【Grant-in-Aid for Scientific Research (S)】

Integrated Disciplines (Informatics)

Title of Project : Next Generation Speech Translation Research

Satoshi Nakamura
(Nara Institute of Science and Technology, Graduate School of Information Science, Professor)

【Purpose and Background of the Research】

The conventional speech translation, which translates a sentence after each utterance in a simple domain such as travel conversation, is advancing for a practical use, but the simultaneous translation still faces difficulties for practical use. Especially for the translation between Japanese and English, as they have different sentence structures, the system cannot start translation until the verb or negation is recognized. The goal of this project is to develop a next generation speech translation system that is able to start translation instantly without waiting for a sentence end as human interpreters do. It conveys the message of speakers by avoiding a breakdown caused by the structure sentence difference.
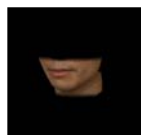
【Research Methods】

①A) Noise reduction based on DNN by using the noise database, integration with independent low-rank matrix analysis, and development of target voice extraction and enhancement with multiple distributed microphone arrays, B) improvement of continuous speech recognition system for lectures based on the system for national congress, C) upgrading simultaneous translation by predicting a potential tree structure of the next utterance from the current parsing analysis result, and utilizing a tree-to-string model, D) trial to compress input/ output layers of attentional NMT based on LSTM, E) introduction of dialogue management technologies

②A) Paralinguistic translation of speech emphases and emotional expressions, B) speech synthesis with paralinguistic information

③A) captioning of speech translation based on summarization technologies, B) multimodal translation by using lip sync for videos
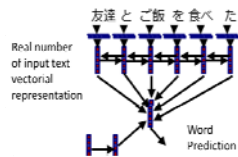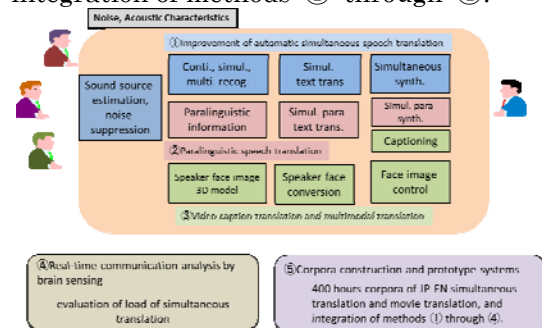
Model Fitting   Mouth Model

④Real-time communication analysis by brain sensing for evaluation of load of translation
⑤A) Corpora construction, B) prototype systems by integration of methods ① through ④.

【Expected Research Achievements and Scientific Significance】

①Continuous noisy speech recognition of multiple speakers and automatic incremental speech translation with consideration of sentence structure difference between languages. ② Paralinguistic speech translation to extract, preserve and reproduce speaker's emotion, emphasis, and individuality. ③Video captioning and translation for lectures and movies. ④Data analysis of brain activity sensing during translation. ⑤ 400 hours corpora of JP-EN simultaneous translation and movie translation, prototype systems

【Publications Relevant to the Project】

・Yusuke Oda, Satoshi Nakamura et al., "Syntax-based Simultaneous Translation through Prediction of Unseen Syntactic Constituents", ACL 2015, pp. 198-207. July 2015.
・Satoshi Nakamura, "Speech Translation Technologies, Past and Beyond," The Journal of the IEICE, vol.96, no.11, pp865-873, 2013.11

【Term of Project】 FY2017-2021
【Budget Allocation】 157,100 Thousand Yen
【Homepage Address and Other Contact Information】 Not yet available