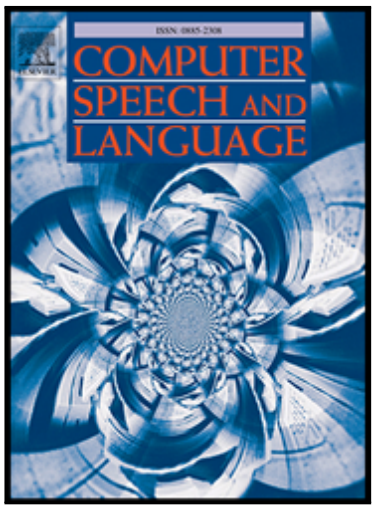

An Information Fusion Framework with Multi-Channel Feature Concatenation and Multi-Perspective System Combination for the Deep-Learning-Based Robust Recognition of Microphone Array Speech

Yan-Hui Tu, Jun Du, Qing Wang, Xiao Bao, Li-Rong Dai, Chin-Hui Lee

PII: S0885-2308(16)30076-6
DOI: 10.1016/j.csl.2016.12.004
Reference: YCSLA 809

To appear in: *Computer Speech & Language*

Received date: 8 April 2016
Revised date: 3 December 2016
Accepted date: 4 December 2016

Please cite this article as: Yan-Hui Tu, Jun Du, Qing Wang, Xiao Bao, Li-Rong Dai, Chin-Hui Lee, An Information Fusion Framework with Multi-Channel Feature Concatenation and Multi-Perspective System Combination for the Deep-Learning-Based Robust Recognition of Microphone Array Speech, *Computer Speech & Language* (2016), doi: 10.1016/j.csl.2016.12.004

Highlights

- The early fusion by using multiple beamformings and feature concatenation.
- The late fusion of subnets from multiple perspectives.
- A simplified and effective MVDR beamforming approach.
- Building the best one-pass single DNN system among all submissions to CHiME-3