

2次統計量を用いたブラインド信号源分離アルゴリズムの検討

坂井良広*

An algorithm using second order statistics for the blind source separation

Yoshihiro Sakai*

On the blind source separation problem, the high order statistics is often used for a separation criterion. Such statistics is useful but a solution is not stable anytime, and we need much data to get the solution. To get a stable solution with fewer data, this study employ a criterion using second order statistics. The algorithm is implemented with an iterative estimation method, and evaluated using computer simulation. The performance of the algorithm is, about 5dB convergence per 10000 iterations, and about 30dB reduction of crosstalk, when the source signal is white noise.

1 はじめに

未知の信号源, 未知の混合系の条件下で信号分離を行なうブラインド信号源分離問題では, 分離の規範として高次統計量が多く用いられている [1]. これは分離された信号の独立性の観点から有用であるが, 一方で解の安定性や計算量, データ量の点で問題のあることが指摘されている. これに対し分離の規範として2次統計量(相関)を用いれば, これらの問題の改善が期待される. そこで本稿では2次統計量を用いる分離アルゴリズムの実現方法を検討し, これを用いた場合の基本的な動作確認を行なう.

本稿では混合系として Fig.1 に示す最も基本的な2チャンネルの場合を仮定する. この系は下記に示す周波数応答

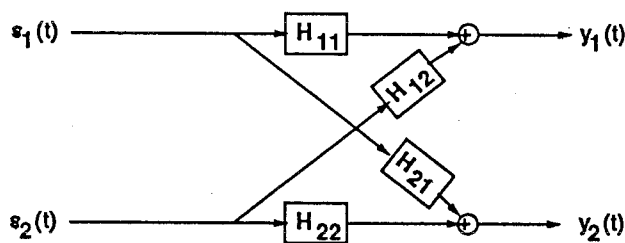


Fig.1 2 信号 2 センサ混合モデル

を持つ.

$$\mathcal{H}(\omega) = \begin{bmatrix} H_{11}(\omega) & H_{12}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) \end{bmatrix} \quad (1)$$

ここで H_{11} および H_{22} は直接経路の伝達関数を表し, H_{12} および H_{21} は混合経路すなわち相互結合効果を表す.

信号強調, すなわち 2 信号分離問題に対しては従来 Widrow らによって提案された手法が用いられている [2].

情報工学科
平成 11 年 8 月 31 日 受付

この手法では Fig.2(a) に示すように, H_{11} と H_{22} は同一であり H_{21} は零であると仮定する. この場合 $s_2(t)$ (妨害信

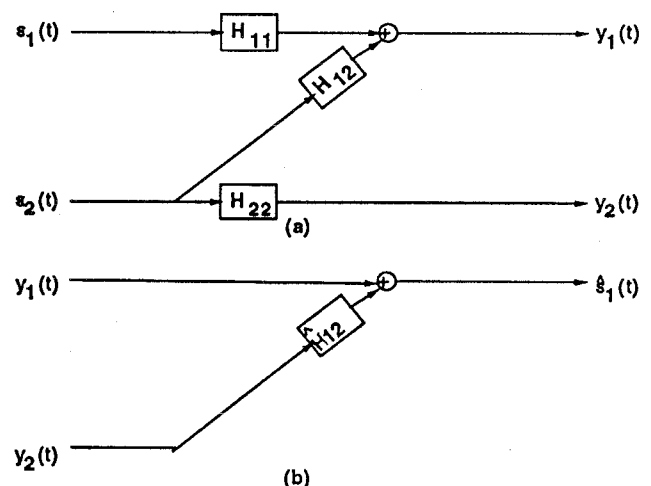


Fig.2 Widrow によるシステム

号) は未知系 H_{12} を通じて第一(主)センサに入力される. H_{12} の入力は第二(参照)センサに入力される信号と同一である. Widrow らは推定された信号の平均パワーを最小化することにより未知系が同定されることを示し, 主センサに入力される妨害信号成分のキャンセルにこの方式を使うことを提案した (Fig.2(b)). 平均パワー最小化は, 主センサ信号中の参照センサ信号成分の二乗の最小化により未知系 H_{12} を推定することに相当し, 最小二乗 (LS) 法と呼ばれる.

以下, 第 2 章では 2 次統計量に基づく分離基準について考察する. 第 3 章では分離基準を達成するアルゴリズムとその評価について述べる. 第 4 章で本検討のまとめを行なう.

2 相関を規範とする信号分離

簡単化のため、 H_{11} と H_{22} が単位伝達関数すなわち $H_{11}(\omega) = H_{22}(\omega) = 1$ である場合を考察する。この仮定は多少限定的であるが、多くの問題に適用可能である。この問題においては所望信号 $s_1(t)$ と $s_2(t)$ は他の信号源が無い時の各センサで測定された信号である。また H_{12} と H_{21} は結合の影響を表す。信号 $y_1(t)$ と $y_2(t)$ は下記の周波数応答を持つ 2×2 線形時不変 (LTI) システムに、 $s_1(t)$ と $s_2(t)$ が入力されたときの出力である。

$$\mathcal{H}(\omega) = \begin{bmatrix} 1 & H_{12}(\omega) \\ H_{21}(\omega) & 1 \end{bmatrix} \quad (2)$$

ここで以下を仮定する。

$$1 - H_{12}(\omega)H_{21}(\omega) \neq 0 \quad \forall \omega \quad (3)$$

もしこの仮定が満たされなければ、 \mathcal{H} は可逆でない。その場合 \mathcal{H} が既知であったとしても、 $s_1(t)$ と $s_2(t)$ は $y_1(t)$ と $y_2(t)$ から回復することができない。残響やマルチパス等の場合で拘束条件がなければ、 \mathcal{H} の逆行列が存在しない場合がある。

$s_1(t)$ と $s_2(t)$ を、無相関な定常ランダム過程で平均が0であると仮定しよう。それらが無相関であるという仮定から、

$$E\{s_1(t)s_2^*(t-\tau)\} = 0 \quad \forall \tau \quad (4)$$

ここで $E\{\cdot\}$ は期待値を表す。また*は複素共役を表す。より一般的な場合 (非零) には共分散を適用する。

系 H_{12} と H_{21} が既知なら $s_1(t)$ と $s_2(t)$ は $y_1(t)$ と $y_2(t)$ から逆フィルタにより推定することができる。しかし、ほとんどの場合 H_{12} と H_{21} は未知であるので、それらを推定する方法または基準が必要となる。

H_{12} と H_{21} の推定 \hat{H}_{12} と \hat{H}_{21} が得られたなら、それを用いて逆フィルタリングを実行する。系の推定が完全であれば、逆フィルタリングによって推定された信号 $\hat{s}_1(t)$ と $\hat{s}_2(t)$ は統計的に無相関である。従って、

$$E\{\hat{s}_1(t)\hat{s}_2^*(t-\tau)\} = 0 \quad \forall \tau \quad (5)$$

信号 $s_1(t)$ と $s_2(t)$ が統計的に無相関であると仮定されるとしても、それは、最小二乗平均誤差基準のような選択された推定基準が統計的に無相関な信号推定を生成するであろうことを含んでいない。ここで必要なことは、信号が無相関であるという仮定を推定基準に変えることである。この基準に基づく解は Widrow らが検討した単純化における LS 解と同じになる。

Fig.3 に示すように、 $\hat{s}_1(t)$ と $\hat{s}_2(t)$ は、入力が $y_1(t)$ と $y_2(t)$ で下記の周波数応答を持つ 2×2 LTI システムの出力である。

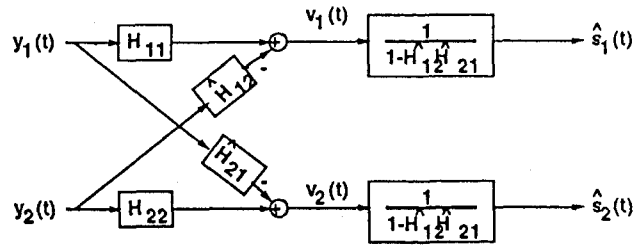


Fig.3 分離回路

$$\hat{\mathcal{H}}^{-1}(\omega) = \frac{1}{1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)} \begin{bmatrix} 1 & -\hat{H}_{12}(\omega) \\ -\hat{H}_{21}(\omega) & 1 \end{bmatrix} \quad (6)$$

ここで、式 (3) のように、下記のことが仮定される。

$$1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega) \neq 0 \quad \forall \omega \quad (7)$$

$\hat{s}_1(t)$ と $\hat{s}_2(t)$ を生成する別構成を Fig.4 に示す。

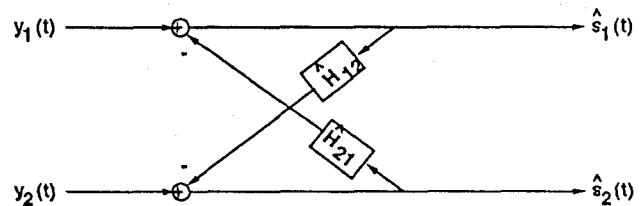


Fig.4 分離回路代替構成

LTI システムの入出力間のパワースペクトルについての関係を用いて、

$$\begin{bmatrix} P_{\hat{s}_1\hat{s}_1}(\omega) & P_{\hat{s}_1\hat{s}_2}(\omega) \\ P_{\hat{s}_2\hat{s}_1}(\omega) & P_{\hat{s}_2\hat{s}_2}(\omega) \end{bmatrix} = \frac{1}{|1 - \hat{H}_{12}(\omega)\hat{H}_{21}(\omega)|^2} \begin{bmatrix} 1 & -\hat{H}_{12}(\omega) \\ -\hat{H}_{21}(\omega) & 1 \end{bmatrix} \begin{bmatrix} P_{y_1y_1}(\omega) & P_{y_1y_2}(\omega) \\ P_{y_2y_1}(\omega) & P_{y_2y_2}(\omega) \end{bmatrix} \begin{bmatrix} 1 & -\hat{H}_{21}^*(\omega) \\ -\hat{H}_{12}^*(\omega) & 1 \end{bmatrix} \quad (8)$$

ここで $P_{\hat{s}_i\hat{s}_j}(\omega)$, $i, j = 1, 2$ は $\hat{S}_1(t)$ と $\hat{S}_2(t)$ の自己およびクロススペクトルである。また $P_{y_iy_j}(\omega)$, $i, j = 1, 2$ は $y_1(t)$ と $y_2(t)$ の自己およびクロススペクトルである。

式 (5) における無相関条件を満たすためには、式 (8) における演算の結果としてすべての ω について $P_{\hat{s}_1\hat{s}_2}(\omega) = 0$ であればよい。したがって、

$$P_{y_1y_2}(\omega) - \hat{H}_{12}(\omega)P_{y_1y_2}(\omega) - \hat{H}_{21}^*(\omega)P_{y_1y_1}(\omega) + \hat{H}_{12}(\omega)\hat{H}_{21}^*(\omega)P_{y_2y_1}(\omega) = 0 \quad (9)$$

ここで、 $P_{y_1y_2}(\omega)$, $i, j = 1, 2$ は実際の観測信号 $y_1(t)$ と $y_2(t)$ から求めることができる。式 (9) を満たす \hat{H}_{12} と \hat{H}_{21} のすべての組み合わせは統計的に無相関な信号推定 $\hat{s}_1(t)$ と $\hat{s}_2(t)$ を実現する。明らかに、この方程式は \hat{H}_{12} と \hat{H}_{21} に

ついでの一義的な解を導かない。 \hat{H}_{21} は任意に選択することができ、この場合 \hat{H}_{12} は下記のように決定される。

$$\hat{H}_{12}(\omega) = \frac{P_{y_1 y_2}(\omega) - \hat{H}_{21}^*(\omega) P_{y_1 y_1}(\omega)}{P_{y_2 y_2}(\omega) - \hat{H}_{21}^*(\omega) P_{y_2 y_1}(\omega)} \quad (10)$$

また逆に、 \hat{H}_{12} を任意に選択することができ、 \hat{H}_{21} は下記のように決定される。

$$\hat{H}_{21}(\omega) = \frac{P_{y_2 y_1}(\omega) - \hat{H}_{12}^*(\omega) P_{y_2 y_2}(\omega)}{P_{y_1 y_1}(\omega) - \hat{H}_{12}^*(\omega) P_{y_1 y_2}(\omega)} \quad (11)$$

特別な場合として、 $\hat{H}_{21} = 0$ とするならば式 (10) は下記のように縮退する。

$$\hat{H}_{12}(\omega) = \frac{P_{y_1 y_2}(\omega)}{P_{y_2 y_2}(\omega)} \quad (12)$$

これは、信号から参照センサへの結合が無いことが仮定している Widrow の LS 解に相当する。したがって LS 解は式 (10) の解に含まれる。LS 解が信号推定値を参照センサ信号から統計的に無相関となるようにすることは既に検証されている。LS 法は多くの事例において効果をあげていが、もし零結合の仮定が満たされなければその性能が大きく低下する。そのため式 (10) は、結合が存在する場合の有効な分離方法を示している。

無相関となる条件を与える式 (9) と真の結合系 H_{12} , H_{21} との関係を検討する。混合系については、

$$\begin{bmatrix} P_{y_1 y_1}(\omega) & P_{y_1 y_2}(\omega) \\ P_{y_2 y_1}(\omega) & P_{y_2 y_2}(\omega) \end{bmatrix} = \begin{bmatrix} 1 & H_{12}(\omega) \\ H_{21}(\omega) & 1 \end{bmatrix} \begin{bmatrix} P_{s_1 s_1}(\omega) & 0 \\ 0 & P_{s_2 s_2}(\omega) \end{bmatrix} \begin{bmatrix} 1 & \hat{H}_{21}^*(\omega) \\ \hat{H}_{12}^*(\omega) & 1 \end{bmatrix} \quad (13)$$

ここで $P_{s_1 s_1}(\omega)$ と $P_{s_2 s_2}$ はそれぞれ $s_1(t)$ と $s_2(t)$ のパワースペクトルである。式 (13) を (9) に代入し整理すると、下記の式を得る。

$$P_{s_1 s_1}(\omega)[1 - \hat{H}_{12}(\omega)H_{21}(\omega)][H_{21}(\omega) - \hat{H}_{21}(\omega)]^* + P_{s_2 s_2}(\omega)[1 - \hat{H}_{21}(\omega)H_{12}(\omega)]^*[H_{12}(\omega) - \hat{H}_{12}(\omega)] = 0 \quad (14)$$

もし $\hat{H}_{21}(\omega) = H_{21}(\omega)$ なら $\hat{H}_{12}(\omega) = H_{12}(\omega)$ であり、このとき式 (3) における条件が満たされれば、 $P_{s_2 s_2}(\omega)$ は正(零を含まない)である。同様に、もし $\hat{H}_{12}(\omega) = H_{12}(\omega)$ なら $\hat{H}_{21}(\omega) = H_{21}(\omega)$ であり、 $P_{s_1 s_1}(\omega)$ は正(零を含まない)である。したがって、もし結合系のうちのひとつが既知なら、無相関基準は他の結合系についての解を与える。

結合系のうちのひとつが先験的に既知か、独立に測定可能なような場合は実際にあり得る。例えば音声強調においては、所望の音声信号か妨害信号のどちらかが固定した位置にあり、したがって、マイクまでの音響的伝達関数は先験的に測定可能である。その場合、式 (10) か (11) のどちらを用いて、他方の結合系を求めることができる。また、競

合話者の分離問題では、一方の話者の無音期間を識別することにより他方の話者についての音響的伝達関数を個々に推定し、両方の話者が発声するときに分離を行なうことができる。

3 推定アルゴリズム

前節で述べたように無相関基準に基づく推定系の解は式 (10),(11) で与えられるが、互いに推定系自身を含んでいるため、混合系が全く未知の場合は解が得られない。そこで無相関基準による逐次推定アルゴリズムを用い、これによって式 (10),(11) を達成することとした。また推定対象の変動への対応を考慮し処理は時間領域で行なう。これらの点を反映して推定アルゴリズムは下記のとおりとした。

$$\begin{aligned} \hat{h}_{12(k)}(n+1) &= \hat{h}_{12(k)}(n) + \mu E\{\hat{s}_1(n)\hat{s}_2(n-k)\}/E\{s_1^2(n)\} \\ \hat{h}_{21(k)}(n+1) &= \hat{h}_{21(k)}(n) + \mu E\{\hat{s}_2(n)\hat{s}_1(n-k)\}/E\{s_2^2(n)\} \end{aligned}$$

式に示すように推定系は分離回路出力の相関を用いて逐次推定を行なう。アルゴリズムの安定のために出力信号のパワーを用いて正規化を行なっている。また μ は逐次推定の速度を設定するパラメータである。

上記に示したアルゴリズムの評価を行なうため、計算機実験を行なった。混合系は Fig.1 に示すように信号源、相互結合が 2×2 とした。信号源は双方とも白色雑音、および双方とも音声の2種類の場合について評価を行なった。相互結合系のインパルス応答は長さ4とし、乱数を用いて発生させた。分離回路は Fig.4 に示す構成とし、 $\mu = 0.0001$ 、推定系の係数は初期状態0から開始する。

信号源が白色雑音である場合の評価結果を示す。信号源のレベルは2個の信号源が同一レベルである。この評価では信号源が定常であるので、逐次推定アルゴリズムの収束状態を定量評価する。推定系の係数値が0である初期状態からの、分離回路出力における残留クロストークの変化を Fig.5(図中の Residual) に示す。図に見られるように初期状態では入力でのクロストークと同じレベルであるが、逐次推定の進行とともに直線的にクロストークが減少し、定常状態では入力でのレベルに対し約-30dB に達する。残留クロストークの減少速度は約5dB/10000回である。これらの特性は ch.1 と ch.2 が同様の特性を示しており、安定な動作を示している。

次に両方の信号源が音声信号の場合の特性を示す。信号源は Fig.6 に示すように、互いに独立で同程度のレベルの音声サンプルを用いた。実時間では約2[s]に相当する。この場合の初期状態からの逐次推定に伴うクロストークの変化を Fig.7 に示す。残留クロストーク 図には片側 (s_1 側) について入力におけるクロストーク (Crosstalk) と分離回路出力における残留クロストーク (Residual) を示している。図に見られるように逐次推定開始直後からクロストークは入

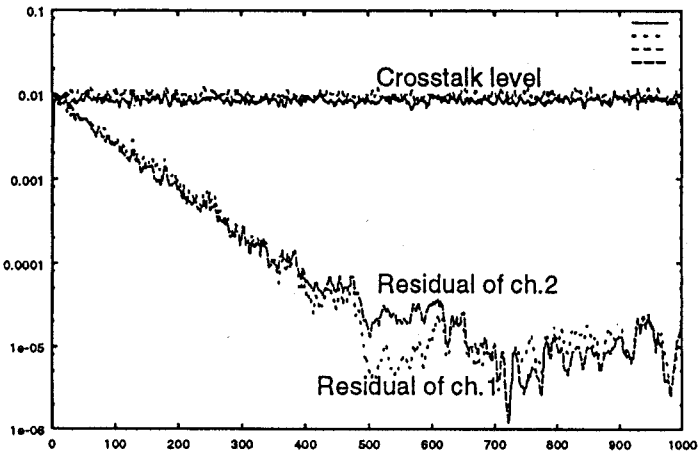


Fig.5 収束特性

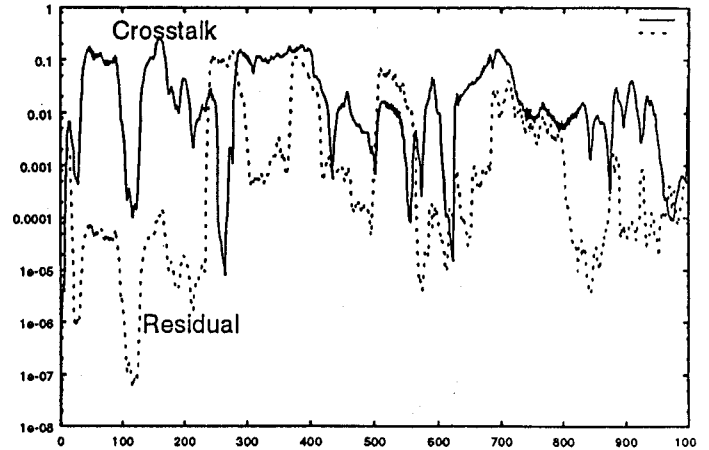


Fig.7 残留クロストークの変化

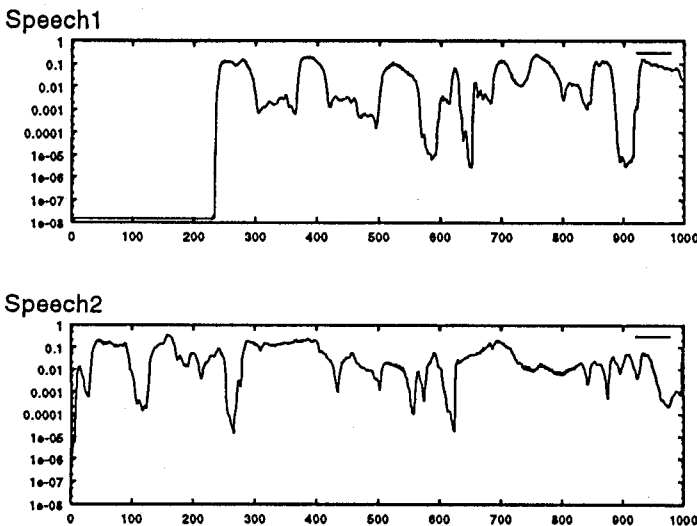


Fig.6 音声入力信号

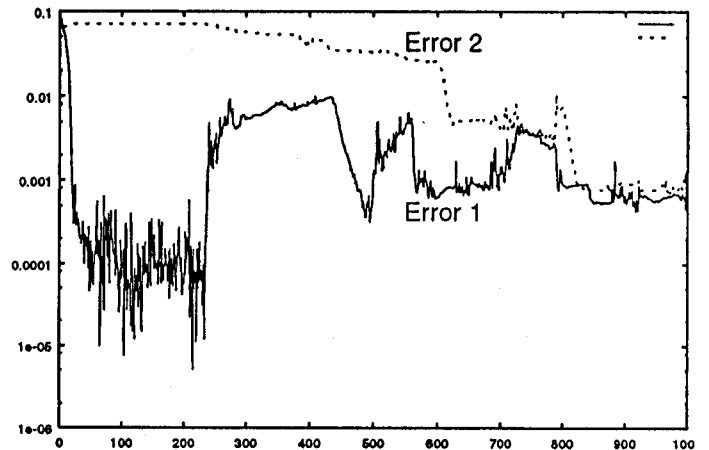


Fig.8 係数誤差の変化

力よりも削減されており、ほぼすべての期間を通じて入力より低いレベルにある。また、同じ条件下での結合系と推定系との係数値誤差の推移を Fig.8 に示す。信号源のレベルは各時点で異なるので誤差は s_1 側 (Error 1) と s_2 側 (Error 2) の変化は異なるが、双方とも最終的には誤差が約 1/10 になる。この結果から、音声信号に対しても逐次推定アルゴリズムが機能し分離が可能であることがわかる。

4 おわりに

相関関数を分離の規範とするブランド信号分離手法を検討し、その規範に基づく逐次推定アルゴリズムを提案、基本的な動作を計算機実験により確認した。その結果によれば、白色雑音を用いた信号源に対しては、クロストークの抑圧量は約 5dB/10000 回で増加し最終的に約 30dB に達した。また音声信号源に対しても約 10dB の抑圧を実現した。

本稿では相関に基づくアルゴリズムの基本動作を確認し

たが、限定された条件下での実験的検討に限られているので、今後は実験条件を一般化した上で高次統計量を用いたアルゴリズムとの比較検討を行ない、両者の相違を明確化したい。また解析的な検討を加えることにより、より一般的な指針を得ることも重要である。これらの検討の中で、音響環境条件の調査や、逐次推定アルゴリズム中の各パラメータ値の推定速度や安定性に関する最適化も検討したい。

参考文献

- [1] 坂井, 高橋, 岩倉: 第 13 回デジタル信号処理シンポジウム (1998) 13-18
- [2] E.Weinstein: IEEE Trans. on SAP, Vol.1, No.4 (1993) 405-413