

音線法による室内音場の簡易バイノーラル可聴化に関する研究

37-246075 大林 紅音

1. はじめに

室内空間の VR 体験システムとして HMD を用いた視覚提示は既に実用に供されているが、それに連動する聴覚提示では室内音響特性の反映など再現精度に課題がある。

一方、音源や視聴者の移動に対応したリアルタイムでの音響シミュレーションや、波動音響による精度の高い音再現には現在技術的な限界がある。そのため、室内音響特性を反映しつつ簡易化することにより、リアルタイムの移動に対応した音再現システムを構築する必要がある。

そこで本研究では、まず VR 体験システムへの実装を目的とした簡易的なバイノーラル可聴化手法を構築し、被験者実験により再現音の評価を行う。さらに、直接音・反射音それぞれの処理について手法の再検討を行い、被験者実験によって検証した。

2. 動的可聴化システムの概要

音源のドライソースに室内インパルス応答 (RIR) を畳み込み、頭部伝達関数 (HRTF) を反映させる可聴化手順において、音源系・受音系の水平方向の指向性を前後左右の 4ch 信号により簡易的に反映する手法を構築する (図 1)。

2.1 ドライソースの録音・反射音インパルス応答の計算

ドライソースについて、無響室の中央に音源、全指向性マイクを水平面内直交 4 方向の音源中心から 1 m に設置し、4ch 同時録音を行う [1]。

また、室内の格子点上に受音点を配置した上で、各受音点において音源系・受音系 $4 \times 4 = 16\text{ch}$ の反射音 RIR を音線法により計算する。音源系・受音系ともに音圧カージオイド特性の指向性を直交 4 方向に設定する。図 2 に示す会議室をモデル化し、音源 1 点と受音点 60 点 (赤点) を配置し、COMSOL Multiphysics の音線法解析モジュールを用いて RIR を算出した。

2.2 ドライソースの 4 方向音源への分配

音線法における 4 方向の音圧カージオイド音源各々の正面の音圧を S (添字 F, B, L, R : 前後左右) とすると、それらを同位相と仮定して合成した 4 方向の音圧 S' は次式で表される。

$$\begin{pmatrix} S'_F \\ S'_R \\ S'_B \\ S'_L \end{pmatrix} = \begin{bmatrix} 1 & 0.5 & 0 & 0.5 \\ 0.5 & 1 & 0.5 & 0 \\ 0 & 0.5 & 1 & 0.5 \\ 0.5 & 0 & 0.5 & 1 \end{bmatrix} \begin{pmatrix} S_F \\ S_R \\ S_B \\ S_L \end{pmatrix} \quad (1)$$

一方、ドライソース 4 方向の音圧を \hat{S} とすると、これらが上式の合成音圧と等しくなるように、ドライソース 4ch 信号をカージオイド音源の各方向の信号に変換する必要がある。そのため、次式で示される Moore-Penrose 擬似逆行列を用いてドライソースの分配を行った。

$$\begin{pmatrix} S_F \\ S_R \\ S_B \\ S_L \end{pmatrix} = \begin{bmatrix} 0.625 & 0.125 & -0.375 & 0.125 \\ 0.125 & 0.625 & 0.125 & -0.375 \\ -0.375 & 0.125 & 0.625 & 0.125 \\ 0.125 & -0.375 & 0.125 & 0.625 \end{bmatrix} \begin{pmatrix} \hat{S}_F \\ \hat{S}_R \\ \hat{S}_B \\ \hat{S}_L \end{pmatrix} \quad (2)$$

2.3 直接音・反射音の処理

直接音については音源・受音点間の距離と方向関係から直接的に 4ch 受音信号を生成する [2]。

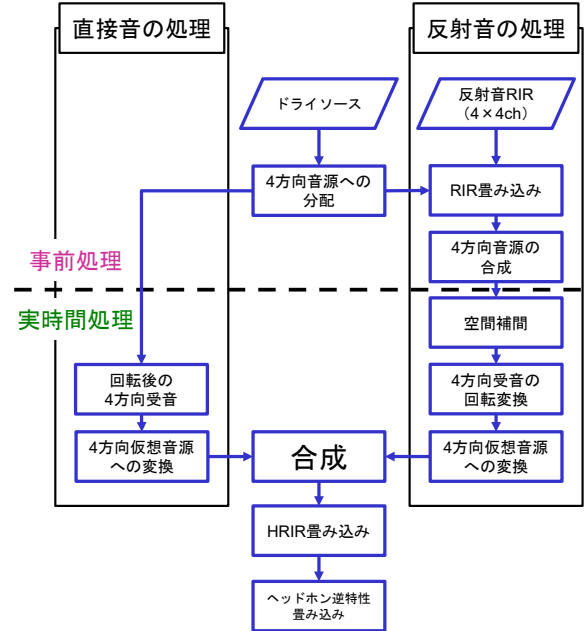


図 1 動的可聴化システムのフロー

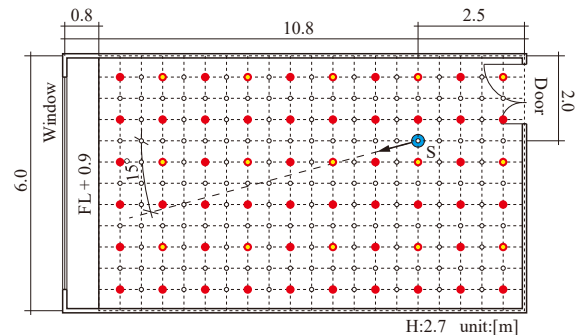


図 2 会議室における音源・受音点の配置

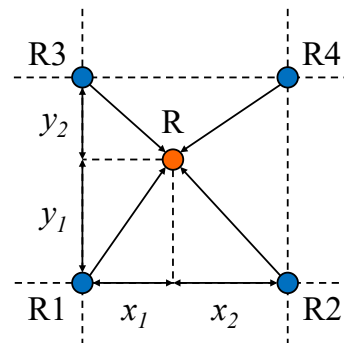


図 3 格子点からの空間補間

任意位置の受音点に対しては格子内の受音点 R の音圧信号 s を周辺 4 つの格子点 $R1 \sim R4$ の音圧信号 s_n から、エネルギー上の線形補間を仮定して、次式により作成する (図 3)。

$$s = \frac{\sqrt{x_2 y_2} s_1 + \sqrt{x_1 y_2} s_2 + \sqrt{x_2 y_1} s_3 + \sqrt{x_1 y_1} s_4}{\sqrt{x_1 y_1} + x_2 y_1 + x_1 y_2 + x_2 y_2} \quad (3)$$

2.4 バイノーラル信号への変換

得られた直接音と反射音の仮想音源信号を合成して頭部インパルス応答 (HRIR) を畳み込み、バイノーラル信号に変換する。最後にヘッドホン逆特性を畳み込む[2]。

3. 可聴化システムの再現性に関する聴感実験

3.1 VR空間における再現音の評価実験

3.1.1 実験概要

図 4 に示す会議室内の音源点 S にダミーヘッドを配置し、被験者は立位で実空間の実音（人工ロスピーカからの女性朗読音の再生音）とヘッドホンからの再現音の順に 13 秒ずつ聴取した。被験者数は 15 名（20～50 代の男性 11 名、女性 4 名）である。

(1) 実験 1

受音位置は図4に示す2カ所(R1, R2)と4方向(音源に対して正面・右・背面・左)の計8条件とした。被験者は実音に対する再生音の相対的な印象を5項目(音量・残響感・距離・音像の明瞭性・左右/前後方向)について7段階評定尺度法で評価した。

(2) 実験 2・3

被験者は音声が生産されている間、その場での回転（実験 2）もしくは所定経路の移動（実験 3）を行う。回転位置は 2 か所（R1, R2）である。経路は 3 通りであり、経路 1 では音源距離、経路 2 では音源指向性、経路 3 では壁との距離について検討する（図 4）。

被験者は実音と比較した際の再現音全体の違和感と、回転/移動中の再現音の変化の違いに関する 6 項目（音量・音色・響き・音像の距離・音像の明瞭性・音像の方向感）を 4 段階評定尺度法で評価した。

3.1.2 実験結果

(1) 実験 1

平均評点 0 に対する有意差を t 検定により検定し、平均評点と標準偏差を図 5 に示す。

音量は、全ての条件で小さく感じられ、R2 では音像が実音より遠く感じられた。これらの傾向は、仮想音源の分配方法や、再現音の周波数特性の変化が影響していると考えられる。

(2) 実験 2・3

実験 2・3 の各項目における全被験者の各回答の割合を図 5・図 6 に示す。実験 2 では全項目でほとんどの被験者が「感じない」「やや感じる」と回答した一方、音量では「とても感じる」という回答も見られた。再現音全体の違和感に関しては、実験 2 より実験 3 のほうが「感じる」と回答した割合が多くなっている。自由記述では、再現音は方向感が分かりにくいという回答が見られた。

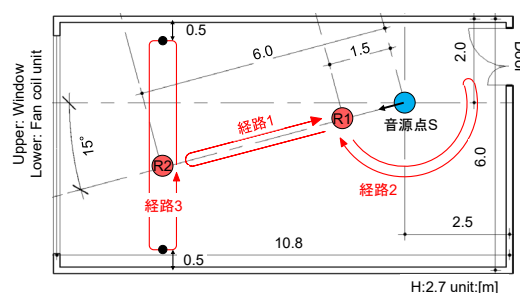


図 4 会議室における音源・受音点の配置と
実験 3 における移動経路

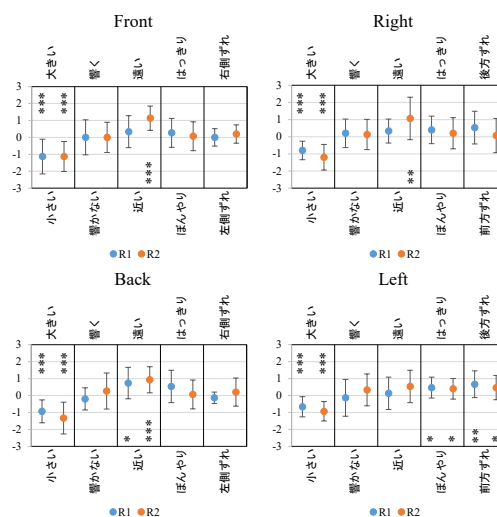


図5 実験1の平均評点と標準偏差

(*: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$)

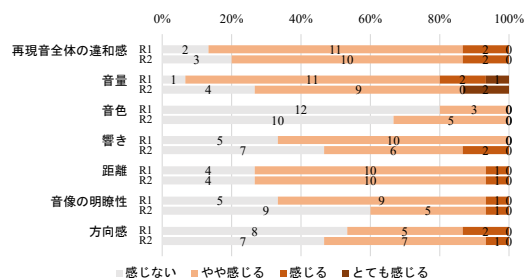


図 6 実験 2 の各回答の割合

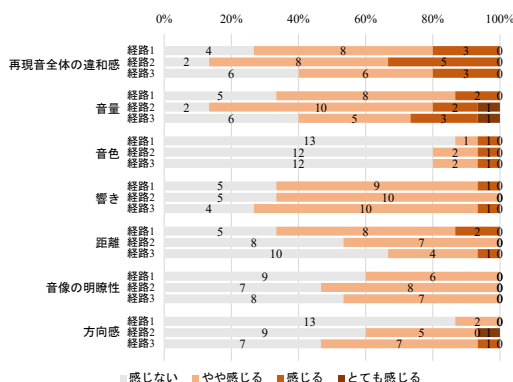


図7 実験3の各回答の割合

4. 直接音処理手法の検討

4.1 直接音の仮想音源変換手法の修正

4 方向音源からなる直接音に対する仮想音源変換について、以下の 4 つの手法を検討する[3]。

- ・擬似逆行列を用いた 4 方向分配 (4ch)

4ch ドライソースから音源と受音点の位置関係と指向性を反映した 4ch カージオイド受音信号を計算する。擬似逆行列を用いて水平面内直交 4 方向の仮想音源信号へ変換する。

- ・擬似逆行列を用いた 3 方向分配 (3ch)

前項同様、4ch カージオイド受音信号を計算後、擬似逆行列を用いて 3 方向の仮想音源信号を計算する。仮想音源の分配方向は受音点の左右に加えて、音源が前方にある場合は正面、後方にある場合は背面とする。

- ・擬似逆行列を用いた 2 方向分配 (2chA)

前項同様、4ch カージオイド受音信号を計算後、擬似逆行列を用いて 2 方向の仮想音源信号を計算する。仮想音源の分配方向は直交 4 方向のうち音源方向を挟む 2 方向とする。

- ・パンニングによる 2 方向分配 (2chB)

4ch ドライソース音源から単純に線形パンニングにより、直交 4 方向のうち音源方向を挟む 2 方向の仮想音源に直接分配する。

4.2 直接音の仮想音源変換に関する方向定位実験

4.2.1 実験概要

仮想音源変換手法の方向定位への影響を検証するため、無響室内の仮想音源方向にスピーカを配置し、4 つの手法による再生音の方向定位実験を行った[3]。

自由音場において受聴者の頭部中心から水平面に 1.5 m 離れた円周上に 30 度間隔で音源が配置されている状況を想定し、12 方向分の仮想音源信号を作成する。音源自体は常に受聴者向きとし、無響室で 4ch 収録した女性音声(5 s)をドライソースとして、12 方向分を計算した。聴感実験ではこの仮想音源信号を直交 4 方向のスピーカに割り当て同時再生した。

被験者は直交 4 方向のスピーカの中心に着席して再生音を聴取し、音の聞こえる方向を 12 方向から選択回答する。冒頭に基準条件として正面 0 度の音を聞いた後、本実験として 12 方向のランダム順の提示を 3 組、計 36 回の提示・回答を繰り返す。この一連のセッションを 4 つの手法およびバイノーラル再生について実施した。被験者数は 7 名(20 ~ 40 代の男性 6 名、女性 1 名)である。

4.2.2 実験結果

全被験者の提示方向に対する回答方向の回答数分布を図 8 に示す。図中の赤線：正答、緑線：前後誤判定、青線：左右誤判定に相当する。手法間の比較では、4ch は左右 90、270 度付近で前後に回答がずれる傾向がみられるが、他の 3 手法では全体的に誤判定が少なくなっている。また、4 方向分配の仮想音源変換による(a)と(e)の比較から、バイノーラル再生では前後誤判定が顕著に増加することが確認できる。これは頭部伝達関数(HRTF)の個人差の影響だと考えられる。

前後正答率の被験者平均値は 2chB、3ch、2chA、4ch の順で低下し、4ch では 65%、他の 3 手法では 75~80% となった。なお、4ch 分配のバイノーラル再生では約 50% と顕著に低下している。

5. 反射音の空間補間についての検証

反射音の妥当な補間方法を明らかにするため、音量に関する数値的検討と聴感実験による検討を行う。

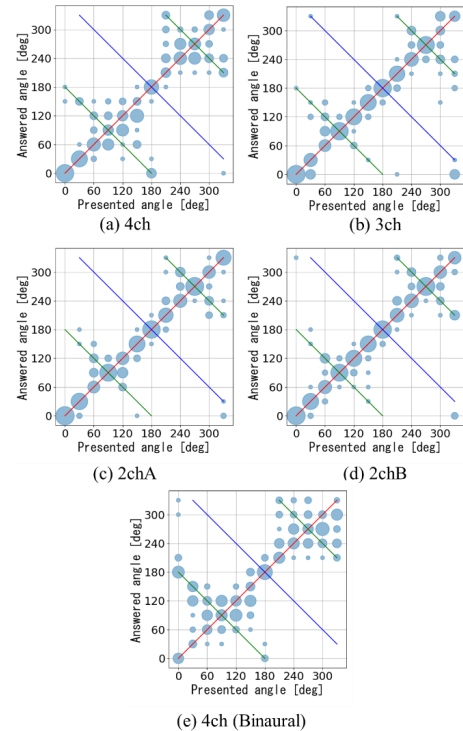


図 8 提示方向に対する回答方向の回答数分布

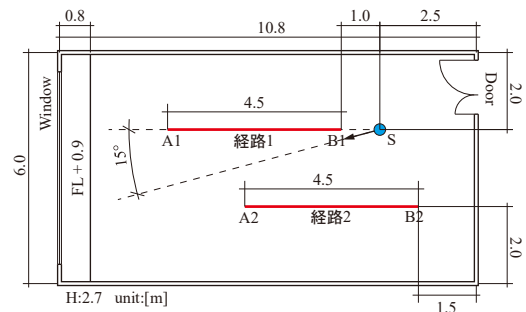


図 9 移動経路

表 1 評価項目

評価対象	評価尺度						
	-3	-2	-1	0	1	2	3
	とても		やや	どちらでもない	やや		とても
音量	小さい				大きい		
響き	響かない				響く		
高音	弱い				強い		
低音	弱い				強い		
なめらかさ	不連続な				連続な		
ノイズ	少ない				多い		

5.1 反射音の空間補間に関する数値的検討

5.1.1 検討概要

図 2 に示す会議室において、高さ 1.5 m に音源点と 50 cm 間隔の格子点 209 点を配置し、反射音 RIR を計算する。

空間補間に関しては 50 cm 間隔の受音信号を基準として、図 2 に示す 1 m (赤) または 2 m (黄) 間隔に間引いた格子点から補間した信号の音圧レベルを比較する。2 m 間隔では下側 2 行と左側 2 列については補間の対象外とする。

5.1.2 結果と考察

全指向性音源点から全指向性受音点の RIR について、オクターブ帯域毎の空間補間の影響を検証する。図 10 に 125、500、2k Hz オクターブ帯域の計算結果を示す。

基準となる 50 cm 格子点上の反射音 RIR の音圧レベルを左列に、左列に対する 1 m・2 m 間隔の格子点から補間した音圧レベルの差を中央列・右列に示す。低音域で最大 5 dB 程度の増減が生じており、1 m 間隔より 2 m 間隔の方が影響は顕著である。波長と同程度の間隔の格子点では、補間に用いる 4 点の音圧の相関性が高く、補間による位相干渉により増減が大きくなるものと考えられる。

5.2 反射音の空間補間に関する聴感実験

5.2.1 実験概要

反射音のみの VR を作成し、被験者は各条件について室内を移動した際の音声を評価する。被験者数は 16 名（20～30 代の男性 5 名、女性 11 名）である。

実験は音源種類 3 条件（女性朗読音、250 Hz バンドノイズ、2 kHz バンドノイズ）、移動経路 2 条件（図 9）、補間 5 条件（50 cm 補間なし・1 m 補間あり/なし・2 m 補間あり/なし）の計 30 条件実施した。

表 1 に評価項目を示す。

5.2.2 実験結果

全項目について、移動経路・音源種類・補間を要因とした三元配置分散分析を行った結果を表 2 に示す。

補間条件についてはノイズのみ有意差が見られ、50 cm 補間なし > 1 m 補間なし > 2 m 補間なし > 1 m 補間あり > 2 m 補間ありの順にノイズが多い結果となった。これは、グリッド間を跨ぐ際に発生するノイズの影響と考

えられる。一方、ノイズ以外の評価項目では補間方法の違いによる有意差はほぼ見られなかった。

6. まとめ

VR 体験システムへの実装を目的とした簡易的なバイノーラル可聴化手法を構築し、実験による検証を行った。

VR 空間においての再現音の評価実験では、実音に対して再現音の音量が小さく、音像の距離感が遠く感じられる傾向が見られた。直接音の方向定位実験では、前後誤判定の傾向が見られ、仮想音源変換処理の変更により改善されたものの、HRTF の影響も大きいことが判明した。反射音の空間補間に関する検証では、ノイズ以外の評価項目では補間方法の違いによる有意差はほぼ見られなかった。

参考文献

- [1] 大林他, 音講論 (秋), 649-652, 2024.
- [2] 小松他, 音講論 (春), 533-536, 2025.
- [3] 小松他, 音講論 (秋), 689-690, 2025.

表 2 分散分析の結果

(* : $p < 0.05$, ** : $p < 0.01$, *** : $p < 0.001$)

評価項目	主効果			交互作用		
	経路	音源	補間	経路 * 音源	経路 * 補間	音源 * 補間
音量	***	***		**		
響き	***	***				
高音	*	***				
低音		***				
なめらかさ		***				
ノイズ	***	***	**	*	*	

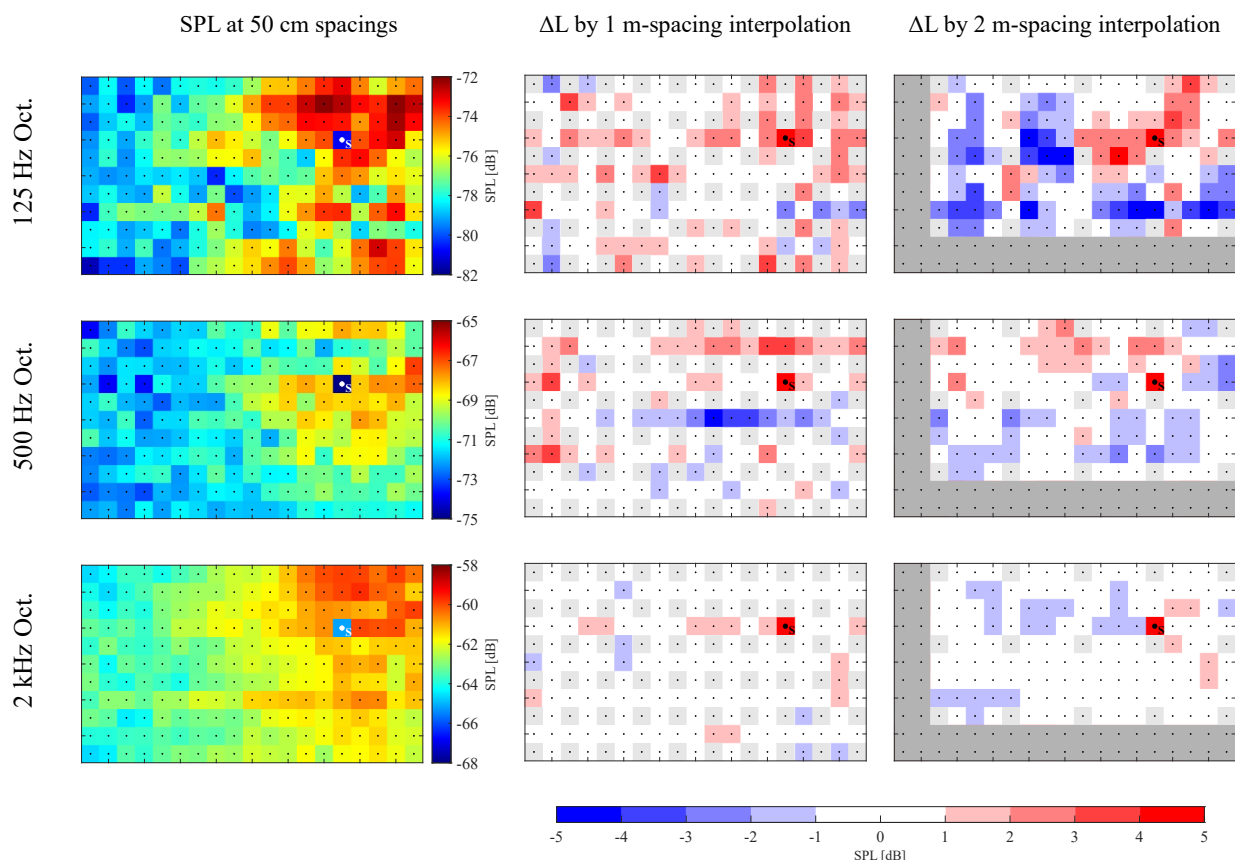


図 10 補間前の音圧レベル（左）、補間前後の音圧レベル差（中央：1 m 間隔、右：2 m 間隔）