

Over-complete Blind Source Separation for Nonlinear Convolutional Mixtures

メタデータ	言語: jpn 出版者: 公開日: 2019-12-06 キーワード (Ja): キーワード (En): 作成者: Nakayama, Kenji メールアドレス: 所属:
URL	https://doi.org/10.24517/00056434

This work is licensed under a Creative Commons
Attribution-NonCommercial-ShareAlike 3.0
International License.



非線形畳み込み混合過程における オーバーコンプリート形ブラインド信号源分離

課題番号 17560335

平成17年度～平成18年度科学研究費補助金
(基盤研究(C))研究成果報告書

平成19年4月

金沢大学附属図書館



1300-04215-X

中山 謙 二

自然科学研究科 教授)

は し が き

信号源やその伝達・混合過程に関する事前情報がなく、それらの統計的な性質のみを仮定して信号源を分離するブラインド形信号源分離 (BSS) に関して研究を行った。

1) オーバーコンプリート形ブラインド信号源分離 (OC-BSS)

OC-BSS に関して、フィードバック形ネットワークを提案した。分離回路において、単一信号源を分離する出力を少なくとも1個有するようにセンサー数を決定する。この状態で1巡目の信号源分離を行う。さらに、単一の信号源のみを含む出力を検出をする方法を提案した。これは、音声のピッチ周波数に基づく方法である。1巡目の信号源分離で分離された単一信号源は分離回路の入力部にフィードバックされ、入力信号 (センサー出力) からキャンセルされる。フィードバックにより信号源を相殺するための新しい学習法を提案した。混合過程に関する情報を使う方法と信号のヒストグラムを使う2つの方法を提案した。さらに、ヒストグラムを使う方法に関して、信号歪みを低減する方法を提案した。信号歪みを雑音と見なして、スペクトルサプレッション法により雑音スペクトルを抑制する方法である。

信号源の数が1個削減された状態で2巡目の信号源分離を行う。以下の操作は1巡目と同じである。このように、提案した方式は、フィードバックにより、信号源の数を等価的に削減しながら OC-BSS を行うものである。混合過程の特性に基づいて、本方式の理論的な限界についても解析した。音声の信号源を3個、センサーを2個としたときのシミュレーションを行い、従来法に比べて、信号対干渉比が約10 dB 改善された。また、理論的な限界にも近い分離特性となっていることを確認した。

2) 非線形混合過程に対するブラインド信号源分離

信号源のグループ分離と線形化を縦続接続する方式を以前に提案したが、その学習法に関して改良を行った。特に、線形化に関して、「初期値の設定法」及び「パラメータの学習法」に関して新しい方法を提案し、信号源の分離特性を大幅に改善した。また、信号源とセンサーの位置関係と必要なセンサー数の関係についても解析し、実用化における指針を与えた。

3) ブラインド信号源分離における信号歪みの低減

一般に BSS では信号歪みが生じる。これに関してフィードフォワード形 (FF-) BSS おフィードバック形 (FB-) BSS において、信号歪みの発生するメカニズムを解析した。さらに、FF-BSS において、信号歪みを抑制する学習法を提案した。信号源が3個以上の場合にも適用可能である。信号源として音声を用いたシミュレーションを行い、従来法よりも信号歪みが抑制されていることを確認した。

フィードフォワード形 (FF-) BSS に対して、信号歪みを抑制する新しい学習法を提案した。従来の学習法に信号歪みを抑制する制約条件を課す方法である。2チャンネルと3チャンネルについてシミュレーションを行い、分離特性と信号歪みを評価した。その結果、従来の信号歪み抑制形学習法に比べて大幅な特性改善を実現した。さらに、フィードバック形 (FB-) BSS と FF-BSS が各々有効に使用できる条件を明らかにした。

研究組織

研究代表者： 中山 謙二 (金沢大学 自然科学研究科 教授)

研究分担者： 平野 晃宏 (金沢大学 自然科学研究科 講師)

交付金額 (配分額)

(金額単位：千円)

	直接経費	間接経費	合 計
平成17年度	2,900	0	2,900
平成18年度	800	0	800
総 計	3,700	0	3,700

研究発表

(1) 学術誌等

◇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", 信号処理学会 論文誌, 2007.5 掲載予定.

◇ B. Jansen and K. Nakayama, "An adaptive penalty-based learning extension for the backpropagation family", IEICE Trans. Fundamentals, Vol.E89-A, No.8, pp.2140-2148, Aug. 2006.

(2) 口頭発表

◇ K. Nakayama, H. Katou and A. Hirano, "A feedback approach and its learning algorithm for over complete blind source separation", Proc. IEEE, ISPACS2006, Yonago, Japan, pp.459-462, Dec. 2006.

◇ 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, Sept 2006.

◇ Boris Jansen and K. Nakayama, "An adaptive penalty-based learning extension for backpropagation and its variants", IEEE&INNS, Proc., IJCNN2006, Vancouver, pp.6427-6432, July 2006.

◇ 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.

◇ B. Jansen and K. Nakayama, "Neural networks following a binary approach applied to the integer prime-factorization problem", IEEE&INNS, Proc., IJCNN2005, Montreal, pp.2577-2582, July-Aug. 2005.

◇ A. Horita, K. Nakayama, A. Hirano and Y. Dejima, "Analysis of signal separation and signal distortion in feedforward and feedback blind source separation based on source spectra", IEEE&INNS, Proc., IJCNN2005, Montreal, pp.1257-1262, July-Aug. 2005.

目次

第1章	序論	1
1.1	ブラインド信号源分離とは	1
1.2	ブラインド信号分離の応用	2
1.3	研究の位置付け	4
1.4	論文の構成	6
第2章	フィードバック形 BSS (FB-BSS)	8
2.1	回路構成と入出力関係	8
2.1.1	回路構成	8
2.1.2	時間領域	10
2.1.3	z 領域	11
2.2	学習アルゴリズムの提案	13
2.2.1	伝搬遅延に関する仮定	13
2.2.2	評価関数	13
2.2.3	$C_{jk}(z)$ の更新式	14
2.2.4	統計的な収束特性	14
第3章	FB-BSS における反響音による影響とその改善法	15
3.1	反響音による影響	15
3.1.1	反響音	15
3.1.2	反響音による影響	15
3.2	指数ステップサイズ方式の提案	17
3.2.1	固定形指数ステップサイズ	17
3.2.2	適応フィルタにおける指数ステップサイズ	18
3.2.3	適応形指数ステップサイズ	19
3.3	シミュレーション	21
3.3.1	シミュレーション条件	21
3.3.2	固定形指数ステップサイズ	24
3.3.3	適応形指数ステップサイズ	24
3.3.4	実測データによるシミュレーション	26
第4章	フィードフォワード形 BSS (FF-BSS)	31
4.1	回路構成と入出力関係	31
4.1.1	回路構成	31
4.1.2	時間領域	33

4.1.3	z 領域	33
4.2	学習アルゴリズム	35
4.2.1	評価関数	35
4.2.2	$W_{jk}(z)$ の時間領域での更新式	35
4.2.3	$W_{jk}(z)$ の周波数領域での更新式	37
第 5 章	信号分離と信号歪み	39
5.1	信号歪みの基準	39
5.1.1	基準として適当である信号	39
5.1.2	本研究で扱う信号歪みの基準	39
5.2	FF-BSS	41
5.2.1	解の自由度	41
5.2.2	信号歪み	42
5.2.3	重み付け効果	43
5.3	FB-BSS	47
5.4	シミュレーション	50
5.4.1	シミュレーション条件	50
5.4.2	信号分離	51
5.4.3	信号歪み	52
第 6 章	従来の信号歪み抑制学習法	63
6.1	Minimal Distortion Principle に基づく BSS (MDP-BSS)	63
6.1.1	Minimal Distortion Principle(MDP)	63
6.1.2	学習アルゴリズム	66
6.1.3	MDP-BSS の問題点	68
6.2	Projection Back 法	69
6.2.1	周波数領域における信号歪み	69
6.2.2	信号の正規化	70
6.2.3	学習アルゴリズム	71
6.3	シミュレーション	73
6.3.1	シミュレーション条件	73
6.3.2	信号スペクトル	73
6.3.3	信号分離と信号歪み	77
第 7 章	信号歪抑制学習法の提案	81
7.1	完全分離かつ無歪みの条件	81
7.1.1	2チャンネルにおける条件	81
7.1.2	条件の一般化	82
7.1.3	順序問題	88
7.2	歪み抑制学習法	89
7.2.1	学習の必要性	89
7.2.2	周波数領域での実装	89
7.2.3	時間領域での実装	90

7.3	シミュレーション	93
7.3.1	表記法	93
7.3.2	シミュレーション条件	93
7.3.3	信号スペクトル	94
7.3.4	信号分離と信号歪み	95
7.3.5	近似法の有効性	97
第8章	各手法の比較	98
8.1	表記法	98
8.2	シミュレーション	98
8.2.1	シミュレーション条件	98
8.2.2	白色信号	102
8.2.3	音声	105
8.2.4	2次ARモデルによって生成された有色信号	111
8.3	考察	118
第9章	混合過程に基づくFB-BSSとFF-BSSの性能比較	119
9.1	遅延差量による性能変化	119
9.1.1	遅延の生成方法	119
9.1.2	対象となる音響空間	124
9.1.3	性能の比較	126
9.2	無遅延差における性能比較	133
第10章	結論	135
	参考文献	138

目 次

1.1	BSS	1
1.2	ビデオ会議システムの市場推移と予測	2
1.3	分離回路	5
2.1	FB-BSS の構成	8
2.2	FB-BSS のブロック図	9
2.3	$C_{21}(z)$ や $C_{12}(z)$ に用いる FIR フィルタ	9
3.1	反響音	15
3.2	反響音 $H'_{11}(z)$, $H'_{22}(z)$ を含む畳み込み形 BSS	16
3.3	各信号成分の遅延の関係 ($C'_{12}(z)$ は式 (3.2) に収束する分離回路)	17
3.4	分離回路の FIR フィルタにおいて式 (3.4), (3.2) の影響の現れるタップ番号 ($C'_{12}(z)$ は式 (3.2) に収束する分離回路)	18
3.5	インパルス応答の概形	19
3.6	反響音を含むモデルに対する伝達関数	22
3.7	反響音を含まないモデルに対する伝達関数	22
3.8	反響音を含むモデルにおける分離回路の理想インパルス応答	23
3.9	反響音を含まないモデルにおける分離回路の理想インパルス応答	23
3.10	白色信号源に対する NRR	24
3.11	適応形指数ステップサイズの指数の底の変化	25
3.12	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	25
3.13	実験に用いた部屋の配置図	26
3.14	測定の回路図	26
3.15	残留誤差	27
3.16	伝達関数	28
3.17	分離特性	29
3.18	適応形指数ステップサイズ方式におけるステップサイズの指数の底	29
3.19	学習終了時の分離回路のフィルタ係数	30
4.1	FF-BSS の構成	31
4.2	FF-BSS の回路構成	32
4.3	$W_{lk}(z)$ に用いる FIR フィルタ	32
5.1	FF-BSS の信号分離と信号歪みの関係	42
5.2	周波数領域で学習する FF-BSS の重み付け効果 (信号源スペクトルが同じ帯域に分布している場合)	45

5.3	周波数領域で学習する FF-BSS の重み付け効果 (信号源スペクトルが違 う帯域に分布している場合)	46
5.4	FB-BSS の信号分離と信号歪みの関係	49
5.5	伝達関数 (インパルス応答)	50
5.6	白色信号における信号分離の比較	51
5.7	音声信号における信号分離の比較	52
5.8	白色信号に対する係数による歪みの比較	53
5.9	音声信号に対する係数による歪みの比較	53
5.10	白色信号に対する信号スペクトルによる歪みの比較	54
5.11	音声信号に対する信号スペクトルによる歪みの比較	55
5.12	白色信号に対する学習後の歪みの比較	56
5.13	音声信号に対する学習後の歪みの比較	56
5.14	白色信号に対するスペクトルの振幅による歪みの比較	57
5.15	音声信号に対するスペクトルの振幅による歪みの比較	58
5.16	白色信号に対する相関で時間遅れを考慮したスペクトルによる歪みの比較	59
5.17	音声信号に対する相関で時間遅れを考慮したスペクトルによる歪みの比較	60
5.18	時間遅れ補正を施したスペクトルによる D_1 の比較	62
6.1	$H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル	73
6.2	FF-TIME の出力のスペクトル	75
6.3	FF-TIME (MDP) の出力信号のスペクトル	75
6.4	FF-FREQ (1) の出力信号のスペクトル	75
6.5	FF-FREQ (2) の出力信号のスペクトル	75
6.6	FF-FREQ (1+PB) の出力信号のスペクトル	76
6.7	FF-FREQ (2+PB) の出力信号のスペクトル	76
6.8	FF-FREQ (1+LPB) の出力信号のスペクトル	77
6.9	FF-FREQ (2+LPB) の出力信号のスペクトル	77
6.10	FF-TIME の SIR_1 の時間変化	78
6.11	FF-TIME (MDP) の SIR_1 の時間変化	78
7.1	$H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル	94
7.2	FF-TIME の出力のスペクトル	95
7.3	FF-TIME (DF) の出力のスペクトル	95
7.4	FF-TIME の SIR_1 の時間変化	96
7.5	FF-TIME (DF) の SIR_1 の時間変化	96
8.1	伝達関数 (インパルス応答)	100
8.2	$S_1(z)$ と $S_2(z)$ のスペクトル	100
8.3	伝達関数 (インパルス応答)	101
8.4	$S_1(z)$, $S_2(z)$ と $S_3(z)$ のスペクトル	102
8.5	白色信号に対する SIR_2 の学習曲線	102
8.6	音声に対する $H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル	105
8.7	FF-TIME の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル . . .	105

8.8	FF-TIME (MDP) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル	105
8.9	FF-TIME (DF) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル	106
8.10	FB の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル	106
8.11	FF-FREQ (1) の出力信号のスペクトル	107
8.12	FF-FREQ (1+PB) の出力信号のスペクトル	107
8.13	FF-FREQ (1+DF) の出力信号のスペクトル	107
8.14	FF-FREQ (2) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル	108
8.15	FF-FREQ (2+PB) の出力信号のスペクトル	108
8.16	FF-FREQ (2+DF) の出力信号のスペクトル	108
8.17	$H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル	111
8.18	$X_1(z)$ と $X_2(z)$ のスペクトル	111
8.19	有色信号に対する FF-TIME の出力 $y_1(n)$, $y_2(n)$ のスペクトル	111
8.20	有色信号に対する FF-TIME (MDP) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	112
8.21	有色信号に対する FF-TIME (DF) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	112
8.22	有色信号に対する FB の出力 $y_1(n)$, $y_2(n)$ のスペクトル	112
8.23	有色信号に対する FF-FREQ (1) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	113
8.24	有色信号に対する FF-FREQ (1+PB) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	114
8.25	有色信号に対する FF-FREQ (1+DF) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	114
8.26	有色信号に対する FF-FREQ (2) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	114
8.27	有色信号に対する FF-FREQ (2) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	115
8.28	有色信号に対する FF-FREQ (1+PB) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	115
8.29	有色信号に対する FF-FREQ (1+DF) の出力 $y_1(n)$, $y_2(n)$ のスペクトル	115
9.1	遅延発生前の信号	121
9.2	零挿後の信号	121
9.3	サンプリング間隔 $T/2$ の信号	122
9.4	通過域	122
9.5	シフト後の信号系列	123
9.6	ダウンサンプル後の信号系列	124
9.7	対象とする混合過程のモデル	125
9.8	伝達関数 (インパルス応答)	125
9.9	伝達関数 (インパルス応答)	126
9.10	想定する音響空間	127
9.11	音声を用いた単純遅延混合における FF-BSS と FB-BSS の性能比較	128
9.12	白色信号を用いた単純遅延混合における FF-BSS と FB-BSS の性能比較	128
9.13	音声を用いた単純遅延混合におけるセンサ間の変化による信号分離性能 の変化	129
9.14	白色信号を用いた単純遅延混合におけるセンサ間の変化による信号分離 性能の変化	129
9.15	音声における FF-BSS と FB-BSS の性能比較	130
9.16	白色信号における FF-BSS と FB-BSS の性能比較	130
9.17	音声におけるセンサ間の変化による信号分離性能の変化	131
9.18	白色信号におけるセンサ間の変化による信号分離性能の変化	131

9.19 音声における遅延時間 τ による信号分離性能の変化	131
9.20 白色信号における遅延時間 τ による信号分離性能の変化	131
9.21 音声における遅延時間 τ による信号分離性能の変化	132
9.22 白色信号における遅延時間 τ による信号分離性能の変化	132
9.23 音声に対する $H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル	133
9.24 FF-TIME (DF) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル	133
9.25 FB の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル	133

表 目 次

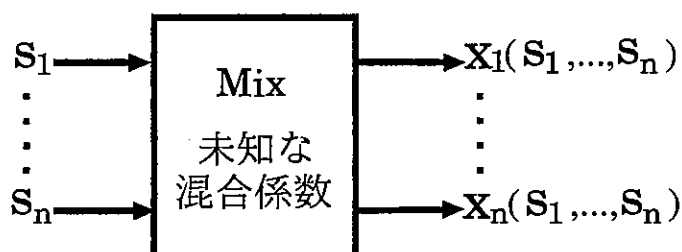
5.1	出力信号と入力または観測信号の相関の最大値	59
5.2	FF-BSS と FB-BSS の比較	62
6.1	各手法の表記方法	74
6.2	音声に対する FF-TIME と FF-TIME (MDP) の性能の比較	79
6.3	$H_{ij}(z)$, $P_{ji}(z)$ 間の相似性	79
6.4	音声に対する FF-FREQ (1),(2) と FF-FREQ (1+PB), (2+PB) の性能の 比較	79
6.5	音声に対する FF-FREQ (1+PB), (2+PB), (1+LPB), (2+LPB) 性能の 比較	80
7.1	各手法の表記方法	93
7.2	歪み抑制しない方式と提案する信号歪み抑制法の比較	97
7.3	近似法による影響の比較	97
8.1	適用する学習アルゴリズムの表記方法	99
8.2	時間領域で学習する BSS における白色信号に対する信号分離と信号歪み に関する評価	103
8.3	白色信号に対する除去すべき信号と信号歪みの基準の相似性	103
8.4	周波数領域で学習する BSS における白色信号に対する信号分離と信号歪 みに関する評価	104
8.5	時間領域で学習する BSS における音声に対する信号分離と信号歪みに関 する評価	106
8.6	周波数領域で学習する BSS における音声に対する信号分離と信号歪みに 関する評価	109
8.7	3 チャンネルの時間領域で学習する BSS における音声に対する信号分離 と信号歪みに関する評価	110
8.8	3 チャンネルの周波数領域で学習する BSS における音声に対する信号分 離と信号歪みに関する評価	110
8.9	時間領域で学習する BSS における有色信号に対する信号分離と信号歪み に関する評価	113
8.10	有色信号に対する除去すべき信号と信号歪みの基準の相似性	113
8.11	周波数領域で学習する BSS における有色信号に対する信号分離と信号歪 みに関する評価	116

8.12 3 チャンネルの時間領域で学習する BSS における有色信号に対する信号 分離と信号歪みに関する評価	116
8.13 3 チャンネルの周波数領域で学習する BSS における有色信号に対する信 号分離と信号歪みに関する評価	116
9.1 2 チャンネルの無遅延状態での信号分離と信号歪みに関する評価	134
9.2 3 チャンネルの無遅延状態での信号分離と信号歪みに関する評価	134

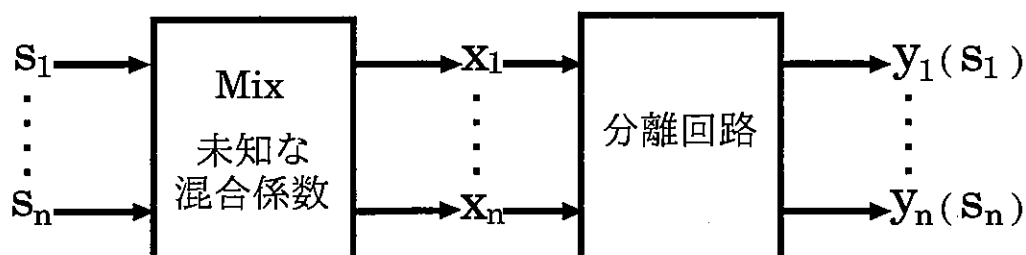
第1章 序論

1.1 ブラインド信号源分離とは

複数の音が混ざった状態から一つの音を取り出したい場合がある (図 1.1(a)). 例えば, 会議中の録音やTV会議で複数の人が同時に話す場合などは, 音声分離できることが望ましい. また, 様々な雑音や干渉音がある中での計測でも所望の信号だけを分離したい. さらに, 音源や音源がマイクで収録されるまでの伝達特性が不明である場合が多い (ブラインド信号処理問題). このような場合に, 音源の統計的な性質に着目して音源を分離する方法としてブラインド信号源分離 (Blind Source Separation: BSS) がある (図 1.1(b)) [1]-[8], [11]-[13], [22].



(a) ブラインド信号源分離の導入前



(b) ブラインド信号源分離の導入後

図 1.1: BSS

1.2 ブラインド信号分離の応用

BSS はさまざまな分野に応用が考えられる。例えば、

- 会議を録音し、話者の音声を分離することによって会議の内容をクリアにする
- カーナビゲーションシステムなどで音声入力をする際に運転手の言葉のみに分離し認識を行なう
- 医療計測器などで状況に応じた有為な波形を分離する

などが考えられる。中でも特にテレビ会議システムやテレビ電話とよばれる映像や音声を遠隔地で相互にやりとりする技術に対する応用が現在注目されている [42]。1984年に商用サービスが開始されて以来、ADSL 網や光ネットワークなどの通信インフラの普及も手伝い、交通費の削減や、移動時間の短縮のためテレビ会議システムの需要は高まっている。また大学などの公開講座を遠隔地で受講する手段としても用いられてきている。また 2004 年 2 月からは NTT コミュニケーションズの CoDen コンセプトに基づくテレビ電話、ドットフォンパーソナルのサービスが開始されテレビ会議システムは個人での使用も増えてきている。

テレビ会議システムを使用する際、複数対複数の会話が行なわれる可能性が非常に高い。その場で行なわれている会話であれば複数の人が会話している中で特定の話者の声を聞き取ることは音源を特定し易いため比較的容易にできるが、一度マイクに集音した音では音源を特定し難いため話者を特定するのが難しくなる。そこでブラインド信号源分離により音声を分離することにより各スピーカに音源を割り当てて、聞き取り易くすることができる。またテレビ会議システムを利用して会議の議事録などを作る際の手助けになる。

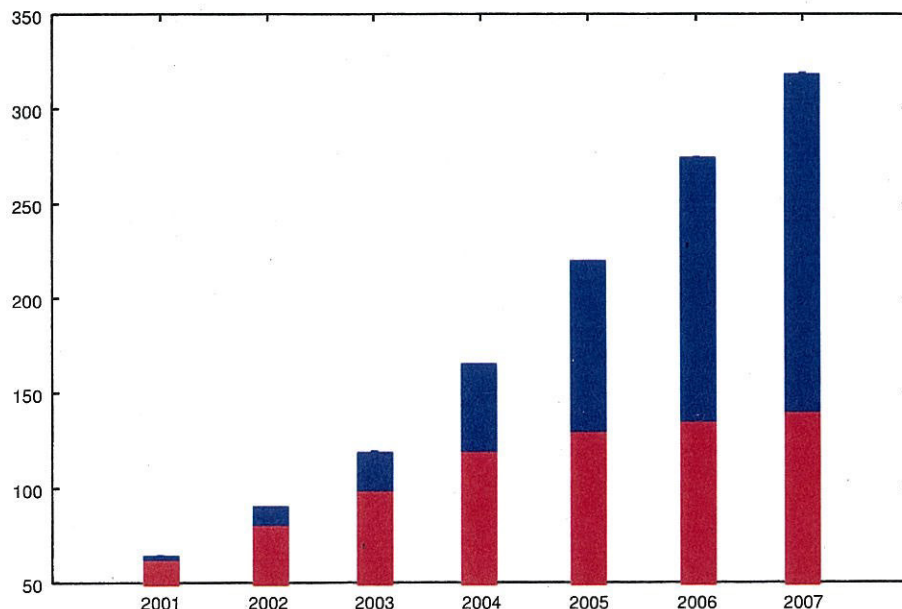


図 1.2: ビデオ会議システムの市場推移と予測

テレビ会議システムには専用のテレビ会議システムのハードウェアを使用するハードウェア型と既存のパソコンやサーバにソフトウェアをインストールして使用するソフトウェア型がある。これらの市場は図1.2にあるように順調に成長し2003年には120億円の市場になっており、4年後には3倍の330億円の市場になると予測されている[37]。

現在は企業や大学など法人の利用が主であるため数十万から数百万する高価な専用ハードウェアを使用するハードウェア型が大半を占めているが、今後は一般向けの高価な機材を必要とせず、既存の安価な機材で利用できるソフトウェア型が主流を占めてくると予測されている。ソフトウェア型では単に通信を行なうだけではなくさまざまなサービスが求められる。その中でも通信の品質を保つためにブラインド信号源分離は必須のサービスとなるだろう。

1.3 研究の位置付け

BSSとは、複数の未知信号が混成した状態を複数のセンサで観測し、それらの混成信号から元の信号を分離することである。しかし、これだけでは問題を解決するのは困難である。よって、BSSには信号源は互いに統計的独立であるという仮定が加えられている。さらに、研究の際には実環境をモデル化し、問題を分割している場合が多い。モデル化の例として

- 瞬時混合
- 畳み込み混合
- 非線形混合

などが挙げられる。

瞬時混合は混合過程を信号源の線形和で表している [1]-[3],[5]-[8],[16],[19],[24],[31]。すなわち、信号源 $s_1(n), \dots, s_I(n)$ に対し、観測信号 $x_1(n), \dots, x_J(n)$ は a_{ij} を用いて

$$x_j(n) = \sum_i a_{ji} s_i(n) + r(n) \quad (1.1)$$

と表すことができる。ここで $r(n)$ はノイズ成分である。瞬時混合は混合過程として最もシンプルなモデル化を行なったものである。したがって、信号源、観測信号、出力信号の関係がもっともシンプルであり、BSSの基礎的な研究を行なう際に有益である。

畳み込み混合は混合過程を FIR フィルタや IIR フィルタなどを使って畳み込みで表現される [10],[14],[15],[17],[20],[21]。すなわち、信号源 $s_i(n), \dots, s_i(n - M_{ji} + 1)$ に対し、観測信号 $x_j(n)$ は $h_{ji}(0), \dots, h_{ji}(M_{ji} - 1)$ を用いて

$$x_j(n) = \sum_{i=1}^N \sum_{m=0}^{M_{ji}-1} h_{ji}(m) s_i(n-m) + r(n) \quad (1.2)$$

と表すことができる [32]。畳み込み混合は比較的単純なモデルで実際の音響空間に近いものを表現できるため実際の状況に近いシミュレーションを行なうことができる。

非線形混合は混合過程に非線形性を有するものである。したがって、信号源 $s_i(n)$ に対し、観測信号 $x_j(n)$ は非線形関数 $f(x)$ を用いて

$$x_j(n) = f(s_1(n), \dots, s_I(n)) + r(n) \quad (1.3)$$

と表すことができる。ただし、 a_{ji} を用いて

$$x_j(n) = f\left(\sum_i a_{ji} s_i(n)\right) + r(n) \quad (1.4)$$

とされる場合が多い。非線形混合はスピーカなどの非線形性を表現することができ、より幅広い応用に対応できるようにする。

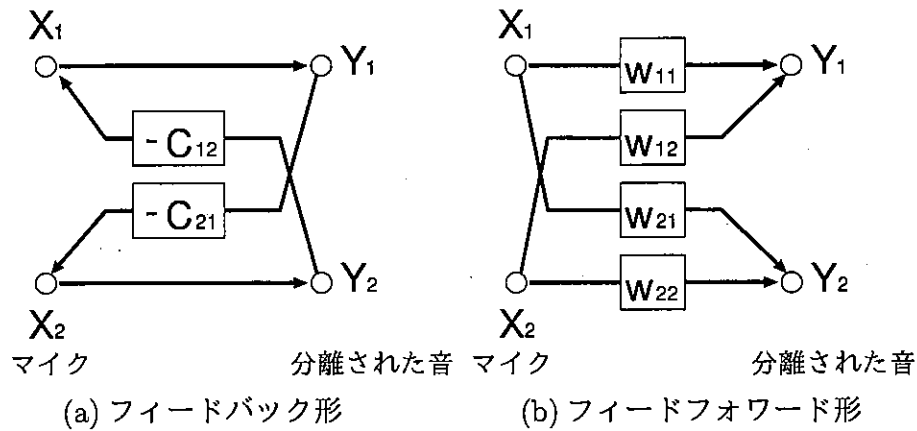


図 1.3: 分離回路

さらに、畳み込み混合において処理の領域として

- 時間領域
- 周波数領域

が考えられる。周波数領域での処理は畳み込みが周波数領域では乗算になることを利用している。このため各周波数で独立に瞬時混合のアルゴリズムを使用することができ、比較的拡張が容易である。しかし、BSSには潜在的に順序問題を内包しており、各周波数で処理を独立におこなって分離を行なった場合でも、結合時に他の周波数成分と結合してしまう恐れがある。一方、時間領域では信号を一括して処理を行なうため、順序問題が信号の成分に影響することはない。しかし、信号を一括で処理するためアルゴリズムが複雑になってしまう。

また現在分離回路としてフィードバック形(図 1.3(a)), フィードフォワード形(図 1.3(b))の2種類が提案されている。フィードバック形の分離回路では出力を観測信号から減ずることによって分離を行う。この回路形式では出力をフィードバックさせているため安定性について考慮する必要がある。また、フィードフォワード形の分離回路では観測信号を順方向に流す際に、分離回路を設けている。この回路形式では全ての経路に対して学習回路を設けているため、出力信号の歪みについて言及する必要がある。

そこで本研究では畳み込み混合形のBSSに対して、時間領域で処理を行なうものを中心に、回路形式による特性の違いというものに着目し、その際問題となる反響音の問題や信号歪みについて検討する。さらに反響音の影響や信号歪みを軽減するための学習法を提案する。

1.4 論文の構成

本論文の構成は以下のようになっている。

まず、現在提案されてるフィードバック形のブラインドソースセパレーションについて、その回路構成、入出力関係、学習方式など基本的事項を第2章で述べる。

フィードバック形のブラインドソースセパレーションは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、この際、各センサへの伝達遅延に十分な差があることが前提となる。しかし、反響音を含む畳み込み混合ではその前提が崩れるため良好なブラインドソースセパレーションを実現することが難しい。そこで、第3章で反響音の影響を抑制する学習法を提案する。まず、反響音の影響について解析を行う。そして、それに基づき反響音の影響が大きく現れる分離回路のタップの後半部分の学習を抑制する指数重み付け(EW) 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、EW ステップサイズの指数重みを自動的に調整する手法を提案する。

フィードバック形のブラインドソースセパレーションと同様に、現在提案されているフィードフォワード形のブラインドソースセパレーションについて、その回路構成、入出力関係、学習方式など基本的事項を第4章で述べる。フィードフォワード形ブラインドソースセパレーションについては時間領域で信号を分離する方式に加え、現在多く研究がなされている短時間フーリエ変換を用いて周波数領域で信号分離を行なう方式についても紹介する。

第5章において、信号歪みの基準について議論し、その議論に基づき適切な信号歪みの基準を定める。さらにその信号歪みの基準に基づき第2, 4章で紹介したフィードバック形ブラインドソースセパレーションとフィードフォワード形ブラインドソースセパレーションについて信号分離と信号歪みの解析を行なう。

第5章で問題になるフィードフォワード形ブラインドソースセパレーションに対する信号歪みに対して従来から信号歪み抑制学習法が提案されている。第6章では時間領域、周波数領域での従来の信号歪み抑制学習法を紹介し、そのアルゴリズムの問題点を指摘する。

続く第7章では、新たにフィードフォワード形ブラインドソースセパレーションに対する信号歪みに対してそれを抑制する信号歪み抑制学習法を提案する。これは完全分離かつ無歪みの条件から導く信号歪みを抑制する制約を従来の学習法に加えるものである。この制約は簡単化のためまず2チャンネルに対して考察する。その後、一般的なマルチチャンネルのブラインドソースセパレーションに対して拡張する。この制約は時間領域で学習する方式、周波数領域で学習する方式のどちらにも付加することが可能である。

そして第8章で、第2, 4, 6, 7章で紹介、提案した手法を比較し、提案した信号歪み抑制学習法の有効性を確認する。

一方、フィードバック形ブラインドソースセパレーションは第5章の解析の結果、信号分離の解と無歪みの解が同一であるため、分離性能が高く、歪みが起き難い。しかし、フィードバック形ブラインドソースセパレーションの学習では混合過程の伝達遅延の差にある仮定をおき、それに基づいてアルゴリズムが形成されているため、その仮定が崩れると性能が著しく劣化する。そこで第9章では、混合過程の伝達遅延の差

に着目し，フィードバック形ブラインドソースセパレーションとフィードフォワード形ブラインドソースセパレーションを比較し，それぞれの有効性について検討を行なう．最後に，以上で得られた結果を元に第10章で本論文の結論を述べる．

第2章 フィードバック形BSS (FB-BSS)

ブラインドソースセパレーションはモデル化方法，処理領域，回路構成などでさまざまな種類に分類することができる．本章では畳み込み形のフィードバック形ブラインドソースセパレーション (FB-BSS) の回路構成や学習アルゴリズムを説明する．

2.1 回路構成と入出力関係

2.1.1 回路構成

Jutten らが提案した FB-BSS の回路構成 [6],[10] を図 2.1 に示す．

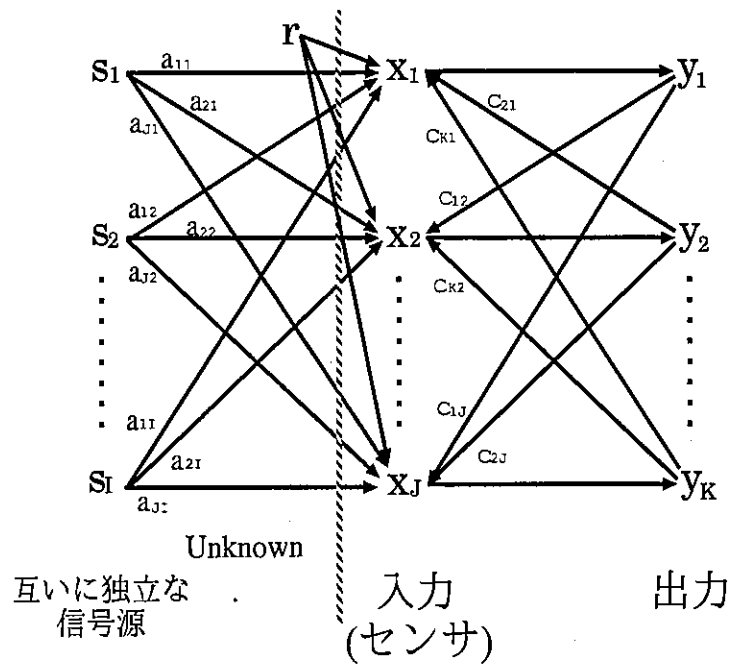


図 2.1: FB-BSS の構成

r はノイズ成分である． I 個の信号源 $s_i(n)$ はそれぞれ混合過程において a_{ji} の影響を受け J 個のセンサで観測される．その際，観測信号には観測雑音 r が含まれる．観測信号 $x_j(n)$ は K 個の独立成分 $y_k(n)$ に分離される．

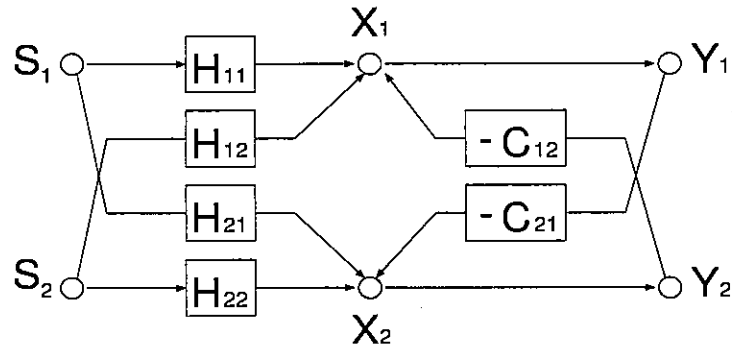
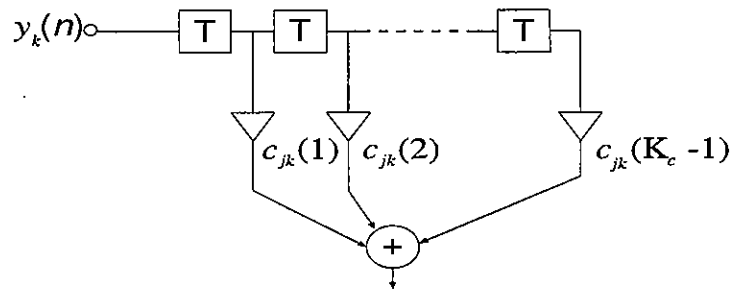


図 2.2: FB-BSS のブロック図

図 2.3: $C_{21}(z)$ や $C_{12}(z)$ に用いる FIR フィルタ

本研究では簡単化のため信号源，センサと出力の数はすべて同じであるとする．また，観測信号にノイズが含まれないものとする．本研究で使用する FB-BSS のモデルにおいて 2 チャンネルの場合のブロック図を図 2.2 に示す．ここで，本研究では畳み込み形 BSS を扱うため，混合過程は畳み込みの形をしていると仮定する．分離回路には図 2.3 に示す FIR フィルタを用いる．

2.1.2 時間領域

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。

$$x_j(n) = \sum_{i=1}^N \sum_{m=0}^{M_{ji}-1} h_{ji}(m) s_i(n-m) \quad (2.1)$$

また、分離過程の出力 $y_j(n)$ は次式で表される。

$$y_j(n) = x_j(n) - \sum_{\substack{k=1 \\ k \neq j}}^N \sum_{l=1}^{L_{jk}-1} c_{jk}(l) y_k(n-l) \quad (2.2)$$

この関係をベクトルと行列で表すと次のようになる。

$$\mathbf{x}(n) = \mathbf{H}^T \mathbf{s}(n) \quad (2.3)$$

$$\mathbf{y}(n) = \mathbf{x}(n) - \mathbf{C}^T \tilde{\mathbf{y}}(n) \quad (2.4)$$

ただし、 $\mathbf{s}(n), \mathbf{x}(n), \mathbf{y}(n), \tilde{\mathbf{y}}(n), \mathbf{H}, \mathbf{C}$ は

$$\mathbf{s}(n) = [s_1^T(n), s_2^T(n), \dots, s_N^T(n)]^T \quad (2.5)$$

$$\mathbf{s}_i(n) = [s_i(n), s_i(n-1), \dots, s_i(n-M_i+1)]^T \quad (2.6)$$

$$\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_N(n)]^T \quad (2.7)$$

$$\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_N(n)]^T \quad (2.8)$$

$$\tilde{\mathbf{y}}(n) = [\mathbf{y}_1^T(n), \mathbf{y}_2^T(n), \dots, \mathbf{y}_N^T(n)]^T \quad (2.9)$$

$$\mathbf{y}_k(n) = [y_k(n), y_k(n-1), \dots, y_k(n-L_{jk}+1)] \quad (2.10)$$

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{21} & \dots & h_{N1} \\ h_{12} & h_{22} & \dots & h_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1N} & h_{2N} & \dots & h_{NN} \end{bmatrix} \quad (2.11)$$

$$\mathbf{h}_{ji} = [h_{ji}(0), h_{ji}(1), \dots, h_{ji}(M_{ji}-1)]^T \quad (2.12)$$

$$\mathbf{C} = \begin{bmatrix} 0 & c_{21} & \dots & c_{N1} \\ c_{12} & 0 & \dots & c_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1N} & c_{2N} & \dots & 0 \end{bmatrix} \quad (2.13)$$

$$\mathbf{c}_{jk} = [c_{jk}(0), c_{jk}(1), \dots, c_{jk}(L_{jk}-1)]^T \quad (2.14)$$

である。

ここで自分自身へのフィードバックがないため、対角要素が0になる。

2.1.3 z 領域

$s_i(n)$, $x_j(n)$, $y_k(n)$ 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表せる.

$$\mathbf{X}(z) = \mathbf{H}(z)\mathbf{S}(z) \quad (2.15)$$

$$\mathbf{Y}(z) = \mathbf{X}(z) - \mathbf{C}(z)\mathbf{Y}(z) \quad (2.16)$$

ここで $\mathbf{S}(z)$, $\mathbf{X}(z)$, $\mathbf{Y}(z)$, $\mathbf{H}(z)$, $\mathbf{C}(z)$, はそれぞれ,

$$\mathbf{S}(z) = [S_1(z), S_2(z), \dots, S_N(z)]^T \quad (2.17)$$

$$\mathbf{X}(z) = [X_1(z), X_2(z), \dots, X_N(z)]^T \quad (2.18)$$

$$\mathbf{Y}(z) = [Y_1(z), Y_2(z), \dots, Y_N(z)]^T \quad (2.19)$$

$$\mathbf{H}(z) = \begin{bmatrix} H_{11}(z) & H_{12}(z) & \dots & H_{1N}(z) \\ H_{21}(z) & H_{22}(z) & \dots & H_{2N}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}(z) & H_{N2}(z) & \dots & H_{NN}(z) \end{bmatrix} \quad (2.20)$$

$$\mathbf{C}(z) = \begin{bmatrix} 0 & C_{12}(z) & \dots & C_{1N}(z) \\ C_{21}(z) & 0 & \dots & C_{2N}(z) \\ \vdots & \vdots & \ddots & \vdots \\ C_{N1}(z) & C_{N2}(z) & \dots & 0 \end{bmatrix} \quad (2.21)$$

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.

$$\begin{aligned} \mathbf{Y}(z) &= (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{X}(z) \\ &= (\mathbf{I} + \mathbf{C}(z))^{-1} \mathbf{H}(z) \mathbf{S}(z) \end{aligned} \quad (2.22)$$

ここで分離行列 $\mathbf{W}(z)$ は以下のように定義できる.

$$\mathbf{W}(z) = (\mathbf{I} + \mathbf{C}(z))^{-1} \quad (2.23)$$

分離性能を評価するために以下の行列を定義する.

$$\mathbf{P}(z) = \mathbf{W}(z)\mathbf{H}(z) \quad (2.24)$$

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. 2 チャンネルの場合をここに示す.

$$\mathbf{P}(z) = \begin{bmatrix} P_{11}(z) & 0 \\ 0 & P_{22}(z) \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 0P_{12}(z) \\ P_{21}(z) & 0 \end{bmatrix} \quad (2.25)$$

$P_{ki}(z)$ のインパルス応答 $p_{ki}(n)$ が

$$p_{ki}(n) = 0 \quad n \neq n_0 \quad (2.26)$$

であったとすれば分離した信号は

$$y_k(n) = p_{ki}(n_0)s(n - n_0) \quad (2.27)$$

となる。これは出力が $s_i(n)$ を時間的にシフトし、定数倍されただけのものであり、分離信号において歪みが生じないことを表している。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。

$$Y_j(z) = P_{ji}(z)S_i(z) \quad (2.28)$$

2.2 学習アルゴリズムの提案

2.2.1 伝搬遅延に関する仮定

簡単化のため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2.3にあるように X_1 のセンサが $s_1(n)$ の近くに、 X_2 のセンサが $s_2(n)$ の近くに配置されていることを意味する．

完全に分離が行なわれる解は次のような場合である．

$$(1) \quad C_{21}(z) = \frac{H_{21}(z)}{H_{11}(z)} \quad C_{12}(z) = \frac{H_{12}(z)}{H_{22}(z)} \quad (2.29)$$

$$y_1(n) = \mathbf{h}_{11}^T \mathbf{s}_1(n) \quad y_2(n) = \mathbf{h}_{22}^T \mathbf{s}_2(n) \quad (2.30)$$

$$(2) \quad C_{21}(z) = \frac{H_{22}(z)}{H_{12}(z)} \quad C_{12}(z) = \frac{H_{11}(z)}{H_{21}(z)} \quad (2.31)$$

$$y_1(n) = \mathbf{h}_{12}^T \mathbf{s}_2(n) \quad y_2(n) = \mathbf{h}_{21}^T \mathbf{s}_1(n) \quad (2.32)$$

$H_{ji}(z)$ における伝搬遅延の仮定から解(1)は因果性を満たすため実現可能である．一方、解(2)は満たさないため、実現できない．

またこれらの解から、分離後の出力 $y_k(n)$ は信号源 $s_i(n)$ と正確には一致せず、 $H_{11}(z)$ や $H_{22}(z)$ の影響が残る．

2.2.2 評価関数

式(2.22)から出力は次のように表せる．

$$\begin{aligned} \begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} &= \frac{1}{1 - C_{12}(z)C_{21}(z)} \begin{bmatrix} 1 & -C_{12}(z) \\ -C_{21}(z) & 1 \end{bmatrix} \\ &\times \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \end{aligned} \quad (2.33)$$

$$\begin{aligned} &= \frac{1}{1 - C_{12}(z)C_{21}(z)} \\ &\times \begin{bmatrix} H_{11}(z) - C_{12}(z)H_{21}(z) & H_{12}(z) - C_{12}(z)H_{22}(z) \\ H_{21}(z) - C_{21}(z)H_{11}(z) & H_{22}(z) - C_{21}(z)H_{12}(z) \end{bmatrix} \\ &\times \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \end{aligned} \quad (2.34)$$

2.2.1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．また $C_{12}(z)$ と C_{21} は正の伝搬遅延をもつ因果性を満たす回路である．したがって式(2.34)の対角要素は0にならない．逆に非対角要素は $C_{12}(z)$ や C_{21} を調整することによって0になり得る．

したがって、コスト関数を以下のように設定できる．

$$J_j(n) = E[q(y_j(n))] \quad (2.35)$$

$q()$ はただ1つの極小値を持つ偶関数である。コスト関数を最小化することによって、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。適応フィルタに対する 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。

$$\hat{J}_j(n) = q(y_j(n)) \quad (2.36)$$

2.2.3 $C_{jk}(z)$ の更新式

$\hat{J}_j(n)$ の勾配は次式のようになる。

$$\begin{aligned} \frac{\partial \hat{J}_j(n)}{\partial c_{jk}(l)} &= \frac{\partial q(y_j(n))}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial c_{jk}(l)} \\ &= \dot{q}(y_j(n)) y_k(n-l) \end{aligned} \quad (2.37)$$

$$y_j(n) = x_j(n) - \sum_{l=1}^{L_{jk}-1} c_{jk}(l) y_k(n-l) \quad (2.38)$$

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。

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \Delta c_{jk}(n, l) \quad (2.39)$$

$$\Delta c_{jk}(n, l) = \mu \dot{q}(y_j(n)) y_k(n-l) \quad (2.40)$$

2.2.4 統計的な収束特性

信号源の確率密度関数 (pdf) が偶関数で、信号源が互いに独立であるとすれば以下の式を満足する [25],[26]。

$$E[f(s_1(n))g(s_2(n))] = E[f(s_1(n))]E[g(s_2(n))] \quad (2.41)$$

$$= 0 \quad (2.42)$$

$f(), g()$: 奇関数

式(2.40)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(2.42)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。

第3章 FB-BSSにおける反響音による影響とその改善法

第2章においてFB-BSSでは混合過程における伝搬遅延に関して満たさなければならない仮定があることを述べた。しかし、反響音が含まれる場合にはその仮定が崩れる場合がある。

本章では反響音による影響を解析し、その影響を軽減する方式を提案する。また、その方式で必要な混合過程の伝達関数の包絡線の推定を自動で行なう方式についても提案する。

3.1 反響音による影響

3.1.1 反響音

反響音とは図3.1に示すように直接マイクなどのセンサに入力されず、壁などに跳ね返ってセンサに入力される音のことである。

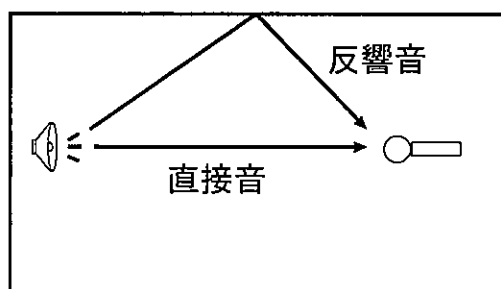
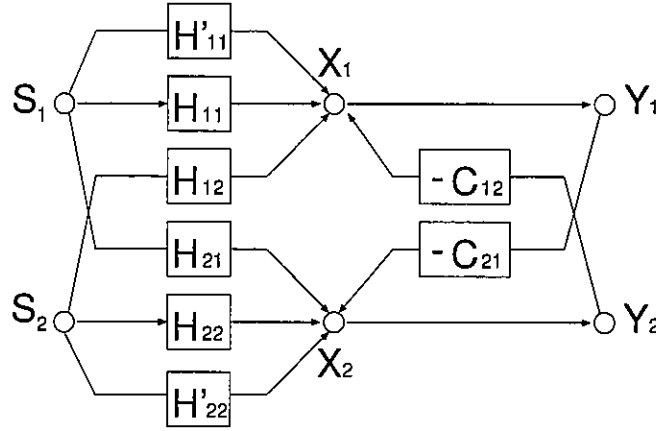


図 3.1: 反響音

図3.1からも分かるように直接音に比べ通る経路が長い。そのため伝搬遅延が大きくなる。また伝達経路として一度壁などに跳ね返っているため直接音に比べパワーが小さくなる。

3.1.2 反響音による影響

反響音を含む場合、振幅の条件と遅延の条件が一致しない可能性がある。反響音を含むモデルを図3.2に示す。

図 3.2: 反響音 $H'_{11}(z)$, $H'_{22}(z)$ を含む畳み込み形 BSS

$H'_{11}(z)$ と $H'_{22}(z)$ は反響音による伝達関数を表していて、これらは遅延時間が直接音に比べて長いとする。 $H'_{12}(z)$ と $H'_{21}(z)$ は今の議論には関係しないのでここでは無視する。 X_1 のパワーを減少させるのは以下の2つの項である。

$$H'_{11}(z)S_1(z) - C_{12}(z)H_{21}(z)S_1(z) \rightarrow 0 \quad (3.1)$$

$$C_{12}(z) \rightarrow \frac{H'_{11}(z)}{H_{21}(z)} \quad (3.2)$$

$$H_{12}(z)S_2(z) - C_{12}(z)H_{22}(z)S_2(z) \rightarrow 0 \quad (3.3)$$

$$C_{12}(z) \rightarrow \frac{H_{12}(z)}{H_{22}(z)} \quad (3.4)$$

また X_2 についても同様である。

$$H'_{22}(z)S_2(z) - C_{21}(z)H_{12}(z)S_2(z) \rightarrow 0 \quad (3.5)$$

$$C_{21}(z) \rightarrow \frac{H'_{22}(z)}{H_{12}(z)} \quad (3.6)$$

$$H_{21}(z)S_1(z) - C_{21}(z)H_{11}(z)S_1(z) \rightarrow 0 \quad (3.7)$$

$$C_{21}(z) \rightarrow \frac{H_{21}(z)}{H_{11}(z)} \quad (3.8)$$

式 (3.4) と式 (3.8) は理想解である。しかし、式 (3.2)、式 (3.6) で与えられる反響音によって $C_{12}(z)$ と $C_{21}(z)$ がこれらの解に近付かなくなってしまう。

3.2 指数ステップサイズ方式の提案

3.2.1 固定形指数ステップサイズ

反響音によって理想解から学習の方向がずれてしまうため、分離性能が劣化してしまう。そこで、反響音の影響を抑える方法を提案する [27],[28],[33]-[35],[40]。

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 3.3 のようになる。式 (3.4)

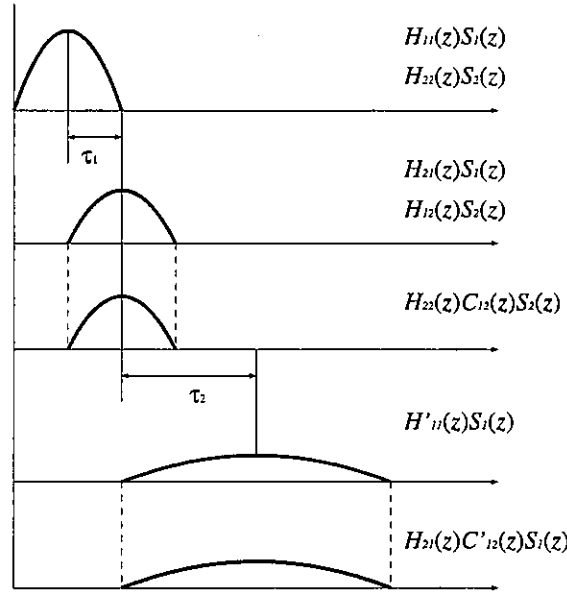


図 3.3: 各信号成分の遅延の関係 ($C'_{12}(z)$ は式 (3.2) に収束する分離回路)

と式 (3.8) 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 (3.2)、式 (3.6) 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。反響音は直接音に比べ振幅が小さく、反響の時間が長いため、 $\tau_1 < \tau_2$ であると考えることができる。したがって、分離回路の FIR フィルタのタップ係数が式 (3.2)、(3.4)、(3.6)、(3.8) により更新される度合は図 3.4 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。

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 フィルタの遅延線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。

分離回路の最適値は式 (3.4)、(3.8) で与えられる。さらにこれらの式は以下のように

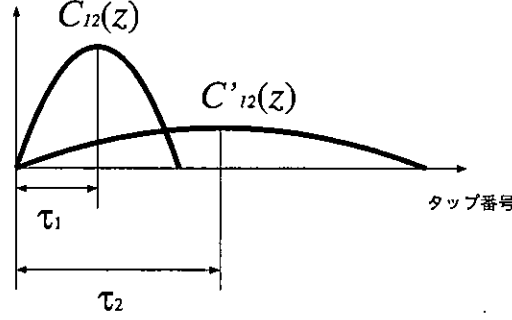


図 3.4: 分離回路の FIR フィルタにおいて式 (3.4), (3.2) の影響の現れるタップ番号 ($C'_{12}(z)$ は式 (3.2) に収束する分離回路)

変形できる.

$$C_{jk}^{opt}(z) = \frac{H_{jk}(z)}{H_{kk}(z)} = H_1(z) + \frac{H_3(z)}{H_2(z)} \quad (3.9)$$

$$= H_1(z) + \sum \frac{1}{1 - z_i z^{-1}} \quad (3.10)$$

$$= H_1(z) + \sum \sum_l z_i^l z^{-l} \quad (3.11)$$

$$= H_1(z) + \sum \sum_l r_i^l e^{j\theta_i l} z^{-l}$$

$H_1(z)$ 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 項は式 (3.9) を部分分数展開し, テーラー展開しており, 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.

$$c_{jk}(n+1, l) = c_{jk}(n, l) + \mu(l) f(y_j(n)) g(y_k(n-l)) \quad (3.12)$$

$$\mu(l) = \mu_0 r^l, \quad 0 < r < 1 \quad (3.13)$$

ステップサイズ $\mu(l)$ は理想解に比例するように決める.

3.2.2 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 [9]. 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.

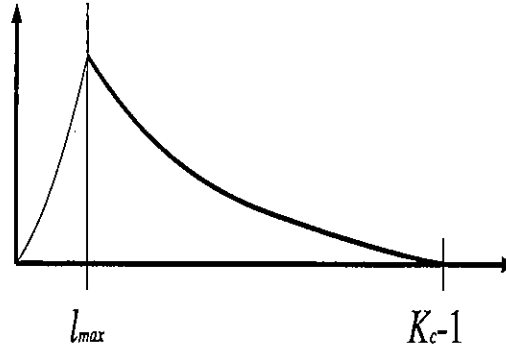


図 3.5: インパルス応答の概形

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 [35]。

3.2.3 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の方式 [9] 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 [28],[30]。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。

部屋のインパルス応答の包絡線はおおよそ図 3.5 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。

$$\mu(n, l) = \mu_0(n)r^l(n) \quad (3.14)$$

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。

$$\begin{aligned} \log \mu(n, l) &= \log \mu_0(n) + l \log r(n) \\ &= x_1(n) + l x_2(n) \end{aligned} \quad (3.15)$$

$$b(n, l) = \log |c_{jk}(n, l)| \quad (3.16)$$

最小2乗法により $x_1(n) + lx_2(n)$ を $b(n, l)$ に近づける.

$$x_1(n) + lx_2(n) = b(n, l) \quad (3.17)$$

$$l = l_{max} \sim K_c - 1$$

$$\mathbf{A}\mathbf{x}(n) = \mathbf{b}(n) \quad (3.18)$$

$$\mathbf{A} = \begin{bmatrix} 1 & l_{max} \\ 1 & l_{max} + 1 \\ \vdots & \vdots \\ 1 & K_c - 1 \end{bmatrix} \quad (3.19)$$

$$\mathbf{x}(n) = \begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} \quad (3.20)$$

$$\mathbf{b}(n) = \begin{bmatrix} \log |c_{jk}(n, l_{max})| \\ \log |c_{jk}(n, l_{max} + 1)| \\ \vdots \\ \log |c_{jk}(n, K_c - 1)| \end{bmatrix} \quad (3.21)$$

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. 最小2乗解は以下のようなになる.

$$\mathbf{x}(n) = \mathbf{A}^+ \mathbf{b}(n) \quad (3.22)$$

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \quad (3.23)$$

この結果から, $r(n)$, $\mu_0(n)$ とステップサイズ $\mu(n, l)$ は以下のようなになる.

$$\mu_0(n) = e^{x_1} \quad (3.24)$$

$$r(n) = e^{x_2} \quad (3.25)$$

$$\hat{r}(n) = \alpha r(n) + (1 - \alpha) \hat{r}(n - 1) \quad (3.26)$$

$$0 < \alpha \ll 1$$

$$\mu(n, l) = \mu_0(n) \hat{r}^l(n) \quad (3.27)$$

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.

3.3 シミュレーション

3.3.1 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．

$$f(y) = \tanh(2.5y) \quad g(y) = \tanh(0.5y) \quad (3.28)$$

分離性能は混合過程の伝達関数 $H(z)$ ，式(2.24)で定義した $P(z)$ を用い，以下のNRR(noise reduction ratio)で評価した．

$$\sigma_{ys}^2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ii}(e^{j\omega T})|^2 d\omega T \quad (3.29)$$

$$\sigma_{yc}^2 = \sum_{j \neq i} \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ji}(e^{j\omega T})|^2 d\omega T \quad (3.30)$$

$$\sigma_{xs}^2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega T})|^2 d\omega T \quad (3.31)$$

$$\sigma_{xc}^2 = \sum_{j \neq i} \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega T})|^2 d\omega T \quad (3.32)$$

$$NRR = 10 \log \frac{\sigma_{ys}^2}{\sigma_{yc}^2} - 10 \log \frac{\sigma_{xs}^2}{\sigma_{xc}^2} \quad [\text{dB}] \quad (3.33)$$

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．

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3.6，3.7に示す．図3.6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3.7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．

これら図より反響音を含む場合，含まない場合の分離回路の理想インパルス応答はそれぞれ図3.8，3.9のようになる．

反響音を含む場合，含まない場合それぞれに対して $C_{12}(z)$ ， $C_{21}(z)$ とともに理想解のインパルス応答は1024タップでカバーできる．

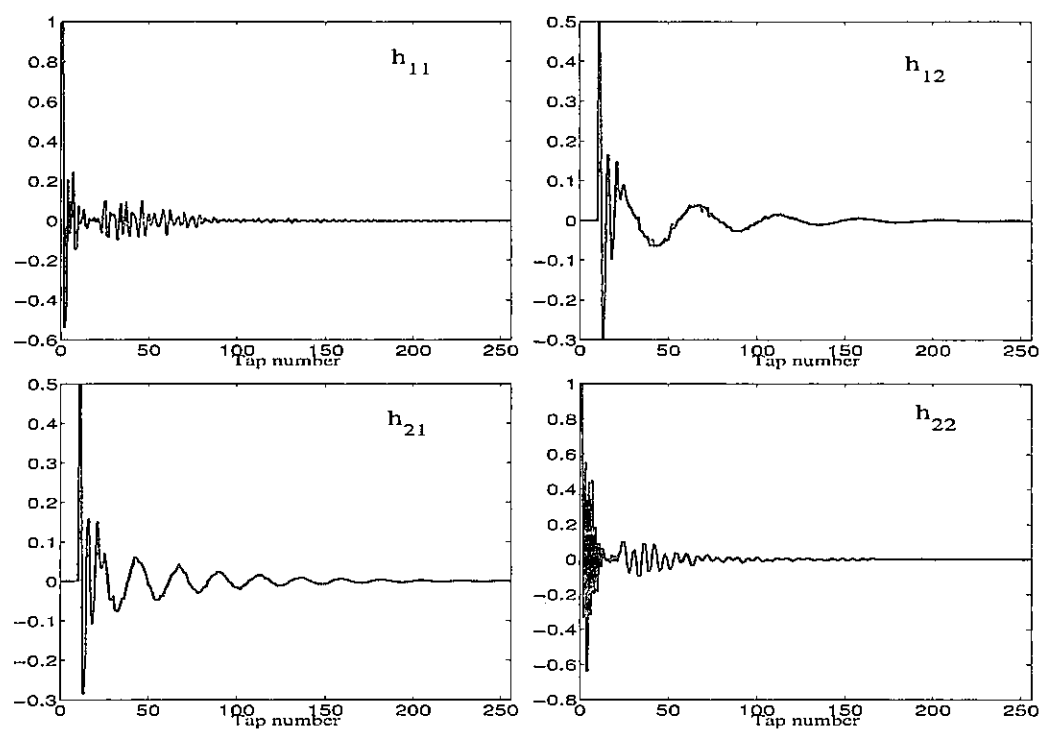


図 3.6: 反響音を含むモデルに対する伝達関数

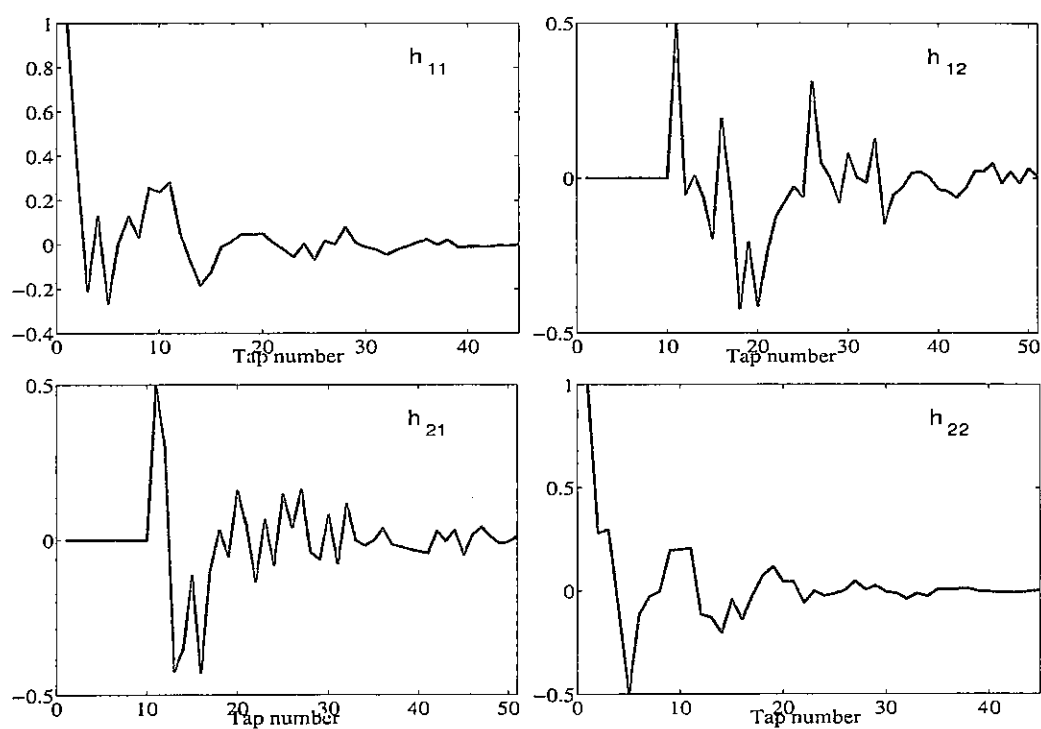


図 3.7: 反響音を含まないモデルに対する伝達関数

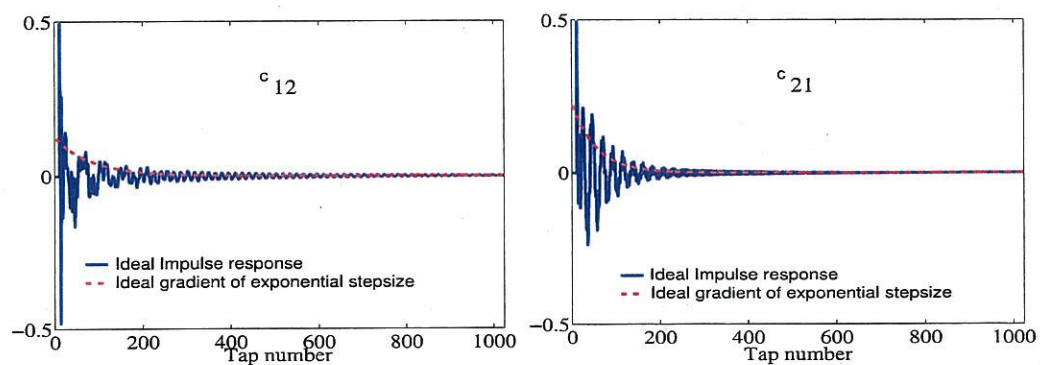


図 3.8: 反響音を含むモデルにおける分離回路の理想インパルス応答

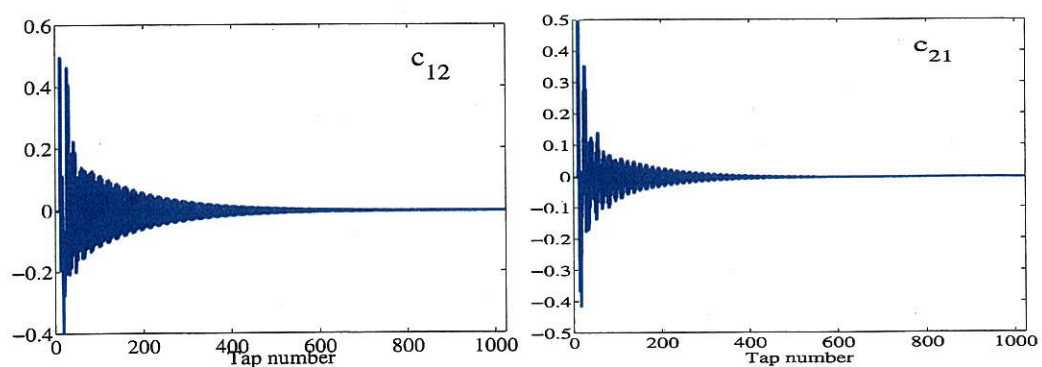


図 3.9: 反響音を含まないモデルにおける分離回路の理想インパルス応答

3.3.2 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 3.10 に示す．信号源には白色信号を使用した．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．

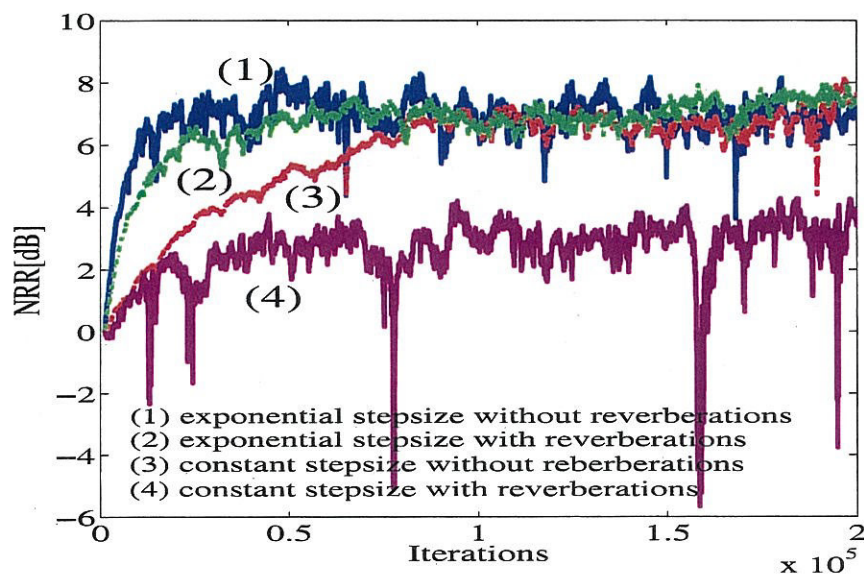


図 3.10: 白色信号源に対する NRR

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，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．

3.3.3 適応形指数ステップサイズ

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．適応形指数の底の変化を図 3.11 に示す．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 3.12 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．

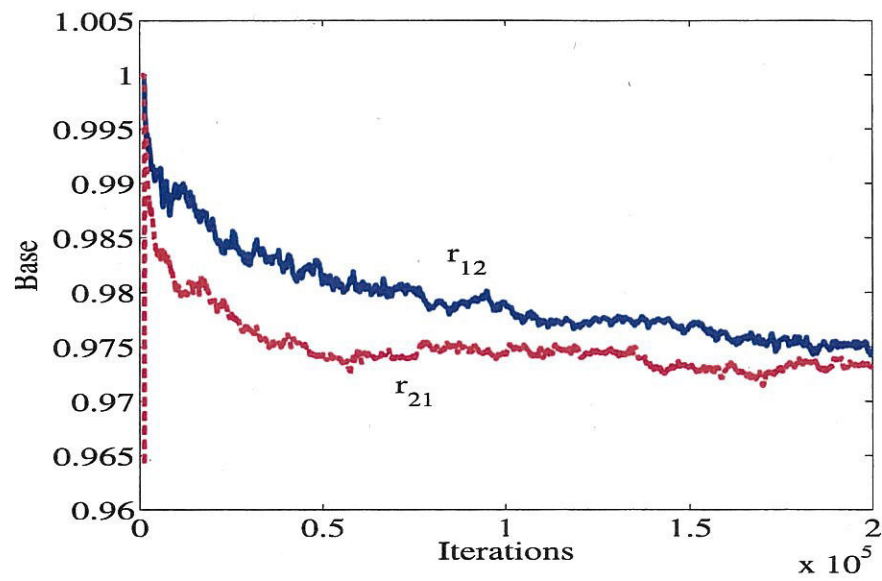


図 3.11: 適応形指数ステップサイズの指数の底の変化

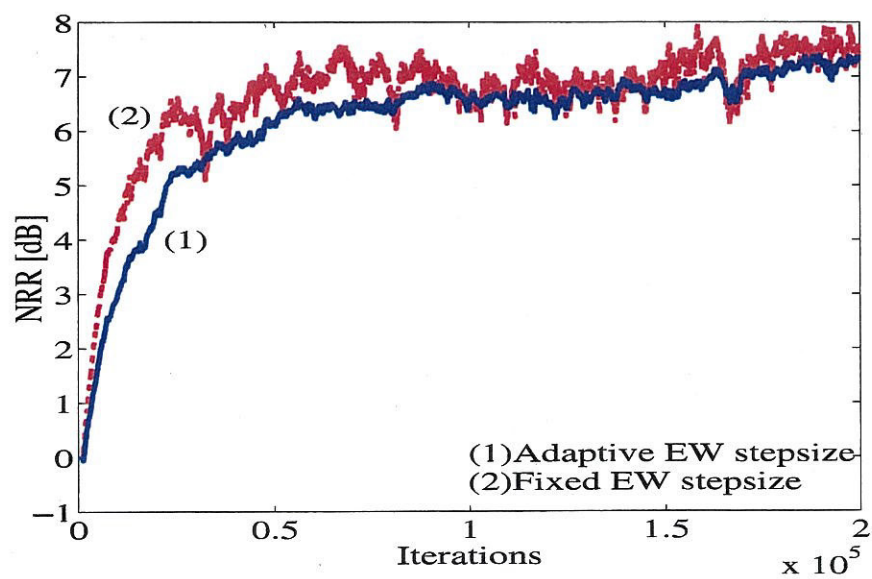


図 3.12: 適応形指数ステップサイズと固定形指数ステップサイズの比較

3.3.4 実測データによるシミュレーション

測定の様子

測定に用いた部屋の配置図を図 3.13 に示す。

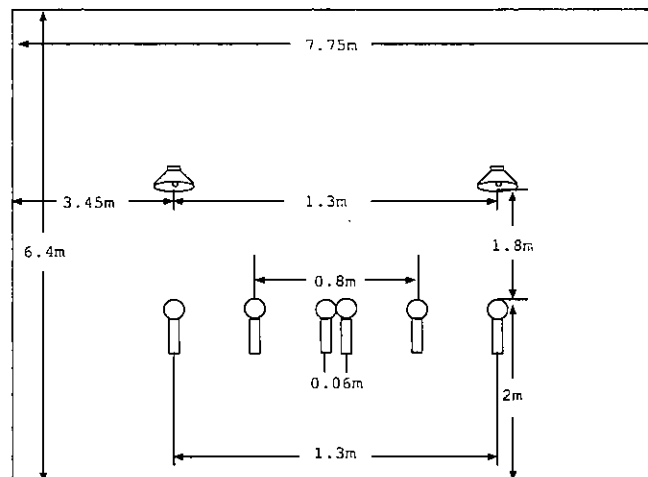


図 3.13: 実験に用いた部屋の配置図

ノートパソコンから信号を出力し、アンプで増幅したのちスピーカから出力する。同時にスピーカに出力する信号とマイクに入力される信号をテープで録音した。図 3.14 にこの構成図を示す。

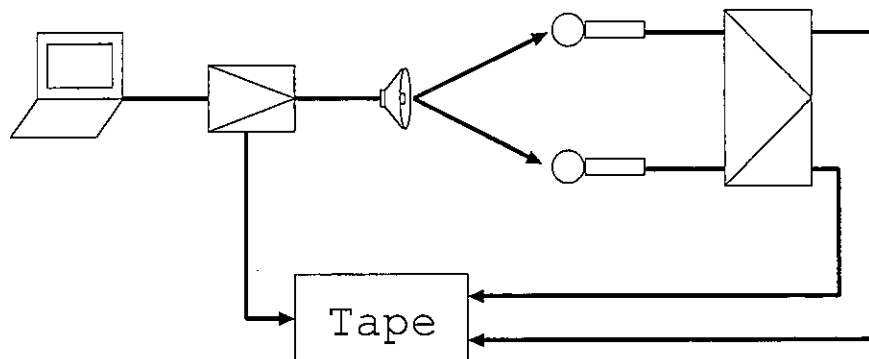


図 3.14: 測定の回路図

信号にはシステム同定用の白色信号と音声信号、音楽を使用した。白色信号はノートパソコンからソフト的に発生させた。音声信号は一旦マイクで録音し、その音声を発生させた。音楽はノートパソコンで CD を再生し行なった。

伝達関数

伝達関数として図 3.13 におけるスピーカから両端のマイクへの経路を使用する。伝達関数の推定には RLS アルゴリズム (忘却係数 1) を使用した。タップ数 4096, 信号に白色信号を用いてシステム同定を行なった。システム同定の残留誤差を図 3.15 に示す。

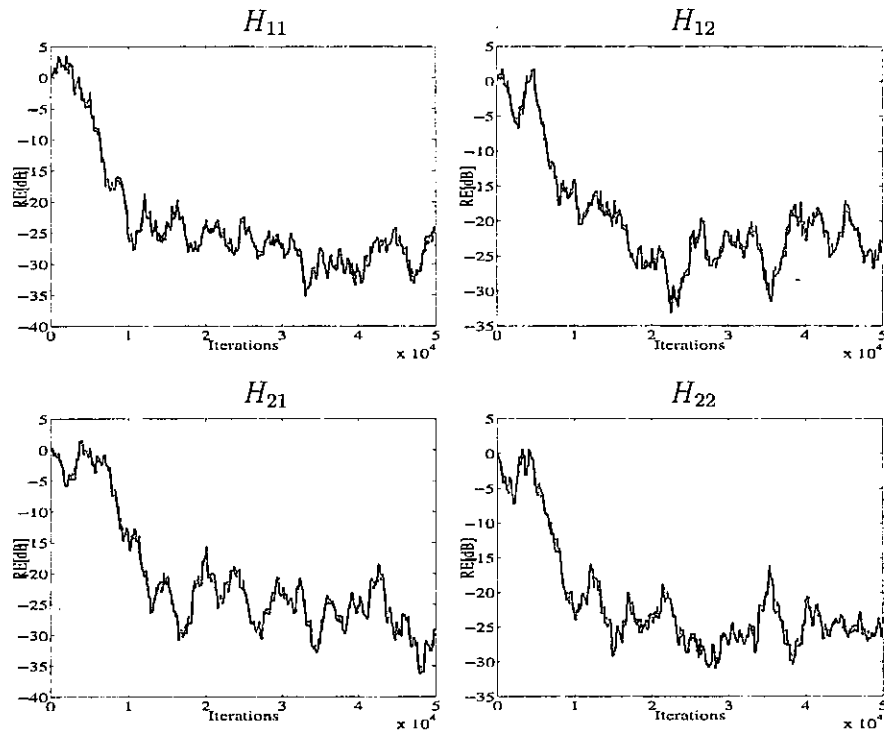


図 3.15: 残留誤差

この図よりどの経路も -25dB から -30dB の誤差に抑えていることが分かる。

シミュレーションには伝達関数は雑音や誤差を考慮して前半 2048 タップを使用する。シミュレーションに使用する伝達関数を図 3.16 に示す。なお、シミュレーションの際には時間変動がないものとして図 3.16 の値で固定して使用する。

また、分離回路の理想インパルス応答は $H_{11}(z), H_{22}$ の根が単位円の外にあるため図示することができない。

パラメータ設定

分離回路のタップ数は 8192 タップとする。信号源は白色信号とし、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 $\mu_0 = 0.0001$, $\gamma_{12} = 0.9993$, $\gamma_{21} = 0.9993$, 適応形指数ステップサイズでは $\mu_0 = 0.0001$, を使ってシミュレーションを行なった。

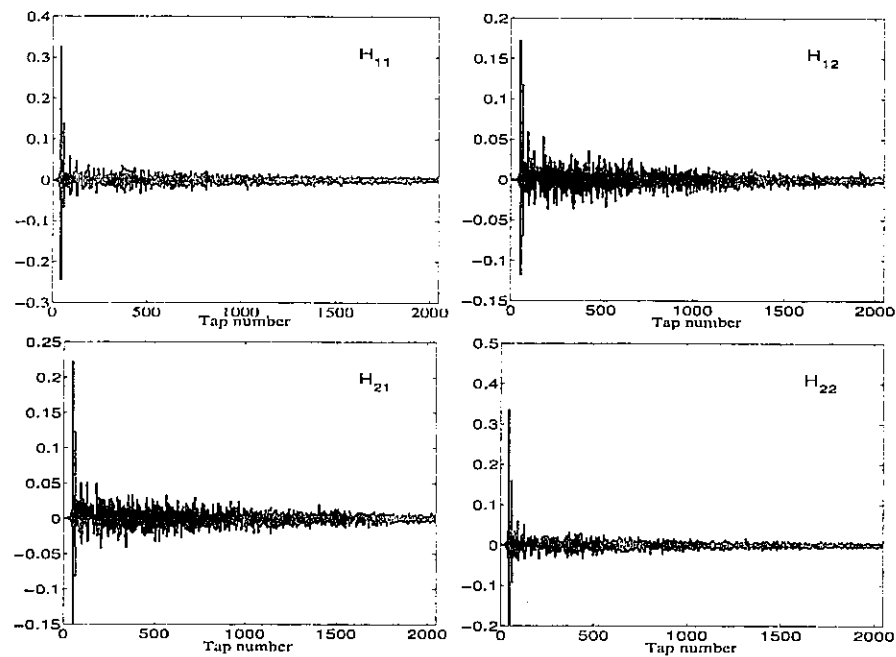


図 3.16: 伝達関数

シミュレーション結果

分離特性を図 3.17 に示す。

固定形指数ステップサイズ方式，適応形指数ステップサイズ方式の方が一定値ステップサイズ方式よりも分離性能が良いことが分かる。

図 3.18 に適応形指数ステップサイズ方式におけるステップサイズの指数の底の変化を示す。

また図 3.19 に学習終了時の分離回路のフィルタ係数を示す。

これまでのシミュレーション同様一定値ステップサイズ方式ではタップの後半まで値が大きい固定形指数ステップサイズ方式，適応形指数ステップサイズ方式では値が小さくなっている。

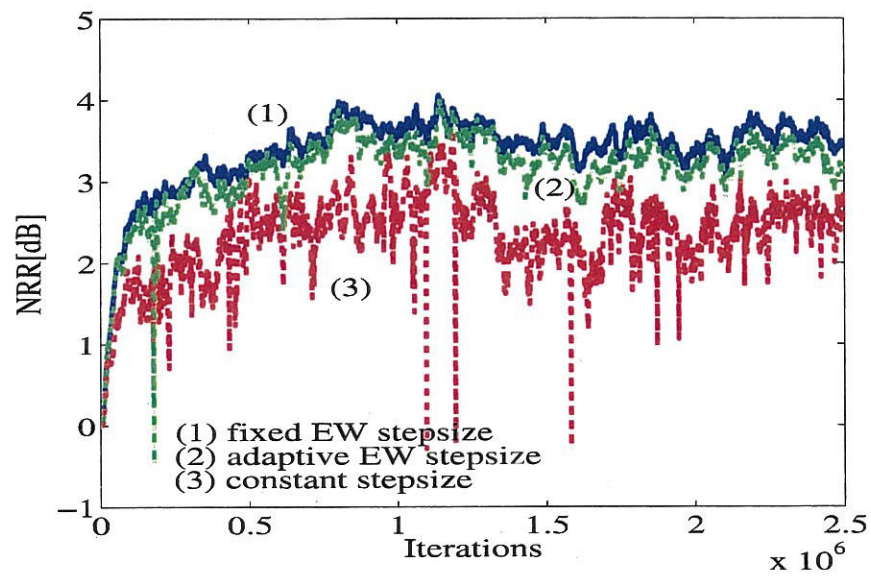


図 3.17: 分離特性

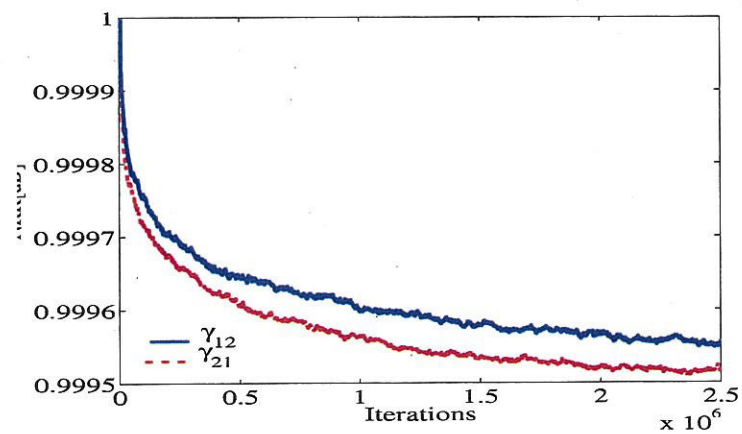
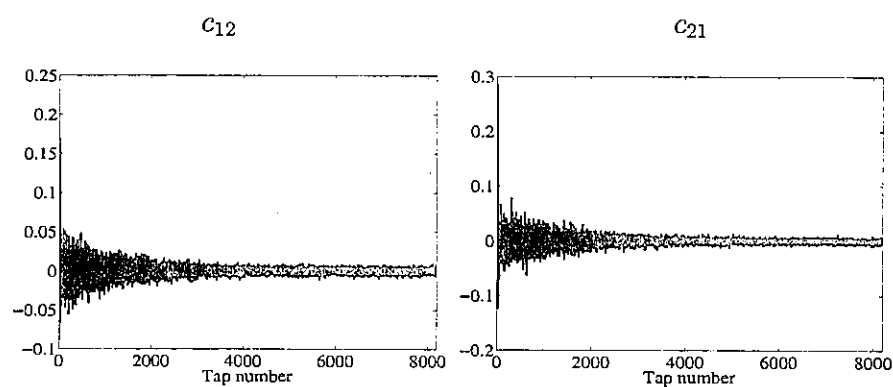
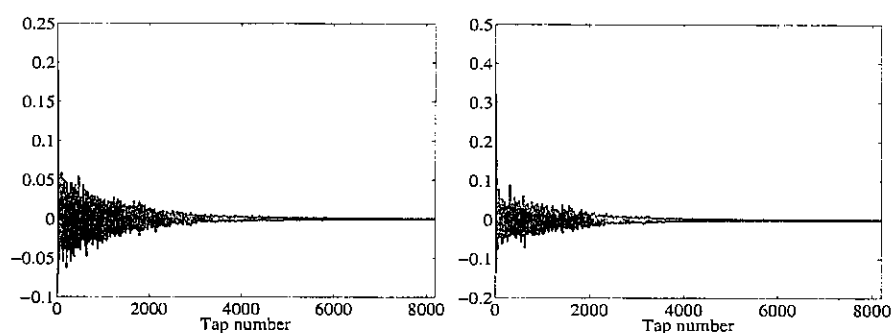


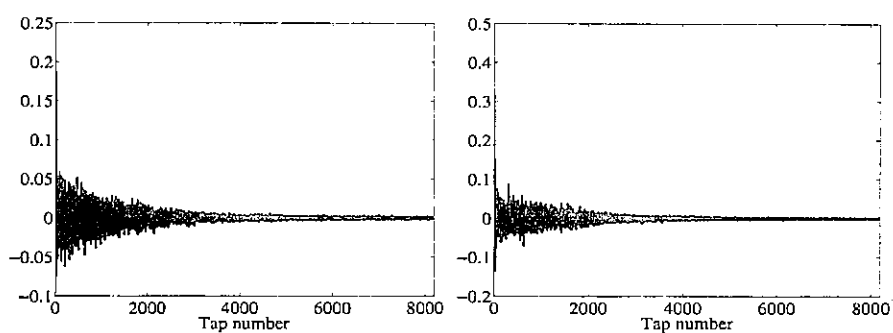
図 3.18: 適応形指数ステップサイズ方式におけるステップサイズの指数の底



(a) 一定値ステップサイズ



(b) 固定形指数ステップサイズ



(c) 適応形指数ステップサイズ

図 3.19: 学習終了時の分離回路のフィルタ係数

第4章 フィードフォワード形BSS (FF-BSS)

畳み込み形のブラインドソースセパレーションでは前章で説明したFB-BSSの他にフィードフォワードの分離回路を使用するフィードフォワード形ブラインドソースセパレーション(FF-BSS)も提案されている。本章ではFF-BSSの回路構成や学習アルゴリズムを説明する。

4.1 回路構成と入出力関係

4.1.1 回路構成

FF-BSSの回路構成を図4.1に示す。

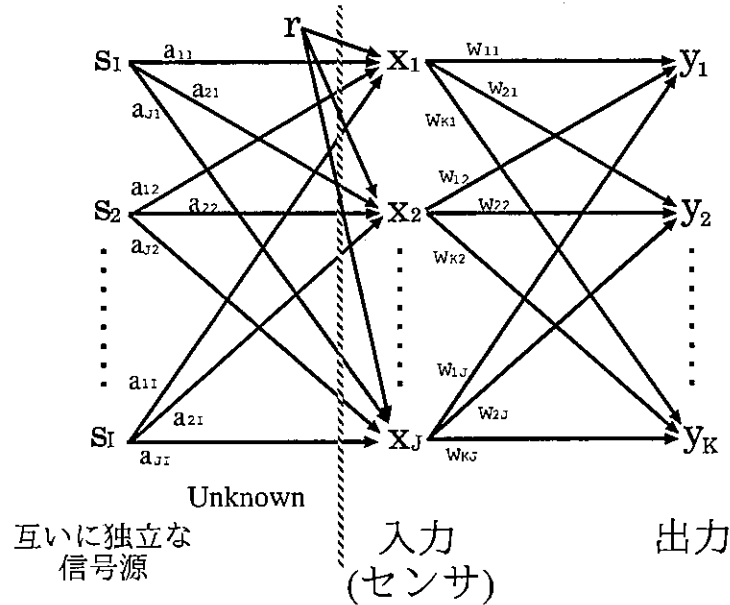


図 4.1: FF-BSS の構成

r はノイズ成分である。 I 個の信号源 $s_i(n)$ はそれぞれ混合過程において a_{ji} の影響を受け J 個のセンサで観測される。その際観測信号には観測雑音 r が含まれる。観測信号 $x_j(n)$ は K 個の独立成分 $y_k(n)$ に分離される。

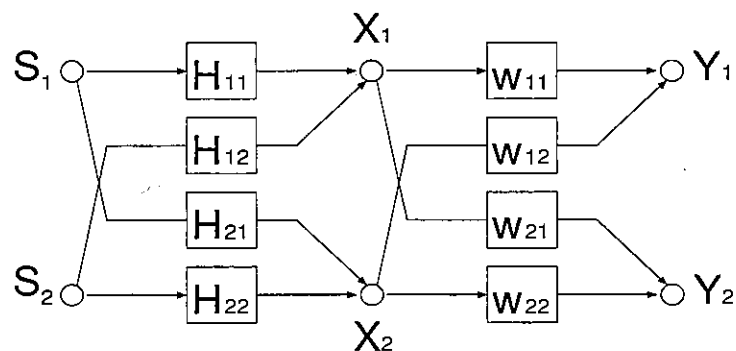
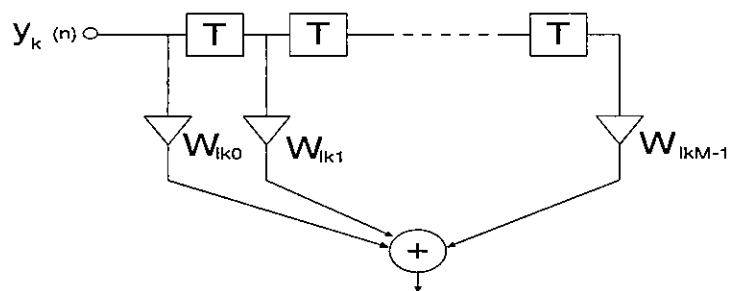


図 4.2: FF-BSS の回路構成

図 4.3: $W_{lk}(z)$ に用いる FIR フィルタ

本研究では簡単化のため信号源，センサと出力の数はすべて同じであるとする．また，観測信号にノイズが含まれないものとする．本研究で使用する FB-BSS のモデルにおいて 2 チャンネルの場合のブロック図を図 4.2 に示す．ここで，本研究では畳み込み形 BSS を扱うため，混合過程は畳み込みの形をしていると仮定する．分離回路には図 4.3 に示す FIR フィルタを用いる．

4.1.2 時間領域

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.

$$x_j(n) = \sum_{i=1}^N \sum_{m=0}^{M_{ji}-1} h_{ji}(m) s_i(n-m) \quad (4.1)$$

また, 分離過程の出力 $y_j(n)$ は次式で表される.

$$y_j(n) = \sum_{k=1}^N \sum_{l=0}^{L_{jk}-1} w_{jk}(l) x_k(n-l) \quad (4.2)$$

この関係をベクトルと行列で表すと次のようになる.

$$\mathbf{x}(n) = \mathbf{H}^T \mathbf{s}(n) \quad (4.3)$$

$$\mathbf{y}(n) = \mathbf{W}^T \mathbf{x}(n) \quad (4.4)$$

ただし, $\mathbf{s}(n), \mathbf{x}(n), \mathbf{y}(n), \mathbf{H}, \mathbf{W}$ は

$$\mathbf{s}(n) = [s_1^T(n), s_2^T(n), \dots, s_N^T(n)]^T \quad (4.5)$$

$$\mathbf{s}_i(n) = [s_i(n), s_i(n-1), \dots, s_i(n-M_i+1)]^T \quad (4.6)$$

$$\mathbf{x}(n) = [x_1(n), x_2(n), \dots, x_N(n)]^T \quad (4.7)$$

$$\mathbf{x}_i(n) = [x_i(n), x_i(n-1), \dots, x_i(n-L_i+1)]^T \quad (4.8)$$

$$\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_N(n)]^T \quad (4.9)$$

$$\mathbf{H} = \begin{bmatrix} h_{11} & h_{21} & \dots & h_{N1} \\ h_{12} & h_{22} & \dots & h_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ h_{1N} & h_{2N} & \dots & h_{NN} \end{bmatrix} \quad (4.10)$$

$$h_{ji} = [h_{ji}(0), h_{ji}(1), \dots, h_{ji}(M_{ji}-1)]^T \quad (4.11)$$

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{21} & \dots & w_{N1} \\ w_{12} & w_{22} & \dots & w_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1N} & w_{2N} & \dots & w_{NN} \end{bmatrix} \quad (4.12)$$

$$w_{jk} = [w_{jk}(0), w_{jk}(1), \dots, w_{jk}(L_{jk}-1)]^T \quad (4.13)$$

である.

4.1.3 z 領域

$s_i(n), x_j(n), y_k(n)$ 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表せる.

$$\mathbf{X}(z) = \mathbf{H}(z) \mathbf{S}(z) \quad (4.14)$$

$$\mathbf{Y}(z) = \mathbf{W}(z) \mathbf{X}(z) \quad (4.15)$$

ここで $S(z), X(z), Y(z), H(z), W(z)$, はそれぞれ,

$$S(z) = [S_1(z), S_2(z), \dots, S_N(z)]^T \quad (4.16)$$

$$X(z) = [X_1(z), X_2(z), \dots, X_N(z)]^T \quad (4.17)$$

$$Y(z) = [Y_1(z), Y_2(z), \dots, Y_N(z)]^T \quad (4.18)$$

$$H(z) = \begin{bmatrix} H_{11}(z) & H_{12}(z) & \dots & H_{1N}(z) \\ H_{21}(z) & H_{22}(z) & \dots & H_{2N}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}(z) & H_{N2}(z) & \dots & H_{NN}(z) \end{bmatrix} \quad (4.19)$$

$$W(z) = \begin{bmatrix} W_{11}(z) & W_{12}(z) & \dots & W_{1N}(z) \\ W_{21}(z) & W_{22}(z) & \dots & W_{2N}(z) \\ \vdots & \vdots & \ddots & \vdots \\ W_{N1}(z) & W_{N2}(z) & \dots & W_{NN}(z) \end{bmatrix} \quad (4.20)$$

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.

$$\begin{aligned} Y(z) &= W(z)X(z) \\ &= W(z)H(z)S(z) \end{aligned} \quad (4.21)$$

分離性能を評価するために以下の行列を定義する.

$$P(z) = W(z)H(z) \quad (4.22)$$

もし $P(z)$ 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. 2 チャンネルの場合をここに示す.

$$P(z) = \begin{bmatrix} P_{11}(z) & 0 \\ 0 & P_{22}(z) \end{bmatrix} \text{ or } \begin{bmatrix} 0 & 0P_{12}(z) \\ P_{21}(z) & 0 \end{bmatrix} \quad (4.23)$$

$P_{ki}(z)$ のインパルス応答 $p_{ki}(n)$ が

$$p_{ki}(n) = 0 \quad n \neq n_0 \quad (4.24)$$

であったとすれば分離した信号は

$$y_k(n) = p_{ki}(n_0)s(n - n_0) \quad (4.25)$$

となる. これは出力が $s_i(n)$ を時間的にシフトし, 定数倍されただけのものであり, 分離信号において歪みが生じないことを表している. しかし, $H(z)$ 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.

$$Y_j(z) = P_{ji}(z)S_i(z) \quad (4.26)$$

4.2 学習アルゴリズム

4.2.1 評価関数

学習には [36] で紹介されている自然勾配法を使用する。信号は定常であるとする。 $\mathbf{Y}(z)$ の同時確率密度関数を次のように置く。

$$p(\mathbf{Y}(z)) = p(Y_1(z), \dots, Y_m(z)) \quad (4.27)$$

分離行列 $\mathbf{W}(z)$ によって観測信号 $\mathbf{X}(z)$ が正しく分離できたならば、 $\mathbf{Y}(z)$ の各成分 $Y_i(z)$ は独立となる。このとき、 $p(Y_i(z))$ を $Y_i(z)$ についての周辺確率密度関数とすれば、 $p(\mathbf{Y}(z))$ はこの周辺分布密度関数の積として、次式のように表される。

$$p(\mathbf{Y}(z)) = \prod_{i=1}^m p(Y_i(z)) \quad (4.28)$$

そこで、式 (4.28) が成り立つように評価関数として $p(\mathbf{Y}(z))$ と $\prod_{i=1}^m p(Y_i(z))$ との間の *Kullback – Leibler divergence* を設定し、これを最小化するように $\mathbf{W}(z)$ を求める。*Kullback – Leibler divergence* の定義は次の通りである。

$$KL(\mathbf{W}(z)) = \int p(\mathbf{Y}(z)) \log \frac{p(\mathbf{Y}(z))}{\prod_{i=1}^m p(Y_i(z))} d\mathbf{Y}(z) \quad (4.29)$$

$$= -H(\mathbf{Y}(z); \mathbf{W}(z)) + \sum_{i=1}^m H(Y_i(z); \mathbf{W}(z)) \quad (4.30)$$

$H(\mathbf{Y}(z); \mathbf{W}(z))$ は同時分布のエントロピーであり、 $H(Y_i(z); \mathbf{W}(z))$ は周辺分布のエントロピーである。ここで、 $p(\mathbf{Y}(z))d\mathbf{Y}(z) = p(\mathbf{X}(z))d\mathbf{X}(z)$ 、 $p(\mathbf{Y}(z)) = \frac{p(\mathbf{X}(z))}{\det \mathbf{W}(z)}$ より、 $H(\mathbf{Y}(z); \mathbf{W}(z))$ と $H(Y_i(z); \mathbf{W}(z))$ は $\mathbf{X}(z)$ と $\mathbf{W}(z)$ を用いて、次式となる。

$$H(\mathbf{Y}(z); \mathbf{W}(z)) = - \int p(\mathbf{Y}(z)) \log p(\mathbf{Y}(z)) d\mathbf{Y}(z) \quad (4.31)$$

$$= - \int p(\mathbf{X}(z)) (\log p(\mathbf{X}(z)) - \log \det \mathbf{W}(z)) d\mathbf{X}(z) \quad (4.32)$$

$$= H(\mathbf{X}(z)) + \log \det \mathbf{W}(z) \quad (4.33)$$

$$H(Y_i(z); \mathbf{W}(z)) = - \int p(\mathbf{Y}(z)) \log p(Y_i(z)) d\mathbf{Y}(z) \quad (4.34)$$

$$= - \int p(\mathbf{X}(z)) \log p(Y_i(z)) d\mathbf{X}(z) \quad (4.35)$$

$$= - \log p(Y_i(z)) \quad (4.36)$$

4.2.2 $W_{jk}(z)$ の時間領域での更新式

$Y_i(z)$ が独立である、すなわち信号源が分離されたとき、 $KL(\mathbf{W}(z))$ は最小である 0 となるため、最急降下法により $\mathbf{W}(z)$ の更新式が求まる。

微分のチェーンルールにより

$$\frac{\partial KL(\mathbf{W}(z))}{\partial w(l)} = \frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{W}(z)} \frac{\partial \mathbf{W}(z)}{\partial w(l)} \quad (4.37)$$

である。ここで $\mathbf{w}(l)$ は z^{-l} の項の係数を集めた行列、すなわち

$$\mathbf{W}(z) = \sum_l \mathbf{w}(l) z^{-l} \quad (4.38)$$

となる行列である。

$$\frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{W}(z)} = -((\mathbf{W}(z)^T)^{-1} - E_{\mathbf{x}}[\varphi(\mathbf{Y}(z))\mathbf{X}(z)^T]) \quad (4.39)$$

$$= -(\mathbf{I} - E_{\mathbf{y}}[\varphi(\mathbf{Y}(z))\mathbf{Y}(z)^T])(\mathbf{W}(z)^T)^{-1} \quad (4.40)$$

$$\varphi(\mathbf{Y}(z)) = -\left(\frac{\partial \log p(Y_1(z))}{\partial Y_1(z)} \dots \frac{\partial \log p(Y_m(z))}{\partial Y_m(z)}\right) \quad (4.41)$$

$$\frac{\partial \mathbf{W}(z)}{\partial \mathbf{w}(l)} = z^{-l} \quad (4.42)$$

実用的には $\varphi(y_j)$ は適当な関数で近似される。

$$\frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{w}(l)} = -(\mathbf{I} - E_{\mathbf{y}}[\varphi(\mathbf{Y}(z))\mathbf{Y}(z)^T])(\mathbf{W}(z)^T)^{-1} z^{-l} \quad (4.43)$$

$$F(\mathbf{Y}(z)) = \frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{w}(l)} \mathbf{W}(z)^T \mathbf{W}(z) \quad (4.44)$$

$$= -(\mathbf{I} - E_{\mathbf{y}}[\varphi(\mathbf{Y}(z))\mathbf{Y}(z)^T]) \mathbf{W}(z) z^{-l} \quad (4.45)$$

ここで

$$E\left[\frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{w}(l)}\right] = 0 \quad (4.46)$$

と

$$E[F(\mathbf{Y}(z))] = 0 \quad (4.47)$$

は等価であるため $F(\mathbf{Y}(z))$ を更新式に使用する。

したがって、 $w_{ij}(l)$ の更新式は次式のようなになる。

$$w_{ij}(n+1, l) = w_{ij}(n, l) + \Delta w_{ij}(n, l) \quad (4.48)$$

$$\Delta w_{ij}(n, l) = \eta \left\{ w_{ij}(n, l) - \sum_{q=0}^{L_{ij}-1} \varphi(y_i(n)) y_p(n-l+q) w_{pj}(n, q) \right\} \quad (p \neq j) \quad (4.49)$$

また、式 (4.38) の $\mathbf{w}(l)$ を用いて更新式を表現すると次式のようなになる。

$$\mathbf{w}(n+1, l) = \mathbf{w}(n, l) + \Delta \mathbf{w}(n, l) \quad (4.50)$$

$$\mathbf{w}(n, l) = \eta \sum_{q=0}^{L_{ij}-1} [\mathbf{I} \delta(n-q) - \langle \Phi(\mathbf{y}(n)) \mathbf{y}^T(n-l+q) \rangle] \mathbf{w}(n, q) \quad (4.51)$$

$$\Phi(\mathbf{y}(n)) = [\Phi(y_1(n)), \dots, \Phi(y_N(n))]^T \quad (4.52)$$

ここで $\delta(n)$ はディラックのデルタ関数であり、 $\delta(0) = 1$ かつ $\delta(n) = 0$ ($n \neq 0$) である。

4.2.3 $W_{jk}(z)$ の周波数領域での更新式

FF-BSSにおいて、周波数領域で学習する方式も提案されている [20], [21], [29], [39], [46]-[52]. 時間領域で学習するものと同様に *Kullback – Leibler divergence* を評価関数とし、勾配法によって最小値を求める.

$$\frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{W}(z)} = -((\mathbf{W}(z)^H)^{-1} - E_{\mathbf{X}}[\Phi(\mathbf{Y}(z))\mathbf{X}(z)^H]) \quad (4.53)$$

$$= -(\mathbf{I} - E_{\mathbf{Y}}[\Phi(\mathbf{Y}(z))\mathbf{Y}(z)^H])(\mathbf{W}(z)^H)^{-1} \quad (4.54)$$

$$\Phi(\mathbf{Y}(z)) = -\left(\frac{\partial \log p(Y_1(z))}{\partial Y_1(z)} \dots \frac{\partial \log p(Y_m(z))}{\partial Y_m(z)}\right) \quad (4.55)$$

実用的には $\Phi(y_j)$ は適当な関数で近似される.

$$F(\mathbf{Y}(z)) = \frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{W}(z)} \mathbf{W}(z)^H \mathbf{W}(z) \quad (4.56)$$

$$= -(\mathbf{I} - E_{\mathbf{Y}}[\Phi(\mathbf{Y}(z))\mathbf{Y}(z)^H]) \mathbf{W}(z) \quad (4.57)$$

ここで

$$E\left[\frac{\partial KL(\mathbf{W}(z))}{\partial \mathbf{W}(z)}\right] = 0 \quad (4.58)$$

と

$$E[F(\mathbf{Y}(z))] = 0 \quad (4.59)$$

は等価であるため時間領域で学習するものと同様に $F(\mathbf{Y}(z))$ を更新式に使用する.

周波数領域での学習のために短時間フーリエ変換を用いる. したがって, 各周波数帯域をフレーム毎に更新することになる. フレームを r , 周波数帯域を m で表すと更新式は

$$\mathbf{W}(r+1, m) = \mathbf{W}(r, m) + \eta[\mathbf{I} - \langle \Phi(\mathbf{Y}(r, m))\mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \quad (4.60)$$

となる. 非線形関数 $\Phi()$ には以下のようなものが使用される.

$$\Phi(\mathbf{Y}(r, m)) = \frac{1}{1 + e^{-\mathbf{Y}^R(r, m)}} + \frac{j}{1 + e^{-\mathbf{Y}^I(r, m)}} \quad (4.61)$$

ここで, $\mathbf{Y}^R(r, m), \mathbf{Y}^I(r, m)$ はそれぞれ $\mathbf{Y}(r, m)$ の実部, 虚部を表す.

ただし, このアルゴリズムでは以下の問題点が指摘されている [29].

式 (4.60) において収束条件は, 左辺の更新量が 0 となることであるので以下の式で表される.

$$\mathbf{I} = \langle \Phi(\mathbf{Y}(r, m))\mathbf{Y}^H(r, m) \rangle \quad (4.62)$$

上式を行列で書き表すと,

$$\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix} = \begin{bmatrix} \langle \Phi(\mathbf{Y}_1)\mathbf{Y}_1^* \rangle & \cdots & \langle \Phi(\mathbf{Y}_1)\mathbf{Y}_N^* \rangle \\ \vdots & \ddots & \vdots \\ \langle \Phi(\mathbf{Y}_2)\mathbf{Y}_1^* \rangle & \cdots & \langle \Phi(\mathbf{Y}_N)\mathbf{Y}_N^* \rangle \end{bmatrix} \quad (4.63)$$

となる．ここで $\mathbf{Y}^*$ は $\mathbf{Y}$ の複素共役を表す．
すなわち，

$$\langle \Phi(\mathbf{Y}_p) \mathbf{Y}_q^* \rangle = 1 \quad (p = q) \quad (4.64)$$

$$\langle \Phi(\mathbf{Y}_p) \mathbf{Y}_q^* \rangle = 0 \quad (p \neq q) \quad (4.65)$$

である．

式 (4.65) は出力を互いに独立にしようと働くのに対して，式 (4.64) は出力の振幅をある値に近付ける制約となっている．

周波数領域に拡張された学習アルゴリズムを用いる場合，学習は各周波数帯域ごとに行なわれる．このことから，出力の各周波数帯域の振幅が一定の値に近付くことは，出力信号のスペクトルが平坦化されることになり，信号が白色化されるという問題がある [39]．

この出力白色化の制約を回避するために式 (4.60) 中の単位行列 $\mathbf{I}$ を対角行列 $\text{diag}(\langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle)$ に置き換えた，以下のようなアルゴリズムがすでに提案されている [29]．

$$\begin{aligned} \mathbf{W}(r+1, m) = & \mathbf{W}(r, m) + \eta [\text{diag}(\langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle) \\ & - \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \end{aligned} \quad (4.66)$$

本稿では，式 (4.60)，(4.66) の学習法をそれぞれ FF-FREQ(1)，FF-FREQ(2) と表記する．

第5章 信号分離と信号歪み

BSS では出力信号を独立な成分に分離する．その際に信号源と出力信号の関係は考慮されない．すなわち出力信号は信号源と同一のものであるという保証はない．つまり BSS による信号歪みが生じる可能性がある．

本章ではまず，信号の歪みの基準を定める．つづいて，FF-BSS と FB-BSS の信号分離と信号歪みの関係について考察する．そして最後にシミュレーションにより両者を比較し，FF-BSS と FB-BSS の信号分離と信号歪みの関係についての考察を検討する．

5.1 信号歪みの基準

5.1.1 基準として適当である信号

まず信号歪みを定義するために歪みの基準となる信号を定義する．信号歪みの基準としては以下の2つが考えられる．

- 信号源
- 観測信号の $S_i(z)$ 成分

信号源を歪みの基準として考える場合，BSS に対して信号の等化までを要求する評価となっている．一方，観測信号の $S_i(z)$ 成分を歪みの基準として考える場合，BSS がどれだけ信号を歪ませるかを評価することになる．これらのそれぞれ D_1, D_2 とすると

$$D_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.1)$$

$$D_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega})S_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.2)$$

で表される．

本章ではこれらの評価方法を用いてフィードフォワード形 BSS やフィードバック形 BSS がどのような振舞をしているか解析する．

5.1.2 本研究で扱う信号歪みの基準

実用的には信号源がそのまま出力される，すなわち出力信号 $Y_i(z)$ が信号源 $S_i(z)$ に一致することが望ましい．しかし BSS は，観測信号を独立な成分に分離するだけであり，混合過程の伝達関数 $H_{ij}(z)$ は推定できないため，本研究では $H_{ii}(z)S_i(z)$ を信号歪みの基準とする [10],[18],[23],[38],[39],[41],[43]-[46]．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。この場合、以下の評価方法が考えられる。

$$\sigma_{d1a}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})S_i(e^{j\omega}) - P_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (5.3)$$

$$\sigma_{d1b}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|H_{ji}(e^{j\omega})S_i(e^{j\omega})| - |P_{ki}(e^{j\omega})S_i(e^{j\omega})|)^2 d\omega \quad (5.4)$$

$$\sigma_1^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (5.5)$$

$$SD_{1x} = 10 \log_{10} \frac{\sigma_{d1x}^2}{\sigma_1^2}, x = a, b \quad (5.6)$$

$$\sigma_{d2a}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega}) - P_{ki}(e^{j\omega})|^2 d\omega \quad (5.7)$$

$$\sigma_{d2b}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|H_{ji}(e^{j\omega})| - |P_{ki}(e^{j\omega})|)^2 d\omega \quad (5.8)$$

$$\sigma_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})|^2 d\omega \quad (5.9)$$

$$SD_{2x} = 10 \log_{10} \frac{\sigma_{d2x}^2}{\sigma_2^2}, x = a, b \quad (5.10)$$

ただし、BSS では信号の大きさを考慮しないため評価には以下の規格化を行なう。

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ki}(e^{j\omega})S_i(e^{j\omega})|^2 d\omega \quad (5.11)$$

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ji}(e^{j\omega})|^2 d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ki}(e^{j\omega})|^2 d\omega \quad (5.12)$$

SD_{1x} は信号を基準に評価を行なう。 SD_{1a} は出力信号 $Y_k(e^{j\omega})$ と観測信号における信号源の成分 $H_{ji}(e^{j\omega})S_i(e^{j\omega})$ を厳密に比較しており、上記の評価方法の中で一番厳密であるといえる。ただし、時間遅れなどの要素が加わった場合、この評価方法ではうまく評価できない。そこで SD_{1b} は信号の大きさのみで評価する。ただしこの評価方法では位相の変化については考慮しておらず位相の変化には対応できない。

また、信号を含める評価は信号の性質に依存してしまう。そこで式 (4.22), (2.24) で定義される分離行列の要素を用いて伝達関数でも評価する。ただしこちらの評価方法に関しても時間遅れなどを考慮する必要があり伝達関数をそのまま評価する SD_{2a} と大きさのみで比較する SD_{2b} の2通りで評価する。

次章以降これらの評価方法を用いて信号歪みを評価する。

5.2 FF-BSS

5.2.1 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 (4.21) より完全分離条件は以下の様になる．

$$\begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} = \begin{bmatrix} W_{11}(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) \end{bmatrix} \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \quad (5.13)$$

$$= \begin{bmatrix} P_{11}(z) & 0 \\ 0 & P_{22}(z) \end{bmatrix} \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \quad (5.14)$$

すなわち，信号源から出力までの伝達行列が $\mathbf{P}(z)$ が

$$\mathbf{P}(z) = \begin{bmatrix} P_{11}(z) & 0 \\ 0 & P_{22}(z) \end{bmatrix} = \begin{bmatrix} W_{11}(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) \end{bmatrix} \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \quad (5.15)$$

$$= \begin{bmatrix} W_{11}(z)H_{11}(z) + W_{21}(z)H_{12}(z) & W_{12}(z)H_{11}(z) + W_{22}(z)H_{12}(z) \\ W_{11}(z)H_{21}(z) + W_{21}(z)H_{22}(z) & W_{12}(z)H_{21}(z) + W_{22}(z)H_{22}(z) \end{bmatrix} \quad (5.16)$$

となる．

いま， $P_{11}(z), P_{22}(z)$ は任意の値であるので分離条件は

$$W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) = 0 \quad (5.17)$$

$$W_{22}(z)H_{21}(z) + W_{21}(z)H_{11}(z) = 0 \quad (5.18)$$

となる．

$W_{ij}(z)$ 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 (5.17), (5.18) の 0 にすべき独立した項はそれぞれ $z^0, z^{-1}, \dots, z^{-(L+M-1)}$ であり，その数は $L+M-1$ 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．一般的には後者が成り立つので FF-BSS では自由度が存在することになる．

同様に N チャネルにおいても

$$\text{diag}[\mathbf{P}(z)] = \mathbf{W}(z)\mathbf{H} \quad (5.19)$$

のとき完全に分離が行われたことになり， $\mathbf{P}(z)$ の対角要素は任意の値で構わないため，分離条件として $N^2 - N$ 本の方程式が成り立つ．また $W_{ij}(z)$ の数は N^2 個ある．2 チャネルのときと同様に $W_{ij}(z)$ 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 1 本の方程式に対し 0 にすべき独立した項の数は 2 チャネルの時と同様に， $L+M-1$ である．それに対し，1 本の方程式に対する未知数，すなわち $W_{ij}(z)$ の項数は NL である．このことからチャネル数が増えるにつれて不定解に陥りやすく，より多くの自由度が存在することが分かる．

5.2.2 信号歪み

式(4.48),(4.49)の学習は相互情報量を最小化しており，出力信号が互いに独立になるように行なわれる．信号の独立性は

- 信号分離
- 信号スペクトルの変化

によって変化する．信号源は独立であると仮定しており，信号を分離することによって信号が互いに独立になるため評価関数を相互情報量としているが，信号スペクトルを互いに離すことによってでも独立性が向上する．FF-BSSでは前述の通り，自由度が存在するため信号を分離するとともに信号スペクトルを変化させ，独立性をより高めることが可能である．したがって，FF-BSSの出力は互いに独立になるように信号スペクトルを変化させてしまい，信号歪みが生じてしまう．白色信号はどんな信号に対しても独立であるため信号が白色化されてしまう場合が多い．信号源が白色信号の場合，混合過程で有色化され，独立性を高めるために白色化され，一見，信号源に戻っているように見えるが，信号源との出力との相関が低く別の系列の白色信号である．

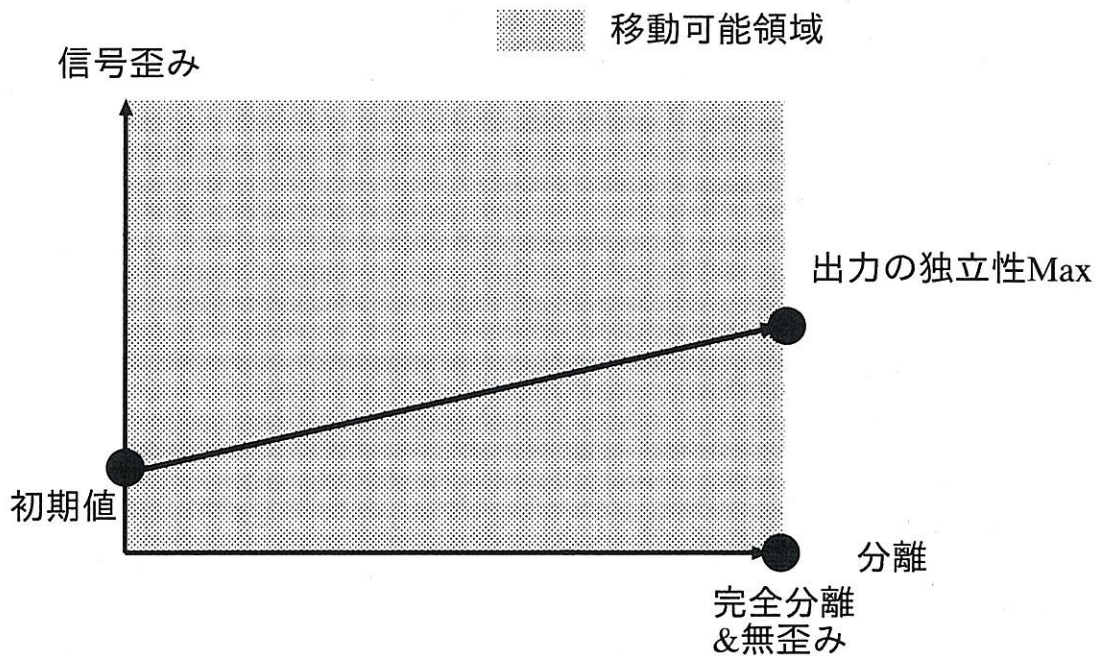


図 5.1: FF-BSS の信号分離と信号歪みの関係

図 5.1 に FF-BSS における信号分離と信号歪みの関係を示している．横軸は分離性能，縦軸は信号歪みの度合を表している．

FF-BSS は自由度があるため信号を分離しながらスペクトルを変化させることができる．すなわち，分離性能と信号歪みの度合で張られる空間の全てを出力は移動することができる．移動可能領域内で出力信号の独立性が最大の点は完全分離かつ無歪みの点から離れ，信号歪みの度合が大きいところに位置する．学習はその独立性が最大となる点に向かって進むため，信号歪みが生じる．

5.2.3 重み付け効果

周波数領域で学習する FF-BSS では、前節までで述べた要因で起こる信号歪みとは別に、重み付け効果と呼ばれる効果によって信号歪みが引き起こされる可能性がある [47], [49].

分離回路 $\mathbf{W}(r, m)$ の学習は式 (4.60), (4.66) の自然勾配法によって行なわれるが、その学習は出力 $\mathbf{Y}(r, m)$ を用いて行なわれる。

式 (4.66) を用いて学習を行なった場合、 $\mathbf{W}$ の更新量は、

$$\begin{aligned} \Delta \mathbf{W}(r, m) = & \eta [\text{diag}(\langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle) \\ & - \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \end{aligned} \quad (5.20)$$

となり、右辺の $\mathbf{W}(r, m)$ の係数を行列で書き直すと以下のようなになる。

$$\eta \begin{bmatrix} 0 & -\langle \Phi(Y_1(r, m)) Y_2^*(r, m) \rangle \\ -\langle \Phi(Y_2(r, m)) Y_1^*(r, m) \rangle & 0 \end{bmatrix} \quad (5.21)$$

上式より、分離回路 $\mathbf{W}(r, m)$ の更新量は各出力の積 $Y_i(r, m) Y_j(r, m)$ ($i \neq j$) に比例するといえる。

このことが分離回路の学習にどのような影響を与えるかを、図 5.2, 5.3 に示す図で考える。

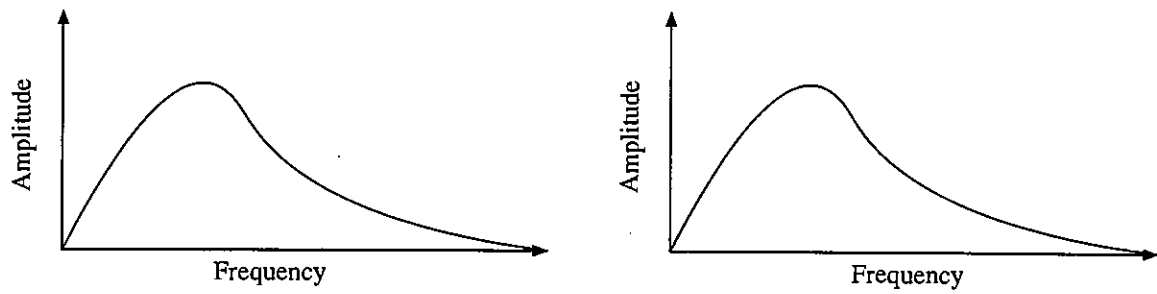
学習初期段階では分離が十分に行なわれていないため、出力 $Y_1(r, m)$, $Y_2(r, m)$ は、どちらも信号源 $S_1(r, m)$, $S_2(r, m)$ 成分の両方が足し合わされたようなものになっている。分離回路の学習は各出力の積に比例するので、分離回路 $\mathbf{W}(r, m)$ は信号源成分の足し合わされたもの、つまり、 $(S_1(r, m) + S_2(r, m))^2$ に近いものに比例する。

また、学習が進んだ段階において、出力 $Y_1(r, m)$, $Y_2(r, m)$ には、それぞれ各信号源成分 $S_1(r, m)$, $S_2(r, m)$ がメインに取り出されているものとする。このとき学習初期段階と同様に考えると、分離回路 $\mathbf{W}(r, m)$ は信号源成分の積 $S_1(r, m) \times S_2(r, m)$ に比例すると言える。これは $S_1(r, m)$, $S_2(r, m)$, $S_1(r, m) \times S_2(r, m)$ の特性によって、帯域毎に学習の偏りが生じることを示している。このことを「重み付け効果」と呼ぶことにする。

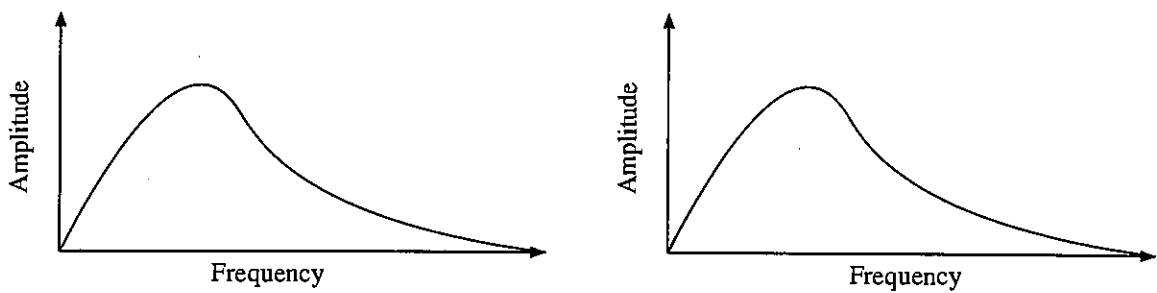
図 5.2(a) で示すように、信号源 $S_1(r)$, $S_2(r)$ のスペクトルが互いに似ている場合について考える。ともに似た周波数特性を持つ信号である場合、 $(S_1(r) + S_2(r))^2$, $S_1(r) \times S_2(r)$ のどちらも、図 5.2(b) に示すような信号源 $S_1(r)$, $S_2(r)$ とスペクトルを持つようになる。これは、主に信号が分布している低域を強調し、あまり信号が分布していない高域を抑える形になる。この場合、分離回路 $\mathbf{W}(r)$ の学習は、低域の更新量が大きく、高域の更新量が小さくなるといった偏りが生じる。これは図 5.2(c) のように出力信号が信号源の形に似てくることになり、特に問題はない。逆に収束を早める効果が期待できる。

しかし、図 5.3 のように信号源の周波数特性が異なる場合においては問題が発生する。図 5.3(a) の信号源に対して、 $(S_1(r) + S_2(r))^2$, $S_1(r) \times S_2(r)$ はどちらも図 5.3(b) に示すような低域と高域の両方を強調した形になっている。この場合、分離回路の学習が $(S_1(r) + S_2(r))^2$, $S_1(r) \times S_2(r)$ に比例して行なわれるということは、分離回路がキャンセルすべき周波数成分も含んだものに比例してしまうことになる。よって、学習によって図 5.3(c) に示すような出力信号に歪みを生じる可能性がある。

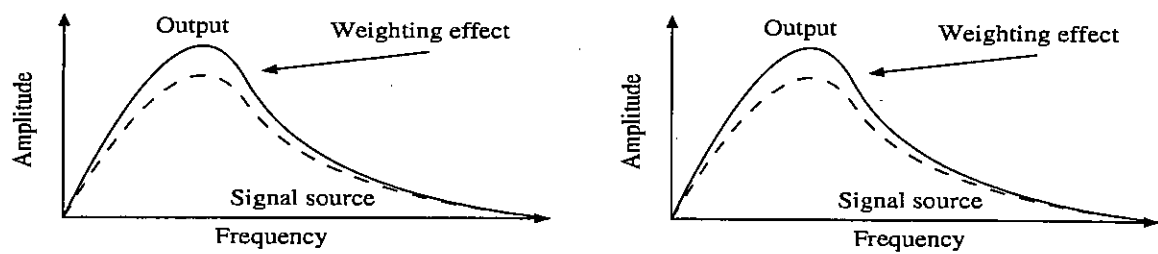
つまり、信号源が互いに特性の似た信号でない場合、学習の際に起こる重み付け効果が信号歪みの原因となることが考えられる。



(a) 同じ帯域に成分が分布している信号源のスペクトル

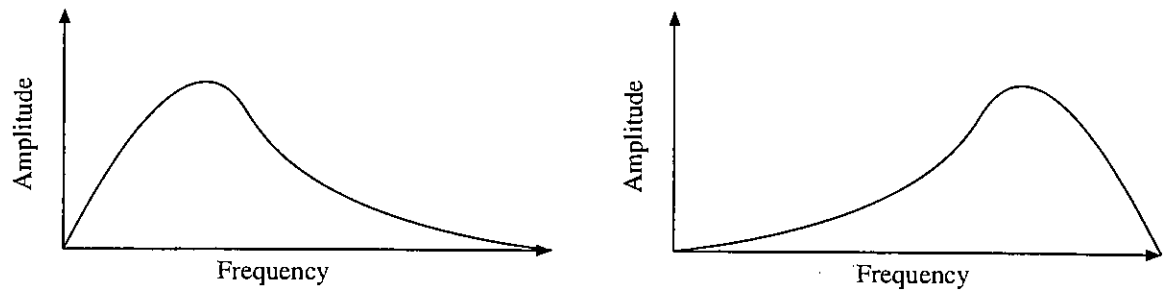


(b) 分離回路における応答

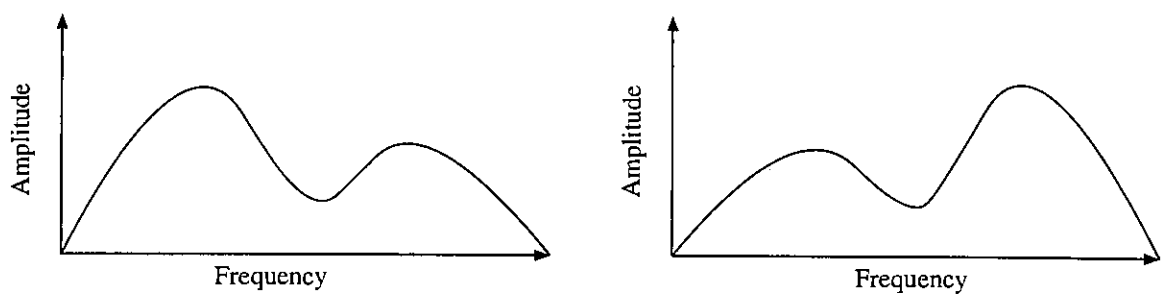


(c) 出力のスペクトル

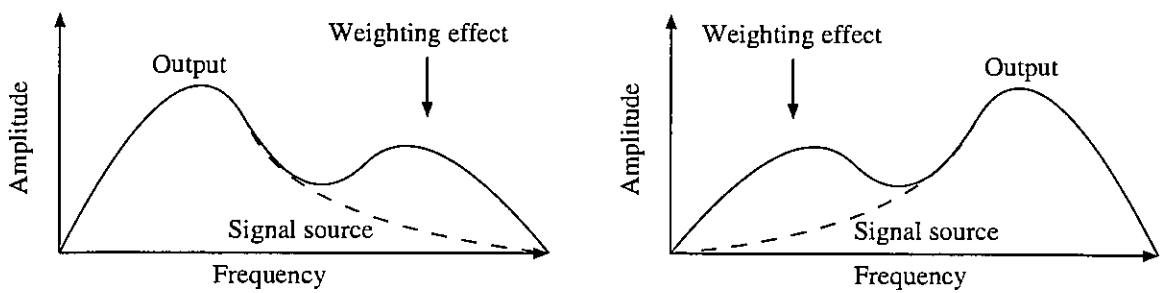
図 5.2: 周波数領域で学習する FF-BSS の重み付け効果 (信号源スペクトルが同じ帯域に分布している場合)



(a) 違う帯域に成分が分布している信号源のスペクトル



(b) 分離回路における応答



(c) 重み付け効果によって影響を受けた出力のスペクトル

図 5.3: 周波数領域で学習する FF-BSS の重み付け効果 (信号源スペクトルが違う帯域に分布している場合)

5.3 FB-BSS

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 (2.22) より完全分離条件は以下の様になる．

$$\begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} = \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & C_{12}(z) \\ C_{21}(z) & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \quad (5.22)$$

$$= \begin{bmatrix} P_{11}(z) & 0 \\ 0 & P_{22}(z) \end{bmatrix} \begin{bmatrix} S_1(z) \\ S_2(z) \end{bmatrix} \quad (5.23)$$

すなわち，信号源から出力までの伝達行列が $\mathbf{P}(z)$ が

$$\mathbf{P}(z) = \begin{bmatrix} P_{11}(z) & 0 \\ 0 & P_{22}(z) \end{bmatrix} \quad (5.24)$$

$$= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & C_{12}(z) \\ C_{21}(z) & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \quad (5.25)$$

$$= \frac{1}{1 - C_{12}(z)C_{21}(z)} \begin{bmatrix} 1 & -C_{12}(z) \\ -C_{21}(z) & 1 \end{bmatrix} \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \quad (5.26)$$

$$= \frac{1}{1 - C_{12}(z)C_{21}(z)} \begin{bmatrix} H_{11}(z) - C_{12}(z)H_{21}(z) & H_{12}(z) - C_{12}(z)H_{22}(z) \\ H_{21}(z) - C_{21}(z)H_{11}(z) & H_{22}(z) - C_{21}(z)H_{12}(z) \end{bmatrix} \quad (5.27)$$

となる．

いま， $P_{11}(z), P_{22}(z)$ は任意の値であるので分離に関する最適解は式 (5.24)-(5.27) の非対角要素より

$$C_{21}(z) = \frac{H_{21}(z)}{H_{11}(z)} \quad C_{12}(z) = \frac{H_{12}(z)}{H_{22}(z)} \quad (5.28)$$

と一意に決定する．

いま分離回路が

$$C_{ij}(z) = \frac{\hat{H}_{ij}(z)}{\hat{H}_{jj}(z)} \quad (5.29)$$

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 (5.24)-(5.27) の対角要素より

$$P_{ii}(z) = \frac{H_{ii}(z) - C_{ij}(z)H_{ji}(z)}{1 - C_{12}(z)C_{21}(z)} \quad (5.30)$$

$$\begin{aligned} &= \frac{\hat{H}_{ii}(z)\hat{H}_{jj}(z)}{\hat{H}_{ii}(z)\hat{H}_{jj}(z) - \hat{H}_{ij}(z)\hat{H}_{ji}(z)} \\ &\times \frac{H_{ii}(z)\hat{H}_{jj}(z) - \hat{H}_{ij}(z)H_{ji}(z)}{\hat{H}_{jj}(z)} \end{aligned} \quad (5.31)$$

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この式は

$$P_{ii}(z) = H_{ii}(z) \quad (5.32)$$

となる。前述の通り信号歪みを $H_{ii}(z)S_i(z)$ からどれだけずれているかで判断するため、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。

同様に N チャネルにおいても完全分離ができるとき

$$\text{diag}[P(z)] = (I - C(z))^{-1}H(z) \quad (5.33)$$

が成り立つ。この式は

$$(I - C(z)) \text{diag}[P(z)] = H(z) \quad (5.34)$$

$$I - C(z) = H(z)(\text{diag}[P(z)])^{-1} \quad (5.35)$$

$$C(z) = I - H(z)(\text{diag}[P(z)])^{-1} \quad (5.36)$$

と変形できる。式 (2.21) にあるように $C(z)$ の対角要素は 0 である。したがって、

$$\text{diag}[C(z)] = \text{diag}[I - H(z)(\text{diag}[P(z)])^{-1}] = 0 \quad (5.37)$$

$$\text{diag}[H(z)(\text{diag}[P(z)])^{-1}] = I \quad (5.38)$$

$$\text{diag}[H(z)](\text{diag}[P(z)])^{-1} = I \quad (5.39)$$

$$\text{diag}[P(z)] = \text{diag}[H(z)] \quad (5.40)$$

$$P_{ii}(z) = H_{ii}(z) \quad (5.41)$$

となる。2 チャネルのときと同様に N チャネルにおいても信号が完全に分離されたときの出力は信号歪みの基準と一致する。つまり、FB-BSS では完全分離できれば信号歪みは発生しない。

ちなみに、完全分離をもたらす $C_{jk}(z)$ の値は

$$\text{non_diag}[C(z)] = \text{non_diag}[H(z)(\text{diag}[H(z)])^{-1}] \quad (5.42)$$

$$C_{jk}(z) = \frac{H_{jk}(z)}{H_{kk}(z)} \quad (5.43)$$

である。ここで、 $\text{non_diag}[C(z)]$ は $C(z)$ の非対角要素を表す。

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。

FF-BSS と同様に信号分離と信号歪みの関係を図 5.4 に示す。FB-BSS では信号分離と信号歪みの関係が一意に定まる。そのため出力は分離特性と信号歪みの度合で張られる空間をある曲線上を動くことになる。その曲線は前述の議論の通り、完全分離が行なわれると歪みがなくなるため初期値と完全分離かつ無歪みの点を通る。図 5.4 では便宜上、簡単に初期値と完全分離かつ無歪みの点を直線で結んでいる。独立性はスペ

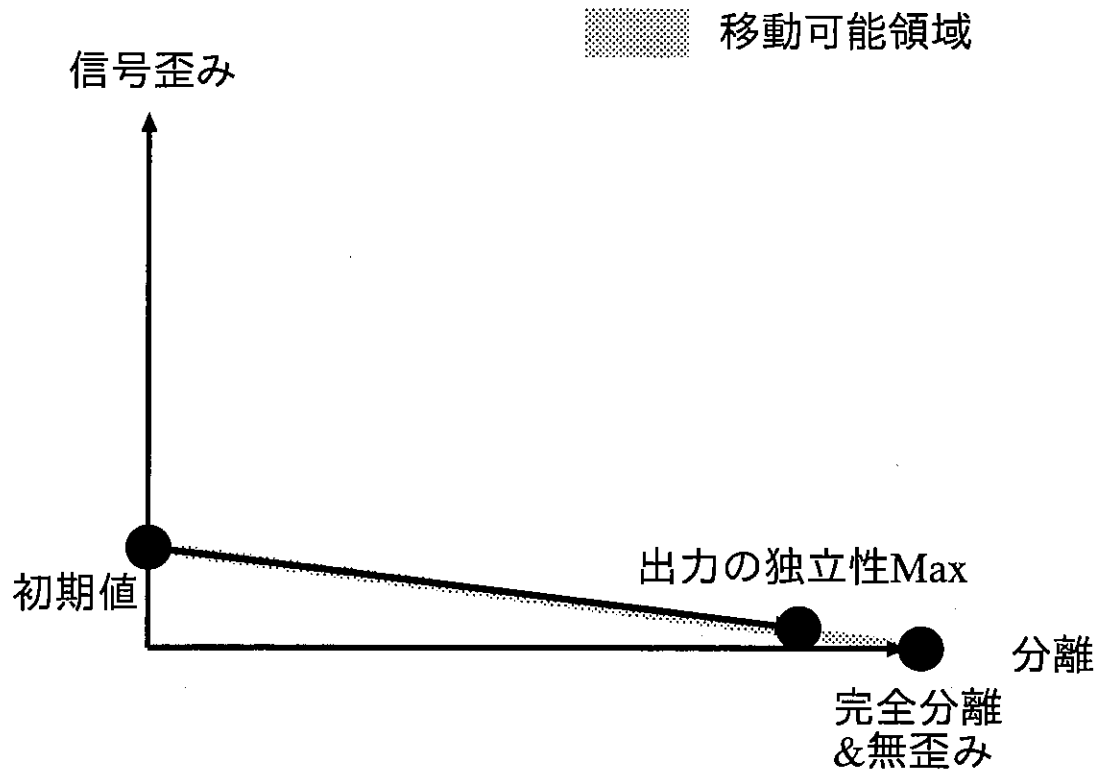


図 5.4: FB-BSS の信号分離と信号歪みの関係

クトルを変化させることによって高まるが、出力が移動可能な領域はスペクトルを少し変化させると大きく完全分離から離れることになり完全分離の近く存在すると考えられる。学習は出力の独立性が大きくなるように学習が行なわれるため、独立性最大の点、すなわち完全分離かつ無歪みの点の近くに収束する。

5.4 シミュレーション

5.4.1 シミュレーション条件

シミュレーションは2信号源, 2センサで行なう. 混合過程のインパルス応答を図 5.5 に示す.

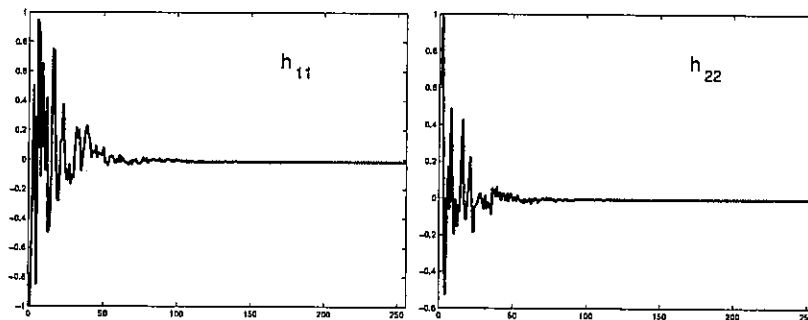


図 5.5: 伝達関数 (インパルス応答)

クロスパス ($H_{12}(z), H_{21}(z)$) と直接パス ($H_{11}(z), H_{22}(z)$) の関係は以下のようにする.
(Mixture-1)

$$H_{21}(z) = 0.9z^{-1}H_{11}(z) \quad (5.44)$$

$$H_{12}(z) = 0.9z^{-1}H_{22}(z) \quad (5.45)$$

(Mixture-2)

$$H_{21}(z) = 0.5z^{-1}H_{11}(z) \quad (5.46)$$

$$H_{12}(z) = 0.5z^{-1}H_{22}(z) \quad (5.47)$$

Mixture-1 は Mixture-2 に比べて難しい問題設定となっている. 信号歪みの解析は Mixture-1 で行ない, FB-BSS と FF-BSS の比較は Mixture-1 と Mixture-2 の両方で比較する.

信号源には (1) 白色信号-白色信号と (2) 音声-音声の 2 種類を用いた. 音声信号は以下の男性と女性の音声を用いている.

- s_1 : 音声 「人間は豊かな自然を求めています。」 (男性)
 「パンはポルトガル人が伝えたものです。」 (男性)
 s_2 : 音声 「健康のため無理をしないでください。」 (女性)
 「太宰治は青森県で生まれました。」 (女性)

タップ数は 256, 分離回路の初期値は FF-BSS では $W_{11}(z) = W_{22}(z) = 1$, $W_{12}(z) = W_{21}(z) = 0$, FB-BSS では $C_{12}(z) = C_{21}(z) = 0$ である.

信号分離の評価には以下の2種類の SIR(Signal-to-Interference Ratios) を使用する.

$$\sigma_{s1}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|P_{11}(e^{j\omega})S_1(e^{j\omega})|^2 + |P_{22}(e^{j\omega})S_2(e^{j\omega})|^2) d\omega \quad (5.48)$$

$$\sigma_{i1}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|P_{12}(e^{j\omega})S_2(e^{j\omega})|^2 + |P_{21}(e^{j\omega})S_1(e^{j\omega})|^2) d\omega \quad (5.49)$$

$$SIR_1 = 10 \log_{10} \frac{\sigma_{s2}^2}{\sigma_{i2}^2} \quad (5.50)$$

$$\sigma_{s2}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|P_{11}(e^{j\omega})|^2 + |P_{22}(e^{j\omega})|^2) d\omega \quad (5.51)$$

$$\sigma_{i2}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (|P_{12}(e^{j\omega})|^2 + |P_{21}(e^{j\omega})|^2) d\omega \quad (5.52)$$

$$SIR_2 = 10 \log_{10} \frac{\sigma_{s1}^2}{\sigma_{i1}^2} \quad (5.53)$$

$P_{ki}(z)$ は式 (2.24), (4.22) で定義されている信号源から出力までの伝達行列の要素である. すなわち, i 番目の信号源から k 番目の出力までの伝達関数である. 式 (5.48)-(5.53) の場合, 信号源 $S_1(z)$ と $S_2(z)$ はそれぞれ $Y_1(z)$ と $Y_2(z)$ に取り出されることを仮定しているが, 逆に $Y_2(z)$ と $Y_1(z)$ に取り出される場合は負の方向に大きな値を示すので一般性は失われない.

5.4.2 信号分離

Mixture-1 における FB-BSS と FF-BSS の分離性能を比較する. 評価方法には SIR_2 (式 (5.51)-(5.53)) を使用する. 図 5.6 に信号源が白色信号-白色信号の場合の FB-BSS と FF-BSS の比較を示す. ステップサイズは FB-BSS と FF-BSS それぞれ 0.00006, 0.00001 とした.

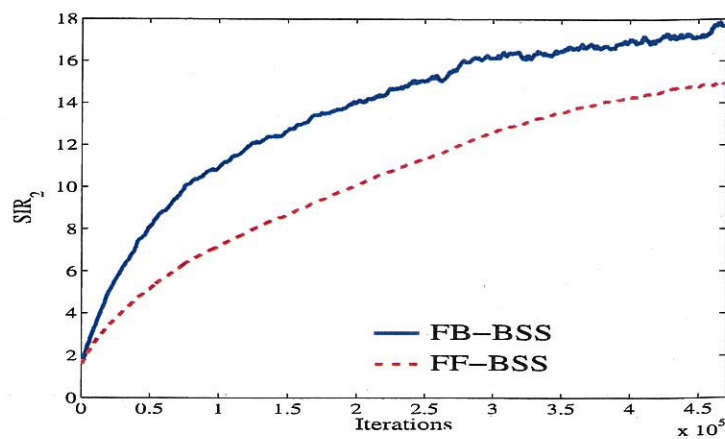


図 5.6: 白色信号における信号分離の比較

FB-BSS に比べ FF-BSS の収束が遅いがこれは学習を安定させるため FF-BSS のステップサイズの値を小さくしているためだと考えられる.

図 5.7 に信号源が音声-音声の場合の FB-BSS と FF-BSS の比較を示す。ステップサイズは FB-BSS と FF-BSS それぞれ 0.000004, 0.000007 とした。

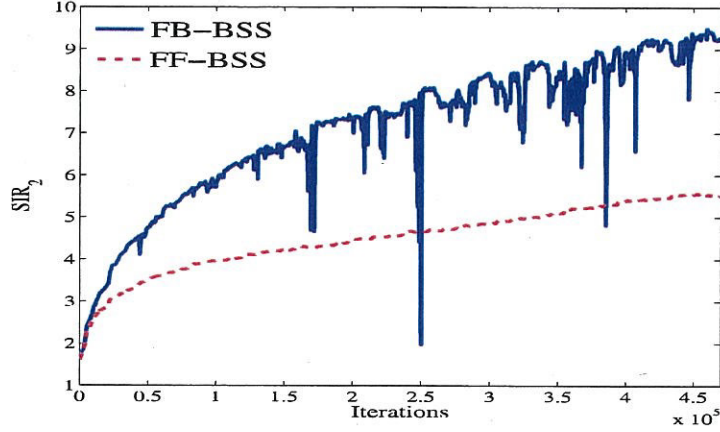


図 5.7: 音声信号における信号分離の比較

信号源が音声のときも白色信号のときと同様 FF-BSS のステップサイズの値を小さくしているため収束速度が FB-BSS に比べ遅くなっている。

5.4.3 信号歪み

さまざまな方法で信号歪みを評価し、FB-BSS と FF-BSS の特徴について解析する。前述の通り、信号歪みの基準を信号源と観測信号の $S_i(z)$ の成分の二通りで評価する。

まず信号の影響を受けずに信号歪みを評価するため伝達関数を使って信号歪みを評価する。歪みの評価 D_1, D_2 の式は以下になる。

$$D_1 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ii}(e^{j\omega}) - \mu_{P_{ii}}|^2 d\omega \quad (5.54)$$

$$D_2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{ii}(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.55)$$

ただし、 $\mu_{P_{ii}}$ は信号源 $S_i(z)$ から出力 $Y_i(z)$ までの伝達関数 $P_{ii}(e^{j\omega})$ の大きさの平均値であり、

$$\mu_{P_{ii}} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |P_{ii}(e^{j\omega})| d\omega \quad (5.56)$$

である。

図 5.8(a) に信号源が白色信号の場合の D_1 を、図 5.8(b) に信号源が白色信号の場合の D_2 を示す。

FB-BSS は D_1 の値ははじめ増加しているが、ある程度のところで収束している。また、 D_2 の値は減少し 0 に近付いている。このことから信号は観測信号の $S_i(z)$ 成分に近付いていることが分かる。今回のシミュレーションでは混合過程の伝達特性が変動しない

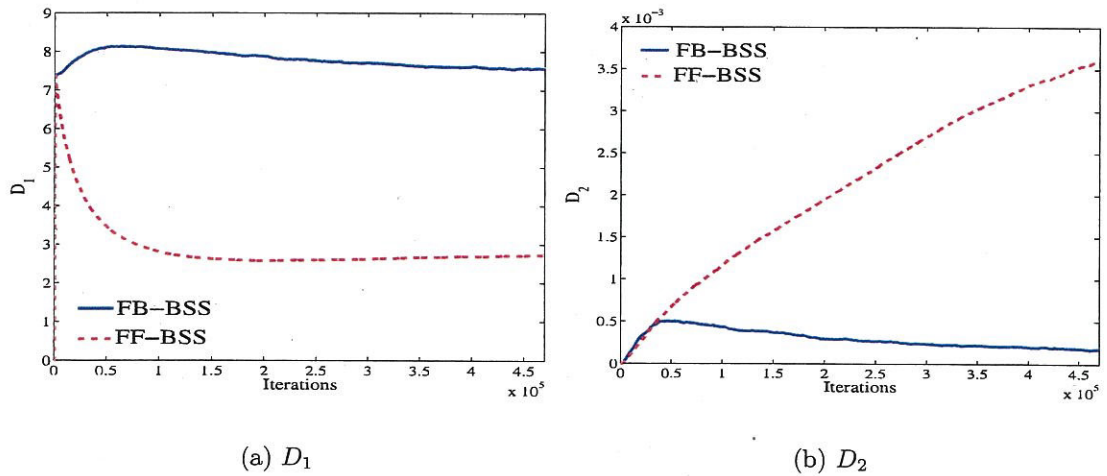


図 5.8: 白色信号に対する係数による歪みの比較

ため出力が観測信号の $S_i(z)$ 成分付近, すなわち信号源から一定の距離で収束したため D_1 の値ははじめ増加しているが, ある程度のところで収束したと考えられる. FF-BSS は FB-BSS とは逆に D_1 の値が減少しているが, D_2 の値が増加し続けている. これは信号源に近付いているように見えるが, D_1 の値はある程度のところで収束しているにも関わらず, D_2 の値が増加し続けていることから信号が白色に近付いているがもとの信号源の白色に戻っていないということが考えられる.

図 5.9(a) に信号源が音声の場合の D_1 を, 図 5.9(b) に信号源が音声の場合の D_2 を示す.

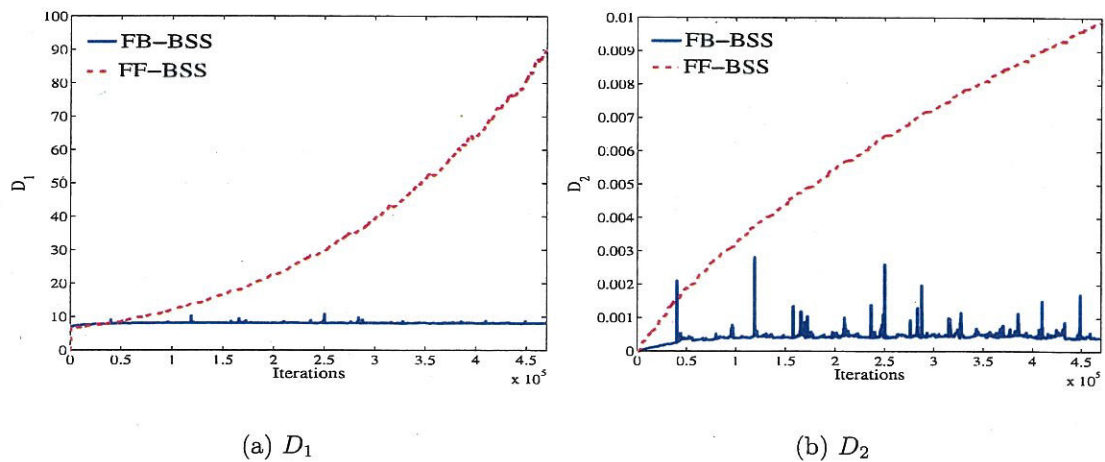


図 5.9: 音声信号に対する係数による歪みの比較

FB-BSS は信号源が白色信号のときと同様の傾向が出ており, 観測信号の $S_i(z)$ 成分に近付いていることが分かる. FF-BSS の D_2 は信号源が白色信号のときと同様, 値が増加し続けているが, D_1 は信号源が白色信号のときとは違い, 値が増加し続けている.

つまり、5.2節で述べたように自由度を利用して出力をより独立な信号にしていることが分かる。

伝達関数を使って歪みを評価したが実際の信号では周波数成分の少ない帯域などがあり、実際の信号の歪みとは異なることが考えられる。そこで信号源 $S_i(z)$ 、観測信号の $S_i(z)$ 成分、出力信号 $Y_i(z)$ を使用して信号から歪みを評価する。

$$D_1 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.57)$$

$$D_2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |XX_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.58)$$

ここで、 $S_i(e^{j\omega})$ 、 $XX_i(e^{j\omega})$ と $Y_i(e^{j\omega})$ はそれぞれ信号源、観測信号の $S_i(z)$ 成分、出力のスペクトルである。ただし、正確には信号全体でスペクトルを計算すべきだが、信号歪みの過程をみるために以下のショートタームでFFTを行なった。

$$S_i(e^{j\omega}) = \text{fft}[s_i(n), s_i(n-1), \dots, s_i(n-N+1)] \quad (5.59)$$

$$XX_i(e^{j\omega}) = \text{fft}[xx_i(n), xx_i(n-1), \dots, xx_i(n-N+1)] \quad (5.60)$$

$$xx_i(n) = \sum_l h_{ii}(l) s_i(n-l) \quad (5.61)$$

$$Y_i(e^{j\omega}) = \text{fft}[y_i(n), y_i(n-1), \dots, y_i(n-N+1)] \quad (5.62)$$

図 5.10(a) に信号源が白色信号の場合の D_1 を、図 5.10(b) に信号源が白色信号の場合の D_2 を示す。 D_2 の傾向は FB-BSS、FF-BSS とともに伝達関数で評価したものと同様だ

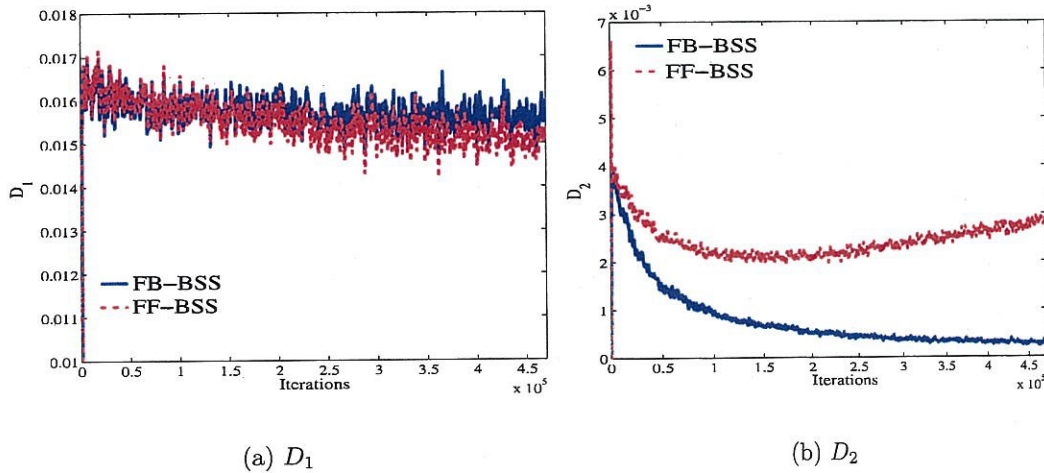


図 5.10: 白色信号に対する信号スペクトルによる歪みの比較

が、 D_1 に関しては伝達関数で評価したものと違い FB-BSS、FF-BSS の両者に違いはあまりない。

また図 5.11(a) に信号源が音声の場合の D_1 を、図 5.11(b) に信号源が音声の場合の D_2 を示すが、信号源に白色信号を用いたときと同様、 D_2 の傾向は FB-BSS、FF-BSS と

もに伝達関数で評価したものと同様だが、 D_1 に関しては伝達関数で評価したものと違い FB-BSS, FF-BSS の両者に違いはあまりない。

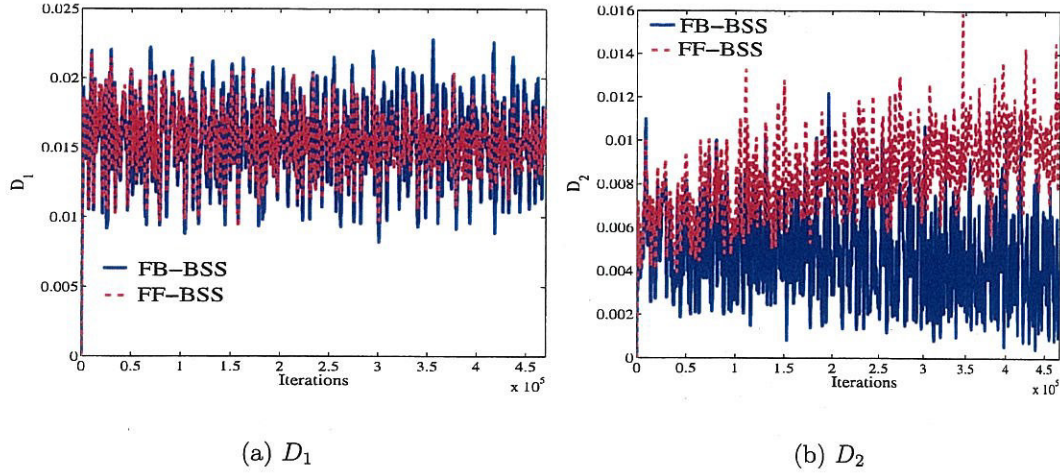


図 5.11: 音声信号に対する信号スペクトルによる歪みの比較

これについては以下の3通りの考え方ができる。

- 学習初期段階では信号が分離されておらず、干渉成分が評価に悪影響を与えている
- 時間遅れにより位相にずれが生じている
- 信号スペクトルの評価では D_1 に差がでない

以降、これらについて検討していく。

まずは学習初期段階では信号が分離されておらず、干渉成分が評価に悪影響を与えているかどうかを解析する。

干渉成分の影響を除去するために完全分離後の分離回路の値に固定して使用し、信号歪みを計算する。すなわち以下の式で評価する。

$$D_1 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_i(e^{j\omega}) - \hat{Y}_i(e^{j\omega})|^2 d\omega \quad (5.63)$$

$$D_2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |X X_i(e^{j\omega}) - \hat{Y}_i(e^{j\omega})|^2 d\omega \quad (5.64)$$

$$S_i(e^{j\omega}) = \text{fft}[s_i(n), s_i(n-1), \dots, s_i(n-N+1)] \quad (5.65)$$

$$X X_i(e^{j\omega}) = \text{fft}[x x_i(n), x x_i(n-1), \dots, x x_i(n-N+1)] \quad (5.66)$$

$$x x_i(n) = \sum_l h_{ii}(l) s_i(n-l) \quad (5.67)$$

$$\hat{Y}_i(e^{j\omega}) = \text{fft}[\hat{y}_i(n), \hat{y}_i(n-1), \dots, \hat{y}_i(n-N+1)] \quad (5.68)$$

$$\hat{y}_i(n) = x_i(n) - \sum_m c_{ij}(\infty, m) y_j(n-m) \quad (5.69)$$

ここで、分離回路のフィルタ係数は実際には分離がある程度進んだと思われる有限回の学習後の値に固定して使用する。

図 5.12(a) に信号源が白色信号の場合の D_1 を、図 5.12(b) に信号源が白色信号の場合の D_2 を、図 5.13(a) に信号源が音声の場合の D_1 を、図 5.13(b) に信号源が音声の場合の D_2 を示す。

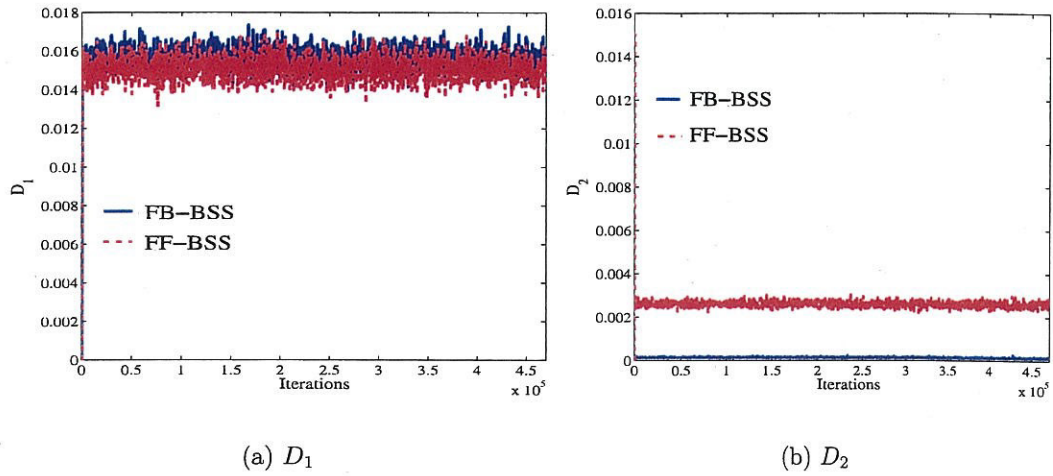


図 5.12: 白色信号に対する学習後の歪みの比較

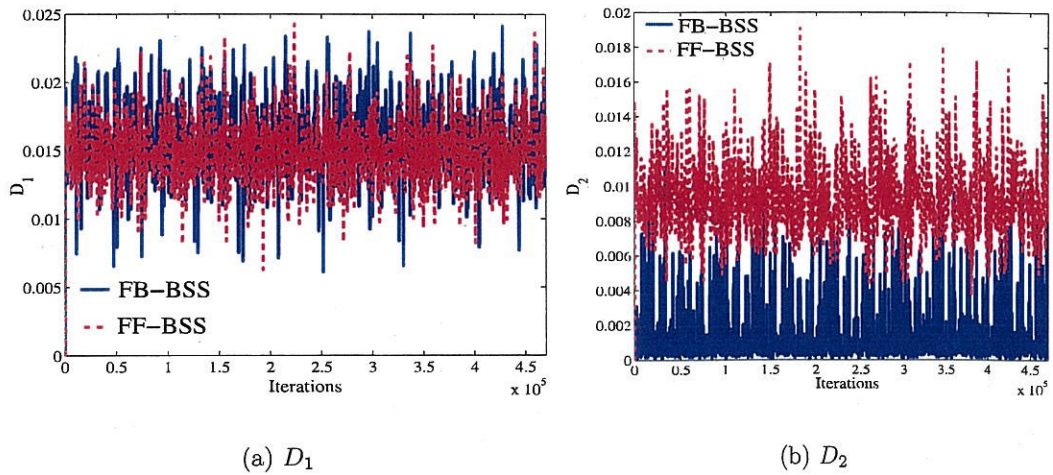


図 5.13: 音声信号に対する学習後の歪みの比較

図 5.12(b), 5.13(b) の D_2 に関してはそれぞれ単純に信号スペクトルを比較したときの D_2 の図 5.10(b), 5.11(b) の学習後半のように差がでている。しかし図 5.12(a), 5.13(a) の D_1 の評価についてはあまり差が見られない。これは単純に信号スペクトルを比較したときと同様の傾向で干渉成分が評価に悪影響を与えているわけではないことが分かる。

続いて時間遅れの影響について検討する。

時間遅れは直線位相となってスペクトルに現れる。そこで位相の情報を無視してスペクトルの大きさのみで比較を行なう。すなわち以下の式で評価を行なう。

$$D_1 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} (|S_i(e^{j\omega})| - |Y_i(e^{j\omega})|)^2 d\omega \quad (5.70)$$

$$D_2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} (|XX_i(e^{j\omega})| - |Y_i(e^{j\omega})|)^2 d\omega \quad (5.71)$$

ここで、 $S_i(e^{j\omega})$ 、 $XX_i(e^{j\omega})$ と $Y_i(e^{j\omega})$ はそれぞれ信号源、観測信号の $S_i(z)$ 成分、出力のショートタームでFFTを行なったスペクトルである。

$$S_i(e^{j\omega}) = \text{fft}[s_i(n), s_i(n-1), \dots, s_i(n-N+1)] \quad (5.72)$$

$$XX_i(e^{j\omega}) = \text{fft}[xx_i(n), xx_i(n-1), \dots, xx_i(n-N+1)] \quad (5.73)$$

$$xx_i(n) = \sum_l h_{ii}(l) s_i(n-l) \quad (5.74)$$

$$Y_i(e^{j\omega}) = \text{fft}[y_i(n), y_i(n-1), \dots, y_i(n-N+1)] \quad (5.75)$$

図 5.14(a) に信号源が白色信号の場合の D_1 を、図 5.14(b) に信号源が白色信号の場合の D_2 を示す。

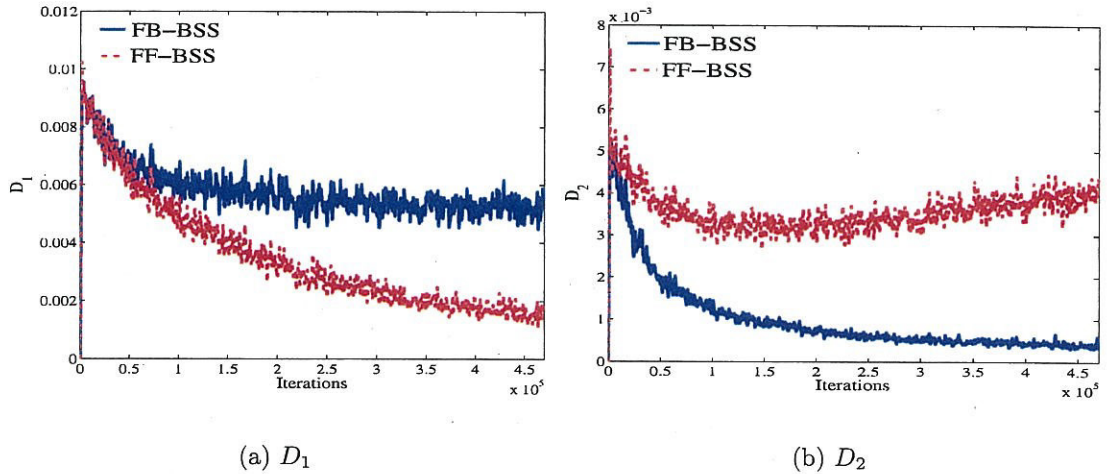


図 5.14: 白色信号に対するスペクトルの振幅による歪みの比較

D_1 はFB-BSSがある程度のところで収束しており、FF-BSSは値が減少している。また D_2 は D_1 の逆にFB-BSSの値が減少しているして0に近付き、FF-BSSは増加し続けている。これは5.2節の議論と一致する。また伝達関数で歪みを評価したときの傾向とも一致する。

図 5.15(a) に信号源が音声の場合の D_1 を、図 5.15(b) に信号源が音声の場合の D_2 を示す。

信号源が音声の場合にも白色信号のときと同様5.2節の議論、また伝達関数で歪みを評価したときの傾向とも一致する。つまり、なんらかの位相変化によって正しく評価で

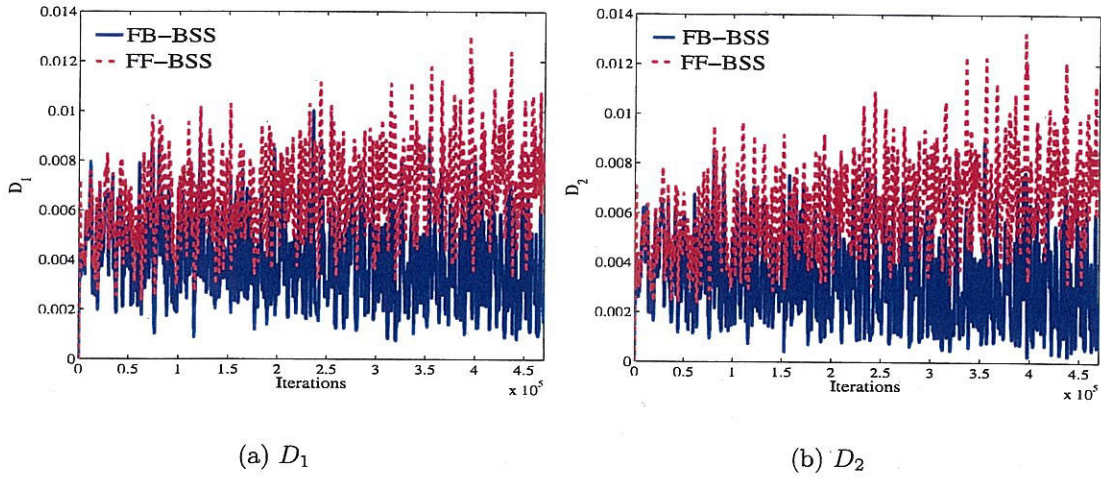


図 5.15: 音声信号に対するスペクトルの振幅による歪みの比較

きていなかったことになる。

時間遅れによる位相変化によって正しく評価できていなかった場合には時間遅れ分を補正して評価すれば正しく評価できる。そこで評価する時間をずらして評価を行なう。

$$D_1 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.76)$$

$$D_2 = \sum_i \frac{1}{2\pi} \int_{-\pi}^{\pi} |XX_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.77)$$

ただし、 $Y_i(e^{j\omega})$ は出力のショートタームでFFTを行なったスペクトル、 $S_i(e^{j\omega})$ 、 $XX_i(e^{j\omega})$ は出力と最も相関が高い点でショートタームでFFTを行なった信号源と観測信号の $S_i(z)$ 成分のスペクトルである。

$$S_i(e^{j\omega}) = \text{fft}[s_i(n - N_{max}^{s_i}), s_i(n - 1 - N_{max}^{s_i}), \dots, s_i(n - N + 1 - N_{max}^{s_i})] \quad (5.78)$$

$$XX_i(e^{j\omega}) = \text{fft}[xx_i(n - N_{max}^{x_i}), xx_i(n - 1 - N_{max}^{x_i}), \dots, xx_i(n - N + 1 - N_{max}^{x_i})] \quad (5.79)$$

$$xx_i(n) = \sum_l h_{ii}(l) s_i(n - l) \quad (5.80)$$

$$Y_i(e^{j\omega}) = \text{fft}[y_i(n), y_i(n - 1), \dots, y_i(n - N + 1)] \quad (5.81)$$

$$N_{max}^{s_i} = \arg \max_l \left(\sum_n s_i(n - l) y_i(n) \right) \quad (5.82)$$

$$N_{max}^{x_i} = \arg \max_l \left(\sum_n x_i(n - l) y_i(n) \right) \quad (5.83)$$

表 5.1 に相関が最大の点を示す。

信号源が白色信号のときと音声のときで相関が最大になる点がずれているがこれは音

表 5.1: 出力信号と入力または観測信号の相関の最大値

Methods	Sources	$Y_1(z) : X X_1(z)$	$Y_2(z) : X X_2(z)$	$Y_1(z) : S_1(z)$	$Y_2(z) : S_2(z)$
FF-BSS	white	0	0	5	1
	speech	1	1	6	1
FB-BSS	white	0	0	5	1
	speech	0	0	6	0

声自身に相関性があり，誤差が生じたためだと考えられる．この表から出力と観測信号の $S_i(z)$ 成分には時間遅れがほとんどないことがわかる．また出力と信号源にはそれぞれ5サンプル，1サンプル程度の時間遅れが発生していることが分かる．これにより信号歪みが正しく評価できていないことが考えられる．そこでこの時間遅れを補正して評価したものを図 5.16, 5.17 に示す．

図 5.16(a) に信号源が白色信号の場合の D_1 を，図 5.16(b) に信号源が白色信号の場合の D_2 を示す．

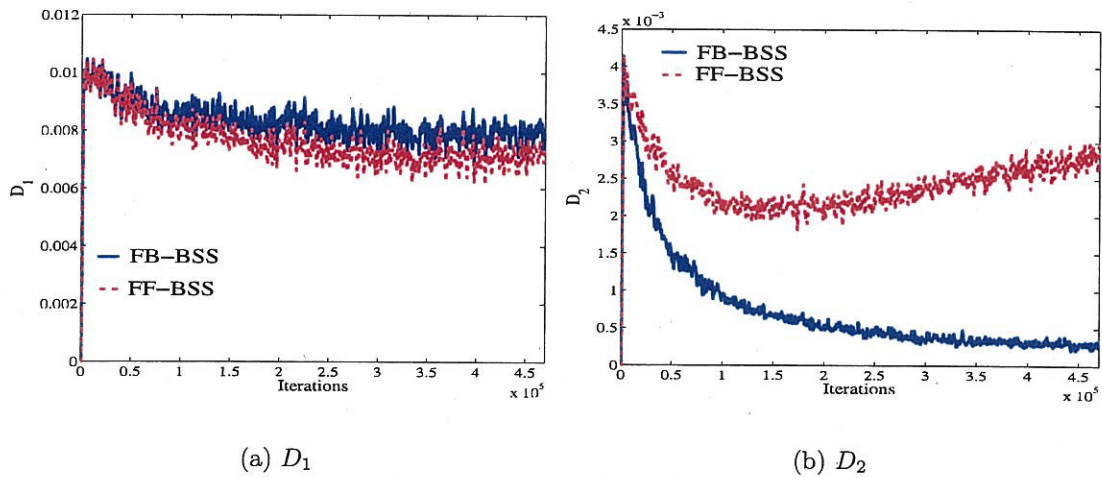


図 5.16: 白色信号に対する相関で時間遅れを考慮したスペクトルによる歪みの比較

D_1 に関して図 5.16(a) はスペクトルをそのまま比較した図 5.10(a) に比べ振幅のみで比較した図 5.14(a) に近いが，図 5.14(a) よりも FF-BSS の歪みが低減されていない結果となっている． D_2 については時間遅れが発生していないのでスペクトルをそのまま比較した図 5.10(b) と同じになる．

図 5.17(a) に信号源が音声の場合の D_1 を，図 5.17(b) に信号源が音声の場合の D_2 を示す．

音声についても白色の時と同様の傾向で D_1 はペクトルをそのまま比較した図 5.11(a) に比べ振幅のみで比較した図 5.15(a) に近いが，図 5.15(a) よりも FB-BSS の歪みが低減されていない結果となっている．また D_2 はスペクトルをそのまま比較した図 5.11(b) よりも FF-BSS と FB-BSS の差が大きくなっている． D_1 の振舞については

- サンプルをスライドさせて時間遅れを補正したがまた π 分のずれが正しい評価を

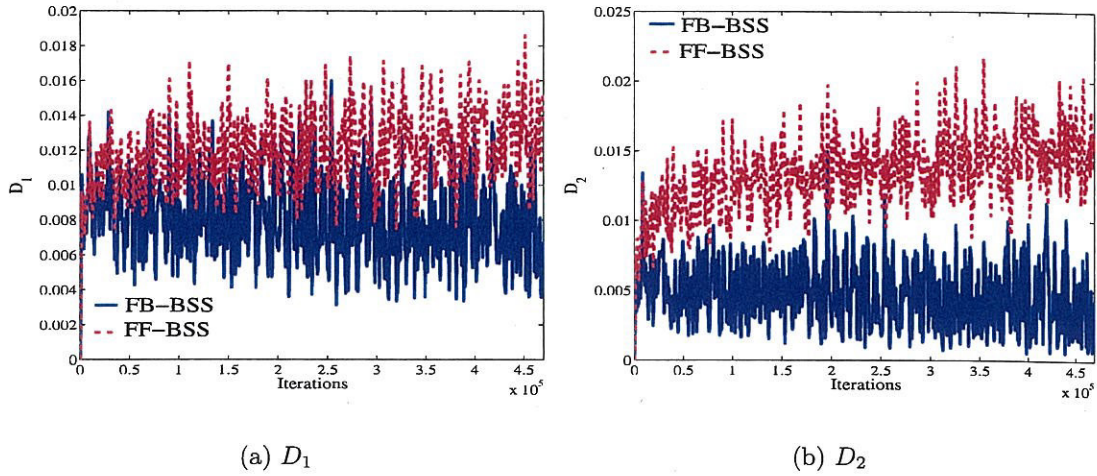


図 5.17: 音声信号に対する相関で時間遅れを考慮したスペクトルによる歪みの比較

妨げている

- 位相の歪みが大きい

の2点が考えられる。

そこで位相を最小2乗近似することによって時間遅れを補正する。まず、位相差が 2π 以上あると位相を正しく求められないため相関が最も高い点に信号をスライドさせて $[-\pi, \pi]$ に収まるようにする。

$$N_{max}^{s_i} = \arg \max_l \left(\sum_n s_i(n-l)y_i(n) \right) \quad (5.84)$$

$$N_{max}^{x_i} = \arg \max_l \left(\sum_n x_i(n-l)y_i(n) \right) \quad (5.85)$$

最終的な位相の補正は信号をスライドさせて行なえるわけではないので $[-\pi, \pi]$ に収まるように補正する際も以下のように位相を回転させる。

$$S_i^{tmp}(e^{j\omega}) = \text{fft}[s_i(n), s_i(n-1), \dots, s_i(n-N+1)] \times \exp\left(\frac{-2j\omega N_{max}^{s_i}}{N}\right) \quad (5.86)$$

$$XX_i^{tmp}(e^{j\omega}) = \text{fft}[xx_i(n), xx_i(n-1), \dots, xx_i(n-N+1)] \times \exp\left(\frac{-2j\omega N_{max}^{x_i}}{N}\right)$$

$$xx_i(n) = \sum_l h_{ii}(l)s_i(n-l) \quad (5.87)$$

$$Y_i(e^{j\omega}) = \text{fft}[y_i(n), y_i(n-1), \dots, y_i(n-N+1)] \quad (5.88)$$

求めたスペクトルの位相差を計算する。

$$YS(\theta) = \angle Y_i(\omega) - \angle S_i^{tmp}(e^{j\omega}) \quad (5.89)$$

$$YX(\theta) = \angle Y_i(\omega) - \angle XX_i^{tmp}(e^{j\omega}) \quad (5.90)$$

ここで最小2乗近似を用いて位相の傾きを計算する.

$$\mathbf{x}^s(n) = \mathbf{A}^+ \mathbf{Y} \mathbf{S}(\theta) \quad (5.91)$$

$$\mathbf{x}^x(n) = \mathbf{A}^+ \mathbf{Y} \mathbf{X}(\theta) \quad (5.92)$$

$$\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \quad (5.93)$$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & N \end{bmatrix} \quad (5.94)$$

$$\mathbf{x}^s(n) = \begin{bmatrix} x_1^s \\ x_2^s \end{bmatrix} \quad (5.95)$$

$$\mathbf{x}^x(n) = \begin{bmatrix} x_1^x \\ x_2^x \end{bmatrix} \quad (5.96)$$

求めた傾きを使用してさらに位相を回転させたスペクトルを求める.

$$S_i(e^{j\omega}) = \text{fft}[s_i(n), s_i(n-1), \dots, s_i(n-N+1)] \times \exp(-j\omega(\frac{2N_{\max}^{s_i}}{N} + x_2^s)) \quad (5.97)$$

$$\begin{aligned} XX_i(e^{j\omega}) &= \text{fft}[xx_i(n), xx_i(n-1), \dots, xx_i(n-N+1)] \times \exp(-j\omega(\frac{2N_{\max}^{x_i}}{N} + x_2^x)) \\ xx_i(n) &= \sum_l h_{ii}(l) s_i(n-l) \end{aligned} \quad (5.98)$$

$$Y_i(e^{j\omega}) = \text{fft}[y_i(n), y_i(n-1), \dots, y_i(n-N+1)] \quad (5.99)$$

最後に時間遅れを補正したスペクトルを用いて信号歪み D_1, D_2 を計算する.

$$D_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.100)$$

$$D_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |XX_i(e^{j\omega}) - Y_i(e^{j\omega})|^2 d\omega \quad (5.101)$$

図5.18(a)に信号源が白色信号の場合の D_1 を, 図5.18(b)に信号源が音声の場合の D_1 を示す. D_2 に関しては図5.16(b), 図5.17(b)と同じなので割愛する.

サンプル点をずらすことによって時間遅れを補正した図5.16(a), 図5.17(a)と比べ, さらに振幅のみで評価した図5.10(b), 図5.11(b)の傾向に近付いている. しかし信号源が白色信号の場合において値の減少が収まりある程度で収束している点で図5.10(b)とは異なる. つまりこれは5.2節で述べたように出力をより独立にしようとしたために白色に近い振幅スペクトルになったがそれは信号源に戻すわけではなく別の系列の白色信号になっただけであると考えられる.

信号源に音声を, 混合過程に Mixture-1 と Mixture-2 を用いて式(5.48)-(5.53)の SIR_1 , SIR_2 , 式(5.3)-(5.10)の SD_{1a} , SD_{1b} , SD_{2a} , SD_{2b} で信号分離, 信号歪みを評価したものを表5.2に示す. SIR_2 に関しては若干FB-BSSの方が性能がよい. 最も厳密な信

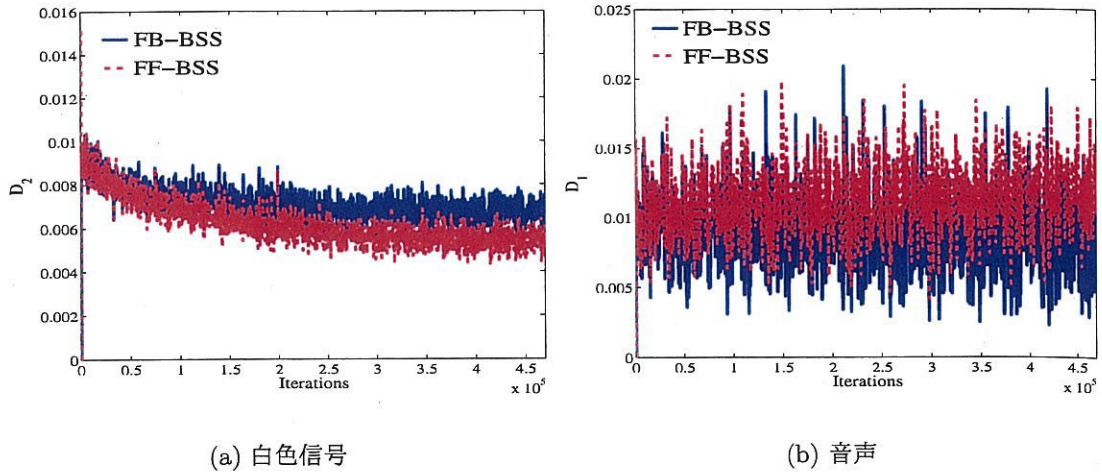
図 5.18: 時間遅れ補正を施したスペクトルによる D_1 の比較

表 5.2: FF-BSS と FB-BSS の比較

Methods	Mix	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-BSS	1	12.2	5.56	0.25	-2.94	0.57	-3.82
	2	19.3	12.1	0.27	-2.95	0.73	-3.69
FB-BSS	1	14.1	9.24	-14.5	-17.9	-14.7	-17.3
	2	22.9	16.0	-28.7	-31.5	-22.9	-23.9

号歪みの評価である SD_{1a} は FF-BSS がほぼ 0dB であり FB-BSS に比べて歪みが大きい。振幅のみで評価をする SD_{1b} ではどちらも SD_{1a} に比べて値が良くなっているが傾向としては SD_{1a} と同じ傾向である。伝達関数で評価している SD_{2a} と SD_{2b} も SD_{1a} , SD_{1b} 同様 FF-BSS がほぼ 0dB であり FB-BSS に比べて歪みが大きい。

第6章 従来の信号歪み抑制学習法

第5章で述べたように FF-BSS では自由度が存在し，その自由度によって出力信号をより独立にしようとするため信号歪みが生じてしまう．実用上，信号歪みは深刻な問題であり，信号歪みを軽減する方法が必要となる．

本章では信号歪みの抑制を行なう従来の手法として Minimal Distortion Principle に基づく BSS (MDP-BSS)[18] と Projection Back 法 [23] を紹介する．MDP-BSS は時間領域で学習を行なう FF-BSS に対して適用される信号歪み抑制のアルゴリズムであり，Projection Back 法は周波数領域で学習を行なう FF-BSS に対して適用される．

6.1 Minimal Distortion Principle に基づく BSS (MDP-BSS)

MDP-BSS は時間領域で学習を行なう FF-BSS に対し，歪みの基準からの距離を評価関数に加えることにより信号歪みを抑制する方式である．この手法で用いられる歪みの基準は第5章で定義した基準と同様のものを使用している．

本章ではこの手法により信号歪みを低減できることを確認する．さらにこの手法の問題点を指摘し，シミュレーションによりその問題点を確認する．

6.1.1 Minimal Distortion Principle(MDP)

最適有効分離回路

以下の形の分離回路を有効分離回路と呼ぶことにする．

$$W(z) = PD(z)H^{-1}(z) \quad (6.1)$$

ここで P は順序行列， $D(z)$ は任意の対角行列

$$D(z) = \begin{bmatrix} d_1(z) & 0 & \cdots & 0 \\ 0 & d_2(z) & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & d_N(z) \end{bmatrix} \quad (6.2)$$

である．ただし簡単化のため

$$P = I \quad (6.3)$$

とする．すなわちここでは

$$W(z) = D(z)H^{-1}(z) \quad (6.4)$$

を有効分離回路と呼ぶことにする。

もし分離回路が有効分離回路であれば信号源は線形伝達特性 $d_i(z)$ を通って、ある1つの出力のみに現れる。

ブラインドソースセパレーションにおいてすべての有効分離回路は等価であると考えられるが、以下の分離回路は特別な意味を持つ。

$$\mathbf{W}^*(z) = \text{diag } \mathbf{H}(z) \mathbf{H}^{-1}(z) \quad (6.5)$$

これを最適有効分離回路と呼ぶことにする。ここで最適有効分離回路は一意に定まるものでないことに注意する。任意の対角行列 $\mathbf{E}(z)$ に対して以下のことが成り立つ。

$$\text{diag } \mathbf{H}(z) \mathbf{E}(z) (\mathbf{H}(z) \mathbf{E}(z))^{-1} = \text{diag } \mathbf{H}(z) \mathbf{H}^{-1}(z) \quad (6.6)$$

したがって、信号源の定義における不確定性には関係なく最適有効分離回路は定義できる。

最適有効分離回路に関する定理

最適有効分離回路 $\mathbf{W}^*(z)$ に関して以下のことが言える。

定理1：最適有効分離回路 $\mathbf{W}^*(z)$ は

$$\|\mathbf{W}(z) \mathbf{H}(z) - \mathbf{H}(z)\|^2 \quad (6.7)$$

を最小化する有効分離回路である。

(証明)

$$\|\mathbf{W}(z) \mathbf{H}(z) - \mathbf{H}(z)\|^2 = \|\mathbf{D}(z) - \mathbf{H}(z)\|^2 \quad (6.8)$$

$$= \int_{-\frac{1}{2}}^{\frac{1}{2}} \|\tilde{\mathbf{D}}(f) - \tilde{\mathbf{H}}(f)\|^2 df \quad (6.9)$$

ここで、上式において f について $\tilde{\mathbf{D}}(f)$ を変化させて $\|\tilde{\mathbf{D}}(f) - \tilde{\mathbf{H}}(f)\|^2$ を最小化することを考えると $\tilde{\mathbf{D}}(f) = \text{diag } \tilde{\mathbf{H}}(f)$ ，すなわち $\mathbf{D}(z) = \text{diag } \mathbf{H}(z)$ が上式を最小化させることがわかる。これを式(6.4)に代入することによって式(6.5)を得ることができる。

定理2：最適有効分離回路 $\mathbf{W}^*(z)$ は

$$E [\|\mathbf{y}(n) - \mathbf{x}(n)\|^2] \quad (6.10)$$

を最小化する有効分離回路である。

(証明)

$$E [\|\mathbf{y}(n) - \mathbf{x}(n)\|^2] = \int_{-\frac{1}{2}}^{\frac{1}{2}} \text{tr}(\tilde{\mathbf{D}}(f) - \tilde{\mathbf{H}}(f)) \Phi_s(f) (\tilde{\mathbf{D}}(f) - \tilde{\mathbf{H}}(f))^H df \quad (6.11)$$

ここで $\Phi_s(f)$ は $s(n)$ のパワースペクトラムである。すなわち、自己相関 $E[s(n)s(n+\tau)]$ のフーリエ変換したものである。上式において f について $\tilde{\mathbf{D}}(f)$ を変化させて $\text{tr}(\tilde{\mathbf{D}}(f) -$

$\tilde{\mathbf{H}}(f)\Phi_s(f)(\tilde{\mathbf{D}}(f) - \tilde{\mathbf{H}}(f))^H$ を最小化することを考える. $\Phi_s(f)$ は対角行列であるので, $\tilde{\mathbf{D}}(f) = \text{diag } \tilde{\mathbf{H}}(f)$, すなわち $\mathbf{D}(z) = \text{diag } \mathbf{H}(z)$ が上式を最小化させることがわかる.

これらの2つの定理は Minimal Distortion Principle を2通りの方法で記述している. つまり最適有効分離回路は $\mathbf{W}(z)\mathbf{H}(z)$ を可能な限り $\mathbf{H}(z)$ に近づける, またその出力 $\mathbf{y}(n)$ を可能な限り $\mathbf{x}(n)$ に近づけるものである.

定理3: 最適有効分離回路 $\mathbf{W}^*(z)$ は

$$\text{diag } \mathbf{W}^{-1}(z) = \mathbf{I} \quad (6.12)$$

を満たす有効分離回路である.

(証明)

上式が成り立つとき式(6.4)より

$$\text{diag } \mathbf{H}(z)\mathbf{D}^{-1}(z) = \mathbf{I} \quad (6.13)$$

となる. この式はつまり

$$\mathbf{D}(z) = \text{diag } \mathbf{A}(z) \quad (6.14)$$

となることを示す.

最適有効分離回路の性質

最適有効分離回路は実際の BSS への実装において有益なさまざまな特徴を持つ.

1. 最適有効分離回路において分離された出力信号は

$$\mathbf{Y}(z) = \mathbf{W}^*(z)\mathbf{H}(z)\mathbf{S}(z) \quad (6.15)$$

$$= \text{diag } \mathbf{H}(z)\mathbf{H}^{-1}(z)\mathbf{H}(z)\mathbf{S}(z) \quad (6.16)$$

$$= \text{diag } \mathbf{H}(z)\mathbf{S}(z) \quad (6.17)$$

となる. これは第5.1.2節で定義した歪みの基準と一致する. つまり, 最適有効分離回路において分離された出力信号は観測信号における唯一の信号源成分となる. この性質は信号分離の解釈と後処理に役立つ.

2. 最適有効分離回路は信号源の性質に依存せず, 混合過程 $\mathbf{H}(z)$ にのみ依存する. そのため音声のような非定常な信号に対しても混合過程が一定であれば最適有効分離回路は変化しない.
3. 実際のインプリメントでは分離回路には FIR フィルタを必要とする. フィルタ次数はできる限り低くしたい. Minimal Distortion Principle に基づいた分離回路は出力をできる限り観測信号に近づけるので低次で実現できる.

6.1.2 学習アルゴリズム

MDP の実装

前述の通り, Minimal Distortion Principle は定理 1, 定理 2 の 2 通りに記述できる. 定理 1 は Minimal Distortion Principle を伝達関数を使用して表現しているものであるが, 混合過程の伝達関数 $H(z)$ を使用しており, 実際の学習には使用できない. 一方, 定理 2 は出力信号, 観測信号を使用し表現しているため学習中に使用できる. そこで定理 2 により学習に Minimal Distortion Principle を導入する.

すなわち

$$\psi = Y(z) - X(z) \quad (6.18)$$

とし, $E[||\psi||^2]$ の最小値を求めることにより Minimal Distortion Principle を実現する. ただし, 実際には信号の特性を決定論的に得ることはできないため瞬時値 $||\psi(n)||^2$ を勾配法を用いて最小値を求めることになる.

ここで, $||\psi||^2$ を因数分解すると

$$||\psi||^2 = ||Y(z) - X(z)||^2 \quad (6.19)$$

$$= ||W(z)X(z) - X(z)||^2 \quad (6.20)$$

$$= (W(z)X(z) - X(z))^T (W(z)X(z) - X(z)) \quad (6.21)$$

$$= X^T(z)W^T(z)W(z)X(z) - X^T(z)W^T(z)X(z) - X^T(z)W(z)X(z) + X^T(z)X(z) \quad (6.22)$$

となる.

各項の $W(z)$ に対する勾配は

$$\frac{\partial}{\partial W(z)} X^T(z)W^T(z)W(z)X(z) = 2W(z)X(z)X^T(z) \quad (6.23)$$

$$\frac{\partial}{\partial W(z)} X^T(z)W^T(z)X(z) = X(z)X^T(z) \quad (6.24)$$

$$\frac{\partial}{\partial W(z)} X^T(z)W(z)X(z) = X(z)X^T(z) \quad (6.25)$$

$$\frac{\partial}{\partial W(z)} X^T(z)X(z) = 0 \quad (6.26)$$

となるので $||\psi||^2$ の $W(z)$ に対する勾配は

$$\frac{\partial ||\psi||^2}{\partial W(z)} = 2W(z)X(z)X^T(z) - 2X(z)X^T(z) \quad (6.27)$$

$$= 2(W(z)X(z) - X(z))X^T(z) \quad (6.28)$$

$$= 2(Y(z) - X(z))X^T(z) \quad (6.29)$$

となる.

そこで, z 多項式における z^{-l} 項の係数 $w(l)$ に対する勾配は微分のチェーンルールを用いて

$$\frac{\partial ||\psi||^2}{\partial w(l)} = \frac{\partial ||\psi||^2}{\partial W(z)} \frac{\partial W(z)}{\partial w(l)} \quad (6.30)$$

$$= 2(Y(z) - X(z))X^T(z)z^{-1} \quad (6.31)$$

となる。

上式の勾配を用いて分離回路の z^{-l} 項の係数 $w(l)$ を更新することによって $\|\psi\|^2$ を最小化する分離回路を求めることができる。

係数の更新

定理2は $\|\psi\|^2$ を最小化する有効分離回路が最適有効分離回路としている。つまり $\|\psi\|^2$ を最小化しても分離回路が有効分離回路でなければ最適有効分離回路を得られない。

そこで従来の学習により有効分離回路を求めるつつ、 $\|\psi\|^2$ を最小化することによって最適有効分離回路を求める。式(4.43)で与えられる *Kullback - Leibler divergence* の傾きに $\|\psi\|^2$ の傾きを付加して分離回路の係数の更新式を得る。なおその際、重み $\beta/2$ により修正の度合を調整する。すなわち

$$\frac{\partial KL(W(z))}{\partial w(l)} + \frac{\beta}{2} \frac{\partial \|\psi\|^2}{\partial w(l)} = -(I - E_y[\varphi(Y(z))Y(z)^T])(W(z)^T)^{-1}z^{-l} + \beta(Y(z) - X(z))X^T(z)z^{-l} \quad (6.32)$$

となる。

ここで従来法と同様、

$$E \left[\frac{\partial KL(W(z))}{\partial w(l)} + \frac{\beta}{2} \frac{\partial \|\psi\|^2}{\partial w(l)} \right] = 0 \quad (6.33)$$

と

$$E[G(Y(z))] = E \left[\left(\frac{\partial KL(W(z))}{\partial w(l)} + \frac{\beta}{2} \frac{\partial \|\psi\|^2}{\partial w(l)} \right) W^T(z)W(z) \right] = 0 \quad (6.34)$$

は等価であるため修正項には $G(Y(z))$ を用いることにする。

$G(Y(z))$ は式(4.45)で表される $F(Y(z))$ を用いて

$$G(Y(z)) = \left(\frac{\partial KL(W(z))}{\partial w(l)} + \frac{\beta}{2} \frac{\partial \|\psi\|^2}{\partial w(l)} \right) W^T(z)W(z) \quad (6.35)$$

$$= F(Y(z)) + \beta(Y(z) - X(z))X^T(z)z^{-l}W^T(z)W(z) \quad (6.36)$$

$$= F(Y(z)) + \beta(Y(z) - X(z))Y^T(z)W(z)z^{-l} \quad (6.37)$$

と表される。

したがって分離回路の係数の更新式は $G(Y(z))$ を逆変換して

$$\begin{aligned} w(n+1, l) &= w(n, l) + \alpha \sum_{m=0}^{K_w-1} [\text{diag}(\langle \Phi(y(n))y^T(n-l+m) \rangle \\ &\quad - \langle \Phi(y(n))y^T(n-l+m) \rangle \\ &\quad - \beta(y(n) - x(n))y^T(n-l+m)]w(n, l) \end{aligned} \quad (6.38)$$

$$\varphi(y(n)) = \frac{1 - e^{-y(n)}}{1 + e^{-y(n)}} \quad (6.39)$$

と求められる。

6.1.3 MDP-BSS の問題点

MDP-BSS では出力

$$Y_i(z) = P_{ii}(z)S_i(z) + P_{ij}(z)S_j(z) \quad (6.40)$$

を観測信号

$$X_i(z) = H_{ii}(z)S_i(z) + H_{ij}(z)S_j(z) \quad (6.41)$$

に近づけるように制約を掛ける．すなわち， $Y_i(z) - X_i(z)$ を 0 にする制約を掛ける． $Y_i(z) - X_i(z)$ は以下のように書き表すことができる．

$$Y_i(z) - X_i(z) = P_{ii}(z)S_i(z) + P_{ij}(z)S_j(z) - (H_{ii}(z)S_i(z) + H_{ij}(z)S_j(z)) \quad (6.42)$$

$$= (P_{ii}(z) - H_{ii}(z))S_i(z) + (P_{ij}(z) - H_{ij}(z))S_j(z) \quad (6.43)$$

ここで，信号源 $S_i(z)$ と $S_j(z)$ は互いに独立であるため恒等的に $Y_i(z) - X_i(z)$ を 0 にするためには

$$P_{ii}(z) = H_{ii}(z) \quad (6.44)$$

$$P_{ij}(z) = H_{ij}(z) \quad (6.45)$$

が成り立つ必要がある．

今，出力 $Y_i(z)$ として信号源 $S_i(z)$ が取り出された場合，理想的には信号源から出力までの伝達行列 $\mathbf{P}(z)$ は

$$\mathbf{P}(z) = \begin{bmatrix} H_{11}(z) & 0 \\ 0 & H_{11}(z) \end{bmatrix} \quad (6.46)$$

となる．したがって，式 (6.44) は信号歪みを抑制する効果を持つ．しかし式 (6.45) は理想的な $\mathbf{P}(z)$ とは異なり，完全分離の条件である $\mathbf{P}(z)$ の各行各列にただ 1 つだけ 0 でない要素を持つという条件に反することになる．すなわち，式 (6.45) は信号分離を妨げることになる．また，出力 $Y_i(z)$ として信号源 $S_j(z)$ が取り出された場合においても，式 (6.45) は信号歪みを抑制する効果を持つが式 (6.44) が信号分離を妨げることになる．

6.2 Projection Back 法

Projection Back 法は周波数領域で学習を行なう FF-BSS に対し、大きさを正規化することにより信号歪みを抑制する方式である。この手法で完全分離が行なわれたときに得られる信号は第5章で定義した信号歪みの基準と一致する。

本章では、Projection Back 法の動作原理を紹介し、2通りの実装を行ないそれぞれについて性能を検討する。

6.2.1 周波数領域における信号歪み

FF-BSS では5.2.1節で述べたように自由度が存在する。一方、FF-BSS の学習は出力の独立性を高めるように行なわれ、5.2.2節で述べたように信号の分離と信号のスペクトルを変化させることによって実現される。本稿では観測信号の信号源の成分 $H_{ii}(z)S_i(z)$ からの信号スペクトルの差を信号歪みとしている。

一方、周波数領域で学習を行なう場合、FF-FREQ(1)、FF-FREQ(2) の係数更新はそれぞれ

$$\mathbf{W}(r+1, m) = \mathbf{W}(r, m) + \eta[\mathbf{I} - \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \quad (6.47)$$

$$\begin{aligned} \mathbf{W}(r+1, m) = & \mathbf{W}(r, m) + \eta[\text{diag}(\langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle) \\ & - \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle] \mathbf{W}(r, m) \end{aligned} \quad (6.48)$$

となり、各周波数帯 m で独立に学習が行なわれる。すなわち、各周波数帯 m 内で成分を独立にするように学習される。独立性は信号の大きさには関係がなく、BSS では大きさに対して任意性が存在する。各周波数に対して任意の大きさになるため信号スペクトルが変化する。

FF-FREQ(1) の場合、第4章で述べたように式(6.47)の収束条件は

$$\mathbf{I} = \langle \Phi(\mathbf{Y}(r, m)) \mathbf{Y}^H(r, m) \rangle \quad (6.49)$$

のように表され、収束時の信号パワー $\langle \Phi(\mathbf{Y}_p) \mathbf{Y}_q^* \rangle$ は

$$\langle \Phi(\mathbf{Y}_p) \mathbf{Y}_p^* \rangle = 1 \quad (6.50)$$

となり、パワーを1に統一する。しかし、これは出力信号のスペクトルが平坦化され、信号が白色化されるということになり、信号歪みの原因となる。

一方、この白色化を回避するアルゴリズムである FF-FREQ(2) においても5.2.3節で述べたように式(6.48)の更新量が

$$\eta \begin{bmatrix} 0 & -\langle \Phi(\mathbf{Y}_1(r, m)) \mathbf{Y}_2^*(r, m) \rangle \\ -\langle \Phi(\mathbf{Y}_2(r, m)) \mathbf{Y}_1^*(r, m) \rangle & 0 \end{bmatrix} \quad (6.51)$$

であり、出力のパワー、すなわち観測信号のパワーに大きく依存し、信号の種類によっては信号歪みが生じてしまう。

つまり、周波数領域で学習を行なう際には、大きさによる任意性を適切な方法により抑えなければならない。

6.2.2 信号の正規化

正規化行列 $D(z)$ を用いて出力信号を正規化し、信号歪みを抑える。正規化した出力信号 $\hat{Y}(z)$ は

$$\hat{Y}(z) = D(z)Y(z) \quad (6.52)$$

となる。以下に正規化行列 $D(z)$ が取るべき値について求める。

式 (4.21) より信号源 $S(z)$ と出力 $Y(z)$ の関係は

$$Y(z) = W(z)H(z)S(z) \quad (6.53)$$

となり、式 (4.22) の伝達行列 $P(z)$ を用いると

$$Y = P(z)S(z) \quad (6.54)$$

と表せる。したがって、正規化した出力信号 $\hat{Y}(z)$ と信号源 $S(z)$ の関係は

$$\hat{Y}(z) = D(z)Y(z) \quad (6.55)$$

$$= D(z)P(z)S(z) \quad (6.56)$$

となる。

ここで、信号が完全に分離されていて、かつ第5章で定義した信号歪みが起こっていないとすると、

$$\hat{Y}(z) = \Lambda(z)S(z) \quad (6.57)$$

が成り立つ。ここで、

$$\Lambda(z) = \text{diag}[H(z)] \quad (6.58)$$

である。式 (6.56) と式 (6.57) から

$$D(z)P(z) = \Lambda(z) \quad (6.59)$$

が成り立つ。

一方、式 (4.22) の定義より信号源から出力までの伝達行列 $P(z)$ は

$$P(z) = W(z)H(z) \quad (6.60)$$

である。これを混合過程における混合行列 $H(z)$ について解くと

$$H(z) = W^{-1}(z)P(z) \quad (6.61)$$

となる。さらに両辺の対角要素のみを比較すると

$$\Lambda(z) = \text{diag}[W^{-1}(z)P(z)] \quad (6.62)$$

のように表せる。

今、分離の学習によって信号が完全に行なわれている場合、 $P(z)$ は対角行列である。したがって上式は

$$\Lambda(z) = \text{diag}[W^{-1}(z)]P(z) \quad (6.63)$$

となる．式 (6.59), (6.63) より

$$D(z)P(z) = \text{diag}[W^{-1}(z)]P(z) \quad (6.64)$$

が成り立つ．すなわち完全分離出来たときの出力が歪まないようにするため正規化行列 $D(z)$ は

$$D(z) = \text{diag}[W^{-1}(z)] \quad (6.65)$$

と設定する．これは分離回路 $W(z)$ のみに依存しており，新たにパラメータを推定する必要がない．

6.2.3 学習アルゴリズム

前節で求めた正規化行列 $D(z)$ を用いて出力の信号歪みを抑える．すなわち，従来の学習アルゴリズム FF-FREQ(1), FF-FREQ(2) によって分離された信号 $Y(r, m)$ に正規化行列 $D(r, m)$ を乗ずることにより歪み補正出力 $\hat{Y}(r, m)$ を得る．つまり，観測信号 $X(r, m)$ に対し，

$$Y(r, m) = W(r, m)X(r, m) \quad (6.66)$$

により分離出力 $Y(r, m)$ を得る．さらに

$$\hat{Y}(r, m) = D(r, m)Y(r, m) \quad (6.67)$$

$$D(r, m) = \text{diag}[W^{-1}(r, m)] \quad (6.68)$$

により歪み補正出力 $\hat{Y}(r, m)$ を得る．分離回路 $W(r, m)$ は従来の学習アルゴリズムにより更新される．すなわち FF-FREQ(1) においては

$$W(r+1, m) = W(r, m) + \eta[I - \langle \Phi(Y(r, m))Y^H(r, m) \rangle]W(r, m) \quad (6.69)$$

FF-FREQ(2) においては

$$\begin{aligned} W(r+1, m) &= W(r, m) + \eta[\text{diag}(\langle \Phi(Y(r, m))Y^H(r, m) \rangle) \\ &\quad - \langle \Phi(Y(r, m))Y^H(r, m) \rangle]W(r, m) \end{aligned} \quad (6.70)$$

をそれぞれ用いて係数の更新を行なう．本章では式 (6.69) を用いて係数を更新し，Projection Back 法を用いて信号歪みを補正した方法を FF-FREQ(1+PB) と記述し，式 (6.70) を用いて係数を更新し，Projection Back 法を用いて信号歪みを補正した方法を FF-FREQ(2+PB) と記述することにする．

また，

$$Y(z) = W(z)X(z) \quad (6.71)$$

であるため

$$\hat{Y}(z) = D(z)Y(z) \quad (6.72)$$

$$= D(z)W(z)X(z) \quad (6.73)$$

$$= \text{diag}[W^{-1}(z)]W(z)X(z) \quad (6.74)$$

と書き表せる.

観測信号から出力までの回路を分離回路と考えると $\text{diag}[\mathbf{W}^{-1}(z)]\mathbf{W}(z)$ 全体を分離回路と考えることができる. すなわち,

$$\hat{\mathbf{Y}}(z) = \hat{\mathbf{W}}(z)\mathbf{X}(z) \quad (6.75)$$

$$\hat{\mathbf{W}}(z) = \text{diag}[\mathbf{W}^{-1}(z)]\mathbf{W}(z) \quad (6.76)$$

となり, 分離回路 $\hat{\mathbf{W}}(z)$ を学習することになる. つまり, 信号歪みの補正を学習に組み込むことができる.

更新式は以下のようなになる.

FF-FREQ(1) を用いる場合

$$\mathbf{W}(r+1, m) = \hat{\mathbf{W}}(r, m) + \eta[\mathbf{I} - \langle \Phi(\mathbf{Y}(r, m))\mathbf{Y}^H(r, m) \rangle] \hat{\mathbf{W}}(r, m) \quad (6.77)$$

$$\hat{\mathbf{W}}(r+1, m) = \text{diag}[\mathbf{W}^{-1}(r+1, m)]\mathbf{W}(r+1, m) \quad (6.78)$$

FF-FREQ(2) を用いる場合

$$\begin{aligned} \mathbf{W}(r+1, m) = & \hat{\mathbf{W}}(r, m) + \eta[\text{diag}(\langle \hat{\Phi}(\mathbf{Y}(r, m))\hat{\mathbf{Y}}^H(r, m) \rangle) \\ & - \langle \hat{\Phi}(\hat{\mathbf{Y}}(r, m))\hat{\mathbf{Y}}^H(r, m) \rangle] \hat{\mathbf{W}}(r, m) \end{aligned} \quad (6.79)$$

$$\hat{\mathbf{W}}(r+1, m) = \text{diag}[\mathbf{W}^{-1}(r+1, m)]\mathbf{W}(r+1, m) \quad (6.80)$$

本章では式 (6.77), (6.78) で更新するものを FF-FREQ(1+LPB) と記述し, 式 (6.79), (6.80) で更新するものを FF-FREQ(2+LPB) と記述することにする.

6.3 シミュレーション

6.3.1 シミュレーション条件

シミュレーションの条件は5.4節と同様に2信号源, 2センサで行ない混合過程の伝達特性には図5.5の $H_{11}(z)$ と $H_{22}(z)$,
(Mixture-1)

$$H_{21}(z) = 0.9z^{-1}H_{11}(z) \quad (6.81)$$

$$H_{12}(z) = 0.9z^{-1}H_{22}(z) \quad (6.82)$$

(Mixture-2)

$$H_{21}(z) = 0.5z^{-1}H_{11}(z) \quad (6.83)$$

$$H_{12}(z) = 0.5z^{-1}H_{22}(z) \quad (6.84)$$

となる $H_{12}(z)$ と $H_{21}(z)$ を用いる.

信号源は以下の男性と女性の音声を繰り返し用いている.

- $s_1(n)$: 音声 「人間は豊かな自然を求めています。」(男性)
 「パンはポルトガル人が伝えたものです。」(男性)
 $s_2(n)$: 音声 「健康のため無理をしないでください。」(女性)
 「太宰治は青森県で生まれました。」(女性)

FFT サイズ, タップ数は256, 分離回路の初期値は $W_{11}(z) = W_{22}(z) = 1$, $W_{12}(z) = W_{21}(z) = 0$ である.

本章で比較に用いる手法の表記方法を表6.1にまとめる.

6.3.2 信号スペクトル

信号歪みの基準となる $H_{ii}(z)S_i(z)$ を図6.1に示す.

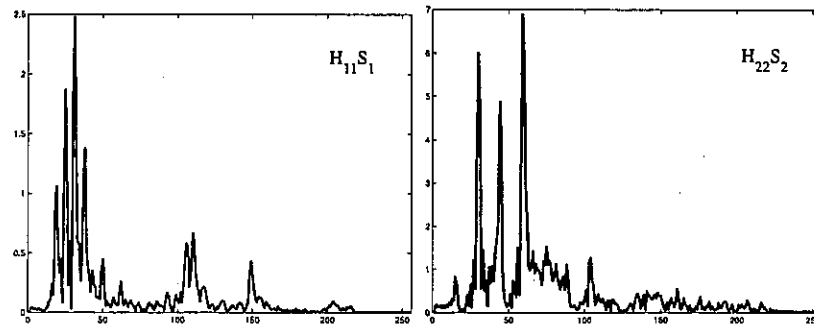


図 6.1: $H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル

また, FF-TIME と FF-TIME (MDP) の出力信号のスペクトルをそれぞれ図6.2, 6.3に示す. なお混合過程には Mixture-1 を使用した.

表 6.1: 各手法の表記方法

FF-TIME	時間領域で学習する, 歪みを抑制しない通常の FF-BSS 式 (4.51)-(4.52) を使用して係数を更新 [12].
FF-TIME (MDP)	Minimal Distortion Principle に基づいた信号歪み抑制法 (時間領域で学習を行なう) 式 (6.39) を使用して係数を更新 [18].
FF-FREQ (1)	周波数領域で学習する, 歪みを抑制しない通常の FF-BSS 式 (4.60)-(4.61) を使用して係数を更新 [12],[21],[29].
FF-FREQ (1+PB)	FF-FREQ(1) で分離した信号に対して Projection Back 法 [23] を適用した FF-BSS 分離した信号を式 (6.65) の正規化行列 $D(z)$ で信号歪みを補正. 式 (4.60)-(4.61) で分離回路の係数を更新.
FF-FREQ (1+LPB)	FF-FREQ(1) に Projection Back 法 [23] を組み込んだ FF-BSS 式 (6.77), (6.78) で係数を更新.
FF-FREQ(2)	第 4.2.3 節で紹介した FF-FREQ(1) から余分な制約を除いたアルゴリズム 式 (4.66) 使用して係数を更新 [29].
FF-FREQ (2+PB)	FF-FREQ(2) で分離した信号に対して Projection Back 法 [23] を適用した FF-BSS 分離した信号を式 (6.65) の正規化行列 $D(z)$ で信号歪みを補正. 式 (4.66) で係数を更新.
FF-FREQ (2+LPB)	FF-FREQ(2) に Projection Back 法 [23] を組み込んだ FF-BSS 式 (6.79), (6.80) で係数を更新.

FF-TIME の出力スペクトルは図 6.1 に示す信号歪みの基準とは大きく違っている. 特に高域の信号を増幅しており, 出力のスペクトルを変化させることによって信号の独立性を高めようとしていることが分かる. 一方, 図 6.2 に示されている, FF-TIME (MDP) [18] の出力のスペクトルでは FF-TIME の出力スペクトルに比べ高域のスペクトルを抑え, 図 6.3 に示す信号歪みの基準に近い形になっている. FF-TIME に比べ, 信号歪みが抑えられているのが視覚的に確認できる.

また, 図 6.4-6.9 に周波数領域で学習を行なう手法で学習を行なった際の出力信号のスペクトルを示す.

図 6.4 に示す FF-FREQ (1) の出力信号スペクトルは図 6.3 に示す信号歪みの基準とは大きく異なっている. これは第 4.2.3 節で述べたように学習アルゴリズムに出力信号のパワーを一定にする制約を内包しているため全周波数帯域で大きさが一定になっている.

一方, 図 6.5 に示す FF-FREQ (2) の出力信号スペクトルは FF-FREQ (1) の出力信号スペクトルに比べ, 図 6.3 に示す信号歪みの基準に近い形になっている. これは, FF-FREQ (2) は FF-FREQ (1) のアルゴリズムから出力パワーを一定の値に近付ける制約を除いたものであるため, 周波数スペクトルが平坦化されることが回避された結果であると考えられる. 重み付け効果により信号スペクトルが歪む可能性も考えられるが,

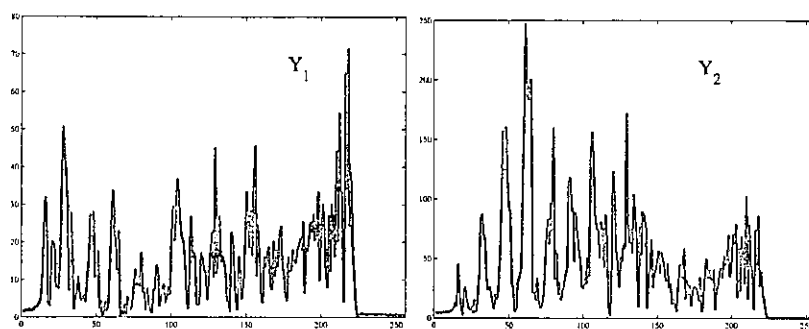


図 6.2: FF-TIME の出力のスペクトル

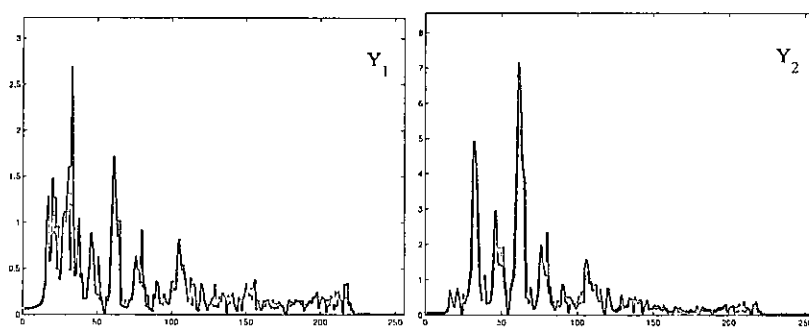


図 6.3: FF-TIME (MDP) の出力信号のスペクトル

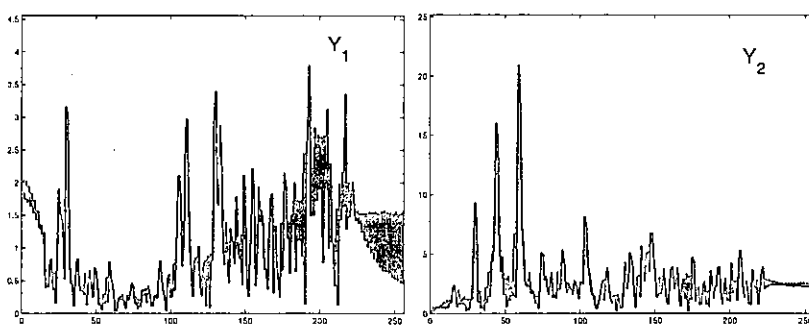


図 6.4: FF-FREQ (1) の出力信号のスペクトル

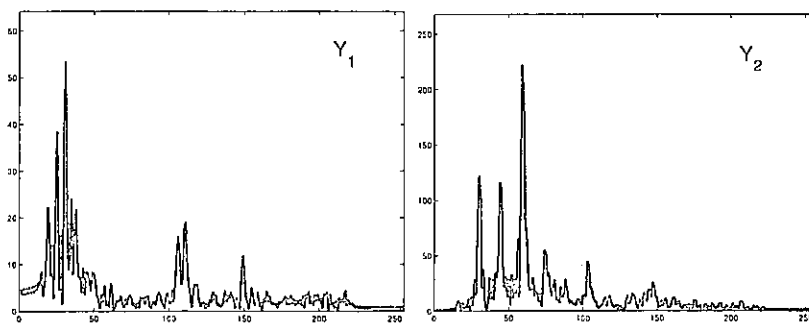


図 6.5: FF-FREQ (2) の出力信号のスペクトル

第5.2.3節でも述べているように信号源が共に音声であるため信号源の周波数分布が似ており、主に信号が分布している低域を強調し、あまり信号が分布していない高域を抑える形になり、分離回路 $W(r)$ の学習は、低域の更新量が大きく、高域の更新量が小さくなるといった偏りが生じ、出力信号が信号源の形に似てくることになり、特に問題はない。逆に収束を早める結果になっていると考えられる。

図6.6にFF-FREQ (1+PB) の出力信号スペクトル、図6.7にFF-FREQ (2+PB) の出力信号スペクトルをそれぞれ示す。

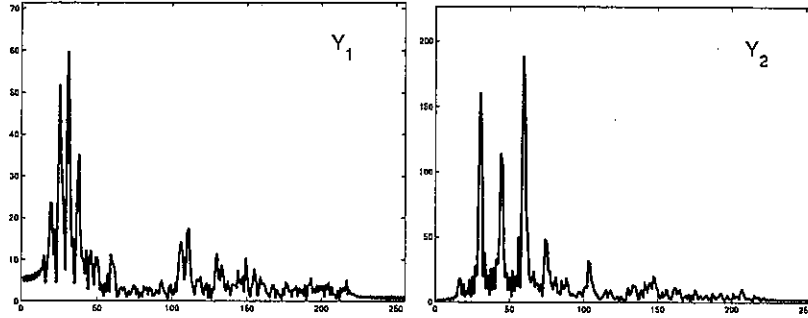


図 6.6: FF-FREQ (1+PB) の出力信号のスペクトル

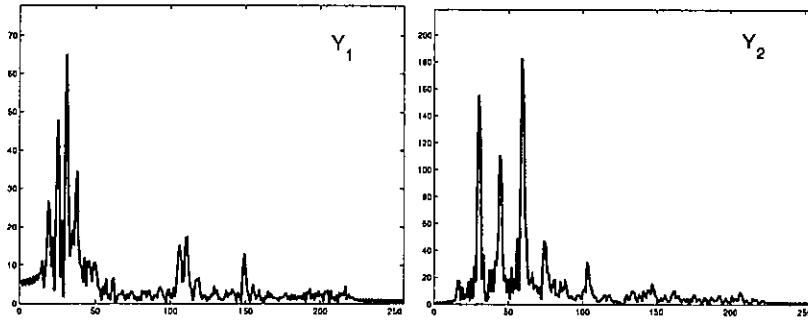


図 6.7: FF-FREQ (2+PB) の出力信号のスペクトル

FF-FREQ (1+PB) の出力信号スペクトル、FF-FREQ (2+PB) の出力信号スペクトルともに図6.3に酷似している。しかしFF-FREQ (1+PB) の出力信号スペクトルに比べ、FF-FREQ (2+PB) の出力信号スペクトルの方が図6.3に示す信号歪みの基準により近い形になっている。これは後述する信号分離の差であると考えられる。FF-FREQ (1+PB)、FF-FREQ (2+PB) は出力信号を正規化行列 $D(z)$ を用いて補正しているだけなので分離の学習には影響を全く及ぼさない。正規化行列 $D(z)$ は完全分離の際に無歪みになるように補正されるため、分離の性能が高いほど信号歪みを抑制する効果が高いと考えられる。そのためFF-FREQ (1+PB)、FF-FREQ (2+PB) の信号分離の性能の差が信号スペクトルに現れたと考えられる。

図6.8にFF-FREQ (1+LPB) の出力信号スペクトル、図6.9にFF-FREQ (2+LPB) の出力信号スペクトルをそれぞれ示す。

FF-FREQ (1+LPB) もFF-FREQ (2+LPB) も同様に図6.3に示す信号歪みの基準と非常に似ている。FF-FREQ (1+LPB) とFF-FREQ (2+LPB) の形を比較してもFF-FREQ (1+PB) とFF-FREQ (2+PB) の形ほどの差は見られない。これは補正が信号

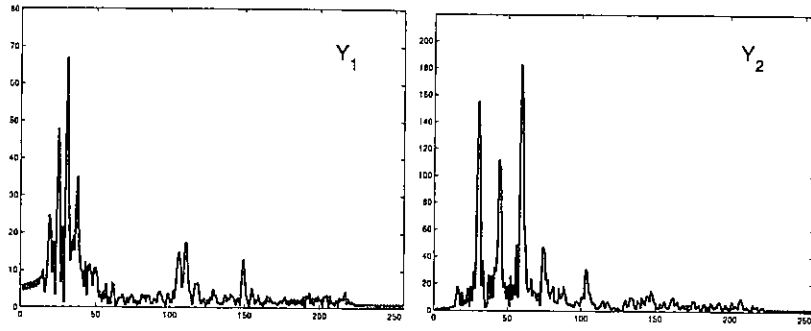


図 6.8: FF-FREQ (1+LPB) の出力信号のスペクトル

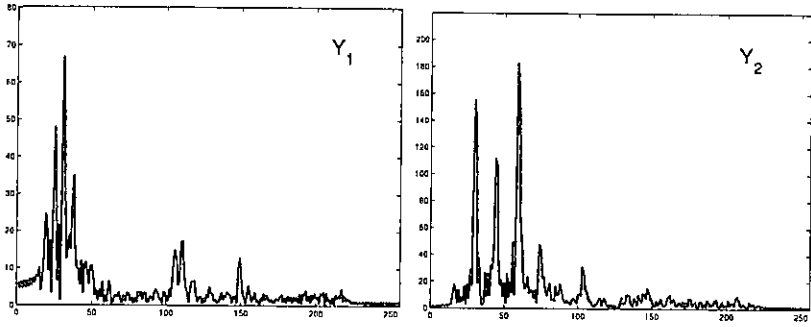


図 6.9: FF-FREQ (2+LPB) の出力信号のスペクトル

分離の学習に組み込まれたため信号分離の差が少なくなったためだと考えられる。信号分離と信号歪みの関係についての考察は後述する。

6.3.3 信号分離と信号歪み

図 6.10, 6.11 に FF-TIME と FF-TIME (MDP) の学習曲線を示す。評価には式 (5.51)-(5.53) で計算される SIR_2 を使用する。

Mixture-1 においても Mixture-2 においても 2 つの学習法の収束速度には大きな違いが見られない。ただし、いずれの混合過程においても FF-TIME の方が最終的な分離性能は優れている。しかしこれは、図 6.10 のように出力信号をより独立にするためにスペクトルを変化させた結果、メインとなる信号を過剰に増幅させた結果であると考えることができる。また、FF-TIME (MDP) は分離性能が低い位置で学習が収束している。これは第 6.1.3 節で述べたように出力信号を観測信号に近付ける制約が分離の妨げになっているためだと考えられる。

また、式 (5.48)-(5.53) で計算される SIR_1 , SIR_2 , 式 (5.3)-(5.10) で表される SD_{1a} , SD_{1b} , SD_{2a} , SD_{2b} を表 6.2 にまとめる。

FF-TIME は Mixture-1 に対しても Mixture-2 に対しても信号歪みの評価 SD_{1a} , SD_{1b} , SD_{2a} , SD_{2b} はすべて悪い。一方、FF-TIME (MDP) では信号歪みが改善されている。しかし、第 6.1.3 節で述べたように出力信号を観測信号に近付ける制約が分離の妨げになっているため信号分離の評価 SIR_1 , SIR_2 は低い。

さらに、FF-TIME (MDP) におけるクロス項の影響について詳しく解析する。第 6.1.3

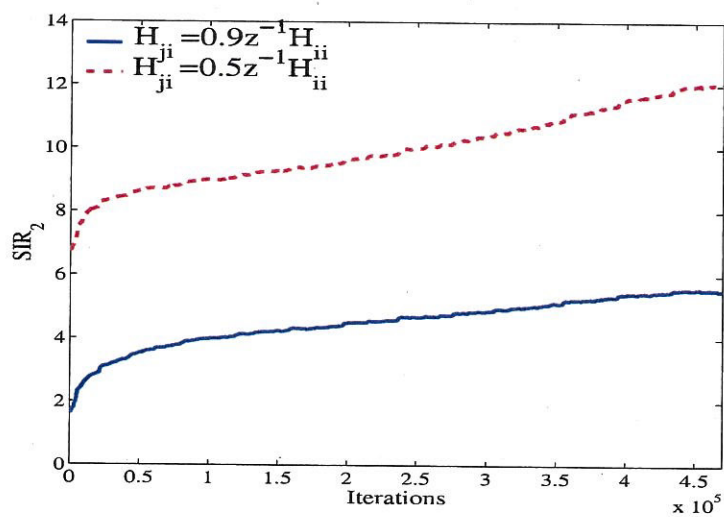
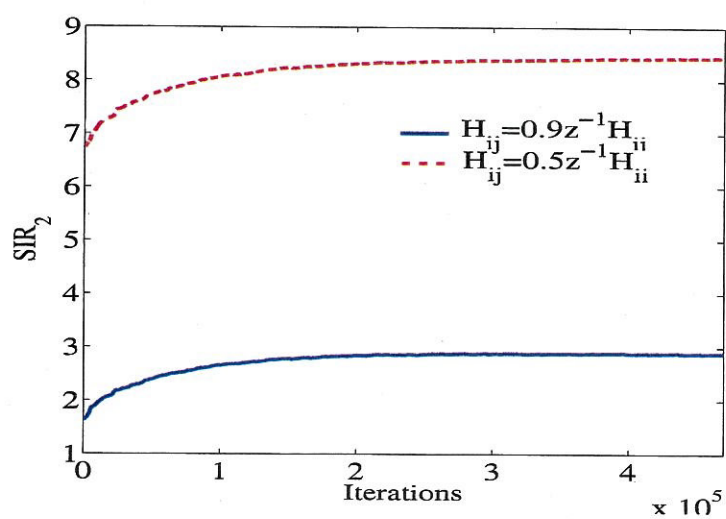
図 6.10: FF-TIME の SIR_1 の時間変化図 6.11: FF-TIME (MDP) の SIR_1 の時間変化

表 6.2: 音声に対する FF-TIME と FF-TIME (MDP) の性能の比較

Methods	Mix	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	1	12.2	5.56	-0.49	-2.71	9.67	8.36
	2	19.3	12.1	-0.52	-2.82	10.4	9.22
FF-TIME (MDP)	1	3.98	2.90	-10.3	-13.6	-8.24	12.3
	2	11.9	8.44	-10.9	-14.0	-9.41	-13.3

節で述べたように FF-TIME (MDP) の学習中, $A_{jj}(z)$ を $H_{jj}(z)$ に近付けようとする一方, $A_{ji}(z)$ を $H_{jj}(z)$ に近付けようとしてしまう. そこで, $A_{ji}(z)$ と $H_{jj}(z)$ の違いを評価する. 図 6.3 は数値データである. このデータは式 (5.3)-(5.10) における i, j, k を $i \neq j = k$ として計算したものである.

表 6.3: $H_{ij}(z)$, $P_{ji}(z)$ 間の相似性

Methods	Mix	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	1	2.44	-0.25	2.03	-0.63
	2	0.17	-0.52	0.37	-1.56
FF-TIME (MDP)	1	-9.36	-10.3	-8.53	-9.39
	2	-5.51	-5.90	-6.88	-7.56

FF-TIME は 0dB 付近の値であるのに対し, FF-TIME (MDP) は非常に小さな値を取っていることが分かる. すなわち, $A_{ji}(z)$ は $H_{jj}(z)$ の形に似ているということであり, 第 6.1.3 節で述べたように余計な制約が働いている. これにより FF-TIME (MDP) の分離性能が低下していると考えられる.

また, 表 6.4 に FF-FREQ (1), FF-FREQ (2), FF-FREQ (1+PB), FF-FREQ (2+PB) の信号分離と信号歪みの性能を示す. 混合過程は Mixture-1 を用いている.

表 6.4: 音声に対する FF-FREQ (1),(2) と FF-FREQ (1+PB), (2+PB) の性能の比較

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-FREQ (1)	7.37	2.61	-6.23	-8.67	-2.82	-2.95
FF-FREQ (1+PB)	15.2	6.59	-23.1	-27.1	-16.9	-18.8
FF-FREQ (2)	13.0	12.9	-9.43	-15.1	-10.9	-13.9
FF-FREQ (2+PB)	14.3	11.6	-16.9	-20.6	-16.9	-18.9

FF-FREQ (1) は信号歪みの評価 SD_{ix} の値が大きく信号歪みが発生している. これは第 4.2.3 節で述べたように学習アルゴリズムに出力信号のパワーを一定にする制約を内包しているため白色化が起こったためだと考えられる. 一方, その白色化を回避するアルゴリズムである FF-FREQ (2) は FF-FREQ (1) に比べ, 信号歪みに関する評価が良くなっている. また, 正規化により信号歪みを低減する FF-FREQ (1+PB), FF-FREQ (2+PB) はさらに信号歪みを抑制しているのがわかる. これにより正規化が信号

歪みに有効に働いていることが確認できる。また、信号分離についてもそれぞれ正規化前のものより同等がそれ以上になっている。しかし、FF-FREQ (1+PB) と FF-FREQ (2+PB) を比較すると FF-FREQ (2+PB) の方が FF-FREQ (1+PB) よりも分離性能が高い。これは出力信号を正規化しているだけであるため分離の学習に影響を与えないため、学習法の差が出たものだと考えられる。一方、信号歪みについては伝達関数を用いて評価をした SD_{2x} については同等であるのに対し、信号で評価した SD_{1x} については FF-FREQ (1+PB) の方が FF-FREQ (2+PB) に比べて良い結果になっている。これについてはさらに検討が必要である。

表 6.5 に FF-FREQ (1+PB), FF-FREQ (2+PB), FF-FREQ (1+LPB), FF-FREQ (2+LPB) の信号分離と信号歪みの比較を示す。

表 6.5: 音声に対する FF-FREQ (1+PB), (2+PB), (1+LPB), (2+LPB) 性能の比較

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-FREQ (1+PB)	15.2	6.59	-23.1	-27.1	-16.9	-18.8
FF-FREQ (1+LPB)	14.2	11.5	-16.8	-20.6	-16.8	-18.9
FF-FREQ (2+PB)	14.3	11.6	-16.9	-20.6	-16.9	-18.9
FF-FREQ (2+LPB)	14.2	11.5	-16.8	-20.6	-16.8	-18.9

表より FF-FREQ (1+LPB), FF-FREQ (2+PB), FF-FREQ (2+LPB) は信号分離、信号歪み共に同等の性能を示している。正規化を分離の学習に組み込むことにより分離の学習に信号歪みの抑制が反映された結果だと考えられる。一方、FF-FREQ (1+PB) は第 4.2.3 節で述べたように学習アルゴリズムに出力信号のパワーを一定にする制約の影響が残るため他の結果と異なっていると考えられる。

第7章 信号歪抑制学習法の提案

本章ではFF-BSSにおいて信号源から出力までの伝達関数から完全分離かつ無歪みの条件を導きだし，学習にその条件を加味して信号歪みを抑制するアルゴリズムを提案する [48]–[52]．そしてシミュレーションにより提案したアルゴリズムの有効性を確認する．

7.1 完全分離かつ無歪みの条件

7.1.1 2チャンネルにおける条件

5.1.2節で述べてように歪みの基準を観測信号の $S_i(z)$ 成分としているため，完全分離ができてかつ無歪みの場合の信号源から出力までの伝達行列 $\mathbf{P}(z)$ (式 (4.22)) は以下のようなになる．

$$\mathbf{P}(z) = \begin{bmatrix} W_{11}(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) \end{bmatrix} \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} \quad (7.1)$$

$$= \begin{bmatrix} W_{11}(z)H_{11}(z) + W_{12}(z)H_{21}(z) & W_{11}(z)H_{12}(z) + W_{12}(z)H_{22}(z) \\ W_{21}(z)H_{11}(z) + W_{22}(z)H_{21}(z) & W_{21}(z)H_{12}(z) + W_{22}(z)H_{22}(z) \end{bmatrix} \quad (7.2)$$

$$= \begin{bmatrix} H_{11}(z) & 0 \\ 0 & H_{22}(z) \end{bmatrix} \quad (7.3)$$

この式の非対角要素が完全分離の条件となり，対角要素が無歪みの条件である．この非対角要素より以下の条件を得る．

$$H_{12}(z) = -\frac{W_{12}(z)}{W_{11}(z)}H_{22}(z) \quad H_{21}(z) = -\frac{W_{21}(z)}{W_{22}(z)}H_{11}(z) \quad (7.4)$$

この式を対角要素に代入することにより完全分離かつ無歪みの条件を得る．

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{22}(z) \quad (7.5)$$

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{11}(z) \quad (7.6)$$

これらの式より

$$W_{11}(z) = W_{22}(z) = W_{ii}(z) \quad (7.7)$$

が簡単に導かれる．したがってこれらの解は

$$W_{ii}^2(z) - W_{ii}(z) - W_{12}(z)W_{21}(z) = 0 \quad (7.8)$$

$$W_{ii} = \frac{1 \pm \sqrt{1 + 4W_{12}(z)W_{21}(z)}}{2} \quad (7.9)$$

と求められる。この式が完全分離かつ無歪みの条件となる。

7.1.2 条件の一般化

3チャンネル以上の場合において完全分離かつ無歪みの条件を考える。2チャンネルの場合と同様に完全分離ができてかつ無歪みの場合の信号源から出力までの伝達行列 $P(z)$ (式(4.22)) は以下の条件を満たす。

$$P(z) = W(z)H(z) = \Lambda(z) \quad (7.10)$$

ただし

$$\Lambda(z) = \text{diag}[H(z)] \quad (7.11)$$

である。

式(7.10)を整理すると以下ようになる。

$$W(z)(\Lambda(z) + \Gamma(z)) = \Lambda(z) \quad (7.12)$$

$$\Gamma(z) = (W^{-1}(z) - I)\Lambda(z) \quad (7.13)$$

ただし Γ は

$$\Gamma(z) = H(z) - \Lambda(z) \quad (7.14)$$

である。

ここで $W(z)$ は出力と観測信号を式(4.15)で関係付けており、もし $W(z)$ の行が従属していた場合、同一の信号源を抽出していることになり分離ができていないことになる。つまり分離ができる場合には $W(z)$ の逆行列は存在する。

$\Gamma(z)$ は式(7.14)で定義される、つまり対角要素が0である行列であるため式(7.13)の対角要素は

$$\text{diag}[\Gamma(z)] = 0 = \text{diag}[(W^{-1}(z) - I)\Lambda(z)] \quad (7.15)$$

となる。

ここで、 $\Lambda(z)$ は対角行列であるため上式は以下のように書き表せる。

$$\text{diag}[(W^{-1}(z) - I)] = 0 \quad (7.16)$$

上式は $W^{-1}(z)$ の対角要素が1であるという条件を示している。

逆行列は一般に

$$W^{-1}(z) = \frac{\text{adj } W(z)}{\det W(z)} \quad (7.17)$$

で表せる。ここで $\text{adj } W(z)$ は $W(z)$ の余因子行列である。また、簡略化のために $\text{adj } W(z)$ について

$$\hat{W}(z) = \text{adj } W(z) \quad (7.18)$$

という表記を用いる。これより以下の関係が導かれる。

$$\frac{\text{diag}[\hat{\mathbf{W}}(z)]}{\det \mathbf{W}(z)} = \mathbf{I} \quad (7.19)$$

$$\frac{\hat{W}_{jj}(z)}{\det \mathbf{W}(z)} = 1 \quad (7.20)$$

$$\hat{W}_{jj}(z) - \det \mathbf{W}(z) = 0 \quad (7.21)$$

ここで $\hat{W}_{jj}(z)$ は $\mathbf{W}(z)$ の余因子である。

余因子は一般に $\mathbf{W}(z)$ の第 i 行、第 j 列を除いてできる小行列 $\mathbf{M}_{ij}(z)$ を用いて

$$\hat{W}_{ij}(z) = (-1)^{i+j} \det \mathbf{M}_{ij}(z) \quad (7.22)$$

と表せる。したがって

$$\hat{W}_{ii}(z) = (-1)^{2i} \det \mathbf{M}_{ii}(z) = \det \mathbf{M}_{ii}(z) \quad (7.23)$$

であり、式 (7.21) は

$$\det \mathbf{M}_{ii}(z) - \det \mathbf{W}(z) = 0 \quad (7.24)$$

と書き直せる。

この式は2チャンネルの場合の条件式 (7.5)-(7.6) と一致しており、この連立方程式を求めることで条件が求まる。しかし、この連立方程式を解くためには非常に大きな計算量が要求されるため近似計算法を考える。

行列式 $\det \mathbf{W}(z)$ は一般に以下のように表記できる。

$$\det \mathbf{W}(z) = \sum_{j=1}^N W_{ij}(z) (-1)^{i+j} \det \mathbf{M}_{ij}(z) \quad (7.25)$$

したがって、式 (7.24) は

$$\det \mathbf{M}_{ii}(z) - \sum_{j=1}^N W_{ij}(z) (-1)^{i+j} \det \mathbf{M}_{ij}(z) \quad (7.26)$$

と書き表せる。この式で $W_{ii}(z)$ を抽出すると

$$(1 - W_{ii}(z)) \det \mathbf{M}_{ii}(z) = \sum_{\substack{j=1 \\ j \neq i}}^N W_{ij}(z) (-1)^{i+j} \det \mathbf{M}_{ij}(z) \quad (7.27)$$

となる。ここで行列式 $\det \mathbf{M}_{ij}(z)$ も式 (7.25) と同様に余因子展開すると

$$\det \mathbf{M}_{ij}(z) = \sum_{\substack{l=1 \\ l \neq i}}^N W_{li}(z) \kappa(i, j, l) \det \mathbf{m}_{li}(z) \quad (7.28)$$

$$\kappa(i, j, l) = \begin{cases} (-1)^{l+i} & (l < j) \\ -(-1)^{l+i} & (l \geq j) \end{cases} \quad (7.29)$$

となる。ここで $m_{kj}(z)$ は $M_{ij}(z)$ から l 行, k 列を除いた小行列である。
したがって、式 (7.27) の右辺は

$$-\sum_{\substack{j=1 \\ j \neq i}}^N W_{ij}(z) \sum_{\substack{l=1 \\ l \neq i}}^N W_{li}(z) \kappa(j, j, l) \det m_{li}(z) \quad (7.30)$$

$$= -\mathbf{w}_{row,i}^T(z) \text{adj } M_{ii}(z) \mathbf{w}_{col,i}(z) \quad (7.31)$$

と変形できる。ただし $\mathbf{w}_{col,i}(z), \mathbf{w}_{row,i}(z)$ は

$$\mathbf{w}_{col,i}(z) = [W_{1i}(z), W_{2i}(z), \dots, W_{yi}(z), \dots, W_{Ni}(z)]^T \quad (y \neq i) \quad (7.32)$$

$$\mathbf{w}_{row,i}(z) = [W_{i1}(z), W_{i2}(z), \dots, W_{iy}(z), \dots, W_{iN}(z)]^T \quad (y \neq i) \quad (7.33)$$

である。

したがって、式 (7.27) 全体は

$$(1 - W_{ii}(z)) \det M_{ii}(z) = -\mathbf{w}_{row,i}^T(z) \text{adj } M_{ii}(z) \mathbf{w}_{col,i}(z) \quad (7.34)$$

$$1 - W_{ii}(z) = -\mathbf{w}_{row,i}^T(z) \frac{\text{adj } M_{ii}(z)}{\det M_{ii}(z)} \mathbf{w}_{col,i}(z) \quad (7.35)$$

$$1 - W_{ii}(z) = -\mathbf{w}_{row,i}^T(z) M_{ii}^{-1}(z) \mathbf{w}_{col,i}(z) \quad (7.36)$$

$$W_{ii}(z) = 1 + \mathbf{w}_{row,i}^T(z) M_{ii}^{-1}(z) \mathbf{w}_{col,i}(z) \quad (7.37)$$

となる。

ここで式 (7.37) において、 $M_{ii}^{-1}(z)$ には $W_{jj}(z)$ ($j \neq i$) を含んでおり、 $W_{ii}(z)$ を陽に解くものではない。つまり、任意の i に対して $W_{ii}(z)$ を求める必要があるのに対し、式 (7.37) は特定の $W_{ii}(z)$ しか求められない。このことに注意する必要がある。

では実際に3チャンネルの場合における条件を考える。

式 (7.24) より

$$\det \mathbf{W}_{jj}(z) = \det \mathbf{W}_{kk}(z) \quad (j \neq k) \quad (7.38)$$

である。式 (7.38) において、 $j = 3, k = 1$ のとき、 $W_{33}(z)$ は

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) = W_{22}(z)W_{33}(z) - W_{23}(z)W_{32}(z) \quad (7.39)$$

$$W_{22}(z)W_{33}(z) = W_{11}(z)W_{22}(z) - (W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z)) \quad (7.40)$$

$$W_{33}(z) = W_{11}(z) - \frac{W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z)}{W_{22}(z)} \quad (7.41)$$

と表せる。また、 $j = 2, k = 3$ の時にも同様の式が成り立つ。

$$W_{11}(z)W_{33}(z) - W_{13}(z)W_{31}(z) = W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) \quad (7.42)$$

そこで上式に式(7.41)の $W_{33}(z)$ を代入し、 $W_{11}(z)$ と $W_{22}(z)$ の関係式を導く.

$$W_{11}(z) \left(W_{11}(z) - \frac{W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z)}{W_{22}(z)} \right) - W_{13}(z)W_{31}(z) = W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) \quad (7.43)$$

$$W_{11}^2(z) - \frac{W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z)}{W_{22}(z)} W_{11}(z) - W_{13}(z)W_{31}(z) = W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) \quad (7.44)$$

$$W_{22}(z)W_{11}^2(z) - (W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z))W_{11}(z) - W_{13}(z)W_{31}(z)W_{22}(z) = W_{11}(z)W_{22}^2(z) - W_{12}(z)W_{21}(z)W_{22}(z) \quad (7.45)$$

$$W_{22}(z)W_{11}^2(z) - (W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z))W_{11}(z) = W_{11}(z)W_{22}^2(z) - (W_{12}(z)W_{21}(z) + W_{13}(z)W_{31}(z))W_{22}(z) \quad (7.46)$$

このように右辺と左辺が対称のである式を導くことができる.

また、式(7.43)より、 $W_{11}(z)$ を $W_{22}(z)$ を用いて表すと、

$$W_{11}^2(z) - \frac{W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z) + W_{22}^2(z)}{W_{22}(z)} W_{11}(z) + W_{12}(z)W_{21}(z) - W_{13}(z)W_{31}(z) = 0 \quad (7.47)$$

$$W_{11}(z) = \frac{W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z) + W_{22}^2(z)}{2W_{22}(z)} \pm \sqrt{\left(\frac{W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z) + W_{22}^2(z)}{W_{22}(z)} \right)^2 - (W_{12}(z)W_{21}(z) - W_{13}(z)W_{31}(z))} \quad (7.48)$$

となる.

ここで、 $j=3$ としたときの式(7.24)は

$$W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z) - (W_{11}(z)W_{22}(z)W_{33}(z) + W_{21}(z)W_{32}(z)W_{13}(z) + W_{12}(z)W_{23}(z)W_{31}(z) - W_{13}(z)W_{22}(z)W_{31}(z) - W_{12}(z)W_{21}(z)W_{33}(z) - W_{11}(z)W_{23}(z)W_{32}(z)) = 0 \quad (7.49)$$

$$(W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z))(1 - W_{33}(z)) + W_{21}(z)W_{32}(z)W_{13}(z) + W_{12}(z)W_{23}(z)W_{31}(z) - (W_{13}(z)W_{31}(z)W_{22}(z) + W_{23}(z)W_{32}(z)W_{11}(z)) = 0 \quad (7.50)$$

であり、そこに式(7.41)を代入する. すると式(7.24)は以下のように変形できる.

$$(W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z))(1 - W_{11}(z) + \frac{W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z)}{W_{22}(z)}) + W_{21}(z)W_{32}(z)W_{13}(z) + W_{12}(z)W_{23}(z)W_{31}(z) - (W_{13}(z)W_{31}(z)W_{22}(z) + W_{23}(z)W_{32}(z)W_{11}(z)) = 0 \quad (7.51)$$

$$-W_{22}(z)W_{11}^2(z) + (W_{22}(z) + 2W_{12}(z)W_{21}(z))W_{11}(z) - W_{12}(z)W_{21}(z) - \frac{W_{12}(z)W_{21}(z)(W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z))}{W_{22}(z)} + W_{12}(z)W_{23}(z)W_{31}(z) - (W_{13}(z)W_{31}(z)W_{22}(z) + W_{23}(z)W_{32}(z)W_{11}(z)) = 0 \quad (7.52)$$

さらに、上式に式 (7.46) を代入する.

$$\begin{aligned} & (W_{22}(z) - W_{22}^2(z) + W_{12}(z)W_{21}(z) + W_{23}(z)W_{32}(z))W_{11}(z) \\ & (W_{22}(z) - 1)W_{12}(z)W_{21}(z) - \frac{W_{12}(z)W_{21}(z)(W_{12}(z)W_{21}(z) - W_{23}(z)W_{32}(z))}{W_{22}(z)} \\ & - (W_{13}(z)W_{31}(z)W_{22}(z) + W_{23}(z)W_{32}(z)W_{11}(z)) = 0 \quad (7.53) \end{aligned}$$

ここで、いくつかの変数を以下のように置き換える.

$$\alpha = W_{12}(z)W_{21}(z) \quad (7.54)$$

$$\beta = W_{23}(z)W_{32}(z) \quad (7.55)$$

$$\gamma = W_{32}(z)W_{21}(z)W_{13}(z) + W_{12}(z)W_{23}(z)W_{31}(z) \quad (7.56)$$

$$\delta = W_{12}(z)W_{21}(z) - W_{32}(z)W_{23}(z) \quad (7.57)$$

$$\varepsilon = W_{12}(z)W_{21}(z) - W_{13}(z)W_{31}(z) \quad (7.58)$$

したがって、式 (7.48), (7.53) はそれぞれ以下のように表せる.

$$W_{11}(z) = \frac{W_{22}^2(z) + \delta}{2W_{22}(z)} \pm \sqrt{\left(\frac{W_{22}^2(z) + \delta}{W_{22}(z)}\right)^2 - \varepsilon} \quad (7.59)$$

$$(W_{22}(z) - W_{22}^2(z) + \alpha + \beta)W_{11}(z) + (W_{22}(z) - 1)\alpha - \frac{\alpha\delta}{W_{22}(z)} - \gamma = 0 \quad (7.60)$$

さらに、式 (7.60) に式 (7.59) を代入し、 $W_{22}(z)$ の多項式で表す.

$$\begin{aligned} & (W_{22}(z) - W_{22}^2(z) + \alpha + \beta) \left(\frac{W_{22}^2(z) + \delta}{2W_{22}(z)} \pm \sqrt{\left(\frac{W_{22}^2(z) + \delta}{W_{22}(z)}\right)^2 - \varepsilon} \right) \\ & + (W_{22}(z) - 1)\alpha - \frac{\alpha\delta}{W_{22}(z)} - \gamma = 0 \quad (7.61) \end{aligned}$$

$$\begin{aligned} & (W_{22}(z) - W_{22}^2(z) + \alpha + \beta) \frac{W_{22}^2(z) + \delta}{2W_{22}(z)} + (W_{22}(z) - 1)\alpha - \frac{\alpha\delta}{W_{22}(z)} - \gamma \\ & = \mp (W_{22}(z) - W_{22}^2(z) + \alpha + \beta) \sqrt{\left(\frac{W_{22}^2(z) + \delta}{W_{22}(z)}\right)^2 - \varepsilon} \quad (7.62) \end{aligned}$$

$$\begin{aligned} & \left((W_{22}(z) - W_{22}^2(z) + \alpha + \beta) \frac{W_{22}^2(z) + \delta}{2W_{22}(z)} + (W_{22}(z) - 1)\alpha - \frac{\alpha\delta}{W_{22}(z)} - \gamma \right)^2 \\ & = \mp (W_{22}(z) - W_{22}^2(z) + \alpha + \beta)^2 \left(\left(\frac{W_{22}^2(z) + \delta}{W_{22}(z)}\right)^2 - \varepsilon \right) \quad (7.63) \end{aligned}$$

$$\begin{aligned}
& (\varepsilon - \alpha)W_{22}^6(z) + (2\alpha + \gamma + 2\varepsilon)W_{22}^5(z) \\
& + (2\alpha^2 + \alpha\beta - \alpha - \gamma + (1 - 2(\alpha + \beta))\varepsilon)W_{22}^4(z) \\
& - (2\alpha^2 + 2\alpha\gamma - 2(\alpha + \beta)\varepsilon + (\alpha + \gamma)(\alpha + \beta - \delta))W_{22}^3(z) \\
& + ((\alpha + \gamma)^2 - (\alpha + \gamma)\delta - 2\alpha^2\delta + \alpha\delta^2 + (\alpha + \beta)^2\varepsilon)W_{22}^2(z) \\
& + (2(\alpha^2\delta + \alpha\gamma\delta) - (\alpha\delta^2 + \alpha(\alpha + \beta)\delta) + (\alpha + \beta)\gamma\delta)W_{22}(z) \\
& - \alpha\beta\delta^2 = 0 \quad (7.64)
\end{aligned}$$

$$\begin{aligned}
& (\varepsilon - \alpha)W_{22}^6(z) + (2(\alpha - \varepsilon) + \gamma)W_{22}^5(z) \\
& + 2(\alpha + \beta - \frac{1}{2})(\alpha - \varepsilon) - \alpha\beta - \gamma)W_{22}^4(z) - 2(\alpha - \varepsilon + \gamma)(\alpha + \beta)W_{22}^3(z) \\
& + ((\alpha + \beta + \gamma)\gamma - (\alpha + \beta)^2(\alpha - \varepsilon) + 2(\alpha + \beta + \frac{1}{2})\alpha\beta)W_{22}^2(z) \\
& + \gamma\delta^2W_{22}(z) - \alpha\beta\delta^2 = 0 \quad (7.65)
\end{aligned}$$

ここで簡単化のため、以下の変数を定義する。

$$\alpha - \varepsilon = W_{13}(z)W_{31}(z) = \lambda \quad (7.66)$$

$$\alpha + \beta = W_{12}(z)W_{21}(z) + W_{23}(z)W_{32}(z) = \nu \quad (7.67)$$

これらにより、式 (7.65) は以下のように表せる。

$$\begin{aligned}
& -\lambda W_{22}^6(z) + (2\lambda + \gamma)W_{22}^5(z) + 2(\nu - \frac{1}{2})\lambda - \alpha\beta - \gamma)W_{22}^4(z) - 2(\lambda + \gamma)\nu W_{22}^3(z) \\
& + ((\nu + \gamma)\gamma - \nu^2\lambda + 2(\nu + \frac{1}{2})\alpha\beta)W_{22}^2(z) + \gamma\delta^2W_{22}(z) - \alpha\beta\delta^2 = 0
\end{aligned}$$

したがって、制約の条件である $W_{11}(z)$, $W_{22}(z)$, $W_{33}(z)$ は式 (7.68), (7.48), (7.41) の多項式を順次解いていくことで求めることができる。

一方、式 (7.37) の近似法を使用した場合、

$$W_{33}(z) = 1 + \begin{bmatrix} W_{31}(z) & W_{32}(z) \end{bmatrix} \begin{bmatrix} W_{11}(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) \end{bmatrix}^{-1} \begin{bmatrix} W_{13}(z) \\ W_{23}(z) \end{bmatrix} \quad (7.68)$$

$$\begin{aligned}
& = 1 + \frac{\begin{bmatrix} W_{31}(z) & W_{32}(z) \end{bmatrix} \begin{bmatrix} W_{22}(z) & -W_{12}(z) \\ -W_{21}(z) & W_{11}(z) \end{bmatrix} \begin{bmatrix} W_{13}(z) \\ W_{23}(z) \end{bmatrix}}{W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z)} \quad (7.69)
\end{aligned}$$

$$\begin{aligned}
& = 1 + \frac{1}{W_{11}(z)W_{22}(z) - W_{12}(z)W_{21}(z)} (W_{22}(z)W_{13}(z)W_{31}(z) - W_{12}(z)W_{23}(z)W_{31}(z) \\
& \quad + W_{11}(z)W_{23}(z)W_{32}(z) - W_{21}(z)W_{13}(z)W_{32}(z)) \quad (7.70)
\end{aligned}$$

となる。同様に $W_{22}(z)$, $W_{11}(z)$ を求めると

$$\begin{aligned}
W_{22}(z) & = 1 + \frac{1}{W_{11}(z)W_{33}(z) - W_{13}(z)W_{31}(z)} (W_{33}(z)W_{12}(z)W_{21}(z) - W_{13}(z)W_{32}(z)W_{21}(z) \\
& \quad + W_{11}(z)W_{32}(z)W_{23}(z) - W_{31}(z)W_{12}(z)W_{23}(z)) \quad (7.71)
\end{aligned}$$

$$\begin{aligned}
W_{11}(z) & = 1 + \frac{1}{W_{22}(z)W_{33}(z) - W_{23}(z)W_{32}(z)} (W_{33}(z)W_{21}(z)W_{12}(z) - W_{23}(z)W_{31}(z)W_{12}(z) \\
& \quad + W_{22}(z)W_{31}(z)W_{13}(z) - W_{32}(z)W_{21}(z)W_{13}(z)) \quad (7.72)
\end{aligned}$$

となり、式の導出の簡単化され、計算量が大きく削減されることがわかる。

7.1.3 順序問題

これまでは対角要素に出力が取り出される場合についてのみ考えてきたがBSSではどの出力にどの信号源成分が取り出されてもよく、 $\mathbf{P}(z)$ は各行各列に1つだけ0でない要素を持てばよい。2チャンネルの場合、

$$\mathbf{P}(z) = \begin{bmatrix} 0 & H_{12}(z) \\ H_{21}(z) & 0 \end{bmatrix} \quad (7.73)$$

となっても歪みなく信号が分離されたとされる。この場合以下の列の入れ替えを行ない0でない要素を対角要素に移動させる。

$$\begin{bmatrix} W_{11}(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) \end{bmatrix} \begin{bmatrix} H_{11}(z) & H_{12}(z) \\ H_{21}(z) & H_{22}(z) \end{bmatrix} = \begin{bmatrix} 0 & H_{12}(z) \\ H_{21}(z) & 0 \end{bmatrix} \quad (7.74)$$

$$\begin{bmatrix} W_{11}(z) & W_{12}(z) \\ W_{21}(z) & W_{22}(z) \end{bmatrix} \begin{bmatrix} H_{12}(z) & H_{11}(z) \\ H_{22}(z) & H_{21}(z) \end{bmatrix} = \begin{bmatrix} H_{12}(z) & 0 \\ 0 & H_{21}(z) \end{bmatrix} \quad (7.75)$$

$$\mathbf{W}(z)\mathbf{H}(z) = \mathbf{\Lambda}(z) \quad (7.76)$$

この $\mathbf{H}(z)$ 、 $\mathbf{\Lambda}(z)$ を使用して条件を導き出せばよい。nチャンネルの条件が $n!$ 次の多項式になるのは各行各列に0でない要素が存在するパターンすべてをこの式で扱っているためである。

7.2 歪み抑制学習法

7.2.1 学習の必要性

前節の条件は式 (7.13) の対角要素のみから求めている。つまり非対角要素は考慮されていない。非対角要素は

$$H_{ij}(z) = H_{ii}(z) \frac{\det \mathbf{W}_{ij}(z)}{\det \mathbf{W}(z)} \quad (7.77)$$

となり $H_{ij}(z)$, $H_{ii}(z)$ を考慮しなければならない。混合過程の伝達関数 $H_{ij}(z)$, $H_{ii}(z)$ は不明であるのでなんらかの学習で求める必要がある。

式 (7.9), (7.37) の制約は $\mathbf{W}(z)$ の対角要素を非対角要素で表しており, $\mathbf{W}(z)$ の対角要素と $\mathbf{W}(z)$ の非対角要素の関係式であると考えることができる。つまり, $\mathbf{W}(z)$ の非対角要素が決定されれば自動的にそれに対する $\mathbf{W}(z)$ の対角要素が求まる。

ここでもう一度 FF-BSS における完全分離の条件について考える。信号源 $\mathbf{S}(z)$ から出力 $\mathbf{Y}(z)$ までの伝達行列 $\mathbf{P}(z)$ が各行各列に 1 つだけ 0 でない要素を持つとき完全分離が可能となる。その際, 0 でない要素がどのような値でも分離はできているとする。そのため完全分離の条件は式 (4.22) より

$$\sum_{k=1}^N W_{jk}(z) H_{ki}(z) = 0 \quad (7.78)$$

となる。ただし, $i = 1, 2, \dots, N$, $j = 1, 2, \dots, N$ であり $i \neq j$ である。ここでは $Y_i(z)$ にそれぞれ $S_i(z)$ が取り出されるとした。方程式の数は $N^2 - N$ に対し, 変数にあたる分離回路 $W_{jk}(z)$ の数は N^2 であるため, 完全分離を行ってもその解には自由度が存在していた。ここで, 完全分離かつ無歪みの条件が $\mathbf{W}(z)$ の対角要素と $\mathbf{W}(z)$ の非対角要素の関係式であるを利用すれば変数を減らすことが可能である。 $\mathbf{W}(z)$ の対角要素の数 N 個なので変数の数を $N^2 - N$ 個にすることができる。つまり, 方程式の本数と変数の数を一致させることができる。また, 完全分離かつ無歪みの条件であるので必ず完全分離かつ無歪みの点を通る。したがって, 分離の学習に完全分離かつ無歪みの条件を付加させればよい。

7.2.2 周波数領域での実装

式 (4.48), (4.49) で与えられる従来の学習法は式 (7.9) のような完全分離かつ無歪みの条件を満たさず, 式 (7.3), (7.10) の非対角要素のみを近似的に満たす。対角要素に関しては保証されない。したがって一般的には信号歪みを抑制することはできない。

そこで学習を式 (7.9) のような完全分離かつ無歪みの制約をつけて行なう方式を提案する。

完全分離かつ無歪みの制約は時間領域と周波数領域のどちらでも実装が可能である。まず, 周波数領域での実装について述べる。

基本的な考え方は以下ようになる。まず, $W_{jj}(z)$ と $W_{jk}(z)$ を信号分離のため従来の方式で更新する。そして, $W_{jj}(z)$ を式 (7.9) で与えられる制約にしたがって更新する。

さらに学習初期の段階では $W_{jk}(z)$ は十分に収束していないため、式 (7.9) の条件は満たさない。そのため制約を徐々に効かせていく。

学習アルゴリズムは以下のように表せる。

Step 1 : 式 (4.60), または (4.66) で $W_{jj}(r, m)$, $W_{jk}(r, m)$ を更新し, それぞれ $W_{jj}(r+1, m)$, $W_{jk}(r+1, m)$ とする。

Step 2 : $W_{jj}(r+1, m)$ を以下の式で補正する:

$$W_{jj}(r+1, m) = (1 - \alpha)W_{jj}(r+1, m) + \alpha \frac{1 + \sqrt{1 + 4W_{12}(r, m)W_{21}(r, m)}}{2} \quad (7.79)$$

$(0 < \alpha \leq 1)$

ただし, 左辺の $W_{jj}(r+1, m)$ は補正後の値である。

n チャンネルの場合には n 次多項式を毎回解くことになる。しかし計算量が非常に大きくなってしまう。式 (4.60) または式 (4.66), (7.79) のステップサイズ η, α は非常に小さな値で時刻 n と $n+1$ で w_{ij} の値が大きく変化しないことから式 (7.37) を使って制限を掛ける。したがって, 学習アルゴリズムは以下のように書き表せる。

Step 1 : 式 (4.60), または (4.66) で $W_{jj}(r, m)$, $W_{jk}(r, m)$ を更新し, それぞれ $W_{jj}(r+1, m)$, $W_{jk}(r+1, m)$ とする。

Step 2 : $W_{jj}(r+1, m)$ を以下の式で補正する:

$$W_{jj}(r+1, m) = (1 - \alpha)W_{jj}(r+1, m) + \alpha \tilde{W}_{jj}(r, m) \quad (7.80)$$

$(0 < \alpha \leq 1)$

ただし, $\tilde{W}_{jj}(r, m)$ は式 (7.37) で計算された $W_{jj}(z)$, 左辺の $W_{jj}(r+1, m)$ は補正後の値である。

7.2.3 時間領域での実装

次に, 時間領域での実装について述べる。

時間領域では式 (7.9) を直接計算することができない。そこで時間領域で計算できるように変形する必要がある。

式 (7.9) は以下のように書き直すことができる。

$$W_{jj}(z)^2 - W_{jj}(z) = W_{12}(z)W_{21} \quad (7.81)$$

この式を学習中に考慮する。まず $W_{12}(z), W_{21}(z)$ を与え, 式 (7.81) の関係を満たすように $W_{ii}(z)$ を決定する。

現在考慮しているシステムは線形システムであるので $W_{ii}(z), W_{12}(z), W_{21}(z)$ はそれぞれ

れ以下のように表すことができる.

$$W_{ii}(z) = \sum_{d_{ii}=-\infty}^{\infty} w_{ii}(d_{ii})z^{-d_{ii}} \quad (7.82)$$

$$W_{12}(z) = \sum_{d_{12}=-\infty}^{\infty} w_{12}(d_{12})z^{-d_{12}} \quad (7.83)$$

$$W_{21}(z) = \sum_{d_{21}=-\infty}^{\infty} w_{21}(d_{21})z^{-d_{21}} \quad (7.84)$$

ただし $w_{ii}(-1), \dots, w_{ii}(-\infty), w_{12}(-1), \dots, w_{12}(-\infty), w_{21}(-1), \dots, w_{21}(-\infty)$ は全て 0 とみなす.

したがって, 式 (7.81) の右辺は以下のように書くことができる.

$$(\text{右辺}) = \sum_{d_{12}=0}^{\infty} w_{12}(d_{12})z^{-d_{12}} \sum_{d_{21}=0}^{\infty} w_{21}(d_{21})z^{-d_{21}} \quad (7.85)$$

$$= \sum_{d_{12}=0}^{\infty} \sum_{d_{21}=0}^{\infty} w_{12}(d_{12})w_{21}(d_{21})z^{-(d_{12}+d_{21})} \quad (7.86)$$

$$= \sum_{d=0}^{\infty} \sum_{d_{12}=0}^d w_{12}(d_{12})w_{21}(d-d_{12})z^{-d} \quad (7.87)$$

また左辺は

$$(\text{左辺}) = \sum_{d_{ii1}=0}^{\infty} \sum_{d_{ii2}=0}^{\infty} w_{ii}(d_{ii1})w_{ii}(d_{ii2})z^{-(d_{ii1}+d_{ii2})} - \sum_{d=0}^{\infty} w_{ii}(d)z^{-d} \quad (7.88)$$

$$= \sum_{d=0}^{\infty} \sum_{d_{ii}=0}^{\infty} w_{ii}(d_{ii})w_{ii}(d-d_{ii})z^{-d} - \sum_{d=0}^{\infty} w_{ii}(d)z^{-d} \quad (7.89)$$

$$= w_{ii}(0)^2 - w_{ii}(0) + \sum_{d=1}^{\infty} \left(2w_{ii}(0)w_{ii}(d) - w_{ii}(d) + \sum_{d_{ii}=1}^{d-1} w_{ii}(d_{ii})w_{ii}(d-d_{ii}) \right) z^{-d} \quad (7.90)$$

と表すことができる. これらはまた z の多項式となっている. そこで z^{-d} の係数ごとに係数比較を行ない $w_{ii}(d_{ii})$ の値を求めていく. すなわち

$d=0$ のとき

$$w_{ii}(0)^2 - w_{ii}(0) = w_{12}(0)w_{21}(0) \quad (7.91)$$

$d>0$ のとき

$$2w_{ii}(0)w_{ii}(d) - w_{ii}(d) + \sum_{d_{ii}=1}^{d-1} w_{ii}(d_{ii})w_{ii}(d-d_{ii}) = \sum_{d_{12}=0}^d w_{12}(d_{12})w_{21}(d-d_{12}) \quad (7.92)$$

を $w_{ii}(d)$ について解く.

式 (7.91) の解は

$$w_{ii}(0) = \frac{1 \pm \sqrt{1 + w_{12}(0)w_{21}(0)}}{2} \quad (7.93)$$

である. また式 (7.92) の解は

$$w_{ii}(d) = \frac{\sum_{d_{12}=0}^d w_{12}(d_{12})w_{21}(d - d_{12}) - \sum_{d_{ii}=1}^{d-1} w_{ii}(d_{ii})w_{ii}(d - d_{ii})}{2w_{ii}(0) - 1} \quad (7.94)$$

である. 右辺の分子の第2項には $w_{ii}(d_{ii}), w_{ii}(d - d_{ii})$ が含まれるが, すべての $d_{ii}, d - d_{ii}$ に対して $d > d_{ii}, d > d - d_{ii}$ が成り立つため低次項から比較を行えば計算が可能である.

前述の通り, 完全分離かつ無歪みの式から信号歪みなしの条件を導いたが, 学習の初期値は適当な値に設定しているため学習初期の段階では信号の分離ができていない. つまり条件における完全分離が満たされておらず制約がうまく機能しない. そこで式 (7.81) の制約を徐々に適用する. 学習アルゴリズムは以下のように表せる.

Step 1 : 式 (4.51) と (4.52) で $w_{jj}(n, l)$ や $w_{jk}(n, l)$ を更新し, それぞれ $w_{jj}(n + 1, l)$, $w_{jk}(n + 1, l)$ とする.

Step 2 : $w_{jj}(n, l)$ を以下の式で補正する. :

$$w_{jj}(n + 1, l) = (1 - \alpha)w_{jj}(n + 1, l) + \alpha\tilde{w}_{jj}(n, l) \quad (7.95)$$

$$(0 < \alpha \leq 1)$$

ただし, $\tilde{w}_{ii}(n + 1)$ は式 (7.93), 式 (7.94) から計算された $w_{ii}(n + 1, l)$ であり, 左辺の $w_{jj}(n + 1, l)$ は補正後の値である.

周波数領域での実装時と同様に n チャンネルの場合には n 次多項式を毎回解くことになる. 係数比較などを行なった場合でも計算量が非常に大きくなってしまう. 式 (4.51), (7.95) のステップサイズ η, α は非常に小さな値で時刻 n と $n + 1$ で w_{ij} の値が大きく変化しないことから時間領域でも式 (7.37) を使って制限を掛ける. すなわち, 学習アルゴリズムは以下のようにする.

Step 1 : 式 (4.51) と (4.52) で $w_{jj}(n, l)$ や $w_{jk}(n, l)$ を更新し, それぞれ $w_{jj}(n + 1, l)$, $w_{jk}(n + 1, l)$ とする.

Step 2 : $w_{jj}(n, l)$ を以下の式で補正する. :

$$w_{jj}(n + 1, l) = (1 - \alpha)w_{jj}(n + 1, l) + \alpha\tilde{w}_{jj}(n, l) \quad (7.96)$$

$$(0 < \alpha \leq 1)$$

ただし, $\tilde{w}_{ii}(n + 1)$ は式 (7.93), 式 (7.37) から計算された $w_{ii}(n + 1, l)$ であり, 左辺の $w_{jj}(n + 1, l)$ は補正後の値である.

ただし, 時間領域では式 (7.37) を直接計算することは難しい. したがって, 一旦周波数領域に変換し, 式 (7.37) を計算した後, 逆変換して求めるなど工夫が必要になる.

7.3 シミュレーション

7.3.1 表記法

シミュレーションでさまざまな手法を比較する．そのため，各手法を表すのに簡略化した表記方法をとる．表 7.1 にその表記方法を示す．

表 7.1: 各手法の表記方法

FF-TIME	時間領域で学習する，歪みを抑制しない通常の FF-BSS 式 (4.51)-(4.52) を使用して係数を更新 [12].
FF-TIME (DF)	時間領域で学習する，本章で提案した信号歪み抑制学習法 式 (4.51)-(4.52) で係数を更新し，式 (7.95) で信号歪みを補正.
FF-TIME (ADF)	時間領域で学習する，本章で提案した近似法を用いた信号歪み抑制学習法 式 (4.51)-(4.52) で係数を更新し，式 (7.96) で信号歪みを補正.
FF-FREQ (1)	周波数領域で学習する，歪みを抑制しない通常の FF-BSS 式 (4.60)-(4.61) を使用して係数を更新 [12],[21],[29].
FF-FREQ (1+DF)	FF-FREQ(1) に本章で提案した信号歪み抑制の制約を付加した方式 式 (4.60)-(4.61) で係数を更新し，式 (7.79) で信号歪みを補正.
FF-FREQ (1+ADF)	FF-FREQ(1) に本章で提案した近似法を用いた信号歪み抑制の制約を付加した方式 式 (4.60)-(4.61) で係数を更新し，式 (7.80) 信号歪みを補正.
FF-FREQ(2)	第 4.2.3 節で紹介した FF-FREQ(1) から余分な制約を除いたアルゴリズム 式 (4.66) 使用して係数を更新 [29].
FF-FREQ (2+DF)	FF-FREQ(2) に本章で提案した信号歪み抑制の制約を付加した方式 式 (4.66) で係数を更新し，式 (7.79) 信号歪みを補正.
FF-FREQ (2+ADF)	FF-FREQ(2) に本章で提案した近似法を用いた信号歪み抑制の制約を付加した方式 式 (4.66) で係数を更新し，式 (7.80) 信号歪みを補正.

7.3.2 シミュレーション条件

シミュレーションの条件は 5.4 節と同様に 2 信号源，2 センサで行ない混合過程の伝達特性には図 5.5 の $H_{11}(z)$ と $H_{22}(z)$ ，

(Mixture-1)

$$H_{21}(z) = 0.9z^{-1}H_{11}(z) \quad (7.97)$$

$$H_{12}(z) = 0.9z^{-1}H_{22}(z) \quad (7.98)$$

(Mixture-2)

$$H_{21}(z) = 0.5z^{-1}H_{11}(z) \quad (7.99)$$

$$H_{12}(z) = 0.5z^{-1}H_{22}(z) \quad (7.100)$$

となる $H_{12}(z)$ と $H_{21}(z)$ を用いる.

信号源は以下の男性と女性の音声を用いている.

- s_1 : 音声 「人間は豊かな自然を求めています。」(男性)
「パンはボルトガル人が伝えたものです。」(男性)
 s_2 : 音声 「健康のため無理をしないでください。」(女性)
「太宰治は青森県で生まれました。」(女性)

タップ数は 256, FFT サイズは 256, 分離回路の初期値は $W_{11}(z) = W_{22}(z) = 1$, $W_{12}(z) = W_{21}(z) = 0$ である.

ステップサイズは FF-TIME は 0.000007, FF-TIME (DF) は 0.0000018 とする. また, FF-FREQ (1), FF-FREQ(2), FF-FREQ (1+DF), FF-FREQ(2+DF) は 0.0000004 とした.

7.3.3 信号スペクトル

信号歪みの基準となる $H_{ii}(z)S_i(z)$ を図 7.1 に示す.

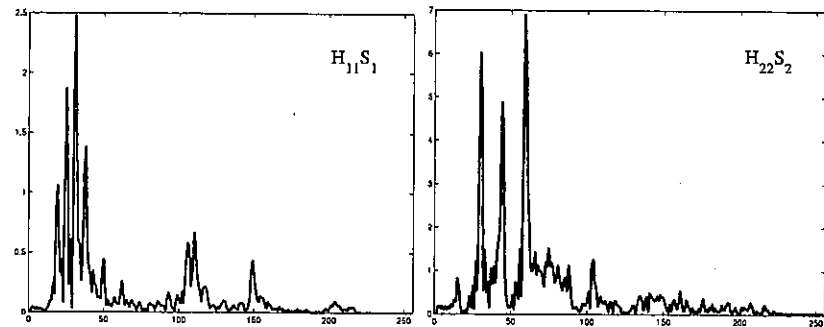


図 7.1: $H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル

また, FF-TIME と提案する信号歪み抑制学習法である FF-TIME (DF) の出力信号のスペクトルをそれぞれ図 7.2, 7.3 に示す. なお混合過程には Mixture-2 を使用した.

FF-TIME の出力スペクトルは図 7.1 に示す信号歪みの基準とは大きく違っている. 特に高域の信号を増幅しており, 出力のスペクトルを変化させることによって信号の

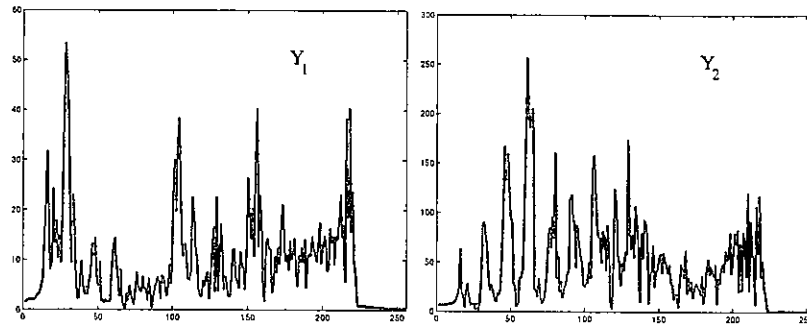


図 7.2: FF-TIME の出力のスペクトル

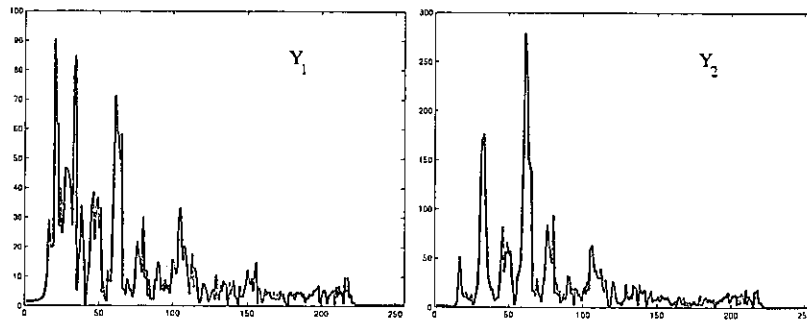


図 7.3: FF-TIME (DF) の出力のスペクトル

独立性を高めようとしていることが分かる．一方，図 7.3 に示されている，提案する信号歪抑制学習法 FF-TIME (DF) の出力のスペクトルでは FF-TIME の出力スペクトルのような変化は見られず，図 7.1 に示す信号歪みの基準と酷似している．これらの図から FF-TIME (DF) では信号歪みが十分に抑制されていることが視覚的に確認できる．

7.3.4 信号分離と信号歪み

図 7.4, 7.5 に学習曲線を示す．評価には式 (5.51)-(5.53) で計算される SIR_2 で行なう．

Mixture-1 においても Mixture-2 においても 2 つの学習法の収束速度には大きな違いが見られない．ただしいずれの混合過程においても FF-TIME の方が最終的な分離性能は優れている．しかしこれは，図 7.2 のように出力信号をより独立にするためにスペクトルを変化させた結果，メインとなる信号を過剰に増幅させた結果であると考えられることができる．

また，式 (5.48)-(5.53) で計算される SIR_1 , SIR_2 ，式 (5.3)-(5.10) で表される SD_{1a} , SD_{1b} , SD_{2a} , SD_{2b} を表 7.2 にまとめる．

信号分離の評価 SIR_1 , SIR_2 は，FF-TIME の方が良い．しかし，すべての信号歪みに関する評価において，FF-TIME はかなり悪い結果になっている．一方，提案した信号歪抑制学習法 FF-TIME (DF) では信号歪みが大きく改善されていることが分かる．

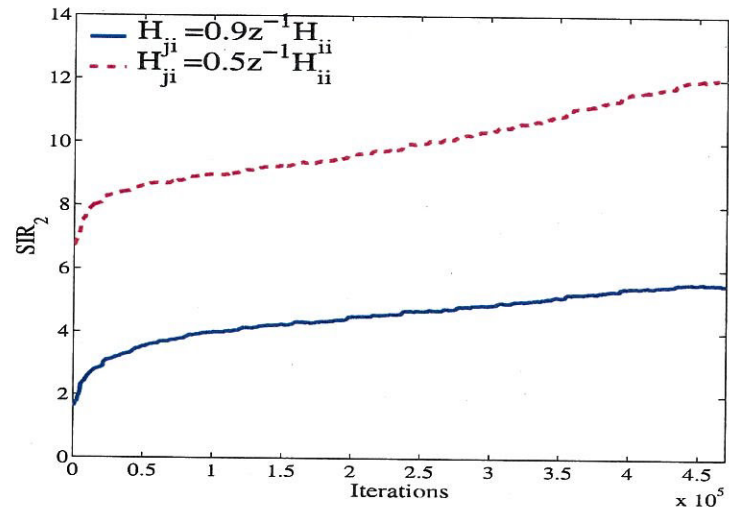
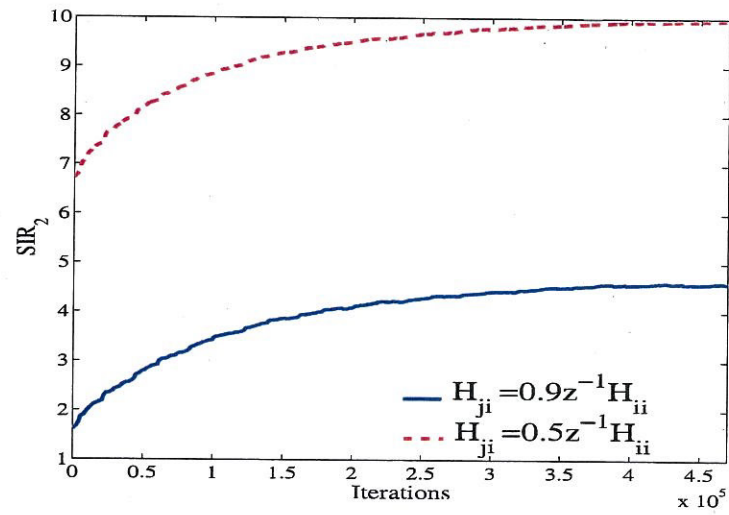
図 7.4: FF-TIME の SIR_1 の時間変化図 7.5: FF-TIME (DF) の SIR_1 の時間変化

表 7.2: 歪み抑制しない方式と提案する信号歪み抑制法の比較

Methods	Mix	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	1	12.2	5.56	-0.49	-2.71	9.67	8.36
	2	19.3	12.1	-0.52	-2.82	10.4	9.22
FF-TIME (DF)	1	8.33	4.33	-6.20	-10.0	-11.1	-15.9
	2	15.6	9.98	-13.8	-16.8	-19.2	-22.8
FF-FREQ (1)	1	7.24	2.55	-7.71	-10.2	-6.41	-6.70
	2	16.2	9.47	-8.92	-10.7	-7.33	-7.38
FF-FREQ (1+DF)	1	9.68	6.38	-13.5	-18.1	-15.1	-18.3
	2	18.6	14.1	-28.0	-31.7	-25.1	-28.4
FF-FREQ (2)	1	18.8	9.68	-10.9	-16.5	-11.9	-15.0
	2	28.3	17.5	-21.3	-27.2	-20.7	-24.3
FF-FREQ (2+DF)	1	18.7	10.1	-25.8	-30.0	-18.4	-21.2
	2	28.3	17.3	-38.9	-40.2	-29.2	-33.7

7.3.5 近似法の有効性

計算量を考慮すると信号歪み抑制の制約には第7.1.2節の式(7.37)の近似法を使用することになる。しかし、第7.2.2節、7.2.3節で述べたように式(7.37)の $M_{ii}^{-1}(z)$ には $W_{jj}(z)$ ($j \neq i$)を含んでおり、 $W_{ii}(z)$ を陽に解くものではないが、ステップサイズ η, α は非常に小さな値で時刻 n と $n+1$ で w_{ij} の値が大きく変化しないため w_{ij} を前のサンプルのものをを使うという意味で近似を行なっている。この近似の影響を調べる。表7.3に2チャンネルのBSSにおいて、近似法を使用したものと厳密に制約を計算した方法での比較を示している。

表 7.3: 近似法による影響の比較

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME (DF)	8.33	4.33	-12.1	-16.2	-15.4	-19.9
FF-TIME (ADF)	8.31	4.33	-12.5	-16.7	-15.4	-19.8
FF-FREQ (1+DF)	9.68	6.38	-13.5	-18.1	-15.1	-18.3
FF-FREQ (1+ADF)	9.68	6.37	-13.5	-18.1	-15.1	-18.3
FF-FREQ (2+DF)	18.7	10.1	-25.8	-30.0	-18.4	-21.1
FF-FREQ (2+ADF)	18.7	10.1	-25.8	-30.0	-18.4	-21.2

この表から、信号歪み抑制の制約に近似を使った方式、FF-TIME (ADF), FF-FREQ (1+ADF), FF-FREQ (2+ADF) は信号分離、信号歪み共にそれぞれの近似法を使用しない方法と同等の性能が得られている。前述の通り、FF-TIME (ADF), FF-FREQ (1+ADF), FF-FREQ (2+ADF) は式Eq(7.37)における $M_{ii}^{-1}(z)$ で用いられる $W_{jj}(z)$ を定数として計算している。このシミュレーション結果は近似法がうまく機能していることを示している。

第8章 各手法の比較

前章までに時間領域、周波数領域それぞれで学習を行なう、さまざまな信号歪みに関する手法を紹介、提案してきた。FB-BSS は分離が進むにつれて信号歪みが抑制される構造であることも既に述べている。

本章では、それらの手法を信号分離、信号歪みの観点から比較することにより、それぞれの手法の特徴を解析するとともに、提案した手法の有効性を示す。

8.1 表記法

本章では多くの手法を比較する。それらの手法の表記方法を表 8.1 にまとめる。

8.2 シミュレーション

8.2.1 シミュレーション条件

シミュレーションは2信号源, 2センサ, または3信号源, 3センサで行なう。

2チャンネル

2チャンネルで使用する混合過程のインパルス応答を図8.1に示す.

信号源は白色信号、音声と定常有色信号を用いる。

音声は以下の男性と女性の音声を繰り返し用いている。

$s_1(n)$: 音声 「人間は豊かな自然を求めています。」(男性)
「パンはポルトガル人が伝えたものです。」(男性)

$s_2(n)$: 音声 「健康のため無理をしないでください。」(女性)
「太宰治は青森県で生まれました。」(女性)

定常有色信号は白色信号 $s_i(n)$ を

$$s_i(n) = \varsigma_i(n) + r \cos \theta_i s_i(n-1) - r^2 s_i(n-2) \quad (i=1,2) \quad (8.1)$$

で表される 2 次 AR で生成した有色信号を使用する. ここで定数 r , θ_1 , θ_2 の値はそれぞれ

$$r = 0.9 \quad (8.2)$$

$$\theta_1 = \frac{1}{4}\pi \quad (8.3)$$

$$\theta_2 = \frac{3}{4}\pi \quad (8.4)$$

表 8.1: 適用する学習アルゴリズムの表記方法

FF-TIME	時間領域で学習する, 歪みを抑制しない通常の FF-BSS 式 (4.51)-(4.52) を使用して係数を更新 [12].
FF-TIME (MDP)	Minimal Distortion Principle に基づいた信号歪み抑制法 (時間領域で学習を行なう) 式 (6.39) を使用して係数を更新 [18].
FF-TIME (DF)	時間領域で学習をする, 本研究で提案する信号歪み抑制学習法 式 (4.51)-(4.52) で係数を更新し, 式 (7.95) で信号歪みを補正.
FF-TIME (ADF)	時間領域で学習する, 本研究で提案する近似法を用いた信号歪み抑制学習法 式 (4.51)-(4.52) で係数を更新し, 式 (7.96) で信号歪みを補正.
FF-FREQ (1)	周波数領域で学習する, 歪みを抑制しない通常の FF-BSS 式 (4.60)-(4.61) を使用して係数を更新 [12],[21],[29].
FF-FREQ (1+PB)	FF-FREQ(1) で分離した信号に対して Projection Back 法 [23] を適用した FF-BSS 分離した信号を式 (6.65) の正規化行列 $D(z)$ で信号歪みを補正. 式 (4.60)-(4.61) で分離回路の係数を更新.
FF-FREQ (1+DF)	FF-FREQ(1) に本研究で提案する信号歪み抑制の制約を付加した方式 式 (4.60)-(4.61) で係数を更新し, 式 (7.79) で信号歪みを補正.
FF-FREQ (1+ADF)	FF-FREQ(1) に本研究で提案する近似法を用いた信号歪み抑制の制約を付加した方式 式 (4.60)-(4.61) で係数を更新し, 式 (7.80) 信号歪みを補正.
FF-FREQ(2)	第 4.2.3 節で紹介した FF-FREQ(1) から余分な制約を除いたアルゴリズム 式 (4.66) 使用して係数を更新 [29].
FF-FREQ (2+PB)	FF-FREQ(2) で分離した信号に対して Projection Back 法 [23] を適用した FF-BSS 分離した信号を式 (6.65) の正規化行列 $D(z)$ で信号歪みを補正. 式 (4.66) で係数を更新.
FF-FREQ (2+DF)	FF-FREQ(2) に本研究で提案する信号歪み抑制の制約を付加した方式 式 (4.66) で係数を更新し, 式 (7.79) 信号歪みを補正.
FF-FREQ (2+ADF)	FF-FREQ(2) に本研究で提案する近似法を用いた信号歪み抑制の制約を付加した方式 式 (4.66) で係数を更新し, 式 (7.80) 信号歪みを補正.
FB	時間領域で学習を行なう FB-BSS 式 (2.40) を使用して係数を更新.

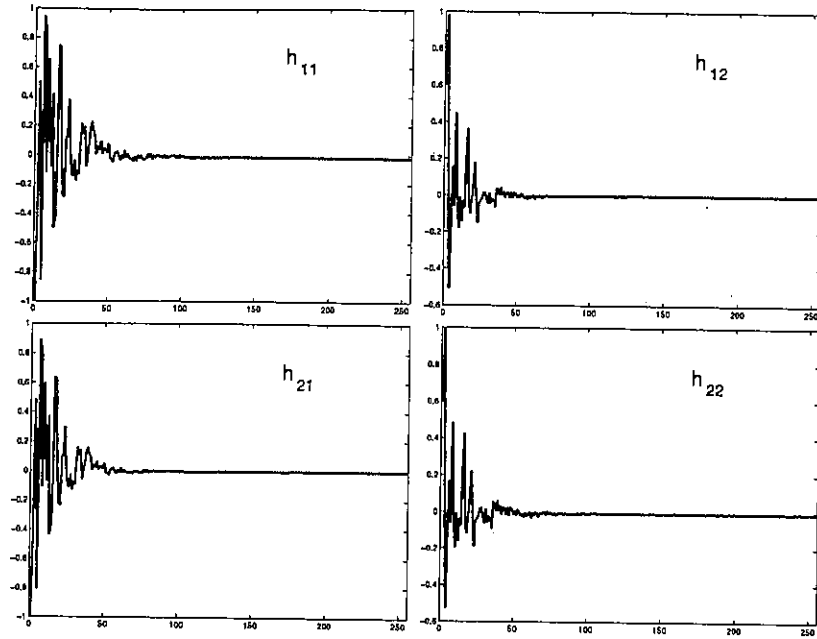
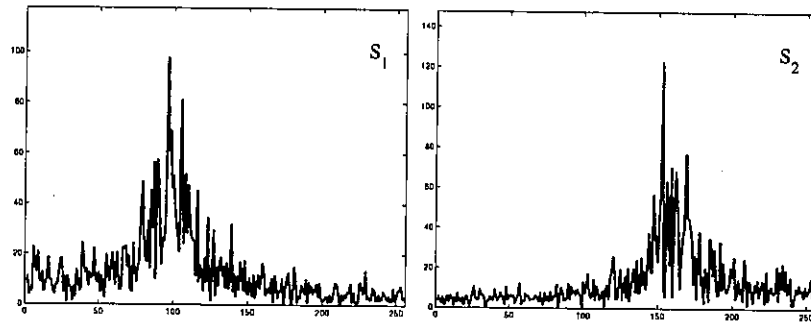


図 8.1: 伝達関数 (インパルス応答)

のように設定した. $s_1(n)$, $s_2(n)$ で θ の値を変えているので図 8.2 のように主成分が別の帯域に存在する.

図 8.2: $S_1(z)$ と $S_2(z)$ のスペクトル

時間領域における FF-BSS においてタップ数は 256, 分離回路の初期値は $W_{11}(z) = W_{22}(z) = 1$, $W_{12}(z) = W_{21}(z) = 0$ である. FB-BSS においてはタップ数は 256, 分離回路の初期値は $C_{12}(z) = C_{21}(z) = 0$ である. さらに周波数領域における FF-BSS においては FFT 点数 256, 分離回路の初期値は $W_{11}(z) = W_{22}(z) = 1$, $W_{12}(z) = W_{21}(z) = 0$ である.

3 チャンネル

3 チャンネルで使用する混合過程のインパルス応答を図 8.3 に示す. 信号源は音声と定常有色信号を用いる.

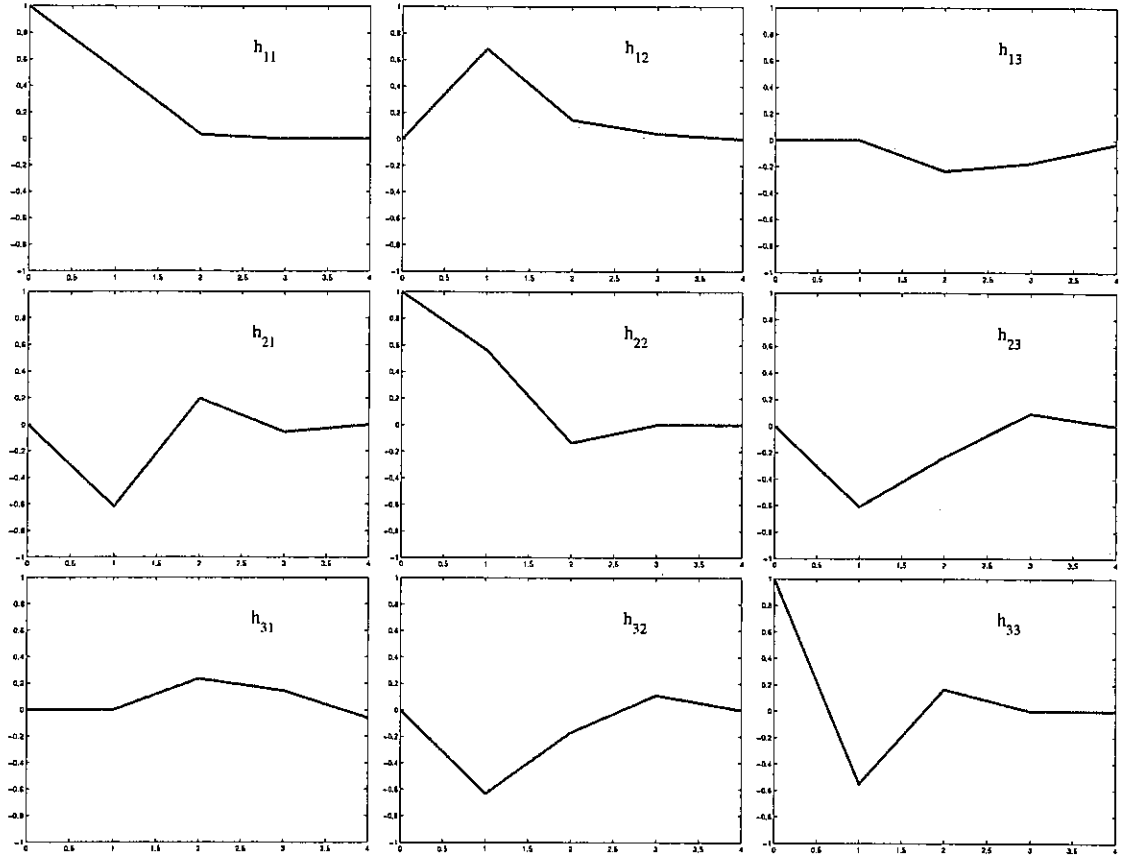


図 8.3: 伝達関数 (インパルス応答)

音声は以下の男性と女性の音声を繰り返し用いている.

$s_1(n)$: 音声 「人間は豊かな自然を求めています。」 (男性)

$s_2(n)$: 音声 「パンはポルトガル人が伝えたものです。」 (男性)

$s_3(n)$: 音声 「太宰治は青森県で生まれました。」 (女性)

定常有色信号は白色信号 $\varsigma_i(n)$ を

$$s_i(n) = \varsigma_i(n) + r \cos \theta_i s_i(n-1) - r^2 s_i(n-2) \quad (i = 1, 2, 3) \quad (8.5)$$

で表される 2 次 AR で生成した有色信号を使用する. ここで定数 r , θ_1 , θ_2 , θ_3 の値はそれぞれ

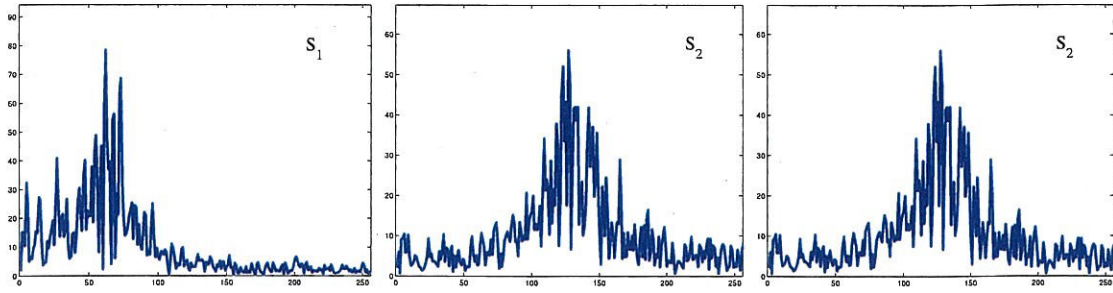
$$r = 0.9 \quad (8.6)$$

$$\theta_1 = \frac{1}{4}\pi \quad (8.7)$$

$$\theta_2 = \frac{2}{4}\pi \quad (8.8)$$

$$\theta_3 = \frac{3}{4}\pi \quad (8.9)$$

のように設定した. $s_1(n)$, $s_2(n)$, $s_3(n)$ で θ の値を変えているので図 8.4 のように主成分が別の帯域に存在する.

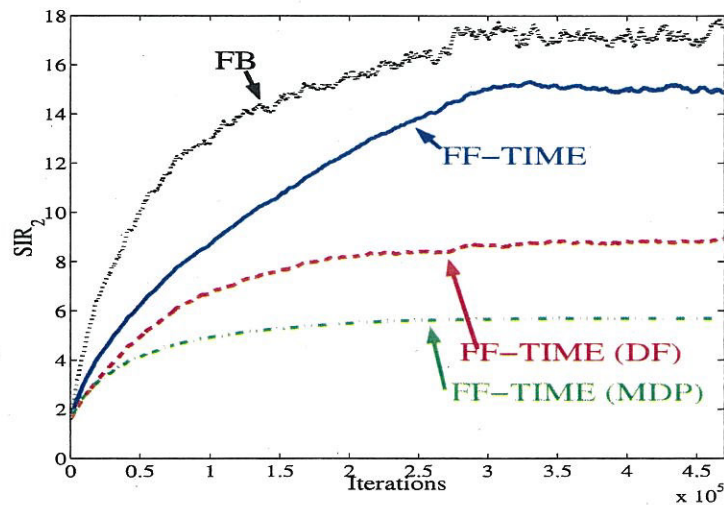
図 8.4: $S_1(z)$, $S_2(z)$ と $S_3(z)$ のスペクトル

時間領域における FF-BSS においてタップ数は 256, 分離回路の初期値は $W_{ii}(z) = 1$, $W_{ij}(z) = 0$ ($i \neq j$) である. FB-BSS においてはタップ数は 256, 分離回路の初期値は $C_{ij}(z) = 0$ ($i \neq j$) である. さらに周波数領域における FF-BSS においては FFT 点数 256, 分離回路の初期値は $W_{ii}(z) = 1$, $W_{ij}(z) = 0$ ($i \neq j$) である.

8.2.2 白色信号

時間領域で学習する BSS の性能

2 チャンネルにおける白色信号に対する SIR_2 の学習曲線を図 8.5 に示す.

図 8.5: 白色信号に対する SIR_2 の学習曲線

FB の収束速度は他の FF-BSS の収束速度よりも速く, かつ最終的な SIR_2 で評価される分離性能も最も高い. FF-BSS の中では FF-TIME が最も良い結果となっている. FF-TIME (DF) は FF-TIME (MDP) に比べ信号分離の性能が高くなっている. さらに FF-TIME (MDP) は既に収束しており, これ以上の分離性能の向上を見込めない.

学習後の信号分離, 信号歪みに関する評価を表 8.2 にまとめる.

SIR_1 , SIR_2 は, FB が最も良い性能を示している. 最も厳密な信号歪みの評価である SD_{1a} は FB が他の手法よりも高い. 振幅のみを比較する SD_{1b} では, その差が小さ

表 8.2: 時間領域で学習する BSS における白色信号に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	14.8	14.9	-4.72	-8.44	-4.77	-8.54
FF-TIME (MDP)	5.70	5.71	-12.7	-16.6	-12.7	-16.6
FF-TIME (DF)	8.97	8.93	-14.9	-18.5	-14.9	-18.4
FB	17.5	18.6	-19.1	-22.3	-19.1	-22.1

くなっている。しかし、FB が依然、最も高い性能を示している。伝達関数を使った評価 SD_{2a} , SD_{2b} でも大きな違いは見られない。したがって、どの信号歪みの評価に関しても FB が最も良いことが分かる。一方、提案する信号歪み抑制法である FF-TIME (DF) と従来の信号歪み抑制法である FF-TIME (MDP) の信号歪みは信号歪みを抑制しない方式である FF-TIME に比べて、劇的に小さくなっている。しかし、FF-TIME (MDP) は信号分離の性能が低下している。

FF-TIME (MDP) では第 6.1.3 節で述べたように、 $A_{ij}(z)$ を $H_{ij}(z)$ に近付けようとするために信号分離が妨げられる。このことを示すために表 8.3 に除去すべき信号と信号歪みの基準の相似性を示す。これは式 (5.3)-(5.10) の SD_{ix} ($i = 1, 2$) ($x = a, b$) の評価を $j \neq k$ として評価したものである。

表 8.3: 白色信号に対する除去すべき信号と信号歪みの基準の相似性

Methods	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	2.45	-1.23	2.42	-1.20
FF-TIME (MDP)	-10.4	-13.5	-10.3	-13.3
FF-TIME (DF)	-3.77	-10.0	-3.78	-9.98
FB	2.27	-3.79	2.23	-3.59

FF-TIME (MDP) の評価が他の手法に比べて圧倒的に小さくなっている。振幅の大きさのみを比較する SD_{1b} , SD_{2b} に関しては FF-TIME (DF) の値も小さくなっている。これは学習をすべての周波数帯域に対して一定のステップサイズで行なっているため、すべての帯域で同じ比率で信号の削除を行なっているために起こっている現象であると考えられる。

周波数領域で学習する BSS の性能

また、周波数領域で学習する方式について信号分離と信号歪みの評価を表 8.4 にまとめる。

FF-FREQ (1) はその他の手法に比べ、信号歪みの評価が悪い。提案する信号歪み抑制学習法である FF-FREQ (1+DF) や従来の信号歪み抑制学習法である FF-FREQ (1+PB) は信号歪みが大きく改善されている。さらに、FF-FREQ (1+DF) と FF-FREQ (1+PB) を比較すると FF-FREQ (1+DF) の方が大きく信号歪みが抑制されている。ま

表 8.4: 周波数領域で学習する BSS における白色信号に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-FREQ (1)	3.33	3.32	-8.97	-10.2	-7.91	-9.08
FF-FREQ (1+PB)	2.79	2.80	-13.6	-18.2	-12.8	-17.4
FF-FREQ (1+DF)	3.68	3.69	-15.3	-19.1	-14.6	-18.1
FF-FREQ (2)	3.11	3.11	-12.6	-16.2	-12.0	-15.6
FF-FREQ (2+PB)	3.35	3.34	-13.2	-17.3	-12.4	-16.4
FF-FREQ (2+DF)	3.38	3.38	-13.1	-17.2	-12.2	-16.1

た, FF-FREQ (1+PB) は信号分離の性能が低下している.

一方, FF-FREQ (2) に関しては FF-FREQ (1) に比べて信号歪みは低減されている. しかし, FF-FREQ (1+DF) よりも信号歪みは大きい. また, FF-FREQ (1+DF) や FF-FREQ (1+PB) の信号歪み抑制学習法を適用した場合にも, 大きな改善が見られない. 定常信号に対しては周波数領域で学習する手法に対して適用される信号歪み抑制学習法の効果が薄いようである. これについては更に検討が必要である. また, 定常な信号に対しては各フレームで信号が変化しないため分離の学習自体が難しく全体的に分離性能はあまり良くない.

8.2.3 音声

2チャンネルの時間領域で学習する BSS の性能

2チャンネルの音声に対する，歪みの基準である $H_{ii}(z)S_i(z)$ のスペクトルを図 8.6 に示す．音声は主成分が互いに同じ周波数帯域に存在しており，スペクトルの形が似ている．

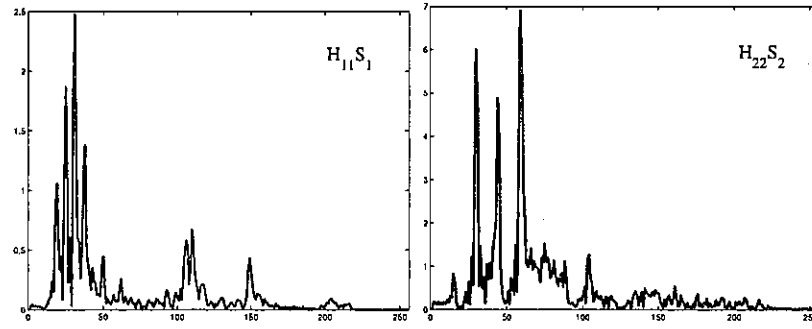


図 8.6: 音声に対する $H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル

図 8.7-8.10 に時間領域で学習した BSS の出力のスペクトルを示している．

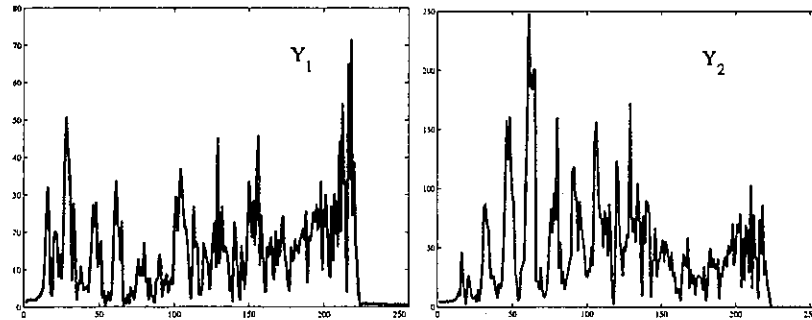


図 8.7: FF-TIME の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル

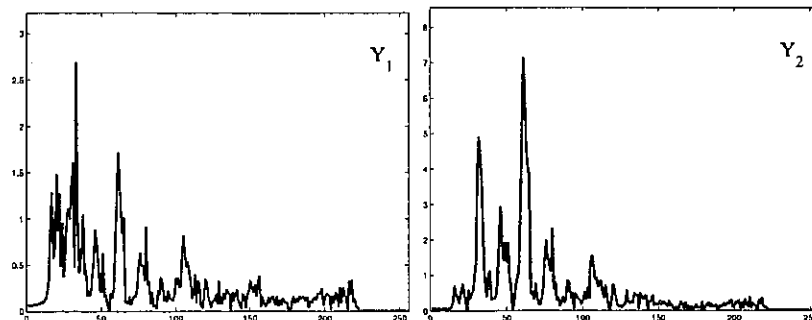
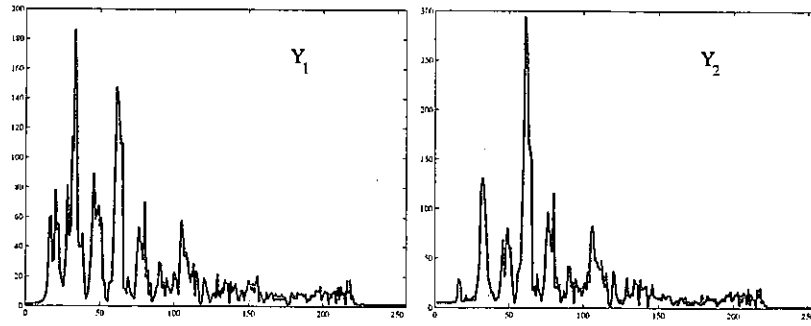
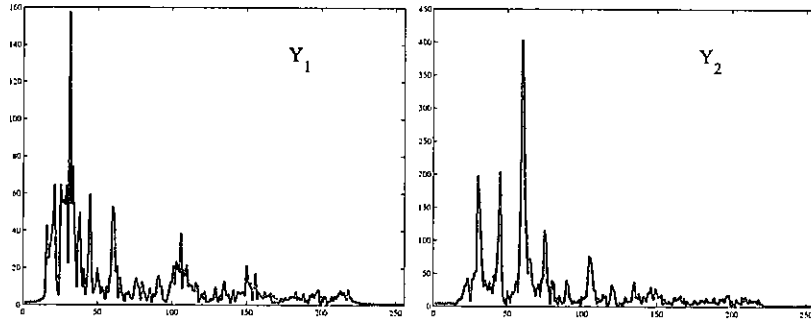


図 8.8: FF-TIME (MDP) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル

FF-TIME の出力スペクトルは図 8.6 に示す信号歪みの基準とは大きく違っている．特に高域の信号を増幅しており，出力のスペクトルを変化させることによって信号の独立

図 8.9: FF-TIME (DF) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル図 8.10: FB の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル

性を高めようとしていることが分かる．一方，図 8.8 に示されている FF-TIME (MDP) の出力のスペクトルでは FF-TIME の出力スペクトルに比べ高域のスペクトルを抑え，図 8.6 に示す信号歪みの基準に近い形になっている．また，図 8.9 に示されている，提案する信号歪み抑制学習法である FF-TIME (DF) の出力のスペクトルも FF-TIME の出力スペクトルのような変化は見られず，図 8.6 に示す信号歪みの基準と酷似している．

2 チャンネルの音声に対して時間領域で学習した BSS の信号分離と信号歪みの性能を表 8.5 に示す．

表 8.5: 時間領域で学習する BSS における音声に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	12.2	5.56	0.25	-2.94	0.57	-3.82
FF-TIME (MDP)	3.98	2.90	-10.3	-13.6	-8.24	-12.3
FF-TIME (DF)	8.33	4.33	-12.1	-16.2	-15.4	-19.9
FB	14.1	9.24	-14.5	-17.9	-14.7	-17.3

FF-TIME は信号歪みに対して性能が悪い．FF-TIME における信号分離の評価 SIR_i は他のものよりも高いが，信号歪みに起因するものであるので大きな意味を持たない．FF-TIME (MDP) は信号歪みを改善している．しかし，信号分離の性能が低い．提案する信号歪み抑制学習法である FF-TIME (DF) は信号分離の性能を保ったまま信号歪みを抑制している．FB は第 5.3 節で述べたように完全分離の解と無歪みの解が同じで

あるため信号分離の評価 SIR_i , 信号歪みの評価 SD_{ix} ともに高い性能を示している.

2 チャンネルの周波数領域で学習する BSS の性能

図 8.11-8.16 に周波数領域で学習を行なう手法で学習を行なった際の出力信号のスペクトルを示す.

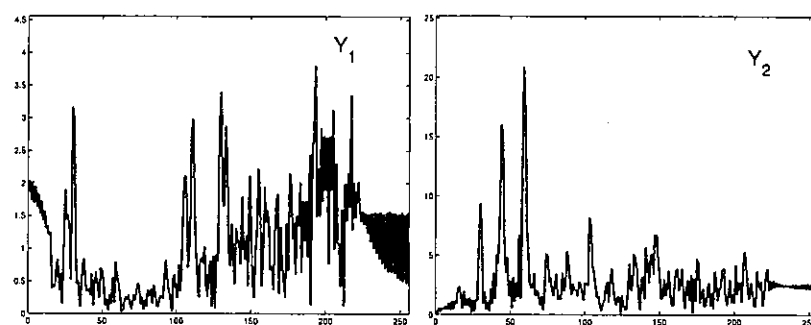


図 8.11: FF-FREQ (1) の出力信号のスペクトル

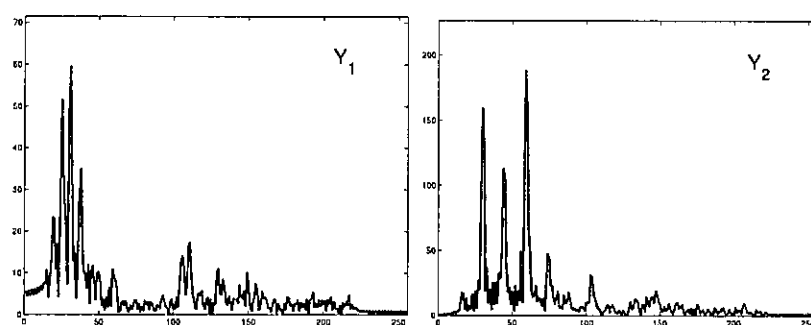


図 8.12: FF-FREQ (1+PB) の出力信号のスペクトル

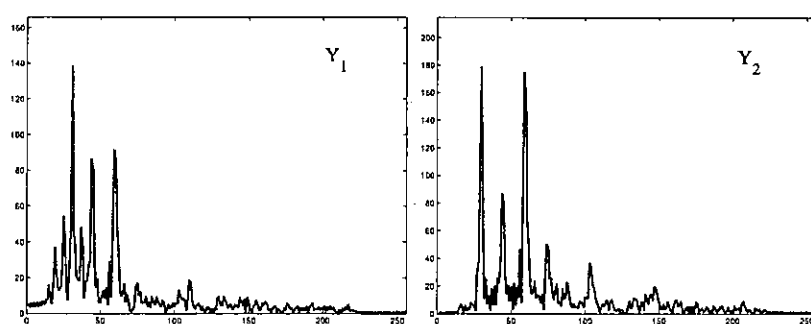


図 8.13: FF-FREQ (1+DF) の出力信号のスペクトル

図 8.11 に示す FF-FREQ (1) の出力信号スペクトルは図 8.8 に示す信号歪みの基準とは大きく異なっている. これは第 4.2.3 節で述べたように学習アルゴリズムに出力信号のパワーを一定にする制約を内包しているため全周波数帯域で大きさが一定になっている.

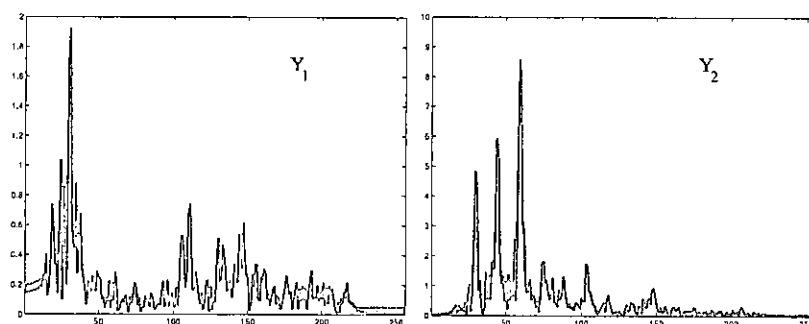
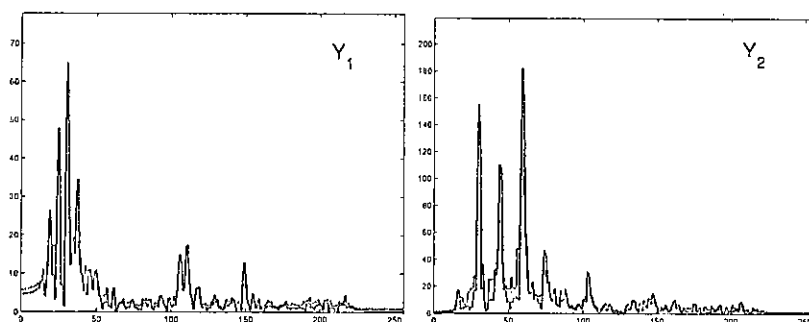
図 8.14: FF-FREQ (2) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル

図 8.15: FF-FREQ (2+PB) の出力信号のスペクトル

一方、図 8.14 に示す FF-FREQ (2) の出力信号スペクトルは FF-FREQ (1) の出力信号スペクトルに比べ、図 8.8 に示す信号歪みの基準に近い形になっている。これは、FF-FREQ (2) は FF-FREQ (1) のアルゴリズムから出力パワーを一定の値に近付ける制約を除いたものであるため、周波数スペクトルが平坦化されることが回避された結果であると考えられる。重み付け効果により信号スペクトルが歪む可能性も考えられるが、第 5.2.3 節でも述べているように信号源が共に音声であるため信号源の周波数分布が似ており、主に信号が分布している低域を強調し、あまり信号が分布していない高域を抑える形になり、分離回路 $W(r)$ の学習は、低域の更新量が大きく、高域の更新量が小さくなるといった偏りが生じ、出力信号が信号源の形に似てくることになり、特に問題はない。逆に収束を早める結果になっていると考えられる。

FF-FREQ (1+PB) の出力信号スペクトル、FF-FREQ (2+PB) の出力信号スペクトル

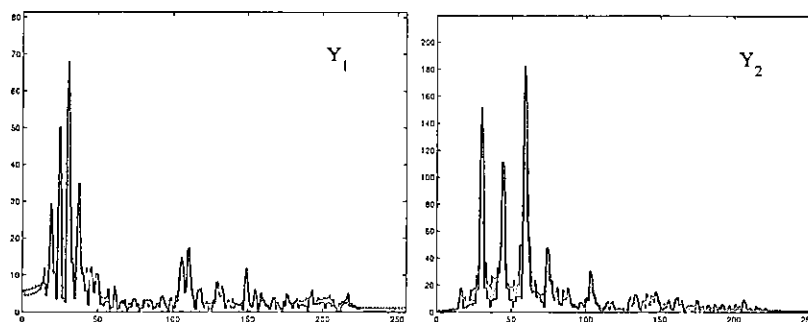


図 8.16: FF-FREQ (2+DF) の出力信号のスペクトル

ルともに図 8.8 に酷似している。しかし FF-FREQ (1+PB) の出力信号スペクトルに比べ、FF-FREQ (2+PB) の出力信号スペクトルの方が図 8.8 に示す信号歪みの基準により近い形になっている。これは後述する信号分離の差であると考えられる。FF-FREQ (1+PB), FF-FREQ (2+PB) は出力信号を正規化行列 $D(z)$ を用いて補正しているだけなので分離の学習には影響を全く及ぼさない。正規化行列 $D(z)$ は完全分離の際に無歪みになるように補正されるため、分離の性能が高いほど信号歪みを抑制する効果が高いと考えられる。そのため FF-FREQ (1+PB), FF-FREQ (2+PB) の信号分離の性能の差が信号スペクトルに現れたと考えられる。

FF-FREQ (1+DF), FF-BSS (2+DF) の出力信号のスペクトルも図 8.8 に酷似している。したがって、信号歪みの抑制が十分に行なわれていることが視覚的に確認できる。

2 チャンネルの音声に対して周波数領域で学習した BSS の信号分離と信号歪みの性能を表 8.6 に示す。

表 8.6: 周波数領域で学習する BSS における音声に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-FREQ (1)	7.37	2.61	-6.23	-8.67	-2.82	-2.95
FF-FREQ (1+PB)	15.2	6.59	-23.1	-27.1	-16.9	-18.8
FF-FREQ (1+DF)	9.68	6.38	-13.5	-18.1	-15.1	-18.3
FF-FREQ (2)	13.0	12.9	-9.43	-15.1	-10.9	-13.9
FF-FREQ (2+PB)	14.3	11.6	-16.9	-20.6	-16.9	-18.9
FF-FREQ (2+DF)	19.8	11.8	-24.6	-28.1	-17.9	-20.6

FF-FREQ (1) は信号分離、信号歪みのどちらの評価も良くない。これに対し、FF-FREQ (1+PB) は信号歪み、信号分離のどちらの評価も改善している。これは提案する FF-FREQ (1+DF) よりも高い性能となっている。また、FF-FREQ (2) も FF-FREQ (1) に比べ、分離性能、信号歪みともに良い。しかし、信号歪みの抑制は十分ではない。FF-FREQ (2+PB) は FF-FREQ (2) の信号歪みを改善している。提案する信号歪み抑制学習法 FF-FREQ (2+DF) はこれよりさらに信号分離、信号歪みともに改善している。そして、これは今回比較した手法で最も良い性能である。

3 チャンネルの時間領域で学習する BSS の性能

3 チャンネルの音声に対して時間領域で学習した BSS の信号分離と信号歪みの性能を表 8.7 に示す。

2 チャンネルのときと同様に FF-TIME では低い信号歪みの性能が FF-TIME (MDP), FF-TIME (ADF) が改善されている。改善の量は提案する FF-TIME (ADF) の方が大きい。また、信号分離の評価に関しても FF-TIME (ADF) の方が FF-TIME (MDP) よりも高い。FB は信号分離、信号歪みのどちらに対しても高い性能が得られている。

表 8.7: 3 チャンネルの時間領域で学習する BSS における音声に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	13.3	7.58	0.50	-2.48	-0.71	-4.65
FF-TIME (MDP)	6.25	4.36	-8.02	-11.2	-10.2	-14.1
FF-TIME (ADF)	8.00	5.24	-14.0	-17.9	-17.1	-20.5
FB	13.6	6.82	-15.4	-18.9	-17.0	-19.7

3 チャンネルの周波数領域で学習する BSS の性能

3 チャンネルの音声に対して周波数領域で学習した BSS の信号分離と信号歪みの性能を表 8.8 に示す。

表 8.8: 3 チャンネルの周波数領域で学習する BSS における音声に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-FREQ (1)	11.2	5.32	-10.4	-16.1	-4.39	-4.52
FF-FREQ (1+PB)	15.2	6.59	-23.1	-27.1	-16.9	-18.8
FF-FREQ (1+ADF)	8.50	7.36	-14.7	-21.1	-16.3	-18.7
FF-FREQ (2)	17.2	10.3	-13.4	-19.4	-13.9	-16.4
FF-FREQ (2+PB)	16.3	9.68	-17.3	-21.1	-18.0	-19.9
FF-FREQ (2+ADF)	16.1	9.56	-24.1	-28.4	-20.1	-21.7

こちらも 2 チャンネルのときと同様に FF-FREQ (1), FF-FREQ (2) から FF-FREQ (1+PB), FF-FREQ (2+PB) がそれぞれ性能を改善させている。さらに, FF-FREQ (1+ADF), FF-FREQ (2+ADF) が高い性能を示している。

8.2.4 2次ARモデルによって生成された有色信号

2チャンネルの時間領域で学習するBSSの性能

2チャンネルの有色信号に対する歪みの基準である $H_{ii}(z)S_i(z)$ のスペクトルを図8.17, 観測信号のスペクトルを図8.18に示す. この有色信号は, 主成分が互いに違う周波数帯域に存在しており, スペクトルの形が似ていない.

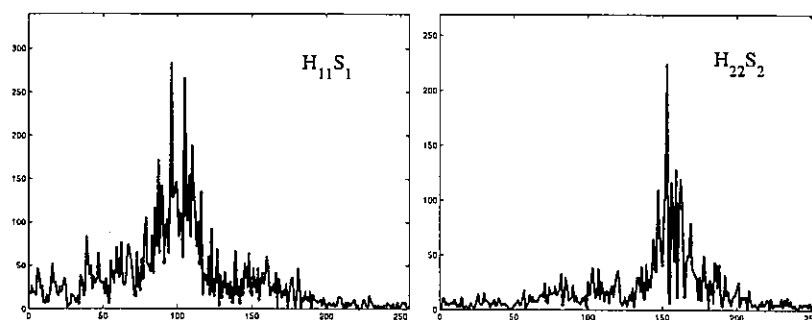


図 8.17: $H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル

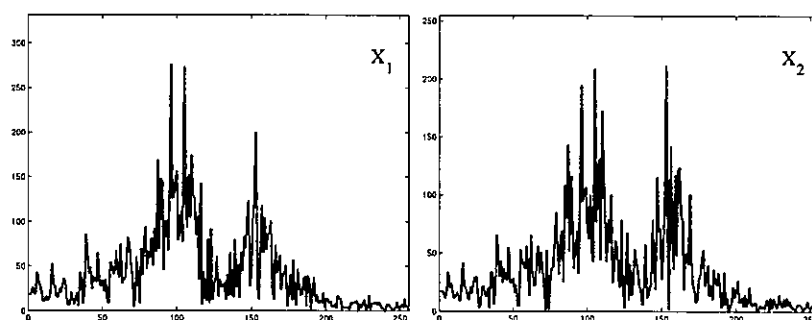


図 8.18: $X_1(z)$ と $X_2(z)$ のスペクトル

図 8.19-8.22 に時間領域で学習した BSS の出力のスペクトルを示している.

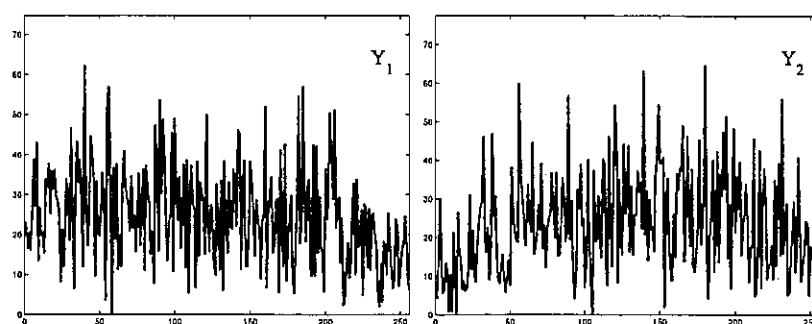
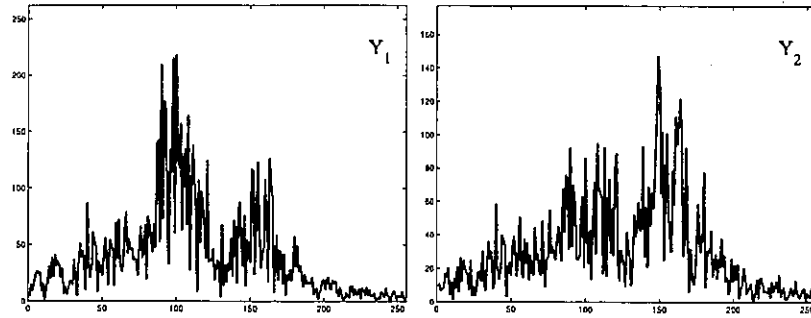
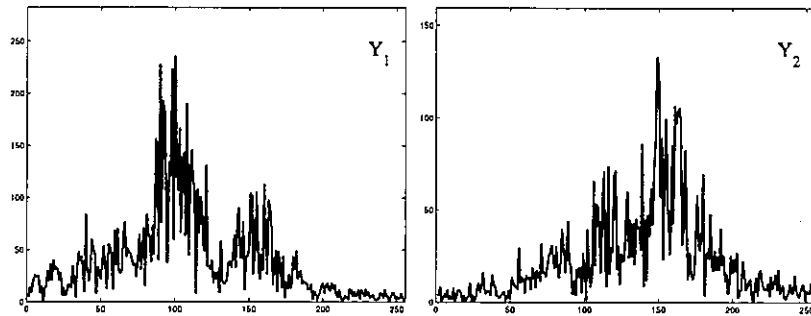


図 8.19: 有色信号に対する FF-TIME の出力 $y_1(n)$, $y_2(n)$ のスペクトル

図 8.20: 有色信号に対する FF-TIME (MDP) の出力 $y_1(n)$, $y_2(n)$ のスペクトル図 8.21: 有色信号に対する FF-TIME (DF) の出力 $y_1(n)$, $y_2(n)$ のスペクトル

音声のときと同様に図 8.19 に示す FF-TIME のスペクトルはより独立にしようとして、本来信号が存在しない周波数帯域を増幅している。図 8.20 に示す FF-TIME (MDP), 図 8.21 に示す FF-TIME (DF) は余計な部分の増幅は抑えられている。どちらも図 8.17 に示す信号歪みの基準より、図 8.18 に示す観測信号そのものに近い形になっているが、FF-TIME (DF) の方が信号歪みの基準に近い形になっている。これは両者とも十分な分離性能が得られていないために干渉成分が残っているためだと考えられるが、FF-TIME (MDP) は第 6.1.3 節で述べた問題により FF-TIME (DF) よりも分離性能が劣化している。FB は信号歪みの基準と似ている。

2 チャンネルの有色信号に対して時間領域で学習した BSS の信号分離と信号歪みの性能を表 8.9 に示す。

FF-TIME (MDP) は高い信号歪み抑制性能を示しているが分離性能が低い。これも

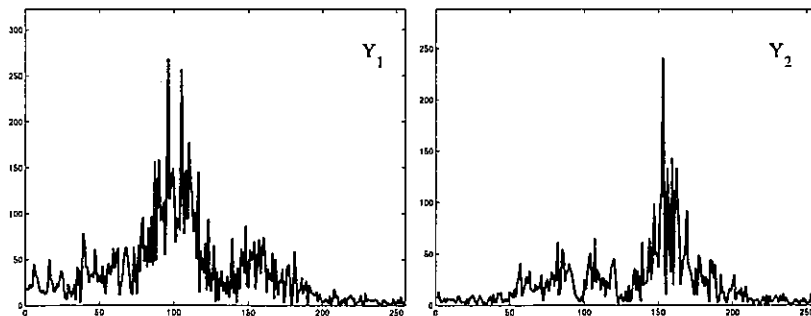
図 8.22: 有色信号に対する FB の出力 $y_1(n)$, $y_2(n)$ のスペクトル

表 8.9: 時間領域で学習する BSS における有色信号に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	7.07	9.49	-0.28	-3.11	-0.69	-4.99
FF-TIME (MDP)	2.20	4.49	-15.7	-18.8	-13.7	-16.5
FF-TIME (DF)	4.07	8.05	-14.4	-16.6	-10.4	-13.2
FB	7.19	16.5	-16.7	-18.6	-10.4	-13.6

白色信号や音声信号に対してみられたのと同様に, $A_{ij}(z)$ を $H_{ij}(z)$ に近づけようとするために信号分離が妨げられたためである. これは表 8.10 において FF-TIME (MDP) の値のみが小さくなっていることから分かる. 一方, FF-TIME (DF), FB は分離性能, 信号歪み抑制性能ともに高い性能が得られている.

表 8.10: 有色信号に対する除去すべき信号と信号歪みの基準の相似性

Methods	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	1.93	-0.65	1.71	-0.99
FF-TIME (MDP)	-8.79	-11.1	-9.48	-12.2
FF-TIME (DF)	-0.59	-2.61	-2.85	-4.74
FB	1.02	-1.43	-0.09	-4.08

2 チャンネルの周波数領域で学習する BSS の性能

2 チャンネルの有色信号に対する周波数領域で学習した BSS の出力信号のスペクトルを図 8.23- 8.29 に示す.

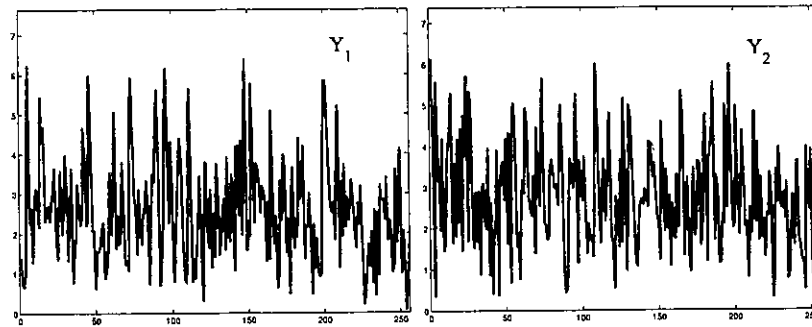
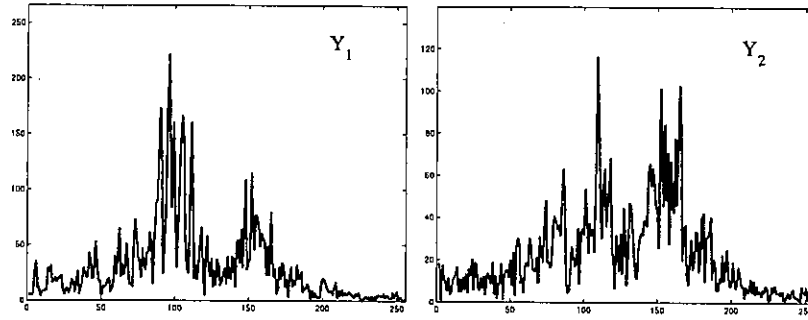
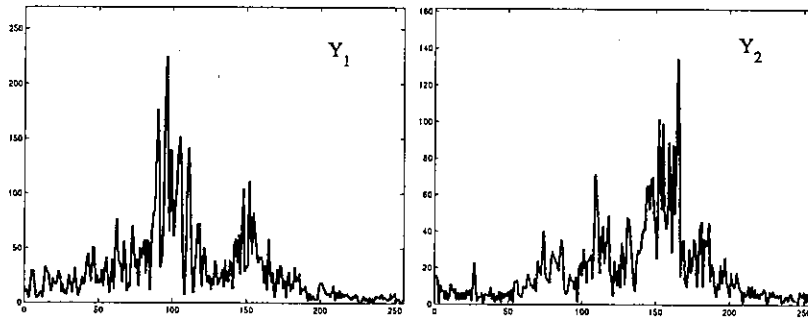
図 8.23: 有色信号に対する FF-FREQ (1) の出力 $y_1(n)$, $y_2(n)$ のスペクトル

図 8.23 に示す FF-FREQ (1) は信号パワーを一定にする制約のため全体的に白色化している.

第 5.3 節で述べたように周波数領域で学習する FF-BSS では重み付け効果が存在する. これは信号源が互いに似ている場合には信号歪みを抑制したが, 信号源が互いに

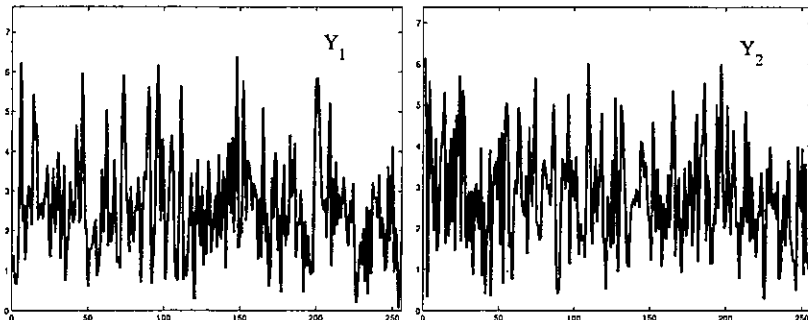
図 8.24: 有色信号に対する FF-FREQ (1+PB) の出力 $y_1(n)$, $y_2(n)$ のスペクトル図 8.25: 有色信号に対する FF-FREQ (1+DF) の出力 $y_1(n)$, $y_2(n)$ のスペクトル

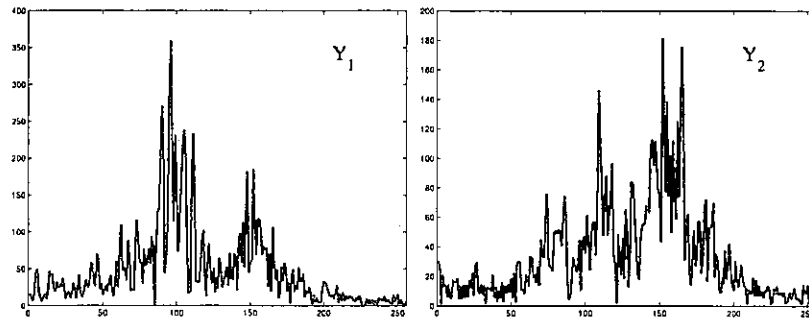
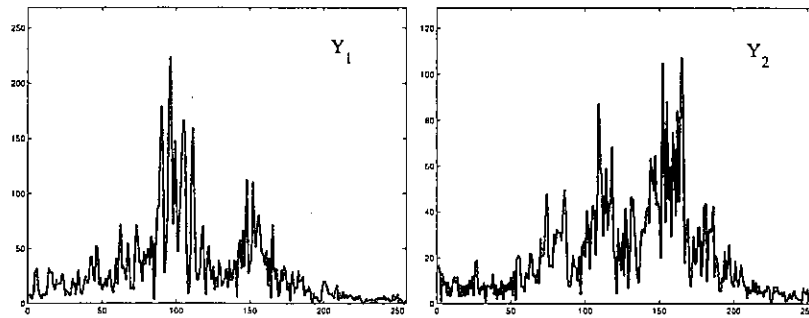
似ていない場合には逆に信号歪みを助長してしまう。図 8.27 に示す FF-FREQ (2) は干渉音のスペクトルがそのまま残っており、重み付け効果が現れている。

図 8.24-8.29 に示した信号歪み抑制を適用した学習法では信号歪みが改善され信号歪み抑制法を使用しないものに比べて図 8.17 に示す歪みの基準に近付いている。しかし、干渉信号の主成分がある部分に信号が残っており、重み付けの影響を除去し切れていない。

2 チャンネルの有色信号に対して周波数領域で学習した BSS の信号分離と信号歪みの性能を表 8.11 に示す。

FF-FREQ (1) は信号分離、信号歪みともに性能が良くない。FF-FREQ (1+PB) は信号分離の性能が低いままであるが、信号歪みに関して改善されている。一方、提案する方式である FF-FREQ (1+DF) は信号分離、信号歪みともに高い性能が得られて

図 8.26: 有色信号に対する FF-FREQ (2) の出力 $y_1(n)$, $y_2(n)$ のスペクトル

図 8.27: 有色信号に対する FF-FREQ (2) の出力 $y_1(n)$, $y_2(n)$ のスペクトル図 8.28: 有色信号に対する FF-FREQ (1+PB) の出力 $y_1(n)$, $y_2(n)$ のスペクトル

いる。

FF-FREQ (2) は FF-FREQ (1) に比べて信号歪みの評価である SD_{ix} が改善されている。FF-FREQ (2+PB), FF-FREQ (2+DF) は僅かに信号歪みの評価が改善しているがあまり大きな改善ではない。主成分が互いに違う周波数帯域に存在する定常な有色信号の場合、FF-BSS FREQ (1+DF) が信号分離、信号歪み両方に関して一番高い性能を示している。

3 チャンネルの時間領域で学習する BSS の性能

3 チャンネルの有色信号に対して時間領域で学習した BSS の信号分離と信号歪みの性能を表 8.12 に示す。

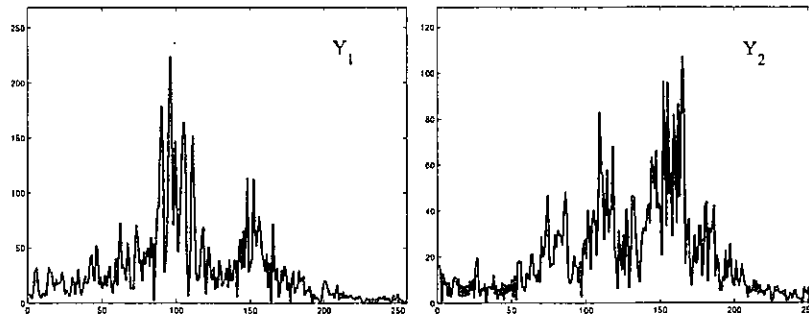
図 8.29: 有色信号に対する FF-FREQ (1+DF) の出力 $y_1(n)$, $y_2(n)$ のスペクトル

表 8.11: 周波数領域で学習する BSS における有色信号に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FREQ (1)	4.93	4.32	-3.62	-5.03	-3.67	-4.28
FREQ (1+PB)	3.22	1.51	-12.7	-14.3	-7.83	-9.38
FREQ (1+DF)	7.63	4.06	-18.4	-20.8	-10.8	-14.1
FREQ (2)	5.22	2.82	-12.2	-14.3	-7.42	-9.44
FREQ (2+PB)	5.35	2.62	-14.1	-15.8	-8.39	-10.7
FREQ (2+DF)	5.67	2.72	-13.7	-15.7	-8.27	-11.1

表 8.12: 3 チャンネルの時間領域で学習する BSS における有色信号に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-TIME	3.20	3.64	-4.12	-7.80	-5.57	-9.27
FF-TIME (MDP)	3.89	3.63	-6.23	-9.87	-6.47	-6.84
FF-TIME (ADF)	8.07	3.92	-21.1	-24.2	-13.9	-16.9
FB	14.6	3.91	-21.9	-25.0	-17.5	-20.3

FF-TIME は信号分離, 信号歪みの両方に対して性能が悪い. また, FF-TIME (MDP) においても大きな改善がみられない. 一方, FF-TIME (DF) では信号分離, 信号歪みの両方で大きな改善が見られる. FB は信号分離, 信号歪みの性能がともに非常に良い.

3 チャンネルの周波数領域で学習する BSS の性能

3 チャンネルの有色信号に対して周波数領域で学習した BSS の信号分離と信号歪みの性能を表 8.13 に示す.

表 8.13: 3 チャンネルの周波数領域で学習する BSS における有色信号に対する信号分離と信号歪みに関する評価

Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
FF-FREQ (1)	10.1	4.06	-9.92	-15.5	-6.25	-6.64
FF-FREQ (1+PB)	10.2	4.01	-19.6	-26.8	-12.8	-15.3
FF-FREQ (1+DF)	12.2	5.20	-23.8	-30.3	-14.7	-17.7
FF-FREQ (2)	11.7	4.81	-22.6	-28.8	-14.1	-16.8
FF-FREQ (2+PB)	11.5	4.76	-20.2	-27.5	-12.9	-16.0
FF-FREQ (2+ADF)	11.6	4.77	-22.4	-28.8	-13.9	-16.6

FF-FREQ (1) は信号歪みに関する性能が良くない. FF-FREQ (1+PB) は信号歪みに関して性能が改善されている. 一方, 提案する方式である FF-FREQ (1+DF) は信

号歪みの性能が改善されているとともに信号分離に関しても高い性能が得られている。

FF-FREQ (2) は信号歪み、信号分離ともに高い性能が得られている。信号歪み抑制学習法である FF-FREQ (2+PB), FF-FREQ (2+DF) は FF-FREQ (2) から性能が向上していない。これは設定した伝達関数が簡単なものであり、信号歪み抑制法を使用しない段階で十分に信号歪みが低減されているためだと考えられる。

8.3 考察

信号歪みの抑制を行なわない FF-TIME は、出力信号をより独立にしようとするために自由度を利用して周波数スペクトルを変形させてしまうため、信号の性質に関わらず信号歪みが生じる。

従来の信号歪みの抑制法 FF-TIME (MDP) は FF-TIME に比べ、信号歪みを抑制することが可能である。しかし、第 6.1.3 節で述べたように出力信号を観測信号に近付ける制約が分離の妨げになっているため信号分離性能は低い。つまり、FF-TIME (MDP) の学習中、 $A_{jj}(z)$ を $H_{jj}(z)$ に近付けようとする一方、 $A_{ji}(z)$ を $H_{jj}(z)$ に近付けようとしてしまう。 $A_{ji}(z)$ を $H_{jj}(z)$ に近付けることは干渉信号の成分が残ることになり信号分離の妨げになる。

提案する信号歪み抑制学習法である FF-TIME (DF) は高い信号分離性能を保ったまま信号歪みを抑制することが可能である。信号歪み抑制能力は FF-TIME (MDP) よりもさらに高い。

FB は第 5.3 節で述べたように完全分離の解と無歪みの解が同じであるため信号分離、信号歪みの評価ともに高い性能を示す。しかし、FB は第 2.2.1 節で述べたように混合過程が遅延の仮定を満たさなければならない。本章のシミュレーションはその仮定を満たす条件で行なっており、さらに反響音の影響も少なくなっている。この意味で本章のシミュレーションは FB にとって有利な条件であったと言える。この仮定に関しての詳しい解析は第 9 章で行なう。

時間領域で学習する BSS に関して、FB が信号分離、信号歪みともに最も高い性能が得られる。FF-BSS に関しては提案した信号歪み抑制学習法 FF-TIME (DF) が信号分離、信号歪みともに最も高い性能を示している。

どの信号源に対しても FF-FREQ (1) はその他の手法に比べ、信号歪みの評価が悪い。これは第 4.2.3 節で述べたように学習アルゴリズムに出力信号のパワーを一定にする制約を内包しているため全周波数帯域で大きさが一定になっていることに起因していると考えられる。これに対し、提案する信号歪み抑制学習法である FF-FREQ (1+DF) や従来の信号歪み抑制学習法である FF-FREQ (1+PB) は信号歪みを大きく改善する。

一方、FF-FREQ (2) に関しては FF-FREQ (1) に比べて信号歪みは低減される。FF-FREQ (2+PB)、F-BSS FREQ (2+DF) は音声のように非定常信号の場合には FF-FREQ (2) の信号歪みを改善する。しかし、定常信号の場合にはその効果があまりみられない。定常信号の場合には、FF-FREQ (2) に信号歪み抑制を適用したものより FF-FREQ (1) に信号歪み抑制を適用したものの方が信号歪みを抑制する。これについてはさらなる検討が必要である。また、定常な信号に対しては各フレームで信号が変化しないため分離の学習自体が難しく全体的に分離性能はあまり良くない。

周波数領域で学習する方式では定常信号に対しては FF-FREQ (1+DF) がもっとも信号分離、信号歪みどちらの性能も最も良い。また非定常信号に対しては FF-FREQ (2+DF) が最も信号分離、信号歪みの性能が良い。

第9章 混合過程に基づくFB-BSSとFF-BSSの性能比較

これまでFB-BSSはどの実験においても優れた分離性能、信号歪み抑制性能を示してきた。FB-BSSの学習は第2.2節にあるように信号の混合過程に対しての仮定の上になりたっている。この仮定は実際の音響空間では信号源 S_I から観測点 X_J , X_K の距離の差によって遅延差が発生することを利用している。ただし、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。

本章では実際の環境を想定して、伝達関数の遅延を変化させる。そのさまざまな遅延の中でFB-BSSの性能を解析し、FF-BSSと比較することによってFB-BSSの性能的境界を検証する[53],[54]。

9.1 遅延差量による性能変化

9.1.1 遅延の生成方法

これまで行ってきたシミュレーションにより8KHzサンプリングにおける音声信号、定常白色信号、定常有色信号に対して1サンプル、すなわち $1/8000\text{Sec}$ の遅延差において、FB-BSSは十分な分離性能、信号歪み抑制性能が得られることが分かっている。

実際の音響空間では1サンプルよりも更に少ない遅延差が起き得る。しかし、既に標本化された信号では1サンプルよりも小さい遅延を単純に発生させることはできない。そこで下記の3ステップで遅延を発生させることにする。

1. アップサンプリング
2. サンプルシフト
3. ダウンサンプル

すなわち、サンプル間を補間した後、遅延を発生させる。最終的に、サンプリング周波数をこれまでのデータと一致させるために間引きを行う。

アップサンプリング

連続信号 $u(t)$ をサンプリングした時に信号 $s(t)$ は

$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)u(t) \quad (9.1)$$

$$= u(t) \sum_{n=-\infty}^{\infty} \delta(t - nT) \quad (9.2)$$

と表すことができる。

さらに連続信号の周波数特性 $U(\omega)$ に対し、離散信号 $s(t)$ の周波数特性 $S(\omega)$ は

$$\Delta(\omega) = \mathcal{F} \left(\sum_{n=-\infty}^{\infty} \delta(t - nT) \right) = \sum_{n=-\infty}^{\infty} e^{-jnT\omega} \quad (9.3)$$

$$= \sum_{k=-\infty}^{\infty} \delta \left(\omega - \frac{k}{T} \right) \quad (9.4)$$

$$S(\omega) = U(\omega) * \Delta(\omega) \quad (9.5)$$

$$= \int_{-\infty}^{\infty} U(w) \Delta(w - \omega) dw \quad (9.6)$$

$$= \sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} U(w) \delta \left(w - \omega - \frac{k}{T} \right) dw \quad (9.7)$$

$$= \sum_{k=-\infty}^{\infty} U \left(\omega + \frac{k}{T} \right) \quad (9.8)$$

のように表せる。すなわち、元の連続信号の周波数特性 $U(\omega)$ のをサンプリング周波数 ($f_s = 1/(2\pi T)$) 間隔で周期的に繰り返すことになる。逆に周波数特性におけるこの周期を伸ばすことによってサンプリング周波数を大きくすることが可能である。

今、図9.1(a)のような信号系列

$$s(n) = \delta(t - nT)u(t) \quad (9.9)$$

を離散時間フーリエ変換した場合、図9.1(b)で示される周波数特性

$$S(\omega) = \sum_{k=-\infty}^{\infty} U \left(\omega + \frac{k}{T} \right) \quad (9.10)$$

$$= U(\omega) \quad \left(-\frac{1}{T} < \omega < \frac{1}{T} \right) \quad (9.11)$$

が得られるとする。

一方、図9.2(a)に示すように、図9.1(a)で示される信号系列に $M-1$ 点、0を内挿した信号系列

$$s(n) = \begin{cases} \delta(t - nT)u(t) & (n = Mm) \\ 0 & (n \neq Mm) \end{cases} \quad (9.12)$$

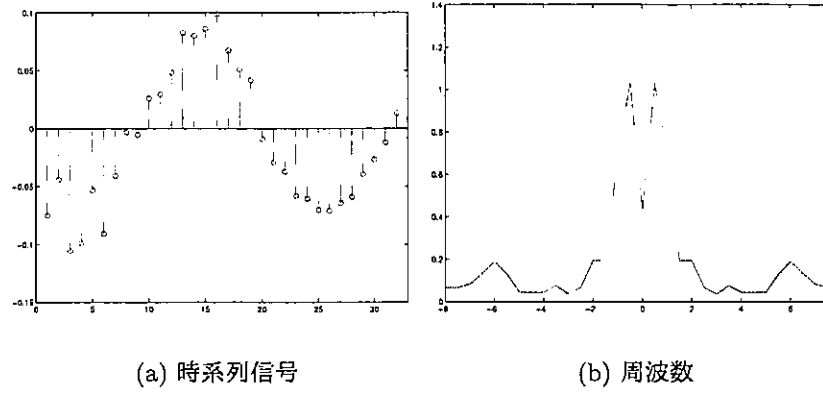


図 9.1: 遅延発生前の信号

を考える．式 (9.9) と式 (9.12) の系列はどちらも同じ信号式 (9.2) で表される $s(t)$ であると考えることができる．すなわち同じ周波数

$$S(\omega) = \sum_{k=-\infty}^{\infty} U\left(\omega + \frac{k}{T}\right) \quad (9.13)$$

を持つことになる．しかし，離散時間フーリエ変換を考えた場合，サンプリング周波数が M 倍になっているため，図 9.2(b) にあるように M 周期分表現する．

$$S(\omega) = \sum_{k=-\lfloor M/2 \rfloor}^{\lfloor M/2 \rfloor} U\left(\omega + \frac{k}{T}\right) \quad \left(-\frac{M}{T} < \omega < \frac{M}{T}\right) \quad (9.14)$$

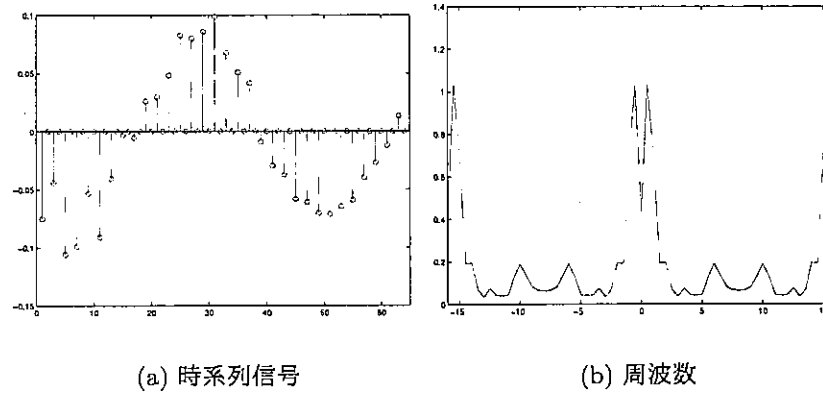


図 9.2: 零挿後の信号

また，図 9.3(a) に示すように信号 $u(t)$ をサンプリング間隔 T/M でサンプリングした信号

$$s(n) = \delta\left(t - \frac{nT}{M}\right) u(t) \quad (9.15)$$

の周波数特性は

$$S(\omega) = \sum_{k=-\infty}^{\infty} U\left(\omega + \frac{kM}{T}\right) \quad (9.16)$$

であり，離散時間フーリエ変換としては図9.3(b)に示すように

$$S(\omega) = U(\omega) \quad \left(-\frac{M}{T} < \omega < \frac{M}{T}\right) \quad (9.17)$$

と表現される．式(9.16)と式(9.14)を比較した場合， $U(\omega)$ の周期関数であるかない

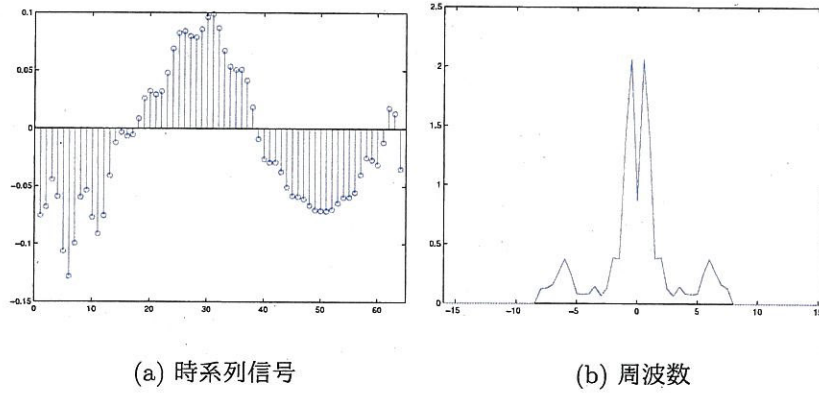


図 9.3: サンプル間隔 $T/2$ の信号

かのみが相違点である．つまり，式(9.14)の信号から式(9.16)の信号を生成するには $U(\omega)$ を取り出せばよい．そこで図9.4にあるように $-1/T < \omega < 1/T$ のみを通過させるようなローパスフィルタをかければ良い．

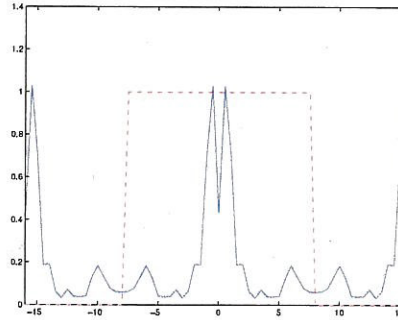


図 9.4: 通過域

M 倍アップサンプルさせる方法をまとめる．

1. 0 を $M - 1$ 個，内挿する

$$s_{i.p.}(n) = \begin{cases} s(m) & (n = Mm) \\ 0 & (n \neq Mm) \end{cases} \quad m = 0, 1, 2, \dots \quad (9.18)$$

2. カットオフ周波数が零挿前のサンプリング周波数の LPF をかける

$$S_{up}(\omega) = W_{LPF}(\omega)S_{i.p.}(\omega) \quad (9.19)$$

$$S_{i.p.}(\omega) = \mathcal{F}(s_{i.p.}(n)) \quad (9.20)$$

$$S_{up}(\omega) = \mathcal{F}(s_{up}(n)) \quad (9.21)$$

$$W_{LPF} = \begin{cases} 0 & (-\frac{M}{T} \leq \omega \leq -\frac{1}{T}) \\ 1 & (-\frac{1}{T} < \omega < \frac{1}{T}) \\ 0 & (\frac{1}{T} \leq \omega \leq \frac{M}{T}) \end{cases} \quad (9.22)$$

となる。ただし、サンプリング周波数はエリアジングを起こさないよう十分に高いものとする。

サンプルシフト

前述のアップサンプルによりサンプリング間隔が T から T/M となっている。つまり、 M 倍の分解能で遅延を発生させることができる。

$s(n) = 0$ ($n < 0$) であるとする、 L/M サンプル分の遅延を発生させるためには信号系列の始めに L 個の 0 を挿入すればよい。

$$s_{shift}(n) = \begin{cases} s_{up}(n - L) & (n \geq 0) \\ 0 & (n < 0) \end{cases} \quad (9.23)$$

サンプルシフト後の波形と周波数特性を図 9.5 に示す。ただし、図 9.5(b) に示す周波数

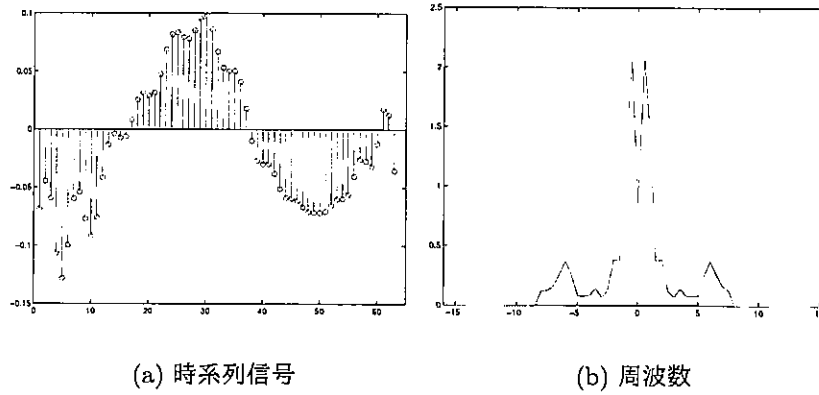


図 9.5: シフト後の信号系列

特性は振幅特性であり、時間シフトした信号は位相が回転するだけなので、図 9.3(b) と変化はみられない。

ダウンサンプル

サンプルシフトにより遅延を発生させた信号を得ることができたが、そのままではサンプリング周波数が違うため比較ができない。そこでダウンサンプルし、サンプリ

ング周波数を戻す。

これまでアップサンプルによりサンプリング周波数は M 倍に上げられていた。 M は整数であるのでそのまま M サンプル毎に $M - 1$ サンプル間引けばよい。すなわち、

$$s_{down}(n) = s_{shift}(Mn) \quad (9.24)$$

とする。間引いた信号を図 9.6 に示す。図 9.6(a) に示す信号系列は図 9.5(a) に比べ、サ

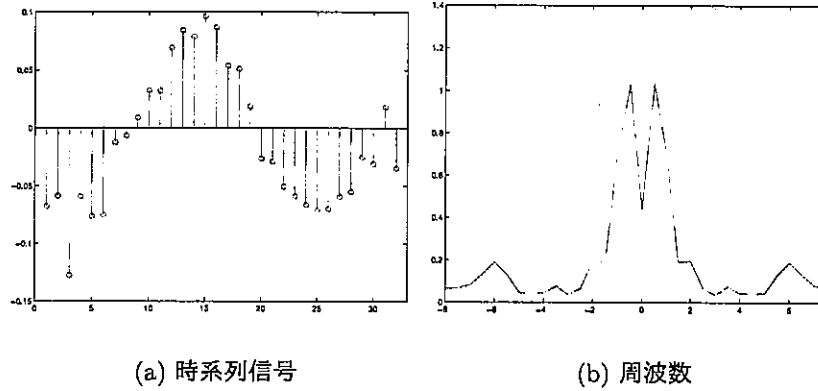


図 9.6: ダウンサンプル後の信号系列

ンプル数が減っているのが分かる。それに伴い周波数応答の周期が短くなっているのが図 9.6(b) から分かる。

挿入する 0 の数 M とシフトさせるサンプル数 L を適当にとることによって任意の遅延を挿入することができる。

9.1.2 対象となる音響空間

図 9.7 の混合過程で τ の値を変化させることによって、遅延時間を変化させて遅延時間による性能の変化を調べる。前節で紹介した生成方法で信号を生成することによって 1 サンプルより細かい精度で遅延を発生させることができる。すなわち、図 9.7 における τ を整数の精度ではなく実数の精度で変化させることが可能である。

混合過程 $H(z)$ には式 (9.25), (9.26) に示す単純に遅延と振幅の減衰のみのものと図 9.8, 9.8 に示す伝達関数の 2 通りの混合過程について検討する。

$$H(z) = \begin{bmatrix} 1 & 0.9 \\ 0.9 & 1 \end{bmatrix} \quad (9.25)$$

$$H(z) = \begin{bmatrix} 1 & 0.75 & 0.5 \\ 0.75 & 1 & 0.75 \\ 0.5 & 0.75 & 1 \end{bmatrix} \quad (9.26)$$

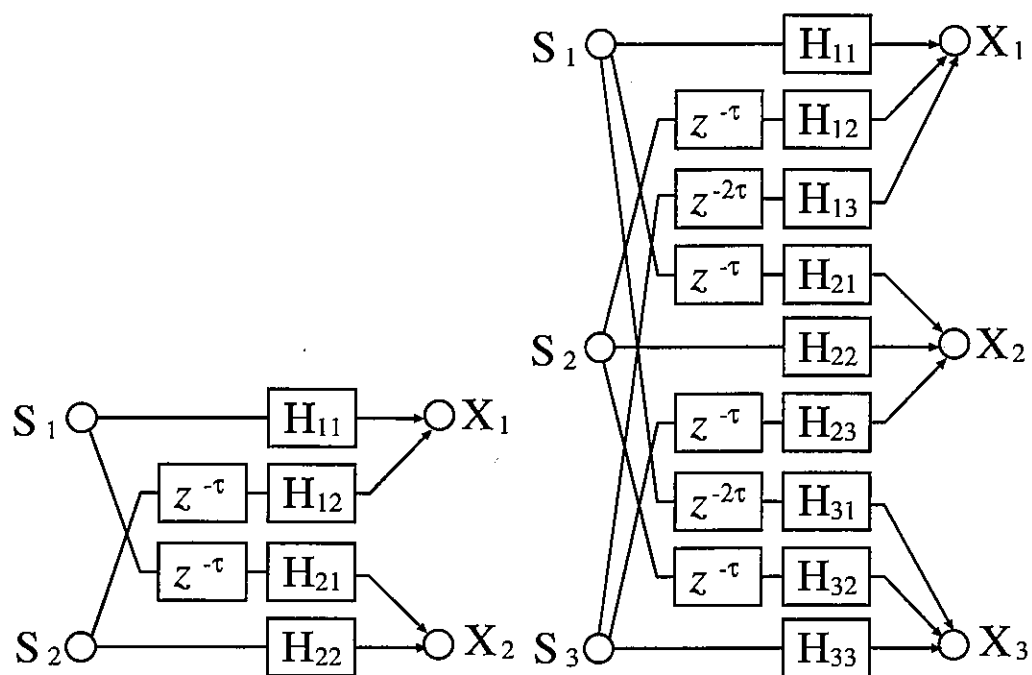


図 9.7: 対象とする混合過程のモデル

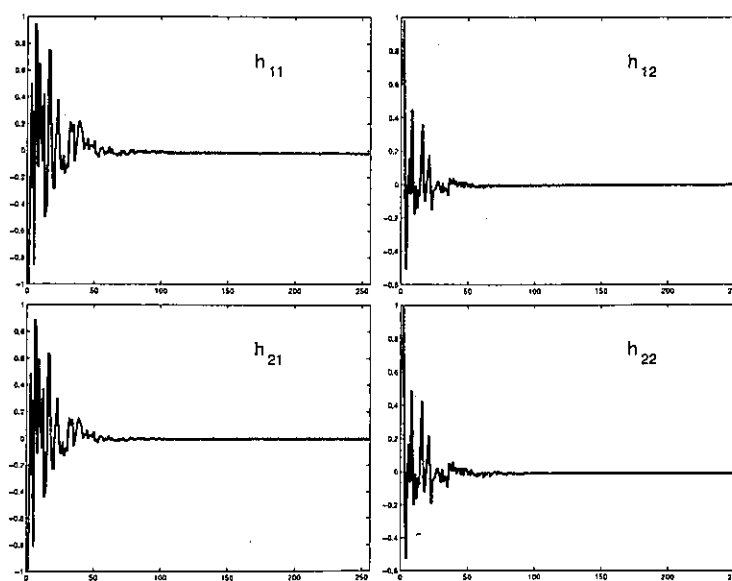


図 9.8: 伝達関数 (インパルス応答)

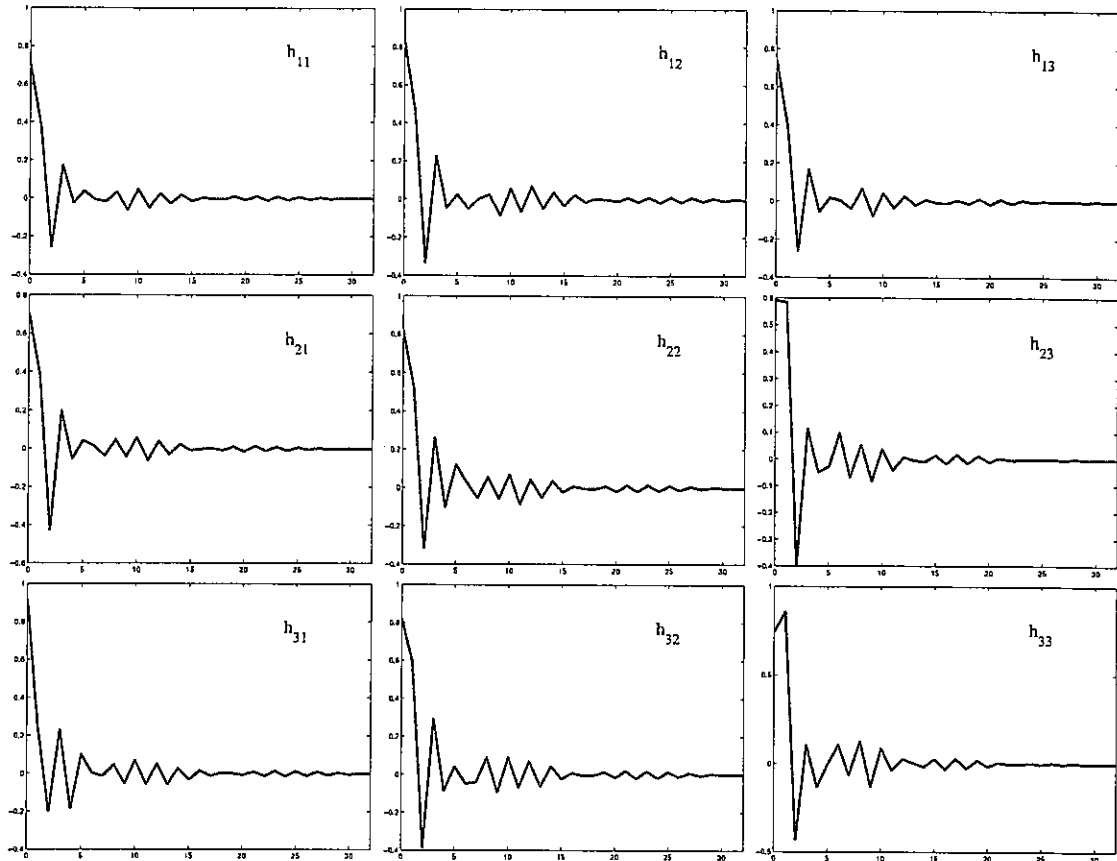


図 9.9: 伝達関数 (インパルス応答)

遅延量の変化がセンサ間の距離が変化したことによって引き起こされたと考えることができる。そこで図 9.10 における配置でセンサ間 L が変化したものとしてセンサ間 L と分離性能の変化の関係も検討する。

信号源には音声と白色信号を用いる。ただし、白色信号は厳密には最大周波数 $4kHz$ の帯域制限された信号である。

9.1.3 性能の比較

2 チャンネルの単純遅延

図 9.11, 図 9.12 に式 (9.25) の単純遅延の混合過程を用いた時の結果を示す。図 9.11 は信号源が音声, 図 9.12 は白色信号を用いている。それぞれ, 分離性能と同時に観測信号におけるそれぞれの相関を示してある。

FB-BSS は音声, 白色信号ともに遅延時間 τ が 0 のところで分離性能が最も悪く, 遅延時間 τ が大きくなるにつれて分離性能が向上している。一方, FF-BSS は遅延時間 τ に関係なく分離性能がほぼ一定である。FB-BSS では遅延が小さいところでは, 消去できないと仮定されている信号が消去されていて分離性能が低下しているが, 遅延が大きくなると仮定が十分満たされて性能が改善している。FF-BSS では混合過程の遅延に影

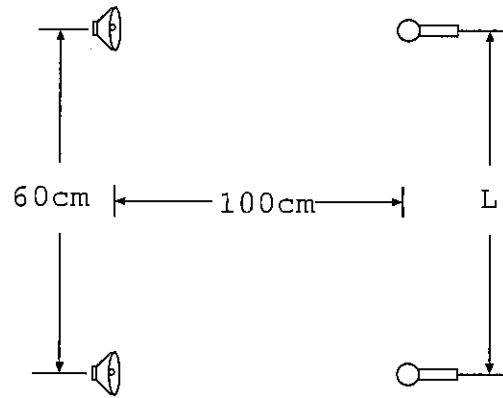


図 9.10: 想定する音響空間

響しないため分離性能が一定である。また、図 9.11(a), 9.11(b) と図 9.12(a), 9.12(b) を比較して分かるように音声の方が白色信号に比べて遅延が大きいところでも相関が高くなっている。FB-BSS の分離性能はそれに伴い、音声信号の方が白色信号より遅延時間 τ が大きいところでも性能があまり良くない。これは、遅延が存在しても信号の相関性によって信号が消去できるため消去できないと仮定している信号も削除できてしまうためである。したがって相関が小さくなるにつれて分離性能が向上していったものと考えられる。ここで、白色信号の相関が遅延時間 τ が 0 のところ以外でも値を持つてるのは最大周波数が 4kHz に帯域制限されているためである。また、白色信号に対して遅延時間が $n \times 125\mu\text{s}$ のとき高い分離性能が得られているのは、信号の相関が 0 であることに加え、遅延サンプルが整数サンプルになるため逆特性をディジタルシステムで生成できるために正確な逆特性を生成できるためであると考えられる。

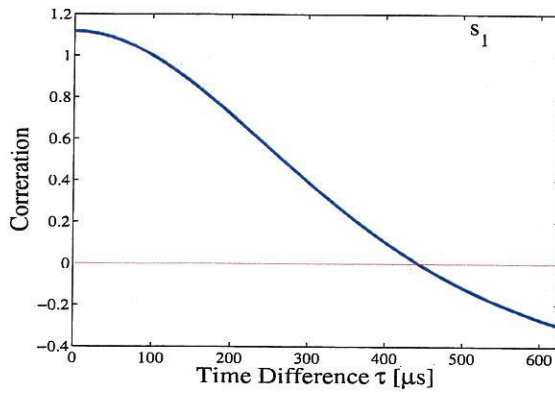
図 9.13, 9.14 に単純遅延における遅延が図 9.10 の音響空間でセンサ間 L の変化によって起きたものと仮定し、FF-BSS と FB-BSS の分離性能の比較を示す。

これらの図において信号源が音声のとき FF-BSS と FB-BSS の性能が同じになるのは 4.26cm のときであり、白色信号のときは 11.7cm のときである。つまり、音源の間が 1.3m 、センサと音源が 1.8m である状況を想定して設計されるシステムでは音源に音声を用いた場合には 4.26cm 以上、白色信号を用いたときは 11.7cm 以上で、FB-BSS が、逆にそれ以下では FF-BSS を用いた方がよいことになる。

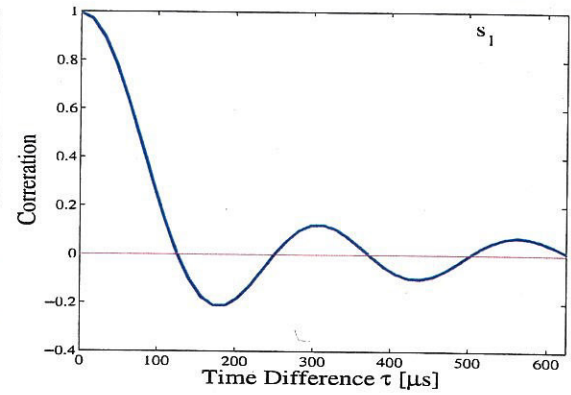
2 チャンネルの音響空間を模した伝達関数

図 9.15, 9.16 に図 9.8 の実際の音響空間を模した伝達関数を使用した際の、結果を示す。図 9.15 は信号源が音声、図 9.16 は白色信号を用いている。それぞれ、分離性能と同時に観測信号におけるそれぞれの相関を示してある。

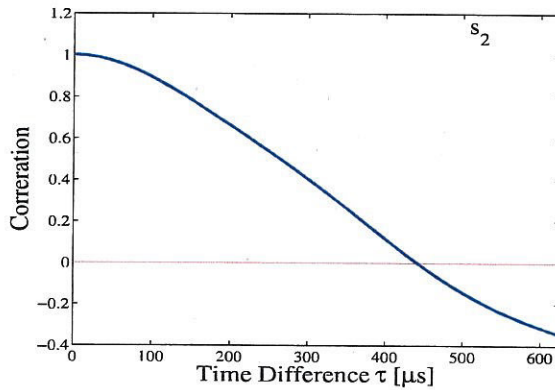
単純遅延のときと同様に、FB-BSS は音声、白色信号ともに遅延時間 τ が 0 のところで分離性能が最も悪く、遅延時間 τ が大きくなるにつれて分離性能が向上して行く。一方、FF-BSS は遅延時間 τ に関係なく分離性能がほぼ一定である。信号の相関は複雑な伝達関数を通っているため単純遅延のときのように信号の特性がはっきりとは現れていないが、白色信号に比べ音声は遅延時間 τ が大きいところまで相関の値が高い。ま



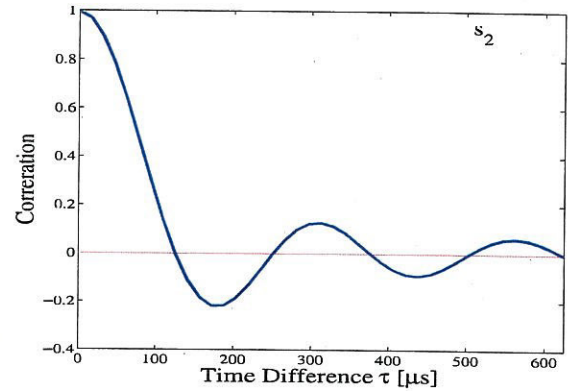
(a) 信号源が音声の時の $H_{11}(z)S_1(z)$ と $H_{21}(z)S_1(z)$ の相関



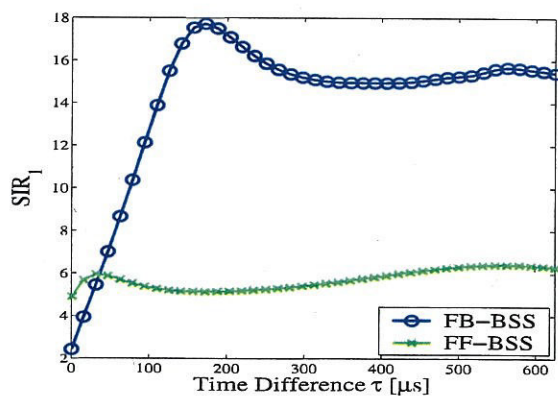
(a) 信号源が白色信号の時の $H_{11}(z)S_1(z)$ と $H_{21}(z)S_1(z)$ の相関



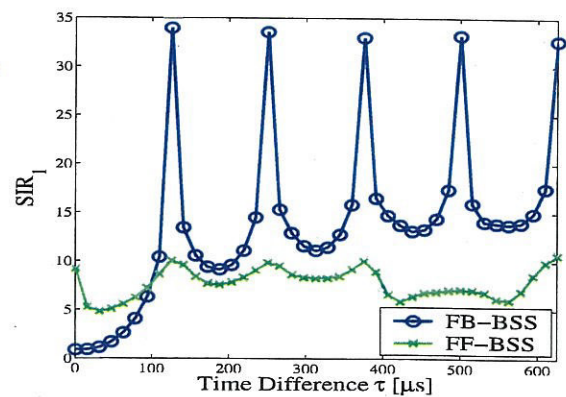
(b) 信号源が音声の時の $H_{22}(z)S_2(z)$ と $H_{12}(z)S_2(z)$ の相関



(b) 信号源が白色信号の時の $H_{22}(z)S_2(z)$ と $H_{12}(z)S_2(z)$ の相関



(c) 音声における遅延時間 τ による信号分離性能の変化



(c) 白色信号における遅延時間 τ による信号分離性能の変化

図 9.11: 音声を用いた単純遅延混合における FF-BSS と FB-BSS の性能比較

図 9.12: 白色信号を用いた単純遅延混合における FF-BSS と FB-BSS の性能比較

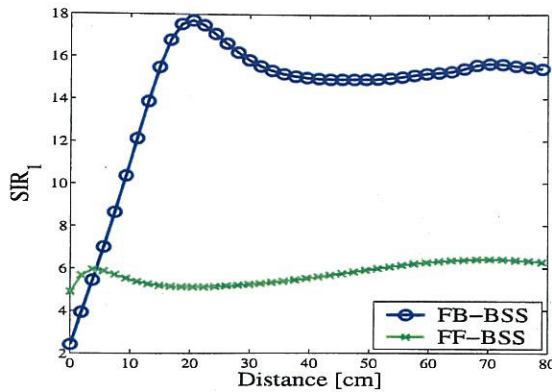


図 9.13: 音声を用いた単純遅延混合におけるセンサ間の変化による信号分離性能の変化

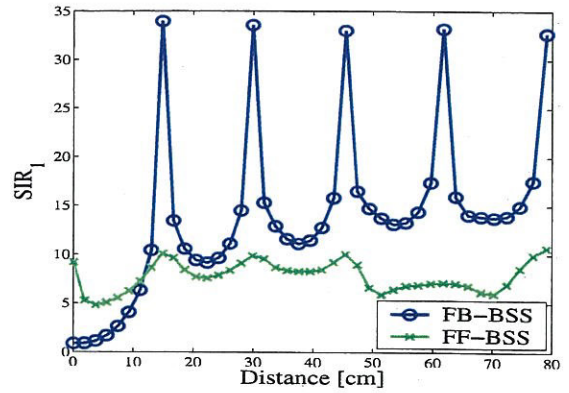


図 9.14: 白色信号を用いた単純遅延混合におけるセンサ間の変化による信号分離性能の変化

た、それに伴ってFB-BSSの分離性能も τ が大きいところまで劣化している。

図 9.17, 9.18 に混合過程が図 9.8 の実際の音響空間を模した伝達関数を使用した時の結果に関して遅延 τ が図 9.10 の音響空間でセンサ間 L の変化によって起きたものと仮定し、FF-BSS と FB-BSS の分離性能の比較を示す。

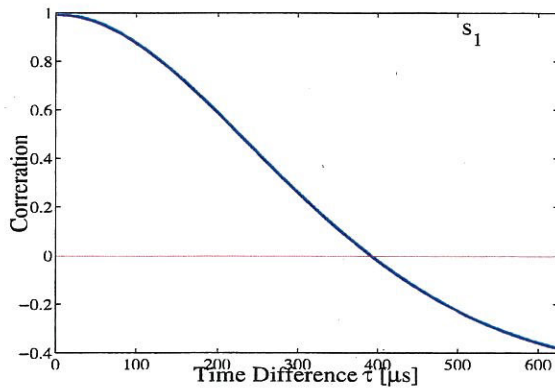
これらの図において信号源が音声のとき FF-BSS と FB-BSS の性能が同じになるのは 10.6cm のときであり、白色信号のときは 8.09cm のときである。つまり、音源の間が 1.3m、センサと音源が 1.8m である状況を想定して設計されるシステムでは音源に音声を用いた場合には 10.6cm 以上、白色信号を用いたときは 8.09cm 以上で、FB-BSS が、逆にそれ以下では FF-BSS を用いた方がよいことになる。これらの値は先ほどの単純遅延のときと異なっているため混合過程の伝達関数によって変わってくるものだと考えられる。いずれにしても、センサ間の距離が短いところでは FF-BSS が優位であり、長いところでは FB-BSS が優位である。

3 チャンネルの単純遅延

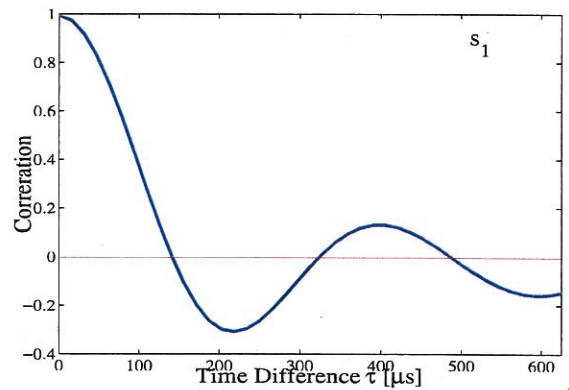
図 9.19, 図 9.20 に式 (9.26) の単純遅延の混合過程を用いた時の結果を示す。図 9.19 は信号源が音声、図 9.20 は白色信号を用いている。

2 チャンネルのときと同様に FB-BSS は音声、白色信号ともに遅延時間 τ が 0 のところで分離性能が最も悪く、遅延時間 τ が大きくなるにつれて分離性能が向上している。一方、FF-BSS は遅延時間 τ に関係なく分離性能がほぼ一定である。FF-BSS では混合過程の遅延に影響しないため分離性能が一定である。これにより 3 チャンネルにおいても FB-BSS は遅延の問題が発生していることが分かる。

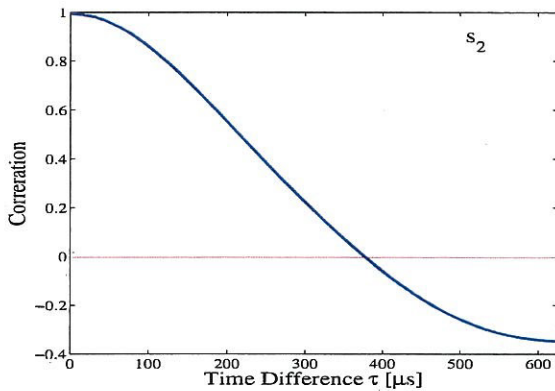
また、音声の方が白色信号に比べて十分な分離性能を得るために必要な遅延時間 τ が大きい。これは、図には示していないが音声の方が遅延が大きい部分でも相関性が高く、遅延が存在しても信号の相関性によって信号が消去できるため消去できないと仮定している信号も削除できてしまうためである。したがって相関が小さくなるにつれて分離性能が向上していったものと考えられる。また、2 チャンネルに比べて分離性能を得るために必要な遅延時間 τ が大きい。これは 2 チャンネルに比べて自由度が高



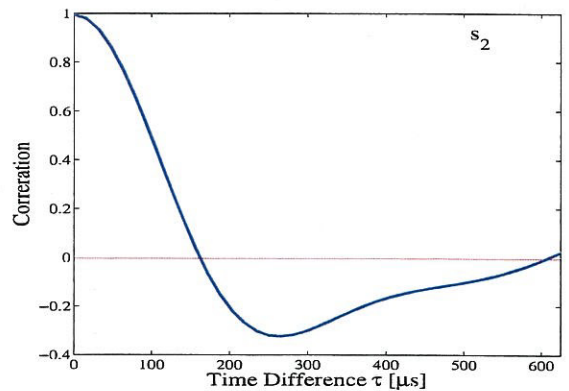
(a) 信号源が音声の時の $H_{11}(z)S_1(z)$ と $H_{21}(z)S_1(z)$ の相関



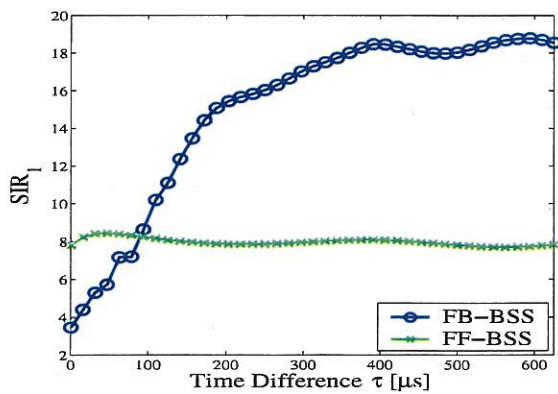
(a) 信号源が白色信号の時の $H_{11}(z)S_1(z)$ と $H_{21}(z)S_1(z)$ の相関



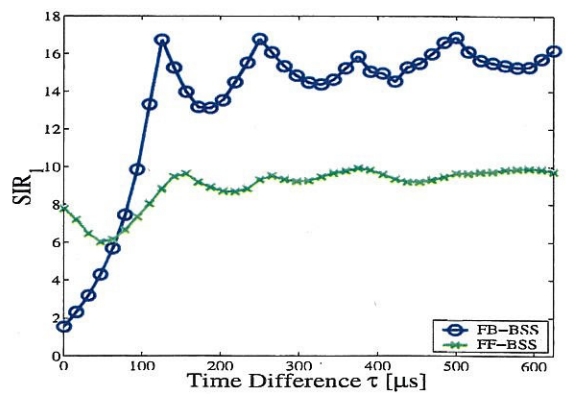
(b) 信号源が音声の時の $H_{22}(z)S_2(z)$ と $H_{12}(z)S_2(z)$ の相関



(b) 信号源が白色信号の時の $H_{22}(z)S_2(z)$ と $H_{12}(z)S_2(z)$ の相関



(c) 音声における遅延時間 τ による信号分離性能の変化



(c) 白色信号における遅延時間 τ による信号分離性能の変化

図 9.15: 音声における FF-BSS と FB-BSS の性能比較

図 9.16: 白色信号における FF-BSS と FB-BSS の性能比較

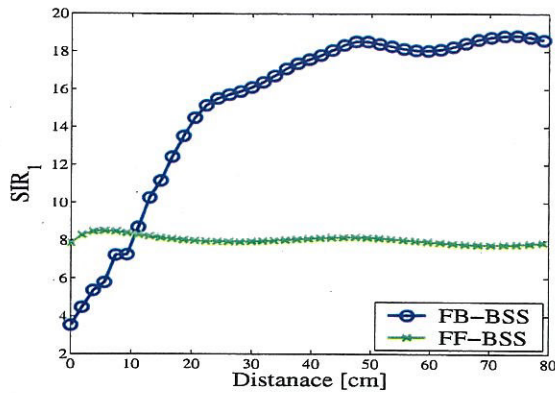


図 9.17: 音声におけるセンサ間の変化による信号分離性能の変化

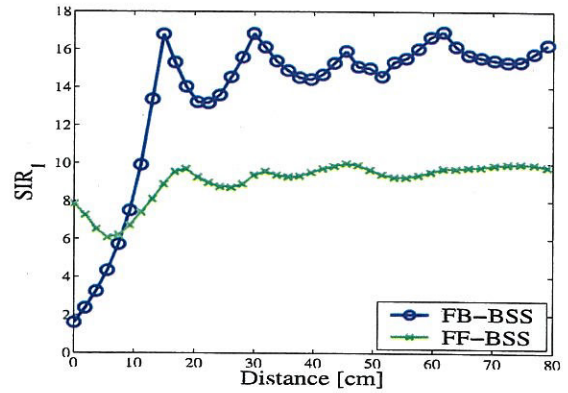
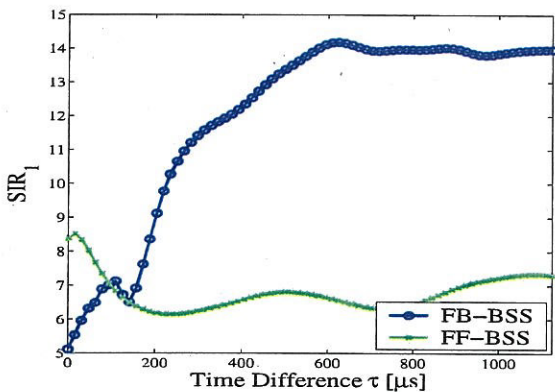
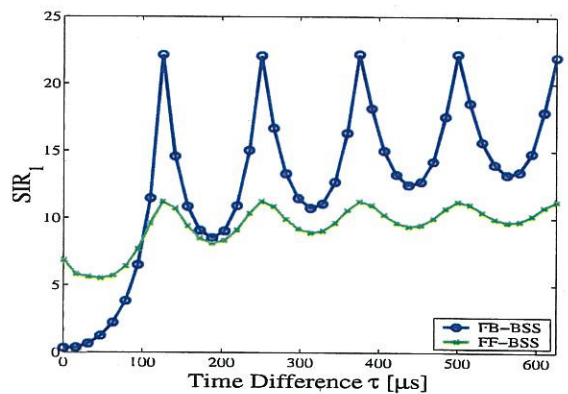


図 9.18: 白色信号におけるセンサ間の変化による信号分離性能の変化

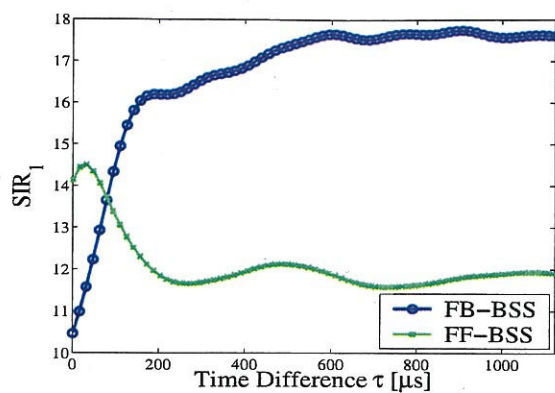
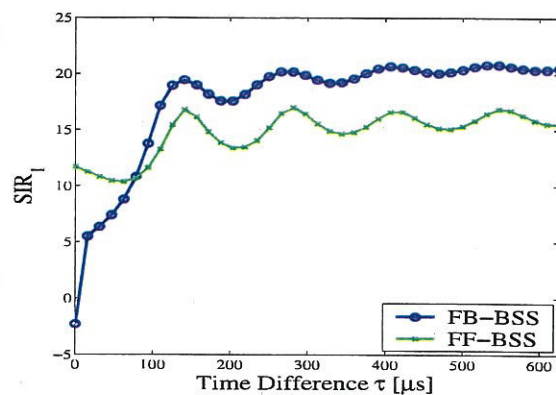
図 9.19: 音声における遅延時間 τ による信号分離性能の変化図 9.20: 白色信号における遅延時間 τ による信号分離性能の変化

いため、相関の影響が τ の大きい部分でも残っているためだと考えられる。また、白色信号に対して遅延時間が $n \times 125\mu\text{s}$ のとき高い分離性能が得られているのは、2チャンネルのときと同様に信号の相関が0であることに加え、遅延サンプルが整数サンプルになるため逆特性をディジタルシステムで生成できるために正確な逆特性を生成できるためであると考えられる。

3 チャンネルの音響空間を模した伝達関数

図 9.21, 9.22 に図 9.9 の実際の音響空間を模した伝達関数を使用した際の結果を示す。図 9.21 は信号源が音声、図 9.22 は白色信号を用いている。

これまでのシミュレーションと同様に、FB-BSS は音声、白色信号ともに遅延時間 τ が0のところで分離性能が最も悪く、遅延時間 τ が大きくなるにつれて分離性能が向上して。一方、FF-BSS は遅延時間 τ に関係なく分離性能がほぼ一定である。白色信号に関しては、単純遅延のときほどは明確に現れていないが、遅延時間が $n \times 125\mu\text{s}$ 付近で高い分離性能が得られている。これは混合過程により相関が0でなくなったため、それほど分離性能の改善には繋がっていないが、整数サンプルの遅延でないところで

図 9.21: 音声における遅延時間 τ による信号分離性能の変化図 9.22: 白色信号における遅延時間 τ による信号分離性能の変化

完全な逆システムが形成できないため分離の精度が劣化しており、波となって現れていると考えられる。また、遅延時間 τ が十分大きいところでのFB-BSSとFF-BSSの分離性能の差が若干小さくなっている。

9.2 無遅延差における性能比較

実際の音響空間では遅延差がないというのはあり得ないが、観測点が無限に近付いたと仮定し、その際の FB-BSS と FF-BSS の性能を比較することによって、それぞれの特性を解析する。

2チャンネルの音声に対する歪みの基準である $H_{ii}(z)S_i(z)$ のスペクトルを図 9.23, に示す。

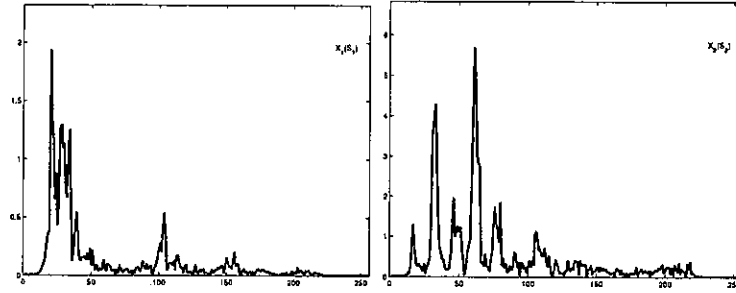


図 9.23: 音声に対する $H_{11}(z)S_1(z)$ と $H_{22}(z)S_2(z)$ のスペクトル

図 9.24, 9.25 に FF-TIME (DF) と FB の出力のスペクトルを示している。

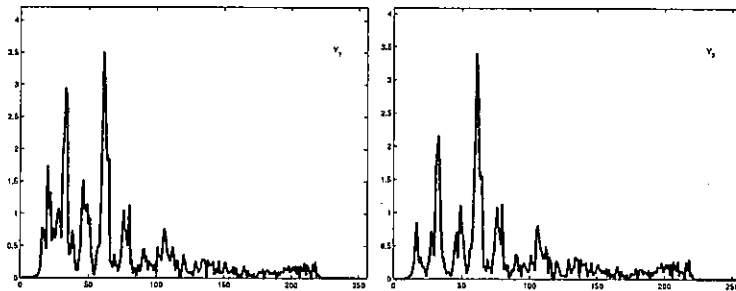


図 9.24: FF-TIME (DF) の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル

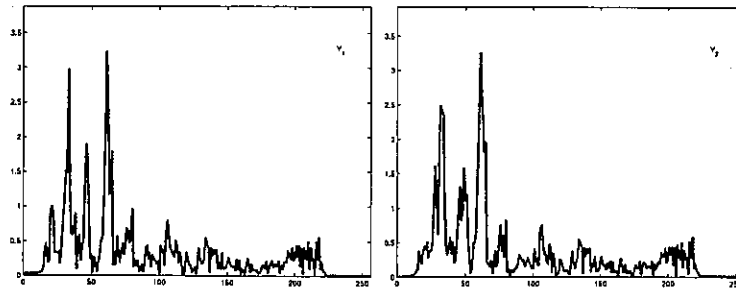


図 9.25: FB の音声信号に対する出力信号 $y_1(n)$, $y_2(n)$ のスペクトル

図 9.25 に示す FB の出力のスペクトルに比べ、図 9.24 に示す FF-TIME (DF) の出力のスペクトルの方が信号歪みに基準である $H_{ii}(z)S_i(z)$ のスペクトルの形に近い形と

なっている。これは、混合過程の伝搬遅延時間に差がないため FB で学習が進まず、信号が混ざりあったままの状態であるためスペクトルに差が出たものと考えられる。

次に音声、白色信号に対して 2 チャンネルの FF-TIME (DF) と FB の信号分離と信号歪みの性能を表 9.1 に示す。

表 9.1: 2 チャンネルの無遅延状態での信号分離と信号歪みに関する評価

Signal Source	Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
white signal	FF-TIME (DF)	4.23	8.24	-17.1	-21.5	-17.2	-21.2
	FB	4.73	0.55	-7.65	-10.3	-2.84	-7.94
speech	FF-TIME (DF)	7.90	7.89	-21.3	-24.2	-21.0	-23.4
	FB	1.61	1.49	6.60	10.3	-5.53	-9.97

これまで行なってきた混合過程に十分な遅延差がある場合には FB は分離性能、信号歪み抑制の性能ともに高い性能が得られていたが、遅延差がない場合の結果である表 9.1 によると FB は音声、白色信号のどちらの信号源に対しても信号分離の性能が低く、それに伴い信号歪みの評価も低い。

さらに表 9.2 に音声、白色信号に対して 3 チャンネルの FF-TIME (DF) と FB の信号分離と信号歪みの性能を示す。

表 9.2: 3 チャンネルの無遅延状態での信号分離と信号歪みに関する評価

Singal Source	Methods	SIR_1	SIR_2	SD_{1a}	SD_{1b}	SD_{2a}	SD_{2b}
white signal	FF-TIME (DF)	14.3	8.83	-23.2	-26.2	-24.0	-27.2
	FB	8.59	2.18	-17.7	-21.1	-10.4	-14.5
speech	FF-TIME (DF)	11.7	11.7	-26.5	-28.7	-26.5	-28.8
	FB	4.15	4.49	3.01	2.92	-12.8	-18.1

これも 2 チャンネルのときと同様に FB は信号分離、信号歪みのどちらも十分な性能が得られていない。つまり、混合過程に遅延が生じない場合には、信号の分離が十分に行なわれず、それに伴い信号歪みも増大してしまう。

第10章 結論

本研究では畳み込み混合形のBSSに対して、時間領域で処理を行なうものを中心に、回路形式による特性の違いというものに着目し、さまざまな解析を行なった。

まずFB-BSSに関して、第2章で混合過程の伝搬遅延に着目し、解析を行ない、出力パワーを最小化するアルゴリズムを導いた。このアルゴリズムはJuttenらが提案した出力信号の独立性を高めるアルゴリズムと等価である。伝搬遅延に着目した解析によりFB-BSSでは反響音による影響を受ける場合には信号分離性能が劣化することが明らかになった。そこで、第3章において反響音の影響を軽減する方式を提案した。この章ではまず反響音の影響は分離回路のタップの後半に現れることからタップ後半の学習を抑える固定形EWステップサイズ方式を提案した。さらに、固定形EWステップサイズ方式ではステップサイズの形を決定する際に分離回路の最適値の包絡線の情報が必要となるためステップサイズの形を自動的に調整する適応形EWステップサイズ方式を提案した。シミュレーションによりこれらの方式が従来適応フィルタで提案されている学習の高速化に加え反響音の影響を軽減できることを確認した。

第5章において、信号歪みの基準について議論し、その議論に基づき適切な信号歪みの基準を定めた。さらにその信号歪みの基準に基づき第2、4章で紹介したFB-BSSとFF-BSSについて信号分離と信号歪みの解析を行なった。FB-BSSは完全分離と無歪みの解が同一であるため完全に分離が行なわれると信号歪みが起こらない。一方、FF-BSSは自由度が存在するため、独立になるように学習を行ない信号歪みを引き起こす可能性が高い。

従来、FF-BSSに対する信号歪みに対してそれを抑制する信号歪み抑制学習法が提案されていた。第6章ではそのアルゴリズムを紹介し、そのアルゴリズムの問題点を指摘した。

第7章では、新たにFF-BSSに対する信号歪み抑制学習法を提案した。これは完全分離かつ無歪みの条件から導く信号歪みを抑制する制約を従来の学習法に加えるものである。この制約は簡単化のためまず2チャンネルに対して考察し、その後、一般的なマルチチャンネルのブラインドソースセパレーションに対して拡張した。この制約は時間領域で学習する方式、周波数領域で学習する方式のどちらにも付加することが可能である。

続く第8章で、第2、4章で紹介した基本的なFB-BSSとFF-BSS、第6章で紹介した従来の信号歪み抑制法、さらに第7章で提案した信号歪み抑制法をシミュレーションにより性能を比較した。FF-BSSの中では提案した信号歪み抑制法が信号分離、信号歪みの両面で最も良い結果が得られている。この提案した信号歪み抑制法を用いたFF-BSSとFB-BSSを比べると信号歪みに関しては同等の性能が得られるが、信号分離はFB-BSSの方が高い。

しかし、FB-BSSの学習では混合過程の伝達遅延の差にある仮定をおき、それに基づ

いてアルゴリズムが形成されているため、その仮定が崩れると性能が著しく劣化する。そこで第 9 章では、混合過程の伝達遅延の差に着目し、FB-BSS と FF-BSS を比較し、それぞれの有効性について検討を行なった。この結果、伝達遅延の差が小さい場合には FF-BSS が有効であり、十分に伝達遅延に差がある場合には FB-BSS が有効であることが分かった。

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.2655-2658.
- [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 III: Stability analysis", Signal Processing, 24, pp.21-29, 1991.
- [9] S.Makino, Y.Kaneda, 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.
- [10] H.L.Nguyen Thi and C.Jutten, "Blind source separation for convolutive mixtures", Signal Processing, vol.45, no.2, pp.209-229, March 1995.
- [11] S.Haykin, Adaptive Filter Theory, 3rd ed., Prentice-Hall, Inc. 1996.

- [12] A.Cichocki,S.Amari,M.Adachi,W.Kasprzak,"Self-adaptive neural networks for blind separation of sources",Proc.ISCAS'96,Atlanta,pp157-161,1996.
- [13] 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.
- [14] 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.IV2093-2112,May 1998.
- [15] S.Cruces and L.Castedo,"A Gauss-Newton methods for blind source separation of convolutive mixtures",IEEE Proc.ICASSP'98 Seattle,pp.IV2093-2096,May 1998.
- [16] K.Nakayam,A.Hirano and M.Nitta,"A constraint learning algorithm for blind source separation",IEEE INNS,Proc.IJCNN'2000,Como,Italy,pp.24-27,July,2000.
- [17] L.Parra and C.Spence, "Convolutive blind separation of nonstationary source," IEEE Trans. Speech Audio Processing, vol.8, pp.320-327, May 2000.
- [18] K.Matsuoka and S.Nakashima, "Minimal dsitortion principle for blind source separation," Proc. ICA2001, pp.722-727, 2001.
- [19] K.Nakayam,A.Hirano and T.Sakai,"A pairchannel learning algorithm with constraints for multichannel blind separation",IEEE INNS,Proc.IJCNN'01,Washington DC,July,2001.
- [20] 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.
- [21] 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.
- [22] H.Mathis and S.C.Douglas,"On optimal and universal nonlinearities for blind signal separation",IEEE Proc.ICASSP'01,MULT-P3.3,May,2001.
- [23] 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.
- [24] 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.
- [25] 堀田 明秀, 中山 謙二, 平野 晃宏 "混合過程の伝搬遅延を考慮した畳み込み形ブラインドソースセパレーション" 電子情報通信学会総合大会, 東京, Mar. 2002.

- [26] K.Nakayama,A.Hirano and A.Horita,"A Learning algorithm for convolutive blind source separation with transmission delay constraint",IEEE INNS,Proc.IJCNN'02,May 2002.
- [27] 堀田 明秀, 中山 謙二, 平野 晃宏 "反響音を含む畳み込み形ブラインドソースセパレーションにおけるフィルタ次数とステップサイズの検討" 電気関係学会北陸支部連合大会, 福井, Sep. 2002.
- [28] 堀田 明秀, 中山 謙二, 平野 晃宏 "A Learning Algorithm with Adaptive Exponential Stepsize for BSS of Convolutional Mixtures with Reverberations" 電子情報通信学会総合大会, 仙台, Mar. 2003.
- [29] S.Araki, R.Mukai, S.Makino, T.Nishikawa, H.Saruwatari, "The fundamental limitation of frequency domain blind source separation for convolutional mixtures of speech," IEEE Trans. Speech and Audio Processing, vol.11, no.2, pp.109-116, March 2003.
- [30] K.Nakayama, A.Hirano and A.Horita, "A learning algorithm with adaptive exponential stepsize for blind source separation of convolutional mixtures with reverberations",IEEE&INNS Proc. IJCNN'03 Portland, Oregon, July 2003.
- [31] K. Nakayama, A.Hirano and T.Sakai, "An adaptive nonlinear function controlled by estimated output pdf for blind source separation", 4th International Symposium on Independent Component Analysis and Blind Source Separation (ICA2003), Nara, Japan, pp.427-432, April 2003.
- [32] 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, Oregon, July 2003.
- [33] 堀田 明秀, 中山 謙二, 平野 晃宏 "反響音を含む畳み込み形 BSS における重み付きステップサイズ制御の一方法" 電気関係学会北陸支部連合大会, 富山, Sep. 2003.
- [34] 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 の学習法ー学習率に重みを付けて制御する方法ー", 第 18 回デジタル信号処理シンポジウム, 志摩, Nov. 2003.
- [35] 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, Mar. 2004.
- [36] H.Saruwatari, T.Takatani, H.Yamajo, T.Sishikawa and K.Shikano, "Blind separation and deconvolution for real convolutional mixture of temporally correlated acoustic signals using SIMO-model-based ICA", ICA'03 pp.549-554 April 2003
- [37] テレコミュニケーションホームページ http://www.telecomi.biz/pdfs/tc_0408_part3.pdf

- [38] 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形 BSS の収束性に関する比較”, 信号処理研究会, pp.57-62 2004,7
- [39] K.Nakayama, A.Hirano and Y.Dejima, "Analysis of signal separation and distortion analysis in feedforward blind source separation for convolutive mixture", IEEE Proc. The 47th International Midwest Symposium on Circuits and Systems (MWSCAS2004), Hiroshima, Japan, pp.III207-III210, Sept. 2004.
- [40] 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 Precessing, Yokohama, Japan, pp.523-528, Aug. 2004.
- [41] 堀田 明秀, 中山 謙二, 平野 晃宏 “時間領域でのフィードフォワード形ブラインドソースセパレーションの収束性解析” 電気関係学会北陸支部連合大会, 金沢, Sep. 2004.
- [42] 堀田 明秀, 中山 謙二, 平野 晃宏 “多種音源に対するブラインド信号源分離装置の開発” 金沢大学ベンチャービジネスプランコンテスト, 金沢, Oct. 2004.
- [43] 堀田 明秀, 中山 謙二, 平野 晃宏 “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”, 第 19 回信号処理シンポジウム, ハケ岳, Nov. 2004.
- [44] 堀田明秀, 中山謙二, 平野晃宏, “フィードフォワード形ブラインド信号源分離における信号歪み抑制学習法”, 電子情報通信学会, 信学技法, SIP 研究会 (名古屋), SIP2004-111, pp.7-12 2005,1
- [45] 堀田明秀, 中山謙二, 平野晃宏, ” フィードフォワード形およびフィードバック形ブラインド信号源分離における信号分離および信号歪の解析”, 電子情報通信学会, 信学技法, SIP 研究会 (仙台) , SIP2005-29, pp.25-30, 2005.6.
- [46] A.Horita, K.Nakayama, A.Hirano and Y.Dejima, "Analysis of signal separation and signal distortion in feedforward and feedback blind source separation based on source spectra", IEEE&INNS, Proc., IJCNN2005, Montreal, pp.1257-1262, July-Aug.
- [47] 出島康宏, 堀田明秀, 中山謙二, 平野晃宏, “ 周波数領域フィードフォワード形ブラインド信号源分離における信号歪みを抑制する学習法”, 電子情報通信学会, 信学技法, Vol.105 No.349 pp.37-42, SIP2005-102, 2005.10.
- [48] 堀田 明秀, 中山 謙二, 平野 晃宏 , 出島康宏, “Anarysis of A Learning Algorithm with Distortion Free Constrant for Convolutive Blind Source Separation in Time Domain”, 第 20 回信号処理シンポジウム, 高知, Nov. 2005.
- [49] 出島康宏, 堀田 明秀, 中山 謙二, 平野 晃宏, “周波数領域フィードフォワード形ブラインド信号源分離における信号歪み抑制法の性能解析”, 第 20 回信号処理シンポジウム, 高知, Nov. 2005.

- [50] A. Horita, K. Nakayama, A. Hirano and Y. Dejima, "A distortion free learning algorithm for feed-forward BSS with convolutive mixture and multi-channel signal sources", Technical Report of IEICE, Vol.SIP2006-15, pp.17-22, 2006.5.
- [51] 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.
- [52] 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, July 2006.
- [53] 堀田明秀, 中山謙二, 平野晃宏, " 畳み込み形 BSS における伝達関数の遅延差による影響", 電子情報通信学会 2006 年ソサイエティ大会, 金沢, A-4-34, p101, 2006-9.
- [54] 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, 2006.11.
- [55] 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", Journal of Signal Processing Vol. 11 No. 3, 2007.5 掲載決定