



Implementação de Algoritmo de Compressão ISO/MPEG-2 Layer 3 em MATLAB®

Autor: Teodoro Orlow Wey (teowey@gmail.com)

Orientador: Renato da Rocha Lopes

Financiado pelo PIBIC

Palavras-chave: Compressão de arquivos, MP3, Psicoacústica



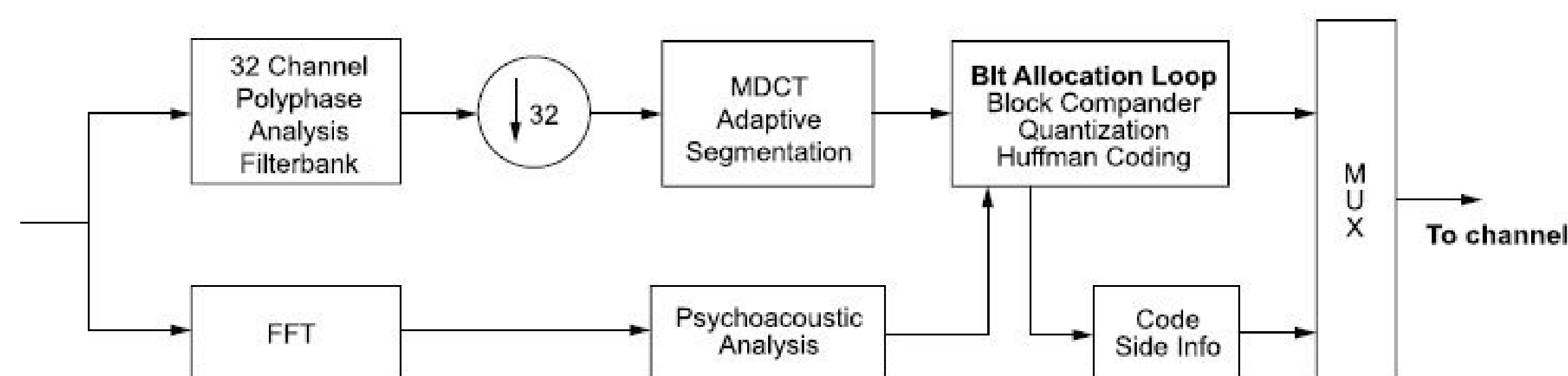
Introdução

Compressão de áudio é uma técnica amplamente aplicada para representar áudio no formato digital em uma quantidade menor de bits, com uma porcentagem de perda que não afete a fidelidade da gravação processada, ou seja, que essa representação não seja distinguível do sinal de áudio original.

A diminuição da quantidade de informação especificada na norma do MPEG-2 Layer III (mais conhecida como *MP3*) implica em economia de espaço de armazenamento, o que faz com que a compressão tenha aplicações em transmissão, armazenamento e outros meios de comunicação que requerem maior rapidez, como a Internet (*streaming*). Com o advento e popularização da rede de computadores, esse se tornou o formato mais barato e acessível, sendo comumente utilizado para codificar música.

O algoritmo

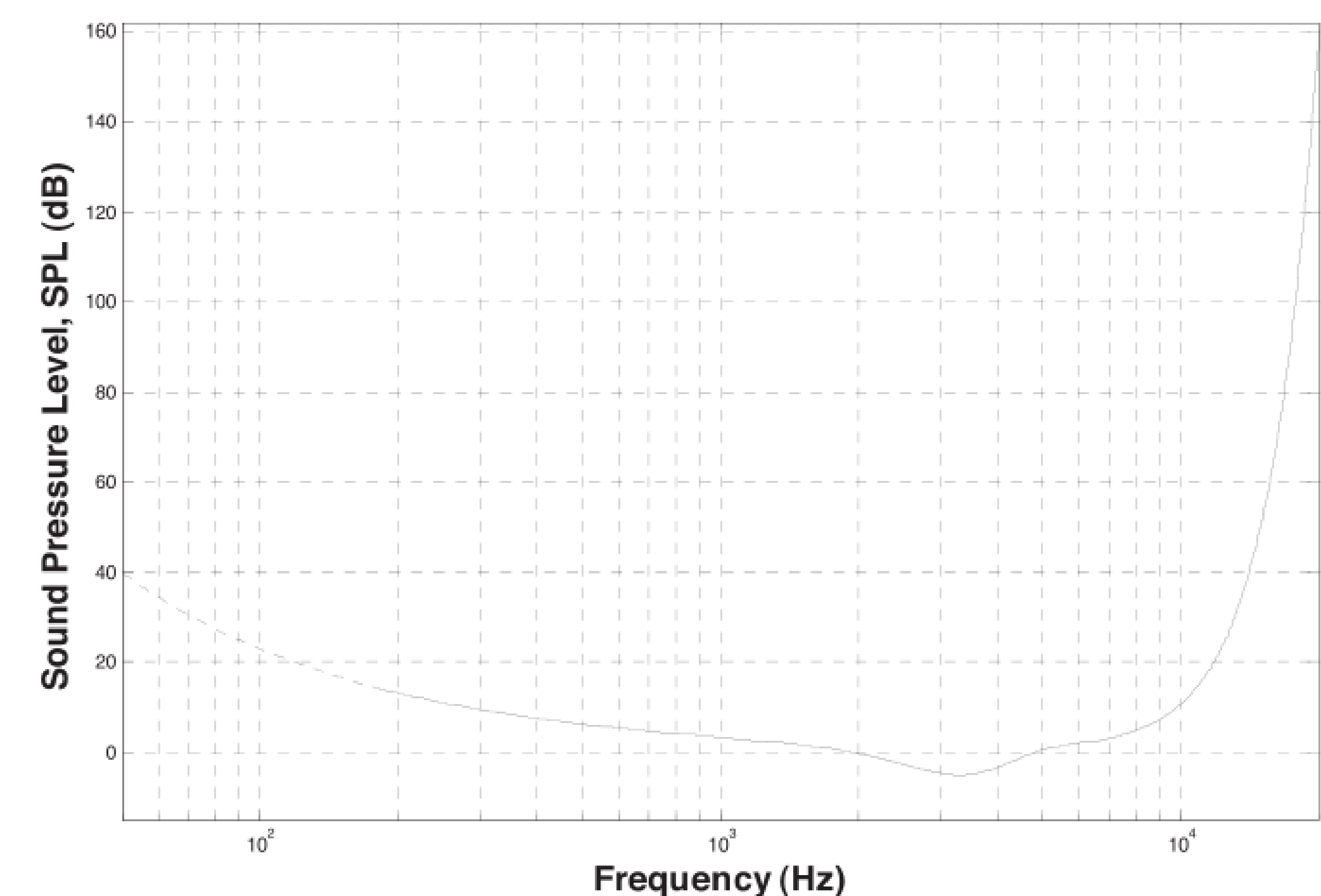
O algoritmo *MPEG-2 Layer III* é separado em 2 fases: Codificação e Decodificação. Sua fase de codificação é aberta para implementação, enquanto que o decodificador deve seguir as funções declaradas no padrão ISO. A Figura abaixo ilustra o fluxograma com as etapas de codificação do sinal de áudio.



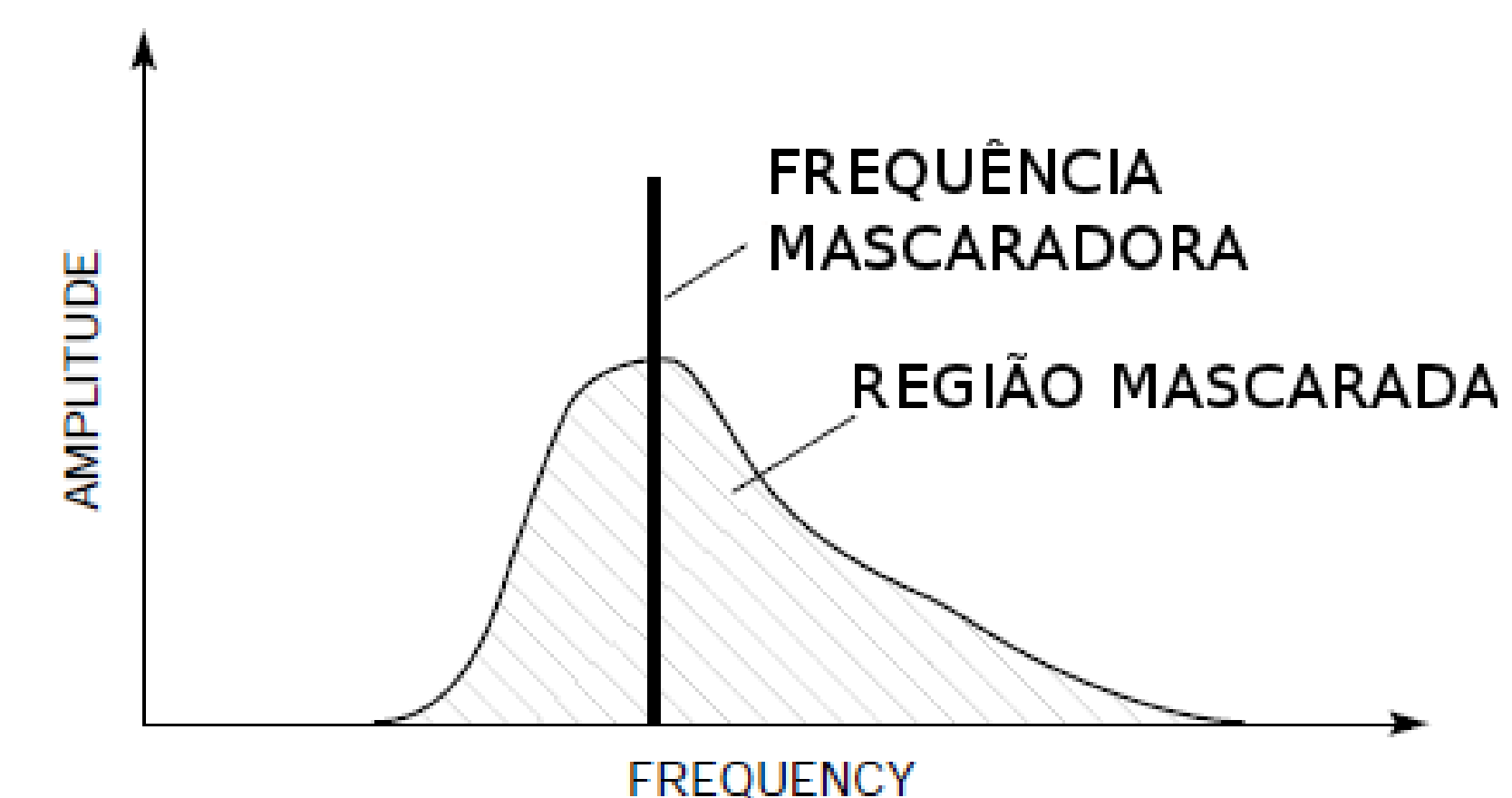
- **Banco de Filtros:** A norma do *MP3* utiliza um banco de filtros para separar o sinal e codificá-lo em sub-bandas, para otimizar a análise e o processamento. O bloco psicoacústico faz uma análise do sinal entrada em função da frequência em função da percepção auditiva humana.
- **MDCT:** A MDCT (*Modified Discrete Cosine Transform*) é uma transformada aplicada para obter uma boa resolução dependendo da quantidade de energia calculada pelo modelo psicoacústico e que vai definir se a transformação de tempo para frequência será em uma janela pequena ou grande.
- **Quantização:** O bloco de quantização e bitstream se encarrega da alocação de bits da informação contida no sinal. São realizados 2 ciclos diferentes sobre os 576 valores de frequência saída do bloco MDCT, *Rate Control Loop* e *Distortion Loop*, que computa os bits necessários para guardar as amostras e checa se ocorreram ruídos no processamento, respectivamente.
- **Decodificação:** O processo de decodificação do *MP3* trata a sequência de bits do bitstream recebido da última etapa do codificador tem três partes principais: decodificação de sequências de bits, requantização e mapeamento frequência-tempo. A primeira parte sincroniza a sequência de bits codificada e extrai os coeficientes de frequência quantificados e qualquer outra informação sobre cada frame.

Modelo Psicoacústico

O sinal digital que entra no bloco psicoacústico é passado para o domínio da frequência através da aplicação de uma FFT. Os dados em frequência seguem para medição dos componentes tonais e não tonais. Segundo estudos em Psicoacústica, sinais sonoros muito fortes mascaram outros sinais mais fracos em frequências próximas, dificultando a audição dos últimos. A figura abaixo com a curva de limiar de audição humano mostra o espectro de frequências audíveis.



Explora-se este fato para ignorar esses sinais mais fracos na compressão, diminuindo assim o número de bits necessários para representar o sinal. A figura a seguir mostra o fenômeno de mascaramento:



Dessa forma, o modelo utilizado para compressão em *MP3* descarta qualquer frequência que possua energia inferior ao limiar global obtido para os blocos de amostras de áudio. Esses blocos são transmitidos para serem quantizados, determinando a alocação de bits apropriada para armazenar a informação. Logo, no final dessa etapa há uma perda significativa de conteúdo sonoro, o que significa uma qualidade pior em termos de áudio (algoritmos com perda na compressão) enquanto economiza espaço de armazenamento, sendo essa versatilidade na memória a grande vantagem dos arquivos de música em *MP3*.

Referências

- [1] M. Ruckert (2005) Understanding MP3
- [2] P.P. Vaidyanathan (1993) Multirate Systems and Filter Banks
- [3] J. J. Thiagarajan and A. Spanias (2011) Analysis of the MPEG-1 Layer 3 (MP3) Algorithm using MATLAB
- [4] T. Painter and A. Spanias (2008) Perceptual Coding of Digital Audio *Proceedings IEEE*