Document Driven Machine Translation Enhanced ASR

M. Paulik (1), Christian F gen (1), Sebastian St ker (1), Tanja Schultz (2), Thomas Schaaf (2), Alex Waibel (1)

(1) Universit t Karlsruhe, Germany; (2) Carnegie Mellon University, Pittsburgh, PA, USA

In human-mediated translation scenarios a human interpreter translates between a source and a target language using either a spoken or a written representation of the source language. In this paper we improve the recognition performance on the speech of the human translator spoken in the target language by taking advantage of the source language representations. We use machine translation techniques to translate between the source and target language resources and then bias the target language speech recognizer towards the gained knowledge, hence the name Machine Translation Enhanced Automatic Speech Recognition. We investigate several different techniques among which are restricting the search vocabulary, selecting hypotheses from n-best lists, applying cache and interpolation schemes to language modeling, and combining the most successful techniques into our final, iterative system. Overall we outperform the baseline system by a relative word error rate reduction of 37.6%.

[Full Paper](#)

Bibliographic reference. Paulik, M. / F gen, Christian / St ker, Sebastian / Schultz, Tanja / Schaaf, Thomas / Waibel, Alex (2005): "Document driven machine translation enhanced ASR", In *INTERSPEECH-2005*, 2261-2264.