Towards Unsupervised Training of the Classifier-Based Speech Translator

Emil Ettelaie, Panayiotis G. Georgiou, Shrikanth S. Narayanan
University of Southern California, USA

Concept classification has been proven to be a useful translation method for speech-to-speech translation applications. However, preparing training data for classifier is a cumbersome task for human annotators. An unsupervised training method is introduced here that is based on utterance clustering. A technique to measure the distance between two utterances, based on the concepts they express, along with an appropriate clustering method has been adapted.

Full Paper