

## 審査の結果の要旨

氏名 李宝潔

本論文は「Adaptation of Acoustic Models for Speech Recognition with a Focus on Intra-Speaker Variation (話者内変動に着目した音声認識のための音響モデルの適応)」と題し、種々の発話に対応した話者内の音響的特徴の変動にも着目して、音声認識の音響モデルの適応手法を開発して、実験的検討を行ったものであって、全6章からなり、英文で記述されている。

第1章は「Introduction」であって、音声認識の現状の概観を述べた上で、音響モデルの問題点として環境の変化による認識性能の低下を指摘し、話者間の違いのみならず発話の違いによる話者内の変動にも対処するために、頑強な音響モデルあるいは有効な適応手法が必要であるとしている。このような観点から、本論文の内容が、音素ペアモデルと2段階 MLLR (Maximum Likelihood Linear Regression) の提案、感情音声認識によるその有効性の実証にあるとしている。最後に、第2章以降の論文構成を説明している。

第2章は「Basics of Automatic Speech Recognition」と題し、まず、音声認識が音響モデルと言語モデルの尤度計算から成立することを述べ、本論文では音響モデルとしての HMM (Hidden Markov Model) を対象とするとしている。その上で、音声認識の問題として、音響的特徴の変動に起因する認識率の低下があり、それに対処するために不特定話者用の音響モデルを特定話者用のモデルに適応することが一般的であるとしている。

第3章は「Speaker Adaptation Techniques」と題し、音響モデルを新しい話者に適応する技術として、MAP (Maximum A posteriori Prediction) 推定、MLLR 推定を中心に概説している。また、次章で提案するモデルと関連して、音素の関係を利用する適応手法について説明している。

第4章は「Phone Pair Model for Speaker Adaptation」と題して、発話の変動に対して頑健な特徴である音素間の特徴パラメータ空間中での相対的位置関係をHMMに組み込んだ音素対モデルを提案し、少量の音声資料でも有効な適応手法が可能であることを示している。さらに、このモデルは、母音に限定して用いているが、それ自体、話者の変動に強い頑健なモデル化であることを示している。

第5章は「Two-Level MLLR Adaptation for Inter- and Intra-Speaker Acoustic Variation」と題して、例えば感情音声のように話者内でも変動の大きい音声を対象とした場合の適応手法として、2段階の MLLR を提案している。この手法は、不特定話者モデルを出発点として、まず、新しい話者の種々の感情音声で MLLR 推定を行い、特定話者モデルを構築する。次に、個々の感情音声でさらに MLLR 推定を行い、特定感情モデルを構築する。特定感情モデルと特定話者モデルを組み合わせて用いることにより、特定話者モデル単独よりも高い認識率が得られることを実験的に示している。この2段階 MLLR 手法は、MLLR 推定のための音声資料が非常に少ない時に特に有効である。また、特定感情モデルを用いて入力音声の感情の識別が可能なることも示している。

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

以上を要するに、本論文は、音声認識で問題となる話者あるいは発話の違いによる音響的特徴の変動に対処する有効な適応手法として、音素対モデルと2段階 MLLR (Maximum Likelihood Linear Regression) 手法を提案し、特に後者については、従来、高精度の認識が困難であった感情音声の認識にも有効であることを示している。現在、音声認識で問題となっている自発発話音声では、発話の違いによる変動が大きいと考えられ、本論文で提案した手法は、それに対する有効な1つの回答となる可能性がある。このように、本論文は音声認識の将来的な発展に結びつくものであり、電子工学、情報工学に貢献するところが少なくない。

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。