

豊橋技術科学大学長 殿

平成 21 年 3 月 3 日

審査委員長 堀川 順生



論文審査及び最終試験の結果報告書

このことについて、下記の結果を得ましたので報告いたします。
記

学位申請者	Huda Mohammad Nurul	学籍番号	第 069301号
申請学位	博士(工学)	専攻名	電子・情報工学専攻
論文題目	A study on articulatory feature extraction for robust speech recognition (頑健な音声認識のための調音特徴抽出に関する研究)		
公開審査会の日	平成 21 年 2 月 23 日		
論文審査の期間	平成 21 年 1 月 29 日～平成 21 年 3 月 2 日	論文審査の結果	合格
最終試験の日	平成 21 年 2 月 23 日	最終試験の結果	合格

論文内容の要旨

本論文では、頑健な音声認識を困難にしているパターン変動の問題を、特徴抽出に対する斬新なアイデアを基に解決することを目指している。現在、多くの音声認識(ASR)システムは、音声波形から MFCC(mel-frequency cepstrum coefficient)など時間-スペクトラム情報を抽出して用いている。本論文では、音声の原信号と考えられる調音指令を直接抽出し、ASR システムに組み込む方式を提案している。評価実験では、調音情報抽出のため、弁別的音素素性(Distinctive Phonetic Feature; DPF) を再定義すると共に、多層ニューラルネットワーク (Multi-layer neural network; MLN) に基づく様々な DPF 抽出器を構成し検出性能を評価している。また、隠れマルコフモデル(HMM)に基づく音素認識器と組み合わせ、認識性能を評価すると共に、性能に多大な影響を与える変動要因に対し、複数の DPF 抽出器を用いた正準化処理を提案し、実装上の様々な工夫を通して、計算コストの少ない方式を実現している。




本論文では、第1章の序論で研究の背景、目的、および章構成を説明している。第2章では、DPF を MLN から抽出する際のネットワーク構成を、リカレント型、ハイブリッド型を含め評価している。第3章では、DPF を HMM ベースの音素認識器に用いた際の認識性能と計算量に対する評価を行い、続いて第4章で、この結果を受け、特に音素境界で生じる DPF 系列歪みを抑制/強調ネットワークを導入して改良している。第5章では、言語モデルに日本語の全音節を含むサブワードモデルを入れ、音素認識評価と改良を行っている。第6章では、様々なパターン変動要因に対して、複数の DPF 抽出器を用いた正準化処理を導入し、性別と騒音に対する実験を通して、計算量を含めた評価を行っている。第7章は本論文のまとめと今後の課題について述べている。




審査結果の要旨

音声認識では、確率的識別器である隠れマルコフモデル (HMM) が世界的に主流となっている。しかし近年、HMM の限界が議論され始めるに伴い、これを越える新しい音声認識手法への期待が年々高まっている。本研究は、「人間は音声を音響信号としてでなく調音運動として聴取している」という近年の脳科学等が示す知見を入れ、これまでスペクトル分析器に過ぎなかった音声特徴抽出器に、人間の聴覚系が発話系と連携することで実現している調音運動抽出機能を組み込み、頑健で高精度の音素認識を実現した、新規性のある研究である。本研究では、調音情報を調音クラス、すなわち弁別的音素素性(Distinctive Phonetic Feature; DPF) を用いて表現している。本研究で提案された、二段の多層ニューラルネットワーク (Multi-layer neural network; MLN) による DPF 抽出器は、これまで標準的に用いられてきた MFCC 特徴と比べ、一段高い音素認識性能を与えることが、静かな環境下と騒音環境下の双方の実験から示されている。また、音素境界でのより正確な DPF 抽出を目指して、Inhibition/ Enhancement (抑制/強調) ネットワークを導入し、DPF 系列の整形を実現し、一段高い音素認識性能を達成している。調音特徴を用いた同様の研究は、内外の研究機関でも広く行われているが、多くは従来の音響特徴(MFCC)を併用していることを考えると、調音特徴単独でこれらの成果を得たことは特筆に値する。

また、性別、音響環境など多様な隠れ因子に基づく変動吸収に、複数の DPF 抽出器を用いた正準化処理を導入し、頑健な ASR システムを実現する方式を提案している。提案方式は、性能面だけでなく、計算コストの面からも、実用価値の高いものである。本研究は、音声認識に斬新なパラダイムを導入し、学術的にも評価されており、工学的発展性からも高く評価できる。以上により、本論文は博士(工学)の学位論文に相当するものと判定した。

審査委員

堀川 順生  中川 聖一  章 忠 

新田 恒雄  印  印  印

(注) 論文審査の結果及び最終試験の結果は「合格」又は「不合格」の評語で記入すること。