

学位論文題名

Speech dereverberation based on multi-channel linear prediction

(多チャンネル線形予測を用いる音声残響除去)

学位論文内容の要旨

話者から離れた位置にあるマイクロホンで収録した音声は、室内残響の影響を受けて歪んだ信号になる。この様な残響歪みは、音声の特性や明瞭性を劣化させ、自動音声認識や電話会議など、音声処理を用いるサービスの妨げになる。残響歪みを収録音声から取り除く方法の一つにブラインド残響除去方法がある。本論文では、多チャンネル線形予測を用いる新しい残響除去方法 LIME (Linear-predictive Multi-input Equalization) について述べる。線形予測を用いるだけでは、従来技術と同様に、残響歪みだけでなく音声特性まで等化してしまう、いわゆる、白色化問題を生じる。LIME の特徴は、音声特性を複数マイクロホンの出力信号に共通な特性として抽出し、抽出した特性を一旦白色化された音声に付与する事で、残響歪みの低減された原音声を回復する点にある。

前書きと用語説明 (第 1 章) に続き、第 2 章では上記音声特性の計算原理を述べる。音声特性は、多チャンネル線形予測を元に得られる射影行列の特性多項式に一致する事を示す。第 3 章では、室の残響時間が長くなると射影行列のサイズも大きくなり、特性多項式の計算精度が劣化する問題を取り上げ、計算方法の簡略化を検討する。残響時間 0.3 秒程度の室を用いたシミュレーション実験では、2つのマイクロホンを用いる方法で、20dB 以上の残響低減効果を得た。第 4 章では、より長い残響時間に対処するため、(3, 4 以上の) 多数マイクロホンを用いる方法を検討する。少数のマイクロホンを使う場合、各マイクロホンと音源との間の信号伝達特性の類似性は高くなり (すべての伝達関数に共通な非最小位相零点を生じ)、音声特性の推定精度は劣化する。この問題は、多数のマイクロホンを利用し、伝達特性のばらつきを大きくする事によって回避・軽減出来る。残響時間 0.5 秒程度の室を用いたシミュレーション実験では、4つ以上のマイクロホンを用いる事で、20dB 以上の残響低減効果を得た。第 5 章では、残響と雑音源のある環境へ LIME を適用する。単純に適用すると、雑音を低減する事は出来るが、音声特性の推定値に雑音特性が重畳し、周波数特性を持つ有色雑音の影響を取り除く事は出来ない。音声の含まれない雑音区間を検出し、雑音特性を求め、上記の雑音特性の重畳された特性から差し引く事により音声特性を求める処理を追加した。残響時間 0.2 秒、SN 比 0 dB の室を用いたシミュレーション実験では、20dB の残響低減効果と、18dB の雑音抑圧効果を得た。

本論文で述べた LIME は、(従来法では回避する事が難しかった) 音声の白色化歪みを起こすことなく、雑音源の存在する場合にも、遠隔収録された音声に含まれる残響歪みを高精度に取り除く方法である。残響や雑音のため適用が難しかった環境 (分野) へ、自動音声認識などの利用範囲を拡大する礎を成す計算原理と期待される。

学位論文審査の要旨

主査	助教授	三好正人
副査	教授	荒木健治
副査	教授	長谷川美紀
副査	教授	山本強
副査	教授	宮永喜一

学位論文題名

Speech dereverberation based on multi-channel linear prediction

(多チャンネル線形予測を用いる音声残響除去)

話者から離れた位置にあるマイクロホンで収録した音声は、室内残響の影響を受けて歪んだ信号になる。このような残響歪みは、音声の特徴や明瞭性を劣化させ、自動音声認識や電話会議など、音声処理を用いるサービスの妨げになる事が知られている。残響歪みを収録音声から取り除く方法の一つにブラインド残響除去方法がある。本論文は、多チャンネル線形予測を用いる新しい残響除去方法 LIME (Linear-predictive Multi-input Equalization) について述べている。線形予測を用いるだけでは、従来技術と同様に、残響歪みだけでなく音声特性まで等化してしまう、いわゆる、白色化問題を生じる。LIME の特徴は、音声特性を複数マイクロホンの出力信号に共通な特性として抽出し、抽出した特性を一旦白色化された音声に付与する事で、残響歪みの低減された原音声を回復する点にある。

前書きと用語説明 (第1章) に続き、第2章では上記音声特性の計算原理を述べている。音声特性は、多チャンネル線形予測を元に得られる射影行列の特性多項式に一致する事が示されている。第3章では、室の残響時間が長くなると射影行列のサイズも大きくなり、特性多項式の計算精度が劣化する問題を取り上げ、計算方法の簡略化を検討している。残響時間 0.3 秒程度の室を用いたシミュレーション実験では、2つのマイクロホンを用いる方法で、20dB 以上の残響低減効果を得ている。第4章では、より長い残響時間に対処するため、(3, 4 以上の) 多数マイクロホンを用いる方法を検討している。少数のマイクロホンを使う場合、各マイクロホンと音源との間の信号伝達特性の類似性は高くなり (すべての伝達関数に共通な零点を生じ)、音声特性の推定精度は劣化する。この問題は、多数のマイクロホンを利用し、伝達特性のばらつきを大きくする事によって回避・軽減されている。残響時間 0.5 秒程度の室を用いたシミュレーション実験では、4つ以上のマイクロホンを用いる事で、20dB 以上の残響低減効果が得られている。第5章では、残響と雑音源のある環境へ LIME を適用している。単純に適用すると、雑音を低減する事は出来るが、音声特性の推定値に雑音特性が重畳し、周波数特性を持つ有色雑音の影響を取り除く事は出来ない。この点に鑑み、音声の含まれない雑音区間を検出し、雑音特性を求め、上記の雑音特性の重畳された特性から差し引く事により音声特性を求める処理を追加している。残響時間 0.2 秒、SN 比 0 dB の室を用いたシミュレーション実験では、20dB

の残響低減効果と、18dBの雑音抑圧効果を得ている。

本論文で述べたLIMEは、(従来法では回避する事が難しかった)音声の白色化歪みを起こすこの残響低減効果と、18dBの雑音抑圧効果を得ている。

本論文で述べたLIMEは、(従来法では回避する事が難しかった)音声の白色化歪みを起こすことなく、雑音源の存在する場合にも、遠隔收音された音声に含まれる残響歪みを高精度に取り除く方法である。残響や雑音のため適用が難しかった環境(分野)へ、自動音声認識などの利用範囲を拡大する礎を成す計算原理と期待される。

これを要するに、著者は音声信号処理における残響除去手法に関する新知見を得たものであり、その効果は音質改善にとどまらず、自動音声認識の認識率改善をはじめとして多岐にわたるものである。よって著者は北海道大学博士(情報科学)の学位を授与される資格あるものと認める。