

音声の局在時間を考慮した複数音源定位

Multiple Sound Source Localization in Consideration of
Local Existence Time of Speech Signals前田 隼一朗†
Souichirou Maeda染谷 貴徳†
Takanori Someya陶山 健仁†
Kenji Suyama

1 はじめに

人間は、人混みなど周囲の環境が未知の状況においても、先験的知識なしに特定の音源からの音を聴取可能である。このような能力はカクテルパーティー効果と呼ばれ、生体の持つ優れた能力として知られており、これには複数音源定位能力が関わっていると考えられている。したがって、生体並の機能を有するセンシングシステムの実現を目指すならば複数音源定位の実現が重要な課題となる。本研究ではマイクロホンアレーによる複数音源定位について考える。

代表的な複数音源定位としては、MUSIC(MUltiple SIgnal Classification)[1][2]が知られている。しかし、この手法はマイクロホン数以上の音源の定位が不可能である。一方、マイクロホン数に依存しない複数音源定位法として、音声のスパース性を用いた手法が注目されている[3]~[6]。文献[3][4]の手法は音声信号エネルギーが「時間一周波数」平面上でまばらに分布する性質を利用し、単一音源エネルギーのみを有する「時間一周波数」領域の推定結果のみを利用する手法である。特に文献[3]では、「時間一周波数」平面上の各点における定位結果を、ガウスカネルを用いて平滑化後ヒストグラムを作成し、そのピーク位置から音源方向を推定している。これらの手法では、単一音源定位法として演算負荷の小さい瞬時位相差を用いることが多い。

これに対し我々は、「時間一周波数」平面上の各点においてRoot-MUSIC[1]を使用する手法を検討中である[7]。検討手法では「時間一周波数」平面上の各点において、複数の忘却係数を用いて相関行列を作成し、Root-MUSICにより定位結果を算出する。そのため、音声信号の調波構造により、特定の周波数にエネルギーが局在する性質に基づいて、異なる局在時間に応じた相関行列の算出が可能である。その結果、局在時間に合致する忘却係数を用いた相関行列を用いた推定結果は、高精度な推定結果であることが期待できる。さらに、Root-MUSICは2マイクロホンの場合、単純な代数演算で陽に解を求めることが可能であり、演算負荷も小さい。そして定位結果のヒストグラムを作成する。その結果、瞬時位相差を用いる手法と比べ演算負荷を大きくすることなしに、正解方向に鋭いピークを有するヒストグラムの形成が期待できる。最終的に、作成したヒストグラムに確率分布をEMアルゴリ

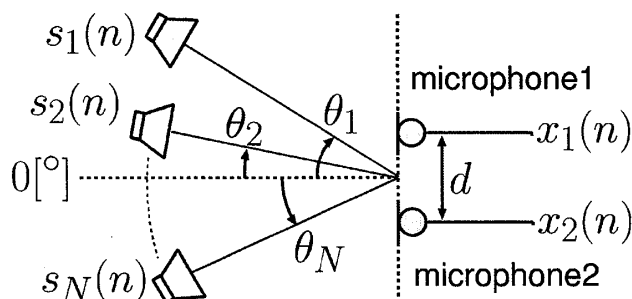


図1 2マイクロホンによる複数音源定位問題

ズムによりフィッティングし、音源方向を推定する。一般に、フィッティングには確率分布として混合正規分布が使用されるが、提案法では外れ値に頑健な混合コーシー分布[8]を使用する。

また本研究では、複数の忘却係数を用いたRoot-MUSICが単一音源定位法として有効であることを確認するために、実環境実験により、2マイクロホンを用いた音声のスパース性に基づく手法である文献[3]の手法と比較し、提案法の有効性を示す。

2 2マイクロホンによる複数音源定位問題

図1に示すように、空間中に存在する N 個の音源から発せられる音声信号 $s_i(n)$ ($i = 1, 2, \dots, N$)を2個のマイクロホンで受音する。ここで N は未知である。マイクロホン間隔 d に対し、マイクロホンから音源までの距離が十分に離れている場合、音源信号はマイクロホン付近では平面波モデルで近似できる。このときマイクロホン m の受信信号 $x_m(n)$ は、

$$x_m(n) = \sum_{i=1}^N s_i(n - (m-1)\tau_i) + \gamma_m(n) \quad (1)$$

となる。 τ_i ($i = 1, 2, \dots, N$)は、 $s_i(n)$ に対するマイクロホン間の到達時間差であり、次式で表すことができる。

$$\tau_i = \frac{d \sin \theta_i}{c} \quad (2)$$

ここで c は音速、 d はマイクロホン間隔である。また $\gamma_m(n)$ はマイクロホン m の観測雑音で、平均が0、パワーが σ^2 のガウスノイズであり、各マイクロホンにおいて独立、 $s_i(n)$ とは無相関であるとする。

† 東京電機大学工学部電気電子工学科
School of Engineering, Tokyo Denki University

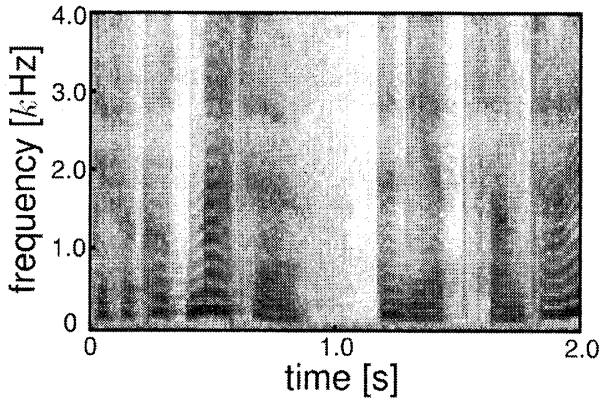


図2 音声のスペクトログラム

$x_m(n)$ を DFT (Discrete Fourier Transform) し、「時間一周波数」平面に展開した信号 $X_m(t, k)$ は次式で表すことができる。

$$X_m(t, k) = \sum_{i=1}^N S_i(t, k) e^{-j\omega_k(m-1)\tau_i} + \Gamma_m(t, k) \quad (3)$$

ここで t はフレーム番号, k は周波数帯域番号, ω_k は k に対応する角周波数である。また $S_i(t, k)$, $\Gamma_m(t, k)$ はそれぞれ $s_i(n)$, $\gamma_m(n)$ の DFT である。(3) 式をベクトル表記すると, 次式で表せる。

$$\begin{aligned} \mathbf{X}(t, k) &= [X_1(t, k), X_2(t, k)]^T \\ &= \mathbf{A}(k) \mathbf{S}(t, k) + \mathbf{\Gamma}(t, k) \end{aligned} \quad (4)$$

ただし, T は転置を表す。ここで $\mathbf{S}(t, k)$, $\mathbf{\Gamma}(t, k)$ はそれぞれ $S_i(t, k)$, $\Gamma_m(t, k)$ を要素にもつベクトルであり, 次式で示される。

$$\mathbf{S}(t, k) = [S_1(t, k), S_2(t, k), \dots, S_N(t, k)]^T \quad (5)$$

$$\mathbf{\Gamma}(t, k) = [\Gamma_1(t, k), \Gamma_2(t, k)]^T \quad (6)$$

また $\mathbf{A}(k)$ は, ステアリングベクトル $\mathbf{a}_k(\theta_i)$ を列ベクトルにもつ行列であり, 次式で表される。

$$\mathbf{A}(k) = [\mathbf{a}_k(\theta_1), \mathbf{a}_k(\theta_2), \dots, \mathbf{a}_k(\theta_N)] \quad (7)$$

$$\mathbf{a}_k(\theta_i) = [1, e^{-j\omega_k \tau_i}]^T \quad (8)$$

図2に音声のスペクトログラムを示す。この図は「時間一周波数」平面上の音声エネルギーの分布を表しており, 色が濃いほど音声エネルギーが大きいことを示している。図2に示すように, 音声エネルギーは「時間一周波数」平面上で疎らに存在する。この性質を音声のスプース性と呼ぶ。また音声には, 「時間一周波数」平面上で特定の周波数において, ある時間区間連続的にエネルギーが存在する性質がある。この

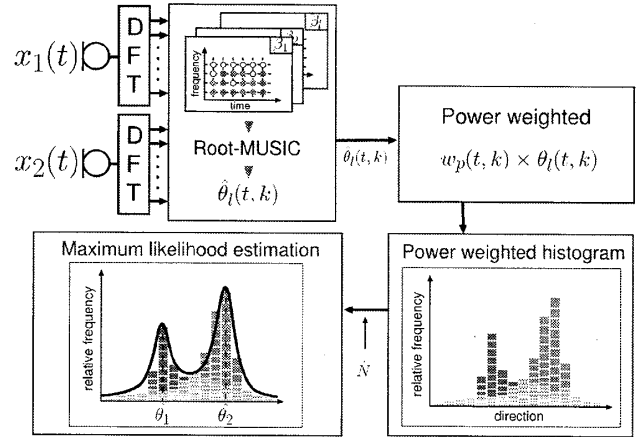


図3 提案法の流れ

性質を本研究では周波数局在性と呼ぶことにする。2個のマイクロホンによる複数音源定位問題は, $x_m(n)$ ($m = 1, 2$) から, θ_i ($i = 1, 2, \dots, N$) を推定する問題である。

3 提案法

提案法の流れを図3に示す。まず2マイクロホンの受信信号 $x_m(n)$ ($m = 1, 2$) を DFT (Discrete Fourier Transform) により, 「時間一周波数」平面上に展開する。次に「時間一周波数」平面上の各点において単一音源定位法である Root-MUSIC により音源定位を行う。Root-MUSIC に必要な相関行列 $\mathbf{R}_l(t, k)$ は, 複数の忘却係数 β_l ($l = 1, 2, \dots, 10$) を用いて算出する。

定位結果 $\theta_l(t, k)$ を (t, k) の信号パワー $w_p(t, k)$ で重み付け, ヒストグラムを作成する。最後に, 作成したヒストグラムに対して混合コーシー分布を当てはめ, 最尤推定により音源方向 $\hat{\theta}_i$ ($i = 1, 2, \dots, \hat{N}$) を推定する。

3.1 音声のスプース性を用いた複数音源定位

音声エネルギーは「時間一周波数」平面上で疎らに分布しており, 音声毎にその分布も異なる。そのため, 2つの音声信号が同時に存在した場合でも, 「時間一周波数」平面上では1音声エネルギーが支配的である領域が多く含まれていると考えられる。そのような領域では, 受信信号ベクトル $\mathbf{X}(t, k)$ は,

$$\mathbf{X}(t, k) \simeq S_i(t, k) \mathbf{a}_k(\theta_i) + \mathbf{\Gamma}(t, k) \quad (9)$$

と近似できる。そこで, 単一音源定位法を「時間一周波数」平面上の各点に対して適用し, 各点に対する推定結果 $\theta(t, k)$ を算出すると, 1音声エネルギーから求めた推定結果は正しく, 各 (t, k) においては1音声エネルギーが支配的であるため, $\theta(t, k)$ の多くはどちらか一方の音声信号の方向の推定結果となる。その結果, $\theta(t, k)$ のヒストグラムを作成すると, 正解方向に対応する音源方向にピークが形成されるため, ピークの位置を調べることで, 複数音源定位が可能となる。ピーク的位置は, DUET[4] などでは混合正規分布をあてはめ, その平均値

パラメータから推定するが、提案法では、EM アルゴリズムを用いてヒストグラムに混合コーシー分布を当てはめることで、最終的な音源定位結果を算出する。

3.2 単一音源定位法

各 (t, k) に対して 1 音源定位を行う手法として、一般に瞬時位相差が用いられる。しかし本研究では、高解像度音源定位法である Root-MUSIC を用いる。そこでまず瞬時位相差について説明し、次に Root-MUSIC について説明する。

3.2.1 瞬時位相差

瞬時位相差は、周波数領域での信号の位相差から音源方向を推定する方法である。音声のスパース性より (9) 式を仮定し (4) 式を書き直すと、

$$\begin{bmatrix} X_1(t, k) \\ X_2(t, k) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{-j\omega_k \tau_i} \end{bmatrix} S_i(t, k) + \begin{bmatrix} \Gamma_1(t, k) \\ \Gamma_2(t, k) \end{bmatrix} \quad (10)$$

となる。ここで (10) 式のマイクロホン m の観測雑音 $\Gamma_m(t, k)$ は一般的に $S_i(t, k)$ に対し小さな値であるため、マイクロホン m の受信信号はそれぞれ

$$X_1(t, k) \simeq S_i(t, k) \quad (11)$$

$$X_2(t, k) \simeq e^{-j\omega_k \tau_i} S_i(t, k) \quad (12)$$

と表すことができる。このとき、 $X_2(t, k) = X_1(t, k)e^{-j\omega_k \tau_i}$ となるため、次式より $X_m(t, k)$ の位相 $\angle X_m(t, k)$ から位相差 $\phi = -\omega_k \tau_i$ を次式で求める。

$$\phi = \angle X_2(t, k) - \angle X_1(t, k) \quad (13)$$

ここで ω_k は既知であるため、 ϕ より音声信号の到達時間差 τ_i が求まる。そして次式より、音源方向 θ_i を推定する。

$$\theta_i = \sin^{-1} \left(\frac{c\tau_i}{d} \right) \quad (14)$$

3.2.2 Root-MUSIC

Root-MUSIC[1] は MUSIC(Multiple Signal Classification) に基づく手法であるため、まず MUSIC について説明し、次に Root-MUSIC について説明する。

MUSIC は信号部分空間と雑音部分空間の直交性を利用する手法で、次式の MUSIC スペクトルを最大にする θ から音源方向を推定する。

$$P_{MU}(\theta) = \frac{\mathbf{a}_k^H(\theta) \mathbf{a}_k(\theta)}{\mathbf{a}_k^H(\theta) \mathbf{q}_k \mathbf{q}_k^H \mathbf{a}_k(\theta)} \quad (15)$$

ここで $\mathbf{a}_k(\theta)$ はステアリングベクトル、 \mathbf{q}_k は雑音部分空間に対応する受信信号の相関行列の固有ベクトル、 H はエルミート転置である。MUSIC は高解像度推定法であるが、 $P_{MU}(\theta)$ のピークサーチに多くの演算を要する。

これに対し Root-MUSIC は、MUSIC スペクトルの分母項から閉じた形で解を求めるため、ピークサーチ無しに高精度の音源方向推定が可能な手法である。Root-MUSIC では、(15) 式分母項より、

$$\mathbf{a}_k^H(\theta) \mathbf{q}_k \mathbf{q}_k^H \mathbf{a}_k(\theta) = 0 \quad (16)$$

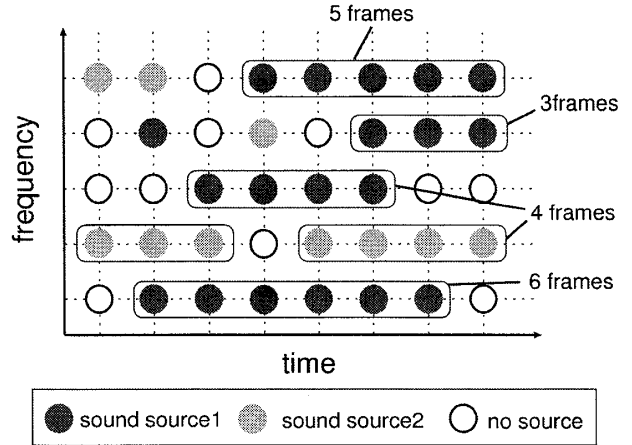


図4 「時間一周波数」平面

を満たす θ を直接求める。そのため、まずステアリングベクトル $\mathbf{a}_k(\theta)$ を次式で表す。

$$\mathbf{a}_k(\theta_i) = [1, e^{-j\omega_k \tau_i}]^T = [1, z]^T \equiv \mathbf{p}(z) \quad (17)$$

ただし、 $z = e^{-j\omega_k \tau_i}$ である。次に $\mathbf{p}(z)$ を用いて次式の Root-MUSIC 多項式を定義する。

$$Q(z) = z \mathbf{p}^T(z^{-1}) \mathbf{q}_k \mathbf{q}_k^H \mathbf{p}(z) = z \mathbf{U}(z) \mathbf{U}^*(1/z^*) \quad (18)$$

これは z の多項式となり、 $z = \exp(-j \frac{2\pi f}{c} d \sin \theta_i)$ のとき、 $Q(z) = 0$ となる。そのため (18) 式を解くことで、 $Q(z) = 0$ となる z から音源方向を陽に得ることができる。

3.2.3 忘却係数

Root-MUSIC を用いる場合、受信信号の相関行列 $\mathbf{R}(t, k)$ が必要である。音声信号は非定常信号であるため、時間変化により音声の性質も変化する。そのため精度の良い $\mathbf{R}(t, k)$ を算出するには、 $\mathbf{R}(t, k)$ 算出に反映するデータの長さを調整することが望ましい。そこで $\mathbf{R}(t, k)$ を算出する際に、次式のように忘却係数 $\beta (0 \leq \beta \leq 1)$ を用いて、 $\mathbf{R}(t, k)$ に使用するデータ数を調節する。

$$\mathbf{R}(t, k) = \mathbf{X}(t, k) \mathbf{X}^H(t, k) + \beta \mathbf{R}(t-1, k) \quad (19)$$

周波数局在性により、 $\mathbf{R}(t, k)$ を算出する際に、図4に示すように局在した単一音源エネルギーのみを使用することで、Root-MUSIC により信頼度の高い定位結果が得られると考えられる。また $\mathbf{R}(t, k)$ に使用するデータ数は、 β を調整することである特定の局在時間に対応させることができる。しかし、図4に示すように音声エネルギーの局在時間長は対象とする音声に依存するため、忘却係数値を1つに固定した場合、音声エネルギーの局在時間毎の $\mathbf{R}(t, k)$ の算出は困難である。そこで複数の忘却係数 β_i を用いて複数の $\mathbf{R}(t, k)$ を算出することで、異なる局在時間に対応させる。

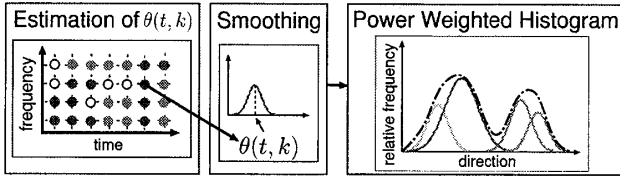


図5 比較手法の流れ

提案法では、 $R(t, k)$ を求める際に複数の忘却係数 $\beta_l (l = 1, 2, \dots)$ を用い、音声の異なる局在時間に対応することを考える。異なる局在時間に対応した相関行列 $R_l(t, k)$ を次式で表す。

$$R_l(t, k) = X(t, k)X^H(t, k) + \beta_l R_l(t, k) \quad (20)$$

(20) 式のように複数の忘却係数を用いることで、様々な局在時間をもつ音声エネルギーから相関行列を算出することができる。

3.3 最尤推定

本手法では Root-MUSIC の推定結果にパワーを重み付けてヒストグラムを作成する。作成したヒストグラムは音源方向に鋭いピークを形成するため、尖度の高い分布となり正規分布によるフィッティングが困難となる場合がある。そこで、正規分布より鋭い分布をもつコーシー分布をヒストグラムにフィッティングさせる。コーシー分布を決定するパラメータは、最頻値と半値半幅である。最頻値は正規分布のパラメータである平均値に比べ、外れ値に頑健である [9]。以上の理由より、提案法ではヒストグラムに次式で示す混合コーシー分布モデルを当てはめ、EM アルゴリズムにより最尤推定を行う。

$$F(\theta) = \sum_{i=1}^N \rho_i \left\{ \frac{1}{\pi} \left(\frac{\lambda_i}{(\theta - \hat{\theta}_i)^2 + \lambda_i^2} \right) \right\} \quad (21)$$

$$\sum_{i=1}^N \rho_i = 1 \quad (22)$$

ここで i は音源のインデックス、 ρ_i は混合比、 $\hat{\theta}_i$ は最頻値、 λ_i は半値半幅を表す。EM アルゴリズムにより得られた混合コーシー分布の最頻値パラメータ $\hat{\theta}_i$ を最終的な音源方向推定結果とする。

4 実環境実験

最初に実環境実験により提案法で使用する忘却係数値の検討を行い、定位に適した忘却係数値を算出する。次に実環境実験により提案法と比較手法を比較し、提案法の性能評価を行う。比較手法には、音声のスパース性を用いた手法である文献 [3] の手法を用いた。比較手法の流れを図5に示す。まず「時間-周波数」平面上の各点に対して瞬時位相差による音源定位を行い、定位結果 $\theta(t, k)$ の頻度を次式により平滑化してパワー重み付きヒストグラムを作成し、そのピークから音源方向を推定する。

表1 実験条件

サンプリング周波数	8[kHz]
使用マイクロホン数	2
マイクロホン間隔 d	40[mm]
フレーム長	512
使用周波数帯域	2[kHz] ~ 4[kHz]
提案法セル幅	1[°]
比較手法セル分割数	1000

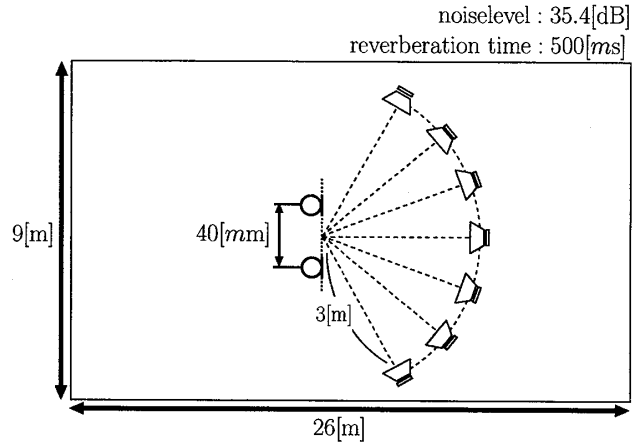


図6 実験環境 (部屋の高さ:2.5m, マイクロホン及び音源の高さ:1.0m)

$$H_{\theta, s}(\theta) = \sum_{k=1}^K |\hat{S}(t, k)|^2 \frac{W_\phi}{\sqrt{\pi}} e^{-W_\phi \|\theta - \theta(t, k)\|^2} \quad (23)$$

ここで W_ϕ は観測雑音の分散 σ^2 に反比例する値 ($\frac{1}{2\sigma^2}$)、 $\hat{S}(t, k)$ は信号パワーの推定値である。ただし観測雑音の分散 σ^2 は未知であるため、適切な W_ϕ を事前実験により選択する必要がある。提案法ではこのような事前実験は必要ない。

表1に実験条件を示す。 W_ϕ は、936 パターンの音源配置に対する事前実験より適切な値として $10^{-1.5}$ を選択した。実験は図6に示すような通常の会議室で行い、音源として男声と女声を用い、マイクロホンアレーの中心から約3[m]の位置に複数の音源を配置した。また音声は交互に発話するのではなく、多くの部分で発話区間が重畳する音声を用いた。本研究では音源定位性能のみ評価したいため音源数を与えて実験を行った。

4.1 忘却係数値の検証

忘却係数値が過度に大きい場合、多くの時間フレームを用いて相関行列を算出するため、スパース性が成り立たない。提案法は音声のスパース性に基づく手法であるため、スパース性が成立する適切な忘却係数値を決定する必要がある。そこで、忘却係数値 $\beta_l (l = 1, 2, \dots)$ を検証するために実環境実験

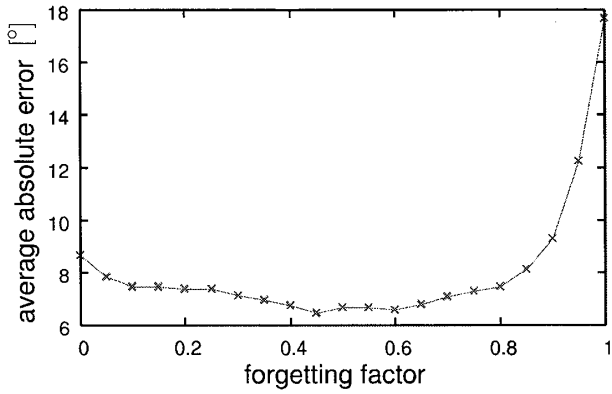


図7 忘却係数を固定した場合の平均絶対値誤差

を行った。

相関行列を算出する際に忘却係数値を1つだけ使用し、その忘却係数を0～1の範囲で0.05刻みに変化させ、936パターンの2音源の組み合わせに対して実験を行った。図7に、各忘却係数に対する絶対値誤差の平均値を示す。図7の実験結果より、忘却係数が0～0.5の範囲では誤差がほぼ単調減少していることが確認できた。また忘却係数が0.5以上の場合、誤差が増加することを確認した。以上の結果より、忘却係数値が0～0.5まではスパース性および周波数局在性が成立していると考え、使用する忘却係数の値を $\beta_l = 0.05l$ ($l = 1, 2, \dots, 10$)とした。

4.2 提案法の性能評価

提案法の性能評価を行うために、実環境実験により提案法と比較手法を比較する。図8、図9に比較手法と提案法の実験結果を示す。ここで実線は音源方向の真値、点線は推定結果である。推定誤差は、比較手法では 4.85° 、提案法では 0.91° となった。次に近接音源に対する結果を図10、図11に示す。図10に示す比較手法の結果では、音源方向を示すピークが形成されず1方向のみ推定している。しかし図11に示す提案法の結果では、音源方向に2つのピークを形成し、2方向を推定していることが確認できる。そのため提案法は、近接音源に対しても有効な手法であることが確認できた。また音源の種類と音源配置を組み合わせる936パターンの実験を行い、提案した音源定位法の性能評価を行った。図12、図13に、実験に用いた2音源の方向の差に対する推定値の絶対値誤差を示す。ここで図12は比較手法、図13は提案法の実験結果である。図12より、比較手法は音源間隔が約 $30^\circ \sim 60^\circ$ の範囲では比較的高精度な推定が行えているが、その範囲より音源間隔が狭い場合や音源間隔が広すぎる場合は、推定誤差が増加することがわかる。近接音源に対する推定誤差の増加は、図10に示したように、平滑化により片方の音源方向に対応するピークが形成できないためと考えられる。また音源間隔を広げた場合の誤差の増加は、推定結果の外れ値がピーク形成に寄与することが原因であると考えられる。これに対し提案法では、

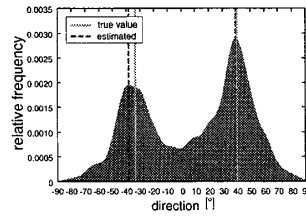


図8 比較手法

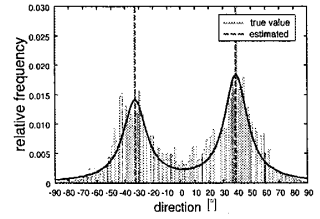


図9 提案法

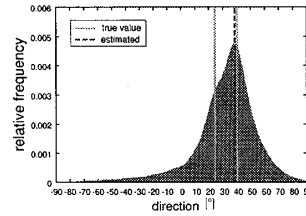


図10 比較手法

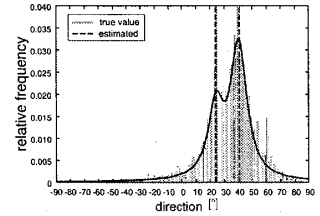


図11 提案法

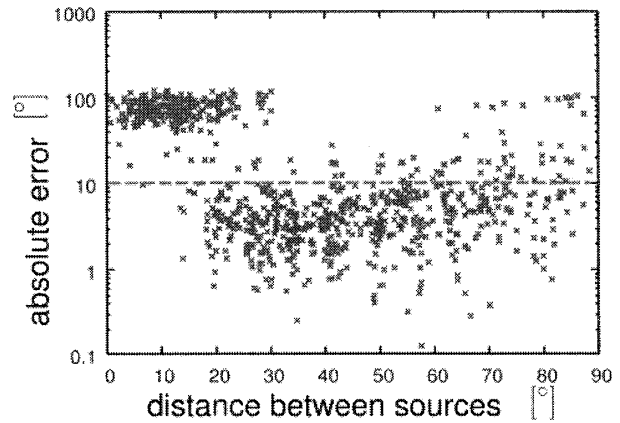


図12 比較手法の音源間隔に対する誤差分布

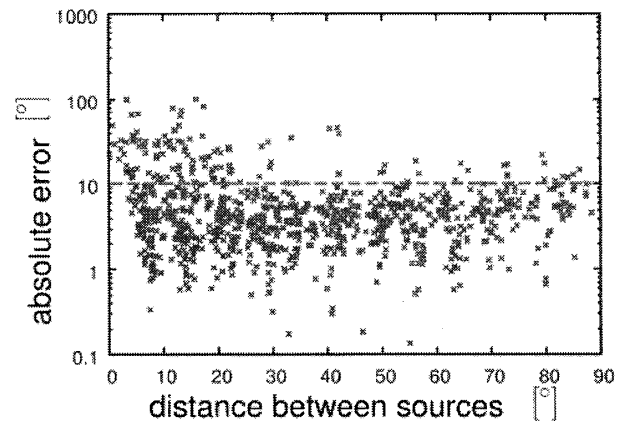


図13 提案法の音源間隔に対する誤差分布

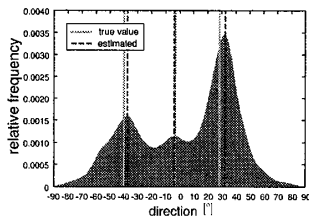


図14 比較手法

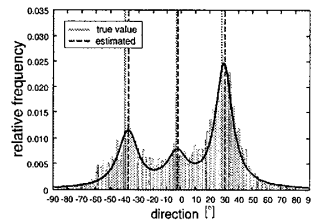


図15 提案法

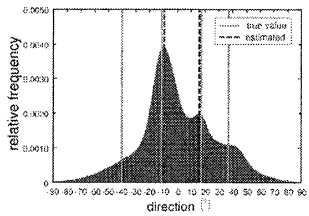


図16 比較手法

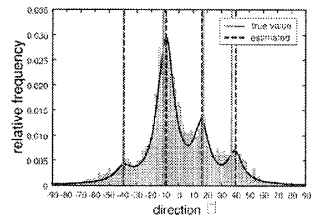


図17 提案法

図13に示すように、音源間隔が 30° 以下の範囲では高精度な推定が行えていないが、音源間隔が 30° より大きい範囲では、推定誤差をほぼ 10° 以内に抑えられることが確認できる。これは複数の忘却係数を用いたことで、比較手法のように平滑化を行わずともヒストグラムの分布がある程度滑らかになったため、音源方向に対応するピークを潰すことなく、ヒストグラムを作成できていることが大きく影響していると考えられる。これに加え、ヒストグラムに対し、外れ値に頑健な混合コーシー分布を用いて最尤推定を行うことで、明らかな外れ値の影響を最小限に抑えていることも、この要因のひとつであると考えられる。また936パターンの実験結果の絶対値誤差の平均値は、提案法では 6.95° 、比較手法では 27.0° となり、この結果からも提案法の推定精度が高いことを確認した。

次に実験に用いる音源数を増した場合の推定結果を比較する。図14、図15に音源数が3音源の結果、図16、図17に音源数が4音源の結果を示す。図14、図15に示した比較手法と提案法の推定誤差は、それぞれ 6.72° 、 5.53° であった。この結果から、3音源の場合でも音源間隔が十分に離れていれば、両手法で定位可能であることが確認できた。図16、図17に示した結果では、比較手法では4つの音源方向に対して、ピーク形成に成功したのは2方向のみである。これに対し提案法では、4方向にヒストグラムのピークが形成できており、4音源に対しても定位可能であることが確認できた。この結果から、提案手法は近接音源に対しても有効であるため、音源数が増加し、音源間隔が狭くなった場合においても有効な手法である。

以上の結果より、提案法が従来の瞬時位相差を利用した手法と比べ、高精度に複数音源定位が行えるといえる。

5 まとめ

本研究では、単一音源定位法としてRoot-MUSICを用い、最尤推定に混合コーシー分布を用いて複数音源定位を行った。また相関行列算出に適した忘却係数値の検討を行った。そして、提案法の性能評価を行うため、実環境実験により他の複数音源定位法との性能比較を行った。実験結果より、適切なパラメータを選択した比較手法と比べても、提案法は推定精度が高いことを確認した。これより、複数の忘却係数を用いたRoot-MUSICを単一音源定位法として適用し、最尤推定に混合コーシー分布を用いることが、音源定位性能の向上に有効であることを示した。

参考文献

- [1] 菊間信良, “アレーアンテナによる適応信号処理,” 科学技術出版, pp.191-209, 1998.
- [2] 大賀寿郎, 山崎芳男, 金田豊, “音響システムとデジタル信号処理,” コロナ社, 1995.
- [3] T. Melia, A. Moni, S. Rickard, “Maximum Likelihood Estimation in Sparse Systems and the Power Weighted Histogram Estimation Tool,” Proc. of IMA Conf. on Mathematics in Signal Processing, Cirencester, 2006.
- [4] O. Yilmaz, S. Rickard, “Blind Separation of Speech Mixtures via Time-Frequency Masking,” IEEE Trans. Signal Processing, vol. 52, no. 7, pp 1830-1847, July 2004.
- [5] Y. Izumi, N. Ono, S. Sagayama, “Sparseness-based 2ch BSS using the EM Algorithm in Reverberant Environment,” Proc. of WASPAA2007, pp.147-150, Oct. 2007.
- [6] S. Araki, H. Sawada, R. Mukai and S. Makino, “DOA estimation for multiple sparse sources with normalized observation vector clustering,” Proc. of ICASSP2006, vol.5, pp.33-36, May 2006.
- [7] 染谷貴徳, 陶山健仁, “信頼性重み付け分布による複数音源定位の性能改善,” 信学技報, SIP2008-181, pp.101-106, 2009.
- [8] 石樽彩乃, 吉田裕亮, “EM アルゴリズムの混合コーシー分布への応用とその改良,” 情報学数理モデル化と問題解決研報, vol.2006, no.135, pp.85-88, 2006.
- [9] 小柳義夫, “ロバスト推定法とデータ解析,” 物理学誌, vol.34, no.10, pp.884-891, 1979.