

反響音を有する畳み込み形混合過程に対する高次 BSS の学習法 - 学習率に重みを付けて制御する方法 -

A Learning Algorithm for BSS of Convolutional Mixtures with Reverberations
-By Using Adaptive Weighted Stepsize-

堀田 明秀[†] 中山 謙二[‡] 平野 晃宏[‡]
Akihide Horita Kenji Nakayama Akihiro Hirano

[†]金沢大学大学院 自然科学研究科 電子情報システム専攻
Graduate School of Natural Science & Technology, Kanazawa Univ.

[‡]金沢大学工学部情報システム工学科
Dept. of Information and Systems Eng., Faculty of Eng., Kanazawa Univ.
E-mail: nakayama@t.kanazawa-u.ac.jp

あらまし

ブラインドソースセパレーション (BSS) において、多くの場合、混合過程は周波数特性を持つ。特に、反響音を含む場合には遅延と信号パワーの関係から性能が低下する。

その反響音を含む畳み込み形混合過程に対する BSS 方式において、分離回路の学習における学習率をフィルタ係数に対して可変する方式を以前に提案した [17]。本稿では分離回路のフィルタ次数が高い場合について白色、有色、音声でその性能を解析した。反響音が生じて分離回路のフィルタ次数が高くなる場合や、入力信号の有色性が増すと分離特性は劣化する。しかし、指数重み付きステップサイズは固定ステップサイズより常に有効であり、その差は有色性の影響をあまり受けない。

ABSTRACT

In many applications, mixing processes have a certain frequency characteristic in blind source separation (BSS). When reverberations occur, separation performances are especially worse because of relation between delay and signal power.

We proposed the learning method using an exponentially weighted stepsize for adaptive filters in the unmixing blocks applied to BSS of convolutional with reverberations [17]. Order of adaptive filters used in the unmixing blocks are increased due to the reverberations. In this paper, separation performance is evaluated for very high-order adaptive filters used in

the unmixing block. The separation performance is decreased due to high-order filters and colored input signals. However, the exponentially weighted stepsize is always useful compared with the conventional fixed stepsize.

1 まえがき

雑音除去、エコー除去、回線等化、信号の推定や復元といった信号処理は近年重要になってきている。しかしいずれの場合においても、信号や雑音についての十分な情報が得られない。さらに前もって混合過程や伝達過程を知ることができない。このような状況で、信号の統計的な性質を使って信号を分離するブラインドソースセパレーション (BSS) はとても重要である [1]-[7],[13],[14]。

多くの場合、混合過程において畳み込みが生じるため、分離過程において FIR もしくは IIR フィルタが必要になる。これまで時間領域や周波数領域においてさまざまな手法が提案されてきた。しかし高次のフィルタが必要になる、不安定である、分離性能が低いという問題があった [8]-[12]。さらに反響音では悪条件のため収束に時間がかかってしまい、これまで有効な方法がなかった。

我々は以前、反響音を含む畳み込み形混合過程に対する BSS 方式において、分離回路の学習における学習率をフィルタ係数に対して可変する方式を提案し、簡単なシミュレーションにおいてその有効性を示した [17]。

本稿では分離回路のフィルタ次数が高く、より複雑な問題に対してその有効性を検討する。

2 BSS の回路と学習アルゴリズム

2.1 回路構成と式

Jutten アルゴリズムによる回路構成を図 1 に示す [3] . 混合過程は畳み込みの形をしている . 分離回路は図 2 に示す FIR フィルタを用いる .

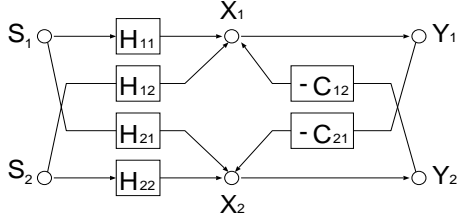


図 1: リカーレント形 BSS のブロック図

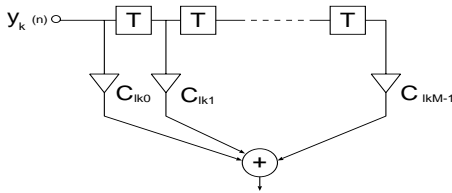


図 2: $C_{21}(z)$ や $C_{12}(z)$ に用いる FIR フィルタ

信号源 $s_i(n), i = 1, 2, \dots, N$ はインパルス応答が $h_{ji}(m)$ である未知畳み込み混合過程を通して, N 点の $x_j(n)$ として観測される .

$$x_j(n) = \sum_{i=1}^N \sum_{m=0}^{M_{ji}-1} h_{ji}(m) s_i(n-m) \quad (1)$$

また, 分離過程の出力 $y_j(n)$ は次式で表される .

$$y_j(n) = x_j(n) - \sum_{\substack{k=1 \\ k \neq j}}^N \sum_{l=0}^{L_{jk}-1} c_{jk}(l) y_k(n-l) \quad (2)$$

この関係をベクトルと行列で表すと次のようになる .

$$\mathbf{x}(n) = \mathbf{H}^T \mathbf{s}(n) \quad (3)$$

$$\mathbf{y}(n) = \mathbf{x}(n) - \mathbf{C}^T \mathbf{y}(n) \quad (4)$$

$$\mathbf{s}(n) = [s_1^T(n), s_2^T(n), \dots, s_N^T(n)]^T \quad (5)$$

$$\mathbf{s}_i(n) = [s_i(n), s_i(n-1), \dots, s_i(n-M_i+1)]^T \quad (6)$$

$$\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_N(n)]^T \quad (7)$$

$$\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_N(n)]^T \quad (8)$$

$$\mathbf{y}(n) = [\mathbf{y}_1^T(n), \mathbf{y}_2^T(n), \dots, \mathbf{y}_N^T(n)]^T \quad (9)$$

$$\mathbf{y}_k(n) = [y_k(n), y_k(n-1), \dots, y_k(n-L_{jk}+1)] \quad (10)$$

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{21} & \dots & h_{N1} \\ h_{12} & h_{22} & \dots & h_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1N} & h_{2N} & \dots & h_{NN} \end{bmatrix} \quad (11)$$

$$\mathbf{h}_{ji} = [h_{ji}(0), h_{ji}(1), \dots, h_{ji}(M_{ji}-1)]^T \quad (12)$$

$$\mathbf{C} = \begin{bmatrix} 0 & c_{21} & \dots & c_{N1} \\ c_{12} & 0 & \dots & c_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1N} & c_{2N} & \dots & 0 \end{bmatrix} \quad (13)$$

$$\mathbf{c}_{jk} = [c_{jk}(0), c_{jk}(1), \dots, c_{jk}(L_{jk}-1)]^T \quad (14)$$

$s_i(n), x_j(n), y_k(n)$ をそれぞれ z 変換したものを $S_i(z), X_j(z), Y_k(z)$ とすると以下のように表せる .

$$\mathbf{X}(z) = \mathbf{H}(z) \mathbf{S}(z) \quad (15)$$

$$\mathbf{Y}(z) = \mathbf{X}(z) - \mathbf{C}(z) \mathbf{Y}(z) \quad (16)$$

$$\mathbf{S}(z) = [S_1(z), S_2(z), \dots, S_N(z)]^T \quad (17)$$

$$\mathbf{X}(z) = [X_1(z), X_2(z), \dots, X_N(z)]^T \quad (18)$$

$$\mathbf{Y}(z) = [Y_1(z), Y_2(z), \dots, Y_N(z)]^T \quad (19)$$

この表現から信号源と分離過程の出力との関係を以下のように表せる .

$$\begin{aligned} \mathbf{Y}(z) &= (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{X}(z) \\ &= (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{H}(z) \mathbf{S}(z) \end{aligned} \quad (20)$$

分離性能を評価するために以下の行列を定義する .

$$\mathbf{P}(z) = (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{H}(z) \quad (21)$$

もし $\mathbf{P}(z)$ の各行各列に 0 でない要素を一つだけ持つならば信号源 $s_i(n)$ は出力 $y_k(n)$ に完全に分離されているといえる . しかし, $\mathbf{H}(z)$ の等化までは保証しないので分離された信号は次のような形になる .

$$Y_j(z) = P_{ji}(z) S_i(z) \quad (22)$$

2.2 学習アルゴリズム

畳み込み形の BSS に対するの概要は [15] で説明している . 簡単化のため 2 チャンネルにおいて考える . 完全に分離が行なわれる解は次のような場合である .

$$(1) C_{21}(z) = \frac{H_{21}(z)}{H_{11}(z)} \quad C_{12}(z) = \frac{H_{12}(z)}{H_{22}(z)} \quad (23)$$

$$y_1(n) = \mathbf{h}_{11}^T \mathbf{s}_1(n) \quad y_2(n) = \mathbf{h}_{22}^T \mathbf{s}_2(n) \quad (24)$$

$$(2) C_{21}(z) = \frac{H_{22}(z)}{H_{12}(z)} \quad C_{12}(z) = \frac{H_{11}(z)}{H_{21}(z)} \quad (25)$$

$$y_1(n) = \mathbf{h}_{12}^T \mathbf{s}_2(n) \quad y_2(n) = \mathbf{h}_{21}^T \mathbf{s}_1(n) \quad (26)$$

ここで $H_{11}(z)$ や $H_{22}(z)$ は $H_{21}(z)$ や $H_{12}(z)$ より遅延時間が短いとする . これは図 2 において X_1 が $s_1(n)$ に近く, X_2 が $s_2(n)$ に近いことを表している . この仮定から解 (1) は因果性を満たすが, 解 (2) は満たさない .

式 (20) から出力は次のように表せる .

$$\begin{aligned} \begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} &= \frac{1}{1 - C_{12}(z)C_{21}(z)} \begin{bmatrix} 1 & -C_{12}(z) \\ -C_{21}(z) & 1 \end{bmatrix} \\ \times \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} & \quad (27) \\ &= \frac{1}{1 - C_{12}(z)C_{21}(z)} \\ \times \begin{bmatrix} H_{11}(z) - C_{12}(z)H_{21}(z) & H_{12}(z) - C_{12}(z)H_{22}(z) \\ H_{21}(z) - C_{21}(z)H_{11}(z) & H_{22}(z) - C_{21}(z)H_{12}(z) \end{bmatrix} \\ \times \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} & \quad (28) \end{aligned}$$

式 (25) は因果性を満たす回路では実現できないため , 式 (28) の対角要素は 0 にならない . 逆に非対角要素は 0 になり得る . したがって , コスト関数を以下のように設定できる .

$$J_j(n) = E[q(y_j(n))] \quad (29)$$

$q(\cdot)$ はただ 1 つの極小値を持つ偶関数である . コスト関数を最小化することによって , $C_{12}(z)$ と $C_{21}(z)$ を式 (23) に近づけることができる . $E[q(y_j(n))]$ の代わりに , 瞬時値 $q(y_j(n))$ を使い , 勾配法を適用する .

$$\hat{J}_j(n) = q(y_j(n)) \quad (30)$$

$\hat{J}_j(n)$ の勾配は次式のようになる .

$$\begin{aligned} \frac{\partial \hat{J}_j(n)}{\partial c_{jk}(l)} &= \frac{\partial q(y_j(n))}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial c_{jk}(l)} \\ &= \dot{q}(y_j(n)) y_k(n-l) \end{aligned} \quad (31)$$

$$y_j(n) = x_j(n) - \sum_{l=0}^{L_{jk}-1} c_{jk}(l) y_k(n-l) \quad (32)$$

ここで $\dot{q}(\cdot)$ は $q(\cdot)$ の導関数である . $k=1$ なら $j=2$ であり , $k=2$ なら $j=1$ である . したがって , $c_{jk}(l)$ の更新式は次式で与えられる .

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \Delta c_{jk}(n, l) \quad (33)$$

$$\Delta c_{jk}(n, l) = \mu \dot{q}(y_j(n)) y_k(n-l) \quad (34)$$

信号源の確率密度関数 (pdf) が偶関数で , 互いに独立であるとすれば以下の式を満足する .

$$\begin{aligned} E[f(s_1(n))g(s_2(n))] &= E[f(s_1(n))]E[g(s_2(n))] \\ &= 0 \end{aligned} \quad (35)$$

$f(\cdot), g(\cdot)$: 奇関数

式 (34) において小さな値のステップサイズ μ を使用したとすると , 修正項は $E[\dot{q}(y_j(n))y_k(n-l)]$ 見なせる . $\dot{q}(y_j(n))$ と $y_k(n-l)$ は奇関数であるから式 (35) を満たす . これは修正項が小さくなるにつれ , $y_1(n)$ や $y_2(n)$ がそれぞれ $h_{11}^T s_1(n)$ や $h_{22}^T s_2(n)$ に近づくということである .

3 重み付き学習率

3.1 収束解析

反響音を含む場合 , 遅延に関する仮定が満たされない可能性がある . 反響音を含むモデルを図 3 に示す .

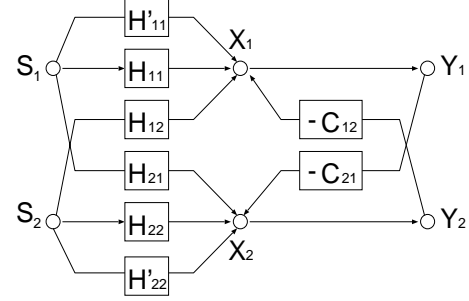


図 3: 反響音 $H'_{11}(z)$ and $H'_{22}(z)$ を含む畳み込み形 BSS

$H'_{11}(z)$ と $H'_{22}(z)$ は反響音による伝達関数を表している , これらは遅延時間が長いとする . $H'_{12}(z)$ と $H'_{21}(z)$ はここでは無視する . X_1 のパワーを減少させるのは以下の 2 つの項である .

$$H'_{11}(z)S_1(z) - C_{12}(z)H_{21}(z)S_1(z) \rightarrow 0 \quad (36)$$

$$C_{12}(z) \rightarrow \frac{H'_{11}(z)}{H_{21}(z)} \quad (37)$$

$$H_{12}(z)S_2(z) - C_{12}(z)H_{22}(z)S_2(z) \rightarrow 0 \quad (38)$$

$$C_{12}(z) \rightarrow \frac{H_{12}(z)}{H_{22}(z)} \quad (39)$$

また X_2 についても同様である .

$$H'_{22}(z)S_2(z) - C_{21}(z)H_{12}(z)S_2(z) \rightarrow 0 \quad (40)$$

$$C_{21}(z) \rightarrow \frac{H'_{22}(z)}{H_{12}(z)} \quad (41)$$

$$H_{21}(z)S_1(z) - C_{21}(z)H_{11}(z)S_1(z) \rightarrow 0 \quad (42)$$

$$C_{21}(z) \rightarrow \frac{H_{21}(z)}{H_{11}(z)} \quad (43)$$

式 (39) と式 (43) は理想解である . しかし , 式 (37) , 式 (41) で与えられる反響音によって $C_{12}(z)$ と $C_{21}(z)$ がこれらの解に近付かなくなってしまう .

3.2 指数制御の学習アルゴリズム

反響音の遅延は長く , 式 (37) , 式 (41) から反響音の影響はインパルス応答の後半部分に出てくる . このことから後半部分の修正を抑える必要がある . これは FIR フィルタのディレイラインにあわせてステップサイズ μ を指数的に制御することによって実現できる . したがって , 更新式を次のようにする .

$$\begin{aligned} c_{jk}(n+1, l) &= c_{jk}(n, l) \\ &+ \mu(l) f(y_j(n)) g(y_k(n-l)) \end{aligned} \quad (44)$$

$$\mu(l) = \mu_0 \gamma^l, \quad 0 < \gamma < 1 \quad (45)$$

$\mu(l)$ は理想解に比例する .

4 シミュレーション

4.1 シミュレーション条件

2チャンネルのシミュレーションを行なった . 非線形関数には以下のものを使用した .

$$f(y) = \tanh(2.5y) \quad g(y) = \tanh(0.5y) \quad (46)$$

分離性能は式 (21) で定義した $P(z)$ を用い , 以下の SNR で評価した .

$$\sigma_s^2 = \sum_{i=1}^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ii}(e^{j\omega T})|^2 d\omega T \quad (47)$$

$$\sigma_c^2 = \sum_{j \neq i} \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ji}(e^{j\omega T})|^2 d\omega T \quad (48)$$

$$SNR = 10 \log \frac{\sigma_s^2}{\sigma_c^2} \quad [\text{dB}] \quad (49)$$

σ_s^2 は残すべき信号のパワー , σ_c^2 は分離すべき信号のパワーを表す .

$c_{12}(z), c_{21}(z)$ の理想インパルス応答を図 4 に示す . この図から 1000 次の FIR フィルタが必要である . また反響音を含まない場合の $c_{12}(z), c_{21}(z)$ の理想インパルス応答を図 5 に示す . この図から反響音を含まない場合は 50 次の FIR フィルタが必要である .

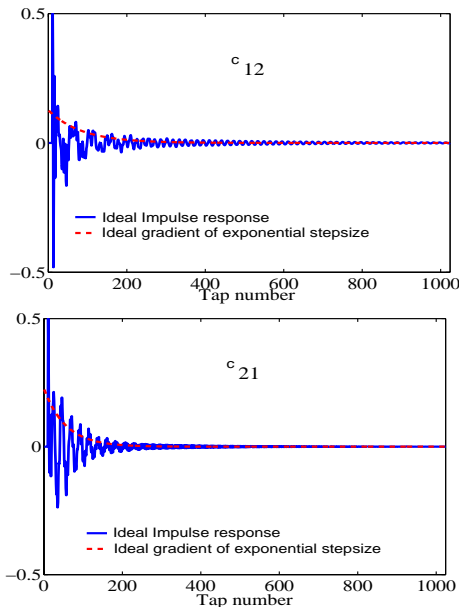


図 4: 理想インパルス応答

4.2 白色信号に対する分離特性

固定ステップサイズ $\mu = 0.0003$, 指数制御ステップサイズ $\mu_0 = 0.00078$, $\gamma_{12} = 0.9883$, $\gamma_{21} = 0.9822$ (図 4 の破

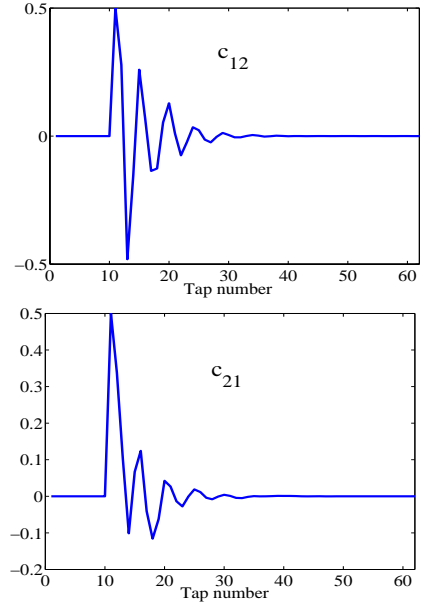


図 5: 反響音を含まない場合における理想インパルス応答

線) を使ってシミュレーションを行なったときの SNR を図 6 に示す . さらに反響音を含まない場合における SNR も同時に示す . 反響音を含まない場合のステップサイズは $\mu = 0.0008$ である .

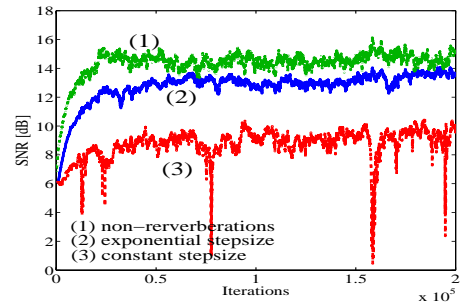


図 6: 白色入力に対する分離特性

この図からステップサイズを固定した場合 , 反響音の影響を受け性能が著しく劣化しているのが分かる . また , 指数制御ステップサイズでは反響音を含まない状況に近い性能が得られていることが分かる .

4.3 有色信号に対する分離特性

次式の 2 次 AR モデルで r を変化させることにより有色度を変化させて有色度による性能の違いを図 7 , 8 に示す .

$$S(z) = \frac{1}{1 - 2r \cos(\frac{\pi}{4}) + r^2} \quad (50)$$

反響音のない場合のステップサイズを 0.00025 とし , その他は白色信号に対するときと同様のものを使用した .

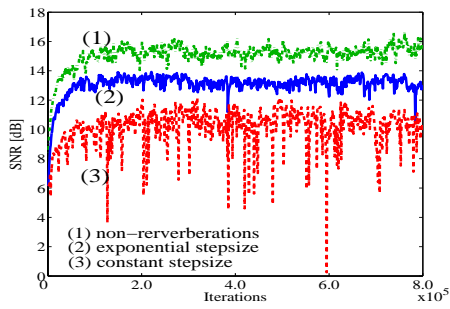


図 7: 有色入力 ($r = 0.8$) に対する分離特性

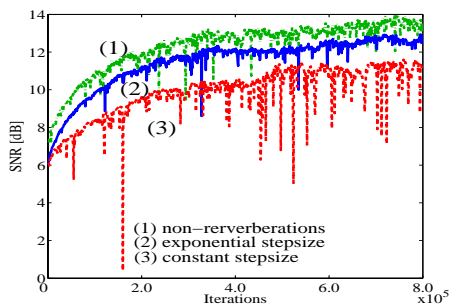


図 8: 有色入力 ($r = 0.99$) に対する分離特性

この図から有色の割合が高くなると全体的に分離性能が悪くなってしまふ．また有色度が高くなると収束が遅くなってしまふことが分かる．

4.4 音声信号に対する分離特性

音声信号に対して，シミュレーションを行なったときの SNR を図 8 に示す．ステップサイズは白色色に対するときと同様のものを使用した．

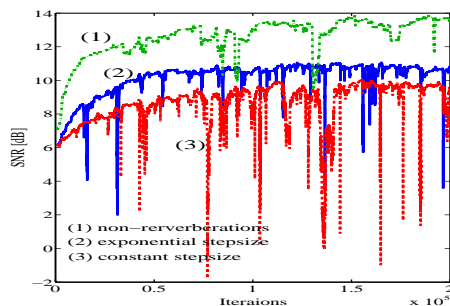


図 9: 音声入力に対する分離特性

音声信号に対しては反響音を含まない場合に比べて指数制御ステップサイズ方式の分離性能が低下している．しかし，固定ステップサイズの分離性能を上回っている．

4.5 分離特性に対する考察

勾配法の残留誤差は適応フィルタの次数とステップサイズに比例する．反響音がない場合の適応フィルタは 50 次で，ある場合は 1000 次である．また，ステップサイズは各々 0.0008 と 0.00025 ~ 0.0003 である．この分だけ反響音がある場合の分離特性が劣化する．しかし，指数重み付きがされている点にも配慮する必要がある．有色信号では固有値広がりが大きくなり，分離特性は白色信号の場合よりも劣化しているが，3 通りの方法の比較においては白色信号とほぼ同じ傾向を示している．一方，非定常信号である音声の場合は，反響音の影響による劣化が少し大きくなっている．この点についてはさらに解析を進める必要がある．

5 結論

反響音が長く分離回路のフィルタ次数が高い場合において指数制御ステップサイズ方式の有効性を検討した．白色信号のように信号間の相関が低く有色性の低い信号に対しては反響音を含まない場合の性能に近い分離性能が得られた．信号の有色性が強まると収束の速度が遅くなり性能が悪化してしまう．音声信号に対してはさらに性能が低下してしまう．しかし，いずれの場合においても指数制御ステップサイズ方式は固定ステップサイズ方式の分離性能を上回っており，指数制御ステップサイズ方式が有効であるといえる．

参考文献

- [1] C.Jutten, J.Herault and A.Guerin, "IIN.C.A: An independent components analyzer based on an adaptive neuromimetic network", in: J.Degmogeot, T.Herve, V.Raille and C.Roche, eds., Artificial Intelligence and Cognitive Science, Manhester Univ. Press, Manchester, 1988.
- [2] J.F.Cardoso, "Eigen structure of the 4th order cumulant tensor with application to the blind source separation problem", ICASSP Proc. pp. 2655-1658.
- [3] 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.
- [4] P.Comon, C.Jutten and J.Herault, "Blind separation of sources, Part II: Problems statement", Signal Processing, 24, pp.11-20, 1991.
- [5] 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.
- [6] K.Nakayama, A.Hirano and M.Nitta, "A constraint learning algorithm for blind source separation", Proc. IJCNN'2000, pp.24-27, July, 2000.

- [7] K.Nakayama, A.Hirano and T.Sakai, "A pair-channel learning algorithm with constraints for multi-channel blind separation", Proc. IJCNN'01, July 2001.
- [8] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures", Signal Processing, vol.45, no.2, pp.209-229, March 1995.
- [9] C.Simon, G.d' Urso, C.Vignat, Ph.Loubaton and C.Jutten, "On the convolutive mixture source separation by the decorrelation approach", Proc. ICASSP'98, pp.IV-2109-2112, May 1998.
- [10] S.Cruces and L.Castedo, "A Gauss-Newton methods for blind source separation of convolutive mixtures", ICASSP'98, pp.IV2093-2096, May 1998.
- [11] S.Araki, S.Makino, T.Nishikawa and H.Saruwatari, "Fundamental limitation of frequency domain blind source separation for convolutive mixture of speech", Proc. ICASSP'01, MULT-P2.3, May 2001.
- [12] I.Kopriva, Z.Devcic and H.Szu, "An adaptive short-time frequency domain algorithm for blind separation of nonstationary convolved mixtures", Proc. IJCNN'01, pp.424-429, July 2001.
- [13] H.Mathis and S.C.Douglas, "On optimal and universal nonlinearities for blind signal separation", Proc. ICASSP'01, MULT-P3.3, May 2001.
- [14] K.Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by kurtosis for blind source separation", Proc. IJCNN'2002, pp.1234-1239, May 2002.
- [15] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm for convolutive blind source separation with transmission delay constraint", Proc. IJCNN'2002, pp.1287-1292, May 2002.
- [16] S.Makino, Y.Kaneda, N.Koizumi, "Exponentially weighted stepsize NLMS adaptive filter based on the statistics of a room impulse response", IEEE Trans. Speech and Audio Processing, vol.1, no.1, pp.101-108, Jan. 1993.
- [17] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm with adaptive exponential stepsize for blind source separation of sonvolutive mixtures with reverberations", Proc. IJCNN'2003 July 2003.
- [18] K. Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output pdf for blind source separation", 4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA2003), Nara, Japan, pp.427-432, April 2003.
- [19] K.Nakayama, A.Hirano and T.Nishiwaki", A cascade form blind source separation connecting source separation and linearization for nonlinear mixtures", IEEE&INNS Proc. IJCNN'03, Portland, Oregon, July 2003.