

伝播時間差に基づく音源定位の高速反復解法*

小野順貴 (東大院・情報理工)、嵯峨山茂樹 (東大院・情報理工)

1 はじめに

伝播時間差に基づく音源の位置推定法は最も基本的な音源定位手法の1つであり、同様の原理は音源定位に限らず、電磁波を用いたレーダやGPSにも用いられている。非線形な目的関数の最小化のために従来は、Newton法のような一般的最適化手法の適用[1]、近似的な解析解の導出[2, 3]等が試みられてきている。我々は分散型マイクロフォンの枠組みの中でこの種の問題に対する補助関数法の適用を検討してきた[4, 5]。今回、簡便な反復解法を導出し、またその高速化について検討したので報告する。

2 伝播時間差に基づく音源定位の問題設定

ここでは単一音源を仮定し、音源から到来する信号を L 個のマイクロフォンで観測するものとする。音源、マイクロフォンの位置をそれぞれ $s = (x \ y \ z)^t$, $r_m = (u_m \ v_m \ w_m)^t$ ($1 \leq m \leq L$) で表し、マイクロフォン位置は既知と仮定する。ただし t は転置を表す。音源から m 番目、 n 番目 (ただし $1 \leq m \leq L$, $1 \leq n \leq L$ のマイクロフォンまでの距離差は、

$$d_{mn} = c\tau_{mn} = |s - r_m| - |s - r_n| \quad (1)$$

と表される。ただし c は音速、 τ_{mn} は音源から m 番目、 n 番目のマイクロフォンまでの音波の伝播時間差であり、 τ_{mn} は、マイクロフォンで取得した信号間の相互相関関数などから観測量として得ることができる。以下では τ_{mn} と等価な d_{mn} を観測量として扱い、十分な数の距離差観測情報から二乗誤差:

$$J(s) = \sum_{m=1}^L \sum_{n=1}^L (|s - r_m| - |s - r_n| - d_{mn})^2 \quad (2)$$

を最小とする音源位置 s を求める問題を考える。この問題は、距離差に基づく音源定位 (range-differenced-based source localization) とも呼ばれる。式(2)は絶対値関数を含み、位置座標 x, y, z に関しては平方根演算を含む非線形な形であるため、 $\partial J/\partial s = 0$ を解析的に解くことができないことに注意する。

3 補助関数の設計と反復更新式の導出

補助関数法とは、非線形な目的関数 $f(\Theta)$ に対し、 $f(\Theta) = \min_{\tilde{\Theta}} \tilde{f}(\Theta, \tilde{\Theta})$ を満たすような、補助変数 $\tilde{\Theta}$ を含む補助関数 $\tilde{f}(\Theta, \tilde{\Theta})$ を設計し、

$$\Theta \leftarrow \min_{\Theta} \tilde{f}(\Theta, \tilde{\Theta}) \quad (3)$$

$$\tilde{\Theta} \leftarrow \min_{\tilde{\Theta}} \tilde{f}(\Theta, \tilde{\Theta}) \quad (4)$$

を交互に反復的に適用することにより $f(\Theta)$ を単調減少させ、 $f(\Theta)$ を極小とするパラメータ Θ を得る最適化手法である。ここで、式(3)、式(4)が解析的に求まるような、都合のよい補助関数 $\tilde{f}(\Theta, \tilde{\Theta})$ が見つ

かるかどうかは、目的関数 $f(\Theta)$ の具体的な形状に依存する。

紙面の制約上証明は省略するが、我々は式(2)の目的関数に対して

$$\bar{r} + \bar{d}_m \geq 0 \quad (5)$$

の条件下で、

$$\tilde{J}(s, \tilde{\Theta}) = 2L \sum_{m=1}^L |s - (r_m + (\bar{r} + \bar{d}_m)e_m)|^2 + C \quad (6)$$

が補助関数となることを導出した。ただし、

$$\bar{d}_m = \frac{1}{L} \sum_{n=1}^L d_{mn}, \quad (7)$$

は観測量から求まる定数、 C は最適化に関係しない定数、 $\tilde{\Theta} = \{\bar{r}, e_1, \dots, e_L\}$ は補助変数であり、 $\bar{r} = \sum_{m=1}^L |s - r_m|$, $e_m = (s - r_m)/|s - r_m|$ のとき \tilde{J} は最小で、 $\tilde{J}(s, \tilde{\Theta}) = J(s)$ となる。この補助関数を用いると式(3)、式(4)にあたる更新式は、

$$\bar{r} \leftarrow \sum_{m=1}^L |s - r_m| \quad (8)$$

$$e_m \leftarrow \frac{s - r_m}{|s - r_m|} \quad (9)$$

$$s \leftarrow \sum_{m=1}^L (r_m + (\bar{r} + \bar{d}_m)e_m) \quad (10)$$

となる。ここで、 \bar{r} は音源とマイクロフォンの平均距離、 e_m はマイクロフォンから音源へ向かう単位方向ベクトル、 $\bar{r} + \bar{d}_m$ はマイクロフォン m と音源の間の距離を表しており、 $r_m + (\bar{r} + \bar{d}_m)e_m$ ($m = 1, \dots, L$)それぞれが音源位置 s の推定値の意味をもつ。すなわちこの反復更新式は、

- 音源推定位置とマイクロフォン間の距離の平均演算 (r_m の更新)
- 音源推定位置の平均演算 (e_m の更新に基づく s の更新)

という単純な平均演算の繰り返しである。クラスタリング手法としてよく知られている k -means法との類似性から、以下ではこれをr(range)-means法と呼ぶ。r-means法は、1) 行列演算を含まない簡単な反復演算である、2) 収束性が保証されている、3) 勾配法におけるステップサイズのようなチューニングパラメータを持たない、4) 反復更新1回あたりの演算量はマイクロフォン数 L に対し線形オーダーである、といった特長をもつ。

なお式(5)の条件は、マイクロフォン m と音源の間の距離の推定値が非負であることを意味しており、通常は満足される。観測データ d_{mn} が大きな雑音を含み式(5)が満たされない場合においても、軽微な修正により、ほぼ同様の目的関数と更新式を導くことができるが、詳細は紙面の制約のため省略する。

*Fast Iterative Algorithm of Sound Source Localization Based on Time Difference of Arrival by Nobutaka ONO and Shigeki SAGAYAMA (The University of Tokyo)

4 加速法の適用

r-means 法の反復更新式は簡単な演算のみからなり、高速に計算することができるが、収束するまでに必要な反復演算の回数は必ずしも少ないとはいえない。そこで数値計算でよく用いられる加速法の適用を検討した。加速法とは、ある収束する数列 a_n からより収束の早い数列 b_n をつくり出す算法のことであり、数値解析の分野でよく用いられている。補助関数法で反復的に生成される推定値 $s^{(n)}$ は (n は反復回数) は、ある収束値に 1 次収束する性質をもつため、こうした加速法と相性がよいと考えられる。本研究では、最も基本的な加速法である Aitken 加速を繰り返し適用する Steffensen 加速を試みた。

Aitken 加速では、 $s^{(n)}$ が s^* に収束するときの誤差モデルとして $s^{(n)} - s^* = \lambda^n c$ を仮定し (ただし $|\lambda| < 1$, c は定数ベクトル)、これに基づき

$$p^{(n)} = s^{(n)} + \alpha \Delta s^{(n)} \quad (11)$$

により収束を加速した数列 $p^{(n)}$ を生成する。ただし、 $\alpha = \lambda / (1 - \lambda)$, $\Delta s^{(n)} = s^{(n)} - s^{(n-1)}$ である。 $p^{(n)}$ の収束は十分早く $p^{(n)} \simeq p^{(n-1)}$ であると仮定すると、未知の係数 α は

$$\Delta s^{(n)} = -\alpha \Delta^2 s^{(n)} \quad (12)$$

より求まる。ただし $\Delta^2 s^{(n)} = \Delta s^{(n)} - \Delta s^{(n-1)}$ である。 s がベクトルの場合には式 (12) は厳密には満たされないで、最小二乗誤差の意味で式 (12) を満たす α を求め、式 (11) を適用する。また、単に加速数列 $p^{(n)}$ を求めるだけでなく、 $s^{(1)}$, $s^{(2)}$, $s^{(3)}$ から式 (11) を用いて $p^{(3)}$ を求め、これを $s^{(4)}$ とし、引き続き $s^{(5)}$, $s^{(6)}$ を求め、また式 (11) を用いて $p^{(6)}$ を求め、これを $s^{(7)}$ と置き換え、というように、反復的に Aitken 加速を適用する手法が Steffensen 加速である。ただし加速法の適用により、目的関数の単調減少という補助関数法のもつ性質は失われるため、目的関数が減少するかどうかは、式 (11) を適用する度に確認する必要がある点に注意する。

5 実験による検証

r-means 法による音源定位の収束性、収束速度をシミュレーション実験により検証した。5m×4mの空間内にランダムに音源とマイクロフォンを配置し、各マイクロフォン対で伝播時間差から距離差が観測されたと仮定し、r-means 法によって式 (2) を最小とする音源位置を推定した。マイクロフォン数は 5 個とし、簡単のため、音源とマイクロフォンは xy 平面内にあると仮定して 2 次元問題として扱った。観測される伝播時間差には、距離差に換算して 0.01m の標準偏差をもつ白色雑音を与えた。

まず、あるマイクロフォン配置に対し 5 通りの音源位置をランダムに設定し、r-means 法により、それぞれに対し単一音源定位を行った結果を図 1 に示す。初期値は全てマイクロフォンの重心位置とし、反復回数は 300 回とした。推定値の軌跡は必ずしも直線とは限らないが、反復とともに推定値が真値に近づいていくことが確認できる。

次に加速法による効果を確認するために、マイクロフォン 5 個と音源 1 個の配置をランダムに 1000

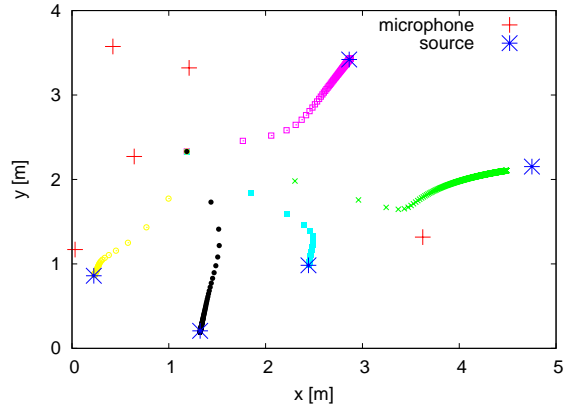


Fig. 1 音源推定位置の軌跡：赤がマイクロフォン位置、青が音源位置、その他の色が、マイクロフォン重心を初期値とした反復による推定位置の軌跡

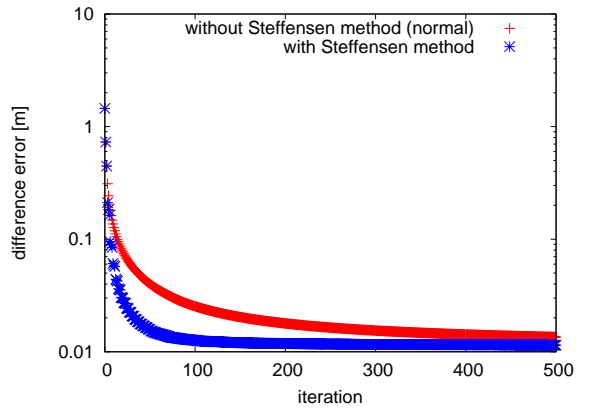


Fig. 2 反復推定に伴う目的関数の変化：赤が r-means 法、青が Steffensen 加速を組み合わせたもの

通り生成し、反復演算に伴う平均的な目的関数の変化 (正確には、距離差 1 つあたりの標準偏差にするため、 $\sqrt{J/(L^2 - L)}$) を、Steffensen 加速を適用しない場合とした場合で比較した。結果をを図 2 に示す。Steffensen 加速は、反復回数が 3 回毎に式 (11) を適用し、目的関数が減少するときに限って、 $p^{(n)}$ を $s^{(n+1)}$ に置き換える、という形で適用した。加速法の適用により急速に目的関数が減少し、ほぼ 100 回程度で収束していることがわかる。なお、Steffensen 加速を併用した場合の反復 100 回に要する演算時間は、通常の PC で 0.1~0.3ms 程度であり、実時間処理にも十分用いることができる計算量と考えられる。

参考文献

- [1] W. H. Foy, IEEE Trans. Aerosp. Electron. Syst., vol. 12, pp. 187-194, Mar. 1976.
- [2] J. Smith and J. Abel, IEEE Trans. ASSP, vol. 35, pp. 1661-1669, Dec. 1987.
- [3] P. Stoica and J. Li, IEEE Sig. Proc. Magazine, pp. 63-69, Nov. 2006.
- [4] 河野, 伊藤, 小野, 嵯峨山, 音講論 (春), pp.703-704, 2009.
- [5] N. Ono, H. Kohno, N. Ito, and S. Sagayama, Proc. WASPAA, Oct. 2009. (to appear)