

令和 2 年 6 月 26 日現在

機関番号：17301

研究種目：基盤研究(C)（一般）

研究期間：2017～2019

課題番号：17K00408

研究課題名（和文）肺音中の副雑音の分離技術と統計的モデル化手法を融合した肺疾患患者識別手法の研究

研究課題名（英文）Detection of unhealthy subjects based on stochastic modeling and extraction of adventitious sounds

研究代表者

松永 昭一（MATSUNAGA, Shoichi）

長崎大学・工学研究科・教授

研究者番号：90380815

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：肺音の異常音と正常音の識別性能を向上させるための事前処理の手法として、スパースモデリングを用いる方式を検討した。聴診音をスパースモデリングにより連続性の音響特徴を持つ成分（時系列信号）と断続性の音響特徴を持つ成分（時系列信号）の二つに分離した。この分離の際に用いる音響特徴として、連続性としてはフーリエ成分を、断続性としてはウェーブレット成分を用いて分離を行う処理系を構築できた。さらに、ニューラルネットにより音響特徴パラメータを学習する枠組みを作成し、さらにその学習した特徴パラメータを用いて、正常肺音と異常肺音を後段のニューラルネットにより識別する枠組みを作成した。

研究成果の学術的意義や社会的意義

我々は日常の生活で、肺音に異常な音（通常の呼吸音と異なる音）を聴取した場合には、呼吸器系の疾患を疑い、専門家の診察・治療を受け、重篤化を避けることができる。これは呼吸器系の疾患を患うと肺音中に呼吸器系臓器の異常に起因する副雑音と呼ばれる疾患音を観測する機会が多いことを経験上利用している。このような肺疾患に起因する疾患音の有無の推定に関することを、機械学習の手法を用いてアルゴリズム化し、かつ、雑音の混入する家庭環境で採取する肺音に対しても、頑健にかつ高精度に疾患音を検出できれば、早期の肺疾患の検出の手段として役立つことができ、重篤な病気に至る前の段階で早期の治療へと繋げることが可能となる。

研究成果の概要（英文）：We introduced a sparse modeling technique to increase classification performance between normal respiration and abnormal respiration. According to the acoustic features of two types of abnormal lung sounds, continuous adventitious sounds and discontinuous adventitious sounds, and essential difference of these acoustic features, we adopt a sinusoidal basis for the resonant-cry sounds and a wavelet basis for the short noise sound. We devised a training technique of weighting factors of two types of basis using neural networks.

研究分野：音情報処理

キーワード：呼吸音 肺疾患 機械学習

様式 C - 19, F - 19 - 1, Z - 19 (共通)

1. 研究開始当初の背景

肺音中の疾患音の検出に関する研究は従来より世界的に行われてきた。初期の段階では、複数の標準パターン(マルチテンプレート)を用いたマッチングによる判定が主に行われてきた。しかし多様な被験者の肺音に対して高い精度で識別を行うことは難しく、雑音に弱いという問題があった。また、海外では肺音データベースを作成し、肺音中の特定の疾患音の検出に関して研究が行われた。これらは、典型的な疾患音の検出は可能であるが、疾患音の音響的特徴や出現するタイミングの多様性への対応が難しかった。これに対し我々は、隠れマルコフモデル(HMM)が時間軸の伸縮に柔軟に対応できる統計的手法であることに着目し、肺音中の音響的特徴の顕著な区間を肺音セグメント(連続性副雑音セグメント等)と定義し、「音響セグメントHMMを用いた肺音中の異常音の検出」の研究を行い、最尤の音響セグメント系列を求めることで、健常者の正常な肺音と患者の明瞭な疾患音が含まれる肺音を識別できること、及び、患者と健常者のある程度高精度に識別することができることを示した。しかし、混入する心音や雑音による音響的特徴の重畳が識別精度をより向上できない原因となっており、継続時間による後処理により対策を施してきたものの、必ずしも十分な識別精度とは言えなかった。

一方で、異なる音源から発せられ音が重畳した場合に、それらの音を分離する研究として、非負値行列因子分解やスパースモデリング等による手法が発展してきた。そのような状況において、肺音中に重畳する疾患音と呼吸音の特徴の違いを利用して、スパースモデリングを行うことで、呼吸音と副雑音(連続性と断続性)を分離できることを示した。しかし、現状のレベルでは、あらゆる肺音から副雑音を精度よく分離できる訳ではなかった。

また、肺音データは複数の聴診箇所から聴取しているが、聴診箇所ごとに肺音の音響特徴が異なるため、評価する肺音データと同じ聴診箇所のみデータを用いて音響モデルを学習していた。また、より多くの肺音データを収集するには、医師や患者の協力が必要となるため大規模なデータベースを整備することは困難であった。そのため少ないデータ数で学習を行うこととなり、認識精度が低い原因となっていた。

2. 研究の目的

本研究では、聴診音(肺音)を用いた、呼吸器疾患の高精度でかつ頑健な検出系の実現を目的とする。その中で、スパースモデリングの手法を用いて肺音中の疾患音を高い精度で分離できる手法を開発し、疾患音及び患者を高精度に検出することが主な目的である。また、本研究の特徴は、副雑音の分離技術と肺音の統計的モデル化手法を統合した患者の識別手法の融合であり、さらに深層学習を加えることで、学習データの拡張や適切な音響特徴量の検討を行って患者及び健常者の識別性能の向上を図ることも目的である。

3. 研究の方法

従来、疾患音の検出精度の低下を招いていた混入雑音の影響を軽減するために、まず、スパースモデリングの手法を用いて肺音中の疾患音(副雑音、例えば喘息における笛音)を分離し、分離した音を統計的識別手法の入力とし、疾患音及び患者を検出する。また、副雑音の分離時に効果の大きい音響パラメータ(基底)を統計モデルの音響特徴量とすることで高精度化を狙う。これらの特徴量を用いることで、小規模な学習データにおいて、深層学習の適用を効率的に行う。また、小規模な学習データ(聴診箇所に依存した学習データ)を音響データの類似性を利用して拡張することで、識別精度の向上を狙う。

4. 研究成果

まず、肺音の異常音と正常音の識別性能を向上させるための事前処理の手法として、スパースモデリングを用いる方式を検討した。聴診音をスパースモデリングにより連続性の音響特徴を持つ成分(時系列信号)と断続性の音響特徴を持つ成分(時系列信号)の二つに分離した。この分離の際に用いる音響特徴として、連続性としてはフーリエ成分を、断続性としてはウェーブレット成分を用いて分離を行う処理系を構築できた。次に、この処理系の評価方式として、それぞれの時系列信号に対して、連続性の副雑音区間及び断続性の副雑音区間に対して、分離前の原音(聴診音)と比較してどの程度のパワー値の変化が生じているかを測る処理系を構築した。この結果、連続性の音響特徴を持つ時系列信号では断続性の副雑音区間でのパワー値が大きく減少し、断続性の音響特徴を持つ時系列信号に対しては連続性の副雑音区間のパワー値が大きく減少し、スパースモデリングが聴診音を効率よく断続性の音と連続性の音に分離できることを示せた。一方で、スペクトグラムを詳しく調査すると断続音と連続音が分離できていない箇所(特に連続音の立ち上がり部分)があることがわかった。

従来、スパースモデリングを適用していた聴診箇所が限られていたため、聴診箇所を増やして副雑音の抽出を行った結果、一部の聴診箇所ではその効果を、異常肺音と正常肺音の識別実験において確認できた。しかしながら、心音が混入する箇所等、その効果を確認できない聴診箇所も存在した。これは、上記のスパースモデリングでは、心音やパルス性の雑音等が分離した信号に残るためである。そこで、あらかじめ肺音中に混入する(検知できる)心音の割合を推定し、その混入の頻度が高ければ心音に関する統計モデルを用い、そうでなければ用いないという手法と組み合わせる効果について検討した。聴診箇所に適した手法を用いることで肺疾患者の識別率の向上を目指した。また、これまでは片側の聴診箇所のみで肺疾患者識別を行って

たが、一方の肺で肺疾患患者識別を行うともう一方の肺に疾患があった場合、異常音が検出しづらいつらいつらと考え、左右の聴診箇所を肺音を用いて総合的に肺疾患患者識別を行った。心音の混入が考えられる聴診箇所では心音検出率を調べ、心音モデルと通常モデルの使い分けを行い、それぞれの聴診箇所ごとに発生しやすい副雑音に頑強なモデルを作るために、上側の聴診箇所ではスパースモデリングによって分離した連続性の肺音成分を、下側の聴診箇所では断続性の肺音成分を学習データに加えた。聴診箇所ごとの特徴を考慮し、複数の聴診箇所の肺音を用いて肺疾患患者識別を行った結果、疾患患者の識別率は向上したが、健康者の識別率は低下した。しかし、左右の聴診箇所を用いての肺疾患患者識別は疾患患者、健康者共に識別率が向上し、聴診箇所の音響特性に応じた音響モデルを用いることの有効性を示した。

また、スパースモデリングを用いて連続性と断続性の音を分離する際に、フーリエ成分とコサイン成分のパラメータを事前に与えているが、これをニューラルネットにより学習する枠組みを作成し、さらにその学習した特徴パラメータを用いて、正常肺音と異常肺音を後段のニューラルネットにより識別する枠組みを作成した。具体的には、時間信号を入力とし、スパースモデリングによる肺音分離と、CNN(Convolutional Neural Network)を接続した特徴抽出から識別までを一括で行なう肺音識別器の深層学習の構築を行った。まず、ハイパーパラメータ(基底行列、正則化パラメータ)の学習可能性を検討した。従来の研究では、正則化パラメータは非負である必要があるにもかかわらず、その制約が考慮されていなかった。基底行列は、初期値にフーリエ係数とウェーブレット係数を用いて、周波数成分と時間成分を学習することを意図しているが、学習の過程で大きく初期値から外れていく可能性があった。正則化パラメータの非負性を担保するための制約として2つの方法を提案した。正則化パラメータを学習時、負の要素を強制的に0にする方法と、正則化パラメータの各要素をあらかじめ2乗する方法である。基底行列が周波数成分と時間成分を学習するための制約としては2つの方法を提案した。Weight decayを用いることで基底行列が大きくなることを防ぐ方法と、基底行列における、パラメータ更新量に閾値を設け、更新量が閾値を超えると学習を打ち止めるという方法である。制約をつけた場合と、制約をつけていない場合の肺音識別(正常肺音/異常肺音)の実験の結果、スパースモデリングによる分離部分を学習していないCNNによる識別精度が最も高くなった。また、そのときの識別精度はGMM-HMMより高い精度となった。一方で、制約をつけて学習を行なった肺音識別器の識別精度は低下した。スパースモデリングにおけるパラメータを学習したことによる精度向上は見られなかったが、制約をつけていない場合と制約をつけた場合の肺音識別器の識別精度を比較すると、制約をつけた場合がより高くなり、スパースモデリングによる信号分離とCNNを接続した肺音識別器において、学習時にパラメータに制約をつけることにより精度が向上することを示した。

次に、重み付けされた正則化パラメータベクトルの導入と、それを学習可能にする展開技法の検討を行った。正則化パラメータをベクトル化することで、各肺音成分の重要性を表す尺度として用い、逆誤差伝搬法が利用できるように、スパースモデリングの問題を計算グラフ化することで正則化パラメータベクトルを学習可能にした。正則化パラメータベクトルは不必要な要素を押しつぶす作用があるので、その値が0として学習される成分は重要であり、大きい値になったときは不要であると仮定した。正則化パラメータベクトルの学習の効果を検証するために断続性副雑音の分離の学習実験と断続性副雑音区間の検出の実験を行った。その結果、実験断続性副雑音を含む肺音のみの学習では学習断続性肺音区間を検出精度の向上につながることはなかったものの、肺音の成分を学習することで断続性副雑音の検出の際、重要と思われる部分(約600Hz以降の領域)を発見することができた。

また、音響特徴量の有効性の検討としてバンドパスフィルタの出力を用いる場合とFFTケプストラムを用いる場合の異常音を含む呼吸音の検出精度の比較を行った。この結果、FFTケプストラムを用いた方が、安定的に良い精度が得られることがわかった。

さらに、異常音検出や疾患患者の検出における十分な学習データ量の確保は重要な問題である。評価箇所と異なる箇所の肺音データにも学習に有効なデータが存在すると考え、複数の聴診箇所のデータを利用して学習データを拡張する二つの手法を提案した。学習データの拡張を行い、より多くのデータを用いて学習を行うことで、異常音の識別精度向上を目指した。一つ目の手法(選択手法)は肺音データにおける音響特徴量の類似性を考慮した選択手法である。これは評価箇所の肺音データと音響特徴量が類似している肺音データを学習に用いることで識別精度が向上するという考えに基づく。肺音データの音響特徴量としてメルケプストラム係数(MFCC)とパワーを利用し、評価箇所の肺音データ分布と評価箇所と異なる箇所の肺音データとの類似性を求める。類似性の計算には、評価箇所の肺音データの分布と他の聴診箇所の肺音データとの分布間の距離を求めるためにパタチャリヤ距離を用いた。この距離が小さいほど同様の音響特徴がある。類似性があらかじめ設定した閾値以上であれば、評価箇所のデータと類似しているものと考え、そのデータを学習データとして採用した。二つ目の手法(選択手法)は候補データの正常と異常の尤度差による信頼性を考慮した選択手法を行った。これは尤度差が大きいほど正常音と異常音を明確に区別することができ、このようなデータを学習に用いることで識別精度が向上するという考えに基づいた。まず、評価箇所と異なる聴診箇所の肺音データである候補データについて、一つ一つの呼吸気ごとに正常と異常の尤度を計算し、その二つの尤度差を計算する。この尤度差について事前に定義した閾値よりも大きい場合、候補データを学習データとして選択した。二つの提案手法の比較対象として2種類のBaselineを準備した。一

つは、評価箇所と同一箇所のみ肺音データを利用し学習を行った音響モデル(Baseline 1)であり、もう一つは、全ての聴診箇所からの全ての肺音データを利用し学習を行った音響モデル(Baseline 2)を用いた。これらの音響モデルを用いて異常肺音と正常肺音の識別実験を行った。Baseline 1の識別率はBaseline 2の識別率から有意に低下した。この結果から、他の聴診箇所の肺音データを全て用いることによる学習データ数の拡張は識別精度向上につながらないことを示した。また提案手法Aと手法Bの識別率は、Baseline 2の識別率から向上し、提案手法の有効性を示した。次に、選択手法Cと手法Dについての比較を実験により行った。その結果、音響特徴量の類似性を用いた手法Cの方が手法Dより高精度な識別ができる可能性があることを示した。しかし、手法Dにおいて、正常と異常を正しく分類することのできた全ての呼吸を拡張データとして利用した場合でも高い精度を維持することができたのに対して、手法Eでは拡張数が大きくなると識別率が大幅に低下した。この結果から、手法Dの方が、閾値の変動に対して頑健であることがわかった。この結果、事前に設定する必要のある閾値に対してよりロバストである手法Eは手法Dより実用的であることが判明した。

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 4件/うち国際共著 0件/うちオープンアクセス 0件）

1. 著者名 Umeno Naoki, Yamashita Masaru, Takada Hiroyuki, Matsunaga Shoichi	4. 巻 -
2. 論文標題 Training Data Expansion for Classification between Normal and Abnormal Lung Sounds	5. 発行年 2019年
3. 雑誌名 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	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） 10.1109/APSIPAASC47483.2019.9023022	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 Ohkawa Kimitake, Yamashita Masaru, Matsunaga Shoichi	4. 巻 -
2. 論文標題 Classification Between Abnormal and Normal Respiration Through Observation Rate of Heart Sounds Within Lung Sounds	5. 発行年 2018年
3. 雑誌名 2018 26th European Signal Processing Conference (EUSIPCO)	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） 10.23919/EUSIPCO.2018.8553364	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 Kurokawa Takayuki, Miura Tasuku, Yamashita Masaru, Sakai Tomoya, Matsunaga Shoichi	4. 巻 -
2. 論文標題 Emotion-Cluster Classification of Infant Cries Using Sparse Representation	5. 発行年 2018年
3. 雑誌名 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	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） 10.23919/APSIPA.2018.8659553	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 Yamashita Masaru, Miura Tasuku, Matsunaga Shoichi	4. 巻 -
2. 論文標題 Distinction between healthy individuals and patients with confident abnormal respiration	5. 発行年 2017年
3. 雑誌名 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	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） 10.1109/APSIPA.2017.8282199	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 -

〔学会発表〕 計6件（うち招待講演 0件／うち国際学会 1件）

1. 発表者名 Wint Yee, Masaru Yamashita, Hiroyuki Takada, Shoichi Matsunaga
2. 発表標題 Acoustic Features for Detection of Emotion in Infant Cry and Abnormal Lung Sounds
3. 学会等名 2019年度 第27回電子情報通信学会九州支部学生講演会
4. 発表年 2019年

1. 発表者名 梅野直幹, 山下優, 高田寛之, 松永昭一
2. 発表標題 異常肺音検出のための聴診音のスペクトルの類似性を考慮した学習データの拡張
3. 学会等名 2019年度 第27回電子情報通信学会九州支部学生講演会
4. 発表年 2019年

1. 発表者名 増田紗依, 山下優, 高田寛之, 酒井智弥, 松永昭一
2. 発表標題 聴診箇所ごとの音響特徴を考慮した肺疾患患者識別
3. 学会等名 2019年度 第27回電子情報通信学会九州支部学生講演会
4. 発表年 2019年

1. 発表者名 Wint Yee, Yoshinori Hirano, Masaru Yamashita, Hiroyuki Takada, Shoichi Matsunaga
2. 発表標題 Effective Acoustic Feature Parameters for Emotion Estimation of Infants' Cry Sound and Detection of Abnormal Lung Sound
3. 学会等名 The Tenth International Conference on Science and Engineering 2019 (国際学会)
4. 発表年 2019年

1. 発表者名 三浦亮, 黒川貴之, 山下優, 酒井智弥, 松永昭一
2. 発表標題 乳児の泣き声を用いた情動推定におけるスパースモデリングの有効性
3. 学会等名 日本音響学会2018年秋季研究発表会
4. 発表年 2018年

1. 発表者名 三浦亮, 山下優, 松永昭一, 西辻臯勢, 志波舜平, 酒井智弥
2. 発表標題 統計モデルに基づく肺音分類のためのスパースモデリングの有効性
3. 学会等名 日本音響学会2017年秋季研究発表会
4. 発表年 2017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	酒井 智弥 (SAKAI Tomoya) (30345003)	長崎大学・工学研究科・准教授 (17301)	