22.2 マルチチャンネル音響の枠型スピーカによる バイノーラル再生法の開発

Development of Binaural Reproduction Method for 22.2 Multichannel Sound with Loudspeaker Frame

学生会員 松井 健太郎^{†1,†2}, 大石 諭^{†3}, 正会員 杉本 岳大^{†1},
森田 雄一^{†4}, 足立 修一^{†2}

Kentaro Matsui^{†1,†2}, Satoshi Oishi^{†3}, Takehiro Sugimoto^{†1}, Yuichi Morita^{†4} and Suichi Adachi^{†2}

あらまし フラットパネルディスプレイに一体化される枠型 12 スピーカを用いた、22.2 マルチチャンネル音響のバイノーラル再生法を開発した。バイノーラル再生法は、人間の聴覚特性を利用し、音像を任意の方向に定位知覚させる技術である。この技術を援用し、スピーカの置かれていない側方、後方のチャンネルをその所定方向に音像として定位知覚させることにより、22.2 マルチチャンネル音響を擬似的に再生する。試作した枠型 12 スピーカを用いた測定実験により、クロストークを十分に抑圧可能であること、複数のスピーカユニットを使用することによって制御器の安定性が向上することを確認した。また、主観評価実験により、22.2 マルチチャンネル音響の擬似再生として十分な空間印象を提示可能であることを確認した。

キーワード：スーパーハイビジョン、頭部伝達関数、バイノーラル再生、予測誤差法、枠型スピーカ、フラットパネルディスプレイ

1. ま え が き

8 K スーパーハイビジョン（以下、8K）の音響方式として、22.2 マルチチャンネル音響（以下、22.2ch 音響）方式の開発を進めている¹⁾。22.2ch 音響方式は、視聴者を取り囲むように配置されたフルレンジの 22 チャンネルと、低音効果（Low Frequency Effects: LFE）用の 2 チャンネルから構成され、その場にいるかのような高い臨場感と 3 次元的な音響空間構築を実現することができる方式である²⁾。3 層からなるそのチャンネル配置を図 1 に、チャンネル番号とラベル、名称の一覧を表 1 に示す。

既存のチャンネルベース音響方式のチャンネル配置を包含し、高臨場感音響方式に求められる要求条件³⁾を満たす本方式は、8 K の音響方式として十分な性能を有しているといえる。一方、ライフスタイルの多様化に伴ってテレビの視聴スタイルも変化する中で⁴⁾、多くの場合、各家庭で 24 個のスピーカを配置することは困難と考えられる。したがって、22.2ch 音響を、より少ないスピーカ数で、より簡

図 1 22.2 マルチチャンネル音響のチャンネル配置

表 1 チャンネル番号、ラベル、名称

番号	ラベル	名称	番号	ラベル	名称
1	FL	Front Left	13	TpFL	Top Front Left
2	FR	Front Right	14	TpFR	Top Front Right
3	FC	Front Center	15	TpFC	Top Front Center
4	LFE1	LFE-1	16	TpC	Top Center
5	BL	Back Left	17	TpBL	Top Back Left
6	BR	Back Right	18	TpBR	Top Back Right
7	FLc	Front Left Center	19	TpSiL	Top Side Left
8	FRc	Front Right Center	20	TpSiR	Top Side Right
9	BC	Back Center	21	TpBC	Top Back Center
10	LFE2	LFE-2	22	BtFC	Bottom Front Center
11	SiL	Side Left	23	BtFL	Bottom Front Left
12	SiR	Side Right	24	BtFR	Bottom Front Right

2015年8月19日受付、2015年11月20日再受付、2015年12月8日採録
†1 NHK 放送技術研究所

（〒157-8510 世田谷区砧 1-10-11, TEL 03-5494-3361）

†2 慶應義塾大学大学院理工学研究科

（〒223-8522 横浜市港北区日吉 3-14-1, TEL 045-566-1822）

†3 NHK 広島放送局

（〒730-8672 広島市中区大手町 2-11-10, TEL 082-504-5241）

†4 フォスター電機株式会社

（〒196-8550 昭島市つつじが丘 1-1-109, TEL 042-546-2445）

易に聴取することができる再生法を 8K 視聴の選択肢として提案することは、重要な課題の一つである。

筆者らは、フラットパネルディスプレイ (Flat Panel Dis-

play: FPD) のベゼルに外接して一体化される棒型のスピーカと、この棒型スピーカを用いたバイノーラル再生法の研究、開発を進めている⁵⁾。バイノーラル再生法は、人間の聴覚特性を利用し、音像を任意の方向に定位知覚させる技術である。この技術を援用し、側方、後方のチャンネルをその所定方向に音像として定位知覚させることによって、棒型スピーカのような前方に配置したスピーカのみで 22.2ch 音響を擬似再生することが可能となる。

本稿では、まず 2 章において、このバイノーラル再生法とその基礎となる頭部伝達関数 (Head-Related Transfer Function: HRTF) について概説する。続く 3 章では、従来長時間を要していた HRTF の測定を短時間化する、多方向同時推定法を報告する。さらに 4, 5 章では、開発を進める棒型 12 スピーカによる 22.2ch 音響のバイノーラル再生法と、その有効性を検証するために実施した評価実験について報告する。

2. バイノーラル再生法

2.1 ヘッドホンによるバイノーラル再生法

HRTF は、自由音場において、「頭がない状態での頭部中心に相当する両耳中心位置から、頭のある状態での両耳鼓膜位置もしくは外耳道入口までの音響伝達関数」として、または近似的に、「頭外音源位置から両耳鼓膜位置もしくは外耳道入口までの音響伝達関数」として定義される、音源位置と周波数にかかる伝達関数である⁶⁾。この HRTF には、両耳間時間差やレベル差、周波数特性上のスペクトルキューなど、音像の定位知覚にかかる多くの特徴量が含まれている。そのため、HRTF を音源信号に作用させることにより、すなわち HRTF の時間領域表現である頭部インパルス応答 (Head-Related Impulse Response: HRIR) を音源信号に畳み込むことにより、任意の方向にその音像を定位知覚させることができる。このようにして生成される信号をバイノーラル信号、バイノーラル信号を主にヘッドホンにより提示する再生法をバイノーラル再生法という。

2.2 スピーカによるバイノーラル再生法

両耳に直接信号を提示することができるヘッドホン受聴では、所望方向の左右耳 HRIR を畳み込んだ音源信号（以下、所望信号）を、直接左右のドライバユニットから再生すればよい（図 2）。しかし、スピーカ受聴の場合には、スピーカから同側耳への信号の伝搬に加え、対側耳への信号の漏洩（クロストーク）が発生する（図 3）。したがって、このクロストークを抑圧し、所望信号のみをそれぞれの耳に伝送するための補償処理が必要となる。この補償処理をクロストーク・キャンセレーションという。クロストーク・キャンセレーションにより実現されるスピーカによるバイノーラル再生法^{*}を、ヘッドホンによる再生法と区別してトラ

図 2 ヘッドホンによるバイノーラル再生

図 3 クロストーク

図 4 スピーカによるバイノーラル再生（トランスオーラル再生）

ンスオーラル再生法ということもある。図 4 に、2 個のスピーカによるバイノーラル再生の概念図を示す。図 4 において、 G はそれぞれのスピーカから左右耳までの HRTF、 x は左右耳に提示する所望方向の HRTF、 H はクロストーク・キャンセレーションのための制御器を表す^{**}。入力信号 u から出力信号 y までは、次式のような関係で表される。

$$y = GHxu$$

ただし

$$\begin{aligned} y &= \begin{pmatrix} y_l \\ y_r \end{pmatrix} \\ G &= \begin{pmatrix} g_{ll} & g_{rl} \\ g_{lr} & g_{rr} \end{pmatrix} \\ H &= \begin{pmatrix} h_{ll} & h_{rl} \\ h_{lr} & h_{rr} \end{pmatrix} \\ x &= \begin{pmatrix} x_l \\ x_r \end{pmatrix} \end{aligned} \quad (1)$$

制御器は、聴取者の耳の位置において所望信号、すなわち入力信号に所望方向の HRTF を作用させた信号が合成されるように設計される。その結果、入力信号と出力信号の関係は

^{*} 一般にはヘッドホンによる提示をバイノーラル再生というが⁶⁾、スピーカによる提示も含めてバイノーラル再生という場合もあり、本稿でもそれにならうことにする。

^{**} これらは設計アプローチにより時間応答、周波数応答、有理伝達関数など様々な表現となりうるが、ここでは便宜上 HRTF とし、周波数領域で説明している。

$$\mathbf{y} = \mathbf{x}u$$

となる。したがって、制御器 $\mathbf{H}$ は $\mathbf{G}$ の逆システムとして設計される。すなわち

$$\mathbf{H} = \mathbf{G}^{-1}$$

なお、制御対象となる聴取者の耳の位置を制御点、制御音源となるスピーカを 2 次音源という。

この制御器の設計に関しては、様々なアプローチが検討されている^{7)~20)}。筆者らも、畳み込み演算に基づく時間領域での処理²¹⁾、特異値分解法による周波数領域での処理²²⁾、HRTF の低次元モデリングと非干渉化制御による方法²³⁾²⁴⁾など、複数の角度から検討を進めてきた。本稿で報告する再生法では、計算時間、占有する作業領域などの計算負荷を考慮し、周波数領域での処理を採用している。この場合、制御器は、離散周波数ビン ω_k ごとに逆行列を計算し、逆フーリエ変換により時間領域に戻すことによって、FIR (Finite Impulse Response) フィルタとして求められる。

$$\mathbf{H}(\omega_k) = \mathbf{G}^{-1}(\omega_k) \quad (2)$$

ただし、 k は離散周波数インデックスを表す。

ここまで 2 個のスピーカによるバイノーラル再生法について概説したが、これは 3 個以上の複数スピーカを用いた方法に容易に拡張できる。2 次音源数を増やすとその数に比例して制御点の数も増やすことができるため、複数スピーカによる制御は、聴取範囲を広げる有力な手段である。このとき (1) は、次のようになる。

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}$$

$$\mathbf{G} = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1q} \\ g_{21} & g_{22} & \cdots & g_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ g_{p1} & g_{p2} & \cdots & g_{pq} \end{pmatrix}$$

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1p} \\ h_{21} & h_{22} & \cdots & h_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ h_{q1} & h_{q2} & \cdots & h_{qp} \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

ここで、 p は制御点の数を、 q は 2 次音源の数を表す。

3. 頭部伝達関数の多方向同時推定

2 次音源、制御点を多数配置したバイノーラル再生法では、その数に応じて測定を必要とする HRTF も増加するため、

測定作業が煩雑となり、長時間を要するようになる。そこで、多方向の HRIR を MISO (Multi Input Single Output) FIR モデルとみなし、システム同定理論に基づいて推定する、同時推定法を開発した²⁵⁾。

3.1 予測誤差法による MISO FIR モデルの推定

一般に、 m 入力 1 出力の n 次 FIR モデルは

$$y(k) = \theta_1^T \mathbf{u}_1(k) + \theta_2^T \mathbf{u}_2(k) + \cdots + \theta_m^T \mathbf{u}_m(k) + w(k) \quad (3)$$

と表される。ここで $y(k)$ は離散時間の時刻 k における出力、 $w(k)$ は正規性白色雑音を表す。また、 $\mathbf{u}_i(k)$ は i 番目の入力 $u_i(k)$ から構成される入力ベクトル、 θ_i は i 番目の FIR モデルのパラメータベクトルを表し、それぞれ次式で与えられる。

$$\mathbf{u}_i(k) = \begin{pmatrix} u_i(k-1) & u_i(k-2) & \cdots & u_i(k-n) \end{pmatrix}^T$$

$$\theta_i = \begin{pmatrix} \theta_{i1} & \theta_{i2} & \cdots & \theta_{in} \end{pmatrix}^T$$

さらに、(3) を時刻 $k = 1 \cdots N$ まで並べると

$$\begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{pmatrix} = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_m \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_m \end{pmatrix} + \begin{pmatrix} w(1) \\ w(2) \\ \vdots \\ w(N) \end{pmatrix} \quad (4)$$

と書き表すことができる。ここで $\mathbf{U}_i$ は各時刻における入力ベクトル $\mathbf{u}_i(k)$ を並べた行列であり、次式で与えられる。

$$\mathbf{U}_i = \begin{pmatrix} u_i(0) & u_i(-1) & \cdots & u_i(1-n) \\ u_i(1) & u_i(0) & \cdots & u_i(2-n) \\ \vdots & \vdots & \ddots & \vdots \\ u_i(N-1) & u_i(N-2) & \cdots & u_i(N-n) \end{pmatrix}$$

簡単化のため

$$\mathbf{y} = \begin{pmatrix} y(1) & y(2) & \cdots & y(N) \end{pmatrix}^T$$

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}_1 & \mathbf{U}_2 & \cdots & \mathbf{U}_m \end{pmatrix}$$

$$\boldsymbol{\theta} = \begin{pmatrix} \theta_1^T & \theta_2^T & \cdots & \theta_m^T \end{pmatrix}^T$$

$$\mathbf{w} = \begin{pmatrix} w(1) & w(2) & \cdots & w(N) \end{pmatrix}^T$$

とおくと、(4) は

$$\mathbf{y} = \mathbf{U}\boldsymbol{\theta} + \mathbf{w} \quad (5)$$

と書き表すことができる。(5) は、時刻 $k-1$ までの過去の入出力と現時刻 k の出力との関係を表している。この関係式に基づく出力 $\mathbf{y}$ の 1 段先予測値 $\hat{\mathbf{y}}$ は、次式により求められる。

$$\hat{\mathbf{y}} = \mathbf{U}\boldsymbol{\theta}$$

予測誤差法とは、この予測値に基づく予測誤差から構成

される評価規範を最小化する推定値を計算する，パラメータ推定法の総称である²⁶⁾．予測誤差の大きさの測度として2次関数を用いると最小二乗法となる．このとき，評価規範 J_N は

$$J_N(\theta) = \frac{1}{N} \|y - U\theta\|_2^2$$

となる．これを最小化する最小二乗推定値 $\hat{\theta}$ は，次式より計算される．

$$\hat{\theta} = R^{-1}f \quad (6)$$

ここで

$$R = \frac{1}{N} U^T U$$

$$f = \frac{1}{N} U^T y$$

とおいた．

3.2 入力信号の設計

(6) が最小二乗推定値として解を持つためには， R が正定値行列である必要がある．この条件を満たす入力信号は，以下の手続きにより作成することができる²⁷⁾．

- (i) 周期 T の M 系列から信号 $m(k)$ を作成し，これを1番目の入力

$$u_1(k) = m(k)$$

とする．ここで，周期 T は次式を満たす．

$$l > n \quad (7)$$

$$T > n + (m-1)l \quad (8)$$

- (ii) 1番目の入力 $u_1(k)$ を時間軸上で l サンプルだけ負の方向に巡回シフトしたものを2番目の入力

$$u_2(k) = \begin{cases} m(k+l) & k=1, \dots, T-l \\ m(k+l-T) & k=T-l+1, \dots, T \end{cases}$$

とする．

- (iii) 以下同様に， i 番目の入力 $u_i(k)$ は， $i-1$ 番目の入力 $u_{i-1}(k)$ を時間軸上で l サンプルだけ負の方向に巡回シフトしたものとす．

$$u_i(k) = \begin{cases} m(k+(i-1)l) & k=1, \dots, T-(i-1)l \\ m(k+(i-1)l-T) & k=T-(i-1)l+1, \dots, T \end{cases}$$

上記の手続きは，入力に対する応答がモデルの次数内に収束することを仮定している．この仮定が成り立たない場合，手続き中 (7)(8) を以下のように修正する²⁸⁾．

$$l \geq \max\{n_s, n\} \quad (9)$$

$$T \geq \max\{n_s, n\} + (m-1)l \quad (10)$$

ただし， n_s は応答の収束時間（サンプル数）を表す．

3.3 頭部伝達関数推定への応用

m 方向の HRIR の集合を，入力を m 方向の測定信号，出

力を片耳の外耳道入口で収音される応答とした MISO FIR モデルとみなし，提案法により推定する．

まず，3.2 節で示した手続きにしたがって測定信号を作成する．48kHz のサンプリング周波数において HRIR を推定するとき，FIR モデルの次数 n は 128 から 512 に設定される．音響無響室など反射のない環境で測定を実施する場合には，そのまま手続きにしたがって測定信号を作成すればよい．防音室など反射の無視できない環境で測定を実施する場合には，その収束時間（サンプル数）を n_s とし，(9)(10) を満たす巡回シフト量 l および周期 T を定める．次に，推定する m 方向から作成した測定信号を同時に印加し，左右耳の外耳道入口で応答をそれぞれ収音する．左右耳の HRIR は独立した MISO FIR モデルとみなすことができるため，応答を同時に収音しても問題はない．最後に，印加した測定信号を入力，収音された応答を出力として (6) に代入し計算することにより，各方向の FIR モデルのパラメータ推定値を求める．

4. 枠型スピーカによる 22.2ch 音響のバイノーラル再生

4.1 ディスプレー一体型枠型スピーカ

家庭での 8K の視聴を考えると，スピーカの数も然ることながら，部屋の広さや家具などにより，その設置場所も制限されることが予想される．そこで筆者らは，設置に場所をとらないスピーカの形態として，FPD 一体型の枠型スピーカの研究，開発を進めている．

85 型液晶ディスプレイのために試作した枠型 12 スピーカを図 5 に示す．図からわかるように，枠の上辺と下辺には各 5 個のスピーカユニットが等間隔で配置され，左右の側辺にはその中間位置に各 1 個のスピーカユニットが配置されている．この配置は，22.2ch 音響方式の前方チャンネルに対応している．枠上の各スピーカユニットは，相互の干渉と混変調を防ぐために，独立したキャビティに収められている．それぞれのスピーカユニットは高い耐入力と低い歪み特性を有しており，振動板の直径は 7cm と小サイズであるにもかかわらず，最大音圧レベルは 92 dB となっている．さらに，図 6 に示す振動板のエッジ構造は NHK の標準モニタ用の技術を小型化して導入したものであり，大振幅時の高い直進性と逆共振の抑圧を実現することにより，従来の同口径スピーカユニットに比べ，中域で 20 dB 程度の歪み低減に成功している．

左右の側辺にはそれぞれ 2 個のサブウーファユニットが備えられており，LFE チャンネルの再生により広がり感や包まれ感などの空間印象を向上させると共に，22.2ch 音響の低域成分の再生により，低域音響特性を改善している．

4.2 枠型 12 スピーカによるバイノーラル再生

22.2ch 音響方式の各チャンネルのうち，ディスプレイの位置に重なる FLc チャンネル，FRc チャンネル，FC チャンネルを除いた前方チャンネルは，枠型 12 スピーカ上の各

図 5 枠型 12 スピーカ

図 6 高調波歪みを低減するエッジ構造

スピーカユニットに割り当てられる。一方、除外された 3 チャンネルは、振幅パンニングにより仮想音像として合成される。実験によれば、水平方向のパンニングと比較して、垂直方向のパンニングは前方により安定した仮想音像を合成することが可能である²⁹⁾。したがって、上記 3 チャンネルの合成には、各チャンネルを挟む上辺と下辺のスピーカユニットを用いている。側方、後方のチャンネルは、12 個すべてのスピーカユニットを 2 次音源としたバイノーラル再生により、規格上の所定方向に音像として合成される。聴取位置は視距離 1.5H (H: 画面高さ) の 1 点とした。したがって、2 次音源は 12 点、制御点は 2 点となる。このとき (2) は劣決定問題となり、解が一意に定まらない。そのため、制御器を設計する際に計算する逆行列の条件数が最も低くなり、制御器が安定となる最小ノルム解を採用している。

制御器の設計に使用する HRTF, すなわち 2 次音源から制御点までの HRTF は、3 章で述べた同時推定法により求めた。NHK 放送技術研究所の音響無響室での測定の様子を図 7 に示す。枠型 12 スピーカの各スピーカユニットに測定信号を同時に印加し、聴取位置に配置した HATS (Head and Torso Simulator) * で応答を收音した。測定信号の信号長は $2^{17} - 1$ サンプル (シフトレジスタ 17), サンプル周波数は 48 kHz とした。スピーカユニットは枠上に配

図 7 測定の様子

置されているため、聴取位置からの距離はそれぞれ異なっている。この距離差により生じる遅延と、スピーカキャビネットでの反射を吸収するため、FIR モデルの次数 n は 1,024 に、収束時間 n_s は 1,200 サンプルとした。測定音圧は、HATS の頭部位置において 70 dB とした。

所望信号の合成に使用する HRTF, すなわち 22.2ch 音響方式の側方や後方のチャンネルから聴取位置までの HRTF は、音響無響室内に 22.2ch 音響の再生環境を構築することが困難であったため、同じ HATS を用いて 1 方向ずつ測定した。測定信号には、信号長 2^{17} サンプル、サンプリング周波数 48 kHz の LogTSP (Logarithmic Time-Stretched Pulse) 信号³⁰⁾ を用いた。測定音圧は、HATS の頭部位置においてピークを 70 dB とした。

測定例として、左右側辺中央のスピーカユニットから左右耳への HRIR を図 8 に、その周波数振幅応答を図 9 に示す。ただし、HRIR はピークがフルスケールを基準として -2 dB となるように正規化を行ったのち、512 サンプルの矩形窓で切り出している。なお、信号長 2^{17} サンプル、サンプリング周波数 48 kHz の LogTSP 信号を用いての 1 ユニットごとの測定も実施しており⁵⁾、これを基準としたときの平均スペクトル歪みは、左耳で 0.425 dB, 右耳で 0.504 dB であった。同時推定法のより一般的な検証については、筆者らによる既報論文²⁵⁾²⁸⁾ を参照されたい。また、聴取位置側方、すなわち SiL チャンネル, SiR チャンネルから左右耳への HRIR を図 10 に、その周波数振幅応答を図 11 に示す。

次に、枠型 12 スピーカと制御器を接続し、入力信号として単位インパルスを制御器に印加したときの応答を測定した。聴取位置でのインパルス応答を図 12 に、その周波数振幅応答を図 13 に示す。ただし、単位インパルスを直接印加することは困難であるため、LogTSP 信号を印加し、応答に逆信号を畳み込むことによりインパルス応答とした。図 12 と図 13 において、「left」と「right」は、左右それぞれの入力端から単位インパルスを印加した場合の応答であることを示す。この場合、同側耳では遅延を伴う単位イン

* 人工耳を備えた人工の人体の上半身。人間に音響機器を装着した状態を模擬した音響研究用のマネキンであり、平均的な男性、女性の頭部と胴体の寸法をもとに製作されている。

図 8 スピーカユニットから左右耳への HRIR の例

図 10 側方チャンネルから左右耳への HRIR の例

図 9 スピーカユニットから左右耳への HRTF 振幅応答の例

図 11 側方チャンネルから左右耳への HRTF 振幅応答の例

図 12 聴取位置で測定されたインパルス応答

図 13 聴取位置で測定された周波数振幅応答

パルスが、対側耳では無信号が観察されることが期待される。図 13 を見ると、同側耳で観察される信号は所望の全域通過特性を精度よく近似しており、対側耳で観察されるクロストークはおおよそ全周波数帯域で 15 dB 以上抑圧されていることがわかる。いずれも、低周波および高周波帯域において精度が低下しているが、これはスピーカユニットの再生周波数帯域外の帯域となるためと考えられる。

5. 評価実験

枠型 12 スピーカによるバイノーラル再生法の有効性を検証するため、2 次音源として使用するスピーカユニット数を変えたときの制御器の安定性を評価した。また、22.2ch 音響の擬似再生法という位置付けから、高臨場感音響方式の要求条件³⁾としても挙げられている再生音の 3 次元的な空間印象を、有効性の一指標として評価した。

5.1 制御器の安定性評価

実験条件

制御器の安定性は、その設計時に計算する逆行列の条件数を尺度として評価した。条件数は、外乱やデータの摂動に対する感度の測度として用いられ、条件数の大きい問題は「悪条件である」といい、微小なノイズや計算過程の誤差混入に弱いとされている。ここでは、枠型 12 スピーカのうちバイノーラル再生に使用するスピーカユニット数とレイアウトを図 14 に示すように変化させ、離散周波数ごとに逆行列の条件数を計算した。なお、HRTF の測定にかかる諸条件は 4 章と同様とした。

実験結果

スピーカユニット数、離散周波数ごとの条件数を図 15 に示す。図からわかるように、周波数ごとに傾向はやや異なるものの、条件数がスピーカユニット数におおよそ反比例して減少する結果となった。スピーカユニットの再生周波数帯域外となる 200Hz 以下を除いて、全周波数帯域での平均は 2.13 (two-units) から 1.53 (12-units) に、最大値は 7.11 (two-units) から 3.21 (12-units) にまで減少した。これは、バイノーラル再生に 2 次音源として使用するスピーカの数を増やすことにより、制御器の安定性、ひいては合成される音響の安定性を向上させることができることを示している。また、高周波帯域に見られる条件数の急峻なピークは、聴感上不快なアーティファクトを生じる遠因となっている。スピーカの数を増やすことにより、これらのピークが大幅に抑圧されることも見て取れる。

5.2 空間印象の主観評価

実験条件

再生音の空間印象は、音の到来方向、上下感、音による包まれ感といった様々な観点から総合的に評価した。評価法には、Rec. ITU-R BS.1534-2³¹⁾ に規定される MUSHRA (Multiple Stimuli with Hidden Reference and Anchor) 法を応用した。MUSHRA 法は音質評価などに用いられる主観評価法の一つで、隠れ基準音、隠れアンカー音を含む

図 14 安定性評価に使用したスピーカユニット配置

図 15 条件数の推移

表 2 評価した再生法		
再生法	スピーカ配置	
—	22.2	基準音, 隠れ基準音
ダウンミックス再生 ¹	5.1 2	
モノフォニック再生 ²	2	隠れアンカー音
バイノーラル再生	12 (枠型) 2	

¹ ARIB STD-B32 準拠³²⁾
² 2 チャンネルダウンミックスを平均化して算出
2 個のスピーカにより同一の評価音を再生

図 16 5 段階劣化尺度

複数の評価音を同時に比較することができるという特徴を持つ。評定者は表 2 に示す各再生法を自由に切り替えて聴くことができ、図 16 に示す 5 段階の劣化尺度により、基準音からの劣化度合いを一括して評価する。ただし本実験では、評価音を通しての空間印象を総合的に判断するため、再生の途中では切り替えないように教示した。

実験は Rec. ITU-R BS.1116-3³³⁾ 準拠の音質評価室で実施した。評価音の再生には、前方のチャンネルについては

表 3 空間印象評価に使用した音源		
ラベル	内容	持続時間 (ms)
Kohaku (紅白)	音楽番組, ヴォーカル	12,190
Yae (八重の桜)	ドラマ, ダイアログ, 背景音	10,460
Chorus	ドラマ, ナレーション	10,730
ThisIsSHV	ジングル, 効果音	10,830

図 17 音源ごとの評価結果

試作した枠型 12 スピーカを用い、側方、後方のチャンネルについては、同一のスピーカユニットを使用して作成したブックシェルフ型のスピーカを用いた。視距離は 1.5H を想定し、枠型 12 スピーカは聴取位置正面に、側方、後方のスピーカは聴取位置を中心とする同心円上に等距離、等間隔で配置した。

評価音の音源には、表 3 に示す 22.2ch 音響のクリップを、典型的な 8K 番組として選定した。評定者は 20 代から 40 代の男女 24 名である。ただし、隠れ基準音に対して 80 点未満の評点をつけた評定者は、事後スクリーニングにより除外した。結果、最終的な評定者は 21 名となった。

実験結果

図 17 に音源ごとの評価結果を、図 18 にすべての音源の評点平均を示す。図中、エラーバーは 95% 信頼区間を表している。

隠れアンカー音として加えたモノフォニック再生を除くと、2 チャンネルバイノーラル再生の評点平均が最も低くなった。次いで低かったのが 2 チャンネルダウンミックス再生であり、この 2 再生法は、基準音からの空間印象の違いが平均して「気になる」という評価となった。

提案する枠型 12 スピーカによるバイノーラル再生は、すべての音源について 2 チャンネルダウンミックス再生より有意に高い評点となった。5.1 チャンネルダウンミックス再生と比較すると、Chorus, ThisIsSHV については同等かそれ以上の評点がついた一方で、Yae については有意に低い評点となった。Yae には側方、後方に明瞭に定位するダイアログが多く含まれており、その方向に実スピーカを有する 5.1 チャンネル配置が優位に働いたと考えられる。

図 18 全音源の評点平均

評定者の内観からは、クリップにナレーションやヴォーカルが含まれる音源 (Kohaku, Yae, Chorus) では、その到来方向、定位感を主な手掛かりとしている傾向が伺えた。そういった手掛かりのない音源 (ThisIsSHV) では、全体的な上下感、広がり感、移動感を判断材料としたようである。後者の音源では、一部の評定者から基準音と枠型 12 スピーカによるバイノーラル再生の弁別が困難との報告も挙げられた*。

すべての音源の平均でみると、基準音からの空間印象の違いが「わかるが気にならない」という評価となった。これは 5.1 チャンネルダウンミックス再生と同じ評価カテゴリであり、その劣化度合いが許容される程度であることを示唆している。したがって、22.2ch 音響の擬似再生としては十分な提示性能であると考えられる。

6. む す び

22.2ch 音響を家庭で簡単に聴取できる再生法として、FPD 一体型の枠型スピーカによるバイノーラル再生法を開発した。試作した枠型 12 スピーカを用いた測定実験により、おおよそ全周波数帯域でクロストークを 15 dB 以上抑圧できていること、複数のスピーカユニットを使用することによって制御器の安定性が向上することを示した。また、主観評価実験により、スイートスポットとなる聴取位置においては、22.2ch 音響の擬似再生として十分な空間印象を提示可能であることを示した。

2 次音源、制御点を多数配置したバイノーラル再生法では、必要となる HRTF が増加し測定作業も煩雑となる。これに対し、HRTF の多方向同時推定法を開発した。

今回の試作では聴取位置を 1 点としたが、これを多点へ、さらにはある限らえた範囲へと拡張することが今後の課題である。また、信号処理量の削減など、実用化に資する検

討も進める予定である。

〔文 献〕

- 1) Y. Shishikui, Y. Fujita, and K. Kubota: "Super Hi-Vision - the star of the show!", EBU Tech. Rev., pp. 4-16 (2009)
- 2) K. Hamasaki, T. Nishiguchi, R. Okumura, Y. Nakayama, and A. Ando: "A 22.2 Multichannel Sound System for Ultrahigh-Definition TV (UHDTV)", SMPTE Mot. Imag. J., **117**, 3, pp. 40-49 (2008)
- 3) Rec. ITU-R BS.1909-0: "Performance requirements for an advanced multichannel stereophonic sound system for use with or without accompanying picture", International Telecommunication Union (2012)
- 4) 澤谷郁子: "家庭におけるマルチチャンネル音響再生技術", NHK 技研 R&D, **128**, pp. 11-17 (2011)
- 5) K. Matsui, S. Oishi, T. Sugimoto, S. Oode, Y. Nakayama, H. Okubo, H. Sato, K. Mizuno, Y. Morita, and S. Adachi: "Binaural Reproduction of 22.2 Multichannel Sound with Flat Panel Display-Integrated Loudspeaker Frame for Home Use", 映情学誌, **68**, 10, pp. J447-J456 (2014)
- 6) "新版 音響用語辞典", 日本音響学会編 (2003)
- 7) B. B. Bauer: "Stereophonic Earphones and Binaural Loudspeakers", J. Audio Eng. Soc., **9**, 2, pp. 148-151 (1961)
- 8) B. S. Atal and M. R. Schroeder: "Apparent Sound Source Translator", US Patent 3,236,949
- 9) M. R. Schroeder: "Digital Simulation of Sound Transmission in Reverberant Spaces", J. Acoust. Soc. Am., **47**, 2(1), pp. 424-431 (1970)
- 10) P. A. Nelson, F. Orduña-Bustamante, and D. Engler: "Experiments on a System for the Synthesis of Virtual Acoustic Sources", J. Audio Eng. Soc., **44**, 11, pp. 990-1007 (1996)
- 11) H. Tokuno, O. Kirkeby, P. A. Nelson, and H. Hamada: "Inverse Filter of Sound Reproduction Systems Using Regularization", IE-ICE Trans. Fundamentals, **E80-A**, 5, pp. 809-820 (1997)
- 12) O. Kirkeby, P. A. Nelson, and H. Hamada: "Local sound field reproduction using two closely spaced loudspeakers", J. Acoust. Soc. Am., **104**, 4, pp. 1973-1981 (1998)
- 13) O. Kirkeby and P. A. Nelson: "Digital Filter Design for Inversion Problems in Sound Reproduction", J. Audio Eng. Soc., **47**, 7/8, pp. 583-595 (1999)
- 14) T. Takeuchi and P. A. Nelson: "Optimal source distribution for virtual acoustic imaging", ISVR Tech. Rep., 288, University of Southampton (2000)
- 15) S. Miyabe, M. Shimada, T. Takatani, H. Saruwatari, and K. Shikano: "Multi-Channel Inverse Filtering with Selection and Enhancement of a Loudspeaker for Robust Sound Field Reproduction", in Proc. IWAENC 2006, pp. 1-4 (2006)
- 16) N. Kamado: "Sound Field Reproduction Integrating Multi-Point Sound Field Control and Wave Field Synthesis", Doctoral Dissertation (2012)
- 17) H. Hokari, Y. Furumi, and S. Shimada: "A Study on Loudspeaker Arrangement in Multi-Channel Transaural System for Sound Image Localization", in Proc. AES 19th International Conference (2001)
- 18) M. R. Bai, C. W. Tung, and C. C. Lee: "Optimal design of loudspeaker arrays for robust cross-talk cancellation using the Taguchi method and the genetic algorithm", J. Acoust. Soc. Am., **117**, 5, pp. 2802-2813 (2005)
- 19) T. Samejima, Y. Sasaki, I. Taniguchi, and H. Kitajima: "Robust transaural sound reproduction system based on feedback control", Acoust. Sci. & Tech., **31**, 4 (2010)
- 20) Y. Huang, J. Benesty, and J. Chen: "On Crosstalk Cancellation and Equalization With Multiple Loudspeakers for 3-D Sound Reproduction", IEEE Sig. Proc. Letters, **14**, 10, pp. 649-652 (2007)
- 21) K. Matsui and A. Ando: "Binaural Reproduction of 22.2 Multichannel Sound over Loudspeakers", 129th Conv. Audio Eng. Soc., Prepr. 8272 (2010)
- 22) K. Matsui and A. Ando: "Binaural Reproduction of 22.2 Multichannel Sound with Loudspeaker Array Frame", 135th Conv. Audio Eng. Soc., Prepr. 8954 (2013)
- 23) K. Matsui, Y. Nakayama, M. Sugaya, and S. Adachi: "Binaural Reproduction over Loudspeakers Using Low-Order Modeled HRTFs", 137th Conv. Audio Eng. Soc., Prepr. 9128 (2014)
- 24) 菅谷真帆, 松井健太郎, 中山靖茂, 足立修一: "トランスオーラル再生のための対角化法による逆システムの構成", 信学技報 EA2014-123, pp. 285-290 (2015)
- 25) 石川健太郎, 徳住ゆりか, 丸田一郎, 足立修一, 松井健太郎, 安藤彰男:

* 評定者はどの評価音がどの再生法によるものなのか知ることはできないが、弁別が困難として高い評点がつけられていたのは、いずれも隠れ基準音と枠型 12 スピーカによるバイノーラル再生であった。

- “システム同定理論を用いた頭部伝達関数の三次元多方向同時推定”, 音響学誌, **69**, 7, pp. 321–330 (2013)
- 26) 足立修一: “MATLAB による制御のためのシステム同定”, 東京電機大学出版局 (1996)
- 27) 竹中裕司, 足立修一: “最小二乗法による多入力システム同定のための同定入力の生成法”, 計測自動制御学会論文集, **47**, 6, pp. 291–293 (2011)
- 28) 金井関利, 松井健太郎, 中山靖茂, 足立修一: “入力設計の改良による頭部伝達関数の多方向同時推定の高精度化”, 音響学誌, **71**, 7, pp. 316–318 (2015)
- 29) T. Sugimoto, S. Oode, Y. Nakayama, and H. Okubo: “Subjective Evaluation of Reproduction Method for Frontal Channels of 22.2 Multichannel Sound over a Direct-View Display”, ITE Trans. MTA, **3**, 1, pp. 67–75 (2015)
- 30) 藤本卓也: “低域バンドでの SN 比改善を目的とした TSP 信号に関する検討-高調波歪の除去-”, 音講論 3-P-8 (2000)
- 31) Rec. ITU-R BS.1534-2: “Method for the subjective assessment of intermediate quality level of audio systems”, International Telecommunication Union (2014)
- 32) ARIB STD-B32 2.2 版: “デジタル放送における映像符号化、音声符号化及び多重化方式”, 電波産業会 (2010)
- 33) Rec. ITU-R BS.1116-3: “Methods for the subjective assessment of small impairments in audio systems”, International Telecommunication Union (2015)

まつい けんたろう
松井健太郎 1998 年, 名古屋大学大学院工学研究科修士課程修了。同年, NHK に入局。2001 年より, 同放送技術研究所にて, 高臨場感音響, 音響信号処理に関する研究に従事。2013 年より, 慶應義塾大学大学院理工学研究科博士後期課程在学。学生会員。

おおいし さとし
大石 諭 2005 年, 早稲田大学大学院理工学研究科修士課程修了。同年, NHK に入局。2010 年より, 同放送技術研究所にて, 音響信号処理, 音響再生技術に関する研究に従事。2014 年より, 同広島放送局にて勤務。

すぎもと たけひろ
杉本 岳大 2001 年, 東京大学大学院工学研究科修士課程修了。同年, NHK に入局。2004 年より, 同放送技術研究所にて, 音響トランスデューサ, 音声圧縮符号化, 三次元音響の研究および MPEG, ARIB の標準化活動に従事。博士 (工学)。正会員。

もりた ゆういち
森田 雄一 2003 年, フォスター電機株式会社入社。戦略開発本部戦略商品開発部にて, スピーカシステム, 音響信号処理の開発に従事。

あだち しゅういち
足立 修一 1986 年, 慶應義塾大学大学院工学研究科電気工学専攻博士課程修了 (工学博士)。同年, (株) 東芝入社, 総合研究所に勤務。1990 年, 宇都宮大学工学部電気電子工学科助教授, 2002 年, 同教授。2003–04 年, ケンブリッジ大学客員研究員。2006 年, 慶應義塾大学理工学部物理情報工学科教授となり, 現在に至る。システム同定理論・制御理論とそれらの産業応用に関する研究に従事。1998 年, 日本機械学会賞 (論文) 受賞。2007, 2014 年, 計測自動制御学会賞 (著述賞) など受賞。計測自動制御学会, 電気学会, 日本機械学会, IEEE などの会員。