

フィードフォワード形 BSS とフィードバック形 BSS における 収束性及び信号歪みの比較

Comparative Study of Convergence Performance and Signal Distortion in Feed-forward BSS and Feedback BSS

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

[†]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

[‡]金沢大学大学院 自然科学研究科 電子情報工学専攻
Division of Electrical and Computer Engineering
Graduate School of Natural Science and Technology, Kanazawa Univ.

あらまし

本稿では、ブラインド形信号源分離 (BSS) において、フィードフォワード (FF) 信号源の形回路とフィードバック (FB) 形回路の分離性能、収束速度、及び信号歪みについて比較検討を行った。BSS の分離回路は出力が統計的に独立になるように学習される。白色信号のように独立性が高い場合は、出力を独立にすることと、信号源分離は一致するが、音声信号のように互いに相関性を持つ場合は、これらは一致しない。FB-BSS では、信号源の完全分離に対して分離回路が唯一の解を持ち、その時の分離回路出力は観測信号に含まれる 1 個の信号源成分と同一になる。この状態を信号歪み無しと見なすことができる。一方、FF-BSS では信号源の完全分離に対して分離回路が自由度を有し、出力が独立となるようにスペクトルが調整され、信号歪みが生じる。信号源として白色信号と音声信号を用いたシミュレーションを行い、上記の解析結果を確認した。

ABSTRACT

This paper analyzes and compares signal separation, convergence speed and signal distortion in both a feed-forward blind source separation (FF-BSS) and a feedback (FB-) BSS. In BSS, a separation block is trained so as to make its output signals to be statistically independent. When the signal sources are highly independent such as white noise, making the output signals to

be statistically independent is the same as separating the signal sources. However, when the signal sources are somewhat correlated to each other such as speech signals, they are not exactly the same. In the FB-BSS, when complete separation is achieved, the separation block has a unique solution, and the output signals are equal to the single signal source included in the observation. On the other hands, in the FF-BSS, the separation block has some degree of freedom under the complete separation condition. This degree of freedom is used to change the output spectrum so as to make them to be independent, causing signal distortion. Through simulations, using white signals and speech signals as the signal sources, the properties above mentioned are confirmed.

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]-[10]。

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。また、フィードフォワード形とフィードバック形という2つの回路構造が提案されている。

BSSは出力信号が統計的に独立になるように学習される。したがって、線形な伝達関数によって引き起こされる歪みについて制御できない[5],[6]。しかし、実際には観測した信号をそのまま出力されることが望ましい。

そこで本稿では、時間領域におけるフィードフォワード形BSS(FF-BSS)とフィードバック形BSS(FB-BSS)をという2種類の回路構造を信号の分離性能の比較を行なうとともに信号歪みという観点からも2者の比較を行なう。

2 FB-BSS

2.1 回路構成と入出力関係

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。

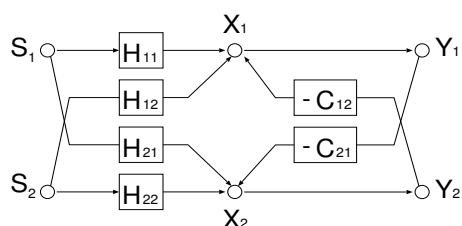


図1: FB-BSSの回路構成

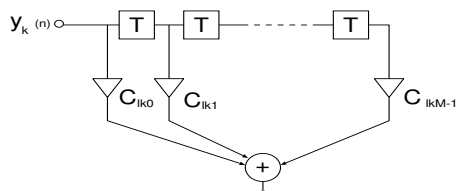


図2: $C_{21}(z)$ や $C_{12}(z)$ に用いる FIR フィルタ

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。

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

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

$$\mathbf{S}(z) = [S_1(z), S_2(z), \dots, S_N(z)]^T \quad (5)$$

$$\mathbf{X}(z) = [X_1(z), X_2(z), \dots, X_N(z)]^T \quad (6)$$

$$\mathbf{Y}(z) = [Y_1(z), Y_2(z), \dots, Y_N(z)]^T \quad (7)$$

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。

$$\begin{aligned} \mathbf{Y}(z) &= (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{X}(z) \\ &= (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{H}(z) \mathbf{S}(z) \end{aligned} \quad (8)$$

分離性能を評価するために以下の行列を定義する。

$$\mathbf{P}(z) = (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{H}(z) \quad (9)$$

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。

$$Y_j(z) = P_{ji}(z) S_i(z) \quad (10)$$

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[11]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。コスト関数を以下のような出力のパワーとする。

$$J_j(n) = E[q(y_j(n))] \quad (11)$$

$q()$ はただ1つの極小値を持つ偶関数である。 $E[q(y_j(n))]$ の代わりに、瞬時値 $q(y_j(n))$ を使い、勾配法を適用する。

$$\hat{J}_j(n) = q(y_j(n)) \quad (12)$$

$\hat{J}_j(n)$ の勾配は次式のようにになる。

$$\begin{aligned} \frac{\partial \hat{J}_j(n)}{\partial c_{jk}(l)} &= \frac{\partial q(y_j(n))}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial c_{jk}(l)} \\ &= \dot{q}(y_j(n)) y_k(n-l) \end{aligned} \quad (13)$$

$$y_j(n) = x_j(n) - \sum_{l=0}^{L_{jk}-1} c_{jk}(l) y_k(n-l) \quad (14)$$

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。

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \Delta c_{jk}(n, l) \quad (15)$$

$$\Delta c_{jk}(n, l) = \mu \dot{q}(y_j(n)) y_k(n-l) \quad (16)$$

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.

$$\begin{aligned} E[f(s_1(n))g(s_2(n))] &= E[f(s_1(n))]E[g(s_2(n))] \\ &= 0 \end{aligned} \quad (17)$$

$f(), g() : \text{奇関数}$

式 (16) 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 $E[\dot{q}(y_j(n))y_k(n-l)]$ と見なせる. $\dot{q}(y_j(n))$ と $y_k(n-l)$ は奇関数であるから式 (17) 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 $\mathbf{h}_{11}^T \mathbf{s}_1(n)$ や $\mathbf{h}_{22}^T \mathbf{s}_2(n)$ に近付くということである.

2.3 信号分離と信号歪み

S_1 と S_2 がそれぞれ Y_1 と Y_2 に取り出されたとする. いま分離回路が

$$C_{ij}(z) = \frac{\hat{H}_{ij}(z)}{\hat{H}_{jj}(z)} \quad (18)$$

であるとする. 信号源 $S_i(z)$ から出力 $Y_i(z)$ への経路 $P_{ii}(z)$ は式 (9) より

$$\begin{aligned} P_{ii}(z) &= \frac{H_{ii}(z) - C_{ij}(z)H_{ji}(z)}{1 - C_{12}(z)C_{21}(z)} \quad (19) \\ &= \frac{\hat{H}_{ii}(z)\hat{H}_{jj}(z)}{\hat{H}_{ii}(z)\hat{H}_{jj}(z) - \hat{H}_{ij}(z)\hat{H}_{ji}(z)} \\ &\times \frac{H_{ii}(z)\hat{H}_{jj}(z) - \hat{H}_{ij}(z)H_{ji}(z)}{\hat{H}_{jj}(z)} \quad (20) \end{aligned}$$

となる. ここで分離回路 $C_{ij}(z)$ が完全分離の条件をみたすとするとこの式は

$$P_{ii}(z) = H_{ii}(z) \quad (21)$$

となる. BSS では、混合過程の伝達関数 $H_{ij}(z)$ は推定できないため、本稿では $H_{ii}(z)S_i(z)$ を信号歪みの基準とする. すなわち、信号歪みを $H_{ii}(z)S_i(z)$ からどれだけずれているかで判断することにする. したがって、FB-BSS では完全分離できれば信号歪みは発生しないことになる.

信号源が有色の場合は完全分離よりも $y_1(n)$ と $y_2(n)$ のスペクトルをずらした方が独立性が高くなることも考えられる. この場合は、出力の独立性は高くなるが、分離性能、信号歪みとも悪くなる. 信号源が白色の場合は、 $H_{ii}(z)$ によって、 $Y_i(z)$ は白色ではなくなる. $H_{11}(z)$ と $H_{22}(z)$ が似ている場合は、 $Y_i(z)$ が $H_{ii}(z)S_i(z)$ からずれて白色に近づいた方が $Y_1(z)$ と $Y_2(z)$ の独立性が高くなる. しかし、自由度が少ないため信号歪みは小さいと考えられる.

3 FF-BSS

3.1 回路構成と入出力関係

FF-BSS の回路構成を図 3 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 4 に示す FIR フィルタを用いる.

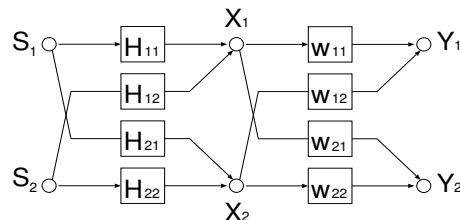


図 3: FF-BSS の回路構成

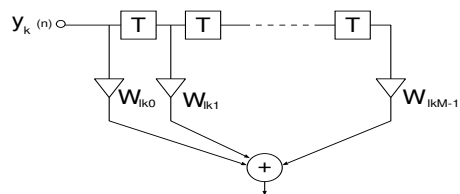


図 4: $W_{lk}(z)$ に用いる FIR フィルタ

混合過程は FB-BSS と同様のモデル化を行なっているため観測信号は式 (1), z 変換し、ベクトルと行列で表したものは式 (3) と表せる.

分離過程の出力 $y_j(n)$ は次式で表される.

$$y_j(n) = \sum_{k=1}^N \sum_{l=0}^{L_{jk}-1} w_{jk}(l)x_k(n-l) \quad (22)$$

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.

$$\mathbf{Y}(z) = \mathbf{W}(z)\mathbf{X}(z) \quad (23)$$

$$\mathbf{S}(z) = [S_1(z), S_2(z), \dots, S_N(z)]^T \quad (24)$$

$$\mathbf{X}(z) = [X_1(z), X_2(z), \dots, X_N(z)]^T \quad (25)$$

$$\mathbf{Y}(z) = [Y_1(z), Y_2(z), \dots, Y_N(z)]^T \quad (26)$$

$$\mathbf{W}(z) = \begin{bmatrix} W_{11}(z) & W_{12}(z) & \dots & W_{1N}(z) \\ W_{21}(z) & W_{22}(z) & \dots & W_{2N}(z) \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1}(z) & W_{N2}(z) & \dots & W_{NN}(z) \end{bmatrix} \quad (27)$$

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.

$$\begin{aligned} \mathbf{Y}(z) &= \mathbf{W}(z)\mathbf{X}(z) \\ &= \mathbf{W}(z)\mathbf{H}(z)\mathbf{S}(z) \end{aligned} \quad (28)$$

分離性能を評価するために以下の行列を定義する.

$$\mathbf{P}(z) = \mathbf{W}(z)\mathbf{H}(z) \quad (29)$$

FB-BSS と同様, 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.

3.2 学習アルゴリズム

学習には [12] で紹介されている自然勾配法を使用する. 簡単化のため 2 チャンネルにおいて考える. 信号は独立した状態が最もエントロピーが低く, 逆に均一に互いが従属している場合にはエントロピーが高い. したがって, エントロピーをコスト関数として勾配法を適用することによって信号を独立成分に分離することができる.

エントロピー $I(\mathbf{y})$ の勾配は次式のようになる.

$$\frac{\partial I(\mathbf{y})}{\partial \mathbf{W}} = ((\mathbf{W}^T)^{-1} - E[\varphi(\mathbf{y})\mathbf{x}^T]) \quad (30)$$

$$= (\mathbf{I} - E[\varphi(\mathbf{y})\mathbf{y}^T])(\mathbf{W}^T)^{-1} \quad (31)$$

$$\varphi(\mathbf{y}) = \left[\frac{\partial \log p(y_1)}{\partial y_1}, \frac{\partial \log p(y_2)}{\partial y_2} \right] \quad (32)$$

ただし, $p(y_j)$ は y_j の確率密度関数である. 実用的には $\varphi(y_j)$ は適当な関数で近似される. したがって, $w_{ij}(l)$ の更新式は次式のようになる.

$$w_{ij}(n+1, l) = w_{ij}(n, l) + \Delta w_{ij}(n, l) \quad (33)$$

$$\Delta w_{ij}(n, l) = \eta \{ w_{ij}(n, l) - \sum_{q=0}^{L_{ij}-1} \varphi(y_i(n)) y_p(n-l+q) w_{pj}(n, q) \}, p \neq j \quad (34)$$

3.3 信号分離と信号歪み

3.3.1 解の自由度

S_1 と S_2 がそれぞれ Y_1 と Y_2 に取り出されたとする. 式 (28) より完全分離成功条件は以下の様になる.

$$W_{ii}(z)H_{ij}(z) + W_{ij}(z)H_{jj}(z) = 0 \quad (35)$$

$W_{ij}(z)$ と $H_{ij}(z)$ を FIR フィルタで表現するとし, それぞれのフィルタ次数を $L-1$ と $M-1$ とすると式 (35) の 0 にすべき独立した項の数は $L+M-1$ である. 一方, 未知数である $W_{ij}(z)$ の項数は $2L$ である.

$L+M-1 > 2L$ のとき方程式の解は不能解となり, 完全な分離が不可能となる. 一方, $L+M-1 < 2L$ のとき分離回路の係数は不定解となり, 自由度が存在する.

3.3.2 歪みの評価

厳密には完全分離を満たす解の中で $|S_i(z) - Y_i(z)|$ を最小にする解が歪みなく信号を分離できていると言える. ただし, $H_{ii}(z)$ と $H_{ij}(z) (i \neq j)$ の特性が非常に近いとき, $H_{ii}(z)S_i(z)$ を信号歪みの基準と考えることもできる. したがって, 信号歪みを

$$D_1 = |S_i(z) - Y_i(z)|^2 \quad (36)$$

$$D_2 = |H_{11}(z)S_i(z) - Y_i(z)|^2 \quad (37)$$

で評価する. 式 (36) の評価を厳密法, 式 (37) の評価を簡便法と呼ぶことにする.

3.3.3 白色信号

信号源に白色信号を用いた場合, 信号源が完全に独立であるため, 分離回路の出力を統計的に独立にすることと信号源分離が一致する. したがって, 自由度があっても, 信号源が分離され, かつ出力が独立になるように学習されるので信号歪みは生じない. また, このときは分離回路に対して唯一の解がある. すなわち $P_{ii} \rightarrow 1, P_{ij} \rightarrow 0$ となる. これは簡便法による信号歪みが大きいと厳密法による信号歪みが小さくなる.

また, $P_{ii} \rightarrow 1, P_{ij} \rightarrow 0$ となるとき, 以下の式を満たすことになる.

$$\begin{bmatrix} W_{11}(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) \end{bmatrix} = \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix}^{-1} \quad (38)$$

これを実現させるためには分離回路のタップ数を十分大きくする必要がある.

3.3.4 有色信号

信号源に音声などの有色信号を用いた場合, 分離回路の出力を統計的に独立にすることと, 信号源分離が一致しない. 分離と独立性最大が一致せず, 出力が独立となるように学習が進む. これは一般に信号源からも観測信号の S_i 成分からもずれが生じる可能性がある.

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.

$$f(y) = \tanh(2.5y) \quad g(y) = \tanh(0.5y) \quad (39)$$

$$\varphi(y) = \tanh(0.5y) \quad (40)$$

混合過程の伝達関数のインパルス応答は図 5 のように設定した.

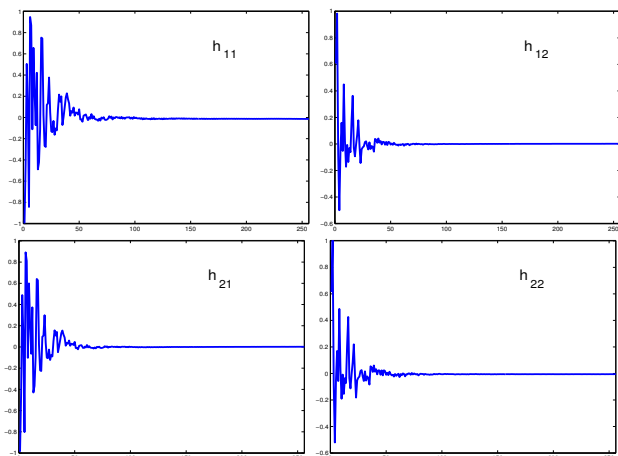


図 5: 混合過程の伝達関数 (インパルス応答)

分離性能は式 (9),(29) で定義した $\mathbf{P}(z)$ を用い、以下の SNR で評価した。

$$\sigma_s^2 = \sum_{i=1}^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ii}(e^{j\omega T})|^2 d\omega T \quad (41)$$

$$\sigma_c^2 = \sum_{j \neq i} \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ji}(e^{j\omega T})|^2 d\omega T \quad (42)$$

$$SNR = 10 \log \frac{\sigma_s^2}{\sigma_c^2} \quad [\text{dB}] \quad (43)$$

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。

また今回シミュレーションで用いた H_{ii} と H_{ij} の特性が似ているため FF-BSS の信号歪みの評価には簡便法を用いる。ただし、信号源が白色の場合には先にも述べたように簡便法による信号歪みが大きくなり、厳密法の信号歪みが小さくなるため厳密法を使用する。

4.2 白色信号

4.2.1 FB-BSS

ステップサイズは 0.00006 とし、タップ数を 256 とし、分離性能と歪みの変化を見た。

図 6 に分離性能を示す。図 7 に学習回数が 100000 回目、400000 回目の信号源 \mathbf{S} から出力 \mathbf{Y} までの伝達特性と混合過程の周波数特性を示す。

分離が進につれ、混合過程の伝達関数 H_{ii} に近付いていることが分かる。

4.2.2 FF-BSS

ステップサイズは 0.00001 とし、タップ数を 64, 128, 256 と変化させてその特徴を見た。

図 8 にそれぞれの分離性能を示す。図 9 にそれぞれの信号源 \mathbf{S} から出力 \mathbf{Y} までの伝達特性を示す。分離性能は

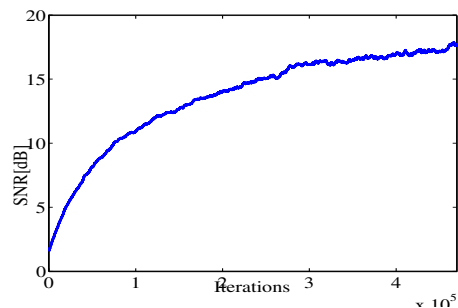


図 6: 白色入力に対する分離特性

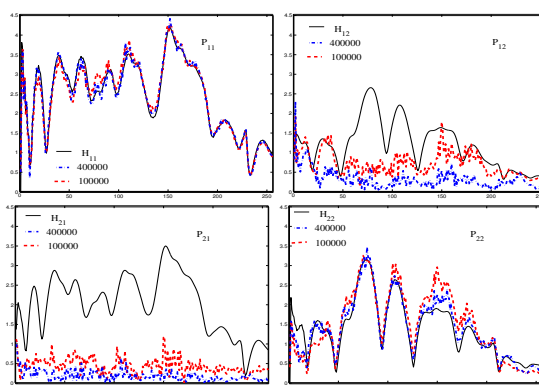


図 7: 白色入力に対する信号歪み

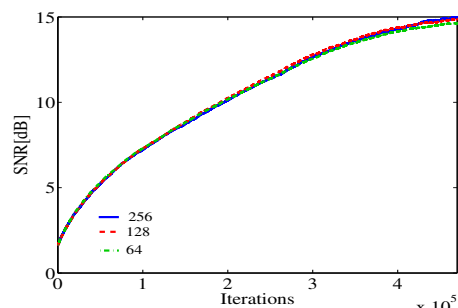


図 8: 白色入力に対する分離特性

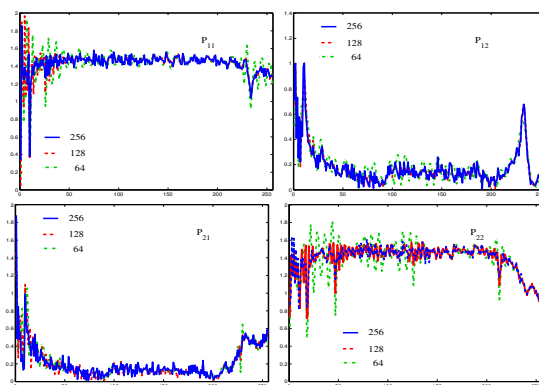


図 9: 白色入力に対する信号歪み

全てのタップではほぼ同等であるが、タップ数が少ない方が周波数応答がフラットでなく、歪んでいることが分かる。

4.3 音声信号

ステップサイズはFB-BSSは0.00004とし、FF-BSSは0.000007として、タップ数が256のときの特徴を見た。図10にそれぞれの分離性能を示す。図11にそれぞれの

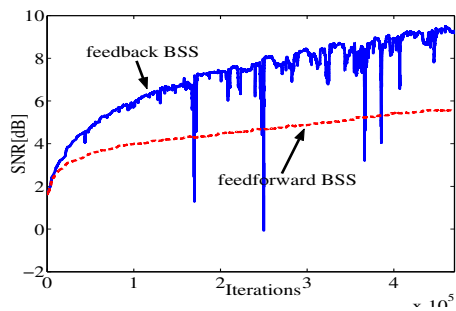


図 10: 音声入力に対する分離特性

信号源 S から出力 Y までの伝達特性を示す。

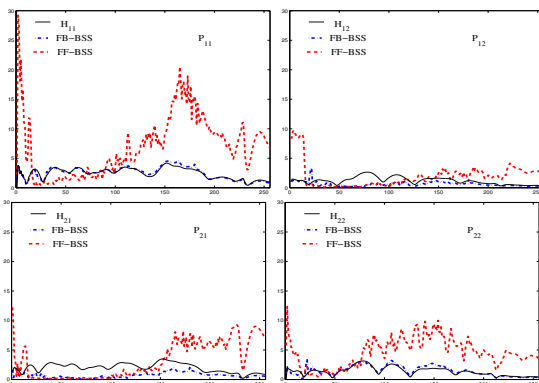


図 11: 音声入力に対する信号歪み

音声に対してもFB-BSSは混合過程の伝達関数 $H_{ii}(z)$ に近付いており、観測信号から信号が歪んでいないことがわかる。一方FF-BSSは低域の一部と高域を強調していることがわかる。この部分は音声にはあまり含まれない成分であり、周波数をフラットに、すなわち白色化しようとしているのが分かる。

5 結論

時間領域のFB-BSSとFF-BSSを信号分離と信号歪みの観点から解析した。FB-BSSでは信号分離と信号歪みを低減する方向が同じため学習が進につれ信号が分離され歪みが低減される。一方、FF-BSSでは信号源が白色であれば信号歪みを抑えるためにタップを十分に長くすればよいが、音声などの有色信号に対して、十分に

信号分離を行なおうとタップを長くすると信号が白色化され、信号が歪んでしまうことがシミュレーションにより確認された。

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, Manchester 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] A.Cichocki, S.Amari, M.Adachi, W.Kasprzak, "Self-adaptive neural networks for blind separation of sources", Proc. ISCAS'96, Atlanta, pp.157-161, 1996.
- [6] 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.
- [7] K.Nakayama, A.Hirano and M.Nitta, "A constraint learning algorithm for blind source separation", Proc. IJCNN'2000, pp.24-27, July, 2000.
- [8] K.Nakayama, A.Hirano and T.Sakai, "A pair-channel learning algorithm with constraints for multi-channel blind separation", Proc. IJCNN'01, July 2001.
- [9] H.Mathis and S.C.Douglas, "On optimal and universal nonlinearities for blind signal separation", Proc. ICASSP'01, MULT-P3.3, May 2001.
- [10] 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.
- [11] 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.
- [12] H.Saruwatari, T.Takatani, H.Yamajo, T.Sishikawa and K.Shikano, "Blind separation and deconvolution for real convolutive mixture of temporally correlated acoustic signals using SIMO-model-based ICA", ICA'03 pp.549-554 April 2003
- [13] 堀田明秀, 出島康宏, 中山謙二, 平野晃宏, "フィードフォワード形及びフィードバック形 BSS の収束性に関する比較", 信号処理研究会, pp.57-62 2004,7