

Foucher, S., Laliberté, F., Boulianne, G. and L. Gagnon. "A Dempster-Shafer Based Fusion Approach for Audio-Visual Speech Recognition with Application to Large Vocabulary French Speech" In *Proceedings of the 2006 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2006)*, pp. I-597 - I-600. Toulouse, France, May 14-19, 2006.

Abstract

This work explores a new way of fusing audio and visual information for audiovisual automatic speech recognition in the context of a large vocabulary application. Mouth shape information is extracted off-line and integrated into a speech recognition system using a phoneme-based Dempster-Shafer fusion approach. The fusion methodology assumes that the audio information about the phonemes is a precise Bayesian source while the visual information is an imprecise evidential source. This ensures the visual information does not degrade significantly the audio information in situation where the audio performs well in controlled noiseless environment. Bayesian and simple consonance belief structures are explored and compared, along with standard stack-based fusion.