

# Slepá extrakce pohybujícího se zdroje (lineárně se měnící mixovací transformace)

Jaroslav Čmejla <jaroslav.cmejla@tul.cz>, Zbyněk Koldovský

Tato práce se zabývá problémem slepé extrakce zdroje ze směsi. Cílem je vytvoření metody pro extrakci zdroje ze směsí, u kterých je předpokládána lineárně v čase se měnící mixovací transformace. V práci je vysvětlena teorie problému slepé separace/extrakce zdrojů, která je použita jako základ k odvození nového dynamického modelu. Součástí práce je návrh algoritmu pro extrakci zdroje z dynamických směsí, jeho otestování a porovnání s konvenčními metodami EFICA a OGICE. V závěru práce je část věnována diskusi budoucího postupu pro řešení problému.

**Klíčová slova:** slepá extrakce zdroje, slepá separace zdrojů, analýza nezávislých komponent, lineární model

## Úvod

V telefonii, při streamování videokonferencí, ovládání chytrých asistentů, ale i při prostém pořizování záznamů je žádoucí, aby byla zaručena čistota signálu požadovaného zdroje (řečníka). Při přenosu hlasu požadujeme čistý příjem, i když jeden z účastníků komunikuje ze zaručeného prostředí. Správnou odezvu od hlasových asistentů požadujeme i v případech, kdy při ovládání asistenta posloucháme muziku, nebo v pozadí hraje výstup z televize. Ve všech těchto situacích lze využít systémů zpracování signálů pro zlepšování řečových nahrávek (z angl. speech enhancement), případně systémů určených pro separaci zdrojů nebo přímo k extrakci určitého zdroje.

Tato práce se zabývá problémem extrakce požadovaného zdroje, SOI (Source of Interest), z dynamických směsí, které byly nahrávány skupinou synchronizovaných mikrofonů. Metoda, která je odvozena v této práci, využívá statistické nezávislosti jednotlivých zdrojů a je modifikací metody určené k extrakci požadovaného zdroje zvané OGICE (Orthogonally Constrained Independent Component Extraction) [1] a její blokované verze BOGICE (Blockwise OGICE) [2]. Úprava ICE modelu je provedena za účelem extrakce zdroje z dynamických směsí, u kterých lze předpokládat lineární změnu mixovacích parametrů požadovaného zdroje.

## Metodika

V případě použití ICA | ICE metod je typické, že se separace | extrakce zdrojů provádí v časově-frekvenční oblasti s využitím STFT (Short Time Fourier Transform,

krátkodobá Fourierova transformace). Proces separace | extrakce se poté provádí sekvenčně po jednotlivých frekvencích. Dynamický mixovací (1a) a demixovací (1b) proces v rámci jedné frekvence lze popsat jako:

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{s}_t, \quad (1a) \quad \mathbf{s}_t = \mathbf{W}_t \mathbf{x}_t, \quad (1b)$$

kde  $t$  značí časový index,  $\mathbf{x}$  značí vektor směsi,  $\mathbf{s}$  značí zdrojové signály,  $\mathbf{A}$  označuje mixovací transformaci a  $\mathbf{W}$  označuje demixovací transformaci. Za použití ICE modelu (viz [1]) a za předpokladu modelování reálných signálů lze mixovací a demixovací matice v čase  $t$  vyjádřit v následující struktuře:

$$\mathbf{A}_t = \begin{pmatrix} \mathbf{a}_t & \mathbf{Q}_t \end{pmatrix} \quad (2a) \quad \mathbf{W}_t = \begin{pmatrix} \mathbf{w}^T \\ \mathbf{B}_t \end{pmatrix} \quad (2b)$$

kde  $\mathbf{a}_t$  značí mixovací vektor požadovaného zdroje ( $s_t$ ),  $\mathbf{Q}_t$  značí mixovací transformaci ostatních zdrojů,  $\mathbf{w}$  značí vektor extrahující požadovaný zdroj ( $s_t = \mathbf{w}^T \mathbf{x}_t$ ) a  $\mathbf{B}_t$  značí transformaci separující ostatní zdroje ( $\mathbf{z}_t = \mathbf{B}_t \mathbf{x}_t$ ). V tomto modelu nejsou ostatní zdroje předmětem optimalizace a matice  $\mathbf{B}_t$  a  $\mathbf{Q}_t$  jsou zvoleny tak, aby platilo:  $\mathbf{A}_t^{-1} = \mathbf{W}_t$  a  $\mathbf{B}_t \mathbf{a}_t = 0$ .

Hlavním předpokladem nového modelu je lineární změna mixovacího vektoru mezi počátečním bodem ( $\mathbf{a}_1$ , v čase 1) a koncovým bodem ( $\mathbf{a}_N$ , v čase  $N$ ). Mixovací vektor  $\mathbf{a}_t$  lze tedy vyjádřit jejich lineární kombinací:

$$\mathbf{a}_t = \left(1 - \frac{t-1}{N-1}\right) \mathbf{a}_1 + \left(\frac{t-1}{N-1}\right) \mathbf{a}_N. \quad (3)$$

Tento předpoklad zaručí redukci počtu odhadovaných pa-

rametrů. Zde je nutno podotknout, že vektor  $\mathbf{w}$  není závislý na čase, tj. předpokládáme existenci jednoho extrahujícího vektoru pro celou délku signálu. Pro zachování nezkresleného výstupu je  $\mathbf{w}$  a  $\mathbf{a}_t$  (pro všechna  $t$ ) provázáno přes tzv. distortionless constraint (podmínka nezkresleného výstupu)  $\mathbf{w}^T \mathbf{a}_t = 1$ .

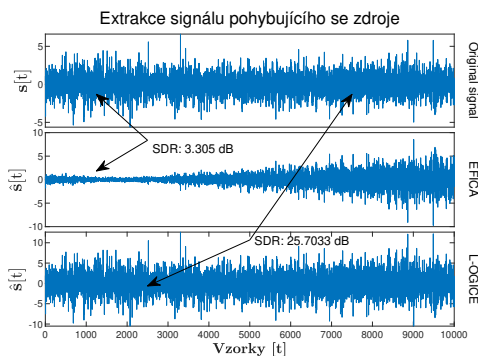
Odhad extrahujícího vektoru je založen na použití **metody maximalizace věrohodnosti**. Rozepsáním statistického modelu směsi a nahrazením skutečných hustot pravděpodobnosti jednotlivých signálů vhodnou náhradou (SOI - typicky Laplaceovo rozložení, ostatní zdroje - Gaussovo rozložení) vznikne účelová kontrastní funkce:

$$\mathcal{C}(\mathbf{a}_1, \mathbf{a}_N, \mathbf{w}) = \log f(\mathbf{w}^T \mathbf{x}_t) - (\mathbf{x}_t^T \mathbf{B}_t^T \mathbf{C}_z^{-1} \mathbf{B}_t \mathbf{x}_t) / 2 + (d - 2) \log |\gamma_t|, \quad (4)$$

kde  $\mathbf{C}_z$  je kovarianční matice ostatních zdrojů  $E[\mathbf{Z}\mathbf{Z}^T]$ ,  $\gamma_t$  je první prvek mixovacího vektoru a  $f(\cdot)$  je vhodná náhrada skóre funkce (podle rozložení SOI). Gradient této funkce (podle  $\mathbf{w}$ ) je poté využit k odhadu extrahujícího vektoru metodou největšího spádu (resp. růstu). Vzniklý gradientní algoritmus je pracovně nazván **L-OGICE**.

## Výsledky a diskuze

Vytvořený algoritmus byl podroben testování na simulovaných datech vygenerovaných tak, aby přesně odpovídaly struktuře modelu. Požadovaný zdroj byl simulován signálem s Laplaceovským rozložením, interference byla simulována signálem s Gaussovým rozložením (bílý šum). Výstup z algoritmů byl srovnán podle zkrácení extrahovaného zdroje, SDR (Signal-to-Distortion Ratio).

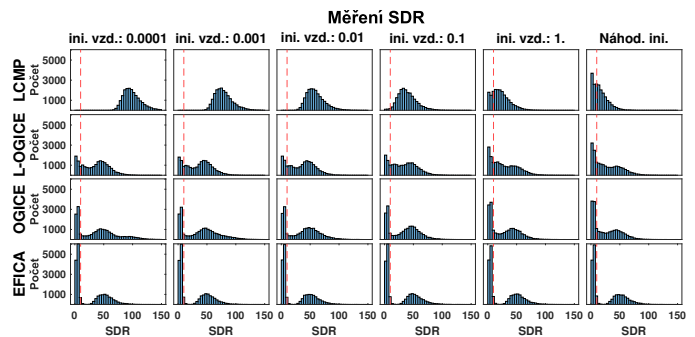


Obrázek 1: Výsledek extrakce pohyblivého zdroje. 1. graf - původní zdroj (SOI), 2. graf - extrahovaný signál metodou EFICA, 3. graf - extrahovaný signál metodou L-OGICE.

Vybraný výsledek extrakce zdroje (signálu dlouhého 10000 vzorků) z dynamické směsi je znázorněn na ob-

rázku č. 1. V případě extrakce konvenční neblokovanou metodou (EFICA [3]) je možné pozorovat zaměření algoritmu na jednu pozici zdroje. Při vzdálení od zaměřené pozice dojde k útlumu požadovaného zdroje. Extrahovaný signál pomocí metody L-OGICE netrpí tímto fenoménem.

Z rozsáhlejšího testování (viz obrázek č. 2) je u konvenčních metod (EFICA, OGICE) možné pozorovat zvýšené množství případů, kdy se nepodařilo vyextrahovat požadovaný zdroj (SDR < 10dB). V případě L-OGICE dochází ke snížení počtu těchto případů.



Obrázek 2: Histogramy výstupního SDR rozsáhlejšího testování (10000 opakování, signály délky 2500 vzorků). Zleva doprava jsou vykresleny výsledky s různou inicializací (vlevo přesná). Srovnávané metody: 1. řádek - LCMP beamformer, 2. řádek - L-OGICE, 3. řádek OGICE, 4. řádek EFICA

## Závěr

V práci došlo k odvození nového algoritmu určeného k extrakci zdroje z dynamického prostředí (pohyby zdrojů). Prvotní experimenty ukazují přínos použití algoritmu a je plánováno rozšíření experimentů na složitější scénáře. Dále je nutné provést důkladnější analýzu případů, kdy nová metoda selhala.

## Poděkování

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

## Reference

- [1] KOLDOVSKÝ, Z.; TICHAVSKÝ, P. Gradient Algorithms for Complex Non-Gaussian Independent Component/Vector Extraction, Question of Convergence. *IEEE Transactions on Signal Processing*.
- [2] KOLDOVSKÝ, Z.; MÁLEK, J.; JANSKÝ, J. Extraction of Independent Vector Component from Underdetermined Mixtures through Block-wise Determined Modeling. In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*
- [3] KOLDOVSKÝ, Z.; TICHAVSKÝ, P.; OJA, E. Efficient Variant of Algorithm FastICA for Independent Component Analysis Attaining the Cramér-Rao Lower Bound. *IEEE Transactions on Neural Networks*