# MORE AUDIO COMPRESSION

LPC
Physical Models – Analysis/Synthesis
Music Notation

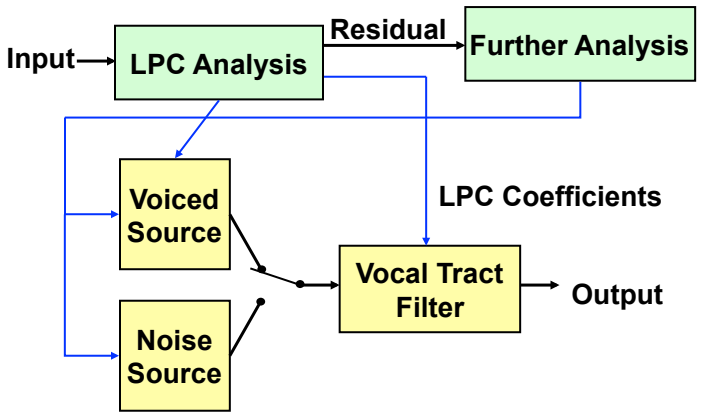




## More Audio Compression

- LPC: Linear Predictive Coding
  - voice is source (voiced/unvoiced) + filter
  - voice pitch and filter coefficients change slowly
  - 1-2K bits/s (Sambur), typical 2400-7200 (Rabiner & Schafer)
- Another example of intersample redundancy
  - Here, we have an "object model" that is used to estimate samples

# Linear Predictive Coding in Practice







# Physical Models / Speech Synthesis

- Future speech compression technology
- Muscles control is low bandwidth
- Speech sounds highly constrained
- Idea: transmit control signals for speech then simulate acoustics of speech production
- May need one-time transmission of parameters to characterize speaker

## Audio Compression Examples

*Note: This slide is not in the video lecture. If you would like to hear some speech compression examples, there is a link to them in the online course page containing the video lecture.*

- Reference (uncompressed, 16-bit)
- Downsampled to 11kHz, 16-bit
- 8-bit uLaw
- IMA ADPCM
- TrueSpeech 8.5 (8.5 Kbits/sec, proprietary coder)




## Music

- Music Notation
  - compact, symbolic representation
  - does not capture performance information
  - expressive "performance" not fully automated
- Performance Information
  - MIDI bandwidth is 3KB/s, or 180KB/min
  - More typical: 3KB/minute, 180KB/hour
    - Complete Scott Joplin: 1MB
    - Output of 50 Composers (400 days of music): 500MB (1 CD-ROM)
  - synthesis of acoustic instruments is a problem

# Summary

- Three kinds of redundancy:
  - Coding
  - Intersample
  - Psycho-Perceptual
- uLaw, ADPCM, etc. – simple, fast, but not high quality or high compression
- MP3 and related schemes – general, high quality and compression, fairly high computation
- Analysis to Model and Resynthesis – special purpose (e.g. speech, musical instruments), potentially very high quality and compression, but still research