

空中に浮かぶ仮想音源の直接操作感覚に関する研究

Investigation of perception of directly manipulating mi-air virtual sound source



羽田 陽一 (Yoichi HANEDA, Ph. D.)

電気通信大学大学院 情報学専攻 教授

(Professor, The University of Electro-communications)

日本音響学会, アメリカ音響学会, IEEE, AES, IEICE

受賞: 電気通信普及財団賞 (2015 年), 日本音響学会技術開発賞 (2007 年),

電子情報通信学会論文賞 (2002 年), 日本音響学会佐藤論文賞 (2002 年),

電子情報通信学会業績賞 (1998 年) 他

著書: 音響学入門ペディア コロナ社 (2017 年)

研究専門分野: 音響信号処理

あらまし

重さも実体もない空中に浮かぶ音像を動かしたり、広げたりするような直接的な操作を行うと、人はどんな感覚を得るのだろうか、という素朴な疑問から本研究はスタートしている。従来から目の前に飛び出すような音像を生成する波面合成法と呼ばれる手法はあったが、これを用いてユーザの動作に合わせて、リアルタイムに仮想音源を移動することはあまり行われてこなかった。しかし近年、モーションキャプチャや計算機能力の高度化により、空中音像をリアルタイムに操作できるようになってきた。このような背景のもと、私たちは、幅が 3m 超の直線スピーカアレイとモーションキャプチャ、人が入れる半径 1.125m の球面スピーカアレイとモーションキャプチャ、の 2 つのインタラクティブオーディオシステムを構築してきており、本研究ではこれらを用いて空中音像に対する直接操作感に関するいくつかの実験を行ったので報告する。

1. 研究の背景

近年、モーションキャプチャを利用して人の動きに同期して映像を変化させたり、ユーザ自身をアバター

に変化させたりするバーチャルリアリティ技術が進展してきている。これらの技術は主に視覚情報をユーザにフィードバックしており、付随的に音に変化を加えることなどがなされている。一方で、近年複数のスピーカを用いた立体音響技術が発展してきており、ユーザ動作に同期して空間的な音の位置を移動させることもできるようになってきた[1]。

しかし、一般的な立体音響技術は、左右や上下といった音の到来する方向を空間的に再現している場合が多く、例えば目の前に音像が迫ってくる、手の位置から音が聞こえる、といった奥行きや距離感まで含めた立体音響技術はまだ一般的とはなっていない。

これに対して私たちは以前から、波面合成法*1 [2] と呼ばれる技術に着目し、奥行き方向を含めた立体音響技術に関する研究を行ってきた。この技術は音が従うべき物理方程式に則って音の波面を再現する技術であり、空中のその場所に音源があるのと同様の波面を合成することで、奥行きを含めた音像定位を実現する。すなわち、波面合成技術を使えば、目の前に物体がなくてもリスナーの目の前に音像を浮かばせることができる。

このような技術的な背景の下、リスナーが実体のない空中音像を移動させたり、広げたりする操作を行った場合[3]に、どのような感覚が生じるかについてはいまだ明らかになっていない。

2. 研究の目的と進め方

本研究は、空中に浮かぶ仮想的な音源を移動させたり、広げたりする行為により重さも実体もない“音”という存在に対して、直接操作している感覚が想起されるのか、また想起させるための条件はなにかを明らかにすることを目的とする。

仮想音源生成には、32 個のスピーカを直線上に等間隔 (10 cm 間隔) に並べた直線スピーカアレイ、および半径 1.125 m の球面上に 42 個のスピーカを配置した球面スピーカアレイを使用する。仮想音源は、直線スピーカアレイでは焦点音源法[4]、球面スピーカアレイでは Near-Field Compensated Higher Order Ambisonics (NFC-HOA)[5]を用いて再現する。また、本研究では、これらの仮想音源の位置を手の動きに合

空中に浮かぶ仮想音源の直接操作感覚に関する研究

Investigation of perception of directly manipulating mi-air virtual sound source

わせて滑らかに移動させるために、時間領域表現した手法をそれぞれに導入する[6]。手の位置を取得するためのモーションキャプチャには、Azure Kinect を使用する。これらの装置を組み合わせ、手の位置に応じてリアルタイムに移動する仮想音源を受聴者に聴取させる。

本研究では、大きく2つの実験を行う。1つめは、直線スピーカアレイを用いて、受聴者の手の動きに合わせて仮想音源を移動させた際に、仮想音源と手の位置関係によって操作感覚が異なるかを調査する。2つめは、仮想音源を両手で広げる動作をしながら、見かけの音源幅を変化させた音を聴取する実験を行う。この実験は球面スピーカアレイを用いて実施し、上下方向を含めてどのように音の広がりを感じるかについて調査する。

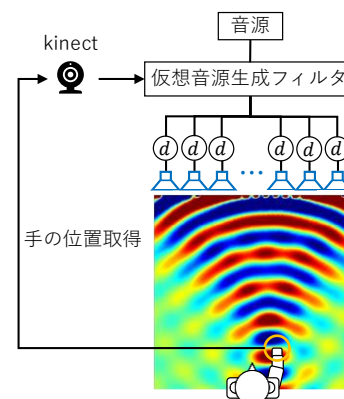


図1 直線スピーカアレイ波面合成システム

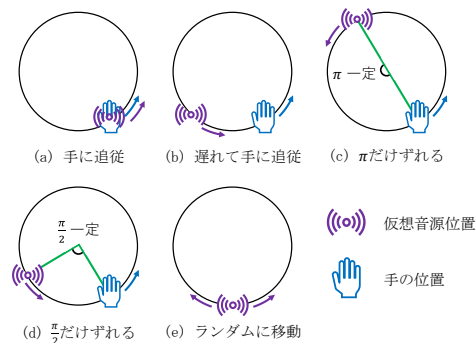


図2 仮想音源移動パターン

3. 手への追従性の違いによる仮想音源の操作感

まず第1の実験では、手の動作に応じて移動する仮想音源[7]に対して、手との位置関係および移動経路の違いが操作感に与える影響について調査する。図1に実験に用いる直線スピーカアレイ波面合成システムと合成される波面の模式図を示す。この装置を用いて、手の動きに同期しながら移動する仮想音源と、投げ動作によって手の位置から離れて移動する仮想音源の2つについてそれぞれ主観評価実験を行う。

3.1 手に追従する仮想音源に関する知覚実験

被験者は直線スピーカアレイ正面の所定位置に立ち、水平面で円を描くように手を動かしながら仮想音源による空中音像を聴取する。この実験では、手の同じ動作に対して、図2に示すように仮想音源を異なる位置で移動するように合成する。具体的には、仮想音源の移動パターンとして、(a) 手の位置に仮想音源が追従、(b) 手の動きに遅れて追従、(c) 円周上で手の位置から180度（ π ）ずれた位置を移動、(d) 円周上で手の位置から90度（ $\pi/2$ ）ずれた位置を移動、(e) 手の動きとは無関係に移動、の合計5つを用意した。これらのパターンはランダムに選択され、被験者はどのパターンが提示されたかを知らない状態で手を動かしながら空中音像を聴取し、操作感覚に対して、操作している

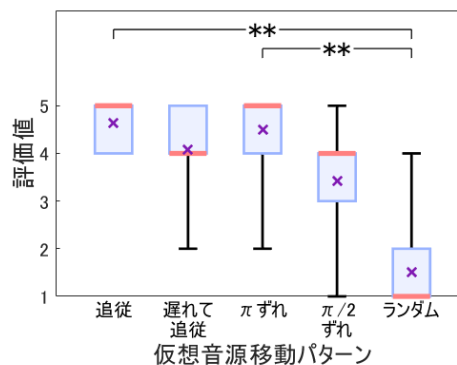


図3 操作感の回答結果

と思う場合は5、そう思えない場合は1という5段階で回答した。

図3にそれぞれの移動パターンに対する評価値を示す。赤色の線は中央値、紫色の「×」は平均値を示す。結果より、仮想音源が手の位置に一致して移動（「追従」）

空中に浮かぶ仮想音源の直接操作感覚に関する研究

Investigation of perception of directly manipulating mi-air virtual sound source

する場合に限らず、「遅れて追従」、「180度(π)ずれ」、「90度($\pi/2$)ずれ」のように手から離れて追従する場合においても操作感が得られることが分かった。以上より、手と仮想音源の位置が一致していない場合でも、手と仮想音源の動きに同期性がある場合には操作感が得られることがわかった。これは直接操作という観点で言うと、普段操作することのない音像が自分の意思に基づいて移動しているため、時間あるいは空間にずれがあったとしても操作している感覚が想起されたと考えられる。

3. 2 投げ動作と仮想音源の移動に関する実験

本節では、図4に示す投げ動作により手元から離れて移動する仮想音源を扱う。このインタラクションでは、前節と異なり、手の動作と仮想音源の移動が同期するのは仮想音源が手から離れるまでであり、手から離れた後の仮想音源にはインタラクションは発生しない。本節では、このような手から離れて移動する音像の移動知覚に着目する。

先行研究より、投げ動作によって手から離れて移動する仮想音源を生成した際に、投げ感覚が得られることが報告されている[8]。ただし先行研究では、投げ動作後の仮想音源の移動の様子がどう感覚に影響するかまでは調査されていない。そこで、本研究では、被験者に対して、「〇m 先に音を投げるような動作」を指示し、一方で指示した投げ動作とは異なる位置で仮想音源を止める、という実験を行った。主観評価実験では、指示とは独立に仮想音源がどこまで到達したかを回答する。具体的には、被験者に仮想音源を1m先、2m先、4m先のいずれかの距離に到達するように仮想音源を投げてもらい、実際には仮想音源の終点位置を1m、2m、4mのいずれかに設定する。これらの組み合わせの9条件に対して主観評価実験を行った。

図5に回答結果を示す。図5において、横軸は組み合わせ条件で、「(指示した距離、知覚した距離)」であり、縦軸は評価値である。結果より、被験者が想定した終点位置よりも、実際に提示した終点位置に仮想音源位置を知覚することが分かった。この結果は手から離れて移動しはじめた音像は、もはや手の動作とは無関係な音像となってしまったことを意味する。3.1



図4 投げ動作による仮想音源の移動

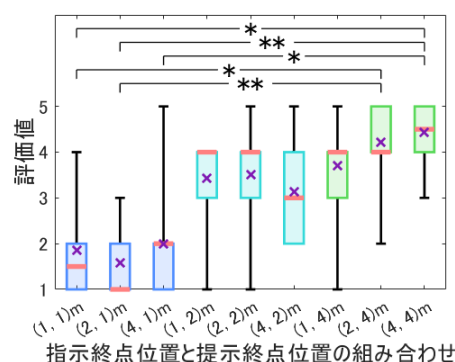


図5 投げ動作後の音像位置の回答結果

節の実験では手の動きと音像の移動が同期していたため、空間的に離れていても操作感が継続したと考えられる。一方で、本実験では手から離れてしまった仮想音源は、手の動作と同期していないため、操作感がなくなり、独立した音像として知覚されたためと考えられる。すなわち、操作感を得るためには、操作に対して同期したインタラクションが重要と考えられる。

4. 広げ動作に応じた音像の広がり方の知覚

本節では第2の実験として、ユーザが両手で音像を広げるような動作を行った際の、音像の広がり知覚を調査する。従来、左右に置かれたステレオスピーカから同じ音、すなわち信号間相関*2が強い信号を流すと、ステレオスピーカの間に1つのシャープな音像が知覚され、一方で相互相関が弱いステレオ信号を放射すると、音像が広がって聞こえることが報告されている[9]。本研究ではこの特性に基づき、ユーザの両手の動きに応じて2つの音源の信号間相関を変化させるインタラ

空中に浮かぶ仮想音源の直接操作感覚に関する研究

Investigation of perception of directly manipulating mi-air virtual sound source

クションを構築し、音像の操作に音の広がり感を与える。ここで、本研究では左右だけでなく上下の手の動きも想定し、仮想的に上下に音源を配置可能な半径1.125mの42チャンネル球面スピーカアレイを実験に用いる（図5参照）。なお、両手の動作のモーショントラッキングには前節と同様 Azure Kinect を用いる。

4.2 音像を広げる操作と信号間相関制御

2つの音源は仮想音源として、図6に示すように被験者から見て開き角が 20° になるように左右と上下に固定する。信号間相関は両手間距離に合わせて相関の強い信号から弱い信号へと変化させることで、徐々に音像が広がる様子を表現する。また、2つの音源の配置方法（左右配置と上下配置）は、そのまま相関信号の物理的な変化方向に対応している。

実験では、図7に示すように被験者に両手を左右方向または上下方向の2つの方向に広げ動作を行ってもらい、左右方向または上下方向に信号間相関を連続的に変化させた際の音像の広がり知覚を主観評価した。各試行後に、見かけの音源の幅（ASW: Apparent Source Width）[9] について、左右方向と上下方向の幅の変化として感じたか、感じなかったかを5段階尺度で評価した。評定語は「1: 感じない」「2: あまり感じない」「3: どちらともいえない」「4: やや感じる」「5: 感じる」とした。条件は、2つの手の動かし方（水平（左右）、垂直（上下））、2つの仮想音源配置（水平、垂直）、相関制御手法の違いとし、各2回ずつランダム順に提示した。原信号には、オーケストラのモノラル音源を用いた。

代表的な結果として、無相関 FIR フィルタを用いて相関制御を行った際の結果のまとめを表2に示す。

表2 両手動作と相関方向別広がり知覚（まとめ）

両手動作	水平（左右）		垂直（上下）	
	水平	垂直	水平	垂直
水平広がり知覚	◎	×	○	×
垂直広がり知覚	×	×	◎	△

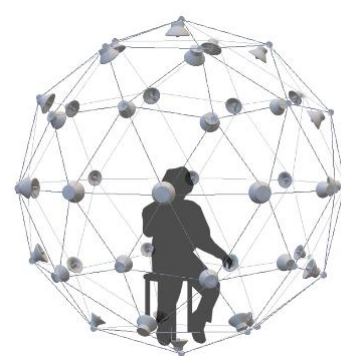
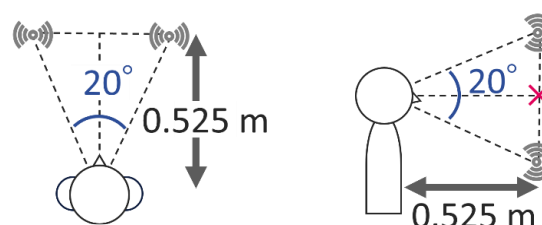


図5 球面アレイのスピーカ配置図



(a) 水平配置

(b) 垂直配置

図6 仮想音源の配置方向



(a) 水平（左右）方向



(b) 垂直（上下）方向

図7 両手を広げる方向

表2から、手を水平（左右）に動かした場合、水平配置の音源では水平方向への広がり知覚された。これは手の動作がなくとも従来から知られている結果と同じである。また、音源を上下に水平配置した場合には、左右の耳に入る音は同じであるため、広がり感の変化が知覚されないことが想定されるが、実際そのような結果が見られた。

一方、手を上下に動かした場合、水平配置の音源においては左右の広がりだけではなく、上下方向の広がり知覚された。これは、実際には上下方向には音は

空中に浮かぶ仮想音源の直接操作感覚に関する研究

Investigation of perception of directly manipulating mi-air virtual sound source

広がっていないにも関わらず、身体動作あるいは手を見ているという視覚による非聴覚の手がかりが音像の知覚に影響を与えた可能性を示唆している。ただし、音源を水平配置したパターン、すなわち左右の耳に入る信号の相関が同じままであるパターンでは、広がりさはほど見られなかった。つまり、身体動作あるいは広げている手の視覚的な情報のみで音が上下（垂直）方向に広がったわけではないこともわかる。この実験から、実際には左右方向に音が広がっている知覚をもたらす状況にも関わらず、上下方向に手を動かすと、上下方向にも音が広がったように感じるという新たな発見をもたらした。以上の結果は、音像の広がりをインタラクティブに制御する際には、音源は左右方向に配置するだけで制御可能であることを示唆している。

5. まとめと将来展望

波面合成や NFC-HOA と呼ばれる技術は、複数のスピーカから放射された音の干渉を利用し、空中に配置した仮想音源から放射されるであろう物理的な波面を合成することで、リスナーの目の前や手の届く位置に音像を提示することができる技術である。本研究は、この仮想音源の生成位置を、バーチャルリアリティで用いられるモーションキャプチャ装置で取得した人の手の位置に追従してリアルタイムに変更することで操作可能とし、その際の操作感について調査を行ったものである。その結果、操作する手の位置と提示された音像の位置関係に時間的・空間的にずれがあったとしても動作と同期した移動がある場合には操作感が得られることが分かった。一方で、手から離れて操作できない状態になった場合には、音像の位置は操作から想定される位置ではなく、直前の動作とは無関係に実際の音像の位置に定位した。これは、操作感を継続するためには、常に操作対象の音像が人からのインタラクションと同期している必要があることを示唆している。

また、これとは独立な実験として、音像を左右あるいは上下に広げる動作をしながら、これに合わせて左右あるいは上下に信号間相関が変化するステレオ仮想音源を聴取する実験を行った。この実験において、左右に手を広げる動作に合わせて、左右の信号間相関を弱めた場合には、音像が左右に広がって聞こえるとい

う従来の知見から予測された結果が得られた。一方、手を上下に広げながら、左右の信号間相関が低下する音を聞いた場合には、左右だけではなく上下方向にも音像が広がって知覚されることがわかった。これは、手を上下に動かすというインタラクションが、本来は左右にしか音像が広がって聞こえないはずの音像に対して上下にも広がって聞こえる、という新たな知覚現象の発見を示唆している。

以上のように、これまで静止して聞くことが多かった音に対して、動かす、投げる、広げる、といった積極的なインタラクションを加えることで新たな感覚が得られる可能性があることが分かった。本実験では、操作している感覚までは得られることがわかったが、実体感の有無までは明らかにすることはできていない。今後は、仮想音源に実体感を持たせるための要件の明確化や、インタラクションをとおして音に対して新たな価値を付加することで、新しいユーザインタフェースの提供や、エンターテインメント分野への応用を進める。

用語解説

*1 波面合成

音源（スピーカや話者、楽器など）がその場に作り出す波面を、複数のスピーカから出た音の干渉を利用して合成する技術。システムとしては、概ね 10cm 間隔で 32 個程度のスピーカを必要とする。

*2 信号間相関

厳密な言葉ではないが、本報告では、2 つのスピーカから流れる元の 2 つの音がどれくらい似ているかを表している。例えば、ステレオスピーカの左右から同じ音を流すと 2 つのスピーカの間に 1 個の音像として知覚することが知られている。この場合、左右の音は同じ、すなわちとても似ているので相関は強いと表現する。一方で、まるっきり似ていない信号だと左右に分離してしまい、広がったという印象にならないため、疑似的に相関を弱める手法を採用している。

空中に浮かぶ仮想音源の直接操作感覚に関する研究

Investigation of perception of directly manipulating mi-air virtual sound source

参考文献

- [1] 粕谷貴司, 塚田学, 菰原裕, 高坂茂樹, 水野拓宏, 野村譲誉, 江崎浩, “インタラクティブな遠隔ライブ VR 配信プラットフォーム,” 情処論文誌: デジタルコンテンツ (DCON) トランザクション, 7(2), pp. 1-14 (2019).
- [2] A. J. Berkhout, D. de Vries, P. Vogel, “Acoustic control by wave field synthesis,” *The Journal of the Acoustical Society of America*, Vol. 93 (5), pp. 2764-2778 (1993).
- [3] J. Müller, M. Geier, C. Dicke, S. Spors, “The boomRoom: mid-air direct interaction with virtual sound sources,” In *Proc. of SIGCHI conf. on human factors in computing systems*, pp. 247-256 (2014, April).
- [4] S. Spors, H. Wierstorf, M. Geier, J. Ahrens, “Physical and perceptual properties of focused sources in wave field synthesis,” In *127th AES Convention*. Audio Engineering Society (AES) (2009, October).
- [5] J. Daniel, M. Sébastien, “Further study of sound field coding with higher order ambisonics,” *AES Convention*. Vol. 116. No. 5. Berlin: Audio Engineering Society (2004).
- [6] 羽田陽一. “インタラクティブオーディオに向けた焦点音源のリアルタイム生成——空中音像とその直接操作——”, 日本音響学会誌, vol. 79 (10), pp. 502-509 (2023).
- [7] M. Hirohashi, Y. Haneda, “Subjective Evaluation of a Focused Sound Source Reproducing at the positions of a Listener’s Moving Hand,” 2023 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), Taipei, Taiwan, pp. 2395-2401 (2023) .
- [8] 山崎萌恵, 羽田陽一, “投動作によって投げられたように移動する焦点音源の受聴について”, 日本音響学会春季研究発表会講演論文集, pp. 243-244 (2023 年 3 月).
- [9] J. Brauert, 森本政之, 後藤敏幸, “空間音響,”

鹿島出版, 1986.

関連文献

- [10] K. Sawajiri, Y. Haneda, “Subjective Evaluation of Interactive 3D Space Virtual Sound Source with Hand Tracking,” In *2024 IEEE 13th Global Conference on Consumer Electronics (GCCE)*, pp. 425-426 (2024 October).
- [11] T. Suefuji, K. Sawajiri, Y. Haneda, “Perceptual evaluation of a virtual sound source reflected by hand using a linear loudspeaker array,” *The Journal of the Acoustical Society of America*, 158, A143-A144 (2025).
- [12] K. Sawajiri, Y. Haneda, “Influence of interactive interaural correlation control on the apparent source width in loudspeaker-based reproduction,” *The Journal of the Acoustical Society of America*, 158, A143-A143 (2025).
- [13] 澤尻 晃大, 羽田 陽一, “相関制御に基づくインタラクティブ仮想音源の見かけの音源の幅制御の検討”, 日本音響学会第 153 回 (2025 年春季) 研究発表会講演論文集 (2025 年 3 月) .
- [14] 澤尻 晃大, 羽田 陽一, “手で音像を広げる操作における音像知覚に対する手の位置情報と能動性の影響について”, 日本音響学会第 154 回 (2025 年秋季) 研究発表会講演論文集 (2025 年 9 月).
- [15] 末藤 立己, 澤尻 晃大, 羽田 陽一, “仮想壁に反射するインタラクティブ波面合成音源の移動知覚と跳ね返し操作について”, 日本音響学会第 154 回 (2025 年秋季) 研究発表会講演論文集 (2025 年 9 月).

この研究は、令和 5 年度 S C A T 研究助成の対象として採用され、令和 6 ～ 7 年度に実施されたものです。