

WS97 Acoustic Processing Group

Andreas G. Andreou (CLSP-JHU/USA)

Hynek Hermansky (OGI/USA)

Christian J. Wellekens (EURECOM/France)

Yasuhiro Minami (NTT/Japan)

Terri Kamm (DOD/USA)

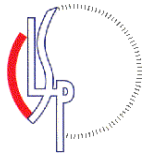
Juergen Luetin (IDIAP/Switzerland)

Dan Fain (Caltech/USA)

Sarel van Vuuren (OGI/USA)

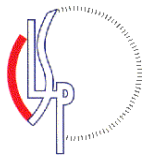
08/19/97

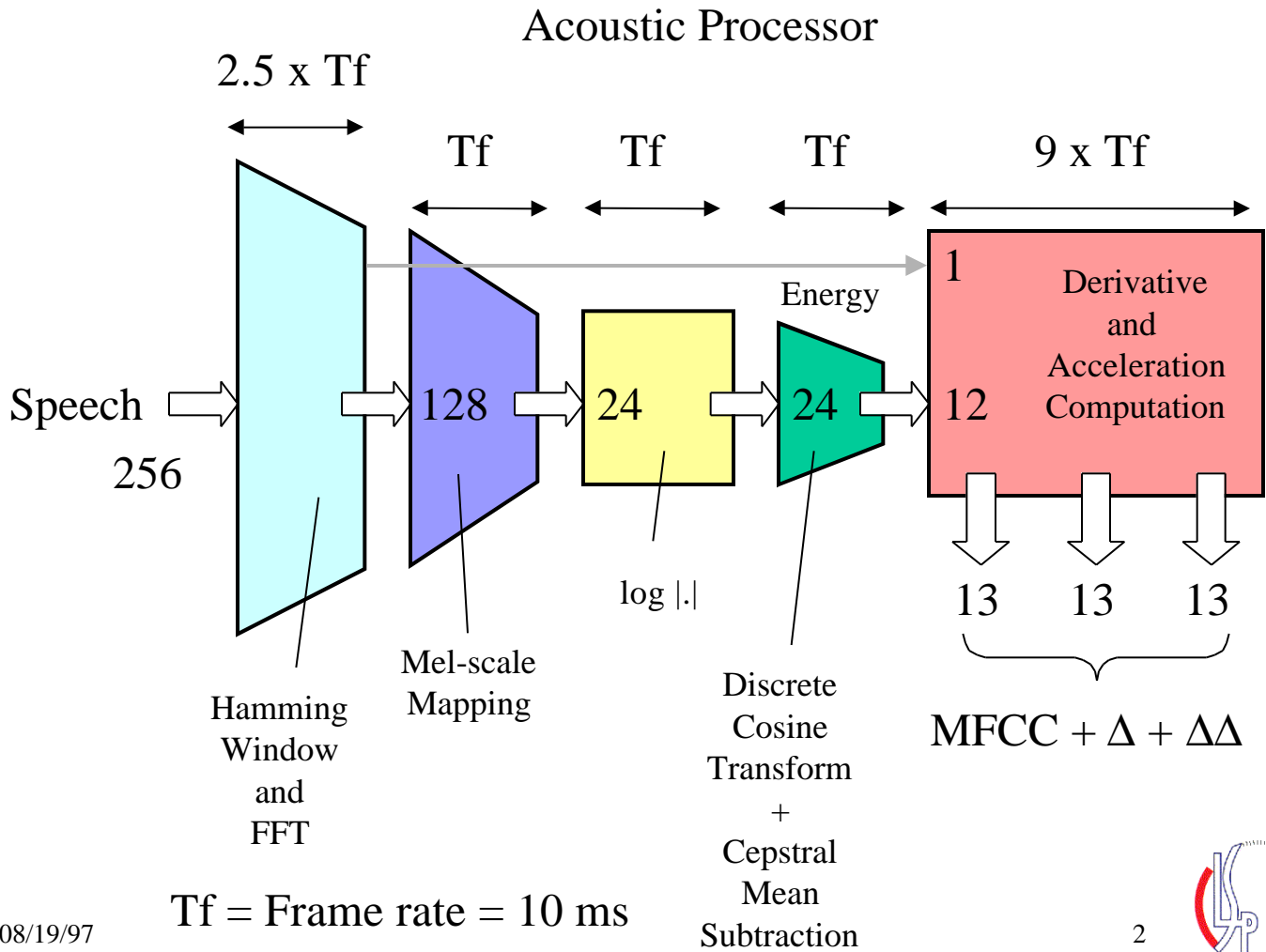
1



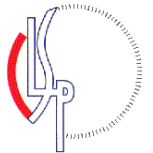
Basic Question

Can we learn from the data
how to extract features?



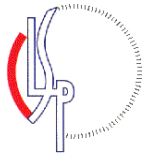


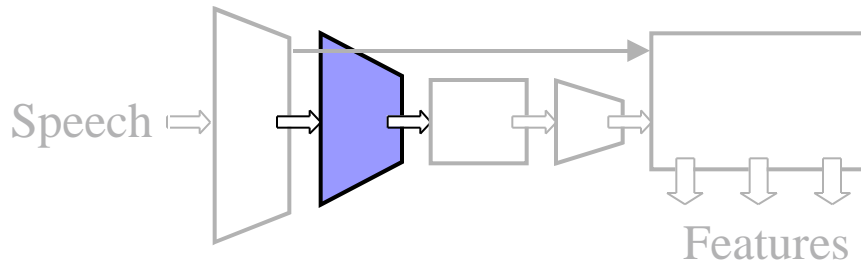
08/19/97



We have investigated

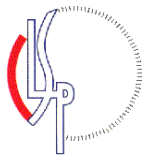
1. Non-linear frequency mapping known as Mel, and Bark scales (10 ms).
2. Optimal linear transforms on features with moderate context lengths (100 ms).
 - a. Linear Discriminant Analysis
 - b. Heteroscedastic Discriminant Analysis
3. Linear filtering of feature trajectories with long context lengths (1000 ms).
 - a. LDA-derived kernels
 - b. ML-derived kernels
 - c. HMM class-dependent kernels
4. Cepstrum-bias on per utterance basis (500 ms - 4000 ms)

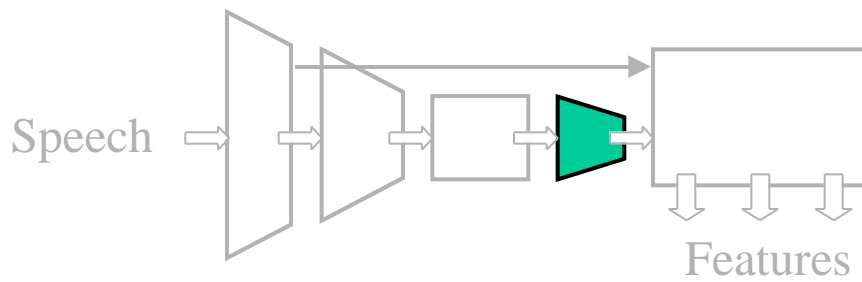




Learning the Mel-scale and Optimal VTN Mapping

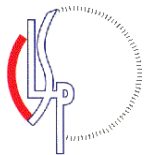
Terri Kamm, Andreas G. Andreou, Hynek Hermansky



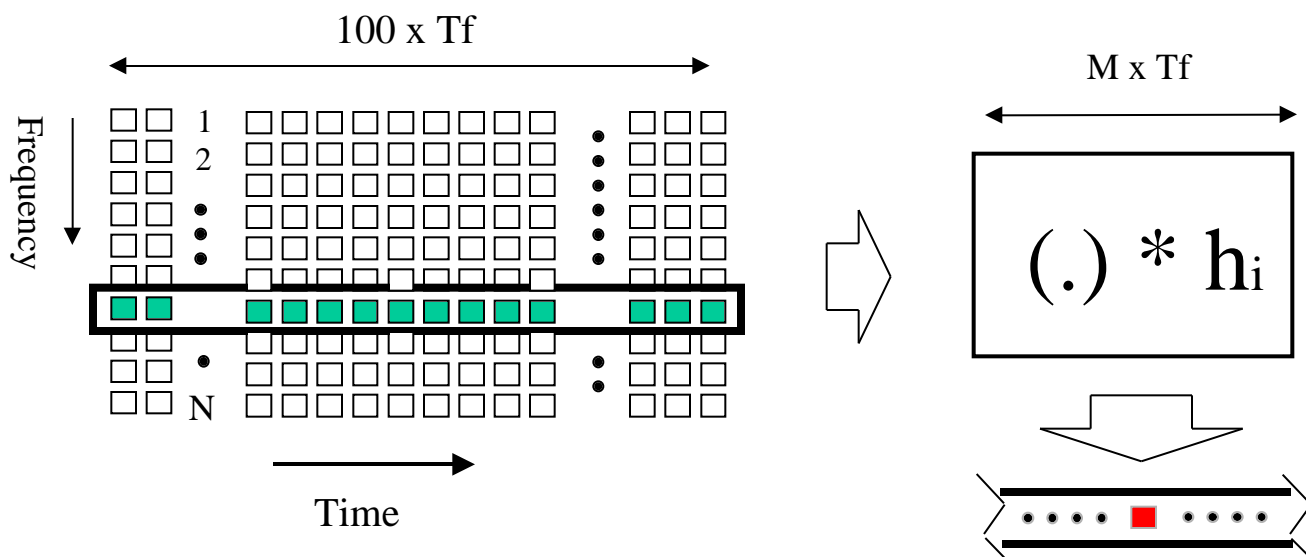


Unsupervised Learning of Cepstrum-Bias on per Utterance Basis

Yasuhiro Minami

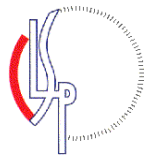


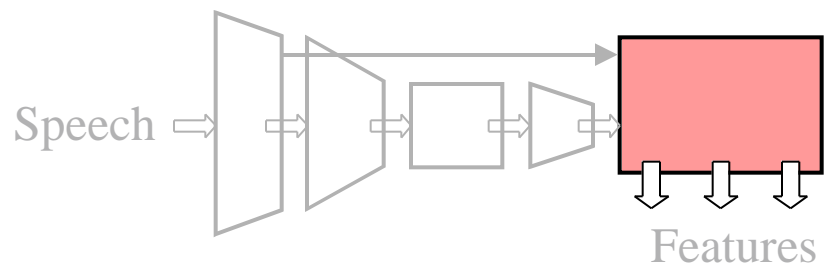
Trajectory Filtering



h_i : Convolution kernel: low pass, band pass ...

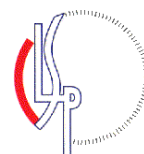
M : Kernel size

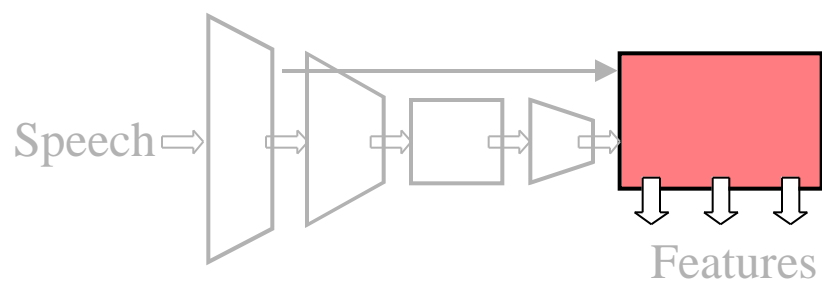




LDA-Derived Feature Trajectory Filtering

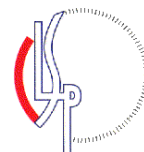
Hynek Hermansky, Sarel van Vuuren

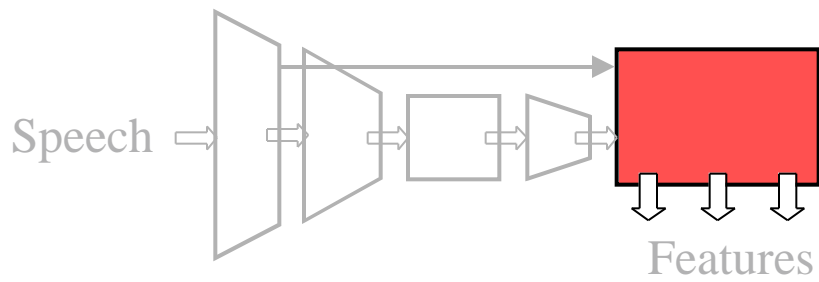




Joint Learning of HMMs and Feature Trajectory Filters

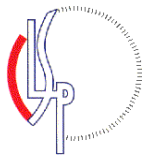
Christian Wellekens, Hynek Hermansky, Terri Kamm



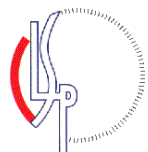
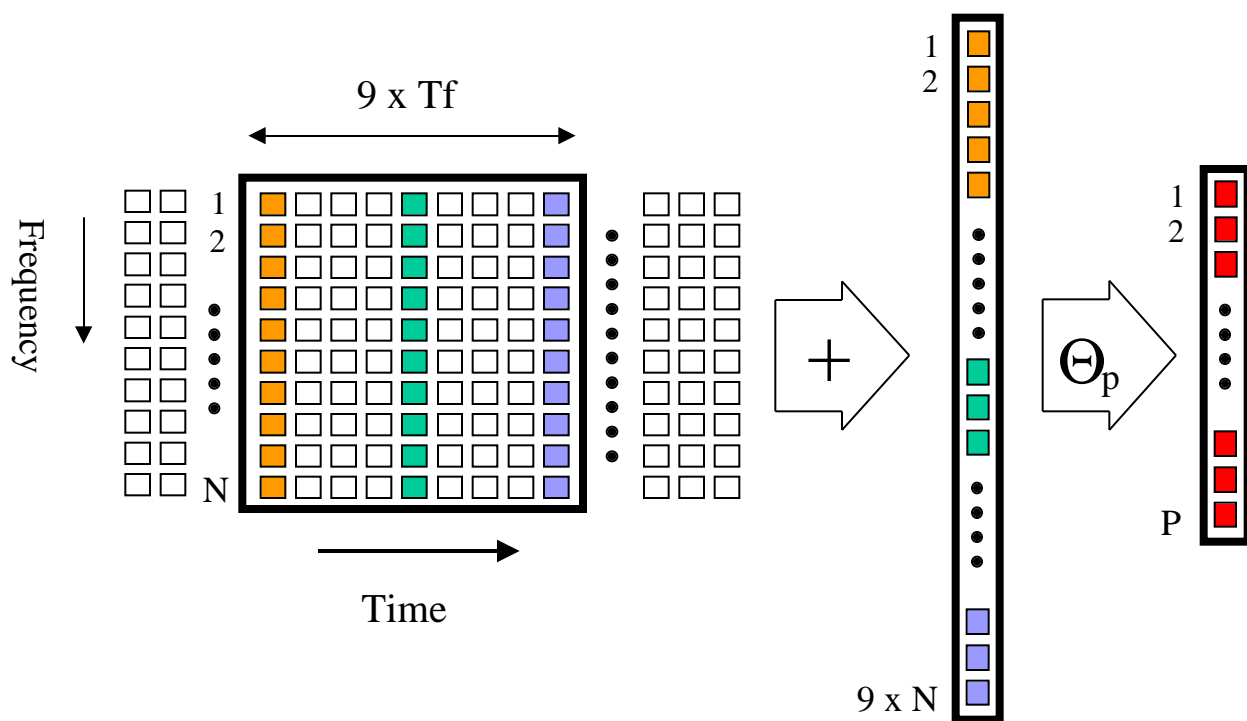


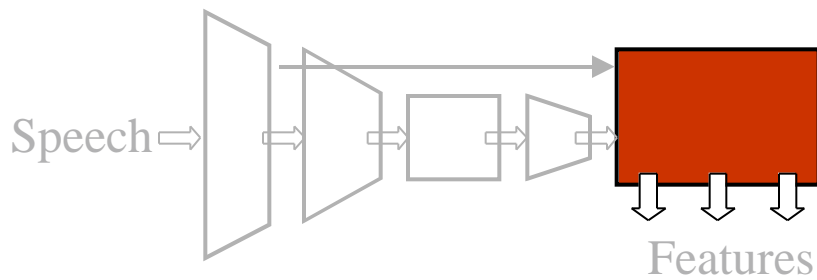
RASTA-HMMs

Juergen Luettin, Sarel van Vuuren, Hynek Hermansky



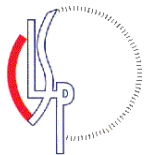
LDA and HDA





Heteroscedastic Discriminant Analysis: Optimal Feature Filtering

Dan Fain, Andreas G. Andreou, Terri Kamm



Bibliography

- [1] *J. Cohen, T. Kamm and A.G. Andreou*, “Vocal tract normalization in speech recognition: compensating for systematic speaker variability,” 129th Meeting of the Acoustical Society of America, Abstract in *J. Acoust. Soc. Am.*, Vol. 97, page 3246-3247, No.5, Pt. 2, May 1995; also Proceedings of the Fifteenth Annual Speech Research Symposium, pp. 175-178, Baltimore MD, June 12-14, 1995.
- [2] *N. Kumar and A.G. Andreou*, “Generalization of linear discriminant analysis in the maximum likelihood framework,” Proceedings of Joint Statistical Meeting, Chicago, August 1996; JHU-CLSP Technical Report #16.
- [3] *S. van Vuuren and H. Hermansky*, “Data-driven design of RASTA-like filters,” Eurospeech 97.

