Abstract

In speech & audio applications, short-term signal spectrum is often represented using mel-frequency cepstral coefficient (MFCC) computed from a windowed discrete Fourier transform (DFT). Windowing reduces spectral leakage but variance of the spectrum estimate remains high. An extension to windowed DFT is called multitaper method which uses multiple time domain windows which are called as tapers with frequency domain averaging. Then detailed statistical analysis of MFCC bias & variance is done.

For speaker verification the extracted feature is used to design a model using classifier (GMM), which implements likelihood ratio test to decide whether to accept or deny the registered speaker.

References

1. Kinnunen T., Li.,H. An overview of Text Independent Speaker recognition:from feature to
supervectors Speech communication (2009), doi:10.1016/j.specom.2009.08.009


3. Patrick Kenny1, Douglas O'Shaughnessy2, Study of Low-variance Multi-taper Features for Distributed Speech Recognition, INRS-EMT, University of Quebec, Montreal, Canada Speech Conference (2008).


6. Puming zhan, Martin Westphal, Speaker Normalization Based On Frequency Warping, Article in Interactive system laboratories, Carnegie University Germany.


10. Mohd Zaizu Ilyas, Member, IEEE, Salina Abdul Samad, Senior Member, IEEE, Aini Hussain, Member, IEEE and Khairul Anuar Ishak, Member, IEEE, “Speaker Verification using Vector Quantization and Hidden Markov Model”, the 5th student conference on research and development – scored 2007 11-12 december 2007, Malaysia.


Index Terms

Computer Science

Signal Processing

Keywords
Mel-frequency cepstral coefficient, multitaper, GMM, speaker verification, tapers.