

## 信号源の組合せが変動するブラインド信号源分離の 回路構成と学習法

### A Circuit Structure and A Learning Algorithm for Blind Source Separation Applied to Variable Signal Source Combinations

磯端謙介                      中山謙二                      平野晃宏  
金沢大学大学院 自然科学研究科 電子情報工学専攻

Kensuke ISOBATA                      Kenji NAKAYAMA                      Akihiro HIRANO  
Division of Electrical and Computer Engineering,  
Graduate School of Natural Science and Technology, Kanazawa Univ.  
E-mail : nakayama@t.kanazawa-u.ac.jp

あらまし 信号源の数と組合せが変動する場合のブラインド信号源分離 (BSS) の回路構成と学習法を提案している。本方法では、瞬時混合を対象とし、「センサ数 = 信号源の最大数」という条件を設けている。信号源の数が最大の場合の信号源の組合せに対して一つの分離回路を用いる。分離回路の数は信号源の組合せの数と同じになる。信号源の数が最大数より少ない場合もこの分離回路を用いる。分離回路の学習法は、信号源の数が最大の時に学習するステップ1、及び、ステップ1の学習収束後に信号源の数が減少した状態で学習するステップ2から構成される。本方法では、複数の分離回路を用いるため、分離回路の出力側では、分離回路の選択とその回路における出力信号の選択が必要となる。本稿では、出力信号の相関係数と信号レベルに基づいて選択する方法を提案している。シミュレーションでは、4人の話者を想定し、同時に話す人数を1～3人として分離回路の学習、及び、分離性能の解析を行なった。その結果、学習においてはステップ2の学習が安定していること、分離性能においては正しい分離回路と出力信号が選択されていることが確認された。

**ABSTRACT** A circuit structure and a learning algorithm are proposed for Blind Source Separation (BSS) system applied to the following situation, the number and combination of signal sources are changed. The following conditions are assumed, an instantaneous mixing process and the number of sensors is equal to the maximum number of signal sources. A separation block is assigned to a combination of the signal sources, the number is maximum. A plural number of the separation blocks are used. When the number of signal sources

is less than the maximum, the same separation block, described above, is used. First, the separation block is trained by using the signal sources, the number is maximum (Step 1). After the learning in Step 1 converges, the separation block is trained by using the signal sources, the number is reduced (Step 2). A new learning algorithm is proposed for Step 2. Since the proposed method employs a plural number of the separation blocks, they must be selected, and also the output signals should be selected. A selection algorithm based on correlation coefficients and signal level of the output signals is proposed. In simulation, 4 speakers are employed. The number of the speakers, who speak at the same time, is changed from 1 to 3. The learning process in Step 2 is stable. The correct separation block and output signals are selected.

#### 1 はじめに

TV会議システムにおける話者分離、音響や通信システムにおける雑音除去、エコー除去、回線等化、信号の推定や復元といった信号処理が近年重要になってきている。しかし、信号や雑音に関する十分な情報が得られない場合がある。また、前もって混合過程や伝達過程を知ることができない場合も多い。このような状況下で、信号の統計的な性質を利用して信号を分離する手法としてブラインドソースセパレーション (BSS) がある。BSSでは、複数の未知信号が混合過程を経て混ざり合った混合信号を複数のセンサで観測し、分離回路を通すことにより元の信号を分離する。BSSでは「信号源は互いに統計的独立である」という仮定の下に回路構成や学習法が検討されている [1]-[3]。

一般にBSSでは、センサ数＝信号源の数とする場合が多い。しかし、実際の応用では信号源の数を推定できない場合や、途中で信号源の数が増える場合が考えられる。センサ数が信号源の数よりも少ないオーバーコンプリート形BSSも検討されているが難しい課題である [4]-[7]。

本稿では、信号源の数と組合せが変動する場合を想定し、かつ、瞬時混合を対象とし、「センサ数＝信号源の最大数」という条件下で動作するBSSを対象とする [8]-[10]。すなわち、「信号源の数 ≤ センサ数」という条件と、信号源の組合せによって混合過程が変化するという条件下で動作するBSSの回路構成と学習法を検討する。また、分離回路とその出力信号の選択方法も検討する。信号源として、話者を考えた場合、「話者の数 > センサ数」及び「同時に話す話者の最大数＝センサ数」という条件下で検討することに相当する。シミュレーションでは、4人の話者を想定し、同時に話す話者の数を1～3人として評価実験を行なう。

## 2 新しいBSSの回路構成と分離回路の構成

### 2.1 BSSの全体構成と入出力関係

信号源として、話者を考えた場合、本稿で提案する方式は「話者の数 > センサ数」及び「同時に話す話者の最大数＝センサ数」という条件を前提としている。また、「同時に話す話者の最大数 < 話者の数」なので、同時に話す最大数の話者の組み合わせは複数ある。これらの組み合わせに対して独立に分離回路を構成する。以下、回路構成を具体例で説明する。

フィードフォワード形の分離回路を用いて回路を構成する。混合過程は瞬時混合を想定し、信号の減衰のみを考え、遅延を考えないものとした。回路構成例を図1に示す。信号源(話者)の数を4個、センサ数を3個としている。同時に話す話者の数は1～3人(センサ数以下)で変動するものとする。

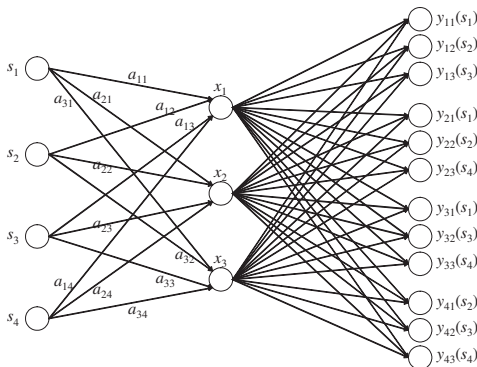


図1: 本稿におけるFF形BSSの回路構成

独立な信号源を  $s_k(n)$  ( $k = 1, 2, \dots, n$ ), 混合過程は瞬時混合であるとし、 $k$  番目の信号源から  $i$  番目のセンサへの伝達係数を  $a_{ik}$  とすると、センサへの入力信号  $x_i(n)$  ( $i =$

$1, 2, \dots, n$ ) は次のように表される。

$$x_i(n) = \sum_{k=1}^n a_{ik} s_k(n) \quad (\text{i.e. } \mathbf{x}(n) = \mathbf{A}\mathbf{s}(n)) \quad (1)$$

ここで  $\mathbf{A}$  は  $a_{ik}$  を要素とする混合行列である。このとき出力  $\mathbf{y}(n)$  は、

$$\mathbf{y}(n) = \mathbf{W}(n)\mathbf{x}(n) \quad (2)$$

$$= \mathbf{W}(n)\mathbf{A}\mathbf{s}(n) \quad (3)$$

$$= \mathbf{P}(n)\mathbf{s}(n) \quad (4)$$

となる。ここで  $\mathbf{W}(n)$  は分離行列である。分離行列  $\mathbf{W}(n)$  を学習させることにより、線形に混ざった信号  $\mathbf{A}\mathbf{s}(n)$  から信号源  $\mathbf{s}(n)$  を分離する。このとき式 (4) の  $\mathbf{P}(n)$  を分離成功の評価関数として用いる。もし  $\mathbf{P}(n)$  の各行各列に0でない要素を1つだけみつければ、信号源  $\mathbf{s}(n)$  は出力  $\mathbf{y}(n)$  に完全に分離されているといえる。

### 2.2 分離回路の構成

センサ数を  $K$  とした場合に  $K \times K$  の分離回路を構成する。例えば、図1において、 $s_1 \sim s_4$  を話者とし、かつ、同時に話す話者の数を1～3人とした場合、 $(s_1, s_2, s_3)$ ,  $(s_1, s_2, s_4)$ ,  $(s_1, s_3, s_4)$ ,  $(s_2, s_3, s_4)$  をそれぞれ分離するための分離回路  $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3, \mathbf{W}_4$  が構成される。この  $3 \times 3$  で構成した分離回路を「同時に話す話者の数 < センサ数」となった場合にも同様に用いる。例えば、 $(s_1, s_2, s_3)$  を分離できる回路  $\mathbf{W}_1$  は同時に話す話者が  $(s_1, s_2)$  や  $(s_1, s_3)$  となったときに分離を行なうためにも用いられる。

このように構成することによって、話者のすべての組合せに対して分離回路を構成するよりも、分離回路の個数を低減することができる。本構成法では計4個の回路で分離を行なうが、全組合せに対して構成した場合は11個もの分離回路が必要となり、回路規模が膨大になる。

## 3 学習アルゴリズム

### 3.1 分離行列の更新式

分離回路の学習には、式 (6) で示される Kullback-Leibler divergence を誤差関数とする自然勾配法を用いる。

$$KL(\mathbf{W}) = \int p(\mathbf{y}) \log \frac{p(\mathbf{y})}{\prod_{i=1}^m p(y_i)} d\mathbf{y} \quad (5)$$

$$= -H(\mathbf{y}; \mathbf{W}) + \sum_{i=1}^m H(y_i; \mathbf{W}) \quad (6)$$

このときの分離回路の更新分は、この K-L divergence を最小化するように更新される。分離回路の更新式は、

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \eta[\mathbf{I} - \langle \varphi(\mathbf{y}(n))\mathbf{y}^T(n) \rangle] \mathbf{W}(n) \quad (7)$$

である。(\*) は時間平均を表す。また、 $\varphi(\mathbf{y}(n))$  は活性化関数であり、式 (8) を用いる。

$$\varphi(\mathbf{y}(n)) = \tanh(0.1\mathbf{y}(n)) \quad (8)$$

### 3.2 初期学習とオンライン学習

この節では、学習すべき分離回路及び出力信号が正しく選択されたものとして、その学習法について説明する。学習する分離回路の選択については5節で述べる。

#### ステップ1：初期学習

「信号源(同時に話す話者)の数 = センサ数」となった場合に分離回路を式(7), (8)で更新する。

#### ステップ2：オンライン学習

初期学習が収束した後は、「信号源(同時に話す話者)の数 ≤ センサ数」の条件下で常時学習を行なう。話者の位置変動、伝達特性の変動等に対応するために常時学習を行なうことを想定している。

初期学習で構成された  $K \times K$  に対応する分離回路を、センサ数未満になった場合にも学習させるため、同時に話す話者の数に応じて以下のように係数の更新を行っていく。

i) 同時に話す話者の数 = 1 になった場合

$$\mathbf{W}_j(n+1) = \mathbf{W}_j(n) \quad (\text{学習しない})$$

ii)  $2 \leq$  同時に話す話者の数  $< K$  になった場合

$s_\alpha$  がなかったとき  $y_{j\alpha} = 0$  にした上で

$\mathbf{W}_j$  の  $\alpha$  行の学習率  $\eta$  を 0 にして更新する

iii) 同時に話す話者の数 =  $K$  になった場合

$\mathbf{W}_j$  を式(7)で更新 (初期学習と同じ)

同時に話す話者の数がセンサ数未満である場合(ii)に初期学習と同じ方法で学習を行なうと、信号の漏れ込みを含んで係数を更新してしまう。さらに、式(3)において  $s_i = 0$  とした場合に混合行列  $\mathbf{A}$  の第  $i$  列が 0 と見なされ、その逆行列を求めようとして分離回路の係数が  $\infty$  に発散してしまう可能性もある。したがって、同時に話す話者の数がセンサ数未満の場合に対しても初期学習と同じ手法を用いて係数更新を行なうことは難しい。

### 3.3 分離性能の評価

式(9)で表される SNR を用いて、分離性能の評価を行なう。SNR は高ければ高いほど分離性能は高くなる。

$$\text{SNR}_j = 10 \log_{10} \frac{\sum_u p_{juu}^2}{\sum_{u \neq v} p_{juv}^2} [\text{dB}] \quad (9)$$

また、各行毎に対しても SNR を求め、評価を行なう。

## 4 分離回路学習のシミュレーション

図1における  $(s_1, s_2, s_3)$  を分離する分離回路  $\mathbf{W}_1$  が選択された後の、分離回路の学習とその分離性能の解析を行なう。

### 4.1 シミュレーション条件

- ・ 音声1:  $s_1$  「人間は豊かな自然を求めています」
- ・ 音声2:  $s_2$  「パンはポルトガル人が伝えたものです」
- ・ 音声3:  $s_3$  「健康のため無理をしなで下さい」

(標本化周波数: 8kHz)

・ 学習における話者の組合せ:

$$\text{COM} = \begin{matrix} 123 & 12 & 13 & 23 & 1 & 2 & 3 \end{matrix}$$

$$\begin{pmatrix} s_1 & s_1 & s_1 & & s_1 & & \\ s_2 & s_2 & & s_2 & & s_2 & \\ s_3 & & s_3 & s_3 & & & s_3 \end{pmatrix} \quad (10)$$

・ 組合せ1つあたりのサンプル数: 40000 sample (5秒)

・ ステップサイズ:  $\eta = 0.0003$

・ 学習の時間平均に用いるサンプル数: 500 sample

・ 分離行列の初期値:  $\mathbf{W}_1 = \mathbf{I}$  ( $3 \times 3$  の単位行列)

・ 混合過程の係数:

$$\mathbf{A}_1 = \begin{pmatrix} 1.0 & 0.9 & 0.8 \\ 0.9 & 1.0 & 0.9 \\ 0.8 & 0.9 & 1.0 \end{pmatrix}$$

### 4.2 分離回路の初期学習と性能評価

まず、入力の組合せを  $\text{COM} = (123)$  として同時に話す話者の数が3の状態での学習をさせる。図2に各行のSNRと分離行列、図3に分離回路  $\mathbf{W}_1$  全体のSNRの推移をそれぞれ示し、表1に40000サンプル時点での値を示す。

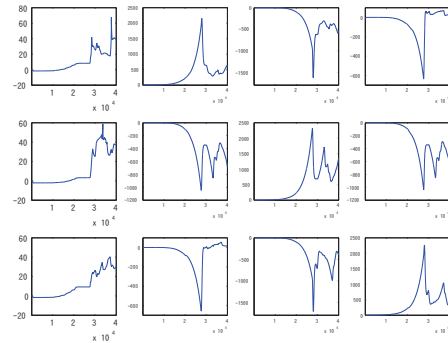


図2: 初期学習における各行のSNRと分離行列の推移

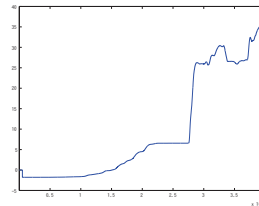


図3: 初期学習における分離回路  $\mathbf{W}_1$  全体のSNRの推移

表1: 初期学習における分離行列、伝達関数、SNRの値

$\mathbf{W}_1$			$\mathbf{P}_1$			$\text{SNR}_{1\alpha}$	$\text{SNR}_1$
570.5	-537.5	27.98	1	0.010	0.006	38.46	
-682.6	1356.4	-672.8	-0.001	1	0.013	37.53	35.38
20.58	-395.5	417.1	-0.022	-0.021	1	30.30	

図2-3において、25000サンプル付近は無音区間であるため分離回路の係数が大きく振れてしまっている。しかし、SNRで見れば値が維持されており、その後も増大していく傾向が見られる。また、表1より、伝達関数  $\mathbf{P}_1$  の対角要素と非対角要素(信号の漏れ込みを示す)を比較し

ても 50 倍程度の差が見られることから、分離は成功していると言える。

### 4.3 分離回路のオンライン学習と性能評価

この節では、初期学習を行なった後の学習の性能評価を行なう。オンライン学習として学習法を不変とした場合と提案法の場合の 2 種類のシミュレーションを行ない、提案法の有効性を確認する。

まず、COM = (123, 12) の組合せに対し、初期学習と同じ方法を用いて係数更新を行なった場合の各行の SNR と分離行列を図 4 と表 2 に示す。

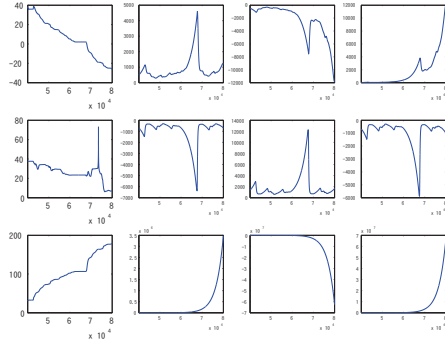


図 4: 学習法不変の各行の SNR と分離行列の推移

表 2: 学習法不変の分離行列, 伝達関数, SNR の値

$10^7 \times W_1$			$P_1$			$SNR_{1\alpha}$	$SNR_1$
0.001	-0.012	0.012	0.799	0.011	16.60	-26.34	
-0.001	0.002	-0.001	-0.003	1	-0.440	7.13	74.84
3.480	-62.65	66.13	$-9 \cdot 10^{-5}$	$4 \cdot 10^{-6}$	91752	179.7	

図 4 より、ほぼ無音が出力されるはずの  $y_{13}$  を大幅に更新し、3 人の信号を分離する係数からは逸れてしまっている。これは、式 (3) において  $s_3 = 0$  とした場合に  $A$  の第 3 列が 0 と見なされ、その逆行列を求めようとして  $\infty$  に発散してしまうためであると考えられる。したがって、同時に話す話者の数が 2 の場合にも初期学習と同じ手法を用いて係数更新を行なうことは難しい。

次に、提案法を用いて COM = (123, 12) の場合に学習した際の分離性能を評価する。図 5 に各行の SNR と分離行列の推移を、図 6 に分離回路  $W_1$  全体の SNR をそれぞれ示し、表 3 に 80000 サンプル時点での具体的な値を示す。

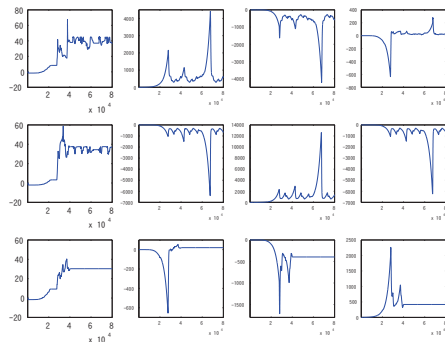


図 5: 提案法における各行の SNR と分離行列の推移

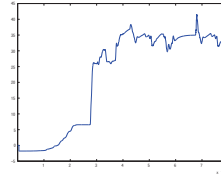


図 6: 提案法における分離回路  $W_1$  全体の SNR の推移

表 3: 提案法における分離行列, 伝達関数, SNR の値

$W_1$			$P_1$			$SNR_{1\alpha}$	$SNR_1$
569.0	-534.1	26.15	1	0.014	0.006	36.46	
-684.4	1358.5	-673.4	-0.003	1	0.013	37.37	34.94
20.58	-395.5	417.1	-0.022	-0.021	1	30.30	

40000 サンプルから同時に話す話者が  $(s_1, s_2)$  になるので、 $s_3$  に対応する出力  $y_3 = 0$  としている。更に、分離行列の第 3 行を学習しないために、第 3 行の  $\eta = 0$  として制御している。図 5 の下段の通り、40000 サンプルからは分離係数を更新していない。それによって信号の漏れ込みを用いた学習を抑制することができ、表 3 より  $SNR_{11}, SNR_{12}$  の値を保って学習し、かつ分離回路  $W_1$  全体としての  $SNR_1$  も保持していることを確認できる。

他の組合せにおいても提案法が有効であることを確認するために、COM = (123, 13) と COM = (123, 23) においても同様の学習を行なった。表 4, 表 5 に分離性能結果を示す。

表 4: 提案法における COM = (123, 13) での分離性能

$W_1$			$P_1$			$SNR_{1\alpha}$	$SNR_1$
570.3	-530.3	20.24	1	0.011	-0.007	37.72	
-682.6	1356.4	-672.8	-0.001	1	0.013	37.53	36.54
32.11	-406.5	417.7	0.006	-0.021	1	33.32	

表 5: 提案法における COM = (123, 23) での分離性能

$W_1$			$P_1$			$SNR_{1\alpha}$	$SNR_1$
570.5	-537.5	27.98	1	0.010	0.006	38.46	
-683.1	1358.2	-674.4	-0.001	1	0.012	38.61	36.29
18.24	-391.0	414.9	-0.022	-0.015	1	31.44	

表 1 の初期学習時に比べて、COM = (123, 13) と COM = (123, 23) ではいずれも分離性能が向上している。したがって、同時に話す話者の数がセンサ数未満の場合であっても提案法を用いて学習することにより、その分離性能を更に向上させられる可能性もあり、その有効性を確認することができた。

## 5 分離回路の選択

前節までは、1 つの分離回路に着目して、その学習法について述べてきた。図 1 に示す通り、本稿では分離回路を複数構成する必要がある。話者の数が 4 であり、同時に話す話者の最大数が 3 であるため、 $(s_1, s_2, s_3)$ ,  $(s_1, s_2, s_4)$ ,  $(s_1, s_3, s_4)$ ,  $(s_2, s_3, s_4)$  を分離する計 4 個の分離回路が構成される。

これらすべての組に対しての学習を同時に話す話者の数が 3 の状態で行なった後は、各分離回路における出力

レベルに閾値を設けて、必要な組合せでのみ相関係数を計算し、分離回路を選択するという手法を取る。

### 5.1 出力レベルの閾値

出力信号の過去  $N$  サンプルを用いて 2 乗平均を計算し、入力信号に比例した閾値を設ける。閾値を求める方法としては出力と同じサンプル数を用いて、各センサ  $x_i$  ごとの 2 乗平均を相加平均したものにスケーリング係数を乗算している。

$$y_{th} = Const. \times \frac{\overline{x_1^2} + \overline{x_2^2} + \overline{x_3^2}}{3} \quad (11)$$

### 5.2 相関係数

式 (11) の閾値を用いて、出力信号の独立性を評価するために、出力組  $(y_{j\alpha}, y_{j\beta})$  の相関係数を必要な組合せでのみ計算する。

$$cor_{j\alpha\beta} = \frac{\sigma_{y_{j\alpha}y_{j\beta}}}{\sigma_{y_{j\alpha}}\sigma_{y_{j\beta}}} \quad (\text{if } y_{j\alpha} > y_{th}, y_{j\beta} > y_{th}) \quad (12)$$

更に、分離回路  $j$  ごとに式 (12) の相関係数の絶対値平均を求める。

$$cor_j = \frac{1}{u} \sum |cor_{j\alpha\beta}| \quad (13)$$

ここで  $u$  は分離回路  $j$  で求めた相関係数の個数である。この  $cor_j$  が最も小さいものを分離回路  $j$  として選ぶことで、相関性が低い出力信号を得ることができる。

同時に話す話者の数が 1 であった場合は、相関係数  $cor_{j\alpha\beta}$  が計算されないか、もしくは  $cor_{j\alpha\beta} = \pm 1$  になるため、これを検出して出力の 2 乗平均の比が最も高かったものを分離回路として選択する。

その後、式 (12) により求めた各出力の相関係数を見ることによって出力信号を選択する。

## 6 分離回路及び出力信号選択のシミュレーション

### 6.1 シミュレーション条件

- ・音声 4 :  $s_4$  「太宰治は青森県で生まれました」
  - ・話者の組合せ :  
 $COM = (123, 124, 134, 234, 12, 13, 14, 23, 24, 34, 1, 2, 3, 4)$  (14)
  - ・組合せ 1 つあたりのサンプル数 : 25000 sample (約 3 秒)
  - ・相関係数算出に用いるサンプル数 :  $N = 5000$  [sample]
  - ・出力の閾値に用いるスケーリング係数 :  $Const. = 0.005$
- 他の条件は学習のシミュレーション条件と同一である。

### 6.2 分離回路の出力切替

式 (14) による組合せで信号源が入力されてきたときの出力切替についての評価を行なう。まず、分離回路の出力切替に用いる相関係数とその絶対値平均を表 6 に示す。

表 6 において、入力が  $(s_1, s_2, s_3)$  の場合、 $s_4$  に関する出力のレベルが閾値を下回るため  $(s_*, s_4)$  に対応する相関係数は計算しない。また、相関係数の絶対値平均を見る

と 0.0504 が一番小さいため、 $W_1$  を分離回路として選択する。そして、個々の相関係数に対応したところを見ると、 $y_{11}, y_{12}, y_{13}$  がそれぞれ  $s_1, s_2, s_3$  を分離するための出力として選択されている。入力が  $(s_1, s_2)$  の場合は、想定される回路選択は  $W_1, W_2$  の 2 通りが考えられる。いずれも相関係数の絶対値平均は低くなっているが、ここでは絶対値平均がより低い  $W_2$  が選択された。個々の相関係数を見ると  $y_{21}, y_{22}$  がそれぞれ  $s_1, s_2$  を分離するための出力として選択されている。

図 7 に各分離回路の相関係数の絶対値平均の推移を示す。上から順に  $cor_1, cor_2, cor_3, cor_4$  である。

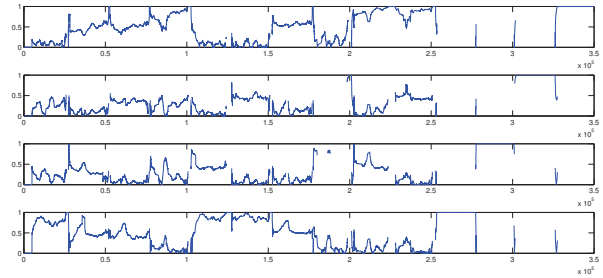


図 7: 各出力組における相関係数の絶対値平均の推移

この 4 つの相関平均を比較し、最も小さな値をとる分離回路が選択される。ただし、同時に話す話者の数が 1 (COM = 1, 2, 3, 4) であるときは、各分離回路における出力の 2 乗平均の比が最も大きいものが選択される。理想の回路切替を図 8 に、実際の回路切替を図 9 に示す。

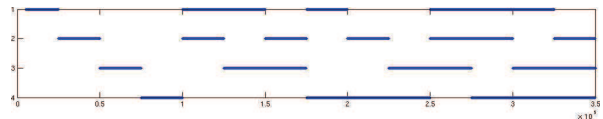


図 8: 理想の回路切替

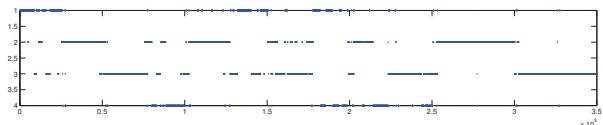


図 9: 実際の回路切替

図 8 と図 9 を比較すると、回路切替は概ねうまくできていると言える。しかしながら、図 8 と図 9 を比較すると、適切な分離回路の切替がされていない部分もあることがわかる。

### 6.3 回路切替のずれに対する考察

実際の回路切替 (図 9) の 25000 サンプルまでを拡大したものを図 10 に、それに対応する信号源  $(s_1, s_2, s_3)$  の波形を図 11 に示す。

表 6: 出力切替に用いる相関係数とその絶対値平均

出力	$W_1(s_1, s_2, s_3)$			$W_2(s_1, s_2, s_4)$			$W_3(s_1, s_3, s_4)$			$W_4(s_2, s_3, s_4)$			選択された回路
	$(s_1, s_2)$ $(y_{11}, y_{12})$	$(s_1, s_3)$ $(y_{11}, y_{13})$	$(s_2, s_3)$ $(y_{12}, y_{13})$	$(s_1, s_2)$ $(y_{21}, y_{22})$	$(s_1, s_4)$ $(y_{21}, y_{23})$	$(s_2, s_4)$ $(y_{22}, y_{23})$	$(s_1, s_3)$ $(y_{31}, y_{32})$	$(s_1, s_4)$ $(y_{31}, y_{33})$	$(s_3, s_4)$ $(y_{32}, y_{33})$	$(s_2, s_3)$ $(y_{41}, y_{42})$	$(s_2, s_4)$ $(y_{41}, y_{43})$	$(s_3, s_4)$ $(y_{42}, y_{43})$	
$(s_1, s_2, s_3)$	0.045	-0.027	0.079	-0.215			0.236			-0.777			$W_1$
		0.050			0.215			0.236			0.777		
$(s_1, s_2, s_4)$	-0.151	0.264	-0.887	0.030	-0.013	0.026	0.390	-0.093	-0.462	-0.865	0.232	-0.345	$W_2$
		0.434			0.023			0.315			0.485		
$(s_1, s_3, s_4)$	-0.319	0.209	-0.868	-0.445	-0.098	0.417	0.027	-0.013	0.031	-0.893	0.283	-0.168	$W_3$
		0.465			0.320			0.024			0.448		
$(s_2, s_3, s_4)$			-0.774			0.253			-0.233	0.081	-0.028	0.047	$W_4$
		0.774			0.253			0.233			0.052		
$(s_1, s_2)$	0.080			0.048			0.409			-0.864			$W_2$
		0.080			0.048			0.409			0.864		
$(s_1, s_3)$		0.005		-0.435			0.040			-0.893			$W_1$
		0.005			0.435			0.040			0.893		
$(s_1, s_4)$	-0.308	0.282	-0.999		0.005			0.006		-0.999	0.303	-0.331	$W_2$
		0.530			0.005			0.006			0.545		
$(s_2, s_3)$			0.100							0.101			$W_1$
		0.100								0.101			
$(s_2, s_4)$			-0.885			0.041			-0.453		0.008		$W_4$
		0.885			0.041			0.453			0.008		
$(s_3, s_4)$			-0.868			0.435					0.051	0.084	$W_3$
		0.885			0.435			0.453			0.008		

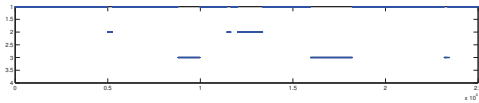


図 10: 実際の回路切替 (25000 サンプル拡大)

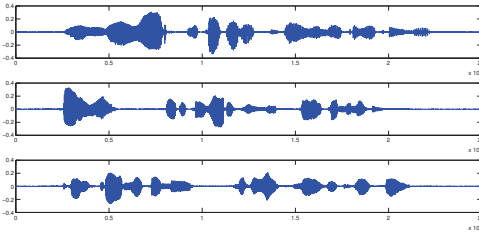


図 11: 信号源の信号波形

図 10 の 5000 サンプル付近で分離回路  $W_2$  が選ばれた理由について考察する。図 10 において 4000 サンプル付近から  $s_3$  の信号がほぼ無音になる。5000 サンプルで 2 乗平均を求め、相関を計算しているため、回路の切替は若干遅れる。出力の 2 乗平均のうち閾値を超えるものが  $(s_1, s_2)$  になるため、 $(s_1, s_2, s_4)$  の分離ができる分離回路  $W_2$  が選択されたと考えられる。また、出力の 2 乗平均が閾値を超えたとしても本来ほぼ無音であった  $s_3$  の区間には  $s_1$  と  $s_2$  の信号が漏れ込む。そこから相関係数を計算すると 1 に近い値が得られるため、結果として分離回路  $W_1$  が選ばれなくなったとも考えられる。

図 10 の 9000 サンプル付近に対しても同様であり、 $s_2$  が 6000 サンプル付近からほぼ無音になることにより、 $(s_1, s_3, s_4)$  の分離ができる分離回路  $W_3$  が選択されたと考えられる。

## 7 まとめ

本稿では、信号源の数と組合せが変動する場合のブラインド信号源分離 (BSS) の回路構成と学習法を提案した。瞬時混合を対象とし、「センサ数 = 信号源の最大数」という条件を設けている。信号源の数が最大の場合の信号源の

組合せに対して 1 つの分離回路を用いる方法である。分離回路の学習法は、信号源の数が最大の時に学習するステップ 1、及び、ステップ 1 の学習収束後に信号源の数が減少した状態で学習するステップ 2 から構成される。分離回路の選択とその回路における出力信号の選択は、出力信号の相関係数と信号レベルに基づいて行なっている。シミュレーションでは、4 人の話者を想定し、同時に話す人数を 1~3 人として分離回路の学習、及び、分離性能の解析を行なった。その結果、学習においてはステップ 2 の学習が安定していること、分離性能においては正しい分離回路と出力信号が選択されていることが確認された。

## 参考文献

- [1] S.Amari, T.Chen and A.Cichocki, "Stability analysis of learning algorithms for blind source separation", Neural Networks, vol.10, no.8, pp.1345-1351, 1997.
- [2] A. Horita, K. Nakayama and A. Hirano, "A distortion-free learning algorithm for feedforward multi-channel blind source separation", IEICE Trans. Fundamentals, Vol.E90-A, No.12, pp.2835-2845, Dec., 2007.
- [3] 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, Vol.11, No.3, pp.237-246, May., 2007.
- [4] A.Jung, F.J.Theis, C.G.Puntonet and E.W.Lang, "FastGeo - A histogram based approach to linear geometric ICA", Proc. of ICA2001, pp.418-423, Apr., 2001.
- [5] K.Waheed and F.M.Salem, "Algebraic overcomplete independent component analysis", ICA2003, Nara, Japan, Apr., 2003.
- [6] 加藤治雄, 中山謙二, 平野晃宏, "オーバーコンプリート形ブラインド信号源分離の構成法と学習法", 電子情報通信学会, 信学技報, Vol.105, No.293, pp.43-48, SIP2005-92, Sep., 2005.
- [7] 加藤治雄, 中山謙二, 平野晃宏, "オーバーコンプリート形 BBS のフィードバック構成における信号歪み低減法", 第 21 回信号処理シンポジウム (京都), C4-4, Nov., 2006.
- [8] 本田京実, "ダイナミック・オーバーコンプリート形ブラインド信号源分離の一検討", 金沢大学, 卒業論文, Feb., 2006.
- [9] 堀田明秀, "畳み込み混合過程に対するブラインド信号源分離の学習法に関する研究", 金沢大学大学院, 博士論文, Jan., 2007.
- [10] 李ソソミン, "フィードフォワード回路によるダイナミックオーバーコンプリート形 BSS", 金沢大学, 卒業論文, Feb., 2007.