

機関番号：17301

研究種目：基盤研究（C）

研究期間：2008～2010

課題番号：20500157

研究課題名（和文） 生体音に含まれる特徴的音情報の統計的手法による検出—高精度な識別器を目指して—

研究課題名（英文） Detection of acoustic features in human biological sounds based on statistical approach

研究代表者

松永 昭一（MATSUNAGA SHOICHI）

長崎大学・工学部・教授

研究者番号：90380815

研究成果の概要（和文）：本研究では統計的手法に基づき正常肺音と異常肺音を識別する手法の研究を行った。本手法の特徴は異常肺音を正確に検出するために、音響特徴を扱うための統計モデルと、呼気/吸気に含まれる音響的特徴セグメントの生起順序に関する統計モデルを用いる。また、韻律情報を用いて乳児の情動クラスを推定する手法の研究を行った。本手法の特徴は泣き声のセグメントと無音セグメントの継続時間の割合を韻律情報として用いることであり、従来のスペクトル情報を用いた手法より識別性能が大きく向上した。

研究成果の概要（英文）：In this research project, we devised a novel classification procedure for distinguishing between normal and abnormal respiratory sounds on the basis of stochastic approach. The main characteristic of our procedure is that two stochastic models are used to detect abnormal respiratory sounds precisely: hidden Markov models for acoustic spectral features and bigram models for the occurrence of acoustic segments in each respiratory period. We also devised an approach to the classification of emotion clusters using prosodic features. In our approach, we use the duration ratios of specific acoustic segments—resonant cry segments and silence segments—in the infants' cries as prosodic features. The classification performance of our approach using the segment duration ratios was significantly better than that of the method using spectral features.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	1500000	450000	1950000
2009年度	1000000	300000	1300000
2010年度	900000	270000	1170000
総計	3400000	1020000	4420000

研究分野：音声言語情報処理

科研費の分科・細目：情報学・知覚情報処理・知能ロボティクス

キーワード：生体音, 肺音, 情動推定, 特徴パラメータ

## 1. 研究開始当初の背景

我々は日常の生活において、肺音、心臓音、乳児の泣き声などの生体音において、平常と異なる音を聴取した場合に異常を感じ、その要因（例えば病気等）を推察する場合がある。これらの音は人間から発せられる非言語音であるが、可聴帯域の音であり、その要因の違いにより音響的特徴が異なる

ために、要因の推察に経験的に利用している。特に、医師は肺音や心臓音の違いにより病気等の要因を推察し、保母や母親は乳児の泣き声の違いにより乳児の状態（空腹や怒り等）を知る。これらの医師や保母らの所謂エキスパートは、推察した要因に対処するために患者や乳児に対応する処置を行う。このようなエキスパートが行う要因

の推察(本研究では「肺音における疾患音の検出」と「乳児の泣き声からの情動識別」)に関することを、統計的手法を用いてアルゴリズム化し高精度に自動識別を行うことを目的として本研究を行う。血圧計や体脂肪計のように、家庭で使える簡便な機器として実現することができれば、「肺音中の疾患音の検出」は医療施設が十分でない環境(離島や過疎地等)における早期の異常検出の手段として、「泣き声の情動検出」は核家族化された社会における親子間コミュニケーションの支援機器として大いに役立つ。

## 2. 研究の目的

本研究では、家庭において簡易に利用できる生体音の識別装置(具体的には肺音中に含まれる疾患音の検出、乳児の泣き声における情動の識別装置)の実現を目指した、高精度に生体音を識別できる統計的モデルの作成法の研究と頑健で高速な識別アルゴリズムとその実装化に向けた基礎研究を行う。本研究を行うことで、肺音中に含まれる疾患音(例えば喘息における笛音等)を高精度に検出したり、乳児の泣き声を入力することで情動(例えば、乳児がお腹を空かせて泣いている等)を実時間で識別するための基盤となる手法を構築する。本研究の最も大きな特徴は、統計的アプローチにより生体音の識別を実現することにある。

## 3. 研究の方法

本研究では、生体音として、肺音(成人)と乳児(10カ月前後)の泣き声の2種を用い、肺音における副雑音の検出、泣き声における情動の推定を実時間で高精度に行う手法の開発を行う。本研究では、(1)生体音の音響モデル化のための基礎検討として、①生体音の音響的特徴の調査(どのような音が副雑音や情動を識別するために有効であるか)、②生体音識別のための音響特徴パラメータの検討(スペクトル特徴や韻律特徴と継続時間)、③特徴パラメータの統計的モデル化(音響セグメントのHMMによるモデル化やセグメント生起確率のセグメントbigramによるモデル化)と、(2)生体音の実時間識別アルゴリズムの研究として、④生体音の頑健で高精度な識別・検出法(雑音による影響の軽減)、⑤生体音のFPGAを用いた実時間識別法の検討を行う。平成20年度には(1)の音響的基礎検討を重点的に行い、この結果を用いて(2)の生体音の識別実験を開始する。平成21年度以降は生体音モデルの高精度化の検討、生体音の検出手法及び実時間での識別法の研究を主として行う。

## 4. 研究成果

### 肺音中の疾患音の検出

肺音中の異常音(副雑音)を自動検出する研究は既に広く行われており、隠れマルコフモデルを用いた音響的特徴の統計モデルが異常音検出に有効であることや、副雑音を種類によって細分化することで、検出率が向上することが確認されている。

しかし音響的特徴の統計モデルの接続規則が全て等確率で記述されているという問題があった。本論文ではその接続規則に着目し、呼気・吸気内の各音響モデル間の生起確率を利用した「肺音列モデル」を作成し、異常音検出に利用することで、より高度なシステムの構築を目指した。肺音列モデルに重みを加え、検出結果の与える影響を制御しながら最適な重みを求めた。接続規則として文法を用いたときと肺音列モデルを用いたときの識別精度を表1に示す。

表1 文法を用いたときと肺音列モデルを用いたときの識別精度の比較[%]

接続規則	識別精度
文法	85.3
肺音列モデル	87.3

肺音列モデルを使用することで正解精度を高めることに成功し、最適な重みを得ることができた。

次に異常音検出実験によって識別された呼気・吸気を利用することで、被験者を肺疾患患者か、健常者かの識別を行った。識別された呼気・吸気の関連性を利用するために呼気・吸気間の生起確率を利用した「呼吸列モデル」を作成し、肺疾患患者判定実験に利用した。呼吸列モデルはバイグラムモデルとディスタンスバイグラムモデルの2種類を作成し、バイグラムモデルで呼気・吸気と吸気・呼気の生起確率を、ディスタンスバイグラムモデルで呼気・呼気と吸気・吸気の生起確率を考慮した。また2種類のモデルそれぞれに重みを付けることで識別精度が最大となる重みを求めた。呼吸列モデルを用いた肺疾患患者の識別率を表2に示す。

表2: 呼吸列モデルを用いた肺疾患患者の識別率[%]

疾患患者	健常者	平均
85.3	95.5	90.0

呼吸列モデルが肺疾患患者の識別に有効であることを示し、呼吸列モデルの重みを変化させ制御することで、90%の精度で被験者を疾患患者か健常者か識別することに成功した。

また、隠れマルコフモデル(HMM)を用いた音響的特徴の統計的モデルが異常音検出に有効であることを示してきた。しかし、肺音の収録において体内の他の器官から発せられる音や様々な雑音を副雑音として誤認識するという課題があった。また、雑音と副雑音のスペクトル情報が類似しているため、スペクトル情報だけで雑音と副雑音を区別するのは難しいという課題もあった。そこで、雑音と副雑音の違いを調査し、雑音と副雑音の継続時間が異なっていることに着目した。そこで、雑音と副雑音の継続時間分布を用いることで、スペクトル情報だけでは難しかった短時間の雑音と副雑音の誤認識を減らし、HMMを用いた肺音の正常と異常の識別率をさらに高めることを目的とする。雑音と副雑音の継続時間分布から閾値を定め、HMMを用いた統計的手法によって出力された副雑音の継続時間が閾値以下なら、副雑音区間の尤度にペナルティを与えることを提案した。提案手法を取り入れる前と後の識別率表3に示す。

表 3 スペクトル情報のみ雑音と副雑音の継続時間分布を用いた識別精度の比較[%]

識別に用いるパラメータ	識別精度
スペクトル情報雑音のみ	85.5
スペクトル情報+雑音副雑音の継続時間分布	89.9

有意水準 5%の片側検定を行った結果、提案手法の識別率の統計的向上が示され、雑音と副雑音の継続時間分布の違いを考慮した異常肺音の検出が有効であることが示された。しかし、識別結果の中には、検出した副雑音の継続時間が非常に短い区間になっている場合があり、正確な継続時間を表せていないものもある。そのため、雑音と副雑音を継続時間のみで区別することができていないことが考えられる。今後は継続時間をさらに正確に求める手法を検討する必要がある。また、提案手法では正常音の識別率を上げることのみが可能なため、今後は、雑音の音響的特徴の統計的モデルを導入し、雑音のスペクトル情報と時間情報を考慮した実験を行う必要がある。

#### 乳児の泣き声からの情動識別

一方、乳児の泣き声を用いた情動推定に関する研究は歴史が浅く、有効な音響特徴量は未だ明確にはわかっていない。これまでは、その音響特徴量としてスペクトル情報を用いることが多かったが、情動に関する情報が泣き声のスペクトル情報だけに含まれるとは限らず、その他の音響特徴量に乳児の情動が表れている可能性もあると考えられる。

そこで、プロジェクトでは泣き声の韻律情報を調査し、その中でも泣き声データを構成するセグメントの中で、泣き声音と無音の時間割合を表4に示すように情動によって異なる傾向があることを示した。

表 4 甘え、怒りにおける泣き声音、無音の時間割合[%]

	泣き声音	無音
甘え	33.6	29.6
怒り	53.0	16.3

次に、泣き声音、無音を利用した情動推定を行うための前準備として、パワーとピッチ周波数を用いてセグメンテーションを行ったところ、セグメンテーション精度はF値が約75%となり良好な結果であった。そして得られたセグメンテーション結果から時間割合を算出し、正規分布で近似した時間割合分布モデルを用いて情動推定を行った結果、71%の精度で情動推定できた。このことから、泣き声音、無音の時間割合が情動推定に有効であることを示した。また、人手で付与したラベルを用いて、泣き声音、無音の時間割合を求め、情動推定を行ったところ、同程度の正解率であった。

また、スペクトル情報を用いた情動推定手法の改善も行い、スペクトル情報、韻律情報の両情報を用いた情動推定手法では、72%の精度で情動推定できたことから、それぞれ単独で推定するよりも、高精度に情動を推定できることを示した。

次に、スペクトル情報を用いてセグメンテーションを行い、各情動の最尤のセグメンテーション結果から泣き声音と無音のセグメントの時間割合を求める。そして、泣き声音と無音のガウス分布から時間割合の値を確率値に変換し、対数を取り尤度を算出する。次に、セグメントの時間割合から求めた尤度とスペクトル情報から求めた尤度を足し合わせる。このようにして算出した各情動の尤度を比較して高い方を情動推定結果とする。

提案手法を評価するために、スペクトル情報のみを用いた場合と韻律情報のみを用いた場合、スペクトル情報と韻律情報を用いた場合の3種の識別実験を行った。この結果を表5に示す。スペクトル情報のみを用いた場合、学習データとして複数の乳児を用いた方が良いことがわかった。

表 5 スペクトル情報のみを用いた情動推定結果[%]

評価乳児	学習に用いた乳児	正解率
乳児 D	D	64.3
乳児 X	X	63.6
乳児 Y	Y	66.7
全乳児	複数の乳児	62.8
乳児 D	複数の乳児	66.7
乳児 X	複数の乳児	72.2
乳児 Y	複数の乳児	66.7

本研究では、FPGA を用いる利点である小型化・高速化を活かしたりリアルタイム生体音認識システムを目指して、特徴分析処理と音認識処理を音認識における処理単位であるフレームに同期して行う、フレーム同期型音認識システムの提案と検証を行った。

フレーム同期において最大の問題となり得るのが、1 クロックあたりの処理工程が増加することで発生する、システム全体の処理能力の低下である。この問題に関して、HMM の処理工程を分割して、1 クロックあたりの処理を減らすことで対処を行った。HMM に着目した理由として、複数のモデルを同時に動かす必要があるため1クロックあたりの処理が大きな負荷になっていた点と、音データ入力を待つ動作する特徴分析処理と違い、HMM そのものは実行時間が短時間で済む点がある。そのため、HMM の出力確率の計算式を変形し、1 クロックあたりの処理を簡略化することで、その問題に対処した。

最終的に、制御カウンタの実装を行ってロジックエレメント数 33, 216 の FPGA 上でフレーム同期型音認識システムの動作が可能となり、システムの評価でも同様のアルゴリズムを C 言語の double 型 (64bit 浮動小数点) で実装した後比較を行ったところ、変数型による性能差は出たものの、良好な精度を得ることができた。

一般的に音声認識に特徴量のパワーは有効だと確認されているが、先行研究は特徴量に 12 次元のケプストラム係数を使用し、パワーは使用していない。その理由として、パワーはダイナミックレンジが大きく、オーバーフローしてしまうためである。本研究ではパワーのレンジ変換方法を検討し、FPGA 上でのケプストラムとパワーを特徴量として用いた、音認識アルゴリズムの実装を目的とする。本研究では、パワーのレンジを変更するのに 3 つの手法を提案した。変換手法 1 はパワーを定数  $\alpha$  で割る方法、変換手法 2 はパワーの対数をとった後に定数  $\alpha$  をかける方法、変換手法 3 はパワーを定数  $\beta$  で引いた後に定数

$\alpha$  で割る方法である。この 3 つの手法と従来の手法を FPGA 上で、同様のアルゴリズムで行う C 言語のプログラムで認識を行った。認識率の向上は見られなかったものの、3 つの手法とも FPGA 上で動作可能となり、解決すべき課題が明確になった。

課題の一つとして HMM パラメータ、FFT ケプストラムのオーバーフローがある。解決方法として 16bit 固定小数点のフォーマットを HMM パラメータ、FFT ケプストラムがともにオーバーフローしないように見直す必要がある。認識精度以外の課題として、特徴量抽出部と認識部の同時実装がある。先行研究により FPGA に特徴量抽出部と認識部と一緒に載る回路規模に達している。本研究の結果で変換手法 1 が FPGA 上での計算量が少なさ、認識率がともに他の変換手法より高かったため、変換手法 1 を採用して特徴量抽出部を実装すると、認識率が向上し回路規模も大幅に大きくならずに済み、FPGA 上への実装することが可能であると考えている。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 11 件)

- ① 山内勝也, 山下優, 松永昭一, 宮原末治: 統計モデルを利用した高精度な異常肺音検出手法の検討, 第 47 回生体医工学会大会, 2008 年 5 月, 神戸
- ② 山内勝也, 松江駿一, 山下優, 松永昭一, 宮原末治: 副雑音の発生タイミングを考慮した異常肺音検出手法の検討, 第 33 回肺音(呼吸音)研究会, 2008 年 11 月, 東京
- ③ 山下優, 山内勝也, 松永昭一, 宮原末治: 統計モデルを用いた異常肺音検出における音響モデル詳細化の効果, 日本音響学会 2009 年春季研究発表会, 2009 年 3 月, 東京
- ④ Shoichi Matsunaga, Katsuya Yamauchi, Masaru Yamashita, Sueharu Miyahara: Classification between normal and abnormal respiratory sounds based on maximum likelihood approach, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2009 年 4 月, 査読有, 台北
- ⑤ 山本仁, 山下優, 山内勝也, 松永昭一, 宮原末治: 副雑音クラスの連鎖生起確率を用いた異常肺音の検出, 日本音響学会 2009 年秋季研究発表会, 2009 年 9 月, 郡山
- ⑥ 道脇慎司, 山内勝也, 松永昭一, 山下優, 篠原一之: 泣き声を用いた乳児の情動推定のための有効な韻律的特徴の検討, 日本音響学会 第 8 回学生のための研究発表会, 2009 年 11 月, 熊本
- ⑦ 辻恭志, 山内勝也, 山下優, 松永昭一, 小栗清: FPGA を用いた FFT ケプストラム係数の抽出法の検討, 日本音響学会 第 8 回学

生のための研究発表会, 2009年11月, 熊本  
⑧ Hitoshi Yamamoto, Shoichi Matsunaga, Masaru Yamashita, Katsuya Yamauchi and Sueharu Miyahara, Classification between normal and abnormal respiratory sounds based on stochastic approach, International Congress on Acoustics 2010 (ICA), 2010年8月, 査読有, Sydney (Australia)

⑨ 辻恭志, 山内勝也, 松永昭一, 小栗清: FFT ケプストラムと HMM を用いた単語音声認識の FPGA 上への実装, 電子情報通信学会九州支部学生大会, 2010年9月, 福岡

⑩ 北原一樹, 道脇慎司, 松永昭一, 山内勝也, 山下優, 篠原一之: 乳児の泣き声を構成する音響特徴セグメントの継続時間割合を考慮した情動推定, 日本音響学会 2010 年秋季研究発表会, 2010年9月, 大阪

⑪ 山下優, 松永昭一, 宮原末治, 正常音と副雑音を含む呼気吸気音の尤度差を利用した肺疾患患者の検出法の検討, 日本音響学会 2011 年春季研究発表会, 2011年3月, 東京

## 6. 研究組織

### (1) 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

### (2) 研究分担者

山内勝也 (YAMAUCHI KATSUYA)

長崎大学・工学研究科・助教

研究者番号: 10380718

### 研究分担者

小栗清 (OGURI KIYOSHI)

長崎大学・工学研究科・教授

研究者番号: 80325670