

視覚・聴覚を併用した複合現実感システムの開発(2) — 体験者に合った頭部伝達関数の選択と接合法の検討 —

A Mixed Reality System that uses Audio and Visual Senses (2)
-Discussion about Selection and Piecing Method of Head-Related Transfer Function to Adapt Users-

湊佳彦, 西浦敬信, 木村朝子, 柴田史久, 田村秀行
Yoshihiko Minato, Takanobu Nishiura, Asako Kimura, Fumihisa Shibata and Hideyuki Tamura

立命館大学 情報理工学部
(〒525-8577 滋賀県草津市野路東 1-1-1)

Abstract: In order to reproduce 3-dimensional sound field with high accuracy, measurements of Head-Related Transfer Function (HRTF) suitable for each person are necessary. However, it is not easily applicable, since too much time and special equipments are needed to measure HRTF. Generally, we utilize one customarily used HRTF for every person or select the most befits HRTF for each person from large database of HRTFs. However, these methods could provide or select a mismatching HRTF, if there were no befits HRTF for the person. Therefore, this paper proposes a new method to produce a suitable HRTF by selecting plural HRTFs and piecing them together.

Key Words: Head-Related Transfer Function, Piecing Method, 3-Dimension Sound Field, Mixed Reality

1. はじめに

近年, 現実空間と仮想空間を違和感なく融合する複合現実感 (Mixed Reality; MR) 技術は, 様々な分野での応用が期待され, アートやエンターテインメント分野でも注目を集めている[1]. しかし, MR 研究の多くは視覚的な融合に留まっており, 聴覚における「複合現実」はほとんど着手されていない. そこで我々は, MR 空間で視覚と聴覚を同時に扱う視覚・聴覚融合型 MR システムの研究を行ってきた[2].

一方, 音響分野では, 頭部伝達関数(Head-Related Transfer Function; HRTF)による3次元音場の再現を目的とした研究[3]が数多くなされてきた. HRTF とは音源から聴取者の耳に届くまでの伝達関数のことを指し, HRTF の利用により水平面における方向定位は, ある程度正確に制御できることが知られている. バーチャルリアリティ分野でも, この技術を利用して, 視覚・聴覚両面で仮想空間を実現する研究[4]が行われている.

しかし, 3次元音場の正確な再現には個々人の HRTF を計測する必要があり, 多大な時間を要する. 例えば体験者参加型の MR アトラクションなど, 次々と新しい体験者が利用する状況では, HRTF の計測に多くの時間を割くことができない. 少数のまばらな点のみ HRTF の実測を行い, 実測していない部分の HRTF は計算によって補間するという研究[5]も行われているが, HRTF の計測には, 防音室や特殊な装置を要するため, 新しい体験者が次々と利用する

場面では, 即応は難しい. また, 計算機の性能の向上により, 体験者の頭部 CAD データから HRTF を数値解析により求める研究[6]も行われているが, 計測, 計算に時間がかかるという点では前述した方法と同様である. 逆に, HRTF を実測しない簡易的な方法として, 既に計測された他者の HRTF を利用するという方法があるが, 正確な3次元音場の再現が難しいという問題がある[7].

そこで本稿では, HRTF を実測することなく, しかも簡易に各体験者に適した HRTF を構築するために, HRTF をいくつかの領域に分け, あらかじめ用意した複数の HRTF の中からそれぞれの領域に適した HRTF を選択し, 接合する方法を提案する.

2. 従来手法

本章では, 従来から利用されている HRTF を選択する方法について, その利点・欠点を述べる.

(a) 標準的, 汎用的な HRTF を使用[7]

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 を使用する方法で, HRTF の選択に時間を取られないという利点を持つ. しかし, 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もあり, その場合は正しく3次元音場を知覚できないという欠点がある.

(b) 複数ある HRTF の中から自分に最も合うものを選択[7]

体験者が, 事前いくつかの HRTF を適応した音を実際

に受聴し、その中から体験者自身が自分に最も適すると判断する HRTF を使用する方法である。あらかじめ特徴的な HRTF を複数用意することで、体験者にとって使いやすいものを選択できる可能性が高いという利点があるが、用意された HRTF がどれも体験者に適さない場合もあり、その場合は正しく3次元音場を知覚できないという欠点がある。また、沢山の HRTF を用意すれば体験者に適した HRTF を選択できる確率も高くなるが、HRTF を選択するための作業負担が大きくなるという問題がある。

(c) 頭や耳などの大きさ、形状から計算機が最も合う HRTF を選択[8]

体験者の耳や頭部周りなど、頭部のいくつかの部位の長さを測り、それを基に計算機が体験者に適する HRTF を選択する方法である。利点として、先に紹介した2つの方法よりも高い定位感を得やすいことがあげられる。しかし、計測箇所が多く計測自体に時間がかかる、計測者の技能が結果に影響する、体験者によっては知らない人に計測されることを嫌う場合がある、などの欠点がある。

3. 提案手法

そこで本研究では、それぞれの体験者に適した HRTF を簡易的に構築する方法を提案する。

3.1 HRTF 領域分割モデル

HRTF とは、音源から聴取者の鼓膜までの頭部や耳介での回折を含めた音の伝達特性を表したものであり、2章(b)の方法は、顔や耳の形状が似ている人の HRTF を探して利用することと同じである。全ての部位が似ている人を探すのは非常に困難であるが、モンタージュ写真を作るように、それぞれの部位が似ている人を探して組み合わせることは、それほど難しくはない。そこで、本研究では、HRTF をいくつかの領域に分け、あらかじめ用意した複数の HRTF の中からそれぞれの領域に適した HRTF を選択し、接合する方法を提案する。

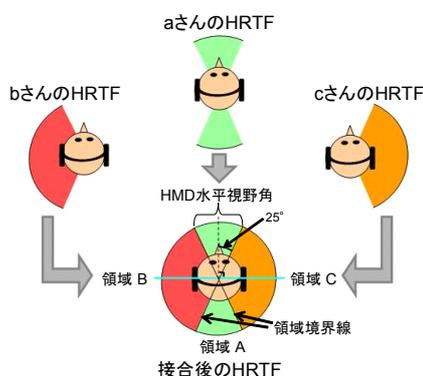


図1 HRTF 領域分割モデル

表1 Canon 製 VH-2002 の仕様

| | |
|-------|-----------------------|
| 視野角 | 水平方向 51 度 / 垂直方向 37 度 |
| 重量 | 325 g |
| 表示解像度 | VGA (640×480 画素) |
| 画像信号 | NTSC |

HRTF の領域分割であるが、領域を細かく分割しすぎると、接合箇所が増加し接合部での違和感につながると予想される。そこで、本研究では図1に示すように領域を前後、左、右の3つに分割する。3つの領域を分割する方位角は前後線対称とし、その方位角は、MR で使用する Head Mounted Display (HMD) の左右視野角に合わせる。今回は、Canon 製の HMD である VH-2002 (表1) を想定し、被験者の視線方向を中心として±25度で分割する。

3.2 HRTF 接合方法

図1のそれぞれの領域で選択された HRTF を、単純に切り替えて利用するだけでは、その接合部分(領域の境界部分)で違和感が生じる。そこで、境界部分では、2つの領域の HRTF をそれぞれ重み付けし、足し合わせることで、滑らかな接合を行う。2つの HRTF を足し合わせる範囲は、3つの領域が重なることがないように40度とする。本研究で利用する HRTF のデータベース[9]が5度の方位間隔で計測されているため、計9方位に対して接合処理を行うことになる。

以下、領域 A と B の接合を行う場合を例に取り、説明する。まず、A、B それぞれの領域で選択された HRTF を HRTF A、HRTF B とすると、図2のように一方の HRTF のパワースペクトル (HRTF A) をもう片方 (HRTF B) の値に近づけることで、それぞれの HRTF の周波数特性を保持しつつ接合処理を行うことができる。この際、HRTF A を HRTF B にどの程度近づけるかを、重み係数 γ_n で調節する。

次に、重み係数 γ_n を求める方法について説明する。領域境界線の方位で $n=0$ 、時計回りを正、逆を負で、 $-4 \leq n \leq 4$ (n は整数) とする (図3) と、 γ_n は、式(1)で求めることができる。但し、 α_n 、 β_n は重み係数、 $|A(\theta_n, \omega)|^2$ 、 $|B(\theta_n, \omega)|^2$ は、HRTF A、B のパワースペクトルを、 θ_n は求める HRTF の方位角を、 ω は周波数を表す。

$$\gamma_n = \alpha_n \times |A(\theta_n, \omega)|^2 + \beta_n \times |B(\theta_n, \omega)|^2 \quad (1)$$

(2)(3)に、 α_n 、 β_n を求める式を示す。 $n < 0$ の場合には HRTF B よりも HRTF A の波形に近づくように、 α_n の値が大き

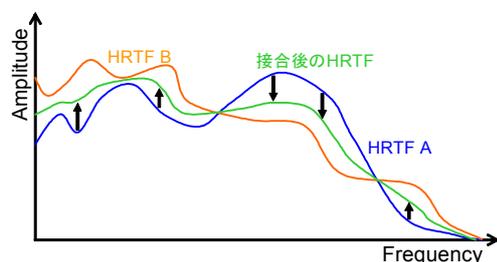


図2 2種類の HRTF の接合

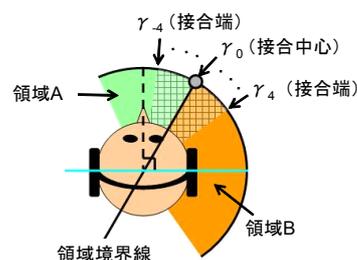


図3 HRTF 接合時の重み付け係数

なり、 β_n の値は小さくなる。 $n>0$ の場合は、逆に β_n の値が大きくなり、 α_n の値が小さくなる。

$$\alpha_n = \begin{cases} 0.85 + 0.0375 \times n (n = -4, -3, -2, -1, 0) \\ 0.0375 \times n (n = 1, 2, 3, 4) \end{cases} \quad (2)$$

$$\beta_n = 1 - \alpha_n \quad (3)$$

但し、式(2)(3)では、 $n=0$ のときに α_n 、 β_n がそれぞれ 0.5 とはならず、 $n=0$ と $n=1$ の間で、 HRTF A と B の接合比率 (α_n 、 β_n) の変化が他の場合よりも大きくなる。これは、実際に α_n 、 β_n を 0.5 に近づけて接合したところ、ミュージカルノイズが発生し、このノイズをできる限り抑えるためには、 α_n 、 β_n の比率がどちらかに偏っている必要があったからである。式(2)(3)では、経験的に α_n 、 β_n 値が 0.0~0.15 および 0.85~1.0 の範囲となるようにしている。また式(2)では、 α_n が上記の範囲の値となるよう、0.15 を 4 (n の最大・最小値の絶対値) で割った 0.0375 を n の係数としている。

最後に、図 3 の接合処理を行う。式(4)(5)に示すとおり、接合する各方位について、HRTF の実部 $A(\theta_n, \omega)_{Re}$ 、虚部 $A(\theta_n, \omega)_{Im}$ と γ_n の積を計算することで接合後の HRTF の実部 $A'(\theta_n, \omega)_{Re}$ 、虚部 $A'(\theta_n, \omega)_{Im}$ が求まる。実部、虚部に同じ値をかけているため、複素平面上での位相を保持したまま HRTF を接合できる。

$$A'(\theta_n, \omega)_{Re} = A(\theta_n, \omega)_{Re} \times \gamma_n \quad (4)$$

$$A'(\theta_n, \omega)_{Im} = A(\theta_n, \omega)_{Im} \times \gamma_n \quad (5)$$

4. 実験

4.1 実験目的

複数ある HRTF の中から被験者に最も合うものを選択する従来からの方法と、本稿で提案する HRTF を接合する方法を比較し、本手法によって構築した HRTF を利用することで、被験者がより高い定位感を得られるかどうかを検討する。

4.2 実験環境

- ・アトラクション等で利用することを想定し、防音室ではなく、研究室環境にて実験する（実験時の暗騒音は 50dBA）
- ・提示音は、440Hz 正弦波、2000Hz 正弦波、ヘリコプターのアイドリング音（アトラクション等で使う可能性のある音）の 3 種類
- ・被験者を中心として図 4 に示す 12 方位、距離 1520mm（実験で使用する HRTF が計測されたときの音源と被験者との距離）から音を提示する

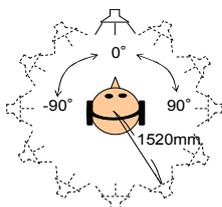


図 4 音源の提示位置

- ・MR で利用することを想定し、聴覚的 MR に適した開放型ヘッドホン（SONY MDR-F1）を使用する[2]
- ・実験中、被験者は HMD を装着せず、目は閉じる
- ・被験者は正常な聴力を持つ男子学生 3 名

4.3 実験手順

- 名古屋大学板倉（武田）研究室が提供している 96 種類の HRTF[9]の中から、特徴的な HRTF を 4 種類選択。
- (i) で選択した、4 つの HRTF を適応した合成音（音源に HRTF を畳み込んだ音）を、それぞれ図の 12 方位から提示し、被験者に音を知覚した方位を回答させる。音源方位の提示順序はランダムとし、各方位に対して、「合図音の提示⇒合成音を 3 秒間提示⇒3 秒間インターバル⇒合図音の提示⇒合成音を 3 秒間提示⇒被験者が知覚した方位を回答する」という手順で実験を行う。この実験を 3 種類の提示音に対して行うので、1 人の被験者が 3 種類×12 方位＝計 36 回の方位推定を行う。
- (ii) の実験結果から、4 つの HRTF の中で最も正答率の高い HRTF を選択し、この HRTF のときの(ii)の実験結果を、従来手法を利用した場合の方位推定結果とする。
- (ii) の実験結果から、図 1 の A、B、C の領域内で、最も正答率の高い HRTF をそれぞれ選択し、提案手法を用いて各 HRTF を接合する。
- (iv) で接合した HRTF を利用して、(ii)と同様の音源方位推定実験を行い、その結果を、提案手法を利用した場合の方位推定結果とする。
- また、(iii) で選択された HRTF と(iv)で接合された HRTF それぞれを利用した場合について、被験者の頭の周りをヘリコプターが旋回する音を提示する。被験者は連続音の定位感と音質の変化について回答する。

4.4 結果と考察

4.3節(iv)の段階で、図 1 の A、B、C 各領域に対してそれぞれの被験者に最も適すると判断された HRTF を表 2 に示す。HRTF 1、2、3、4 は4.3節(i)で選択した 4 種類の HRTF を指す。表より、今回の実験では、全領域で同じ HRTF が選ばれた被験者はいなかったことがわかる。このことから、多くの場合、聴取者に適している HRTF が領域によって変わることがわかる。

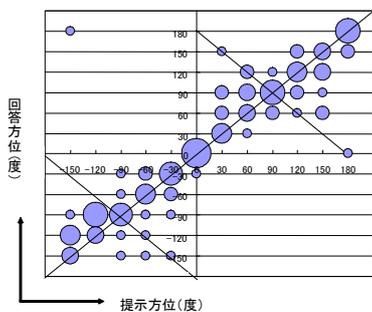
次に、各被験者の音源方位推定の正答率を表 3 に、全被験者の方位推定の回答（3 人×音源 3 種類×12 方位＝計 108 回答）結果を図 5 のバブルグラフに示す。今回の実験では、音源の違いによる方位推定結果の差はほとんど見られなかったため、これらの結果をまとめて示している。図 5(a)

表 2 各領域で選択された HRTF

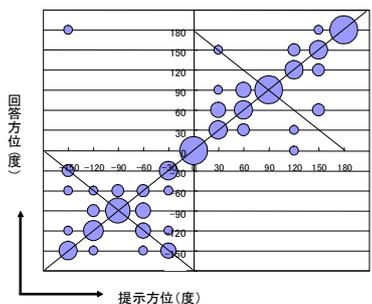
| | 領域 A | 領域 B | 領域 C |
|-------|--------|--------|--------|
| 被験者 1 | HRTF 4 | HRTF 2 | HRTF 2 |
| 被験者 2 | HRTF 1 | HRTF 4 | HRTF 4 |
| 被験者 3 | HRTF 3 | HRTF 3 | HRTF 4 |

表 3 方向推定の正答率

| | 被験者 1 | 被験者 2 | 被験者 3 | 平均 |
|------|-------|-------|-------|-------|
| 従来手法 | 41.6% | 50.0% | 58.3% | 50.0% |
| 提案手法 | 50.0% | 63.8% | 66.6% | 60.2% |



(a) 従来手法で選択された HRTF



(b) 提案手法で接合した HRTF

図 5 方位推定実験の結果

は4.3節(iii)の実験手順で選択された、4つの HRTF の中で最も正答率の高かった HRTF を用いた場合、(b)は提案手法を用いて接合した HRTF を用いた場合の結果であり、横軸は音源を提示した方位、縦軸は被験者が回答した方位である。右肩上がりの直線は回答方位と提示方位が一致した場合を示しており、この直線上に円がある場合は被験者が正しい回答をしたことを示している。また、左肩下がりの2本の直線は音像が折り返した場合（前後を誤判定した場合）を示している。円の大きさは、音源を提示した方位に対して、被験者が回答する方位角度の出現頻度を表している。音源をある方位から提示し、3人の被験者が、3種類の音源すべてで、それを同じ方位と推定した場合に円の大きさは最大となる。

まず表 3 より、全ての被験者について、提案手法の正答率の方が、約 10%程度高いことがわかる。図 5(a)(b)を比較しても、(b)の方が右肩上がりの直線上により大きな円が描かれていることがわかる。前後での誤判定の数（左肩下がりの直線上に描かれた円の大きささと数）は、図 5(a)と(b)で大きな違いはないが、(a)よりも(b)の方が、被験者の正面から左方向において音像の前後誤判定が多く発生している。この原因については、接合に利用する HRTF の選定が原因か、領域分割の方法に問題があるのか今後検討する必要がある。

一方、ヘリコプターの旋回音を提示した実験（4.3節(vi)）の結果、音の定位感については、従来手法、提案手法ともに連続して定位しているという回答を得た。しかし、音質に関しては、提案手法では HRTF の接合を行っている部分にノイズ（ミュージカルノイズなど）が混入しているという回答や、図 1の A, B, C の領域が切り替わる部分で音質に違いを感じるという回答があった。

以上の結果から、本稿で提案する手法を利用することで、被験者に合った、より正確な 3次元音場を提示できることがわかった。また、HRTF の接合部分での音質の違和感がなくなるよう改良する必要があることが確認された。

5. むすび

本稿では、各体験者に適した HRTF を構築する簡易な方法として、HRTF をいくつかの領域に分け、あらかじめ用意した複数の HRTF の中からそれぞれの領域に適した HRTF を選択し、接合する方法を提案した。従来方法との比較実験により、その有効性を確認したが、現状では音源が連続的に定位する場合に、接合部分で音質やノイズなどによる違和感が残ることがわかった。今後は、この接合部分での音質の違和感をなくす方法について検討するとともに、HRTF 領域分割モデルの妥当性についても再度検討する。

謝辞 本研究を進めるにあたってご協力いただいた立命館大学大学院の中山雅人氏、傳田遊亀氏に感謝の意を表す。本研究の一部は文部省科学研究費補助金・基盤研究(A)「三感融合型複合現実空間の構成法に関する研究」による。

参考文献

- [1] 横矢直和, 竹村治雄(編): “特集: 複合現実感 3”, 日本バーチャルリアリティ学会論文誌, Vol.10, No.3, 2005.
- [2] 比嘉恭太, 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開発(1) —視覚的 MR と聴覚的 MR の同時提示の実現—”, 本大会, 2006.
- [3] 渡邊祐子, 川名和博, 浜田春夫: “HRTF を用いた 3次元音場再生システムを利用した通信会議システムについて”, 日本音響学会講演論文集, pp.521-522, 2000.
- [4] 大久保洋幸, 大谷眞道, 小野一穂, 正岡頭一郎, 池沢龍, 小宮山撰, 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”, 日本バーチャルリアリティ学会論文誌, Vol.5, No.3, pp.911-918, 2000.
- [5] 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.
- [6] B.F.G.Katz, “Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation,” Journal of the Acoustical Society of America, vol.110, pp.2440-2448, 2001.
- [7] 大谷真, 伊勢史郎: “頭部伝達関数の数値解析とその応用”, 日本音響学会誌, Vol.61, No.7, pp.404-409, 2005.
- [8] Sensaura Virtual Ear™ technology:
<http://www.sensaura.com/technology/index.php?article=virtualEar.htm>
- [9] 名古屋大学板倉(武田)研究室:
<http://www.sp.m.is.nagoya-u.ac.jp/HRTF>