

# 畳み込み混合過程に対するブラインド信号源分離の学習法に関する研究

メタデータ	言語: jpn 出版者: 公開日: 2017-10-05 キーワード (Ja): キーワード (En): 作成者: メールアドレス: 所属:
URL	<a href="http://hdl.handle.net/2297/26700">http://hdl.handle.net/2297/26700</a>

氏名	堀田 明秀
学位の種類	博士(工学)
学位記番号	博甲第862号
学位授与の日付	平成19年3月22日
学位授与の要件	課程博士(学位規則第4条第1項)
学位授与の題目	畳み込み混合過程に対するブラインド信号源分離の学習法に関する研究
論文審査委員(主査)	中山 謙二(自然科学研究科・教授)
論文審査委員(副主査)	船田 哲男(自然科学研究科・教授), 西川 清(自然科学研究科・教授), 堀田 英輔(自然科学研究科・講師), 平野 晃宏(自然科学研究科・講師)

## 1 Abstract

In this thesis, learning algorithms for convolutive Blind Source Separation (BSS) are proposed, which can suppress effects of reverberation in a mixing process and signal distortion in a separation block. Furthermore, two types of BSS systems, including Feedback (FB-) and Feedforward (FF-), are compared.

In the learning algorithm for FB-BSS, a condition on transmission delay in the mixing process is required. However, when reverberations occur, this condition is not satisfied, and separation performance is degraded. An exponentially weighted step-size is proposed to stabilize the learning process for the reverberations. Furthermore, the step-size is automatically adjusted by approximating an envelop of the FIR filter coefficients. Through simulations, good separation performance, which is the same as in no reverberations condition, can be achieved by the proposed method.

FF-BSS systems have some degree of freedom in the solution space. Therefore, signal distortion is likely to occur. A criterion for the signal distortion is discussed. Properties of conventional methods proposed to suppress the signal distortion are analyzed. A general condition for complete separation and distortion free is derived for multi-channel FF-BSS systems. This condition is incorporated in learning algorithms as a distortion free constraint. Computer simulations using speech signals and stationary colored signals demonstrate that the proposed method can well suppress signal distortion, while maintaining a high source separation performance.

## 2 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, 信号や雑音, 伝達過程についての十分な情報が得られない場合が多い。このような状況では, 信号の統計的な性質に基づいて信号を分離するブラインド信号源分離(BSS)は重要である。

これまでBSSに関して, フィードバック形(FB-)やフィードフォワード(FF-)といった分離回路が提案されている。また, 時間領域, 周波数領域に対してさまざまな学習法が提案されている。

多くの場合, 混合過程に畳み込みが生じる。畳み込み混合過程に対しては, 反響音が生じ, 信号の分離が困難になる場合がある。また, 畳み込みにより信号の周波数特性が変化し, 信号歪みが生じる可能性がある。

本研究では, 第一にFB-BSSにおける反響音の影響

を解析し, 反響音を抑制する学習法を提案する。次に, FF-BSSに対して起こる信号歪みを抑制する学習法を提案する。最後に両者を比較し, それぞれの特徴を解析する。

## 3 FB-BSSの構成

### 3.1 回路構成と入出力関係

Juttenらが提案した回路構成を図1に示す[1]。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

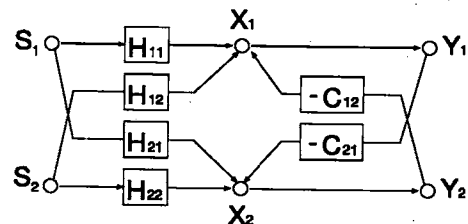


図1: FB-BSSの回路構成

して観測される。

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

また, 分離過程の出力  $y_j(n)$  は次式で表される。

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

これらを  $z$  変換し, ベクトルと行列で表すと次のようになる。

$$\mathbf{X}(z) = \mathbf{H}(z)\mathbf{S}(z) \quad (3)$$

$$\mathbf{Y}(z) = \mathbf{X}(z) - \mathbf{C}(z)\mathbf{Y}(z) \quad (4)$$

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。

$$\mathbf{Y}(z) = (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{H}(z)\mathbf{S}(z) \quad (5)$$

$$= \mathbf{A}(z)\mathbf{S}(z) \quad (6)$$

### 3.2 学習アルゴリズム

各センサが十分に離れている場合、遅延差の関係から因果性を満たす回路ではセンサに一番近い信号源を削除することができない。このため出力のパワーを最小化して信号を分離する方式が提案されている [4],[9]。\$c\_{jk}(n, l)\$ の更新式は次式で与えられる。

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \Delta c_{jk}(n, l) \quad (7)$$

$$\Delta c_{jk}(n, l) = \mu f(y_j(n))g(y_k(n-l)) \quad (8)$$

ここで、\$f(y\_j(n))\$、\$g(y\_k(n-l))\$ は異なる奇関数である。

## 4 反響音による影響の軽減法の提案

### 4.1 FB-BSS における反響音の影響

FB-BSS では遅延の関係から因果性を満たす回路では 1 つの成分を消すことができないことを利用して出力の信号パワーを最小化することによって信号を分離している。しかし、反響音を含む場合、直接音に比べ反響音が遅れてセンサに入力されるため、本来消すことのできない信号成分が消えてしまい、最適な解から遠ざかってしまう。そのため FB-BSS は反響音を含む場合には分離性能が低下してしまう。

### 4.2 固定形 EW ステップサイズ

反響音の影響はタップの後半部分に現れてくる。後半部分の更新を抑える必要がある。これは FIR フィルタの遅延線にあわせてステップサイズ \$\mu\$ を制御することによって実現できる。分離回路の最適値の包絡線は指数的に減衰する。したがって、フィルタ係数更新も理想形に比例した形で行なうこととする [4],[9]。

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \mu(l)f(y_j(n))g(y_k(n-l)) \quad (9)$$

$$\mu(l) = \mu_0 r^l, \quad 0 < r < 1 \quad (10)$$

ステップサイズ \$\mu(l)\$ は理想解に比例するように決める。

### 4.3 適応形 EW ステップサイズ

固定形 EW ステップサイズの方式では指数の底を部屋のインパルス応答から推定しなければならない。実際には、事前に部屋のインパルス応答を推定することは難しく、実用的ではない。適応的に指数の値を推定する方法を提案する [4],[9]。推定法の基本的な考え方は、次のようになる。分離回路の係数の対数とり、最小 2 乗誤差を与える直線を求める。これらの冪乗を取ることによりステップサイズで使用する指数の底を求める。まず、最小 2 乗法により最小 2 乗誤差を与える直線を求める。

$$\mathbf{x}(n) = \mathbf{A}^+ \mathbf{b}(n) \quad (11)$$

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \quad (12)$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & K_c - 1 \end{bmatrix}^T \quad (13)$$

$$\mathbf{b}(n) = \log |c_{jk}(n)| \quad (14)$$

この結果から、\$r(n)\$、\$\mu\_0(n)\$ とステップサイズ \$\mu(n, l)\$ は以下ようになる。

$$\mu_0(n) = e^{x_1} \quad (15)$$

$$r(n) = e^{x_2} \quad (16)$$

$$\hat{r}(n) = \alpha r(n) + (1 - \alpha) \hat{r}(n-1) \quad (17)$$

$$0 < \alpha \ll 1$$

$$\mu(n, l) = \mu_0(n) \hat{r}^l(n) \quad (18)$$

指数の底は徐々に更新される。また、\$\hat{r}(n)\$ の初期値は 1 とする。

## 5 FF-BSS の構成

### 5.1 回路構成と入出力関係

FF-BSS の回路構成を図 2 に示す。観測信号は式 (1)、式 (3) と表せる。

分離過程の出力 \$y\_j(n)\$ は次式で表される。

$$y_j(n) = \sum_{k=1}^N \sum_{l=0}^{K_w-1} w_{jk}(n, l) x_k(n-l) \quad (19)$$

これらを \$z\$ 変換し、ベクトルと行列で表すと次のようになる。

$$\mathbf{Y}(z) = \mathbf{W}(z) \mathbf{X}(z) \quad (20)$$

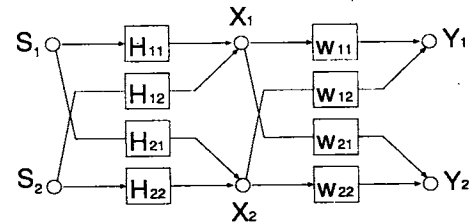


図 2: FF-BSS の回路構成

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。

$$\mathbf{Y}(z) = \mathbf{W}(z) \mathbf{H}(z) \mathbf{S}(z) \quad (21)$$

$$= \mathbf{A}(z) \mathbf{S} \quad (22)$$

### 5.2 学習アルゴリズム

学習には出力の相互情報量を最小化するアルゴリズムを使用する。時間領域、周波数領域での更新式はそれぞれ以下ようになる。

$$w(n+1, l) = w(n, l) + \eta \sum_{q=0}^{K_w-1} [\mathbf{I} \delta(n-q) - \langle \Phi(\mathbf{y}(n)) \mathbf{y}^T(n-l+q) \rangle] w(n, q) \quad (23)$$

$$\mathbf{W}(r+1, m) = \mathbf{W}(r, m) + \eta [\text{diag}(\langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle) - \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \quad (24)$$

ここで \$r\$ は FFT のフレーム番号、\$m\$ は周波数点を表す。

## 6 信号歪みの評価

### 6.1 信号歪みの基準

実用的には信号歪みの基準を信号源とした方がよい。しかし、BSSでは信号を独立するように学習を行ない、混合過程を推定するわけではない。したがって、観測信号における信号源の成分を基準とし、BSSによって引き起こされた信号歪みを対象とする。

### 6.2 完全分離と無歪みの条件

式(6)、(22)において、各行各列に1つだけ0でない要素を持つとき、完全分離しているといえる。また、前節で定めた基準より0でない要素が $H_{ij}(z)$ であったとき無歪みであるといえる。

FB-BSSでは完全分離状態の $A_{ij}(z)$ は $H_{ij}(z)$ となるため信号を分離すると共に信号歪みを抑制している。一方、FF-BSSの完全分離と無歪みの条件はそれぞれ信号分離： $A(z)$ の非対角要素が0

$$A_{12}(z) = W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (25)$$

$$A_{21}(z) = W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) = 0 \quad (26)$$

無歪み： $A(z)$ の対角要素が $H_{ii}(z)$ 。

$$A_{11}(z) = W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) = H_{11}(z) \quad (27)$$

$$A_{22}(z) = W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z) = H_{22}(z) \quad (28)$$

となる。従来の方式では式(25)と(26)のみをコスト関数としている。信号分離の方程式が2つであるが、変数となる分離回路のパラメータが4つであり、不定解となる。これは出力の周波数を変化させて出力をより独立にしようとして信号歪みが生じる可能性が高い。

## 7 信号歪み抑制学習法の提案

今節では、FF-BSSで起こる信号歪みを抑える学習法を提案する[5]-[8]。

### 7.1 完全分離かつ無歪みの条件

完全分離かつ無歪みの条件は以下ようになる。

$$W(z)H(z) = A(z) \quad (29)$$

$$A(z) = \text{diag}[H(z)] \quad (30)$$

ここから、混合過程の伝達関数 $H(z)$ を削除していくと

$$\text{diag}[(W^{-1}(z) - I)] = 0 \quad (31)$$

となり、この方程式を解くことで完全分離と無歪みの関係が得られる。しかし、この式を解くのは複雑なので以下の近似式を考える。

$$W_{jj}(z) = 1 + w_{row,j}^T(z)M_{jj}^{-1}(z)w_{col,j}(z) \quad (j=1, \dots, N) \quad (32)$$

$M_{jj}(z)$ は $W(z)$ から $j$ 行 $j$ 列を除いた小行列を表し、 $w_{col,j}(z)$ と $w_{row,j}(z)$ は以下の式で与えられる。

$$w_{col,j}(z) = [W_{1j}(z), W_{2j}(z), \dots, W_{Nj}(z)]^T \quad (33)$$

$$w_{row,j}(z) = [W_{j1}(z), W_{j2}(z), \dots, W_{jN}(z)]^T \quad (34)$$

ただし、いずれも $W_{jj}(z)$ を含まない。式(32)は $M_{jj}(z)$ に $W_{jj}(z)$ を含むため式(31)を陽に解いたものではない。

### 7.2 完全分離かつ無歪みの条件と学習アルゴリズムの融合

信号歪みと無歪みの条件は $W_{jj}(z)$ と $W_{jk}(z)$ の関係を表しているだけであり、最適値が得られている訳ではない。したがって、信号分離の学習法と組み合わせる必要がある。提案法の基本的な考え方は、まず $W_{jj}(z)$ や $W_{jk}(z)$ を従来法で更新し、その後、式(32)の条件により $W_{jj}(z)$ を補正する。

提案するアルゴリズムは以下ようになる。

**Step 1:** 式(24)を使用する従来の学習法で $W_{jj}(r, m)$ や $W_{jk}(r, m)$ を更新し、 $W_{jj}(r+1, m)$ や $W_{jk}(r+1, m)$ とする。

**Step 2:**  $W_{jj}(r+1, m)$ を式(32)で計算し、 $\tilde{W}_{jj}(r+1, m)$ とする。

**Step 3:**  $W_{jj}(r+1, m)$ を以下のように補正する。

$$W_{jj}(r+1, m) = (1 - \alpha)W_{jj}(r+1, m) + \alpha\tilde{W}_{jj}(r+1, m) \quad (0 < \alpha \leq 1) \quad (35)$$

前述の通り、式(32)はすべての $W_{jj}(r+1, m)$ 、 $j=1, \dots, N$ に対して陽に解いているわけではない。しかし、 $W_{kk}(z)$ の更新が非常に小さいため $M_{jj}^{-1}(z)$ 内の $W_{kk}(z)$ は定数とみなせる。

時間領域で学習を行なう際も同様の方法で行なうが、Step 2の計算を時間領域で行なうと計算量が大きいため、一旦、周波数領域に変換して計算を行ない、時間領域に戻して分離学習と組み合わせる。

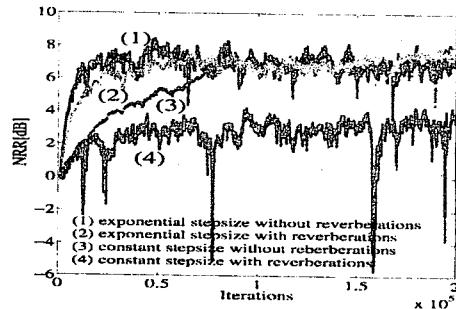


図 3: 白色信号源に対する NRR

### 7.3 2チャンネルにおける完全分離かつ無歪みの条件

2チャンネルにおいては式(31)の条件は簡単に陽に解ける。式(31)は以下のように表せる。

$$W_{jj}^2(z) - W_{jj}(z) - W_{jk}(z)W_{kj}(z) = 0 \quad (36) \quad j=1, 2, k=1, 2, j \neq k$$

$W_{jj}(z)$ について解くと以下ようになる。

$$W_{jj}(z) = \frac{1 \pm \sqrt{1 + 4W_{12}(z)W_{21}(z)}}{2}, \quad j=1, 2 \quad (37)$$

2チャンネルの場合にはこちらを学習に組み込む。

## 8 シミュレーション

### 8.1 シミュレーション条件

8.2節ではFB-BSSを、8.3節ではFB-BSS,FF-BSSに関して議論する。信号分離の評価として出力信号の抽出信号と干渉信号のパワーの比から観測信号の抽出信号と干渉信号のパワーの比を引いた $NRR[dB]$ と、出力信号の抽出信号と干渉信号のパワーの比である $SIR[dB]$ を使用する。また、信号歪みの評価には基準信号と出力信号の差を基準信号で正規化した $SD[dB]$ を用いる。 $NRR$ ,  $SIR$ は値が大きい方が、 $SD$ は値が小さい方が性能が高い。

### 8.2 反響音による影響の軽減法

2チャンネルにおいて、反響音がある場合、ない場合においてそれぞれ一定値ステップサイズ、固定形指数ステップサイズを使ったときの $NRR$ を図3に示す。分離回路のタップ数は1024とした。固定形指数ステップサイズは理想フィルタ係数の包絡線から求めている。

反響音がない場合、指数ステップサイズの収束は速いが、最終的な $NRR$ は指数ステップサイズも一定値ステップサイズも同様である。この場合は、指数ステップサイズにより学習の高速化が達成されている。

反響音がある場合、先ほどと同様に収束の高速化の効果は認められる。しかし、指数ステップサイズは一定値ステップサイズに比べて収束が速いのに加え、最終的な $NRR$ が高くなっている。これは反響音がある場合、指数ステップサイズは高速化のみではなく、反響音によるフィルタ係数の変動を抑えることにより、分離性能の改善も達成していることを示している。さらに、反響音がある場合でも、指数ステップサイズにより、反響音がない場合と同じ分離性能が得られており、

表 1: 音声に対する各学習法の性能比較

Methods	2ch		3ch	
	$SIR$	$SD$	$SIR$	$SD$
FF-TIME	12.2	0.25	13.3	0.50
FF-TIME (MDP)	3.98	-10.3	6.25	-8.02
FF-TIME (DF)	8.33	-12.1	8.00	-14.0
FF-FREQ	13.0	-9.43	17.2	-13.4
FF-FREQ (PB)	11.6	-16.9	16.3	-17.3
FF-FREQ (DF)	11.8	-24.6	16.1	-24.1
FB	14.1	-14.5	13.6	-15.4

フィルタ係数の学習において反響音の影響がほとんど抑制されていることが分かる。また、本稿では図示しないが適応形指数ステップサイズ方式は固定形とほぼ同等の分離性能が得られる。

### 8.3 信号歪み抑制学習法

表1にさまざまな手法の信号分離、信号歪みの評価をまとめる。時間領域の学習におけるフィルタのタップ数、周波数領域の学習におけるFFTサイズは256とする。FF-TIME,FF-FREQはそれぞれ時間領域、周波数領域で学習する信号歪み抑制法を用いない従来の方式である。FF-TIME (DF),FF-FREQ(DF)は提案する信号歪み抑制学習法であり、FF-TIME(MDP), FF-FREQ(PB)はそれぞれ[2],[3]で提案されている信号歪み抑制学習法である。FBは時間領域で学習するFB-BSSである。時間領域での学習に関して、FF-TIMEは信号歪みの評価が悪い。一方、信号分離の評価は高い

がこれは信号歪みによって過剰に見積もられたものだと考えられる。FF-TIME (MDP)は信号歪みが改善されている。しかし、信号分離の性能が低い。それに対し、提案した信号歪み抑制学習法FF-TIME (DF)は高い分離性能を保ったまま信号歪みを抑制している。一方、FBは信号分離、信号歪みに対して高い性能が得られている。

周波数領域での学習に関してFF-FREQ (PB), 提案した信号歪み抑制学習法FF-FREQ (DF)ともにFF-FREQと同等の分離性能を保ったまま、信号歪みを改善している。しかし、FF-FREQ (PB)に比べてFF-FREQ (DF)の改善が大きい。

### 8.4 FB-BSSとFF-BSSの性能比較

最後に混合過程に基づくFB-BSSとFF-BSSの性能比較を行なった。図4にFF-BSSのFF-TIME(DF)とFB-BSSの $SIR$ を示す。混合過程の伝達関数 $H_{ii}(z)$ ,  $H_{ij}(z)$ の遅延差 $\tau$ を変化させている。この図よりFF-BSSは伝達遅延の差に影響を受けない。一方、FB-BSSは遅延差に大きく影響を受け、遅延差が小さいところでは $SIR$ の値が悪くなっている。この遅延差は観測点 $X_J$ ,  $X_K$ の距離が非常に小さくなることによって十分に小さくなることが考えられる。そこで、混合過程の伝達遅延を基にFB-BSSとFF-BSSの特徴を解析した。FB-BSSは観測点が十分に離れて設計できるとき有効である。一方、FF-BSSはシステムを小型化したいときに有効である。

## 9 結論

FB-BSSにおける反響音の影響を解析し、反響音を抑制する固定形EWステップサイズ方式を提案した。また、指数の底を自動的に制御する適応EWステップ

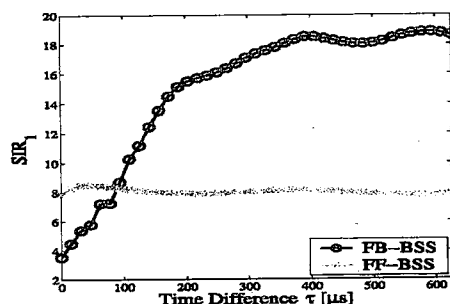


図 4: 音声における伝達遅延による $SIR$ の変化

サイズ方式を提案した。また、BSSにおける信号歪みについて解析を行ない、FF-BSSに対して起こる信号歪みを抑制する学習法を提案した。最後にFB-BSSとFF-BSSを遅延に基づいて比較した。

## 参考文献

- [1] C.Jutten and Jeanny Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture", Signal Processing, 24, pp.1-10, July 1991.
- [2] K.Matsuoka and S.Nakashima, "Minimal distortion principle for blind source separation", Proc. ICA2001, San Diego, CA, USA, pp.722-727, Dec. 2001.

- [3] N. Murata, S. Ikeda and A. Ziehe, "An approach to blind source separation based on temporal structure of speech signals", Neurocomputing, Vol. 41, pp. 1-24, Oct. 2001.
- [4] A. Horita, K. Nakayama and A. Hirano, "A blind source separation with exponentially weighted stepsize and its convergence analysis in convolutive mixtures with reverberations", Proc. IFAC Workshop on Adaptation and Learning Control and Signal Processing, Yokohama, Japan, pp.523-528, Aug. 2004.
- [5] A. Horita, K. Nakayama, A. Hirano and Y. Dejima, "A distortion free learning algorithm for feedforward BSS and its comparative study with feedback BSS", IEEE&INNS, Proc. IJCNN2006, Vancouver, pp.7642-7649, July 2006.
- [6] A. Horita, K. Nakayama, A. Hirano and Y. Dejima, "A learning algorithm with distortion free constraint and comparative study for feedforward and feedback BSS", Proc. EUSIPCO2006, Florence, Italy, CD-ROM, Sept. 2006.
- [7] 堀田明秀, 中山謙二, 平野晃宏 "畳み込み形多チャンネル BSS における信号歪み抑制学習法及び回路形式に基づく考察", 第 21 回信号処理シンポジウム (京都), C4-1, CD-ROM, 2006.11.
- [8] A. Horita, K. Nakayama and A. Hirano, "A distortion free learning algorithm for feedforward multi-channel blind source separation", IEICE Trans. Fundamentals, 投稿中
- [9] 堀田明秀, 中山謙二, 平野晃宏, "反響音を有する畳み込み形混合過程に対するブラインドソースセパレーションの学習法", 信号処理, 投稿中

## 学位論文審査結果の要旨

平成19年1月31日に開催された第1回学位論文審査委員会, 及び平成19年2月2日に行われた口頭発表と第2回学位論文審査委員会で審査した結果, 以下の通り判定した。

本論文では, 複数の信号が伝達・混合され, 観測された信号から元の信号源を分離する問題を扱っている。特に, 信号源や混合過程に関する情報がない場合に, 信号源の独立性を仮定して信号源分離を行うブラインド信号源分離 (Blind Source Separation: BSS) を研究対象としている。混合過程が周波数特性を有する畳み込み形の場合は, 反響音により BSS における分離回路の学習が遅くなったり, 不安定になるという問題がある。これに対して, 本論文では反響音の影響を抑える新しい学習法を提案している。分離回路におけるフィルタ係数の更新を制御するステップサイズを自動的に最適化する方式である。シミュレーションによれば, 学習の高速化と分離特性の向上が達成されている。次に, 分離回路で発生する信号歪みを抑制する学習法を提案している。信号源分離と信号無歪みという2つの条件を同時に満たす制約条件を求め, これにより分離回路の学習を制御する方法である。シミュレーションにより従来法に比べて信号源分離と信号歪みの双方で良好な特性を実現している。最後に, フィードバック形 BSS とフィードフォワード形 BSS の比較を行い各々が有効に適用できる条件を明らかにしている。

以上の研究成果は, 畳み込み混合過程に対する BSS の学習に関する重要な問題を解決するものであり, BSS の実用化にも貢献しており, 学術的及び実用的価値が高く, 博士論文に値するものと判定する。