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

Minds in Motion

The Heritage of the QMF-Transform in aacPlus and MPEG Surround

Reducing Complexity by Exploring Synergies

Thomas Ziegler

coding technologies

Overview

- Introduction
- aacPlus Recap
- The MPEG Surround Standard
  - Technology
  - Operating Modes
- aacPlus with MPEG Surround Support
  - Synergy Effects
  - Applications
- Audio Demo
- Conclusions

Introduction

- aacPlus – State-of-the-Art audio codec
  - CD-near stereo quality at 24 kbps
  - MPEG standard since 2003
  - Adopted by lots of other standardization bodies
  - Paradigm shift in high quality audio doding
- MPEG surround follows the new paradigm
  - MPEG standardization finalized end 2006
  - Common proposal:
    FhG, Agere, Philips and Coding Technologies

Overview

- Introduction
- aacPlus Recap => see TIDC-2006
- The MPEG Surround Standard
  - Technology
  - Operating Modes
- aacPlus with MPEG Surround Support
  - Synergy Effects
  - Applications
- Audio Demo
- Conclusions

aacPlus

- MPEG Standardization
  - 2003: aacPlus-v1, MPEG-4 Amd 1
  - 2004: aacPlus-v2, MPEG-4 Amd 2
  - 2005: 3rd MPEG-4 edition

- Adopted by
  - Digital Radio Mondiale (DRM), DAB
  - 3GPP, 3GPP2
  - ...

- Widely used in digital broadcast, mobile phones, …
- Presented in detail at TIDC06
- => aacPlus should be well known

aacPlus

- Paradigm Shift: Hybrid approach
  - Use traditional waveform codec only for psychoacoustic difficult parts
  - Exploit (co)relation between difficult parts & the rest
  - Parameterize (co)relation, transmit as side info
  - Guided reconstruction at the decoder side

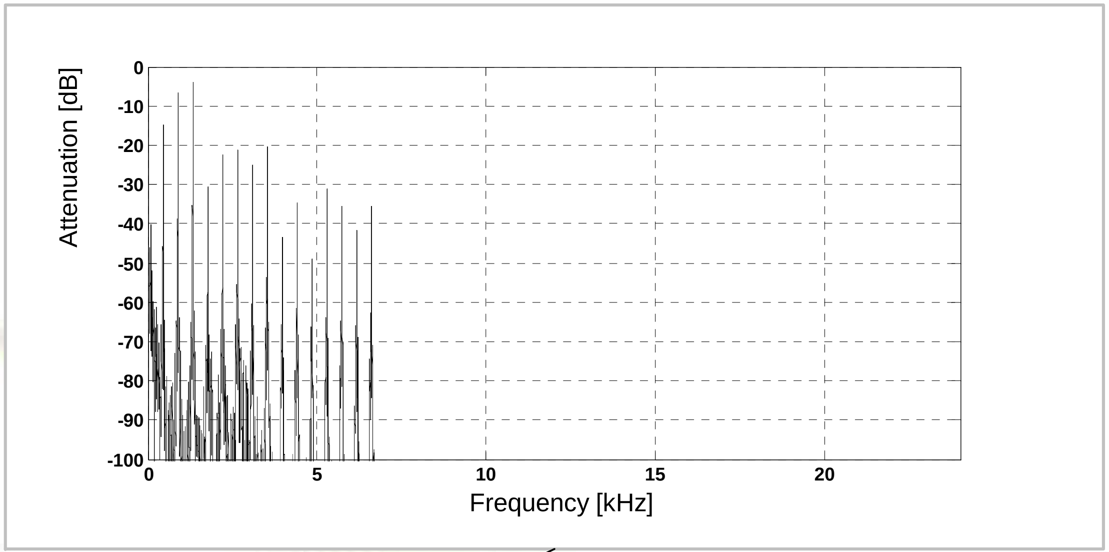
- Parametric Tools
  - Spectral Band Replication (SBR)
  - Parametric Stereo (PS)

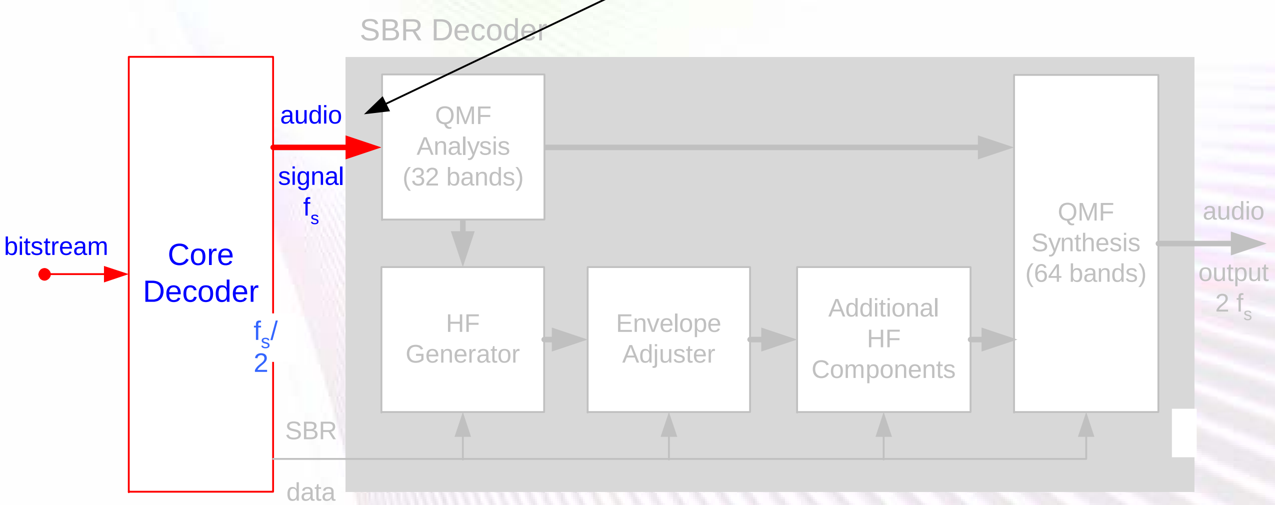
Spectral Band Replication (SBR)

- Exploit correlation between low band and high band
- Transmit only low band with a conventional codec
- Reconstruct the missing high band in a perceptually accurate manner
- "Guided" reconstruction using small amount of guidance data in the bit stream (1-3 kbps)

SBR Decoder

- Bit stream demux
  - Core bit stream
  - SBR data
- Core decoder output is band limited
- Core coder operates at ½ nominal sampling rate f_s.
- Frame size 1024 samples

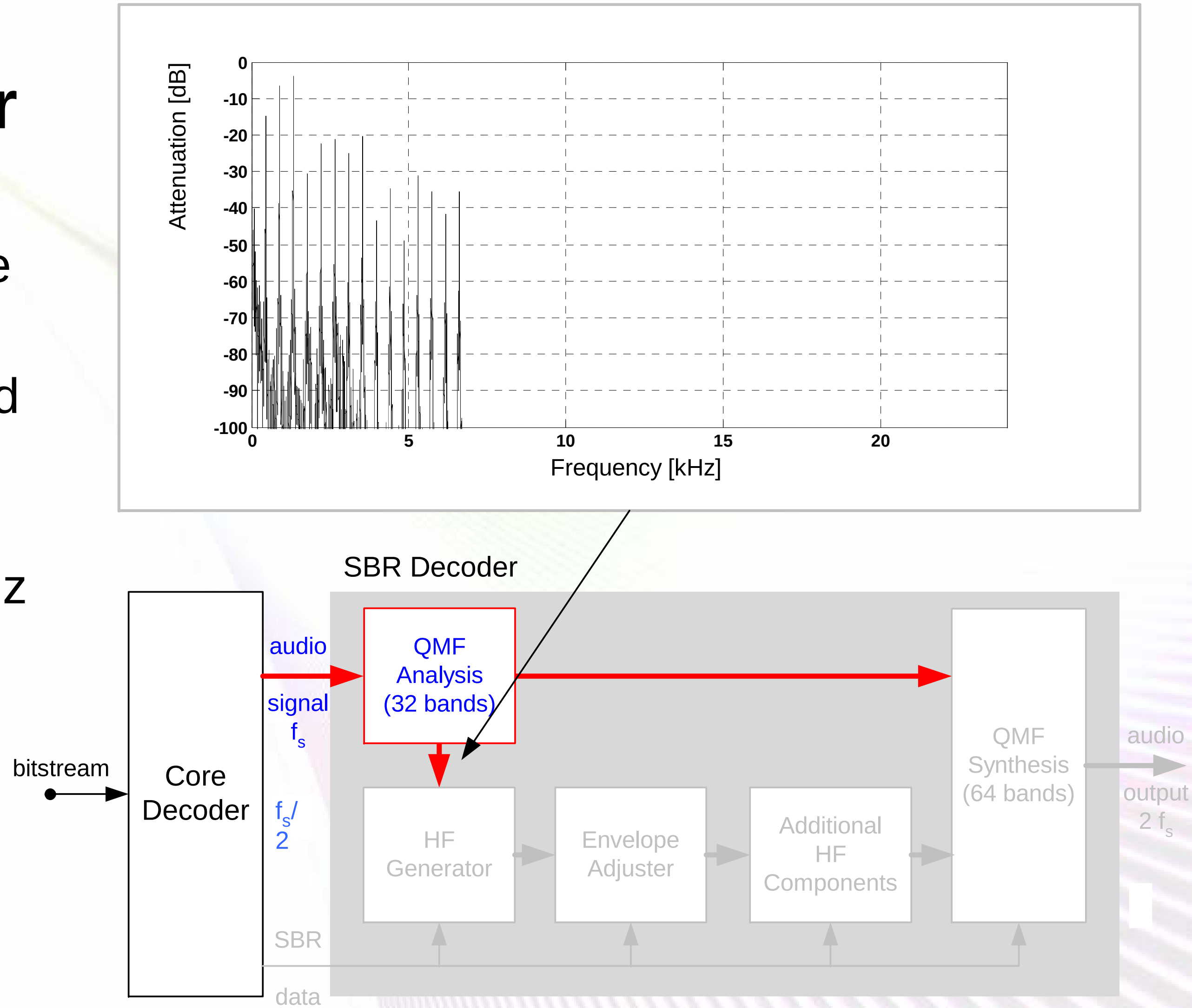




Only one channel is depicted, second channel equivalent

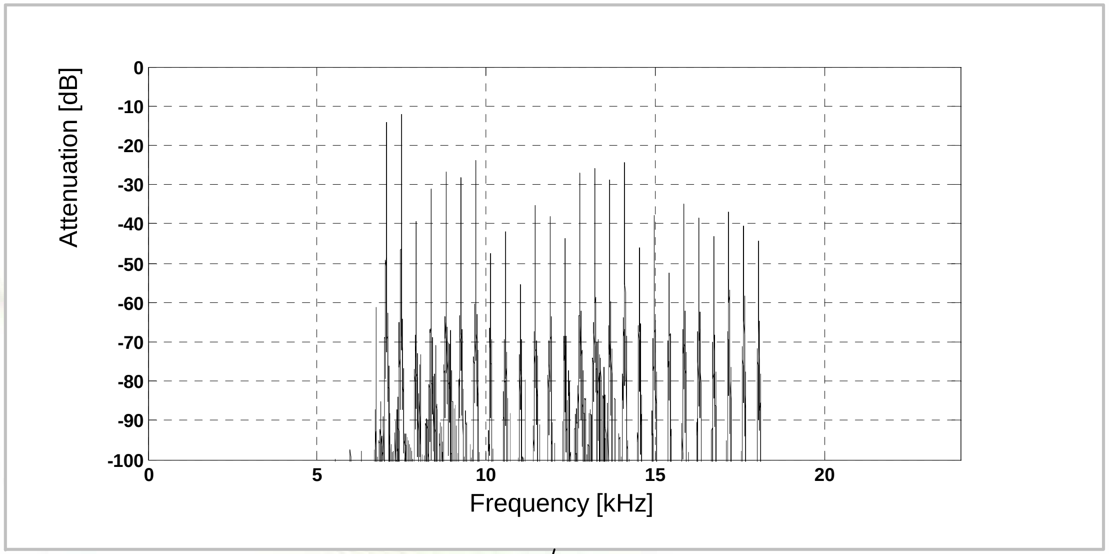
SBR Decoder

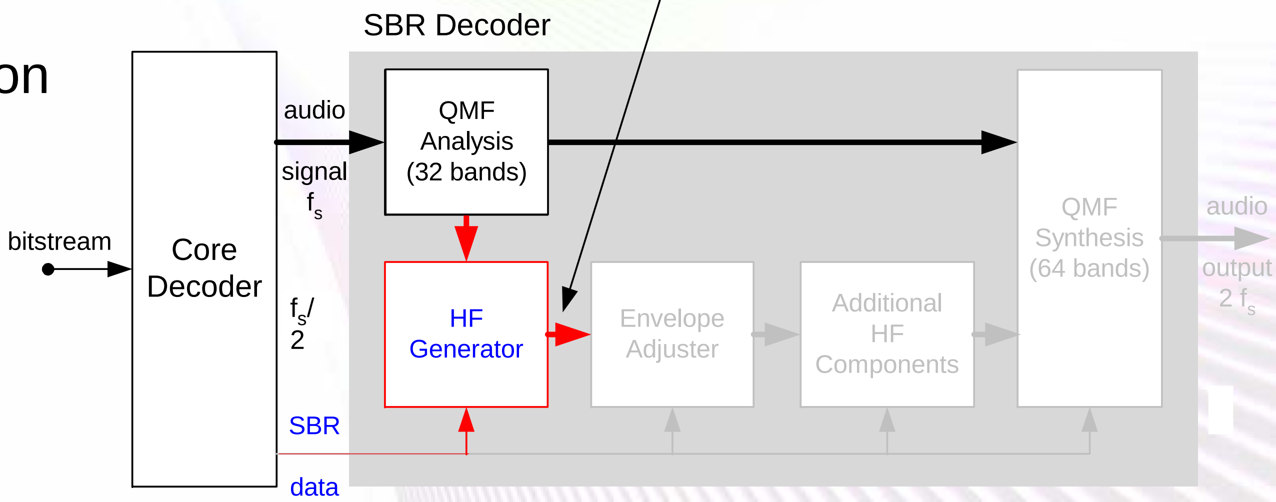
- Complex QMF enables alias-free processing
- 32 equally spaced bands
- Bandwidth of QMF-band 345 Hz @ 44.1 kHz



SBR Decoder

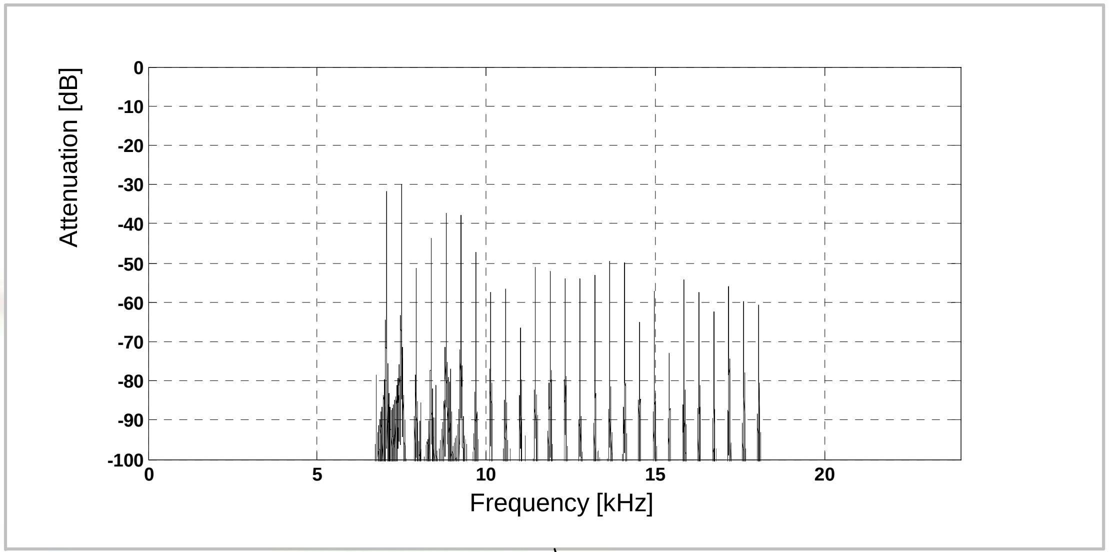
- Generation of high band by transposition
- Apply inverse filtering tool if necessary
  - Autocorrelation
  - LPC-Filtering

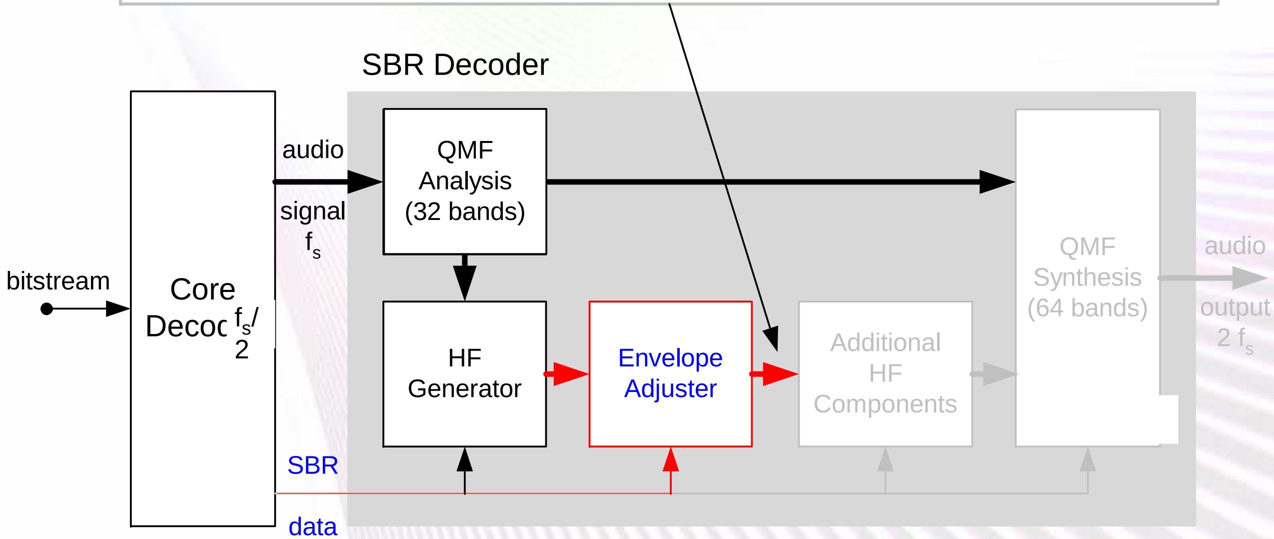




SBR Decoder

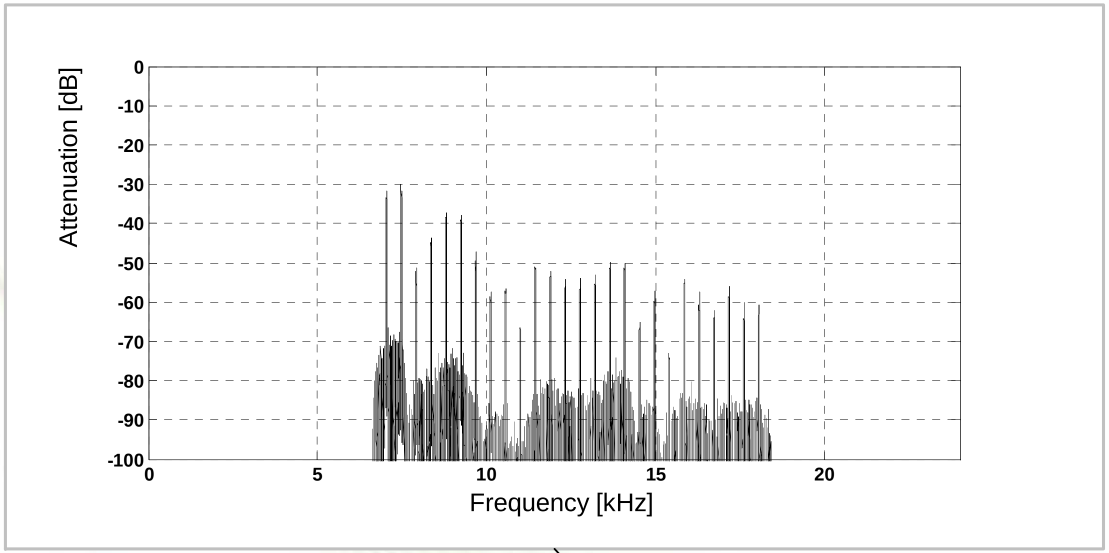
- Calculate energy of transposed signal
- Compare with transmitted envelope
- Calculate gain factors and adjust spectral shape to original
- Up to five envelopes per AAC-frame

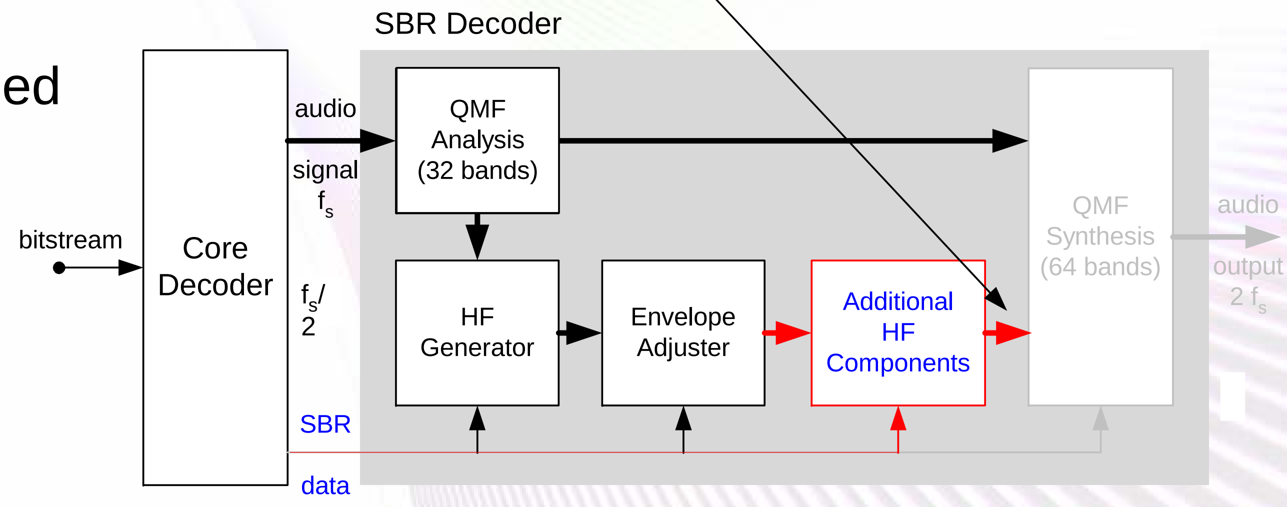




SBR Decoder

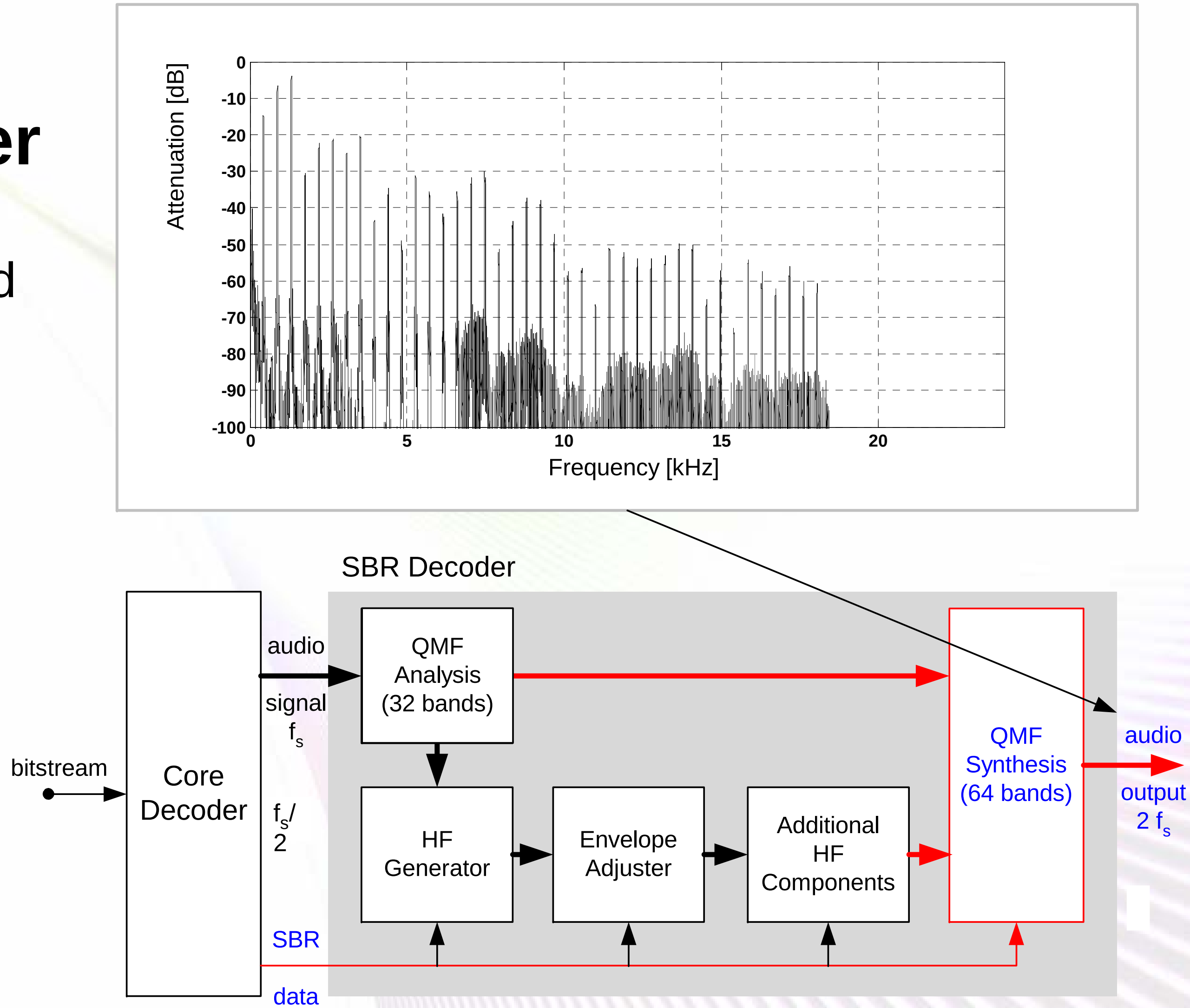
- Adaptive noise addition unit
  - Random noise injection
- Sinusoidal Regenerator
- Both tools controlled by SBR-data





SBR Decoder

- Combine low and high band
- Perform implicit upsampling
- Frame size 2048 samples

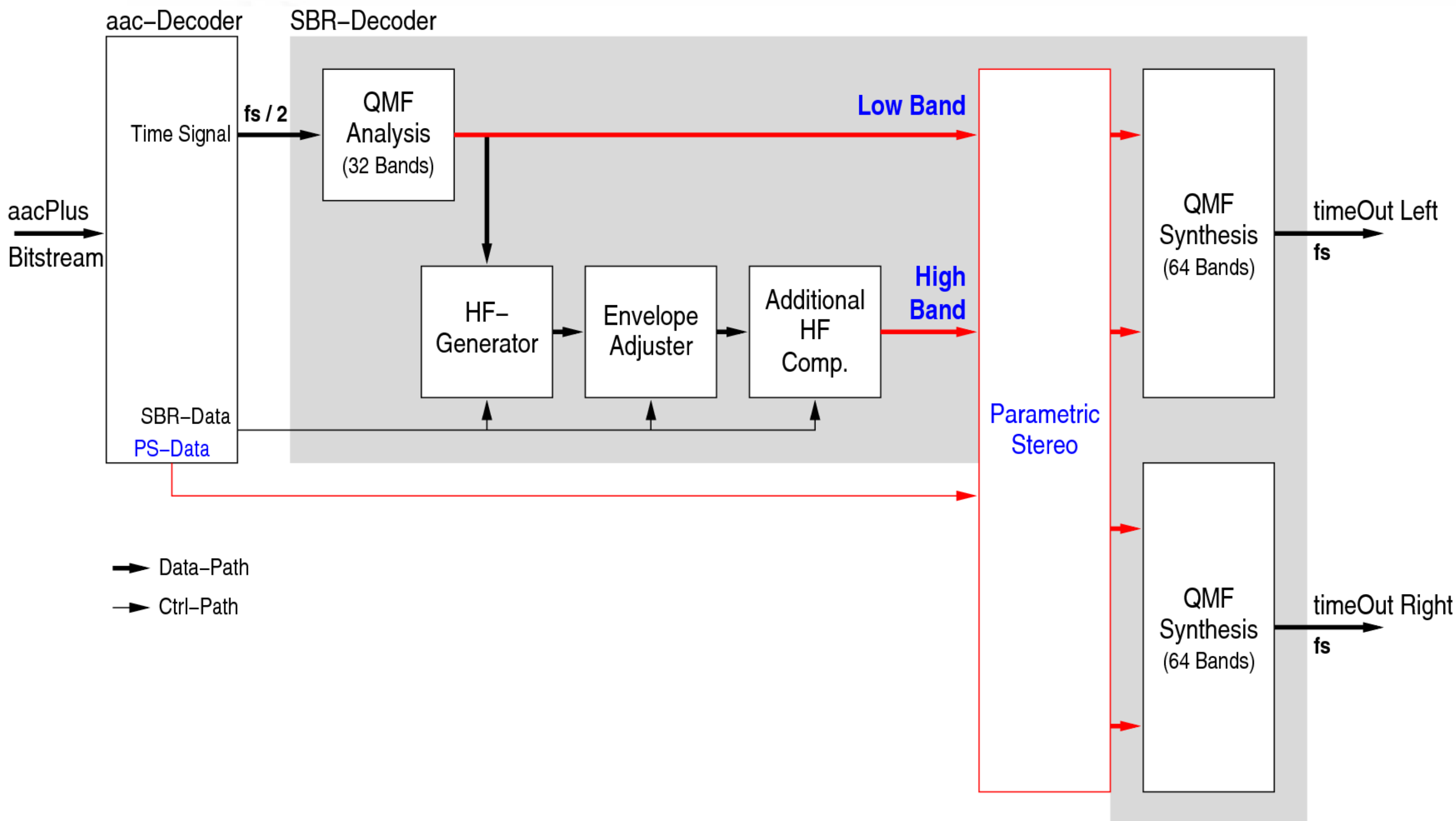


Parametric Stereo (PS)

- Exploit high correlation between right and left channel
- Transmit only one channel: left + right (mid) signal
- “Guided Reconstruction” of stereo signal using small amount of guidance data (3-5 kbps)
  - Inter-channel Intensity Differences (IID)
  - Inter-channel Cross Correlation (ICC)

- PS operates in the complex QMF-domain
  => Reuse of SBR filter bank

PS-Decoder

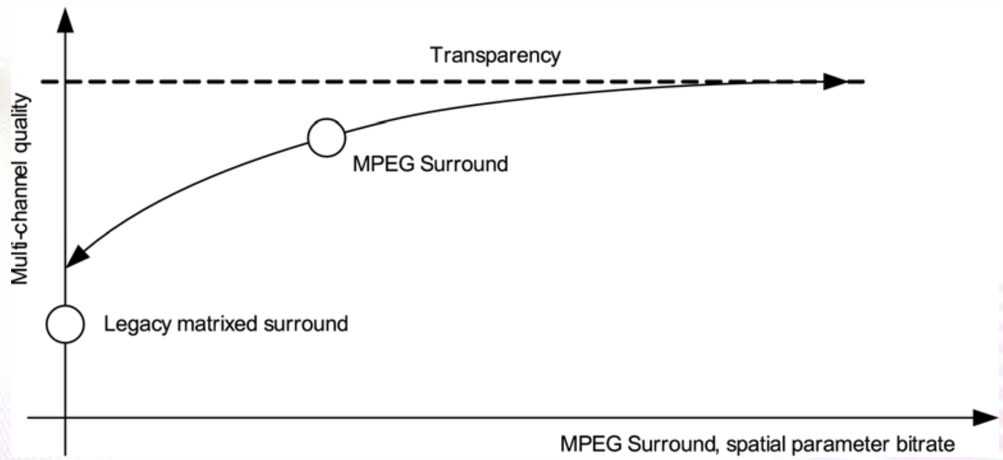

aac–Decoder
SBR–Decoder
Time Signal
fs / 2
QMF Analysis (32 Bands)
Low Band
aacPlus Bitstream
HF–Generator
Envelope Adjuster
Additional HF Comp.
High Band
SBR–Data
PS–Data
Parametric Stereo
QMF Synthesis (64 Bands)
timeOut Left
fs
QMF Synthesis (64 Bands)
timeOut Right
fs
Data–Path
Ctrl–Path

Overview

- Introduction
- aacPlus
- The MPEG Surround Standard
  - Technology
  - Operating Modes
- aacPlus with MPEG Surround Support
  - Synergy Effects
  - Applications
- Audio Demo
- Conclusions

MPEG Surround (MPS)

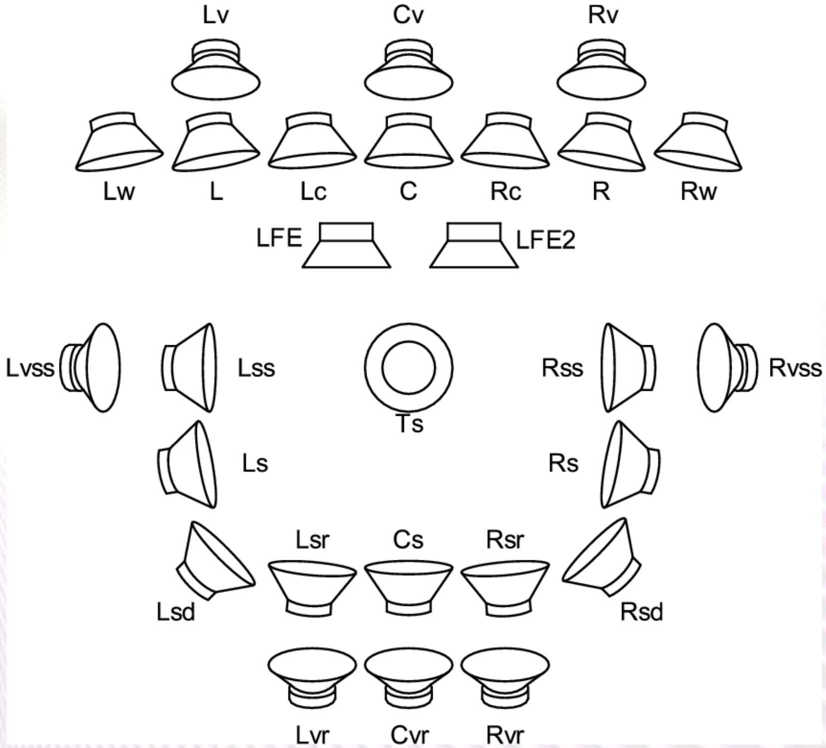
- Follows the PS paradigm
- Transmit 2 channels, reconstruct 5 channels
- Toolbox concept
  - Full Quality-vs-Bitrate scalability
  - Low bit rate solutions
  - Transparency



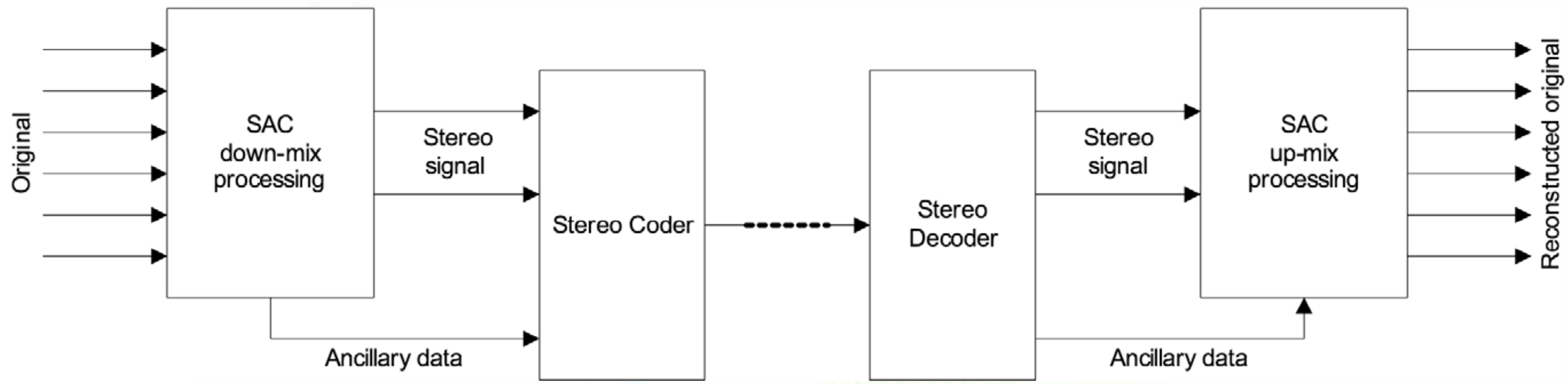
- Design is independent from core codec

MPS: Speaker Positions

- Reconstruct N channels, transmit only M<=N channels
- Maximum: 27 Channels
- Speaker positions

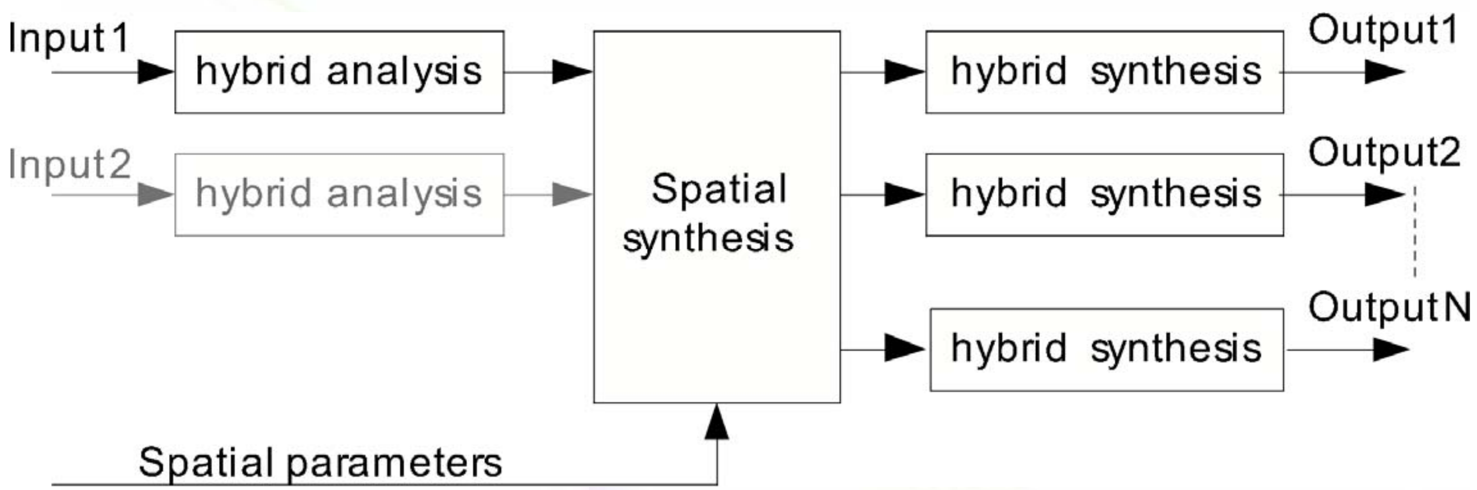


MPS: General Structure



- Downmix to stereo
- Extract parametric description of spatial image (spatial cues)
- Backwards compatible bitstream embedding
- Up-mix process guided by spatial cues

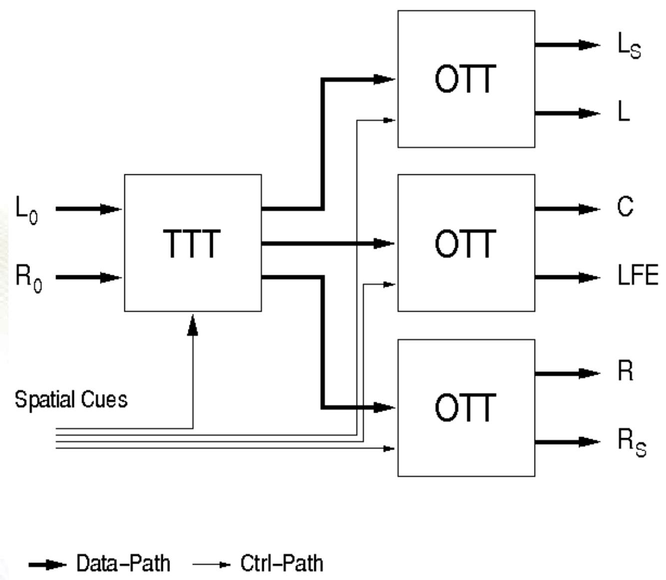
MPS: Spatial Up-mix



- QMF-based hybrid filterbank
- Perceptual accurate, non-uniform frequency resolution
- Identical to Parametric Stereo hybrid filterbank
- Input 1...M channels, output N>=M channels

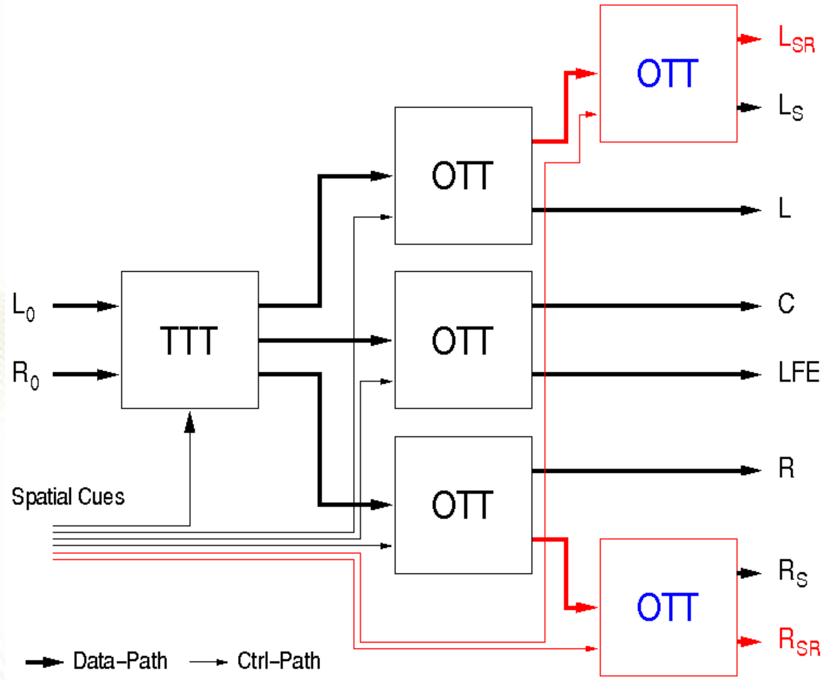
MPS: Spatial Synthesis

- Tree-like organization
  - One-to-Two (OTT)
  - Two-to-Three (TTT)



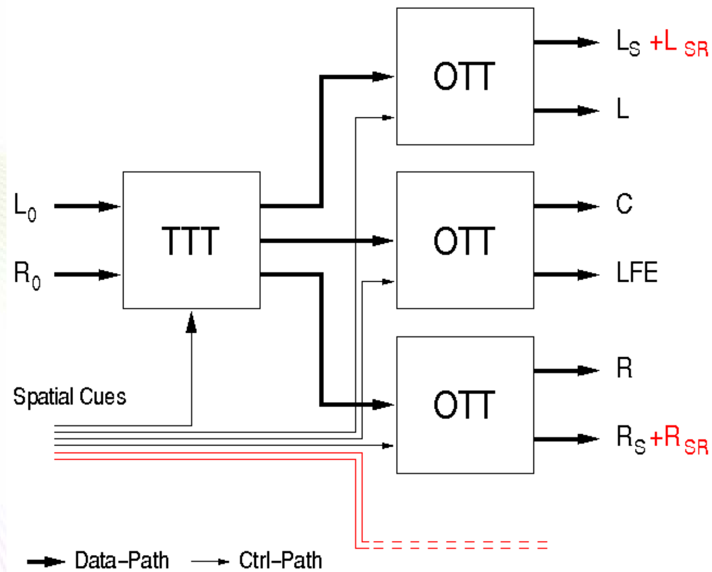
MPS: Spatial Synthesis

- Tree-like Organization
  - One-to-Two (OTT)
  - Two-to-Three (TTT)
- Allows arbitrary combination of N channels based on M<=N input channels

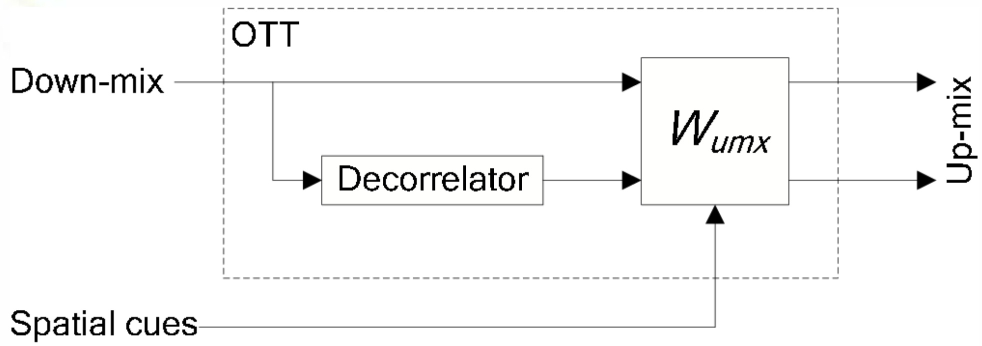


MPS: Spatial Synthesis

- Tree-like Organization
  - One-to-Two (OTT)
  - Two-to-Three (TTT)
- Allows arbitrary combination of N channels based on M<=N input channels
- Tailor up-mix process to receiver environment



MPS: OTT-Element



- Generates two output channels based on one input channel
- Spatial parameters
  - Channel Level Differences (CLD)
  - Inter-channel Coherence/Cross-correlation (ICC)
- Up-mix controlled by matrix W_{umx}
- Corresponding OTT-Elements at encoder side perform down-mix while extracting spatial cues

MPS: TTT-Element

- Generates three output channels based on two input channels
- Estimation of third channel can alternatively be controlled by Channel Prediction Coefficients (CPC)
- An additional ICC parameter compensates for prediction loss due to only partially correlated original signals

MPS: Temporal Shaping Tools

- Guided Envelope Shaping (GES)
  - Shaping controlled by encoder
- Subband Temporal Processing (STP)
  - Shaping based on decoder data
  - Encoder only activates/deactivates

MPS: Quality-vs-Bitrate Scalability

- Bit rate control of parametric side info controls quality of reconstruction
- Frequency resolution
  - Grouping of spectral bands
- Time resolution
  - Update rate of spatial cues
- Quantization
  - Coarse/Fine resolution
  - Adaptive Parameter Smoothing in time direction
- Residual Coding

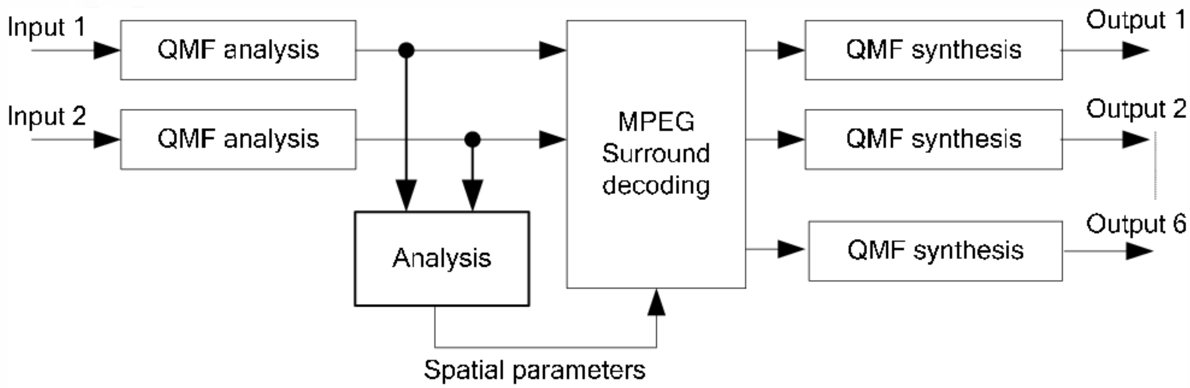
MPS: Residual Coding

- Required to reach transparency
- Compensates for the limitations of parametric coding
- Encoder
  - Transmit error signal after down-mix
  - Simplified AAC waveform encoding
- Decoder
  - MDCT-to-QMF transform
  - Decorrelator output (=synthetic residual) replaced by true residual signal
- Simplifies design of hierarchical systems

Overview

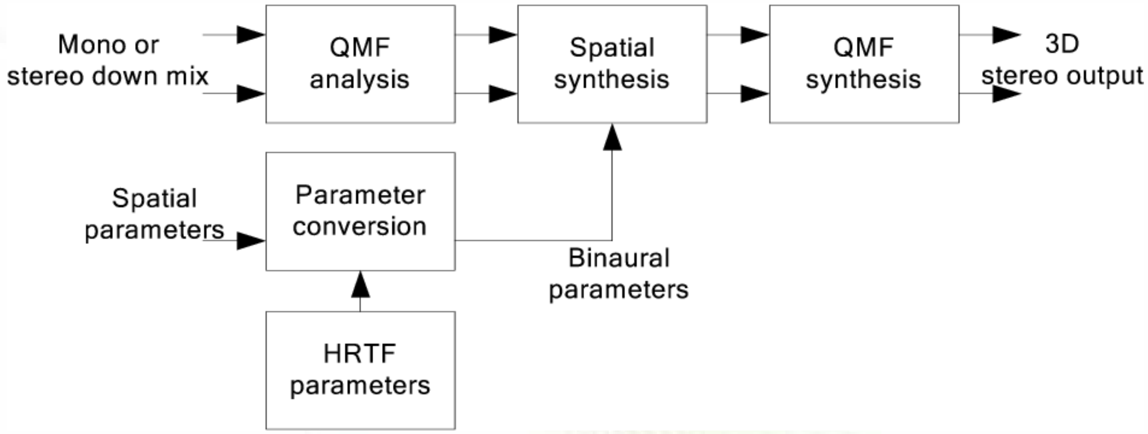
- Introduction
- aacPlus
- The MPEG Surround Standard
  - Technology
  - Operating Modes
- aacPlus with MPEG Surround Support
  - Synergy Effects
  - Applications
- Conclusions

MPS: Enhanced Matrix Mode



- MPEG Surround sound rendering engine is generic
- Parameter estimation based on analysis of down-mixed stereo signal
- Lookup table based estimation => Training
- Quality significantly superior to legacy matrixed-surround sound systems (Source MPEG Tests)

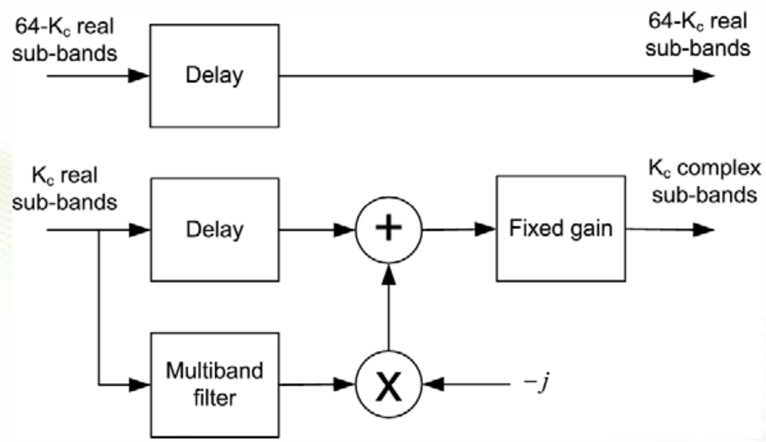
MPS: Binaural Decoding



- Binaural rendering
  => out-of-head localization of sound sources
- Common: Apply HRTFs to 5.1 time domain signal
- HRTF processing in the parameter domain saves computational complexity
- Binaural decoding is a pure MPS-decoder feature
- Might be combined with Enhanced Matrix Mode

MPS: Low Power Decoding

- Partial complex processing
- Real-to-complex converter
  - Lower 8 hybrid bands
  - 11-Tap FIR



- Aliasing reduction tool
- Limited or no residual decoding
- Shorter decorrelators; PS-decorrelators
- Low Power mode is a pure decoder mode

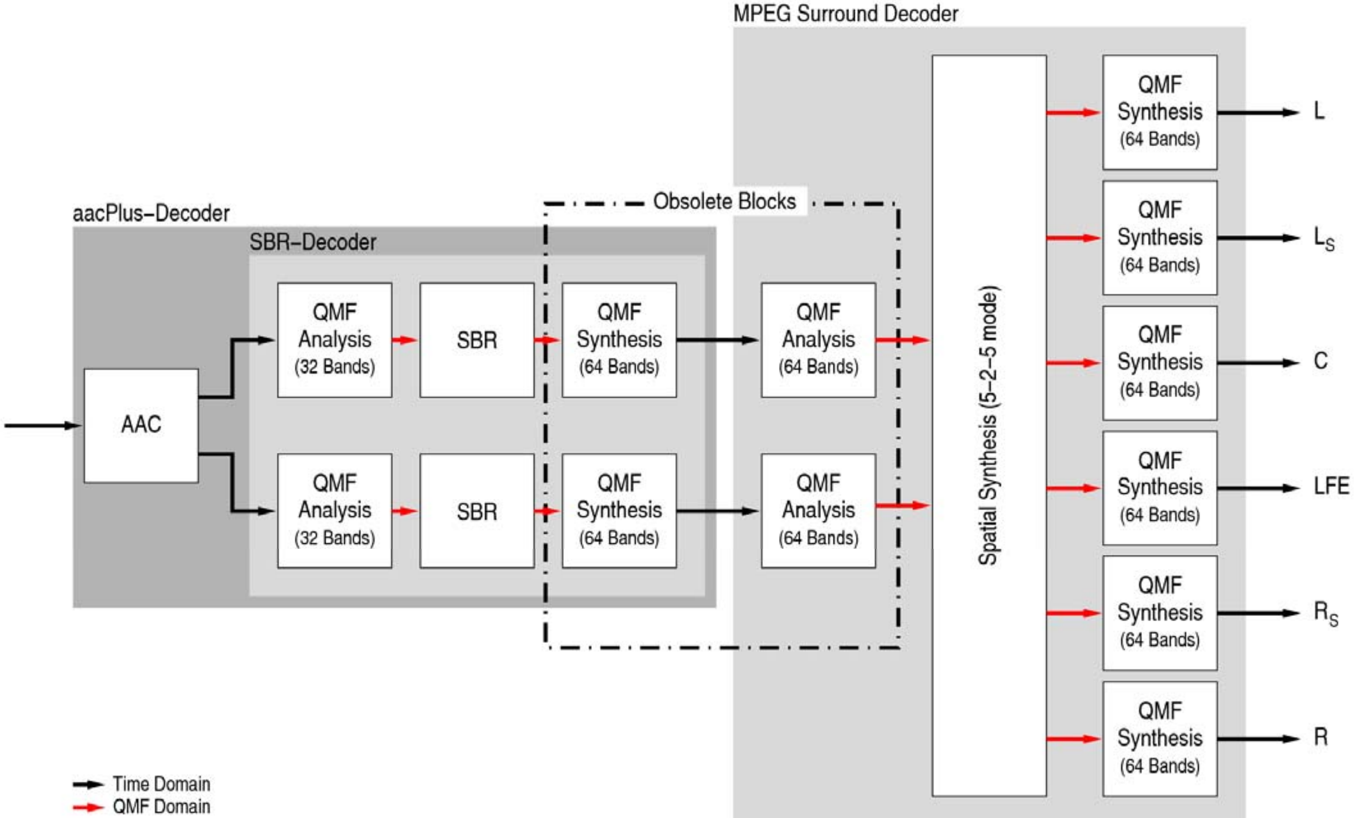
Overview

- Introduction
- aacPlus
- The MPEG Surround Standard
  - Technology
  - Operating Modes
- aacPlus with MPEG Surround Support
  - Synergy Effects
  - Applications
- Audio Demo
- Conclusions

aacPlus with MPEG Surround Support

- Ideal combination
- Coding efficiency
- Quality-vs-total bitrate scalability
- Code sharing
  - QMF-based hybrid filterbank
  - PS-decorrelators
  - Residual decoder
- Synergies reduce computational complexity

aacPlus with MPEG Surround Support



- Synthesis-Analysis steps are obsolete
- Structure similar to Parametric Stereo processing
- Less complex than AAC with MPEG Surround Support

Implementation Challenges

- Number of output channels
  - Increases computational complexity and memory footprint
  - Increased working set may not fit into common cache sizes
- Decorrelators
  - Lattice structure (standardized) or direct form IIRs
  - Cascaded 2nd order IIRs (PS decorrelators)
- Matrix multiplication
  - Large dimension
- Real-to-Complex converters
  - 11-Tap FIR

aacPlus with MPS: Expected Requirements

- 5.1-Channel mode
  - Computational complexity in the same range as discrete aacPlus-5.1
  - Memory requirements in the same range as discrete aacPlus-5.1
- Binaural
  - Depends on complexity of underlying HRTF
  - Computational complexity compared to regular stereo + 20%…50%
  - Memory requirements compared to regular stereo up to 2x
- Please attend the presentation for detailed profiling

aacPlus with MPS: State of Development

- Fixed-point Firmware Reference Code
  - 1st prototype: now
  - Product: Q1 2007
- Embedded SDKs
  - Directly derived from reference code
  - C64, C55, C6722, others, ...
  - Available Q1/Q2 2007

MPEG Surround Applications

- Digital Audio Broadcasting
  - WorldDMB specified
    - aacPlus (efficiency)
    - MPS (optional multi channel extension)
  - Automotive
    - true multi channel
    - Enhanced Matrix Mode for regular FM stations
  - Kitchen radio: stereo compatible
  - Portable device: binaural

MPEG Surround Applications

- Digital Video Broadcasting
  - STB
    - true multi channel
    - hierarchical coding
    - Binaural via blue tooth headset
  - DVB-H
    - Built-in mini speakers: stereo compatible
    - Headphones: binaural

MPEG Surround Applications

- Music Download
  - Upgrade existing (stereo) infrastructure
  - Bitrate comparable low
    => no degradation of stereo compatible part
  - Mobile phones: binaural
  - PC / Home stereo: true multi channel

Overview

- Introduction
- aacPlus
- The MPEG Surround Standard
  - Technology
  - Operating Modes
- aacPlus with MPEG Surround Support
  - Synergy Effects
  - Applications
- Audio Demo
- Conclusions

Audio Demo

- aacPlus with MPEG Surround, 5-2-5
  - Total bit rate: 64 kbps
  - Pop / Classic
- Binaural via headphones
  - Total bit rate: 64 kbps
  - On request after the presentation

Conclusions

- MPEG Surround: generic audio rendering engine
- Combination of aacPlus & MPEG Surround
  - Perfect candidate for
    - Upgrading current stereo systems to surround sound
    - Surround sound systems via speakers
    - Binaural surround sound via head phones
  - First embedded SDKs available now
- Future application scenarios on the horizon
  - Object oriented scene description
  - Gaming, Video conferencing

References

[1] Thomas Ziegler, Michael Beer, and Gustavo Hoffmann. aacPlus – High Efficiency Audio Coding for Broadcast and Mobile Applications. In *TI Developer Conference, Dallas, Feb. 28 - March 2, Birmingham, March 30, Munich, April 4*, 2006.

[2] Thomas Ziegler. aacPlus – Theory and Practice behind the Successful High Efficiency Audio Codec. In *IEEE Workshop on Multimedia Compression, Bangalore, Oct 27 - 28*, 2005.

[3] Lars Villemoes, Jürgen Herre, Jeroen Breebaart, Gerard Hotho, Sascha Disch, Heiko Purnhagen, and Kristofer Kjörling. MPEG Surround: The Forthcoming ISO Standard for Spatial Audio Coding. In *28th International AES Conference, The Future of Audio Technology – Surround and Beyond, Piteå, Sweden, June 30-July 2*, 2006.

[4] Jeroen Breebaart, Jürgen Herre, Lars Villemoes, Craig Jin, , Kristofer Kjörling, Jan Plogisties, and Jeronen Koppens. Multi-Channels goes Mobile: MPEG Surround Binaural Rendering. In *29th International AES Conference, Audio for Mobile and Handheld Devices, Seoul, Sept 2-4*, 2006.

[5] Thomas Ziegler, Alexander Gröschel, and Andreas Ehret. Extending aacPlus with MPEG Surround Support. In *Global Signal Processing Conference (GSPx), Santa Clara, Oct. 30 - Nov 2*, 2006.

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

Minds in Motion

The Heritage of the QMF-Transform in aacPlus and MPEG Surround

Reducing Complexity by Exploring Synergies

Thomas Ziegler
Zie@CodingTechnologies.com
http://www.CodingTechnologies.com

coding technologies

IMPORTANT NOTICE

Texas Instruments Incorporated and its subsidiaries (TI) reserve the right to make corrections, modifications, enhancements, improvements, and other changes to its products and services at any time and to discontinue any product or service without notice. Customers should obtain the latest relevant information before placing orders and should verify that such information is current and complete. All products are sold subject to TI's terms and conditions of sale supplied at the time of order acknowledgment.

TI warrants performance of its hardware products to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are used to the extent TI deems necessary to support this warranty. Except where mandated by government requirements, testing of all parameters of each product is not necessarily performed.

TI assumes no liability for applications assistance or customer product design. Customers are responsible for their products and applications using TI components. To minimize the risks associated with customer products and applications, customers should provide adequate design and operating safeguards.

TI does not warrant or represent that any license, either express or implied, is granted under any TI patent right, copyright, mask work right, or other TI intellectual property right relating to any combination, machine, or process in which TI products or services are used. Information published by TI regarding third-party products or services does not constitute a license from TI to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property of the third party, or a license from TI under the patents or other intellectual property of TI.

Reproduction of information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations, and notices. Reproduction of this information with alteration is an unfair and deceptive business practice. TI is not responsible or liable for such altered documentation.

Resale of TI products or services with statements different from or beyond the parameters stated by TI for that product or service voids all express and any implied warranties for the associated TI product or service and is an unfair and deceptive business practice. TI is not responsible or liable for any such statements.

Following are URLs where you can obtain information on other Texas Instruments products and application solutions:

Products		**Applications**	
Amplifiers	amplifier.ti.com	Audio	www.ti.com/audio
Data Converters	dataconverter.ti.com	Automotive	www.ti.com/automotive
DSP	dsp.ti.com	Broadband	www.ti.com/broadband
Interface	interface.ti.com	Digital Control	www.ti.com/digitalcontrol
Logic	logic.ti.com	Military	www.ti.com/military
Power Mgmt	power.ti.com	Optical Networking	www.ti.com/opticalnetwork
Microcontrollers	microcontroller.ti.com	Security	www.ti.com/security
Low Power Wireless	www.ti.com/lpw	Telephony	www.ti.com/telephony
		Video & Imaging	www.ti.com/video
		Wireless	www.ti.com/wireless

Mailing Address: Texas Instruments
Post Office Box 655303 Dallas, Texas 75265