



**UFSC**

**Ciências da  
Computação**

# **CAP 3. COMPRESSÃO DE DADOS MULTIMÍDIA**

**INE5431 Sistemas Multimídia**  
**Prof. Roberto Willrich (INE/UFSC)**  
**roberto.willrich@ufsc.br**

**Aula 6: Padrões de compressão  
multimídia - JPEG**

# Compressão de Dados Multimídia



- **Conteúdo:**

- Necessidade de compressão
- Princípios da compressão
- Classificação das técnicas de compressão
- Medição do desempenho de compressão
- Técnicas de compressão sem perdas
  - RLE, Huffman, LZW (GIF), Codificação Preditiva
- Técnicas de compressão de áudio, vídeo e imagens
- **Padrões de compressão multimídia**
  - **JPEG**, MPEG, MPEG-4, H.261, H.263

# Padrões de Compressão Multimídia



- **Várias técnicas e produtos para compressão são disponíveis**
  - Utilização de padrões promove a compatibilidade entre diferentes equipamentos/aplicações (interoperabilidade)
- **Exemplos de padrões**
  - TIFF: padrão independente de fabricante para imagens
  - PNG: padrão de imagens alternativo ao GIF
  - ISO JPEG para compressão de imagens;
  - ISO JBIG para compressão sem perda de imagens bi-níveis (1 bit/píxel) para transmissão fac-símile
  - ITU-TS H.261 para videofonia e aplicações de teleconferências na taxa de bits múltiplos de 64 Kbps;
  - ITU-TS H.263 para aplicações de videofonia na taxa abaixo de 64 Kbps;
  - ISO MPEG para compressão de vídeo e áudio associado;

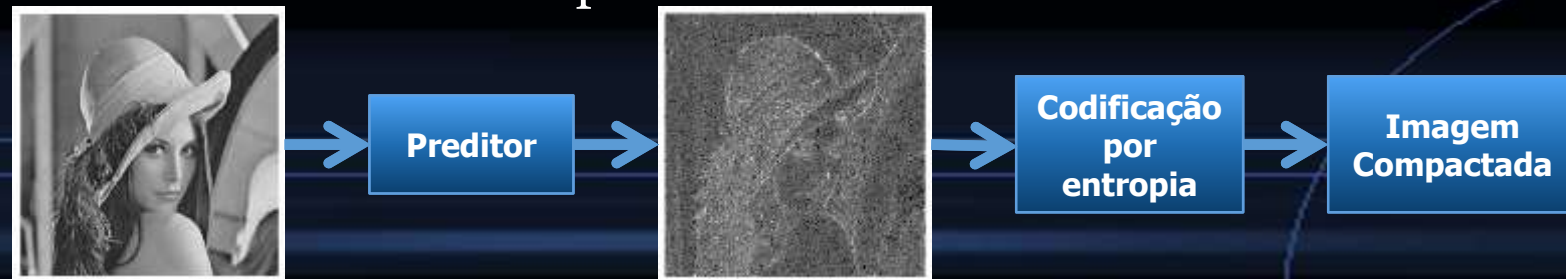
# Padrão de Compressão de Imagens JPEG



- **JPEG colaboração entre a ISO/IEC e a ITU-TS (1992)**
  - Uma das melhores tecnologia de compressão de imagem
  - Implementado em software e hardware
  - Codificação/decodificação JPEG tempo-real tem sido implementada para vídeo (Motion JPEG - MJPEG)
- **Possui versões diferentes:**
  - Versões para compressão sem perdas
  - Versões para compressão com perdas

# Padrão de Compressão de Imagens JPEG

- **Codificação JPEG sem perda**
  - Reprodução é exata
  - Necessária em aplicações que não toleram perdas (médicas e legais)
  - Existem variações
    - JPEG sem perdas original, que se baseia no DPCM e o uso de codificação por entropia (de Huffman ou aritmética)
    - JPEG-LS utiliza a técnica de codificação de Golomb-Rice e RLE
    - JPEG 2000 utiliza técnica de compressão wavelets





# Padrão de Compressão de Imagens JPEG



- **JPEG com perdas**
  - Se baseia nas limitações da percepção humana e na codificação por entropia
  - Compressão parametrizável
    - JPEG cobre grande faixa de qualidades de imagens e permite especificar o comportamento do codificador a partir de parâmetros
    - Quatro modos de operação:
      - Codificação sequencial (baseline)
      - Codificação progressiva
      - Codificação sem perda
      - Codificação hierárquica

# Padrão de Compressão de Imagens JPEG



- **Codificação Sequencial (baseline)**

- Suportado por toda implementação JPEG
- Modo com perdas baseada em DCT
- Componentes de imagem são codificados em uma única varredura da esquerda para direita e de cima para baixo

- **Codificação progressiva**

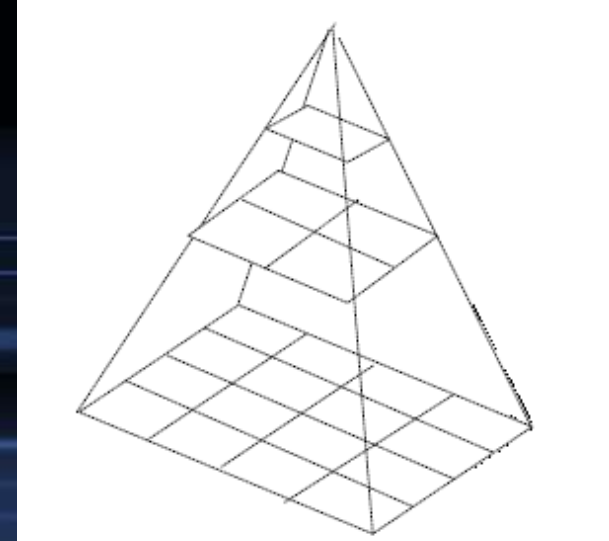
- Com perdas baseada em DCT expandido
  - Fornece avanços ao modo baseline
- Varreduras sucessivas
  - imagem é compactada em um processo de múltiplas linhas de varredura
- Geralmente utilizada em arquivos que são transmitidos pela Internet
  - pois possibilita a visualização da imagem inteira, em menor resolução, enquanto o restante da imagem esta sendo enviada

# Padrão de Compressão de Imagens JPEG



## ■ Codificação hierárquica

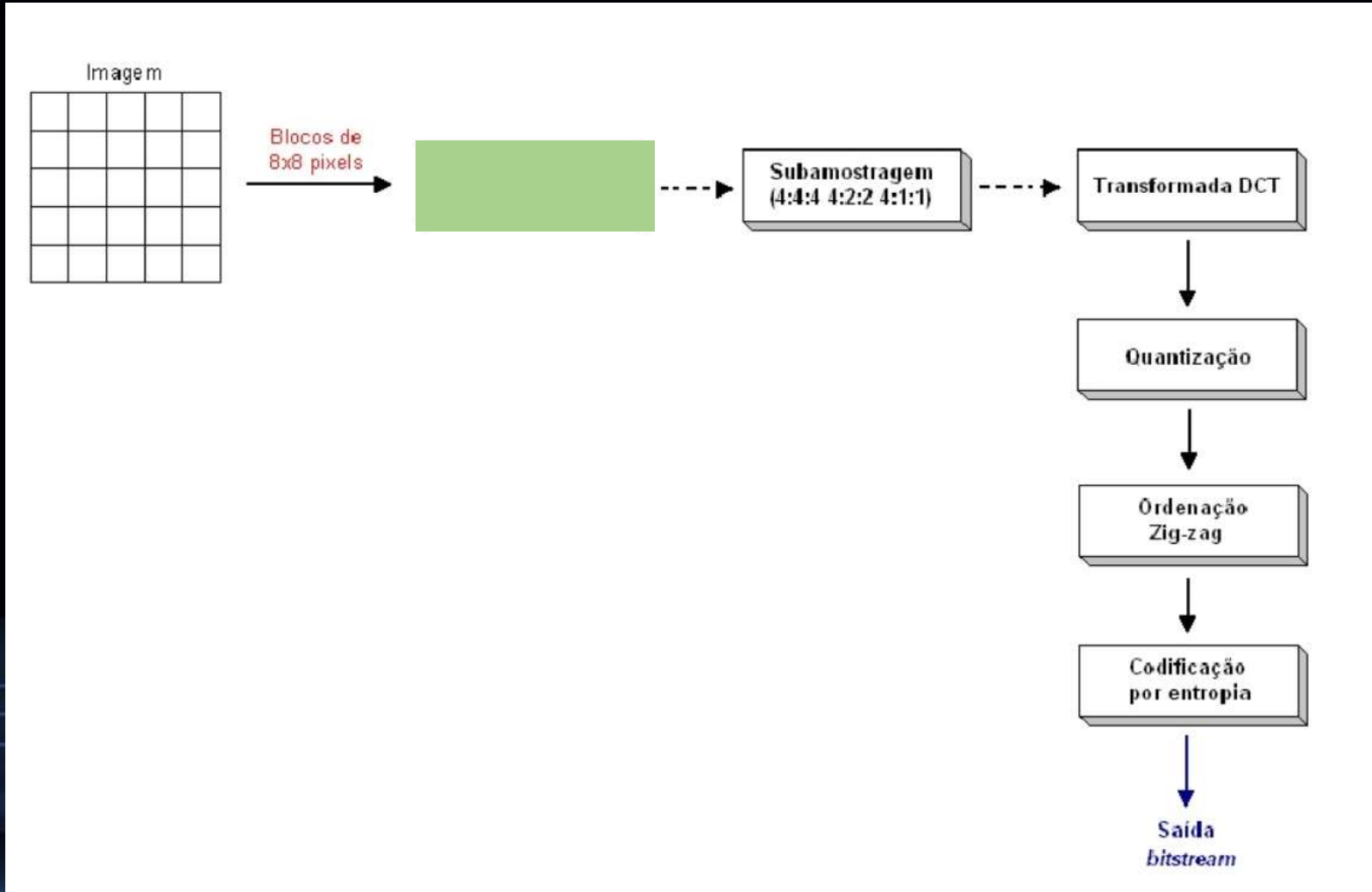
- Oferece uma codificação progressiva que aumenta de resolução espacial entre estágios progressivos
- Versões podem ser acessadas sem a necessidade de primeiro descompactar a imagem na resolução completa
- Os elementos de imagem das resoluções já recebidas são utilizados na próxima resolução, diminuindo desta forma o tamanho do arquivo
- Taxa de compressão é mais baixa que ter uma resolução única





# Algoritmo de compressão JPEG

- Operações a compressão JPEG (Sequencial)



# Algoritmo de compressão JPEG



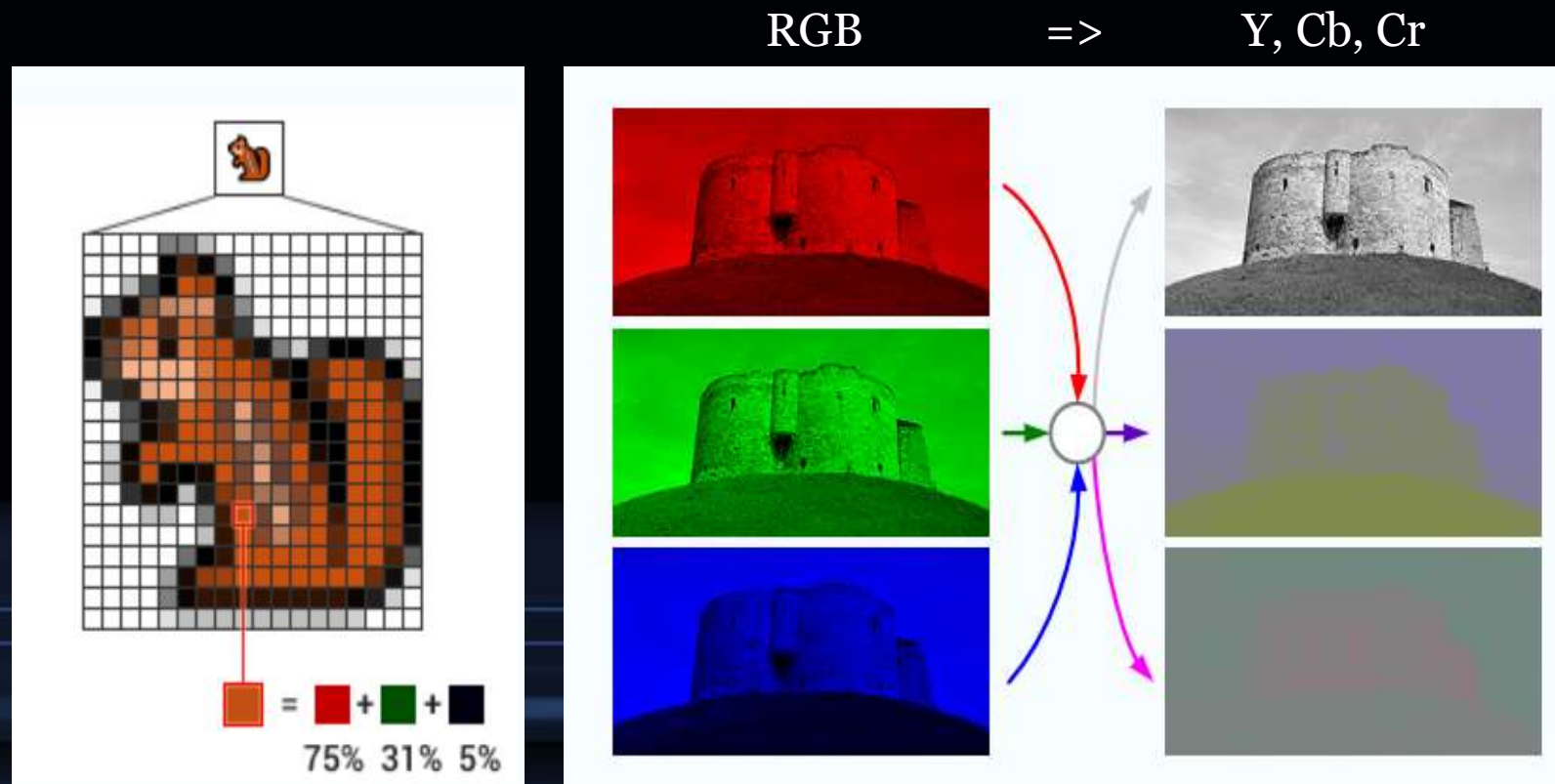
- **Transformação do espaço de cores para YCrCb**
  - Componentes “RGB” da imagem são convertidos para componentes de luminância (“Y”) e cromaticidade (“Cr” e “Cb”)
    - Y: Luminância é uma escala de representação numérica do cinza,
    - CrCb: Cromaticidade são duas escalas numéricas, que juntas representam as cores.

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ -0.1687 & -0.3313 & 0.5000 \\ 0.5000 & -0.4187 & -0.0813 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- YCbCr permite uma maior compressão sem um efeito significativo na qualidade da imagem percebida.

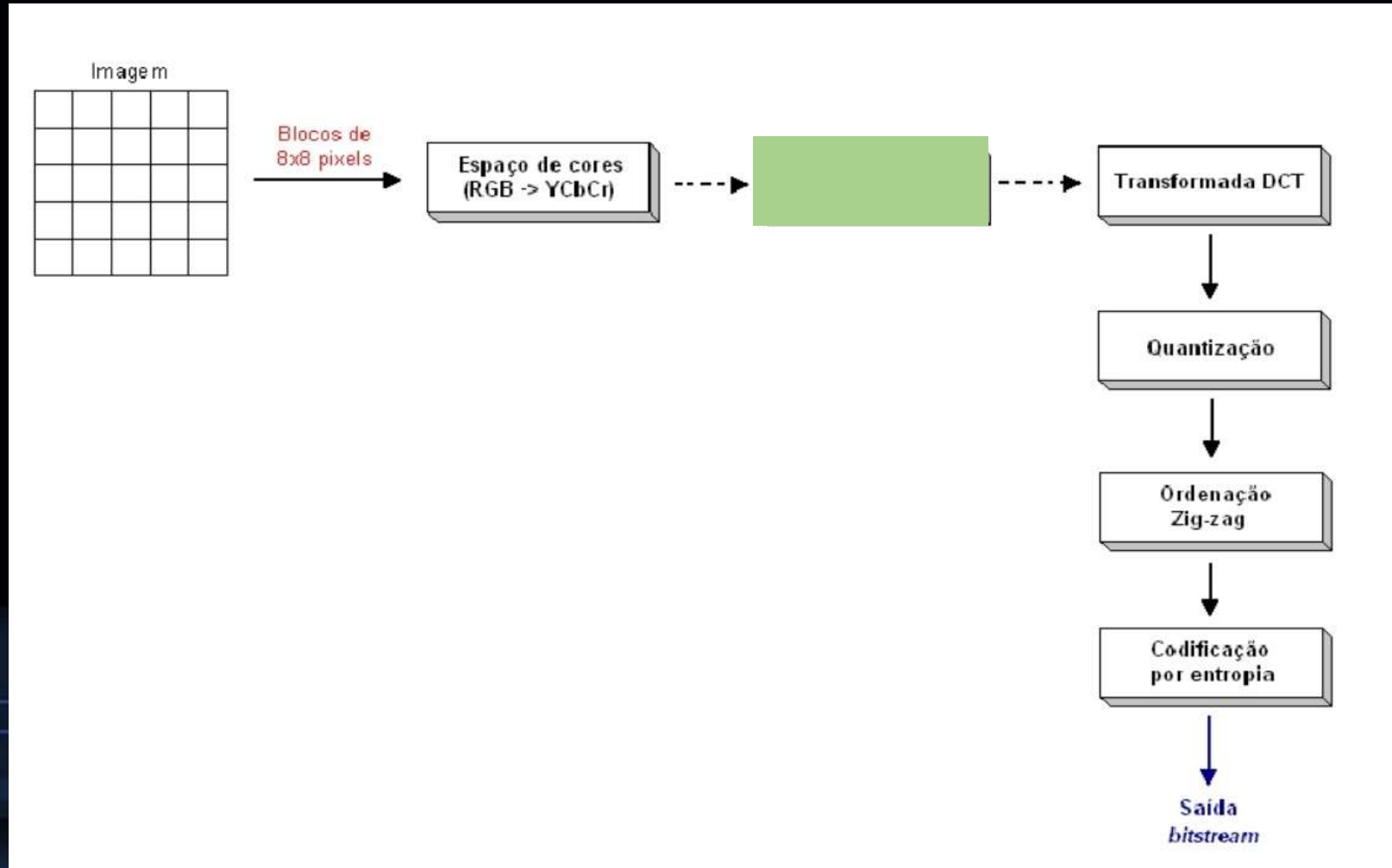
# Algoritmo de compressão JPEG

- **Transformação do espaço de cores para YCrCb**
  - Veja o tutorial em <https://cgjennings.ca/articles/jpeg-compression/>



# Algoritmo de compressão JPEG

## ■ Operações a compressão JPEG (Sequencial)



# Algoritmo de compressão JPEG



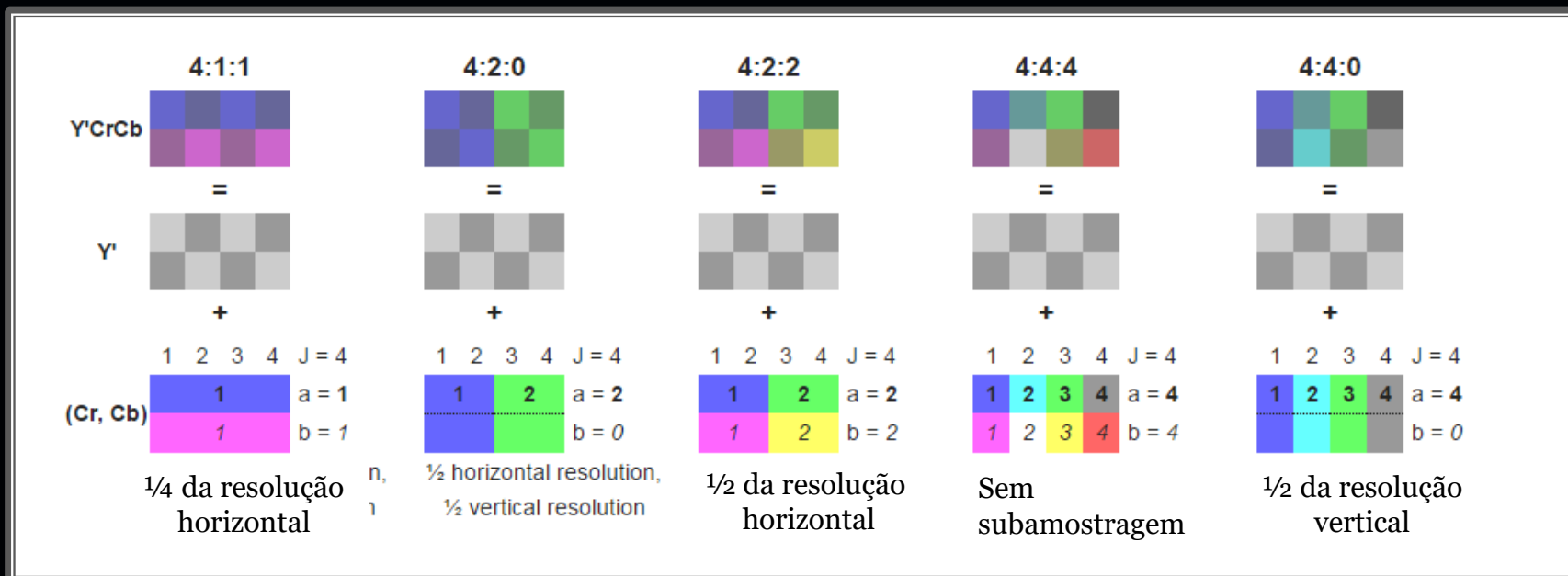
## ■ Subamostragem

- Onde é feita uma redução da resolução das matrizes YCbCr.
- Taxas de subamostragem que são normalmente aplicados no JPEG
  - 4:4:4 (sem subamostragem)
  - 4:2:2 onde as matrizes de crominância são reduzidas na taxa de 2:1 horizontalmente (cada duas linhas é convertida em uma),
  - 4:2:0 mais comumente adotada, onde a uma redução do fator 2 nas direções horizontais e verticais.
- A matriz de luminância geralmente não é reduzida
  - pois o olho humano é mais sensível à luminância (tonalidade de cinza) do que à crominância (tonalidades das cores), o que permite maior taxa de perda de crominância sem que esta perda seja percebida
- No resto do processo de compressão, Y, Cb e Cr são processadas separadamente de maneira muito similar.



# Padrão ISO/IEC MPEG-1 Vídeo

## ■ Subamostragens YCbCr



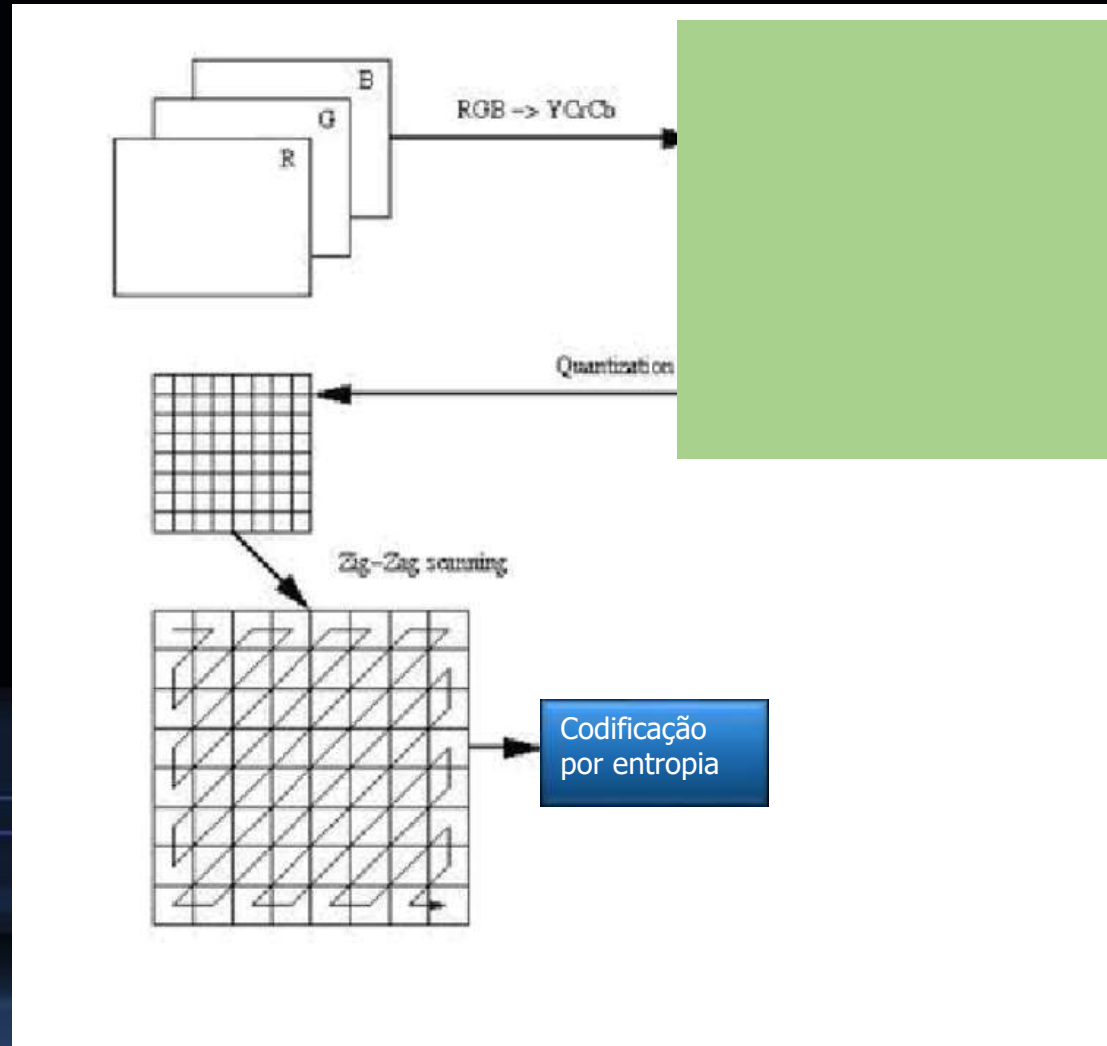
Original sem sub-amostragem de crominância, 200% de zoom



Original após sub-amostragem de crominância, 200% de zoom

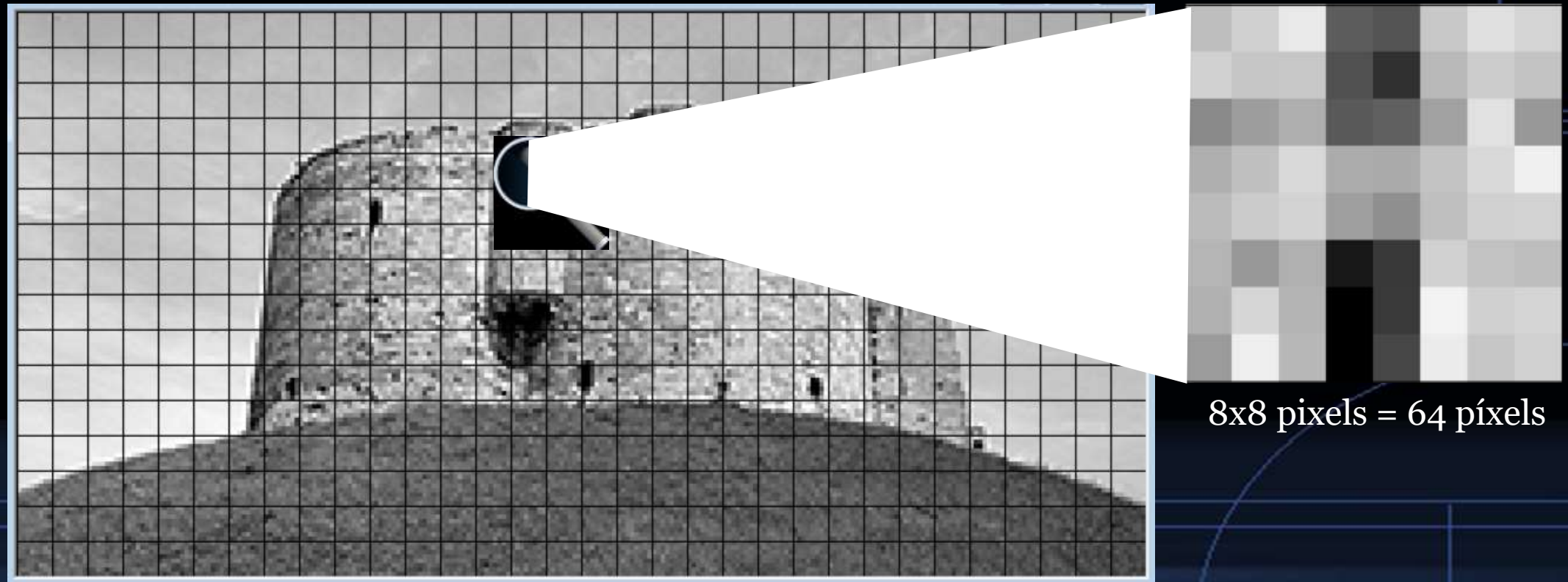
# Algoritmo de compressão JPEG

- **Operações a compressão JPEG**



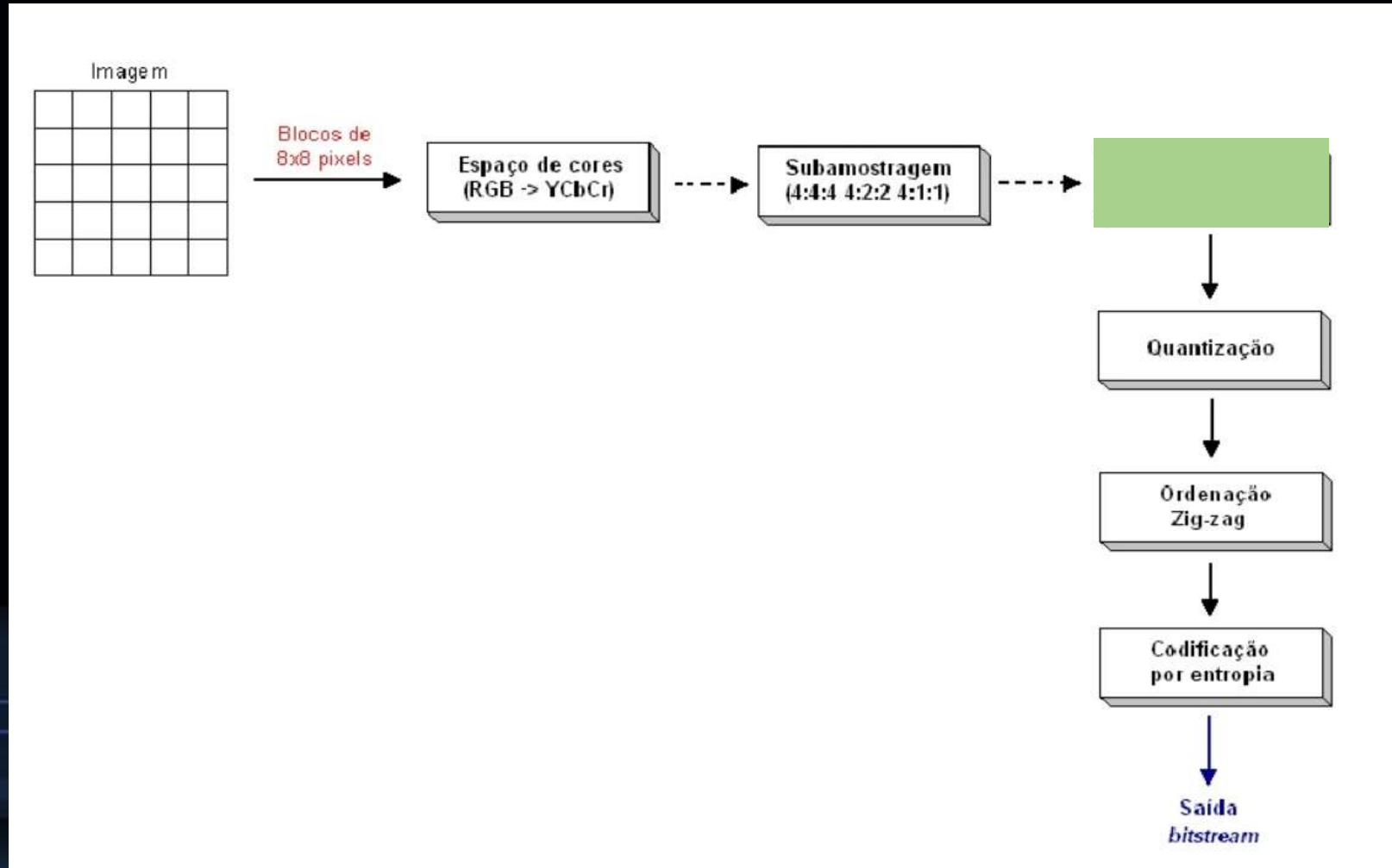
# Algoritmo de compressão JPEG

- **Decomposição das matrizes Y,Cb,Cr**
  - Decompostas em blocos de 8x8 píxels



# Algoritmo de compressão JPEG

## ■ Operações a compressão JPEG (Sequencial)



# Algoritmo de compressão JPEG



- **Transformação discreta de co-seno (DCT) dos blocos**
  - Blocos 8x8 são transformado para o domínio da frequência espacial usando a transformada DCT
  - Similar a transformada de Fourier (representação de sinal com somatório de senos de diferentes frequências, fases e amplitudes)

No espaço bidimensional de uma imagem de 8x8 pixels, a transformada discreta de co-senos (*FDCT: Forward Discrete Cosine Transform*) é dada por:

$$F(u,v) = \frac{1}{4} C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right]$$

$$C(w) = \frac{1}{\sqrt{2}} \text{ para } w = 0$$

$$C(w) = 1 \text{ para } w = 1, 2, \dots, 7$$

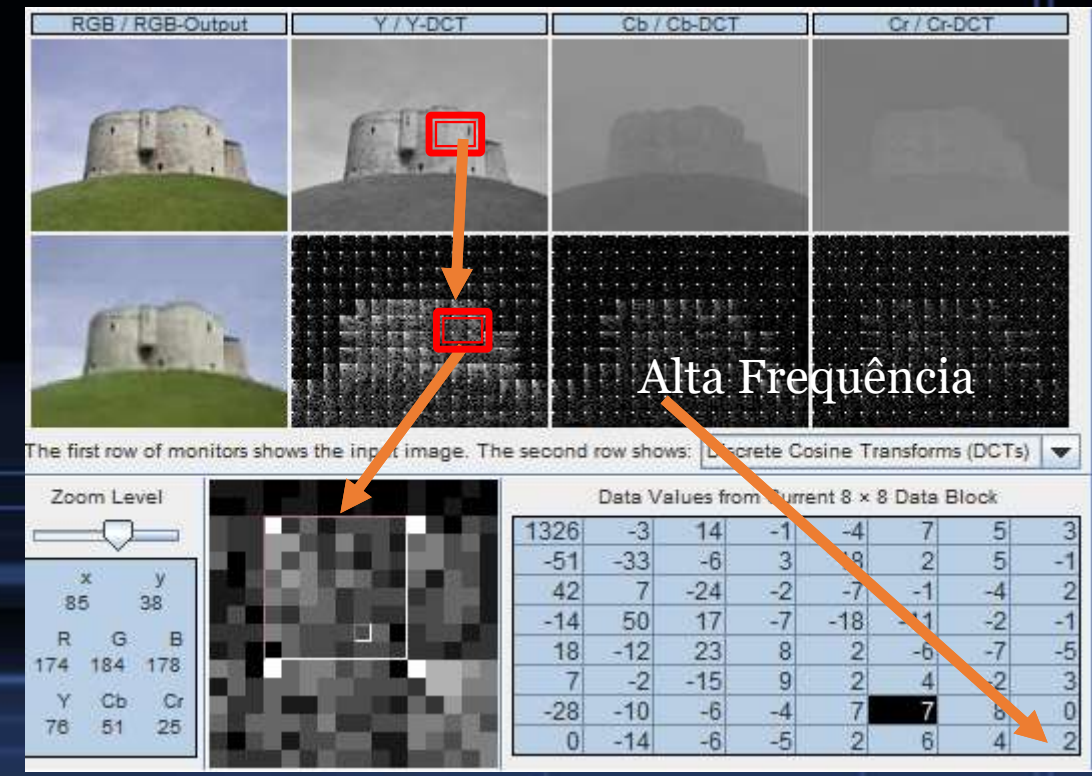
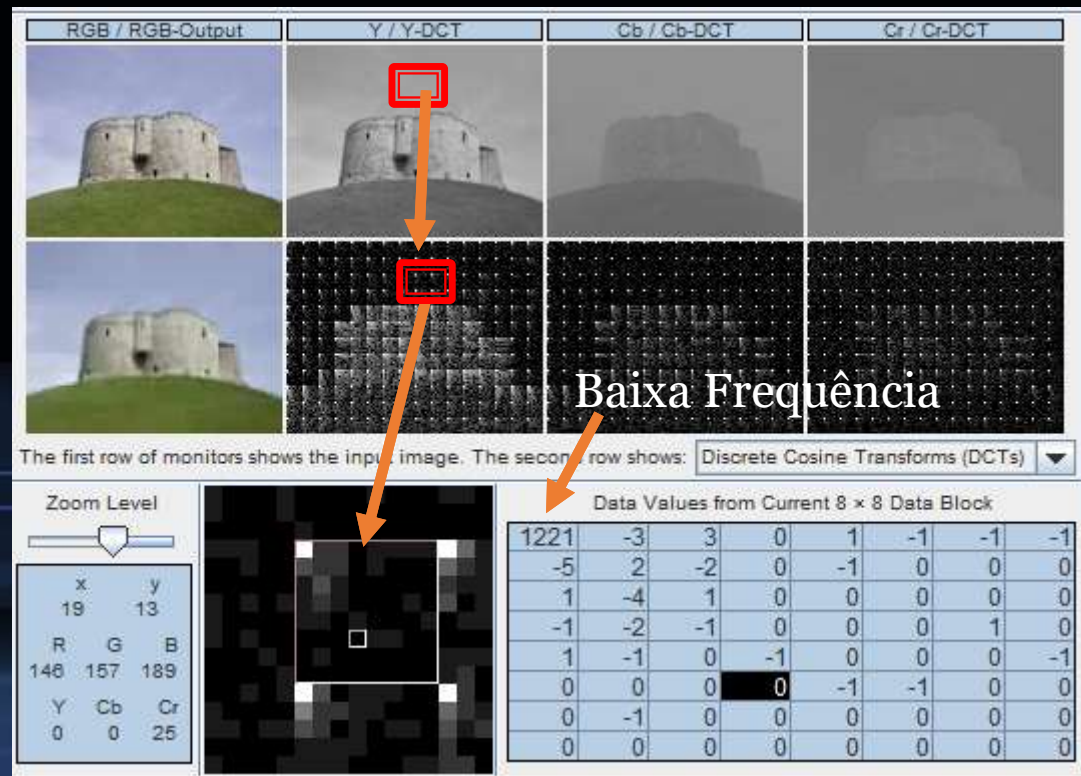
E a transformada inversa (*IDCT: Inverse Discrete Cosine Transform*) por:

$$f(x,y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v) F(u,v) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right]$$



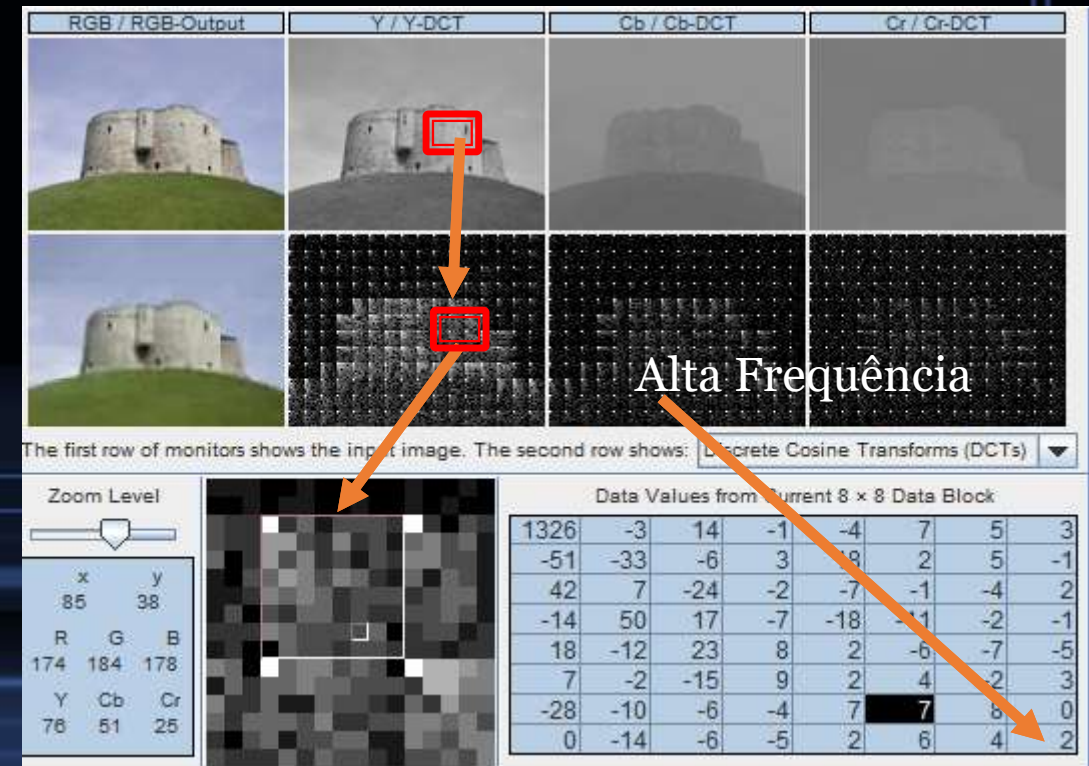
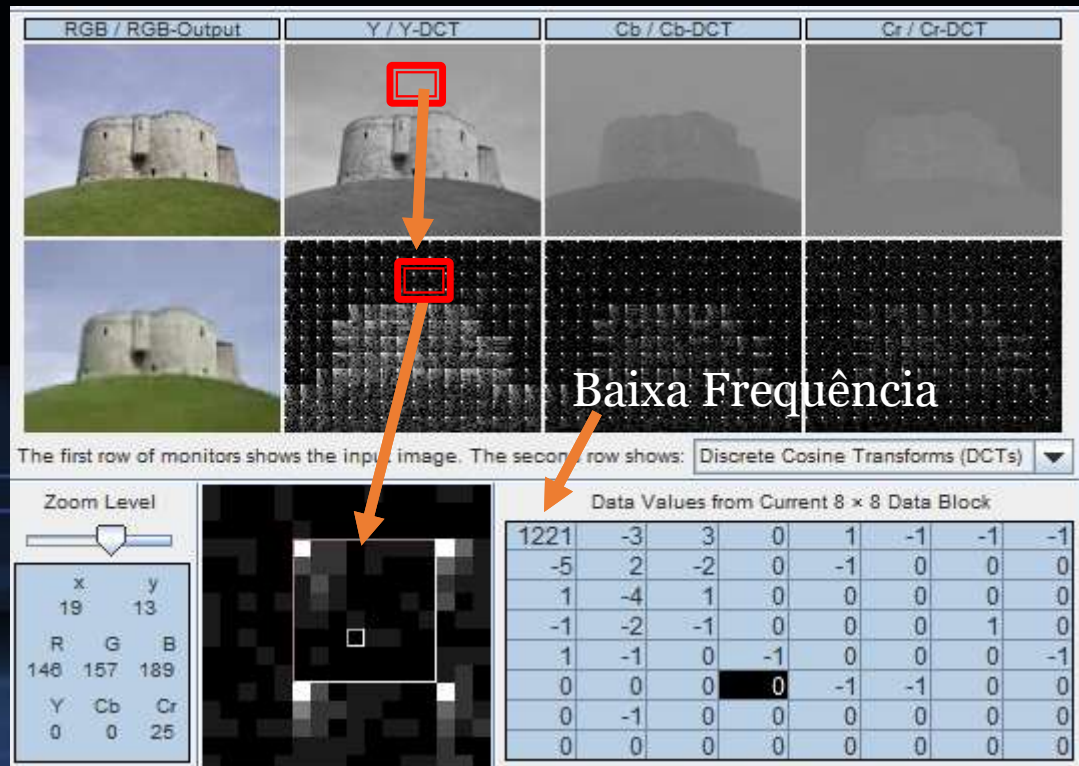
# Algoritmo de compressão JPEG

- **Transformação discreta de co-seno (DCT) dos blocos**
  - Sinal discreto de 64 pontos (um para cada bloco) transformado é uma função de duas dimensões espaciais, x e y
    - estas componentes são chamadas de frequências espaciais ou coeficientes DCT



# Algoritmo de compressão JPEG

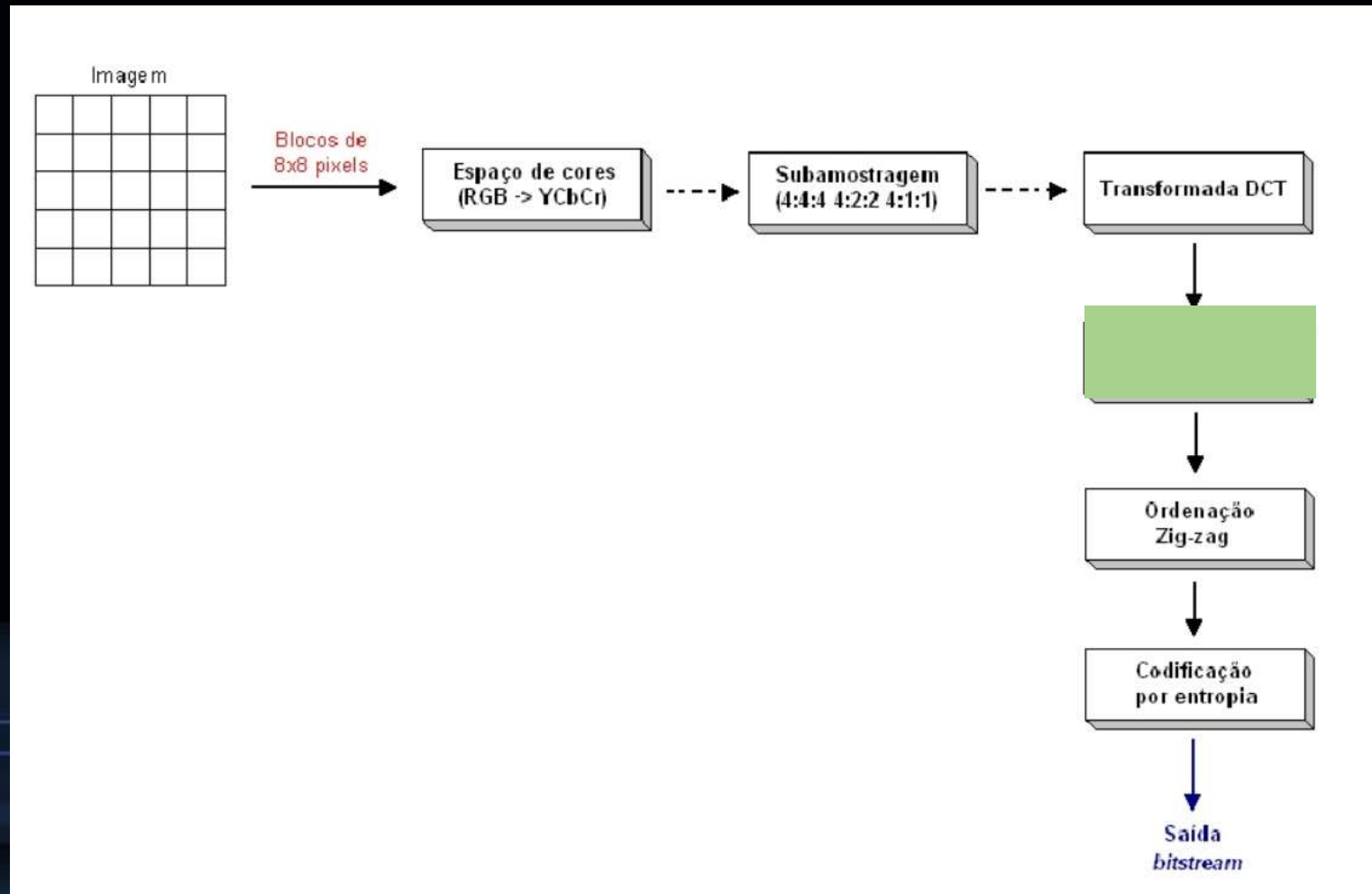
- **Transformação discreta de co-seno (DCT) dos blocos**
  - Mudanças abruptas que acontecem nos contornos de uma figura estão concentradas nas frequências mais altas.
    - uma imagem com poucos contornos deve concentrar seus coeficientes nas frequências baixas.
  - Coeficientes das frequências altas são menos importantes e perdas nesses coeficientes podem diminuir um pouco a nitidez da imagem, mas para muitas aplicações isto pode ser aceitável





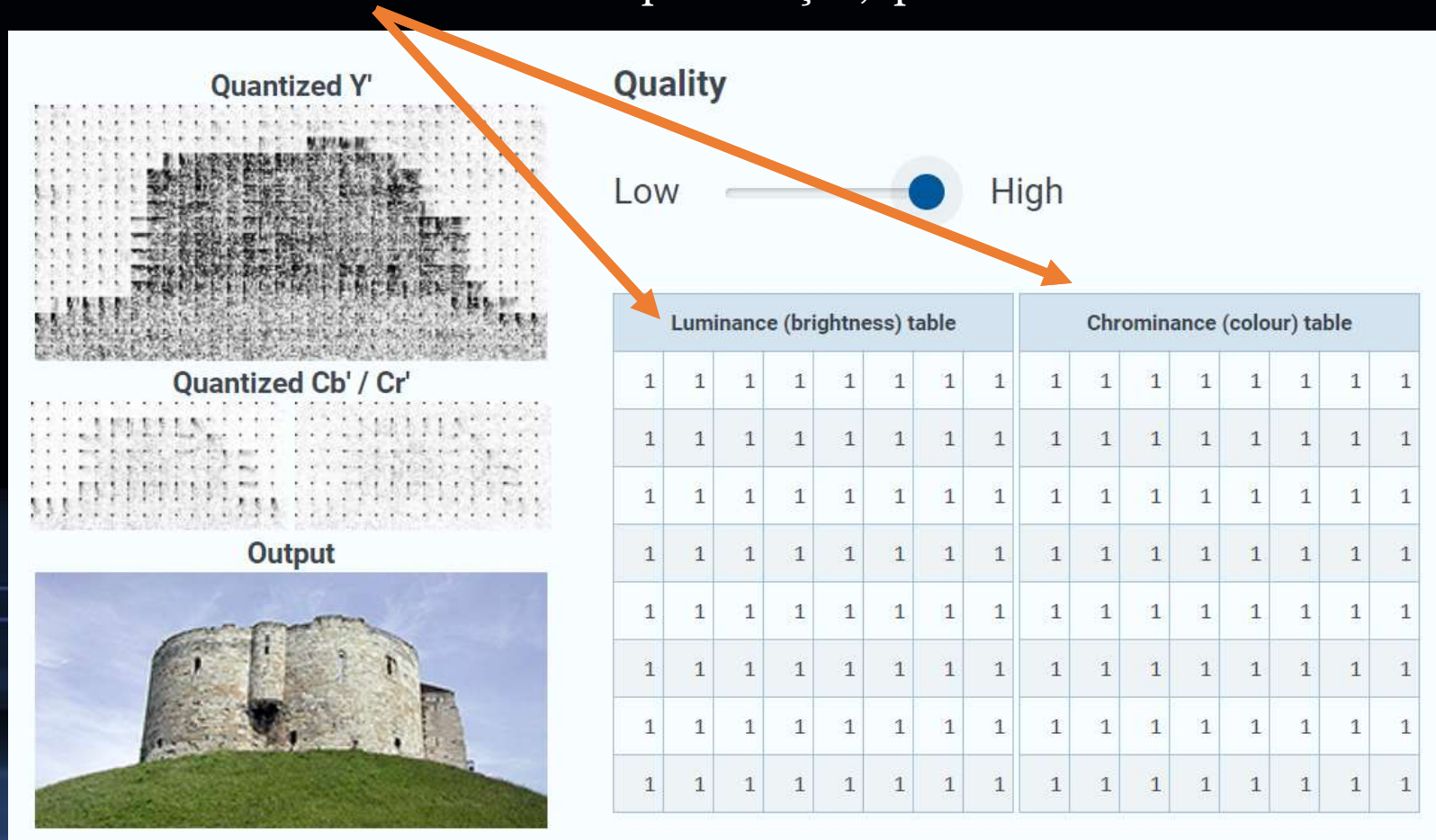
# Algoritmo de compressão JPEG

## ■ Operações a compressão JPEG



# Algoritmo de compressão JPEG

- **Quantificação (escolha da qualidade)**
  - Alta frequência é desprezada de acordo com a taxa de compressão desejada:
    - Utilizando duas tabelas de quantização, que variam de acordo com a qualidade desejada



# Algoritmo de compressão JPEG



## ■ Quantificação (escolha da qualidade)

- Tabelas de quantização são divisores: Em cada bloco 8x8 terá seus coeficientes DCT divididos pelo número correspondente em sua tabela de quantização.
- O resultado de cada divisão é arredondado para o número inteiro mais próximo e as partes fracionárias são jogadas fora

Coeficientes DCT

55	50	45	20	12	2	2	1
55	48	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

Tabela de quantização

1	2	3	3	3	6	6	12
1	3	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

Coeficientes DCT quantizado

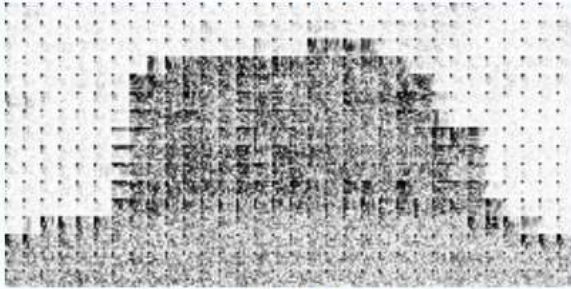
55	25	15	20	6	0	0	0
55	16	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...




# Algoritmo de compressão JPEG

- **Quantificação (escolha da qualidade)**
  - Alta frequência é desprezada de acordo com a taxa de compressão desejada:
    - maior é taxa de compressão, maior é o número de componentes de alta frequência desprezados


Quantized Y'



Quantized Cb' / Cr'



Output



Quality

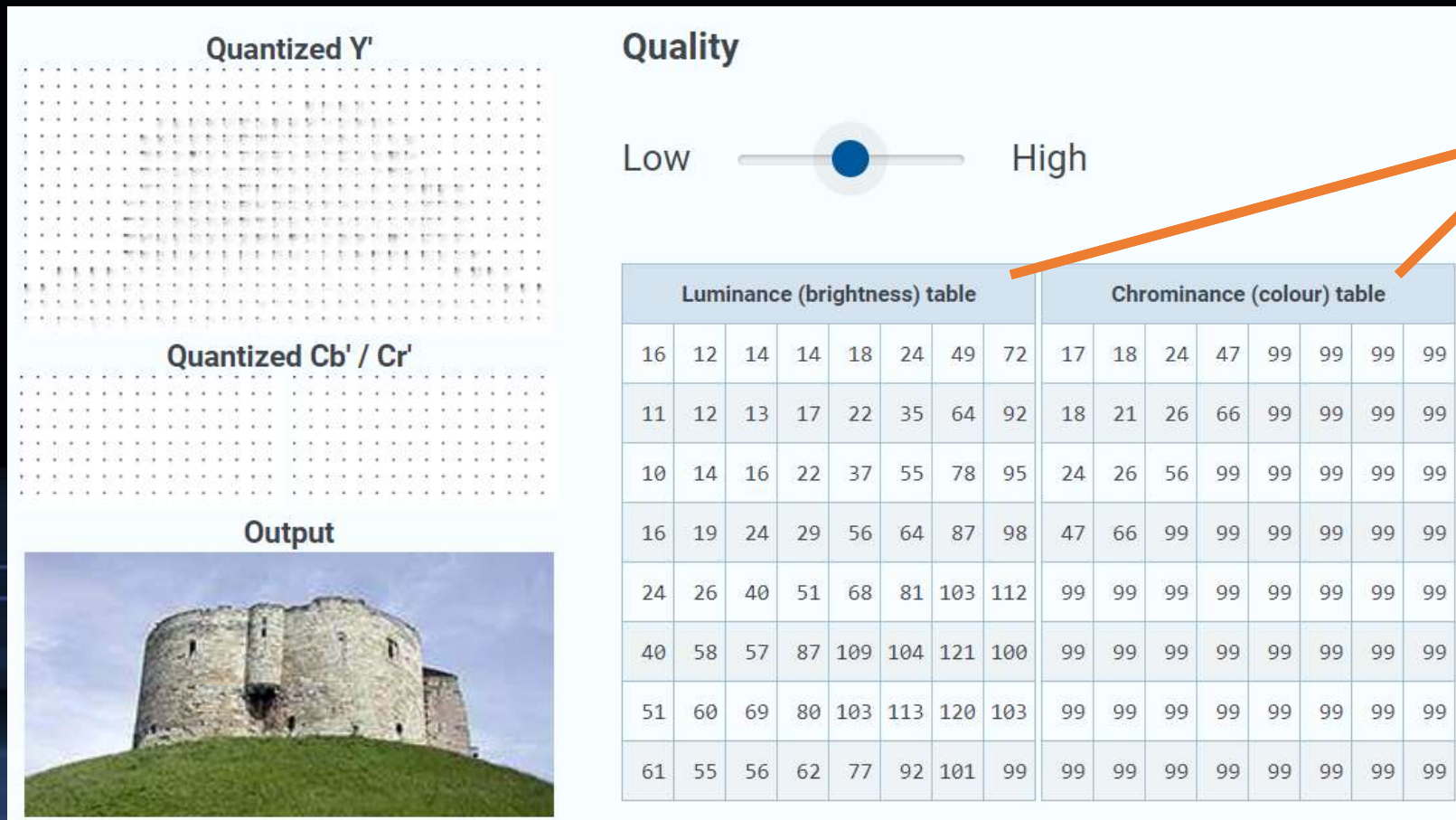
Low ☐ High

Luminance (brightness) table								Chrominance (colour) table							
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	

Alta qualidade preserva os coeficientes de alta e baixa frequência

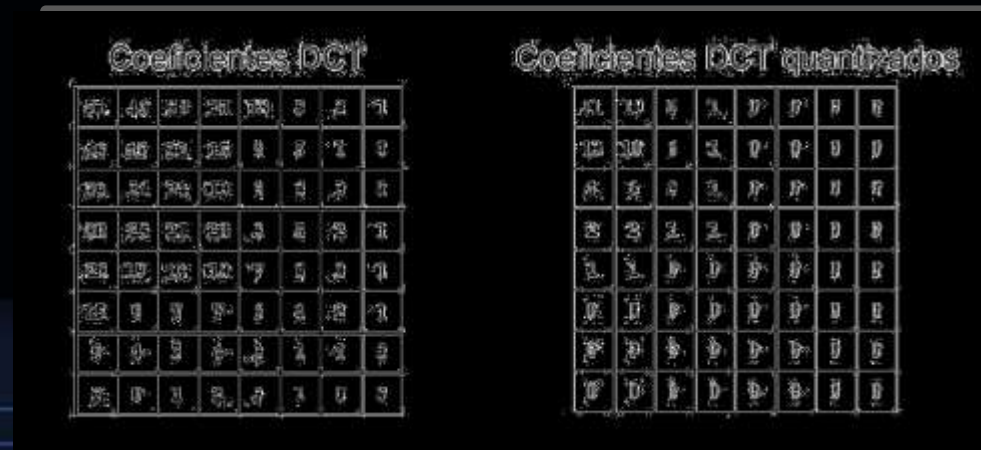
# Algoritmo de compressão JPEG

- **Quantificação (escolha da qualidade)**
  - Alta frequência é desprezada de acordo com a taxa de compressão desejada:
    - maior é taxa de compressão, maior é o número de componentes de alta frequência desprezados



# Algoritmo de compressão JPEG

- **Quantificação (escolha da qualidade)**
  - Quantificação prioriza a baixa frequência
    - os coeficientes gerados são quantizados de forma diferenciada, usando uma maior precisão para as frequências mais baixas.





# Algoritmo de compressão JPEG

- **Quantificação (escolha da qualidade)**
  - Alta frequência é desprezada de acordo com a taxa de compressão desejada:
    - maior é taxa de compressão, maior é o número de componentes de alta frequência desprezados

Quantized Y'

Quantized Cb' / Cr'

Output

Quality

Low  High

Luminance (brightness) table								Chrominance (colour) table							
112	84	98	98	126	168	343	504	119	126	168	329	693	693	693	693
77	84	91	119	154	245	448	644	126	147	182	462	693	693	693	693
70	98	112	154	259	385	546	665	168	182	392	693	693	693	693	693
112	133	168	203	392	448	609	686	329	462	693	693	693	693	693	693
168	182	280	357	476	567	721	784	693	693	693	693	693	693	693	693
280	406	399	609	763	728	847	700	693	693	693	693	693	693	693	693
357	420	483	560	721	791	840	721	693	693	693	693	693	693	693	693
427	385	392	434	539	644	707	693	693	693	693	693	693	693	693	693

Baixa qualidade descarta muitos coeficientes DCTs, gerando o efeito bloco

# Algoritmo de compressão JPEG

- **Quantificação (escolha da qualidade)**
  - Alta frequência é desprezada de acordo com a taxa de compressão desejada:
    - maior é taxa de compressão, maior é o número de componentes de alta frequência desprezados

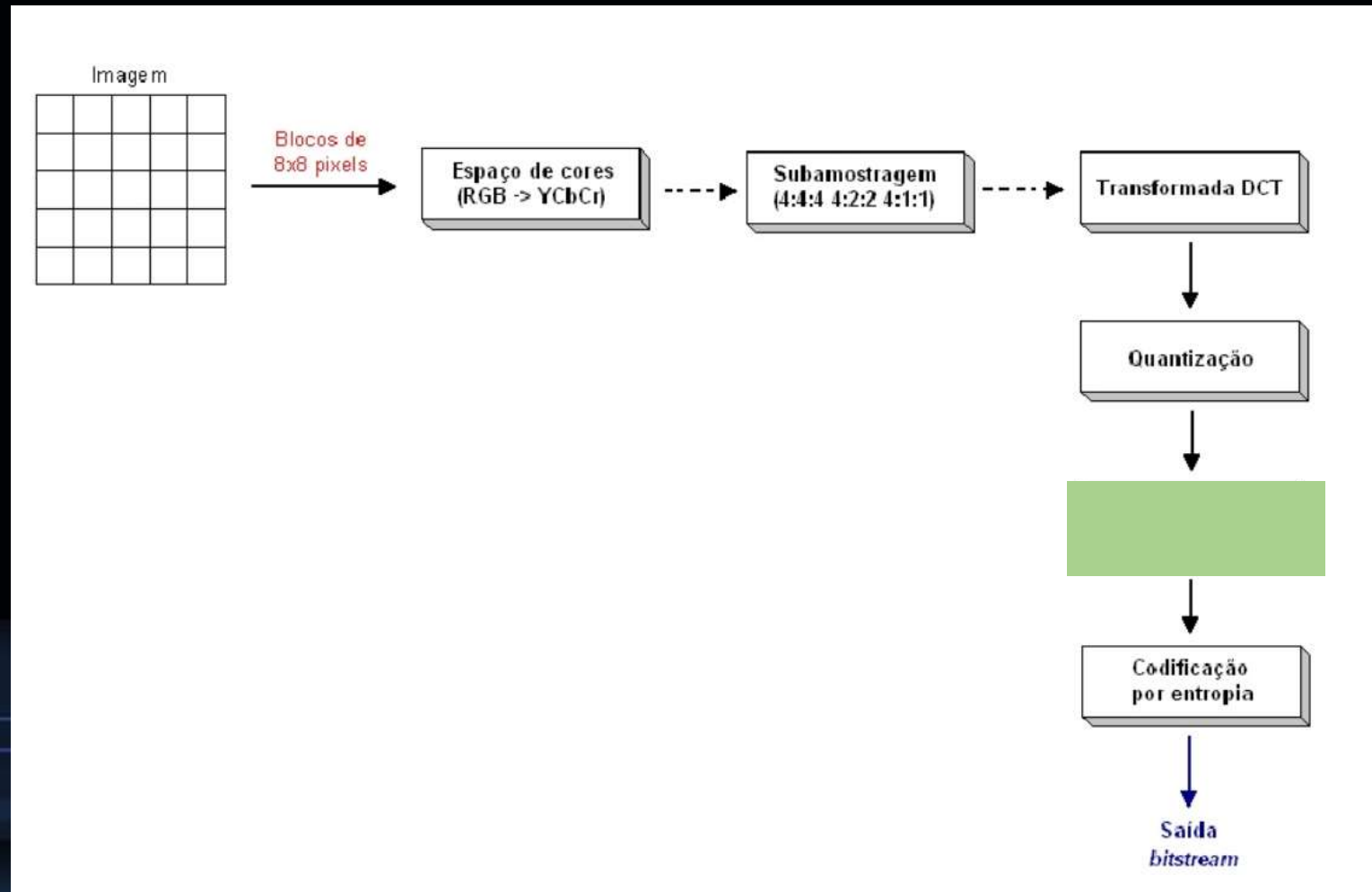


Baixa qualidade  
descarta muitos  
coeficientes DCTs,  
gerando o efeito bloco



# Algoritmo de compressão JPEG

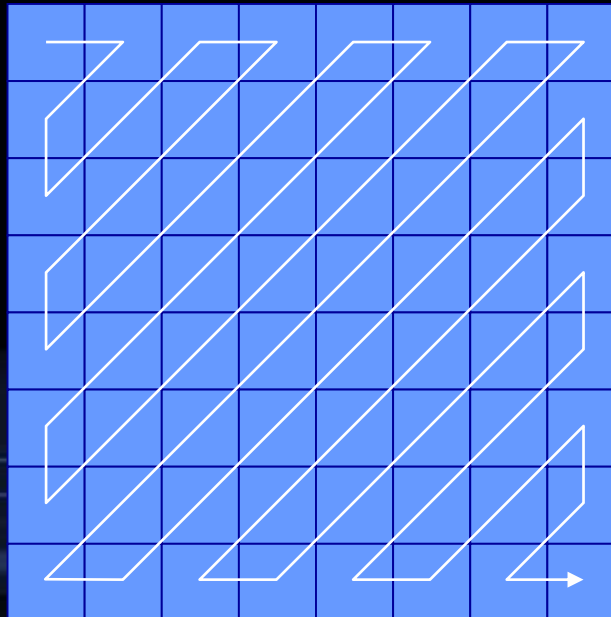
## ■ Operações a compressão JPEG



# Algoritmo de compressão JPEG

- **Ordenação dos coeficientes DCT**
  - Coeficientes DCT são ordenados em uma sequência zig-zag
    - para obter uma sequência unidimensional de dados para ser usado na codificação por entropia

coeficiente 0

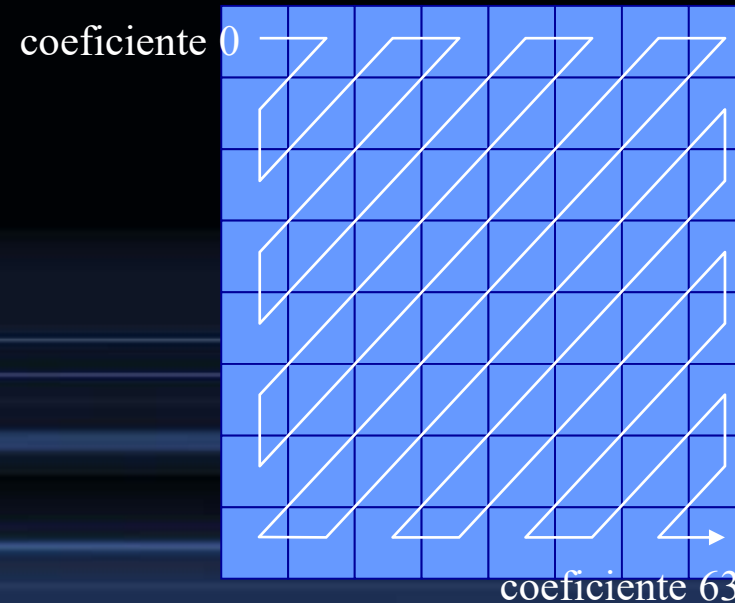


coeficiente 63

# Algoritmo de compressão JPEG

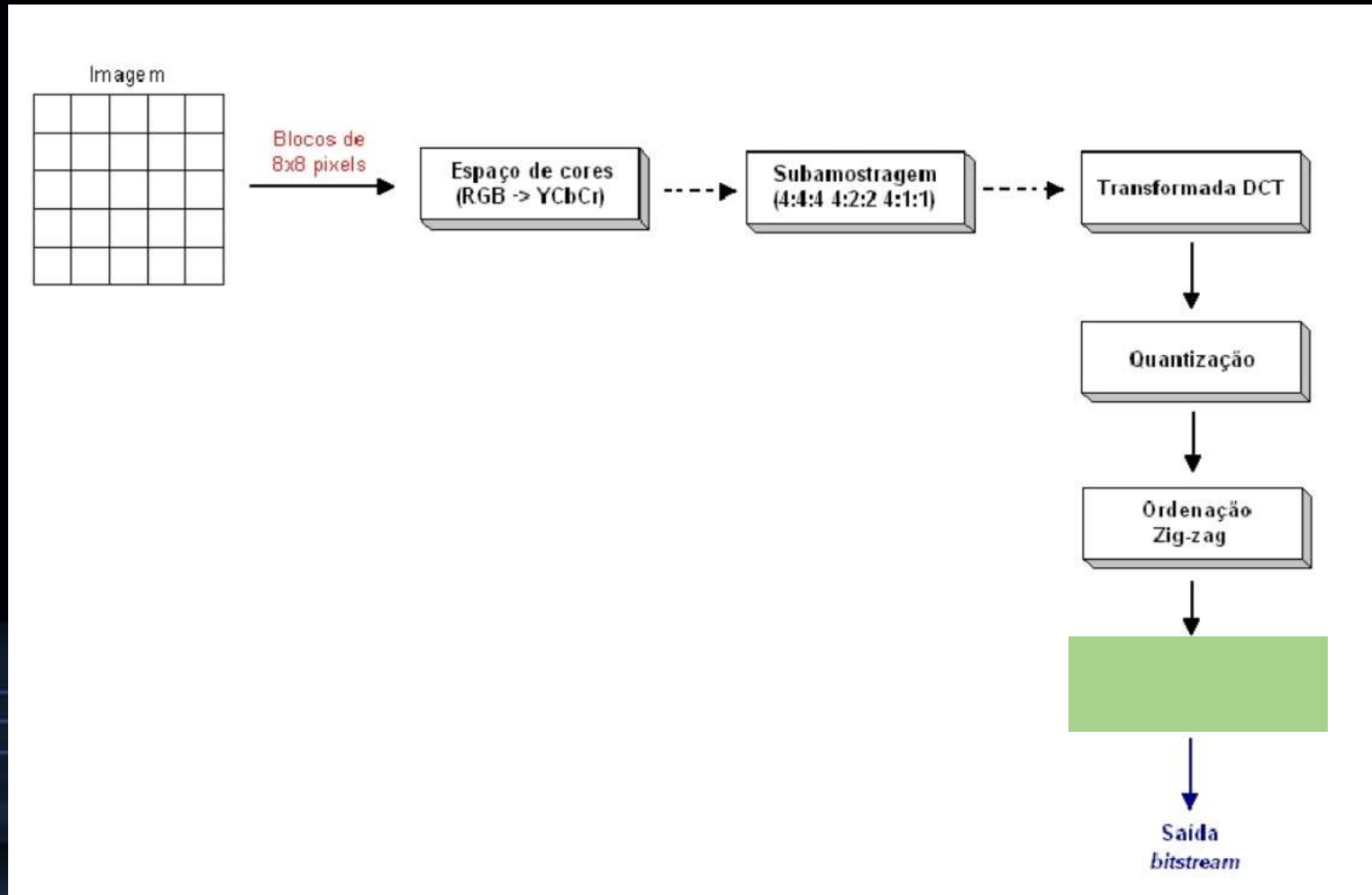
- **Ordenação dos coeficientes DCT**
  - Propósito do escaneamento zig-zag é ordenar os coeficientes em ordem decrescente de frequências espectral
    - coeficientes de alta frequências (no canto direito inferior) tem valores mais próximos a zero
      - isto leva a uma maior eficiência da codificação por entropia

1055	86	40	22	15	10	7	5
53	37	25	17	11	8	6	4
21	21	19	13	9	7	5	4
12	12	11	9	7	5	4	3
7	7	7	7	5	4	3	3
5	5	5	4	4	3	3	3
3	3	3	3	3	3	3	3
3	3	3	3	3	3	3	2



# Algoritmo de compressão JPEG

- **Operações a compressão JPEG**





# Algoritmo de compressão JPEG



- **Codificação por entropia**
  - Esta etapa fornece uma compressão adicional
  - JPEG define dois métodos de codificação por entropia
    - Codificação de Huffman
      - única especificado no modo baseline
    - Codificação aritmética
      - normalmente 10% mais eficiente que a codificação de Huffman

# Padrão de Compressão de Imagens JPEG



- **Taxas de compressão obtidas**
  - Quanto maior for a taxa de compressão maior será o número de componentes de alta frequência desprezados
    - para obter taxas de compressão muito elevadas é descartado um número significativo de componentes de alta frequência
      - levando ao aparecimento do efeito de bloco (perda de definição nos contornos das imagens).
  - Valores médios
    - Taxas de compressão de 10:1 a 20:1 – Alta qualidade de imagem
    - Taxas de compressão de 30:1 a 50:1 – Média qualidade de imagem
    - Taxas de compressão de 60:1 a 100:1 – Fraca qualidade de imagem

# Algoritmo de compressão JPEG

- **Taxas de compressão obtidas e qualidades**

- Alta qualidade

- Taxa de 2.6:1



- Boa qualidade

- Taxa de 15:1



# Algoritmo de compressão JPEG

- **Taxas de compressão obtidas e qualidades**

- Qualidade média

- Taxa de 23:1



- Baixa qualidade

- Taxa de 46:1





# Algoritmo de compressão JPEG

- **Taxas de compressão obtidas e qualidades**

- Mais baixa qualidade
  - Taxa de 144:1



- **Demonstração:**

- <https://cgjennings.ca/articles/jpeg-compression/>

# Padrão de Compressão de Imagens JPEG



- **JPEG é para imagens fotográficas**
  - JPEG apresenta ótimas taxas de compressão para imagens fotográficas naturais multitonais
  - Qualidade diminui consideravelmente quando aplicado a
    - imagens gráficas com contornos e áreas bem definidas de cor, ou
    - imagens com texto, como é o caso dos logotipos

# Padrão de Compressão de Imagens JPEG

- **JPEG é para imagens fotográficas**
  - JPEG apresenta ótimas taxas de compressão para imagens fotográficas naturais multitonais

JPEG (50kB)




PNG (177KB)



# Padrão de Compressão de Imagens JPEG

- **JPEG é para imagens fotográficas**
  - Qualidade diminui consideravelmente quando aplicado a
    - imagens gráficas com contornos e áreas bem definidas de cor, ou
    - imagens com texto, como é o caso dos logotipos



PNG vs JPEG

The image shows the text 'PNG vs JPEG' in a large, bold, black font. The 'PNG' part is sharp and clear, while the 'JPEG' part is significantly blurred and pixelated, illustrating the loss of detail when compressing text-based images with JPEG.



# Padrão de Compressão de Imagens JPEG



- **Para imagens gráficas e com texto**
  - JPEG introduz ruído nas zonas de imagem compostas por cores sólidas
    - pode distorcer o aspecto geral da imagem
  - Imagem PNG compactam mais eficazmente que JPEG e apresenta uma melhor definição dos contornos do texto



# Pontos Importantes

## Algoritmo JPEG

- Saber descrever cada etapa do algoritmo de compressão JPEG



**UFSC**

**Ciências da  
Computação**

# **CAP 3. COMPRESSÃO DE DADOS MULTIMÍDIA**

**INE5431 Sistemas Multimídia**  
**Prof. Roberto Willrich (INE/UFSC)**  
**roberto.willrich@ufsc.br**

**Aula 7: Padrões de compressão  
multimídia – Codec de Voz**

# Compressão de Dados Multimídia



- **Conteúdo:**

- Necessidade de compressão
- Princípios da compressão
- Classificação das técnicas de compressão
- Medição do desempenho de compressão
- Técnicas de compressão sem perdas
  - RLE, Huffman, LZW (GIF), Codificação Preditiva
- Técnicas de compressão de áudio, vídeo e imagens
- **Padrões de compressão multimídia**
  - JPEG, **Codecs de Voz**, MPEG, MPEG-4, H.261, H.263



# Recomendações ITU-TS para Voz



Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM não linear	3,4	8	64	20	160
G.721	ADPCM	3,4	8	32		
G.722	ADPCM sub-banda	7	16	48, 56, 64		
G.723.1m	MP-MLQ	3,4	8	6.4	30	24
G.723.1a	ACELP	3,4	8	5,3	30	20
G.726	ADPCM	3,4	8	16, 24, 32, 40	15	60
G.728	LD-CELP	3,4	8	16	20	40
G.729A	CS-CELP	3,4	8	8	20	20

# Codecs e Quadro de Voz



- **Quadro de voz**
  - A maioria dos codificadores de voz se baseia em quadros
    - Quadros de voz ou pacotes de voz
    - Cada quadro de voz tem uma duração de 1 a 30 ms
  - Codecs compactam quadros de voz
    - Contendo um número fixo de amostras
    - Número de amostras depende do codec utilizado

# CODECS E QUADRO DE VOZ



- **Montagem do quadro**
  - Fluxo de dados de áudio precisa ser acumulado
    - até que ele atinja o tamanho do bloco antes de ser processado pelo codificador
  - Acumulação de amostra leva tempo
    - soma-se ao atraso fim-a-fim.

# Codecs e Quadro de Voz





# Codecs e Quadro de Voz



# Codecs e QUADRO de Voz



- **Tamanho do quadro de voz (payload)**
  - Tamanho do payload (em bytes)
    - Taxa do Codec (em bits/sec) x tempo do quadro de voz (s)
    - Exemplo 1:
      - Codec G.711 => 64 kbps
      - Tamanho do pacotes = 20ms
      - Tamanho do payload =  $(64000 \times 0,02)/8 = 160$  bytes
      - Teríamos  $(1000/20) = 50$  pacotes de 160 B de dados a cada segundo
    - Exemplo 2:
      - Codec G.711 => 64 kbps
      - Tamanho do pacotes = 30ms
      - Tamanho do payload =  $(64000 \times 0,03)/8 = 240$  bytes
      - Teríamos 33 pacotes de 240 B de dados a cada segundo

# Codecs e QUADRO de Voz



- **Problema da sobrecarga de protocolos**
  - Para ser transmitido na rede, o quadro de voz deve ser encapsulado em diversos protocolos
    - Até chegar à camada de enlace, aos pacotes de voz vão ser adicionados 40 bytes: RTP (12 bytes) + UDP (8 bytes) + IP (20 bytes) = 40 bytes.
    - Se o quadro é pequeno, a sobrecarga de protocolos é maior:
      - Quadro de voz de 160B: pacote IP será de 40+160=200B, sobrecarga de  $40/200 = 20\%$
      - Quadro de voz de 240B: pacote IP será de 280B, sobrecarga de  $40/280 = 14\%$



# Codecs, Quadros e Pacotes de Voz



- **Relação entre tamanho de quadro de voz e atraso**

- Para redução do atraso, o codec escolhido deveria ter um quadro de quadro pequeno
- Exemplo 1:
  - Codec G.711 => 64 kbps
  - Tamanho do pacotes = 20ms
  - Tempo de empacotamento será 20ms
- Exemplo 2:
  - Codec G.711 => 64 kbps
  - Tamanho do pacotes = 30ms
  - Tempo de empacotamento será de 30ms



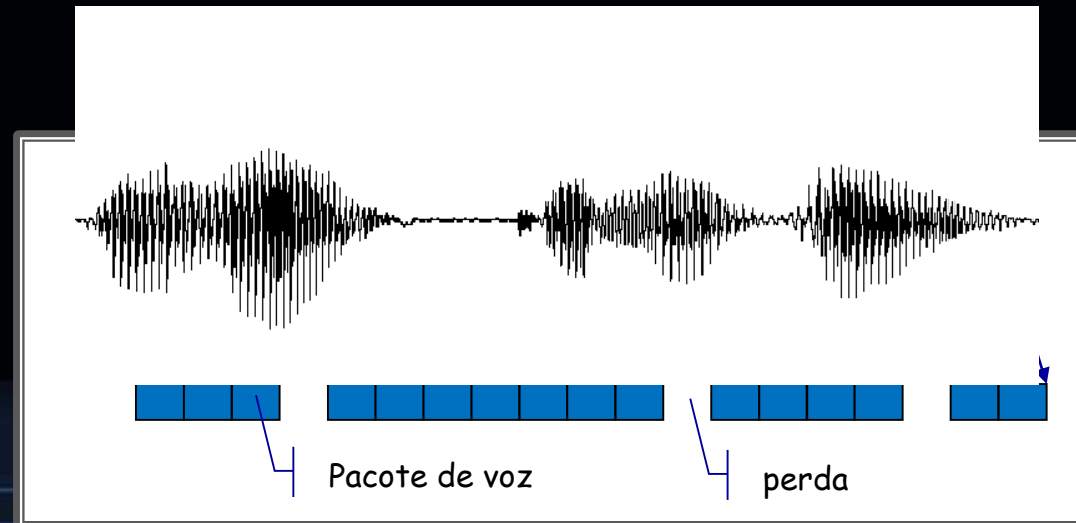
# Codecs, Quadros e Pacotes de Voz



- **Relação entre tamanho de quadro de voz e taxa de transmissão**
  - Quadro pequeno (i.e. menor atraso) gera uma maior taxa de bits devido a sobrecarga dos protocolos
  - Exemplo 1: Codec G.711 => 64 kbps
    - Tamanho do pacote de voz = 20ms
    - Tamanho do Pacote IP  $40 + 160 = 200$  B
    - Teríamos 50 pacotes IP de 200 B de dados a cada segundo
      - Taxa de bits é de  $50 * 200 * 8 = 80$  kbps
  - Exemplo 2: Codec G.711 => 64 kbps
    - Tamanho do pacote de voz = 30ms
    - Tamanho do pacote IP =  $40 + 240 = 280$  B
    - Teríamos 33 pacotes de 280 B de dados a cada segundo
      - Taxa de bits é de  $33 * 280 * 8 = 73,9$  kbps

# Codecs, Quadros e Pacotes de Voz

- **Relação entre tamanho de quadro de voz e qualidade na ocorrência de perda de pacotes**
  - Se o pacote de voz é maior, a perda gera um tempo maior de ausência de som na saída
  - Maior o pacote, pior a tolerância à perda de pacotes.



# Recomendações ITU-TS para Voz



## ■ G.711

- Usa PCM compandido (escala não linear)
  - Serve para aumentar a resolução de sinais de baixa amplitude
    - Mais importante para os humanos
  - Operando de forma análoga ao ouvido humano
- Dois tipos de escala
  - A-law (Europa e Brasil)
  - $\mu$ -law (EUA e Japão)
- Usado na maioria dos backbones telefônicos digitais
- Fluxo de bits de 64 kbps
  - 8 bits por amostra, 8000 amostras/s (uma amostra a cada 125 $\mu$ s)
- Supressão de silêncio é opcional
  - Reduz a taxa de bits gerada

Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM não linear	3,4	8	64	20	160

# Recomendações ITU-TS para Voz



## ■ **Tamanho do Payload**

- Quanto menor o payload de voz maior é a sobrecarga dos diversos protocolos de transmissão da voz
- Quanto maior o payload maior é o atraso na aplicação
  - para aguardar a montagem do payload

Recomendação	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	64	20	160
G.723.1m	6.4	30	24
G.723.1a	5.3	30	20
G.726	16, 24, 32, 40	15	60
G.728	16	20	40
G.729A	8	20	20



# Recomendações ITU-TS para Voz



## ■ G.721

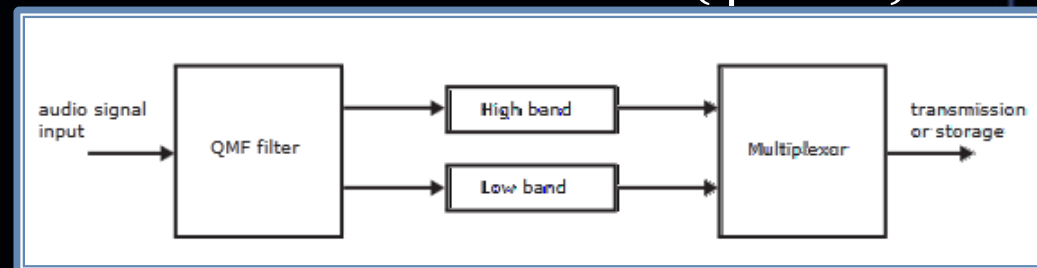
- Converte um fluxo de 64 kbps em um fluxo de 32 kbps aplicando uma compressão ADPCM
- A previsão e o tamanho do passo altera com o histórico do sinal
- Está obsoleto e substituído pelo G.726

Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM não linear	3,4	8	64	20	160
G.721	ADPCM	3,4	8	32		
G.722	ADPCM sub-banda	7	16	48, 56, 64		

# Recomendações ITU-TS para Voz

## ■ G.722

- Fornece uma melhor qualidade que o G.711 e G.721: Utiliza 14 bits por amostra
- ADPCM Sub-banda: sinal de voz é dividido em duas sub-bandas: alta (4-8kHz) e baixa (0-4kHz), no 64kbps:
  - 2 bits/amostra para banda alta (16 kbps)
  - 6 bits/amostra para banda baixa (48 kbps)
- Próprio para aplicações de videoconferência uma vez que telefones comuns não respondem na faixa de 7kHz



Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM não linear	3,4	8	64	20	160
G.721	ADPCM	3,4	8	32		
G.722	ADPCM sub-banda	7	16	48, 56, 64		

# Recomendações ITU-TS para Voz



## ■ G.723.1

- Opera a 6,4 kbps (Multipulse-Maximum Likelihood Quantification) e a 5,3 kbps (Algebraic-Code-Excited Linear Prediction)
- Em cada janela de 30 ms do sinal de voz
  - são analisadas 240 amostras de 16 bits do sinal de voz (tomadas a 8kHz) para identificação de padrões repetitivos (pitches) e são gerados 12 ou 10 códigos de 16 bits, conforme o algoritmo esteja configurado para uma taxa de 6,3 ou 5,3 kbps
- Valor típico de tamanho do pacote de voz (payload) é de 30ms (20 ou 24 bytes)

Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM não linear	3,4	8	64	20	160
G.723.1m	MP-MLQ	3,4	8	6.4	30	24
G.723.1a	ACELP	3,4	8	5.3	30	20

# Recomendações ITU-TS para Voz



## ■ G.726

- O G.726 utiliza o ADPCM a 40, 32, 24 e 16 kbps
- Sinal de voz é amostrado a 8kHz, codificado em 8 bits (leis A ou m) e são transmitidas diferenças entre amostras com 5, 4, 3 ou 2 bit em quantificação adaptativa
- Valor típico de tamanho do pacote de voz (payload) é de 15ms (60 bytes)

Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM ñ linear	3,4	8	64	20	160
G.723.1m	MP-MLQ	3,4	8	6.4	30	24
G.723.1a	ACELP	3,4	8	5.3	30	20
G.726	ADPCM	3,4	8	16, 24, 32, 40	15	60
G.728	LD-CELP	3,4	8	16	20	40
G.729A	CS-CELP	3,4	8	8	20	20



# Recomendações ITU-TS para Voz



## ■ G.728

- Técnica de codificação LD-CELP (Low-Delay, Code-Excited Linear Prediction), gerando uma taxa de bits de 16 kbps
- Tabela (codebook) utilizada é formada por 1024 ( $2^{10}$ ) valores
  - contém os valores de códigos (vetores) que representam as possíveis amostras do sinal de voz
- Em cada janela de 0,625ms do sinal de voz são analisadas 5 amostras de 8 bits e é gerado 1 código de 10 bits

Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM ã linear	3,4	8	64	20	160
G.723.1m	MP-MLQ	3,4	8	6.4	30	24
G.723.1a	ACELP	3,4	8	5.3	30	20
G.726	ADPCM	3,4	8	16, 24, 32, 40	15	60
G.728	LD-CELP	3,4	8	16	20	40
G.729A	CS-CELP	3,4	8	8	20	20

# Recomendações ITU-TS para Voz



## ■ G.729

- Bastante popular em aplicações de voz sobre frame relay e em modems V.70 para voz e dados
- G.729 Técnica de codificação LD-CELP gerando uma taxa de bits de 8 kpbs e G.729A a codificação CS-ACELP (Algebraic-ACELP)

Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM ñ linear	3,4	8	64	20	160
G.723.1m	MP-MLQ	3,4	8	6.4	30	24
G.723.1a	ACELP	3,4	8	5.3	30	20
G.726	ADPCM	3,4	8	16, 24, 32, 40	15	60
G.728	LD-CELP	3,4	8	16	20	40
G.729A	CS-CELP	3,4	8	8	20	20

# Recomendações ITU-TS para Voz



## ■ G.729

- Em cada janela de 10ms do sinal de voz são analisadas 80 amostras de 8 bits para geração de 10 códigos de 8 bits
- Valor típico de tamanho do pacote de voz é de 20ms (20 bytes)

Recomend.	Técnica de Compressão	Largura de banda da voz (kHz)	Taxa de amostragem (kHz)	Taxa de bit compactado (kbps)	Duração do payload (ms)	Tamanho do payload (bytes)
G.711	PCM ñ linear	3,4	8	64	20	160
G.723.1m	MP-MLQ	3,4	8	6.4	30	24
G.723.1a	ACELP	3,4	8	5.3	30	20
G.726	ADPCM	3,4	8	16, 24, 32, 40	15	60
G.728	LD-CELP	3,4	8	16	20	40
G.729A	CS-CELP	3,4	8	8	20	20

# Supressão de Silêncio e remoção de sons repetitivos

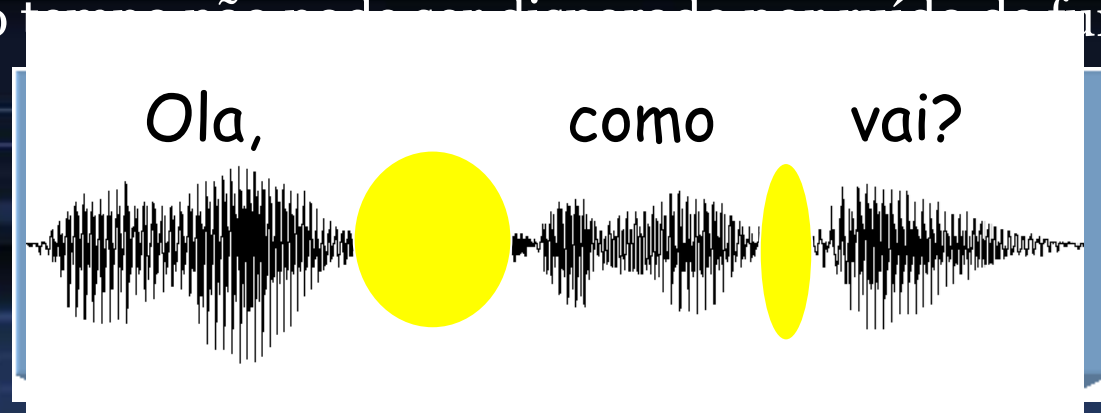


- **Compressão da voz via remoção dos períodos de silêncio e de informações redundantes encontradas na fala humana**
  - Existem informações na fala humana que não são necessárias para que uma comunicação efetiva exista através de uma rede
- **Sons repetitivos, inerentes à voz, são causados pela vibração das cordas vocais**
  - transmissão destes sons idênticos não é necessária para efetivação da comunicação e a sua remoção resulta em um aumento de eficiência na utilização da banda de rede
- **Composição da fala**
  - 22% do que se fala são componentes essenciais da comunicação
    - devem ser transmitidos para o entendimento do diálogo
  - 22% são padrões repetitivos
  - 56% representa as pausas entre falas



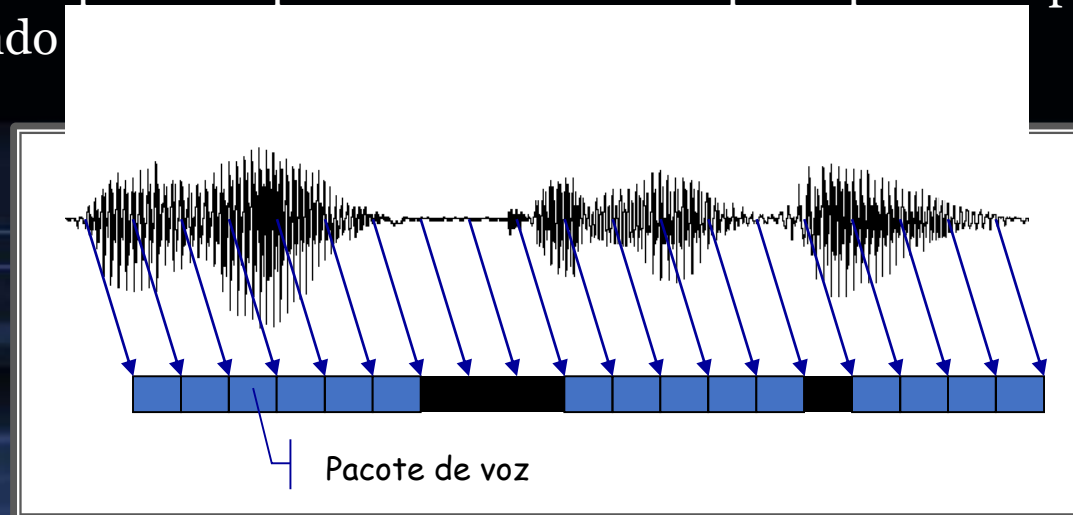
# Supressão do silêncio: Componentes

- **VAD (Detector de Presença de Voz)**
  - Responsável por determinar quando o usuário está conversando e quando ele está em silêncio
  - É útil para economizar energia no caso de dispositivos que funcionam a bateria
  - Deve ser bastante sensível
    - Caso contrário, o início das palavras podem ser perdidas e um silêncio inútil pode ser incluído no final das sentenças
    - Mas ao mesmo tempo, não pode ser tão sensível a ruídos de fundo.



# Supressão do silêncio: Componentes

- **DTX (Discontinuous Transmission)**
  - Capacidade de um codec de parar de transmitir quadros quando o VAD tiver detectado um período de silêncio
  - VAD + DTX: modo eficiente de liberar dinamicamente a banda
    - proporcionando uma economia de até 50% da banda
  - Alguns codecs avançados não vão interromper a transmissão completamente
    - Em vez disso, vão para um modo de silêncio no qual usam muito menos largura de banda e enviam apenas os parâmetros mínimos para que o receptor possa restituir o ruído de fundo



# Supressão de Silêncio e remoção de sons repetitivos



- **Alguns pontos devem ser considerados na supressão do silêncio**
  - Quando a fala é muito frequente, contínua, os ganhos com a supressão do silêncio não são alcançados;
  - Como a detecção da presença de voz na transmissão não é imediata
    - Pode ocorrer o corte das primeiras sílabas da locução
    - Fenômeno é denominado de clipping;
  - Quando o ruído de fundo é muito alto
    - Torna-se difícil distinguir entre o que é ruído e o que realmente é fala
    - Corre-se o perigo de empacotamento de ruído.

# Pontos Importantes

## Codecs de Voz

- Entender o que são pacotes de voz e a relação do tamanho do pacote de voz com taxa de bits, atraso e impacto na qualidade quando da perda de pacotes

## Supressão de Silêncio

- Entender as vantagens e limitações



**UFSC**

**Ciências da  
Computação**

# **CAP 3. COMPRESSÃO DE DADOS MULTIMÍDIA**

**INE5431 Sistemas Multimídia**  
**Prof. Roberto Willrich (INE/UFSC)**  
**roberto.willrich@ufsc.br**

**Aula 8: Padrões de compressão  
multimídia – MPEG Áudio**



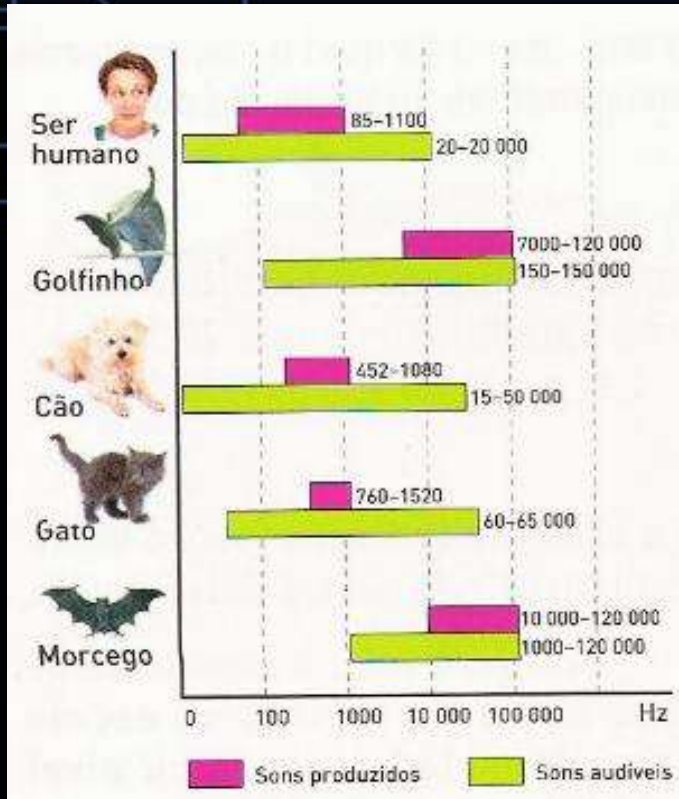
# Compressão de Dados Multimídia



- **Conteúdo:**

- Necessidade de compressão
- Princípios da compressão
- Classificação das técnicas de compressão
- Medição do desempenho de compressão
- Técnicas de compressão sem perdas
  - RLE, Huffman, LZW (GIF), Codificação Preditiva
- Técnicas de compressão de áudio, vídeo e imagens
- **Padrões de compressão multimídia**
  - JPEG, Codecs de Voz, **MPEG Áudio**, MPEG Vídeo, MPEG-4, H.261, H.263

# MPEG-Audio



- **Padrão de compressão de áudio genérico (até 20 kHz)**

- E não apenas para voz (de 3,4 a 7 kHz)
- Explora a percepção humana e não as características da fonte do áudio

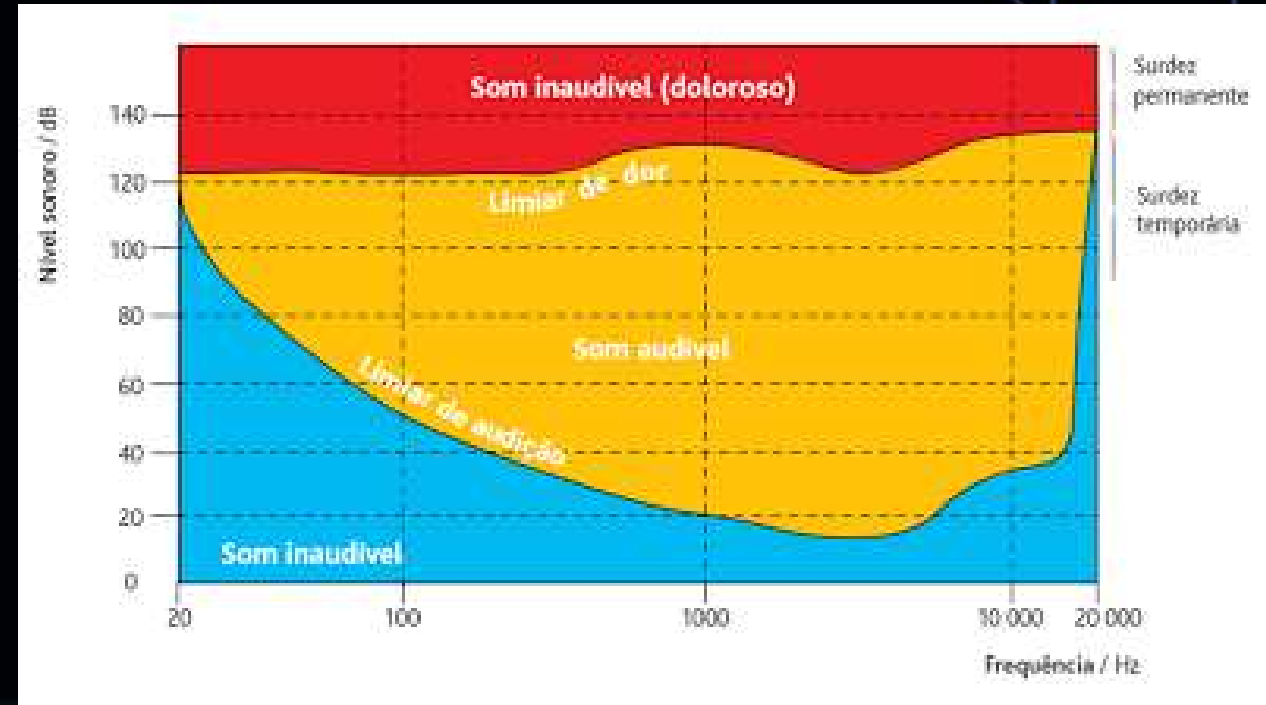
- **Princípio de Compressão**

- Faixa de frequência audível humana
  - Filtra sons acima de 20 KHz

# MPEG-Audio

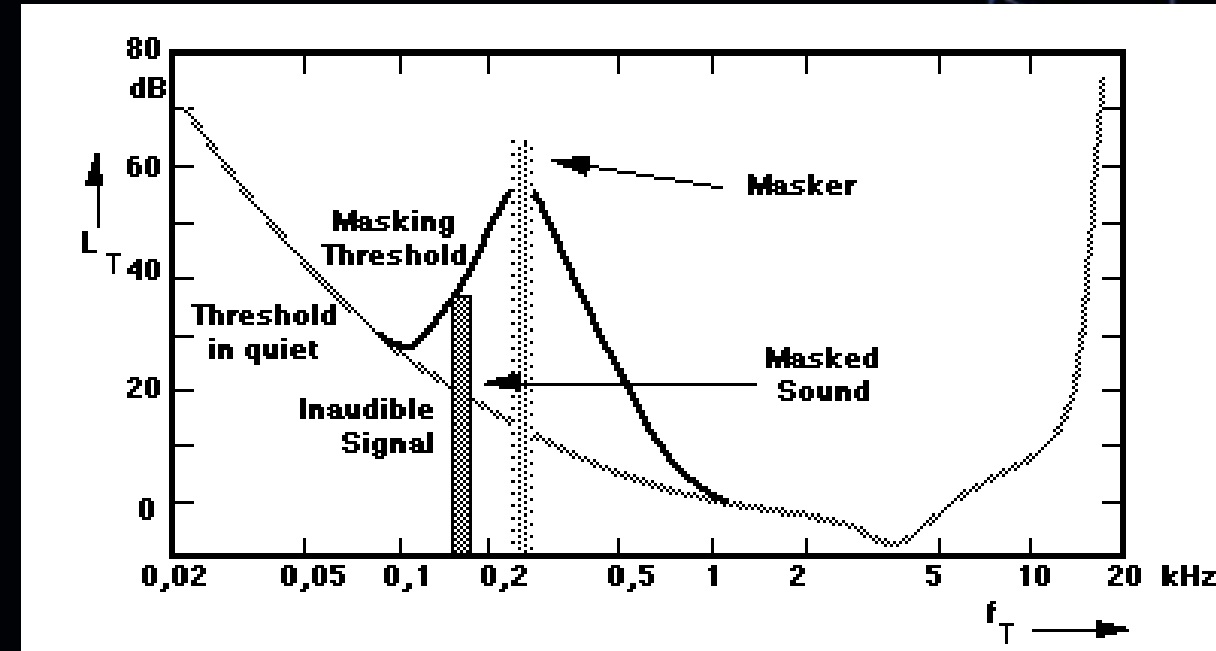
## ■ Princípio de Compressão

- Limiar de audição na faixa de frequência audível
  - Explora a curva de percepção da audição humana dentro da faixa de frequências audíveis (limiar de audição)
    - Sensibilidade para sons dentro desta faixa não é uniforme (depende da frequência)
    - O que se faz é descartar amostras que se encontrem abaixo deste limiar.



# MPEG-Audio

- **Princípio de Compressão**
  - Mascaragem: um som pode tornar outro impossível de ser ouvido, ou pode tornar o outro sem peso
    - tipos de mascaragem: total ou parcial
    - sons mascarados podem ser descartados (não são audíveis)
    - característica explorada pelo padrão MPEG-Áudio
      - explora as limitações perceptivas do sistema auditivo humano



# MPEG-Audio



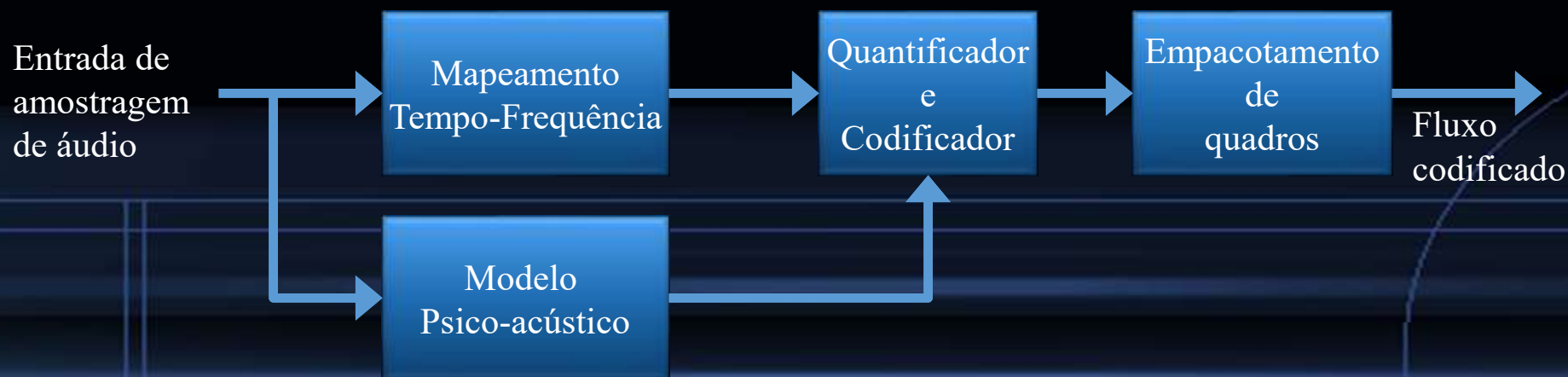
- **Principais características do MPEG-1 Audio:**
  - Sequência de bits compactada pode suportar um ou dois canais
    - um canal único
    - dois canais independentes
    - um sinal estéreo
  - Três taxas de amostragens
    - 32, 44.1 ou 48 kHz
    - MPEG-2.5 (não oficial) 8, 11.025, 12, 16, 22.05 e 24 kHz.
  - Fluxo compactado pode ter uma das várias taxas de bits fixas e predefinidas variando de 32 a 320 kbps
  - Padrão MPEG-2.5 (não oficial)
    - Taxa de bits de 8, 16, 24, e 144 kbps
  - Razão de compressão: 2,7 a 24 (depende da taxa de amostragem)
    - 6:1 ouvintes experientes não detectam diferenças



# Um codificador básico MPEG-Áudio



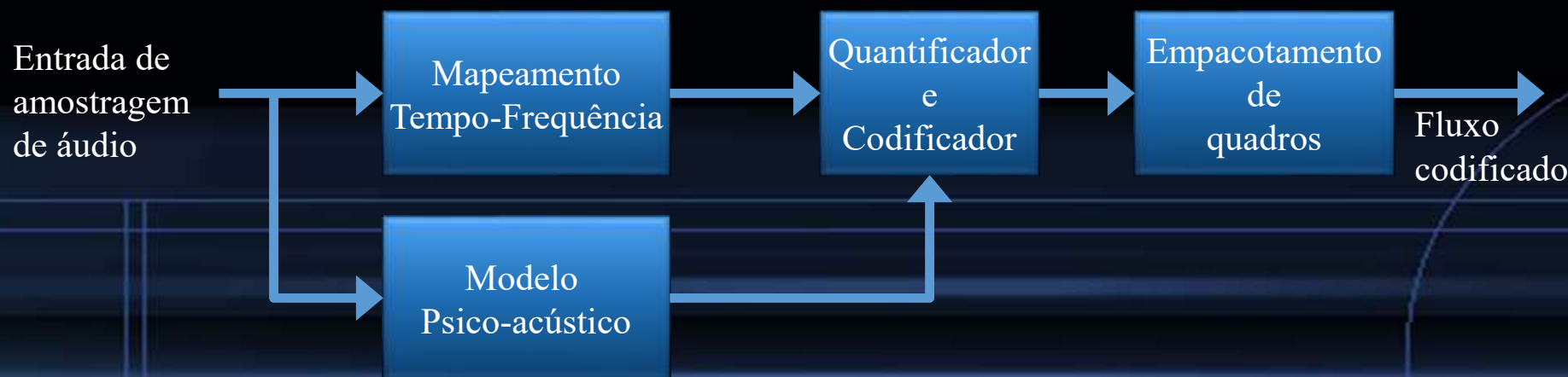
- **Bloco mapeamento tempo-frequência**
  - Divide a entrada em sub-bandas de frequências múltiplas
- **Bloco modelo psico-acústico**
  - Cria um conjunto de dados para controlar a operação do bloco quantificador e codificador
    - Considera limiar de audição, sons mascarados, etc.



# Um codificador básico MPEG-Áudio



- **Bloco quantificador e codificador**
  - Cria um conjunto de símbolos de código
    - sub-bandas menos importantes e áudios inaudíveis são removidos
- **Bloco Empacotamento de quadros**
  - Monta e formata os símbolos de código e adiciona outras informações



# MPEG-1 Audio



- **MPEG Audio especifica uma família de 3 esquemas de codificação de áudio**
  - Chamadas de Layer-1, Layer-2 e Layer-3
    - de Layer-1 a Layer-3, a complexidade e desempenho (qualidade de som e taxa de bits) aumentam
  - Os três codificadores são compatíveis no modo hierárquico
    - decodificador Layer-N é capaz de decodificar um fluxo de bits fluxo codificado com codificador Layer-N e abaixo de N
  - MP3 é MPEG-1 Layer-3
- **Padrão especifica o formato do fluxo de bits e o decodificador para cada esquema de codificação**
  - não especifica o codificador para avanços futuros

# MPEG-2 Áudio



- **Estende as funcionalidades do MPEG-1 Áudio**
  - cinco canais (esquerdo, direito, centro, e dois canais surround)
  - mais um canal de baixa frequência
  - ou sete canais multilíngues/comentários
  - tem taxas de amostragens adicionais

# Pontos Importantes

## MPEG Áudio

- Entender os princípios gerais da compressão





**UFSC**

**Ciências da  
Computação**

# **CAP 3. COMPRESSÃO DE DADOS MULTIMÍDIA**

**INE5431 Sistemas Multimídia**  
**Prof. Roberto Willrich (INE/UFSC)**  
**roberto.willrich@ufsc.br**

**Aula 9: Padrões de compressão  
multimídia – MPEG Vídeo**

# Compressão de Dados Multimídia



- **Conteúdo:**

- Necessidade de compressão
- Princípios da compressão
- Classificação das técnicas de compressão
- Medição do desempenho de compressão
- Técnicas de compressão sem perdas
  - RLE, Huffman, LZW (GIF), Codificação Preditiva
- Técnicas de compressão de áudio, vídeo e imagens
- **Padrões de compressão multimídia**
  - JPEG, Codecs de Voz, MPEG Áudio, **MPEG Vídeo**, H.261, H.263

# Padrão ISO/IEC MPEG



- **Motion Picture Expert Group**

- Grupo de padrões de representação codificada de vídeos, áudios e suas combinações
  - armazenados e recuperados em Digital Storage Media (DSM)
    - dispositivos de armazenamento convencionais, CD-ROMs, drivers de fita, HDs, drivers ótico escrevíveis e canais de telecomunicação (redes de longa distância, locais, etc.)

- **Características**

- MPEG usa a compressão intra e inter-quadros de vídeos
  - obtém altas taxas de compressão devido a alta redundância dos vídeos
- Especificações MPEG também incluem um algoritmo para compressão de áudio
  - Compressão do áudio associado e a sincronização áudio-visual não podem ser independente da compressão do vídeo

# Padrão ISO/IEC MPEG



## ■ **Vários itens de trabalho**

### ▣ MPEG-1 (1993)

- Vídeo pode codificar imagens de até 4096x4096 pixels e 60 fps.
- A maioria das aplicações usam o formato SIF, com 240x352 pixels e 30fps, e subamostragem de croma 4:2:0.

### ▣ MPEG-2 (1994)

- Pode codificar imagens de até 16.383 x 16.383 pixels
- Padrão organizado em perfis e níveis. Exemplos:
  - Nível baixo (240 x 352 pixels x 30 fps – idêntico ao SIF MPEG 1),
  - Nível principal, visando a codificação com qualidade de TV (720 x 480 e 30 fps), e
  - Níveis alto, visando a TV de alta resolução - HDTV, e a produção de filmes (em geral 1280 x 720 e 30fps; 1920 x 1080 e 30 fps ou 1440 x 1152 e 30 fps).
- O padrão permite subamostragem de croma 4:2:0, 4:2:2 e 4:4:4.

# Padrão ISO/IEC MPEG



- **Vários itens de trabalho**

- MPEG-3

- Para vídeo com qualidade HDTV na taxa de 40 Mbps
    - interrompido em julho 1992

- MPEG-4 (1998)

- Objetivo inicial: codificação para audiovisual com muito baixa taxa de transmissão (variando de 4,8 a 64 Kbps)
    - Hoje: oferece soluções para vários tipos de aplicações com qualidades diferentes

- MPEG-7 (2001)

- Interface de Descrição de Conteúdo Multimídia: um padrão de descrição de dados multimídia (informações audiovisuais)
    - Permitindo a busca e filtragem



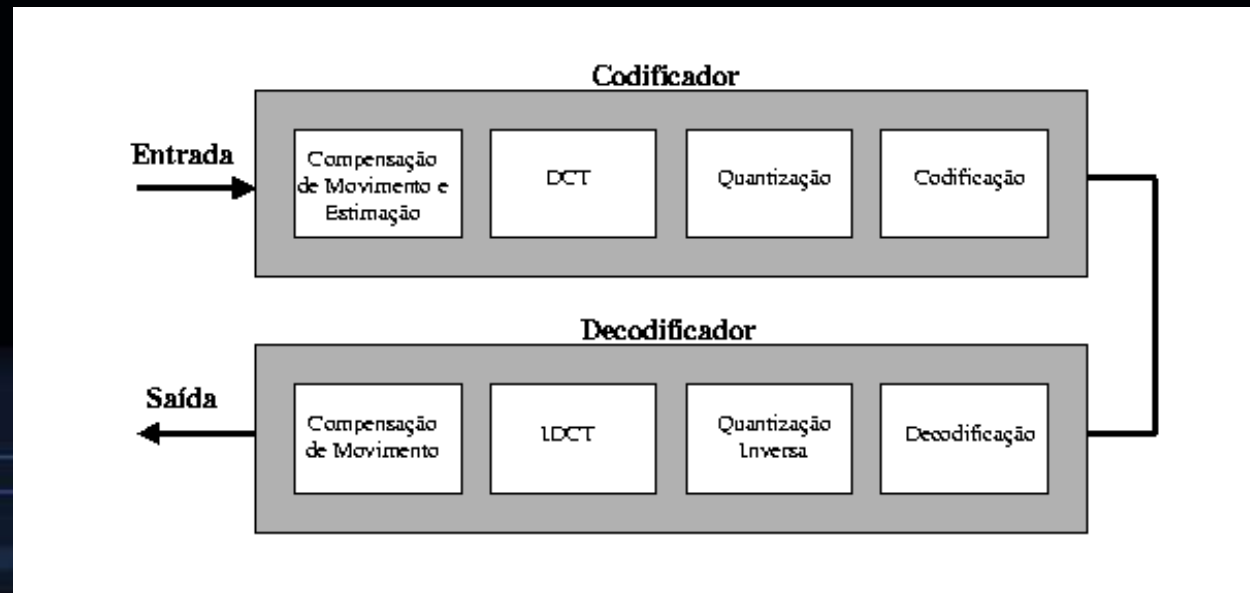
# Padrão ISO/IEC MPEG 1



- **MPEG 1 (ISO/IEC-11172) tem cinco partes**
  - ISO/IEC-11172-1 MPEG-Sistemas
    - Sincronização e multiplexação de fluxos de áudio e vídeo compactados
  - ISO/IEC-11172-2 MPEG-Vídeo
    - Compressão de sinais de vídeo;
  - ISO/IEC-11172-3 MPEG-Áudio
    - Compressão de um sinal de áudio digital
  - ISO/IEC-11172-4 Teste de Conformidade
    - Especifica procedimento para determinar as características dos fluxos codificados e para testar a conformidade com os requisitos identificados no Áudio, Vídeo e Sistemas
  - ISO/IEC-11172-5 Simulação de Software
    - Oferece uma implementação de referência.
- **MPEG separa áudio e vídeo**
  - Compressão em três camadas: a camada de sistema, de áudio e de vídeo

# Padrão ISO/IEC MPEG-1 Vídeo

- **MPEG especifica a sintaxe dos fluxos codificados para que decodificadores possam decodificar**
  - como gerar o bitstream não é padronizado
    - permite inovações no projeto e implementação de codificadores



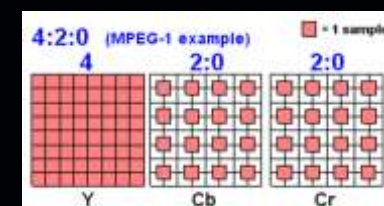
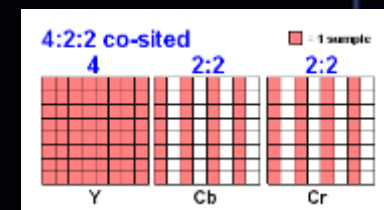
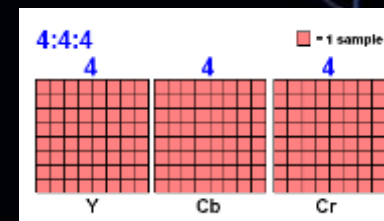
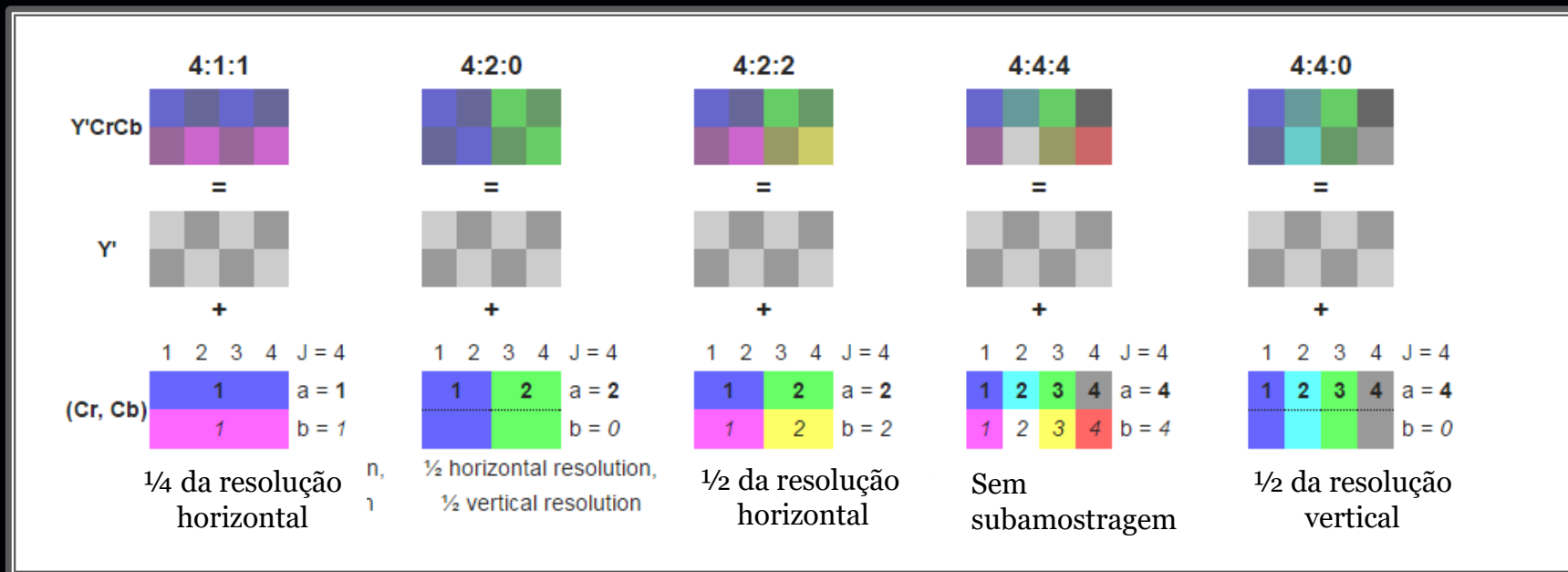
# Padrão ISO/IEC MPEG-1 Vídeo

- **Imagem MPEG (quadro do vídeo)**
  - É a unidade elementar para a codificação do vídeo
  - Imagens são representadas no espaço de cores YCbCr
    - Grupo de três matrizes retangulares que representam a luminância (Y) e a croma (Cr e Cb)
    - É preferível YCrCb pois o olho é mais sensível a luminosidade que a croma
      - Podem ser feita subamostragens nas matrizes Cb e Cr



# Padrão ISO/IEC MPEG-1 Vídeo

## ■ Subamostragens YCbCr



Original sem sub-amostragem de croma, 200% de zoom



Original após sub-amostragem de croma, 200% de zoom

# Padrão ISO/IEC MPEG-1 Vídeo

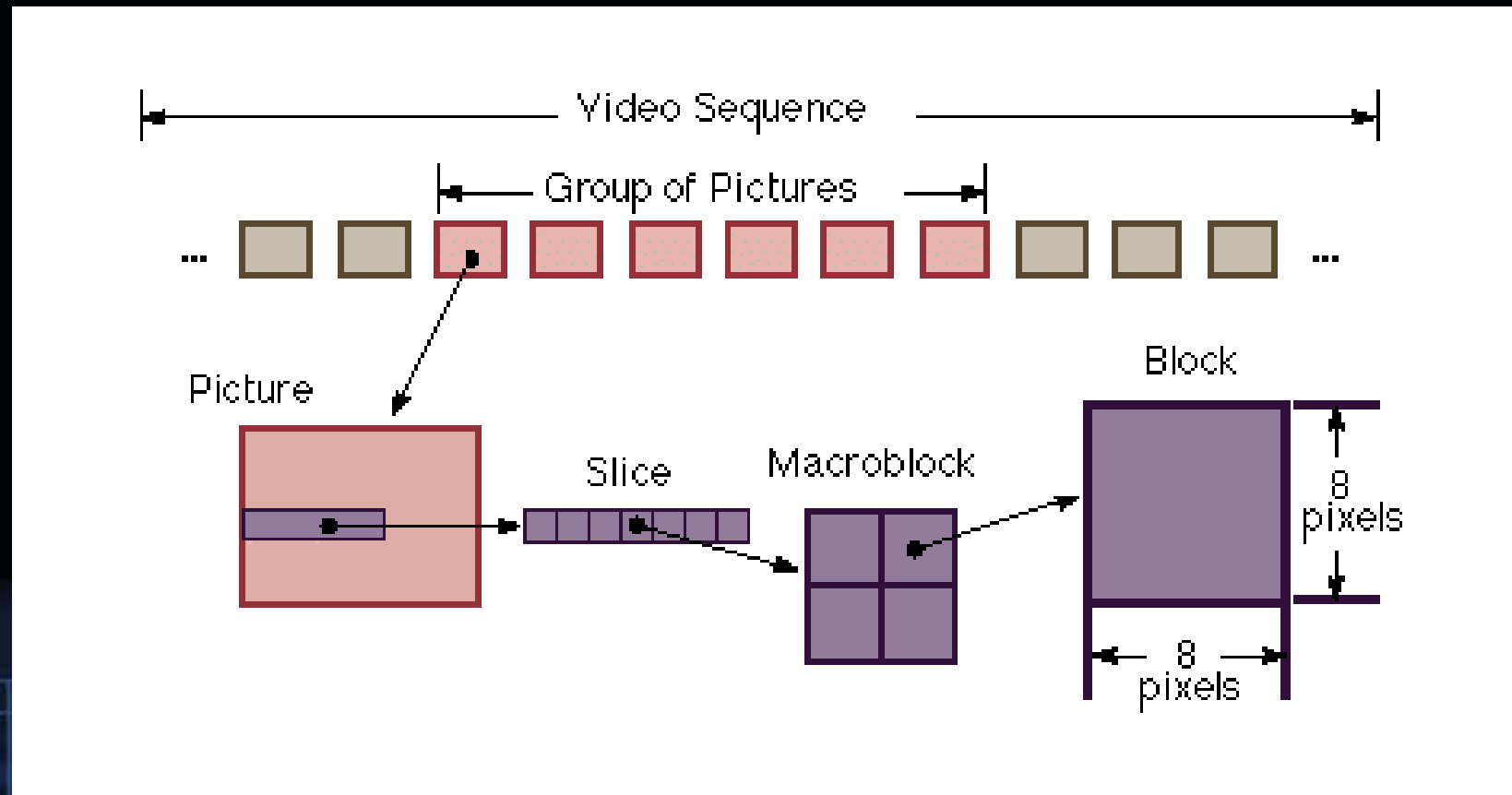


- **Compactação MPEG-1 Video**
  - Em vídeo existem dois tipos de redundância: espacial e temporal
    - MPEG-1 explora estas duas redundâncias
- **Redundância espacial**
  - Pode ser explorada pela simples codificação em separado de cada quadro
  - Blocos de 8x8 pixels são compactados similar a uma compressão JPEG



# Padrão ISO/IEC MPEG-1 Vídeo

- Hierarquia do fluxo de dados MPEG-1**



# Padrão ISO/IEC MPEG-1 Vídeo

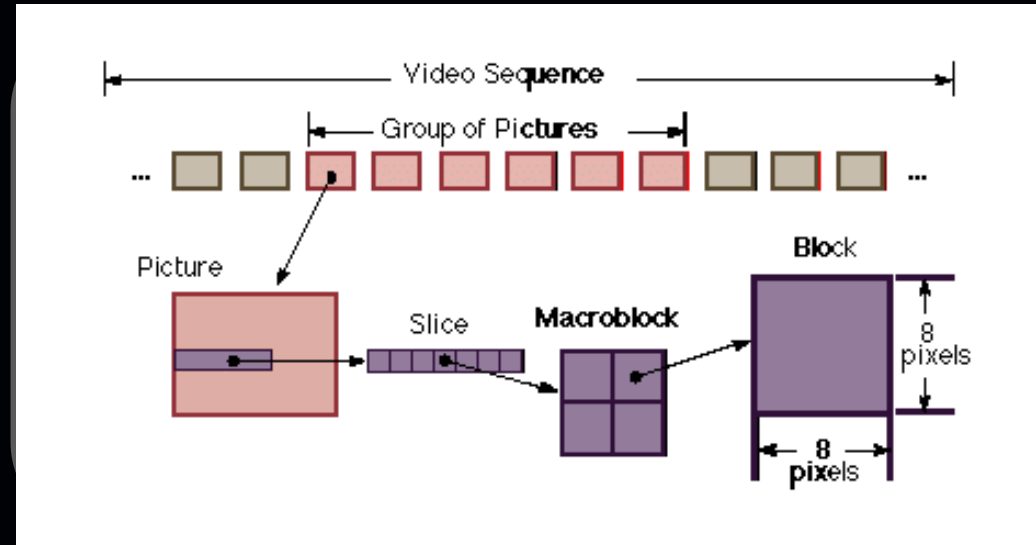
## ■ Fluxo de vídeo MPEG-1:



- GOP (Grupo de imagens): fornece um ponto de acesso aleatório
- Camada de imagem contem todas as informações codificadas de uma imagem
  - cabeçalho contem a referência temporal de uma imagem, o tipo de codificação, etc..

# Padrão ISO/IEC MPEG-1 Vídeo

## ■ Fluxo de vídeo

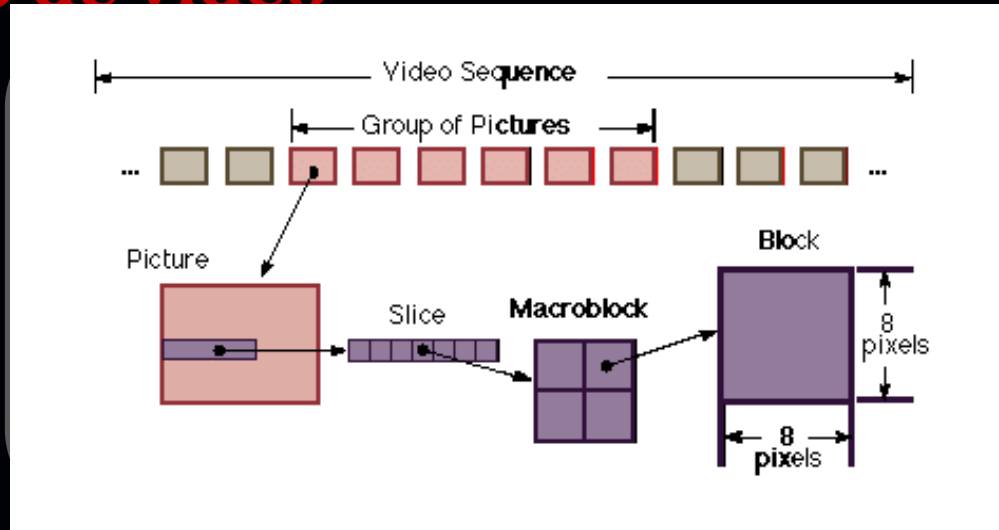


## ■ Imagens (Quadro do Vídeo)

- Podem ser codificadas com diferentes I-Frame, P-Frame, B-Frame
  - Compactados eliminando redundância espacial (I-Frame) ou espacial/temporal (P e B-Frames)
  - Compactados descartando informações pouco relevantes para a percepção humana

# Padrão ISO/IEC MPEG-1 Vídeo

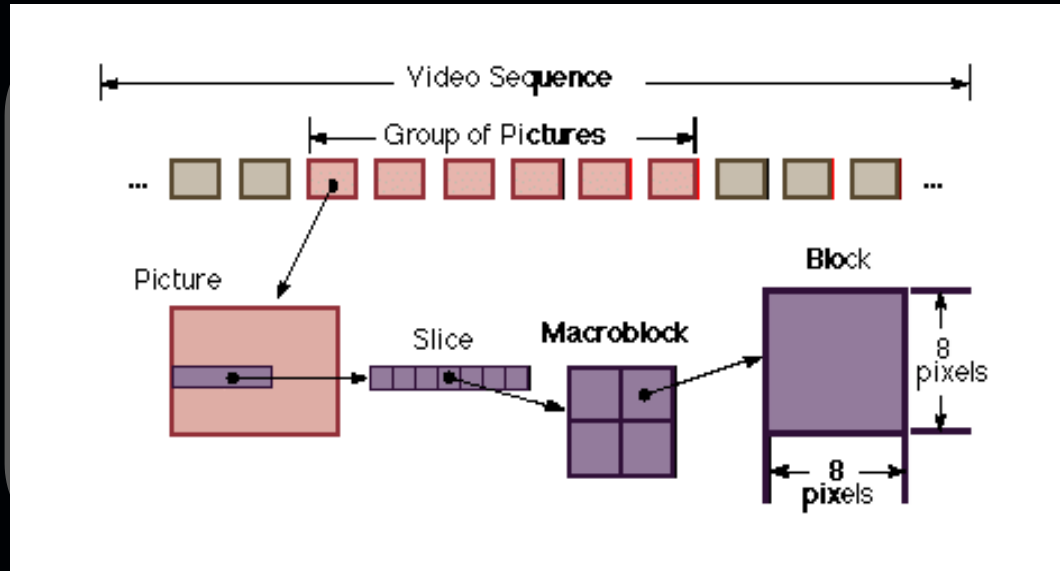
## ■ Fluxo de vídeo



- Imagens são divididas em pedaços (slices):
  - Número de pedaços podem ser variáveis
  - Cada pedaço consiste de um número variável de macroblocos (16x16 pixels)
  - Cada pedaço pode ser codificados de maneira diferente
  - Importante para o controle de erro
    - Se existe um erro no fluxo de dados, o decodificador pode saltar um pedaço
    - Maior o número de pedaços, melhor é o tratamento de erro

# Padrão ISO/IEC MPEG-1 Vídeo

## ■ Fluxo de vídeo



### □ Macrobloco

- Usado na estimativa e compensação do movimento

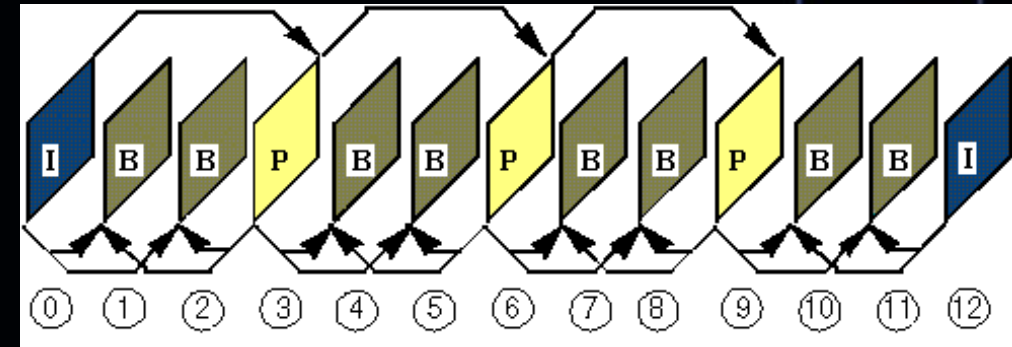
### □ Bloco de Imagem

- um bloco é uma matriz 8x8 pixels tratados como unidades e entrada para o DCT



# Padrão ISO/IEC MPEG-1 Vídeo

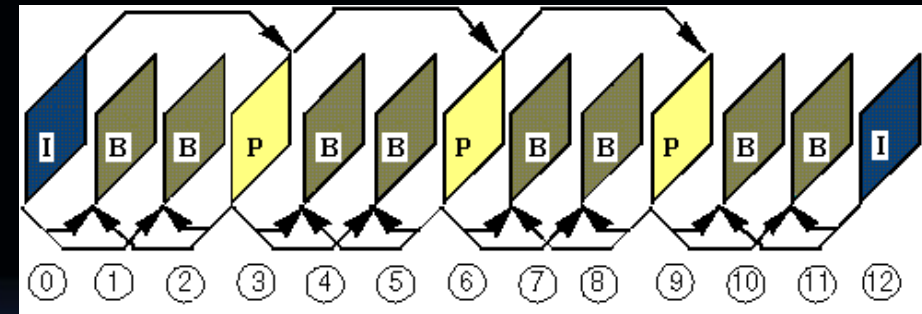
- **Grupo de Imagem consiste de quatro tipos de quadros:**
  - Quadros I (Intracoded)
    - imagens estáticas, independentes e codificadas com o JPEG.
  - Quadro P (Predictive)
    - diferença bloco a bloco com o quadro I ou P anterior
  - Quadro B (Birectional)
    - diferença com o último quadro e com o quadro seguinte
  - Quadro D (DC-coded)
    - Médias de bloco usadas para o avanço rápido (fast forward).



# Padrão ISO/IEC MPEG-1 Vídeo

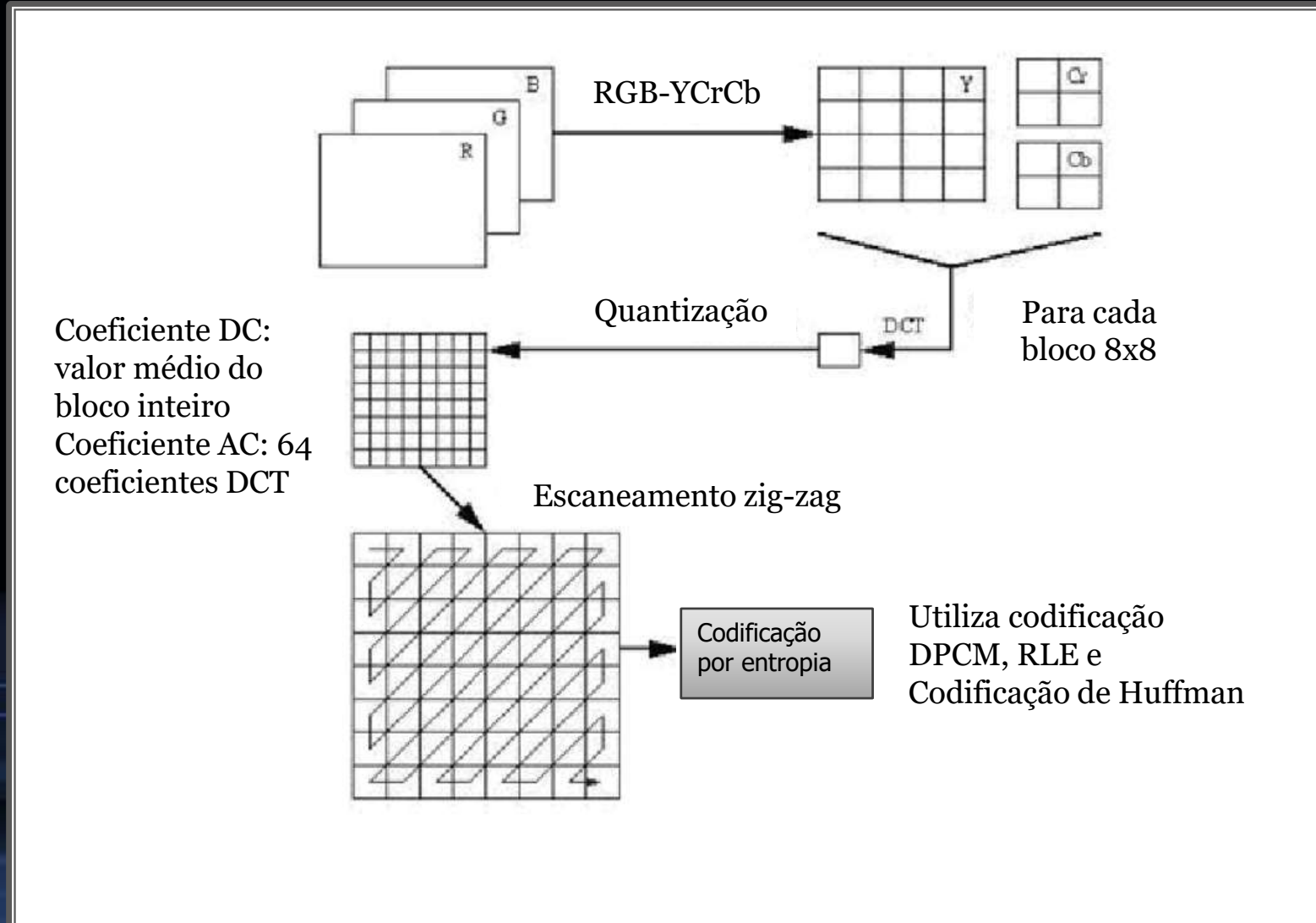
## ■ Quadro I

- Imagens estáticas, independentes e codificadas com o JPEG
- É necessário que quadros I apareçam periodicamente no fluxo de saída
  - no caso de transmissão multicast
    - receptores podem entrar no grupo em tempos distintos, requerendo um quadro I para começar a decodificação MPEG-1
  - se um quadro for recebido com erro
    - decodificação não será mais possível
- Quadros I são inseridos na saída uma ou duas vezes por segundo



# Padrão ISO/IEC MPEG-1 Vídeo

- **Compressão dos blocos 8x8 píxeis**



# Padrão ISO/IEC MPEG-1 Vídeo

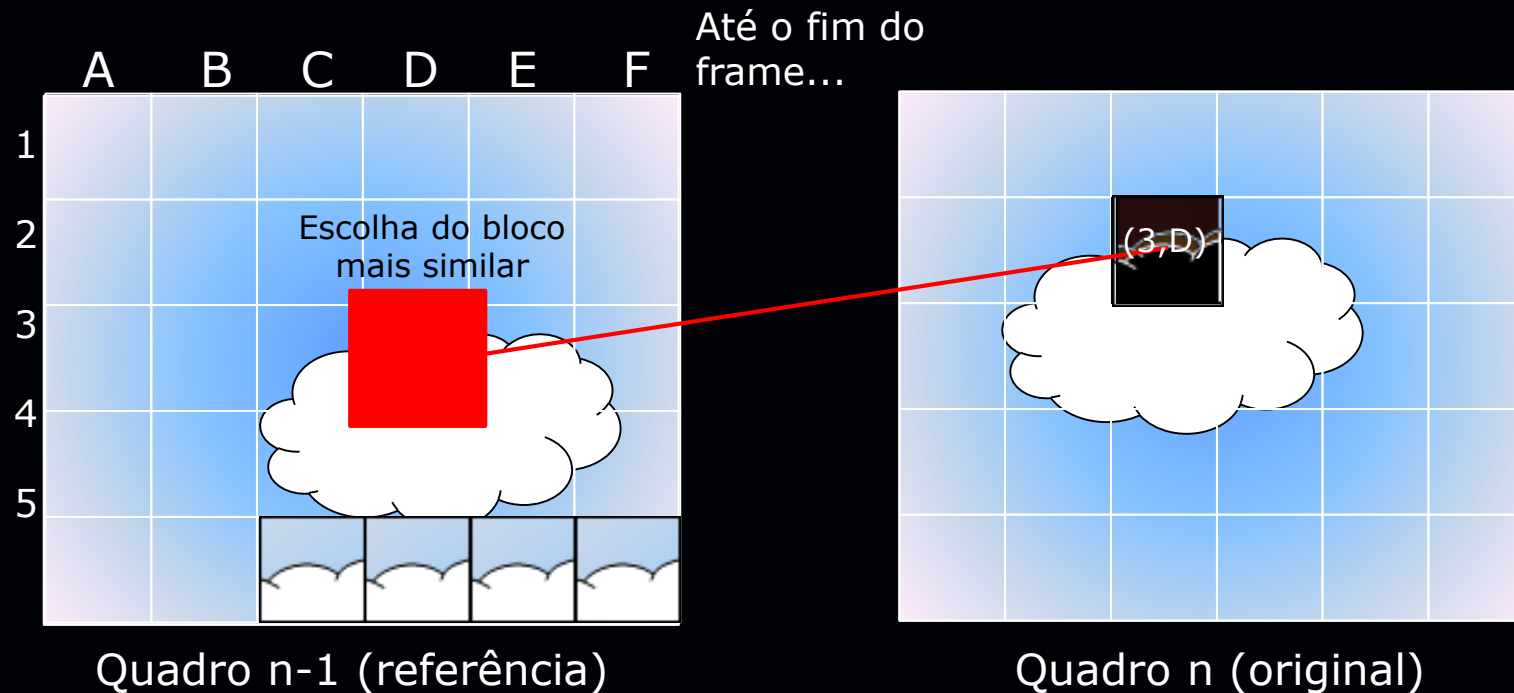
- **Quadros P e B exploram a redundância temporal**
  - Compactação adicional pode ser obtida explorando o fato de que dois quadros consecutivos são, com frequência, quase idêntico
  - MPEG faz compensação de movimento
    - Calcula o vetor de movimento dos macroblocos e a diferença macrobloco a macrobloco





# Padrão ISO/IEC MPEG-1 Vídeo

- **Exemplo simples: Compara a similaridade entre blocos**



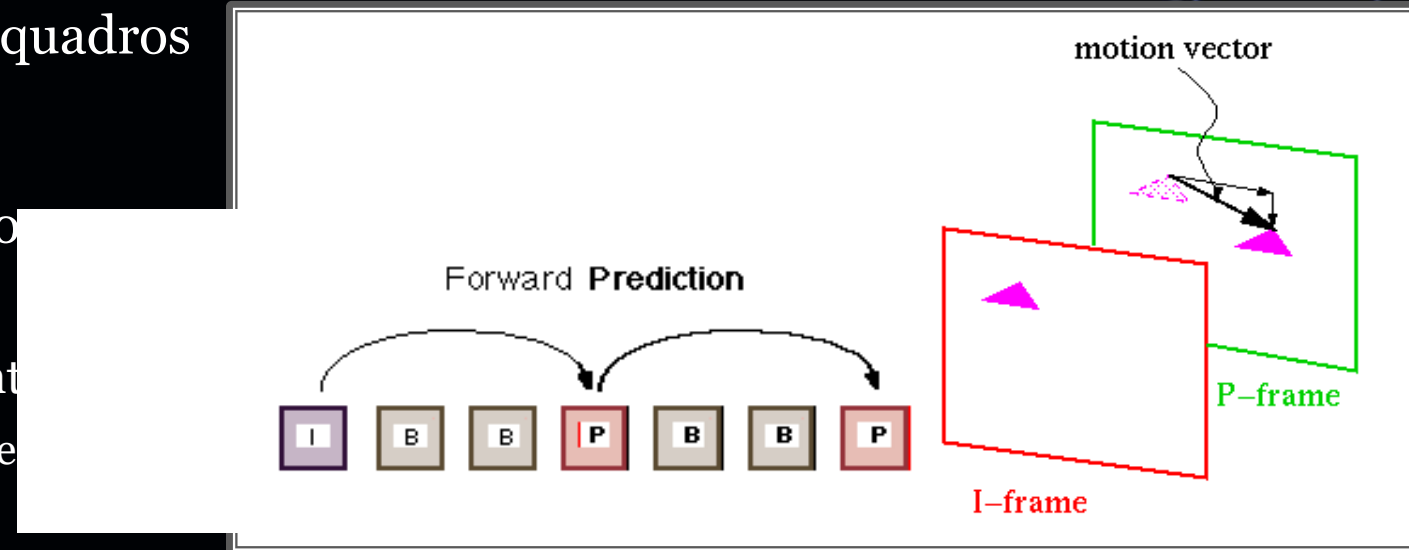
- ▣ Mantém a diferença entre os blocos (resíduo);
- ▣ Cria o vetor de movimento, referenciando o bloco do quadro anterior;



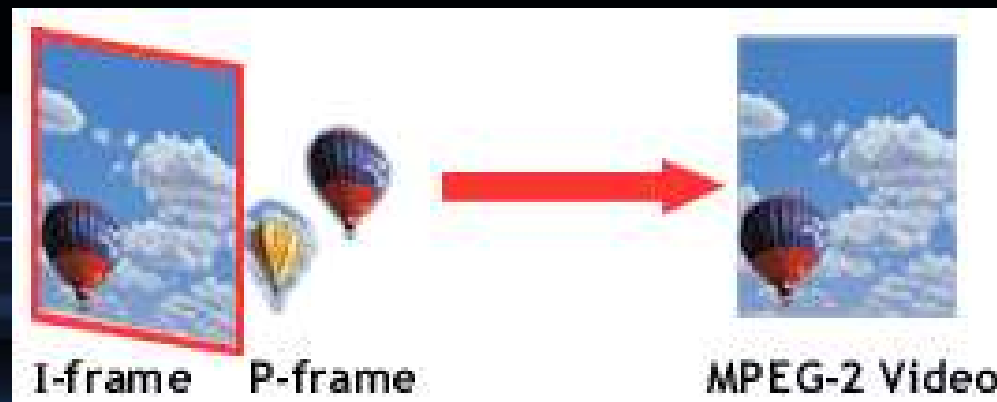
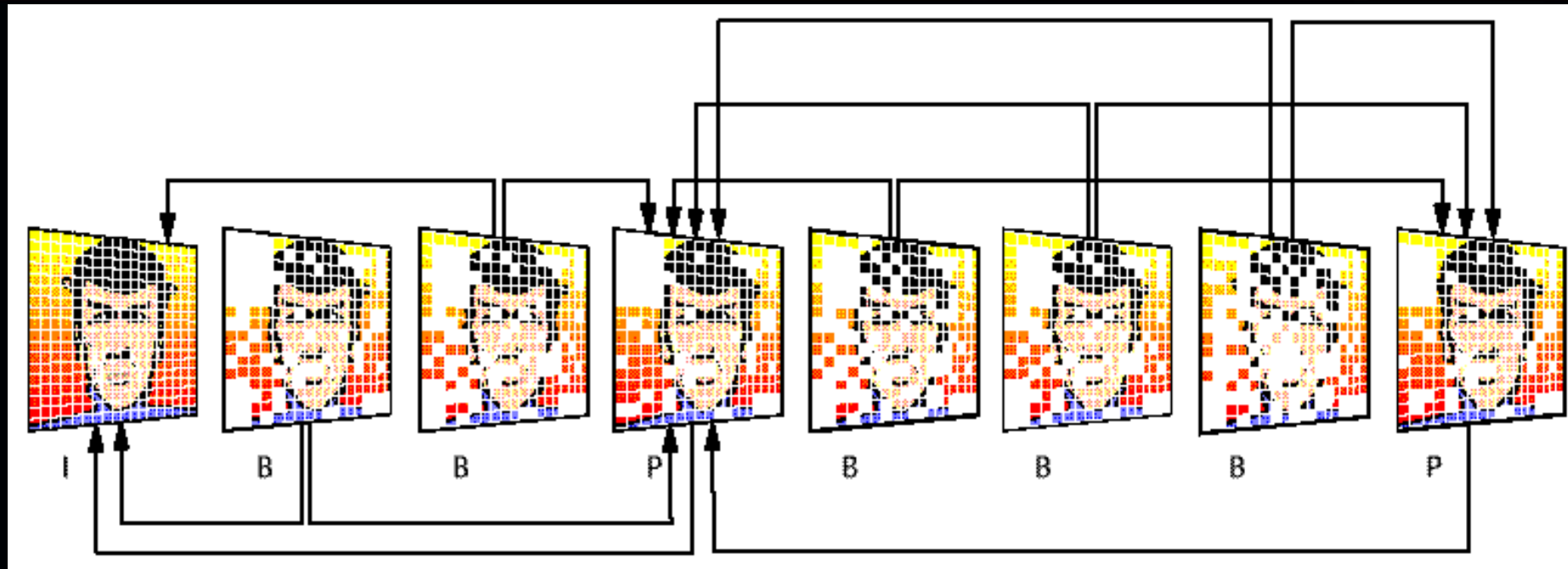
# Compactação MPEG-1 Video

## ■ Quadro P

- Codificam as diferenças entre os quadros
  - 50% do tamanho de um quadro I
- Se baseiam na idéia dos macrobloco de 16x16 pixels
  - macrobloco é codificado da seguinte forma:
    - tentando-se localizá-lo, ou algo parecido com ele, no quadro anterior
- Decodificar quadros P requer que o decodificador armazene o quadro I ou P anterior em um buffer
  - a partir do qual o novo quadro é construído baseado em macroblocos completamente codificados e macroblocos contendo diferenças com o quadro anterior

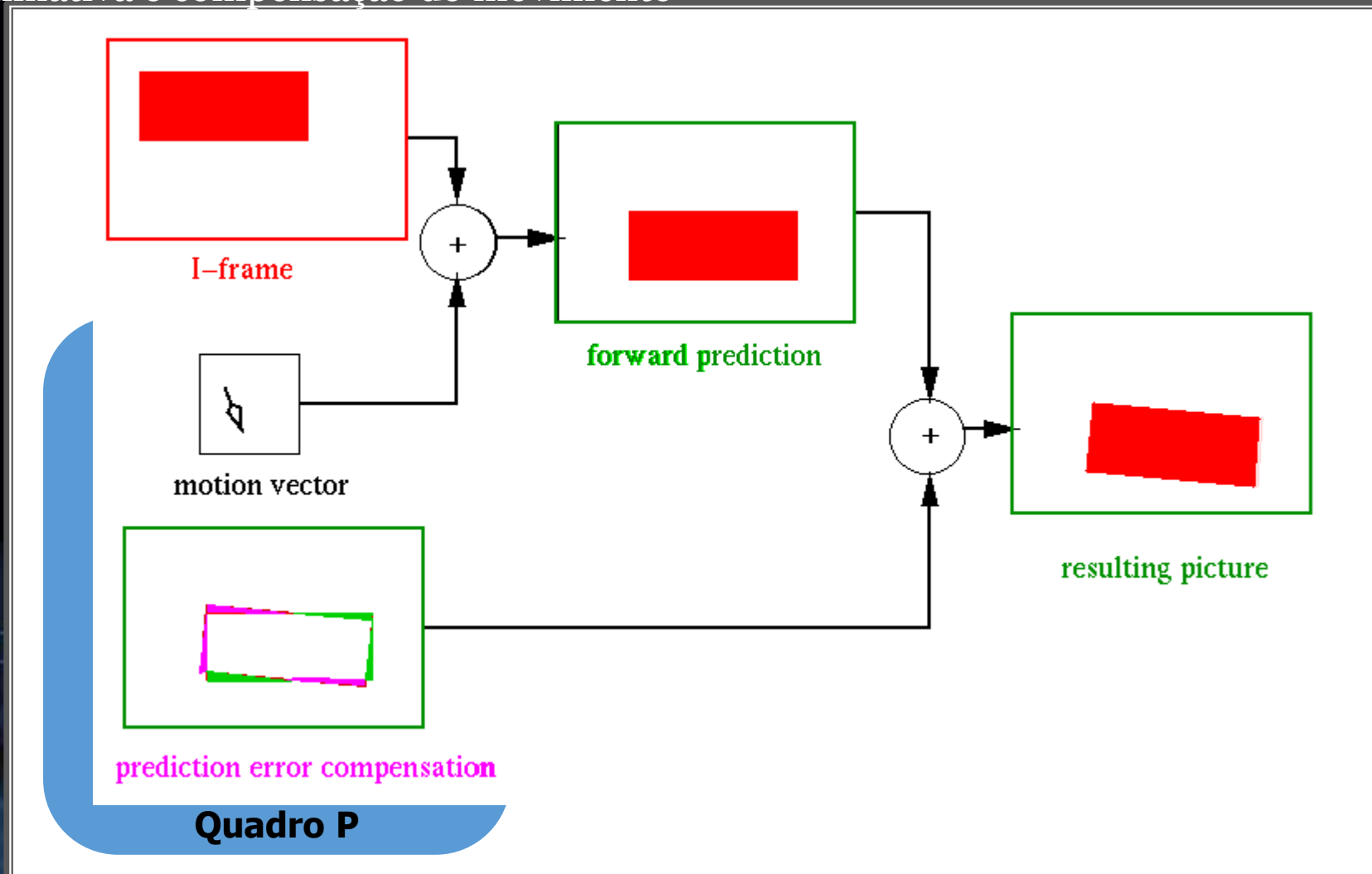


# Compactação MPEG-1 Video



# Compactação MPEG-1 Video

- **Quadro P**
  - Estimativa e compensação do movimento



# Compactação MPEG-1 Video

- **Quadro P**
  - Estimativa e compensação do movimento



I-frame



B-frame



B-frame



B-frame



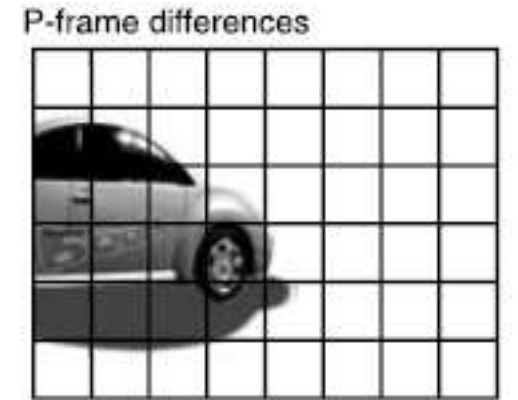
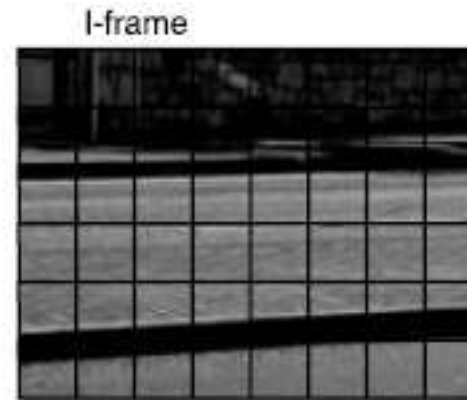
B-frame



B-frame



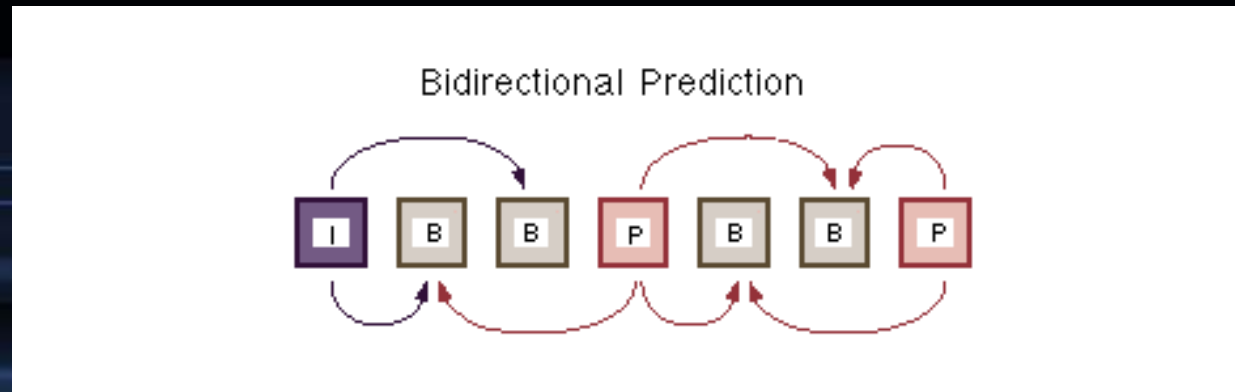
P-frame



# Compactação MPEG-1 Video

## ■ Quadro B

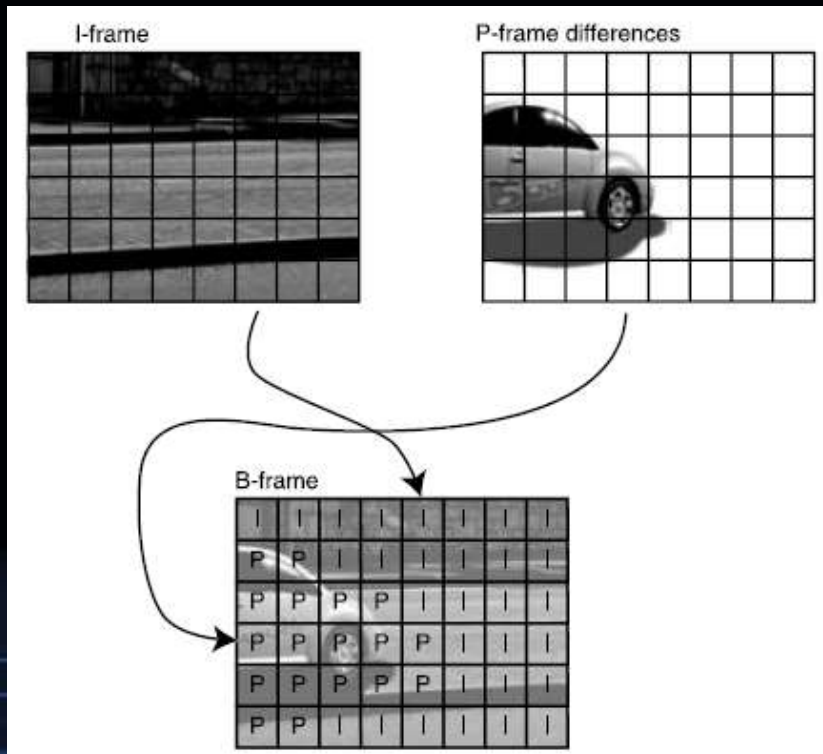
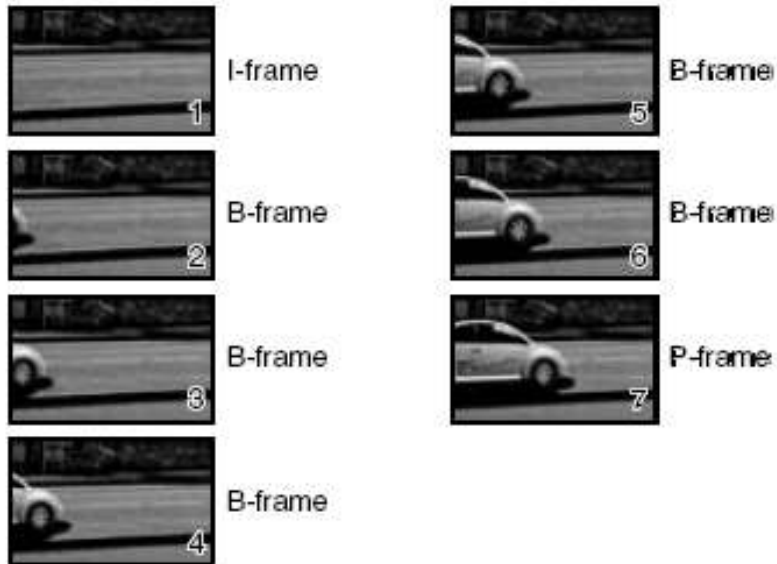
- Codificam as diferenças com o último quadro I ou P e com o quadro seguinte
  - 15% do tamanho de um quadro I
  - permitem que o macrobloco de referência esteja tanto no quadro anterior quanto no quadro seguinte
  - acarreta uma melhoria na compensação do movimento
- Para decodificar quadros B
  - decodificador precisa manter três quadros decodificados na memória ao mesmo tempo: o quadro anterior, o atual e o próximo





# Compactação MPEG-1 Video

- **Quadro B**
  - Estimativa e compensação do movimento



# Compactação MPEG-1 Video



- **Quadro D**

- Só são usado para possibilitar a apresentação de uma imagem de baixa resolução quando um avanço rápido ou um retrocesso

- **Um fluxo MPEG-1**

- Uma sequência de quadros codificados teria a seguinte forma:
  - IBBPBBPBBPBBIBBPBBPBBPB.....

- **Codificação MPEG-2**

- É fundamentalmente semelhante à codificação MPEG-1
  - com quadros I, P e B
  - quadros D não são aceitos
- Transformação discreta de co-seno é de 10x10 em vez de 8x8
  - para proporcionar mais 50 por cento de coeficientes
    - melhor qualidade

# MPEG-1 Sistemas



- **Define uma estrutura para:**
  - Combinar fluxos elementares, incluindo áudio, vídeo e outros fluxos de dados
    - chamado de Fluxo MPEG
    - até 32 fluxos de áudio MPEG e 16 fluxos de vídeo MPEG podem ser multiplexados juntamente com 2 fluxos de dados de diferentes tipos
- **Especifica o modo de representar as informações temporais necessárias para reprodução de sequências sincronizadas em tempo real**
  - sincronização de fluxos elementares
  - gerenciamento de buffer nos decodificadores
  - acesso aleatório
  - identificação do tempo absoluto do programa codificado

# Pontos Importantes

## MPEG Vídeo (1 e 2)

- Conhecimentos dos princípios da compressão: redundância espacial (similar a JPEG) e temporal (estimativa e compensação do movimento)



**UFSC**

**Ciências da  
Computação**

# **CAP 3. COMPRESSÃO DE DADOS MULTIMÍDIA**

**INE5431 Sistemas Multimídia**  
**Prof. Roberto Willrich (INE/UFSC)**  
**roberto.willrich@ufsc.br**

**Aula 10: Padrões de compressão  
multimídia – MPEG-4 e H.26\***



# Compressão de Dados Multimídia



- **Conteúdo:**

- Necessidade de compressão
- Princípios da compressão
- Classificação das técnicas de compressão
- Medição do desempenho de compressão
- Técnicas de compressão sem perdas
  - RLE, Huffman, LZW (GIF), Codificação Preditiva
- Técnicas de compressão de áudio, vídeo e imagens
- **Padrões de compressão multimídia**
  - JPEG, Codecs de Voz, MPEG Áudio, MPEG Vídeo, **MPEG-4, H.26\***

# MPEG-4



- **Padrão MPEG-4**

- Começou a ser concebido em julho de 1993, tendo sido aprovado como padrão internacional em 2000.
- MPEG-4 absorve muita das características do MPEG-1 e MPEG-2 e outros padrões relacionados, adicionando novas características

- **Uso**

- Vários vídeos transmitidos pela Internet fazem uso deste padrão, assim como telefones celulares que utilizam imagens
- Também é utilizado em diversos padrões de transmissão de TV digital, especialmente os de alta definição (HDTV)

# MPEG-4



## ■ Partes (padrões)

Parte	Data	Título	Descrição
<b>Part 1</b>	2010	Systems	Descreve sincronização e multiplexação de áudio e vídeo. Especifica o MPEG-TS (Transport Stream)
<b>Part 2</b>	2004	Visual	Compressão de vídeo.
<b>Part 3</b>	2009	Audio	Formatos de compressão de áudio: AAC, ALS, SLS.
<b>Part 4</b>	2004	Conformance testing	Procedimento para testes de conformidade
<b>Part 5</b>	2001	Reference software	Software de referência.
<b>Part 6</b>	2000	Delivery Multimedia Integration Framework (DMIF)	Interface entre a aplicação e o transporte
<b>Part 7</b>	2004	Optimized reference software for coding of audio-visual objects	Exemplos de como melhorar implementação.
<b>Part 8</b>	2004	Carriage of ISO/IEC 14496 contents over IP networks	Métodos para transportar conteúdo MPEG-4 em redes IP.
<b>Part 9</b>	2009	Reference hardware description	Provê projetos de hardware
<b>Part 10</b>	2012	Advanced Video Coding (AVC)	Um formato de compressão para vídeo (ITU-T H.264).

# MPEG-4

## ■ Partes (padrões)



Parte	Data	Título	Descrição
<b>Part 11</b>	2005	Scene description and application engine	BIFS, XMT, MPEG-J. Define posicionamento de objetos, representação de objetos sintéticos 2D e 3D, ...
<b>Part 12</b>	2012	ISO base media file format	Um formato de arquivo para armazenar conteúdo de mídia baseado em tempo
<b>Part 13</b>	2004	Intellectual Property Management and Protection (IPMP) Extensions	Gerenciamento de propriedade intelectual e proteção
<b>Part 14</b>	2003	MP4 file format	Formato de arquivo MPEG-4 versão 2
<b>Part 15</b>	2010	Advanced Video Coding (AVC) file format	Formato de arquivo MPEG-4 Parte 10
<b>Part 16</b>	2011	Animation Framework eXtension (AFX)	Especifica o modelo MPEG-4 AFX para representação de conteúdos gráficos 3D
<b>Part 17</b>	2006	Streaming text format	Formato de legenda
<b>Part 18</b>	2004	Font compression and streaming	
<b>Part 19</b>	2004	Synthesized texture stream	
<b>Part 20</b>	2008	Lightweight Application Scene Representation (LAsER) and Simple Aggregation Format (SAF)	Baseado no SVG Tiny. Para portais interativos, TV móvel, desenhos 2D, mapas gráficos interativos, etc.

# MPEG-4

## ■ Partes (padrões)

Parte	Data	Título	Descrição
<b>Part 21</b>	2006	MPEG-J Graphics Framework eXtensions (GFX)	Descreve um ambiente programático leve para aplicações multimídia interativas.
<b>Part 22</b>	2009	Open Font Format	
<b>Part 23</b>	2008	Symbolic Music Representation (SMR)	
<b>Part 24</b>	2008	Audio and systems interaction	
<b>Part 25</b>	2009	3D Graphics Compression Model	
<b>Part 26</b>	2010	Audio Conformance	
<b>Part 27</b>	2009	3D Graphics conformance	
<b>Part 28</b>	2012	Composite font representation	
<b>Part 29</b>	2014	Web video coding	
<b>Part 30</b>	2014	Timed text and other visual overlays in ISO base media file format	
<b>Part 31</b>	Em desenvol.	Video Coding for Browsers (VCB)	Codec para browsers
<b>Part 33</b>	Em desenvol.	Internet video coding	



# MPEG-4



## ■ MPEG-4 Parte 2

- Um padrão de compressão de vídeo DCT similar aos padrões MPEG-1 e MPEG-2
- 21 Perfis (Profiles)
  - Agrupam características em perfis (profiles) e níveis.
    - Para permitir seu uso em várias aplicações, variando de câmeras de segurança de baixa qualidade, baixa resolução a HDTVs e DVDs,
  - Perfil Simple Profile (SP): usado em situações onde a baixa taxa de bits e baixa resolução são mandatórios devido a largura de banda da rede, tamanho do dispositivo, etc
    - telefones celulares, sistemas de segurança, etc.
  - Perfil Advanced Simple Profile (ASP): muito similar ao H.263, incluindo suporte para a quantificação do estilo MPEG, suporte a vídeo entrelaçado, suporte a imagens do tipo B, compensação de movimento QPel (Quarter Pixel) e Global (GMC).

# MPEG-4



## ■ MPEG-4 Parte 2

### ▫ Perfil Simple Studio Profile (SStP)

- Tem 6 níveis indo de SDTV até a resolução 4K
- Permite até profundidade de píxel de 12-bits e subamostragem de croma 4:4:4

Nível	Max bit depth and chroma subsampling	Max resolution and frame rate	Max data rate (Mbit/s)
1	10-bit 4:2:2	SDTV (e.g. 704x480)	180
2	10-bit 4:2:2	1920×1080 30p/30i	600
3	12-bit 4:4:4	1920×1080 30p/30i	900
4	12-bit 4:4:4	2K×2K 30p	1.350
5	12-bit 4:4:4	4K×2K 30p	1.800
6	12-bit 4:4:4	4K×2K 60p	3.600

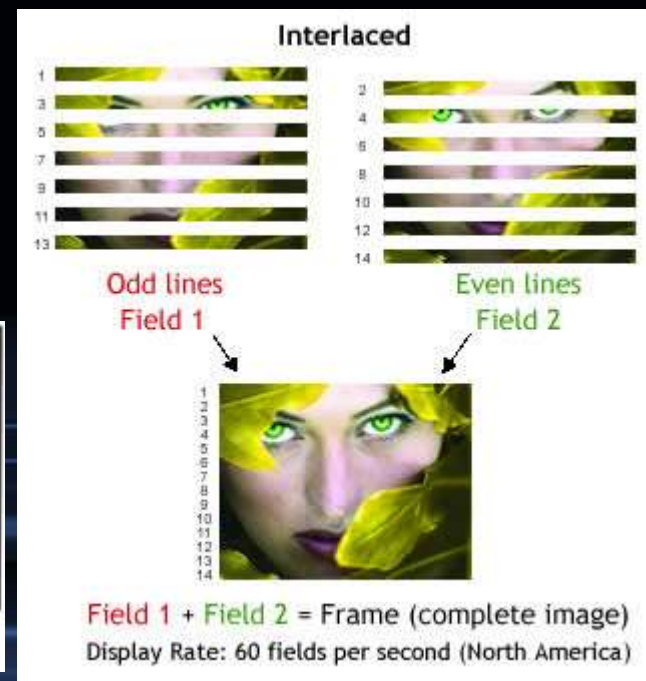
# MPEG-4

- **Varredura Progressiva (p)**

- “Varre” a tela inteira em uma única passada, transmitindo e exibindo todas as linhas da tela a cada atualização

- **Varredura entrelaçada (i)**

- Monta em cada passagem metade das linhas da tela, as linhas pares ou ímpares



# MPEG-4



## ■ MPEG-4 Parte 10

- Também conhecidos como H.264 ou AVC (Advanced Video Coding)
- Um padrão de codec de vídeo digital que tem a característica de alta taxa de compressão
- O padrão define 7 perfis, voltada a classes de aplicações específicas. Por exemplo:
  - Baseline Profile (BP) é voltado para aplicações de custo mais baixo com limitado recursos computacionais, usado em aplicações de videoconferência e móveis.
  - Extended Profile (XP) é voltado para streaming de vídeo, com alta taxa de compressão e robustez para perda de dados.
  - High Profile (HiP) é o principal perfil para aplicações de armazenamento em disco e broadcast, particularmente para aplicações de HDTV e adotado pelos discos HD-DVD e Blu-ray.



# MPEG-4 BIFS

- **BIFS – Binary Format for Scenes**
  - MPEG-4 é um sistema baseado em objetos.
  - BIFS permite a organização no tempo e espaço de vários tipos de mídia:
    - Uma descrição de cena compõe estes objetos
    - Descreve interatividade com objetos
    - Anima objetos

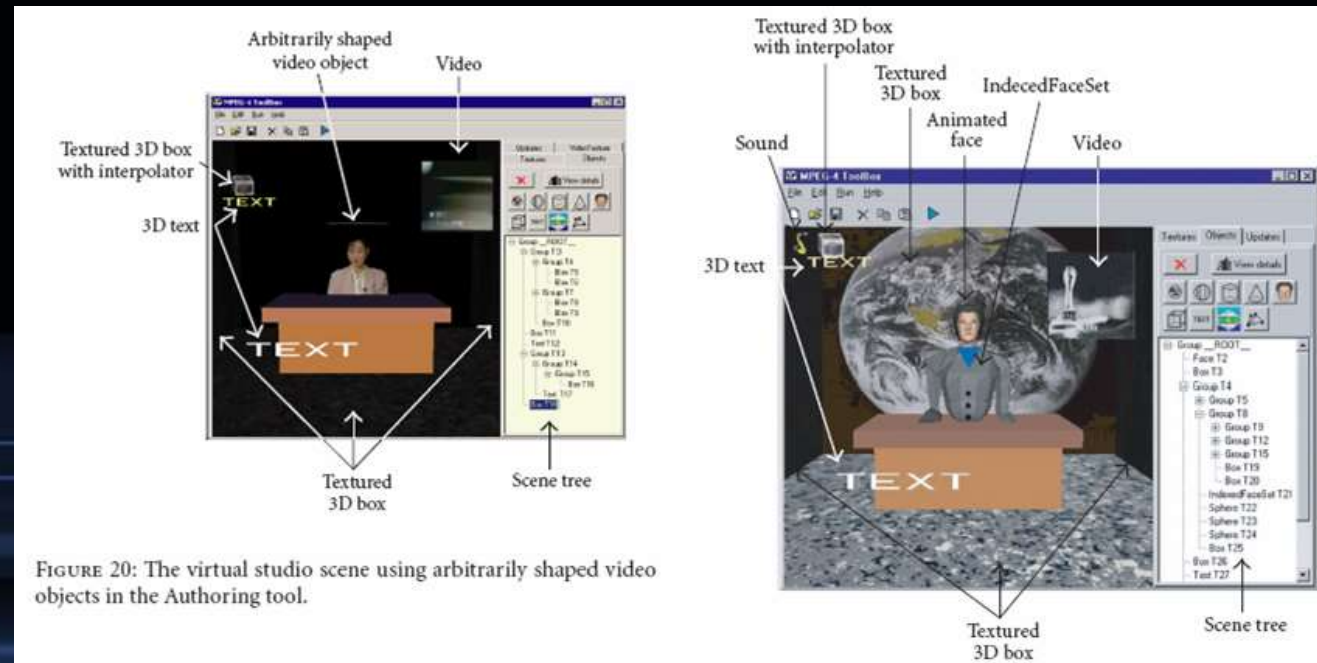


FIGURE 20: The virtual studio scene using arbitrarily shaped video objects in the Authoring tool.



# ITU-T H.261



## ■ Origem

- Necessidade de fornecer serviços de vídeo onipresentes na Rede Digital de Serviços Integrados (ISDN)
- ISDN: Tecnologia de transmissão digital de voz, vídeo e dados e outros serviços sobre a rede pública de telefonia comutada
  - Cada circuito garantindo 64 kbps sem variação de atrasos.
- Um dos padrões da família H.320 para videofonia e teleconferência na taxa de 64 Kbps a 2 Mbps

## ■ Padrão ITU-T

- Ratificado em novembro de 1988
- Influenciou o H.263, MPEG 1 a 4, etc.

# ITU-T H.261

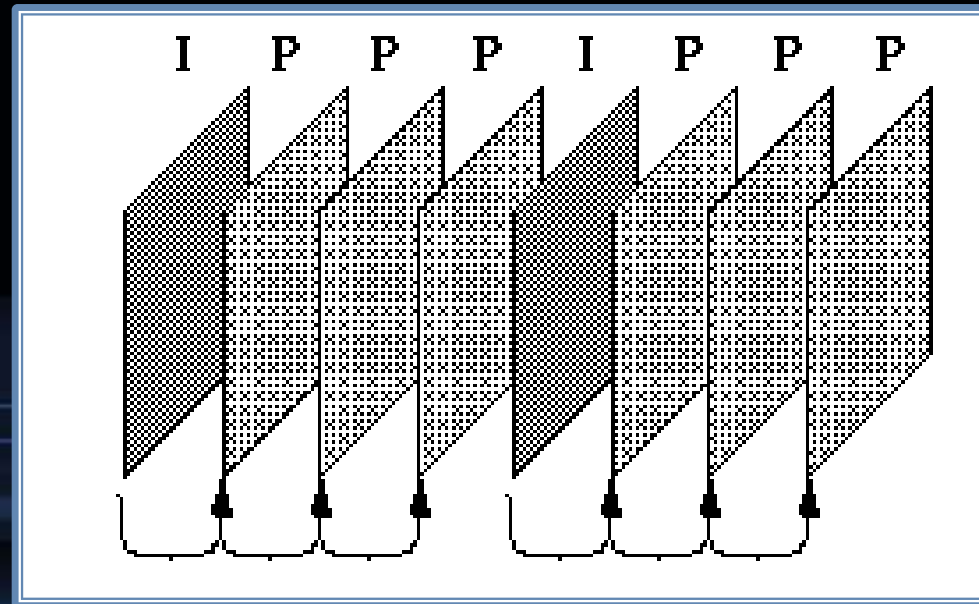


- **H.261 foi projetado para usar toda a capacidade do canal ISDN**
  - $p \times 64$  Kbps ( $p=1$  a  $30$ )
  - $p = 1$  ou  $2$  é apropriado para comunicação visual face-a-face e baixo movimento (videofonia)
  - $p > 5$  melhor qualidade (videoconferência)
  - Máxima taxa de bits disponível é 1,92 Mbps ( $p=30$ )
    - suficiente para obter imagens de qualidade VHS

- **Principais características**
  - Para aplicações de videofonia e teleconferência
  - Algoritmo de compressão de vídeo opera em tempo-real com atraso mínimo
  - Fornece uma resolução cerca de oito vezes mais baixa que a qualidade TV PAL/SECAM
  - É para aplicações usualmente sem movimentos intensos
    - algoritmo usa uma limitada estratégia de busca e estimação de movimento para obter taxas de compressão mais altas

# ITU-T H.261

- **Algoritmo de compressão**
  - Dois tipos de quadros: intraquadros (quadros I) e interquadros (quadros P)
    - I fornece um ponto de acesso e usa basicamente JPEG
    - P usa estimativa e compensação do movimento do quadro anterior

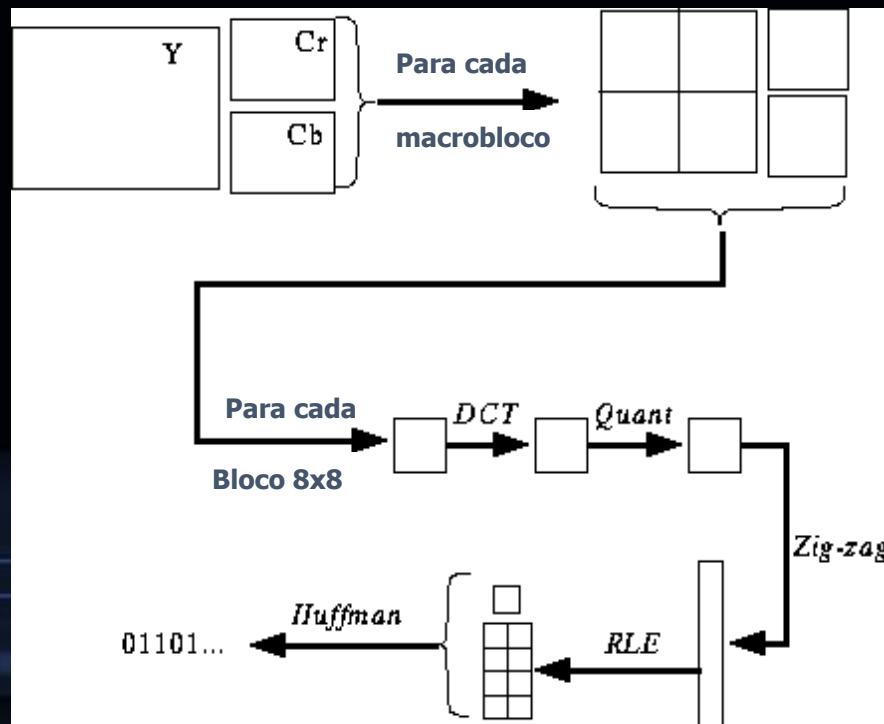


# ITU-T H.261

- **Algoritmo de compressão**

- ▣ Quadros I

- Usa o conceito de macrobloco: área de 16x16 píxeis no Y e 8x8 no Cb e Cr (4:2:0)

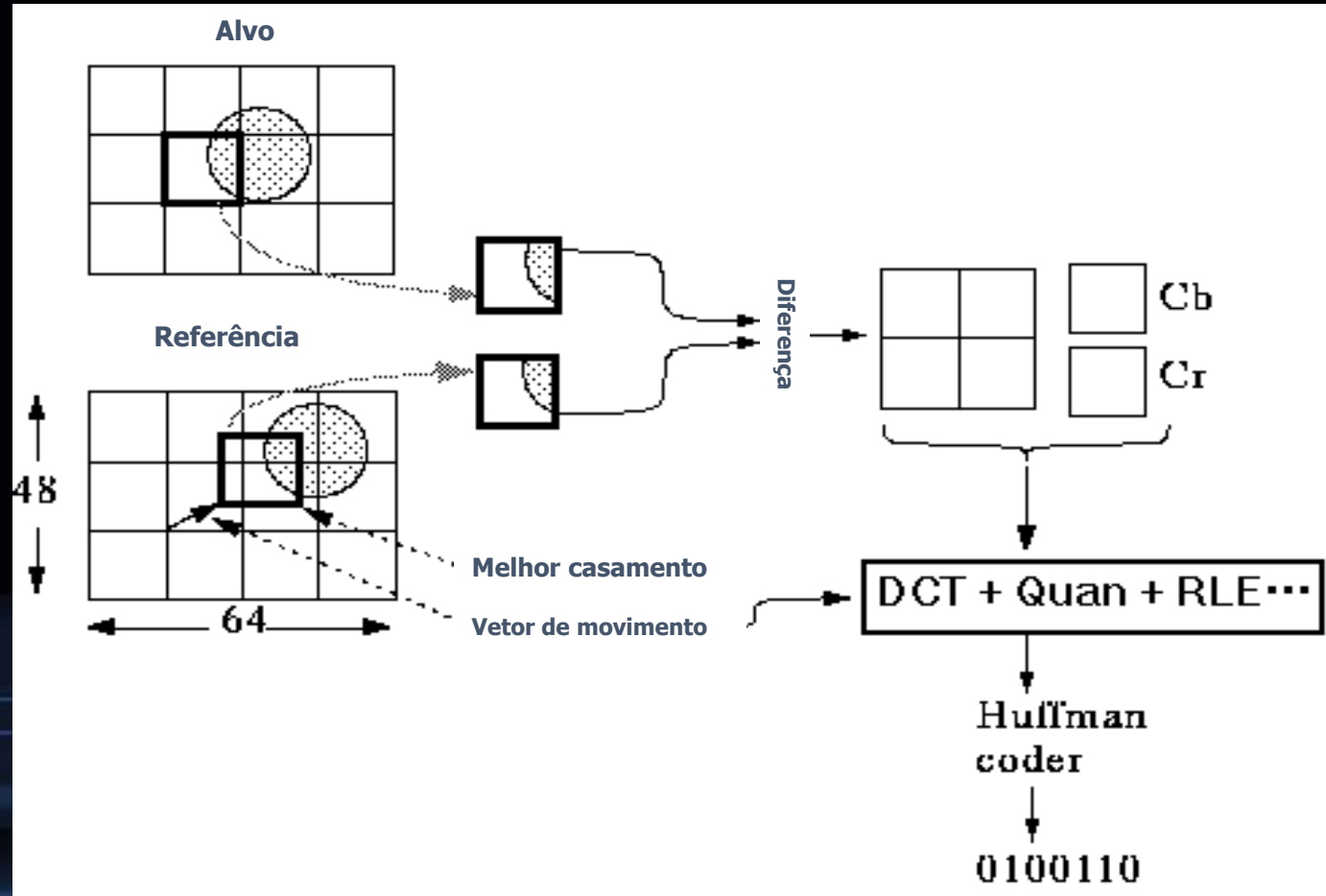




# ITU-T H.261

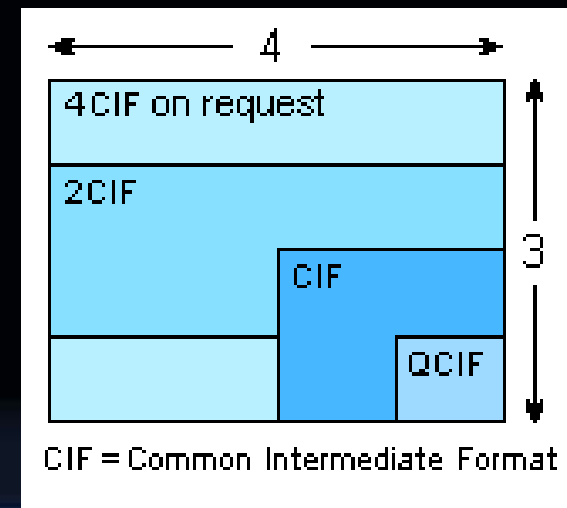
- ▣ Quadros P

- ▣ Usa o conceito de macrobloco: área de 16x16 pixels no Y e 8x8 no Cb e Cr (4:2:0)



# ITU-T H.261

- **Formatos de imagens**
  - H.261 opera com dois formatos de imagem
    - CIF (Common Intermediate Format) - 320x288
      - permite usar um formato único dentro e entre regiões usando padrões de TV de 625 e 525 linhas
    - QCIF (quarter-CIF) - 160x144
      - mais útil em taxas de bit menores ( $p < 6$ ).



# ITU-T H.26\*



- **H.262**
  - MPEG-2 pela ISO/IEC
- **H.263 (1995)**
  - Padrão de vídeo a baixa taxa de bits para aplicações de teleconferência que opera a taxas abaixo de 64 Kbps
- **H.263 v1 H.263 v2 (H.263+, 1997),**
- **H.263 v3 (H.263++, 2000), H.26L (2002)**
- **H.264/AVC (2003)**
- **H.265/HEVC (MPEG-H Part 2)**
  - 25% a 50% mais compressão que AVC na mesma qualidade