

学 位 論 文 題 名

A Study of Noise-robust Speech Analysis Using
Running Spectrum Filtering

(ランニングスペクトルフィルタリングによる音声分析に関する研究)

学位論文内容の要旨

この数年、数多くの音声分析手法が提案されている。しかし、それらの多くは比較的静かな環境下で用いることを想定したものが大半を占めており、実環境で背景雑音の影響が大きい場合、音声信号の推定性能が劣化する問題がある。そのため、背景雑音にロバストな音声分析方法が必要とされている。

雑音にロバストな音声分析法を確立するためのアプローチとして、雑音が重畳した音声から雑音成分を取り除き、クリーンな音声を抽出して分析を行う方法が考えられる。

雑音除去の従来法としては、Spectral Subtraction (SS) 法が定常雑音の場合には有効な手法であり、広く用いられている。SS 法は雑音が重畳した音声のパワースペクトルから、雑音の推定パワースペクトルを減算することにより、音声のパワースペクトルを推定する手法である。そのため、雑音は定常であることを前提としており、重畳雑音のパワースペクトルを事前情報として推定する必要がある。従って、SS 法は非定常雑音に本質的には適用できない問題点が存在している。

そこで本論文では、ランニングスペクトルフィルタリング (Running Spectrum Filtering、以下 RSF) による音声低減手法を提案する。RSF 法は事前に雑音スペクトルの推定は行わず、ランニングスペクトル上で、バンドパスフィルタを用い、重要な音声の特徴を抽出することで、雑音を低減し、音声を強調する。つまり、非定常雑音のランニングスペクトルが音声と違うことにより、RSF 法は音声の特徴のみを抽出することができる。従って、RSF 法は定常雑音だけでなく、非定常雑音でも有効となる方法である。

本論文で用いた音声分析方法は、拡張回帰的二乗法 (Extended Recursive Least Mean Square、以下 ELS) である。従来の音声分析では、有限個パラメータモデルとして、AR モデル (Autoregressive Model)、もしくは、ARMA モデル (Autoregressive and Moving-average Model、ARMA モデル) が用いられている。AR モデルは、システムの出力信号のみからモデルの係数を推定できるため、音声信号のみを用いた分析に適している。また、AR モデルの伝達関数は共振点をもち、周波数領域でのエネルギーの集中を表現できる。これは、ホルマントで特徴づけられる母音の分析に適している。これらより、AR モデルは古くから音声の線形予測分析に用いられてきた。一方、ARMA モデルは、周波数領域でエネルギーの集中と消散を同時に表現できる。これは、ホルマントとアンチホルマントをもつ鼻音や摩擦音の分析に適している。すなわち、音声の鼻子音等に存在する零点や母

音などにおける放射特性を考慮する場合、音声生成系を極一零型モデル (ARMA モデル) とする必要がある。ELS 法は最小二乗推定による ARMA モデルを用いた音声スペクトルの適応的分析手法である。本論文で提案する手法は、RSF 後のスペクトルを用いて音声を合成し、ELS 法でより正確な推定スペクトルを求める手法である。

本論文は、全 6 章より構成される。

第 1 章では緒論。

第 2 章では人間の音声生成、知覚過程を述べる。

第 3 章では RSF 手法について述べる。

3.1 節では、RSF 手法を取り扱う準備として、短時間フーリエ変換について説明している。

3.2 節では従来法の SS 法について説明している。SS 法は事前に推定された雑音スペクトルを信号スペクトルから減ずることで、加法性の雑音成分を取り除くので、非定常雑音に効果がないことを示す。

3.3 節では RSF 雑音低減手法を提案する。まず、短時間フーリエ変換を用い、ランニングスペクトルを定義する。音声と雑音のランニングスペクトル上の時間変化の違いを利用し、雑音成分と音声成分が分離できることを示す。次に、短時間パワースペクトルの時間変化に対するスペクトルである変調スペクトル (Modulation Spectrum) を定義する。雑音信号の変調スペクトルは直流成分 (1 Hz 以下) にそのエネルギーが集中している。一方、音声信号の変調スペクトルは 1Hz 以上の低域に分布している。この特徴を利用し、バンドパスフィルタをランニングスペクトルの各周波数列にかけていくことで、加法性の雑音成分を大きく低減することができるのを示した。それに、比較的变化のゆるやかな非定常雑音の変調スペクトルも直流成分の近くにエネルギーが集中しているので、同様にバンドパスフィルタによるフィルタリングを行うと、雑音低減効果があることがわかる。

第 4 章では ELS 音声分析手法について述べる。ARMA モデルを用いた音声分析手法を取り扱う準備として、最小二乗法により推定される ARMA モデルの性質を、近似されるシステムとの関係から説明する。ここでは、確率過程モデルを ARMA モデルと考え、そのパラメータをある時刻で回帰的に求める最尤推定法を導く。

第 5 章ではロバスト音声分析アルゴリズムを示す。本章では、ARMA パラメータのスペクトル改善効果を評価する尺度を定義した。提案するスペクトル歪尺度は、音声のエネルギーが大きく、同時にスペクトル歪が大きい時、聴感上の特性が劣化することを考慮している。

第 6 章では提案手法と従来法による比較実験結果を示した。実験結果より提案手法の有効性を確かめた。特に非定常加法性雑音に対して、提案手法を用いることで、より正確な ARMA スペクトルを推定できることがわかる。

最後の章では本論文の結論について述べた。

学位論文審査の要旨

主査	教授	宮永喜一
副査	教授	野島俊雄
副査	教授	小柴正則
副査	教授	小川恭孝
副査	助教授	吉田則信

学位論文題名

A Study of Noise-robust Speech Analysis Using Running Spectrum Filtering

(ランニングスペクトルフィルタリングによる音声分析に関する研究)

雑音にロバストな音声分析の従来法として、Spectral Subtraction (SS) 法が提案されている。この手法は定常雑音の場合に有効な手法であり、様々な応用分野で広く用いられている。しかし、SS 法は非定常雑音に本質的には適用できない問題点が存在している。

本論文では、ランニングスペクトルフィルタリング (Running Spectrum Filtering、以下 RSF) による音声低減手法を提案している。RSF 法は事前に雑音スペクトルの推定は行わず、ランニングスペクトル上で、バンドパスフィルタを用い、重要な音声の特徴を抽出することで、雑音を低減し、音声を強調する。つまり、非定常雑音のランニングスペクトルが音声と違うことにより、RSF 法は音声の特徴のみを抽出することができる。従って、RSF 法は定常雑音だけでなく、非定常雑音でも有効な方法となっている。

また、本論文で用いた音声分析方法は、拡張回帰的二乗法 (Extended Recursive Least Mean Square、以下 ELS) である。ELS により推定される特徴パラメータは ARMA パラメータであり、これらは周波数領域でエネルギーの集中と消散を同時に表現できる。従って、ホルマントとアンチホルマントをもつ鼻音や摩擦音の分析に適している。すなわち、音声の鼻子音等に存在する零点や母音などにおける放射特性も含めた音声生成系の同定が可能となる。ELS 法は最小二乗推定による ARMA モデルを用いた音声スペクトルの適応的分析手法である。本論文で提案する手法は、RSF 後のスペクトルを用いて音声を合成し、ELS 法でより正確な推定スペクトルを求める手法である。

本論文は、全 6 章より構成されている。

第 1 章は緒論。

第 2 章は人間の音声生成、知覚過程を述べている。

第 3 章は RSF 手法について述べている。

3.1 節では、RSF 手法を取り扱う準備として、短時間フーリエ変換について説明している。

3.2 節では従来法の SS 法について説明している。SS 法は事前に推定された雑音スペクトルを信号スペクトルから減ずることで、加法性の雑音成分を取り除くので、非定常雑音に効果がないことを示している。

3.3 節では RSF 雑音低減手法を提案している。まず、短時間フーリエ変換を用い、ランニングスペクトルを定義する。音声と雑音のランニングスペクトル上の時間変化の違いを利用し、雑音成分と音声成分が分離できることを示している。次に、短時間パワースペクトルの時間変化に対するスペクトルである変調スペクトル (Modulation Spectrum) を定義する。雑音信号の変調スペクトルは直流成分 (1 Hz 以下) にそのエネルギーが集中している。一方、音声信号の変調スペクトルは 1 Hz 以上の低域に分布している。この特徴を利用し、バンドパスフィルタをランニングスペクトルの各周波数列にかけていくことで、加法性の雑音成分を大きく低減できることを示した。それに、比較的变化のゆるやかな非定常雑音の変調スペクトルも直流成分の近くにエネルギーが集中しているので、同様にバンドパスフィルタによるフィルタリングを行うと、雑音低減効果があることも述べている。

第 4 章では ELS 音声分析手法について述べている。ARMA モデルを用いた音声分析手法を取り扱う準備として、最小二乗法により推定される ARMA モデルの性質を、近似されるシステムとの関係から説明している。ここでは、確率過程モデルを ARMA モデルと考え、そのパラメータをある時刻で回帰的に求める最尤推定法を導いた。

第 5 章ではロバスト音声分析アルゴリズムを示している。本章では、ARMA パラメータのスペクトル改善効果を評価する尺度を定義した。提案するスペクトル歪尺度は、音声のエネルギーが大きく、同時にスペクトル歪が大きい時、聴感上の特性が劣化することを考慮している。

第 6 章では提案手法と従来法による比較実験結果を示している。実験結果より提案手法の有効性を確かめた。特に非定常加法性雑音に対して、提案手法を用いることで、より正確な ARMA スペクトルを推定できることがわかる。

最後の章では本論文の結論について述べている。

これを要するに、筆者は、雑音環境下での新しい音声分析を提案し、実音声信号に対してこれを応用しその有効性を示した。これにより、ロバスト音声分析手法に関する多くの有益な知見を得ており、音声情報科学の分野に貢献するところ大なるものがある。

よって筆者は、北海道大学博士(工学)の学位を授与される資格あるものと認める。