基礎論文

X-Media Galaxy における移動音像実現のための音像補間

伊藤仁一[*1*3]，中山雅人[*1]，西浦敬信[*1]，木村朝子[*1]，柴田史久[*1]，田村秀行[*2]

Sound Image Interpolation for Moving Sound Image Implementation in X-Media Galaxy

Yoshikazu Ito[*1*3], Masato Nakayama[*1], Takanobu Nishiura[*1]
Asako Kimura[*1], Fumihisa Shibata[*1], and Hideyuki Tamura[*2]

Abstract – We have developed an MR system that merges the real and virtual worlds both audibly and visually. To achieve the audio MR, we developed the "Acoustic Planetarium," which is composed of multiple parametric loudspeakers. In this paper, we propose a method to position moving sounds using this system. Firstly, in order to select the appropriate positioning method, we compared an energy sum constant method and an amplitude sum constant method, which interpolate the sound position between multiple sounds located by the parametric loudspeakers. We also used the full MR system, which uses both audio and visual information, to verify the accuracy in perception of the moving sound. In order to do so, we displayed a CG image at the origin of the sound with the help of a head mounted display. The results showed that the energy sum constant method was appropriate for the parametric loudspeaker and the accuracy of the perceived sound was adequate for the full MR system.

Keywords: Mixed Reality, Audio and Visual, Parametric Loudspeaker, Moving Sound Image

1. はじめに

現実世界と仮想世界を実時間で融合する複合現実感（Mixed Reality; MR）の研究の多くは，視覚的な融合に限られているが，我々は視覚的 MR だけでなく聴覚的 MR を同時に実現し，視覚・聴覚両面で現実世界と仮想世界を矛盾なく融合する「2×2 方式視聴覚併用 MR システム」の研究･開発を行ってきた．これを実現するには，(1) 視覚的には現実世界の光景と電子的に生成した CG 映像の重畳合成，(2) 聴覚的には現実世界で発した実音とコンピュータで生成した人工音（生成音）の混合，(3) 上記の映像と音像の複合現実空間における実時間幾何学的整合性を達成する必要がある．

その第 1 歩として開発したシステム [1] では，視覚的 MR にはビデオシースルー方式の HMD (Head Mounted Display)，聴覚的 MR には開放型ヘッドホンを用いたバイノーラル再生方式を採用した．一般にバイノーラル再生方式は，予め受聴者の頭部形状を計測し，頭部伝達関数（HRTF）を用意しておく必要がある．体験者が限定される場合は，この方法で問題ないが，不特定多数の参加者に MR システムを体験させたい場合には，その場で瞬時に個々の体験者全員の HRTF を用意することは事実上不可能である. 多数の受聴者用に一般化された HRTF を用意し，体験者全員が同じ HRTF を使用する方法 [2][3] を試みたが，正しく 3 次元音場を知覚できない体験者が多かった. 次に, 複数の HRTF を用意し，体験者の好みに合う HRTF を選択・接合する方式 [4] を試みた．かなりの改善はあったものの，実用的に不特定多数を満足させられるレベルには達しなかった．

そこで我々は，この問題を抜本的に解決すべく，全く新しい音像定位方式を導入することにし，複数の超音波スピーカから放射される音を, 壁面, 天井, 反射板などの実物体に反射させることで，その反射面上に音像を構築する「音像プラネタリウム方式」[5] を考案し，その基幹ユニット実機を開発した．

この方式はスポットライトのように狭く限定された範囲に音を放射可能で，壁面などに反射させても指向特性は失わないという超音波スピーカの性質を利用したものである．ヘッドホンを装着する必要が

*1 立命館大学大学院理工学研究科
*2 立命館大学総合科学技術研究機構
*3 現在, 株式会社ニコン
*1 Graduate School of Science & Engineering, Ritsumeikan University
*2 Research Organization of Science and Technology, Ritsumeikan University
*3 Nikon Corp.

なく，頭部伝達関数のような個人性による問題が生じない，複数人が実音と人工音を同時に受聴することを可能にする．また，超音波の反射を利用することで，スピーカを配置していない方向や場所にも音像を構築することが可能である．

このような長所をもつ同方式を採用した視聴覚併用 MR システムを「X-Media Galaxy」（クロスメディア・ギャラクシーと読む）と呼ぶことにし，その効果を体験できる応用事例を開発した [6].

音像プラネタリウム方式は，全く新しい原理に基づく，自由度の高い 3 次元音場生成方式として，注目を集め，今後の発展が期待されているが，現状ではまだ欠点がいくつか存在する．例えば，音像を反射面に構築するため，音像位置を自由に移動させることができず，現段階では固定音の実現に留まっている．我々は，大きな可能性をもつ「音像プラネタリウム方式」を一歩ずつ前進させるべく，現有の X-Media Galaxy で，体験者に移動音を感じさせる方法を検討した．

超音波スピーカを用いて音像を移動させる方法として，まず超音波スピーカの放射方向を動的に制御することが考えられる [7-10]．例えば，木村ら [7] は回転放物面（パラボラ）に向けて超音波スピーカから音を放射し，スピーカ方向を動的に制御することで，任意の位置に音像提示する大画面スクリーン向けの放射拡散方式音像提示システムを提案している．パラボラによる音の拡散とスピーカ方向制御により良好な音像提示が実現可能であるが，方向制御のためのモータ駆動音やモータ制御による遅延などの問題がある．

一方，スピーカ自体を動的に制御せずに 3 次元音場を実現する方法として，従来のダイナミックスピーカ（以降，スピーカと表記する）を用いた場合では，5.1ch サラウンドシステムに代表されるステレオ効果を利用したものがある．近年では超音波スピーカを用いたステレオ再生 [11] や複数台の超音波スピーカを用いた音像位置補間提示 [12] が研究されている．これらは，スピーカと同様に音波の位相差やレベル差を付与することで音像位置を移動させており，水平方向に関して音像位置を移動させることが可能なことがわかっている．しかし，これらの研究では，左右または上下の補間ができることは示されているが，任意の位置に音像を提示したい場合に複数の超音波をどの程度の強さで射出すればよいかについて明らかになっていない．

音像位置を補間する手法であれば，モータ制御が不要であり，実時間性が要求される MR システムには都合が良い．更に，視聴覚併用 MR システムとして利用する際は，**図 1** のように，実物から発せられる実音（ギター）に加えて，音像プラネタリウム装置を利用して音像（小鳥の鳴き声）を定位させ，その位置にその音源の仮想物体（小鳥の CG）を重畳描画することができるため，視覚と聴覚の相乗効果により音像の定位性能が向上することが期待できる．

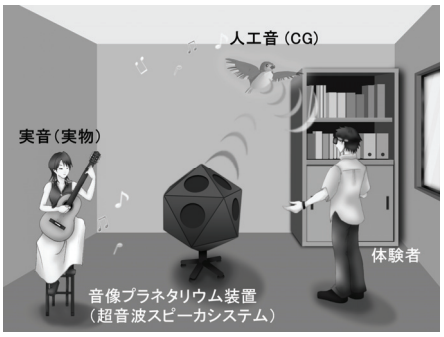


図 1　音像プラネタリウムを用いた視聴覚併用 MR システムのイメージ（体験者は実物のギターからの実音と CG の小鳥からの人工音を同時に受聴可能）

Fig. 1 Image of the audio-visual MR system using acoustic planetarium

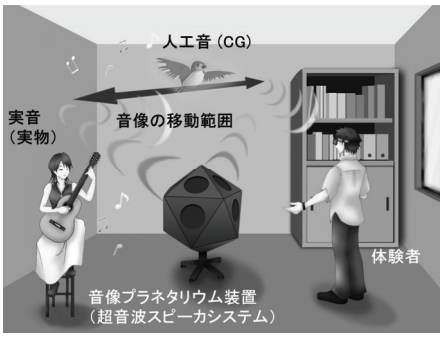


図 2　音像を補間して移動音像を実現

Fig. 2 Implementation of moving sound image by interpolating multiple sound sources

そこで本論文では，複数台の超音波スピーカを利用する音像プラネタリウムを用いて音像位置補間を行い，3 台の超音波スピーカで構成される三角形を組み合わせていくことで，全天周での移動音像を実現する方法を提案する．また，任意の位置に音像を提示には，複数の超音波をどの程度の強さで射出すればよいかや視覚と聴覚の相乗効果について明らかにするために，超音波スピーカに適した音像の補間方法と MR 視覚刺激を付与した場合の定位性能の変化を確認する基礎実験を行う．本論文では，まず**図 2** に示すような体験者 1 名，体験位置固定という条件下での音像の移動を目指す．

以下，2 章で移動音像の実現方法と補間方法，について述べ，3 章では 4，5 章の実験準備と，後の実験と比較するために実施した，スピーカの音像位置を回答させる予備実験について説明する．4 章では，超音波スピーカに適した補間方法を確認するために，

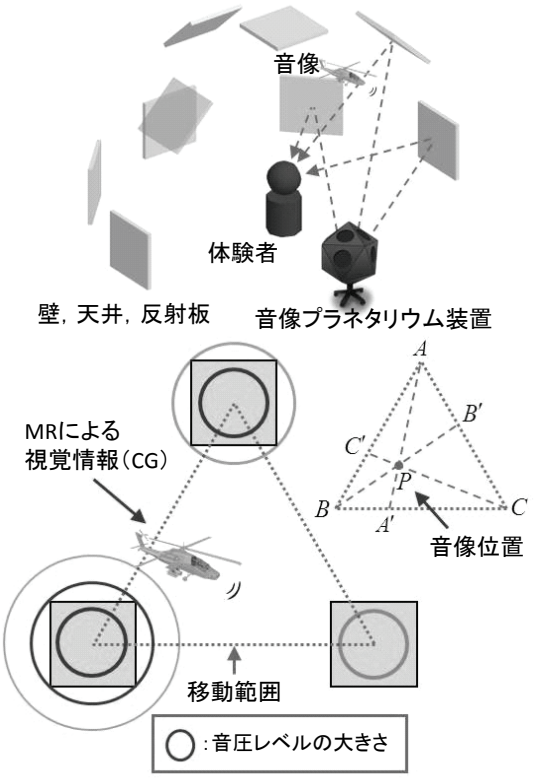


図 3　反射面の構築例と 3 点間補間

Fig. 3 Construction example of the reflecting points and interpolation between three points

補間方法の違いによる音像位置の定位性能を調べる基礎実験（実験 1）を行う．更に 5 章で MR による視聴覚刺激を併用した場合の相乗効果について基礎実験（実験 2）を行い，6 章で提案手法を実現する上での最小構成となる 3 台の超音波スピーカを利用して，移動音像を実現した結果について述べる．

2. 移動音像の実現方法

2.1 基本方針

全天周を移動する音像を実現するため，まず体験者を囲む全天周に複数の反射面を構築し，超音波スピーカからこれらの反射面に音を放射する．その反射面間の音像位置を補間することで，反射面間を自由に移動可能な音像を提示する（**図 3**）．この際，音像の反射面は，受聴者を囲むように複数の反射板を配置するか，またはドーム形状の壁面を利用することで構成する．

具体的には，3 つの反射面を頂点とする三角形平面上に対して音像位置補間を行い，この 3 点間での補間を複数組み合わせることにより音像を多面体の平面上に移動させる．

2.2 音像位置の補間方法

一般に，人間が音像方向を判断する手がかりは頭部形状による両耳間時間差および両耳間レベル差であることが知られており [13][14]，これらを再現することで音像方向を制御可能である [15]．ただし，時間差を利用した音像位置の補間の場合，良好な補間を行うために受聴者の耳の位置を固定する必要がある．この制約は MR 空間内の体験者の行動を大きく制限してしまい，好ましくない．一方レベル差を利用した方法では，体験位置こそ固定されるものの，体験者の頭部移動にもある程度耐えうると考えられ，MR アトラクション作成時の自由度も位相差と比べ高い．そこで，本研究では音像位置の補間方法として，複数の反射面間での音圧レベル差を用いる方法を採用する．

具体的には，複数の超音波スピーカから各反射面（壁など）上に音を放射することで，構築される音像の音圧レベルを調節し，反射面間の音像位置を補間する．この音像位置の補間は 3 つの反射面を基準として行い，3 点間での補間を組み合わせる．

3 点間での音像位置の補間は，3 つの基準点と提示する音像位置との距離により算出される比をもとに，基準となる各反射面での音圧レベルを調節する．このとき，まず一般的な補間方法として，音波のエネルギー和が一定になるように考慮する方法が考えられる [16]．これは，単純に左右の音波の振幅の和が一定になるように振り分けてしまうと音像が中央に定位したときに，聴覚上の音量が下がってしまうという問題を解決するためである．一方，頭部伝達関数を用いた立体音響では未測定の音源方向の音を頭部伝達関数から推定し求める．このときの補間には線形補間が良い場合がある [2]．このように，従来のスピーカにおいてはそれぞれの補間法が有効な場面が知られている．

図 3 で示した提示する音像位置 P と 3 つの基準点 A, B, Cでの入力信号の増幅率を S_A, S_B, S_Cとしたとき，エネルギー和を考慮した補間式は，

$$S_A = \sqrt{\frac{PA'}{AA'}} \quad S_B = \sqrt{\frac{PB'}{BB'}} \quad S_C = \sqrt{\frac{PC'}{CC'}} \qquad (3.1)$$

と定義できる（以降，エネルギー和一定の式と表記する）．また，単純に振幅の和を一定にすることを考慮した補間式は，

$$S_A = \frac{PA'}{AA'} \quad S_B = \frac{PB'}{BB'} \quad S_C = \frac{PC'}{CC'} \qquad (3.2)$$

となる（以降，振幅和一定の式と表記する）．

本研究では通常のスピーカと放射特性が異なり，鋭い指向性を持つ超音波スピーカを用いている [17]（**図 4**）．超音波スピーカを用いた場合に適した補間方法についてはほとんど報告がないため，まずスピーカと同様エネルギー和を一定にする補間式と

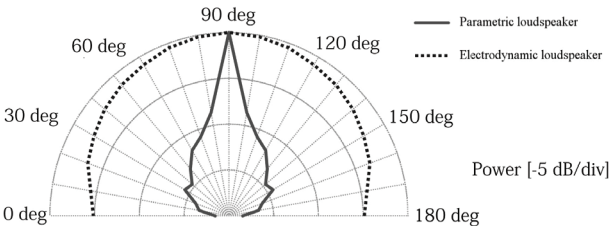
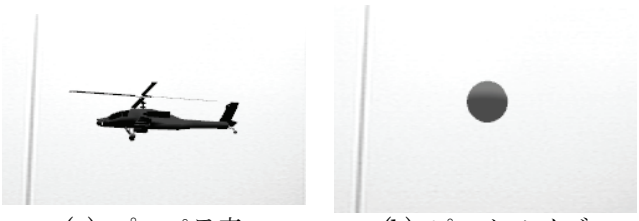


図 4　放射特性

Fig. 4 Radiation characteristics



(a) プロペラ音　　(b) ピンクノイズ

図 5　音源に対する視覚情報

Fig. 5 Visual information for sound source

単純に振幅の和を一定にする補間式で，それぞれの音像提示に差異があるのか，どのような傾向があるのかを，実験 1（4 章）で確認し，超音波スピーカにより適した補間方法を採用することにする．

3. 実験準備と予備実験

3.1 実験準備

【実験条件】

3 点間での補間は 2 点を基準とした補間の拡張であるため，実験 1，2 では 2 点間での補間実験を行う．また，移動音は固定音の連続的な遷移であると考えられるため, 固定音による主観評価実験を行う．3 点間での補間による移動音源の実現については, 6 章で述べる．

人間の音像定位能力は，水平方向と比べ垂直方向が劣っていることが知られており，我々の先行研究 [18] でも同様の結果が得られている．そこで，本実験では人間の音像定位能力がより高く，手法の違いによる差が大きく影響すると考えられる水平方向に対して実験を行う．

予備実験および実験 1，2 での音像提示方法は，被験者の前方に配置されたスピーカをそれぞれ実際に出力する場合（複数スピーカ（切替)），被験者前方の左右端のスピーカ 2 台を用いて音像位置補間をする場合（2ch スピーカ（補間)），被験者前方の 2 台の超音波スピーカを用いて音像位置補間をする場合（2ch 超音波スピーカ（補間)）の 3 通りである．

提示する音像位置は，スピーカの設置位置と同様 7 箇所で, 音像は提示時間 1 秒の音を 0.5 秒間隔で 3 回繰り返して提示する．この 3 回の提示を 1 セットの音像提示とする．

音像位置の補間式は，例えば図 3 の BC 2 点間での補間の場合，2.2 のエネルギー和一定の式 (3.1) と振幅和一定の式 (3.2) をもとに，それぞれ以下のようになる．

$$S_B = \sqrt{\frac{PC}{BC}} \qquad S_C = \sqrt{\frac{PB}{CB}} \tag{3.1'}$$

$$S_B = \frac{PC}{BC} \qquad S_C = \frac{PB}{CB} \tag{3.2'}$$

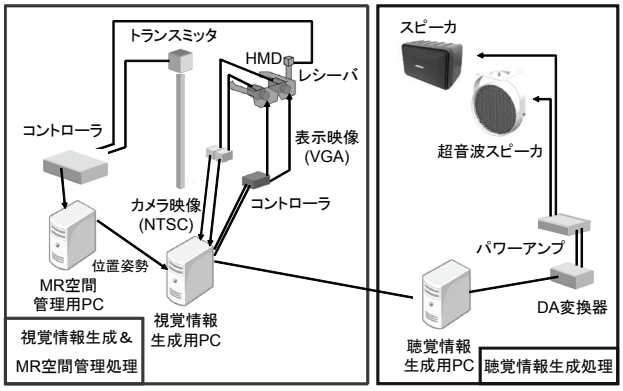


図 6　システム構成

Fig. 6 System configuration

実験中，被験者は常に HMD を装着するが，視覚提示の条件として， MR による視覚情報 (CG) を提示しない場合（CG なしの場合）と 提示する場合（CG ありの場合）の 2 つの条件を設けている．使用する音源は，ヘリコプタのプロペラ音とピンクノイズの 2 種類で，プロペラ音にはヘリコプタの CG を，ピンクノイズには灰色の球の CG を視覚情報として提示する（**図 5**）.

【実験環境】

次節で述べる予備実験および実験 1，2 で用いる視聴覚併用 MR システムの構成を**図 6** に示す．システムは大きく分けて視覚的 MR の処理と MR 空間を管理する視覚情報生成兼 MR 空間管理処理部と聴覚的 MR を処理する聴覚情報生成処理部の 2 つの処理部から構成される．聴覚的 MR を提示する機器としてスピーカと超音波スピーカを利用する．また，実際に視聴覚併用 MR システムを体験する状況を想定し，全ての実験で，被験者はビデオシースルー型 HMD (Canon 製 VH-2002) を装着することとする．

実験時の機材配置を**図 7** に示す．このとき，被験者の前方にスピーカと超音波スピーカの両方を配置し，実験における被験者の視界（HMD 越し）に入る視覚情報を統一している．これは音像提示機器の種類を変更した際に，機器の配置の変更による視覚的な音像定位の影響を防ぐためである．スピーカの高さは 1.15 [m], 超音波スピーカの高さは 1.36 [m]

である．各スピーカから提示される音は，ピンクノイズ，プロペラ音が，被験者の頭の位置でそれぞれ 43 [dB]，46.7 [dB]となるように設定している．

実験で使用したシステムの処理速度は約 60 [fps]，実験環境は暗騒音 43 [dB]のオフィス環境で，被験者は正常な聴力をもつ学生 12 名（男性 9 名，女性 3 名）である．

【実験手順】

実験の手順はすべて共通で，次の手順で行う．

(1) 被験者を椅子に座らせる
(2) 被験者の耳の高さを一定に調整
(3) 回答位置について説明
(4) 頭を動かさない・目を瞑らないように指示
(5) ランダムに音像提示を行う
(6) 被験者に音像の位置を回答させる
(7) 条件を変えて (5)(6) を繰り返す
(8) 全パターン終了後，コメントを聴取

手順 (2) では，椅子の高さを調整することで被験者の耳の高さを音像提示機器と同じ高さに調整する．手順 (4) では，実験条件を同じにするために，被験者に音像提示中は頭を動かさない，また目を開いた状態で音像を聞くよう指示した．

3.2 予備実験

4 章，5 章の実験結果と比較する基準とするために，実際の音源位置に設置した複数台のスピーカから音を提示する「複数台スピーカ（切替）」方式により予備実験を行う．

【実験条件】

音像提示は 7 台のスピーカで行い，左のスピーカ位置から 1～7 と番号をつけ，被験者に音が出力されていると感じるスピーカを番号で回答させる．被験者は HMD を装着し音像位置を回答するが，MR による視覚提示（CG 提示）は行わない．また，1 セットの音像提示で必ず音像位置を回答させる（音像位置がわからないという回答を許さない）．音像を提示するスピーカ（位置）は，7 台のスピーカから各々2 回ずつ提示され，提示順はランダムとなるようにする．以上の実験を，被験者の半数はプロペラ音，ピンクノイズの順，残りの半数はピンクノイズ，プロペラ音の順で行う．また，被験者には実験開始前に練習として 3 セット分の音像提示を行う．

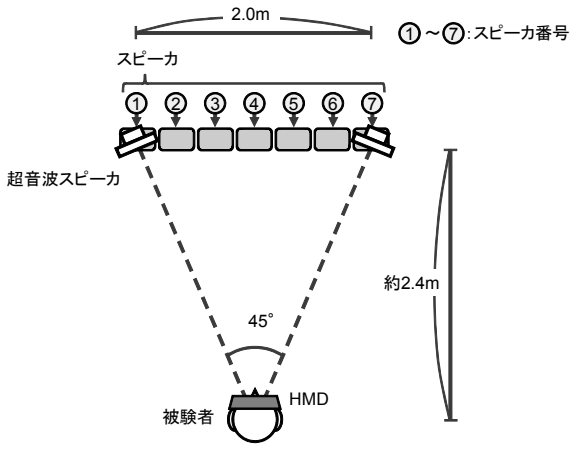


図 7　機材配置

Fig. 7 Equipment placement

【結果と考察】

音源ごとの音像の提示位置と被験者の回答位置の分布を**図 8** に示す．図 8 は，被験者 12 名，各スピーカからの提示回数 2 回，各提示位置につき計 24 回の回答を母数としたバブルグラフで，白色の円は正解を灰色の円は不正解を表している．円の大きさは回答数に比例しており，円に重ねて記載した数値は，各回答位置と回答された回答数である．図より，多少のばらつきはあるものの，被験者は音源の種類によらず概ね音像の位置を正しく回答できていることがわかる．

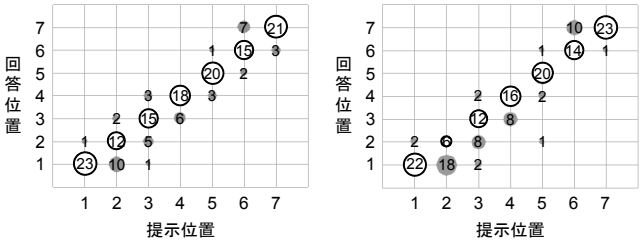
以降の実験で，この実験結果と同程度の定位性能が得られれば，現実世界とほぼ同等の（十分な）定位性能を実現できていると考えることができる．

4. 音像位置補間に関する主観評価実験（実験 1）

4.1 目的と条件

音圧レベル差を利用した音像位置制御においてスピーカを用いた場合と超音波スピーカを用いた場合でどの程度定位性能に差が生じるのか，また補間式の違いによりどのような傾向があるのかを確認する．

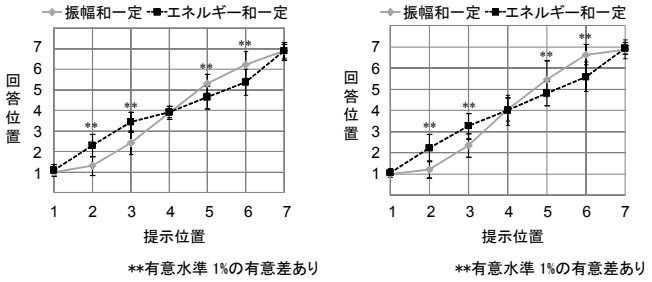
「2ch スピーカ（補間）」および「2ch 超音波スピーカ（補間）」方式により 7 台のスピーカの位置にランダムに音像を提示し，被験者には予備実験と同様音像が定位していると感じる位置のスピーカ番号を回答させる．被験者には HMD を装着させるが，CG 提示は行わず，音源の変更順も予備実験と同様である．ただし，「2ch スピーカ（補間）」と「2ch 超音波スピーカ（補間）」の実験は，それぞれ別日に行った．また，エネルギー和一定の式 (3.1') と振幅和一定の式 (3.2') の 2 種類の補間式による音像提



(a) プロペラ音　　(b) ピンクノイズ

図 8　音像位置と回答の分布（複数スピーカ（切替））

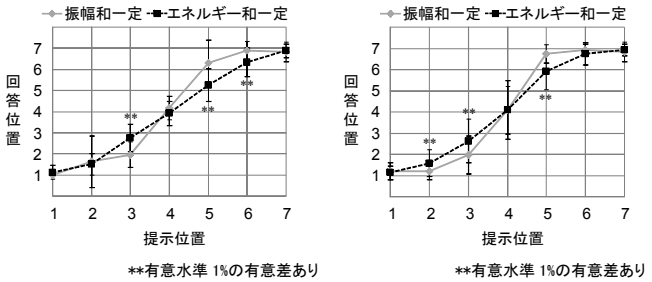
Fig. 8 Distribution of answers to the sound image position (multi dynamic loudspeakers (switching))



(a) プロペラ音　　(b) ピンクノイズ

図 9　音像位置と回答の平均
（2ch スピーカ（補間））

Fig. 9 The average of the answers
to the sound image position
(2ch dynamic loudspeakers (interpolation))



(a) プロペラ音　　(b) ピンクノイズ

図 10　音像位置と回答の平均
（2ch 超音波スピーカ（補間））

Fig. 10 The average of the answers
to the sound image position
(2ch parametric loudspeakers (interpolation))

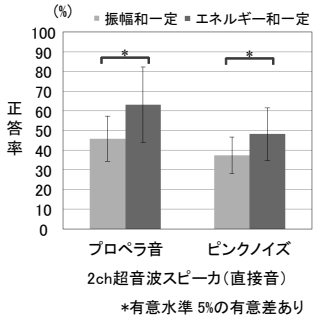
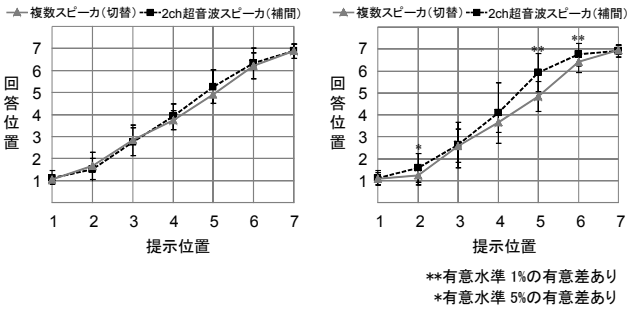


図 11　正答率の平均
（2ch 超音波スピーカ（補間））

Fig. 11 The average percentage of correct answers
(2ch parametric loudspeakers (interpolation))



(a) プロペラ音　　(b) ピンクノイズ

図 12　音像位置と回答の平均
（複数スピーカ　vs 2ch 超音波スピーカ（補間））

Fig. 12 The average of the answers
to the sound image position
(multi dynamic loudspeakers (switching) vs
2ch parametric loudspeakers (interpolation))

示，および音像の提示位置はランダムに提示され，式と提示位置の全組合せ（2 種類×7 箇所）に対して，それぞれ 2 回ずつ提示する．

4.2 結果と考察

まず補間式の違いによる傾向を比較するため，補間式ごとの被験者の回答位置の平均，標準偏差，t 検定の結果を**図 9, 10** に示す．図 9 はスピーカ（補間），図 10 は超音波スピーカ（補間）の実験結果である．まず，図 9，10 より，左右端（スピーカの位置，提示位置 1 と 7）および中央（提示位置 4）の定位性能については，いずれの手法でも比較的高いことが分かる．また，被験者の回答の平均および有意差検定の結果をみると，2ch スピーカ（補間）の場合，エネルギー和一定の式では中央に，振幅和一定の式では左右端に音像が定位する傾向がみられる．しかし，2ch 超音波スピーカ（補間）の場合では，図 10 から補間式にかかわらず左右端に音像が定位しやすくなっていることがわかる．このことから，超音波スピーカの放射特性による音像定位への影響が顕著にあらわれていることがわかる．また，振幅和一定の式よりもエネルギー和一定の式が音像を提示した位置に被験者の回答の平均が近く，理想的な結果であるといえる．2ch 超音波スピーカ（補間）の場合の被験者の正答率の平均を比較した結果についても，振幅和一定とエネルギー和一定では有意差が得られている（**図 11**）．以上より，超音波スピーカにおける音像位置補間でも，エネルギー和一定の式が有効であるといえる．

次に定位性能の高かった 2ch 超音波スピーカ（補間）でエネルギー和一定の式を採用した場合の結果と予備実験（複数スピーカ（切替））の結果を比較する．それぞれの被験者平均，標準偏差，t 検定を行った結果を**図 12** に示す．プロペラ音に関しては 2ch 超音波スピーカ（補間）の場合と複数スピーカ（切替）の結果が非常に近く，理想的な音像提示ができていることがわかる．しかし，ピンクノイズに関しては 2ch 超音波スピーカ（補間）の場合，一部の提示位置で左右端に音像が定位しやすくなる傾向がみられる．この原因としては，プロペラ音とピンクノイズは周波数特性が異なるため，スピーカと超音波スピーカの放射特性の違いが顕著に表れたのではな

いかと考えられる．

以上から，超音波スピーカにおける音像位置補間にはエネルギー和一定の式が有効であること，スピーカと超音波スピーカでは音像知覚に差異が存在し，超音波スピーカの場合，定位性能に劣化が生じる音源が存在することがわかった．

5. 視聴覚併用時の定位性能に関する主観評価実験（実験 2）

5.1 目的と条件

視聴覚を統合することによる刺激の定位（空間的腹話術効果）に関する研究はこれまでにも多く発表されており [19]，視聴覚刺激がほぼ同時に提示されると聴覚刺激は視覚の位置に向かって誤定位されること [20][21]，視聴覚刺激間の空間的，時間的分離が増すほど空間的腹話術効果が弱くなることが知られている [22-27]．

そこで，実験 2 では，聴覚併用 MR システム下で視覚と聴覚を併用した場合に，その相乗効果により，実験 1 で不十分であった音像の定位性能がどの程度向上するか，基礎実験を通して確認する．

実験 1 と同様，「2ch 超音波スピーカ（補間）」方式でスピーカの位置にランダムに音像を提示し，被験者には音像を知覚した位置をスピーカ番号で回答させる．音源の提示は，7 箇所の提示位置に対してそれぞれ 2 回，ランダムに提示する．音像位置の補間処理には実験 1 で性能の高かったエネルギー和一定の式 (3.1') を用いる．被験者は HMD を装着し，音像提示時には音源の CG 提示（図 5）を行う．このとき，被験者が提示された CG の位置のみを手がかりに回答するのを防ぐために，提示する音像位置と CG の位置が一致しないダミー提示も混在させる．このダミー提示はプロペラ音，ピンクノイズ，各音源に対して 7 回（全試行の 33.3%）がランダムに提示される．

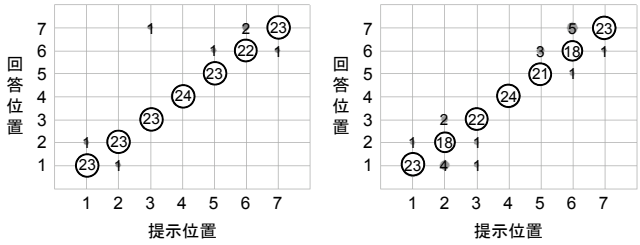
5.2 結果と考察

まず，ダミーを提示（音像位置と異なる位置に CG を提示）した際の，正答率（音像の位置を正しく回答した割合）と，CG の位置の方を回答した割合を**表 1** に示す．表より，プロペラ音とヘリコプタの CG を組み合わせた場合の方が，ピンクノイズと球の CG を組み合わせた場合より視覚刺激に影響を受けている傾向が伺われる．これは視覚刺激と聴覚刺激の意味的な一致性 [28][29] が，ヘリコプタの方がより高かったためと考えられるが，いずれも CG の位置を回答している割合は半数以下であり，被験者は単純に CG の位置を手がかりに回答しているのではなく，腹話術効果の影響下で音像位置を聞き分けて回答していることがわかる．

表 1　ダミー提示に対する回答の割合

Table 1 Percentage of answers in dummy presentation

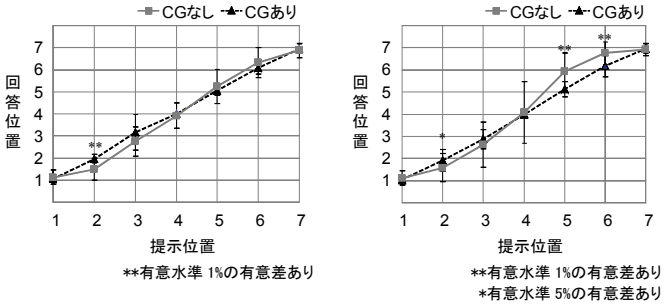
音源の種類	正解	不正解	
	音像の位置を回答	CG の位置を回答	その他
プロペラ音	25%	45%	30%
ピンクノイズ	33%	36%	31%



(a) プロペラ音　　(b) ピンクノイズ

図 13　音像位置と回答の分布（CG あり）

Fig. 13 Distribution of answers to the sound image position (presented CG)



(a) プロペラ音　　(b) ピンクノイズ

図 14　音像位置と回答の平均

Fig. 14 The average of the answers to the sound image position

次に，音源ごとの音像の提示位置と被験者の回答位置の分布を**図 13** に示す．図 13 は図 8 と同様に，被験者 12 名，各音像位置での提示回数 2 回，各提示位置につき計 24 回の回答を母数としたバブルグラフである．図より，プロペラ音，ピンクノイズともに音像の提示位置に回答が集中しており，回答の分散も小さく，定位性能が向上していることがわかる．実験 1 の CG なしの場合の結果と比較した音像提示位置に対する被験者の回答位置の平均を**図 14** に，被験者の正答率の平均と分散を**図 15** に示す．図からも，どちらの音源も定位性能が向上していることがわかる．実験 1 で定位性能の劣化がみられたピンクノイズに関しても，CG を提示したことで大きく定位性能が向上しており，CG なしと CG ありでは，有意差が得られた．

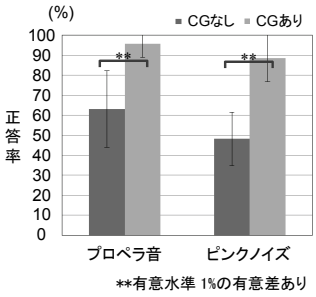


図 15　正答率の平均

Fig. 15 The average percentage of correct answers

予備実験の結果と比較しても，音源の種類にかかわらず，エネルギー和一定の式による音像位置補間と MR による視覚刺激を併用することで十分な定位性能が実現できていることがわかる．

Jackson [23] は，音源そのものから発せられる聴覚刺激が視覚刺激の位置に向かって誤定位されること，誤定位される聴覚刺激と視覚刺激の距離は 20～30 度程度であることを示している．本研究で提示しているのは，位置補間された（定位性能の不十分な）音像であるが，図 14 の CG なしのグラフから回答位置と提示位置の差は概ね 2～3 か所（20 度）程度であるので，それらの音像定位位置が CG の位置に定位したという今回の結果は，従来研究の知見とも符合する．

6. 移動音像の実現

4 章, 5 章では，音像の移動方向を水平方向についてのみ実験を行っているが，先行研究 [18] において，音像定位の劣る垂直方向でも，視覚提示があれば左右方向と同程度の定位性能が期待できることが分かっている．

更に，3 点間での音像補間が可能であれば，これを組み合わせることで，図 3 上図に示すように，より広範な移動音像を実現することができるはずである．そこで，これまでの実験結果を踏まえ，実際に 3 枚の反射板を配置し 3 点間での移動音像が可能であるか確認した．超音波スピーカ，反射板間，体験者の配置は，**図 16** の通り．各反射板間の距離は，**図 17** に示す通りで，底辺が床と水平，三角形の面が壁面と並行となるよう配置している．視聴覚併用 MR システムは図 6 と同様である．

提示する音像の動きは，3 点で構成される三角形内で，破線で示す直線，または円を描く（図 17）音像を提示した．CG は提示する音源と同じ位置に提示するが，被験者には同じとは限らないと伝えた．音像および CG の移動速度は, 0.25 [m/s], 0.5 [m/s], 1.0 [m/s] の 3 種類とした．

以上の条件下で，3 名の体験者に移動音像を体験

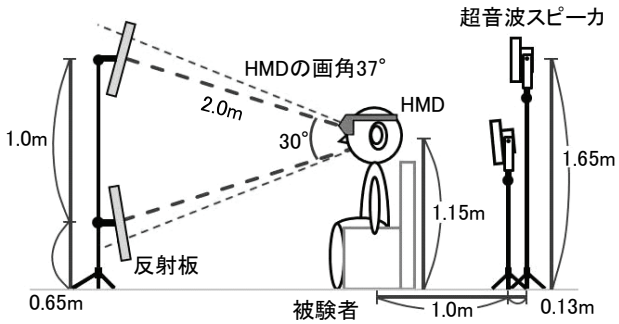
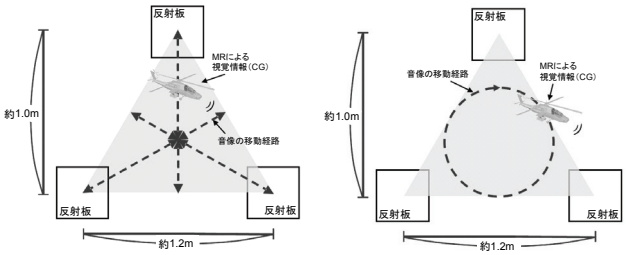


図 16　機材配置

Fig. 16 Equipment placement



(a) 直線的な動き　　(b) 内接円を旋回する動き

図 17　音像の動き

Fig. 17 Movement of sound image

させたところ，直線的な動き，円を旋回する動き（時計回り・反時計回り），いずれの場合でも，3 人全員が音像の移動を正しく知覚できることを確認した．

7. むすび

音像プラネタリウムを用いた視聴覚併用 MR において，スピーカ方向を制御するのではなく，複数の超音波スピーカを用いて，複数の反射面から提示される音像の位置を補間することで任意の位置を自由に移動する音像を実現した．超音波スピーカに適した音像補間式を確認するために，2 種類の補間式を比較する主観評価実験を行った．その結果，超音波スピーカでもエネルギー和一定の式が有用であることがわかった．MR による視覚情報を同時に提示することで，音像位置の定位性能が実世界に近いレベルまで向上することも確認した．

今後は，音像位置補間と MR による視覚刺激を併用することによって正しく音像定位可能な，反射板間の距離や部屋の形状による影響などの設計パラメータを明らかにしていく予定である．また，本論文では，体験者 1 名，体験位置固定という条件下での音像の移動を実現したが，今後は，体験者の移動や複数人体験にも対応できるよう，手法を拡張していく予定である．

本研究の一部は，科研費・基盤研究 S「複合現実型情報空間の表現力基盤強化と体系化」による．

参考文献

[1] 比嘉恭太，西浦敬信，木村朝子，柴田史久，田村秀行：“視覚・聴覚の現実と仮想を融合する 2×2 方式複合現実感システムの実現”，日本 VR 学会論文誌，Vol. 13, No. 2, pp. 227 - 237, 2008.

[2] 西野隆典，梶田将司，武田一哉，板倉文忠：“水平面上の頭部伝達関数の補間”，日本音響学会誌 Vol. 55, No. 2, pp. 91 - 99, 1999.

[3] 大谷真，伊勢史郎：“頭部伝達関数の数値解析とその応用”，同上，Vol. 61, No. 7, pp.404 - 409, 2005.

[4] 森勢将雅，湊佳彦，西浦敬信，木村朝子，柴田史久，田村秀行：“視覚・聴覚を併用した複合現実感システムのための頭部伝達関数の個人化”，電気学会論文誌 C，Vol. 130, No. 8, pp. 1466 - 1467 , 2010.

[5] 森勢将雅，杉林裕太郎，栗元総太，西浦敬信：“音像プラネタリウム：超音波スピーカを利用した 3 次元音場再生方式”，日本 VR 学会論文誌，Vol. 16, No. 4, pp. 687 - 693, 2011.

[6] 杉山孝之，木村朝子，柴田史久，田村秀行：“音像プラネタリウム方式を用いた視聴覚併用 MR システム X-Media Galaxy (1) ～基幹システムの設計と体験事例の実装～”，2012 信学総大，A-16-8, p. 257, 2012.

[7] 木村健太郎，宝珠山治，谷川智洋，廣瀬通孝：“超指向性スピーカによる大画面上音像提示システム”，映像情報メディア学会技術報告，Vol. 34, No. 25, pp. 1 - 6, 2010.

[8] 中垣拳，筧康明：“指向性スピーカを用いた空間拡張デバイス SonalShooter の基礎検討”，インタラクション 2011, pp. 371 - 374, 2011.

[9] K. Ishii, Y. Yamamoto, M. Imai, K. Nakadai: “A navigation system using ultrasonic directional speaker with rotating base,” The 12th Int. Conf. on Human-Computer Interaction, Vol. 4558, pp. 526 - 535, 2007.

[10] 石井健太郎，谷口祐司，大澤博隆，中臺一博，今井倫太：“身体性遠隔コミュニケーションにおけるユーザとアバタの視点の一致”，インタラクション 2012, pp. 49 - 56, 2012.

[11] S. Aoki, M. Toba and N. Tsujita: “Characteristics of stereo reproduction with parametric loudspeakers,” AIP Conf. Proc. 1433, 47 2012.

[12] K. Maeda, T. Nishino, and H. Naruse: “Horizontal and vertical sound image control using multiple parametric speakers,” Proc. the acoustics 2012, Paper Number: 1aHT11, 2012.

[13] L. Rayleigh: “Acoustical observations,” Phil. Mag. 3, 6th series, pp. 456 - 464, 1877.

[14] L. Rayleigh: “On our perception of sound direction,” Phil. Mag. 13, 6th series, pp. 214 - 232, 1907.

[15] M. Morimoto, K. Iida and M. Itoh: “Upper hemisphere sound localization using head-related transfer functions,” Acount. Sci & Tech. 24, pp. 267 - 275, 2003.

[16] 大久保洋幸，大谷眞道，小野一穂，正岡顕一郎，池沢龍，小宮山摂，浅山宏，湯山一郎：“CG 画像と同期したインタラクティブ音場再生システムについて”，日本 VR 学会論文誌，Vol. 5, No. 3, pp. 965 - 974, 2000.

[17] 鎌倉友男，酒井新一：“パラメトリックスピーカの原理と応用”, 信学会技術研究報告 EA, 105 (556), pp. 25 - 30, 2006.

[18] 伊藤仁一，森勢将雅，西浦敬信，木村朝子，柴田史久，田村秀行：“超音波スピーカによる移動音源の実現と視覚補助による定位性能の向上(2)―音像の移動範囲の拡張―”，第 16 回日本 VR 学会大会論文集, pp. 151 - 154, 2011.

[19] スペンスチャールズ，五十嵐由夏，北川智利：“視聴覚統合”，日本音響学会誌, vol. 63, No. 2, pp. 83 - 92, 2007.

[20] R. B. Welch and D. H. Warren: “Immediate perceptual response to intersensory discrepancy,” *Psychol. Bull.*, 88, pp. 638 - 667, 1980.

[21] P. Bertelson and B. de Gelder: “The psycology of multimodal perception,” i*n Crossmodal Space and Crossmodal Attenion,* C. Spence and J. Driver, Eds. (Oxford University Press, Oxford, 2004), pp. 141 - 177.

[22] M. Radeau and P. Bertelson: “Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations,” *Percept. Psychophys.*, 22, pp. 137 - 14, 1977.

[23] C. V. Jackson: “Visual factors in auditory localization,” *Q. J. Exp. Psyco*l., 5, pp. 52 - 65, 1953.

[24] R.I. Bermant and R.B. Welch: “Effect of degree of separation of visual-auditory stimulus and eye position upon spatial interaction of vision and audition,” *Percept. Mot. Skills*, 42, pp. 487 - 493, 1976.

[25] W.D. Hairston, M.T. Wallace, J.W. Vaughan, B.E. Stein, J.L. Norris and J.A. Schirillo: “Visual localization ability influences cross-modal bias,” *J. Cognit. Neurosci.*, 15, pp. 20 - 29, 2003.

[26] C.E. Jack and W.R. Thurlow: “Effects of degree of visual association and angle of displacement on the “ventriloquism” effect,” *Percept. Mot. Skills*, 37, pp. 967 - 979, 1973.

[27] D.A. Slutsky and G.H. Recanzone: “Temporal and spatial dependency of the ventriloquism effect,” *Neuroreport,* 12, pp. 7 - 10, 2001.

[28] P.J. Laurienti, R.A. Kraft, J.A. Maldjian, J.H. Burdette and M.T. Wallance: “Semantic congruence is a critical factor in multisensory behavioral performance,” *Exp. Brain Res.,* 158, pp. 405 - 414, 2004.

[29] S. Molholm, W. Ritter, D.C. Javitt and J.J. Foxe: “Multisensory visual-auditory object recognition in humans: A high-density electrical mapping study,” *Cerebral Cortex,* 14, pp. 452 - 465, 2004.

(2012 年 12 月 19 日受付)

[著者紹介]

伊藤 仁一（学生会員）



2011 年立命館大学情報理工学部メディア情報学科卒．2013 年同 大学院理工学研究科博士前期課程修了．同年 4 月，（株）ニコン入社．2010 年より 2013 年まで，超音波スピーカを用いた視聴覚併用複合現実感システムの研究に従事．

中山 雅人（非会員）



2001 年 近畿大・生物理工・電気システム情報工卒．2003 年 和歌山大大学院・システム工学研究科・博士前期課程修了. 2008 年 立命館大大学院・理工学研究科・博士後期課程満了．博士（工学）．立命館大・情報理工・特任助教および近畿大・生物理工・非常勤講師．音響信号処理，音声情報処理に関する研究に従事．電子情報通信学会，日本音響学会，各会員．

西浦 敬信（正会員）



1997 年 奈良高専・専攻科・電子情報卒．1999 年 奈良先端大・情報科学研究科・博士前期課程修了．2001 年 同大博士後期課程修了. 同年 和歌山大・シス工・助手. 2004 年 立命館大・情報理工・助教授. 2007 年 同准教授，現在に至る．博士（工学）．音響信号処理， 主として音環境の解析・理解・再現・生成に関する研究に従事．日本音響学会，情報処理学会，日本騒音制御工学会，電子情報通信学会，各会員．

木村 朝子（正会員）



1996年大阪大学基礎工学部卒．1998年同大学院基礎工学研究科博士前期課程修了．同大学助手，立命館大学理工学部助教授，科学技術振興機構さきがけ研究員等を経て，2009年4月より立命館大学情報理工学部メディア情報学科准教授．現在，同教授．博士（工学）．実世界指向インタフェース，複合現実感，ハプテックインタフェースの研究に従事．電子情報通信学会，情報処理学会，ヒューマンインタフェース学会，ACM，IEEE各会員．本学会学術奨励賞・論文賞，情報処理学会山下記念研究賞等受賞．

柴田 史久（正会員）



1996 年大阪大学大学院基礎工学研究科博士前期課程修了．1999 年同研究科博士後期課程修了．大阪大学産業科学研究所助手を経て，2003 年 4 月より立命館大学理工学部助教授．現在，同情報理工学部情報コミュニケーション学科教授．博士（工学）．モバイルコンピューティング，複合現実感等の研究に従事.本学会複合現実感研究委員会委員．IEEE，電子情報通信学会，日本ロボット学会，情報処理学会等の会員．本学会学術奨励賞・論文賞を受賞．

田村 秀行（正会員）



1970 年京都大学工学部電気工学科卒．工業技術院電子技術総合研究所，キヤノン（株）等を経て，2003 年 4 月より立命館大学理工学部教授，同情報理工学部メディア情報学科教授．現在，同大学総合科学技術研究機構教授．工学博士．1997 年より 2001 年まで，MR システム研究所にて「複合現実感研究プロジェクト」を率いた．本学会フェロー，元理事，現在，評議員，複合現実感研究委員会顧問．編著書「Mixed Reality」(Ohmsha & Springer)「コンピュータ画像処理」（オーム社）など．電子情報通信学会フェロー，IEEE，ACM，情報処理学会，人工知能学会，映像情報メディア学会等の会員．本学会及び情報処理学会論文賞，人工知能学会功労賞等を受賞．