

WS97 Acoustic Processing Group

Members: Andreas G. Andreou (JHU/USA)
 Hynek Hermansky (OGI/USA)
 Minami, Yasuhiro (NTT/Japan)
 Christian J. Wellekens (EURECOM/France)
 Terri Kamm (DOD/USA)
 Juergen Luettin (IDIAP/Switzerland) - HH mentor
 Dan Fain (Caltech/USA) - AGA mentor

Objectives

Main theme: *explore architectures and algorithms for acoustic processing/modeling extracting information from speech at different time scales.*

- 1a. Understand why the technique of Vocal Tract Normalization works well and place it in a firm theoretical framework.
- 1b. Revisit the notion of “Mel-scale” with fixed parameters
2. Apply recent theoretical formulation of Heteroscedastic LDA for optimal dimensionality reduction to the switchboard problem.
3. Investigate further recent ideas related to multiple streams/multiple channels acoustic processing.

Preparations

- AGA, HH, TK, DF had a meeting in February with some discussion of issues pertaining to the group.
- DF the eight week student spend the Spring semester at JHU learning the tools
- AGA, TK, DF participated in an advanced speech processing seminar course whose goal was to bring members of the group up to speed in the areas of expertise of the other members (For example MLP, LDA, VTL).
- Code preparation for HLDA and VTL

Plans until the workshop

- Continue work on writing code that will be used during the workshop (DF, TK)
- Discussion and possible visit at JHU by CW to discuss further tool requirements (for example tools for MLPs and other VSMs).

References related to work in the group

- [1] H. Bourlard, S. Dupont, and C. Ris "Temporal Integration of Multi-Stream Inputs in Speech Recognition Systems" Working paper, <http://www.clsp.jhu.edu/ws96/ris/multi/multi.html>.
- [3] N. Kumar and A.G. Andreou, "Generalization of linear discriminant analysis in the maximum likelihood framework," Proceedings of Joint Statistical Meeting, Chicago, August 1996.
- [3] H. Hermanski, Working paper, the Keele Workshop on Auditory Basis of Speech Perception, held in July 1996.
- [4] 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.