EUROSPEECH 2003 - INTERSPEECH 2003

8th European Conference on Speech Communication and Technology

**Geneva, Switzerland
September 1-4, 2003**

Noise Robustness in Speech to Speech Translation

Fu-Hua Liu, Yuqing Gao, Liang Gu, Michael Picheny

IBM T.J. Watson Research Center, USA

This paper describes various noise robustness issues in a speech-to-speech translation system. We present quantitative measures for noise robustness in the context of speech recognition accuracy and speech-to-speech translation performance. To enhance noise immunity, we explore two approaches to improve the overall speech-to-speech translation performance. First, a multi-style training technique is used to tackle the issue of environmental degradation at the acoustic model level. Second, a pre-processing technique, CDCN, is exploited to compensate for the acoustic distortion at the signal level. Further improvement can be obtained by combining both schemes. In addition to recognition accuracy for speech recognition, this paper studies and examines how closely speech recognition accuracy is related the overall speech-to-speech recognition. When we apply the proposed schemes to an English-to-Chinese translation task, the word error rate for our speech recognition subsystem is substantially reduced by 28% relative, to 13.2% from 18.9% for test data of 15dB SNR. The corresponding BLEU score improves to 0.478 from 0.43 for the overall speech-to-speech translation. Similar improvements are also observed for a lower SNR condition.

Full Paper

Bibliographic reference. Liu, Fu-Hua / Gao, Yuqing / Gu, Liang / Picheny, Michael (2003): "Noise robustness in speech to speech translation", In *EUROSPEECH-2003*, 2797-2800.