

研究終了報告書

「分散配置アレイによる音空間の記録・再生技術基盤の構築」

研究期間：2018年10月～2022年3月

研究者：小山 翔一

1. 研究のねらい

音空間を記録あるいは伝送し、再生しようという技術的な試みは、遠隔地間でのコミュニケーションや音楽などの記録・保存を目的とし、古くから取り組まれてきた。多数のマイクロフォンとスピーカを用いて、広い領域の音空間を忠実に記録・再生することを目指す音場收音・再現技術は、聴覚におけるバーチャルリアリティを実現する上での究極的な目標であり、旧来のサラウンドステレオやバイノーラル再生と比べても、その優位性は大きい。このような技術の概念自体は古くから存在しているが、音響センサ・アクチュエータの小型化や低コスト化、A/D・D/A 変換器の多チャンネル化、計算機の高性能化に伴ってシステムとしての実現可能性が高まったことで、急速に注目を集めている研究分野である。

音場收音・再現技術は、マイクロフォン・スピーカアレイを用いた音場の再構成を目的とし、従来は音場の境界値積分表現のような物理方程式を理論的根拠としたものがほとんどであった。このような従来技術では、マイクロフォン・スピーカ素子を平面や直線、球面や円形といった単純なアレイ形状で配置する必要がある。さらに、高い周波数帯域までの再現精度を保つためには多くの素子数を必要とし、再現可能な周波数の上限値である空間ナイキスト周波数以上では、空間エイリアシングと呼ばれる誤差が生じる。このようなアレイ素子配置の制約は、システムの大規模化や高コスト化を招き、音場收音・再現技術の適用範囲を限定してしまう要因となると考えられる。

本研究では、アレイ素子を任意の位置に配置可能な音場收音・再現技術について検討を行う。このような技術が可能になれば、システムの大規模化・高コスト化を避けることができ、より柔軟なシステム構成を用いて広いアプリケーションに適用可能となるほか、将来的にはスマートフォンやタブレット PC、スマートスピーカなどの身の回りにあるマイクロフォン・スピーカを搭載した様々なデバイスをアレイ素子として利用できるようになると期待できる。また、学術的な面においても、従来の波動論のみに依拠したような決定論的なアプローチを適用することができないため、分散配置アレイ信号のための新たな方法論を構築する必要がある。よって本研究では、分散配置したアレイ素子を用いた音場收音・再現の基盤となる技術の構築を目的とする。

2. 研究成果

(1) 概要

音場分析・合成の基礎技術、およびその応用技術において、幅広い研究成果が得られた。研究開始当初はバーチャルリアリティ(VR)としての音場收音・再現への応用のみを想定していたが、COVID-19などによる社会情勢の変化から、応用範囲を拡大して研究に取り組んだ。

まず、音場分析・合成の基礎技術に関しては、無限次元波動関数展開に基づく音場分析法、スパースモデリングに基づく音場分析法、多点振幅制御による音場合成法について、理

論的な進展が得られた。無限次元波動関数展開については、音源方向の事前情報を組み入れることによる高精度化を実現したほか、マルチカーネル学習の枠組みを取り入れるための基礎検討も行った。スパースモデリングについては、対象領域の格子点への離散化を必要としない、Reciprocity Gap Functional と呼ばれる逆問題的な手法を提案しており、annihilation filter 法によるロバスト化の検討も行った。多点振幅制御法は、制御点上の振幅と位相を所望のものとは一致するようにスピーカ駆動信号を求める従来の多点音圧制御法と異なり、振幅のみを所望のものとは一致させ、位相は任意とする手法である。これにより、マルチゾーン音場合成による音のゾーニングなどに応用可能な技術を実現した。

音場分析・合成のためのセンサ・アクチュエータ配置の最適化についても取り組み、特に多点音圧制御法による音場合成において、スピーカ・制御点の配置を最適化する手法について、提案する empirical interpolation method と従来のセンサ配置最適化法を含めた網羅的な比較評価を行った結果を総説論文としてまとめた。さらに、音場のカーネル補間法を応用して新たに設計したコスト関数による音場推定のためのセンサ配置最適化法についても様々に検討を行った。

応用技術については、VR のためのバイノーラル再現や音場再現、あるいは領域的な騒音抑圧を目的とした空間アクティブ騒音制御について検討を行った。特に実環境において収録した音場をバイノーラル再現する技術については、複数の小規模アレイによる実験システムを構築し、客観・主観評価も行ったほか、国際会議においてデモ発表なども行っている。空間アクティブ騒音制御では、従来誤差マイク位置付近のみに限られていた騒音抑圧効果を領域的に実現するため、音場のカーネル補間法を応用した手法を構築した。

(2) 詳細

研究開始当初は音場收音・再現というバーチャルリアリティへの応用のみを想定していたが、COVID-19 などによる社会情勢の変化も鑑み、応用範囲を拡大して研究に取り組んだ。よって、分散配置アレイによる音場の分析・合成の基盤技術とその応用技術における研究成果について

以下に示す(図1)。



図1 音場分析・合成とその応用技術

研究テーマA「音場分析・合成の基盤技術の確立」

音場の分析と合成それぞれにおいて、様々な応用技術における基盤となる基礎技術を確立した。主に、(1) 無限次元波動関数展開による音場分析法、(2) スパースモデリングに基づく音場分

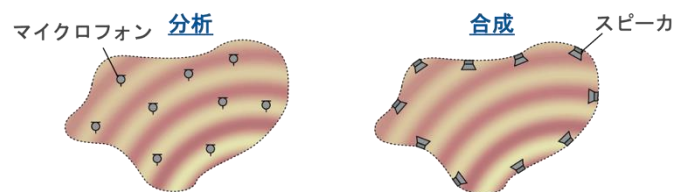


図2 音場の分析と合成

析法、(3) 多点振幅制御による音場合成法、について検討を行った。

(1) 無限次元波動関数展開による音場分析法

音場分析の基本的な問題は、離散配置されたマイク信号から、対象領域内の音場を推定する、というものである。ここで、推定された音場は周波数領域においてヘルムホルツ方程式に従っているはずであり、これは物理法則由来の事前情報と言える。このようなヘルムホルツ方程式制約化での音場推定を実現するには、その要素解である平面波関数や球波動関数などに基底展開することがよく行われる。通常は基底展開の展開次数を有限次元で打ち切ることが行われるが、我々はこれを無限次元化することによる手法を提案してきた。これにより、展開次数や展開中心などを経験的に定める必要がなくなり、より広い条件下で高い推定精度を達成することが可能となる。基本的な結果は研究開始前に得られていた [Ueno+ IEEE SPL 2018, IWAENC 2018]が、音源方向の事前情報を加えることによる高精度化[(1)-3]などの理論的な進展が新たに得られた。さらに、音源方向の事前情報を組み入れる場合にいくつかのパラメータを定めることが必要となるが、これらを観測信号に対して適応的に求める手法も提案した[(3)-1]。

(2) スパースモデリングに基づく音場分析法

音場分析では、対象領域内に音源を含まない場合は上記のような基底関数展開による方法を用いることが可能であるが、音源を含む場合には、音源分布がどのような関数となるか未知であるため、より難しい問題となる。これまで、音源が空間的に疎に分布するという仮定の下、スパースモデリングを用いた音場分解による推定法を提案してきた [Koyama+ JASA 2018, IEEE JSTSP 2019, Murata+ IEEE TSP 2018]。これらの手法は対象領域内に音源を含まない場合においても、高周波数帯域における高精度化を可能とする。しかしながら、従来の方法では対象領域を格子点として離散化することが必要となり、真の音源位置次第では誤差が大きくなるほか、反復法によるアルゴリズムを用いるために計算負荷も大きい。そこで、Reciprocity Gap Functional と呼ばれる逆問題的な方法論により、格子点への離散化を必要とせず、閉形式解として音場分解を可能とする手法を提案した [Takida+ Signal Process. 2021]。また、そのロバスト化のための手法として、複数の時間周波数における同時スパース性を利用した、annihilating filter による手法も提案した [Takida+ IEEE ICASSP 2019]。

(3) 多点振幅制御による音場合成法

従来の音場合成法では、ある所望の音場を設定し、それを合成するためのスピーカ駆動信号を求める。例えば、多点音圧制御法と呼ばれる手法では、対象領域を離散化することで得られる制御点において、合成される音圧が所望のものと一致するように最適化問題を解く。マルチゾーン音場合成という応用においては、対象領域を複数設定し、音が聞こえる領域と聞こえない領域を合成する、ということが行われる。このような応用においては、音が聞こえる領域において、どの方向から音が聞こえるかはあまり重要でない場合も多く、また、所望音場の設定によっては物理的に実現が困難な場合もあるため、所望の音圧、つまり振幅と位相の両方を合成することは必ずしも必要ではない。そこで、対象領域内に所望の振幅を合成し、位相は任意とする、多点振幅制御法を提案した[(3)-3]。多

点音圧制御の最適化問題は閉形式で解けるのに対し、多点振幅制御は目的関数が非線形かつ非凸となるため、非線形最適化アルゴリズムを用いる必要がある。しかしながら、目的関数が微分できないために一般的な勾配法を用いることはできない。そこで、Majorization-minimization 法あるいは交互方向乗数法に基づく多点振幅制御のアルゴリズムを導出した。本技術は、マルチゾーン音場合成としての応用を見据え、企業との共同研究に発展した。

研究テーマ B「音場分析・合成におけるセンサ・アクチュエータ配置最適化」

音場の分析・合成において、センサ・アクチュエータ(マイク・スピーカ)の配置はその精度に大きく影響を与える。従来は、ヘルムホルツ方程式の境界値積分表現に基づく手法が主流であったため、平面や球のような単純な形状の境界面上を等間隔に離散化することが行われてきた。分散配置アレイを用いる場合、アレイ素子を配置可能な領域の中で、どこに実際に配置するのが最適かわかれば、実用上も有用であると考えられる。特に、マイクやスピーカが設置可能な領域が限られている自動車の車室内などでは、その配置最適化の効果は大きいと期待できる。

これまでに、多点音圧制御による音場合成において、スピーカと制御点の配置を最適化する手法として、empirical interpolation method と呼ばれる関数補間の手法を応用した配置最適化法を提案している[Koyama+ ICASSP 2018]。この手法は、音場合成の問題を、スピーカの伝達関数を補間関数、制御点をサンプリング点とみなした場合に関数補間の問題と等価であることから、empirical interpolation method による補間関数・サンプリング点の最適化を適用するというものである。一方で、機械学習の文脈において、センサ配置最適化の手法が様々な提案されている。研究期間には、empirical interpolation method を含め、従来のセンサ配置最適化法を網羅的に比較評価する実験を行い、その結果を総説論文としてまとめた[(3)-4]。本成果はフランス パリ大学および CentralSupélec の研究者らとの共同研究による。

また、音場推定のためのセンサ配置最適化問題において、音場のカーネル補間法[Ueno+ IWAENC 2018]に基づく新しい目的関数設計による手法を提案した[Ariga+ IEEE ICASSP 2020, Nishida+ EUSIPCO 2020]。特に、[Nishida+ EUSIPCO 2020]においては、センサ配置候補領域と、推定対象領域が異なる状況においても配置最適化が可能な手法を提案した。音場合成においても、一般的な線形逆問題に基づく手法に適用可能かつ、所望音場の事前情報を組み入れることが可能な手法を提案している[Kimura+ WASPAA 2021]。

研究テーマ C「バーチャルリアリティへの応用」

実環境で収録した音場を、受聴者に対して再現する VR への応用に関して、ヘッドフォンでの受聴を想定したバイノーラル再現、スピーカアレイによる音場再現のそれぞれについて検討を行った。

実環境の音場をバイノーラル信号として再現する技術としては、Ambisonics と



図 3 マイクアレイ信号からのバイノーラル再現

呼ばれる方式が知られており、Youtube や Facebook などの 360° 動画において採用されているものの、その空間解像度は著しく低く、音質も大きく劣化する。高い空間解像度を実現するために球状マイクロフォンアレイを用いる場合が多いが、広い領域を收音するためには大規模なアレイが必要となってしまう。そこで、無限次元はどう関数展開に基づく音場推定手法を応用し、小規模なアレイを複数配置することによる、柔軟性とスケーラビリティを持つ收音システムによるバイノーラル再現手法を提案した[(1)-2]。また、推定した球波動関数展開係数からバイノーラル信号を合成する際に、頭部伝達関数の測定距離を考慮することによって高精度化する手法も合わせて提案している[Iijima+ IEEE MMSP 2020]。本手法はデモシステムの構築も行っており、IEEE WASPAA 2021 においてデモ発表を行った。

スピーカアレイを用いる音場再現については、重み付きモードマッチング法[Ueno+ IEEE/ACM Trans. ASLP 2019]と呼ばれる手法を提案している。研究期間では、無限次元波動関数展開に基づく音場推定をスピーカ伝達関数の展開係数推定に応用し、実環境における評価を行った[(3)-2]。実環境実験においては、32ch のスピーカアレイシステムを構築し、対象領域内を格子点状に離散化した点でのインパルス応答を測定することを行った。このインパルス応答データは様々な研究において有用であると考えられるため、データセットとしての公開も行った[Koyama+ WASPAA 2021]。

研究テーマ D「空間アクティブ騒音制御への応用」

スピーカ信号によって騒音を抑圧するアクティブ騒音制御では、誤差マイクによってモニタリングする騒音をキャンセルするように二次スピーカの駆動信号を計算する。従来の手法では、誤差マイク位置での音圧パワーのみをコスト関数とするため、騒音抑圧の効果は誤差マイク位置付近のみに限られてしまう。誤差マイク

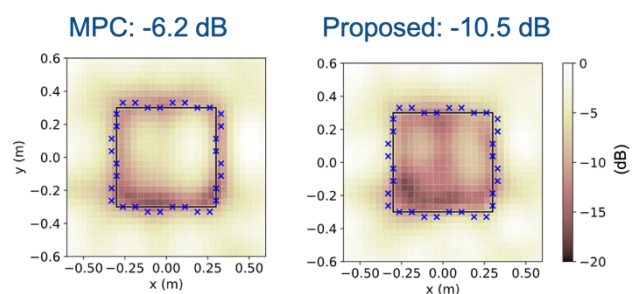


図 4 空間アクティブ騒音制御

位置以外も含む対象領域全体にわたって騒音を抑圧する空間アクティブ騒音制御を実現するには、誤差マイク信号から対象領域全体の音場を予測し、それをキャンセルするように二次スピーカ信号を得る必要がある。そこで、音場のカーネル補間法を応用した空間アクティブ騒音制御技術を提案した[(1)-1]。基本的な理論を発表した IEEE ICASSP[(3)-5]においては、Best Student Paper Award も受賞している。実時間で動作するシステムの構築には至っていないが、実環境で測定したインパルス応答を用いた実験的検証も行った。

3. 今後の展開

応用技術として最も実用化に近いと考えられるのは、マイクアレイ信号からのバイノーラル再現技術である。現在は Ambisonics 方式が主に採用されているが、6DoF での VR の重要性が増していることから、広い領域の音場を効率的に收音し、再現することが可能な提案技術の有用性は高いと考えられる。より高い周波数での音質の改善やリアルタイム実装など、

いくつかの課題はあるものの、基本的な技術としてはある程度完成している。ただし、より少数のマイクによって同様の技術を実現するには、データ駆動型の音場推定手法などを今後3～4年のうちに検討することが必要と考える。

また、スピーカアレイによる音場再現技術に関しても、残響がある程度抑えられた室内であればすでに高い効果が期待できる。室内環境の変動に対する追従などの課題は残るものの、アプリケーション次第ではすぐに適用可能と考える。また、多点振幅制御によるマルチゾーン音場合成も同様であるが、対象領域の外部への放射を抑圧する枠組みが必要になると考えられ、今後1年ほどで技術的な実証を行いたい。スピーカアレイを用いる技術は複数のスピーカを配置可能な応用先をうまく見つけることがまずは必要と考えている。

空間アクティブ騒音制御は、幅広く応用が期待できる技術であるが、現状はシミュレーションによる実験的評価にとどまっている。実環境での測定データを用いた場合でもある程度の効果が確認されているが、今後1年程度で実時間動作するシステムを構築し、技術を可視化していくことが必要と考える。その上で、課題点をさらに洗い出し、2～3年で技術をよりブラッシュアップすることで、広く社会実装につながることを期待している。

4. 自己評価

当初の研究目的については予想を上回る成果が得られたと考えている。基礎技術については論文などの多数の学術的成果が得られた他、その一部に関して IEEE ICASSP と呼ばれる信号処理分野のトップ国際会議における Tutorial や、IEEE Signal Processing Society 主催の Webinar を行うなど、学術的に広く注目を集めた。また、応用範囲を音場収音・再現以外にも拡大したことで、様々な課題に応用できる可能性のある技術を多数創出できた。COVID-19 の影響による社会情勢の変化も影響し、広く波及効果を持つ成果であると考え

5. 主な研究成果リスト

(1) 代表的な論文(原著論文)発表

研究期間累積件数:6件

1. Shoichi Koyama, Jesper Brunström, Hayato Ito, Natsuki Ueno, and Hiroshi Saruwatari, "Spatial Active Noise Control Based on Kernel Interpolation of Sound Field," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 29, pp. 3052–3063, 2021.

音場のカーネル補間法に基づく空間アクティブ騒音制御を提案した。従来のアクティブ騒音制御は、誤差マイク信号のパワーをコスト関数とするような適応フィルタによって実現されていたため、誤差マイク位置付近のみに騒音低減効果が限られていた。誤差マイク位置以外の領域も含む対象領域全体の騒音を抑圧するため、領域内における音圧の二乗積分値をコスト関数とした。誤差マイク信号から領域的音圧パワーを推定するため、音場のカーネル補間法を用い、Filtered-X Least Mean Square に基づく適応フィルタアルゴリズムを導出した。

2. Naoto Iijima, Shoichi Koyama, and Hiroshi Saruwatari, "Binaural Rendering from Microphone Array Signals of Arbitrary Geometry," Journal of Acoustical Society of America,

vol. 150, no. 4, pp. 2479–2491, 2021.

マイクアレイ信号からのバイノーラル再現技術を提案した。従来手法では球状マイクアレイを用いる場合が多いが、広い領域を收音するにはマイクアレイを大規模にする必要があるため、実現が困難であった。そこで、複数の小規模アレイ、例えばアンビソニックマイクを用いて、広い領域を効率的に收音できれば、柔軟性とスケーラビリティを持つ收音システムによる再現が可能となる。收音側では無限次元波動関数展開を用い、再現側では頭部伝達関数の測定距離を考慮した合成法により、先述の複数小規模マイクアレイを用いた高精度なバイノーラル再現を実現した。

3. Natsuki Ueno, Shoichi Koyama, and Hiroshi Saruwatari, “Directionally Weighted Wave Field Estimation Exploiting Prior Information on Source Direction,” IEEE Transactions on Signal Processing, vol. 69, pp. 2383–2395, 2021.

波動場の推定において、無限次元波動関数展開に基づく手法に対し、音源方向の事前情報を組み入れることで高精度化を行う手法を提案した。波動場のノルムが事前情報を反映する重み関数を使って定義され、重み関数としては von-Mises Fisher 分布を用いた。閉形式として解が得られる他、線形推定器として実現可能であるため、応用上も極めて重要な基礎理論となる。実験では従来法と比較した場合の優位性が示された他、事前情報に誤差を含む場合でも頑健に推定できることを示した。

(2) 特許出願

特になし

(3) その他の成果(主要な学会発表、受賞、著作物、プレスリリース等)

1. Ryosuke Horiuchi, Shoichi Koyama, Juliano G. C. Ribeiro, Natsuki Ueno, and Hiroshi Saruwatari, “Kernel Learning for Sound Field Estimation With L1 and L2 Regularizations,” in Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2021.
2. Shoichi Koyama, Keisuke Kimura, and Natsuki Ueno, “Sound Field Reproduction With Weighted Mode Matching and Infinite-Dimensional Harmonic Analysis: An Experimental Evaluation,” in Proc. International Conference on Immersive and 3D Audio (I3DA), 2021. (招待あり)
3. Shoichi Koyama, Takashi Amakasu, Natsuki Ueno, and Hiroshi Saruwatari, “Amplitude Matching: Majorization Minimization Algorithm for Sound Field Control Only With Amplitude Constraint,” in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2021.
4. Shoichi Koyama, Gilles Chardon, and Laurent Daudet, “Optimizing Source and Sensor Placement for Sound Field Control: An Overview,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 686–713, 2021. (総説論文)
5. Hayato Ito, Shoichi Koyama, Natsuki Ueno, and Hiroshi Saruwatari, “Feedforward Spatial Active Noise Control Based on Kernel Interpolation of Sound Field,” in Proc.

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP),
2019. (Best Student Paper Award, IEEE SPS Japan Student Conference Paper Award)