ECE462 – Lecture 27

MPEG Layers

- MPEG audio offers three compatible *layers* :
  - Each succeeding layer able to understand the lower layers
  - Each succeeding layer offering more complexity in the psychoacoustic model and better compression for a given level of audio quality
  - each succeeding layer, with increased compression effectiveness, accompanied by extra delay
- The objective of MPEG layers: a good tradeoff between quality and bit-rate

MPEG Layers

- Layer 1 quality can be quite good provided a comparatively high bit-rate is available
  - Digital Audio Tape typically uses Layer 1 at around 192 kbps
- Layer 2 has more complexity; was proposed for use in Digital Audio Broadcasting
- Layer 3 (MP3) is most complex, and was originally aimed at audio transmission over ISDN lines
- Most of the complexity increase is at the encoder, not the decoder – accounting for the popularity of MP3 players

MPEG Audio Strategy

- **MPEG approach to compression** relies on:
  - Quantization
  - Human auditory system is not accurate within the width of a critical band (perceived loudness and audibility of a frequency)
- **MPEG encoder** employs a bank of filters to:
  - Analyze the frequency ("spectral") components of the audio signal by calculating a frequency transform of a window of signal values
  - Decompose the signal into subbands by using a bank of filters (Layer 1 & 2: "quadrature-mirror"; Layer 3: adds a DCT; psychoacoustic model: Fourier transform)

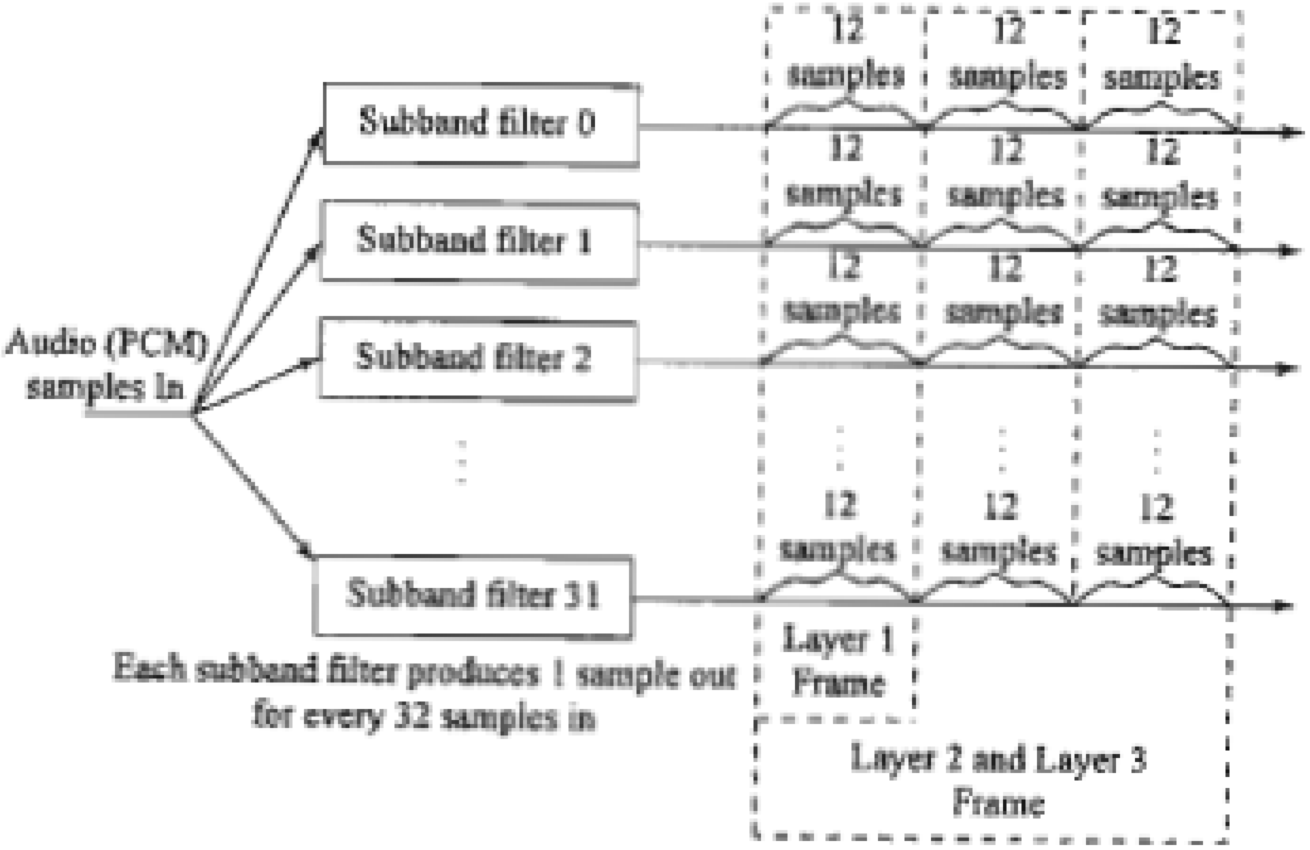
MPEG Audio Strategy

- **Frequency masking**: by using a psychoacoustic model to estimate the just noticeable noise level:
  - Encoder balances the masking behavior and the available number of bits by discarding inaudible frequencies
  - Scaling quantization according to the sound level that is left over, above masking levels
- May take into account the actual width of the critical bands:
  - For practical purposes, audible frequencies are divided into 25 main critical bands (Table 14.1)
  - To keep simplicity, adopts a *uniform width* for all frequency analysis filters, using 32 overlapping subbands

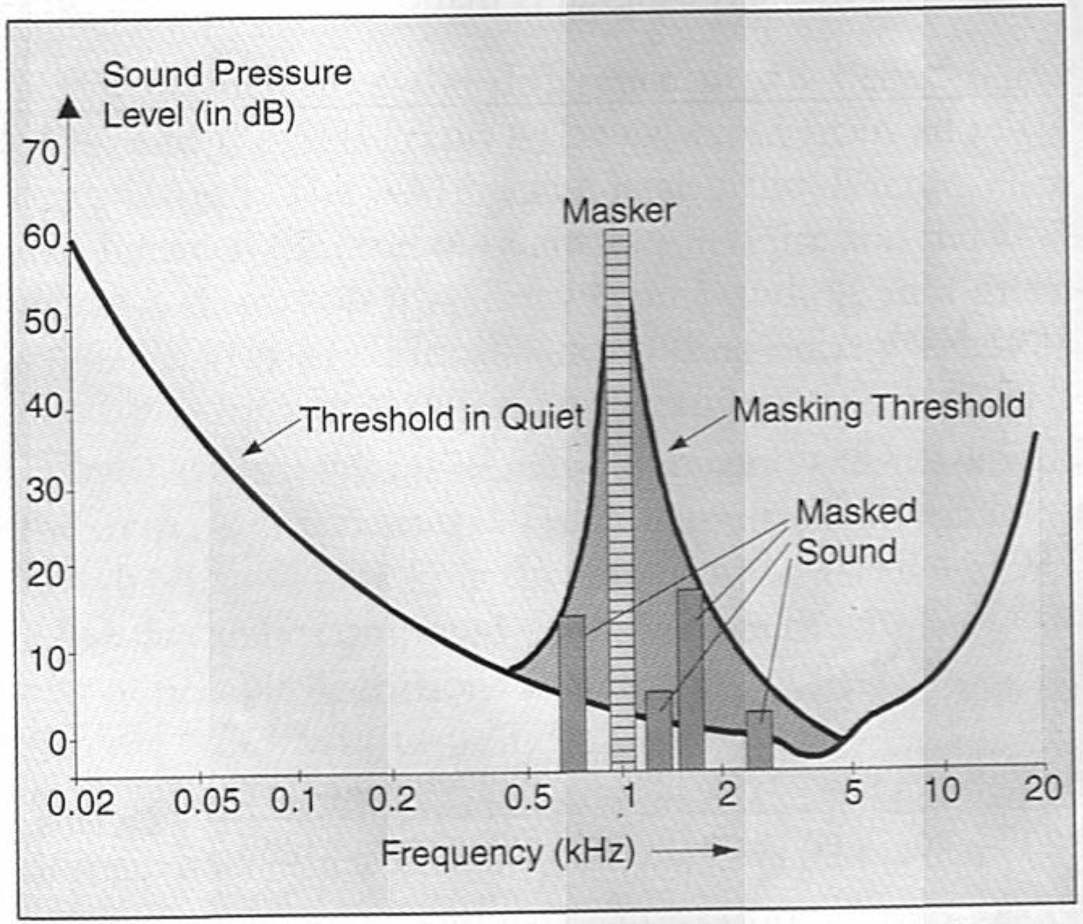
Basic Algorithm

- The algorithm proceeds by dividing the input into 32 frequency subbands, via a filter bank
  - A linear operation taking 32 PCM samples, sampled in time; output is 32 frequency coefficients
- In the Layer 1 encoder, the sets of 32 PCM values are first assembled into a set of 12 groups of 32s
  - an inherent time lag in the coder, equal to the time to accumulate 384 (i.e., 12×32) samples
- Fig.14.11 shows how samples are organized
  - A Layer 2 or Layer 3, frame actually accumulates more than 12 samples for each subband: a frame includes 1,152 samples

Basic Algorithm



Basic Algorithm



▲ *1. Threshold in quiet and masking threshold (acoustical events in the gray areas will not be audible).*

- A masked signal (below the threshold level may consist of low level signal contributions, quantization noise, aliasing distortion or transmission errors …..

Bit Allocation Algorithm

- **Aim**: ensure that all of the quantization noise is below the masking thresholds
- **One common scheme**:
  - For each subband, the psychoacoustic model calculates the *Signal-to-Mask Ratio* (SMR)in dB
  - Then the "Mask-to-Noise Ratio" (MNR) is defined as the difference (as shown in Fig.14.12):

$$MNR_{dB} \equiv SNR_{dB} - SMR_{dB} \tag{14.6}$$

  - The lowest MNR is determined, and the number of code-bits allocated to this subband is incremented
  - Then a new estimate of the SNR is made, and the process iterates until there are no more bits to allocate

Bit Allocation Algorithm

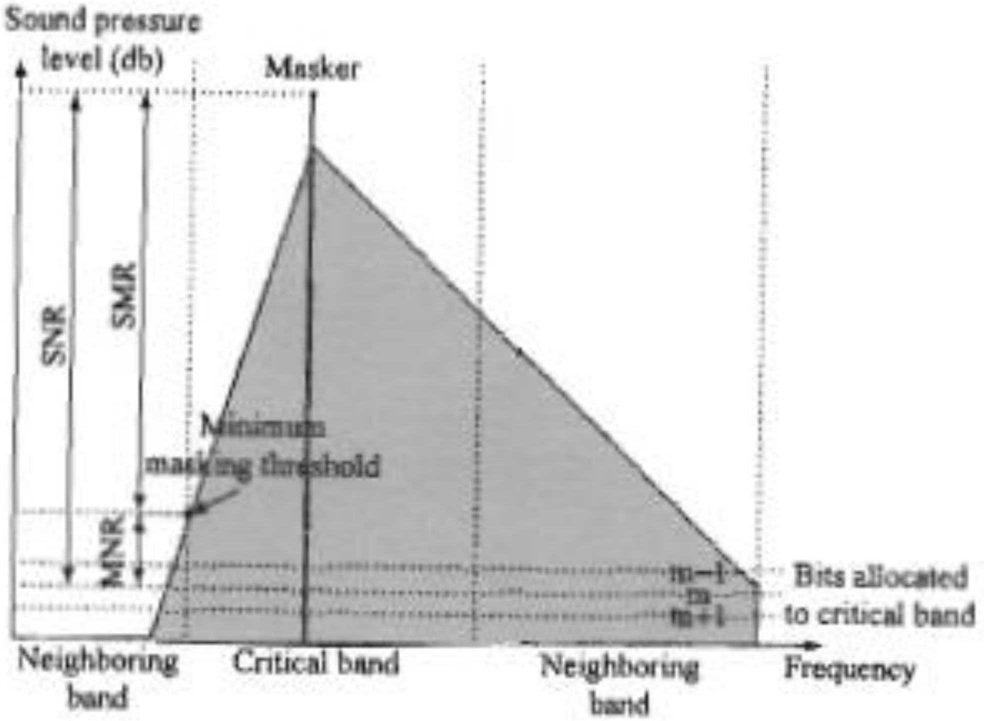
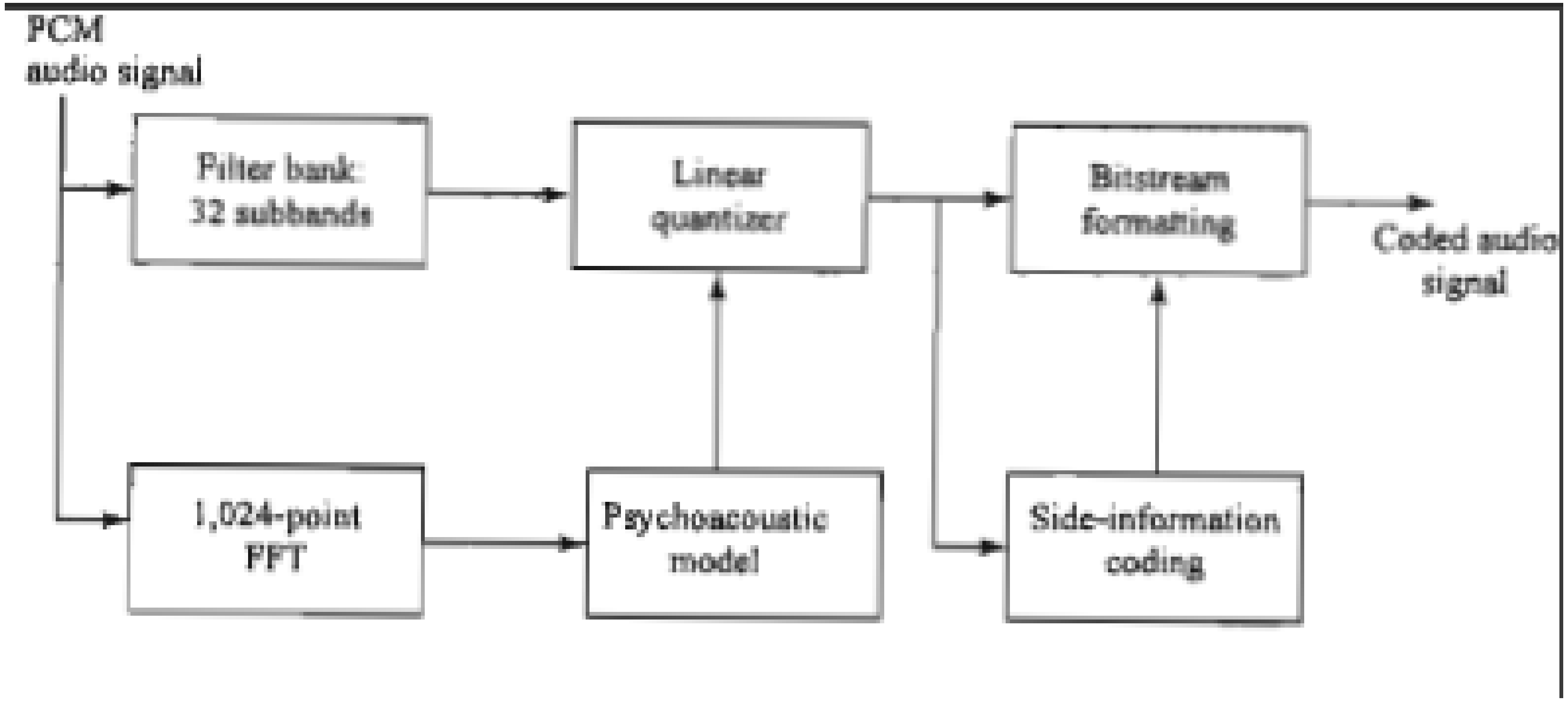


Fig. 14.12: MNR and SMR. A qualitative view of SNR, SMR and MNR are shown, with one dominate masker and m bits allocated to a particular critical band.

MPEG-1 Audio Layers 1 and 2



MPEG-1 Audio Layer 3 coding

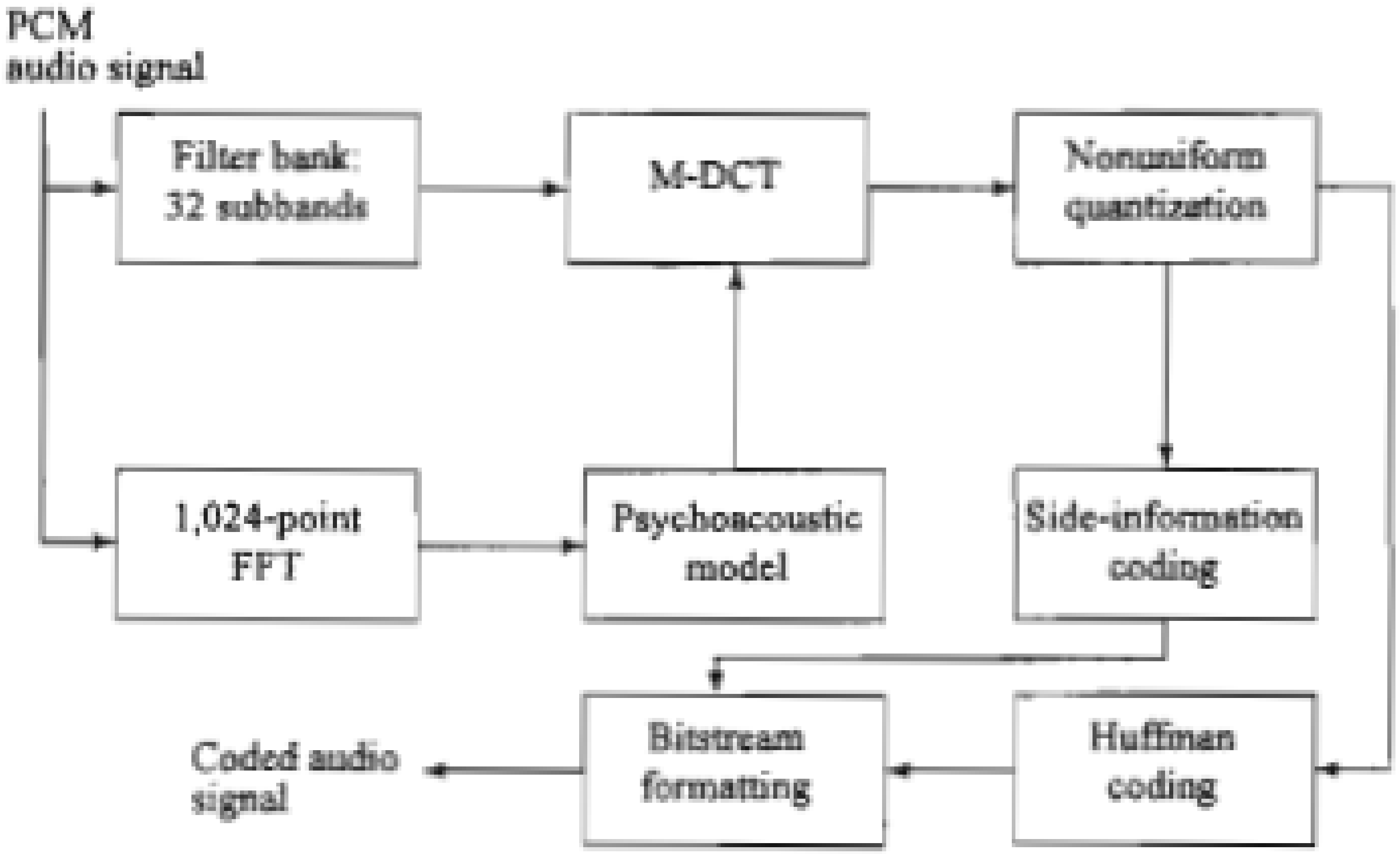
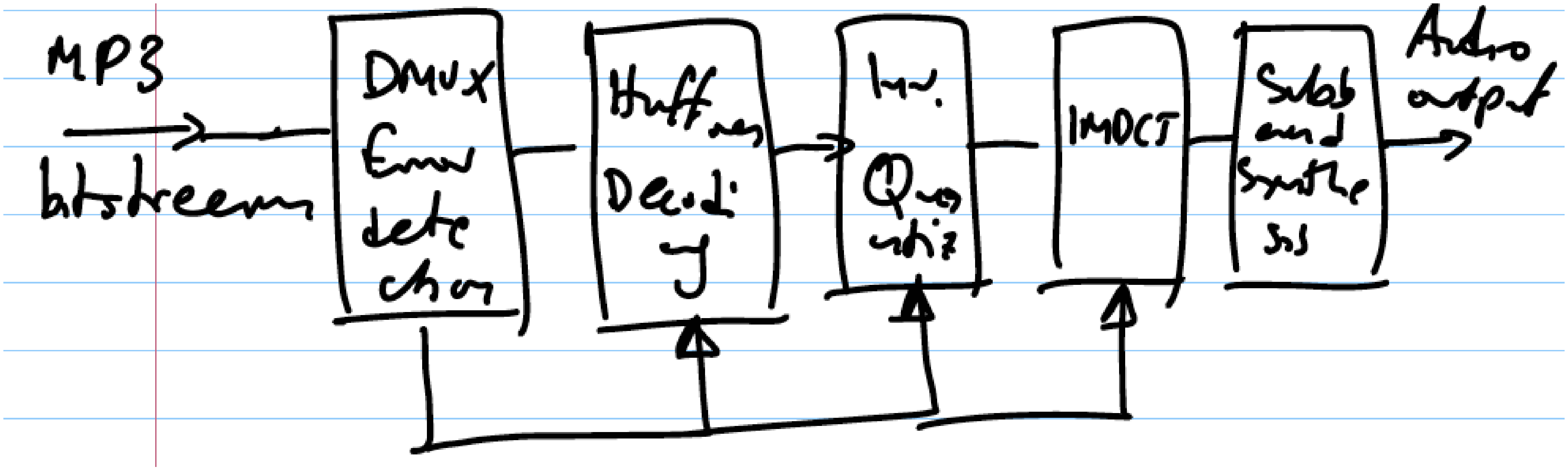


Table 14.2: MP3 compression performance

Sound Quality	Bandwidth	Mode	Compression Ratio
Telephony	3.0 kHz	Mono	96:1
Better than Short-wave	4.5 kHz	Mono	48:1
Better than AM radio	7.5 kHz	Mono	24:1
Similar to FM radio	11 kHz	Stereo	26 - 24:1
Near-CD	15 kHz	Stereo	16:1
CD	>15 kHz	Stereo	14 - 12:1

MPEG Audio Decoder (MP3)



MPEG Audio Decoder (MP3)

- As already mentioned MPEG-audio standard specifies only the decoder, which need to be as simple as possible.
- The sophistication and complexity of the encoder can be lighter as long as it satisfies the decoding standards
- How do we assess sound quality ?
- In general, there are no satisfactory objective tests.
- Listening tests are conducted with a lot of trained listeners.

MPEG Audio Decoder (MP3)

- Evaluation scores (perceptual) :

5 -> transparent

4 -> perceptible, but not annoying

3 -> slightly annoying

2 -> annoying

1 -> very annoying

At low bit rates (60-64 Kbps)

Layer 2 scores between 2.1 to 2.6

Layer 3 scores between 3.6 to 3.8

MPEG-2 AAC (Advanced Audio Coding)

- The standard vehicle for DVDs:
  - Audio coding technology for the DVD-Audio Recordable (DVD-AR) format, also adopted by XM Radio
- Aimed at transparent sound reproduction for theaters
  - Can deliver this at 320 kbps for five channels so that sound can be played from 5 different directions: Left, Right, Center, Left-Surround, and Right-Surround
- Also capable of delivering high-quality stereo sound at bit-rates below 128 kbps

MPEG-2 AAC (Advanced Audio Coding)

- Support up to 48 channels, sampling rates between 8 kHz and 96 kHz, and bit-rates up to 576 kbps per channel
- Like MPEG-1, MPEG-2, supports three different "profiles", but with a different purpose:
  - *Main* profile
  - *Low Complexity*(LC) profile
  - *Scalable Sampling Rate* (SSR) profile