

審査の結果の要旨

氏名 大淵 康成

本論文では、音声認識機能を携帯機器に搭載することを目的として、その目的を達成するために扱うべき課題を明らかにし、それらに対する有効な解決手段を論じている。

第一章では、音声認識機能を携帯機器に搭載することの社会的価値を論じ、その実現形態として有効性が高いと思われるアプリケーションの種類と特性を論じている。その結果に基づき、既存の技術を活用すべき部分と新規技術を開発しなければならない部分とを明らかにしている。さらに、開発すべき新規技術の位置付けを明確化し、「コンパクトな音声認識プログラムの開発」、「低処理量の耐雑音特徴補償方式」、「複数マイクの活用による雑音環境下音声認識性能の向上」という三つの課題を解決することにより、音声認識機能を携帯機器に搭載するという目標を達成することができるという結論を導き出した。

第二章では、上記の三つの課題を解決するにあたって活用できる可能性のある従来技術を整理し、その長所と短所とを論じている。従来の音声認識研究において大きな成功を収めている隠れマルコフモデルに基づく方式の基本部分を踏襲することの意義を論じつつ、その枠組みの中で用いられる様々な手法を理論的に検討し、そのまま活用できるもの、本研究において新規に開発する際の土台となるもの、活用する必要を認めないものに分類している。これらの検討結果により、次章以降での論述にあたっての出発点を明確にすることができた。

第三章では、「コンパクトな音声認識プログラムの開発」という目標のもと、これが市場において受け入れられる水準を念頭に置き、「汎用マイクロプロセッサ用音声認識ミドルウェア」という形で具体化し、処理速度やメモリ使用量などの要求仕様を明確化した。その上で、対象となるプロセッサの特性に対して有効な高速化手法を複数導入した。それらの高速化手法の中でも特に重要であるサブベクトル量子化については、様々な条件下での処理速度と音声認識精度の関係を実験的に調査し、それらの結果を利用して多数のパラメータの値を最適化する方式を論じた。これらの結果により、与えられた処理速度およびメモリ使用量の目標を達成し、音声認識ミドルウェアとして市場に提供、さまざまな応用製品で活用されることとなった。

第四章では、「低処理量の耐雑音特徴補償方式」という目標を掲げている。屋外を含めた様々な環境で用いられる携帯機器においては、多種多様な雑音環境でも高い音声認識性能を維持することが重要である。一方、前章で達成した高速性・省メモリ性も保たなければならない。そこでこの目標に対して従来手法の中で有効性の高いヒストグラム同等化（HEQ）に着目し、そこで生じるオーバーフィッティングの問題を解決することによって性能を向上させるデルタケプストラム正規化（DCN）方式を新たに開発した。ここでは、DCNが従来手法の問題点を解決しうることを概念的に説明するとともに、詳細な定式化により、パラメータ値の最適化を含めた実装方法についても論じている。また、実データを用いた評価実験により、音声認識率および処理速度の両面において、この方式の有効性を確認した。

第五章では、「複数マイクの活用による雑音環境下音声認識性能の向上」という目標を掲げている。近い将来のプロセッサ性能の向上を念頭に置き、複数マイクからの音声入力を十分に活用しうるだけの処理性能が得られるという前提のもと、より困難な雑音環境ないしは対象タスクを仮定し、音声認識性能を向上させるための方式について論じている。特に、汎用ミドルウェアという提供形態の特性を鑑み、システムとしての動作安定性やプログラムの再利用可能性を保つことの重要性を述べ、そのような条件下で、チャンネル選択と特徴量合成という二つのアプローチが有効であることを示した。さらに、チャンネル選択というアプローチでは、前章で扱った特徴補償方式の活用が有効であること、特徴量合成というアプローチでは、各チャンネルに対する認識仮説の活用が有効であることをそれぞれ示し、前者に対しては「デコーダに基づくチャンネル選択（DBCS）」、後者に対しては「仮説に基づく特徴量合成（HBC）」という具体的な新規アルゴリズムとして定式化した。これらのアルゴリズムの有効性について、実際の音声データを用いた評価実験による検証も合わせて行っている。

これらの成果により、携帯機器における音声認識の主要な問題点が明らかとなるとともに、新規に提案したアルゴリズムの活用によって音声認識の性能が向上することが示された。さらにこれらの結果により、携帯機器における音声認識の更なる性能向上へ向けて、明確な方向付けを与えることができた。

以上を要するに、本論文は博士（情報理工学）の学位請求論文として合格と認められる。