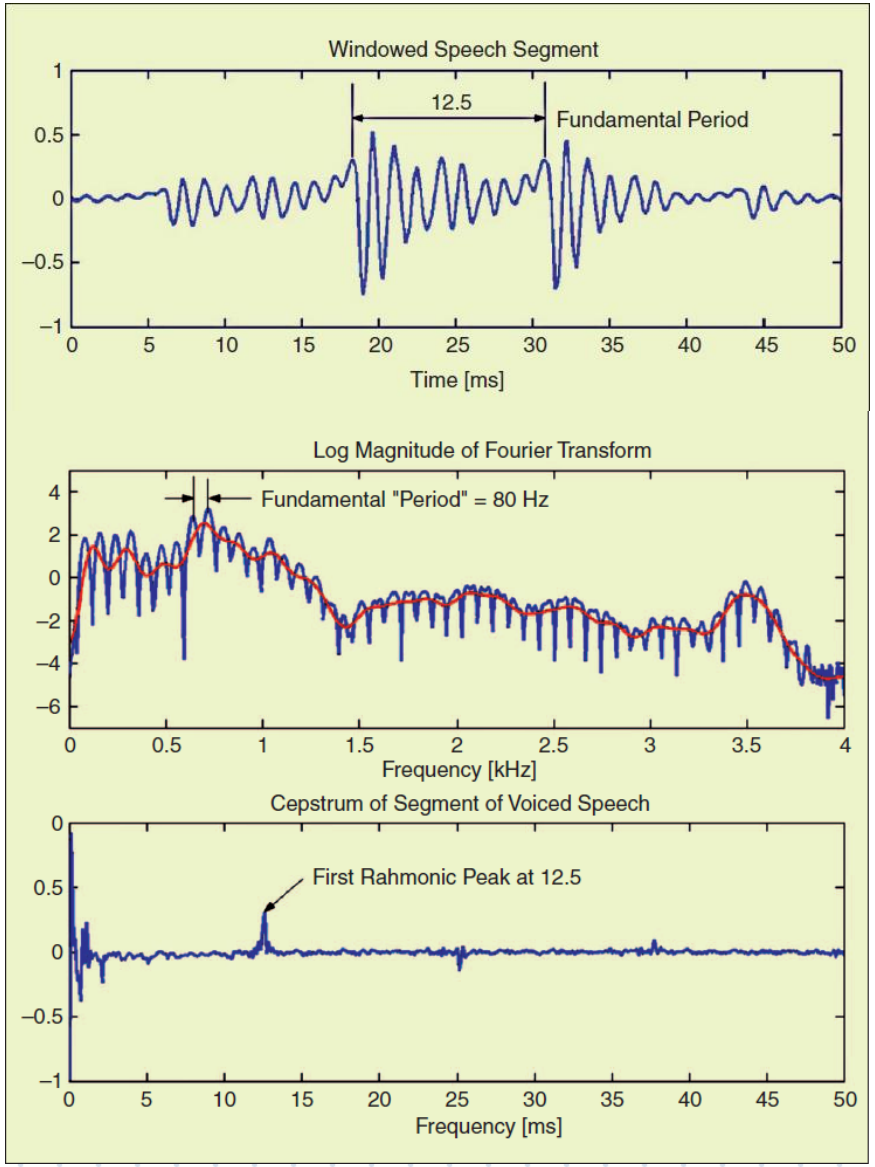
Cepstrum Overview



▲ *1. A segment of voiced speech with a Hamming window during a voiced (vowel-like) time interval of 50 ms. Short-time analysis of speech involves the analysis of a succession of such segments taken sequentially in time. At this particular time in a vowel sound, the pitch period is approximately 12.5 ms.*

▲ *2. Log magnitude (in blue, upper graph), lowpass liftered (linearly filtered) log magnitude (in red, upper graph), and cepstrum (in blue, bottom graph) of the segment of voiced speech illustrated in Figure 1. The rapidly varying curve in the upper graph is the log magnitude of the discrete-time Fourier transform of the segment of speech in Figure 1. The lower plot is the cepstrum, i.e., the inverse discrete Fourier transform of the log magnitude in the upper plot. Note the peaks at "rahmonics" of 1/80 = 12.5 ms, the fundamental quefrency of the quasi-periodic ripples in the upper graph. As can be seen by comparing the speech segment in Figure 1 to this cepstrum, this fundamental quefrency is also the period (pitch period) of the time waveform. The search for such peaks in the cepstrum was the basis for Noll's pitch detection algorithm.*

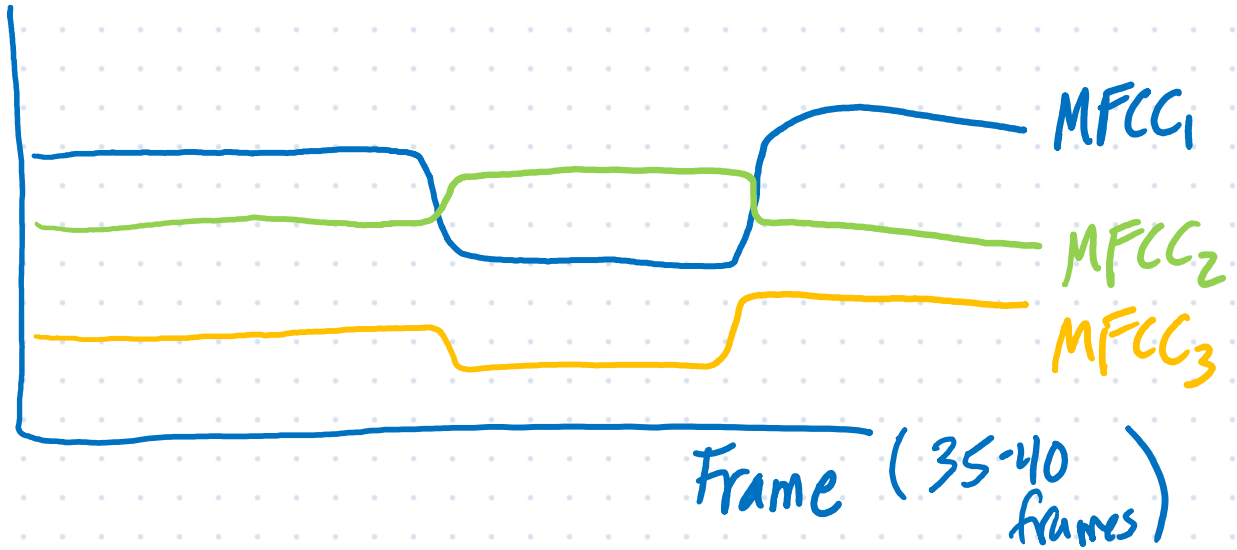
Figures from: Alan V. Oppenheim and Ronald W. Schafer, "From Frequency to Quefrency: A History of the Cepstrum," *IEEE Signal Processing Magazine*, **21**(5), pp. 95-99;106 (September 2004).

Spoken Digits 0 through 9 (in Arabic!)

[UCI Machine Learning Repository: Spoken Arabic Digit Data Set](https://archive.ics.uci.edu/ml/datasets/Spoken+Arabic+Digit)
(https://archive.ics.uci.edu/ml/datasets/Spoken+Arabic+Digit)

Be sure to read the data set description – it explains how to parse the data!

More complete project guidance will be provided by early next week



← Model this (> 2D)