

遅延和アレー法を用いた複数移動音源追尾の性能改善

Performance Improvement of Multiple Sound Sources Tracking Method
Using Delayed Sum Array Method押田 賢浩† 大和田 昇† 陶山 健仁†
Masahiro Oshida Noboru Owada Kenji Suyama

1 はじめに

移動音源追尾技術は、ロボット聴覚、遠隔会議システムなどにおいて重要な要素技術の1つである。これを実現する方法として、マイクロホンアレーを用いた手法 [1] が有効である。これまで、移動音源追尾手法としてパーティクルフィルタなど、さまざまな手法が提案されてきた [2]~[4]。一方、我々は、高分解能推定法である MUSIC (MUltiple SIgnal Classification)[5] に基づいて、逐次更新アルゴリズムである PAST (Projection Approximation Subspace Tracking)[6] と IPLS (Interior Point Least Square)[7] を用いた単一移動音源追尾手法について検討してきた。PAST は固有値分解なしに受信信号の相関行列の固有ベクトルを RLS アルゴリズムにより逐次更新則で推定する手法であり、IPLS はその固有ベクトルに基づいた MUSIC スペクトル関数の分母を評価関数とし、その極小解を逐次更新で求める手法である。PAST-IPLS により、低演算量かつ高精度な移動音源追尾を実現した [8]。

また、遠隔会議などのアプリケーションへ応用を考えた場合、話者は複数存在することが考えられる。その際に、PAST-IPLS を用いて高精度な複数移動音源追尾を実現するには、IPLS の実行可能領域を適切に設定する必要がある。音声が無音状態である場合や音声エネルギーが小さい場合、IPLS 評価関数はその音源方向に極小解をもたない。IPLS 評価関数は多峰性関数であるために、音源方向とは異なる局所最小解に収束する場合がある。その後、発話状態に復帰し、IPLS 評価関数が音源方向に極小解を有しても、実行可能領域が不適切であるために音源方向を推定することは困難である。そのため、IPLS 評価関数の極小解に対応する音源数が増加したときに、実行可能領域の再設定が高精度追尾に対して不可欠となる。

実行可能領域を再設定するには、発話音源数と実行可能領域の手がかりとなる音源方向を推定する必要がある。これらを推定するために、文献 [9] では、MUSIC スペクトルの粗いピークサーチ、文献 [10] では「時間一周波数」平面上の各点で瞬時位相差により推定した音源方向のパワー重み付けヒストグラムを用いて、実行可能領域を再設定した。しかし、MUSIC スペクトルの粗いピークサーチは MUSIC スペクトル関数を用いるため、演算量が高負荷である。また、瞬時位相

差を用いたヒストグラムは MUSIC スペクトルの粗いピークサーチと比べ演算量を低減しているものの、少ないデータ数からヒストグラムを作成しなければならず、発話音源数の推定精度が低い。そのため、追尾精度が劣化する場合があった。

そこで、本研究では、実行可能領域の再設定手法として空間エイリアシングを考慮した遅延和アレー法 [1],[11] を用いて、PAST-IPLS による複数移動音源追尾の性能改善を行なう。遅延和アレー法に用いるマイクロホンは、演算量低減と空間解像度向上を図るため、マイクロホン数を限定し、マイクロホン間隔を広げる。また、マイクロホン間隔を広げることで生じる空間エイリアシングの影響を PAST により算出した信号部分空間を基に、音源選択を行なうことで回避する。

実環境実験より、提案法が実行可能領域の再設定にパワー重み付けヒストグラムを用いた手法と比べ同等の演算量で追尾可能であり、追尾精度が向上することを示す。

2 問題設定

図1のように、空間中に2つの音源 $s_i(n)$ ($i = 1, 2$) が $\theta_i(n)$ 方向に存在し、その音響信号を間隔 d で等間隔に配置した M 個のマイクロホンで受信する。ここで、 n は離散時間である。 d に対し、マイクロホンと $s_i(n)$ の距離が十分に離れているならば、音響信号は平面波モデルとして仮定でき、マイクロホン間の到達時間差 $\tau_{i,n}$ は、 $\tau_{i,n} = d \sin \theta_i(n) / c$ となる。ここで、 c は音速である。マイクロホン1を基準とすると、 m 番目のマイクロホンの受信信号 $x_m(n)$ は、次式のようになる。

$$x_m(n) = \sum_{i=1}^2 s_i(n - (m-1)\tau_{i,n}) + \gamma_m(n) \quad (1)$$

ここで、 $\gamma_m(n)$ はマイクロホン m の観測雑音であり、平均が0、パワーが σ^2 とする。また、 $\gamma_m(n)$ は $s_i(n)$ と無相関であると仮定する。(1)式を離散フーリエ変換 (DFT: Discrete Fourier Transform) により、周波数領域に変換すると、次式のようになる。

$$\mathbf{X}_t(k) = \sum_{i=1}^2 S_{i,t}(k) \mathbf{a}_k(\tau_{i,t}) + \mathbf{\Gamma}_t(k) \quad (2)$$

$$= [X_{t,1}(k), X_{t,2}(k), \dots, X_{t,M}(k)]^T \quad (3)$$

ここで、 T は転置、 k は周波数番号、 t はフレーム番号、 $S_{i,t}(k)$ は $s_i(n)$ の k 番目の周波数における複素振幅である。また、 $\mathbf{a}_k(\tau_{i,t})$ はステアリングベクトル、 $\mathbf{\Gamma}_t(k)$ は雑音ベクトルであ

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

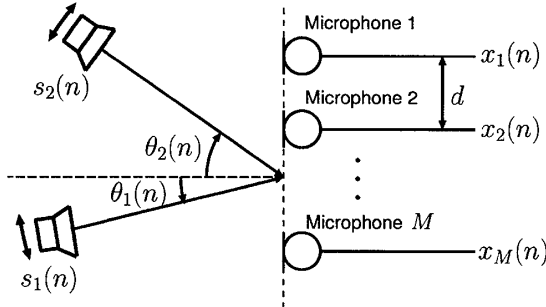


図1 複数移動音源追尾問題設定

り、次式のように表わせる。

$$\mathbf{a}_k(\tau_{i,t}) = [1, e^{-j\omega_k \tau_{i,t}}, \dots, e^{-j\omega_k (M-1)\tau_{i,t}}]^T \quad (4)$$

$$\mathbf{\Gamma}_t(k) = [\Gamma_{t,1}(k), \Gamma_{t,2}(k), \dots, \Gamma_{t,M}(k)]^T \quad (5)$$

ここで、 $\tau_{i,t} = d \sin \theta_{i,t} / c$ であり、 $\theta_{i,t}$ はフレーム t における音源方向である。 $\tau_{i,t}$ を求めることにより、 $\theta_{i,t} = \sin^{-1}(c\tau_{i,t}/d)$ から $\theta_{i,t}$ を推定することが可能である。また、 $\omega_k = 2\pi f_s k / N$ であり、 f_s はサンプリング周波数、 N はフレーム長である。

複数移動音源追尾問題は、 $\mathbf{X}_t(k)$ から $\theta_{i,t}$ を推定する問題である。

3 複数移動音源追尾

本手法は MUSIC[5] に基づく手法である。MUSIC は受信信号の相関行列 $\mathbf{R}_{k,t} = \overline{\mathbf{X}_t(k)\mathbf{X}_t^H(k)}$ の固有値分解によって算出する信号部分空間 $\mathbf{Q}_S(k,t)$ と雑音部分空間 $\mathbf{Q}_N(k,t)$ の直交性に基づく手法である。ここで、 $\overline{\cdot}$ は時間平均、 H はエルミート転置を表わす。MUSIC は次式の MUSIC スペクトル関数 $P_{\text{MU}}(\tau, k)$ から、 $\mathbf{a}_k(\tau)$ と $\mathbf{Q}_N(k,t)$ の直交性を評価し、 $P_{\text{MU}}(\tau, k)$ のピークに対応する τ から算出する θ が音源方向となる。

$$P_{\text{MU}}(\tau, k) = \frac{\mathbf{a}_k^H(\tau)\mathbf{a}_k(\tau)}{\mathbf{a}_k^H(\tau)\mathbf{Q}_N(k,t)\mathbf{Q}_N^H(k,t)\mathbf{a}_k(\tau)} \quad (6)$$

しかし、MUSIC には、部分空間を算出するための固有値分解と $P_{\text{MU}}(\tau, k)$ のピークサーチに高い演算量を必要とする。そこで、固有値分解の代わりに PAST を用いて部分空間を推定し、ピークサーチの代わりに IPLS を用いて音源方向を推定する。これにより、PAST-IPLS は低演算量で高精度な移動音源追尾を可能とする。

図2に提案法の流れを示す。まず、 $x_m(n)$ を DFT により周波数領域に変換する。次に、周波数ごとに PAST から $\mathbf{Q}_S(k,t)$ を推定する。また、遅延和アレー法により音源数と実行可能領域の手がかりとなる音源方向 $\tilde{\theta}_{i,t}$ を推定する。最後に、PAST によって算出した $\mathbf{Q}_S(k,t)$ を用いて IPLS により音源方向を推定する。IPLS はフレーム t ごとに音源方向 $\hat{\theta}_{i,t}$ を逐次更新して推定する。しかし、遅延和アレー法により

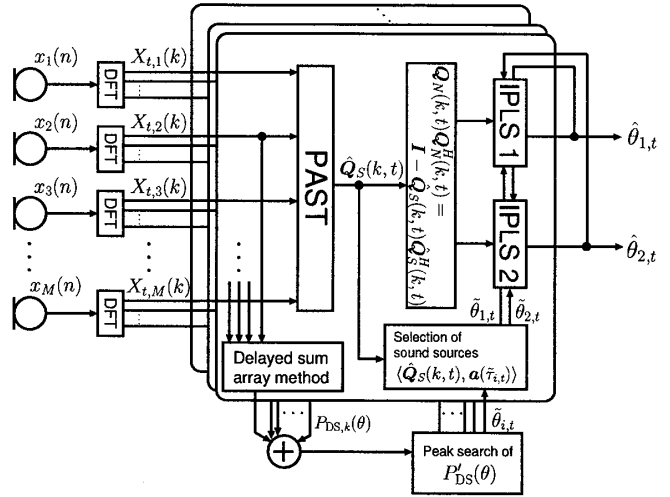


図2 提案手法の流れ

推定した音源数が前フレームと比べ、増加した場合には実行可能領域を再設定するために IPLS は $\tilde{\theta}_{i,t}$ を用いて推定する。一方、音源数が減少または変化がない場合には前フレームの推定結果 $\hat{\theta}_{i,t-1}$ を用いる。最終的な音源方向の推定結果は帯域ごとに IPLS により推定した $\hat{\theta}_{i,t}$ の平均値とする。

3.1 PAST

音源の位置が変化した場合、部分空間もそれに伴い変動する。そのため、観測ごとに固有値分解による部分空間推定が必要となり、高負荷な演算となる。そこで、演算量を低減するために PAST[6] を用いて、固有値分解なしで $\mathbf{Q}_S(k,t)$ を逐次更新して求める。PAST は 1 フレームごとに次式の評価関数 $J_P(\mathbf{Q}_S(k,t))$ を最小にする $\hat{\mathbf{Q}}_S(k,t)$ を逐次更新則で求める。

$$J_P(\mathbf{Q}_S(k,t)) = \sum_{i=1}^t \beta^{t-i} \|\mathbf{X}_i(k) - \mathbf{Q}_S(k,t)\mathbf{h}(i)\|^2 \quad (7)$$

ただし、 $\mathbf{h}(i) = \hat{\mathbf{Q}}_S^H(k, i-1)\mathbf{X}_i(k)$ 、 β は忘却係数である。PAST によって推定した $\hat{\mathbf{Q}}_S(k,t)$ から、 $\mathbf{Q}_N(k,t)\mathbf{Q}_N^H(k,t) = \mathbf{I} - \hat{\mathbf{Q}}_S(k,t)\hat{\mathbf{Q}}_S^H(k,t)$ とし、IPLS に適用し移動音源追尾を行なう。PAST の更新則は文献 [9] に従う。

3.2 IPLS

MUSIC スペクトルのピークサーチによる高負荷な演算の代わりに、IPLS[7] を用いて逐次的に低演算量で音源方向を推定する。そこで、(6) 式の分母を用いて、次式の評価関数 $J_{I,k}(\tau)$ を最小にする $\hat{\tau}_{i,t}$ を求める。

$$J_{I,k}(\tau) = \mathbf{a}_k^H(\tau)\mathbf{Q}_N(k,t)\mathbf{Q}_N^H(k,t)\mathbf{a}_k(\tau) \quad (8)$$

しかしながら、図3のように $J_{I,k}(\tau)$ は多峰性の関数であり、局所最小解を多数もつ。そこで、IPLS では制約条件を設け、探索範囲を極小解付近に限定する。すなわち、次式のような制約付最小化問題を考える。

$$\begin{aligned} \min. \quad & J_{I,k}(\tau) \\ \text{sub.to} \quad & J_{I,k}(\tau) \leq \zeta_{i,t}, \tau^2 \leq \eta^2 \end{aligned} \quad (9)$$

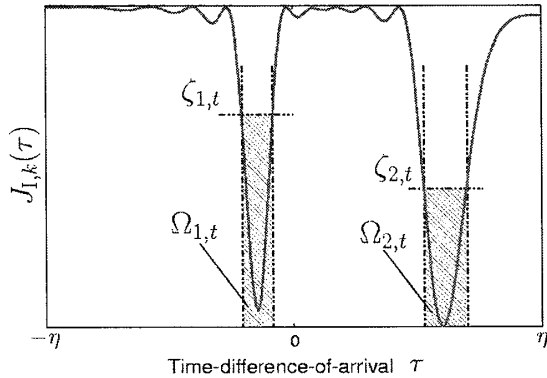


図3 IPLS 評価関数と実行可能領域

ここで、 $\zeta_{i,t}$ は $J_{1,k}(\tau)$ の上限である。また、 η は $|\tau|$ が取りうる範囲であり、 $\eta = d/c$ である。(9) 式に対する実行可能領域 $\Omega_{i,t}$ は次式のように表わせる。

$$\Omega_{i,t} = \{\tau \in \mathbb{R} | J_{1,k}(\tau) \leq \zeta_{i,t}, \tau^2 \leq \eta^2\} \quad (10)$$

$\Omega_{i,t}$ を設定することにより、安定した音源方向推定が可能である。また、 $\zeta_{i,t}$ を次式のように更新し、 $J_{1,k}(\tau)$ の変動に応じて $\Omega_{i,t}$ を伸縮する。

$$\zeta_{i,t} = J_{1,k}(\hat{\tau}_{i,t-1}) + \mu \frac{\eta}{\sqrt{2}} |\nabla J_{1,k}(\hat{\tau}_{i,t-1})| \quad (11)$$

ただし、 μ は伸縮のステップサイズ、 ∇ は微分演算子である。また、 $\hat{\tau}_{i,t-1}$ は $t-1$ フレームで推定した音源方向 $\hat{\theta}_{i,t-1}$ から算出した到達時間差である。 $J_{1,k}(\tau)$ を最小にする $\hat{\tau}_{i,t}$ は、ニュートン法の1ステップにより近似して求める。IPLS の更新則は文献 [9] に従う。

3.3 実行可能領域の再設定

音声が発話状態では、通常 $J_{1,k}(\tau)$ はその音源方向に対して極小解を有する。しかし、音声が無音状態の場合や音声エネルギーが小さい場合には、極小解が得られない場合がある。また、PAST の β により、追従性能の追従性を制御可能であるが、追従性と安定性にはトレードオフの関係がある。そのため、 β を小さく設定して高速な追従を目指した場合、PAST の更新が数値的不安定に陥りやすいという RLS アルゴリズムと共通の性質をもつ。そのとき、 $\hat{Q}_S(k, t)$ の更新は不安定となり、音源が発話中においてさえ、音源方向に極小解が得られない場合がある。このように、音源方向付近に極小解が存在しないとき、 $J_{1,k}(\tau)$ が多峰性評価関数であるが故に、IPLS で推定する τ は音源方向とは異なる局所解に収束する。その後、 $J_{1,k}(\tau)$ が再び音源方向に極小解を有しても、その音源方向を推定することは困難である。この問題を回避するために、 $J_{1,k}(\tau)$ の変動と共に、IPLS の $\Omega_{i,t}$ を再設定する必要がある。

3.3.1 従来用いた実行可能領域の再設定手法

IPLS の $\Omega_{i,t}$ を再設定するために、これまでに2つの手法が提案されている [9],[10]。文献 [9] では、MUSIC スペクト

ルの粗いピークサーチを用いて、ピーク数とピークに対応する方向から、音源数と $\Omega_{i,t}$ の初期値を推定したが、(6) 式の MUSIC スペクトル関数を扱うため、膨大な演算量を必要とした。また、演算量を低減する目的として、文献 [10] では、音声のスパース性に着目し、「時間一周波数」平面上の各点から瞬時位相差により音源方向を推定し、そのパワー重み付けヒストグラムから $\Omega_{i,t}$ を再設定した。しかしながら、ヒストグラムは少ないデータサンプルで作成するため、その精度は低く、2音源が発話時に1音源方向のみピークを形成し、もう一方の音源方向に対してピークの形成が困難な場合がある。その結果、 $\Omega_{i,t}$ が再設定されずに追従誤差が大きくなる場合があった。

3.3.2 遅延和アレー法

本研究では、低演算量かつ高精度に音源数と音源方向を推定するために、遅延和アレー法 [11] を用いる。遅延和アレー法は任意の方向に指向性ビームを向け、ビーム方向に対するアレーの出力パワー $P_{DS,k}(\theta)$ を算出する。そのため、 $P_{DS,k}(\theta)$ がピークとなる θ が音源方向となる。 $P_{DS,k}(\theta)$ は次式で算出する。

$$P_{DS,k}(\theta) = \mathbf{a}_k^H(\tau) \mathbf{X}_t(k) \quad (12)$$

音源が音声の場合、スパース性により、音声エネルギーをもたない帯域が存在する。そこで、次式のように全ての使用周波数帯域の $P_{DS,k}(\theta)$ を加算したパワースペクトル $P'_{DS}(\theta)$ を用いて音源方向を推定する。

$$P'_{DS}(\theta) = \sum_k P_{DS,k}(\theta) \quad (13)$$

遅延和アレー法は、MUSIC と比べ低演算量であるが、従来法で用いた瞬時位相差と比べると演算量が高負荷である。そこで、遅延和アレー法に用いるマイクロホン数を限定することで演算量を低減するとともに、マイクロホン間隔を広げ解像度を向上する。しかし、マイクロホン間隔 d を広げた場合、 $d \leq c/(2f)$ を満たさない高周波数帯域において空間エイリアシングが生じ、偽りのピークが現れる。一方、 $d \leq c/(2f)$ を満たす低周波数帯域では、パワースペクトルの解像度が低く、また残響の影響により、推定精度が低い。そこで、高解像度である高周波数帯域のみ用いて、 $P'_{DS}(\theta)$ のピークを次節の音源選択により、音源方向に対応したピークであるか判断する。 $P'_{DS}(\theta)$ は図4のように、粗い間隔でピークサーチし、傾きが正から負に変わったところを音源方向の候補 $\tilde{\theta}_{i,t}(i = 1, 2, \dots)$ とした。

3.3.3 音源選択

マイクロホン間隔を広げることにより、高周波数帯域で空間エイリアシングが生じる。その結果、遅延和アレー法により推定したパワースペクトルには、偽りのピークが存在する。そこで、音源に対応するピークを判別するために、帯域ごとに音源選択 [10] を行なう。音源選択では、次の関係を仮定する。

$$\text{span}\{\mathbf{Q}_S(k, t)\} = \text{span}\{\mathbf{a}_k(\tau_{1,t}), \mathbf{a}_k(\tau_{2,t})\} \quad (14)$$

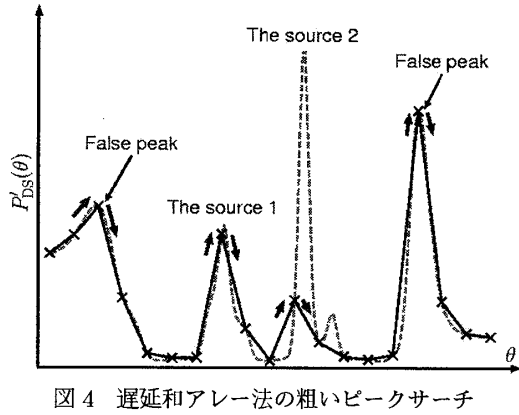


図4 遅延和アレー法の粗いピークサーチ

ただし, $\text{span}\{\mathbf{a}_k(\tau_{1,t}), \mathbf{a}_k(\tau_{2,t})\}$ はベクトル $\mathbf{a}_k(\tau_{1,t})$, $\mathbf{a}_k(\tau_{2,t})$ によって張られる部分空間である。すなわち, PAST によって推定した $\hat{\mathbf{Q}}_S(k, t)$ とパワースペクトルのピークに対応する $\hat{\theta}_{i,t}$ から算出する $\mathbf{a}_k(\hat{\tau}_{i,t})$ を用いて音源選択を行なう。 $\mathbf{a}_k(\hat{\tau}_{i,t})$ が $\hat{\mathbf{Q}}_S(k, t)$ 上に存在するならば, $\hat{\theta}_{i,t}$ は音源に対応するピークであり, 逆に $\mathbf{a}_k(\hat{\tau}_{i,t})$ が $\hat{\mathbf{Q}}_S(k, t)$ 上に存在しなければ, $\hat{\theta}_{i,t}$ は音源に対応しない偽りのピークであると判断する。音源選択では, 内積 $\langle \hat{\mathbf{Q}}_S(k, t), \mathbf{a}_k(\hat{\tau}_{i,t}) \rangle$ を算出し, 内積値が閾値 C_{th} より高い場合, $J_{1,k}(\tau)$ は $\hat{\theta}_{i,t}$ 方向付近に極小解をもつと考える。そして, C_{th} を越えた $\hat{\theta}_{i,t}$ の個数をフレーム t での音源数とする。なお, C_{th} を越えた $\hat{\theta}_{i,t}$ の数が2以上の場合, 内積値の大きい $\hat{\theta}_{i,t}$ から $\Omega_{i,t}$ の再設定に採用し, 音源数を2とする。音源数が増加したとき, 次のように IPLS の $\Omega_{i,t}$ を再設定する。

- (1) 音源数が0, 1 から2に増加した場合, 小さい $\hat{\theta}_{i,t}$ から IPLS1, IPLS2 に $\Omega_{i,t}$ を再設定する。
- (2) 音源数が0 から1に増加した場合, 前フレームの IPLS による推定結果 $\hat{\theta}_{i,t-1}$ と $\hat{\theta}_{i,t}$ を比較し, 差が小さい方の IPLS に $\Omega_{i,t}$ を再設定する。

なお, 音源数が減少, または変化がない場合には $\Omega_{i,t}$ を再設定せず, IPLS による更新を継続する。

図5に遅延和アレー法により推定したパワースペクトルと IPLS 評価関数の一例を示す。 $J_{1,k}(\tau)$ の○は $\hat{\theta}_{i,t}$ から算出する内積値が C_{th} より高い場合を表している。すなわち, $\hat{\theta}_{i,t}$ を極小解とみなした場合である。一方, ×は C_{th} より低い場合であり, $\hat{\theta}_{i,t}$ を局所解とみなした場合である。(a)は遅延和アレー法により, 4つのピークが生じていることが確認できる。一方, $J_{1,k}(\tau)$ には極小解が $\hat{\theta}_{3,t}$ 方向付近にあり, このとき $\hat{\theta}_{3,t}$ のみ音源に対応する方向と判断し, 音源数は1とする。また, (b)では, 5つのピークが生じ, $J_{1,k}(\tau)$ は極小解を $\hat{\theta}_{3,t}$, $\hat{\theta}_{4,t}$ 付近に2つ有することが確認できる。よって, このとき音源数は2とし, 前フレームで音源数が1以下の場合, $\hat{\theta}_{3,t}$ から算出した $\hat{\tau}_{3,t}$ を $\hat{\tau}_{1,t-1}$, $\hat{\theta}_{4,t}$ から算出した $\hat{\tau}_{4,t}$ を $\hat{\tau}_{2,t-1}$ とし, (11)式から $\Omega_{i,t}$ を再設定する。

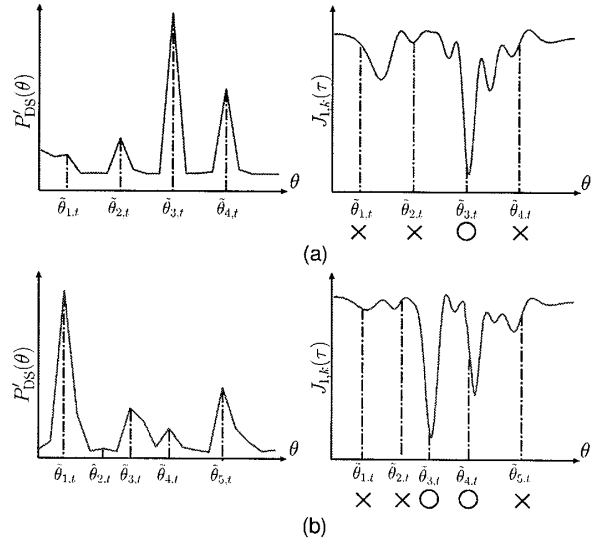


図5 パワースペクトルと IPLS 評価関数

4 実環境実験

提案法の有効性を確認するために実環境実験を行なった。実験に用いた部屋は騒音レベルが 37.6 [dB], 残響時間が 0.8 [s], 大きさが 10.5 [m]×9.0 [m]×2.4 [m] の通常の会議室である。表1に実験条件を示す。PAST-IPLS は $M = 16$, $d = 40$ [mm] のマイクロホンアレーを使用した。使用周波数帯域は残響の影響を受けにくく,かつ IPLS の評価関数が高解像度である 2~4 [kHz] を使用した。また, IPLS の初期値 $\hat{\tau}_{i,0}$ は 0° に設定した。遅延和アレー法に用いるマイクロホンは, メインアレーの受信信号 $x_m(n)$ に対して, $m = 2, 6, 10, 14$ を使用した。すなわち, $M = 4$, $d = 160$ [mm] のサブアレーである。遅延和アレー法の使用周波数帯域は前述の理由により 3~4 [kHz] とした。また, $P'_{DS}(\theta)$ のピークサーチ間隔は予備実験により 10° , 音源数判定の閾値 C_{th} は 0.7 と設定した。

提案法の有効性を検証するために, 従来法として $\Omega_{i,t}$ の再設定にパワー重み付けヒストグラムを用いた手法 [10] と比較した。実験に用いた音声データは図6の4つの音声 ((a),(b)-女声, (c),(d)-男声) の内, 2つの音声を用いてスピーカから出力した。また, 話者の位置をカメラアレーを用いて画像計測し, 話者方向の真値 $\theta_{i,t}$ とした。複数移動音源追従性能の評価方法は次式により誤差 ε を算出した。

$$\varepsilon = \sqrt{\frac{1}{2} \sum_{i=1}^2 |\theta_{i,t} - \hat{\theta}_{i,t}|^2} \quad (15)$$

ここで, ε は音声が発話状態にあるフレームのみで評価した。また, PAST の β , IPLS の μ は ε を最小にする値を用いた。

4.1 追従結果

実験では4通りのパターンで移動した。表2に4つの実験の使用音声と β , μ , ε を示す。また, 図7に全フレーム, 全

表1 実験条件

Condition	Value
f_s	8 [kHz]
N	256
M	16
d	40 [mm]
Frequency band used for the PAST-IPLS	2~4 [kHz]
Initial value of IPLS $\hat{\tau}_{i,0}$	0
The number of microphones for delayed sum array method	4
Microphone distance for delayed sum array method	160 [mm]
Frequency band used delayed sum array method	3~4 [kHz]
Width of peak search	10°
C_{th}	0.7

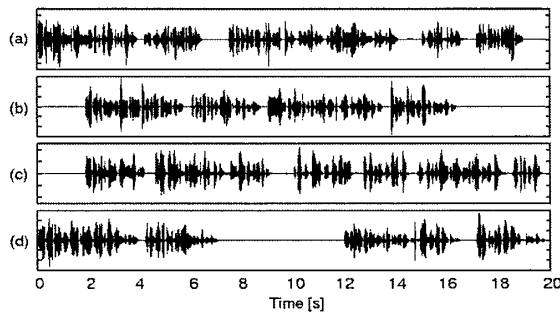


図6 実験に使用した音声波形

使用周波数帯域に対する $\Omega_{i,t}$ を再設定したフレーム、周波数の割合を示す。全ての実験において、提案法は従来法と比べ $\Omega_{i,t}$ の再設定の回数が2~10ポイント増え、提案法の精度が向上していることが確認できた。また、 β が小さい値であり $\hat{Q}_S(k,t)$ が不安定な更新であっても、 $\Omega_{i,t}$ の再設定を遅延和アレー法で行なうことによって、高精度な追尾が実現可能であることがわかる。

次に、追尾結果例として、移動パターン1、移動パターン2で実験したときの結果を示す。図8、図9に移動パターン1での提案法、従来法による追尾結果をそれぞれ示す。なお、追尾結果における図中の真値の断線部分は音源が無音状態またはパワーが小さい区間であり、 ε の評価に含めていない。従来法では、2~6 [s] 間の2音源発話時に IPLS1 の追尾結果が、音源方向から外れていることがわかる。一方、提案法では、 $J_{1,k}(\tau)$ の極小解が変化する度に、 $\Omega_{1,t}$ を再設定することによって、安定して音源方向を追尾していることが確認できた。

図10、図11に移動パターン2での提案法、従来法による追尾結果を示す。移動パターン2では、2音源の追尾開始時の方向はそれぞれ -20° 、 20° であり、IPLS の初期値 0° から離れている。また、5 [s] や 16 [s] 付近において、従来法は IPLS1 の追尾結果が音源方向から外れてる。しかし、提案法は従来法と比べ、 $J_{1,k}(\tau)$ の極小解の変化に対応して $\Omega_{i,t}$ を再設定しているため、音源方向に高速に追従していることが確認できる。

最後に、2 [s] ごとに提案法と従来法の ε を計算し、追尾精度の改善量 δ を調べた。図12に各時間における δ を示す。提

表2 従来法、提案法の最適な β , μ と ε

No.	$s_i(n)$	Conventional method			Proposed method		
		β	μ	ε	β	μ	ε
1	(a),(b)	0.08	0.01	7.99	0.07	0.01	4.91
2	(a),(d)	0.01	0.01	9.09	0.10	0.01	5.78
3	(a),(c)	0.75	0.03	7.00	0.89	0.03	5.08
4	(c),(d)	0.10	0.01	12.12	0.12	0.01	8.24

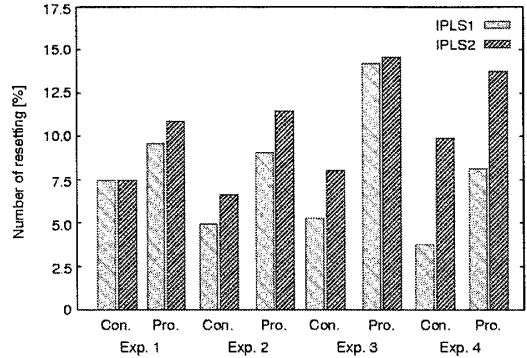


図7 従来法と提案法の実行可能領域の再設定の割合

案法が全ての時間において追尾性能が向上しており、また提案法による追尾性能の改善は10 [s] 以内に顕著に現れていることが確認できる。これは提案法が追尾開始時に $\Omega_{i,t}$ の再設定によって高速に音源方向に追従することが可能なためである。また、最大 9.34° の改善が実験1の4~6 [s] 間で確認できた。

4.2 計算時間

従来法と提案法の演算量の差を確認するために、計算時間を調査した。処理時間の測定には、Intel[®] Core[™]2 Duo2.6 [GHz]、メモリ2 [GByte] を搭載したPCを使用した。信号長が20 [s] のデータを処理するのに、従来法の計算時間は平均11.5 [s]、提案法の計算時間は平均11.9 [s] であった。1 [s] 当りの計算時間はそれぞれ、0.58 [s]、0.60 [s] となり、従来法と提案法に演算時間の差がないことを確認済である。

5 むすび

本研究では、PAST と IPLS を用いた複数移動音源追尾の性能改善のために、実行可能領域の新たな再設定手法について検討した。提案手法では、空間エイリアシングを考慮した遅延和アレー法を用いて、音源方向と音源数を推定した。実験結果より、提案手法は従来法と比べ同等の演算時間で処理し、追尾精度が向上することを示した。

参考文献

- [1] 大賀寿郎, 山崎芳男, 金田豊, 音響システムとデジタル処理, pp.173-218, コロナ社, 1995.
- [2] A. Quinlan, M. Kawamoto, Y. Matsusaka, H. Asoh and F. Asano, "Tracking intermittently speaking multiple speakers using a particle filter," EURASIP

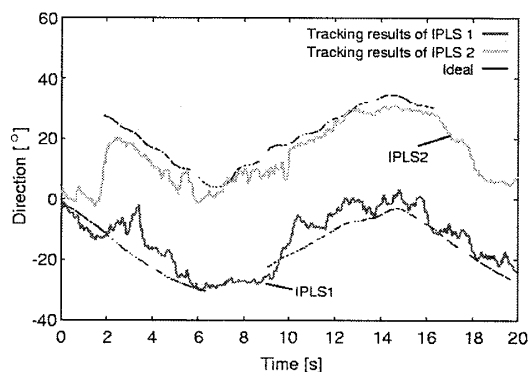


図8 提案法による追尾結果 (移動パターン 1)

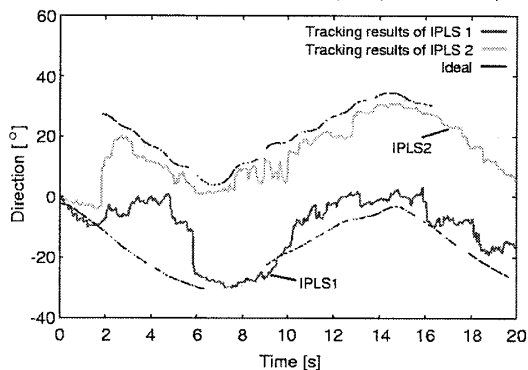


図9 従来法による追尾結果 (移動パターン 1)

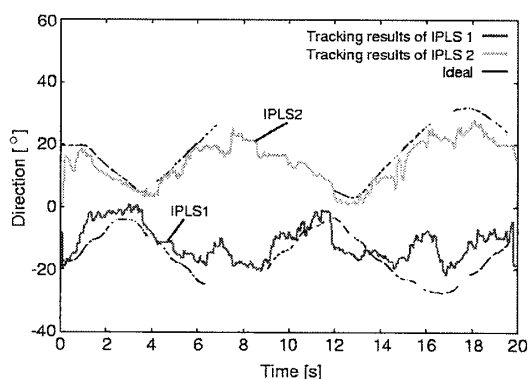


図10 提案法による追尾結果 (移動パターン 2)

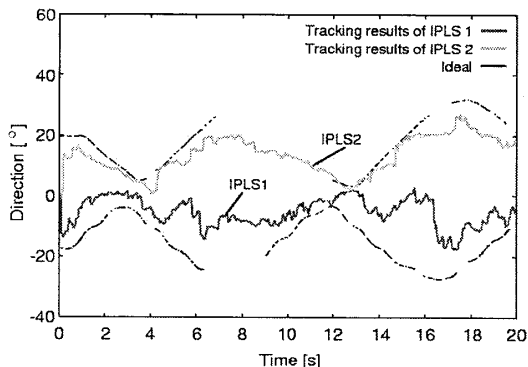


図11 従来法による追尾結果 (移動パターン 2)

Journal on Audio, Speech and Music Process., vol.2009, Article ID 673202, pp.1-9, 2009.

- [3] J.S. Hu, C.Y. Huang, C.H. Yang, C.K. Wang and M.T. Lee, "Sound source tracking and speech enhancement by microphone array on an embedded dual-core processor platform," Proc. IEEE ARSO2008, pp.1-6, 2008.
- [4] V. Cevher, R. Velmurugan and J.H. McClellan, "Acoustic multitarget tracking using direction-of-arrival bathces," IEEE. Trans. Signal Process., vol.55, no.6, pp.2810-2825, 2007.
- [5] R.O. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas Propag., vol.34, no.3, pp.276-280, 1986.
- [6] B. Yang, "Projection approximation subspace tracking," IEEE Trans. Signal Process., vol.43, no.1, pp.95-107, 1995.
- [7] K.H. Afkhaime, Z.-Q. Luo and K.M. Wong, "Adaptive liner filtering using interior point optimization techniques," IEEE Trans. Signal Process., vol.48, no.6, pp.1637-1648, 2000.
- [8] D. Tsuji and K. Suyama, "A moving sound source tracking based on two successive algorithms," Proc. IEEE ISCAS2009, pp.2577-2580, 2009.

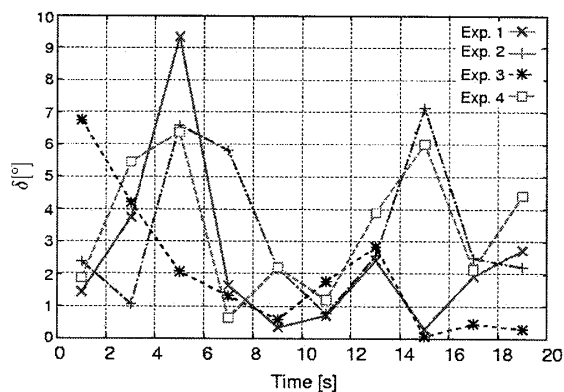


図12 提案法と従来法の誤差

- [9] 大和田昇, 陶山健仁, "多峰性評価関数に対する逐次解法に基づく複数移動音源追尾," 信学論 (A), vol.J93-A, no.4, pp.248-259, 2010.
- [10] N. Ohwada and K. Suyama, "Multiple sound sources tracking method based on subspace tracking," Proc. IEEE WASPAA2009, pp.217-220, 2009.
- [11] 田中雅史, 金田豊, 小島順治, "音源方向推定法の室内残響下での性能評価," 音響誌, vol.50, no.7, pp.540-548, 1994.