



Audiocodierung

Gerd Brohasga M.Eng., Dipl.-Ing. (FH) Medientechnik

Audio ich höre (lat.)
Video ich sehe (lat.)



Fraunhofer Institut für Integrierte Schaltungen IIS

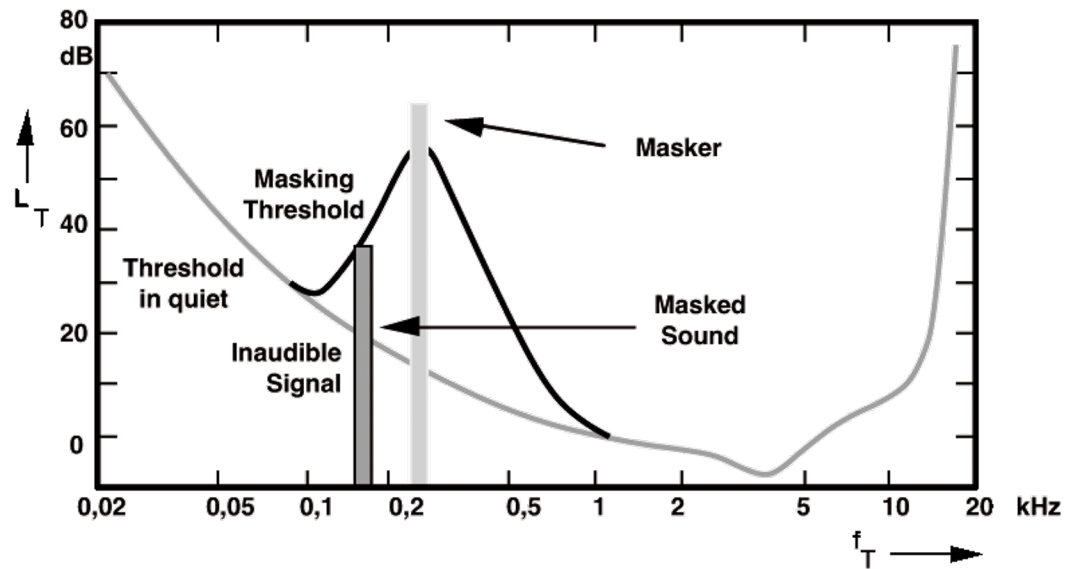
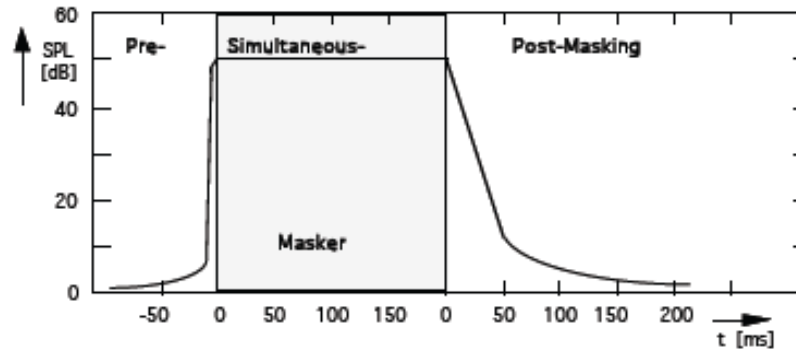


- seit 1981: Forschung im Bereich Audiocodierung in Erlangen (Universität Erlangen-Nürnberg, Prof. Seitzer)
- 1987: Erste Echtzeitcodierung von Stereosignalen am Fraunhofer IIS
- 1992: MPEG-Audio Layer-3 (= MP3) fertiggestellt
- 2020: MP3 ist immer noch das beliebteste Audioformat



seit 1994: Gemeinsame MPEG AAC
Entwicklungsarbeit des Fraunhofer IIS,
AT&T, Dolby und Sony

Warum Audiocodierung sowie Grundlagen



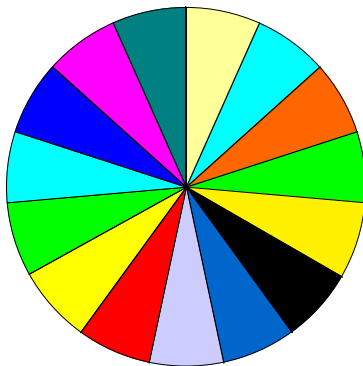
Warum Audiocodierung (»Audiokompression«)?

Das Ziel war Musik über zwei ISDN-Kanäle bei 128 kbit/s zu übertragen
fünf Minuten Musik (bei 44,1 kHz, 16 Bit)...

... ohne Kompression

42 MB Platzbedarf auf CD

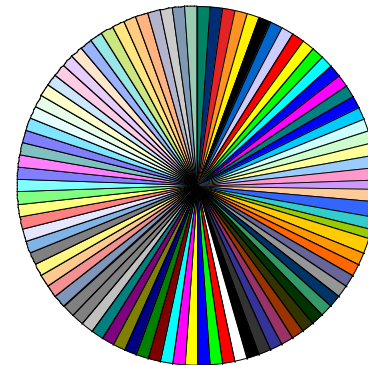
16 Musikstücke auf einer CD



... mit Kompression bei 128 kBit/s

3,8 MB Platzbedarf auf CD

184 Musikstücke auf CD



Wie funktioniert 's?

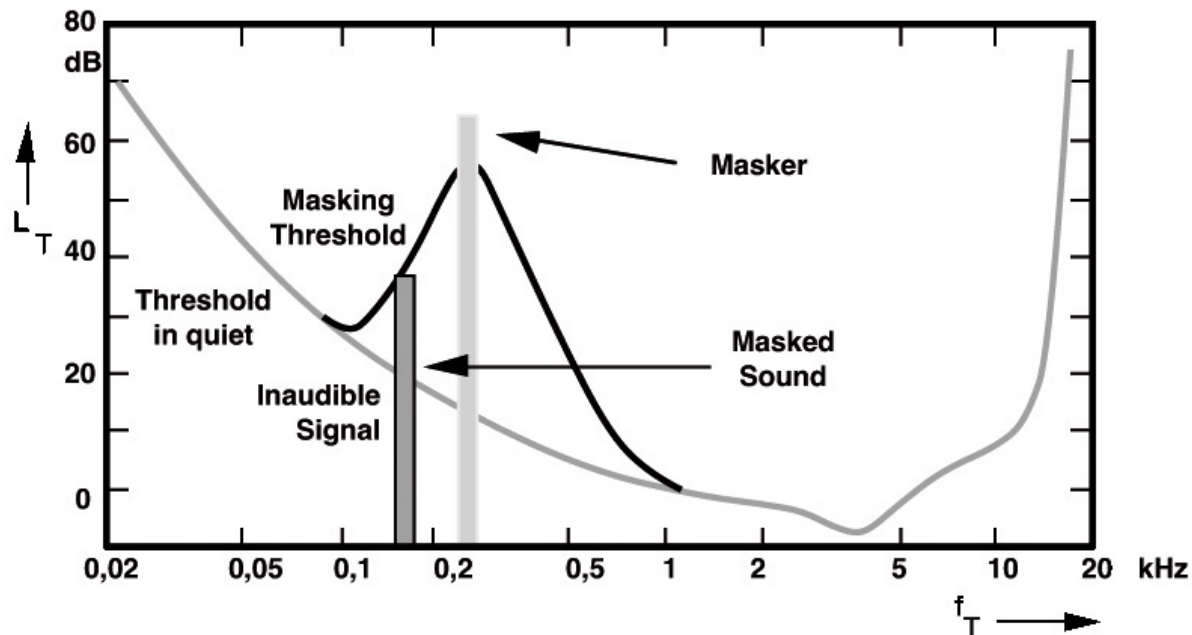
- »Was ich nicht weiß, macht mich nicht heiß«
 - der Mensch hört nicht alles gleich gut
- Audiocodierung nutzt Wahrnehmung des Menschen
 - **wichtige** Signalanteile muss man erhalten
 - bei **unwichtigen** Signalanteilen kann man Daten sparen
- hoher Aufwand (einige 10 Millionen Rechenschritte pro Sekunde)
- seit etwa 20 Jahren erlauben leistungsfähige »Chips« die Anwendung in Produkten aller Art

Redundanz und Irrelevanz

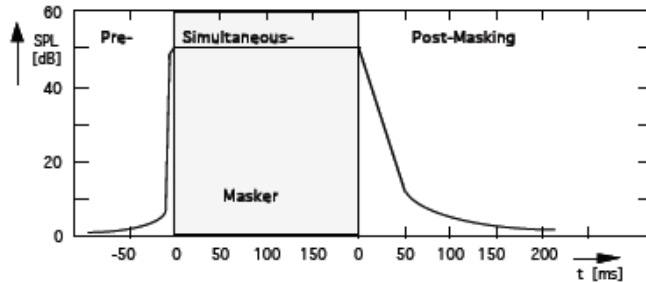
- **redundante** Signal-Anteile kann man ohne Informations-verlust verringern
 - »lossless audio coding«
 - mittlerer Reduktionsfaktor 2 bis 3 erreichbar
 - FLAC, Apple Lossless, HD-AAC
- **irrelevante** Signal-Anteile kann man nur mit Informations-verlust verringern; ein nicht-wahrnehmbarer Verlust bleibt für den Empfänger ohne Bedeutung
 - »lossy audio coding«
- Moderne Audiocodierverfahren (z.B. mp3 oder AAC) nutzen beide »Sparmaßnahmen« und erreichen Reduktionsfaktoren von 8 bis 32 (und deutlich mehr) bei sehr guter Tonqualität.
- Weitere Formate sind z.B. .wma (Windows), .rm (Real Media), .ogg (Ogg Vorbis), .mp4/.m4a (AAC)

Psychoakustik: Die Ruhehörschwelle

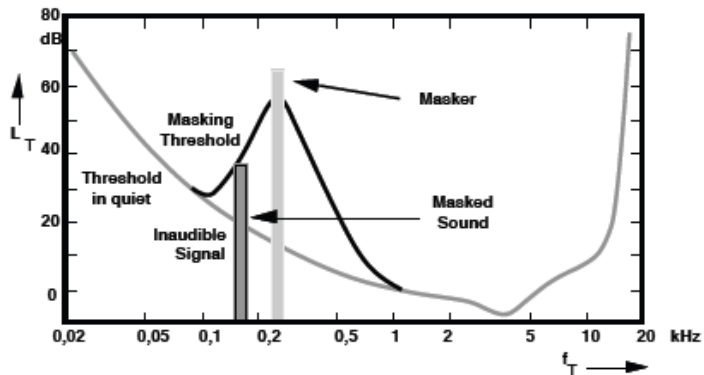
- Psychoakustische Modelle beschreiben die Wahrnehmungsfähigkeit des menschlichen Ohres
- Alles was unterhalb der Ruhehörschwelle ist unhörbar
- Teile des Audiosignals, die nicht mehr wahrnehmbar sind, können ineffizient gespeichert werden



Wahrnehmungseffekte



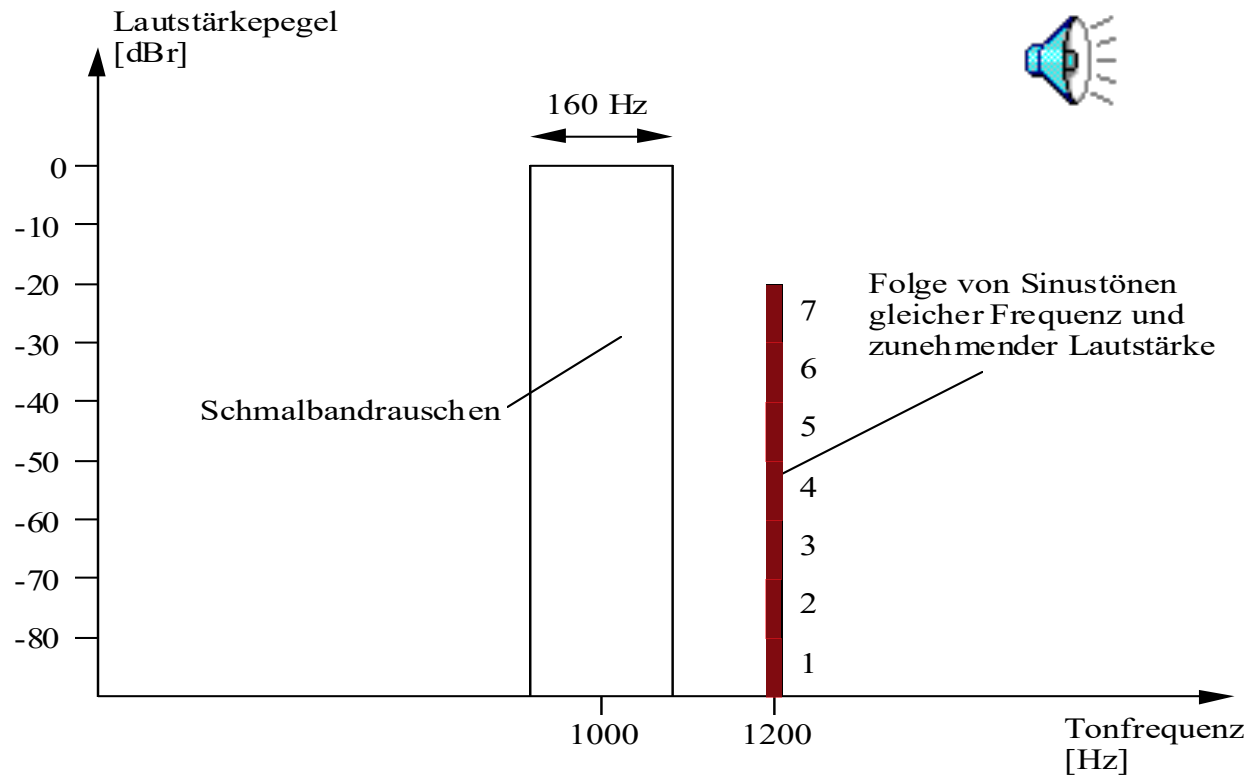
- Verdeckung im Zeitbereich
 - Nach einem Knall ist man hinterher »etwas taub«.



Example: Narrow band noise masker at 250 Hz, masked sound at 175 Hz

- Verdeckung im Frequenzbereich
 - Laute Töne »maskieren« die Wahrnehmung leiserer Töne ähnlicher Frequenz.

Beispiel: Ein lautes Geräusch verdeckt einen leisen Ton



Stereo Audiocodec – MP3



- ISO/IEC MPEG Standard (MPEG-1 Layer 3) seit 1992
- Populärstes Audioformat in der Unterhaltungselektronik
- Namensgebung .mp3 entstand 1995 bei einer internen Umfrage
- Erster MP3 Player 1998 MP-Man (64 MB Speicherkapazität)
- Über eine Milliarde tragbare MP3-Spieler
- MP3 CDs, Flash-Speichergeräte, HDD Player, PC/Macs, DVD Player, PDAs, Mobiltelefone, Autoradios, Spielekonsolen...
- WorldSpace Satellite Radio (Südamerika, Afrika)

AAC Steckbrief

- AAC ist eine Abkürzung:
 - Advanced Audio Coding (ISO/IEC MPEG Standard)
- wird verwendet
 - für Mono- und Stereo-Signale
 - für Surround-Sound (z.B. »5.1 Lautsprecher«)
- Reduktionsfaktor und erreichbare Tonqualität:
 - 12: »Studioqualität« (= 128 kbit/s)
 - 16..24: »gute Qualität«
 - 32..48: »gute Qualität« mit *HE-AAC*
- Signallaufzeit ca. 200 ms



AAC-Entwicklungspartner

Diese vier Firmen entwickelten seit 1994 den neuen MPEG-Standard, der 1997 / 2000 erfolgreich standardisiert wurde.

AAC: beste Tonqualität bei geringsten Datenraten!



- AAC-Software ist auf vielen PCs oder Macs (u.a. iTunes, Nero, Quicktime, Real, Winamp)
- AAC ist das Basisformat von Apples iPod
 - 50 Mrd verkaufte Lieder im *iTunes Store* (2003-2013)
- AAC kann man abspielen
 - in vielen portablen »mp3-Spielern«
 - in Mobiltelefonen oder PDAs aller Art
 - in Sonys PSP (Playstation Portable)

Apples iPod (seit 2001)

Die verwendeten Audio-Decoder nutzen die Software des **Fraunhofer IIS**.

AAC im Bereich Streaming und TV



- AAC / HE-AAC kann man hier empfangen:
 - im europäischen digitalen Radio (DAB+)
 - im mobilen digitalen Fernsehen (DMB, DVB-H)
 - im mobilen digitalen US-Fernsehen (MediaFLO)
 - beim Internet-Fernsehen (IPTV, ISMA)
- AAC / HE-AAC ist in verschiedenen Standards z.B. DVB enthalten:

HE-AAC: Hören Sie den Unterschied?

- Jedes ? ist
 - entweder O (Original, 1536 kbps stereo)
 - oder C (Codiert, 48 kbps stereo)
- Trompeten-Solo

–

O



–

?



–

?



–

?



HE-AAC: Hören Sie den Unterschied?

- O (Original, 1536 kbps stereo)
- C (Codiert, 48 kbps stereo)

- Das codierte Signal braucht also nur ca. 3 % der ursprünglichen Datenmenge!

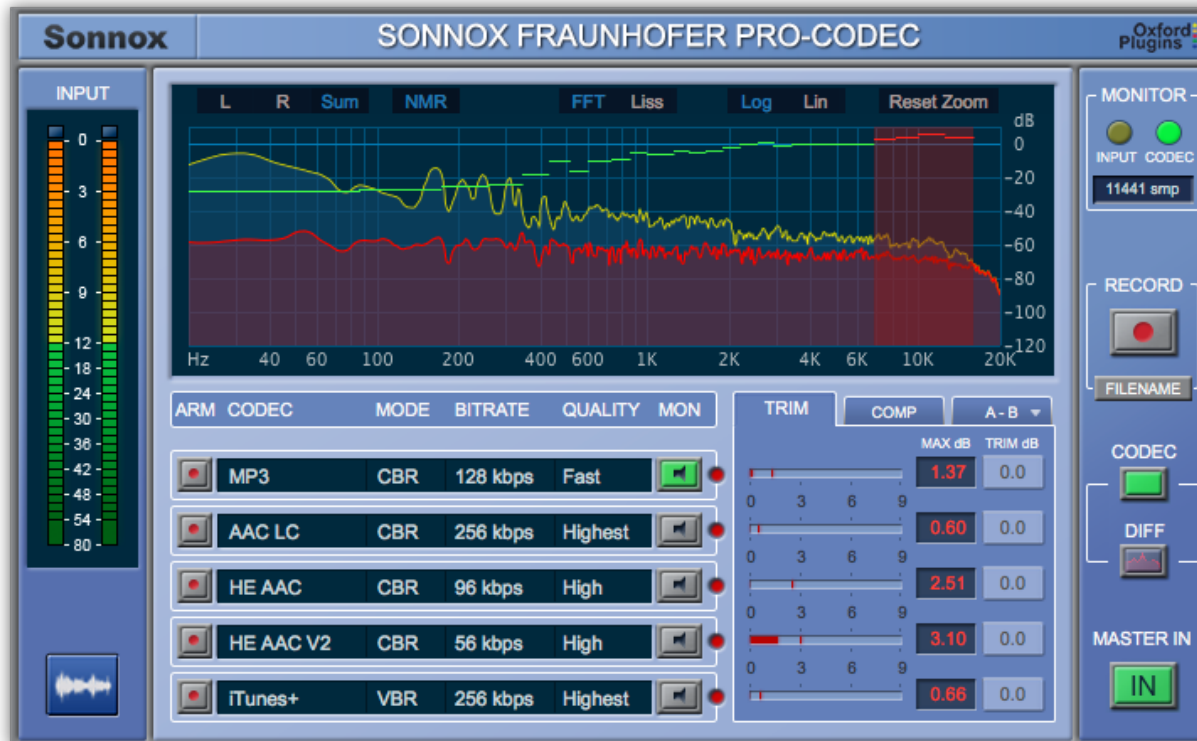
- O
- C
- C
- O

HE-AAC: Flavours

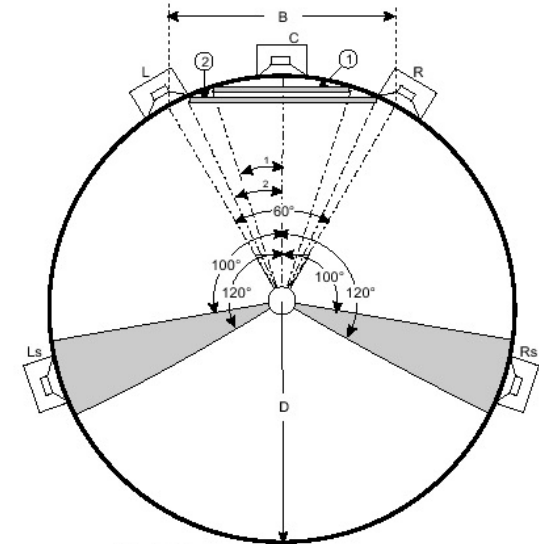
Codec	Features	Typical Applications	Typical Bit rate
AAC-LC (Low Complexity AAC)	High performance audio codec for excellent audio quality at low bit-rates	<ul style="list-style-type: none"> – Apple iPod – iTunes – ISDB television broadcasting (Japan) 	128 kbit/s (Stereo)
HE-AAC (High Efficiency AAC)	High performance audio codec for good quality at bit-rates of 32-48 kbit/s per channel	<ul style="list-style-type: none"> – XM Radio – Mobile music download – Digital Radio Mondiale 	64 kbit/s (Stereo)
HE-AAC v2	Highest performance audio codec for good quality at bit-rates of 16-24 kbit/s per channel	<ul style="list-style-type: none"> – 3GPP music download – Digital radio DAB+ – Internet radio streaming to mobile devices such as iPhone 	48 kbit/s (Stereo)
HD-AAC (High Definition AAC)	Lossless audio codec for better-than-CD-quality with 24 bit and up to 192 kHz	<ul style="list-style-type: none"> – Home networks – Music distribution / production 	Roughly half the bit-rate of the uncompressed file
AAC-LD (Low Delay AAC)	AAC encoding with 20 ms algorithmic delay	<ul style="list-style-type: none"> – Video conferencing – Telepresence – VoIP telephony – Broadcast gateway 	128 kbit/s (Stereo)
AAC-ELD (Enhanced Low Delay AAC)	Low delay full audio bandwidth codec at data-rates down to 24 kbit/s per channel and 15 ms delay	<ul style="list-style-type: none"> – Video conferencing – Telepresence – VoIP telephony – Broadcast gateway 	64 kbit/s (Stereo)
MPEG Surround	Surround Sound extension e.g. for AAC-LC and HE-AAC	<ul style="list-style-type: none"> – Digital broadcasting – Mobile applications with binaural surround sound – Music distribution 	64...192 kbit/s (5.1 channels)

Sonnox Fraunhofer Pro-Codec

- Plug-In für AAX, VST und AU
- Echtzeitvergleich von bis zu fünf MPEG Codecs
- Stapelverarbeitung für den Export



Surround-Audiocodierung

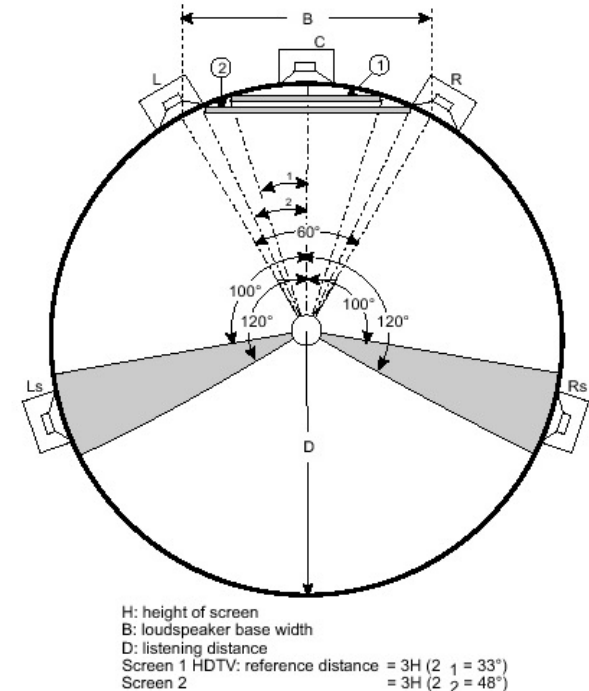


H: height of screen
 B: loudspeaker base width
 D: listening distance
 Screen 1 HDTV: reference distance = $3H$ ($2_1 = 33^\circ$)
 Screen 2 = $3H$ ($2_2 = 48^\circ$)



Surround Sound: Wiedergabe

- häufigste Lautsprecher-Anordnung:
 - »5.1«
 - Vorne: Links – Mitte – Rechts
 - Hinten: Links – Rechts
 - plus LFE (Low Frequency Effects oder Low Frequency Enhancement)
 - private Nutzung
 - am PC oder Mac
 - im Wohnzimmer
 - im Auto
- (SACDs, DVD Audio, DVD Video, BluRay Disc, DVB)



Normaufstellung nach ITU-775

Surround Sound: Produktionen

- Kinofilme
- Computerspiele
- Fernseh- und Radioproduktionen (DVB-S, DVB-C)
- entweder: Stereo- und 5.1-Tonspuren sind gleichzeitig vorhanden, z.B.
 - Stereo MPEG Layer-2, ca. 224 kbps
 - 5.1 Dolby Digital, ca. 512 kbps
 - DTS, 1,4 Mbit/s
- oder: es gibt nur eine Stereo-Tonspur mit Surround-Matrix
 - Dolby Surround PL II

Surround Sound: Digitale Formate

- »Dolby Digital« (AC-3)
 - gute Tonqualität bei 512 kbit/s (für 5.1)
- »Dolby Digital TrueHD«
- »Digital Theater Sound« (DTS)
 - gute Qualität bei 1400 kbit/s (für 5.1)
- »Digital Theater Sound HD Master« (DTS-Master HD)
- AAC
 - gute Qualität bei 320 kbit/s (für 5.1, diskret)
 - Anwendung im japanischen Rundfunk, UK
- Dolby Digital Plus (E-AC3)
 - 32 kbps–6 Mbps, scalable; typically 768 kbps–1.5 Mbps on HD optical discs; 256 kbps for broadcast and online

Dolby TrueHD

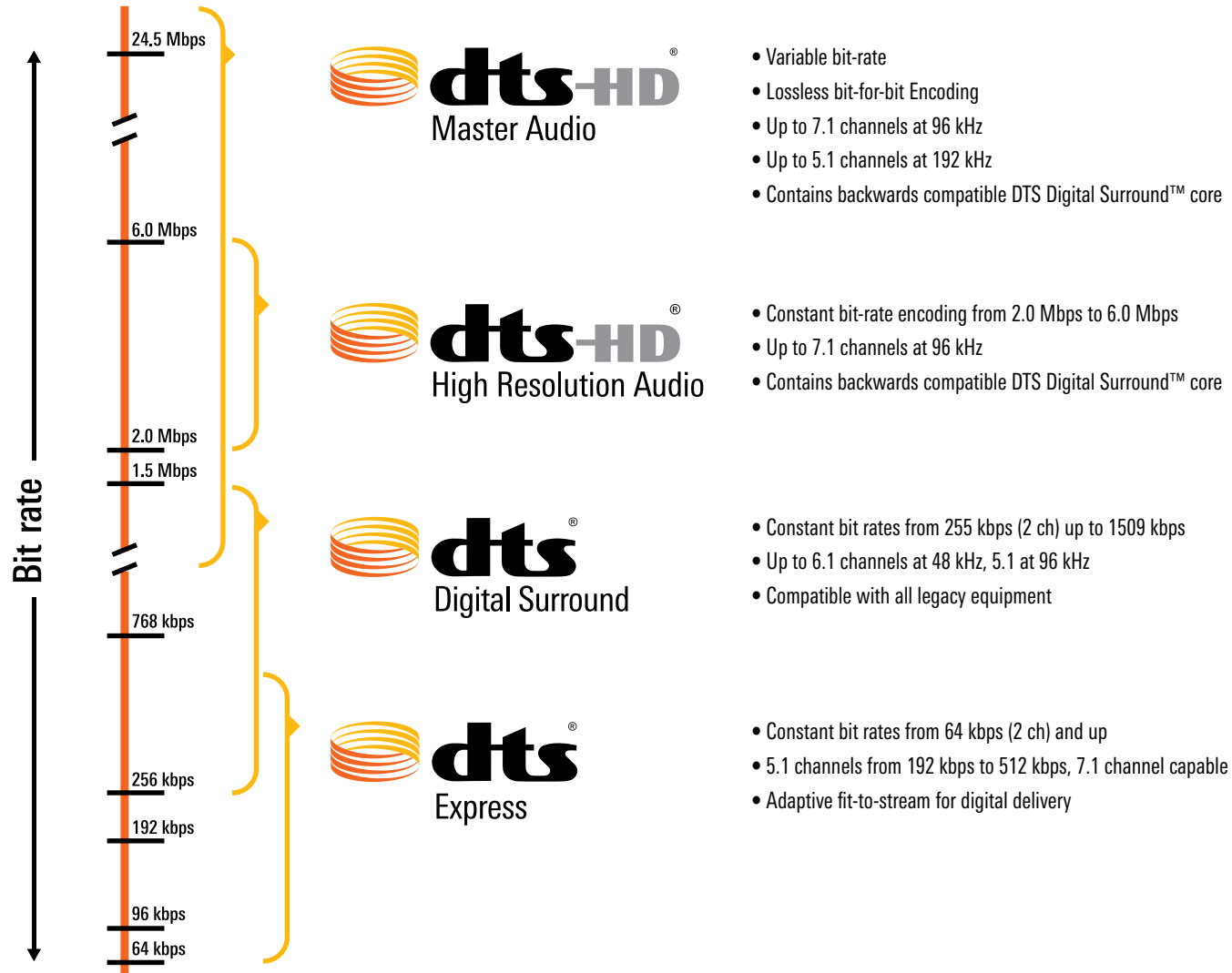


- **Dolby TrueHD Lossless Audio Specifications**
- **Format:** Lossless compression
- **Sampling frequency:** 44.1 kHz to 192 kHz
- **Word length:** Up to 24 bits
- **Data rate:** Variable, up to 18 Mbps
- **Maximum channel support:** 16 channels, as presently deployed
- **Blu-ray Disc channel support:**
 - Up to eight channels of 96 kHz/24-bit audio
 - Six channels (5.1): up to 192 kHz/24-bit
 - Two- to six-channel support: up to 192 kHz/24-bit maximum bit rate, up to 18 Mbps on Blu-ray Disc

Dolby Vergleich

Feature	Dolby Digital Plus	Dolby Digital	Competing Technologies	Dolby TrueHD
7.1 Channels	•		•	•
5.1 Channels	•	•	•	•
Streaming/download support	•	•	•	Limited
Multiple BonusView support for Blu-ray	•		•	
Automatic stereo downmix	•	•	•	•
Support for mobile networks	•			Limited
Scalable for bandwidth-constrained applications	•			
Compression	Lossy	Lossy	Lossy	Lossless
Data rates supported: Mono–multichannel	32 kbps–6 Mbps	32–640 kbps		Up to 18 Mbps

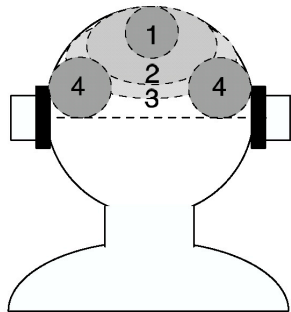
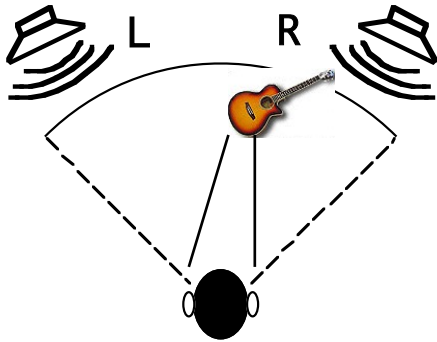
DTS Audio Codec Overview



Surround Sound: Binaural Cue Coding

- neue digitale Darstellung eines Multikanal-Tonsignals als:
 - Mono- oder Stereosignal
 - plus kleines Zusatzsignal für den Raumklang
- 100% rückwärtskompatibel:
 - einfache Empfänger ignorieren das Zusatzsignal und liefern hervorragenden Mono- oder Stereo-Ton
 - neue Empfänger nutzen das Zusatzsignal und erzeugen den Raumklang
- Schlüsseltechnik für datensparenden Raumklang!

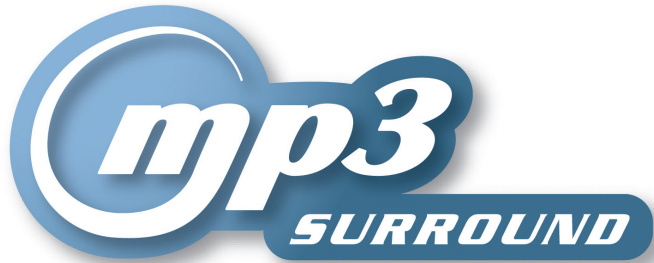
Binaural Cue Coding: Wahrnehmungseffekte



[Faller 2004]

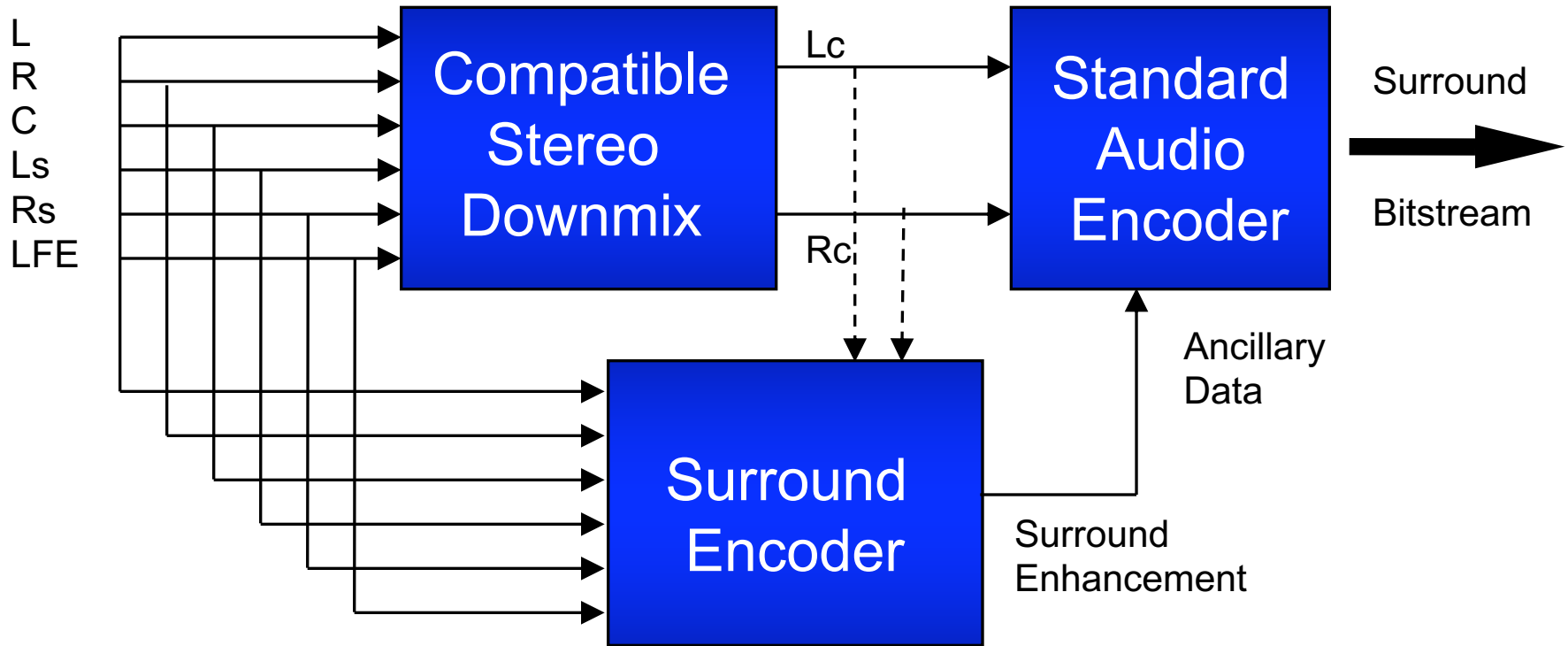
- räumliche Schallwahrnehmung ist ein vielschichtiger Vorgang
- wesentliche Anhaltspunkte (»*spatial cues*«) sind:
 - Pegelunterschiede der beiden Ohrsignale
 - zeitliche Unterschiede der beiden Ohrsignale
 - Kohärenz-Effekte bei der Auswertung:
 - je weniger ähnlich die beiden Ohrsignale sind, um so größer wird die empfundene räumliche Breite des Schallereignisses
 - im Bild dargestellt von 1 → 4
- sehr geringe Datenmenge reicht für guten Raumklang!

Binaural Cue Coding: MP3 Surround

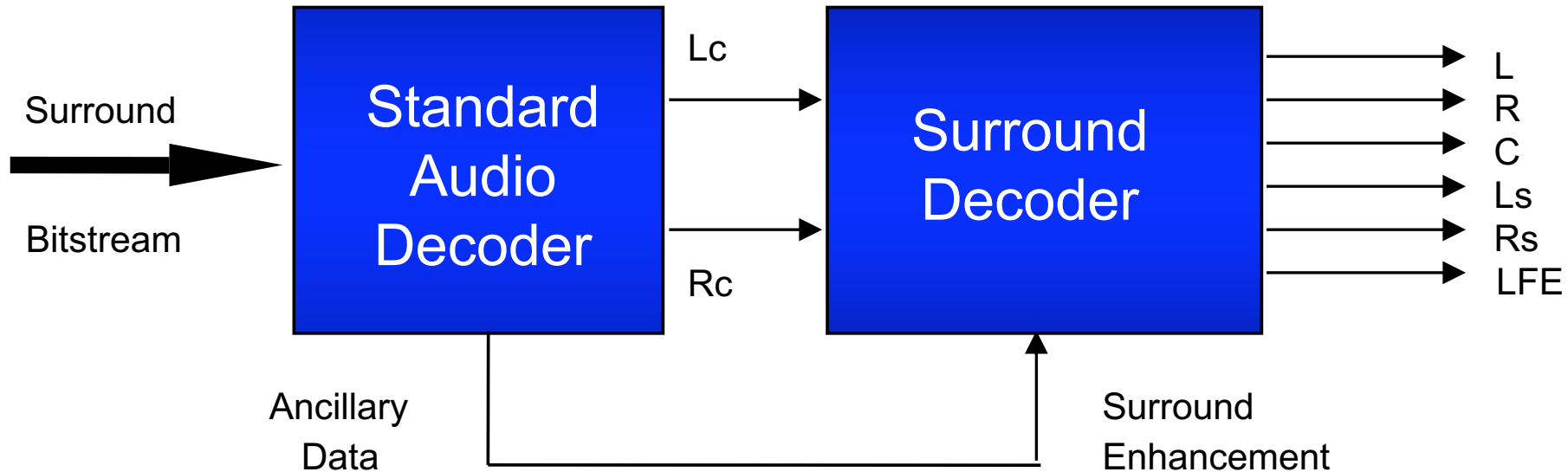


- Kompression von 5.1 Mehrkanalton-Material (Kompressionsrate von 1 zu 23)
- Bitrate vergleichbar mit Stereo MP3 (z.B. 256 kBit/s)
- Vollständig rückwärtskompatibel und damit auf allen Stereo MP3-Playern abspielbar

Binaural Cue Coding: Technologie



Binaural Cue Coding: Technologie



Immersive Audioformate, 3D Audio



3D Audio: Wiedergabe

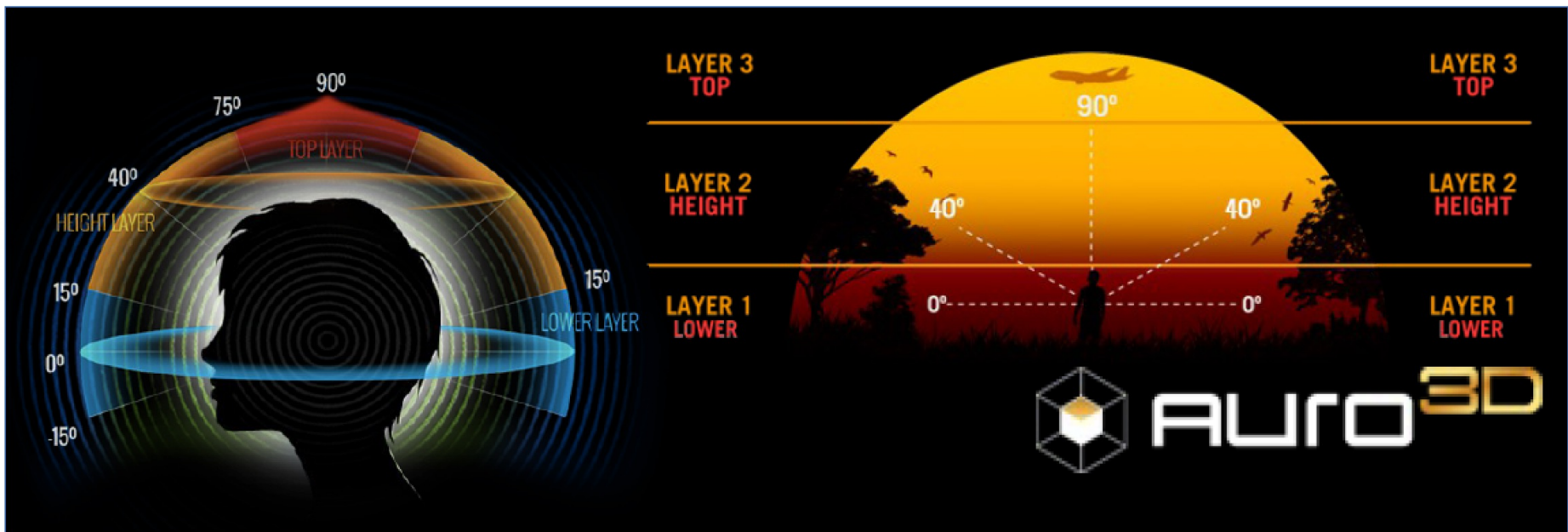
Verfahren zum räumlichen Abbilden von Schallquellen.

3D Audio ist ein akustisches Verfahren, mit dem die realistische räumliche Abbildung von Schallquellen erfolgt.

Der Zuhörer erfährt dabei eine akustische Atmosphäre, die der des Aufnahmeortes nachempfunden ist.

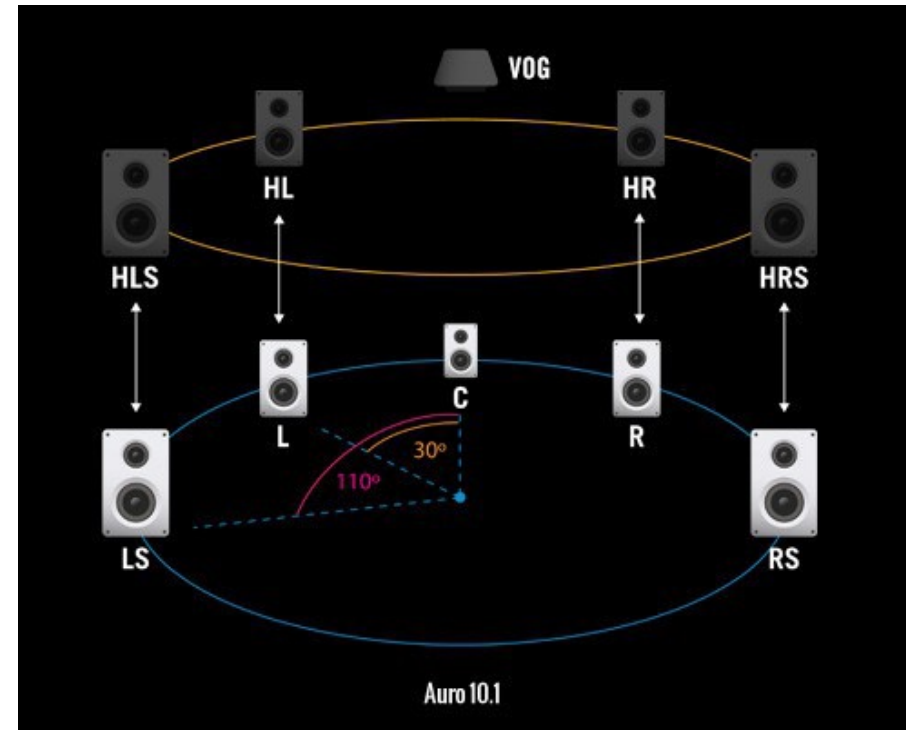
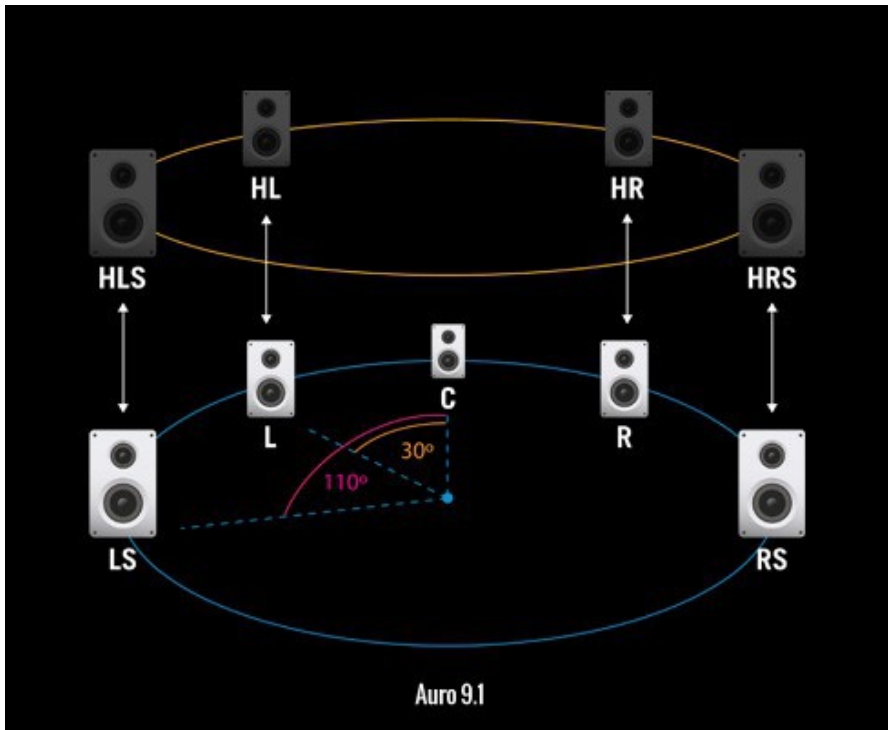
Erreicht wird dies über den Einsatz von Höhenlautsprechern.

Auro 3D



Ziel der Auro-3D-Technologie ist eine realitätsnahe dreidimensionale Audio-Reproduktion durch eine zusätzliche Höhen-Information, welche als dritte Dimension zur ersten Dimension der Breite (Stereo) und zur zweiten Dimension der Tiefe (herkömmliches Surround) ergänzend hinzutritt.

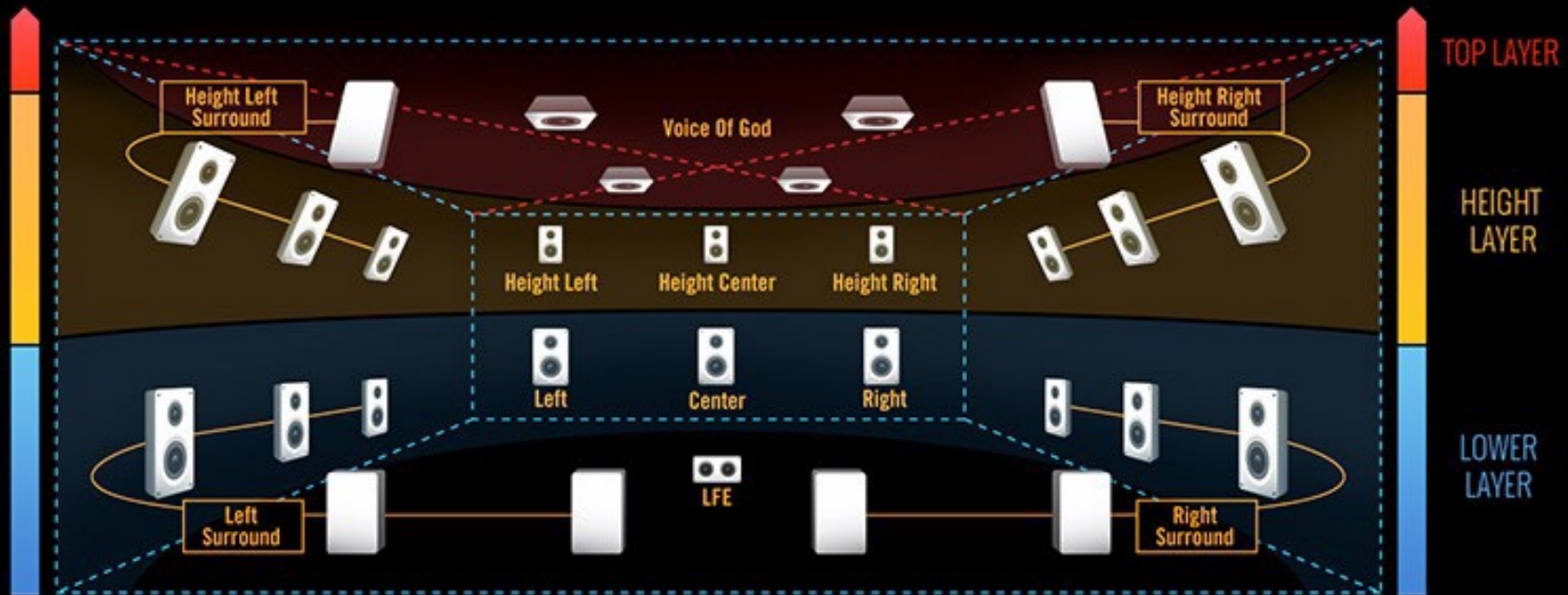
Auro 3D



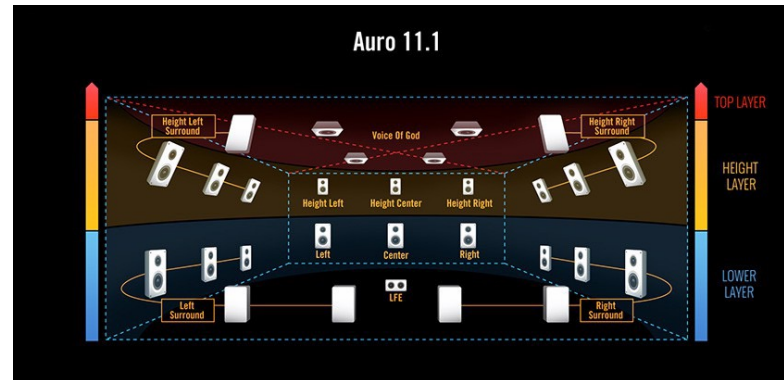
Auro 9.1 verzichtet sowohl auf den Höhenkanal über dem Center-Lautsprecher als auch auf den Deckenlautsprecher für die „Voice of God“. Möchte man auf die „Stimme Gottes“ direkt von der Decke nicht verzichten und kann dort einen Lautsprecher unterbringen, benötigt jedoch für den kleinen Bildschirm zu Hause keinen Center- Höhenkanal, so ist Auro 10.1 die richtige Wahl.

Auro 3D

Auro 11.1



Auro 3D



Auro 11.1 basiert auf einem herkömmlichen 5.1-System, dessen Kanäle (links, rechts, center, surround-links und surround-rechts) jeweils um einen darüberliegenden Höhenkanal ergänzt werden, sowie zusätzlich durch einen Kanal für Lautsprecher direkt von oben, der sogenannten „Voice of God“.

Je nach Bedarf und Größe des Raums kann die Anzahl der Lautsprecher pro Kanal erhöht werden. Die einem Kanal zugeordneten Lautsprecher erhalten jedoch das identische Signal.

Dolby Atmos



Dolby Atmos ist zugleich ein objektorientiertes Verfahren kombiniert mit einem kanalorientiertem Verfahren. Kanalbasierte sogenannte „Beds“ bis 9.1. eignen sich für die Wiedergabe von Atmos.

Dynamische Objekte werden, von Metadaten gesteuert, in Echtzeit auf verschiedene Lautsprecheranordnungen mit bis zu 64 Lautsprechern gerendert und ermöglichen spektakuläre Panning-Effekte. Deckenlautsprecher sind ausdrücklich vorgesehen.

Dolby Atmos



5.1.2: Regular 5.1 setup with 2 ceiling or 2 Dolby Atmos-enabled speakers



5.1.4: Regular 5.1 setup with 4 ceiling or 4 Dolby Atmos-enabled speakers



7.1.4: Regular 7.1 setup with 4 ceiling or 4 Dolby Atmos-enabled speakers

POSSIBLE DOLBY ATMOS SPEAKER SETUPS



7.1.2: Regular 7.1 setup with 2 ceiling or 2 Dolby Atmos-enabled speakers



9.1.2: Regular 9.1 setup with 2 ceiling or 2 Dolby Atmos-enabled speakers

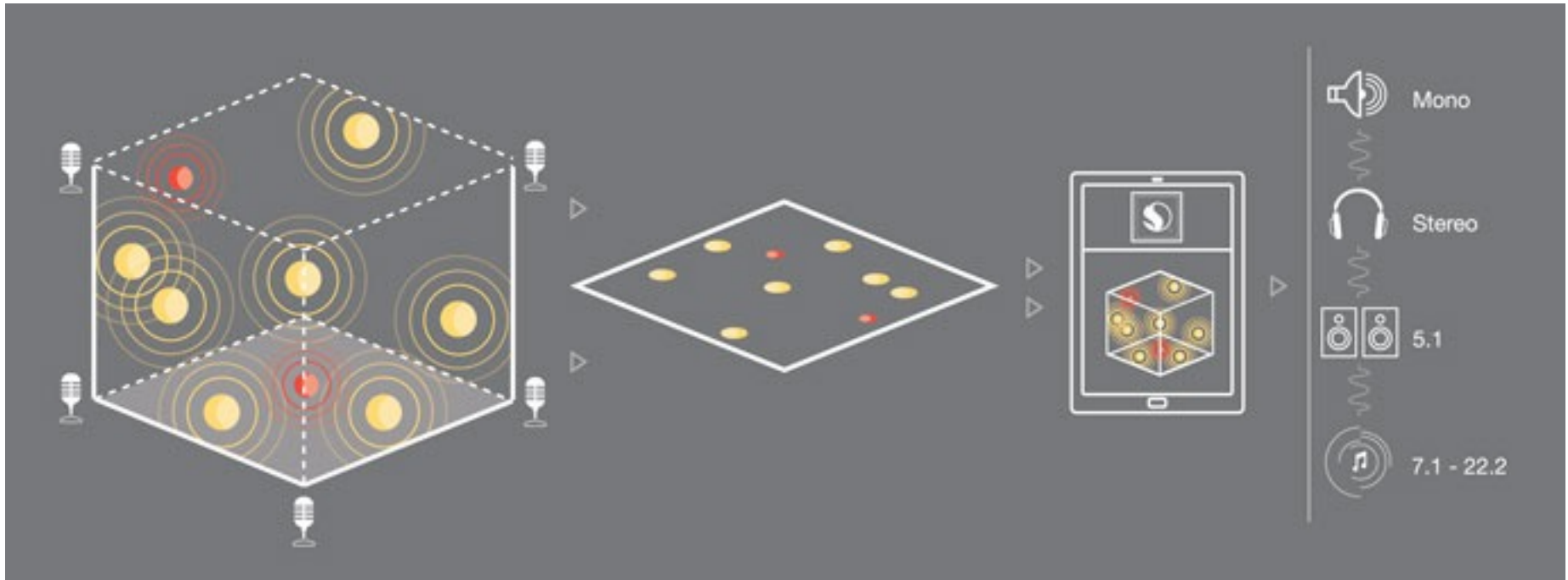
DTS X



Mit DTS X verabschiedet man sich bei DTS auch vom kanalbasierten Verfahren und geht über zur **objektbasierten Berechnung**.

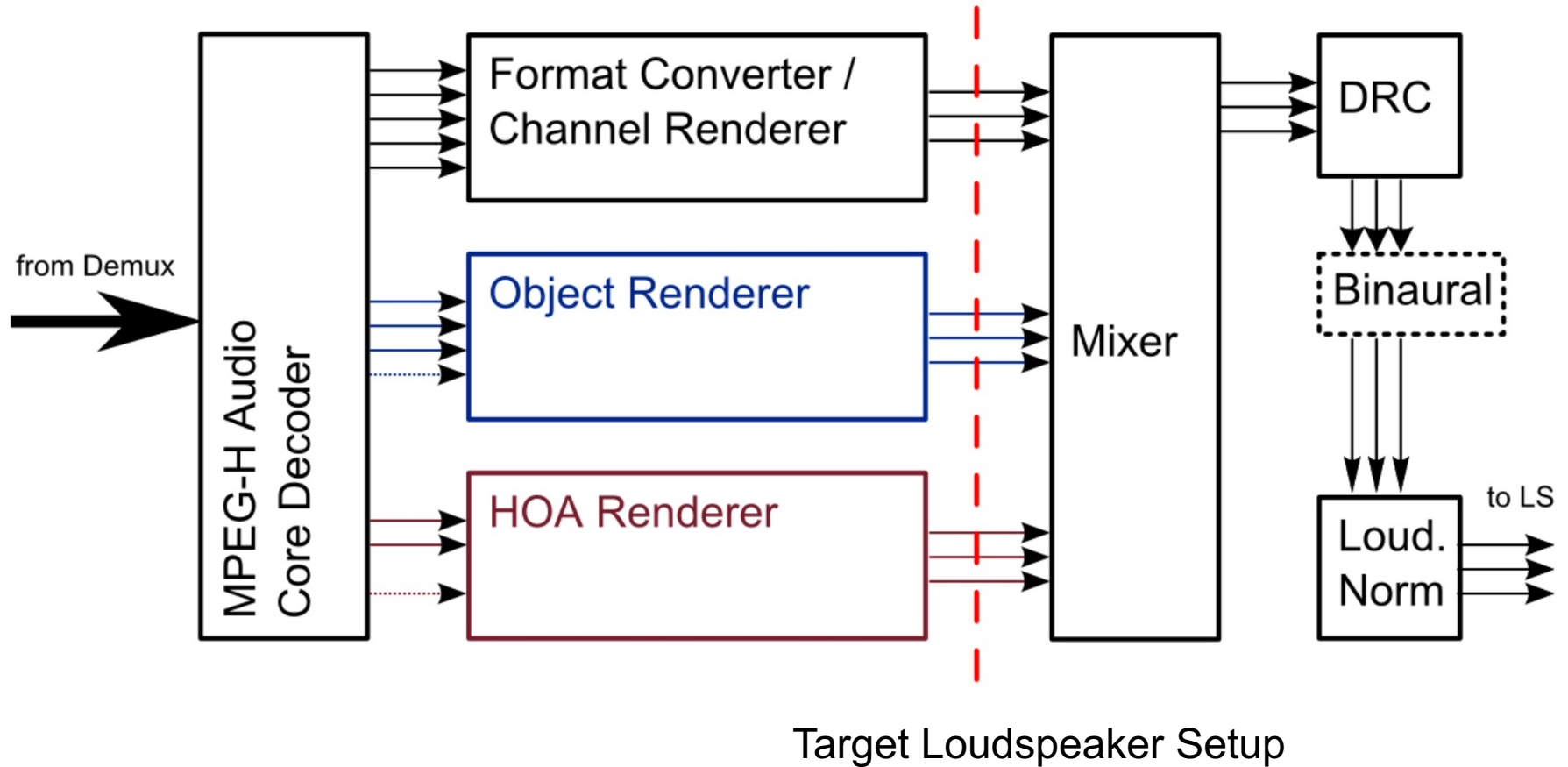
Mit der Verwendung von dynamischen Audio Objekten ähnelt DTS X auf dem ersten Blick sehr Dolby Atmos

MPEG H



Ein offener MPEG Audio-Standard. Um MPEG-H Audio voranzutreiben, haben sich die drei Institute Fraunhofer IIS, Qualcomm Technologies und Technicolor zur MPEG-H Audio Alliance zusammengeschlossen..

MPEG H



Vielen Dank für Ihre Aufmerksamkeit

Fragen?



Die Weiterverbreitung des Skripts außerhalb des Vorlesungsbetriebs ist nicht erlaubt.