

独立成分分析の基礎と応用

近畿大学大学院	五反田 博 (Hiromu Gotanda)
熊本高等専門学校	石橋 孝昭 (Takaaki Ishibashi)
九州工業大学大学院	岩崎 宣生 (Nobuo Iwasaki)
	井上 勝裕 (Katsuhiko Inoue)

1 まえがき

喧騒とした街中では、車のエンジン音やタイヤ音、通行人の話声、店頭から流れる音楽や宣伝アナウンスなどの様々な音が混じって耳に入る。このように複数の音が入り混じる中、我々は一つ一つの音を聞き分けることができる。例えば、サイレンを鳴らしながら遠くを走る救急車に目をやったり、おしゃべりしながら歩み進む中、「チリン」という背後からの音に思わず振り向いて自転車をよけた経験は誰にでもある。このように特定の声や音を聞き分ける能力（音源分離能力）はカクテルパーティ効果として昔からよく知られている。このような分離能力、つまり、様々な信号源から来た信号が入り混じって観測されたデータから元の信号源を分離して取り出そうと云うのがブラインド信号分離（BSS: Blind Source Separation）である。BSSは、従来の信号処理技術では解決困難な問題に対する新たな方法として、音源分離だけでなく、脳波解析、通信路推定、画像処理、振動解析、時系列予測など多岐の分野で、その発展が期待されている。BSSのアプローチには、独立成分分析法（ICA: Independent Component Analysis）[1, 2, 3]、時間周波数マスクキング法（TFM: Time Frequency Masking）[4, 5]、スパースコーディング[6]、非負値行列分解[7]、ウェーブレット変換に基づく方法[8]などがある。

本稿では、ICAに焦点をあてて解説する。具体的には、まず、観測過程を瞬時混合モデルで定式化して、ICAの基本原則を明らかにする。また、代表的なICAアルゴリズムとして知られた自然勾配法（NG: Natural Gradient）[9]とFastICA法[10]の導出原理を述べて、両者の共通点と相違点を明らかにする。その中でFastICA法の収束がなぜ速いかを述べる。次に、実環境下における実際的な混合モデルとして、時間領域畳込みモデルと周波数領域瞬時混合モデルを紹介し、収束の観点からは後者が有利であることを述べる。さらに、ICA特有の課題として、スケールの不定性と成分置換の問題を述べて、これらの問題が分割スペクトルを導入することで解決できることを示す。

2 独立成分分析法

独立成分分析法（ICA）は、複数の信号源が統計的に独立であることを前提に、信号源からセンサーまでの伝達特性が未知のもとで、センサーでの観測データのみを用いて元の信号源を推定する統計的方法である。ICAは定常信号だけでなく話者音声や音楽などの非定常信号も分離できるという従来技術にない特徴を持っており、信号源の推定だけでなく、観測データの背後に潜む構造や特徴の抽出にも利用されている。

2.1 混合モデルと分離モデル

統計的に独立な $N (\geq 2)$ 個の信号源 $\mathbf{s}(t)=[s_1(t), s_2(t), \dots, s_N(t)]^T$ から出た信号が N 個¹のセンサーで

$$\mathbf{x}(t) = A\mathbf{s}(t) \quad (1)$$

と観測される場合（混合モデル）を考える．ここに， $\mathbf{x}(t)=[x_1(t), x_2(t), \dots, x_N(t)]^T$ は観測データ（混合信号）で， T は転置記号を表す．また， A は a_{mn} を要素とする混合行列 ($N \times N$)， a_{mn} は n 番目の信号源から出た信号が m 番目のセンサーに到達するまでの伝達特性を表す未知の定数である．このとき，瞬時混合モデルに対して，分離モデルを

$$\mathbf{u}(t) = W\mathbf{x}(t) \quad (2)$$

と考える．ここに， $\mathbf{u}(t)=[u_1(t), u_2(t), \dots, u_N(t)]^T$ は分離信号， W は w_{nm} を要素とする分離行列 ($N \times N$) である．

以上のもとで，混合信号 $\{\mathbf{x}(t)|t=1, 2, \dots\}$ だけをデータとして使用して，分離信号が統計的に独立となるように分離行列 W を逐次更新しながら，分離信号 $\{u_n(t)|t=1, 2, \dots\}$ を生成しようと云うのがICA アルゴリズムである．

2.2 独立性の尺度としてのKL 情報量

信号源 $\mathbf{s}$ の同時分布を $p(s_1, \dots, s_N) (=p(\mathbf{s}))$ ，周辺分布を $p(s_n)$ と表記する²．このとき， $p(s_1, \dots, s_N) = \prod_{n=1}^N p(s_n)$ が成り立てば，信号源は統計的に独立である，と云う．したがって，分離信号 $\mathbf{u}$ を統計的に独立となるようにするには，

$$p(u_1, \dots, u_N) = \prod_{n=1}^N p(u_n) \quad (3)$$

と分離信号の同時分布が周辺分布の積と等しくなるように，式 (2) の W を更新していけば良いことになる．そこで，式 (3) の両辺の分布の差を調べるため，分布 $p(\mathbf{u})$ と $q(\mathbf{u})$ の差を量る尺度として知られている KL (Kullback-Leibler) 情報量 $\int p(\mathbf{u}) \log \frac{p(\mathbf{u})}{q(\mathbf{u})} d\mathbf{u}$ において， $q(\mathbf{u})$ を周辺分布の積 $\prod_{n=1}^N p(u_n)$ で置き換えて，

$$KL(\mathbf{u}) = \int p(\mathbf{u}) \log \frac{p(\mathbf{u})}{\prod_{n=1}^N p(u_n)} d\mathbf{u} \quad (4)$$

と定義する．KL 情報量は， $KL(\mathbf{u}) \geq 0$ と非負の値をとり，式 (3) が成立するとき $KL(\mathbf{u}) = 0$ となって最小の値をとる．それゆえ，分離信号が統計的に独立であるか否かは，式 (4) の KL 情報量により判断できる．

¹センサーと信号源の数は異なるとして論じることは可能であるが，ここでは混乱を避けるため，両者は同数として議論する．

²以降，信号を時系列とみるとき $\mathbf{s}(t)$ ，確率変数とみるとき $\mathbf{s}$ のように表記する．

2.3 KL 情報量の分離行列 W による表現

上述のことより、統計的に独立な分離信号を生成するには、KL 情報量を W に関して最小化すれば良い。そこで以下では、KL 情報量を W の陽な関数として定式化する。

式 (4) の KL 情報量は、エントロピーを用いて

$$KL(\mathbf{u}) = \sum_{n=1}^N \mathcal{H}(u_n) - \mathcal{H}(\mathbf{u}) \quad (5)$$

と書き改められる。ここに、 $\mathcal{H}(\mathbf{u}) = -\int p(\mathbf{u}) \log p(\mathbf{u}) d\mathbf{u}$ と $\mathcal{H}(u_n) = -\int p(u_n) \log p(u_n) du_n$ はそれぞれ分離信号の同時エントロピーと周辺エントロピーである。また、式 (2) に基づいて分布の変換を行うと、 $p(\mathbf{x}) = p(\mathbf{u})/|W|$ なる関係が得られる。ここに、 $|\cdot|$ は行列式を表す記号である。この関係を同時エントロピーの式に代入すると、 $\mathcal{H}(\mathbf{u}) = \mathcal{H}(\mathbf{x}) + \log |W|$ となつて、最終的に、KL 情報量が W の陽な関数として

$$KL(W) = \sum_{n=1}^N \mathcal{H}(u_n) - \log |W| - \mathcal{H}(\mathbf{x}) \quad (6)$$

と表現されることになる。

2.4 自然勾配アルゴリズム

$KL(W)$ を評価として、勾配法（最急降下法）を適用すれば、分離行列 W の更新式が

$$W \leftarrow W - \eta E[\varphi(\mathbf{u})\mathbf{u}^T - I]W^{-T} \quad (7)$$

と得られる。ここに、 η は探索ステップ幅、 I は単位行列、 $\varphi(\mathbf{u}) = [\varphi(u_1), \varphi(u_2), \dots, \varphi(u_N)]^T$ で、その要素 $\varphi(u_n)$ はスコア関数と呼ばれる非 2 次的関数である³。

ところで、勾配法は、元来、ユークリッド空間（ピタゴラスの定理が成立する空間）における探索法である。しかし、分離行列 W の N^2 個の要素 w_{nm} の張る空間はユークリッド空間ではない。つまり、各要素 w_{nm} のなす軸は互いに直交するユークリッド空間ではなく、曲がった（リーマン）空間である。そこで Amari は [9]、勾配の概念をリーマン空間に拡張し、そこでの W の更新を自然勾配（NG: Natural Gradient）アルゴリズムとして

$$W \leftarrow W - \eta E[\varphi(\mathbf{u})\mathbf{u}^T - I]W \quad (8)$$

と定式化した。したがって、式 (8) による更新によって $KL(\mathbf{u})$ を最小にする解 W を求め、それを式 (2) に代入することで、統計的に独立な分離信号 $\mathbf{u}(t)$ が生成されることになる。

2.5 自然勾配法と同種の ICA 法

分離信号と混合信号の相互情報量を最小化する観点から導かれる ICA 法として、InfoMax 法がある [12]。相互情報量は $\mathcal{I}(\mathbf{u}; \mathbf{x}) = \mathcal{H}(\mathbf{u}) - \mathcal{H}(\mathbf{u}|\mathbf{x})$ と定義される情報量であるが、最終的に $\mathcal{I}(\mathbf{u}; \mathbf{x}) = \sum_{n=1}^N \mathcal{H}(u_n) - \mathcal{H}(\mathbf{u})$ と展開されて式 (5) と等しくなる。その中で同時エン

³スコア関数としては $\varphi(u) \approx -d \log p(u)/du$ が望ましい。 $p(u)$ の最終的な分布（等価的に信号源の分布）がスーパーガウシアンするとき $\varphi(u) = \tanh(u)$ 、サブガウシアンとき $\varphi(u) = u^3$ がよく用いられる。

トロピー $\mathcal{H}(\mathbf{u}) = \mathcal{H}(\mathbf{x}) + \log |W|$ を評価として勾配法により最大化を図ったのが InfoMax 法である。また、勾配法の代わりに自然勾配法を適用するなどして InfoMax 法をさらに発展させたのが Extended-InfoMax 法である [13]。

対数尤度 $\ell(W) = \sum_{n=1}^N \ln\{|W| \prod_{n=1}^N r(\mathbf{w}_n^T \mathbf{x})\}$ を評価関数とする最尤法からも ICA アルゴリズムを導くことができる [14]。ここに、 $\mathbf{w}_n$ は W の n 列目のベクトル、 $r(\cdot)$ は信号源の分布である。ここで、尤度の対象は、式 (4) の KL 情報量で用いた分離信号の分布 $p(\cdot)$ ではなく、信号源の分布 $r(\cdot)$ であることに注意されたい。両者には、 $r(\cdot)$ が固定した分布であるのに対し、 $p(\cdot)$ は W の更新に伴って変化する分布、と云う違いがある。しかし、 W の収束後の両者の分布は本質的に同じと見なせば、 $KL(W) \approx -\ell(W)$ と近似できる。

以上のことから、InfoMax 法や最尤法に基づく ICA アルゴリズムは、自然勾配法による式 (8) と本質的に等価と考えられる。すなわち、等価な評価関数を対象に、基本的には勾配原理に基づいて探索を行っている点で、三者は同じ範疇にあると云える。

3 FastICA 法

KL 情報量の式 (6)、つまり、 $KL(W) = \sum_{n=1}^N \mathcal{H}(u_n) - \log |W| - \mathcal{H}(\mathbf{x})$ の中で、混合信号のエントロピー $\mathcal{H}(\mathbf{x})$ は $KL(W)$ の最小化に寄与しない。また、混合モデルの自由度ゆえに⁴、 $KL(W)$ を最小にする分離行列 W は 1 つでなく多数（厳密には無限個）存在する。したがって、 W を行列式が $|W| = 1$ となる直交行列 ($W^T W = I$) のクラスに限定しても差し支えない。この場合、KL 情報量は

$$KL(W) \approx \sum_{n=1}^N \mathcal{H}(u_n) \quad (9)$$

のように個々の分離信号 u_n の周辺エントロピーの和で近似できる。言い換えると、分離行列を直交行列に絞り込んだ場合、KL 情報量を最小化することは、分離信号の個々のエントロピーを最小化することと等価になる。Hyvärinen は、この考えを発展させて、高速な ICA アルゴリズムとして知られている FastICA 法を導いた [10][11]。以下ではその概要を述べる。

3.1 混合信号の前処理

分離行列 W を上述のように直交行列に絞り込めば、個々の分離信号のエントロピーを最小化することで、統計的に独立な分離信号を生成できる。ただし、 W を直交行列に絞り込むには、あらかじめ混合信号に対して前処理（中心化と白色化）を行う必要がある。中心化とは $\hat{\mathbf{x}} = \mathbf{x} - E[\mathbf{x}]$ のように混合信号の平均（中心）を原点に移動させる処理のことである。また、白色化とは、 $E[\hat{\mathbf{x}}\hat{\mathbf{x}}^T]$ の固有値と固有ベクトルをそれぞれ λ_n と $\mathbf{c}_n$ として定義される $\Lambda = \text{diag}[\lambda_1, \dots, \lambda_n]$ と $\Gamma = [\mathbf{c}_1, \dots, \mathbf{c}_n]$ に基づいて、 $\tilde{\mathbf{x}} = \Lambda^{-1/2} \Gamma^T \hat{\mathbf{x}}$ と変換し、 $E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] = I$ となるように規格化する処理のことである。

このとき、原信号 s_m の平均をゼロ、分散を 1 と仮定して⁵、白色化後の混合信号 $\tilde{\mathbf{x}}$ と原信号 $\mathbf{s}$ の関係を見てみると、

$$\tilde{\mathbf{x}} = \Lambda^{-1/2} \Gamma^T A \mathbf{s} \quad (10)$$

⁴式 (1) は、 N 個の式に対して未知数は $(N^2 + N)$ 個であるから、 N^2 の自由度となる。

⁵後述するように、ICA は分離信号が原信号と同じスケールで得られることは保証していない。それゆえ、原信号の分散を 1 と仮定しても問題ない。

となる。したがって、白色化後の混合行列は、 $\tilde{A} = \Lambda^{-1/2} \Gamma^T A$ と表現され、 $\tilde{A} \tilde{A}^T = I$ を満たすことから、結果的に直交行列となる。このことは、式 (2) の分離モデルを式 (1) の混合モデルの逆変換過程とみれば、分離行列 $W = [\mathbf{w}_1, \dots, \mathbf{w}_N]^T$ が直交行列のクラスに絞り込めることを示唆している。また、 W が直交行列の場合、 $\|\mathbf{w}_n\|^2 = 1$ となって、探索空間は超曲面に絞り込まれるため⁶、探索アルゴリズム (ICA アルゴリズム) の収束は容易になると考えらる。以上のことから、混合信号に対する前処理は、分離行列 W を直交行列に絞り込むための準備、と位置づけられる。

3.2 FastICA アルゴリズム

中心化と白色化の前処理を行えば、式 (9) が成立することから、分離信号の統計的独立性は、個々の分離信号

$$u_n = \mathbf{w}_n^T \tilde{\mathbf{x}} \quad (11)$$

のエントロピー $\mathcal{H}(u_n)$ を最小化することで保証されることになる。しかし、 $\mathbf{w}_n$ を更新する度に、分離信号の分布を推定して $\mathcal{H}(u_n)$ を求めることや、 W が直交行列であることを制約条件に取り込んで最小化を図ることは容易でない。そこで、Hyvärinen は [2]、エントロピー $\mathcal{H}(u_n)$ を最小化する代わりに、

$$J(u_n) = \mathcal{H}(\nu) - \mathcal{H}(u_n) \geq 0 \quad (12)$$

と定義されるネグエントロピーを最大化することで、分離信号を統計的に独立させるアプローチをとった。ここに、 ν は平均が 0 で分散が 1 のガウス分布に従う確率変数である⁷。

ネグエントロピーは、 $J(u_n) \geq 0$ と非負の値をとり、 u_n がガウス分布のとき最小の 0 となって、 u_n の分布がガウス分布から遠ざかるほど大きくなる、と云う性質をもっている。そのため、ネグエントロピーは非ガウス性の尺度として利用できる。この観点から、 $J(u_n)$ を近似して、制約条件 $\|\mathbf{w}_n\|^2 = 1$ を取り込むと、条件付き評価関数が⁸

$$L(\mathbf{w}_n) = \{E[G(u_n)] - E[G(\nu)]\}^2 - \beta \{\|\mathbf{w}_n\|^2 - 1\} \quad (13)$$

のように導かれる。ここに、 $G(\cdot)$ はコントラスト関数と呼ばれる非 2 次的関数で、 β はラグランジェの未定定数である。そして、式 (13) の不動点 ($\partial L(\mathbf{w}_n)/\partial \mathbf{w}_n = 0$ となる点) における関係を求めると、

$$E[\tilde{\mathbf{x}} g(\mathbf{w}_n^T \tilde{\mathbf{x}})] - \beta \mathbf{w}_n = 0 \quad (14)$$

なる恒等式が得られる。ここに、 $g(\cdot)$ は $G(\cdot)$ の導関数で、式 (8) の自然勾配アルゴリズムにおける $\varphi(\cdot)$ に相当するスコア関数である。さらに、この恒等式にニュートン法⁸を適用することで、最終的に、分離荷重 $\mathbf{w}_n$ を

$$\mathbf{w}_n^+ = E[\tilde{\mathbf{x}} g(\mathbf{w}_n^T \tilde{\mathbf{x}})] - E[g'(\mathbf{w}_n^T \tilde{\mathbf{x}})] \mathbf{w}_n \quad (15)$$

$$\mathbf{w}_n = \frac{\mathbf{w}_n^+}{\|\mathbf{w}_n^+\|} \quad (16)$$

⁶例えば、 $\mathbf{w}_n$ が 2 次元ベクトルの場合、探索範囲は 2 次元平面から単位円周上に絞り込まれることになる。

⁷厳密には、 ν は平均と分散が u_n と等しいガウス分布に従う確率変数と定義される。しかし FastICA の場合、中心化と白色化の前処理により、 u_n は平均がゼロで、分散は $E[u_n^2] = E[\mathbf{w}_n^T \tilde{\mathbf{x}} \tilde{\mathbf{x}}^T \mathbf{w}_n] = \mathbf{w}_n^T E[\tilde{\mathbf{x}} \tilde{\mathbf{x}}^T] \mathbf{w}_n = \|\mathbf{w}_n\|^2 = 1$ と規格化されるため、 ν は平均 0 分散 1 のガウス分布として良い。

⁸関数 $f(x) = 0$ の根を $x \leftarrow x - f(x)/f'(x)$ と求める数値計算法

のように更新する FastICA アルゴリズムが導かれる。この更新式 (15)(16) は、

$$|\mathbf{w}_{n,\text{old}}^T \mathbf{w}_{n,\text{new}}| \simeq 1 \quad (17)$$

のように更新前後の $\mathbf{w}_n$ の方向が一致したとき、収束したと判定される。ここに、添え字の old と new はそれぞれ更新の前後を指す。

上述の手順で、最初 ($n = 1$) に得られる分離荷重 $\mathbf{w}_1$ をもとに、 $u_1 = \mathbf{w}_1^T \mathbf{x}$ と生成される分離信号は、 N 個の信号源の中で非ガウス性が最大の信号源 s_m を分離したものとなる⁹。そして、 $n = 2$ として得られる分離荷重 $\mathbf{w}_2$ をもとに生成される分離信号は、非ガウス性が2番目に大きい信号源を分離したものとなる。以下、信号源は非ガウス性の大きいものから順に分離されることになる。

ただし、 $n \geq 2$ の手順においては、分離荷重 $\mathbf{w}_n$ が先に推定された分離荷重 $\mathbf{w}_i (i \leq n-1)$ と等しくなるのを避けるため、グラムシュミットの方法で

$$\mathbf{w}_n = \mathbf{w}_n - \sum_{i=1}^{n-1} \mathbf{w}_i^T \mathbf{w}_n \mathbf{w}_i \quad (18)$$

のように直交化させて、式 (16) で規格化する必要がある。

FastICA アルゴリズムには、以上のように $\mathbf{w}_n$ を1つずつ更新するアルゴリズム (Deflationary FastICA) の他に、 $\mathbf{w}_n (n = 1, 2, \dots, N)$ を同時にまとめた行列 $W = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N]^T$ を一括して更新するアルゴリズム (Symmetric FastICA) もある。その詳細については [2] を参照されたい。

3.3 FastICA 法と他の ICA 法との違い

FastICA 法は、自然勾配法や InfoMax 法、最尤法と同じく KL 情報量の流れから説明できる方法であるが、アルゴリズムの様相は大きく異なり、その収束は文字通り速い。以下では、その違いを順に述べる。

まず、中心化と白色化の前処理が前提となるが、分離行列 $W = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_N]^T$ の探索範囲を直交行列のクラスに絞り込んだことは、他の方法には見られない FastICA 法の大きな特徴である。これにより、式 (4) のように同時分布 $p(u_1, \dots, u_N)$ と周辺分布 $\prod_{n=1}^N p(u_n)$ の2つの分布の比較評価 (KL 情報量) から、式 (9) のように個々の分布 $p(u_n)$ の単体評価 (エントロピー) へと、より簡便な評価法に移行できる。このことが FastICA 法とその他の方法との分岐点になっている。

エントロピーは分布の形状によって異なる値をとり、ガウス分布のとき最大となる。一方、中心極限定理より、平均が0で分散の等しい独立な確率変数の和はガウス分布に漸近する。言い換えると、独立成分 (信号源) の混じり合った混合信号 x_n はガウス分布に漸近する。これを裏返せば、ガウス分布から遠ざかるように (非ガウスの分布になるように) 分離荷重 $\mathbf{w}_n$ を決めることができれば、独立成分 (信号源) が取り戻せる、と云う FastICA の基本原理にたどり着く。非ガウス性の尺度としてはネゲントロピーが知られている。以上のことから、分布の形状を量る評価としては、エントロピーよりも、式 (12) で等価的に定義されるネゲントロピーの方に利がある。

⁹脚注4で指摘したように式 (1) の混合モデルには自由度があるため、ICA 解には後述するスケールの不定性や成分置換の問題がある。この成分置換に起因して、最初に得られた分離信号 u_1 は必ずしも信号源 s_1 を反映した信号とはならない。

また、FastICA 法の収束が速い理由は、 W の直交行列への絞り込みを、ネグエントロピーの条件付き最適化問題として式 (13) のように取り込んで定式化できたことにある¹⁰。この定式化により、分離荷重 w_n の探索範囲を全空間から超球面に絞り込むことが数値計算的に可能になるとともに、勾配に代わって不動点に基づく探索アプローチ（ニュートン法）の適用が可能となる。ニュートン法は、勾配法に比べて収束が速い。以上のことから、FastICA アルゴリズムの高速性が裏付けられる。

さらに、FastICA 法には、収束の速さに加えて、アルゴリズムを実装する上での利点もある。まず、ニュートン法を探索アプローチとすることで、勾配法には必須の探索ステップ幅 η の設定が不要になる。 η を適切な値に設定するには、事前に多数の試行が必要である。次に、式 (14) にニュートン法を適用して式 (15) と式 (16) を導出してみれば分かるように、条件付き最適化問題には必須のラグランジェの未定定数 β の求解は不要である。このことは計算負荷の軽減につながる。さらに、収束判定基準を式 (17) のように定式化できたことは、アルゴリズムを実装する上で極めて大きなメリットと云える。これにより、実用的な収束判定が行えるようになった。

4 実環境下での音源分離

本節では音響信号を対象に議論を進める。そのため、信号源は音源、センサーはマイクと読み替える。この場合、2 節で述べた式 (1) の瞬時混合モデルは、各音源から出た音波が個々のマイクに同時に到達すると云う非現実的なモデルになってしまう。現実の環境では、個々の音源からの直接波に加えて壁や天井からの反射波もマイクに入るため、

$$x_m(t) = \sum_{n=1}^N a_{mn}(t) * s_n(t) = \sum_{n=1}^N \sum_{t'=0}^{T'-1} a_{mn}(t') s_n(t-t') \quad (19)$$

のように観測される。ここに、 $s_n(t)$ ($n = 1, 2, \dots, N$) は音源、 $\{x_m(t)|t = 0, 1, 2, \dots\}$ ($m = 1, 2, \dots, N$) は m 番目のマイクでの観測信号（混合信号）、 $a_{mn}(t')$ は n 番目の音源から m 番目のマイクまでのインパルス応答、 t' は遅れ時間、 T' はインパルス応答長、 $*$ は畳込みを表す。式 (19) は時空間混合モデルあるいは時間領域畳込み混合モデルと呼ばれる。

4.1 時間領域 ICA

時間領域畳込み混合モデルに対して、その分離モデルを

$$u_m(t) = \sum_{n=1}^N w_{nm}(t) * x_m(t) = \sum_{n=1}^N \sum_{t''=0}^{T''-1} w_{nm}(t'') x_m(t-t'') \quad (20)$$

のようにタップ長が T'' の分離フィルター $w_{nm}(t)$ で構成する。このとき、混合信号のデータ $\{x_m(t)|t = 0, 1, 2, \dots\}$ ($m = 1, 2, \dots, M$) から元の音源を式 (20) のように分離・復元する方法を時間領域 ICA（TDICA: Time Domain ICA）と云う。

TDICA は、反射が弱く残響時間 $T_{60}[\text{sec}]$ ¹¹ が小さい場合や、 $T_{60}[\text{sec}]$ が大きくてもマイクから音源までの距離が近い場合、良好に機能する。しかし、マイクが音源から数十セン

¹⁰ 自然勾配法や InfoMax 法、最尤法の場合、中心化やと白色化の前処理を行ったとしても、 W が直交行列に絞り込まれたことにはならないため、収束性が改善されるとは限らない。

¹¹ 音源の発音が止まってから、残響音が 60dB 減衰するまでの時間。

チメートルも離れると、反射の影響を強く受けるため、良好な結果を得るのが難しくなる。これは、例えば、残響時間が $T_{60}=50[\text{msec}]$ と小さい場合でも、8KHz サンプリングのときのインパルス応答長は $T' = 400$ となって¹²、これと同程度のタップ長の分離フィルターを推定しなければならないからである。つまり、逆フィルターのタップ長を単純に $T'' = T'$ と考えたとしても、1個の逆フィルターにつき400個と極めて多数のパラメータ $\{w_{nm}(t'') | t'' = 0, 1, 2, \dots, T'' - 1\}$ を推定しなければならず、残響時間が大きくなるほど、その個数は増えることになって、収束が難しくなる¹³。

4.2 周波数領域ICA

上述のことから、実環境下では、式(19)を $x_m(\omega_l, k) = \sum_{n=1}^N a_{mn}(\omega_l) s_n(\omega_l, k)$ のように短時間フーリエ変換して得られる周波数領域混合モデル

$$\mathbf{x}(\omega_l, k) = A(\omega_l) \mathbf{s}(\omega_l, k) \quad (21)$$

を考えるのが一般的である。ここに、 ω_l と k はそれぞれ後述する l 番目の規格化周波数とフレーム時刻、 $A(\omega_l)$ は $a_{mn}(t')$ のフーリエ変換 $a_{mn}(\omega_l)$ を要素とする未知の周波数伝達関数行列、 $\mathbf{s}_n(\omega_l, k)$ は未知の音源の短時間スペクトルである。また、 $\mathbf{x}_m(\omega_l, k)$ は、次式で実際に計算して得られる混合信号の短時間スペクトル（以下、混合スペクトルと呼ぶ。）である。すなわち、混合信号 $\{x_m(t) | t = 0, 1, 2, \dots\}$ を L 個ずつ切り出して、 $\{x_m(l + kR) | l = 0, 1, 2, \dots, L-1\}$ と得られる k フレーム目のデータを窓がけして、短時間フーリエ変換により、

$$x_m(\omega_l, k) = \sum_{l=0}^{L-1} x_m(l + kR) h(l) e^{-j\frac{2\pi}{L}lk} \quad l = 0, 1, 2, \dots, L-1 \quad (22)$$

と計算して混合スペクトル $x_m(\omega_l, k)$ を求める。ここに、 $\omega_l = 2\pi l/L$ ($l = 0, 1, 2, \dots, L-1$) は規格化周波数、 R はフレーム周期（シフト幅）、 $h(l)$ は窓関数、 $j = \sqrt{-1}$ である。

また、式(21)の周波数領域混合モデルに対し、式(20)を $u_n(\omega_l, k) = \sum_{m=1}^N w_{nm}(\omega_l) x_m(\omega_l, k)$ と短時間フーリエ変換して得られる式

$$\mathbf{u}(\omega_l, k) = W(\omega_l) \mathbf{x}(\omega_l, k) \quad (23)$$

を周波数領域分離モデルと定義する。ここに、 $\mathbf{u}(\omega_l, k)$ は分離信号の短時間スペクトル（以下、分離スペクトルと呼ぶ。）、 $W(\omega_l)$ は分離行列である。このとき、個々の周波数 ω_l において、混合スペクトル $\{\mathbf{x}_m(\omega_l, k) | k = 0, 1, 2, \dots, K\}$ から分離行列 $W(\omega_l)$ を推定して、分離スペクトル $\mathbf{u}(\omega_l, k)$ を求め、それを逆短時間フーリエ変換することにより、元の音源に対応する分離信号 $u_n(t)$ ($n = 1, 2, \dots, N$) を生成しよう、と云うのが周波数領域ICA (FDICA: Frequency Domain ICA) である。

FDICA の場合、データとして用いる混合スペクトル $\mathbf{x}_m(\omega_l, k) | k = 0, 1, 2, \dots, K\}$ や推定すべき分離行列 $W(\omega_l)$ の要素は複素数であるから、2節で述べた実数版のICAアルゴリズムは複素数版に変更する必要がある。

¹²インパルス応答長 T' は残響時間 $T_{60}[\text{sec}]$ とサンプリング周波数 $f_s[\text{Hz}]$ の積で $T' \approx T_{60}f_s$ と近似できる [15]。

¹³インパルス応答は必ずしも最小位相推移とは限らないので、安定な逆フィルタが得られるかと云うことも問題になる。

まず、周波数領域自然勾配 (NG) アルゴリズムの場合、式 (8) の複素数版は

$$W(\omega_l) \leftarrow W(\omega_l) - \eta E[\varphi(\mathbf{u}(\omega_l, k)) \mathbf{u}(\omega_l, k)^H - \text{diag}(E[\varphi(\mathbf{u}(\omega_l, k)) \mathbf{u}(\omega_l, k)^H]) W(\omega_l)] \quad (24)$$

と与えられる [16]. ここに、 $\varphi(\mathbf{u}(\omega_l, k)) = \varphi(\Re(\mathbf{u}(\omega_l, k)) + j\Im(\mathbf{u}(\omega_l, k)))$ で、 $\Re$ と $\Im$ はそれぞれ実数部と虚数部を表し、 H はエルミート転置記号である.

また、周波数領域 FastICA アルゴリズムの場合、式 (15)(16) の複素数版は

$$\begin{aligned} \mathbf{w}_n^+(\omega_l, k) \leftarrow & E[\tilde{\mathbf{x}}(\omega_l, k) \tilde{u}_n(\omega_l, k) g(|u_n(\omega_l, k)|^2)] \\ & - E[g(|u_n(\omega_l, k)|^2) + |u_n(\omega_l, k)|^2 g'(|u_n(\omega_l, k)|^2)] \mathbf{w}_n(\omega_l, k) \end{aligned} \quad (25)$$

$$\mathbf{w}_n(\omega_l, k) \leftarrow \frac{\mathbf{w}_n^+(\omega_l, k)}{\|\mathbf{w}_n^+(\omega_l, k)\|} \quad (26)$$

と与えられる [17]. ここに、 $u_n(\omega_l, k) = \mathbf{w}_n^H(\omega_l, k) \tilde{\mathbf{x}}(\omega_l, k)$ で、 $\tilde{\mathbf{x}}(\omega_l, k)$ は各周波数ビン ω_l で $\{\mathbf{x}(\omega_l, k) \mid k = 1, 2, \dots, K\}$ に対して中心化と白色化の前処理を行った後の混合スペクトルである. また、“ $-$ ” は複素共役を表す.

式 (24) の $\varphi(\cdot)$ と式 (25) の $g(\cdot)$ および $g'(\cdot)$ の比較から分かるように、スコア関数は周波数領域 NG アルゴリズムでは複素数値をとるのに対し、周波数領域 FastICA アルゴリズムでは実数値をとる. この差はアルゴリズムを実装する際、メモリー使用量や演算速度に影響する. CPU パワーやメモリー容量の制限される携帯機器に実装する場合、実数は複素数よりメモリーや演算回数が少なくて済むので有利である.

4.3 スケールの不定性と成分置換問題

周波数領域 NG や周波数領域 FastICA などの周波数領域 ICA (FDICA) アルゴリズムで分離行列 $W(\omega_l)$ を推定した場合、式 (21) に自由度があるため、

$$W(\omega_l) A(\omega_l) = P(\omega_l) D(\omega_l) \quad (27)$$

のようにスケールの不定性と成分置換の問題が残る. ここに、 $D(\omega_l)$ は対角行列、 $P(\omega_l)$ は置換行列¹⁴である. 言い換えると、 $W(\omega_l)$ の推定値を式 (23) に代入して $\mathbf{u}(\omega_l, k) = W(\omega_l) \mathbf{x}(\omega_l, k)$ と生成される分離スペクトルは、必ずしも音源のスペクトルと等しくならない. つまり、必ずしも $u_n(\omega_l, k) = s_n(\omega_l, k)$ とはならず、 n 番目の分離スペクトルは $u_n(\omega_l, k) = d_i(\omega_l) s_i(\omega_l, k)$ のように n 番目の音源ではなく $i (\neq n)$ 番目の音源を $d_i(\omega_l)$ 倍した値となる. このように分離スペクトルの順番 n と音源のスペクトル順番 i が一致しないことを成分置換という. また、スケール $d_i(\omega_l)$ は周波数ビン l 毎に異なる. これをスケールの不定性という.

したがって、スケールの不定性と成分置換が解消されなければ、式 (23) から得られる分離スペクトル $\{\mathbf{u}(\omega_l, k) \mid l = 0, 1, \dots, L-1, k = 1, 2, \dots, K\}$ を逆短時間フーリエ変換して時間領域に戻しても、 $u_n(t)$ の期待する $s_n(t)$ は復元できない.

4.3.1 分割スペクトル

ここでは、スケールの不定性と成分置換の問題の本質を明らかにするため、周波数ビン ω_l やフレーム番号 k は外して議論する. また、分離スペクトルは成分が置換されている

¹⁴各行と各列において、1 の値をとる一個の要素を除いて、その他の要素はすべてゼロとなる行列.

ことを明示するため、分離スペクトル $\mathbf{u}$ の成分は $u_{\tilde{n}}$ のように成分番号を示す添字の n に“ $\sim$ ”を付けて、 $\mathbf{u} = [u_{\tilde{1}}, u_{\tilde{2}}, \dots, u_{\tilde{N}}]^T$ と表記し直す。補足すると、分離スペクトル $\mathbf{u}$ の第 $\tilde{n}$ 番目の成分 $u_{\tilde{n}}$ は N 個の音源 $\{s_1, s_2, \dots, s_N\}$ のどれか1つと排他的に対応するが、どれに対応するか定かでない。つまり、集合 $\{u_{\tilde{1}}, u_{\tilde{2}}, \dots, u_{\tilde{N}}\}$ は集合 $\{s_1, s_2, \dots, s_N\}$ と1対1の関係にあるが、どれがどれに対応するか、具体的な対応は不明である。

以上の準備のもとで、以下では、Murata ら [18] により

$$\xi_{\tilde{n}} = \mathbf{W}^{-1} [0, \dots, 0, u_{\tilde{n}}, 0, \dots, 0]^T \quad (28)$$

と定義されるスペクトル $\xi_{\tilde{n}} = [\xi_{1\tilde{n}}, \xi_{2\tilde{n}}, \dots, \xi_{j\tilde{n}}, \dots, \xi_{N\tilde{n}}]^T$ を活用することで、スケールの不定性や成分置換が解消できることを述べる。便宜のため、このスペクトルを分割スペクトル (Decomposed Spectrum; 元の成分に戻されたスペクトル) と呼ぶ。この分割スペクトルについては、付録で証明するように、第 $\tilde{n}$ 番目の分割スペクトル $\xi_{\tilde{n}}$ の第 m 要素 $\xi_{m\tilde{n}}$ と第 n 番目の音源 s_n との間に次の関係が成り立つ [19][20]。

$$\xi_{m\tilde{n}} = a_{mn}s_n \quad m = 1, 2, \dots, N \quad (29)$$

ここで、 $u_{\tilde{n}}$ は、第 n 番目の音源 s_n が分離スペクトル $\mathbf{u}$ では順番が入れ替わって第 $\tilde{n}$ 番目の成分として算出されたもので、 $\tilde{n}$ は未知であることに注意されたい。つまり、 s_n が $\{u_{\tilde{1}}, u_{\tilde{2}}, \dots, u_{\tilde{N}}\}$ のどれに対応するか判然としないが、とりあえず $u_{\tilde{n}}$ と表記しているにすぎない。したがって、 $u_{\tilde{n}}$ から式 (28) のように誘導される第 $\tilde{n}$ 番目の分割シンボル $\xi_{\tilde{n}}$ についても、それが具体的に何番目の音源に対応しているか分からない。

式 (29) は、たとえ音源の順番 n と分割スペクトルの順番 $\tilde{n}$ の対応が未知でも、分割スペクトル $\xi_{\tilde{n}}$ には以下の性質があることを主張している [20]。

【性質 1】

分割スペクトルの要素 $\xi_{m\tilde{n}}$ は、音源 s_n から各マイク ($m = 1, 2, \dots, N$) への入力分を表している。

【性質 2】

$\xi_{m\tilde{n}}$ と a_{mn} は、第2添字が $\tilde{n}$ と n のように異なるとしても、第1添字は同じ m をとる。つまり、成分置換があつたとしても、 $\xi_{m\tilde{n}}$ の第1添字 m は a_{mn} の第1添字 m の順番をそのまま継承する。

【性質 3】

分離スペクトル $u_{\tilde{n}}$ から分割スペクトル $\xi_{m\tilde{n}}$ ($m = 1, 2, \dots, N$) が生成される際のスケール (a_{mn}) は、音源 s_n からマイク x_m ($m = 1, 2, \dots, N$) までの伝達特性 (a_{mn}) に等しい。

性質 2 と 3 をまとめると、 n 番目の音源から各マイク ($m = 1, 2, \dots, N$) への伝達メカニズムは、 $u_{\tilde{n}}$ から分割スペクトル $\xi_{m\tilde{n}}$ ($m = 1, 2, \dots, N$) を生成するメカニズムに継承される、と云える。言い換えると、例え成分置換が起きていたとしても、分割スペクトルの要素を調べることによって、信号源からセンサーへの (未知の) 伝達メカニズム (混合過程) に関する情報が得られる、と云うことになる。また、このことは ICA アルゴリズムの種類に依存しない。

4.3.2 スケールの不定性と成分置換の解消

スケールの不定性については、分割スペクトルを導入することで以下のように自ずと解決される。式 (29) を周波数ビン ω_l やフレーム番号 k を復活して

$$\xi_{m\hat{n}}(\omega_l, k) = a_{mn}(\omega_l) s_n(\omega_l, k) \quad m = 1, 2, \dots, N \quad (30)$$

と表示すれば分かるように、分割スペクトル $\xi_{m\hat{n}}(\omega_l, k)$ は音源を $a_{mn}(\omega_l)$ 倍した値となる。厳密に云うと、分割スペクトル $\xi_{m\hat{n}}(\omega_l, k)$ は、第 n 番目の音源のみを活性させその他の音源を不活性にした状況で、第 m 番目のマイクで観測される値である。この場合、 $a_{mn}(\omega_l)$ は、本来、音源とマイク間の周波数特性であることから、各周波数 ω_l でのスケール（倍率）は、音場の周波数特性で規定された値となることに注意されたい。以上のことから、分割スペクトル $\xi_{m\hat{n}}(\omega_l, k)$ にはスケールの不定性はないと結論づけられる [19][21]。

成分置換についても式 (30) に基づいて解決できる。まず、音源とマイクの位置関係が先験的に与えられる場合について述べる。すなわち、簡単のため、2 個の音源と 2 個のマイクが対向して並んでおり、 $n = 1$ 番目の音源は $m = 1$ 番目のマイクに近く、 $n = 2$ 番目の音源は $m = 2$ 番目のマイクに近い、と云う先験情報がある場合を考える。この場合、伝達関数 $a_{mn}(\omega_l)$ のゲインと位相について

$$|a_{nn}(\omega_l)| > |a_{mn}(\omega_l)|, \angle a_{nn}(\omega_l) > \angle a_{mn}(\omega_l) \quad \text{for } m \neq n \quad (31)$$

なる不等式が成り立ち、これを式 (30) に反映させると、

$$|\xi_{11}(\omega_l, k)| > |\xi_{21}(\omega_l, k)|, |\xi_{22}(\omega_l, k)| > |\xi_{12}(\omega_l, k)| \quad (32)$$

$$\angle \xi_{11}(\omega_l, k) > \angle \xi_{21}(\omega_l, k), \angle \xi_{22}(\omega_l, k) > \angle \xi_{12}(\omega_l, k) \quad (33)$$

なる関係が得られる [19][22]。したがって、生成された分割スペクトル $\xi_{\hat{n}}$ が式 (32) のゲイン条件を満たすとき成分置換はないと判定され、満たさないとき成分置換が起きていると判定できる。また、式 (33) の位相条件からも同様な判定が行える。ゲイン条件と位相条件は理論上は全く等価であるが、実際の応用では音響信号の伝達特性が周波数帯域で異なるため、帯域毎に使い分けることで精度の高い成分置換の修正が可能となる [22]。

次に、音源が音声と雑音の場合、前者の分布は非ガウス的で、後者の分布はガウス分布に近いことが知られている。これをエントロピーの観点から焼き直して分割スペクトルのエントロピーの大小関係として得られる先験情報に基づいて、成分置換を解決する方法も提案されている [23]。この方法は音源とマイクの配置に依存しない点に特徴があり、単一話者の音声の抽出を目的とする雑音除去法としては実用的である。

また、2 つのマイクの中心から見た音源 $s_1(t)$ と $s_2(t)$ の到来方向をそれぞれ $\theta_1(\omega_l)$ と $\theta_2(\omega_l)$ とするとき、これらの推定値は分離行列 $W(\omega_l)$ の逆行列をもとに、

$$\hat{\theta}_1(\omega_l) = \cos^{-1} \left(\frac{c(\angle[W^{-1}(\omega_l)]_{21} - \angle[W^{-1}(\omega_l)]_{11})}{2dF_s\omega_l} \right) \quad (34)$$

$$\hat{\theta}_2(\omega_l) = \cos^{-1} \left(\frac{c(\angle[W^{-1}(\omega_l)]_{22} - \angle[W^{-1}(\omega_l)]_{12})}{2dF_s\omega_l} \right) \quad (35)$$

と与えられるが式 (30) から導かれる。ここに、 c は音速、 d はマイク間距離、 F_s はサンプリング周波数で、 $[W^{-1}(\omega_l)]_{nm}$ は $W(\omega_l)$ の逆行列 $W(\omega_l)^{-1}$ の (n, m) 要素の位相である。したがって、到来方向の推定値 $\hat{\theta}_1(\omega_l)$ と $\hat{\theta}_2(\omega_l)$ を用いて、周波数毎に成分置換を是正でき

る。この方法は先験情報を必要としない点に特徴があり、2個以上の音源に対する拡張も容易である。

式(30)に基づいて成分置換の是正は、信号源の個数が $N = 50$ と多い場合にも適用できる[20]。地上デジタル放送や無線LANの規格として採用されている直交周波数分割多重(OFDM: Orthogonal Frequency Division Multiplexing)方式で、送信局と受信局の搬送周波数がドップラー効果等によりオフセット(ずれ)が生じた場合、サブキャリア間に干渉が起きて伝送誤りが起きる。これを回避するには、オフセットを推定するとともに、通信路の伝送特性を推定する必要がある[24]。OFDM伝送の性質より、サブキャリア間の干渉は、

$$|\xi_{n\bar{n}}| = |a_{nn}s_n| > |a_{mn}s_n| = |\xi_{m\bar{n}}| \quad \text{for } m \neq n \quad (36)$$

となることが導かれる。ここに、 a_{mn} は n 番目のサブキャリアから m 番目のサブキャリアへの干渉度、 s_n は n 番目のサブキャリアに割り当てられた送信シンボルである。式(36)は、 $\xi_{m\bar{n}}$ は、その絶対値が $\xi_{m\bar{n}}$ の第一添字 m が $m = n$ となるときの最大となることを示しており、この事実を利用することでサブキャリア間の成分置換を是正できる。

5 むすび

本稿では、独立成分分析(ICA)の代表的手法である自然勾配法とFastICA法について、瞬時混合モデルの枠内で導出原理を述べた。その中で、Infomax法や最尤法に基づくICA法が自然勾配法と同類の方法であることを示すと同時に、FastICA法の収束が自然勾配法等に比べて速い理由を明らかにした。また、実環境下においては、時間領域畳込みモデルに基づく時間領域ICA(TDICA)と、周波数領域瞬時混合モデルに基づく周波数領域ICA(FDICA)の2つのアプローチが考えられるが、アルゴリズムの収束は前者に比べて後者が有利となることを明らかにした。また、ICA特有の問題として知られるスケールの不定性と成分置換問題に言及し、これらの問題が分割スペクトルを導入することで解決できることを示した。

実環境下での音響信号を対象にする場合、1フレームを数十[msec]として、混合信号の短時間スペクトルを求めることが一般的である。この場合、インパルス応答長が1フレーム以内に収まれば、式(21)は近似モデルとして十分に意味をなす。しかし、インパルス応答長がフレーム長を越えると、式(21)の近似は崩れる。そのため、残響時間が長くなった場合、周波数領域の畳込みモデルに基づく方法[25]などがあるが、今後、さらなる改善が望まれる。また、適用分野によって、スコア関数をどう選ぶかで、ICAの分離能力は異なる。この観点からスコア関数の推定や分布の推定を含めたICAアプローチもあるが[26][27]、これについても発展が望まれる。

参考文献

- [1] A. Cichocki and S. Amari: Adaptive blind signal and image processing, learning algorithm and applications; *John Wiley & Sons* (2002)
- [2] A. Hyvärinen, J. Karhunen and E. Oja: Independent component analysis; *John Wiley & Sons* (2001)
- [3] T. W. Lee: Independent component analysis; *Kluwer Academic* (1998)

- [4] O. Yilmaz and S. Rickard: Blind separation of speech mixtures via time-frequency masking; *IEEE Trans. Signal Processing*, Vol. 52, No. 7, pp. 1830-1847 (2004)
- [5] K. Fujita: Remarks on a method of blind source separation; *Information*, Vol. 13, No. 3(B), pp. 829-834 (2010)
- [6] B. A. Olshausen and D. J. Field: Sparse coding of sensory inputs; *Current Opinion in Neurobiology*, Vol. 14, pp. 481-487 (2004)
- [7] A. Cichocki, R. Zdunek and S. Amari: Nonnegative matrix and tensor factorization; *IEEE Signal Processing Magazine* Vol. 25, No. 1, pp. 142-145 (2008)
- [8] R. Ashino, T. Mandai and A. Morimoto: Blind source separation of spatio-temporal mixed signals using phase information of analytic wavelet transform; *Int. J. Wavelets Multiresolut. Inf. Process.*, Vol. 8, No. 4 pp. 575-594 (2010)
- [9] S. Amari: Natural Gradient Works Efficiently in Learning; *Neural Computation*, Vol. 10, No. 2, pp. 251-276 (1998)
- [10] A. Hyvärinen and E. Oja: Independent component analysis: algorithms and applications; *Neural Networks*, Vol. 13, No. 4-5, pp. 411-430 (2000)
- [11] A. Hyvärinen: Fast and Robust Fixed-Point Algorithms for Independent Component Analysis; *IEEE Trans. Neural Networks*, Vol. 10, No. 3, pp. 626-634 (1999)
- [12] A. J. Bell and T. J. Sejnowski: An information maximization approach to blind separation and blind deconvolution; *Neural Computation*, Vol. 7, No. 6, pp. 1129-1159 (1995)
- [13] T. W. Lee, M. Girolami and T. J. Sejnowski: Independent Component Analysis Using an Extended Infomax Algorithm for Mixed Subgaussian and Supergaussian Sources; *Neural Computation*, Vol. 11, No. 2, pp. 417-441 (1999)
- [14] D. T. Pham and P. Garat: Blind separation of mixture of independent sources through a quasi-maximum likelihood approach; *IEEE Trans. Signal Processing*, Vol. 45, No. 7, pp. 1712-1725 (1997)
- [15] E. A. P. Habets, S. Gannot, I. Cohen and P. C. W. Sommen: Joint Dereverberation and Residual Echo Suppression of Speech Signals in Noisy Environments; *IEEE Trans. Audio Speech and Language Processing*, Vol. 16, No. 8, pp. 1433-1451 (2008)
- [16] P. Smaragdis: Blind separation of convolved mixtures in the frequency domain; *Neurocomputing*, Vol. 22, pp. 21-34 (1998)
- [17] E. Bingham and A. Hyvärinen: A fast fixed-point algorithm for independent component analysis for complex valued signals; *Int. J. Neural Systems*, Vol. 10, No. 1, pp. 1-8 (2000)
- [18] N. Murata, S. Ikeda and A. Ziehe: An approach to blind source separation based on temporal structure of speech signals; *Neurocomputing*, Vol. 41, Issue 1-4, pp. 1-24 (2001)

- [19] H. Gotanda, K. Nobu, T. Koya, K. Kaneda, T. Ishibashi, N. Haratani: Permutation correction and speech extraction based on split spectrum through FastICA; *Proc. ICA2003*, pp. 379-384 (2003)
- [20] 中河史成, 高瀬成史, 白土浩, 五反田博: ICA による OFDM 周波数オフセットの推定とシンボル復元; 電子情報通信学会論文誌 A, Vol. J91-A, No. 4, pp. 448-457 (2008)
- [21] K. Nobu, T. Koya, K. Kaneda, N. Haratani and H. Gotanda: Noise Reduction Using Locational Information on Target Sound Source; *J. Robotics and Mechatronics*, Vol. 15, No. 1, pp. 15-23 (2003).
- [22] 石橋孝昭, 井上勝裕, 五反田博, 熊丸耕介: 実環境下での伝達特性を利用した周波数領域 ICA の成分置換問題の解決; システム制御情報学会論文誌, Vol. 19, No. 12, pp. 471-478 (2006)
- [23] 金田圭市, 古屋武志, 五反田博: 分割スペクトルのエントロピーに基づく成分置換解消法; 電子情報通信学会論文誌 A, Vol. J87-A, No. 7, pp. 1065-1069 (2004)
- [24] サイビシット ヴィタヤ, 木村 哲也, 中河 史成, 白土 浩, 原谷 直実, 五反田博: QAM-OFDM における周波数オフセットと伝送路のブラインド推定; 電子情報通信学会論文誌 A, Vol. J92-A, No. 3, pp. 141-149 (2009)
- [25] 古屋武志, 石橋孝昭, 白土浩, 五反田博: 周波数領域畳込みモデルに基づく高残響環境下での音源分離; システム制御情報学会論文誌, Vol. 22, No. 8, pp. 287-294 (2009)
- [26] J. Karvanen and V. Koivunen: Blind separation methods based on Pearson system and its extensions; *Signal Processing*, Vol. 82, No. 4, pp. 663-673 (2002)
- [27] F. R. Bach and M.I. Jordan: Kernel independent component analysis; *J. Machine Learning Research*, Vol. 3, pp. 1-48 (2002)

付録 分割信号と信号源の関係

ICA により信号源 s_n が成分置換されて, 分離信号が $u_{\tilde{n}}$ と得られるとき, 分離信号 $u_{\tilde{n}}$ から誘導される分割信号 $\xi_{\tilde{n}}$ の第 m 番目の成分 $\xi_{m\tilde{n}}$ は, 信号源 s_n と

$$\xi_{m\tilde{n}} = a_{mn}s_n \quad m = 1, 2, \dots, N \quad (37)$$

なる関係で対応づけられる.

【証明】

N 個の信号源 $\mathbf{s} = [s_1, s_2, \dots, s_N]^T$ を N 個のセンサで観測して得られる混合信号を $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ とする. このとき, ICA により得られる分離信号 $\mathbf{u} = [u_1, u_2, \dots, u_N]^T$ は,

$$\mathbf{u} = P\mathbf{D}\mathbf{s} \quad (38)$$

と与えられる。ここに、 $D = \text{diag}[d_1, d_2, \dots, d_N]$ はスケールの不定性を表す対角行列で、 P は各行各列にただ一つの値 1 をもつ置換行列である。

一方、混合行列を A 、分離行列を W とするとき、分離信号は

$$\mathbf{u} = W\mathbf{x} = WAs \quad (39)$$

と関係づけられる。また、式 (39) は両辺を展開すると次のようになる。

$$\begin{bmatrix} u_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ u_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ u_{\tilde{N}} \end{bmatrix} = W \sum_{n=1}^N \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{N,n} \end{bmatrix} s_n \quad (40)$$

以上の準備のもとで、第 $\tilde{n}$ 番目の分離信号 $u_{\tilde{n}}$ は、第 n 番目の信号源 s_n が成分置換されて

$$u_{\tilde{n}} = d_n s_n \quad (41)$$

と得られたものと仮定する。このとき、第 n 番目の信号源 s_n のみが活性 ($s_n \neq 0$) で、その他の信号源 s_n は不活性 ($s_n = 0$) とすると、式 (40) より、

$$\begin{bmatrix} 0 \\ \vdots \\ 0 \\ u_{\tilde{n}} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = W \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{Nn} \end{bmatrix} s_n \quad (42)$$

となる関係が得られる。

更に、第 $\tilde{n}$ 番目の分離信号 $u_{\tilde{n}}$ から誘導される分割信号 $\boldsymbol{\xi}_{\tilde{n}} = [\xi_{1\tilde{n}}, \xi_{2\tilde{n}}, \dots, \xi_{N\tilde{n}}]^T$ を

$$\boldsymbol{\xi}_{\tilde{n}} = W^{-1}[0, \dots, 0, u_{\tilde{n}}, 0, \dots, 0]^T \quad (43)$$

と定義すると、式 (42) より

$$\begin{bmatrix} \xi_{1\tilde{n}} \\ \xi_{2\tilde{n}} \\ \vdots \\ \xi_{N\tilde{n}} \end{bmatrix} = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{N,n} \end{bmatrix} s_n \quad (44)$$

となる。したがって、これを要素表示すれば、

$$\xi_{m\tilde{n}} = a_{mn} s_n \quad m = 1, 2, \dots, N \quad (45)$$

となって、式 (37) が成り立つことが証明される。式 (45) は、成分置換が起きている場合でも、信号源 s_n から各センサ ($m = 1, 2, \dots, N$) までの伝達特性 a_{mn} の番号 m は、分割信号 $\boldsymbol{\xi}_{\tilde{n}}$ の要素 $\xi_{m\tilde{n}}$ の番号 m と一致することを示している。したがって、各信号源から N 個のセンサーまでの伝達構造 a_{mn} は、たとえ成分置換があっても、保存される。