

氏名	かたぎり しげる 片桐 滋
授与学位	工学博士
学位授与年月日	昭和57年3月25日
学位授与の根拠法規	学位規則第5条第1項
研究科, 専攻の名称	東北大学大学院工学研究科 (博士課程)情報工学専攻
学位論文題目	短区間線形予測分析法とその音声分析への応用に 関する研究
指導教官	東北大学教授 城戸 健一
論文審査委員	東北大学教授 城戸 健一 東北大学教授 木村 正行 東北大学教授 星子 幸男 東北大学教授 鈴木 篤

## 論文内容要旨

### 概要

本論文は、母音型声道伝達特性の“正確な”推定法の確立を目指した研究に関するものである。ここで“正確な”という表現は、①声帯音源の周期性に影響されないことと、②声帯音源の周波数特性に影響されないこと、③時間分解能が高いことの3点を意味している。対象を母音型音声に限定してはいるが、その声道伝達特性が“正確に”推定されることの意義は大きい。これまで観測が困難であった、女声の声帯音源波や有声破裂音の声道伝達特性の時間変化が明らかになり、こうした知見は音声合成や音声認識の品質向上にもつながる為である。

本論文の内容は、各章毎に以下に概説する。各章の検討を経て、従来の手法の比較及び体系的整理を行ない、さらに新手法をも提案する。線形予測定式化の一つである改良共分散法や時間窓位置の決定法であるインパルス入力モデルは、本論文で提案する手法であり、夫々の有効性も確かめられる。本論文の最後では、所期の目的を満足する母音型声道伝達特性推定法が確立される。

### 第1章 序 論

本論文の概要を述べる章であり、線形予測の理論や音声研究の現状について概説し、問題点の指摘と本論文の研究目的の提示とを行なっている。

なお母音型声道伝達特性推定に線形予測を用いるのは、生成モデルの声道フィルタと分析モデル

の声道フィルタと分析モデルの予測フィルタとが、いずれも全極型である点で基本的に一致し、しかも処理が高速であり、時間分解能も高い為である。

## 第2章 線形予測による母音型声道特性推定の基礎

本章では、母音型声道伝達特性推定における従来の手法の問題点を具体的に考察し、本論文における研究の方向付けを行ない、次章以下の準備としている。

自然音声の持つ不規則性や声門閉鎖の不完全さを考慮した、本章で提起する“より自然な”生成モデルを用いて考えると、従来用いられてきた確定的で声門閉鎖も完全である“理想的な”生成モデルに関する以下の諸項目は成立たなくなる。①線形予測定式化の一つである共分散法は生成モデルと一致する。②予測次数の決定は容易である。③プリアンフェサスは不要である。④時間窓は十分長い声門閉鎖区間に合わせれば良い。本論文では、上記の項目が一般の実音声では成立ち難いことを指摘し、現実における分析条件の適切な設定を試みるのである。

## 第3章 種々の線形予測定式化の理論的関連

ここでは線形予測定式化の理論的体系化を行なっている。定式化に関する従来の検討は、分析合成系のフィルタを構成する立場によるものや統計的性質の比較を行なったものがあるにすぎず、短区間分析に用いる為の実用的検討は無い。本章は、各定式化における時間窓長の影響などの実用上の特徴を明らかにするものであり、母音型声道伝達特性推定に用いるべき定式化の選択にとどまらず、夫々の定式化の効果的利用にも役立つ。なおここで扱っている定式化は、共分散法と自己相関法、Burg法、PARCOR法、共分散ラティス法、不偏自己相関法である。

検討の結果、これらは時間窓内で仮定される定常性の程度によって区別されることがわかった。各定式化を大別して分類すると次の様になる。①自己相関法とBurg法、PARCOR法、共分散ラティス法は、形式的には定常性の仮定を満足し、予測フィルタも安定である。これらは、定常過程の時系列の分析に用いるべきである。②不偏自己相関法は、Toeplitz型の自己相関行列を持つが定常性の仮定を満たさない。また予測フィルタも安定とは限らない。この為、本定式化は実用的でない。③共分散法は、自己相関行列が非負性を満足するだけであり、明らかに非定常な解法である。本定式化は、特に過渡的不規則過程（非定常）の分析に適している。

さらにこれらの定式化を音声分析に用いる時は、次の様に使い分けるべきである。共分散法は母音型声道伝達特性推定に、自己相関法とBurg法、PARCOR法、共分散ラティス法は分析合成系に用いるのが良いことがわかった。

本章ではまた、改良共分散法という新しい定式化を提案している。これは、過渡的不規則過程の非常に短い時系列を分析する時（例：高ピッチ音声のピッチ内分析）共分散法では不十分な為、この対策として提案したものである。

## 第4章 音声の線形予測における分析条件の実験的検討

本章では、自然性をも考慮した合成音声（“より自然な”生成モデルによる音声等）の分析実験

により、①プリエンファシスと、②予測次数、③時間窓の長さや位置、④線形予測定式化の各項目毎に適切な設定を試みる。なお①と②、③に関する検討は、定式化に共分散法を用いて行なっている。ここで共分散法を用いるのは、3章においてこれが理論的に適切であると予想される為である。

実験の結果、次の点が明らかになった。①短区間分析であってもプリエンファシスの効果が認められ、しかも1次差分による固定的（差分係数が一定）なものが良い。②予測次数は、予想されるホルマント数の2倍強であれば良い。③時間窓長は1周期より短い（ピッチ内分析）のものが良い。しかし必要以上に短くすべきではなく、声帯音源周期の0.5倍程度が良い。また声道伝達特性の推定誤差は、時間窓内における声門閉鎖時点の有無に強く影響され、正確な声道伝達特性を得る為には声門閉鎖時点を避けて時間窓を設定すべきである。④線形予測定式化の理論的検討結果（3章）はここで実験的にも確認された。母音型声道伝達特性推定には共分散法が適している。また3章で提案した改良共分散法の、非常に短い時間窓を用いる場合の有効性も確かめた。

なお、母音型声道伝達特性推定の正確さを高める為には、時間窓の位置に関するより詳細な検討をすべきであり、これが5章の論点となる。

## 第5章 母音型声道特性推定における時間窓位置の決定法

本章では、4章で指摘される時間窓位置の最適な設定について、より詳細に検討している。4章によれば、声道伝達特性の推定誤差は、時間窓内における声門閉鎖時点の有無に強く影響される。しかし声帯音源波の観測は不可能な為、時間窓位置を僅かずつずらせながら分析した結果をなんらかの指標によって評価し、推定誤差の小さいと思われる分析結果を選択することによって、時間窓位置を決定する。なお、音声波形から声門閉鎖時点を予想することは困難であり、しかも予め分析位置を限定してしまうことの危険度が大きい為、予め適切と思われる位置のみで分析を実行する方法は避けた。

指標としては、推定誤差の小さい分析結果を容易に選択できるものが望まれる。そこで時間窓位置決定法は、次の2種類の実験によって検討した。①まず推定誤差と、従来提案されていたものを含む種々の決定法で用いられる指標との相関を検討した。時間窓の移動に伴う推定誤差と指標との時間変化パターンが一致すれば（両者の相関係数が1であれば）、指標を調べることで推定誤差は正確に把握できる為である。②また上述の指標の値を用いて、推定誤差が小さいと思われる分析結果のみを選択し、その推定誤差を検討した。これは、実用上持つべき性質に特に注目した検討法である。

検討は、まず従来から提案されてきた決定法（予測フィルタの安定性と予測残差2乗和、自己相関行列の行列式、白色雑音入力モデルを用いるもの）の比較から行なっている。これらに関する従来の検討が、“理想的な”生成モデルによる合成音声の分析実験や自然音声分析例による定性的説明にとどまっており、各決定法の優劣も明らかでない為である。なお本章の検討で、“より自然な”生成モデルを用いるのはいうまでもない。

検討の結果、従来の決定法の中では白色雑音入力モデルを用いるものが優れていることがわかった。しかしこの決定法も、高ピッチ音声の分析において性能の劣化が見られる為、実用的には満足

できない。“正確な”母音型声道伝達特性推定法は、高ピッチ音声の分析に用いる時にこそ効果的であるべきことを考えると、さらに決定法の改善が望まれる。

本章では、決定法の改善を目指した新しい提案も行なう。4章によれば、推定誤差が小さい分析結果を見出す為には、声門閉鎖時点と時間窓との位置関係に注目する次の2方法を提案している。すなわち、定常性を利用するものとインパルス入力モデルを利用するものである。前者は、時間窓の前後いずれの側に声門閉鎖時点があるかを、前向及び後向予測残差2乗和を用いた指標で予想しようとするものである。また後者は、声門閉鎖時点にインパルスモデルをあてはめ、これを利用して声道フィルタ係数の推定誤差を求めるものである。

本章で提案したこれらの決定法の優劣についても、従来のものと同様の評価を行ない、その結果本論文中の全ての決定法の中で、インパルス入力モデルが最も有効でありしかも実用に十分耐え得ることがわかった。ここでインパルス入力モデルの特徴をまとめておく。①白色雑音入力モデル同様、声道フィルタ係数と予測フィルタ係数との2乗誤差を指標として求める。この為、指標と推定誤差との関係が明瞭である。②白色雑音入力モデルが時間領域で白色雑音入力の仮定を行なうのに対し、ここでは周波数領域でインパルス入力の仮定を行なう。③インパルス位置は予測フィルタの逆フィルタリングにより、インパルス振幅(ここで、インパルスとは $\delta$ 関数ではなく、ある1点でのみ非零の有限値をとるものを指す)は声道フィルタの入出力時系列のエネルギー関係式より求める。

本章におけるインパルス入力モデルの有効性と4章で設定された分析条件とによって、本論文の研究目的である母音型声道伝達特性推定法が確立された。そこで最後に、この確立された推定法を用いて自然音声(女声)を分析し、本推定法が実際に有用な知見をもたらすことを示している。

## 第6章 結 論

本章では、本論文のまとめと残された問題点の指摘とを行なっている。本論文で確立された母音型声道伝達特性推定法が今後の音声研究に有用であることは、前章までに明らかにされている。しかし、その有用性をさらに高める為に関連技術の改善として、プリアンファシスやピッチ抽出法などの技術の向上が期待される。

## 審査結果の要旨

線形予測分析法が音声の分析合成系に応用された結果、音声の情報圧縮技術が画期的に進展したのみならず、音声生成の音響学的過程を、音声分析することによって推定する能率のよい方法が生み出された。この推定法の基盤となる全極形モデルは、声帯の振動によって励振され、かつ、鼻音化を伴わない母音形音声によく適合するが、その場合でも、女声のような声帯振動の周波数の高い音声（高ピッチ音声）からの声道伝達特性の推定は困難な問題であった。本論文は、短区間の音声を対象とする線形予測分析の手法を確立し上述の問題の解決に資することを目的として行った研究の成果を述べたもので、全文6章からなる。

第1章は序論である。第2章では、母音形音声の声道伝達特性の線形予測分析法による推定のための音声生成モデルについて論じ、その数学的背景を述べて、本論文の基礎を明らかにしている。

第3章では、既に開発されている種々の線形予測分析法の間の理論的関連を体系づけ、各定式化の差異が分析区間内の定常性の仮定に基づいて生じていることを明らかにし、本研究で対象とする母音形声道伝達特性の正確な記述のためには共分散法が適していることを示している。さらに、非常に短い時系列の分析に適する改良共分散法を提案している。

第4章では前章の理論的検討の成果を、不規則性や声門閉鎖の不完全さなどの自然性を考慮した合成音声を用いて検証し、適切な分析条件の設定法を示すと共に、分析区間内に声門閉鎖が行われない場合にも良好な結果が得られることを示している。

第5章では、短区間分析に極めて重要な母音形音声の声道伝達特性推定における分析区間の決定について、定常性を利用する方法とインパルス入力モデルによる方法とを提案し、合成音声によって比較検討して、インパルス入力モデルによる方法がすぐれていることを明らかにしている。最後に、各章で得られた結果を総合して自然音声进行分析し、本論文で与えた分析法により、初めて、従来の分析法では知ることができなかった高ピッチ自然音声の過渡区間の特性が明らかにされることを示している。

第6章は結論である。

以上要するに、本論文は高ピッチ音声から声道伝達特性を推定することを目的として、音声波の短い区間の全極形線形予測分析に必要な条件を明らかにし、具体的な分析手法を確立して音声情報処理に新しい可能性を加えたもので、情報工学に寄与するところが少なくない。

よって、本論文は工学博士の学位論文として合格と認める。