

豊橋技術科学大学長 殿

平成27年 2月27日

審査委員長 増山 繁



## 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。

学位申請者	木村 優志	学籍番号	第033711号
申請学位	博士(工学)	専攻名	電子・情報工学専攻
論文題目	マルチモーダルタスク推定のための調音運動HMM音声認識合成方式		
公開審査会の日	平成27年 2月12日		
論文審査の期間	平成27年 1月22日～平成27年 2月23日	論文審査の結果	合格
最終試験の日	平成27年 2月12日	最終試験の結果	合格

論文内容の要旨  
音声認識合成技術の発展に伴い、近い将来、人間が遂行するタスクを支援するロボットにも音声対話技術が利用されることが期待されている。その際には音声認識・合成、および、対話技術のみでなく、周辺環境から対話者の状態、取り組むタスク、あるいは行動を推定する技術が必要になると考えられる。本論文は、こうした応用が現実となる状況に備えるべく、特に音声合成とタスク推定技術に焦点を当て、それらの性能向上に取り組んだものである。

第1章では、人間と共生するロボットが備えるべき音声認識合成技術について認知心理学的な面から分析すると共に、ロボットが備えるべきタスク推定の機能を明らかにしている。第2章では、画像情報と音声言語の情報に潜在意味解析を適用し、タスクを推定する手法を提案すると共に、短時間で周辺の人物が従事するタスクを同定可能なことを示している。第3章では、人間の脳で行っていると推定される、音声認識と合成を調音運動のワンモデル（認識と合成で共通の音響モデル）として実現するシステムに基づき、特に音声合成に焦点を当てた基本アルゴリズムを提案すると共に、主観・客観評価を実施している。続いて第4章では、前章の音声合成法のうち駆動音源に残差符号帳を導入することで、より高品質な合成音が可能であることを示している。第5章は結論であり、本論文の中で提案したタスク推定とワンモデル合成の手法が、人間と共生するロボットを実現する上で重要な技術となることを示唆している。

審査結果の要旨  
本研究は、人間と共生する近未来のロボット実現に必要なマルチモーダルタスク推定法、および、その要素技術の一つである調音運動 HMM (Hidden Markov Model) 音声認識合成方式について提案している。

タスク推定に関しては、タスク毎の音声言語情報と画像情報を同一の行列上に表現すると共に、潜在意味解析を適応して得た特異ベクトルからタスク毎の類似度を定義し、未知タスクを判定している。提案方法は、シンプルな手法ながら高精度かつ高速にタスクを推定できることが評価実験から示されており、ロボットによるタスク推定の実現に向けて重要な技術になり得ると考えられる。また音声認識合成方式については、主に音声合成について検討しており、話者非依存な調音運動 HMM モデル（このモデルは音声認識にも共通して利用できる）を用いた合成方式を提案する等、従来にはない手法を提案している。人間の発音器官に相当する話者依存の部分については、調音運動から声道を模擬する LSP (Line Spectral Pairs) パラメータに変換する多層ニューラルネットワークに機能を集約しているため、工学的には少量のデータで音声合成を実現することができるという利点があり、その有用性は高い。また駆動音源の改良に対しては、HMM の各状態に残差符号帳中最小歪を与える符号を貼り付けるという、新規性の高い手法を提案し実装している。

これらの研究成果は査読付き学術論文2編、査読付き国際会議論文1編として採録され、学術的価値が認められている。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	増山 繁	堀川 順生	新田 恒雄
	桂田 浩一		

(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。