots, m + 1\}$, přičemž hodnoty 2 nabývá, když jeden sloupec je násobkem jiného. Čím menší je spark , tím řidší musí vektor \mathbf{x} být pro zajištění jednoznačnosti tohoto řešení.

Definice 16 [19] Pokud má soustava $\mathbf{Ax} = \mathbf{b}$ řešení \mathbf{x} splňující

$$\|\mathbf{x}\|_0 < \frac{\text{spark}(\mathbf{A})}{2}, \quad (4.1)$$

pak \mathbf{x} je nutně nejřidší možné řešení a žádné jiné řešení se stejnou řídkostí neexistuje.

Jestliže tedy nalezneme řešení splňující (4.1), je tím nalezeno i řešení původního problému [19]. Nalezení $\text{spark}(\mathbf{A})$ je bohužel také výpočetně náročné, je potřeba hledat jednodušší způsob ověření jednoznačnosti.

Definice 17 [19] *Vzájemná koherence (mutual coherence) matice \mathbf{A} je definována jako největší absolutní normalizovaný skalární součin dvou různých sloupců matice \mathbf{A} ,*

$$\mu(\mathbf{A}) = \max_{1 \leq j, k \leq N, j \neq k} \frac{|\mathbf{a}_j^\top \mathbf{a}_k|}{\|\mathbf{a}_j\|_2 \cdot \|\mathbf{a}_k\|_2},$$

kde \mathbf{a}_j označuje j -tý sloupec matice \mathbf{A} .

Vzájemná koherence vyjadřuje míru lineární závislosti sloupců matice. Pouze unitární matice (sloupce jsou po dvou ortogonální a čitatel je tedy nulový) mají nulovou koherenci. Pro náš případ $m < N$ mají matice $m \times N$ plné hodnosti vždy nenulovou koherenci.

Tvrzení 3 [19] Pro libovolnou matici \mathbf{A} platí

$$\text{spark}(\mathbf{A}) \geq 1 + \frac{1}{\mu(\mathbf{A})}.$$

Tímto jednoduše získáme dolní ohraničení pro $\text{spark}(\mathbf{A})$. Pak platí:

Tvrzení 4 [19] Pokud má soustava $\mathbf{Ax} = \mathbf{b}$ řešení \mathbf{x} splňující

$$\|\mathbf{x}\|_0 < \frac{1}{2} \left(1 + \frac{1}{\mu(\mathbf{A})} \right), \quad (4.2)$$

pak je \mathbf{x} nutně nejřidší možné a jediné takové. Navíc lze tohoto řešení dosáhnout ℓ_1 -minimalizací (část 4.2.1).

Jestliže se blíží koherence nule, pravá strana (4.2) roste nad všechny meze. Snahou je tedy používat maximálně nekoherentní slovníky.

4.2. Algoritmy hledání řídkých reprezentací

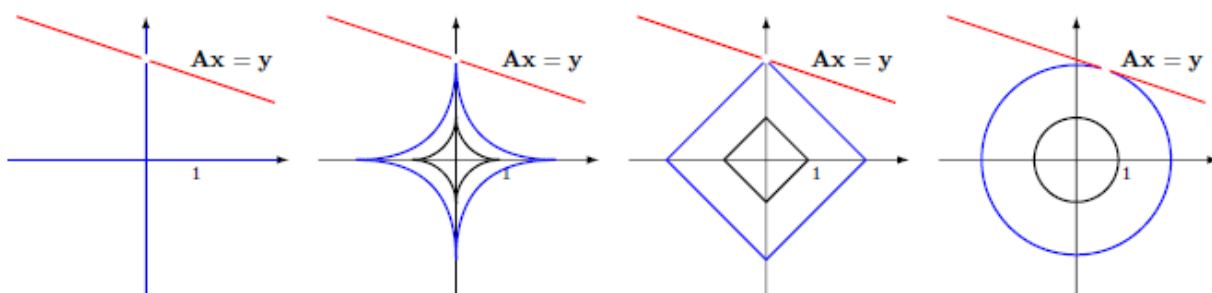
Předpokládáme $\text{spark}(\mathbf{A}) > k_0$ a existenci k_0 -řídkého řešení soustavy lineárních rovnic. Toto řešení je pak nejřidší možné a jediné takové. Pro nalezení přesného řešení je nutné projít $\binom{N}{k_0}$ kombinací podmnožin sloupců matice \mathbf{A} , což je NP-složitá úloha. Existují aproximativní metody, které dělíme zhruba do dvou skupin podle jejich principu.

4.2.1. ℓ_1 relaxace

Protože ℓ_0 -norma není konvexní funkce, nemůžeme pro úlohu (P0) použít žádnou metodu konvexní optimalizace. Normy jsou ovšem pro $p \geq 1$ konvexními funkcemi (obr. 4.2), nabízí se tedy možnost využití ℓ_1 -normy. Úloha (P0) by tedy přešla na

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k} \quad \mathbf{Ax} = \mathbf{y}. \quad (\text{P1})$$

Za jistých podmínek lze opravdu použít ℓ_1 -normu místo ℓ_0 .



Obrázek 4.2: „Nafukující“ se koule v normách $\ell_0, \ell_{0.5}, \ell_1, \ell_2$ a jejich dotyk s nadrovinou $\mathbf{Ax} = \mathbf{y}$ [19].

Podmínky ekvivalence řešení ℓ_0 - a ℓ_1 -minimalizace

Definice 18 [19] Chyba nejlepší aproximace vektoru $\mathbf{x} \in \mathbb{C}^N$ k -řádkým vektorem v normě ℓ_p je definována jako

$$\sigma(\mathbf{x})_p := \sigma_k^N(\mathbf{x})_p := \inf_{\mathbf{z} \in \Sigma_k^N} \|\mathbf{x} - \mathbf{z}\|_p,$$

kde $\Sigma_k^N := \{\mathbf{x} \in \mathbb{C}^N \mid \|\mathbf{x}\|_0 \leq k\}$ označuje množinu všech k -řádkých vektorů o délce N .

Pro $T \subset \{1, \dots, N\}$ označíme $\mathbf{x}_T \in \mathbb{C}^N$ vektor odvozený z $\mathbf{x} \in \mathbb{C}^N$ tak, že prvky na pozicích patřících do T ponecháme a zbytek vynulujeme. Doplněk T označíme $T^c = \{1, \dots, N\} \setminus T$.

Definice 19 [19] Řekneme, že matice $\mathbf{A} \in \mathbb{C}^{m \times N}$ splňuje vlastnost nulového prostoru (Null Space Property, NSP) řádu k s konstantou $\gamma \in (0, 1)$, pokud platí

$$\|\boldsymbol{\eta}_T\|_1 \leq \gamma \|\boldsymbol{\eta}_{T^c}\|_1,$$

pro všechny množiny $T \subset \{1, \dots, N\}$, $|T| \leq k$ a pro všechny vektory $\boldsymbol{\eta} \in \ker(\mathbf{A})$.

Vlastnost nulového prostoru zajišťuje nalezení k -řádkého řešení ℓ_1 -minimalizací i jeho jednoznačnost. Následující dává odhad chyby i v ostatních případech:

Tvrzení 5 [19] Nechť matice $\mathbf{A} \in \mathbb{C}^{m \times N}$ splňuje NSP řádu k s konstantou $\gamma \in (0, 1)$. Nechť $\mathbf{x} \in \mathbb{C}^N$, $\mathbf{y} = \mathbf{Ax}$ a $\mathbf{x}^* \in \mathbb{C}^N$ je řešení ℓ_1 -minimalizace. Potom

$$\|\mathbf{x} - \mathbf{x}^*\|_1 \leq \frac{2(1 + \gamma)}{1 - \gamma} \sigma_k(\mathbf{x})_1.$$

4.2. ALGORITMY HLEDÁNÍ ŘÍDKÝCH REPREZENTACÍ

Pokud existuje nějaké nejvýše k -řádké řešení $\mathbf{y} = \mathbf{A}\mathbf{x}$ při splnění předpokladů tvrzení, pak $\sigma_k(\mathbf{x})_1 = 0$ a proto $\mathbf{x} = \mathbf{x}^*$. Tedy toto řídké řešení nalezneme také ℓ_1 -minimalizací. Opačně platí, že jestliže lze z $\mathbf{A}\mathbf{x} = \mathbf{y}$ získat všechny k -řádké vektory \mathbf{x} ℓ_1 -optimalizací, pak matice \mathbf{A} splňuje NSP řádu k s nějakou jinou konstantou $\gamma \in (0, 1)$ [19].

Tzv. *vlastnost zeslabené isometrie* (Restricted Isometry Property, RIP) přináší řešení stabilní i pod vlivem šumu a je výpočetně přijatelnější.

Definice 20 [19] *Konstanta omezené isometrie δ_k matice $\mathbf{A} \in \mathbb{C}^{m \times N}$ je nejmenší číslo takové, že platí*

$$(1 - \delta_k) \leq \frac{\|\mathbf{A}\mathbf{z}\|_2^2}{\|\mathbf{z}\|_2^2} \leq (1 + \delta_k)$$

pro všechny vektory $\mathbf{z} \in \sum_k^N$. Řekneme, že matice \mathbf{A} splňuje RIP řádu k s konstantou δ_k , pokud $\delta_k \in (0, 1)$.

Isometrie je lineární zobrazení zachovávající délku vektorů, zeslabením se omezíme na všechny podmatice \mathbf{A} s k sloupci a povolujeme malou odchylku δ_k . Tedy všechny tyto podmatice musí být přibližně ortogonální.

Předem nevíme, které prvky vektoru \mathbf{x} budou nenulové, tedy které sloupce matice \mathbf{A} se na signálu \mathbf{y} podílejí, a proto je nutné do definice RIP zahrnout *všechny* matice o k sloupcích.

RIP lze spočítat přímo [19]:

$$\delta_k = \max_{T \subset \{1, \dots, N\}, |T| \leq k} \|\mathbf{A}_T^\top \mathbf{A}_T - \mathbf{I}\|_{2 \rightarrow 2},$$

kde norma matice $\|\cdot\|_{2 \rightarrow 2}$ je definována

$$\|\mathbf{B}\|_{2 \rightarrow 2} = \max_{\mathbf{x}} \frac{\|\mathbf{B}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}.$$

Jako u NSP lze i u RIP určit ℓ_1 -minimalizací chybu aproximace. Také lze najít vzájemný vztah mezi těmito vlastnostmi.

Tvrzení 6 [19] *Nechť $\mathbf{A} \in \mathbb{C}^{m \times N}$ splňuje RIP řádu $K = k + h$ s konstantou $\delta_k \in (0, 1)$. Pak \mathbf{A} splňuje NSP řádu k s konstantou*

$$\gamma = \sqrt{\frac{k}{h} \frac{1 + \delta_k}{1 - \delta_k}}.$$

Odchylku řešení můžeme omezit shora:

Tvrzení 7 [19] *Nechť $\mathbf{A} \in \mathbb{C}^{m \times N}$ splňuje RIP řádu $3k$ s konstantou $\delta_{3k} < \frac{1}{3}$. Pro $\mathbf{x} \in \mathbb{C}^N$, nechť $\mathbf{y} = \mathbf{A}\mathbf{x}$ a $\mathbf{x}^* \in \mathbb{C}^N$ je řešení ℓ_1 -minimalizace. Pak*

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C \frac{\sigma_k(\mathbf{x})_1}{\sqrt{k}}$$

a

$$\|\mathbf{x} - \mathbf{x}^*\|_1 \leq C \sigma_k(\mathbf{x})_1$$

pro nějakou konstantu C .

Dále vlastnosti řešení pro případ šumu:

Tvrzení 8 [19] Předpokládejme, že matice $\mathbf{A} \in \mathbb{C}^{m \times N}$ vyhovuje RIP řádu $2k$ s konstantou

$$\delta_k < \frac{2}{3 + \sqrt{7/4}} \approx 0,4627.$$

Následující platí pro $\forall \mathbf{x} \in \mathbb{C}^N$. Nechť měření jsou zatížena šumem: $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}$, $\|\mathbf{e}\|_2 \leq \epsilon$, \mathbf{x}^* je řešením úlohy

$$\min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{vzhledem k} \quad \|\mathbf{A}\mathbf{z} - \mathbf{y}\|_2 \leq \epsilon.$$

Potom

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C_1 \frac{\sigma_k(\mathbf{x})_1}{\sqrt{k}} + C_2 \epsilon$$

pro kladné konstanty C_1, C_2 závisející pouze na δ_{2k} .

Chyba je tedy rozdělena na část závisející na řídkosti a část podmíněnou energií šumu.

Deterministické matice, které by splňovaly RIP s předem danými parametry se dosud nepodařilo zkonstruovat, pouze matice, které tuto vlastnost splňují s vysokou pravděpodobností [19].

Relaxační algoritmy tedy využívají předchozích podmínek k předpokladu, že se dostaneme k řešení alespoň relativně blízkému přesnému řešení.

Basis Pursuit (BP)

[8],[19]. Basis pursuit hledá nejlepší reprezentaci signálu ℓ_1 -minimalizací složek vektoru \mathbf{x} . Řeší tedy problém (P1)

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k} \quad \mathbf{A}\mathbf{x} = \mathbf{y}.$$

Užití ℓ_1 -normy nám umožní předepsat váhu každému atomu, který v reprezentaci používáme. Například váhu přiřazujeme proporcionálně velikosti koeficientu. Protože máme dodatečnou podmínku k řešení systému rovnic, můžeme problém přepsat jako úlohu lineární optimalizace tvaru

$$\min_{\mathbf{x}} \mathbf{c}^\top \mathbf{x} \quad \text{vzhledem k} \quad \mathbf{A}\mathbf{x} = \mathbf{y}, \quad \mathbf{x} \geq 0,$$

kde $\mathbf{c}^\top \mathbf{x}$ je účelová funkce, $\mathbf{A}\mathbf{x} = \mathbf{y}$ systém omezení typu rovnost a $\mathbf{x} \geq 0$ množina ohraničení.

V našem případě vypustíme omezení $\mathbf{x} \geq 0$ a naše účelová funkce je $\|\mathbf{x}\|_1 = |x_1| + \dots + |x_n|$. Toto sice ještě není úloha LP, ale nelinearitu můžeme převést na množinu omezení přidáním n proměnných t_1, \dots, t_n . Účelová funkce tedy přejde na $t_1 + \dots + t_n$ a omezení budou tvaru $|x_i| \leq t_i$, $\mathbf{A}\mathbf{x} = \mathbf{y}$. Dalšími úpravami získáme úlohu LP [8], řešitelnou mnoha algoritmy (např. simplexovou metodou).

4.2. ALGORITMY HLEDÁNÍ ŘÍDKÝCH REPREZENTACÍ

Iterative Reweighted Least Squares (IRLS)

[14]. Tato metoda reprezentuje ℓ_p -normu po pevně dané $0 < p \leq 1$ jako váženou ℓ_2 -normu. V iterativním algoritmu za předpokladu aktuálního přibližného řešení \mathbf{x}_{k-1} položíme $\mathbf{X}_{k-1} = \text{diag}(|\mathbf{x}_{k-1}|^q)$. Dále předpokládejme, že je tato matice invertibilní, $\|\mathbf{X}_{k-1}^{-1}\mathbf{x}\|_2^2 = \|\mathbf{x}\|_{2-2q}^{2-2q}$. Tedy zvolením $q = 1 - p/2$ tento výraz napodobuje ℓ_p -normu $\|\mathbf{x}\|_p^p$.

Namísto inverze \mathbf{X}_{k-1} výše použijeme pseudoinverzi $\|\mathbf{X}_{k-1}^+\mathbf{x}\|_2^2$. Dále se pokusíme řešit problém

$$\min_{\mathbf{x}} \|\mathbf{X}_{k-1}^+\mathbf{x}\|_2^2 \quad \text{vzhledem k} \quad \mathbf{A}\mathbf{x} = \mathbf{y}.$$

Pro $q = 1$ je toto varianta (P0) úlohy. Tento problém je však již možné řešit pomocí klasické lineární algebry a Lagrangeových multiplikátorů, poněvadž účelová funkce používá klasickou ℓ_2 -normu. Tedy

$$\begin{aligned} L(\mathbf{x}) &= \|\mathbf{X}_{k-1}^+\mathbf{x}\|_2^2 + \lambda^\top(\mathbf{y} - \mathbf{A}\mathbf{x}), \\ \Rightarrow \frac{\partial L(\mathbf{x})}{\partial \mathbf{x}} &= \mathbf{0} = 2(\mathbf{X}_{k-1}^+)^2\mathbf{x} - \mathbf{A}^\top\lambda. \end{aligned}$$

Pokud předpokládáme, že \mathbf{X}_{k-1}^+ je invertibilní, má nulové prvky na hlavní diagonále [14], a platí pak $(\mathbf{X}_{k-1}^+)^{-1} = \mathbf{X}_{k-1}$. V tomto případě je řešení

$$\mathbf{x}_k = \frac{1}{2} \cdot \mathbf{X}_{k-1}^2 \mathbf{A}^\top \lambda.$$

Obecně, pokud jsou některé prvky na hlavní diagonále \mathbf{X}_{k-1} nulové, řešení nuluje odpovídající prvky \mathbf{x}_k . Protože kvůli šíření řídkosti tyto nulové složky \mathbf{x}_k mohou zůstat v následujících iteracích, je tento vzorec stále platný. Řešení \mathbf{x}_k pak dosadíme do $\mathbf{A}\mathbf{x} = \mathbf{y}$ a dostáváme

$$\frac{1}{2} \cdot \mathbf{A}\mathbf{X}_{k-1}^2\mathbf{A}^\top\lambda = \mathbf{y} \quad \Rightarrow \quad \lambda = 2(\mathbf{A}\mathbf{X}_{k-1}^2\mathbf{A}^\top)^{-1}\mathbf{y}.$$

Zde je opět nutné nahradit inverzi pseudoinverzí [14]. Potom

$$\mathbf{x}_k = \mathbf{X}_{k-1}^2 \mathbf{A}^\top (\mathbf{A}\mathbf{X}_{k-1}^2 \mathbf{A}^\top)^+ \mathbf{y}.$$

Aproximované řešení dosadíme do diagonální matice \mathbf{X}_k a pokračujeme další iterací.

Tento algoritmus konverguje, ovšem ne nutně k optimálnímu řešení. Také z poslední rovnice plyne, že ve chvíli, kdy je složka \mathbf{x}_k vynulována, již tak zůstane. Proto by počáteční aproximace měla mít všechny složky nenulové.

Ve chvíli, kdy je původní úloha (P0) nahrazena relaxovanou verzí, je potřeba řešit normalizaci sloupců matice \mathbf{A} . Zatímco ℓ_0 -norma nereflktuje velikost nenulových složek, ℓ_p -normy penalizují složky s větší velikostí. Algoritmus tím pádem ovlivňuje řešení tak, že vybírá za nenulové složky ty, které v ℓ_p -normě násobí největší atomy matice \mathbf{A} . Proto je matice před výpočtem normována. Pro zajištění ekvivalence s řešením původní matice je potřeba řešení na konci po složkách vydělit příslušnými normami matice \mathbf{A} .

4.2.2. Hladové (greedy) algoritmy

[14]. Hladové algoritmy jsou algoritmy iterační. V každé iteraci algoritmus zvolí jeden nebo skupinu atomů, které v dané iteraci nejlépe aproximují signál. Vybraný atom (atomy) se

přidá do aktuální aproximace a poté se počítá aproximace další (již s tímto atomem). Ukončující podmínkou bývá nastavení počtu iterací, popřípadě při poklesu chyby aproximace pod danou prahovou hodnotu.

Výhodou je výpočetní jednoduchost, nevýhodou je však neschopnost zaručení dosažení globálního optima.

Uvažujeme problém (P0). Předpokládáme, že $\text{spark}(\mathbf{A}) > 2$ a existuje jediné 1-řádké řešení systému $\mathbf{Ax} = \mathbf{y}$, tzn. \mathbf{y} je skalárním násobkem nějakého sloupce matice \mathbf{A} . Tento sloupec dokážeme identifikovat v m krocích. j -tý test tedy provedeme minimalizací odchylky $\epsilon(j) = \|\mathbf{a}_j z_j - \mathbf{y}\|_2$, což dává $z_j^* = \mathbf{a}_j^\top \mathbf{y} / \|\mathbf{a}_j\|_2^2$. Jestliže je po dosazení do $\epsilon(j) = \min_{z_j} \|\mathbf{a}_j z_j - \mathbf{y}\|_2^2$ chyba rovna nule, našli jsme přesné řešení [14].

Obecněji předpokládáme $\text{spark}(\mathbf{A}) > 2k_0$ a optimalizační problém má k_0 -řádké řešení. Bohužel nalezení přesného řešení je výpočetně velmi náročné. Hladové algoritmy využijí lokálně optimálního výběru. Počínaje v $\mathbf{x}^0 = 0$ v každé iteraci přidáváme vybraný atom, který má nejmenší reziduální ℓ_2 -chybu při aproximaci \mathbf{y} . Ve chvíli, kdy celková reziduální ℓ_2 -chyba klesne po daný práh, algoritmus se ukončí (popřípadě ukončení nastane po provedení daného počtu iterací).

Matching Pursuit (MP)

Tato metoda konverguje k řešení, často ovšem pomalu. Je jedna z prvních metod použitých k hledání řídkých řešení [14].

Na začátku je nastaveno řešení $\mathbf{x}^0 = 0$, počáteční reziduum $\mathbf{r}^0 = \mathbf{y} - \mathbf{Ax} = \mathbf{y}$ a výchozí nosič $S^0 = \text{supp}(\mathbf{x}) = \emptyset$.

- V rámci prvního kroku minimalizujeme chybu $\epsilon(j) = \min_{z_j} \|\mathbf{a}_j z_j - \mathbf{r}^{k-1}\|_2^2$ užitím $z_j^* = \mathbf{a}_j^\top \mathbf{r}^{k-1} / \|\mathbf{a}_j\|_2^2$.
- Poté aktualizujeme nosič S^k přidáním pozice optimálního atomu (ozn. j_0) do S^{k-1} .
- Vektor řešení nastavíme $\mathbf{x}^k = \mathbf{x}^{k-1}$ a aktualizujeme $x^k(j_0) = x^{k-1}(j_0) + z_{j_0}^*$.
- Vypočítáme nové reziduum $\mathbf{r}^k = \mathbf{y} - \mathbf{Ax}^k = \mathbf{r}^{k-1} - z_{j_0}^* \mathbf{a}_{j_0}$.
- Nakonec testujeme ukončovací kritérium $\|\mathbf{r}^k\|_2 < \epsilon_0$.

Nevýhodou je, že pokud algoritmus vybere z nějakých důvodů v prvním kroku špatný atom, v následných iteracích toto nelze opravit. Existují také signály, které tímto algoritmem nelze optimálně řídkce reprezentovat.

Orthogonal Matching Pursuit (OMP)

[14]. Tento algoritmus je z hladových algoritmů asi nejpoužívanější. Je variantou předchozího MP, kdy navíc v každé iteraci provádíme ortogonalizaci rezidua.

- Vektor řešení aktualizujeme na \mathbf{x}^k minimalizací $\|\mathbf{Ax} - \mathbf{y}\|_2^2$ vzhledem k $\text{supp}(\mathbf{x}) = S^k$:

4.3. PROXIMÁLNÍ ALGORITMY

- Označíme \mathbf{A}_{S^k} jako matici obsahující sloupce matice \mathbf{A} , které patří do nosiče S^k . Tedy minimalizujeme $\|\mathbf{A}_{S^k}\mathbf{x}_{S^k} - \mathbf{y}\|_2^2$, kde \mathbf{x}_{S^k} je nenulová část vektoru \mathbf{x} . Řešení je dáno

$$\mathbf{A}_{S^k}^\top (\mathbf{A}_{S^k}\mathbf{x}_{S^k} - \mathbf{y}) = -\mathbf{A}_{S^k}^\top \mathbf{r}^k = 0.$$

Reziduum je totiž $\mathbf{r}^k = \mathbf{y} - \mathbf{A}\mathbf{x}^k = \mathbf{y} - \mathbf{A}_{S^k}\mathbf{x}_{S^k}$.

Sloupce \mathbf{A} , které jsou součástí nosiče S^k jsou nutně ortogonální k reziduu \mathbf{r}^k . V dalších iteracích tedy tyto sloupce pro nosič nebudou vybrány. Z této ortogonalizace plyne název metody.

Další metody

Mezi ostatní metody, které se neřadí do jedné z předchozích skupin, patří například algoritmy založené na *prahování* (thresholding). Další jsou pak tzv. *hybridní*, které využívají kombinace relaxace a hladových metod [14].

4.3. Proximální algoritmy

Proximální algoritmy spadají do skupiny algoritmů využívajících ℓ_1 -relaxaci. Poněvadž jsou ale právě tyto používány v experimentální části, věnujeme jim zde samostatnou sekci.

Mnoho úloh zpracování signálu lze formulovat jako problém konvexní optimalizace tvaru

$$\min_{\mathbf{x}} f_1(\mathbf{x}) + f_2(\mathbf{x}) + \dots + f_m(\mathbf{x}), \quad (4.3)$$

kde f_1, \dots, f_m jsou konvexní funkce z \mathbb{R}^N do \mathbb{R} . Problém při řešení této úlohy nastává, pokud některé z těchto funkcí nejsou diferencovatelné, a tedy nemůžeme použít klasické optimalizační metody (např. metoda sdružených gradientů a další). Proximální metody rozdělí účelovou funkci a pracují z každou z funkcí f_1, \dots, f_m individuálně. Nazývají se *proximální*, protože každou nehladkou funkci nahrazují jejím proximálním operátorem. Před definováním proximálního operátoru uvedeme několik pojmů pro tuto definici potřebných.

Definice 21 [11] *Nechť C je neprázdná konvexní podmnožina prostoru \mathbb{R}^N . Charakteristická (indikátorová) funkce této množiny je*

$$v_C : \mathbf{x} \mapsto \begin{cases} 0 & \mathbf{x} \in C \\ +\infty & \mathbf{x} \notin C \end{cases}$$

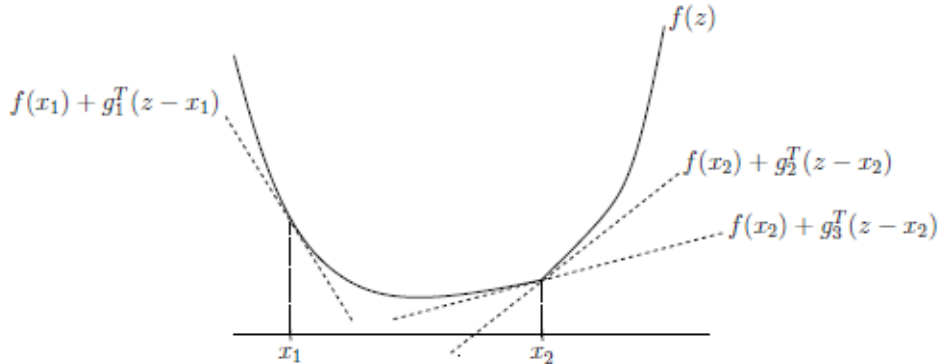
Definice 22 [11] *Nechť vzdálenost bodu $\mathbf{x} \in \mathbb{R}^N$ od množiny C je $d_C(\mathbf{x}) = \inf_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$. Pokud je C uzavřená a konvexní, projekce $\mathbf{x} \in \mathbb{R}^N$ na C je jediný vektor $P_C\mathbf{x} \in C$ takový, že*

$$d_C(\mathbf{x}) = \|\mathbf{x} - P_C\mathbf{x}\|.$$

Definice 23 [6] *Nechť $f : U \rightarrow \mathbb{R}$ je reálná konvexní funkce definovaná na konvexní otevřené podmnožině prostoru \mathbb{R}^N . Vektor $\mathbf{u} \in \mathbb{R}^N$ se nazývá subgradient funkce f v bodě \mathbf{x}_0 množiny U , pokud $\forall \mathbf{x} \in U$*

$$\mathbf{u}^\top (\mathbf{x} - \mathbf{x}_0) + f(\mathbf{x}_0) \leq f(\mathbf{x}).$$

Subgradient je zobecněným pojmem gradientu funkce, v bodech, kde je f diferencovatelná, se subgradient a gradient rovnají. Ilustrace na obr. 4.3.



Obrázek 4.3: V bodě x_1 je funkce f diferencovatelná, subgradient g_1 je jediným subgradientem v x_1 . V bodě x_2 není f diferencovatelná, má zde tedy mnoho subgradientů: na obrázku jsou znázorněny například subgradienty g_2, g_3 [6].

Definice 24 [32] Uvažujme funkci $f : \mathbb{R}^N \rightarrow \mathbb{R}$ a vektor $\mathbf{x}_0 \in \mathbb{R}^N$. Řekneme, že funkce f je v bodě \mathbf{x}_0 polospojité zdola, jestliže platí

$$f(\mathbf{x}_0) \leq \liminf_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}).$$

Definice 25 [6] Nechť $f : \mathbb{R}^N \rightarrow \mathbb{R}$ je zdola polospojité konvexní funkce s neprázdným definičním oborem. Subdiferenciál funkce f je operátor

$$\partial f : \mathbb{R} \rightarrow 2^{\mathbb{R}^N} : \mathbf{x} \mapsto \{\mathbf{u} \in \mathbb{R}^N \mid (\forall \mathbf{y} \in \mathbb{R}^N) \mathbf{u}^\top (\mathbf{y} - \mathbf{x}) + f(\mathbf{x}) \leq f(\mathbf{y})\}.$$

Subdiferenciál $\partial f(\mathbf{x}_0)$ je množina všech subgradientů funkce f v \mathbf{x}_0 .

Například, pokud $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|$, subdiferenciál v bodě $x = 0$ je interval $[-1, 1]$, v bodě $x < 0$ je to jednoprvková množina $\{-1\}$.

Proximální operátor

Definice 26 [11] Nechť funkce $f : \mathbb{R}^N \rightarrow \mathbb{R}$ je zdola polospojité, konvexní a s neprázdným definičním oborem. Pro každé $\mathbf{x} \in \mathbb{R}^N$ má minimalizační problém

$$\min_{\mathbf{y} \in \mathbb{R}^N} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + f(\mathbf{y}) \quad (4.4)$$

právě jedno řešení, které se označuje $\text{prox}_f \mathbf{x}$. Takto definovaný operátor $\text{prox}_f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ nazýváme proximální operátor funkce f .

Proximální operátor lze charakterizovat inkluzí:

$$\mathbf{p} = \text{prox}_f(\mathbf{x}) \Leftrightarrow \mathbf{x} - \mathbf{p} \in \partial f(\mathbf{p}), \quad \forall (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^N \times \mathbb{R}^N.$$

Pokud je funkce f diferencovatelná, inkluze se redukuje na

$$\mathbf{p} = \text{prox}_f(\mathbf{x}) \Leftrightarrow \mathbf{x} - \mathbf{p} \in \nabla f(\mathbf{p}), \quad \forall (\mathbf{x}, \mathbf{p}) \in \mathbb{R}^N \times \mathbb{R}^N.$$

4.3. PROXIMÁLNÍ ALGORITMY

Proximální operátor je zobecněním projekce $P_C \mathbf{x}$ na uzavřenou konvexní podmnožinu $C \subset \mathbb{R}^N$. Ta je řešením problému

$$\min_{\mathbf{y} \in \mathbb{R}^N} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \iota_C(\mathbf{y}).$$

Funkce ι_C je zdola polospojité konvexní funkce, splňuje tedy i předpoklady proximálního operátoru.

Vlastnosti proximálních operátorů je dělájí vhodnými prostředky pro užití v iteračních minimalizačních algoritmech. Dvě z těchto vlastností si uvedeme, více [11]:

Tvrzení 9 (Vlastnosti proximálního operátoru [11]) Předpokládejme, že funkce $\varphi : \mathbb{R}^N \rightarrow \mathbb{R}$ je zdola polospojité konvexní funkce s neprázdným definičním oborem, $C \subset \mathbb{R}^N$ je neprázdná, uzavřená a konvexní, $\mathbf{x} \in \mathbb{R}^N$.

- Posunutí: $f(\mathbf{x}) = \varphi(\mathbf{x} - \mathbf{z})$, $\mathbf{z} \in \mathbb{R}^N$, $\text{prox}_f(\mathbf{x}) = \mathbf{z} + \text{prox}_\varphi(\mathbf{x} - \mathbf{z})$.
- Změna měřítka: $f(\mathbf{x}) = \varphi(\mathbf{x}/\rho)$, $\rho \in \mathbb{R} \setminus \{0\}$, $\text{prox}_f(\mathbf{x}) = \rho \text{prox}_{\varphi/\rho^2}(\mathbf{x}/\rho)$.

Z pohledu zpracování signálů mají proximální operátory přirozenou interpretaci při odstraňování šumu ze signálu [11]. Uvažujme klasickou úlohu odstraňování šumu: hledáme původní signál $\tilde{\mathbf{x}} \in \mathbb{R}^N$ z pozorování

$$\mathbf{y} = \tilde{\mathbf{x}} + \mathbf{w},$$

kde $\mathbf{w} \in \mathbb{R}^N$ představuje šum. Problém lze formulovat jako (4.4), kde $\|\cdot - \mathbf{y}\|_2^2 / 2$ zajišťuje blízkost dat a f modeluje požadavek na $\tilde{\mathbf{x}}$.

Proximální operátor pro ℓ_1 -normu

Pro získání proximálního operátoru pro ℓ_1 -normu v definici (4.4) dosadíme za funkci $f(\mathbf{y}) = \lambda \|\mathbf{y}\|_1$:

$$\min_{\mathbf{y}} \underbrace{\frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{y}\|_1}_{h(\mathbf{y})},$$

a tedy $\text{prox}_{\lambda \|\cdot\|_1}(\mathbf{x})$ je řešení tohoto problému. Pro nalezení minima je potřebujeme kořeny gradientu funkce $h(\mathbf{y})$. Oba členy funkce jsou separovatelné v \mathbf{y} , a lze tedy minimalizovat každý prvek vektoru \mathbf{y} zvlášť. Zjednodušeně pak můžeme napsat

$$h(y_i) = \frac{1}{2}(y_i - x_i)^2 + \lambda |y_i|, \quad \forall i = 1, \dots, N.$$

Abychom našli derivaci $|y_i|$, máme 2 případy:

1. $y_i > 0$:

$$\begin{aligned} y_i - x_i + \lambda &= 0, \\ y_i = x_i - \lambda > 0 &\Rightarrow x_i > \lambda \end{aligned}$$

2. $y_i < 0$:

$$y_i - x_i - \lambda = 0,$$

$$y_i = x_i + \lambda < 0 \quad \Rightarrow \quad x_i < -\lambda$$

Případ $y_i = 0$ je pak minimum pro $-\lambda \leq x_i \leq \lambda$.

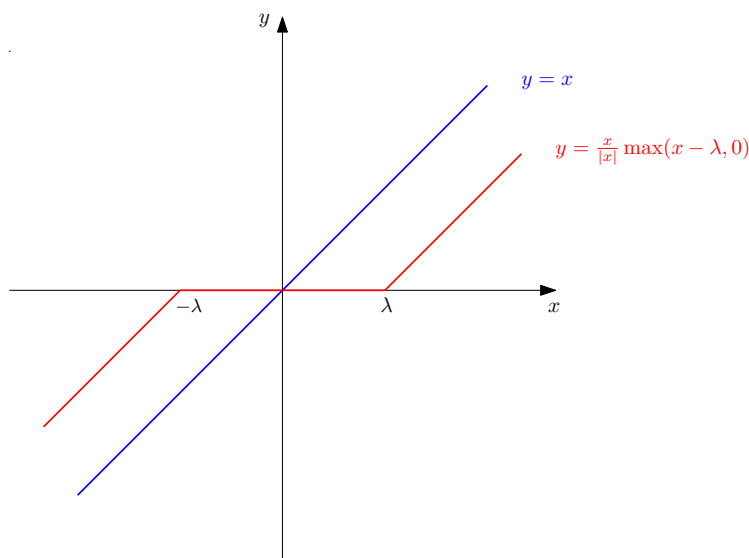
Dohromady

$$y_i = \begin{cases} x_i - \lambda & x_i > \lambda \\ x_i + \lambda & x_i < -\lambda \\ 0 & -\lambda \leq x_i \leq \lambda, \end{cases}$$

což je ekvivalentní s

$$y_i = \text{prox}_h(x_i) = \text{sgn}(x_i) \cdot \max(|x_i| - \lambda, 0) = \frac{x_i}{|x_i|} \max(|x_i| - \lambda, 0).$$

Proximální operátor pro ℓ_1 -normu je tedy funkcí měkkého prahování (obr. 4.4).



Obrázek 4.4: Červeně funkce měkkého prahování.

Proximální operátor pro smíšenou normu

Uvažujme, že prvky \mathbf{y} jsou rozděleny do 2 skupin a bereme normu $\|\cdot\|_{2,1}$. Problém je pak

$$\min_{\mathbf{y}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \left(\|y_1, \dots, y_k\|_2, \|y_{k+1}, \dots, y_N\|_2 \right)_1,$$

tedy

$$\min_{\mathbf{y}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \left\| \left(\sqrt{y_1^2 + \dots + y_k^2}, \sqrt{y_{k+1}^2 + \dots + y_N^2} \right) \right\|_1,$$

a nakonec

$$\min_{\mathbf{y}} \frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \left(\sqrt{y_1^2 + \dots + y_k^2} + \sqrt{y_{k+1}^2 + \dots + y_N^2} \right).$$

4.4. STRUKTUROVANÁ ŘÍDKOST

Proximální operátor pro prvky ze skupiny ($i = 1, \dots, k$ nebo $i = k + 1, \dots, N$) je měkké prahování na každé skupině, tj.

$$y_i = \frac{x_i}{\|\mathbf{x}_g\|_2} \max(\|\mathbf{x}_g\|_2 - \lambda, 0),$$

\mathbf{x}_g jsou složky tvořící jednu skupinu [38].

Proximální gradientní metoda (Forward-Backward splitting)

Nyní uvedeme metodu tzv. *dopředného-zpětného dělení* [11]. Uvažujme případ $m = 2$ funkcí v (4.3), přičemž jedna z nich je hladká.

Máme danou úlohu

$$\min_{\mathbf{x}} f_1(\mathbf{x}) + f_2(\mathbf{x}). \quad (4.5)$$

Funkce $f_2 : \mathbb{R}^N \rightarrow \mathbb{R}$ je zdola polospojitá konvexní funkce a $f_1 : \mathbb{R}^N \rightarrow \mathbb{R}$ je konvexní a diferencovatelná s β -Lipschitzovským spojitým gradientem, tj.

$$\|\nabla f_1(\mathbf{x}) - \nabla f_1(\mathbf{y})\| \leq \beta \|\mathbf{x} - \mathbf{y}\|, \quad \forall (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^N \times \mathbb{R}^N,$$

kde $\beta \in \mathbb{R}^+$. Lze ukázat [11], že úloha má alespoň jedno řešení, pokud $f_1(\mathbf{x}) + f_2(\mathbf{x}) \rightarrow +\infty$ pro $\|\mathbf{x}\| \rightarrow +\infty$. Pro nějaké $\gamma \in \mathbb{R}^+$ je řešení charakterizováno

$$\mathbf{x} = \text{prox}_{\gamma f_2}(\mathbf{x} - \gamma \nabla f_1(\mathbf{x})).$$

Toto vede na možnost iterace

$$\mathbf{x}^{k+1} = \underbrace{\text{prox}_{\gamma^k f_2}}_{\text{zpětný krok}} \left(\underbrace{\mathbf{x}^k - \gamma^k \nabla f_1(\mathbf{x}^k)}_{\text{dopředný krok}} \right) \quad (4.6)$$

pro vhodně ohraničené velikosti kroku γ^k . Schéma se tedy dělí na dopředný (explicitní) gradientní krok užitím funkce f_1 a zpětný (implicitní) krok užitím funkce f_2 . Pokud $f_2 = 0$, získáme *gradientní metodu*

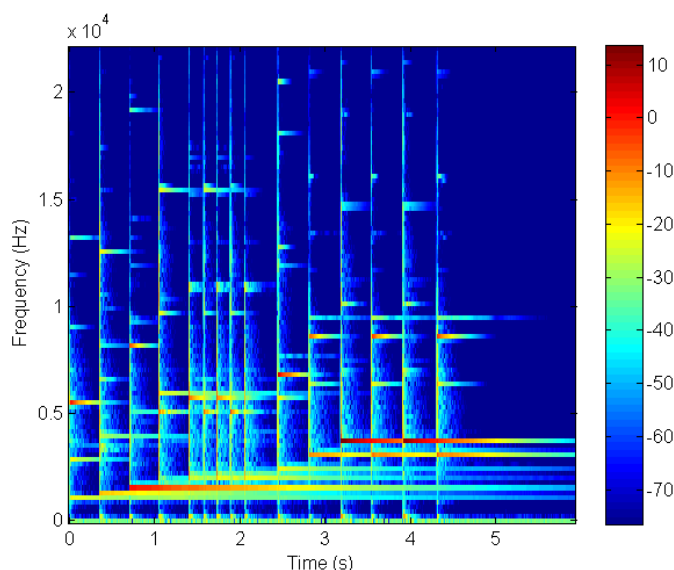
$$\mathbf{x}^{k+1} = \mathbf{x}^k - \gamma^k \nabla f_1(\mathbf{x}^k)$$

pro minimalizaci funkce s Lipschitzovským spojitým gradientem, a pokud naopak $f_1 = 0$, získáme *metodu proximálního bodu*

$$\mathbf{x}^{k+1} = \text{prox}_{\gamma^k f_2}(\mathbf{x}^k)$$

pro minimalizaci nediferencovatelné funkce [11].

Pokud je délka kroku $\gamma^k < \frac{2}{\beta}$, posloupnost $\{\mathbf{x}^k\}$ generována algoritmem (4.6) konverguje k řešení úlohy (4.5) [11].

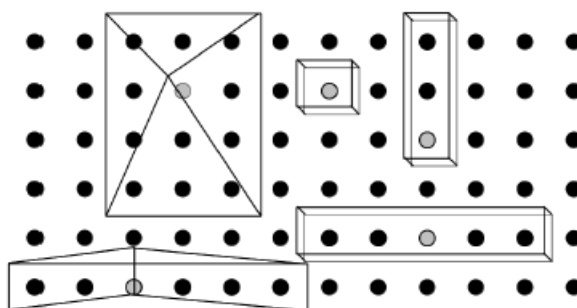


Obrázek 4.5: Signál zvonkohry ('glockenspiel'). Pozorujeme strukturu v čase (úderů) i kmitočtu (harmonické složky).

4.4. Strukturovaná řídkost

Strukturovaná řídkost na rozdíl od klasické řídkosti využívá faktu, že většina zvukových signálů (řeč, hudba) je strukturovaná v čase i kmitočtu [33]. Ve vertikálním směru pozorujeme strukturu v čase způsobenou přechodovými jevy (transienty), v horizontálním směru pak frekvenční strukturu v harmonických složkách (obr. 4.5).

V předchozí části textu jsme chápali řídkost jako koeficient po koeficientu, přístup strukturované řídkosti uvažuje koeficienty zformované do skupin, a dále pracuje pouze s těmito skupinami (obr. 4.6). Platí totiž, že jestliže je jeden koeficient významný (nenulový), bývají významné i ty v jeho blízkosti.



Obrázek 4.6: Různé možnosti vytváření skupin v časově-kmitočtové rovině [33].

Strukturovaná řídkost se dělí na několik typů [5],[23],[24]:

- *Elitářská řídkost (Elitist)* – z každé skupiny vybrání pouze nejvýznamnější zástupci
- *Skupinová řídkost (Group)* – pouze několik skupin s největší energií významných
- *Simultánní řídkost (Joint)* – více signálů se stejnou strukturovanou řídkostí v dané reprezentaci

4.5. KOMPRIMOVANÉ SNÍMÁNÍ

- *Sociální řídkost (Social)* – jeden koeficient leží zároveň ve více skupinách, tedy se mohou skupiny překrývat

Při využití strukturované řídkosti můžeme problém formulovat podobně jako pro řídkost koeficientů po koeficientu a používat například upravené hladové algoritmy či relaxační algoritmy.

Tento pohled na řídkost bere v úvahu dvourozměrnost časově-kmitočtových reprezentací. Při relaxaci využíváme smíšených norem aplikovaných na skupiny koeficientů [33]. Pokud jsme schopni srovnat koeficienty do matice \mathbf{X} , relaxační úloha (P1) je pak tvaru

$$\min_{\mathbf{X}} \|\mathbf{X}\|_{1,2} \quad \text{vzhledem k} \quad \mathbf{A} \cdot \text{vec}(\mathbf{X}) = \mathbf{y},$$

kde $\text{vec}(\mathbf{X})$ je vektorizace matice \mathbf{X} , převádějící matici na sloupcový vektor [37]. V tomto případě ($p = 1, q = 2$) máme elitářskou řídkost - ℓ_2 -norma aplikovaná na skupiny (sloupce) jejich řídkost nevynucuje, avšak ℓ_1 -norma vynucuje řídkost v řádcích, tedy v každé skupině je vybrán nejvýznamnější zástupce.

Volba $p = 2, q = 1$ dává skupinovou řídkost - ℓ_2 -norma aplikovaná na jednotlivé koeficienty (řádky) řídkost nevynucuje, avšak ℓ_1 -norma by nutí řídkost sloupců, tedy skupin.

4.5. Komprimované snímání

Nyní krátce shrneme principy jedné z důležitých aplikací řídkých reprezentací signálů. Běžná metoda komprese dat byla doposud tato: Nejdříve naměříme všechna data, provedeme nějakou vhodnou transformaci a vyhodnotíme získané koeficienty. Většinu z těchto koeficientů dále nevyužijeme, poněvadž z nějakého důvodu nesou málo informace. V kompresním formátu JPEG například získáme hodnoty všech pixelů fotografie, provedeme dvojrozměrnou diskretní kosinovou transformaci DCT a poté koeficienty kvantujeme. Nejčastěji pak ponecháme pouze koeficienty s největší velikostí a zakódujeme jejich velikost a pozici.

Komprimované snímání (compressed sensing, compressive sampling, CS) využívá jiný přístup. Snímá signál lineárně a neadaptivně pouze tolikrát, kolik je potřeba. Rekonstrukce je již časově náročnější kvůli své nelinearitě.

V komprimovaném snímání řešíme stejný problém, jako v předchozí kapitole

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{vzhledem k} \quad \mathbf{A}\mathbf{x} = \mathbf{y},$$

prakticky pak aproximovaný problém

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k} \quad \mathbf{A}\mathbf{x} = \mathbf{y},$$

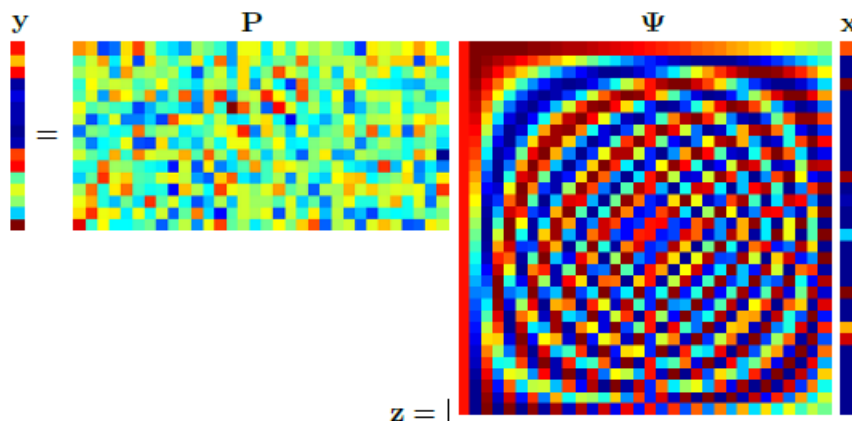
nyní ovšem bude mít matice \mathbf{A} speciální tvar [18].

Předpokládáme, že signál je řídký v nějakém slovníku (zde ortonormální báze). Báze je označena Ψ , signál \mathbf{z} má pak vyjádření $\mathbf{z} = \Psi\mathbf{x}$, kde vektor \mathbf{x} je k -řídký. Chceme provést „malý počet“ neadaptivních měření, které mají charakter skalárních součinů se signálem, toto můžeme zapsat jako $\mathbf{y} = \mathbf{P}\mathbf{z} = \mathbf{P}\Psi\mathbf{x}$. \mathbf{P} označuje tzv. měřicí matici rozměrů $m \times N$,

jednotlivé složky \mathbf{y} jsou výsledky měření vzniklé jako lineární kombinace vzorků signálu. Počet měření je $m \ll N$. Problém (P1) lze tedy zapsat

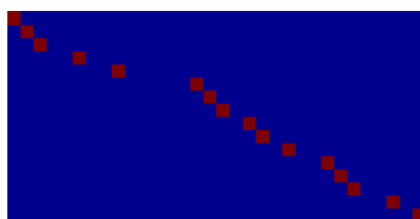
$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k} \quad \mathbf{y} = \mathbf{P}\Psi\mathbf{x}.$$

Řešení tohoto problému označíme \mathbf{x}^1 , signál \mathbf{z} lze pak získat jednoduše $\mathbf{z} = \Psi\mathbf{x}^1$ (obr. 4.7).



Obrázek 4.7: Vektor měření \mathbf{y} je součinem měřicí matice $\mathbf{P} = \mathbf{R}\Phi$, matice Ψ a vektoru \mathbf{x} , přičemž $\mathbf{z} = \Psi\mathbf{x}$ (signál je řádký ve slovníku Ψ , zde ortonormální báze zpětné DCT) [18].

Měřicí matice \mathbf{P} se většinou uvažují tvaru $\mathbf{P} = \mathbf{R}\Phi$ [18]. Φ je nějaká matice rozměrů $N \times N$ a matice \mathbf{R} vznikne z matice jednotkové rozměrů $N \times N$ ponecháním m náhodně vybraných řádků. Tato matice tedy funguje jako výběr řádků z matice Φ , obr. 4.8. Nejčastěji se tento výběr řídí rovnoměrným rozdělením pravděpodobnosti. Tedy matice $\mathbf{A} = \mathbf{R}\Phi\Psi$.



Obrázek 4.8: Příklad matice \mathbf{R} [18].

Nyní je potřeba řešit, kolik měření m provést a jak volit matici Φ , abychom mohli signál rekonstruovat ℓ_1 -optimalizací. Je nutné uvést pojem vzájemné koherence μ pro matici, která je složena ze dvou ortonormálních bází Φ a Ψ , tj. $[\Phi, \Psi]$ [18].

$$\mu([\Phi, \Psi]) = \max_{1 \leq i, j \leq N} |\psi_i^\top \phi_j|,$$

hodnota μ se pohybuje mezi $\frac{1}{\sqrt{N}}$ a 1. Následuje podmínka, která zaručuje přesnou rekonstrukci z m měření.

Tvrzení 10 [7] *Nechť je dán signál \mathbf{z} , který má v bázi Ψ k -řádkou reprezentaci \mathbf{x} . Pak řešení ℓ_1 -minimalizace*

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k} \quad \mathbf{y} = \mathbf{R}\Phi\Psi\mathbf{x},$$

4.5. KOMPRIMOVANÉ SNÍMÁNÍ

kde \mathbf{y} jsou měření, je současně s vysokou pravděpodobností nejřídší možné, pokud je m zvoleno

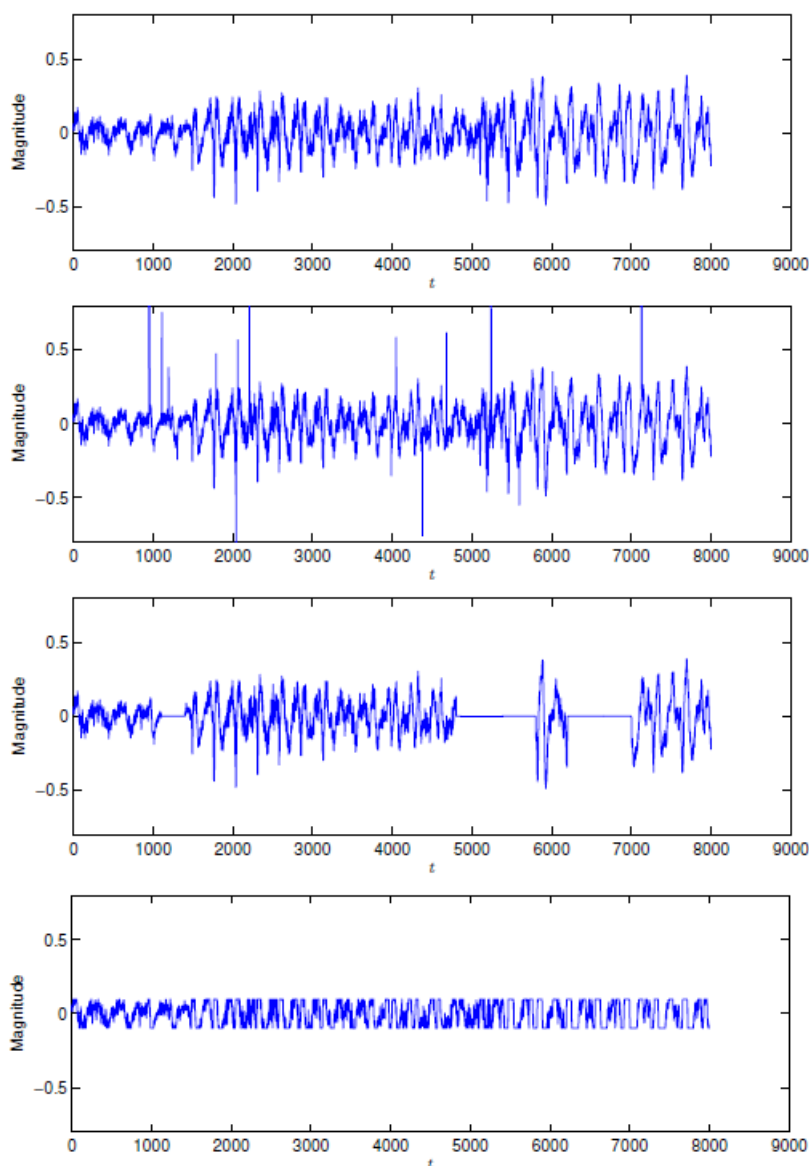
$$m \geq C \cdot \mu^2([\Phi, \Psi]) \cdot k \cdot N \cdot \ln N$$

pro nějakou kladnou konstantu C .

Tedy počet měření na řídkosti k závisí pouze lineárně, na koherenci kvadraticky – proto se opět snažíme hledat minimálně koherentní dvojice. Pokud koherence roste, využívání této podmínky pro měření přestává být výhodné. V jistém okamžiku pak přestává mít smysl, protože počet měření přeroste počet vzorků signálu N [7].

5. Doplnování chybějících dat do zvukového signálu

Zvukové signály jsou často při jejich zpracování nebo přenosu nechtěně ovlivněny. Vznikají v nich pak různé na poslech velmi nepříjemné artefakty. Například přehrávání poškrábané gramofonové desky či CD způsobí praskání (impulzní šum), chyby přenosu v bezdrátových telefonech či ve VoIP (Voice over IP) vedou ke ztrátě paketů, přebuzený zesilovač (nucený k většímu zesílení, než dokáže jeho napájení dodat) způsobí clipping – saturaci (zesilovač ořezává signál na své maximální zesílení – sinusoidální signál začíná připomínat obdélníkový) a další [1],[17]. Ilustrace na obr. 5.1.



Obrázek 5.1: Nahoře původní signál, dále narušený impulzním šumem (praskance), ztráta paketů a saturovaný signál [21].

5.1. Formulace problému

Pro daný signál $\mathbf{y} \in \mathbb{R}^N$ známe, které jeho vzorky jsou spolehlivé a které jsou porušené či chybějící. Tedy jsme schopni rozdělit nosič vektoru na dvě množiny: \mathcal{I}^r obsahující indexy spolehlivých vzorků a \mathcal{I}^m obsahující indexy chybějících vzorků. Platí $\mathcal{I}^r \cup \mathcal{I}^m = \{1, \dots, N\}$ a $\mathcal{I}^r \cap \mathcal{I}^m = \emptyset$.

Audio inpainting problém (doplňování dat do zvukového signálu) [1] pak spočívá v rekonstrukci vzorků \mathbf{y}^m na základě:

- spolehlivých vzorků \mathbf{y}^r ,
- dělení $\{\mathcal{I}^m, \mathcal{I}^r\}$,
- dodatečné informace o pozorovaném signálu,
- případné informace o chybějících datech.

V jednotkové $N \times N$ matici vymažeme řádky s indexy v \mathcal{I}^m , a tím získáme tzv. měřicí matici $\mathbf{M}^r \in \mathbb{R}^{|\mathcal{I}^r| \times N}$, která vybírá spolehlivé vzorky ze signálu

$$\mathbf{y}^r = \mathbf{M}^r \mathbf{y}.$$

Podobně, data, která chceme rekonstruovat, jsou $\mathbf{y}^m = \mathbf{M}^m \mathbf{y}$, kde $\mathbf{M}^m \in \mathbb{R}^{|\mathcal{I}^m| \times N}$.

Uvažujeme rekonstrukci chybějících vzorků v jednobandovém zvukovém signálu.

5.2. Dosavadní metody

V této části uvedeme metody používané k řešení tohoto problému [1],[21].

5.2.1. AR modely

První metoda spočívá ve využití *autoregresních modelů* [15],[16].

Autoregresní (AR) model signálu \mathbf{y} předpokládá, že hodnotu y_n signálu lze předpovědět z předchozích hodnot [30]:

$$y_n = \sum_{k=1}^p a_k y_{n-k} + e_n,$$

kde p je řád modelu, a_k koeficienty lineární predikce a e_n chyba mezi odhadem \tilde{y}_n a skutečným y_n . Tento typ modelování lze využít pro rekonstrukci dat interpolováním pomocí zpětné (směrem zprava) a dopředné (směrem zleva) predikce do chybějící části signálu. Tedy AR analýza je provedena před a za chybějící částí. Koeficienty lineární predikce a_k jsou přitom odhadovány klasickými algoritmy minimalizace střední kvadratické chyby (MSE) [29],[30]. Celkově je tedy postup

- Odhad parametrů AR modelu *před* chybějícím úsekem.
- „Dopředná“ lineární predikce signálu, zisk y_n^{left} .
- Odhad parametrů AR modelu *za* chybějícím úsekem.

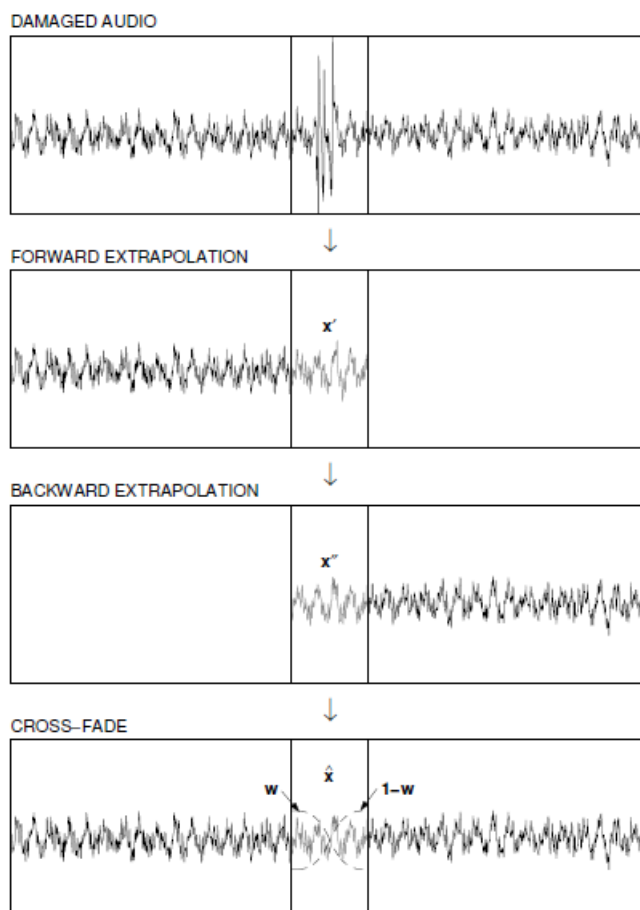
5. DOPLŇOVÁNÍ CHYBĚJÍCÍCH DAT DO ZVUKOVÉHO SIGNÁLU

- „Zpětná“ lineární predikce signálu, zisk y_n^{right} .
- Spojení obou predikcí a užití tzv. křížového prolnutí (cross-fading)

$$\tilde{y}_n = w_n y_n^{\text{left}} + (1 - w_n) y_n^{\text{right}},$$

kde w_n je funkce křížového prolnutí [20],[30].

Ilustrováno na obr. 5.2.



Obrázek 5.2: Poškozený úsek, predikce zleva, predikce zprava a nakonec spojení obou predikcí užitím funkce křížového prolnutí [20].

Sinusoidální modelování

Modifikace metody popsané výše spočívá v užití sinusoidálního modelování [25]. To reprezentuje signál jako sumu sinusoidálních komponent – složek proměnlivých frekvencí $f_k(t)$, fází $\phi_k(t)$ a amplitud $A_k(t)$ [26].

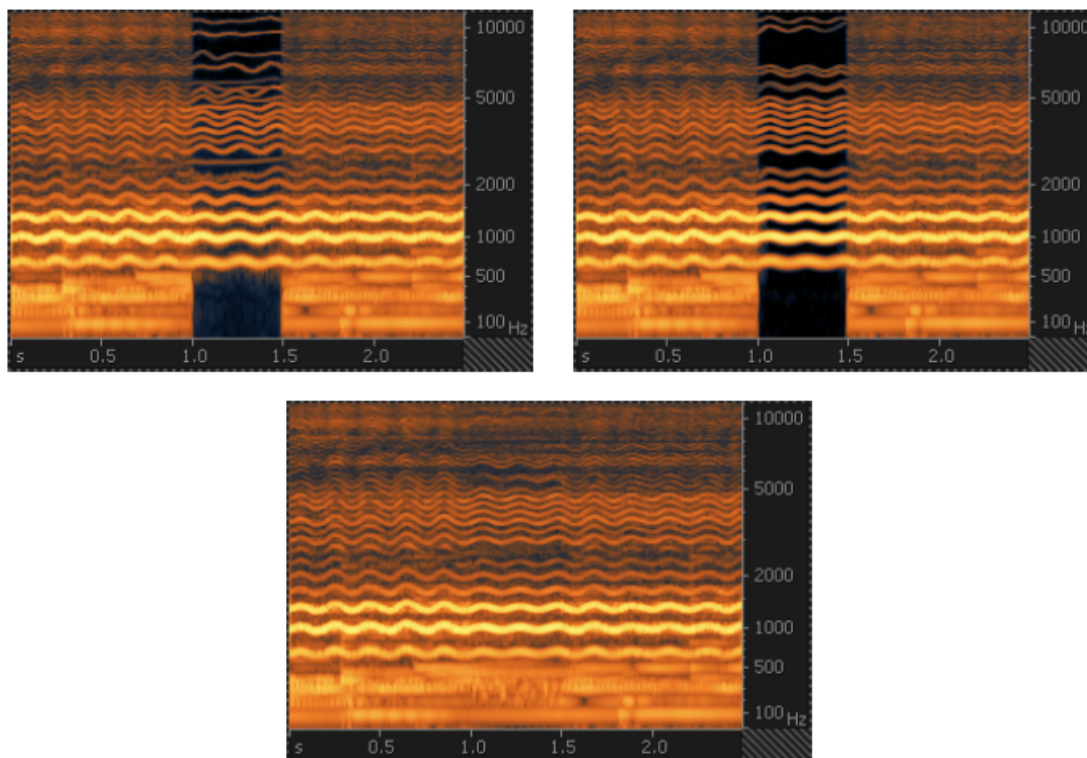
$$\mathbf{y}(t) = \sum_{k=1}^K A_k(t) \cos(\phi_k(t)),$$

$$\phi_k(t) = \phi_k(0) + 2\pi \int_0^t f_k(u) du.$$

5.2. DOSAVADNÍ METODY

AR modelem jsou odhadovány tyto parametry. Z nich je pak možné signál generovat. Signál v chybějícím úseku pak predikujeme stejným způsobem jako výše.

Publikace [26] uvádí vylepšení této metody zakomponováním reziduálního šumu do interpolace a prováděním odhadu parametrů z obou stran chybějícího úseku zároveň. Výsledek ilustrován na obr. 5.3. Navrhované změny přináší vyšší kvalitu interpolace a umožňují aplikaci tohoto algoritmu na širší spektrum audionahrávek.



Obrázek 5.3: Souběžná interpolace navržená publikací [26] vpravo, vs. předchozí metoda iterpolace vlevo. Dole pak kompletní výsledek nového algoritmu zahrnující interpolaci reziduálního šumu.

5.2.2. Neuronové sítě

Společným znakem metod umělé inteligence je jejich schopnost učit se z příkladů. Tím je postaven základ znalostí rozpoznávající známé situace. Tyto algoritmy mohou zobecnit získané znalosti ve smyslu rozpoznání dat z určité třídy, přestože jsou vztahy mezi nimi skryté. Proto je možné k předchozím metodám doplnění chybějících úseků přidat učící algoritmy [12].

Lineární predikce v AR modelech je v případě neuronových sítí nahrazena nelineárním užitím neuro-prediktorové funkce. Je zpracovávána posloupnost vzorků za účelem vygenerování následujícího vzorku v případě dopředné predikce a předcházejícího v případě zpětné predikce. Získaný vzorek je spojen s již existujícími a procedura se opakuje. Za předpokladu dostatečného počtu příkladů nenarušeného signálu v procesu učení může metoda interpolace využívající neuronové sítě zaručit dostatečnou přesnost. Podrobnější popis této metody je uveden například v [12],[36].

5.2.3. Metody využívající řídkou reprezentaci

Předpokládáme, že je signál pomocí řídké reprezentace dobře aproximován ve vhodném slovníku

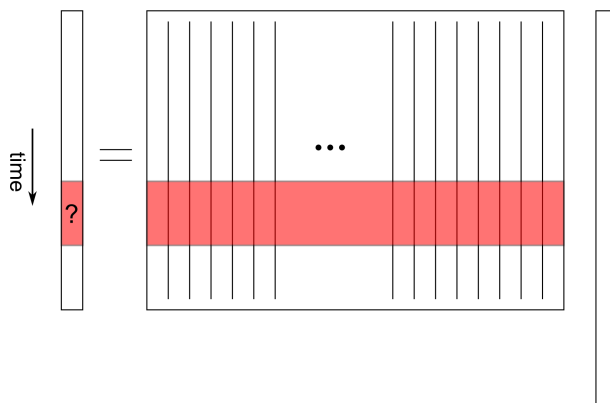
$$\mathbf{y} \approx \mathbf{D}\mathbf{x},$$

přičemž jeho koeficienty \mathbf{x} v tomto slovníku jsou řídké, tj. $\|\mathbf{x}\|_0 \ll N$.

Pro spolehlivé vzorky platí

$$\mathbf{y}^r = \mathbf{M}^r \mathbf{y} \approx \mathbf{M}^r \mathbf{D}\mathbf{x},$$

znázorněno na obr. 5.4.



Obrázek 5.4: Schéma řídké reprezentace signálu, maskované poškozené vzorky v původním signálu a ve slovníku (červený pás, obsahuje nuly místo původních dat).

Chceme rekonstruovat neznámé vzorky \mathbf{y}^m odhadnutím řídké reprezentace $\hat{\mathbf{x}}$ vektoru z nenarušených pozorování \mathbf{y}^r (a slovníku $\mathbf{M}^r \mathbf{D}$) a následném dopočítání chybějící části

$$\hat{\mathbf{y}}^m = \mathbf{M}^m \mathbf{D}\hat{\mathbf{x}}.$$

Celkový rekonstruovaný signál je pak $\hat{\mathbf{y}} = \mathbf{D}\hat{\mathbf{x}}$. Potřebujeme tedy vhodně odhadnout $\hat{\mathbf{x}}$, což lze zapsat například jako

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{vzhledem k} \quad \mathbf{M}^r \mathbf{D}\mathbf{x} = \mathbf{y}^r,$$

v praxi při zašuměném signálu

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{vzhledem k} \quad \|\mathbf{M}^r \mathbf{D}\mathbf{x} - \mathbf{y}^r\|_2 \leq \delta. \quad (5.1)$$

Řešení je možné hledat např. hladovými algoritmy [1],[2] nebo pomocí ℓ_1 -relaxace, jak bylo uvedeno v kapitole 4. Pokud uvažujeme ℓ_1 -relaxaci, úloha (5.1) je ve tvaru

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{vzhledem k} \quad \|\mathbf{M}^r \mathbf{D}\mathbf{x} - \mathbf{y}^r\|_2 \leq \delta. \quad (5.2)$$

Jedná se o tzv. omezenou úlohu (minimalizujeme $\|\mathbf{x}\|_1$ na omezené oblasti), označovanou také BPDN (Basis Pursuit Denoising). Úlohu (5.2) je možné přepsat do neomezeného

5.2. DOSAVADNÍ METODY

tvaru, tzv. LASSO (Least Absolute Shrinkage and Selection Operator) (5.3). Toto je vhodnější, poněvadž většina algoritmů konvexní optimalizace je schopna řešit pouze neomezený tvar.

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left(\underbrace{\frac{1}{2} \|\mathbf{M}^T \mathbf{D} \mathbf{x} - \mathbf{M}^T \mathbf{y}\|_2^2}_{f_1} + \underbrace{\lambda \|\mathbf{x}\|_1}_{f_2} \right). \quad (5.3)$$

Funkce f_1 vyžaduje malou odchylku od dat, f_2 vynucuje řídkost. f_1 je diferencovatelná, máme tedy splněny podmínky pro použití proximálního algoritmu.

Celkově je pak odhad chybějících vzorků

$$\hat{\mathbf{y}}^m = \mathbf{M}^m \mathbf{D} \cdot \arg \min_{\mathbf{x}} \left(\frac{1}{2} \|\mathbf{M}^T \mathbf{D} \mathbf{x} - \mathbf{M}^T \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1 \right). \quad (5.4)$$

Poněvadž jsou dostupné pouze spolehlivé vzorky, v algoritmu je potřeba váhovat koeficienty normou atomů ve zkráceném slovníku [1].

Syntetizující a analyzující model

Tvar optimalizační úlohy pro hledání řídké reprezentace signálu (5.3) nazýváme také jako *syntetizující model*. Toto vyjadřuje, že minimalizujeme koeficienty, které ve vhodném slovníku dobře aproximují signál. Druhou možností je uvažovat tzv. *analyzující model* (5.5) [28], kdy hledáme signál \mathbf{z} , který dobře aproximuje náš původní signál a je po analýze řídký:

$$\arg \min_{\mathbf{z}} \left(\frac{1}{2} \|\mathbf{M}^T \mathbf{y} - \mathbf{M}^T \mathbf{z}\|_2^2 + \lambda \|\mathcal{T} \mathbf{z}\|_1 \right), \quad (5.5)$$

kde \mathcal{T} značí analyzující operátor. V experimentech užíváme syntetizující model.

Slovníky

Hojně užívanými slovníky pro řídkou reprezentaci je slovník DCT (diskrétní kosinové transformace) a Gaborův slovník (DGT). Další možností je například slovník vícerozměrné DCT nebo pro danou úlohu slovník tzv. naučit, aby se tím zajistilo co nejřidší vyjádření signálu. Takovýto modifikovaný slovník lze například získat pomocí algoritmu K-SVD [17].

Pracujeme se slovníkem DGT. V experimentech používáme LTFAT Matlab/Octave toolbox [31] a odpovídající definici této transformace. Uvažujeme vhodnou okénkovou funkci \mathbf{g} a signál $\mathbf{y} \in \mathbb{C}^L$. a označuje časový posun, M počet kmitočtových kanálů a definujeme $N = L/a$. Výstupem DGT je pole $M \times N$ koeficientů

$$x_{m+1,n+1} = \sum_{j=0}^L \mathbf{y}_{j+1} \mathbf{g}_{j-na+1}^* e^{-2\pi i \frac{m}{M} j},$$

kde $m = 0, \dots, M-1, n = 0, \dots, N-1$ a $j-na$ je počítáno modulo L [31]. Pro vhodné \mathbf{g}, a, M tvoří posunuté a modulované verze okna frame. Uvažujeme těsný frame, pak atomy framu tvoří sloupce matice \mathbf{D} .

5.2.4. Výhody a nevýhody dosavadních metod

Metoda	Výhody	Nevýhody
AR modely	rychlé, vhodné pro stacionární signály s dírami do 20 ms	nevhodné pro delší chybějící úseky a nestacionární signály
Sinusoidální modely	lze používat pro stacionární signály s dírami až 1 s	nelze použít pro nestacionární signály
Neuronové sítě	vhodné pro detekci impulzního šumu a jeho odstranění, rozlišení nechtěného šumu (praskance) od chtěného (transienty, nárazové zvuky)	znalosti získané sítě nejsou aplikovatelné univerzálně - proces učení potřeba opakovat při změně charakteru audio materiálu nebo impulzních narušení
Řídká reprezentace (hladové algoritmy)	jednoduchá implementace, rychlé	dosažení globálního optima není zaručeno
Řídká reprezentace (ℓ_1 -relaxace)	při splnění podmínek existence a jednoznačnosti výpočetně efektivní a praktické	ztráta energie koeficientů v rekonstrukci, tím větší, čím více daný atom zasahoval do díry

5.3. Využití strukturované řídkosti

Jak již bylo zmíněno, každý hudební signál má jistou strukturu a tedy předpoklad nezávislosti jednotlivých koeficientů neodpovídá realitě. Proto jsou algoritmy uvažující informace o této struktuře při zpracování signálu výhodnější oproti těm klasickým, které chápou řídkost koeficient po koeficientu. Problém doplnění chybějících dat (5.4) přechází na

$$\hat{\mathbf{y}}^m = \mathbf{M}^m \mathbf{D} \cdot \arg \min_{\mathbf{x}} \left(\frac{1}{2} \|\mathbf{M}^r \mathbf{D} \mathbf{x} - \mathbf{M}^r \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_{p,q} \right).$$

Pro strukturu skupina-člen použijeme indexování (g, m) :

$$\hat{\mathbf{y}}^m = \mathbf{M}^m \mathbf{D} \cdot \arg \min_{\mathbf{x}} \left[\frac{1}{2} \|\mathbf{M}^r \mathbf{D} \mathbf{x} - \mathbf{M}^r \mathbf{y}\|_2^2 + \lambda \left(\sum_{g=1}^G \left(\sum_{m=1}^M |x_{g,m}|^p \right)^{q/p} \right)^{1/q} \right].$$

V aplikacích se pracuje se skupinovou a elitářskou řídkostí, tedy $p, q \in \{1, 2\}$. Proximální algoritmy řešící tyto úlohy pak využívají zobecněnou funkci prahování pro dané smíšené normy [22],[33].

V úloze bez strukturované řídkosti měl proximální operátor tvar

$$\text{prox}_{\lambda \|\cdot\|_1}(x_{g,m}) = \frac{x_{g,m}}{|x_{g,m}|} \cdot \max(|x_{g,m}| - \lambda, 0).$$

Pro $p = 2, q = 1$ se úloha transformuje na Group-LASSO (GL)

5.3. VYUŽITÍ STRUKTUROVANÉ ŘÍDKOSTI

$$\text{prox}_{\lambda \|\cdot\|_{21}}(x_{g,m}) = \frac{x_{g,m}}{\|\mathbf{x}_g\|_2} \cdot \max(\|\mathbf{x}_g\|_2 - \lambda, 0),$$

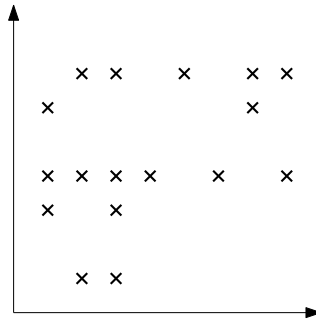
kde $\|\mathbf{x}_g\|_2 = \sqrt{\sum_m |x_{g,m}|^2}$.

Volba $p = 1, q = 2$ dává Elitist-LASSO (EL)

$$\text{prox}_{\lambda \|\cdot\|_{12}}(x_{g,m}) = \frac{x_{g,m}}{|x_{g,m}|} \cdot \max\left(|x_{g,m}| - \frac{\lambda}{1 + M_g \lambda} \|\tilde{\mathbf{x}}_g\|_1, 0\right),$$

kde $\tilde{\mathbf{x}}_g = (\tilde{x}_{g,1}, \dots, \tilde{x}_{g,M_g})$ a $\{\tilde{x}_{g,\tilde{m}}\}_{\tilde{m}}$ označuje pro každou skupinu klesající posloupnost absolutních hodnot $|x_{g,m}|$. Přirozené číslo M_g je závislé na velikosti koeficientů ve skupině $(x_{g,1}, \dots, x_{g,M})$ [22].

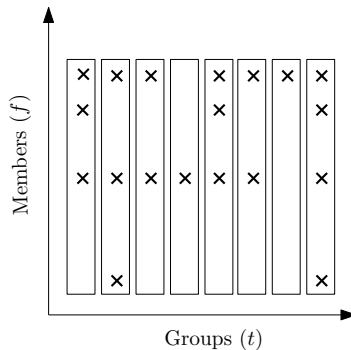
Na obrázku 5.5 znázornění příkladu řídkých koeficientů bez využití skupin.



Obrázek 5.5: Příklad řídkých koeficientů bez použití skupin.

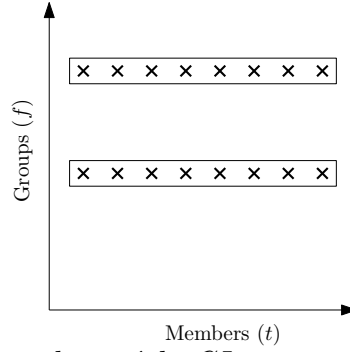
Tvoříme skupiny v časově-kmitočtové oblasti. Pro účely doplňování chybějících dat používáme skupiny pro zachycení harmonické struktury signálu (horizontální). Vertikální struktura nám totiž nemůže pomoci při rekonstrukci, jelikož v díře nemáme žádnou informaci o těchto významných koeficientech. Naopak v horizontální struktuře mohou významné koeficienty přesahovat levý či pravý okraj díry a tím pádem máme údaje pro rekonstrukci.

Proto můžeme používat buď EL se skupinovým indexem g odpovídajícím časovému indexu (na obr. 5.6), nebo lze použít GL se skupinovým indexem g odpovídajícím kmitočtovému indexu (na obr. 5.7).



Obrázek 5.6: Koeficienty ve skupinách, EL vynucuje řídkost mezi členy skupin.

5. DOPLŇOVÁNÍ CHYBĚJÍCÍCH DAT DO ZVUKOVÉHO SIGNÁLU

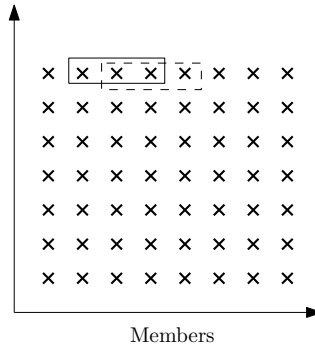


Obrázek 5.7: *Koeficienty ve skupinách, GL vynucuje řídkost mezi skupinami.*

Pro lepší zachycení struktury byl zaveden flexibilnější způsob určení významných koeficientů, a to pomocí posuvného okna [22], [33]. Významnost koeficientu je posuzována na základě energie jeho *časově-kmitočtového okolí*. Každému indexu $k = (g, m)$ je přiřazena skupina sousedů $\mathcal{N}(k)$, a dále pracujeme pouze s tímto okolím. Tedy např. pro GL

$$\text{prox}_{\lambda \|\cdot\|_{21}}(x_{g,m}) = \frac{x_{g,m}}{\|\mathbf{x}_k\|_{2;(\mathcal{N}(k))}} \cdot \max(\|\mathbf{x}_k\|_{2;(\mathcal{N}(k))} - \lambda, 0),$$

kde $\mathbf{x}_k = \{x_{k'}, k' \in \mathcal{N}(k)\}$. Získáváme tím pak tzv. Windowed-Group-LASSO (WGL), pro EL získáme Persistent-Elitist-LASSO (PEL). Příklad posuvného okna je znázorněn na obr. 5.8.



Obrázek 5.8: *Příklad posuvného okna.*

Jednotlivým prvkům okolí je také možné přiřazovat váhy $w_k(k')$ (nejčastěji nejvyšší váhu prvku v centru skupiny a stále menší váhy prvkům vzdálenějším od něj). Okolí definujeme

$$\mathcal{N}(k) = \{k' : w_k(k') \neq 0\},$$

w_k splňují $w_k(k) > 0, w_k(k') \geq 0, \forall k'$ a $\sum_{k' \in \mathcal{N}(k)} w_k(k') = 1$.

V následující části jsem testovala kvalitu rekonstrukce pro různé velikosti okolí.

6. Experiment

V této části jsem provedla rekonstrukci pomocí strukturované řídkosti s postupně rostoucí velikostí horizontálního okolí. Použila jsem zrychlenou verzi proximálního algoritmu (tzv. FISTA) upravenou pro strukturovanou řídkost – structFISTA. Regularizační parametr λ zůstal během testování různých délek okolí nezměněn.

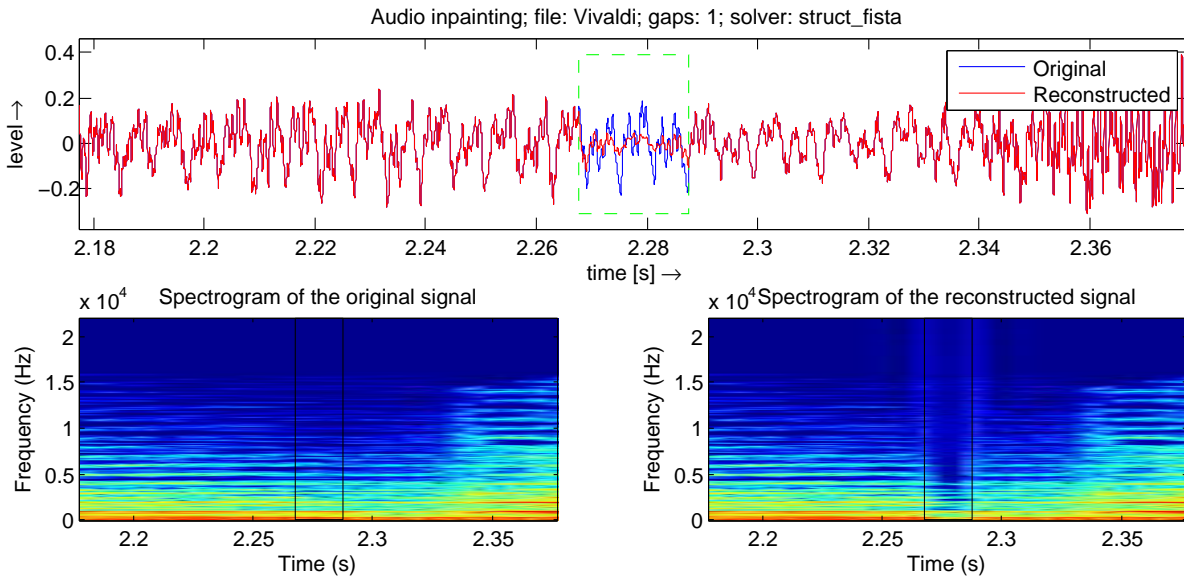
SNR

Často využívaným kritériem pro porovnání kvality rekonstrukce je tzv. *Signal-to-noise ratio* (SNR), neboli odstup signálu od šumu. Výpočet je dán

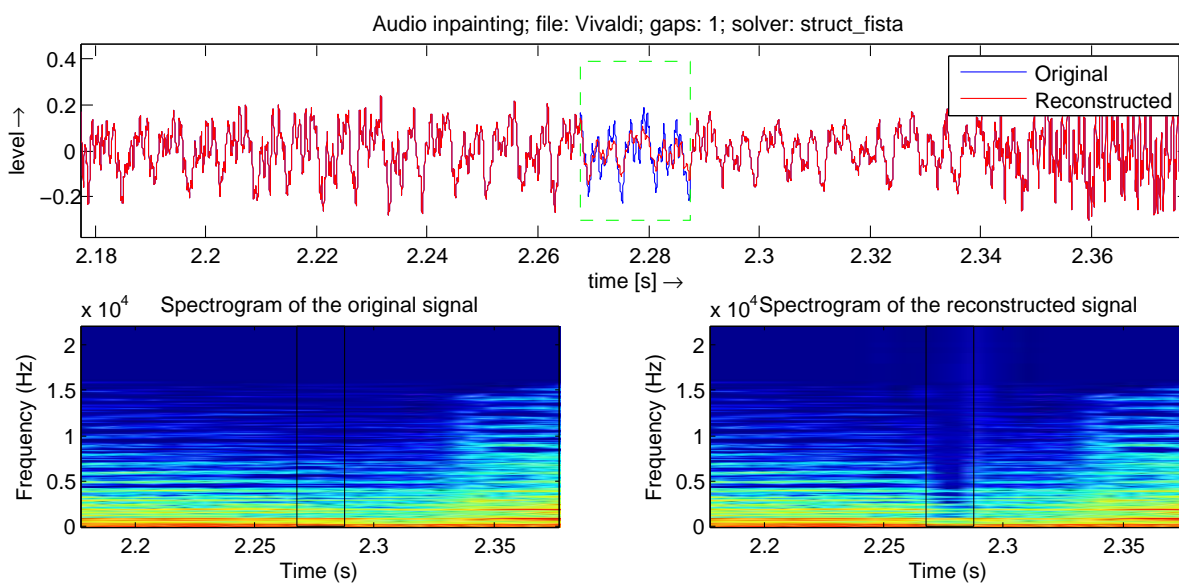
$$\text{SNR}(\mathbf{y}, \hat{\mathbf{y}}) = 10 \cdot \log_{10} \left(\frac{\text{var}(\mathbf{y})}{\text{var}(\mathbf{y} - \hat{\mathbf{y}})} \right),$$

kde var označuje rozptyl. Poněvadž původní signál \mathbf{y} známe (díra je vytvořena uměle), můžeme toto kritérium použít.

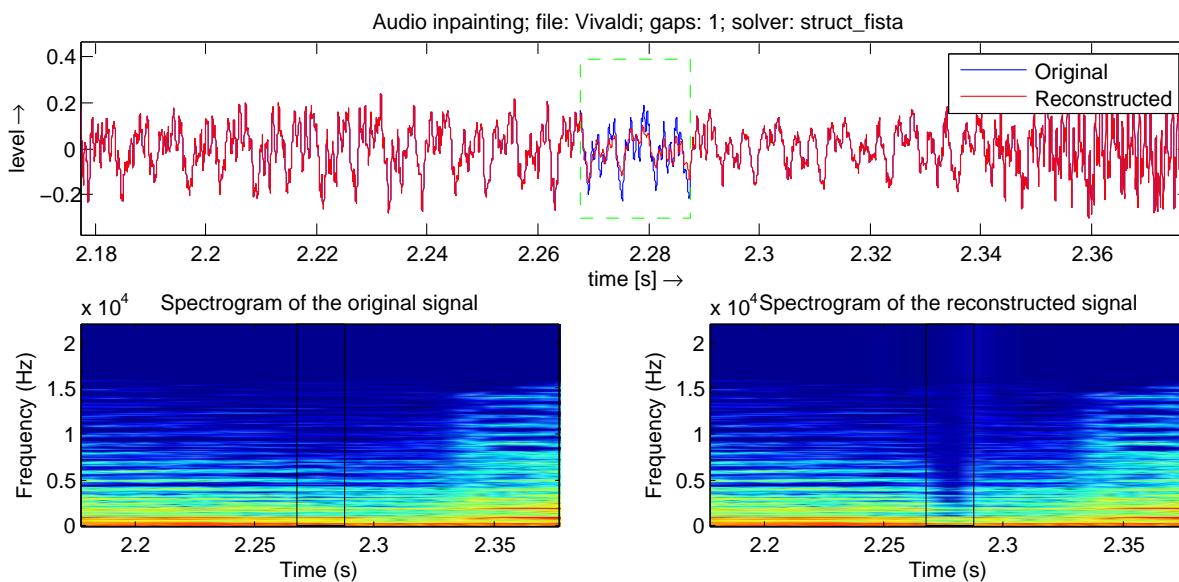
Nejdříve byl v signálu smyčcového orchestru (vzorkovací kmitočet 44,1 kHz) doplňován chybějící úsek o délce 20 ms (882 vzorků). Pro reprezentaci byl použit těsný Gaborův frame generovaný Hannovou okénkovou funkcí o délce 2049 vzorků, časovým posunem 683 vzorků a s 2049 rovnoměrně rozloženými kmitočtovými kanály. Na obrázcích 6.1 - 6.9 je zobrazena rekonstrukce v časové i v časově-kmitočtové oblasti, postupně pro počet koeficientů okolí 1 (bez struktury) až po 17.



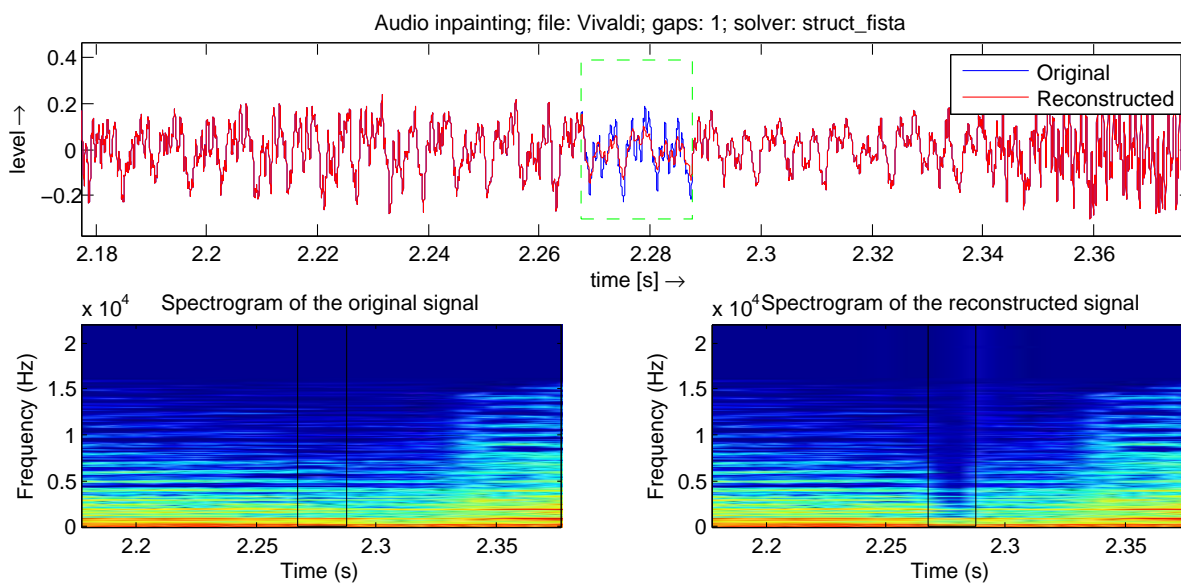
Obrázek 6.1: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 1 (bez struktury).



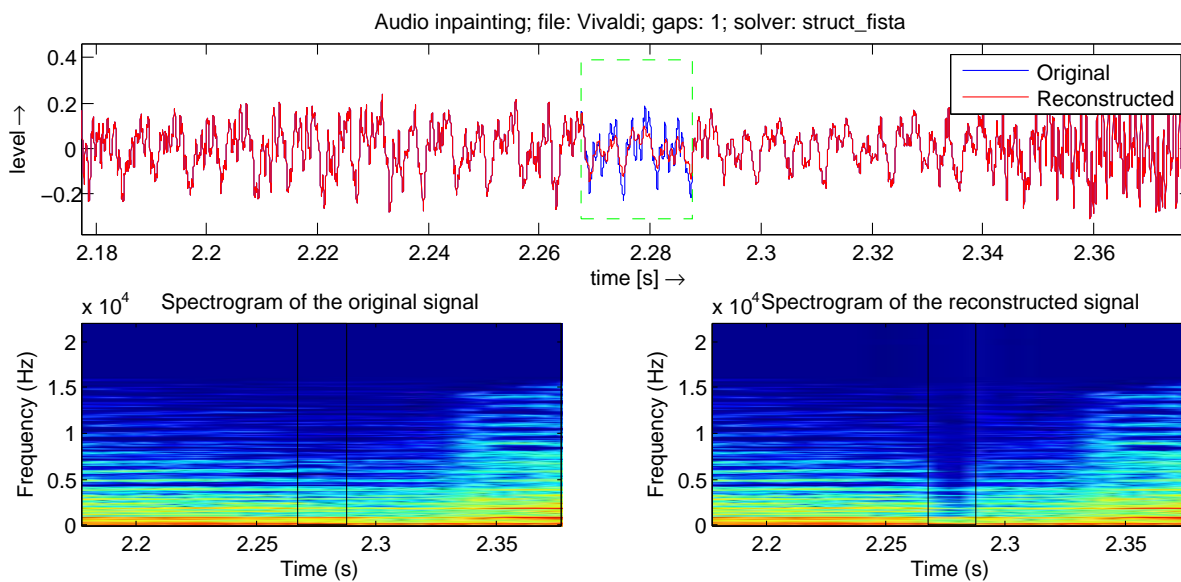
Obrázek 6.2: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 3.



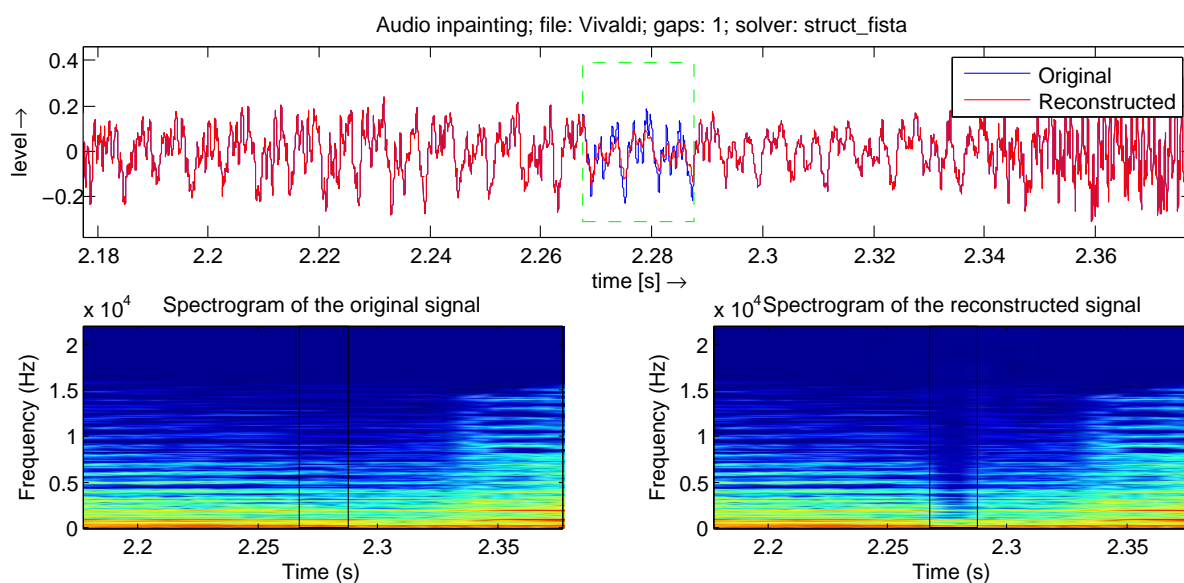
Obrázek 6.3: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 5.



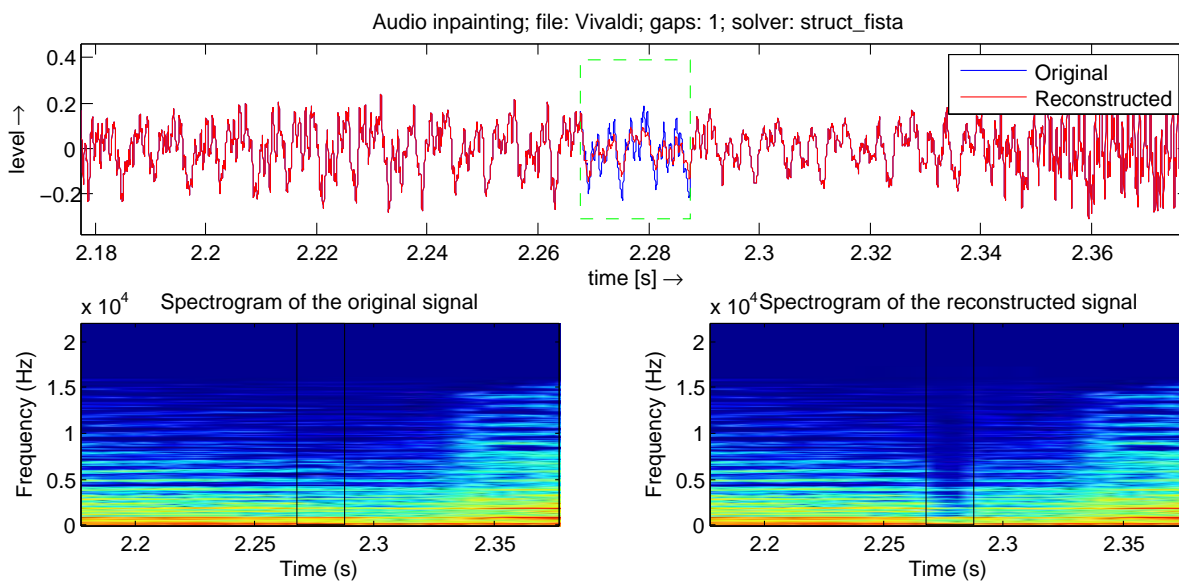
Obrázek 6.4: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 7.



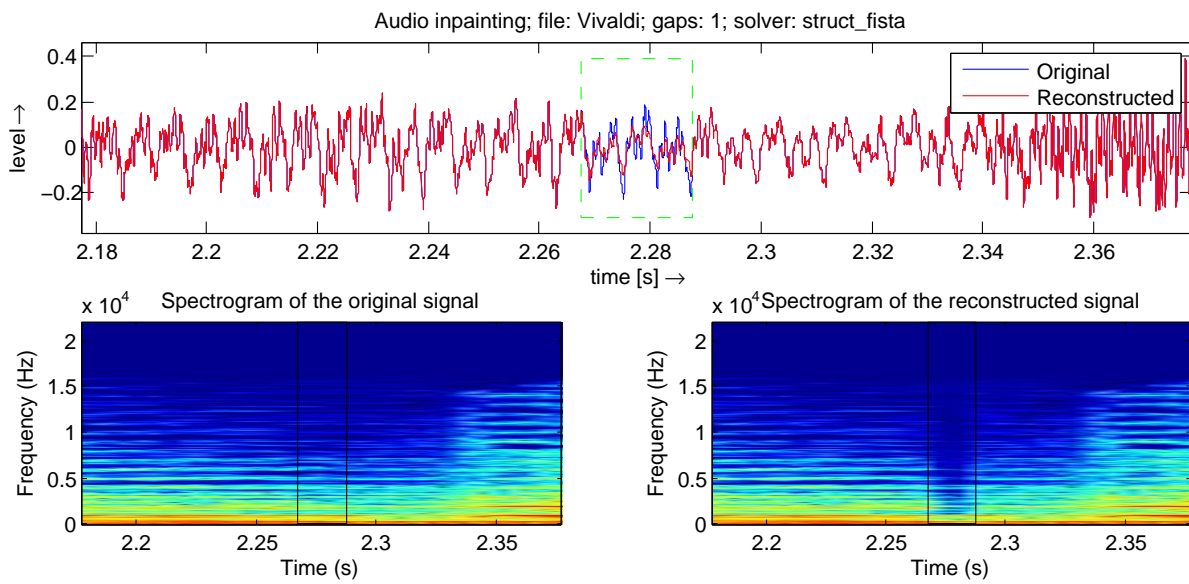
Obrázek 6.5: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 9.



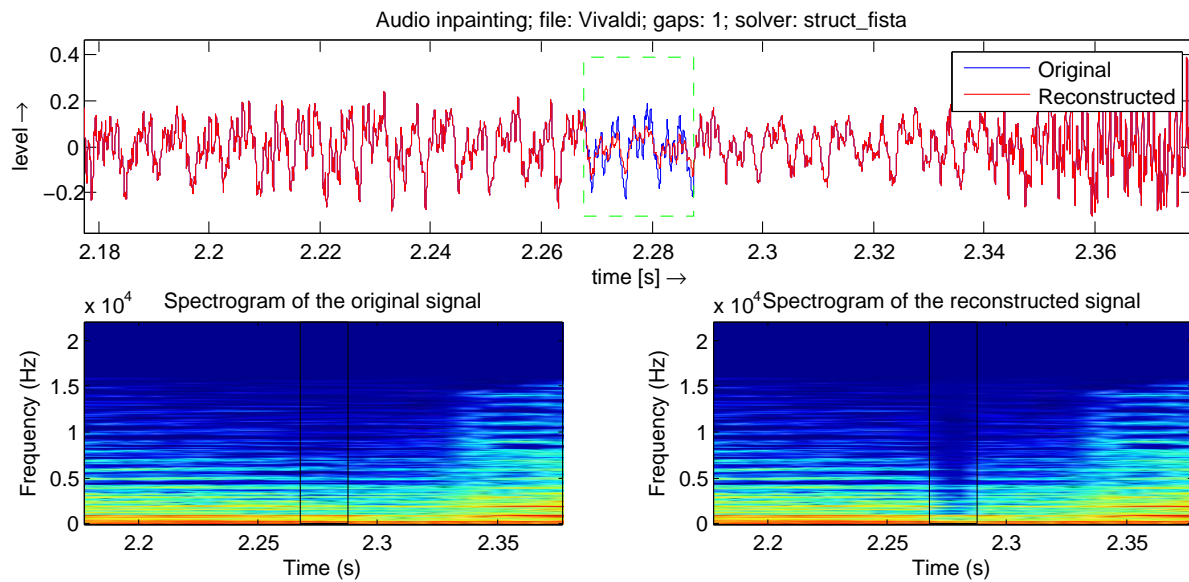
Obrázek 6.6: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 11.



Obrázek 6.7: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 13.

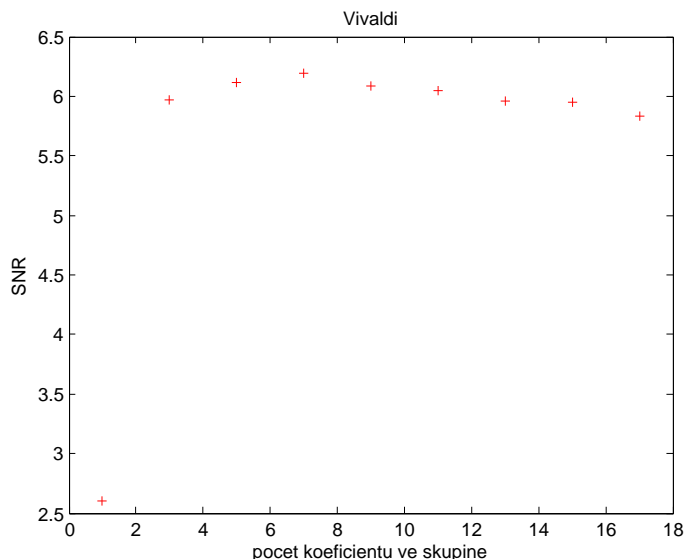


Obrázek 6.8: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 15.



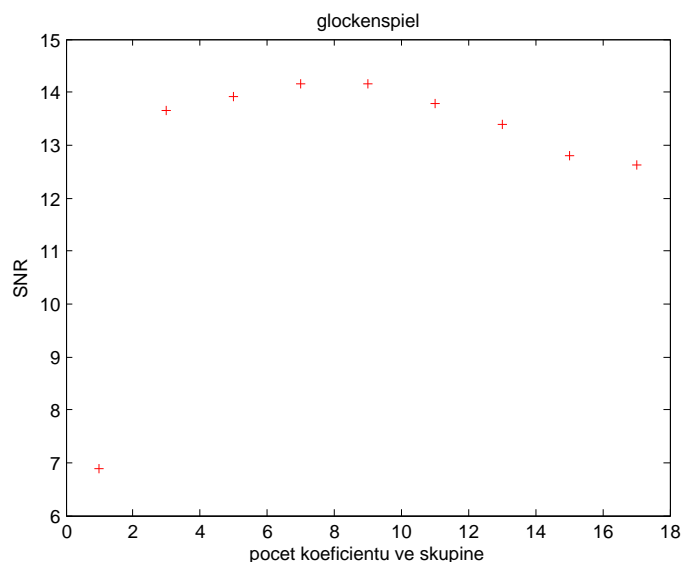
Obrázek 6.9: Rekonstrukce chybějícího úseku pro počet koeficientů skupiny 17.

Kvalita rekonstrukce pro různé skupiny byla vyhodnocena také pomocí SNR (obr. 6.10). Je patrné, že použití strukturované řídkosti přináší výrazné zlepšení. Nejlepší výsledky odpovídají skupinám o délce 5 a 7 koeficientů (délka nosiče skupiny okolo 140 ms), poté však úspěšnost rekonstrukce opět klesá.



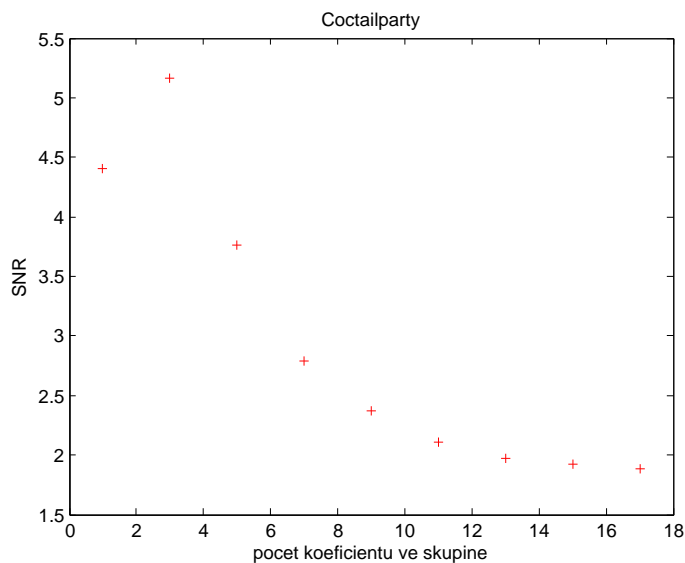
Obrázek 6.10: Vývoj SNR pro různé velikosti skupin, smyčcový orchestr, $f_s = 44,1$ kHz. Nejlepší rekonstrukce pro skupinu 7 koeficientů - což odpovídá nosiči o délce cca 140 ms.

Dále byla stejným způsobem testována rekonstrukce dalších 4 signálů. Délka gaborovského okna a parametry translace a modulace a, M byly použity stejné jako u prvního signálu. Chybějící úsek měl také stále délku 20 ms (počet vzorků se liší podle vzorkovacího kmitočtu f_s daného signálu). Na obr. 6.11 signál zvonkohry, $f_s 44,1$ kHz. Nejlepší výsledek přináší volba skupin o 7 a 9 koeficientech.



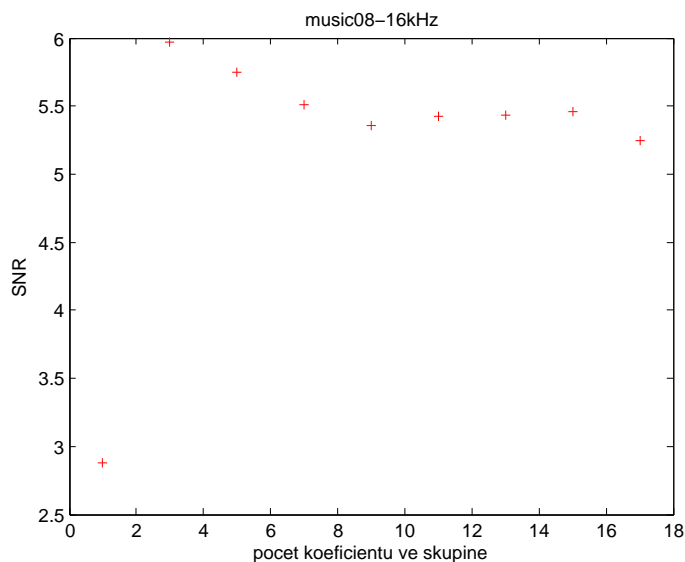
Obrázek 6.11: Vývoj SNR pro různé velikosti skupin, zvonkohra, $f_s = 44,1$ kHz. Nejlepší rekonstrukce pro skupinu 7 koeficientů - což odpovídá nosiči o délce cca 140 ms.

Na obr. 6.12 řeč – věta pronesená Angličanem. f_s je 44,1 kHz, díra je tedy dlouhá 882 vzorků. Zde je nejlepší SNR získáno se skupinou délky 3 koeficienty a poté se výrazně zhoršuje.



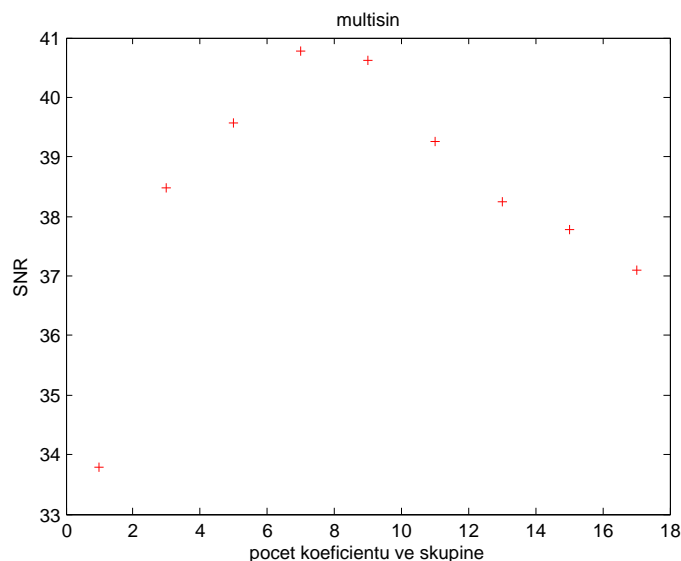
Obrázek 6.12: Vývoj SNR pro různé velikosti skupin, řeč, $f_s = 44,1$ kHz. Nejlepší rekonstrukce pro skupinu 3 koeficientů – což odpovídá nosiči o délce cca 77 ms.

Na obr. 6.13 signál poprockové hudby, f_s 16 kHz (320 chybějících vzorků). Nejlépe je rekonstrukce provedena algoritmem užívajícím skupiny 3 a 5 koeficientů.



Obrázek 6.13: Vývoj SNR pro různé velikosti skupin, poprocková hudba, $f_s = 16$ kHz. Nejlepší rekonstrukce pro skupinu 3 koeficientů - což odpovídá nosiči o délce cca 213 ms.

Na obr. 6.14 signál složený z více sinusovek, f_s 16 kHz, tedy 320 chybějících vzorků. Nejlepší výsledek dávají skupiny 7 a 9 koeficientů.



Obrázek 6.14: Vývoj SNR pro různé velikosti skupin, několik sinusovek, $f_s = 16$ kHz. Nejlepší rekonstrukce pro skupinu 7 koeficientů - což odpovídá nosiči o délce cca 384 ms.

Shrnutí

Provedené experimenty prokazují, že strukturovaná řídkost přináší znatelné zlepšení. Z vyhodnocení pomocí SNR dává nejlepší výsledky použití skupin od velikosti 3 po 7 koeficientů, délka nosiče v ms samozřejmě záleží na vzorkovacím kmitočtu. Je zřejmé, že hudební signály různých žánrů a řeč mají trochu rozdílnou strukturu a tím odpovídající optimální délku nosiče použitého okolí. Pro řeč byla nejlepší rekonstrukce získána se skupinami s nosičem okolo 77 ms, pro hudební signály se pohybovala od 140 po 213 ms. Signál tvořený z několika sinusovek byl nejlépe rekonstruován skupinou s nosičem cca 384 ms, odpovídá 9 koeficientům okolí.

Pro skupiny tvořené více koeficienty kvalita rekonstrukce klesá, poněvadž se harmonické struktury v hudebních signálech pohybují okolo délky 200 ms a nepřesahují 0,5 s.

6.1. Maximalizace efektivity výpočtu

V této části textu se zabývám výpočetní náročností. Délka úseku signálu zpracovávaného metodami založenými na řídkosti může znatelně ovlivnit délku výpočtu. Autoři současných publikací vždy uvažují před použitím algoritmu pouze „dostatečně dlouhý úsek signálu kolem díry“. Zrychlit výpočet však můžeme zkrácením daného úseku na nutné minimum.

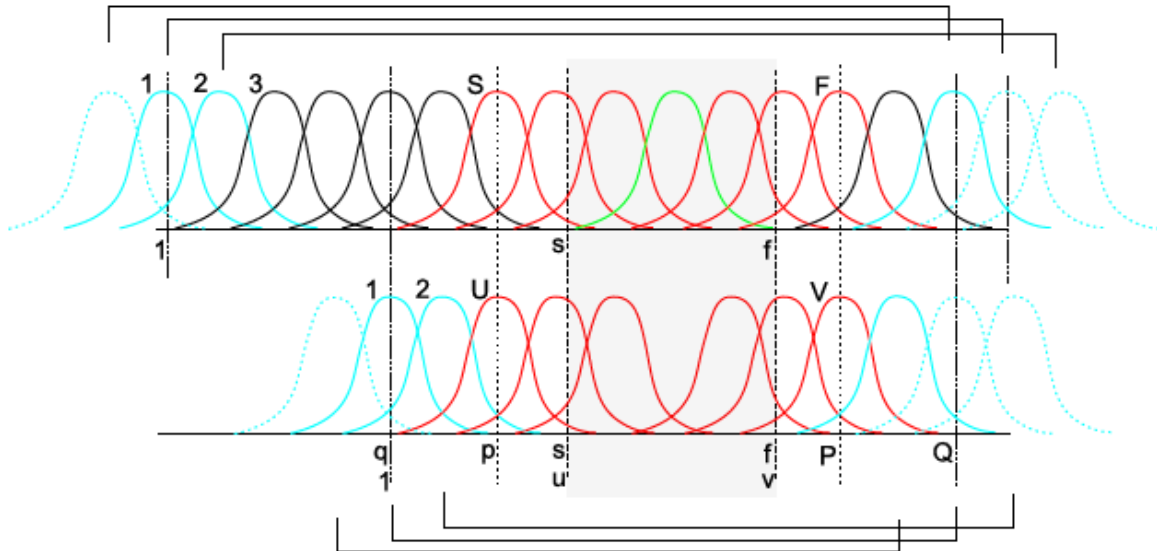
Máme tedy otázku

Jaký je minimální počet vzorků signálu vlevo, resp. vpravo od díry, které je nutné vložit do rekonstrukčního algoritmu, aby poskytoval shodné výsledky jako rekonstrukce s použitím jakkoliv delšího úseku?

Zabýváme se rekonstrukcí dle úlohy

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{M}^r \mathbf{D} \mathbf{x} - \mathbf{M}^r \mathbf{y}\|_2^2 + \lambda \|\mathbf{x}\|_1, \quad (6.1)$$

6.1. MAXIMALIZACE EFEKTIVITY VÝPOČTU



Obrázek 6.15: *Nahoře: Originální, dlouhý úsek signálu a pozice díry s, f v rámci něj. Schematicky jednotlivé okénkové funkce DGT. Dole: Nový úsek signálu s minimálním potřebným nosičem. Pouze červeně označená okna budou použita pro regularizaci a fitování dat – tzv. „užitečná“. Azurová barva označuje okna, která budou použita výhradně kvůli periodicitě DGT, ale nebudou použita v optimalizační úloze. Černě označená okna jsou pro rekonstrukci nadbytečná a těch se „zbavíme“ díky vhodnému zkrácení signálu. Okno označené zeleně celé spadá do díry a tudíž příslušný atom nemůže nést žádnou informaci relevantní pro rekonstrukční úlohu.*

tj. řídkost bez vnitřní struktury. Jako slovník předpokládáme gaborovský systém, avšak následující odvození lze přizpůsobit i jiným reprezentačním systémům.

V úloze (6.1) pro dlouhý signál hledáme řídkou reprezentaci. Ovšem, při doplňování chybějící části (syntéza $\hat{y} = \mathbf{D}\hat{x}$) pouze zlomek nese užitečnou informaci. Pro rekonstrukci jsou totiž podstatné *pouze* atomy (příp. skupiny atomů), které mají s dírou přesah.

Diskrétní Gaborova transformace (DGT) je periodická: existují atomy, které zasahují jak do levého, tak do pravého konce signálu (obr. 6.15). Koeficienty příslušné těmto atomům je nutné při DGT počítat, ale při doplňování dat do díry je nemůžeme využít. Obsahují informaci z obou konců signálu a tyto úseky spolu často nijak nesouvisejí.

6.1.1. Značení

Prvky vektorů jsou číslovány počínaje jedničkou. Výchozí signál bude mít délku N , tzn. prvky na pozicích $1, \dots, N$.

Délka gaborovského okna odpovídajícího použitému slovníku bude označena w , přičemž může být sudá či lichá. Tedy především uvažujeme pouze okna konečné délky (FIR).

Jako „prostřední“ prvek okna chápeme index $\lfloor \frac{w}{2} \rfloor + 1$. Nalevo od prostředního prvku je $\lfloor \frac{w}{2} \rfloor$ hodnot, napravo od něj je $\lceil \frac{w}{2} \rceil - 1$ hodnot. Pokud je tedy w liché, „prostřední“ prvek leží skutečně vprostřed okna, pro w sudé je vychýlen o jeden vzorek doprava. První okno transformace má centrální prvek umístěn na indexu 1, a tedy toto okno končí nenulovou hodnotou na indexu $\lceil \frac{w}{2} \rceil$. Okno s pořadovým číslem $n + 1$ má prostřední prvek na pozici $1 + na$, první prvek na pozici $1 + na - \lfloor \frac{w}{2} \rfloor$ a poslední prvek na pozici $na + \lceil \frac{w}{2} \rceil$ (pokud

by jakákoliv z těchto hodnot padla mimo rozsah $\{1, \dots, N\}$, počítáme je modulo N díky periodicitě transformace).

Celé číslo a značí posun okna po časové ose, M je počet kanálů (rovnoměrně rozložených spektrálních koeficientů) transformace. Symboly s a f budou zastupovat pozice začátku a konce díry v rámci *originálního* signálu.

Indexy q, Q označují první, resp. poslední vzorek hledaného zkráceného úseku vzhledem k původním signálu. Indexy p, P označují podobně prostřední prvky prvního, resp. posledního okna, která mají s dírou překryv (a tudíž přináší smysluplnou informaci); čísla S, F vyjadřují pořadí těchto dvou oken v transformaci originálního signálu. Symboly u, v, U, V mají význam analogický k s, f, S, F pro zkrácený signál.

Obr. 6.15 ukazuje popsanou situaci.

6.1.2. Odvození algoritmu

Je zřejmé, že uvedené veličiny jsou svázány přes parametr a (translace okna) tak, že pro vhodné $k, l, m \in \mathbb{N}$ platí

$$p = q + ka, \quad P = q + la, \quad Q = q + ma. \quad (6.2)$$

Díky tomu můžeme postupně dojít až k finálnímu algoritmu.

Jak nalezneme prostřední index p prvního okna zleva zasahujícího do díry?

Protože na indexu 1 v signálu leží právě prostřední prvek prvního okna transformace, posouváme se v násobcích a od tohoto indexu až k takovému oknu, které má s dírou překryv alespoň jeden vzorek. To tedy znamená, že hledáme takové nejmenší k , že platí

$$\begin{aligned} p + \left\lceil \frac{w}{2} \right\rceil - 1 &\geq s \\ 1 + ka + \left\lceil \frac{w}{2} \right\rceil - 1 &\geq s \\ k &\geq \left(s - \left\lceil \frac{w}{2} \right\rceil \right) / a. \end{aligned}$$

Takové k zároveň souvisí s pořadím tohoto okna v původní transformaci, a platí $S = \left\lceil \left(s - \left\lceil \frac{w}{2} \right\rceil \right) / a \right\rceil + 1$. Proto hledané

$$p \leftarrow 1 + (S - 1)a.$$

Jak nalezneme první index q hledaného minimálního úseku?

Aby první užitečné okno nebylo v DGT periodizováno (jinak příslušné koeficienty ponесou mylnou informaci), musí být zajištěno, aby index q neležel vpravo od nejlevějšího prvku prvního užitečného okna. Hledáme nejmenší k takové, že

$$\begin{aligned} q &\leq p - \left\lfloor \frac{w}{2} \right\rfloor \\ p - ka &\leq p - \left\lfloor \frac{w}{2} \right\rfloor \\ k &\geq \left\lfloor \frac{w}{2} \right\rfloor / a. \end{aligned}$$

6.1. MAXIMALIZACE EFEKTIVITY VÝPOČTU

Proto

$$q \leftarrow p - \left\lfloor \left\lfloor \frac{w}{2} \right\rfloor / a \right\rfloor \cdot a.$$

Jak nalezneme prostřední index P prvního okna zprava zasahujícího do díry?

Od indexu p se posouváme v násobcích a až k oknu, které má s dírou překryv zprava nejméně jeden vzorek. Tedy hledáme takové největší k , pro které platí

$$\begin{aligned} P - \left\lfloor \frac{w}{2} \right\rfloor &\leq f \\ p + ka - \left\lfloor \frac{w}{2} \right\rfloor &\leq f \\ k &\leq \left(f + \left\lfloor \frac{w}{2} \right\rfloor - p \right) / a. \end{aligned}$$

Takové k opět souvisí s pořadím F tohoto okna v původní transformaci, a platí $F = S + \left\lfloor (f + \left\lfloor \frac{w}{2} \right\rfloor - p) / a \right\rfloor$. Hledané P pak

$$P \leftarrow p + (F - S)a.$$

Jak nalezneme poslední index Q hledaného minimálního úseku?

Poslední index hledaného úseku získáme obdobným způsobem jako index q , tj. poslední užitečné okno nesmí být v DGT periodizováno. Q je poslední index kratšího signálu a na tomto indexu musí být „umístěno“ poslední okno. Přesněji, poslední okno je totéž jako první okno díky periodicitě, a proto na indexu Q musí ležet nikoliv prostřední prvek, ale levý soused prostředního prvku. Tzn. máme $Q = P + ka - 1$. Toto poslední okno je vlastně virtuální a proto je ve schématu tečkovaně. Nejpravější prvek posledního užitečného okna je na indexu $P + \left\lceil \frac{w}{2} \right\rceil - 1$, což vede na

$$\begin{aligned} Q &\geq P + \left\lceil \frac{w}{2} \right\rceil - 1 \\ P + ka - 1 &\geq P + \left\lceil \frac{w}{2} \right\rceil - 1 \\ k &\geq \left\lceil \frac{w}{2} \right\rceil / a. \end{aligned}$$

Proto

$$Q \leftarrow P + \left\lceil \left\lceil \frac{w}{2} \right\rceil / a \right\rceil \cdot a.$$

Volitelně můžeme zkontrolovat dělitelnost nové délky signálu $Q - q + 1$ konstantami a , M a případné posunutí indexu Q . V případě, že by nový posunutý index Q přesahoval délku původního signálu N , lze signál jednoduše doplnit $Q - N$ nulami.

Pokud je v našem zájmu změnit počet přesahujících vzorků prvního okna, můžeme zvolit korekční parametr c , který je implicitně nastaven na 0. (Čísla S, F budou dávat smysl jen když $c = 0$.) Můžeme to udělat například tak, že nalezneme minimální signál s výchozím $c = 0$, z výstupu algoritmu získáme přesahy oken s dírou, a pokud je budeme chtít korigovat, nastavíme tomu odpovídající c . Příslušně stačí posunout shodně indexy q a Q , není potřeba algoritmus spustit znovu s novým c .

Pozice díry v rámci zkráceného signálu určíme $u = s - q + 1$ a $v = f - q + 1$. Podobně pořadí prvního a posledního okna zasahujícího do díry $U = (p - q) / a + 1$ a $V = U + (F - S)$.

Algoritmus (Nalezení optimální délky signálu kolem díry).

Vstup: w, a, M, s, f (příp. korekce c)

Výstup: $q, Q, p, P, S, F, u, v, U, V$

1. $S \leftarrow \lceil (s - \lceil \frac{w}{2} \rceil) / a \rceil + 1$
2. $p \leftarrow 1 + (S - 1) \cdot a$
3. (Volitelně) $p \leftarrow p + c, c \in [0, w - 1]$
4. $q \leftarrow p - \lceil \lceil \frac{w}{2} \rceil / a \rceil \cdot a$
5. $F \leftarrow S + \lfloor (f + \lfloor \frac{w}{2} \rfloor - p) / a \rfloor$
6. $P \leftarrow p + (F - S) \cdot a$
7. $Q \leftarrow P + \lceil \lceil \frac{w}{2} \rceil / a \rceil \cdot a$
8. Pokud $a \cdot M \nmid (Q - q + 1)$
 $Q \leftarrow Q + a \cdot M - [(Q - q + 1) \bmod a \cdot M]$
9. $u \leftarrow s - q + 1, \quad v \leftarrow f - q + 1$
10. $U \leftarrow (p - q) / a + 1, \quad V \leftarrow U + (F - S)$
11. (Volitelně) výpočet vektoru *overlap*.

Výstupem algoritmu může být i vektor *overlap* nesoucí informaci o velikosti překryvů jednotlivých *užitečných* oken s dírou. Obsahuje počet vzorků daného okna, které spadají do díry. Vektor *overlap* tedy obsahuje $F - S + 1$ hodnot. Získán je jednoduše: Je provedena DGT indikátorového signálu díry (již zkráceného dle předchozího postupu, tedy na indexech q až Q původního signálu), a to s obdélníkovým oknem (obsahujícím jako prvky samé jedničky). Zajímá nás pouze DC složka nesoucí hledanou informaci. Výstup DGT obsahuje i nulové překryvy neužitečných, periodizujících se oken, tedy do vektoru *overlap* bereme pouze nenulové hodnoty.

Experiment

Pro ověření funkčnosti a užitečnosti algoritmu byl proveden experiment: Nejdříve byl rekonstruován signál s délkou 64000 vzorků (což odpovídá délce 4 s) a s chybějícím úsekem délky 320 vzorků (tj. 20 ms). Pro rekonstrukci jsme použili těsný Gaborův systém s Hanovskou okénkovou funkcí dlouhou $w = 2049$, parametrem translace $a = 683$ a modulace $M = 2049$. Na rekonstrukci byl použit algoritmus dopředného zpětného dělení. Výpočet trval cca 16.69 s. Následně byl rekonstruován stejný úsek s využitím optimálně zkráceného signálu. Tato část měla délku 4782 vzorků. Pro stejný výsledek experimentu užitím dopředného-zpětného dělení (stejně SNR) byl upraven regularizační parametr algoritmu λ . Výpočet trval cca 2.04 s.

6.1.3. Modifikace pro strukturovanou řídkost

Algoritmus pro výběr minimálního úseku potřebného pro rekonstrukci lze jednoduchým způsobem modifikovat i pro strukturovanou řídkost. Potřebujeme pouze počet prvků okolí $neig$, se kterým pracujeme. Indexy q, Q pak posuneme doleva, resp. doprava o $(neig - 1)$ -násobek translačního parametru a . Dodatečným vstupem algoritmu je tedy pouze velikost užívaného okolí. V modifikovaném algoritmu má číslo S význam posledního okna první skupiny, která má přesah do díry zleva, podobně F označuje první okno poslední skupiny s přesahem zprava v originálním signálu. Indexy p, P opět vyjadřují prostřední indexy těchto oken. Čísla U, V mají význam analogický k S, F ve zkráceném signálu. Čísla S', F', U', V' označují pořadí *skupiny*, která má první překryv zleva, resp. poslední překryv zprava.

Modifikovaný algoritmus (Nalezení optimální délky signálu kolem díry při použití strukturované řídkosti).

Vstup: $w, a, M, s, f, neig$ (příp. korekce c)

Výstup: $q, Q, p, P, S, F, S', F', u, v, U, V, U', V'$

1. $S \leftarrow \lceil (s - \lfloor \frac{w}{2} \rfloor) / a \rceil + 1$
2. $p \leftarrow 1 + (S - 1) \cdot a$
3. $S' \leftarrow \lfloor p - 1 - (neig - 1) \cdot a \rfloor / a + 1$
4. (Volitelně) $p \leftarrow p + c, c \in [0, w - 1]$
5. $q \leftarrow p - \lceil \lfloor \frac{w}{2} \rfloor / a \rceil + (neig - 1) \cdot a$
6. $F \leftarrow S + \lfloor (f + \lfloor \frac{w}{2} \rfloor - p) / a \rfloor$
7. $F' \leftarrow F$
8. $P \leftarrow p + (F - S) \cdot a$
9. $Q \leftarrow P + \lceil \lceil \lfloor \frac{w}{2} \rfloor / a \rceil + (neig - 1) \rceil \cdot a$
10. Pokud $a \cdot M \nmid (Q - q + 1)$
 $Q \leftarrow Q + a \cdot M - [(Q - q + 1) \bmod a \cdot M]$
11. $u \leftarrow s - q + 1, v \leftarrow f - q + 1$
12. $U \leftarrow (p - q) / a + 1, V \leftarrow U + (F - S)$
13. $U' \leftarrow \lfloor p - q - (neig - 1) \cdot a \rfloor / a + 1, V' \leftarrow U' + (F' - S')$

7. Závěr

Řídká reprezentace má při zpracování signálů široké spektrum využití. V této práci je nejprve shrnut teoretický základ ohledně reprezentačních systémů, tj. bází a framů, speciálně pak byly zavedeny Gaborovy framy. Poté byla formulována optimalizační úloha hledající řídkou reprezentaci signálu a principy algoritmů, které ji řeší. Poněvadž „norma“ určující řídkost daného vektoru není konvexní funkcí (a tím pádem není možné použít metody konvexní optimalizace), nabízí se její nahrazení nejbližší normou, která konvexní je. Toto vede na tzv. ℓ_1 -relaxaci. Byly uvedeny podmínky existence a ekvivalence řešení původní a relaxované úlohy. Zvláštní pozornost byla věnována proximálním algoritmům, které byly užívány v experimentech užíváme. Byly zmíněny základní principy důležité aplikace řídkých reprezentací nazývané komprimované snímání.

Hlavní aplikací prezentovanou v této práci bylo doplňování chybějících úseků do audio-signálů. Formulovala jsem problém, uvedla jsem dosavadní metody a shrnula jejich výhody a nevýhody. Dále jsem testovala nejnovější přístup využití řídkých reprezentací pro doplnění dat uvažující harmonickou strukturu zpracovávaných signálů. Testování probíhalo v programovém prostředí Matlab. Optimální velikosti skupin koeficientů pro hudební signály různých žánrů či řeč se trochu liší, ovšem z provedených experimentů je jednoznačně patrné, že využití strukturované řídkosti přináší výrazné zlepšení oproti řídkosti chápané koeficient po koeficientu.

V poslední části jsem odvodila algoritmus zaručující maximalizaci efektivity výpočtu a testováním prokázala znatelné zrychlení. Také jsem odvozený algoritmus modifikovala pro strukturovanou řídkost.

Dále se nabízí testování strukturované řídkosti na větším počtu signálů a následné statistické vyhodnocení.

Literatura

- [1] ADLER, A., V. EMIYA a M. G. JAFARI. Audio Inpainting. *IEEE Transactions on Audio, Speech, and Language Processing*. 2012, vol. 20, issue 3, s. 922-932. ISSN 1558-7916. DOI: 10.1109/TASL.2011.2168211.
- [2] ADLER, A., V. EMIYA, M. G. JAFARI, aj. A constrained matching pursuit approach to audio declipping. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2011. DOI: 10.1109/icassp.2011.5946407.
- [3] BALAZS, P., M. DÖRFLER, N. HOLIGHAUS, aj. Theory and Implementation of Nonstationary Gabor Frames. *J. Comput. Appl. Math.* 2011, vol.256, issue 6.
- [4] BAZARAA, M. a J. J. JARVIS. *Linear programming and network flows*. New York: Wiley, 1977, 565 s. ISBN 04-710-6015-1.
- [5] BOUFOUNOS, P. T., P. SMARAGDIS a B. RAJ. Joint sparsity models for wide-band array processing, *Proceeding of SPIE 8138, Wavelets and Sparsity XIV*, 2011. DOI:10.1117/12.893870.
- [6] BOYD, S. a L. VANDENBERGHE. *Subgradients*. Convex Optimization II. Stanford University. Dostupné z: http://see.stanford.edu/materials/lsocoe364b/01-subgradients_notes.pdf.
- [7] CANDÈS, E.J. a M.B. WAKIN. An Introduction To Compressive Sampling. *IEEE Signal Processing Magazine*. 2008, vol. 25, issue 2, s. 21-30. DOI: 10.1109/MSP.2007.914731. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4472240>.
- [8] CHEMAN, K. M. *Optimization techniques for solving basis pursuit problems*. [Master thesis.] Raleigh: North Carolina State University, 2006. 39 s.
- [9] CHEN, S. S., D. L. DONOHO a M. A. SAUNDERS. Atomic Decomposition by Basis Pursuit. *SIAM Review*. 2001, vol. 43, issue 1, s. 129-159. DOI: 10.1137/s003614450037906x.
- [10] CHRISTENSEN, O. *Frames and bases: an introductory course*. Boston: Birkhauser, 2008, 313 s. ISBN 978-0-8176-4678-3.
- [11] COMBETTES, P. L. a J.-C. PESQUET. Proximal Splitting Methods in Signal Processing. In BAUSCHKE, H. H.: *Fixed-point algorithms for inverse problems in science and engineering*. New York: Springer, 2011, s.185-212. ISBN 978-1-4419-9568-1.
- [12] CZYZEWSKI, A. Learning Algorithms for Audio Signal Enhancement, Part 1: Neural Network Implementation for the Removal of Impulse Distortions. *Journal of the Audio Engineering Society*. New York, NY: AES, 1997, vol.45, issue 10, s. 815-831. Dostupné z: <http://maircsoft.com/maircsoft/papers/learning.pdf>
- [13] DONOHO, D. L. a P. B. STARK. Uncertainty principles and signal recovery. *SIAM J. Appl. Math.* 1989, vol. 49, issue 3, s. 906-931. DOI: 10.1137/0149053.

- [14] ELAD, M. *Sparse and redundant representations: from theory to applications in signal and image processing*. New York: Springer, 2010, 376 s. ISBN 978-1-4419-7011-4.
- [15] ESQUEF, P. A. A. Interpolation of Long Gaps in Audio Signals Using LineSpectrum Pair Polynomials. Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Report 72. *IEEE Transactions on Speech and Audio Processing*, submitted for publication. 2004.
- [16] GODSILL, S. J. a P. J. W. RAYNER. *Digital Audio Restoration: A Statistical Model-Based Approach*. London: Springer, 1998. ISBN 3 540 76222 1.
- [17] GUICHAOUA, C. Dictionary Learning for Audio Inpainting. Robotics. 2012. Dostupné z: <http://dumas.ccsd.cnrs.fr/dumas-00725263>
- [18] HRBÁČEK, R., P. RAJMIC, V. VESELÝ a J. ŠPIŘÍK. Řídké reprezentace signálů: komprimované snímání. *Elektrorevue*. 2011, roč. 2011, č. 67. ISSN 1213-1539. Dostupné z: <http://www.elektrorevue.cz/cz/clanky/zpracovani-signalu/10/ridke-reprezentace-signalu-komprimovane-snimani/>
- [19] HRBÁČEK, R., P. RAJMIC, V. VESELÝ a J. ŠPIŘÍK. Řídké reprezentace signálů: úvod do problematiky. *Elektrorevue*. 2011, roč. 2011, č. 50. ISSN 1213-1539. Dostupné z: <http://www.elektrorevue.cz/cz/clanky/zpracovani-signalu/0/ridke-reprezentace-signalu-uvod-do-problematiky/>
- [20] KAUPPINEN, I. a K. ROTH. Audio signal extrapolation - Theory and applications. *Proc. of the 5th Int. Conference on Digital Audio Effects*. 2002.
- [21] KITIC, S. Audio inpainting by sparse regularization methods. *Journées Bordelaises d'Analyse Mathématique des Images*. 2014.
- [22] KOWALSKI, M. a B. TORRÉSANI. Sparsity and persistence: mixed norms provide simple signal models with dependent coefficients. *Signal, Image and Video Processing*. 2009, vol. 3, issue 3, s. 251-264. DOI: 10.1007/s11760-008-0076-1. Dostupné z: <http://link.springer.com/10.1007/s11760-008-0076-1>
- [23] KOWALSKI, M. a B. TORRÉSANI. Structured Sparsity: from Mixed Norms to Structured Shrinkage. *SPARS'09 - Signal Processing with Adaptive Sparse Structured Representations*. 2009.
- [24] KOWALSKI, M., K. SEIDENBURG a M. DÖRFLER. Social Sparsity! Neighborhood Systems Enrich Structured Shrinkage Operators. *IEEE transactions on signal processing*. 2013, vol.61, issue 10. DOI: 10.1109/TSP.2013.2250967.
- [25] LAGRANGE, M., S. MARCHAND a J. B. RAULT. Long Interpolation of Audio Signals Using Linear Prediction in Sinusoidal Modeling. *Journal of the Audio Engineering Society*. New York, NY: AES, 1953-, roč. 2005, vol.53, issue 10, s. 891-905. Dostupné z: <http://webhome.cs.uvic.ca/lagrange/papers/lagrangeJaes05.pdf>
- [26] LUKIN, A. a J. TODD. Parametric Interpolation of Gaps in Audio Signals. *125th Audio Engineering Society Convention*. 2008. Dostupné z: <http://audio.rightmark.org/lukin/pub/LukinTodd10.pdf>

LITERATURA

- [27] MALLAT, S.G., Z. ZHANG. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*. 1993, vol.41, issue 12, s. 271-293. DOI: 10.1016/b978-0-08-052084-1.50018-1.
- [28] Nam, S., M.Davies, M. Elad, aj. The cosparse analysis model and algorithms. *Applied and Computational Harmonic Analysis*. 2013, vol 34, issue 1, s. 30–56. ISSN 1063-5203. DOI:<http://dx.doi.org/10.1016/j.acha.2012.03.006>.
- [29] OUDRE, L. Interpolation of missing samples in audio signals based on autoregressive modeling. *Image Processing On Line*. Draft. 2014. Dostupné z: http://dev.ipol.im/oudre/interpolation_3.pdf.
- [30] PRIETO-GUERRERO, A., C. MAILHES a F. CASTANIÉ. Lost Sample Recovering of ECG Signals in e-Health Applications. *Conference of the IEEE EMBS*. 2007. Dostupné z: <http://perso.ens-lyon.fr/patrick.flandrin/00100263.pdf>.
- [31] PRŮŠA, Z., P. SONDERGAARD, P. BALASZ, aj. LTFAT: A Matlab/Octave toolbox for sound processing. in *Proceedings of the 10th International Symposium on Computer Music Multidisciplinary Research (CMMR 2013)*, Laboratoire de Mécanique et d'Acoustique, Marseille, France: Publications of L.M.A., October 2013. ISBN 978-2-2909669-23-6, ISSN 1159-09747.
- [32] Semicontinuous function. *Encyclopedia of Mathematics*. Dostupné z: http://www.encyclopediaofmath.org/index.php/Semicontinuous_function
- [33] SIEDENBURG, K. a M. DÖRFLER. Structured sparsity for audio signals. In: *Proceeding of 14th conference on digital audio effects*. 2011.
- [34] ŠPIŘÍK, J., P. RAJMÍČ a V. VESELÝ. Reprezentace signálů: od bází k framům. *Elektrorevue*. 2010, roč.2010, č. 111. ISSN 1213-1539. Dostupné z: <http://www.elektrorevue.cz/cz/clanky/zpracovani-signalu/20/reprezentace-signalu-od-bazi-k-framum/>
- [35] TIBSHIRANI, Robert. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1996, vol.58, issue 1, s. 267-288.
- [36] UNCINI, A. Audio signal processing by neural networks. *Neurocomputing: Evolving Solution with Neural Networks*. 2003, vol.55, issues 3-4, s. 593-625.
- [37] Vectorization (mathematics). In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2015-04-13]. Dostupné z: [http://en.wikipedia.org/wiki/Vectorization_\(mathematics\)](http://en.wikipedia.org/wiki/Vectorization_(mathematics)).
- [38] YUAN, L., J. LIU a J. YE. Efficient Methods for Overlapping Group Lasso. *Advances in Neural Information Processing Systems*. 2011, č. 24. Dostupné z: <http://papers.nips.cc/paper/4275-efficient-methods-for-overlapping-group-lasso.pdf>

Seznam použitých zkratek a symbolů

k	řidkost vektoru
\mathbf{I}	jednotková matice
$\ker(\mathbf{A})$	jádro lineárního zobrazení určeného maticí \mathbf{A}
$\text{spark}(\mathbf{A})$	spark matice \mathbf{A}
$\text{supp}(\mathbf{x})$	nosič vektoru \mathbf{x}
$\mu(\mathbf{A})$	vzájemná koherence
$\sigma(\mathbf{x})_p$	chyba nejlepší aproximace vektoru \mathbf{x} k -řidkým vektorem
γ	konstanta nulového prostoru
δ_k	konstanta zeslabené izometrie
f_s	vzorkovací kmitočet
E_b	operátor modulace
T_a	operátor translace
ι_C	indikátorová funkce množiny C
∂f	subdiferenciál funkce f
prox_f	proximální operátor funkce f
$\ \cdot\ _p$	p -norma
$\ \cdot\ _{p,q}$	(p, q) -smíšená norma
AR	Autoregresní
BP	Basis Pursuit
BPDN	Basis Pursuit Denoising
DCT	Diskrétní Kosinová Transformace
DGT	Diskrétní Gaborova Transformace
EL	Elitist Lasso
FISTA	Fast Iterative Shrinkage Thresholding Algorithm
GL	Group Lasso
IRLS	Iterative Reweighted Least Squares
LASSO	Least Absolute Shrinkage and Selection Operator

LITERATURA

MP	Matching Pursuit
NSP	Vlastnost nulového prostoru
OMP	Orthogonal Matching Pursuit
PEL	Persistent Elitist Lasso
RIP	Vlastnost zeslabené izometrie
SNR	Signal to noise ratio
VoIP	Voice over Internet Protocol
WGL	Windowed Group Lasso

A. Obsah CD

Příložené CD obsahuje následující zdrojové kódy v programovém prostředí Matlab (R2011b). Pro funkčnost skriptů je potřeba ltfat toolbox [31].

<i>min_sig_supp.m</i>	funkce obsahující algoritmus pro zjištění minimální potřebné délky signálu
<i>overlap.m</i>	funkce určující počet překrývajících se vzorků jednotlivých užitečných oken s dírou
<i>test_min_sig_supp_overl.m</i>	test výše uvedených funkcí – ověření, že koeficienty signálu, jenž je zkrácený od indexu q po index Q podle výstupu funkce <i>min_sig_supp.m</i> , jsou podmnožinou koeficientů původního signálu; funkce <i>overlap.m</i> zobrazí překryvy jednotlivých užitečných oken s dírou