

3次元マイクロホンアレイを用いた多音源ブラインド分離

Blind Source Separation of Many Speech Signals Using Small 3-D Microphone Array

向井 良
Ryo Mukai

澤田 宏
Hiroshi Sawada

荒木 章子
Shoko Araki

牧野 昭二
Shoji Makino

日本電信電話株式会社, NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, NTT Corporation

1 はじめに

多音源ブラインド分離 (Blind Source Separation, BSS) のための試作システムを開発した [1]。本システムは 4cm 角の立方体の頂点に位置する 8 つのマイクロホンから成るアレイを用い, 3 次的に配置された多数の音源から到来する信号の分離を実現する (図 1)。観測された混合信号は周波数領域における独立成分分析 (Independent Component Analysis, ICA) によって処理され, 最大 8 つの信号に分離される。

2 ICA を用いた周波数領域 BSS

ICA は BSS を実現するために用いられる主要な統計的手法の一つである。我々は周波数領域に変換した信号に対して ICA を適用する周波数領域 BSS の手法を用いている。本システムにおける処理の流れを図 2 に示す (簡単のため, 音源数=マイク数=2 の場合を示している)。時間領域において畳み込み混合された観測信号を, 短時間フーリエ変換によって, 周波数領域における複数の瞬時混合信号に変換する。次に, 各周波数ビンにおいて ICA を適用し, 分離行列と分離信号を得る。ICA の解には出力チャンネルに関する順序のあいまいさがあるため, 結果を時間領域に戻す前に, 全周波数ビンで順序を揃える必要がある。これはパーミュテーションの問題と呼ばれる。パーミュテーションの問題を解決するための代表的手法として, 音源到来方向 (DOA) を用いる方法や出力信号のエンベロープの相関を用いる方法がある。どちらの方法も単独では十分な性能が得られなかったり, 性能が不安定であったりするため, 本システムでは両者を併用することにより, 頑健かつ正確にパーミュテーションを解決している [2]。また, DOA 推定部分では, 3 次的に配置された音源に対応するための推定手法 [3] を用いている。

3 実験

本システムの性能を評価するために, 実環境における分離実験を行った。実験では, 図 1 に示すように, 通常の室内において 6 つの音源をマイクロホンアレイの周りに 3 次的に配置し, 全ての音源から音声信号を同時に再生し, 観測した混合信号を分離した。評価尺度には SIR (Signal to Interference Ratio) を用いた。各音源ごとの評価結果を表 1 に示す。入力 SIR が非常に低いにもかかわらず, 出力 SIR は高く, SIR 改善量の平均値は 20dB 以上と高い分離性能を示した。

4 まとめ

本稿では多音源ブラインド分離のための試作システムを紹介した。本システムは一般の PC 上に Matlab と C

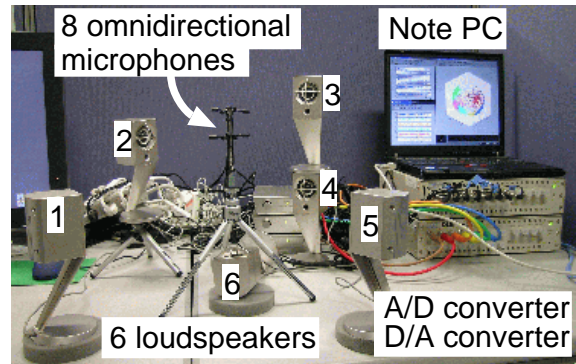


図 1 試作システムの外観と音源・マイクの配置

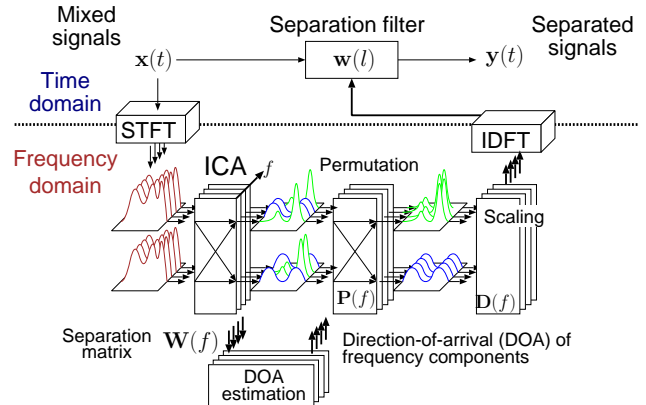


図 2 周波数領域 BSS の流れ

表 1 実験結果 (dB)

	SIR ₁	SIR ₂	SIR ₃	SIR ₄	SIR ₅	SIR ₆	ave.
入力 SIR	-11.6	-9.0	-9.0	-6.6	-6.9	-2.5	-7.6
出力 SIR	7.6	12.2	16.4	14.4	13.6	13.7	13.0

言語によって実装されており, 8ch の A/D コンバータを除いて特殊なハードウェアを必要としない。アルゴリズムの詳細については参考文献を参照願いたい。また, 分離音のサンプルは [4] から聴くことができる。

参考文献

- [1] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Blind source separation and DOA estimation using small 3-D microphone array," in *Proc. HSCMA 2005*, 2005, pp. d.9-10.
- [2] H. Sawada, R. Mukai, S. Araki, and S. Makino, "A robust and precise method for solving the permutation problem of frequency-domain blind source separation," *IEEE Trans. Speech Audio Processing*, vol. 12, no. 5, pp. 530-538, 2004.
- [3] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Frequency domain blind source separation for many speech signals," in *Proc. ICA2004 (LNCS 3195)*, pp. 461-469. Springer-Verlag, 2004.
- [4] <http://www.kecl.ntt.co.jp/icl/signal/mukai/demo/>