Audio Coding
Quantization and Coding Methods

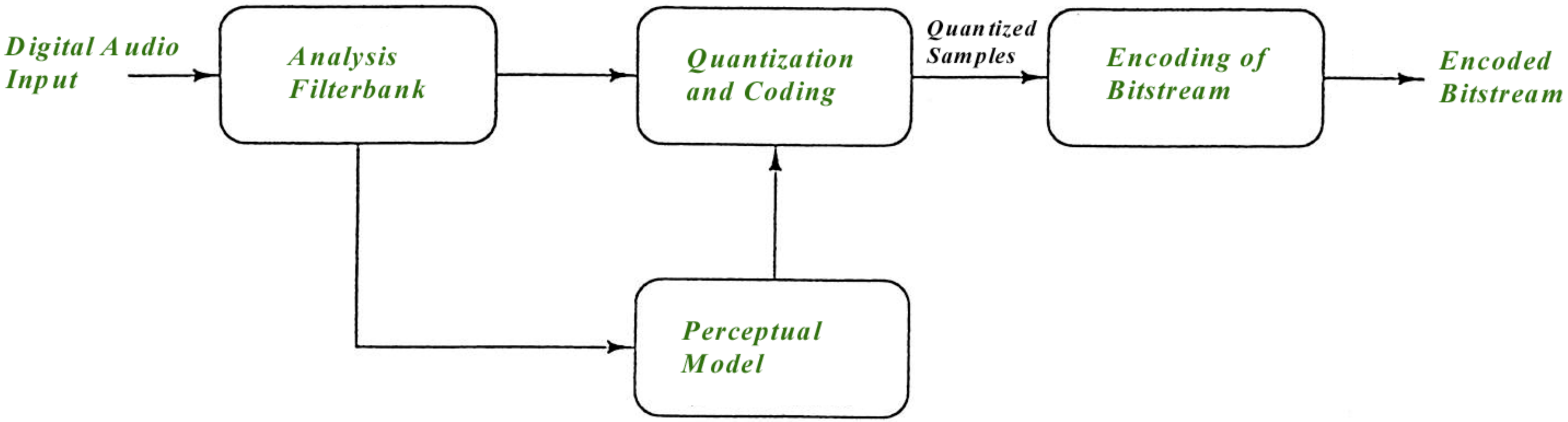
Prof. Dr.-Ing. Karlheinz Brandenburg
karlheinz.brandenburg@tu-ilmenau.de

© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Quantization and Coding Methods (1)



© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Quantization and Coding Methods (2)

Objective:
- „Good“ representation of spectral data
  - Compactness (low bit rate)
  - Smallest possible perceptible distortion (high subjective quality)

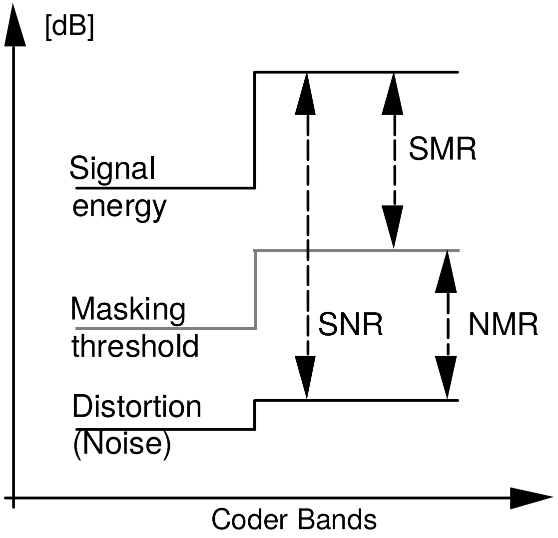
Overview:
- Quantization
- Noiseless Coding
- Joint Quantization/ Coding Techniques
- Encoding Strategies

© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Quality and Rate Measures



- Signal-to-Noise Ratio (SNR)
  - Quadratic distortion metric
  - Does not allow prediction about quality !
- Noise-to-mask Ratio (NMR)
  - Ratio of distortion with respect to masking threshold
  - Determines audibility of distortions
  - Should be as small as possible($\leq$ 0 dB)
- Signal-to-Mask Ratio (SMR)
  - Relation between signal and masking threshold
  - Gives indication of bit demand

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

© Fraunhofer IDMT

Quantization (1)

Basics:

- Data reduction by removing irrelevance
- Explicit control of quantization distortion according to time/frequency-dependent masking threshold (perceptual coder)
- High variability in local SNR
  - (e.g. 0db ... >30db)
- Most popular case: Scalar quantization

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU

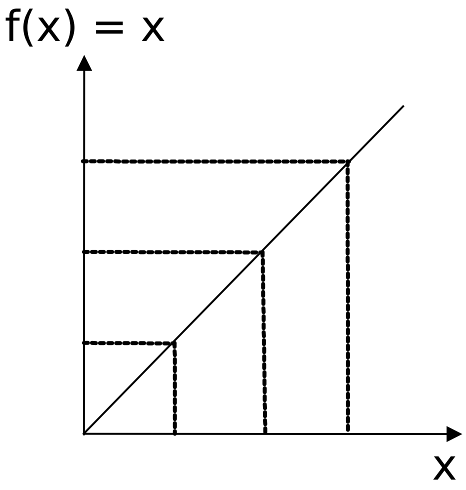
Fraunhofer IDMT

Quantization (2)

Simpler:

- Uniform Quantization
  - MPEG-1/2 Layer I and II, ATRAC, AC-3
  - Average distortion independent on size of coefficients
  - Quantization stepsize is constant
- more precise control of quantization noise:
- Stepsize determines the quantization noise:

$$E(e^2) = \frac{\Delta^2}{12}$$



© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Quantization (3)

More sophisticated:

- Non-uniform quantization
- MPEG-1/2 Layer 3, MPEG-2/4 AAC:
  -

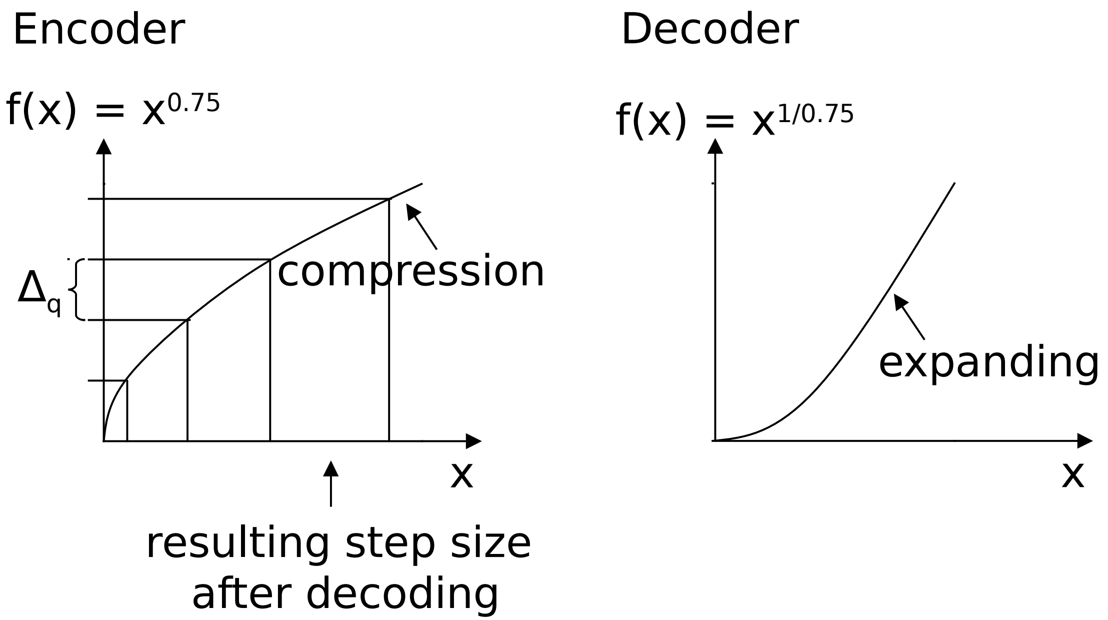
$$S_i = round\left(\frac{x_i^{0.75}}{\Delta_q}\right)$$

- More distortion for larger coefficients (subband signal x)
- In comparison: For uniform quantization, the exponent is 1

© Fraunhofer IDMT

Quantization (4)

Non-uniform quantization:



© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Quantization (5)

Approach #1:

- Group of spectral coefficients (subband signal x) is normalized by means of a common multiplier („scalefactor“)
- Side information: Scalefactor and quantizer resolution (bits/sample)
- „Block companding“ / „Block floating point“
- MPEG-1/2 Layers I and II, ATRAC, AC-3

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de
© Fraunhofer IDMT

Quantization (6)

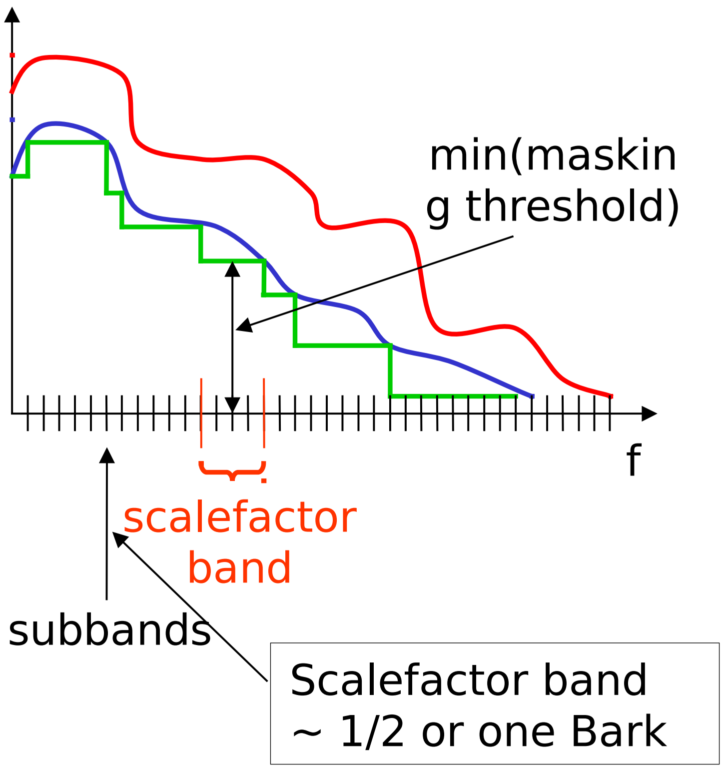
Approach #2:

- Group of spectral coefficients is scaled („scalefactor“) and subsequently quantized by a fixed quantizer
- Scalefactor can be seen as a quantization stepsize
- Side information: Scalefactor
- Used with entropy coding
- Quant. Precision controlled by scalefactor
- MPEG-1/2 Layer 3, MPEG-2/4 AAC

© Fraunhofer IDMT

Scalefactor Bands and Masking Threshold

Shaping of the noise according to psycho-acoustic model. Several subbands in each scalefactor band, with same step size.



Signal
Masking Threshold
Each step: scalefactor band with minimum value of masking threshold

Approximation of Masking Threshold by a step function
→ fewer bits for parameterization

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

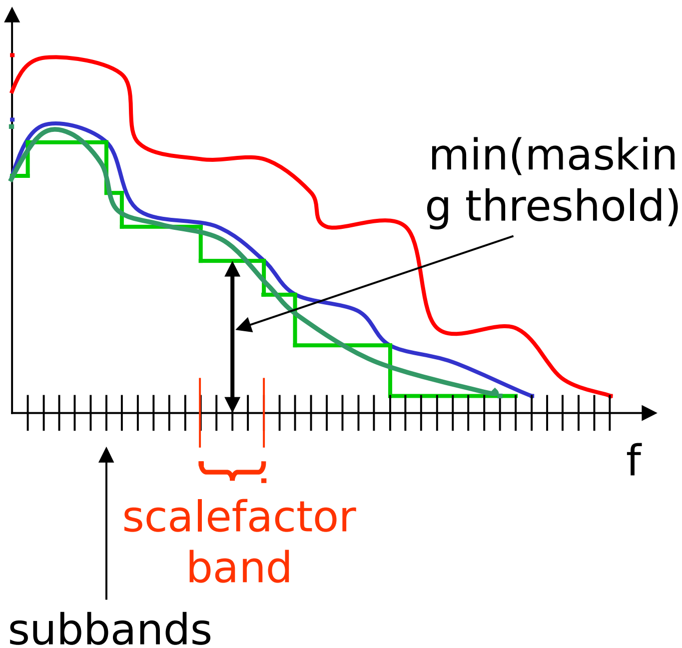
© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Effect of non-uniform Quantization

in each scalefactor band with



Signal
Masking Threshold
Each step: scalefactor band with minimum value of masking threshold

Additional noise shaping with non-uniform quantizer ($x^{0.75}$): quantization noise follows signal x

Approximation of Masking Threshold by a step function
→ fewer bits for parameterization

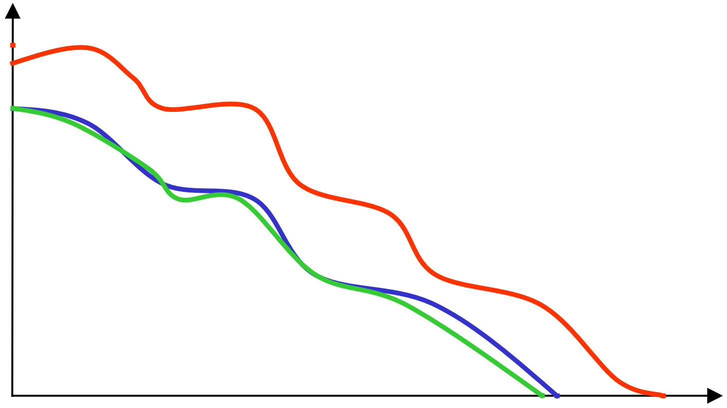
© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Polynomial Approximation of Masking Threshold

Different way to approximate the Masking Threshold, use polynomial approximation instead of step-function of scale factor bands.

Goal: Approximation of Masking Threshold as close as possible to reduce bit demand



Signal

Masking Threshold

Polynomial Approximation (example: 12 coefficients)

Used in Ultra Low Delay Coder, predictive coder, no subbands.
Is frequency response of linear filter.

© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Scalefactor bands

- MDCT subbands which fall into one scalefactor band → their power sums up
- Conversion of DFT/MDCT frequency bands into 1/2 or one Bark, at low frequencies broader
- After quantization more efficient representation of the codewords is needed → noiseless coding

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de
© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU

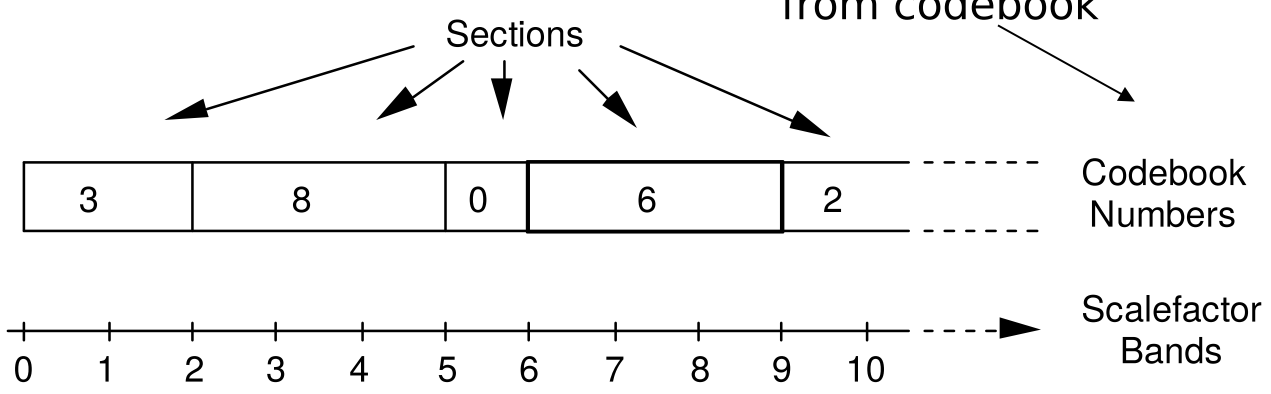
Fraunhofer IDMT

Noiseless Coding (1)

Even more efficient:

- Higher gain by using entropy coding
- Efficient representation of the symbols after quantization
  - Huffman coding (MPEG-1/2 Layer III, MPEG-2/4 AAC)
  - Arithmetic coding (MPEG-4 version 2 BSAC, USAC)

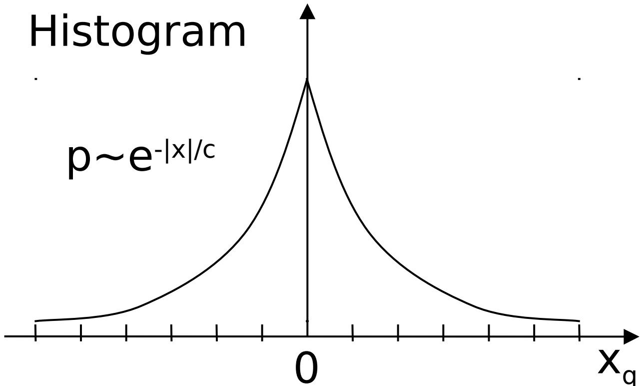
choose „best“ Huffman code from codebook



Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de
© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Noiseless Coding (2)

- Huffman coding uses the probability of the symbol to determine the codeword
- typical probability distribution of audio samples: Laplace distribution



0: highest probability
→ reason for compression gain with entropy coding
→ shorter codewords for the more likely symbols (e.g. 0)
→ shortest codeword for 0

© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Noiseless Coding (3)

Example for state-of-the art coding kernel (MPEG-2/4 AAC)

- Multi-dimensional (2 or 4-dim.) entropy coding -> exploiting joint statistics of vector comp.
- Huffman coding, several codebooks
- Signal information part of codebook or separate escape mechanism for large quant. values
- Choice of Huffman codebook for arbitrary groups of scalefactor bands („sections“)
- can obtain less than 1 bit per sample with 2 or 4 dimensional vectors

© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

MPEG Audio Layer-3: Huffman Code Tables, Example

Table B.7 -- Huffman codes for Layer III

Huffman code table for quadruples (A)

v w x y	hlen	hcod
0 0 0 0	1	1
0 0 0 1	4	0101
0 0 1 0	4	0100
0 0 1 1	5	00101
0 1 0 0	4	0110
0 1 0 1	6	000101
0 1 1 0	5	00100
0 1 1 1	6	000100
1 0 0 0	4	0111
1 0 0 1	5	00011
1 0 1 0	5	00110
1 0 1 1	6	000000
1 1 0 0	5	00111
1 1 0 1	6	000010
1 1 1 0	6	000011
1 1 1 1	6	000001

Huffman code table for quadruples (B)

v w x y	hlen	hcod
0 0 0 0	4	1111
0 0 0 1	4	1110
0 0 1 0	4	1101
0 0 1 1	4	1100
0 1 0 0	4	1011
0 1 0 1	4	1010
0 1 1 0	4	1001
0 1 1 1	4	1000
1 0 0 0	4	0111
1 0 0 1	4	0110
1 0 1 0	4	0101
1 0 1 1	4	0100
1 1 0 0	4	0011
1 1 0 1	4	0010
1 1 1 0	4	0001
1 1 1 1	4	0000

© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

MPEG Audio Layer-3: Huffman Code Tables(2)

Huffman code table 0

linbits=0

x	y	hlen
0	0	0

Huffman code table 1

linbits=0

x	y	hlen	hcod
0	0	1	1
0	1	3	001
1	0	2	01
1	1	3	000

Huffman code table 2

linbits=0

x	y	hlen	hcod
0	0	1	1
0	1	3	010
0	2	6	000001
1	0	3	011
1	1	3	001
1	2	5	00001
2	0	5	00011
2	1	5	00010
2	2	6	000000

Huffman code table 3

linbits=0

x	y	hlen	hcod
0	0	2	11
0	1	2	10
0	2	6	000001
1	0	3	001
1	1	2	01
1	2	5	00001
2	0	5	00011
2	1	5	00010
2	2	6	000000

Huffman code table 4

not used

© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

MPEG Audio Layer-3: Huffman Code Tables(3)

Huffman code table 5

linbits=0

x	y	hlen	hcod
0	0	1	1
0	1	3	010
0	2	6	000110
0	3	7	0000101
1	0	3	011
1	1	3	001
1	2	6	000100
1	3	7	0000100
2	0	6	000111
2	1	6	000101
2	2	7	0000111
2	3	8	00000001
3	0	7	0000110
3	1	6	000001
3	2	7	0000001
3	3	8	00000000

Huffman code table 6

linbits=0

x	y	hlen	hcod
0	0	3	111
0	1	3	011
0	2	5	00101
0	3	7	0000001
1	0	3	110
1	1	2	10
1	2	4	0011
1	3	5	00010
2	0	4	0101
2	1	4	0100
2	2	5	00100
2	3	6	000001
3	0	6	000011
3	1	5	00011
3	2	6	000010
3	3	7	0000000

Huffman code table 7

linbits=0

x	y	hlen	hcod
0	0	1	1
0	1	3	010
0	2	6	001010
0	3	8	00010011
0	4	8	00010000
0	5	9	000001010
1	0	3	011
1	1	4	0011
1	2	6	000111
1	3	7	0001010
1	4	7	0000101
1	5	8	00000011
2	0	6	001011
2	1	5	00100
2	2	7	0001101
2	3	8	00010001
2	4	8	00001000
2	5	9	000000100
3	0	7	0001100

⋮

© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

MPEG Audio Layer-3: Huffman Code Tables(4)

Huffman code table 8

linbits=0

x	y	hlen	hcod
0	0	2	11
0	1	3	100
0	2	6	000110
0	3	8	00010010
0	4	8	00001100
0	5	9	000000101
1	0	3	101
1	1	2	01
1	2	4	0010
1	3	8	00010000
1	4	8	00001001
1	5	8	00000011
2	0	6	000111
2	1	4	0011
2	2	6	000101
2	3	8	00001110
2	4	8	00000111
2	5	9	000000011
3	0	8	00010011

⋮

Huffman code table 9

linbits=0

x	y	hlen	hcod
0	0	3	111
0	1	3	101
0	2	5	01001
0	3	6	001110
0	4	8	00001111
0	5	9	000000111
1	0	3	110
1	1	3	100
1	2	4	0101
1	3	5	00101
1	4	6	000110
1	5	8	00000111
2	0	4	0111
2	1	4	0110
2	2	5	01000
2	3	6	001000
2	4	7	0001000
2	5	8	00000101
3	0	6	001111

⋮

Huffman code table 10

linbits=0

x	y	hlen	hcod
0	0	1	1
0	1	3	010
0	2	6	001010
0	3	8	00010111
0	4	9	000100011
0	5	9	000011110
0	6	9	000001100
0	7	10	0000010001
1	0	3	011
1	1	4	0011
1	2	6	001000
1	3	7	0001100
1	4	8	00010010
1	5	9	000010101
1	6	8	00001100
1	7	8	00000111
2	0	6	001011
2	1	6	001001

⋮

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

© Fraunhofer IDMT

Joint Quantization / Coding Techniques (1)

Vector quantization

- Excellent coding efficiency at very low rates (<< 1 bit / sample), but
- Perceptual control of distortion difficult

→ control only total distortion for "group"

Application range:

- Used for intermediate quality / very low bit rate coding
  - e.g. MPEG-4 TwinVQ

Nice theoretical concept, today not in use in widely used standards!

The excitation codebook in CELP speech coding follows this paradigm.

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de
© Fraunhofer IDMT

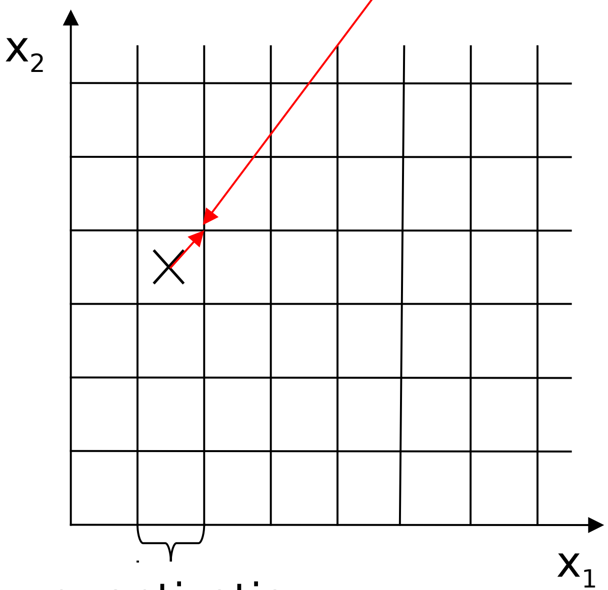
TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Joint Quantization / Coding Techniques (2)

resulting quantization error for 2 subbands

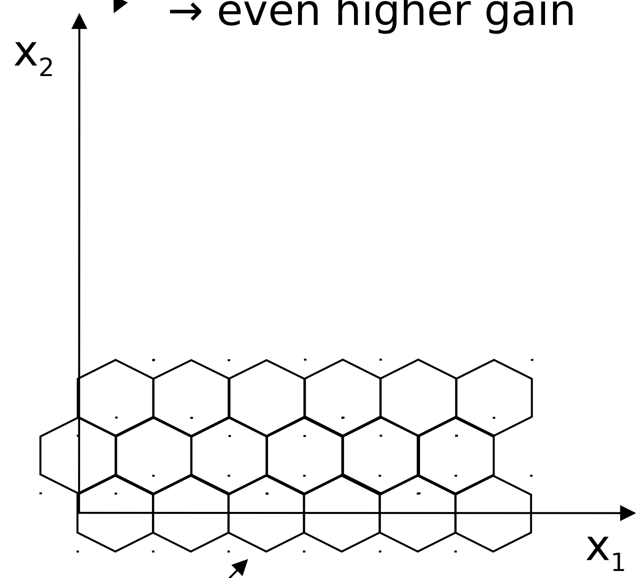
subband value 2



quantizatio n step size

subband value 1

also possible for higher dimensions → even higher gain



denser packing of quantization „spheres“ → reduced average quantization errors

Encoding Strategies

- Today‘s coding schemes provide large degree of flexibility
  - Quantization noise profile over frequency
  - Trade-off audio bandwidth vs. overall distortion
  - Bit rate & coding mode
  - Usage of optional tools (Temporal Noise Shaping, prediction)
- Encoding strategy „intelligent“ part of encoding; determines quality
- Arena for specific know-how („secrets of audio coding“)

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de
© Fraunhofer IDMT

Constant Quality Coding

Goal:
- Coding with a pre-selected constant quality

Procedure:
- Estimate time & frequency dependent masking threshold
- Adjust quantizer step size to meet target distortion

Constant quality -> Variable bit rate

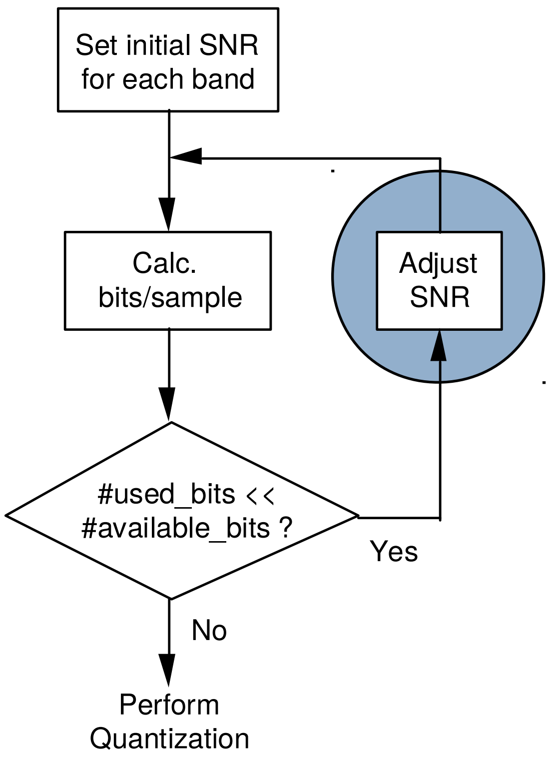
Applications:
- Storage media / transmission channels supporting variable rate e.g.
  - Storage on digital media; music on internet

- Needs accurate perceptual model, often constant bit-rate coders sound better

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de
© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Constant Rate Coding (1)



Paradigm 1 „Bit Allocation“

- Used typically in coding schemes with block companding
- Direct translation between bit rate and local SNR available

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de

© Fraunhofer IDMT

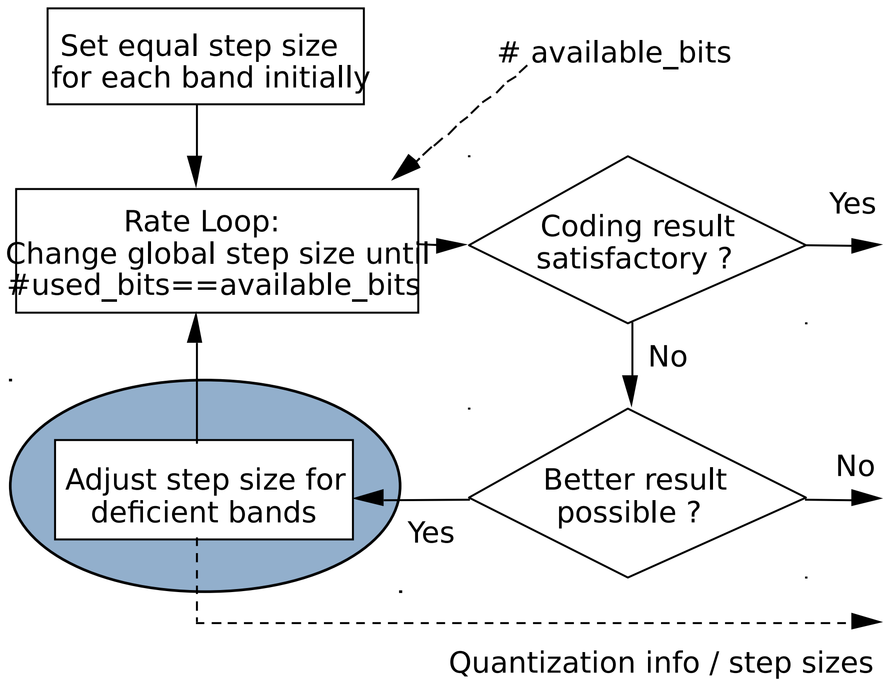
TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Constant Rate Coding (2)

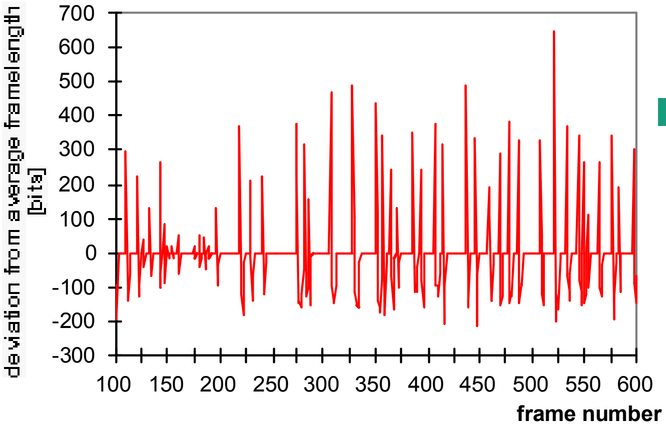
Paradigm 2 „Noise Allocation“

- Used in codecs with entropy encoding



© Fraunhofer IDMT

Constrained Variable Bit Rate Coding (Bit Reservoir)



- Constant quality vs. Constant bit rate
  - Constant quality → variable bit rate (VBR)
  - Constant bit rate → variable quality
  - ? How to achieve both ?
- Compromise:
  - Bit reservoir = constrained VBR Coding
  - Limit accumulated deviation of bit consumption from average („bit reservoir size“)
  - Additional delay
  - Used in MPEG-1/2 Layer III, MPEG-2/4 AAC

© Fraunhofer IDMT

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Constant vs. Variable Bit Rate Example

- Take our Python example:

```
python psycho-acoustic-modelDFT_gs.py
```

Constant Quality, variable bit-rate:
If we **don't shift the masking threshold** with the arrow key, we get a **variable bit-rate,** which depends on the audio signal.
Observe: The suitable **adjustment of the masking threshold is critical for the audio quality**.

Constant bit-rate, variable quality:
We constantly **adjust the masking threshold** such that we obtain a **constant bit-rate**. Here the adjustment of the masking theshold is taken over by the bit-rate requirement. But we obtain a variable audio quality.

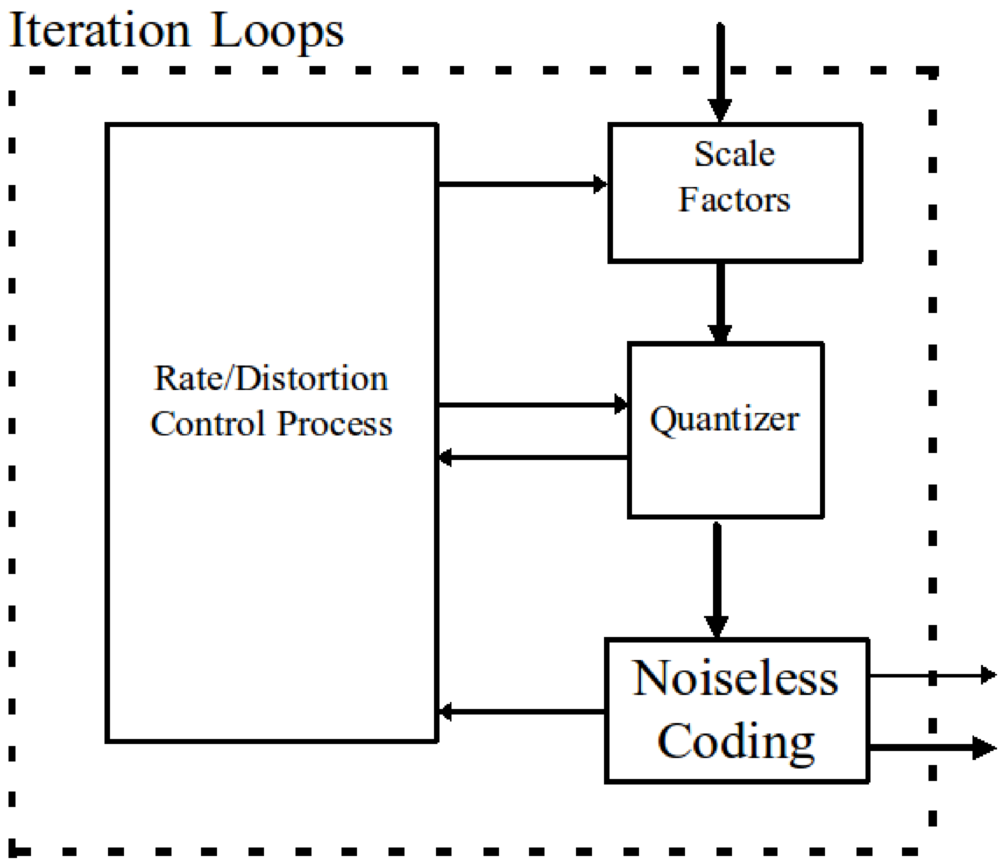
© Fraunhofer IDMT

Quantization / Coding: Codec Overview

	Filterbank channels	Quantization	Coding
MPEG-1/2 Layer I / II	32	uniform	block companding
MPEG-1/2 Layer III	576 / 192	non-uniform	Huffman coding
MPEG-2 AAC	1024 / 128	non-uniform	Huffman coding
MPEG-4 TwinVQ	1024 / 128 (960/120)	VQ	VQ
AC-3	256 / 128	uniform	block companding
ATRAC	96 .. 512	uniform	block companding

© Fraunhofer IDMT
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Layer-3 Iteration Loops



Layer-3 : Outer Loop

- Distortion Loop (control of the distortion)
- Saves the unquantized spectral values
- Compares the reconstructed values with the original
- Builds the actual distortion in the frequency domain
- Scaling by frequency groups with the amount of distortion
- Convergence of the iteration is not guaranteed

Layer-3 : Inner Loop

- Rate Loop (Data rate controller)
- Entropy coding: Data rate depends on actual data set
- Buffer Control: Controls the necessary bits
- Convergence through iterations: is always possible
- Beginning level: Calculated from SFM (Spectral Flatness Measure)

Prof. Dr.-Ing. K. Brandenburg, karlheinz.brandenburg@tu-ilmenau.de Prof. Dr.-Ing. G. Schuller, gerald.schuller@tu-ilmenau.de
© Fraunhofer IDMT