Reading assignment:

“Combined monaural and binaural localization of sound sources,” W. Chau & R.O. Duda. Add a summary to your web page, including any questions or aspects you found particularly interesting.


Practical assignment:

This week we will use time differences between microphones to detect activity of different sources.

In the Meeting Recorder section of my Sound Examples web site, you will find the stereo soundfiles pzm12a.wav and pzm34a.wav, each containing two channels of a 5 minute recording made from a set of four microphones arranged along the middle of a conference table during a meeting. The function xcorrpeak.m will take two waveform segments and return the sample index of the peak of their cross-correlation i.e. if the second signal is a delayed version of the first, the result will be the delay in samples.

(a) Use this function to calculate the best-fitting timing skew for every half-second frame from the two waveforms, and make a scatter plot of skew(chan1→chan2) versus skew(chan3→chan4).

(b) Can you see the distinct clusters corresponding to individual speakers? Make a rough classification to generate ‘active speaker’ labels for every time frame.

(c) How well do your labels agree with the manual transcription in transcript.txt? Be quantitative.

Project:

Be sure to keep up your steady project progress as the semester moves into its last few weeks.