b

UNIVERSITÄT
BERN

UNIVERSITÄT BERN

Advanced Networking and Future Internet

IX. Audio Compression

Prof. Dr. Torsten Braun, Institut für Informatik

Bern, 09.11.2020

Advanced Networking and Future Internet: Audio Compression

UNIVERSITÄT BERN

Table of Contents

UNIVERSITÄT BERN

1. Introduction

– Uncompressed images, audio and video

→ High storage and bandwidth requirements

Examples

- Stereo audio signal with CD quality
  - 2 channels, sampling rate: 44.1 kHz, quantization: 16 bits per sample
  - Bandwidth: 1.4 Mbps
  - Memory for 1 hour: 605 MB
- Video sequence
  - HDTV: 1920 • 1080 pixels, 24 bits per pixel, 30 frames / s
  - Bandwidth: 1.5 Gbps

UNIVERSITÄT BERN

1. Introduction

1. Entropy Encoding

- is lossless,
- is used for different media, and
- does not consider media dependent characteristics.
- Data to be compressed are considered as a sequence of digital values.

Examples

- Run Length Encoding (RLE)
- Statistical encoding
  - Encoding of characters with different numbers of bits
    - Frequent characters → short bit sequences
    - Infrequent characters → long bit sequences
  - Examples: Huffman and arithmetic encoding
- Directory methods
  - Patterns are often known a priori, e.g., in data bases.
  - Ordering of patterns in table and replacement by index prior to transmission
  - Example: Vector Quantization

UNIVERSITÄT BERN

1. Introduction

2. Source Encoding

- is mostly lossy and uses semantics of information to be encoded.
- Original and decoded data is similar but not identical.

Examples

- Differential encoding
  - encodes differences between individual samples,
  - is useful for large amplitudes, but small differences between samples
  - can be lossy or lossless.
- Transformation encoding
  - Transformation between types of information, e.g., discrete cosine transform: pixel matrix → frequency matrix
  - Transformation of data from one mathematical domain into another one that is more suitable for further compression
  - allows frequency-specific encoding, sub-band coding
  - Example: Discrete Cosine Transformation
    - Audio: one-dimensional DCT (in time)
    - Images: two-dimensional DCT (in space)

1. Introduction

3. Digitization of Audio Data

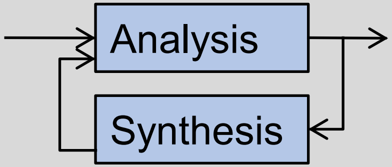
- Periodic measurement of amplitudes
  → sequence of samples
- Analog/Digital Converter
  - Transformation of analog audio signal to digital samples
    - Sampling
      - Sampling theorem: sampling rate > 2 • highest frequency
    - Quantization
      - Resolution depends on available bits for encoding the digital sample.
- Digital/Analog Converter
  - Transformation of digital samples into analog audio signal
- ISDN: Pulse Code Modulation: 8 bit sample every 125 µs

1. Introduction

4. Voice Encoding

Voice encoding consists of

- Analysis
  - Encoding of voice as a compact set of parameters
    - Open loop
      - No feedback in encoding process
    - Closed loop
      - Extraction of parameters and encoding of differential values
      - *analysis by synthesis*
  - Parameter presentation
    - Non voice-specific: ***waveform*** coding
    - Voice-specific: ***vocoding*** based on voice modeling
- Synthesis
  - Decoding of parameters and voice reconstruction



UNIVERSITÄT BERN

2. Waveform Coding

- Encoding in time domain
- Approximation of original waveform
- Better quality and more robust than vocoders, but less efficient (higher bandwidth)
- tries to determine relation between different samples

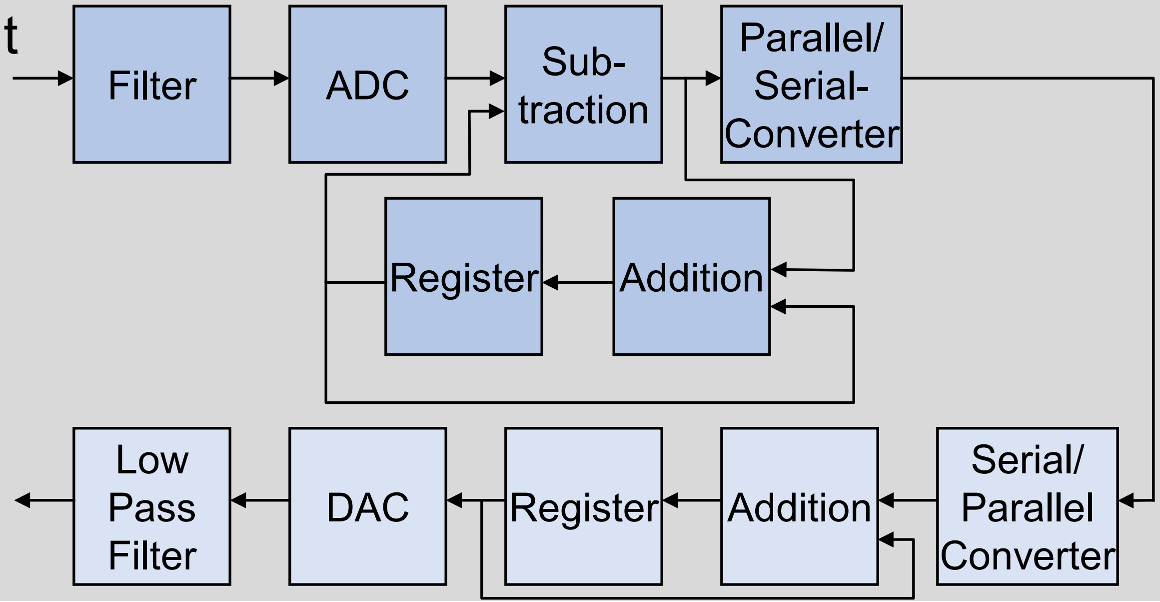
Options

1. Differential coding: encoding relative to previous value; Calculation of difference can result in smaller values.
2. (Linear) Prediction: coding of difference between waveform and predicted one
3. Vector quantization

2. Waveform Coding

1. Differential PCM

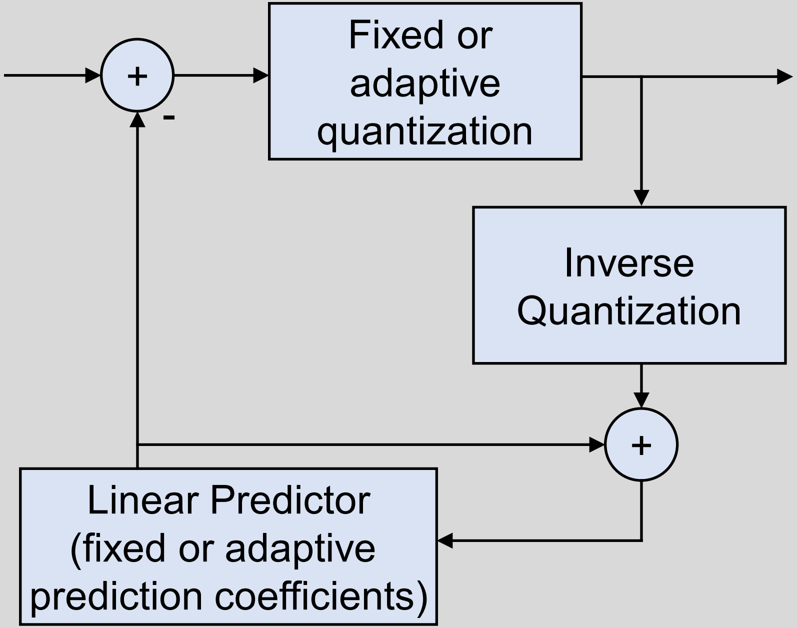
- Difference of subsequent audio samples is often very small.
- Coding of difference values requires a few bits only.
- Delta Modulation: Encoding of difference values in 1 bit



2. Waveform Coding

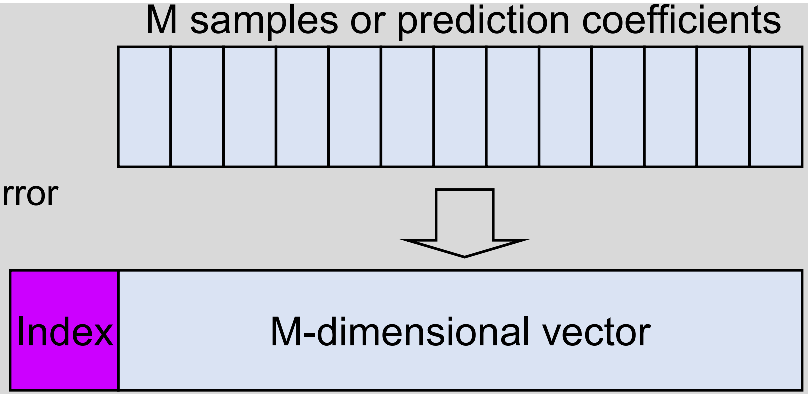
2. Linear Prediction

- Calculation at time j:
  $e(j) = y(j) - [a_1 y(j-1) + a_2 y(j-2) + \ldots + a_p y(j-p)]$
  - y(i): sample
  - a_i: prediction coefficients
  - [predicted value]
- e(j) have a small value range,
  which allows encoding using a few bits only.
- Optimization of prediction coefficients a_i to minimize
  $E = e^2(0) + e^2(1) + \ldots + e^2(p)$, e.g., every 10-30 ms.



2. Waveform Coding

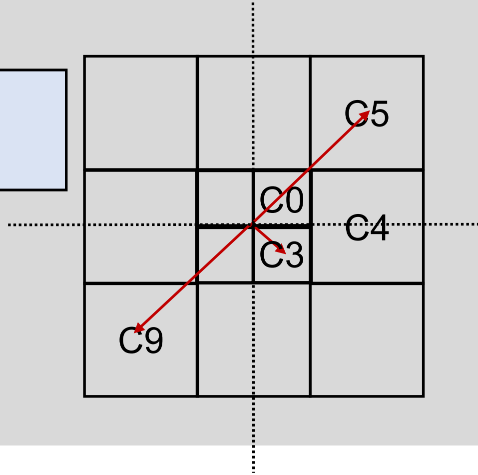
3. Vector Quantization



- Vector table = codebook
- Calculation of the closest vector from codebook (smallest sum of error squares or mapping into space):

$$\epsilon(s,v) = \sum_{k=0}^{M-1} \left(s(k) - v(k)\right)^2$$

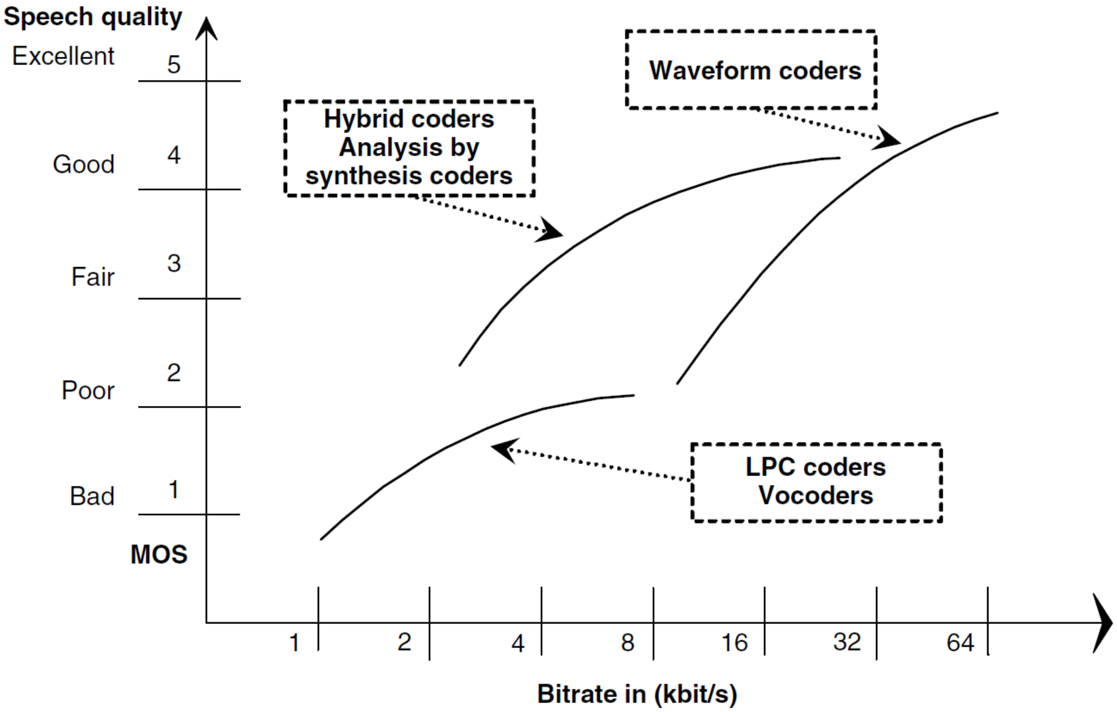
- Transmission of vector index from table
- Problem: delay
- Codebook design:
  a priori (fixed) after training phase or adaptive



3. Vocoding

- Transformation into frequency domain and analysis of frequency spectrum
- Make use of rather slow rate of signal change
- Characterization of waveforms based on various parameters such as
  - voiced / unvoiced
  - signal energy
  - signal duration
  - frequency
- Reconstruction (synthesis) at receiver of a completely new waveform from transmitted parameters
  - sounds often synthetic
  - extremely low bit rates in the area of a few kbps
- Example: Code-Excited Linear Prediction
  - Pre-calculation of a finite set of waveform templates and storing them in template codebooks at sender and receiver
  - Assignment of waveform to best suited template
  - Accumulation and processing of several samples creates significant (algorithmic/processing) delay.

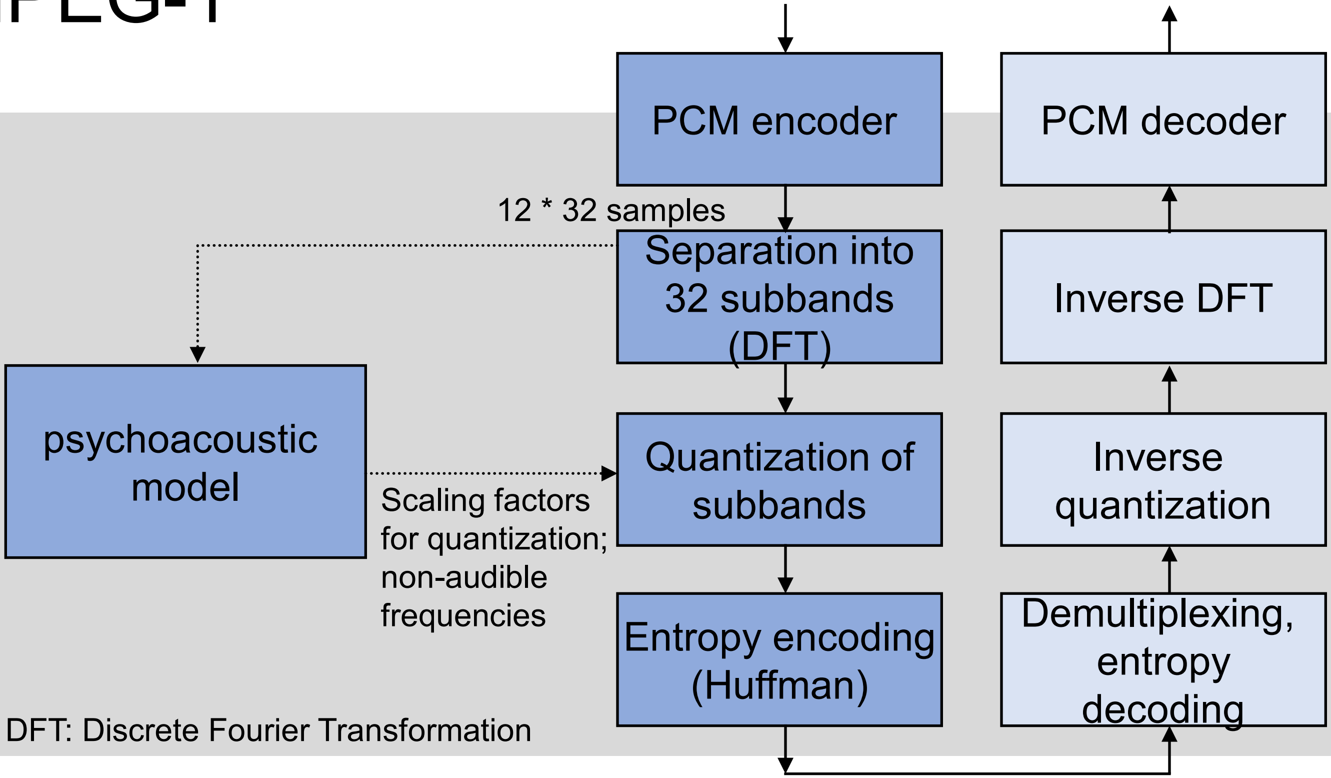
4. Quality of Audio Encoding



UNIVERSITÄT BERN

5. MPEG-Audio

1.1 MPEG-1



DFT: Discrete Fourier Transformation

5. MPEG-Audio

1.2 MPEG-1 Layers

Layers with increasing complexity and performance

I. Digital Audio Cassette
   – HiFi quality at 192 kbps per channel, delay: 20 ms

II. Digital Audio and Video Broadcasting
   – Near CD quality at 128 kbps per channel, delay: 40 ms

III. CD quality over low bit rate channels
   – CD quality at 64 kbps per channel, delay: 60 ms

5. MPEG-Audio

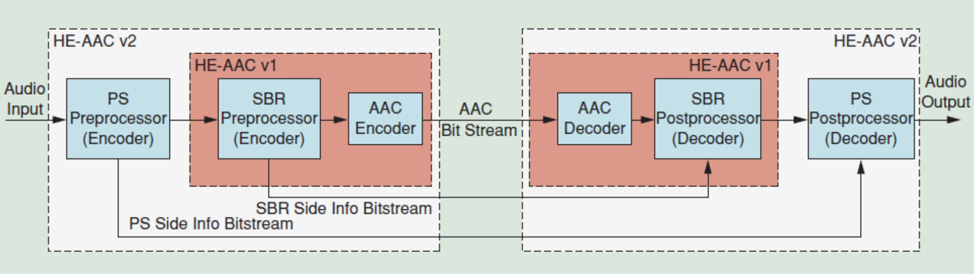
2. MPEG-4

- Efficient coding mechanisms for low bit rates, e.g., based on CELP
- Scalability
- Audio objects, e.g., background sound
- Synthetic / natural / hybrid encoding
- Text-to-speech
- Variants
  - Advanced Audio Coding
  - Audio Lossless Coding

5. MPEG-Audio

2.1 MPEG-4 Advanced Audio Coding

- AAC encoder
  - Waveform codec using modified DCT (≡ 1024 PCM input samples), stereo processing, temporal noise shaping, quantization and coding
- Spectral Band Replication
  - Regeneration (instead of transmission) of higher frequencies from lower frequencies using low-bit rate guidance data

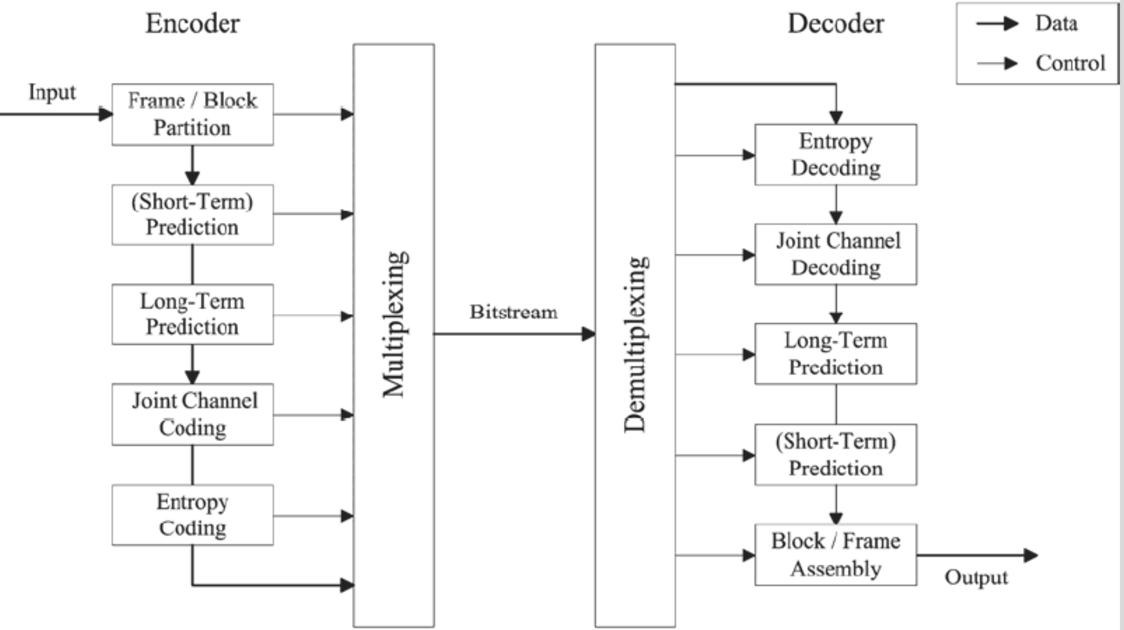


- Parametric Stereo
  - Mono-downmix transmission, sophisticated stereo processing considering time and phase differences

UNIVERSITÄT BERN

5. MPEG-Audio

2.2 MPEG-4 Audio Lossless Coding



- Frame Block Partition
  - Division of audio data stream into frames and blocks
- Short/Long-Term Prediction
  - Calculation of residual error signal using linear prediction and forward adaptive prediction
- Joint Channel Coding
  - Prediction based on similarity of channels
- Entropy Coding

Thanks

for Your Attention

b UNIVERSITÄT BERN

Prof. Dr. Torsten Braun, Institut für Informatik

Bern, 09.11.2020