没入型聴覚ディスプレイ装置”音響樽”における逆システム設計法の検討

○伊勢史郎（東京電機大／JST, CREST）

1　はじめに

境界音場制御の原理に基づく 3 次元音場再現システムを用いた没入型聴覚ディスプレイ装置が実用レベルに近づきつつある。現在，我々は 80ch のマイクロホンアレイによって収録した音場を再生するために，96 個のスピーカを壁面に取り付けた樽型の音場再生室“音響樽”を開発しているが，その逆システムの設計法が性能を左右する。80ch 入力 96ch 出力となるような大規模な逆システムを合理的に計算する方法は提案されていない。そこで本稿では最適な正則化パラメータをオクターブバンド毎に求めて，逆システムを求める方法を提案する。

2　理論

2.1 BoSC 原理による音場再現

図 1 のように，ある空間に境界面Sの領域Vの音場（原音場）と，それとは別の空間に領域Vと合同となる領域V'の音場（再生音場）を想定する。境界音場制御の原理に基づけば，原音場においてある領域を囲む境界面 S 上の音圧と粒子速度(音圧勾配) を計測し，それらが再生音場の境界面S' 上において相対的に同じ位置で再生されたとき，原音場における領域V内の音場は再生音場の領域V' 内に完全に再現される[1]。ここで原音場における境界面 S 上の音圧および粒子速度はM個のマイクロホンで計測した音圧信号により再現可能と仮定し，その位置座標を$\boldsymbol{q_j}$ $(j = 1 \cdots M)$とする。同様に再生音場に設置するマイクロホンの位置座標を$\boldsymbol{q'_j}$とする。原音場でのマイクロホン出力信号から得られる逆システムの入力信号ベクトルを$[X_j](\in \mathbb{C}^{1\times M})$，再生音場における$L$個のスピーカからマイクロホンへの伝達関数マトリクスを$[G_{ij}](\in \mathbb{C}^{L\times M})$，逆システムの伝達関数マトリクスを$[H_{ji}](\in \mathbb{C}^{M\times L})$，再生音場におけるマイクロホンからの出力信号ベクトルを$[Y_j](\in \mathbb{C}^{1\times M})$とすると次式が成り立つ。

$$[Y_j] = [X_j][H_{ji}][G_{ij}] \qquad (1)$$

ただし，$X_j = P(\boldsymbol{q_j})$, $Y_j = P(\boldsymbol{q'_j})$である。ここで式(1)が成立するためには$[Y_j] = [X_j]$となる$[H_{ji}]$を求めればよい。

2.2 逆システムの計算方法

マイクロホンよりもスピーカの数が多い場合は逆システムを時間領域で求めることにより FIR システムとして設計できるが[2]，本システムのように多チャンネルシステムの逆システムを時間領域で求めることは困難である。周波数領域で求める場合には式(1)を解くことにより逆システムを求めることができる。しかし，M<L の場合には逆行列を一意に求めることができないため，最小ノルム解により求める方法が提案されている[3]。最小ノルム解を与えるムーアペンローズ（MP）一般逆行列は二次音源からの出力を最小化するため，比較的安定した逆システムの設計が期待でき

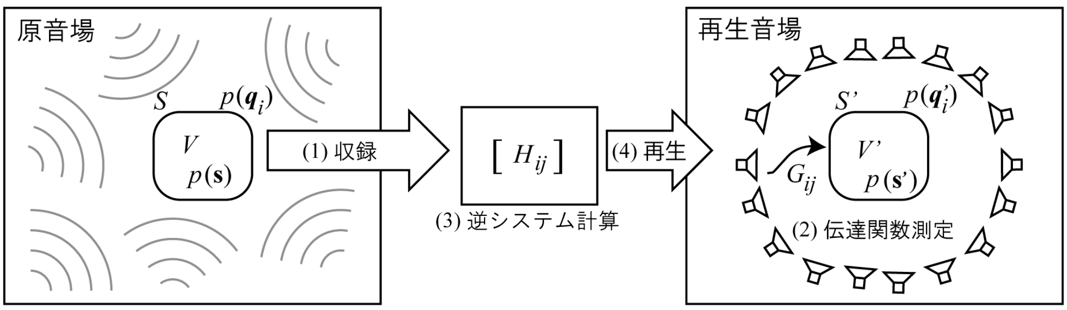

図 1　BoSC 原理による音場再現

* Study on designing an inverse system for the immersive auditory display system “Sound Cask”, ISE, Shiro (Tokyo Denki University/CREST, JST).

る。しかし，チャンネル数が増えれば条件数が過度に大きくなる可能性が増え，想定した時間範囲で収束する逆システムを設計することが困難となる。そこで安定した逆システムを設計する以下の方法が提案されている。

2.2.1. 正則化パラメータ法

次式のように一般化逆行列を求める際に正則化パラメータβを乗じた単位行列を加える方法が提案されている[4]。

$$[H_{ji}] = \left([G_{ij}]^\dagger [G_{ij}] + \beta I_M\right)^{-1} [G_{ij}]^\dagger \qquad (2)$$

ただし$[\cdot]^\dagger$は行列の共役転置，βは正則化パラメータ，I_MはM次元単位行列である。すなわち,正則化パラメータを加えることにより，行列の対角成分が大きくなるためその逆行列から安定した FIR フィルタを設計することが可能となる。

2.2.2. 打ち切り特異値分解による方法

$[G_{ij}]'(\in \mathbb{C}^{M\times L})$を特異値分解すると次のようになる。

$$[G_{ij}]' = [U_{ij}][D_j \quad O_{ij}][V_{ij}]'$$
$$[D_j] = diag[d_1 \cdots d_M]$$

ただし，$[d_j](\in \mathbb{C}^{M\times M})$は M 個の特異値$d_j(j = 1 \cdots M,\ d_j > d_{j+1})$を対角要素とする行列，$[O_{ij}](\in \mathbb{C}^{M\times L-M})$は零行列，$[U_{ij}](\in \mathbb{C}^{M\times M})$，$[V_{ij}]'(\in \mathbb{C}^{L\times L})$は各特異値に対応する固有ベクトルを列ベクトルにもつユニタリ行列で、$[\,]'$は行列の転置である。このとき$[G_{ij}]'$の MP 一般逆行列$[G_{ij}]^-$は次のように表される。

$$[G_{ij}]^- = [V_{ij}]' \begin{bmatrix} E_j \\ O_{ij} \end{bmatrix} [U_{ij}]^H$$

ただし，$[E_j](\in \mathbb{C}^{M\times M})$は特異値の逆数$1/d_j$を対角要素とする行列である。打ち切り特異値分解による方法とは特異値の逆数が最も大きいものから順に 0 に置き換えることにより，$[G_{ij}]^-$が過度に大きくなることを防ぐ方法である。これは線形独立性が小さい音源からの出力を，その周波数成分に関しては出力しないことを意味する。このとき MP 一般逆行列の真値からは離れてしまうことになるが，音場制御の全体的な精度は向上することが期待される。

2.2.3. 条件数抑制による方法

特異値の値が小さい場合にはベクトルの線形独立性が小さくなり，逆システムの安定性が減少するが，これは他の特異値との相対的な比率によるものである。そこで最大特異値と最小特異値の比率，すなわち

$$c = \frac{max(d_j)}{min(d_j)}$$

を条件数としてシステムの安定性を評価する方法が提案されている。前述の打ち切り特異値分解による方法では，強制的にランクを落とすことにより，最小特異値が大きくなるため条件数は小さくなる。しかし，条件数を小さくするという目的であれば，ランクを落とさなくても$max(d_j)/c$よりも小さな特異値を$max(d_j)/c$に置き換えれば，すなわち

$$d_k = max(d_j)/c$$
$$\forall k \in d_k < max(d_j)/c$$

とすれば一定の条件数 c を保つことができる[5]。

2.2.4. 比較検討

上述の 3 手法はいずれも近似的な逆システムを求めるため，周波数領域上では誤差は大きくなるが，時間領域に変換したときに時間幅の制約により生じる誤差を軽減できるため総合的には音場制御システムの性能は向上する。また，特異値を操作する方法は周波数毎に異なる特異値の要素を変更するため，個別のスピーカ出力に注目した場合に周波数軸上での連続性が失われることが懸念される。一方，正則化パラメータ法は例えば周波数帯域毎に正則化パラメータを決めるなどの方法により周波数軸上での連続性をある程度維持できるが，提案されている最急降下法により正則化パラメータを決める方法[6]は現在の BoSC システムのチャンネル数や PC の能力などの条件を考えると計算に数か月かかるため現実的ではない。

3 提案手法

3.1 正則化パラメータの決定方法

オクターブ帯域毎に最適な正則化パラメータを決定する方法を提案する。中心周波数f

のオクターブバンドがk_1からk_2番目（FFT ポイント数はN）までの周波数成分を含む場合，システムの伝達関数の時間信号は

$$g_{ij}[n] = \mathfrak{F}^{-1}\left(G_{ij}[k]\right)$$

ただし，

$$G_{ij}[k] = \begin{cases} \left[g_{ij}\right]'_k & k_1 \le k \le k_2 \\ 0 & k < k_1, k_2 < k \end{cases}$$

となる。ここでNはバンドパスフィルタの時間信号が巡回畳み込みによる時間軸上折り返し誤差が十分小さくなる大きさとする。同様に逆システムの時間信号は

$$h_{ij}[n] = \mathfrak{F}^{-1}\left(H_{ij}[k]\right)$$

ただし，

$$H_{ij}[k] = \begin{cases} \left[h_{ij}\right]_k & k_1 \le k \le k_2 \\ 0 & k < k_1, k_2 < k \end{cases}$$

となる。また逆システムの遅延をN_1として，さらに時間幅$N_2 (< 2N_1)$の窓関数を乗じるため最終的な逆システムの時間信号$\hat{h}_{ij}[n]$は

$$\hat{h}_{ij}[n] = h_{ij}[n - N_1]_{mod\ N} \cdot \mathrm{w}[n]$$

ただしハニング窓を用いる場合，

$$w[n] = 0.5 - 0.5\ \cos\pi\frac{2n - 2N_1 + N_2}{N_2}$$

$$N_1 - N_2/2 \le n \le N_1 + N_2/2$$

となる。システム$g_{ij}[n]$と逆システム$\hat{h}_{ij}[n]$の畳み込みのマトリクス，すなわち

$$p_{jk}[m] = \sum_{i=1}^{L}\sum_{n=0}^{m} g_{ij}[m-n]\hat{h}_{ik}[n]$$

$$m = 0 \cdots 2N - 1$$

が単位行列に相当する次式

$$q_{jk}[m] = \begin{cases} \mathfrak{F}^{-1}\left(Q_{jk}[n]\right) & j = k \\ 0 & j \ne k \end{cases}$$

ただし，

$$Q_{jk}[n] = \begin{cases} e^{-j2\pi n N_1/N} & k_1 \le n \le k_2 \\ 0 & n < k_1, k_2 < n \end{cases}$$

に等しくなれば望みの逆システムが実現できる。したがって，ある正則化パラメータβで中心周波数fの帯域に関する評価式は次のようになる。

$$J_{\beta,f} = \sum_{j,k=1}^{M}\sum_{m=0}^{2N-1}\left(p_{jk}[m] - q_{jk}[m]\right)^2$$

レベルで表すと次のようになる。

$$SNR_{\beta,f} = 10\log_{10}\frac{M\sum_{m=0}^{2N-1} q_{11}[m]^2}{J_{\beta,f}} \quad [\mathrm{dB}]$$

4　実験

4.1 実験条件

開発した没入型聴覚ディスプレイ装置“音響樽”内で 96 個のスピーカから 80 個のマイクロホンをもつ BoSC マイクロホンへの合計 7680 のインパルス応答$\left[G_{ij}\right]'$をサンプリング周波数 48kHz で計測した。すべてのインパルス応答は 2048 点でほぼ収束することを確認した。FFT ポイント数N=8192，逆システムの遅延$N_1 = 4096$，窓幅$N_2 = 4096$とした。また，計算する中心周波数は 125（ただし低域は 20Hz まで含む)，250，500，1k，2k，4k，8k，16k（ただし高域は 20kHz まで）の 8 種類とした。それぞれの帯域分割の周波数を Table 1 に示す。

Table 1 帯域分割の周波数範囲

f[Hz]	f_1 [Hz]	f_2 [Hz]	k_1	k_2
≤125	20.0	176.8	4	30
250	176.8	353.6	31	60
500	353.6	707.1	61	120
1k	707.1	1414.2	121	241
2k	1414.2	2828.4	242	482
4k	2828.4	5656.9	483	965
8k	5656.9	11313.7	966	1930
16k	11313.7	20000.0	1931	3413

正則化パラメータβは以下の 21 種類（$i = 1 \ldots 21$）について計算した。

$$\beta_i = 10^{(i-17)/4}$$

4.2 実験結果

上記の条件で各正則化パラメータ，各帯域のすべての組み合わせについて PC（Intel Corei7-3770k, CPU3.5GHz, Memory16GB）および数値計算ソフト matlab を用いて計算したところ約 96 時間かかった。

計測したインパルス応答の一例を図２に示す。また中心周波数 1kHz におけるβと SNR の関係を図３と図４に示す。中心周波数 1kHz 以下ではβが小さいとき SNR は低く，βを増やすと SNR は最大値をとり，さらにβを増やすと SNR は低くなる。また中心周波数 2kHz 以上ではβが小さい場合も SNR は高く特に 4kHz 以上では明確なピークを持たないことがわかる。

SNR 最大値をとる各中心周波数に関するβの値とそのときの SNR を Table 2 に示す。Table 2 のβを用いて正則化パラメータ法に

より逆システムを設計した。逆システムの時間波形を図５に示す。比較のために正則化パラメータ法を用いずに（β=0）設計した逆フィルタを図６に示す。低域の周波数が強調されていることが時間波形からもわかる。低域周波数ではすべてのマイクロホンにおいて信号の相関が高くなるためである。また全帯域（20～20kHz）にわたる SNR は最適化 β の場合は 6.7 dB と正則化パラメータを用いない場合（β＝０）の－3.9 dB や直感的に正則化パラメータを決めた場合の 4.8 dB と比較して向上した。

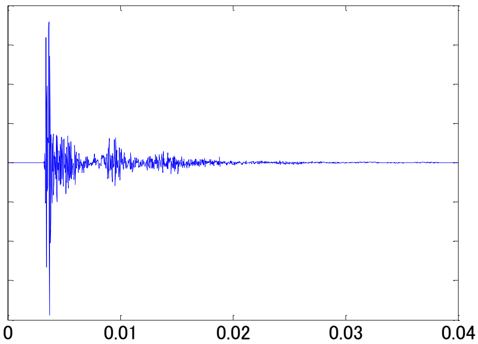

図２　音響樽内におけるインパルス応答

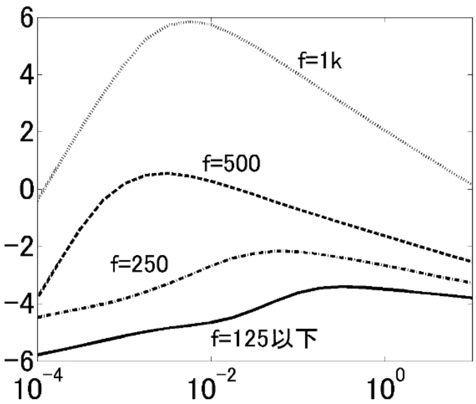

図３　β と SNR の関係（f=1k 以下）

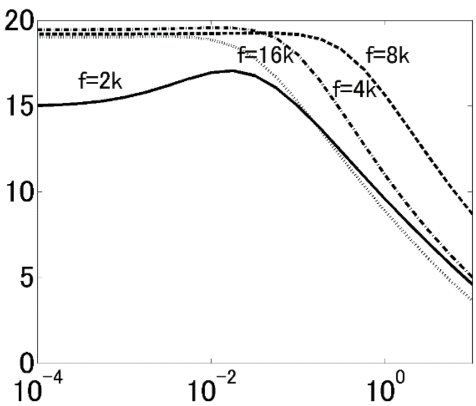

図４　β と SNR の関係（f=2k 以上）

Table 2　最適化 β と SNR

f [Hz]	β	SNR [dB]
≦125	0.3162	-3.4
250	0.0562	-2.2
500	0.0032	0.6
1k	0.0056	5.9
2k	0.0178	17.1
4k	0.0100	19.6
8k	0.0316	19.2
16k	0.0032	19.0

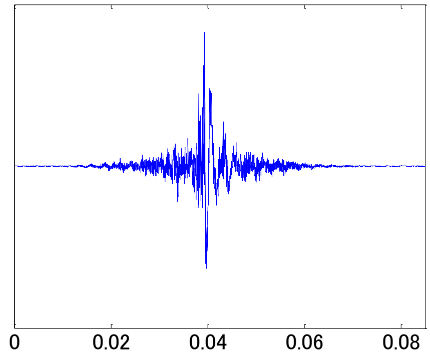

図５　逆システムの時間波形（最適化 β）

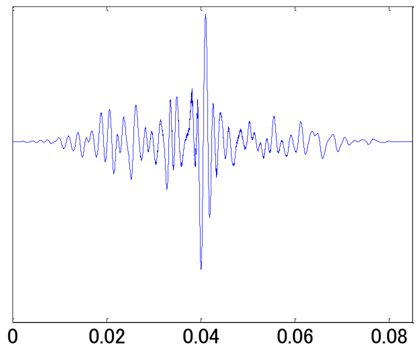

図６　逆システムの時間波形（β＝０）

5　おわりに

オクターブバンド毎の周波数軸上で最適な正則化パラメータを見つけることにより逆システムを設計する方法を提案し，その有効性を示した。

参考文献

[1]　伊勢，音学誌, 53(9), 706-713, 1997.

[2]　Miyoshi et al., IEEE Trans. ASSP, 36, 2, 145-152, 1988.

[3]　神沼他，音学誌, 57(3), 175-183, 2001.

[4]　Tokuno et al. IEICE Trans. Fundamentals, E80-A, 5, 809-820, 1997.

[5]　開原他，信学技報, 105(556), 7-12, 2006.

[6]　李他，音学誌, 69(6), 276-183, 2013.