

複数人 AR 体験を目的とした仮想オブジェクトの 実在感向上支援システム

下田 桂輔^{1,a)} 福島 拓^{1,b)}

概要：近年、スマートフォン端末で利用できる AR サービスが普及している。複数人でのスマートフォン AR 体験において、他ユーザとの空間共有感を維持しながら仮想オブジェクトの実在感を向上させることは臨場感のある AR 体験には重要であるが、これを評価している研究は少ない。本稿では音源定位による仮想キャラクタの実在感の向上や、他ユーザやキャラクタとの空間共有感の向上を目的として、超指向性スピーカによる音源定位手法とイヤホンによる音源定位手法を比較する AR ライブの予備評価実験を行った。被験者へのアンケートの結果、超指向性スピーカのほうが、キャラクタの実在感を得られる傾向がみられた。

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る [1]。利用シーンは多岐にわたり、スマートフォンゲームや音楽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・訓練などが考えられる。また、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レスの AR システムの開発が今後も進められていくと考えられ、AR 機能を搭載するスマートフォンにより手軽に AR 体験ができるようになりつつある。しかしスマートフォン AR においては、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り、仮想オブジェクトの出現位置が把握しづらい。さらに、AR コンテンツの仮想オブジェクトは AR 空間上を自由に動き回るため、出現場所やコンテンツの演出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スマートフォンを向けるよう促すために、仮想オブジェクトの音が出た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。一般的な HRTF (Head Related Transfer Function) を用いた手法では、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。そのため、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周囲の状況を把握できず危険である。さらに、仮想空間の音だけしか聞こえないことにより、実空間の光景と聞いている音が一致せず、仮想オブジェクトと自分自身が同じ空間を共有しているような感覚（以下、空間共有感と呼称する）が得られにくいと考えられる。

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], [3]。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源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]。特に、仮想オブジェクトが自分の周囲を横切るときに、実在感が得られやすいことが分かった。また、音源定位精度の向上を目的として、反射面素材の比較実験を行った結果、乱反射が発生しやすいカーペットなどの素材において音源定位感が比較的高まることが示された [3]。

しかし、このシステムは未だ実運用的な評価がされておらず、複数人で AR 体験を同時に行う状況下において、他ユーザや仮想オブジェクトとの空間共有感が得られるかどうかは不明である。本稿では実際に AR コンテンツを視聴してもらうことで、イヤホンを利用した視聴と比べて、超指向性スピーカを利用した手法のほうが、空間共有感が優

¹ 大阪工業大学

^{a)} m1m19a15@st.oit.ac.jp

^{b)} taku.fukushima@oit.ac.jp

位に働くかを検証する。

2. 関連研究

2.1 超指向性スピーカによる音源定位手法

超指向性スピーカとは、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。音源定位を可能とした従来研究には、音像プラネタリウム方式を用いた手法 [4] や視聴者に追尾して適切な音を提示する手法 [5] が存在するが、いずれも複数人が同時に視聴する環境での利用には課題が残っている。

フレキシブルパラメトリックスピーカを用いてオーディオスポットを制御する手法 [6] や Multi-PAL を用いて 3 次元音場を生成する手法 [7] では、複数人への音像提示を可能としている。しかし、仮想音源の位置をある程度固定する必要があり、AR 仮想オブジェクトのような自由に動き回る音源には対応していない。

北岸らは、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 [8]。プロジェクタで影と移動軌跡の描画、超指向性スピーカと 2 軸の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。本稿では同様の手法を AR 体験に適用し、エンタテインメント分野での有用性を検討する。

2.2 周辺環境の影響に関する研究

森川らは仮想エージェントとの会話において、雑音環境下や会話に参加しない非参加者の存在により、ユーザの発話が促進されるかどうかを検証している [9]。その結果、非参加者が存在する場合、環境音や喧騒音を発生させることでユーザの発話が促進されることが示唆されている。本稿でも周辺環境の音や他人の存在が、AR 体験に影響を及ぼすかを検証し、耳を塞がない超指向性スピーカの有用性を補強する。

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。AR ライブや対戦ゲームなど複数人の体験共有型リアルタイムコンテンツに対応するため、仮想オブジェクトにリア

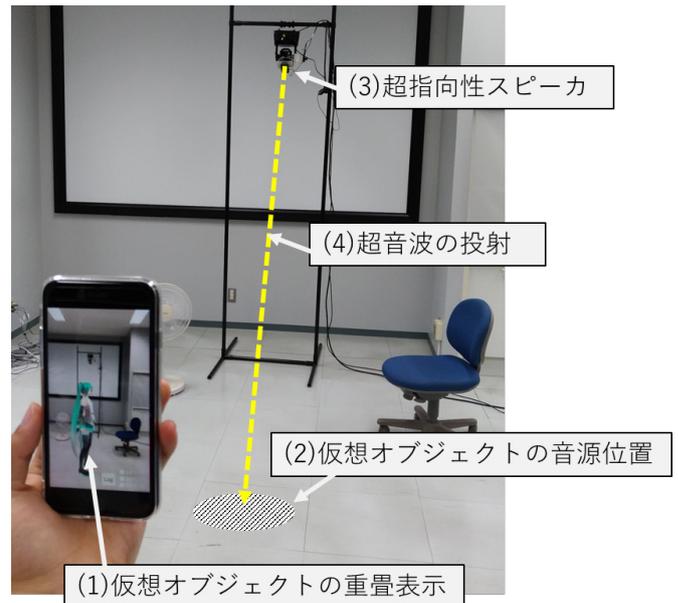


図 1 本システムの利用例

ルタイムで追従が可能な、超指向性スピーカとムービングヘッドを用いた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、同時にイヤ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 の制御部分は AR フレームワークである ARKit [10] を使用した。最初に iOS 端末のカメラで空間キャ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。音声は後述の音声発生装置で再生されるが、実験での比較

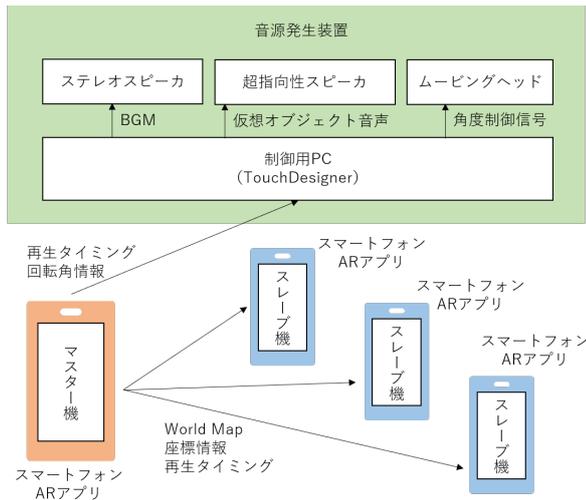


図 2 システム構成

用としてスマートフォン上で HRTF を用いた立体音響 [11] も再生することができる。また、アニメーション再生前に、超指向性スピーカの位置を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、World Map を複数端末と共有できる。1 台をマスター機として設定することで、スレーブ機の画面に表示されている仮想オブジェクトの座標やモーションの再生タイミングがマスター機のものと同期される。

3.3 音声発生装置

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, 一般的なステレオスピーカから構成される。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タイミングを取得し、制御用 PC からムービングヘッドに回転角情報を、超指向性スピーカに音声を送る。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が可能となる。ムービングヘッドの制御は TouchDesigner (Derivative) を使用している。今回の実験では、超指向性スピーカからキャラクターの歌声を再生し、ステレオスピーカから音楽を再生した。この 2 つの音声は予め同期されており、音楽に合わせて歌声が流れる。

4. 予備評価実験

4.1 実験概要

本システムは今までに複数人での AR 体験を評価してい

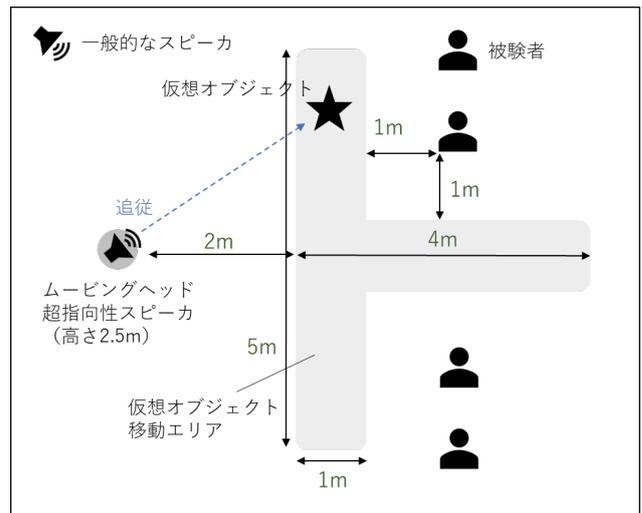


図 3 実験環境

なかつたため、実運用上のコンテンツとして「AR ライブ」を想定して実験を行った。AR ライブとはバーチャルキャラクターが AR 空間上で歌って踊るライブのことである。本稿では 4 人の被験者に AR ライブを視聴してもらい、超指向性スピーカ手法がイヤホン手法より、1) キャラクターの実在感が向上するか、2) 空間共有感が向上するか、を検証した。

実験環境を図 3 に示す。仮想オブジェクト移動エリアや被験者周辺には床面にカーペットを敷いている。被験者は着席した状態で音源を視聴する。被験者の周囲に仮想オブジェクト移動エリアを設定し仮想オブジェクトが自由に移動できるようにする。音源の照射角度によっては被験者に直接超指向性スピーカからの反射音が当たることが予想される。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 の位置に設置し、被験者に直接超指向性スピーカからの音声が当たらないようにする。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。

まず、被験者にはスマートフォンで周囲を見渡してもらい、空間キャリブレーションを行う。その状態であらかじめ設定しておいた World Map を適用し、座標空間をすべての端末で同期する。AR ライブは 2 回行い、1 回目は HRTF を用いたイヤホン手法 (従来手法)、2 回目は超指向性スピーカ手法 (提案手法) で行う。1 回のライブはどちらも 4 分程度で、仮想オブジェクトは特定のタイミングで被験者の視野外と予想される場所に瞬間移動する演出を含めた。また、被験者にはタスクを設定し、1 回目は 2 人に対してペンライトを振る、2 回目は別の 2 人に対してキャラクターに手を振る、というタスクとした。

1 回の AR ライブごとにアンケート (リッカートスケール, 5 段階評価) を行い、2 回のライブの結果を比較した。

表 1 「キャラクタが存在する方向が分かった」の主観評価

比較条件	評価段階*1*2					平均値	中央値
	1	2	3	4	5		
イヤホン	0	1	1	0	2	3.75	4
超指向性スピーカ	0	0	0	2	2	4.5	4.5

表 2 「キャラクタを見失ってもすぐに発見できた」の主観評価

比較条件	評価段階*1*2					平均値	中央値
	1	2	3	4	5		
イヤホン	1	2	0	1	0	2.25	2
超指向性スピーカ	0	0	1	3	0	3.75	4

表 3 「キャラクタがその場に実在していたように感じた」の主観評価

比較条件	評価段階*1*2					平均値	中央値
	1	2	3	4	5		
イヤホン	0	1	2	1	0	3	3
超指向性スピーカ	0	0	1	2	1	4	4

表 4 「同時にライブを見ている周りの観客の存在感があった」の主観評価

比較条件	評価段階*1*2					平均値	中央値
	1	2	3	4	5		
イヤホン	0	1	0	2	1	3.75	4
超指向性スピーカ	0	0	0	3	1	4.25	4

4.2 実験結果

アンケートによる主観評価を表 1～表 4 に示す。

表 1 より超指向性スピーカにおいてキャラクタが存在する方向が分かりやすい傾向にあることが分かった。さらに表 2 より、キャラクタが AR ライブ中に瞬間移動して見失ったときにキャラクタがすぐに発見できた被験者は超指向性スピーカのほうが多かった。これらの結果から、イヤホンより超指向性スピーカのほうが音源の定位感が比較的得られることがわかった。それに伴い、キャラクタの実在感も超指向性スピーカのほうが高い傾向にあった(表 3)。しかし、同時にライブを見ている周りの観客の存在感については、イヤホンと超指向性スピーカの比較では大きな差はなかった(表 4)。これは、被験者が実際の一般的な音楽ライブと比較した評価をしたことや、被験者がタスクを積極的に行わなかったことで今回の比較では差が出なかったと考えられる。正しく評価するために被験者に課すタスクの内容や質問項目を見直す必要がある。

5. おわりに

本稿では、複数人でのスマートフォン AR 体験において、超指向性スピーカによる音源定位手法とイヤホンによる音源定位手法を比較する AR ライブの予備評価実験を行った。被験者へのアンケートの結果、超指向性スピーカのほ

うが、キャラクタの定位感を得られる傾向があった。それに伴い、キャラクタの実在感も得られる傾向がみられた。今後は実験設定を見直し、サンプル数を増やした実験を行うことで、本システムについてより詳細に評価する。

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 2020/12/1)
- [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] 下田桂輔, 福島拓: 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, 情報処理学会関西支部大会, C-05 (2020).
- [4] 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, Vol.112, No.385, pp.93-98 (2013).
- [5] 池田輝政, 遠藤正隆, 中嶋裕一, 三浦哲郎, 菱田隆彰: 顧客の意思を忖度するデジタルサイネージ広告システム, 情報処理学会エンタテインメントコンピューティングシンポジウム 2017 論文集, Vol.2017, pp.439-440 (2017).
- [6] 小森慎也, 益永翔平, 生藤大典, 中山雅人, 西浦敬信: フレキシブルパラメトリックスピーカを用いたオーディオスポットの制御, 日本バーチャルリアリティ学会論文誌, Vol.20, No.3, pp.189-198 (2015).
- [7] 小森慎也, 大上佳範, 中山雅人, 西浦敬信: マルチパラメトリックアレイスピーカを用いた音像ホログラムの構築, 電子情報通信学会論文誌, Vol.J101-D, No.3, pp.578-587 (2018).
- [8] 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).
- [9] 森川大輔, 石井裕, 渡辺富夫: VR 空間における複数非参加者を含む雑音環境を利用した発話支援システム, ヒューマンインタフェース学会論文誌, Vol.22, No.4, pp.403-410 (2020).
- [10] ARKit4, Apple, <https://developer.apple.com/jp/augmented-reality/arkit/> (参照 2020/12/1).
- [11] ResonanceAudio, Google, <https://resonance-audio.github.io/resonance-audio/> (参照 2020/12/1).

*1 表中の評価段階の数字は人数を表す。

*2 評価段階: 1:強く同意しない, 2:同意しない, 3:どちらともいえない, 4:同意する, 5:強く同意する