MPEG-4 General Audio Coding

Jürgen Herre
Fraunhofer Institute for Integrated Circuits (IIS)

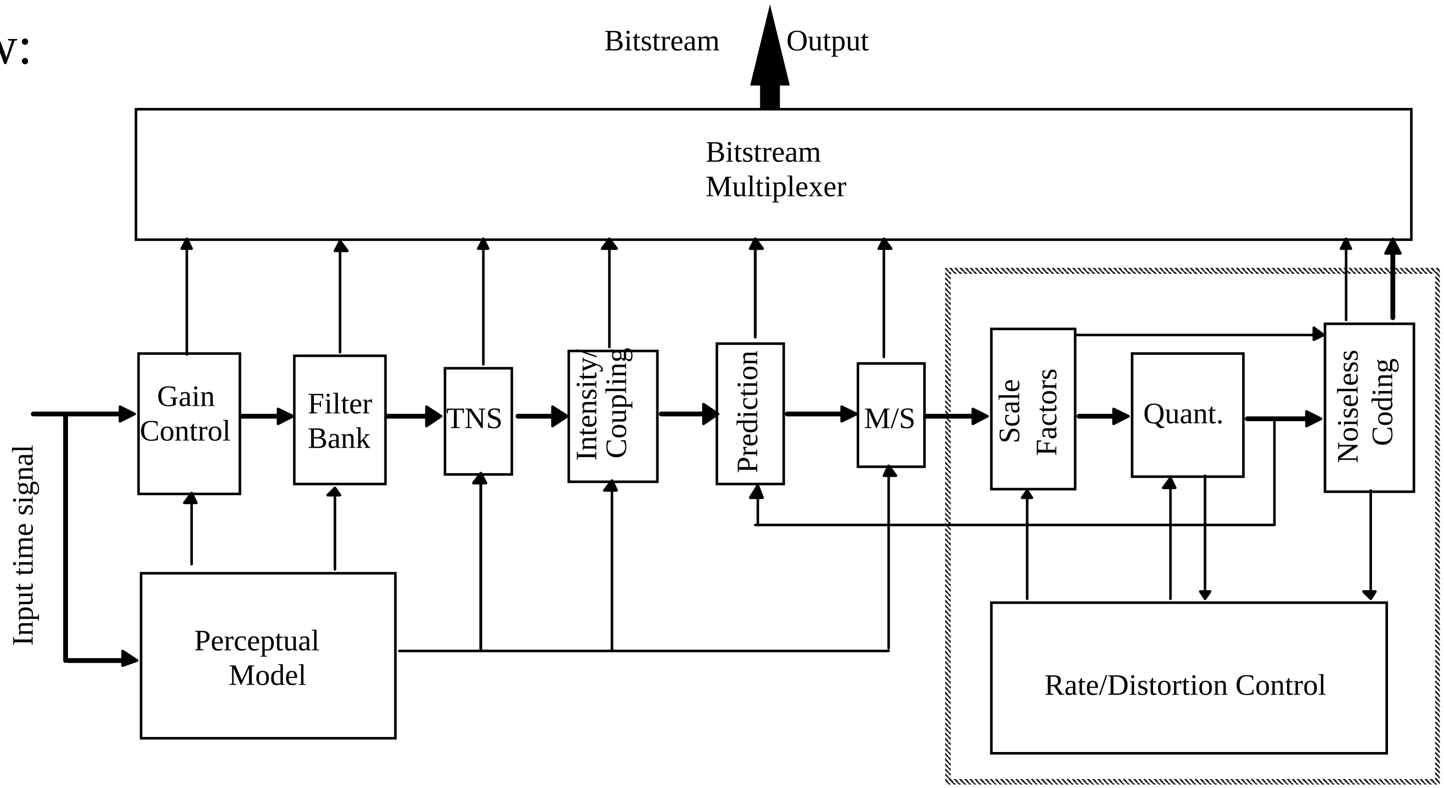
General Audio Coding

- Solid state players, Internet audio, terrestrial and satellite digital audio broadcasting (HDTV), ...
- MPEG-2 Advanced Audio Coding (AAC)
- MPEG-4 Extensions to AAC:
  - Perceptual Noise Substitution (PNS)
  - Long Term Prediction
  - TwinVQ Coding Core
  - Bit Slice Arithmetic Coding (BSAC) (v2)
  - Low Delay AAC (v2)

Fraunhofer Institut Integrierte Schaltungen

MPEG-2 AAC (1)

Encoder Overview:



MPEG-2 AAC (2)

Basic features:

- High frequency resolution filterbank-based coder (1024 lines MDCT with 50% overlap)
- 1:8 block switching (8 *128 lines MDCT)
- Non-uniform quantizer
- Noise shaping in half critical bands (scalefactor bands)
- Huffman coding of scalefactors and spectral coefficients

MPEG-2 AAC (3)

Advanced coding tools:

- Window shape adaptation
- Temporal noise shaping (TNS)
- Gain control (SRS profile only)
- Backward adaptive prediction

Joint stereo coding tools:

- Mid/Side stereo (MS) per scalefactor band
- Intensity stereo coding between channel pairs
- Coupling channel(s)

Other features:

- Flexible bitstream format for up to 48 audio channels, up to 16 Low Frequency Enhancement (LFE) channel(s)

MPEG-2 AAC Performance

Test Results:

- Broadcast quality at 320 kbit/s for 5 channels (better than MPEG-2 Layer II at 640 kbit/s)
- Broadcast quality at 128 kbit/s stereo
- Comparison to other codecs:
  AAC 96 kbit/sec stereo comparable to
  - AC-3 at 160 kbit/s
  - Layer II at 192 kbit/s
  - Layer III at 128 kbit/s
- Very low bitrates (comparison within MPEG):
  AAC best audio coder at bitrates down to 16 kbit/s for mono and stereo

Perceptual Noise Substitution

Principle:

- Noise-like signal components are detected on a scalefactor band basis
- Corresponding groups of spectral coefficients are excluded from quantization/coding
- Instead, only a "noise substitution flag" plus total power of the substituted band is transmitted in the bitstream
- Decoder inserts pseudo random vectors with desired target power as spectral coefficients

⇒ Highly compact representation for noise-like spectral components

Fraunhofer Institut Integrierte Schaltungen

Demonstration: Perceptual Noise Substitution

- Contemporary Pop Spot:

w/o PNS	with PNS	PNS parts only

AAC coder @ 32 kbit/s, fs=48 kHz, bandwidth 14 kHz

Note: This is not "normal" operation mode

Fraunhofer Institut Integrierte Schaltungen

Transform-Domain Weighted Interleave VQ (1)

Background:

- Desire audio coding at extremely low bitrates (6 kbit/s)
- CELP speech coders do not perform well for music

MPEG-4:

- *Transform-Domain Weighted Interleave Vector Quantization (TwinVQ)* as additional coding kernel
- Fully integrated into MPEG-4 AAC coding system:
  - Uses same spectral representation as AAC coder
  - Makes use of other MPEG-4 tools (e.g. LTP, TNS, joint stereo coding)
  - Possible core coder for ⇒ MPEG-4 scalable coding

Fraunhofer Institut Integrierte Schaltungen

Transform-Domain Weighted Interleave VQ (2)

Structure:

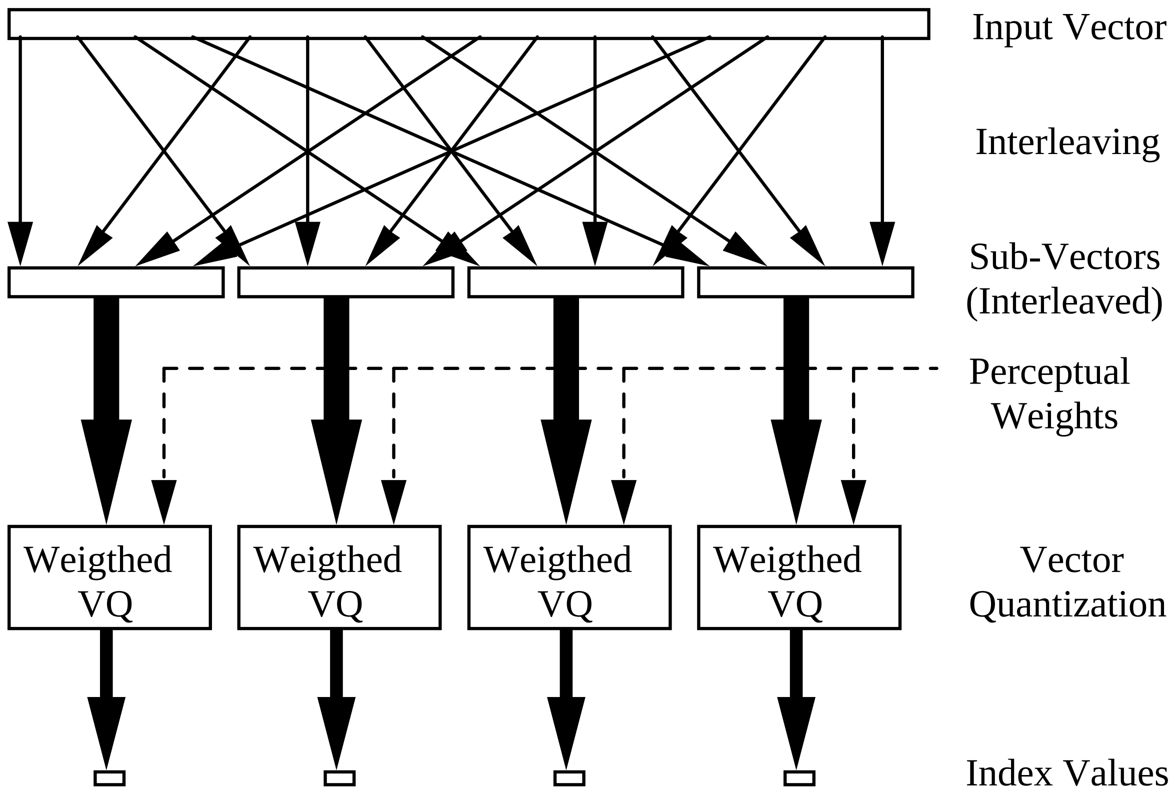
- Normalization of spectral coefficients:
  - LPC envelope (overall spectral shape)
  - Periodic component coding (harmonic components)
  - Bark-scale envelope coding (additional flattening)

- Vector Quantization (VQ) process:
  - Interleaving of spectral coefficients into new sub-vectors
  - Vector quantization
    (two sets of codebooks, weighted distortion measure allows distortion control by perceptual model)

⇒ no bit/noise allocation or rate control iteration

Transform-Domain Weighted Interleave VQ (3)

Interleave VQ Structure:



Dr. Jürgen Herre, hrr@iis.fhg.de

Fraunhofer Institut Integrierte Schaltungen

Conclusions

- MPEG-4 General Audio Coding:
  The "all-round" coding system among the MPEG-4 audio schemes, providing a set powerful tools
- Based on MPEG-2 Advanced Audio Coding kernel
- MPEG-4 enhancements
  - add functionality
  - improved coding efficiency

Fraunhofer Institut Integrierte Schaltungen