

Seminar Kompressionsalgorithmen: MP3 & Ogg Vorbis

Erik Noorman

June 6, 2012

Inhalt

- 1 Grundlagen
- 2 Psychoakustik
- 3 MP3
- 4 Ogg Vorbis
- 5 Vergleich

Digitalisierung von Audiosignalen



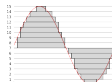
Quelle



Schall



Mikrofon

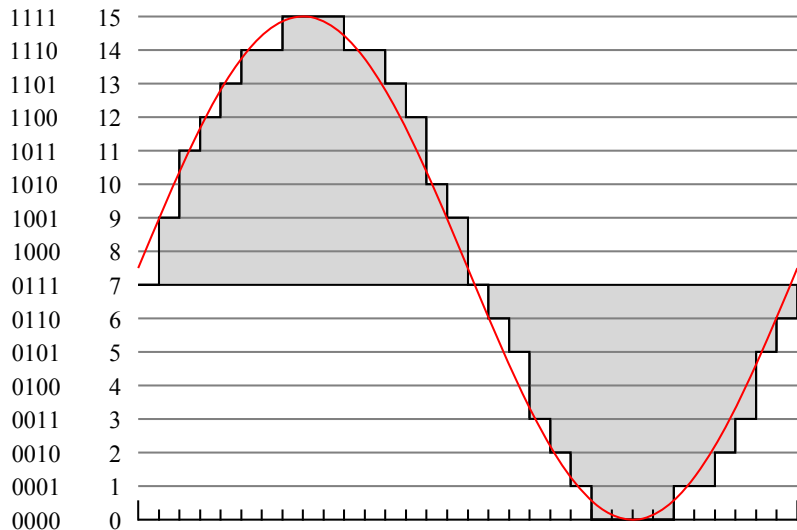


AD-Wandler

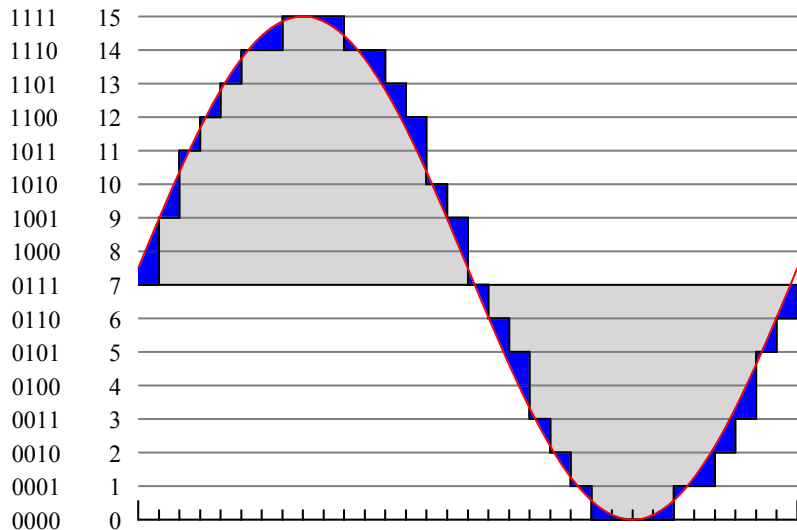


PC

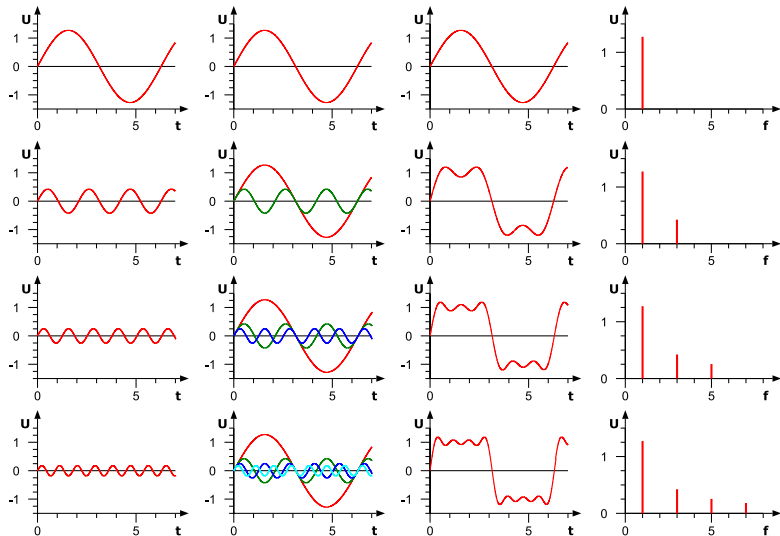
Digitalisierung: Abtastung, Quantisierung, Codierung



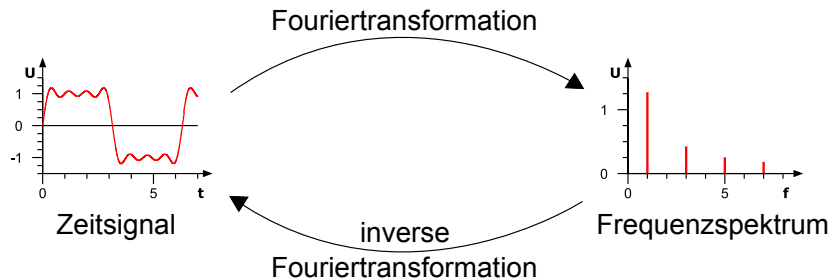
Quantisierungsrauschen



Frequenzspektrum



Fouriertransformation



- Zeitsignal und Frequenzspektrum sind eindeutig ineinander überföhrbar
- Annahme: Bandbegrenztes Signal

Motivation

Welches von beiden Beispielen ist komprimiert?

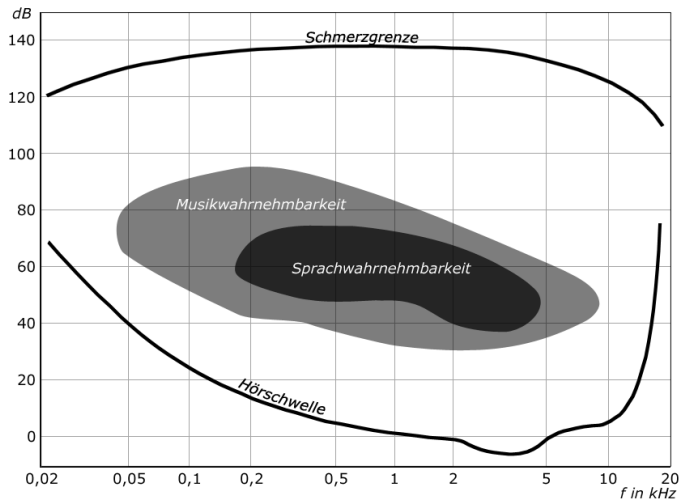


Motivation

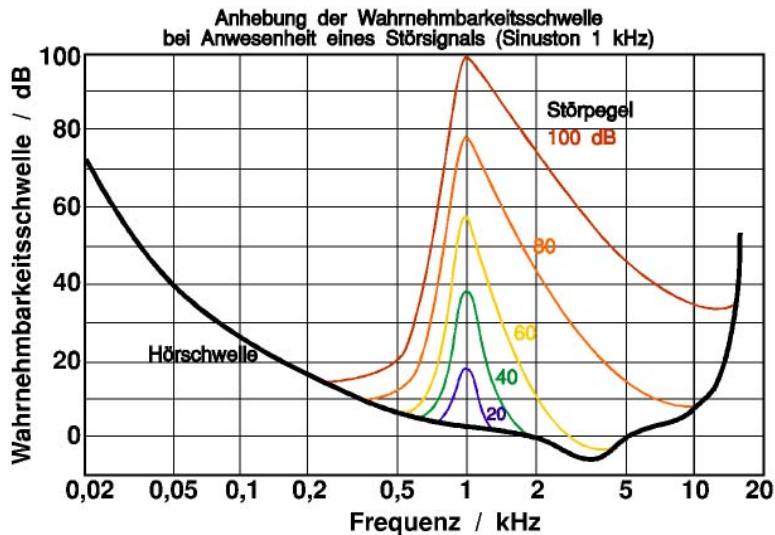
Welches von beiden Beispielen ist komprimiert?

- 1  MP3 (192kb/s = 317KB)
- 2  WAV (unkomprimiert = 2340KB)

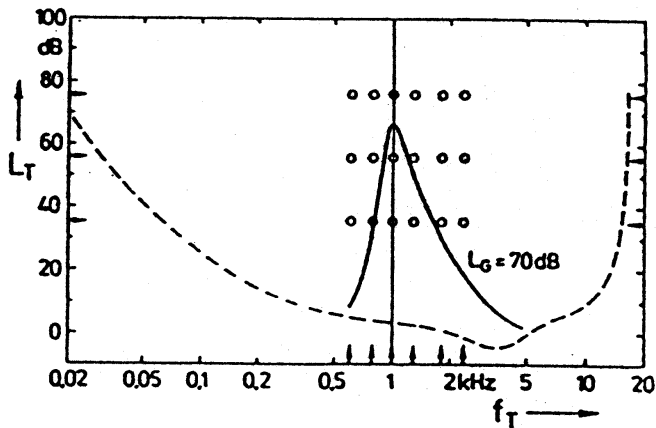
Hörschwelle



Maskierung im Frequenzbereich

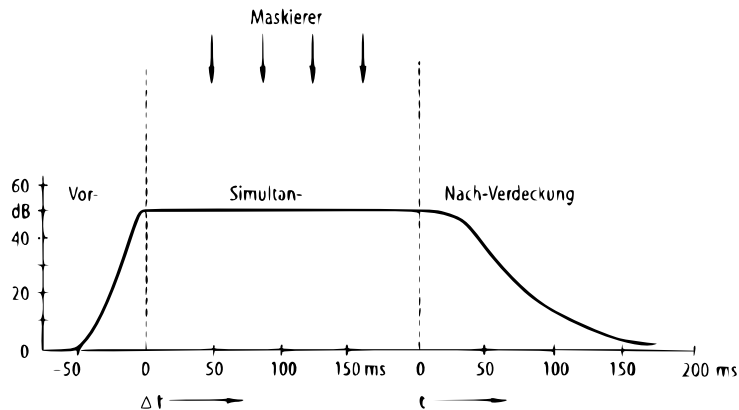


Mithörschwelle bei Schmalbandrauschen



Beispiel: 

Maskierung im Zeitbereich

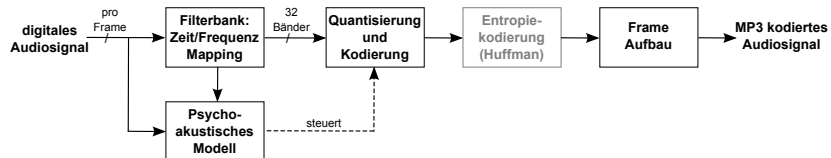


MP3

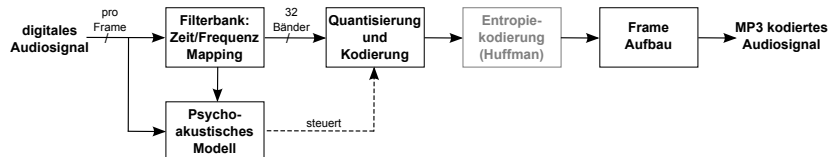
MP3

- verlustbehafteter Audiokompressionsalgorithmus
- macht sich psychoakustische Effekte zunutze
- 1987: entwickelt vom Fraunhofer Institut für Integrierte Schaltungen (IIS)
- 1992: Teil des MPEG-1-Standards
- 1998: verlangt FhG/Thomson Lizenzgebühren
- heute: weit verbreitet

Kodierer



Kodierer

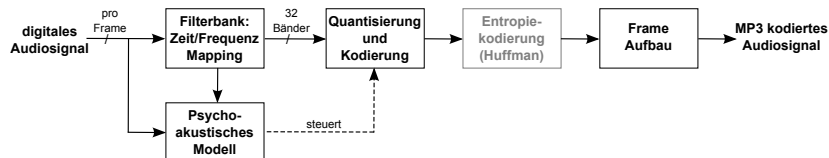


Filterbank: Zeit/Frequenz Mapping

- Zeit-Signal → Frequenzspektrum
- Unterteilung des Frequenzspektrums in 32 Bänder



Kodierer



Filterbank: Zeit/Frequenz Mapping

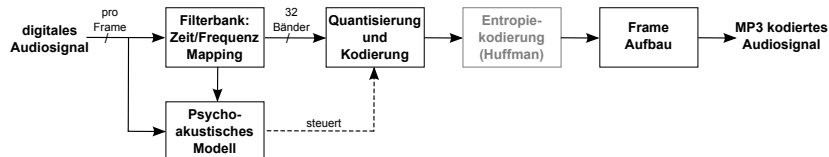
- Zeit-Signal $\rightarrow$ Frequenzspektrum
- Unterteilung des Frequenzspektrums in 32 Bänder



Psychoakustisches Modell

- berechnet Signal-Maskierungs-Verhältnis (SMV)

Kodierer



Filterbank: Zeit/Frequenz Mapping

- Zeit-Signal $\rightarrow$ Frequenzspektrum
- Unterteilung des Frequenzspektrums in 32 Bänder



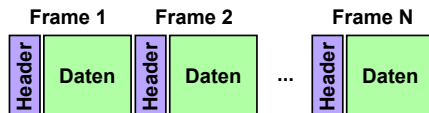
Psychoakustisches Modell

- berechnet Signal-Maskierungs-Verhältnis (SMV)

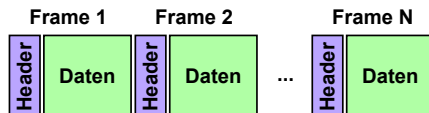
Quantisierung und Kodierung

- nutzt SMV zur Berechnung der minimalen Bitrate bei der das Quantisierungsrauschen gerade nicht mehr hörbar ist

Frame Aufbau

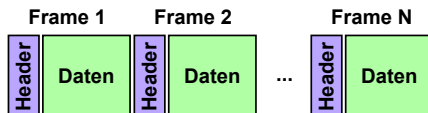


Frame Aufbau



Position	Aufgabe	Länge in Bit
A	Frame-SYNC (1111111111)	11
B	MPEG Audioversion	2
C	MPEG Layer (1=Layer III, 2=Layer II, 3=Layer I)	2
D	Protection	1
E	Bitrate-Index (*)	4
F	Samplingfrequenz (0=44100 1=48000 2=32000)	2
G	Padding Bit	1
H	Private Bit	1
I	Channelmode (0=Stereo, 1=Joint Stereo, 2=Dual, 3=Mono)	2
J	Mode-Extension	2
K	Copyright	1
L	Original	1
M	Emphasis	2

Frame Aufbau



Position	Aufgabe	Länge in Bit	
A	Frame-SYNC (1111111111)	11	1=32
B	MPEG Audioversion	2	2=40
C	MPEG Layer (1=Layer III, 2=Layer II, 3=Layer I)	2	3=48
D	Protection	1	4=56
E	Bitrate-Index (*)	4	5=64
F	Samplingfrequenz (0=44100 1=48000 2=32000)	2	6=80
G	Padding Bit	1	7=96
H	Private Bit	1	8=112
I	Channelmode (0=Stereo, 1=Joint Stereo, 2=Dual, 3=Mono)	2	9=128
J	Mode-Extension	2	10=160
K	Copyright	1	11=192
L	Original	1	12=224
M	Emphasis	2	13=256
			14=320

Kodierung von Stereosignalen

Stereo

linker (L) und rechter (R) Kanal werden einzeln kodiert

Kodierung von Stereosignalen

Stereo

linker (L) und rechter (R) Kanal werden einzeln kodiert

Joint-Stereo

hohe Korrelation von L und R wird ausgenutzt

$$Mid = \frac{L + R}{2}$$

$$Side = \frac{L - R}{2}$$

Kodierung von Stereosignalen

Stereo

linker (L) und rechter (R) Kanal werden einzeln kodiert

Joint-Stereo

hohe Korrelation von L und R wird ausgenutzt

$$Mid = \frac{L + R}{2}$$

$$Side = \frac{L - R}{2}$$

Rekonstruktion:

$$R = Mid - Side$$

$$L = Mid + Side$$

Dekodierer

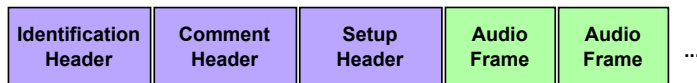


Ogg Vorbis

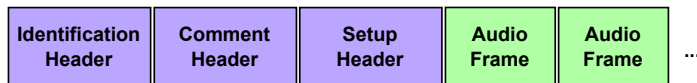
Ogg Vorbis

- besteht aus zwei Verfahren
 - 1 Ogg Containerformat
 - 2 Vorbis verlustbehaftetes Audiokompressionsverfahren
- entwickelt von Xiph.Org Foundation
- Patent- und abgabefrei
- 1993: Xiph.Org Gründung
- 1998: Vorbis wird entwickelt aufgrund der Lizenzeinführung bei MP3

Aufbau der Vorbisdaten



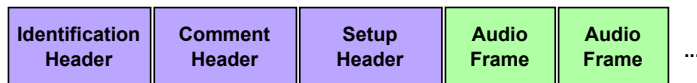
Aufbau der Vorbisdaten



Identification Header

allgemeine Informationen: Vorbis Version, Anzahl der Kanäle, Samplingrate, etc.

Aufbau der Vorbisdaten



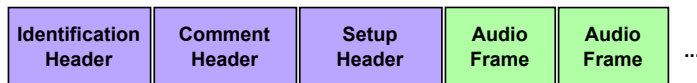
Identification Header

allgemeine Informationen: Vorbis Version, Anzahl der Kanäle, Samplingrate, etc.

Comment Header

User-Kommentare (tags)

Aufbau der Vorbisdaten



Identification Header

allgemeine Informationen: Vorbis Version, Anzahl der Kanäle, Samplingrate, etc.

Comment Header

User-Kommentare (tags)

Setup Header

Information zum Dekodierer Setup: akustische Modell, Huffman- und Vektorquantisierungstabellen (+ über 100 Einträge zum Finetuning des Dekodierers)

Audio-Frame Aufbau

Type
Mode
Window

Floor
curves

Residue
vectors

Audio-Frame Aufbau

Type
Mode
Window

Floor
curves

Residue
vectors

Floor curves

- **grob** aufgelöste Repräsentation des Signals
- werden Huffman- und VQ-kodiert abgespeichert

Audio-Frame Aufbau

Type
Mode
Window

Floor
curves

Residue
vectors

Floor curves

- **grob** aufgelöste Repräsentation des Signals
- werden Huffman- und VQ-kodiert abgespeichert

Residue vectors

- **fein** aufgelöste Repräsentation des Signals
- werden Huffman- und VQ-kodiert abgespeichert

Audio-Frame Aufbau

Type
Mode
Window

Floor
curves

Residue
vectors

Floor curves

- **grob** aufgelöste Repräsentation des Signals
- werden Huffman- und VQ-kodiert abgespeichert

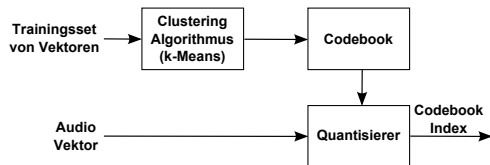
Residue vectors

- **fein** aufgelöste Repräsentation des Signals
- werden Huffman- und VQ-kodiert abgespeichert

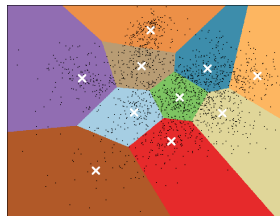
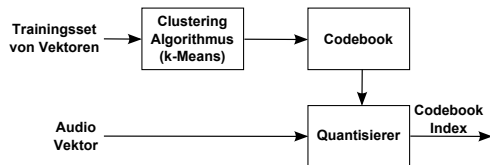
Codebooks

- konfiguriert Vektorquantisierung und Huffman-Kodierung
- bietet abstrahierten Zugriff auf Huffman-Tabelle

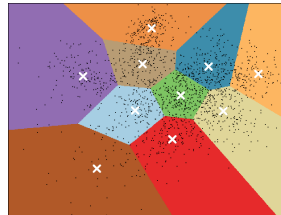
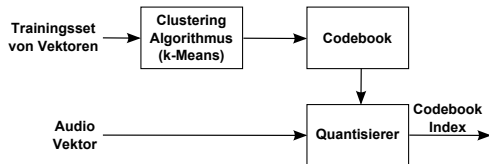
Vektorquantisierung



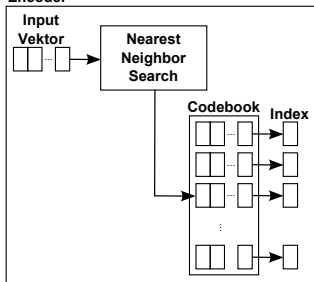
Vektorquantisierung



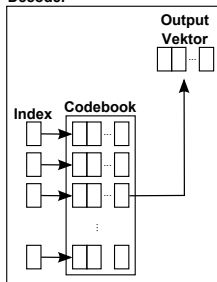
Vektorquantisierung



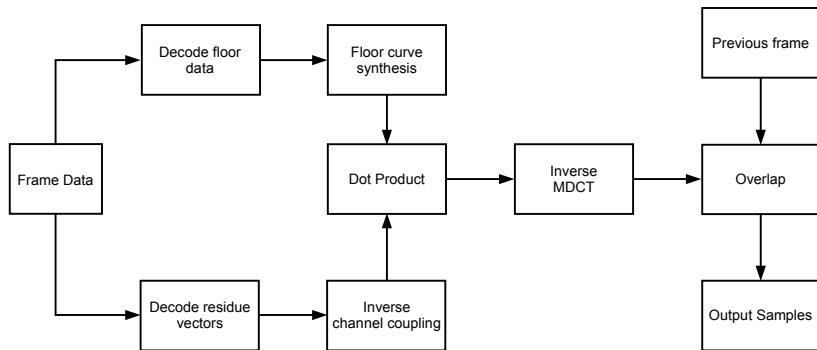
Encoder



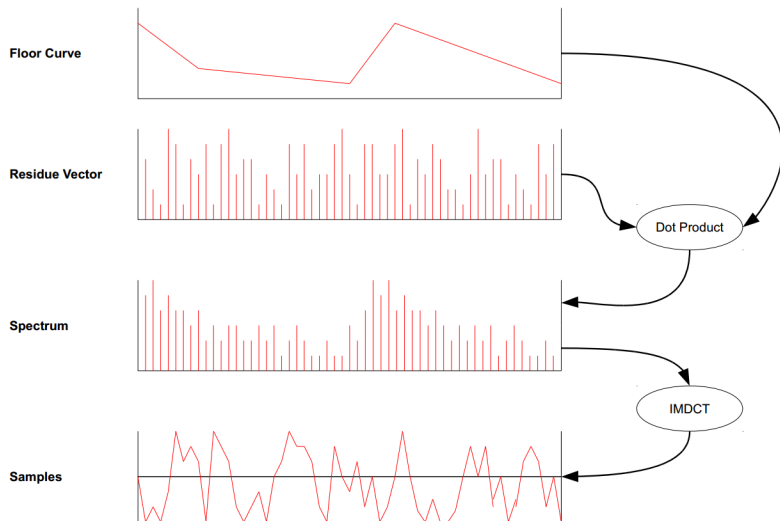
Decoder



Dekodierer



Floor Residue Dotproduct



Soundqualität

	MP3	Ogg Vorbis
64 kb/s	105 KB 	111 KB 
128 kb/s	209 KB 	220 KB 
192 kb/s	318 KB 	334 KB 
320 kb/s	505 KB 	567 KB 

Song: Michael Jackson - Beat it (©Epic Records)

Vor- und Nachteile

	MP3	Ogg Vorbis
Patentbehaftet	- ja	+ nein
Verbreitung Unterstützung	+	-
Bitraten	- 32-320 kb/s	+ 32-500 kb/s
Abtastfrequenz	- 32-48 kHz	+ 8-192 kHz
Kanäle	- 1-2	+ 1-255
Kodierer Speicherverbrauch	+	-
Dekodierer Komplexität	-	+
Soundqualität	-	+

Zusammenfassung

MP3 und Ogg Vorbis:

- sind verlustbehaftete Audiokompressionsverfahren
- machen sich psychoakustische Effekte zunutze

Zusammenfassung

MP3 und Ogg Vorbis:

- sind verlustbehaftete Audiokompressionsverfahren
- machen sich psychoakustische Effekte zunutze

Wichtigste Unterschiede:

- MP3 ist patentbehaftet
- Ogg Vorbis ist nicht so weit verbreitet wie MP3
- Ogg Vorbis ist wesentlich flexibler