

Compressão MP3

Autores: Anderson Braga de Avila, Felipe Leivas Teixeira
Patrícia Teixeira Davet e Thiago Ferreira Pontes

Professor: Marilton Sanchotene de Aguiar

Instituição: UFPEL – PPGC

Junho de 2014

- Introdução
- Conceitos
- Processo de Compressão MP3
- Processo de Descompressão MP3
- Referências

□ Formato WAV

- Formato padrão para arquivo de som em PCs
- Ocupa em média 10 MB para armazenar um minuto de música, em uma gravação de som de 16 bits estéreo com 44.1 KHz

□ Surgimento MP3

- Formato que permitiu armazenar músicas no computador de forma a ocupar menos espaço e com a mesma qualidade sonora
- Um minuto de música, corresponde a cerca de 1 MB em MP3
- Tornou-se um popular meio de distribuição de músicas pela internet

- MP3 – MPEG-1 Layer 3

- MPEG (*Moving Picture Experts Group*)
 - Comissão técnica criada em 1988 pela ISO e IEC, encarregada de estabelecer normas para a representação de áudio, imagem e vídeo.

- MPEG-1
 - Codifica Áudio e Vídeo a uma taxa de 1,5Mbps
 - Define 3 níveis de compressão
 - Nível mais conhecido: layer-3
 - MP3 –MPEG-1 Layer 3

- MPEG-2
 - Broadcast
 - Pode ser utilizado em transmissões a taxas de 4 a 9 Mbps
 - Uma versão modificada do MPEG-2 é usada pelo padrão HDTV e também nos DVDs.
- MPEG-3
 - Formato morto, pois o formato MPEG-4 o suplantou com muitas vantagens. Associado a uma taxa de 60 Mbps (mais tarde reduzido para 40 Mbps e então cancelado).
- MPEG-4
 - Compressão de Vídeo
 - Derivados: DivX, XViD
 - Podem oferecer qualidade semelhante à MPEG-2 sem ocupar tanto espaço

■ MPEG-1 Layers (Camadas)

	TC
Layer I	1:4
Layer II	1:6 .. 1:8
Layer III (MP3)	1:10 .. 1:12

- Compressão MP3 - MPEG-1 Layer3
 - Terceira camada do formato MPEG-1
 - Destinada a codificação de áudio
 - Taxa de compressão: 10:1 podendo chegar até 12:1
 - Uso de elementos de psicoacústica

- Codec (COder-DECoder)
- Frames
 - Forma de dividir um arquivo de áudio
 - Uma unidade de codificação
 - **MP3 -> 1152 amostras PCM** □
- Bitrate(taxa de bits)
 - Número de bits usados para codificar um frame
 - Medida em kbps
 - CBR (constant bitrate) -A mesma quantidade de bits para representar qualquer frame (WAV)
 - ABR (average bitrate) -Frames menos complexos deixam bits para os próximos (MP3)
 - VBR (variable bitrate) -O número de bits para cada frame deve garantir a qualidade definida no início da codificação (AAC, VORBIS)

□ Psicoacústica

- O ramo da psicoacústica estuda o modo que o ser humano escuta o som. Utilizando os conceitos deste ramo, a tecnologia MP3 suprime o som original sons que estão aquém da percepção humana. Essa técnica não causa perda substancial da qualidade do áudio. Os princípios de supressão são:
 - Faixa de frequência audível dos seres humanos
 - Limiar de audição na faixa de frequência audível
 - Mascaramento em frequência e mascaramento temporal

- Faixa de frequência audível dos seres humanos
 - Faixa de frequência audível varia de 20 Hz a 20 KHz
 - Desta forma a compressão MP3 não armazena dados referentes a sons fora desta faixa de frequência, pois ao serem reproduzidos, os mesmos não serão percebidos por um ser humano

- Limiar de audição na faixa de frequência audível
 - Outro fator utilizado pela compressão MP3 é a curva de percepção da audição humana dentro da faixa de frequências audíveis, ou Limiar de Audição.
 - A sensibilidade para sons dentro desta faixa não é uniforme. Ou seja, a percepção da intensidade de um som varia com a frequência em que este se encontra
 - A compressão MP3 descarta amostras que se encontrem abaixo deste limiar, pois não são perceptíveis ao ouvido humano

- Mascaramento em frequência e mascaramento temporal
 - O mascaramento em frequência ocorre quando um som que normalmente poderia ser ouvido é mascarado por outro, de maior intensidade, que se encontra em uma frequência próxima
 - O mascaramento em frequência depende da frequência em que o sinal se encontra, podendo variar de 100 Hz a 4 KHz

- Mascaramento em frequência e mascaramento temporal
 - Um sinal forte num dado instante torna inaudíveis sinais fracos em janelas temporais imediatamente anteriores (pré-mascaramento) ou posteriores (pós-mascaramento)
 - O pré-mascaramento existe só por um curto momento, cerca de 20ms, enquanto que o pós-mascaramento tem efeito por até 200 ms

Processo de Compressão MP3

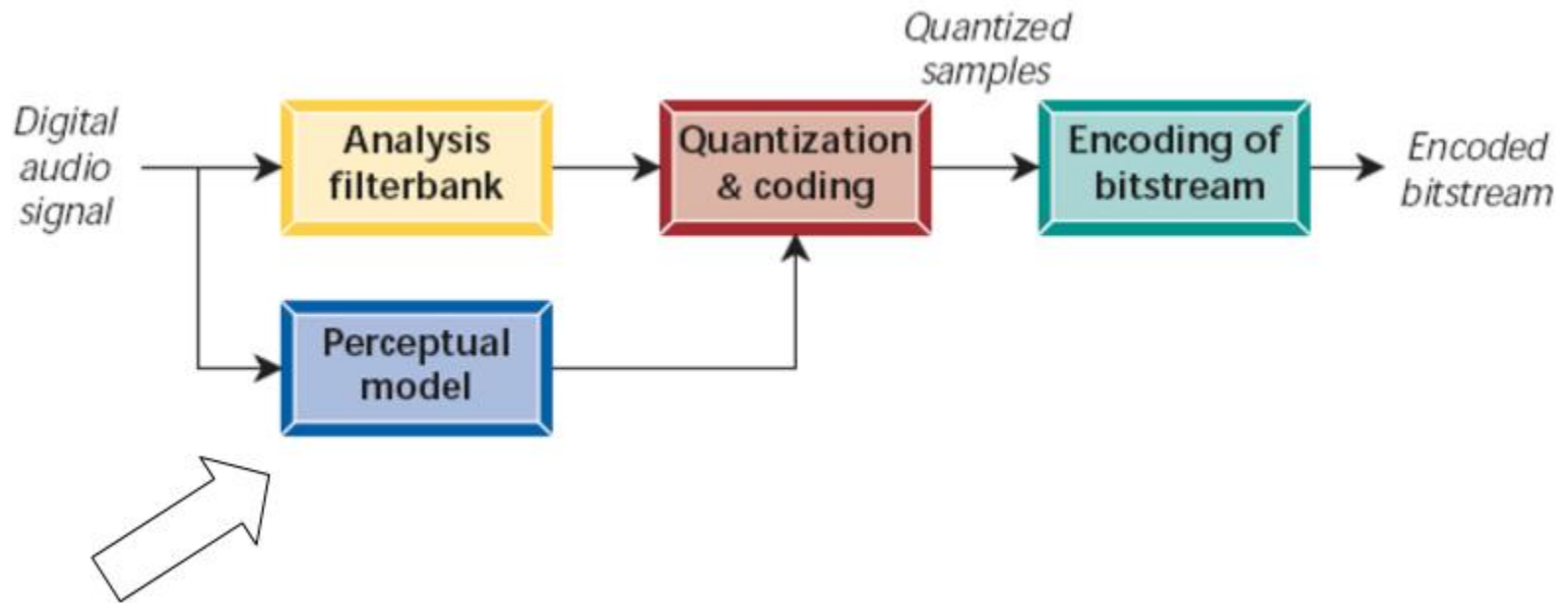


Diagrama em Blocos do Processo de Compressão MP3

Processo de Compressão MP3

Modelo Perceptual

■ Objetivo

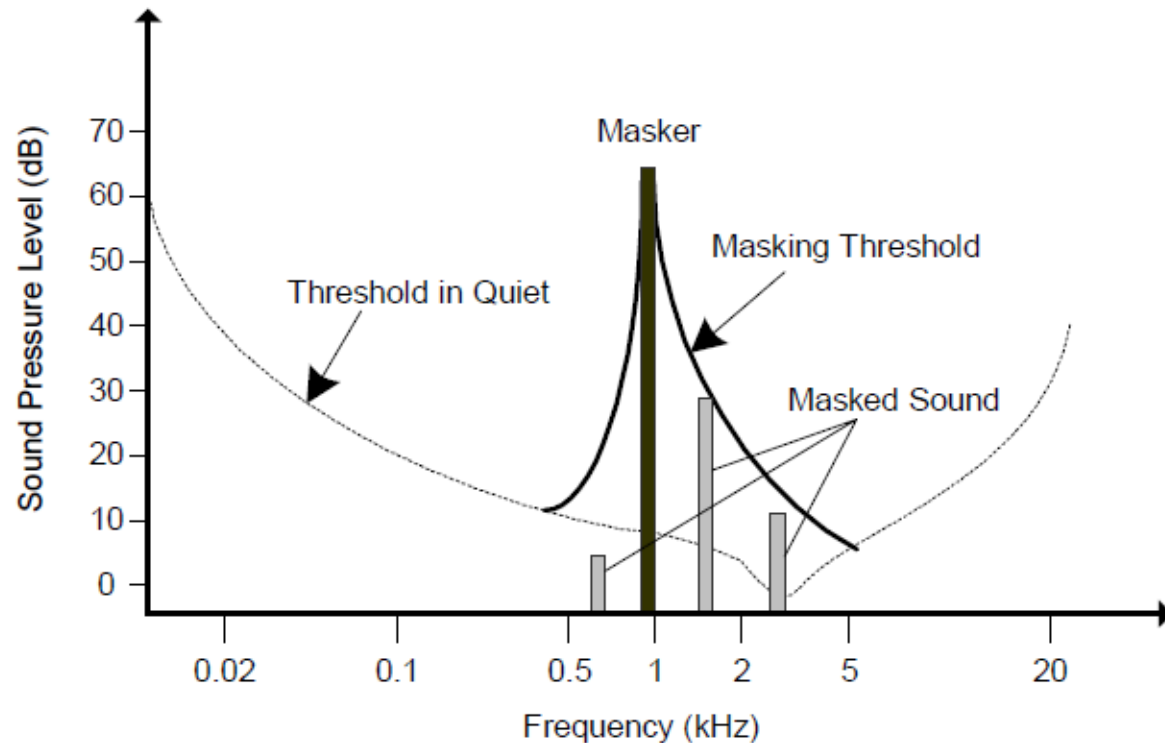
- Analisar o que pode ser considerado redundância

■ Critérios

- Limiar de audibilidade
- Mascaramento de sons

Processo de Compressão MP3

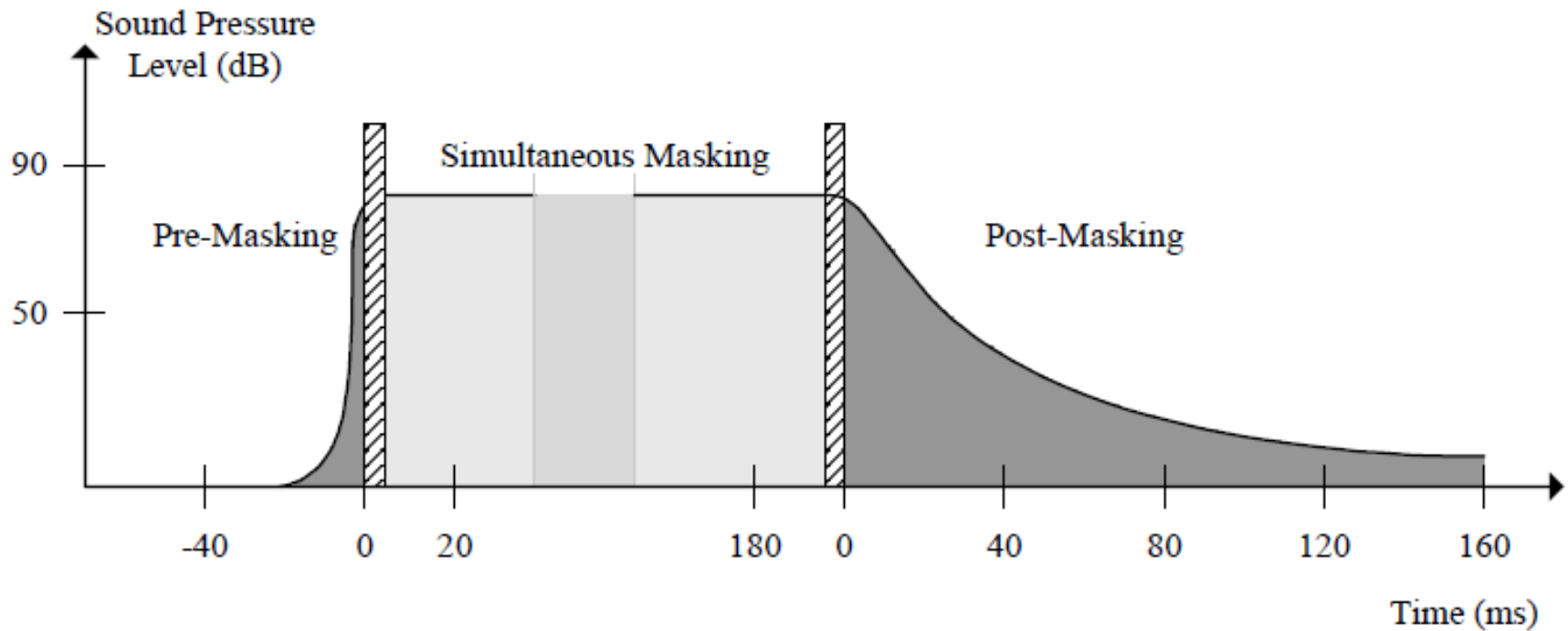
Modelo Perceptual



Limiar de audibilidade e Mascaramento em frequência

Processo de Compressão MP3

Modelo Perceptual



Mascaramento Temporal

Processo de Compressão MP3

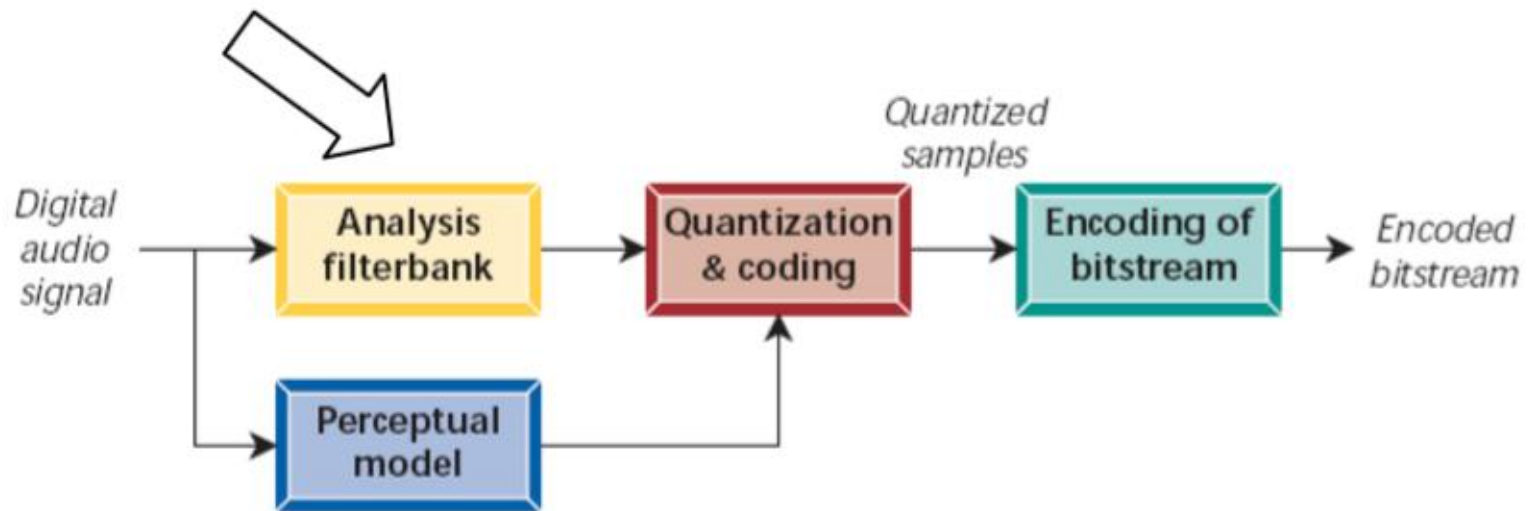


Diagrama em Blocos do Processo de Compressão MP3

Processo de Compressão MP3

□ Analysis Filterbank

- Amostragem do sinal de áudio de acordo com o Teorema de Nyquist. Este diz que para garantir a reprodução de um sinal, temos de amostrá-lo pelo menos a duas vezes sua frequência máxima
- Ou seja, neste caso, como a frequência máxima de interesse é 20 KHz, basta amostrar a 40 KHz

Processo de Compressão MP3

Analysis Filterbank

- Utiliza-se 44.100 Hz como taxa de amostragem, pois leva-se em consideração uma tolerância e busca-se um valor produto dos quatro primeiros números primos ao quadrado
 - Obs: $(2 \times 3 \times 5 \times 7)^2 = 44100$

- Desta forma, esta taxa de amostragem funciona como um filtro passa-baixa, que remove todos os componentes de frequência fora da faixa de interesse, neste caso, acima de 20 KHz

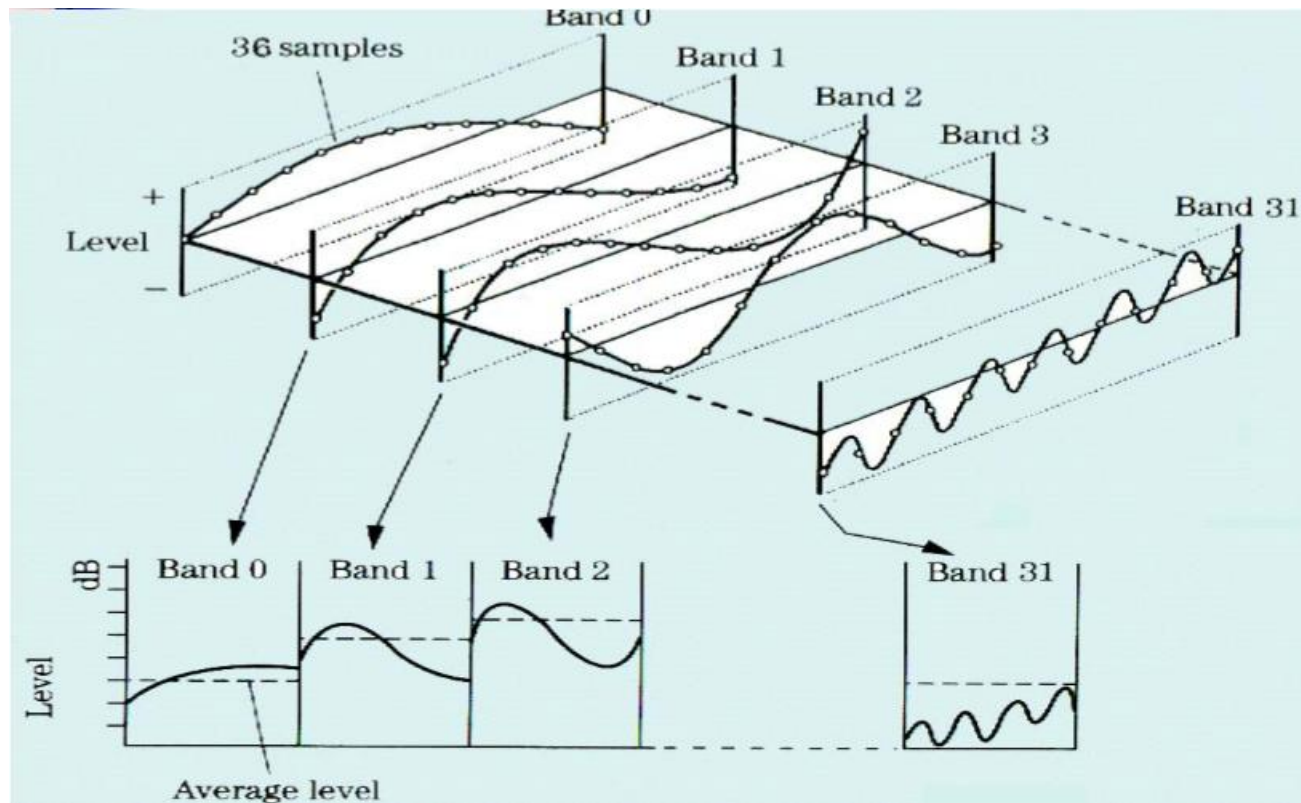
Processo de Compressão MP3

Analysis Filterbank

- Uma sequência de 1152 amostras PCM são filtrados por um banco de filtros passa-banda que divide o sinal de áudio em 32 sub-bandas de igual largura espectral ($22.050/32 = 689$ Hz)
- □ banco de filtros e o seu inverso não são transformações sem perdas. Mesmo sem quantização, a transformação inversa não pode perfeitamente recuperar o sinal original. No entanto, pelo projeto o erro introduzido pelo banco de filtros é pequeno e inaudível

Processo de Compressão MP3

Analysis Filterbank



Filtragem em Sub-bandas

Processo de Compressão MP3

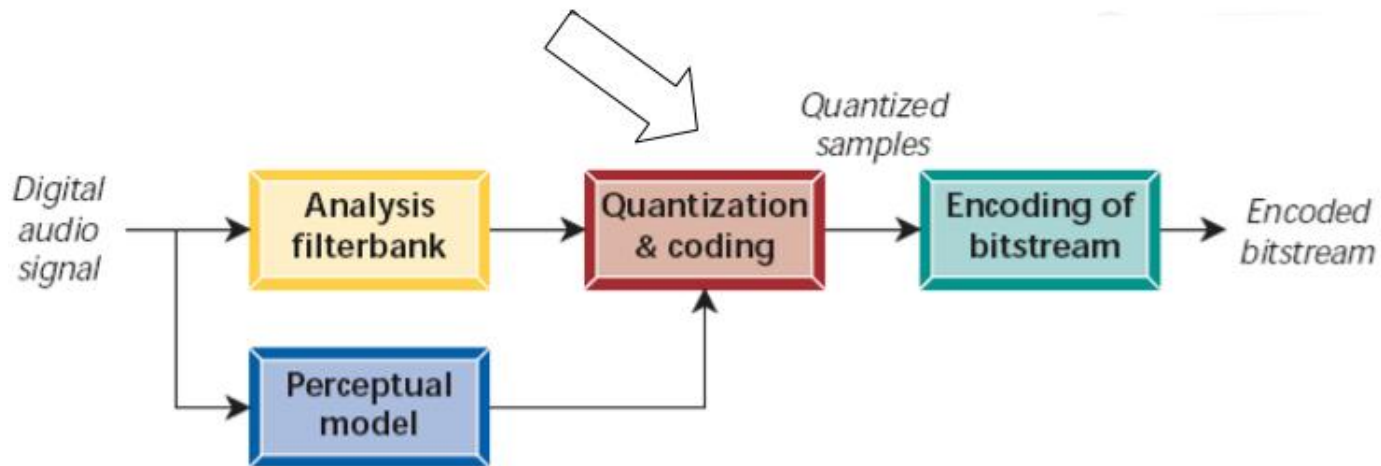


Diagrama em Blocos do Processo de Compressão MP3

Processo de Compressão MP3

Quantização e Codificação

□ Quantização e Codificação

■ Quantização

- As informações irrelevantes (detectadas na fase de modelagem perceptual) são efetivamente retirada
- O número de bits para cada banda é determinado

■ Codificação

- Dissociação inter-canaís (caso estéreo)
 - Os Canais E e D são transformados em média ($E+D/2$) e side ($E-D/2$)
 - Codificação de Huffman

Processo de Compressão MP3

Codificação Huffman

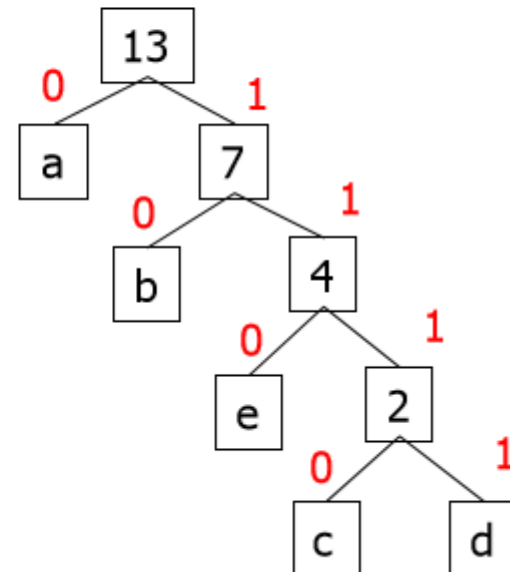
- Mensagem

- aaaaaabbcbdee
- 104 bits

símbolo	ocorrência	código
a	6	0
b	3	10
c	1	1110
d	1	1111
e	2	110

- Mensagem comprimida

- 00000010101011101111110110
- 26 bits. TC = 75%



Processo de Compressão MP3

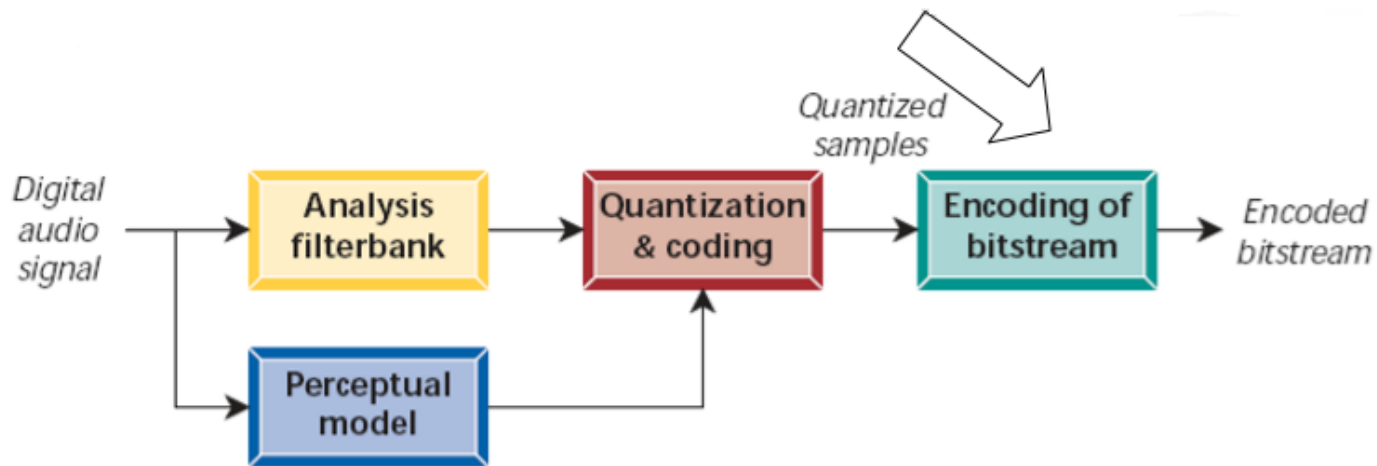
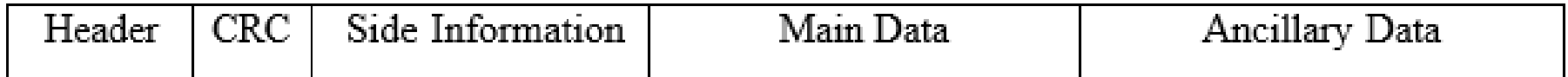


Diagrama em Blocos do Processo de Compressão MP3

Processo de Compressão MP3

Codificação do bitstream

□ Codificação do bitstream



Processo de Compressão MP3

Codificação do bitstream

□ Header (32)

■ Sync (12-11)

□ Todos bits setados.

■ ID (1-2)

□ Especifica a versão do MPEG.

00	MPEG-2.5 (Later extension of MPEG-2)
01	Reserved
10	MPEG-2
11	MPEG-1

	Sync		
ID	Layer	Prot. bit	
Bitrate			
Frequency	Pad. bit	Priv. bit	
Mode	Mode extesion		
Copy Home	Emphasis		

Processo de Compressão MP3

Codificação do bitstream

Header (32)

■ Layer (2)

00	reserved
01	Layer III
10	Layer II
11	Layer I

■ Protection bit (1)

Se '1', usar CRC.

■ Bitrate (4)

Informa o decodificador em qual bitrate o frame foi codificado;

Mesmo valor em todos os frames, se CBR.

ID	Layer	Prot. bit
Bitrate		
Frequency	Pad. bit	Priv. bit
Mode	Mode extesion	
Copy Home	Emphasis	

Processo de Compressão MP3

Codificação do bitstream

bits	MPEG-1, layer I	MPEG-1, layer II	MPEG-1, layer III	MPEG-2, layer I	MPEG-2, layer II	MPEG-2, layer III
0 0 0 0						
0 0 0 1	32	32	32	32	8	
0 0 1 0	64	48	40	64	48	16
0 0 1 1	96	56	48	96	56	24
0 1 0 0	128	64	56	128	64	32
0 1 0 1	160	80	64	160	80	64
0 1 1 0	192	96	80	192	96	80
0 1 1 1	224	112	96	224	112	56
1 0 0 0	256	128	112	256	128	64
1 0 0 1	288	160	128	288	160	128
1 0 1 0	320	192	160	320	192	160
1 0 1 1	352	224	192	352	224	112
1 1 0 0	384	256	224	384	256	128
1 1 0 1	416	320	256	416	320	256
1 1 1 0	448	384	320	448	384	320
1 1 1 1						

Processo de Compressão MP3

Codificação do bitstream

Header (32)

■ Frequency (2)

Bits	MPEG1	MPEG2	MPEG2.5
00	44100 Hz	22050 Hz	11025 Hz
01	48000 Hz	24000 Hz	12000 Hz
10	32000 Hz	16000 Hz	8000 Hz
11	reserv.	reserv.	reserv.

Frequency	Pad. bit	Priv. bit
Mode	Mode extension	
Copy Home	Emphasis	

■ Padding bit (1)

Usado para informar que o frame tem ou não um bit extra.

■ Private bit (1)

Processo de Compressão MP3

Codificação do bitstream

□ Header (32)

■ Mode (2)

00	Stereo
01	Joint Sereo
10	Dual Channel
11	Single Channel

Mode		Mode extesion	
Copy	Home	Emphasis	

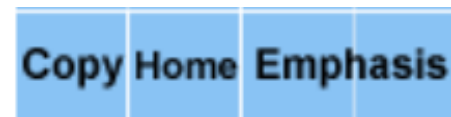
■ Mode extension (2)

Bits	Intensity stereo	MS stereo
00	Off	Off
01	On	Off
10	Off	On
11	On	On

Processo de Compressão MP3

Codificação do bitstream

Header (32)



■ Copyright bit (1)

Se `1`, é ilegal copiar o conteúdo.

■ Home bit (1)

Se `1`, o frame está em sua mídia original.

■ Emphasis (2)

Informa o modelo de supressão de ruído a ser usado na descompressão.

00	None
01	50/15 ms
10	Reserved
11	CCITT J.17

Processo de Compressão MP3

Codificação do bitstream

CRC (0-16 bytes)

- Existirá se o bit de proteção for setado;
- Verifica o frame por erros de transmissão;
- Caso seu valor seja incorreto, o decodificador irá corromper o frame;
- Soluções:
 - Não processar o frame (mute);
 - Substituí-lo pelo frame anterior.

CRC	Side Information	Main Data	Ancillary Data
-----	------------------	-----------	----------------

Processo de Compressão MP3

Codificação do bitstream

□ Side Information (17,32 bytes)

- Consiste de informações para a decodificação do Main Data;
- Depende da codificação dos canais;
 - 17 bytes, se single channel;
 - 32 bytes, caso contrário.



Processo de Compressão MP3

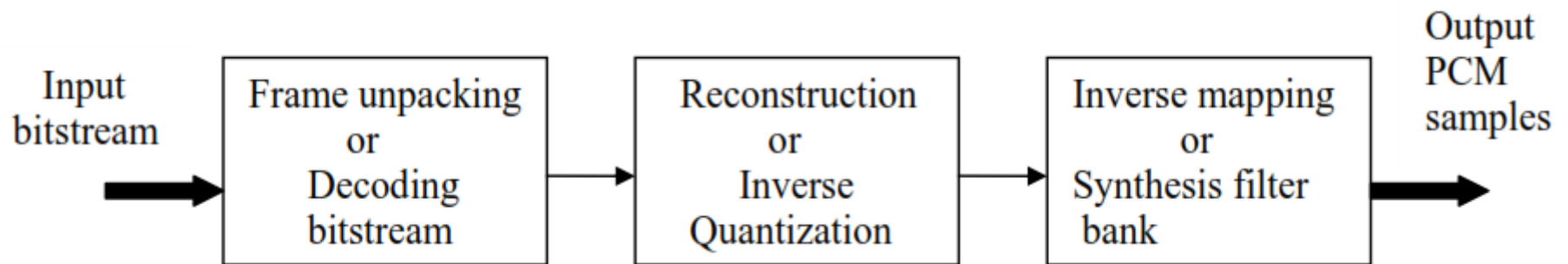
Codificação do bitstream

□ Main Data

- Consiste dos fatores de escalonamento, dos bits codificados por Huffman e de informações auxiliares;



Processo de Descompressão MP3



Processo de Descompressão MP3

Desempacotamento do Frame

- ❑ Encontrar o cabeçalho do bitstream;
- ❑ Decodificar informação adicional;
- ❑ Decodificar os fatores de escalonamento;
- ❑ Decodificar os dados (Huffman).

Processo de Descompressão MP3

Reconstrução

- Requantização;
- Reordenamento do Espectro.

Processo de Descompressão MP3

Mapeamento Inverso

- ❑ Processamento de joint stereo;
- ❑ Redução de alias;
- ❑ Síntese via IMDCT;
- ❑ Banco de filtragem polifásica.

Referências

- [1] SRIPADA, Praveen . **MP3 DECODER in Theory and Practice**. Tese de Mestrado. Department of Signal Processing and Telecommunications . Blekinge Institute of Technology, 2006.
- [2] RAISSI, Rassol. **The Theory Behind Mp3**. December 2002.
- [3] YANG, Xiang; WAXMAN, Corey. **MP3 Coding Scheme**. University of Pennsylvania.
- [4] MATTOS, Carlos Ignácio de. **MPEG Layer 3 – MP3**. Departamento de Engenharia Elétrica. UFPR, 2007.
- [5] MELLO, Carlos Alexandre. **Técnicas de Codificação de Áudio e Vídeo**. Centro de Informática. UFPE.

Compressão MP3

Autores: Anderson Braga de Avila, Felipe Leivas Teixeira
Patrícia Teixeira Davet e Thiago Ferreira Pontes

Professor: Marilton Sanchotene de Aguiar

Instituição: UFPEL – PPGC

Junho de 2014
