

# Sistemas Multimídia

## O Áudio (Parte IV)



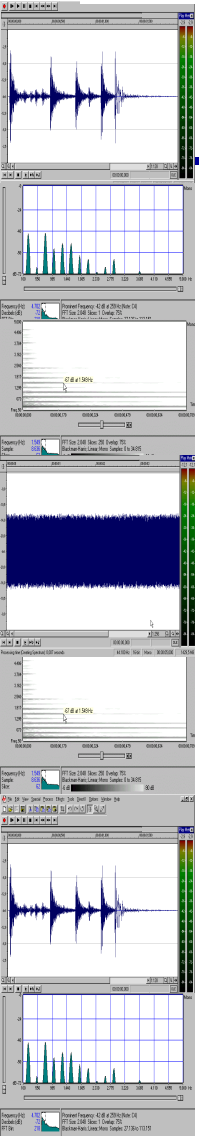


# Sistemas Multimídia

---

## Tópico:

- Aspectos relevantes da compressão de áudio

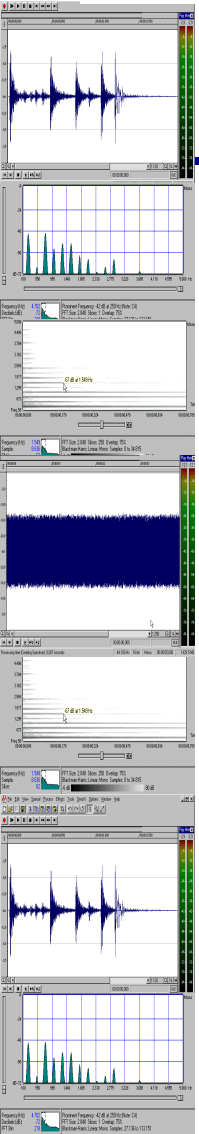




# Compressão de Áudio

---

- Qual a largura de banda necessária para transmitir áudio codificado a uma frequência de amostragem de 44.1 kHz, 16 bits por amostra, estéreo (qualidade de CD)?





# Compressão de Áudio

---

$44100 \times 16 \times 2 = 1.411.200$  bits por segundo.

- Isto significa que para transmitir tal arquivo por uma rede, é necessária uma largura de banda de **1,41 Mbits/s**.

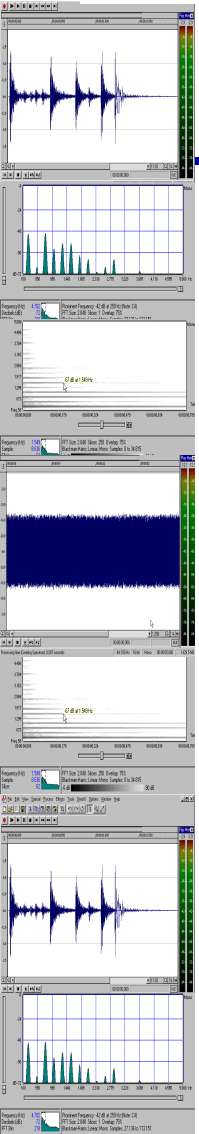




# Compressão de Áudio

---

- Uma música de três minutos (180 segundos) de áudio nestas condições (frequência de amostragem de 44.1 kHz, 16 bits por amostra, estéreo) utiliza quanto (MB) de armazenamento em um computador?





# Compressão de Áudio

---

$180 \times 44100 \times 16 \times 2 = 254.016.000$  bits

- ou mais de 30 MB de armazenamento no computador.
- Estes dois exemplos dão uma idéia da importância da compressão de som.

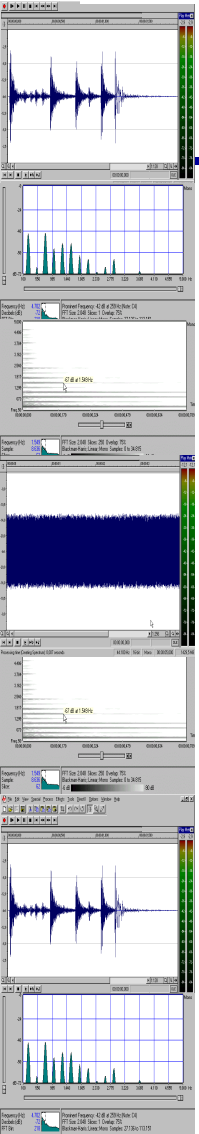




# Compressão de Áudio

---

- O sistema MP3 aproveita, além das técnicas habituais de compressão, o conhecimento das imperfeições ou limitações na audição, para eliminar certas informações sem afetar o que ouvimos, conseguindo níveis de compressão de até 12 vezes.



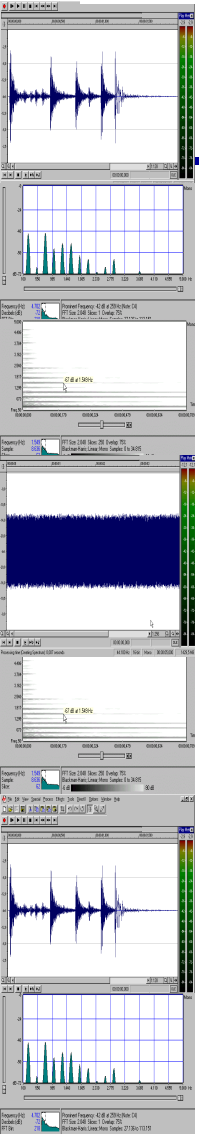


# Compressão de Áudio

---

## Características da Audição Humana

- A audição humana não é perfeita. Além das limitações físicas do ouvido, depois o som tem que viajar através dos nervos até o córtex auditivo do cérebro onde é transformado em diferentes percepções das quais tomamos consciência.





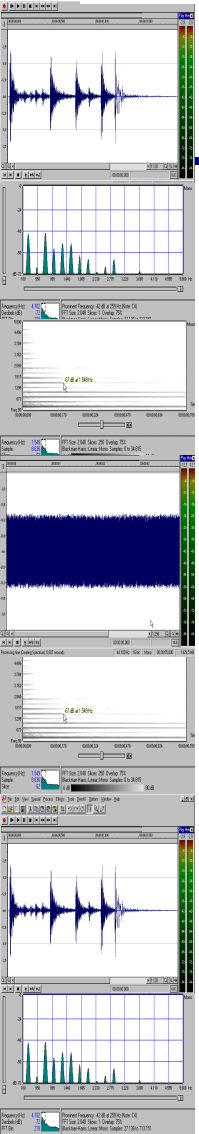


# Compressão de Áudio

## Características da Audição Humana

### ▣ Sonoridade (*Loudness*):

- Dois sons com a mesma amplitude podem ser percebidos com intensidade distinta dependendo das freqüências que têm.
- A percepção da intensidade de um som não é constante com a freqüência.
- O ouvido humano tem maior sensibilidade ao som entre os 1000 e os 5000 Hz.





# Compressão de Áudio

---


## Características da Audição Humana

### □ Alcance de Freqüências

- O ser humano pode perceber sons na faixa de freqüências de 20 Hz a 20 kHz devido às limitações físicas do ouvido.
  - O alcance de freqüências muda com a idade, perdemos a habilidade de ouvir as freqüências mais altas ao envelhecer.

### □ Alcance Dinâmico

- A menor variação de pressão do ar que um humano pode detectar (20 micro Pascal) medido nas freqüências na qual somos mais sensíveis, é usada como referência (0 dB) para medir a intensidade de outros sons.



Relembrando: Uma conversa normal: entre 50-60 dB e som do trânsito de carros: é aproximadamente de 80 dB. O máximo som que o ouvido pode tolerar é 130 dB, o que dá um alcance dinâmico de 0 a 130 dB.

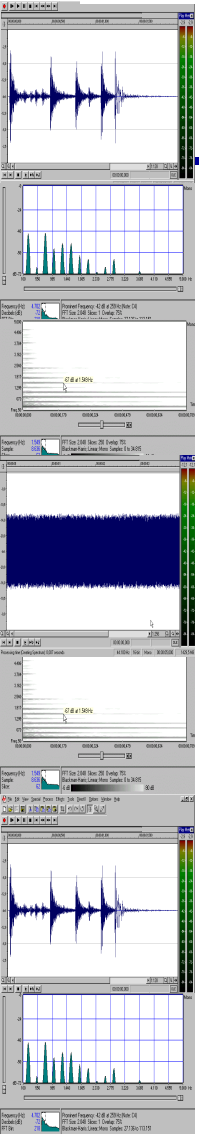


# Compressão de Áudio

---

## Mascaramento Auditivo

- ❑ Definição: corresponde à "audibilidade diminuída de um som devido à presença de outro".
- ❑ O mascaramento auditivo consiste de mascaramento em frequência e mascaramento temporal.





# Compressão de Áudio

---

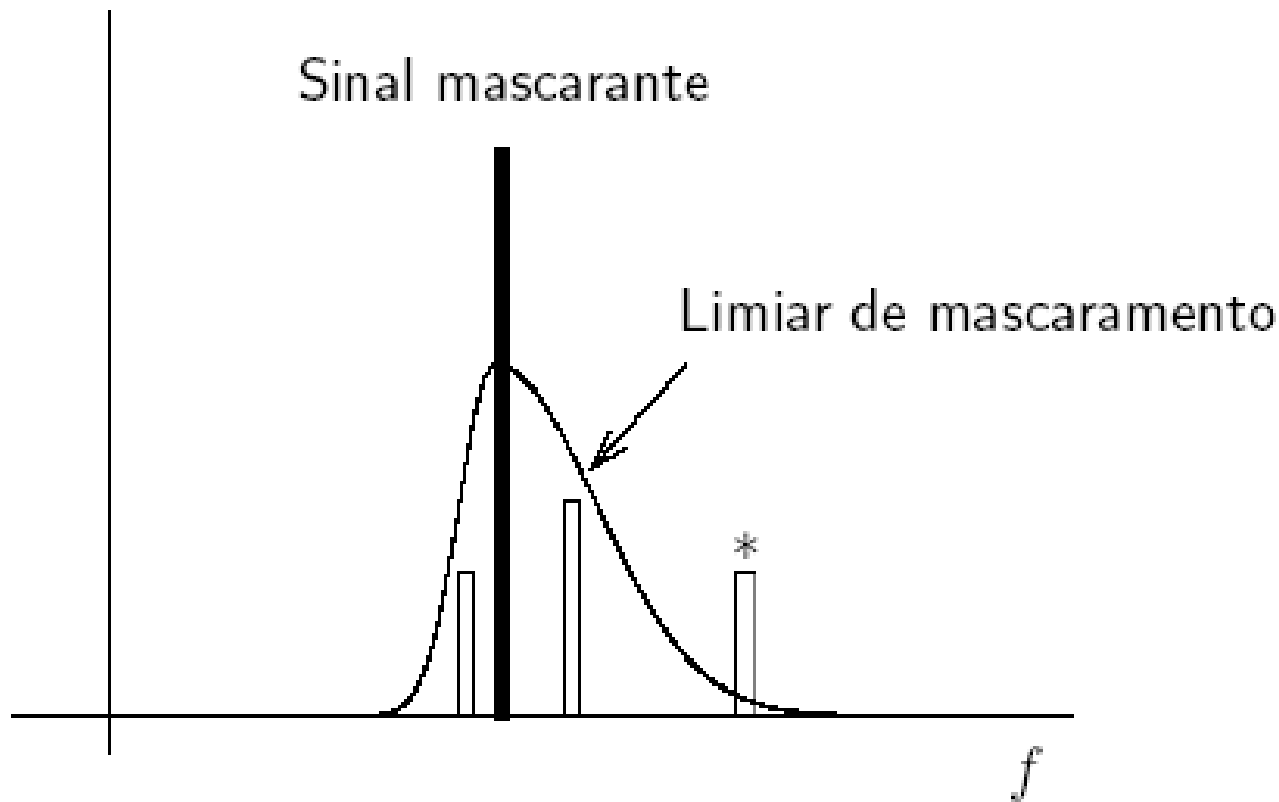
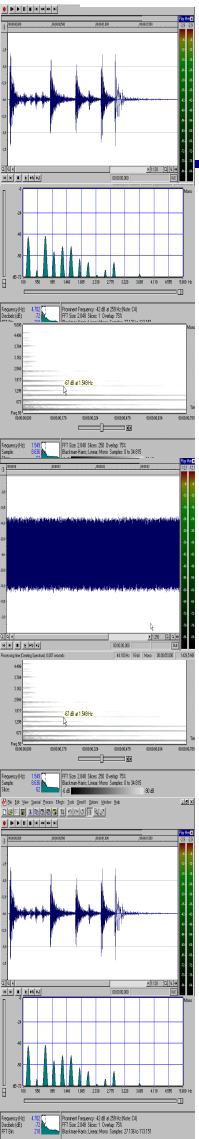
## Mascaramento Auditivo

### □ Mascaramento em Frequência:

- Também chamado mascaramento simultâneo.
- **Exemplo:** Dado um som forte com uma frequência de 1000 Hz, e também um som na frequência de 1100 Hz que está a 18 dB por baixo do anterior, o som de 1100 Hz não pode ser ouvido porque está sendo mascarado pelo som mais forte de 1000 Hz.
- Isto ocorre porque o som de 1000 Hz é mais forte e está perto em frequência. Quanto mais perto em frequência estão, mais fortes são os sons que podem ser mascarados pelo som mais forte.



# Compressão de Áudio



**Mascaramento Psicoacústico:** o sinal marcado com \* não será mascarado.





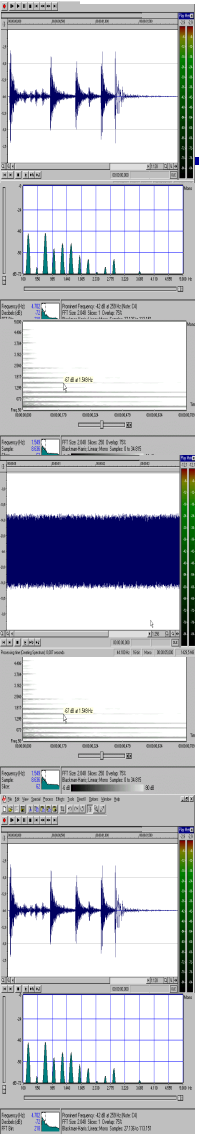
# Compressão de Áudio

---

## Mascaramento Auditivo

### □ Mascaramento Temporal:

- ocorre antes e depois de um som forte.
- Se um som é mascarado depois de um som mais forte é chamado pós-mascaramento, e se é mascarado antes em tempo é chamado pré-mascaramento.
- O pré-mascaramento existe só por um curto momento (20 ms).
- O pós-mascaramento tem efeito até por 200 ms.





# Compressão de Áudio

---

## Mascaramento Auditivo

**Explorando os mascaramentos  
(em frequência e temporal)  
é possível reduzir substancialmente a informação de  
áudio, sem uma mudança audível.**



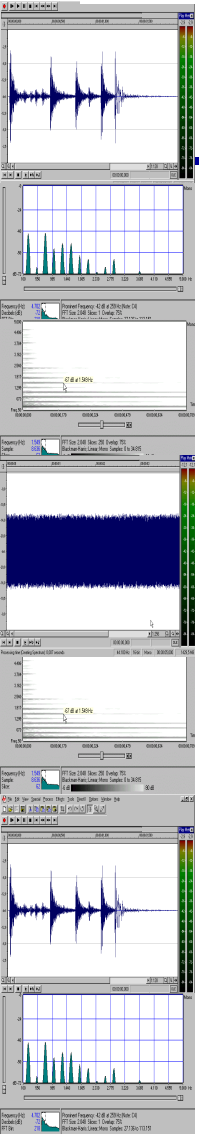


# Compressão de Áudio

---

## Redundância Estereofônica

- O fato do ouvido humano não poder detectar a direção das baixas frequências é chamado Redundância Estereofônica.
- Estas imperfeições, ou limitações na audição, possibilitam "jogar fora" certa informação de áudio, sem afetar o que ouvimos.







# Compressão de Áudio

---

## Codificadores por Forma de Onda e Perceptuais

### Tipos de codificadores de áudio:

#### ❑ Codificadores por forma de onda

- tentam reconstruir o sinal tão exatamente quanto seja possível depois de codificar e decodificar.

#### ❑ Codificadores perceptuais

- não tentam manter o sinal exatamente como era antes da etapa de codificação e decodificação.
- Procuram assegurar que a saída seja percebida como o original pelo ouvido humano.
- Aproveita conhecimentos sobre as propriedades do ouvido e as limitações da audição humana, para eliminar parte do sinal que nós não podemos perceber.



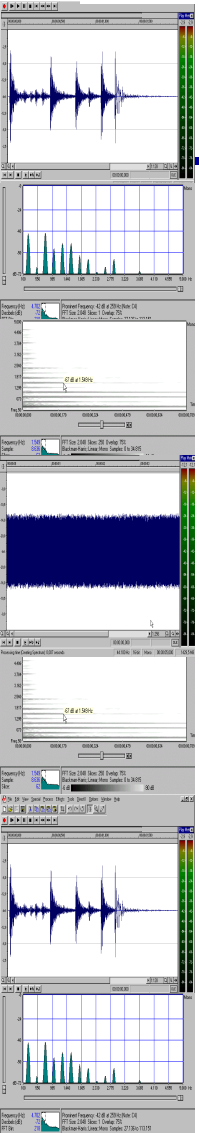


# Compressão de Áudio

---

## Codificadores Perceptuais

- ❑ Quase todos transformam o som do domínio do tempo para o domínio da freqüência, e separaram as diferentes freqüências em sub-bandas.
- ❑ Usam o conhecimento de como o ouvido funciona para eliminar informação não necessária.
- ❑ O efeito de mascaramento é o fenômeno auditivo mais habitualmente explorado.
- ❑ **Exemplos de codificadores perceptuais:**
  - todos os codificadores de áudio MPEG e
  - o codificador ATRAC da Sony, usado nos seus sistemas MiniDisc.





# Compressão de Áudio

---

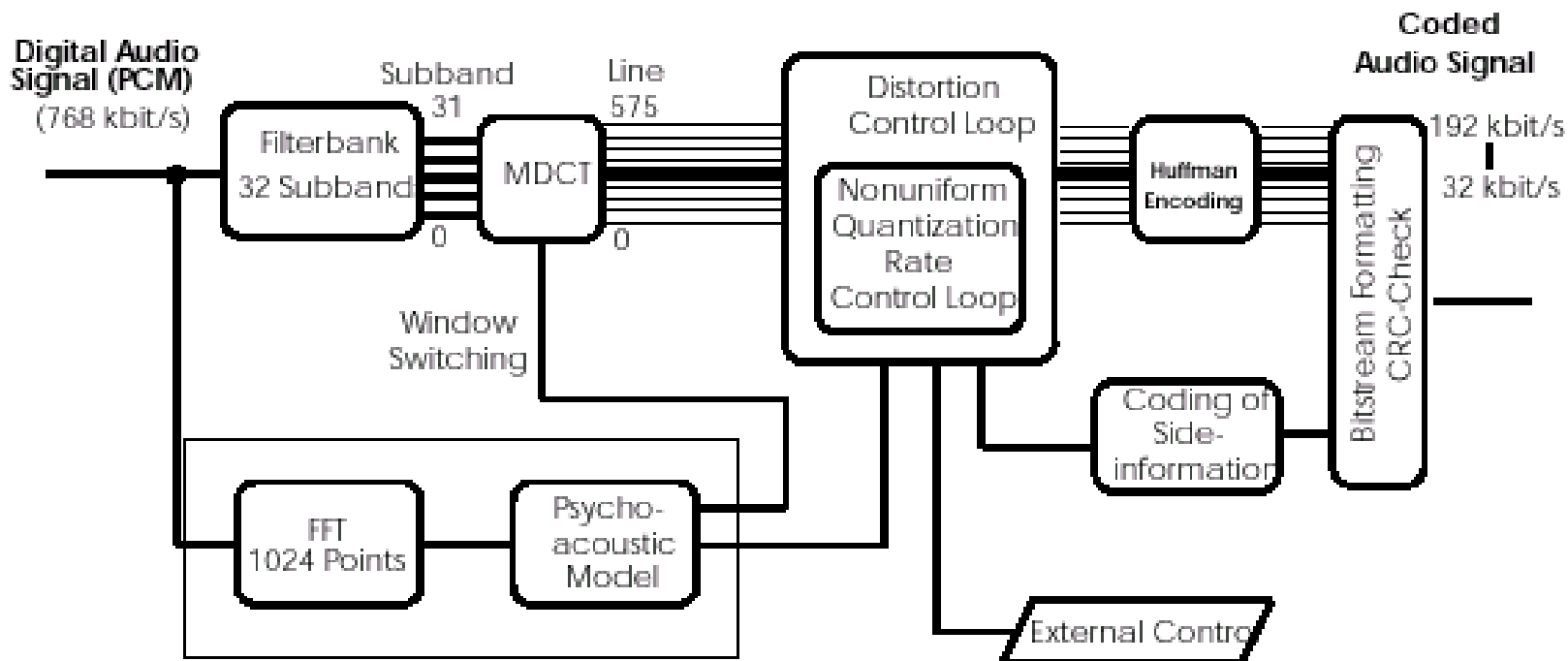
## MPEG-1 Layer III (MP3)

- Algumas das razões pelas quais MP3 é a principal ferramenta para a distribuição de áudio pela Internet:
  - Padrão aberto: Qualquer pessoa interessada pode implementar o *standard* e o formato esta bem definido.
  - Disponibilidade de codificadores e decodificadores.
  - Existência de tecnologia de apoio: amplo uso de placas de som para computadores, computadores suficientemente rápidos para rodar programas de codificação e decodificação de áudio e acesso rápido a Internet .



# Compressão de Áudio

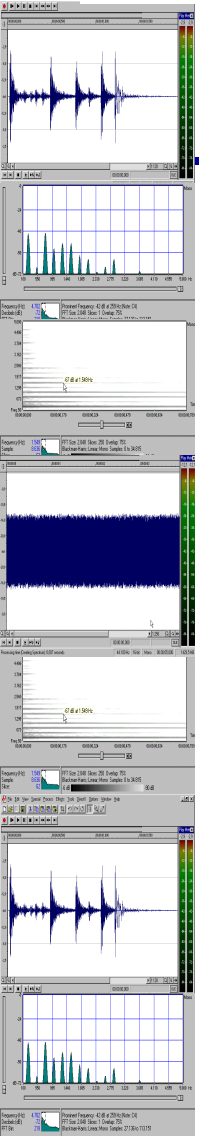
- Visão geral do algoritmo codificador de MP3 (MPEG-1 layer III)



**Objetivo:** produzir um "stream" de bits de acordo com o padrão MPEG.



# Compressão de Áudio



- ❑ Cabeçalho do quadro do MPEG Audio Layer I/II/III (MPEG Layer III)
- ❑ Um arquivo de áudio MPEG é formado de uma sucessão de partes menores chamadas quadros (*frames*).
- ❑ Um quadro é um bloco de dados com o seu próprio cabeçalho e informação de áudio.
- ❑ No caso de Layer III os quadros não são sempre independentes devido ao possível uso do "byte reservoir" (reservatório de bytes), uma espécie de "buffer". No pior caso 9 quadros podem ser necessários antes da decodificação de um quadro.
- ❑ O cabeçalho do quadro tem um comprimento de 32 bits (4 bytes) . Os primeiros onze bits são sempre posicionados em 1 e são chamados "sincronismo de quadro".

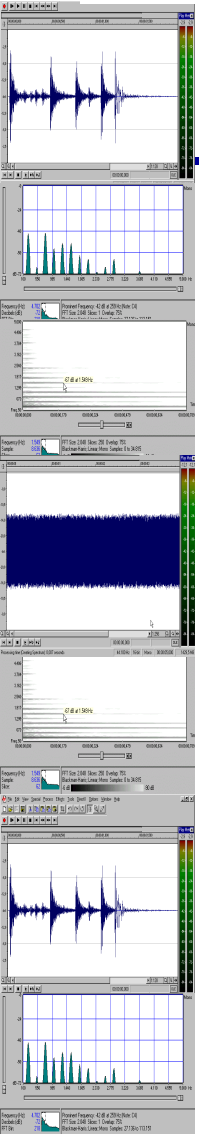




# Compressão de Áudio

- ▣ Cabeçalho do quadro do MPEG Audio Layer I/II/III (MPEG Layer III)
- ▣ Os caracteres A até M do cabeçalho são usados para indicar os diferentes campos.

**Forma: AAAAAAAAA AAABBCCD EEEFFGH IIJJKLMM**





# Compressão de Áudio

---

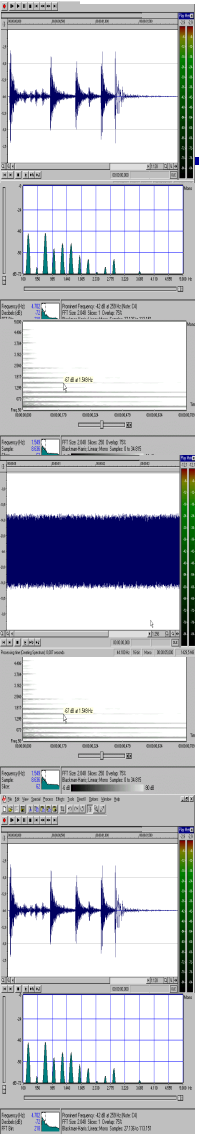
▣ Descrição do cabeçalho do (MPEG Layer III):

Símbolo	Comprimento (bits)	Posição (bits)	Descrição
A	11	(31-21)	Sincronismo de quadro (todos em ?1?)
B	2	(20, 19)	Identificação da versão de MPEG Audio  00 - MPEG Versão 2.5 01 - reservado 10 - MPEG Versão 2 (ISSO/IEC 13818-3) 11 - MPEG Versão 1 (ISSO/IEC 11172-3)
C	2	(18, 17)	Descrição da camada (Layer)  00 - reservado 01 - Layer III 10 - Layer II 11 - Layer I





# Compressão de Áudio



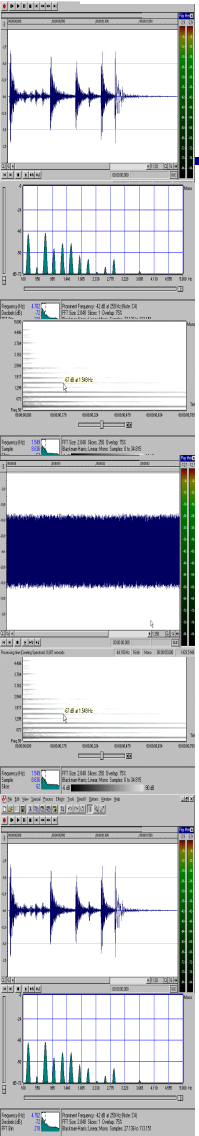
D	1	(16)	Bit de proteção  0 - Protegido por CRC (CRC de 16 bits a continuação do cabeçalho) 1 - Não protegido
E	4	(15, 12)	Indicador de taxa de bits (bitrate) em kbps Se mostra só para MPEG-1 Layer III (mp3)  0000 - formato livre 0001 - 32 0010 - 40 0011 - 48 0100 - 56 0101 - 64 0110 - 80 0111 - 96 1000 - 112 1001 - 128 1010 - 160 1011 - 192 1100 - 224 1101 - 256 1110 - 320 1111 - não permitido







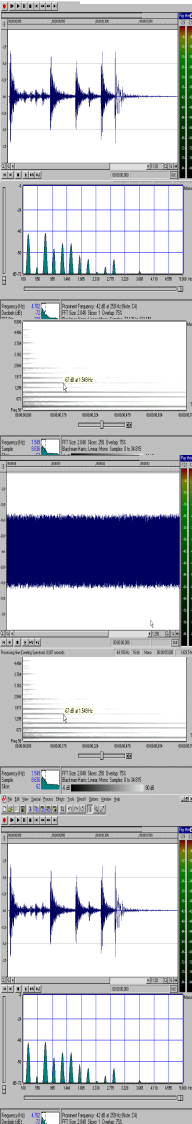
# Compressão de Áudio



F	2	(11, 10)	Taxa de amostragem  Se mostra só para MPEG-1 (MPEG-1 Layer III = mp3)  00 - 44100 Hz 01 - 48000 Hz 10 - 16000 Hz 11 - reservado
G	1	(9)	"Padding bit":  0 - o quadro não está completo ("not padded") 1 - o quadro está completo ("padded")  <input type="checkbox"/> "Padding" é usado para completar exatamente a taxa de bits.
H	1	(8)	Bit privado (informativo)
I	2	(7, 6)	Modo de canal  00 - Estéreo 01 - Estéreo conjunto (Estéreo) 10 - Canal dual (2 canais mono) 11 - Canal único (Mono)



# Compressão de Áudio



J	2	(5, 4)	Modo de extensão (Só usado em estéreo conjunto) (mostra-se só a parte de mp3)		
			valor	Estéreo de intensidade	MS estéreo
			00	Não	não
			01	Sim	não
			10	Não	sim
			11	sim	sim
K	1	(3)	Direitos autorais (Copyright) 0 - O áudio não tem direitos autorais 1 - O áudio tem direitos autorais		
L	1	(2)	Original 0 - Cópia do original 1 - Original		
M	2	(1, 0)	Ênfases: 00 - nenhum 01 - 50/15 ms 10 - reservado 11 - CCIT J.17		





# Compressão de Áudio

---

Mais informações:

- <http://www.mp3dev.org/mp3/>
- <http://www.mp3-tech.org/index.html>
- <http://www.mpeg.org/MPEG/index.html>

