

数理・情報のフロンティア
2019 年度採択研究代表者

2020 年度 年次報告書

木村 直紀

東京大学 大学院学際情報学府
大学院生(博士課程)

口腔内超音波画像と深層学習を用いた無声発話認識に関する研究

§ 1. 研究成果の概要

本研究は、発音動作を捉える口腔内超音波画像を中心としたセンサ情報を用いて、発話内容を推定する無声発話認識に関する研究である。2019年度は超音波画像から音響特徴量を直接推定する手法を中心に研究を行っていたが、より高い精度を達成するために2020年度は音声認識のパイプラインを使用する Visual Speech Recognition (VSR) へターゲットを変更した。また、口腔内超音波画像に加えて口元画像が同期されたデータセットに焦点を絞り、既存の公開されたデータセットを用いて、深層学習を用いた VSR パイプラインの基盤を作成した。この時得られた成果は国際会議にて発表された。またその結果から、より大きなデータセットを構築することによって高い精度が得られる可能性が示唆されたため、データセット収集を行うための準備を行なった。