

時間周波数解析によるブラインド信号源分離

Blind source separation using time-frequency analysis *

大阪教育大学・情報科学 守本 晃 (Akira Morimoto) [†]

Division of Information Science, Osaka Kyoiku University

概要

パーティ会場のような雑音や他の音声信号が入り交じった環境でも我々は人の話し声を聞き分けることができる。我々は背景雑音から特定の音声を識別できる能力を持っている。このような聴覚の特徴をカクテルパーティ効果とよぶ。カクテルパーティ効果を工学的に考察すると、複数のセンサーを使って得られた観測信号のみから、信号源の個数・位置および混合モデルのパラメータを推定し、最終的に元の信号を再構成しようという逆問題になる。この逆問題をブラインド信号源分離問題とよぶ。本講究録では、時間周波数解析の立場からブラインド信号源分離を考察する。

キーワード：ブラインド信号源分離，時間周波数情報，ウェーブレット変換

Keywords: blind source separation, time-frequency information, wavelet transform

2000 *Mathematics Subject Classification*. Primary: 42C40; Secondary: 94A12, 65T60

1 はじめに

パーティ会場では、いろいろな話し声や音楽や食器の音などの様々な音声信号が混在している。そのような周囲の喧噪にもかかわらず、我々は会話を楽しむことができる。我々は、複雑に重なり合った音の中から、特定の音声信号のみを選択的に聞き分けることができる。この聴覚系の能力は心理学者の Cherry [8] によって最初に問題視され、**カクテルパーティ効果**と名付けられた。カクテルパーティ効果は心理学から工学にわたる広い分野で研究されている。Cherry は、センサーである左右の耳の位置・指向性・感度の違いおよび文法・文脈などの特徴を利用すれば、特定の音声信号を分離する機械（フィルター）を作れる可能性がある

*この研究は部分的に科学研究費補助金 (C) 20540168 の補助を受けている。

[†]e-mail: morimoto@cc.osaka-kyoiku.ac.jp

いっている。工学的には、複数個の観測信号から、信号源と観測信号の間の数理モデルを利用して、信号源の情報（個数・方向・位置）とモデルパラメータの推定、それぞれの信号源の出力を決定する逆問題を**ブラインド信号源分離**とよぶ。

1990 年あたりから、**独立成分分析**という手法が開発されて、信号源分離が実現できるようになってきた。この手法は信号源が確率的に独立であるとの仮定を基に作られている。つまり、各信号源の確率的な独立性が増えるように数理モデルのパラメータを少しずつ変更して、元の信号を推定する方法である。この少しずつ変更するやり方は、1986 年に [26] で提案されたニューラルネットワークのバックプロパゲーションの学習アルゴリズムを用いて実現される。

本講究録では、**時間周波数解析**を用いたブラインド信号源分離問題を取り扱う。本講究録で提案する手法の特徴は、独立成分分析を用いないこと、および最初の解析で信号源の数が決定できることである。信号源分離は複数個の観測信号を同等な信号源（元の信号）に分離する問題である。したがって、情報処理の代表的な分野であるデータの圧縮やノイズ除去など、一つの信号を主要部分と重要度の低い部分に分ける問題とは本質的に異なる。

準備として、第 2 節で時間周波数解析について述べる。とくに短時間フーリエ変換と連続ウェーブレット変換について述べる。第 3 節でブラインド信号源分離問題を定式化する。第 4 節でカクテルパーティ効果の歴史についてふれて、独立成分分析を用いた標準的な解法と時間周波数解析を用いた従来の解法を振り返る。第 5 節では我々が研究に用いたウェーブレット関数について述べる。第 6 節で空間的混合問題に対して、時間周波数解析を用いた解法とそのアルゴリズムを提案する。第 7 節で、一番簡単な時空間的混合問題の解法とそのアルゴリズムを提案する。第 8 節では、第 7 節の解法の問題点と位相情報を利用して改善した解法とそのアルゴリズムについて述べる。

ホームページ [32] に数値計算で用いた元の信号や観測信号などを置いておく。

2 時間周波数解析

$f(t), g(t) \in L^2(\mathbb{R})$ の内積とノルムは

$$\langle f, g \rangle = \int_{\mathbb{R}} f(t) \overline{g(t)} dt, \quad \|f\| = \sqrt{\langle f, f \rangle}$$

である。フーリエ変換・逆フーリエ変換を

$$\begin{aligned} \mathcal{F}[f] &= \widehat{f}(\xi) = \int_{\mathbb{R}} f(t) e^{-it\xi} dt, \\ \mathcal{F}^{-1}[g] &= \check{g}(t) = \frac{1}{2\pi} \int_{\mathbb{R}} g(\xi) e^{it\xi} d\xi \end{aligned}$$

で定義する。このときパーセバルの等式は、

$$\langle f, g \rangle = \frac{1}{2\pi} \langle \widehat{f}, \widehat{g} \rangle$$

である。本講究録では、フーリエ変換を取ったときの角周波数を表す [rad/sec] 単位の変数として ξ を選ぶ。 ω は [Hz] 単位の周波数を表す。 $\xi = 2\pi\omega$ である。

2.1 窓関数

$g(t) \neq 0$ で $g(t), tg(t), \xi\hat{g}(\xi) \in L^2(\mathbb{R})$ を満たす関数を**窓関数**とよぶ。窓関数に対しては、以下の 5 種類の積分が計算できる。最初の積分は、エネルギーである。

$$E[g] := \|g\|^2 = \int_{\mathbb{R}} |g(t)|^2 dt = \frac{1}{2\pi} E[\hat{g}]. \quad (2.1)$$

時間窓の中心 $c[g]$ と幅 $\Delta[g]$ を

$$c[g] := \frac{1}{E[g]} \int_{\mathbb{R}} t |g(t)|^2 dt, \quad (2.2)$$

$$\Delta[g] := \sqrt{\frac{1}{E[g]} \int_{\mathbb{R}} (t - c[g])^2 |g(t)|^2 dt} \quad (2.3)$$

で定義し、周波数窓の中心 $c[\hat{g}]$ と幅 $\Delta[\hat{g}]$ を

$$c[\hat{g}] := \frac{1}{E[\hat{g}]} \int_{\mathbb{R}} \xi |\hat{g}(\xi)|^2 d\xi, \quad (2.4)$$

$$\Delta[\hat{g}] := \sqrt{\frac{1}{E[\hat{g}]} \int_{\mathbb{R}} (\xi - c[\hat{g}])^2 |\hat{g}(\xi)|^2 d\xi} \quad (2.5)$$

で定義する。式 (2.2), (2.3), (2.4), (2.5) は、次の確率密度関数

$$\mu_g(t) := \frac{|g(t)|^2}{E[g]}, \quad \mu_{\hat{g}}(\xi) := \frac{|\hat{g}(\xi)|^2}{E[\hat{g}]}$$

に対する平均と標準偏差である。

定義 2.1. 窓関数 $g(t)$ の**時間周波数窓**を

$$[c[g] - \Delta[g], c[g] + \Delta[g]] \times [c[\hat{g}] - \Delta[\hat{g}], c[\hat{g}] + \Delta[\hat{g}]] \quad (2.6)$$

で定義する。前半の $[c[g] - \Delta[g], c[g] + \Delta[g]]$ を時間窓、後半の $[c[\hat{g}] - \Delta[\hat{g}], c[\hat{g}] + \Delta[\hat{g}]]$ を周波数窓とよぶ。

信号 $f(t) \in L^2(\mathbb{R})$ と窓関数 $g(t)$ の内積 $\langle f, g \rangle$ は、信号 $f(t)$ の時間周波数窓 (2.6) の情報にアクセスしている。窓関数 $g(t)$ を固定して、時間窓と周波数窓を適当に動かす 2 パラメータ α, β を導入した関数系 $g^{(\alpha, \beta)}(t)$ を作成し、信号 $f(t)$ と $g^{(\alpha, \beta)}(t)$ の内積 $\langle f, g^{(\alpha, \beta)} \rangle$ を用いて、信号の情報を探る手法が**時間周波数解析**である。

注意 2.2. 任意の窓関数 $g(t)$ に対して、不確定性原理より、

$$\Delta[g] \times \Delta[\hat{g}] \geq \frac{1}{2}$$

である。よって、時間窓の幅 $\Delta[g]$ と周波数窓の幅 $\Delta[\hat{g}]$ を同時に小さくすることはできない。

2.2 短時間フーリエ変換

窓関数 $g(t)$ を固定し、時間を表す**平行移動パラメータ** $x \in \mathbb{R}$ と角周波数を表す**モジュレーションパラメータ** $\xi \in \mathbb{R}$ (ω [Hz] なら $\xi = 2\pi\omega$) を導入した関数系

$$g^{(x,\xi)}(t) := g(t-x) e^{i\xi t}$$

を考える。信号 $f(t)$ とこの関数系との内積を用いて、信号の情報を引き出す方法が短時間フーリエ変換である。

定義 2.3. 信号 $f(t)$ の窓関数 g を用いた**短時間フーリエ変換**を

$$\begin{aligned} V_g f(x, \xi) &:= \langle f, g^{(x,\xi)} \rangle = \int_{\mathbb{R}} f(t) \overline{g(t-x)} e^{it\xi} dt \\ &= \int_{\mathbb{R}} f(t) \overline{g(t-x)} e^{-it\xi} dt \end{aligned} \quad (2.7)$$

で定義する。

式 (2.7) は信号 $f(t)$ に時刻 x を中心とした窓 $\overline{g(t-x)}$ をかけて切り出した関数 $f(t) \overline{g(t-x)}$ をフーリエ変換して角周波数 $\xi = 2\pi\omega$ の値を取ったと考えることもできる。短時間フーリエ変換 $V_g f(x, \xi)$ は、信号 $f(t)$ の時間周波数窓

$$\left[x + c[g] - \Delta[g], x + c[g] + \Delta[g] \right] \times \left[\xi + c[\widehat{g}] - \Delta[\widehat{g}], \xi + c[\widehat{g}] + \Delta[\widehat{g}] \right]$$

の情報にアクセスしている。短時間フーリエ変換の時間周波数窓の大きさは、パラメータ x, ξ の値によらず時間幅 $\Delta[g]$, 周波数幅 $\Delta[\widehat{g}]$ で一定である。また、時間窓の中心は $x + c[g]$ [sec] で周波数窓の中心は $\xi + c[\widehat{g}]$ [rad/sec] である。

定義 2.4. 信号 $f(t)$ の時刻 t [sec], 周波数 ω [Hz] の**時間周波数情報** $F^g(t, \omega)$ を

$$F^g(t, \omega) := V_g f(t - c[g], 2\pi\omega - c[\widehat{g}]) \quad (2.8)$$

で定義する。

命題 2.5 (逆短時間フーリエ変換).

$$f(s) = \frac{1}{2\pi E[g]} \int_{\mathbb{R}} \int_{\mathbb{R}} V_g f(x, \xi) g^{(x,\xi)}(s) dx d\xi. \quad (2.9)$$

証明 右辺の積分を計算すると、

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}} V_g f(x, \xi) g^{(x,\xi)}(s) dx d\xi &= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(t) \overline{g(t-x)} e^{-it\xi} g(s-x) e^{i\xi s} dt dx d\xi \\ &= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t) \overline{g(t-x)} e^{-it\xi} dt \right) e^{i\xi s} d\xi \right] g(s-x) dx \\ &= 2\pi \int_{\mathbb{R}} f(s) \overline{g(s-x)} g(s-x) dx = 2\pi E[g] f(s) \end{aligned}$$

である。ただし、2 行目の $[\dots]$ 内は関数 $f(t) \overline{g(t-x)}$ のフーリエ変換・逆変換なので、 $2\pi f(s) \overline{g(s-x)}$ になる。□

この証明では、積分順序の交換ができることの部分が一番難しいが省略する。詳しくは、Cohen [11], Gröchenig [15]などを参照。

2.3 連続ウェーブレット変換

定義 2.6. 窓関数 $\psi(t)$ が次の **admissible 条件**

$$C_\psi := \int_{\mathbb{R}} \frac{|\widehat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty \quad (2.10)$$

を満たすとき、 $\psi(t)$ を**ウェーブレット関数**とよぶ。

ウェーブレット関数 $\psi(t)$ を固定して、**平行移動パラメータ** $b \in \mathbb{R}$ と**ダイレーションパラメータ** $a \in \mathbb{R}, a \neq 0$ を導入した関数系

$$\psi^{(b,a)}(t) := \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right)$$

を考える。信号 $f(t)$ とこの関数系との内積を用いて、信号の情報を引き出す方法が連続ウェーブレット変換である。

定義 2.7. 信号 $f(t)$ の**連続ウェーブレット変換**を

$$W_\psi f(b, a) := \langle f, \psi^{(b,a)} \rangle = \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt, \quad b, a \in \mathbb{R}, a \neq 0 \quad (2.11)$$

で定義する。

$c[\widehat{\psi}] > \Delta[\widehat{\psi}] > 0$ かつダイレーションパラメータ $a > 0$ の場合を考えよう。連続ウェーブレット変換 $W_\psi f(b, a)$ は、信号 $f(t)$ の時間周波数窓

$$\left[b + a(c[\psi] - \Delta[\psi]), b + a(c[\psi] + \Delta[\psi]) \right] \times \left[\frac{c[\widehat{\psi}] - \Delta[\widehat{\psi}]}{a}, \frac{c[\widehat{\psi}] + \Delta[\widehat{\psi}]}{a} \right]$$

の情報にアクセスしている。連続ウェーブレット変換の時間周波数窓の大きさは、ダイレーションパラメータ a の値に依存して変化する。時間窓の幅が $a\Delta[\psi]$ で周波数窓の幅が $\Delta[\widehat{\psi}]/a$ なので、 $a > 0$ が小さい（周波数窓の中心 $c[\widehat{\psi}]/a$ が高くなる）と、時間窓の幅が小さくなって周波数窓の幅が増える。逆に、 a が大きい（周波数窓の中心 $c[\widehat{\psi}]/a$ が低周波数になる）と、時間窓の幅が大きくなって周波数窓の幅が短くなる。時間窓の中心は $b + a c[\psi]$ [sec] で周波数窓の中心は $c[\widehat{\psi}]/a$ [rad/sec] である。したがって、 ω Hz を周波数窓の中心に合わせるためには、 $2\pi\omega = c[\widehat{\psi}]/a$ なので、 $a = c[\widehat{\psi}]/(2\pi\omega)$ と取ればよい。

定義 2.8. 信号 $f(t)$ の時刻 t [sec], 周波数 ω [Hz] の**時間周波数情報** $F^\psi(t, \omega)$ を

$$F^\psi(t, \omega) := W_\psi f \left(t - \frac{c[\psi] c[\hat{\psi}]}{2\pi\omega}, \frac{c[\hat{\psi}]}{2\pi\omega} \right) \quad (2.12)$$

で定義する.

注意 2.9. 上の時間周波数情報 $F^\psi(t, \omega)$ の定義では, ダイレーションパラメータ $a > 0$ に加えて $c[\hat{\psi}] > \Delta[\hat{\psi}] > 0$ を仮定しているので $\omega = c[\hat{\psi}]/(2\pi a) > 0$ である. 本講究録の数値実験の時に使うウェーブレット関数 (第 5 節の Ausher のウェーブレットの解析信号) ではこれらの条件は満たされる.

命題 2.10 (逆連続ウェーブレット変換).

$$f(s) = \frac{1}{C_\psi} \int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi f(b, a) \psi^{(b,a)}(s) \frac{dad b}{a^2}. \quad (2.13)$$

証明 右辺の $W_\psi f(b, a)$ に式 (2.11) を代入して計算すると,

$$\begin{aligned} \int_{\mathbb{R}} \int_{\mathbb{R}} W_\psi f(b, a) \psi^{(b,a)}(s) \frac{dad b}{a^2} &= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \psi\left(\frac{s-b}{a}\right) \frac{dad b}{|a|^3} \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(t) \overline{\psi\left(\frac{t+b}{a}\right)} \psi\left(\frac{s+b}{a}\right) \frac{dbdad t}{|a|^3} \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(t) \overline{|a| \hat{\psi}(a\xi) e^{it\xi}} |a| \hat{\psi}(a\xi) e^{is\xi} \frac{d\xi dad t}{|a|^3} \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \frac{|\hat{\psi}(a\xi)|^2}{|a|} da \right) f(t) e^{-it\xi} e^{is\xi} dt d\xi = C_\psi f(s) \end{aligned}$$

である. 最初の変形は $-b$ を b と置き直した. 次に b 積分をパーセバルの等式で周波数 ξ の積分に変更した. このとき, 次のフーリエ変換を用いた.

$$\mathcal{F}_b \left[\psi\left(\frac{t+b}{a}\right) \right] (\xi) = \int_{\mathbb{R}} \psi\left(\frac{t+b}{a}\right) e^{-ib\xi} db = |a| \hat{\psi}(a\xi) e^{it\xi}.$$

最後の式で, admissible 条件から,

$$\int_{\mathbb{R}} \frac{|\hat{\psi}(a\xi)|^2}{|a|} da = C_\psi$$

である. 残りの部分は, 信号 $f(t)$ のフーリエ変換・逆変換である. $\square$

この証明では, 積分順序の交換ができることが一番難しいが省略する. 詳しくは, Daubechies [13], Mallat [22] など参照.

注意 2.11. ウェーブレット関数が次の admissible 条件

$$C'_\psi = \int_0^\infty \frac{|\widehat{\psi}(\xi)|^2}{|\xi|} d\xi = \int_{-\infty}^0 \frac{|\widehat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty$$

を満たす場合には、ダイレーションパラメータ $a \in \mathbb{R}_+ := (0, \infty)$ の範囲を考えれば十分である。このとき、逆ウェーブレット変換は、

$$f(s) = \frac{1}{C'_\psi} \int_{b \in \mathbb{R}} \int_{a \in \mathbb{R}_+} W_\psi f(b, a) \psi^{(b,a)}(s) \frac{da db}{a^2}$$

である。たとえば、 $\psi(t)$ が実数値関数の時は、 $\widehat{\psi}(-\xi) = \overline{\widehat{\psi}(\xi)}$ なので、この admissible 条件を満たす。

2.4 解析ウェーブレット変換

第 5 節で述べるように、我々の最近の研究では解析ウェーブレット変換を用いている。ここで、解析ウェーブレット変換の定義と性質を述べる。

定義 2.12. 信号 $f(t) \in L^2(\mathbb{R})$ に対して、**解析信号**を対応させる作用素 $\mathcal{A}f$ を

$$\mathcal{A}f(t) := \frac{1}{\pi} \int_{\mathbb{R}_+} \widehat{f}(\xi) e^{it\xi} d\xi \quad (2.14)$$

で定義する。ただし、 $\mathbb{R}_+ := (0, \infty)$ である。

信号 $f(t)$ のフーリエ変換 $\widehat{f}(\xi)$ の周波数が正の部分だけを取り出して、逆フーリエ変換して 2 倍すると、解析信号 $\mathcal{A}f$ が得られる。実数値関数 $f(t)$ では、 $f(t) = \Re \mathcal{A}f(t)$ である。ただし、複素数 z の実部を $\Re z$ と書く。

補題 2.13. 窓関数 $g(t)$ に対して、 $\mathcal{A}g \neq 0$ が窓関数になるための必要十分条件は $\widehat{g}(0) = 0$ になることである。

系 2.14. ウェーブレット関数 $\psi(t)$ の解析信号 $\mathcal{A}\psi$ が $\mathcal{A}\psi \neq 0$ を満たすならば、 $\mathcal{A}\psi$ はウェーブレット関数になる。

なぜなら、 $\widehat{\psi}(0) = 0$ なので、 $\mathcal{A}\psi$ は窓関数になる。さらに、admissible 条件は自動的に満たすから、 $\mathcal{A}\psi$ はウェーブレット関数である。

したがって、ウェーブレット関数 $\mathcal{A}\psi$ に関する連続ウェーブレット変換 $W_{\mathcal{A}\psi} f$ を考えることができる。この変換を**解析ウェーブレット変換**とよぶ。このとき次の定理が成立する。

定理 2.15. ダイレーションパラメータ $a \in \mathbb{R}_+$ の場合には、信号 $f(t)$ の連続ウェーブレット変換に対して、

$$\mathcal{A}(W_\psi f(\cdot, a)) = W_\psi \mathcal{A}f = W_{\mathcal{A}\psi} f = \frac{1}{2} W_{\mathcal{A}\psi} \mathcal{A}f \quad (2.15)$$

が成立する。

注意 2.16. ダイレクションパラメータが $a < 0$ のときは, $(W_{\mathcal{A}\psi}\mathcal{A}f)(b, a) = 0$ である. したがって, 逆解析ウェーブレット変換で信号 f の解析信号 $\mathcal{A}f$ を再構成する場合には, $a \in \mathbb{R}_+$ の範囲の情報だけで十分である. つまり,

$$\begin{aligned}\mathcal{A}f(s) &= \frac{1}{C_{\mathcal{A}\psi}} \int_{b \in \mathbb{R}} \int_{a \in \mathbb{R}_+} (W_{\mathcal{A}\psi}\mathcal{A}f)(b, a) (\mathcal{A}\psi)^{(b,a)}(s) \frac{dad b}{a^2} \\ &= \frac{2}{C_{\mathcal{A}\psi}} \int_{b \in \mathbb{R}} \int_{a \in \mathbb{R}_+} (W_{\mathcal{A}\psi}f)(b, a) (\mathcal{A}\psi)^{(b,a)}(s) \frac{dad b}{a^2}.\end{aligned}$$

実数値信号 $f(t)$ を解析ウェーブレット変換で扱う場合には, $f = \Re \mathcal{A}f$ なので, ダイレクションパラメータ $a \in \mathbb{R}_+$ の範囲のみ考えればよい.

詳しい証明については, 文献 [3, 4]などを参照のこと.

3 ブラインド信号源分離問題

N 個の信号源から出力した信号 (元の信号とよぶ) を, M 個のセンサーで捉えて観測信号が得られたとしよう. ここで, $M \geq N$ を仮定する. 元の信号から観測信号を作る数理モデルが線形な場合に, 観測信号から信号源の個数 N や数理モデルのパラメータを決定して, 最終的に元の信号を推定する逆問題を**ブラインド信号源分離問題**とよぶ. ブラインド信号源分離問題は数理モデルの違いにより, **時間的混合問題**, **空間的混合問題**, **時空間的混合問題**の3種類に分類できる. 元の信号と観測信号を縦ベクトルで

$$s(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T, \quad (3.1)$$

$$x(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T \quad (3.2)$$

と記述しよう. どちらも実数値関数であると仮定する.

3.1 時間的混合問題

$M = N = 1$ の場合で, 観測信号 $x(t)$ が元の信号 $s(t)$ とインパルス・レスポンス $h(t)$ の畳み込み

$$x(t) = h * s(t) = \int_{-\infty}^{\infty} h(u) s(t-u) du$$

で記述できる場合に, 観測信号から数理モデルのパラメータであるインパルス・レスポンス $h(t)$ と元の信号 $s(t)$ を求める逆問題を**時間的混合問題**とよぶ. 時間的混合問題は本講究録では取り扱わない.

3.2 空間的混合問題

時刻 t によらない $M \times N$ の定数行列 A が取れて、観測信号 $x(t)$ と元の信号 $s(t)$ の間の関係が

$$x(t) = As(t) \quad (3.3)$$

で記述できる数理モデルを仮定したブラインド信号源分離問題を**空間的混合問題**または**瞬時混合問題**とよぶ。また行列 A を**混合行列**とよぶ。この場合、観測信号から信号源の個数と混合行列 A を求めて、逆行列（一般化逆行列）をかけることにより元の信号を推定する。第 6 節で空間的混合問題を取り扱う。

3.3 時空間的混合問題

時刻 t に依存した $M \times N$ の行列 $A(t)$ が取れて、観測信号 $x(t)$ と元の信号 $s(t)$ の間の関係が

$$x(t) = A * s(t) = \int_{-\infty}^{\infty} A(u) s(t-u) du$$

で記述できる数理モデルを仮定したブラインド信号源分離問題を、**時空間的混合問題**とよぶ。本講究録では、行列 $A(t)$ の (j, k) 成分が

$$A(t)_{j,k} = a_{j,k} \delta(t - c_{j,k})$$

と δ 関数の定数倍で表される**一番簡単な時空間的混合問題**を考える。このとき、観測信号 $x(t)$ と元の信号 $s(t)$ の間の数理モデルは、

$$x_j(t) = \sum_{k=1}^N a_{j,k} s_k(t - c_{j,k}), \quad j = 1, 2, \dots, M \quad (3.4)$$

になる。 $A = (a_{j,k})$ を**混合行列**、 $C = (c_{j,k})$ を**遅延行列**または**時間遅れ行列**とよぶ。第 7, 8 節でこの数理モデルを取り扱う。一番簡単な時空間的混合問題で、信号の伝達速度が無限大の場合、あるいは全てのセンサーが同じ場所に設置されている場合には、空間的混合問題に帰着できる。

4 カクテルパーティ効果の歴史

パーティ会場やうるさい講義中のような、いろいろな音声信号が飛び交ったり、さまざまな雑音が入っているような環境下でも、我々は会話ができる。人間は、そのような背景雑音の中にあっても少なくとも一つの音声信号は識別できる能力を持っている。この人間の聴覚に関する能力を**カクテルパーティ効果**とよぶ。カクテルパーティ効果に最初に気づいて、問題にしたのが 1953 年の Cherry [8] である。1978 年の Cherry [9] の本にも記述がある。カクテルパーティ効果を、2005 年にレビューした Haykin-Chen [16] も参照のこと。

4.1 Cherry の考え方と実験

Cherry [8] は, 1953 年の論文で, 次の 5 つの条件を考慮に入れると, 混じり合った音声信号 (の内少なくとも一つ) を分離できる機械や “フィルター” を作れるのではないかと指摘している.

- (a) The voices come from different directions.
- (b) Lip-reading, gestures, and the like.
- (c) Different speaking voices, mean pitches, mean speeds, male and female, and so forth.
- (d) Accents differing.
- (e) Transition-probabilities (subject matter, voice dynamics, syntax...).

時間周波数解析を用いた本講究録の方法では, 条件 (c) と条件 (a) や信号がセンサーに届く時間差などに着目している. 一方, 独立成分分析を用いた方法では, 分離した信号源が確率的に独立であるという基準を用いるので, 条件 (e) に注目しているのだろうか?

Cherry は, 被験者に音声を分離させる実験を 2 パターン行った. どちらの実験も, 同一人物の発声した Message 1 と Message 2 を録音して用いる. これは, 条件 (b), (c), (d) の影響をなくすためである.

最初の実験では, Message 1 と Message 2 を混合したモノラル音声を作成し, それを被験者に聞かせて分離させた. これは, 人間は条件 (e) のみでどの程度分離できるかを調べるための実験である. 結論は, 話の内容や文法・慣用句などと片側のメッセージの無音部分を利用してある程度の精度で分離可能であるが, かなりの困難を伴うというものである.

2 番目の実験では, 左の耳で Message 1 を右の耳で Message 2 を同時に聞かせて, 被験者に分離させる実験である. これは, 左右の耳に異なった音声が入るという条件 (a) の極端な場合を再現した実験である. こちらの結論は, 片方の耳に意識を集中すれば, 簡単に分離できるというものである. 人間は簡単に片方の耳の音声入力をマスクできるのである.

Cherry の実験は, 複数の観測信号を用いて分離した信号を推定することは比較的簡単にできる可能性があることを示唆している. Cherry の実験を試してみたい人は, ホームページ [32] を参照のこと.

4.2 独立成分分析

独立成分分析を用いたブラインド信号源分離は, 次のアルゴリズム 4.1 にしたがって行われる.

アルゴリズム 4.1 (独立成分分析). 以下のアルゴリズムで目的関数を極小にする解を探す.

1. 目的関数とよばれる信号源（元の信号）の確率的な独立性を評価する関数（最小値を取れば独立性が高いと仮定する）を作成する.
2. 信号源の数 N を適当に定める, 最終結果が悪ければ N を取り直す.
3. 数理モデルのパラメータを適当に初期化する.
4. 数理モデルにしたがって, 観測信号から元の信号を推定する.
5. 目的関数を評価して, 推定した信号源間の独立性が満足できれば終了する. あるいは, 目的関数の値の減少幅が基準値以下になれば, 手順 2. へ
6. 独立性が上がるようにモデルパラメータを調整して（パラメータで目的関数を微分して降下方向にパラメータを動かして）, 手順 5. へ

目的関数をパラメータで微分して降下方向に少しずつ動かしていく方法は, ニューラルネットワークのバックプロパゲーションの学習アルゴリズムと同じ種類の考え方である. バックプロパゲーションのアルゴリズムは Rumelhart-Hinton-Williams [26] らによって 1986 年に発表されたので, 独立成分分析を用いたブラインド信号源分離問題の研究はそれ以降である.

1991 年の Signal Processing の 3 部作, Jutten-Herault [21], Comon-Jutten-Herault [12], Sorouchyari [27] あたりが英語で書かれた最初の方の論文である. フランス語のプローシーディングスでは, 1988 年まで溯ることができる. 空間的混合問題で音声信号を分離するには, 信号源の独立性を評価する目的関数として 1 次と 3 次のクロスモーメントを用いて作成すると, うまくできるらしい.

独立成分分析を用いた信号源分離の文献は, Cichocki-Amari [10] に詳しい. また日本語では村田 [29], Hyvärinen-Karhunen-Oja（根本幾・川勝真喜訳）[19] などがある.

4.3 時間周波数解析からのアプローチ

独立成分分析は, 信号源が確率的に独立であるという仮定の下に, 元の信号への分離を行う手法である. 時間周波数解析で, 信号源の独立性に代わる仮定として, 2000 年に Jourjine-Rickard-Yilmaz [20] らが, 窓関数 g を用いた短時間フーリエ変換を使った g -disjoint orthogonality という条件を提案した.

ℓ 番目の元の信号 $s_\ell(t)$ の時間周波数情報を $S_\ell^g(t, \omega)$ と書く. 元の信号 $s_\ell(t)$ が活動している時間周波数位置 (t, ω) の集合を

$$E_\ell = \{(t, \omega); S_\ell^g(t, \omega) \neq 0\}$$

とすると、 g -disjoint orthogonality は、 $\ell \neq k$ の時、 $E_\ell \cap E_k = \phi$ という条件である。つまり、同時に 2 個以上の信号源は活動していないという仮定である。これは、厳密な意味では数学的に成り立たない仮定であるが、 E_ℓ の定義を $|S_\ell^g(t, \omega)|$ がある程度大きいに变えておけば問題ない。

空間的混合問題を考える。2 個の観測信号 $x_1(t) = \sum a_{1,k} s_k(t)$, $x_2(t) = \sum a_{2,k} s_k(t)$ の時間周波数情報の商を考える。 $(t, \omega) \in E_\ell$ に対して、

$$Q_2^g(t, \omega) = \frac{X_2^g(t, \omega)}{X_1^g(t, \omega)} = \frac{\sum a_{2,k} S_k^g(t, \omega)}{\sum a_{1,k} S_k^g(t, \omega)} = \frac{a_{2,\ell} S_\ell^g(t, \omega)}{a_{1,\ell} S_\ell^g(t, \omega)} = \frac{a_{2,\ell}}{a_{1,\ell}}$$

と ℓ に応じた一定の実数値 $a_{2,\ell}/a_{1,\ell}$ になることが分かる。 $a_{2,\ell}/a_{1,\ell}$ は全て異なると仮定する。

アルゴリズム 4.2. 信号源は g -disjoint orthogonality の仮定を満たすとする。このとき、信号源分離は以下のように行われる。

1. 観測信号の時間周波数情報が十分小さい (t, ω) を無視すると、商 $Q_2^g(t, \omega)$ が N 種類の異なる実数値 (q_ℓ , $\ell = 1, 2, \dots, N$) を取れば、この N が信号源の個数の推定値である。
2. ℓ 番目の信号源が活動している領域を求める。 $\tilde{E}_\ell = \{(t, \omega); Q_2^g(t, \omega) = q_\ell\}$ 。
3. ℓ 番目の元の信号の時間周波数情報 $\tilde{S}_\ell^g(t, \omega)$ (順序と定数倍の自由度がある) を次で定める。

$$\tilde{S}_\ell^g(t, \omega) = \begin{cases} X_1^g(t, \omega), & (t, \omega) \in \tilde{E}_\ell, \\ 0, & \text{otherwise.} \end{cases}$$

4. $\tilde{S}_\ell^g(t, \omega)$ を逆短時間フーリエ変換して、元の信号 $\tilde{s}_\ell$ を推定する。

g -disjoint orthogonality の仮定は非常に強いので、観測信号が 2 個あれば信号源分離ができる。時間周波数平面で余り混ざっていない元の信号を少ない観測信号から分離できる有用な方法として、この考え方は**時間周波数マスキング**とよばれて、盛んに研究されている (Yilmaz-Rickard [28] など)。

2000 年の同時期に、Balan-Rosca [7] が g -disjoint orthogonality とよく似た仮定を使って、2 個の信号源の時空間的混合問題を時間周波数情報の商を用いた方法で扱った。

さて、図 1 は、第 6 節の空間的混合問題の数値実験で用いる 4 個の音声信号を同じ強さで混ぜたときの信号源の活動状態を表した図である。黒色を塗ったところ (全体の約 4 %) が、1 個の信号源しか活動していない時間周波数領域で、灰色 (全体の約 80 %) が 2 個以上の信号源が活動している時間周波数領域で、白色 (全体の約 16 %) が信号源が活動していない時間周波数領域である。 g -disjoint

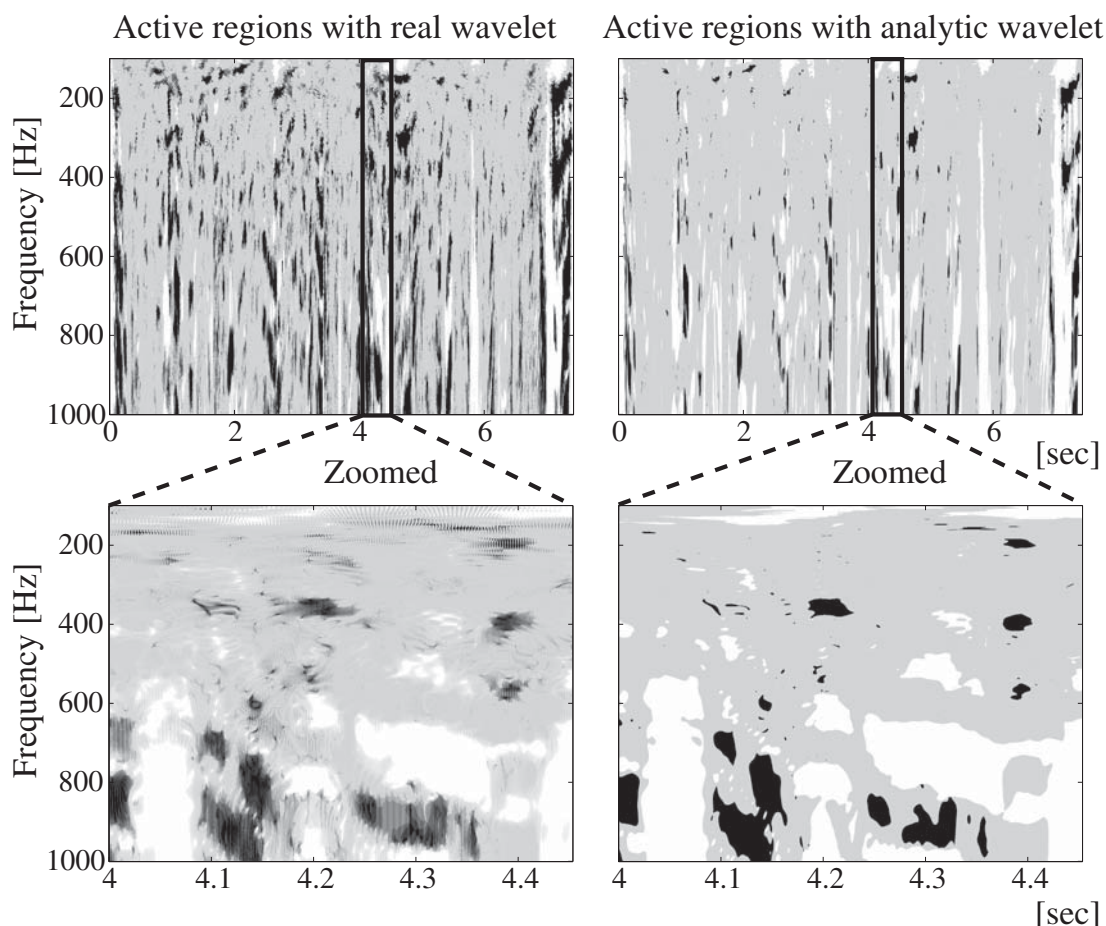


図 1: 信号源の活動領域, 左: 実数値ウェーブレット, 右: 解析ウェーブレット. 第 6 節の 4 音源, 全体 (7.4 [sec]) とズーム (4 [sec] から 4.5 [sec]). 黒: 一つの信号源だけ活動 (4 %), 灰色: 2 つ以上活動 (80 %), 白: 活動無し (16 %).

orthogonality の仮定は, 図 1 の灰色部分の面積が 0 という仮定で, 音声信号の分離に使うには強すぎる.

2002 年, Napoletani-Berenstein-Krishnaprasad [24] は, 図 1 のような信号源に対しても時間周波数情報の商を用いて, 空間的混合問題のモデルパラメータの推定が可能であることを示した. ただし, 元の信号を推定するには, 信号源以上の観測信号が必要であった. 2005 年に, Napoletani-Berenstein-Krishnaprasad-Struppa [25] では, 信号源の個数より少ない観測信号からの信号源分離を考えている.

我々は, Napoletani と Berenstein との共同研究を経て, 空間的混合問題で, 時間周波数情報の商を用いた方法の数学的なバックグラウンドを示した [2, 14]. その後, 一番簡単な時空間的混合問題を [3, 4] で扱った. ただし, 図 1 のように時間周波数平面でよく混ざった信号源を扱い, 信号源の数より観測信号の数が多いという仮定を付ける. 我々の信号源に対する仮定は, 図 1 の黒色部分の面積が十

分大きい（全体の数 %）というものである。

5 我々の研究で使っているウェーブレット関数

第 6 節以降で詳しく述べるが，時間周波数解析を用いた信号源分離の研究では，観測信号の時間周波数情報の商を用いて，一つの信号源のみが活動している時間周波数領域（図 1 の黒色の部分）を探ることが重要である．本講義録で行う数値実験では，この領域が探査する時間周波数領域全体の数 % しかないので，観測信号 $f(t)$ を複数種類の実数値ウェーブレット関数 $\psi_\ell(t)$ で連続ウェーブレット変換した時間周波数情報 $F^{\psi_\ell}(t, \omega)$ を利用すると効率がよい．たとえば，図 1 の左側が実数値の Ausher のウェーブレット関数を用いた場合で，右側が Ausher のウェーブレットの解析信号である複素数値ウェーブレット関数（実部と虚部のどちらもウェーブレット関数なので，実数値 2 個分である）を用いた場合である．下図のズームエリアでみると，左図の黒色の部分が霜降り状態になっている実数値ウェーブレットの場合に比べ，右図の複素数値ウェーブレットの場合は単純な領域になっていて解析しやすいことがわかる．

先行研究での短時間フーリエ変換の利用は，実部と虚部が使えるので，都合がよかったのである．最初，我々 [30, 31, 2, 14] はガボールウェーブレット $\psi(t) = e^{-t^2/2} e^{i2\pi\omega_0 t}$ を使った．その後，論文 [5] では，3 種類の実数値ウェーブレット関数

$$\begin{aligned}\psi_1(t) &= \sin(2\pi t) \exp(-(t/8)^2/2), \\ \psi_2(t) &= \sin(2\pi t) \exp(-(t/10)^2/2), \\ \psi_3(t) &= \sin(2\pi t) \exp(-(t/12)^2/2)\end{aligned}$$

を使った．最近 [3, 4] では，Meyer のウェーブレット関数あるいは Ausher [6] のウェーブレット関数の解析信号を複素数値ウェーブレット関数として利用している．

以下で，Ausher のウェーブレット関数およびその解析信号の作成方法を解説する．これは，Meyer のウェーブレット関数をフーリエ空間で高周波数方向へずらした関数に相当する．

まず，図 2 左のような，次の 3 条件を満たす補助関数 $\theta(t) \geq 0$ を用意する．

1. $\theta(t)$ は単調増大で滑らか．
2. $\theta(t) = 0, (\theta \leq 0), \quad \theta(t) = 1, (\theta \geq 1)$.
3. $(1/2, 1/2)$ で点対称，つまり， $\theta(t) + \theta(1-t) = 1$.

注意 5.1. 数値計算するときには，たとえば，

$$\theta(t) = \frac{\int_0^t x^2(1-x)^2 dx}{\int_0^1 x^2(1-x)^2 dx} = 6t^5 - 15t^4 + 10t^3, \quad 0 \leq t \leq 1$$

と取れば，2 階微分まで連続で上記 3 条件を満たす $\theta(t)$ ができる．

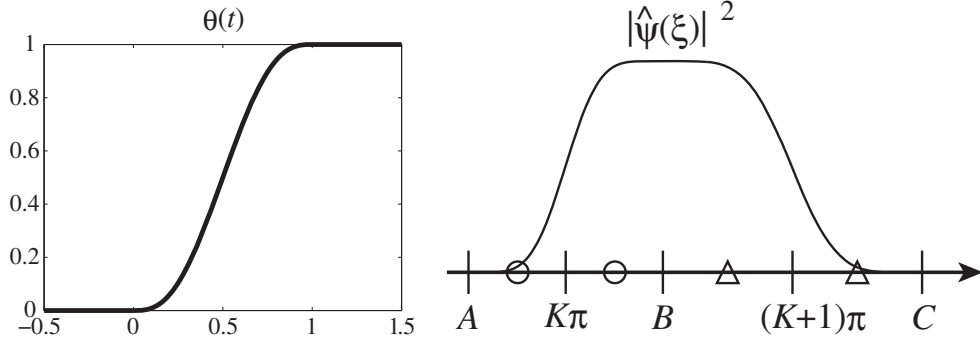


図 2: 補助関数 $\theta(t)$ と Ausher のウェーブレットのフーリエ変換 $|\hat{\psi}(\xi)|^2$.

次に、整数値のパラメータ $K > 0$ を固定する．次節以降の信号源分離の数値実験では、 $K = 10$ と取った． $\alpha = (K + 1)/K > 1$ と置く．図 2 右にあるように、

$$(K + 1)\pi = \alpha K\pi, \quad B = \alpha A, \quad C = \alpha B, \quad (A + B)/2 = K\pi, \quad (B + C)/2 = (K + 1)\pi$$

を満たすように定数 A, B, C をとる．このとき、

$$A = \frac{2\pi K^2}{2K + 1}, \quad B = \frac{2\pi K(K + 1)}{2K + 1}, \quad C = \frac{2\pi(K + 1)^2}{2K + 1}.$$

そして、ウェーブレット関数 $\psi(t)$ のフーリエ変換 $\hat{\psi}(\xi)$ を

$$\hat{\psi}(\xi) = \begin{cases} 0, & \xi \leq A, \\ 2e^{-i\xi/(2K)} \sin\left(\frac{\pi}{2}\theta\left(\frac{\xi-A}{B-A}\right)\right), & A \leq \xi \leq B, \\ 2e^{-i\xi/(2K)} \cos\left(\frac{\pi}{2}\theta\left(\frac{\xi-B}{C-B}\right)\right), & B \leq \xi \leq C, \\ 0, & C \leq \xi \end{cases}$$

で定める．このウェーブレット関数の実部 $\Re\psi(t)$ は、関数系

$$\{\alpha^{j/2} \Re\psi(\alpha^j t - k) ; j, k \in \mathbb{Z}\}$$

が $L^2(\mathbb{R})$ の完全正規直交系になり、**Ausher の正規直交ウェーブレット関数**とよばれている． $K = 1$ の場合の実部が **Meyer の正規直交ウェーブレット関数**である． $K = 1, 4$ の場合のウェーブレット関数のグラフ（実部と虚部）を図 3 にあげる．また、窓関数としての性質を表 1 にあげる．

このウェーブレットを用いる利点は、フーリエ空間でコンパクトサポートを持つ簡単な関数になるので、解析ウェーブレット変換が信号のどの周波数にアクセスしているかが分かりやすいことと、解析ウェーブレット変換の数値計算を高速フーリエ変換を用いて速くできることである．

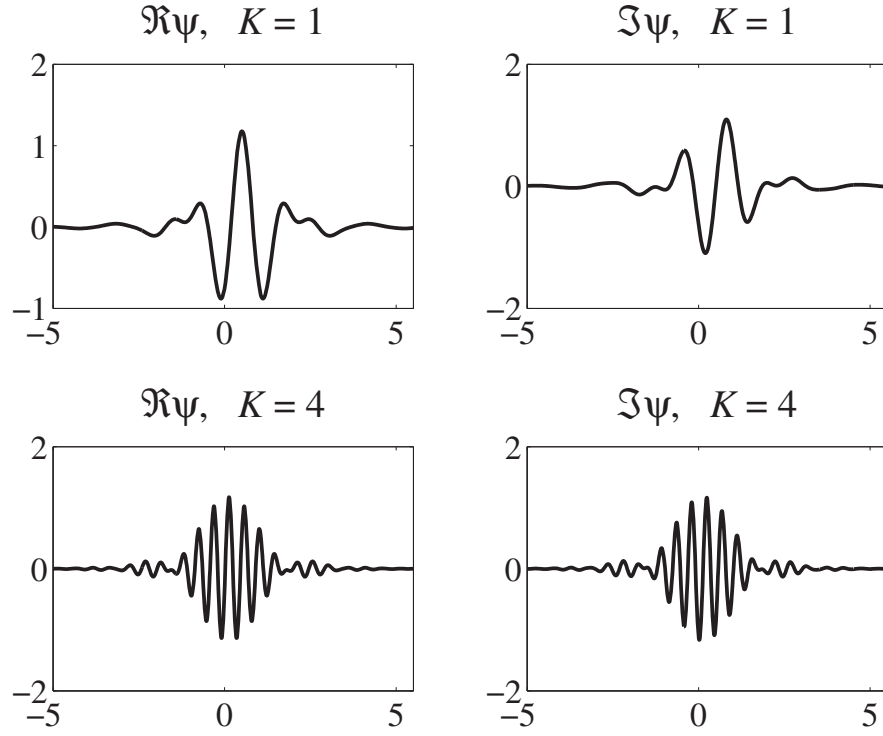


図 3: Ausher のウェーブレット関数の解析信号 $\psi(t)$, $K = 1, 4$ のグラフ.

K	$c[\psi] = 1/(2K)$	$\Delta[g]$	$c[\hat{\psi}]$ [Hz]	B [Hz]	$\Delta[\hat{\psi}]$ [Hz]	supp ψ [Hz]
1	0.5	0.63387	0.75566	0.66667	0.15973	[0.33333, 1.3333]
2	0.25	0.60994	1.2534	1.2	0.15885	[0.8, 1.8]
4	0.125	0.60134	2.2519	2.2222	0.15851	[1.7778, 2.7778]
6	0.083333	0.59939	3.2513	3.2308	0.15843	[2.7692, 3.7692]
8	0.0625	0.59865	4.251	4.2353	0.1584	[3.7647, 4.7647]
10	0.05	0.59829	5.2508	5.2381	0.15838	[4.7619, 5.7619]
12	0.041667	0.59809	6.2507	6.24	0.15837	[5.76, 6.76]
14	0.035714	0.59797	7.2506	7.2414	0.15837	[6.7586, 7.7586]
16	0.03125	0.59789	8.2505	8.2424	0.15837	[7.7576, 8.7576]
18	0.027778	0.59783	9.2505	9.2432	0.15836	[8.7568, 9.7568]
20	0.025	0.59779	10.2504	10.2439	0.15836	[9.7561, 10.7561]

表 1: $K > 0$ に対する Ausher のウェーブレットの解析信号の窓の大きさ. この表では周波数の単位は [Hz] を用いる. $B = \frac{K(K+1)}{2K+1}$ [Hz] は $|\hat{\psi}|$ が最大値を取る位置.

6 空間的混合問題

この節では，空間的混合問題の解法の仕方 [2, 14, 30, 31] を数値実験（実験 1 とよぶ）に沿って説明しよう．まず，空間的混合問題のモデルパラメータを設定して，信号源から観測信号を作成する．つぎに，観測信号の時間周波数情報の商のヒストグラムを描いて，信号源の数とモデルパラメータを推定する．最後に，分離した信号と元の信号の誤差評価を行う．ただし，時間周波数情報の作成を，Ausher のウェーブレットを用いた解析ウェーブレット変換で行っている点が，上に上げた文献（ガボールウェーブレットを使用）と異なっている．

6.1 数理モデルと観測信号の作成

$N = 4$ 個の信号源から， $M = 5$ 個の観測信号を作成する．実数値を取る元の信号 $s(t)$ と観測信号 $x(t)$ を

$$s(t) = [s_1(t), s_2(t), \dots, s_N(t)]^T, \quad x(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$$

と書く．元の信号 $s(t)$ から観測信号 $x(t)$ を作成するには，数理モデル

$$\begin{aligned} x(t) &= A s(t), \\ x_j(t) &= \sum_{k=1}^N a_{j,k} s_k(t), \quad j = 1, 2, \dots, M \end{aligned} \quad (6.1)$$

を用いる．ここで， $M \times N = 5 \times 4$ の混合行列 $A = (a_{j,k})$ として，

$$A = \begin{pmatrix} 1.0000 & 0.4472 & 0.4472 & 0.4851 \\ 0.2000 & 0.1857 & 0.4472 & 0.3714 \\ 0.4000 & 0.2031 & 0.2981 & 0.6325 \\ 0.4472 & 0.2425 & 0.1857 & 0.2169 \\ 0.2325 & 0.6325 & 0.2209 & 0.1789 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.2000 & 0.4152 & 1.0000 & 0.7656 \\ 0.4000 & 0.4541 & 0.6667 & 1.3038 \\ 0.4472 & 0.5423 & 0.4152 & 0.4472 \\ 0.2325 & 1.4142 & 0.4939 & 0.3688 \end{pmatrix}$$

を選ぶ．この混合行列は，第 7 節の一番簡単な時空間的混合問題の数値実験の混合行列と同じである．観測信号の時間周波数情報の商を取って解析するので，各行を第 1 行で割って正規化した行列 $\tilde{A}$ が推定したいモデルパラメータである． $\tilde{A}$ の 4 行目の第 1 列と，第 4 列が同じ値になっていることに注意せよ．

信号源として，日本建築学会の音声信号 [1] から取ってきた表 2 の 4 種類の音声信号を用いた．サンプリングレート $f_1 = 44100$ [Hz] の信号を $1/5$ にダウンサンプリングした音源を用いる．元の信号と観測信号のグラフを図 4 にあげる．サンプリングレート $f_0 = 8820$ [Hz] で約 7.4 [sec] (2^{16} 点) の音声信号である．ホームページ [32] に，数値実験で用いた音声ファイルを置いておく．

信号源	情報	
s_1	男性英語	Coffee is the most commonly ...
s_2	女性英語	Coffee is addictive ...
s_3	男性日本語	ワインは人類が最初に出会ったお酒 ...
s_4	女性日本語	あのヘミングウェイがこよなく愛したワイン ...

表 2: 信号源の説明, サンプルングレート $f_0 = 8820$ [Hz], または $f_1 = 44100$ [Hz].

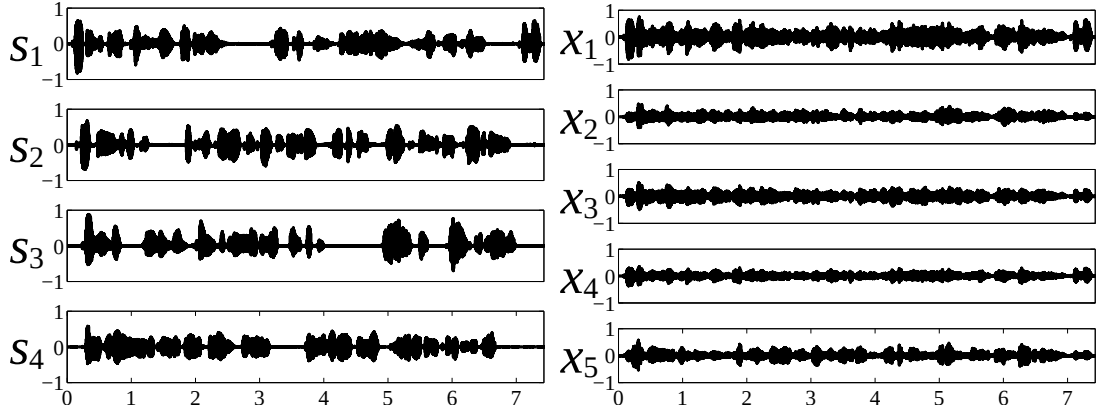


図 4: 実験 1: 空間的混合問題, 約 7.4 [sec] の元の信号 s と観測信号 x .

6.2 時間周波数情報の商を用いた解法

元の信号から観測信号を作る数値モデル (6.1) の両辺の解析ウェーブレット変換を取る. 元の信号と観測信号それぞれの, 時刻 t [sec], 周波数 ω [Hz] の時間周波数情報 $S_k(t, \omega)$ と $X_j(t, \omega)$ の関係式になおすと,

$$X_j(t, \omega) = \sum_{k=1}^N a_{j,k} S_k(t, \omega), \quad j = 1, 2, \dots, M$$

になる. 次に, 観測信号の時間周波数情報の商を取る.

$$Q_j(t, \omega) := \frac{X_j(t, \omega)}{X_1(t, \omega)} = \frac{\sum_{k=1}^N a_{j,k} S_k(t, \omega)}{\sum_{k=1}^N a_{1,k} S_k(t, \omega)} \in \mathbb{C}, \quad j = 2, \dots, M. \quad (6.2)$$

E_k を k 番目の信号源だけが活動している時間周波数領域とする.

$$E_k := \{(t, \omega); S_k(t, \omega) \neq 0, S_\ell(t, \omega) = 0, (\ell \neq k)\}.$$

実際は, $|S_k(t, \omega)|$ がある程度大きくその他の $|S_\ell(t, \omega)|$ が十分小さい領域を E_k とすればよい. 時間周波数位置 $(t, \omega) \in E_k$ の場合, 商は

$$Q_j(t, \omega) = \frac{\sum_{k=1}^N a_{j,k} S_k(t, \omega)}{\sum_{k=1}^N a_{1,k} S_k(t, \omega)} = \frac{a_{j,k} S_k(t, \omega)}{a_{1,k} S_k(t, \omega)} = \frac{a_{j,k}}{a_{1,k}} \in \mathbb{R}$$

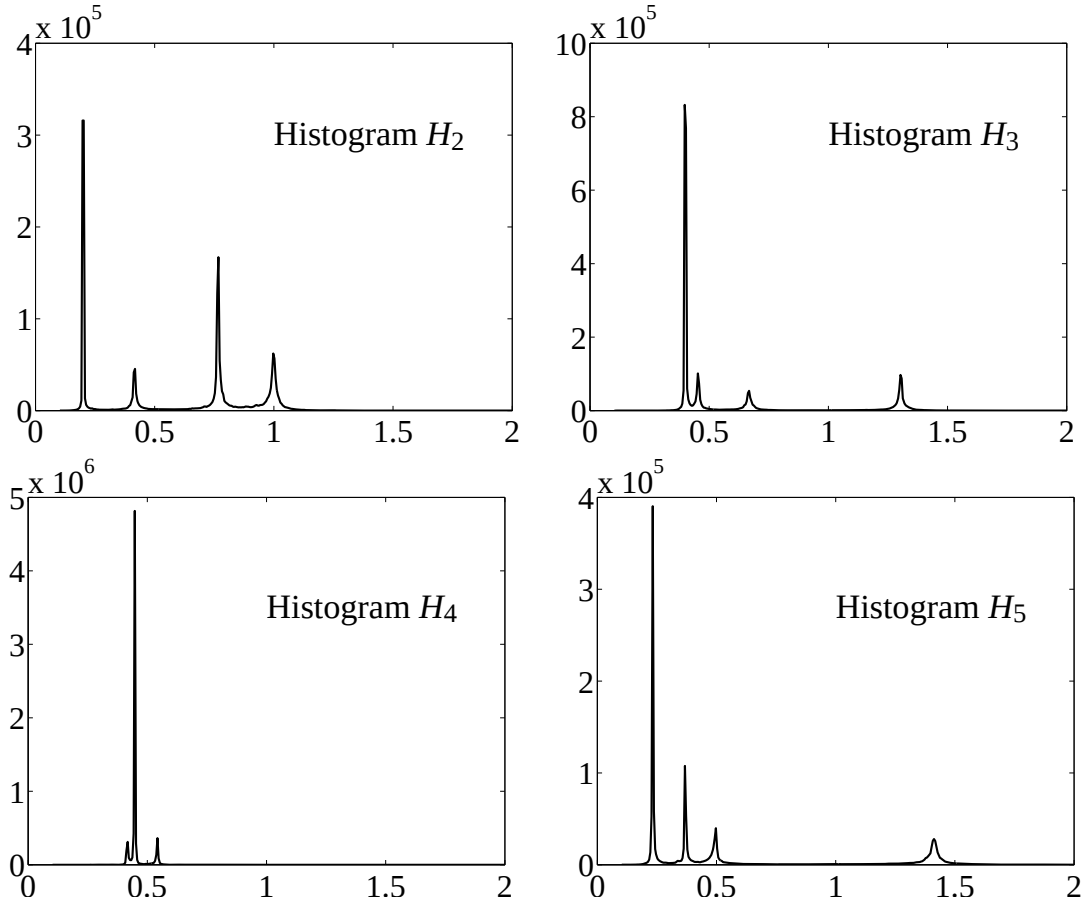


図 5: 実験 1：商のヒストグラム $H_j(x) = \#\{(t, \omega); Q_j(t, \omega) = x\}$ のグラフ.

になる．複素数値関数 $Q_j(t, \omega)$ が，同じ実数値 $a_{j,k}/a_{1,k}$ を少なくとも領域 E_k の面積分は取る．領域 E_k がある程度大きければ， $Q_j(t, \omega)$ が取る値の分布を調べれば，実数値 $a_{j,k}/a_{1,k}$ の所にピークが現れるので識別可能である．具体的には次のアルゴリズムを用いる．

アルゴリズム 6.1 (商のヒストグラムの作成). 以下の方法で，商 $Q_j(t, \omega)$, $j = 2, \dots, M$ のヒストグラムを描き，混合行列の各成分を推定する．

1. 実数値 x に対して，商が x に等しくなる時間周波数位置を数える．

$$H_j(x) := \#\{(t, \omega); Q_j(t, \omega) = x\}.$$

2. 横軸に x を取り，縦軸に個数 $H_j(x)$ をプロットし，ヒストグラムを図示する．
3. ヒストグラム $H_j(x)$ のピークの個数が，信号源の数 N の推定値になる．
4. $H_j(x)$ のピークを取る位置 x が混合行列の $j = 2$ 行目の推定値 $b_{2,k}$ と $j \geq 3$ 行目以降の推定値 $\tilde{b}_{j,k}$ を与える．

商 $Q_2 (H_2(x))$			商 $Q_3 (H_3(x))$		
Peaks	$b_{2,k}$	$\sharp(t, \omega)$	Peaks	$\tilde{b}_{3,\tilde{k}}$	$\sharp(t, \omega)$
$k = 1$	0.2025	314684	$\tilde{k} = 1$	0.3975	830760
$k = 2$	0.7675	166217	$\tilde{k} = 2$	0.4525	100078
$k = 3$	0.9975	61492	$\tilde{k} = 3$	1.3025	96989
$k = 4$	0.4175	44612	$\tilde{k} = 4$	0.6675	52652

商 $Q_4 (H_4(x))$			商 $Q_5 (H_5(x))$		
Peaks	$\tilde{b}_{4,\tilde{k}}$	$\sharp(t, \omega)$	Peaks	$\tilde{b}_{5,\tilde{k}}$	$\sharp(t, \omega)$
$\tilde{k} = 1$	0.4475	4814521	$\tilde{k} = 1$	0.2325	388205
$\tilde{k} = 2$	0.5425	362462	$\tilde{k} = 2$	0.3675	106718
$\tilde{k} = 3$	0.4175	310859	$\tilde{k} = 3$	0.4975	39104
			$\tilde{k} = 4$	1.4125	27537

表 3: 実験 1：空間的混合問題，ヒストグラム $H_j(x)$ のピークの数と位置と頻度．

注意 6.2. アルゴリズム 6.1 を実行する上で次のことに注意せよ．

1. 数値計算上は，複素数 $Q_j(t, \omega)$ に対して，虚部 $|\Im Q_j(t, \omega)|$ が絶対値 $|Q_j(t, \omega)|$ に比べて十分小さい時間周波数位置 (t, ω) をピックアップして，実部 $\Re Q_j(t, \omega)$ が $x - \varepsilon/2$ と $x + \varepsilon/2$ の間に入る (t, ω) の数を数えた．ただし， ε はヒストグラムの bin size である．
2. 混合行列の $j \geq 3$ 行目以降の推定値 $\tilde{b}_{j,\tilde{k}}$ がどの列（信号源）に対応する成分になるかは，次のアルゴリズム 6.3 で定める． $j = 2$ の時に $b_{2,k}$ のようにチルダを付ける必要がないのは， $b_{2,k}$ を基準にして $j \geq 3$ の列を定めるからである．

実験 1 で，各 $j = 2, \dots, M$ に対して，商のヒストグラム $H_j(x)$ を描くと図 5 を得る．この図からピークの個数は 4 なので，信号源の数 $N = 4$ と推定する．図 5 の左下の $H_4(x)$ の場合に，ピークは 3 個しか捉えられていないが，これは正規化した混合行列 $\tilde{A}$ の $j = 4$ 行目の第 1 列と，第 4 列が同じ値になっていることが原因である．各 $j = 2, \dots, M$ の商のヒストグラム $H_j(x)$ のピークを取る $x = b_{2,k}$ ， $\tilde{b}_{j,\tilde{k}}$ とピーク値 $\sharp(t, \omega)$ をピークの高い順にピックアップすると，表 3 を得る．

このヒストグラムを実数値ウェーブレット関数を用いた連続ウェーブレット変換で作成すると図 6 になる．図 5 の複素数値関数 $Q_j(t, \omega)$ が特定の実数値になる場合を探すのと比較して，実数値関数 $Q_j(t, \omega)$ が特定の実数値を取るのを探すのは困難で，図 6 では山の裾野が高く広がってピークが読み取り辛い．この図 6 では，Ausher $K = 10$ の正規直交ウェーブレット関数を用いた．

次に，各 $j = 2, \dots, M$ の商のヒストグラム $H_j(x)$ 間で，ピークの対応付けを行うアルゴリズム 6.3 を説明する．このアルゴリズムを使うことにより，推定し

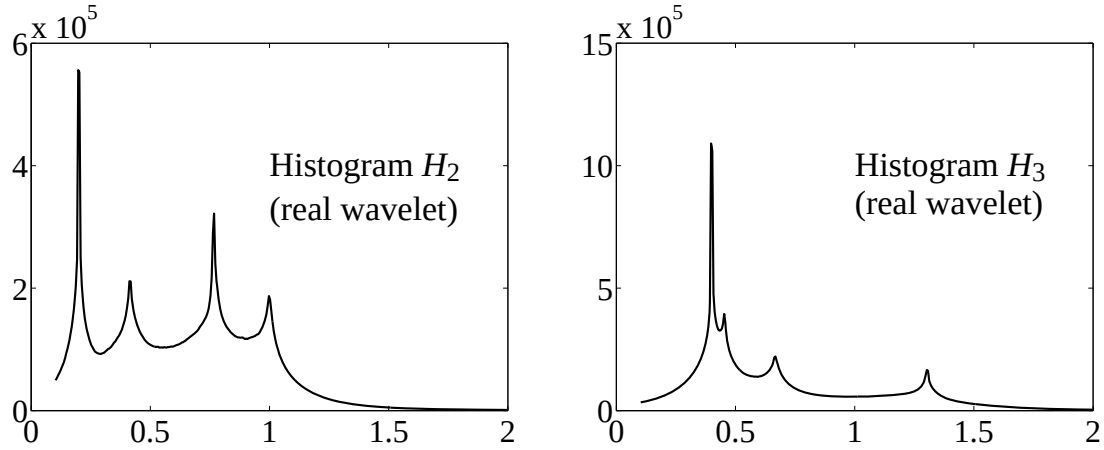


図 6: 実験 1: 空間的混合問題, 実数値ウェーブレット関数を用いた時間周波数情報から作成した商のヒストグラム $H_j(x)$, $j = 2, 3$ の場合.

た混合行列の各成分を正しい位置に配置することができる.

アルゴリズム 6.3 (ピークの対応付け). 商 $Q_2(t, \omega)$ のヒストグラム $H_2(x)$ がピークを取る時間周波数位置が商 $Q_j(t, \omega)$, $j \geq 3$ のヒストグラム $H_j(x)$ のどのピークに対応するかを以下の手順で調べる.

1. $H_2(x)$ が k 番目のピーク $b_{2,k}$ を取る時間周波数位置 (t, ω) を 1000 点選ぶ.

$$Y_k := \{(t, \omega); Q_2(t, \omega) = b_{2,k}, 1000 \text{ 点選ぶ}\}.$$

2. $j \geq 3$ の商 $Q_j(t, \omega)$ に $(t, \omega) \in Y_k$ を代入して次の集合の個数を数える.

$$\#\left\{(t, \omega) \in Y_k; Q_j(t, \omega) = \tilde{b}_{j,k}\right\}.$$

3. 個数の一番大きい集合に対応する $\tilde{b}_{j,k}$ を $b_{j,k}$ にとる.
4. 混合行列を $B = (b_{j,k})$ と推定する. ただし, B の第 1 行は全て 1 とする.

アルゴリズム 6.3 にしたがってヒストグラム $H_2(x)$ の各ピークが $H_j(x)$, $j \geq 3$ のどのピークに対応するかを調べたのが表 4 である. 表中の数値は, アルゴリズム 6.3 の手順 2. で数えた集合の要素数である. $b_{2,4}$ のピークは, $\tilde{b}_{3,2}$, $\tilde{b}_{4,2}$, $\tilde{b}_{5,4}$ に対応していることが分かる. こうして混合行列の第 4 列が推定できる. ピークが 3 個しか捉えられなかった H_4 では, $b_{2,1}$ と $b_{2,2}$ が H_4 の同じピーク $\tilde{b}_{4,1}$ に対応していることが分かる.

$H_2 \setminus H_3$	$\tilde{b}_{3,\bar{1}}$	$\tilde{b}_{3,\bar{2}}$	$\tilde{b}_{3,\bar{3}}$	$\tilde{b}_{3,\bar{4}}$
$b_{2,1}$	1000	0	0	0
$b_{2,2}$	0	0	980	0
$b_{2,3}$	0	0	0	939
$b_{2,4}$	0	893	0	0

$H_2 \setminus H_4$	$\tilde{b}_{4,\bar{1}}$	$\tilde{b}_{4,\bar{2}}$	$\tilde{b}_{4,\bar{3}}$
$b_{2,1}$	1000	0	0
$b_{2,2}$	1000	0	0
$b_{2,3}$	0	0	1000
$b_{2,4}$	33	963	0

$H_2 \setminus H_5$	$\tilde{b}_{5,\bar{1}}$	$\tilde{b}_{5,\bar{2}}$	$\tilde{b}_{5,\bar{3}}$	$\tilde{b}_{5,\bar{4}}$
$b_{2,1}$	1000	0	0	0
$b_{2,2}$	0	904	0	0
$b_{2,3}$	0	0	929	0
$b_{2,4}$	0	0	0	474

表 4: 実験 1: 空間的混合問題, 各商のピーク位置の対応.

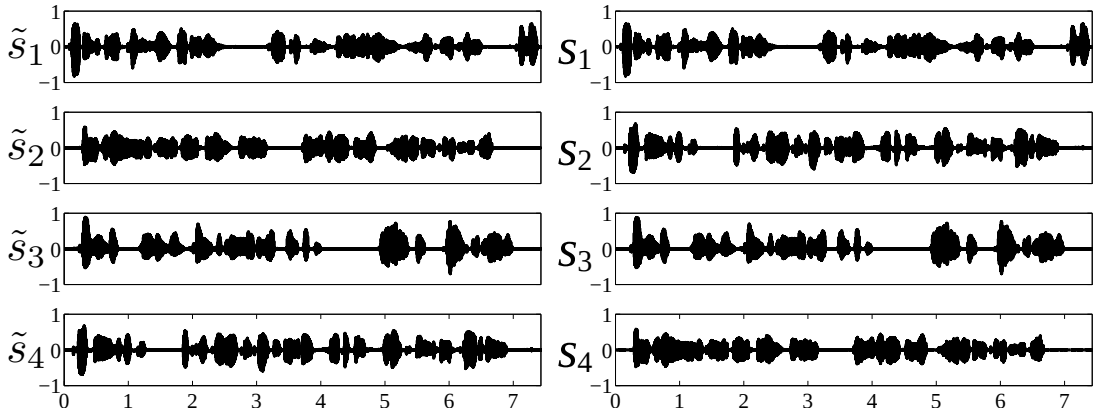


図 7: 実験 1: 分離した信号 $\tilde{s}$ と元の信号 s .

混合行列 B を推定する. 横に正規化したモデルパラメータ $\tilde{A}$ をならべる.

$$B = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.2025 & 0.7675 & 0.9975 & 0.4175 \\ 0.3975 & 1.3025 & 0.6675 & 0.4525 \\ 0.4475 & 0.4475 & 0.4175 & 0.5425 \\ 0.2325 & 0.3675 & 0.4975 & 1.4125 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.2000 & 0.4152 & 1.0000 & 0.7656 \\ 0.4000 & 0.4541 & 0.6667 & 1.3038 \\ 0.4472 & 0.5423 & 0.4152 & 0.4472 \\ 0.2325 & 1.4142 & 0.4939 & 0.3688 \end{pmatrix}.$$

推定した混合行列 B の第 1, 2, 3, 4 列は, それぞれモデルパラメータ $\tilde{A}$ の第 1, 4, 3, 2 列に対応していることが分かる. 両者は小数点 2 桁までは一致している. 信号の分離は, 分離する信号 $\tilde{s}(t)$ と観測信号 $x(t)$ の関係式

$$x(t) = B\tilde{s}(t)$$

ESS	SS	Max error	L^1 error	L^2 error	SNR
$\tilde{s}_1$	s_1	0.018 %	0.020 %	0.017 %	75.280 dB
$\tilde{s}_2$	s_4	1.381 %	1.162 %	1.051 %	39.565 dB
$\tilde{s}_3$	s_3	1.380 %	2.360 %	1.667 %	35.560 dB
$\tilde{s}_4$	s_2	0.643 %	0.640 %	0.510 %	45.847 dB

表 5: 実験 1: 空間的混合問題の誤差評価. ESS は分離した信号, SS は元の信号.

に混合行列 B の一般化逆行列をかけて解けば良い. こうして分離した信号 $\tilde{s}_k$ のグラフを図 7 に誤差評価を表 5 にのせる. ただし, この誤差評価では, 分離した信号 $\tilde{s}_k(t)$ と対応する元の信号 $s_{n(k)}(t)$ の ℓ^p , $p = \infty, 1, 2$ ノルムを 1 に修正した後, 各ノルムでの誤差評価を行った. つまり, 誤差は,

$$\left\| \frac{\tilde{s}_k}{\|\tilde{s}_k\|_{\ell^p}} - \frac{s_{n(k)}}{\|s_{n(k)}\|_{\ell^p}} \right\|_{\ell^p} \times 100 \%, \quad p = \infty, 1, 2$$

を計算して表にのせた. SNR も

$$20 \log_{10} \left(\left\| \frac{\tilde{s}_k}{\|\tilde{s}_k\|_{\ell^2}} - \frac{s_{n(k)}}{\|s_{n(k)}\|_{\ell^2}} \right\|_{\ell^2} \right)$$

で計算した. ただし, k 番目の推定信号源は, $n(k)$ 番目の信号源に対応している.

6.3 独立成分分析による解法

Cichocki-Amari [10] のサポートホームページ [33] から, ICALAB という MATLAB で空間的混合問題の独立成分分析による解法を行うプログラムが入手できる.

本小節では, このプログラムを用いて, 実験 1 の観測信号から信号源の分離を行う. 実験 1 では 5 個の観測信号を使って 4 個の信号源を取り出す. そして独立成分分析では最初に信号源の数を与えないといけないので, 信号源の数が観測信号の数 5 以下に設定できるアルゴリズムでないとうまく働かない. ICALAB では, 18 種類のアルゴリズムが実装されていて, その内の “2. EVD 2” というアルゴリズムを用いた.

分離するときに推定した混合行列 B_1 と正規化したモデルパラメータ $\tilde{A}$ は,

$$B_1 = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.1590 & 0.7008 & 0.9513 & 0.4622 \\ 0.3552 & 1.3282 & 0.7110 & 0.5449 \\ 0.4469 & 0.4385 & 0.4131 & 0.5231 \\ 0.2105 & 0.2089 & 0.4103 & 1.2140 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.2000 & 0.4152 & 1.0000 & 0.7656 \\ 0.4000 & 0.4541 & 0.6667 & 1.3038 \\ 0.4472 & 0.5423 & 0.4152 & 0.4472 \\ 0.2325 & 1.4142 & 0.4939 & 0.3688 \end{pmatrix}.$$

である. B_1 の第 1, 2, 3, 4 列は, モデルパラメータ $\tilde{A}$ の第 1, 4, 3, 2 列に対応している. ただし, 独立成分分析では, B_1 ではなく, B_1 の一般化逆行列を推定し

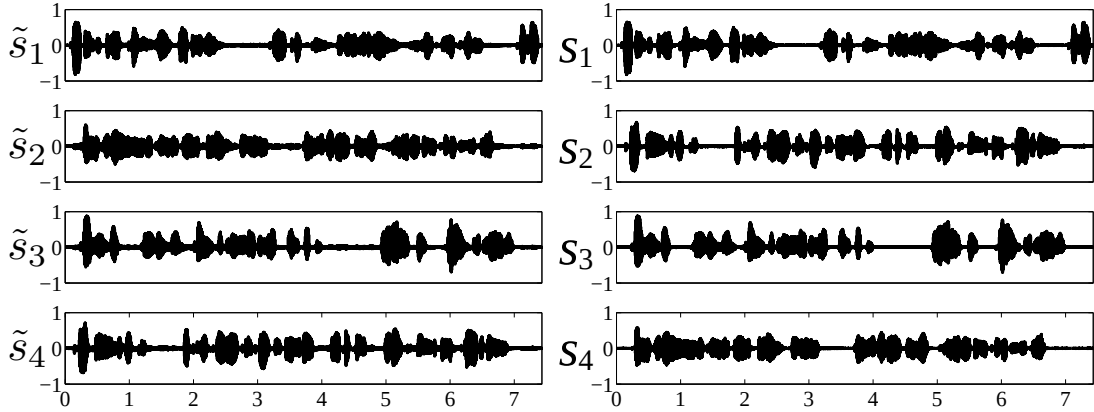


図 8: 実験 1 : ICALAB で分離した信号 $\tilde{s}$ と元の信号 s .

ESS	SS	Max error	L^1 error	L^2 error	SNR
$\tilde{s}_1$	s_1	6.064 %	10.731 %	8.369 %	21.546 dB
$\tilde{s}_2$	s_4	22.395 %	20.149 %	16.522 %	15.639 dB
$\tilde{s}_3$	s_3	7.465 %	16.332 %	11.016 %	19.159 dB
$\tilde{s}_4$	s_2	12.072 %	18.315 %	14.034 %	17.057 dB

表 6: 実験 1 : ICALAB の誤差評価. ESS は分離した信号, SS は元の信号.

ているので, B_1 の誤差が大きく感じられるのは仕方ない. 分離した信号の外形を図 8 に描く. また, 誤差評価を表 6 にのせる. ICALAB を用いた場合にはかなりの誤差があることがわかる. この程度の誤差でも, 分離した信号を耳で聞く場合には, 他の音声信号が小さいレベルで混じっている程度で,それほど問題なく聞き取れる. ICALAB の実行時間は, 1 [sec] 程度で非常に早く, 実時間処理も可能である. 一方, 我々のアルゴリズムでは, 280 [sec] ほどかかっている.

7 一番簡単な時空間的混合問題

この節では, 一番簡単な時空間的混合問題の解法アルゴリズム [3, 4] を数値実験 (実験 2 とよぶ) に沿って説明しよう. 最初に, 信号源の位置と観測地点を設定して平面上の音の伝播モデルにしたがって, モデルパラメータを設定する. そして, 元の信号から数理モデルにしたがって観測信号を作成する.

作成した観測信号と数理モデルの仮定から, 時間遅れの入った観測信号の時間周波数情報の商の 3D ヒストグラムを描いて, 信号源の個数とモデルパラメータを推定する. 信号が到着する時間差から信号源の位置推定を行う. 信号を分離し分離した信号の誤差評価を行う.

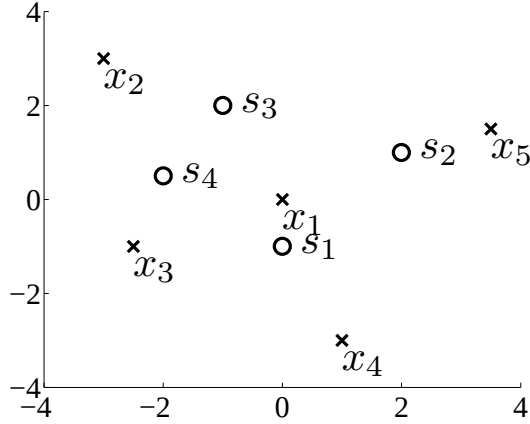


図 9: 実験 2：一番簡単な時空間的混合問題の信号源の位置と観測地点.

7.1 数理モデルと観測信号の作成

図 9 にあるように $N = 4$ 個の信号源と $M = 5$ 個の観測地点を配置しよう. 観測地点 P_{x_j} と信号源の位置 P_{s_k} は,

$$P_{x_1} = (0, 0), P_{x_2} = (-3, 3), P_{x_3} = (-2.5, -1), P_{x_4} = (1, -3), P_{x_5} = (3.5, 1.5),$$

$$P_{s_1} = (0, -1), P_{s_2} = (2, 1), P_{s_3} = (-1, 2), P_{s_4} = (-2, 0.5).$$

音速を $V = 330$ [m/sec] とし、音声信号が距離に反比例して弱まると仮定すると、元の信号から観測信号を作る数理モデルは,

$$x_j(t) = \sum_{k=1}^N a_{j,k} s_k(t - c_{j,k}), \quad j = 1, 2, \dots, M \quad (7.1)$$

で書き表すことができる. ただし, j 番目の観測地点 P_{x_j} と k 番目の信号源の位置 P_{s_k} の間の距離を $r_{j,k} = \|P_{x_j} - P_{s_k}\|$ としたとき, $a_{j,k} = 1/r_{j,k}$, $c_{j,k} = r_{j,k}/V$ である. $A = (a_{j,k})$ を混合行列, $C = (c_{j,k})$ を時間遅れ行列または遅延行列とよぶ.

混合行列 $A \in \mathbb{R}^{M \times N}$ と A の各行を第 1 行で割って正規化した $\tilde{A}$ は,

$$A = \begin{pmatrix} 1.0000 & 0.4472 & 0.4472 & 0.4851 \\ 0.2000 & 0.1857 & 0.4472 & 0.3714 \\ 0.4000 & 0.2031 & 0.2981 & 0.6325 \\ 0.4472 & 0.2425 & 0.1857 & 0.2169 \\ 0.2325 & 0.6325 & 0.2209 & 0.1789 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.2000 & 0.4152 & 1.0000 & 0.7656 \\ 0.4000 & 0.4541 & 0.6667 & 1.3038 \\ 0.4472 & 0.5423 & 0.4152 & 0.4472 \\ 0.2325 & 1.4142 & 0.4939 & 0.3688 \end{pmatrix}$$

である. 離散信号である元の信号を数理モデル (7.1) にしたがって混ぜ合わせる必要があるので, 遅延行列はサンプリング間隔 $1/f_0$ [sec] の整数倍に丸める必要が

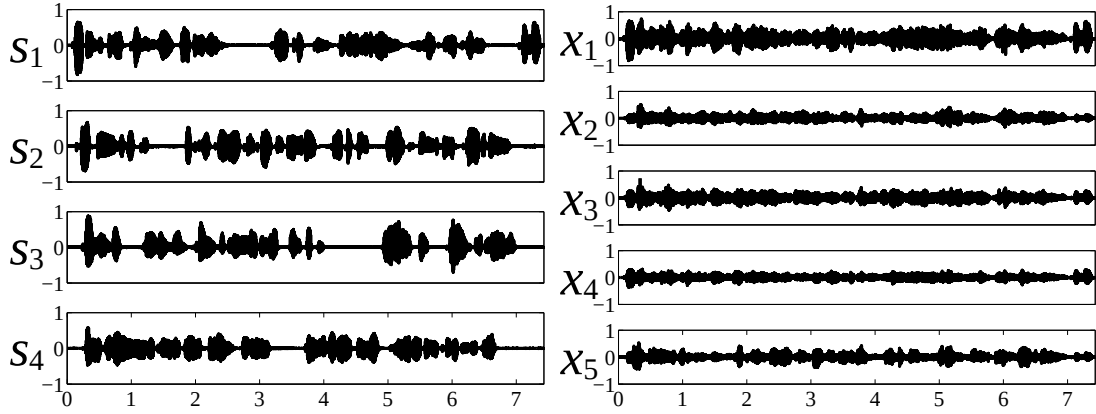


図 10: 実験 2 : 約 7.4 [sec] の元の信号 s と観測信号 x .

ある．ただし， $f_0 = 8820$ [Hz] はサンプリングレートである．この丸めた遅延行列 $C \in \mathbb{R}^{M \times N}$ と C の各行から第 1 行を引いて正規化した遅延行列 $\tilde{C}$ は，

$$C = \frac{1}{f_0} \begin{pmatrix} 27 & 60 & 60 & 55 \\ 134 & 144 & 60 & 72 \\ 67 & 132 & 90 & 42 \\ 60 & 110 & 144 & 123 \\ 115 & 42 & 121 & 149 \end{pmatrix}, \quad \tilde{C} = \frac{1}{f_0} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 107 & 84 & 0 & 17 \\ 40 & 72 & 30 & -13 \\ 33 & 50 & 84 & 68 \\ 88 & -18 & 61 & 94 \end{pmatrix}$$

である．時間周波数情報の商を使った解析方法では，モデルパラメータとして $\tilde{A}$ と $\tilde{C}$ を推定する．前節の実験 1 と同じ表 2 の信号源を用いて，観測信号を数理モデル (7.1) にしたがって作成する．すると，図 10 の観測信号を得る．元の信号と観測信号はともに，サンプリングレート $f_0 = 8820$ [Hz]，約 7.4 [sec] の 2^{16} 点からなる離散信号である．

7.2 3 D ヒストグラムを用いたモデルパラメータの推定

数理モデル (7.1) の両辺を解析ウェーブレット変換し，観測信号と元の信号それぞれの時間周波数情報 $X_j(t, \omega)$ と $S_k(t, \omega)$ の関係式に書き直すと，

$$X_j(t, \omega) = \sum_{k=1}^N a_{j,k} S_k(t - c_{j,k}, \omega), \quad j = 1, 2, \dots, M. \quad (7.2)$$

注意 7.1. 数理モデル (7.1) の両辺を短時間フーリエ変換して，観測信号と元の信号それぞれの時間周波数情報 $\tilde{X}_j(t, \omega)$ と $\tilde{S}_k(t, \omega)$ で書き直すと，次式のようにモジュレーションが入って複雑になる．

$$\tilde{X}_j(t, \omega) = \sum_{k=1}^N e^{-ic_{j,k}2\pi\omega} a_{j,k} \tilde{S}_k(t - c_{j,k}, \omega).$$

時間遅れ δ の入った、次の観測信号の時間周波数情報の商を考える．

$$Q_j(\delta, t, \omega) := \frac{X_j(t + \delta, \omega)}{X_1(t, \omega)} \in \mathbb{C}, \quad j = 2, \dots, M. \quad (7.3)$$

遅延行列 C を考慮した上で、信号源 s_k のみが活動している時間周波数領域

$$E_{C,k} := \{ (t, \omega); S_k(t - c_{j,k}, \omega) \neq 0, S_l(t - c_{j,l}, \omega) = 0 \text{ for } \forall l \neq k, \forall j \}$$

を定義する．このとき、 $(t, \omega) \in E_{C,k}$, $\delta = c_{j,k} - c_{1,k} = \tilde{c}_{j,k}$ に対して、複素数値を取る商 $Q_j(\delta, t, \omega)$ は、

$$\begin{aligned} Q_j(\delta, t, \omega) &= \frac{X_j(t + \delta, \omega)}{X_1(t, \omega)} = \frac{\sum_{k=1}^n a_{j,k} S_k(t + \delta - c_{j,k}, \omega)}{\sum_{k=1}^n a_{1,k} S_k(t - c_{1,k}, \omega)} \\ &= \frac{a_{j,k} S_k(t - c_{1,k}, \omega)}{a_{1,k} S_k(t - c_{1,k}, \omega)} = \frac{a_{j,k}}{a_{1,k}} \in \mathbb{R} \end{aligned} \quad (7.4)$$

と同一の実数値 $a_{j,k}/a_{1,k}$ を取る．これは正規化した混合行列 $\tilde{A}$ の (j, k) 成分に相当する．したがって、領域 $E_{C,k}$ の面積が十分大きいと以下のアルゴリズム 7.2 により、時間遅れ $\tilde{c}_{j,k}$ と実数値 $a_{j,k}/a_{1,k}$ は推定可能である．

アルゴリズム 7.2 (商の 3 D ヒストグラム). 以下の手順で、商 Q_j , $j = 2, \dots, M$ の 3 D ヒストグラムを作成し、遅延行列と混合行列の各成分を推定する．

1. 時間遅れ δ と実数 x に対して、商 $Q_j(\delta, t, \omega)$ が実数値 x に等しくなる時間周波数位置の個数を数える．

$$H_j(\delta, x) := \# \{ (t, \omega); Q_j(\delta, t, \omega) = x \}.$$

2. $\delta - x$ 平面に $H_j(\delta, x)$ をプロットして、商の 3 D ヒストグラムを作成する．
3. ヒストグラム $H_j(\delta, x)$ のピークの個数が信号源の個数 N の推定値である．
4. $H_j(\delta, x)$ は、 $\tilde{k}$ 番目のピークを $(\tilde{\delta}_{j,\tilde{k}}, \tilde{b}_{j,\tilde{k}})$ で取るとする．このとき、 $\tilde{\delta}_{j,\tilde{k}}$ は正規化した遅延行列 $\tilde{C}$ の $(j, \tilde{k})$ 成分 $\tilde{c}_{j,\tilde{k}}$ の推定値であり、 $\tilde{b}_{j,\tilde{k}}$ は正規化した混合行列 $\tilde{A}$ の $(j, \tilde{k})$ 成分 $a_{j,\tilde{k}}/a_{1,\tilde{k}}$ の推定値である．

注意 7.3. $j = 2$ の場合の時間遅れ $\delta_{2,k}$ と混合係数 $b_{2,k}$ には $\sim$ を付けない．

実験 2 で、アルゴリズム 7.2 にしたがって、商 Q_j の 3 D ヒストグラム $H_j(\delta, x)$ を描くと図 11 を得る．各 $j = 2, \dots, M$ に対して、4 個のピークが確認できるので、信号源の数は $N = 4$ と推定する．各 j に対して、ピークの高い方から番号 $k = 1, \dots, 4$, $\tilde{k} = \tilde{1}, \dots, \tilde{4}$ を打って、そのピークに対応する時間遅れ $\tilde{\delta}_{j,\tilde{k}}$ と混合係数 $\tilde{b}_{j,\tilde{k}} = a_{j,\tilde{k}}/a_{1,\tilde{k}}$ を読み取ると、表 7 を得る．ただし、時間遅れ δ はサンプリング間隔 $1/f_0 = 1/8820$ [sec] の整数倍しか動かせない．

遅延行列と混合行列を求めるためには、各 j の商 Q_j の 3 D ヒストグラム $H_j(\delta, x)$ のピークを対応付けて、各行の正しい列に、表 7 で求めた成分をはめ込まなければならない．このために次のアルゴリズム 7.4 を用いる．

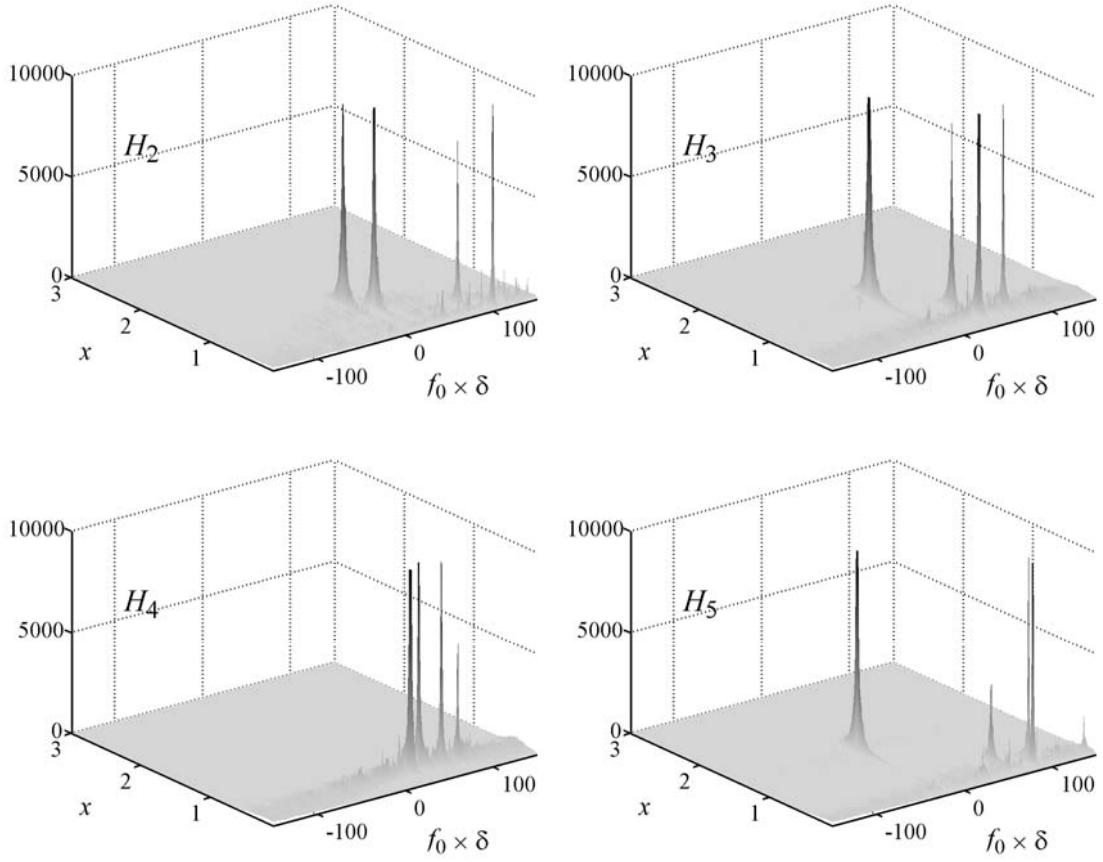


図 11: 実験 2 : 3 D ヒストグラム $H_j(\delta, x) = \# \{(t, \omega); Q_j(\delta, t, \omega) = x\}$.

アルゴリズム 7.4 (ピークに対応付け). 以下の手順で, 商の 3 D ヒストグラム $H_j(\delta, x)$, $j = 2, \dots, M$ のピークを対応付けて, 遅延行列と混合行列を完成させる.

1. $H_2(\delta, x)$ が k 番目のピークを取る, 時間周波数位置 (t, ω) を 1000 点選ぶ.

$$Y_k := \{(t, \omega); Q_2(\delta_{2,k}, t, \omega) = b_{2,k}, 1000 \text{ 点選ぶ}\}.$$

2. $j \geq 3$ の商 $Q_j(\tilde{\delta}_{j,\tilde{k}}, t, \omega)$ に $(t, \omega) \in Y_k$ を代入して次の集合の個数を数える.

$$\# \left\{ (t, \omega) \in Y_k; Q_j(\tilde{\delta}_{j,\tilde{k}}, t, \omega) = \tilde{b}_{j,\tilde{k}} \right\}.$$

3. 個数の一番大きい集合に対応する $\tilde{\delta}_{j,\tilde{k}}$ を $\delta_{j,k}$ に $\tilde{b}_{j,\tilde{k}}$ を $b_{j,k}$ にとる.
4. 遅延行列 $\Delta = (\delta_{j,k})$ と混合行列 $B = (b_{j,k})$ を定める. ただし, Δ の第 1 行は全て 0, B の第 1 行は全て 1 とする.

商 $Q_2 (H_2)$				商 $Q_3 (H_3)$			
Peaks	$\delta_{2,k}$	$b_{2,k}$	$\sharp(t, \omega)$	Peaks	$\tilde{\delta}_{3,\tilde{k}}$	$\tilde{b}_{3,\tilde{k}}$	$\sharp(t, \omega)$
$k = 1$	$107/f_0$	0.1975	184538	$\tilde{k} = 1$	$40/f_0$	0.4025	376585
$k = 2$	$17/f_0$	0.7675	65630	$\tilde{k} = 2$	$-13/f_0$	1.3075	51871
$k = 3$	$0/f_0$	0.9975	19111	$\tilde{k} = 3$	$72/f_0$	0.4575	11973
$k = 4$	$84/f_0$	0.4125	8376	$\tilde{k} = 4$	$30/f_0$	0.6625	9678

商 $Q_4 (H_4)$				商 $Q_5 (H_5)$			
Peaks	$\tilde{\delta}_{4,\tilde{k}}$	$\tilde{b}_{4,\tilde{k}}$	$\sharp(t, \omega)$	Peaks	$\tilde{\delta}_{5,\tilde{k}}$	$\tilde{b}_{5,\tilde{k}}$	$\sharp(t, \omega)$
$\tilde{k} = 1$	$33/f_0$	0.4475	772296	$\tilde{k} = 1$	$88/f_0$	0.2325	208034
$\tilde{k} = 2$	$68/f_0$	0.4475	51913	$\tilde{k} = 2$	$94/f_0$	0.3675	26337
$\tilde{k} = 3$	$50/f_0$	0.5425	25089	$\tilde{k} = 3$	$-18/f_0$	1.4025	15199
$\tilde{k} = 4$	$84/f_0$	0.4175	6282	$\tilde{k} = 4$	$61/f_0$	0.4925	4740

表 7: 実験 2：一番簡単な時空間的混合問題，3 D ヒストグラムのピーク位置．

$H_2 \setminus H_3$	$\tilde{b}_{3,\tilde{1}}$	$\tilde{b}_{3,\tilde{2}}$	$\tilde{b}_{3,\tilde{3}}$	$\tilde{b}_{3,\tilde{4}}$
$b_{2,1}$	1000	0	0	0
$b_{2,2}$	0	517	0	0
$b_{2,3}$	0	0	0	524
$b_{2,4}$	0	0	708	0

$H_2 \setminus H_4$	$\tilde{b}_{4,\tilde{1}}$	$\tilde{b}_{4,\tilde{2}}$	$\tilde{b}_{4,\tilde{3}}$	$\tilde{b}_{4,\tilde{4}}$
$b_{2,1}$	1000	0	0	0
$b_{2,2}$	0	800	0	0
$b_{2,3}$	0	0	0	606
$b_{2,4}$	0	0	873	0

$H_2 \setminus H_5$	$\tilde{b}_{5,\tilde{1}}$	$\tilde{b}_{5,\tilde{2}}$	$\tilde{b}_{5,\tilde{3}}$	$\tilde{b}_{5,\tilde{4}}$
$b_{2,1}$	1000	0		0
$b_{2,2}$	0	501	0	0
$b_{2,3}$	0	0	0	371
$b_{2,4}$	0	0	366	0

表 8: 実験 2：一番簡単な時空間的混合問題，各商のピークの対応付け．

実験 2 で，アルゴリズム 7.4 に沿って，ピークの対応付けを行うと，表 8 の結果を得る．表中の数値は，手順 2. で数えた集合の要素数である．たとえば， H_2 の 3 番目のピークは， H_3, H_4, H_5 の 4 番目のピークに対応していることが分かる．こうして，推定した時間遅れ行列 Δ とモデルパラメータ $\tilde{C}$ をならべて書くと，

$$\Delta = \frac{1}{f_0} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 107 & 17 & 0 & 84 \\ 40 & -13 & 30 & 72 \\ 33 & 68 & 84 & 50 \\ 88 & 94 & 61 & -18 \end{pmatrix}, \quad \tilde{C} = \frac{1}{f_0} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 107 & 84 & 0 & 17 \\ 40 & 72 & 30 & -13 \\ 33 & 50 & 84 & 68 \\ 88 & -18 & 61 & 94 \end{pmatrix}$$

である． Δ の第 1, 2, 3, 4 列が， $\tilde{C}$ の第 1, 4, 3, 2 列に対応している．両者は完全に一致している．推定した混合行列 B とモデルパラメータ $\tilde{A}$ をならべて書くと，

$$B = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.1975 & 0.7675 & 0.9975 & 0.4125 \\ 0.4025 & 1.3075 & 0.6625 & 0.4575 \\ 0.4475 & 0.4475 & 0.4175 & 0.5425 \\ 0.2325 & 0.3675 & 0.4925 & 1.4025 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.2000 & 0.4152 & 1.0000 & 0.7656 \\ 0.4000 & 0.4541 & 0.6667 & 1.3038 \\ 0.4472 & 0.5423 & 0.4152 & 0.4472 \\ 0.2325 & 1.4142 & 0.4939 & 0.3688 \end{pmatrix}$$

を得る． B の第 1, 2, 3, 4 列が， $\tilde{A}$ の第 1, 4, 3, 2 列に対応している．両者は，小数点第 2 位まで一致している．商の 3D ヒストグラムを用いれば，数理モデルのパラメータを高精度で推定可能である．

7.3 時間遅れからの信号源の位置の推定

k 番目の推定信号源（位置 $P_{\tilde{s}_k}$ ）からの信号が，1 番目の観測地点 P_{x_1} と j 番目の観測地点 P_{x_j} に到着するときの時間差が $\delta_{j,k}$ として推定できた． $\delta_{j,k}$ は $\tilde{c}_{j,k} = c_{1,k} - c_{j,k}$ の推定値である．したがって， $P_{\tilde{s}_k}$ は P_{x_1} と P_{x_j} を焦点とし距離の差が $\delta_{j,k}V$ の双曲線上にある．ただし， $V = 330$ [m/sec] は音速である．

信号源の位置の推定は，双曲線の交点を求める Mellin-Pachter-Raquet [23] の方法を用いて行った．観測地点の凸包内に信号源があるので，信号源の位置は一意に定まる（文献 [4]）．求めた信号源の位置は，

$$\begin{aligned} P_{\tilde{s}_1} &= (0.0058, -0.9988), & P_{\tilde{s}_2} &= (-1.9821, 0.5032), \\ P_{\tilde{s}_3} &= (-0.9911, 1.9991), & P_{\tilde{s}_4} &= (2.0106, 1.0007) \end{aligned}$$

になり，図 12 左にのせた．図 9 の実際の信号源の位置と信号源の番号は異なるが，小数点 2 桁目がわずかに違っている程度の精度で位置の推定が可能である．

推定した信号源の位置 $P_{\tilde{s}_k}$ と第 1 観測地点 $P_{x_1} = (0, 0)$ の距離を求めて，その距離を音速 V で割って，遅延行列の 1 行目の成分 $c_{1,k}$ の推定値 $d_{1,k}$ を計算する．ただし， $d_{1,k}$ はサンプリング間隔 $1/f_0$ [sec] の整数倍になるように丸める．推定値 $d_{1,k}$ をもとに，遅延行列を修正した行列が $D = (d_{j,k})$ ， $d_{j,k} = d_{1,k} + \delta_{j,k}$ であり，数理モデルの最初の遅延行列 C とならべて書き出す．

$$D = \frac{1}{f_0} \begin{pmatrix} 27 & 55 & 60 & 60 \\ 134 & 72 & 60 & 144 \\ 67 & 42 & 90 & 132 \\ 60 & 123 & 144 & 110 \\ 115 & 149 & 121 & 42 \end{pmatrix}, \quad C = \frac{1}{f_0} \begin{pmatrix} 27 & 60 & 60 & 55 \\ 134 & 144 & 60 & 72 \\ 67 & 132 & 90 & 42 \\ 60 & 110 & 144 & 123 \\ 115 & 42 & 121 & 149 \end{pmatrix}.$$

遅延行列の推定は，列の置換を除いて，完全に一致することがわかる．

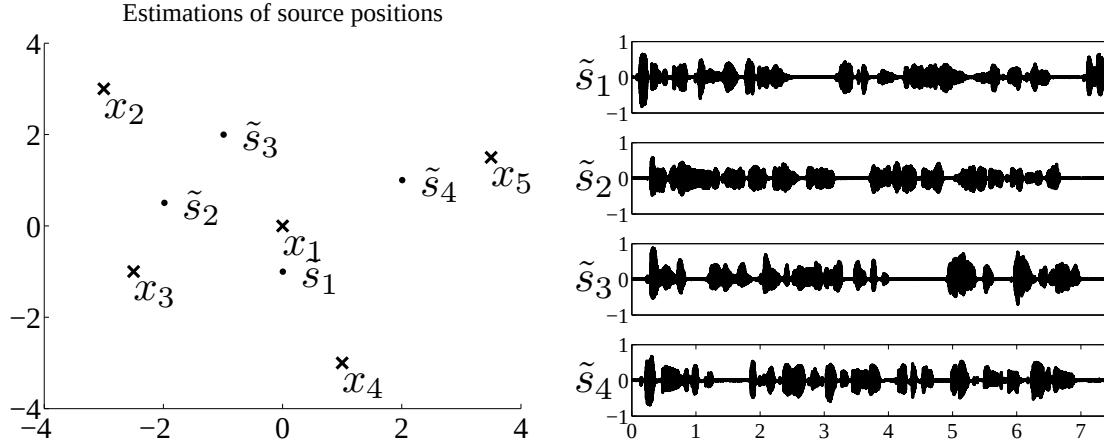


図 12: 実験 2: 推定した信号源の位置 $P_{\tilde{s}_k}$ と分離した信号 $\tilde{s}$.

7.4 信号の分離

正規化した混合行列 $\tilde{A}$ と正規化した遅延行列 $\tilde{C}$ の推定行列 $B = (b_{j,k})$ と $\Delta = (\delta_{j,k})$ が求まったので, 信号源と観測信号の間の数理モデル式 (7.1) は,

$$x_j(t) = \sum_{k=1}^N b_{j,k} \sigma_k(t - \delta_{j,k}), \quad j = 1, 2, \dots, M \quad (7.5)$$

と推定できる. この連立方程式 (7.5) から $\sigma_k(t)$ を求めればよい. ただし, 元の信号 $s_k(t)$ と比較できるのは, 分離した信号 $\tilde{s}_k = \sigma_k(t + d_{1,k})$ である.

式 (7.5) の両辺のフーリエ変換を取ると,

$$\hat{x}_j(\xi) = \sum_{k=1}^N b_{j,k} e^{-i\delta_{j,k}\xi} \hat{\sigma}_k(\xi), \quad j = 1, 2, \dots, M \quad (7.6)$$

を得る. これを行列を使って表現する. $M \times N$ 行列 $B_1(\xi)$ を $(B_1(\xi))_{j,k} = b_{j,k} e^{-i\delta_{j,k}\xi}$ と置く. 観測信号 $x(t)$ のフーリエ変換を列べた M 次縦ベクトル $\hat{x}(\xi)$ と $\hat{\sigma}_k(\xi)$ を列べた N 次縦ベクトル $\hat{\sigma}(\xi)$ を

$$\hat{x}(\xi) = [\hat{x}_1(\xi), \hat{x}_2(\xi), \dots, \hat{x}_M(\xi)]^T, \quad \hat{\sigma}(\xi) = [\hat{\sigma}_1(\xi), \hat{\sigma}_2(\xi), \dots, \hat{\sigma}_N(\xi)]^T$$

とすると, 式 (7.6) は,

$$\hat{x}(\xi) = B_1(\xi) \hat{\sigma}(\xi)$$

と書ける. $M \geq N$ だったので, $B_1(\xi)$ の一般化逆行列をかけると, $\hat{\sigma}(\xi)$ が求まる. 逆フーリエ変換を行えば, $\sigma_k(t)$ が求まる. 分離した信号は, $\tilde{s}_k = \sigma_k(t + d_{1,k})$ である. 図 12 右の分離した信号の誤差評価は, 表 9 にある. 実験 2 の実行時間は, 2 時間程度である.

ESS	SS	Max error	L^1 error	L^2 error	SNR
$\tilde{s}_1$	s_1	2.181 %	0.679 %	0.540 %	45.344 dB
$\tilde{s}_2$	s_4	1.079 %	0.845 %	0.709 %	42.984 dB
$\tilde{s}_3$	s_3	0.453 %	1.016 %	0.721 %	42.838 dB
$\tilde{s}_4$	s_2	0.572 %	0.519 %	0.394 %	48.097 dB

表 9: 実験 2：誤差評価. ESS は分離した信号, SS は元の信号を表す.

8 位相情報の応用

第 7 節で信号源の離散音声信号から観測信号を作る際に、時間遅れはサンプリング間隔 $1/f_0$ [sec] の整数倍であるとの仮定を置いた. ただし, $f_0 = 8820$ [Hz] はサンプリングレートである. この仮定は実際に観測地点にマイクをおいて録音したときなどには満たされない条件である. さらに, 図 11 の 3D ヒストグラム $H_j(\delta, x)$ が時間遅れ δ 方向に薄いのが気にかかる. δ の最小刻み幅はサンプリング間隔 $1/f_0$ [sec] であって, 3D ヒストグラム $H_j(\delta, x)$ は, ピークから δ が $1/f_0$ ずれば, ほぼ 0 になってしまう (一方, 実数 x 方向にはある程度の厚みがある).

そこで本節では, サンプリングレート $f_1 = 44100$ [Hz] の音声信号を使い, 前節と同じ図 9 の地点に信号源と観測地点を配置した数理モデルを用いて, f_1 レートの観測信号を作成する. その後, この観測信号を $1/5$ にダウンサンプリングして, サンプリングレート $f_0 = 8820$ [Hz] の新しい観測信号を作り, サンプリングレート f_0 の観測信号から信号源分離を試みる (実験 3 とよぶ).

8.1 ダウンサンプリングした観測信号の作成

信号源と観測地点は, 第 7 節の図 9 と同じ位置に置くとする. 信号源と観測信号の間の数理モデルも同一であるとする. j 番目の観測地点と k 番目の信号源の間の距離を $r_{j,k}$ として, 混合行列 A の (j, k) 成分は, 距離 $r_{j,k}$ に反比例して $a_{j,k} = 1/r_{j,k}$ だったので, 混合行列は同じになる. 遅延行列は, 音速を V として, $r_{j,k}/V$ をサンプリング間隔 $1/f_0$ [sec] の整数倍で近似していた.

したがって, サンプリングレートを $f_0 = 8820$ [Hz] から $f_1 = 44100$ [Hz] に変更すると, 遅延行列 C と各行から第 1 行を引いて正規化した遅延行列 $\tilde{C}$ は,

$$C = \frac{1}{f_1} \begin{pmatrix} 134 & 299 & 299 & 275 \\ 668 & 720 & 299 & 360 \\ 334 & 658 & 448 & 211 \\ 299 & 551 & 720 & 616 \\ 575 & 211 & 605 & 747 \end{pmatrix}, \quad \tilde{C} = \frac{1}{f_1} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 534 & 421 & 0 & 85 \\ 200 & 359 & 149 & -64 \\ 165 & 252 & 421 & 341 \\ 441 & -88 & 306 & 472 \end{pmatrix}$$

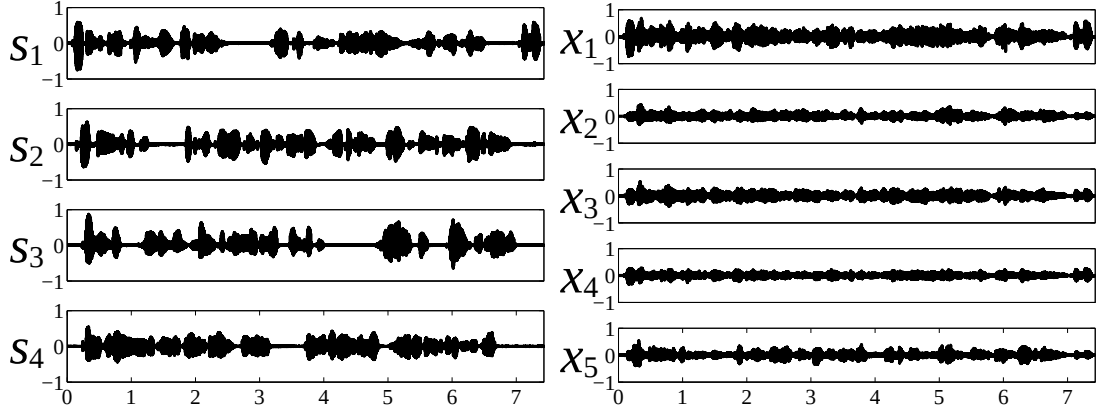


図 13: 実験 3: サンプルングレート f_1 の元の信号 s とレート f_0 の観測信号 x .

になる．この状況で，サンプルングレート f_1 の元の離散信号 $s_k[n]$, $k = 1, \dots, N$ からサンプルングレート f_1 の離散観測信号 $\tilde{x}_j[n]$, $j = 1, \dots, M$ を次で作成する．

$$\tilde{x}_j[n] = \sum_{k=1}^N a_{j,k} s_k[n - f_1 \times c_{j,k}], \quad j = 1, \dots, M.$$

ただし，信号源の数 $N = 4$ ，観測信号の数 $M = 5$ であり， $f_1 \times c_{j,k}$ は遅延行列 C の (j, k) 成分のサンプルングレート f_1 倍で，整数値になる．

こうして作った，サンプルングレート f_1 の観測信号 $\tilde{x}_j[n]$, $j = 1, \dots, M$ にローパスフィルターをかけてから， $1/5$ にダウンサンプルングすると，サンプルングレート $f_0 = 8820$ [Hz] の観測信号 $x_j[n]$, $j = 1, \dots, M$ が得られる．図 13 にサンプルングレート f_1 の元の信号 $s(t)$ とサンプルングレート f_0 の観測信号 $x(t)$ のグラフをのせる．聞き比べてみても，前節の観測信号との違いは分からなかった．

この観測信号から，信号源分離を行ってみよう．時間遅れ変数 δ は， $1/f_0$ [sec] 刻みでしか動かさないので，遅延行列 C を $1/f_0$ 倍で書き直すと，各成分を 5 で割れば良いから，

$$C = \frac{1}{f_0} \begin{pmatrix} 26.8 & 59.8 & 59.8 & 55.0 \\ 133.6 & 144.0 & 59.8 & 72.0 \\ 66.8 & 131.6 & 89.6 & 42.2 \\ 59.8 & 110.2 & 144.0 & 123.2 \\ 115.0 & 42.2 & 121.0 & 149.4 \end{pmatrix}, \quad \tilde{C} = \frac{1}{f_0} \begin{pmatrix} 0.0 & 0.0 & 0.0 & 0.0 \\ 106.8 & 84.2 & 0.0 & 17.0 \\ 40.0 & 71.8 & 29.8 & -12.8 \\ 33.0 & 50.4 & 84.2 & 68.2 \\ 88.2 & -17.6 & 61.2 & 94.4 \end{pmatrix}$$

になる．次小節でアルゴリズム 7.2 にしたがって，商の 3D ヒストグラム $H_j(\delta, x)$ を作成して，時間遅れ $\tilde{C}$ が 0.2, 0.4, 0.6, 0.8 などの小数部分を含むときに，3D ヒストグラムがどうなるかを調べてみよう．

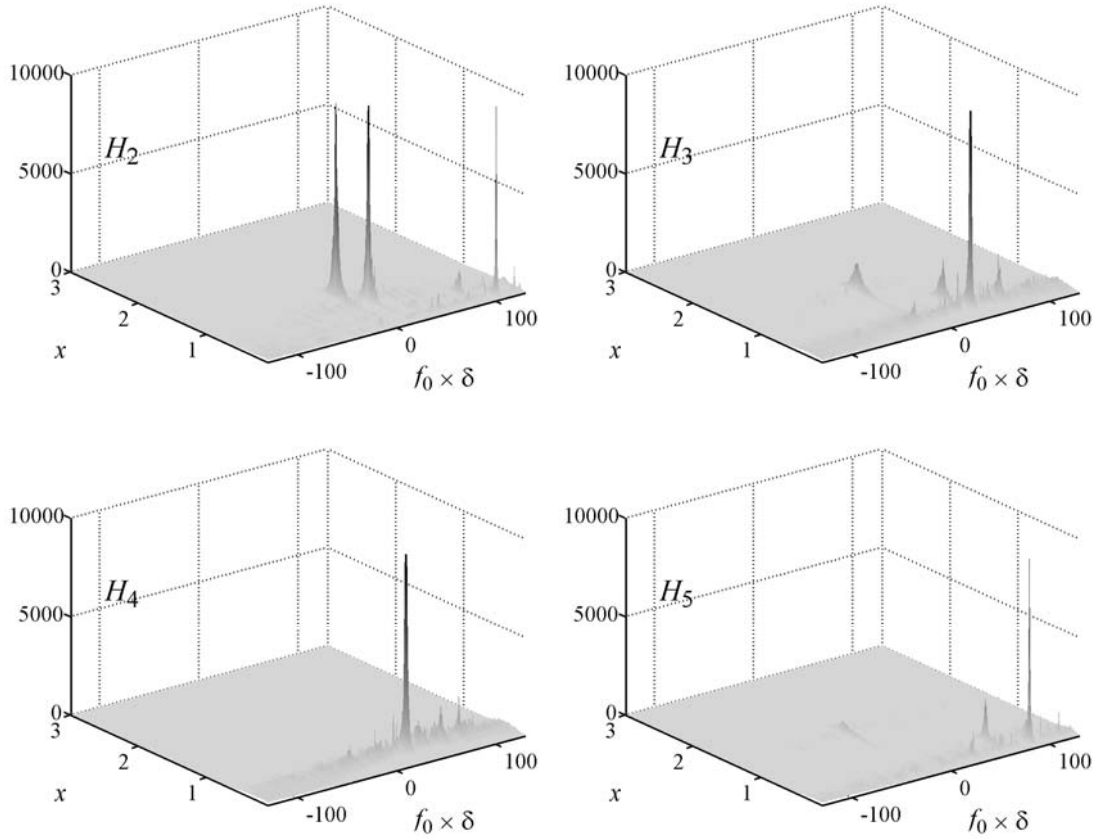


図 14: 実験 3 : 3 D ヒストグラム $H_j(\delta, x)$. δ がサンプリング間隔 $1/f_0$ [sec] の整数倍にならない信号源に対応するピークはほとんど判別できない.

8.2 3 D ヒストグラムの限界

アルゴリズム 7.2 を使って、観測信号の時間周波数情報の商 Q_j の 3 D ヒストグラム $H_j(\delta, x)$, $j = 2, \dots, M$ を作成して図示すると、図 14 を得る. 各 3 D ヒストグラム $H_j(\delta, x)$ のピークを高い方から列記しよう. $H_2(\delta, x)$ から判別できるピークは $\delta_{2,k} = 17/f_0, 0/f_0, 107/f_0$ の 3 個である. $H_3(\delta, x)$ から判別できるピークは $\tilde{\delta}_{3,\tilde{k}} = 40/f_0$ の 1 個のみである. $H_4(\delta, x)$ から判別できるピークは $\tilde{\delta}_{4,\tilde{k}} = 33/f_0$ の 1 個のみである. $H_5(\delta, x)$ から判別できるピークは $\tilde{\delta}_{5,\tilde{k}} = 88/f_0$ の 1 個のみである. 16 個あるはずのピークの内 6 個のピークしか判別できない.

遅延行列 $\tilde{C}$ がサンプリング間隔 $1/f_0$ [sec] の整数倍になっている 4 個のピークは、全て図 14 から確認できる. しかし整数倍でない 12 個の成分の内、3 D ヒストグラムからピークが確認できるのは、(2, 1) 成分の $106.8/f_0$ と (5, 1) 成分の $88.2/f_0$ のみである. この図から信号源の数 $N = 4$ 個のピークを読み取ることは、不可能である.

3D ヒストグラム $H_j(\delta, x)$ を時間遅れ軸方向と実数 x 方向にそれぞれ射影して周辺分布ヒストグラム $H_j^1(\delta)$ と $H_j^2(x)$ を作る．

$$H_j^1(\delta) = \sum_x H_j(\delta, x) = \int H_j(\delta, x) dx,$$

$$H_j^2(x) = \sum_\delta H_j(\delta, x) = \int H_j(\delta, x) d\delta.$$

周辺分布ヒストグラム $H_j^1(\delta)$ はある程度使える． $H_j^1(\delta)$ では，4 個のピーク $\tilde{\delta}_{j,\tilde{k}}$ が捉えられて，それらは $f_0 * \tilde{C}$ を整数値に丸めた値に一致する．つぎに 3D ヒストグラム $H_j(\delta, x)$ をピークに対応する $\delta = \tilde{\delta}_{j,\tilde{k}}$ でスライスしたヒストグラム $H_j(\tilde{\delta}_{j,\tilde{k}}, x)$ がピークを取る x から，混合係数 $\tilde{b}_{j,\tilde{k}}$ の推定もある程度可能である．しかしながら，異なる j 間のピークの対応が付けにくいのである．

実験 3 の観測信号から信号源の分離は一応可能であるが，途中でかなり無理なことをしている．商が特定の実数値を取る時間周波数位置にのみ着目していると，重要な情報を見逃しているのである．そこで，次小節で時間ずれと位相ずれの関係を調べてみよう．

8.3 時間遅れと位相ずれの関係

第 7.2 小節で，式 (7.3) の時間遅れの入った商 $Q_j(\delta, t, \omega)$ を考えた．遅延行列 C を考慮した上で，信号源 s_k のみが活動している時間周波数領域を $E_{C,k}$ と置いた．このとき，時間周波数位置 $(t, \omega) \in E_{C,k}$ とぴったりの時間遅れ $\delta = c_{j,k} - c_{1,k} = \tilde{c}_{j,k}$ を商 $Q_j(\delta, t, \omega)$ に代入すれば，以下のように変形 (式 (7.4)) できて，実数値 $a_{j,k}/a_{1,k}$ になったのである．つまり，

$$\begin{aligned} Q_j(\delta, t, \omega) &= \frac{X_j(t + \delta, \omega)}{X_1(t, \omega)} = \frac{\sum_{k=1}^n a_{j,k} S_k(t + \delta - c_{j,k}, \omega)}{\sum_{k=1}^n a_{1,k} S_k(t - c_{1,k}, \omega)} \\ &= \frac{a_{j,k} S_k(t - c_{1,k}, \omega)}{a_{1,k} S_k(t - c_{1,k}, \omega)} = \frac{a_{j,k}}{a_{1,k}} \in \mathbb{R}. \end{aligned}$$

しかしながら，この節の実験 3 では δ は $1/f_0$ の整数倍しか動けないので，観測信号をダウンサンプリングした場合には，ぴったりの時間遅れ $\delta = c_{j,k} - c_{1,k} = \tilde{c}_{j,k}$ が選べない．そこで， δf_0 を $\tilde{c}_{j,k} f_0$ に一番近い整数になるように取ると， $\delta = \tilde{c}_{j,k} + \varepsilon$ ， $|\varepsilon| \leq 1/(2f_0)$ と取れる．この δ を商 $Q_j(\delta, t, \omega)$ に代入すると，

$$Q_j(\delta, t, \omega) = \frac{\sum_{k=1}^n a_{j,k} S_k(t + \delta - c_{j,k}, \omega)}{\sum_{k=1}^n a_{1,k} S_k(t - c_{1,k}, \omega)} = \frac{a_{j,k} S_k(t - c_{1,k} + \varepsilon, \omega)}{a_{1,k} S_k(t - c_{1,k}, \omega)} \quad (8.1)$$

になる．そこで，信号 $f(t)$ の時間周波数情報 $F(t, \omega)$ と ε [sec] だけずれた時間周波数情報 $F(t + \varepsilon, \omega)$ の関係を考えよう．信号 $f(t)$ が $\omega > 0$ [Hz] の複素三角関数

$f(t) = e^{i2\pi\omega t}$ の場合には, ε [sec] の時間ずれで, $f(t + \varepsilon) = e^{i2\pi\omega(t+\varepsilon)} = e^{i2\pi\omega\varepsilon} f(t)$ と位相が $2\pi\omega\varepsilon$ [rad] ずれるので, 時間周波数情報の関係は

$$F(t + \varepsilon, \omega) \doteq e^{i2\pi\omega\varepsilon} F(t, \omega), \quad \omega > 0$$

であると思われる. 実験 3 の場合には, 十分小さな時間ずれ ε [sec] ($|\varepsilon| \leq 1/(2f_0)$) に対して, この近似は正当化できることを以下に示しておこう.

定義 2.8 にあるように, $F(t, \omega)$ は信号 $f(t)$ の連続ウェーブレット変換を用いた時間周波数情報なので, 周波数 $\omega > 0$ [Hz] に対して, スケールパラメータ $a = c[\hat{\psi}]/(2\pi\omega) > 0$ を取って,

$$F(t + ac[\psi], \omega) = W_\psi f(t, a) = \frac{\sqrt{a}}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) \overline{\hat{\psi}(a\xi)} e^{it\xi} d\xi \quad (8.2)$$

であった. 一方, ε [sec] ずれた時間周波数情報は,

$$F(t + ac[\psi] + \varepsilon, \omega) = W_\psi f(t + \varepsilon, a) = \frac{\sqrt{a}}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) \overline{\hat{\psi}(a\xi)} e^{i(t+\varepsilon)\xi} d\xi \quad (8.3)$$

である. 第 5 節で述べたように, 連続ウェーブレット変換には, $K = 10$ の Ausher の正規直交ウェーブレット関数の解析信号 $\psi(t)$ を用いたので, そのフーリエ変換 $\hat{\psi}(\xi)$ のサポートは, $\text{supp } \hat{\psi} \subset [A, C]$ である. ただし, $A = 2\pi K^2/(2K+1)$, $C = 2\pi(K+1)^2/(2K+1)$ であった. 表 1 より, 周波数窓の中心は $c[\hat{\psi}] \doteq 5.25 \times 2\pi$ [rad/s] で, 周波数窓の中心はサポートの中心 $D = (A+C)/2$ とほぼ同じになり, サポートの大きさは $C - A = 2\pi$ である. このことから積分 (8.2), (8.3) の ξ を $\eta = a\xi - D$ に置き換えると, 積分範囲は $\eta \in [-\pi, \pi]$ で近似できるので,

$$\begin{aligned} F(t + ac[\psi], \omega) &\doteq \frac{1}{2\pi\sqrt{a}} \int_{-\pi}^{\pi} \hat{f}\left(\frac{\eta + D}{a}\right) \overline{\hat{\psi}(\eta + D)} e^{it\frac{\eta + D}{a}} d\eta, \\ F(t + ac[\psi] + \varepsilon, \omega) &\doteq \frac{1}{2\pi\sqrt{a}} \int_{-\pi}^{\pi} \hat{f}\left(\frac{\eta + D}{a}\right) \overline{\hat{\psi}(\eta + D)} e^{i(t+\varepsilon)\frac{\eta + D}{a}} d\eta \\ &= \frac{e^{i\varepsilon\frac{D}{a}}}{2\pi\sqrt{a}} \int_{-\pi}^{\pi} \hat{f}\left(\frac{\eta + D}{a}\right) \overline{\hat{\psi}(\eta + D)} e^{it\frac{\eta + D}{a}} e^{i\varepsilon\frac{\eta}{a}} d\eta \end{aligned}$$

である. 数値実験 3 のパラメータの範囲では, $e^{i\varepsilon\frac{\eta}{a}}$ の部分が無視できることを示そう. 時間遅れは, せいぜい $|\varepsilon| \leq 1/(2f_0) = 1/(2 \times 8820)$ [sec] で, 積分範囲 $|\eta| \leq \pi$ であった. また周波数 ω は, 1000 [Hz] 以下を見ていた. よって,

$$\left| \frac{\varepsilon\eta}{a} \right| = \left| \varepsilon \eta \frac{2\pi\omega}{c[\hat{\psi}]} \right| < \frac{1}{2 \times 8820} \times \pi \times \frac{1000}{5.25} = 0.03393$$

である. このことから, $e^{i\varepsilon\frac{\eta}{a}}$ の 1 からのずれは,

$$\left| 1 - e^{i\varepsilon\frac{\eta}{a}} \right| < 0.03393$$

になって無視できるほど小さい。したがって

$$F(t + ac[\psi] + \varepsilon, \omega) \doteq e^{i\varepsilon \frac{D}{a}} F(t + ac[\psi], \omega) \doteq e^{i2\pi\omega\varepsilon} F(t + ac[\psi], \omega)$$

である。ここで、 $D \doteq c[\hat{\psi}]$, $a = c[\hat{\psi}]/(2\pi\omega)$ になることを使った。まとめると、

命題 8.1. 実験 3 で使う信号・パラメータに対して、次の近似が成り立つ。

$$F(t + \varepsilon, \omega) \doteq e^{i2\pi\omega\varepsilon} F(t, \omega). \quad (8.4)$$

ただし、 $\varepsilon \leq 1/(2f_0)$ [sec], $\omega \leq 1000$ [Hz], ウェーブレット関数として $K = 10$ の Ausher の正規直交ウェーブレット関数の解析信号を用いる。

式 (8.1) の時間周波数情報の商 Q_j に命題 8.1 を用いれば次の系を得る。

系 8.2. 時間周波数位置 $(t, \omega) \in E_{C,k}$ と、時間遅れ $\delta = \tilde{c}_{j,k} + \varepsilon$, $|\varepsilon| \leq 1/(2f_0)$ に対して、時間周波数情報の商は

$$Q_j(\delta, t, \omega) = \frac{a_{j,k} S_k(t - c_{1,k} + \varepsilon, \omega)}{a_{1,k} S_k(t - c_{1,k}, \omega)} \doteq \frac{a_{j,k}}{a_{1,k}} e^{i2\pi\omega\varepsilon} \quad (8.5)$$

で近似できる。 $(\delta$ は $1/f_0$ の整数倍である.)

8.4 偏角ヒストグラムを用いたモデルパラメータの推定

第 8.2 小節の 3D ヒストグラム $H_j(\delta, x)$ では、商 $Q_j(\delta, t, \omega)$ が実数値になる時間周波数位置 (t, ω) を探していた。これは、系 8.2 の式 (8.5) では $\varepsilon = 0$ に相当する。実際、図 14 の 3D ヒストグラムでは、 $\varepsilon = 0$ に相当する時間遅れが $1/f_0$ の整数倍に取れる場所全てにピークが現れていた。

本小節では、商 $Q_j(\delta, t, \omega)$ が式 (8.5) を満たす、つまり $Q_j(\delta, t, \omega) \doteq (a_{j,k}/a_{1,k}) e^{i2\pi\omega\varepsilon}$ のような、時間遅れ δ と時間周波数位置 (t, ω) を探す。この (8.5) 式を $Q_j(\delta, t, \omega)$ の偏角が周波数 ω [Hz] に比例して変化する部分と読み替えて、次のアルゴリズムにしたがって、偏角ヒストグラム $H_j(\delta, \omega, \theta)$ を作成する。そして、偏角ヒストグラムから大まかな時間遅れ $\bar{\delta}_{j,k}$ と信号源の数 N を推定する。

アルゴリズム 8.3. 大まかな時間遅れ $\bar{\delta}_{j,k}$ と信号源の数 N を推定する。

1. $1/f_0$ [sec] の整数倍になる時間遅れ δ [sec] と周波数 ω [Hz] と偏角 θ [rad] に対して、次を満たす時刻 t の集合の数を数えて**偏角ヒストグラム**

$$H_j(\delta, \omega, \theta) := \#\{t; \arg(Q_j(\delta, t, \omega)) = \theta\}, \quad j = 2, \dots, M$$

を作成する ($\arg z$ は複素数 z の偏角をあらわす)。

2. 各時間遅れ δ を固定して, $\omega - \theta$ 平面に偏角ヒストグラム $H_j(\delta, \omega, \theta)$ をプロットする.
3. δ を動かして, アニメーションを作成し, 偏角ヒストグラムの周波数 ω 方向に連なる山が偏角 $\theta = 0$ [rad] のラインと重なるときの時間遅れ $\bar{\delta}_{j,\tilde{k}}$ を大まかな時間遅れとよぶ. また, 連なる山が偏角 $\theta = 0$ [rad] のラインと重なる回数が信号源の数 N の推定値である.

注意 8.4. 正規化した混合係数 $\tilde{A}$ が負の成分を持つ場合に, 負の成分に対応する山は偏角 $\theta = \pi$ のラインに重なる時間遅れを探す.

アルゴリズム 8.3 を用いて, 商 $Q_5(\delta, t, \omega)$ の偏角ヒストグラム $H_5(\delta, \omega, \theta)$ を時間遅れ $\delta = 85/f_0, 86/f_0, 87/f_0, 88/f_0, 89/f_0, 90/f_0$ の範囲で描くと図 15 を得る. 周波数 ω 方向に連なった山が下から上に動いて, $\delta = 88/f_0$ 辺りで偏角 0 のラインと重なる. よって, 大まかな時間遅れ $\bar{\delta} \doteq 88/f_0$ を持つ信号源があることを示唆している (正確には時間遅れ $\delta = 88.2/f_0$ に対応する). また, $\delta = 90/f_0$ の図には, 原点からのびて 1000 [Hz] 辺りで偏角 $-\pi$ へ連なる右下がりの山の連なりもうっすら見えている. この山の連なりは $\delta = 94.4/f_0$ に対応するものである. またどの偏角ヒストグラム $H_j(\delta, \omega, \theta)$, $j = 2, \dots, M$ のアニメーションを見ても山の連なりが偏角 $\theta = 0$ [rad] のラインと 4 回重なるので, 信号源の数は $N = 4$ であると推定できる. 各 $j = 2, \dots, M$ に対して, 大まかな時間遅れを小さい順にならべて, $\bar{\delta}_{j,\tilde{k}}$, $\tilde{k} = 1, \dots, N$ と番号を打った大まかな遅延行列は,

$$\bar{\Delta} = \frac{1}{f_0} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 17 & 84 & 107 \\ -13 & 30 & 40 & 72 \\ 33 & 50 & 68 & 84 \\ -18 & 61 & 88 & 94 \end{pmatrix} \quad (8.6)$$

である. ただし, 時間遅れを測るときの基準になる $\bar{\Delta}$ の 1 行目の成分は全て 0 とする. 他のアニメーションは, ホームページ [32] にあげておく. 次のアルゴリズムで, 詳細な時間遅れと対応する混合係数を求める.

アルゴリズム 8.5. 以下の手順で大まかな時間遅れ $\bar{\delta}_{j,\tilde{k}}$ を補正して, 詳細な時間遅れ $\tilde{\delta}_{j,\tilde{k}}$ と対応する混合係数 $\tilde{b}_{j,\tilde{k}}$ を求める.

1. 大まかな時間遅れ $\bar{\delta}_{j,\tilde{k}}$ を固定して, 偏角ヒストグラム $H_j(\bar{\delta}_{j,\tilde{k}}, \omega, \theta)$ を描く. 真上から見た偏角ヒストグラムの方が傾きを求めやすい.
2. 周波数 $700 \leq \omega \leq 1000$ [Hz] ごとに, 偏角 $-0.1\pi \leq \theta \leq 0.1\pi$ [rad] 内で, 偏角ヒストグラムが最大値を取る $\theta(\omega)$ を計算する. データ $(\omega, \theta(\omega))$ が原点を通る直線上 ($\theta = 2\pi\epsilon\omega$) にのるように傾き $2\pi\epsilon$ を最小 2 乗法で計算する. 直線近似の誤差が大きいところをはぶいて, 2 回ほど繰り返し計算する.

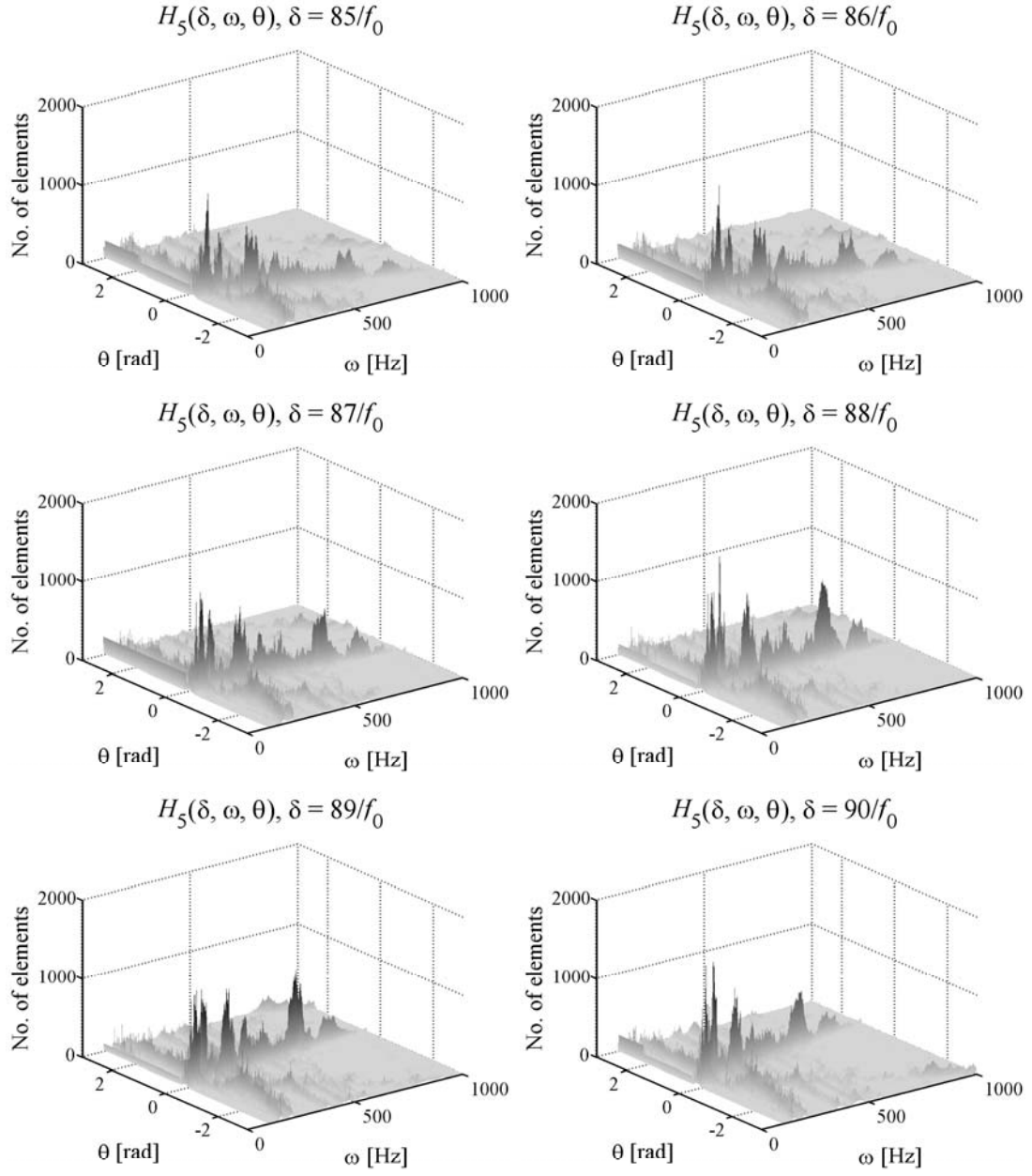


図 15: 実験 3: 商 Q_5 の偏角ヒストグラム $H_5(\delta, \omega, \theta) = \#\{t; \arg(Q_5(\delta, t, \omega)) = \theta\}$ のアニメーション. $\delta = 85/f_0, 86/f_0, 87/f_0, 88/f_0, 89/f_0, 90/f_0$.

3. 系 8.2 では, “ $\delta = \tilde{c}_{j,k} + \varepsilon$ ” の関係があった. このアルゴリズムでの変数の対応は, δ が大まかな時間遅れ $\bar{\delta}_{j,\tilde{k}}$ に相当し, $\tilde{c}_{j,k}$ が推定したい詳細な時間遅れ $\tilde{\delta}_{j,\tilde{k}}$ である. したがって,

$$\tilde{\delta}_{j,\tilde{k}} := \bar{\delta}_{j,\tilde{k}} - \varepsilon.$$

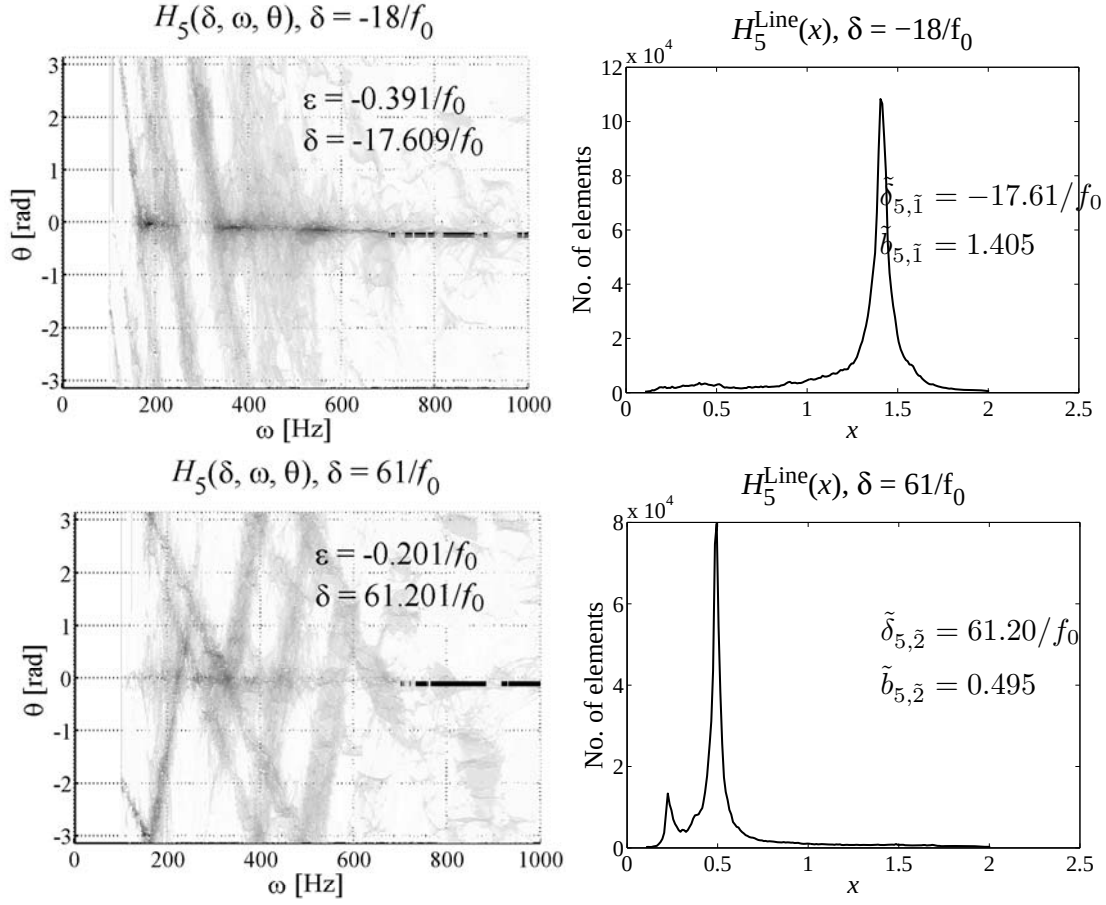


図 16: 実験 3：真上から見た偏角ヒストグラム $H_5(\delta, \omega, \theta)$ と直線上のヒストグラム $H_5^{\text{Line}}(x)$ ($\delta = -18/f_0, 61/f_0$) .

4. 2. で求めた直線上にのっている時間周波数位置の集合を求める.

$$E_{j,\tilde{k}} := \{(t, \omega) ; \arg(Q_j(\bar{\delta}_{j,\tilde{k}}, t, \omega)) = 2\pi\epsilon\omega\}.$$

5. 直線上の $|Q_j(\bar{\delta}_{j,\tilde{k}}, t, \omega)|$ のヒストグラムを描く. つまり, 実数値 x に対して,

$$H_j^{\text{Line}}(x) := \sharp \{(t, \omega) \in E_{j,\tilde{k}} ; |Q_j(\bar{\delta}_{j,\tilde{k}}, t, \omega)| = x\}.$$

6. 直線上のヒストグラム $H_j^{\text{Line}}(x)$ が最大値を取る x が混合係数の推定値 $\tilde{b}_{j,\tilde{k}}$ である.

注意 8.6. $j = 2$ の場合の詳細な時間遅れ $\delta_{2,k}$ と混合行列 $b_{2,k}$ には $\sim$ を付けない.

アルゴリズム 8.3 から, 偏角ヒストグラム $H_5(\delta, \omega, \theta)$ の δ 軸方向のアニメーションから, 信号源の数 $N = 4$ と, 大まかな時間遅れ $\bar{\delta}_{5,\tilde{k}} = -18/f_0, 61/f_0, 88/f_0, 94/f_0$

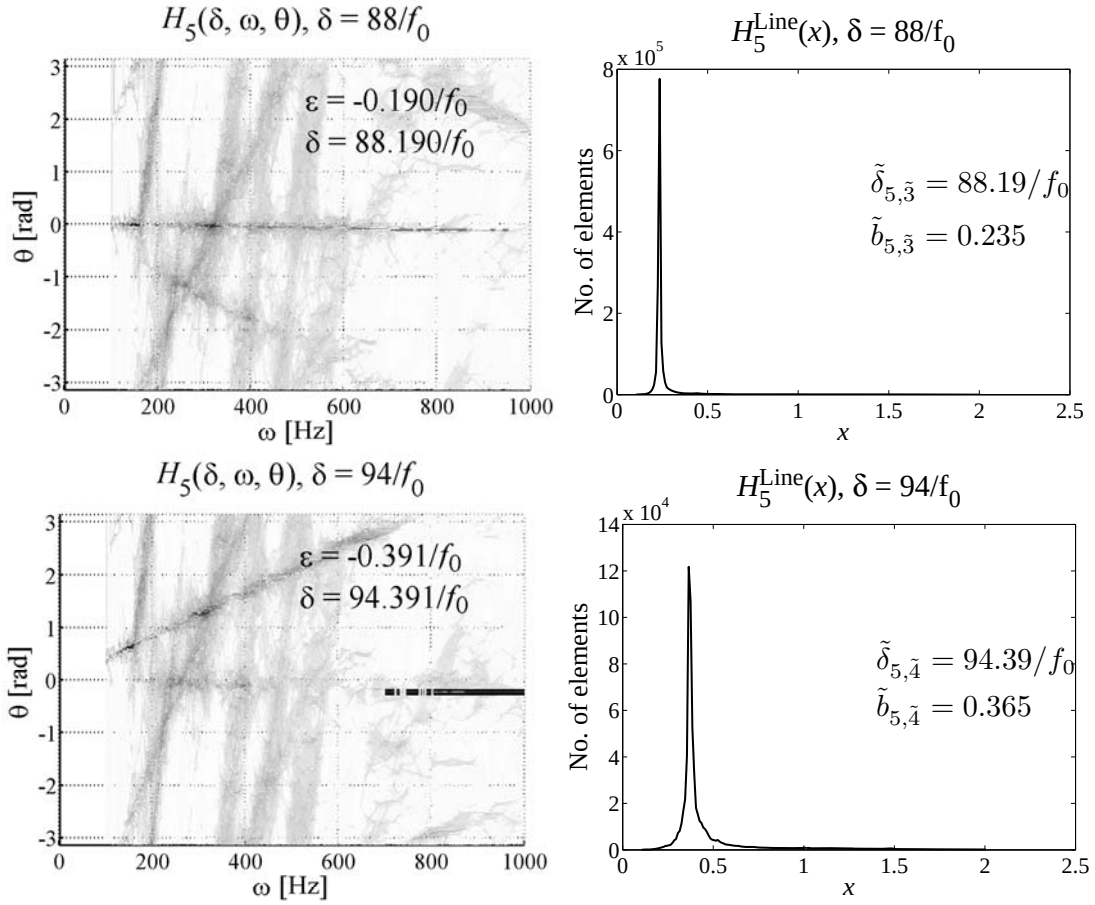


図 17: 実験 3: 真上から見た偏角ヒストグラム $H_5(\delta, \omega, \theta)$ と直線上のヒストグラム $H_5^{\text{Line}}(x)$ ($\delta = 88/f_0, 94/f_0$) .

を推定した. 次にアルゴリズム 8.5 を用いて, 詳細な時間遅れ $\tilde{\delta}_{5,k}$ と対応する混合係数 $\tilde{b}_{5,k}$ を求めてみよう.

アルゴリズム 8.5 の手順 1. にそって, 大まかな時間遅れ $\bar{\delta} = -18/f_0$ に対応する偏角ヒストグラム $H_5(\delta, \omega, \theta)$ を真上から見た図を作ると, 図 16 左上になる. 次に, 手順 2. を使って, 最小 2 乗法で水平に近い直線 ($\theta = 2\pi\epsilon\omega$) の傾き ϵ を求める. 傾きを求めるために使ったデータは, 図 16 左上図の水平に近い直線上で $\omega \geq 700$ Hz で黒点を打った場所である. 直線の傾きは $\epsilon = -0.391/f_0$ になった. 手順 3. より, 詳細な時間遅れは, $\tilde{\delta}_{5,1} = -18/f_0 - \epsilon = -17.609/f_0$ である.

この詳細な時間遅れ $\tilde{\delta}_{5,1}$ に対応する混合係数 $\tilde{b}_{5,1}$ を求めよう. 手順 4. を用いて, 商 $Q_5(-18/f_0, t, \omega)$ がこの直線上に来る時間周波数位置の集合 $E_{5,1}$ を作成し, 手順 5. に沿って, 直線上のヒストグラム $H_5^{\text{Line}}(x)$ を描くと図 16 右上になる. 手順 6. で直線上のヒストグラム $H_5^{\text{Line}}(x)$ のピークに対応する x 座標を読むと, 混合係数の推定値 $\tilde{b}_{5,1} = 1.405$ が得られる.

商 $Q_2 (H_2)$			
Peaks	$\delta_{2,k}$	$b_{2,k}$	$\sharp(t, \omega)$
$k = 1$	$0.0007/f_0$	0.995	132122
$k = 2$	$17.0000/f_0$	0.765	274874
$k = 3$	$84.1968/f_0$	0.415	110168
$k = 4$	$106.7931/f_0$	0.205	556175

商 $Q_3 (H_3)$			
Peaks	$\tilde{\delta}_{3,\tilde{k}}$	$\tilde{b}_{3,\tilde{k}}$	$\sharp(t, \omega)$
$\tilde{k} = 1$	$-12.8061/f_0$	1.305	222566
$\tilde{k} = 2$	$29.7921/f_0$	0.665	105104
$\tilde{k} = 3$	$40.0000/f_0$	0.395	917379
$\tilde{k} = 4$	$71.8253/f_0$	0.455	136971

商 $Q_4 (H_4)$			
Peaks	$\tilde{\delta}_{4,\tilde{k}}$	$\tilde{b}_{4,\tilde{k}}$	$\sharp(t, \omega)$
$\tilde{k} = 1$	$32.9879/f_0$	0.445	1691786
$\tilde{k} = 2$	$50.4004/f_0$	0.545	209335
$\tilde{k} = 3$	$68.1925/f_0$	0.445	249495
$\tilde{k} = 4$	$84.1858/f_0$	0.415	92724

商 $Q_5 (H_5)$			
Peaks	$\tilde{\delta}_{5,\tilde{k}}$	$\tilde{b}_{5,\tilde{k}}$	$\sharp(t, \omega)$
$\tilde{k} = 1$	$-17.6091/f_0$	1.405	108315
$\tilde{k} = 2$	$61.2009/f_0$	0.495	79734
$\tilde{k} = 3$	$88.1905/f_0$	0.235	776227
$\tilde{k} = 4$	$94.3906/f_0$	0.365	121725

表 10: 実験 3：アルゴリズム 8.5 で求めた詳細な時間遅れと対応する混合係数.

$j = 5$ の残り 3 個の大まかな時間遅れ $\bar{\delta}_{5,\tilde{k}} = 61/f_0, 88/f_0, 94/f_0$ に対して、アルゴリズム 8.5 を用いて、同様の解析をしたときのグラフを、図 16 下図と図 17 にあげる. $j = 2, 3, 4$ の大まかな時間遅れに対しても、アルゴリズム 8.5 を用いて、詳細な時間遅れと対応する混合係数を求めると、表 10 を得る.

次に、こうして求めた詳細な時間遅れと混合係数から $M \times N$ の遅延行列と混合行列を推定しなければならない. そこで、各 $j = 2, \dots, M$ 行の詳細な時間遅れがどの列に入るかを次のアルゴリズム 8.7 を用いて決定する.

アルゴリズム 8.7. 以下の手順に沿って、2 行目の詳細な時間遅れ $\delta_{2,k}$ と混合係数 $b_{2,k}$ を j 行目の $\tilde{\delta}_{j,\tilde{k}}$ と $\tilde{b}_{j,\tilde{k}}$ へ対応付ける.

1. アルゴリズム 8.5 の手順 4. で求めた偏角ヒストグラムが直線上になる時間周波数位置の集合 $E_{2,k}$ に属し、さらに $|Q_2(\bar{\delta}_{2,k}, t, \omega)|$ が対応する混合係数と等しい時間周波数位置 (t, ω) を 1000 点選び、集合

$$Y_k := \{(t, \omega) ; \arg(Q_2(\bar{\delta}_{2,k}, t, \omega)) = 2\pi\varepsilon\omega, |Q_2(\bar{\delta}_{2,k}, t, \omega)| = b_{2,k}, 1000 \text{ 点}\}$$

を作る. これは、直線上のヒストグラム $H_2^{\text{Line}}(x)$ がピーク $b_{2,k}$ を取る場所に対応する時間周波数位置である.

2. 各 $j \geq 3$ に対して、商の絶対値 $|Q_j(\bar{\delta}_{j,\tilde{k}}, t, \omega)|$ に $(t, \omega) \in Y_k$ を代入して次の集合の個数を数える.

$$\sharp \left\{ (t, \omega) \in Y_k ; |Q_j(\bar{\delta}_{j,\tilde{k}}, t, \omega)| = \tilde{b}_{j,\tilde{k}} \right\}.$$

3. 個数の一番多い集合に対応する詳細な時間遅れ $\tilde{\delta}_{j,\tilde{k}}$ が $\delta_{j,k}$ に、混合係数 $\tilde{b}_{j,\tilde{k}}$ が $b_{j,k}$ になるように番号 k を付けなおす.

$H_2 \setminus H_3$	$\tilde{b}_{3,\bar{1}}$	$\tilde{b}_{3,\bar{2}}$	$\tilde{b}_{3,\bar{3}}$	$\tilde{b}_{3,\bar{4}}$
$b_{2,1}$	0	758	0	0
$b_{2,2}$	898	0	0	0
$b_{2,3}$	0	0	0	816
$b_{2,4}$	0	0	1000	0

$H_2 \setminus H_4$	$\tilde{b}_{4,\bar{1}}$	$\tilde{b}_{4,\bar{2}}$	$\tilde{b}_{4,\bar{3}}$	$\tilde{b}_{4,\bar{4}}$
$b_{2,1}$	0	0	0	949
$b_{2,2}$	0	0	963	0
$b_{2,3}$	0	944	0	0
$b_{2,4}$	1000	0	0	0

$H_2 \setminus H_5$	$\tilde{b}_{5,\bar{1}}$	$\tilde{b}_{5,\bar{2}}$	$\tilde{b}_{5,\bar{3}}$	$\tilde{b}_{5,\bar{4}}$
$b_{2,1}$	0	942	0	0
$b_{2,2}$	0	0	0	328
$b_{2,3}$	790	0	0	0
$b_{2,4}$	0	0	1000	0

表 11: 実験 3: アルゴリズム 8.7 による $b_{2,k}$ と $\tilde{b}_{j,\bar{k}}$, $j = 3, \dots, M$ の対応付け.

4. 遅延行列 $\Delta = (\delta_{j,k})$ と混合行列 $B = (b_{j,k})$ を作成する. ただし, Δ の第 1 行は全て 0, B の第 1 行は全て 1 と置く.

このアルゴリズム 8.7 を用いて, $j = 2$ の詳細な時間遅れ $\delta_{2,k}$ と混合係数 $b_{2,k}$ を $j \geq 3$ 行目の $\tilde{\delta}_{j,\bar{k}}$ と $\tilde{b}_{j,\bar{k}}$ へ対応付けると, 表 11 のような結果を得る. 表の中の数値は, 手順 2. で数えた集合の要素数である. たとえば, 混合係数 $b_{2,1}$ には, $\tilde{b}_{3,\bar{2}}$, $\tilde{b}_{4,\bar{4}}$, $\tilde{b}_{5,\bar{2}}$ が対応していて, これらが混合行列 B の第 1 列になる. この対応付けをもとに, 混合行列 B を推定すると,

$$B = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.9950 & 0.7650 & 0.4150 & 0.2050 \\ 0.6650 & 1.3050 & 0.4550 & 0.3950 \\ 0.4150 & 0.4450 & 0.5450 & 0.4450 \\ 0.4950 & 0.3650 & 1.4050 & 0.2350 \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} 1.0000 & 1.0000 & 1.0000 & 1.0000 \\ 0.2000 & 0.4152 & 1.0000 & 0.7656 \\ 0.4000 & 0.4541 & 0.6667 & 1.3038 \\ 0.4472 & 0.5423 & 0.4152 & 0.4472 \\ 0.2325 & 1.4142 & 0.4939 & 0.3688 \end{pmatrix}$$

である. 右にならべて書いた $\tilde{A}$ は正規化したモデルパラメータである. 推定した混合行列 B の第 1, 2, 3, 4 列は, それぞれモデルパラメータ $\tilde{A}$ の第 3, 4, 2, 1 列に対応している. また小数点第 2 位程度の精度で正確に混合行列は推定されている. 同様に推定した遅延行列 Δ は,

$$\Delta = \frac{1}{f_0} \begin{pmatrix} 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 17.00 & 84.20 & 106.79 \\ 29.79 & -12.81 & 71.83 & 40.00 \\ 84.19 & 68.19 & 50.40 & 32.99 \\ 61.20 & 94.39 & -17.61 & 88.19 \end{pmatrix}, \quad \tilde{C} = \frac{1}{f_0} \begin{pmatrix} 0.0 & 0.0 & 0.0 & 0.0 \\ 106.8 & 84.2 & 0.0 & 17.0 \\ 40.0 & 71.8 & 29.8 & -12.8 \\ 33.0 & 50.4 & 84.2 & 68.2 \\ 88.2 & -17.6 & 61.2 & 94.4 \end{pmatrix}$$

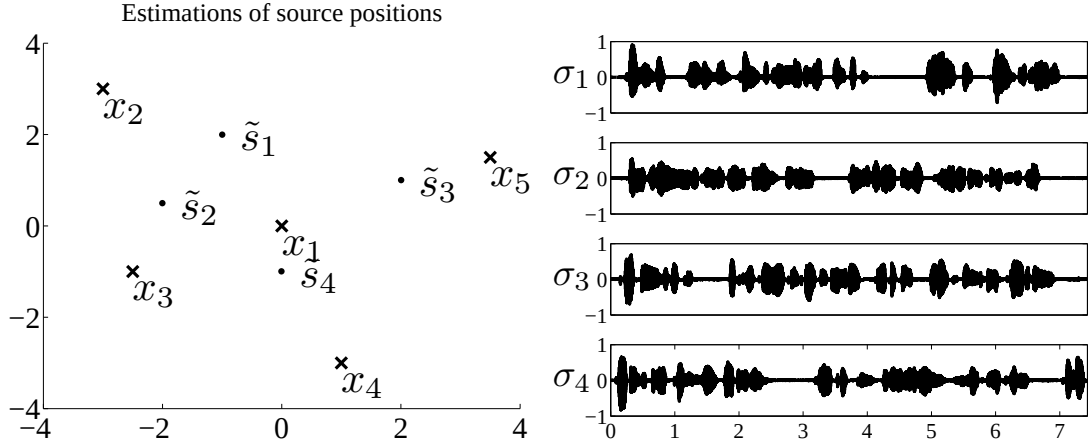


図 18: 実験 3：信号源の推定位置と分離した信号 $\sigma(t)$.

になる．右にならべて書いた $\tilde{C}$ は正規化したモデルパラメータである．推定した遅延行列 Δ の第 1, 2, 3, 4 列は，それぞれモデルパラメータ $\tilde{C}$ の第 3, 4, 2, 1 列に対応している．また小数点第 1 位までは正確に推定されている．

8.5 信号源の位置の推定

第 7.3 小節と同じく，双曲線の交点を求める Mellin-Pachter-Raquet [23] の方法を用いて信号源の位置の推定を行うと，

$$\begin{aligned} P_{\tilde{s}_1} &= (-1.0002, 2.0008), & P_{\tilde{s}_2} &= (-2.0020, 0.4987), \\ P_{\tilde{s}_3} &= (2.0014, 0.9989), & P_{\tilde{s}_4} &= (-0.0019, -0.9951) \end{aligned}$$

になり，図 18 左にのせた．図 9 の実際の信号源の位置と信号源の番号は異なるが，小数点 3 桁目がわずかに違っている程度の精度で位置の推定が可能である．

信号源の推定位置 $P_{\tilde{s}_k}$ が求まったので，第 1 観測地点 $P_{x_1} = (0, 0)$ との距離が求まる．この距離を音速 V で割って，遅延行列の 1 行目の成分 $c_{1,k}$ の推定値 $d_{1,k}$ が計算できる．そして遅延行列を $D = (d_{j,k})$, $d_{j,k} = d_{1,k} + \delta_{j,k}$ で補正すると，

$$D = \frac{1}{f_0} \begin{pmatrix} 59.7862 & 55.1422 & 59.7843 & 26.5976 \\ 59.7869 & 72.1422 & 143.9811 & 133.3907 \\ 89.5783 & 42.3361 & 131.6096 & 66.5976 \\ 143.9720 & 123.3347 & 110.1847 & 59.5855 \\ 120.9871 & 149.5328 & 42.1752 & 114.7880 \end{pmatrix}$$

である．最初の数理モデルの遅延行列 C は，

$$C = \frac{1}{f_0} \begin{pmatrix} 26.8 & 59.8 & 59.8 & 55.0 \\ 133.6 & 144.0 & 59.8 & 72.0 \\ 66.8 & 131.6 & 89.6 & 42.2 \\ 59.8 & 110.2 & 144.0 & 123.2 \\ 115.0 & 42.2 & 121.0 & 149.4 \end{pmatrix}$$

であった．推定した遅延行列 D の第 1, 2, 3, 4 列は，それぞれ C の第 3, 4, 2, 1 列に対応している． $d_{1,4} = 26.5976/f_0$ と対応する $c_{1,1} = 26.8/f_0$ の誤差が大きいのは， $P_{x_0} = (0, 0)$ と $P_{s_1} = (0, -1)$ の間の時間遅れは $r_{1,1}/V = 133.64/f_1$ [sec] なので， $1/f_1$ の整数倍に丸めるときに $134/f_1 = 26.8/f_0$ [sec] とした．一方，推定信号源の位置 $P_{s_4} = (-0.0019, -0.9951)$ と原点の距離は 0.9951 になって，時間遅れは $26.5976/f_0$ [sec] である．したがって， $1/f_1$ の整数倍に丸めたときの誤差が現れてくるのである．

8.6 信号の分離

実験 3 も，信号源から観測信号を作る数理モデルは実験 2 と同じ (7.5) 式で表される一番簡単な時空間的混合問題である．つまり，

$$x_j(t) = \sum_{k=1}^N b_{j,k} \sigma_k(t - \delta_{j,k}), \quad j = 1, 2, \dots, M.$$

したがって，両辺のフーリエ変換を取ると，式 (7.6) になる．つまり，

$$\hat{x}_j(\xi) = \sum_{k=1}^N b_{j,k} e^{-i\delta_{j,k}\xi} \hat{\sigma}_k(\xi), \quad j = 1, 2, \dots, M.$$

第 7.4 小節では，時間遅れ $\delta_{j,k}$ がサンプリング間隔 $1/f_0$ [sec] の整数倍であったが，この連立方程式を解くときには $\delta_{j,k}$ は実数値だったら何でも良いので，本節の実験 3 でも問題なく解くことができる．この連立方程式を解いて得られた分離した信号 $\sigma(t)$ を図 18 にのせる．実験 1 および 2 で用いたサンプリングレート $f_0 = 8820$ [Hz] の信号源と聞き比べても違いは分からなかった．

分離した信号の評価を行うためには，同じサンプリングレート f_0 [Hz] になるように $1/5$ にダウンサンプリングした元の信号を作成しなければならない．そこで， k 番目の信号源の音声のみを出力して，1 番目の観測地点で観測した音声を $1/5$ にダウンサンプリングして比較可能な元の信号 $\tilde{\sigma}_k(t)$, $k = 1, \dots, N$ を作成する．そして，分離した信号 $\sigma(t)$ と比較可能な元の信号 $\tilde{\sigma}(t)$ で誤差を評価すると，表 12 を得る．実験 3 の計算時間は，6 時間くらい必要である．

ESS	SS	Max error	L^1 error	L^2 error	SNR
σ_1	$\tilde{\sigma}_3$	1.018 %	1.694 %	1.196 %	38.443 dB
σ_2	$\tilde{\sigma}_4$	1.688 %	1.400 %	1.196 %	38.450 dB
σ_3	$\tilde{\sigma}_2$	1.132 %	1.059 %	0.824 %	41.685 dB
σ_4	$\tilde{\sigma}_1$	0.570 %	1.025 %	0.832 %	41.601 dB

表 12: 実験 3：誤差評価. ESS は分離した信号, SS は比較可能な元の信号.

8.7 計算の高速化について

位相角ヒストグラムを使った実験 3 の解法では, 計算時間が 6 時間ほどかかっているが, 以下の方法で計算時間を 1/3 程度に減らせることができる.

位相角ヒストグラムは, 偏角 θ [rad] 方向は 2π 周期なので, 図 16 左と図 17 左では上辺と下辺を同一視する. 周波数 ω [Hz] 方向で右下がりの山の連なりがあれば, 見ている δ より大きい時間遅れで水平になる. 右上がりの場合は, 見ている δ より小さい時間遅れで水平になる. たとえば, 図 16 左上の $\delta = -18/f_0$ では右下がりの山の連なりしか見えないので, これより小さい δ を見る必要はない. 傾きの急な山の連なりしかないので, 次の δ をだいたい大きくとっても良いことも分かる. 図 16 左下の $\delta = 61/f_0$ では, 急な右上がりが 1 本とやや急な右下がりが 2 本が見えているので, $\delta = 61/f_0$ より小さい時間遅れと大きい時間遅れ (2 個以上) をたどれる.

さらに, 図 17 左下の時間遅れ $\delta = 94/f_0$ の位相角ヒストグラムでは, 詳細な時間遅れ $\tilde{\delta} = 88.2/f_0$ に相当する右上がりの山の連なりが捉えられている. この右上がりの山に対して, アルゴリズム 8.5 を用いると, 図 19 が得られ, 詳細な時間遅れ $\tilde{\delta}_{5,3} = 88.224/f_0$ と対応する混合係数 $\tilde{b}_{5,3} = 0.235$ が推定できる. この推定は, 大まかな時間遅れ $\bar{\delta} = 88/f_0$ の図 17 で推定した詳細な時間遅れ $\tilde{\delta}_{5,3} = 88.190/f_0$ と混合係数 $\tilde{b}_{5,3} = 0.235$ と同程度の精度を持つ.

これらのことから, 位相角ヒストグラムを作る必要のある大まかな時間遅れを減らすことができ, 高速化が可能になる.

9 まとめと今後の課題

本講究録ではブラインド信号源分離の空間的混合問題と一番簡単な時空間的混合問題を時間周波数解析を使って解く方法を提案した. 提案した手法は, 時間周波数情報の商から一つの信号源のみ活動している時間周波数領域を迫りかけて, 数理モデルのパラメータを同定する方法である.

空間的混合問題と一番簡単な時空間的混合問題に対しては, ここで提案した方法で, 計算時間はかかるが対処可能である. 時間については, 7.4 [sec] の観測信号の分離に対して, 空間的混合問題の実験 1 で 280 [sec] 程, 一番簡単な空間的混

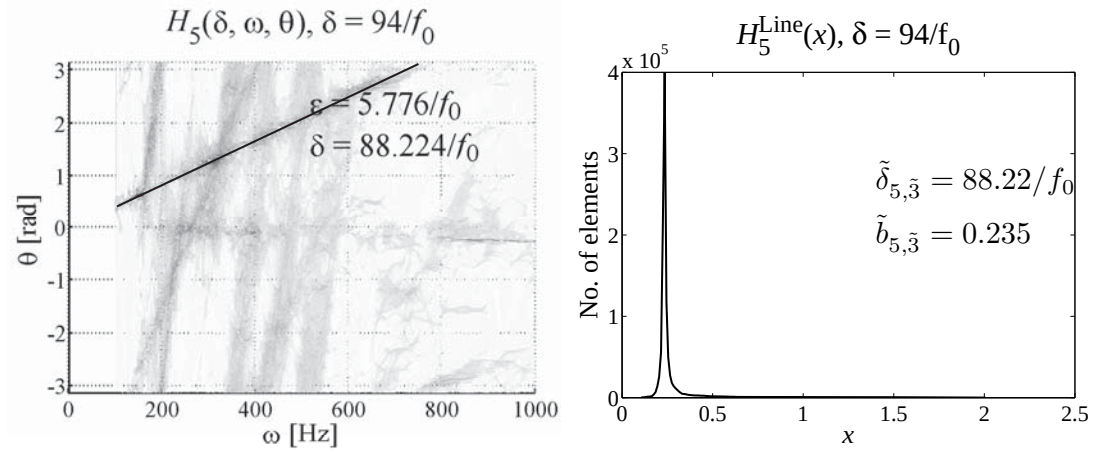


図 19: 実験 3: 偏角ヒストグラム $H_5(94/f_0, \omega, \theta)$ で黒線の右上がりの山の連なりを追跡 (左), 右は黒線上のヒストグラム $H_5^{\text{Line}}(x)$ ($\delta = 94/f_0$) .

合問題の実験 2 で 2 時間程度, 位相角ヒストグラムを使った実験 3 の場合で 6 時間ほどかかっている.

今後の目標は, 実際に観測地点にマイクを設置して録音した観測信号から信号源分離を行うことである. しかしながら, 実際の録音をもとに図 15 の位相角ヒストグラムのアニメーションを作ると, 大まかな時間遅れ $\bar{\delta}$ を動かしたとき, 雲状のかたまりが下から上に移動するのが見える程度で, これを直線で近似するのは不可能であろう.

実際のマイクの録音では, 信号源と観測信号の数理モデルはマイクのインパルス・レスポンスを考慮して, 時空間的混合問題になる. これは, 同じ特性の無指向性マイクを用いて, 信号源の替わりに, 一つずつの信号源をマイクで受けた信号を新しい信号源だと思えば, 我々の方法で分離できるかもしれない.

おそらく, マイクの周波数レスポンスも考慮に入れられないと思われる. この場合, 数理モデルを時間周波数情報の関係になおした際に, 注目する周波数 ω [Hz] による影響が入るはずである.

謝辞

この講究録は, 以下の方々との共同研究を元に作成したので感謝します. アルファベット順に, 大阪教育大学・数理科学の芦野隆一さん, University of Maryland, Mathematics and the Institute for Systems Research の Carlos A. Berenstein さん, 佐賀大学・文化教育学部 of 藤田景子さん, 大阪電気通信大学・工学部の萬代武史さん, 国際基督教大学の森本光生さん, 大阪教育大学の西原清顕さん, George Mason University, Mathematics の Domenico Napoletani さん, 東京理科大学・建

築学科の佐々木文夫さん，京都大学・数理解析研究所の竹井義次さんである．特に，芦野さんと萬代さんには，本講究録を読んでもらって意見を伺ったので感謝します．研究集会で話す機会とこんなに長い講究録を書く機会を与えてくれた京都大学・数理解析研究所の山田道夫さんに感謝します．

参考文献

- [1] ARCHITECTURAL INSTITUTE OF JAPAN, *Sound material in living environment*, Gihodoshuppan, Chiyodaku, Tokyo, Japan, 2004.
- [2] 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*, Appli. Anal., **86** (5), 577–609, 2007.
- [3] R. ASHINO, T. MANDAI, A. MORIMOTO, AND F. SASAKI, *Blind source separation using time-frequency analysis*, Fields Institute Communications, **52**, 401–414, 2007.
- [4] R. ASHINO, T. MANDAI, A. MORIMOTO, AND F. SASAKI, *Blind source separation of spatio-temporal mixed signals using time-frequency analysis*, submitted.
- [5] R. ASHINO, K. FUJITA, T. MANDAI, A. MORIMOTO, AND K. NISHIHARA, *Blind source separation using time-frequency information matrix given by several wavelet transforms*, Information, **10** (5), 555–568, 2007.
- [6] P. AUSCHER, *Wavelet bases for $L^2(\mathbb{R})$ with rational dilation factor*, in Wavelets and their applications, 439–451, Jones and Bartlett, Boston, MA, 1992.
- [7] R. BALAN AND J. ROSCA, *Statistical properties of STFT ratios for two channel systems and applications to blind source separation*, in Proceedings of the 2nd International Workshop on Independent Component Analysis and Blind Signal Separation, 19–22, Helsinki, Finland, June 2000.
- [8] C. CHERRY, *Some experiments on the recognition of speech, with one and with two ears*, Journal of Acoustical Society of America, **25** (5), 975–979, 1953.
- [9] C. CHERRY, *On human communication*, Third ed., The MIT Press, Cambridge, MA, 1978.

- [10] A. CICHOCKI AND S. AMARI, *Adaptive blind signal and image processing: Learning algorithms and applications*, John Wiley & Sons, Chichester, West Sussex, England, 2002.
- [11] L. COHEN, *Time-frequency analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [12] P. COMON, C. JUTTEN, AND J. HERAULT, *Blind separation of sources, part II: Problems statement*, Signal Processing, **24** (1), 11–20, 1991.
- [13] I. DAUBECHIES, *The wavelet transform, time-frequency localization and signal analysis*, IEEE Trans. Inform. Theory, **36** (5), 961–1005, 1990.
- [14] K. FUJITA, Y. TAKEI, A. MORIMOTO, AND R. ASHINO, *Mathematical view of a blind source separation on a time frequency space*, Appl. Math. Comput., **187**, 153–162, 2007.
- [15] K. GRÖCHENIG, *Foundations of time-frequency analysis*, Birkhäuser Boston, Inc., Boston, MA, 2001.
- [16] S. HAYKIN AND Z. CHEN, *The cocktail party problem*, Neural Computation, **17** (9), 1875–1902, 2005.
- [17] C. HEIL AND D. WALNUT, *Continuous and discrete wavelet transforms*, SIAM Rev., **31** (4), 628–666, 1989.
- [18] C. HEIL AND D. WALNUT, *Fundamental Papers in Wavelet Theory*, Princeton University Press, William Street, Princeton, NJ, 2006.
- [19] A. HYVÄRINEN, J. KARHUNEN, AND E. OJA, *Independent component analysis*, John Wiley & Sons, New York, NY, 2001. 根本幾・川勝真喜訳, **詳解 独立成分分析 信号解析の新しい世界**, 東京電機大学出版局, 2005.
- [20] A. JOURJINE, S. RICKARD, AND O. YILMAZ, *Blind separation of disjoint orthogonal signals: Demixing n sources from 2 mixtures*, in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 2985–2988, June 5-9, 2000, Istanbul, Turkey, IEEE Press, 2000.
- [21] C. JUTTEN AND J. HERAULT, *Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture*, Signal Processing, **24** (1), 1–10, 1991.
- [22] S. MALLAT, *A wavelet tour of signal processing*, Second ed., Academic Press, New York, NY, 1999.

- [23] G. MELLIN, II, M. PACHTER, AND J. RAQUET, *Closed-form solution for determining emitter location using time difference of arrival measurements*, IEEE Trans. Aerospace and Electronic Systems, **39** (3), 1056–1058, 2003.
- [24] 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.
- [25] D. NAPOLETANI, C. A. BERENSTEIN, P. S. KRISHNAPRASAD, AND D. C. STRUPPA, *Quotient signal estimation*, in Harmonic Analysis, Signal Processing, and Complexity, Progress in Mathematics, **238**, 151–162, Birkhäuser, Boston, MA, 2005.
- [26] D. E. RUMELHART, G. E. HINTON, AND R. J. WILLIAMS, *Learning internal representations by error propagation*, In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, **1**, 318–362, Cambridge, MA, MIT Press, 1986.
- [27] E. SOROUCHYARI, *Blind separation of sources, part III: Stability analysis*, Signal Processing, **24** (1), 21–29, 1991.
- [28] 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.
- [29] 村田昇, **入門 独立成分分析**, 東京電機大学出版局, 2004.
- [30] 守本晃・藤田景子・芦野隆一, **時間周波数情報を用いたブラインド信号源分離—実例を中心に—**, 信学技報, EA2005, SIP2005, 31–36, 2005.
- [31] 藤田景子・竹井義次・守本晃・芦野隆一・森本光生, **時間周波数情報を用いたブラインド信号源分離—数学的背景—**, 信学技報, EA2005, SIP2005, 37–42, 2005.
- [32] 守本晃, 本講究録の実験で使った音声などをおいたホームページ,
<http://www.osaka-kyoiku.ac.jp/~morimoto/researches/BSS/RIMS08BSS/>
- [33] A. Cichocki et al., ICALAB のホームページ,
<http://www.bsp.brain.riken.go.jp/ICALAB/>