

論文

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

堀田 明秀, 中山 謙二, 平野 晃宏

金沢大学大学院 自然科学研究科, 石川県金沢市角間町

A Learning Algorithm for Blind Source Separation of Convolutive Mixtures with Reverberations

Akihide Horita, Kenji Nakayama and Akihiro Hirano

Graduate School of Natural Science & Technology, Kanazawa University

Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan

E-mail: horita@leo.ec.t.kanazawa-u.ac.jp, nakayama@t.kanazawa-u.ac.jp

Abstract An stable learning algorithm is proposed for a feedback blind source separation applied to convolutive mixtures. First, convergence property is analyzed. A condition on transmission delay times in the mixture, under which delay time of direct paths are less than that of cross paths, is required for stable convergence of learning FIR filters in a separation block. However, reverberations in a room, where the BSS system is used, have long transmission delay times, which do not satisfy the above condition. In this paper, an exponentially weighted step-size is proposed to stabilize the learning process for the reverberations. The corrections of the FIR filters are exponentially scaled along the tap number. 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.

Keywords: blind source separation, convolution, reverberations, stepsize, convergence

1. はじめに

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, 信号や雑音についての十分な情報が得られない場合が多い。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とも多い。このような状況で, 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ソースセパレーション (BSS) は重要である [1]-[13]。

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(Finite Impulse Response) フィルタもしくは IIR(Infinite Impulse Response) フィルタが必要になる。畳み込み混合過程に対する BSS として, これまでに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, 分離性能が低いなどの問題があった [14]-[18]。これらの問題は混合過程に反響音を含む場合はさらに顕著になり, 良好な BSS を実現することが難しくなっている。

本稿では畳み込み混合のモデルを時間領域で処理す

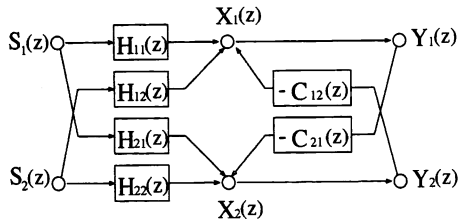


図1 フィードバック形ブラインドソースセパレーションのブロック図

Fig. 1 Block diagram of feedback blind source separation

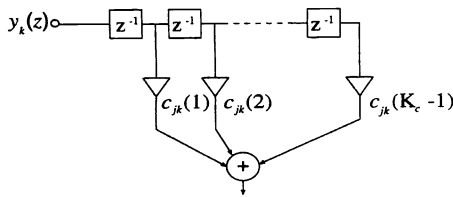


図2 $C_{21}(z)$ や $C_{12}(z)$ に用いる FIR フィルタ

Fig. 2 FIR filter used for $C_{21}(z)$ and $C_{12}(z)$ in feedback

るフィードバック形 BSS において, Jutten らが提案した学習アルゴリズム [6], [14] に対し, 混合過程の伝達遅延の観点から解析を行なう。それに基づき, 学習アルゴリズムに対し, 反響音の影響を抑制するため指数変化するステップサイズによる学習アルゴリズムを導入する。さらに, 指数重みを自動的に調整する方法を提案する。最後に, シミュレーションにより提案法の有効性を検討する。

2. 回路構成と入出力関係

2.1 回路構成

Jutten らが提案したフィードバック形 (FB-)BSS の回路構成 [6], [14] を図 1 に示す。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。信号源, センサと出力の数はすべて同じであるとする。

2.2 時間領域の入出力関係

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

$x_j(n), j = 1, 2, \dots, N$ として観測される。

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

ここで, n は時刻, m はインパルス応答 $h_{ji}(m)$ のサンプル番号, K_h は混合過程のインパルス応答の長さであり, 伝達関数を FIR フィルタとしたときのタップ数である。また, 分離過程の出力 $y_j(n)$ は次式で表される。

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

ここで, K_c は分離回路における FIR フィルタのタップ数を表す。この関係をベクトルと行列で表すと次のようになる。

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

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

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

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

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

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

$$\tilde{\mathbf{y}}(n) = [y_1(n), y_2(n), \dots, y_N(n)]^T \quad (9)$$

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

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

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

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

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

式 (14) の最初の 0 は z^0 の係数に相当する。式 (2) では l が 1 から始まっていることに相当する。 \mathbf{C} の対角要素が零であることは, $y_i(n)$ が $x_i(n)$ にフィードバックされないことに相当する。式 (4) の右辺第 2 項の計算では, 式 (14) の \mathbf{c}_{jk} と式 (10) の $\mathbf{y}_k(n)$ が畳み込まれる。要素の数は K_c 個で一致している。 \mathbf{C} は $N \times NK_c$

行列で、 $\tilde{\mathbf{y}}(n)$ は NK_c 次元ベクトルであり、 $\mathbf{C}\tilde{\mathbf{y}}(n)$ は N 次元ベクトルとなる。

2.3 z 領域の入出力関係

$s_i(n)$, $x_j(n)$, $y_k(n)$ 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。

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

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

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

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

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

$$\mathbf{H}(z) = \begin{bmatrix} H_{11}(z) & H_{12}(z) & \dots & H_{1N}(z) \\ H_{21}(z) & H_{22}(z) & \dots & H_{2N}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}(z) & H_{N2}(z) & \dots & H_{NN}(z) \end{bmatrix} \quad (20)$$

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

ここで、 $\mathbf{H}(z)$, $\mathbf{C}(z)$ はそれぞれ、式 (11) のインパルス応答 h_{ji} , 式 (13) のフィルタ係数 c_{jk} の z 変換を要素とする行列である。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。

$$\begin{aligned} \mathbf{Y}(z) &= (\mathbf{I} + \mathbf{C}(z))^{-1}\mathbf{X}(z) \\ &= (\mathbf{I} + \mathbf{C}(z))^{-1}\mathbf{H}(z)\mathbf{S}(z) \end{aligned} \quad (22)$$

ここで分離行列 $\mathbf{W}(z)$ を以下のように定義する。

$$\mathbf{W}(z) = (\mathbf{I} + \mathbf{C}(z))^{-1} \quad (23)$$

信号源から出力までの伝達関数を表す行列を次のように定義する。

$$\mathbf{P}(z) = \mathbf{W}(z)\mathbf{H}(z) \quad (24)$$

もし、 $\mathbf{P}(z)$ の各行各列が 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 $y_k(n)$ に完全に分離される。2 チャンネルの例を示す。

$$\mathbf{P}(z) = \begin{bmatrix} P_{11}(z) & 0 \\ 0 & P_{22}(z) \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 0 & P_{12}(z) \\ P_{21}(z) & 0 \end{bmatrix} \quad (25)$$

3. 学習アルゴリズム

反響音を含む混合過程に対する学習アルゴリズム導入のために、本節では伝搬遅延時間に着目して学習アルゴリズムを導く [19], [20]。

3.1 伝搬遅延に関する仮定

単純化のため 2 チャンネルにおいて考える。 $H_{11}(z)$ や $H_{22}(z)$ の遅延は $H_{21}(z)$ や $H_{12}(z)$ に比べて短いものとする。これは図 1 にあるように $X_1(z)$ のセンサが $S_1(z)$ の近くに、 $X_2(z)$ のセンサが $S_2(z)$ の近くに配置されていることを意味する。

分離が完全に行なわれるとき、 $C_{jk}(z)$ と $y_k(n)$ は次式で表される 2 通りの解を持つ。

$$(a) \quad C_{21}(z) = \frac{H_{21}(z)}{H_{11}(z)} \quad (26)$$

$$C_{12}(z) = \frac{H_{12}(z)}{H_{22}(z)} \quad (27)$$

$$y_1(n) = \mathbf{h}_{11}\mathbf{s}_1^T(n) \quad (28)$$

$$y_2(n) = \mathbf{h}_{22}\mathbf{s}_2^T(n) \quad (29)$$

$$(b) \quad C_{21}(z) = \frac{H_{22}(z)}{H_{12}(z)} \quad (30)$$

$$C_{12}(z) = \frac{H_{11}(z)}{H_{21}(z)} \quad (31)$$

$$y_1(n) = \mathbf{h}_{12}\mathbf{s}_2^T(n) \quad (32)$$

$$y_2(n) = \mathbf{h}_{21}\mathbf{s}_1^T(n) \quad (33)$$

$H_{ji}(z)$ における伝搬遅延の仮定から (a) の $C_{21}(z)$, $C_{12}(z)$ は正の時間遅れを有し、因果性を満たすため実現可能である。一方、(b) の $C_{21}(z)$, $C_{12}(z)$ は負の時間遅れ、すなわち時間進みを有し、因果性を満たさないため、実現できない。

また、分離後の出力 $y_k(n)$ は信号源 $s_i(n)$ と \mathbf{h}_{11} や \mathbf{h}_{22} との畳み込みとなるが、BSS では混合過程そのものを推定できないため、 $s_i(n)$ に戻すことはできない。

3.2 評価関数

式 (22) から出力は次のように表せる。

$$\begin{aligned} \begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} &= \frac{1}{1 - C_{12}(z)C_{21}(z)} \begin{bmatrix} 1 & -C_{12}(z) \\ -C_{21}(z) & 1 \end{bmatrix} \\ &\times \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \\ &= \frac{1}{1 - C_{12}(z)C_{21}(z)} \\ &\times \begin{bmatrix} H_{11}(z) - C_{12}(z)H_{21}(z) & H_{12}(z) - C_{12}(z)H_{22}(z) \\ H_{21}(z) - C_{21}(z)H_{11}(z) & H_{22}(z) - C_{21}(z)H_{12}(z) \end{bmatrix} \\ &\times \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \end{aligned} \quad (34)$$

3.1節で述べたように、 $H_{ji}(z), j \neq i$ の伝搬遅延時間は $H_{ii}(z)$ のものより長いと仮定する。 $C_{12}(z)$ と $C_{21}(z)$ は正の伝搬遅延時間をもつ因果性を満たす回路である。したがって、式(34)の対角要素は($H_{11}(z)$ の伝搬遅延時間 $< C_{12}(z)H_{21}(z)$ の伝搬遅延時間)となり、0にならない。逆に非対角要素は $C_{12}(z)$ や $C_{21}(z)$ を調整することによって0になり得る。つまり、出力パワーを最小化するように分離回路 $C_{12}(z)$ 、 $C_{21}(z)$ を調整することで対角要素を一定のレベルに保ったまま、非対角要素を最小化することができる。すなわち、1つの信号を抽出し、信号を分離することができる。今、出力パワーを

$$J_j(n) = E[q(y_j(n))] \quad (35)$$

とする。ただし、 $q()$ はただ1つの極小値を持つ偶関数である。適応フィルタにおけるLMSアルゴリズムのように $E[q(y_j(n))]$ の代わりに、瞬時値 $q(y_j(n))$ を使用する。すなわち、評価関数として以下の $\hat{J}_j(n)$ を使用する。

$$\hat{J}_j(n) = q(y_j(n)) \quad (36)$$

3.3 $C_{jk}(z)$ の更新式

$\hat{J}_j(n)$ の勾配は次式ようになる。

$$\begin{aligned} \frac{\partial \hat{J}_j(n)}{\partial c_{jk}(l)} &= \frac{\partial q(y_j(n))}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial c_{jk}(l)} \\ &= -\dot{q}(y_j(n))y_k(n-l) \end{aligned} \quad (37)$$

$$y_j(n) = x_j(n) - \sum_{l=1}^{K_c-1} c_{jk}(l)y_k(n-l) \quad (38)$$

ここで $\dot{q}()$ は $q()$ 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。

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \Delta c_{jk}(n, l) \quad (39)$$

$$\Delta c_{jk}(n, l) = \mu \dot{q}(y_j(n))y_k(n-l) \quad (40)$$

上記の更新式において、奇関数 $\dot{q}(y_j(n))$ と $y_k(n-l)$ が用いられている。学習の高速化や安定化のためにこれらを別の奇関数に置き換えることができる[19], [20]。したがって、式(40)は

$$\Delta c_{jk}(n, l) = \mu f(y_j(n))g(y_k(n-l)) \quad (41)$$

と拡張できる。 $f()$ と $g()$ は異なる奇関数である。ただし、定数倍を除き、少なくとも1つが非線形関数である。

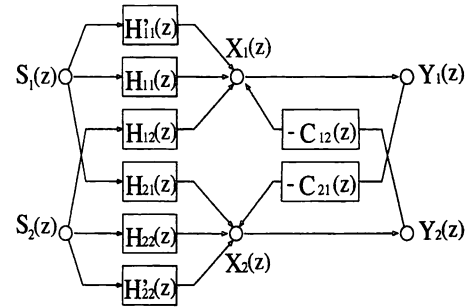


図3 反響音経路 $H'_{11}(z), H'_{22}(z)$ を含む畳み込み形BSS

Fig. 3 Convolutional BSS model with reverberation paths $H'_{11}(z)$ and $H'_{22}(z)$

この更新式はJuttenらが提案した学習アルゴリズム[6], [14]の更新式と一致するが、伝搬遅延時間に着目することにより、反響音を有する場合の学習安定化法の導入を容易にしている。

4. 反響音を含む畳み込み形BSSの学習アルゴリズム

4.1 学習の収束性

反響音を含む場合、学習が収束する条件である遅延に関する仮定が満たされない可能性がある。反響音を含む場合のBSSのモデルを図3に示す。 $H'_{11}(z)$ と $H'_{22}(z)$ は反響音に対する伝達関数を表している。直接音に比べると反響音の遅延時間は長い。3.1節で述べた遅延時間に対する条件は($H_{ii}(z)$ の遅延時間 $< H_{ji}(z)$ の遅延時間)である。 $S_1(z)$ から $X_2(z)$ への経路、 $S_2(z)$ から $X_1(z)$ への経路において発生する反響音はこの条件を満たしているため、 $H_{21}(z)$ と $H_{12}(z)$ を含めて考えることにする。

前述の通り、BSSの学習は出力パワーを最小化することによって行なわれる。 $Y_1(z)$ のパワーを減少させる可能性として、次の2通りが考えられる。

$$H_{12}(z)S_2(z) - C_{12}(z)H_{22}(z)S_2(z) \rightarrow 0 \quad (42)$$

$$H'_{11}(z)S_1(z) - C_{12}(z)H_{21}(z)S_1(z) \rightarrow 0 \quad (43)$$

式(42)はこれまでと同じであるが、式(43)は反響音によって新たに加わった項である。式(42), (43)に対し $C_{12}(z)$ はそれぞれ次のように収束する。但し、式(44)と(45)を同時に満たす $C_{12}(z)$ は存在しない。

$$C_{12}(z) \rightarrow \frac{H_{12}(z)}{H_{22}(z)} \quad (44)$$

$$C_{12}(z) \rightarrow \frac{H'_{11}(z)}{H_{21}(z)} \quad (45)$$

式 (44) は信号源分離として理想解の方向を示しており、式 (45) は反響音のために理想解からずれる方向を示している。すなわち、学習では $C_{12}(z)$ は式 (44) の方向に収束することが望ましく、式 (45) の方向への変動は抑制する必要がある。

$C_{21}(z)$ についても同様である。 $Y_2(z)$ のパワーを減少させる可能性として、次の2通りが考えられる。

$$H_{21}(z)S_1(z) - C_{21}(z)H_{11}(z)S_1(z) \rightarrow 0 \quad (46)$$

$$H'_{22}(z)S_2(z) - C_{21}(z)H_{12}(z)S_2(z) \rightarrow 0 \quad (47)$$

これらに対して $C_{21}(z)$ の修正方向として、次の2通りがある。

$$C_{21}(z) \rightarrow \frac{H_{21}(z)}{H_{11}(z)} \quad (48)$$

$$C_{21}(z) \rightarrow \frac{H'_{22}(z)}{H_{12}(z)} \quad (49)$$

式 (48) は理想解であるが、式 (49) の方向へも変動する可能性がある。以上の解析から分かるように、分離回路の出力パワーを最小化するように $C_{12}(z)$, $C_{21}(z)$ の学習を行っても、反響音がある場合には理想解に収束しない。

4.2 固定形指数ステップサイズ

ここでは、式 (45), 式 (49) による変動を小さく抑える学習法を提案する [22]-[25]。前節で述べた各信号成分の信号の大きさ、遅延の関係は図 4 のようになる。式 (44) と式 (48) の学習は $H_{ii}(z)$ の遅延と $H_{ij}(z)$ の遅延の差 τ_1 だけ遅れて行なわれる。一方、式 (45), 式 (49) の影響も $H_{ji}(z)$ の遅延と $H'_{ii}(z)$ の遅延の差 τ_2 の後に現れる。 $C_{12}(z)$ は式 (42), (44) により学習される伝達関数を表し、 $C'_{21}(z)$ は式 (43), (45) により学習される伝達関数を表す。

反響音は直接音に比べ振幅が小さく、反響の時間が長いので、 $\tau_1 < \tau_2$ であると考えられる。したがって、分離回路の FIR フィルタのタップ係数が式 (44), (45), (48), (49) により更新される度合は図 5 のようになる。つまり、FIR フィルタのタップの前半では直接音の方が反響音に比べて振幅が大きいため支配的にフィルタ係数を更新する。タップの後半では直接音の影響がなくなり、代わりに遅延時間の長い反響音が支配的になり、フィルタ係数を更新する。しかし、この更新はフィルタ係数が理想値から変動することを意味する。

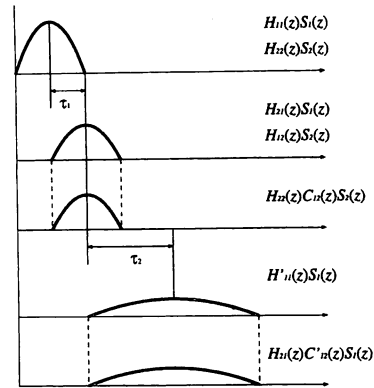


図 4 各信号成分の遅延の関係

Fig. 4 Relations between delay of each signal components

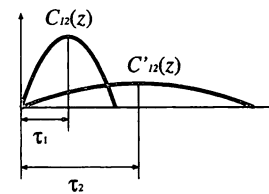


図 5 式 (44), (45) により形成される分離回路の FIR フィルタ係数

Fig. 5 FIR filter coefficients in separation block formulated by Eqs.(44) and (45)

以上のことから、後半部分の更新を抑える必要がある。これは FIR フィルタのタップ番号にあわせてステップサイズ μ を制御することによって実現できる。

分離回路の最適な伝達関数は式 (44), (48) で与えられる。さらにこれらの式は以下のように変形できる。

$$\begin{aligned} C_{jk}^{opt}(z) &= \frac{H_{jk}(z)}{H_{kk}(z)} = H_1(z) + \frac{H_3(z)}{H_2(z)} \\ &= H_1(z) + \sum_i \frac{1}{1 - z_i z^{-1}} \\ &= H_1(z) + \sum_i \sum_l z_i^l z^{-l} \\ &= H_1(z) + \sum_i \sum_l r_i^l e^{j\theta_i l} z^{-l} \quad (50) \end{aligned}$$

$H_1(z)$ は $H_{jk}(z)$ と $H_{kk}(z)$ の遅延差、すなわち、次数差から成る有限個の項数を有する z^{-1} の多項式であり、FIR フィルタの遅延の大きな部分に影響を及ぼさない。第 2 項は指数関数の和になっている。すなわち、分離回路のフィルタ係数の理想値は指数的に変化して

いる。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。

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \mu(l)f(y_j(n))g(y_k(n-l)) \quad (51)$$

$$\mu(l) = \mu_0 r^l, \quad 0 < r < 1 \quad (52)$$

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 適応フィルタにおける指数ステップサイズ

指数重み付けステップサイズの方法が NLMS で学習する適応フィルタにおいて提案されている [21]。これは学習の高速化を目的に提案されたアルゴリズムである。誤差の大きいタップの学習を高速に行なうためにステップサイズをフィルタ係数の理想解の包絡線に比例させている。

一方、本稿で提案する指数ステップサイズ方式は適応フィルタにおけるステップサイズ $\mu(l)$ と形は同じであるが、FIR フィルタの係数が反響音により最適値から変動することを抑えるため、反響音の影響が大きい部分の学習を抑制する方式である。すなわち、指数ステップサイズの作用が大きく異なる [24]。

5. 適応形指数ステップサイズ

これまで提案されている指数ステップサイズ的方式 [21] では指数の底を部屋のインパルス応答を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。本稿では、適応的に指数の値を推定する方法を提案する [22]。これは指数ステップサイズをフィルタ係数の学習中に自動的にフィルタ係数の包絡線に近づけるものである。

分離回路の包絡線はおおよそ図6のように表すことができる。

時刻 n における l タップ目のステップサイズを $\mu(n, l)$ とし、次のように表す。

$$\mu(n, l) = \mu_0(n)r^l(n) \quad (53)$$

ステップサイズとフィルタ係数を以下のように変換する。

$$\begin{aligned} \log \mu(n, l) &= \log \mu_0(n) + l \log r(n) \\ &= v_1(n) + lv_2(n) \end{aligned} \quad (54)$$

$$b(n, l) = \log |c_{jk}(n, l)| \quad (55)$$

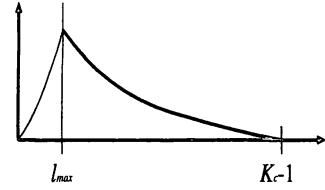


図6 分離回路のフィルタ係数の包絡線の概形

Fig. 6 A rough sketch of envelop of FIR filter coefficients in separation block

最小2乗法により $v_1(n) + lv_2(n)$ と $b(n, l)$ の2乗誤差を最小にする $v_1(n)$, $v_2(n)$ を求める。

$$v_1(n) + lv_2(n) = b(n, l) \quad (56)$$

ただし、 l は $l_{max} \sim K_c - 1$ を考える。ここで l_{max} はフィルタ係数の大きさが最大になるタップを表している。これらを行列、ベクトルを使って表現すると以下のようなになる。

$$A\mathbf{v}(n) = \mathbf{b}(n) \quad (57)$$

$$A = \begin{bmatrix} 1 & l_{max} \\ 1 & l_{max} + 1 \\ \vdots & \vdots \\ 1 & K_c - 1 \end{bmatrix} \quad (58)$$

$$\mathbf{v}(n) = \begin{bmatrix} v_1(n) \\ v_2(n) \end{bmatrix} \quad (59)$$

$$\mathbf{b}(n) = \begin{bmatrix} \log |c_{jk}(n, l_{max})| \\ \log |c_{jk}(n, l_{max} + 1)| \\ \vdots \\ \log |c_{jk}(n, K_c - 1)| \end{bmatrix} \quad (60)$$

最小2乗解は以下のようなになる。

$$\mathbf{v}(n) = A^+ \mathbf{b}(n) \quad (61)$$

$$A^+ = (A^T A)^{-1} A^T \quad (62)$$

この結果から、 $r(n)$, $\mu_0(n)$ は以下のように決定される。

$$\mu_0(n) = e^{v_1} \quad (63)$$

$$r(n) = e^{v_2} \quad (64)$$

ステップサイズの急激な変化により学習が不安定になることを防ぐため、指数の底は徐々に更新する。

$$\hat{r}(n) = \alpha r(n) + (1 - \alpha) \hat{r}(n - 1) \quad (65)$$

$$0 < \alpha \ll 1$$

したがって、ステップサイズ $\mu(n, l)$ は式 (58)~(65) を用いて以下のように計算される。

$$\mu(n, l) = \mu_0(n) \hat{r}^l(n) \quad (66)$$

$\hat{r}(n)$ の初期値は 1 とする。

6. シミュレーション

6.1 シミュレーション条件

信号源 2 個、センサ 2 個の FB-BSS について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。

$$f(y) = \tanh(2.5y) \quad g(y) = \tanh(0.5y) \quad (67)$$

分離性能は混合過程の伝達関数 $H(z)$ 、式 (24) で定義した $P(z)$ を用い、以下の NRR(noise reduction ratio) で評価した。

$$\sigma_{ys}^2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ii}(e^{j\omega T})|^2 d\omega T \quad (68)$$

$$\sigma_{yc}^2 = \sum_{j \neq i} \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ji}(e^{j\omega T})|^2 d\omega T \quad (69)$$

$$\sigma_{xs}^2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega T})|^2 d\omega T \quad (70)$$

$$\sigma_{xc}^2 = \sum_{j \neq i} \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega T})|^2 d\omega T \quad (71)$$

$$NRR = 10 \log \frac{\sigma_{ys}^2}{\sigma_{yc}^2} - 10 \log \frac{\sigma_{xs}^2}{\sigma_{xc}^2} \quad [\text{dB}] \quad (72)$$

σ_{ys}^2 , σ_{xs}^2 は取り出したい信号のパワー、 σ_{yc}^2 , σ_{xc}^2 は抑圧したい信号のパワーを表す。式 (72) 右辺第 1 項は分離回路出力における SNR で、第 2 項は分離回路入力における SNR である。NRR は分離回路の入力から出力まで至る過程で SNR がどの程度改善されたかを示している。

混合過程の伝達関数は実際の音響空間を模して疑似的に作成した。反響音は直接音に比べ、遅延が大きく大きさが小さいものとした。混合過程の伝達関数を図 7, 8 に示す。図 7 の伝達関数は反響音を含んでおり、遅延が大きい部分で減衰率が小さくなっている。一方図 8 の伝達関数は反響音を含んでおらず減衰率が一定である。

分離回路の FIR フィルタのタップ数は 1024 とする。

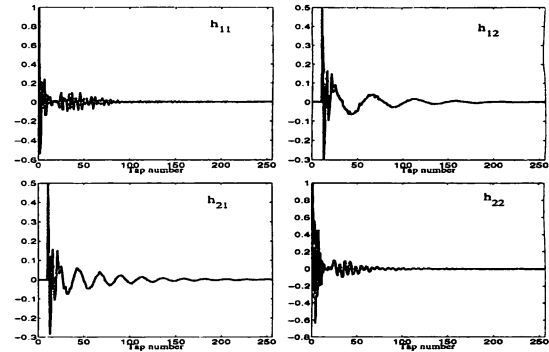


図 7 反響音を含むモデルに対する伝達関数

Fig. 7 Impulse responses in convolutive mixture with reverberations

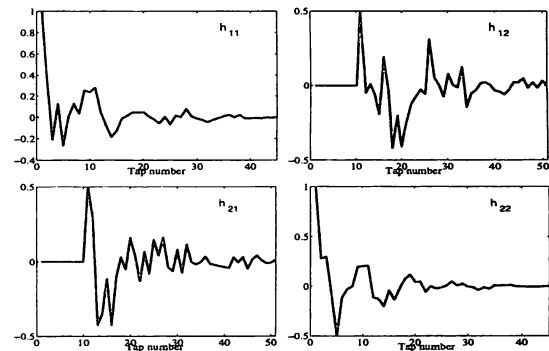


図 8 反響音を含まないモデルに対する伝達関数

Fig. 8 Impulse responses in convolutive mixture without reverberations

6.2 固定形指数ステップサイズ

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 9 に示す。信号源には白色信号を使用した。一定値ステップサイズ方式のステップサイズは分離性能が最も高い値を示し、かつ最も収束が速くなるように経験的に決定している。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。それぞれの値は、反響音がない場合において一定値ステップサイズ $\mu = 0.0001$ 、固定形指数ステップサイズ $\mu_0 = 0.00078$, $r_{12} = 0.9933$, $r_{21} = 0.9913$ 、反響音を有する場合において一定値ステップサイズ $\mu = 0.0003$ 、固定形指数ステップサイズ $\mu_0 = 0.00078$, $r_{12} = 0.9883$, $r_{21} = 0.9822$ である。

図 9 から分かるように、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

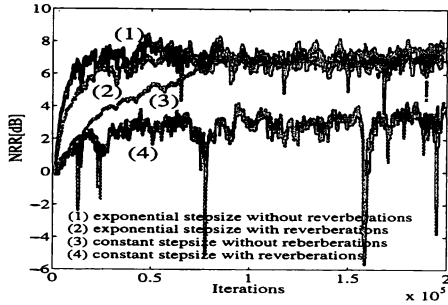


図9 白色信号源に対するNRR

Fig. 9 NRR for white signal sources. Convolutional mixtures with and without reverberations, and constant and Exponential Weighted (EW) stepsizes are used

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を用いることにより、適応フィルタとしての高速化 [21] が可能であると同時に、反響音によってフィルタ係数が最適値から変動することを抑制していることを示している。具体的には、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、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。

6.3 適応形指数ステップサイズ

前節の反響音がある場合と同じ条件で適応形指数ステップサイズを用いたシミュレーションを行なった。式(66)において $\hat{\mu}$ で示される、適応形指数ステップサイズの底の変化を図10に示す。 r_{12} , r_{21} はそれぞれ $C_{12}(z)$, $C_{21}(z)$ の学習に用いるステップサイズの底を表している。これらの値は最適な分離回路のフィルタ係数から求めた指数の底 $r_{12} = 0.9883$, $r_{21} = 0.9822$ に近づいている。また、適応形指数ステップサイズと固定形指数ステップサイズのNRRを図11に示す。固定形は理想フィルタ係数の包絡線から求めており、理想的な指数関数である。一方、適応形指数ステップサイズは理想フィルタ係数の包絡線を推定する必要がな

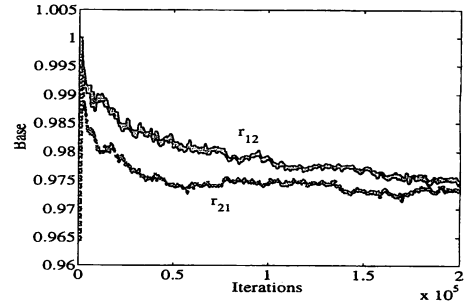


図10 $c_{12}(n)$, $c_{21}(n)$ の更新に用いる $r_{12}(n)$, $r_{21}(n)$ の適応過程

Fig. 10 Adjusting process of geometric ratios $r_{12}(n)$ and $r_{21}(n)$ used in updating $c_{12}(n)$ and $c_{21}(n)$, respectively

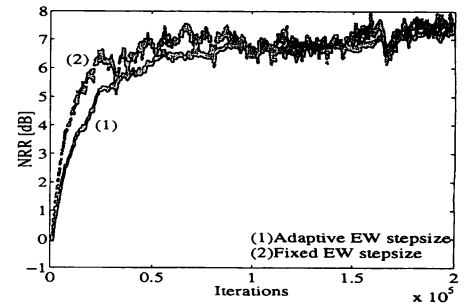


図11 適応形指数ステップサイズと固定形指数ステップサイズの比較

Fig. 11 Separation performance using adaptive EW stepsize and fixed EW stepsize

く、実用性が高い。適応形は固定形に比べ、収束速度が僅かに遅いがほぼ同等の分離性能が得られており、その有効性が確認できた。

6.4 実測データによるシミュレーション

図12に示す環境で測定した混合過程のデータを用いてシミュレーションを行なった。分離回路のFIRフィルタのタップ数は8192タップとする。信号源には白色信号を用いる。ステップサイズは一定値ステップサイズ方式では0.00007, 固定形指数ステップサイズでは $\mu_0 = 0.0001$, $r_{12} = 0.9993$, $r_{21} = 0.9993$ とする。その他の条件は6.1節と同じとする。分離性能を図13に示す。

固定形指数ステップサイズ方式は従来の一定値ステップサイズ方式より高い分離性能が得られており、反響音を抑制していることが分かる。また、適応形指数ス

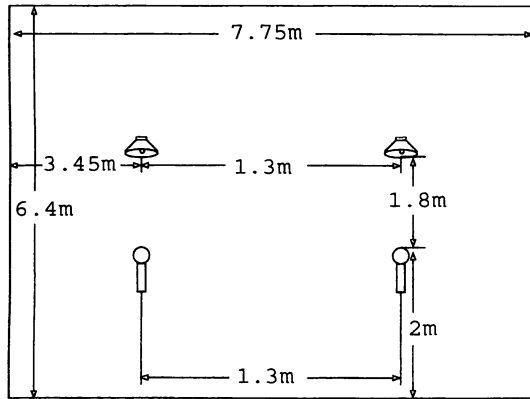


図 12 測定環境

Fig. 12 Room environment in which some signals are measured

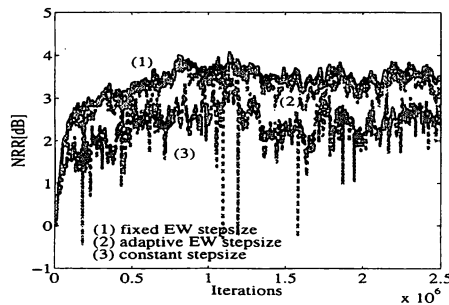


図 13 実測データによる各手法の分離性能

Fig. 13 Separation performance using adaptive EW stepsize, fixed EW stepsize and constant stepsize for survey data

ステップサイズ方式は固定形指数ステップサイズ方式とほぼ同等の分離性能が得られており、十分にステップサイズを学習できていることが分かる。

7. まとめ

畳み込み混合のモデルを時間領域で学習する FB-BSS において、従来の Jutten らが提案した学習アルゴリズム [6], [14] に対し、混合過程の伝達遅延の観点から解析を行なった。その中で、反響音を含む畳み込み形 BSS では、反響音のため分離回路のフィルタ係数が理想値からずれてしまうことが分かった。反響音の影響はフィルタ係数の後半部分に現れる。そこで指数重みステップサイズの学習アルゴリズムを導入し、さらに、指数の底を自動的に制御する方法を提案した。2チャンネルの FB-BSS のシミュレーションにより、

これらの方式の有効性を確認した。また、指数ステップサイズの効果を反響音の抑制と学習の高速化の観点より解析を行ない、そのどちらの効果も得られることを確認した。

参考文献

- [1] C.Jutten and J.Herault and A.Guerin: IN.C.A.: An independent components analyzer 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, pp145-152, Feb.-March 1989.
- [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 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, Aug. 1989.
- [5] J.F.Cardoso: Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem, ICASSP Proc. Albuquerque, USA, pp.2655-1658. March 1990.
- [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, July 1991.
- [7] P.Comon, C.Jutten and J.Herault: Blind separation of sources, Part II :Problems statement, Signal Processing, 24, pp.11-20, July 1991.
- [8] E.Sorouchyari: Blind separation of sources, Part III :Stability analysis, Signal Processing, 24, pp.21-29, July 1991.
- [9] S.Haykin: Adaptive Filter Theory, 3rd ed., Prentice-Hall, Inc.1996.
- [10] A.Cichocki, S.Amari, M.Adachi and W.Kasprzak: Self-adaptive neural networks for blind separation of sources, Proc.ISCAS'96, Atlanta, pp157-161, May 1996.
- [11] S.Amari, T.Chen and A.Cichocki: Stability analysis of learning algorithms for blind source separation, Neural Networks, Vol.10, No.8, pp1345-1351, Nov. 1997.
- [12] H.Mathis and S.C.Douglas: On optimal and universal nonlinearities for blind signal separation, IEEE Proc.ICASSP'01, Salt Lake City, UT, USA, MULT-P3.3, May, 2001.
- [13] K.Nakayama, A.Hirano and T.Sakai: An adaptive nonlinear function controlled by kurtosis for blind source separation, IEEE&INNS, Proc. IJCNN'2002, Honolulu, Hawaii, pp.1234-1239, May 2002.
- [14] H.L.Nguyen Thi and C.Jutten: Blind source separation for convolutive mixtures, Signal Processing, Vol.45, No.2, pp.209-229, March 1995.

- [15] C.Simon, G.d'Urso, C.Vignat, Ph.Loubaton and C.Jutten: On the convolutive mixture source separation by the decorrelation approach, IEEE Proc.ICASSP'98, Seattle, pp.IV 2093-2112, May 1998.
- [16] S.Cruces and L.Castedo: A Gauss-Newton methods for blind source separation of convolutive mixtures, IEEE Proc.ICASSP'98 Seattle, pp.IV 2093-2096, May 1998.
- [17] S.Araki, S.Makino, T.Nishikawa and H.Saruwatari: Fundamental limitation of frequency domein blind source separation for convolutive mixture of speech, IEEE Proc.ICASSP'01, Solt Lake City, MULT-P2.3, May 2001.
- [18] I.Kopriva, Z.Devcic and H.Szu: An adaptive short-time frequency domain algorithm for blind separation of nonstationary convolved mixtures, IEEE&INNS Proc.IJCNN'01, pp.424-429, July 2001.
- [19] A.Horita, K.Nakayama and A.Hirano: A learning algorithm for convolutive blind source separation with transmission delay constraint, IEICE General Conference, Sendai, A-4-43, March 2002.
- [20] K.Nakayama, A.Hirano and A.Horita: A learning algorithm for convolutive blind source separation with transmission delay constraint, IEEE&INNS, Proc. IJCNN'2002, Honolulu, Hawaii, pp.1287-1292, May 2002.
- [21] S.Makino, Y.Kaneda and 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.
- [22] A.Horita, K.Nakayama and A.Hirano: A learning algorithm with adaptive exponential stepsize for BSS of convolutive mixtures with reverberations, IEICE General Conference, Sendai, SA-3-7, March 2003.
- [23] 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高次 BSS の学習法—学習率に重みを付けて制御する方法—, 第 18 回デジタル信号処理シンポジウム, 志摩, A2-4, 2003-11.
- [24] 堀田明秀, 中山謙二, 平野晃宏: 反響音を有する混合過程の BSS における指数重みステップサイズの効果, 電子情報通信学会総合大会, 東京, A-4-38, 2004-3.
- [25] 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.



堀田 明秀 平 14 金沢大・工・電気情報卒。平 16 同大学院大学院博士前期課程了。現在, 同大学院大学院博士後期課程在学中。ブラインドソースセパレーションに関する研究に従事。



中山 謙二 昭 46 東工大・工・電子卒。昭 46~47 同大学研究生。昭 58 工博(東工大)。昭 47 日本電気(株)入社。伝送通信事業部および C & C システム研究所に勤務し, 通信用各種フィルタ及びデジタル信号処理の研究開発に従事。昭 63 金沢大・工・電気情報工学科助教授, 平 2 同教授, 平 12 情報システム工学科(新設)教授, 平成 16 自然科学研究科教授となり現在に至る。平成 15~16 留学生センター長を併任。最近の研究テーマは主として適応信号処理及びニューラルネットワーク。昭 62 IEEE Circuits & Devices Mag. 論文賞受賞。著書「SC 回路網の設計と応用」(東海大学出版会)他。電子情報通信学会会員, IEEE シニア会員, INNS 会員。



平野 晃宏 昭 62 金沢大・工・電子卒。平 1 同大学院修士課程了。平 12 工博(金沢大)。平 1 NEC 入社。研究開発グループにてエコーキャンセラの研究開発に従事。平 10 金沢大・工・電気・情報工学科助手, 平 13 講師, 平 16 同大学院・自然科学研究科・電子情報科学専攻講師となり現在に至る。最近の研究テーマは主として適応信号処理およびニューラルネットワーク。平 7 信学会学術奨励賞受賞。IEEE 会員。

(2006 年 10 月 23 日受付, 2007 年 1 月 10 日再受付)