MP3 (MPEG Audio Layer III)

Audio Coding, Psychoacoustic Models and Bitrate Settings

DETI - IC

João Sousa - 103415
João Monteiro - 102690
João Gaspar - 107708

What is MP3?

Created in the 1990s, MP3 it’s a lossy digital audio compression format developed by the MPEG (Moving Picture Experts Group). It reduces the size of audio files and facilitates storage and online distribution.

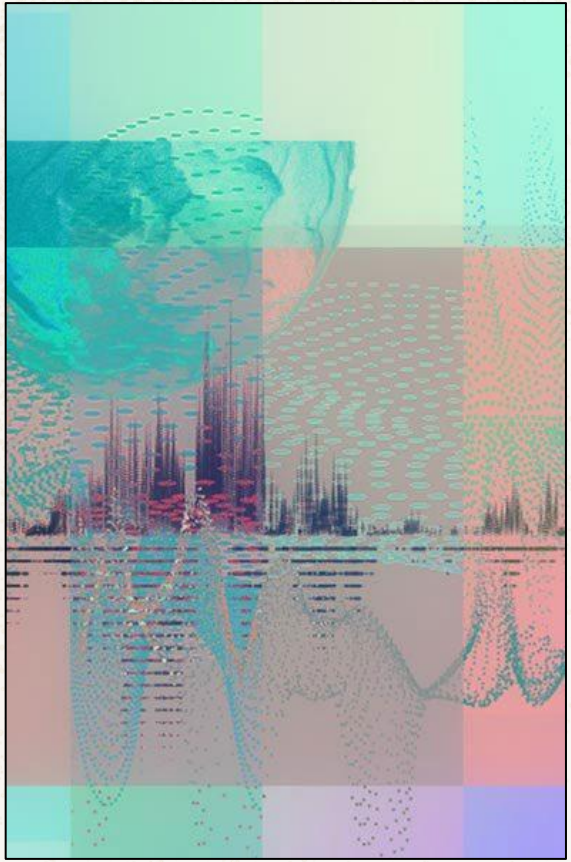
Previously, formats such as WAV took up a lot of space and were not ideal for mass distribution.

Figure 1: Audio Formats Over the Years

The MP3 Revolution

Global impact:

- Democratized access to digital music.
- Reduced physical space needed for music storage.
- Emergence of platforms like Spotify, iTunes, and MP3 Players.



Overview of the Encoding Process

Primary Steps in MP3 Encoding:

- Polyphase Analysis Filter bank
- FFT (Fast Fourier Transform)
- Psychoacoustic Analysis
- MDCT (Modified Discrete Cosine Transform)
- Bit Allocation Loop
- Bitstream Formatting

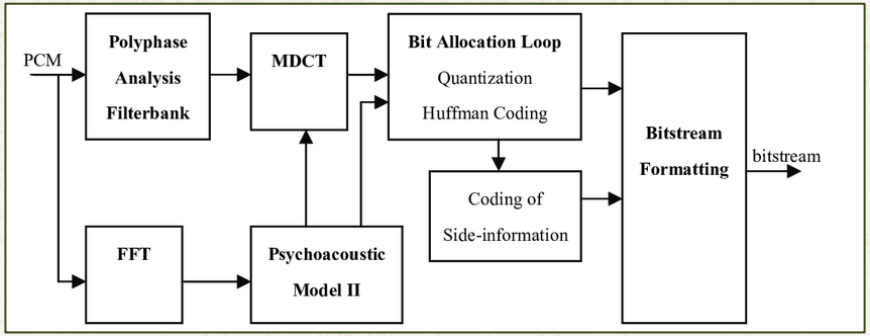


Figure 2: MP3 Compression Process

Step 1: Polyphase Analysis Filter Bank

Splits the input PCM signal into multiple subbands for better frequency resolution.

How It Works:

- Uses polyphase filters to isolate different frequency ranges.
- Prepares subband data for further processing by the MDCT.

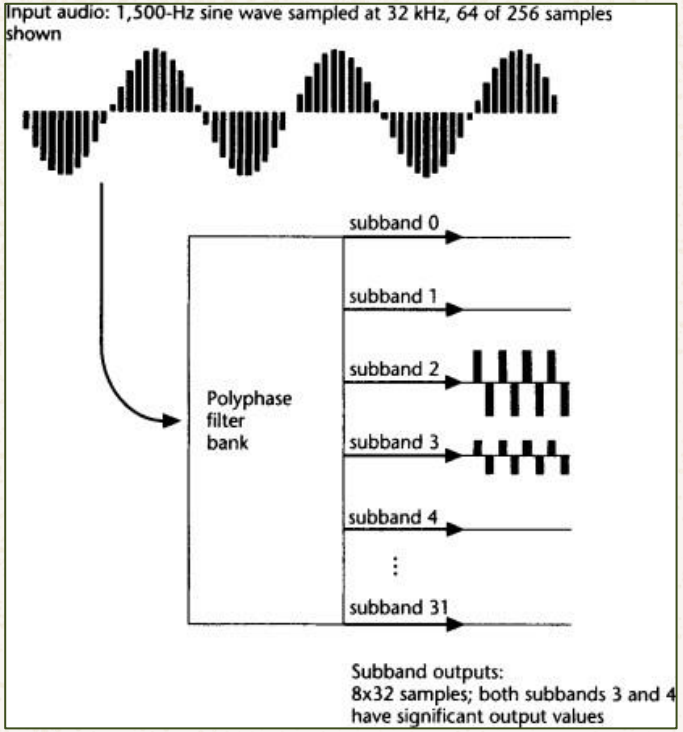


Figure 3: Synthesis Filter Bank

Step 2: FFT and Psychoacoustic Analysis

FFT (Fast Fourier Transform) converts the time-domain signal into a frequency-domain representation for easier analysis.

Psychoacoustic models identify which parts of the audio are less perceptible to the human ear and applies auditory masking effects:

- Frequency masking: Strong sounds mask weaker nearby frequencies.
- Temporal masking: Strong sounds mask weaker sounds occurring just before or after.

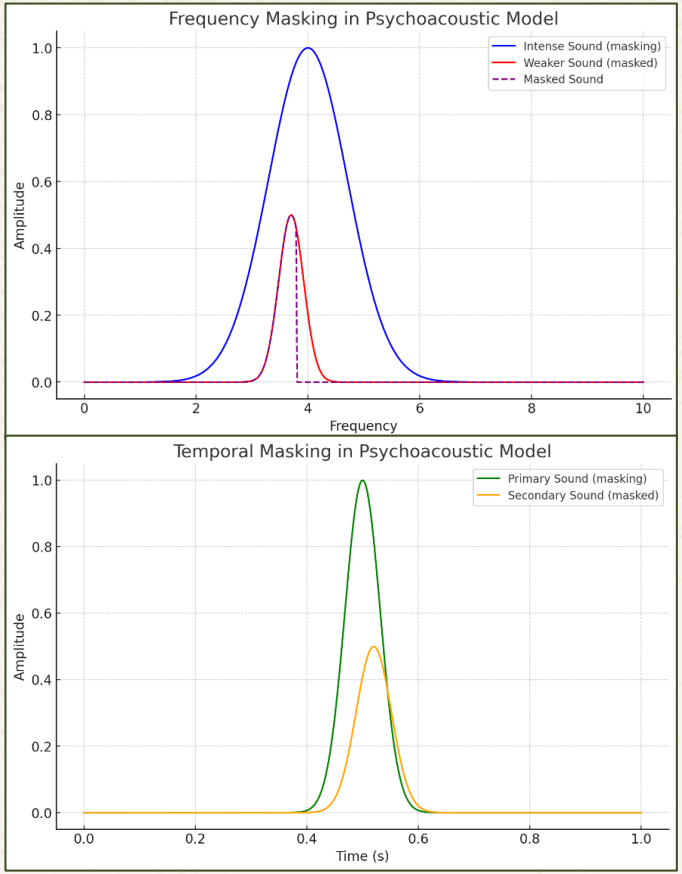


Figure 4: Psychoacoustic Models in MP3

Step 3: How MDCT works?

MDCT transforms subband data into a frequency-domain representation for compression.

First, the continuous audio signal is divided into fixed-size overlapping blocks (e.g. 1024 samples).

Then, the MDCT formula is applied to convert the time-domain data into frequency coefficients.

$$X_k = \sum_{n=0}^{N-1} x_n \cdot \cos\left[\frac{\pi}{N}\left(n + \frac{1}{2} + \frac{N}{2}\right)\left(k + \frac{1}{2}\right)\right]$$

X_k: The k-th MDCT coefficient.

x_n: The input sample at position n.

N: The block size (e.g., 1024 samples).

Figure 5: MDCT Formula

Step 3: Advantages of MDCT

Reduces Temporal Redundancy:

- Exploits the correlation between adjacent samples, minimizing repetitive information.

Reduces Spectral Redundancy:

- Groups frequency coefficients more compactly, optimizing the data representation for further compression.

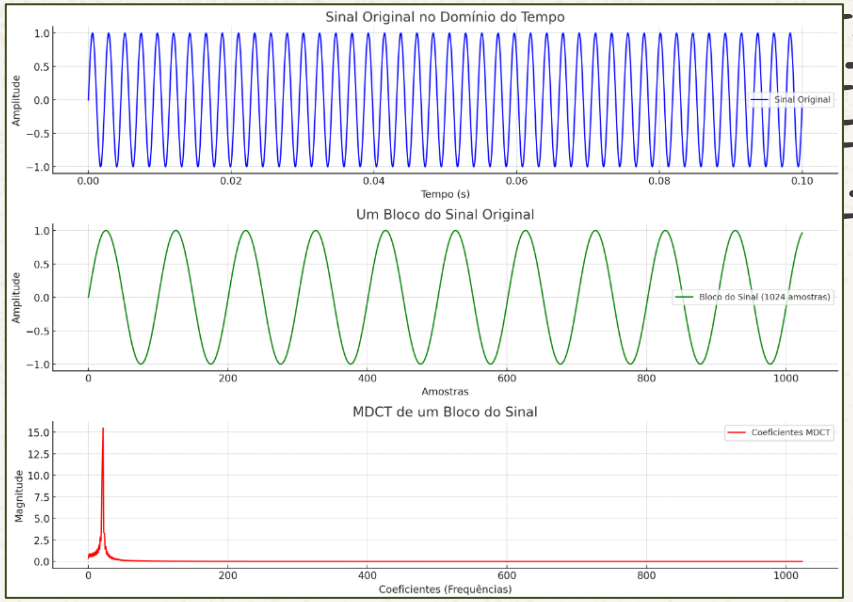


Figure 6: Transformation of an Audio Signal with 440 Hz Frequency

Step 4: Bit Allocation Loop

It’s purpose is to reduces file size further while retaining perceptual quality.

Steps in the Loop:

- Quantization: Approximates frequency coefficients to reduce data.
- Huffman Coding: Compresses data by assigning shorter codes to frequent values.

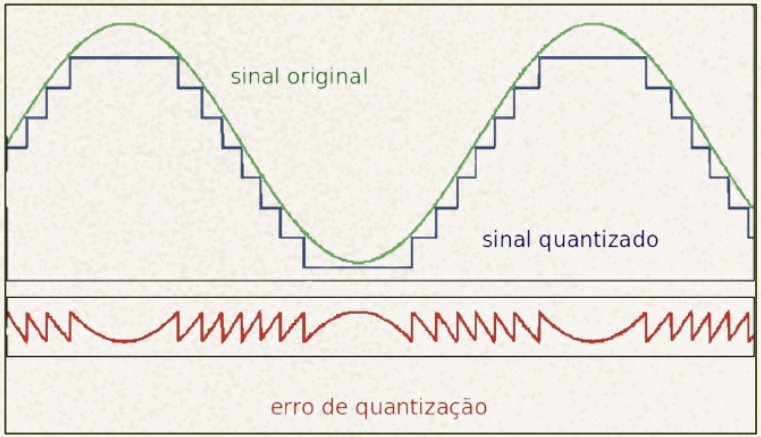


Figure 7: Quantization of an Audio Signal and Possible Errors

Step 5: Bitstream Formatting

It structures the compressed data into a standardized MP3 format.

Includes:

- Encoded audio frames.
- Side information (e.g., bit allocation and scale factors).

Final Output:

- A complete MP3 file ready for playback, storage, or streaming.

Figure 8: MP3 File

Bitrate Configuration

Bitrate refers to the amount of data (in kbps) used to represent one second of audio.

Common configurations:

- Low quality: 96 kbps (compressed audio, noticeable loss).
- Medium quality: 128–192 kbps (balanced).
- High quality: 256–320 kbps (near-transparent).

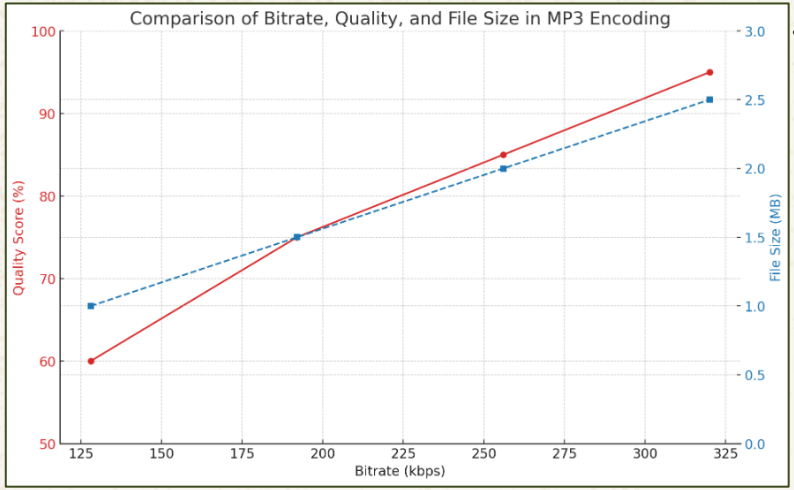


Figure 9: Comparison of Bitrate, Quality, and MP3 File Size

Bitrate Modes: CBR, VBR, and ABR

CBR (Constant Bitrate):

- Fixed rate for the entire file.
- Less flexible, more predictable.

VBR (Variable Bitrate):

- Rate varies based on audio complexity.
- Delivers better overall quality.

ABR (Average Bitrate):

- Combines CBR and VBR for an average value.

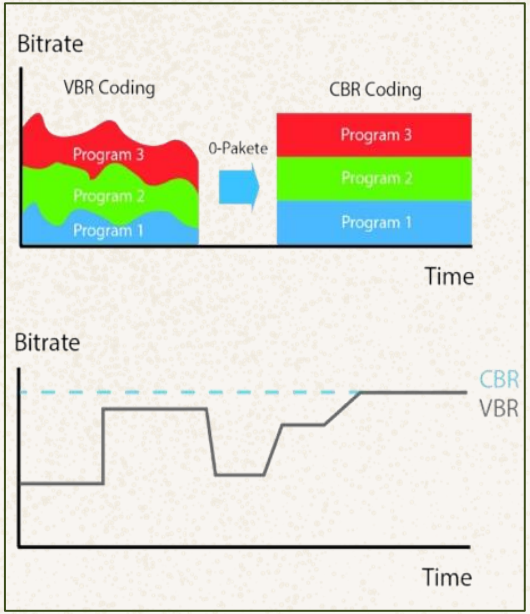


Figure 10: Comparison Between VBR and CBR Methods

MP3 Benefits and Limitations

Benefits:

- High compatibility: Supported by almost all devices.
- Efficient compression: Smaller files without significant quality loss.
- Ease of transmission: Ideal for streaming and quick downloads.

Limitations:

- Lossy compression: Irreversible data loss.
- Reduced quality at low bitrates:
    - Audible artifacts in heavily compressed files.
- Partial replacement: New formats (AAC, Opus) offer better performance.

Conclusion

The MP3 enabled efficient audio compression with minimal quality loss, becoming the standard format of the digital age.

Despite its limitations, the MP3 remains a milestone in the history of digital music and continues to be used.

Thank you!

Questions?

João Sousa - 103415
João Monteiro - 102690
João Gaspar - 107708