Lab 4

Data Compression

- **Reduces the number of bits required to represent a signal**
- **Exploits redundancies in a signal, i.e., statistical correlations**
  - Can’t compress noise!

Two types

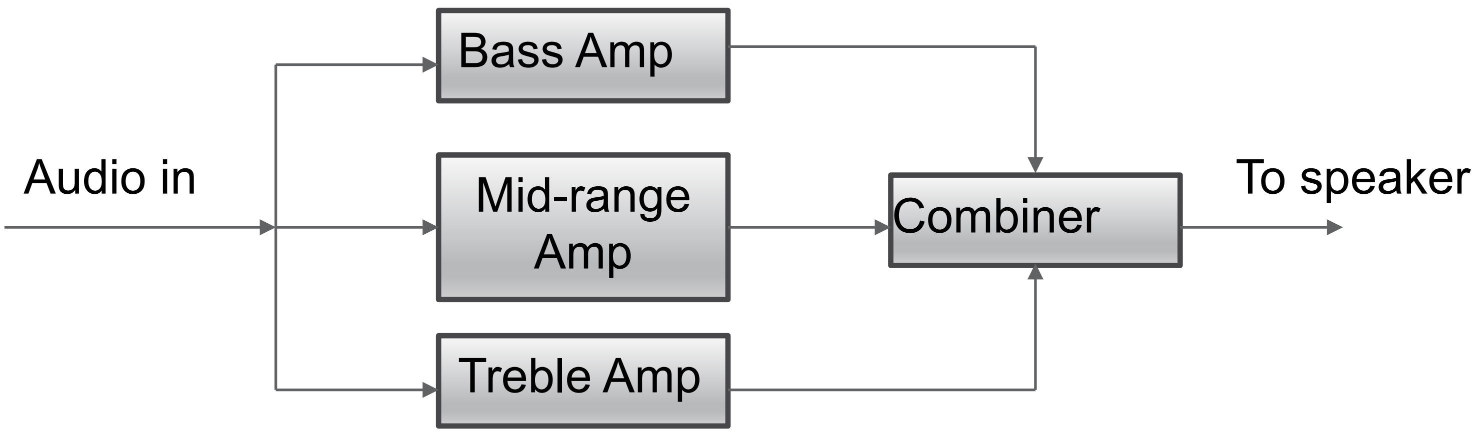
- **Lossless**
  - Can reconstruct the original signal from the compressed bistream
    - ✓ Banking data, emails, etc.
- **Lossy**
  - Cannot perfectly reconstruct the original signal
    - ✓ Audio, video

Compression Techniques

- **Lossless**
  - Huffman coding, arithmetic coding, Lempel-Ziv, etc.
- **Lossy**
  - Resolution reduction, quantization, predictive coding, transform coding

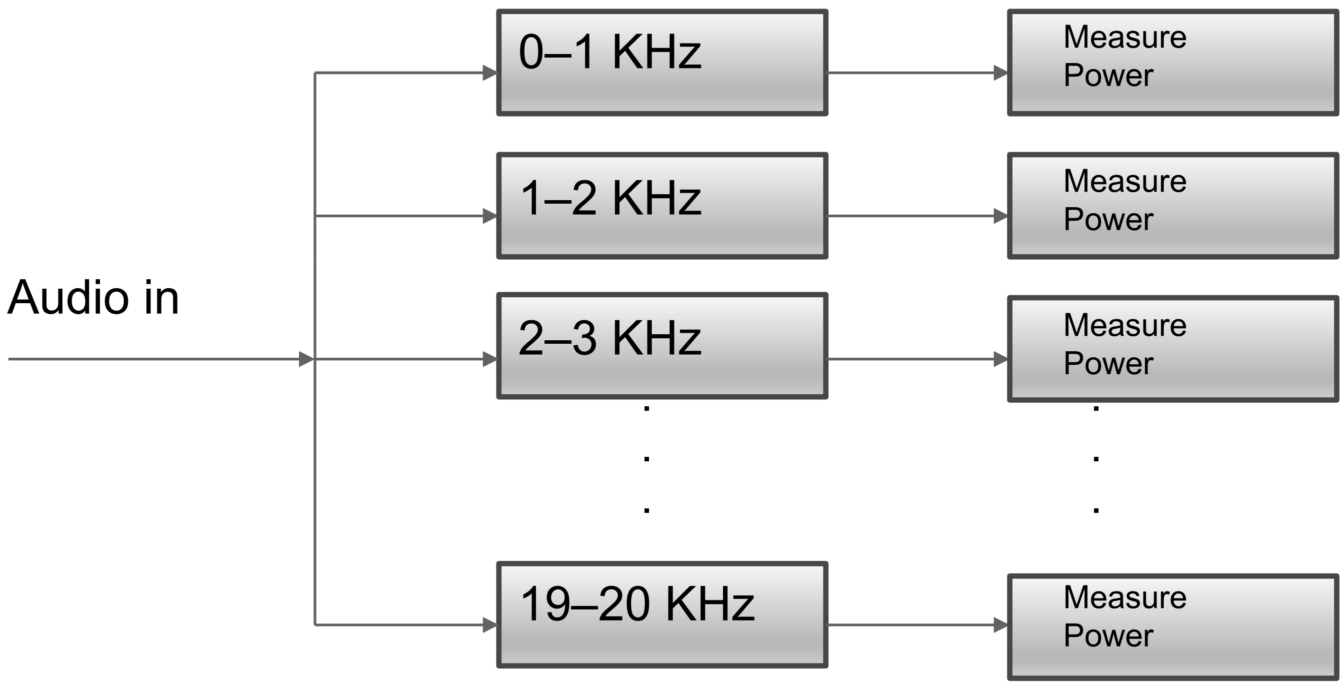
Sub-band Audio Coding

Warmup: Audio equalizer



Warmup: spectrum analyzer

CU Boulder

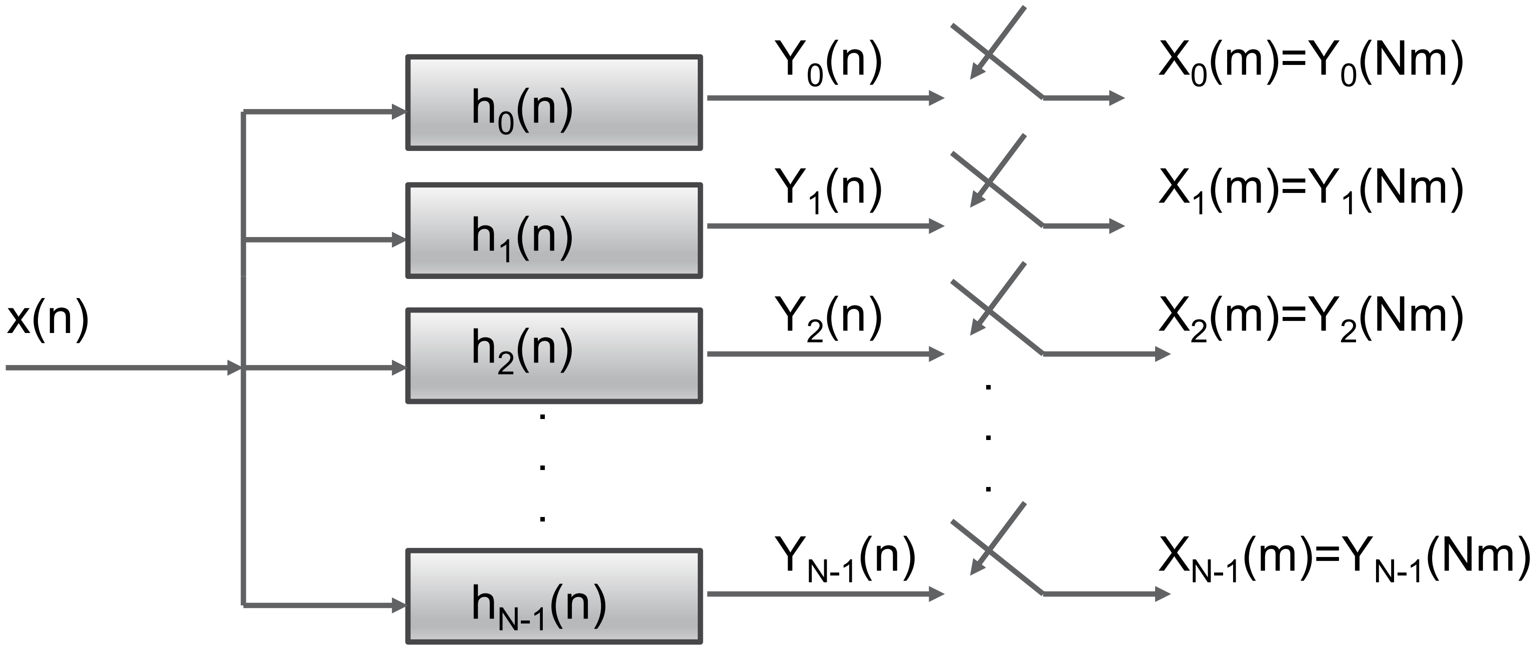


The DFT as a Filter Bank

- **The length-N DFT viewed as a bandpass filter bank**

Impulse response of filters: $h_k(n) = e^{-\frac{i2\pi kn}{N}}$

Sample every N samples (sub-sample by N)

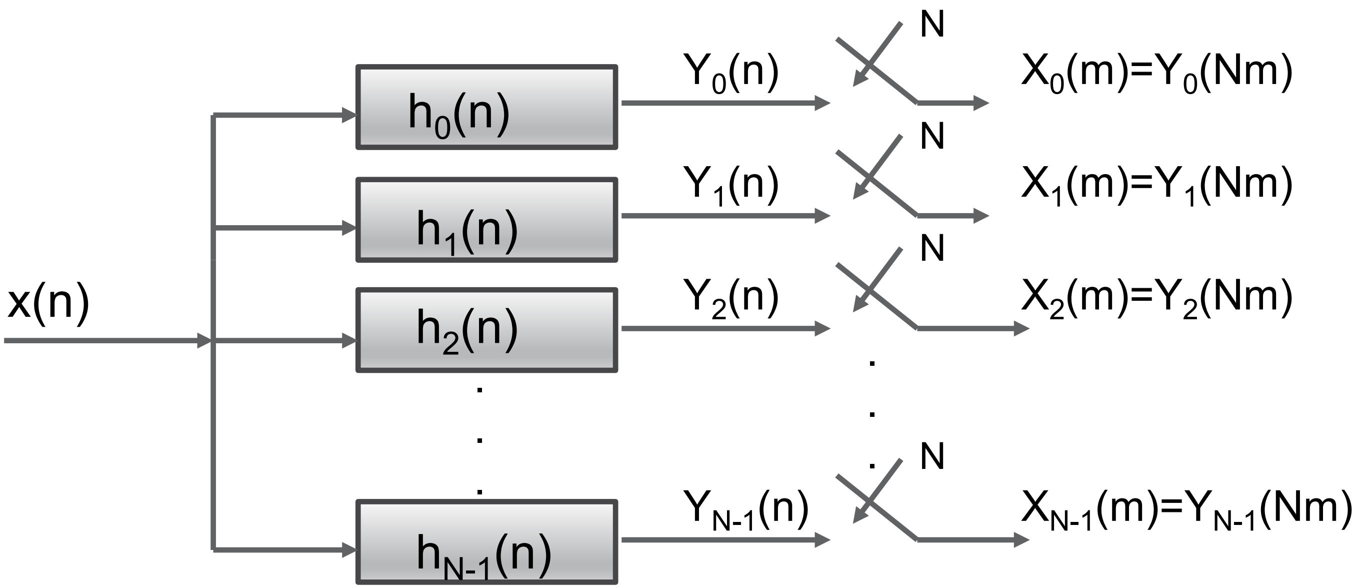


Filters (convolve with impulse response)

$$X_k = \sum_{n=0}^{N-1} x(n) e^{-\frac{i2\pi kn}{N}}$$

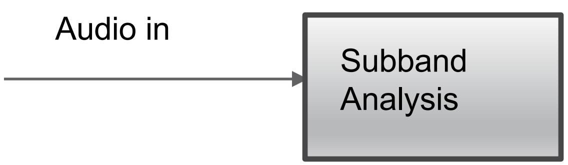
Observations / Extensions

- **The impulse responses, sinusoids, implement bandpass filters**
- **The filter responses could be longer than length *N***
  - Aliasing cancellation on reconstruction must be considered



Generic perceptual audio codec

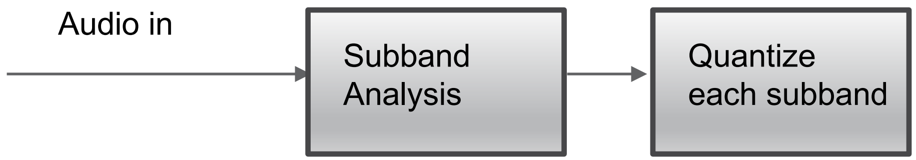
- **Step 1: "Analyze" spectrum into sub-bands with a filter bank**



Generic perceptual audio codec

CU Boulder

Step 3: Quantize sub-bands

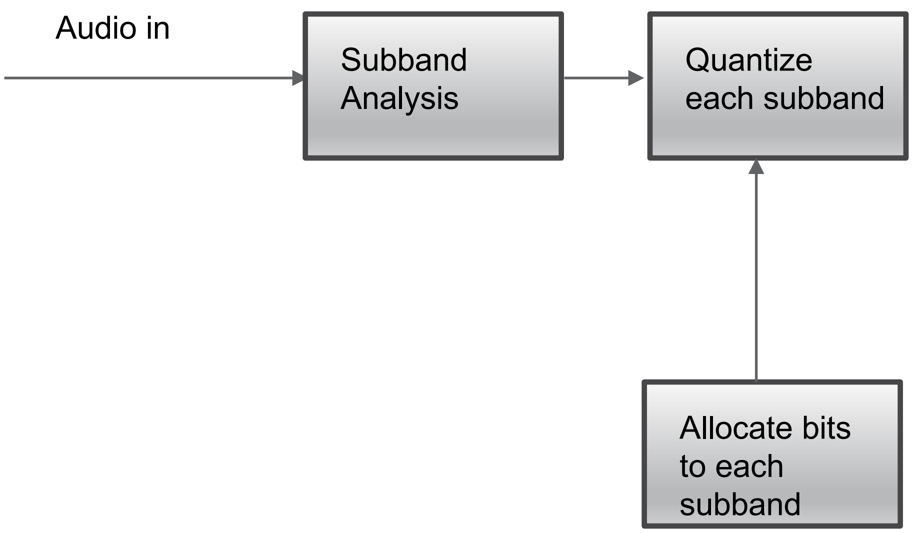


- But, are all sub-bands equally important?
  - ✓NO
    - Some carry more power
    - Some are more perceptually significant

So we first need Step 2 on the next slide!

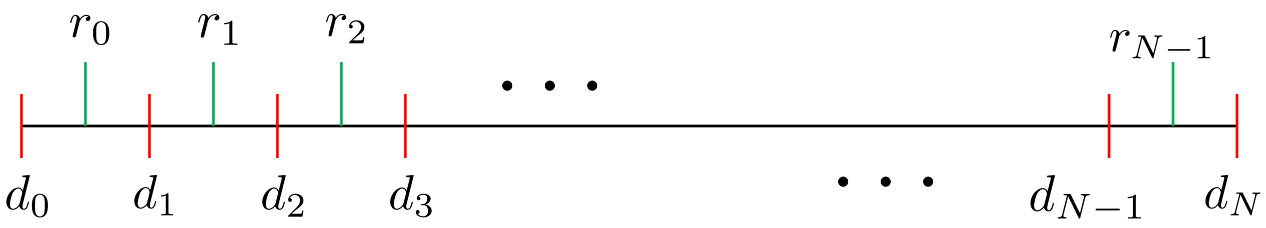
Generic perceptual audio codec

- **Step 2: Allocate bits to sub-bands**
  - How should we choose to allocate bits?
    - ✓Minimize error power?
    - ✓Minimize perceptability?



Uniform quantizer reminder

- **All values between two decision levels are mapped to the same reconstruction level**
  - Fewer bits means a coarser representation of the signal

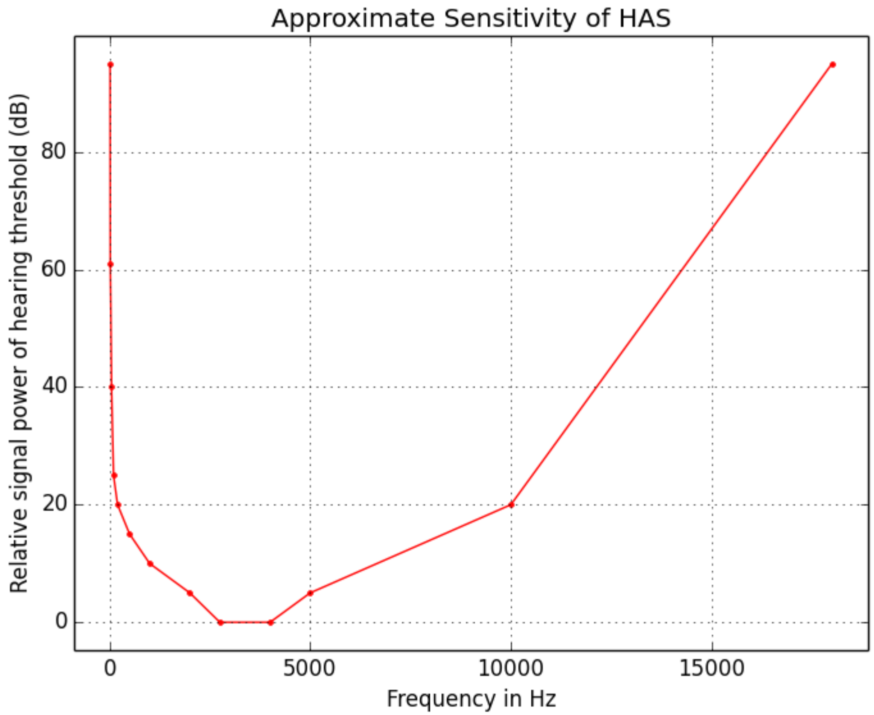


Generic perceptual audio codec

- **Step 3: Allocate bits to sub-bands**
  - We have at least two choices when allocating bits
    - ✓ Minimize error power
      - – May be appropriate if subsequent signal processing steps will be required later for other purposes
        - » Think studio processing, not consumer listening
    - ✓ Minimize perceptibility
      - – If the goal is to listen to the music, perceptual coding is justified and offers significant gains
        - » Consumer listening

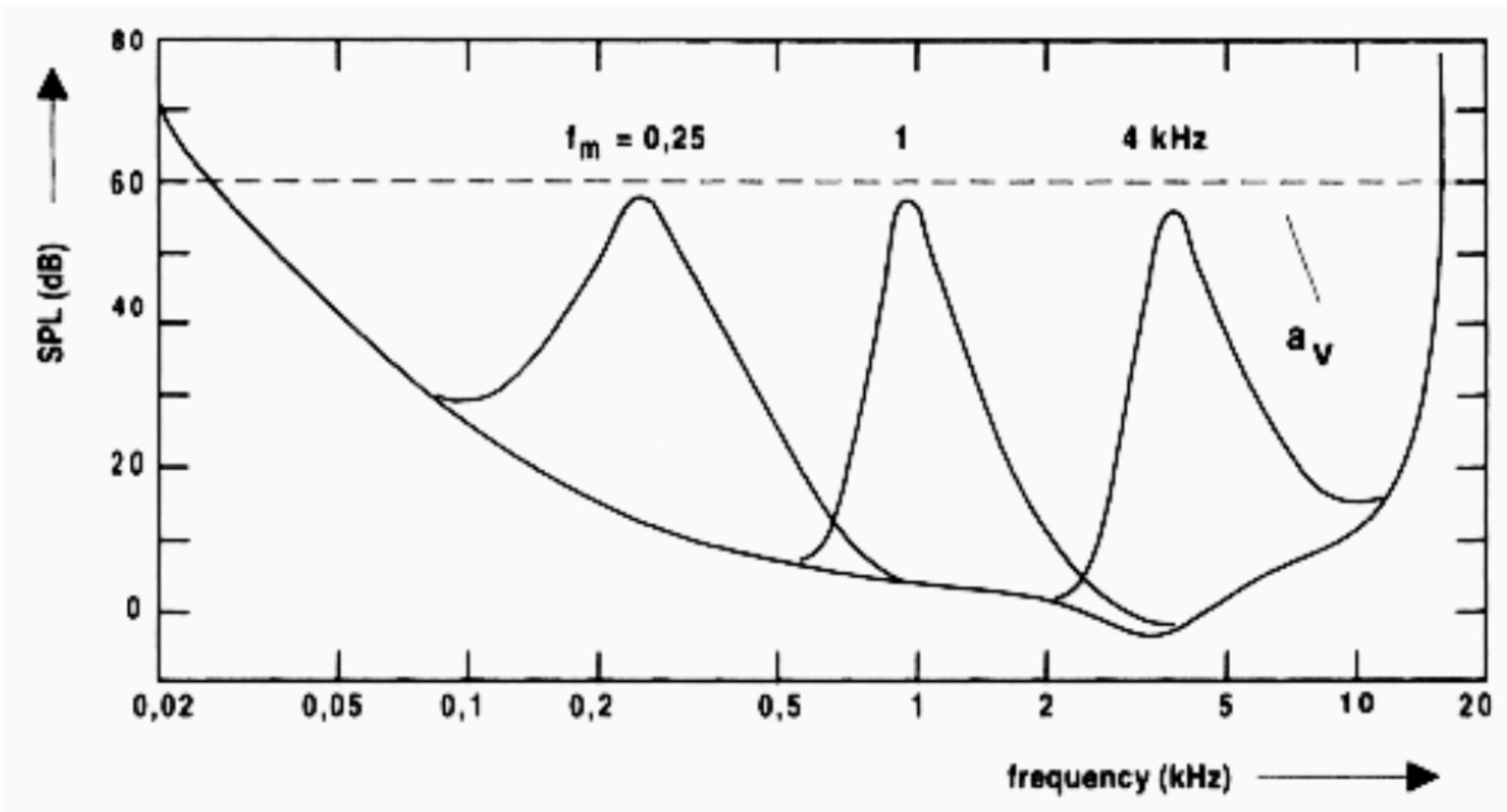
Generic perceptual audio codec

- **Human Auditory Sensitivity (HAS) as a function of frequency**



Generic perceptual audio codec

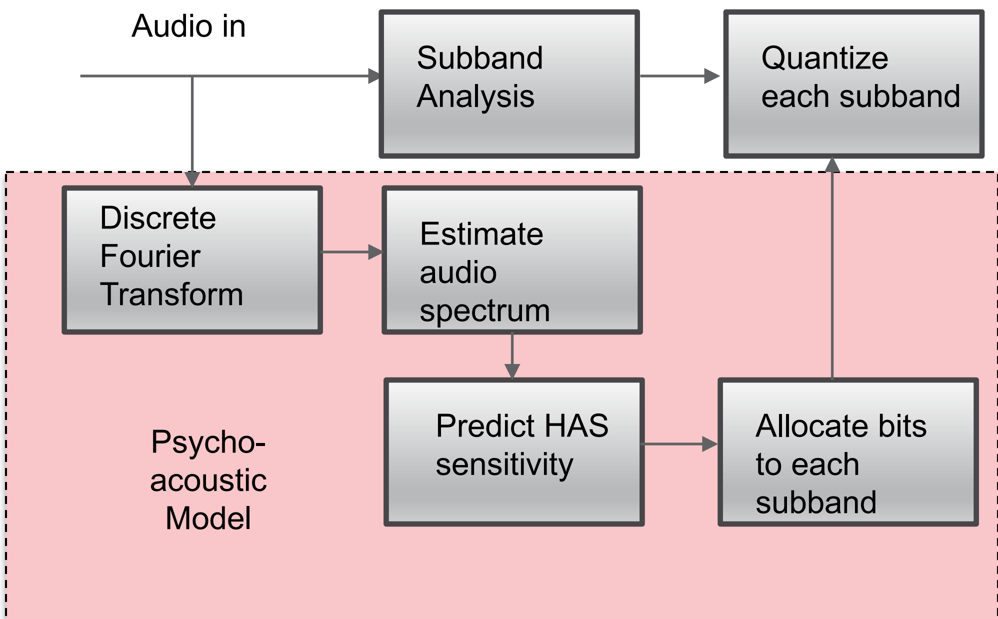
- **Audio masking**
  - Sensitivity depends on power of nearby signals



Generic perceptual audio codec

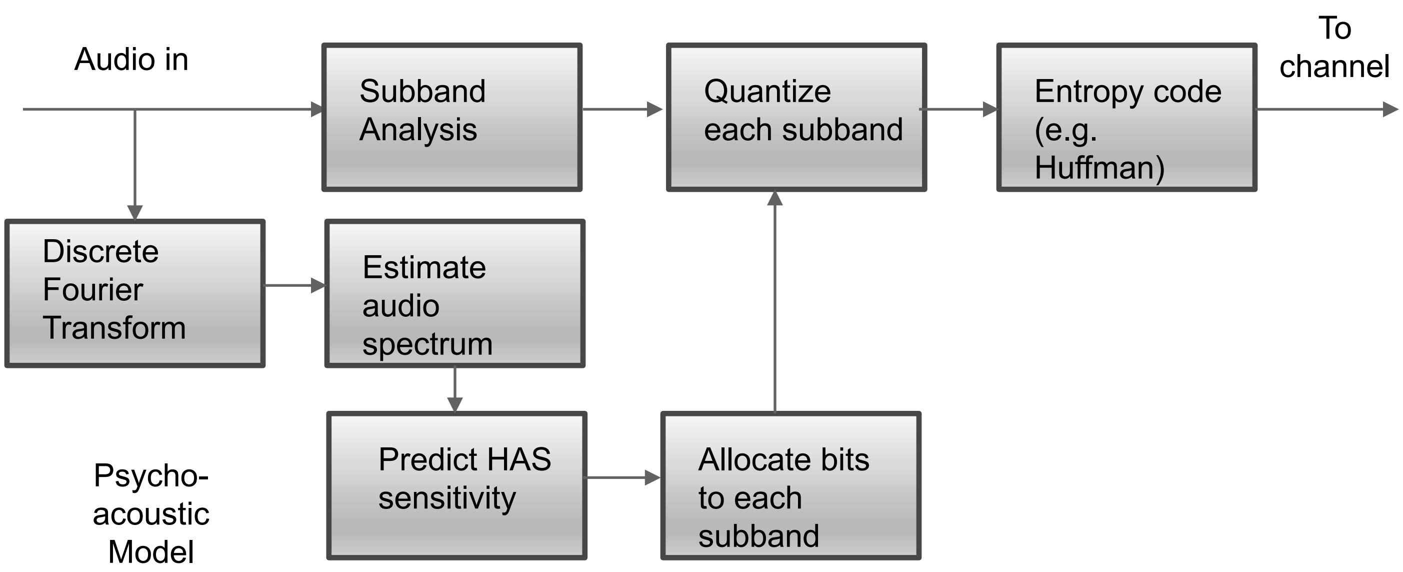
CU Boulder

- **To allocate bits we estimate the spectrum and take into account HAS sensitivity, including masking effects**



Generic perceptual audio codec

- **Step 4: Huffman coding**
  - Lossless coding to further reduce the bitrate



Huffman coding

- **Allocate fewer bits to the most probable symbols**
  - The theoretically optimal codeword length

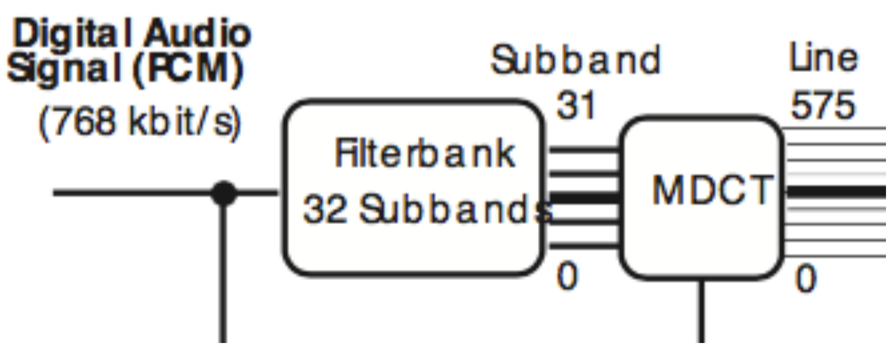
$$l(x) = \log_2\left(\frac{1}{p(x)}\right)$$

  - The expected value of the number of bits per symbol when the theoretical limit is achieved is the entropy

$$H(X) = \sum_x p(x) \log_2\left(\frac{1}{p(x)}\right)$$

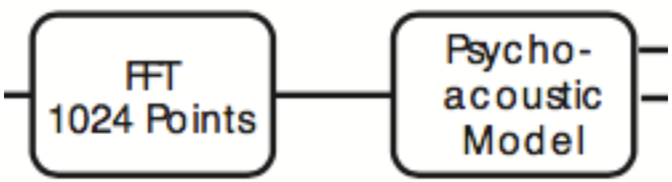
MP3 is a perceptual audio coder

- **Utilizes analysis / synthesis**
  - Hybrid system
    - ✓ First a 32 bank subband analysis
    - ✓ Each subband is then further analyzed into 18 frequency channels using an MDCT
      - – Think of MDCT as like the DFT



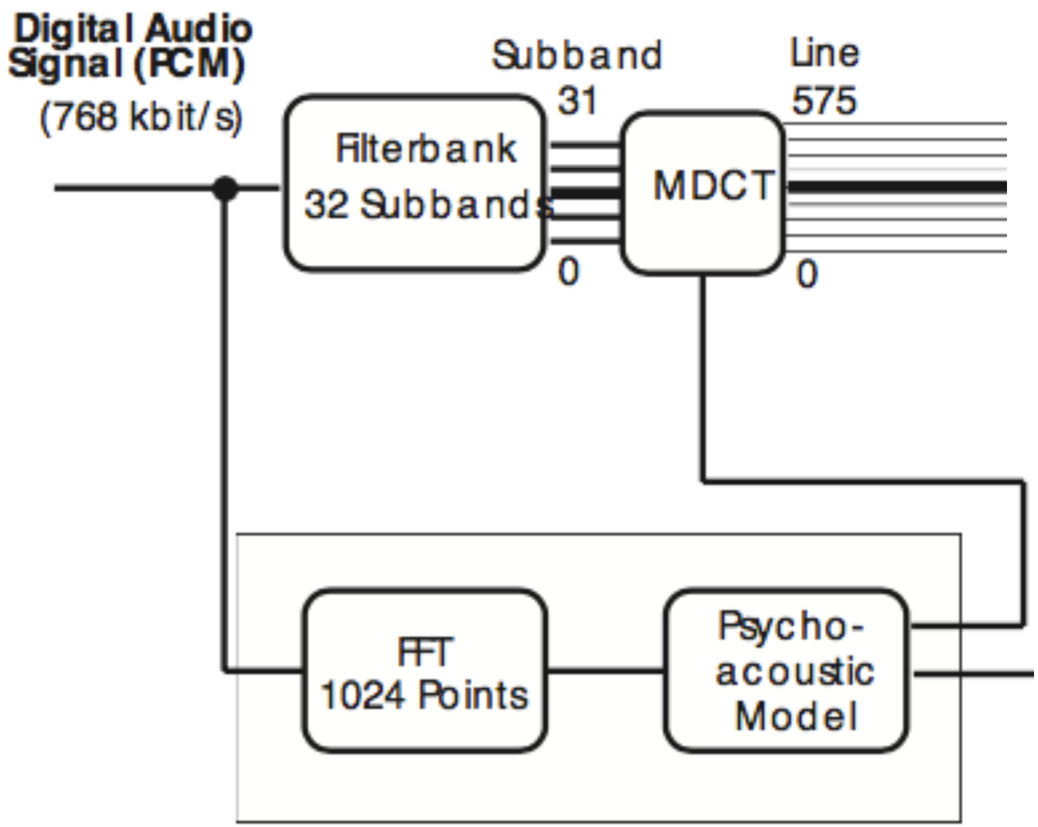
Psycho-acoustic model

- **Psycho-acoustic model**
  - Determines amount of noise that can be allowed in each analysis sub-band
  - Usually driven by a separate frequency analyzer (DFT)
  - Temporal masking and stereo considerations are exploited



Analysis and perceptual model

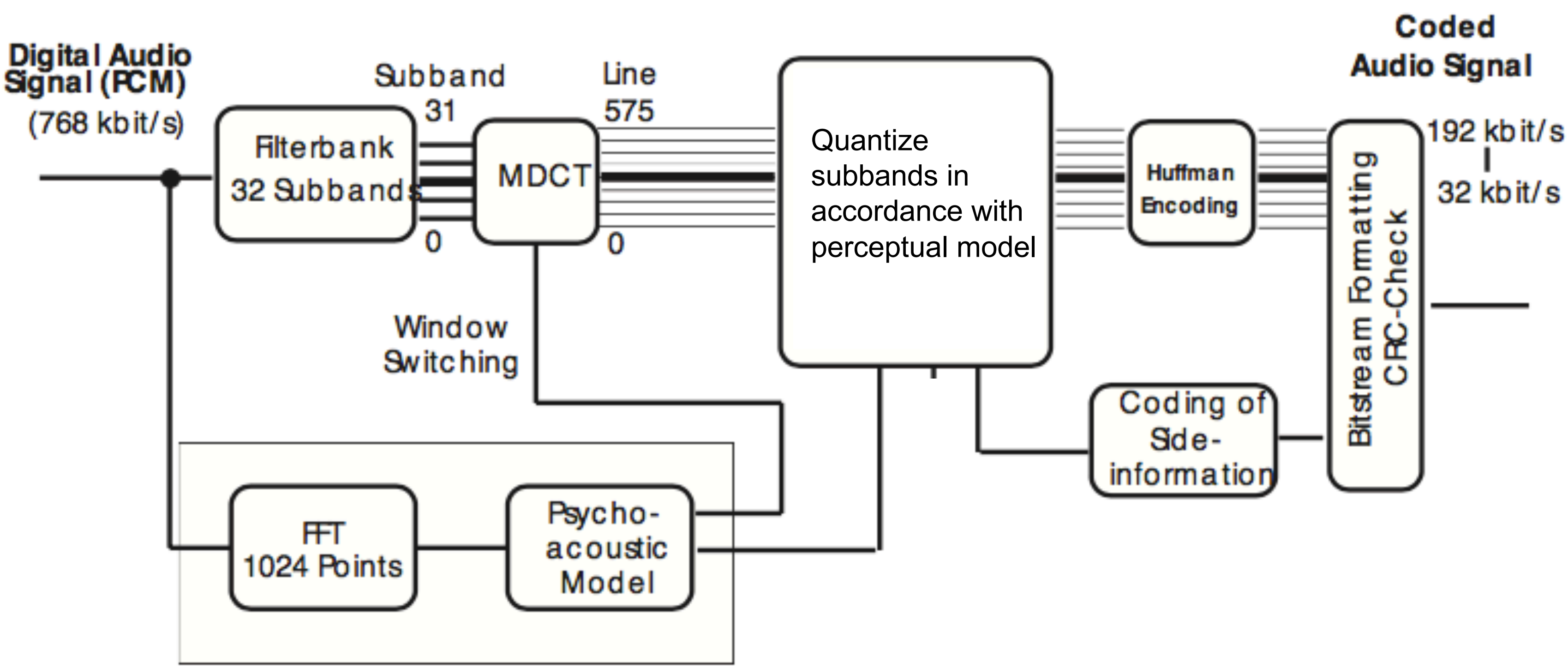
- **Front-end analysis and perceptual model components**



Quantization and Coding

- **Uses a power-law non-linear quantizer**
  - Larger sound intensity values are quantized more coarsely
- **Different subbands are quantized with different quantizer step sizes (“scalefactors”)**
  - Goal is to drive quantization noise just below threshold of perception
- **Quantizer outputs are Huffman encoded**
  - Optimum Huffman table is adaptively selected depending on local music statistics
  - Different tables for different frequency bands

MP3 entire encoder



Iterative approach to subband quantization and coding