

## 音声情報処理における適応 ARMA 分析手法の応用に関する研究

### 学位論文内容の要旨

近年、音声情報処理の分野では、一括型 AR モデルに基づく従来の LPC 分析手法よりも高精度なスペクトル分析が可能な適応 ARMA 分析アルゴリズムがいくつか提案されている。これらの適応 ARMA 分析アルゴリズムは高精度音声分析が行えることが既に報告されているが、得られた高精度スペクトルを応用した音声認識システムや符号化・合成システムに関する研究は少ない。

本論文は、上述の高精度適応 ARMA スペクトル分析アルゴリズムを音声認識と音声符号化・合成の分野に応用し、実音声を対象とする音声処理システムを構築して、それらの有効性を実験により検証したものである。すなわち本論文は、音声情報処理分野への適応 ARMA 分析手法の総合的な応用について論じたものである。

本論文ではまず、認識・合成・符号化の基礎となる音声のスペクトル分析手法について、既存の適応 ARMA 分析アルゴリズムの問題点を改善するための、アルゴリズム選択型スペクトル分析システムを構築している。次に音声認識の分野において、認識の基礎となるセグメンテーション候補点検出システムを構築している。符号化・合成の分野では適応 ARMA 分析アルゴリズムを導入した CELP 符号化システムを構築し、種々の条件のもとで実験・評価を行っている。

論文は全 15 章からなり、音声分析・認識に関する第 1 部と、音声符号化・合成に関する第 2 部とから構成されている。

第 1 章は全体の序論である。第 1 部は、第 2 章から第 8 章までにより構成されている。第 2 章では、本論文全体で用いられる高精度適応 ARMA スペクトル推定手法とモデルについて述べる。すなわち重み付き RLS 法、MIS 法、改良適応 ARMA ラティスマデル、規格化 ARMA ラティスマデル、コンパクト ARMA ラティスマデルのそれぞれについて、アルゴリズムおよびモデルを表現する重要な式を示し、それらの手法の特徴について簡略に述べる。

第 3 章から第 7 章までは、従来よりさらに高精度な音声分析システムを構築することを目的とする。そのために従来理論的・数式的であった信号処理に、経験的・知識工学的な手法を導入する。有声音声と無声音声は別個のアルゴリズムで分析すべきであるという考えから、まず有声音声について、最適なアルゴリズムを検討する。次に有声音声と無声音声の分離を行い、有聲・無聲それぞれに最適なアルゴリズムを選択して分析する新しい音声分析システムを構築する。

第3章では、有声音声に対し理論的に高精度な直接的分析法を提案する。この直接的分析法は、相関操作を全く用いずに原信号から直接スペクトルを求める信号処理手法である。モデル実験の結果より、直接的分析法は理論的には高精度であるが、量子化雑音等のわずかな雑音にもきわめて敏感であることが示される。

第4章では、直接的分析法よりも実際の有声音声分析に適する手法として重み付き RLS 法を取り上げ、その中で用いられている重み係数 $\lambda$ の周波数特性を新しく明らかにする。 $\lambda$ の周波数特性から、重み付き RLS 法が有声音声に適した分析法であることが確認される。

第5章では、重み付き RLS 法において分析結果の選択が必要であることを述べ、if-then ルールを用いた選択システムと、ニューラルネットのノードを用いた選択システムを構築し、それらのシステムによって良好なスペクトル選択が行えることを示す。

第6章では、有声音声と無声音声に別個の分析アルゴリズムを適用するために、有声音声・無声音声の分離を行う。分離には、1次の適応 AR 係数が用いられ、分離点の決定には、経験的ルールを用いる知識工学的手法が用いられる。実音声による実験結果を示し、提案する手法が有効であることが示される。

第7章では、第2章から第6章までで述べた手法を統合することにより、有声音声・無声音声それぞれに適したアルゴリズムで高精度な適応 ARMA スペクトル分析を行う、新しい音声スペクトル分析システムを構築する。

続く第8章では、高精度適応 ARMA スペクトルを用いた音声認識への第一歩として、第6章で述べた有声音声・無声音声分離システムの出力結果と、MIS 法のスペクトル変化量を用いて、連続音声のセグメンテーション候補検出システムを作成する。実音声による実験から、提案するセグメンテーション候補検出システムは音素の脱落の少ない良好な結果が得られることを示す。

第2部は、第9章から第14章までにより構成され、音声の符号化・合成について述べる。すなわち、従来の音声符号化システムのスペクトル推定フィルタ(短期予測フィルタ、STP)部に適応 ARMA スペクトル分析を導入し、より高精度なスペクトルを用いることによって、低ビット音声符号化・合成システムの性能向上をめざす。

第9章では、まず音声符号化システムにおける適応 ARMA 分析の必要性について述べる。すなわち、高精度なスペクトル推定によって残差信号の情報量が減少し、合成音声の品質が改善される可能性があることを述べる。また本論文では合成音声の品質評価法として、客観評価尺度の一つであるセグメンタル S/N 比を用いることを述べる。

第10章では、スペクトル推定フィルタの残差波形を単純に線形量子化した場合について述べる。実験により、適応 ARMA スペクトルを用いた場合、アンチフォルマントを含むと考えられる鼻子音等で、従来の一括 AR 分析手法である PARCOR を用いた場合より良好な結果が得られることを示す。このことは第9章で述べた、符号化・合成における適応 ARMA 分析の必要性の裏付けとなる。

第11章では、以下に続く各章で用いる CELP 符号化手法、長期予測フィルタ(LTP)、および聴覚重み付けフィルタ(PWF)について述べる。

第12章と第13章では、CELP符号化手法に適応ARMAスペクトル分析アルゴリズムを導入した符号化システムを構築し、様々な条件下で行った多くの実験の詳細について述べる。すべての実験の中で、スペクトル推定にARモデルの重み付きRLS、符号化手法にClosed loop CELPを用いた符号化システムが最も良い結果を示し、従来のPARCORに基づくClosed loop CELPシステムと比較して、平均セグメンタルS/N比で0.5~1.5dBの音質改善が得られたことを述べる。また、スペクトル推定にARMAモデルを用いた場合、概して良好な結果が得られないことが示され、CELP符号化方式とMA部の関係からその原因について考察する。

第2部の最後となる第14章では、第13章で示されたMA部とCELPとの不整合を避けるため、高精度ARMAスペクトルから抽出したフォルマント情報を用いるCELP符号化システムを構築し、実験によりその有効性を検討する。

第15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。まず音声の分析において、信号の性質に応じて最適な適応ARMA分析アルゴリズムを選択する分析システムの有効性が示された。認識の分野において、適応ARMA分析によって得られたスペクトルが、連続音声のセグメンテーションに有効であることを示した。符号化の分野では、ARモデルに基づく適応スペクトル推定アルゴリズムを用いたCELP符号化システムが、音声符号化に有効であることを示した。

以上より本研究によって、適応ARMA分析手法の音声情報処理における有効性が示された。

# 学位論文審査の要旨

主査	教授	栃内	香次
副査	教授	小川	吉彦
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

### 音声情報処理における適応 ARMA 分析手法の応用に関する研究

本論文では、近年提案された高精度適応 ARMA 分析アルゴリズムを、音声認識と音声符号化・合成の分野に応用し、実音声を対象とする音声処理システムを構築して、それらの有効性を実験により検証している。すなわち本論文は、従来十分な研究がなされていなかった、適応 ARMA 分析手法の音声情報処理分野への総合的な応用について論じたものである。

論文は全 15 章からなり、音声分析・認識に関する第 1 部と、音声符号化・合成に関する第 2 部とから構成されている。

第 1 章は全体の序論である。第 1 部は、第 2 章から第 8 章までにより構成されている。第 2 章では、本論文全体で用いられる高精度適応 ARMA スペクトル推定手法について、アルゴリズムを表現する重要な式を示し、それらの手法の特徴について簡略に述べている。

第 3 章では、有声音声を対象とした直接的分析法を提案している。この直接的分析法は、関連操作を全く用いずに原信号から直接スペクトルを求める信号処理手法である。モデル実験の結果より、直接的分析法は理論的には高精度であるが、量子化雑音等のわずかな雑音にもきわめて敏感であることが示されている。

第 4 章では、直接的分析法よりも実際の有声音声分析に適する手法として重み付き RLS 法を取り上げ、その中で用いられている重み係数 $\lambda$ の周波数特性を新しく明らかにしている。 $\lambda$ の周波数特性より、重み付き RLS 法が有声音声に適した分析法であることを確認している。

第 5 章では、重み付き RLS 法において分析結果の選択が必要であることを述べ、ニューラルネットワークのノードを用いた選択システムによって、良好なスペクトル選択を行っている。

第 6 章では、有声音声と無声音声に別個の分析アルゴリズムを適用するために、有声音声・無声音声の分離を行っている。分離には、1 次の適応 AR 係数が用いられ、分離点の決定には、経験的ルールを用いる知識工学的手法を用いている。実音声による実験により、提案する手法が有効であることが示されている。

第7章では、第2章から第6章までで述べた手法を統合することにより、有声・無声に応じてアルゴリズムを選択する適応 ARMA 分析システムを構築し、高精度なスペクトルを得ている。

続く第8章では、第6章で述べた有声・無声分離システムの出力結果と、MIS 法より得られるスペクトル変化量を用いて、連続音声のセグメンテーション候補検出システムを作成している。実音声による実験において、提案するシステムは、音素の脱落の少ない良好な結果を示している。

第2部は、第9章から第14章までにより構成され、適応 ARMA 分析を用いた音声の符号化・合成について述べている。第9章では、まず音声符号化システムにおける適応 ARMA 分析の必要性について述べている。

第10章では、スペクトル推定フィルタの残差波形を単純に線形量子化した場合について検討している。実験によると、適応 ARMA 分析手法を用いた場合、鼻音音等で従来の PARCOR 分析を用いた場合より良好な結果が得られている。

第11章では、以下の章で用いる CELP 符号化手法、長期予測フィルタ (LTP)、および聴覚重み付けフィルタ (PWF) について述べている。

第12章と第13章では、CELP 符号化手法に適応 ARMA 分析アルゴリズムを導入した符号化システムを構築し、様々な条件下で多くの実験を行っている。すべての実験の中で、スペクトル推定に AR モデルの重み付き RLS、符号化手法に Closed loop CELP を用いた符号化システムが最も良い結果を示し、従来の PARCOR に基づくシステムと比較して、平均セグメンタル S/N 比で 0.5~1.5dB の音質改善が得られている。また、スペクトル推定に ARMA モデルを用いた場合、CELP との不整合から、概して良好な結果が得られないことが示されている。

第14章では、第13章で示された MA 部と CELP との不整合を避けるため、フォルマント情報を用いる CELP 符号化システムを構築し、実験により有効性を検討している。

第15章は結論として、本論文で述べた研究を総括し、今後の課題を述べている。

本研究の成果は以下のように要約される。まず音声の分析において、信号の性質に応じて最適な適応 ARMA 分析アルゴリズムを選択する分析システムの有効性が示された。認識の分野において、適応 ARMA 分析によって得られたスペクトルが、連続音声のセグメンテーションに有効であることが示された。符号化の分野では、AR モデルに基づく適応スペクトル推定アルゴリズムを用いた CELP 符号化システムが、音声符号化に有効であることが示された。

以上のように著者は本論文において、適応 ARMA 分析手法の音声情報処理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、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。よって著者は、博士(工学)の学位を授与される資格あるものと認める。