Speech Processing

Speech Processing

References (Figures shown in class and posted on TritonEd from (1)-(3))

(1) J.L. Flanagan. Speech Analysis, Synthesis, and Perception. Springer-Verlag (2nd Ed) (1972).

(2) L. Rabiner and B. Gold. Theory and Application of Digital Signal Processing. Prentice Hall (1975).

(3) L. Rabiner and R. Schafer. Digital Processing of Speech Signals. Prentice Hall (1978).

(4) J. Markel, "Digital Inverse Filtering - A New Tool for Formant Trajectory Estimation," IEEE Trans. Audio and Electroacoustics, AU-20: 129-137 (1972).

Speech Processing

Concatenated tube model of the speech production process

Reflection coefficients characterize the acoustic impedance contrast between adjacent tubes with different diameters.

Speech Processing

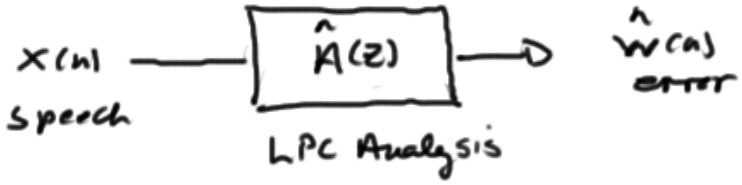
Speech Transmission

Direct digitizing of speech: $f_s \sim 8$ kHz, ~ 8 bits/sample

—> ~ 64 kb/s data rate.

(speech bandlimited to ~ 4-5 kHz)

LPC Analysis



segments ~ 25 ms or ~ 40 segments/sec ("frames")

For each segment or frame need to obtain:

- $\hat{A}$ vector (~ 10 - 14 coefficients)
- voiced/unvoiced decision
- pitch period
- amplitude

Results in bit rates of ~ 1.0 - 2.4 kb/s data rate