

オーバーコンプリート・ブラインド信号源分離の フィードバック形構成とその学習アルゴリズム

加藤 治雄[†] 中山 謙二^{††} 平野 晃宏^{††}

[†] 金沢大学大学院 自然科学研究科 電子情報工学専攻

^{††} 金沢大学大学院 自然科学研究科 電子情報科学専攻

E-mail: ^{††}nakayama@t.kanazawa-u.ac.jp

あらまし 本稿では、信号源の数がセンサ数より多いオーバーコンプリート・ブラインド信号源分離においてフィードバック形構成法と分離回路の学習アルゴリズムを提案する。まず、1巡目で信号源の分離を行う。ここでは、一つの信号源が複数の出力に含まれないことを分離の条件とする。このための学習法を提案している。センサ数を信号源数の約半分以上とすることにより、1巡目の信号源分離で少なくとも1個の出力に単一信号源を分離できる。この出力を単一信号源の特徴を利用して検出する。更に、この出力をフィードバックして観測信号からキャンセルすることにより、等価的に信号源の数を低減する。当該出力と混合過程の情報を使ってキャンセルする他、観測信号と当該出力のヒストグラムを使ってキャンセルする方法を組み合わせることにより、条件不足の問題を解消する。2巡目では、観測信号に含まれる信号源が1個少ない状態で1巡目と同じ処理を行う。このように、提案法では、観測信号における信号源の数を1個ずつ減らしながら信号源分離を繰り返す。信号源として音声を用いたシミュレーションにより、従来法との比較を行い、提案方法の有効性を確認している。

キーワード 信号源分離, オーバーコンプリート, フィードバック, 音声信号, 学習アルゴリズム, 基本周期, ヒストグラム

A Feedback Approach and Its Learning Algorithm for Overcomplete Blind Source Separation

Haruo KATOU[†], Kenji NAKAYAMA^{††}, and Akihiro HIRANO^{††}

[†] Division of Electrical and Computer Engineering

^{††} Division of Electrical Engineering and Computer Science

Graduate School of Natural Science and Technology, Kanazawa Univ.

E-mail: ^{††}nakayama@t.kanazawa-u.ac.jp

Abstract A feedback structure and its learning algorithm are proposed for overcomplete blind source separation, where the number of sources is larger than that of sensors. In the first phase, the signal sources are separated so as to satisfy the condition, under which one signal source is not included in different outputs. By setting the number of the sensors to be more than a half of the sources, at least one output includes a single source. This output is selected based on property of a single source. Furthermore, this output is fed back and cancelled from observed signals. The selected single source, information about a mixing process and histogram of the observations and the separated source are used for this cancellation. In the second phase, the same process is carried out by using the modified observations, in which the source, separated in the first phase, is cancelled. Like this, in the proposed method, the source separation is repeated by reducing the number of equivalent sources. Simulations using speech signals demonstrate usefulness of the proposed method compared to the conventional methods.

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

1. まえがき

近年、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場合も多い。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信号源分離 (Blind Source Separation: BSS) が重要となる [1],[2],[5],[7]。

実際のアプリケーションでは、信号源の数を予測することが難しい場合がある。このような状況下では、一般的にセンサ数は信号源数とは異なる。特に、信号源数がセンサ数より多い場合の BSS はオーバーコンプリート BSS と呼ばれ、信号源数に対して観測信号の情報が不足することになり信号源分離が難しくなる。信号源分離を行うためには、観測信号以外に信号源に関する他の情報を追加する必要がある。

新たな情報として主に観測信号のヒストグラムを用いるいくつかの方法が既に提案されている [12],[13],[14]。しかし、十分な分離特性は得られていない。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、出力が一度に M 個の信号源となるように分離する。これらの方法は大きく分類して、混合行列推定過程と信号源復元過程からなり、未知の混合行列を推定する精度により、分離特性が大きく左右される。

本稿では、フィードバック形ネットワーク構成と分離回路の学習法を提案する。単一分離された信号源をフィードバックして観測信号からキャンセルする方法である。1巡目の信号源分離と単一信号源を分離する出力の検出方法は以前に提案した [16]。本稿では、分離した単一信号源をフィードバックにより観測信号からキャンセルする方法を提案する。音声信号源を用いたシミュレーションを行い、分離特性を従来法と比較する。

2. フィードバック形オーバーコンプリート BSS

2.1 ネットワーク構成

信号源数を 3、センサ数を 2 としたときのフィードバック形オーバーコンプリート BSS のブロック図を図 1 に示す。

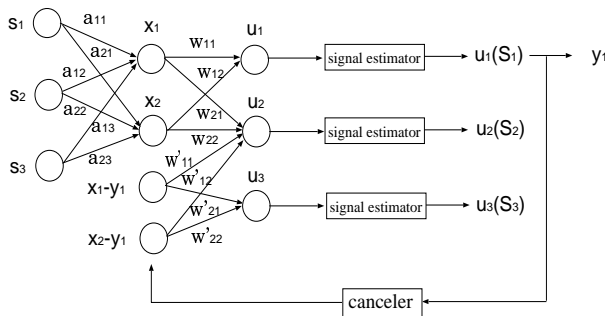


図 1 フィードバック形オーバーコンプリート BSS の回路構成

信号源数を M 、センサ数を N とすると、センサへの入力 \mathbf{x} と分離回路の出力 \mathbf{u} はそれぞれ以下のように表される。 h_{ji}

は信号源から分離回路出力までの伝達特性を表す。

$$\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1M} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NM} \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_M \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \\ = \begin{bmatrix} h_{11} & \cdots & h_{1M} \\ \vdots & \ddots & \vdots \\ h_{N1} & \cdots & h_{NM} \end{bmatrix} \begin{bmatrix} s_1 \\ \vdots \\ s_M \end{bmatrix} \quad (2)$$

2.2 必要なセンサ数

センサ数は「分離回路における少なくとも 1 個の出力に単一の信号源を分離する」条件から決めることにする。BSS の学習は出力信号が統計的に独立になるように学習されるため、学習が進むと、1 個の信号源は 1 個の出力に含まれると仮定する。この仮定の精度に関しては後で議論する。このとき、上記の条件を満たすためのセンサ数は

$$N \geq \left\lceil \frac{M}{2} + 1 \right\rceil \quad (3)$$

となる。演算子 $\lceil X \rceil$ は X を超えない整数を表す。

2.3 フィードバックによる等価的な信号源の削減

1 例として、 u_1 は s_1 を含み、 u_2 は s_2, s_3 を含む場合を考える。全ての出力に対して単一音声か否かの検出を signal estimator のブロックで行う。 u_1 が単一音声のみを含むので、これが検出される。 u_1 は最終的な出力 y_1 となる。同時に、 y_1 は分離回路の入力にフィードバックされ、入力信号 (= 観測信号) からその成分を差し引く。すなわち、観測信号 x_1, x_2 に含まれる s_1 成分をキャンセルすることになる。結果として、観測信号 x_1, x_2 は信号源 s_2, s_3 成分のみを含むことになる。 s_1 成分がキャンセルされた観測信号を x'_1, x'_2 と表す。これらは w'_{ji} と表す別の分離回路で分離される。この例では、この段階でセンサ数と信号源数が等しくなるので、2 巡目では従来の信号源分離と同じになる。

3. 一巡目の信号源分離

3.1 信号源分離の理論的境界

図 1 に示したネットワークに基づいて信号源分離の理論的境界を検討する。式 (1), (2) より、 u_1 において s_2, s_3 を含まない条件は

$$a_{12}w_{11} + a_{22}w_{12} = 0 \quad (4)$$

$$a_{13}w_{11} + a_{23}w_{12} = 0 \quad (5)$$

これは、ベクトル $\mathbf{a}_2 = [a_{12}, a_{22}]^T$ と $\mathbf{a}_3 = [a_{13}, a_{23}]^T$ に直交する重みベクトル $\mathbf{w}_1 = [w_{11}, w_{12}]^T$ を求める問題となる。従って、ベクトル \mathbf{a}_2 と \mathbf{a}_3 が近い方が上記の条件が良く成り立つことになる。

一方、 u_2 において、 s_1 を含まない条件は

$$a_{11}w_{21} + a_{21}w_{22} = 0 \quad (6)$$

となり、これを満たす重みベクトル \mathbf{w}_2 は必ず存在する。

信号源が5個、センサが3個の場合は全ての出力において複数の信号源を抑制することになるため、性能解析は上記の u_1 の場合と同じになる。

2.2節において、少なくとも1個の出力に単一信号源を分離することを仮定したが、この精度は式(4),(5)の精度に依存する。

3.2 最小2乗法による理論的境界の計算

混合過程は瞬時混合とし、 a_{ji} は時間遅れを含まないものとする。観測信号、分離回路の出力は式(1),(2)として表せる。これを行列とベクトルで表すと以下ようになる。

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (7)$$

$$\mathbf{u} = \mathbf{W}\mathbf{x} = \mathbf{W}\mathbf{A}\mathbf{s} = \mathbf{H}\mathbf{s} \quad (8)$$

$$\mathbf{H} = \mathbf{W}\mathbf{A} \quad (9)$$

ここで、 u_1 に s_1 成分が、 u_2 に s_2, s_3 成分が分離されるものと仮定し、出力 $u_1(n), u_2(n)$ の信号対干渉信号比(SIR_i)を以下のように評価する。

$$SIR_1 = 10 \log_{10} \left(\frac{h_{11}^2}{h_{12}^2 + h_{13}^2} \right) \quad [dB] \quad (10)$$

$$SIR_2 = 10 \log_{10} \left(\frac{h_{22}^2 + h_{23}^2}{h_{21}^2} \right) \quad [dB] \quad (11)$$

SIR_1, SIR_2 を最大化するためには、次の条件を満たさなければならない。

$$h_{11}^2 = (w_{11}a_{11} + w_{12}a_{21})^2 \quad \text{to be constant} \quad (12)$$

$$h_{12}^2 = (w_{11}a_{12} + w_{12}a_{22})^2 \quad \text{to be minimized} \quad (13)$$

$$h_{13}^2 = (w_{11}a_{13} + w_{12}a_{23})^2 \quad \text{to be minimized} \quad (14)$$

$$h_{21}^2 = (w_{21}a_{11} + w_{22}a_{21})^2 \quad \text{to be minimized} \quad (15)$$

$$h_{22}^2 = (w_{21}a_{12} + w_{22}a_{22})^2 \quad \text{to be constant} \quad (16)$$

$$h_{23}^2 = (w_{21}a_{13} + w_{22}a_{23})^2 \quad \text{to be constant} \quad (17)$$

分離行列の重み w_{kj} を上記の条件を満足するように決めることにより、 SIR_1, SIR_2 の理論的境界を導出することができる。また、 h_{22}^2 と h_{23}^2 のバランスも重要であり、それを以下のように評価する。

$$R = \frac{|h_{23}|}{|h_{22}|} \quad (18)$$

信号源分離として、 R は1に近いことが望ましい。

3.3 学習アルゴリズム

従来法による学習アルゴリズムは、基本的に s_1 と $s_2 + s_3$ に分離するようなグループ分離にも適用することができる。信号源 s_1 と $s_2 + s_3$ 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ことから、分離回路の学習法は、相互情報量を式(19)のように誤差関数とする自然勾配法を用いる[9]。

$$l(\mathbf{x}, \mathbf{W}) = -\log |\det(\mathbf{W})| - \sum_{k=1}^M \log p_k(u_k) \quad (19)$$

この時の分離行列の更新分は出力 \mathbf{u} の各要素間の相互情報量を最小化するように更新される。分離回路の更新式は

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \eta[\mathbf{A}(n) - \langle \varphi(\mathbf{u}(n)) \mathbf{u}^T(n) \rangle] \mathbf{W}(n) \quad (20)$$

$$\varphi(\mathbf{u}(n)) = \frac{p'(\mathbf{u}(n))}{p(\mathbf{u}(n))} = \frac{1 - \exp(-5\mathbf{u}(n))}{1 + \exp(-5\mathbf{u}(n))} \quad (21)$$

となる。ここで、 $\langle \cdot \rangle$ は時間平均、 p_k は u_k の確率密度関数を表す。また、 p の1次微分である p' も確率密度関数であり、学習過程を安定化するために、非線形関数 φ は式(21)を満たす必要がある[8]。このような目的に対して、いくつかの方法が既に提案されている[3],[6],[11]。本稿では、非線形関数 φ は式(22)に従って制御する。

$$\varphi(\mathbf{u}(n)) = a \tanh(\mathbf{u}(n)) + (1 - a)\mathbf{u}(n)^3 \quad (22)$$

$$a = \frac{1 - \exp(-2.1\kappa_4 - 2.5)}{1 + \exp(-2.1\kappa_4 - 2.5)} \quad (23)$$

ここで、 κ_4 は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 である。

3.4 観測信号のヒストグラムによる学習制御

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[14]。このアルゴリズムを信号源分離の学習法に取り入れる。

ここでは簡単のために2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 $\mathbf{z} = \pi(\mathbf{x})$ とする。 \mathbf{z} のヒストグラムを用いることで直接的に観測信号の投影上の位置を以下のように求めることができる。

$$\psi \rightarrow \frac{l_1(\psi) + l_2(\psi)}{2} - \left(\psi + \frac{\pi}{2} \right) \quad (24)$$

$$l_i(\psi) = f_z^{-1} \left(\frac{f_z(\psi + i\frac{\pi}{2}) + f_z(\psi + (i-1)\frac{\pi}{2})}{2} \right) \quad (25)$$

f_z は \mathbf{z} 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。このようにヒストグラムを用いて観測信号 $\mathbf{x}(n)$ を投影したものを $\mathbf{v}(n)$ とする。図2に $\mathbf{z}(n)$ のヒストグラムの1例を示す。

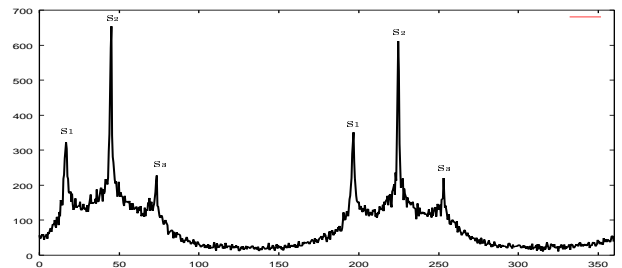


図2 観測信号を超球上に投影した時のヒストグラムの一例

学習アルゴリズムは以下のように修正する。

$$\mathbf{w}^*(n) = \operatorname{argmax}_{\mathbf{w}_j} |\mathbf{v}^T(n) \mathbf{w}_j(n)| \quad (26)$$

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

ここで、 pr は更新された分離行列を再び超球上に投影する関数である。また、学習アルゴリズムとして、観測信号ベクトルに最も接近している重みベクトルを更新する。

3.5 シミュレーションと考察

シミュレーションで用いた信号源は男声 2 個，女声 1 個とする。また，混合過程の係数 (混合行列) は， $a_{11} = a_{23} = 1$ ， $a_{13} = a_{21} = 0.3$ ， $a_{12} + a_{22} = 1.4$ とする。ここで， a_{12} と a_{22} の比を以下のように α として定義する。

$$\alpha = \frac{a_{12}}{a_{22}}, \quad 0 < \alpha \leq 1 \quad (27)$$

ここで， $\alpha = 1$ の時，信号源 s_2 の位置は 2 個のセンサの中間にあり， $u_2(n)$ に s_2 成分を出力することは非常に困難となる。一方で， α が小さい時，つまり信号源 $s_2(n)$ の位置が 2 個目のセンサ ($x_2(n)$) に近い場合， $s_2(n)$ 成分を $u_2(n)$ に分離することは容易である。

α を横軸とした時の， SIR_1 ， SIR_2 ， R のシミュレーション結果をそれぞれ図 3，図 4，図 5 に示す。

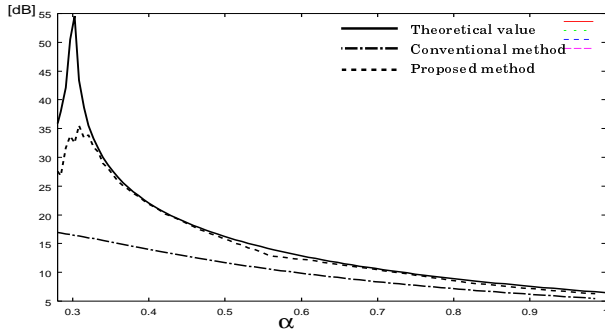


図 3 式 (10) で定義された SIR_1 (3-2-2BSS)

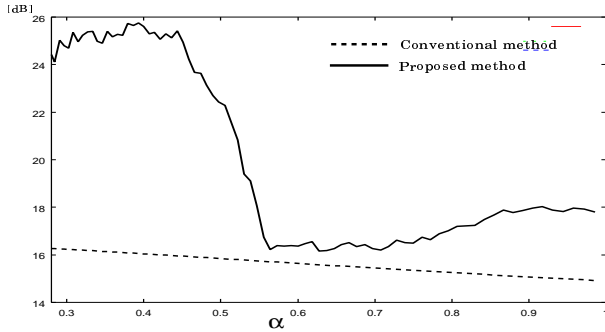


図 4 式 (11) で定義された SIR_2 (3-2-2BSS)

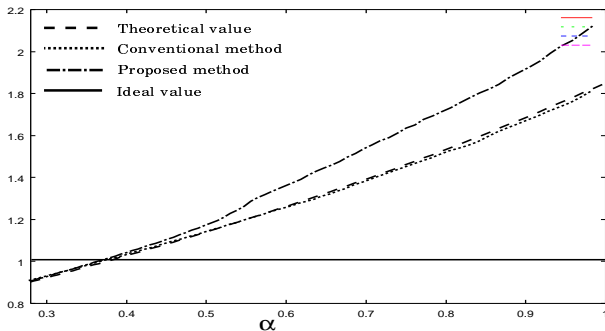


図 5 式 (18) で定義された R (3-2-2BSS)

従来法は式 (20)，(21) による学習法での結果，提案法は式 (22) から式 (26) を考慮した学習法での結果を意味している。提案法での結果は従来法での結果と比較しても良好な分離特性が得られており， SIR_1 は理論的限界値のカーブとほぼ等しいことが分かる。次に， SIR_2 は α に依存するといえ

るが，16dB以上の結果が得られており，良好な分離となっている。ここで， SIR_2 の理論的限界値は無量大であるので省略してある。また，バランス R は理想的には 1 に近い事が望ましい。しかし，理論的限界値は 1 ではなく，多少のアンバランスさは避けられないと考えられる。提案法による R は α が大きくなるにつれてわずかに増加している。

4. 単一音声識別

信号源が音声の場合について考える。1 巡目の信号群分離後，2 個の出力 $u_1(n)$ ， $u_2(n)$ が得られる。仮に 1 巡目の信号源分離が完全なものとすると，出力のうち 1 個は単一音声のみを含んでいる。次のステップとして，単一音声を有する出力を選択する方法について検討する。

本稿では，ピッチ周波数である基本周期に基づいて，出力を分析する [15]。各音声はそれぞれの基本周期をもっており，単一音声といくつかの音声混合された複数音声の違いを識別することができる。

基本周期 T は以下の式を用いて推定される。

$$d_t(\tau) = \sum_{j=1}^D (x(j) - x(j+T))^2 \quad (28)$$

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

なお，基本周期 T は $d'_t(\tau)$ の平均値が最小になるように決定される [16]。

単一音声と複数音声によるシミュレーションを行ない，単一音声識別の動作評価を行なう。使用する音声は 3 章で用いた音声とし，複数音声はこれらの音声を複数混合して作成するものとする。基本周期の推定を各音声で行なった結果，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 $d'(T)$ の平均値の値で評価する。表 1 はそのシミュレーション結果である。

表 1 基本周期 T における $d'(T)$ の平均値の比較

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

単一音声の場合，表 1 から分かるように $d'(T)$ の平均値は小さな値となる。一方，複数音声の場合，単一音声と比較すると，その平均値は大きな値となる。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 $d'(T)$ の平均値を比較すれば識別が可能となる。

この方法を分離回路の出力に適用する。表 2 はそのシミュレーション結果である。

表 2 1 巡目の信号源分離後の出力での $d'(T)$ の平均値

α	$\alpha = 0.4285$	$\alpha = 0.5555$	$\alpha = 0.7073$			
出力	u_1	u_2	u_1	u_2	u_1	u_2
平均値	0.1685	0.2273	0.1987	0.2270	0.2249	0.2227

表 2 から， α が小さい場合では出力 u_1 は出力 u_2 と比較して， $d'(T)$ の平均値の値は小さいことが分かる。しかし， α 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 u_1 と出力 u_2 の $d'(T)$ の平均値の値はほぼ等しくなり，それ

が単一音声か複数音声かを見分けるのは難しい。また表1から、信号源により $d'(T)$ の平均値の値に大小関係があると分かるので、信号源の性質も多少検出結果に影響していると考えられる。

5. フィードバックによる音声のキャンセル

5.1 混合行列の推定

基本的に混合行列 A は未知である。しかし、観測信号のヒストグラムを用いれば、推定は可能となる [14]。混合行列は以下のように推定される。

$$\hat{a}_i(n+1) = \hat{a}_i(n) + \eta \cdot \text{sgn}(v_i(n) - \hat{a}_i(n)) \quad (29)$$

この式で、 $\hat{a}_i(n)$ は $a_i(n)$ の推定結果を表している。ここで、図6は観測信号 x の分布の一例である。

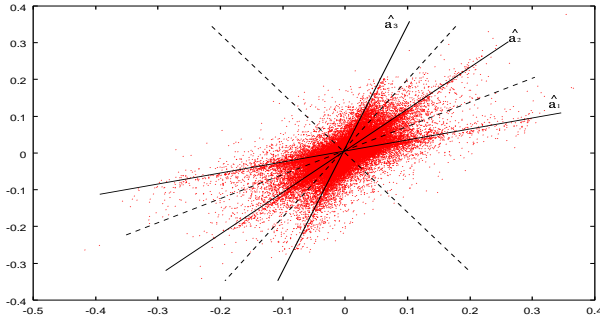


図6 $\alpha = 0.428$ の時の観測信号 x の分布。横軸: $x_1(n)$, 縦軸: $x_2(n)$

横軸、縦軸はそれぞれ $x_1(n), x_2(n)$ を表している。なお、 $\hat{a}_i(n)$ の初期値は図6での実線方向成分(図2のヒストグラムが凸な部分)とする。 $\alpha = 0.428$ の時、未知の混合行列 A を推定した混合行列を A' とすると、

$$A = \begin{bmatrix} 0.957 & 0.393 & 0.287 \\ 0.287 & 0.919 & 0.957 \end{bmatrix}$$

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

となり、 α に依存するが、良好な推定精度が得られている事が分かる。 α が1に近づくにつれて、観測信号の分布も鮮明な分布になり、混合行列の推定精度は良好になる。

5.2 単一音声のキャンセル

5.2.1 直接的な音声除去

混合行列が正確に推定でき、 $s_1(n)$ が単一音声として $u_1(n)$ に含まれている状況を仮定する。 $u_1(n)$ は図1の最終的な出力である $y_1(n)$ として選ばれる。ここで、 $x_1(n)$ と $y_1(n)$ は以下のように表現できる。

$$x_1(n) = a_{11}s_1(n) + a_{12}s_2(n) + a_{13}s_3(n) \quad (30)$$

$$y_1(n) = (a_{11}w_{11} + a_{21}w_{12})s_1(n) \quad (31)$$

このような関係から、次式により $s_1(n)$ を $x_1(n)$ からキャンセルすることができる。

$$x'_1(n) = x_1(n) - \frac{a_{11}y_1(n)}{a_{11}w_{11} + a_{21}w_{12}} \quad (32)$$

ここで、理想的には $x'_1(n)$ は信号源 $s_1(n)$ 成分を含んでいないことになるが、混合行列の推定精度により $s_1(n)$ 成分が残る場合がある。

5.2.2 ヒストグラムに基づいた音声除去

$y_1(n)$ に含まれていた $s_1(n)$ 成分を観測信号 $x_1(n), x_2(n)$ のヒストグラムに基づいて $x_2(n)$ からキャンセルする。音声の確率密度関数は Super-Gaussian となることを考慮して、消したい成分である $s_1(n)$ 成分の影響がよく表れている部分のヒストグラムをキャンセルする。キャンセル後のヒストグラムが0となっている部分は $x_2(n)$ の振幅を0にする。また、ヒストグラムが変化している部分に対しては、ヒストグラムの値に沿うように $x_2(n)$ のサンプルのうち、振幅を0にする部分をランダムに決定する。このようにして得られた観測信号を $x'_2(n)$ とする。

$x_2(n)$ のサンプルを0にすることで、 $x'_2(n)$ の周波数成分全体に歪みが生じる。元々、音声の周波数成分は低周波成分が大きく、高周波成分は小さい。つまり、ヒストグラムを変化させたことで、高周波成分に対する歪みの影響は大きいと考えられる。この問題に対して、式(32)により得られた $x'_1(n)$ の高周波成分を $x'_2(n)$ の高周波成分の代わりに用いることにする。

5.3 2巡目の信号源分離

フィードバック後、観測信号 $x_1(n), x_2(n)$ は $x'_1(n), x'_2(n)$ に置き換えられ、分離行列は式(20),(22)によって更新される。例として、信号源 $s_2(n)$ は $u_2(n)$ に、信号源 $s_3(n)$ は $u_3(n)$ に分離される。1巡目の信号源分離後の $u_1(n)$ に対する $x_1(n), x_2(n)$ の分離行列の重みは固定される。2巡目の信号源分離では、 $u_2(n), u_3(n)$ に対する $x'_1(n), x'_2(n)$ の重みは再び調整される。そして、学習終了後の出力 $u_2(n), u_3(n)$ に対する基本周期を求めて、もし条件を満たすようであれば、それらは最終的な出力 $y_2(n), y_3(n)$ として選ばれる。

5.4 シミュレーションとその考察

分離行列の初期値は以下のように単位行列とする。

$$W_{initial} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (33)$$

評価として、信号対干渉信号比 SIR'_i を以下のように定義する。

$$SIR'_i = 10 \log_{10} \left(\frac{\sum s_i^2(n)}{\sum (s_i(n) - y_i(n))^2} \right) \quad [\text{dB}] \quad (34)$$

ここで、 $s_i(n)$ と $y_i(n)$ を正規化して、評価する。図7に s_1 成分をキャンセルした後のヒストグラムを示す。

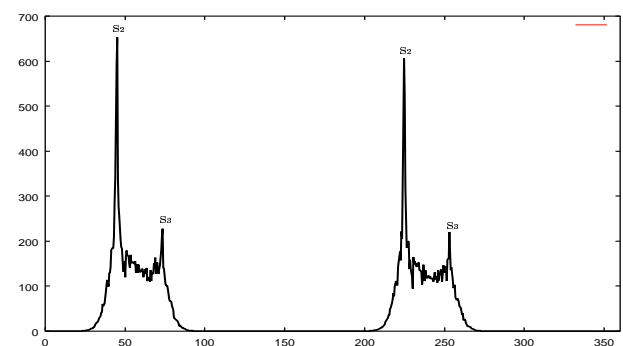


図7 フィードバックで s_1 成分をキャンセルしたヒストグラム

図2のグラフと比較すると、信号源 s_1 成分の影響が強い部分

のヒストグラムの値が低減されていることが分かる。次に、 SIR'_i の評価結果をそれぞれ図8, 図9, 図10に示す。

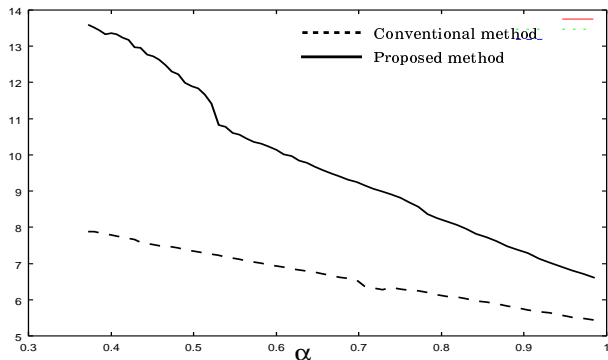


図8 式(34)で定義された SIR'_1 (3-2-2BSS)

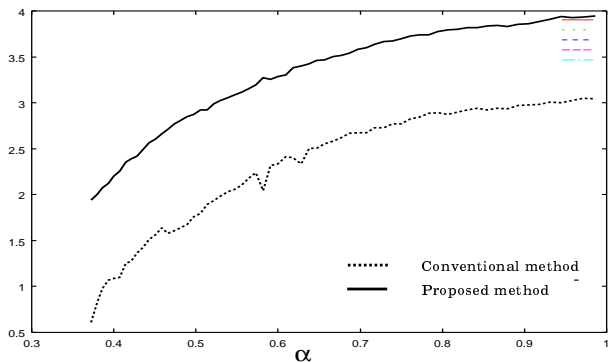


図9 式(34)で定義された SIR'_2 (3-2-2BSS)

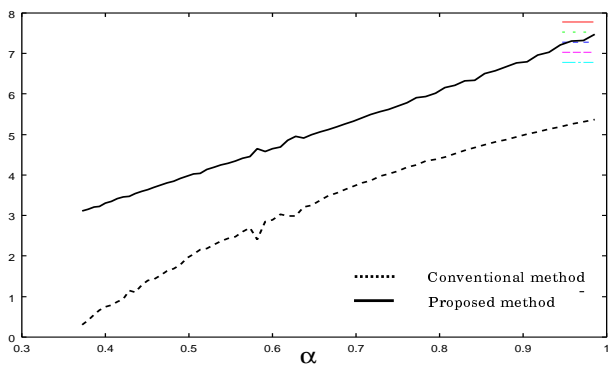


図10 式(34)で定義された SIR'_3 (3-2-2BSS)

従来法とは、観測信号のヒストグラムに基づいた Shortest-Path method[14] による結果を表している。

SIR'_1 は SIR_1 と等価であり、式(34)により評価法を変化させただけである。 SIR'_1 は α が大きくなるにつれて、減少することが分かる。1巡目の信号源分離では、 $s_1(n)$ と $s_2(n) + s_3(n)$ のグループに分離するので、 α が小さければ良好な分離となる。一方、 SIR'_2, SIR'_3 は α の増加とともに大きくなる。その理由として、2巡目の信号源分離は $s_2(n)$ と $s_3(n)$ に分離するので、 α が大きくなるにつれて、 $s_2(n)$ と $s_3(n)$ の位置は遠くなり、分離しやすくなるためである。提案法と従来法を SIR'_i で比較すると、1~3dB改善していることが分かる。

6. まとめ

本稿では、オーバーコンプリート形 BSS に対するフィー

ドバック構成を提案した。センサ数を(信号源数/2+1)を超えない整数とすることにより、少なくとも1つの出力に単一の信号源を分離できる。そのために、相互情報量とヒストグラムを用いる学習法を提案した。分離された単一信号源は観測信号にフィードバックされ、キャンセルされる。このために、波形除去とヒストグラムに基づく除去という2種類のキャンセル法を提案した。その結果、従来法と比較して、分離性能(信号対干渉信号比)において1~3dB程度改善していることを確認した。

文献

- [1] C.Jutten and Jeanny Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, 24, pp.1-10, 1991.
- [2] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures," *Signal Processing*, vol.45, no.2, pp.209-229, March 1995.
- [3] K.Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by kurtosis for blind source separation," *IEEE INNS, Proc. IJCNN'2002, Honolulu*, pp.1234-1239, May 2002.
- [4] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm for convolutive blind source separation with transmission delay constraint," *IEEE INNS, Proc. IJCNN'2002, Honolulu*, pp.1287-1292, May 2002.
- [5] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm with adaptive exponential stepsize for blind source separation of convolutive mixtures with reverberations," *IEEE INNS, Proc. IJCNN'2003 July 2003*.
- [6] K. Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output pdf for blind source separation," *Proc. ICA2003, Nara, Japan*, pp.427-432, April 2003.
- [7] A.Horita, K.Nakayama, A.Hirano and Y.Dejima, "Analysis of signal separation and signal distortion in feed-forward and feedback blind source separation based on source spectra," *IJCNN'05, Montreal*, Aug. 2005.
- [8] 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.
- [9] S.Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol.10, no2, pp.251-275, Feb. 1998.
- [10] A.J.Bell, T.J.Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol.7, pp.1129-1159, 1995.
- [11] H. Mathis, S. C. Douglas, "On optimal and universal nonlinearities for blind signal separation," *IEEE Proc. ICASSP2001, MULT-P3.3, Salt Lake City, Feb. 2001*.
- [12] K.Waheed and F.M.Salem, "Algebraic overcomplete independent component analysis," *ICA2003, Nara, Japan*, April 2003.
- [13] M.S.Lewicki and T.J.Sejnowski, "Learning overcomplete representations," *Neural Computation*, 2001.
- [14] A. Jung, F.J.Theis, C.G.Puntonet and E.W.Lang, "Fast-Geo - A histogram based approach to linear geometric ICA," *Proc. of ICA2001*, pp.418-423, April 2001.
- [15] A.D.Cheveigne and H.Kawahara, "YIN, A fundamental frequency estimator for speech and music," *J.Acoust. Soc. Am.*, 2001. and its learning algorithm for overcomplete blind source separation," *SIP2005, Nagoya, Japan*, September.
- [16] 加藤治雄, 中山謙二, 平野晃宏, "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", *電子情報通信学会, 信学技法, Vol.105 No.293 pp.43-48, SIP2005-92, 2005.09*.