BIEN 2011

Delta-Spectral Cepstral Coefficients for Robust Speaker Recognition

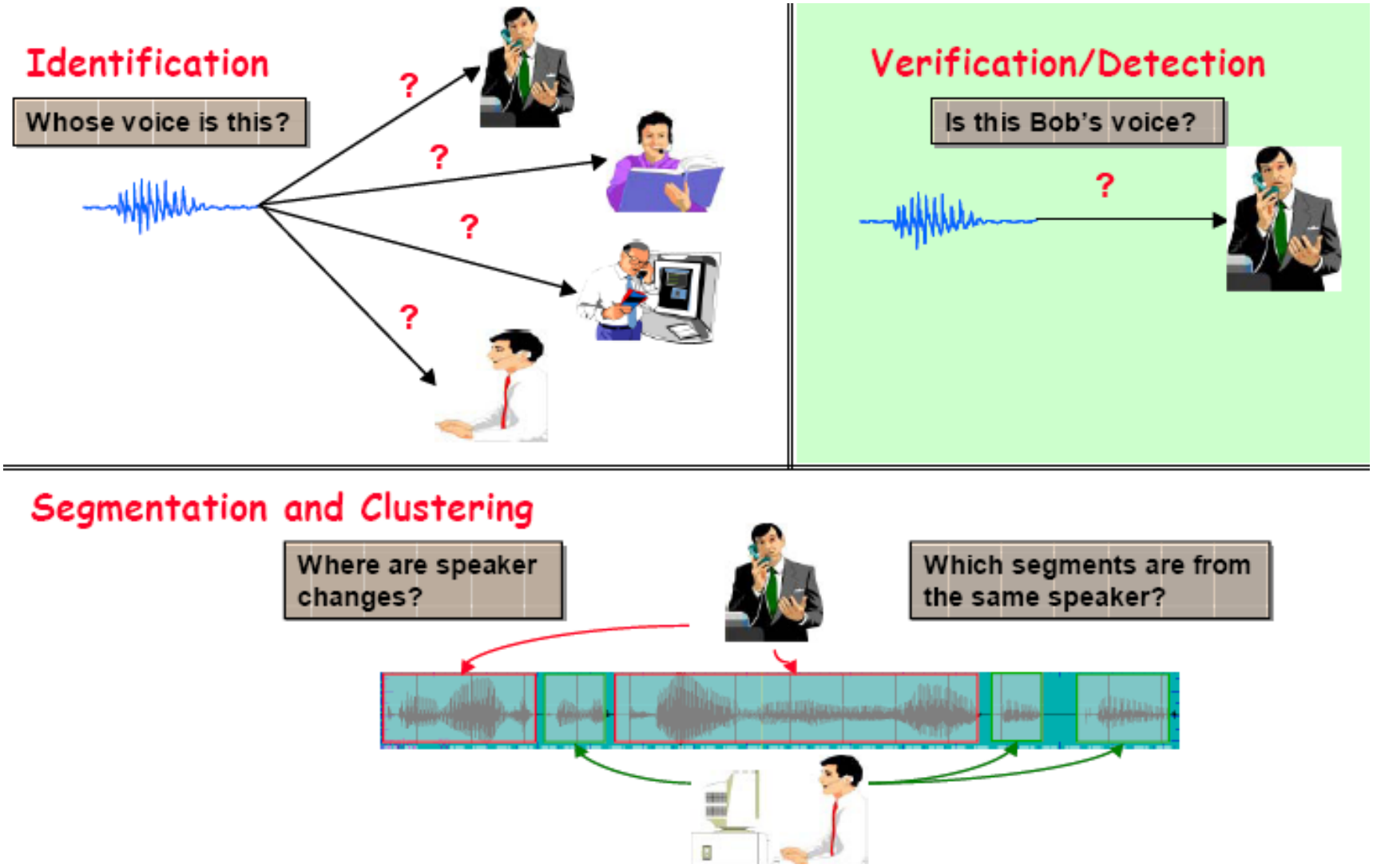
Jonathan Deutsche

Mentors:

Xinhui Zhou

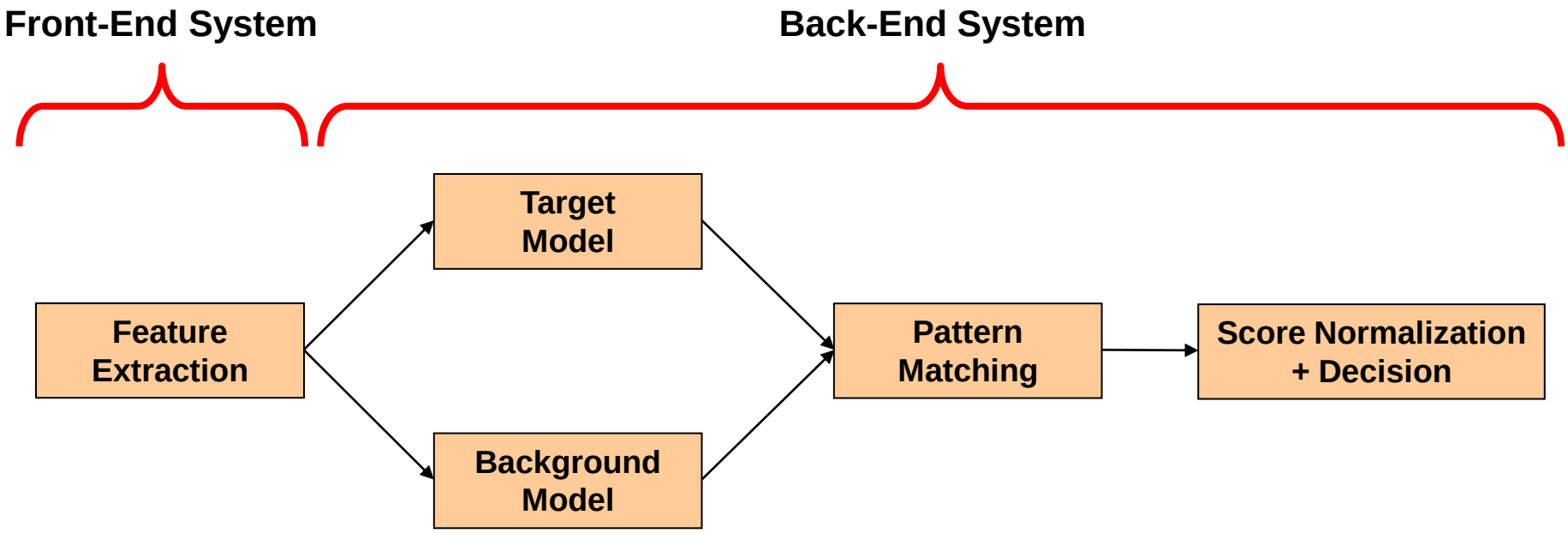
Carol Espy-Wilson

Why Speaker Recognition?



From Garcia-Romero, (2006)

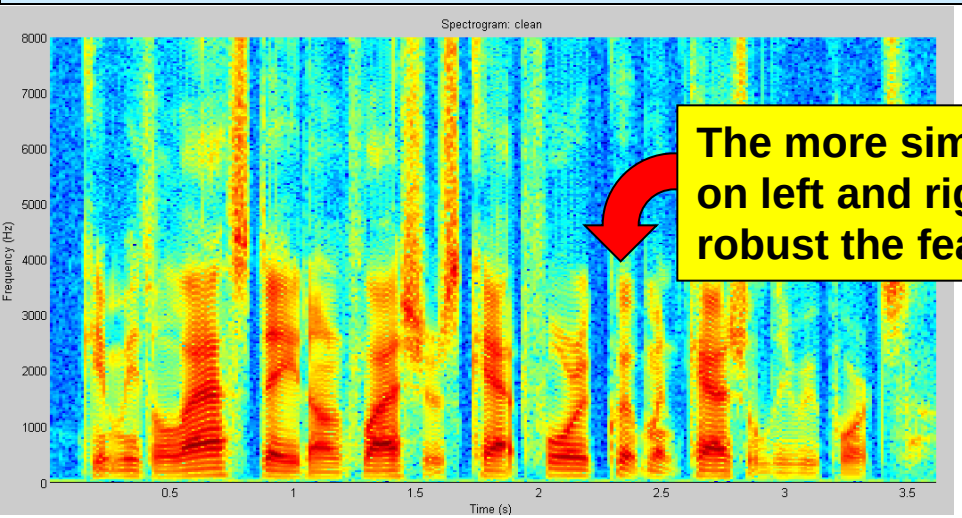
What Is Speaker Recognition?

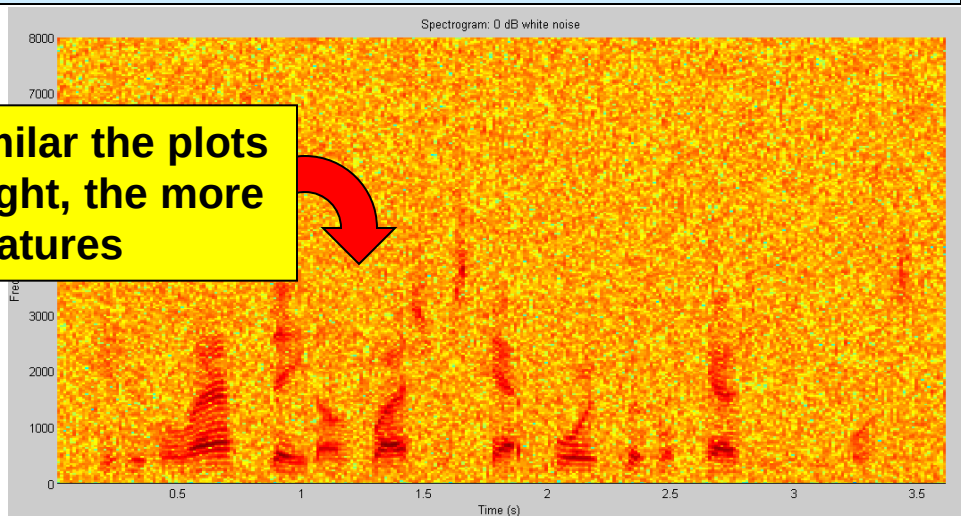


- Commonly used features
  - MFCCs
  - DCCs

BIEN 2011

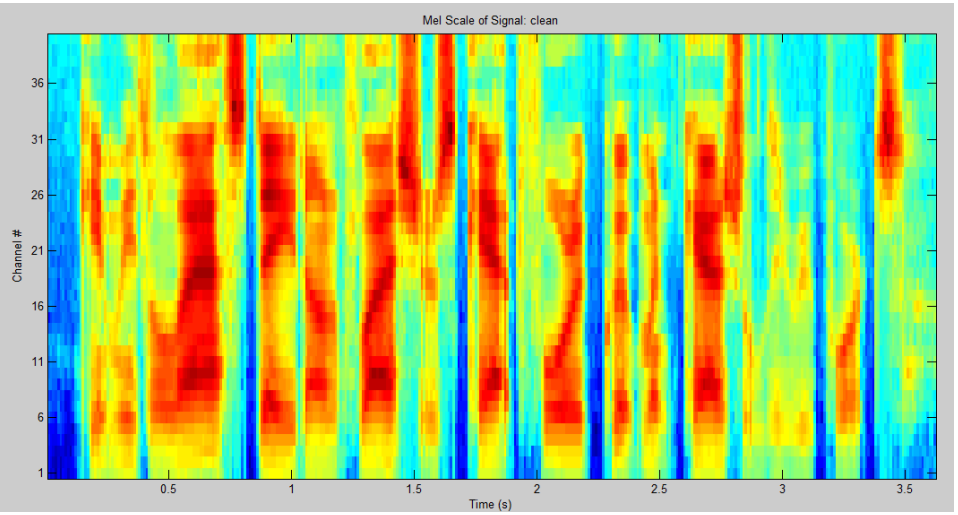
Motivation: More Robust Features

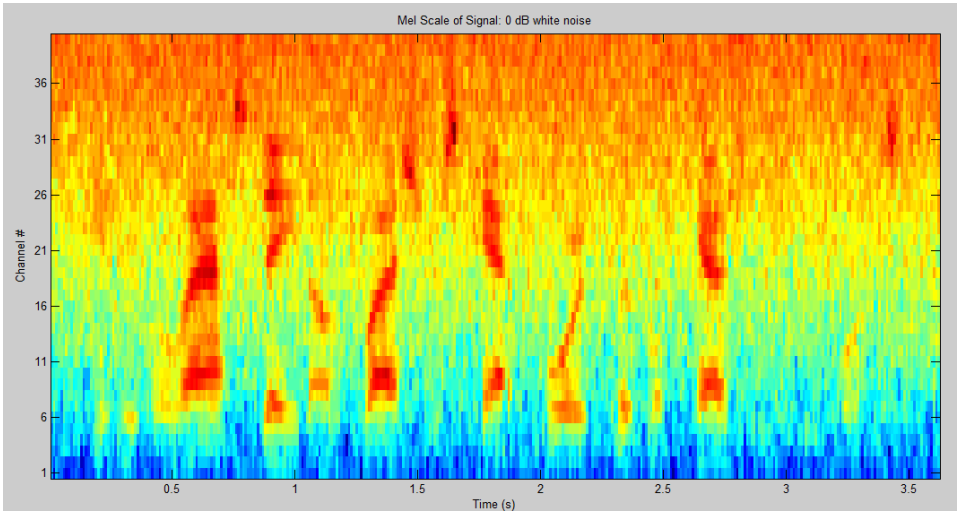




The more similar the plots on left and right, the more robust the features

Spectrogram: Clean vs. Noisy Speech



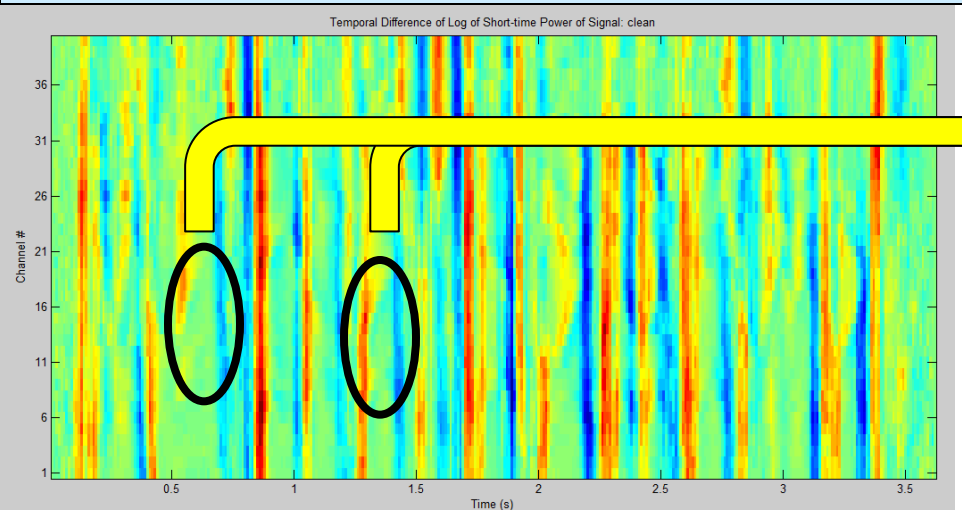


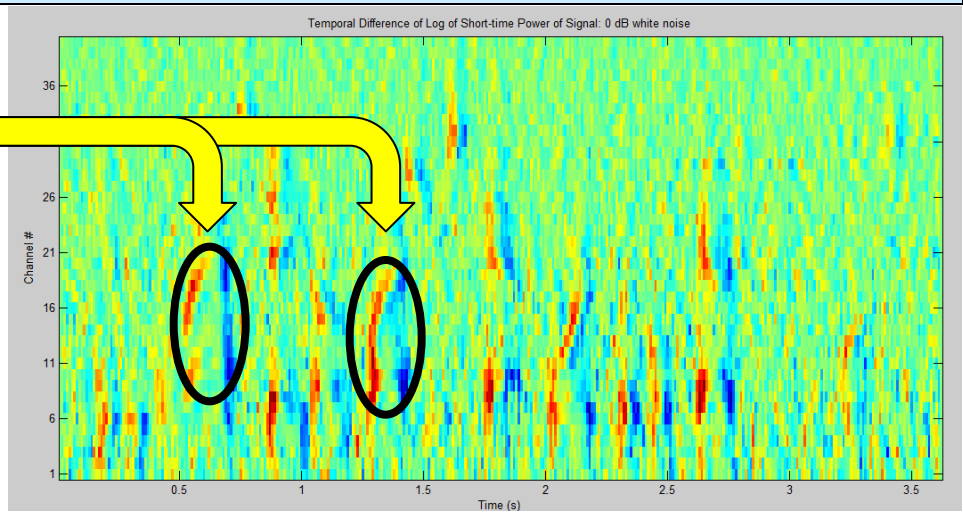
Mel-Filtered Signal: Clean vs. Noisy Speech

BIEN 2011

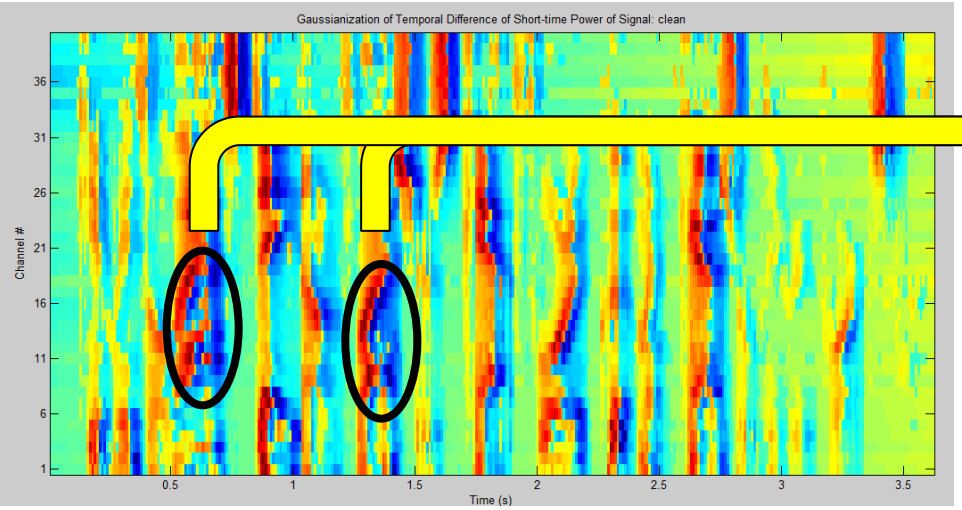
Motivation: More Robust Features -

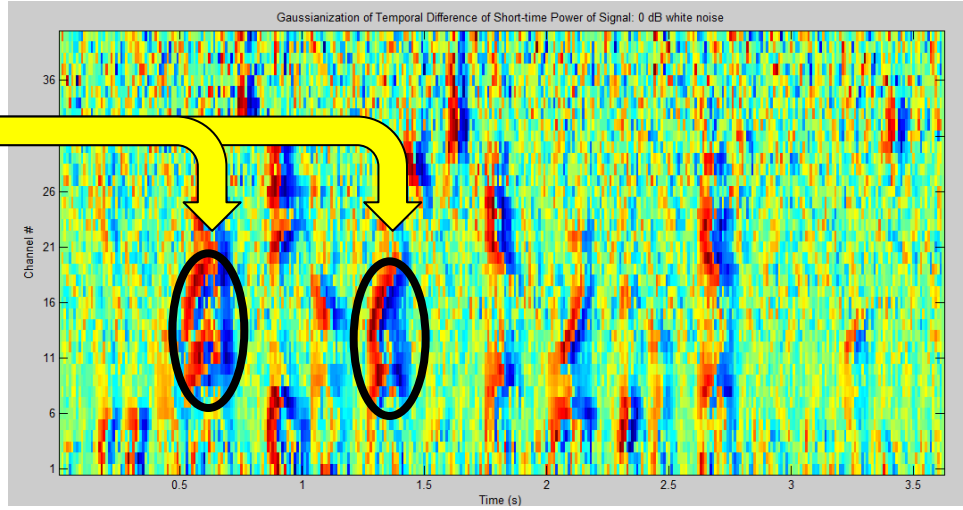
DCCs vs DSCCs





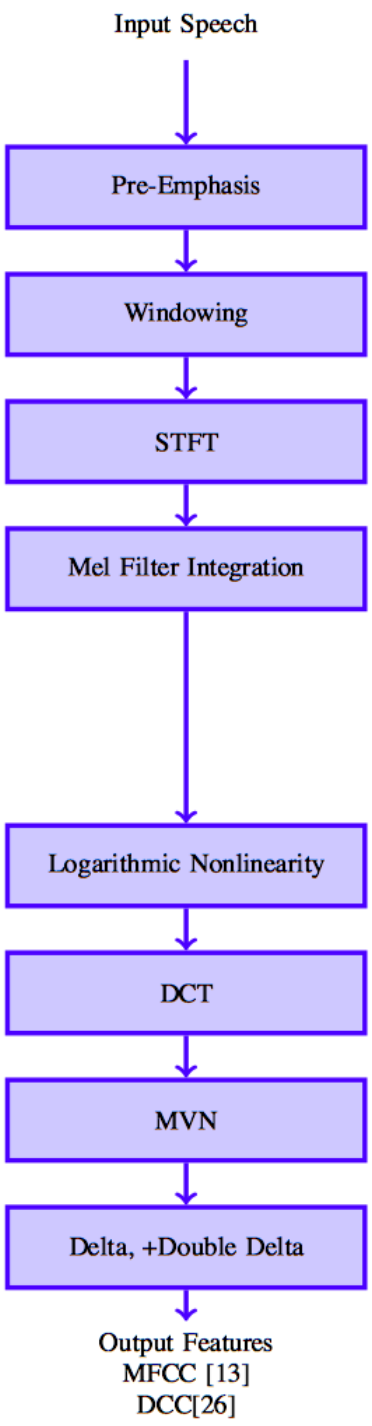
Short-time Power Plot (~DCC): Clean and Noisy Speech

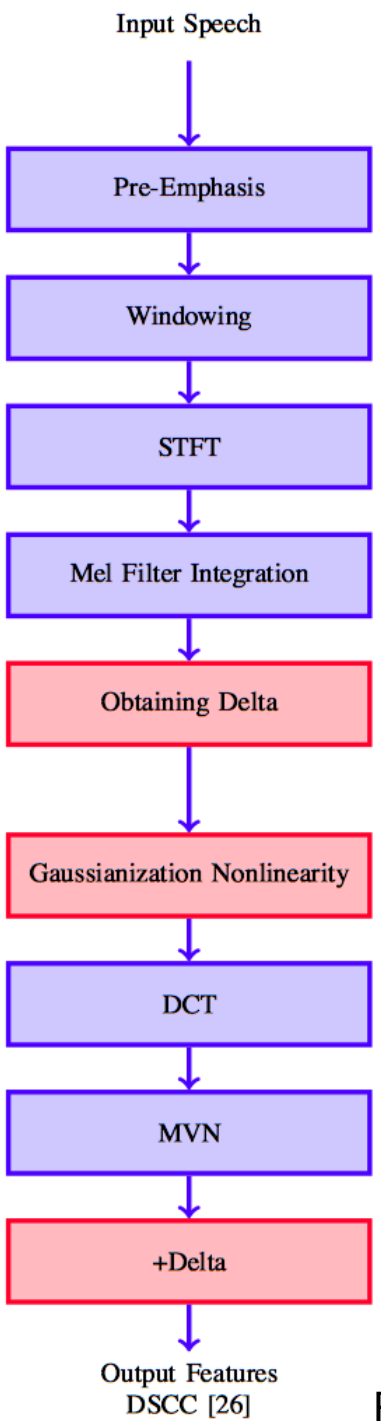




Short-time Power Plot (~DSCC): Clean and Noisy Speech

DCCs





DSCCs

From Kumar, et al. (2011)

Testing Procedures

- NIST 2008 Evaluation Plan
  - Training/testing data
    - Clean audio
    - Robustness of MFCC+DSCC's vs. MFCC+DCC's evaluated by adding white noise, babble noise, and reverberation to test files
  - Two conditions
    - Same mic used in training and test
    - Different mic used in training and test

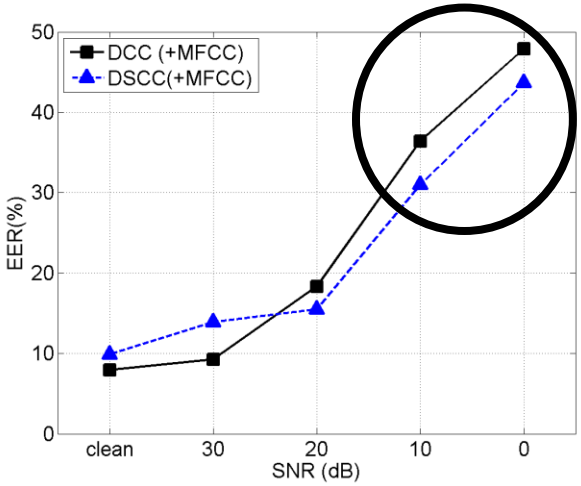
- Performance measure
  - Equal error rate (EER)

Results

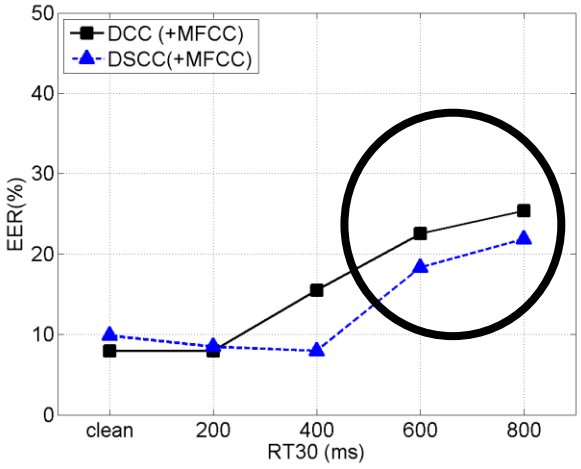
Same Mic in Training and Test

MFCC+DSCC (blue line) – lower EER than MFCC+DCC (black line) – DSCC more robust

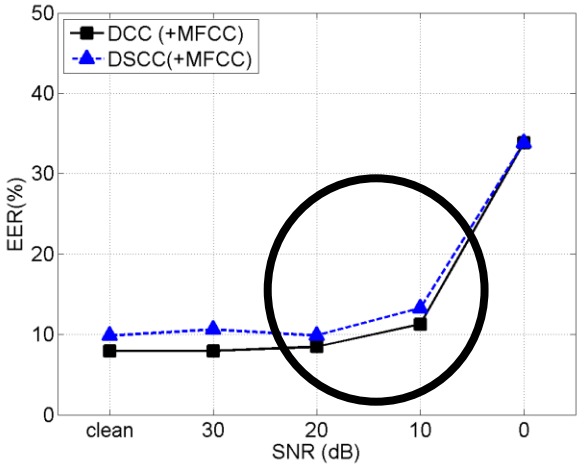
MFCC+DSCC (blue line) – higher EER than MFCC+DCC (black line) – no improvement



White Noise Added



Reverberation Added

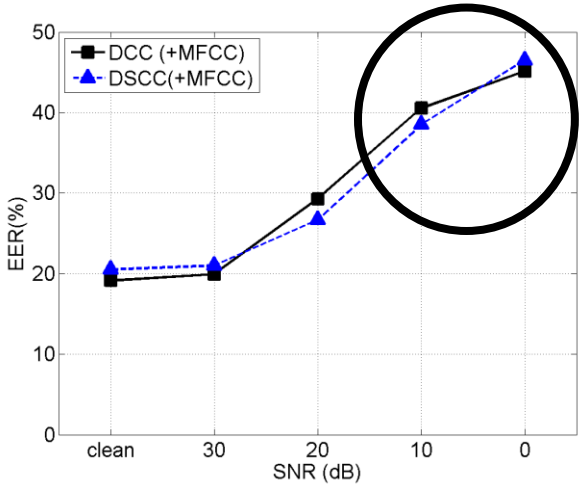


Babble Noise Added

Results (Continued)

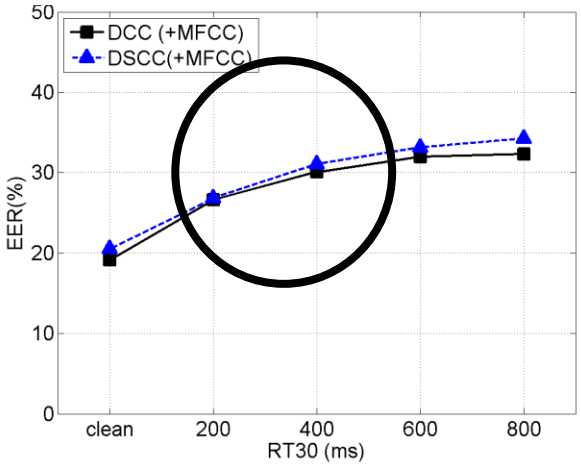
Different Mic in Training and Test

No improvement overall – MFCC+DCC and MFCC+DSCC lines are almost identical



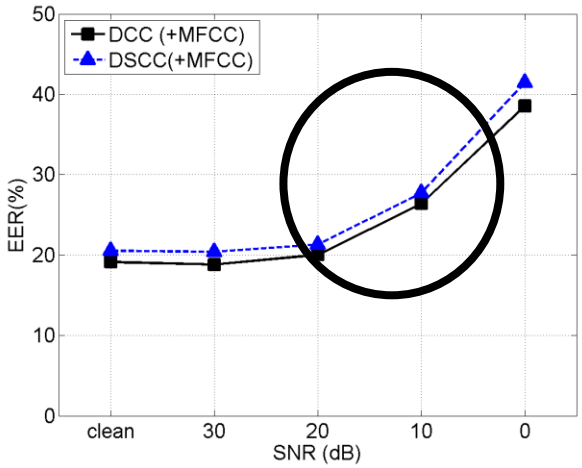
White Noise Added

X



Reverberation Added

X



Babble Noise Added

X

Future Research

- DCCs (with logarithmic nonlinearity) may be more robust to channel mismatch than DSCCs (with Gaussianization nonlinearity)
- Optimize DSCC algorithm
- Test DSCCs in conjunction with other feature types

Acknowledgements

- National Science Foundation OCI award #1063035
- Daniel Garcia-Romero
- Tarun Pruthi