

# 信頼度重みづけを考慮した非同期録音の ブラインドアライメント\*

長谷川圭介 小野順貴 宮部滋樹 嵯峨山茂樹 (東大院・情報理工)

## 1 はじめに

マイクロホンアレー信号処理における新しい枠組みに分散型マイクロホンアレーがあり、大開口のアレーを簡単に構成できる手法として研究が進められている [1],[2] . 分散型アレーにおいては素子位置が未知であること、録音が非同期であることを前提として信号処理を行う必要があり、これについて我々は非同期の観測信号のみから音源位置、マイク素子位置ならびに各素子の位置を推定する問題についての解法の研究を進めている [4],[5] . 特に実環境においては、各音源に対する時間差検出には誤差が含まれ、また誤差の大きさも一様ではなく、音源やマイクロフォン対ごとに異なることが多い . 例えば音量が小さな音源に対してはSN比の低下から、間隔が大きいマイクロフォン対に対しては、観測信号間の相関の低下から、時間差検出精度が低下する . 本稿ではこのように、時間差検出に誤差が含まれる場合、その誤差分散を既知、もしくは推定可能と仮定し、重み付けを用いた目的関数を最適化することで、推定精度を向上させる手法について述べ、基礎的な実験結果とともに報告する .

## 2 ブラインドアライメント問題

マイクが  $L$  個、音源が  $K$  個存在する状況を考える . 本稿では各マイクの観測信号のみから音源位置  $s_i, i = 1, \dots, K$ , マイク位置  $r_m, m = 1, \dots, L$ , ならびにマイクの録音開始時刻  $t_m, m = 1, \dots, L$  を推定する問題をブラインドアライメント問題とよぶ . 観測信号の持つ情報のうち、われわれはマイク間で観測される信号の到来時間差に着目して問題を解決することを考えた . ここで各マイクの観測から直接求められるのは図 1 に示すように見かけの時間差  $\tau_{imn}, i = 1, \dots, K, m, n = 1, \dots, L$  である . 以下ではこの情報のみを手掛かりにパラメータ  $\Theta := \{s_i, r_m, t_m | i = 1, \dots, K, m = 1, \dots, L\}$  の推定を行う方法について述べる .

## 3 信頼度重みつき目的関数

### 3.1 最小 2 乗誤差基準に基づく目的関数

見かけの時間差  $\tau_{imn}$  を用いて本来の時間差は  $t_m - t_n + \tau_{imn}$  と表わせ (時間差 = 距離の差 ÷ 音速) の関係を使うと、音速を  $c$  [m/s] として

$$\tau_{imn} = \frac{|s_i - r_m| - |s_i - r_n|}{c} - (t_m - t_n) \quad (1)$$

がすべての  $i, m, n$  について成立する . なお、 $|\cdot|$  はベクトルの絶対値を表すものとする . ここで、左辺は観測から直接得られるパラメータであり、右辺は推定されるべきパラメータである . よって、両辺の  $i, m, n$  に関する和の二乗誤差

$$\sum_{i,m,n} \{|s_i - r_m| - |s_i - r_n| - c(t_m - t_n + \tau_{imn})\}^2 \quad (2)$$

を目的関数として設定し、それを最小化するような  $\Theta$  を求めることが考えられる .

また、[5] によれば 3 次元空間内では  $(K - 4)(L - 4) \geq 9$  でなければブラインドアライメント問題は解く

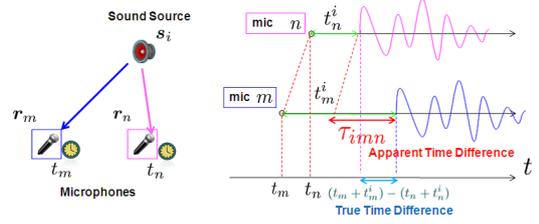


Fig. 1 本来の時間差と見かけの時間差

ことができない . 2 次元空間内では  $(K - 3)(L - 3) \geq 5$  であることが必要である .

### 3.2 観測時間差の信頼度による重みづけ

式 (1) は  $\tau_{imn}$  に誤差がない場合に成立する関係式である . そのため、誤差の大きい場合は目的関数 (2) を最小化するような  $\Theta$  を求めることができたとしても、それが十分良い推定になっているとはいえない . そこで、観測  $\tau_{imn}$  について、各  $i, m, n$  ごとに信頼度  $w_{imn}$  が得られたと仮定し、以下のような目的関数を考える .

$$J(\Theta) := \sum_{i,m,n} \{|s_i - r_m| - |s_i - r_n| - c(t_m - t_n + \tau_{imn})\}^2 w_{imn} \quad (3)$$

この目的関数は各  $i, m, n$  についての二乗誤差について信頼度に応じて目的関数全体に与える影響度を重みづけしていると捉えることができる . 例えば時間差が零平均のガウス分布に従う場合、 $w_{imn}$  を誤差分散の逆数にとれば、(3) の最小化は最尤法に一致することに注意する . この  $J(\Theta)$  を最小化するような  $\Theta$  を求める手法について考える . まず、 $t_m$  については  $\frac{\partial J(\Theta)}{\partial t_m} = 0$  を適当な拘束条件のもとで解くことで解析解を求めることができる . ところが  $s_i, r_m$  についてはベクトルの絶対値演算を含むために直接偏微分 = 0 の式を解くことができない . 本稿では以降、 $s_i, r_m$  の推定法のみについて焦点を合わせる .

## 4 補助関数法による位置パラメータの反復推定

### 4.1 補助関数法の原理

補助関数法とは、ある目的関数  $J(\Theta)$  について、 $J(\Theta) = \min_{\tilde{\Theta}} \tilde{J}(\Theta, \tilde{\Theta})$  を満たす補助関数  $\tilde{J}(\Theta, \tilde{\Theta})$  を設計し、これについて  $\Theta$  についての最小化と  $\tilde{\Theta}$  (補助変数) についての最小化を交互に行うことで  $\tilde{J}$  を単調非増加させることができるという最適化の手法である [3] . なお、性質の良い補助関数  $\tilde{J}$  の設計がどのような  $J$  に対してうまくできるとは限らない .

### 4.2 補助関数の設計

(3) は勾配法でも最小化可能であるが、ここでは [4],[5] の拡張として、補助関数を用いた更新則を導出する .  $d_{imn} := c(t_m - t_n + \tau_{imn})$  とおくと、以下の

\*Blind Alignment for Asynchronous Recording with Reliability Weighting by Keisuke HASEGAWA, Nobutaka ONO, Shigeki MIYABE and Shigeki SAGAYAMA (The University of Tokyo)

形が得られる．

$$J(\Theta) = \sum_{i,m,n} (|s_i - r_m| - |s_i - r_n| - d_{imn})^2 w_{imn} \quad (4)$$

この関数に対し，以下のようにベクトルおよび行列を設定する．

- $W_i$  : 第  $m, n$  成分が  $w_{imn}$  である行列
- $W_i^D := \text{diag}(\sum_n w_{i1n} \dots \sum_n w_{iLn})$
- $b_i := [\sum_{n=1}^L d_{i1n} w_{i1n} \dots \sum_{n=1}^L d_{iLn} w_{iLn}]^T$
- $Q_i : Q_i^T Q_i = W_i$  となる行列

これにより (4) について  $\beta_{im} \geq 0$  かつ  $W_i, i = 1, \dots, K$  が半正定値であるという条件下で以下の補助関数が設計できることを導出した．

$$\tilde{J}(\Theta, \tilde{\Theta}) = 2 \sum_{i,m} w_{im}^D \left| s_i - r_m - \frac{1}{w_{im}^D} \beta_{im} e_{im} \right|^2 + C \quad (5)$$

ここで，

$$\beta_{im} := \sum_{k=1}^{\text{rank}(W_i)} q_{ikm} p_{ik} + b_{im} \quad (6)$$

であり， $q_{ikm}$  は  $Q_i$  の  $k, m$  成分を， $b_{im}$  は  $b_i$  の第  $m$  成分を表す．また， $C$  は最適化に関係しない定数を表し， $\Theta := \{p_{ik}, e_{im}\}$  は補助変数である．さらに， $p_{ik}$  を  $p_i$  の第  $k$  成分であるとすると， $J(\Theta) = \tilde{J}(\Theta, \tilde{\Theta})$  となる条件は  $p_i = Q_i \hat{r}_i$ ,  $\hat{r}_i := [|s_i - r_1| \dots |s_i - r_L|]^T$  および  $e_{im} = (s_i - r_m) / |s_i - r_m|$  である．この補助関数の設計指針は，[6] の r-means 法と呼ばれる音源定位反復解法に基づいている．

#### 4.3 各パラメータの更新式

式 (5) を  $s_i, r_m$  で微分することによりそれぞれの更新式を得ることができる．また，補助変数に関しては上に述べた等号成立条件を満たすように更新すればよい．最終的な更新式は， $W_i^D$  の  $m, m$  成分を  $w_{im}^D$  として以下のようにまとめられる．

$$s_i \leftarrow \frac{\sum_m w_{im}^D (r_m + \frac{1}{w_{im}^D} \beta_{im} e_{im})}{\sum_m w_{im}^D} \quad (7)$$

$$r_m \leftarrow \frac{\sum_i w_{im}^D (s_i - \frac{1}{w_{im}^D} \beta_{im} e_{im})}{\sum_i w_{im}^D} \quad (8)$$

$$p_i \leftarrow Q_i \hat{r}_i, \quad \hat{r}_i = [|s_i - r_1| \dots |s_i - r_L|]^T \quad (9)$$

$$e_{im} \leftarrow \frac{s_i - r_m}{|s_i - r_m|} \quad (10)$$

以上の議論は  $t_m$  についての推定を更新のステップに交えても成立するため，時間原点も含めた推定を行う場合はさらに  $d_{imn} = c(t_m - t_n + \tau_{imn})$  と置き換えて  $\frac{\partial J}{\partial t_m} = 0$  から更新則を得ればよい．

この議論は  $W_i = Q_i^T Q_i$  のように  $W_i$  に対して平方根がとれることを前提としているが，目的関数 (3) について  $W_i$  の対角成分に当たる項は必ず 0 になるため， $W_i$  の対角成分をいくらでも大きくすることができる．よって， $W_i$  の固有値のうち最小のものを  $\lambda_{\min}$  とすると， $W_i \leftarrow W_i + (\lambda_{\min} + \varepsilon)I$  として  $W_i$  を最小の固有値が  $\varepsilon$  になるフルランクの行列に作り替えることができる．この操作により， $W_i = Q_i^T Q_i$  の分

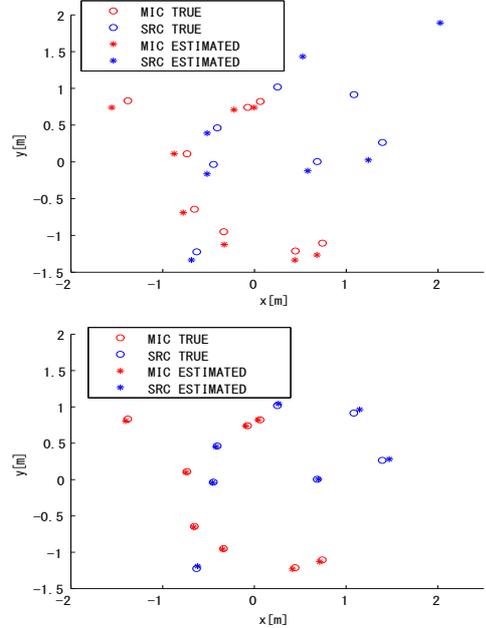


Fig. 2 重みづけを行わない場合 (上) 行った場合 (下) の音源およびマイクの位置推定結果

解可能性については保証される．なお，反復計算に入る前に一度  $Q_i$  を求めてしまえばよいため，固有値分解による計算時間の大幅な増加は起こらない．また，各  $\beta_{im}$  についてはその非負性を反復ステップごとに確認し， $\beta_{im} < 0$  となった場合は初期値を再設定し反復をやり直す必要がある．

## 5 実験

ノイズの乗っている時間差に対する，提案手法の有効性をシミュレーション実験により検証した． $3 \text{ m} \times 3 \text{ m}$  の 2 次元空間上にランダムに 8 つのマイクと 7 つの音源を配置し，その配置 (正解配置) から各音源ごとに各マイクペア間の到来時間差を計算した．観測されるべき時間差は全部で  $7 \cdot \frac{8 \cdot (8-1)}{2} = 196$  個あるが，そのうち 40 個について，標準偏差が距離差に換算して 1m であるような白色雑音を与えた．この場合に提案手法において，重みづけを行わずすべての観測の重みを 1 とした場合と，誤差の乗っている観測の重みのみを 0 とした場合についての位置推定を行った例を図 2 に示す．重みづけを行うことによる推定精度の向上が確認できる．また，同様の誤差の条件のもとで，重みづけがある場合とない場合のそれぞれに対しランダムに 10 通りのマイク 8 個と音源 7 個の配置を生成し，5000 回の反復により求められた位置推定の誤差平均を算出した．重みがすべて 1 である場合は 0.4352 m，重みづけがしてある場合は 0.0454 m であり，精度向上に重みづけが寄与していることが確認できる．

この手法を実環境に適用するためには，観測信号からいかに信頼度を推定するかが重要であり，今後，観測信号間の相互相関関数のピークの大きさ，鋭さ，複数フレームからの時間差検出結果の分散，等を用いた信頼度推定について検討していく予定である．

## 参考文献

- [1] Lienhart, et al., Proc. ICASSP, pp. 840–843, 2003.
- [2] Bertrand and Moonen, Proc. ICASSP, pp. 85–88, 2010.
- [3] Kameoka et al., Proc. ICASSP, pp. 29–32, 2008.
- [4] 長谷川ら，電子情報通信学会全国大会，p.150, 2010.
- [5] Ono et al., Proc. WASPAA, pp.161–164, 2009.
- [6] Ono and Sagayama, Proc. of ICASSP, pp.2718–2721, 2010.