

# 複数の超音波スピーカを用いた 音像位置の補間と視覚補助に基づく移動音像の実現

伊藤 仁一 中山 雅人 西浦 敬信 木村 朝子 柴田 史久 田村 秀行

立命館大学 〒525-8577 滋賀県草津市野路東 1-1-1

E-mail: itou@rm.is.ritsumei.ac.jp

あらまし 我々は視覚・聴覚の両方で現実世界と仮想世界を融合する複合現実感 (MR) システムの開発を行っている。本研究では、聴覚的 MR の提示方式として音像プラネタリウム方式を用いた際の移動音像の実現方法を提案した。提案手法は超音波スピーカの特性を利用したもので、複数の反射面から提示される音像位置を補間し、その音像位置に視覚情報を提示するという手法である。音像位置の補間処理は従来のダイナミックスピーカで用いられる手法に基づいて実装し、超音波スピーカを用いた場合でも有効であるのかを確認するため主観評価実験を行った。本稿では、その実験結果と考察を述べる。

キーワード 複合現実感, 視聴覚併用, 超音波スピーカ, 移動音像

## Implementation of Moving Sound Image Based on Interpolation of the Sound Image Position and Visual Assistance with Parametric Loudspeakers

Yoshikazu ITO Masato NAKAYAMA Takanobu NISHIURA Asako KIMURA  
Fumihisa SHIBATA and Hideyuki TAMURA

Ritsumeikan University 1-1-1 Nojihigashi, Kusatu, Shiga, 525-8577 Japan

E-mail: itou@rm.is.ritsumei.ac.jp

**Abstract** We develop an MR system that merged the real and virtual worlds in both the audio and visual sense. In this study, we suggested the method to realize the moving sound image when using "Acoustic Planetarium" as an audio presentation. This method is to interpolate the sound image position between multiple sound images on reflecting points which located by parametric loudspeakers, and showing the sight information at the sound image position. The sound image position interpolation processing was implemented based on conventional methods used for dynamic loudspeakers. And, we had subjective evaluation experiments to confirm whether it is effective even in the case of using the parametric loudspeakers. This paper presents the results and consideration of these experiments.

**Keyword** 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継ぎ目なく実時間で融合する複合現実感 (Mixed Reality; MR) の研究は、多くが視覚的な融合を目指している。これに対して、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視聴覚併用 MR システムの研究・開発を行ってきた。

視聴覚併用 MR システムを実現するためには、(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、これまでに我々は「2×2 方式視聴覚併用 MR システム

[1]」を実現した。同システムは、聴覚的 MR の達成に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ている。一般にバイノーラル再生方式は、受聴者ごとに頭部伝達関数 (HRTF) を用意する必要があるが、不特定多数の人が体験する MR アトラクションの運用を考慮すると、個々の体験者用の HRTF を用意することは難しい。そこで、多くの受聴者に対して一般化された HRTF を用意し、アトラクションの体験者全員が同じ HRTF を使用する方法を適用した。ただし、前述の通り HRTF には個人差があるため、用意された HRTF が体験者に適さない場合、正しく 3 次元音場を知覚できないという問題がある。

そこで我々は、新たに、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 [2]」を提案・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し、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、部屋の中央に複数の超音波スピーカを搭載した音像プラネタリウムの筐体を置くだけで、空間中の任意の場所に音像を定位させ、複数人が実音と人工音を同時に自然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。

しかし、音像プラネタリウム方式では音像を反射面に構築するため、自由に音像位置を移動させることができず、現段階では固定音の実現にとどまっている。そこで、本研究では音像プラネタリウムを用いた視聴覚併用 MR システムにおいて、ヘッドホンを利用する従来システムのように、任意の位置を自由に移動可能な音像の実現を目指す。

## 2. 関連研究

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。木村ら[3]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）を用いた場合には、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 [4] 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。

この手法であれば、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 のように音像を定位させた位置にその音源の仮想物体 (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。

そこで本研究では、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、

全天周での移動音像を実現する方法を提案・実装する。また、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について確認する実験を行う。本研究では、まず図 1 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。

## 3. 移動音像の実現方法

### 3.1. 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 1）。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。

### 3.2. 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と知られている。本研究では MR アトラクション作成時の自由度が高い音圧レベル差を用いた音像位置の補間処理を採用する。

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。

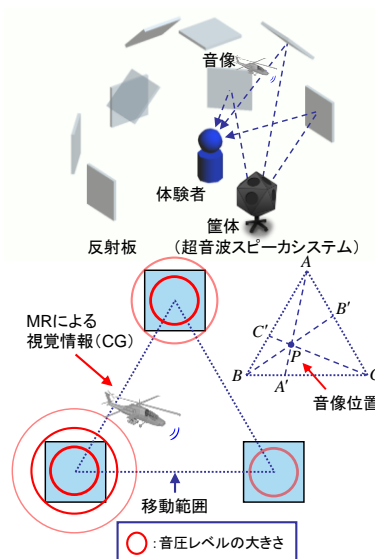


図 1 反射面の構築例と 3 点間補間

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。このときの補間方法として、まず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[5]。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声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[6]。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1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、

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

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、

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

となる（以降、振幅和一定の式と表記する）。

本研究では従来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。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。

## 4. 実験準備と予備実験

### 4.1. 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、本実験では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。

予備実験および本実験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ヶ所で、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。音像位置の補間処理に使用する補間式は、3.2で述べた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、例えば図1の  $B$  と  $C$  の2点間での音像補間の式は、それぞれ以下ようになる。

$$S_B = \sqrt{\frac{PC}{BC}} \quad S_C = \sqrt{\frac{PB}{CB}} \quad (3.1')$$

$$S_B = \frac{PC}{BC} \quad S_C = \frac{PB}{CB} \quad (3.2')$$

視覚提示の条件として、被験者は常にHMDを装着するが、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2)。

#### 【実験環境】

次節で述べる予備実験および本実験で用いる視聴覚併用MRシステムの構成を図3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場合を想定し、全ての実験で、被験者はビデオシースルー型HMD(Head Mounted Display)(Canon製VH-2002)を装着することとする。

実験時の機材配置を図4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、本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、ホワイトノイズを用いて被験者の頭の位置でそれぞれ約50[dB]となるように調節する。

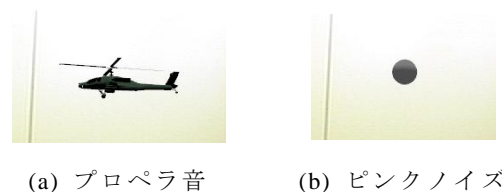


図2 音源に対する視覚情報

実験で使用したシステムの処理速度は約 60 [fps]である。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 の繰り返し
- (8) 全パターン終了後、自由コメント回答

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、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に指示する。

4.2. 予備実験

そもそも人（被験者）がどの程度音像位置を正確に判断できるのかを確認するため、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。

【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 5 に示す。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ると考えることができる。

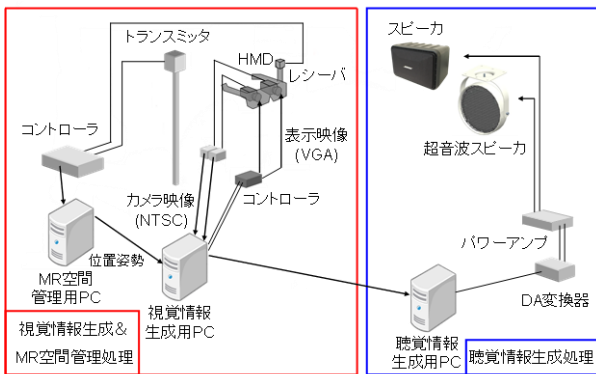


図 3 システム構成

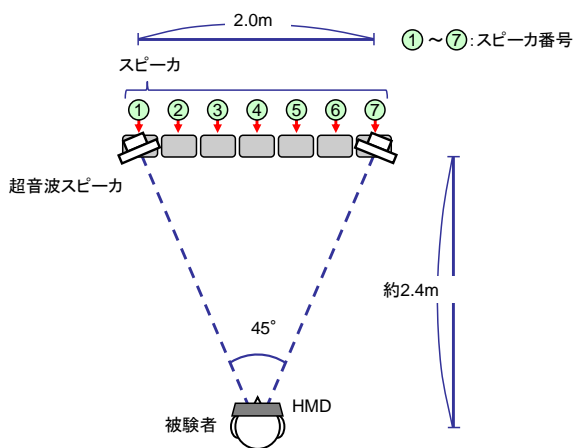


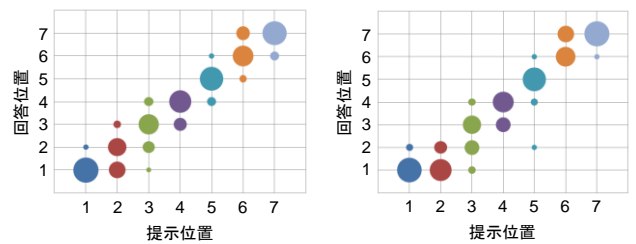
図 4 機材配置

5. 音像位置補間に関する主観評価実験(実験 1)

5.1. 目的と条件

音像位置補間を用いた音像提示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。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の日に行った。また、エネルギー和一定の式 (3.1') と振幅和一



(a) プロペラ音

(b) ピンクノイズ

図 5 音像位置と回答の分布（複数スピーカ（切替））

定の式 (3.2') の 2 種類の補間式による音像提示, 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 回ずつ提示される.

## 5.2. 結果と考察

まず補間式の違いによる傾向を確認するため, 補間式ごとの被験者の回答位置の平均を比較する (図 6, 7). 被験者の回答の平均をみると, 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 7 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. また, 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, 振幅和一定とエネルギー和一定では有意差 ( $p < 0.05$ ) が得られた (図 8). 以上より, 超音波スピーカにおける音像位置補間でも従来通りエネルギー和一定の式が有効であるといえる.

次に定位性能の良かった 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の結果の被験者平均を比較する (図 9). 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, プロペラ音が低域帯中心の音に対してピンクノイズは高周波を含むため, 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ことが考えられる.

以上から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では定位性能に劣化が生じる音源が存在することがわかる.

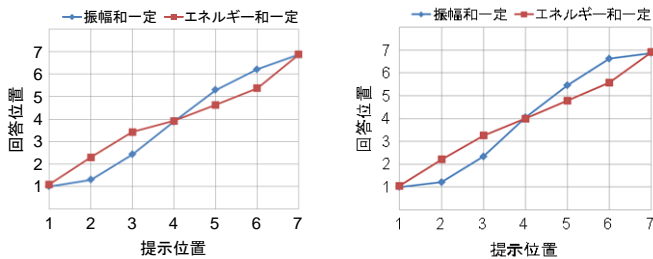
## 6. 視聴覚併用における定位性能に関する

### 主観評価実験 (実験 2)

#### 6.1. 目的と条件

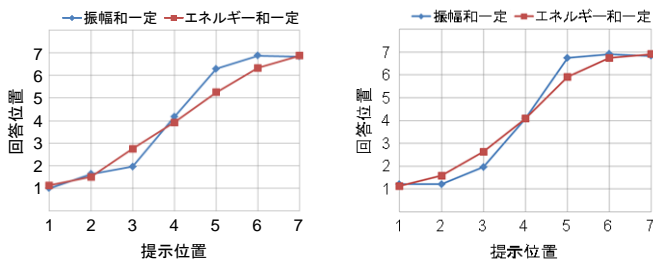
実験 2 では, 聴覚併用 MR システム下で視覚と聴覚を併用した場合に, その相乗効果により, 実験 1 で不十分であった音像の定位性能がどの程度向上するか, 基礎実験を通して確認する.

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 2) 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



(a) プロペラ音 (b) ピンクノイズ

図 6 音像位置と回答の平均 (2ch スピーカ (補間))



(c) プロペラ音 (d) ピンクノイズ

図 7 音像位置と回答の平均 (2ch 超音波スピーカ (補間))

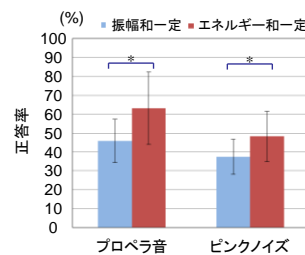
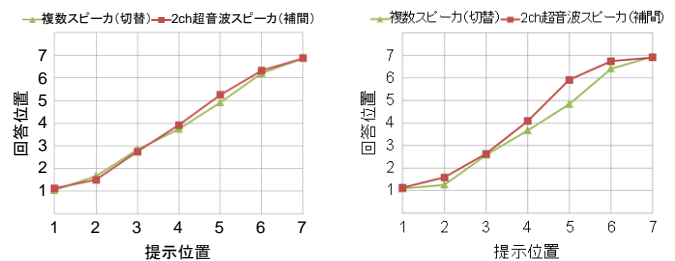


図 8 正答率の平均



(c) プロペラ音 (d) ピンクノイズ

図 9 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))

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。

## 6.2. 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、プロペラ音の方が視覚刺激により影響を受けている傾向がみられるものの、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 CG の位置を手がかりに回答しているのではなく、音像位置を聞き分けて回答していることがわかる。

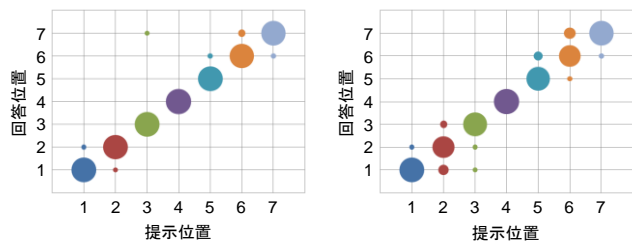
次に、音源ごとの音像の提示位置と被験者の回答位置の分布を図 10 に示す。図より、プロペラ音、ピンクノイズともに音像の提示位置に回答が集中しており、回答の分散も小さく、定位性能が向上していることがわかる。実験 1 の CG なしの場合の結果と比較した音像提示位置に対する被験者の回答位置の平均を図 11 に、被験者の正答率の平均と分散を図 12 に示す。これらの図からも、どちらの音源も定位性能が向上していることがわかる。実験 1 で定位性能の劣化がみられたピンクノイズに関しても、CG を提示したことで大きく定位性能が向上しており、CG なしと CG ありでは有意差 ( $p < 0.01$ ) が得られた。予備実験の結果と比較しても、音源の種類にかかわらず、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。

## 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

表 1 ダミー提示に対する回答の割合

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 10 音像位置と回答の分布 (CG あり)

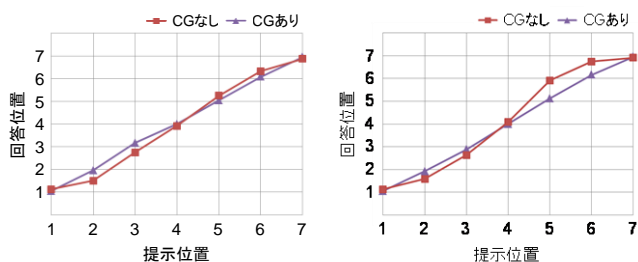
像の位置を補間することで任意の位置を自由に移動する音像を実現した。超音波スピーカに適した音像補間式を確認するために、2種類の補間式を比較する主観評価実験を行った。その結果、従来スピーカ同様、超音波スピーカでもエネルギー和一定の式が有用であることがわかった。MR による視覚情報を同時に提示することで、音像位置の定位性能が実世界に近いレベルまで向上することも確認した。

本研究では、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に、手法を拡張していく予定である。

謝辞 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。

## 文献

- [1] 比嘉恭太ほか：“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] 森勢将雅ほか：“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.
- [3] 木村健太郎ほか：“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.
- [4] S. Aoki *et al.*：“Characteristics of stereo reproduction with parametric loudspeakers,” AIP Conf. Proc. 1433, 47 2012.
- [5] 大久保洋幸ほか：“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.
- [6] 西野隆典ほか：“水平面上の頭部伝達関数の補間”，日本音響学会誌 55 (2), pp. 91 - 99, 1999.



(a) プロペラ音

(b) ピンクノイズ

図 11 音像位置と回答の平均

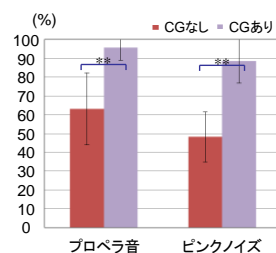


図 12 正答率の平均