

超指向性スピーカを用いた実空間と仮想空間の 共有感向上手法の検討

Examination of a Method for Improving Shared Senses Between Real and Virtual Spaces Using Parametric Speaker

下田 桂輔†
Keisuke Shimoda

福島 拓†
Taku Fukushima

1. はじめに

近年、AR 市場は拡大し続けており、AR サービスもそれに伴い普及していくと予測されている[1]。利用シーンはスマートフォンゲームやバーチャルライブなどのエンターテインメント、商品のプロモーション、教育・訓練など多用途にわたり、日常的に AR 技術に触れる頻度が高まってきた。特に最近では、カメラで取得した画像の特徴量から空間認識を行うマーカースレスの AR システムの開発が進められており、AR 機能を搭載するスマートフォンにより手軽に AR 体験ができる環境が整いつつある。

ユーザが自身のスマートフォンを用いて AR 体験をする場合、仮想空間を描画する画面領域が小さく視野角が狭いといった特徴がある。そのため、仮想オブジェクトの出現位置が把握しづらく、出現場所によっては発見できなかったり、見失ったりする問題が発生する。

そこで、仮想オブジェクトの出現位置に AR デバイスを向けるよう促すために、仮想オブジェクトの音はその出現位置から発生しているかのように知覚するさまざまな音源定位手法が存在している。ユーザの位置や向きから HRTF (Head Related Transfer Function) を用いて音声を動的生成する手法では、多くの場合イヤホンやヘッドホンを使用する必要がある。イヤホン等をつけたユーザは現実空間上の周囲の音が聞こえにくいため、他ユーザとコミュニケーションをとることや、周囲の状況に気づくことが難しい。特に周囲の音が聞こえない状況でデバイスの画面を注視しているため危険が伴う。さらに、仮想空間の音だけしか聞こえないことにより、実空間と仮想空間を重ねて見ている光景と聞いている音が一致せず、空間共有感を得られないと考えられる。

そこで我々は、イヤホン等で耳を塞がずに音源定位を行う手法として、超指向性スピーカとムービングヘッドを用いた音声提示手法を採用して研究を行っている[2]。これは任意の場所に超指向性スピーカを向け、床面もしくは天井面からの反射音を発生させることで、あたかもその場所から音が聞こえるシステムである。スマートフォン AR アプリと連動させることにより、仮想オブジェクトの出現位置に音像を提示することができる。このシステムでは、動き回る仮想オブジェクトや連続的な音を発する仮想オブジェクトにおいて、実在感や定位感が得られる可能性があることが先行研究より示唆されている[2]。しかし、利用シーンが限定的であり、すべての AR コンテンツにこのシステムを適用するのは難しい。特に静止している仮想オブジェクトや足音等の断続的な音を発する仮想オブジェクト、歌声など高音質が求められる仮想オブジェクトの実在感には本システムには適していないことが一部指摘されている。

本稿ではより多くの仮想オブジェクトや音声の種類への対応や実在感の向上を図るため、超指向性スピーカの反射特性と実空間の環境音に着目した。今回は実験により 1) 音源定位精度の向上を目的とした音の反射面素材の検討、2) 実空間の環境音を仮想空間上の音と同時に視聴することによる空間共有感の優位性の評価、をそれぞれ行う。

2. 関連研究

2.1 音源定位に関する課題

前述した HRTF を用いる音源定位手法では、イヤホン等で耳を塞ぐ必要があるため、実空間における周囲の音が聞こえにくくなる。そこで、周囲の音も聞こえる音源定位手法が求められる。耳元にマイクを設置し、実空間の音を集音した音声と仮想空間上から出力される音をミキシングすることで、実空間の環境音を聞くことも可能だが、構成デバイスが煩雑になる(2.3 節で後述)。イベント等で活用されている音の AR を用いた手法では、GPS やセンサによる位置情報を利用して、パーソナライズした音を手元デバイスや設置されたスピーカからユーザに提示できるが、ユーザの動きに連動した音源定位まではあまり行われていない。複数スピーカによりサラウンド環境を構成する手法では、ユーザがスピーカに囲まれているため、音源が複数聞こえるといった問題も指摘されている[3]。同期した複数のスマートフォンから音声を提示し音源定位を動的に行う手法[4]についても同様に音の定位感に問題があることがわかっている。

2.2 超指向性スピーカを用いて音源定位を行う研究

超指向性スピーカ[5]とは、超音波を使って音に直進性をもたせることができるスピーカであり、パラメトリックスピーカとも呼称されている。可聴音を直進性の高い超音波に変調して空気中に放射すると、空気中で復調されスピーカを向けた方向に可聴音を届けることができる。伊藤らの音像プラネタリウム方式を用いた際の移動音源の実現についての研究[6]では、音を提示する空間の周囲に複数の反射面を用意し、複数の超指向性スピーカで反射面へ音声を投射することで仮想音源を生成している。これにより、ユーザの周りの音源定位を実現しているが、音像を提示できる範囲に限界があり、AR 体験中などユーザが自由に動く用途や複数人での用途には適さない。また、北岸ら[7]は、講師・受講生間などの一人対多人数コミュニケーションを目的として、影と移動軌跡の描画による視覚定位と、移動する足音音源による音源定位で表現される講師の分身エージェントを提案している。プロジェクトで影と移動軌跡の描画、超指向性スピーカとサーボモータで足音音源の提示を

行っているため、分身エージェントと音源の動的な提示が可能となっている。この研究では、講義での利用を想定しているために影と移動軌跡、足音のみで分身エージェントの位置を表現している。音源定位手法は我々の研究と同様の仕組みだが、足音だけでなく様々な音源への対応、音源提示精度の向上、実空間における環境音の影響の調査を行い、実在感の向上が可能な利用シーンの拡張を検討する。

2.3 骨伝導を用いて音源定位を行う研究

実空間の周囲の音が聞こえる音源定位手法として、骨伝導イヤホンで仮想空間の音を提示し、耳内部の蝸牛で実空間の環境音とミキシングする方法がある。Robertら[8]は、サラウンドスピーカ、骨伝導ヘッドホン、集音マイク付きヘッドホンの3つの環境音付き音声提示手法を比較し、骨伝導ヘッドホンの音源定位感の検証を行っている。集音マイク付きヘッドホンは、ミキシング用のデバイスが必要となるほか、場合によってはノイズキャンセリングが必要となるため、手軽にAR体験を行う環境としては適さないと考えられる。高井ら[9]は、誘導案内音声とユーザの周囲環境の環境音を同時聴取する手法としてHRTFが畳み込まれた骨伝導による誘導案内システムの検討を行っている。これらの骨伝導を使用したシステムは、実験により音源定位感を感じられることが示されており、本稿では骨伝導イヤホンとHRTFを用いた音源定位手法を比較対象として採用する。

3. 提案システム

3.1 概要

本研究では仮想オブジェクトの実在感向上を目的として、スマートフォンARアプリ(3.2節)と音源発生装置(3.3節)で構成されるAR体験システムを開発している。これは、ARライブや対戦ゲームなど体験共有型リアルタイムコンテンツに対応するため、①ユーザが所有するスマートフォンのみで手軽に利用できる、②エリア内の任意の場所で利用できる、③複数人で同時に利用できる、の3つの要件を元に設計した。検証する音源定位手法として超指向性スピーカとムービングヘッドを用いた仮想オブジェクト追従型の音源定位手法を採用する。システムの利用例を図1に示す。アプリに重畳表示される仮想オブジェクト(図1(1))の実空間での位置を計算し、その場所を仮想オブジェクトの音源位置(図1(2))とする。仮想オブジェクトの音源位置に超指向性スピーカ(図1(3))を向け、超音波をその場所に投射する(図1(4))ことで、あたかもその場に音源があるように知覚させる。また、仮想オブジェクトの移動に合わせて超指向性スピーカの投射方向・角度を動的に変更することで、オブジェクトベースの音響システムを物理的に実現している。

システム構成を図2に示す。ユーザはスマートフォンARアプリを使用してAR体験を行う。このときアプリは角度情報と音声の再生信号を音源発生装置の制御用PCに送信する。また、実験に対応するため、同時にヘッドホン等に立体音響を出力する機能も実装した。

3.2 スマートフォンARアプリ

Unityで作成したアプリケーションであり、iOSアプリとして動作する。AR動作部分はARフレームワークであるARKit[10]を使用した。最初にiPhoneのカメラで空間キャ

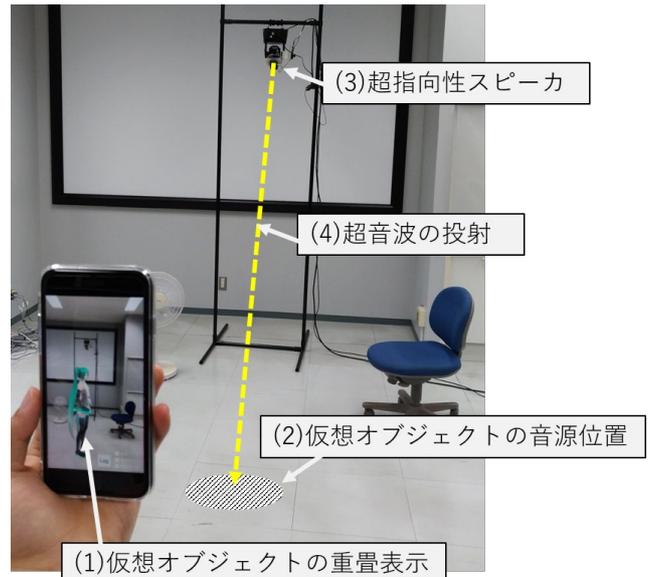


図1 本システムの利用例

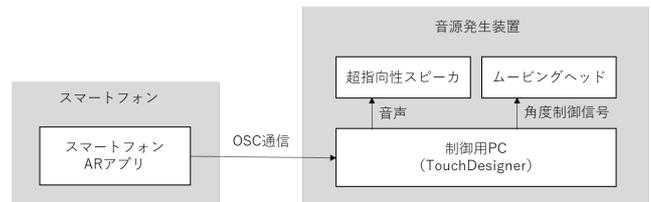


図2 システム構成

リブレーションを行っておくことで、World Mapが生成され、任意の場所に選択した仮想オブジェクトを出現させることができる。出現させた仮想オブジェクトはあとから位置や向きを微調整する機能も実装した。仮想オブジェクトには予めアニメーションが設定されており、再生ボタンを押すことでアニメーションと音声再生される。音声はResonanceAudio[11]を用いて、HRTFが畳み込まれたものを再生することができる。アニメーション再生前に、超指向性スピーカの位置もアプリ上で設定しておくことで、超指向性スピーカと仮想オブジェクトとの仰俯角や方位角が計算される。計算された数値と音声の再生情報はOSC(OpenSound Control)通信を使用して、外部に出力できる。このアプリは複数人での利用に対応しており、1台をマスター機として設定することで、各ユーザの画面に表示されている仮想オブジェクトの位置やモーションがマスター機のものと同様される。

3.3 音源発生装置

超指向性スピーカ K-02617(有限会社トライステート, FM変調, 周波数特性: 400Hz~5kHz)とムービングヘッド LM70S(BETOPPER, 可動域: 540度(方位角)・180度(仰俯角)), 制御用のPCから構成される。スマートフォンARアプリから送信される、超指向性スピーカの仰俯角・方位角、仮想オブジェクトの音声の再生情報を取得し、制御用PCから超指向性スピーカとムービングヘッドにそれぞれ信号を送る。これにより、音像を仮想オブジェクトの移動に合わせて追従させることができ、移動するユーザや複数人ユーザで音像を共有することも可能となる。

ムービングヘッド制御は TouchDesigner (Derivative) を使用している。

4. 実験

先行研究[2]では定位感を得るのが難しい仮想オブジェクトが存在することがわかった。そこで、定位感の精度向上を検討するため、1)音の反射面の素材によって反射面からの音の聞こえ方に違いがあるか(4.1節)を検証する。また、2)実空間の環境音を仮想空間上の音と同時に視聴する場合とそうでない場合の違い(4.2節)を検証することで、実空間の環境音が空間共有感の向上にどれだけ寄与するかを考察し、仮想オブジェクトの実在感向上につなげる。

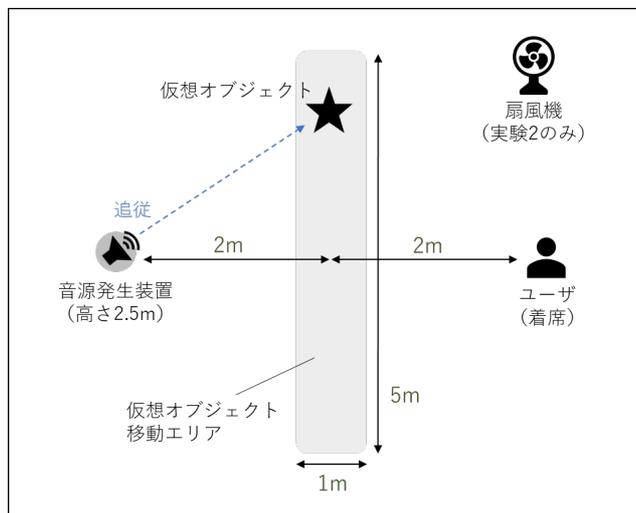
今回は被験者一人ひとりに対して実験を行い、複数人での同時利用は評価しない。実験環境を図3に示す。ユーザは着席した状態で音源を視聴する。ユーザから2m離れた位置に1m*5mの仮想オブジェクト移動エリアを設定し、ユーザから見て仮想オブジェクトが左右に移動できるようにする。音源の照射角度によってはユーザに直接超指向性スピーカからの反射音が当たるため、仮想オブジェクト移動エリアから2m離れた位置にムービングヘッドと超指向性スピーカを高さ2.5m(天井に近い高さ)の位置に設置する。超指向性スピーカは仮想オブジェクト移動エリア内に存在する仮想オブジェクトの位置に向けられ、仮想オブジェクトの移動に合わせて超指向性スピーカの向きも変化する。

4.1 音の反射面の素材を比較する実験(実験1)

超指向性スピーカの反射特性を変化させ、定位感の精度向上につなげるため、仮想オブジェクト移動エリアの床面の素材を変えて、反射面からの音の聞こえ方を比較する実験を行う。比較対象は硬い床、コルクマット、ダンボール、スポンジ、カーペットなどを予定している。被験者にはAR映像の視聴なしに、超指向性スピーカから再生した朗読音声と足音を聞いてもらう。提示された聴覚情報のみで、仮想オブジェクトがどこに位置しているか、どのように移動しているか、実在感を得られるかななどを、レーザポインタを用いた指示やアンケートにより回答してもらう。

4.2 実空間の環境音の影響を比較する実験(実験2)

この実験では実空間の環境音と仮想空間で発生した音が混在する音空間共有状態と、仮想空間で発生した音のみし



🔊 実験教室外スピーカ(実験2のみ)

図3 実験環境

か聞こえない状態とを比較し、実空間の環境音が空間共有感や仮想オブジェクトの実在感の向上に寄与するかを調べる。実空間の環境音を発生させるために、実験教室内に扇風機を設置し風量を強にする。加えて実験教室外の廊下に人がいるという状況を再現するため、実験教室の廊下側にスピーカを設置し予め録音した学生の話し声と足音を自然な音量で再生する。

音源提示要因として、超指向性スピーカ+ムービングヘッド、骨伝導イヤホン+HRTF、ヘッドホン+HRTFの3種類、環境音要因として環境音の有り、無しの2種類を用意し、計5条件(ヘッドホン+HRTFでは環境音無しのみ)で感じ方に違いがあるかを調べる。被験者にはスマートフォンARアプリを使用して仮想オブジェクトを提示し、条件を変えたときに実空間における環境音が仮想オブジェクトの存在感に影響するか、空間共有感に違いがあるかをアンケートにより評価する。

5. おわりに

AR体験において仮想オブジェクトの音源定位を行うことは、仮想オブジェクトの実在感向上に有用である。本稿は超指向性スピーカとムービングヘッドを利用した音源定位手法において、より多くの仮想オブジェクトや音源の種類で実在感を向上させることを目的として、1)音の反射面の素材の検討、2)実空間の環境音による空間共有感の評価を行うものである。今後は実験を行い、分析結果から仮想オブジェクトの実在感向上に適したARコンテンツを考察する。また、適切な音源定位手法の切替、ARコンテンツ視聴中における他ユーザ同士のコミュニケーション誘導などに応用することを検討する。

参考文献

- [1] 情報通信白書令和元年版, 総務省, <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nd112130.html> (参照 2020/07/12).
- [2] 下田桂輔, 福島拓: 超指向性スピーカを用いた仮想オブジェクトの実在感向上手法の検討, 情報処理学会研究報告デジタルコンテンツクリエーション(DCC), vol.2020-DCC-24, No.28, pp.1-6 (2020).
- [3] 越後宏紀 小林稔: アウェアネス支援のための足音伝達手法の検討, 情報処理学会研究報告グループウェアとネットワーク(GN), vol. 2019-GN-106, No.20, pp.1-8 (2019).
- [4] 金子翔麻, 渡邊恵太: SymPhonicAudio: 複数のスマートフォンを同期し現実空間への音像定位を動的に行う手法の提案, 情報処理学会インタラクション2018, 1B39(2018).
- [5] 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, 信学技報, Vol.105, No.556, pp.25-30(2005).
- [6] 伊藤仁一, 中山雅人, 西浦敬信, 木村朝子, 柴田史久, 田村秀行: 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, 電子情報通信学会技術研究報告, Vol112, No.385, pp.93-98(2013).
- [7] 北岸佑樹, 田中友樹, 米澤朋子: 多人数講義における講師分身エージェントを用いたコミュニケーションの促進, 電子情報通信学会技術研究報告, Vol.118, No.404, pp.107-114(2018).
- [8] Lindeman, Robert, Noma, Haruo, De Barros, Paulo: Hear-Through and Mic-Through Augmented Reality: Using Bone

Conduction to Display Spatialized Audio, Sixth {IEEE/ACM} International Symposium on Mixed and Augmented Reality, pp.173-176(2007).

[9] 高井一希, 朝倉巧: 骨導音を用いた誘導案内システムに関する基礎的検討, 第 23 回日本バーチャルリアリティ学会大会論文集, 11B-6 (2018)

[10] ARKit4, Apple, <https://developer.apple.com/jp/augmented-reality/arkit/> (参照 2020/7/12) .

[11] ResonanceAudio, Google, <https://resonance-audio.github.io/resonance-audio/> (参照 2020/7/19) .