創造工学専攻特別研究||発表審査会@マルチメディア教室 12:50~16:00

深層パーミュテーション解決法に基づく ブラインド音源分離 Blind Source Separation Based on Deep Permutation Solver

香川高専専攻科 創造工学専攻 電気情報コース 2年 北村研究室 蓮池 郁也

はじめに

音源分離とは

- 音声, 雑音, 歌声, 楽器音, 機械音等の音源を個々に分離





- ・音源分離の応用先
 - 音声認識
 - 自動採譜
 - 補聴器の高機能化
 - ノイズキャンセリング etc.



ブラインド音源分離

ブラインド音源分離 (blind source separation: BSS) - 混合系Aが未知の条件で分離系Wを推定



- <mark>優決定</mark>BSS(マイク数≧分離したい音源の数)
 - 混合系が正方行列にできるので逆行列が定義可能
 - 線形分離が可能なため分離音の音質が良い

例. 独立成分分析(ICA) [Comon, 1994] 独立ベクトル分析 (IVA) [Hiroe, 2006], [Kim+, 2006] 独立低ランク行列分析 (ILRMA) [Kitamura+, 2016]

・本研究では、優決定BSSについて取り扱う - 高音質であり、様々な分野に適用可能

BSSの歴史



本発表の目次

- 従来手法
 - 独立成分分析(ICA)と周波数領域ICA(FDICA)におけるパー
 ミュテーション問題
 - パーミュテーションを回避する手法
 - 従来の深層パーミュテーション解決法(DPS)
- 提案DPS
 - 概要とパーミュテーション行列の推定方法
 - 推定分離信号の作成と損失の計上
 - テストデータに対する処理
- 実験
 - 提案DPSの性能を調査するための実験
 - FDICAの分離誤差に対する実験
- まとめ

本発表の目次

- 従来手法
 - 独立成分分析(ICA)と周波数領域ICA(FDICA)におけるパー
 ミュテーション問題
 - パーミュテーションを回避する手法
 - 従来の深層パーミュテーション解決法(DPS)
- 提案DPS
 - 概要とパーミュテーション行列の推定方法
 - 推定分離信号の作成と損失の計上
 - テストデータに対する処理
- 実験
 - 提案DPSの性能を調査するための実験
 - FDICAの分離誤差に対する実験
- まとめ

音源間の独立性に基づくBSS:ICA

- 独立成分分析(independent component analysis: ICA)[Comon, 1994]
 - 混合行列Aが未知の条件で分離行列Wを推定



- 2つの仮定を用いて分離行列Wを推定
 - 1. 独立成分は互いに独立(音源は多くの場合独立)
 - 2. 混合行列は可逆で時不変(優決定, 音源やマイクは移動しない)
- 分離信号の順番(パーミュテーション)は決定できない



ICAに基づくBSSの耐残響性の向上

- **周波数領域ICA(FDICA)**[Smaragdis, 1998]
 - 各周波数ビンの複素時系列に対して独立なICAを適用



FDICAで生じるパーミュテーション問題

周波数領域独立成分分析(FDICA)におけるパーミュ
 テーション問題



Time

パーミュテーション問題を回避する手法

- 独立ベクトル分析(IVA)
 - 各信号源は全周波数成分の 強弱が同期すると仮定
- 独立低ランク行列分析(ILRMA)
 - 各信号源は時間周波数構造が
 低ランクな構造(繰り返しを多分
 に含む)を持つと仮定







深層パーミュテーション解決法(DPS)の動機

11

 ・ 音源ごとの時間周波数構造の違い



 - 音源モデルが音源に適していない場合, IVAやILRMAのBSS

 の精度は低下

- さまざまな音源に適応する万能な音源モデルの作成は困難

• パーミュテーション問題の解決のみをDNNを用いて実装

- さまざまな音源に適応するモデルを作成できる可能性あり



DPS [Yamaji+, 2020] が提案される(2音源のみ適用可能)

本発表の目次

- 従来手法
 - 独立成分分析(ICA)と周波数領域ICA(FDICA)におけるパー
 ミュテーション問題
 - パーミュテーションを回避する手法
 - 従来の深層パーミュテーション解決法(DPS)
- 提案DPS
 - 概要とパーミュテーション行列の推定方法
 - 推定分離信号の作成と損失の計上
 - テストデータに対する処理
- 実験
 - 提案DPSの性能を調査するための実験
 - FDICAの分離誤差に対する実験
- まとめ

提案深層パーミュテション解決法(DPS)

パーミュテーション行列をDNNを用いて予測





[、] DNNを用いて推定

前処理

• パーミュテーション不整合信号 Y_n に対して正規化処理



DNNの構造

・ DNNは入力層, BiLSTM層3層, 出力層の計5層で構成



推定パーミュテーション行列の導出

• DNNの出力を用いてパーミュテーション行列 \hat{P}_{ij}^{-1} を作成





16

局所時間推定分離信号の導出

行列積を用いて推定分離信号を作成



損失の導出方法

- ・損失関数の設計
 - 推定分離信号と完全分離信号との間で平均二乗誤差(mean squared error: MSE)を導入
 - 分離信号の順序は予測の対象としないため, 順序不変学習 (permutation invariant training: PIT)[Yu+, 2017]を導入

$$\text{Loss} = \min \left(||\hat{\tilde{Z}}_{j1} - \check{Z}_{j1}||_2^2 + ||\hat{\tilde{Z}}_{j2} - \check{Z}_{j2}||_2^2 , ||\hat{\tilde{Z}}_{j1} - \check{Z}_{j2}||_2^2 + ||\hat{\tilde{Z}}_{j2} - \check{Z}_{j1}||_2^2 \right)$$



本発表の目次

- 従来手法
 - 独立成分分析(ICA)と周波数領域ICA(FDICA)におけるパー
 ミュテーション問題
 - パーミュテーションを回避する手法
 - 従来の深層パーミュテーション解決法(DPS)
- 提案DPS
 - 概要とパーミュテーション行列の推定方法
 - 推定分離信号の作成と損失の計上
 - テストデータに対する処理
- 実験
 - 提案DPSの性能を調査するための実験 - FDICAの分離誤差に対する実験
- まとめ

実験条件

- 本実験の目的
 - 提案DPSの性能を調査する(クリーンな信号を使用)
- 比較手法
 - 局所時間に基づくDPS [Yamaji+, 2020]
 - 提案DPS
- 評価指標
 - 信号対歪み比(source-to-distortion ratio: SDR)[Vincent+, 2006]
- 実験データ
 - SiSEC2011より男女の音声及びドラムとギターの音楽信号

音響の種類	音響信号	ファイル名	信号長 [s]
 	男性	dev2_male4_inst_src_2	10.0
目	女性	dev3_female4_inst_src_2	10.0
音楽ドラムdev1ぜターdev1	dev1_wdrums_src_3	11.0	
	ギター	dev1_wdrums_src_2	11.0

実験条件

- 学習データ
 - 音声信号及び音楽信号の時間周波数信号を周波数ビン単位 でランダムにシャッフルしたデータ
 - シャッフルパターンは150
- テストデータ

 デストデータ
 学習データとは重複しない10パターン を用いてランダムに入れ替えたデータ



- 2つのモデル作成
 - 音声モデル:男女の音声信号(2種類)を用いて作成
 - 音楽モデル:ギターとドラムの音楽信号(2種類)を用いて作成
- 2種類のテスト条件: in-domainとout-of-domain
 - In-domain: 学習データとテストデータの音源が同じ
 - Out-of-domain: 学習データとテストデータの音源が異なる

実験条件

テスト条件: in-domain



テスト条件:out-of-domain

学習データと重複しない パーミュテーションパターン



実験結果(in-domainデータセット)

 どちらの条件においても、提案DPSのSDRの改善が従 来DPSを上回った

音声信号のin-domainに対するSDR値 [dB]

テストデータ パターン	観測信号	従来DPS	提案DPS
1	-6.25	3.60	44.5
2	-6.85	4.65	44.5
3	-5.40	3.60	44.5
4	-6.45	3.55	44.5
5	-6.60	4.70	44.5
6	-6.45	4.65	44.5
7	-6.35	3.60	44.5
8	-5.50	4.65	44.5
9	-5.85	3.60	44.5
10	-5.55	4.65	44.5

音楽信号のin-domainに対するSDR値 [dB]

23

テストデータ パターン	観測信号	従来手法	提案DPS
1	-0.95	2.95	64.75
2	2.00	2.95	64.75
3	0.55	2.95	155.00
4	1.25	2.95	64.75
5	-1.00	2.95	66.65
6	-1.00	2.95	61.15
7	-0.85	2.95	66.65
8	-0.15	2.95	64.75
9	0.60	2.95	64.75
10	-0.35	2.95	61.15

実験結果(out-of-domainデータセット)

- ・ 音楽信号のout-of-domainに対して提案DPSはSDRの
 の善が見られた
- ・ 音声信号のout-of-domainに対して提案DPSでSDR値
 の改善は見られたものの従来DPSが優勢

音声信号のout-of-domainに 対するSDR値 [dB]

			-
テストデータ パターン	観測信号	従来DPS	提案DPS
1	-0.95	5.05	3.35
2	2.00	5.05	1.75
3	0.55	5.05	3.35
4	1.25	11.35	3.35
5	-1.00	11.35	3.35
6	-1.00	11.35	3.35
7	-0.85	11.35	3.35
8	-0.15	5.05	3.35
9	0.60	5.05	3.35
10	-0.35	5.05	1.75

音楽信号のout-of-domainに 対するSDR値 [dB]

テストデータ パターン	観測信号	従来DPS	提案DPS
1	-6.25	-8.00	33.55
2	-6.85	-5.85	22.85
3	-5.40	-7.20	33.85
4	-6.45	-7.60	23.50
5	-6.60	-7.40	22.00
6	-6.45	-7.25	24.05
7	-6.35	-1.40	23.60
8	-5.50	-7.65	26.65
9	-5.85	-6.40	25.15
10	-5.55	-7.90	24.05

本発表の目次

- 従来手法
 - 独立成分分析(ICA)と周波数領域ICA(FDICA)におけるパー
 ミュテーション問題
 - パーミュテーションを回避する手法
 - 従来の深層パーミュテーション解決法(DPS)
- 提案DPS
 - 概要とパーミュテーション行列の推定方法
 - 推定分離信号の作成と損失の計上
 - テストデータに対する処理
- 実験
 - 提案DPSの性能を調査するための実験
 - FDICAの分離誤差に対する実験
- まとめ

FDICAにおける分離誤差

実際のFDICAの出力は推定誤差を含む



DPSでは、FDICAの推定誤差の影響で分離精度が低下



FDICAの推定誤差を考慮した学習データの作成が必要

FDICAの推定誤差を考慮した学習データ

27

• FDICAの推定誤差量を行列Rを用いて模倣



実験条件

- 比較手法
 - パーミュテーション解決法(PS)を用いないFDICA(PS: none)
 - 音源の到来方向情報(DOA)によるPSを用いたFDICA(PS: DOA)[Saruwatari+, 2006]
 - 提案DPSを用いたFDICA(PS: DPS)
 - 独立ベクトル分析(IVA)[Ono, 2011]
 - 理想的なPS(IPS)を用いたFDICA(PS: IPS)
 - FDICAに基づくBSSの上限性能
- 評価指標
 - SDR (Source-to-distortion ratio) [Vincent+, 2006]

実験条件

- ・学習データ
 - SiSEC2011より, ドラムとギターの11秒程度の2種類の 音楽信号のみ使用
- テストデータ
 - JVSコーパスより男女の100セット分の音声信号を使用 - 部屋の生成条件は学習データと同一



実験結果

FDICAの分離誤差を含んだ信号に対してもある程度
 パーミュテーション問題を解決できている



まとめ

目的

- さまざまな音源に適応可能なPSの構築
- 省サンプルデータで汎用性が高いモデルの作成
- 提案手法
 - 周波数方向再帰に基づくDPSを提案
- 結果
 - ワンショットの音楽信号で学習したモデルが,良好な精度で音 声信号のパーミュテーション問題を解決できた

• 研究業績

- 蓮池郁也,北村大地,渡辺瑠伊,川口翔也,"周波数双方向再帰に基づく深層パーミュテーション解決法," *電子情報通信学会 第*37回信号処理 シンポジウム,pp. 308-313, 2022.
- Fumiya Hasuike, Daichi Kitamura, and Rui Watanabe, "DNN-based frequency-domain permutation solver for multichannel audio source separation," *in Proc. APSIPA ASC*, pp. 872–877, 2022.
- 蓮池郁也,北村大地,渡辺瑠伊,"深層パーミュテーション解決法の汎化性能に関する実験的評価,"日本音響学会 2022 年秋季研究発表会講 演論文集,1-Q-28, pp. 351-354, 2022.
- 蓮池郁也,渡辺瑠伊,北村大地,"深層ニューラルネットワークに基づくパーミュテーション解決法の基礎的検討," 電子情報通信学会 技術研究 報告, EA2022-13, vol. 122, no. 20, pp. 62-67, 2022.
- 蓮池郁也,北村大地,"深層パーミュテーション解決法に基づくブラインド音源分離の性能評価,"日本音響学会 2024 年秋季研究発表会講演論 文集,1-R-25, pp. 235-238, 2024.