

多音源に対する周波数領域ブラインド音源分離

Blind source separation of many sounds in the frequency domain

澤田 宏, 向井 良, 荒木 章子, 牧野 昭二

Hiroshi Sawada, Ryo Mukai, Shoko Araki and Shoji Makino
日本電信電話 (株) NTT コミュニケーション科学基礎研究所
NTT Communication Science Laboratories, NTT Corporation
{sawada, ryo, shoko, maki}@cslab.kecl.ntt.co.jp

Abstract

This paper describes the frequency-domain approach to the blind source separation (BSS) of convolutively mixed acoustic signals. The advantage of the frequency-domain approach is that convolutive mixtures in the time domain can be approximated as multiple simple mixtures in the frequency domain. However the permutation ambiguity should be solved to group the frequency components of the same source together. This paper presents effective methods to align the permutation ambiguity. Based on the methods, we succeeded in separating many sources in real-world situations.

1 はじめに

複数の音が混ざり合った複数マイクでの観測信号から目的の音を取り出す音源分離技術には、雑音下での音声認識など、様々な応用が期待できる。もし、目的音源の方向などを事前に知っていれば、ビームフォーミング [1] により分離はある程度達成できる。しかし、それらの事前情報が得られない、あるいは得られたとしても正確でない場合には、事前情報を必要としない、いわゆるブラインド音源分離 (BSS: Blind Source Separation [2]) の技術が重要となる。

独立成分分析 (ICA: Independent Component Analysis [3]) は、BSS にとって主要な統計的処理の一つであり、源信号の非ガウス性と独立性に着目して分離を達成する。音声など有益な音信号は、多くの場合非ガウス性を持ち、ICA が効率良く適用できる。ただし、実環境で音が混ざる場合は、単なる混合ではなく、時間遅れと残響を伴った畳み込み混合となるため、いかにして ICA を適用するかが問題となる。

第一の方法は、時間領域において、畳み込み混合を直接 ICA で解くものである [4–6]。これは、正しい収束点では精度良く分離を達成できるが、複雑な畳み込み混合を扱うため、収束までの計算時間が大きいという困難さがある。第二の方法は、観測信号を短時間フーリエ変換し、周波数領域で ICA を適用するものである [7–10]。周波数領域では、畳み込み混合が周波数ビン毎の単純混合に近似できるため、ICA 自体の収束は速い。しかし、ばらばらに分離された各音源の周波数成分を音源毎にグループ化するという、いわゆる permutation の問題を解かなければならない。

我々は、第二の周波数領域での方法に関して精力的に研究を行っており [11–16]、特に、permutation の問題に対する効率的な手法を開発した。その結果、実環境において、2 音源 2 マイクの基本形を初めとし、3 次元的に配置された 6 音源の分離や、無数の背景雑音の中での主要 3 音源の分離を達成した。次章以後、その技術を説明し、実験結果を示す。

2 周波数領域ブラインド音源分離の流れ

N 個の音源 $s_1(t), \dots, s_N(t)$ が空間で畳み込み混合され、 M 個のマイク $x_1(t), \dots, x_M(t)$ で観測されるとする。

$$x_j(t) = \sum_{k=1}^N \sum_l h_{jk}(l) s_k(t-l) \quad (1)$$

ここで、 h_{jk} は音源 k からマイク j までのインパルス応答である。従来の BSS 研究では、音源数 N を知っており、なおかつそれがマイク数 M 以下であるという状況に限る場合が多い。本稿で提案する手法は、そのような状況では当然有利ではあるが、それ以外の状況でも動作するように設計されている。ただし、例えば音源数 N がマイク数 M より大きい場合、すべての音源を分離する [16] ことはせず、主要な音源に対応するものから順に M 個の分離信号を出力する。

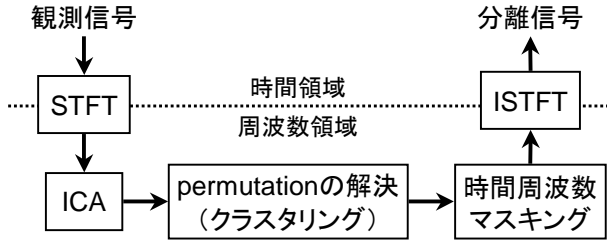


Figure 1: 本稿で説明する周波数領域 BSS の流れ

本章では、以下、周波数領域ブラインド音源分離の流れを説明する。まず、マイクでの観測信号に対してフレーム長 L の短時間フーリエ変換 (STFT: Short-Time Fourier Transform) を適用する。

$$x_j(f, \tau) = \sum_{r=-L/2}^{L/2-1} x_j(\tau + r) \text{win}(r) e^{-j2\pi fr} \quad (2)$$

ここで、 $f \in \{0, \frac{1}{L}f_s, \dots, \frac{L-1}{L}f_s\}$ は周波数、 $\text{win}(r)$ はハニング窓 $\frac{1}{2}(1 + \cos \frac{2\pi r}{L})$ など両端が 0 に収束する長さ L の窓、 τ は時間を表現する新たな変数である。STFT により、式 (1) の畳み込み混合は、各周波数 f での単純混合

$$x_j(f, \tau) \approx \sum_{k=1}^N h_{jk}(f) s_k(f, \tau) \quad (3)$$

に近似される。ここで、 $h_{jk}(f)$ は音源 k からマイク j までの周波数応答、 $s_k(f, \tau)$ は式 (2) を用いて同様に得られる音源の時間周波数表現である。ベクトル表記 $\mathbf{x} = [x_1, \dots, x_M]^T$ 、 $\mathbf{h}_k = [h_{1k}, \dots, h_{Mk}]^T$ により要素をまとめると、式 (3) のベクトル表現

$$\mathbf{x}(f, \tau) \approx \sum_{k=1}^N \mathbf{h}_k(f) s_k(f, \tau) \quad (4)$$

が得られる。周波数領域で解釈すると、ブラインド音源分離の目的は、観測信号ベクトル $\mathbf{x}(f, \tau)$ から、音源 k 毎にすべての周波数成分 $\mathbf{h}_k(f) s_k(f, \tau)$ を求めることにある。

そのためにまず、周波数毎に独立成分分析 (ICA)

$$\mathbf{y}(f, \tau) = \mathbf{W}(f) \mathbf{x}(f, \tau), \quad (5)$$

を適用する。ここで、 $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_N]^H$ 、 $\mathbf{w}_i = [w_{1i}, \dots, w_{Mi}]^T$ は $N \times M$ の分離行列、 $\mathbf{y} = [y_1, \dots, y_N]^T$ は分離信号の要素をまとめたベクトルである。もし音源数 N を知っているかつ $N < M$ であれば、主成分分析により次元削減を行うことは有効であるが、それ以外の場合は $N = M$ 、すなわち正方行列として \mathbf{W} を求める。なお、複素数に対する ICA のアルゴリズム、および、そこで用いられる非線形関数に関する議論は、[11] を参照されたい。

ICA により分離信号 y_i の非ガウス性が高められるため、音源が非ガウス性であり、互いに独立であれば、周波数毎

に分離が達成される。しかし、ICA の解には permutation の任意性 (分離信号の順序を入れ換えても ICA の解となる) と scaling の任意性 (分離信号を定数倍しても独立性は保たれる) が存在するため、これらを解決する必要がある。permutation の解決については次章で詳しく説明する。scaling は、次章の (7) 式により基底ベクトルを求めた後、ある着目したマイク J に対応する要素 $a_{Ji}(f)$ を用いて、

$$y_i(f, \tau) \leftarrow a_{Ji}(f) y_i(f, \tau), \quad (6)$$

として解決できる。これは、マイク J での観測信号に scaling を合わせることであり、minimal distortion principle [4] や projection back [8] と呼ばれるものと等価である。

次に、図 1 の流れに示す時間周波数マスキングを適用する。式 (5) に示す ICA による分離は、線形フィルタによる分離のため、マイク数 M が音源数 N 以上でない場合には、干渉音の残留成分が分離信号にどうしても残る。時間周波数マスキングは、この残留成分を減らす効果がある。その方法の詳細は [15] を参照されたい。本稿では、4.2 章でその効果のみを示す。

以上の処理を経た後、ISTFT: Inverse STFT

$$y_i(\tau + r) = \frac{1}{L \cdot \text{win}(r)} \sum_{f \in \{0, \frac{1}{L}f_s, \dots, \frac{L-1}{L}f_s\}} y_i(f, \tau) e^{j2\pi fr}$$

により、分離信号の周波数成分を集めて時間領域の信号に戻す。

3 permutation の解決

3.1 基底ベクトル

ICA の解の任意性を解決するにあたり、分離行列 \mathbf{W} の逆行列 ($N < M$ の場合は疑似逆行列) を計算することが非常に有益である。以下、 $N = M$ 、すなわち \mathbf{W} が正方行列である場合を説明する。本稿では、分離行列 \mathbf{W} の逆行列により得られるベクトル \mathbf{a}_i を基底ベクトル (basis vector)

$$[\mathbf{a}_1, \dots, \mathbf{a}_M] = \mathbf{W}^{-1}, \quad \mathbf{a}_i = [a_{1i}, \dots, a_{Mi}]^T. \quad (7)$$

と呼ぶ。これは、この逆行列を式 (5) の両辺に掛け合わせることで、観測信号ベクトルがその線形和

$$\mathbf{x}(f, \tau) = \sum_{i=1}^M \mathbf{a}_i(f) y_i(f, \tau) \quad (8)$$

で表現されることによる。この式は周波数領域 BSS において非常に重要である。ICA の解が良好に得られていれば、式 (8) のある i に関する項 $\mathbf{a}_i y_i$ が式 (4) のある k に関する項 $\mathbf{h}_k s_k$ に対応するからである。その対応関係を求めることが permutation を解くことに相当する。その際に使えるのは、基底ベクトル \mathbf{a}_i と分離信号 y_i の情報である。 \mathbf{W} も使えるが、 \mathbf{a}_i と逆行列の関係にあるため、基本的には同じ情報になる。 \mathbf{a}_i と y_i の双方を使うことでより精度

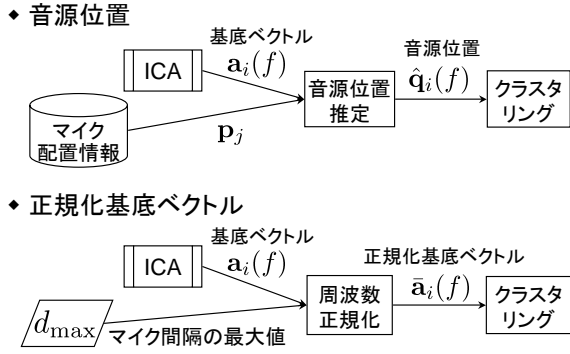


Figure 2: permutation 解決のための 2 種類の方法

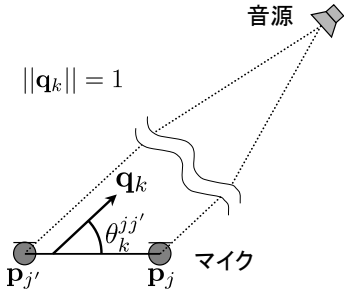


Figure 3: 遠距離場モデル

良く permutation を解決できる [12]. 以下では, 基底ベクトル \mathbf{a}_i の情報をどのように用いるかを説明する. 図 2 に示すように 2 種類の方法があるが, 以下の章でそれぞれ説明する.

3.2 推定音源方向のクラスタリング

一つ目は, 基底ベクトルから音源位置を推定し, その値をクラスタリングすることで permutation を解決する方法である. 本章では, 話を簡単化するため, 遠距離場モデル

$$h_{jk}(f) \approx \exp [j2\pi f c^{-1} \mathbf{p}_j^T \mathbf{q}_k] \quad (9)$$

により音源 k からマイク j への周波数応答 $h_{jk}(f)$ を近似して音源方向 (DOA: Direction-Of-Arrival) を推定する. ここで, c は音の速度, \mathbf{p}_j はマイク j の位置を示す 3 次元ベクトル, \mathbf{q}_k は音源 k の方向を示す長さ 1 ($\|\mathbf{q}_k\| = 1$) の 3 次元ベクトルである. 図 3 に示すように, 二つのマイク j と j' を考えると,

$$\frac{h_{jk}(f)}{h_{j'k}(f)} \approx \exp [j2\pi f c^{-1} (\mathbf{p}_j - \mathbf{p}_{j'})^T \mathbf{q}_k] \quad (10)$$

$$= \exp [j2\pi f c^{-1} \|\mathbf{p}_j - \mathbf{p}_{j'}\| \cos \theta_k^{jj'}] \quad (11)$$

が得られる. このように, 音源方向は, 座標系に対して決まるもの \mathbf{q}_k と, マイクペア (j, j') に対して相対的に決まる角度 $\theta_k^{jj'}$ の 2 種類の表現がある. なお, より一般的な近距離場モデルを仮定して, 方向だけでなく距離も含めた位置を推定する方法に関しては [13, 14] を参照されたい.

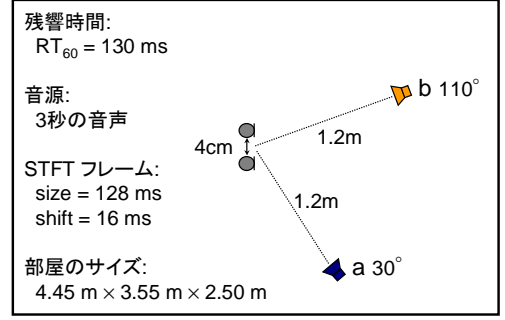


Figure 4: 実験条件 (音源方向推定)

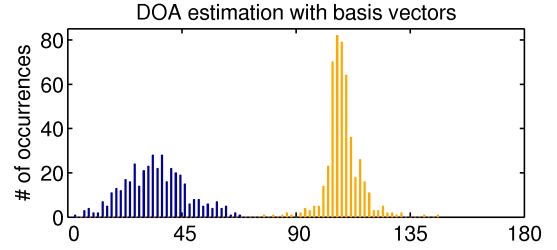


Figure 5: 音源方向推定結果

ICA の解が良好に得られて式 (4) と式 (8) の項に対応関係があると仮定すると, 基底ベクトル $\mathbf{a}_i(f)$ の要素 $a_{ji}(f)$ と $a_{j'i}(f)$ の比は以下のように表現できる.

$$\frac{a_{ji}(f)}{a_{j'i}(f)} = \frac{a_{ji}y_i}{a_{j'i}y_i} \approx \frac{h_{jk}s_k}{h_{j'k}s_k} = \frac{h_{jk}(f)}{h_{j'k}(f)} \quad (12)$$

ここで, 添字 i と k が異なる可能性があることが, permutation の任意性を示している. 式 (12) と式 (11) の偏角に着目すると, 角度の推定値

$$\hat{\theta}_i^{jj'}(f) = \arccos \frac{\arg[a_{ji}(f)/a_{j'i}(f)]}{2\pi f c^{-1} \|\mathbf{p}_j - \mathbf{p}_{j'}\|} \quad (13)$$

を求めることができる.

図 4 に示す実験条件で周波数領域 BSS を実行し, 式 (13) により音源の角度を推定した結果を図 4 に示す. STFT のフレーム長は $L = 1024$ であったため, その約半分の周波数ビンでの推定をヒストグラムで示している. 二つのクラスタが存在することがわかり, そのクラスタリング結果に基づいて permutation を解決することができる.

次に, 座標系に対して決まる方向 \mathbf{q}_k を考える. 式 (12) と式 (10) の偏角に着目すると,

$$2\pi f c^{-1} (\mathbf{p}_j - \mathbf{p}_{j'})^T \mathbf{q}_k \approx \arg[a_{ji}(f)/a_{j'i}(f)] \quad (14)$$

が得られる. u 個のマイクペア $(j_1, j'_1), \dots, (j_u, j'_u)$ に対し同様に考えると, 連立方程式

$$2\pi f c^{-1} \mathbf{D} \mathbf{q}_k = \mathbf{r}_i(f) \quad (15)$$

が得られる. ここで,

$$\mathbf{D} = [\mathbf{p}_{j_1} - \mathbf{p}_{j'_1}, \dots, \mathbf{p}_{j_u} - \mathbf{p}_{j'_u}]^T, \\ \mathbf{r}_i(f) = [\arg(a_{j_1 i}/a_{j'_1 i}), \dots, \arg(a_{j_u i}/a_{j'_u i})]^T.$$

である．実際の環境では推定誤差等の影響により，この連立方程式を厳密に満たす解 \mathbf{q}_k は存在しにくい．そのため，Moore-Penrose 疑似逆行列 \mathbf{D}^+ を用いてその推定値

$$\hat{\mathbf{q}}_i(f) = \frac{\mathbf{D}^+ \mathbf{r}_i(f)}{2\pi f c^{-1}}, \quad \hat{q}_i(f) \leftarrow \frac{\hat{\mathbf{q}}_i(f)}{\|\hat{\mathbf{q}}_i(f)\|} \quad (16)$$

を近似的に得る．4.1 章では，3次元配置の音源とマイクに対して周波数領域 BSS を実行し，式 (16) に従って音源方向を推定してクラスタリングした結果を紹介する．

3.3 正規化基底ベクトルのクラスタリング

次に，基底ベクトル $\mathbf{a}_i(f)$ の周波数依存性をできるだけ除去した正規化基底ベクトル $\bar{\mathbf{a}}_i(f)$ をクラスタリングすることで permutation を解決する方法を示す．これは，上記に示した音源方向による方法ほど直観的ではないが，マイクの配置情報が不要であるという利点がある．まず手順を説明する．正規化基底ベクトル $\bar{\mathbf{a}}_i(f) = [\bar{a}_{1i}(f), \dots, \bar{a}_{Mi}(f)]^T$ の要素 $\bar{a}_{ji}(f)$ は，

$$\bar{a}_{ji}(f) \leftarrow |a_{ji}(f)| \exp \left[j \frac{\arg[a_{ji}(f)/a_{Ji}(f)]}{4fc^{-1}d_{\max}} \right] \quad (17)$$

により計算される．ここで， J はある基準マイクの添字， d_{\max} はある正の実数値であり，基準マイク J と他のマイクとの距離の最大値とすれば良い．この式により，周波数依存性が除去される．次に，scaling の任意性を除去するために，長さを 1 に正規化する．

$$\bar{\mathbf{a}}_i(f) \leftarrow \bar{\mathbf{a}}_i(f) / \|\bar{\mathbf{a}}_i(f)\| \quad (18)$$

これらの操作により，正規化基底ベクトル $\bar{\mathbf{a}}_i(f)$ は，式 (9) の遠距離場モデルに従うと，周波数に依存せず音源の方向 \mathbf{q}_k およびマイクの位置 $\mathbf{p}_1, \dots, \mathbf{p}_M$ のみに依存する．実際，式 (14) の関係を利用すると，

$$\bar{a}_{ji}(f) \approx \frac{1}{\sqrt{M}} \exp \left[j \frac{\pi (\mathbf{p}_j - \mathbf{p}_J)^T \mathbf{q}_k}{2 d_{\max}} \right],$$

であることがわかる．なお，より一般化された近距離場モデルに従った場合でも，周波数 f に依存しないことが証明できる [15]．

次に，正規化基底ベクトル $\bar{\mathbf{a}}_i$ のクラスタリングを行い，クラスタ C_1, \dots, C_M を求める．クラスタ C_k のセントロイド \mathbf{c}_k は， $|C_k|$ をクラスタ C_k のメンバ数として，

$$\mathbf{c}_k \leftarrow \sum_{\bar{\mathbf{a}} \in C_k} \bar{\mathbf{a}} / |C_k|, \quad \mathbf{c}_k \leftarrow \mathbf{c}_k / \|\mathbf{c}_k\|, \quad (19)$$

と計算する．クラスタリングの基準は，クラスタのメンバ $\bar{\mathbf{a}} \in C_k$ とセントロイド \mathbf{c}_k との自乗距離の総和 \mathcal{J} を最小化することである．

$$\mathcal{J} = \sum_{k=1}^M \mathcal{J}_k, \quad \mathcal{J}_k = \sum_{\bar{\mathbf{a}} \in C_k} \|\bar{\mathbf{a}} - \mathbf{c}_k\|^2. \quad (20)$$



Figure 6: 3次元に配置された6音源と4cm立方体の各頂点に配置された8個のマイク

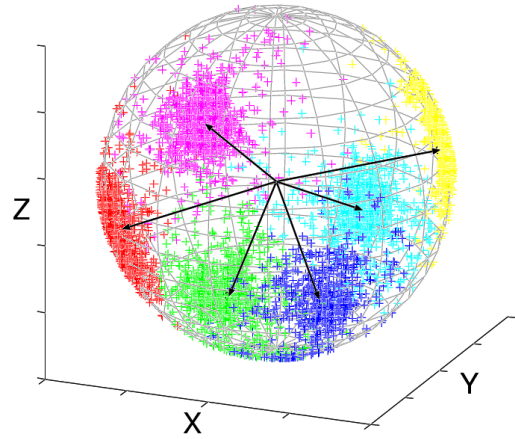


Figure 7: 音源方向推定結果

クラスタリングが終了すれば，各周波数 f で出力の順序を入れ換える順列 Π_f は

$$\Pi_f = \operatorname{argmin}_{\Pi} \sum_{k=1}^M \|\bar{\mathbf{a}}_{\Pi(k)}(f) - \mathbf{c}_k\|^2. \quad (21)$$

として計算できる．4.2 章では，無数の背景雑音の中での主要 3 音源の分離において，正規化基底ベクトルがどのようにクラスタリングされるかを示す．

4 実験

これまでに述べた手法を用いて，実環境で多音源を分離した実験結果を 2 種類示す．

4.1 3次元的に配置された6音源の方向推定と分離

図 6 に示すような 3 次元的配置の 6 音源と 8 マイクを用いて周波数領域 BSS を実行した．音源は，8 秒の英語音声で 6 個用いた．BSS 処理に要した時間は，2GHz の Pentium M を搭載したノート PC で 25 秒程度であった．

式 (16) によりすべての周波数ビンで音源方向を推定した結果を図 7 に示す．音源方向の推定値であるベクトル

Table 1: 6 音源分離: SIR 改善量 (dB)

	SIR ₁	SIR ₂	SIR ₃	SIR ₄	SIR ₅	SIR ₆	平均
入力 SIR	-11.6	-9.0	-9.0	-6.6	-6.9	-2.5	-7.6
SIR 改善量	19.2	21.2	25.4	21.0	20.5	16.2	20.6

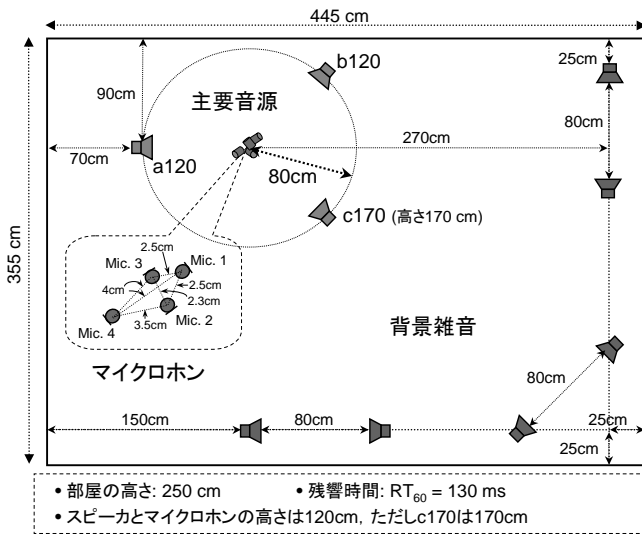


Figure 8: 実験条件 (主要 3 音源の分離)

$\hat{q}_i(f)$ の長さは 1 に正規化されているため、単位球の表面に推定値が乗る。これら推定値に対してクラスタリングを行い、その結果を 6 色で示した。方向の推定値にある程度の分散はあるが、permutation が正しく解ける程度には推定できている。この情報を元に permutation を解き、分離信号を作成した。SIR: Signal-to-Interference Ratio による分離性能の評価を表 1 に示す。SIR の平均改善量は 20dB 程度であり、高精度に分離が達成できたと言える。

4.2 無数の背景雑音の中での主要音源の分離

次に、図 8 に示す状況で実験を行った。ここでは、マイクから遠い背景雑音が 6 個あり、雑踏など無数の背景雑音が存在する状況を模倣している。マイクの近くには 3 個の主要音源があり、これらの音を分離することを目的とした。9 個の音源すべてに音声を用いた。

マイクの数 が 4 個のため、各周波数ビン毎に 4 つの周波数成分を ICA で求めた。permutation の問題は、3.3 章に示した正規化基底ベクトルのクラスタリングで行った。クラスタリング結果を図 9 と図 10 に示す。マイクの数と同じ 4 個のクラスタがある。正規化基底ベクトルは M 次元の複素ベクトルであるため、クラスタの様子を可視化することは難しいが、ここでは、それぞれのクラスタについて、セントロイドとメンバとの自乗距離を示している。

4 つのクラスタのうち、どのクラスタがマイクに近い主要音源に対応し、どのクラスタが背景雑音に対応するかを決定するために、我々はクラスタの分散 $J_k/|C_k|$ に着目

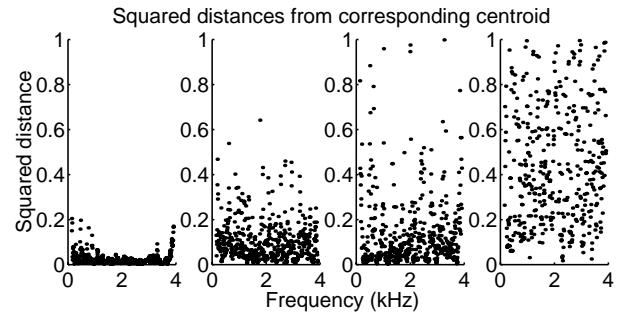


Figure 9: 正規化基底ベクトルのクラスタリング結果 (主要 1 音源の場合)

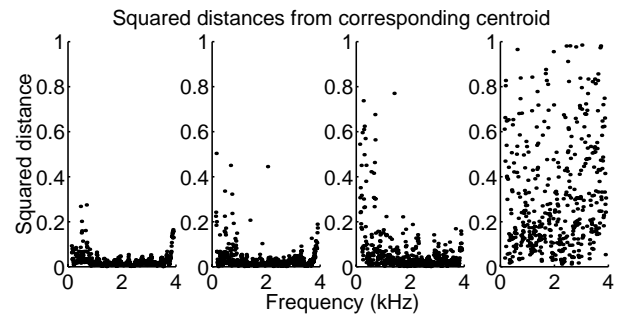


Figure 10: 正規化基底ベクトルのクラスタリング結果 (主要 3 音源の場合)

し、分散が小さいクラスタを主要音源として選んでいる。図 9 は、マイクに近い主要音源のうち 1 つだけを鳴らし、他の二つを鳴らさなかった場合のクラスタリング結果である。左端のクラスタの分散が小さく、主要音源に対応していることがわかる。図 10 は、3 個すべての主要音源を鳴らした場合のクラスタリング結果である。分散が小さいクラスタが 3 個あり、これらが主要音源に対応していることがわかる。このようにクラスタの分散から、主要音源の数を推定することもできる。

次に、時間周波数マスクングの効果を図 11 に示す。これは、マイクに近い主要音源のうち 1 つだけを鳴らした場合の一例である。主要音源 1 個と背景雑音 6 個の合計 7 個の音源があるため、4 個のマイクを用いた ICA による分離では、図 11 の 2) に示すように干渉音の残留成分がどうしても残る。[15] の方法で時間周波数マスクングを適用すると、3) に示すようにその残留成分が抑圧され、4) の目的信号のみのスペクトrogram に近付いた。

最後に、3 個すべての主要音源を鳴らした場合の分離性能を、SIR 改善量で表 2 に示す。音源の組合せを変化させて 10 回試行した平均である。4 個のマイクに対して 9 個の音源 (主要音源 3 個と背景雑音 6 個) という非常に厳しい条件ではあるが、10dB 以上の SIR 改善量が得られた。時間周波数マスクングを併用することで、その改善量は更に高まった。分離音のサンプルは、<http://www.kecl.ntt.co.jp/icl/signal/sawada/demo/dominant/> で聞くことができる。なお、このような処理をリアルタイム

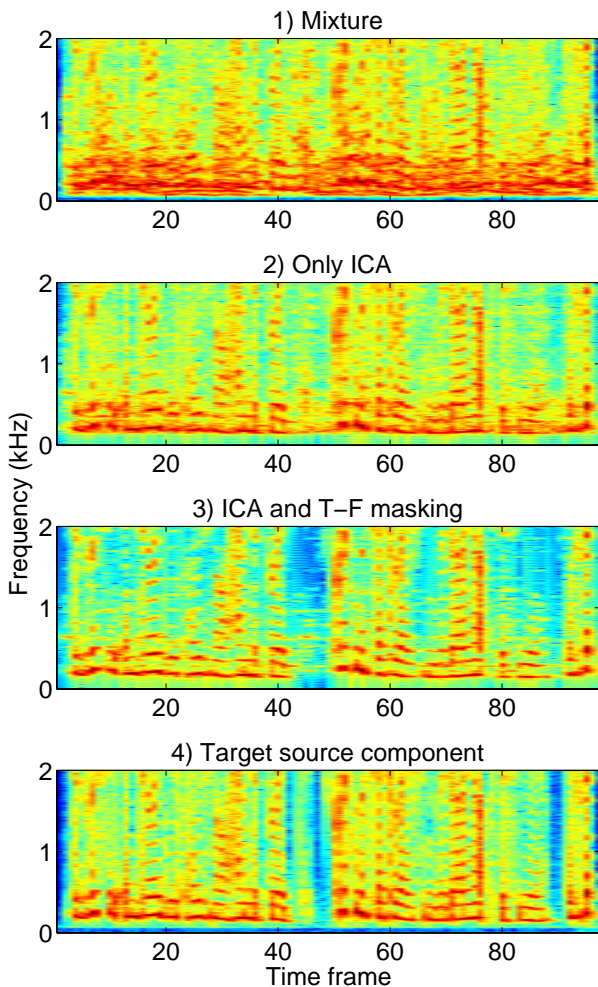


Figure 11: スペクトログラムの例: 1) 混合信号, 2) ICA による分離信号, 3) ICA と時間周波数マスキング (T-F masking) による分離信号, 4) 目的信号のみ (正解)

Table 2: 9 音源中の主要 3 音源分離: SIR 改善量 (dB)

主要音源の位置	a120	b120	c170
入力 SIR	-3.9	-3.6	-5.9
ICA のみ	12.5	13.6	14.5
ICA と時間周波数マスキング	15.1	16.5	17.6

で行うシステムも実現している。

5 おわりに

周波数領域 BSS について説明し、音源数が多い状況での 2 種類の実験でその有効性を示した。ポイントとなる技術は permutation の解決であるが、本稿で説明したように、音源方向や正規化基底ベクトルをクラスタリングすることで効率的に解決できる。なお、マイク数よりも多い音源をすべて分離するものとして、ICA を用いない周波数領域 BSS が知られているが、そのような手法でも、本稿で述べた permutation の解決法と同様の技術が適用できる [16]。

参考文献

- [1] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, pp. 4–24, Apr. 1988.
- [2] S. Haykin, Ed., *Unsupervised Adaptive Filtering (Volume I: Blind Source Separation)*. John Wiley & Sons, 2000.
- [3] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. John Wiley & Sons, 2001.
- [4] K. Matsuoka and S. Nakashima, "Minimal distortion principle for blind source separation," in *Proc. ICA 2001*, Dec. 2001, pp. 722–727.
- [5] S. C. Douglas and X. Sun, "Convolutional blind separation of speech mixtures using the natural gradient," *Speech Communication*, vol. 39, pp. 65–78, 2003.
- [6] S. C. Douglas, H. Sawada, and S. Makino, "A spatio-temporal FastICA algorithm for separating convolutional mixtures," in *Proc. of 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2005)*, vol. V, Mar. 2005, pp. 165–168.
- [7] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomputing*, vol. 22, pp. 21–34, 1998.
- [8] N. Murata, S. Ikeda, and A. Ziehe, "An approach to blind source separation based on temporal structure of speech signals," *Neurocomputing*, vol. 41, no. 1–4, pp. 1–24, Oct. 2001.
- [9] F. Asano, S. Ikeda, M. Ogawa, H. Asoh, and N. Kitawaki, "Combined approach of array processing and independent component analysis for blind separation of acoustic signals," *IEEE Trans. Speech Audio Processing*, vol. 11, no. 3, pp. 204–215, May 2003.
- [10] H. Saruwatari, S. Kurita, K. Takeda, F. Itakura, T. Nishikawa, and K. Shikano, "Blind source separation combining independent component analysis and beamforming," *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 11, pp. 1135–1146, Nov. 2003.
- [11] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Polar coordinate based nonlinear function for frequency domain blind source separation," *IEICE Trans. Fundamentals*, vol. E86-A, no. 3, pp. 590–596, Mar. 2003.
- [12] —, "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, Sept. 2004.
- [13] R. Mukai, H. Sawada, S. Araki, and S. Makino, "Frequency domain blind source separation for many speech signals," in *Proc. of the 5th International Conference on Independent Component Analysis and Blind Signal Separation (ICA 2004 / LNCS 3195)*. Springer-Verlag, Sept. 2004, pp. 461–469.
- [14] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Frequency-domain blind source separation," in *Speech Enhancement*, J. Benesty, S. Makino, and J. Chen, Eds. Springer, Mar. 2005, pp. 299–327.
- [15] H. Sawada, S. Araki, R. Mukai, and S. Makino, "Blind extraction of a dominant source from mixtures of many sources using ICA and time-frequency masking," in *Proc. of 2005 IEEE International Symposium on Circuits and Systems (ISCAS 2005)*, May 2005, pp. 5882–5885.
- [16] S. Araki, H. Sawada, R. Mukai, and S. Makino, "A novel blind source separation method with observation vector clustering," in *Proc. 2005 International Workshop on Acoustic Echo and Noise Control (IWAENC 2005)*, Sept. 2005.