

氏 名 井本 桂右

学位(専攻分野) 博士(情報学)

学位記番号 総研大甲第 1930 号

学位授与の日付 平成29年3月24日

学位授与の要件 複合科学研究科 情報学専攻
学位規則第6条第1項該当

学位論文題目 音情報の多様な観測を活用した音響シーン分析の研究

論文審査委員 主 査 准教授 小野 順貴
准教授 山岸 順一
准教授 相原 健郎
教授 佐藤 真一 国立情報学研究所
教授 篠田 浩一 東京工業大学

論文の要旨

Summary (Abstract) of doctoral thesis contents

本論文では、様々な条件で観測された音情報を活用して、音が観測された場所や周囲の状況、周囲にいるユーザの行動などの音響シーンを分類する問題を扱う。近年、スマートフォンやウェアラブルデバイス、IoT 機器、街頭センサなど、気軽に利用可能なマイクロホンが急速に増加しており、これまでは活用されていなかった音声以外の様々な音を積極的に活用することを目指し、音響イベント検出や音響シーン分析といった研究が盛んに進められている。一方で、気軽に様々な音が収録可能になるにつれ、音の収録環境も多様化している。本論文では、これまでに音響シーン分類問題の観測条件として扱われていない、逐次的な観測や間欠的な欠損を有する観測、時間同期誤差を持つ多チャンネル観測に対する音響シーン分類のための新たなアプローチを提案する。提案手法では、音のスパース性をモデルに導入でき、限られたデータからでも過学習することなく音響シーンをモデル化可能な音響トピックモデルに着目し、様々な観測条件に適用可能な新たな音響トピックモデルに基づくシーン分類を可能とする。

本論文ではまず、逐次的に新たな音データが得られる場合において、音響シーンをモデル化、分類する問題について検討を行う。逐次的に新たな音データが得られる場合、音データが得られる度に音響シーンモデルを再学習する必要があり計算コストが大きくなるという課題や、学習初期には限られた音データを用いて音響シーンをモデル化する必要があるため過学習に陥りやすいという課題がある。そこで本論文では、音響シーンに関する情報をパラメータとしてモデルに保存しておき、逐次的にこのパラメータを更新可能である点、また、音のスパース性をモデルに導入でき限られた音データから過学習することなく音響シーンをモデル化可能である点に着目し、音響トピックモデルに基づく逐次学習手法を提案する。提案手法では、従来の音響トピックモデルの学習手法を逐次学習可能な手法として拡張することに加え、計算コストとパラメータ推定精度それぞれに利点を持つ、崩壊型変分ベイズ法の近似手法と崩壊型ギブスサンプリングを組合せたハイブリッド型パラメータ推定手法も併せて導入する。実環境収録音を用いた評価実験により、提案手法で学習されたモデルパラメータを用いることによりバッチ型学習法で学習した場合と同等の音響シーン分類が実現できることを示す。

続いて、観測の一部が欠損している音データから音響シーンを分類し、さらに欠損した観測を同時に推定する問題について検討を行う。本論文では、欠損した観測を潜在的な確率変数とみなすことで、欠損していない音響ワードの観測と欠損した観測を同一の生成モデルで扱うことができる点に着目し、音響トピックモデルを拡張した新たな音響シーン分類手法を提案する。提案モデルでは、音の時間連続性に基づき観測の時間遷移を教師あり音響トピックモデルに組み込むことで、欠損した観測を推定しながら音響シーンの分類を行うことを可能とする。具体的には、1) 音響ワードの時間遷移関係を単純マルコフモデルによりモデル化し、教師あり音響トピックモデルに組み込んだ、音響ワード遷移型教師あり音響トピックモデルと、2) 音響トピックの時間遷移関係を隠れマルコフモデルによりモデル化し、教師あり音響トピックモデルに組み込んだ、音響トピック遷移型教師あり音響

(別紙様式 2)
(Separate Form 2)

トピックモデルを提案する。また、崩壊型ギブスサンプリングによる提案モデルのパラメータ推定方法も合わせて示す。実環境収録音を用いた性能評価実験により、提案手法では音響ワードの欠損率が 50%程度になった場合でも、欠損がない場合と同程度の音響シーン分類性能が実現可能であることを示す。

さらに、時間同期誤差を持つ分散マイクロホンアレイによる多チャンネル観測から音響シーンを分類する問題についても検討を行う。本論文では、音響トピックモデルにより空間情報を扱う方法について議論を行う。また、マイクロホン間の時間同期誤差に頑健な音響特徴抽出手法として、複数のマイクロホンで得られた振幅情報の対数に主成分分析を行うことで得られる、空間ケプストラムという新たな特徴抽出方法を提案する。空間ケプストラムは従来のマイクロホンアレイ処理と異なり、マイクロホンやスピーカの位置情報が不要であり、マイクロホン間の時間同期誤差にも頑健であるため、スマートホンなどが持つ複数のマイクロホンを用いた音響シーン分類に適していると言える。また、空間ケプストラムを一般化し、周波数情報と空間情報を同時に抽出可能な、一般化周波数-空間ケプストラムも併せて提案する。実環境収録音を用いた音響シーン分類実験により、空間ケプストラムおよび一般化周波数-空間ケプストラムは周波数特徴量同様に音響シーン分類に効果的であることを示す。また、提案手法を用いることで、チャンネル間に時間同期誤差がある場合でも頑健に音響シーン分類可能であることも併せて示す。

Summary of the results of the doctoral thesis screening

本博士論文は、多様な条件で観測された音情報を活用し、音が観測された場所や周囲の状況、ユーザの行動などの音響シーンを分類する問題について論じたものである。

第1章では序論として、音響シーン分類の意義と、スマートホン等録音機器の普及により、音の観測形態が多様化している研究背景について述べている。

第2章では、音響シーン分類という問題の定義と用語の定義が述べられ、続いて関連研究の概要と、音響シーン分類に対する有効なアプローチである音響トピックモデルという手法についての説明がなされている。

第3章では、音響信号の逐次的な観測を用いて音響シーンをモデル化し、分類する手法を提案している。まず従来法に対し、データが逐次的に得られる度にモデルを再学習するための計算コストや、学習初期に過学習しやすい点を問題として指摘している。次にこれを解決する手法として、オンライン型の音響トピックモデルを提案している。音響トピックモデルでは、過去の観測情報を潜在変数として保持でき、スパース性等の事前情報の導入により過学習を回避しやすい利点を論じた上で、変分ベイズ法とギブスサンプリングを組合せたハイブリッド型のパラメータ推定手法を提案し、実環境収録音を用いた評価実験により、提案手法がバッチ型学習法と同程度の精度でシーン分類が実現できることを示している。

第4章では、観測の一部が欠損している音情報に基づき音響シーンを分類する手法を提案している。具体的には、欠損した観測を潜在的な確率変数として音響トピックモデルに新たに導入し、音響ワード、ないし音響トピックの時間遷移をマルコフ的にモデル化することにより、欠損した観測を推定しながら音響シーンを推定する手法を提案している。実環境収録音を用いた評価実験により、提案手法では、観測の50%程度が欠損した場合でも、欠損がない場合と同程度の音響シーン分類性能が実現できうることを示している。

第5章では、分散したマイクロホンから得られる多チャンネル観測に基づき音響シーンを分類する問題を検討している。分散配置された別々のマイクロホンから得られる録音は同期誤差をもつことを想定し、同期誤差に頑健な振幅情報のみを用い、対数振幅を主成分分析することにより得られる空間ケプストラムという新たな特徴抽出方法を考案している。また、これを一般化し、スペクトル情報と空間情報を同時に抽出可能な一般化周波数・空間ケプストラムという特徴量も導出している。実環境収録音を用いた音響シーン分類実験により、空間ケプストラムおよび一般化周波数・空間ケプストラムが、従来のスペクトル特徴量同様に音響シーン分類に効果的であること、およびチャンネル間の同期誤差に頑健であることを示している。

第6章では、結論として本論文全体の内容がまとめられ、今後の課題についても述べられている。

審査会では出願者から論文全体について発表がなされ、その後の質疑においても適切な回答がなされた。本博士論文は、音響シーン分類において、逐次的、欠損、多チャンネルと

(別紙様式 3)

(Separate Form 3)

いった多様な観測形態を、音響トピックモデルにより統一的に扱うための新しいアルゴリズムと特徴抽出法を提案し、実収録音を用いた評価実験により提案法の有効性を示しており、当該分野に大きく貢献している。また、本論文の内容は当該分野のトップジャーナルに

Keisuke Imoto and Nobutaka Ono, "Spatial Cepstrum as a Spatial Feature Using Distributed Microphone Array for Acoustic Scene Analysis," IEEE/ACM Trans. Audio, Speech and Language Processing

という査読付きジャーナル論文として採録が確定している他、出願者が主著者である3編の査読付き国際会議論文がすでに掲載されている。

以上より本論文は、博士学位を与えるに十分な水準に達していると、審査委員全員一致で認められた。