

Neural Networks Based Text-Independent Speaker Identification with Breath and Ey

Wenbo Zhao

School of Electrical and

Computer Engineering

Carnegie Mellon University

Pittsburgh, PA 15213

Email: wzhao1@andrew.cmu.edu

Abstract—In text-independent scenarios and with short-time utterances, identifying speakers is difficult due to rich acoustical variations, lack of contextual information and data. To represent speech features, prior speaker models often use Gaussian Mixture Models. While effective, these models rely on prior data distribution assumptions. Moreover, training these models usually requires large amount of data and iterative procedures. To overcome these problems, this study proposes a neural networks based speaker identification framework. The framework extracts constant-Q spectrogram features from speech signals and models them using convolutional network and Long Short-Term Memory network. The framework is tested by identifying speakers using small amount of breath or ‘Ey’ sound recordings. Results show that the framework is accurate and converges fast.