

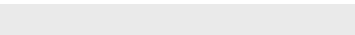
•
•
•
•
•
•
•
•
•
•

18-796



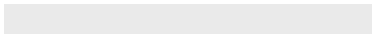
Multimedia Communications:
Coding, Systems, and Networking

Prof. Tsuhan Chen
tsuhan@ece.cmu.edu



•
•
•
•
•
•
•
•
•
•

MPEG Audio



Outline

- Basics
 - Psychoacoustics
 - Subband coding
- MPEG-1 audio
 - Layer I and II
 - Layer III
 - Frame structure and packetization
- MPEG-2 audio
 - Multichannel audio
 - Backward compatible coding
 - Non backward compatible coding

18-796/Spring 1999/Chen

Digital Audio

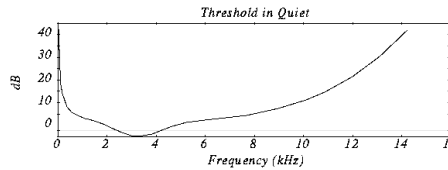
	Frequency Band (Hz)	Sampling Rate (kHz)	Bits per Sample	Raw Bitrate (kbits/s)
Telephone Speech	300~3400	8	8	64
Wideband Speech	50~7000	16	8	128
Mediumband Audio	10~11000	24	16	384
Wideband Audio	10~22000	48	16	768

– CD: $44.1 \text{ kHz} \times 16 \text{ bits} \times 2 \text{ channels} = 1.411 \text{ Mbits/s}$

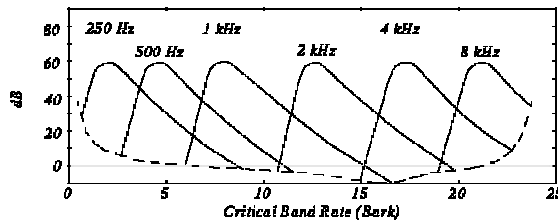
18-796/Spring 1999/Chen

Psychoacoustics

- Threshold in quiet



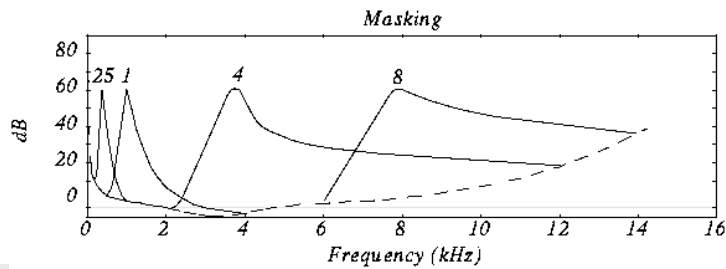
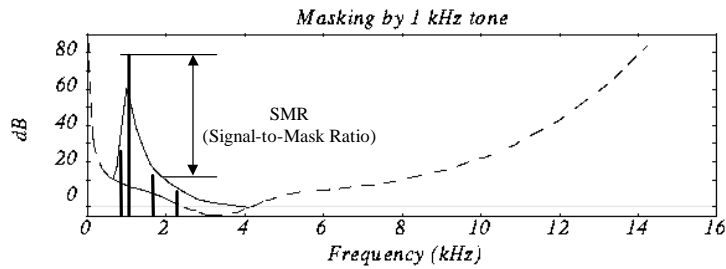
- 26 critical bands 0~24 kHz



- Frequency masking in the same critical band

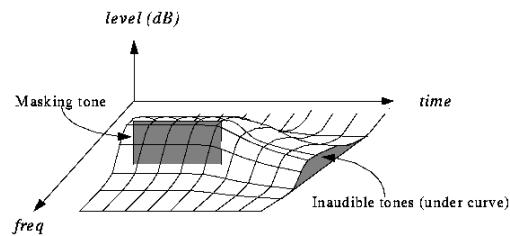
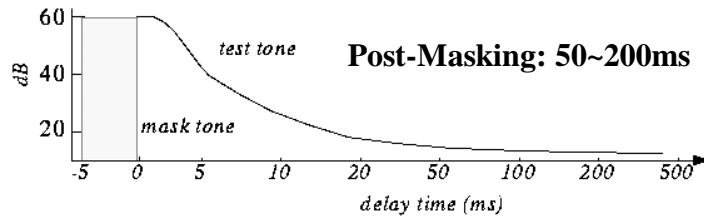
18-796/Spring 1999/Chen

Frequency Masking



18-796/Spring 1999/Chen

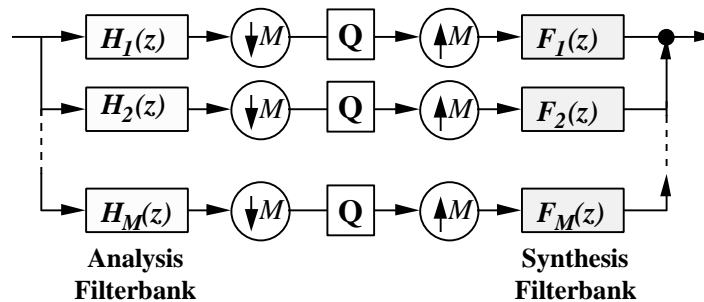
Temporal Masking



- Pre-Masking: 1/10 of post-masking

18-796/Spring 1999/Chen

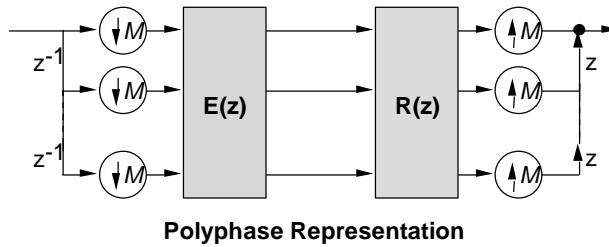
Subband Coding



- Maximal downsampling
- Q should be based on signal-to-masking ratio (SMR)
- Ear's critical bands are not uniform, but logarithmic
 - The filter bank should match the critical bands
 - Tree-structure filter bank (to be derived on board)

18-796/Spring 1999/Chen

Subband Coding vs. DCT



- When $E(z) = \text{DCT matrix}$, this becomes DCT
 - No overlap; blocking artifact
- Modified DCT (MDCT)
 - 50% overlap; less blocking artifact

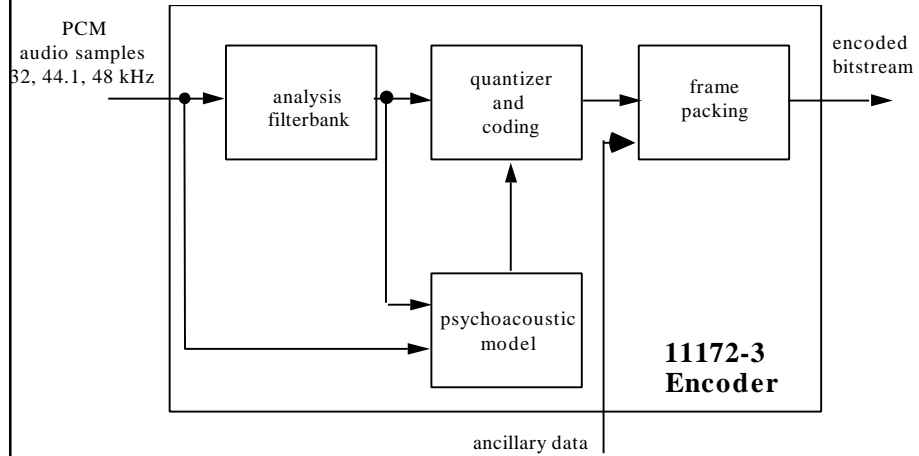
18-796/Spring 1999/Chen

MPEG-1 Audio

- ISO/IEC 11172-3 (1988~1991)
 - First high quality audio compression standard
 - Sampling rates: 32, 44.1, 48 kHz
 - CD quality two-channel audio at ~256 kbits/s
 - CD: $44.1 \text{ kHz} \times 16 \text{ bits} \times 2 = 1.411 \text{ Mbits/s}$
- Quality demonstration (MPEG-1 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

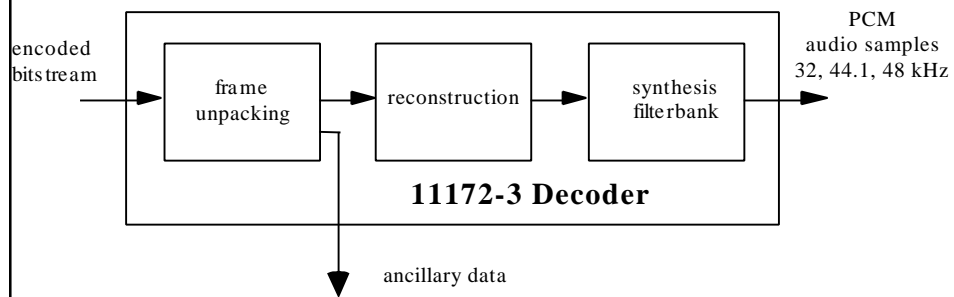
18-796/Spring 1999/Chen

Encoder Block Diagram



18-796/Spring 1999/Chen

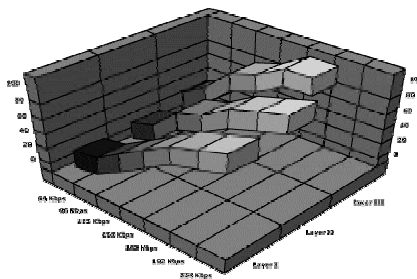
Decoder Block Diagram



18-796/Spring 1999/Chen

Layers

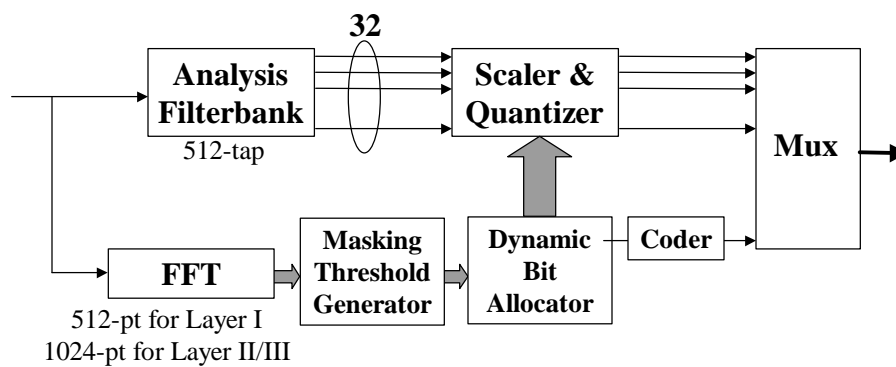
- Increasing complexity, delay, and quality
 - Layer I: ~384 kbits/s for perceptually lossless quality (4:1)
 - Layer II: ~192 kbits/s for perceptually lossless quality (8:1)
 - Layer III: ~128 kbits/s for perceptually lossless quality (12:1)
(for two channels)



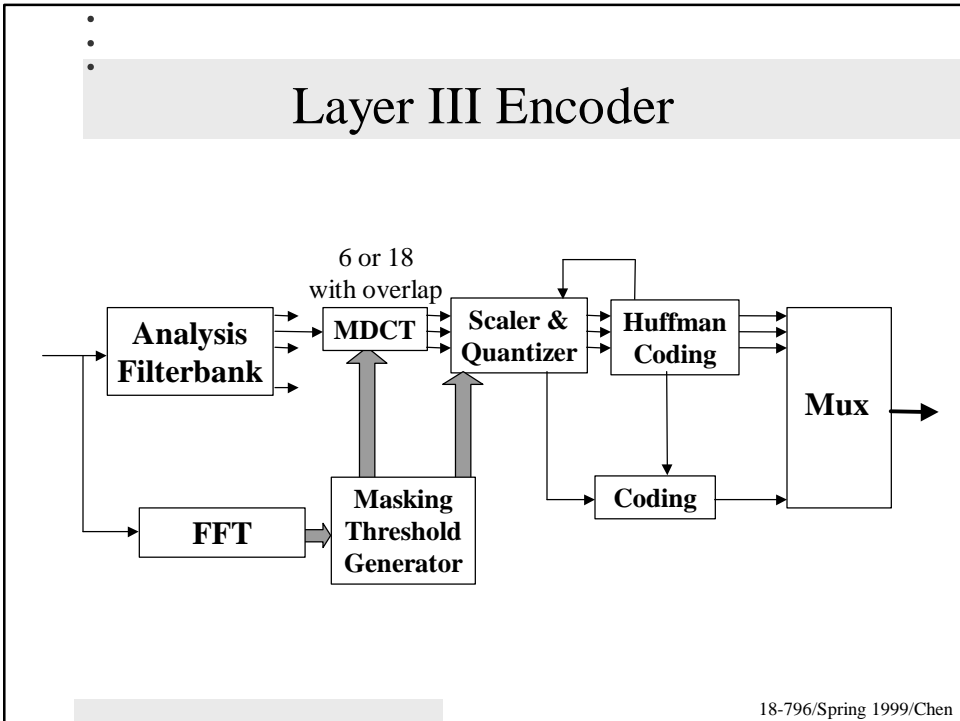
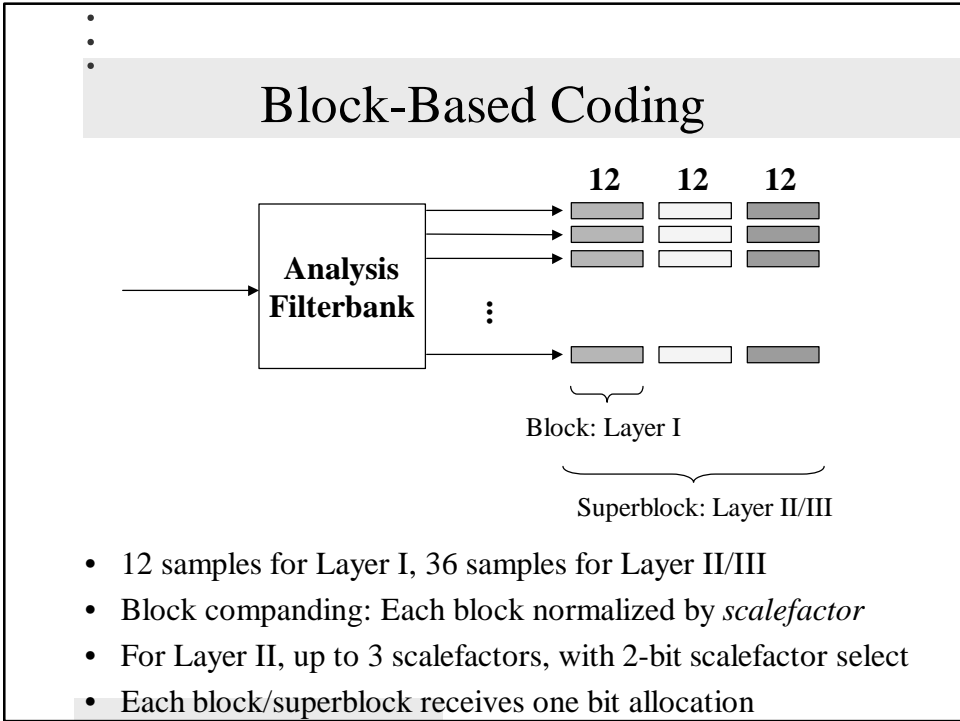
100% \approx
perceptual lossless

18-796/Spring 1999/Chen

Layer I and II Encoder



18-796/Spring 1999/Chen



•
•

Features in Layer III

- Hybrid filterbank
 - MDCT with filterbank
- 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)

18-796/Spring 1999/Chen

•
•

Frame Structure

Header Info	Side Info	Subband Samples	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

18-796/Spring 1999/Chen

-
-

Stereo Redundancy Coding

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

18-796/Spring 1999/Chen

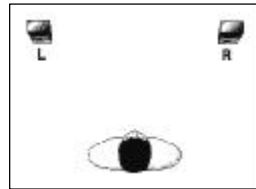
-
-

MPEG-2 Audio

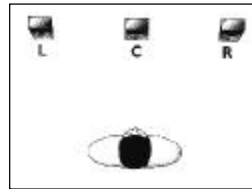
- ISO/IEC 13818-3
 - Allows lower sampling rates
 - 16, 22.05, and 24 kHz: about half of MPEG-1
 - From wideband speech to mediumband audio
 - Higher frequency resolution
 - Layer I, II, and III
- Multichannel coding
 - 2~5 channels; surround sound, multilingual, for visual/hearing-impaired
- Backward compatible and non-backward compatible coding (13818-7: MPEG-2 AAC)

18-796/Spring 1999/Chen

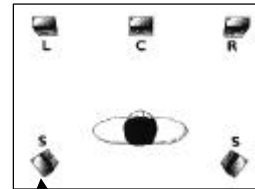
Multichannel Audio



2/0-stereo

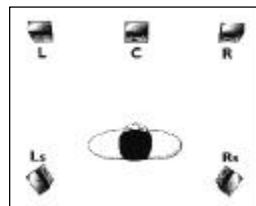


3/0

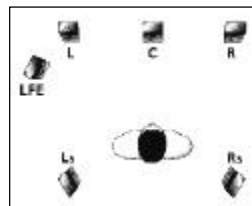


3/1

Surround



3/2



3/2 with woofer
(5.1 system)

LFE: Low-frequency enhancement (woofer)

- 15~120 Hz
- Can be anywhere

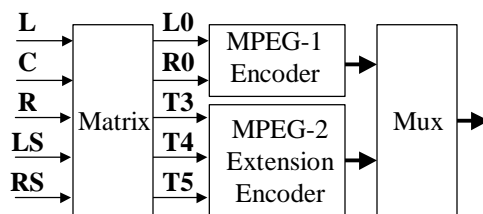
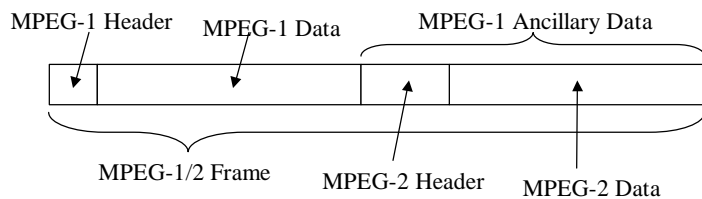
18-796/Spring 1999/Chen

Compatibility

- Forward compatibility
 - A new decoder can decode an old bitstream
 - Usually simple to achieve
- Backward compatibility
 - An old decoder can decode a new bitstream, at least partially
 - Usually limits the coding efficiency

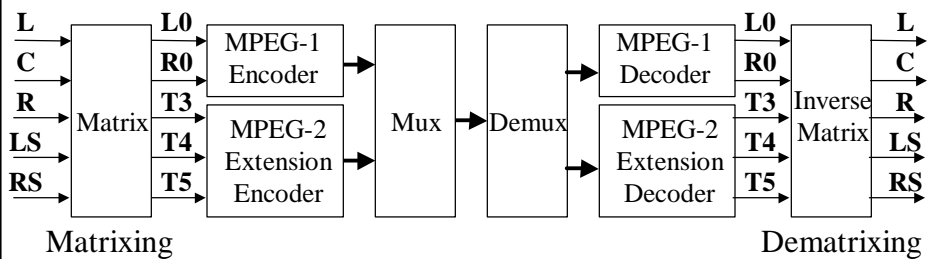
18-796/Spring 1999/Chen

MPEG-2 Backward Compatible Audio Coding



$$\begin{cases} L0 = a(L + b \cdot C + d \cdot LS) \\ R0 = a(R + b \cdot C + d \cdot RS) \end{cases} \quad a = \frac{1}{1+\sqrt{2}}; b = d = \frac{1}{\sqrt{2}} \text{ or } a = 1; b = d = 0$$

Backward Compatible Audio Coding (cont.)



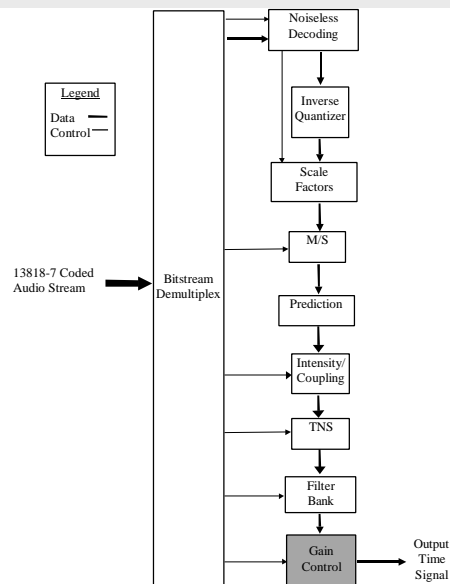
Non Backward Compatible (NBC) Coding

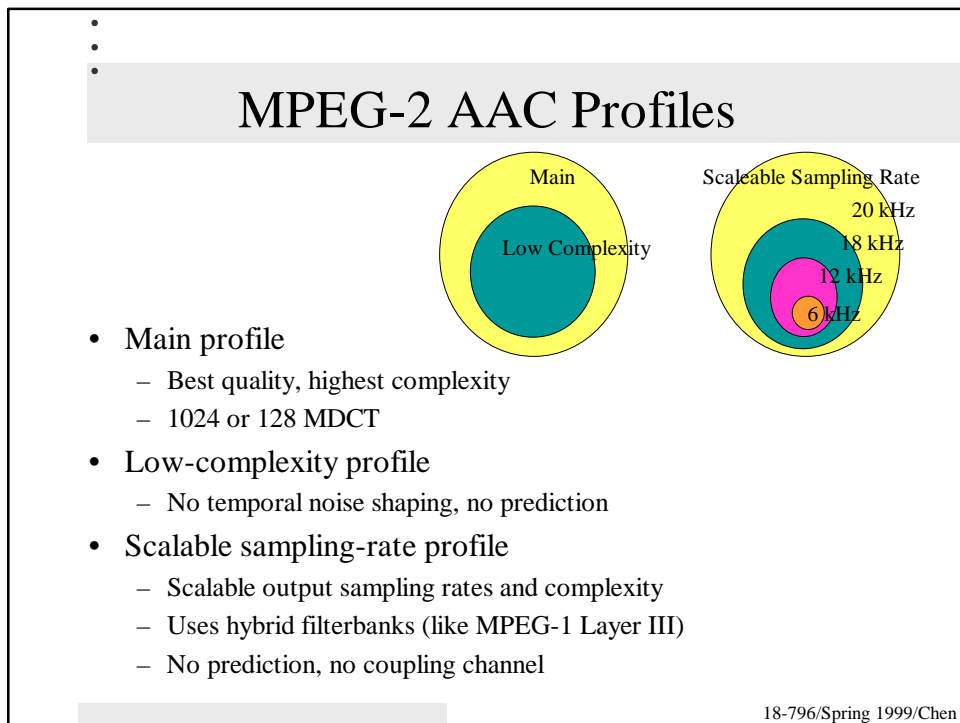
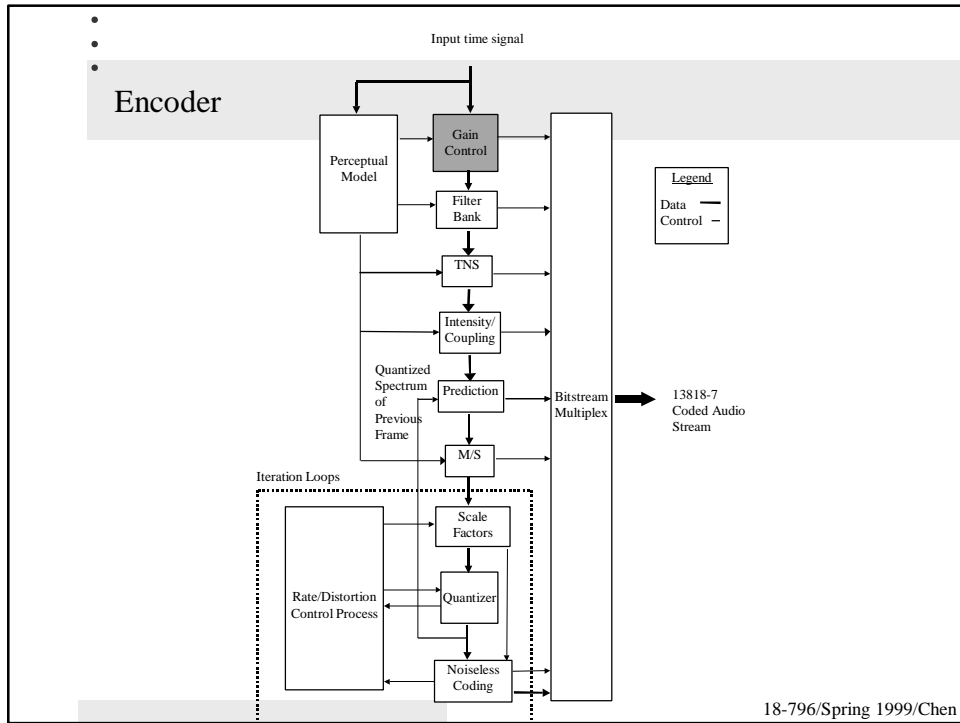
- MPEG-2 Advanced Audio Coding (AAC)
 - ISO/IEC 13818-7 (April 1997)
 - 320~384 kbits/s for 5 channels, 64kbits/channel
 - NBC at 320 kbits/s as good as BC coding at 640 kbits/s
 - 1~48 audio channels, 0~16 LFEs, 0~16 data streams
 - Same framework (perceptual subband coding) as MPEG-1, with some enhancements

18-796/Spring 1999/Chen

MPEG-2 AAC

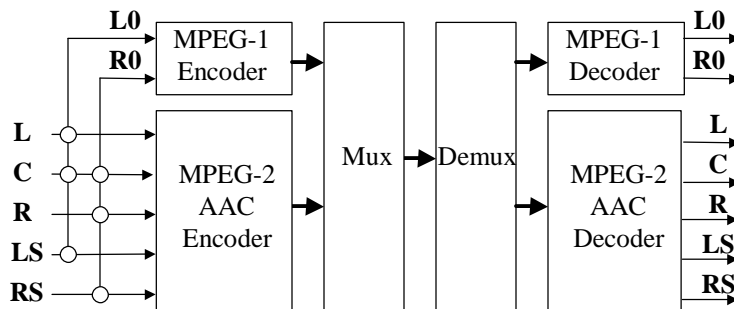
- Enhancements
 - Preprocessing
 - High resolution filterbanks
 - 1024-line MDCT / 128
 - Temporal noise shaping (TNS): time-dependent quantization
 - Coupling channel
 - Intensity multichannel coding
 - Backward adaptive prediction in subbands
 - M/S stereo coding
 - Noiseless coding (entropy coding): Huffman coding





Simcast

- To achieve backward compatibility at the cost of higher bitrate



18-796/Spring 1999/Chen

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 Multimedia*, v. 2, no. 2, 1995, pp. 60-74
- <http://www.mpeg.org/MPEG/audio.html>
- <http://www.cselt.it/mpeg/faq/faq-audio.htm>
- <http://www.tnt.uni-hannover.de/project/mpeg/audio/>

18-796/Spring 1999/Chen