



Title	Sign Language Translation Using Wearable Motion Capture System and Machine Learning Methods [an abstract of dissertation and a summary of dissertation review]
Author(s)	Gu, Yutong
Citation	北海道大学. 博士(工学) 甲第15177号
Issue Date	2022-09-26
Doc URL	http://hdl.handle.net/2115/87182
Rights(URL)	https://creativecommons.org/licenses/by/4.0/
Type	theses (doctoral - abstract and summary of review)
Additional Information	There are other files related to this item in HUSCAP. Check the above URL.
File Information	Yutong_Gu_review.pdf (審査の要旨)



[Instructions for use](#)

学位論文審査の要旨

博士の専攻分野の名称 博士 (工学) 氏名 Gu Yutong

審査担当者 主査教授 東藤 正浩
副査教授 大橋 俊朗
副査教授 清水 裕樹

学位論文題名

Sign Language Translation Using Wearable Motion Capture System and Machine Learning Methods

(装着型運動計測システムと機械学習による手話翻訳手法に関する研究)

手話は聴覚障害者の間の主要なコミュニケーション手段である。話し言葉や書き言葉の機械翻訳が非常に正確に行われるようになった一方、手話に関しては今日の主流の研究テーマには至っていない。しかし、深層学習モデルを使用した機械翻訳の研究は、手話翻訳タスクの開発の方向性と革新的な方法を提供することが可能となる。エンドツーエンドの翻訳に関する研究をさらに進めるには、深層学習モデルの適用を検討する必要がある。手話翻訳に関するこれまでの研究では、主に視覚ベースとウェアラブルセンサベースの2つのカテゴリに分類される。視覚ベースの方法では、カメラを利用して手の特徴をキャプチャする。一方、ウェアラブルセンサベースの研究では、データグローブ、腕時計、アームバンドなどのデバイスがデータ収集の主流となっている。

本論文ではウェアラブルセンサを用いた手話翻訳手法を提案している。本研究では、独立したジェスチャ認識から始まり、最終的にウェアラブルセンサを使用した全文のエンドツーエンドの翻訳を行っている。手話パフォーマンスの表情も筋電信号としてデバイスにより収集した。自然言語処理とウェアラブルセンサの組み合わせにより新しい手話翻訳技術を提供している。収集したデータセットは、より多くの手話翻訳と機械学習の研究にも貢献するものである。

本論文は全5章で構成されており、以下にそれぞれの章の概要を示す。

第1章では、手話翻訳作業の背景と必要性について説明している。翻訳方法と関連技術に関する文献を紹介し、通常使用されるデータセットについてもまとめている。

第2章では、単独のジェスチャ認識に焦点を当てている。いくつかの広く使用されているハンドジェスチャ認識データセットを導入し、従来の機械学習モデルでテストを行っている。また、カスタマイズされたモデルを構築して、分類精度を向上させた。最終的に、17のジェスチャのデータセットがハンドモーションキャプチャシステムによって収集するとともに、その精度について述べている。

第3章では、アメリカ手話の指のジェスチャを紹介している。アメリカ手話文字のパフォーマンス中に右手の慣性データを収集し、文字レベルと単語レベルの両方で分類タスクを実行している。機械学習モデルには、特徴抽出のための畳み込みニューラルネットワーク層、時系列特性を学習するための長期短期記憶層、およびモデル出力とグラウンドトゥルースとの間の整合問題を解決するためのコネクショニスト時間分類層が含まれている。

第4章では、両手の動きと顔の表情を考慮したウェアラブルセンサベースの手話翻訳方法について述べている。慣性センサユニットと筋電図信号は前処理され、変換モデルの入力として一連のフ

フレームに分割し、筋電図信号のみで表情を分類している。信号からテキスト文へのエンドツーエンドの手話翻訳を実現するためのエンコーダ/デコーダモデルを構築した。2種類のエンドツーエンドモデルが、収集されたデータセットによってトレーニングおよび評価している。どちらのモデルも40個のアメリカ手話文を高い精度で翻訳でき、変換ベースのモデルは優れたパフォーマンスを発揮することを示している。筋電信号の役割は、顔の表情の分類とモデルのパフォーマンスの両方で検証している。

第5章では、本研究で得られた成果を述べている。

これを要するに、著者はウェアラブルセンサを用いた深層学習による手話翻訳手法を提案したものであり、生体医工学の発展に貢献するところ大なるものがある。よって著者は、北海道大学博士(工学)の学位を授与される資格あるものと認める。