E6820 Speech & Audio Processing & Recognition **Homework 5**

Assigned: Tuesday 2001-02-13 Due: Tuesday 2001-02-20

Dan Ellis <dpwe@ee.columbia.edu>

Background reading:

Read chapters 14 through 18 in Gold & Morgan, the section on auditory perception. We will not dwell long on physiology, though it's important to be aware of the basics. Pitch perception is a particular interest of Gold's, so it gets perhaps more space than it deserves, but it's an interesting case study of the difficulty of tying down perceptual phenomena.

Reading assignment:

"Critical Band Width in Loudness Summation," E. Zwicker, G. Flottorp and S.S. Stevens. This is a classic paper describing further experiments into the phenomenon of 'critical bands'. In this case, two sounds with equal *power* (a physical quantity) have different *loudnesses* (a perceptual quantity) if the bandwidth occupied exceeds some minimum. This paper also gives you some idea of the level of care and detail required for psychoacoustic experiments. Add a summary to your web page.

http://www.ee.columbia.edu/~dpwe/courses/e6820-2001-01/papers/ZwicFS57-crband.pdf

Practical assignment:

This week you will compare different feature sets and classification schemes using the formant frequency example files from last week, fmtU.txt, fmtO.txt and fmtA.txt. The course web site contains the Matlab scripts for training simple one-hidden layer MLPs (nntrain.m) and Gaussian mixture models (gmmest.m), as well as associated scripts for applying the models and visualizing the results.

First, divide your data into two halves, one for training, the other for test. (You should select these subsets at random because the data in each file is ordered so that the extreme cases tend to lie near to one another, but you want them equally represented in training and test). Then, construct two different classifiers from the training data, one using a neural net, the other using Gaussians or Gaussian mixtures. Measure the error rate of your models by using them to classify the test-set portion of the data (which they have not seen during training) and counting the number of correct and incorrect classifications on this set. Initially, use just the F1 and F2 columns from the data file (as with last week's assignment), then see how including the 3rd column (F3) improves results.

Suggest an explanation for variations in performance that you observe.

http://www.ee.columbia.edu/~dpwe/courses/e6820-2001-01/matlab/

Problems:

From the chapter-end exercises in Gold and Morgan:
Problems 8.1, 8.3 and 9.3.