Advanced Speech Lab

Course 11-753

Tanja Schultz & Thomas Schaaf

Goals

- Learn how a large vocabulary speech recognition system works
- Learn the steps involved in building an acoustic model
- Learn how to build a language model
- Learn how to decode an audio signal
- Learn how to measure performance

General overview

- Learn basic Tcl/Tk
- Build a training database
  - What data?
- Components of a speech recognition system
  - Preprocessing
  - Phonemes, Dictionary
  - HMM topology
  - Probability density for acoustic modeling
- Choose preprocessing of audio signal
- Initial description for context independent system

General overview (2)

- Train a context independent system
  - Linear Discriminant Analysis
  - Sample extraction
  - K-means initialization
  - Label/Viterbi training
- Build a 3-gram language model
  - What is back-off
  - How to use classes
- Do a first decoding with context independent acoustic models

General overview (3)

- Move to context dependent phoneme models
- Build a context dependent phonetic decision tree
  - Optimization criteria
  - Statistic required
  - How splitting is performed
- Train context dependent acoustic model
- Test context dependent model

General overview (4 optional)

- Adaptation methods
  - VTLN
  - MLLR
  - Cross Adaptation
- Speaker Adaptive Training (SAT)
- Confidence Measure
- Rover

Schedule (tentative)

11/17/2005		Introduction, get started
11/24/2005	-------	Thanksgiving
12/01/2005		Janus and Tcl/Tk Introduction, Objects
12/08/2005		Pronunciation dictionary, database, word error rate, more Tcl/Tk
12/15/2005		Preprocessing
12/22/2005		Preprocessing, Acoustic Modeling
12/29/2005	-------	Winter Break
01/05/2006		Acoustic Modeling
01/12/2006		Training CI
01/19/2006		Training CD
01/26/2006		Language Modeling
02/02/2006		Search, Scoring
02/09/2006		Advanced Topics (optional)