

マイクロホンアレイを用いたドラムセット音源分離の データセット収録・公開

☆森末 結, 北村 大地(香川高専)

1. 研究背景

・ドラムセット収録時のマイキング

- バスドラムやスネア等に
マイクを近接させる
→ マルチトラック録音と呼ぶ

- 近接させたドラムパーツの音
(目的音)のみを録音するため

・被り音問題

- マイクを近接させても他のパーツ
の音が混入
→ この混入音を被り音と呼ぶ

- 被り音はミキシングの音質を劣化

- 目的音はそのままに, 被り音のみを
抑圧することが重要

・被り音抑圧 ≡ 音源分離問題

- マルチトラック録音はマイクの間隔が広い
→ ブラインド音源分離(BSS)が困難

- マイクアレイを使えばBSSで被り音を
抑圧できるかはあまり調査されていない → そもそも実験用
データセットもない

研究目的

- ・ マイクアレイを用いてドラムセットの演奏を収録・公開
- ・ BSSでドラムセットの被り音抑圧がどの程度可能か調査

2. データセットの収録・公開

・収録の条件

- 使用したドラムパーツ

- バスドラム(BD)
- スネア(SD)
- ハイハット(HH)
- クラッシュ(CC)

- 使用したマイク(計16ch同期録音)

- 4chマイクアレイ(演奏者の右肩付近)
- 8chマイクアレイ(ドラムセット正面1m付近)
- BD・SD・HH・CCの各々の近接マイク

- 同時に演奏するドラムパーツ

- BD・SD・HH・CCの全て(同時演奏)
- BDのみ
- SDのみ
- HHのみ
- CCのみ

音源分離における
正解の信号を取得
するため

- 演奏パターン(drums1/2/3/4の4種)

- drums1: BD・SD・クローズHHの3種のみの8ビート
- drums2: drums1にオープンHHとCCを加えたもの
- drums3: BPM80のスローテンポな8ビート
- drums4: BPM150のアップテンポな16ビート



・データセットの概要

- フォルダ構成

```
drumsBss
├── dataset
│   ├── drums1
│   │   ├── take1~3 (64 files each)
│   │   └── drums2~4
│   │       ├── take1~3 (80 files each)
│   └── score (4 files)
├── fig (10 files)
└── README.md
```



Zenodoにて公開!

- ファイル名

➢ real_scoreType_takeNum_sorceType_micNum.wav

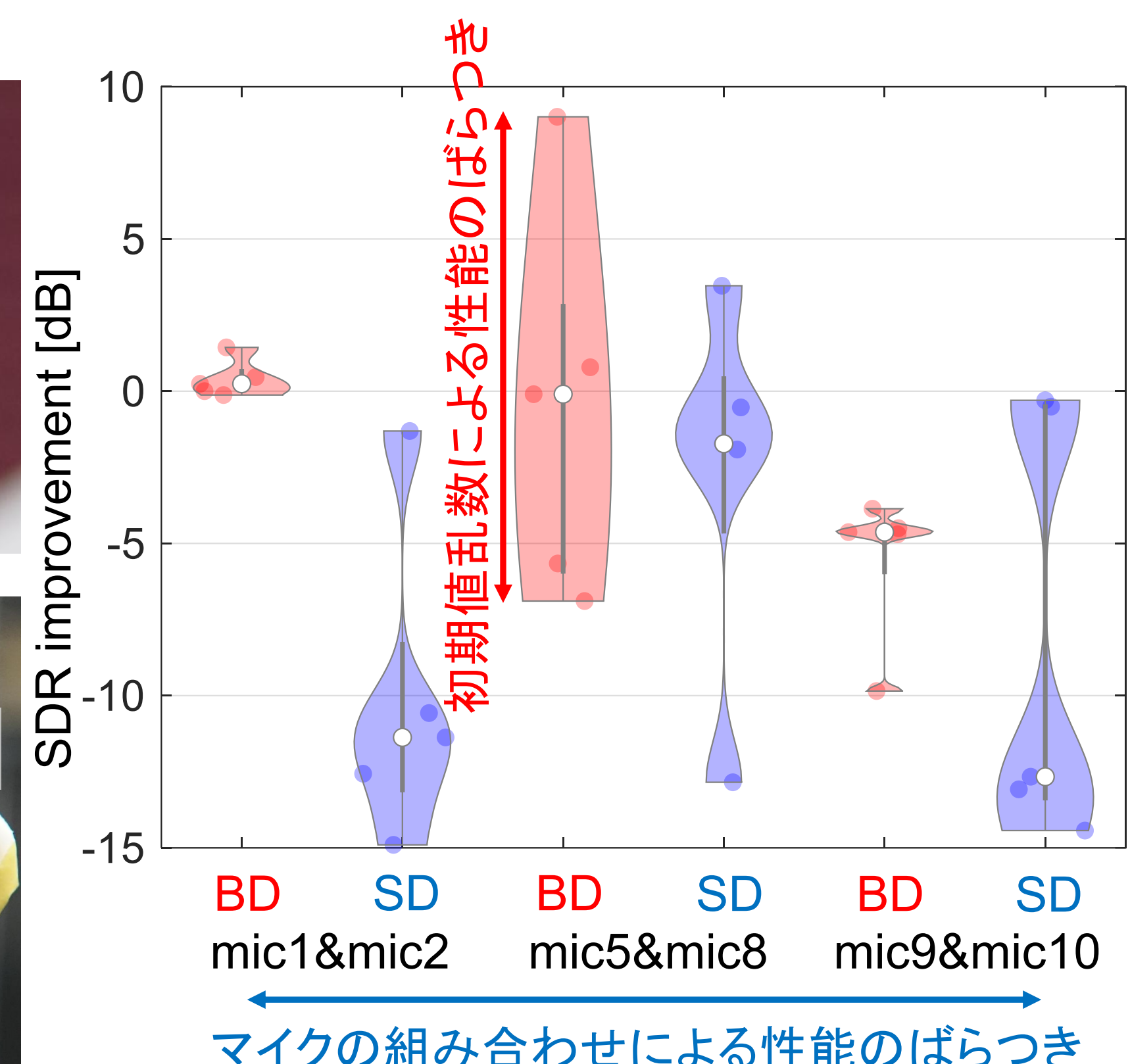
scoreType	演奏パターン	drums1/drums2/drums3/drums4
takeNum	演奏のテイク数	take1/take2/take3
sorceType	演奏する ドラムパーツ	src0: 全ドラムパーツ src1: BDのみ src2: SDのみ src3: HHのみ src4: CCのみ
micNum	マイクのチャンネル	mic1~mic4: 4chマイクアレイ mic5~mic12: 8chマイクアレイ mic13: BD近接マイク mic14: SD近接マイク mic15: HH近接マイク mic16: CC近接マイク

3. 音源分離実験

・収録したデータセットにBSSを適用

- 独立低ランク行列分析(ILRMA) [Kitamura+, 2016]
- 信号対歪み比(SDR) [Vincent+, 2006] を用いて
分離性能を評価
➢ SDRは「分離された目的音の品質」と
「被り音の抑圧量」を加味した指標
- 2音源(BDとSD)2マイクでBSSを実行

サンプリング 周波数	44.1 kHz
使用した ドラムパターン	drums1
使用したテイク数	take3
窓長	371.5 ms
シフト長	92.9 ms
ILRMAの基底数	1音源当たり4本 最適化の 反復更新回数
	100



マイクの 組み合わせ	観測SDR (BD) [dB]	観測SDR (SD) [dB]	乱数 シード	推定SDR (BD) [dB]	推定SDR (SD) [dB]	SDR改善量 (BD) [dB]	SDR改善量 (SD) [dB]
mic1 & mic2	-11.1	11.0	1	-10.7	0.4	0.5	-10.6
			2	-9.7	9.7	1.4	-1.3
			3	-11.3	-3.9	-0.1	-14.9
			4	-10.9	-0.4	0.2	-11.4
			5	-11.1	-1.6	0.0	-12.6
mic5 & mic8	-5.4	6.2	1	-11.1	4.5	-5.7	-1.7
			2	3.6	9.7	9.0	3.5
			3	-12.3	4.3	-6.9	-1.9
			4	-4.7	5.7	0.8	-0.5
			5	-5.5	-6.6	-0.1	-12.8
mic9 & mic10	-5.4	10.1	1	-15.3	9.8	-9.8	-0.3
			2	-9.9	-4.4	-4.5	-14.4
			3	-10.1	-2.6	-4.7	-12.7
			4	-9.3	9.5	-3.9	-0.5
			5	-10.1	-3.0	-4.6	-13.1

- SDR改善量が負の場合が多い → 分離性能は高くなく, 不安定
- 条件による性能のばらつき

・理想的な分離性能を検討

- 各音源信号の(分離された)パワースペクトログラムを音源モデルに
与えるILRMA(理想ILRMA)

マイクの 組み合わせ	観測SDR (BD) [dB]	観測SDR (SD) [dB]	推定SDR (BD) [dB]	推定SDR (SD) [dB]	SDR改善量 (BD) [dB]	SDR改善量 (SD) [dB]
mic1&mic2	-11.1	11.0	-1.2	11.9	9.9	0.9
mic1&mic4	-11.0	8.1	-0.7	12.1	10.3	4.0
mic5&mic6	-5.4	6.2	3.7	9.7	9.1	3.5
mic5&mic8	-5.4	6.2	3.5	9.5	9.0	3.3
mic9&mic10	-5.4	10.1	-1.9	9.3	3.5	-0.8
mic9&mic12	-5.3	10.1	-2.5	8.3	2.9	-1.8

大きく性能が向上

性能があまり
向上していない

- 理想ILRMAでは分離性能が全体的に向上

➢ マイクの組み合わせによってはSDR改善量が負の場合もあり
様々な条件・方法によりマイクロホンアレイを用いた
ドラムセット音源分離を研究する余地がある!!