

# 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	2011307	提出日	令和 4年 1月 19日
学生氏名	胡 尤佳		
論文題目	End-to-end Speech Translation Leveraged by ASR Posterior Distribution (音声認識出力の曖昧性を考慮するEnd-to-End音声翻訳)		
要旨			
<p>End-to-end speech translation (ST) translates source language speech directly into target language without an intermediate automatic speech recognition (ASR) output, as in a cascading approach. End-to-end ST has the advantage of avoiding error propagation from the intermediate ASR results, but its performance still lags behind the cascading approach. A recent effort to increase performance is multi-task learning using an auxiliary task of ASR. However, previous multi-task learning for end-to-end ST using cross entropy (CE) loss in ASR-task targets one-hot references and does not consider ASR hypotheses. In this study, we propose a novel end-to-end ST training method using ASR loss against ASR posterior distributions given by a pre-trained model, named as ASR Posterior-based Loss (ASR-PBL). The proposed method is expected to consider possible ASR hypotheses due to competing hypotheses with similar pronunciations. Experimental results in Fisher Spanish-to-English show that the proposed method resulted in better BLEU scores than the baseline with the standard CE loss with label smoothing. And we also conducted the experiment of ASR Sequence-based loss (ASR-SBL), which uses the ASR one-best predictions into the ST training, in Fisher and MuST-C dataset. The results ASR-SBL shows that improvement in Fisher in most of the cases under various ASR models, and both ASR-PBL and ASR-SBL has the ability of improvement compared to baselines. However, there are no improvement in MuST-C data.</p>			