

# FUNDAMENTOS DE PROCESSAMENTO DE IMAGENS

- Modelo de Câmera *Pinhole*
- Amostragem, Quantização, Digitalização
- A Geometria da Formação de Imagens

# Aquisição de imagens

Cenário

+

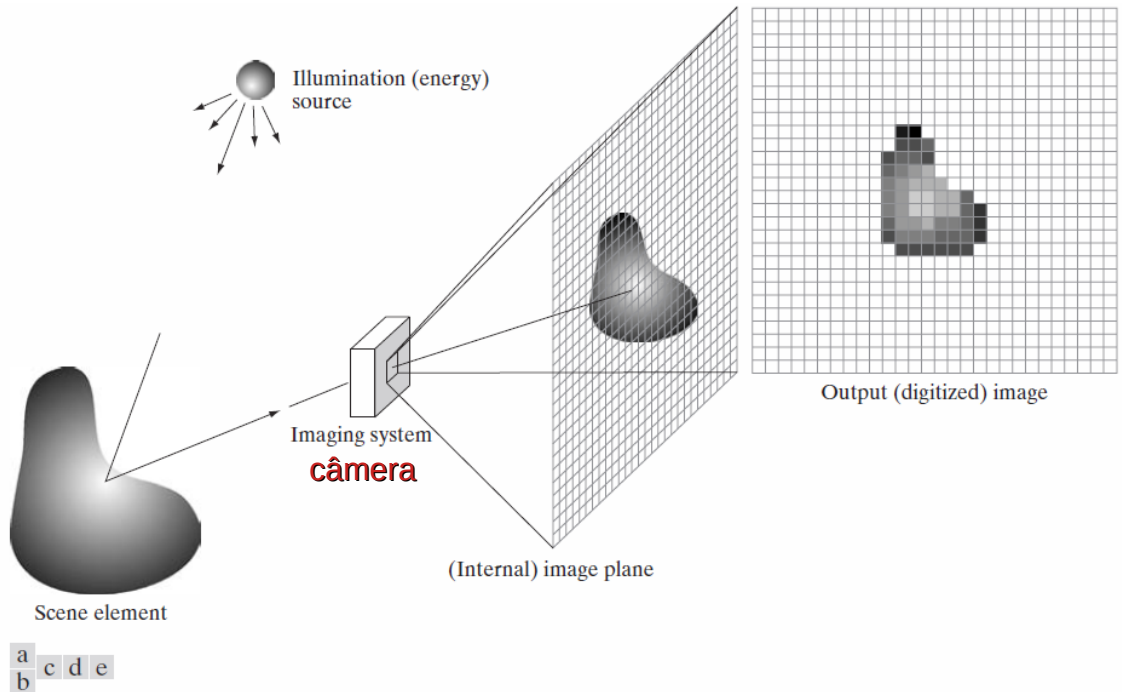
Fonte de Luz

+

Câmera

-Analógica

-Digital

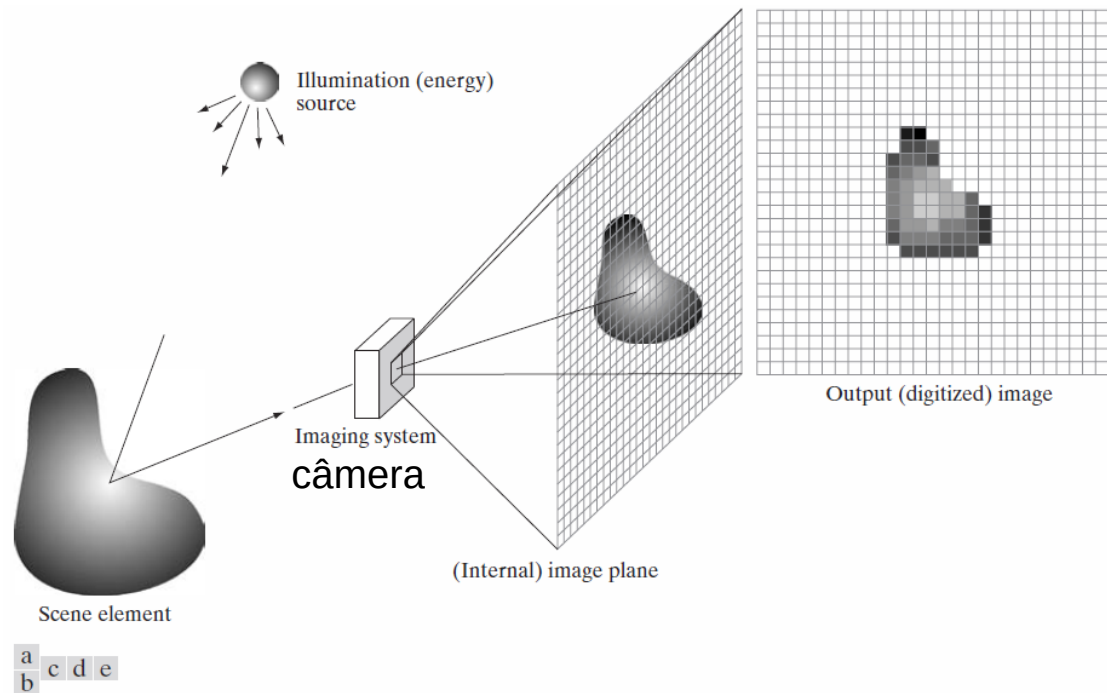


**FIGURE 2.15** An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Nosso foco vai ser na aquisição digital de imagens

# Aquisição de imagens

## Cenário + Fonte de Luz + Câmera

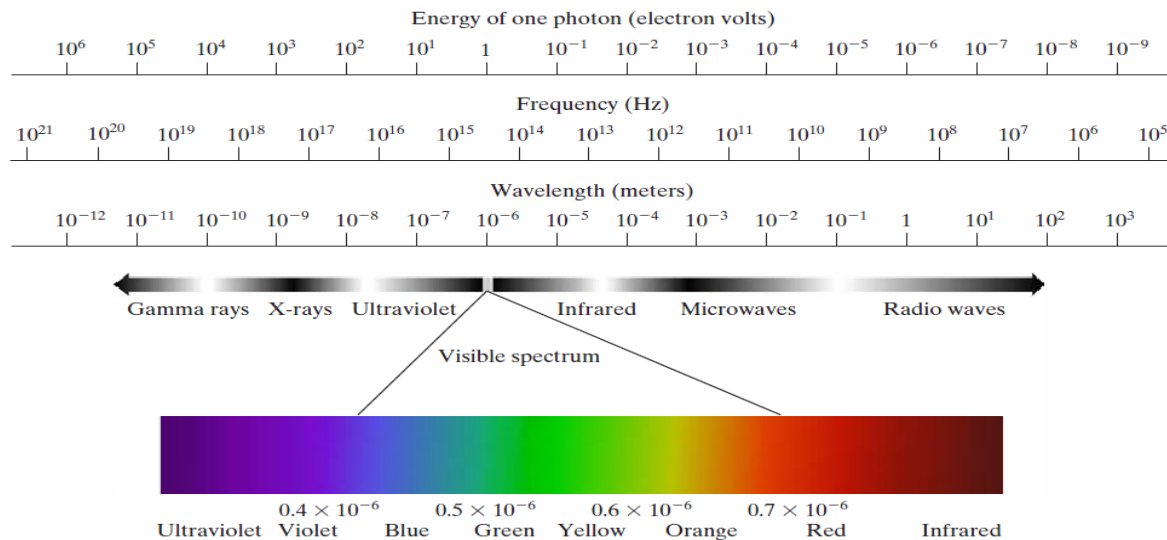


**FIGURE 2.15** An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

Vamos fazer uma breve discussão sobre a fonte de luz, a câmera e seu sensor (digitalização) e a geometria de formação de imagens.

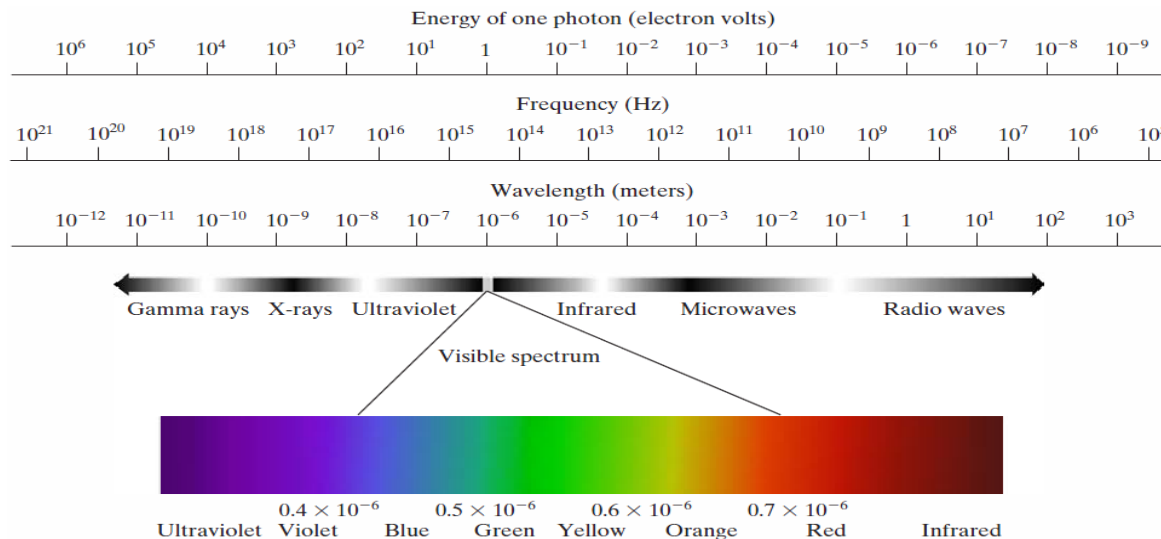
# “Fonte de Luz”

A fonte de luz pode estar em qualquer ponto do espectro eletromagnético...



# “Fonte de Luz”

Para a radiação ionizante, com capacidade de atravessar materiais. (RX, por exemplo, apresenta  $\lambda$  em torno de  $10^{-9}\text{m} = 10\text{ nm}$ ) as imagens não são capturadas sobre um sensor comum



## “Fonte de Luz”

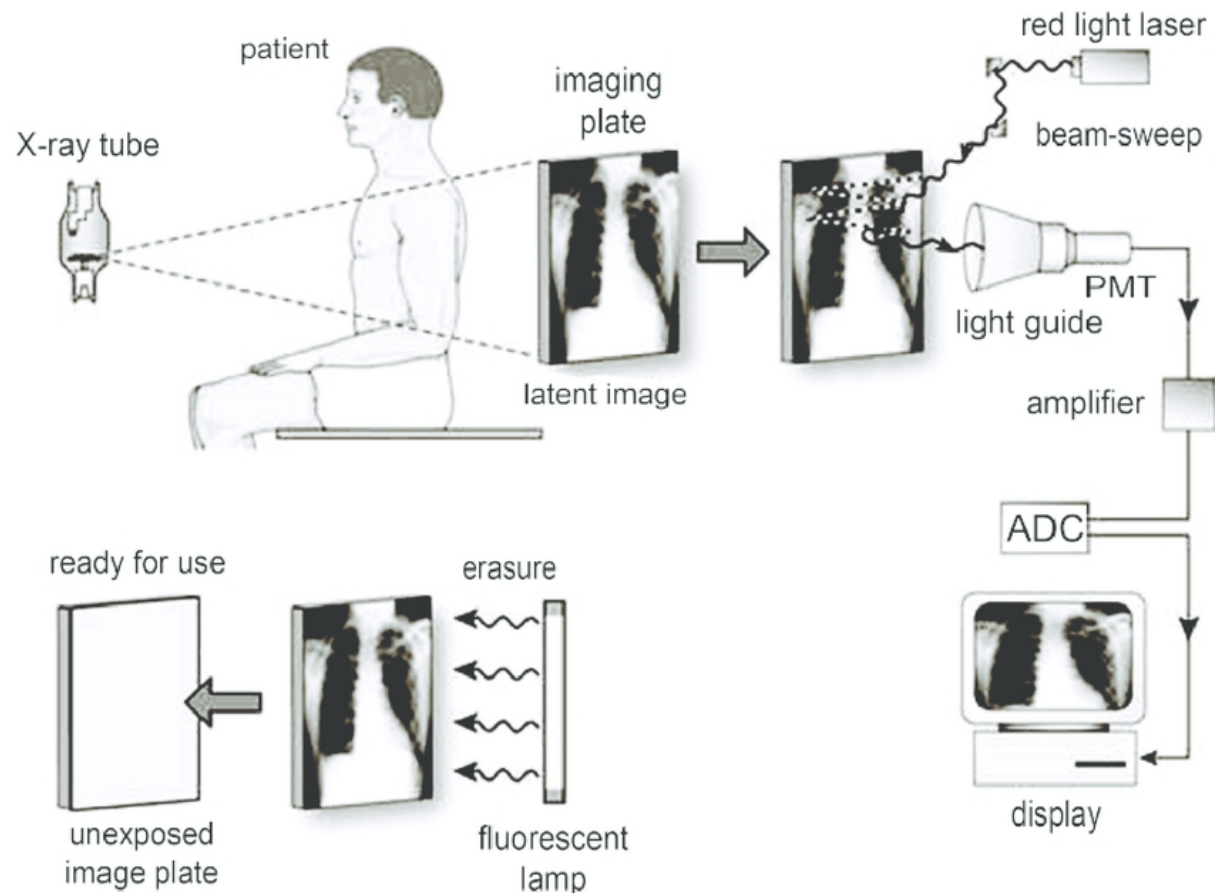
É possível a captura da imagem a partir de uma fonte de RX, por exemplo:

- CCD with phosphor: not direct detection (radiation hardness), phosphor ruins spatial resolution, good for tomography.
- CMOS: direct detection, good spatial resolution, good sensitivity (single electron detection).

LINK

# “Fonte de Luz”

Captura a partir de uma fonte de RX:  $\lambda_{RX}$  situa-se em torno de  $10^{-9}\text{m} = 10\text{ nm}$   
O usual é aplicar-se uma etapa de coleta direta pela incidência de RX sobre uma placa tratada com fósforo (“apagável”) seguida da captura da imagem por um sistema de laser aplicado sobre essa placa



Fonte da figura.

Link:  
CMOS para captui  
ra de RX

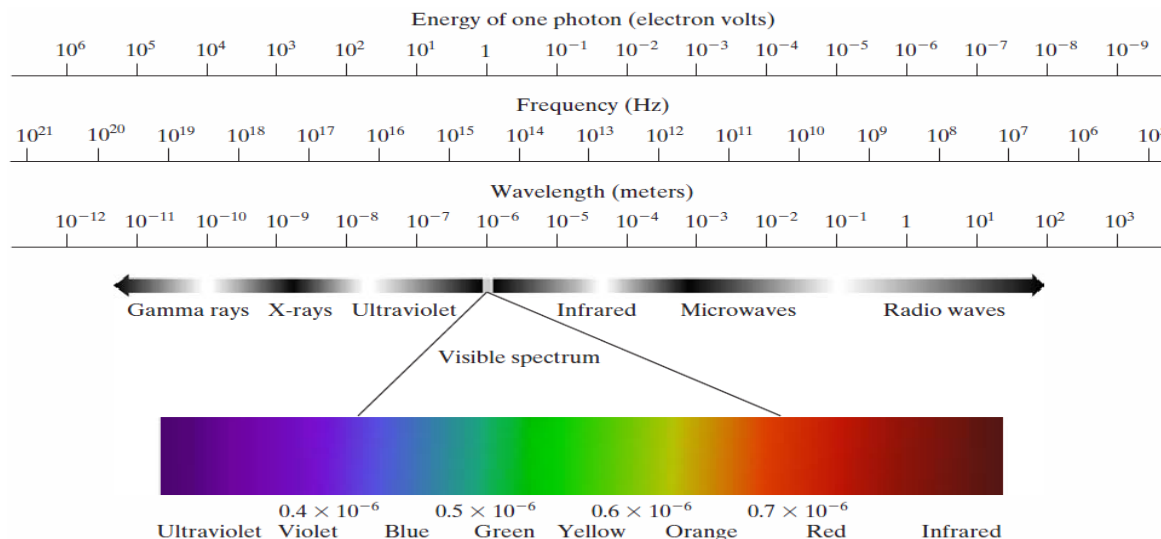
# “Fonte de Luz”

- A onda eletromagnética que irá detectar o objeto-alvo “sensoreado” deve apresentar comprimento de onda ( $\lambda$ ) menor ou igual às dimensões do objeto;
- Para capturar a imagem de uma molécula de  $\text{H}_2\text{O}$ , a qual mede aproximadamente 0,1nm é preciso uma fonte de radiação com  $\lambda$  na faixa  $\leq$  ultravioleta;
- Também é possível capturar imagens por outros meios que não a radiação de ondas eletromagnéticas. Por exemplo. Via ultrassonografia;

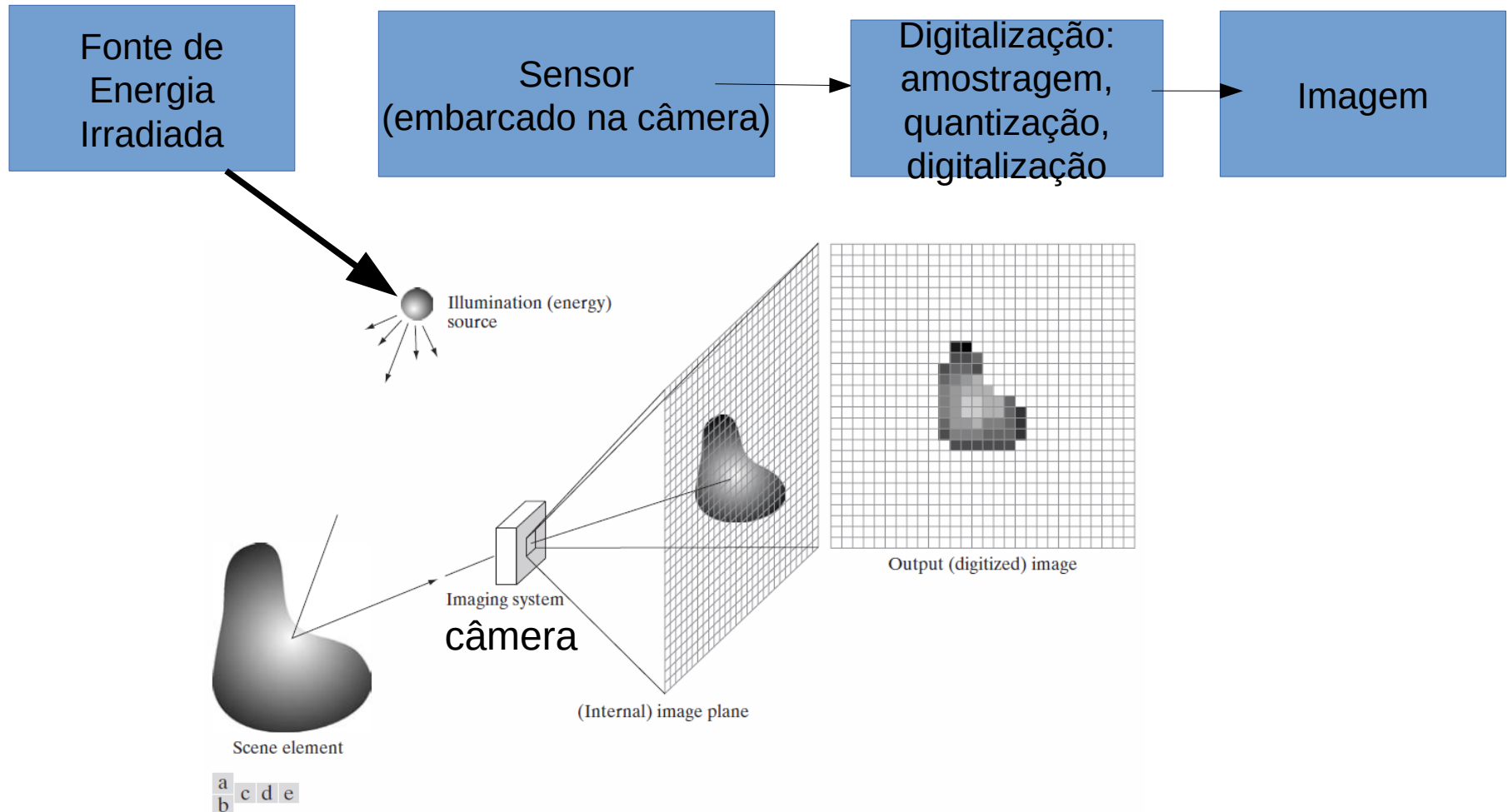


# “Fonte de Luz”

- Estaremos mais interessados em fontes de luz no **espectro visível**
- Luz acromática (sem cor) também chamada de monocromática: é percebida apenas pela sua intensidade. Corresponde a variações de cinza do branco ao preto;
- Luz cromática ou colorida: é percebida pela radiância (energia emitida da fonte), luminância (energia percebida pelo observador) e brilho (subjetivo).



# O sensor e a digitalização



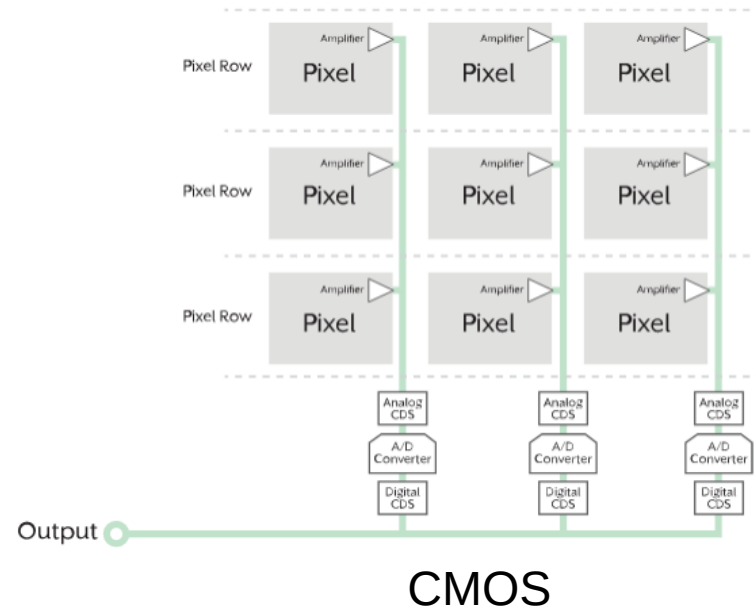
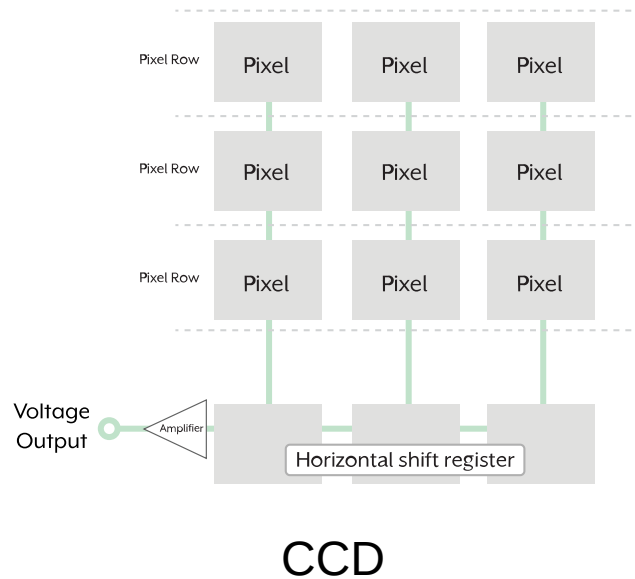
**FIGURE 2.15** An example of the digital image acquisition process. (a) Energy (“illumination”) source. (b) An element of a scene. (c) Imaging system. (d) Projection of the scene onto the image plane. (e) Digitized image.

# O sensor e a digitalização

Há basicamente duas tecnologias de sensores: CMOS e CCD

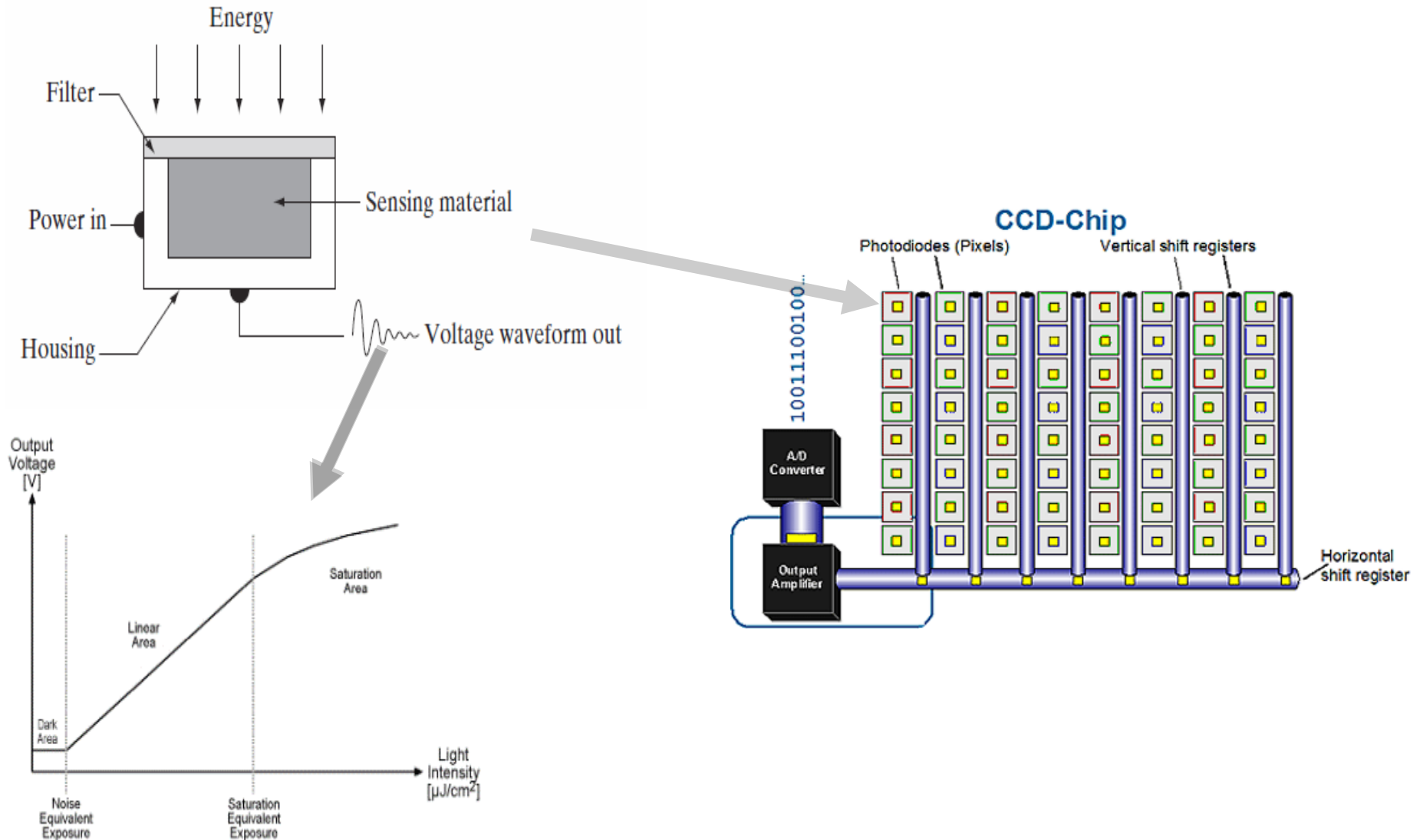
CMOS (Complementary Semiconductor)

Metal-Oxide



# O sensor e a digitalização

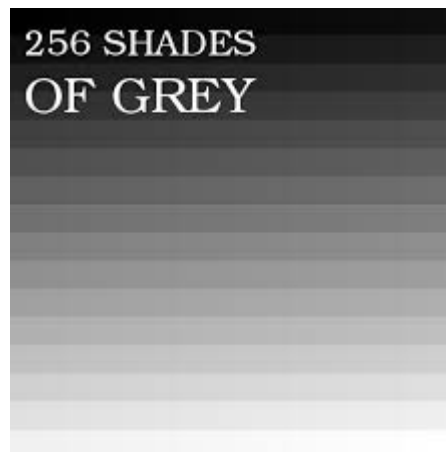
CCD (Charge-Coupled Device), nos referiremos mais a este tipo e tratando de imagens em tons de cinza (monocromáticas).



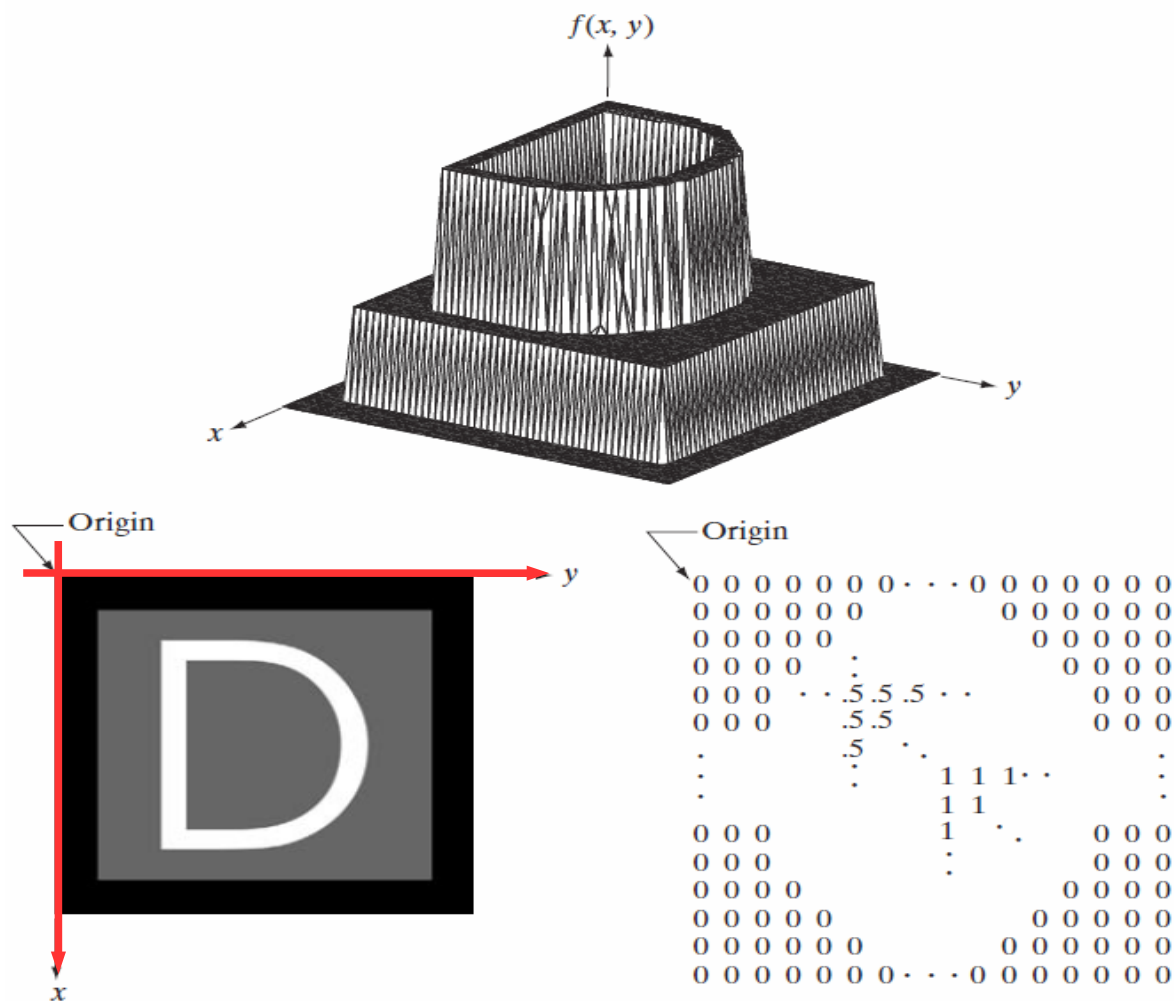
Observação: na imagem monocromática, ou em **tons de cinza**, cada pixel(x,y) apresenta um certo brilho (sem cor).

Na escala de cinza comum esse brilho corresponde ao valor no intervalo  $[0, 2^k - 1]$ . Onde,  $k$  é a quantidade de bits disponível por pixel na representação de cada nível de cinza.

Usualmente,  $k=8$  e a escala é de 256 tons de cinza no intervalo  $[0, 255]$ , zero corresponde o preto e 255 ao branco



Você pode também interpretar a imagem em tons de cinza como uma superfície



a  
b c

**FIGURE 2.18**

(a) Image plotted as a surface.

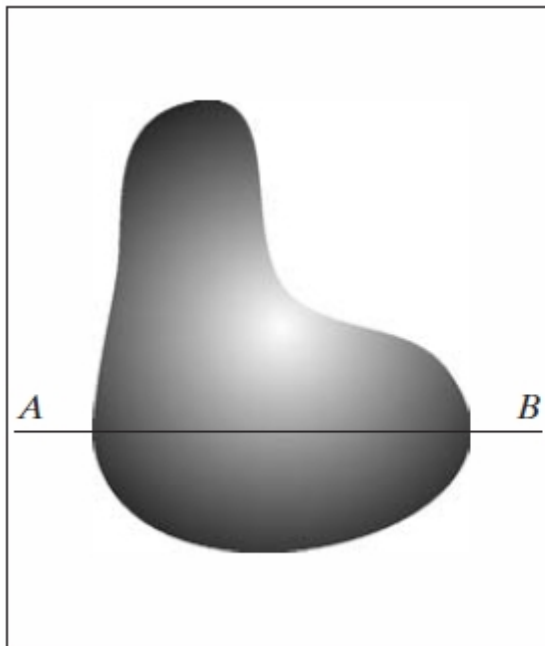
(b) Image displayed as a visual intensity array.

(c) Image shown as a 2-D numerical array (0, .5, and 1 represent black, gray, and white, respectively).

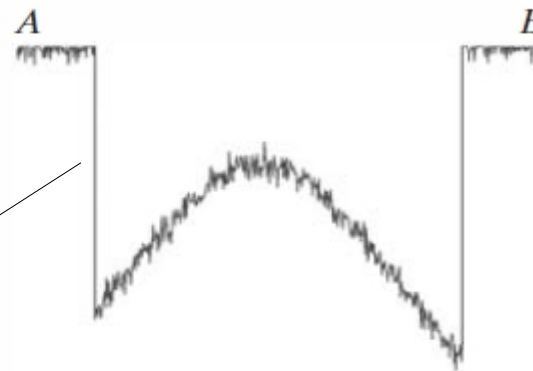
# O sensor e a digitalização

- Imagem em tons de cinza

Cena



Representação de uma hipotética leitura contínua das intensidades de brilho da linha AB na cena

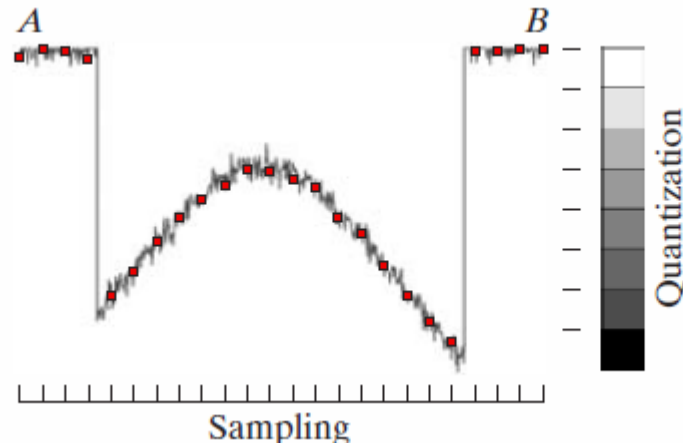
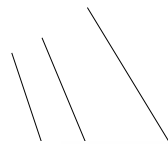
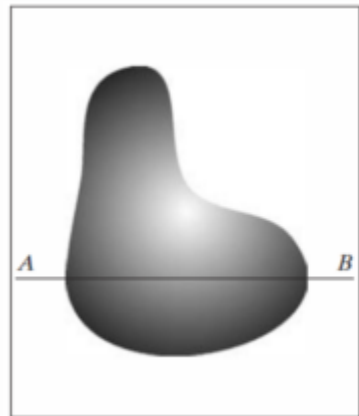


A representação contínua é apenas ilustrativa pois a leitura é de fato uma amostragem em pixels (discretos)...

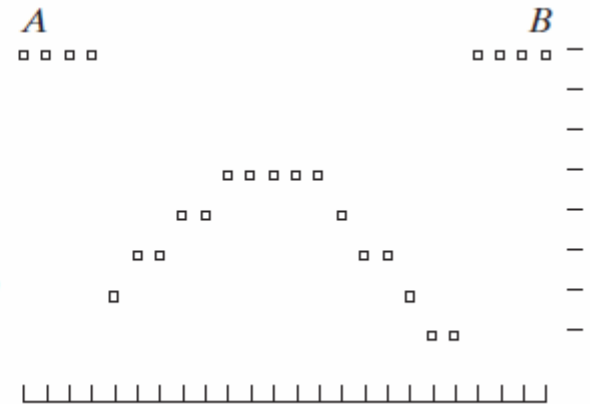
# O sensor e a digitalização

- Imagem em tons de cinza

amostragem das intensidades de brilho da cena (linha AB)



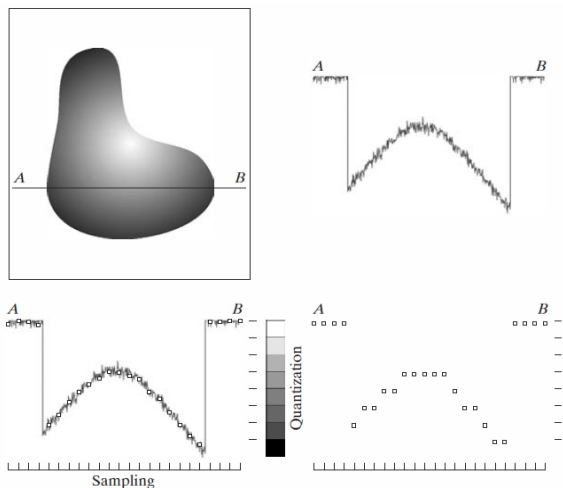
... no caso, as amostras foram quantizadas em 8 níveis de cinza





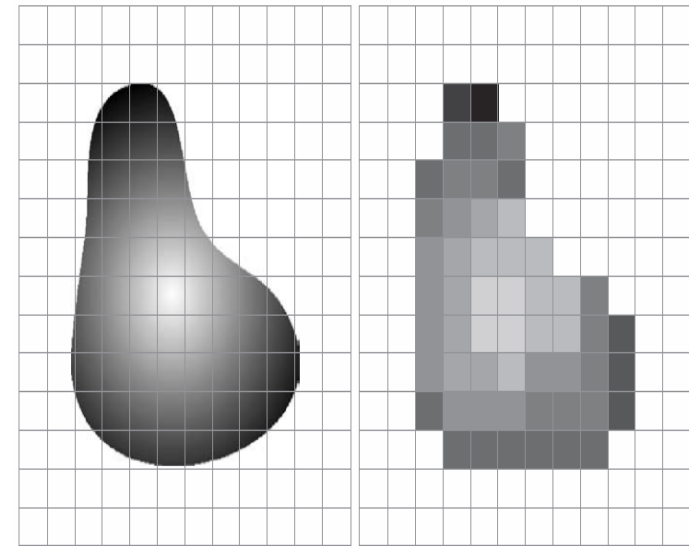
# O sensor e a digitalização

- Imagem em tons de cinza



“Digitalização” :

Especificação da quantidade de bits necessários ao atendimento da quantização. Ou seja, quantos bits serão necessários para a representação de cada nível discretos do sinal quantizado. No exemplo ao lado, tomando como referência a linha AB, há 8 níveis de cinza portanto são necessários pelo menos 3 bits por pixel.



a b

**FIGURE 2.17** (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

# Exemplo: digitalização em diferentes quantizações

Nível de tensão (mV)	Digitalização na escala $[0, 2^k - 1]$			
	K=1 $[0,1]$	K=2 $[0,3]$	K=3 $[0,7]$	K=4 $[0,15]$
0	0	00	000	0000
0,125	0	00	001	0010
0,25	0	01	010	0100
0,375	0	01	011	0110
0,5	0/1?	10	100	1000
0,625	1	10	101	1010
0,75	1	10/11?	110	1100
0,825	1	11	111	1110
1	1	11	111	1111

Escala de cinza:  $[0, 2^k - 1]$ ,  $k$  = número de bits por pixel (bpp).

# O sensor e a digitalização

Suponha que a amostragem realizada por um sensor digital produziu a matriz abaixo, valores em mV (milivolts). Cada valor representa uma intensidade de brilho (tons de cinza) vinda da cena adquirida.

Considerando que haja restrição de memória no seu sistema, discuta o caso e faça a discretização e digitalização dessa imagem, produzindo uma imagem digital de saída.

6.8	6.8	6.8	6.8	6.8	6.8
1.8	1.8	1.8	1.8	1.8	1.8
6.8	6.8	6.8	6.8	6.8	6.8
1.8	1.8	1.8	1.8	1.8	1.8
6.8	6.8	6.8	6.8	6.8	6.8
1.8	1.8	1.8	1.8	1.8	1.8

# O sensor e a digitalização

Suponha que a amostragem realizada por um sensor digital produziu a matriz abaixo, valores em mV (milivolts). Cada valor representa uma intensidade de brilho (tons de cinza) vinda da cena adquirida.

Considere que a sua câmera possa realizar digitalizações com as seguintes profundidades de pixels (diferentes resoluções radiométricas): 1bpp, 2 bpp, 4 bpp, 8 bpp, 12 bpp. Compare os resultados obtidos para discretização e digitalização dessas imagens em termos da resolução de detalhes das imagens. Discuta um critério de escolha, entre as alternativas de digitalização, se o sistema de processamento de imagens (posterior à captura da cena) tiver restrições de memória.

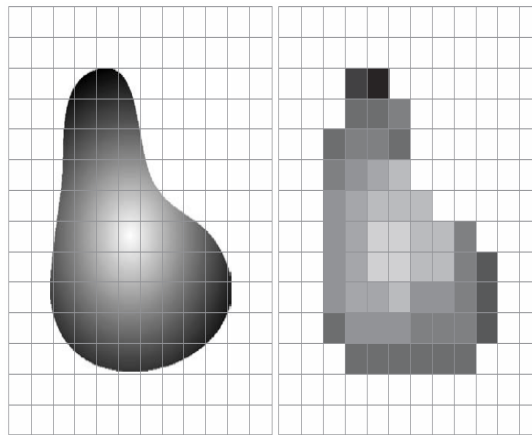
6.8	6.87	6.89	6.89	6.87	6.8
5.8	5.87	5.89	5.89	5.87	5.8
4.8	4.87	4.89	4.89	4.87	4.8
3.8	3.87	3.89	3.89	3.87	3.8
2.8	2.87	2.89	2.89	2.87	2.8
1.8	1.87	1.89	1.89	1.87	1.8

Observação:

bpp significa bit(s) por pixel.

# O sensor e a digitalização

- Resolução espacial x radiométrica
  - A amostragem/quantização estão relacionadas a duas características importantes, quais sejam a resolução espacial (densidade de pixels) e a resolução radiométrica ou profundidade de pixel (quantidade de bits por pixel)

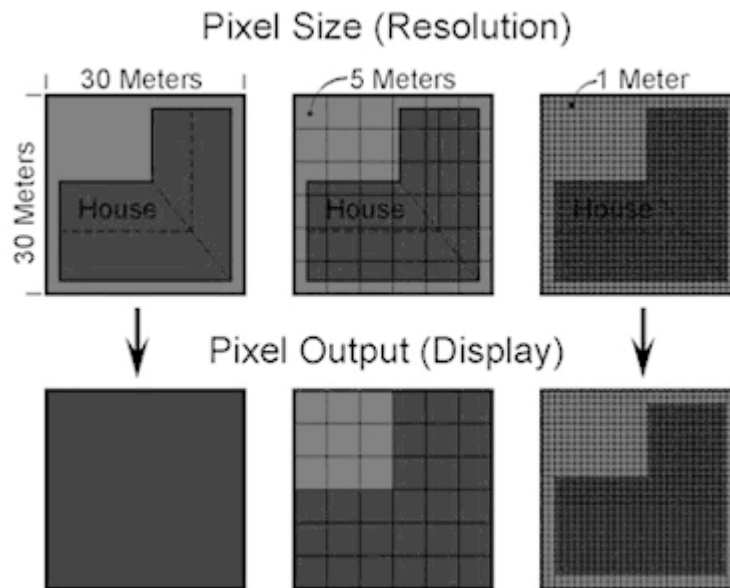


a b

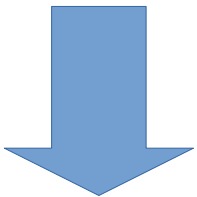
**FIGURE 2.17** (a) Continuous image projected onto a sensor array. (b) Result of image sampling and quantization.

# O sensor e a digitalização

- Resolução espacial (densidade de pixels - dpi)

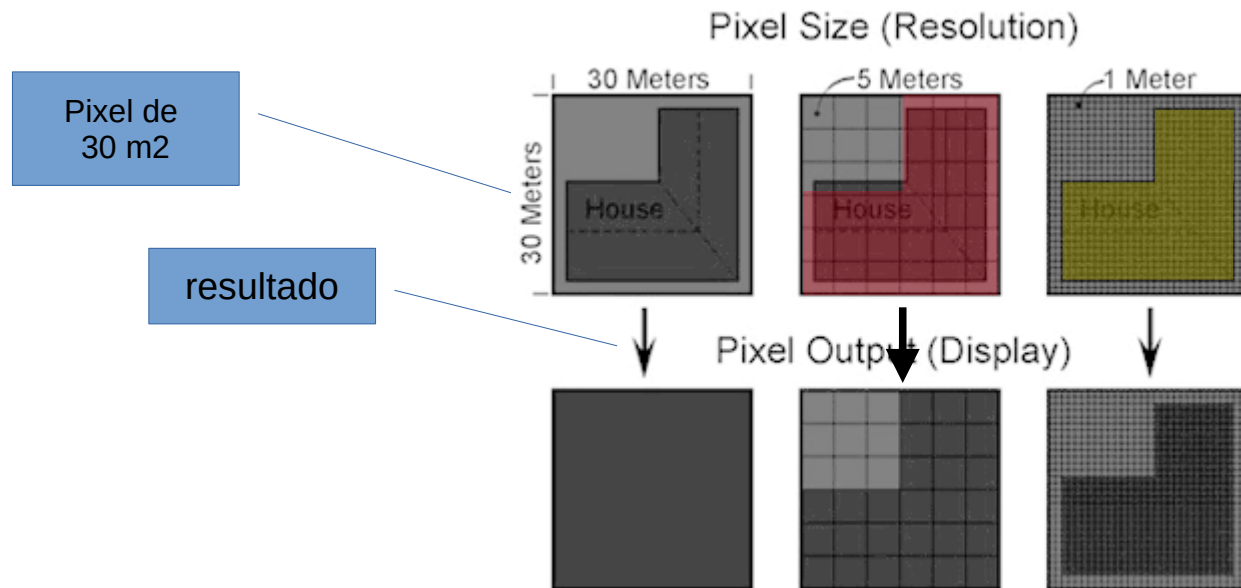


A ativação de cada pixel depende das características de resposta do sensor



# O sensor e a digitalização

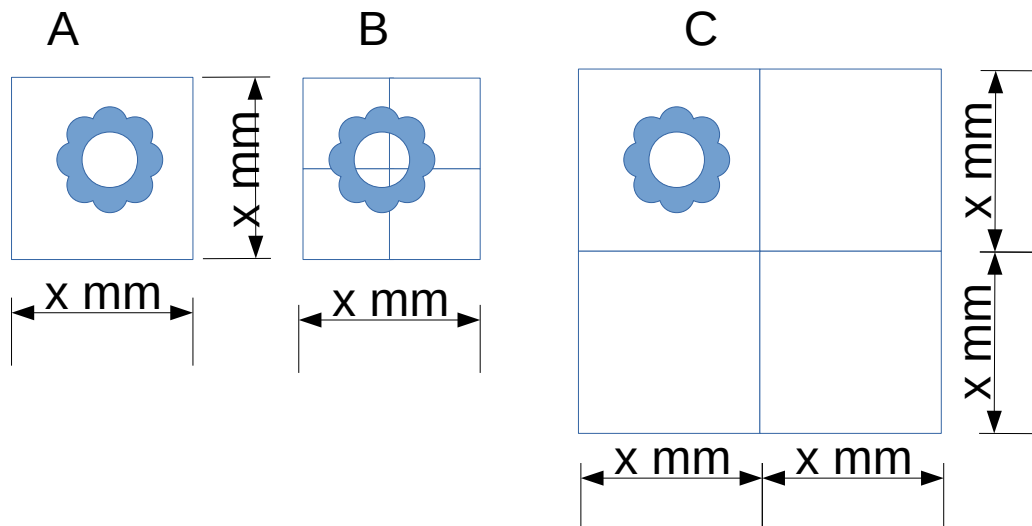
- Resolução espacial (densidade de pixels – dpi);
- Quanto mais pixels houver em uma mesma área, maior será a resolução espacial.



A ativação de cada pixels depende das características de resposta do sensor

# O sensor e a digitalização

- Resolução espacial (densidade de pixels)
- Discuta sobre a resolução das imagens abaixo:
  - Em A, B e C O que muda é apenas a densidade de pixels do sensor, a distância da câmera ao alvo (flor) bem como os parâmetros da câmera (distância focal, abertura, etc) são sempre os mesmos.
  - A = 1x1 pixel, B = 2x2, C=2x2
  - $X=0,0076\text{mm}$



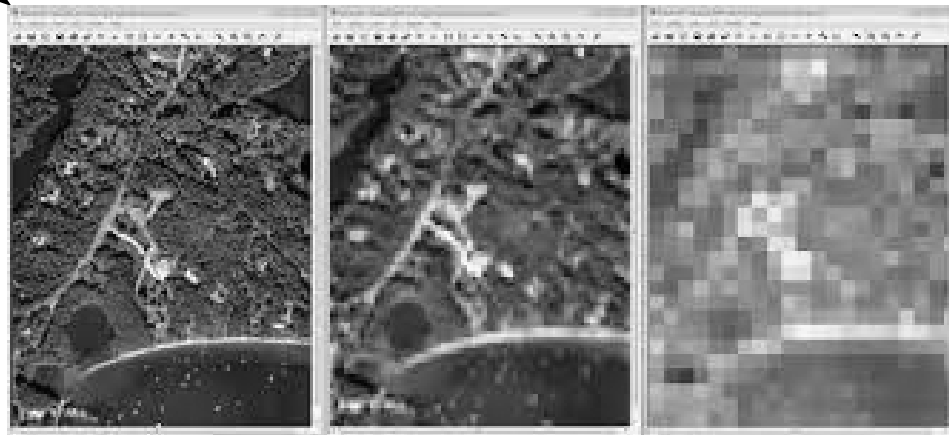


# O sensor e a digitalização

- Resolução espacial (densidade de pixels)

Mais amostras  
Mais informação

Menos amostras  
Menos informação



Mesma área

# O sensor e a digitalização

- A resolução radiométrica (profundidade de pixel) está relacionada ao conceito de range dinâmico
  - Discuta sobre a resolução das imagens abaixo

(A) 8 bits (256 levels)



(B) 4 bits (16 levels)



(C) 2 bits (4 levels)

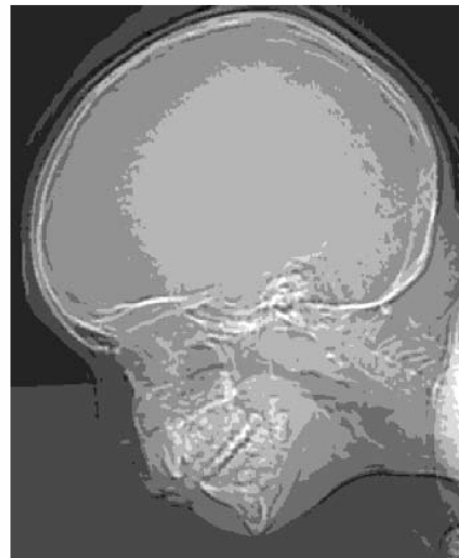


(D) 1 bit (2 levels)



# Resolução radiométrica

Mais bits por pixel  
Pixels,  $k=5$



e f  
g h

**FIGURE 2.21**  
(Continued)  
(e)–(h) Image displayed in 16, 8, 4, and 2 intensity levels. (Original courtesy of Dr. David R. Pickens, Department of Radiology & Radiological Sciences, Vanderbilt University Medical Center.)



Menos bits por pixel  
Pixels,  $k=1$   
Imagem binária

# Range Dinâmico



Range dinâmico=  $A/B$

A: nível máximo representável pelo sistema

B: nível mínimo representável pelo sistema

A: saturação, B: ruído

Contraste: diferença entre a maior e a menor  
Intensidade de brilho na imagem

Se uma expressiva quantidade de pixels  
apresenta alto range dinâmico, então há alto  
contraste na imagem

O contrário indica baixo contraste.

# O sensor e a digitalização

O impacto da velocidade de amostragem da câmera na formação de vídeos (sequências de imagens);

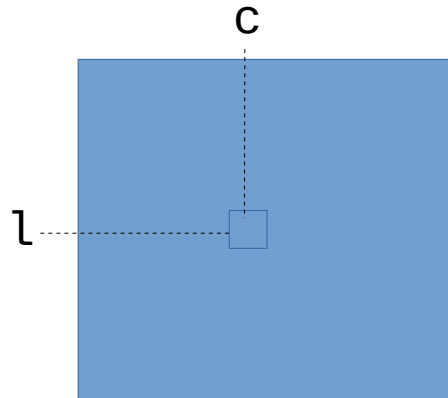
Aqui uma explicação bem ilustrativa: [LINK](#)

- A velocidade da câmera diz respeito à quantidade de frames (imagens) que podem ser produzidas por segundo (amostragem -> quantização -> digitalização → saída).
- Exemplo, em câmeras comuns temos 25 fps (frames por pixels);
- Porém, há exemplos de câmeras especiais de 1.000 fps e até mais ([LINK](#))

Em PIM iremos lidar principalmente com  
imagens em tons de cinza e coloridas em  
formato  RGB

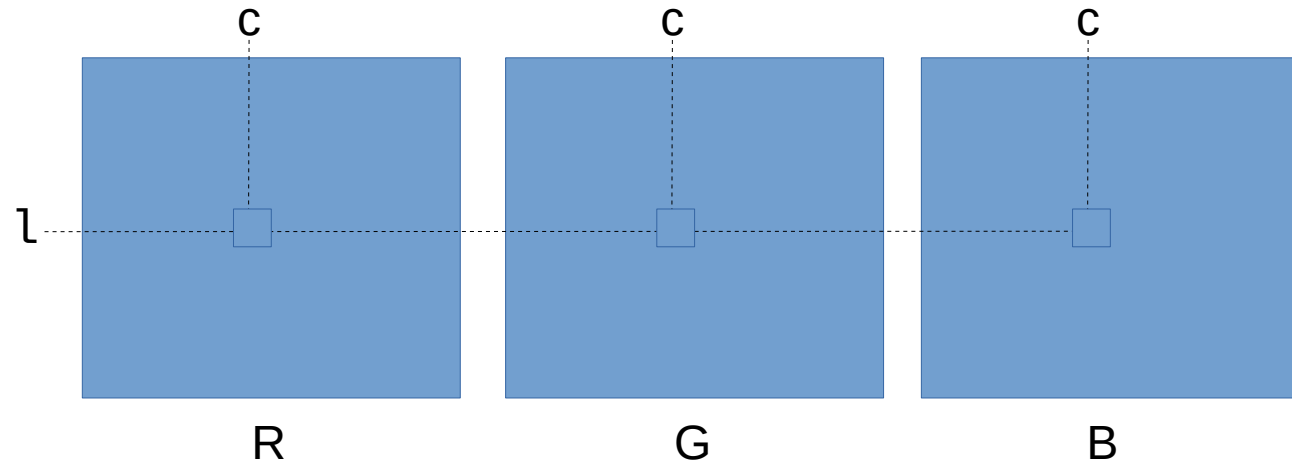
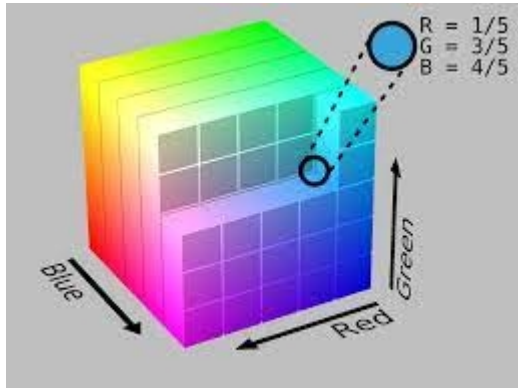
# Imagem – RGB

Tons de cinza



pixel(l,c) apresenta um valor no intervalo [0-255]

RGB



$\text{pixelRGB}(l,c) = (\text{pixR}(l,c); \text{pixG}(l,c); \text{pixB}(l,c))$

$\text{pixR}(l,c)$ ,  $\text{pixG}(l,c)$  e  $\text{pixB}(l,c)$  apresentam valor no intervalo [0-255]

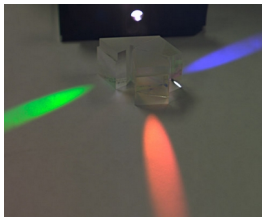
# Imagem – RGB

## Três Sensores CCD

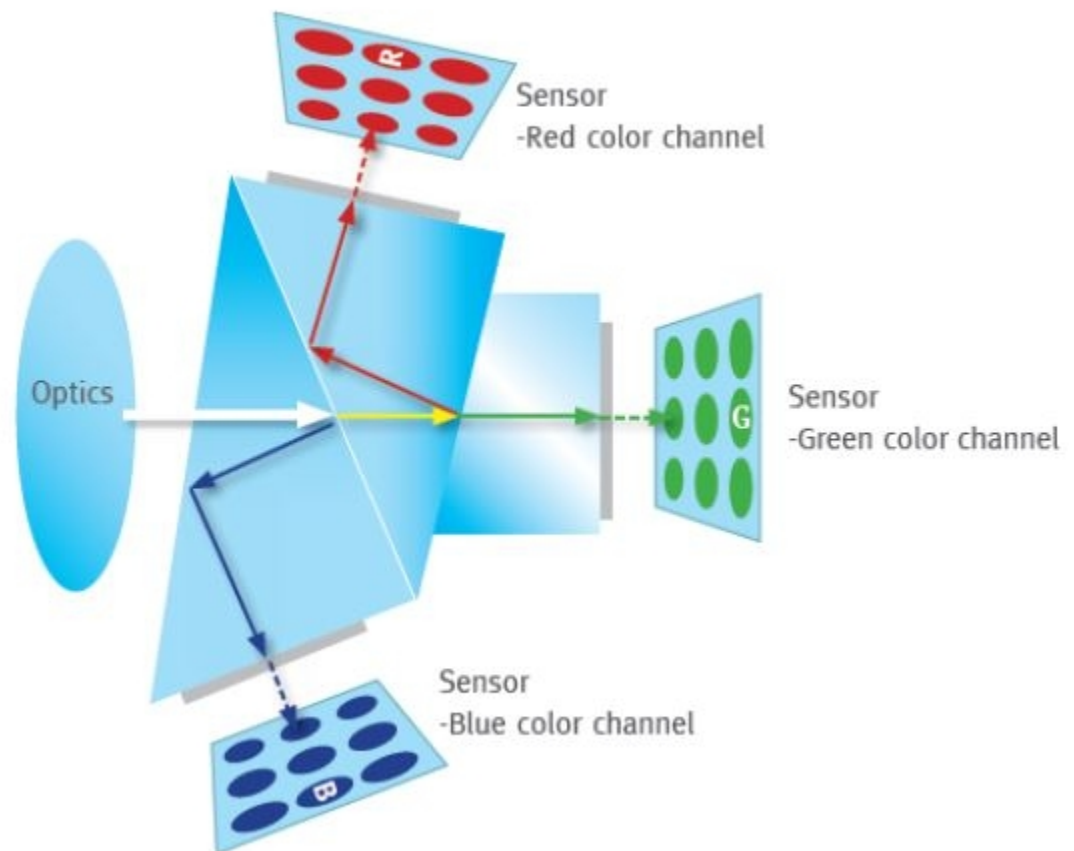
3 sensores (CCDs) distintos, cada um recebe um valor específico de R, G ou B.

Melhor acurácia de cor.

A informação de cor é separada em R, G e B por meio de um prisma óptico.



*Principle: Prism-based 3-CCD camera:*





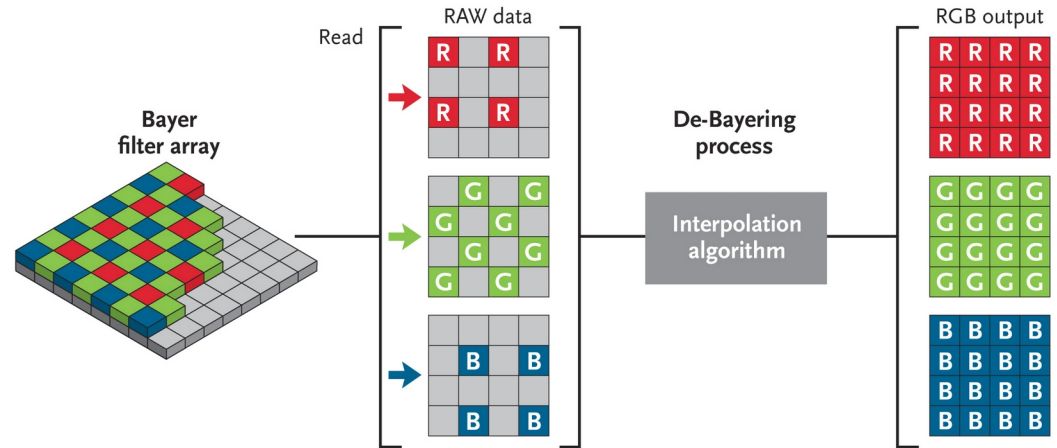
# Imagem – RGB

## Um Sensor CCD / Bayer Pattern + interpolação

Nesse caso, um sistema de filtro (Bayer Pattern) é aplicado e o sensor único recebe um mosaico de valores RGB;

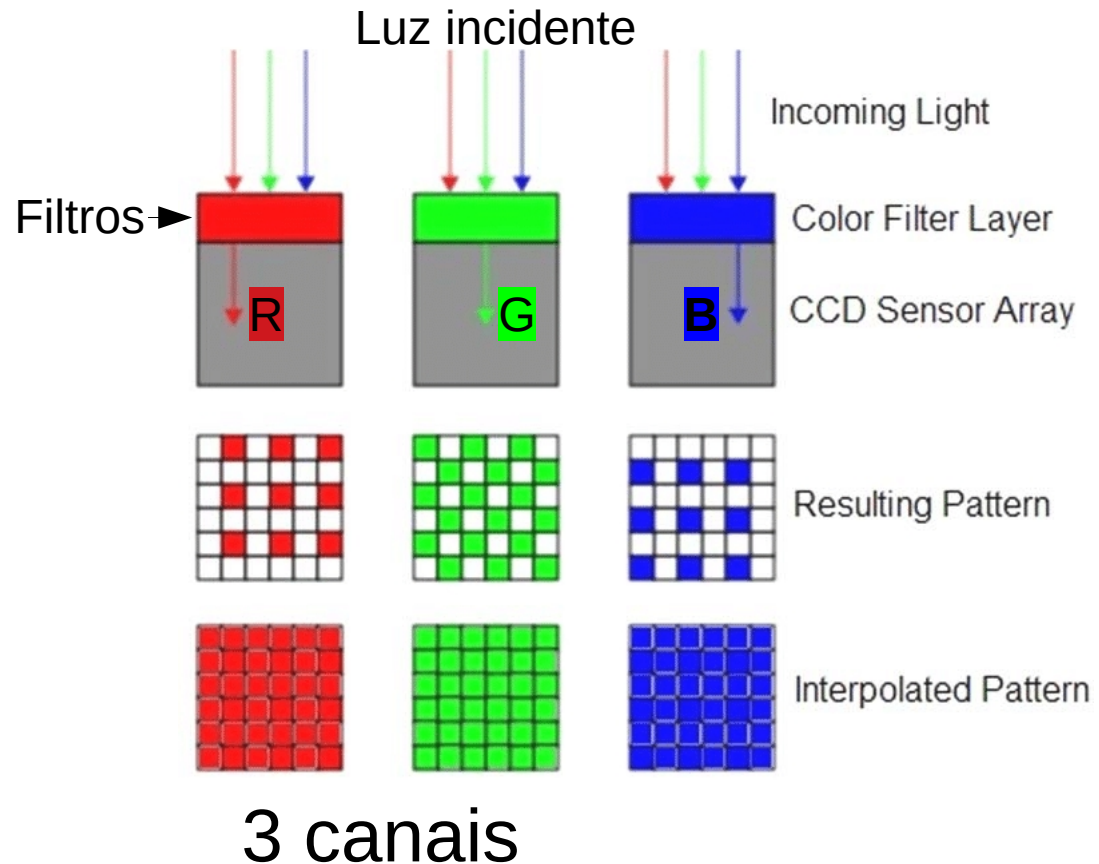
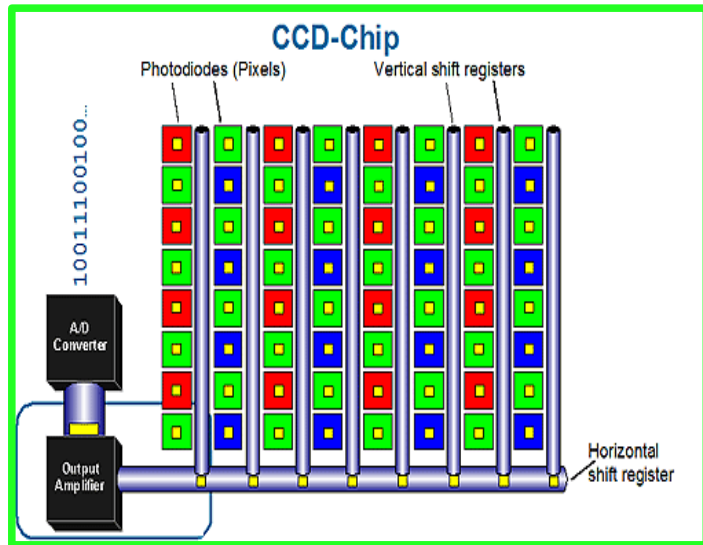
Não há a informação específica RGB de cada pixel, O filtro distribui valores individuais de R G e B seguindo um padrão;

Cada canal de cor (RGB) apresenta lacunas, as quais são preenchidas por interpolação.

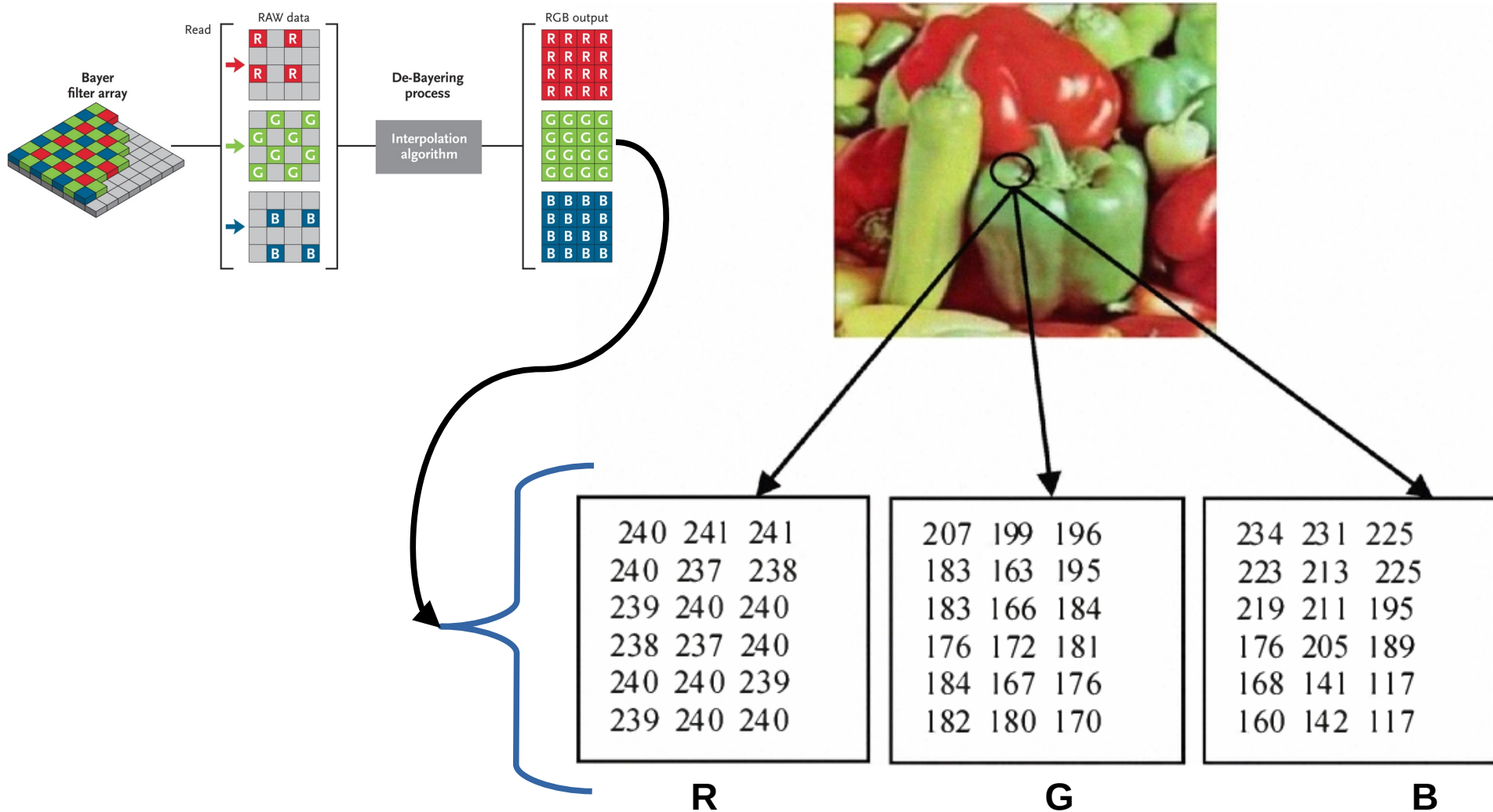


# Imagem – RGB

## Sensor CCD Colorido



# Imagem – RGB



# Imagem – RGB

## Sensor CCD Colorido

Determine os canais R, G e B a partir do padrão Bayer abaixo :

10	130	15	110	15	120
215	40	250	30	250	40
15	255	15	255	15	230
210	30	255	45	250	45
10	115	10	110	10	115
110	30	110	35	115	45

# Imagem – vs.

Além dos 3 canais (RGB) a imagem **pode** ainda contemplar um canal extra chamado de canal alfa, por meio do qual é possível gerar a sensação de opacidade (ou transparência) de uma imagem.

Portanto, uma imagem colorida pode eventualmente estar codificada na forma **RGBA**

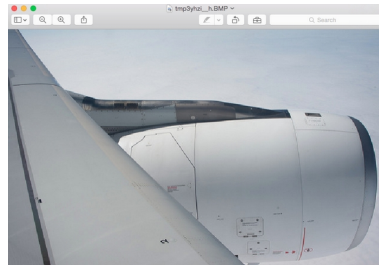
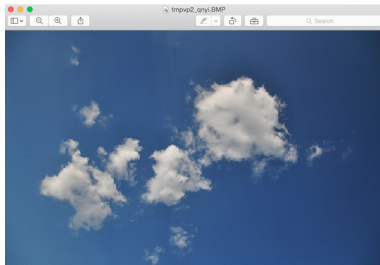
<http://www.fazendovideo.com.br/artigos/alpha-channel-a-transparencia-das-imagens.html>

# Imagem – RGB vs. RGBA

RGB



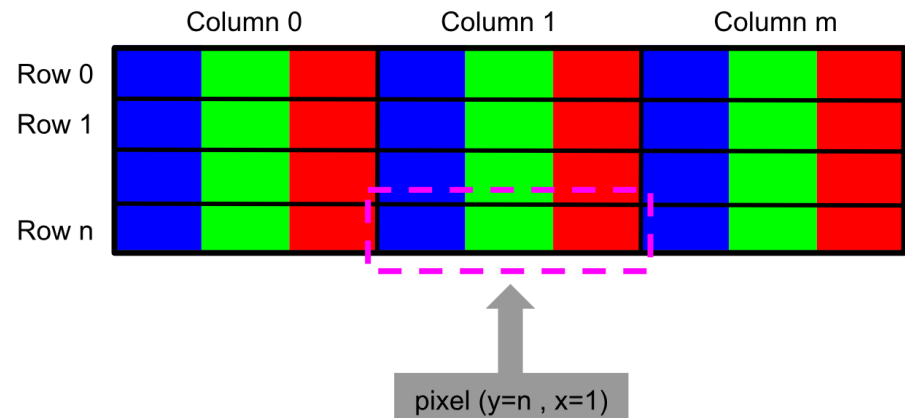
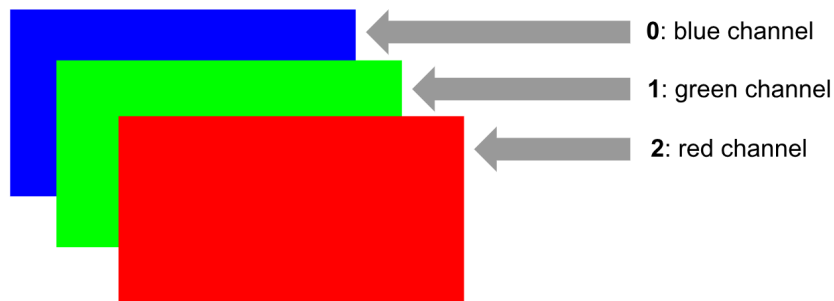
RGBA



[https://jdhao.github.io/2019/03/07/pillow\\_image\\_alpha\\_channel/](https://jdhao.github.io/2019/03/07/pillow_image_alpha_channel/)  
<https://pythontic.com/image-processing/pillow/alpha-composite>

Atenção: Se desejar separar os canais de cor via OpenCV, entenda que o OpenCV usa o sistema BGR ao invés do RGB.

```
blue,green,red = cv2.split(img)
```

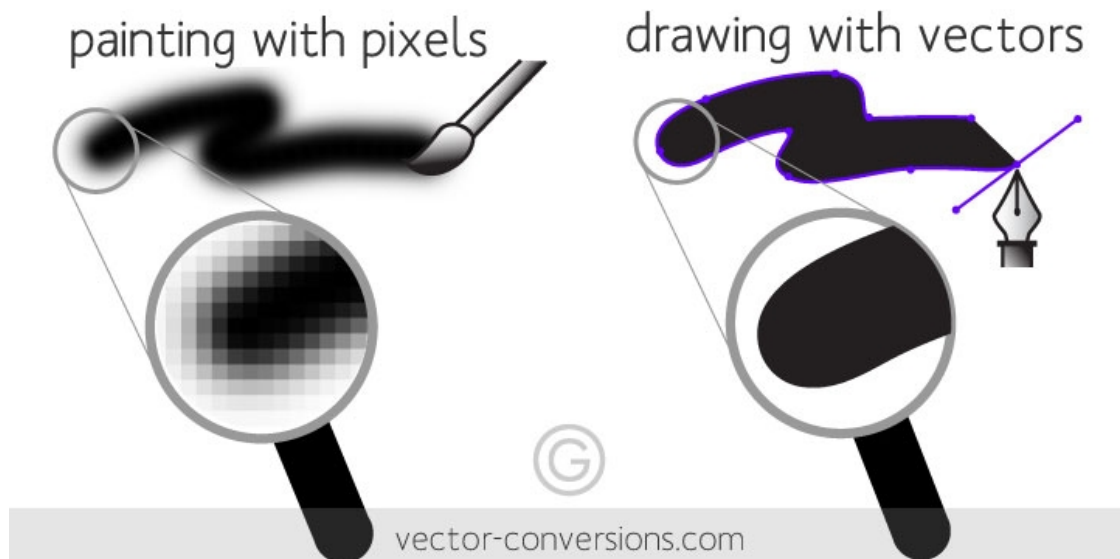


## Formato direto em Pixels versus formato vetorial:

O formato de imagens que utilizamos em PIM são do tipo “Raster images”, ou seja, são matrizes de pixels.

Como um pixel é o menor elemento em um dispositivo de exibição. Se você continuar fazendo “zoom” em uma imagem raster, em algum momento ela ficará “pixelizada”.

Imagens vetoriais são criadas por cálculos matemáticos de um ponto a outro que formam linhas e formas. Ao ampliar um gráfico vetorial ele sempre terá a mesma aparência, pois as fórmulas matemáticas permanecem as mesmas, renderizando o mesmo gráfico visual, independentemente do tamanho sem perder a qualidade.





Cameras CMOS Canon:

<https://canon-cmos-sensors.com/canon-35mmfhdxsca-19um-cmos-sensor/>  
<https://canon-cmos-sensors.com/cmos-sensors/>

Material sobre especificação de câmeras no site:

<http://guide.directindustry.com/choosing-the-right-machine-vision-camera/>

Também há **sites** de fornecedores que ajudam na **escolha da câmera** em função de suas necessidades:

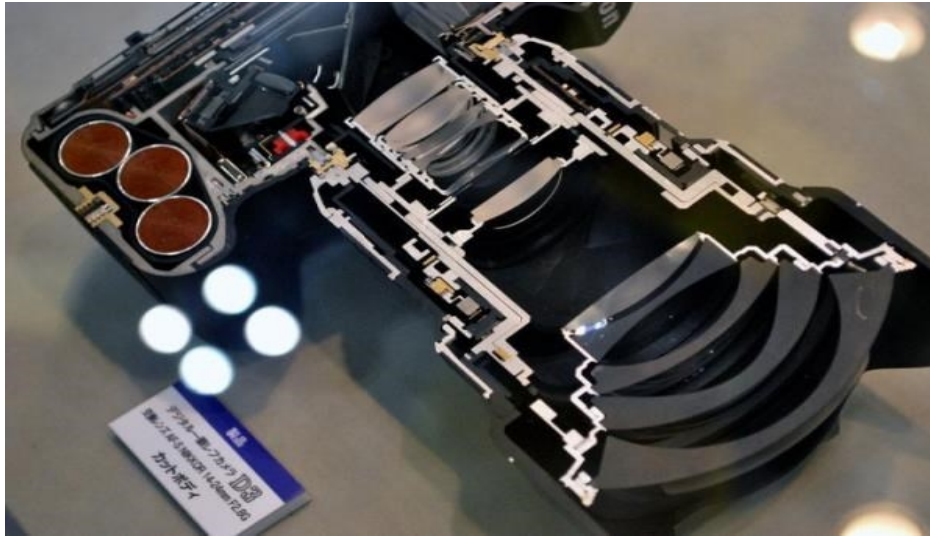
Exemplo:

<https://www.cognex.com/resources/interactive-tools/lens-advisor>

<https://thinklucid.com/tech-briefs/understanding-digital-image-sensors/>

<http://howthingswork.org/electronics-how-digital-camera-works/>

## Geometria de formação da imagem



As estruturas de câmeras reais apresentam componentes bastante complexos, por exemplo: obturador/abertura (controlam a exposição à luz) e conjuntos de lentes (as quais contribuem com qualidade da imagem, foco, tratamento de aberrações cromáticas, zoom, etc...).

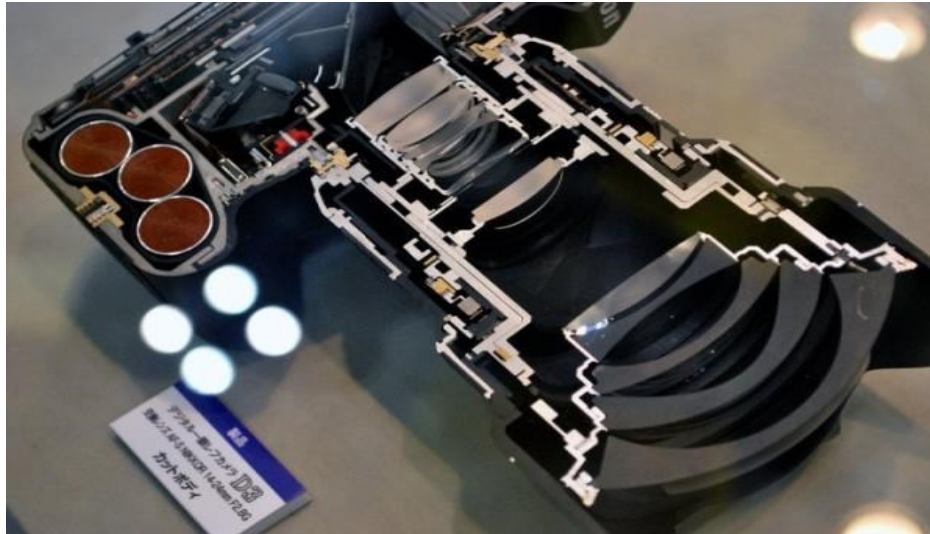
Por outro lado, desejamos uma câmera mais simples, que guarde apenas as características essenciais e possa gerar um modelo mais intuitivo nos permitindo discussões sobre a geometria de formação de imagens. Essa câmera desejada é a chamada “câmera pinhole”.

Quem tiver interesse em descrições sobre lentes pode procurar em:

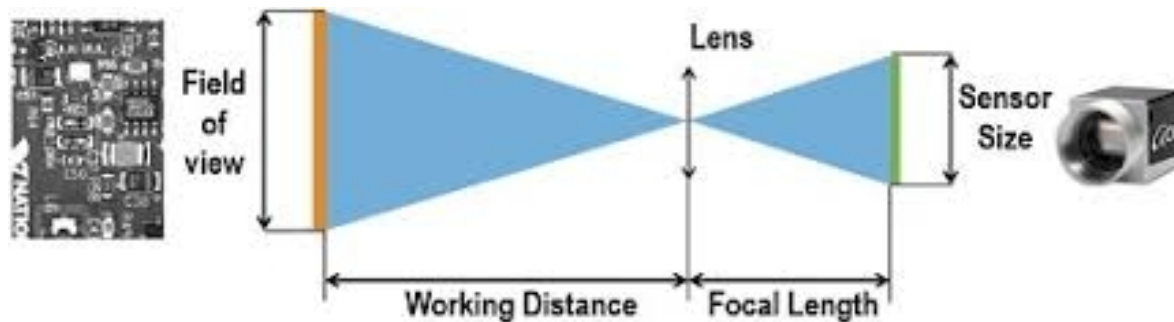
<https://www.khanacademy.org/computing/pixar/virtual-cameras/depth-of-field/v/optics7-final>

<https://www.stemmer-imaging.com/fr-fr/donnees/aberrations-of-optics/>

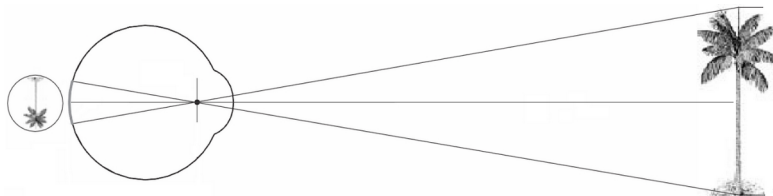
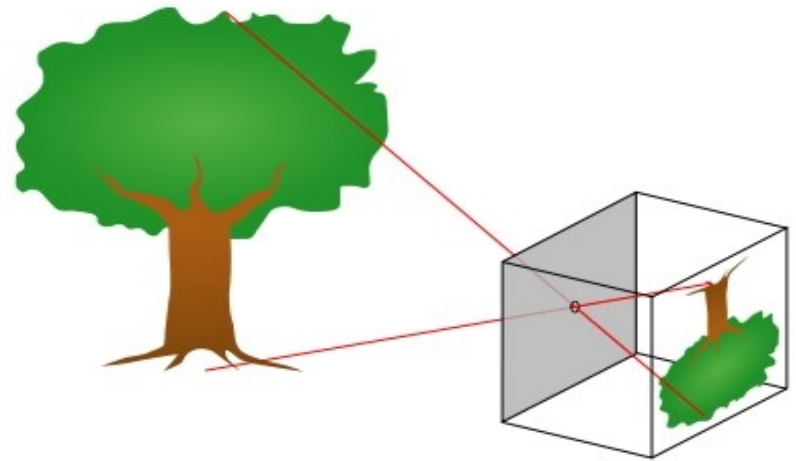
<https://oneslidephotography.com/facts-and-myths-about-camera-lenses/>



**Simplificando:**  
**Câmera de orifício (*Pinhole*)**

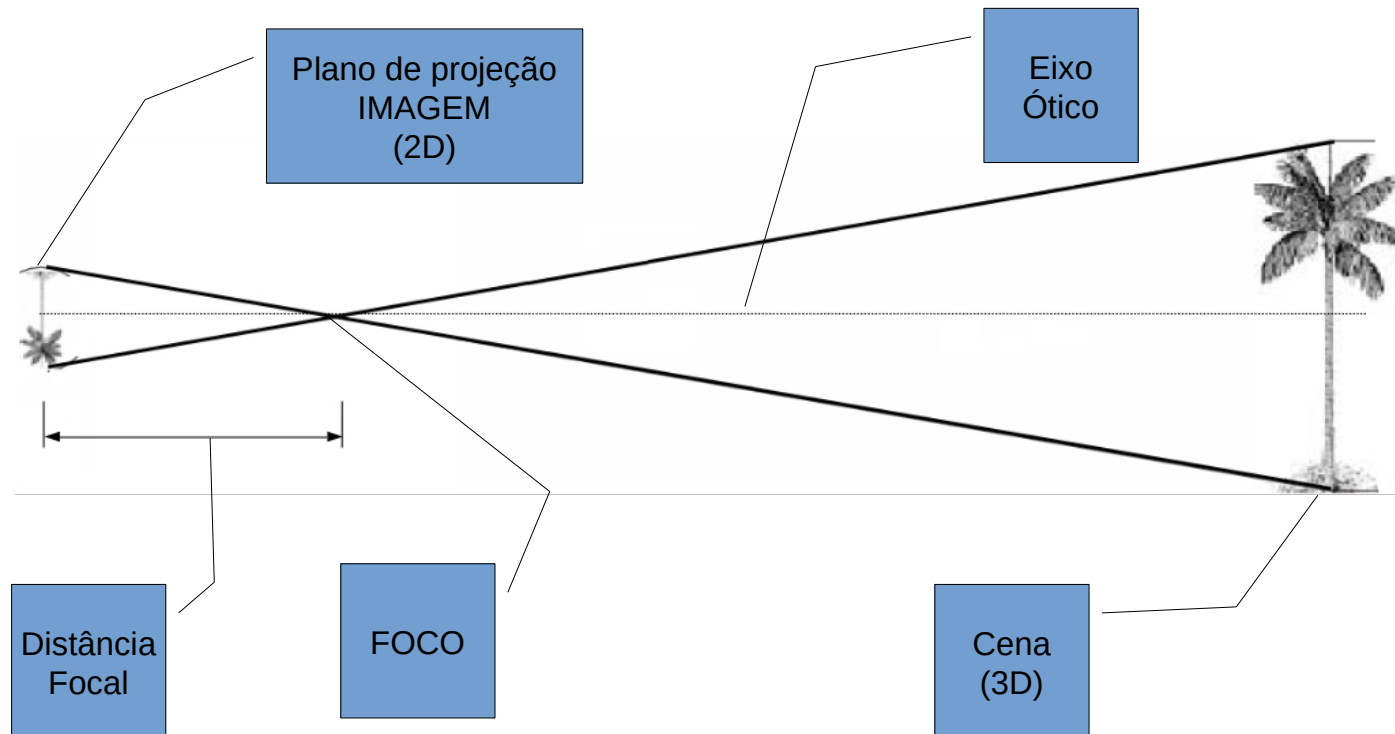


A geometria de formação de imagem será analisada por meio na câmera de orifício (Pinhole). Mas, antes de entrar nesse tópico, iremos discutir um pouco Sobre a amostragem, quantização e digitalização

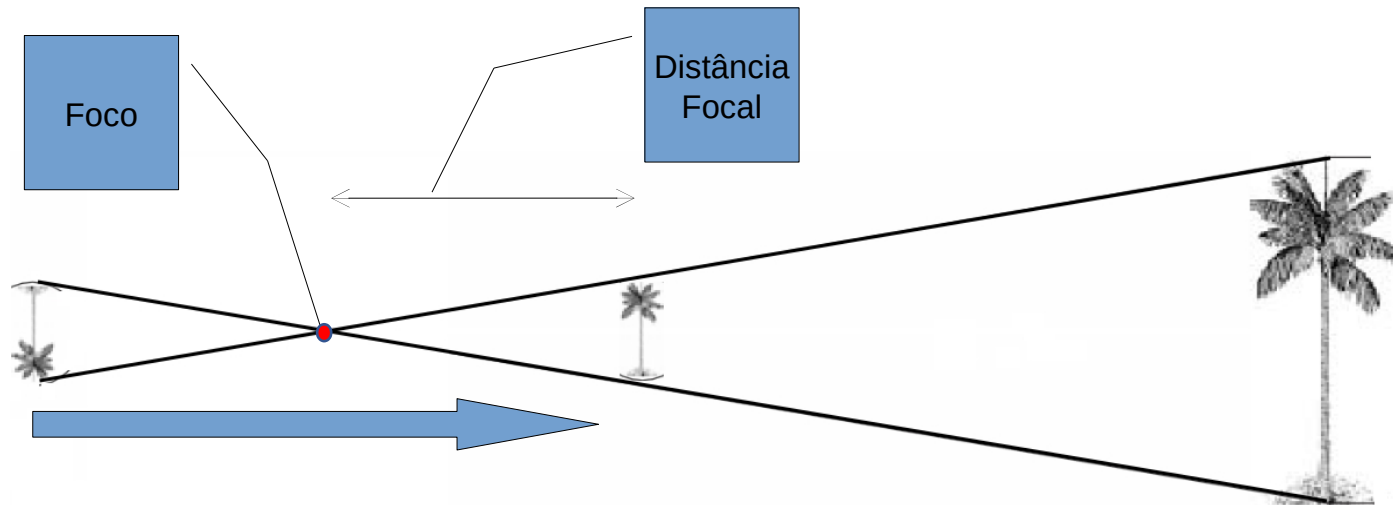


GEOMETRIA DE FORMAÇÃO DA  
IMAGEM  
COM UM SISTEMA DE COORDENADAS  
NA CÂMERA

# A geometria da formação de imagem na retina é similar à da câmera de orifício (Pinhole)

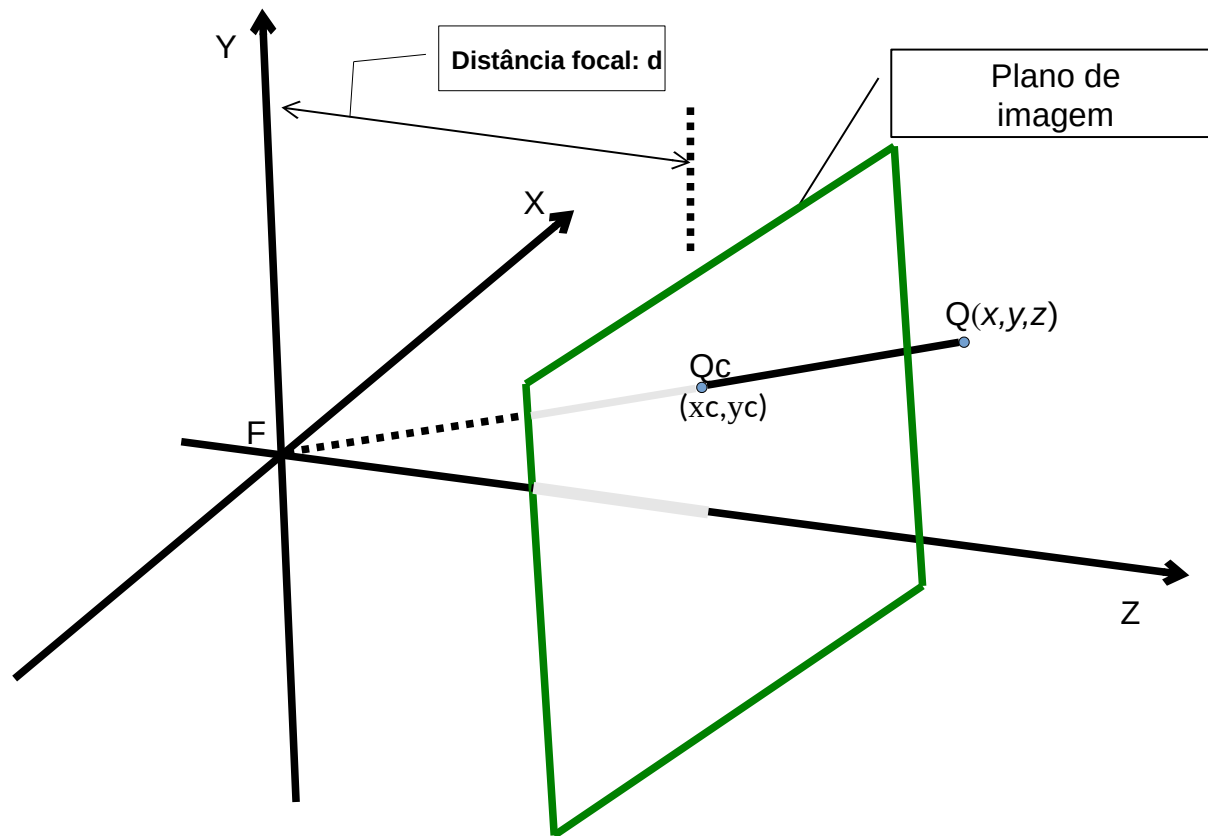


# Formação de imagem na câmera Pinhole



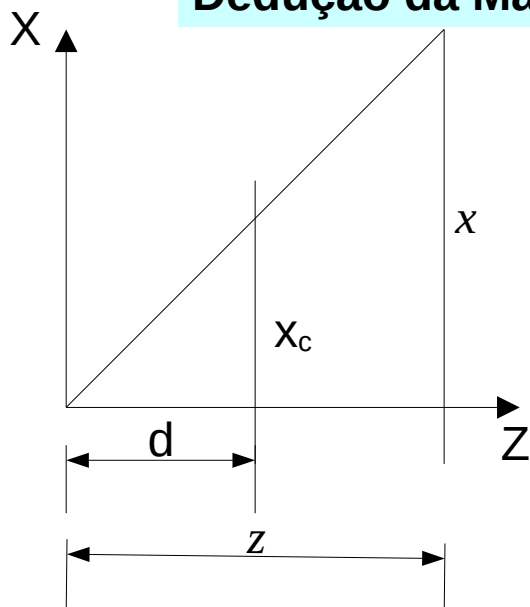
Por conveniência matemática, movemos o plano de imagem para a frente do foco.  
Isso permite uma modelagem mais simples.

**Projeção perspectiva: plano de projeção deslocado para a frente do centro de projeção sobre o eixo +Z**



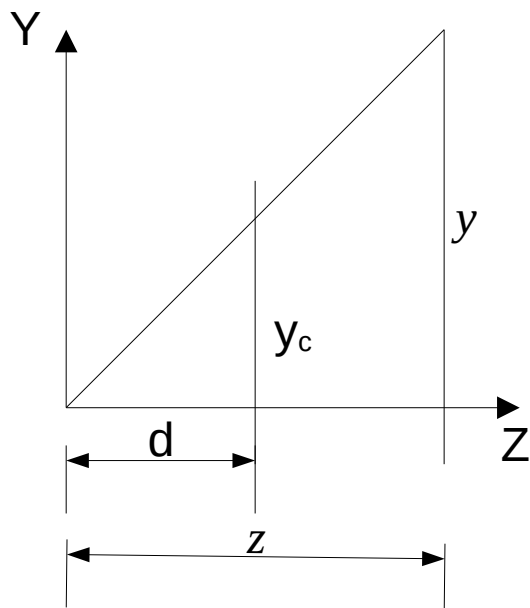


## Dedução da Matriz de projeção perspectiva

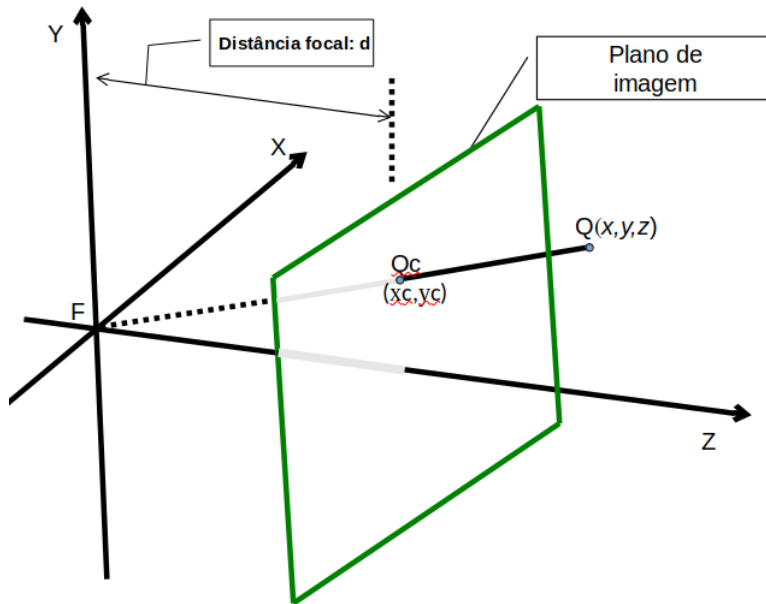


$$\frac{x_c}{d} = \frac{x}{z} \Rightarrow x_c = \frac{x}{z/d}$$

$$\frac{y_c}{d} = \frac{y}{z} \Rightarrow y_c = \frac{y}{z/d}$$



**A) Matriz de projeção perspectiva considerando:**  
Plano de projeção normal ao eixo Z a uma distância  $d$  em relação a origem.



$$M_{per} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/d \end{bmatrix}$$

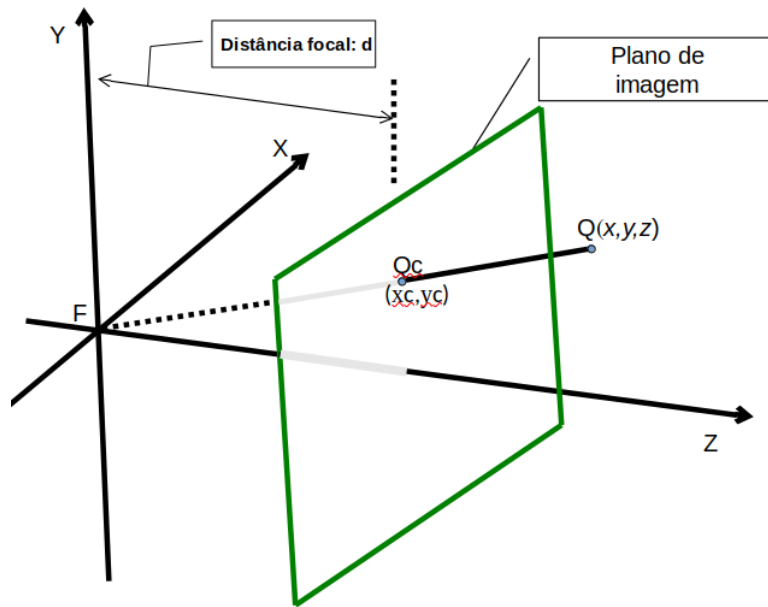
$x_c$  e  $y_c$ :

são coordenadas da projeção de Q sobre a imagem.

estão em unidades métricas iguais às de Q (milímetros, por exemplo)

### A) Matriz de projeção perspectiva considerando:

Plano de projeção normal ao eixo Z a uma distância  $d$  em relação a origem.



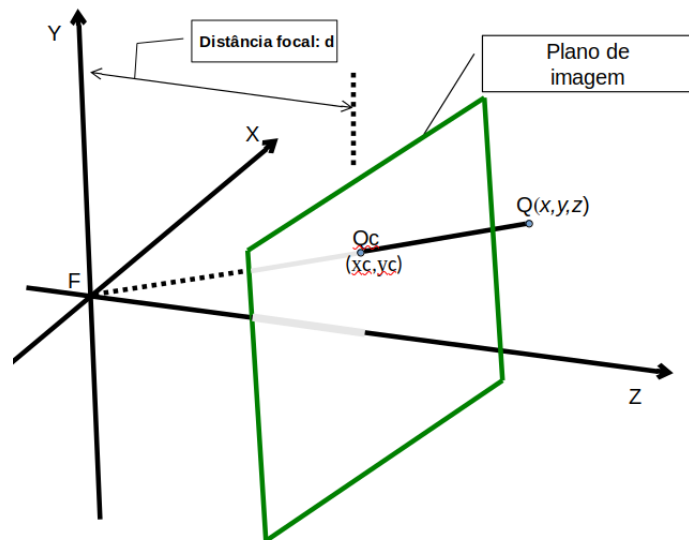
$Q_h$  é o ponto no espaço homogêneo, com  $w=z/d$

$Q_c$  é o ponto no espaço cartesiano =  $Q_h/w$

$Q_p$  é o ponto em pixels

$$Q_h = M * Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/d \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z/d \end{bmatrix}$$

## A) Exemplo



$Q_h$  é o ponto no espaço homogêneo, com  $w=100$

$Q_c$  é o ponto no espaço cartesiano =  $Q_h/w$

$Q_p$  é o ponto em pixels

$$d=5\text{mm} \quad Q = \begin{bmatrix} 150,75 \\ 1500 \\ 500 \end{bmatrix} \text{mm}$$

$$Q_c = \begin{bmatrix} ? \\ ? \end{bmatrix} \text{mm} \quad Q_p = \begin{bmatrix} ? \\ ? \end{bmatrix} \text{pixels}$$

$$Q_h = M * Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/d \end{bmatrix} * \begin{bmatrix} 150,75 \\ 1500 \\ 500 \end{bmatrix} = \begin{bmatrix} 150,75 \\ 1500 \\ 100 \end{bmatrix} = \begin{bmatrix} x_h \\ y_h \\ w \end{bmatrix}$$

$$Q_c = \begin{bmatrix} x_h/w \\ y_h/w \\ w/w \end{bmatrix} = \begin{bmatrix} 1,51 \\ 15 \\ 1 \end{bmatrix} \text{mm}$$

$$Q_p = ???$$

O modelo representado na matriz  $M$  é restrito a um único sistema de referência de coordenadas.

A operação realizada pela matriz  $M$  não fornece pontos em pixels, fornece em unidades métricas (ex: cm, mm) as quais não estão intuitivamente associadas a conceitos tais como o de resolução;

Por outro lado, a imagem é uma matriz de pixels (sensor CCD ou CMOS)

As coordenadas em pixels são referenciadas no canto superior esquerdo da matriz de pixels do sensor, ao passo que as coordenadas fornecidas pela matriz  $M$  referem-se a coordenadas em relação ao centro do sensor;

Discutiremos o aprimoramento dessa modelagem em outro momento...

Mas, apesar da simplicidade a modelagem de projeção fornecida pela matriz  $M$  é útil para certas aplicações, um exemplo pode ser encontrado na 9ª questão da 1ª parte da primeira lista de exercícios.

Segue...