

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	知能コミュニケーション (中村 哲 (教授))		
学籍番号	2011191	提出日	令和 4年 1月 19日
学生氏名	徳山 太顕		
論文題目	パラ言語情報の強調表現を利用したニューラル音声翻訳		
要旨			
<p>音声コミュニケーションにおいて、話し手は自分のメッセージを感情や強調などのパラ言語情報とともに言葉で伝える。すべての話し言葉の情報を文字や言葉の形に変換する際、特に言語をまたいで変換しなければならない場合は、言語だけで表せない非常に多くの情報が必要とある。しかし、既存の音声からテキストへの翻訳システムにおいては、言語情報の翻訳にのみ焦点を当てており、パラ言語情報は無視しているものがほとんどである。近年の研究では、言語とパラ言語の認識と翻訳を別々にモデル化する手法が提案されているが、隠れマルコフモデル (HMM) ベースの自動音声認識 (Automatic Speech Recognition; ASR) とテキスト音声合成 (Text-to-Speech; TTS) にニューラルネットワークベースの機械翻訳 (Machine Translation; MT) を使用しており、複雑で最適とは言えないものであった。さらに、パラ言語情報は音響的な形で保持されているため言語的に表現ができていないことが多い。そこで本研究では、ターゲット言語のテキストに含まれる強調のパラ言語情報をテキストに埋め込むことに焦点を当てた。具体的には、カスケード型および直接型のニューラルネットワークを用いて音声からテキストへの翻訳モデルを構築し、ターゲット言語のテキストに強調情報を表現する様々な方法を検討した。そして日英言語間、およびパラ言語間の音声からテキストへの翻訳のフレームワークを用いて実験を行った。その結果、提案手法は、標準的な言語翻訳と同様の性能を維持したまま、言語情報とパラ言語情報の両方を翻訳できることが明らかになった。</p>			