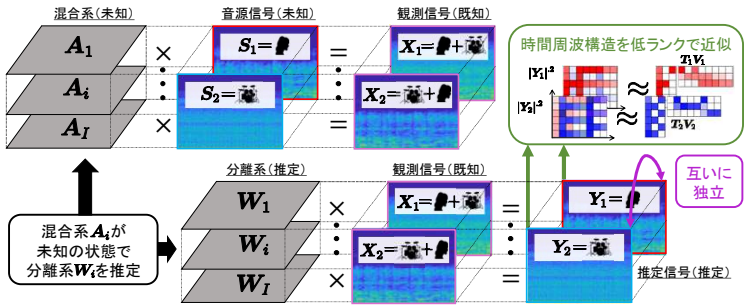
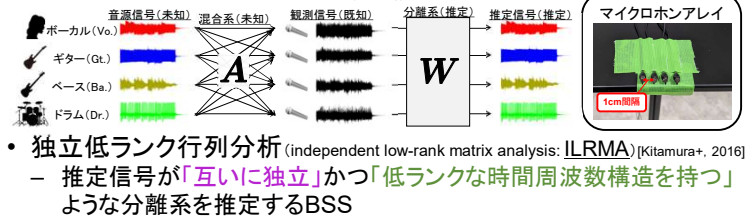
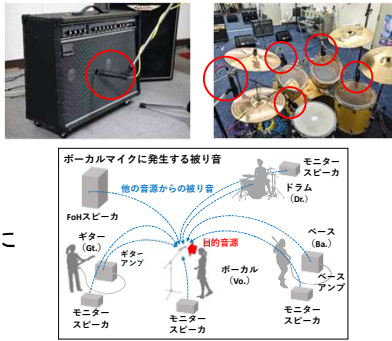


セミブラインド型独立低ランク行列分析による 音楽ライブ演奏時の被り音抑圧

☆谷野宮 蒼士, 北村 大地(香川高専), 猿渡 洋, 高宗 典玄, 山岡 洸瑛(東京大学),
高橋 祐, 山川 颯人, 近藤 多伸(ヤマハ株式会社)

1. 研究背景

- 被り音 (bleeding sound)
 - 音楽ライブ演奏の録音時, 目的の音のみを録音するために楽器やアンプにマイクを近接
 - 実際には目的音以外の音である, 被り音が混入
 - 被り音はミキシングや演奏に悪影響を及ぼすため, 抑圧することが望まれる
- ブラインド音源分離 (blind source separation: BSS)
 - 音源やマイクの空間的な配置 (混合系 A) が不明な状態で分離系 W を推定する技術
 - 通常はマイクロホンアレイの観測信号に対して適用
- 独立低ランク行列分析 (independent low-rank matrix analysis: ILRMA) [Kitamura+, 2016]
 - 推定信号が「互いに独立」かつ「低ランクな時間周波数構造を持つ」ような分離系を推定するBSS



ILRMAでの被り音抑圧の問題点
音楽ライブ演奏時は, 各マイクが数メートル離れている
位相差が正確に測定できず, 抑圧性能が低下してしまう

ILRMAでは
実用的なレベルの
被り音抑圧は
難しい

本研究の内容

- 音楽ライブ演奏時のマイキングで実際に発生している被り音を収録し, 被り音の影響の大きさについて解析する
- 従来手法に代わる, より被り音抑圧性能の高い新たな被り音抑圧手法を開発し, その性能を調査する

2. 被り音調査実験

- インパルス応答測定実験
 - 実験手順
 - ライブハウスにて音楽ライブ演奏を模したステージをセッティングし, 次の2つのケースにおける各マイクのインパルス応答を測定する

マイク近接スピーカー
Vo: Foxtex 6301DT
Gt: Marshall 1980
Ba: Ampeg SVT-810E
Dr: Foxtex 6301DT

その他のスピーカー
FoH: NEXO GEOM12
モニター: YAMAHA DXR12

マイク
Ch. 1 (Vo): SM58
Ch. 2 (Gt): SM57
Ch. 3 (Ba): SM57
Ch. 4 (Dr): SM58

アンプ (付属スピーカー無し)
Gt: Benchmark AHB2
Ba: Benchmark AHB2

Case A (FoH・モニター無し)

近接マイク基準の相対パワー	音源1 (Vo.)	音源2 (Gt.)	音源3 (Ba.)	音源4 (Dr.)
マイク1 (Vo.)	0.0	-17.3	-12.9	1.2
マイク2 (Gt.)	-56.3	0.0	-25.3	-29.4
マイク3 (Ba.)	-59.7	-34.8	0.0	-31.2
マイク4 (Dr.)	-72.7	-44.6	-40.6	0.0

Case B (FoH・モニター有り)

近接マイク基準の相対パワー	音源1 (Vo.)	音源2 (Gt.)	音源3 (Ba.)	音源4 (Dr.)
マイク1 (Vo.)	0.0	-9.8	-6.9	1.3
マイク2 (Gt.)	-31.5	0.0	-16.9	-19.8
マイク3 (Ba.)	-32.5	-21.5	0.0	-23.1
マイク4 (Dr.)	-43.9	-34.1	-31.4	0.0

Vo.マイクの被り音相対エネルギーは大きい
Dr.マイクの被り音相対エネルギーは小さい

3. 提案手法

- セミブラインド型独立低ランク行列分析 (semi-blind independent low-rank matrix analysis: Semi-blind ILRMA)
 - 「ある特定のマイクに, ある1つの音源からの音のみが入っている」状況に特化したILRMA
 - 音楽ライブ演奏時に, 被り音が小さいマイクが存在することに着目

2音源かつ X_2 のマイクの被り音が小さい場合

S_1 の音が X_2 のマイクにあまり混入していないため音源 S_2 を既知として扱える

被り音が小さいマイクをリファレンスマイクとして, ILRMAを適用

リファレンスにより, 推定する分離系は $W_i = \begin{bmatrix} w_{1,i} & w_{2,i} \\ 0 & 1 \end{bmatrix} V_i$ となる

分離系の一部のパラメータが固定されるため, 推定すべきパラメータが減る

被り音が小さいため $X_2 \approx S_2$ とみなす

X_2 をリファレンスとするため $Y_2 = X_2$ とする

性能調査の観点

- 被り音抑圧性能 (非リファレンスマイクのSDR改善量)
- モデルの破綻に対する頑健性
 - リファレンスマイクの被り音を大きくしていき, 抑圧性能の変化を確認

4. 被り音抑圧実験

- Semi-blind ILRMA適用実験
 - 実験手順
 - 問題の簡単化のため, Vo.マイクとDr.マイクの2chでの被り音抑圧を実施

Case A (FoH・モニター無し)

近接マイク基準の相対パワー	音源1 (Vo.)	音源2 (Dr.)
マイク1 (Vo.)	0.0	1.2
マイク2 (Dr.)	-72.7	0.0

Case B (FoH・モニター有り)

近接マイク基準の相対パワー	音源1 (Vo.)	音源2 (Dr.)
マイク1 (Vo.)	0.0	1.3
マイク2 (Dr.)	-43.9	0.0

被り音が小さいため Dr.マイクをセミブラインドILRMAのリファレンスマイクとする
Vo.マイクの被り音を抑圧

実験結果

- 従来手法 (ILRMA) と提案手法 (Semi-blind ILRMA) とで Vo.マイクの被り音抑圧性能を比較
- SDR改善量が高いほど被り音抑圧性能が高い
- いずれの窓長においても **提案手法が従来手法の性能を上回っている**ことが確認できた

- セミブラインドモデルの頑健性調査
 - 実験手順
 - リファレンスマイクに僅かに混入している被り音を人工的に増幅させていく
 - セミブラインド仮定の破綻が被り音抑圧性能に与える影響を調査する

Dr.マイクに混入する Vo.音源の音 (被り音) を増幅する
→ Dr.マイクのSNRを低下させる

増幅無し (0dB増幅) 60dB増幅

近接マイク基準の相対パワー	音源1 (Vo.)	音源2 (Dr.)
マイク1 (Vo.)	0.0	1.3
マイク2 (Dr.)	-43.9	0.0

被り音を10dBずつ増幅

Case B. STFT窓長 1145[ms]

- 従来手法と提案手法とで Dr.マイクのSNR低下に対する被り音抑圧性能の変化を比較
- 提案手法は **抑圧性能の低下が始まるまでにある程度余裕がある**ことが確認できた

ILRMAをセミブラインドの形にして推定パラメータを減らすことで, 被り音抑圧にも適用できることを確認