

3次元音場再生システムにおける移動音源定位評価に関する基礎的検討*

☆浅井拓朗, 尾本章 (九大・芸工/JST CREST)

1 はじめに

境界音場制御の原理に基づく3次元音場再生システム(BoSCシステム)が提案されており, 没入型聴覚ディスプレイ装置である音響樽^[1]が開発されている。その定位性能は小林ら^[2]によって主観評価実験が行われ報告されている。これらの実験は静止音源が用いられているが, 音響樽では任意の位置に仮想音源を連続的に配置できることや, 現実の音場では音源が静止していることは稀であることから移動音源定位に関してパーティクルフィルタを用いた評価を試みている。前稿^[3]ではパーティクルフィルタを用いた移動音源定位による逆フィルタの評価の可能性が示され, 本稿では鉛直面でも評価を行い, 主観評価も含めた検討を行った。

2 逆フィルタ

BoSCシステムでは, 再生音場の特性を打ち消すための逆フィルタが再現性能やシステムの安定性に影響する。これまでその設計手法に関して様々な提案がなされている^[4]。

制御点数 M , 二次音源数 N の逆システムを設計する場合, 制御点で得られる信号 $\mathbf{Y}(\omega) = [Y_1(\omega), \dots, Y_M(\omega)]^T$ は以下のように表される。

$$\mathbf{Y}(\omega) = \mathbf{G}(\omega)\mathbf{H}(\omega)\mathbf{X}(\omega) \quad (1)$$

$\mathbf{X}(\omega) = [X_1(\omega), \dots, X_N(\omega)]^T$ は再現したい音源の信号ベクトル, $\mathbf{H}(\omega)$ は N 行 M 列の逆フィルタ, $\mathbf{G}(\omega)$ は制御点 i , 二次音源 j 間の伝達関数 $G_{ij}(\omega)$ を要素とする M 行 N 列の伝達関数行列である。

制御点での信号 $\mathbf{Y}(\omega)$ を再現したい信号 $\mathbf{X}(\omega)$ と一致させるためには, 逆フィルタ $\mathbf{H}(\omega) = \mathbf{G}(\omega)^{-1}$ とする必要がある。 $\mathbf{G}(\omega)$ が非正方行列の場合, 一意に $\mathbf{H}(\omega)$ が定まらないため, 特異値分解や最小ノルム解を用いた擬似逆行列を用いる。また, システムの安定化のため特異値の打ち切りや, 置換えを行ったもの, 正則化パラメータを用いたものがある。今回の実験に用いた逆フィルタを Table 1 に示す。

3 パーティクルフィルタ

音源が移動する場合には, 短時間に音源位置を推定する必要がある。マイクロホンアレイを用いた音源

Table 1 Design principle of inverse filter

invtype	説明
normal	通常の特異値分解を用いた Moore-Penrose 型逆行列
heuristic	正則化パラメータを用いた二乗誤差とノルムを最小にした逆行列
optimum	正則化パラメータを用いた二乗誤差とノルムを最小にした逆行列
suppress	最大特異値の (1/30) 以下の特異値をすべて置換した Moore-Penrose 型逆行列
trunc	打ち切り特異値分解を用いた Moore-Penrose 型逆行列

定位推定はビームフォーミングや MUSIC 法などの方法があるが, 周波数やマイクロホン配置に依存したり, 推定に一定の統計量が必要となる。今回は少ない統計量で推定することが移動音源では重要であることを鑑みて, パーティクルフィルタ^[5]を用いることで過去のマイクロホンからの観測信号 \mathbf{z}_t を基にした内部状態である角度 x_t の事後確率計算を効率的に行なった。

現在の観測信号のブロック t をフレーム数 J で分割し, 各マイクロホンでフレームごとの短時間フーリエ変換をまとめた行列が $\mathbf{z}(t, j)$ である。これを J 個のフレームにまとめたものを入力ベクトル \mathbf{z}_t とする。

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{z}(t, 1) \\ \vdots \\ \mathbf{z}(t, J) \end{bmatrix}^T = \begin{bmatrix} \mathbf{Z}_1(t, 1) & \cdots & \mathbf{Z}_1(t, J) \\ \vdots & \ddots & \vdots \\ \mathbf{Z}_M(t, 1) & \cdots & \mathbf{Z}_M(t, J) \end{bmatrix}$$

パーティクルフィルタアルゴリズムでは推定角度である内部状態 x_t と重み w_t の2つのパラメータを持つ粒子が後述する尤度関数 $p(\mathbf{z}_t | \mathbf{x}_t)$ から求められる重みに基づいて増殖・淘汰する手続きを繰り返すことで, 次の時刻 $t+1$ の事後確率を求める。

事後確率密度関数 $\hat{p}(\mathbf{x}_k | \mathbf{Z}_{1:k})$ はモンテカルロ法, 重点サンプリングにより下式に従って近似的に計算される。

$$\hat{p}(\mathbf{x}_k | \mathbf{Z}_{1:k}) = \sum_{i=1}^{N_p} \tilde{w}_k^{(i)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}) \quad (2)$$

*Study on the evaluation of moving sound source localization for 3-D sound field reproduction systems by ASAI, Takuro, OMOTO, Akira (Faculty of Design, Kyushu University, JST/CREST)

$\tilde{w}_k^{(i)}$ はインポートンス重みであり、 $\tilde{w}_k^{(i)} \propto \tilde{w}_k^{(i-1)} p(\mathbf{y}_k | \mathbf{x}_k^{(i)})$ という関係から尤度関数 $p(z_k | \mathbf{x}_k^{(i)})$ と一時刻前の重み $w_{k-1}^{(i)}$ を用いて再帰的に求められる。

3.1 尤度関数

確率分布関数である尤度関数は観測モデルと観測信号 $z(t, j)$ との適合度を表している。各マイクロホンの観測信号を短時間周波数変換した行列 z_t の空間相関行列を尤度関数の算出に用いた。空間相関行列の各要素は振幅、位相の情報を含む各マイクロホン間の分散関係を表している。この観測信号の空間相関行列 C_z に対して、音源位置とマイクロホン配置の伝達関数を用いた観測モデルの空間相関行列 K_z を用いることで、尤度関数 $p(z_t | \mathbf{x}_t)$ を表す [4]。観測信号 z_t の空間相関行列は以下のように表される。

$$C_z = \frac{1}{J} \sum_{j=1}^J z(t, j) z^H(t, j) \quad (3)$$

次に M 個のマイクロホン、 N 個の音源、背景雑音があると仮定した観測モデル z_{model} を考える。

$$z_{model} = \sum_{i=1}^N \mathbf{a}_i \mathbf{S}_i + \mathbf{n} = \mathbf{A} \mathbf{s} + \mathbf{n} \quad (4)$$

$\mathbf{a}_i = [A_{1,i} e^{-j\omega\tau_{1,i}}, \dots, A_{M,i} e^{-j\omega\tau_{M,i}}]$ であり、音源からマイクロホンまでの伝達関数で表す。 \mathbf{s} 及び \mathbf{n} はそれぞれ音源、背景雑音のスペクトルから成るベクトルである。相関行列を求めると、モデル相関行列は以下のように表される。

$$K_z = \mathbf{A}(\boldsymbol{\theta}_j) \boldsymbol{\Gamma}_j \mathbf{A}^H(\boldsymbol{\theta}_j) + \sigma_n^2 \mathbf{I} \quad (5)$$

音源信号 \mathbf{s} 、背景雑音 \mathbf{n} が正規分布を仮定すると、尤度 $p(z_t | \mathbf{x}_t)$ は以下のように表される。

$$p(z_t | \mathbf{x}_t) = \psi_x \exp\left(-\frac{1}{2} \text{tr}[C_z K_z^{-1}]\right) \quad (6)$$

今回は自由空間を仮定して観測信号をモデル化することで、現在の観測信号に対する尤度関数を算出している。

3.2 手続き

具体的なパーティクルフィルタの手続きについて述べる。パーティクルの変化を表したものを Fig.1 に示す。

1. **サンプリング** 事前遷移確率密度 $p(\mathbf{x}_t | \mathbf{x}_{t-1})$ から粒子を発生させる。各粒子には内部状態 \mathbf{x}_k と重み N_p^{-1} が与えられる。
2. **重み** 新たな観測値 z_t と予測値 $x_k^{(i)}$ から尤度関数 $p(z_k | x_k^{(i)})$ を求め、各粒子の内部状態 $\tilde{x}_k^{(i)}$ から式 (2) における重み $\tilde{w}_k^{(i)}$ の更新を行なう。

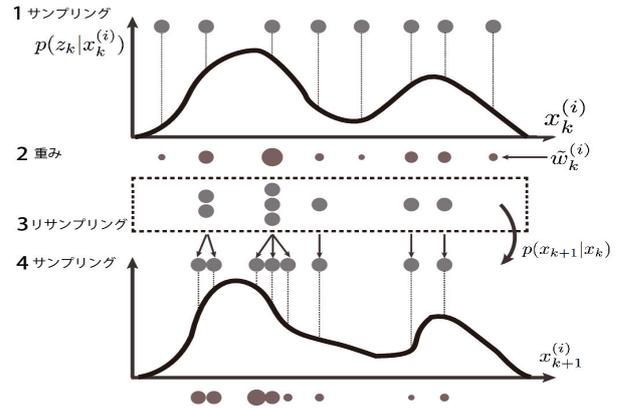


Fig. 1 Concept of particle filter

3. **リサンプリング** 各粒子の重みに従って粒子を増殖または淘汰し、均一な重み N_p^{-1} を持つ $\tilde{x}_k^{(j)}$ を発生させる。この分布が事後確率密度 $p(\mathbf{x}_k | \mathbf{Z}_{1:k})$ の近似 $\hat{p}(\mathbf{x}_k | \mathbf{Z}_{1:k})$ である。

4. **サンプリング** 次に状態遷移を行ない、均一な重みの粒子を分布させる。今回は $N(x_t^{(i)}, \sigma^2 = 2)$ の正規分布を遷移確率とした。これにより縮退が回避される。

以上の手続きを繰り返すことで事後確率を求めることで、音源方向に粒子 \mathbf{x}_t が集まり、音源位置が推定される。これらの手続きの数学的な結びつきは参考文献を参照頂きたい [6]。

今回用いたパーティクルフィルタのパラメータ設定を Table 2 に示す。

Table 2 Parameter of particle filter

パラメータ	値
方向の初期値	一様分布 [-180, +180]
式 (5) における σ	0.05
粒子数 N_p	500

4 実験

原音場を無響室とし、80 ch のフラーレン型マイクロホンをターンテーブルに乗せ、スピーカを固定した状態で水平・仰角方向について以下の 2 パターンの測定を行なった。

1. マイクロホンから半径 1m の円周上を音源が 24 [s/round] の速さで 1 回転する。
2. マイクロホンから半径 1m の円周上を音源が下図 (Fig.2) のように回転する。ただし、初期値を $-30^\circ, 90^\circ, 210^\circ$ の 3ヶ所から開始した。

実線の区間は音源が駆動していることを示している。使用した音源は純音 (125, 250, 500, 1000, 2000, 4000, 8000 Hz)、帯域雑音 (中心周波数 125, 250, 500, 1000, 2000, 4000, 8000 Hz) を用いた。パーティクルフィルタによる音源推定に用いたマイクロホンは前

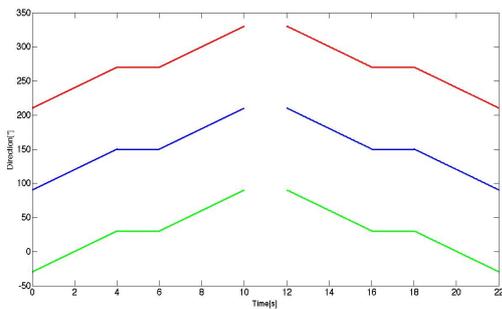


Fig. 2 Assumed rotation pattern

稿では同一平面上に存在するもののみであったが、今回は 80ch すべてを使用した。

原音場と比較を行う再現音場には 80ch マイクロホンで測定した音源を先述の設計方法が異なる 5 つの逆フィルタを畳込んだ上で、音響樽内で再生した音場とした。この樽内の再現音場においても原音場と同様に音源推定のための録音を行なった。

5 実験結果

5.1 原音場

原音場ではすべての帯域において正しい音源位置推定が可能であった。低域では尤度関数が音源推定に十分に先鋭にならない部分もみられたが、音源信号によらず水平方向、垂直方向共に一致した良好な推定であった。また、推定に用いたマイクロホンの数が増加したことで、これまでの誤推定がなくなりより正確な推定が可能であることが確認された。原音場において純音 1000 Hz で初期値 90° から Fig. 2 のように回転させた試行を例に同一平面上マイクのみを使用した場合 (Fig. 3) と全てのマイクをした場合 (Fig. 4) の尤度関数を示す。

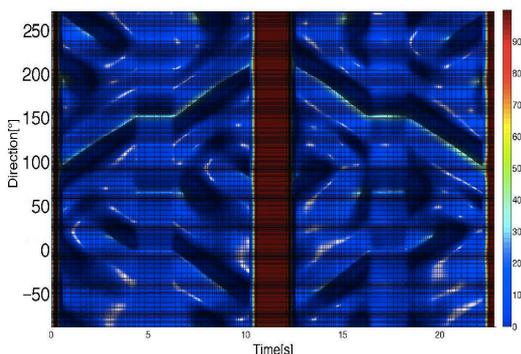


Fig. 3 Likelihood function (Primary field, Puretone 1000 Hz, Horizontal direction, using microphones in the same plane)

5.2 再現音場

再現音場では、鉛直方向で逆フィルタごとの再現精度の差が現れた。例として純音 1000Hz, 90° から回転したパターンにおける原音場と再現音場での逆フィ

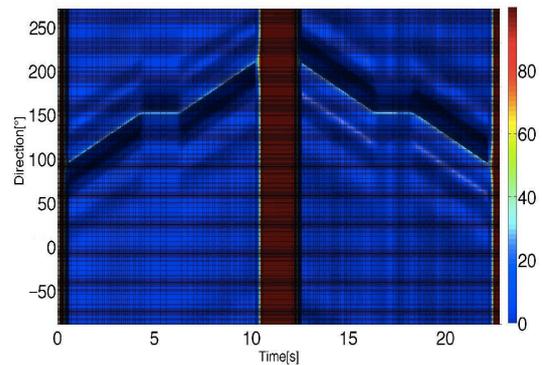


Fig. 4 Likelihood function (primary field, puretone 1000 Hz, horizontal direction, using 80ch microphones)

ルタ別のパーティクルの中央値を以下に示す。ただし、特異値解を利用し安定化を行わない逆フィルタ (normal) は SN 比が十分ではなく、推定ができなかったため省略する。

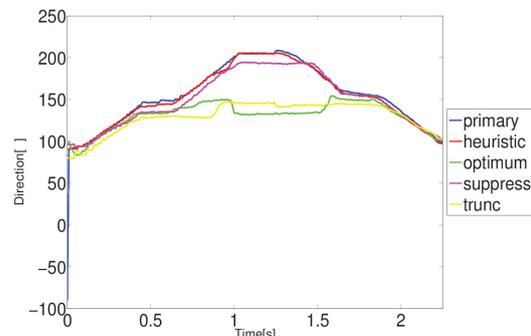


Fig. 5 Sound source trajectory (pure tone 1000 Hz, vertical direction, initial angle 90°)

heuristic, suppress では、原音場 (primary) とほぼ同様の音場が再現されている。特に heuristic では原音場との差が常に 5° 以内となっており、精度の高い提示が可能であることが確認できた。それに対し、optimum, trunc では途中から尤度関数が大きく広がり、原音場のような位置に音源が推定されていない。このことから、最も波面が再現されている逆フィルタは heuristic であることが分かる。また、水平方向ではすべての逆フィルタで概ね原音場と一致した推定がなされた。

6 主観評価アンケート

物理的な測定結果と主観評価との対応がなされているかを検討するために簡単なアンケートを行った。

6.1 刺激

刺激音は全て無響室にて実測で作成した。音源信号は 1kHz を中心周波数とした 1/1 オクターブバンドノイズを原音とし、水平面内で半径 1 m の円周上 360° において、真正面から右回りに音源を等速 (24 s/round) で移動させた。収録した音源に normal 以外の 4 つ逆フィルタを畳込んだものを刺激とした。刺

激の音圧レベルは、全ての条件において被験者の頭部中心位置で 70 dB(A) となるように調整した。

6.2 手続き

聴覚健全な成人男女 12 名にどの位置から音源が提示されたかを図示で回答させた。また、刺激ごとの印象についても同時にアンケート (5 段階評価) を行った。刺激提示中は頭部の動きを許容し、提示する順番はランダムとした。アンケート項目は音源移動の自然さ (なめらかさ)、音像の大きさ、ノイズの大きさ、総合的な知覚しやすさの 5 項目である。

6.3 結果

Fig. 6 に回答した音源の軌跡が原音場と完全一致していたものを正答とした全実験参加者の平均正答率、Fig. 7 に全実験参加者のアンケート結果の平均値を示す。

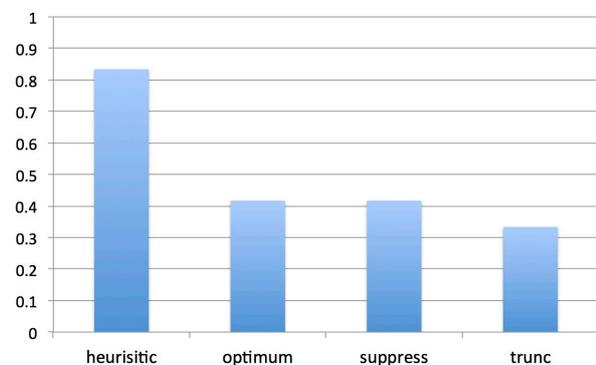


Fig. 6 Percentage of correct answers

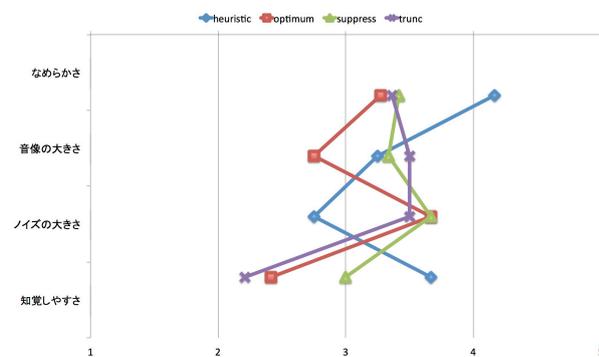


Fig. 7 Questionnaire results

heuristic は 80%以上の正答率があり、他に比べ顕著に高いことが分かる。それに対し、その他のフィルタでは 40%ほどであり大きな差異は見られなかった。アンケート結果の特徴的な傾向としては heuristic が他に比べ「なめらか」で「ノイズが少ない」という点である。

6.4 考察

正答率では heuristic の空間再現性能が高く知覚されやすいという点はパーティクルフィルタによる物理測定による評価と一致したが、suppress の正答率が低い点は相違がある。今回は完全に回答が一致した時

のみを正答としており、部分的な一致は判別が難しいことから全て不正解としたためと考えられるが、アンケートの「知覚しやすさ」では suppress が optimum, trunc よりも知覚しやすいと判断されている。また、部分的な正答も多い傾向であったことから、パーティクルフィルタによる物理測定による結果と主観評価アンケートとの傾向はよく一致することと言える。

7 まとめ

パーティクルフィルタを用いた移動音源定位推定により、再現音場の空間再現性の評価を行った。推定に用いるマイクロホン数を同一平面上のマイクロホンから 80ch に増加させることで、誤推定を低減し高い精度での推定が可能となった。空間再現性能が高い逆フィルタは順に heuristic, suppress であった。normal, optimum, trunc に関しては誤った位置に音源が推定されたため、波面が正しく再現されていない可能性がある。

移動音源に関する主観評価アンケートでも、物理測定による結果と同様の傾向がみられた。このことより移動音源推定による逆フィルタの評価は有効である可能性が示された。今後は推定角の 3 次元への拡張や、厳密な主観評価実験による検証が必要となる。

参考文献

- [1] Omoto *et al.*, “Sound field reproduction and sharing system based on the boundary surface control principle,” *Acoust. Sci. Tech*, Vol. 36, 1-11, 2015
- [2] KOBAYASHI *et al.*, “Subjective evaluation of a virtual acoustic system: Trials with three-dimensional sound field reproduced by the ‘Sound Cask,’” *Proc ICA*, 2013
- [3] 浅井 他, “3次元音場再生システムにおける移動音源定位評価に関する基礎的検討,” *日本音響学会講演論文集*, 2-P-31, 2015
- [4] 高以良 他, “多チャンネル音場再生システムにおける伝達関数の周波数特性を用いた逆フィルタ安定化法の提案,” *日本音響学会講演論文集*, 2-P-33, 2015
- [5] 浅野 太, 麻生 英樹, “パーティクルフィルタを用いた移動音源の追跡技術,” *日本音響学会誌*, 61 巻 12 号, 720-727, 2005.
- [6] 浅野 太, “音のアレイ信号処理 -音源の定位・追跡と分離-”, コロナ社, 2011.