

## 科学研究費助成事業 研究成果報告書

平成 29 年 6 月 13 日現在

機関番号：82626

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26330216

研究課題名(和文) 駆動源HMMのトポロジ 自動生成を用いた病的音声の疾患検知

研究課題名(英文) Voice-pathology analysis based on automatic topology generation of glottal source HMM

研究代表者

佐宗 晃 (SASOU, AKIRA)

国立研究開発法人産業技術総合研究所・知能システム研究部門・主任研究員

研究者番号：50318169

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：声帯疾患を伴った病的音声の検知を音響的に非侵襲で実現する音声分析評価システムの開発が望まれている。このようなシステムを実現するためには、観測音声から声帯音源を高精度に推定する音声分析手法が不可欠である。本研究課題では、調音フィルタをAuto-Regressive過程で表し声帯音源をHidden Markov Modelで表したAR-HMMに基づいて、そのパラメータ推定と、HMMの状態をMinimum Description Length基準に基づき逐次的に分割するトポロジ最適化とを組み合わせた新しい音声分析法を構築する。

研究成果の概要(英文)：Voice-pathology detection from a subject's voice is a promising technology for the pre-diagnosis of larynx diseases. Glottal source estimation in particular plays a very important role in voice-pathology analysis. To more accurately estimate the spectral envelope and glottal source of the pathology voice, we propose a method that can automatically generate the topology of the Glottal Source Hidden Markov Model, as well as estimate the Auto-Regressive (AR)-HMM parameter by combining the AR-HMM parameter estimation method and the Minimum Description Length-based Successive State Splitting algorithm.

研究分野：音声・音響信号処理

キーワード：音声分析 声帯音源 AR-HMM

## 1. 研究開始当初の背景

音声は最も重要なコミュニケーション手段であり、高齢者のみならず人が充実した社会生活を送るために重要な要素である。しかし、喉頭がんの進行により喉頭全摘手術を余儀なくされ、自分の声を失う高齢者は少なくない。このような状況を改善するには、声帯疾患や喉頭がんの初期症状の1つとしてあらわれる嚔声を早期に発見することが重要である。従来、言語聴覚士のような専門家が実施している病的音声の検知や嚔声の聴覚印象評価に匹敵する機能を、例えば、スマートフォンのアプリケーションとして実装し、ユーザの通話音声などを常に監視することで、声帯疾患を伴った病的音声の検知をより早期に、そしてより低コストに実現できる音声分析評価システムの開発が望まれている。

## 2. 研究の目的

声帯疾患を伴った病的音声の検知を、音響的に非侵襲で実現する音声分析評価システムを構築するには、調音器官の共振・反共振特性を受けた観測音声からは直接観測できない声帯音源を、高精度に推定できる音声分析手法が不可欠である。本研究課題では、調音フィルタを Auto-Regressive 過程で表し、声帯音源を Hidden Markov Model で表した AR-HMM に基づいて、声帯疾患を伴った声帯音源を観測音声から高精度に推定する音声分析手法を構築する。そして、得られる音響的な特徴と各声帯疾患や嚔声の声質との対応関係を明らかにすることで、声帯疾患を伴った病的音声の検知や嚔声の客観的な声質評価を実現する音声分析基盤技術の確立を目的とする。

## 3. 研究の方法

本研究課題では、(1) トポロジ自動生成 AR-HMM 分析、(2) 病的音声の検知、(3) 嚔声の性質評価に関する各手法の研究開発に取り組んだ。以下に各手法の研究方法について述べる。

### (1) トポロジ自動生成 AR-HMM 分析

研究代表者は、過去の研究で、調音フィルタを AR 過程で表し、声帯音源を HMM で表した AR-HMM による音声分析手法が、従来の線形予測法では基本周波数の高い音声から声道特性の推定が困難となる問題を、大幅に改善できることを示した。声帯振動による駆動源は周期波形となるため、HMM のトポロジは状態間を周回するリング状に固定して、AR-HMM のパラメータを推定していた。一方、声帯疾患を伴った病的音声の場合、必ずしも周期性を表すリング状トポロジが最適とは限らない。声帯疾患を伴った声帯音源をより高精度に推定するには、各話者の声帯音源に対して HMM の最適なトポロジを適応的に生成しながら音声进行分析する必要がある。これを実現するために、Minimum Description Length 基準

に基づいて逐次的に HMM の状態分割を繰り返し、AR-HMM パラメータの推定だけでなく、HMM の最適なトポロジの自動生成も同時に行うアルゴリズムを構築する。

### (2) 病的音声の検知

トポロジ自動生成 AR-HMM 分析で得た結果から声帯疾患を伴った病的音声の検知を実現するために、以下の方法を考える。健常者音声は声帯の閉鎖期における呼気漏れは殆どないが、声帯疾患のある病的音声では声帯の閉鎖期において呼気漏れが生じ、これにより閉鎖期の声帯音源により多くの雑音成分が発生すると仮定する。一方、AR-HMM 分析で得られる HMM の各出力分布の分散は、推定した声帯音源の雑音成分の強弱を反映する。先の仮定に基づけば、AR-HMM 分析結果から得た HMM の出力分布の中で、声帯の閉鎖期に該当する出力分布の分散に着目すれば、声帯疾患の有無を検知できる可能性がある。これを確かめるために、病的音声だけでなく健常者音声もトポロジ自動生成 AR-HMM に基づいた分析実験を行い、得られた HMM の声帯閉鎖期に該当する出力分布の分散を比較する。もし明らかな違いが見出されれば、それを手掛かりとする病的音声の検知手法を構築する。

### (3) 嚔声の声質評価

声帯疾患に基づく病的音声の検知だけでなく、言語聴覚士が主観的に行っている病的音声の聴覚印象評価を客観的に行う音声分析評価手法の構築も検討する。具体的には、嚔声の4性質(粗さう性、気息性、無力性、努力性)のそれぞれについて、AR-HMM 分析結果に基づいて聴覚的印象を識別する手法を構築する。これら4性質を識別するためには、AR-HMM 分析で得られる AR 過程の調音フィルタよりも声帯音源に関する情報の方が重要である。AR-HMM 分析では声帯音源を確率統計処理したモデルである HMM で表現するが、4性質の識別においてはよりリッチな情報を含んでいると思われる声帯音源そのものをベースとすべきと考える。本研究課題では、AR-HMM 分析で得られる声帯音源から求めた局所自己相関係数を入力とし、言語聴覚士が行った4性質の識別結果をターゲット出力として、ニューラルネットワークを学習することで聴覚印象評価の自動化を試みる。

## 4. 研究成果

### (1) トポロジ自動生成 AR-HMM 分析

構築したアルゴリズムは、HMM の状態数が1と2のトポロジから開始して、それぞれのトポロジの AR-HMM パラメータを推定する。そして、確率統計モデルの評価指標である MDL に基づき、良いモデルを選択する。状態数が1のモデルが良い場合はそこで終了する。もし状態数が2のモデルの方が良いと判断された場合は、尤度に基づいて選択したどちらか一方の状態を、並列分割と直列分割の2種

類の方法で分割する。これにより状態数が3のトポロジが2種類得られ、それぞれについてAR-HMMパラメータを再推定する。そしてMDLに基づいて良いモデルを選択することで、状態数が3のトポロジが決定する。更に、先の状態数が2のモデルとMDLに基づいた比較を行い、状態数が2のモデルが良い場合はそこで終了する。もし状態数が3のモデルの方が良いと判断されたら、尤度に基づき選択した状態を分割して状態数が4のモデルを検討する。これを最適なモデルが選択されるまで繰り返す。構築したアルゴリズムの妥当性を評価するために、合成音声を用いた実験を行った。合成音声の音源が非定常・非周期的になるように、喉頭全摘出者の代替発声法の1つである食道発声音源から音源を抽出した。そしてその音源を、健常者音声から推定した調音フィルタに通すことで音声合成を行った。この合成音声に対してトポロジ自動生成AR-HMM分析を行い、推定した調音フィルタ特性と合成に用いた調音フィルタ特性の間の距離を求めることで、提案法の分析精度を評価した。図1にその結果を示す。提案法のトポロジ自動生成AR-HMM分析(青)の他、比較のために従来の線形予測法(緑)とリングトポロジ固定のAR-HMM分析(赤)の結果も示す。これからトポロジ自動生成AR-HMM分析の誤差が最小であることを確認した。

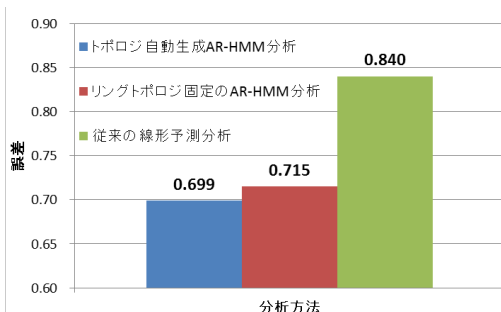


図1 分析精度の評価結果

### (2) 病的音声の検知

AR-HMM分析の結果から、声門が最大に開いた状態から閉じるまでの閉鎖期に該当する状態が、最大分散となる傾向が見られた。閉鎖期に該当する状態を含めた他の状態の分散を最大分散で正規化することで、他の音声の分析結果と比較できるようにする。病的音声と健常者音声をAR-HMMで分析し、それぞれ最大分散で正規化した正規化分散(最大分散を除く)の生起確率分布を図2に示す。病的音声と健常者音声を状態数2のAR-HMMで分析した結果を、それぞれ赤と青の実線で示す。健常者音声の場合は、最大分散に対して他の分散は0に近い領域に分布している。これに対し病的音声の場合は、最大分散の半値あたりをピークに比較的広範囲に分布している。図中、点線は状態数が10のAR-HMMで分析した結果を示す。状態数を増やすと、声帯が開いている開放期に該当する状態の分割も進

むため、特に健常者音声の正規化分散の生起確率分布が最大分散に近い領域へ広がり、病的音声の分布との重なりが大きくなる。病的音声と健常者音声の識別において、この正規化分散を特徴量として使用する場合は、病的音声と健常者音声でその生起確率分布が、できるだけ分離している方が有利となる。状態数2のAR-HMM分析で得られる正規化分散は、病的音声検知の手掛かりとして有望である。

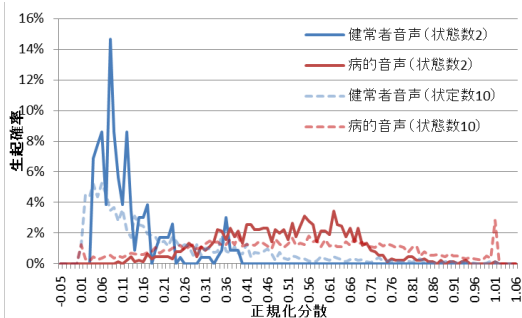


図2 正規化分散の生起確率分布

健常者音声と病的音声からAR-HMM分析により求めた正規化分散の生起確率分布を、それぞれ対数正規分布と正規分布でモデル化し、観測された正規化分散に対して尤度の高いモデルを選択することで、病的音声と健常者音声を識別する手法を構築した。そして、状態数が2から10までの各AR-HMMを用いて識別実験を行った結果を図3に示す。図中、赤(青)の実線が、病的音声(健常者音声)を正しく病的音声(健常者音声)と識別した割合を示す。また、緑の実線はその平均を示す。状態数が2の場合に、最大識別率が得られることを確認した。

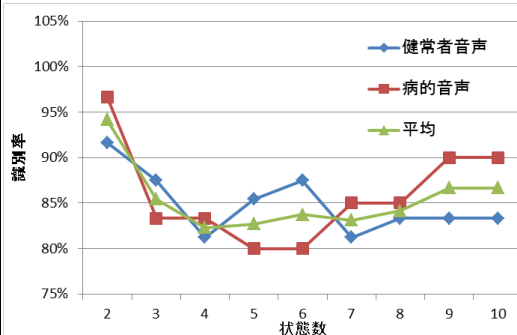


図3 AR-HMMの状態数に対する識別率

### (3) 嚔声の声質評価

声質評価は、200msのフレーム毎にAR-HMM分析により音声から声帯音源を抽出し、その高次局所自己相関値をニューラルネットワークに入力することで嚔声4性質を識別する短時間評価と、対象話者の全発話から得た全短時間評価を総合して評価する総合評価の2段階評価により実施する。短時間評価で用いる嚔声4性質の識別器は性質毎に用意し、それぞれ最適な構成を実験的に求めた。総合評価では、対象話者から得た全短時間評価の中で、それぞれの性質について陽性と識別された短時間評価の割合が設定閾値以上とな

つたら総合評価において陽性と識別した。評価実験には病的音声のみを使用し、各性質について陽性な病的音声を正しく陽性と識別した異常標本精度と、陰性な病的音声を正しく陰性と識別した正常標本精度を、各設定閾値について求めた。そして、異常標本精度と正常標本精度が一致する分岐点精度を、各識別器の性能評価指標として求めた結果を図4に示す。嘔声4性質の平均として88%の識別精度が得られることを確認した。

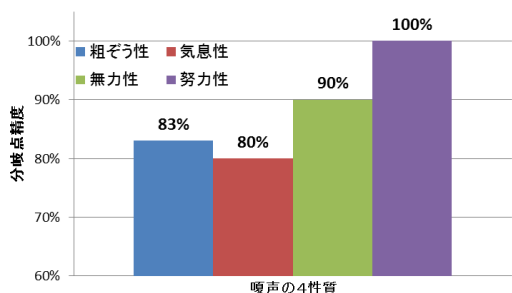


図4 嘔声4性質の識別精度

## 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)

〔雑誌論文〕(計 件)

〔学会発表〕(計 4件)

Akira Sasou, "Voice-pathology analysis based on AR-HMM," Proc. of Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2016), Dec. 2016.

Akira Sasou, "An evaluation of the automatically topology-generated auto-regressive hidden Markov model with regard to an esophageal voice enhancement task," Proc. of 5th joint meeting of the acoustical society of America and acoustical society of Japan, Nov. 2016.

佐宗 晃, "駆動源 HMM の出力分布に基づいた病的音声識別," 日本音響学会 2015 年春季研究発表会 ,1-R-3, Mar. 2015.

Akira Sasou, "Accuracy Evaluation of Esophageal Voice Analysis Based on Automatic Topology Generated-Voicing Source HMM," Proc of Interspeech 2014, Sep. 2014

〔図書〕(計 件)

〔産業財産権〕

出願状況(計 件)

名称：  
発明者：  
権利者：  
種類：  
番号：  
出願年月日：  
国内外の別：

取得状況(計 1件)

名称：AR-HMM のトポロジ自動生成  
発明者：佐宗晃  
権利者：国立研究開発法人産業技術総合研究所  
種類：特許  
番号：特許第 6078872 号  
取得年月日：平成 29 年 1 月 27 日  
国内外の別：国内

〔その他〕  
ホームページ等

## 6. 研究組織

### (1) 研究代表者

佐宗晃 (SASOU AKIRA)  
国立研究開発法人産業技術総合研究所・知能システム研究部門・主任研究員  
研究者番号：50318169

### (2) 研究分担者

( )

研究者番号：

### (3) 連携研究者

( )

研究者番号：

### (4) 研究協力者

( )