

Compressão

A compressão visa a redução do volume de dados especialmente para fins de transmissão e armazenamento.

As áreas de aplicação são variadas:

1. Imagens médicas (ressonância, tomografia);
2. Digitalização de documentos em alta resolução (documentos históricos, livros raros, mapas, fotografias, coleções de jornais, revistas, etc);
3. Digitalização de pinacotecas inteiras, coleções de filmes e documentários raros, etc...

Compressão

Supondo um vídeo RGB de qualidade HD, onde cada frame representa 1280x720 pixels de 8 bits, teríamos:

$$3 \times 1280 \times 720 = 2.764.800 \text{ bytes por frame}$$

Se o mesmo vídeo tiver 2 h de duração a uma taxa de 30 fps:

$$\text{Tempo total} = (2 \times 60) \text{ min} = (2 \times 60 \times 60) \text{ s} = 7200 \text{ s}$$

$$\text{Total de frames} = 7200 \times 30 = 216.000 \text{ frames}$$

$$\text{Total de memória: } 2.764.800 \times 216.000 \approx 556 \text{ Gbytes}$$

Quando pensamos em streaming com resoluções e taxas de fps maiores (full HD e 60fps, por exemplo), temos a intuição da compressão desses dados.

Compressão

O objetivo da compressão de imagens é reduzir o volume de dados;

Dados são representações de informações.
Informações são conceitos e ideias passadas ao Observador;

A mesma informação pode ser passada na forma de dados compactados;

Vejamos a seguir, como separar esses dois conceitos para compactar os dados mantendo o principal da informação...

Compressão

- 1) “A grande caixa retangular, a qual estava localizada no canto da linda mesa redonda, caiu rapidamente sobre o piso e rolou sobre este por alguns segundos”
- 2) “A caixa retangular sobre a mesa redonda caiu sobre o piso”

As frases acima são representações diferentes levando a mesma mensagem. A frase 1 usou $d1=28$ palavras enquanto a frase 2 usou $d2=11$

A taxa de compressão foi de $C=d1/d2=2,54$

$C=2,54$ é o ganho por usar a segunda frase para informar que a caixa caiu.

Compressão

A compressão de dados é baseada na quantidade de dados redundantes. Para as frases citadas ($d_1 > d_2$) a redundância relativa percentual R será:

$$R = \frac{d_1 - d_2}{d_1} * 100$$

$$R = \frac{C - 1}{C} * 100$$

60,7% das palavras da primeira frase são redundantes e podem ser eliminadas sem risco de perda da informação.

Compressão de Imagens

Nas imagens a representação da informação ocorre por meio de bits/bytes,

O objetivo da compressão de imagens é reduzir a quantidade de bytes necessários à representação das *features* (características visuais) presentes na imagem eliminando redundâncias de dados.

Compressão de Imagens

Existem 3 tipos de redundância:

1. **Redundância de Código**: desperdício de bits para representação de informação pouco frequente;
2. **Redundância Espacial**: *pixels* vizinhos são similares e podem ser representados com o mesmo valor;
3. **Redundância Visual**: a visão humana não necessita de muitos dados para interpretar a informação visual.

Redundância de Código

Usualmente, cada pixel de uma imagem é representada por uma quantidade constante de bits, em tons de cinza são 8 por pixel. Porém, nem todos os tons/cores “aparecem” nas imagens.

O ideal é utilizar uma representação (código) menor (menos bits) para tons mais frequentes e códigos maiores para tons menos frequentes.

Tabela 8.1 Exemplo de codificação de tamanho variável.

r_k	$p_r(r_k)$	Código 1	$l_1(r_k)$	Código 2	$l_2(r_k)$
$r_{87} = 87$	0,25	01010111	8	01	2
$r_{128} = 128$	0,47	10000000	8	1	1
$r_{186} = 186$	0,25	11000100	8	000	3
$r_{255} = 255$	0,03	11111111	8	001	3
r_k para $k \neq 87, 128, 186, 255$	0	—	8	—	0

Código de Huffman

Graylevels	n_i	h_i
0	3441	0.21
1	4423	0.27
2	3932	0.24
3	1802	0.11
4	1311	0.08
5	819	0.05
6	492	0.03
7	164	0.01
n_t	16384	

Código de Huffman

Graylevel	Probability	3	4	5	6	7	8
g_1	0.27	0.27	0.27	0.27	0.28	0.45	0.55
g_2	0.24	0.24	0.24	0.24	0.27	0.28	0.45
g_0	0.21	0.21	0.21	0.21	0.24	0.27	
g_3	0.11	0.11	0.11	0.17	0.21		
g_4	0.08	0.08	0.08	0.11			
g_5	0.05	0.05	0.09				
g_6	0.03	0.04					
g_7	0.01						

Graylevel	Probability	3	4	5	6	7	8
g_1	0.27 01	0.27 01	0.27 01	0.27 01	0.28 00	0.45 1	0.55 0
g_2	0.24 10	0.24 10	0.24 10	0.24 10	0.27 01	0.28 00	0.45 1
g_0	0.21 11	0.21 11	0.21 11	0.21 11	0.24 10	0.27 01	
g_3	0.11 001	0.11 001	0.11 001	0.17 000	0.21 11		
g_4	0.08 0000	0.08 0000	0.08 0000	0.11 001			
g_5	0.05 00010	0.05 00010	0.09 0001				
g_6	0.03 000110	0.04 00011					
g_7	0.01 000111						

Redundância Espacial

Refere-se ao alto grau de similaridade entre *pixels* vizinhos.

A intensidade da redundância espacial (calculada via função de autocorrelação) depende do nível de detalhes na imagem:

A) imagens de alta redundância de *pixels*: grandes regiões contendo muita similaridade de *pixels* (baixa frequência espacial, larga resposta da função de autocorrelação);

B) imagens com uma menor redundância relativa de *pixels*: poucas regiões contendo muita similaridade de *pixels* (alta frequência espacial, resposta estreita para a função de autocorrelação).



(A)



(B)

Redundância Espacial

Exemplo:

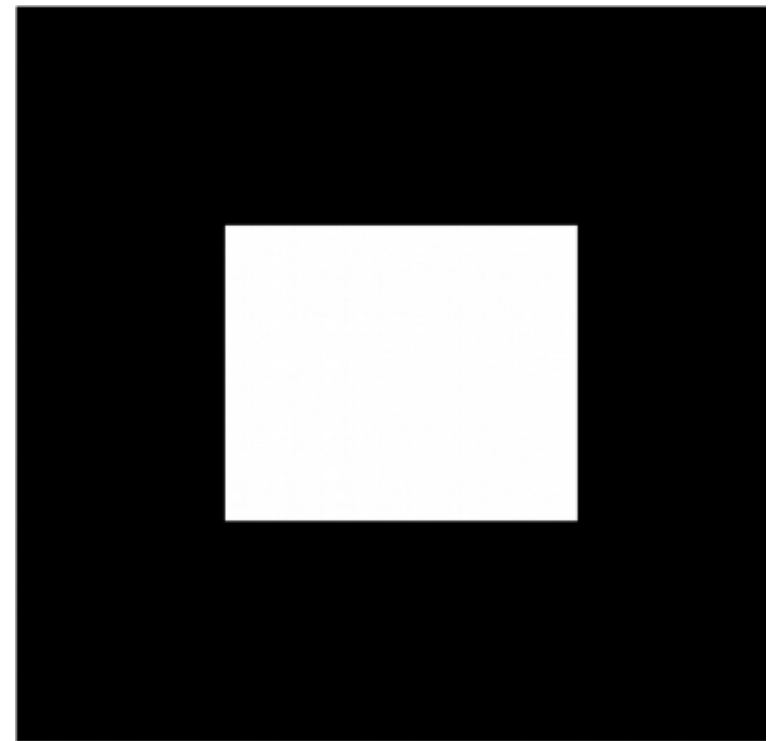
A técnica consiste em descrever a imagem utilizando sequencias
(G,#)=(tom de cinza, quantas vezes ocorre esse tom de cinza)

A imagem binária A com dimensões 128x128 possui um quadrado interno de

dimensões 64x64;

Se a profundidade de pixel for de 1 bit por pixel:

total= $128 \times 128 \times 1 = 16.384$ bits;

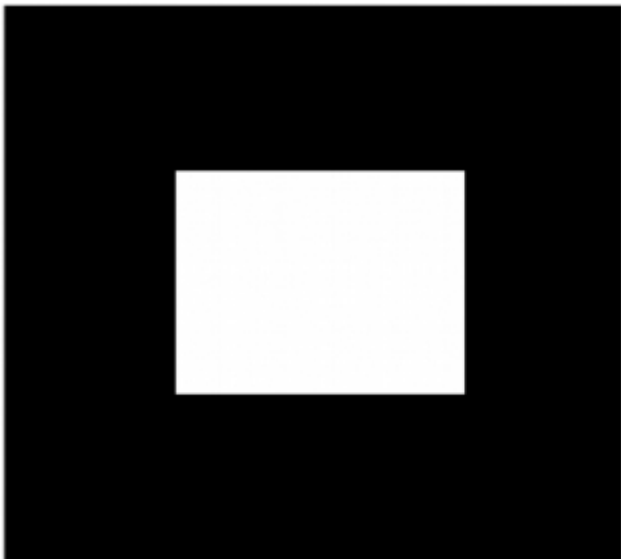


Redundância Espacial

Há apenas dois níveis Gb e Gw.

Utilizando um esquema de representação que trabalhe sobre a alta redundância espacial:

- A representação (G,#) demanda 8 bits:
 - 1 bit para a intensidade e
 - 7 bits para a quantidade de ocorrências;
- Tamanho da compressão = 32 linhas de (Gb,128) + 64 linhas de ((Gb,32)(Gw,64)(Gb,32)) + 32 linhas de (Gb,128) = $32 \cdot 8 + 64 \cdot (3 \cdot 8) + 32 \cdot 8 = 2048$ bits
- Taxa de compressão = $16348/2048 = 8$ para 1



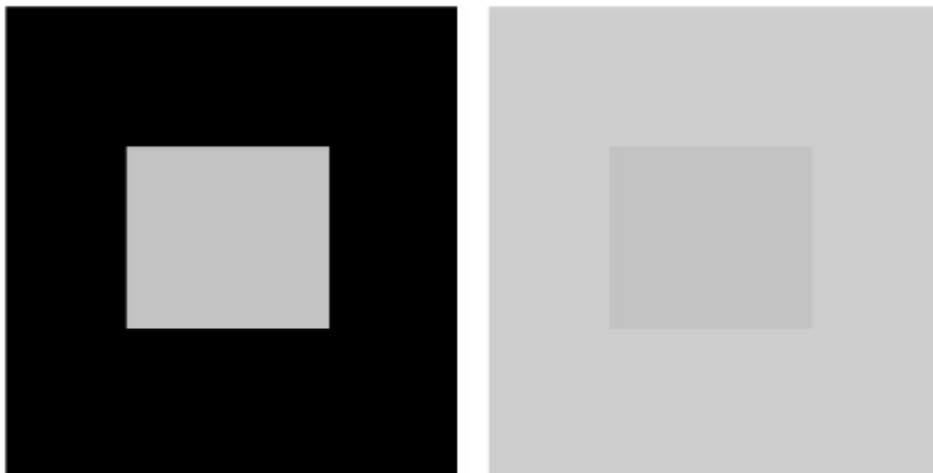
```
0000000...0000000...0000000...0000000...000
0000000...0000000...0000000...0000000...000
0000000...0000000...0000000...0000000...000
0000000...0000000...0000000...0000000...000
...
0000000...00011111...11110000000...0000000
0000000...00011111...11110000000...0000000
0000000...00011111...11110000000...0000000
0000000...00011111...11110000000...0000000
0000000...00011111...11110000000...0000000
...
0000000...0000000...0000000...0000000...000
0000000...0000000...0000000...0000000...000
0000000...0000000...0000000...0000000...000
0000000...0000000...0000000...0000000...000
```

```
(Gb,128)
(Gb,128)
...
(Gb,128)
...
(Gb,128)
(Gb,32)(Gw,64)(Gb,32)
...
(Gb,32)(Gw,64)(Gb,32)
...
(Gb,32)(Gw,64)(Gb,32)
(Gb,32)(Gw,64)(Gb,32)
...
(Gb,128)
(Gb,128)
...
(Gb,128)
```

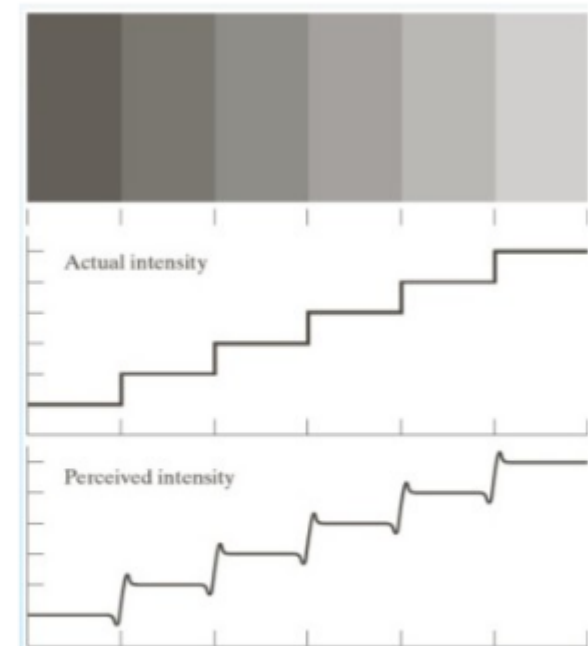

Redundância Visual

O sistema de visão humano apresenta certas peculiaridades:

- Nosso sistema é pobre em determinar o brilho absoluto de objetos na imagem;
- Para o olho humano a informação de brilho depende do tamanho e intensidade relativa dos objetos adjacentes na imagem (vizinhança).



Distorção de tamanho devido ao brilho relativo



Mach Band: distorção na percepção de Brilho na vizinhança de arestas.

Redundância Visual

Na verdade nem todos os dados da imagem são utilizados na interpretação visual humana;

Nosso sistema elimina “redundâncias visuais” fazendo atalhos para obter a representação;

O olho humano é sensível principalmente a arestas e cantos e insensível às variações de intensidade ao redor dessas regiões;

A compressão baseada em redundância visual produz uma nova imagem que é uma versão relativamente degradada da original:

- Diferentemente dos dois casos anteriores, esta classe de técnicas produz uma compressão com perdas de informações (a imagem descomprimida não será uma cópia da original).

Redundância Visual

Quantização não uniforme: imagens de 147x240

A) 8 bpp (≈ 34 kB) \rightarrow reduzida a \rightarrow B) 4 bpp (≈ 17 kB)

8 bpp = $[p_7 \ p_6 \ p_5 \ p_4 \ p_3 \ p_2 \ p_1 \ p_0] = 256$ níveis

4 bpp = $[p_7 \ p_6 \ p_5 \ p_4] = 16$ níveis



(A)



(B)

Redundância Visual

Duas técnicas simples:

Amostragem não uniforme:

- As arestas são visualmente importantes, então utiliza-se uma amostragem espacial mais alta em regiões de arestas e uma amostragem menor em regiões suaves da imagem. Com isso a resolução espacial será maior nas arestas.

Quantização não uniforme:

- O sistema visual humano não é muito sensível a tons de cinza próximos às arestas. Ao invés de quantizar uniformemente a imagem, utiliza-se uma quantidade menor de tons de cinza (bits por pixel) nas regiões de arestas (3 ao invés de 8 bits, por exemplo);

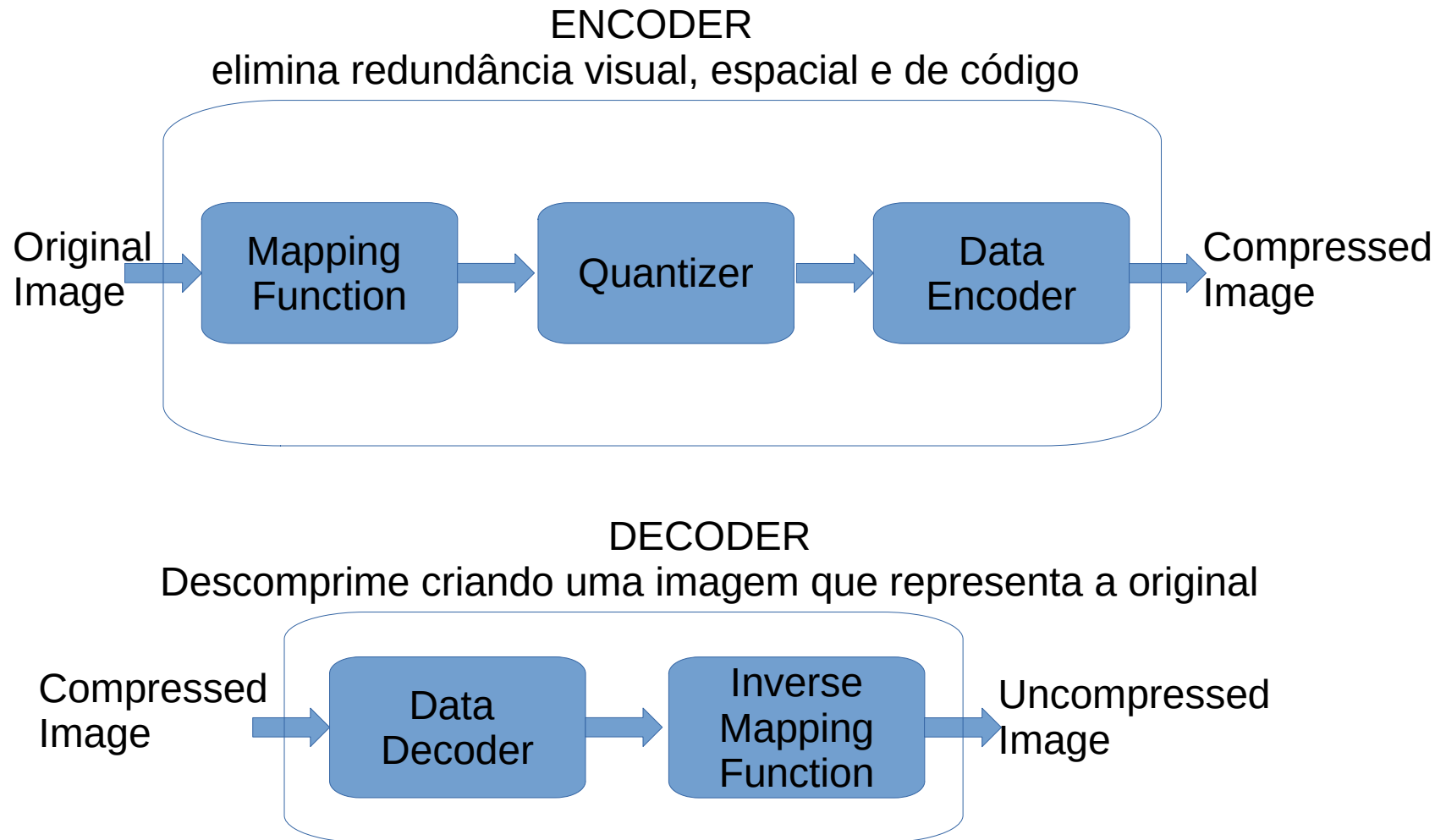
Métricas

1. Redundância de Código: sem perdas;
2. Redundância Espacial: sem perdas;
3. Redundância Visual: com perdas.

Existem medidas para aferir a fidelidade da imagem comprimida em relação a original:

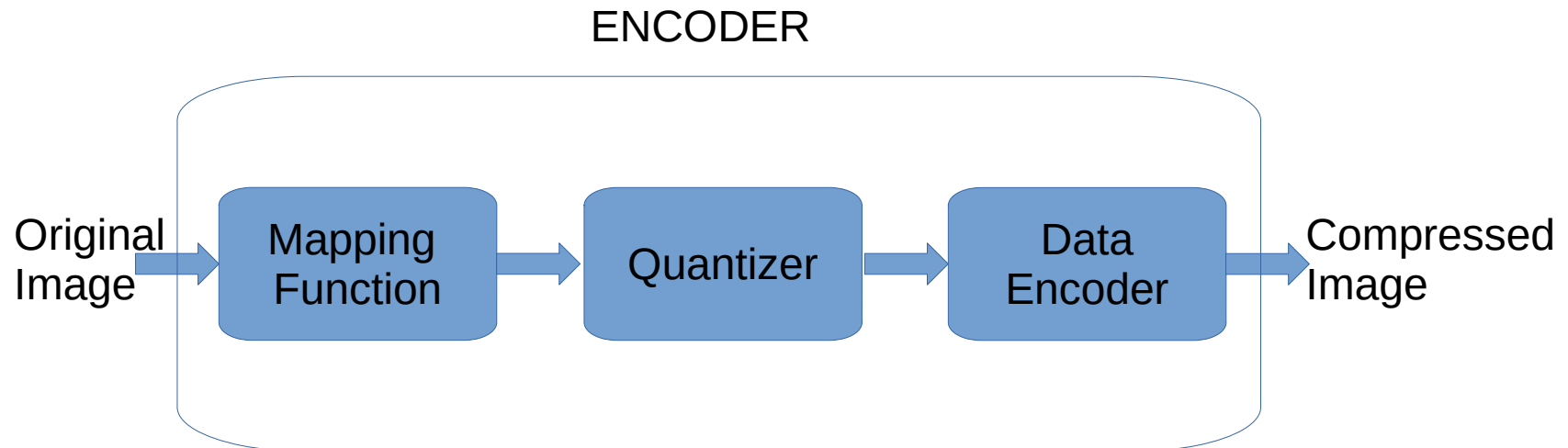
- MSE (Mean Square Error), quanto menor o MSE mais próximas serão as imagens original e comprimida;
- SNR (Signal to Noise Error), quanto maior o SNR mais próximas serão as imagens original e comprimida;
- PSNR (Peak to Peak Signal to Noise Error), etc.

Compressão de Imagens



Um método de compressão sem perdas não apresenta o estágio Quantizer;

Compressão de Imagens

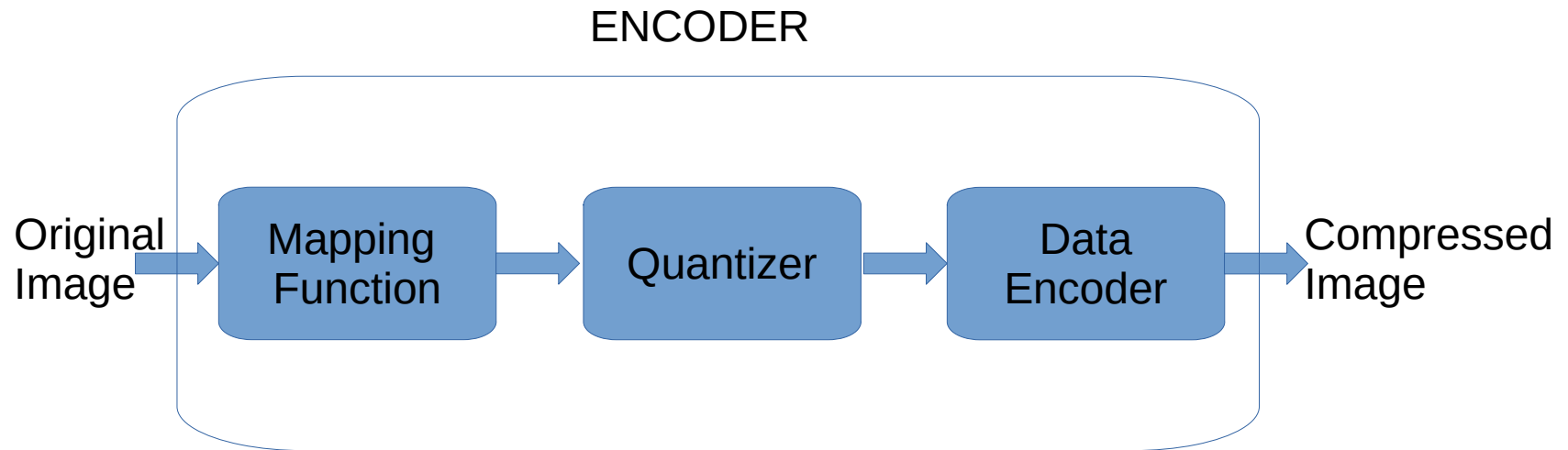


Mapping function: transforma a imagem para um novo domínio mais apropriado à eliminação de redundância espacial;

Quantizer: reduz a quantidade de valores que são gerados pela Mapping Function digitalizando esses valores para um subconjunto menor, é nesse estágio que a redundância visual da imagem original é reduzida;

Data Encoder: constrói uma representação comprimida aplicando os códigos de tamanho variável ao resultado do quantizer (reduz a redundância de código);

Compressão de Imagens

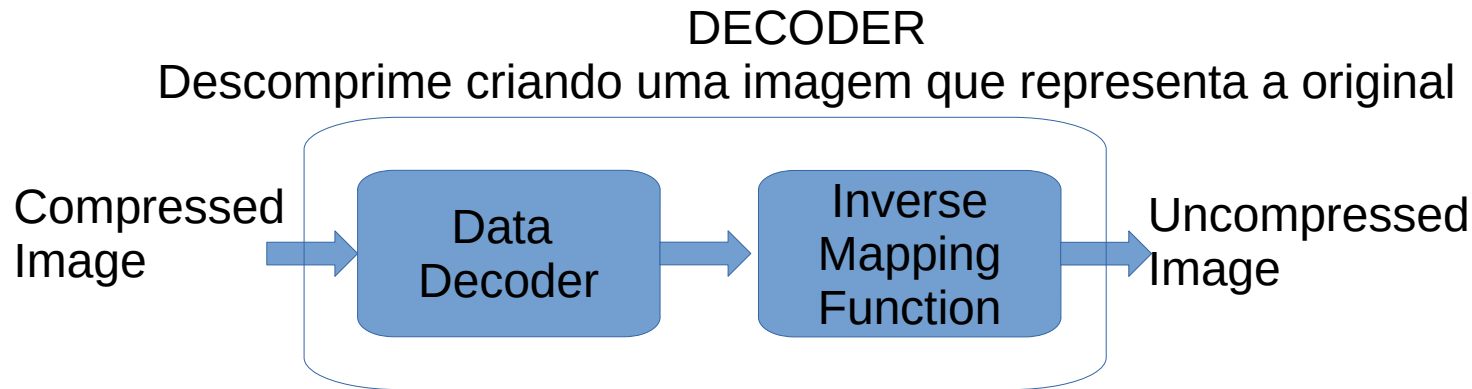


Os estágios Mapping Function e Data Encoder são totalmente reversíveis, eles não removem informação da imagem original;

O Quantizer remove informação da imagem resultando em uma descompressão que exibe uma versão degradada da imagem Original;

Um método de compressão sem perdas não apresenta o estágio Quantizer;

Compressão de Imagens



Data Decoder: reverte a operação do data encoder, produzindo valores iguais a saída do quantizer ou a saída do estágio mapping function a depender do método utilizado ter sido com ou sem perdas;

Inverse Mapping Function: produz a imagem descomprimida.

Compressão de Imagens

GIF	CompuServe	<i>Graphic Interchange Format.</i> A file format that uses lossless LZW coding [8.2.4] for 1- through 8-bit images. It is frequently used to make small animations and short low resolution films for the World Wide Web.
PDF	Adobe Systems	<i>Portable Document Format.</i> A format for representing 2-D documents in a device and resolution independent way. It can function as a container for JPEG, JPEG 2000, CCITT, and other compressed images. Some PDF versions have become ISO standards.
PNG	World Wide Web Consortium (W3C)	<i>Portable Network Graphics.</i> A file format that losslessly compresses full color images with transparency (up to 48 bits/pixel) by coding the difference between each pixel's value and a predicted value based on past pixels [8.2.9].
TIFF	Aldus	<i>Tagged Image File Format.</i> A flexible file format supporting a variety of image compression standards, including JPEG, JPEG-LS, JPEG-2000, JBIG2, and others.
JPEG	ISO/IEC/ ITU-T	A <i>Joint Photographic Experts Group</i> standard for images of photographic quality. Its lossy <i>baseline coding system</i> (most commonly implemented) uses quantized discrete cosine transforms (DCT) on 8×8 image blocks [8.2.8], Huffman [8.2.1], and run-length [8.2.5] coding. It is one of the most popular methods for compressing images on the Internet.

Compressão de Imagens

Bibliografia utilizada:

Gonzalez, R. e Woods, R. Digital Image Processing, 3rd edition.

Weeks, A. Fundamentals of Eletronic Image Processing, 2ª edição.