非線形混合過程に対する縦続形BSSにおける線形化処理の改善と動作解析

メタデータ	言語: Japanese
	出版者:
	公開日: 2017-10-03
	キーワード (Ja):
	キーワード (En):
	作成者: 竹多, 裕也, 中山, 謙二, 平野, 晃宏
	メールアドレス:
	所属:
URL	http://hdl.handle.net/2297/18229

非線形混合過程に対する縦続形 BSS における線形化処理の改善と動作解析

Improvement of Linearization in Cascade Form BSS for Nonlinear Mixture

竹多 裕也‡　中山 謙二†　平野 晃宏†

Yuya Takeda　Kenji Nakayama　Akihiro Hirano

‡金沢大学大学院　自然科学研究科　電子情報工学専攻
Division of Electronics and Computer Science
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

信号源の統計的性質を利用して、混成信号から信号源を分離する方法をブラインド信号原分離という。実際の混合過程では非線形歪みを生じる場合も多く、分離がより複雑になる。このような場合に対して分離と線形化を縦続に構成する方法が提案されている。本研究ではこの縦続形構成における線形化に対して安定な手法を提案し、従来法より良い性能で線形化できていることを確認した。また、信号分離と結合した場合の性能評価、回路構成の検討を行ない、信号分離前に線形化を精度良く行なうことができれば、センサ数の削減が可能であることがわかった。

ABSTRACT

A cascade form blind source separation (BSS) has been proposed for nonlinear mixture. It consists of a signal separation block and a linearization block in this order from the input to the output. In this paper, a learning process for the linearization block is improved. New ideas are introduced for initialization and adjusting of parameters. Performance of linearization is well improved. Furthermore, by ordering the new linearization block and the signal separation block in this order, the number of sensors can be well reduced. For instance, 3 sensors are enough to 3 sources.

1　まえがき

多くの場合、マイクロホンなどのセンサによって収集される信号は、話者の音声とノイズ等の混成信号である。このとき一般に信号源及び信号源とセンサ間の伝達関数が未知であるが、信号源の統計的性質を利用して信号を分離する手法をブラインド信号源分離 (BSS) という。

実際の混合過程では非線形歪みを生じる場合も多く、混成信号が非線形を含む複雑な形となるので分離がより複雑になり、線形混合に対する BSS[1]~[8] では分離が困難となる。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 (Post-nonlinear 混合) がある。

Post-nonlinear 混合のために，ミラー構造 BSS が主に使用されている [6]．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 [3],[4]。さらに，最尤推定法を用いる方法 [5] 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 [7]。

非線形として 2 次の多項式を考えた場合に、信号源分離と線形化を含む縦続形構成の BSS が提案されている [9]~[11] 。信号源分離の段階で非線形を含む信号を分離し、後の線形化の段階で非線形成分を抑制する構成となっている。

本研究では、この縦続形構成における線形化回路について初期値の決定法、安定な学習法の提案を行ない、その効果を確認する。また、分離回路と結合した場合の回路全体の動作を解析する。簡単化のため、混成信号は

瞬時混合で発生することとする。

2 縦続形 BSS

2.1 回路構成

縦続形 BSS の回路構成を図 1 に示す [9]~[11].

F_iは非線形関数であり、本稿では 2 次まで ($F_i = au_i + bu_i^2$) を対象とする。信号源分離の段階では、信号源 s_i が s_iと s_i^2を含む形で出力される。信号源ごとに分離された後、非線形項 s_i^2が線形化過程で抑圧される。

z_iと z_jに分離される信号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、クロス項 $s_i s_j$は信号源に相関があり分離が不可能であるため消去する必要がある。信号源とセンサ数が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、センサ数を信号源の数より増やし条件を多くする必要がある。よってセンサ数を信号源より多くする構成となっている。

信号源分離の段階において出力される信号はお互いに独立した形になっているが、信号源の高次項が残っている。z_iを線形化回路 G_iに通すことにより線形化を行ない、高次項の抑制を行なう。

また、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 m は式 (1) で求められる。信号源が 2 個だとセンサは 4 個、信号源が 3 個だとセンサは 8 つ必要となる。

$$m = \frac{n(n+1)}{2} + n - 1 \tag{1}$$

2.2 学習アルゴリズム

[9]~[11] 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 段階に分けて行なう。

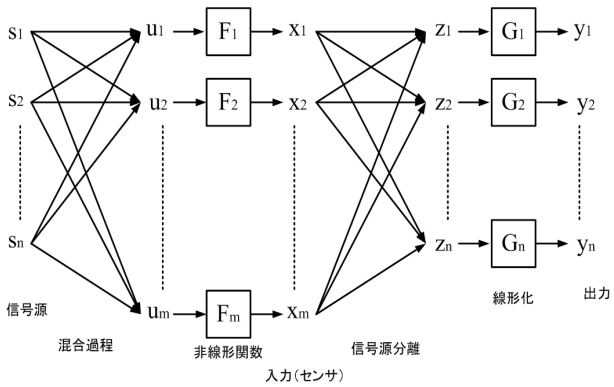

図 1:縦続形 BSS の回路構成

2.2.1 信号源分離

信号源はその高次項を含んでいても互いに独立である。すなわち $s_i + s_i^2$と $s_j + s_j^2$は互いに独立である。したがって。従来の相互情報量を誤差関数とし、自然勾配法を用いることによりこれらを分離することが可能である [8]。その誤差関数、更新式は式 (2)、(3) になる。

$$l(\boldsymbol{z}, \boldsymbol{W}) = -\log|\det(\boldsymbol{W})| - \sum_{i=1}^{m} \log p_i(z_i) \tag{2}$$

$$\boldsymbol{W}(n+1) = \boldsymbol{W}(n) + \eta(n)[\Lambda(n) - f(\boldsymbol{z}(n))\boldsymbol{z}^T(n)]\boldsymbol{W}(n) \tag{3}$$

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

$$f(z) = \tanh(z) \tag{4}$$

2.2.2 線形化

非線形関数 F_iとして 2 次までを想定しているので、信号源分離の出力は式 (4) のような形になっていると仮定できる。

$$z_i = a_i s_i + b_i s_i^2 \tag{5}$$

この形から信号源 s_iだけを出力するために解の公式を用いる。

$$y_i(n) = G_i(z_i) = -\frac{\alpha_i}{2} \pm \sqrt{\frac{\alpha_i^2}{4} + z_i(n)\beta_i} \tag{6}$$

$$\alpha_i = \frac{a_i}{b_i} \tag{7}$$

$$\beta_i = \frac{1}{b_i} \tag{8}$$

信号源として音声や音楽を対象とした場合、これらの波形の平均値はほぼ零である。したがって、非線形特性として 2 次までを考える場合には、線形化回路の出力 $y_i(n)$ 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。

$$E_i(n) = |\frac{1}{M}\sum_{j=0}^{M-1} y_i(n-j)| \tag{9}$$

パラメータ$\alpha_i(n)$, $\beta_i(n)$ は勾配法で学習する.

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

$$\beta_i(n) = \beta_i(n-1) - \eta\frac{\partial E_i(n)}{\partial \beta_i} \quad (11)$$

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

$$\begin{aligned}\frac{\partial E_i(n)}{\partial \beta_i(n)} &= \frac{1}{M}\sum_{i=0}^{M-1}\frac{\partial y_i(n-i)}{\partial \beta_i(n)} \\ &= \frac{1}{M}\sum_{i=0}^{M-1}(\pm\frac{z_i(n-i)}{2}(\frac{\alpha_i^2}{4} \\ &\quad +\beta_i z_i(n-i))^{-\frac{1}{2}}) \quad (13)\end{aligned}$$

解の公式における±の符合の判定において次のような条件を定める。

- 非線形成分より線形成分の方が大きい。
- 波形の振幅が 0.5 以下である。

この条件は実用的にも適用可能である。1 つ目の条件は信号分離後出力である式 (5) において、$a > b$ であるということである。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。本研究ではこの条件が成り立っているものとし、+のみを使うことにする。

3 線形化回路に対する提案

3.1 初期値の設定法

[9]∼[11] では、線形化におけるα、βの初期値の設定法については述べられていない。よってこの初期値の推定法を提案する。

初期値は式 (5) における b の正負、歪みで決定される。a の符号は初期値設定、分離特性には影響しない。なぜなら BSS では定数倍のスケーリングは補正できないものとして残るからである。以下 $b > 0$ として議論する。また、このとき信号はbs^2によって+方向に歪む。この時「正の歪みがある」と呼ぶことにする。

初期値設定の第 1 段階として、振幅の最大値、最小値から a、b を推定することを考える。線形化は一定サンプルの平均値を誤差関数とするため、学習には必ずある程度のサンプルを必要とする。図 2 のように正の歪みによって信号は+に持ち上げられたような形になる。サンプル内の最大値、最小値を z_{max}、z_{min}とすると式のように as、bs^2を推定できる。ここで $a = 1$ と仮定することにより、b を求め初期値とする。

$$as = \frac{|z_{max}| + |z_{min}|}{2} \quad (14)$$

$$bs^2 = \frac{|z_{max}| - |z_{min}|}{2} \quad (15)$$

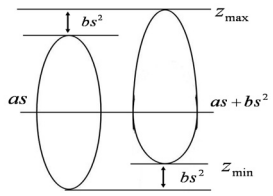

図 2:非線形による波形の歪み

初期値設定の第 2 段階として、準備時間として長い時間を設定できる場合、誤差関数 (信号の平均値 E) を用いて初期値を補正することを考える。第 1 段階で推定した初期値を用いて線形化を行い、誤差関数 E の値を用いてα、βを一定の割合で補正する。式 (6)∼(9) より、$E > 0$ のときαを小さく、βを大きく補正することにより最適値 ($E = 0$) に近付く。また、$E < 0$ のときαを大きく、βを小さく補正することにより最適値に近付くことがわかる。このように操作することにより、学習前にある程度最適値に近付けることができる。

3.2 学習法への提案

[9]∼[11] ではαとβを両方勾配法で学習していたが、この学習率η に差をつけて学習することを提案する。実際は、両方学習する必要はなく、βを適した値に学習することができれば線形化が完了し、定数倍にスケーリングされた s が出力されるのだが、初期値としてのαの設定が適切でないと学習に時間がかかってしまう。よってαも学習した方が収束が早いが、αの変化は学習に大きく影響してしまうのでαの学習率をβの学習率の 1/5 程度に設定することにする。

混合過程や非線形成分が時間とともに変化しない場合は、一定のサンプルを繰り返し用いて学習することにする。これによって、基準となる誤差関数の値 E が歪み以外の要素で変動しなくなるため学習も安定し収束も速くなる。

3.3 線形化後のスケーリング

「線形化→信号分離」において、線形化の出力において定数倍の自由度があるために、分離前に信号のパワー

に差が生じる。これは分離性能にも影響を与える。そこで、線形化後にパワーを揃えるスケーリングを行なう。

3.4 線形化回路の動作確認

提案した手法を用いて、線形化回路の動作確認を行なった。信号源 1 つ、センサ数 1 つ、出力 1 つとし、混合過程は用いずセンサ入力に非線形成分のみを加えた。($F = u + 0.6u^2$) 混合過程、非線形成分は変化しないものとし前半 5000 サンプルだけで、初期値推定、学習を行ない信号全体を線形化した。

初段の初期値推定では $(\alpha, \beta) = (1.37, 1.37)$ と推定され、この時の誤差関数 E=-5.57 となり、非線形成分を強く推定し過ぎていることがわかる。2 段目の補正によって (1.39,1.21) に補正され、学習によって (1.40,1.19) に収束した。E=-0.14 となり誤差を 0 にするように学習できていることが確認できた。図 3、図 4 からも正の歪みがとれて上下対称の形に線形化できていることが確認できた。

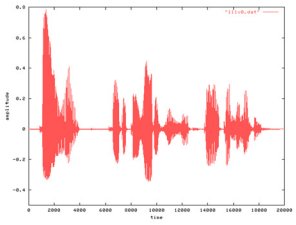

図 3:線形化前の波形

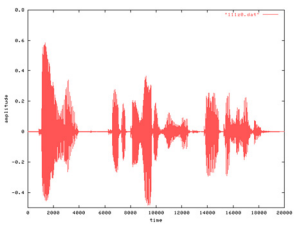

図 4:線形化後の波形

次に同じ条件で [9]~[11] の線形化手法との比較を行なった。評価は 4.1 に示す SIR で行なう。従来法でもある程度の線形化はできているが、提案法では学習がより進み大きな SIR の改善が見られる。

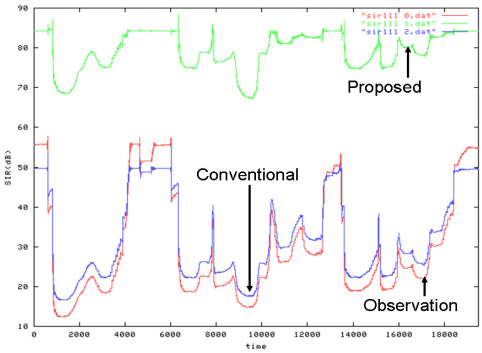

図 5:SIR の比較

4 シミュレーション

分離回路と提案した線形化回路を結合して、回路全体の動作を解析する。縦続形 BSS においては余分な成分を消去するためにセンサ数が増加してしまうことが大きな問題であり、センサ数の削減が課題の 1 つに挙げられる。よって、非線形成分、回路構成、センサ数を考慮しシステムの性能を調べる。

シミュレーションの結果を式 (16) で求める SIR(dB) で評価する。

$$SIR = \log_{10} \sum_{i=0}^{2} \sum_{i=0}^{M-1} \frac{(s_{n-i})^2}{(s_{n-i} - \gamma y_{n-i})^2} \tag{16}$$

ここで s_iは信号源、y_iは出力信号である。またγは振幅の差を補正する係数であり勾配法で分母が最小になるように求める。

4.1 回路構成と非線形成分の影響

今回用いている Post-nonlinear 混合モデルにおいては、式 (17) のようにセンサ入力が信号のグループの 2 次式という形になっているので分離の前に線形化を行なう構成を適用することができる。

$$u_j = a_{j1}s_1 + a_{j2} + \ldots + a_{jn}s_n \tag{17}$$

$$x_j = u_j + u_j^2 \tag{18}$$

「信号分離→線形化」と「線形化→信号分離」の 2 つの構成について動作を調べる。この際、非線形の大きさを変えて、歪みが分離に与える影響を調べる。

「線形化→信号分離」のシミュレーションにおいて、線形化後のスケーリングの効果を調べる。図 7 のようにスケーリングを行なうことにより大きく性能が向上する。

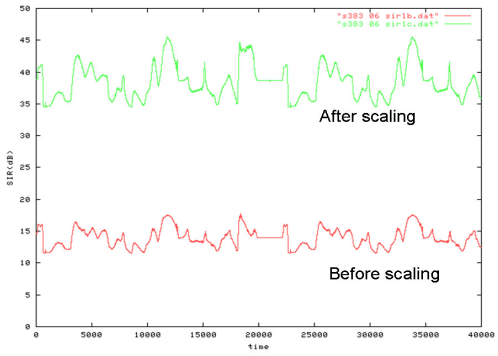

図 7:スケーリングの効果

- 信号源 3 つ、センサ数 8 つ、出力 3 つ
- 非線形歪み $b = 0.4, 0.6, 0.8$
- サンプル数 40000

下にシミュレーション結果を示す。図中の数字は歪み b の大きさである。図 8 においては非線形成分が大きくなるほど分離性能が低下しているのがわかる。図 9 においては図 8 の結果に基づいて、線形化によって SIR が 7dB 程度向上しているのが分かる。図 10 においては歪みの大きさに関係なく性能が一定となった。分離前に完全に線形化に成功していると、2 次項とクロス項は消去され、分離回路には線形成分のみが入力される。この結果より分離に非線形成分が影響していない、非線形成分が消去できていることが確認できる。

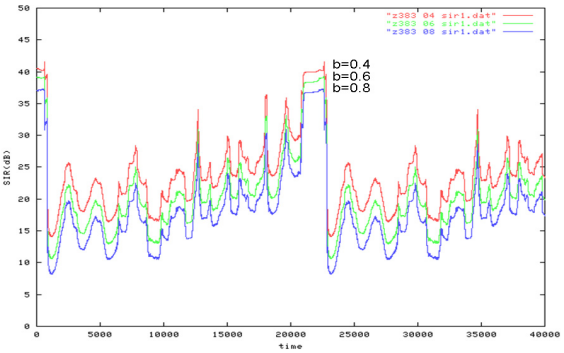

図 8:信号分離後の SIR

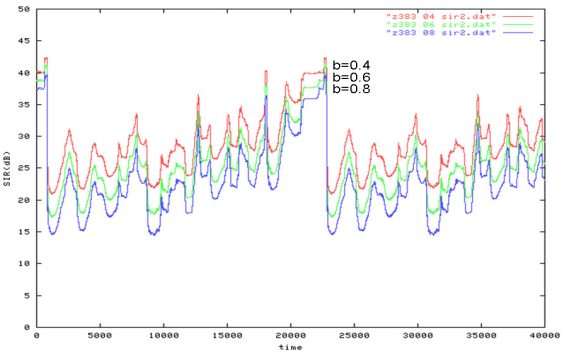

図 9:「信号分離→線形化」の SIR

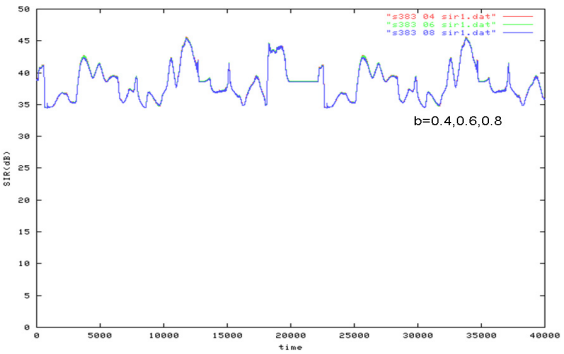

図 10:「線形化→信号分離」の SIR

4.2 センサ数低減の影響

「信号分離→線形化」においては分離できない成分のために性能が落ちると考えられる。「線形化→信号分離」においては信号分離前に非線形成分を十分に低減できるならば、センサ数が低減できると考えられる。以上のことを歪みが小さい場合と大きい場合 (b=0.4,0.8) に分けてシミュレーションを行った。

シミュレーション結果を以下に示す。図 11 から、「信号分離→線形化」はセンサ数が 8、7、6 個のときはほぼ同じ性能であり、5 個以下で大きく性能が低下していることがわかる。よって非線形成分が小さいときはセンサ数を少々削減しても分離特性に影響を与えないことがわかる。

図 12 は非線形成分が大きい例である。センサが多い方が性能が良いが、あまり差が見られない結果になった。非線形成分が大きいときはセンサ数を確保していても、分離の性能が落ち、そのため線形化の性能も下がっていると考えられる。

図 13,14 において、「線形化→信号分離」では非線形成分に関係なくセンサ数 8、7、6、3 のとき約 35dB 以上、センサ数 5、4 のときも 25dB 以上の分離に成功している。

分離前に線形化を精度良く行なうことができればセンサ数を信号源の数まで削減しても良好な特性が得られることが分かる。

5 まとめ

非線形混合過程における縦続形 BSS の線形化回路の学習法として新しい方法を提案し、従来法よりも良い性能で線形化ができていることを確認した。信号分離と結合した場合の性能を評価し、信号分離の前段に線形化を配置することによりセンサ数を削減できることが可能であることがわかった。

今後は遅延を含む混合過程への適用を検討する。

本研究の一部は科学研究費補助金（基盤研究 (C)17560335 ）による。

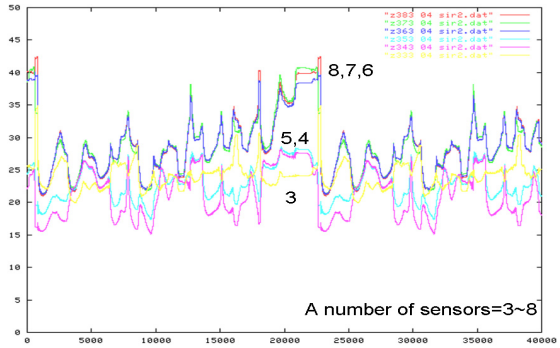

図 11:「信号分離→線形化」$b = 0.4$

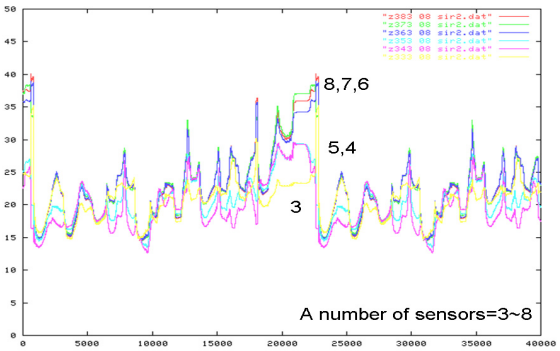

図 12:「信号分離→線形化」$b = 0.8$

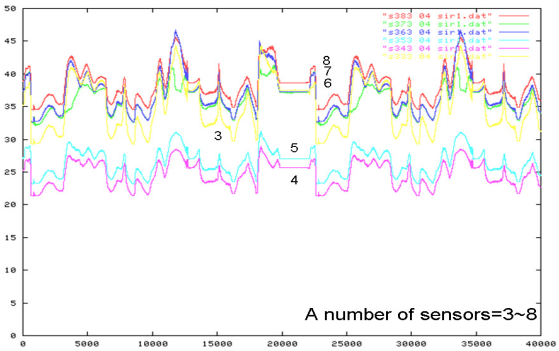

図 13:「線形化→信号分離」$b = 0.4$

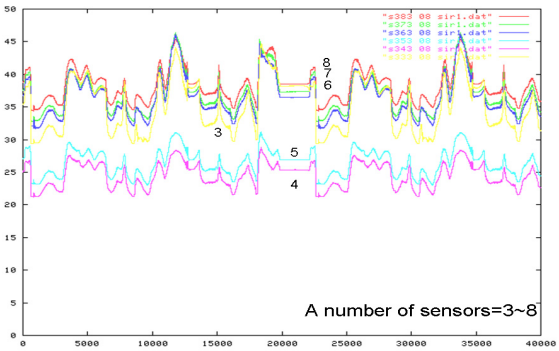

図 14:「線形化→信号分離」$b = 0.8$

参考文献

[1] C.Jutten and Jeanny Herault, "Blind separation of sources,Part I:Anadaptive algorithm based on neuromimetic architecture", SignalProcessing,24, pp.1-10,1991.

[2] P.Comon,C.Jutten and J.Herault, "Blind separation of sources, Part Ⅱ:Problems statement", Signal Processing,24,pp.11-20,1991.

[3] 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.

[4] 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.

[5] A.Koutras ,"Blind separation of non-linear convolved speech mixtures", IEEE Proc.ICASSP'2002, Orlando, Florida,pp.913-916,May 2002.

[6] 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.

[7] 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.

[8] S.Amori,T.Chen and A.Cichocki,"Stability analysis of learning algorithms for blind source separation",Neural Networks,vol.10,no.8,pp1345-1351,1997.

[9] 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.

[10] 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.

[11] 西脇 貴之、中山 謙二、平野 晃宏 "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"、第 19 回信号処理シンポジウム、山梨、2004.11