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

Negjmedin Fazlija
Fraunhofer IIS
faz@iis.fraunhofer.de
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

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

58 Fraunhofer Institutes in 40 locations in Germany
Fraunhofer IIS
Institute for Integrated Circuits

• 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

• 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.
• 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
• 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

<table>
<thead>
<tr>
<th>Codec</th>
<th>Typical Delay (HW or DSP)</th>
<th>Typical Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP3</td>
<td>140 ms</td>
<td>Music Player</td>
</tr>
<tr>
<td>AAC</td>
<td>210 ms</td>
<td>Music Player, Broadcasting</td>
</tr>
<tr>
<td>HE-AAC</td>
<td>360 ms</td>
<td>Mobile Music, Satellite Radio</td>
</tr>
</tbody>
</table>
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:

ITU G.107 E-model Quality Score

G.107 Table 2 Default Values, except echo case

- Very Satisfied
- Satisfied
- Some Users Dissatisfied
- Many Users Dissatisfied
- Nearly All Users Dissatisfied

One-Way Delay \( T_a = T \), ms

R

G.711, G.728, 16 kb/s
G.726, 32 kb/s
G.729
G.711, -40dB Echo

Nearly All
Users
Dissatisfied
Many Users
Dissatisfied
Some Users
Dissatisfied
Satisfied
Very Satisfied
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
Example: Next-Generation Video Conference

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

- **CD like Quality (BW >15 kHz)**
  - ULD 80 kbps
  - AAC-LD 64 kbps
  - AAC-LC 56 kbps
- **Wide Band (BW 7 kHz)**
  - MP3 64 kbps
  - HE-AAC 28 kbps
- **Narrow Band (BW 3.5 kHz)**
  - Telephone codec
  - Speech codec (Speech)
  - Speech codec (Music)

Audio quality: 10 20 130 30 40 50

Algorithmic delay in ms: 100 130

TI Developer Conference

Minds in Motion

Technology for Innovators™

Texas Instruments
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

- **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’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.
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
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

<table>
<thead>
<tr>
<th>Codec</th>
<th>Delay Sources</th>
<th>Delay at 48kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC-LD</td>
<td>MDCT+IMDCT</td>
<td>20 ms</td>
</tr>
<tr>
<td>AAC-LD+SBR</td>
<td>MDCT+IMDCT (dual rate)</td>
<td>40 ms</td>
</tr>
<tr>
<td></td>
<td>QMF</td>
<td>12 ms</td>
</tr>
<tr>
<td></td>
<td>SBR look ahead</td>
<td>8 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>60 ms</strong></td>
</tr>
<tr>
<td>AAC-LD+LD-SBR</td>
<td>MDCT+IMDCT (dual rate)</td>
<td>40 ms</td>
</tr>
<tr>
<td></td>
<td>QMF</td>
<td>12 ms</td>
</tr>
<tr>
<td></td>
<td>SBR look ahead</td>
<td>0 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>52 ms</strong></td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Codec</th>
<th>Delay Sources</th>
<th>Delay at 48kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAC-LD</td>
<td>MDCT+IMDCT</td>
<td>20 ms</td>
</tr>
<tr>
<td>AAC-LD+LD-SBR</td>
<td>MDCT+IMDCT</td>
<td>40 ms</td>
</tr>
<tr>
<td></td>
<td>QMF</td>
<td>12 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>52 ms</strong></td>
</tr>
<tr>
<td>AAC-ELD</td>
<td>LD-Filterbank</td>
<td>30 ms</td>
</tr>
<tr>
<td>(LD-Filterbank + LD-SBR)</td>
<td>QMF</td>
<td>12 ms</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>42 ms</strong></td>
</tr>
</tbody>
</table>
ELD AAC Listening Test

Average and 95% Confidence Intervals

RESULTS. session 01. 9 subjects

1. original
5. AACLD, 48kBit/s

2. anchor 3.5 kHz
6. Enhanced AACLD, 32kBit/s

3. anchor 7 kHz
7. check #1

4. AACLD, 32kBit/s
8. check #2

Technology for Innovators™
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 → delay < 20 ms
- Low data rate for robust wireless transmission

→ 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

Encoder

- Audio Sig.
- Pre-filter
- Psych. Model
- Irrelevance reduction
- Time domain signal
- Q
- Lossless coding
- Redundancy reduction

Decoder

- Lossless decoding
- Post-filter
- Audio Sig.
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

<table>
<thead>
<tr>
<th>Sampling rate</th>
<th>Algorithmic Delay</th>
<th>Real-time delay / hardware</th>
</tr>
</thead>
<tbody>
<tr>
<td>32 kHz</td>
<td>8 ms</td>
<td>17 ms/Demonstrator on TI 320C6713</td>
</tr>
<tr>
<td>44.1 kHz</td>
<td>5.8 ms</td>
<td></td>
</tr>
<tr>
<td>48 kHz</td>
<td>5.3 ms</td>
<td>10 ms / suitable hardware</td>
</tr>
<tr>
<td>48 kHz</td>
<td>1.3 ms (low delay psych)</td>
<td></td>
</tr>
</tbody>
</table>
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
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

- Floating-point Reference
  - Memory and runtime optimization
  - Fixed-point knowhow
- CDK Fixed-point Template Code
  - Assembler optimization
  - Platform specific cache and memory management
  - Optimized transcendent functions
  - Quality & stability tests
- Work done by
  - Fraunhofer
  - Licensee or Fraunhofer
- TI-C6416
- TMS320DM643
- TMS320DM644x

Technology for Innovators™
Fixed-Point Audio Coding Basics

- Heavy use of floating-point arithmetic in audio coding → needs to be 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

<table>
<thead>
<tr>
<th></th>
<th>Encoder</th>
<th>Decoder</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TI-C6416</td>
<td>TI-C5510</td>
</tr>
<tr>
<td>MP3</td>
<td>45</td>
<td>130</td>
</tr>
<tr>
<td>MPEG-4 AAC-LC (TNS + PNS)</td>
<td>27</td>
<td>140</td>
</tr>
<tr>
<td>MPEG-4 HE-AAC (TNS + PNS)</td>
<td>60</td>
<td>tbd</td>
</tr>
<tr>
<td>MPEG-4 AAC Low Delay</td>
<td>18</td>
<td>(160 estim.)</td>
</tr>
</tbody>
</table>
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


Copies of these documents are available from the links provided, and in some cases, from your Fraunhofer representative.
Low Delay Audio Coding for TI DSPs

Negjmedin Fazlija
Fraunhofer IIS
faz@iis.fraunhofer.de