



# 韓国語を対象とした音声分解技術

梨和 美佑, 大島 風雅, 中山 仁史

広島市立大学 大学院情報科学研究科

## 1. はじめに

- 一般的な音声認識・・・音素単位の音響マッチング

従来の音素アライメントは各音素の時間的重複を考慮せず音声波形を切り出している

音素 { 子音 ... 音源や調音構造の違いにより周波数特性は異なる  
母音

我々は音素や音声記号単位で音源が独立していると着想し、音声を対象とした音源分離の可能性と可視化を試みる[1].

- なぜなら、各音素を別音源とみなした分離が実現すれば
  - 一時系列的なオーバーラップを含む本来収録時可能な音声信号の抽出および解析が期待できる
  - 発声メカニズムの更なる解明や、新たな音響特徴量として利用可能
- 非負値因子分解(Non-negative Matrix Factorization ; NMF)[1]
  - 周波数スペクトルと各音色のアクティベーションに分解できる.
- 単音節に含まれる音素数・・・閉音節 ≥ 開音節
  - 本研究では閉音節言語である韓国語に注目し、各音素へ分解し音響的特徴を明らかにする

## 2. 韓国語の音節構造・特徴

単音節 { 開音節・・・最終音素が母音  
閉音節・・・最終音素が子音

韓国語の音節構造

|     |       |           |
|-----|-------|-----------|
| 開音節 | V     | 例 이 [i]   |
| 閉音節 | C V   | 例 파 [pa]  |
| 閉音節 | V C   | 例 알 [al]  |
| 閉音節 | C V C | 例 팔 [pal] |

日本語にほぼ存在しない  
※撥音, 拗音など一部例外あり

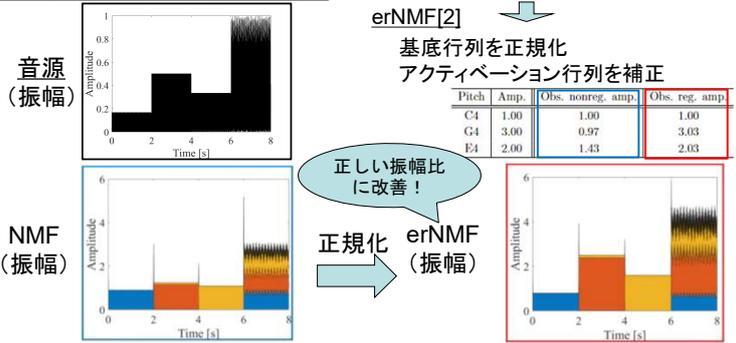
韓国語は他の言語と比較して,  
①単音節の種類, ②単音節に含まれる音素の数が多い

## 3. 非負値行列因子分解

従来のNMF・・・振幅の任意性

erNMF[2]

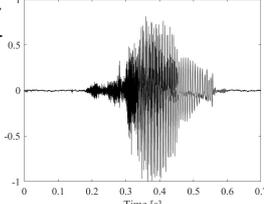
基底行列を正規化  
アクティベーション行列を補正



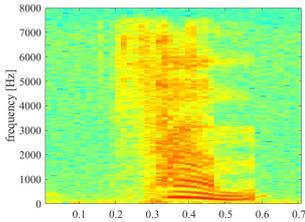
## 4. 韓国語単音節の音響モデリング

1. 対象音源: 単音節 /sam/  
KOR\_F2\_KOR\_W3.wav

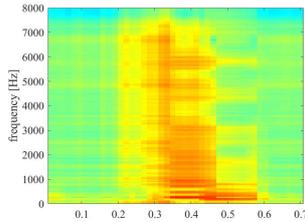
基底数: 3



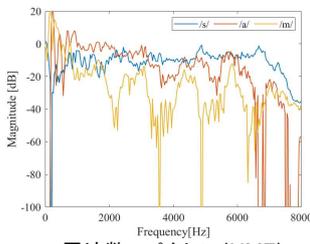
音声波形



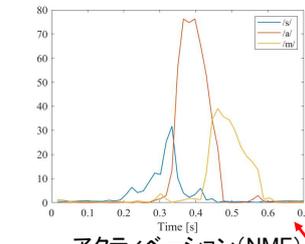
スペクトログラム(DGT)



スペクトログラム(NMF)



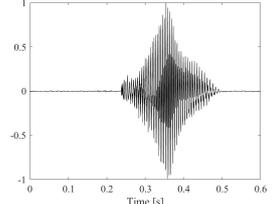
周波数スペクトル(NMF)



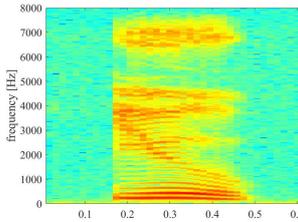
アクティベーション(NMF)

2. 対象音源: 単音節 /yuk/  
KOR\_F1\_KOR\_W6.wav

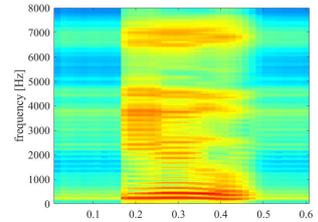
基底数: 4



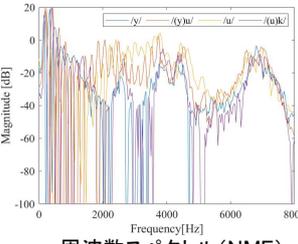
音声波形



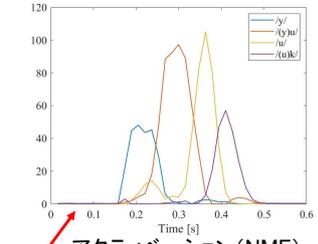
スペクトログラム(DGT)



スペクトログラム(NMF)



周波数スペクトル(NMF)



アクティベーション(NMF)

- アクティベーションの振幅比は元の音声波形と等しい  
→正規化を用いることで、各音素との大きさが比較できる
- 音素が重複した区間を音素の混合比として可視化  
→単音節に含まれる音素数が増えても音素分解の音源分離が可能

時間的な重複を考慮した分解を実現  
(周波数特性の混合比や強度を可視化)

## 5. おわりに

- 閉音節言語である韓国語単音節においても時間的な重複を考慮した分解の可能性を確認した
- 単音節の種類によって再構成できる基底数が異なる
- 分解結果が実音源と再合成音との音響的差を有する  
→音源を増加させて確認。人間の聴覚に適した音声解析ツールの構築.

## 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 他, "非負値行列因子分解に基づく音響的变化の可視化手法の基礎検討", 2023年日本音響学会秋季研究発表会, 2023.  
 [3] 筑波大学 知能情報生体工学研究室, "平成9-13年度 筑波大学特別プロジェクト「東西言語文化の類型論」多言語音声コーパス (UT-ML)", 国立情報学研究所 音声資源コンソーシアム, 2006.