

Oponentský posudek disertační práce

Název práce: Inpainting of Missing Audio Signal Samples

Autor práce: Ing. Václav Mach

Předložená práce se zabývá doplňováním chybějících vzorků v audio záznamech a redukcí šumu. Z praktického hlediska je cíleno na restaurování digitalizovaných záznamů z nejstarších zařízení, jako je například fonograf. Digitální archivy takových záznamů obsahují velké množství nahrávek, o jejichž zrestaurování a zkvalitnění lze usilovat v současnosti a znovu i v budoucnosti, přičemž neustále bude možné využívat nejnovějších poznatků z oblasti teorie informace a zpracování signálů. Téma práce a použité metody jsou z hlediska současného stavu aktuální. Opírají se o nejnovější metody rekonstrukce řídkých signálů pomocí konvexních optimalizací. Aktuálnost dokládá i stáří citované literatury, která má k obsahu práce nejblíže.

První kapitola práce obsahuje úvod a praktickou motivaci. Popisuje jaké typy poškození a nekvality obsahují nahrávky ze starých zařízení a poskytuje přehled. Druhá kapitola popisuje konvenční metodiky pro rekonstrukci chybějících vzorků, které jsou postaveny na autoregresních modelech signálů v časové oblasti nebo na autoregresních parametrických modelech a na metodách minimalizace kvadratických kritérií. Třetí až pátá kapitola pak popisují nové metody, které jsou v práci použity za účelem překonání konvenčních postupů. Zavedeny jsou nové pojmy a značení, popsána je problematika řídkých reprezentací signálů a jejich rekonstrukce z neúplné informace, dále metody učení slovníků pro řídké reprezentace a nakonec hladové a konvexní optimalizační algoritmy pro výpočet řešení.

Tím končí teoretická část práce, ke které mám dvě důležité výhrady. První se týká nedostatku detailů, hloubky popisu a značení. Metody jsou popsány poměrně povrchně a čtenář je převážně odkázán na literaturu. Například podstata známé a poměrně jednoduché metody K-SVD není řádně vysvětlena. Značení často není plně definované a je nepřehledné. Příkladem za všechny budiž popis zmíněné metody K-SVD na straně 51. Kapitola 5 popisuje nejdůležitější metody, kterými je v práci dosaženo hlavních výsledků. Bohužel i zde detaily chybí. Například definice strukturované řídkosti není srozumitelná, čtenář se prakticky nedozvídá, jak konkrétně autor problém řeší, co je sociální řídkost atp. Druhým nedostatkem je fakt, že autor použil pouze stávající metodiky a v teoretické části nenavrl žádné vlastní řešení ani žádný vlastní nápad na zlepšení. Znamená to, že přínosem práce je pouze experimentální studie a konkrétní implementace metod. To dokládá i šestá kapitola, ve které jsou tyto cíle zformulovány.

Původní výsledky práce obsahuje rozsáhlá sedmá kapitola. Autor v ní použil všechny metody zmíněné v teoretické části. Každou metodu testuje na jednotlivých signálech při různých nastaveních a výsledky vyhodnocuje pomocí objektivního kritéria, které nazval SNR. Jedná se ve skutečnosti o normalizovanou kvadratickou vzdálenost mezi rekonstruovaným a skutečným signálem. Správnější označení pro toto kritérium je tedy NMSE. Autor výsledky v každém experimentu vyhodnocuje a diskutuje. V závěru kapitoly se pak krátce věnuje experimentu, kde byl redukován šum v digitalizovaných nahrávkách, které byly zaznamenány fonografem na voskované válečky. Poslední experiment pak porovnává všechny metody na konkrétním vzorku signálu pomocí kritérií SNR a PEMO-Q.

Zvolený způsob experimentální studie není příliš vhodný. Jedná se o dílčí pokusy na jednotlivých vzorcích různých druhů hudebních signálů (kytara, bicí, zpěv, apod.). Výsledky jsou často velmi nejednoznačné, protože výsledné grafy mají velký rozptyl a jsou citlivé na změnu parametrů. Často se těžko vyvozuje nějaký

věcný závěr než jen konstatování, že výsledek je takový, jaký je. Autor tento postup mnohokrát opakoval. Čtenář jeho textu se musí probírat nezáživným výčtem výsledků a je pro něj těžké udržet pozornost. Závěr, že nové metody jsou lepší než stávající, je sice evidentní z ukázky na přiloženém CD, ale přesně objektivně to výsledky práce potvrzují málo. Metody měly být testovány na co největším počtu vzorků a vyhodnoceny pomocí průměrů nebo mediánů vhodných kritérií. Závěry by po takovém testování byly, v případě pozitivních výsledků, mnohem relevantnější.

Na přiloženém CD je kompletní balík funkcí a skriptů v jazyce Matlab. Kódy jsou přehledné a komentované. Chybí zde však strukturovaný návod na použití balíku a příklady použití. Namátkou jsem vyzkoušel skript `inpainting_structured_sparsity.m`, bohužel Matlab (R2013b) skončil běhovou chybou, jejíž příčinu neznám. K balíku schází grafické uživatelské rozhraní, které by uživateli zjednodušilo používání. V práci není uvedeno, jaký je podíl autora na tomto balíku (autoři jsou tři).

Práce je napsaná v anglickém jazyce, což je dobré, protože tím se stává dostupná i čtenářům v zahraničí. Bohužel je však úroveň jazyka velmi nízká. Text obsahuje nesčetně mnoho gramatických chyb, jako jsou chybějící členy, čárky ve větách, nevhodná slova, pořádek slov ve větách apod.

Typograficky je práce na vysoké úrovni díky tomu, že je napsaná v LaTeXu. Zde bych pouze vytkl nedostatečné popisky obrázků, kde často chybí zmínka experimentu, ke kterému se výsledek vztahuje. V práci také chybí seznam symbolů, který by pomohl zlepšit srozumitelnost.

Předložená práce má své kvality a přínos. Nejdůležitějším výsledkem je srovnání analyzujícího modelu se syntetizujícím a poukázání na fakt, že analyzující model dosahuje srovnatelných výsledků při 2 až 4 krát menší výpočetní zátěži. Podle mého hodnocení je ale práce celkově průměrná. Mohla být lepší, ale zřejmě se na ní negativně projevil spěch. Experimentální části mělo být věnováno mnohem více úsilí, už proto, že v ní spočívá těžiště přínosu.

Publikační činnost autora je dostačující. Zároveň je zřejmé, že autor je odborníkem v oblasti zlepšování hudebních nahrávek, což dokládá jeho publikační činnost, předchozí praxe a účast na projektech.

Přes zmíněné výhrady **práci považuji za odpovídající obecně uznávaným požadavkům k udělení akademického titulu Ph.D. a doporučuji ji k obhajobě.**

V Liberci dne 9. června 2016

Doc. Ing. Zbyněk Koldovský, Ph.D.

Ústav informačních technologií a elektroniky

Proděkan pro vědu a výzkum a doktorské studium

Vedoucí skupiny pro analýzu a zpracování akustických signálů