

令和 5 年 5 月 11 日現在

機関番号：17301

研究種目：基盤研究(C) (一般)

研究期間：2020～2022

課題番号：20K12045

研究課題名(和文) 生体音中の特徴音の明瞭性を考慮した識別手法の検討

研究課題名(英文) Classification methods taking account clarity of characteristics in biological sounds.

研究代表者

松永 昭一 (Matsunaga, Shoichi)

長崎大学・情報データ科学部・客員研究員

研究者番号：90380815

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：本研究では、生体音に含まれる特徴音の識別(「聴診音(肺音)を用いた患者の検出」、及び「乳児の泣き声に含まれる情動の検出」の二つのタスク)において、被験者に直観的で分かりやすい認識結果を特徴音の明瞭性に基づいて高い精度で提供することを目的とした。特徴音の主観的な聴取しやすさを明瞭性と捉え、識別クラスの設定を行った。従来1つのクラスとしていたものを、明瞭性の高いクラスとその他のクラスに分類して識別を行うことが直観的に分かりやすいことが判明した。また、機械学習を用いて高精度化を行うための学習データの拡張方法、及び有効な機械学習方式とその入力パラメータを明らかにできた。

研究成果の学術的意義や社会的意義

生体音に含まれる特徴的な情報の識別では、人間の主観に応じた判断結果を必要とする場合が多い。しかしながら、識別対象とする識別クラスの設定に関しては、従来は多くの場合、識別されるクラスが Top-down に、かつ、決定論的に定められるものであり、利用者に必ずしも有益な情報を与えていない場合も存在した。この問題を解決するために、本研究では人間の主観に応じた新たな識別クラスを提示することを検討したことに意義がある。また、識別クラスの増加に伴い機械学習が学習データの不足により、学習不能となることを防ぐことができる機械学習法と学習データの増強法の一例を示せたことに意義がある。

研究成果の概要(英文)：In this study, we aimed to provide subjects with intuitive and easy-to-understand recognition results (detection classes) with higher accuracy based on the clarity of characteristic sounds in the biomedical sounds. We conducted this research through two topics, "detection of adventitious sounds in auscultation sounds (lung sounds)" and "discrimination of emotions contained in infant crying" to obtain general results. The classification classes were designed by considering the subjective easiness of detecting the characteristic sounds as clarity. We found that it is intuitively easy to understand for the subjects by classifying into a class with high clarity and other class. In addition, we developed the increasing method of training data for high accuracy using machine learning approach, effective machine learning methods and its acoustic input parameters.

研究分野：生体音処理

キーワード：肺音 泣き声 識別

1. 研究開始当初の背景

被験者(もしくはユーザ)の生体音に含まれる特徴的な音情報を用いて被験者に関する情報(例えば、疾患の有無等)を検出し被験者に提供する場合に、検出法としては従来より動的計画法や統計的手法に基づくパターン認識技術が用いられてきており、どのような識別対象のクラスに属するか(識別結果)を提供することで実現されてきた。そのため、時系列の情報から識別のための音響特徴を抽出する手法や機械学習を用いた識別手法が幅広く検討されてきた。我々も、統計的手法として隠れマルコフモデル(HMM, Hidden Markov Model)や深層学習(DNN, Deep Neural Network)を用いた識別方式を検討し、識別精度の向上を図ってきた。しかしながら、識別対象とする識別クラスの設定に関しては、例えば異常クラスと正常クラスの識別の様に、多くの場合、識別されるクラスが Top-down に、かつ、決定論的に定められるものであった。しかし、生体音に含まれる特徴的な情報の識別では、人間の主観に応じた識別クラス(たとえば、「やや異常」や「非常に異常」等)を提示することが有益である場合が多い。そのため、主観に応じた識別クラスの設定方法や識別すべきクラスが増加した場合の機械学習方法及び学習データの増強法が求められていた。

2. 研究の目的

近年の深層学習に代表される機械学習技術の進展により、いわゆるパターン認識による識別精度は、飛躍的に向上してきた。生体音の識別に関しても、機械学習の技術を取り入れることにより高精度化してきた。しかし、生体音に含まれる特徴的な情報を識別では、人間の主観に応じた判断結果を必要とする場合が多い。そこで、本研究では、生体音に含まれる特徴音の識別(において、人間の主観に応じた被験者により有用な認識結果を提供することを目的とした識別クラスの設定と、そのクラスを高精度で識別できる機械学習の方式の検討を目的とする。また、タグ付けされた生体音のデータ量は限られる場合が多いため、クラス数の増加に応じた学習データの増加方法の検討を目的とする。

3. 研究の方法

生体音に含まれる特徴音を用いた識別として本研究では「聴診音(肺音)を用いた疾患音の検出」、及び「乳児の泣き声に含まれる情動の識別」の二つの課題を対象とする。この二つの課題を並行して行う理由は、より汎用的で有益な識別クラスの設定とその識別法の確立を目指すためである。前者の課題では、聴診器で患者と健常者から収録した肺音を用いて、従来は正常音と異常音として識別していたものを、人間による異常音としての主観評価値を付与して、その評価値を用いて異常音のクラスの分離の可能性について検討する。クラスの識別手法としては、時系列を扱う機械学習である Convolutional Neural Network(CNN), Long Short Term Memory(LSTM), Transformer 等を比較検討する。また、学習データの増加に関しては副雑音を含んだ区間の切り出し位置を変化させて行い、その効果を検証する。一方、後者の乳児の情動識別では、「甘え」の情動と「怒り」の情動の識別を課題として行う。情動の強さに応じて情動のクラスを増加させることについて検討する。また、識別法としては従来は HMM や DNN を用いていたが LSTM を用いて精度の向上を図る。また、泣き声全体ではなく、情動を多く含むと想定される区間のみを情動の評価区間もしくは学習データの増強に用いることで精度の向上を図る。

4. 研究成果

聴診音を用いた疾患音の検出に関しては、副雑音を含む疾患肺音を聴取すると、副雑音のパワー値、呼吸音の強弱、雑音の混入等により、副雑音の聴取しやすさに違いがある。そこで、このような様々な副雑音の主観による聴取しやすさを「明瞭性」と定め、異常肺音のクラスの分類を行った。具体的には副雑音の明瞭性が高い順に「明瞭性:高」、「明瞭性:中」、「明瞭性:低」の3段階のラベルを付与し、その特徴の分布を調査したところ、相対的に「明瞭性:低」と「明瞭性:中」が極めて類似していたため、「明瞭性:中」と「明瞭性:低」を合わせた「明瞭性:中低」のクラスと、「明瞭性:高」のクラス間の識別を行うことが被験者に直観的でわかりやすいことがわかった。識別実験を行った結果、明瞭性の違いをある程度識別でき、「明瞭性:中低」の識別精度が「明瞭性:高」より高いことがわかった。次に、疾患音のパワー値と明瞭性に関する相関が高いことに着目し「確からしい異常音」と「異常音」の識別クラスを用いる方法を考案した。確からしい異常音の検出判定基準には、疾患音の調査の結果から副雑音区間のパワーの最大値を用いた。検出実験の結果、「確からしい異常音」を検出した精度は、右第二肋間の聴診箇所から得た肺音データでは良好であったものの、「確からしい異常音」と「異常音」の識別精度は十分ではなかった。次に、識別精度を高精度化するために、従来の一つの呼吸吸気区間ではなく、時間的により長い区間を検出対象区間とすることで、副雑音が周期的に生起することを捉えることで検出精度の向上を行った。また、従来用いていた波形を入力とする CNN を用いた識別法では計算量が膨大になり学習が困難となるため、入力を MFCC(メル周波数ケプストラム係数)とパワーとする Transformer を開発し検出精度の向上を行った。また、副雑音を含む区間を考慮して

肺音の学習データの拡張を行うことで、CNN, LSTM, Transformer を用いた機械学習を可能にした。

一方、乳児の泣き声に含まれる情動の検出に関しては、母親や保育士等が泣き声を聞いて作成した情動の主観値を用いて、それらの情動間の相関に着目し、2次元のヒストグラムを作成し、その結果、強い相関がみられた「悲しみ」と「怒り」を「怒りクラス」、「眠い」と「甘え」を「甘えクラス」とまとめて1クラスとした。「甘え」と「怒り」の2種の情動の識別を行う際に、被験者にとって重要な情報は「怒り」と「強い怒り」の分離(「強い怒り」と感じる場合は乳児への早急な対応が必要なため)と考えた。この「強い怒り」に共通する特徴的な声は「裏返った声」と「割れた声」であり、「裏返った声」はピッチが短時間で急激に上昇するという特徴があり、ピッチの変化量が閾値を超えるか否かで検出できると考えた。「割れた声」はピッチが抽出できないという特徴があったため、ピッチが抽出できない区間の時間長が設定した閾値を超えるか否かで検出した。この結果、「強い怒り」の高い検出率を達成できた。また、泣き声の中でどの区間に情動の情報がより含まれるかの検討を行った結果、単純な泣き声区間が最も情報が多いことが示せた。次に、乳児の情動がこの泣き声区間の音の立ち上がり部分に多く存在すると仮定し、従来は泣き声の収録音の全体を情動の検出対象区間として検出していたものを、パワーを用いて泣き声の立ち上がり部分を抽出し、その立ち上がりから一定時間の区間のみを情動の検出対象区間とした。また、学習音声に関してはラベルデータを用いて、泣き声の立ち上がり区間が必ず含まれる様に時間的に変化を施して切り出しを複数回行い、学習データ数を増加させた。学習及び識別手法に関しては、従来の LSTM による識別精度をより向上させるために、新たに Bidirectional Long Short Term Memory(BLSTM)により構築した。BLSTM を用いた場合に、泣き始めを考慮しない先行研究と比較して、情動の識別率を大きく向上させることができた。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計7件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 清本優希, 松永昭一, 高田寛之
2. 発表標題 乳児の泣き声を用いた情動推定における泣き声発声区間のみの効果
3. 学会等名 第29回電子情報通信学会九州支部学生会講演会
4. 発表年 2021年

1. 発表者名 河原泰雅, 高田寛之, 松永昭一
2. 発表標題 異常肺音検出におけるLSTMとCNNの比較
3. 学会等名 第29回電子情報通信学会九州支部学生会講演会
4. 発表年 2021年

1. 発表者名 弥吉大輔, 高田寛之, 松永昭一
2. 発表標題 肺音中の副雑音の明瞭性を考慮した異常肺音分類
3. 学会等名 第28回電子情報通信学会九州支部学生会講演会
4. 発表年 2020年

1. 発表者名 小川大河, 高田寛之, 松永昭一
2. 発表標題 乳児の泣き声を用いた情動推定のための 主観評価値の相関に基づくクラスタリング
3. 学会等名 第28回電子情報通信学会九州支部学生会講演会
4. 発表年 2020年

1. 発表者名 毛利文哉, 高田寛之, 松永昭一, 酒井智弥
2. 発表標題 スパースモデリングを用いた肺音分離学習による断続性副雑音検出
3. 学会等名 第28回電子情報通信学会九州支部学生会講演会
4. 発表年 2020年

1. 発表者名 阿利迪洋, 高田寛之, 松永昭一
2. 発表標題 DNN-HMM を用いた乳児の情動推定におけるセグメント時間情報付与の効果
3. 学会等名 第28回電子情報通信学会九州支部学生会講演会
4. 発表年 2020年

1. 発表者名 梅野直幹, 松永昭一, 高田寛之
2. 発表標題 複数の聴診箇所からの聴診音から選択した学習データを用いた異常肺音検出
3. 学会等名 第19回情報科学技術フォーラム(FIT)
4. 発表年 2020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	高田 寛之 (Takada Hiroyuki) (10297616)	長崎大学・情報データ科学部・助教 (17301)	

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