

学位論文題名

音声の特徴抽出のための新しい ARMA モデルと
その適応的係数推定に関する研究

学位論文内容の要旨

音声情報処理の研究は古くから行われているが、近年のデジタル信号処理技術の急速な発展にともない、音声認識、合成、低ビット符合化などの研究が盛んに行われるようになってきている。しかし、音声認識の現状をみると、高性能の認識用 LSI も開発されているが、特定話者あるいは有限語彙というように機能が限定されたものとなっている。音声認識における認識精度を向上させるためには、音声の特徴量を対応する文字に変換する処理の確さはもちろんであるが、音声波形の特徴抽出を一層高度にすることも重要である。音声の重要な特徴の一つとして音韻情報があるが、これは短時間スペクトルの時間変化パターンでとらえられる。従来から用いられている代表的なスペクトル推定手法としては、LPC 分析法があるが、LPC 分析では音声の生成モデルとして AR モデルを用いている。AR モデルは周波数領域でのホルマントを精度良く近似できるが、鼻子音のようにアンチホルマントが存在する音声の場合には推定精度に限界がある。この問題を解決するため、ホルマントとアンチホルマントの同時表現が可能である ARMA モデルを音声生成モデルとして用いる手法も提案されている。しかしながら、ARMA モデルを用いた場合でも、音声信号は時間と共にその特徴が変化する時変信号であるため、分析手法の時間分解能を考慮すれば、子音から母音へと移行するトランジェントな部分での特徴表現は難しい。

本論文は、従来の ARMA モデルでは表現することが難しかった子音から母音（または、母音から子音）へと移行するトランジェントな部分の特徴表現を考え、有声音声と無声音声とで異なる生成モデルを持つ ARMA モデルを考え、このモデルに基づいた適応的係数推定法を提案し、それらの有効性を実験により検証したものである。

本論文ではまず、上述の新しい ARMA モデル (ARMA-D モデル) を定義し、従来の手法を用いてこのモデルの係数を同定する手法について述べている。次に、従来の適応的 ARMA 係数推定手法を拡張し、より時間分解能がよい適応的係数推定手法を構築し、この手法により精度の高い ARMA-D モデルの係数を得ることに成功している。さらに、デジタル計算機上の有限語長の影響を受けにくくするため、この適応的 ARMA-D 係数推定アルゴリズムを複素数に拡張し、複素信号の一種である解析信号処理による分析精度の向上も達成している。

本論文は全 6 章からなっており、各章を要約すると以下ようになる。

第 1 章は全体の序論である。第 2 章では、第 3 章以降で述べられる音声分析法の基礎理論となっている適応的 ARMA 係数推定法が紹介され、その特徴が簡単に述べられる。すなわち、モデル同定システム (MIS 法) と指数重み係数を導入した回帰的最小 2 乗法 (RLS 法) による ARMA 係数推定アルゴリズムが紹介され、それぞれの手法が現時刻の信号に対してどのような影響を過去の信号に与えて、時変処理を行っているかが簡単に説明される。

第3章では、最初に ARMA-D モデルの定義を与え、従来の音声分析において用いられている代表的な確率モデルとの違いを説明する。次に、第2章で紹介した手法を用いて ARMA-D モデルの係数を適応的に同定する手法について述べる。また、ARMA-D モデルのパワースペクトルの計算法についても述べられる。一般に、ARMA モデルを用いて信号の推定スペクトルを表現するときには、モデルの伝達関数の対数パワースペクトルを求めることによって信号の推定パワースペクトルを表現する。しかしながら、ARMA-D モデルにおいては伝達関数が存在しないので、ARMA モデルのときのように推定パワースペクトルを計算できない。ここでは、ARMA-D モデルのパワースペクトルを2種類のパワースペクトルの利得付き和で表現する。そして、合成音声と実音声に対する実験を行い、提案するモデルの有効性を示す。まず、ARMA-D モデルから生成される合成信号に対して分析実験を行い、推定モデルとしての ARMA-D モデルの次数に関する考察を行う。次に、実音声に対する分析実験を行い、提案するモデルが有声音から無声音へと移行変わるトランジェントな部分で ARMA モデルより精度のよいスペクトルを推定できることを示す。

第4章では、非定常 ARMA 係数推定アルゴリズムとして提案されている MIS 法を拡張する。最初に、ここで提案する手法の概要を説明する。音声信号には、連続母音中の緩やかな変化に代表される時変量（連続時変量）と、音素の変わり目に代表される時変量（不連続時変量）が存在すると考えられる。ここで述べる時変係数推定アルゴリズムは、それら二つの評価を同時に行ない、信号生成モデルが連続的に変化する場合でも急激に不連続に変化する場合でも対応できるアルゴリズムである。具体的には、MIS 法のイノベーション記述に基づいたアルゴリズムを拡張することで RLS 法の指数重み係数に対応する連続時変量を導入し、更に、MIS 法の時変量である不連続時変量の認定規準を新しくしている。次に、上述の方針を数式上で議論し、具体的にアルゴリズムを構築する。そして、合成信号と実音声に対する分析実験を行ない、本手法が RLS 法や MIS 法と比較して、より正確なスペクトルを与えることを示し、本手法の有効性を検証している。

第5章では、第4章で述べた手法を複素信号が扱えるように拡張している。実際の音声信号処理システムはデジタル計算機上に実現されるため、有限語長による制限を受ける。この有限語長による影響を受けにくくする手法として、複素信号の一種である解析信号を用いる手法がある。最初に、解析信号とそれに間引きを施した信号の特徴について述べる。すなわち、解析信号に間引きを施した複素信号に基づく処理の有効性について理論的に論じる。本章で提案する複素アルゴリズムは、前処理を導入することで推定音声スペクトルの低・高域での誤差を減少させている。そして、間引かれた信号も利用できるようにアルゴリズムを修正することで、信号の間引きによって生じる適応的アルゴリズムのトラッキング能力の悪化という問題を改善している。最後に、合成音声と実音声に対して実験を行い、本章で提案する手法は第4章で述べている手法より精度の高いスペクトルを推定できることを示す。

第6章は結論として、本論文で述べた研究を総括し、今後の課題を述べる。

本研究の成果は以下のように要約される。まず、音声の特徴抽出のための新しい ARMA モデル (ARMA-D モデル) の有効性が示された。さらに、従来の適応的 ARMA 係数推定手法を拡張することで、より時間追従性のよい適応的 ARMA (ARMA-D) 係数推定手法が構築され、その手法が音声信号の時間的に変動する特徴の抽出に有効であることが示された。そして、この係数推定手法をデジタル計算機上での有限語長の影響を受けにくくするため複素信号が扱えるように拡張した場合、実数の手法に比べより正確な係数を推定できることが示された。

以上より本研究によって、音声の特徴抽出のための新しい ARMA モデル (ARMA-D モデル) に基づく音声分析手法の有効性が示された。

学位論文審査の要旨

主査 教授 栃内 香次
副査 教授 永井 信夫
副査 教授 新保 勝
副査 助教授 宮永 喜一

学位論文題名

音声の特徴抽出のための新しい ARMA モデルと その適応的係数推定に関する研究

音声の情報処理において、音波形よりスペクトル情報を高精度に求めることは重要な問題であるが、時変なスペクトル推定は非常に難しいとされている。本研究は、音声の特徴抽出において従来の自己回帰移動平均モデル (ARMA モデル) では表現することが難しかった子音から母音 (または母音から子音) へと移行する部分の特徴表現のために、有声音声と無声音声とで異なる生成モデルを持つ新しい ARMA モデルを定義し、このモデルに対する適応的係数推定法を提案し、それらの有効性を実験により検証したものである。

本論文は全 6 章からなっており、各章を要約すると以下ようになる。

第 1 章は全体の諸論である。

第 2 章では、第 3 章以降で述べられる音声分析法の基礎理論となっている適応的 ARMA 係数推定法が紹介され、その特徴が簡単に述べられている。すなわち、モデル同定システム (MIS 法) と指数重み係数を導入した回帰的最小 2 乗法 (RLS 法) による ARMA 係数推定アルゴリズムが紹介された。

第 3 章では、まず ARMA-D モデルの定義を与え、従来の音声分析において用いられている代表的な確率モデルとの違いを説明した。次に、第 2 章で紹介した手法を拡張し、ARMA-D モデルの係数を適応的に同定する手法を開発した。また、ARMA-D モデルのパワースペクトルを 2 種類のパワースペクトルの利得付き和で表現することで、このモデルのパワースペクトルを計算法を明らかにした。さらに、合成音声と実音声に対する実験結果より、提案するモデルが有声音声から無声音声へと移行変わるトランジェントな部分で精度のよいスペクトルを推定できることを示した。

第 4 章では、時変 ARMA 係数推定アルゴリズムとして提案されている MIS 法を拡張した。音声信号には、連続母音中の緩やかな変化に代表される時変量 (連続時変量) と、音素の変わり目に代表される時変量 (不連続時変量) が存在すると考えられる。ここで述べられた時変係数推定アルゴリズムは、それら二つの評価を同時に行い、信号生成モデルが連続的に変化する場合でも急激に不連続に変化する場合でも対応できるアルゴリズムである。具体的には、MIS 法のイノベーション記述に基づいたアルゴリズムを拡張することで RLS 法の指数重み係数に対応する連続時変量を導入し、更に、MIS 法の時変量である不連続時変量の

認定規準を新しくした。さらに、合成信号と実音声に対する実験結果より、本手法が RLS 法や MIS 法と比較してより正確なスペクトルを与えることを示した。

第 5 章では、第 4 章で述べた手法を複素信号が扱えるように拡張した。実際の音声信号処理システムはデジタル計算機上に実現されるため、有限語長による制限を受ける。この有限語長による誤差を減少させる手法として、複素信号の一種である解析信号を用いる手法がある。本文ではまず、解析信号とそれに間引きを施した信号の特徴について述べた。すなわち、解析信号に間引きを施した複素信号に基づく処理の有効性について理論的に論じた。本章で提案している複素アルゴリズムは、前処理を導入することで推定音声スペクトルの低・高域での誤差を減少させている。そして、間引かれた信号も利用できるようにアルゴリズムを修正することで、信号の間引きによって生じる適応的アルゴリズムのトラッキング能力の悪化という問題を改善している。さらに、合成音声と実音声に対する実験結果より、本章で提案する複素化された手法は第 4 章で述べている実信号のみを扱う手法より高精度であることを示した。

第 6 章は結論として本論文で述べた研究を総括し、今後の課題を述べた。

以上のように本論文は、音声の特徴抽出のための新しい ARMA モデルとその適応的係数推定法に関して、音声処理の分野における有効性を理論的検討ならびに実験の両面にわたって示したもので、信号処理工学ならびに音声情報処理工学に寄与するところ大である。よって筆者は、博士（工学）の学位を授与される資格があるものと認める。