Transcription System using Automatic Speech Recognition in the Japanese Parliament (Diet)

Tatsuya Kawahara

Professor Academic Center for Computing and Media Studies, Kyoto University

Sakyo-ku, Kyoto 606-8501, Japan - Member of the Scientific Committee of Intersteno

The House of Representatives of the Japanese Parliament (Diet) is introducing a new-generation transcription system, using automatic speech recognition (ASR) technologies, for producing their verbatim records. The system is to replace the conventional manual shorthand scheme, in order to streamline the process and also to reduce the burden and training periods of the stenographers. We have developed a speaker-independent ASR system dedicated to the meetings in the Diet, by collecting over 300 hours of meeting speech aligned with the official verbatim records of the proceedings. Please note that there are a number of mismatches between verbatim records and what is actually uttered, because of the removal of fillers and disfluencies and correction of colloquial expressions. In the Japanese language, differences between spoken language and document-style language are large and typing involves conversion of kana (phonetic) input to orthographic kanji (Chinese) characters. These issues make it hard to real-time typing and re-speaking to dictation software, which are common in US and several other countries.

Even the state-of-the-art ASR systems do not realize accuracy of usable level (>80%), without construction of a large-scale speech database, which needs much cost and expert knowledge. With these efforts together with some innovative techniques, our system has achieved accuracy of over 90% for plenary sessions and 85% for committee meetings with turn-around of almost real-time. The post-editing software for correcting speech recognition errors and cleaning the verbatim transcripts, which was designed by stenographers, also incorporates the advanced ICT, including quick replay of audio/video recordings indexed with every utterance of the transcriptions. It is integrated with the overall system including job assignment, editorial proof-reading and final publication.

The system is currently being tried and evaluated, and it is confirmed that the system is usable with this level of accuracy (85%), though the accuracy of over 90% is preferred. The system will be deployed within this fiscal year, and we expect that it will be improved with more experiences and data being accumulated.

Presentation slides in INTERSTENO 2009 available at:

http://www.intersteno.it/materiale/Beijing2009/Conferences/Int09%20-%20sessions%20-%20Kawahara%20-%20presentation.ppt