

補助関数法に基づくサブサンプル時間差推定

Sub-Sample Time Delay Estimation for Two Signals with Auxiliary-Function-Based Optimization

山岡洸瑛[†] シャイブラーロビン[†] 小野順貴[†] 若林佑幸[†]
[†] 首都大学東京大学院

Kouei Yamaoka[†] Robin Scheibler[†] Nobutaka Ono[†] Yukoh Wakabayashi[†]
[†] Graduate School of Systems Design, Tokyo Metropolitan University

アブストラクト 本稿では、効率的な更新則により、2つの観測信号間の時間差推定をサブサンプルの精度で行う手法を提案する。我々は、サブサンプル時間差推定を連続の時間変数をもつ一般化相互相関 (continuous generalized cross correlation; cGCC) の最大化問題として定式化した。cGCC を最大化するという最適化問題に対して、これを解く閉形式の解は未だ得られていない。そこで本稿では、補助関数法に基づいて cGCC を効率的に最大化する手法を提案する。まず、cGCC が正弦波の和で表せることから、cGCC の最大化において2次の補助関数を設計できることを示す。次に、設計した補助関数を用いて cGCC を最大化する更新則を示す。実験では、提案法は目的関数の単調増加が保証されていることを示し、また、良い初期値が得られた場合には高速に収束することを示す。さらに、提案法が雑音にロバストであること、時間差の推定値がクラメル・ラオの下限を達成することも示す。

1 はじめに

音声強調や音源分離、音源定位などといったマイクロフォンアレイ信号処理において、空間的な特徴はそれらの性能向上に寄与する重要な情報である [1]。空間的な特徴は、異なるマイクロフォン間（チャンネル間）の振幅の変位及び時間差（もしくは位相差）によって表される。特に時間差は重要な情報であり、これを用いたマイクロフォンアレイ信号処理は、音源分離 [2] だけでなく、音源定位 [3, 4]、サンプリング周波数ミスマッチ推定 [5]、録音信号の同期 [6] などといった多くの技術を実現する。一般に、時間差の推定精度の改善は、これらの信号処理の性能向上に直結する。

2チャンネルのサンプリングされた信号について、チャンネル間の時間差を推定することを考える。単純には、2チャンネルの信号の離散点に対して計算された相互相関 (cross correlation; CC) の最大値を求めることで、最大値に対応する時間差が推定値として得られる。このとき、何らかの

後処理を適用しない限り、時間差推定の精度はサンプリング周波数によって制限される。これは、特に小さなマイクロフォンアレイを使う場合には、無視できない問題である。例として、4 cm 間隔で配置された2チャンネルマイクロフォンにおける最大の到来時間差 (time difference of arrival; TDOA) は 0.12 ms 未満であり、例えば、サンプリング周波数が 16 kHz の信号に対しては2サンプル未満である。このようなサンプルの精度の時間差推定の精度は、音響信号処理だけでなく、ソナー [7]、レーダー [8]、反射地震学 [9] などといった多くの重要な技術にとって明らかに不十分である。

時間差推定において広く使われる手法は、一般化相互相関 (generalized cross correlation; GCC) [10, 11] に基づく推定である。ここで、時間差の推定値は2チャンネルの信号間の離散の時間変数をもつ一般化相互相関 (discrete GCC; dGCC) の最大値の位置により与えられる。次に、得られた最大値の近傍において、サンプル間の GCC の値を補間することで、サブサンプルの精度の推定値を得ることができる。補間手法として、2次関数のフィッティングによるもの [12] やガウス関数のフィッティングによるもの [13] など、多くの手法が提案されている [14–16]。また、周波数領域におけるゼロ詰め（ゼロパディング）も有効な手法であり、これは GCC のディリクレカーネル補間 [17] に対応する。ゼロ詰めにより達成可能な推定精度は、ゼロ詰め未適用の信号と適用後の信号の比率に相当する。例えば、10 倍の推定精度が欲しい場合は、周波数領域において信号長が 10 倍になるようにゼロ詰めする必要がある。補間に基づく手法に対して、連続な GCC の最大値を直接求めることも可能である。帯域制限された信号の場合、標本化定理 [18, 19] に従い、連続の時間変数をもつ一般化相互相関 (continuous GCC; cGCC) は dGCC の sinc 補間により得られる。cGCC の最大化は非凸問題であり、一般に閉形式の解は知られていない。しかし、黄金分割探索 (golden-section search; GSS) [20] などの探索アルゴリズム

を使うことで、局所解を得ることは可能である。GSS による推定では、最大値が存在するとわかっている区間を狭めることにより、単峰性関数の最大値（または最小値）を探索する。

本稿では、補助関数法（もしくは majorization-minimization (MM) アルゴリズム [21, 22]) に基づく新たなサブサンプル時間差推定法を提案する。提案法では、補助関数の反復最大化により、効率的に cGCC を最大化する。我々はまず、目的関数が2次の補助関数によって大域的に抑えられることを示す。このとき、補助関数は目的関数の単調増加を保証しながら反復的に最大化することができる。実験による提案法の評価では、初期値を真の最大値の近くにとることで、数回の反復で最適解に収束することを示す。また、ガウスノイズの存在下でクラメール・ラオの下限 (Cramér–Rao lower bound; CRLB) を達成すること、さらに、残響環境における性能も示す。GSS を用いた推定と比較すると、提案法は通常、より少ない反復で収束することが確認できる。

1 章では本稿の概要を述べた。2 章では、まず時間差推定の定式化を示し、3 章で提案するアルゴリズムについて説明する。実験とその結果は 4 章で示し、結論を 5 章で述べる。

2 問題設定

x_n 及び y_n を離散信号とする。このとき、 x_n と y_n の GCC は以下で表される。

$$\Phi^{(xy)}(t) = \frac{1}{N} \sum_{k=-N/2+1}^{N/2} W_k S_k^{(xy)} e^{j2\pi kt/N}, \quad (1)$$

$$S_k^{(xy)} = X^*(k)Y(k), \quad (2)$$

ここで、 k は離散周波数を表すインデックス、 $W_k \in \mathbb{R}^+$ は任意の重み関数であり、 $t \in \mathbb{R}$ は x_n と y_n 間の時間差を表す時間変数である。また、 $S_k^{(xy)}$ は x_n と y_n のクロススペクトルである。ただし、 $X(k), Y(k)$ は、それぞれ x_n, y_n の離散フーリエ変換 (discrete Fourier transform; DFT) であり、 $\{\cdot\}^*$ は複素共役を表す。 t は通常、離散の時間変数として扱われ、本稿ではこれを dGCC と呼ぶ。

GCC [10, 11] は、(1) を最大化する離散時間差 t を推定するために広く使われる。また、適切な重み関数 W_k を利用した手法として、例えば、(3) に示すような重みを用いた GCC-phase transform (PHAT) や GCC-smoothed coherence transform (SCOT) などが提案されている。

$$W_k^{\text{PHAT}} = |S_k^{(xy)}|^{-1}, \quad W_k^{\text{SCOT}} = \left(S_k^{(xx)} S_k^{(yy)} \right)^{-\frac{1}{2}}, \quad (3)$$

また、 $W_k = 1$ のとき、GCC は通常の CC と一致する。一般に、上記 dGCC は入力信号のサンプリング周波

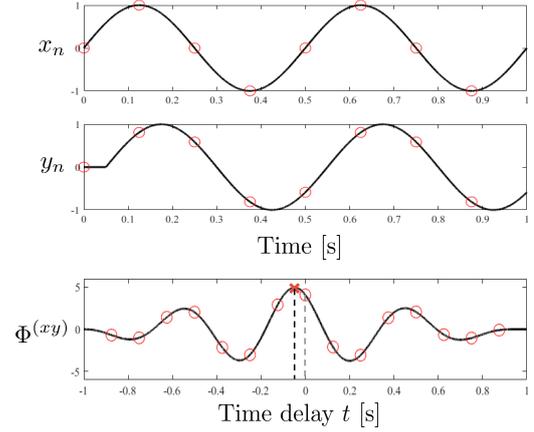


図 1: x_n, y_n に対する GCC の例。○は各サンプルに対するの値を表し、×はサブサンプルの精度の時間差を表す。

数 F_s により与えられる離散の時間差、すなわち、 $t \in \left\{ \frac{k}{F_s} \mid k = -\frac{N}{2} + 1, \dots, \frac{N}{2} \right\}$ に対してのみ計算される (図 1 参照)。従って、時間差の推定精度はサンプリング周波数に依存した限界があり、性能を向上させるためにはサブサンプルの精度での時間差推定が必要不可欠である。

dGCC に対し、(1) の時間変数 t が連続値を取ると考え、これを連続関数であると考えることができる。そこで、連続の時間変数をもつ一般化相互相関 (continuous GCC; cGCC) を最大化する連続の時間変数 $t \in \mathbb{R}$ の推定を考える。(1) は、 $S_k^{(xy)}$ の複素共役対称性を利用することで、以下のように正弦波の和の形に変形することができる。

$$\Phi^{(xy)}(t) = \sum_{k=0}^{N/2} A_k \cos(\omega_k t + \phi_k), \quad (4)$$

ここで、 $A_k = \frac{\beta_k}{N} |W_k S_k^{(xy)}|$, $\omega_k = 2\pi \frac{k}{N}$, $\phi_k = \angle S_k^{(xy)}$ であり、 $\beta_0 = \beta_{N/2} = 1$, $\beta_k = 2$ ($k \notin \{0, N/2\}$) である。従って、我々の目的は、次式で示されるサブサンプルの精度の時間差推定値 \hat{t} を求めることである。

$$\hat{t} = \arg \max_{t \in \mathbb{R}} \Phi^{(xy)}(t). \quad (5)$$

しかし、この最適化問題に対する閉形式の解は未だ得られていない。

$\Phi^{(xy)}(t)$ は正弦波の和で表されることから、短い区間においては単峰性の関数であると仮定することができる。実際に、 $\Phi^{(xy)}(t)$ はその最大値付近において単峰性関数である。この時、GSS [20] などの探索アルゴリズムを適用することで、探索区間内の最大値を見つけることができる。しかし、GSS により最大値を見つけるためには、探索区間が最大値を含んでいる必要がある。具体的には、探索区間の左端点と右端点は、それぞれ最大値付近の単峰範囲の増加部分と減少部分から選択する必要がある。

これに対し本稿では、閉形式の解を持つ補助関数を用いることで、(4)を最大化する新たな更新アルゴリズムを提案する。このとき最大値を見つける条件は、初期値が真の最大値が存在する単峰区間に含まれることのみである。

3 補助関数法に基づくサブサンプル時間差推定

補助関数法 (MM アルゴリズム [21,22] としても知られる) は、expectation-maximization (EM) アルゴリズムの一般化ともいえる枠組みであることや、様々なアルゴリズム [23–25] が提案されていることから広く知られた手法である。提案する最適化問題 (5) を解くためには、以下を満たす補助関数 $Q(t, \theta)$ を設計する必要がある。

- 任意の t 及び θ に対して、 $\Phi^{(xy)}(t) \geq Q(t, \theta)$ を満たす
- 任意の t に対して、 $\Phi^{(xy)}(t) = Q(t, \theta)$ を満たすような $\theta = f(t)$ が存在する

ただし、 $\theta = (\theta_0, \theta_1, \dots, \theta_{N/2})$ は補助変数である。上記を満たす $Q(t, \theta)$ が存在し、初期値 \hat{t}_0 が与えられたとき、次式による一連の更新は局所解に収束することが保証される。

$$\theta^{(\ell)} = f(\hat{t}^{(\ell)}), \quad \hat{t}^{(\ell+1)} = \arg \max_{t \in \mathbb{R}} Q(t, \theta^{(\ell)}), \quad (6)$$

ここで ℓ は反復回数を表すインデックスである。

3.1 連続の時間変数をもつ相互相関関数に対する 2 次の補助関数

本節では、(4) で表される目的関数 $\Phi^{(xy)}(t)$ に対する補助関数を示す。

定理 1 次式は (4) で表される $\Phi^{(xy)}(t)$ に対する補助関数である。

$$Q(t, \theta) = \sum_{k=0}^{N/2} -\frac{A_k}{2} \cdot \frac{\sin \theta_k}{\theta_k} (\omega_k t + \phi_k + 2n_k \pi)^2 + C, \quad (7)$$

ただし、 C は t を含まない定数項を表し、 n_k は $|\omega_k t + \phi_k + 2n_k \pi| \leq \pi$ を満たす整数である。補助変数は θ_k 及び n_k であり、 $\Phi^{(xy)}(t) = Q(t, \theta)$ となる等号成立条件は次式である。

$$\theta_k = \omega_k t + \phi_k + 2n_k \pi \quad (8)$$

この定理は、余弦関数に関する次の不等式から直接的に導かれる。

補題 1 $|\theta_0| \leq \pi$ とする。任意の実数 θ について、次の不等式が満たされる。

$$\cos \theta \geq -\frac{1}{2} \frac{\sin \theta_0}{\theta_0} \theta^2 + \left(\cos \theta_0 + \frac{1}{2} \theta_0 \sin \theta_0 \right). \quad (9)$$

$|\theta_0| < \pi$ のとき、等号成立の必要十分条件は $|\theta| = |\theta_0|$ であり、 $|\theta_0| = \pi$ のとき、等号成立の必要十分条件は $\theta = (2n+1)\pi, n \in \mathbb{Z}$ である。

証明:

$$f(\theta) = \cos \theta + \frac{1}{2} \frac{\sin \theta_0}{\theta_0} \theta^2 - \left(\cos \theta_0 + \frac{1}{2} \theta_0 \sin \theta_0 \right) \quad (10)$$

とする。このとき

$$f'(\theta) = -\sin \theta + \frac{\sin \theta_0}{\theta_0} \theta = -\theta \left(\frac{\sin \theta}{\theta} - \frac{\sin \theta_0}{\theta_0} \right) \quad (11)$$

である。ここで、次の 3 つの場合、それぞれについて検討する。

Case 1: $0 < |\theta_0| < \pi$

$\sin \theta / \theta$ は $0 \leq \theta \leq \pi$ の範囲において単調減少するため、

$$f'(\theta) \begin{cases} < 0 & (0 \leq \theta < |\theta_0|) \\ = 0 & (\theta = |\theta_0|) \\ > 0 & (|\theta_0| < \theta \leq \pi) \end{cases} \quad (12)$$

である。従って、 $f(\theta)$ は $\theta = |\theta_0|$ で最小値を取る。さらに、 $f(\theta_0) = 0$ であるから、 $0 \leq \theta \leq \pi$ の範囲において $f(\theta) \geq 0$ である。 $f(\theta)$ は偶関数であるから、 $-\pi \leq \theta \leq \pi$ の範囲においてもその最小値は 0 である。 $\cos \theta$ は周期関数だが、一方で $-\theta^2$ は周期的ではないため、任意の θ ($-\pi \leq \theta \leq \pi$) と整数 $n \neq 0$ に対して、 $f(\theta + 2n\pi) > f(\theta)$ である。以上より $f(\theta) \geq 0$ であり、等号成立の必要十分条件は $|\theta| = |\theta_0|$ である。

Case 2: $\theta_0 = 0$

このとき、 $0 \leq \theta \leq \pi$ において、 $f'(\theta)$ は

$$f'(\theta) \begin{cases} = 0 & (\theta = |\theta_0| = 0) \\ > 0 & (|\theta_0| < \theta \leq \pi) \end{cases} \quad (13)$$

となる。ここで、 $f(\theta)$ は偶関数であるから、 $-\pi \leq \theta \leq \pi$ の範囲において、最小値 $f(0) = 0$ をとる。従って Case 1 と同様に、 $f(\theta) \geq 0$ を得る。ただし等号成立の必要十分条件は $\theta = 0$ である。

Case 3: $\theta_0 = \pm \pi$

このとき、 $f(\theta) = \cos \theta + 1$ である。従って、 $f(\theta) \geq 0$ であり、等号成立の必要十分条件は任意の $n \in \mathbb{Z}$ について $\theta = (2n+1)\pi$ である。■

このとき、定理 1 は次のように証明される。

証明: $n_k \in \mathbb{Z}$ に対して、 $\cos(\omega_k t + \phi_k) = \cos(\omega_k t + \phi_k + 2\pi n_k)$ であり、また、 $A_k \geq 0$ である。従って、命題 1 を (4) の各項毎に (k 毎に) 適用することができる。■

3.2 補助関数と更新則の導出

$Q(t, \theta)$ は 2 次関数であるから, t について微分して 0 と置くことで, 簡単に最大化することができる.

$$\frac{\partial Q}{\partial t} = - \sum_{k=0}^{N/2} A_k \omega_k \frac{\sin \theta_k}{\theta_k} (\omega_k t + \phi_k + 2n_k \pi) = 0. \quad (14)$$

従って,

$$\hat{t} = \frac{\sum_{k=0}^{N/2} A_k \omega_k^2 (\sin \theta_k / \theta_k) (-(\phi_k + 2n_k \pi) / \omega_k)}{\sum_{k=0}^{N/2} A_k \omega_k^2 (\sin \theta_k / \theta_k)} \quad (15)$$

となる. ここで, (8) に示す等号条件の下で, $\phi_k + 2n_k \pi = \theta_k - \omega_k t$ であるから, 以下の最終的な更新則を得る.

$$n_k^{(\ell)} \leftarrow \arg \min_{n \in \mathbb{Z}} \left| \omega_k t^{(\ell)} + \phi_k + 2n\pi \right|, \quad (16)$$

$$\theta_k^{(\ell)} \leftarrow \omega_k t^{(\ell)} + \phi_k + 2n_k^{(\ell)} \pi, \quad k = 0, \dots, \frac{N}{2}, \quad (17)$$

$$t^{(\ell+1)} \leftarrow t^{(\ell)} - \frac{\sum_{k=0}^{N/2} A_k \omega_k^2 \left(\frac{\sin \theta_k^{(\ell)}}{\theta_k^{(\ell)}} \right) \frac{\theta_k^{(\ell)}}{\omega_k}}{\sum_{k=0}^{N/2} A_k \omega_k^2 \left(\frac{\sin \theta_k^{(\ell)}}{\theta_k^{(\ell)}} \right)}. \quad (18)$$

ある初期推定値 t_0 に対して, 上記更新則を反復的に適用することで, 最終的なサブサンプル時間差推定値 \hat{t} を得る. 興味深いことに, (18) の第 2 項は, 周波数によってスケールされた補助変数, すなわち $\theta_k^{(\ell)} / \omega_k$ による重み付き和となっている.

4 シミュレーションによる提案法の評価

提案法の有効性を確認するため, サブサンプル時間差推定の性能を, 4.1 節および 4.2 節では収束性について, 4.3 節では収束に必要な反復回数と雑音に対する頑健性について評価した. 従来法としては GSS に基づく推定 [20] を用いた. 提案法の初期値は, dGCC の最大値に対応する時間差とし, その左右の点に挟まれた区間を GSS の初期探索区間とした. また, GCC の重み W_k はすべての周波数において 1 とした.

4.1 収束性

本節の実験では, -5 から 5 サンプルの範囲の時間差をもつ 2 チャネルの観測信号を用意し, 各マイクロフォンに対して, 信号対雑音比 (signal-to-noise ratio; SNR) が 10 dB となるように白色ガウス雑音を重畳した. ただし, 時間差は周波数領域における位相差として, シミュレーションにより与えた. 時間差推定の対象となる信号 (目的信号) としては, 日本語及び英語の男性及び女性の発話, 計 4 種類の異なる信号を使用した. なお, サンプリング周波数は 16 kHz とした.

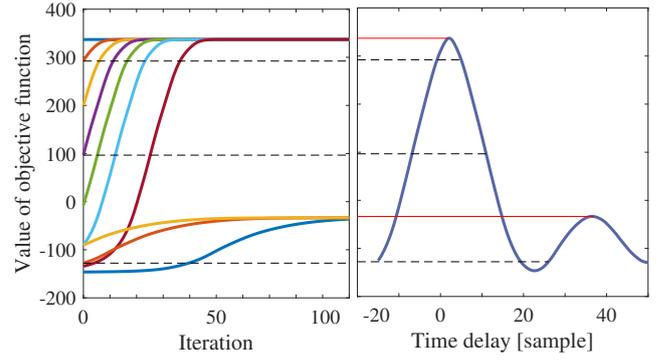


図 2: 左: 提案法の更新による目的関数の値の変化. 各曲線は異なる初期値に対する目的関数の変化を表し, 2 から 29 サンプルまで, 3 サンプル毎に初期値として使用した. 右: 目的関数

図 2 は, 様々な初期値に対して, 提案法による目的関数の収束の様子例である. この例では, 時間差の真値は 2.0966 サンプルである. 提案法の初期値としては, 2 から 29 まで 3 サンプルおきのサンプル値を用いた. まず, 図より提案法の時間差推定値の更新により, 目的関数が単調増加することが確認できる. 従って, 初期値が最大値を含む単峰区間にある場合, 提案法により最適解が必ず求まるといえる. なお, これは, 図 2 においては -20 から 22 サンプル程度の範囲となる. 一方で, そうでない場合, 例えば図中の初期値が 26 または 29 の場合などは必ず局所解が得られることに注意されたい. また, 収束速度について, 初期値が真値に近ければ近いほど, 収束が速くなることも確認できる. これは, 例えば初期値の設定方法として, 2 次関数のフィッティングなどを用いることで, 非常に高速な収束が期待できることを意味する. 一方で, GSS の初期探索区間は, 単峰区間であり, かつ最大値を含む必要がある. GSS の収束を早めるためには, 初期探索区間を狭くする必要があるが, その方法は自明ではなく, 特に雑音環境においては困難な問題である.

4.2 残響環境における収束性

次に, 4.1 節で使用した信号と, インパルス応答の畳込みにより得られた信号を用いて, 残響環境における性能評価を行った. ここで, 鏡像法によるインパルス応答の生成は, Python パッケージである *pyroomacoustics* [26] を利用して行った. 時間差と, 残響時間などのインパルス応答に依存するパラメータを除いた実験条件は 4.1 節と同一である. 実験では, 残響時間 300 ms の環境において, 3 つの異なる音源到来方向 (direction of arrival; DOA) に対応する TDOA の推定を行った. ここで, DOA は 30° , 70° , 135° とし, マイク間隔は 4 cm, 音源とマイク間の距離は

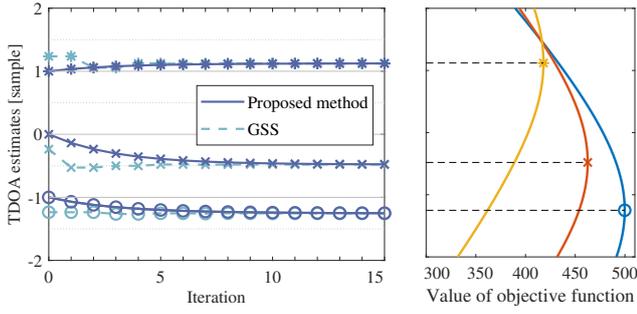


図 3: 左: 残響環境における TDOA 推定値 (SNR: 10 dB). \circ , \times , $*$, は異なる DOA を表し, それぞれ 30° , 70° , 315° である. 右: 各 DOA に対応する目的関数

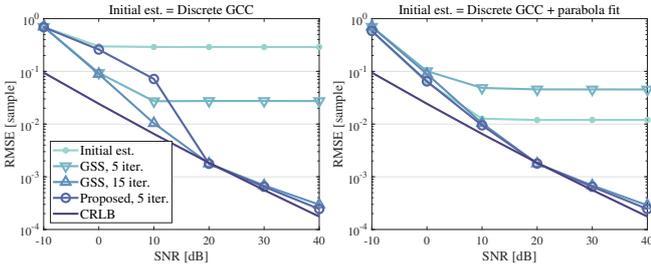


図 4: 雑音環境における TDOA 推定値の RMSE

1.5 m とした. 実験は 4 つの目的信号それぞれに対して行い, その平均を結果として示す.

図 3 に, 提案法と GSS による TDOA 推定値を反復回数毎に示した. 図中の記号 ' \circ ', ' \times ', ' $*$ ' はそれぞれ DOA が 30° , 70° , 135° であったときの結果を示す. 残響環境下における TDOA の推定は困難な問題であるが [27], 図より提案法及び GSS は両者ともに最適解に収束していることがわかる.

4.3 雑音に対する頑健性

この実験では, 雑音環境における提案法と GSS による時間差推定について, 初期値と反復回数の観点から性能の比較を行った. 使用した信号は 4.1 節と同一であり, 時間差は $t \sim U[0, 1]$ の範囲からランダムにシミュレートした. 2 チャンルの観測信号には, それぞれ白色ガウス雑音を SNR が -12 dB から 40 dB になるように重畳した. このとき, \hat{t} に対する CRLB は, 与えられた入力信号と雑音の分散 σ^2 を用いて以下で求まる.

$$\text{Var}\{\hat{t}\} \geq \left\{ 8\pi^2 \sum_{k=0}^{N/2} \frac{(|S_k|^2 / (N\sigma^2))^2}{1 + 2(|S_k|^2 / (N\sigma^2))} (kF_s / N)^2 \right\}^{-1},$$

ここで, S_k は入力信号の DFT の k 番目の点であり, F_s はサンプリング周波数である. 提案法の初期化方法としては, 1. dGCC の最大値に対応する時間差, 2. 2 次関数の

フィッティング [12] による結果, の 2 通りを試した. 実験結果は, 異なる時間差, 異なる雑音に対する 100 通りの試行の結果を平均して示す.

図 4 に, 雑音環境における TDOA 推定値の二乗平均平方根誤差 (root mean square error; RMSE) を示した. 図より, SNR が高い場合, 提案法と GSS の両者共に CRLB を達成することがわかる. しかし, 両者の CRLB に到達するまでの反復回数は異なる. SNR が低い場合, GSS は, 5 回の反復で 15 回の反復時と同等の性能が得られているが, SNR が高い場合は 15 回を超える更新が必要になる. この結果は初期値に依存しない. 一方で提案法は, 2 次関数のフィッティングなどにより良い初期値が得られた場合, 常に 5 回の反復で高い推定精度を得ることができる. 以上より, 雑音環境においても, 提案法により, 高速に CRLB に近い精度の時間差推定を求めることができるといえる.

5 結論

本稿では, 補助関数法に基づく効率的な反復更新により, 連続の時間変数をもつ GCC を最大化することで, サブサンプルの精度の時間差推定を行う新たな手法を提案した. 目的関数, すなわち cGCC は正弦波の和で表すことができ, 従って, それに対する 2 次の補助関数を導出することができる. 補助関数法に基づき, 時間差の推定値と補助変数を交互に更新することで, 目的関数は単調に増加する. 提案する最適化は, 従来の GSS に基づく推定が信号のペアに対してのみ適用可能な一方で, 適切な目的関数を設計することにより, 多チャンネル信号への拡張を考えることができる.

実験では, 提案法が GSS に対して少ない更新回数で収束し, また, CRLB を達成することを示した. 以上より, 提案法がサブサンプル時間差推定法として有効だと言える.

謝辞

本研究は JSPS 科研費 JP16H01735, JP19J20420 の助成を受けたものである.

参考文献

- [1] M. Brandstein and D. Ward, *Microphone arrays: signal processing techniques and applications*, Springer, 2001.
- [2] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 9, pp. 1626–1641, 2016.
- [3] A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," *IEEE Trans.*

- Signal Process.*, vol. 56, no. 5, pp. 1770–1778, Apr. 2008.
- [4] T.-K. Le and N. Ono, “Closed-form and near closed-form solutions for TDOA-based joint source and sensor localization,” *IEEE Trans. Signal Process.*, vol. 65, no. 5, pp. 1207–1221, Dec. 2016.
- [5] S. Miyabe, N. Ono, and S. Makino, “Blind compensation of interchannel sampling frequency mismatch for ad hoc microphone array based on maximum likelihood estimation,” *Signal Processing*, vol. 107, pp. 185–196, Feb. 2015.
- [6] A. J. Coulson, “Maximum likelihood synchronization for OFDM using a pilot symbol: algorithms,” *IEEE J. Sel. Areas Commun.*, vol. 19, no. 12, pp. 2486–2494, Dec. 2001.
- [7] G. Carter, “Time delay estimation for passive sonar signal processing,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 29, no. 3, pp. 463–470, June 1981.
- [8] P. Protiva, J. Mrkvica, and J. Machac, “Estimation of wall parameters from time-delay-only through-wall radar measurements,” *IEEE Trans. Antennas Propag.*, vol. 59, no. 11, pp. 4268–4278, Nov. 2011.
- [9] J. Capon, “Applications of detection and estimation theory to large array seismology,” *Proc. IEEE*, vol. 58, no. 5, pp. 760–770, May 1970.
- [10] C. H. Knapp and G. C. Carter, “The generalized correlation method for estimation of time delay,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 24, no. 4, pp. 320–327, 1976.
- [11] I. J. Tashev, *Sound Capture and Processing, Practical Approaches*. John Wiley & Sons, 2009.
- [12] G. Jacovitti and G. Scarano, “Discrete time techniques for time delay estimation,” *IEEE Trans. Signal Process.*, vol. 41, no. 2, pp. 525–533, 1993.
- [13] L. Zhang and X. Wu, “On the application of cross correlation function to subsample discrete time delay estimation,” *Dig. Signal Process.*, vol. 16, pp. 682–694, 2006.
- [14] F. Viola and W. F. Walker, “Computationally efficient spline-based time delay estimation,” *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 55, no. 9, pp. 2084–2091, 2008.
- [15] B. Qin, H. Zhang, Q. Fu, and Y. Yan, “Subsample time delay estimation via improved GCC PHAT algorithm,” *International Conference on Signal Process.*, pp. 2579–2582, 2008.
- [16] R. Tao, X.-M. Li, Y.-L. Li, and Y. Wang, “Time-delay estimation of chirp signals in the fractional fourier domain,” *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2852–2855, 2009.
- [17] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of signal processing*, Cambridge Univ. Press, 2014.
- [18] C. E. Shannon, “Communication in the presence of noise,” *Proc. IRE*, vol. 37, pp. 10–21, 1949.
- [19] H. Ogawa, “Sampling theory and Isao Someya: A historical note,” *Sampling theory in signal and image process.*, vol. 5, no. 3, pp. 247–256, 2006.
- [20] J. C. Kiefer, “Sequential minimax search for a maximum,” *Proc. of the American Mathematical Society*, vol. 4, no. 3, pp. 502–506, 1953.
- [21] D. R. Hunter and K. Lange, “A tutorial on MM algorithms,” *The Am. Stat.*, vol. 58, no. 1, pp. 30–37, 2004.
- [22] 小野 順貴, “補助関数法による最適化アルゴリズムとその音響信号処理への応用,” *日本音響学会誌*, vol. 68, no. 11, pp. 566–571, 2012.
- [23] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Proc. NIPS*, 2000, pp. 556–562.
- [24] H. Kameoka, T. Nishimoto, and S. Sagayama, “A multipitch analyzer based on harmonic temporal structured clustering,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 3, pp. 982–994, Mar. 2007.
- [25] N. Ono, “Stable and fast update rules for independent vector analysis based on auxiliary function technique,” *Proc. WASPAA*, pp. 189–192, Oct. 2011.
- [26] R. Scheibler, E. Bezzam, and I. Dokmanić, “Pyroomacoustics: A Python package for audio room simulations and array processing algorithms,” in *Proc. IEEE ICASSP*, Calgary, CA, Apr. 2018, pp. 351–355.
- [27] B. Champagne, S. Bedard, and A. Stephenne, “Performance of time-delay estimation in the presence of room reverberation,” *IEEE Trans. Speech Audio Process.*, vol. 4, no. 2, pp. 148–152, Mar. 1996.