

論文の内容の要旨

論文題目	深層学習に基づく音源情報推定のための確率論的目的関数の研究
学 位 申 請 者	小泉 悠馬

本研究は、マイクロホンで観測した音響信号から、源信号などの音に関する情報である「音源情報」を推定する研究である。特に本研究では、従来から推定対象とされてきた音源方向や位置などの顕在的な音源情報に加え、音源の種類や状態などの潜在的な音源情報を推定することを目指す。

潜在的な音源情報推定の題材として、源信号と雑音が重畳した観測信号から源信号を推定する「音源強調」と、観測信号に含まれる環境音の種類や状態を推定して周囲の危険を予測/察知する「異常音検知」に焦点を当てる。音源の種類や状態などの潜在的な音源情報を考慮しながら音源強調ができれば、大歓声に包まれたサッカースタジアムで、特定の選手の声やボールのキック音を推定でき、まるでサッカースタジアムに潜り込んだようなコンテンツ視聴の方法をユーザに提供可能になる。観測信号に含まれる環境音の種類や状態を推定する異常音検知が実現すれば、機器の動作音から、その機器の動作が正常か異常か（状態）を推定できるようになり、製造/保守業務の効率化ができる。

近年、深層学習を音源情報推定に適用する研究が盛んに行われている。この方法では、ニューラルネットワークを観測信号から所望の音源情報への非線形写像関数として用いる。本研究で取り上げる音源強調と異常音検知では、ニューラルネットワークはそれぞれ、観測信号から所望の音源の振幅スペクトルなどへの写像関数と、観測信号から状態や音響特徴量などへの写像関数として用いられる。そして、音源情報の推定精度を評価する「目的関数」の値を最大化/最小化するように、ニューラルネットワークは学習される。多くの深層学習において目的関数には、二乗誤差関数や交差エントロピー関数などの決定論的な目的関数が用いられる。

音源情報推定において目的関数の設計とは、所望の音源情報の性質や推定精度を定義することと等価である。音源情報の中には、決定論的な目的関数では音源情報の性質や推定精度を定義できないものや、もしくは定義することが妥当ではないものも存在する。例えば、人間の主観的な音質評価を最大化する源信号を推定したい場合、二乗誤差関数で推定精度を評価することは妥当ではない。また、異常音（ラベルデータ）が収集できない状態は、交差エントロピー関数を目的関数に利用できない。この問題を解決するためには、ニューラルネットワークの学習に用いる目的関数を進化させる必要がある。

本研究では、決定論的な関数で目的関数を設計できない音源情報を推定するた

めに、深層学習に基づく音源情報推定のための目的関数の研究を行う。所望の音源情報の特性や解きたい問題に応じて入出力値がとるべき値の確率分布や集合として定義することで、目的関数を、ニューラルネットワークの入出力が満たすべき統計的な性質として記述するという着想からこの問題に取り組む。

本論文の構成は以下のとおりである。

1章は序論であり、研究の背景と論文の構成をまとめ、2章では、音源強調と異常音検知に関する従来研究と、これらの分野における深層学習の適用法、およびその課題についてまとめる。

3章では、スポーツの競技音など、ラベルデータが十分に存在しない源信号を強調する課題に対し、所望の源信号を推定するために適切な音響特徴量を選択するための目的関数を提案する。定量評価試験では、従来の音響特徴量選択法と比べSDR (Signal-to-Distortion Ratio) が向上することを示し、また主観評価試験では、従来法と比べ源信号の明瞭性が向上することを示す。この成果により、これまで推定が困難とされていた、学習データが十分に得られないような源信号や、これまで源信号の推定対象とされてこず、適切な音響特徴量が未知な源信号も推定できるようになる。

4章では、音源強調の出力音の主観品質を向上させるために、主観評価値を最大化するための目的関数を提案する。ラベルデータを用意するのではなく、主観評価値と相関の高い音質評価値（聴感評点）を最大化する。主観評価試験では、従来の二乗誤差最小化に基づく目的関数を利用した音源強調よりも高い主観品質で音源強調できることを示す。この成果により、これまで音源強調の学習に利用できなかった聴感評点や人間の評価などの、より高次の評価尺度を目的関数として利用できるようになり、ニューラルネットワークを用いた音源強調の応用範囲を広げることができる。

5章では、異常音検知のための新たな目的関数を提案する。正常音が従う確率分布と統計的に差異がある音を異常音と定義することで異常音検知を仮説検定とみなし、仮説検定の最適化基準であるネイマン・ピアソンの補題から「ネイマン・ピアソン指標」を導出する。定量評価試験では、従来法と比べ安定して異常音が検知できることを示す。また実環境実験を行うことで3Dプリンタや送風ポンプ、ベアリングの傷などの異常音を検知できることを示す。この成果により、異常音データの集まらない状態識別問題を安定的に解くことが可能になり、銃声検知や未知話者検出などのセキュリティのための音源情報推定技術など、負例データの収集が困難な様々な音源情報推定へと応用ができる。

6章は結論であり、本研究の全体のまとめ、および研究の意義と貢献をまとめる。

論文審査の結果の要旨

学位申請者氏名 小泉 悠馬

審査委員主査 羽田 陽一

委員 橋本 直己

委員 柳井 啓司

委員 庄野 逸

委員 南 泰浩

小泉悠馬君提出の学位論文「深層学習に基づく音源情報推定のための確率論的
目的関数の研究」は、マイクロホンで観測した音響信号から音源情報を推定する
ために、深層学習に用いる目的関数を確率論的に拡張するという着想に基づき、
源信号の推定問題である「音源強調」と、環境音の状態の推定問題である「異常
音検知」のための新たな目的関数を提案したものであり、全6章から構成される。

音源情報推定において目的関数の設計とは、所望の音源情報の性質や推定精度
を定義することと等価である。音源情報の中には、決定論的な目的関数では音源
情報の性質や推定精度を定義できないものや、定義することが妥当ではないもの
も存在する。例えば、人間の主観的な音質評価を最大化する源信号を推定したい
場合、二乗誤差関数で推定精度を評価することは妥当ではない。また、異常音（ラ
ベルデータ）が収集できない状態は、交差エントロピー関数を目的関数に利用で
きない。この問題を解決するためには、ニューラルネットワークの学習に用いる
目的関数を進化させる必要がある。

本研究では、決定論的な関数で目的関数を設計できない音源情報を推定するた
めに、深層学習に基づく音源情報推定のための目的関数の研究を行っている。所
望の音源情報の特性や問題に応じて入出力値がとるべき値の確率分布や集合とし
て定義することで、ニューラルネットワークの入出力が満たすべき統計的な性質
として目的関数を記述するという着想からこの問題に取り組んでいる。

第1章は序論であり、研究の背景と論文の構成をまとめている。

第2章では、音源強調と異常音検知に関する従来研究と、これらの分野におけ
る深層学習の適用法、およびその課題についてまとめている。

第3章では、スポーツの競技音など、ラベルデータが十分に存在しない源信号
を強調する課題に対し、所望の源信号を推定するために適切な音響特徴量を選択
するための目的関数を提案している。定量評価試験では、従来の音響特徴量選択
法と比べSDR (Signal-to-Distortion Ratio) が向上することを示し、また主観評
価試験では、従来法と比べ源信号の明瞭性が向上することを示している。この成
果により、これまで推定が困難とされていた、学習データが十分に得られないよ
うな源信号や、これまで源信号の推定対象とされてこず、適切な音響特徴量が未

知な源信号も推定できるようになった。(なおこの成果は「高臨場感プレミアム空間演出ソリューション」の要素技術としてNTT研究所から発表されており、先進的な技術として注目を集めた。)

4章では、音源強調の出力音の主観品質を向上させるために、主観評価値を最大化するための目的関数を提案している。主観評価試験では、従来の二乗誤差最小化に基づく目的関数を利用した音源強調よりも高い主観品質で音源強調できることを示している。この成果により、これまで音源強調の学習に利用できなかった聴感評点や人間の評価などの、より高次の評価尺度を目的関数として利用できるようになり、ニューラルネットワークを用いた音源強調の応用範囲を広げることができた。

5章では、異常音検知のための新たな目的関数を提案している。正常音が従う確率分布と統計的に差異がある音を異常音と定義することで異常音検知を仮説検定とみなし、仮説検定の最適化基準であるネイマン・ピアソンの補題から「ネイマン・ピアソン指標」を導出している。定量評価試験では、従来法と比べ安定して異常音が検知できることを示している。また実環境実験では3Dプリンタや送風ポンプ、ベアリングの傷などの、実環境でも頑健に異常音を検知できることを示している。この成果により、異常音データの集まらない状態識別問題を安定的に解くことが可能になり、銃声検知や未知話者検出などのセキュリティのための音源情報推定技術など、負例データの収集が困難な様々な音源情報推定へと応用ができる。(なおこの成果は「異音検知ソリューション」の要素技術として実用化されている。)

6章は結論であり、本研究の全体のまとめ、および研究の意義と貢献をまとめている。

以上、小泉悠馬君提出の学位論文は、決定論的な目的関数では音源情報の性質や推定精度を定義できず、これまで推定が困難とされてきた潜在的かつ高次の音源情報を推定するために、目的関数を確率論的に拡張するという概念を作り出した点で極めて先駆的である。またその成果は、実装され実環境で動作しており社会的なインパクトも大きい。

よって、学術上・産業上寄与することが大きく、学位論文としてふさわしい内容であるため、本審査委員会は、本論文は博士(工学)の学位請求論文として十分な価値を有するものと認める。