

令和元年6月5日現在

機関番号：12601

研究種目：若手研究(A)

研究期間：2015～2018

課題番号：15H05312

研究課題名（和文）高精細音響システムのための超解像型音場収録・再現の研究

研究課題名（英文）A study on super-resolution sound field recording and reproduction for high-fidelity audio systems

研究代表者

小山 翔一（Koyama, Shoichi）

東京大学・大学院情報理工学系研究科・講師

研究者番号：80734459

交付決定額（研究期間全体）：（直接経費） 18,600,000円

研究成果の概要（和文）：本研究では、高精細な音響システムのための音空間の再構成を実現する、音場収録・再現の超解像化に関して研究を行なった。具体的には以下のような成果が得られた。(1) マイクロフォンアレイを用いた音場の再構成において、対象領域が音源を含まない場合に関して、無限次元調和解析に基づく推定法を提案した。また、音源を含む場合に関して、スパース音場分解法を提案し、実験的な検証によってその有効性を確認した。(2) 音源位置あるいは受聴エリアの事前情報を用いた音場収録・再現アルゴリズムを考案した。(3) 音場制御におけるスピーカ・マイクロフォン配置の最適化法について理論構築を行なった。

研究成果の学術的意義や社会的意義

本研究は、音空間を計測あるいは制御する上で、少数のマイクロフォンやスピーカのみを用いて高精度化を実現する信号処理手法について、理論構築と実験的な検証を行なったものである。これは臨場感の高い音響再生システムや、音の空間的な制御などに応用が可能な技術である。従来の方法では、再現可能な周波数の上限値や領域のサイズが、マイクロフォン・スピーカ素子数によって決まるため、システムが大規模化してしまうという問題があった。本研究により、このような素子数の削減が可能となり、より簡易的なシステムによって同様の効果が得られる可能性が示唆された。

研究成果の概要（英文）：In this study, we addressed super-resolution sound field recording and reproduction for high-fidelity audio systems. Our achievements are summarized as follows: (1) For sound field reconstruction using a microphone array, we proposed a method based on harmonic analysis of infinite orders for estimating a sound field inside the source-free region. For estimating a sound field including sources, a sparse sound field decomposition method is proposed, and its efficacy is validated by experiments. (2) We proposed a sound field recording and reproduction method using the prior information of source locations and listening area. (3) A theory on the optimal loudspeaker and microphone placement method for sound field control is investigated.

研究分野：音響信号処理

キーワード：音場制御 音場計測 パーチャルリアリティ 圧縮センシング 機械学習

様式 C-19、F-19-1、Z-19、CK-19 (共通)

1. 研究開始当初の背景

音のバーチャルリアリティを実現する空間音響技術は、従来はステレオフォニック・サラウンド方式によるものや、バイノーラル合成によるものが主流であった。これらの従来の方法は、聴覚の特性に基づくものが多いことから、適切に受聴できる領域が狭いことや、音の空間的な解像度が十分に得られないという問題があった。一方、音空間の忠実な再現を目的とする音場再現技術では、多数のマイクロフォン・スピーカによる音波面の再構成を可能とする。このような技術は、近年の音響デバイスの小型化・低コスト化、A/D・D/A変換装置の多チャンネル化、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り、最近では国内外で盛んに研究が進められている。

2. 研究の目的

本研究の目的は、高精細音場収録・再生システムを実現するため、超解像型音場収録・再現技術を確立することである。音場収録・再現技術とは、マイクロフォン・スピーカアレイを用いて、音空間を物理的に忠実に再構成することを目指すものである。既存の技術では、アレイ素子の数や配置間隔によって、再現可能な周波数の上限値や領域のサイズが制限されてしまうため、現実的な規模のシステムでは、再現音場の音質や音像の定位感が劣化してしまうという問題があった。本研究では、収録場や再現音場の性質を陽に取り入れた信号モデルと、機械学習理論による最適化に基づく新たな信号処理のアプローチにより、聴感上の品質を保つために必要なアレイ素子数がより少ない高精細音響収録・再生システムの実現を目指す。研究開始当初の具体的な目的は、以下の3つであった。(1) 超解像型信号変換に適した音場の生成モデルと信号分離手法の構築、(2) 複数種類のセンサを協調することによる超解像型信号変換手法の構築、(3) 統合アルゴリズムによる信号変換手法の構築及び聴感上の品質との関係性の解明。

3. 研究の方法

本研究では、以下の4つの内容を進めてきた。

(1) マイクロフォンアレイ信号のスパース信号モデル及び信号分解アルゴリズムの検討

より少ないマイクロフォンを用いた音場収録を実現するには、音場のスパース表現が有効と考えられ、これによって音場収録・再現の超解像化が可能となると期待できる。このように、マイクロフォンアレイでの収録信号を、スパースなエネルギー分布を持つような基底関数系に展開するためには、音場の生成モデルや物理的な性質に基づいた基底関数系の設計が重要となる。また、このような信号モデルへの分解を可能にする最適化アルゴリズムについても、同時に検討すべき課題である。

(2) 音源位置・受聴者位置の事前情報を利用した超解像型信号変換に関する検討

実際のシステムでは、マイクロフォン以外のセンサを組み合わせることで、収録側の話者や演奏者などの音源位置や、再現側の受聴者位置をおおまかに取得できると考えられる。このようなセンサ情報を信号変換の際の事前情報として用いることで、マイクロフォン数やスピーカ数が少ない場合においても、高い精度での収録・再現が可能となると期待できる。

(3) 統合アルゴリズムによる信号変換手法と様々なアレイ形状への適用の検討

上記2つのアプローチを統合することで、マイクロフォンやスピーカの素子数が少数の場合でも、高い精度で音場収録・再現を実現可能な、超解像型信号変換手法を構築する。まずは平面または直線状アレイ形状への適用を想定した検討を行い、その後、円筒、球、円状アレイなど、様々なアレイ形状への適用を検討する。

(4) 超解像型信号変換と聴感上の印象との関係性に関する実験的検証

以上のアルゴリズムを適用することで、実環境で動作する超解像型音場収録・再生システムを構築する。システムの客観的・主観的評価により、音場収録・再現システムに要求される品質に対して必要な最低限のアレイ素子数の指標を見出す。

4. 研究成果

(1) 音場再構成のための無限次元調和解析およびスパース音場分解

複数のマイクロフォンから対象領域内の音場を推定する、音場再構成の問題は、対象領域が音源を含むか否かでアプローチが大きく異なる。音源を含まない領域の再構成においては、従来は球面調和関数などの基底関数展開に基づく方法が提案されてきた。これは研究開始当初想定していなかった成果であるが、無限次元の球面調和関数を用いた推定方法を新たに発見し、本手法が従来法に比べて高精度かつマイクロフォン素子配置の自由度が高いことをシミュレーションにより実証した [Ueno et al. IEEE SPL 2018]。さらに、この手法が球ベッセル関数をカーネル関数とするカーネルリッジ回帰に等しいことを理論的に示した [Ueno et al. IWAENC 2018]。一方、対象領域が音源を含む場合は、その再構成は ill-posed な問題となり、音源の分布に対するなんらかの仮定が必要となる。音源の空間的なスパース性を仮定したスパース音場分解では、このような音場再構成を可能とするだけでなく、少ないマイクロフォン素子での高精度な再構成、すなわち超解像化を実現する。このスパース音場分解の理論構築、および音場収録・再現への適用を行なったほか [Koyama et al. JASA 2018]、空間的なスパース性に加えて音源信号の時間周波数領域のスパース性も考慮に入れたスパース表現の新たなアルゴリズム

を提案した [Murata et al. IEEE TSP 2018]。本アルゴリズムは、多次元信号の各次元におけるスパース性を誘導するノルムを定義し、それを正則化項とするような非凸な最適化問題を Majorization-minimization 法と呼ばれる手法に基づいて解くものであり、一般的なスパース分解手法としても新規性が高いものである。さらに、残響環境下においても頑健なスパース音場分解を実現するため、残響音場のモデルを陽に取り入れた手法についても提案した [Koyama and Daudet IEEE JSTSP 2019]。スパース音場分解については実環境で収録した信号に対する評価を行い、これをシミュレーションによってスピーカアレイを用いた再現に適用した結果を図1に示す。従来法ではマイクロフォン数が少ないことによりエイリアシングの誤差が顕著に現れているが、提案法では単一音源からの所望音場を再現できていることがわかる。再現側を実システムで実装し、評価するには至らなかったが、理論的側面では当初の想定を大きく上回る成果が得られたと考えている。

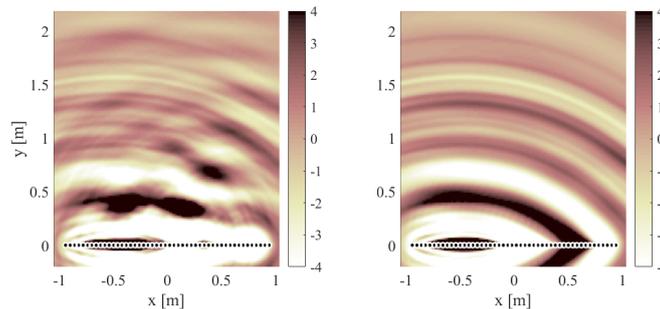


図 1 左：従来法による再現結果，右：スパース音場分解による再現結果。

(2) 音源位置・受聴エリア事前情報を利用した音場收音・再現

従来の音場收音・再現では、マイクロフォンの收音信号とアレイ素子位置のみから再現のためのスピーカ駆動信号を得るものであった。しかしながら、実際のシステムにおいては、おおまかな音源位置や受聴エリアを与えることも可能であり、これを用いることで收音・再現の高精度化が実現できると考えられる。まず、音源位置の事前情報を用いる方法については、与えられた音源位置からの球面波伝播を仮定した場合のスピーカ駆動信号の予測振幅分布を計算し、これを用いた事後確率最大化に基づく駆動信号の推定による手法を提案した [Koyama et al. IEEE JSTSP 2015]。さらに、受聴エリアを事前情報として与えた場合についても、受聴者位置を確率分布として与えることによる期待値最大化に基づく、重み付きモードマッチング法を提案した [Ueno et al. IEEE ICASSP 2017]。本手法はスパース音場分解法とも組み合わせることが可能であるが、統合アルゴリズムとしての検討は行わず、重み付きモードマッチング法が任意のアレイ形状、指向性のスピーカに対して適用可能であることを明らかにし、その実験的な評価をシミュレーションによって行なった。

(3) 音場收音・再現のためのスピーカ・マイクロフォン配置の最適化

音場再現において、特に対象領域が任意形状の場合や残響環境の場合、どのようなスピーカ・マイクロフォン配置が最適かは明らかにされていない。従来の方法では、アレイ形状が直線や円形のような単純な形状を想定し、残響のない自由空間を仮定していたため、対象領域を等間隔にサンプリングすることである程度高い性能が得られていた。これは当初想定していなかった研究内容であるが、任意形状の対象領域の音場を制御する場合に、スピーカ・マイクロフォン配置の最適化を行う方法について新たな発見をし、検討を行なった [Koyama et al. ICASSP 2018]。本手法は音場制御を所望の音場を伝達関数の線形結合によって補間・近似する問題とみなし、Empirical Interpolation Method と呼ばれる関数補間の方法を導入した。シミュレーション実験による評価では従来の方法に比べて極めて高い性能を示した。今後より詳細な実験的検証が必要になると考えている。

5. 主な発表論文等

[雑誌論文] (計6件)

1. S. Koyama and L. Daudet, "Sparse representation of a spatial sound field in a reverberant environment," *IEEE journal of selected topics in signal processing*, vol. 13, no. 1, pp. 172–184, 2019, DOI: 10.1109/JSTSP.2019.2901127.
2. N. Murata, S. Koyama, N. Takamune, and H. Saruwatari, "Sparse representation using multidimensional mixed-norm penalty with application to sound field decomposition," *IEEE transactions on signal processing*, vol. 66, no. 12, pp. 3327–3338, 2018, DOI: 10.1109/TSP.2018.2830318.
3. S. Koyama, N. Murata, and H. Saruwatari, "Sparse sound field decomposition for super-resolution in recording and reproduction," *Journal of the acoustical society of America*, vol. 143, no. 6, pp. 3780–3895, 2018, DOI: 10.1121/1.5042215.
4. N. Ueno, S. Koyama, and H. Saruwatari, "Sound field recording using distributed microphones based on harmonic analysis of infinite order," *IEEE signal processing letters*, vol. 25, no. 1, pp. 135–139, 2018,

DOI: 10.1109/LSP.2017.2775242.

5. S. Koyama, K. Furuya, K. Wakayama, S. Shimauchi, and H. Saruwatari, "Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle," *Journal of the acoustical society of America*, vol. 139, no. 3, pp. 1024-1036, 2016, DOI: 10.1121/1.4942590.
6. S. Koyama, K. Furuya, Y. Haneda, and H. Saruwatari, "Source-location-informed sound field recording and reproduction," *IEEE journal of selected topics in signal processing*, vol. 9, no. 5, pp. 881-894, 2015, DOI: 10.1109/JSTSP.2015.2434319.

[学会発表] (計 17 件)

1. Y. Takida, S. Koyama, N. Ueno, and H. Saruwatari, "Robust gridless sound field decomposition based on structured reciprocity gap functional in spherical harmonic domain," in *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, Brighton, May 2019, pp. 581-585.
2. Y. Takida, S. Koyama, and H. Saruwatari, "Exterior and interior sound field separation using convex optimization: comparison of signal models," in *Proceedings of european signal processing conference (EUSIPCO)*, Rome, Sep. 2018, pp. 2567-2571.
3. N. Ueno, S. Koyama, and H. Saruwatari, "Kernel ridge regression with constraint of Helmholtz equation for sound field interpolation," in *Proceedings of international workshop on acoustic signal enhancement (IWAENC)*, Tokyo, Sep. 2018, pp. 436-440.
4. Y. Takida, S. Koyama, N. Ueno, and H. Saruwatari, "Gridless sound field decomposition based on reciprocity gap functional in spherical harmonic domain," in *Proceedings of IEEE sensor array and multichannel signal processing workshop (SAM)*, Sheffield, Jul. 2018, pp. 627-631.
5. N. Ueno, S. Koyama, and H. Saruwatari, "Sound field reproduction with exterior radiation cancellation using analytical weighting of harmonic coefficients," in *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, Calgary, Apr. 2018, pp. 466-470.
6. S. Koyama, G. Chardon, and L. Daudet, "Joint source and sensor placement for sound field control based on empirical interpolation method," in *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, Calgary, Apr. 2018, pp. 501-505.
7. S. Koyama and L. Daudet, "Comparison of reverberation models for sparse sound field decomposition," in *Proceedings of IEEE workshop on applications of signal processing to audio and acoustics (WASPAA)*, New Paltz, Oct. 2017, pp. 214-218.
8. S. Koyama, N. Murata, and H. Saruwatari, "Effect of multipole dictionary in sparse sound field decomposition for super-resolution in recording and reproduction," in *Proceedings of international congress on sound and vibration (ICSV)*, London, Jul. 2017.
9. N. Murata, S. Koyama, N. Takamune, and H. Saruwatari, "Spatio-temporal sparse sound field decomposition considering acoustic source signal characteristics," in *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, New Orleans, Mar. 2017, pp. 441-445.
10. N. Ueno, S. Koyama, and H. Saruwatari, "Listening-area-informed sound field reproduction based on circular harmonic expansion," in *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, New Orleans, Mar. 2017, pp. 111-115.
11. N. Ueno, S. Koyama, and H. Saruwatari, "Listening-area-informed sound field reproduction with gaussian prior based on circular harmonic expansion," in *Proceedings of hands-free speech communication and microphone arrays (HSCMA)*, San Francisco, Mar. 2017, pp. 196-200.
12. S. Koyama, N. Murata, and H. Saruwatari, "Super-resolution in sound field recording and reproduction based on sparse representation," in *5th joint meeting of the acoustical society of america and acoustical society of japan*, Honolulu, Nov. 2016.
13. S. Koyama, "Source-location-informed sound field recording and reproduction: a generalization to arrays of arbitrary geometry," in *Proceedings of 2016 AES international conference on sound field control*, Guildford, Jul. 2016.
14. S. Koyama and H. Saruwatari, "Sound field decomposition in reverberant

- environment using sparse and low-rank signal models,” in *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, Shanghai, Mar. 2016, pp. 345-349.
15. N. Murata, S. Koyama, H. Kameoka, N. Takamune, and H. Saruwatari, “Sparse sound field decomposition with multichannel extension of complex NMF,” in *Proceedings of IEEE international conference on acoustics, speech, and signal processing (ICASSP)*, Shanghai, Mar. 2016, pp. 395-399.
 16. S. Koyama, A. Matsubayashi, N. Murata, and H. Saruwatari, “Sparse sound field decomposition using group sparse Bayesian learning,” in *Proceedings of asia-pacific signal and information processing association annual summit and conference (APSIPA ASC)*, Hong Kong, Dec. 2015, pp. 850-855.
 17. S. Koyama, K. Ito, and H. Saruwatari, “Source-location-informed sound field recording and reproduction with spherical arrays,” in *Proceedings of IEEE workshop on applications of signal processing to audio and acoustics (WASPAA)*, New Paltz, Oct. 2015.

[その他]

ホームページ等

<http://www.sh01.org>

6. 研究組織

(1) 研究分担者

なし

(2) 研究協力者

研究協力者氏名：ローラン ドデ

ローマ字氏名：Laurent Daudet

研究協力者氏名：ジル シャルドン

ローマ字氏名：Gilles Chardon

※科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。