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	2211224	提出日	令和 6年 1月 19日
学生氏名	西川 勇太		
論文題目	Extraction of Speech Representation for Multilingual Speech Translation		

要旨

In End-to-End speech translation, pre-trained models of speech and text are crucial elements for improving translation quality. Previous models simply con- nected encoders and decoders. However, only the information from final layer of the encoder is forwarded to the decoder, which means that the worthy information contained in the intermediate layers of the speech SSL model is not fully utilized. This thesis proposes two new methods to effectively utilize the information from each layer of the speech SSL model. The first method, Inter-connection, aggre- gates the information from each layer using a static weighted sum and inputs it to the cross attention in the decoder. This approach was particularly successful in improving the BLEU score by about 1.5 points for the En-De, En-Ja, En-Zh language pairs. Additionally, it was observed that slightly different weights are obtained for each language pair. The second method, Layer Attention, aggre- gates information using dynamic weights based on the input. While this method showed a slight improvement in performance, it resulted in a more complex model and an overall performance that was inferior to Inter-connection. This research demonstrates the potential for application in fields beyond End-to-End speech translation. Furthermore, by further developing the methods of utilizing inter- mediate representations and extracting them more effectively, the versatility and effeciency of pre-trained models are expected to be further enhanced.