

# BoSC マイクロホンを用いた仮想音源推定と音場再現への応用 -近接多点法におけるロバスト推定の適用による精度向上- \*

☆木村和基, 渡邊祐子, 伊勢史郎 (東京電機大学)

## 1 はじめに

没入型聴覚ディスプレイ「音響樽」を用いることにより室内音響空間を高い精度で3D音場として再現することが可能となり, 室内の音響伝達関数をリアルタイムで畳み込むシステムを用いて楽器演奏をバーチャル空間で行うことも可能となった [1]. バーチャル空間に生成する室内の音響特性は実際のコンサートホールなどの音場でインパルス応答を測定することにより得られるが, 音場の反射音構造を変更することができれば, より多様なバーチャル空間を提供することができる. ところで室内音響空間における反射音の空間構造を把握する手法として近接4点法 [2], および多チャンネルマイクアレイを用いる方法すなわち「近接多点法」 [3] が提案されている. 後者は多数のマイクロホンを用いることにより高い精度で反射音構造を推定することが狙いだが, 外れ値の影響を受けやすく正確な反射音構造の情報が得られない. そこで本稿では最小二乗法にロバスト推定を導入することにより外れ値の影響を抑制する計算手法を提案し, その有効性を示す.

## 2 ロバスト推定を適用した近接多点法

### 2.1 仮想音源位置推定

ある仮想音源 (位置  $\mathbf{S} = [s_1, s_2, s_3]$ ) から  $N$  チャンネルのマイクロホンアレイの  $j$  番目のマイクロホン (位置  $\mathbf{M} = [m_{j1}, m_{j2}, m_{j3}], j = 1, 2, \dots, N$ ) までの距離を  $r_j$  とすると次式が成り立つ.

$$(s_1 - m_{j1})^2 + (s_2 - m_{j2})^2 + (s_3 - m_{j3})^2 = r_j^2 \quad (1)$$

式 (1) を一次連立方程式として表すと以下のようになる.

$$\mathbf{A}\mathbf{S} = \mathbf{B} \quad (2)$$

ただし,  $\mathbf{A} = [a_1, a_2, \dots, a_N]^T$ ,  $a_j = (m_j - m_i)$ ,  
 $\mathbf{B} = [b_1, b_2, \dots, b_N]^T$ ,

$$b_j = \frac{1}{2}(r_i^2 - r_j^2 + m_{i1}^2 - m_{j1}^2 + m_{i2}^2 - m_{j2}^2 + m_{i3}^2 - m_{j3}^2),$$

$$(i, j = 1, 2, \dots, N, i \neq j).$$

式 (2) に最小二乗法を適用すると, 仮想音源の座標  $\mathbf{S}$  は以下のように表せる.

$$\mathbf{S} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{B} \quad (3)$$

この解の残差ベクトル  $\mathbf{E}$  と  $\mathbf{A}$  のハット行列  $\mathbf{H}$  は,

$$\mathbf{H} = \mathbf{A} (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \quad (4)$$

$$\begin{cases} \mathbf{E} = \mathbf{B} - \mathbf{A}\mathbf{S} \\ \mathbf{E} = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{N-1}]^T \end{cases} \quad (5)$$

である.  $\mathbf{H} = [h_{nm}]$  が  $n = m$  の時の成分を  $h_{jj}$  と呼ぶ. ここで,  $\mathbf{E}$  の  $j$  番目の要素を  $\varepsilon_j$  とすると, 調整済み残差  $\varepsilon_{adjj}$  は式 6 で表せる.

$$\varepsilon_{adjj} = \frac{\varepsilon_j}{\sqrt{1 - h_{jj}}} \quad (6)$$

また, 式 6 を標準化すると,

$$u_j = \frac{\varepsilon_{adjj}}{K\sigma} \quad (7)$$

と表せる. ただし,

$K = 4.685, \sigma = MAD/\sqrt{0.6745}$ ,  $j = 1, 2, \dots, N-1$  とする. ここで  $MAD$  は  $\varepsilon_{adjj}$  の中央絶対偏差である.

式 7 を用いて, 重み値行列  $\mathbf{W}$  を以下に示す.

$$\mathbf{W} = \text{diag}(w_1, w_2, \dots, w_{N-1}) \quad (8)$$

ただし,

$$w_j = \begin{cases} (1 - u_j^2)^2 & |u_j| < 1 \\ 0 & |u_j| \geq 1 \end{cases} \quad (9)$$

であり,  $j = 1, 2, \dots, N-1$  とする.

よって, 重み値行列 8 により外れ値の影響を抑えた重み付き最小二乗法の関係式を以下の式 10 に表す.

$$\mathbf{S}_{\text{weight}} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{B} \quad (10)$$

直前の重み付き最小二乗法で求めた  $\mathbf{S}_{\text{weight}}$  を式 5 の  $\mathbf{S}$  に代入し, 重み付き最小二乗法を解き直し, 反復の上限回数に達するか, 反復の必要がなくなるまで行い, 最終的な  $\mathbf{S}_{\text{weight}}$  の値を音源座標  $\mathbf{S}$  とする.

## 3 最小二乗法とロバスト推定を適用した最小二乗法の仮想音源位置推定精度の比較

### 3.1 実音場における実験

オーディオルーム (東京電機大学千葉 NT キャンパス視聴覚実験室, 幅 5.2 m, 奥行 6.6 m, 高さ 2.7 m) の中心に BoSC マイクロホン (高さ約 1.2 m), 中心から距離 2 m, 見開き左 20° にスピーカ (高さ約 1.2 m) を設置した .80ch の BoSC マイクロホンとその中心に設置した基準マイクロホンを合わせて計 81ch でインパルス応答を測定し, 以下の手順で仮想音源位置推定を行った.

1. 各チャンネルにおける直接音と信号全長の相関から仮想音源候補を導出
2. 基準チャンネルの仮想音源信号と他チャンネルの信号の相関から到達時間差を算出
3. それぞれの推定手法により仮想音源座標を推定 (ロバスト推定の場合, 外れ値チャンネルの到達時間を推定した仮想音源座標により修正)
4. 基準マイクまでの仮想音源到達距離を所望距離とし, 推定距離との誤差が 0.3m 以内であれば仮想音源と判定する
5. 到達時間を用いて同期加算平均し, 各チャンネル信号から減算する.
6. 前回のフレームで推定した仮想音源の位置座標との距離が 1 m 以内であれば同一仮想音源として合成し, エネルギーが大きい方の座標を採用
7. 次のフレームに進み仮想音源抽出を繰り返す

\* Estimation of virtual sound source by using BoSC microphone and its application to sound field reproduction. -Improvement of the localization accuracy by applying robust estimation to closely multi point method- by KIMURA, Kazuki, WATANABE Yuko, and ISE, Shiro (Tokyo Denki University).

上記アルゴリズムで最小二乗法を用いた場合の仮想音源分布を Fig.1 に示す。マイクロホン位置は [0m, 0m], 長方形の実線は室形状である。円は仮想音源であり円の中心が仮想音源の推定位置, 半径がエネルギーの大きさを示す。次にロバスト推定を適用した場合の仮想音源分布を Fig.2 に示す。最小二乗法のみの場合には直接音と上下方向の一次反射音の3つだったが, ロバスト推定を用いた最小二乗法では7つの仮想音源が抽出され, 二次反射音を含めた側方反射音も抽出されていることがわかる。

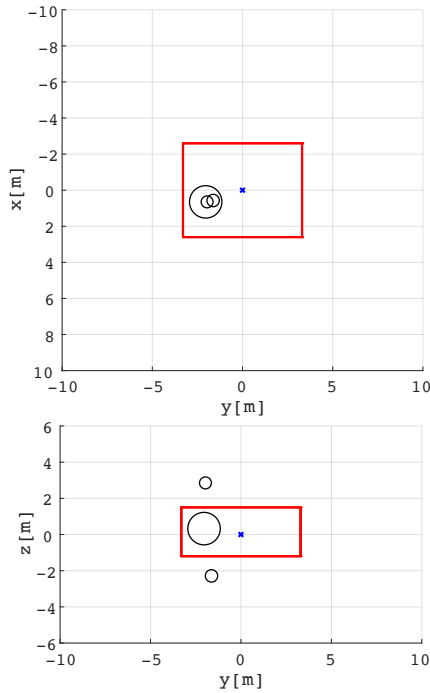


Fig. 1 最小二乗法推定結果

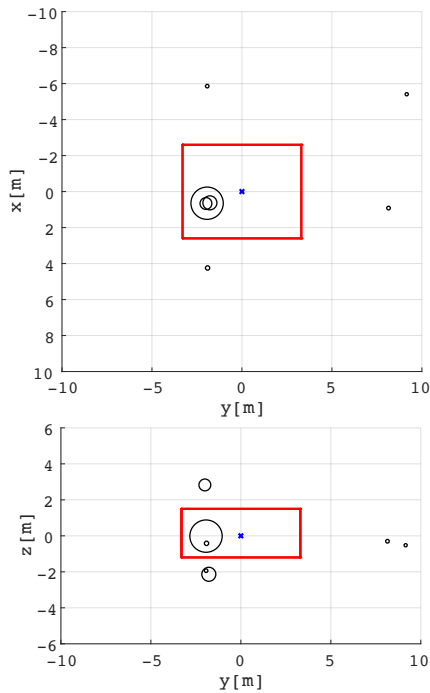


Fig. 2 ロバスト推定による最小二乗法推定結果

#### 4 音場再現への応用の可能性

仮想音源の原信号の波形を算出するために全チャンネル信号を仮想音源毎に切り出し同期加算することにより求めた。例として基準マイクロホンのインパルス応答を Fig.3, 抽出した仮想音源信号波形を Fig.4, 残差信号を Fig.5 に示す。

仮想音源毎に振幅・到達時間・周波数特性などを変更して残差に加算すれば異なる室内音響特性をもつインパルス応答を再合成することが可能となり, さらに再合成したインパルス応答を BoSC システム [4] 等の音場再現システムで用いることで多様な室内音響特性を実現することが可能となる。

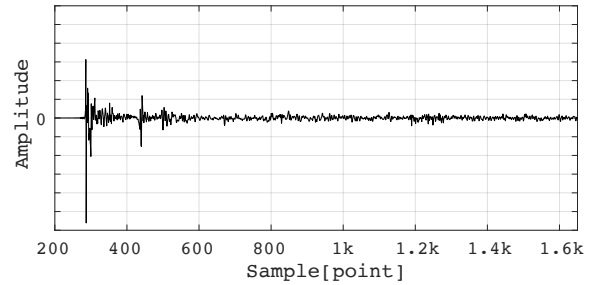


Fig. 3 インパルス応答

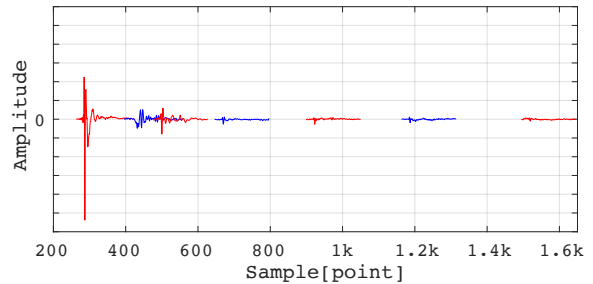


Fig. 4 仮想音源推定信号

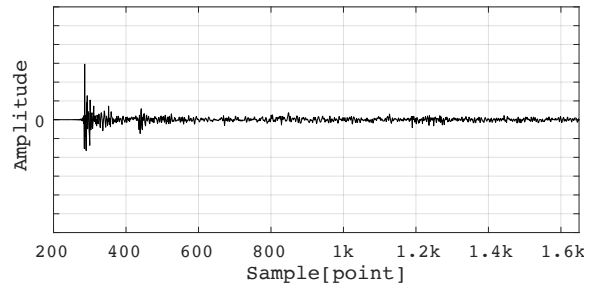


Fig. 5 残差信号

#### 5 むすび

近接多点法の精度向上手法として, ロバスト推定を用いた新手法を提案し, 実音場での測定実験において室の形状に即した仮想音源推定が可能であることを確認した。

#### 参考文献

- [1] 伊勢史郎, 音講論 (秋), 892-894, 2014.
- [2] Yoshio Yamasaki *et al.*, J.ASJ(E)10, 2, 1989.
- [3] 西村昌浩他, 音講論 (春), 789-790, 2012.
- [4] 伊勢史郎, 音学誌, 53(9), 706-713, 1997.