

# 時間周波数解析によるブラインド信号源分離

芦野隆一（大阪教育大学） 萬代武史（大阪電気通信大学）  
守本晃（大阪教育大学） 佐々木文夫（東京理科大学）

第 68 回神楽坂解析セミナー 2008/01/26

概要— 人間がパーティ会場のように背景雑音の中で複数の人が話しているような状況でも自分が注意を注いでいる相手の声を聞き取ることができるのは、人間の聴覚や脳がどのような処理をしているからなのだろうか。この問題はカクテルパーティー問題と呼ばれている。カクテルパーティー問題に対する一つのアプローチが、ブラインド信号源分離と呼ばれる信号処理の問題であり、マイクやコンピュータなどの機器を使ってこの人間の処理をシミュレートすることを目的とする。本講演では、ブラインド信号源分離とはどんな問題かを説明し、最も簡単な時空間混合問題に対して、時間周波数解析を使った解法を概観する。

キーワード： ブラインド信号源分離，時間周波数解析，時空間混合問題

Key Words： blind source separation, time-frequency analysis, spatio-temporal mixture problem

## 1 導入

未知の  $n$  個の信号源を  $\{s_k(t)\}_{k=1}^n$  とし、既知の  $m$  個の観測信号を  $\{x_j(t)\}_{j=1}^m$  とする。これらの信号は実数値であるとし、 $m \geq n$  と仮定する。次のベクトル表記：

$$s(t) = [s_1(t), \dots, s_n(t)]^T, \\ x(t) = [x_1(t), \dots, x_m(t)]^T$$

を使う。ブラインド信号源分離 (blind source separation) における最も簡単な時空間混合問題 (spatio-temporal mixture problem) では、未知の定数行列  $A = (a_{j,k}) \in \mathbb{R}^{m \times n}$  と  $C = (c_{j,k})$ ,  $c_{j,k} > 0$  があって、観測信号  $\{x_j(t)\}_{j=1}^m$  が

$$x_j(t) = \sum_{k=1}^n a_{j,k} s_k(t - c_{j,k}) \quad (1)$$

と表されていると仮定する。行列  $A$  を混合行列 (mixture matrix) といい、行列  $C$  を遅延行列 (delay matrix) という。特に、遅延行列  $C$  が零行列のときは、空間混合問題 (spatial mixture problem) と呼ぶ。一般の時空間混合問題では、時間に依存する混合行列  $A(t)$  により、信号源との合成積： $x(t) = (A * s)(t)$  と表されていると仮定する。観測信号  $x_j$  の観測地点を  $P_{x_j}$  とし、信号源  $s_k$  の発生地点を  $P_{s_k}$  とおく。我々の問題は以下の問題 1 である。

問題 1. (1) において、 $m$ ,  $x_j$ ,  $P_{x_j}$  を既知として、未知の  $n$ ,  $s_k$ ,  $c_{j,k}$ ,  $a_{j,k}$ ,  $P_{s_k}$  を求めよ。

音速を  $v_0 := 330[\text{m/sec}]$  とおくと、 $l_{j,k} := v_0 c_{j,k}$  は 2 点  $P_{x_j}$ ,  $P_{s_k}$  の間の距離を表す。観測信号  $x_j$  の観測地点  $P_{x_j}$  は既知なので、 $l_{j,k}$  が求まれば GPS などにも使われている双曲線航法と同じ方法によって、信号源  $s_k$  の発生地点  $P_{s_k}$  を求めることができる。

まず、空間混合問題を考えてみよう。空間混合問題では、観測信号  $\{x_k(t)\}_{k=1}^m$  が  $x(t) = As(t)$  と表されていると仮定することになるが、難しいのは  $n$  と  $A$  が未知だということである。 $n$  がわかったとしても、与えられた  $x(t)$  に対して、 $x(t) = \tilde{A}\tilde{s}(t)$  となる  $\tilde{A}$  と  $\tilde{s}(t)$  の取り方は無数にある。このような分解の中でできるだけ信号源  $s(t)$  に近い  $\tilde{s}(t)$  を得るには、信号源  $s(t)$  が持つと思われる性質を予め知る必要がある。 $\tilde{s}(t)$  がその性質を満たすように  $x(t) = \tilde{A}\tilde{s}(t)$  と分解するのであ

る。この信号源  $s(t)$  が持つと思われる性質としてどのような条件を信号源  $s(t)$  に仮定するかによって、信号源分離の可否あるいは精度が決まってしまうので、信号源に対する仮定は非常に重要となる。

独立成分分析によるブラインド信号源分離

ブラインド信号源分離の最も強力な道具の一つは独立成分分析 (independent component analysis) である。(独立成分分析については A. Hyvärinen, J. Karhunen, and E. Oja<sup>9</sup>) を参照。独立成分分析を使ったブラインド信号源分離については A. Cichocki and S. Amari<sup>6</sup>) を参照。) 独立成分分析では次の仮定 1 を仮定する。

仮定 1. 信号源は統計的に独立である。

つまり、独立成分分析では分離された信号源  $\tilde{s}(t)$  の統計的独立性が最大になるように  $x(t) = \tilde{A}\tilde{s}(t)$  と分解する。数学的には統計的独立性は確率密度を用いて定義される。

時間周波数情報によるブラインド信号源分離

信号源  $s(t)$  が持つと思われる別の性質として、A. Jourjine, S. Rickard and O. Yilmaz<sup>10</sup>) によって提案された時間周波数領域における直交性 (W-disjoint orthogonality) がある。これは、定義 1 のように定義される。

定義 1 (時間周波数領域における直交性). 信号源の短時間フーリエ変換の時間周波数領域におけるサポートは本質的に重ならない。

信号とそのフーリエ変換のサポートが共に有限であるならば、信号は恒等的に 0 となることが数学によって証明されているので、この定義 1 のような状況が実際に起こるのは非常に稀であろう。彼らは 2 つの信号源に対して、2 つの観測信号の短時間フーリエ変換の比を調べることにより、ブラインド信号源分離を行なう方法を提案した。さらに R. Balan and J. Rosca<sup>3</sup>) は、時空間混合問題が扱える時間周波数領域における直交性を提案した。定義 1 をより現実的な形にした仮定の下で、未知の個数の信号源に対する空間混合問題を短時間フーリエ変換の比を使って解く方法が、D. Napolitano, C. A. Berenstein and P. S. Krishnaprasad<sup>11</sup>), D. Napolitano, C. A. Berenstein, P. S. Krishnaprasad and D. C. Struppa<sup>12</sup>) によって与えられた。K. Fujita, Y. Takei, A. Morimoto, and R. Ashino<sup>7</sup>), R. Ashino,

C. A. Berenstein, K. Fujita, A. Morimoto, M. Morimoto, D. Napoletani, and Y. Takei<sup>1)</sup> では、これらの論文を数学的に吟味している。

本講演では、空間混合問題の結果を (1) の形の時空間混合問題にうまく適用して解く方法を提案した R. Ashino, T. Mandai, A. Morimoto, and F. Sasaki<sup>2)</sup> のアイデアを概説する。

## 2 準備

関数  $f \in L^1(\mathbb{R})$  のフーリエ変換と関数  $g \in L^1(\mathbb{R})$  の逆フーリエ変換をそれぞれ

$$\mathcal{F}[f] = \hat{f}(\omega) = \int_{\mathbb{R}} e^{-it\omega} f(t) dt,$$

$$\mathcal{F}^{-1}[g] = \check{g}(t) = (2\pi)^{-1} \int_{\mathbb{R}} e^{it\omega} g(\omega) d\omega$$

で定義する。時間周波数解析で基本的な 3 つの  $L^2(\mathbb{R})$  上のユニタリ作用素は

$$\text{平行移動} \quad T_y f(t) = f(t - y),$$

$$\text{変調} \quad M_\xi f(t) = e^{it\xi} f(t),$$

$$\text{伸張} \quad D_\rho f(t) = |\rho|^{-1/2} f(\rho^{-1}t),$$

$y, \xi \in \mathbb{R}$ ,  $\rho \in \mathbb{R} \setminus \{0\}$  である。ウェーブレット  $\psi$  に関する  $f \in L^2(\mathbb{R})$  の連続ウェーブレット変換  $W_\psi f(b, a)$  を

$$W_\psi f(b, a) = \langle f, T_b D_a \psi \rangle$$

で定義する。以後、ウェーブレット  $\psi$  として Meyer のウェーブレット  $\psi_M$  に対応する解析信号をとる。このウェーブレット  $\psi$  は複素数値であり、

$$\Re\psi = \psi_M, \quad \Im\psi = \text{Hilbert\_Transform}[\psi_M]$$

を満たす。このウェーブレットの利点は、周波数の解像度が良いこと、伸張パラメータ  $a$  が  $a = \xi^*/2\pi\omega$  によって周波数  $\omega$  と対応することである。ここで、 $\xi^*$  は  $\psi$  の中心周波数である。信号源と観測信号の時間周波数情報として

$$S_k(t, \omega) := W_\psi s_k(t, \xi^*/2\pi\omega),$$

$$X_j(t, \omega) := W_\psi x_j(t, \xi^*/2\pi\omega)$$

を考える。式 (1) の連続ウェーブレット変換は

$$X_j(t, \omega) = \sum_{k=1}^n a_{j,k} S_k(t - c_{j,k}, \omega) \quad (2)$$

となる。式 (1) の短時間フーリエ変換の場合には、 $S_k(t - c_{j,k}, \omega)$  の係数に  $c_{j,k}$  が入った形になることを注意しておく。

集合  $\{(t, \omega) : S_k(t, \omega) \neq 0\}$  を信号源  $s_k(t)$  のアクティブ領域 (active region) と呼ぶ。信号源  $s_k(t)$  だけがアクティブな領域を

$$E_k^\circ := \{(t, \omega) : S_k(t, \omega) \neq 0, S_l(t, \omega) = 0 \ (l \neq k)\}$$

とおく。アクティブな領域が単純な 4 つの信号源の場合を Fig. 1 に示す。

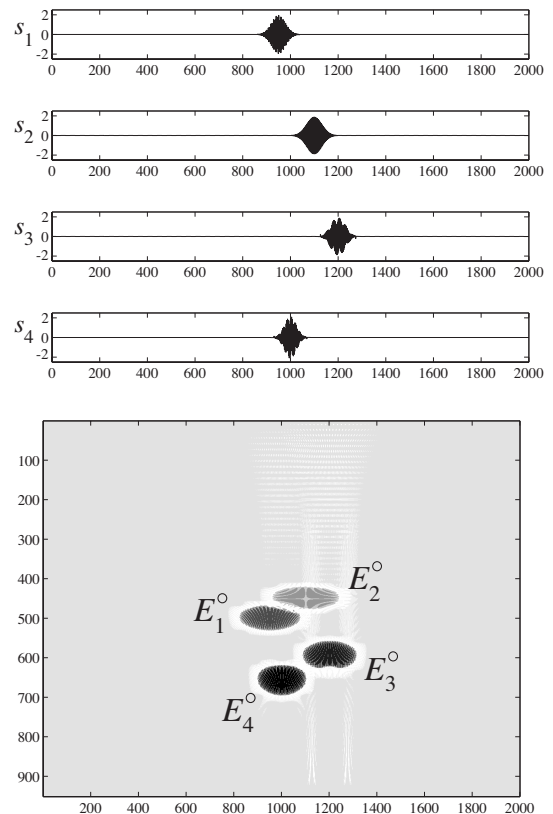


Fig. 1: Simple four sources (top) and their active regions in the time-frequency space (bottom).

## 3 空間混合問題

我々の空間混合問題の解法は次の 2 つのステップからなる。

1. 信号源の個数を評価：2 個の観測信号だけで可能。
2. 信号源の分離： $n$  個の観測信号が必要。  
(個数の評価でつかった 2 個の観測信号を含めて  $n$  個の観測信号を  $m$  個の観測信号から選ぶ。)

信号源の個数評価

2 個の観測信号の時間周波数情報  $X_1(t, \omega)$ ,  $X_2(t, \omega)$  を考えよう。

$$Q(t, \omega) := X_2(t, \omega)/X_1(t, \omega)$$

とおく。複素ウェーブレットを使っているので、一般に 2 個の時間周波数情報の比  $Q(t, \omega)$  は複素数である。しかし、集合  $E_k^\circ$  上では、

$$Q(t, \omega) = a_{2,k}/a_{1,k} \in \mathbb{R}$$

が成り立つことに注意する。もし、集合  $E_k^\circ$  に含まれる点  $(t_k, \omega_k)$  が 1 点でもわかれば、 $a_{2,k}/a_{1,k} = Q(t_k, \omega_k)$  により比  $a_{2,k}/a_{1,k}$  が求まる。我々はこれらの比の個数で信号源の個数を評価するので、次の条件を仮定する。

仮定 2.

$$a_{2,k}/a_{1,k} \neq a_{2,l}/a_{1,l}, \quad k \neq l.$$

問題は集合  $E_k^\circ$  がわからないことである。そこで、作業仮説として次の仮定 3 をおく。

仮定 3. 比  $Q(t, \omega)$  は一般に実数にはならず, 実数になったとしても一定ではない. 複素数値関数の値が特定の実数値が非常に多いというのは関数の特別な構造によるものだろう. 正しい推定値の場合はそのような構造を持っている.

十分小さい正数  $\epsilon > 0$  に対して,

$$\tilde{E}_\epsilon^\circ = \{(t, \omega) : |\Im Q(t, \omega)| < \epsilon\}$$

とおく. この集合  $\tilde{E}_\epsilon^\circ$  は計算で求めることができる. 明らかに,

$$\bigcup_{k=1}^n E_k^\circ \subset \tilde{E}_\epsilon^\circ$$

である. 時間周波数領域における直交性にあたる作業仮説として次の仮定 4 をおく.

仮定 4.  $\tilde{E}_\epsilon^\circ \setminus \bigcup_{k=1}^n E_k^\circ$  は無視できる.

$\Re Q(t, \omega)$ ,  $(t, \omega) \in \tilde{E}_\epsilon^\circ$  を求め, 分布図を描く. 分布図のピークは比  $Q(t, \omega)$  がある実定数を多くの点でとることを示すので, このピークの数で信号源の個数が評価できる. 分布図のピーク数は  $n$  であるとし, それぞれのピークを与える横軸の座標を  $b_{2,k}$ ,  $k = 1, \dots, n$  とおく. この横軸の座標  $b_{2,k}$ ,  $k = 1, \dots, n$  は比  $a_{2,k}/a_{1,k}$ ,  $k = 1, \dots, n$  のどれかに 1 対 1 に対応する.

Fig. 1 の信号源に対して, 混合行列を

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0.8 & 0.9 & 1 & 1.1 \\ 1.2 & 1 & 0.95 & 0.8 \\ 0.8 & 1.3 & 1.0 & 0.7 \end{pmatrix}.$$

として観測信号を構成したときの集合  $\tilde{E}_\epsilon^\circ$  と分布図を Fig. 2 に示す. 分布図のピークは 4 つあり, それらのピークを与える横軸の座標は  $a_{2,k}/a_{1,k}$ ,  $k = 1, 2, 3, 4$  の値,  $0.8/1, 0.9/1, 1/1, 1.1/1$  に対応している.

信号源の分離

上述のステップで信号源の個数  $n$  は既知となったので,  $m$  個の観測信号から  $n$  個の観測信号  $x_j(t)$ ,  $j = 1, \dots, n$  を選んだとする.  $\tilde{x}(t) = [x_1(t), \dots, x_n(t)]^T$  とおく. それぞれの  $j = 3, \dots, n$  に対して, 時間周波数情報の比  $X_j(t, \omega)/X_1(t, \omega)$  を考え, 上述のステップと同様の方法で,  $b_{j,k}$ ,  $k = 1, \dots, n$  を求める. このとき,  $b_{2,k}$  を与えるいくつかの点  $(t, \omega)$  での比  $X_j(t, \omega)/X_1(t, \omega)$  を調べることにより, ピークを与える横軸の座標が  $b_{2,k}$  と同じ  $E_k^\circ$  上でピークを与えるかがわかるので, ピークを与える横軸の座標  $b_{j,k}$ ,  $k = 1, \dots, n$  の並べ方は一意に定まる.  $b_{1,k} = 1$ ,  $k = 1, \dots, n$  とおき,  $B = (b_{j,k}) \in \mathbb{R}^{n \times n}$  と定義すれば,  $(a_{j,k}/a_{1,k}) = (b_{j,k})$  が成り立ち, 避けられない不定性を除き,  $A$  が推定できたことになる. したがって,  $\det B \neq 0$  ならば,  $\tilde{s}(t) = B^{-1}\tilde{x}(t)$  と分離できる.

#### 4 最も簡単な時空間混合問題

空間混合問題の解法を顧みると, 時間周波数情報の比の値が特定の実数値が非常に多いというのは, 構造的理由によるものであって, 正しい推定値の場合はそのような構造を持っているという仮定 3 に基づいていることがわかる. このことが最も簡単な時空間混合問題 (1) にもあてはまると仮定しよう. 適当なステップ

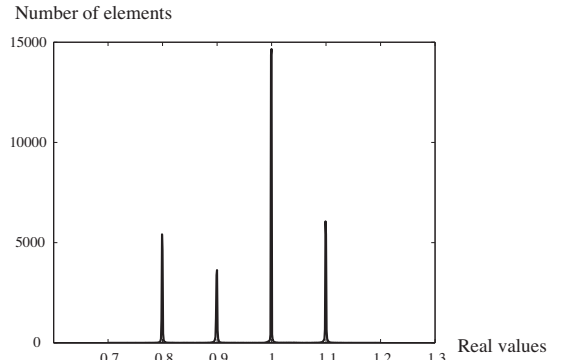
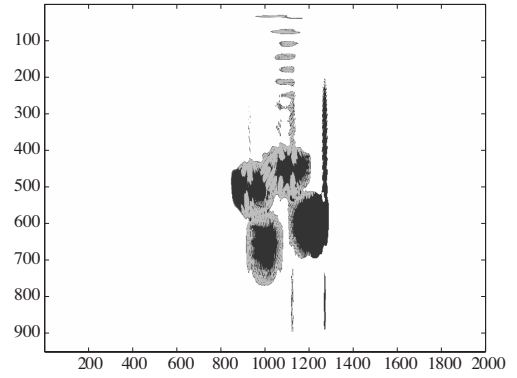


Fig. 2: The set  $\tilde{E}_\epsilon^\circ$  (top) and the distribution of values of  $\Re Q(t, \omega)$  for  $(t, \omega) \in \tilde{E}_\epsilon^\circ$  (bottom).

で予想される未知の遅延行列  $C = (c_{j,k})$  を順に与えて, 空間的混合問題の手法を適用する.

まず,  $X_1(t, \omega)$  を基準とし,  $X_2(t, \omega)$  の時間のずれを  $\tau$  とおき,

$$\tilde{c}_{j,k} := c_{j,k} - c_{1,k}, \quad j = 2, \dots, n, \quad k = 1, \dots, n$$

とおく. 時間のずれ  $\tau$  を持つ時間周波数情報の比:

$$\begin{aligned} & \frac{X_2(t + \tau, \omega)}{X_1(t, \omega)} \\ &= \frac{\sum_{k=1}^n a_{2,k} S_k(t - c_{1,k} + (\tau - \tilde{c}_{2,k}), \omega)}{\sum_{k=1}^n a_{1,k} S_k(t - c_{1,k}, \omega)} \end{aligned}$$

を考える. 空間混合問題の  $E_k^\circ$  にあたる集合  $E_k^\dagger$  を

$$E_k^\dagger := \{(t, \omega) : S_k(t, \omega) \neq 0, \\ S_l(t + c_{j,k} - c_{j,l}, \omega) = 0 \ (\forall j, l \neq k)\}$$

で定義する. このとき,

$$(t - c_{1,k}, \omega) \in E_k^\dagger, \quad \tau = \tilde{c}_{2,k}$$

であるならば,  $X_2(t + \tau, \omega)/X_1(t, \omega) = a_{2,k}/a_{1,k}$  となることに注意する.

適当なステップで  $\tau$  を動かす. それぞれの  $\tau$  ごとに,  $X_2(t + \tau, \omega)/X_1(t, \omega)$  が実数値となる場合の分布図を描く. このとき, 作業仮説として次の仮定 5 をおく.

仮定 5.  $\tau$  が  $\tilde{c}_{2,k}$ ,  $k = 1, \dots, n$  のどれかの正しい推定値であれば分布図にピークが出るが, そうでなければ分布図にピークが出ない.

この仮定の下で、比  $b_{2,k}$  と  $\tilde{c}_{2,k}$  が同時に求まる。同様に  $X_j(t + \tau, \omega)/X_1(t, \omega)$ ,  $j = 3, \dots, n$  を考えることにより、比  $b_{j,k}$  と  $\tilde{c}_{j,k}$  が求まる。  $\tilde{c}_{1,k} = 0$ ,  $k = 1, \dots, n$  とおき、  $\tilde{C} = (\tilde{c}_{j,k}) \in \mathbb{R}^{n \times n}$  と定義すれば、  $(c_{j,k} - c_{1,k}) = (\tilde{c}_{j,k})$  が成り立ち、避けられない不定性を除き、  $C$  が推定できたことになる。

$n = 4$  の場合の実数値となる場合の分布図の例を Fig. 3 に示す。

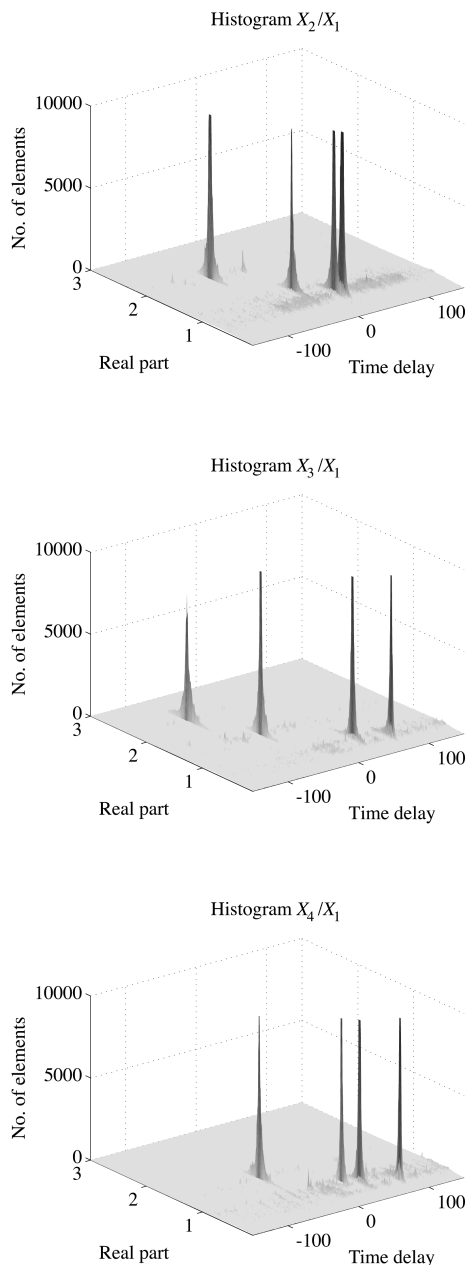


Fig. 3: The distribution of real values  $X_j(t + \tau, \omega)/X_1(t, \omega)$ ,  $j = 2, 3, 4$ .

### 双曲線航法

遅延行列  $C$  が推定できれば、観測地点  $P_{x_j}$  と信号源の位置  $P_{s_k}$  との間の距離  $l_{j,k}$  がわかったことになる。どうやって既知の  $P_{x_j}$  と  $l_{j,k}$  から未知の  $P_{s_k}$  を求めればよいのだろうか。これには、双曲線航法と呼ばれる方法と同じ方法を使えばよい。双曲線航法の例を Fig. 4 に示す。

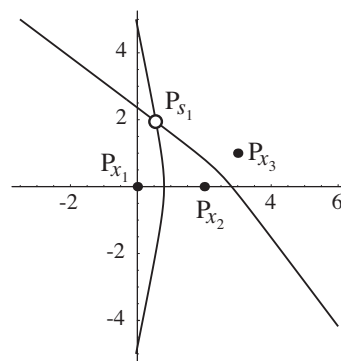


Fig. 4: Hyperbolic navigation system.

### 参考文献

- 1) R. ASHINO, C. A. BERENSTEIN, K. FUJITA, A. MORIMOTO, M. MORIMOTO, D. NAPOLETANI, AND Y. TAKEI, *Mathematical background for a method on quotient signal decomposition*, *Applicable Analysis*, **86** (5), 577–609, 2007.
- 2) R. ASHINO, T. MANDAI, A. MORIMOTO, AND F. SASAKI, *Blind source separation using time-frequency analysis*, *Fields Institute Communications*, **52**, 401–414, 2007.
- 3) R. BALAN AND J. ROSCA, *Statistical properties of STFT ratios for two channel systems and applications to blind source separation*, *Proceedings of the 2nd International Workshop on Independent Component Analysis and Blind Signal Separation*, 19–22, Helsinki, Finland, June 2000.
- 4) C. CHERRY, *On human communication*, the revised edition, Cambridge, MA, The MIT Press, 1980.
- 5) S. CHOI AND A. CICHOCKI, *Blind separation of non-stationary sources in noisy mixtures*, *Electron. Lett.*, **36** (9), 848–849, 2000.
- 6) A. CICHOCKI AND S. AMARI, *Adaptive blind signal and image processing: Learning algorithms and applications*, Chichester, West Sussex, England, John Wiley & Sons, 2002.
- 7) K. FUJITA, Y. TAKEI, A. MORIMOTO, AND R. ASHINO, *Mathematical view of a blind source separation on a time frequency space*, *Applied Mathematics and Computation*, **187**, 153–162, 2007.
- 8) S. HAYKIN AND Z. CHEN, *The cocktail party problem*, *Neural Computation*, **17** (9), 1875–1902, 2005.
- 9) A. HYVÄRINEN, J. KARHUNEN, AND E. OJA, *Independent component analysis*, New York, NY, John Wiley & Sons, 2001.
- 10) A. JOURJINE, S. RICKARD AND O. YILMAZ, *Blind separation of disjoint orthogonal signals: Demixing  $n$  sources from 2 mixtures*, *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE Press, 2000. June 5-9, 2000, Istanbul, Turkey.
- 11) D. NAPOLETANI, C. A. BERENSTEIN AND P. S. KRISHNAPRASAD, *Quotient Signal Decomposition and Order Estimation*, Technical Research Report of University of Maryland, TR 2002-47.
- 12) D. NAPOLETANI, C. A. BERENSTEIN, P. S. KRISHNAPRASAD AND D. C. STRUPPA, *Quotient Signal Estimation*, *Progress in Mathematics*, **238**, Boston, MA, Birkhäuser, 2005.
- 13) O. YILMAZ AND S. RICKARD *Blind source separation of speech mixtures via time-frequency masking*, *IEEE Transactions on Signal Processing*, **52** (7), 1830–1847, 2004.