

令和 4 年 6 月 7 日現在

機関番号：34310

研究種目：若手研究

研究期間：2019～2021

課題番号：19K20304

研究課題名（和文）あらゆる環境音を分析可能とする音響イベント検出手法の検討

研究課題名（英文）Sound event detection method capable of analyzing any environmental sound

研究代表者

井本 桂右（Imoto, Keisuke）

同志社大学・理工学部・准教授

研究者番号：90802116

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究課題は、音響処理技術において重要なトピックの1つである環境音分析において、少量の環境音データのみを用いた場合でも高い分析性能を達成する手法の実現を目指した。とりわけ本課題では、環境音の共起性や、収録環境による音響イベントの発生の遍在性を利用した音響イベント検出について検討を進めた。研究期間内に、音響イベントの共起性を従来の深層学習に導入する手法や、音響イベント検出とシーン分類のマルチタスク深層学習手法を提案した他、音響イベント間のデータに不均衡が生じている場合でも性能低下を引き起こさないモデル学習手法を提案し、高精度に音響イベント検出が可能となることを示した。

研究成果の学術的意義や社会的意義

人間の耳のように、様々な環境音の種類を聞き分ける技術が実現できれば、補聴器などの聴覚補助システムのみならず、公共スペースでの自動監視システム、高齢者や乳幼児の見守りシステム、自動運転の補助、環境の自動モニタリング、知的ロボットなど様々なサービスに広く貢献できる。このように、環境音分析は音響処理の中でも非常に重要な技術として位置づけられる。また、画像/動画の分析などの技術と組み合わせることで、人間の知覚を模した人工知能を実現することも可能となる。

研究成果の概要（英文）：This research project aims to develop a method to achieve reasonable performance in environmental sound analysis, which is one of the most important topics in acoustic processing, even when only a small amount of environmental sound data is available. In particular, we investigated a method for sound event detection (SED) based on the co-occurrence of environmental sounds and the omnipresence of sound events in an acoustic scene. During the research period, we have proposed a method considering the co-occurrence of acoustic events with deep learning methods, a multi-task learning method of SED and acoustic scene classification, and a model learning technique that does not cause performance degradation even when there is a data imbalance between sound event classes, showing that sound events can be detected with high accuracy.

研究分野：音響信号処理

キーワード：環境音分析 音響イベント検出 音響シーン分類

1. 研究開始当初の背景

スマートホンなどの収録機器の普及や機械学習理論や計算機環境の発展に伴い、音響分野で扱われる研究対象が、音声や音楽から、より一般的なあらゆる音事象（音響イベント）へと広がりつつある。人間の耳のように、様々な環境音の種類を聞き分ける技術（音響イベント検出）が実現できれば、高齢者や乳幼児の見守り、公共スペースでの監視システム、補聴器、実環境での人工知能（人と同じ聴覚機能を持つロボット）など様々なサービスに広く貢献できる。

音響イベントなどの環境音分析研究では機械学習、特に深層学習によるアプローチが主流であるが、検出したい環境音の種類に対して学習データを十分に集めることが難しいことが知られている。例えば、現状の音声認識技術では40種類程度の音声（音素）に対して数千時間～数万時間程の音声コーパスを利用して学習を行っている。一方、環境音は数百～数千種類存在するとされており、音声認識手法をそのまま適用するためには数万時間～数十万時間程度の環境音コーパスを用いる必要がある。しかしながら、環境音コーパスを作成する際のラベル付与作業は非常に時間を要することが課題となっている。

2. 研究の目的

本研究の目的は、数百～数千時間程度の少量の環境音コーパスで、実世界に存在する数百から数千の音響イベントを検出可能な手法の実現を目的とする。具体的には、環境音の共起性や、収録環境による音響イベントの発生の遍在性を利用した手法を検討し、他の音の認識タスク（音声認識や音響シーン分類）に匹敵する性能の実現を目指す。

3. 研究の方法

従来の深層学習に基づく音響イベント検出では、それぞれの音響イベントを独立に学習している。環境音には音を収録する場所や周囲の状況などシーンに依存して、発生する音響イベントの種類が変化する性質や、同時に発生しやすい音響イベントが存在する性質がある。本研究では、これらの性質を考慮した音響イベント検出について検討を進めた。具体的には図1のように、音響イベントの共起関係をグラフ表現して得られたグラフラプラシアンを深層学習アルゴリズムの罰則項として組み込み、共起性・スパース性を考慮する手法を提案し、評価を行った。また、音が収録された場所（シーン）と発生する音響イベントの種類が大きく関連しているという事実に基づき、シーンの予測誤差を深層学習アルゴリズムの罰則項として組み込んだ手法についても検討を進めた。なお、後者の手法はシーンと音響イベントのマルチタスク学習手法とも解釈可能である。

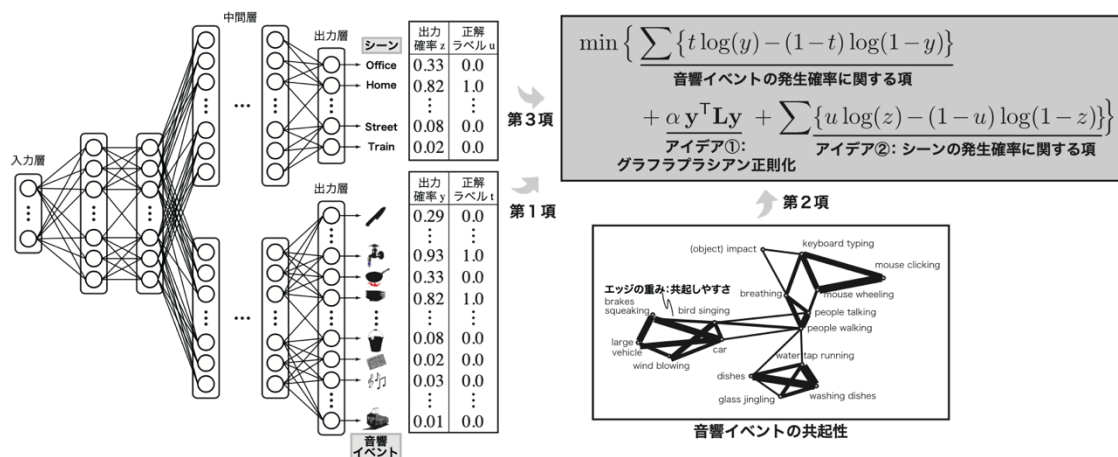


図1: 音響イベントの共起ならびにシーンと音響イベントの関係を考慮したイベント検出

4. 研究成果

平成31年度（令和1年度）は、音響イベントの共起性とスパース性を考慮した手法について検討を進め、音響イベントの共起性をグラフラプラシアン正則化として定式化し、従来の深層学習に導入する手法の検討を進めた。評価実験により、従来法よりも飛躍的に高い性能を発揮することを確認した。また、本検討を進める中で注目した音響信号のグラフ表現を応用し、ケプストラム特徴量をグラフ信号領域に拡張したグラフケプストラムという手法を考案した。

当初は令和2年度の計画としていた、収録環境（シーン）と発生する音響イベントの種類が関連している性質に着目した音響イベント検出手法についても平成31年度（令和1年度）に検討を開始し、音響イベント検出と音響シーン分類のマルチタスク深層学習を行う手法を提案した。

実験的な評価により、提案手法は少量のデータを用いた場合でも数十種類の音響イベントを高精度に検出できることを示した。

令和2年度は、シーンどうしの類似度を考慮するために、Teacher-Student 学習をマルチタスク深層学習に組み込んだ手法を提案し、前年に提案したマルチタスク深層学習に基づく手法よりもさらに高精度に環境音分析できることを実験により確認した。また、シーンと発生する音響イベントの種類の関連を考慮した性能評価手法も提案し、より現実の状況に即した評価実験が可能になることを確認した。

令和2年度の検討により、数十種類の音響イベントを高性能に分析する手法は実現できたものの、各音響イベントの継続長がデータ不均衡問題を引き起こし音響イベントモデル学習に悪影響を及ぼすことが原因となり、音響イベント検出のさらなる高性能化の障害になっていることが明らかになった。そこで令和3年度は音響イベントデータの不均衡に着目して検討を進めた。具体的には、音響イベント間のデータ不均衡が生じている場合でも性能低下を引き起こさないモデル学習手法を提案し、データ不均衡が生じている場合でも高精度に音響イベント検出が可能であることを評価実験により示した。

以上の研究活動を通して、学術論文5編ならびに国際会議発表15件、国内学会発表20件、加えて解説論文2編を発表した。得られた成果のさらなる発展については今後の研究活動を通じて実施し、最終的には高度な補聴器や人工聴覚システムに導入可能な環境音分析技術の実現を目指す。

## 5. 主な発表論文等

〔雑誌論文〕 計7件（うち査読付論文 5件／うち国際共著 0件／うちオープンアクセス 6件）

1. 著者名 Yuki Okamoto, Keisuke Imoto, Shinnosuke Takamichi, Ryosuke Yamanishi, Takahiro Fukumori, and Yoichi Yamashita	4. 巻 -
2. 論文標題 Onoma-to-wave: Environmental Sound Synthesis from Onomatopoeic Words	5. 発行年 2022年
3. 雑誌名 APSIPA Transactions on Signal and Information Processing	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	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 砺波 紀之, 井本 桂右, 岡本 悠希, 福森 隆寛, 山下 洋一	4. 巻 -
2. 論文標題 誤検出の深刻さを考慮した音響イベント検出のための評価指標	5. 発行年 2022年
3. 雑誌名 日本音響学会誌	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 井本 桂右, 川口 洋平	4. 巻 15
2. 論文標題 環境音分析・異常音検知の研究動向	5. 発行年 2022年
3. 雑誌名 電子情報通信学会 基礎・境界ソサイエティ Fundamentals review	6. 最初と最後の頁 268-280
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/essfr.15.4_268	査読の有無 無
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 Noriyuki TONAMI, Keisuke IMOTO, Ryosuke YAMANISHI, Yoichi YAMASHITA	4. 巻 E104.D
2. 論文標題 Joint Analysis of Sound Events and Acoustic Scenes Using Multitask Learning	5. 発行年 2021年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 294 ~ 301
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transinf.2020EDP7036	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 井本 桂右	4. 巻 75
2. 論文標題 環境音分析の研究動向	5. 発行年 2019年
3. 雑誌名 日本音響学会誌	6. 最初と最後の頁 512 ~ 518
掲載論文のDOI (デジタルオブジェクト識別子) 10.20697/jasj.75.9_512	査読の有無 無
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 IMOTO Keisuke	4. 巻 E103.D
2. 論文標題 Graph Cepstrum: Spatial Feature Extracted from Partially Connected Microphones	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 631 ~ 638
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2019EDP7162	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 Keisuke Imoto and Seisuke Kyochi	4. 巻 E103.D
2. 論文標題 Sound Event Detection Utilizing Graph Laplacian Regularization with Event Co-Occurrence	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 1971 ~ 1977
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2019EDP7323	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35件 (うち招待講演 2件 / うち国際学会 15件)

1. 発表者名 Noriyuki Tonami, Keisuke Imoto, Ryotaro Nagase, Yuki Okamoto, Takahiro Fukumori, and Yoichi Yamashita
2. 発表標題 Sound Event Detection Guided by Semantic Contexts of Scenes
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2022年

1. 発表者名 Keisuke Imoto
2. 発表標題 Fundamentals and Recent Advances in Environmental Sound Analysis
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 Kayo Nada, Keisuke Imoto, Reina Iwamae, and Nobutaka Ono
2. 発表標題 Multitask Learning of Acoustic Scenes and Events Using Dynamic Weight Adaptation Based on Multi-Focal Loss
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (国際学会)
4. 発表年 2021年

1. 発表者名 Keisuke Imoto
2. 発表標題 Acoustic Scene Classification Using Multichannel Observation with Partially Missing Channels
3. 学会等名 European Signal Processing Conference (EUSIPCO) (国際学会)
4. 発表年 2021年

1. 発表者名 Keisuke Imoto, Sakiko Mishima, Yumi Arai, and Reishi Kondo
2. 発表標題 Impact of Sound Duration and Inactive Frames on Sound Event Detection Performance
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2021年

1. 発表者名 Noriyuki Tonami, Keisuke Imoto, Yuki Okamoto, Takahiro Fukumori, and Yoichi Yamashita
2. 発表標題 Sound Event Detection Based on Curriculum Learning Considering Difficulty of Events
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2021年

1. 発表者名 井本 桂右, 賀谷 采珠, 椿 俊介
2. 発表標題 音響イベントの強ラベル付与におけるアノテーター間のばらつきの分析
3. 学会等名 日本音響学会 2022年春季研究発表会
4. 発表年 2021年

1. 発表者名 小松 由佳, 井本 桂右, 小松 達也
2. 発表標題 音響シーンとイベントが相互に及ぼす影響の調査
3. 学会等名 日本音響学会 2022年春季研究発表会
4. 発表年 2021年

1. 発表者名 椿 俊介, 宇都 瑛祐, 井本 桂右, 小野 順貴
2. 発表標題 弱ラベルを用いた音響シーンとイベントの同時分析
3. 学会等名 日本音響学会 2022年春季研究発表会
4. 発表年 2021年

1. 発表者名 砺波 紀之, 井本 桂右, 永瀬 亮太郎, 岡本 悠希, 福森 隆寛, 山下 洋一
2. 発表標題 事前定義されていないシーン情報を利用可能な音響イベント検出
3. 学会等名 日本音響学会 2022年春季研究発表会
4. 発表年 2021年

1. 発表者名 高橋 皓大, 井本 桂右, 土屋 隆生
2. 発表標題 グラフ深層学習を用いた音響イベントとシーンの同時分析
3. 学会等名 日本音響学会 2022年春季研究発表会
4. 発表年 2021年

1. 発表者名 岡本 悠希, 井本 桂右, 高道 慎之介, 福森 隆寛, 山下 洋一
2. 発表標題 環境音合成における主観評価手法の検討
3. 学会等名 日本音響学会 2022年春季研究発表会
4. 発表年 2021年

1. 発表者名 岡本 悠希, 井本 桂右, 高道 慎之介, 山西 良典, 福森 隆寛, 山下 洋一
2. 発表標題 Transformerを用いたオノマトベからの環境音合成
3. 学会等名 日本音響学会 2021年秋季研究発表会
4. 発表年 2021年



1. 発表者名 Taiga Kawamura, Ryoichi Miyazaki, Keisuke Imoto, and Nobutaka Ono
2. 発表標題 Experimental Investigation of Robustness of Spatial Cepstrum Features Under Various Recording Conditions
3. 学会等名 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (国際学会)
4. 発表年 2020年

1. 発表者名 Yuki Okamoto, Keisuke Imoto, Shinnosuke Takamichi, Ryosuke Yamanishi, Takahiro Fukumori, and Yoichi Yamashita
2. 発表標題 RWCP-SSD-Onomatopoeia: Onomatopoeic Word Dataset for Environmental Sound Synthesis
3. 学会等名 Detection and Classification of Acoustic Scenes and Events (DCASE) Workshop (国際学会)
4. 発表年 2020年

1. 発表者名 Noriyuki Tonami, Keisuke Imoto, Takahiro Fukumori, and Yoichi Yamashita
2. 発表標題 Evaluation Metric of Sound Event Detection Considering Severe Misdetections by Scenes
3. 学会等名 Detection and Classification of Acoustic Scenes and Events (DCASE) Workshop (国際学会)
4. 発表年 2020年

1. 発表者名 Keisuke Imoto, Noriyuki Tonami, Yuma Koizumi, Masahiro Yasuda, Ryosuke Yamanishi, and Yoichi Yamashita
2. 発表標題 Sound Event Detection by Multitask Learning of Sound Events and Scenes with Soft Scene Labels
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2020年

1. 発表者名 Tatsuya Komatsu, Keisuke Imoto, and Masahito Togami
2. 発表標題 Scene-dependent Acoustic Event Detection with Scene Conditioning and Fake-scene-condition Loss
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2020年

1. 発表者名 岩前 玲那, 白波瀬 壮, 高橋 皓大, 井本 桂右, 土屋 隆生
2. 発表標題 音響イベントとシーンのマルチタスク学習における評価関数の重みの自動調整
3. 学会等名 日本音響学会 2021年春季研究発表会
4. 発表年 2021年

1. 発表者名 井本 桂右, 美島 咲子, 荒井 友督, 近藤 玲史
2. 発表標題 音響イベント長とイベント非活性区間長の不均衡が検出性能に及ぼす影響
3. 学会等名 日本音響学会 2021年春季研究発表会
4. 発表年 2021年

1. 発表者名 岡本 悠希, 井本 桂右, 高道 慎之介, 山西 良典, 福森 隆寛, 山下 洋一
2. 発表標題 Onoma-to-wave: オノマトベからの環境音合成手法の提案
3. 学会等名 日本音響学会 2021年春季研究発表会
4. 発表年 2021年

1. 発表者名 河村 泰雅, 宮崎 亮一, 井本 桂右
2. 発表標題 実環境におけるマイクロホンの移動に対する空間ケプストラムの頑健性の調査
3. 学会等名 日本音響学会 2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 砺波 紀之, 井本 桂右, 福森 隆寛, 山下 洋一
2. 発表標題 音響シーンをを用いて検出誤りの深刻さを考慮したイベント検出の評価指標,
3. 学会等名 日本音響学会 2020年秋季研究発表会
4. 発表年 2020年

1. 発表者名 井本 桂右
2. 発表標題 環境音分析ことはじめ
3. 学会等名 日本音響学会 電気音響研究会/電子情報通信学会 応用音響研究会 (招待講演)
4. 発表年 2020年

1. 発表者名 Keisuke Imoto and Seisuke Kyochi
2. 発表標題 Sound Event Detection Using Graph Laplacian Regularization Based on Event Co-occurrence
3. 学会等名 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (国際学会)
4. 発表年 2019年

1. 発表者名 Noriyuki Tonami, Keisuke Imoto, Masahiro Niitsuma, Ryosuke Yamanishi, and Yoichi Yamashita
2. 発表標題 Joint Analysis of Acoustic Events and Scenes Based on Multitask Learning
3. 学会等名 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA) (国際学会)
4. 発表年 2019年

1. 発表者名 Keisuke Imoto and Nobutaka Ono
2. 発表標題 RU Multichannel Domestic Acoustic Scenes 2019: A Multichannel Dataset Recorded by Distributed Microphones with Various Properties
3. 学会等名 Detection and Classification of Acoustic Scenes and Events (DCASE) Workshop (国際学会)
4. 発表年 2019年

1. 発表者名 Yuki Okamoto, Keisuke Imoto, Naoki Tsukahara, Ken Nagata, Koh Sueda, Ryosuke Yamanishi, and Yoichi Yamashita
2. 発表標題 Crow Call Detection Using Gated Convolutional Recurrent Neural Network
3. 学会等名 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP) (国際学会)
4. 発表年 2020年

1. 発表者名 井本 桂右, 小野 順貴
2. 発表標題 RU Multichannel Domestic Acoustic Scenes 2019: 複数デバイスで構成された分散マイクアレイによる音響シーン分析のための環境音データセット
3. 学会等名 日本音響学会 2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 秋山 大知, 井本 桂右, 山西 良典, 山下 洋一
2. 発表標題 音の継続長の違いと検出難度を考慮した音響イベント検出
3. 学会等名 日本音響学会 2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 田中 良樹, 砺波 紀之, 井本 桂右, 山西 良典, 山下 洋一
2. 発表標題 環境音分析におけるマルチタスク学習の損失関数に対するアニーリングの検討
3. 学会等名 日本音響学会 2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 岡本 悠希, 井本 桂右, 高道 慎之介, 山西 良典, 山下 洋一
2. 発表標題 オノマトペを用いた環境音合成のためのデータセット構築とその分析
3. 学会等名 日本音響学会 2020年春季研究発表会
4. 発表年 2020年

1. 発表者名 砺波 紀之, 井本 桂右, 山西 良典, 山下 洋一
2. 発表標題 音響シーンの知識蒸留を用いた音響イベント検出
3. 学会等名 日本音響学会 2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 岡本 悠希, 柳生 拓巳, 井本 桂右, 小松 達也, 高道 慎之介, 山西 良典, 山下 洋一
2. 発表標題 多様な環境音の合成と変換のための基礎検討
3. 学会等名 日本音響学会 2019年秋季研究発表会
4. 発表年 2019年

1. 発表者名 Keisuke Imoto
2. 発表標題 Acoustic Event and Scene Analysis: Recent Advances and Challenges
3. 学会等名 日本音響学会 2019年春季研究発表会
4. 発表年 2019年

〔図書〕 計0件

〔出願〕 計1件

産業財産権の名称 音響モデル生成方法、音響分析方法、演算装置、及び、コンピュータプログラム	発明者 井本 桂右, 秋山 大知, 岡本 悠希, 山西 良典, 山下 洋一	権利者 同左
産業財産権の種類、番号 特許、特願2020-101291	出願年 2020年	国内・外国の別 国内

〔取得〕 計0件

〔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
---------------------------	-----------------------	----

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
---------	---------