18-899 Special Topics in Signal Processing

Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu

Lecture 8

MPEG-1 Audio

MPEG-1 Audio

- Outline
  - Background
  - Psychoacoustics
  - Subband coding
  - Layer I and II
  - Layer III
  - Frame structure and packetization

MPEG-1 Audio

- ISO/IEC 11172-3 (1988~1991)
- First high quality audio compression standard
- CD quality two-channel audio at 256 kbits/s
  - CD: 44.1 kHz × 16 bits × 2 = 1.411 Mbits/s

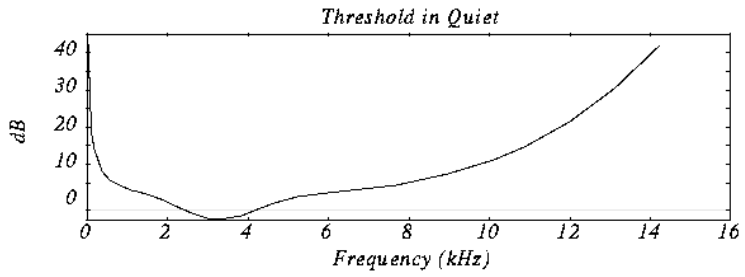
	Frequency Band (Hz)	Sampling Rate	Bits per Sample	Raw Bitrate
Telephone Speech	300~3400	8	8	64
Wideband Speech	50~7000	16	8	128
Mediumband Audio	10~11000	24	16	384
Wideband Audio	10~22000	48	16	768

Quality Demonstration

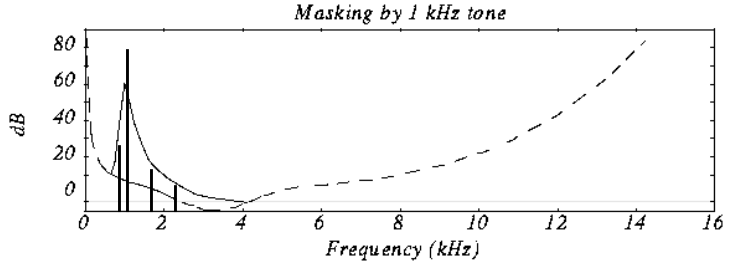
- MPEG-1 Audio (Layer II)
  - Stereo 44.1 kHz at 64 kbits/s
  - Stereo 44.1 kHz at 128 kbits/s
  - Stereo 44.1 kHz at 192 kbits/s
  - Stereo 44.1 kHz at 256 kbits/s

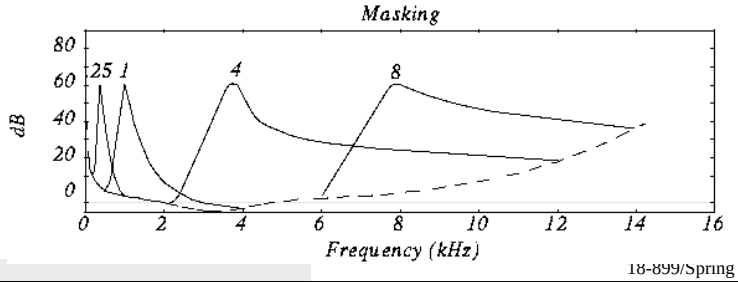
Psychoacoustics

- Threshold in quiet

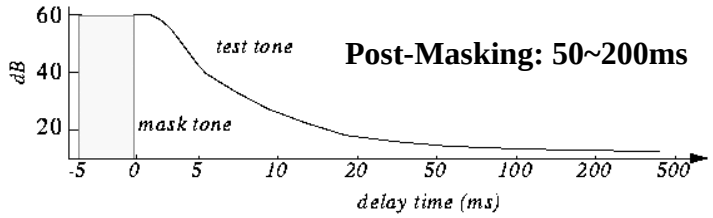


Frequency Masking

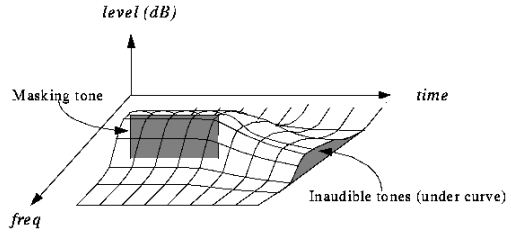




Temporal Masking

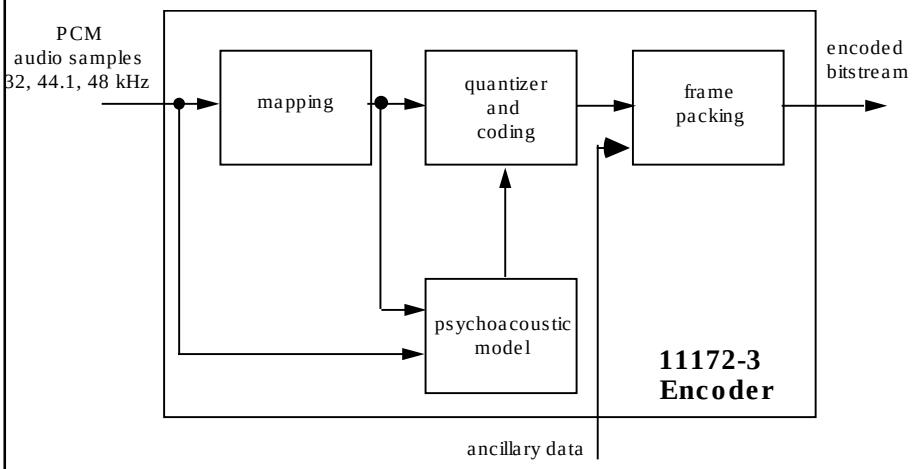


Post-Masking: 50~200ms

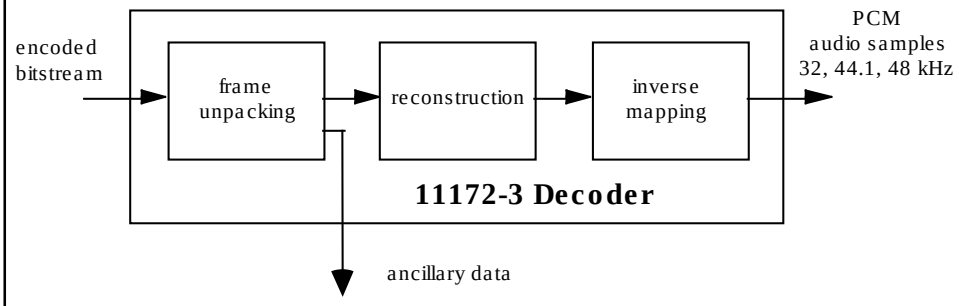


Also Pre-Masking (much shorter)

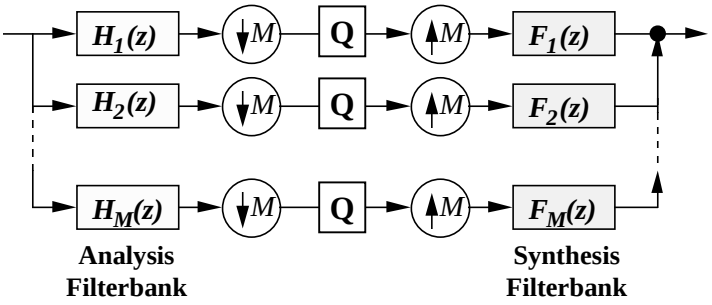
Encoder Block Diagram



Decoder Block Diagram

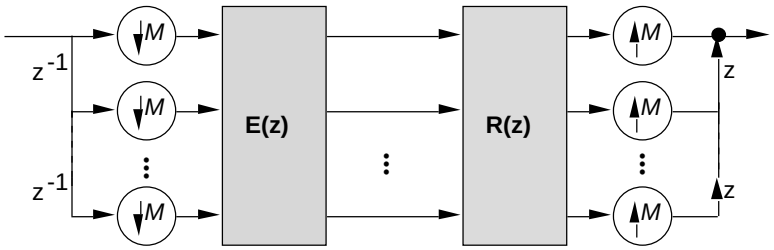


Mapping: Subband Coding



- Critical downsampling
- Q should be based on signal-to-masking ratio (SMR)
- Ear's critical bands are not uniform, but logarithmic

Polyphase Filterbank

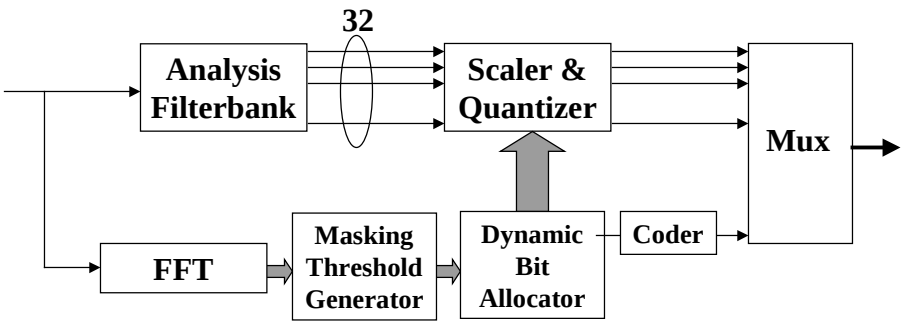


- Alias cancellation and perfect reconstruction

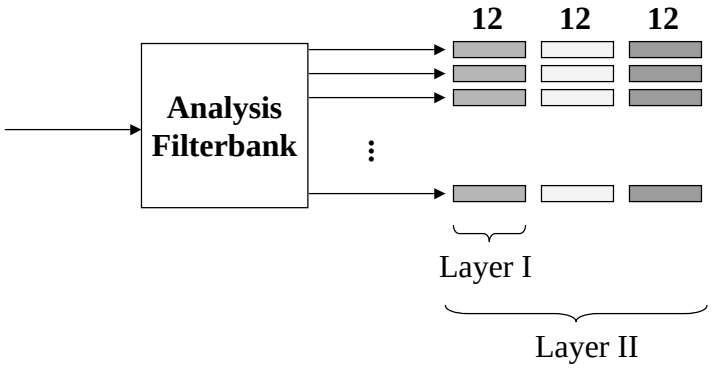
Layers

- Increasing complexity, delay, and quality
- Layer I
  ~384 kbits/s for perceptually lossless quality (4:1)
- Layer II
  ~192 kbits/s for perceptually lossless quality (8:1)
- Layer III
  ~128 kbits/s for perceptually lossless quality (12:1)

Layer I and II Encoder

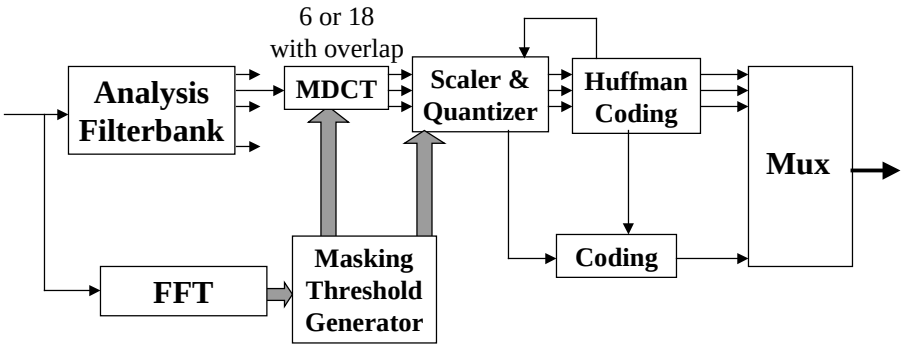


Block-Based Coding



- 12 samples for Layer I, 36 samples for Layer II
- Block companding: Each block normalized by *scalefactor*
- For Layer II, up to 3 scalefactors, with 2-bit scalefactor select
- Each block receives one bit allocation

Layer III Encoder



New Features in Layer III

- Modified DCT (MDCT)
  - DCT with overlap
  - Long/short window switching
    - Short for better temporal resolution (to prevent pre-echoes)
    - Long for better frequency resolution
- Nonuniform quantization
- Entropy coding
  - Run-length and Huffman coding
- Bit reservoir (buffer)

Frame Structure

Header Info	Side Info	Subband Sanples	Aux Data

- Header info: Sync bits, system info, CRC (cyclic redundancy code)
- Side info: bit allocation, scalefactor, (and scalefactor select for Layer II and III)
- Subband samples: 32 × 12 for Layer I, 32 × 36 for Layer II and III
- Packetization: 4-byte header, 184-byte payload

Stereo Redundancy Coding

- Four modes: mono, stereo, dual with two separate channel, joint stereo
- In joint stereo mode
  - Human stereo perception > 2kHz is based on envelope
  - Intensity stereo coding > 2kHz
    - Encode (L + R)
    - Assign independent left- and right- scalefactors
- Layer III supports (L+R) and (L–R) coding

References

- Peter Noll, "MPEG digital audio coding," *IEEE Signal Processing Magazine*, Sept. 1997, pp. 59-81
- D. Pan, "A tutorial on MPEG/Audio compression," *IEEE Trans. on Multimedia*, vol. 2, no. 2, 1995, pp. 60-74