# Audio & Multimedia Realtime Systems Department


# MPEG low delay audio codecs

**Manfred Lutzky**
Group Manager Audio for Communication

# MPEG low delay audio codecs

## Why MPEG codecs on ETSI STQ workshop?

Really?

ETSI work traditionally based on ITU-T codecs

**MPEG**

- Application: Music streaming, broadcasting
- high bitrates, high delay, stereo-> Multichannel

**ITU-T**

- Application: telephony
- low bitrate, low delay, narrow band, low compex

# MPEG low delay audio codecs

## Well known standards

| Codec | Quality | Band-width | Bitrate | Delay | Complexity | application |
|---|---|---|---|---|---|---|
| **MPEG** mp3/AAC | Up to perceptual transaprency | ~16 kHz | 48-92 kbps/ch | >100 ms | Medium | Broadcasting music download |
| | | | | | | |
| | | | | | | |
| G.726,G.729 **ITU-T** | Toll | 3.5kHz | < 32kbps | < 20ms | Low | Telecommunication |

# MPEG low delay audio codecs

## latest enhancements are overlapping

| Codec | Quality | Band-width | Bitrate | Delay | Complexity | application |
|---|---|---|---|---|---|---|
| **MPEG** mp3/AAC | Up to perceptual transparency | ~16 kHz | 48-92 kbps/ch | >100 ms | Medium | Broadcasting music download |
| AAC-LD | Up to perceptual transparency | ~16 kHz | 48-92 kbps/ch | 20 ms | Medium | Conferencing communication |
| AAC-ELD | high | ~16 kHz | 24-48kbps/ch | 15-32ms | Medium | |
| G.722.1-C, G.722.1-E, G.729.1 SWB,G.EV.VBR | high | 14-20kHz | 32 – 64kbps | < 50ms | Medium | |
| G.726,G.729 | Toll | 3.5kHz | < 32kbps | < 20ms | Low | Telecommunication |

**ITU-T**



Fraunhofer Institut Integrierte Schaltungen

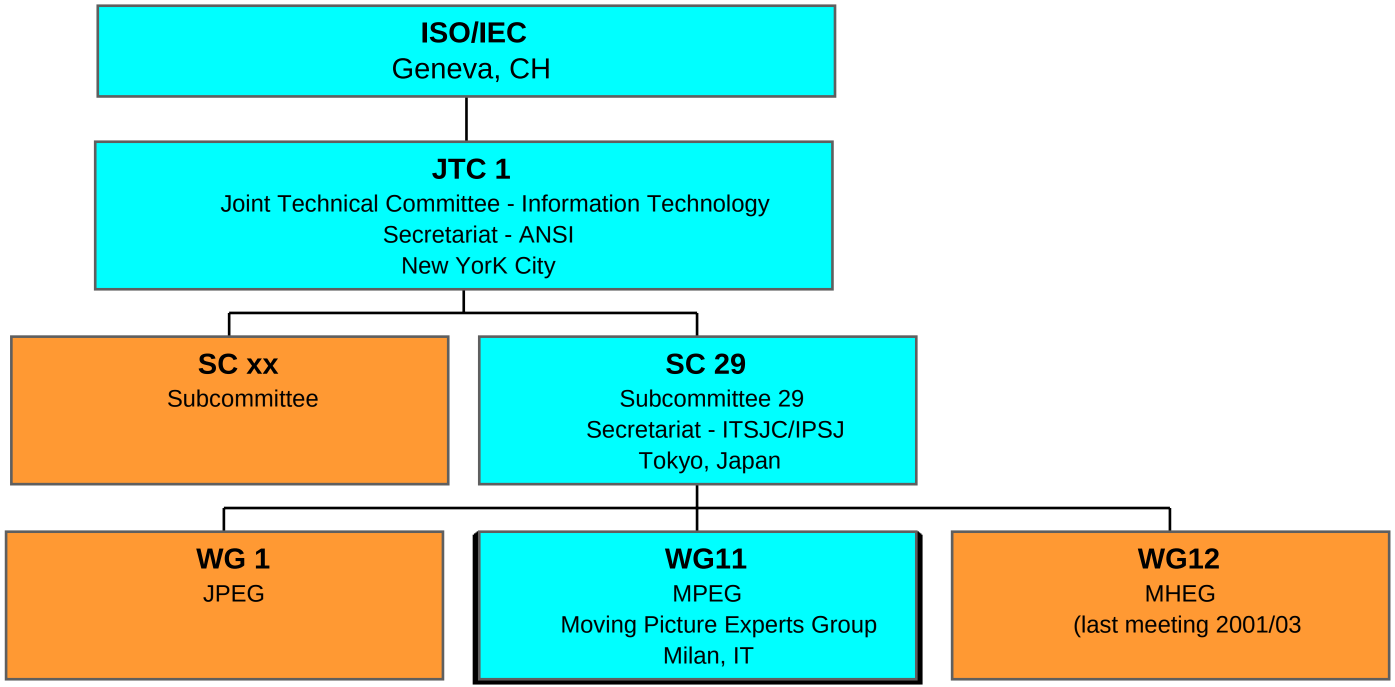
## Overview

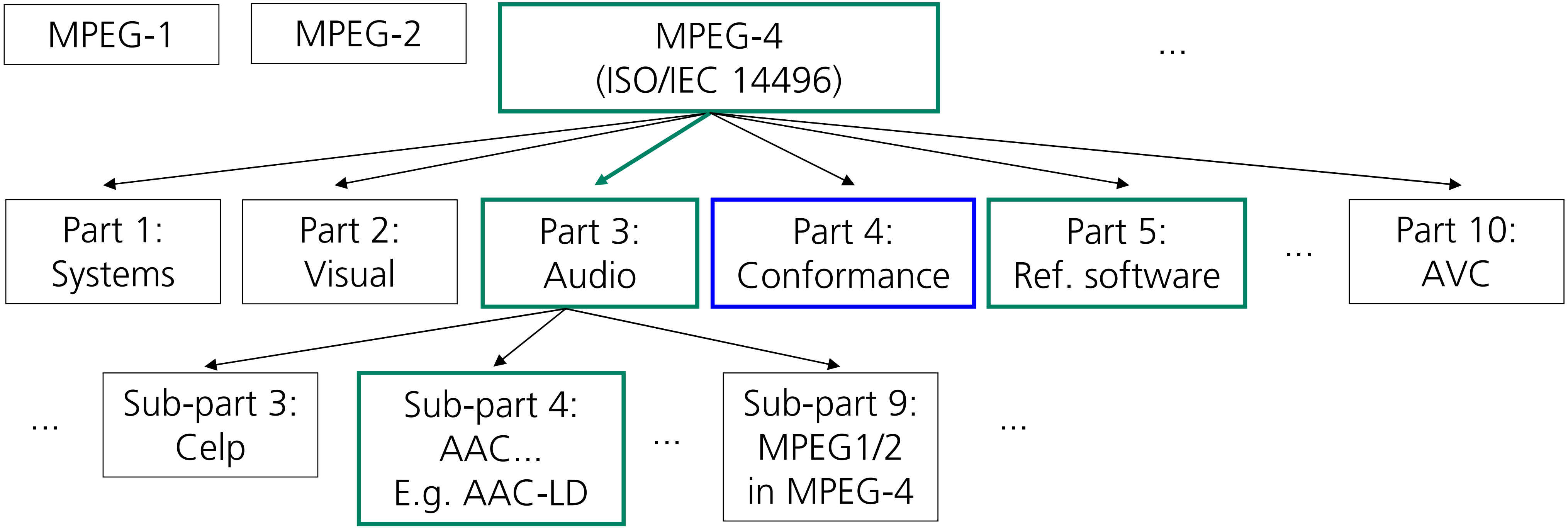
- MPEG introduction
- MPEG-4 low delay AAC (AAC-LD)
- MPEG-4 enhanced low delay AAC (AAC-ELD)
- MPEG Spatial Audio Object Coding (SAOC)



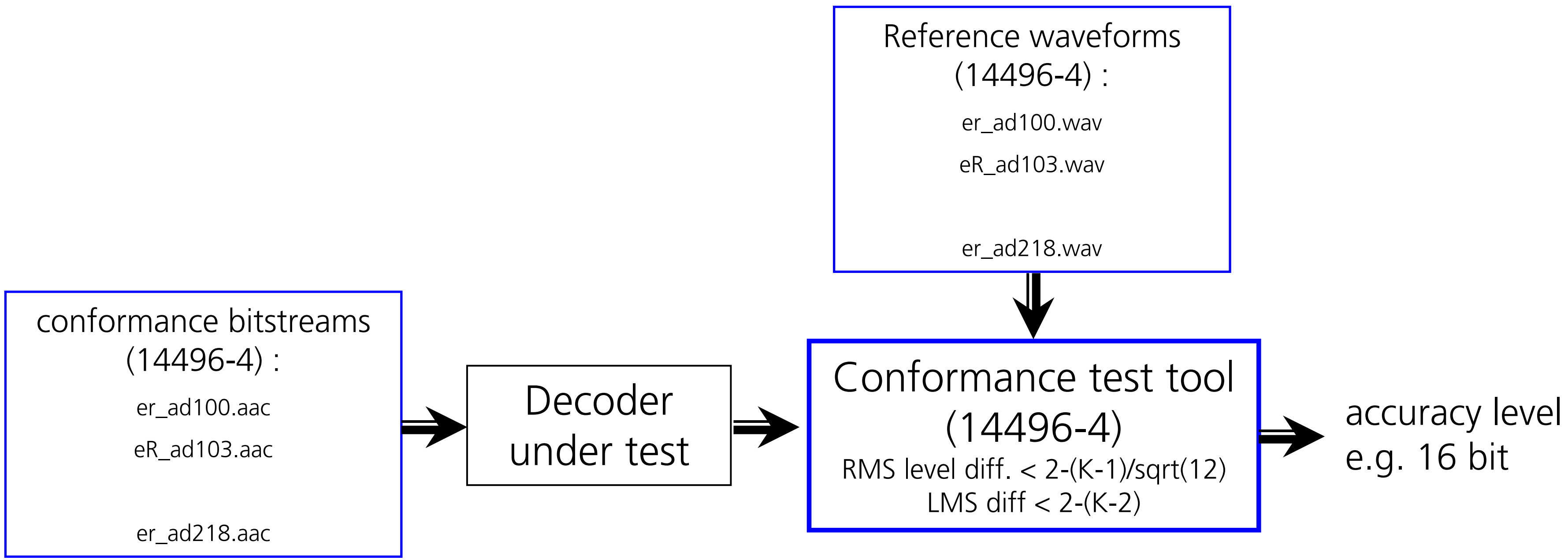
Fraunhofer Institut Integrierte Schaltungen

Moving Picture Experts Group **(MPEG) = ISO/IEC JTC 1/SC 29/WG 11**




MPEG-1
MPEG-2
MPEG-4
(ISO/IEC 14496)
...
Part 1:
Systems
Part 2:
Visual
Part 3:
Audio
Part 4:
Conformance
Part 5:
Ref. software
...
Part 10:
AVC
...
Sub-part 3:
Celp
Sub-part 4:
AAC...
E.g. AAC-LD
...
Sub-part 9:
MPEG1/2
in MPEG-4
...

## MPEG-4 conformance (ISO/IEC 14496-4)



**Specification development time-lines**

No fix duration specified or required

**Exploration**
- 6-12 months depending on extent of search

**Requirements development**
- 6-12 months partly in parallel with exploration

**Competitive phase**
- 3-6 months partly in parallel with requirements

**Collaborative phase**
- 1 year following completion of competitive phase

**-> total approximately 2 years**

## MPEG standardization output

**bitstream syntax description:**

-> Guaranteed interoperability

**decoder behavior defined:**

-> Guaranteed audio quality of given bitsream

**encoder:**

- designed to flexible operation

-> Enables future improvements

-> minimum quality profen



Fraunhofer Institut Integrierte Schaltungen

## AAC family

**Relevant members of MPEG-4 AAC codec family:**

- AAC-LC - high quality codec (iTunes, ISDB…)
- HE-AAC - low bitrate version (XM Radio, 3GPP…)
- SLS – scalable lossless (HD AAC)
- ER AAC-LD delay optimized



Fraunhofer Institut Integrierte Schaltungen

## Facts about AAC-LD

**Status International Standard since 2000**

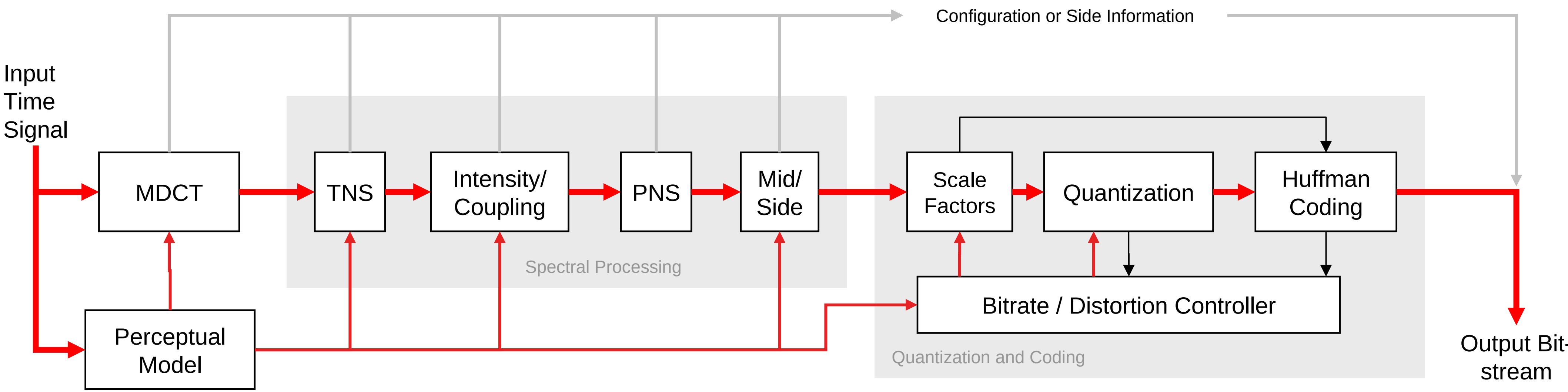
- subpart 4 ISO/IEC 14496-3
- Perceptual audio codec
- Error Resilient bitstream syntax
- flexible configuration:
  - sampling frequency 22.05 – 48 kHz
  - Bitrate typical 32-80 kbps/channel

channel configuration: Mono/ stereo/ multi-channel

**Key Features:**

- Bandwidth up to 16 kHz (and more)
- Algorithmic delay: 20 ms

## Encoder Block diagram




Bitstream Payload Formatter, Bit Reservoir Not Shown

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

Tandberg MXP

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

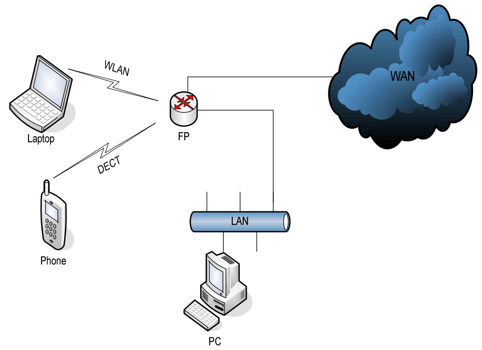
2006: Codian MCU

2006: Comrex Access

## Cisco Telepresence System

audio: 3 channel AAC-LD @ 48kHz
video: 3 x H.264 @ 1080p

## ETSI New Generation DECT standardization

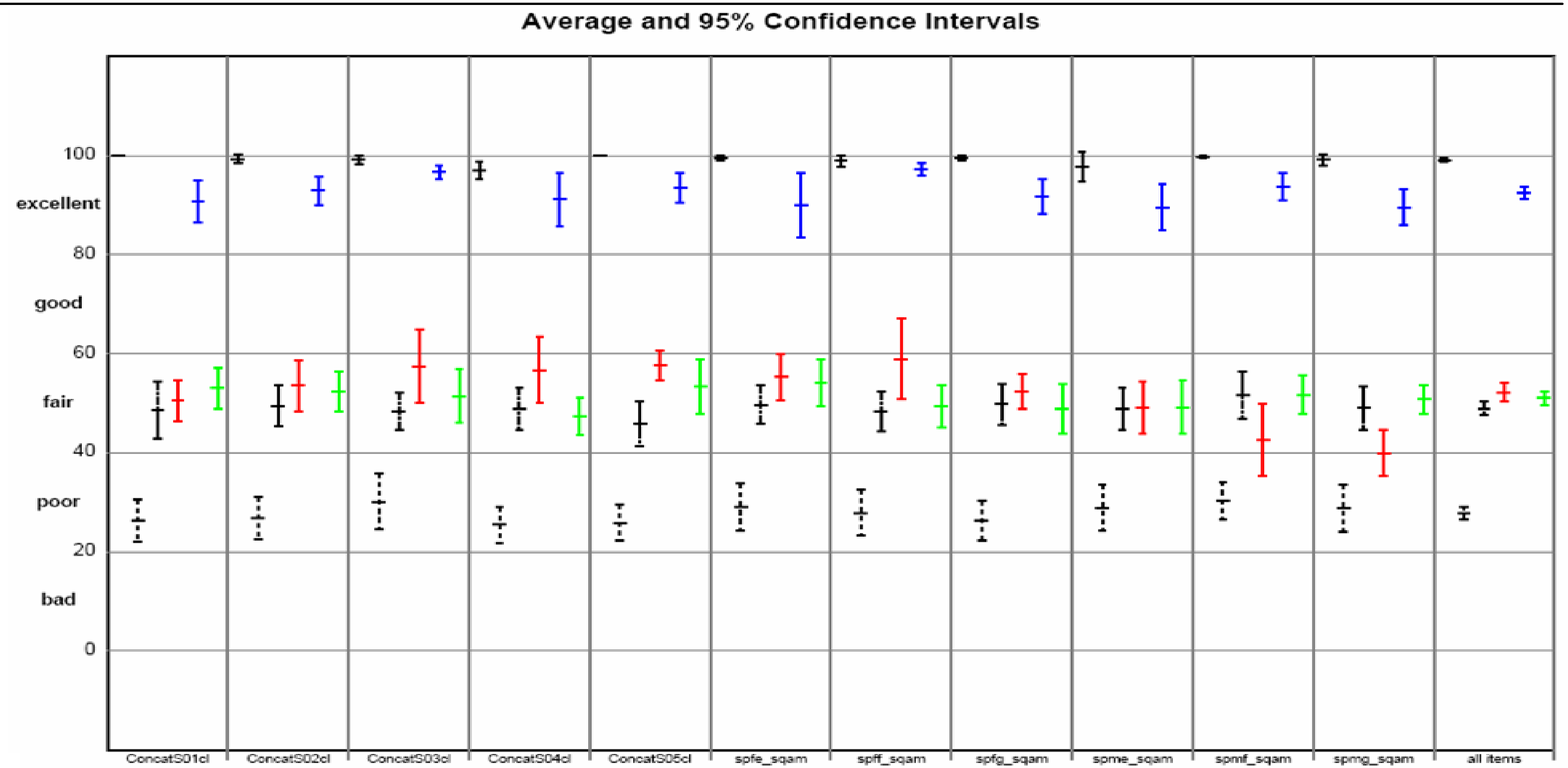



- Part 1 finished 02/2007
  - wide band speech
- New long slot type (64kbps) introduced
- AAC-LD optional codec for super wide band speech
- other codecs:
  - mandatory: G.722, (G.726)
  - optional wb codec: G.729.1

## MUSHRA test on speech items

(12 experienced listeners)

- hidden referece
- anchor 3.5 kHz
- anchor 7.0 kHz
- AAC-LD 32 kbps
- AAC-LD 64 kbps
- G.722 64 kbps




Fraunhofer Institut Integrierte Schaltungen

## Enanced low delay AAC (AAC-ELD)

Status (5/2007): FPDAM

Scheduled International Standard 1/2008

Modifications to AAC-LD:

- Low delay Spectral Bandwidth Replication
- Delay optimized AAC core

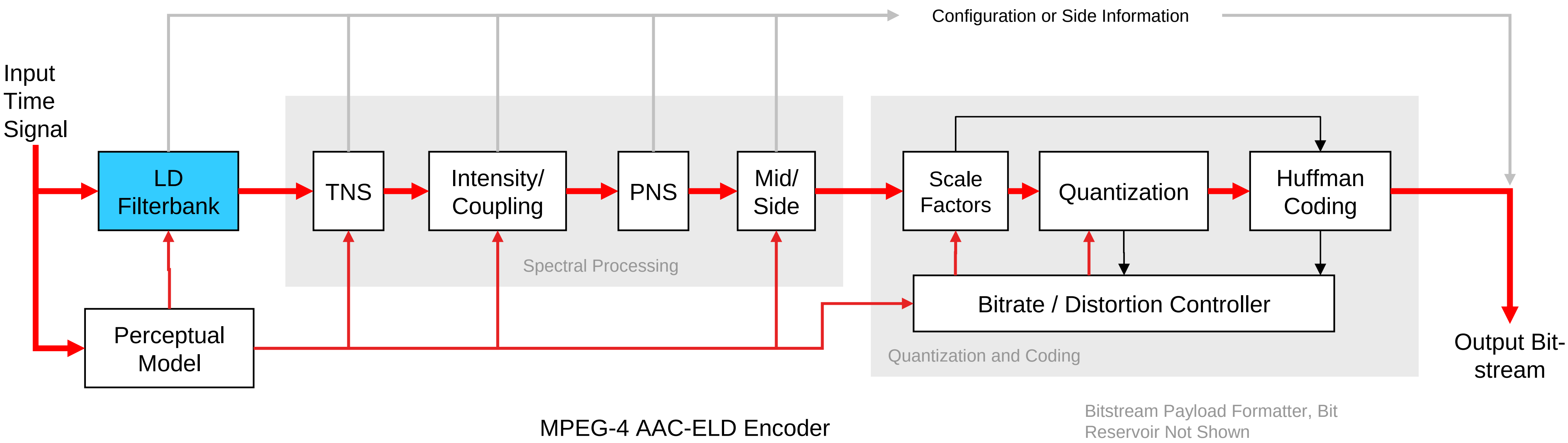
**Key Features:**

- Bandwidth up to 16 kHz (and more)
- Algorithmic delay: 15-32 ms
- Typcial bitrate 24-48 kbps
- Blocklength 20 ms



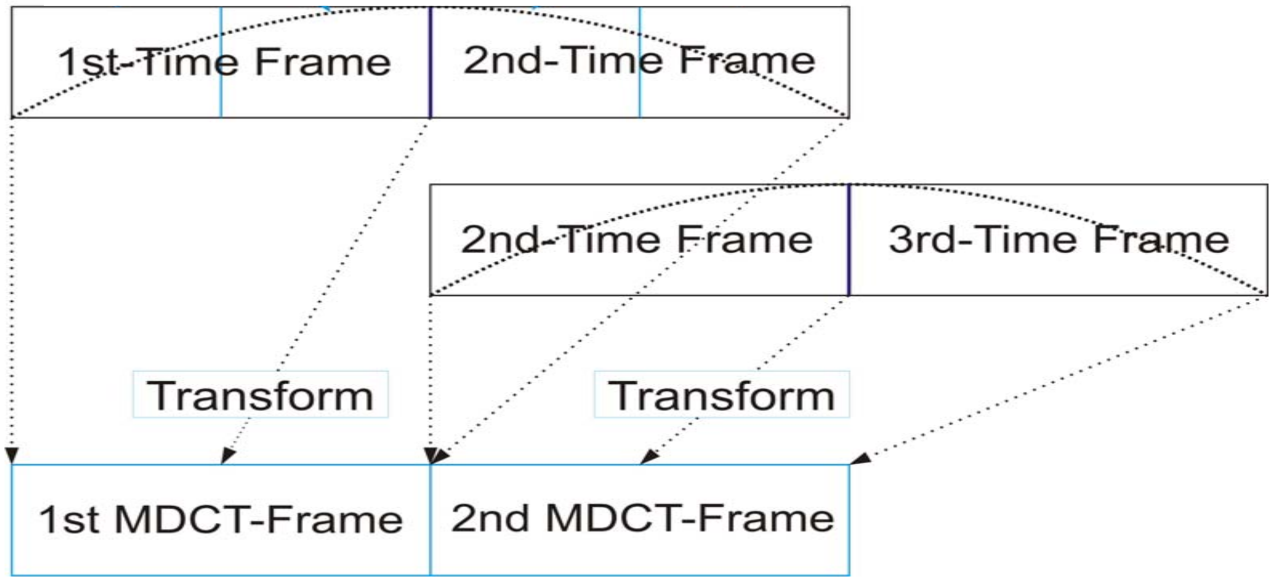
Fraunhofer Institut Integrierte Schaltungen

## Low Delay Filterbank for AAC ELD



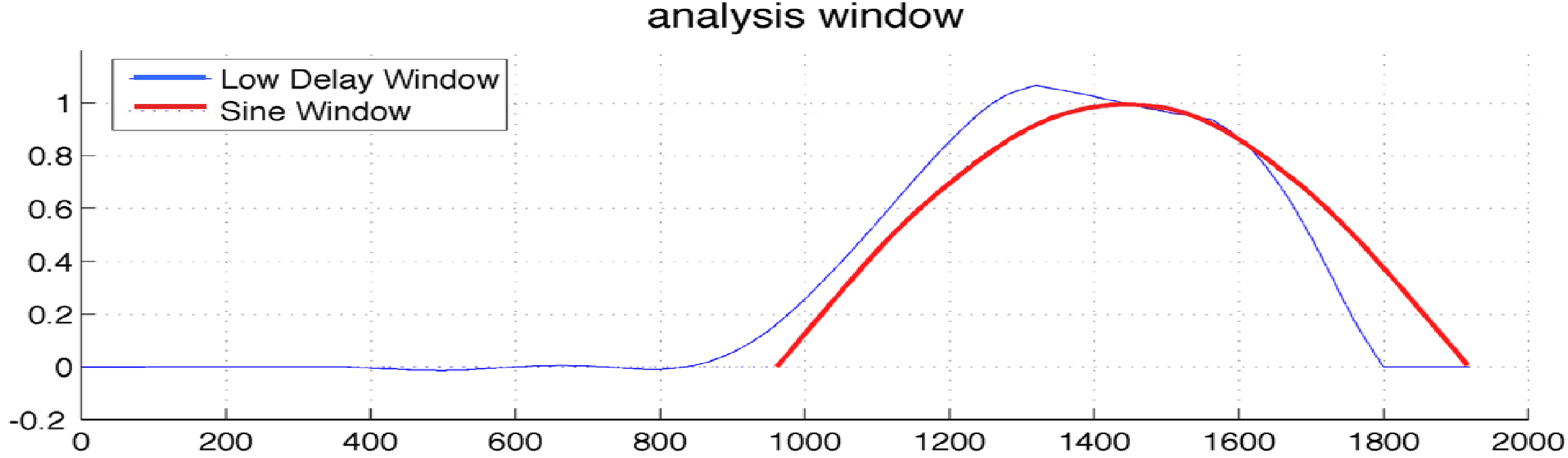

MPEG-4 AAC-ELD Encoder

## Overlap and Add Principle




50% Overlap over past and future samples

## Low Delay Filterbank for AAC-ELD




- Low-delay window reduces delay from 960 samples (MDCT) to 720 samples (for a frame size of 480 samples)
- Low delay filterbank does not increase computational complexity
- Perfect reconstruction and similar frequency response

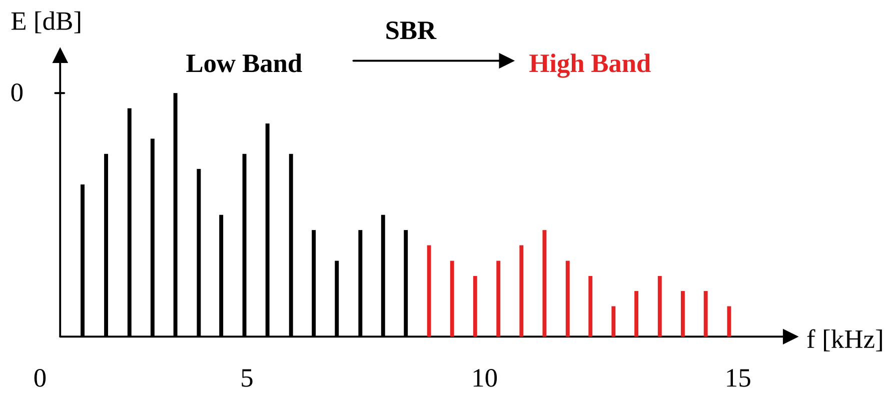
## AAC-ELD Delay Analysis

| Codec | Delay Sources | Delay at 48kHz |
|---|---|---|
| AAC-LD | MDCT+IMDCT | 20 ms |
| AAC-ELD | LD-Filterbank | **15 ms** |



Fraunhofer Institut Integrierte Schaltungen

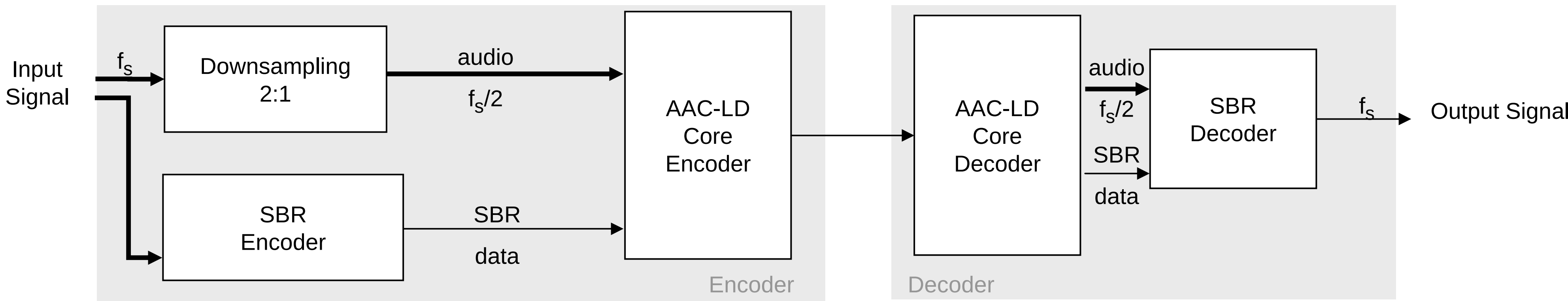
## Low Delay Spectral Bandwidth Replication tool for AAC-LD




### Goal:

- produce good audio quality at bitrates lower than 48kbit/s
- maintaining a reasonable low algorithmic delay

## Low Delay Spectral Band Replication for AAC-LD




delay optimizations SBR tool

## AAC–ELD + LD-SBR Delay Analysis

| Codec | Delay Sources | Delay at 48kHz |
|---|---|---|
| AAC-LD | MDCT+IMDCT | **20 ms** |
| AAC-ELD | LD-Filterbank | **15 ms** |
| AAC-ELD + LD-SBR | LD-Filterbank | 30 ms |
| | QMF -> CLDFB | 12 ms ->1.3ms |
| | **SUM** * updated values 80thMPEG meeting | **32 ms (*)** |



Fraunhofer Institut Integrierte Schaltungen

## Audio quality of AAC-ELD

Listening test results: MPEG2007/M14518, April 2007

MPEG test items: speech, music, single instruments

10 expert listeners

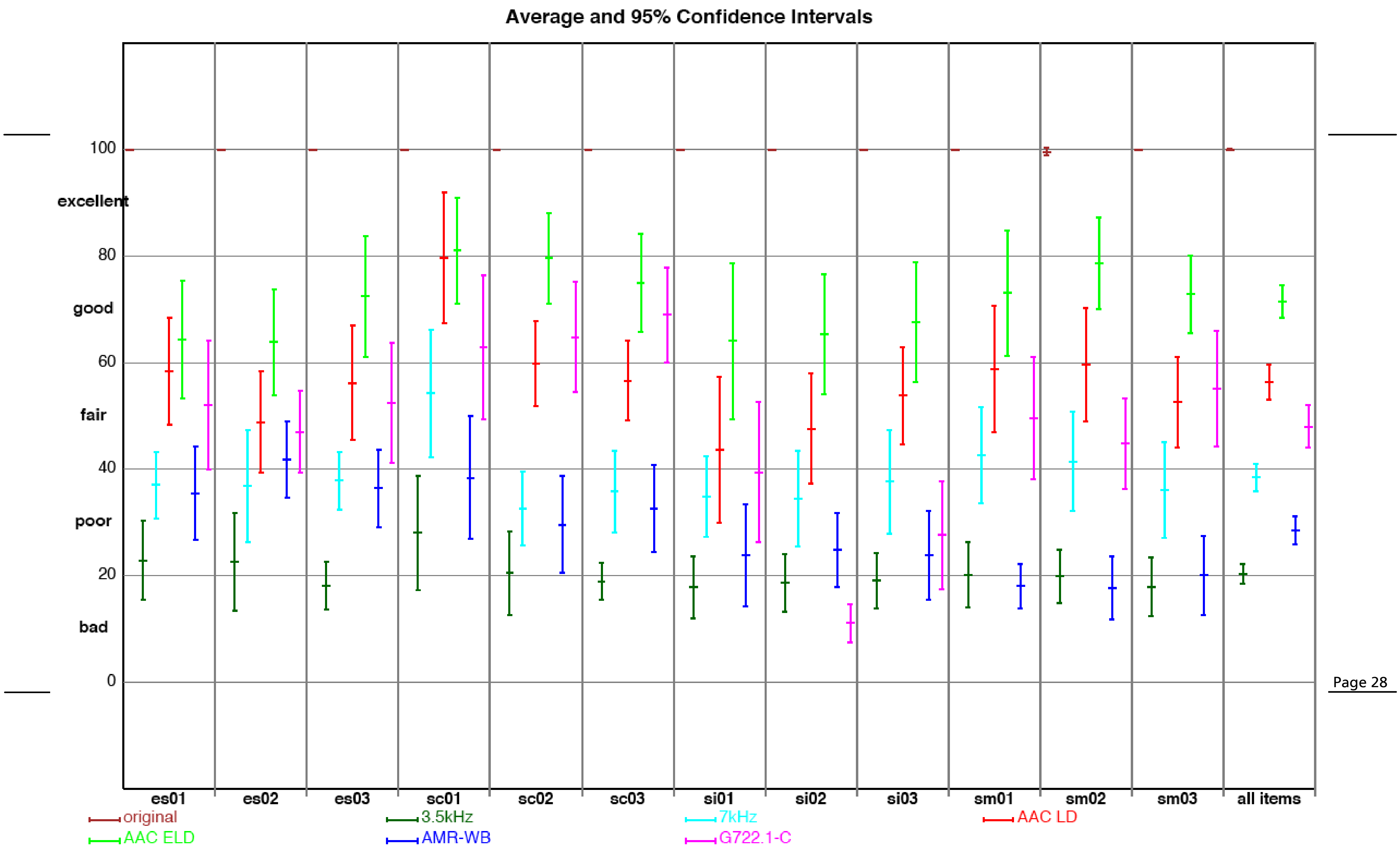
Bitrate: 32 kpbs

Reference codecs:

- MPEG-4 AAC-LD
- ITU-T G.722.1-C
- AMR-WB (G.722.2, @ 24kbps)

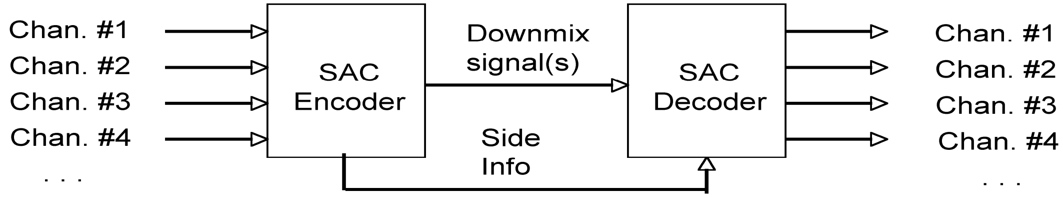


Fraunhofer Institut Integrierte Schaltungen


Average and 95% Confidence Intervals
100
excellent
80
good
60
fair
40
poor
20
bad
0
es01
es02
es03
sc01
sc02
sc03
si01
si02
si03
sm01
sm02
sm03
all items
original
3.5kHz
7kHz
AAC LD
AAC ELD
AMR-WB
G722.1-C

## From Spatial Audio Coding to Spatial Audio Object Coding

Current (“channel-oriented”) Spatial Audio Coding (SAC) – MPEG surround




**MPEG surround**

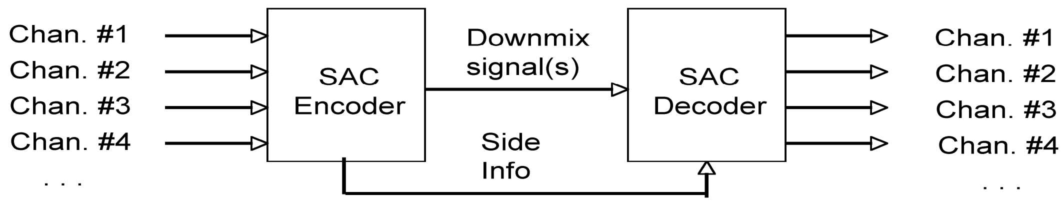
- Mono or stereo downmix

+ spatial parameters

-> multichannel 5.1 down to 48 kbps

## From Spatial Audio Coding to Spatial Audio Object Coding

Current ("channel-oriented") Spatial Audio Coding (SAC) – MPEG surround



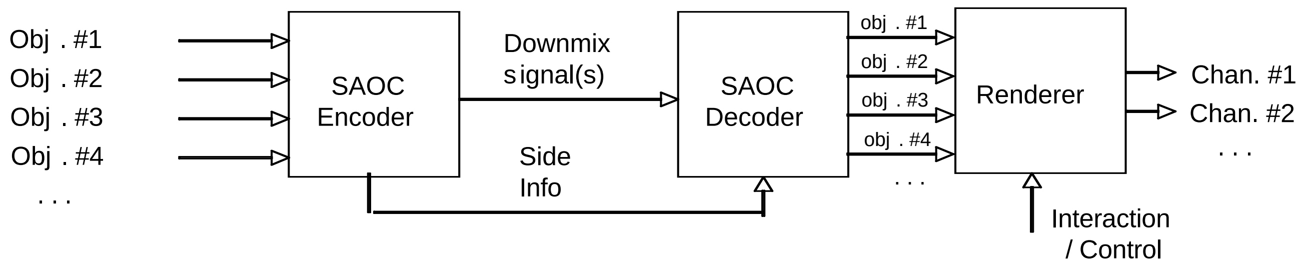

**MPEG surround**

- Mono or stereo downmix

+ spatial parameters

-> multichannel 5.1 down to 48 kbps

Alternative ("object-oriented") Spatial Audio Object Coding (SAOC)




**SAOC**

- based on MPEG surround technology
- Object oriented
- user interaction on renderer

## SAOC - Features

SAOC standardization in MPEG currently CFP

**Features**

- high compression due to mono downmix signal
- backward compatibility

- Interaction on renderer:
  - attenuation of single objects (talkers)
  - Movement of objects in space

- Number of input channels can be different to output channels

**Application**

- multi-participant (Video) conferencing
- games

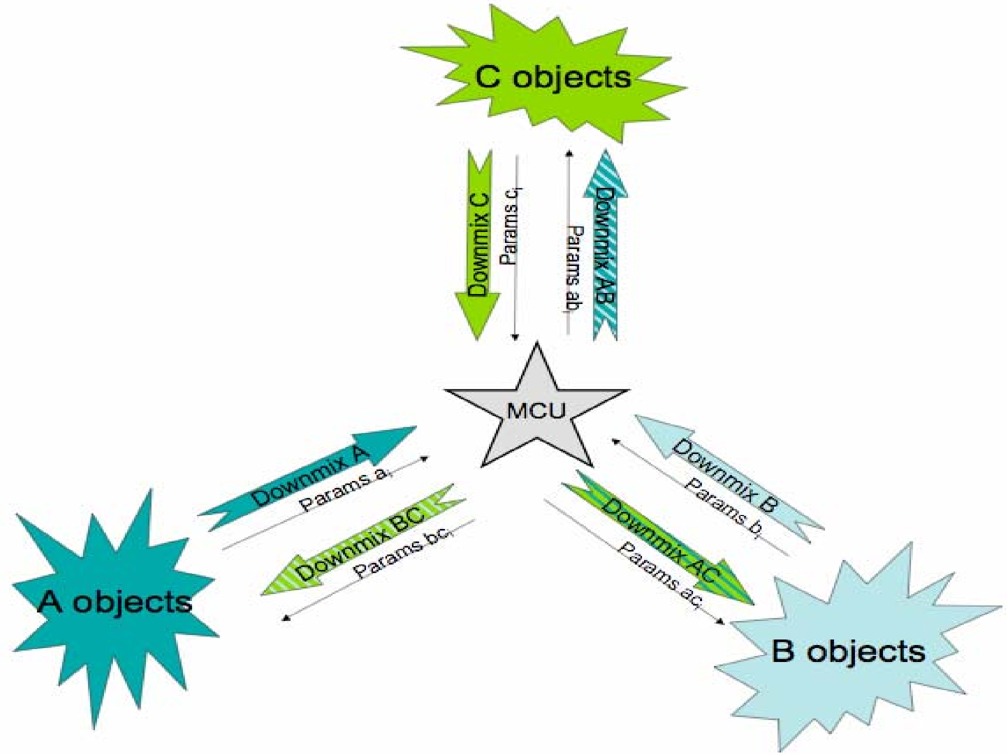
**Spatial Audio Object Coding demonstration**

## MCU: Mixing of SAOC bitstreams

**Scenario:** Teleconferencing with three SAOC stations connected via the MCU

**objects can be:**
- Single talker ->mono bitstream
- Multiple talkers -> downmix + params

**Free of delay Mixing with in MCU**:
- spatial parameters
- AAC-ELD coded downmix bitstreams






Fraunhofer Institut Integrierte Schaltungen

Thank you for your attention!