

# Fundamentos de Imagem Digital

Jefersson Alex dos Santos

jefersson@dcc.ufmg.br



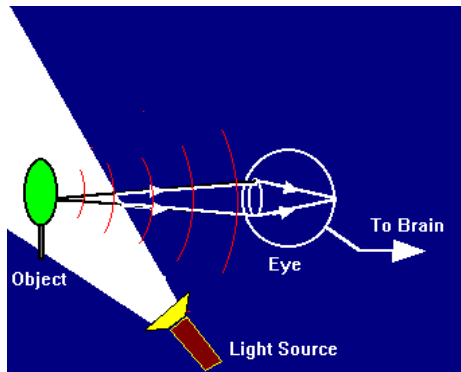
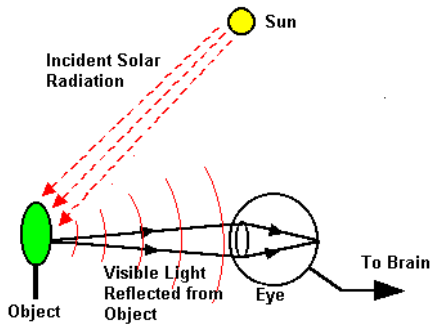
# Roteiro da Aula

- 1** Formação de Imagem
- 2** Espaços de Cor
- 3** Imagens multi-dimensionais e/ou multi-paramétricas
- 4** Formatos de armazenamento de imagem
- 5** Histograma de Cor

# Roteiro da Aula

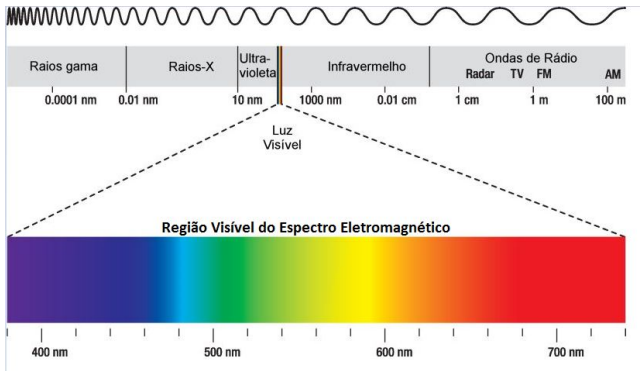
- 1** Formação de Imagem
- 2 Espaços de Cor
- 3 Imagens multi-dimensionais e/ou multi-paramétricas
- 4 Formatos de armazenamento de imagem
- 5 Histograma de Cor

# Visão Humana



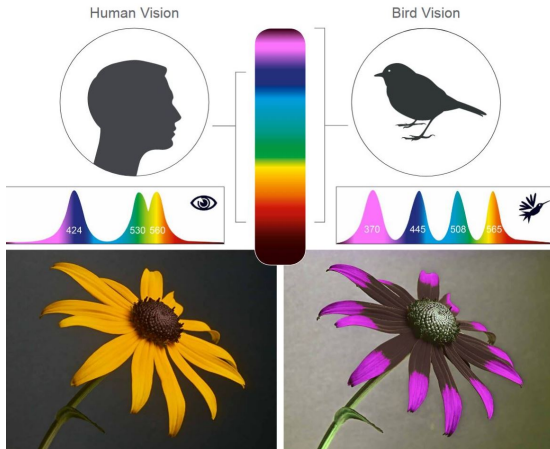
Leitura recomendada: Capítulo 2 do Gonzalez & Woods.

# Visão Humana

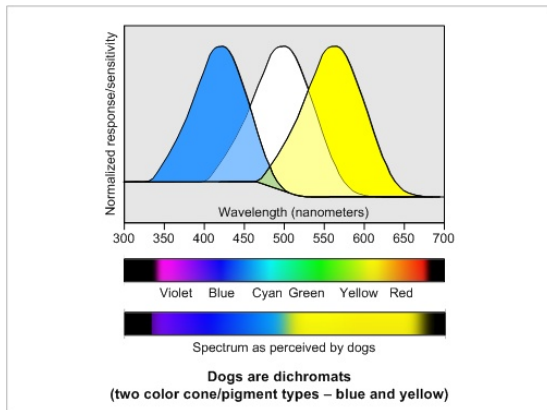


Escala de níveis de intensidade luminosa do sistema visual humano:  $10^{10}$

# Visão Humana



# Visão Humana

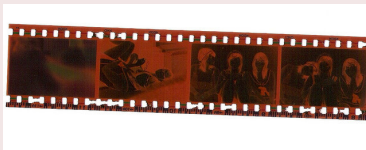


# Sinais Analógicos

Tudo que vem da natureza: sons, imagens, tato, cheiro, etc

## Imagens analógicas

- Fotografias
- Imagens médicas (chapa)
- Televisão



Processamento da imagem é realizado química, fisicamente ou por sinais elétricos!

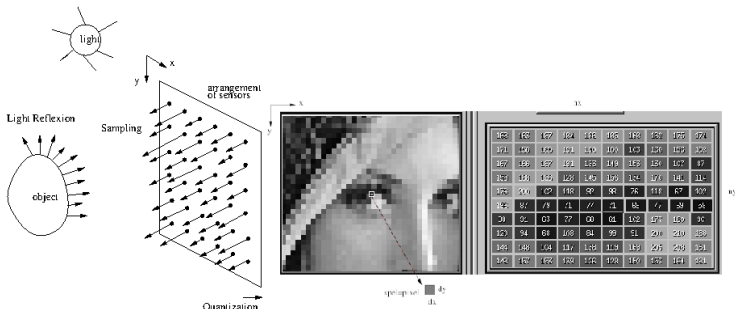
Digital x Analógico: quais as vantagens/desvantagens?



# Imagem Digital Monocromática

## Formação da Imagem

- Amostragem: sensores medem a reflectância da superfícies dos objetos
- Quantização: medidas são convertidas em números inteiros



# Amostragem e Quantização

- Se  $f(x, y)$  representa uma função contínua da reflexão de luz no espaço  $\mathbb{R}^2$ , a **amostragem** gera a função discreta  $I(x, y)$ :

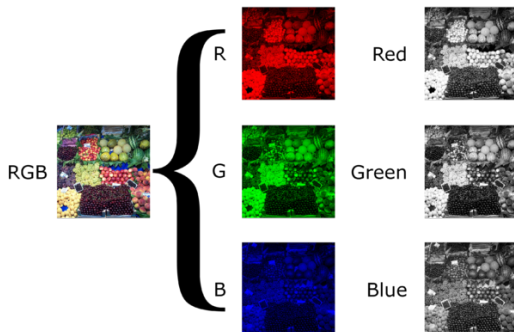
$$I(x, y) = f(x, y) \sum_{j=0}^{n_y-1} \sum_{i=0}^{n_x-1} \delta(x - id_x, y - jd_y)$$

onde  $\delta(x, y)$  é a função delta de Dirac discreta, definida para cada pixel  $p = (x, y)$ ,  $x = 0, \dots, n_x - 1$  e  $y = 0, \dots, n_y - 1$ . As distâncias  $d_x$  e  $d_y$  (normalmente  $d_x = d_y$ ) entre as amostras definem as **dimensões do pixel** (*picture element*).

- A **quantização** mapeia os valores reais dos pixels em números inteiros  $I(p)$  de 0 a  $2^b - 1$ , onde  $b$  é o número de bits necessário para representar cada pixel (**profundidade da imagem**).

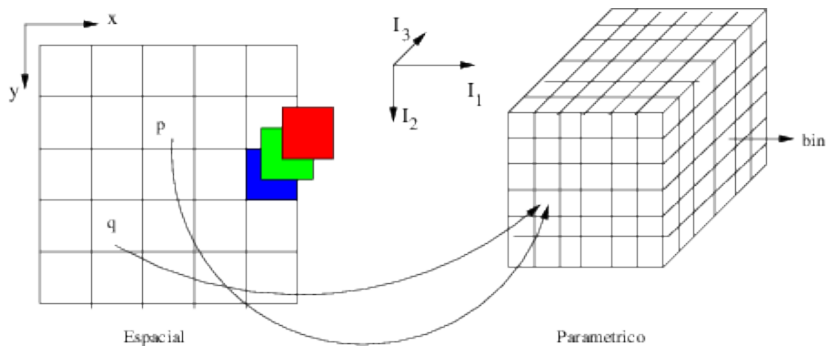
# Imagens Digitais Coloridas

Cada pixel  $p$  terá associado um vetor  $\vec{l}(p) = (l_1(p), l_2(p), l_3(p))$  com as medidas de reflexão de luz nos comprimentos de onda do vermelho, verde, e azul, respectivamente.



# Coordenadas espaciais e paramétricas

Os três componentes de cor de um pixel podem ser vistos como características (atributos/parâmetros), mapeando cada pixel em um ponto do espaço paramétrico correspondente.



\* Para  $b=24$ bits (8 bits por componente de cor), o número de *bins* (cores) é  $2^{24} - 1$ .

# Roteiro da Aula

1 Formação de Imagem

2 Espaços de Cor

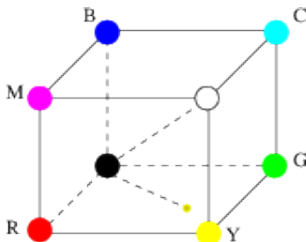
3 Imagens multi-dimensionais e/ou multi-paramétricas

4 Formatos de armazenamento de imagem

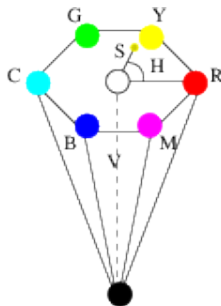
5 Histograma de Cor

# Espaços de cor

- Uma imagem colorida pode ser armazenada em diferentes espaços de cor (RGB, CMYK, HSV, Lab, etc.).
- A conversão da imagem de um espaço de cor para outro é uma operação matemática pixel a pixel (e.g., multiplicação matricial envolvendo  $\vec{I}(p)$ ).



RGB and CMY

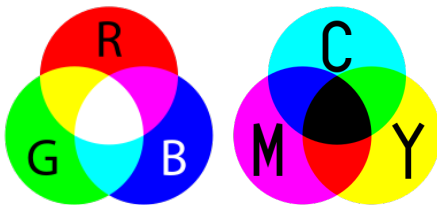


HSV

# Espaços de cor

## CMYK

- Complementar ao RGB
- Cores subtrativas, formado por: Cyan, Magenta, Yellow e Black (Key)
- Empregado por impressoras e fotocopiadoras

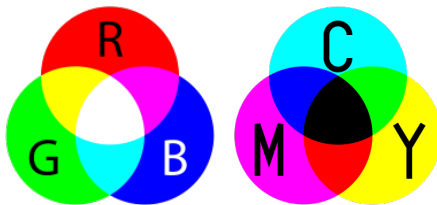


Como seria a conversão de um pixel  $p$  do espaço RGB para CMYK para uma imagem de 8-bits?

# Espaços de cor

## CMYK

- Complementar ao RGB
- Cores subtrativas, formado por: Cyan, Magenta, Yellow e Black (Key)
- Empregado por impressoras e fotocopiadoras



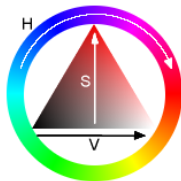
Como seria a conversão de um pixel  $p$  do espaço RGB para CMYK para uma imagem de 8-bits?



# Espaços de cor

## HSV ou HSI

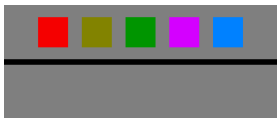
- Representa a matiz/tonalidade (hue), a saturação (saturation) e o brilho (value)
- Vantagem: componentes descorrelacionados. Mudanças de brilho não alteram a matiz



# Espaços de cor

## CIELAB – $L^*a^*b$

- É considerado como sempre uniforme para pequenas distâncias de cor.
- É projetado para aproximar à visão humana (ao contrário de RGB e CMYK)
- Evita o efeito *Helmholtz-Kohlrausch*:



Conversão de RGB para passa por um espaço intermediário, denominado XYZ. Existem várias propostas, mais detalhes sobre conversão em <http://www.bruce.lindbloom.com/>.

# Roteiro da Aula

**1** Formação de Imagem

**2** Espaços de Cor

**3** Imagens multi-dimensionais e/ou multi-paramétricas

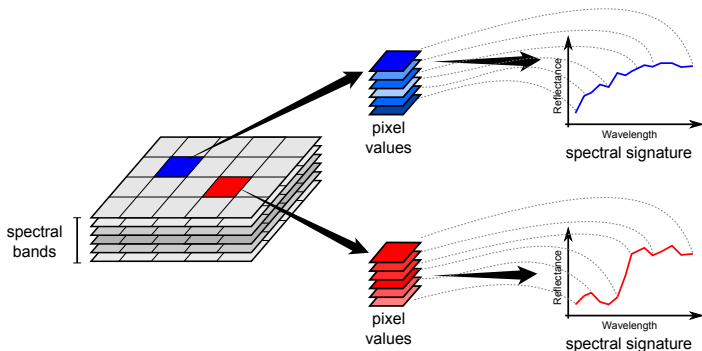
**4** Formatos de armazenamento de imagem

**5** Histograma de Cor

# Formação de imagem multi-banda

## Sensoriamento Remoto

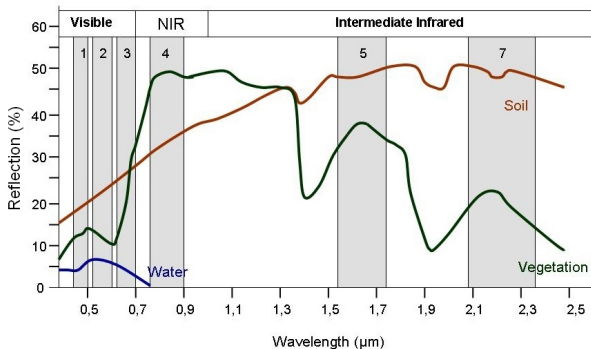
Uma imagem de satélite pode conter para cada pixel  $p$  amostras da radiação refletida na superfície da Terra em vários comprimentos de onda (**imagem multi-bandas**)



# Formação de imagem multi-banda

## Sensoriamento Remoto

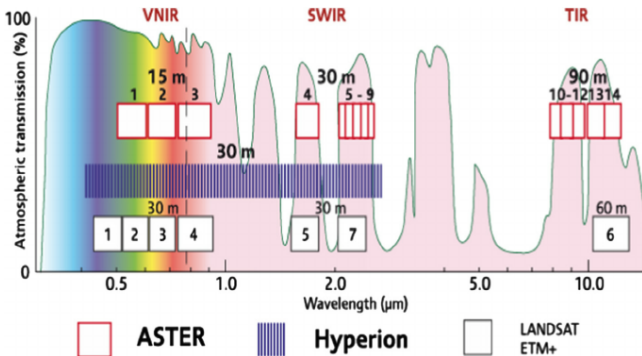
O satélite Landsat 5 (sensor *Thematic Mapper*), por exemplo, gera um vetor  $\vec{l}(p)$  com 7 valores de reflectância.



# Formação de imagem multi-banda

## Sensoriamento Remoto

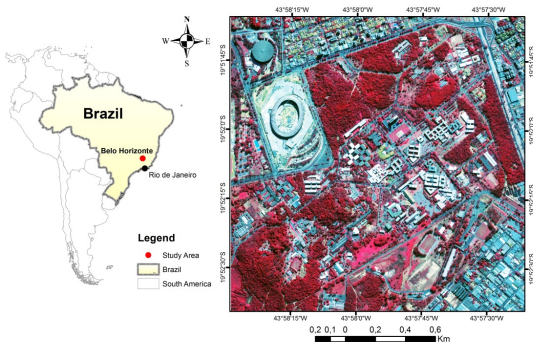
O satélite Landsat 5 (sensor *Thematic Mapper*), por exemplo, gera um vetor  $\vec{l}(p)$  com 7 valores de reflectância.



# Formação de imagem multi-banda

## Sensoriamento Remoto

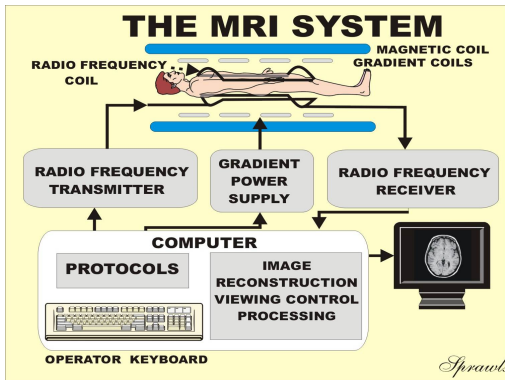
- Monitor do computador mostra apenas RGB
- Visualização de bandas só pode ser feita por combinações de 3 bandas por vez
- Cada banda é mapeada para um componente de RGB ((e.g., infra-vermelho no vermelho, vermelho no verde, e verde no azul))



# Formação de imagens médicas

## Ressonância Magnética

Mede o tempo de relaxação dos spins dos átomos de hidrogênio em uma coordenada ( $x, y, z$ ) após desmagnetização de um campo magnético forte.

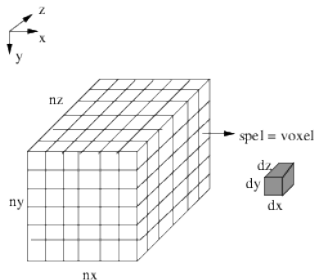
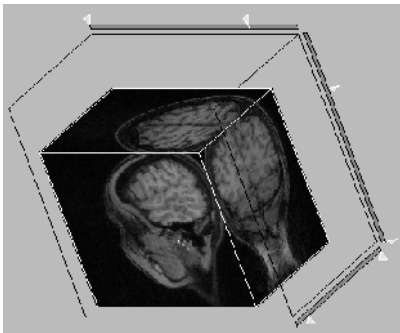




# Formação de imagens médicas

## Ressonância Magnética

Tecidos têm tempos de relaxação distintos, gerando spels (voxels - *volume elements*) com valores  $I(p)$  correspondentes.



\*Tomografia de raios-X mede a atenuação do tecido à passagem de raios-X pelo corpo

# Definição matemática de imagem digital

- Uma imagem digital  $\hat{I}$  pode ser vista como um par  $(D_I, \vec{I})$  onde  $D_I \subset Z^n$  é o domínio da imagem e todo **spel**  $p \in D_I$  tem associado um vetor de valores inteiros  $\vec{I}(p) \in Z^m$ .
- O valor de  $n$  é a dimensão no espaço imagem e o valor de  $m$  é a dimensão no espaço paramétrico.
- Se a imagem for cinza, usamos apenas  $I(p)$  para denotar os valores dos spels.
- Se a imagem sofre uma transformação matemática (e.g., conversão de RGB para  $L^*a^*b$ ), os valores de  $\vec{I}(p)$  podem ser mapeados em  $\Re^m$ .

## Definição matemática de imagem digital

- Uma imagem digital  $\hat{I}$  pode ser vista como um par  $(D_I, \vec{I})$  onde  $D_I \subset Z^n$  é o domínio da imagem e todo **spel**  $p \in D_I$  tem associado um vetor de valores inteiros  $\vec{I}(p) \in Z^m$ .
- O valor de  $n$  é a dimensão no espaço imagem e o valor de  $m$  é a dimensão no espaço paramétrico.
- Se a imagem for cinza, usamos apenas  $I(p)$  para denotar os valores dos spels.
- Se a imagem sofre uma transformação matemática (e.g., conversão de RGB para  $L^*a^*b$ ), os valores de  $\vec{I}(p)$  podem ser mapeados em  $\Re^m$ .

# Definição matemática de imagem digital

- Uma imagem digital  $\hat{I}$  pode ser vista como um par  $(D_I, \vec{I})$  onde  $D_I \subset Z^n$  é o domínio da imagem e todo **spel**  $p \in D_I$  tem associado um vetor de valores inteiros  $\vec{I}(p) \in Z^m$ .
- O valor de  $n$  é a dimensão no espaço imagem e o valor de  $m$  é a dimensão no espaço paramétrico.
- Se a imagem for cinza, usamos apenas  $I(p)$  para denotar os valores dos spels.
- Se a imagem sofre uma transformação matemática (e.g., conversão de RGB para  $L^*a^*b$ ), os valores de  $\vec{I}(p)$  podem ser mapeados em  $\Re^m$ .

# Definição matemática de imagem digital

- Uma imagem digital  $\hat{I}$  pode ser vista como um par  $(D_I, \vec{I})$  onde  $D_I \subset Z^n$  é o domínio da imagem e todo **spel**  $p \in D_I$  tem associado um vetor de valores inteiros  $\vec{I}(p) \in Z^m$ .
- O valor de  $n$  é a dimensão no espaço imagem e o valor de  $m$  é a dimensão no espaço paramétrico.
- Se a imagem for cinza, usamos apenas  $I(p)$  para denotar os valores dos spels.
- Se a imagem sofre uma transformação matemática (e.g., conversão de RGB para  $L^*a^*b$ ), os valores de  $\vec{I}(p)$  podem ser mapeados em  $\Re^m$ .

# Definição matemática de imagem digital

Exemplos de imagens que podem ser multi-dimensionais e/ou multi-paramétricas.

- Fotografia colorida:  $n = 2$  e  $m = 3$ .
- Imagens de ressonância magnética e tomografia de raios-X:  $n = 3$  e  $m = 1$ .
- Vídeo colorido:  $n = 3$  e  $m = 3$ .
- Imagem do Landsat 5:  $n = 2$  e  $m = 7$ .

# Resolução

- Radiométrica:

Considerando um mesmo intervalo em  $f(x, y, z)$ , quanto maior for a profundidade  $b$  (menor será a distância entre os níveis de quantização), maior será a resolução radiométrica de uma imagem cinza. Imagens médicas, por exemplo, costumam ter  $b=12\text{bits}$  (e.g., 4096 tons de cinza).

- Espectral:

Similarmente, maior resolução espectral está associada ao número maior de bandas em um mesmo intervalo do espectro de luz.

- Temporal:

No caso de vídeos, quanto mais quadros por segundo, maior será sua resolução temporal.

# Resolução

- Radiométrica:

Considerando um mesmo intervalo em  $f(x, y, z)$ , quanto maior for a profundidade  $b$  (menor será a distância entre os níveis de quantização), maior será a resolução radiométrica de uma imagem cinza. Imagens médicas, por exemplo, costumam ter  $b=12\text{bits}$  (e.g., 4096 tons de cinza).

- Espectral:

Similarmente, maior resolução espectral está associada ao número maior de bandas em um mesmo intervalo do espectro de luz.

- Temporal:

No caso de vídeos, quanto mais quadros por segundo, maior será sua resolução temporal.



# Resolução

- Radiométrica:

Considerando um mesmo intervalo em  $f(x, y, z)$ , quanto maior for a profundidade  $b$  (menor será a distância entre os níveis de quantização), maior será a resolução radiométrica de uma imagem cinza. Imagens médicas, por exemplo, costumam ter  $b=12\text{bits}$  (e.g., 4096 tons de cinza).

- Espectral:

Similarmente, maior resolução espectral está associada ao número maior de bandas em um mesmo intervalo do espectro de luz.

- Temporal:

No caso de vídeos, quanto mais quadros por segundo, maior será sua resolução temporal.

## Resolução

- Espacial:

O mesmo se aplica ao domínio da imagem. Para uma mesma região do espaço, quanto maior o número de voxels (menor será  $d_x$ ,  $d_y$ ,  $d_z$ ), maior será a resolução espacial da imagem.



Imagem com  $256 \times 256$  pixels e  $b=24$ bits (esquerda), reduzindo para  $b=12$ bits (centro), e reduzindo para  $64 \times 64$  pixels.

## Resolução

- Espacial:

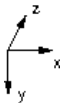
O mesmo se aplica ao domínio da imagem. Para uma mesma região do espaço, quanto maior o número de voxels (menor será  $d_x$ ,  $d_y$ ,  $d_z$ ), maior será a resolução espacial da imagem.



Imagem com  $256 \times 256$  pixels e  $b=24$ bits (esquerda), reduzindo para  $b=12$ bits (centro), e reduzindo para  $64 \times 64$  pixels.

## Representação da imagem

Podemos adotar a representação matricial com  $n$  dimensões (e.g.,  $n = 3$ ) onde cada célula terá  $m$  valores ( $l_1(p), \dots, l_m(p)$ ):



	z=1	x=0	x=1	x=2	x=3
y=0		p=12	p=13	p=14	p=15
y=1		p=16	...		
y=2					p=23

	z=0	x=0	x=1	x=2	x=3
y=0		p=0	p=1	p=2	p=3
y=1		p=4	p=5	...	
y=2					p=11

No entanto, a forma mais comum é armazenar os valores dos pixels linearmente, fatia por fatia (ao longo de  $z$ ), da esquerda para direita (ao longo de  $x$ ), e de cima para baixo (ao longo de  $y$ ).

# Representação da imagem

Para  $n_x \times n_y \times n_z$  voxels:

$$\begin{aligned} p &= x_p + y_p n_x + z_p n_x n_y \\ x_p &= \text{mod}(\text{mod}(p, n_y n_x), n_x) \\ y_p &= \text{mod}(p, n_y n_x) / n_x \\ z_p &= p / (n_y n_x) \end{aligned}$$

# Roteiro da Aula

- 1 Formação de Imagem
- 2 Espaços de Cor
- 3 Imagens multi-dimensionais e/ou multi-paramétricas
- 4 Formatos de armazenamento de imagem**
- 5 Histograma de Cor

# Formatos de armazenamento da imagem em disco

A imagem é sempre armazenada na forma linear em qualquer formato (tif, gif, jpg, png, pgm, ppm, etc)

## Imagem colorida ppm do tipo P6:

P6

$n_x \ n_y$

255

em binário:

$l_1(0)l_2(0)l_3(0)l_1(1)l_2(1)l_3(1)\dots l_1(n_x n_y - 1)l_2(n_x n_y - 1)l_3(n_x n_y - 1)$

# Formatos de armazenamento da imagem em disco

A imagem é sempre armazenada na forma linear em qualquer formato (tif, gif, jpg, png, pgm, ppm, etc)

## Imagem no formato PGM, P2, com 4 bits (valores de 0 a 15):

```
P2
24 7
15
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 3 3 3 3 0 0 7 7 7 7 0 0 11 11 11 11 0 0 15 15 15 15 0
0 3 0 0 0 0 0 7 0 0 0 0 0 11 0 0 0 0 0 15 0 0 15 0
0 3 3 3 0 0 0 7 7 7 0 0 0 11 11 11 0 0 0 15 15 15 15 0
0 3 0 0 0 0 0 7 0 0 0 0 0 11 0 0 0 0 0 15 0 0 0 0
0 3 0 0 0 0 0 7 7 7 7 0 0 11 11 11 11 0 0 15 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

Imagem resultante ampliada:





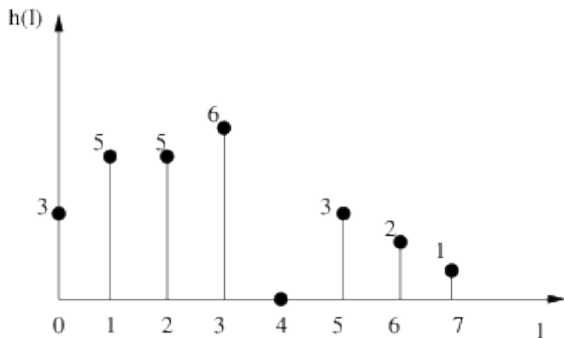
# Roteiro da Aula

- 1 Formação de Imagem
- 2 Espaços de Cor
- 3 Imagens multi-dimensionais e/ou multi-paramétricas
- 4 Formatos de armazenamento de imagem
- 5 Histograma de Cor**

## Algoritmos envolvendo imagem digital

- Um exemplo simples é o cálculo do vetor de frequências (**histograma**) dos valores de uma imagem cinza.
- O histograma  $h(I)$  de uma imagem  $\mathbf{I} = (D_I, I)$  cinza deve armazenar em cada bin  $0 \leq I \leq 2^b - 1$ , o número de spels  $p \in D_I$  com valor  $I(p) = I$ .

5	5	3	3	1
1	1	3	1	1
0	2	6	0	5
0	2	3	2	6
2	3	3	2	7



# Algoritmo para cálculo do histograma

Entrada: Imagem  $\hat{I} = (D_I, I)$ .

Saída: Histograma  $h(I)$  com  $2^b$  bins.

- 1 Para todo  $l = 0, \dots, 2^b - 1$ , faça  $h(l) \leftarrow 0$ .
- 2 Para todo  $p \in D_I$ , faça  $h(I(p)) \leftarrow h(I(p)) + 1$ .

## Exercícios

- 1 Estenda este algoritmo para imagens coloridas com 24bits (8 bits por componente de cor). Note que o número de bins é muito elevado, então você deve dividir cada eixo de cor em  $N$  intervalos de modo a obter um histograma com  $N^3$  bins.
- 2 Qual a profundidade de uma imagem com 65536 níveis de cinza?