学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称　博士（工学）　氏名　Gu Yutong

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)

Sign language is the main communication method among hearing-impaired people. As a kind of natural language, sign language has not become a mainstream research topic in natural language processing, although the machine translation of spoken or written language is highly accurate today. However, the research of machine translation with deep learning models provides development direction and innovative methods for sign language translation tasks. In order to further the research on end-to-end translation, it is necessary to consider the application of deep learning models. Previous works about sign language translation mainly falls to two categories: vision-based and wearable sensors-based. Vision-based methods exploit camera to capture features of hands. In wearable sensors-based research, devices like data glove, wristwatch or armband are the mainstream for data collection. In this dissertation, we will explore the sign language translation using wearable sensors.

Chapter 1 provides the background and necessity of sign language translation task. The literature review of translation methods and devices were introduced. The normally used datasets were also summarized.

Chapter 2 focuses on the isolated gesture recognition. Some widely use hand gesture recognition datasets were introduced. Then, the classification task of Ninapro DB5 dataset was tested with traditional machine learning models. Also, customized model for Ninapro DB5 was built to promote the classification accuracy. Finally, dataset of 17 gestures was collected by hand motion capture system. The quality of data from the device was tested by a deep neural network classifier.

Chapter 3 introduces the finger spell in American sign language. We collected inertial data of right hand during ASL letters performance, and did the classification tasks in both letter and word levels. The machine learning model contains convolutional neural network layers for feature extraction, long short-term memory layers for learning time series characteristics, and connectionist temporal classification layers to solve alignment problem between model output and ground truth.

Chapter 4 presents a wearable sensors-based sign language translation method considering both hands' movements and facial expressions. Inertial measurement units and electromyography signals were preprocessed and segmented into a sequence of frames as the input of translation models. We classified facial expressions with EMG data only. Then we built encoder-decoder models to realize end-to-end sign language translation from signals to text sentences. Two kinds of end-to-end models based on LSTM and transformer were trained and evaluated by the collected dataset. Word error rate and sentence error rate were used to compare the translation ability of models. Both models could translate 40 ASL sentences with high accuracy and the transformer-based model performed better than LSTM. The special role of EMG was verified with both facial expressions' classification and models' perfor-

mance after removing EMG from the input. The translation accuracy in user-independent conditions was evaluated.

Chapter 5 summarizes the works in the dissertation and offers the prospective studies that can be investigated in this research field.

In summary, this study proposed the complete research process of sign language translation technology. The research was started with isolated gesture recognition and finally went to the end-to-end translation of full sentences using wearable sensors. The facial expressions in sign language performances were also collected by EMG device. The combination of natural language processing and wearable sensors provides a new idea for sign language translation task. The datasets we collected will make it easier for more people to start research on sign language translation and machine learning. The works in this study are significant new and may contribute a huge impact for researchers in this field.