

グループ分離と線形化による非線形BSSにおける収束性解析

The Convergent Analysis of a Cascade Form BSS Connecting Source Separation and Linearization for Nonlinear Mixtures

西脇 貴之[†] 中山 謙二[‡] 平野 晃宏[‡]
Takayuki Nisiwaki 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

Division of Electronics and Computer Engineering

Graduate School of Natural Science and Technology, Kanazawa Univ.

[‡] 金沢大学大学院 自然科学研究科 電子情報科学専攻

Division of Electrical Engineering and Computer Science

Graduate School of Natural Science and Technology, Kanazawa Univ.

E-mail: nakayama@t.kanazawa-u.ac.jp

あらまし

2次の非線形混合において、信号群分離と線形化を縦続構成するブラインドソースセパレーション(BSS)が提案されている。信号群分離では、信号源とその高次項を含んだ同一グループの信号群に分離される。残った高次項は、線形化を通すことにより抑制される。本稿では、まず、この非線形BSSの分離特性を解析する。センサの数は信号源以外の干渉成分を消去するために、信号源の数より増やす必要がある。また、干渉成分は非線形成分の大きさによって決まる。非線形成分の割合とセンサ数の関係について解析し、非線形成分の割合が低い場合にセンサ数を低減できることを確認した。次に、学習において、分離行列の初期値依存性について解析した。ランダムに発生する初期値に対して、約30%の確率で良い分離特性が得られた。これにより、比較的少ない探索回数で有効な初期値が求まることが分かった。また、非線形として3次までを考慮した場合の影響を解析した。3次項が1次項に比べて約10%程度では良い分離特性が得られることを確認した。

ABSTRACT

A blind source separation (BSS), cascading a signal group separation block and a linearization block has been proposed for low-order nonlinear mixtures. In the separation block, the signal sources are separated into each group, including its high-order components. The high-order components are further suppressed through

the linearization block. In this paper, separation performance of the nonlinear BSS is analyzed from several view points. The number of the sensors is increased from that of the signal sources in order to cancel the interference. Moreover, the interference components is decided by a ratio of the nonlinear and the linear components. A relation between the ratio of the components and the number of the sensors is analyzed. The number of the sensors can be reduced when the ratio of the nonlinearity is small. Next, effects of the initial guess of the separation matrix is analyzed. The training was carried out using 50 independent random initial guess, and good separation is obtained by a 30% probability. Moreover, effect of including 3rd-order terms is analyzed. When the 3rd-order term is under 10%, good separation performance can be obtained.

1 まえがき

BSSの実際の混合過程として、センサ(マイク)やアンプ等により信号に非線形歪みが生じることがある。この場合、混合信号は非線形を含む複雑な形となり、線形混合に対するBSSの方法[1]-[8]では分離できない。非線形混合過程として、信号源の線形結合後に非線形歪みが生じるようにモデル化する方法がある。Post-nonlinear(PNL)混合のために、ミラ構造BSSが主に使用されている[14]。この場合、非線形歪みは始めの線形化の段階で消去される。スプライン非線形関数やスプラインニューラルネットワークもまた非線形混合に適用

されている [11],[12]. さらに, 最尤推定法を用いる方法 [13]なども提案されている. ニューラルネットワークもまた適用されている [15].

非線形として2次の多項式近似したものを考えた場合, 信号群分離と線形化を含む縦続形BSSが提案されている [16][17]. 線形化の段階では, 線形項とその2次の高次項を含んだ信号に分離される. 次に線形化の段階では, 残った高次項が消去される.

センサの数は欲しい信号以外の干渉成分を消去するために, 信号源の数より増やす必要がある. 理論上必要なセンサの数は決めることができる. また, 干渉成分は非線形の大きさによって決まる. 本稿では, 非線形の割合とセンサ数の関係について解析する.

また, 分離行列のパラメータの初期値の決め方としても重要になってくる. 本稿では, 初期値依存性について解析する.

さらに, 現実的にスピーカを信号源と考えた場合, 非線形として3次までを考慮する必要がある, 3次項を含む場合の影響について解析をする.

2 縦続形ブラインドソースセパレーション

2.1 回路構成

縦続形BSSの回路構成を図1に示す.

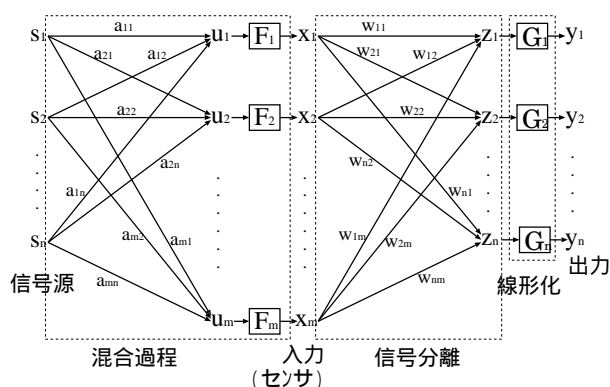


図1:縦続形BSSの回路構成

F_i は非線形関数である. 本稿では, 2次まで ($F_i = a_i u_i + b_i u_i^2$)を対象とする. 信号群分離の過程では, 信号源が自らの非線形項を含む形で分離される. 信号源毎に分離された後, 非線形項が線形化過程で抑圧される.

z_i と z_j に出力される信号 $f_i(s_i)$ と $f_j(s_j)$ は統計的に独立であり従来法で分離が可能である. しかし, s_i と s_j のクロス項 $s_i s_j$ は s_i と s_j の信号源に相関があるので, 分

離が不可能であり消去する必要がある. 信号源とセンサの数と同じ場合は, このクロス項を消去することができない. クロス項を消去するためには, センサの数を信号源の数より多くし, 条件の数を増やす必要がある. よって, センサの数を信号源の数より多くする構成となっている.

また, 信号群分離の段階において出力される信号 $f_i(s_i)$ と $f_j(s_j)$ というのは, お互いに独立した形になっているが, 線形項のほかに同じ信号源の高次項も残っている. z_i を非線形関数 G_i に通すことにより線形化を行い, 高次項の抑制をしている.

2.2 必要なセンサ数

非線形関数が $F_i = a_i u_i + b_i u_i^2$ の場合, 信号源 n 個に対して必要なセンサ数は $n(n+1)/2 + n - 1$ 個である. よって, 信号源2個だとセンサは4個, 信号源3個だとセンサは8つ必要となる.

3 学習アルゴリズム

[25]で提案されているBSS方式では, 第1段階で統計的に独立な成分に分離し, 第2段階で非線形成分を抑制する. 学習過程も第1段階, 第2段階に分けて行う.

3.1 信号群分離

信号源はその高次項を含んでも互いに独立である. すなわち, $s_i + s_i^2$ と $s_j + s_j^2, i \neq j$ は互いに独立である. 従って, 従来の自然勾配法 [9]を用いることによりこれらを分離できる. その誤差関数と更新式は次の式になる.

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

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \eta(n)[\Lambda(n) - f(z(n))z^T(n)]\mathbf{W}(n) \quad (2)$$

$\eta(n)$ は学習率, $\Lambda(n)$ は対角行列, 非線形関数 $f(z)$ として次のものを用いている.

$$f(z) = \tanh(z) \quad (3)$$

3.2 線形化

3.2.1 2次方程式の解

2次までの非線形を想定しているので, 信号を分離した結果は,

$$z_i = a_i s_i + b_i s_i^2 \quad (4)$$

という形になっていると仮定する。この形から元の信号 s だけを出力するために解の公式を用いると、

$$y_i(n) = G_i(z_i) = -\frac{\alpha_i}{2} \pm \sqrt{\frac{\alpha_i^2}{4} + \frac{z_i(n)}{\beta_i}} \quad (5)$$

$$\alpha_i = \frac{a_i}{b_i} \quad (6)$$

$$\beta_i = \frac{1}{b_i} \quad (7)$$

となる。

3.2.2 誤差関数

信号群分離の出力には、1次項と2次項が含まれている。また、信号源として音声や音楽を対象とした場合は、これらの波形の平均値はほぼ零である。従って、非線形特性として2次歪みまで考える場合は、次に示す出力 $y_i(n)$ の平均値を誤差関数とすることにより、非線形歪みを抑制できる。

$$E_i(n) = \frac{1}{M} \sum_{j=0}^{M-1} y_i(n-j) \quad (8)$$

パラメータ $\alpha_i(n)$, $\beta_i(n)$ は勾配法で学習することが出来る。

$$\alpha_i(n) = \alpha_i(n-1) - \eta \frac{\partial E_i(n)}{\partial \alpha_i} \quad (9)$$

$$\beta_i(n) = \beta_i(n-1) - \eta \frac{\partial E_i(n)}{\partial \beta_i} \quad (10)$$

$$\begin{aligned} \frac{\partial E_i(n)}{\partial \alpha_i(n)} &= \frac{1}{M} \sum_{i=0}^{M-1} \frac{\partial y_i(n-i)}{\partial \alpha_i(n)} \\ &= \frac{1}{M} \sum_{i=0}^{M-1} \left(-\frac{1}{2} \pm \frac{\alpha_i}{4} \left(\frac{\alpha_i^2}{4} \right. \right. \\ &\quad \left. \left. + \frac{1}{\beta_i} z_i(n-i) \right)^{-\frac{1}{2}} \right) \quad (11) \end{aligned}$$

$$\begin{aligned} \frac{\partial E_i(n)}{\partial \beta_i(n)} &= \frac{1}{M} \sum_{i=0}^{M-1} \frac{\partial y_i(n-i)}{\partial \beta_i(n)} \\ &= \frac{1}{M} \sum_{i=0}^{M-1} \left(\mp \frac{z_i(n-i)}{2\beta_i^2} \left(\frac{\alpha_i^2}{4} \right. \right. \\ &\quad \left. \left. + \frac{1}{\beta_i} z_i(n-i) \right)^{-\frac{1}{2}} \right) \quad (12) \end{aligned}$$

3.2.3 符号判定

式(5)における+と-の符号の判定において、 s_i の大きさに応じて符号を決定している。しかし、信号源 s_i は未知であるから、この条件を用いることが出来ない。そのため、 s_i の大きさと非線形特性に次のような条件を課し、符号を一意に決めることにしている。

- 非線形成分より線形成分の方が大きい。

- 波形の振幅の絶対値が1以下である。 $|s_i(n)| < 1$

この条件は実用的にも適用可能なものである。この条件下では、信号分離の出力信号

$$z_i(n) = b_{i1}s_i(n) + b_{i2}s_i^2 \quad (13)$$

において、次の等式が常に成り立っているといえる。

$$|b_{i1}s_i(n)| > |b_{i2}s_i^2(n)| \quad (14)$$

これは z_i の符号は $b_{i1}s_i(n)$ の符号と等しいことを意味している。 b_{i1} の符号を除いて、 $y_i(n)$ の符号は $z_i(n)$ の符号と同じように制御することができる。 b_{i1} の符号は分離特性に影響しない。何故ならば、BSSでは定数倍のスケールは補正できないものとして残るからである。

この2つの条件が成り立っているとき+だけを用いることができる。よって本研究では上の2つの条件が成り立っているものとし、+の符号のみを使うことにする。

3.3 評価方法

3.3.1 信号群分離

信号群分離の評価方法は「分離信号/その他の干渉信号」としてSNR[dB]で評価した。式は次のように定義される。

$$SNR = 10 \log_{10} \frac{\sigma_s^2}{\sigma_n^2} \quad (15)$$

σ_s^2 は s_i のパワーで、 σ_n^2 は s_i 以外の干渉成分のパワーである。

3.3.2 線形化

線形化の評価方法も信号群分離と同じように「分離信号/その他の干渉信号」としてSNR[dB]で評価した。しかし、 s_i が直接現れないので、式(16)を使うことはできない。よって、 s_i とその他の干渉成分を次のように区別している。

信号分離された信号 z_i は線形化するために、

$$y_i = -\frac{\alpha_i}{2} + \sqrt{\frac{\alpha_i^2}{4} + \frac{z_i}{\beta_i}} \quad (16)$$

を通す。このとき、

$$\sqrt{\frac{\alpha_i^2}{4} + \frac{z_i}{\beta_i}} = \sqrt{a_i s_i^2 + b_i s_i + c_i} \quad (17)$$

とする。さらに、

$$\sqrt{a_i s_i^2 + b_i s_i + c_i} = d_i s_i + e_i \quad (18)$$

両辺2乗すると、

$$a_i s_i^2 + b_i s_i + c_i = d_i^2 s_i^2 + 2d_i e_i s_i + e_i^2 \quad (19)$$

となる。両辺の係数を比較すると、

$$d_i^2 = a_i \quad (20)$$

$$2d_i e_i = b_i \quad (21)$$

$$e_i^2 = c_i \quad (22)$$

となる。以上のことより性能評価を次の式で行なうことにする。

$$SNR = 10 \log \frac{\sum_{i=1}^2 p_i(n)}{\sum_{i=1}^2 q_i(n)} \quad (23)$$

$$p_i(n) = \frac{1}{M} \sum_{j=0}^{M-1} (y_i(n) + \frac{\alpha}{2} - e_i(n))^2 \quad (24)$$

$$q_i(n) = \frac{1}{M} \sum_{j=0}^{M-1} (e_i - \frac{\alpha_i}{2})^2 \quad (25)$$

4 非線形項の大きさと必要なセンサ数の関係

4.1 非線形とセンサ数への影響

非線形を考慮しない場合は、信号源の数とセンサ数は同じでよい。しかし、非線形を考慮する場合干渉成分が増えるので、センサ数を増やす必要がある。信号群分離の結果出力される信号は、

$$x_i = \alpha_i s_i + \beta_i s_i^2 + v_i$$

となる。 v_i は干渉成分である。この残留干渉成分 v_i は信号群分離、線形化の双方の分離特性に影響を与える。センサ数が少ないと v_i は大きくなり、センサ数が大きいと v_i は小さくなる。

また、干渉成分 v_i は非線形項の大きさによっても決まる。非線形項が大きいと v_i は大きくなり、非線形項が小さいと v_i は小さくなる。

以上のことから非線形項が小さい場合はセンサの数を減らせる可能性がある。よって、非線形の大きさを変えその分離特性を調べ、必要とするセンサの数を調べる。

4.2 線形対非線形比(LNR2)

用いた非線形関数は $F_i = a_i u_i + b_i u_i^2$ とである。線形対非線形比を次のように定義する。

$$LNR2 = 10 \log_{10} \frac{\sum_{i=1}^m a_i^2}{\sum_{i=1}^m b_i^2} (dB) \quad (26)$$

4.3 非線形成分の比率が大きい場合

信号源3チャンネル、センサ8チャンネル、 $LNR2 = 2.97dB$ とする。図2は線形化後の分離特性のグラフである。

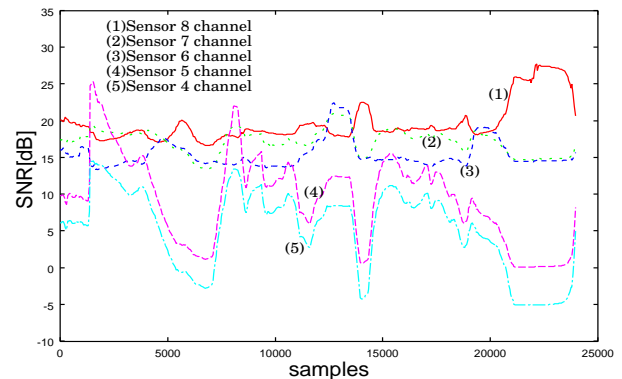


図2:線形化後の SNR(LNR2=2.97dB)

センサ数が8チャンネルの場合が一番分離特性が良く、約20dBまで分離ができている。6,7チャンネルにおいても8チャンネルより分離特性は悪くなっているが、約15dBまで分離ができている。

4.4 非線形成分の比率が小さい場合

次に $LNR2 = 8.62dB$ の場合について分離特性を調べた。図3は線形化後の分離特性のグラフである。

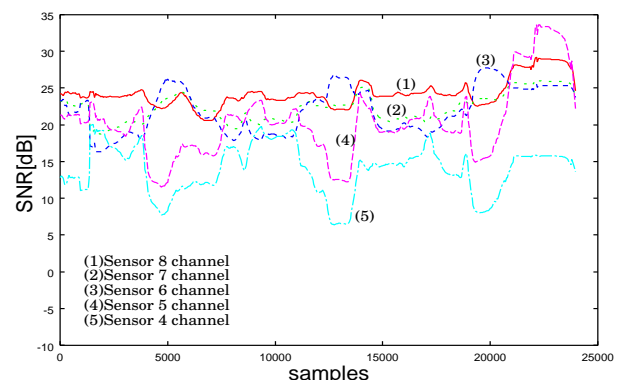


図3:線形化後の SNR(LNR2=8.62dB)

非線形項を小さくした場合、センサ数6,7チャンネルにおいて約20dBまで分離することができた。また5チャンネルにおいても約15dBまで分離することができた。

分離成功のラインを約20dBとすると、非線形の比率が大きい場合は最低でもセンサ8チャンネル必要となる。それに比べ、非線形の比率が小さい場合は6チャンネルあれば十分であることがグラフからわかる。

5 分離行列の初期値依存性

図2は線形化後の分離特性のグラフである。

分離行列のパラメータを決める際は、ある乱数を与えシミュレーションを行い、失敗したらもう一度乱数を与えまたシミュレーションを行う、というのを成功するまで手作業で行っていた。今回分離行列のパラメータの初期値依存性を調べるために、50回にわたり初期値をランダムに与え、その中で何回分離特性が良くなるかについてシミュレーションを行ってみた。信号群分離における分離特性の一番最後の数値を調べ、その値が10dB以上の分離が見られた場合に成功とみなし、それ以外を失敗とみなすことにする。

信号源2チャンネル、センサ4チャンネルとし、混合行列は乱数、非線形関数は次の4つを用いた。

$$\begin{aligned} F_1 &= u_1 + 0.4u_1^2 \\ F_2 &= u_2 + 0.2u_1^2 \\ F_3 &= u_3 - 0.6u_1^2 \\ F_4 &= u_4 + 0.3u_1^2 \end{aligned}$$

シミュレーションした結果、50回のうち14回成功したとの結果が得られた。

6 3次項の有無についての分離性能

6.1 3次項の影響

縦続形ラインドセパレーションにおいて、2次の高次非線形 $F_i = a_i u_i + b_i u_i^2$ を前提としている。本稿では、非線形項を3次項を含む場合 $F_i = a_i u_i + b_i u_i^2 + c_i u_i^3$ とした場合に分離性能にどのような影響を与えるか調べてみる。

6.2 3次対(1次+2次)比LNR3

3次も考慮したときの非線形関数は $F_i = a_i u_i + b_i u_i^2 + c_i u_i^3$ となる。3次項と(1次項+2次項)の比率を次のように定義する。

$$LNR3 = 10 \log_{10} \frac{\sum_{i=1}^m c_i^2}{\sum_{i=1}^m a_i^2 + \sum_{i=1}^m b_i^2} (dB) \quad (27)$$

6.3 シミュレーション

信号源2チャンネル、センサ4チャンネルとし、用いた非線形関数は次の2種類である。

- 3次項小さい (LNR3=-22.982dB)

$$\begin{aligned} F_1 &= u_1 + 0.4u_1^2 + 0.1u_1^3 \\ F_2 &= u_2 + 0.2u_2^2 - 0.03u_2^3 \\ F_3 &= u_3 - 0.6u_3^2 - 0.1u_3^3 \\ F_4 &= u_4 + 0.3u_4^2 + 0.05u_4^3 \end{aligned}$$

- 3次項大きい (LNR3=-8.542)

$$\begin{aligned} F_1 &= u_1 + 0.4u_1^2 - 0.3u_1^3 \\ F_2 &= u_2 + 0.2u_2^2 + 0.1u_2^3 \\ F_3 &= u_3 - 0.6u_3^2 - 0.4u_3^3 \\ F_4 &= u_4 + 0.3u_4^2 + 0.2u_4^3 \end{aligned}$$

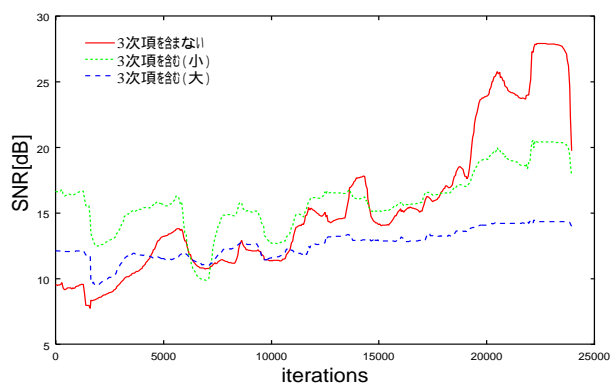


図4:信号群分離後の SNR

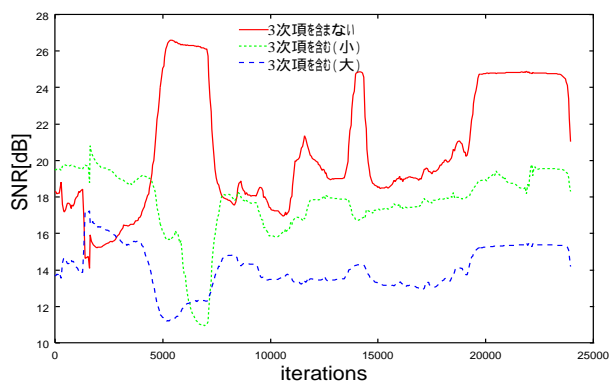


図5:線形化後の SNR

図4は信号群分離後の分離特性で、図5は信号群分離後の結果を固定した場合の線形化の分離特性のグラフである。

図6は3次項を含まない場合、3次項が小さい場合、3次項が大きい場合における信号群分離と線形化の分離特性のグラフである。3次項が小さい場合における分離特性としまして、3次項を含まない場合と比べ、信号群分離においてはほとんど一緒であるが、線形化においては分離特性が悪くなっている。3次項が大きい場合には、信号群分離後の分離特性も線形化後の分離特性も3次項を含まない場合に比べ悪くなっていることが確認できた。

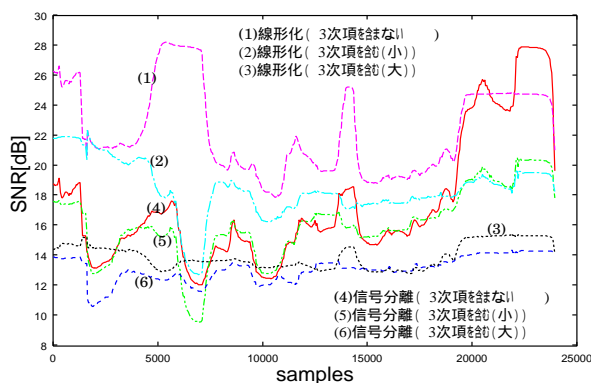


図6:信号群分離後、線形化後の SNR(係数固定)

7 結論

非線形混合過程の BSS について、縦続形構成を用いて収束性解析を行った。

非線形の比率を小さくすることにより、センサ数が低減できることを確認した。分離行列の初期値依存性に関しては、50回にわたり係数の初期値を乱数で決めシミュレーションを行った。その結果、約30%の確率で分離ができるということがわかった。また、3次項を含む場合は含まない場合に比べ分離特性が悪くなるが、多少の3次項が入力されても分離ができるということが確認できた。

参考文献

- [1] C.Jutten and J.Herault and A.Guerin, "IN.C.A.: An independent components analyser based on an adaptive neuromimetic network", in: J.Demongeot, T.Herve, V.Rialle and C.Roche, eds., Artificial Intelligence and Cognitive Sciences, Manchester Univ. Press, Manchester, 1988.
- [2] C.Jutten and J.Herault, "Analog implementation of permanent unsupervised learning algorithm", Proc. NATO Advanced Research Workshop on Neurocomputing, Les Arcs, France, 27 February-3 March 1989, pp.145-152.
- [3] P.Comon, "Separation of stochastic process whose linear mixtures observed", Proc. ONR- NSF-IEEE

Workshop on Higher Spectral Analysis Vail, Colorado, pp.174-179, June 28-30, 1989.

- [4] P.Comon, "Separation of sources using higher order cumulants", SPIE Conference, Vol.1152, Advanced Algorithms and Architectures for Signal Processing IV, San Diego, California, pp.170-181, August 6-11, 1989.
- [5] J.F.Cardoso, "Eigen structure of the 4th order cumulant tensor with application to the blind source separation problem", ICASSP Proc. pp.2655-1658.
- [6] 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.
- [7] P.Comon, C.Jutten and J.Herault, "Blind separation of sources, Part II: Problems statement", Signal Processing, 24, pp.11-20, 1991.
- [8] E.Sorouchyari, "Blind separation of sources, Part I: Stability analysis", Signal Processing, 24, pp.21-29, 1991.
- [9] S.Amori, T.Chen and A.Cichocki, "Stability analysis of learning algorithms for blind source separation", Neural Networks, vol.10, no.8, pp.1345-1351, 1997.
- [10] K.nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output PDF for blind source separation", Proc.4th International Symposium on Independent Component Analysis and Blind signal Separation, ICA'2003, Nara, Japan, April 2003.
- [11] M.Solazzi, F.Piazza and A.Uncini, "Nonlinear blind source separation by spline neural networks", IEEE Proc. ICASSP'2001, Salt Lake City, MULT-P3.4, May 2001.
- [12] F.Milani, M.Solazzi and A.Uncini, "Blind source separation of convolutive nonlinear mixture by flexible spline nonlinear functions", IEEE Proc. ICASSP'2002, Orlando, Florida, pp.1641-1644, May 2002.
- [13] A.Koutras, "Blind separation of non-linear convolved speech mixtures", IEEE Proc. ICASSP'2002, Orlando, Florida, pp.913-916, May 2002.
- [14] A.Ziehe, M.Kawanabe, S.Harmeling and K.R.muller, "Blind separation of post-nonlinear mixtures using Gaussianizing transformations and temporal decorrelation", Proc. ICA'03, Nara, pp.269-274, April 2003.
- [15] R.M.Clements, S.H.Mellado, J.I.Acha, F.Rojas and C.G.Puntone., "MLP-based source separation for MLP-like nonlinear mixtures", Proc. ICA'03, Nara, pp.155-160, April 2003.
- [16] 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, July 2003.
- [17] T.Nishiwaki, K.Nakayama, A.Hirano, "A blind source separation cascading separation and linearization for low-order nonlinear mixtures", IEEE Proc. ICASSP2004, Montreal, pp.V-569-572, May 2004.