

Extrakce zdrojů na lokalizovaných masivně multimikrofonních polích s redukováným počtem demixujících parametrů

Tomáš Kounovský <tomas.kounovsky@tul.cz>, Zbyněk Koldovský <zbynek.koldovsky@tul.cz>

Extrakce zdroje na poli s velkým počtem mikrofonů nabízí vysoký extrakční výkon za cenu velkého množství parametrů k odhadu. Pokud je mikrofonní pole malé velikosti, pak lze potenciálně využít jeho geometrie a blízkosti mikrofonů a nahradit kompletní parametrizaci parametrizací částečnou. Ta by měla popisovat extrakční parametry pro menší podmnožinu mikrofonů celého pole a pomocí jednoduchých transformací tyto parametry adaptovat pro zbytek mikrofonů. Předběžné výsledky ukazují, že za určitých podmínek lze využít tohoto předpokladu a zvýšit extrakční výkon nad úroveň plně parametrizovaného konvenčního algoritmu se stejným počtem extrakčních parametrů.

Klíčová slova: Masivní multimikrofonní pole, pole s malou clonou, extrakce zdroje, beamforming

Úvod

Lokalizovaná masivně multimikrofonní pole (LMMP) jsou pole o velkém počtu mikrofonů situovaných na relativně malém prostoru, která mohou představovat ideální řešení pro případy, kdy je potřeba jak vysoký extrakční výkon, tak skladnost a přenositelnost zařízení. LMMP se v odborné literatuře příliš nevyskytuje, nejbližší je porovnání velikosti multimikrofonních polí [1]. LMMP je realizovatelné např. pomocí MEMS mikrofonů [2]. Problém u konvenčních metod (slepá separace, beamforming) využívajících takové pole může být v odhadu velkého množství extrakčních parametrů, což může být výpočetně náročné. Tato práce se zabývá metodou snížení počtu extrakčních parametrů využitím geometrie a velikosti LMMP.

Metodika

Hlavní myšlenka této práce spočívá v rozdělení celého LMMP na menší množiny mikrofonů (subpole). Pokud je LMMP dostatečně malé, dá se předpokládat, že ideální extrakční vektor bude velmi podobný mezi dvěma subpoli, jenž mají stejnou geometrii a jsou prostorově blízko u sebe. Toho lze využít vytvořením strukturovaného extrakčního vektoru pro všechny mikrofony v LMMP, který obsahuje úplné extrakční parametry pouze pro jedno obecné subpole. Tento obecný extrakční vektor je pak pomocí jednoduchých transformací adaptován na všechna subpole v LMMP, čímž výrazně snižuje celkový počet odhadovaných parametrů.

Jako základ navrhované metody slouží adaptovaný MMSE beamformer, který řeší

$$\operatorname{argmin}_{w_g, \beta_1, \dots, \beta_N} \|S - \sum_{i=1}^N \beta_i^H w_g^H X_i\|^2, \quad (1)$$

kde w_g je obecný extrakční vektor pro jedno subpole, β_i jsou adaptační parametry (komplexní skalár - zesílení a zpoždění) pro i -té subpole, X_i jsou data pro i -té subpole a S je signál extrahovaného zdroje. Řešení vede k iterativnímu algoritmu, který v jednotlivých krocích adaptuje w_g dle

$$w_g = \left(\sum_{i=1}^N \sum_{j=1}^N \beta_i^H \beta_j X_i X_j^H \right)^{-1} \sum_{i=1}^N \beta_i^H X_i S^H \quad (2)$$

a následně dopočítá β_i ze soustavy

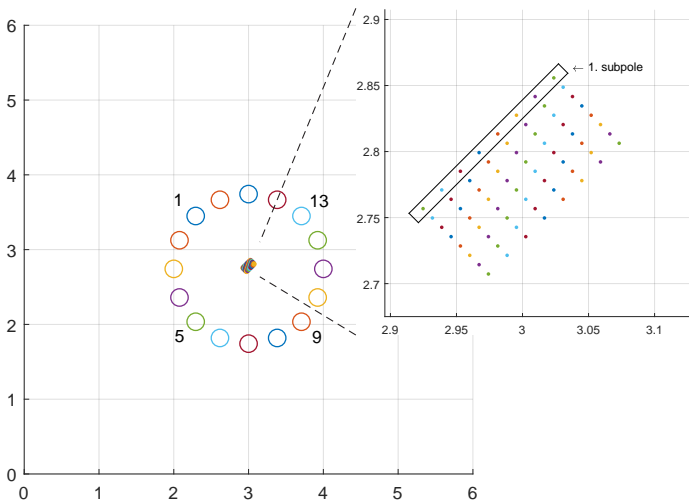
$$\sum_{i=1}^N w_g^H X_i X_i^H w_g \beta_i = w_g^H X_i S^H \quad (3)$$

dokud není dosaženo požadované konvergence (z hlediska střední čtvercové chyby). Tento algoritmus budeme nazývat LMMP beamformer.

Výsledky a diskuze

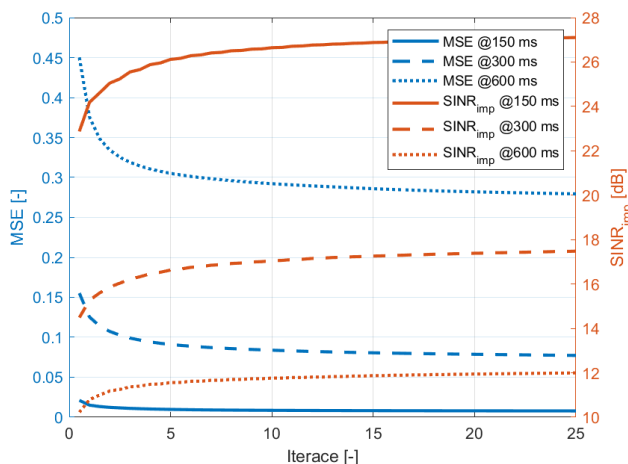
LMMP beamformer byl otestován z hlediska rychlosti konvergence a porovnán s MPDR, MVDR a MMSE beamformery z hlediska extrakčního výkonu v různých simulovaných podmínkách. Všechny experimenty byly simulovány pomocí RIR generátoru [3] v místnosti o rozměrech 6x6x3 metry, viz obrázek 1. Místnost byla simulována ve třech reverberačních časech 150, 300 a 600 ms. Nahrávky

o vzorkovací frekvenci 16 kHz byly zaznamenávány polem zhruba uprostřed místnosti o 64 mikrofonech s mezimikrofonními vzdálenostmi 2x1 cm. Řečové signály pocházely z celkem 16 pozic rozmístěných do kruhu kolem LMMP.



Obrázek 1: Simulované prostředí experimentálních LMMP

Rychlost konvergence můžete vidět v obrázku 2. Algoritmus konverguje v řádu jednotek iterací i v nejtěžších podmínkách.



Obrázek 2: Konvergence LMMP beamformery

V dalším experimentu jsou porovnány výsledky z hlediska zlepšení SINR mezi konvenčními beamformery a LMMP beamformery. Výsledky jsou v tabulce 1. Subskript $MMSE_i$ označuje počet subpolů použitých k extrakci, např. $MMSE_2$ využívá 2 subpole o 8 mikrofonech, tudíž 16 parametrů; LMMP beamformer využívá w_g o 8 parametrech + 8 transformačních parametrů β . Pokud je problém extrakce jednoduchý (1 zdroj a 1 interference - 1v1), pak

LMMP beamformer dosahuje zhruba stejných výsledků jako MMSE beamformer se stejným počtem extrakčních parametrů. Pokud je však úloha těžká (1 zdroj, 14 interference a nekorelovaný bílý šum - 1v14+WGN), pak dosahuje LMMP beamformer lepších výsledků než metoda se stejným počtem extrakčních parametrů. To ukazuje, že pro velmi obtížné úlohy může metoda využívající většího počtu mikrofónů dosahovat lepších výsledků i s menším počtem extrakčních parametrů, a tedy stupňů volnosti.

Tabulka 1: Zlepšení SINR [dB]

| Metoda | $T_{60} - 1v1$ | | | $T_{60} - 1v14+WGN$ | | |
|-------------------|----------------|--------------|--------------|---------------------|-------------|-------------|
| | 150 | 300 | 600 | 150 | 300 | 600 |
| LMMP | 28,02 | 18,34 | 12,53 | 8,87 | 7,35 | 5,94 |
| MVDR | 49,36 | 36,47 | 29,03 | 10,04 | 9,30 | 8,07 |
| MPDR | 6,16 | 5,80 | 5,58 | 2,75 | 2,46 | 2,02 |
| MMSE ₁ | 25,87 | 16,25 | 10,97 | 7,45 | 6,43 | 5,36 |
| MMSE ₂ | 31,24 | 20,58 | 14,20 | 8,08 | 6,92 | 5,76 |
| MMSE ₃ | 32,53 | 22,40 | 15,95 | 8,52 | 7,27 | 6,05 |
| MMSE ₄ | 32,98 | 23,02 | 16,45 | 8,88 | 7,53 | 6,26 |
| MMSE ₅ | 33,20 | 23,24 | 16,64 | 9,21 | 7,76 | 6,43 |
| MMSE ₆ | 33,40 | 23,47 | 16,89 | 9,52 | 7,97 | 6,57 |
| MMSE ₇ | 33,48 | 23,59 | 16,97 | 9,82 | 8,17 | 6,71 |
| MMSE | 33,54 | 23,66 | 16,99 | 10,13 | 8,37 | 6,85 |

Závěr

V této práci byl nastíněn postup pro snížení počtu extrakčních parametrů pro algoritmy extrahující signál z velkého množství mikrofónů na malé ploše pomocí využití malé vzdálenosti mezi mikrofony a geometrie mikrofoniho pole. Výsledky ukazují, že tento postup může najít využití v obtížných úlohách s mnoha řečníky v zašuměném prostředí. V budoucí práci bych rád odvodil a analyzoval tento algoritmus pro metody slepé separace.

Poděkování

Tato práce byla podpořena z projektu Studentské grantové soutěže (SGS) na Technické univerzitě v Liberci v roce 2020.

Reference

- [1] SACHAR, J. et. al. *Large vs small aperture microphone arrays: Performance over a large focal area*. IEEE ICASSP. 2001, 3049–3052.
- [2] FULÍNOVÁ, V. *Aplikace ovládání a vizualizace mikrofoniho pole..* 2016. [Diplomová práce] Technická Univerzita v Liberci.
- [3] HABETS, A. *Room impulse response generator*. 2006.: no. 2.4, vol. 2. Tech. Rep. Technische Universiteit Eindhoven.