

フィードバック形オーバーコンプリート BSS における学習法の検討

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 信号源数がセンサ数より多いオーバーコンプリート形ブラインド信号源分離 (BSS) で、信号源の完全分離は困難という問題に対して、我々はフィードバック形の回路構成と学習法を既に提案している。まず、1 個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して、1 巡目の信号源分離を行ない、少なくとも 1 個の出力に単一信号源を分離する。この出力を観測信号にフィードバックして、観測信号から単一信号源をキャンセルすることにより、等価的に信号源数を低減する。単一信号源のキャンセル法として、当該出力と推定した混合過程の情報を使ってキャンセルする他、観測信号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。本稿ではさらに、フィードバックにより生じる信号歪みの問題に対して、スペクトルサプレッション法を導入することにより、信号歪みを抑制する。2 巡目の信号源分離では、観測信号に含まれる信号源が 1 個少ない状態で 1 巡目とは異なる学習法により処理を行なう。信号源として音声を用いたシミュレーションにより、提案方法の有効性を確認した。

キーワード ブラインド信号源分離、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、スペクトルサプレッション、ヒストグラム

Improvement of A Learning Algorithm for Feedback Over-Complete BSS

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 were proposed for over-complete 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 the 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 this paper, a spectrum suppression method is introduced for reducing signal distortion caused in the cancellation process based on the histogram. Furthermore, a learning algorithm in the second source separation is also modified. Compared to the conventional method we proposed, the signal to interference is well improved.

Key words Blind source separation, Over-complete, Feedback, Speech signal, Learning algorithm, Spectral Suppression, 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形 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、信号源の完全分離はできない。この条件下で良好な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、十分な分離特性が得られているものはない。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。分離過程は大きく分類して、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。

オーバーコンプリート形 BSS に対して、我々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既に提案している [16],[17]。これは、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。観測信号から信号源を単一分離するための学習法には相互情報量とヒストグラムを用いている。また、1 巡目の信号源分離により分離された単一信号源と複数信号源の識別は基本周期を用いて行なわれる。そして、分離された単一信号源は観測信号からキャンセルされ、観測信号に含まれる信号源を等価的に低減した状態で 2 巡目の信号源分離を行なう。なお、単一信号源のキャンセル法として、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する。

フィードバックの問題として、フィードバック後の観測信号の周波数成分全体に信号歪みが生じる。本稿では、フィードバックによる信号源のキャンセルにおいてスペクトルサブレーション法を導入することにより、信号歪みを抑制する。また、2 巡目の学習法を改良することにより分離性能を向上する。そして、音声信号を用いたシミュレーションを行い、分離特性を評価する。

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

瞬時混合の BSS において、センサへの入力 X と分離回路の出力 U は、未知の混合行列 A と信号源 S により、それぞれ

以下のように表される。

$$X = AS \quad (1)$$

$$U = WX = HS \quad (2)$$

ここで、 W は分離行列、 H は信号源から分離回路出力までの伝達特性を表す。

信号源数を 3、センサ数を 2 とした場合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。

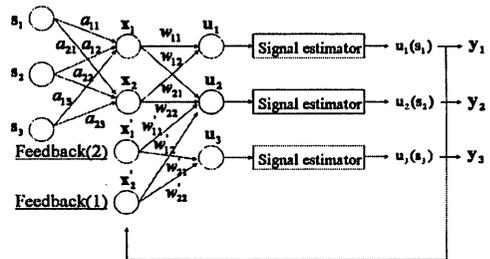


図 1 フィードバック形オーバーコンプリート BSS の回路構成

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。信号源数を M 、センサ数を N とする場合、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

一例として、 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信号源か否かの検出を Signal estimator のブロックで行い、 u_1 が単一信号源のみを含むので、これが検出され、 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) からその信号源成分を差し引く。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り、 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。この観測信号は w'_{ji} と表す別の分離回路で分離される。図 1 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 巡目の信号源分離は従来の構成と同じになり、理論的には全ての信号源を個々に分離することが可能となる。

3. フィードバックによる音声のキャンセル

3.1 直接的な音声除去 (FB(1))

混合行列と 1 巡目の信号源分離後の分離行列を用いて、観測信号から消したい信号源成分を直接的にキャンセルする。ここで、混合行列 A は未知であるので、観測信号のヒストグラムを用いて以下のように推定を行なう [14],[17]。

$$a'_i(n+1) = a'_i(n) + \eta \cdot \text{sgn}(v_i(n) - a'_i(n)) \quad (4)$$

ここで、 v は観測信号を超球上に投影したもので、 A' は A の推定結果を表している。 A' を推定する際の初期値は非常に重要であり、観測信号のヒストグラムが凸な方向成分とする。

未知の混合行列を正確に推定でき、 $s_1(n)$ が単一信号源として $u_1(n)$ に含まれている状況を仮定した場合、 $u_1(n)$ は図1の最終的な出力である $y_1(n)$ として選ばれる。1巡目の信号源分離特性が良好な場合、理想的に $y_1(n)$ には $s_2(n), s_3(n)$ 成分は含まれていないことになる。そこで、 $x_2(n)$ と $y_1(n)$ は以下のように表現できる。

$$x_2(n) = a_{21}s_1(n) + a_{22}s_2(n) + a_{23}s_3(n) \quad (5)$$

$$y_1(n) = (a_{11}w_{11} + a_{21}w_{12})s_1(n) \quad (6)$$

このような関係から、推定した混合行列 A' を用いて、次式により $s_1(n)$ を $x_2(n)$ からキャンセルすることができ、キャンセル後の観測信号を $x'_2(n)$ とする。

$$x'_2(n) = x_2(n) - \frac{a'_{21}y_1(n)}{a'_{11}w_{11} + a'_{21}w_{12}} \quad (7)$$

ここで、理想的には $x'_2(n)$ は信号源 $s_1(n)$ 成分を含んでいないことが望ましいが、未知の混合行列の推定精度や逐次更新される分離回路の変化、分離回路における少なくとも1個の出力に単一信号源を分離する際の性能などの影響を受けて、 $s_1(n)$ 成分が残る場合がある。

FB(1)の問題点として、キャンセルの際に全ての観測信号に対してこの手法を用いることはできない。これは、FB(1)により得られた新たな観測信号は別の観測信号の定数倍になり、このような状況下では信号源分離が非常に難しくなる。しかし、FB(1)でのキャンセル精度は非常に高いので、ある観測信号に対してはFB(1)を使用し、別の観測信号には違うアプローチによって消したい信号源成分をキャンセルする必要がある。

3.2 ヒストグラムに基づいた音声除去 (FB(2))

y_1 に含まれている s_1 成分を観測信号 x_1, x_2 のヒストグラム F に基づいて x_2 からキャンセルする。キャンセル法として、音声の確率密度関数はSuper-Gaussianとなることを考慮して、消したい信号源成分である s_1 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[17]。図2に s_1 成分をキャンセルする前後のヒストグラムを示す。

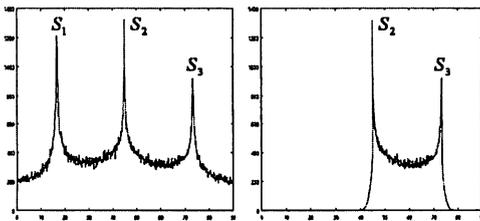


図2 s_1 成分をキャンセルする前後のヒストグラム

図2を見ると、ヒストグラム F のうち、 s_1 成分が大きい部分

にヒストグラムの値が存在しないことが分かる。しかし、必要以上に s_1 成分付近の情報欠如のために、本来あるはずの s_2, s_3 成分の情報も失われてしまうという問題が生じる。

そこで、消したい信号源成分をヒストグラム F からキャンセルする新たな方法を提案する。必要な情報として、1巡目の信号源分離によって得られた単一信号源を含む y_1 と観測信号により求められるヒストグラム \hat{F} を用いる。理想的に y_1 は s_1 成分を多く含むことが考えられ、現実的にもヒストグラム \hat{F} は s_1 成分の部分にピークがあり、 s_2, s_3 成分の情報は少ない。そこで、ヒストグラム \hat{F} を用いれば、ヒストグラム F 中の s_1 成分を低減させ、 s_2, s_3 成分の情報をできるだけ残すことが可能となる。これらの事を考慮に入れて、ヒストグラム F から s_1 成分をキャンセルさせた後のヒストグラム F' を以下のように求める。

$$F'(\theta) = F(\theta) \left(1.0 - \frac{\hat{F}(\theta)}{\hat{F}_{max}} \right)^5 \quad (8)$$

ここで、 \hat{F}_{max} は \hat{F} の最大値であり、 $F'(\theta)$ の値は $\hat{F}(\theta)$ の値が最大の場合0になり、それ以外では常に F の値が低減された値となる。なお、 s_2, s_3 成分のピーク間の情報は重要なので、その場合 F の値は変化させない。このような方法で s_1 成分をキャンセルする前後のヒストグラムを図3に示す。

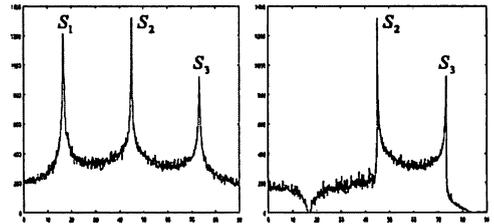


図3 s_1 成分をキャンセルする前後のヒストグラム

ここで、ヒストグラム F と \hat{F} の信号源成分のピークが少しずれるという問題があり、 \hat{F} の位相を F の位相と等しくなるようにシフトさせる必要がある。

FB(2)によって s_1 成分をキャンセルした後のヒストグラム F' とヒストグラム F は位相によって値が大きく異なる。 F に対する F' の変化分は不要な成分なので、ヒストグラム F' に沿うようにその変化分だけ x_1 の振幅を低減させるサンプルを決定する。仮に、ある位相でのヒストグラム F の値が100、 F' の値が40の場合、100サンプルの中から変化分60サンプルをランダムに決定し、そのサンプルの振幅を変化させることになる。なお、振幅を変化させる割合は、 y_1 の振幅に応じて低減させる。このようにして得られた観測信号を $x'_1(n)$ とする。

観測信号 x'_1 は、FB(2)によりサンプル毎に振幅が低減され、 s_1 成分はキャンセルされる。しかし、連続したサンプルの振幅が低減されるなどの原因から、 $x'_1(n)$ の周波数成分全体に歪みが生じる。その信号歪みは本来音声信号にはないもので、音声として聞いた場合に不連続な値となるので雑音のように聞こえてしまう。そこで、スペクトルサブプレッショ

ン法を導入することにより、FB(2)によって生じた信号歪みを抑制し、2巡目の信号源分離の分離特性が良好なものになるよう改善する。

3.3 スペクトルサプレッション法による信号歪み抑制

音声のスペクトル成分は Super-Gaussian という条件下で用いられる Joint MAP 法 [18] により、FB(2) により生じた信号歪みを抑制する。FB(1) 後の観測信号をクリーン音声とし、雑音成分を消したい信号源成分 s_1 とすると、Super-Gaussian の確率密度関数は音声の分散を λ_s として式 (9) で表せられる。

$$p(B) = \frac{\mu^{\tau+1}}{\Gamma(\tau+1)} \frac{B^\tau}{\lambda_s^{\tau+1}} \exp\{-\mu \frac{B}{\lambda_s}\} \quad (9)$$

ただし、式中の B は音声の振幅を表し、 τ と μ は変数で $P(B)$ の形状を決定する。Joint MAP 法では、雑音混入音声 X という条件での音声の振幅 B と偏角 δ における条件付確率 $P(B, \delta|X)$ を最大にするスペクトルゲイン $G(k)$ を以下のように求める。

$$G(k) = m + \sqrt{m^2 + \frac{\tau}{2\gamma(k)}} \quad (10)$$

$$m = \frac{1}{2} - \frac{\mu}{4\sqrt{\gamma(k)\xi(k)}} \quad (11)$$

ここで、 $\xi(k)$ はクリーン音声対雑音比、 $\gamma(k)$ は雑音混入音声対雑音比であり、それぞれ以下のように表される。

$$\xi(k) = \frac{Q(k)^2}{\lambda(k)}, \quad \gamma(k) = \frac{R(k)^2}{\lambda(k)} \quad (12)$$

$Q(k)^2$ は FB(1) 後の観測信号のパワースペクトル、 $R(k)^2$ は FB(2) 後の観測信号のパワースペクトルを用いる。また、 $\lambda(k)$ は本来雑音の分散だが、FB(2) 後の振幅スペクトル - FB(1) 後の振幅スペクトルで求める。分析フレーム長は 512 サンプルのハミング窓を用いており、オーバーラップ長は 256 サンプルとする。なお、式 (10) のスペクトルゲイン $G(k)$ は、 $\xi(k)$ が小さい場合、つまり雑音成分がクリーン音声に対して大きい場合に小さくなり、スペクトル成分は強く補正される。

音声は必ずしも確率密度関数が Super-Gaussian ではないので、確率密度関数の変化に応じた Joint MAP 法のパラメータ設定が重要になる。そこで、FB(2) 後の観測信号の平均値付近の分布を表現する 4 次の統計量 kurtosis の値に応じて、 μ と τ の値を適時変化させ、適切なスペクトルゲインを求める。そして、 $\xi(k)$ が低 SNR 時に残留信号を抑圧し、雑音の少ない強調音声を出力する。また、スペクトルゲインによる過剰抑圧を防止し、強調音声の品質を改善するために、スペクトルゲインの最小値の下限を設定する。

$\xi(k)$ が高 SNR 時に Joint MAP 法を用いる場合、スペクトルゲインは 1 に近いことが望ましいが、スペクトルを過剰に抑圧してしまう場合がある。そこで、フィルタ特性 $Z(k)$ が以下のように表されるウィーナーフィルタを改良したものをを用いる。

$$Z(k) = \frac{P_x(k)}{P_x(k) + P_n(k)} \quad (13)$$

ここで、 $P_x(k)$ はクリア音声のパワースペクトル、 $P_n(k)$ は雑音のパワースペクトルである。この特性を応用して、 $\xi(k)$ が高 SNR 時のスペクトルゲインを Joint MAP 法の代わりに以下のように求める。

$$G(k) = \frac{\xi(k)}{\xi(k) + \delta} \quad (14)$$

なお、 δ は 1 以下の定数である。

4. 2巡目の信号源分離の学習法の検討

4.1 1巡目の信号源分離の学習法

分離回路の学習法は、式 (15) の相互情報量を誤差関数とする自然勾配法を用いる [9]。

$$l(\mathbf{x}, \mathbf{W}) = -\log |\det(\mathbf{W})| - \sum_{k=1}^M \log p_k(u_k) \quad (15)$$

この時の分離回路の更新分は出力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以下のような学習アルゴリズムで更新される [17]。

$$\mathbf{w}^*(n) = \underset{\mathbf{w}_j}{\operatorname{argmax}} |\mathbf{v}^T(n)\mathbf{w}_j(n)| \quad (16)$$

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

ここで、 pr は更新された分離回路を再び超球上に投影する関数、 $\langle \cdot \rangle$ は時間平均、 $\mathbf{v}(n)$ は観測信号のヒストグラムを用いて $\mathbf{x}(n)$ を超球上に投影したものを表す。この学習アルゴリズムでは、観測信号に最も接近している分離回路ベクトル \mathbf{w}^* (勝者ベクトル) のみを更新させる。また、式 (16) 中の非線形関数 φ は学習を安定化させるためのもので、観測信号の kurtosis (κ_4) の値によって以下のように制御する [3],[6]。

$$\varphi(\mathbf{u}(n)) = a \tanh(\mathbf{u}(n)) + (1-a)\mathbf{u}(n)^3 \quad (17)$$

$$a = \frac{1 - \exp(-2.1\kappa_4 - 2.5)}{1 + \exp(-2.1\kappa_4 - 2.5)} \quad (18)$$

4.2 2巡目の信号源分離の学習法

1巡目の信号源分離の学習法を応用して、2巡目の信号源分離の学習法を改良する。まず、式 (16) 中の勝者ベクトルの決定法を少し変更する。1巡目の学習には勝者ベクトル決定の際に $\mathbf{v}(n)$ を用いていたが、2巡目の学習ではヒストグラムを用いず、直接的に $\mathbf{x}(n)$ を超球上に投影したものを $\mathbf{v}(n)$ の代わりに用いる。

理想的にフィードバック後の観測信号には、キャンセルされた信号源 s_1 成分は存在せず、 s_2, s_3 成分のみが残っている状態となる。フィードバック後の観測信号に対する各信号源の等価的な混合行列を $\hat{\mathbf{a}}$ 、2巡目の分離回路を \mathbf{w}' とすると、出力 \mathbf{u}' と信号源 s_2, s_3 成分の関係は以下のように表される。

$$\mathbf{u}'_2 = (\mathbf{w}'_{11}\hat{\mathbf{a}}_{12} + \mathbf{w}'_{12}\hat{\mathbf{a}}_{22})s_2 + (\mathbf{w}'_{11}\hat{\mathbf{a}}_{13} + \mathbf{w}'_{12}\hat{\mathbf{a}}_{23})s_3 \quad (19)$$

$$\mathbf{u}'_3 = (\mathbf{w}'_{21}\hat{\mathbf{a}}_{12} + \mathbf{w}'_{22}\hat{\mathbf{a}}_{22})s_2 + (\mathbf{w}'_{21}\hat{\mathbf{a}}_{13} + \mathbf{w}'_{22}\hat{\mathbf{a}}_{23})s_3 \quad (20)$$

仮に、 \mathbf{u}'_2 に s_2 成分を、 \mathbf{u}'_3 に s_3 成分を分離する場合、

$$\mathbf{w}'_{11}\hat{\mathbf{a}}_{13} + \mathbf{w}'_{12}\hat{\mathbf{a}}_{23} = 0, \quad \mathbf{w}'_{21}\hat{\mathbf{a}}_{12} + \mathbf{w}'_{22}\hat{\mathbf{a}}_{22} = 0 \quad (21)$$

が成り立つ。式(21)から、以下のような分離行列の関係を導出できる。

$$w'_{11} = -\frac{w'_{12}\hat{a}_{23}}{\hat{a}_{13}}, \quad w'_{21} = -\frac{w'_{22}\hat{a}_{22}}{\hat{a}_{12}} \quad (22)$$

式(22)の関係を学習に適用することは、 s_2, s_3 成分を分離するのに有効である。しかし、 \hat{a} が未知という問題がある。そこで、 \hat{a} をフィードバック後の観測信号のヒストグラムから求める。ヒストグラムが凸な部分を ϕ_1, ϕ_2 として、近似的ではあるが以下のように \hat{a} を置き換え、これを学習アルゴリズムに取り入れる。

$$\frac{\hat{a}_{23}}{\hat{a}_{13}} = \tan\phi_1, \quad \frac{\hat{a}_{22}}{\hat{a}_{12}} = \tan\phi_2 \quad (23)$$

式(16)中の分離回路の更新分を q とする。フィードバック後の観測信号と2巡目の信号源分離の分離回路から勝者ベクトルを決定し、勝者ベクトルが1列目の場合、以下のように分離回路は更新される。

$$w_{11}(n+1) = (1.0 - \zeta)q_{11}(n) - \zeta \frac{q_{12}}{\tan\phi_2} \quad (24)$$

$$w_{21}(n+1) = (1.0 - \zeta)q_{21}(n) - \zeta \frac{q_{22}}{\tan\phi_2} \quad (25)$$

同様に、勝者ベクトルが2列目の場合、分離回路は以下のように更新される。

$$w_{12}(n+1) = (1.0 - \zeta)q_{12}(n) - \zeta \tan\phi_2 q_{11}(n) \quad (26)$$

$$w_{22}(n+1) = (1.0 - \zeta)q_{22}(n) - \zeta \tan\phi_2 q_{21}(n) \quad (27)$$

なお、 ζ は1以下の定数とする。

4.3 2巡目の信号源分離の初期値

分離回路の初期値は単位行列としていたが、適切な初期値から分離回路の学習を始めるように考慮する。2巡目の信号源分離は信号源2個、センサ2個のモデルなので、その場合に有効な、未知の混合行列を推定する方法を用いる[19]。未知の混合行列の逆行列は分離回路の最適値であるので、その事を考慮して、推定した混合行列の逆行列を分離回路の初期値とする。この混合行列推定の手法は観測信号の相関係数に基づいた方法で、観測信号の相互相関が大きい場合に初期値設定を行なうと、逆に分離特性が劣化する場合がある。この問題に対して、観測信号の相互相関に閾値を設けて、相互相関が閾値以上であれば分離行列の初期値は単位行列、閾値以下であれば、適切な初期値設定を行なうものとする。

5. シミュレーションとその考察

5.1 シミュレーション条件

信号源として、男声2個と女声1個を用い、センサ数は2個のオーバーコンプリート形BSSのモデルでシミュレーションを行なう。混合行列は $a_{11} = a_{23} = 1.0$, $a_{13} = a_{21} = 0.3$, $a_{12} + a_{22} = 1.4$ とし、横軸を $\alpha = a_{12}/a_{22}$ とする。 α が1の時、信号源 s_2 の位置は信号源 s_1 と s_3 の中間にあり、単一信号源と複数信号源に分離する場合、非常に分離しにくい状況となる。一方、 α が小さい時、信号源 s_2 は s_1 から離れ、 s_3 に近い位置にあり、分離特性は良好なものとなる。

5.2 1巡目の信号源分離から単一信号源識別

1巡目の信号源分離では、信号源 s_1 と $s_2 + s_3$ のグループに分離することを目的としている。単一信号源と複数信号源にグループ分離する際の理論的限界を最小二乗法を用いて計算したものと、1巡目の信号源分離の分離特性を比較し、理論的限界に追従する良好な結果となる。また、単一信号源として s_1 が分離でき、その出力と複数の信号源が混ざった出力を識別することにも成功している[15],[17]。そして、識別された s_1 のみを含む出力 w_1 を最終的な出力 y_1 とする。

5.3 フィードバックから学習の初期値設定までの流れ

フィードバックから2巡目の信号源分離の分離回路の初期値設定までの簡単なブロック図を以下に示す。

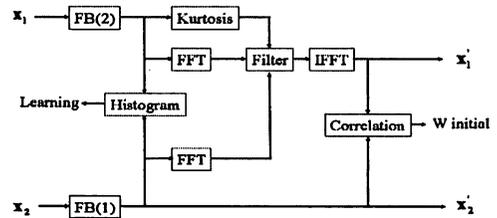


図4 フィードバックから学習の初期値設定までのブロック図

x_1 に対してはFB(2)を、 x_2 に対してはFB(1)を用いて消したい信号源 s_1 成分をキャンセルする。FB(2)によるキャンセルでは信号歪みが生じるので、スペクトルサブプレッション法に基づいた信号歪み除去を行なう。そして、2巡目の信号源分離特性を向上させるために、学習法の改良と適切な初期値設定を行なう。

5.4 評価

1巡目の信号源分離から単一信号源識別、そして3章と4章のフィードバックと2巡目の信号源分離が終了し、最終的な出力をそれぞれ y_2, y_3 とする。評価法として、信号対干渉信号比 SIR_i を以下のように定義する。

$$SIR_i = 10 \log_{10} \left(\frac{\sum s_i^2(n)}{\sum (s_i(n) - y_i(n))^2} \right) \quad [\text{dB}] \quad (28)$$

$\alpha = a_{12}/a_{22}$ を横軸として、出力 y_1 に対して s_1 、出力 y_2 に対して s_2 、出力 y_3 に対して s_3 を式(28)の SIR_i で評価したものを以下に示す。

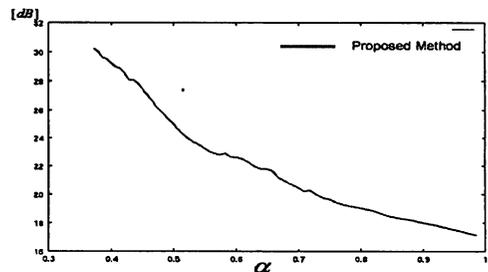


図5 信号対干渉信号比(SIR_1)

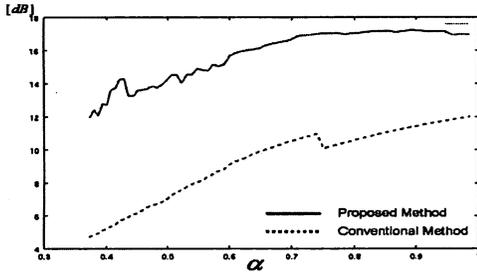


図6 信号対干渉信号比 (SIR_2)

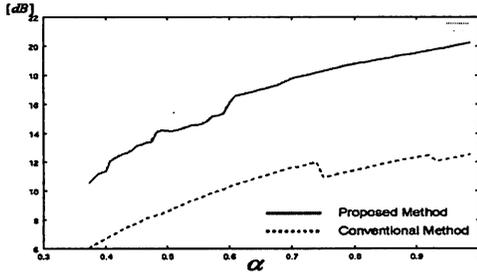


図7 信号対干渉信号比 (SIR_3)

まず、 SIR_1 は α が小さい場合は分離特性が非常に良好であるが、 α が大きくなるにつれて、分離特性が大幅に劣化する。1巡目の信号源分離は信号源 s_1 と $s_2 + s_3$ のグループ分離を行なうので、 α が大きい、つまり s_2 の位置が s_1 に近づくとも分離が難しい状況になる。一方で、 SIR_2 と SIR_3 は SIR_1 に反比例するような形で、 α が大きくなるにつれて分離特性が向上する。これは、信号源が混合される過程で既に各信号源の分離性能が決まるからである。2巡目の信号源分離の学習アルゴリズムでは勝者ベクトルを決定して、勝者ベクトルだけを学習させるが、片方の列に学習が偏ることがある。その問題に対して、分離回路の適切な初期値設定を行なった結果、学習が進まなかった分離行列の列が良好な解に収束した。また、提案法の分離特性は、スペクトルサプレッション法を用いない従来法や、Shortest Path Method[14]での分離特性と比較しても、非常に良好な結果となる。

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形構成法を提案した。1巡目の信号源分離により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、直接的な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でキャンセルされる。ヒストグラムに基づいた除去の際に、キャンセルされた観測信号に信号歪みが生じるが、スペクトルサプレッション法を応用して信号歪みを抑制した。また、2巡目の信号源分離の学習アルゴリズムを改良し、分離回路の初期値を適切な値とすることにより、分離特性が向上した。今後の課題として、フィードバックの際に生じる信号歪みのさらなる除去、2巡目の信号源分離特性のさらなる向上などが挙げられる。

本研究の一部は科学研究費補助金(基礎研究(C)17560335)による。

文 献

- [1] C.Jutten and Jeanny Hérault, "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 non-linear 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 non-linear 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, no.8, pp.1345-1351, 1997.
- [9] S.Amari, "Natural gradient works efficiently in learning," *Neural Computation*, vol.10, no.2, 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ト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.
- [17] 加藤治雄, 中山謙二, 平野晃宏, "オーバーコンプリート・ブライント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", *電子情報通信学会, 信学技法*, Vol.106 No.21 pp.49-54, SIP2006-9, 2006.04.
- [18] T.Lotter, P.Vary "Speech Enhancement by MAP Spectral Amplitude Estimation Using a Super-Gaussian Speech Model", *EURASIP Journal on Applied Signal Processing* 2005:7, 1110-1126.
- [19] J.Van de Laar and P.C.W.Sommen "Closed-form instantaneous blind signal separation based on second order statistics and non-stationarity", 2002.