

# Inhalt

## ■ Multimedia-Daten

### ◆ Audio

- ☞ Codierung (PCM, ADPCM, ...)
- ☞ Kompressionsverfahren (MPEG, Fraunhofer MP3, ATRAC)
- ☞ wichtige Formate (PCM, MPEG)

### ◆ Images

- ☞ Codierung (Farbtiefen, Alpha-Kanal)
- ☞ verlustlose/verlustbehaftete Bildkompression
- ☞ wichtige Formate (GIF, JPEG, TIFF)

### ◆ Video

- ☞ analoges Video-Signal
- ☞ Standards (PAL, NTSC, HDTV, CCIR)
- ☞ digitales Video
- ☞ Standards (DVB)
- ☞ Codierung (YUV, 4:2:2, ...)
- ☞ Kompressionsverfahren (M-JPEG, MPEG, Quicktime)

# Audio

## ■ Codierung

### ◆ abhängig von Bandbreite und Speicherkapazität

#### ☞ PCM Pulse Code Modulation

- am weitesten verbreitet (CD)
- hoher Bandbreitenbedarf
- für hochqualitative Audioübertragung
- lineare und nicht-lineare Quantisierung

#### ☞ Sonderformen

- DPCM, ADPCM (Differential PCM, Adaptive Differential PCM)
  - fehlerbehaftet
  - geringerer Bandbreitenbedarf

#### ☞ verlustbehaftete Verfahren (psychoakustische Codierung)

- MPEG-1 (Layer 1-3)
- ATRAC (Mini Disk)

# Audio

## ■ PM - Pulse Modulation

- ◆ Digitalimpulse stellen durch Position/Anzahl/Dauer die Codierung der Quelldaten dar
- ◆ verschiedene Verfahren

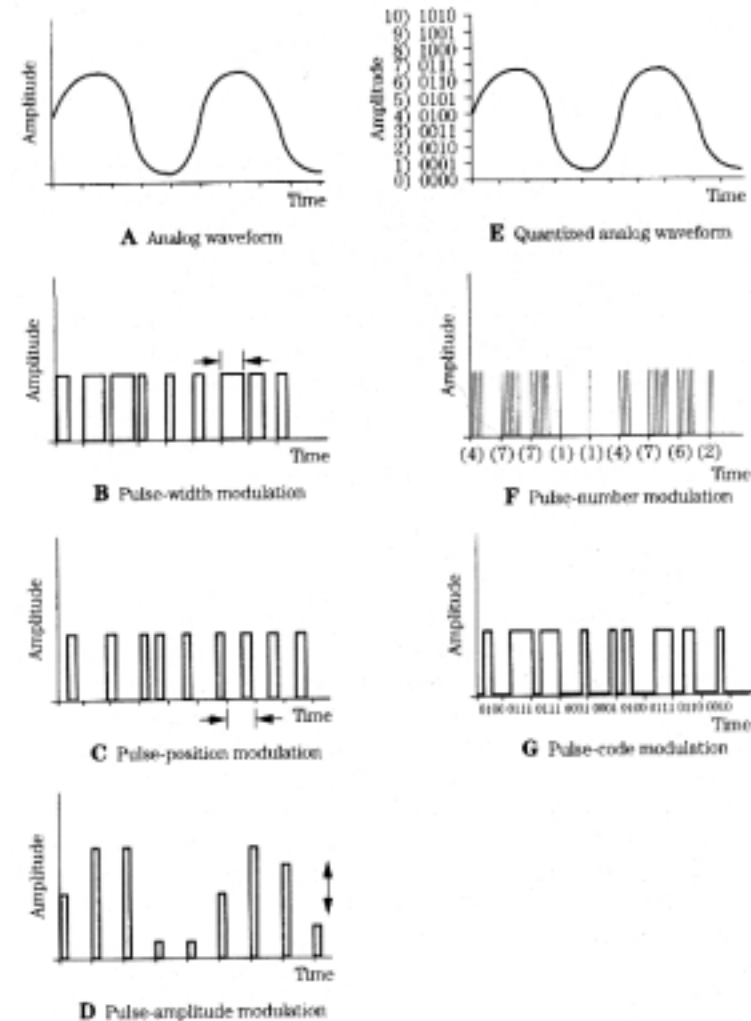


Figure 3.1 PWM, PPM, and PAM are examples of pulse-parameter modulation. PNM and PCM are examples of numerical pulse parameter modulation.

# Audio

## ■ PCM - Pulse Code Modulation - 1

- ◆ 1937 von Alec Reeves beschrieben
- ◆ analoges Signal durch A/D-Wandlung in Digitalworte umgesetzt
- ◆ Binärwort entspricht der Amplitude des Originalsignals
- ◆ auf Empfängerseite wird durch D/A-Wandler wieder entsprechendes analogsignal rekonstruiert
- ◆ Dynamik, Bandbreite, ... (lt. Shannon, Nyquist)
- ◆ hoher Bandbreitenbedarf
- ◆ meist mit linearer Quantisierung (schlechter SNR bei niedriger Auflösung!)
- ◆ z.B. CD-Audio

2 Kanäle, 44,1 kHz Abtastrate, 16 Bit Abtastworte

1,4112 Mbit/s Datenrate

# Audio

## ■ PCM - Pulse Code Modulation - 2

### ◆ nicht lineare Quantisierung

- ☞ Reduktion der Datenrate durch Reduktion der Auflösung
- ☞ durch Compander SNR bei niedrigen Signalamplituden hoch
- ☞ großer Dynamikumfang durch Kompressor möglich
- ☞ z.B. Telefoniebereich  $\mu$ -Law und A-Law Kompression (ITU-T G.711)
  - ermöglicht Dynamikumfang von 12 Bit bei 8-Bit-Codierung
  - 8 kHz Samplingrate, 8 Bit Codewort: 64 kbit/s Bandbreite

$$y = \frac{\log(1 + \mu x)}{\log(1 + \mu)}$$

$\mu$ : Kompressionsfaktor (üblich  $\mu=255$ )

$x$ : Eingangswert ( $x \geq 0$ )

$y$ : komprimierter Ausgangswert

# Audio

## ■ DPCM - Differential Pulse Code Modulation

- ◆ Anstelle der absoluten Codeworte werden nur Differenzen zwischen aufeinanderfolgenden Codeworten übertragen
- ◆ prinzipielle Funktionsweise
  - ☞ aus vergangenen Samplewerten wird “Prädiktor” berechnet
  - ☞ Prädiktor wird vom aktuellen Samplewert abgezogen -> Prädiktionsfehler, dieser wird (mit wenigen Bits codiert) übertragen
  - ☞ auf Empfängerseite wird aus vergangenen rekonstruierten Samples ebenfalls Prädiktor berechnet und durch übertragenen Prädiktionsfehlerwert auf Original-Sample-Wert korrigiert
- ◆ Vorteil:
  - ☞ geringer Bandbreitenbedarf, sehr einfache Hardware
- ◆ Nachteil:
  - ☞ 6dB/Oktave Pegelverlust (SNR sinkt um 6dB/Oktave) -> Höhen schlechter übertragen (Abhilfe: Höhen vor Codierung anheben)
  - ☞ Prädiktionsfunktion beeinflusst Qualität (quellendatenabhängig)

# Audio

## ■ DM - Delta Modulation - 1

- ◆ Extremform der DPCM

- ◆ sehr hohe Abtastrate

- ☞ Audiosignal kann sich in kurzer Zeit nur sehr wenig ändern

- ☞ es reicht daher 1-Bit Quantisierungswert aus

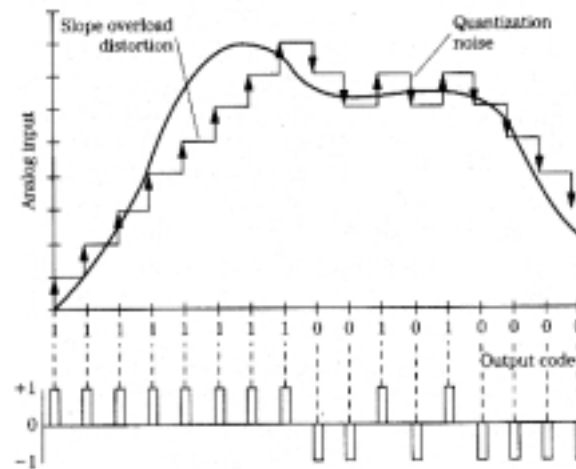
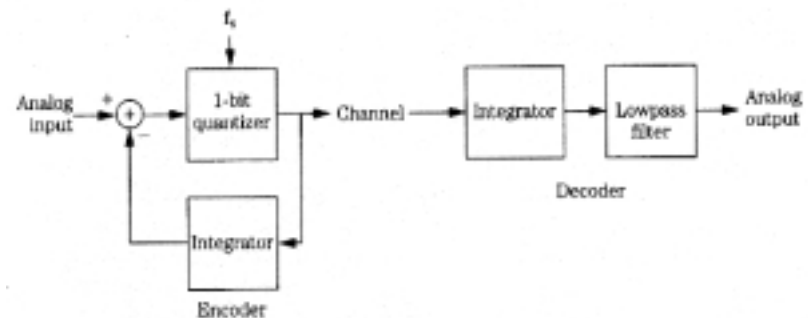


Figure 4.20 In a delta-modulation coder, one differential bit is used to encode the audio signal.



# Audio

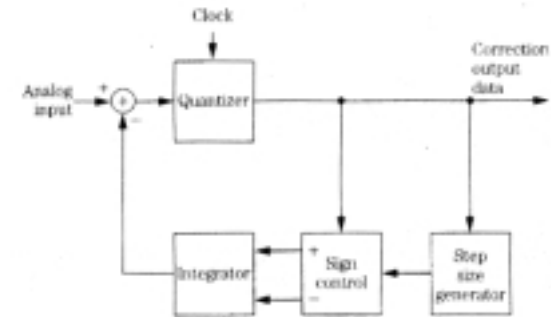
## ■ DM - Delta Modulation - 2

### ◆ Nachteil:

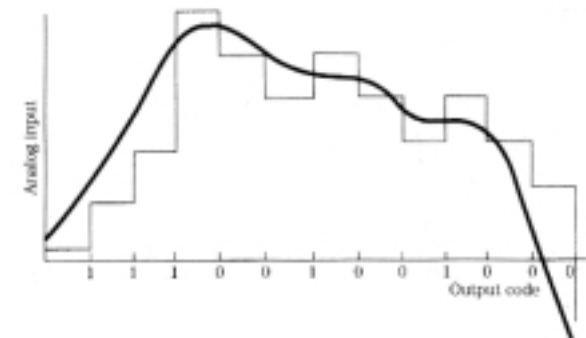
- ☞ Quantisierungsfehler sehr hoch für schnell wechselnde Signale
- ☞ hohe Abtastrate notwendig: 96dB SNR/20kHz Bandbreite benötigt 200MHz DM-Abtastrate

### ◆ ADM - Adaptive Delta Modulation

- ☞ passt Quantisierungsschrittgrößen an
- ☞ Probleme bei hochfrequenten Signalen mit Adaption



A Block diagram of an adaptive delta modulation encoder, with variable step size.



B Changes in step size are triggered by continuous 1s or 0s.



# Audio

- ADPCM - Adaptive Differential PCM - 1
  - ◆ kombiniert PCM mit ADM-Differenzsignal
  - ◆ verschiedene Verfahren, meist jedoch
    - ☞ Differenzsignal wird adaptiv skaliert dann quantisiert
  - ◆ Skalierungsfaktor wird nach unterschiedlichen Kriterien gebildet
    - ☞ meist dient als Grundlage die durchschnittliche Signalamplitude
      - große Signalunterschiede bewirken großen Skalierungsfaktor
      - kleine Signalpegel dagegen einen kleinen Faktor (feinere Schritte -> bessere Auflösung)
  - ◆ ADPCM erreicht damit
    - ☞ Bitratenreduktion abhängig von Amplitude und Spektrum des Audiosignals
    - ☞ 32kbps und weniger Datenrate für Sprache leicht erreichbar

# Audio

## ■ ADPCM - Adaptive Differential PCM - 2

### ◆ Vergleich ADPCM mit PCM

- ☞ 4 Bit ADPCM entspricht Qualitativ 8 Bit PCM (1:2 Kompression)

### ◆ Einsatzgebiete

- ☞ Telekommunikationsbranche
  - Voice-Kompressoren (2-4xVoice über 64 kb/s ISDN)
- ☞ CD-ROM/XA
  - 8 Bit ADPCM für 90dB SNR und 17kHz Bandbreite
  - 4 Bit ADPCM für 60dB SNR und 8.5/17kHz Bandbreite
  - Bitraten von 80kb/s bis 309kb/s
- ☞ Quicktime/WindowsMedia
  - einfacher ADPCM Decoder läßt sich mit Software sehr effizient implementieren (Integer Multiplikationen und Additionen)

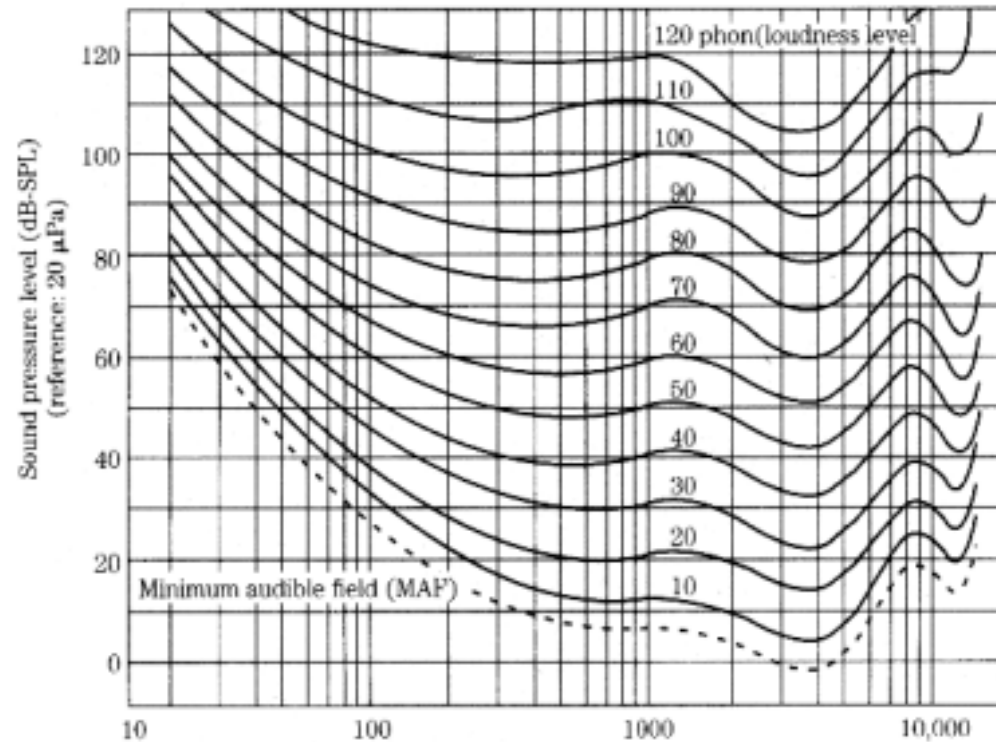
# Audio

- Psychoakustische Verfahren (perceptual coding) - 1
  - ◆ Codec identifiziert nicht-wahrnehmbare (damit irrelevante) Anteile des Audiosignals
  - ◆ reduzierter Datenstrom wird dann redundanzmindernd Codiert
  - ◆ Quantisierungsfehler (-rauschen) soll dabei unter der Wahrnehmungsschwelle bleiben
  
  - ◆ wichtig für geeignete Codierungsverfahren
  - ◆ Eigenschaften des menschlichen Gehörs
    - ☞ Frequenzempfinden logarithmisch (1 Oktave = doppelte Frequenz)
    - ☞ Frequenzempfinden lautstärkeabhängig (lautere Signale werden als Niederfrequenter empfunden, bis zu 10% Toleranz möglich)
    - ☞ Lautstärkeempfinden ist Frequenzabhängig, auch Hörschwelle
    - ☞ nahe beieinander liegende Töne verschmelzen (unter 200Hz)
    - ☞ Dynamikumfang des Ohres ist 120dB

# Audio

## ■ Psychoakustische Verfahren - 2

- ◆ Frequenz/Lautstärkeempfinden des menschlichen Ohres



benötigter Schalldruckpegel abhängig von Frequenz um gleichen Lautstärkeindruck wie bei 1kHz zu erhalten, größte Empfindlichkeit zwischen 1 kHz und 5 kHz

# Audio

## ■ Psychoakustische Verfahren - 3

### ◆ Innenohr (Basiliarmembran + Haarzellen+ Nervenzellen)

- ☞ Spektralanalyse des eintreffenden Schallereignisses
- ☞ Haarzellen reagieren auf stärkste Signal in ihrer Frequenzumgebung
  - es ergeben sich gewisse Frequenzbänder “critical bands” (Fletcher)
- ☞ Frequenzauflösung bei tiefen Frequenzen viel besser wie bei hohen
- ☞ 4/5 der Critical Bands liegen unter 5 kHz
- ☞ Ohr erhält mehr Infos aus tiefen Frequenzen

$$\text{Bandbreite eines Critical Band} = 24.7 (4.37F+1) \text{ [Hz]}$$

F = Bandmittenfrequenz in kHz

Bsp: F = 1kHz B = 160Hz

Liegt Sinus mit 1kHz an, wird alles was im Critical Band von 160Hz um Sinus-Ton liegt nur dann wahrgenommen, wenn es lauter als Sinuston ist (Überdeckungseffekt)

# Audio

- Psychoakustische Verfahren - 4
  - ◆ Modell f. Critical Bands (Zwicker)

Critical band number (Bark)	Center frequency (Hz)	Critical band (Hz)	Lower cutoff frequency (Hz)	Upper cutoff frequency (Hz)
1	50	—	—	100
2	150	100	100	200
3	250	100	200	300
4	350	100	300	400
5	450	110	400	510
6	570	120	510	630
7	700	140	630	770
8	840	150	770	920
9	1000	160	920	1080
10	1170	190	1080	1270
11	1370	210	1270	1480
12	1600	240	1470	1720
13	1850	280	1720	2000
14	2500	320	2000	2320
15	2900	380	2320	2700
16	2900	450	2700	3150
17	3400	550	3150	3700
18	4000	700	3700	4400
19	4800	900	4400	5300
20	5800	1100	5300	6400
21	7000	1300	6400	7700
22	8500	1800	7700	9500
23	10,500	2500	9500	12,000
24	13,500	3500	12,000	15,500
25	18,775	6550	15,500	22,050

wichtig: nicht fix, entstehen immer um hörbaren Ton herum

1 bark=1 critical band (relativiert absolute Frequenz und wahrgenommene Tonhöhe)

# Audio

## ■ Psychoakustische Verfahren - 5

### ◆ neue Designkriterien für Codecs

- ☞ Irrelevanzreduktion muß sich an den Eigenschaften des Gehörs orientieren
- ☞ Dinge die nicht gehört werden, werden auch nicht übertragen
- ☞ Codec kann z.B. bei wenig Informationsgehalt niedrige Bitrate liefern, bei viel Information entsprechend höhere

### ◆ steht im Gegensatz zum bisherigen Audio-System-Design wo

- ☞ linearer Frequenzgang
- ☞ minimales Rauschen
- ☞ ...

### ◆ die Designkriterien waren

### ◆ “PCM ist Overkill” (wir hören sowieso nicht alles)

# Audio

## ■ Psychoakustische Verfahren - 6

### ◆ Hörschwelle und Maskierung

#### ◆ Hörschwelle ist Grenze wo Ohr einen Ton wahrnimmt

- ☞ immer auf 1kHz Ton bezogen (0 dB)
- ☞ -> zwei Töne mit unterschiedlicher Frequenz aber gleichem Schalldruck hören sich nicht gleich laut an
- ☞ -> Empfindlichkeit für Rauschen/Verzerrungen stark  
Frequenzabhängig
- ☞ Codec kann alle Signale, die unter Hörschwelle liegen, streichen

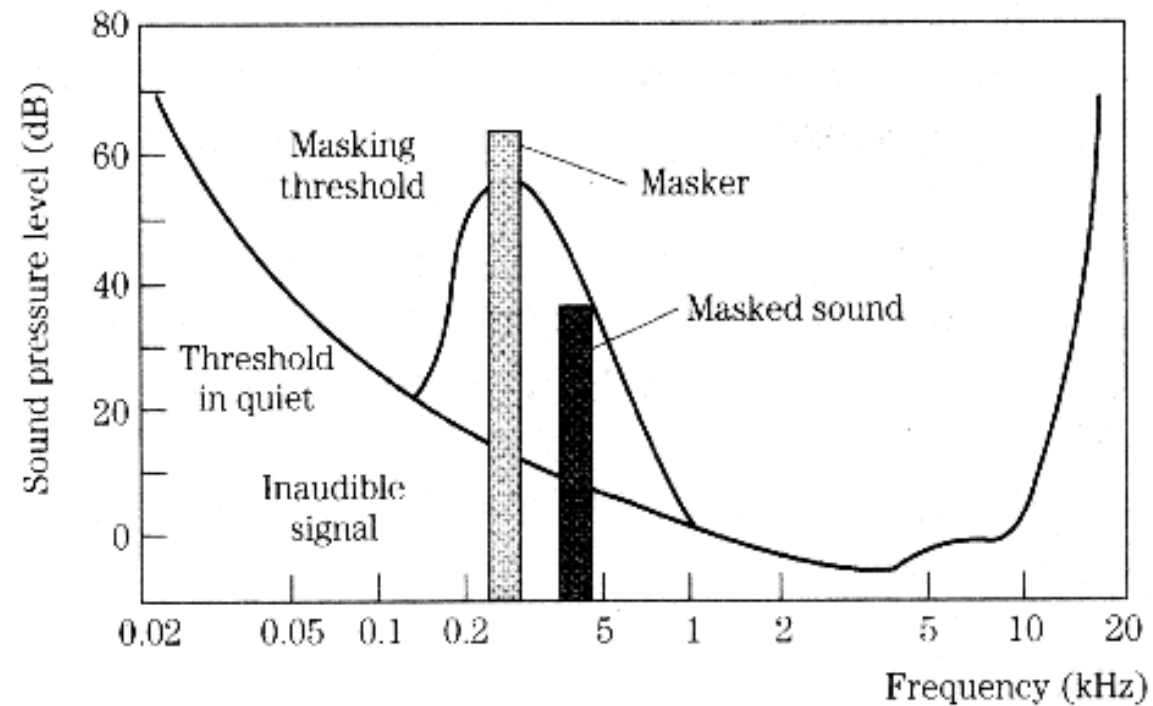
#### ◆ Maskierung tritt immer dann auf wenn

- ☞ lauter Ton in “seinem” critical band, leisere Töne überdeckt
- ☞ Codec braucht leisen Ton nicht zu codieren



# Audio

- Psychoakustische Verfahren - 7
  - ◆ Hörschwelle und Maskierung

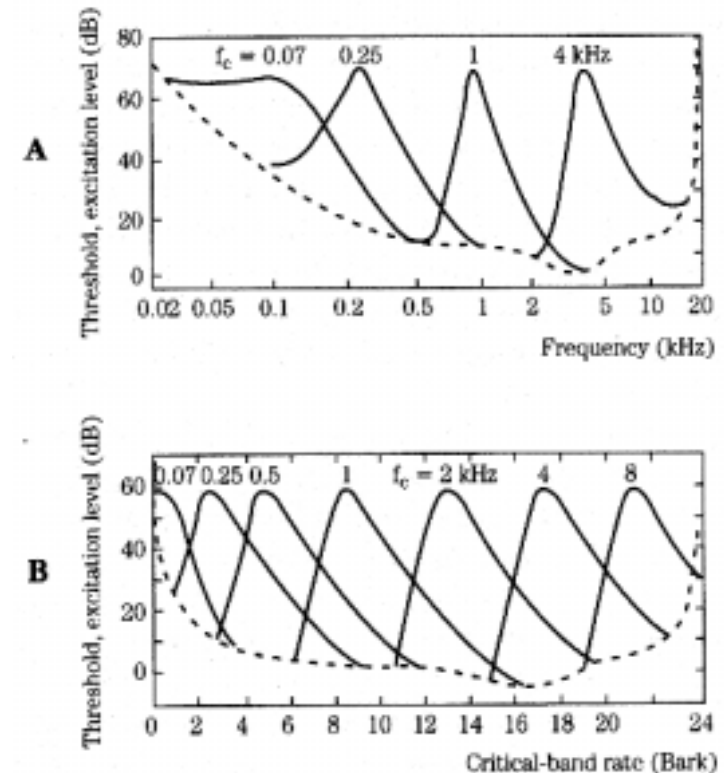


# Audio

## ■ Psychoakustische Verfahren - 8

### ◆ Maskierung

- ☞ frequenzabhängig
- ☞ komplexe Signale (mit Oberwellen) haben breiteres Critical Band und können auch höherfrequente Signale maskieren
- ☞ sehr laute Signale können auch mehrere critical bands überdecken
- ☞ höherfrequente Töne maskieren seltener niederfrequente (Maskierungsschwelle fällt auf niederfrequenter Seite mit ca. 27dB/bark, auf hochfrequenter Seite mit 5-20dB/bark)

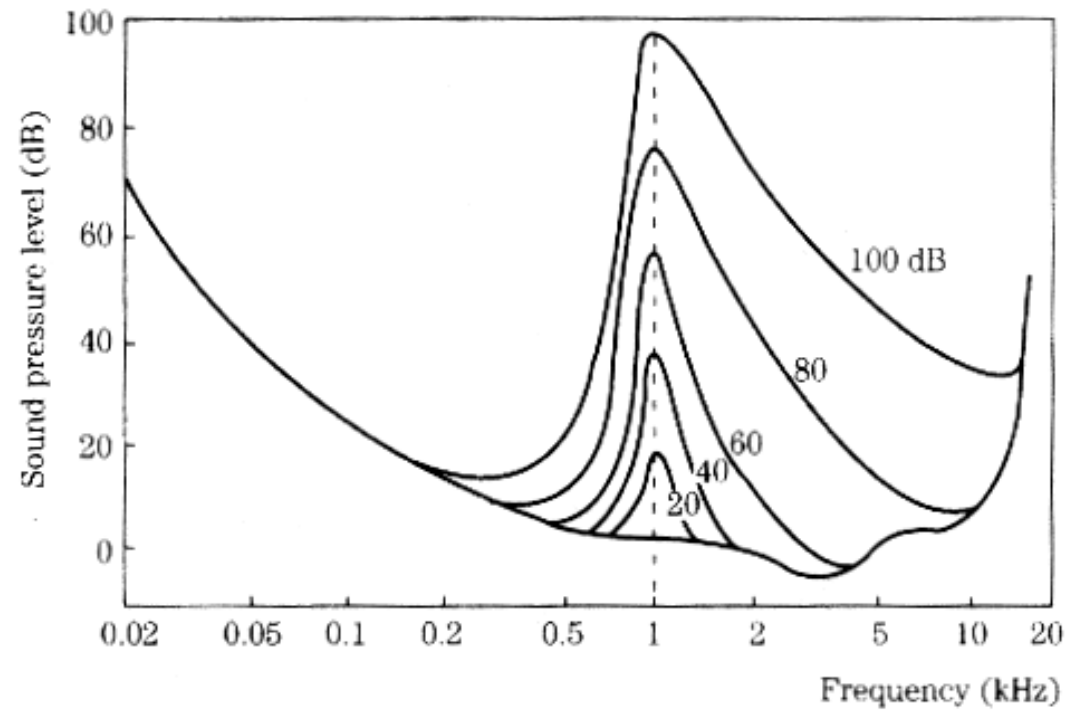


# Audio

## ■ Psychoakustische Verfahren - 9

### ◆ Maskierung

- ☞ pegelabhängig
- ☞ Bsp: 1 kHz Signal mit unterschiedlichen Pegeln

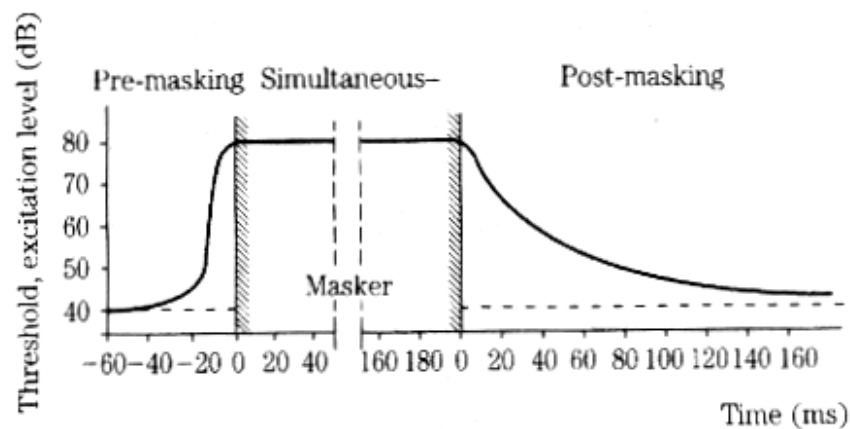


# Audio

## ■ Psychoakustische Verfahren - 10

### ◆ Maskierung

- ☞ zeitabhängig
- ☞ gleichzeitiges Auftreten des Maskierungstones (simultaneous masking)
- ☞ Signal kann auch durch später auftretendes Signal maskiert werden (pre/backward masking) - nur wenige ms
- ☞ Signal kann durch früher endendes Signal weiter maskiert werden (post/forward masking) - einige ms, aber nur von niederfrequenten

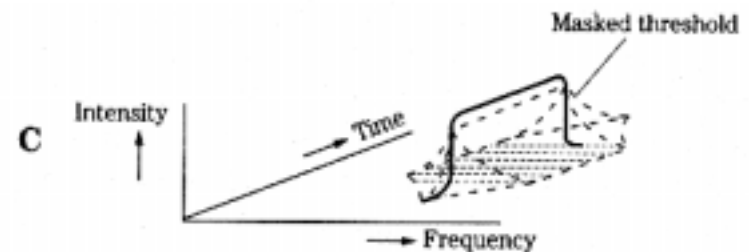
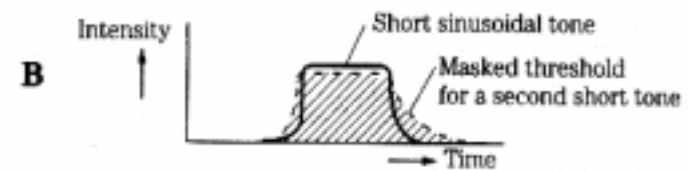
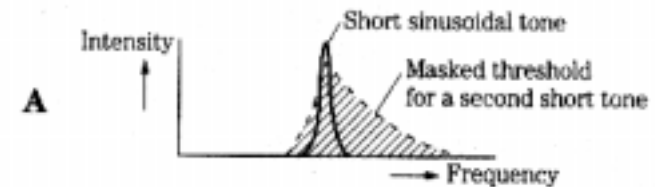


# Audio

## ■ Psychoakustische Verfahren - 11

### ◆ Maskierung

- ☞ Kombination aus zeit- und frequenzabhängiger Maskierung ergibt “Schwellwertfläche” im Zeit/Frequenzbereich, unter deren Pegel nichts wahrgenommen wird
- ☞ dort kann Codec die Quantisierungsfehler “ansiedeln”



### ◆ Stereo/Mehrkanal

- ☞ Kanäle sind stark korreliert
- ☞ Stereosignal ist nicht 2xMono
- ☞ Richtungshören stark amplitudenabhängig, weniger von Phase
- ☞ auch hier Maskierung

# Audio

## ■ Psychoakustische Verfahren - 12

- ◆ Warum nun psychoakustische Verfahren?

- ◆ Bitratenreduktion 2:1 bis 12:1 möglich

  - ☞ 48kHz Signal benötigt dann zwischen 70 und 768 kbps)

- ◆ Störfestigkeit höher wie bei PCM

  - ☞ gestörtes Datenwort trägt zur temporären Erhöhung des Quantisierungsrauschens bei, bei PCM gibt's "Knackser"

- ◆ Aber

  - ☞ heute werden Signale oft mehrmals Codiert/Decodiert

  - ☞ -> steigender Quantisierungsfehler

  - ☞ -> Rauschen kann über Maskierungsgrenzen steigen -> benötigt Bits zur Codierung die für Nutzsignal fehlen

  - ☞ Abhilfe: immer von höchstqualitativem System ausgehen (z.B. zuerst PCM, dann erst andere Systeme)

# Audio

## ■ Psychoakustische Verfahren - Prinzip

- ◆ Reduktionsverfahren beruhen auf Spektralanalyse des Audiosignales
- ◆ Grundlage dafür ist Fourier Transformation (Zeit -> Frequenzbereich)
- ◆ Samples/Spektralanteile durch psychoakustisches Modell gefiltert

## ◆ 2 Arten von Reduktionsverfahren heute üblich

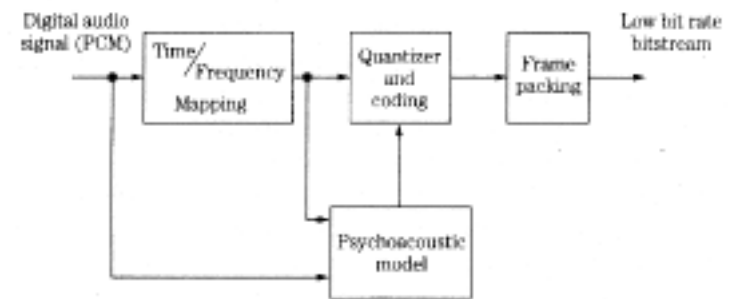
### ☞ Subband-Codecs

- Frequenzteilmäander um Sample-Zahl zu reduzieren

### ☞ Transformations-Codecs

- Spektrum (viele Frequenzbänder) um Frequenzanteile zu Filtern

### ☞ Prinzip der Verfahren eigentlich ident (adaptive Filterbank)



A



B

A: Encoder, B: Decoder

# Audio

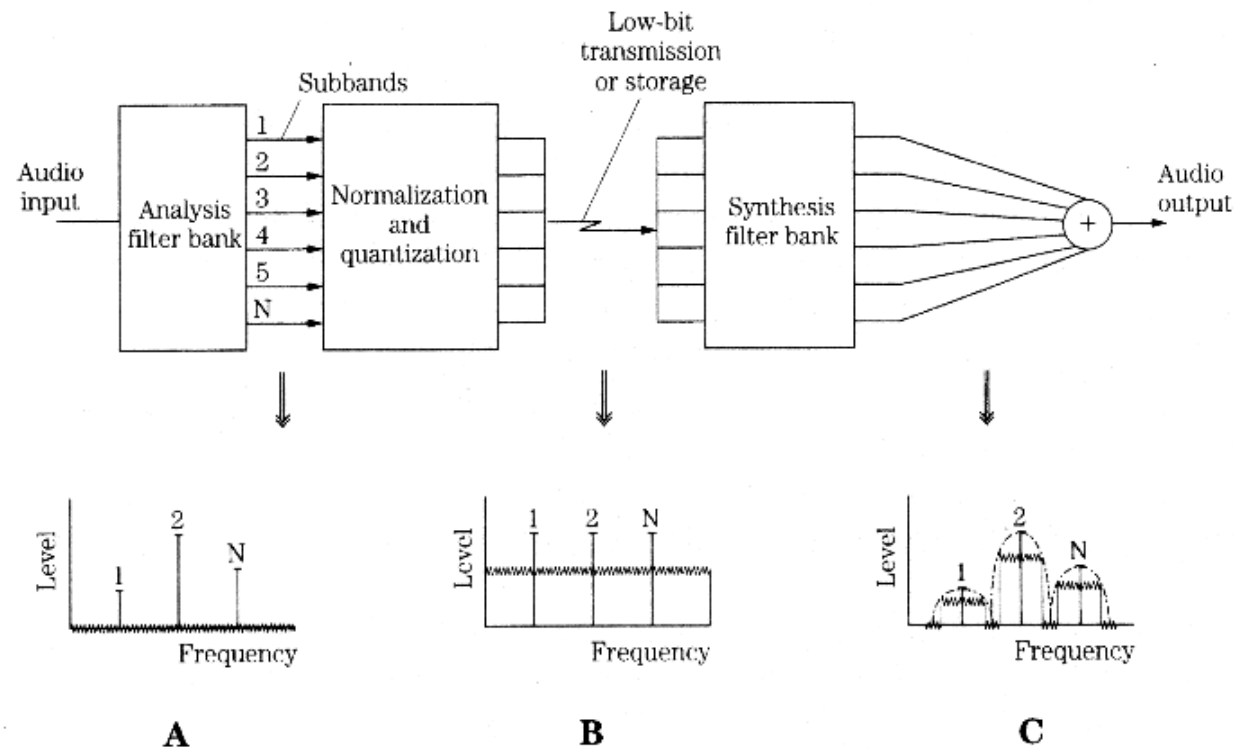
## ■ Subband Codec (Bell Labs, ca. 1980)

- ◆ hybride Technologie aus Zeit/Frequenzbereich-Verarbeitung
- ◆ Funktionsweise
  - ☞ eingehender kurzer Block von Samples (3-4ms Länge) wird durch Bandpaß-Filterbank (Digitalfilter, 100dB Flankensteilheit) in bis zu 32 Teilbänder (entsprechend den “critical bands” des Ohres) aufgespaltet
  - ☞ über “Side Chain” FFT wird spektrale Energie in den Subbändern festgestellt
  - ☞ psychoakustisches Modell bestimmt daraus welche Teilband-Samples wie quantisiert werden
  - ☞ Quantisierung der Teilbänder erfolgt meist dynamisch, abhängig von der zur Verfügung stehenden Übertragungsbandbreite und der Anzahl der zu quantisierenden Teilbänder
  - ☞ gesamter Prozeß wird für jeden Sample-Block wiederholt
  - ☞ auf Decoderseite wird über eine Synthese-Filterbank die Summe der übertragenen Teilbänder zum Gesamtsignal zusammengefügt



# Audio

## ■ Subband Codec - 2



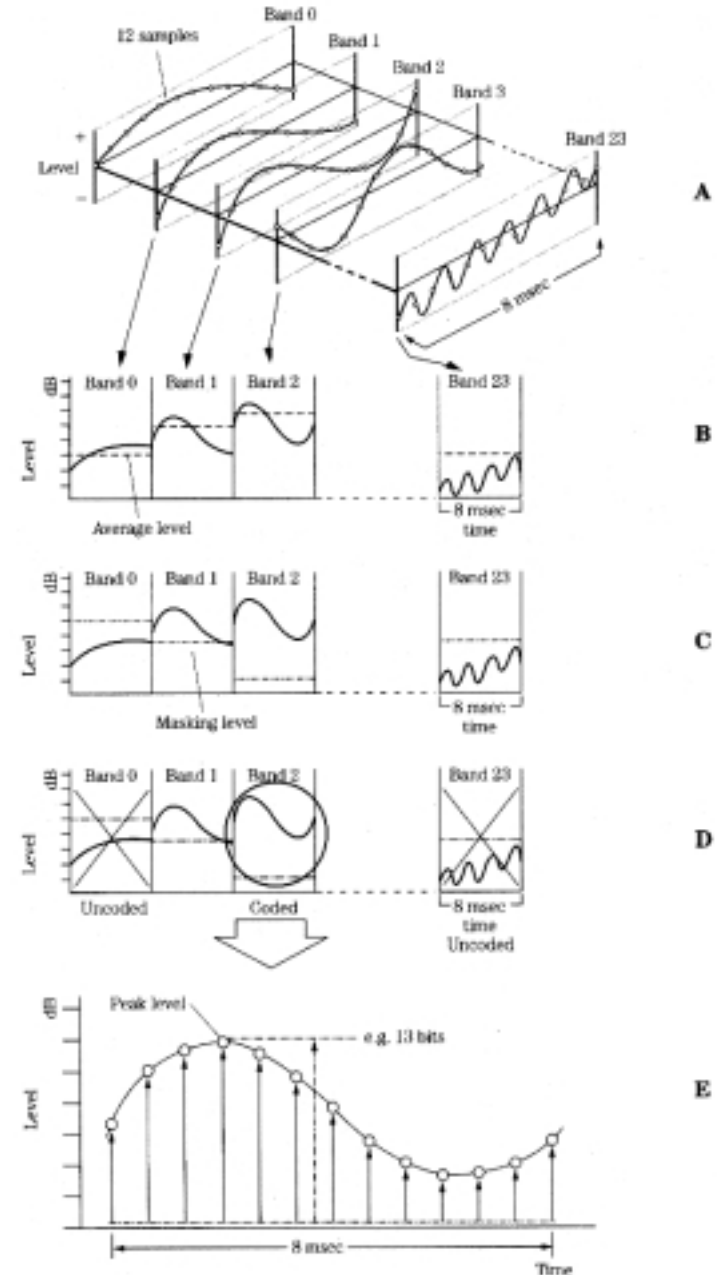
- ◆ zusätzliche Side-Chain mit FFT bringt Vorteil, daß spektrale Energie in Teilbänder und Maskierung zwischen Teilbändern leicht bestimmt werden kann

# Audio

## ■ Subband Codec - 3

### ◆ “Big Picture”

- ☞ A: Samples auf Filterbank
- ☞ B: durchschnittlicher Pegel in jedem Teilband
- ☞ C: Maskierungspegel in Teilbändern
- ☞ D: streichen der maskierten Bänder und Feststellen, welche codiert werden
- ☞ E: Quantisierung in Abhängigkeit von der Amplitude des Signals



# Audio

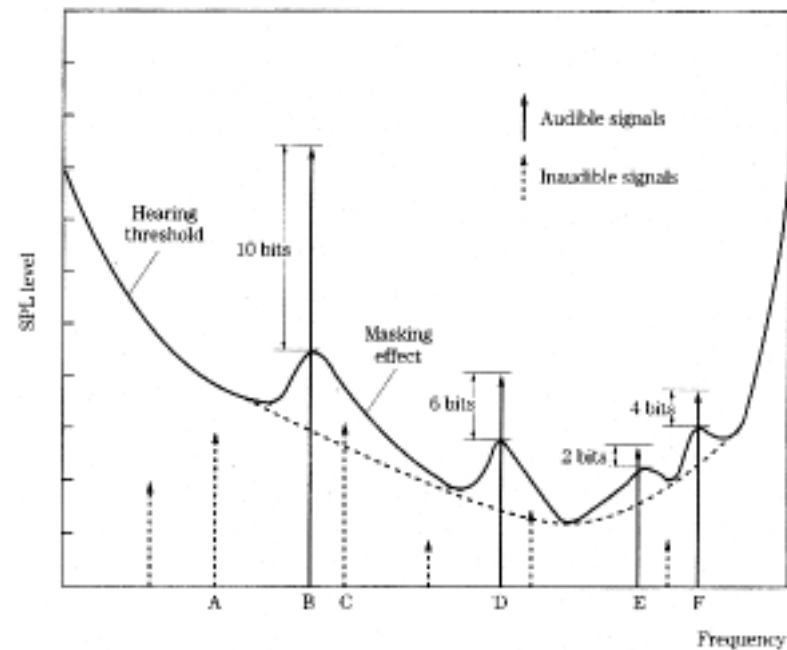
## ■ Subband Codec - 4

### ◆ Bit Allocation - Quantisierung

- ☞ Quantisierung der einzelnen Subband-Signale erfolgt dynamisch
- ☞ allokierte Bitzahl für jedes Teilbandsignal wird berechnet aus
  - Pegel des Signals über Hörschwelle
  - Pegel des Signals in Relation zu anderen zu übertragenden Signalen

### ☞ Verbesserung:

- Bitallokation abhängig von Pegel über Maskierungsschwelle -> reduziert benötigte Bitzahl
- ☞ sollten nach einem Allokierungsdurchlauf noch Bits im "Bandbreitenpool" vorhanden sein, erneute Iteration zur Quantisierung, u.U. kommen auch maskierte Subbänder dann zum Zug (wenn genug Bits da sind)



# Audio

## ■ Transformations Codec - 1

### ◆ Frequenzbereich-Verarbeitung

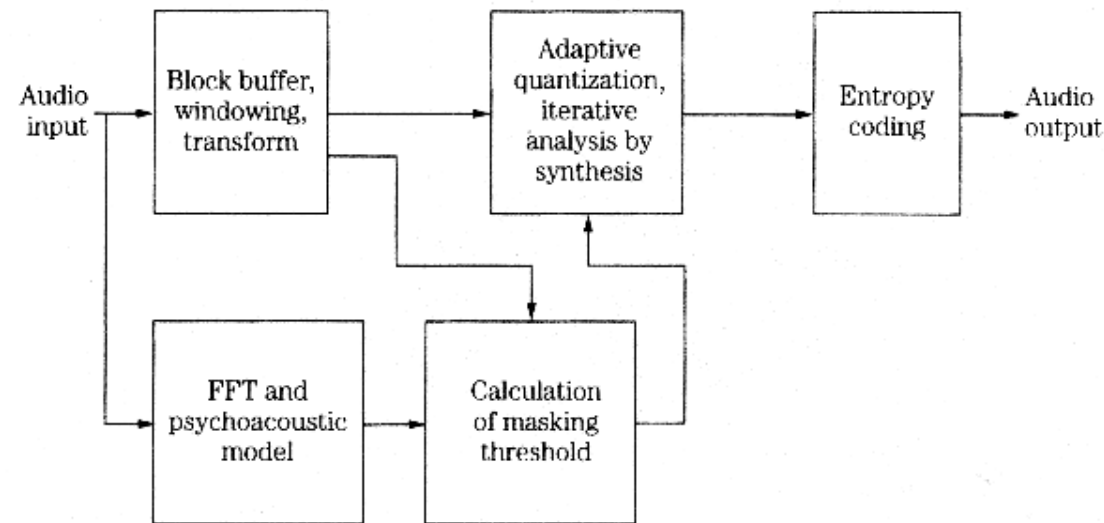
### ◆ Funktionsweise

- ☞ Block von Samples wird in Frequenzbereich transformiert (FFT, DCT, Reduktion von z.B. 1024 Samples in 512-FFT-Punkte)
- ☞ Spektralkoeffizienten werden in Subbänder zusammengefaßt (ca. 32) um eine "critical band" Analyse zu erlauben
- ☞ Spektralkoeffizienten dann über psychoakustisches Modell quantisiert
  - bei adaptiver Quantisierung wird für jedes Subband die Bitanzahl dynamisch festgelegt (innerhalb des Bandes wird aber immer gleich quantisiert)
- ☞ Quantisierung meist wieder iterativ um vorhandene Bandbreite optimal ausnutzen zu können
- ☞ gesamter Prozeß wird für jeden Sample-Block wiederholt
- ☞ auf Decoderseite wird über eine inverse FFT/DCT aus den Spektralkoeffizienten wieder das Gesamtsignal erzeugt

# Audio

## ■ Transformations Codec - 2

### ◆ Blockdiagramm



### ◆ Probleme

- ☞ um gute Frequenzauflösung zu erhalten braucht man lange Sample-Blocks
- ☞ dadurch wird Zeitauflösung sehr schlecht
- ☞ es kann z.B. vor Signal schon "pre-echo" (Störton, Rauschen) auftreten
- ☞ Abhilfe: Überlappung der Sample-Blocks um bis zu 50%

# Audio

## ■ Codecs

### ◆ MPEG-1

- ☞ (ISO/IEC 11172 “Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s”)
- ☞ Moving Pictures Expert Group, 1992
- ☞ erlaubt Audio/Videowiedergabe mit “normaler” Audio-CD-Bitrate
- ☞ erlaubt Codierung von 32, 44.1 und 48kHz Samplerate Mono/Stereo
- ☞ Output-Bitraten meist Vielfache von 64kbit/s
- ☞ MPEG-1 Layer I bis III heute üblich
- ☞ MPEG-1 spezifiziert nur Decoder, Encoder nicht definiert und daher jederzeit neuer Technologie anpaßbar
- ☞ MPEG-2 erweitert um Multi-Channel-Mode (AAC)

### ◆ ATRAC (Adaptive TRansform Acoustic Coding, 1991)

- ☞ 292 kb/s Datenrate (5:1 Kompression)
- ☞ entwickelt für MiniDisk-Aufzeichnung/Wiedergabe
- ☞ fixe Samplerate 44.1 kHz

# Audio

## ■ MPEG-1

### ◆ Layer I

- ☞ einfachste Version
- ☞ 192 kbps und Kanal

### ◆ Layer II

- ☞ basiert auf Layer I
- ☞ 96-128kbps und Kanal
- ☞ Layer IIA: Joint Stereo Version (128-192 kbps für Stereosignal)

### ◆ Layer III (MP3)

- ☞ unterscheidet sich von Layer I/II
- ☞ 64 kbps pro Kanal
- ☞ liefert bei niedrigen Bitraten bestes Ergebnis

### ◆ Layer II, IIA und III sind broadcasttauglich

- ☞ Codierung mit Layer II oder III mit 2x128 kbps oder 1x192kbps Joint Stereo nicht von 16 Bit PCM System zu unterscheiden

# Audio

## ■ MPEG-1

### ◆ Funktionsprinzip der Codecs

- ☞ Audiosignal durch Filterbank zerteilt
- ☞ Frequenzanalyse in Teilbändern
- ☞ psychoakustisches Modell zur Maskierung und Quantisierung
- ☞ Bitallokationsinformation und codierte Daten werden in Frames gepackt und übertragen

### ◆ Joint-Stereo-Coding

- ☞ Subbandinfos für Links/Rechts werden ab einer gewissen Frequenz in einem Kanal zusammengefasst (3, 6, 9 oder 12 kHz)
- ☞ Amplitudeninformationen werden getrennt übertragen
- ☞ Decoder ändert nur Amplituden der Subbänder, Spektrum von L/R ist ident

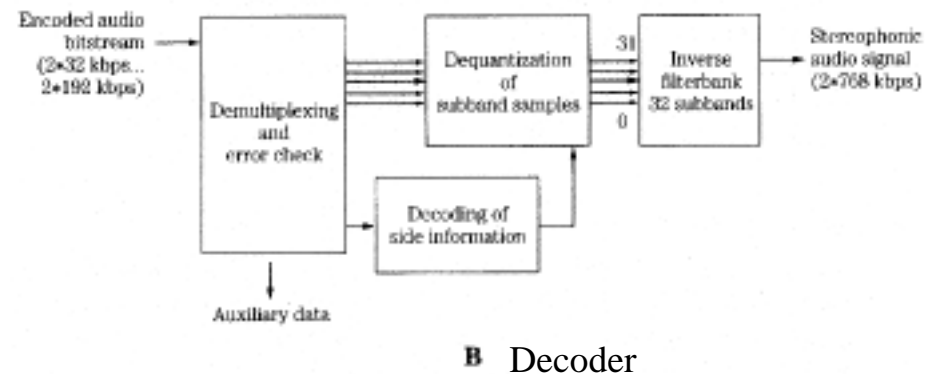
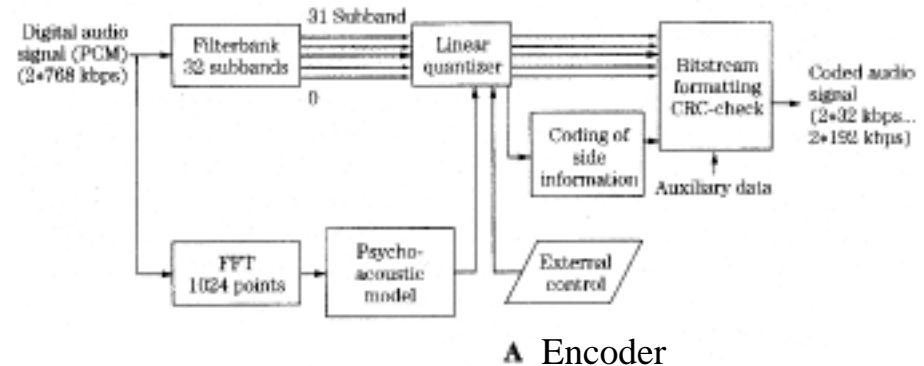


# Audio

## ■ MPEG-1 - Layer I

### ◆ Funktionsprinzip

- ☞ Eingangssignal in 32 gleichbreite Subbänder geteilt
- ☞ 32 Samples liefern 32 Ausgangswerte aus Filter
- ☞ je 12 Werte aus Subband werden gruppiert und gemeinsam behandelt
- ☞ 512-Sample FFT liefert Daten für psychoakustisches Modell
- ☞ Subbanddaten werden dann mit 2-15 Bits quantisiert (nicht Spektraldaten)

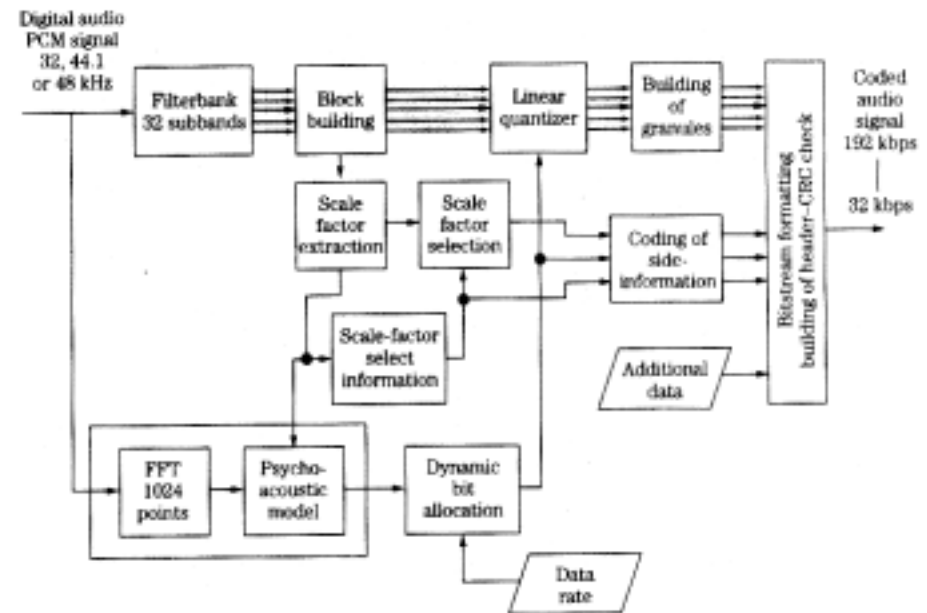


# Audio

## ■ MPEG-1 - Layer II

### ◆ Funktionsprinzip

- ☞ Eingangssignal in 32 gleichbreite Subbänder geteilt
- ☞ je 3x12 Werte aus Subband werden gruppiert und gemeinsam behandelt
- ☞ 1024-Sample FFT liefert Daten für psychoakustisches Modell
- ☞ Subbanddaten werden dann mit 2-15 Bits quantisiert (nicht Spektraldaten)
- ☞ tiefe Subbänder bis zu 15 Bits, mittlere bis zu 7 und hohe nur mit max. 3 Bits quantisiert
- ☞ jeweils 3 Subbandwerte gemeinsamsam quantisiert (granules)



# Audio

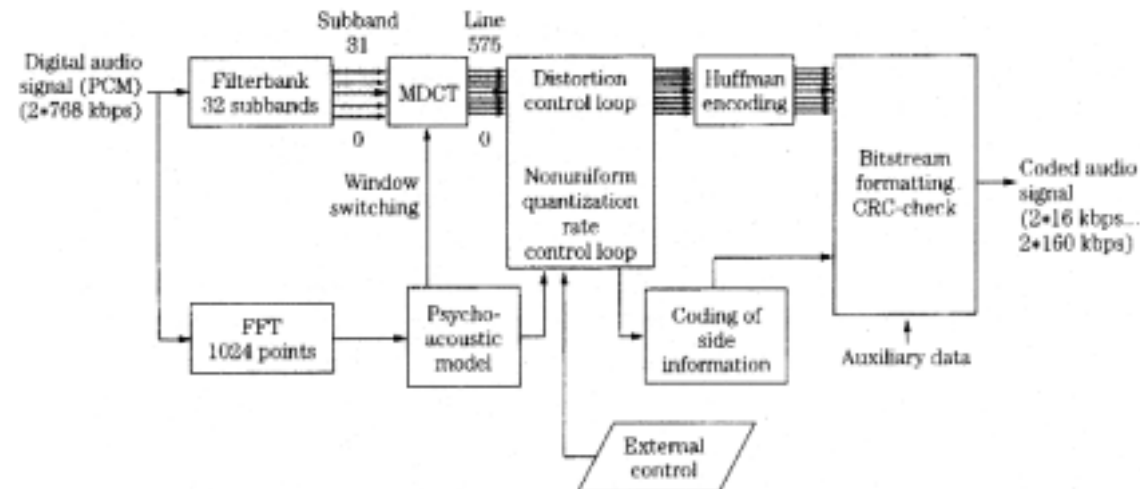
## ■ MPEG-1 - Layer III (MP3)

### ◆ Funktionsprinzip

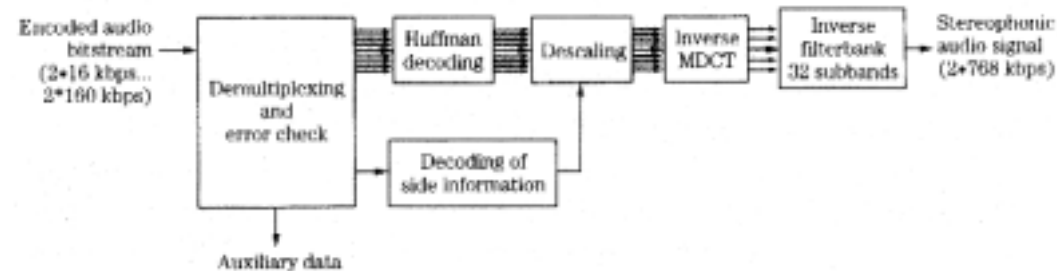
- ☞ Eingangssignal in 32 gleichbreite Subbänder geteilt
- ☞ jedes Subband in max. 18 Spektralkoeffizienten transformiert (MDCT) ergibt bis zu 576 Spektralkoeffizienten entsprechend 41.67 Hz Bandbreite, Zeitauflösung 24 ms bei 48kHz Samplerate
- ☞ dadurch schlechte Zeitauflösung (pre-echo Problem)
  - Lösung: long(36 Samples) / short (12 Samples) DCT-Fenster wird für einzelne Subbänder abhängig von Signaltyp (konstant, transient) umgeschaltet -> min 8ms Zeitauflösung und 192 Spektralkoeffizienten
- ☞ Quantisierung erfolgt nicht-linear (compander) mit 1024-FFT Side-Chain für psychoakustisches Modell
- ☞ Huffman-Codierung der Ausgangsdaten
- ☞ Layer III liefert variable Bitrate am Ausgang -> Problem in Übertragungssysteme -> iteratives Verfahren um "Bandbreitenpool" durch Erhöhung der Quantisierungsauflösung zu verbrauchen

# Audio

- MPEG-1 - Layer III (MP3)
  - ◆ Funktionsprinzip - Encoder / Decoder Blockdiagramm



A



B

# Audio

## ■ MPEG-1 - Layer III (MP3)

### ◆ Stereokodierung (kann pro Frame anders sein)

#### ☞ 4 Arten

- normal (getrennte Kanäle für L(R))
- MS
- intensity (joint stereo)
- intensity (hohe Frequenzen) und MS (tiefe Frequenzen)

#### ☞ MS (mid/side) Stereokodierung

- codiert wird Summe von Links und Rechts und Differenz
- korrelieren L und R sehr stark, ist Differenz minimal und daher geringer Codierungsaufwand

#### ☞ intensity/joint stereo

- stereopanorama wird für subbänder übertragen (allerdings nur für die Skalierungsfaktoren um Bits zu sparen)

# Audio

## ■ MPEG-2 (1994)

### ◆ Erweiterung von MPEG-1

- ☞ Übertragung von 6 Audiokanälen (5.1 = 3 x Front, 2x Surround, 1x Subwoofer bis 120Hz)

### ◆ Auf/Abwärtskompatibel zu MPEG-1

- ☞ MPEG-2 kann alle MPEG-1 Formate wiedergeben
- ☞ MPEG-1 decoder kann Stereoinformationen aus MPEG-2 wiedergeben
- ☞ L+R Kanäle enthalten Matrixcodierte Audioinformation aller Kanäle, diese werden MPEG-1 konform übertragen
- ☞ Zusatzinfos in MPEG-2 Erweiterung übertragen

### ◆ Sampleraten von 32, 44.1 und 48 kHz (auch 16, 22.05 und 24 kHz für Layer I-III, nicht MPEG-1 kompatibel)

### ◆ Gesamt-Bitraten von 386kbps bis 1066 kbps

- ◆ Erweiterung: AAC (Advanced Audio Coding, 1997): 8-96kHz Samplingrate
- ◆ Andere: Dolby Digital (AC3),...

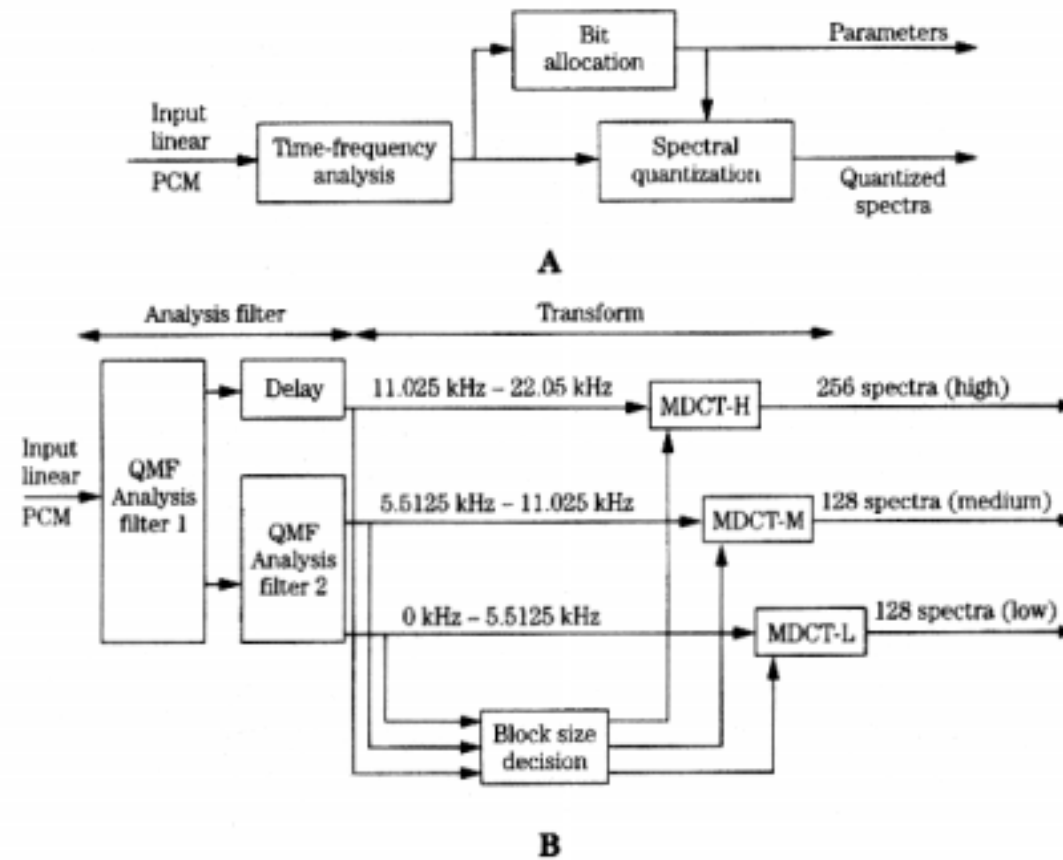
# Audio

## ■ ATRAC (Adaptive TRansform Acoustic Coding)

- ◆ reduziert Datenrate von 1,411 Mbps auf 292 kbps (5:1)
- ◆ 44.1 kHz Samplerate
- ◆ verwendet
  - ☞ unterschiedliche Frequenzbänder
  - ☞ psychoakustisches Modell, daß Dynamikempfinden des Ohres berücksichtigt (grobe Quantisierung wo hohe Pegel, feine Quantisierung bei geringen Pegeln)
- ◆ Funktionsweise
  - ☞ Eingangssignal in Zeitblöcke geteilt und frequenzanalysiert
  - ☞ Analyse erfolgt in 3 Frequenzbereichen und insgesamt 512 Spektralbändern (52 Block Floating Units deren Bandbreiten dem Gehör angepaßt sind)
  - ☞ Quantisierung erfolgt nach Signaldynamik und Maskierungseffekten
  - ☞ wie bei MP3 wird zwischen kurzen und langen Transformationsblöcken umgeschaltet (1,45/2,9ms und 11,6 ms)

# Audio

- ATRAC
  - ◆ Blockdiagramm





# Audio

## ■ Weitere Infos

◆ Ken. C. Pohlmann, Principles of Digital Audio (4th edition), Mc.Graw-Hill, 2000, ISBN 0-07-134819-0

### ◆ Inhalt:

- ☞ Audiogrundlagen
- ☞ Digitale Audio Aufzeichnung und -wiedergabe
- ☞ Psychoakustische Kodierung, Fehlerkorrektur
- ☞ Aufzeichnungsgeräte: Magnetband, DAT, CD, DVD, MD
- ☞ PC-Audio, Internet Audio
- ☞ Digitales Radio und Fernsehen
- ☞ Digitale Signalverarbeitung, AD/DA-Wandlung
  
- ☞ 736 Seiten, ca. öS 900,-

# Inhalt

## ■ Multimedia-Daten

### ◆ Audio ✓

### ◆ Images

- ☞ Codierung (Farbtiefen, Alpha-Kanal)
- ☞ verlustlose/verlustbehaftete Bildkompression
- ☞ wichtige Formate (GIF, JPEG, TIFF)

### ◆ Video

- ☞ analoges Video-Signal
- ☞ Standards (PAL, NTSC, HDTV, CCIR)
- ☞ digitales Video
- ☞ Standards (DVB)
- ☞ Codierung (YUV, 4:2:2, ...)
- ☞ Kompressionsverfahren (M-JPEG, MPEG, Quicktime)