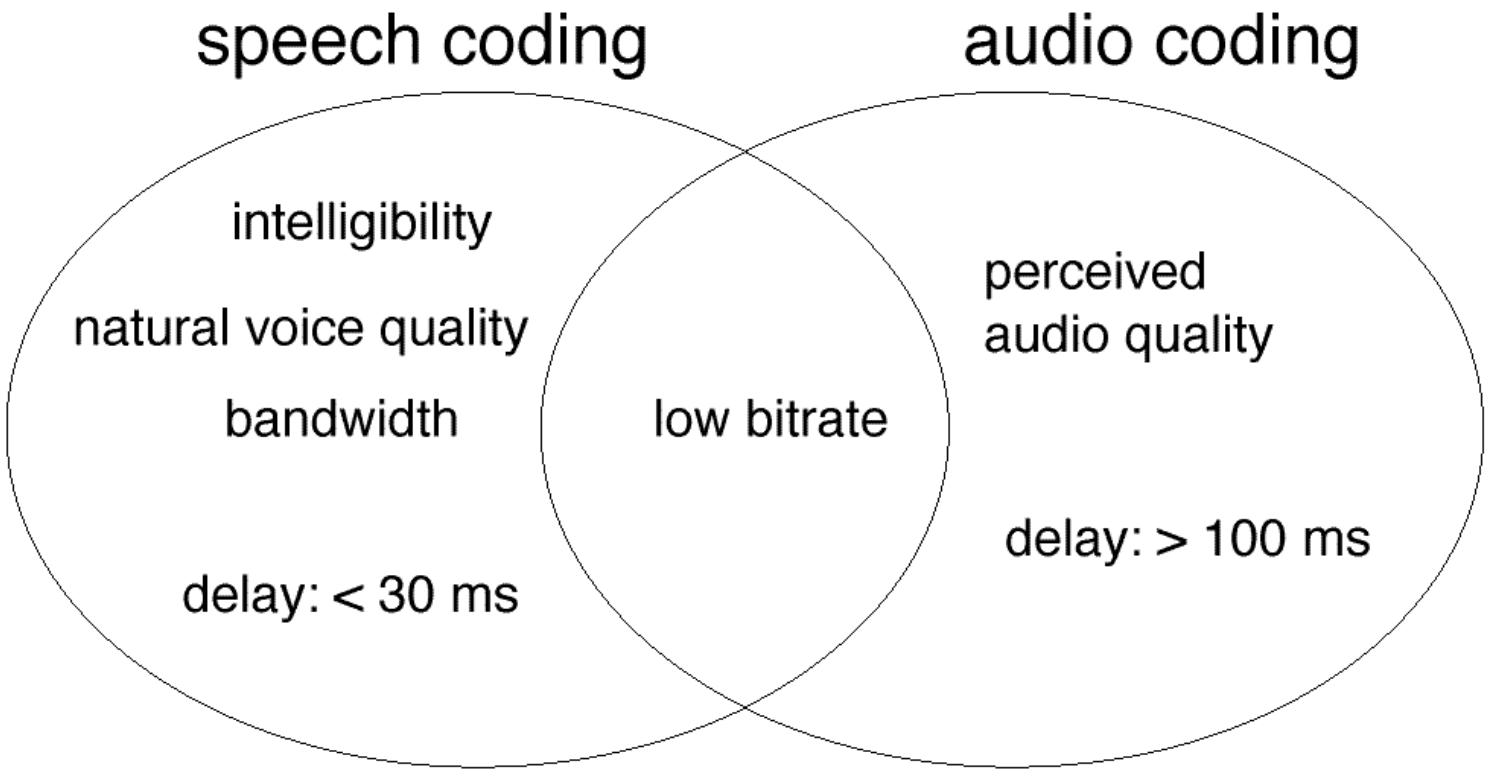
MPEG-4 Low Delay Audio Coding

Jürgen Herre
Fraunhofer Institut for Integrated Circuits (FhG-IIS)
Erlangen, Germany

Overview

- Speech coding vs. perceptual audio coding
- MPEG-4 Version 2 low delay coding
- Delay sources in perceptual audio coding
- Low delay AAC coder
- Test results
- Summary

Speech Coding vs. Audio Coding



MPEG-4 Version 2 Low Delay Audio Coding

Target

- High audio and speech quality *and*
- Low bitrate *and*
- Low algorithmic delay (20 ms)

Solution

MPEG-4 Version 2 Low Delay Audio Coder:

- Derived from MPEG-2/4 "Advanced Audio Coding" (AAC)
- Specific modifications for low-delay operation

Fraunhofer Institut Integrierte Schaltungen

Delay Sources in Perceptual Audio Coding

- Framing delay
- Filterbank delay
- Look-ahead delay for block switching
- Use of bit reservoir

$\Rightarrow$ Overall delay:

$$t_{\text{delay}} = \frac{N_{\text{framing}} + N_{\text{filterbank}} + N_{\text{look-ahead}}}{F_{\text{s}}} + t_{\text{bitres}}$$

Fraunhofer Institut
Integrierte Schaltungen

Example: Delay of AAC Codec (48 kHz / 64 kbps)

- Framing delay : 1024 samples
- Filterbank delay : 1024 samples
- Look-ahead delay for block switching : 576 samples
- Use of bit reservoir : 74.7 ms

$\Rightarrow$ Overall delay:

$$t_{\text{delay}} = \frac{1024+1024+576}{48000} + 74.7\ \text{ms} = 129.4\ \text{ms}$$

Fraunhofer Institut
Integrierte Schaltungen

Low Delay AAC Codec (48 kHz, min. delay mode)

- Reduced framing delay : 480 samples
- Reduced filter bank delay : 480 samples
- No block switching

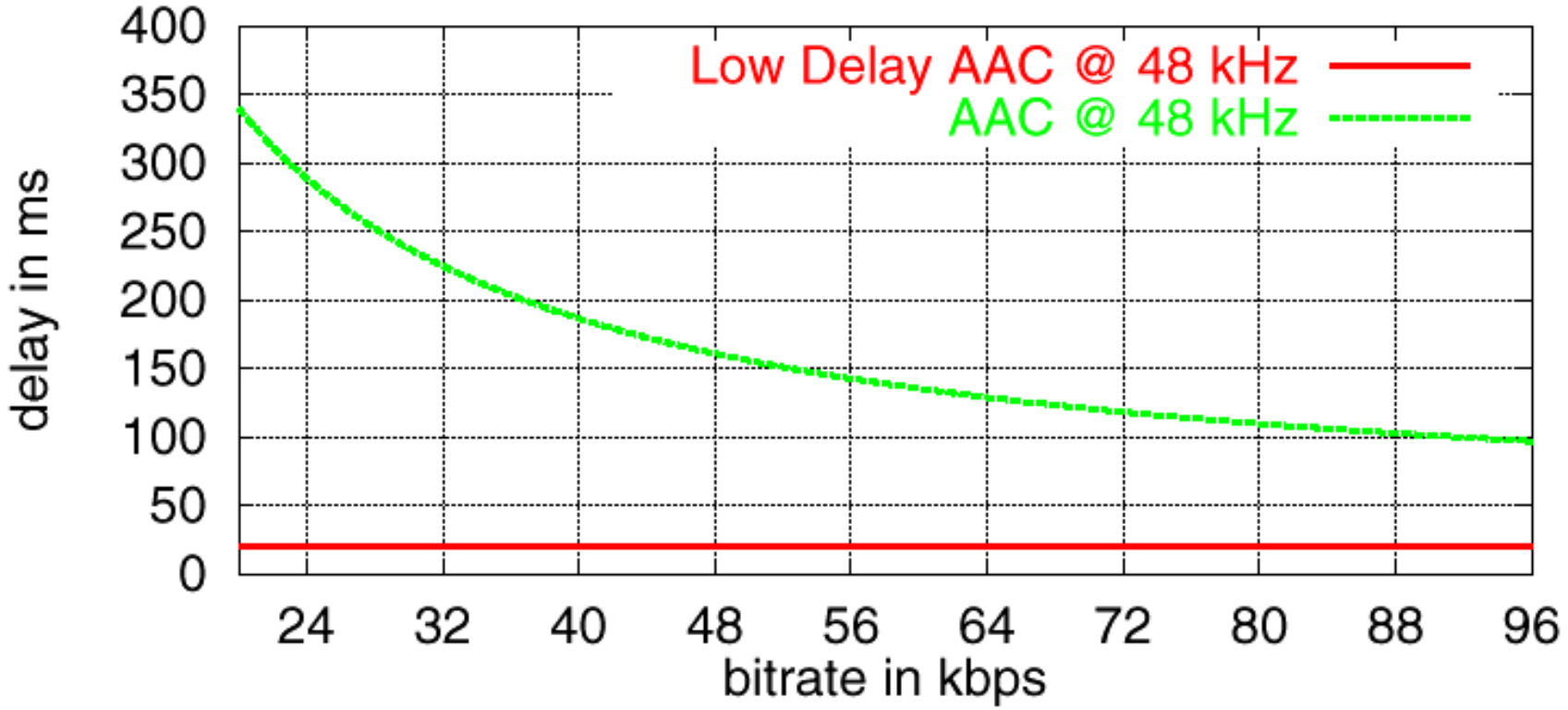
  $\Rightarrow$ no look-ahead delay: 0 samples
- Minimal bit reservoir : 0...32 bits

$\Rightarrow$ Overall delay:

$$t_{\text{delay}} = \frac{480 + 480 + 0}{48000} + 0\,\text{ms} = 20\,\text{ms}$$

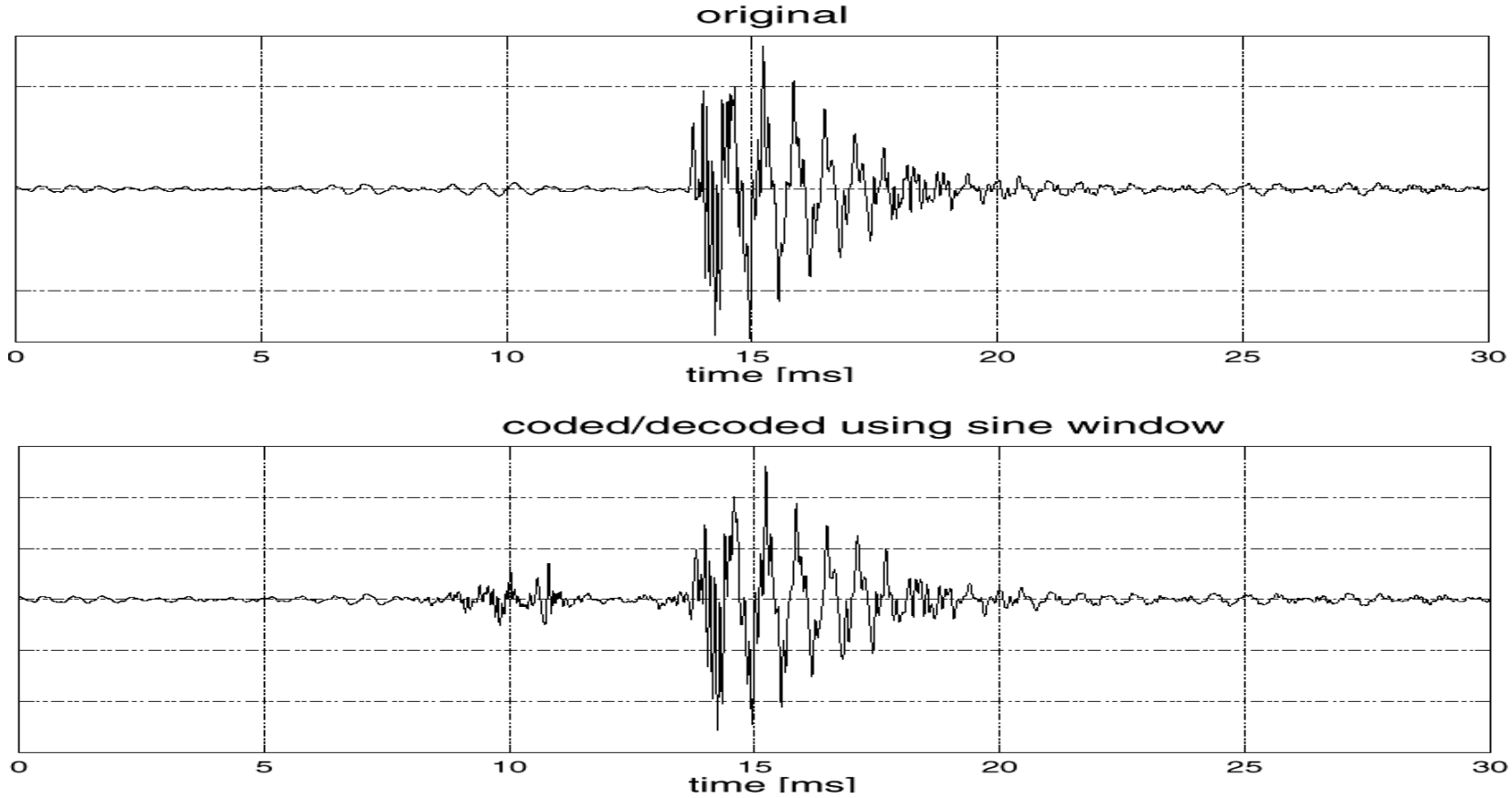
Fraunhofer Institut
Integrierte Schaltungen

Delay vs. Bitrate



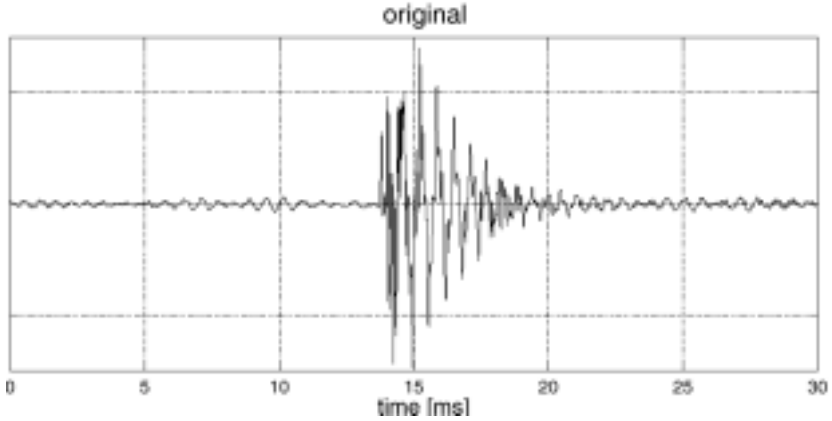
Fraunhofer Institut
Integrierte Schaltungen

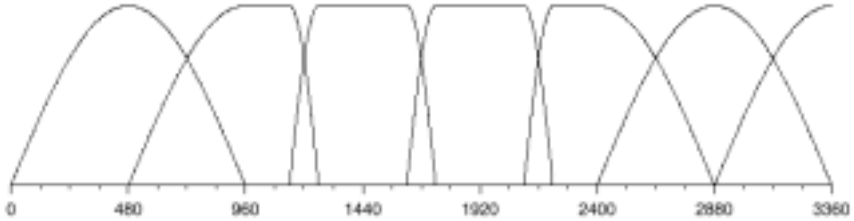
Preecho Behaviour

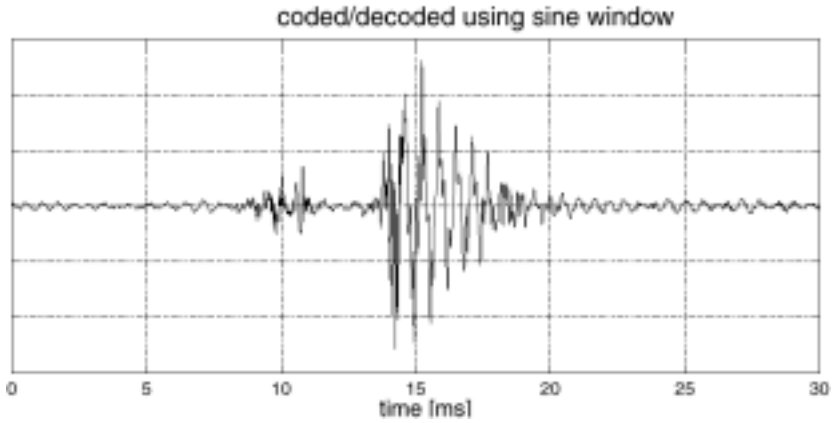


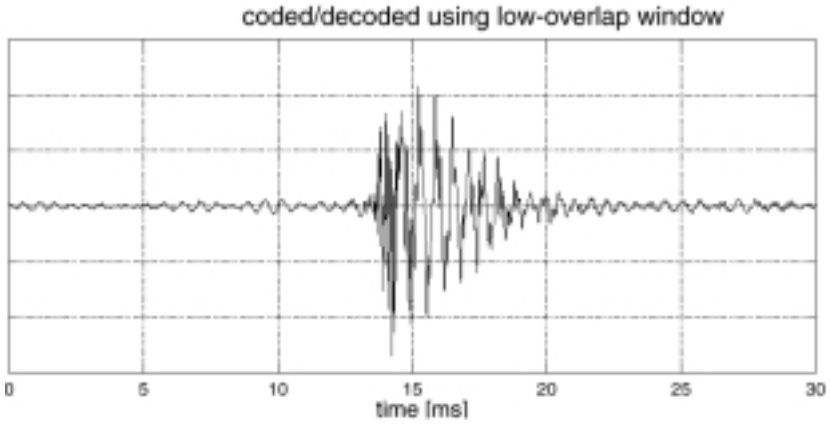
Fraunhofer Institut
Integrierte Schaltungen

Preecho Reduction by Window Shape Adaptation









Fraunhofer Institut
Integrierte Schaltungen

Test Results (1) - MPEG-4 V2 Verification Tests

Test #1

- Comparable performance of 32 kbps LD-AAC and 24 kbps AAC (main profile)
  - Reduction of delay by 91% (323ms → 30 ms)
  - Bitrate increase of 33%

Test #2

- Comparable performance of 64 kbps LD-AAC and 56 kbps AAC (main profile)
  - Reduction of delay by 86% (146ms → 20 ms)
  - Bitrate increase of 14%

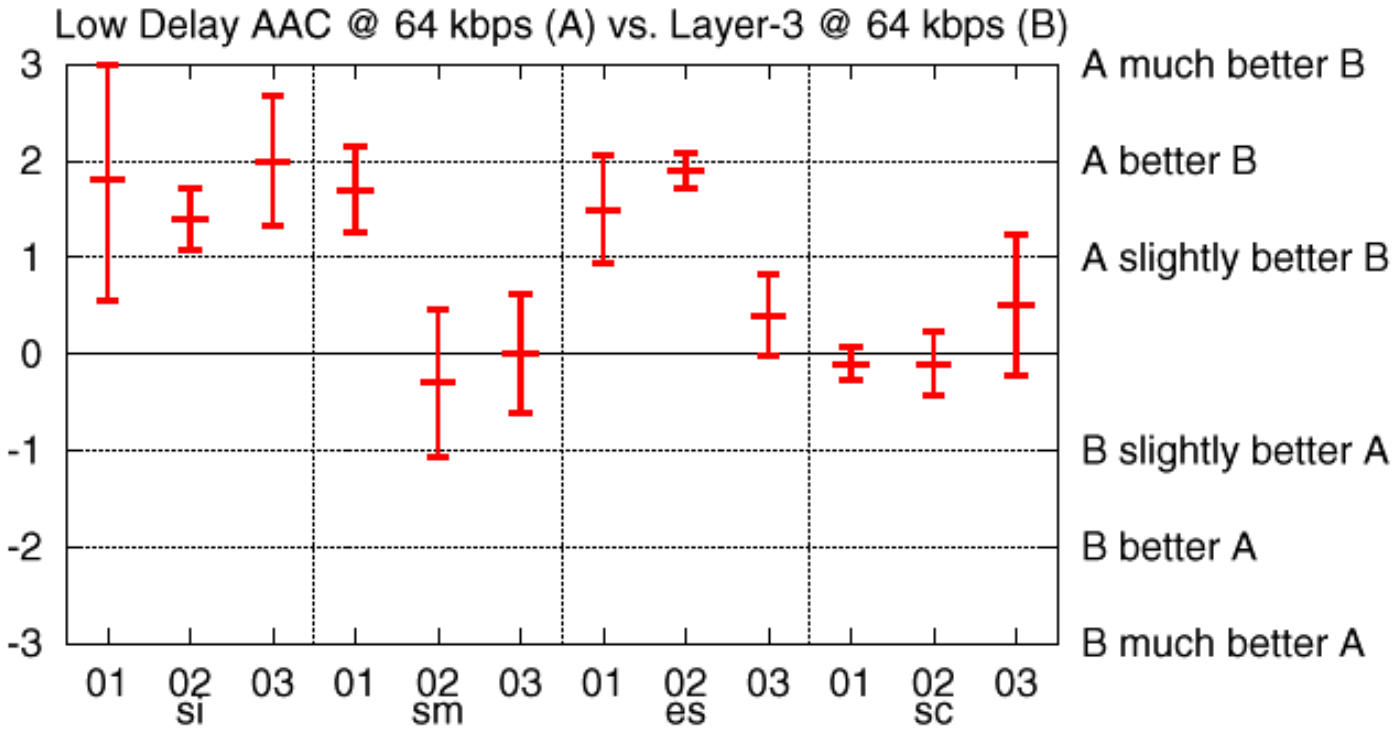
Summary

- Better performance than CELP codecs for applications with both audio and speech

MPEG Test Report

- http://drogo.cselt.it/mpeg/ public/

Test Results (2) - Comparison to "MP3"



⇒ Better than MPEG-1 Layer 3 ("MP3") at 64 kbps

Fraunhofer Institut
Integrierte Schaltungen

Summary

The MPEG-4 V2 low-delay coder provides

- High audio quality for music *and* speech
- Algorithmic delay of 20 ms

  $\Rightarrow$ enables two-way communications
- Audio quality scales with bitrate
- Stereo and multi-channel capabilities (inherited from MPEG-2/4 AAC)
- Compares to well w.r.t. both standard AAC and MP3
- Low computational & memory complexity
- Error robustness

Fraunhofer Institut
Integrierte Schaltungen

Demonstration

Real-time codecs implemented on
563xx-based PCMCIA DSP cards:

- MPEG-2 AAC LC
  (measured delay ca. 150ms;
  64kbps mono, fs = 48 kHz)

- MPEG-4 Version 2 Low-Delay AAC Codec
  (measured delay: ca. 45 ms;
  64kbps mono, fs = 48 kHz)