Improved MPEG Low-Delay Audio Coding on DaVinci and TI C64 series DSPs

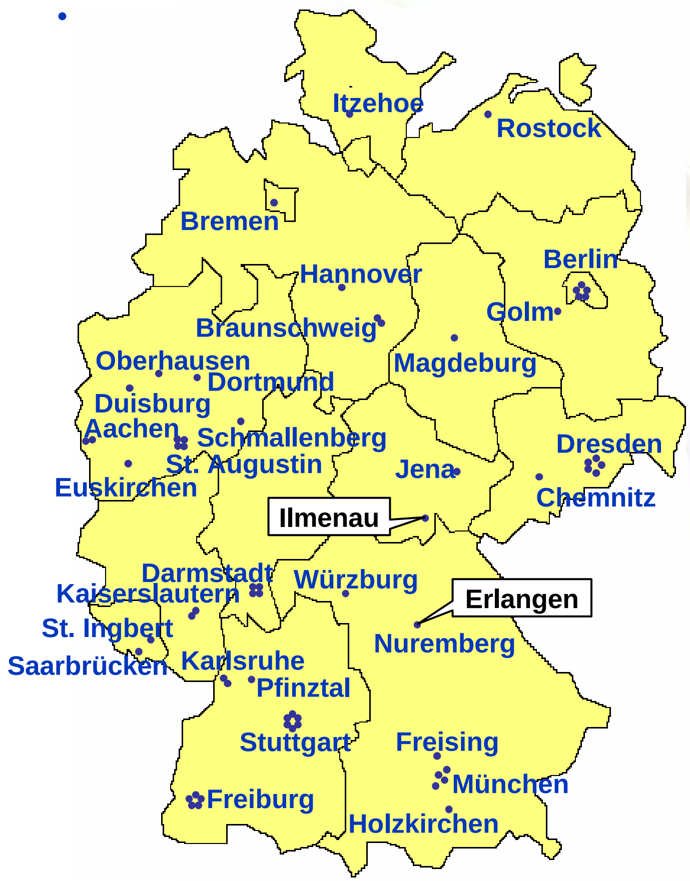
Negjmedin Fazlija
Fraunhofer IIS
faz@iis.fraunhofer.de

IIS
Fraunhofer Institut Integrierte Schaltungen

Agenda

- The Fraunhofer Institute for Integrated Circuits
- What Is Low Delay Audio Coding?
- AAC-Low Delay
- Ultra Low Delay Audio Coding
- MPEG Surround and Spatial Audio Object Coding
- Implementing AAC-Low Delay on TI DSPs

The “Fraunhofer Gesellschaft” - FhG



58 Fraunhofer Institutes in 40 locations in Germany

Largest private research organization in Europe

Non-profit organization, founded in 1949

Offices in Europe, USA and Asia

Permanent staff 12,500 primarily scientists and engineers

Parent Organization of:

- Fraunhofer IIS in Erlangen
- Fraunhofer IDMT in Ilmenau

Minds in Motion

Fraunhofer IIS
Institute for Integrated Circuits

- The Audio & Multimedia Realtime Systems departments of Fraunhofer IIS have joined with Fraunhofer IDMT into an Audio & Multimedia (AMM) departments cluster, employing more than 120 engineers.

- The headquarters of the Fraunhofer Institute for Integrated Circuits IIS are located in Erlangen, Germany. The institute's activities were extended by the establishment of the Fraunhofer Institute for Digital Media Technology (IDMT) in Ilmenau.

- Audio & Multimedia Cluster
  - MPEG Layer-3 (MP3)
  - MPEG-4 Advanced Audio Coding (AAC, HE-AAC, HE-AAC v2)
  - Spatial Audio Coding
  - Low Delay Audio Coding
  - Lossless Audio Coding
  - MPEG-4/AVC Video Coding
  - AV Streaming Technologies
  - Digital Broadcasting
  - Digital Rights Management
  - Metadata, Music Recognition

Fraunhofer Milestones

- **1979** Audio compression developed at Erlangen-Nuremberg University by Seitzer & Brandenburg
- **1987** First real-time stereo coding (LT-ATC) at Fraunhofer IIS in alliance with Erlangen-Nuremberg University
- **1988** FhG IIS begins contributing to MPEG
- **1989** MP3 precursor OCF published
- **1991** MP3 finalized
- **1994** Collaboration with Dolby, AT&T, Sony on AAC begins
- **1995** FhG begins use of ".mp3" file extension
- **1997** Microsoft licenses MP3
- AAC becomes part of MPEG-2
- Liquid Audio adopts AAC for internet music
- FhG provides firmware for first flash-based MP3 player (Saehan)
- **1999** WorldSpace satellite radio begins
- **2000** AAC-LD becomes part of MPEG-4
- **2001** MP3 Pro software released
- XM Satellite Radio begins
- 3GPP adopts AAC for mobile audio services
- **2003** Digital Radio Mondiale begins
- **2004** MP3 Surround introduced
- Windows Media Player 10 released with FhG MP3 encoder
- 3GPP adopts HE-AAC v2
- **2005** Ensonido virtual headphone technology introduced
- MP3 SX "surround from stereo" introduced
- **2006** MPEG Surround standardization finished

Agenda

- The Fraunhofer Institute for Integrated Circuits
- **What Is Low Delay Audio Coding?**
- AAC-Low Delay
- Ultra Low Delay Audio Coding
- MPEG Surround and Spatial Audio Object Coding
- Implementing AAC-Low Delay on TI DSPs

Music Codecs usually have long latency

- "General Audio Codecs" or "Music Codecs" are not typically used in interactive situations, and delay is unimportant

Codec	Typical Delay (HW or DSP)	Typical Application
MP3	140 ms	Music Player
AAC	210 ms	Music Player, Broadcasting
HE-AAC	360 ms	Mobile Music, Satellite Radio

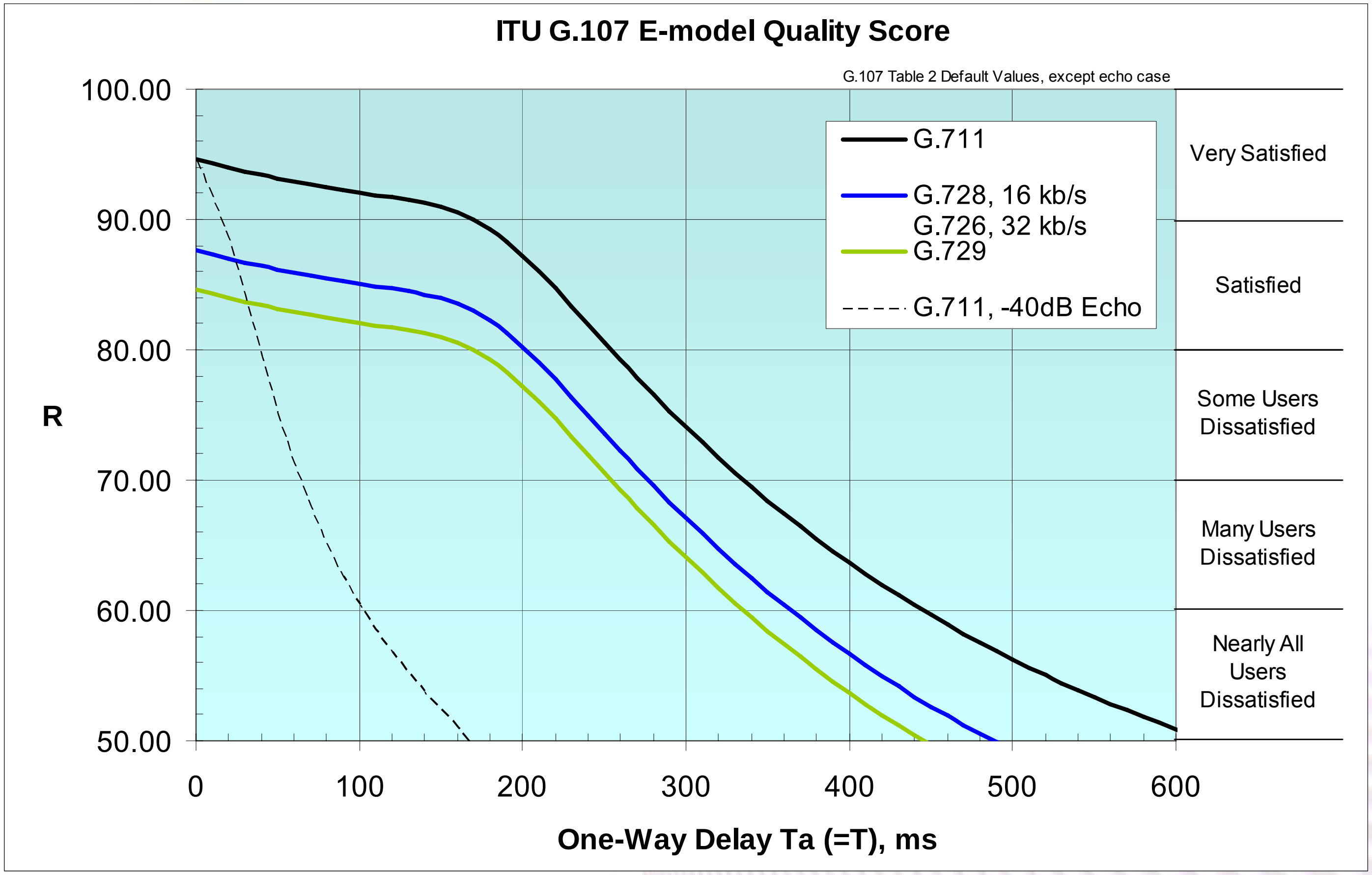
[3]

Why do we need Low Delay coding?

- For two-way interactive communication
  - Long delay feels un-natural
  - Audio must be synchronized with video
  - Example: Video Conferencing
- Where speakers can hear the coded signal
  - Speakers find speech difficult when they hear themselves with > 25 ms delay
  - Example: Phone call without an echo canceller
- Where acoustic delay is important
  - 1 ms ~= 1 foot of sound propagation
  - Example: Wireless speakers and microphones

Standards for Phone Call Delay

ITU-T G.107, G.114 specifies delay impairment for voice circuit:



[7]

Synchronizing Audio and Video

- H-series video codecs for interactive (chat, video conferencing) use have typically had 200-500 ms of delay
- Audio has been delayed to synchronize with the video.
- In these cases, audio codec delay is not so important
- But – Recent H.264 codecs have delays of ~80 ms

- Audio Codec is only one part of system

  ⟶ limit audio codec delay to 40-80ms

Future trends

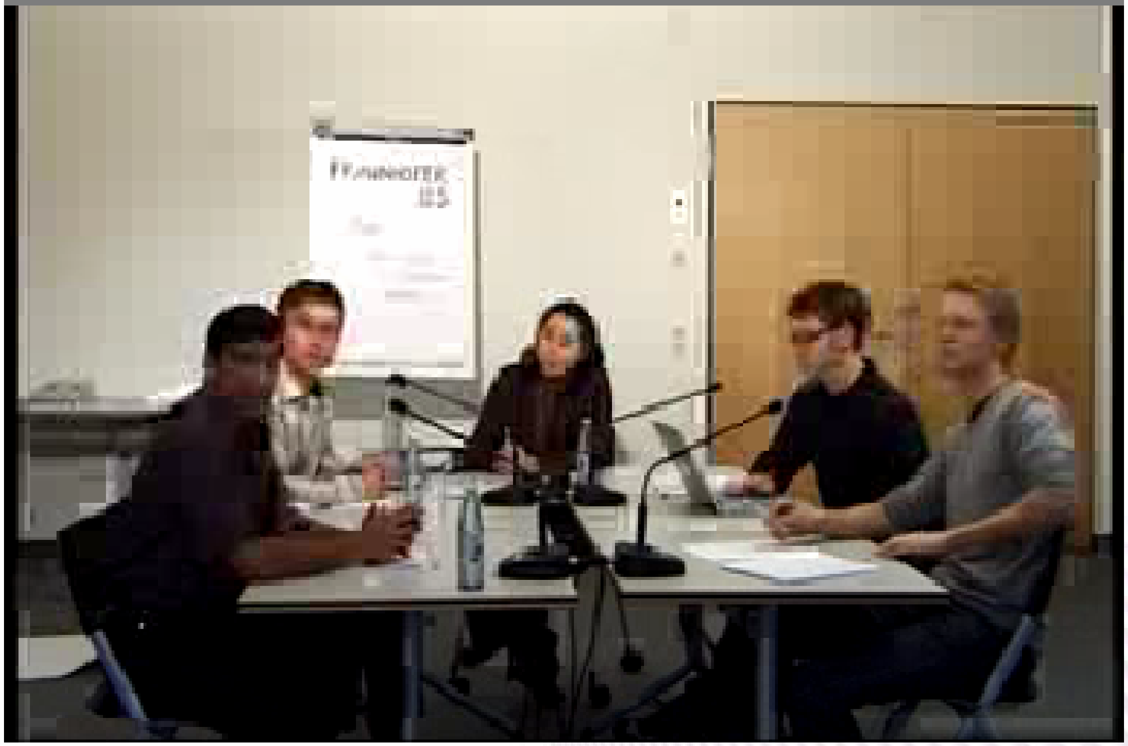
- Consumers want better quality from audio and video conferencing
- Transmission bandwidth is becoming available to give it to them
- For these systems, we need a higher quality audio codec:
  - More audio bandwidth
  - Less noise and coding artifacts
  - Still low delay

Goals of Next-Generation Conferencing

- The vision of these systems is to move from just delivering intelligible signals to high-fidelity, entertainment-grade performance
- Intelligibility
  - Natural sound quality
  - No annoying coding artifacts
  - Full audio bandwidth
  - Robust with tandem coding (cascaded codecs, MCUs)
- Speaker Separation
  - Be able to hear several people talking at once
- Ambience
  - Sounds come from the room, not a speaker

Example: Previous-Generation Video Conference

H.261 (Apple QuickTime) 112 Kb +
G.728 (Intel IPP) 16Kb



Play in QT

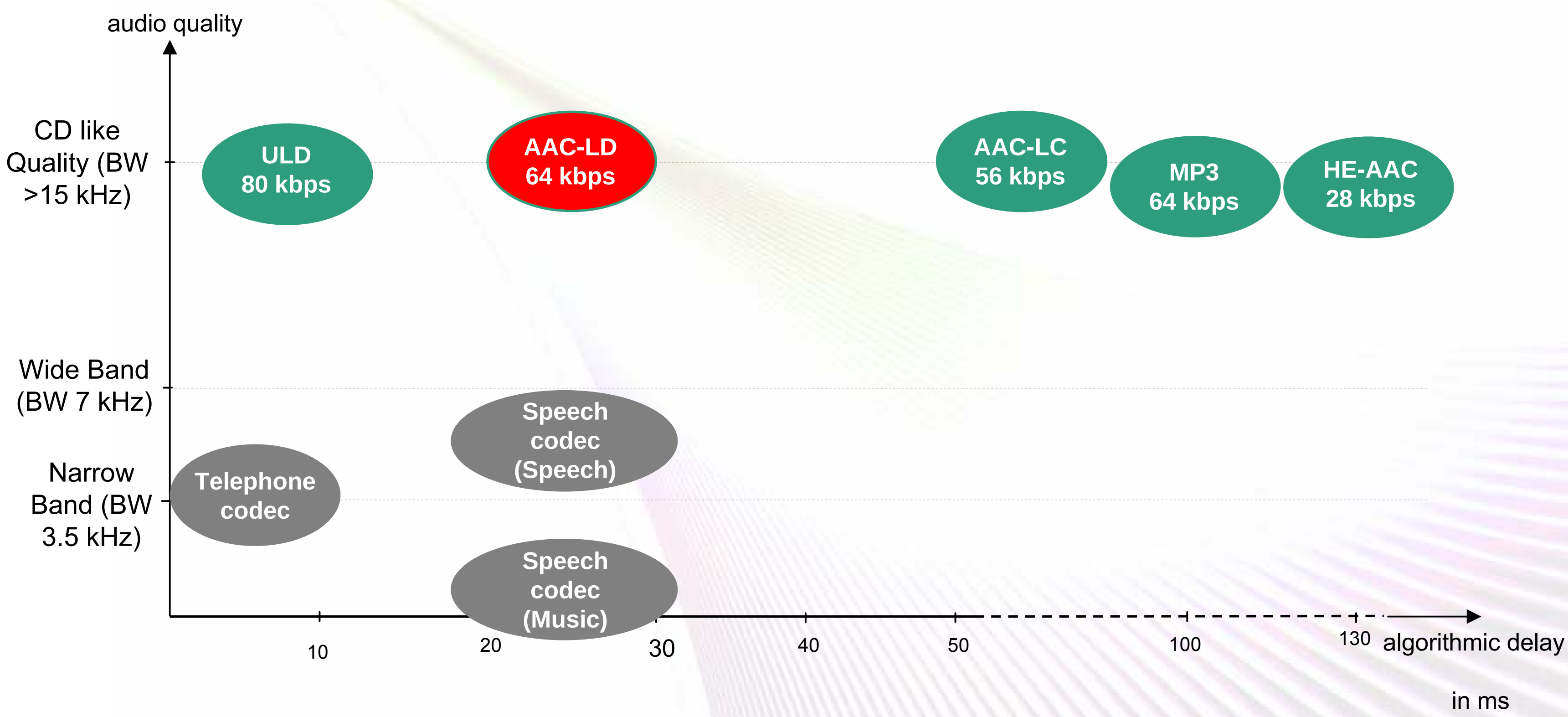
Minds in Motion

Example: Next-Generation Video Conference

AVC (Fraunhofer) 1000 Kb +
AAC-LD (Fraunhofer) 64Kb * 5

Play in QT

Codec Overview



Fraunhofer's Low Delay Codecs

"Realtime communication"

- MPEG-4 AAC Low Delay

  Bidirectional communication, VoIP

  Delay: 20 ms algorithmic,
  31 ms implementation

"Wireless Audio"

- ULD: Ultra Low Delay Audio Codec

  Wireless audio: microphones, loudspeakers, hearing-aids, VoIP, music jam sessions

  Delay: 6 ms algorithmic,
  10 ms implementation

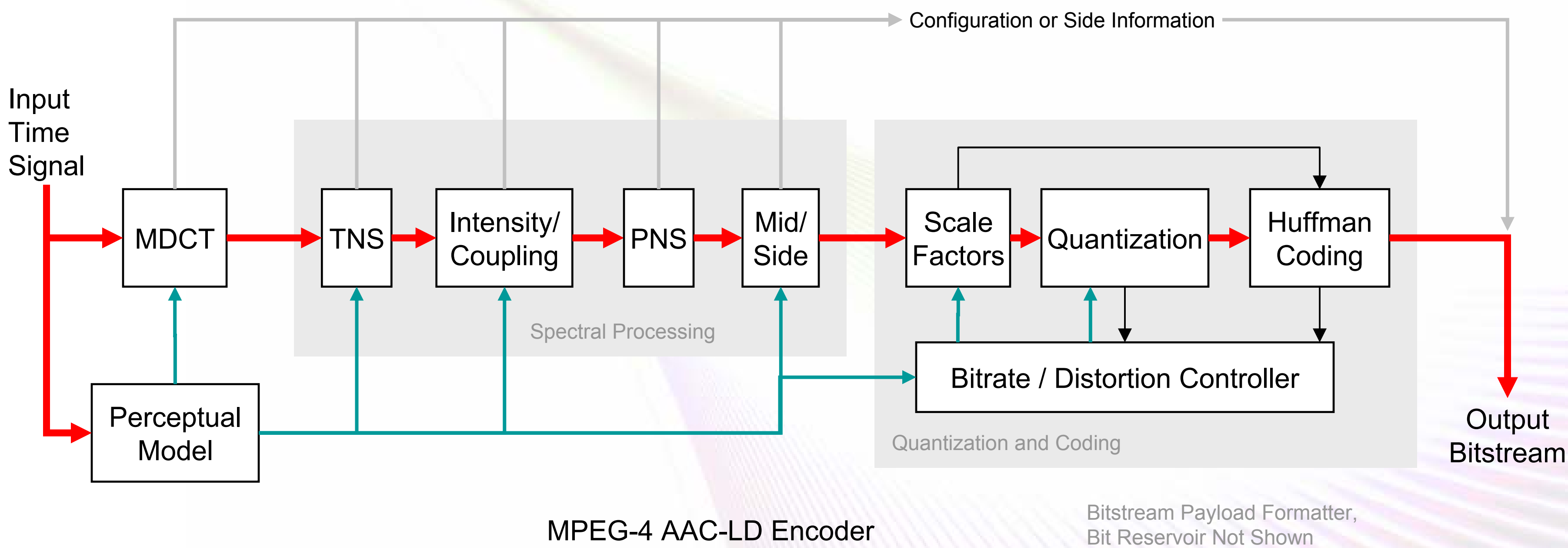
Agenda

- The Fraunhofer Institute for Integrated Circuits
- What Is Low Delay Audio Coding?
- **AAC-Low Delay**
- Ultra Low Delay Audio Coding
- MPEG Surround and Spatial Audio Object Coding
- Implementing AAC-Low Delay on TI DSPs

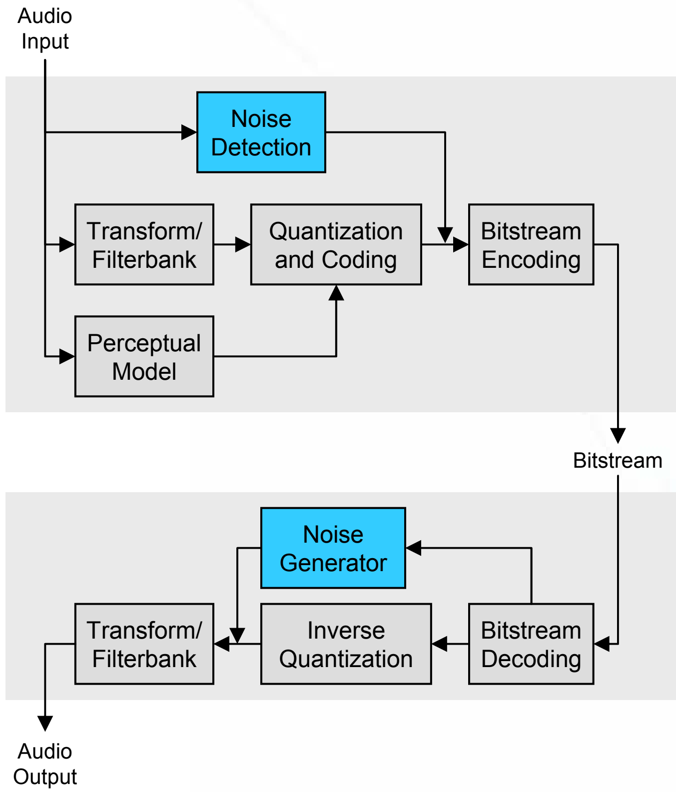
The AAC Codec Family

- Part of ISO MPEG-2 and MPEG-4 Standards:
  - MPEG-2 AAC: (ISDB)
  - AAC-LC: Music Coding (iPod)
  - HE-AAC: Low-bitrate music (XM Radio)
  - HE-AAC v2: Lower-bitrate music (3GPP)
  - AAC-LD: Communications applications

AAC-Low Delay Codec Structure



Coding Tools for AAC-LD



AAC's PNS tool detects noise-like frequency bands in the encoder and sends only a noise power parameter instead of frequency coefficients. A noise generator recreates the noise in the decoder.

- Temporal Noise Shaping (TNS)
  – Avoids pre-echoes for transient signals
- Perceptual Noise Substitution (PNS)
  – Parametric representation of noise-like bands is sent instead of frequency coefficients
- Mid-Side Stereo (M/S)
  – Increased coding gain for correlated channels
  – Avoids stereo unmasking

AAC-LD Delay Optimizations

- Low number of subbands
  → Reduced filter bank delay, only 2 x 480 samples
- No block switching (compensated by TNS)
  → No "look ahead" delay
- Minimal bit reservoir
  – < 100 bits/channel
- Result:
  – 20 ms algorithmic delay (48 kHz)
  – Real-time implementation ~31 ms delay
  – Audio quality comparable to MP3 at same bitrate
  – Up to fs/2 audio bandwidth
    - tuned for best performance at 16 kHz for 64 kb/s/channel
  – Large range of usable bitrates 32 - 80 kb/s/channel

VC/TC systems use or announce AAC-LD/LC

- **Tandberg** MXP
- **Sony** PCS-TL50P
- **Vcon** HD4000/HD5000
- **Lifesize**
- **Telos** Zephyr Xstream
- **Musicam** Netstar
- **Mayah** CENTAURI
- **Source Elements**
- **Codian** MCU 4200
- **Comrex** Access
- Several others in pipe

2006: Codian MCU

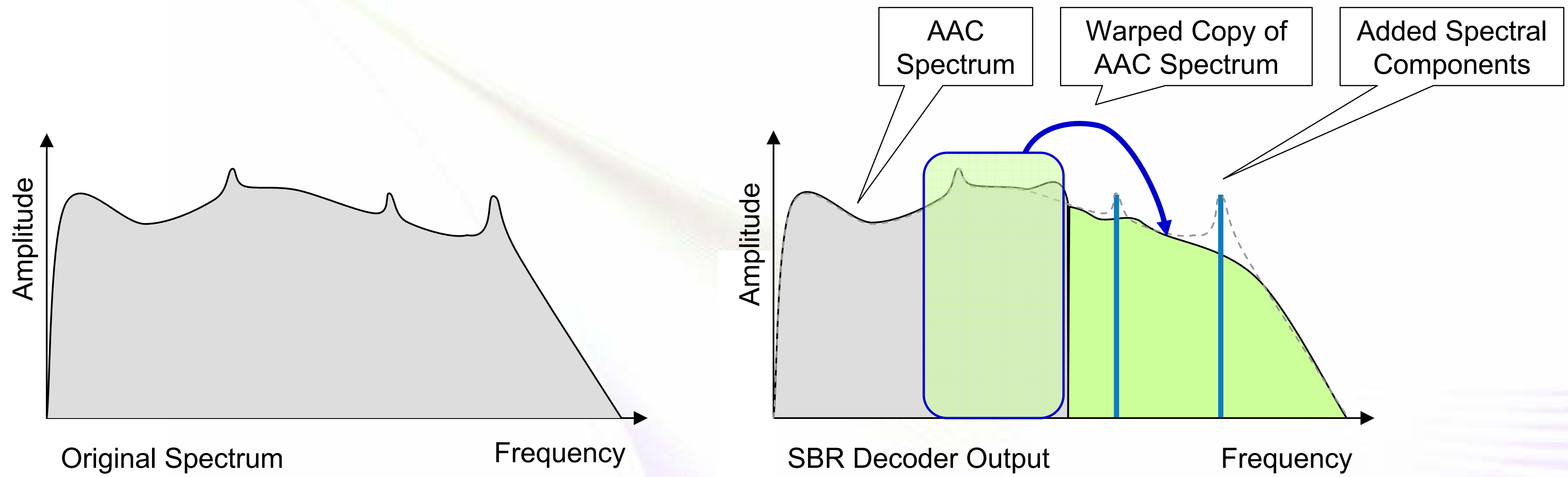
2006: Comrex Access

Enhanced Low Delay AAC

- Good audio quality at low bitrates lower than 48kbit/s
- Maintain a reasonable low algorithmic delay

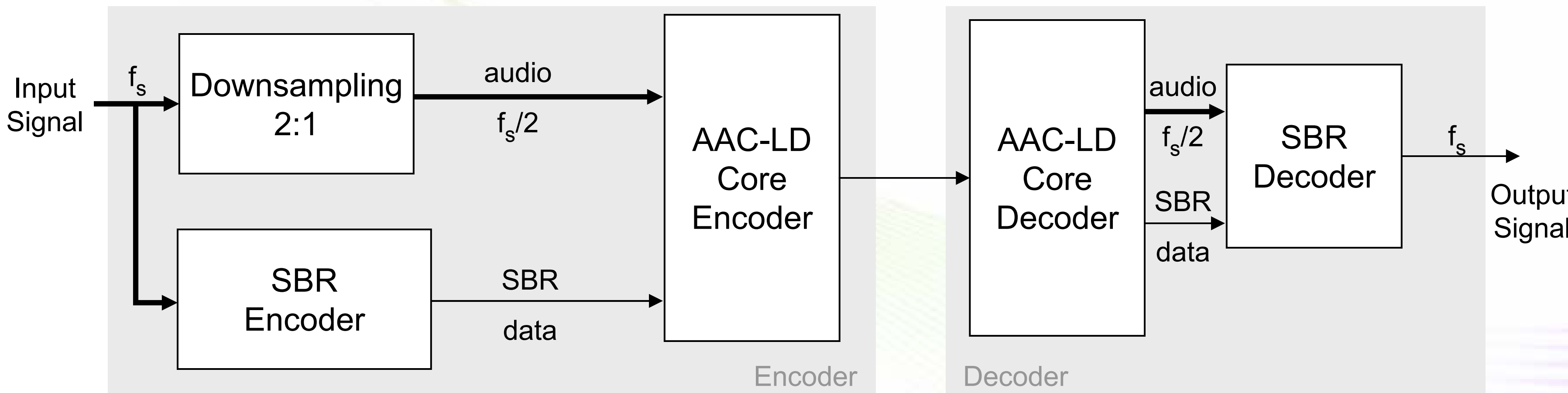
→ Use of SBR tool (used in HE-AAC)

Spectral Band Replication - Principle



- The top octave is not transmitted in the AAC signal. The SBR tool copies portions of the adjacent spectrum into the top octave, and modifies its envelope according to the SBR side data. The SBR decoder can also add noise or additional sinusoidal frequency components that were detected in the SBR encoder.

AAC-LD+SBR



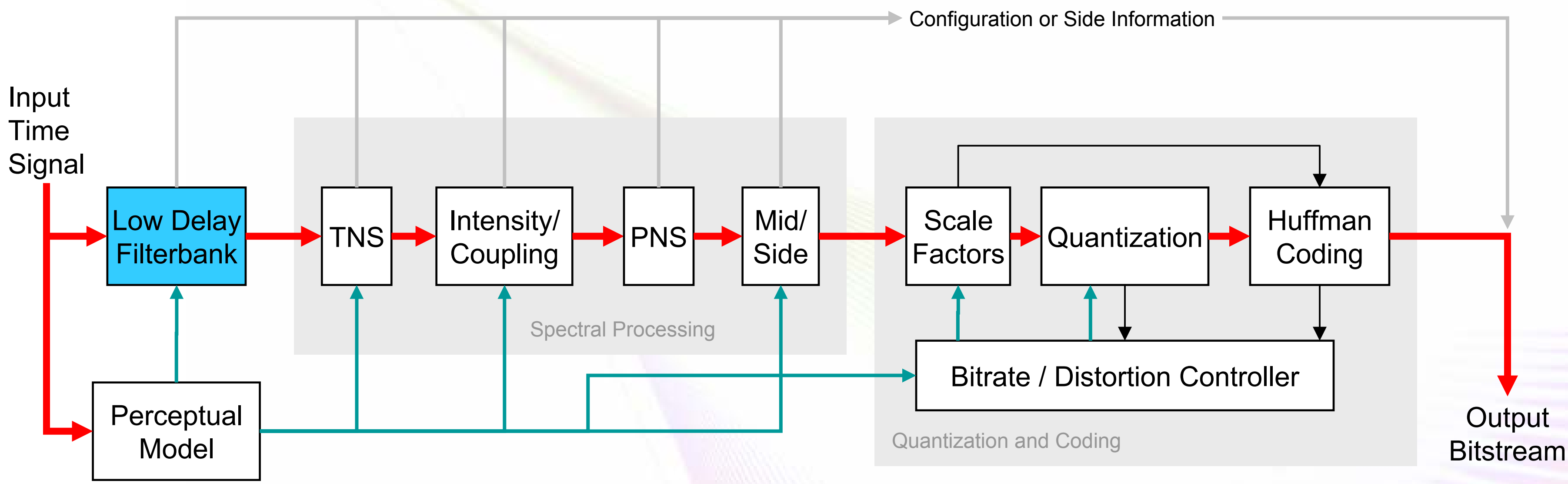
- Dual rate system causes further delay

→ delay optimizations necessary

AAC–LD+SBR Delay Analysis

Codec	Delay Sources	Delay at 48kHz
AAC-LD	MDCT+IMDCT	**20 ms**
AAC-LD+SBR	MDCT+IMDCT (dual rate)	40 ms
	QMF	12 ms
	SBR look ahead	8 ms
		60 ms
AAC-LD+LD-SBR	MDCT+IMDCT (dual rate)	40 ms
	QMF	12 ms
	SBR look ahead	0 ms
		52 ms

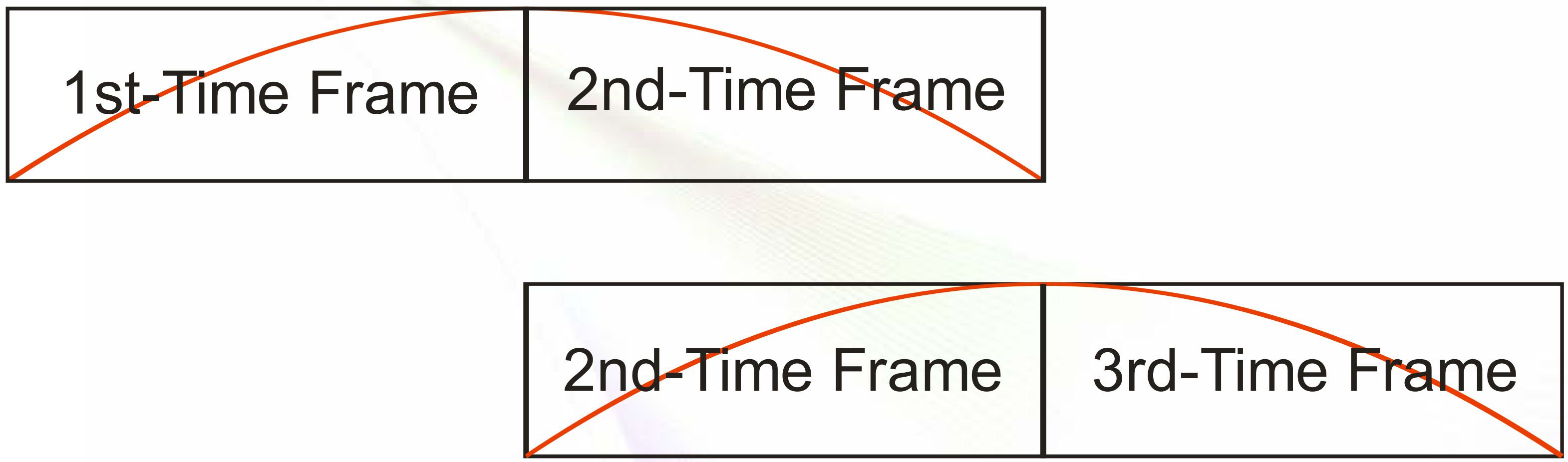
Enhanced Low Delay AAC Core



MPEG-4 AAC-LD Encoder

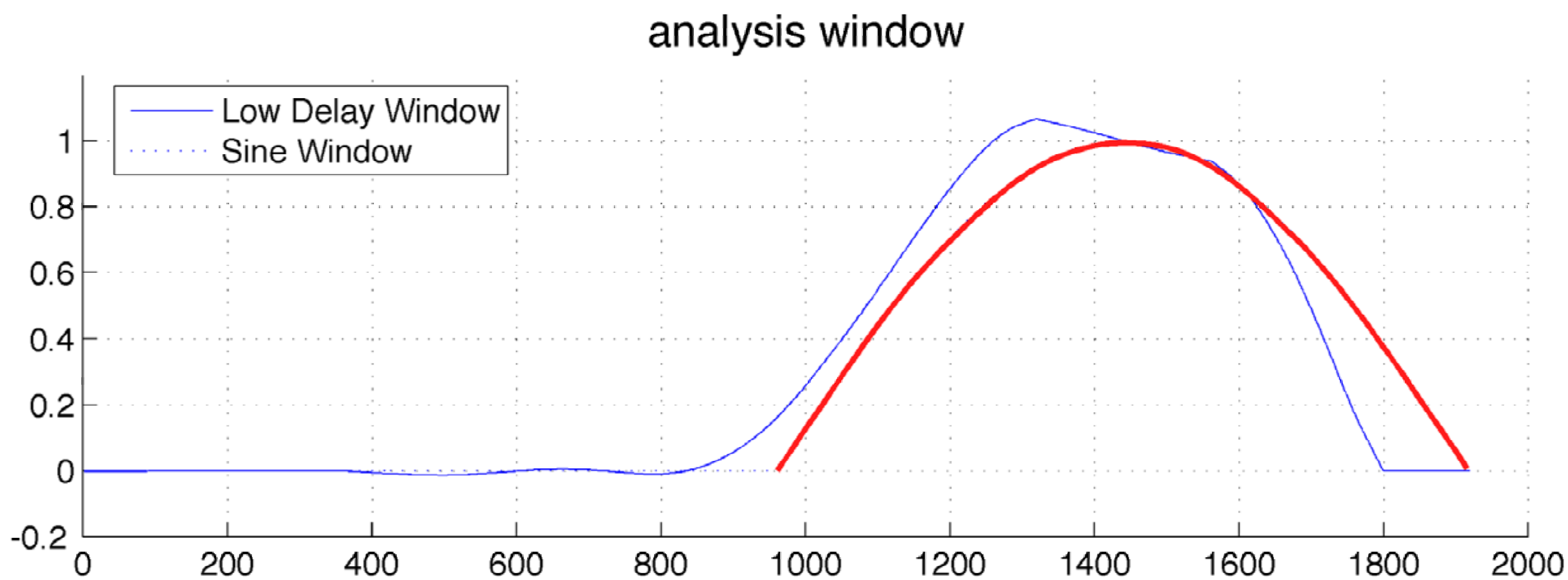
Bitstream Payload Formatter,
Bit Reservoir Not Shown

MDCT Overlap



- 50% window overlap over past and future samples

Low Delay Filterbank for AAC-ELD

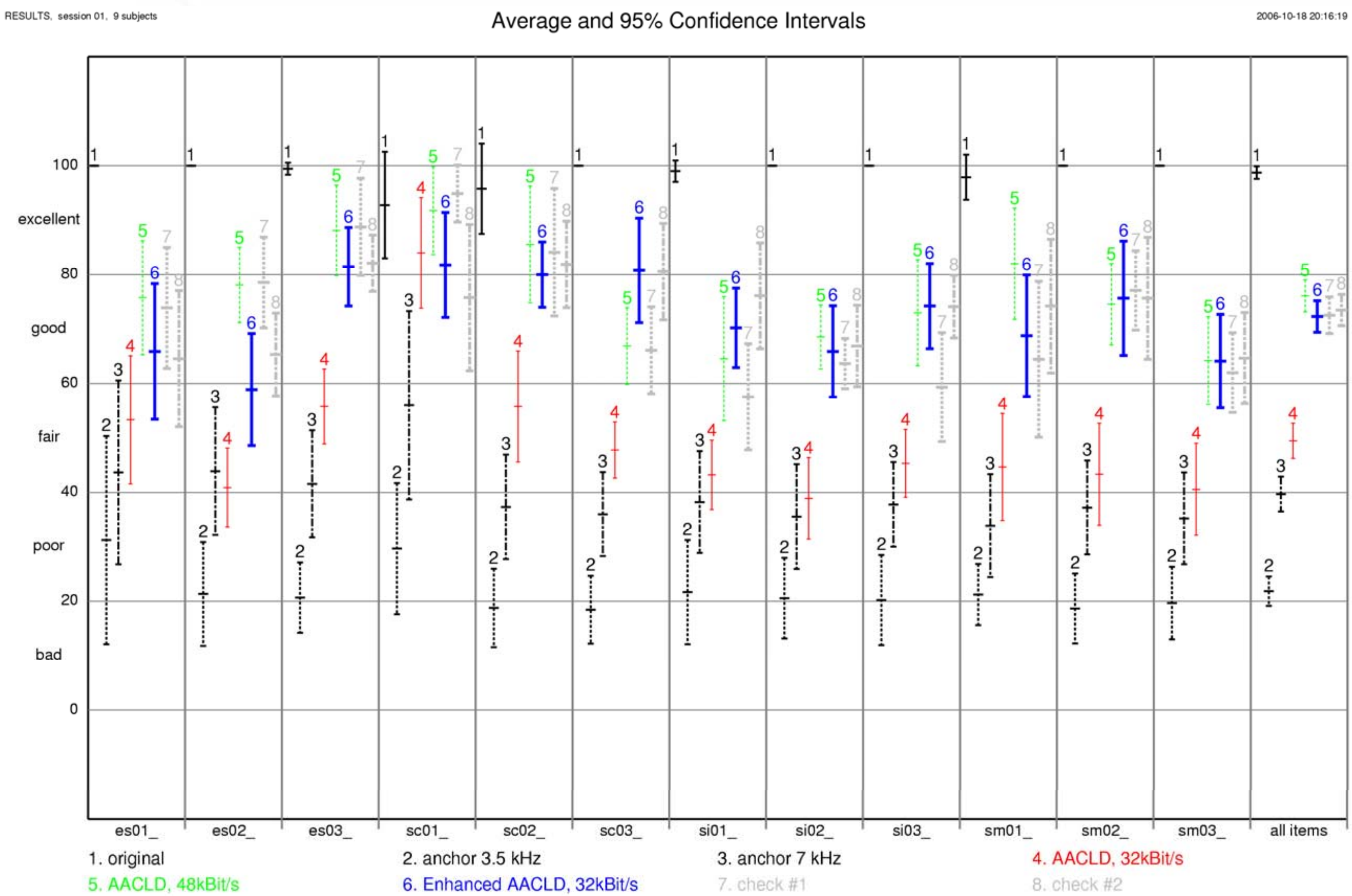


- Reduced overlap with future samples
  → reduced delay
- Audio quality equal

AAC–ELD Delay Analysis

Codec	Delay Sources	Delay at 48kHz
AAC-LD	MDCT+IMDCT	**20 ms**
AAC-LD+LD-SBR	MDCT+IMDCT	40 ms
	QMF	12 ms
		52 ms
AAC-ELD (LD-Filterbank + LD-SBR)	LD-Filterbank	30 ms
	QMF	12 ms
		42 ms

ELD AAC Listening Test



Agenda

- The Fraunhofer Institute for Integrated Circuits
- What Is Low Delay Audio Coding?
- AAC-Low Delay
- **Ultra Low Delay Audio Coding**
- MPEG Surround and Spatial Audio Object Coding
- Implementing AAC-Low Delay on TI DSPs

Applications with very critical delay requirements

- Jam over IP
- Wireless Rear loudspeakers, headphones, or hearing aids frequently need wireless connection
- Voice over IP

Requirements on audio codec:

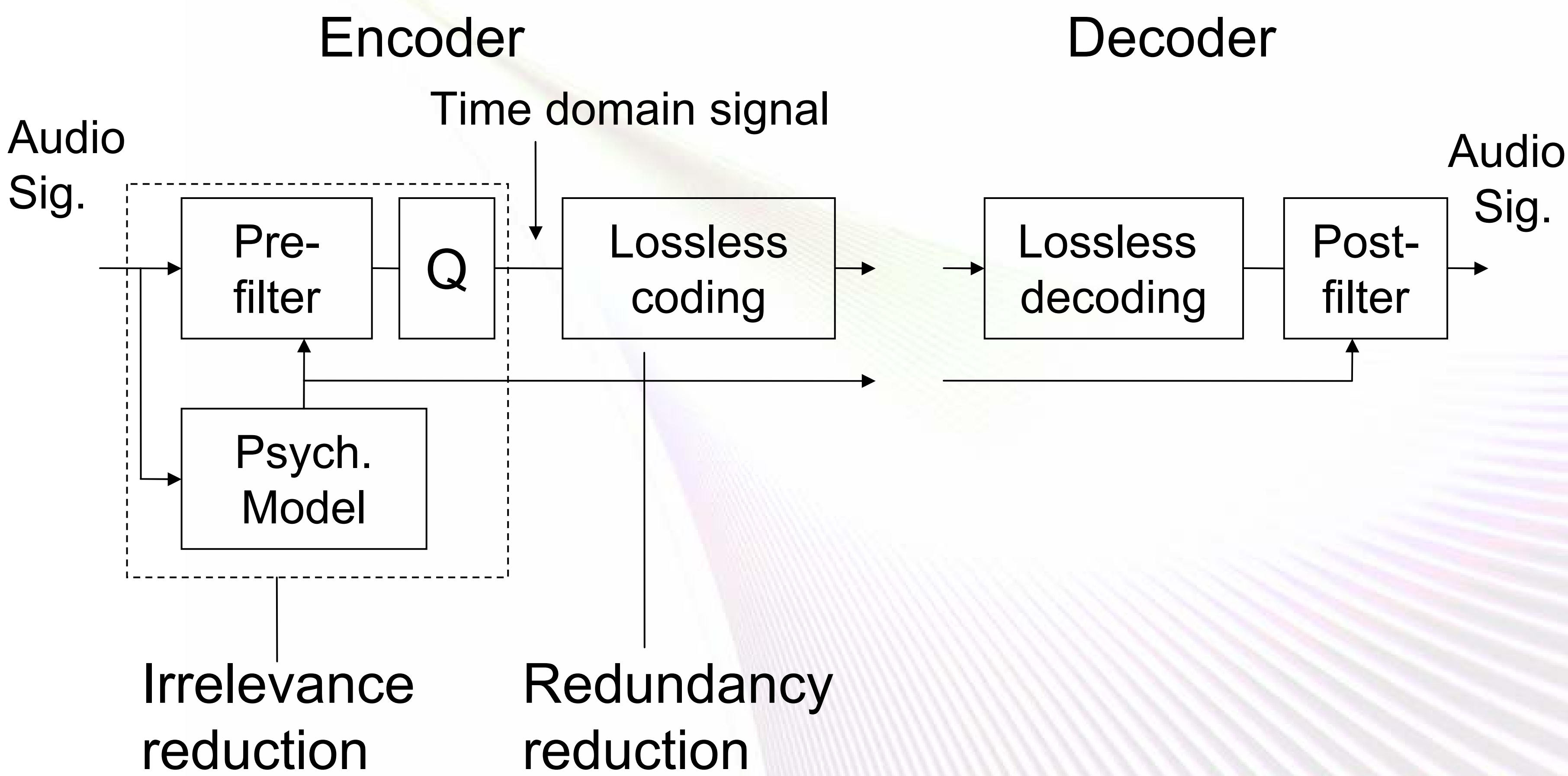
- High audio quality
- Video and lip synchronization
  $\longrightarrow$ delay < 20 ms
- Low data rate for robust wireless transmission

$\longrightarrow$ Ultra Low Delay audio codec

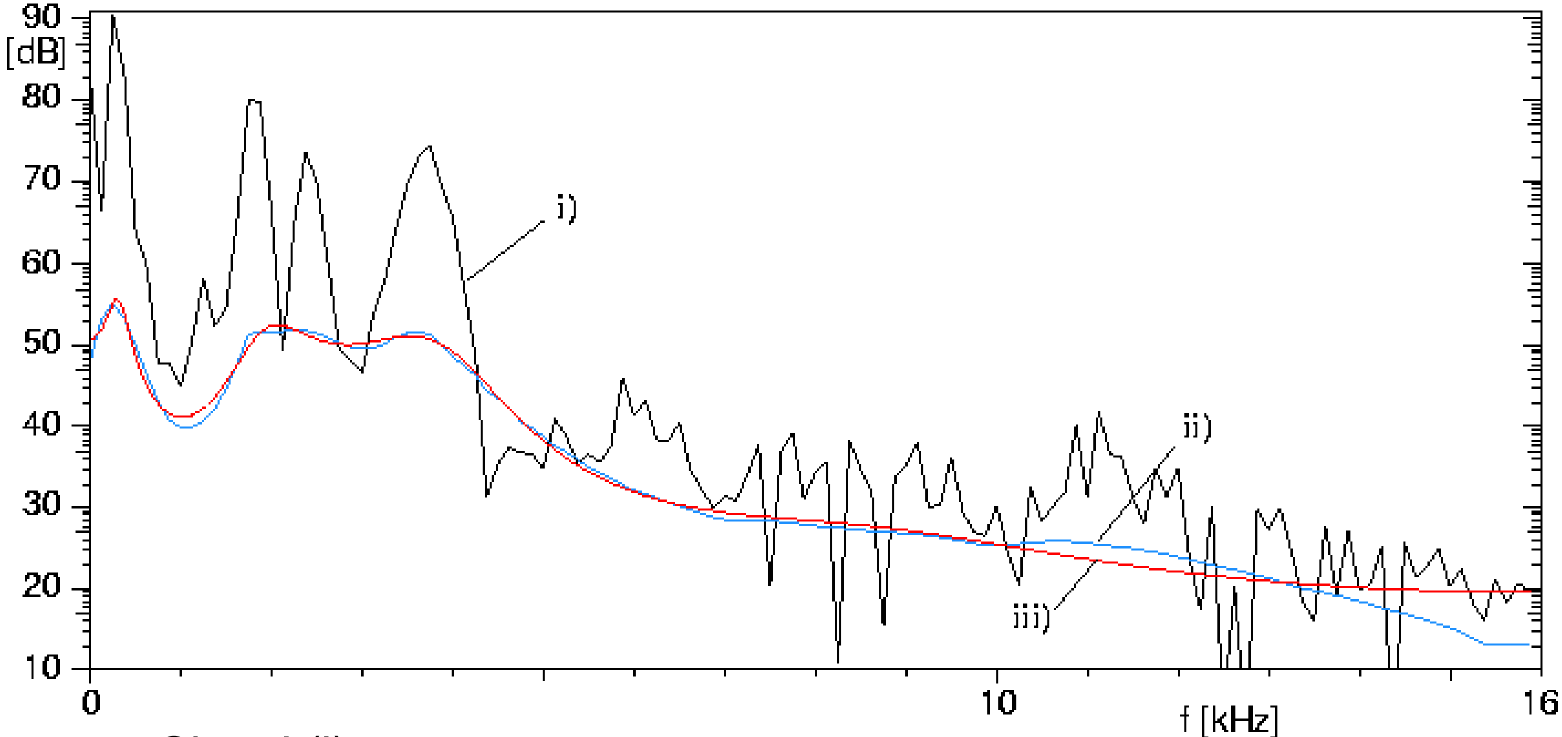
Ultra Low Delay Audio Codec

- Very low algorithmic delay (1.3 ms – 8 ms)
- High Audio Quality at 64-80 kbps/ch
- Complexity (16-Bit DSPs, Harvard Architecture)
  - Encoder ~100MHz,
  - Decoder ~53MHz
- Used for very delay critical applications

Pre- and Post-Filter Approach



Ultra Low Delay Audio Codec



- Signal (i)
- Masking threshold (ii)
- Impulse response of post-filter (iii)

Ultra Low Delay Audio Codec

- Delay depends on sampling rate and some other parameters

Sampling rate	Algorithmic Delay	Real-time delay / hardware
32 kHz	8 ms	17 ms/Demonstrator on TI 320C6713
44.1 kHz	5.8 ms	
48 kHz	5.3 ms	10 ms / suitable hardware
48 kHz	1.3 ms (low delay psych)	

Ultra Low Delay Audio Codec Summary

- Encoding/decoding delay ~6 ms at 44.1 kHz sampling
- High audio quality at 64-80 kbits/ch
- Used for very delay critical audio applications like
  - Jam over IP
  - Parlamential audio systems
  - Wireless loudspeakers
- Licensing of source code or object code through Fraunhofer IIS

Agenda

- The Fraunhofer Institute for Integrated Circuits
- What Is Low Delay Audio Coding?
- AAC-Low Delay
- Ultra Low Delay Audio Coding
- **MPEG Surround and Spatial Audio Object Coding**
- Implementing AAC-Low Delay on TI DSPs

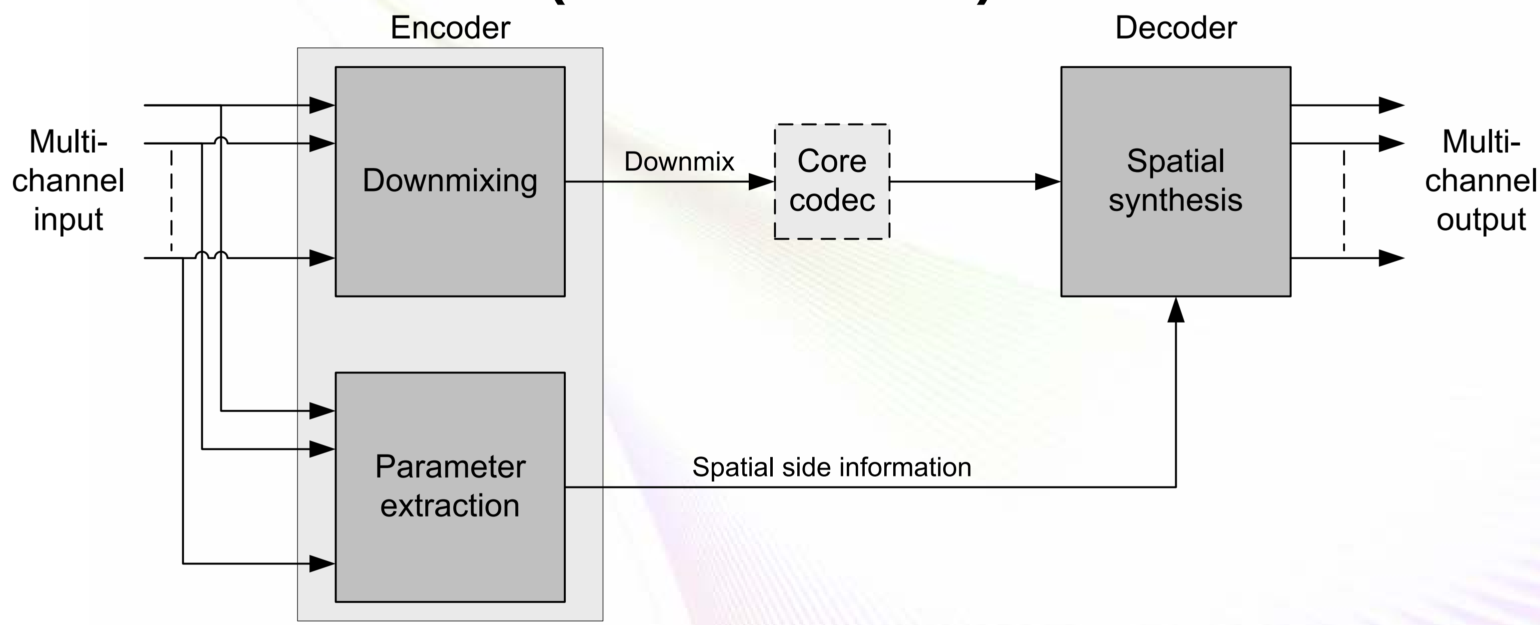
Improving Multi-Point Communications

- Transmit full audio bandwidth
- Keep coding delay low
  → ULD, AAC-LD

- Be able to hear several people talking at once
- Ambience
  – Sounds come from the room, not a speaker
- Adjust spatial position and loudness
- Modify characteristics of individual talkers or objects

Spatial Audio Coding (SAC)

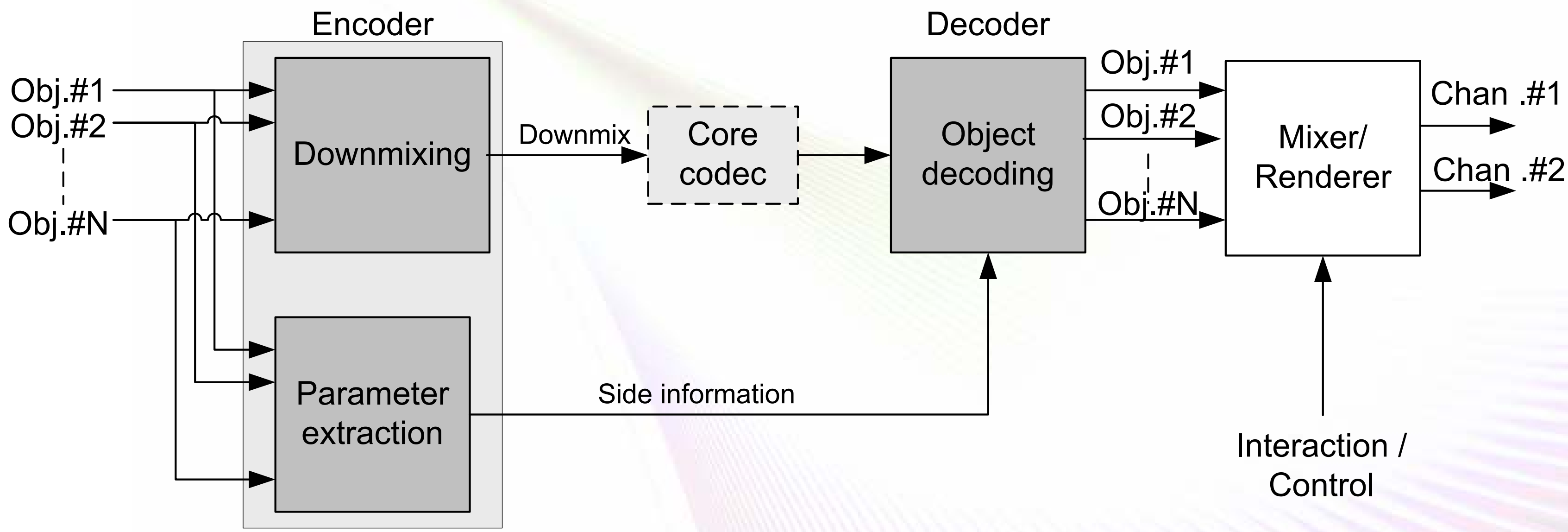
Current Spatial Audio Coding: Channel-oriented (MPEG Surround)



- Bitrate-efficient and compatible extension of existing audio distribution infrastructure towards multi-channel audio/surround sound.

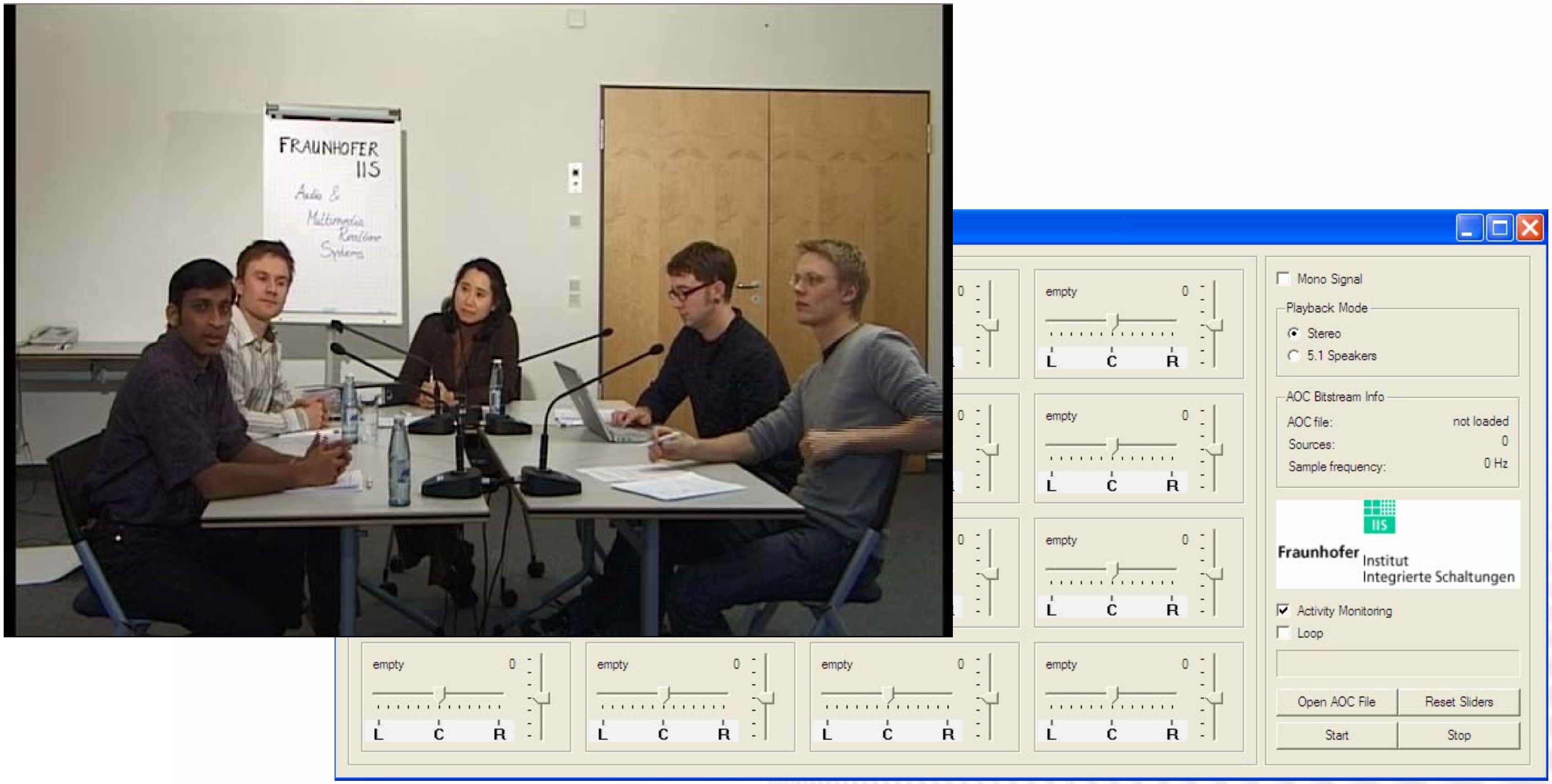
Spatial Audio Object Coding (SAOC)

Alternative Spatial Audio Coding: Object-oriented



- Focus on bitrate-efficient and compatible extension of existing audio distribution infrastructure towards object-based presentation

SAOC Demo


FRAUNHOFER IIS
empty
L C R
Mono Signal
Playback Mode
Stereo
5.1 Speakers
AOC Bitstream Info
AOC file: not loaded
Sources: 0
Sample frequency: 0 Hz
Fraunhofer Institut Integrierte Schaltungen
Activity Monitoring
Loop
Open AOC File
Reset Sliders
Start
Stop

Characteristics of Spatial Audio Object Coding

Efficiency	• **Very high compression efficiency (like SAC)**
Backward Compatibility	• **Backward compatible with existing transmission infrastructures (like SAC)** – Legacy monophonic signals are simply considered a *stream with only one object*
Output Interface / Rendering	• **Object output signals are suitable for feeding them jointly into a mixing/rendering engine** • **Decoded object signals can *in principle* be connected to any external mixing device …**

Agenda

- The Fraunhofer Institute for Integrated Circuits
- What Is Low Delay Audio Coding?
- AAC-Low Delay
- Ultra Low Delay Audio Coding
- MPEG Surround and Spatial Audio Object Coding
- **Implementing AAC-Low Delay on TI DSPs**

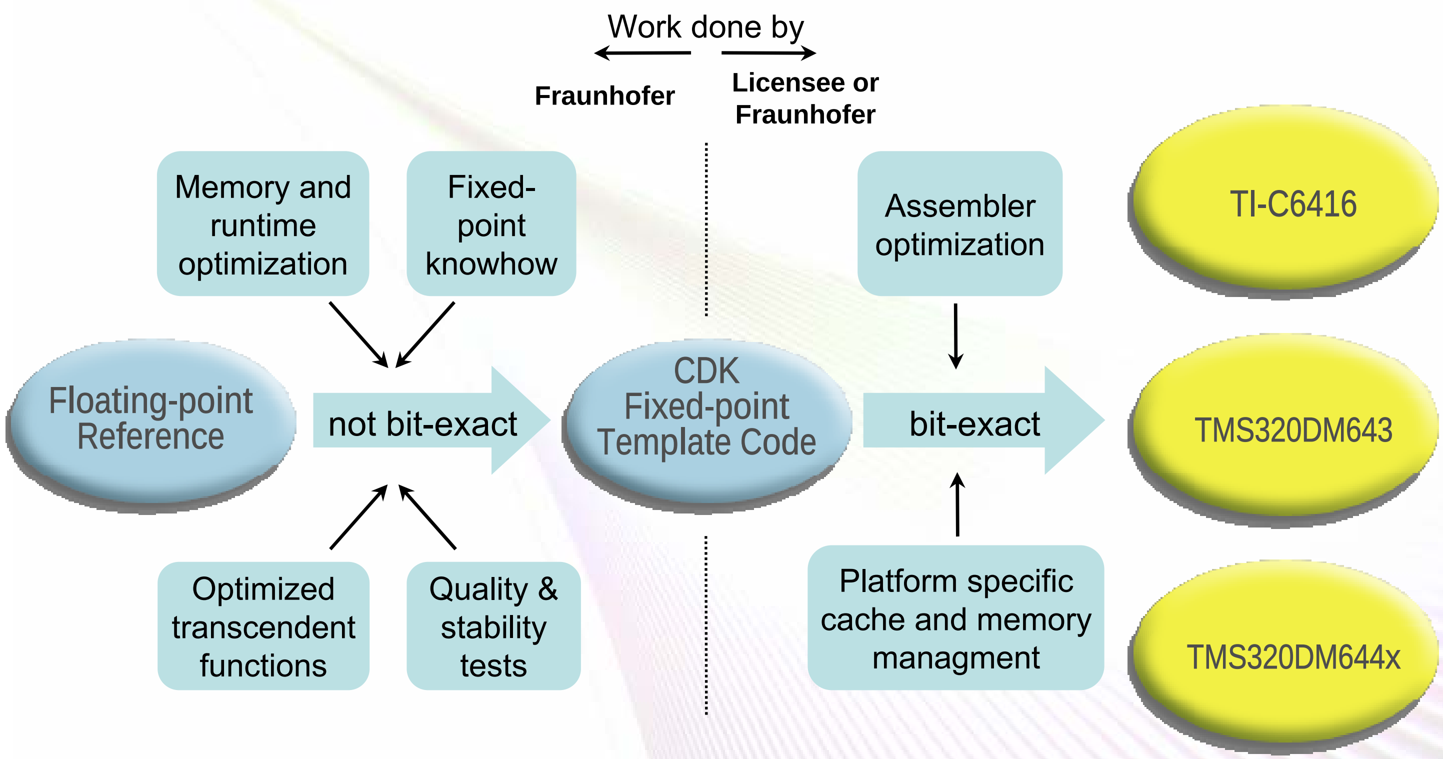
Implementing AAC Low-Delay on DSPs

- Fraunhofer's CDK source code and libraries
- Benchmarks on TI DSPs
- 6416 DSK free demo program
- Demonstration Example

Fraunhofer Target Platforms

- General-Purpose Computers
  - For software products – rippers, audio workstations, music managers and players
  - For operating system libraries
- Embedded Processors
  - For consumer electronics, such as the ARM and OMAP families
- Digital Signal Processors
  - Floating Point, such as the TI C67 family
  - Fixed Point, such as the TI C64x and C55 families

Development Flow



Fixed-Point Audio Coding Basics

- Heavy use of floating-point arithmetic in audio coding
  → needs to removed wherever possible
- Block floating-point
- Avoid divisions and square roots
- Clever use of arithmetic tricks and appropriate data types (16-bit, 32-bit, fractional) to meet dynamic range and precision requirements
- Encoder (MP3, AAC) much more complex than decoder

TI-C6416 and TMS320DM644x

TI-C6416

- Dual multipliers, each with two 16x16-bit mult/cycle
- Frequency range 300 – 1000 MHz (4000 MMAC)
- Normalization, Saturation, Bit-Counting

TMS320DM644x

- 594-MHz C64x+™ Clock Rate
  – Dual multipliers, each with four 16x16-bit mult/cycle
  – Normalization, Saturation, Bit-Counting
- 297-MHz ARM926EJ-S™ Clock Rate
  – Support for 32-Bit and 16-Bit (Thumb Mode) Instruction Set
  – DSP Instruction Extensions and Single Cycle MAC

Processing Power Requirements

All figures in MHz for 44.1 kHz stereo content	Encoder			Decoder		
	TI-C6416	TI-C5510	TI-DM644x (ARM926)	TI-C6416	TI-C5510	TI-DM644x (ARM926)
MP3	45	130	80	20	45	27
MPEG-4 AAC-LC (TNS + PNS)	27	140	58	11	30	21
MPEG-4 HE-AAC (TNS + PNS)	60	tbd	100	23	tbd	47
MPEG-4 AAC Low Delay	18	(160 estim.)	85	14	(40 estim.)	35

Free AAC-LD Demo on 6416 DSK Board

- Demonstration Program (.out file) runs on 6416 DSK Board
- Real-time AAC Low Delay encoding and decoding
- Selectable stereo bit rates of 64, 96, 128, 160 kb/s
- Uses on-board AIC 23 ADC/DAC – plug analog sources and headphones directly into board
- Example main() source code provided, based on TI DSP/BIOS audio pipe example
- Available now to FhG customers, free download coming soon

Available Audio Codecs for TI DSPs

- MPEG-4 AAC
  - AAC Low Complexity and AAC Low Delay
- MPEG-4 HE-AAC
  - High Efficiency AAC v2 (aka. "enhanced AAC+ v2")
- MPEG Layer-3 (MP3)
- MP3 Surround Decoder
- ULD: Ultra Low Delay audio codec
- MPEG Surround

- MP3 products licensed through Thomson Multimedia

References

1. Eric Allamanche, Ralf Geiger, Juergen Herre and Thomas Sporer (FhG-IIS). "MPEG-4 Low Delay Audio Coding based on the AAC Codec", presented at the 106th Convention of the Audio Engineering Society, May 8-11, 1999, Munich, Germany. Preprint number 4929. www.aes.org.
2. Ralf Geiger, Manfred Lutzky, Markus Schnell, Markus Schmidt (FhG-IIS). "Structural Analysis of Low Latency Audio Coding Schemes", presented at the 119th Convention of the Audio Engineering Society, October 7-10, 2005, New York, NY. Preprint number 6601. www.aes.org.
3. Manfred Lutzky, Gerald Schuller, Marc Gayer, Ulrich Krämer, Stefan Wabnik (FhG-IIS, FhG-IDMT)."A Guideline to Audio Codec Delay, presented at the 116th Convention of the Audio Engineering Society, May 8–11, 2004, Berlin, Germany. Preprint number 6062. www.aes.org.
4. Karlheinz Brandenburg (FhG-IIS). "MP3 and AAC Explained", presented at the AES 17th International Conference on High Quality Audio Coding, September 2-5, 1999, Florence, Italy. www.aes.org.
5. Advanced Television Systems Committee. "ATSC Implementation Subcommittee Finding: Relative Timing of Sound and Vision for Broadcast Operations", IS-191, 26 Jun 2003. www.atsc.org.
6. ISO/IEC 14496-3:2001, "Information technology -- Coding of audio-visual objects -- Part 3: Audio". (MPEG-4 Audio) www.iso.ch
7. ITU-T Recommendation G.107 (03/2005), "The E-model, a computational model for use in transmission planning" www.itu.int

Copies of these documents are available from the links provided, and in some cases, from your Fraunhofer representative.

TI | Developer Conference
March 7-9, 2007 • Dallas, TX

Low Delay Audio Coding for TI DSPs

Negjmedin Fazlija
Fraunhofer IIS
faz@iis.fraunhofer.de

Fraunhofer Institut Integrierte Schaltungen

Minds in Motion