

論文内容の要旨

博士論文題目 Noise Reduction Front-End for Robust Speech
Recognition using Multi-Channel Signals and
Harmonic Structure

(マルチチャンネル信号と調波構造を利用したロバスト
音声認識のための雑音除去フロントエンド)

氏名 市川 治

(論文内容の要旨)

現在の機械音声認識システムは、音声信号のみの入力を仮定しているため、背景雑音の存在に大きく影響を受けてしまうことが知られている。しかし、従来の雑音抑圧手法は、いまだその性能が十分ではなく、また実環境雑音は多種多様であるため、単純に一種類の雑音抑圧手法を用いるだけでは現実システムへの対応は不十分である。本論文では、実環境における多様な雑音を高精度に抑圧するため、3種類の手法を新たに提案し、その有効性を音声認識実験等によって検証している。

1つ目の雑音抑圧手法は、プロファイルフィッティングに基づくマイクロフォンアレイ信号処理である。ここでは、従来技術のように雑音と音声の到来方向の違いを利用するが、到来する音声の方位別パワー分布を観測プロファイルとして導入する点に新規性がある。観測プロファイルを既知のテンプレートプロファイルに成分分解することにより、たとえ拡散性の背景雑音が存在しても目的方位の信号成分（音声成分）を高精度に抽出することが可能となった。また、本論文では、上記技術を音源位置推定として用いる場合についても言及している。プロファイルの利用により、音源位置推定の手がかりとなるチャンネル間位相差・強度差などを高精度に推定することができ、高度な音源位置推定が可能であることが実証された。

2つ目の提案手法は、SSEC (Simultaneous adaptation of spectral Subtraction and Echo Cancellation) アルゴリズムに基づく頑健なエコーキャンセラ信号処理である。一般に、自動車内において、カーオーディオからの音楽やカーナビ音声は雑音源である場合には、そ

れらを直接機器内部から得ることができ、それらを参照信号としたエコーキャンセラを構成することができる。しかし、従来のエコーキャンセラは、振幅と位相の両方に関する伝達関数を推定する必要があるため計算量負荷が重く、かつ背景雑音がある場合には上記推定精度が大きく劣化するという問題があった。本論文では、振幅スペクトルに関してのみの減算（スペクトルサブトラクション）に基づく頑健なエコーキャンセラを提案し、その係数の適応と定常雑音成分の推定とを同時に行うアルゴリズムを提案している。実験により、走行雑音とカーオーディオ音とが同時に発生しているという状況下でも、走行雑音成分とエコー成分の両方を効果的に除去可能であることが示された。

3つ目の提案手法は、Local Peak Enhancement に基づく音声強調処理である。これは、音声の有する調波構造を利用することにより、雑音に埋もれた音声成分を強調するものである。本技術に類する従来技術としては櫛型フィルタ等があったが、正確なピッチ周波数推定と有声・無声判定が必要で、高騒音環境下ではしばしばその精度が問題となった。本論文では、それらの情報を一切利用せず、観測パワースペクトルそのものから直接フィルタ係数を推定するアルゴリズムを提案した。音声認識実験においては、特に自動車内における高雑音環境下での雑音抑圧に関して大きな改善を示した。また、従来の標準的な前処理技術であるスペクトルサブトラクションなどと組み合わせることで、さらに大きな改善が得られることも実証された。

以上の理論提案およびその実験的検証により、様々な実雑音環境下におけるより高精度な雑音抑圧の実現およびその音声認識への応用が示された。

氏名	市川 治
----	------

(論文審査結果の要旨)

機械音声認識システムの研究開発においては、雑音の混入が大きな問題であるとされている。しかし、従来の雑音抑圧手法は、その性能が不十分であり、また実環境における雑音の多様性に対応することが困難であった。そこで本論文では、以下で述べる幾つかの新規理論提案を行い、その有効性を実環境データによって検証している。

(1) プロファイルフィッティングに基づくマイクロフォンアレイ信号処理：

到来音声の方位別パワー分布を観測プロファイルとし、これを既知のテンプレートプロファイルに成分分解することにより、高背景雑音下でも目的方位の信号成分（音声成分）を高精度に抽出することが可能となった。また、本技術を頑健な音源位置推定として用いることも可能であることが示された。

(2) SSEC アルゴリズムに基づく頑健なエコーキャンセラ：

一般に、自動車内における従来エコーキャンセラは、複素伝達関数における振幅と位相の両方を推定する必要があるため、計算量負荷の増大かつ背景雑音に敏感であるという欠点があった。本論文では、位相を無視した振幅スペクトル減算に基づくエコーキャンセラを構築し、その係数の適応推定と定常雑音成分の推定とを同時に行う手法を提案した。自動車内実験により、走行雑音とエコー成分の両方を効果的に除去できることが実証された。

(3) Local Peak Enhancement に基づく音声強調処理：

これは、音声の有する調波構造を利用することにより、雑音に埋もれた音声成分を強調するものである。本技術は、従来技術と異なり、ピッチ周波数推定と有声・無声判定が不要である。音声認識実験により、特に自動車内における高雑音環境下での雑音抑圧に関して非常に有効であることが示された。また、同時に、従来の標準的なフロントエンド処理との親和性も良いことが実証された。

以上の理論提案およびその実験的検証により、様々な実雑音環境下におけるより高精度な雑音抑圧の実現およびその音声認識への応用が示された。本研究成果は、4編の原著論文や多数の国際・国内会議にて発表され、本論文の音響・音声信号処理分野への貢献は極めて大きいといえる。以上より、平成20年7月29日に開催した公聴会の結果も参考にして、本博士論文の審査を行い、本論文は博士論文（工学）の学位論文として十分な価値があるものと判断した。