

# オーバーコンプリート形ブラインド信号源分離の構成法と学習法

加藤 治雄<sup>†</sup> 中山 謙二<sup>††</sup> 平野 晃宏<sup>††</sup>

<sup>†</sup> 金沢大学大学院 自然科学研究科 電子情報工学専攻

<sup>††</sup> 金沢大学大学院 自然科学研究科 電子情報科学専攻

E-mail: <sup>††</sup>nakayama@t.kanazawa-u.ac.jp

**あらまし** 本論文では、信号源の数がセンサ数より多い場合のブラインド信号源分離 (BSS) についてのフィードバック構造と学習アルゴリズムを提案する。提案法では、信号源をいくつかのグループに分離し、ある条件下では少なくとも一つの出力を単一の信号源として分離できる。出力に単一の信号源が分離されるように分離回路の出力を求め、観測信号に対して分離された単一の信号源をフィードバックする。そして、観測信号からその成分を差し引くことにより、信号源数を等価的に減らす。

**キーワード** 信号源分離, オーバーコンプリート, 音声信号, 学習アルゴリズム, 基本周期

## A Feedback Structure and Its Learning Algorithm for Overcomplete Blind Source Separation

Haruo KATOU<sup>†</sup>, Kenji NAKAYAMA<sup>††</sup>, and Akihiro HIRANO<sup>††</sup>

<sup>†</sup> Division of Electronics and Computer Science

Graduate School of Natural Science & Technology, Kanazawa Univ.

<sup>††</sup> Division of Electrical Engineering and Computer Science

Graduate School of Natural Science and Technology, Kanazawa Univ.

E-mail: <sup>††</sup>nakayama@t.kanazawa-u.ac.jp

**Abstract** A feedback structure and its learning algorithm are proposed for overcomplete blind source separation (BSS). The number of sources is larger than that of the sensors. In the proposed method, the signal sources are separated in several groups. Among them, at least one output can separate a single signal source under some condition. The outputs of the unmixing block are evaluated, which output provides a single signal source. The separated single signal source is fed back to the input node, and subtracted from the input signal, in order to reduce the number of equivalent signal sources.

**Key words** Signal source separation, Overcomplete, Speech signal, Learning algorithm, Fundamental period

### 1. ま え が き

近年、世界的規模の通信を通して多くの情報量が伝えられ処理されている。通信端末は様々な環境において使用され、同時に高いクオリティーが必要とされている。雑音除去、エコー除去、回線等化などの信号処理において、信号復元とは大変重要な技術となっている。また、信号源や干渉成分について十分な情報が得られない上に、混合過程や伝送過程についても事前に知ることはできない。このような環境では、信号源の統計的性質に基づいて信号源分離を行うブラインドソースセパレーション (BSS) が重要となる。一般に信号源の数を予測できない場合が多く、信号源の数よりセンサ数が少なくなる場合もあり、信号源分離が難しくなる (オーバーコ

ンプリート形 BSS)。

BSS の学習アルゴリズムが Jutten をはじめとして、数多く提案されている [1][2][3][4][5]。そのなかで、BSS の収束性や分離特性は出力信号の PDF と更新式に含まれる非線形関数の関係に大きく依存している。すでに、Kurtosis に基づく適した非線形関数の決定法がいくつか提案されている [6][7]。一般的なオーバーコンプリート形 BSS も数多く提案されているが、まだ完全分離に近い分離特性を得られている報告はないと言える [8][9][10]。これらのシステム構成では信号源数を  $M$ 、センサ数を  $N$  とした場合 ( $M > N$ ) に、出力として一度に  $M$  個の信号源に分離する。これは大きく分けて、混合行列推定過程と信号源復元過程からなり、未知の混合行列を正確に推定しないと、信号源をうまく復元できない。

本稿では、一般的なシステム構成とは異なるシステム構成を新たに提案する。これは、一度に信号源を個々に分離するものではなく、段階的に信号源を分離するものである。そこで、まず、このシステム構成での分離性能の限界を理論的に明らかにし、次に、分離性能を限界に近付ける学習法について検討する。

## 2. 提案するオーバーコンプリート形 BSS

### 2.1 回路構成

本章では、センサのチャンネル数を信号源のチャンネル数より少ない条件で信号源分離を行なう。今回提案する分離回路を図 1 に示す。

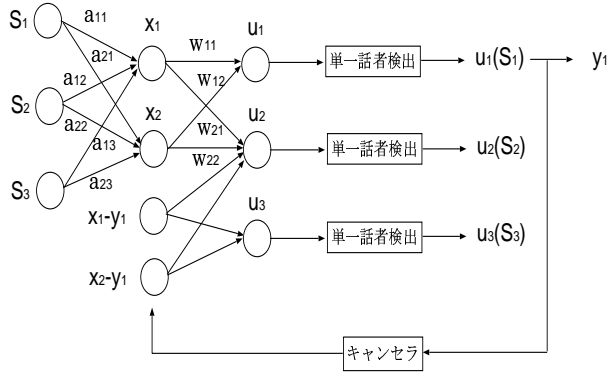


図 1 提案法の回路構成

信号源の  $M$  に対して、センサ数を  $\lfloor M/2 + 1 \rfloor$  以下とする。この条件下で信号源分離が行なわれると、ある出力に一個の信号源が分離される。これを検出し、フィードバックしてキャンセルする。観測信号において、信号源の数を減らした形で、さらに信号源分離を行なう。このような処理を繰り返して、全ての信号源を個々に分離する。信号源が音声である場合は、出力において、単一音声と複数音声を識別する必要がある。このために、音声の性質を利用した方法を検討する。

### 2.2 信号源数とセンサ数の関係

今回提案する回路構成での信号源数とセンサ数の関係について、信号源数 ( $M$ ) がセンサ数 ( $N$ ) より多い場合の例で説明する。センサへの入力  $\mathbf{x}$  と出力  $\mathbf{y}$  はそれぞれ以下のように表される。なお、 $\mathbf{P}$  は信号源から出力までの伝達特性を表す。

$$\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_M \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_N \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} p_{11} & \cdots & p_{1M} \\ \vdots & \ddots & \vdots \\ p_{N1} & \cdots & p_{NM} \end{bmatrix} \begin{bmatrix} \mathbf{S}_1 \\ \vdots \\ \mathbf{S}_M \end{bmatrix} \quad (3)$$

ここで、ある出力に信号源が一つ分離されるためには、 $\mathbf{P}$  の要

素のうち非零の数は  $M$  個、 $\mathbf{P}$  の要素数は  $NM$  個であることから、零の数は  $NM - M$  個でなければならない。また、未知数の数は分離行列  $\mathbf{W}$  の数に等しいので  $N^2$  である。以上のことから、方程式の数は  $NM - M$  個、未知数の数は  $N^2$  個であることから、解が存在するためには、

$$N^2 \geq M(N-1) \rightarrow \frac{N^2}{N-1} \geq M \quad (4)$$

の関係が成り立つ。例として、 $N = 2$  に対して  $M \leq 4$ 、 $N = 3$  に対して  $M \leq 4.5$ 、 $N = 4$  に対して  $M \leq 5.3$  である。これを式 (4) から近似すると、信号源数が  $M$  の場合、センサ数は  $\lfloor \frac{M}{2} + 1 \rfloor$  となり、この条件下で信号源分離を行なうと、ある出力に一つの信号源を取り出すことができる。

### 2.3 一巡目の分離性能の理論的限界

実際のオーバーコンプリート形 BSS では、分離行列の理想解である、混合行列の逆行列は存在しないので、信号源の完全分離は不可能である。本節では、信号源のグループ分離を行ない、分離性能の理論的限界を求めて、提案する回路構成の有効性を確認する。

ここでは簡単のため、信号源を 3 個、センサを 2 個とする。さらに、出力  $\mathbf{y}_1$  には  $\mathbf{S}_1$  が、出力  $\mathbf{y}_2$  には  $\mathbf{S}_2$  と  $\mathbf{S}_3$  が分離される条件で考える。なお、混合過程の係数は、 $a_{11} = a_{23} = 1.0$ 、 $a_{13} = a_{21} = 0.3$  とし、 $a_{12}$  と  $a_{22}$  は  $a_{12} + a_{22} = 1.4$  の制約下で変化するものとする。

#### 2.3.1 出力 $\mathbf{y}_1$ での理論的限界

出力  $\mathbf{y}_1$  では  $\mathbf{S}_2$ 、 $\mathbf{S}_3$  成分がキャンセルされると考えると、

$$\mathbf{S}_2 \rightarrow \mathbf{y}_1: 0 \Rightarrow a_{12}w_{11} + a_{22}w_{12} = p_{12} \quad (5)$$

$$\mathbf{S}_3 \rightarrow \mathbf{y}_1: 0 \Rightarrow a_{13}w_{11} + a_{23}w_{12} = p_{13} \quad (6)$$

の関係式が得られる。理想的には出力  $\mathbf{y}_1 = \mathbf{S}_1$  であることから、

$$\mathbf{S}_1 \rightarrow \mathbf{y}_1: 1 \Rightarrow a_{11}w_{11} + a_{21}w_{12} = p_{11} \quad (7)$$

とする。ここで、出力  $\mathbf{y}_1$  で  $\mathbf{S}_2$ 、 $\mathbf{S}_3$  成分をキャンセルする条件としては、式 (7) より、以下の誤差関数  $E_1$  が最小となるように分離行列  $\mathbf{W}$  を求める。

$$\begin{aligned} E_1 &= (a_{12}w_{11} + a_{22}w_{12})^2 + (a_{13}w_{11} + a_{23}w_{12})^2 \\ &= \alpha w_{11}^2 + \beta w_{11} + \gamma \rightarrow \text{最小} \end{aligned} \quad (8)$$

式 (8) 中の  $\alpha$  は正であるので、誤差関数の偏微分を求めると、分離行列の理想解である  $w_{11}$ 、 $w_{12}$  が得られ、これを用いて出力  $\mathbf{y}_1$  における分離性能の評価を行なう。

#### 2.3.2 出力 $\mathbf{y}_2$ での理論的限界

出力  $\mathbf{y}_2$  では  $\mathbf{S}_1$  成分がキャンセルされ、 $\mathbf{S}_2$ 、 $\mathbf{S}_3$  成分が出力されると考えると、

$$\mathbf{S}_1 \rightarrow \mathbf{y}_2: 0 \Rightarrow a_{11}w_{21} + a_{21}w_{22} = p_{21} \quad (9)$$

$$\mathbf{S}_2 \rightarrow \mathbf{y}_2: 1 \Rightarrow a_{12}w_{21} + a_{22}w_{22} = p_{22} \quad (10)$$

$$\mathbf{S}_3 \rightarrow \mathbf{y}_2: 1 \Rightarrow a_{13}w_{21} + a_{23}w_{22} = p_{23} \quad (11)$$

となる。出力  $\mathbf{y}_2$  で  $\mathbf{S}_1$  成分をキャンセルさせる条件としては、

$$E_2 = \frac{\alpha w_{22}^2 + \beta w_{22} + \gamma}{\alpha w_{22}^2 + \beta w_{22} + c + 1} \rightarrow \text{最小} \quad (12)$$

の誤差関数  $E_2$  が最小となるような  $w_{21}$ 、 $w_{22}$  を求め、出力  $\mathbf{y}_2$  における分離性能の評価を行なう。

### 3. 一巡目の分離回路の学習法

#### 3.1 学習法

信号源  $S_1$  と  $S_2$  と  $S_3$  に分離する場合と、信号源を個々に分離する場合の出力の統計的独立性は同じになる。よって、分離回路の学習法は、相互情報量を式 (13) の誤差関数とする自然勾配法 [3] を用いる。

$$l(\mathbf{x}, \mathbf{W}) = -\log |\det(\mathbf{W})| - \sum_{i=1}^m \log p_i(y_i) \quad (13)$$

この時の分離行列の更新分は、出力  $\mathbf{y}$  の各要素間の相互情報量の最小化を目指すように更新される。分離回路の更新式は、

$$\begin{aligned} \mathbf{W}(n+1) &= \mathbf{W}(n) + \eta[\mathbf{A}(n) - \varphi(\mathbf{y}(n))\mathbf{y}^T(n)]\mathbf{W}(n) \\ \varphi(\mathbf{y}(n)) &= \frac{p'(\mathbf{y}(n))}{p(\mathbf{y}(n))} = \frac{1 - \exp(-5\mathbf{y}(n))}{1 + \exp(-5\mathbf{y}(n))} \end{aligned} \quad (14)$$

とし、これを従来法とする。しかし、従来法だけの学習だけでは、良好な特性が得られないので、検討する必要がある。

#### 3.2 非線形関数の制御

観測データの平均値付近の分布を表現する 4 次の統計量 Kurtosis ( $\kappa_4$ ) によって信号源の分布を調べ、それに応じて非線形関数を制御する [7]。Kurtosis が正であれば、確率密度関数は Super-Gaussian、負であれば、確率密度関数は Sub-Gaussian となる。式 (13) の誤差関数の安定条件から、Kurtosis が負であれば、非線形関数は  $y^3$  が有効で、Kurtosis が正であれば  $\tanh(y)$  が有効となる。Kurtosis は正負を激しく変動し、非線形関数の間には連続性がないので、Kurtosis により非線形関数を切り替えようとする、断続的なものになる。そこで非線形関数を以下のようにモデル化する。

$$\varphi(\mathbf{y}(n)) = a \tanh(\mathbf{y}(n)) + (1 - a)\mathbf{y}(n)^3 \quad (15)$$

$$a = \frac{1 - \exp(-2.1\kappa_4 - 2.5)}{1 + \exp(-2.1\kappa_4 - 2.5)} \quad (16)$$

本論文で用いている信号源は音声なので、音楽ほど Kurtosis の変動は激しくないが、音声でも Kurtosis は正負間を変動するので、非線形関数を制御することは効果があるといえる。

#### 3.3 学習法の改良

センサへの入力である観測データのヒストグラムを用いて、未知の混合行列を推定するアルゴリズムがすでに提案されている [11]。このアルゴリズムを信号源分離の学習法に取り入れることを考える。

ここでは簡単のために二次元の場合を考える。センサへの入力である観測信号ベクトルを  $\mathbf{X}$  とし、 $\mathbf{X}$  を超球上に投影したものを  $\mathbf{Y} = \pi \circ \mathbf{X} = \pi(\mathbf{X})$  とする。 $\mathbf{Y}$  のヒストグラムを用いることで、直接観測データの投影上の位置を以下のように求めることができる。

$$\mu : [0, \pi) \rightarrow \mathbf{R} \quad (17)$$

$$\varphi \rightarrow \frac{l_1(\varphi) + l_2(\varphi)}{2} - \left(\varphi + \frac{\pi}{2}\right)$$

$$l_i(\varphi) = F_Y^{-1} \left( \frac{F_Y(\varphi + i\frac{\pi}{2}) + F_Y(\varphi + (i-1)\frac{\pi}{2})}{2} \right) \quad (18)$$

ここで、 $F_Y$  は  $\mathbf{Y}$  の累積分布関数で、確率変数  $X$  が  $-\infty < x < \infty$  の  $x$  に対して、 $x$  以下の値を取る確率である。以下に  $\mathbf{Y}$  のヒストグラムと累積分布関数の例を示す。

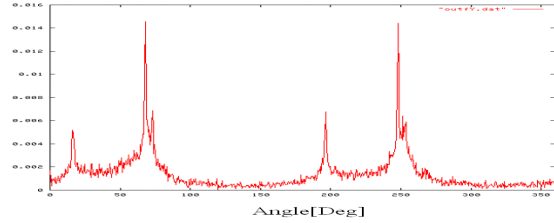


図2 Yのヒストグラム

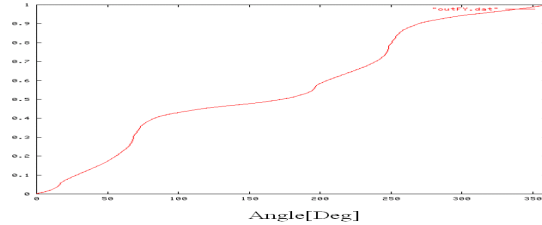


図3 Yの累積分布関数

#### 3.4 提案法

3章で紹介した学習アルゴリズムを提案法とすると、分離行列は以下のように更新される。

$$\mathbf{w}_j^* = \operatorname{argmax}_{\mathbf{w}_j} |\mathbf{Y}_i^T * \mathbf{w}_j| \quad (19)$$

$$\varphi(\mathbf{y}(n)) = a \tanh(\mathbf{y}(n)) + (1 - a)\mathbf{y}(n)^3 \quad (20)$$

$$\mathbf{w}_j^*(n+1) = pr(\mathbf{w}_j^*(n) - \eta[\langle \varphi(\mathbf{y}(n))\mathbf{y}^T(n) \rangle \mathbf{w}_j^*(n)])$$

ここで、式 (19) は、 $\mathbf{Y}$  と分離行列  $\mathbf{w}_j$  の内積を求めて、最も近接した分離行列ベクトルだけを更新する。

#### 3.5 性能評価

本論文で用いた信号源は、以下の音声を繰り返したものを用いる。

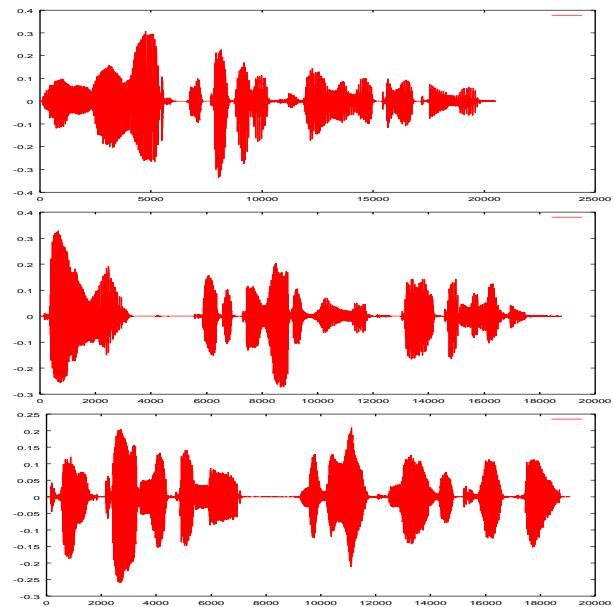


図4 用いる音声データ (男声2個, 女声1個)

出力 $\mathbf{y}_1, \mathbf{y}_2$ での評価方法として、以下の評価式を用いる。

$$SNR_1 = 10 \log_{10} \frac{1}{E_1} = 10 \log_{10} \left( \frac{p_{11}^2}{p_{12}^2 + p_{13}^2} \right) [dB] \quad (21)$$

$$SNR_2 = 10 \log_{10} \frac{1}{E_2} = 10 \log_{10} \left( \frac{p_{22}^2 + p_{23}^2}{p_{21}^2} \right) [dB] \quad (22)$$

次に、出力 $\mathbf{y}_2$ では $\mathbf{S}_2, \mathbf{S}_3$ 成分が分離されるので、それらの大きさのバランスが重要となる。そのために、 $\mathbf{S}_2$ に対する $\mathbf{S}_3$ の大きさの比率として  $R = p_{23}/p_{22}$  を定義する。信号源分離としては、 $R$ が 1 に近いことが望ましい。以下に横軸を  $\alpha = a_{12}/a_{22}$  とした時の  $SNR_1, SNR_2, R$  の理論的境界と提案法による学習結果を示す。なお、 $SNR_2$  の理論的境界は無限大に発散するので割愛する。

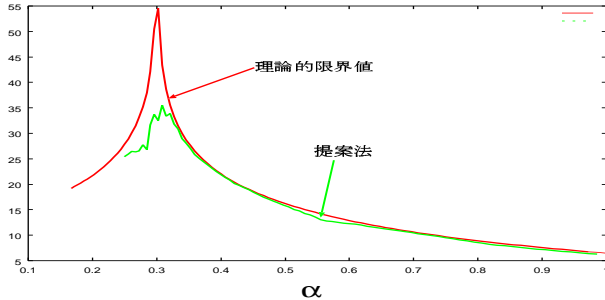


図5  $SNR_1$

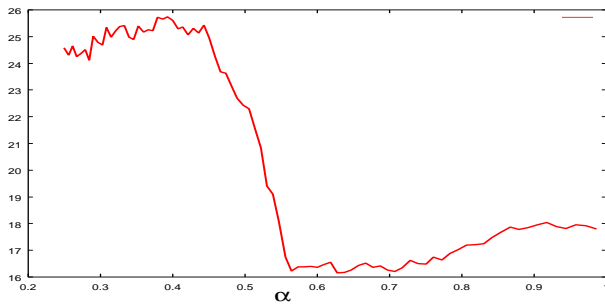


図6  $SNR_2$

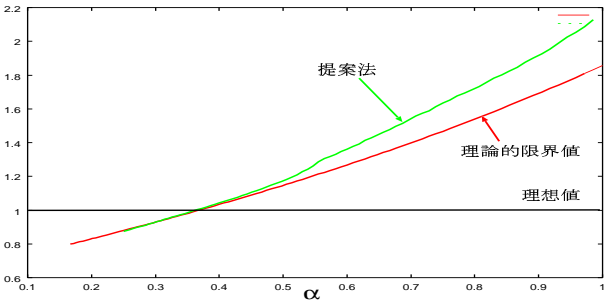


図7  $R$

図5, 図7 から、 $\alpha$ が 1 に近づく ( $\mathbf{S}_2$ が $\mathbf{S}_1$ と $\mathbf{S}_3$ の中間に位置する) ほど分離特性が劣化する。しかし、 $SNR_1$ は多少 $\alpha$ が大きくなっても、10[dB] 程度の特性が得られていることから、今回提案したシステム構成の有効性を確認できる。

## 4. 単一話者検出

### 4.1 基本周期導出過程

一巡目の信号群分離後、単一音声と複数音声を識別する必要がある。そこで、人間の音声や音楽の基本周期 ( $T$ ) を推定して音声識別を行なう [12]。

離散時間信号  $x_t$  の自己相関関数  $r_\tau$  は、

$$r_\tau = \sum_{j=t+1}^{t+D} x_j x_{j+\tau} \quad (23)$$

と表される。ここで、 $\tau$  は時間  $t$  に対する遅延量、 $D$  は遅延の許容範囲である。周期信号に対して、自己相関関数は基本周期でピークを示す。そこで、一定の遅延区間で、遅延 0 でない最も高いピークを抽出し、基本周期を求める。しかし、自己相関関数で基本周期を推定すると誤差が大きい。そこで、信号  $x_t$  の未知の基本周期を式 (24) が 0 になる時の  $\tau$  とすることで推定誤差が軽減される。

$$d_t(\tau) = \sum_{j=1}^D (x_j - x_{j+T})^2 = r_t(0) + r_{t+\tau}(0) - 2r_t(\tau) \quad (24)$$

しかし、式 (24) の  $d_t(\tau)$  は遅延 0 で 0 となるので、低次の項が誤って基本周期として抽出される。そこで、遅延 0 では値 1 をとり、 $d_t(\tau)$  を短い遅延間隔での平均値で除算して得られる式 (25) を以下のように設定する。

$$d'_t(\tau) = \begin{cases} 1, & \text{if } \tau = 0 \\ d_t(\tau) / \left[ (1/\tau) \sum_{j=1}^{\tau} d_t(\tau) \right], & \text{otherwise} \end{cases} \quad (25)$$

ここで、式 (23),(24),(25) をあるサンプルで計算した時の図を以下に示す。

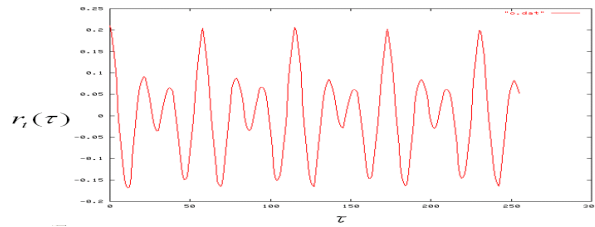


図8  $r_t(\tau)$

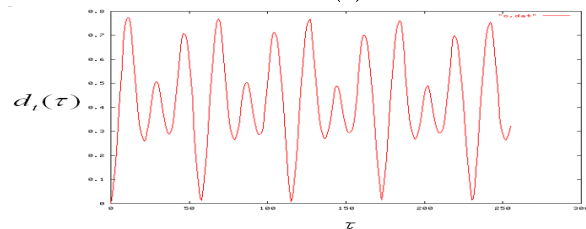


図9  $d_t(\tau)$

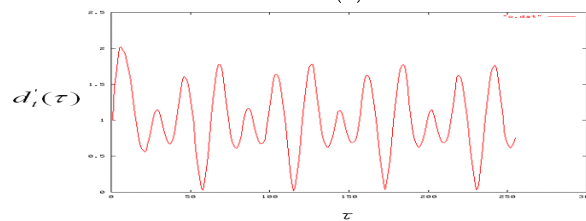


図10  $d'_t(\tau)$

自己相関関数を用いると、高次のピークを誤って選ぶ可能性がある。そこで、 $d'_t(\tau)$  に対して閾値を設定し、閾値より小さい  $d'_t(\tau)$  の極小値の候補の中で、最も小さい  $\tau$  を基本周期とする。極小値は Brent の極小値求解法により求める [13]。

非定常な音声区間(無音区間等)に対して,基本周期の推定は誤差を生じることが多い。そこで,サンプルをある短い区間に分割し,区間中最小の $d_i(T)$ の基本周期を採用する。また,無音区間は長いサンプルで続くことが多いので,自己相関関数を求める際に多大な影響を及ぼす。図11は無音区間が多い例である。

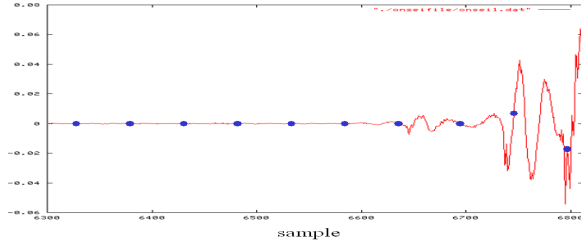


図11 無音区間の例

無音区間の対策として,50点おきに過去の値を10個とり,その絶対値を求める。そして,判別値を0.006とし,絶対値との大小関係を調べ,そのうち6個以上判別値より小さいなら,求まる基本周期は用いないことにするという処理を加える。

#### 4.2 動作評価

単一音声と複数音声によるシミュレーションを行ない,単一話者検出の動作評価をする。用いる音声は図4の音声とし,複数音声はこれらの音声を複数加算して作成するものとする。基本周期の推定結果には各音声での結果ばらつきが見られたので,評価方法として,各サンプルで推定された基本周期 $T$ における $d'(T)$ の値の平均値で評価する。以下に,図4の音声を用いてのシミュレーション結果を示す。

音声	1	2	3	1+2	1+3	2+3	1+2+3
平均値	0.120	0.139	0.084	0.268	0.217	0.231	0.318

表1 基本周期 $T$ における $d'(T)$ の平均値の比較

単一音声の場合,表1から分かるように, $d'(T)$ の平均値小さな値となる。一方,複数音声の場合,単一音声と比較すると,平均値は大きな値となる。音声1+音声2+音声3の場合は用いた音声の中では最大になっている。

以上のことより,単一音声と複数音声を識別するには,各サンプルごとの基本周期 $T$ における $d'(T)$ の平均値を比較できれば,識別が可能となること分かる。

#### 4.3 一巡目の信号源分離後の出力での評価

一巡目の信号源分離後の出力 $y_1, y_2$ でシミュレーションを行った結果を表2に示す。

$\alpha$	$\alpha = 0.4285$		$\alpha = 0.5555$		$\alpha = 0.7073$	
出力	$y_1$	$y_2$	$y_1$	$y_2$	$y_1$	$y_2$
平均値	0.1685	0.2273	0.1987	0.2270	0.2249	0.2227

表2 1巡目の信号源分離後の出力での $d'(T)$ の平均値

表2から, $\alpha$ が小さい時では,出力 $y_1$ は出力 $y_2$ に比べて, $d'(T)$ の平均値は小さいことが分かる。しかし, $\alpha$ が1に近い場合,一巡目の信号源分離も難しくなるので,出力 $y_1$ と出力 $y_2$ の $d'(T)$

の平均値はほぼ等しくなり,単一音声か複数音声を識別するのは難しい。また,表1から,信号源により $d'(T)$ の平均値に大小関係があるので,信号源の性質も多少検出結果に影響している。

### 5. フィードバック構成

仮に出力 $y_1$ が単一音声と識別された場合のフィードバック構成を考える。この場合,出力 $y_1, y_2$ は,信号源 $S_1$ が検出されたと仮定すると,

$$y_1 = (w_{11}a_{11} + w_{12}a_{21})S_1 \quad (26)$$

$$y_2 = (w_{21}a_{11} + w_{22}a_{21})S_1 = 0 \quad (27)$$

となり,出力 $y_1$ は式(28)の形で表せる。

$$y_1 = a_{11} \left( w_{11} - w_{12} \frac{w_{21}}{w_{22}} \right) S_1 \quad (28)$$

実際にセンサへ入力される観測信号中の信号源 $S_1$ 成分は,センサ1では $a_{11}S_1$ ,センサ2では $a_{21}S_1$ であるので,式(28)より,

$$a_{11}S_1 = \frac{y_1}{w_{11} - w_{12} \frac{w_{21}}{w_{22}}} \quad (29)$$

$$a_{21}S_1 = -\frac{w_{21}}{w_{22}} a_{11}S_1 = -\frac{w_{21}}{w_{22}} \cdot \frac{y_1}{w_{11} - w_{12} \frac{w_{21}}{w_{22}}} \quad (30)$$

となる。式(29)と式(30)には混合行列 $A$ の成分が含まれていないので,この式をそれぞれセンサ1とセンサ2の観測信号から差し引くと,理論的に $S_1$ 成分をキャンセルすることが可能となる。 $\alpha = 0.428$ とした場合のフィードバック後の観測信号を以下に示す。

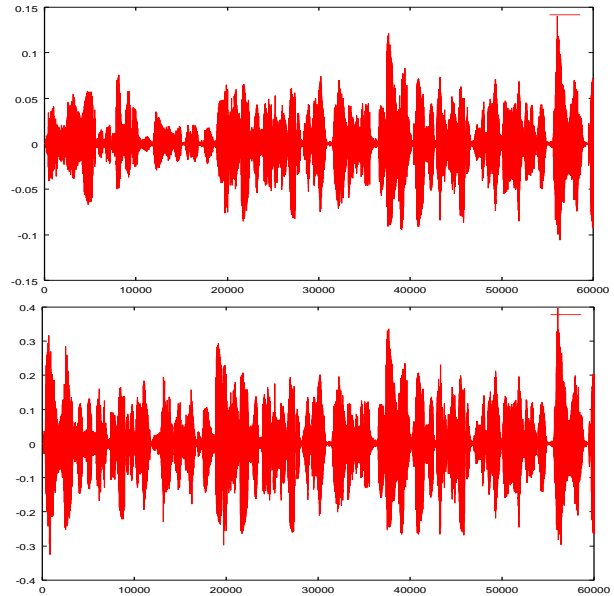


図12 フィードバック後の観測信号

図13から,フィードバック後の観測信号は,分離回路の学習が進むに従って,振幅は異なるがほぼ同じ形状(信号源分離が難しい状態と考えられる)となること分かる。捕捉として,ここでは信号源 $S_1$ の場合を取り上げたが, $S_2$ と $S_3$ の場合でも同様にフィードバックすることは可能であり,出力 $y_2$ が単一音声と識別される場合も同様である。

## 6. フィードバック後の各信号源の比率

フィードバック後の観測信号に対する各信号源の比率を求める。比率を求めるには、未知の混合行列  $\mathbf{A}$  を推定する必要がある。図 13 は  $\alpha = 0.428$  の時の観測信号の分布である。

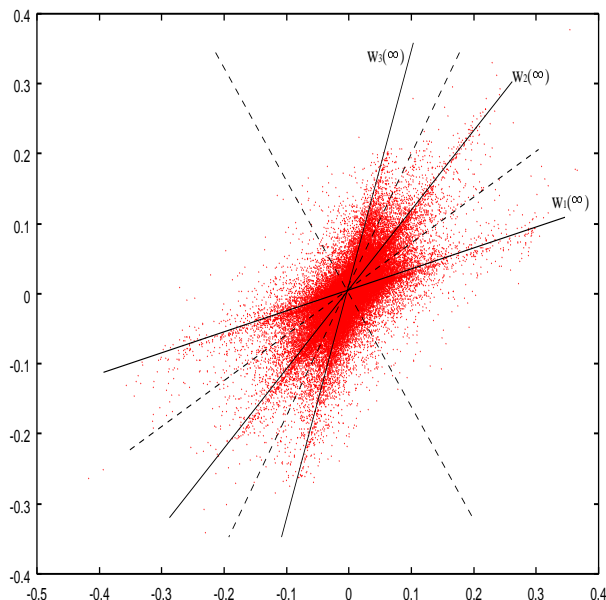


図 13  $\alpha = 0.428$  の時の観測信号の分布

観測信号ベクトル  $\mathbf{Y}$  のあるサンプルを  $y_n$  とすると、推定混合行列の列ベクトル  $w_i$  は式 (18), (19) を用いて以下のように更新される。

$$w_i(n) = w_i(n-1) + \eta(n) \text{sgn}(y_n - w_i(n-1)) \quad (31)$$

ここで、仮に未知の混合行列  $\mathbf{A}$  を、

$$\mathbf{A} = \begin{bmatrix} 1.0 & 0.42 & 0.3 \\ 0.3 & 0.98 & 1.0 \end{bmatrix} \simeq \begin{bmatrix} 0.957 & 0.393 & 0.287 \\ 0.287 & 0.919 & 0.957 \end{bmatrix}$$

とすると、推定された混合行列  $\mathbf{A}'$  は、

$$\mathbf{A}' = \begin{bmatrix} 0.958 & 0.398 & 0.288 \\ 0.284 & 0.917 & 0.957 \end{bmatrix}$$

となり、ある程度の推定精度が得られている事が分かる。この推定結果より得られるフィードバック後の各信号源の割合は以下ようになる。

	信号源 1	信号源 2	信号源 3
センサ 1	-0.100148	0.193119	0.226193
センサ 2	-0.115542	0.839475	0.934177

表 3 フィードバック後の各信号源の割合

フィードバックによって  $S_1$  成分はキャンセルされるが、実際は表 3 から分かるようにフィードバック後でも  $S_1$  成分は残る。また、フィードバック後、両方のセンサに  $S_2, S_3$  成分は同じような割合で干渉する (図 12) ことから、二巡目の信号群分離は難しくなる。

## 7. ま と め

本稿では、オーバーコンプリート形 BSS に対して、回路方式と学習法を提案した。まず一巡目の学習法は、従来の自然勾配法だけで学習を行っても良好な結果が得られなかったことから、新たに学習法を提案し、理論的境界に近付くような結果が得られた。次に、一巡目の信号群分離後の出力が単一音声か複数音声か識別を行なったが、まだまだ完全に識別できていたとは言えず、これからの検討が必要である。また、出力の単一音声を用いてフィードバックを行なったが、完全に単一音声をキャンセルできてはいなく、この成分が二巡目の信号源分離に影響を与えている。今後、二巡目の信号群分離の分離特性を改善する方法を検討する。

### 文 献

- [1] C.Jutten and J.Herault and A.Guerin, "IN.C.A.: An independent components analyser based on an adaptive non-mimetic network", in: J.Demongeot, T.Herve, V.Rialle and C.Roche, eds., Artificial Intelligence and Cognitive Sciences, Manchester Univ. Press, Manchester, 1988.
- [2] S. Amari, T. Chen and A. Cichocki, "Stability analysis of learning algorithms for blind source separation", Neural Networks, vol.10, no8, pp.1345-1351, 1997.
- [3] S.Amari, "Natural Gradient Works Efficiently in Learning", Neural Computation, vol.10, no2, pp251-275, Feb, 1998
- [4] A. J. Bell, T. j. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution", Neural Computation, 7, pp1129-1159, 1995.
- [5] K.Nakayama, A.Hirano, M.nitta, "A Constraint Learning Algorithm for Blind Separation", International Joint Conference on NN, July 24-27, 2000
- [6] H. Mathis, S. C. Douglas, "On optimal and universal nonlinearities for blind signal separation", IEEE Proc.ICASSP'01, MULT-P3.3, ay, 2001.2
- [7] K. Nakayama, A. Hirano and T. Sakai, "An Adaptive Nonlinear Function Controlled by Kurtosis for Blind Source Separation", Proc. IJCNN'02, May 2002.
- [8] Khurram Waheed, Fathi M.Salem, "Algebraic Overcomplete Independent Component Analysis" ICA2003, April 2003, Nara, Japan
- [9] Andreas Jung, Fabian Theis, Carlos Puntonet, Elmar Lang, "Fastgeo A Histogram Based Approach To Linear Geometric Ica", 2003
- [10] Michael S. Lewicki, Terrence J. Sejnowski, "Learning Overcomplete Representations", Neural Computation, 2001
- [11] A. Jung, F.J.Theis, C.G.Puntonet, and E.W.Lang, "Fastgeo - A Histogram Based Approach to Linear Geometric ICA", Proc. of ICA 2001, pp.418-423, 2001
- [12] A.de Cheveigne and H.Kawahara, "YIN, a fundamental frequency estimator for speech and music," J.Acoust. Soc. Am., 2001
- [13] Brent, R.P.1973, Algorithms for Minimization without Derivatives (Englewood Cliffs, NJ: PrenticeHall), Chapter 5. [1]