

論文審査の結果の要旨

氏名 馬 学彬

本論文は「Chinese Dialect-Based Speaker Classification and Pronunciation Assessment Using Structural Representation of Speech (音声構造表象を用いた中国語方言に基づく話者分類と発音評価)」と題し、全十章から成る。年齢や性別といった要因による音声の音響的変動に影響されずに、中国語話者を方言性に基づいて分類する方式を提案し、その技術的及び言語学的妥当性を実験的に検証している。

第一章は序章であり、本研究の背景、目的、及び章構成について述べている。

第二章では、中国語方言の概要についてまとめている。中国語は方言が異なると音声コミュニケーションが困難となるが、ここでは、方言差異を越えた普遍的な言語的特徴と同時に、個々の方言によって異なる言語的特徴について説明している。また、民族の移動や北京語の習得義務などに起因した方言の変容についても述べている。工学的に音声からの方言同定を試みる場合、従来、性別・年齢の多様性を網羅するために、同一方言話者の音声を多数集めて統計モデルを構築することが行われてきたが、中国語の場合、方言の分化が多岐・多様に行われており「同一方言の話者を多数集める」ことが困難な様子も説明している。即ち、集めること以外の解決が必要となる。

第三章では、関連する音声技術と、類似した言語学的な方言研究についてまとめている。音声認識や言語認識で用いられる音響特徴量や、そのモデリング技術、更には、この特徴量が非言語的要因によってどのように変形するのかについて説明している。音声試料を用いた方言分類は言語学においても研究されており、そこで用いられている話者正規化手法について概観している。

第四章では、第五章以降で行う中国語の方言性に基づく話者分類実験において用いられる種々の要素技術について説明している。話者性を消去した形で音声（発音）を表象する技術（音声の構造的表象）は先行研究で提唱された技術であるが、ここでは、これを中国語方言分類に用いるための種々の準備について述べている。特に、方言が異なった場合でも、同じ「情報」を方言話者に発声させる必要がある。彼らが発声する「情報」の同一性を保証し、かつ、方言の違いが明確に観測される発声を得る方法として、ここでは、方言学において用いられてきた「文字セット」の読み上げを採択した。

中国語は大きく七方言に分類されているが、第五章では、その中から四方言・18名と対象として、方言性に基づく話者分類を検討した。文字セットは38文字あり、各文字の母音部を用いて38角形を構成し、この多角形を18名の話者間で比較することで分類した。実験の結果、方言性に基づいた分類が行われていることが示された。18話者のうち1人は二方言を話す話者であるが（即ち声質が同一の異方言音声）、同一話者であるにも拘わらず、方言に基づいた分類が行われた。

第六章では、下位方言（sub-dialect）や、下位下位方言（subsub-dialect）に基づく話者分類を検討した。中国語は大きく七方言に分かれるが、各方言が下位方言を、各下位

方言が下位下位方言を持つ。提案手法は方言性に基づいて個人を分類することを目指しており、これらのより下位区分に基づく話者分類が可能であることを示すことができた。

第七章では、年齢・性別に依存しない話者分類についての技術的妥当性を示す目的で、同一話者による多方言音声を収録した。即ち、多話者・多方言音声を方言学者一人によって「真似させ」、同一話者・多方言音声を収録した（性別・年齢差異の無い多方言音声試料が得られる）。この音声試料に基づく方言分類と、多話者を用いた方言分類とを比較すると、ほぼ等しい結果が得られ、提案手法の話者不変性を示すことができた。

第八章では、逆に、年齢や性別に起因する音響的変動を、音声モーフィング技術を用いて擬似的に最大化し、この場合においても提案手法が頑健に動作する様子を実験的に示した。更に、従来の音響モデリング技術である、スペクトルの統計モデル及びスペクトル照合に基づく話者分類についても検討した。その結果、スペクトルをそのまま比較すると、体格や性別によって話者が分類される結果となった。このように年齢・性別・体格差が極めて大きな音声データであっても、提案手法を用いれば、これらに影響を受けずに、方言性に基づいた話者分類が可能であることが示された。

第九章では、提案手法の更なる応用可能性を検討することを目的として、方言話者の北京語音声について、北京語母語話者との比較を通して、発音矯正すべき部位を特性する実験を行い、この場合においても、良好な結果を示すことができた。

第十章にて、本論文を総括している。本論文では非言語的特徴を消失させながら音声を表象する技術に基づいて、方言性に基づく中国語話者分類、及び、訛りを有する北京語分析を行い、構造表象の方言分析への応用可能性を実験的に示した。本論文では韻律的特徴までを網羅することは出来なかったが、十分に言語学的妥当性のある方言分類を、音声試料に対する自動音響分析のみで生成することに成功した。

以上要するに、本手法は方言的多様性の高い中国語に対して、性別や年齢といった要因を音声データより消失させつつ、方言性に基づく自動話者分類を実現しており、情報学の基盤に貢献するところが少なくない。よって、本論文は博士（科学）の学位請求論文として合格と認められる。