

時間周波数表現の冗長度による独立成分分析の信号分離性能の変化*

☆樋口 裕季, 山田 宏樹, 矢田部 浩平 (農工大), 北村 大地 (香川高専)

1 まえがき

ブラインド音源分離 (BSS) は, 音源情報やマイクロホン配置が未知な状況下で, 複数の音源信号が混ざった観測信号を各音源信号に分離する問題である。BSS を解決する手法の一つに, 周波数領域での独立成分分析 (ICA) がある。時間周波数領域へは短時間フーリエ変換 (STFT) によって変換されるが, STFT には窓の種類・窓長・窓のシフト量・周波数点数などのパラメータが存在する。窓長に関してはインパルス応答長より長い方が良いという目安が広く知られている一方で, その他のパラメータに関してはあまり検証されていない。そこで本稿では, 窓長・窓のシフト量・周波数点数の3種類のパラメータを変化させた際の ICA の分離性能について, STFT の冗長度の観点から検証する。

2 周波数領域独立成分分析

ICA は, 複数の混合信号から元の信号を分離する手法である。ICA は周波数ごとに観測信号に分離フィルタを適用することで分離信号を得る。 J を音源数, X_1, X_2, \dots, X_J を観測信号, Y_1, Y_2, \dots, Y_J を分離信号とすると, 分離信号は時間周波数領域で以下のように得ることができる。

$$Y_i[m, n] = \sum_{j=1}^J W_{j,i}[m] X_j[m, n] \quad (1)$$

ただし, 周波数と時間のインデックスは $m = 0, 1, \dots, M-1, n = 0, 1, \dots, N-1, W_{j,i}[m]$ はチャンネル j から音源 i の周波数 m 番目に対する分離フィルタである。観測信号 $X_j[m, n]$ は, ICA で分離する前に, 時間領域から時間周波数領域に変換された信号である。この変換を行う STFT は以下のように定義される。

$$X_j[m, n] = \sum_{l=0}^{L-1} x_j[l] \overline{w_j[l-am]} e^{-\frac{2\pi i b n l}{L}} \quad (2)$$

ただし, 長さ L の時間領域の観測信号を x , 窓を w , 周波数と時間の間引き量をそれぞれ自然数 $b, a, i = \sqrt{-1}$ とする。

3 STFT の冗長度

STFT の結果である $X[m, n]$ の要素数は $M \times N$ である。時間信号長 L に対するスペクトログラムの点数 $M \times N$ の比を冗長度という。

$$R = \frac{NM}{L} \quad (3)$$

スペクトログラムの周波数点数または時間フレーム数または両方を上げることで, 冗長度が上がる。典型的な設定として, 周波数点数を窓長と合わせ, $1/2$ または $1/4$ オーバーラップで窓をシフトした, 冗長度 2 または 4 の STFT が用いられることが多い。一方, STFT のパラメータは任意に設定できるので, それ以外の冗長度の変換を ICA に利用することも可能である。

4 性能評価

本稿では, SiSEC の2音源の混合音声のデータセット4種類と, RWCP のインパルス応答2種類 (E2A, JR2) を使用して, 冗長度を変化させたときの ICA の分離性能の変化を検証した。ICA は AuxICA を用い, パーミュテーションソルバは文献 [1] の手法を用いた。サンプリング周波数は 16 kHz , 窓のシフト量は窓長の $1/2, 1/4, 1/8$, 周波数点数は窓長の $1, 2, 4, 8$ 倍, 窓長は $256, 512, 1024, 2048, 4096, 8192$ とした。

冗長度を変化させたときの SDR 改善量の平均を図-1に, パーミュテーションの正答率の平均を図-2, 図-3に示す。ただし, 図-2は素の正答率を, 図-3は振幅で重みづけた正答率を求めた。窓長が 512 以外の場合は冗長度を上げることで分離性能が向上した。図-3を見ると, エネルギーの大きい周波数ビンにおけるパーミュテーションの正答率が向上したことが一因であると考えられる。他にも, 各周波数での ICA の分離性能自体

*Changes in signal separation performance of independent component analysis due to the redundancy of time-frequency representation By Yuki HIGUCHI, Koki YAMADA, Kohei YATABE (Tokyo University of Agriculture and Technology) and Daichi KITAMURA (National Institute of Technology, Kagawa College)

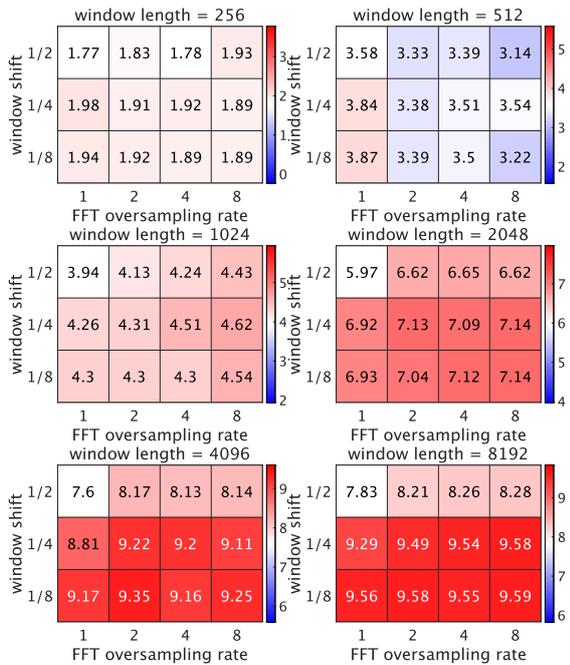


図-1 冗長度を変化させたときの SDR 改善量の平均。冗長度 2 (窓のシフト量 1/2, 周波数点数 1 の窓長に対する比) での結果を基準として, SDR 改善量が 2dB 大きい場合赤色, 2dB 小さい場合青色で着色した。

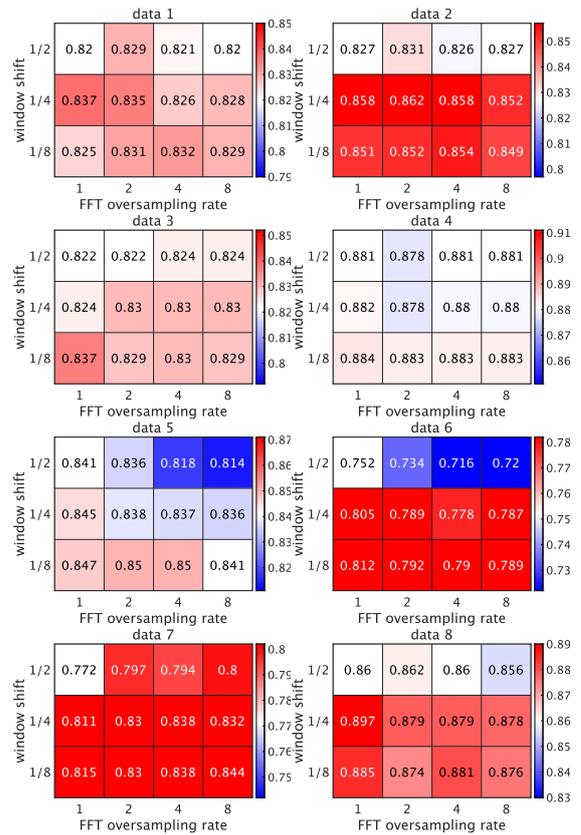


図-3 窓長を 2048 としたときの振幅で重みづけたパーミュテーション正答率。冗長度 2 (窓のシフト量 1/2, 周波数点数の窓長に対する比 1) での結果を基準として, 正答率が 3% 高い場合赤色, 3% 低い場合青色で着色した。

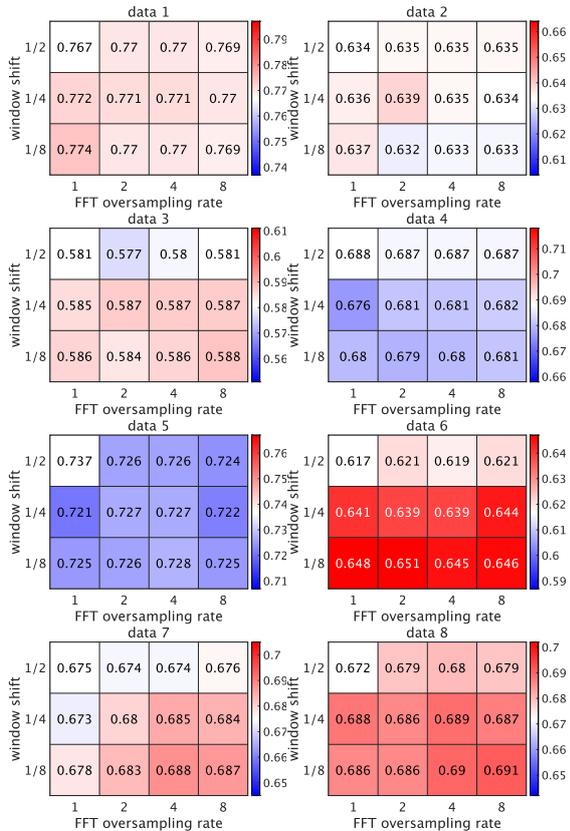


図-2 窓長を 2048 としたときのパーミュテーションの正答率。冗長度 2 (窓のシフト量 1/2, 周波数点数の窓長に対する比 1) での結果を基準として, 正答率が 3% 高い場合赤色, 3% 低い場合青色で着色した。

の向上や, 冗長表現による頑健性の向上などの要因も考えられるので, それらについて今後検証を行う必要がある。

5 むすび

本研究では時間周波数表現の冗長度を変化させることで ICA の分離性能がどのように変化するかを検証した。SDR やパーミュテーション正答率による検証の結果, 冗長度を上げることでインパルス応答長に対して窓長が十分に長いときは SDR 改善量が向上することが分かった。分離性能の向上は, パーミュテーション解決の改善や ICA の分離性能自体の向上, 冗長表現によるパーミュテーションのミスに対する頑健性の向上などによるものと考えている。今後は, さらなる検証を通じて ICA の分離性能が向上した原因を明らかにする。

参考文献

[1] H. Sawada, S. Araki and S. Makino, "Measuring Dependence of Bin-wise Separated Signals for Permutation Alignment in Frequency-domain BSS," IEEE Int. Symp. Circuits Syst., pp. 3247-3250, 2007.