

審査の結果の要旨

氏名 王 曉東

本論文は「Neural-Network-Based Tone Recognition of Continuous Speech of Standard Chinese Using Tone Nucleus Model (声調核モデルに基づくニューラルネットワークを用いた標準中国語連続音声の声調認識)」と題し、ニューラルネットワークを用いた中国語音声の声調型識別器において、音節の中で安定した基本周波数パターンの特徴を示す声調核に着目することにより、従来よりも高い声調型識別率を達成したものであり、全6章からなり英文で書かれている。

第1章は「Introduction」であって、まず、中国語音声における標準中国語音声の位置付けをした後、その発音と表記について概説している。次に、声調言語の概説を行って、標準中国語の音節声調型の特徴を説明している。引き続き、現在までの、標準中国語の声調型に関する研究を概観し、最後に本論文の章立てを示している。

第2章は「Researches of Tone Recognition of Chinese」と題し、これまでの中国語声調型識別の手法を、音声認識システムで暗に取り扱うものと、明に取り扱うものとに分けて整理した上で、それぞれの問題点を指摘し、本論文では、音節の韻律的特徴に着目して明に声調型識別を行う手法を開発するとしている。特に、これまでの声調型識別手法として、声調核に着目する手法とニューラルネットワークを用いた手法を取り上げ、本論文では、それらを統合する形で、新しい声調型識別手法を開発するとしている。

第3章は「Tone Nucleus Model and Neural Network (Multi-Layer Perceptron)」と題し、本論文で開発した手法の基となる声調核モデルとニューラルネットワークについて述べている。声調核モデルは、音節基本周波数パターンを、先行音節からの過渡部、声調核、後続音節への過渡部に分けるもので、声調核の基本周波数パターンは隣接音節の声調型の影響を受けにくく、安定した形状を示すという特徴がある。音節全体でなく、声調核の基本周波数パターンに着目することで、隣接音節の声調型に頑健な声調型識別が実現できる。また、声調核の始点終点の基本周波数を、隣接音節のそれらと比較したパラメータも識別に有効である。与えられた基本周波数パターンから声調核を自動的に検出する手法として、ビット探索に基づくものを新しく提案し、99.4% (従来法では 97.5%) の検出性能を得たとしている。声調識別に、一般的に利用される隠れマルコフモデル (HMM) は、音素持続時間といった特徴を識別に直接利用することが困難である。これに対し、ニューラルネットワークでは、基本周波数と持続時間のような、性格の異なる特徴を、同時に識別に利用することが容易である。ニューラルネットワークの種々のタイプを概説し、本論文では、構造が簡単な Multi-Layer (3層) Perceptron を用いるとしている。

第4章は「Tone Recognition Based on Tone Nucleus Model and Multi-Layer Perceptron」と題して、本論文で提案・開発した声調型識別手法とそれによる1名の女性話者連続音声についての声調型識別実験を行っている。性能比較のため、音節全体の基本周波数パターンの特徴を用いた Multi-Layer Perceptron による声調型識別実験を行い、提案手法の有効性を示している。また、声調核モデルに基づいた声調型識別手法として、従来報告された HMM を用いた場合の識別結果と比較しても、良い結果が得られたとしている。

第5章は「Feature Normalizations of Speaker Independent Tone Recognition」と題して、不特定話者の連続音声を対象とした声調型識別について述べている。話者による基本周波数レンジの違いに対処する手法として、平均の基本周波数で正規化する手法を用いるとし、正規化の手法について比較検討した結果について述べている。その結果、発話全体の基本周波数の平均値を用いるよりも、直近の10文程度の平均値 (Moving Average) を用いること

が有効であるとしている。音節全体の基本周波数パターンの特徴を用いた **Multi-Layer Perceptron** との性能比較を行い、声調核モデルを用いることの有効性を確認している。さらに、認識対象話者の基本周波数のヒストグラムを、学習データのそれにマッピングすることにより、識別率が向上することを示した。

第6章は「Summary」であって、本研究で得られた成果を要約し、将来の課題・展望について言及している。

以上を要するに、本論文は、標準中国語連続音声認識で問題となっている声調型識別を有効に行い得る手法として、声調核の韻律的特徴を用いた **Multi-Layer Perceptron** による手法を新しく提案し、その、従来手法と比較した優位性を、特定話者、不特定話者の場合の声調型識別実験を行って実証したものであり、電子工学、情報工学に貢献するところが少ない。

よって、本論文は博士（工学）の学位請求論文として合格と認められる。