

1. はじめに

<研究背景>

- 非負値行列因子分解 (nonnegative matrix factorization: NMF) [1]
 - 時間周波数領域の信号に適用
 - 周波数スペクトルとそのパワーに行列分解することができる
 - 解釈性の高い表現を有する各行列それぞれに着目することで新たな音響的特徴を得ることができる
- Euclidean metric regulated (erNMF) [2]
 - 従来のNMFではスケールの任意性が生じる
 - 得られた音響的特徴を基底ごとに比較することができない
 - 基底行列に対して正規化処理を行うことでNMFの行列分解時におけるスケールの任意性を解決

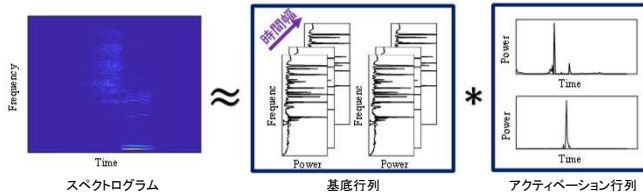
<研究目的>

- 厳密な音響的特徴を表現するためには非定常音である音声に適した動的なモデリングを実現させる必要がある
- 非定常音である音声でも対応が期待できるConvolutional NMFに着目
 - NMFと同様に強度の任意性が問題
- これを解決するためにユークリッドノルムの正規化を用いたConvolutional NMFを提案する

2. Convolutional NMF [3]

<Convolutional NMF>

- 基底行列に時間変化成分を想定するようにNMFを拡張した手法
 - 時間変化するパターンの抽出も可能となる



スペクトログラムの近似には影響しないが、基底・アクティベーションのどちらに大きさが振り分けられるか不明⇒基底間の音響的特徴が比較できない

⇒視覚化による比較のためには、基底行列の正規化が必要

3. 提案手法

<基底の正規化>

- 基底ベクトル $t_{k,0} \in T_0$ についてユークリッドノルムを計算
- T_c をユークリッドノルムで正規化
- T_0 のユークリッドノルムをアクティベーション行列にかけて補正する

<定式化>

$$Z \approx \sum_{c=0}^{C-1} T_c \vec{V}^c$$

$$= \sum_{c=0}^{C-1} \frac{T_c}{\|T_0\|} (\|T_0\| \circ \vec{V}^c)$$

4. 音響的可視化のための評価実験

- 提案手法がどの程度の性能及び有用性を実験を通して確認

<実験条件>

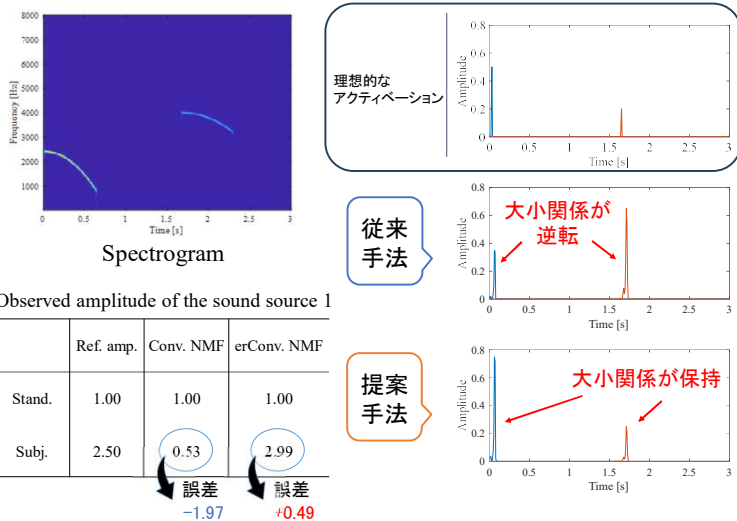
実験条件は右に示す表のとおりである。

また、NMFの基底数は音源数、Convolutional NMFの時間幅は予備実験で最も良好な分離性能を示したものを使用した。

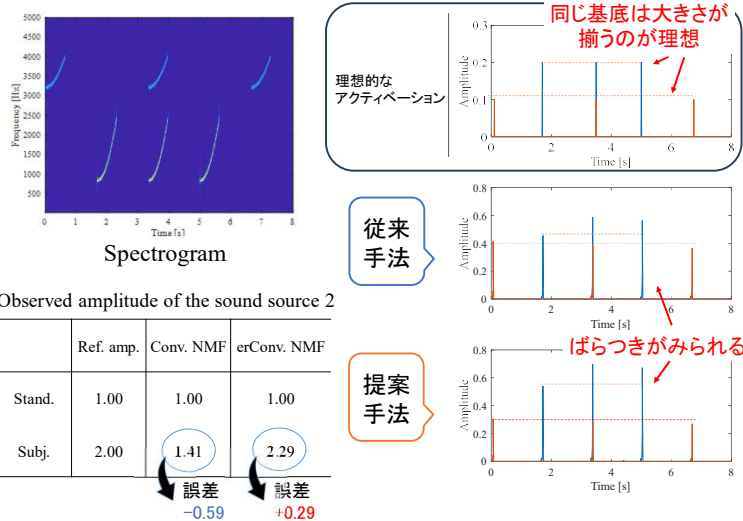
Experimental conditions

Parameter	Value
Sampling frequency	16,000 Hz
FFT length	512 samples
Shift length	128 samples
Number of iteration	200 times

<実験1: 音源の重複を含まない分離信号>



<実験2: 音源の重複を含む分離信号>



⇒提案手法を用いることでスケールの任意性が解消、各行列において基底間の音響的特徴の比較が可能となった

5. おわりに

<まとめ>

- 基底の正規化処理を有した新たなerConvolutional NMFを提案
- 実験から、提案手法は従来手法と比較して推定信号における各基底の大小関係について安定した表現を得ること確認

<今後の予定>

- アクティベーションの振幅の大きさについて安定した表現を実現するための検討
- 複雑な音源における有用性の確認

参考文献

- [1] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization", *Nature*, Vol. 401, No. 6755, pp. 788–791, 1999.
- [2] 大村美結, 大島風雅, 中山仁史, "非負値行列因子分解に基づく音響的変化の可視化手法の基礎検討" 日本音響学会2023年秋季研究発表会講演論文集, pp.1349–1352, 2023
- [3] A. Ozerov and C. Févotte, "Multichannel nonnegative matrix factorization in convolutional mixtures for audio source separation", *IEEE Trans. ASLP*, vol. 18, no. 3, pp. 550–563, 2010.