

RESTAURACE POŠKOZENÝCH AUDIOSIGNÁLŮ POMOCÍ ŘÍDKÝCH REPREZENTACÍ

ONDŘEJ MOKRÝ

ABSTRAKT. V tomto článku se zabýváme problematikou doplnění chybějícího úseku vzorků v audiosignálu. Problém formulujeme jako konvexní minimalizační úlohu, kterou řešíme vhodným iterativním algoritmem. U rekonstruovaného signálu přitom požadujeme řídkost jeho reprezentace ve vybraném Gaborově systému. Následně navrhuje modifikaci metody za účelem kompenzace poklesu energie v rekonstruovaném úseku signálu. Porovnání základní a modifikované metody stručně ilustrujeme vybranými experimentálními výsledky¹.

1. ÚVOD

Restaurace audiosignálů je aktuální problematikou a ačkoliv k řešení mnohých poruch existují rozšířené nástroje, objevují se nové přístupy, které cílí buď na větší přesnost rekonstrukce, nebo naopak na co nejvyšší rychlost a zpracování v reálném čase. Jedním z běžných modelů poruchy je chybějící úsek vzorků. Při přenosu signálu může dojít ke chvilkové poruše, čímž vznikne výpadek vzorků, nebo máme k dispozici signál s chybějícími či znehodnocenými úseky – příkladem je záznam na gramofonové desce, která je fyzicky poškozená a digitalizovaný signál tak obsahuje rušivé praskání.

Motivace pro použití řídkých reprezentací vychází z fyzikální podstaty audiosignálu, konkrétně z harmonické struktury hudebního tónu. Díky tomu lze audiosignál lokálně vyjádřit jako součet nemnoha harmonických funkcí (sinus, kosinus) o různé periodě a amplitudě, tedy s vhodnou množinou vektorů (příkladem je níže popsáný Gaborův systém) lze získat řídkou reprezentaci audiosignálu, čehož při řešení problému restaurace audiosignálu využijeme.

Poznamenejme, že ačkoliv tento příspěvek pojednává pouze o doplnění chybějícího úseku signálu, koncept řídkosti audiosignálu lze využít i pro odstraňování šumu nebo tzv. *audio declipping*, při němž je cílem doplnit vzorky, jejichž amplituda přesáhla dovolený dynamický rozsah a signál je v těchto místech omezen určitou hodnotou výchyly.

2010 MSC. Primární 42Cxx.

Klíčová slova. Audiosignál, doplňování chybějících dat, Gaborova transformace, řídké reprezentace, Douglasův-Rachfordův algoritmus, kompenzace poklesu energie, inpainting.

¹Článek vznikl na základě bakalářské práce autora v oboru Matematické inženýrství na FSI VUT v Brně. Vedoucím práce byl Pavel Rajmic z Ústavu telekomunikací FEKT VUT v Brně.

2. ŘÍDKÉ REPREZENTACE

Jak již bylo nastíněno v úvodu, prezentovaný algoritmus bude vyžadovat řídkou reprezentaci signálů. Digitální signál budeme chápat jako vektor $\mathbf{y} \in \mathbb{C}^p$, kde p je počet vzorků signálu². Nyní popíšeme takovou množinu vektorů $G = \{\mathbf{g}_i, i = 1, \dots, q\} \subset \mathbb{C}^p$, která bude generovat prostor \mathbb{C}^p a daný audiosignál $\mathbf{y} \in \mathbb{C}^p$ bude možné vyjádřit jako lineární kombinaci malého počtu prvků z G .

2.1. Gaborovy systémy

Gaborův systém je množina konstruovaná na základě zvoleného vektoru \mathbf{g} , který nazveme oknem a jehož nosič je kompaktní a délky $w \ll p$. Toto okno slouží k časové lokalizaci spektra signálu³. Systém G pak získáme translacemi a modulacemi okna \mathbf{g} a můžeme jej úplně popsat pomocí okna \mathbf{g} a parametrů a (posun mezi sousedními okny, neboli parametr translace) a M (počet modulací okna).

Pro systém G definujeme (lineární) operátor syntézy $\mathcal{G}: \mathbb{C}^q \rightarrow \mathbb{C}^p$,

$$\mathbf{y} = \mathcal{G}\mathbf{c} = \sum_{i=1}^q c_i \mathbf{g}_i, \quad (2.1)$$

a k němu adjungovaný operátor analýzy $\mathcal{G}^*: \mathbb{C}^p \rightarrow \mathbb{C}^q$. Poznamenejme, že protože předpokládáme $q > p$, může pro daný signál \mathbf{y} existovat více vektorů koeficientů \mathbf{c} takových, že $\mathbf{y} = \sum_{i=1}^q c_i \mathbf{g}_i$. Bližší teoretický rozbor operátorů \mathcal{G} a \mathcal{G}^* je nad rámec tohoto textu, na tomto místě proto pouze uvedeme, že pro vhodné okno \mathbf{g} a nastavení parametrů a a M je systém G tzv. Parsevalovým framem (blíže viz [4, 3]), který je z výpočetních důvodů výhodný. Dále budeme pracovat právě s Parsevalovým framem a jeho prvky budeme nazývat atomy.

3. FORMULACE PROBLÉMU A ALGORITMUS

Úlohu doplnění chybějícího úseku signálu budeme nyní formulovat jako optimalizační problém, ve kterém budeme hledat nejjednodušší vektor koeficientů $\mathbf{x} \in \mathbb{C}^q$ rekonstruovaného signálu (samotný signál pak snadno dostaneme jako $\mathcal{G}\mathbf{x}$). Množina přípustných řešení Γ je množina takových reprezentací, že z nich syntetizovaný signál se od původního neliší v nepoškozených úsecích. Formálně

$$\mathbf{x} = \arg \min_{\mathbf{z}} \|\mathbf{z}\|_0 \quad \text{vzhledem k } \mathbf{z} \in \Gamma. \quad (3.1)$$

Protože minimalizace ℓ_0 -pseudonormy (neboli maximalizace řídkosti) je NP-těžký problém, řešíme relaxovanou úlohu

$$\mathbf{x} = \arg \min_{\mathbf{z}} \|\mathbf{z}\|_1 \quad \text{vzhledem k } \mathbf{z} \in \Gamma. \quad (3.2)$$

²Signál chápeme jako komplexní vektor z důvodu Gaborovy transformace, která se obvykle zavádí pro komplexní vektory [3]. V aplikacích, kde jsou signály reálné, se pak omezujeme pouze na reálnou složku \mathbf{y} .

³Proto hovoříme o tzv. časově-frekvenční reprezentaci signálu.

Úloha (3.2) je již úlohou konvexní optimalizace, k jejímu řešení tudíž známe efektivní algoritmy.

3.1. Algoritmus

Definujeme-li *indikátorovou funkci* ι_Γ množiny Γ vztahem

$$\iota_\Gamma(\mathbf{x}) = \begin{cases} 0 & \mathbf{x} \in \Gamma, \\ \infty & \mathbf{x} \notin \Gamma, \end{cases}$$

můžeme úlohu (3.2) formulovat v neomezeném tvaru

$$\mathbf{x} = \arg \min_{\mathbf{z}} \|\mathbf{z}\|_1 + \iota_\Gamma(\mathbf{z}). \quad (3.3)$$

Pro řešení úlohy (3.3) použijeme proximální Douglasův-Rachfordův algoritmus (viz [1]), jehož zjednodušená verze pro náš problém je v algoritmu 1. Funkce soft_τ (tzv. měkké prahování) a proj_Γ (projekce na množinu Γ) jsou tzv. proximální operátory ℓ_1 -normy, resp. indikátorové funkce (blíže viz [1]).

Algoritmus 1: Douglasův-Rachfordův algoritmus pro zaplnění chybějícího úseku signálu

```

1 zvolíme  $\tau > 0, \mathbf{q}_0 \in \mathbb{C}^q$ 
2 for  $n = 0, 1, 2, \dots$  do
3    $\mathbf{x}_n = \text{soft}_\tau(\mathbf{q}_n)$ 
4    $\mathbf{q}_{n+1} = \mathbf{q}_n + \text{proj}_\Gamma(2\mathbf{x}_n - \mathbf{q}_n) - \mathbf{x}_n$ 
5 end
6  $\mathbf{x} = \text{proj}_\Gamma(\mathbf{x}_n)$ 

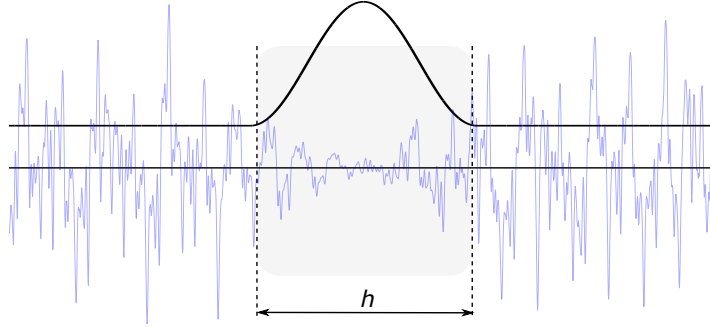
```

Poslední krok algoritmu 1 je oproti obecné formě algoritmu přidán navíc, protože volíme konzervativní přístup, při kterém požadujeme příslušnost výsledného vektoru \mathbf{x} do množiny Γ , i když algoritmus skončí v důsledku zvoleného ukončovacího kritéria dříve, než je dosaženo optimální řešení.

3.2. Kompenzace poklesu energie v rekonstruovaném signálu

Signály rekonstruované výše popsaným algoritmem obecně vykazují pokles energie v místě zaplněné díry (tento jev symbolicky ukazuje obrázek 1). Pro kompenzaci tohoto poklesu energie navrhuje dvě metody: váhování atomů a kompenzaci v časové oblasti.

3.2.1. Váhování atomů. V důsledku ℓ_1 relaxace dochází při optimalizaci kromě požadovaného nulování složek Gaborovy reprezentace též k snižování složek, které zůstanou nenulové a ze kterých tedy syntetizujeme výslednou rekonstrukci. K tomuto dochází realizací operátoru soft_τ , který (zjednodušeně řečeno) zmenšuje složky argumentu o hodnotu τ . Váhování atomů v našem případě znamená přiřazení různých hodnot τ_i jednotlivým složkám Gaborovy reprezentace (namísto konstantního parametru τ). To činíme tak, že atomům, které více přispívají k rekonstrukci díry, přiřadíme menší hodnotu τ_i , aby byly operátorem měkkého prahování



Obrázek 1. Ukázka poklesu energie v místě doplněné díry (díru délky h symbolizuje šedá oblast). Černá křivka zobrazuje návrh na kompenzaci poklesu energie v časové oblasti.

méně penalizovány. Určit váhy splňující tento požadavek lze mnoha způsoby a předeem nelze odhadnout, jaký postup dosáhne nejlepších výsledků. Dále tedy navrhuje čtyři různé možnosti, které reflektují vliv jednotlivých atomů na doplnění chybějícího úseku a jednotlivé možnosti se liší rozptylem vypočítaných hodnot τ_i .

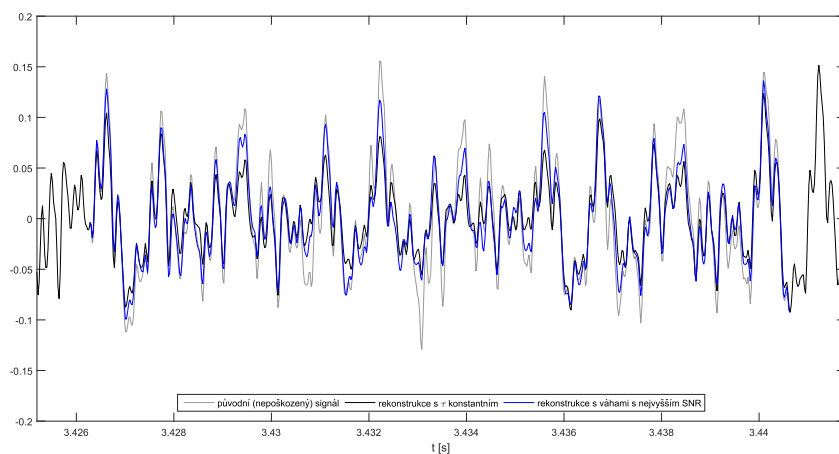
Označme nyní \mathbf{g}_i^r část atomu \mathbf{g}_i , která odpovídá neporušené části signálu. Váhy τ_i pak navrhuje mimo konstantní (neváňované) varianty dle následujících vzorců:

$$\begin{aligned} \text{I. } \tau_i &= \frac{|\text{supp}(\mathbf{g}_i^r)|}{|\text{supp}(\mathbf{g}_i)|}, & \text{II. } \tau_i &= \frac{\|\mathbf{g}_i^r\|_1}{\|\mathbf{g}_i\|_1}, \\ \text{III. } \tau_i &= \frac{\|\mathbf{g}_i^r\|_2}{\|\mathbf{g}_i\|_2}, & \text{IV. } \tau_i &= \frac{\|\mathbf{g}_i^r\|_2^2}{\|\mathbf{g}_i\|_2^2}. \end{aligned}$$

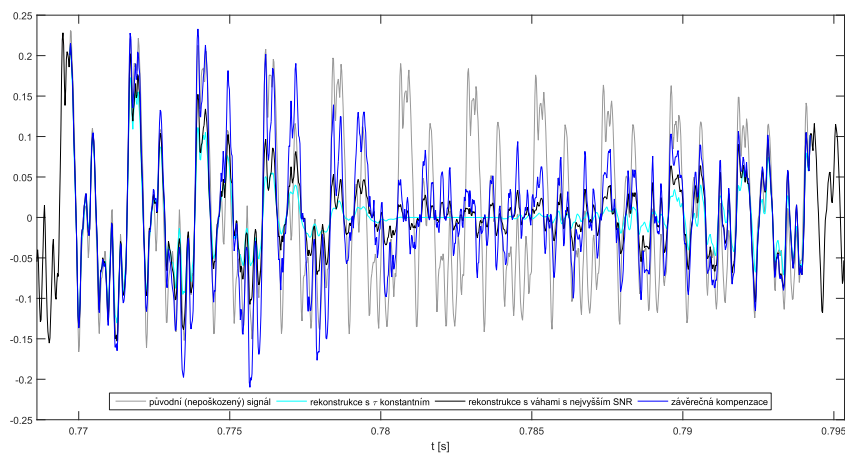
3.2.2. Kompenzace v časové oblasti. Cílem této metody není reagovat přímo na příčinu poklesu energie, jak tomu bylo u váňování atomů. Zde se snažíme amplitudu doplněné části signálu zvýšit prostým vynásobením rekonstruovaného signálu po složkách vhodnou funkcí. Pro ilustraci zde máme na tuto funkci pouze požadavek hladkosti, volíme tedy její hodnoty dle jedné periody (délky h) funkce kosinus posunuté podél osy y tak, aby na hranicích díry hladce navazovala na konstantní funkci s hodnotou 1 (mimo díru signál neměníme). Tento jednoduchý model je ilustrován v obrázku 1.

4. UKÁZKOVÉ VÝSLEDKY

Obrázky 2 a 3 ilustrují použití navržených metod v praxi. V obou je viditelný nárůst amplitudy, je-li použit algoritmus s váňování atomů. Ačkoliv v případě uvedeném v obrázku 3 je pokles energie zřejmý i s použitím obou navržených kompenzačních metod po sobě, dochází zde oproti základní metodě (tyrkysová barva) k eliminaci úseku pouze nulových vzorků a rozdíl mezi nepoškozeným signálem a rekonstrukcí je sluchem nerozeznatelný.



Obrázek 2. Porovnání variant algoritmu (váhování voleno dle varianty IV). Testovaný signál byl záznam zvuku houslí a violy, délka díry $h = 632$ vzorků (se vzorkovací frekvencí 44,1 kHz, tedy přibližně 14 ms).



Obrázek 3. Porovnání variant algoritmu (váhování voleno dle varianty I). Testovaný signál byl záznam zvuku houslí, délka díry $h = 1080$ vzorků (se vzorkovací frekvencí 44,1 kHz, tedy přibližně 24 ms).

5. ZÁVĚR

Prezentovali jsme algoritmus pro doplnění chybějícího úseku audiosignálu založený na předpokladu řídkosti Gaborovy reprezentace signálu. Z dosažených výsledků jsme se omezili na ukázkou dvou jednoduchých experimentů. Širší analýza výsledků

provedená v [2] ukazuje, že z hlediska objektivního hodnocení jsou navržené kompenzační metody zlepšením oproti základnímu algoritmu 1, ovšem ne ve všech případech.

REFERENCE

- [1] P. L. Combettes, J.-C. Pesquet: *Proximal splitting methods in signal processing*, in Fixed-Point Algorithms for Inverse Problems in Science and Engineering, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, H. Wolkowicz (eds.), 185–212, Springer-Verlag, New York, 2011.
- [2] O. Mokrý: *Restaurace poškozených audiosignálů pomocí řídkých reprezentací*, Bakalářská práce, Fakulta strojního inženýrství, Vysoké učení technické v Brně, Brno, 2017.
- [3] G. E. Pfander: *Gabor Frames in Finite Dimensions*, 193–239, Birkhäuser Boston, Boston, 2013.
- [4] P. Rajmic, M. Daňková: *Úvod do řídkých reprezentací signálů a komprimovaného snímání*, Vysoké učení technické v Brně, Brno, 2014.

Ondřej Mokrý, Ústav matematiky, Fakulta strojního inženýrství, Vysoké učení technické v Brně, Technická 2, 616 69 Brno, Česká republika,
e-mail: ondrej.mokry@mail.com