

令和 5 年 5 月 27 日現在

機関番号：32689

研究種目：基盤研究(B)（一般）

研究期間：2019～2021

課題番号：19H04131

研究課題名（和文）音環境の認識と理解のための革新的マイクロホンアレー基盤技術の研究

研究課題名（英文）Research on Innovative Microphone Array Technology for Recognition and Understanding of Acoustic Environments

研究代表者

牧野 昭二（Makino, Shoji）

早稲田大学・理工学術院（情報生産システム研究科・センター）・特任教授

研究者番号：60396190

交付決定額（研究期間全体）：（直接経費） 13,200,000円

研究成果の概要（和文）：昨今では、政府主導により超スマート社会の実現に向けた様々な取り組みが行われている。マルチモーダルコミュニケーションの基盤技術の開発は、それらの問題解決の喫緊の課題である。応用範囲としてのスマートセキュリティ、高齢者見守り、ロボット聴覚、等への応用に関しては、いずれもマルチモーダルの根幹となる音環境を解析・理解するための統計数理、高速信号処理の基盤技術の確立が必要である。本研究は社会実装に向けた音環境を中心としたマルチモーダルコミュニケーションの基盤技術の開発である。具体的には、バーチャルマイクロホンや分散型マイクロホンアレーを用いた音響計測と深層学習による音環境の理解の基盤技術の研究を行った。

研究成果の学術的意義や社会的意義

本研究では、我々がこれまで提案してきたバーチャルマイクロホンという新概念に、音声や音響信号のパワフルな統計モデルや先進的な最適化手法を取り入れ、新しい分散型マイクロホンアレーシステムや信号処理アルゴリズムを開発したことに学術的意義がある。さらに、本研究では、マイクロホンアレー信号処理の積年の問題に立ち向かい、新しいロバストな分散型マイクロホンアレー信号処理アルゴリズムを考案し、包括的で安定な解法を開発したことに社会的意義がある。

研究成果の概要（英文）：In recent years, various efforts have been made under government leadership towards the realization of a highly smart society. Developing fundamental technologies for multimodal communication is also an urgent issue that needs to be addressed to solve these problems. Establishing fundamental technologies for statistical mathematics and fast signal processing to analyze and understand the sound environment, which is the core of multimodal communication, is necessary for applying it to smart security, elderly monitoring, robot audition, and other areas. This research focused on developing fundamental technologies for multimodal communication centered on the sound environment for social implementation. Specifically, we conducted research on the fundamental technologies of acoustic measurement using virtual microphones and distributed microphone arrays, as well as understanding the sound environment through deep learning.

研究分野：音響信号処理

キーワード：ブラインド音源分離 音響イベント検出 音情景解析

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

音響信号処理の分野では、長い間音声が最も重要な対象であったが、近年の音響信号処理の発展に伴い、その対象は音声から音楽、そしてあらゆる音へと拡大してきた。あらゆる音を対象とし、音から環境を解析・理解するための基盤技術は、スマートセキュリティ、高齢者見守り、ロボット聴覚、等、様々な新しい音情報サービスを可能にする技術と期待されている。また、近年では競争型ワークショップ[Detection and Classification of Acoustic Scenes and Events, IEEE AASP Challenges]や国際会議でのスペシャルセッション[Acoustic Event Detection and Classification]等が毎年開催されるなど、音響信号処理分野で、音響イベントの検出と分類は大きな注目を集めている研究トピックである。

2. 研究の目的

音響信号を用いて、音源位置を推定したり、音源の種類を特定したり、雑音の中から目的音源のみを分離・強調したりする技術は、スマートセキュリティ、高齢者見守り、ロボット聴覚、などの幅広い分野で必要とされている技術である。マイクロホンアレーを用いた音源分離・音源検出・音源定位は音響信号処理の中でも最も活発に研究が行われている分野の一つである。本研究では、我々を取り巻く音環境を解析・理解するための統計数理、高速信号処理の基盤技術の確立を目的として、バーチャルマイクロホン、分散型マイクロホンアレーなどの新たな発想に基づくマイクロホンアレー技術の確立および音響信号を対象とした深層学習による音環境の理解技術の確立を目指す。

3. 研究の方法

音環境の解析・理解能力は、機械学習に基づく認知過程に基盤を置き、目的とする音源を検出し、その信号成分を抽出し、周囲の音環境を学習して周囲の状況を把握する。本研究では、音環境からの情報をマイクロホンアレーを用いて取得し、音環境を解析・理解する新しい音響信号処理アルゴリズムを開発した。提案した信号処理アルゴリズムにより、大きな妨害音や残響の多い劣悪な音環境下でも、様々な種類の目的音源を検出し追跡できた。

(1)[検討項目1]

線形ビームフォーマに基づく音声強調は、強力だが劣決定条件下(マイク数 $M <$ 音源数 N)において性能が劣化する。近年普及しているスマートフォンなど、少ないマイクを持つ小型機器でも効果的な音声強調を達成する手法の開発が求められている。

音の伝播の物理的なモデルに基づいて観測信号を補間し、実際には存在しない、いわばバーチャルな観測信号を作り出して素子数を擬似的に増やすことにより、音源数に依存することなく高品質な出力を得るための統一的なアレー信号処理を検討した。擬似観測の振幅は非線形補間により推定した。擬似観測を用いた音声強調の劣決定拡張により、擬似観測の基本的な検証を行った。さらに、バーチャルマイクロホンのさらなる動作原理の解明と高性能化を図った。

(2)[検討項目2]

音環境からの情報を利用した多チャンネル信号処理アルゴリズムを開発した。既存のアルゴリズムをバーチャルマイクロホンに対応できるように一般化し、さらに強力な最適化規範を導入した。バーチャルマイクロホンにおけるサブアレーの同期手法を検討した。ブラインド音源分離/抽出アルゴリズムや多チャンネル残響除去アルゴリズムをバーチャルマイクロホンに対応できるように開発した。さらに、必要なマイクロホンを最小化して演算量を削減しながら、性能を最適化するためのマイクロホン選択手法も検討した。

(3)[検討項目3]

強調された音源信号から抽出した特徴量に基づき、音環境を解析・理解した。音源信号に関する先見知識を利用し、特徴量次元での分類法も利用した。分類精度を向上させるために、深層学習などの最新の音声認識技術を活用した。

4. 研究成果

本研究では、雑誌論文9件、国際会議発表33件、および、国内大会発表29件の研究成果を得た。主な研究成果を以下に示す。

(1)[検討項目1]

バーチャルマイクロホンの内挿と外挿を基に、周波数ごとの最適化を目的とした新たなバーチャルマイクロホンの提案を行った。一般的に、短い波長をもつ高周波の信号においては、短いマイクロホン間隔の方が空間的エイリアシングの影響を低減できるため有利である。反対に、長い波長をもつ低周波の信号においては、長いマイクロホン配置の方が位相差やゲイン差などの空間情報を十分に得ることが出来るため、最適な空間フィルタの構成において有利である。こ

のように、マイクロホンアレイ信号処理において、低周波帯域での性能と高周波帯域での性能にはトレードオフ関係が存在し、マイクロホン間隔を調整することによって制御することが出来る。マイクロホン間隔が波長の半分の長さであれば、空間的エイリアシングを起こさずに最大の位相差を得ることが出来、良好な指向特性を持った空間フィルタを構成することが出来る。しかし、実際のマイクロホンアレイはマイクロホンの位置が固定されているため、周波数ごとにその位置を変えることは現実的ではない。そのため、このトレードオフ関係への対応として、不等間隔マイクロホンアレイが用いられてきた。これは多数のマイクロホンを不等間隔に配置し、周波数帯域ごとに適切なマイクロホン間隔を持ったマイクロホンのペアを選択して信号処理を行う手法である。これにより、周波数帯域ごとに最適化された

マイクロホンアレイ信号処理が可能であるが、必然的に多くのマイクロホンが必要となる。したがって、広く用いられているスマートフォンや IC レコーダーに搭載するのは難しい。

一方、バーチャルマイクロホン技術では、信号処理が各周波数ビンで独立して実行されるため、周波数ごとにバーチャルマイクロホンの位置を変更することができる。これらより本研究では、バーチャルマイクロホンの内挿と外挿に基づいて、周波数ごとに波長に比例した位置にバーチャルマイクロホンを配置するバーチャルマイクロホンの波長比例配置について提案した。また、これを劣決定条件下における MaxSNR ビームフォーマを用いた音声強調に適用し、従来のバーチャルマイクロホン技術を適用するよりも、優れた音声強調性能が得られることを確認した。提案法の特徴として、適切な波長係数 k を設定することで、空間的エイリアシングを起こさずに鋭い指向特性を持った空間フィルタの構成に十分な位相差を全周波数で得ることが出来る事を確認した(図1)。

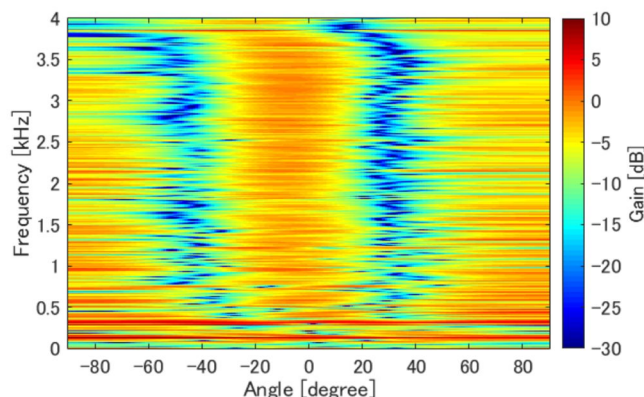


図1 波長比例バーチャルマイクロホンの指向特性。空間的エイリアシングを起こさずに鋭い指向特性を持った空間フィルタの構成に十分な位相差を全周波数で得ることが出来る。

これまでに我々は、深層学習に基づくバーチャルマイクロホン生成技術(neural network-based virtual microphone estimator, NN-VME)を提案した。NN-VMEでは、学習時には運用時よりも多くのマイクロホンが利用可能であることを仮定し、その学習基準として実際のマイクロホン信号と推定したバーチャルマイクロホン信号を比較する損失関数を採用した。この学習基準は後段の処理に依存していないため、推定したバーチャルマイクロホン信号を任意のアレイ信号処理に適用することが可能であると期待される。これに対して、後段で適用するアレイ信号処理があらかじめ定まっている場合、その出力に応じた損失関数を学習時に用いることで更に後段の各信号処理に適したバーチャルマイクロホン信号を推定可能になると期待される。

本研究では、後段の信号処理にビームフォーマを適用し、時間領域でバーチャルマイクロホン信号を推定する NN-VME の学習基準としてバーチャルマイクロホンの推定信号に対する損失関数とビームフォーマの出力信号に対する損失関数を組み合わせたマルチタスク損失関数を提案した(図2)。また、それらの重みづけを変化させて、音声強調性能および音声認識性能が向上することを確認した。

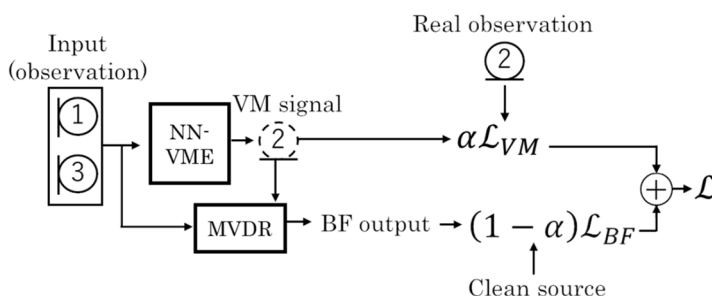


図2 マルチタスク学習深層学習に基づくバーチャルマイクロホン生成。

(2)[検討項目 2]

MVAE (Multichannel Variational Autoencoder) 法の計算コストの削減を目的とし、FastMVAE 法と呼ぶ高速アルゴリズムを提案した。FastMVAE 法は、クラス識別器付き VAE (Auxiliary Classifier VAE: ACVAE) を用いて音源スペクトログラムの生成モデルであるデコーダと共に、音源クラスの分布と潜在変数の事後分布を近似する識別器分布とエンコーダ分布を学習することで、学習で得られた識別器とエンコーダの順伝播により事後分布が最大となるようなデコーダ入力値を近似する手法である。しかし、FastMVAE 法では音源分離アルゴリズムを高速化できた一方で、未知話者や長い残響の場合など、テスト時において学習時と条件が一致しない場合に分離性能が低下する傾向があった。本研究では、さらにこの問題を解決するため、知識蒸留を用いた新たな ACVAE 音源モデルの学習法とモデル構造を提案した(図 3)。提案法で学習された音源モデルは従来よりコンパクトな構造かつ高い汎用性を持ち、未知話者でも高速かつ高精度な音源分離を実現できることを示した。

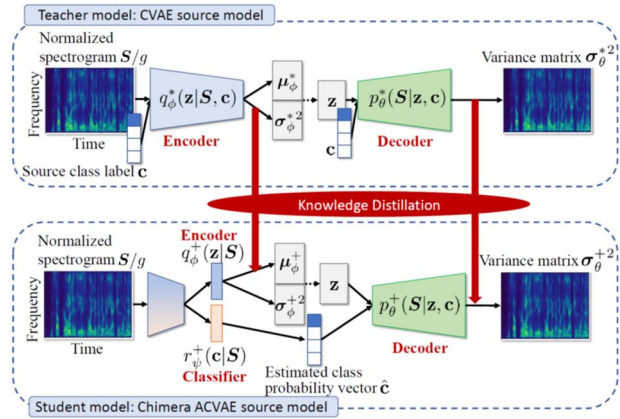


図 3 MVAE (Multichannel Variational Autoencoder) 法の計算コストを削減させた FastMVAE 法。誤差逆伝播法を不要とし、さらに知識蒸留を用いて計算コストを削減させた。

周波数領域 BSS では、観測信号が入力されてからその信号が短時間フーリエ変換 (Short-Time Fourier Transform: STFT) されるまでのアルゴリズム遅延が不可避である。音源分離の補聴器や車室内コミュニケーションシステムへの適用には低遅延化が要求されるため、STFT の窓長を短くする必要がある。しかし、IVA などの手法では、残響による分離性能の劣化を回避する為に STFT の窓長が十分に長い必要がある。一般に車室内の残響時間は短い、STFT の窓長を短くした場合、残響の影響は無視できなくなる。この問題に対して、重み付き予測誤差 (Weighted Prediction Error: WPE) 残響除去法を用いることで分離性能を妨げる残響を取り除くことができる。単純な方法としては、オンライン IVA とオンライン WPE を連結することで、分離性能を維持しつつ窓長を短くすることが期待できる。しかし、この連結した手法は音源分離と残響除去で各々最適化しているため、両方を考慮した同時最適化が保証されていない。バッチ処理による同時最適化手法は提案されているが、残響除去の推定を行う度に巨大な逆行列演算を行う為、計算コストが大きくなる。また、これらのバッチ手法をオンライン化する方法は未だ提案されていない。一方、バッチ処理での同時最適化における計算コストの問題を解決する手法が提案された。この手法は WPE で推定する残響除去フィルタを音源の数だけ増やして各々の音源の残響を取り除くように残響除去フィルタを求める。残響除去フィルタの数が増える代わりに、1つ1つの残響除去フィルタの更新の為に計算する逆行列が従来の WPE と同じ大きさになり、総合的に見ると従来より少ない計算コストで同時最適化を実現できる。本研究ではこの同時最適化手法をオンライン化した手法を提案した。提案法は従来のオンライン手法に簡単な改良を加えるだけで実装でき、実時間動作を十分に実現できる利点がある。雑音を含めた車室内環境下での提案法の有効性を音源分離実験により確認した(図 4)。

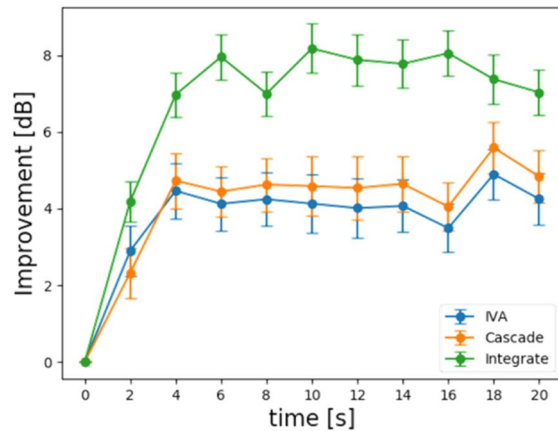


図 4 重み付き予測誤差 (Weighted Prediction Error: WPE) 残響除去法と独立ベクトル分析を同時最適化した提案法の音源分離性能。

BSS手法は音源間の独立性に基づいて分離を行うため話者情報や事前学習が不要であるという利点がある。一方で出力チャンネルの順序に任意性があるため、話者情報や空間情報を用いて分離後に目的音声を選択するための後処理が必要となる。このような出力順序問題と音源分離問題を同時最適化するためのアプローチとして、音源到来方向(Direction Of Arrival:DOA)やセンサ位置といった空間情報を与えて分離行列を求める幾何学的制約付き BSS (Geometrically Constrained BSS: GC-BSS)が研究されている。中でもベクトル座標降下法に基づく幾何学的制約付き AuxIVA (GC-AuxIVA with Vectorwise Coordinate Descent: GC-AuxIVA-VCD)は VCD のアイデアを取り入れて導出した更新アルゴリズムによって AuxIVA の利点を保持しつつ、出力順序問題を効率的に解決する手法である。さらにそのオンライン手法も提案されており、有効性が確認されている。しかし AuxIVA や GC-AuxIVA-VCD における更新則は、逆行列計算が含まれるため大きな計算コストを要し数値計算に不安定さを招くなどの問題があった。

このような問題を克服するため、Iterative Source Steering(ISS)と呼ばれる逆行列演算不要で計算コストが小さい新たな更新アルゴリズムを用いた Aux-IVA(AuxIVA-ISS)が近年提案され、IPに基づく AuxIVA と比較し実行時間を削減したことが示されている。また ISS の更新規則は混合行列の列ベクトルを更新することと同義であるため、移動音源と固定音源が混合された状況では、分離行列のうち移動音源に対応するベクトルのみを逐次更新することでさらに効率的に計算出来ることが示されている。本研究では計算量の削減のため ISS に基づく GC-AuxIVA のオフラインアルゴリズム(GC-AuxIVA-ISS)を提案し、短い演算時間で従来法の GC-AuxIVA-VCD と同等かそれ以上の性能を達成した。更に本研究では、GC-AuxIVA-ISS のオンラインアルゴリズムを提案した。また現実では移動音源の DOA が既知であると仮定することは強い制限となるため、MULTiple Signal Classification(MUSIC)による DOA 推定を取り入れた実験を行い、指向性音声強調における提案法の性能を評価し、有効性を確認した(図5)。

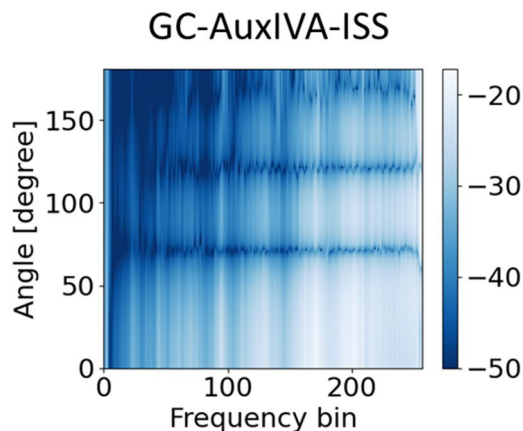


図5 幾何学的制約付き (Geometrically Constrained: GC)音源分離の指向特性。音源到来方向の指向特性を確実に制御することができる。

(3)[検討項目3]

Wave-U-Net と音響シーン識別器のエンドツーエンド学習による音響シーン識別手法を提案した。本来、Wave-U-Net の学習は、目的音と雑音をそれぞれ教師信号とする二乗誤差最小化規範で行われる。一方、提案法では、Wave-U-Net と識別器を連結し、通常は識別器の学習に用いるクロスエントロピー最小化規範でネットワーク全体をエンドツーエンドに学習する(図6)。これにより、波形レベルでの教師信号は不要となり、Wave-U-Net のネットワークは識別器がより識別し易い信号を出力するようになると思われる。家庭内の人の行動を識別する実験を行い、提案法の有効性を検証した。

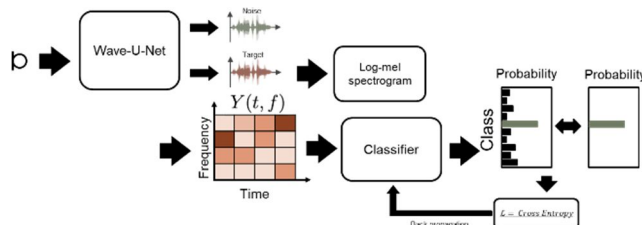


図6 Wave-U-Net と音響シーン識別器のエンドツーエンド学習による音響シーン識別手法。

GAN を用いた半教師あり学習を SED に適用した手法を提案し、その有効性を確認した。この手法では、主タスクとして SED、副タスクとして真偽判別を設定している。本研究ではこの手法の検出精度を更に改善するため、Generator によって弱ラベル付きデータを生成し、そのデータを用いて、真偽判別に加えて直接 SED の学習を行う手法を提案し、その有効性を検証した(図7)。

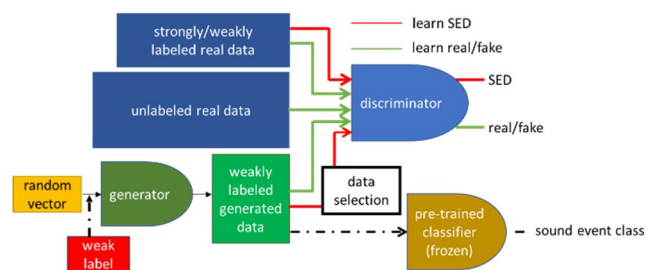


図7 GAN を用いた半教師あり学習を適用した SED 手法。

5. 主な発表論文等

〔雑誌論文〕 計9件（うち査読付論文 9件/うち国際共著 0件/うちオープンアクセス 9件）

1. 著者名 K. Yamaoka, N. Ono, and S. Makino	4. 巻 vol. 29
2. 論文標題 Time-frequency-bin-wise linear combination of beamformers for distortionless signal enhancemen	5. 発行年 2021年
3. 雑誌名 IEEE/ACM Trans. Audio, Speech and Language Processing	6. 最初と最後の頁 3461-3475
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TASLP.2021.3126950	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 N. Murashima, H. Kameoka, L. Li, S. Seki, and S. Makino	4. 巻 vol. 25
2. 論文標題 Single-channel multispeaker separation with variational autoencoder spectrogram model	5. 発行年 2021年
3. 雑誌名 Journal of Signal Processing	6. 最初と最後の頁 145-149
掲載論文のDOI（デジタルオブジェクト識別子） 10.2299/jsp.25.145	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 R. Takahashi, L. Li, S. Makino, and T. Yamada	4. 巻 VOL. 25
2. 論文標題 VMInNet: Interpolation of virtual microphones in optimal latent space explored by autoencode	5. 発行年 2021年
3. 雑誌名 Journal of Signal Processing	6. 最初と最後の頁 245-250
掲載論文のDOI（デジタルオブジェクト識別子） 10.2299/jsp.25.245	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Y. Kaneko, T. Yamada, and S. Makino,	4. 巻 VOL. 25
2. 論文標題 Monitoring of domestic activities using multiple beamformers and attention mechanism	5. 発行年 2021年
3. 雑誌名 Journal of Signal Processing	6. 最初と最後の頁 239-243
掲載論文のDOI（デジタルオブジェクト識別子） 10.2299/jsp.25.239	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 L. Li, H. Kameoka, and S. Makino	4. 巻 vol. 8
2. 論文標題 Majorization-minimization algorithm for discriminative non-negative matrix factorization	5. 発行年 2020年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 227399-227408
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2020.3045791	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 L. Li, H. Kameoka, S. Inoue, and S. Makino	4. 巻 vol. 8
2. 論文標題 FastMVAE: A fast optimization algorithm for the multichannel variational autoencoder method	5. 発行年 2020年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 228740-228753
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2020.3045704	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Yuki Kubo, Norihiro Takamune, Daichi Kitamura, and Hiroshi Saruwatari	4. 巻 VOL. 28
2. 論文標題 Blind Speech Extraction Based on Rank-Constrained Spatial Covariance Matrix Estimation With Multivariate Generalized Gaussian Distribution	5. 発行年 2020年
3. 雑誌名 IEEE/ACM TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING	6. 最初と最後の頁 1948 - 1963
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.3003165	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Yuki Mitsufuji, Norihiro Takamune, Shoichi Koyama, Hiroshi Saruwatari	4. 巻 Volume: 29
2. 論文標題 Multichannel blind source separation based on evanescent-region-aware non-negative tensor factorization in spherical harmonic domain	5. 発行年 2020年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech and Language Processing	6. 最初と最後の頁 607 - 617
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.3045528	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 H. Kameoka, L. Li, S. Inoue, and S. Makino	4. 巻 vol. 31, no. 9
2. 論文標題 Supervised determined source separation with multichannel variational autoencoder	5. 発行年 2019年
3. 雑誌名 Neural Computation	6. 最初と最後の頁 1891-1914
掲載論文のDOI (デジタルオブジェクト識別子) 10.1162/neco_a_01217	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計62件 (うち招待講演 8件 / うち国際学会 33件)

1. 発表者名 T. Yu, T. Ueda, and S. Makino
2. 発表標題 Blind Source Separation of Moving Sound Sources in Reverberant Indoor Environments, '' in Proc
3. 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP) (国際学会)
4. 発表年 2022年

1. 発表者名 K. Ouma, T. Yamada, and S. Makino
2. 発表標題 Semi-Supervised Learning Using Weakly Labeled Data Generated by GAN in Sound Event Detection, '' in Proc
3. 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP) (国際学会)
4. 発表年 2022年

1. 発表者名 S. Nakaoka, L. Li, S. Makino, and T. Yamada
2. 発表標題 Reducing algorithmic delay using low-overlap window for online Wave-U-Net
3. 学会等名 Invited in Proc. APSIPA (招待講演) (国際学会)
4. 発表年 2021年

1 . 発表者名 H. Segawa, L. Li, S. Makino, and T. Yamada
2 . 発表標題 Extension of virtual microphone technique to multiple real microphones and investigation of the impact of phase and amplitude interpolation on speech enhancement
3 . 学会等名 in Proc. APSIPA (国際学会)
4 . 発表年 2021年

1 . 発表者名 S. Misawa, N. Takamune, T. Nakamura, D. Kitamura, H. Saruwatari, M. Une, and S. Makino
2 . 発表標題 Speech enhancement by noise self-supervised rank-constrained spatial covariance matrix estimation via independent deeply learned matrix analysis
3 . 学会等名 in Proc. APSIPA (国際学会)
4 . 発表年 2021年

1 . 発表者名 J. Santoso, T. Yamada, S. Makino, K. Ishizuka, and T. Hiramura
2 . 発表標題 Speech emotion recognition based on attention weight correction using word-level confidence measure
3 . 学会等名 in Proc. INTERSPEECH (国際学会)
4 . 発表年 2021年

1 . 発表者名 T. Ueda, T. Nakatani, R. Ikeshita, K. Kinoshita, S. Araki, and S. Makino
2 . 発表標題 'Low latency online source separation and noise reduction based on joint optimization with dereverberation
3 . 学会等名 Invited in Proc. EUSIPCO (国際学会)
4 . 発表年 2021年

1 . 発表者名 S. Inoue, H. Kameoka, L. Li, and S. Makino
2 . 発表標題 SepNet: A deep separation matrix prediction network for multichannel audio source separation
3 . 学会等名 in Proc. ICASSP2021 (国際学会)
4 . 発表年 2021年

1 . 発表者名 T. Ueda, T. Nakatani, R. Ikeshita, K. Kinoshita, S. Araki, and S. Makino
2 . 発表標題 Low latency online blind source separation based on joint optimization with blind dereverberation
3 . 学会等名 in Proc. ICASSP2021 (招待講演) (国際学会)
4 . 発表年 2021年

1 . 発表者名 S. Nakaoka, L. Li, S. Inoue, and S. Makino
2 . 発表標題 Teacher-student learning for low-latency online speech enhancement using wave-U-net
3 . 学会等名 in Proc. ICASSP2021 (国際学会)
4 . 発表年 2021年

1 . 発表者名 L. Li, H. Kameoka, and S. Makino
2 . 発表標題 FastMVAE2: On improving and accelerating the fast variational autoencoder-based source separation algorithm for determined mixtures
3 . 学会等名 arXiv:2109.13496
4 . 発表年 2021年

1. 発表者名 J. Santoso, T. Yamada, K. Ishizuka, T. Hashimoto, and S. Makino
2. 発表標題 Neutral/Emotional Speech Classification Using Autoencoder and Output of Intermediate Layer in Emotion Recognizer
3. 学会等名 日本音響学会 2022年春 季研究発表会講演論文集
4. 発表年 2022年

1. 発表者名 山田友紀, 山田武志, 牧野昭二
2. 発表標題 Wave-U-Netと識別器のエンドツーエンド学習による音響シーン識別の検討
3. 学会等名 日本音響学会 2022年春季研究発表会講演論文集
4. 発表年 2022年

1. 発表者名 李莉, 亀岡弘和, 牧野昭二
2. 発表標題 ChimeraACVAEによる高速多チャンネル変分自己符号化器法
3. 学会等名 日本音響学会 2021年秋季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 中岡想太郎, 李莉, 牧野昭二, 山田武志
2. 発表標題 Low-overlap window を用いたオンラインWave-U-Net のアルゴリズム遅延の削減
3. 学会等名 日本音響学会 2021年秋季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 瀬川華子, 李莉, 牧野昭二, 山田武志
2. 発表標題 ヴァーチャルマイクロフォンの内挿における位相及び振幅補間の音声強調性能への影響の評価
3. 学会等名 日本音響学会 2021年秋季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 合馬一弥, 山田武志, 牧野昭二
2. 発表標題 音響イベント検出におけるGANを用いた弱ラベルデータ生成による半教師あり学習
3. 学会等名 日本音響学会 2021年秋季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 L. Li, H. Kameoka, S. Makino
2. 発表標題 Online directional speech enhancement using geometrically constrained independent vector analysis
3. 学会等名 IEEE International Workshop on Machine Learning for Signal Processing (MLSP2020) (国際学会)
4. 発表年 2020年

1. 発表者名 L. Li, K. Koishida, and S. Makino
2. 発表標題 Online directional speech enhancement using geometrically constrained independent vector analysis
3. 学会等名 Interspeech2020 (国際学会)
4. 発表年 2020年

1 . 発表者名 M. Une, Y. Kubo, N. Takamune, D. Kitamura, H. Saruwatari, and S. Makino
2 . 発表標題 Multichannel hearing-aid system based on basis-shared semi-supervised independent low-rank matrix analysis
3 . 学会等名 Forum Acusticum2020 (国際学会)
4 . 発表年 2020年

1 . 発表者名 K. Goto, L. Li, R. Takahashi, S. Makino, and T. Yamada
2 . 発表標題 Study on geometrically constrained IVA with auxiliary function approach and VCD for in-car communication
3 . 学会等名 Asia-Pacific Signal and Information Processing Association (APSIPA 2020) (招待講演) (国際学会)
4 . 発表年 2020年

1 . 発表者名 H. Segawa, R. Takahashi, R. Jinzai, S. Makino, and T. Yamada
2 . 発表標題 Applying virtual microphones to triangular microphone array in in-car communication
3 . 学会等名 Asia-Pacific Signal and Information Processing Association (APSIPA 2020) (国際学会)
4 . 発表年 2020年

1 . 発表者名 R. Takahashi, L. Li, S. Makino, and T. Yamada
2 . 発表標題 VMInNet: Interpolation of virtual microphones in optimal latent space explored by autoencoder
3 . 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2021) (国際学会)
4 . 発表年 2021年

1 . 発表者名 N. Murashima, H. Kameoka, L. Li, S. Seki, and S. Makino
2 . 発表標題 Single-channel multi-speaker separation via discriminative training of variational autoencoder spectrogram model
3 . 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2021) (国際学会)
4 . 発表年 2021年

1 . 発表者名 Y. Kaneko, T. Yamada, and S. Makino
2 . 発表標題 Monitoring of domestic activities using multiple beamformers and attention mechanism
3 . 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2021) (国際学会)
4 . 発表年 2021年

1 . 発表者名 Keigo Kamo, Yuki Kubo, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi and Kazunobu Kondo
2 . 発表標題 Joint-Diagonalizability-Constrained Multichannel Nonnegative Matrix Factorization Based on Multivariate Complex Student ' s t-distribution
3 . 学会等名 Asia-Pacific Signal and Information Processing Association (APSIPA 2020) (国際学会)
4 . 発表年 2020年

1 . 発表者名 Keigo Kamo, Yuki Kubo, Norihiro Takamune, Daichi Kitamura, Hiroshi Saruwatari, Yu Takahashi and Kazunobu Kondo
2 . 発表標題 Joint-Diagonalizability-Constrained Multichannel Nonnegative Matrix Factorization Based on Multivariate Complex Sub-Gaussian Distribution
3 . 学会等名 European Signal Processing Conference (EUSIPCO 2020) (国際学会)
4 . 発表年 2021年

1. 発表者名 井上翔太, 亀岡弘和, 李莉, 牧野昭二
2. 発表標題 SepNet: 高速多チャンネル音源分離のための分離行列予測ネットワーク
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 村島允也, 亀岡弘和, 李莉, 関翔悟, 牧野昭二
2. 発表標題 識別的変分自己符号化器学習による特定話者モノラル音声分離
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 上田哲也, 中谷智広, 池下林太郎, 木下慶介, 荒木章子, 牧野昭二
2. 発表標題 低遅延でオンライン動作する残響除去と音源分離の同時最適化
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 中岡想太郎, 李莉, 井上翔太, 牧野昭二
2. 発表標題 Teacher-Student 学習を用いたWave-U-Net による低遅延リアルタイム音声強調
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 瀬川華子, 高橋理希, 李莉, 陣在遼河, 牧野昭二, 山田武志
2. 発表標題 車室内の三角マイクロフォンアレイへのヴァーチャルマイクロフォン技術の適用
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 後藤加奈, 高橋理希, 李莉, 牧野昭二, 山田武志
2. 発表標題 補助関数法に基づく幾何学的制約付き独立ベクトル分析の車室内音声強調への適用
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 陳軼夫, 山田武志, 牧野昭二
2. 発表標題 音響イベント検出と位置推定における転移学習の効果の検証
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 樋口隼太, 李莉, 井上翔太, 牧野昭二, 山田武志
2. 発表標題 車室内環境でのWave-U-Netによる雑音除去の検討
3. 学会等名 電子情報通信学会 2021年 総合大会
4. 発表年 2021年

1. 発表者名 近藤祐斗, 久保優騎, 高宗典玄(東大), 北村大地(香川高専), 猿渡洋(東大)
2. 発表標題 ブラインド音声抽出のためのランク制約付き空間分散行列推定法における雑音欠落ランク空間基底推定
3. 学会等名 日本音響学会2020秋季研究発表会
4. 発表年 2020年

1. 発表者名 加茂佳吾, 久保優騎, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸
2. 発表標題 多変量複素Sub-Gauss分布に基づく同時対角化制約付き多チャネル非負値行列因子分解の様々な残響条件下における実験的評価
3. 学会等名 日本音響学会 2020年秋季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 加茂佳吾, 久保優騎, 高宗典玄, 北村大地, 猿渡洋, 高橋祐, 近藤多伸
2. 発表標題 多変量複素Sub-Gauss分布に基づく同時対角化制約付き多チャネル非負値行列因子分解におけるmajorization-equalizationアルゴリズムを用いた更新則
3. 学会等名 日本音響学会 2021年春季研究発表会講演論文集
4. 発表年 2021年

1. 発表者名 K. Yamaoka, N. Ono, S. Makino, and T. Yamada
2. 発表標題 Time-frequency-bin-wise switching of minimum variance distortionless response beamformer for underdetermined situations
3. 学会等名 International Conference on Acoustics, Speech, and Signal Processing (ICASSP2019) (招待講演) (国際学会)
4. 発表年 2019年

1 . 発表者名 L. Li, H. Kameoka, and S. Makino
2 . 発表標題 Fast MVAE: Joint separation and classification of mixed sources based on multichannel variational autoencoder with auxiliary classifier
3 . 学会等名 International Conference on Acoustics, Speech, and Signal Processing (ICASSP2019) (国際学会)
4 . 発表年 2019年

1 . 発表者名 S. Inoue, H. Kameoka, L. Li, S. Seki, and S. Makino
2 . 発表標題 Joint separation and dereverberation of reverberant mixtures with multichannel variational autoencoder
3 . 学会等名 International Conference on Acoustics, Speech, and Signal Processing (ICASSP2019) (国際学会)
4 . 発表年 2019年

1 . 発表者名 K. Yamaoka, L. Li, N. Ono, S. Makino, and T. Yamada
2 . 発表標題 CNN-based virtual microphone signal estimation for MPDR beamforming in underdetermined situations
3 . 学会等名 European Signal Processing Conference (EUSIPCO 2019) (招待講演) (国際学会)
4 . 発表年 2019年

1 . 発表者名 R. Jinzai, K. Yamaoka, M. Matsumoto, S. Makino, and T. Yamada
2 . 発表標題 Wavelength proportional arrangement of virtual microphones based on interpolation/extrapolation for underdetermined speech enhancement
3 . 学会等名 European Signal Processing Conference (EUSIPCO 2019) (招待講演) (国際学会)
4 . 発表年 2019年

1 . 発表者名 L. Li, K. Yamaoka, Y. Koshino, M. Matsumoto, and S. Makino
2 . 発表標題 Gated convolutional neural network-based voice activity detection under high-level noise environments
3 . 学会等名 International Congress on Acoustics (ICA2019) (国際学会)
4 . 発表年 2019年

1 . 発表者名 S. Inoue, H. Kameoka, L. Li, and S. Makino
2 . 発表標題 Joint separation, dereverberation and classification of multiple sources using multichannel variational autoencoder with auxiliary classifier
3 . 学会等名 International Congress on Acoustics (ICA2019) (招待講演) (国際学会)
4 . 発表年 2019年

1 . 発表者名 L. Li, T. Toda, K. Morikawa, K. Kobayashi, and S. Makino
2 . 発表標題 Improving singing aid system for laryngectomees with statistical voice conversion and VAE-SPACE
3 . 学会等名 Annual Conference of the International Society for Music Information Retrieval (ISMIR2019) (国際学会)
4 . 発表年 2019年

1 . 発表者名 M. Une, Y. Kubo, N. Takamune, D. Kitamura, H. Saruwatari, and S. Makino
2 . 発表標題 Evaluation of multichannel hearing aid system by rank-constrained spatial covariance matrix estimation
3 . 学会等名 Asia-Pacific Signal and Information Processing Association (APSIPA 2019) (招待講演) (国際学会)
4 . 発表年 2019年

1 . 発表者名 J. Santoso, T. Yamada, and S. Makino
2 . 発表標題 Classification of causes of speech recognition errors using attention-based bidirectional long short-term memory and modulation spectrum
3 . 学会等名 Asia-Pacific Signal and Information Processing Association (APSIPA 2019) (国際学会)
4 . 発表年 2019年

1 . 発表者名 T. Ueda, S. Inoue, S. Makino, M. Matsumoto, and T. Yamada
2 . 発表標題 Blind source separation with low latency for in-car communication
3 . 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2020) (国際学会)
4 . 発表年 2020年

1 . 発表者名 R. Takahashi, K. Yamaoka, L. Li, S. Makino, T. Yamada, and M. Matsumoto
2 . 発表標題 Underdetermined multichannel speech enhancement using time-frequency-bin-wise switching beamformer and gated CNN-based time-frequency mask for reverberant environments
3 . 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2020) (国際学会)
4 . 発表年 2020年

1 . 発表者名 Y. Kaneko, R. Kurosawa, T. Yamada, and Shoji Makino
2 . 発表標題 Spatial feature extraction based on convolutional neural network with multiple microphone inputs for monitoring of domestic activities
3 . 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2020) (国際学会)
4 . 発表年 2020年

1. 発表者名 宇根昌和, 久保優騎, 高宗典玄, 北村大地, 猿渡洋, 牧野昭二
2. 発表標題 ランク制約付き空間共分散モデル推定を用いた多チャンネル補聴器システムの評価
3. 学会等名 日本音響学会 2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 サントソジェニファー, 山田武志, 牧野昭二
2. 発表標題 BLSTMと変調スペクトルを用いた発話特徴識別の検討
3. 学会等名 日本音響学会 2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 舒禹清, 山田武志, 牧野昭二
2. 発表標題 BLSTMを用いた音声認識誤り区間推定の検討
3. 学会等名 日本音響学会 2019年秋季研究発表会講演論文集
4. 発表年 2019年

1. 発表者名 上田哲也, 井上翔太, 牧野昭二, 松本光雄, 山田武志
2. 発表標題 車室内コミュニケーション用低遅延音源分離手法の検討
3. 学会等名 日本音響学会 2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 高橋理希, 中谷智広, 落合翼, 木下慶介, 池下林太郎, Marc Delcroix, 荒木章子, 牧野昭二
2. 発表標題 DNNマスク推定に基づく畳み込みビームフォーマによる音源分離・残響除去・雑音除去の同時実現
3. 学会等名 日本音響学会 2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 宇根昌和, 久保優騎, 高宗典玄, 北村大地, 猿渡洋, 牧野昭二
2. 発表標題 基底共有型半教師あり独立低ランク行列分析に基づく多チャンネル補聴器システム
3. 学会等名 日本音響学会 2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 舒禹清, 山田武志, 牧野昭二
2. 発表標題 発話の時間変動に着目した音声認識誤り区間推定の検討
3. 学会等名 日本音響学会 2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 陳鞅夫, 山田武志, 牧野昭二
2. 発表標題 空間特徴と音響特徴を併用する音響イベント検出の検討
3. 学会等名 日本音響学会 2020年春季研究発表会講演論文集
4. 発表年 2020年

1. 発表者名 大野泰己, 山田武志, 牧野昭二
2. 発表標題 空間フィルタの自動推定による音響シーン識別の検討
3. 学会等名 電子情報通信学会 2020年 総合大会
4. 発表年 2020年

1. 発表者名 合馬一弥, 山田武志, 牧野昭二
2. 発表標題 Generative Adversarial Networks を用いた半教師あり学習の音響イベント検出への適用
3. 学会等名 電子情報通信学会 2020年 総合大会
4. 発表年 2020年

1. 発表者名 李莉, 亀岡弘和, 井上翔太, 牧野昭二
2. 発表標題 多チャンネル変分自己符号化器法による任意話者の音源分離
3. 学会等名 電子情報通信学会 2019年 応用音響研究会
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	猿渡 洋 (Saruwatari Hiroshi) (30324974)	東京大学・大学院情報理工学系研究科・教授 (12601)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分 担 者	山田 武志 (Yamada Takeshi) (20312829)	筑波大学・システム情報系・准教授 (12102)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関