MPEG-1

Overview of MPEG-1 Standard

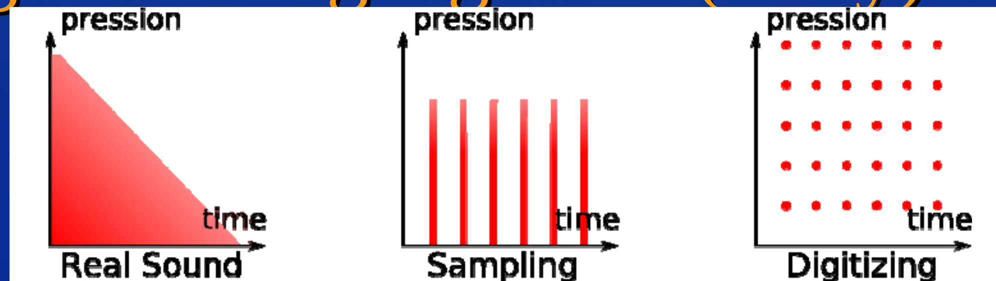
Introduction to perceptual and entropy codings

Contents

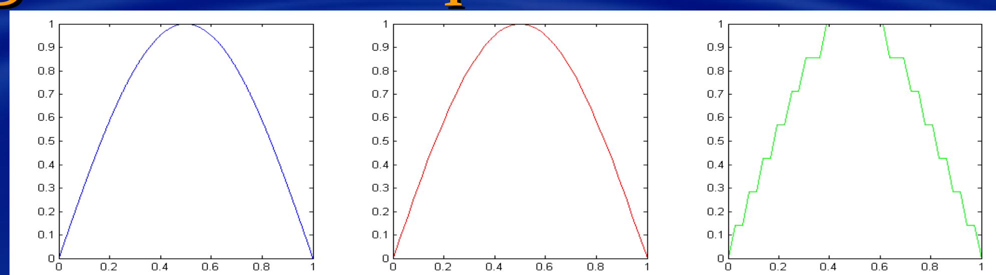
- History
- Psychoacoustics and perceptual coding
- Entropy coding
- MPEG-1
  - Layer I/II
  - Layer III (MP3)
  - Comparison and Audio Quality

Introduction

- Digitizing an analog signal is (lossy) compression



- Digitizing introduces quantization noise



- Quantization noise imply loss of quality
  - Linear quantization > 16 bit (98 dB) → inaudible noise (CD)
  - Linear quantization 4 bit (26 dB)

Contents

- History
- Psychoacoustics and perceptual coding
- Entropy coding
- MPEG-1
  - Layer I/II
  - Layer III (MP3)
  - Comparison and Audio Quality

History

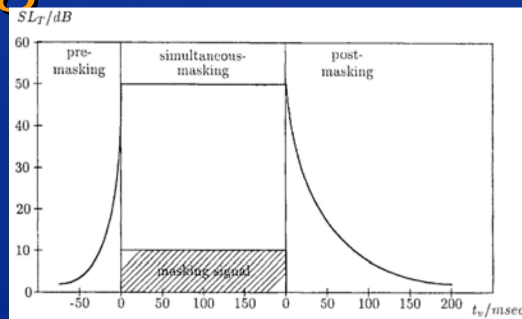
- Moving Picture Expert Group (MPEG)
  - Created in January 1988
  - Starts the development of MPEG-1 in May 1988
  - Publishes the MPEG-1 standard in November 1992 (ISO/IEC 11172-3 for audio)
- MPEG-1 standard
  - Defines bit-stream
  - Defines decoding functions
  - DOES NOT define encoding techniques
  - Inspired by MUSICAM (Masking pattern Universal Subband Integrated Coding And Multiplexing)

Contents

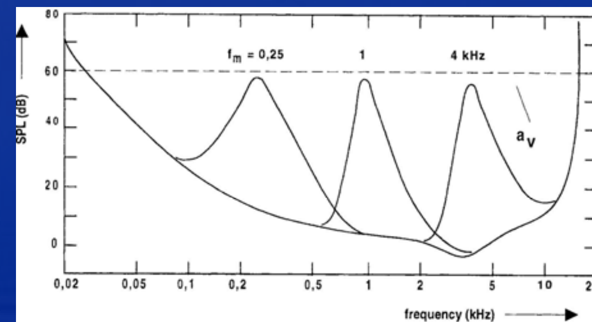
- History
- **Psychoacoustics and perceptual coding**
- Entropy coding
- MPEG-1
  - Layer I/II
  - Layer III (MP3)
  - Comparison and Audio Quality

Psychoacoustics

- Masking effect



Time domain masking (*Pohlmann 2000*)

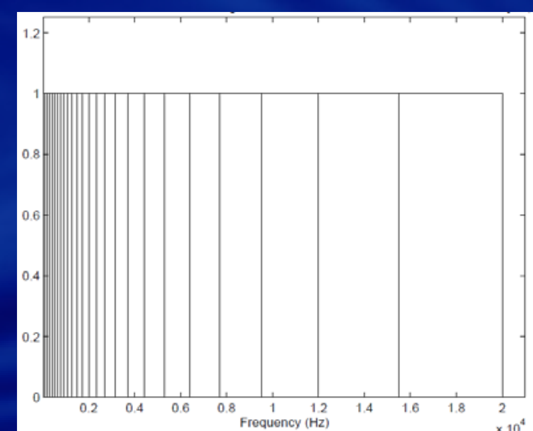


Frequency domain masking (*Pohlmann 2000*)

Critical bands

z/Bark	lower boundary	higher boundary	bandwidth	central frequency
0	0	100	100	50
1	100	200	100	150
2	200	300	100	250
3	300	400	100	350
4	400	510	110	450
5	510	630	120	570
6	630	770	140	700
7	770	920	150	840

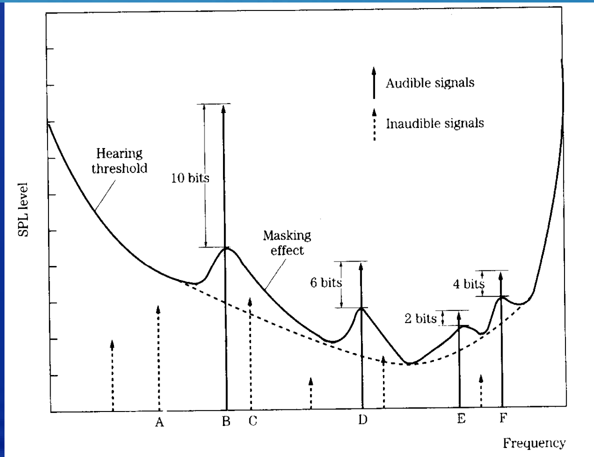
(Brandenburg)



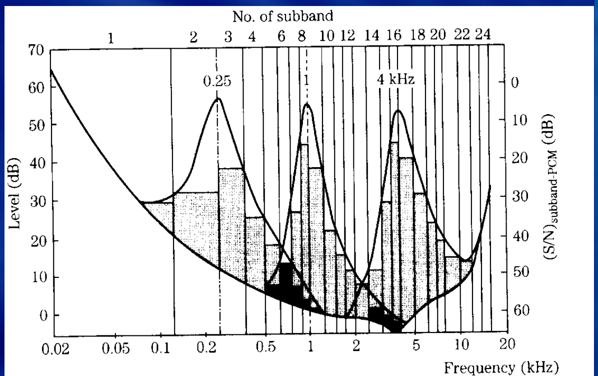
Idealized critical bands (*Painter & Spanias 2000*)

Perceptual Coding

- Dividing the different subbands of a signal
- Ignoring masked audio information
- Introducing inaudible quantization noise

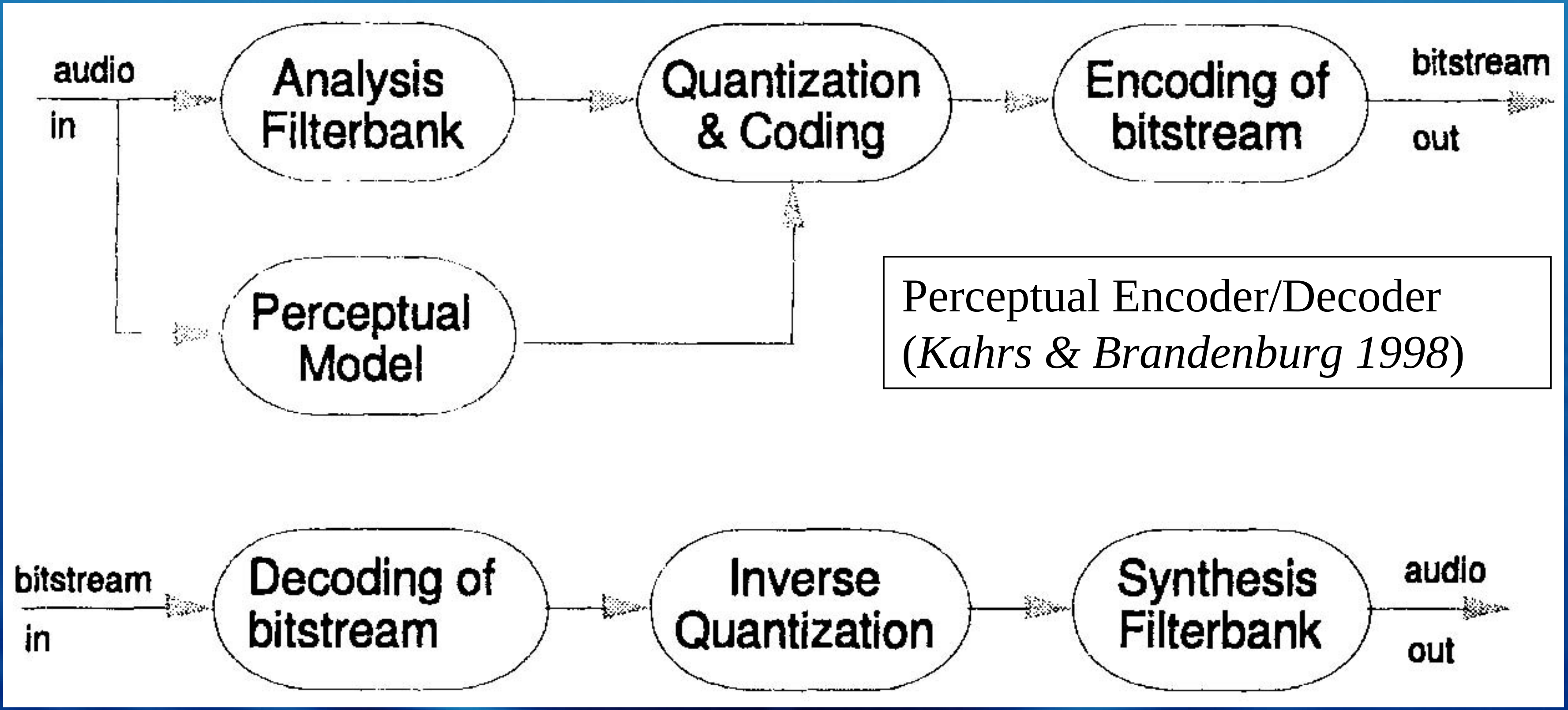


Bits association according to masking threshold (*Pohlmann 2000*)



Quantization noise added according to masking threshold (*Pohlmann 2000*)

Perceptual Coding



Perceptual Encoder/Decoder
(*Kahrs & Brandenburg 1998*)

Contents

- History
- Psychoacoustics and perceptual coding
- **Entropy coding**
- MPEG-1
  - Layer I/II
  - Layer III (MP3)
  - Comparison and Audio Quality

Entropic Coding

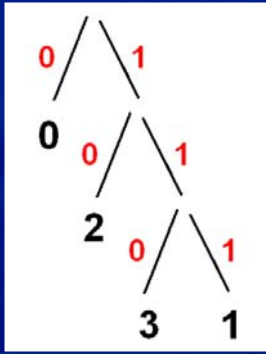
- Use information about the signal to code efficiently
- Entropy of a signal
  - Example 1: {0, 2, 2, 2, 0, 0, 0, 0, 0, 2, 0, 3, 2, 2, 0, 0, 0, 3, 0, 0}
    20 symbols – twelve 0 (0.6), zero 1 (0), six 2 (0.3), two 3 (0.1)
    Entropy $H = 1.30$
  - Example 2: {1, 2, 3, 0, 2, 1, 1, 2, 3, 0, 0, 1, 0, 3, 3, 3, 2, 0, 1, 2}
    20 symbols – five 0 (0.25), five 1 (0.25), five 2 (0.25), five 3 (0.25)
    Entropy $H = 2$
- Shannon theorem

  ➔ It is impossible to code with less than H bits/symbol

  ➔ It is possible to code with less than $H+1$ bits/symbol

Entropic Coding

- ## Huffman coding
  - Example 1: {0, 2, 2, 2, 0, 0, 0, 0, 0, 2, 0, 3, 2, 2, 0, 0, 0, 3, 0, 0}
    20 symbols – twelve 0 (0.6), zero 1 (0), six 2 (0.3), two 3 (0.1)
    Entropy *H* = 1.30
  - Immediate coding:
    0 → "00" 1 → "01" 2 → "10" 3 → "11"
    "0010101000000000001000111010000000110000"
  - Huffman coding:
    0 → "0" 1 → "111" 2 → "10" 3 → "110"
    "010101000000100110101000011000"



- ## Efficiency:
  - Immediate coding: 2 bits/symbol
  - Huffman coding: 1.5 bits/symbol (statistically)

Contents

- History
- Psychoacoustics and perceptual coding
- Entropy coding
- **MPEG-1**
  - Layer I/II
  - Layer III (MP3)
  - Comparison and Audio Quality

MPEG-1

- Sampling rate: 32, 44.1 and 48 kHz
- Four modes:
  - Mono: 1 channel
  - Stereo: 2 channels
  - Dual: 2 channels independent (e.g. bilingual programmes)
  - Joint stereo: 2 channels coded together
- 2 perceptual models
- Floating point quantization (normalization)
- Error checking: Cyclic redundancy check (CRC)

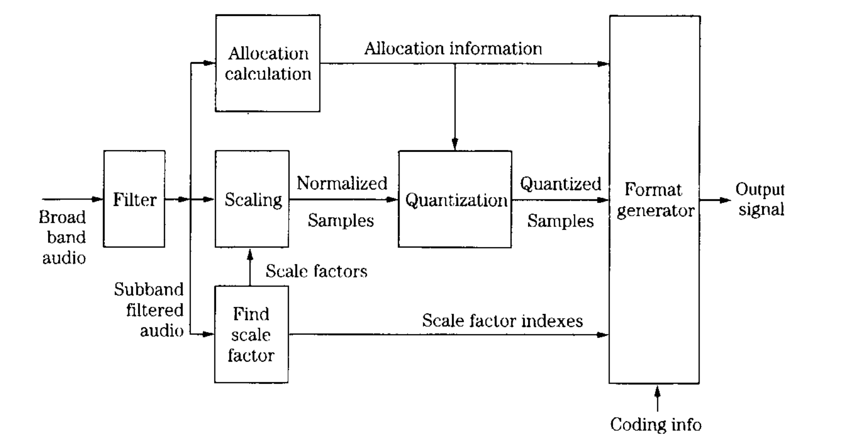
Contents

- History
- Psychoacoustics and perceptual coding
- Entropy coding
- **MPEG-1**
  - **Layer I/II**
  - Layer III (MP3)
  - Comparison and Audio Quality

MPEG-1 Layer I

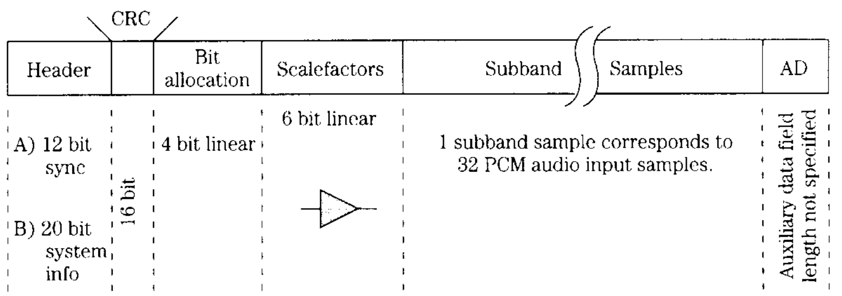
- From 32 to 448 kbps
- 32-subband polyphase filterbank
- Bit allocation (0-15)
  - Max dynamic range > 120 dB
- Linear quantization
- 1 frame ⇔ 384 samples

Example: Philips Digital Compact Cassette



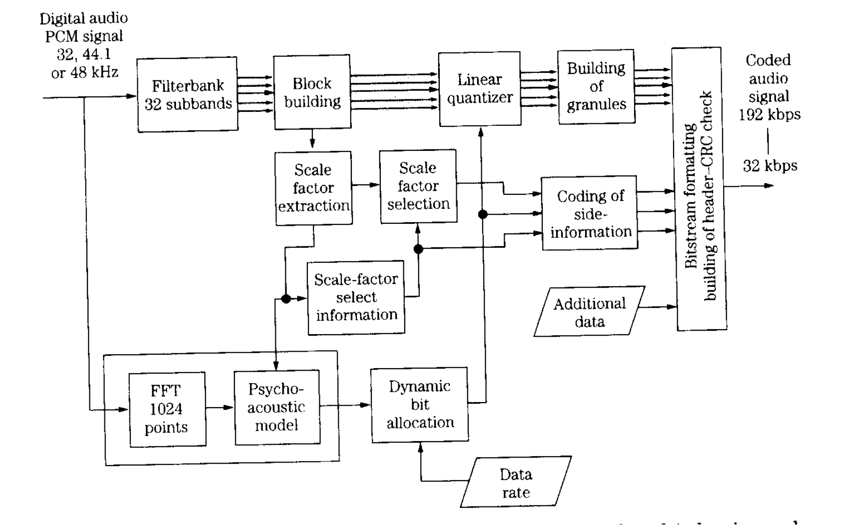
Example of Layer I encoder (*Pohlmann 2000*)

Layer I frame format (*Pohlmann 2000*)



MPEG-1 Layer II

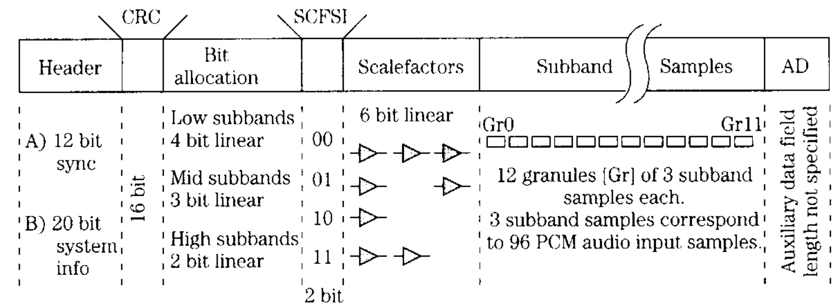
- From 32 to 384 kbps
- Improvement of Layer I
  - Improved FFT analysis
  - Scale factor redundancy
  - Finer quantization
- 1 frame ⇔ 1152 samples



Example of Layer II encoder (*Pohlmann 2000*)

Example: Digital Audio Broadcasting (DAB)

Layer II frame format (*Pohlmann 2000*)

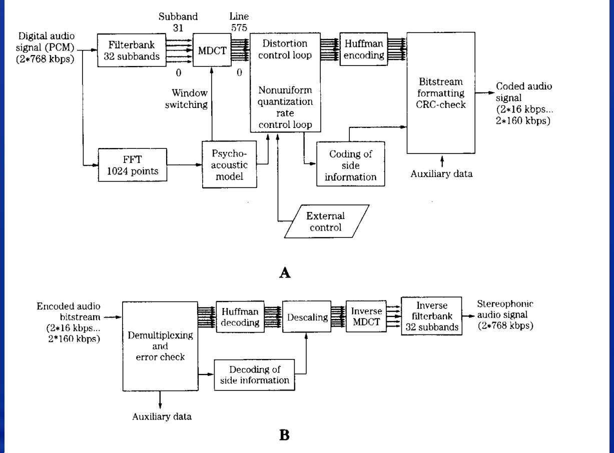


Contents

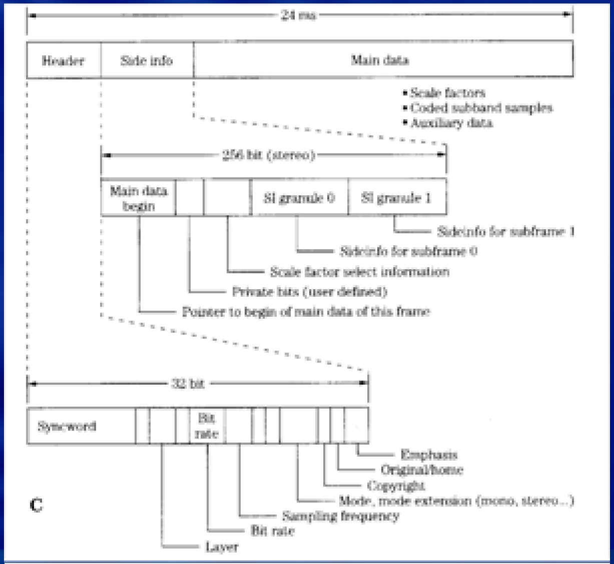
- History
- Psychoacoustics and perceptual coding
- Entropy coding
- **MPEG-1**
  - Layer I/II
  - **Layer III (MP3)**
  - Comparison and Audio Quality

MPEG-1 Layer III (MP3)

- From 32 to 320 kbps
- Improvements:
  - Finer psychoacoustics model
  - Alias reduction (MDCT filters)
  - Nonuniform quantization
  - Entropy coding
  - Adaptative block size
- Only Layer with patents
- Inspired by:
  - ASPEC (audio spectral perceptual entropy coding)
  - OCF (optimal coding in the freq. domain)



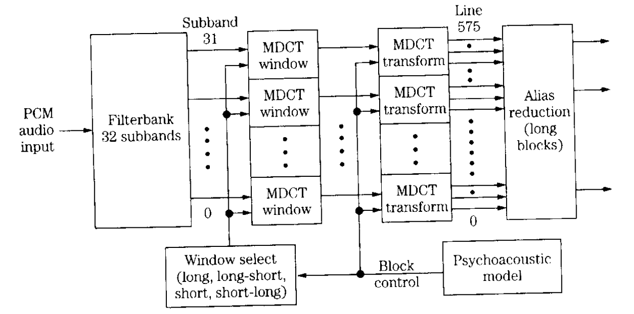
Example of Layer III encoder (*Pohlmann 2000*)



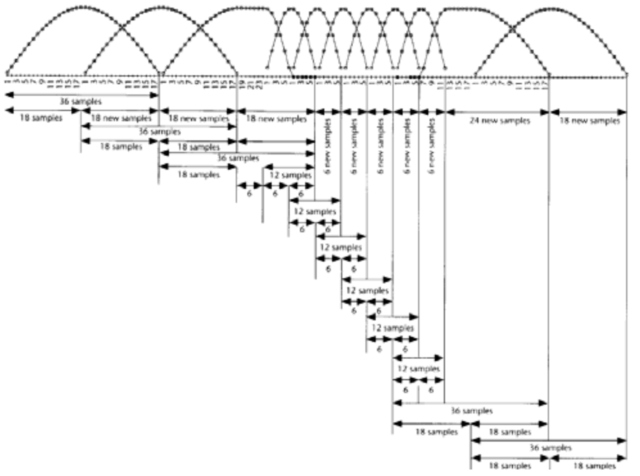
Layer III frame format (*Pohlmann 2000*)

MPEG-1 Layer III (MP3)

- Filtering: Hybrid polyphase filter/MDCT
  - Steady-state signals: 18-point MDCT on every subband
    Frequency resolution: 41.67 Hz
    Time resolution: 24 ms
  - Transient signals: 6-point MDCT
    Frequency resolution: 125 Hz
    Time resolution: 8 ms
  - 3 blocks modes
  - Pre-echo detection
- Quantization : power 3/4
- Entropy coding:
  - Huffman tables
  - Run length coding



Filtering stage in Layer III encoder (*Pohlmann 2000*)



MDCT filterbank in Layer III encoder (*Pohlmann 2000*)

Joint Stereo Coding

- Intensity coding
  - Sum of left/right channels
  - Coding of the sum and of left/right scale factors
  - Usually only for high-frequency subbands
  - Efficient for redundant audio channels
- MS (mid/side) stereo coding
  - Sum and difference of left/right channels
  - Coding of the two values
- Stereo masking

Contents

- History
- Psychoacoustics and perceptual coding
- Entropy coding
- **MPEG-1**
  - Layer I/II
  - Layer III (MP3)
  - **Comparison and Audio Quality**

Audio Quality

Comparison with CD-quality at 48 kHz (16 bit - 1.412 Mbps)

- Layer I:
  - No perceptual difference for 384 kbps (stereo) – 2:1 compression
- Layer II:
  - No perceptual difference for 256 kbps (stereo) – 4:1 compression
- Layer III
  - Increase of mean opinion score compared to Layer II at 256 kbps (stereo) for 128 kbps (stereo) – 8:1 compression

Comparison between Layers

- Layers I and II are very similar
- Each Layer has its defined decoder design
- Encoding/Decoding complexity: Layer I/II
  → broadcasting
- Encoding/Decoding quality: Layer III
  → audio storage

Conclusion

Bibliography:

- Ambikairajah, E. & al. *Auditory masking and MPEG-1 audio compression*, Electronics & Communication Engineering Journal, **1997**
- Brandenburg, K. & Bosi, M. *Overview of MPEG Audio: Current and Future Standards for Low-Bit-Rate Audio Coding*, Journal of the Audio Engineering Society, **1997**, Vol. Vol. 45(No. 1/2)
- Painter, T. & Spanias, A. *Perceptual Coding of Digital Audio*, Proceedings of IEEE, **2000**, Vol. Vol. 88(No. 4)
- Painter, T. & Spanias, A. *A Review of Algorithms for Perceptual Coding of Digital Audio Signals*, Digital Signal Processing, **1997**
- Pan, D. *A Tutorial on MPEG/Audio Compression*, IEEE MultiMedia, IEEE Computer Society, **1995**, Vol. 2(2), pp. 60-74
- Pan, D.Y. *Digital Audio Compression*, Digital Technical Journal, **1993**, Vol. 5
- Kahrs, M. and Brandenburg, K. *Applications of digital signal processing to audio and acoustics*, Kluwer Academic Publishers, **1998**
- Mallat, S. *Traitement du Signal*, Ecole Polytechnique, **2000**
- Pohlmann, K.C. *Principles of Digital Audio*, McGraw-Hill Professional, **2000**