

# 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—

Mixed Reality System using Audio and Visual Senses (1)  
-Implementation of Simultaneous Presentation in Both Audio and Visual MR-

比嘉恭太, 西浦敬信, 木村朝子, 柴田史久, 田村秀行

Kyota HIGA, Takanobu NISHIURA, Asako KIMURA, Fumihisa SHIBATA, and Hideyuki TAMURA

立命館大学大学院 理工学研究科  
(〒525-8577 滋賀県草津市野路東1-1-1)

**Abstract:** In Virtual Reality, there have been many implementations using both audio and visual senses. However, Mixed Reality (MR), which merges real and virtual worlds in real-time, so far deals with the implementations only using visual sense. In this study, we have developed an MR system using both audio and visual senses, in which the geometric consistency of audio sense with visual sense was coordinated. We also have tried two approaches for merging real and virtual worlds in audio sense, one using an open-air headphone and the other using a closed-air headphone. The former corresponds to an optical see-through method and the latter to a video see-through method in case of visual sense.

**Key Words:** *Mixed Reality, Audio and Visual Senses, Geometric Consistency, Open-Air Headphone, Closed-Air Headphone.*

## 1. はじめに

現実世界と仮想世界を融合する複合現実感 (Mixed Reality; MR) [1]は急速な発展を遂げ、人工現実感研究の中でも最も活発な分野の1つとなっている。しかしながら、MR 研究開発の大半は視覚的な融合だけに留まっており、聴覚や触覚の「複合現実」はほとんど着手されていない。我々は、視覚・聴覚・触覚の三感融合型 MR 空間構成法を目指しているが、まずその第一歩として、視覚・聴覚が共存する MR システムの研究開発を進めている。

視覚と聴覚を併用した人工現実感 (Virtual Reality; VR) の実現例は既にいくつもある[2][3]。我々が目指すのは、視覚と聴覚の両方で現実と仮想が融合されていて、かつそれらが体験者の実在する現実空間と矛盾なく同時提示される MR システムである。そのためには、

- (1) 視覚的には現実世界の光景と CG 映像の重畳合成
- (2) 聴覚的には実音と人工音 (生成音) の混合
- (3) 上記の映像と音像の実時間幾何学的整合性

を達成する必要がある。(1)は既に達成されているので、本研究では(2)と(3)に挑戦し、第1期システムを完成した。

(1)の結果は体験者がシースルーHMD (Head Mounted Display) を装着して MR 空間を移動体験することから、(2)の結果は据置のスピーカでなく、ヘッドホンを装着して音像提示するものとした。この場合でも、現実 (実音) と仮想 (人工音) は、開放型ヘッドホンを利用する場合と密閉型ヘッドホンを利用する場合の2つの場合が考えられる、これは視覚的融合における「光学シースルー方式」と「ビ

デオシースルー方式」に相等する。本研究では、聴覚における両方式を実際に試して、その結果を考察した。

以下本稿では、2章で視覚的 MR、聴覚的 MR それぞれの提示方法と両者を同時提示できるシステム構成を、3章で視聴覚での幾何学的整合性の達成方法を、4章でヘッドホンの特性の違いによる音響的整合性の実験方法とその結果に関して述べる。

## 2. 視覚と聴覚の複合現実感

### 2.1 視覚の複合現実感

視覚的な複合現実感の提示方法としては、図1におけるいずれかの方式の HMD を用いるのが一般的である。前者はハーフミラー機能を有した光学的映像合成で、いわば網膜上で2つの映像が合成される。後者は、両眼を代行する一対のビデオカメラを遮蔽型 HMD に付加した構成で、カメラでキャプチャした現実世界の映像にコンピュータで生成した仮想物体を重畳描画し、眼前のディスプレイに表示する。それぞれに長短所があるが、光学的整合性 (画質合わせ) には後者が勝っているため、本研究では「ビデオシースルー方式の HMD」を用いる。

### 2.2 聴覚の複合現実感

コンピュータ内で生成した3次元音場の提示には、ヘッドホンを利用するバイノーラル方式と複数のスピーカを利用するトランスオーラル方式がある。我々は同時に複数人が MR 空間体験することを想定しているため、トランスオーラル方式ではその実現が困難である。よって本研究で

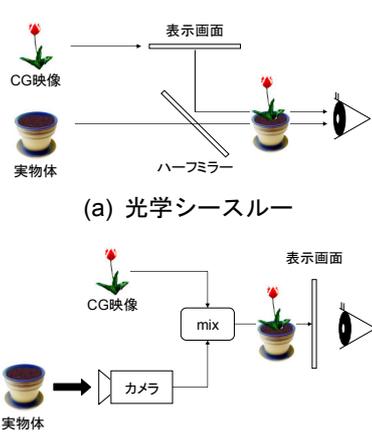


図 1 2つのシースルー方式

は、ヘッドホンを用いて各人に3次元音場を提示するバイノーラル方式を採用する。

現実と仮想の聴覚的な融合にも、視覚におけるHMDと同様、図2に示すような2方式が考えられる。(a)は開放型ヘッドホンを用いる方式で、人工音はヘッドホンから提示されるが、ヘッドホンと耳の間の隙間から実音を聞くことができる。もう一方は、密閉型ヘッドホンの外側にマイクロホンを装着する方法で、マイクロホンで受音した実音とコンピュータ内で生成した人工音の混合結果がヘッドホンから提示される。

本研究では、この両方式をそれぞれ「開放型音独立提示方式」「密閉型混合音提示方式」と呼ぶことにした。

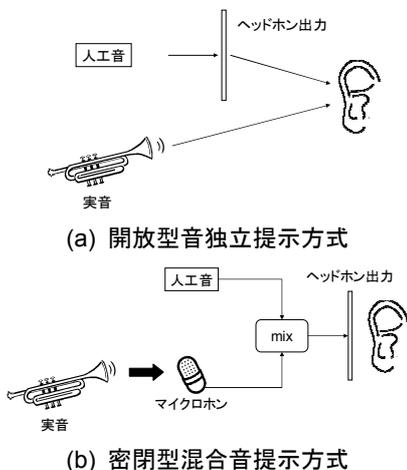


図 2 2つの音提示方式

### 2.3 視聴覚共存MRシステムの構成

現実世界をベースにCG映像空間と3次元音場がそれぞれ幾何学的整合性を保って存在し、体験者の移動・対話的操作に対して応答できれば、視聴覚共存のMR空間が構築できる。本研究では図3に示すようなシステム構成でこれを実現した。

本システムの構成は、3次元映像生成&MR空間管理処理と3次元音場生成処理の2つに大別できる。3次元映像生成&MR空間管理処理では、HMD(Canon VH-2002)に内蔵したカメラでキャプチャした現実世界の映像にコンピュータ内で生成したCG映像を重畳描画しHMDに表示する。このとき体験者の頭部の位置・姿勢はHMDに取り付けた磁気センサ(Polhemus 3SPACE FASTRAK)で検出す

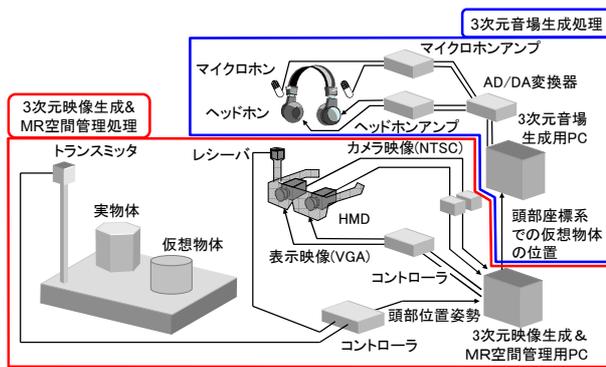


図 3 システム構成

表 1 通信データ

送信データ	説明
オブジェクト ID	仮想物体の ID 番号
サウンド ID	発音する音の ID 番号
相対距離[m]	体験者と仮想物体間の相対距離
相対方位[deg]	体験者と仮想物体間の相対方位

る。3次元音場生成処理では、3次元映像生成&MR空間管理処理用PC(Canon MR Platform System)から送られてくる仮想物体の情報(表1)をもとに3次元音場生成用PC(Dell Precision 670)が3次元音場を生成する。「開放型音独立提示方式」では、DA変換器(Thinknet DF-2032D)を通してヘッドホンより人工音を提示する。「密閉型混合音提示方式」では実音をマイクロホンで受音し、AD変換器(Thinknet DF-2X16-2)を通して3次元音場生成用PCに取り込み、人工音と混合したものを提示する。

### 3. 幾何学的整合性の達成

#### 3.1 視覚における現実と仮想の位置合わせ

現実世界と仮想世界の間で、視覚的な幾何学的整合性をとるためには、図4の実世界座標系(W)、仮想物体のモデル座標系(M)、HMD内蔵カメラ座標系(C)の3つの座標系の幾何学的整合性をとる必要がある。実世界座標系(W)、センサ座標系(S)、モデル座標系(M)の3者の相対位置、頭部座標系(H)、カメラ座標系(C)の相対位置は、いずれも一定である。これらの座標系の位置合わせはセンサとカメラ間の事前キャリブレーションによって行う。またHMDに取り付けた磁気センサの位置・姿勢はセンサ座標系(S)上で常に計測されているので、磁気センサを原点とする頭部座標系(H)とセンサ座標系(S)の対応付けが可能である。以上により、実世界座標系(W)、モデル座標系(M)、カメラ座標系(C)の3つの座標系の幾何学的対応付けが可能となる。

#### 3.2 聴覚における現実と仮想の位置合わせ

現実世界と仮想世界の間で、聴覚的な幾何学的整合性をとるためには、図4の実世界座標系(W)、モデル座標系(M)、

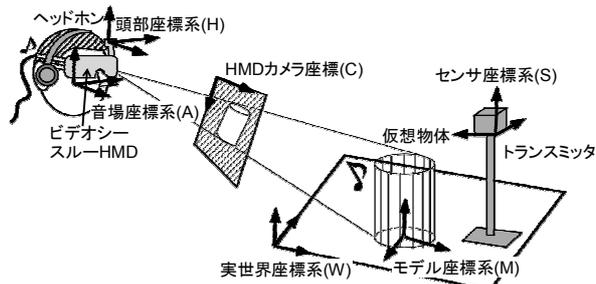


図 4 幾何学的位置合わせを行う座標系

音場座標系(A)の3つの座標系の幾何学的整合性をとる必要がある。音場座標系(A)は体験者の頭部中心を原点としており、体験者の頭の大きさがわかれば、HMD上のセンサを原点とした頭部座標系(H)との対応付けが可能である。頭部座標系(H)、実世界座標系(W)、モデル座標系(M)の対応付けは3.1節で述べたように既知であるため、実世界座標系(W)、仮想物体のモデル座標系(M)、音場座標系(A)の3つの座標系の対応付けが可能である。

### 3.3 3次元音場の生成

体験者が提示音を聴取する様子を図5に模式的に示す。図5において、仮想音源を $s(t)$ 、仮想音源から体験者までの相対距離を $d$ 、仮想音源と体験者の相対方位を $\theta$ 、左右の耳に対する頭部伝達関数(Head-Related Transfer Function; HRTF)をそれぞれ、 $h_l(\theta, t)$ 、 $h_r(\theta, t)$ 、左右の耳で提示される提示音をそれぞれ、 $o_l(t)$ 、 $o_r(t)$ とすると、

$$o_l(t) = (s(t) * h_l(\theta, t)) / d \quad (1)$$

$$o_r(t) = (s(t) * h_r(\theta, t)) / d \quad (2)$$

(\*は畳み込み演算)

となる。本来、距離減衰を厳密に再現するためには、温度や湿度、風速などを考慮して距離減衰量を算出しなければならない。しかし、音の定位感を重視するのであれば、相対距離だけでも十分に距離減衰を表現できるといえ、本システムでは、畳み込み演算で得られた結果を相対距離で割り、簡易的に距離減衰を表現している。本研究では、3.1節~3.3節を実装し、MR空間で視覚・聴覚が幾何学的違和感なく融合することを確認した。

## 4. 実音と人工音の音響的整合性実験

### 4.1 実験目的

「開放型音独立提示方式」「密閉型混合音提示方式」のいずれが視覚・聴覚融合型のMR空間に適しているのか、音響的整合性の観点から2つの実験で評価する。本実験では、環境音として用意した音を実物のスピーカで再生し、これを現実空間から伝わってくる実音とみなす。ここでは、実音源、仮想音源ともに床面に配置する。また実験中は被験者の位置を固定し、さらに視覚的な影響を抑えるために目を閉じてもらう。本実験では無指向性のマイクロホンを使用し実音を受音しているため、「密閉型混合音提示方式」では受音した実音にHRTFの影響を考慮する場合と考慮しない場合の2種類で実験を行う。密閉型ヘッドホンの外観を図6に示す。また、ヘッドホン特性が結果に影響しないようにしており、実験は暗騒音48dBAの環境で行った。

### 4.2 実音のみの提示実験

#### 4.2.1 実験内容

密閉型ヘッドホンは実音を一度マイクで受音するため、音質の劣化やシステム遅延が発生する。一方、開放型ヘッドホンは耳を覆うためHRTFが変化する。そこでヘッドホンを装着しない場合と装着した場合で実音の聴取に及ぼ

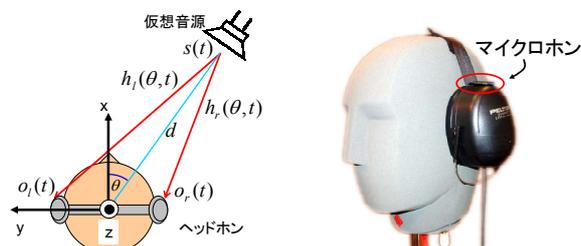


図5 提示音の聴取 図6 密閉型ヘッドホンの外観

す影響を開放型ヘッドホン (SONY MDR-F1)、密閉型ヘッドホン (PELTOR HTM79B-S) の両方で検討する (図7)。ここでは音質に対する違和感と遅延による違和感の2つを評価する。前者は、ヘッドホンを装着して実音を聴取する場合と装着しないで聴取する場合を比べて、どの程度音質に対する違和感を覚えるか3段階 (1:違和感がある, 2:違和感が少しある, 3:違和感はない) で評価する。提示する実音には、楽曲を使用する。一方後者は、ヘッドホンを装着して実音を聴取する場合と装着しないで聴取する場合を比べて、どの程度遅延を感じるか3段階 (1:遅延を感じる, 2:遅延を少し感じる, 3:遅延を感じない) で評価する。提示する実音には、被験者自身の拍手と声の2つを使用する。被験者は正常な聴力を持つ学生5名である。



図7 実音のみの聴取 (ヘッドホン装着)

#### 4.2.2 実験手順

被験者ごとに以下の手順で行う。

- (1) ヘッドホン未装着で実音を聴取する
- (2) 開放型ヘッドホンを装着して実音を聴取する
- (3) ヘッドホン未装着で実音を聴取する
- (4) 密閉型ヘッドホンを装着してHRTFを考慮していない実音を聴取する
- (5) ヘッドホン未装着で実音を聴取する
- (6) 密閉型ヘッドホンを装着してHRTFを考慮した実音を聴取する

#### 4.2.3 実験結果と考察

音質に対する違和感の実験結果を図8に、遅延に対する違和感の実験結果を図9に示す。音質に関する評価では、開放型ヘッドホンはほとんどの被験者が違和感を覚えないのに対し、密閉型ヘッドホンはともにほとんどの被験者が少なからず違和感を覚えている。遅延に関する評価では、開放型ヘッドホンは、遅延を感じないという結果であったが、密閉型ヘッドホンはともに遅延を感じるという結果であった。また、提示音が「被験者の拍手」「被験者自身の声」のどちらでも同じ結果であった。

図8、図9より、実音のみの聴取は音質や遅延に対する違和感において開放型ヘッドホンが明らかに優れている。

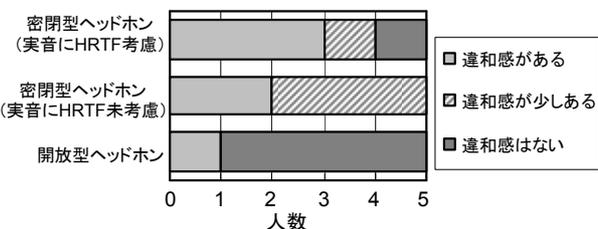


図8 音質に対する評価

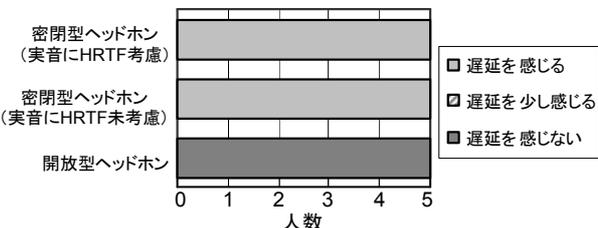


図9 遅延に対する評価

ということがわかる。密閉型ヘッドホンは、受音する実音に HRTF を考慮する場合でも、そうでない場合でも結果は変わらなかった。これは、本実験で使用した HRTF[4] と被験者自身の HRTF との差が原因として考えられる。

また、数人の被験者より「自分自身が発話すると不快感がある」という意見があり、これは被験者自身の声が体を伝わり直接耳に届くものと、マイクロホンで受音したものが両方聴こえることが原因であると考えられる。

#### 4.3 実音と人工音の提示実験

##### 4.3.1 実験内容

「開放型音独立提示方式」と「密閉型混合音提示方式」で、実音と人工音を同時に聴取したときの両者の音質や音の定位感に違いがあるか検討する(図 10)。ここでは、音質に対する違和感と定位感の 2 つを評価する。前者は、各方式について、実音と人工音の音質に対する違和感が少ないと感じた順を相対的な順位で評価する。提示音として、実音と人工音にそれぞれ異なる楽曲を使用した。後者も同様に、人工音の定位感が高いと感じた順を相対的な順位で評価する。提示音として、実音に楽曲、人工音にヘリの音を使用した。被験者は 4.2 の実験と同じである。

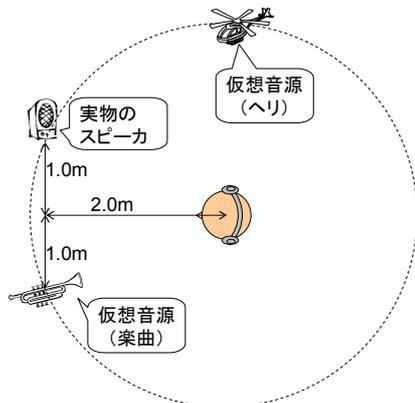


図 10 実音と人工音の聴取 (ヘッドホン装着)

##### 4.3.2 実験手順

被験者ごとに以下の手順で行う。

- (1) 開放型ヘッドホンで、実音と人工音を聴取する
- (2) 密閉型ヘッドホンで、HRTF を考慮していない実音と人工音を聴取する
- (3) 密閉型ヘッドホンで、HRTF を考慮した実音と人工音を聴取する

本実験では、実験順序が結果に影響しないように、被験者に応じて実験順序を入れ替えて行った。

##### 4.3.3 実験結果と考察

音質に対する違和感の実験結果を図 11 に、定位感の実験結果を図 12 に示す。音質に対する違和感と定位感の実験結果は、両方とも開放型ヘッドホンが最も違和感が少ないと回答した被験者が 3 名、実音に HRTF を考慮した密閉型ヘッドホンが最も違和感が少ないと回答した被験者が 2 名となり、双方に差はなかった。

図 11、図 12 より、それぞれのヘッドホンの優劣に差はなく、また音質、定位感ともに回答の傾向が同じであることがわかる。さらに、ほとんどの被験者がどのヘッドホンでも「音質に対する違和感を覚える」「定位感はある」と述べており、実音と人工音を同時に聴取する際には、ヘッドホンの特性の違いによる差はあまりないと考えられる。4.2

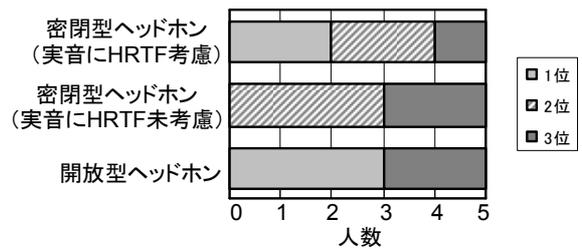


図 11 音質に対する相対順位

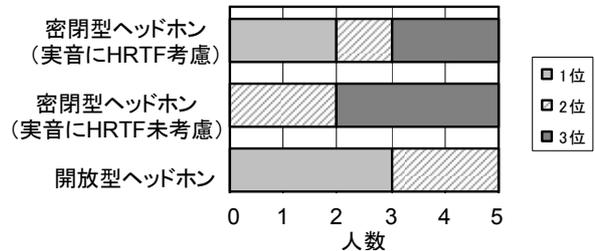


図 12 定位感に対する相対順位

節の実験では、実音のみの聴取はあきらかに開放型ヘッドホンの評価が高く、そのことを踏まえると実音を使用したり体験者自身が音を発する MR 空間では、開放型ヘッドホンを使用する「開放型音独立提示方式」の方が適していると考えられる。

## 5. むすび

本稿では、まず視覚と聴覚の複合現実感について述べた。次に視覚と聴覚を併用した複合現実感システムを実現し、そのシステムにおける幾何学的整合性の達成方法と音響的整合性の検討について述べた。幾何学的整合性の達成では、視覚・聴覚それぞれの現実と仮想の位置合わせを行い、視覚・聴覚に共通する座標系を用いて両者の幾何学的対応付けを行った。音響的整合性実験では、聴覚的 MR の実現方式として「開放型音独立提示方式」「密閉型混合音提示方式」の 2 方式を考えてその優劣を比較実験した。当初の予想では、視覚のビデオシースルー方式と併用するには「密閉型混合音提示方式」の方が適しているかと思われたが、結果は開放型ヘッドホンを用いる前者の方が優れていることが判明した。

今後は、CG 映像空間と 3 次元音場の幾何学的整合性を意図的になくした場合と音のスケール感を意図的に変えた場合に生まれる効果等について検討する。

**謝辞** 本学大学院生中山雅人氏、傳田遊亀氏には、本実験で多くの助言を頂いた。両氏に感謝の意を表します。本研究の一部は、科研費・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] 田村秀行, 大田友一: “複合現実感”, 映像情報メディア論文誌, Vol. 52, No. 3, pp.206-207, 1997.
- [2] 田村祐一, 他: “数値シミュレーションデータ表現のための音情報機能を付加したバーチャルリアリティシステムの開発”, 日本バーチャルリアリティ学会論文誌, Vol. 5, No. 3, pp. 943-948, 2000.
- [3] 大久保洋幸, 他: “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”, 同上, Vol. 5, No. 3, pp. 965-974, 2000.
- [4] 名古屋大学板倉 (武田) 研究室: <http://www.sp.m.is.nagoya-u.ac.jp/HRTF>