

エコーキャンセラ用信号分離回路の置換問題に関する検討

坂井良広*

Study on a permutation problem of Blind Source Separation

Yoshihiro Sakai*

Substitution problem in blind signal isolating circuit was examined on the system in the echo canceller that blind signal separation proposed in advance was pretreated. As FDSOS using actual sound condition and voice sample the algorithm of the blind signal separation the basic performance it evaluates. Next, the combination of the parameter that it arranges parameter in the algorithm and relationship between the required condition variously and maximizes the separating function by the experiment it investigates investigate.

Key words: echo canceller, blind source separation, independent component analysis

1 はじめに

音響的な回り込み妨害を軽減する手段としてエコーキャンセラが用いられる^{1),2)}。通常実用されているエコーキャンセラでは適応アルゴリズムによりエコーを推定し、エコー消去を行う。この際、受信信号と同時に送話信号が発声(ダブルトーク状態)されると適応の精度が低下し相手側にエコーを伝えてしまうという問題点がある。このようなエコー増大を抑制し、ダブルトーク状態でも適応動作を継続し音響経路変動への対応可能なことが望ましい。これらを同時に実現する手法のひとつとして、先にブラインド信号分離回路を用いたエコーキャンセラ方式を提案した³⁾。本稿ではここで用いるブラインド信号分離回路における置換問題について検討した。ブラインド信号分離のアルゴリズムはFDSOSとし、実際の音響条件、音声サンプルを用いて基本性能を評価した。次にアルゴリズムにおけるパラメータと各種要求条件の関係について整理し、実験によって分離性能を最大化するパラメータの組合せを調査した。さらに周波数領域における分離行列の置換不整合の影響を評価した。

2 研究の背景

2.1 エコーキャンセラの前処理に用いるブラインド信号分離回路

近年、様々な形態の音声通話の利便性向上のため、TV会議システムやハンズフリー電話機等が発達している。これらのシステムでは音響的まわりこみを防止し通話品質を確保するために、音響エコーキャンセラは必須である。エコーキャンセラでは近端話者のダブルトーク時に適応精度が劣化しエコーが増大する問題が生じるので、その対策が必要である。そのため先にブラインド信号分離回路を前処理とするエコーキャンセラシステムを提案した³⁾。その構成をFig.1に示す。図の

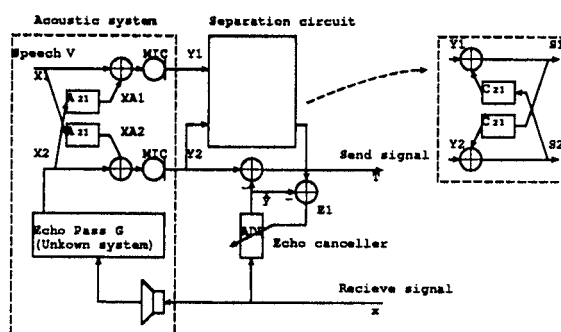


Fig.1 Echo Canceller using Separation Circuit

左側点線内が会議室音響系の等価回路、右側点線内がブラインド信号分離回路、ADFがエコーキャンセラである。会議室にはスピーカを1個、マイクを2個配置する。 $G(z)$ はスピーカからマイクまでの伝達関数、 $A_{12}(z)$ および $A_{21}(z)$ はそれぞれ、エコー経路および話者から反対側マイクまでの伝達関数である。2個のマイクはどちらにもエコーと話者音声とが混合されるが、これを右側のブラインド信号分離回路で処理す

ることにより、端子 $S_2(z)$ にはエコー信号だけが出力される。これをエコーキャンセラに入力すると、ダブルトークに影響されることなく、エコー経路 $G(z)$ を精度良く推定し、良好なエコー消去を実現できる³⁾。

2.2 提案構成における置換問題

ところで提案したシステムでは、ブラインド信号分離回路の $S_1(z)$ 端子に話者音声、 $S_2(z)$ 端子にエコー信号が分離して出力されることを期待しているが、ブラインド信号分離回路では分離規範に信号の独立性だけを用いているため、一般的には出力がどちらになるか不定となる置換問題が存在する⁴⁾。提案システムにおいて $S_2(z)$ に期待した信号と逆の信号が出力されると、エコーキャンセラの動作に大きな悪影響をおよぼす。そこで、提案したシステムにおけるブラインド信号分離回路の置換問題について検討した。

3 定式化

以下、ブラインド信号分離 (BSS) に絞って議論を進める。音響モデルによれば、 I 個の統計的に独立な音声信号 $s_i(n)$, $i = 1, \dots, I$ が J 個のマイクで受信される場合、受信信号は下記のように混合され畳み込みされたものになる。

$$x_j(n) = \sum_{i=1}^I \sum_{p=0}^{P-1} h_{ji}(p) s_i(n-p), \quad j = 1, \dots, J \quad (1)$$

ここで $h_{ji}(n)$ は音源 i からマイク j への P 点インパルス応答である。ベクトル表記に書き換えると、

$$\begin{aligned} \mathbf{x}(n) &= \sum_{p=0}^{P-1} \mathbf{H}(p) \mathbf{s}(n-p) \\ &= \mathbf{H}(n) * \mathbf{s}(n) \end{aligned} \quad (2)$$

ここで $\mathbf{X}(n) = [x_1(n), \dots, x_J(n)]^T$ は $(J \times 1)$ 個の受話信号、 $\mathbf{H}(n)$ は $(J \times I)$ の混合フィルタ行列、 $\mathbf{s}(n) = [s_1(n), \dots, s_I(n)]^T$ は $(I \times 1)$ の音源信号ベクトルであり、 $*$ は畳み込みを表し、 $[\cdot]^T$ は転置操作を表す。BSS の目的は I 個の音源を分離する長さ Q の分離フィルタ $\mathbf{W}(n)$ を見いだすことである。すなわち

$$\hat{\mathbf{s}}(n) = \mathbf{W}(n) * \mathbf{X}(n) \quad (3)$$

ここで $(\hat{\cdot})$ は推定を表す。

ブラインド信号分離の要点は信号源の独立性にある。一般の信号源の場合は独立性を扱うために高次統計量を用いるが、これは計算量が多いうえに推定値にばらつきが存在する。一方音声信号の場合は非定常な性質を合わせ持つので、ブラインド信号分離は二

次統計量で十分であり、比較的計算量が少なくてすむ利点がある。統計的に独立で非定常な信号 $s_i(n)$, および $s_j(n)$ については、

$$E\{s_i(n)s_j^*(n-\tau)\} = 0, \forall n, \tau, i \neq j \quad (4)$$

が成り立つ。ここで $E\{\cdot\}$ は期待値、 $(\cdot)^*$ は複素共役を表す。

式 (3) に示すようにブラインド分離は時間領域では畳み込みのため複雑な問題であるが、周波数領域に変換することにより、各帯域ごとの単純な問題に置き換えられる。周波数領域手法は、時間領域手法に比べて計算複雑性を大きく減少させる⁵⁾。

4 周波数領域におけるブラインド音源分離

簡単のため、図2に示すように2チャンネルの場合を考える。

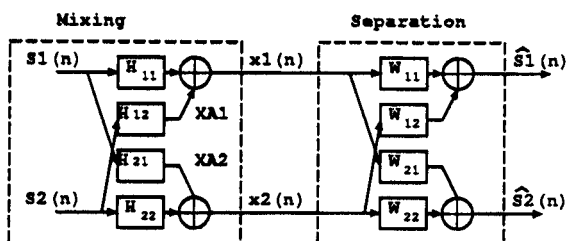


Fig.2 (2x2) convolutive blind source separation problem

出力 $x_1(n)$ と $x_2(n)$ は下記のように与えられる。

$$x_1(n) = h_{11}(n) * s_1(n) + h_{12}(n) * s_2(n) \quad (5)$$

$$x_2(n) = h_{21}(n) * s_1(n) + h_{22}(n) * s_2(n) \quad (6)$$

式 (2) における時間領域畳み込み混合 $\mathbf{x}(n)$ は、 T 点短時間フーリエ変換 (STFT) を計算することにより、周波数領域における瞬時混合に変換することができる。

$$\mathbf{x}(\omega, m) \approx \mathbf{H}(\omega) \mathbf{s}(\omega, m) \quad (7)$$

ここで m はブロックのインデックスであり、相関を持つ要素の畳み込みへの影響を避けるために $T \geq 2P$ とする。 P は混合フィルタの長さである。受信データについて下記の計算を行う。

$$\mathbf{x}(\omega, m) = \sum_{\tau=0}^{T-1} w(\tau) \mathbf{x}(\beta T m + \tau) e^{-j2\pi\omega\tau/T} \quad (8)$$

ここで、 $\omega = 1, \dots, T$, および $m = 0, \dots, (N - T)/(\beta T)$ とし、 $w(\tau)$ は窓関数であり、 $\beta(0 < \beta \leq 1)$

はデータのオーバーラップ係数である。ところで共分散行列 $\mathbf{R}_x(\omega, k)$ は、受信データのエルゴート性を仮定すれば、下記の式を用いて推定できる。

$$\hat{\mathbf{R}}_x(\omega, k) = \frac{1}{M} \sum_{m=0}^{M-1} \mathbf{x}(\omega, Mk + m) \mathbf{x}^H(\omega, Mk + m) \quad (9)$$

ここで $k = 0, \dots, K-1$ である。式 (9) において、スーパーブロック (M 個のブロックの連続) の番号 k における共分散行列を得るために、周波数領域データは $M = [N + T(\beta - 1)] / (K\beta T)$ の連続ブロックに渡って、オーバーラッピングしながら平均する。音源信号が互いに独立であるという仮定により、下記のような共分散行列の同時対角化を行って、推定された音源信号 $\hat{s}_1(n)$ と $\hat{s}_2(n)$ を無相関化する分離フィルタ行列を得る。

$$\Lambda_{\hat{\mathbf{S}}}(\omega, k) = \mathbf{W}(\omega) \hat{\mathbf{R}}_x(\omega, k) \mathbf{W}^H(\omega), \quad k = 1, \dots, K \quad (10)$$

各周波数における K 個の行列を連係して対角化する $\mathbf{W}(\omega)$ を見いだすことが必要である。

一般的な信号の場合は式 (10) の 2 次無相関化基準は $\mathbf{W}(\omega)$ を解くための十分な条件を単独では決定できない。しかし、非定常な信号については、 K 個の十分に離れた時間間隔について、独立な無相関化 (10) を実現することができる⁶⁾。各周波数帯 $\omega (\omega = 1, \dots, T)$ についての分離フィルタ $\mathbf{W}(\omega)$ (これは k 個の無相関化方程式を同時に満足する) は、下記の最小二乗解を用いて得ることができる⁶⁾。

$$\hat{\mathbf{W}}(\omega) = \arg \min_{\mathbf{W}(\omega)} \sum_{k=1}^K \|\mathbf{V}(\omega, k)\|^2 \quad (11)$$

ここで $\|\cdot\|^2$ は二乗ノルムであり、誤差は下記のとおりである。

$$\mathbf{V}(\omega, k) = \mathbf{W}(\omega) \hat{\mathbf{R}}_x(\omega, k) \mathbf{W}^H(\omega) - \text{diag}[\mathbf{W}(\omega) \hat{\mathbf{R}}_x(\omega, k) \mathbf{W}^H(\omega)] \quad (12)$$

ここで $\text{diag}(\cdot)$ は、行列要素の対角要素を抽出して作られる対角行列である。式 (11) に対する最小解は最急降下アルゴリズムを用いて得られる。

$$\begin{aligned} \mathbf{W}^{(l+1)}(\omega) &= \mathbf{W}^{(l)}(\omega) \\ -\mu(\omega) \frac{\partial}{\partial \mathbf{W}^{(l)*}(\omega)} \times \left\{ \sum_{k=1}^K \|\mathbf{V}^{(l)}(\omega, k)\|^2 \right\} \end{aligned} \quad (13)$$

ここで $\omega = 1, \dots, T$ である。文献⁶⁾を参考に、下記のステップサイズを用いる。

$$\mu(\omega) = \frac{\alpha}{\sum_{k=1}^K \|\hat{\mathbf{R}}_x(\omega, k)\|^2} \quad (14)$$

ここで α は正規化されたステップサイズである。各反復において、 $\mathbf{W}(\omega)$ の非対角要素だけを更新する。対角要素はそれらの初期値に保持する。この BSS アルゴリズムを以下周波数領域の二次統計量 (FDSOS) 方式と呼ぶ。これを用いて置換不整合問題を調査した。

5 実験条件および評価尺度

様々な動作条件で実際の音響環境において FDSOS アルゴリズムを実験した。各実験において、データベースの音声サンプルを用いた。オリジナルの音声データは 16kHz サンプルであるが、実験で用いるために 8kHz に間引きした。信号 $s_1(n)$ と $s_2(n)$ は、表 1 にリストしたように、異なる女性話者 (f) および男性話者 (m) によって読まれた文章である。

Table 1 Speech samples

$s_1(n)$ 入力	$s_2(n)$ 入力
f15	m33
f14	m10
f14	m13
f17	m08

音声データは、51.2s の録音である。すなわち、8kHz サンプルレートにおいて $N = 409600$ サンプルである。このデータベースにおける全ての文は、同じ信号強度になるように概略正規化されている。

直接チャネル $[h_{11}(n), h_{22}(n)]$ および交差チャネル $[h_{12}(n), h_{21}(n)]$ のインパルス応答は、Fig.3 のように実際の部屋で測定したものである。

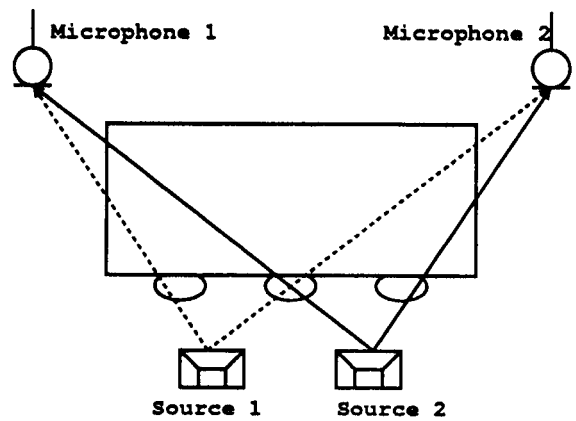


Fig.3 Room configuration where the actual impulse response were measured

部屋前方の左側と右側に 1 対の無指向性マイクがあり、部屋後方中央 (音源 1) と右位置 (音源 2) にスピーカを置いた。

性能評価のため下記の尺度を定義する。

- (1) 音源 1 に関する信号対妨害比 (SIR_1) および音源 2 に関する信号対妨害比 (SIR_2)

出力 1 における音源 1 の SIR は、音源 1 による所望信号のパワーと、音源 2 による非所望信号パワーとの比として定義する。2 つのマイクにおける受信信号については、その比は処理前 (入力音源 1) SIR と呼び、下記の式で表す。

$$SIR_{1i} = \frac{\sum_{\omega=1}^T |H_{11}(\omega)|^2 |s_1(\omega)|^2}{\sum_{\omega=1}^T |H_{12}(\omega)|^2 |s_2(\omega)|^2} \quad (15)$$

受信信号を処理した後の改善量を比較するため、処理後 (出力音源 1) SIR を下記のように定義する。

$$SIR_{1o} = \frac{\sum_{\omega=1}^T |A_{11}(\omega)|^2 |s_1(\omega)|^2}{\sum_{\omega=1}^T |A_{12}(\omega)|^2 |s_2(\omega)|^2} \quad (16)$$

ここで

$$\mathbf{A}(\omega) = \mathbf{W}(\omega)\mathbf{H}(\omega) \quad (17)$$

は総合的な伝達関数である。同様に音源 2 に関する SIR を定義する。

- (2) 平均信号対妨害比 (SIR)

平均 SIR は直接チャネル経由で得られた全体の信号パワーと、交差チャネル経由で受信された全体の信号パワーとの比として定義する。処理前では、下記のように定義する。

$$SIR_i = \frac{\sum_{\omega=1}^T \sum_{i=1}^2 |H_{ii}(\omega)|^2 |s_i(\omega)|^2}{\sum_{\omega=1}^T \sum_{i=1, i \neq j}^2 \sum_{j=1}^2 |H_{ij}(\omega)|^2 |s_j(\omega)|^2} \quad (18)$$

同様に、処理後の平均 SIR は下記のように定義する。

$$SIR_o = \frac{\sum_{\omega=1}^T \sum_{i=1}^2 |A_{ii}(\omega)|^2 |s_i(\omega)|^2}{\sum_{\omega=1}^T \sum_{i=1, i \neq j}^2 \sum_{j=1}^2 |A_{ij}(\omega)|^2 |s_j(\omega)|^2} \quad (19)$$

BSS 方式の目標は大きな SIR 改善を実現することである。SIR の改善とは、すなわち入力 SIR に対する出力の比である。

6 FDSOS 方式の基本的な性能

FDSOS アルゴリズムの基本性能を評価した。実験には音声サンプルペア (fl5, m33) を用いた。これらの音声はインパルス応答と畳み込みを行い、式 (5) と (6) によって混合信号を生成する。ステップサイズは各周波数帯ごとに式 (14) に従って、別々に選択し、 α は正規化された値 $\alpha = 0.5$ を用いた。初期値として $\mathbf{W}^{(0)} = \mathbf{I}_2$ を用いた。 $\mathbf{I}_2$ は、 (2×2) 単位行列である。アルゴリズムにおいては計算を 100 回反復実行した。関連推定値の計算に式 (9) を用いた。また式 (8) の

FFT は $T/2$ サンプルのオーバーラップ ($\beta = 0.5$) を持つハミング窓を用いて計算した。

FFT 計算において、通常は $T = 2P$ を選んだ。その目的は、式 (2) における線形畳み込みを循環畳み込みで近似するためである。しかし、もし P が大きければ、FFT はデータの定常な期間において計算しなければならないという仮定と矛盾する。 $P = 2048$ は 256ms に相当し、256ms 以上の FFT 窓長は一般的に音声データの定常仮定に反する。種々の T の値を用い、二つの音源について平均した FDSOS アルゴリズムの性能を評価した。その結果を Table 2 に示す。

Table 2 FDSOS の性能

T	Q	M	K	SIR_i (dB)	SIR_o (dB)
1024	1024	100	4	1.36	1.28
2048	2048	50	4	1.37	1.04
4096	4096	30	3	1.36	1.65
8192	8192	25	3	1.36	1.58

各場合においては、 M と K の種々の可能な組み合わせを試み、最良なものを選んだ。すべての場合において、分離フィルタ Q の長さは T に等しくなるように選んだ。

T, M , および K のどの組み合わせでも目に見えた分離性能にならないことがわかる。 $T = 1024$ と 2048 については、処理後 SIR は処理前 SIR さえも下回る。

7 置換不整合問題

周波数領域での信号分離は下記のように表現される。

$$\begin{bmatrix} \hat{s}_1(\omega) \\ \hat{s}_2(\omega) \end{bmatrix} = \begin{bmatrix} W_{11}(\omega) & W_{12}(\omega) \\ W_{21}(\omega) & W_{22}(\omega) \end{bmatrix} \begin{bmatrix} x_1(\omega) \\ x_2(\omega) \end{bmatrix} \quad (20)$$

周波数帯 ω における分離行列 $\mathbf{W}(\omega)$ のブラインド推定は、スケールと置換の組合せで表現される⁷⁾。すなわち、

$$\hat{\mathbf{W}}(\omega)\mathbf{H}(\omega) = \mathbf{P}(\omega)\hat{\mathbf{D}}(\omega) \quad (21)$$

ここで $\mathbf{P}(\omega)$ は置換行列、 $\mathbf{D}(\omega)$ は対角行列で、両方とも、 (2×2) 次元の行列である。 $\hat{\mathbf{W}}(\omega)$ は各周波数帯で独立に得られるので、回復された音源信号 $\hat{s}_i$ は、必ずしも、各周波数帯における推定 s_i の首尾一貫した推定であるとは限らない。この場合、すべての周波数帯に渡って置換を整列させるために何らかの方式を分離アルゴリズムに組み込まなければならない。

Para および Spence は分離フィルタ $W_{ij}(n)$ の長さを $Q < T$ となるように制約することにより置換問題

を解くことを提案した [6]。このような時間領域における有限長制約は、周波数領域において解を連続的にする。しかし分離フィルタの長さを制限することは音源を再構成するために利用可能な次数を減少させる。このためトレードオフの課題が発生する。

分離性能に対する制約 $Q < T$ の影響を検討する。5章で述べたように、混合過程 (式 2) を時間から周波数領域に変換する時に、 $T < 2P$ は循環畳み込み誤差をもたらすため、音声については $256\text{ms}(T = 2048)$ 以上の STFT 窓区間は定常仮定に違反する。このトレードオフは、付加的な制約 $Q < T$ については、より複雑である。良好な分離性能のために、理想的には Q は少なくとも P と同じ長さであるべきである。しかし $Q < T$ については、 T は P と等しいかそれ以上もあり得る。 T がとることのできる値と、満すべき要求との関係を Table 3 に示す。

Table 3 FDSOS における要求条件

要求条件	$T \leq P$	$P < T < 2P$	$T \geq 2P$
畳み込み条件	×	×	○
定常仮定 ($P = 2048$)	○	×	×
Q は P 以上	×	○	○

これらの要求を同時に満たす T の値は存在しない。そこで、各種パラメータ値に対する分離性能を評価し、最良となる設定条件を調査した。

Fig.3 に示した音響環境で、 $P = 2048$ として、様々な T, Q , およびブロック長 M について実験した。平均 SIR 性能を Table 4 に示す。

表にあるように下記の場合に性能は劣化する。

- 1) FFT 長 T がフィルタ長 $P(2048)$ より短い時、
 - 2) T が定常仮定に反するようになかなか長い時、
- $T = 2048$ が最良の性能を与えることがわかった。最良の組み合わせでは、約 3.5dB の SIR 改善量になった。これらすべての場合において、 $W(n)$ の長さを $Q = 500$ に制限することによって置換制約が適用される時に最良の SIR 改善を得られことがわかった。これは $P = 2048$ サンプルの混合フィルタが用いられても成り立つ。

8 置換整合誤りの影響

FDSOS アルゴリズムは、分離フィルタの長さを時間領域において制約することにより、周波数に渡る分離フィルタ行列の置換を整理させようと試みるものである。これは周波数領域において円滑性を向上させるが、限界があり、また Table 4 から明らかなように、大域的最適値に導くには不十分である。

Table 4 FDSOS の性能

T	Q	M	K	$SIR_i(\text{dB})$	$SIR_o(\text{dB})$
1024	300	100	4	1.36	4.48
1024	500	100	4	1.36	4.66
1024	500	200	2	1.36	4.66
1024	800	100	4	1.36	4.06
1024	1024	100	4	1.36	1.28
2048	300	50	4	1.37	4.39
2048	500	50	4	1.37	4.72
2048	500	65	3	1.37	4.81
2048	500	100	2	1.37	4.83
2048	800	50	4	1.37	4.62
2048	1000	50	4	1.37	4.57
2048	1500	50	4	1.37	4.24
2048	2048	50	4	1.37	1.04
4096	300	30	3	1.36	4.43
4096	500	30	3	1.36	4.66
4096	500	35	2	1.36	4.33
4096	1000	30	3	1.36	4.31
4096	1500	30	3	1.36	4.29
4096	2500	30	3	1.36	4.26
4096	3500	30	3	1.36	3.82
4096	4096	30	3	1.36	1.65
8192	2000	20	2	1.36	4.10
8192	8192	15	3	1.36	1.58

そのために、下記の実験を行う。実験では、アルゴリズムが一旦収束した後で置換を整理させる操作を行う。その際、比較評価のため、整理前後の SIR を計算する。

検討条件は、 $T = 2048, M = 50, K = 4$, 音声ペア I(f15,m33) とした。置換制約は緩和した。すなわち $Q = P = 2048$ とした。100 回反復計算後の収束時点で各周波数帯において $W(\omega)$ の置換を整理した。これは下記に定義するフリップ関数により実行した。

- (1) 下記の式を用いて音源 j による再構成信号 $i(i, j = 1, 2)$ について、各周波数において SIR_{ij} を計算する。

$$SIR_{11}(\omega) = \frac{|A_{11}(\omega)|^2 |s_1(\omega)|^2}{|A_{12}(\omega)|^2 |s_2(\omega)|^2}$$

$$SIR_{12}(\omega) = \frac{|A_{21}(\omega)|^2 |s_1(\omega)|^2}{|A_{22}(\omega)|^2 |s_2(\omega)|^2}$$

$$SIR_{21}(\omega) = \frac{|A_{12}(\omega)|^2 |s_2(\omega)|^2}{|A_{11}(\omega)|^2 |s_1(\omega)|^2}$$

$$SIR_{22}(\omega) = \frac{|A_{22}(\omega)|^2 |s_2(\omega)|^2}{|A_{21}(\omega)|^2 |s_1(\omega)|^2}$$

- (2) 下記の二値基準によって各周波数において、分離フィルタ $W(\omega)$ をフリップする。

●音源 1

If $SIR_{12}(\omega) > SIR_{11}(\omega)$ → $W(\omega)$ の列を交換する

else

→ $W(\omega)$ を交換しないでおく

●音源 2

If $SIR_{21}(\omega) > SIR_{22}(\omega)$ → $W(\omega)$ の列を交換する

else

→ $W(\omega)$ を交換しないでおく

(3) 音源 1, 音源 2, および平均 SIR を計算する.
結果を Table 5 に示す.

Table 5 フリップ前後の SIR

信号源	入力 SIR	フリップ前	方法 1	方法 2
音源 1	-0.58	-1.34	8.31	8.31
音源 2	3.37	3.56	9.08	9.08
平均	1.37	1.04	8.81	8.81

方法 1 : $SIR_{12} > SIR_{11}$ 方法 2 : $SIR_{21} > SIR_{22}$

$SIR_{12} > SIR_{11}$, $SIR_{21} > SIR_{22}$ のどちらでも同じ結果が得られた。結果として、置換の整列により平均出力 SIR において $8.81 - 1.04 = 7.77\text{dB}$ の改善を得た。この最大到達可能 SIR は、理想性能と言える。

置換制約 $Q = 500$ を適用した場合も検討した。100 回反復終了時点後、制約しない場合と同じ SIR 比較方式を用いて分離フィルタ行列の置換を整列した。その場合の平均出力 SIR は 7.12dB であった。8.81dB から 7.12dB への減少は $W(n)$ のサイズを 500 サンプルに制限したためである。

9 おわりに

先に提案したブラインド信号分離を前処理とするエコーキャンセラにシステムに関し、ブラインド信号分離回路における置換問題について検討した。ブラインド信号分離のアルゴリズムは FDSOS とし、実際の音響条件、音声サンプルを用いて基本性能を評価した。次にアルゴリズムにおけるパラメータと各種要求条件の関係について整理し、実験によって分離性能を最大化するパラメータの組合せを調査した。さらに周波数領域における分離行列の置換不整合の影響を評価した。今後は置換の不整合状態を改善するアルゴリズムについて検討を行う。

参 考 文 献

- 1) 伊藤, 丸山, 古屋 : テレコンファレンス用音声伝送装置の開発, 昭和 59 信学通全大 (1984) No.591
- 2) 古川 : 室内音響エコー消去方式の一検討, 昭和 59 信学通全大 (1984) No.2345
- 3) 坂井良広, 高橋弘太, 岩倉博 : ブラインド信号分離回路を用いた適応フィルタに関する一検討, 第 15 回デジタル信号処理シンポジウム講演論文集, (2000)359-364
- 4) J.F.Cardoso : Blind signal separation; Statistical principles, Pccoc.IEEE, vol.88, no.10 (1998) 2009-2025
- 5) F.Ehlers, H.G.Schuster : Blind separation of convolutive mixtures and and applicatiuon in automatic speech recognition in a noisy environment, IEEE Trans. SP, 45, 10 (1997) 2608-2612
- 6) L.Para, C.Spence : Covolutive blind separation of nonstationary sources : IEEE Trans,SAP, 8, 3 (2000) 320-237
- 7) L.Tong, R.W.Liu, V.C.Sonn : Indetermiacy and identifiability of blind identification : IEEE Trans,CS 38, 5 (1991) 499-509