

**Processamento de Sinais  
Aplicado à Engenharia Mecânica**

**Profs.  
Flavius P. R. Martins  
Flávio Celso Trigo**

## Bibliografia

- Gonzalez, R.C., Woods, R.E., Digital Image Processing, 3<sup>rd</sup> ed. Prentice Hall, 2008.
- Parker, J. R., Algorithms for Image Processing and Computer Vision, 2<sup>nd</sup> ed. John Wiley Sons, 2011.
- Sonka, M., Hlavac, V., Boyle, R, Image Processimg, Analysis and Machine Vision. International Student Edition, 2008.
- Forsyth, D.A., Ponce, J., Computer Vision: A Modern Approach, 2nd ed. Pearson, 2012.
- Davies, E. R., Machine Vision: Theory, Algorithms and Practicalities 3rd ed. Academic Press, 2005.
- Russ, J. C., The Image Processing Handbook, 6<sup>th</sup> ed. CRC Press, 2011.
- Hartley, R., Multiple View Geometry in Computer Vision. Cambridge University Press, 2000.
- Faugeras, O., Three Dimensional Computer Vision: A Geometric Viewpoint. MIT Press, 2001.

# **Bibliografia**

## **Revistas periódicas importantes**

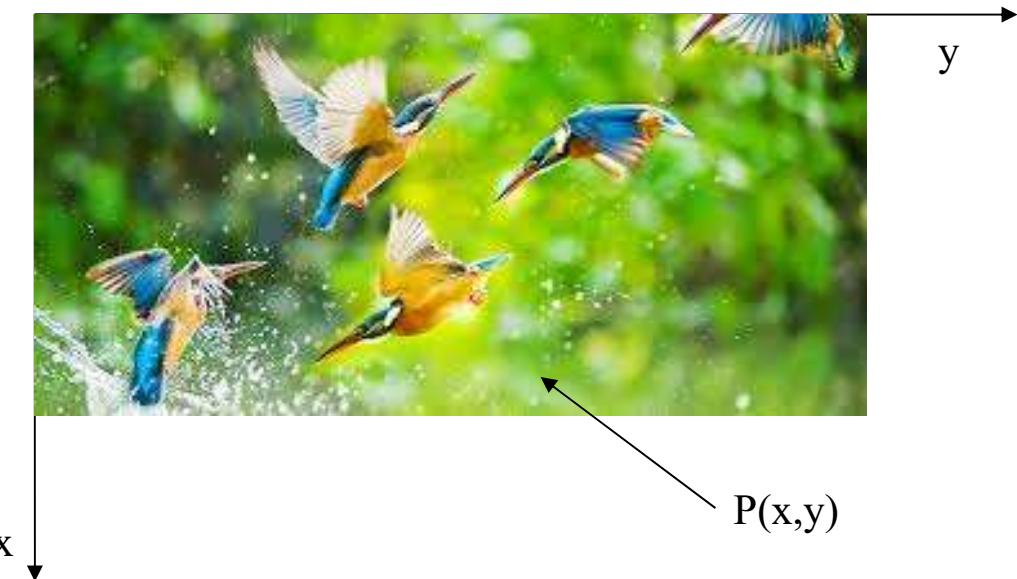
- IEEE Transactions on Image Processing.
- International Journal of Computer Vision
- IEEE Transactions on Pattern Analysis and Machine Intelligence.
- Pattern Recognition.
- Computer Vision and Image Understanding.
- Machine Vision and Applications.
- Medical Image Analysis.
- ISPRS Journal of Photogrammetry and Remote Sensing.
- Signal Processing.
- Pattern Recognition Letters.
- Journal of Visual Communication and Image Representation.

# **Bibliografia**

## **Ferramentas computacionais úteis**

- Matlab Image Processing Toolbox
- OpenCV Library
- Scilab Image Processing Toolbox
- JAI – Java Advanced Imaging
- FFTW Library

# Imagen: conceitos preliminares



Imagens comuns:  $P(x,y)$  é uma função de duas variáveis, representando uma 'intensidade luminosa'.

Seqüência de imagens:  $P(x,y,t)$  é uma função de três variáveis, representando uma coleção de imagens capturadas durante um certo intervalo de tempo.

Imagens coloridas ou multi-espectrais: são descritas por uma função vetorial  $\mathbf{f}(x,y) = (f_1(x,y), f_2(x,y), \dots, f_n(x,y))$ .

Imagens não se limitam a representar cenas do mundo visível; podem também expressar distribuições bidimensionais de outras quantidades físicas, como, por exemplo, temperaturas, pressões e distâncias relativas ao observador.

# Imagen – exemplos



Imagen teste: “Lena”



Imagen de impressora  
telegráfica (1921)



Imagen teste:  
“The camera man”



Imagen médica: ultrassom

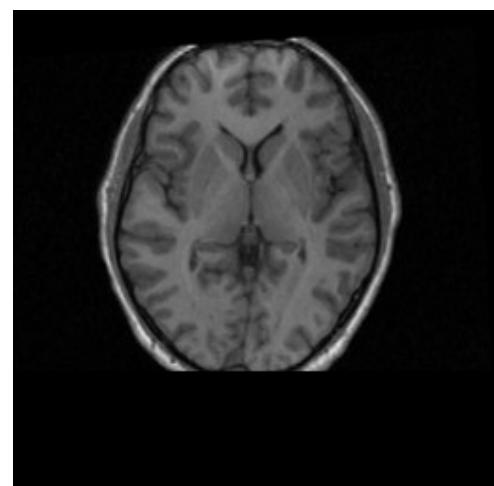


Imagen médica: ressonância magnética  
(Teorema de Radon)

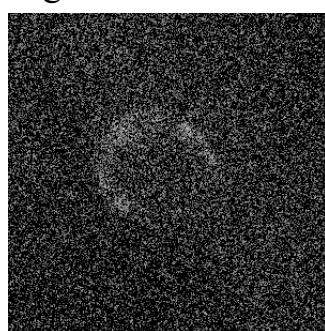


Imagen astronómica: Cygnusloop  
(ultravioleta extrema)

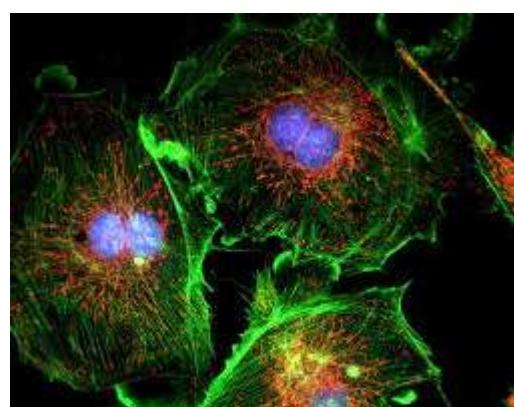


Imagen de fluorescência

# Imagen - exemplos

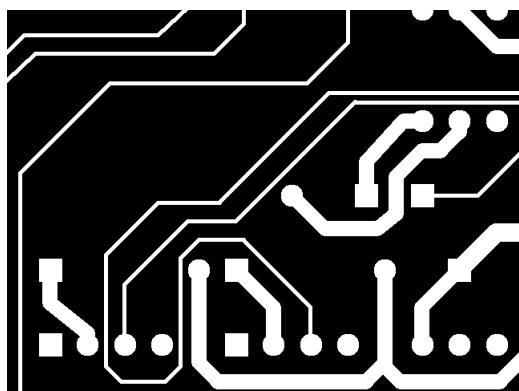


Imagen binaria: artefato industrial



Mapa de profundidade



Seqüência temporal de imagens

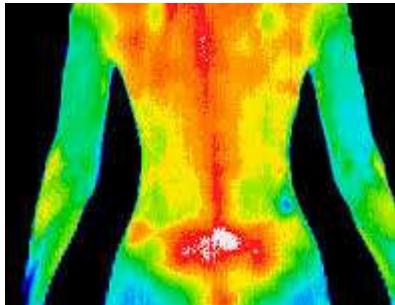


Imagen termográfica  
(uso de falsas cores)



Imagen colorida RGB



Imagen de cena  
sintetizada artificialmente

# Imagen - exemplos



verde



vermelho

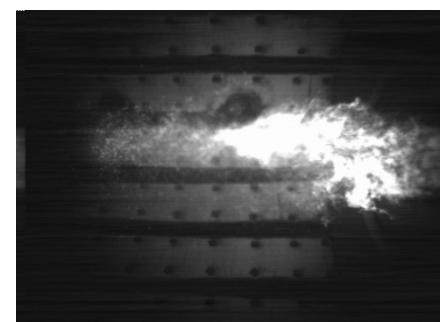
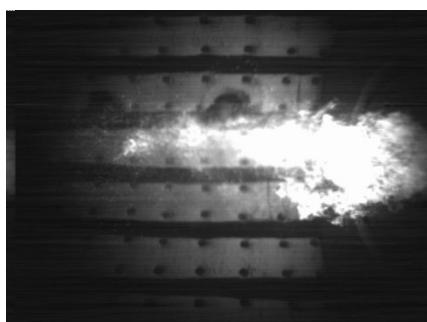
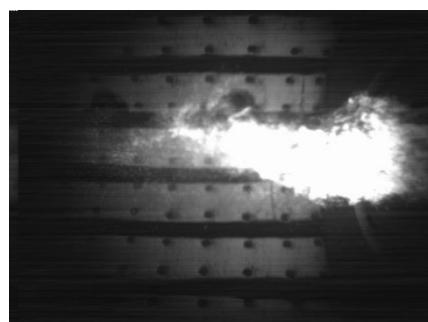
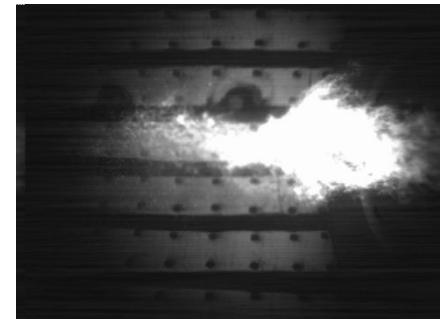
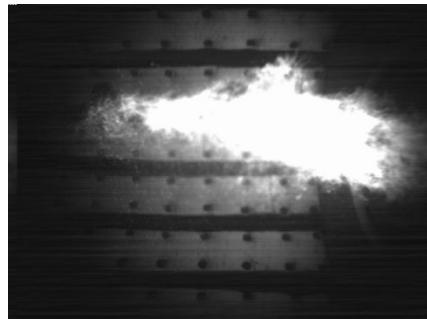
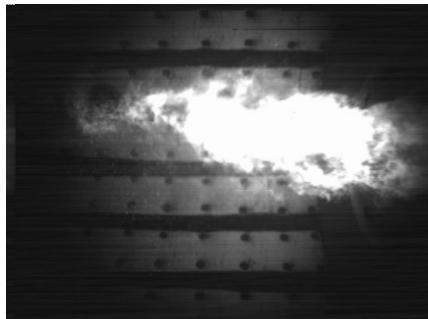


Infravermelho próximo

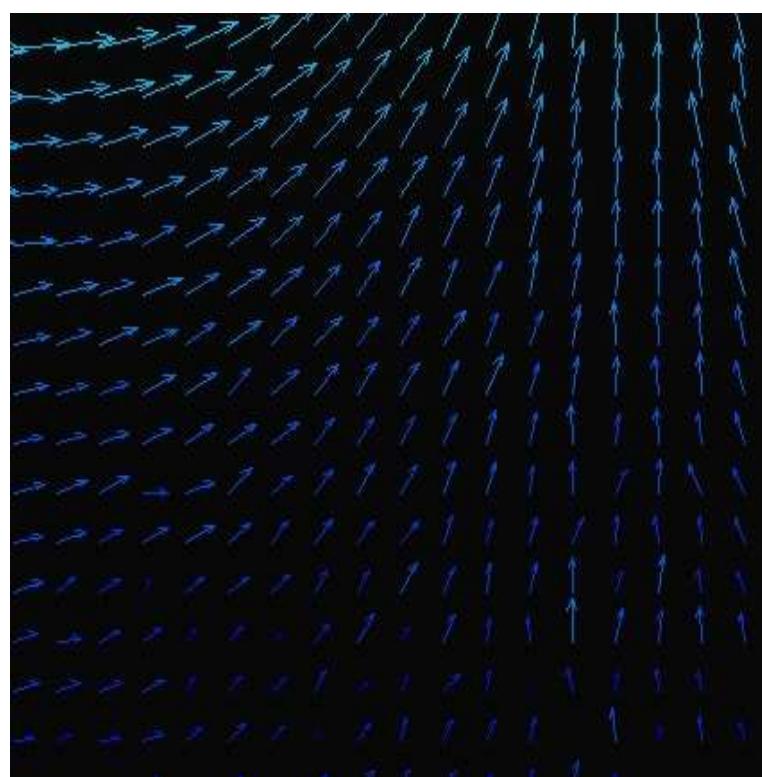


Imagen composta, usando falsas cores

## Imagen - exemplos

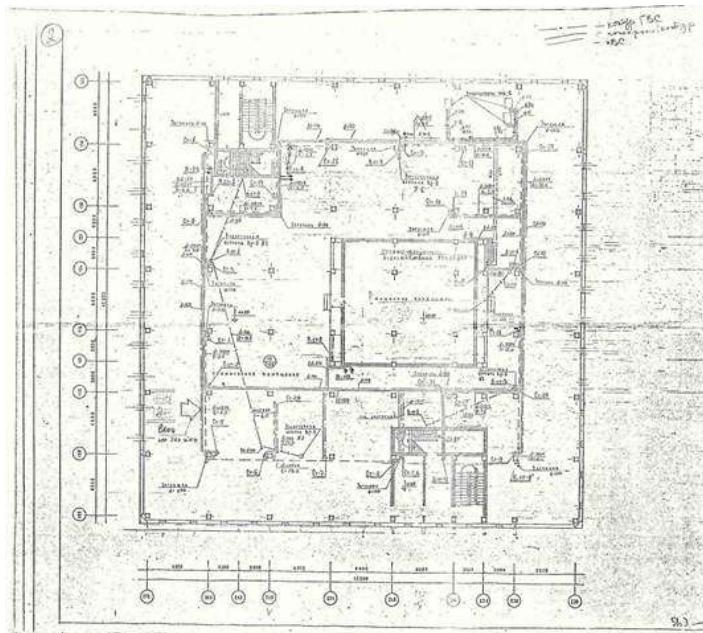


Seqüência temporal de imagens de chama – processo estocástico



Campo de velocidades: pode ser representado como uma função complexa

# Imagen versus desenhos



Desenho de engenharia digitalizado

S Domingos, 10 - 8 - 68 / GUINÉ'

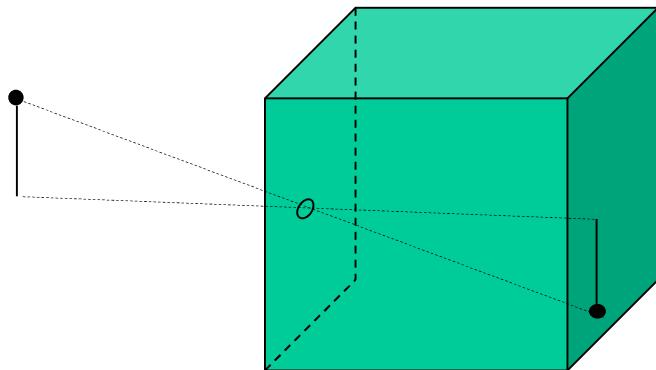
O Aly. Nacido, saídos, rasga a terra  
para retirar uma estrada coloca de  
um artado S.D. → Sussua, e meio  
caminho para NHAMBALI, a cerca  
de 5 Km's de S.Domingos. A estrada  
estava mal montada:  
1º os deitores das ruas que lhe nôdo  
colocados.  
2º o favoriad outros paralelo à artado  
que vê de outras toritonal.  
Foi detectado pelo proceder da  
C.A.T. 1744. A observar o de  
C.A.P. Cardoso C.C.S.

Texto manuscrito digitalizado

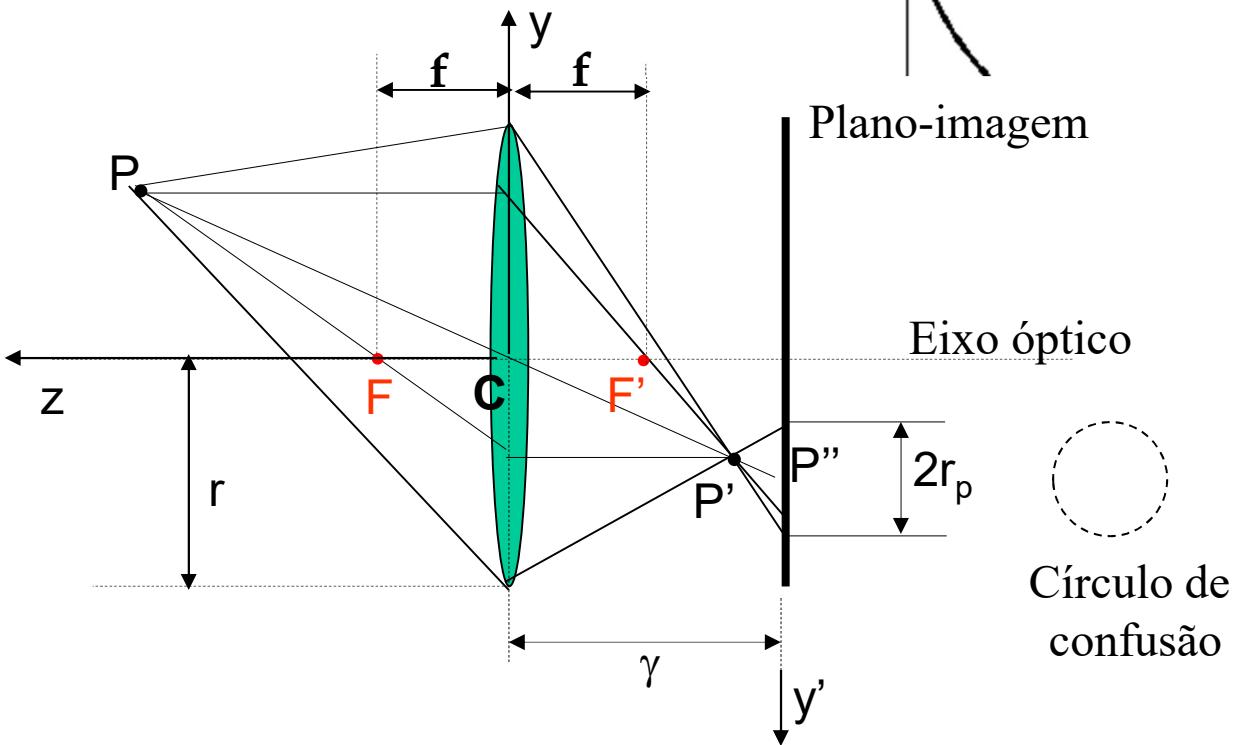
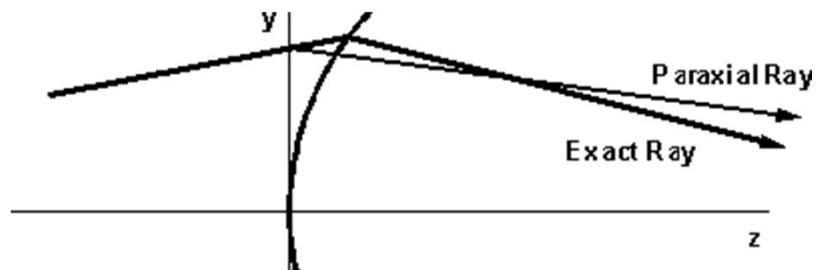
# Modelo opto-geométrico de uma câmera

## Câmera estenopéica ('pinhole')

Nesse modelo de câmera, a imagem de um objeto corresponde à sua projeção a partir de um pequeno furo instalado em uma câmera escura.



**Óptica de uma lente delgada**  
Faz-se a aproximação paraxial.



# Modelo opto-geométrico de uma câmera

## Equações de uma lente delgada

$$\frac{y}{|CP|} = -\frac{y'}{|CP'|}$$

$$\frac{1}{|CP|} + \frac{1}{|CP'|} = \frac{1}{f}$$

$$\begin{cases} x' = -f \frac{x}{z-f} \\ y' = -f \frac{y}{z-f} \\ z' = -f \frac{z}{z-f} \end{cases}$$

$m$  = fator de amplificação

$$m = \frac{f}{z} = \frac{z''}{f}$$

Para que a imagem seja bem focalizada, deve-se ter:

$$\gamma = -z'$$

O raio e o centro do círculo de confusão podem ser determinados a partir das equações a seguir:

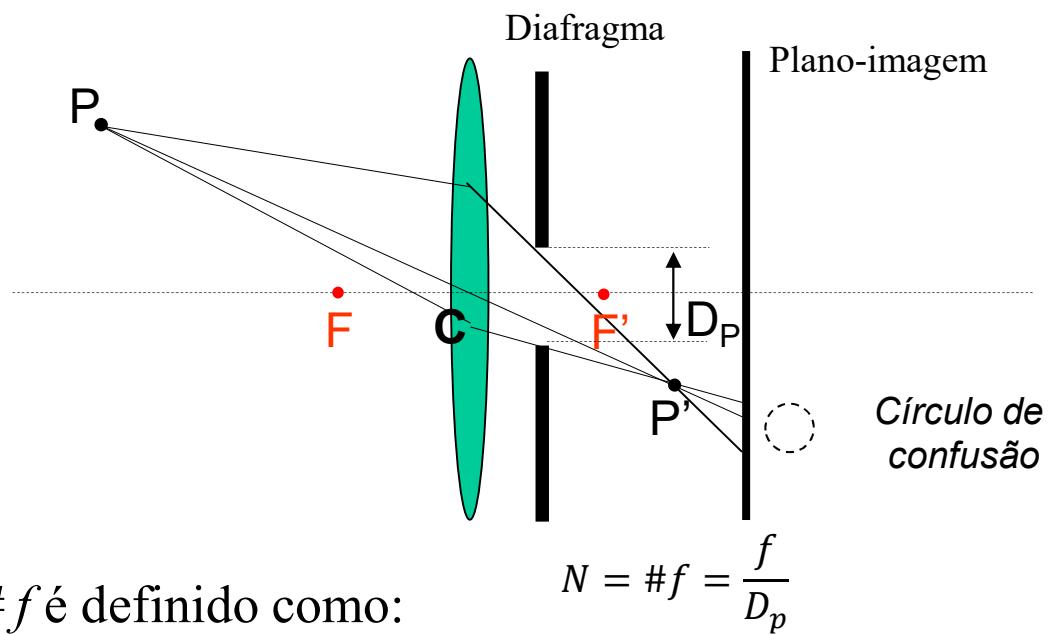
$$r_p = r \cdot \left[ 1 + \gamma \frac{f-z}{f \cdot z} \right]$$

$$\begin{cases} x'' = -\frac{\gamma x}{z} \\ y' = -\frac{\gamma y}{z} \\ z' = -\gamma \end{cases}$$

# Modelo opto-geométrico de uma câmera

## Diafragma / número $f$ / valor de exposição

A desfocalização pode ser atenuada reduzindo-se a abertura da lente por meio de um diafragma devidamente posicionado.



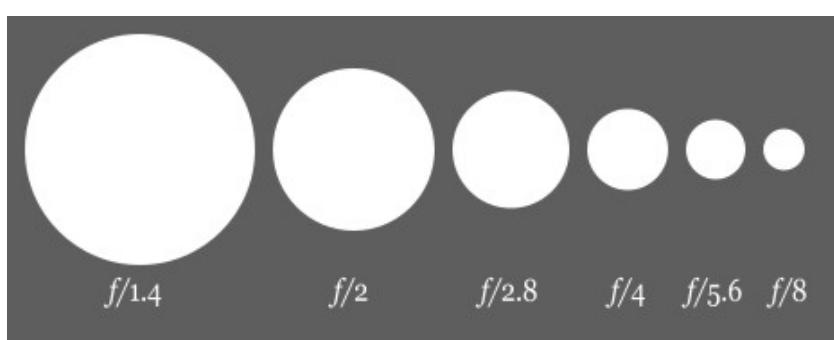
O  $\#f$  é definido como:

O valor de exposição ( $EV$ ) é

$$EV = \log_2 \frac{N^2}{t}$$

$t$  = tempo de ‘shutter’ (de 1s a 1  $\mu$ s)

Aumentando-se o  $\#f$  (com  $f$  fixo e  $D$  decrescente) diminui-se a luminosidade da imagem da cena. Para um valor fixo  $EV$  deve-se aumentar o tempo de *shutter*.



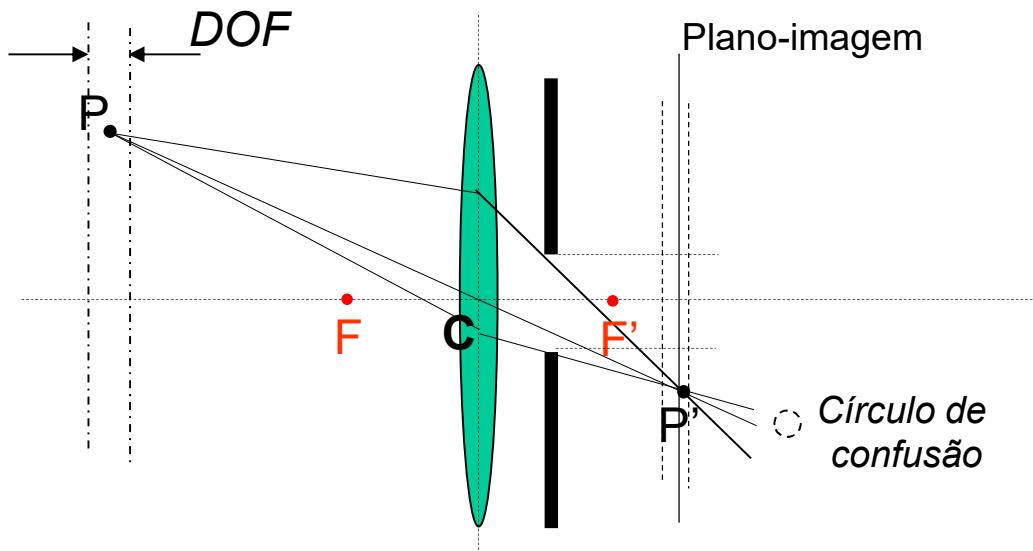
$\#f < 2,8$ : lentes rápidas  
(cenas com pouca luz)

$\#f > 2$ : lentes lentas

# Modelo opto-geométrico de uma câmera

## Profundidade de campo (*DOF*)

É o comprimento da zona em torno do objeto dentro da qual se produz uma imagem aparentemente focalizada.



*DOF* decresce quando

- se aumenta o fator de amplificação, aproximando-se a lente do objeto e mantendo-se  $\#f$  fixo;
  - se aumenta o fator de amplificação aumentando-se a distância focal e mantendo-se  $\#f$  fixo;
  - se diminui o  $\#f$  aumentando-se a abertura  $D_p$ .
- Ações opostas às anteriores incrementam *DOF*.

# Modelo opto-geométrico de uma câmera

## Distância hiperfocal

É a distância além da qual todos os objetos produzem imagens razoavelmente bem focalizadas. Sendo

$2r_p$  = diâmetro máximo admissível para o círculo de confusão,

$D$  = diâmetro do diafragma da lente

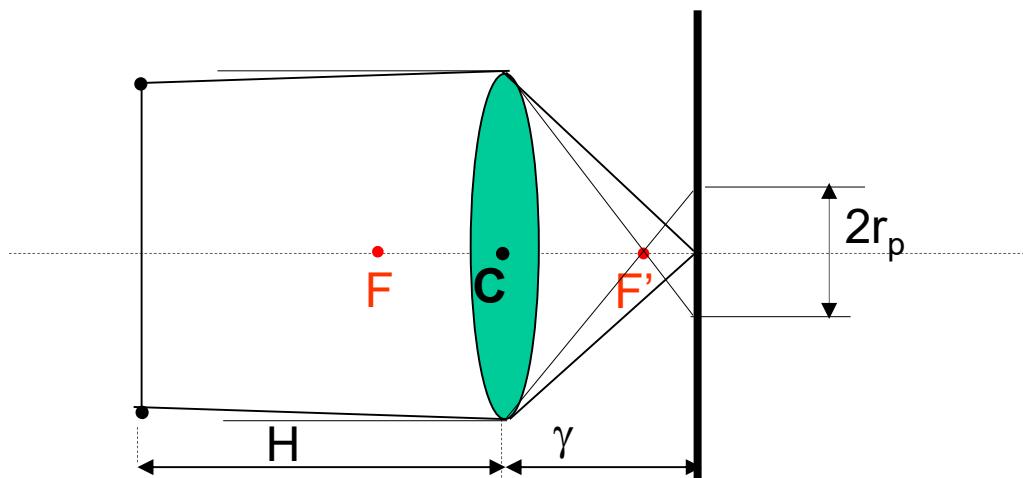
$N = \#f$  da lente

$f$  = distância focal,

$H$  = distância hiperfocal da lente,

tem-se:

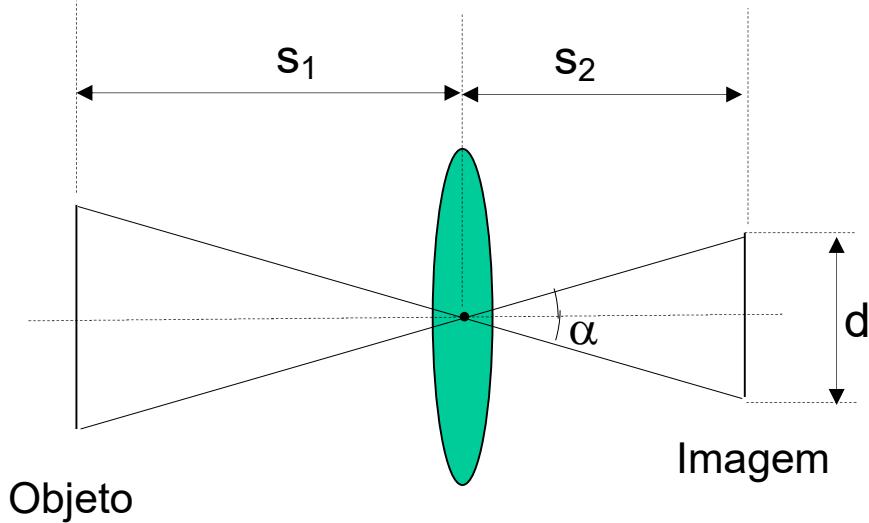
$$H = \frac{D + 2r_p}{2r_p} f = \frac{\frac{f}{N} + 2r_p}{2r_p} f = \frac{f^2}{2Nr_p} + f \approx \frac{f^2}{2Nr_p}$$



# Modelo opto-geométrico de uma câmera

## Campo de visão

É a extensão angular da cena totalmente observada no plano imagem da câmera, ou seja, correspondente à dimensão do objeto tal que sua imagem ocupe toda a extensão do sensor.



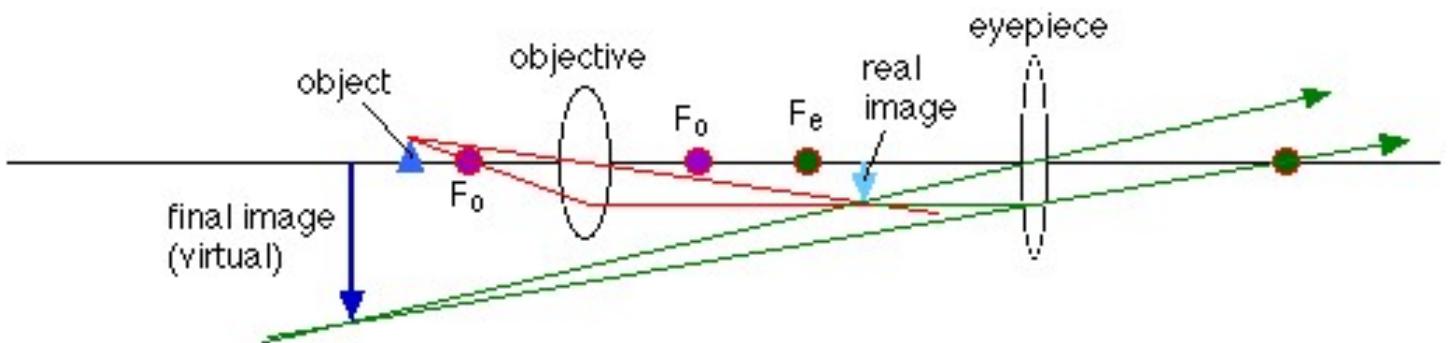
$$\alpha = 2 \arctan \frac{d}{2s_2}$$

# Modelo opto-gemétrico de uma câmera

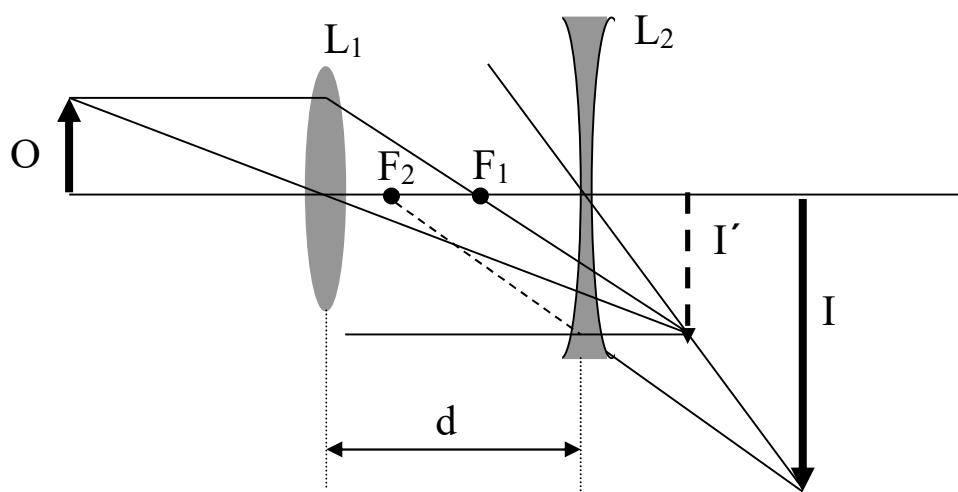
## Composição de lentes

Combinam-se, de acordo com o seguinte princípio: a imagem produzida pela lente  $L_1$ , com fator  $m_1$ , torna-se objeto para a lente  $L_2$ , com fator de ampliação  $m_2$ . O fator de ampliação do par de lentes, resulta igual a  $m_1 \times m_2$ .

Um microscópio compõe-se de um par de lentes convergentes, em que a objetiva cria uma imagem real e invertida e a ocular gera uma imagem virtual e invertida.



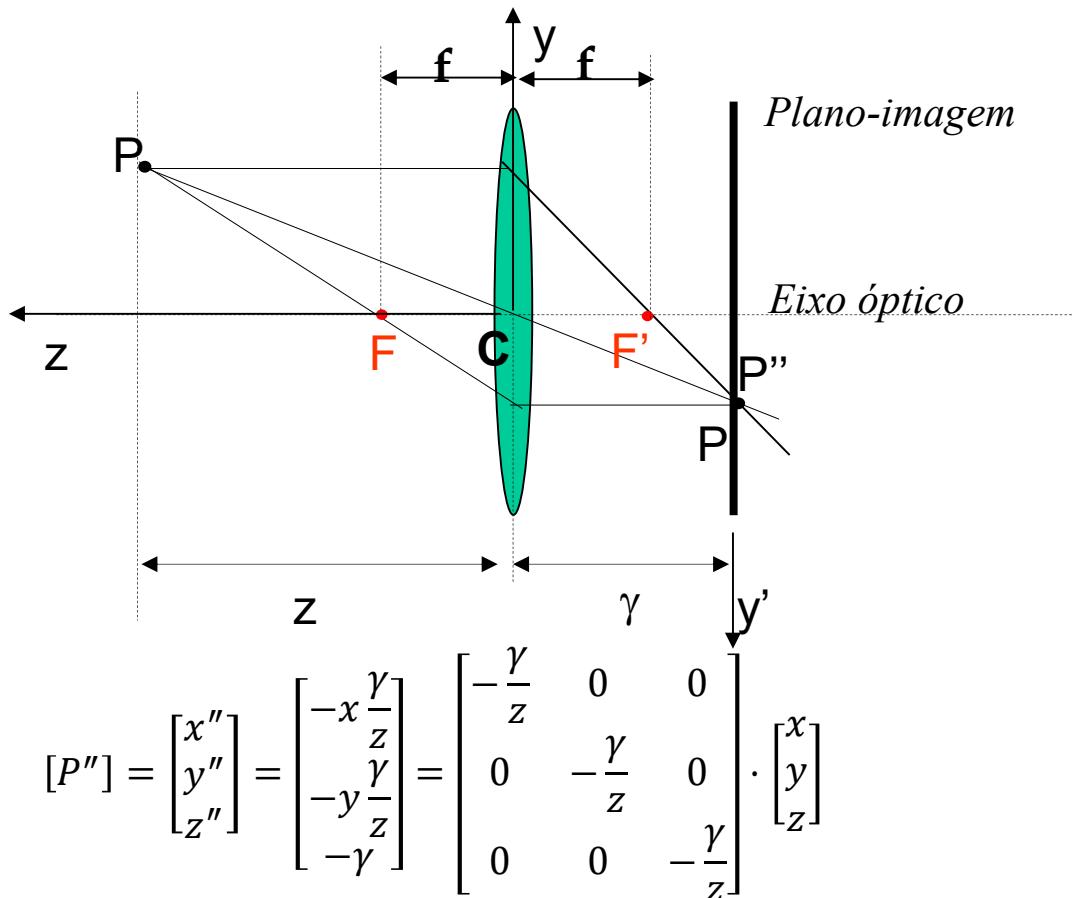
Uma lente *zoom* compõe-se de um par de lentes  $L_1L_2$ , uma convergente, outra divergente. O sistema de lentes  $L_1L_2$  se comporta como uma lente de distância focal variável em função de  $d$ , distância entre  $L_1$  e  $L_2$ .



# Modelo opto-geométrico de uma câmera

## Transformação projetiva

Tomando-se como referência a figura abaixo, tem-se:



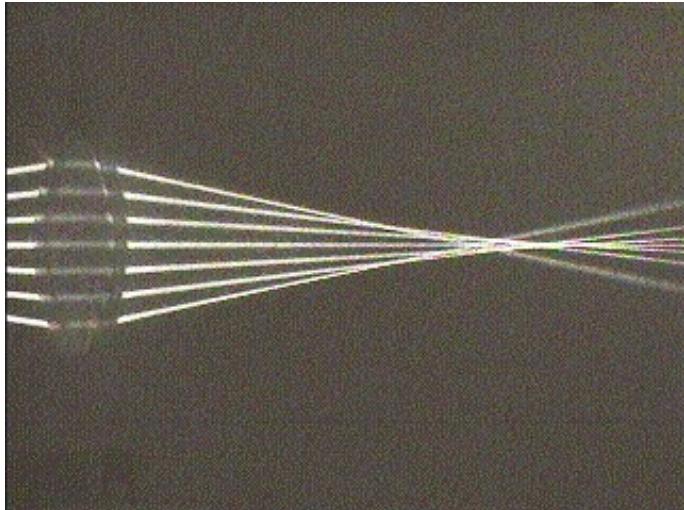
A matriz da transformação projetiva é não-linear em  $z$ . Mediante o uso de coordenadas homogêneas, obtém-se uma transformação linear no espaço quadridimensional:

$$\alpha \cdot \begin{bmatrix} -x'' \\ -y'' \\ 1 \end{bmatrix} = \begin{bmatrix} \gamma & 0 & 0 & 0 \\ 0 & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

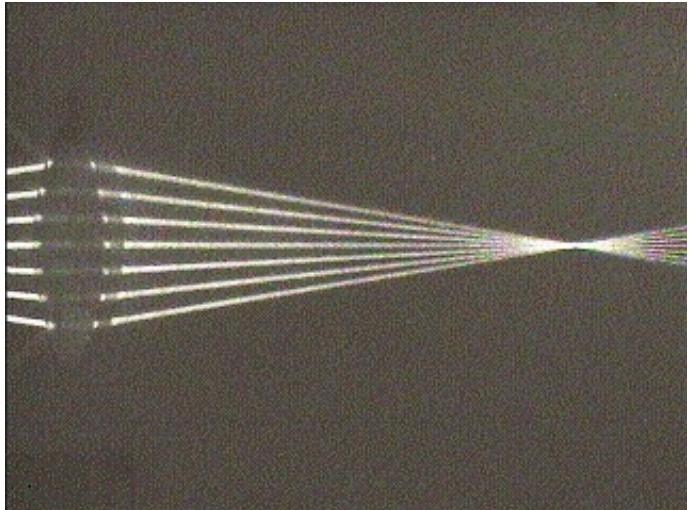
# Modelo opto-geométrico de uma câmera

## Aberraçāo óptica e cromática

Conforme se pode observar nas figuras a seguir, a aberraçāo pode ser eliminada desde que se utilizem lentes não esféricas.

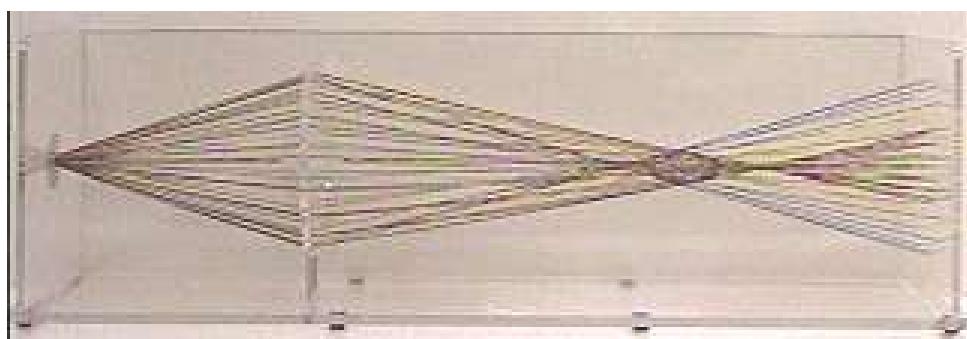


(a) Lente esférica: raios não convergem para o mesmo “foco”



(b) Lente hiperbólica: focalizaçāo ideal

Além da aberraçāo esférica (ilustrada na figura a, acima), lentes esféricas também exibem aberraçāo cromática, ou seja, ondas de diferentes comprimentos são focalizadas em pontos distintos — a banda azul mais próxima do centro óptico que a vermelha.

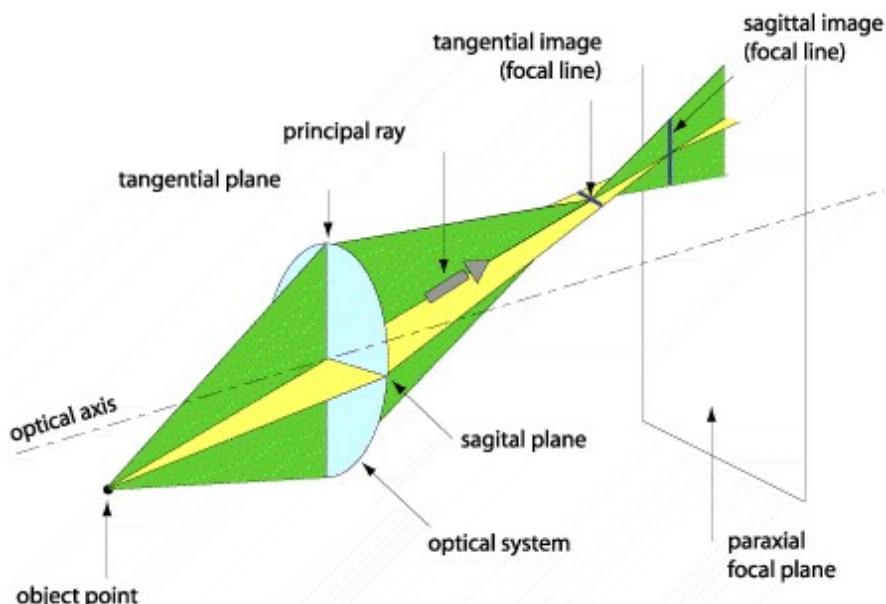


Aberraçāo cromática

# Modelo opto-geométrico de uma câmera

## Astigmatismo

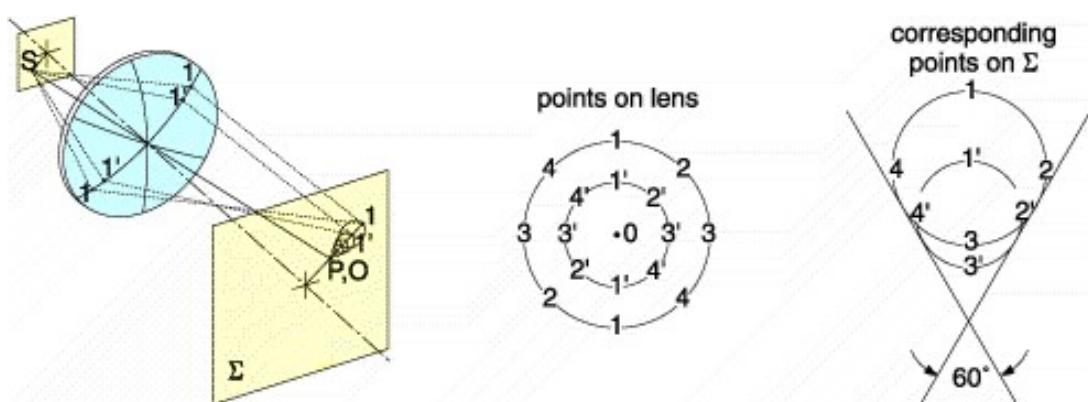
É causado pelo deslocamento lateral do ponto focal em relação ao eixo óptico da lente.



Lente esférica apresentando astigmatismo

## Coma

Deve-se à variação do fator de amplificação com a abertura da lente: a distorção é tanto maior quanto maior a distância ao eixo óptico da lente.

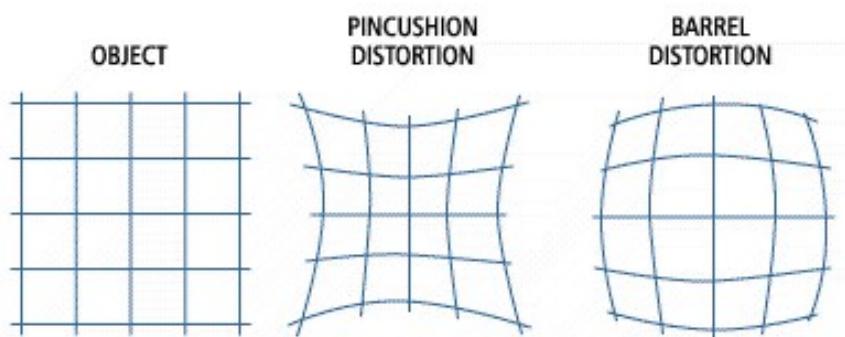


Lente esférica apresentando coma

# Modelo opto-geométrico de uma câmera

## Distorção radial

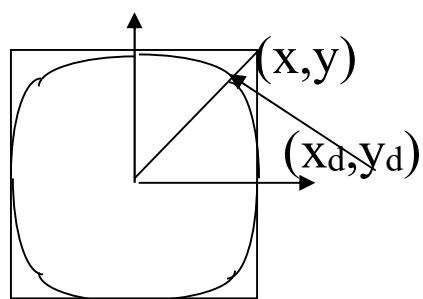
Esse tipo de deformação faz com que um ponto não pertencente ao eixo óptico seja projetado em uma posição diferente daquela prevista pela teoria paraxial. Existem duas formas típicas de distorção radial – ‘pincushion’ e ‘barrel’.



Distorções típicas causadas por lente esférica

Em aplicações de visão computacional que requerem calibração precisa do sistema de captura de imagens, utiliza-se, em geral, o modelo de distorção radial, definido por:

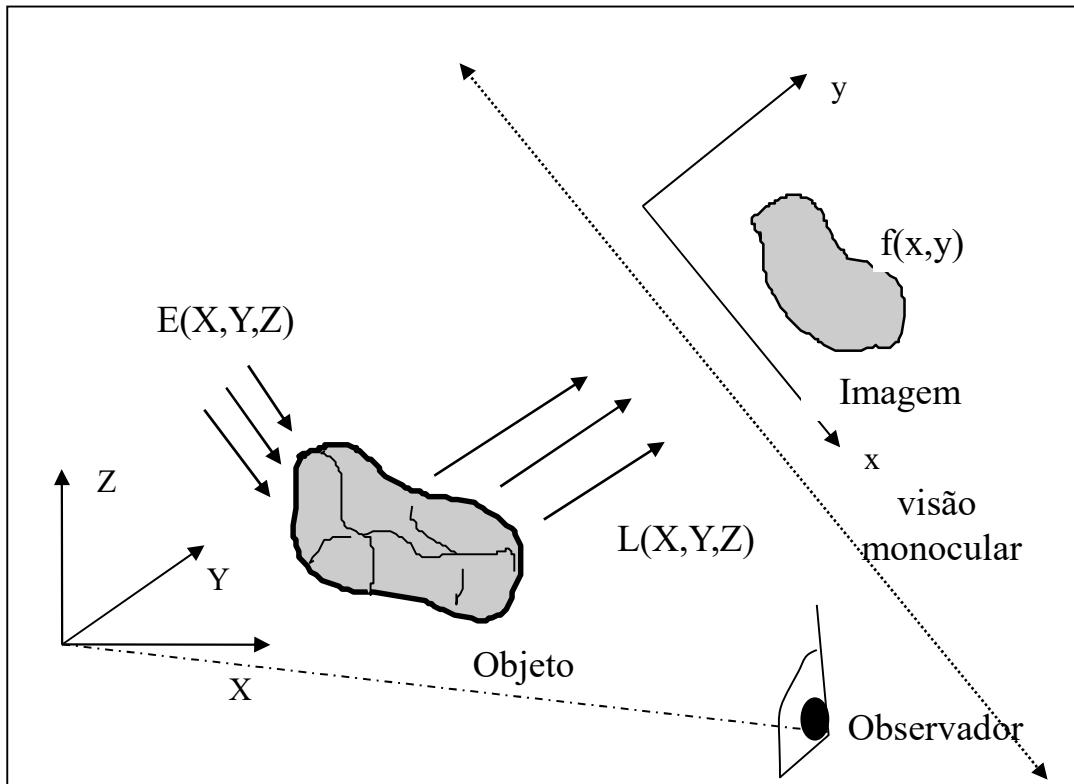
$$x_d = \frac{2 \cdot x}{1 + \sqrt{1 - 4 \cdot k_r \cdot r^2}} \quad y_d = \frac{2 \cdot y}{1 + \sqrt{1 - 4 \cdot k_r \cdot r^2}} \quad r^2 = x^2 + y^2$$



# Formação da imagem

## Fatores físicos que determinam a formação da imagem

- geometria do objeto observado,
- reflectância/refractância de suas superfícies,
- irradiância das fontes luminosas,
- posição/orientação do observador em relação ao objeto,
- características psico-visuais do observador ou opto-eletrônicas da câmera.



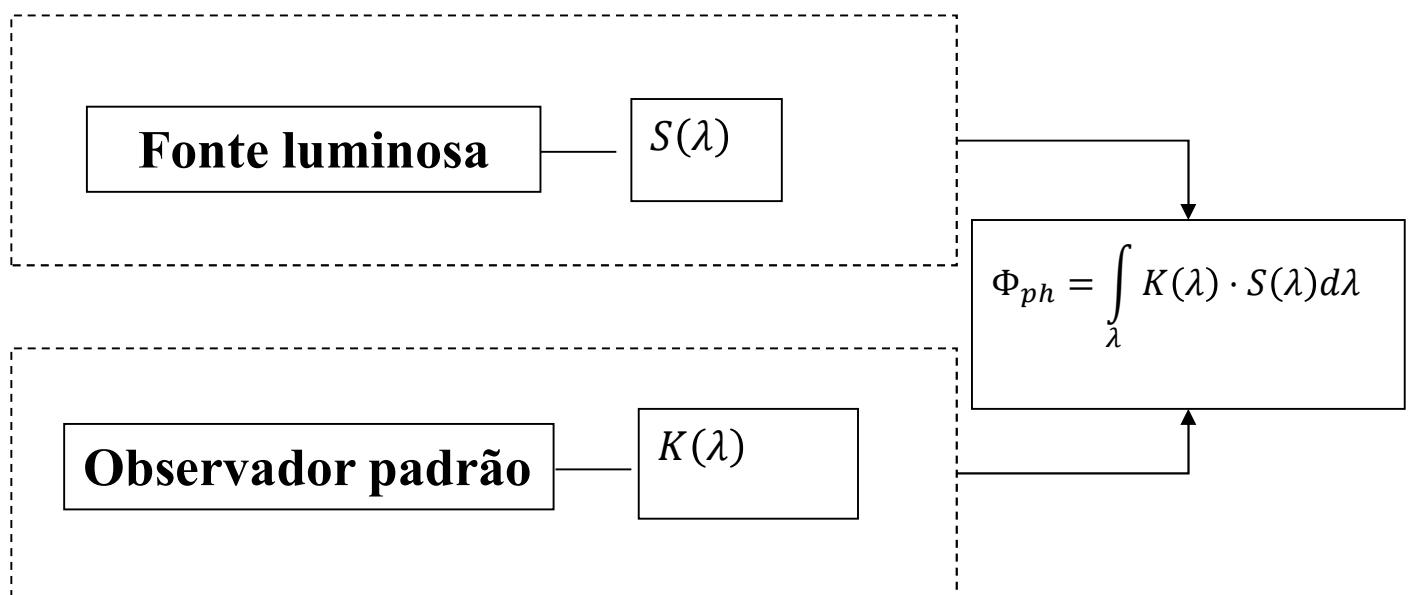
# Elementos de radiometria

**Radiometria:** mede fluxo e transferência da energia radiante.

**Fotometria:** mede fluxo e transferência de energia luminosa em termos do ‘brilho’ percebido pelo olho humano.

**Fluxo radiante ( $\Phi$ ):** potência (em Watts) da fonte radiante, abrangendo todo o espectro electromagnético.

**Fluxo luminoso ( $\Phi_{ph}$ ):** potência (medida em lúmens= $cd \cdot sr$ ) da fonte luminosa, levando em conta seu espectro  $S(\lambda)$  e as características visuais  $K(\lambda)$  (eficácia luminosa: lúmen/Watt) de um ‘observador padrão’ ).



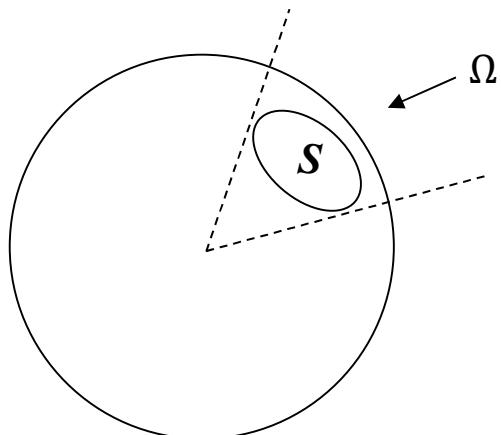
## Observação

Medidas fotométricas dependem do observador !

É sempre preferível utilizar medidas radiométricas

# Elementos de radiometria

**Ângulo sólido ( $\Omega$ ):** razão entre a área  $S$  da superfície de uma esfera de raio unitário limitada por um cone com vértice no centro da esfera e a área da superfície da esfera

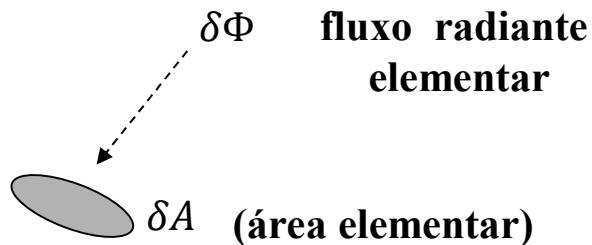


$$\Omega = \frac{S}{4\pi} \quad (\text{eq.1})$$

$$S = 1 \Rightarrow \Omega = \text{stero radiano}$$

**Irradiância ( $E$ ):** medida em  $\text{W/m}^2$ , é o fluxo radiante incidente em uma superfície por unidade de área.

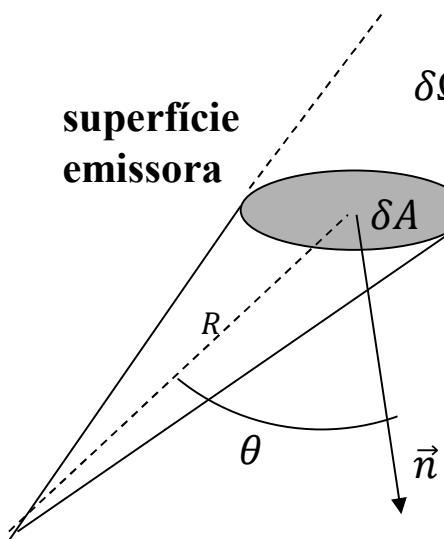
$$E = \frac{\delta\Phi}{\delta A} \quad (\text{eq. 2})$$



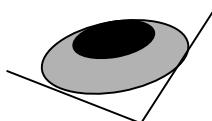
**Iluminância:** medida em lúmen/ $\text{m}^2$ , é o fluxo luminoso incidente em uma superfície por unidade de área.

# Elementos de radiometria

**Radiância ( $L$ ):** medida em  $\text{W/m}^2/\text{sr}$ , é o fluxo radiante emitido, refletido, transmitido ou recebido por uma superfície por unidade de ângulo sólido e por unidade de área projetada (logo, depende da posição do observador).



$$L = \frac{\delta\Phi}{\delta\Omega\delta A \cos} \quad (\text{eq. 3})$$

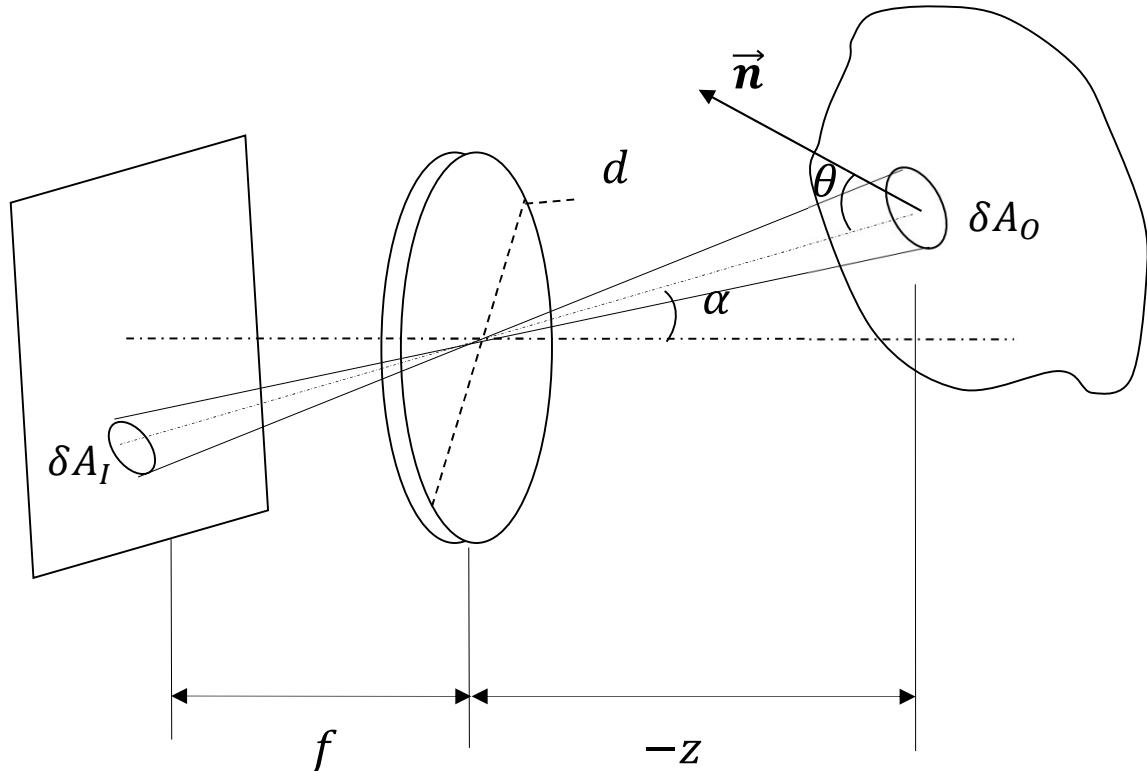


**Brilho ( $L_{ph}$ ) ou luminância:** medida em  $\text{lm/m}^2/\text{sr}$ , é a grandeza fotométrica correspondente à radiância.

# Elementos de radiometria

## Radiância do objeto versus irradiância da imagem

Neste tópico deduz-se a função que determina a irradiância da imagem (fonte de luz) de um objeto cuja radiância é conhecida.



Ângulo sólido subentendido pela área elementar da **imagem**:

$$\Omega_I = \frac{\delta A_I \cos \alpha}{(f/\cos \alpha)^2} \quad (\text{eq. 4})$$

Ângulo sólido subentendido pela área elementar do **objeto**:

$$\Omega_O = \frac{\delta A_O \cos \theta}{(z/\cos \theta)^2} \quad (\text{eq. 5})$$

Logo, tem-se:

$$\Omega_I = \Omega_O \Rightarrow \frac{\delta A_O}{\delta A_I} = \frac{\cos \alpha}{\cos \theta} \left( \frac{z}{f} \right)^2 \quad (\text{eq. 6})$$

# Elementos de radiometria

## Irradiância local da imagem gerada em uma câmera

Ângulo sólido abrangido pelo centro da superfície elementar do objeto e pelo diâmetro da lente:  $\Omega_L$ .

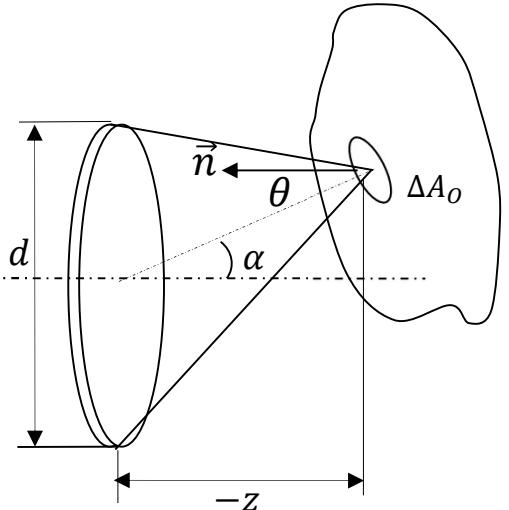
Radiância da superfície elementar do objeto que atravessa a Lente:  $L$ .

Fluxo radiante proveniente da superfície elementar do objeto e que atravessa a lente:  $\delta\Phi$

Irradiação da superfície elementar da imagem:  $E$ .

Da equação 6, tem-se:

$$\Omega = \frac{\left(\frac{\pi d^2}{4}\right) \cos \alpha}{\left(\frac{z}{\cos \alpha}\right)^2} = \frac{\pi}{4} \left(\frac{d}{z}\right)^2 \cos^3 \alpha \quad (\text{eq. 7})$$



Introduzindo-se 7 na eq. 3, obtém-se:

$$\delta\Phi = L \delta A_O \Omega_L \cos \theta = L \delta A_O \frac{\pi}{4} \left(\frac{d}{z}\right)^2 \cos^3 \alpha \cos \theta \quad (\text{eq. 8})$$

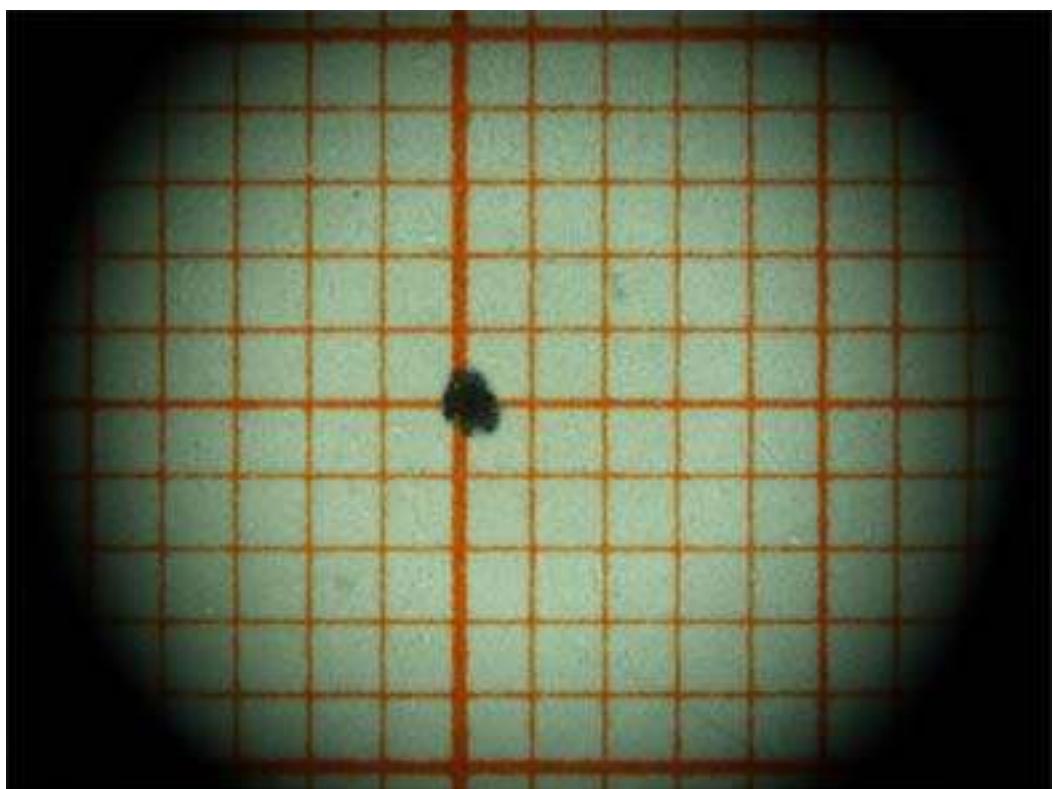
Finalmente, introduzindo-se 8 em 2, resulta:

$$E = \frac{\delta\Phi}{\delta A_I} = L \frac{\pi}{4} \left(\frac{d}{f}\right)^2 (\cos \alpha)^4 \quad (\text{eq. 9})$$

# Elementos de radiometria

## Efeito vinheta

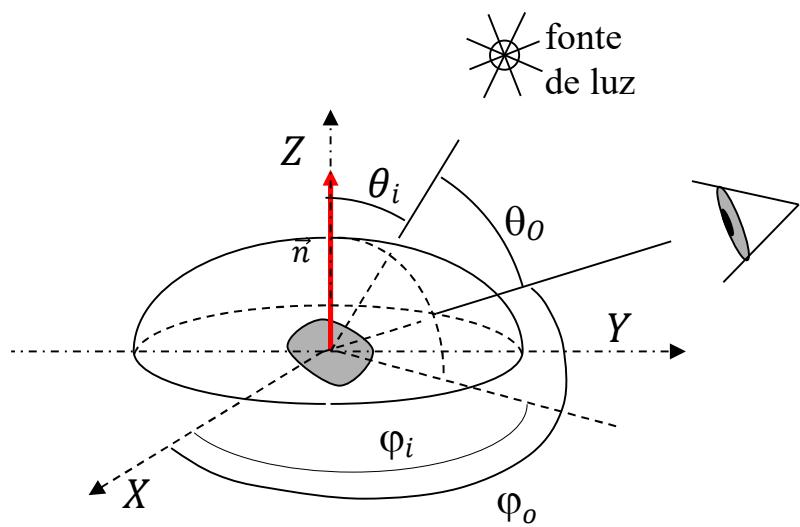
A equação 9 ( $E = E(L)$ ) indica que a irradiância da imagem dos objetos afastados do centro óptico da lente sofre atenuação mais acentuada que a dos objetos que se situam nas imediações do eixo óptico da lente (vide Fig. abaixo).



# Elementos de radiometria

## Reflectância de uma superfície

Se o objeto é opaco, chama-se de **reflectância** à sua radiância. Considere-se o arranjo de iluminação de uma superfície elementar, conforme ilustrado na figura abaixo



A irradiância da superfície elementar é função dos ângulos  $\theta_i, \varphi_i$  de incidência do raio luminoso, ou seja:  $dE(\theta_i, \varphi_i)$

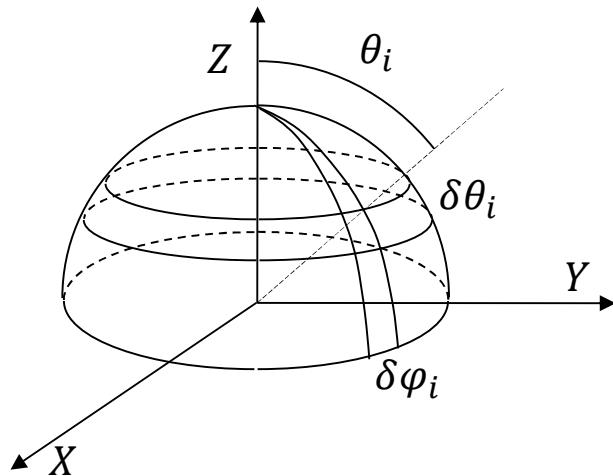
A radiância (reflectância) dessa superfície elementar na direção do observador é função dos ângulos de observação  $\theta_o, \varphi_o$ , ou seja:  $dL(\theta_o, \varphi_o)$

# Elementos de radiometria

## Fontes extensas

A cena pode ser iluminada por múltiplas fontes de luz ou por uma fonte extensa, como, por exemplo, o céu.

Uma fonte elementar contínua de dimensões  $\delta\theta_i, \delta\varphi_i$  delimita um ângulo sólido  $\delta\omega$  dado por:



$$\delta\omega = \sin\theta_i \delta\theta_i \delta\varphi_i \quad (\text{eq.10})$$

Logo, a irradiância da fonte elementar é  $E(\theta_i, \varphi_i) \sin\theta_i \delta\theta_i \delta\varphi_i$

e a irradiância total da fonte extensa é:

$$E_O = \int_{-\pi}^{\pi} \int_0^{\frac{\pi}{2}} E(\theta_i, \varphi_i) \sin\theta_i \cos\theta_i d\theta_i d\varphi_i \quad (\text{eq. 11})$$

É importante destacar que a reflectância de uma superfície depende da:

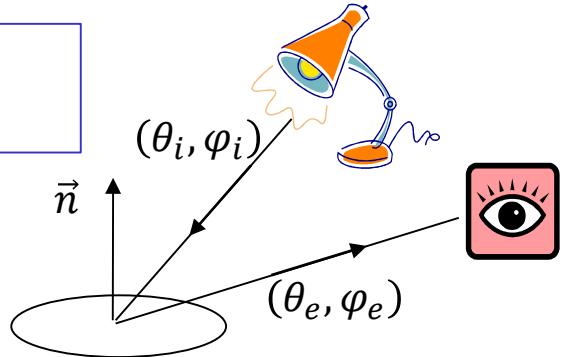
- quantidade de luz incidente;
- quantidade de luz refletida;
- geometria da reflexão: direção de iluminação versus direção de observação.

# Elementos de radiometria

## Função distribuição bidirecional de reflectância

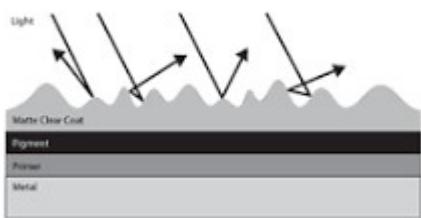
Essa função expressa a intensidade do brilho de uma superfície vista segundo a direção  $(\theta_e, \varphi_e)$  quando iluminada segundo a direção  $(\theta_i, \varphi_i)$ . Tem a seguinte forma:

$$f(\theta_i, \varphi_i, \theta_e, \varphi_e) = \frac{\Delta L(\theta_e, \varphi_e)}{\Delta E(\theta_i, \varphi_i)} \quad (\text{eq. 12})$$



Para muitas superfícies, a radiânci (reflectância) não se altera quando ocorrem rotações em torno da normal  $\vec{n}$ . Nesses casos,  $L$  depende apenas da diferença  $\varphi_e - \varphi_i$ .

### Exemplo 1: superfícies foscas



### Exemplo 2: superfícies espelhadas

É importante destacar que a função distribuição bidirecional de reflectância está sujeita à condição de reciprocidade de Helmholtz:

$$f(\theta_i, \varphi_i, \theta_e, \varphi_e) = f(\theta_e, \varphi_e, \theta_i, \varphi_i) \quad (\text{eq.13})$$

# Elementos de radiometria

## Superfície lambertiana ideal

É a superfície que reflete toda luz incidente e se apresenta igualmente brilhante quando vista de quaisquer direções.

A função distribuição bidirecional de reflectância de uma tal superfície é constante. Para determiná-la, integra-se a radiância em todas as direções e iguala-se esta à irradiância, ou seja:

$$\int_{-\pi}^{\pi} \int_0^{\frac{\pi}{2}} f(\theta_i, \varphi_i; \theta_e, \varphi_e) E \cos \theta_i \sin \theta_e \cos \theta_e d\theta_e d\varphi_e = E \cos \theta_i \quad (\text{eq.14})$$

Para que a expressão acima seja verdadeira, deve-se ter:

$$2\pi \int_{-\pi}^{\pi} \int_0^{\frac{\pi}{2}} \sin \theta_e \cos \theta_e d\theta_e = 1 \quad (\text{eq.15})$$

ou seja,

$$f(\theta_i, \varphi_i; \theta_e, \varphi_e) = \frac{1}{\pi} \quad (\text{eq.16})$$

Da equação 11, resulta que:  $L = \frac{1}{\pi} E_O$     (eq.17)

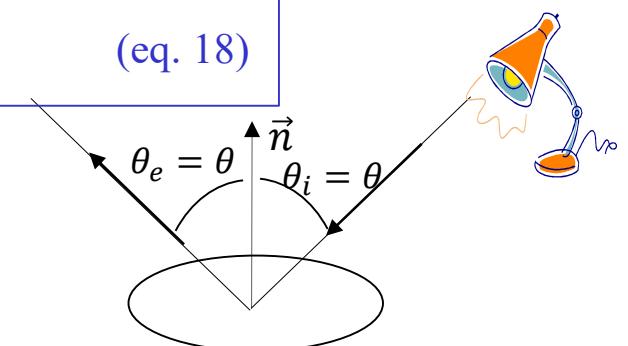
## Superfícies espelhadas

Refletem toda a luz proveniente da direção  $(\theta_i, \varphi_i)$  na direção  $(\theta_i, \varphi_i + \pi)$ . Para tais superfícies, tem-se:

$$f(\theta_i, \varphi_i, \theta_e, \varphi_e) = \frac{\delta(\theta_e - \theta_i)\delta(\varphi_e - \varphi_i - \pi)}{\sin \theta_i \cos \theta_i} \quad (\text{eq. 18})$$

onde  $\delta$  é a função impulsiva bidimensional

$$L(\theta_e, \varphi_e) = E(\theta_e, \varphi_e - \pi) \quad (\text{eq.19})$$

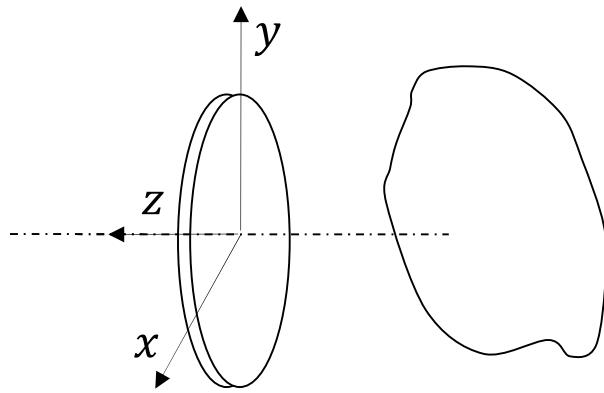
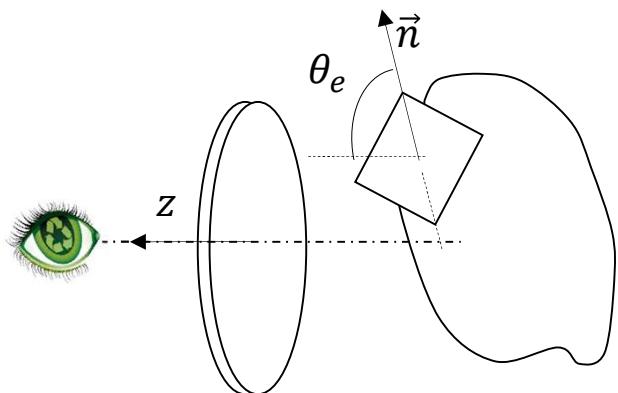


# Elementos de radiometria

## Estimação estéreo-fotométrica da forma da superfície

Integrando-se a função distribuição bidirecional de reflectância ao longo de todas as **direções** de incidência luminosa (conhecidas, portanto), resulta uma função de duas variáveis que pode ser relacionada à orientação da superfície.

Para tanto deve-se adotar um sistema de coordenadas global centrado na posição do observador.



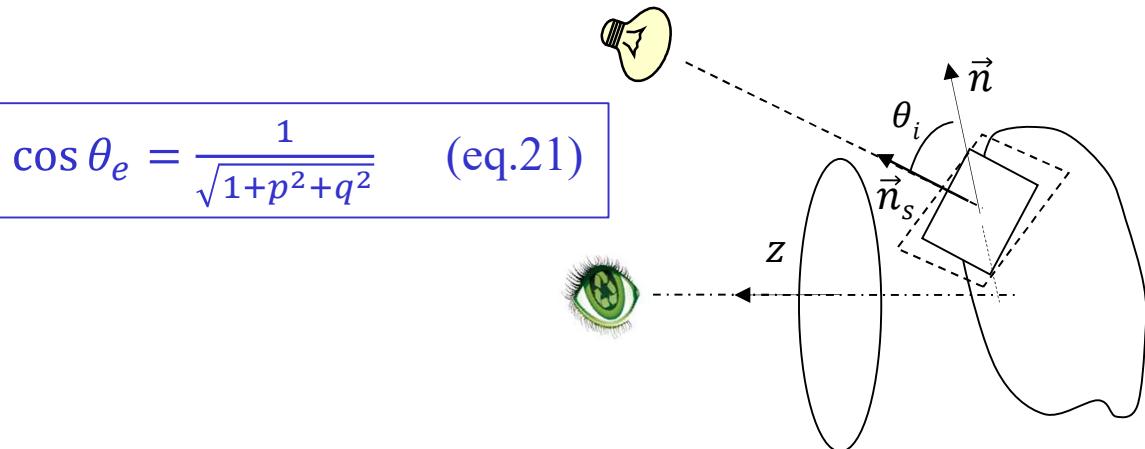
A normal à superfície iluminada  $z = z(x, y)$  é descrita pelo seu gradiente, isto é:

$$\vec{n} = \frac{-\frac{\partial z}{\partial x}\vec{i} - \frac{\partial z}{\partial y}\vec{j} + \vec{k}}{\sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}} = \frac{-p\vec{i} - q\vec{j} + \vec{k}}{\sqrt{p^2 + q^2 + 1}} \quad (\text{eq. 20})$$

# Elementos de radiometria

## Estimação estéreo-fotométrica da forma da superfície (cont)

Portanto, o ângulo  $\theta_e$  formado entre a normal  $\vec{n}$  e a direção do observador  $(0,0,1)$ , é:



Admite-se que a fonte luminosa esteja situada a grande distância do objeto, de modo que um raio luminoso atinja o ponto  $(x, y, z)$  da superfície normalmente a um plano cuja normal é

$$\vec{n}_s = -\frac{p_s \vec{i} - q_s \vec{j} + \vec{k}}{\sqrt{1+p_s^2+q_s^2}} \quad (\text{eq.22})$$

Supondo que a superfície seja lambertiana e iluminada por uma fonte de irradiância  $E$  na **direção** conhecida  $(-p_s, -q_s, 1)$ , tem-se:

$$L = \frac{1}{\pi} E \cos \theta_i \quad (\text{eq.23})$$

Dessa forma, determina-se o ângulo entre a fonte pontual e a normal à superfície, caracterizando-se assim a sua geometria:

$$\cos \theta_i = \frac{1+p_s p + q_s q}{\sqrt{(1+p_s^2+q_s^2)(1+p^2+q^2)}} \quad (\text{eq.24})$$

# Elementos de radiometria

## Estimação estéreo-fotométrica da forma da superfície (cont)

Como a irradiância da imagem é proporcional à radiância da cena, para uma superfície lambertiana iluminada por fonte pontual distante, o mapa de reflectância fica representado por:

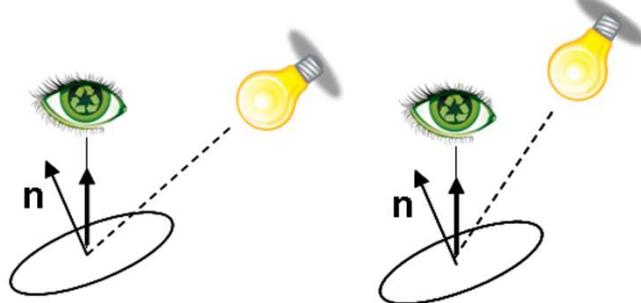
$$R(p, q) = \frac{1 + p_s p + q_s q}{\sqrt{(1 + p^2 + q^2)(1 + p_s^2 + q_s^2)}} \quad (\text{eq.25})$$

onde  $(p_s, q_s, 1)$  caracteriza a direção de iluminação considerada.

É importante destacar que existem **infinitas** superfícies capazes de produzir o mesmo mapa de reflectância, dado que o brilho tem 1 único grau de liberdade enquanto a orientação local da superfície tem 2 graus de liberdade.

Portanto, para estimar essa orientação local a partir do mapa de reflectância deve-se realizar o seguinte procedimento:

1. Utilizar duas imagens capturadas da mesma cena sob diferentes condições de iluminação, conforme ilustrado abaixo.



2. Determinar, para cada pixel da imagem, as funções:

$$\begin{cases} R_1(p, q) = E_1 \\ R_2(p, q) = E_2 \end{cases} \quad (\text{eq.26})$$

# Elementos de radiometria

## Estimação estéreo-fotométrica da forma da superfície (cont)

3. Se as duas equações anteriores forem lineares e independentes, pode-se calcular de forma única os valores de  $p$  e  $q$  para cada ponto da imagem e, assim, obter a geometria da superfície. Caso, por exemplo, se tenha:

$$R_1(p, q) = \sqrt{\frac{1+p_1p+q_1q}{r_1}} \quad R_2(p, q) = \sqrt{\frac{1+p_2p+q_2q}{r_2}} \quad (\text{eq.27})$$

$$\text{com} \quad r_1 = \sqrt{1 + p_1^2 + q_1^2} \quad r_2 = \sqrt{1 + p_2^2 + q_2^2} \quad (\text{eq.28})$$

resulta:

$$\begin{cases} p = \frac{(E_1^2 r_1 - 1)q_2 - (E_2^2 r_2 - 1)q_1}{p_1 q_2 - q_1 p_2} \\ q = \frac{(E_2^2 r_2 - 1)p_1 - (E_1^2 r_1 - 1)p_2}{p_1 q_2 - q_1 p_2} \end{cases} \quad (\text{eq.29})$$

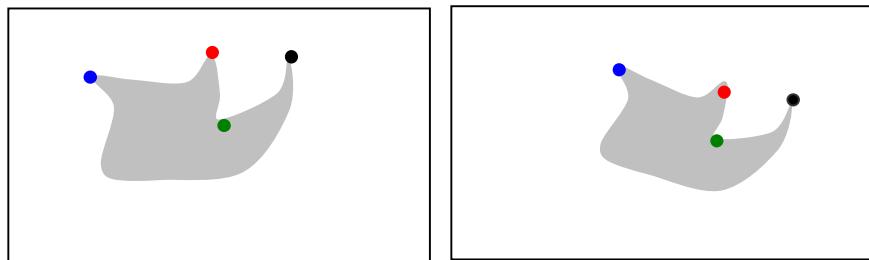
Como o sistema de equações (29) baseia-se em duas imagens da mesma cena, estas imagens devem ser **superpostas** bem como satisfazer à restrição

$$\frac{p_1}{q_1} - \frac{p_2}{q_2} \neq 0 \quad (\text{eq.30})$$

# Elementos de radiometria

## Estimação estéreo-fotométrica da forma da superfície (cont)

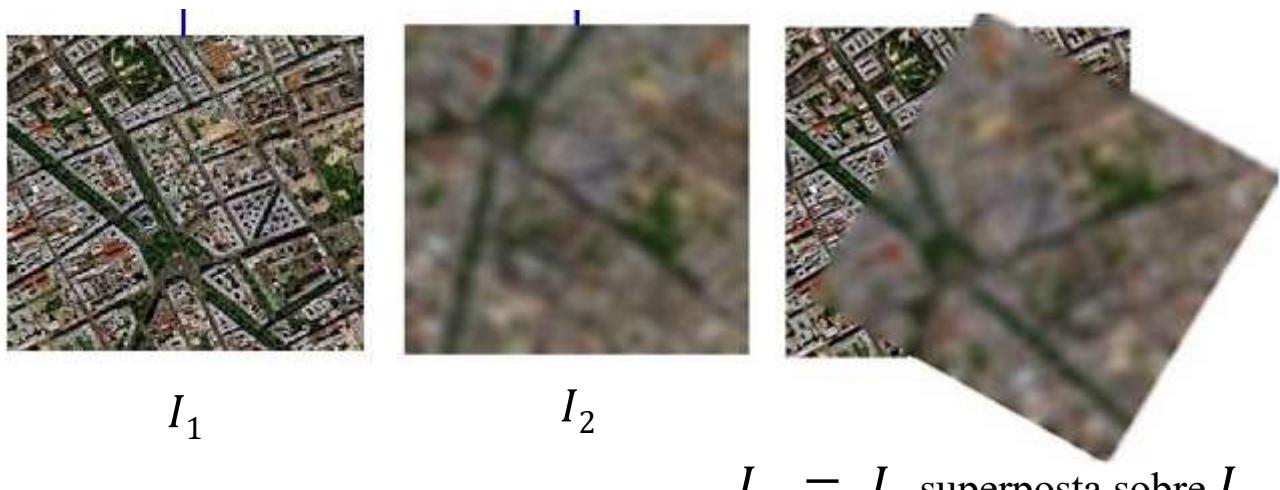
4. Para realizar a superposição de imagens, identificam-se **objetos característicos** (vide figura abaixo) correspondentes em cada uma das imagens  $I_1$  e  $I_2$ .



6. Em seguida, aplica-se uma transformação-afim a uma dessas imagens ( $I_1$  ou  $I_2$ ), de modo a minimizar as diferenças geométricas entre ambas, ou seja:

$$I'_1 = \mathcal{T}(I_1, t_x, t_y, s_x, s_y, \theta) \quad (\text{eq.31})$$

onde  $t_x, t_y, s_x, s_y, \theta$  são, respectivamente, as translações horizontal e vertical, os fatores de escala horizontal e vertical e o ângulo de rotação. Na figura abaixo ilustra-se essa operação.

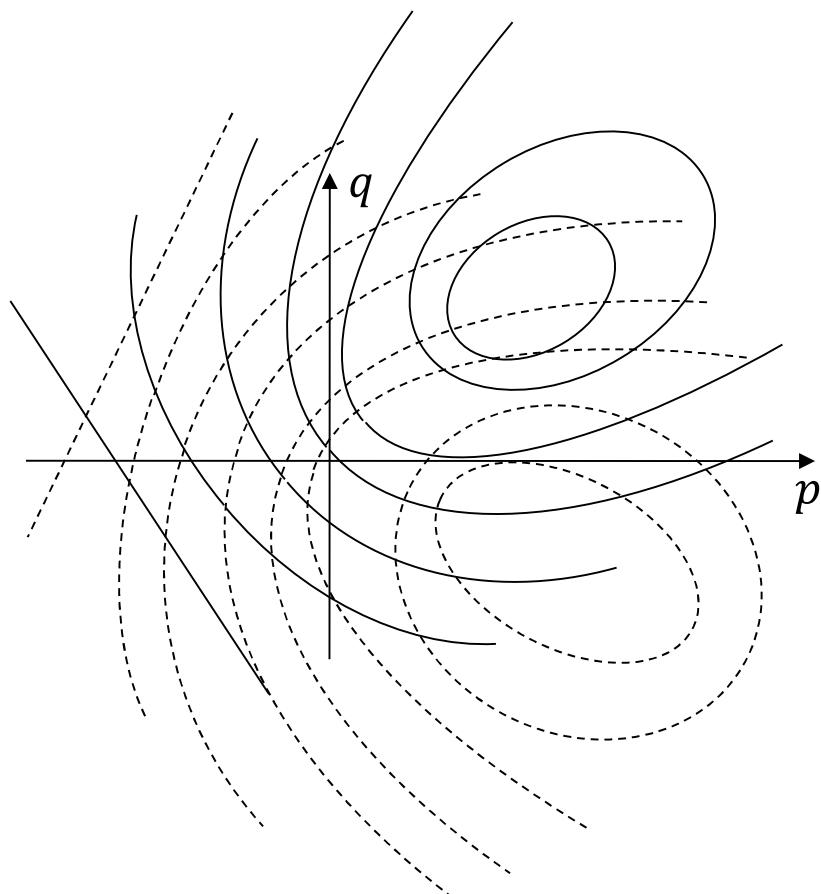


Imagens extraídas de (Yasehim, M.S., 2008)

# Elementos de radiometria

## Estimação estéreo-fotométrica da forma da superfície (cont)

7. Caso as funções radiância de  $I_1$  e  $I_2$  dêem origem a um sistema de duas equações não-lineares, pode-se ter **várias** soluções para  $(p, q)$ , ou, eventualmente, nenhuma. A figura abaixo ilustra essa possibilidade.



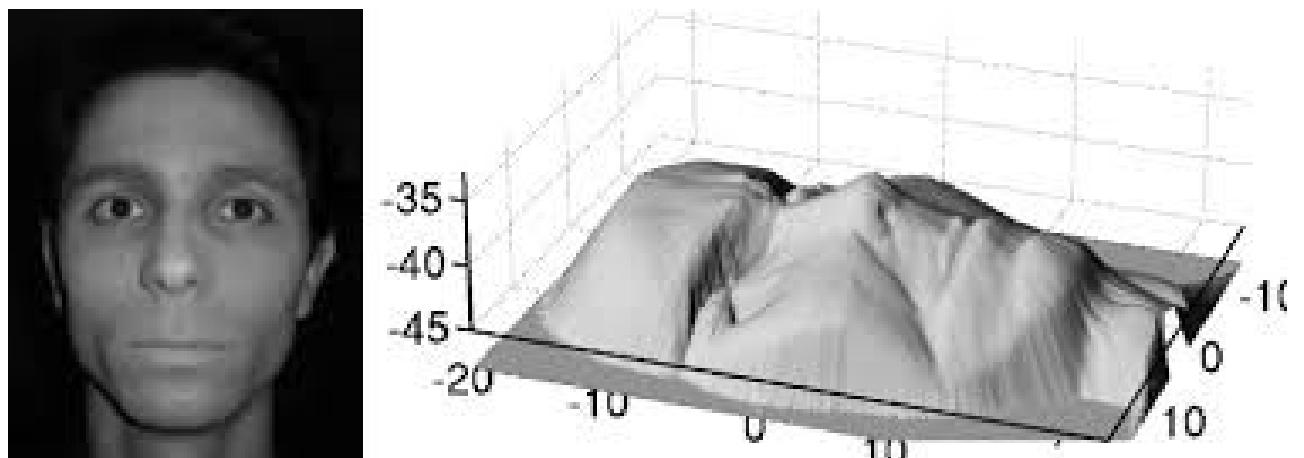
# Elementos de radiometria

## Linhas de pesquisa

- Estéreo-fotometria: superfície 3D estimada a partir de imagens de uma cena sujeita a várias condições de iluminação
- Shape from shading: superfície 3D estimada a partir de uma **única** imagem sujeita a um conjunto de **restrições vinculares**.



Imagens da mesma cena sob duas iluminações distintas (Faugeras, 2006).



Reconstrução baseada na técnica ‘shape from shading’. (Imagens extraídas de Faugeras (2006)).

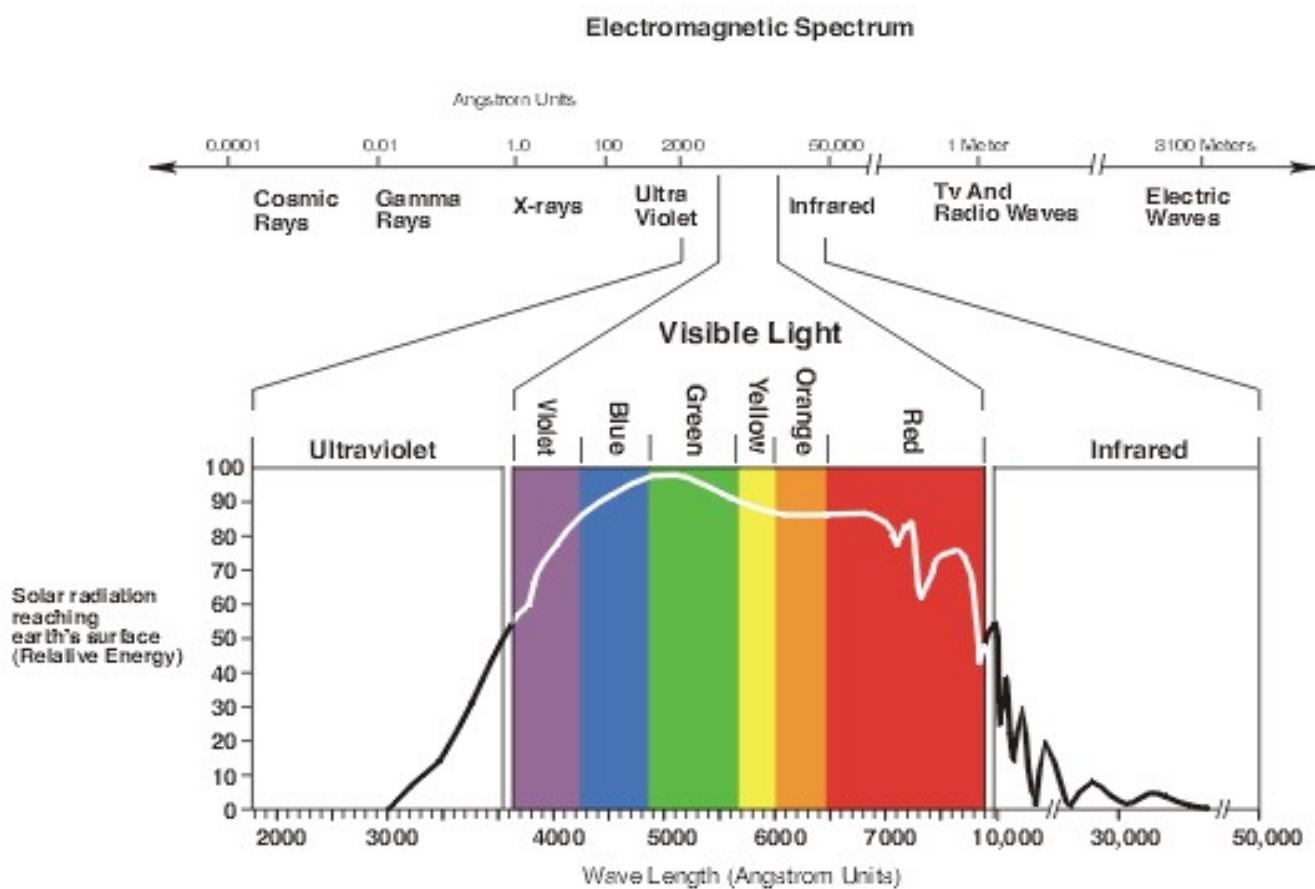
# Elementos de radiometria

## Espectro eletromagnético

A faixa visível abrange comprimentos de onda no intervalo 380-740 nm.

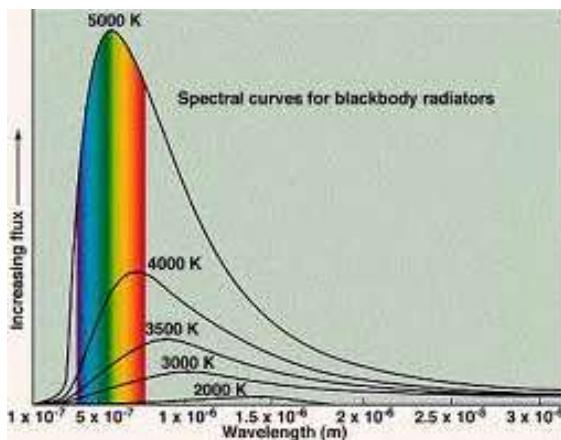
## Características ondulatórias e sua interpretação visual

- Comprimento de onda: interpretado como cor.
- Amplitude: interpretada como intensidade luminosa.
- Fase: auxilia a percepção da profundidade relativa dos objetos.

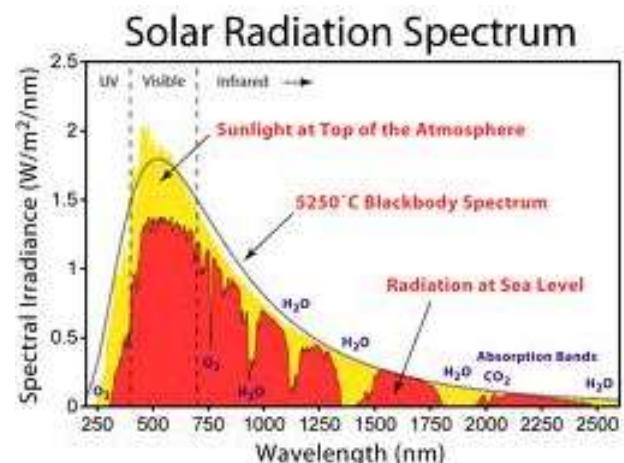


# Elementos de radiometria

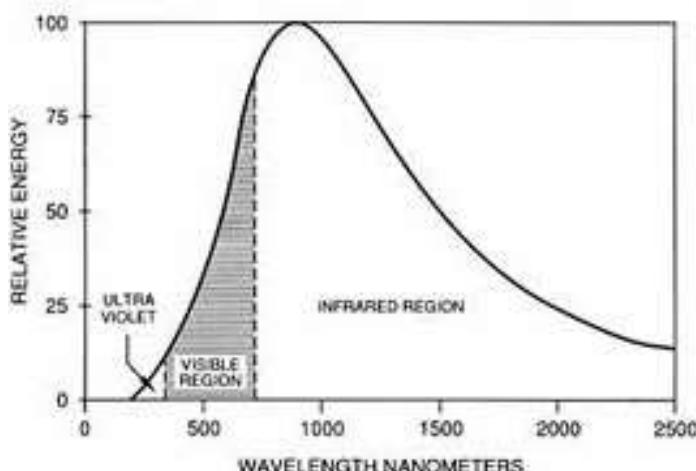
## Espectros de radiação de diferentes fontes luminosas



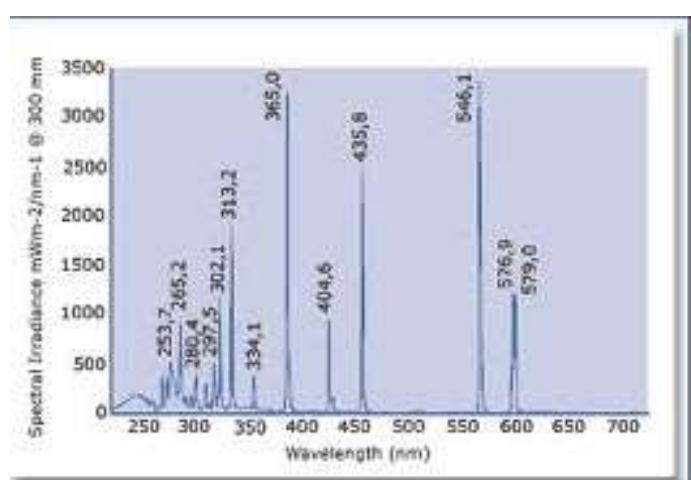
Radiação de corpo negro



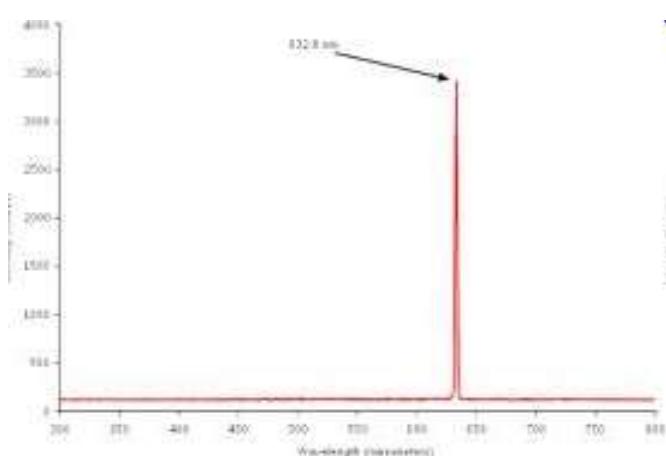
Radiação solar



Radiação de lâmpada de tungstênio



Radiação de lâmpada de vapor de mercúrio



Espectro do laser H-Ne

# Elementos de radiometria

As faixas típicas de interesse do espectro luminoso, são:

- **Azul**, 450-515...520 nm: imagens da atmosfera e águas até a profundidade de 46 m.
- **Verde**, 515..520-590...600 nm: imagens de vegetação e águas até a profundidade de 27 m.
- **Vermelho**, 600...630-680...690 nm: imagens de objetos feitos pelo homem, águas até a profundidade de 9 m, solo e vegetação.
- **Infra-vermelho próximo**, 750-900 nm: vegetação.
- **Infra-vermelho intermediário**, 1550-1750 nm: imagens de vegetação, incêndios florestais, solos úmidos.
- **Infra-vermelho intermediário**, 2080-2350 nm: imagens de solos úmidos, estruturas geológicas, silicatos, argila e incêndios.
- **Infra-vermelho térmico**, 10400-12500 nm: imagens de estruturas geológicas, diferenças térmicas em águas correntes, incêndios, imagens noturnas.

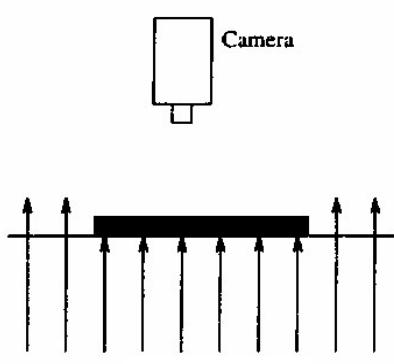
# Elementos de radiometria

## Técnicas de iluminação

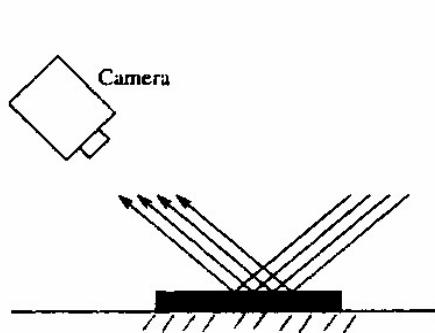
Têm por objetivo facilitar a análise das imagens.

As mais utilizadas, são:

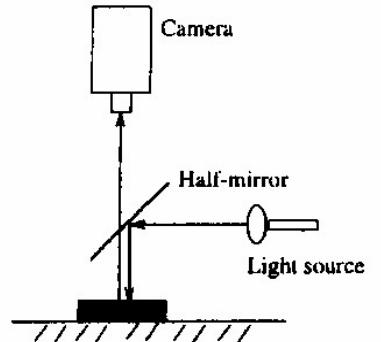
- Iluminação traseira: produz imagens de grande contraste.
- Iluminação direta: adequada para superfícies medianamente refletoras.
- Iluminação difusa: indicada para superfícies muito refletoras.



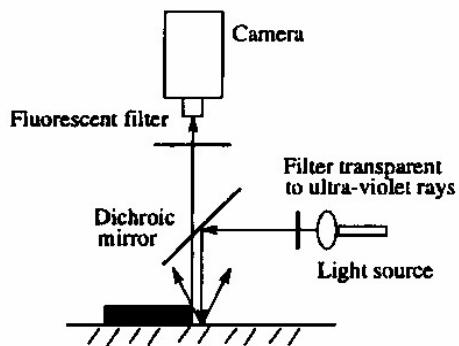
traseira



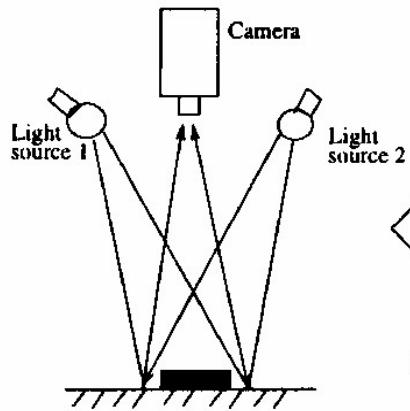
direta



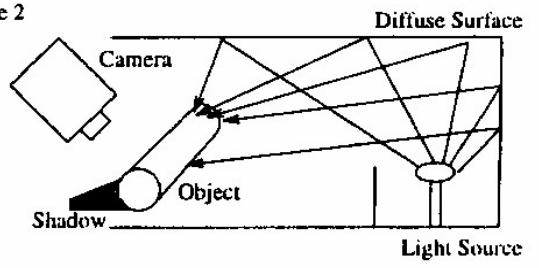
vertical



fluorescente



bi-direcional

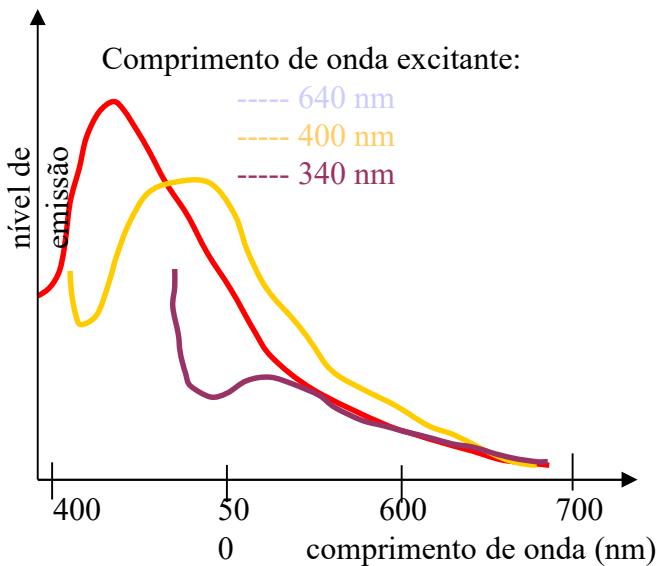


difusa

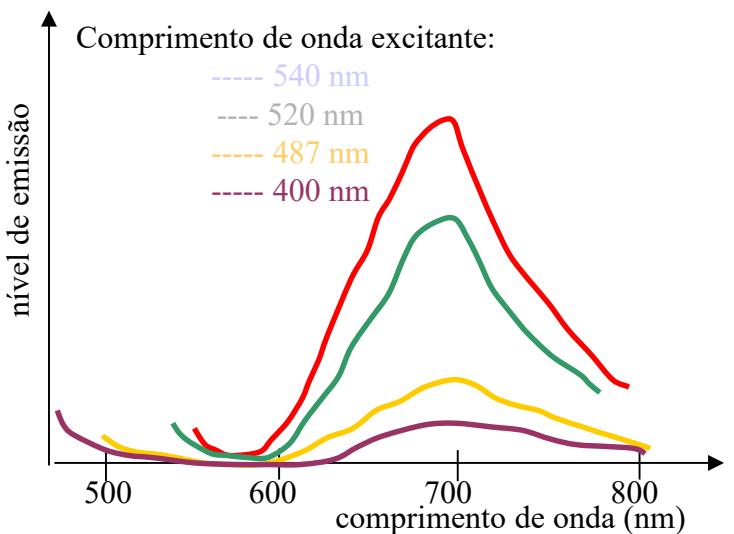
# Elementos de radiometria

## Iluminação fluorescente

Muito útil à observação de objetos feitos de materiais diferentes, mas pouco distintos à luz visível. Como os materiais têm espectros de emissão característicos, excitando-os com radiação apropriada, obtém-se imagens com maior contraste. As figuras abaixo exemplificam a técnica.



Espectro de emissão de fibra de vidro



Espectro de emissão do gel resistente à corrosão



Região da placa iluminada com luz visível



A mesma região iluminada com luz fluorescente

# Cores

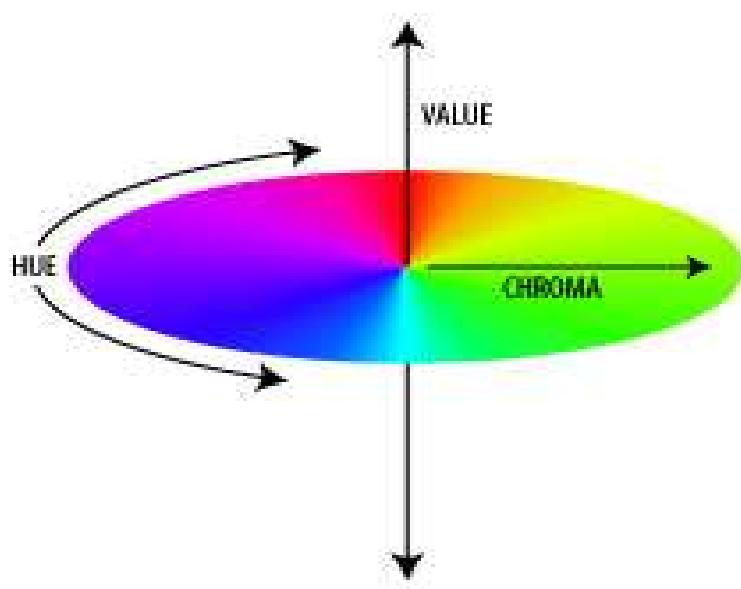
## Descritores

Considerando-se a interação observador / fonte iluminante, a luz é caracterizada por:

- Matiz: distingue luzes de cores distintas, ou seja, associadas a diferentes comprimentos de onda.
- Saturação: mede a ‘tonalidade’ do matiz, desde a mais cinza à mais viva.
- Brilho: mede a luminosidade relativa de uma dada cor, desde o negro até o branco.

## Sistema Munsell

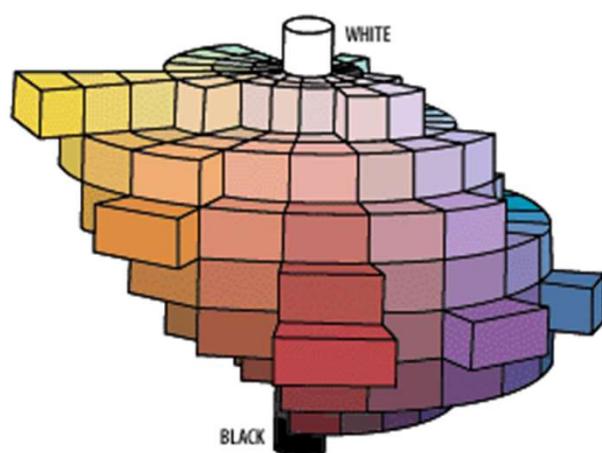
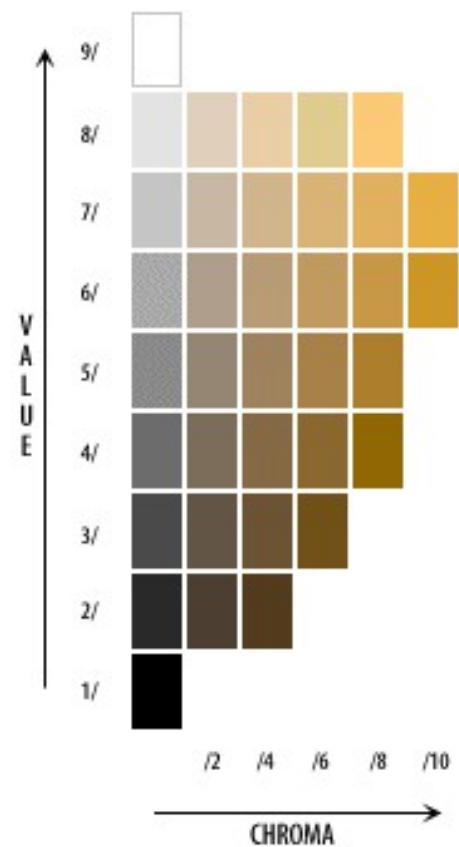
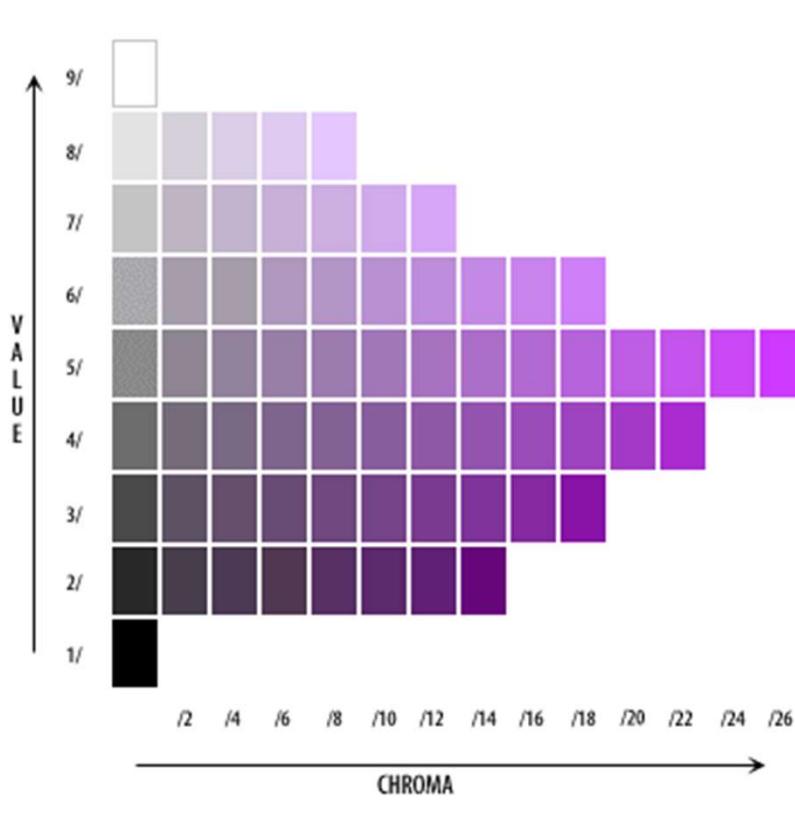
Nesse sistema, matiz ('hue'), saturação ('chroma') e brilho ('value'), são representados em uma figura de forma elipsoidal, onde os matizes se distribuem segundo a direção angular do plano central e os eixos de saturação e brilho apontam na direção das cores menos saturadas e mais luminosas.



# Cores

## Sistema Munsell

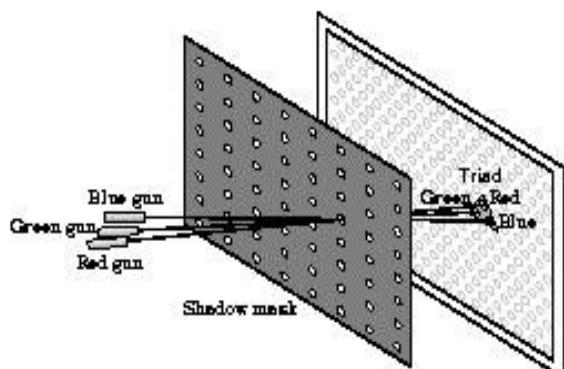
Abaixo, apresentam-se dois perfis de variação de saturação e brilho para duas cores distintas, bem como uma representação tridimensional para o sistema de Munsell



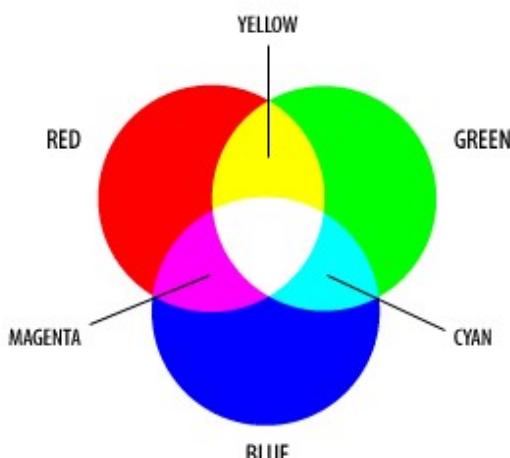
# Cores

## Sistema RGB

Esse é o sistema aditivo, no qual se obtêm cores a partir da superposição de luzes de 3 comprimentos de onda padrão de distintas intensidades. É importante destacar que muitas das cores percebidas pela visão humana **não** podem ser sintetizadas pelo método aditivo.



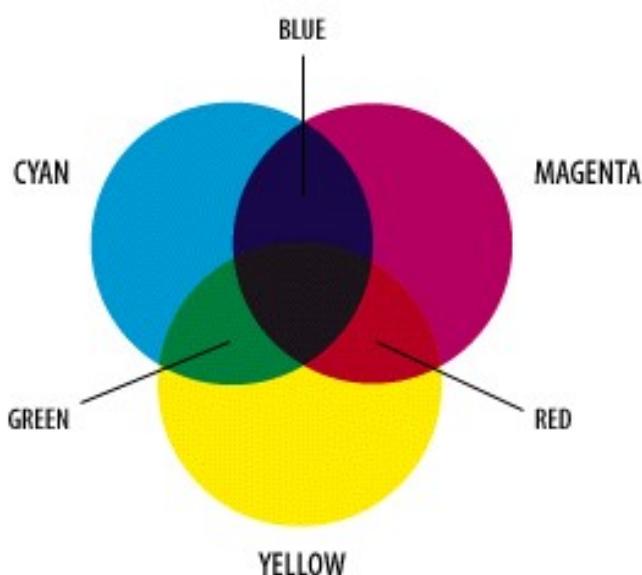
Geração de cores em monitor de vídeo



Cores primárias e secundárias no sistema RGB

## Sistema CMY

Esse é o sistema substrutivo: a superfície do objeto absorve parte do espectro da luz visível refletindo os demais comprimentos de onda. As cores primárias são: ciano, magenta e amarelo; as secundárias são verde, azul e vermelho. Esse sistema não reproduz o brilho das cores RGB e possui menor gama de matizes.



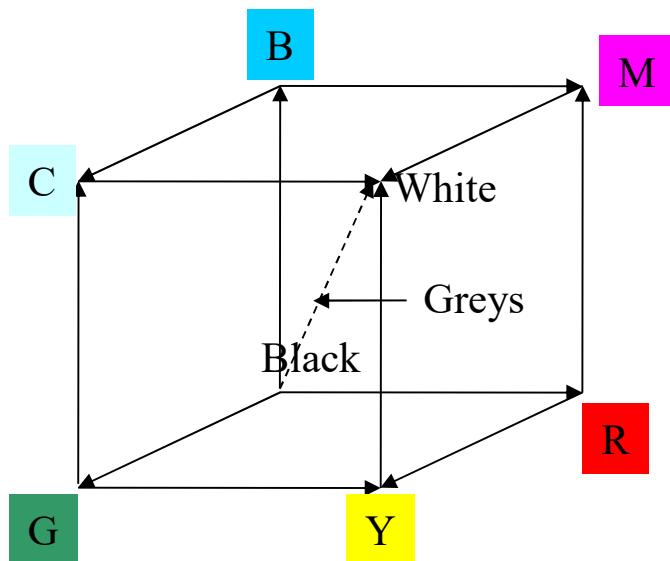
Cores primárias no sistema CMY

# Cores

## Conversão RGB-CMY

Para converter cores entre o sistema *RGB* e *CMY*, basta utilizar as relações a seguir:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{eq.32})$$



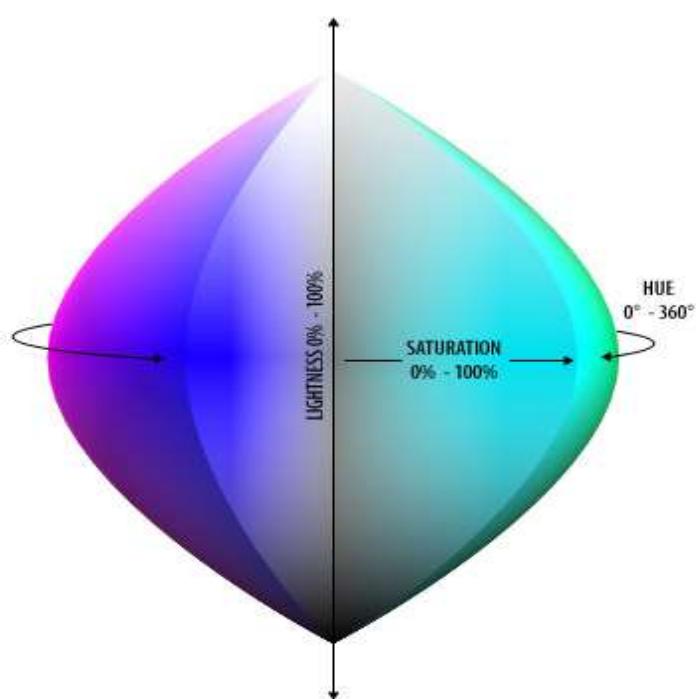
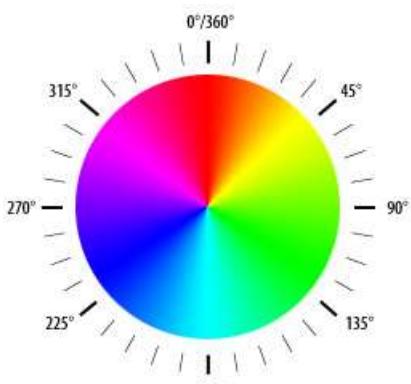
Espaço de cores RGB.

# Cores

## Sistema HLS

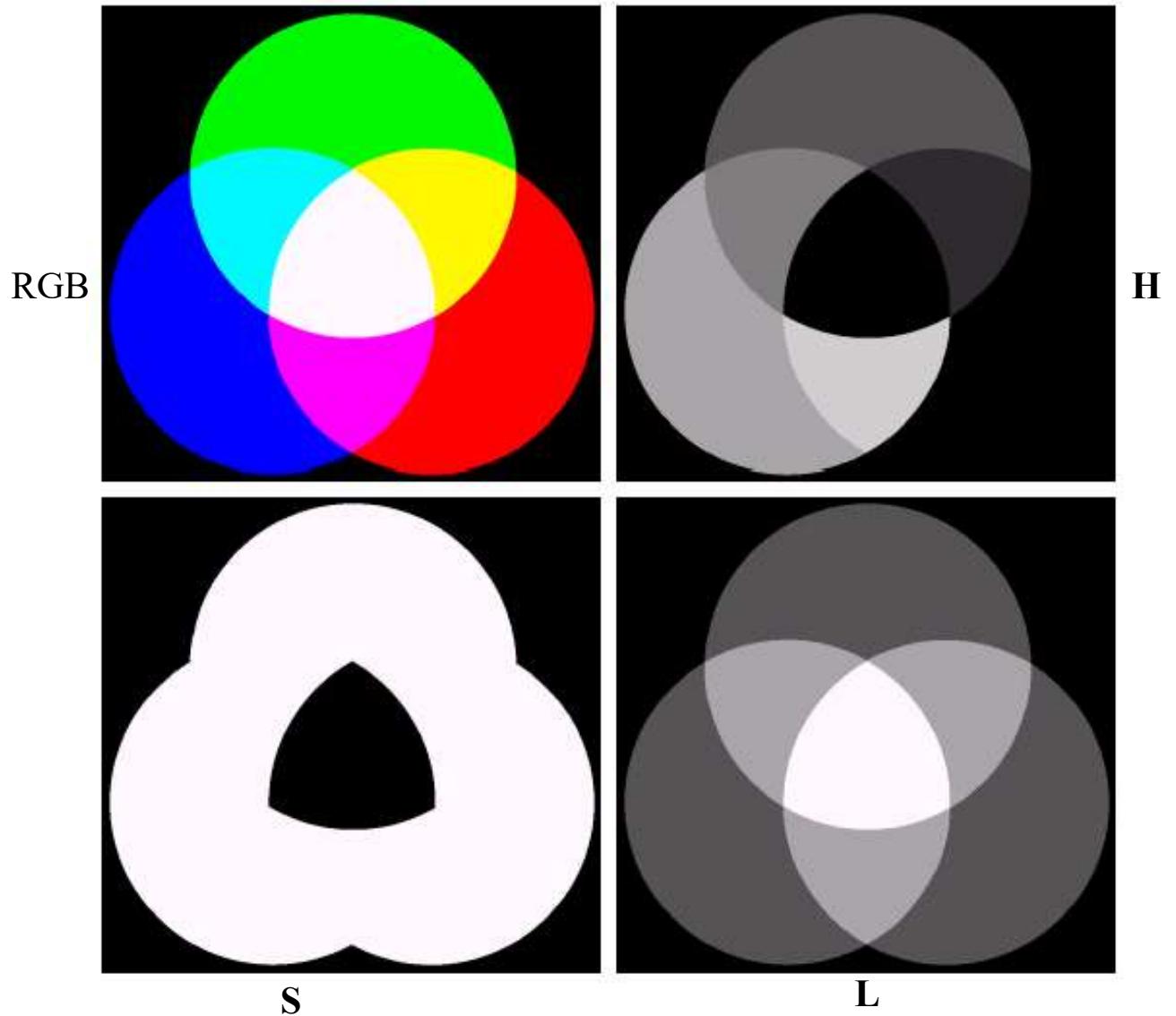
O matiz (H) é representado sobre um eixo radial, variando de 0 a  $360^\circ$ , iniciando e terminando em vermelho. A luminosidade (L) é representada sobre um eixo linear, variando de 0 (negro) até 1 (branco). A saturação é representada sobre um eixo linear, variando de 0 até 1 (mínima e máxima intensidade do matiz).

As figuras abaixo ilustram esse sistema.



# Cores

## Sistema HLS



Espaços de cores RGB e HLS. (Extraída de Gonzalez e Woods)

## Conversão RGB/HLS

Utilizam-se as seguintes fórmulas:

$$\begin{aligned} H &= \frac{1}{2\pi} \left\{ \frac{\pi}{2} - \tan^{-1} \left[ \left( \frac{2R-G-B}{\sqrt{3}} \right) (G - B) \right] \right\} \\ L &= \frac{R+G+B}{3} \\ S &= 1 - \frac{\min(R,G,B)}{3} \end{aligned} \quad (\text{eq.33})$$

# Cores

## Sistema CIE (1931)

Esse é um sistema dito ‘absoluto’. Baseia-se em uma norma que permite comparar cores sob o ponto de vista da percepção visual humana. Para tanto, faz-se uso de fontes luminosas padronizadas e de observadores humanos padrão.

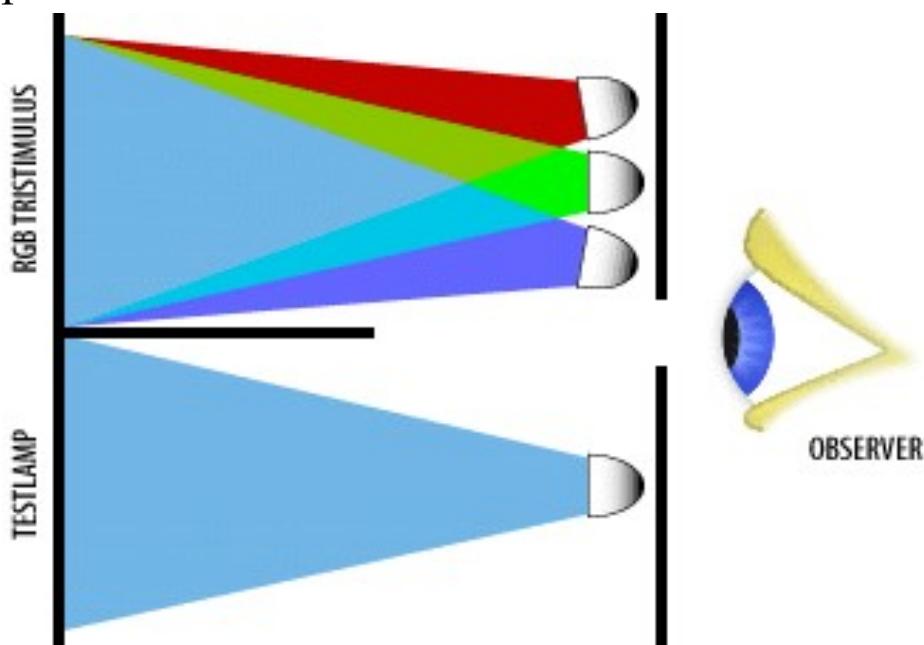
Como iluminantes padrão, utilizam-se:

- lâmpada de tungstênio com temperatura de cor de 2854K,
- fonte emulando a luz do sol ao meio-dia,
- fonte emulando a iluminância média de um dia típico.

Os observadores padrão eram pequenos grupos de pessoas que apresentavam visão considerada normal.

Adotava-se, então, a seguinte metodologia:

1. Apresentavam-se aos observadores duas cores – uma vinda de uma lâmpada emitindo luz em um dado comprimento de onda, outra gerada pela combinação da luz de três lâmpadas emitindo nas bandas do vermelho (700nm), verde (546,1nm) e azul (435,8nm).
2. Quando ambas as cores se tornavam indistinguíveis (metamerismo), os valores de R, G e B da cor composta eram associados à cor espectral pura.



# Cores

Na retina dos observadores realizava-se o seguinte processo:

$$q_i = \int_{\lambda_1}^{\lambda_2} I(\lambda) R_i(\lambda) S(\lambda) d\lambda, i = 1, 2, 3 \quad (\text{eq.34})$$

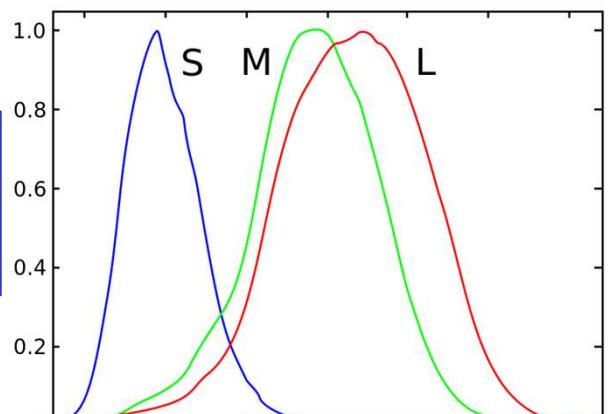
onde:

$I$  : intensidade espectral de iluminação

$R_i$ : sensibilidade espectral do sensor

$S$  : reflectância da superfície

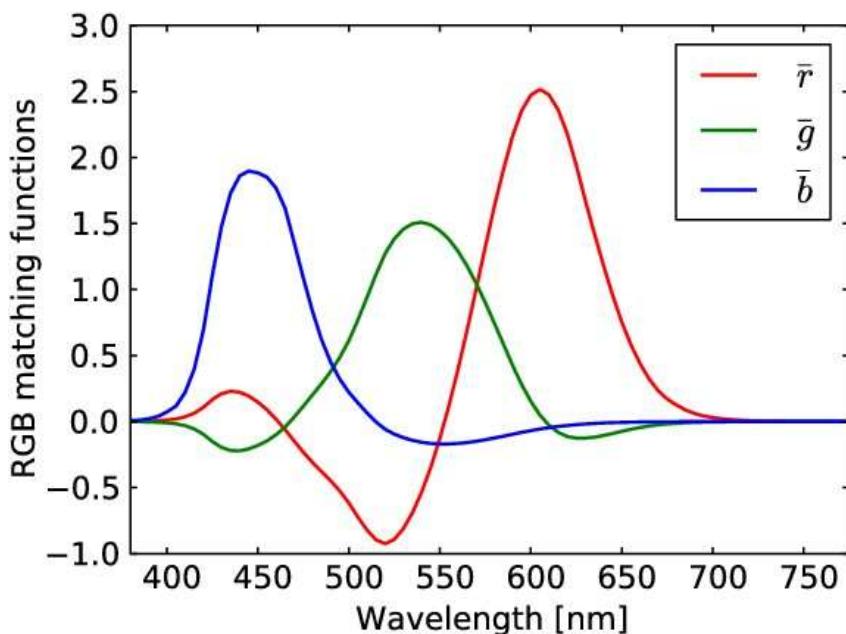
$q_i$  : resposta espectral do sensor  $i$



Sensibilidade normalizada dos cones da retina humana

## Limitação do modelo CIE-1931

Esse modelo era incapaz de reproduzir todas as cores espectrais mediante a combinação de frações R, G e B, a menos que essas pudessem, às vezes, assumir valores **negativos** (vide figura abaixo), o que é contraintuitivo. Na prática, isso se fazia somando-se à cor espectral uma fração de uma das cores R, G ou B e efetuando-se a comparação até a ocorrência do metamerismo).

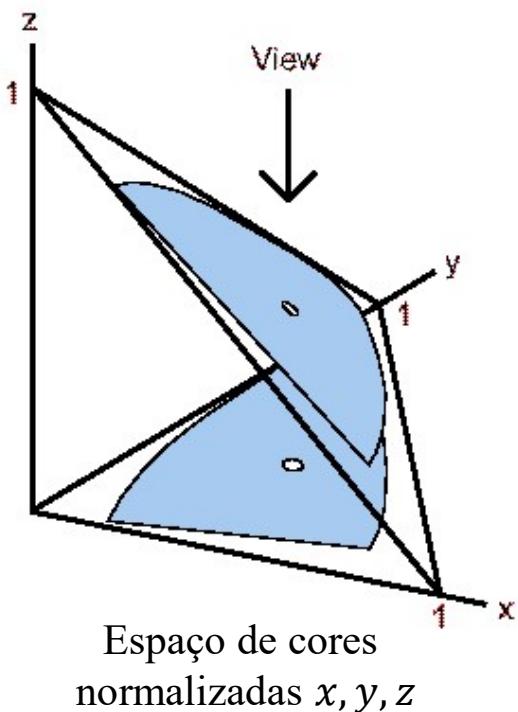


Funções de casamento de cores metaméricas com cores espetrais – padrão CIE-1931.

# Cores

## Novo padrão CIE

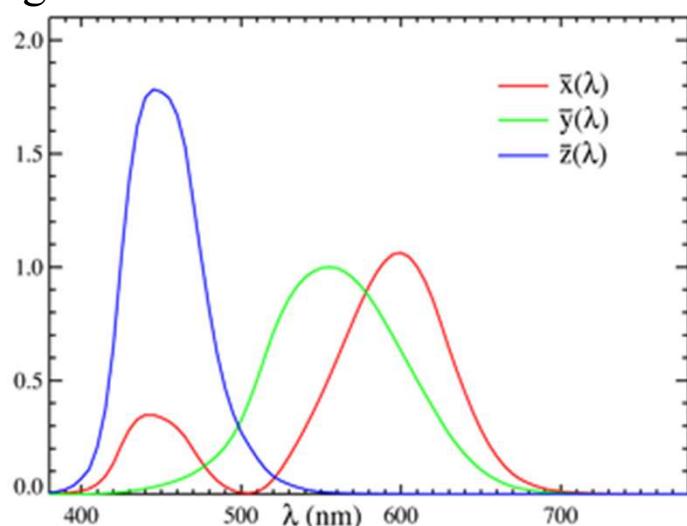
Para resolver o problema apontado no item anterior substituíram-se as variáveis  $R, G, B$  por variáveis artificiais  $X, Y, Z$  **positivas**, tais que  $Z$  correspondia à luminância da cor, e  $X$  e  $Y$ , grosso modo, ao matiz e à saturação (vide figura abaixo).



Utilizando-se valores normalizados  $x, y$  e  $z$  para as variáveis  $X, Y, Z$ , de acordo com as expressões

$$\begin{aligned}x &= \frac{x}{X+Y+Z} \\y &= \frac{y}{X+Y+Z} \\z &= \frac{z}{X+Y+Z}\end{aligned}\quad (\text{eq.35})$$

nota-se que  $x + y + z = 1$  representa um trecho de superfície plana triangular no espaço  $(x, y, z)$ , conforme indicado na figura ao lado.



Na figura à direita ilustram-se as funções de casamento de cores representadas no novo padrão CIE.

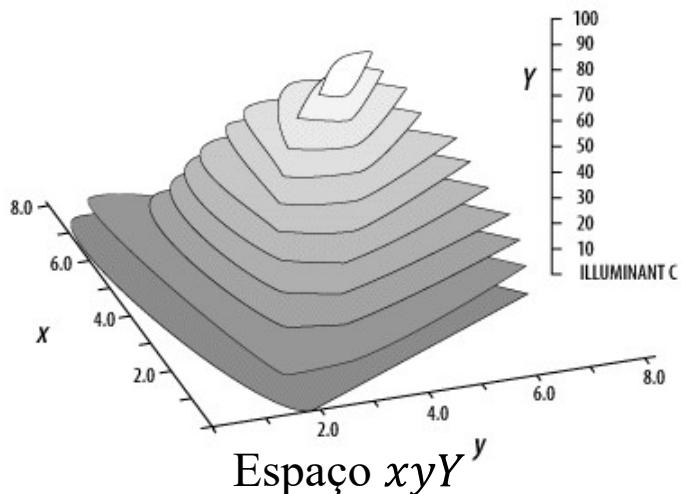
Funções de casamento de cores metaméricas com cores espectrais – novo padrão CIE.

# Cores

## Novo modelo CIE

Nesse espaço normalizado, a projeção das triplas  $(x, y, z)$  sobre o plano  $xy$  é chamada de **diagrama de cromaticidade**, o qual pode ser utilizado para recuperar as coordenadas  $X, Y, Z$  a partir do par de valores  $xy$  e da luminância  $Y$  (vide figura abaixo), pois:

$$\begin{aligned}z &= 1 - x - y \\X &= Y \left( \frac{x}{y} \right) \\Z &= Y \left( \frac{z}{y} \right)\end{aligned}\quad (\text{eq.36})$$



Esse diagrama (vide figura abaixo) contém todas as cores espectrais visíveis pelo ser humano, cores essas que diferem apenas em luminância.

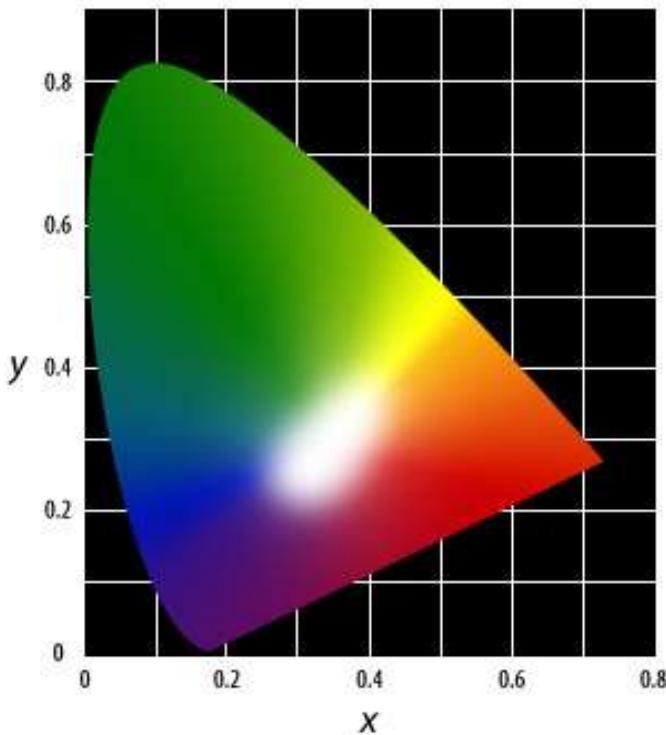
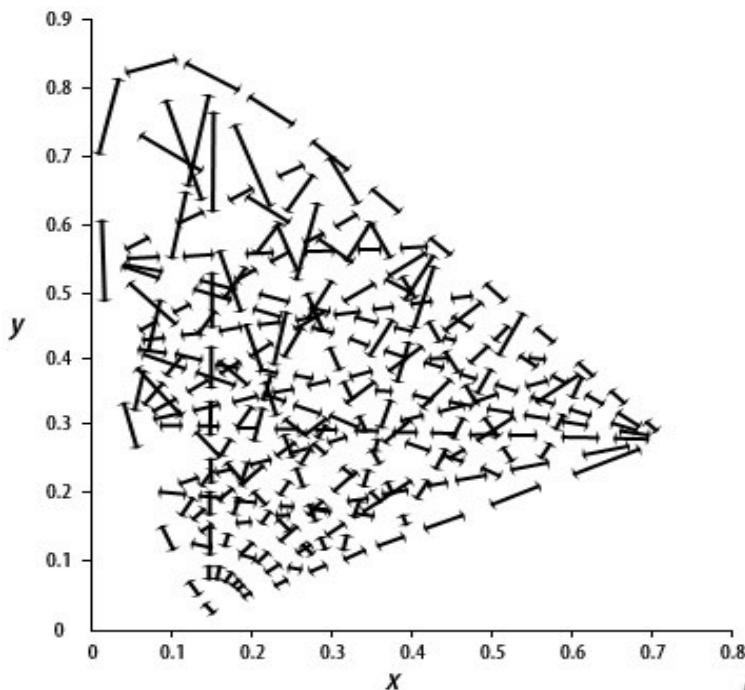


Diagrama de cromaticidade

# Cores

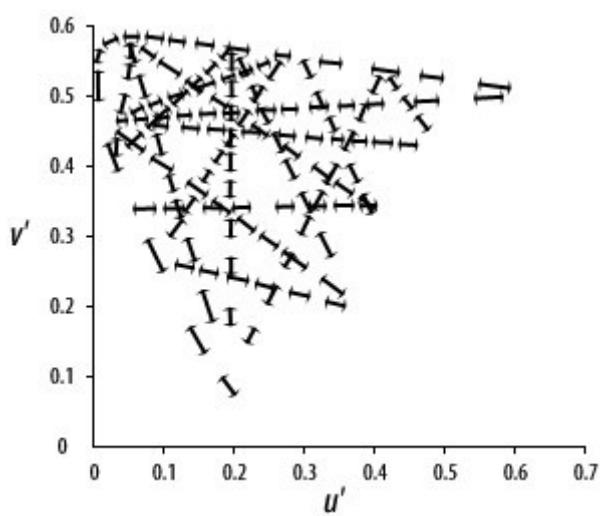
## Diagramas de cromaticidade

O diagrama de cromaticidade original é inadequado para representar a **acuidade** perceptiva de cores. Na figura abaixo os segmentos de reta indicam a distância mínima requerida entre duas cores para que sejam distinguidas. Note-se que a acuidade é máxima para os matizes azuis, menos intensa para os vermelhos e mínima na região do verde.



Acuidade perceptiva do diagrama  $xy$ .

Em 1975, a CIE propôs um novo mapeamento do espaço  $XYZ$  para o plano, no qual as cores são distribuídas de forma a exibirem graus de acuidade perceptiva de matizes aproximadamente homogêneos. Esse sistema é chamado de  $u'v'$  (vide figura abaixo).



Acuidade perceptiva do diagrama  $u'v'$ .

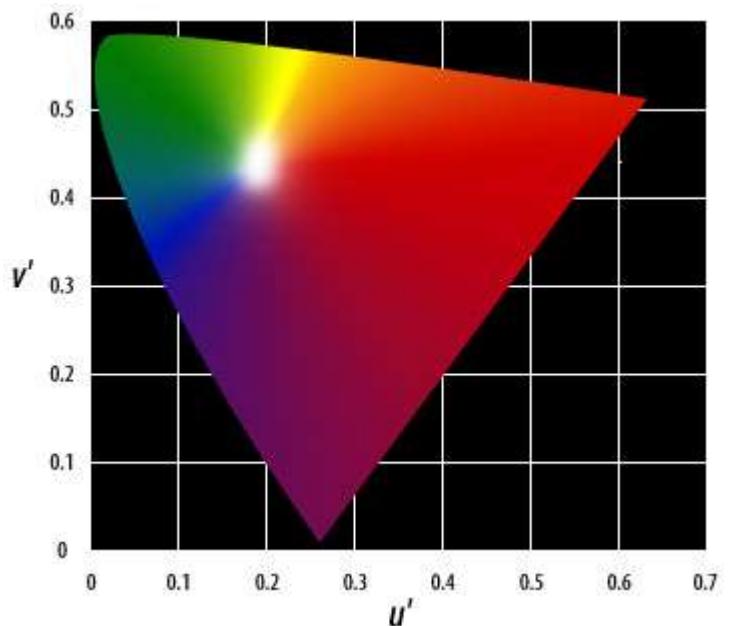


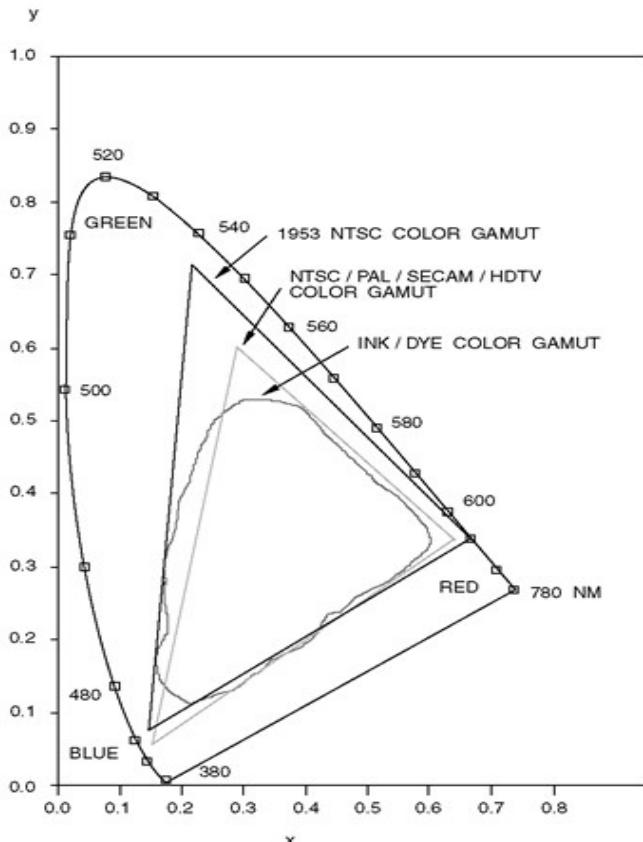
Diagrama  $u'v'$  de cromaticidade

# Cores

## O sistema *YIQ*

Proposto pela *NTSC* (National Television System Committee), para transmissão de imagens coloridas, baseia-se na triade:

- $Y$ : análoga à luminância do padrão *CIE*,
  - $I$  e  $Q$ : componentes que descrevem conjuntamente matiz e saturação.
- Como a banda de transmissão é limitada,  $Y, I, Q$  são definidos de modo a otimizar a percepção de luminância e dos matizes. Dessa forma,
- $Y$  recebe a maior parte (4,2MHz) da banda, pois a visão humana é mais sensível a variações de luminância do que de matiz;
  - $I$ , descriptor do eixo azul-laranja, recebe a segunda maior parcela da banda (1,5MHz), pois a visão humana tem maior acuidade nessa região.
  - $Q$ , descriptor do eixo verde-púrpura, recebe a parte remanescente da banda de transmissão (0,55MHz).



A transformação dos sinais *RGB* em *YIQ* é dada por:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,275 & -0,321 \\ 0,212 & -0,523 & 0,311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (\text{eq.37})$$

As cores geradas no sistema *NTSC*, assim como em todos os demais (*PAL* e *SECAM*, por exemplo), **não** abrangem a gama de cores definidas no padrão *CIE*, mas apenas um certa região triangular do diagrama de cromaticidade (vide figura ao lado). O mesmo ocorre com o sistema *CMY* (região interna ao triângulo *NTSC*).

# Imagen digital

## Definição

Representação numérica de uma imagem (analógica) na forma de uma matriz  $m \times n$  de pixels ('picture elements'), quantificados em uma escala discreta de tons de cinza ou triâdes *RGB* associadas às intensidades vermelho, verde e azul.

## Grades

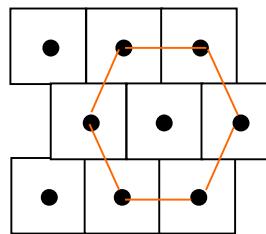
- Retangular: Facilita a representação computacional da imagem.
- Hexagonal: Sua estrutura geométrica é aproximadamente isotrópica. Requer 13% a menos de pixels para a digitalização de uma imagem. Além disso, apresenta conectividade não ambígua (exibe centro de simetria), o que facilita a aplicação de operadores espaciais.

•	•	•
•	•	•
•	•	•

Grade retangular

8	4	8
4	$\times$	4
8	4	8

Conectividades:  
'4' ou '4 e 8'



Grade hexagonal

•	6	6
6	$\times$	6
•	6	6

Conectividade: 6

## Etapas da digitalização

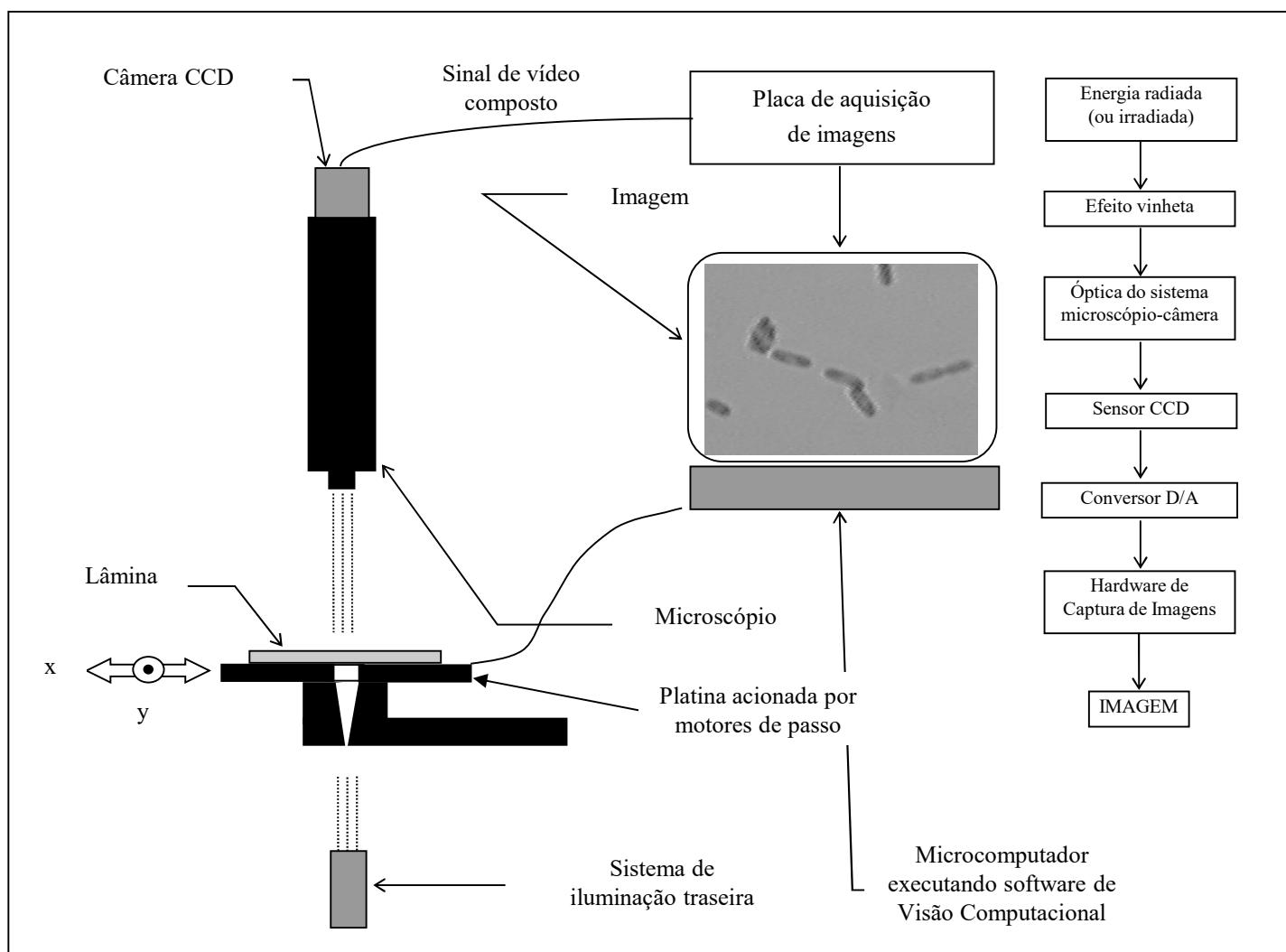
- Varredura e filtragem: depois de gerada, a imagem analógica é varrida por uma abertura óptica (filtro passa-baixa), atenuando as altas freqüências espaciais.
- Amostragem espacial: a intervalos espaciais discretos ao longo de  $x$  e  $y$ , selecionam-se amostras do sinal filtrado pela abertura óptica, transformando-as em um sinal analógico.
- Digitalização e quantização: o sinal analógico gerado na etapa anterior é digitalizado e mapeado para uma escala discreta de  $n$  tons de cinza (ou  $n$  triâdes de tons de cinza se a imagem for colorida).

# Imagen digital

## Captura optoeletrônica da cena

Admitiremos que a imagem analógica seja obtida a partir de um aparato de aquisição de imagens constituído por: sistema de iluminação, microscópio, câmera digital e microcomputador, conforme ilustrado na figura abaixo.

É importante salientar que, nas câmeras atuais, a placa de aquisição de imagens já está incorporada no hardware da câmera.



# Imagen digital

## Varredura e filtragem óptica

As lentes do sistema de captura de imagens são caracterizadas por uma **função de espalhamento de ponto (PSF, de ‘Point Spread Function’)**, ou seja, a resposta óptica à imagem de um ponto ideal.

Abaixo mostram-se representações 3D da *PSF* e da *LSF* (‘Line Spread Function’) de um sistema se lentes, bem como imagens típicas de *PSF*’s.

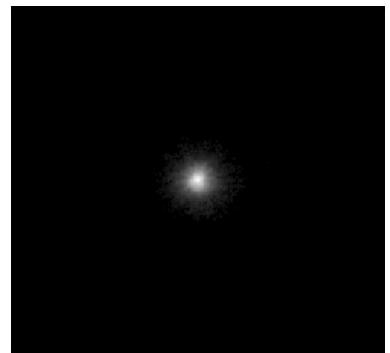
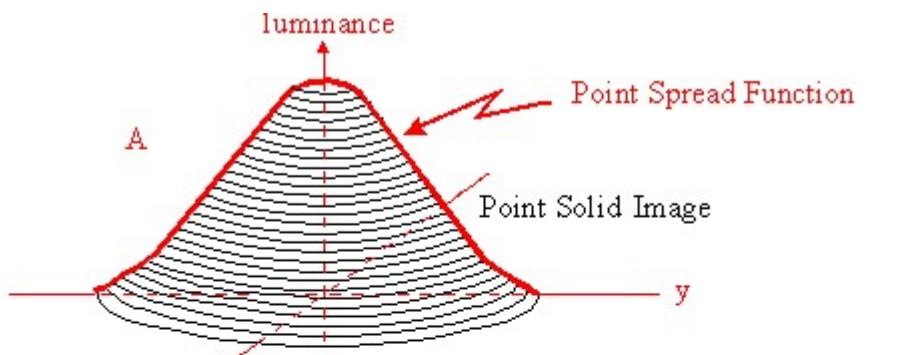
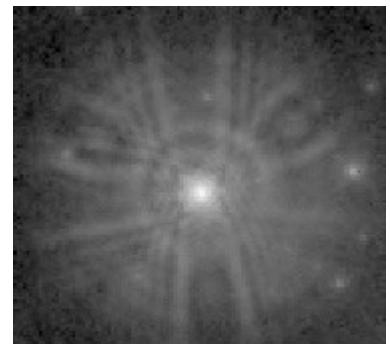
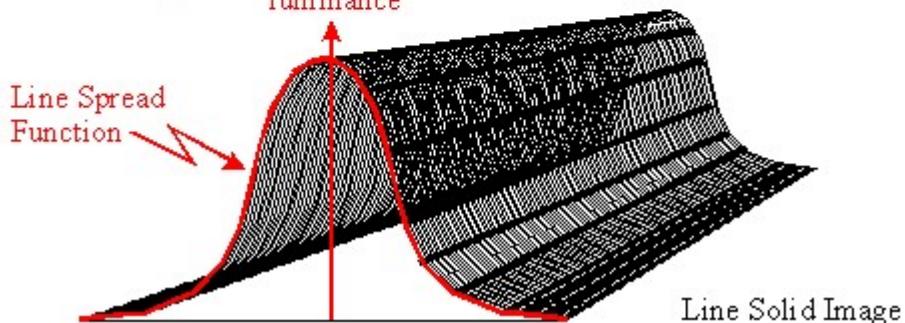


Imagen da *PSF* de um sistema óptico



Representações 3D da *PSF* e *LSF* de um sistema óptico.

; Imagen da *PSF* do sistema de lentes do telescópio Hubble.

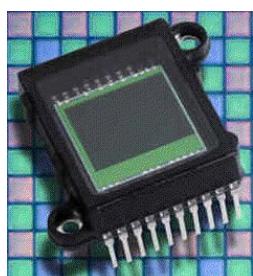
# Imagen digital

## Amostragem espacial

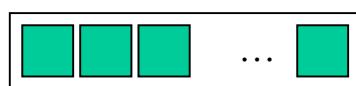
Utilizam-se câmeras *CCD* ou *CMOS*, baseadas em uma matriz de semicondutores fotossensíveis que acumulam cargas proporcionais à intensidade da luz incidente.

A intervalos de tempo controlados, a carga é transferida e convertida em um sinal de voltagem de acordo com as normas vigentes (RS-170, CCIR etc).

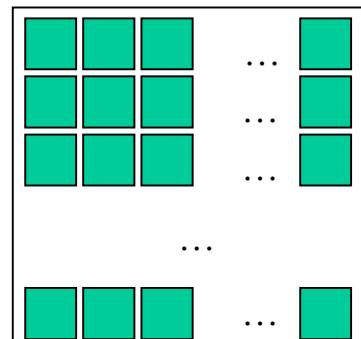
As câmaras *CCD* podem apresentar geometria linear ou matricial. No primeiro caso, o pente de *CCD's* é transportado por um servomecanismo ao longo da cena. No segundo, a imagem é projetada sobre toda a área da matriz.



(a) CCD



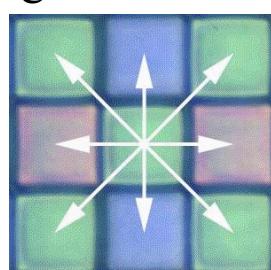
(b) Modelo linear



Modelo matricial

A maioria das matrizes de fotossensores apresentam razões de aspecto variando entre 1:1 e 1:1,5. As mais utilizadas têm 12,5 mm de largura, razão de aspecto 1:1 e resolução de  $512 \times 512$  pixels.

Imagens coloridas são geradas com: a) 3 fotossensores/pixel; b) 1 fotossensor/pixel e 3 exposições (*R, G, B*) por ciclo de ‘shutter’; c) para cada pixel, uma matriz de 9 fotossensores recobertos por filtros *R, G, B* conforme ilustrado na figura abaixo.

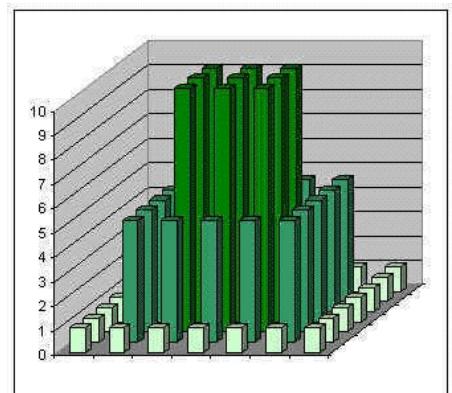


Filtro-mosaico

# Imagen digital

## Amostragem espacial

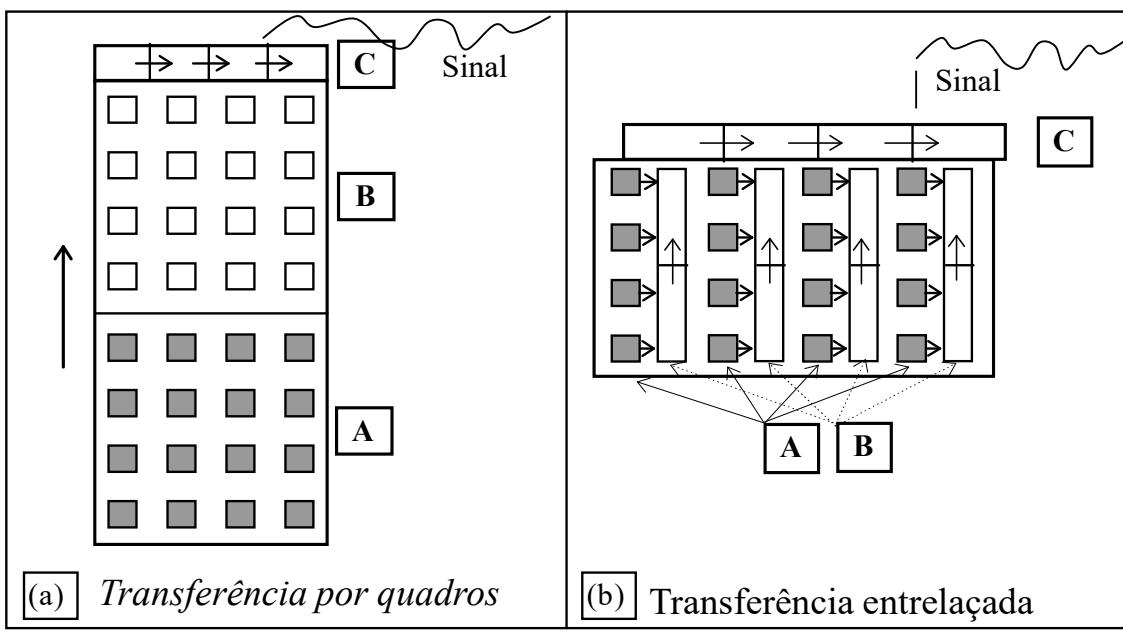
A cada ciclo de ‘shutter’ (freqüências típicas: 60Hz a 800 kHz), cada fotossensor possui uma carga pronta para ser convertida em tensão (vide figura ao lado).



Cargas armazenadas na matriz de fotossensores

A transferência de cargas (vide figura abaixo) pode ser:

- quadro-a-quadro (‘progressive scan’): após a integração temporal do sinal espacial, a matriz  $A$  é transferida para a área  $B$ , que é lida, linha a linha, pelos registradores  $C$ ;
- entrelaçada: apresenta registradores de leitura  $C$  e uma matriz  $A$  assimétrica de sensores adjacentes, coluna a coluna, a um conjunto de registradores  $B$ , os quais são compartilhados alternadamente, em tempos distintos, por sensores de duas linhas contíguas, ímpar e par.

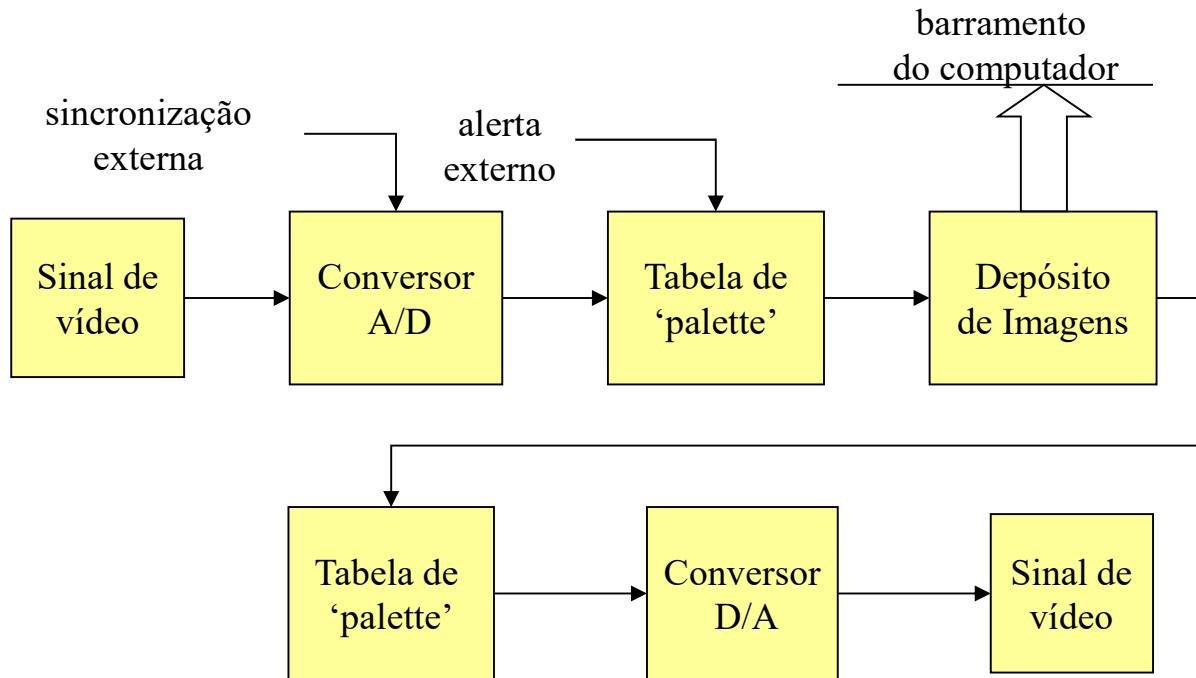


(menos sensível, mais densa)      (mais sensível, menos densa)

# Imagen digital

## Digitalização e quantização

O sinal de vídeo gerado na etapa anterior é digitalizado de acordo com o esquema indicado abaixo.



A conversão A/D (em geral 8 ou 12 bits) utiliza sinais de sincronismo que marcam o início e o término das linhas (e dos campos, se o sinal for entrelaçado).

Tabelas de 'palette' podem ser utilizadas tanto na entrada como na saída. Após o mapeamento, a imagem é transferida para a memória do computador.

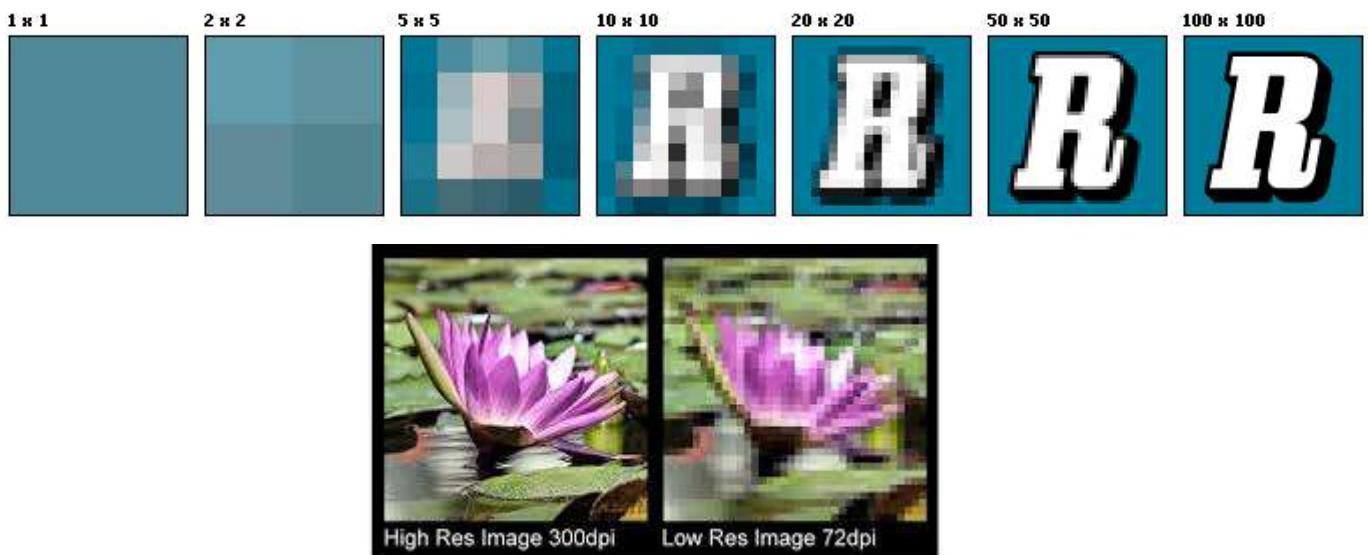
Sinais de sincronismo externo podem ser utilizados durante o processo.

# Imagen digital

## Resolução da imagem digital

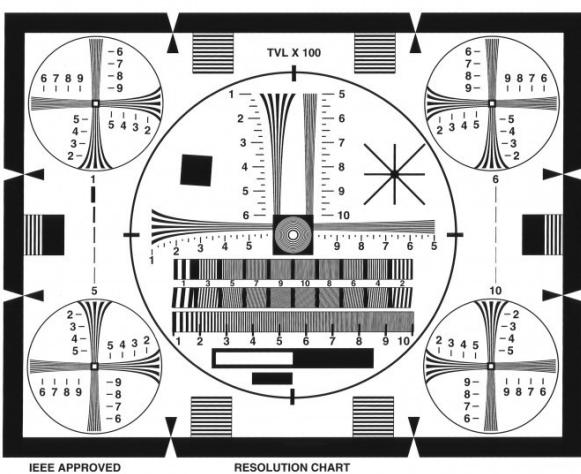
É uma medida (linhas/mm, pixels × pixels, dpi etc) que busca expressar o conteúdo de detalhes de uma imagem.

Nos exemplos abaixo mostra-se como a resolução é uma característica essencial à compreensão da cena.

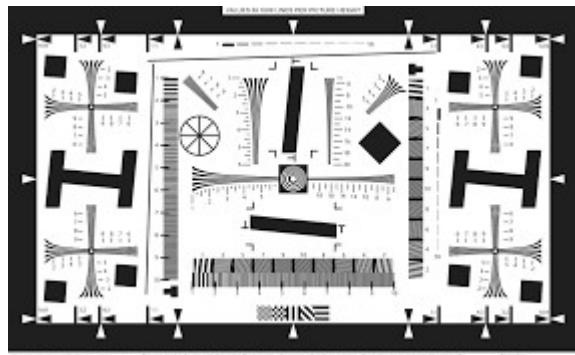


## Padrões de medida de resolução

Alguns desses padrões são apresentados abaixo.



Padrão IEEE

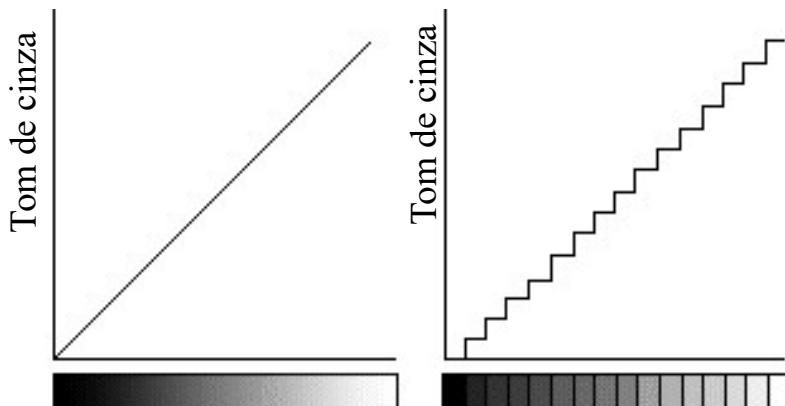


Padrão IRA/ISO

# Imagen digital

## Quantização

Consiste na transformação dos níveis contínuos de intensidade luminosa da imagem em níveis discretos tomados em um intervalo pré-definido.



(a) Perfil contínuo

(b) Perfil discreto com 16 tons de cinza

## Escala de quantização

Para mapear o intervalo de intensidade luminosa  $[0, I_{max}]$  em  $[0, 2^n]$  **não** se adota uma escala linear, ou seja: o tom de cinza  $c_m$  não é calculado a partir de

$$c_m = \frac{I_{max}}{2^n} \quad (\text{eq.38})$$

mas sim a partir de

$$c_m = m \left( \frac{1}{I_0} \right)^{\frac{1}{2^n}} \quad (\text{eq.39})$$

em que  $I_0$  é o menor fração (entre 0 e 1) de intensidade luminosa emitida pela fonte.

Essa escala logarítmica de quantização é utilizada para atender à percepção visual humana, que enxerga igualmente bem variações entre  $0,10I_{max}$  e  $0,11I_{max}$  ou entre  $0,50I_{max}$  e  $0,55I_{max}$ .

Nessa escala, os tons de cinza  $0,1,2, \dots, 2^n$  correspondem a:

$$I_0, rI_0, \dots, r^{2^n}I_0, \text{ com } r = \left( \frac{1}{I_0} \right)^{\frac{1}{2^n}} \quad (\text{eq.40})$$

# Imagen digital

## Efeito da quantização

O nível de quantização utilizado para representar os tons de cinza da imagem afeta significativamente a percepção de detalhes.

No exemplo abaixo os níveis utilizados foram: 16, 8, 4 e 2 bits/pixel.

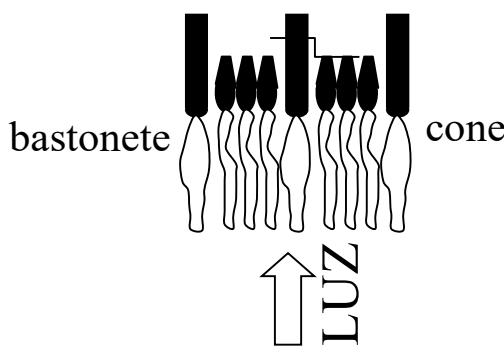
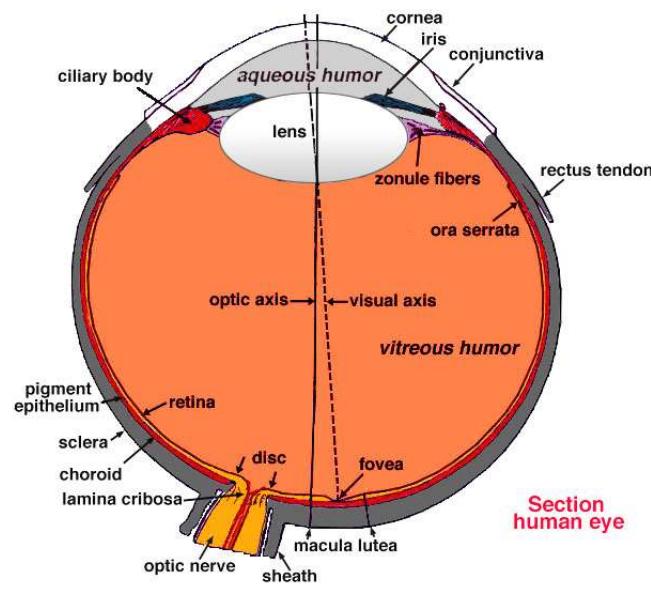


Imagen com diferentes níveis de quantização.

Após a quantização, a imagem é transformada em uma matriz de números inteiros pertencentes ao intervalo  $[0, 2^n]$ , onde  $n$  é o número de bits/pixel. Na figura abaixo mostra-se uma imagem quantizada.

99	71	61	51	49	40	35	53	86	99
93	74	53	56	48	46	48	72	85	102
101	69	57	53	54	52	64	82	88	101
107	82	64	63	59	60	81	90	93	100
114	93	76	69	72	85	94	99	95	99
117	108	94	92	97	101	100	108	105	99
116	114	109	106	105	108	108	102	107	110
115	113	109	114	111	111	113	108	111	115
110	113	111	109	106	108	110	115	120	122
103	107	106	108	109	114	120	124	124	132

# Visão humana



cones:  $7 \times 10^6$  (cores)  
bastonetes:  $120 \times 10^6$

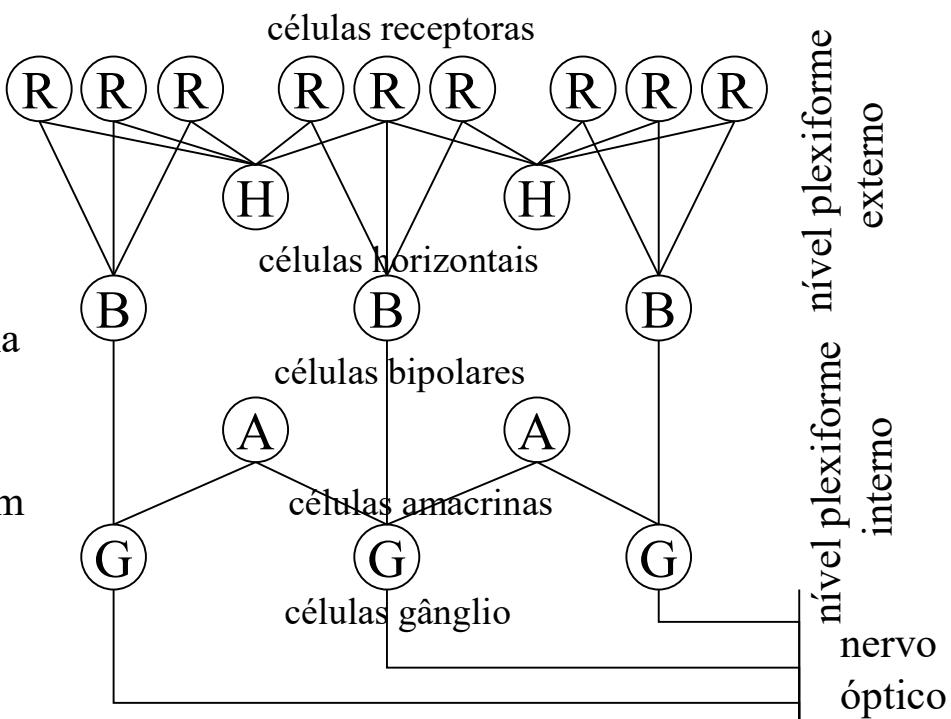
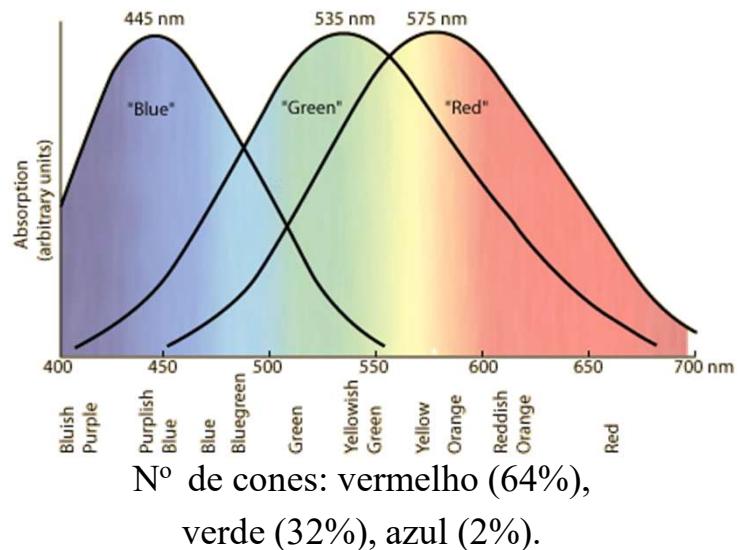
## Compressão de dados

Esse processo se realiza na retina:

$127 \times 10^6$  receptores  
células-gânglio se reduzem  
a  $1 \times 10^6$  receptores

## O olho humano

Atua como um sistema de lentes que projeta a imagem sobre a retina, que é a estrutura fotossensível.



# Visão humana

## Sensibilidade ao contraste

A resposta das células-gânglio a variações de luminância  $\Delta I$  é não-linear,

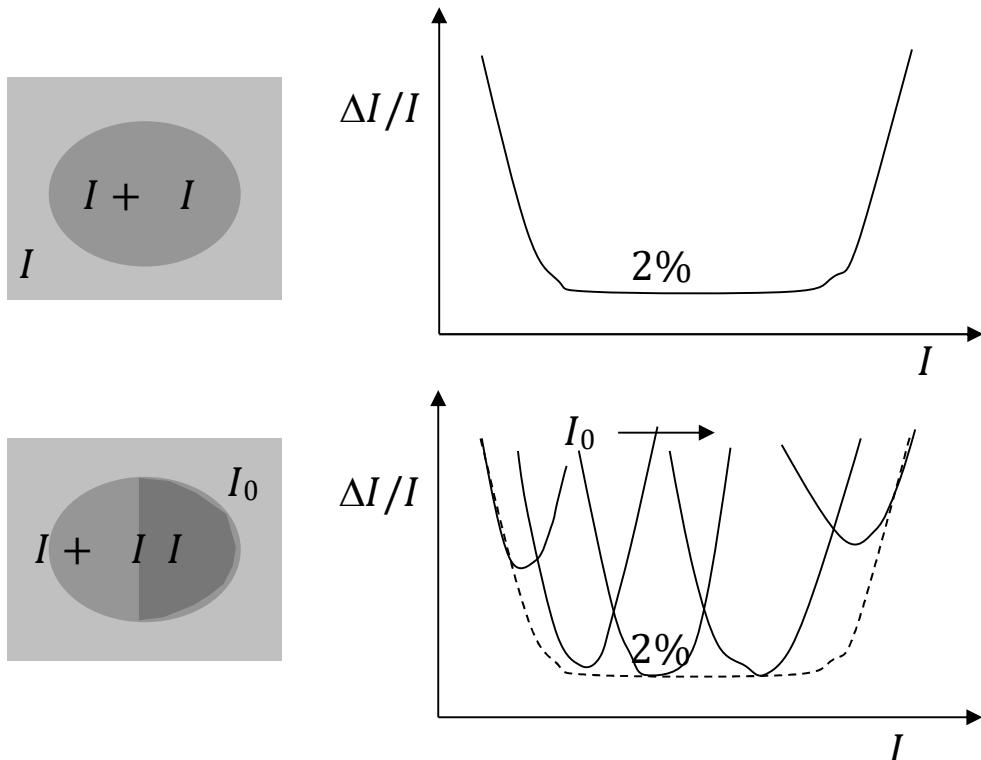
o que desce às leis de Weber, a saber:

$\Delta I/I \approx 0,02$  em um grande intervalo de valores  $I$ .

Para valores baixos e altos de  $I$ , o valor de  $\Delta I/I$  deve ser alto.

A percepção do contraste depende do ‘background’:  $\Delta I/I$  é tanto maior quanto mais escuro ou mais claro for o ‘background’.

O envelope de sensibilidade ao contraste não se altera com o ‘background’.

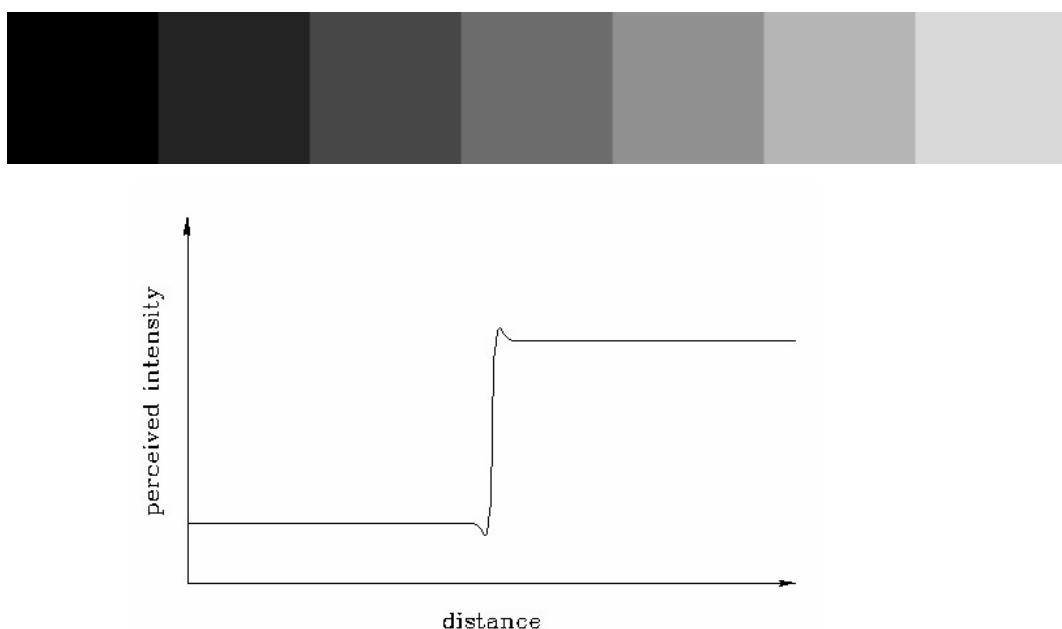


Como  $d(\log I) = dI/I$ , conclui-se pela conveniência de se aplicar os algoritmos de processamento de imagens ao **logaritmo** da intensidade luminosa (ou seja,  $\log I$ ) se se pretende que a imagem resultante seja tão perceptível à visão humana quanto a imagem original.

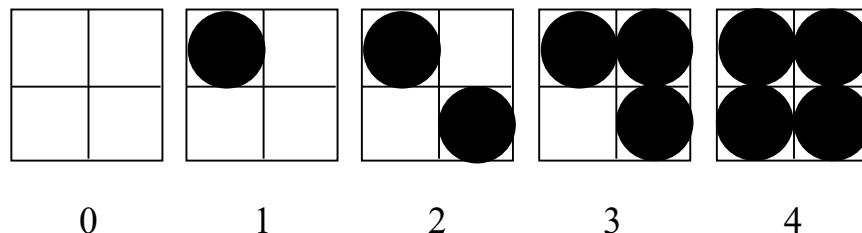
# Visão humana

## Exemplos de processamentos neuro-ópticos

1. Bandas de Mach: A percepção visual humana de uma seqüência de quadrados com níveis homogêneos e linearmente crescentes introduz um realce da intensidade das bordas, conforme ilustrado na figura abaixo.

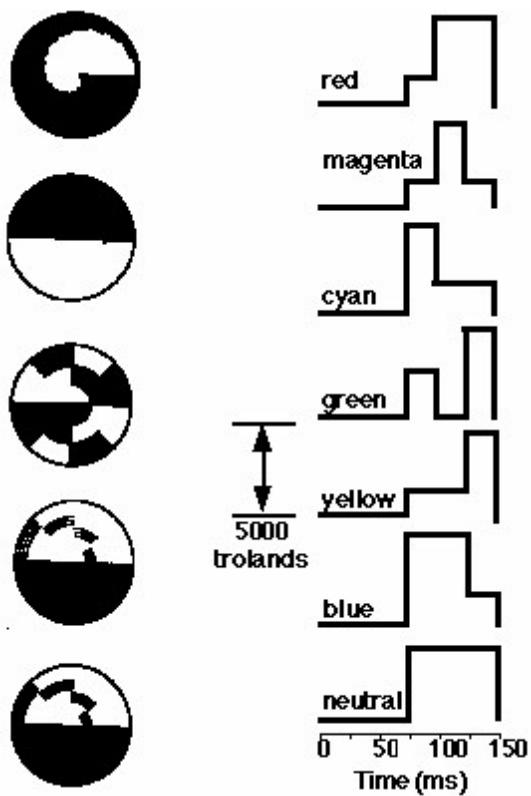


2. Aproximação por meios-tons: O olho humano interpola padrões preto / branco, transformando-os em tons de cinza ('halftoning'), conforme ilustrado a seguir.



## Visão humana

3. Cores subjetivas: Experimentos realizados por Land, Fechner e Helmholtz, utilizando discos giratórios com padrões branco/preto, indicaram que o observador associa o sinal temporal à percepção de uma certa cor, denominada, por essa razão, **cor subjetiva**. A figura a seguir ilustra esse fenômeno.



## Visão humana

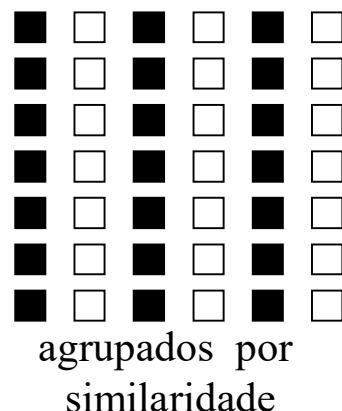
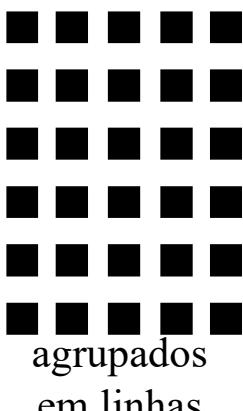
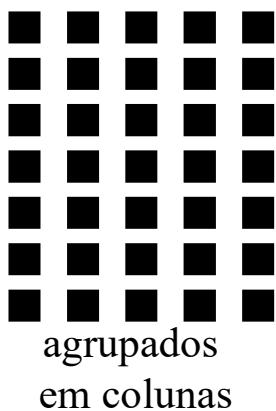
4. Visão estereoscópica: a resolução da disparidade entre duas observações monoculares da mesma cena cria a ilusão de tridimensionalidade. Os estereogramas de Julesz abaixo ilustram esse fenômeno.



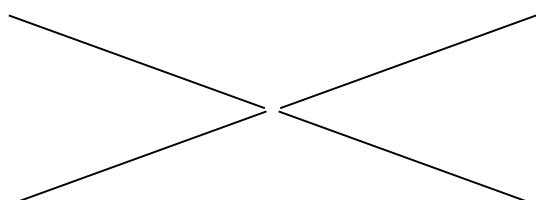
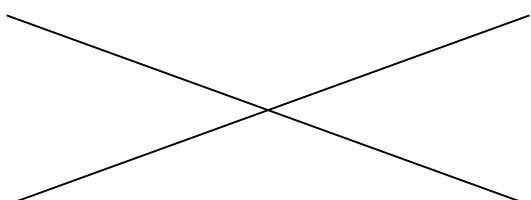
Estereogramas de Julesz: devem ser observados  
com óculos 3D.

# Visão humana

5. Leis da ‘gestalt’ (forma, padrão). Descrevem a organização da informação visual no cérebro. São elas:
- Lei da proximidade: Objetos mais próximos tendem a ser agrupados.
  - Lei da similaridade: Objetos similares tendem a ser agrupados.



- Lei do movimento comum: Objetos que têm a mesma lei de movimento tendem a ser agrupados.
- Lei do continuum: Objetos são agrupados para preservar o continuum das formas.



A interpretação à esquerda é a preferível (lei do continuum).

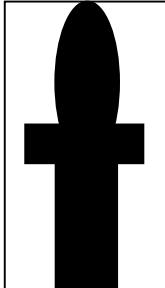
- Lei da clausura: As formas fechadas têm preferência sobre as abertas.
- Lei da simetria: Áreas simétricas são percebidas como objetos sobre um ‘background’.
- Lei de Pragnanz: Das várias interpretações possíveis, prefere-se a mais simples e estável.

# Visão humana

- Lei da clausura: As formas fechadas têm preferência sobre as abertas.
- Lei da simetria: Áreas simétricas são percebidas como objetos sobre um ‘background’.



A interpretação preferível é a de um quadrado ocluído por um losango inteiramente branco (lei da clausura).



A interpretação preferível é a de que existe um objeto preto sobre um ‘background’ branco (lei da simetria).

# Visão humana

## Computador×cérebro

Abaixo faz-se uma breve comparação entre os recursos computacionais do cérebro humano e de um computador.

- Velocidade de processamento

Computador: Freqüência de ‘clock’ da ordem de GHz.

Cérebro: Freqüência de processamento de um neurônio da ordem de kHz.

- Ordem de processamento

Computador: Tipicamente serial.

Cérebro: Massivamente paralela.

- Eficiência energética

Computador: 1

Cérebro: 100.000.

- Abundância e complexidade

Computador: Os maiores supercomputadores têm cerca de  $10 \times 10^6$  núcleos.

Cérebro: Contém cerca de  $86 \times 10^9$  neurônios, cada qual apresentando cerca de 1000 conexões interneuronais.

- Armazenamento

Computador: As informações são armazenadas em endereços fixos da memória; ocorre substituição de informação;

Cérebro: As informações são armazenadas nas interconexões dos neurônios; preserva-se conhecimento pregresso mediante adaptação das sinapses.

- Controle do processamento

Computador: ACPU é o ponto de estrangulamento do processo.

Cérebro: O processamento é totalmente distribuído.

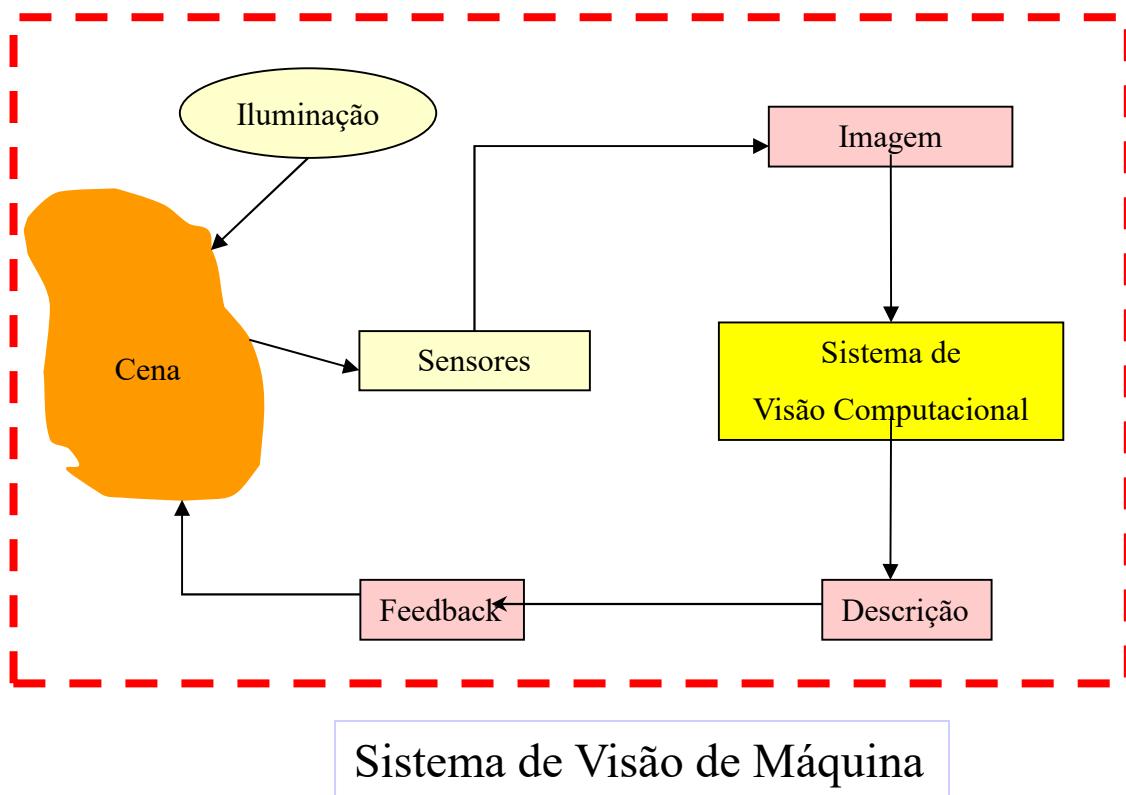
# Visão Computacional

## Escopo da Visão Computacional

Gerar informações úteis e descrições de alto nível de cenas representadas por meio de uma imagem, grupo de imagens ou seqüência temporal de imagem.

## Escopo da Visão de Máquina

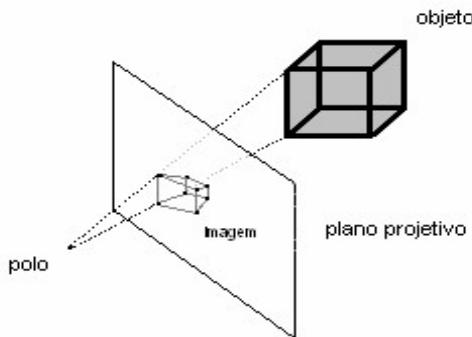
Utilizar as informações e descrições geradas por um sistema de visão computacional como realimentação de uma malha de controle.



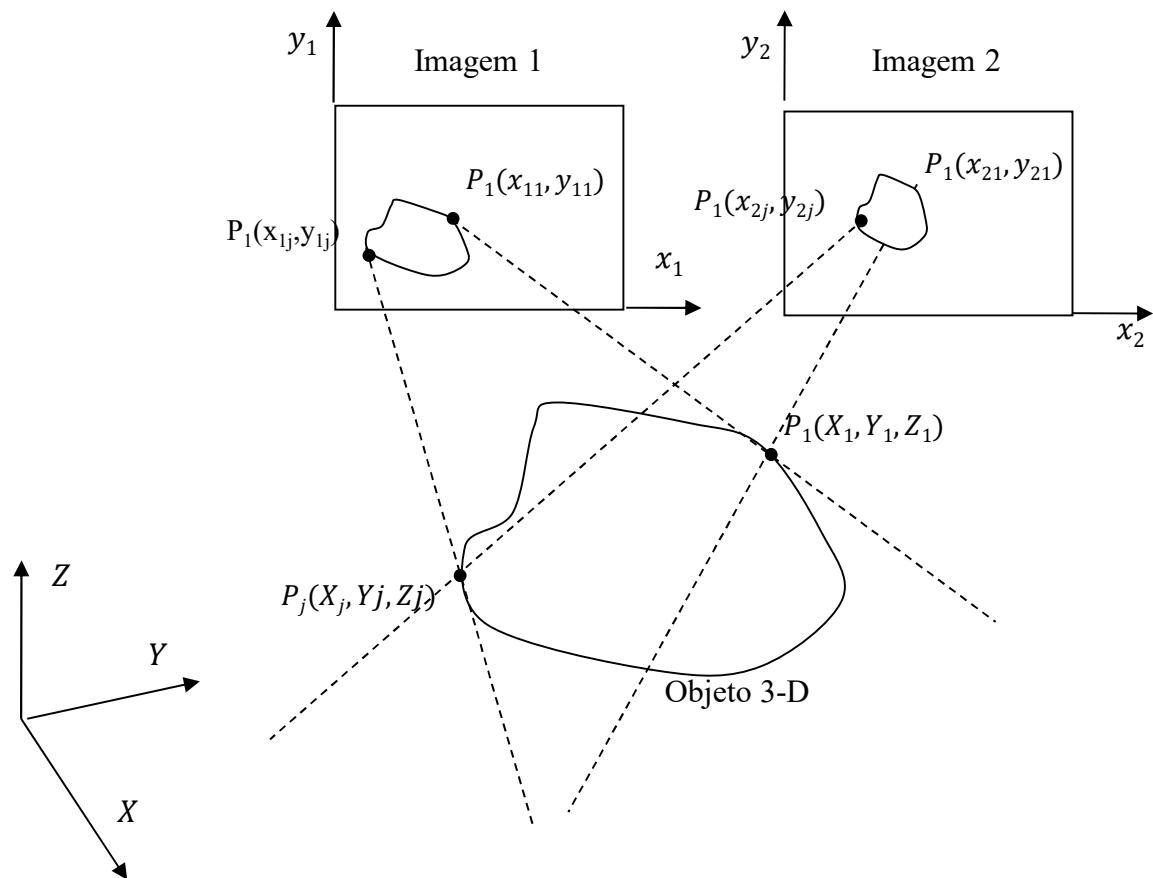
# Visão Computacional

## Algumas dificuldades inerentes à visão computacional

### 1. Medidas baseadas em geometria projetiva



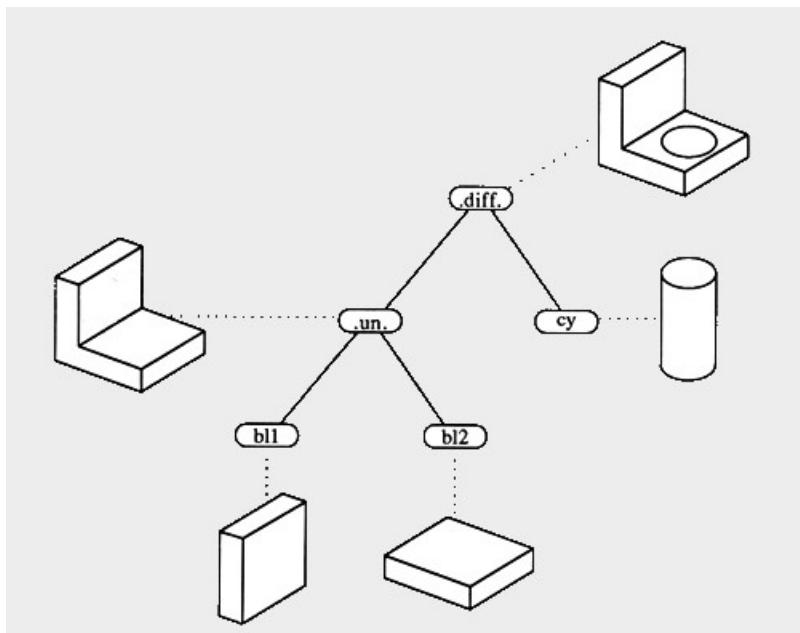
2. Utilização de técnicas estereoscópicas, quais sejam:  
Identificação de pontos característicos.  
Localização de pontos correspondentes.  
Determinação do mapa de disparidade.  
Estimação das coordenadas 3D.



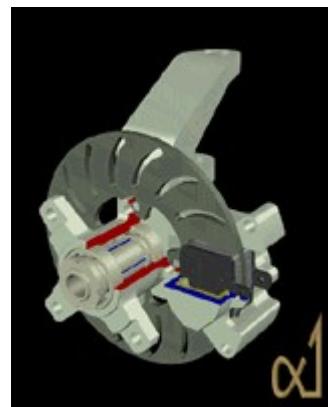
# Visão Computacional

## Algumas dificuldades inerentes à visão computacional

3. Organização e síntese da informação: objetos de computação gráfica são construídos a partir de um rigoroso processo topológico-geométrico (*CSG*, por exemplo).



Uso de primitivas volumétricas para construção do modelo 3D



Representação canônica de objetos 3D

O problema inverso – identificação da estrutura geométrica de um objeto a partir da análise de um par estereoscópico de imagens, apresenta complexidade muitíssimo maior.

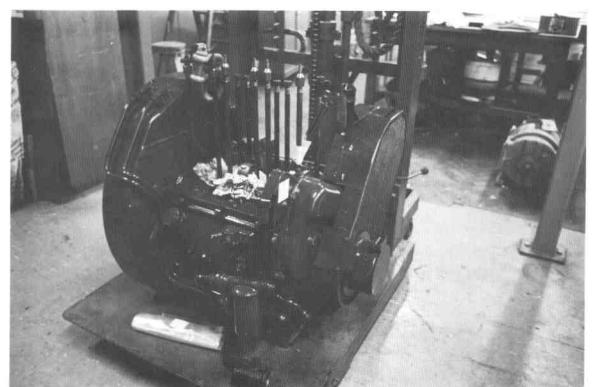


Figure 44(c) - A 3 Cylinder Engine Block before Conversion  
(as many of the original components as possible are used).

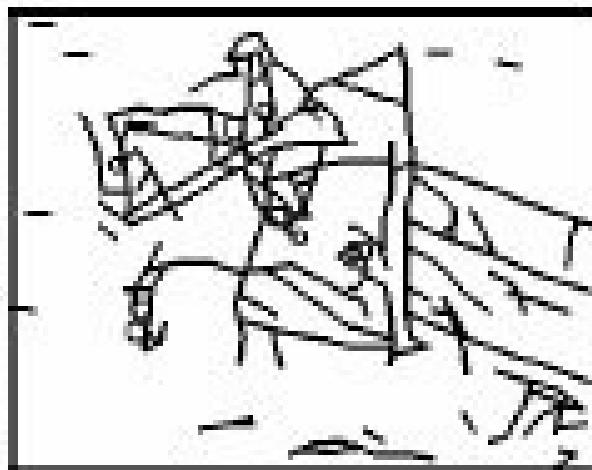
# Visão Computacional

## Algumas dificuldades inerentes à visão computacional

4. Deficiências da representação baseada em estruturas elementares (bordas e vértices, por exemplo), de modo que: 1) é muito difícil organizar essas estruturas; 2) é muito difícil atribuir significado a essas estruturas. O exemplo abaixo é esclarecedor.



Imagen original

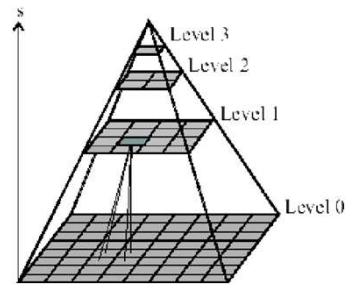


Esboço primal

5. Percepção de detalhes: é necessário armazenar e processar várias representações da imagem, em diferentes escalas espaciais. Em geral utilizam-se pirâmides gaussianas, conforme ilustrado a seguir.



GAUSSIAN PYRAMID



# Visão Computacional

## Algumas dificuldades inerentes à visão computacional

6. Mal-condicionamento: pequenas perturbações, como: a) alterações na intensidade da fonte luminosa e/ou em sua orientação, b) ruídos gerados durante a captura da imagem, translações, c) rotações e mudanças de escala, podem afetar significativamente a resolução do problema.

Ex: Considere-se o problema de reconhecimento de caracteres abaixo, em que, na linha superior, se encontram caracteres-referência e, na inferior, caracteres-teste.

Abordagem: Admite-se que o ‘training set’ seja constituído por bits dispostos em uma grade  $5 \times 5$ .



*training set*



caracteres a serem identificados

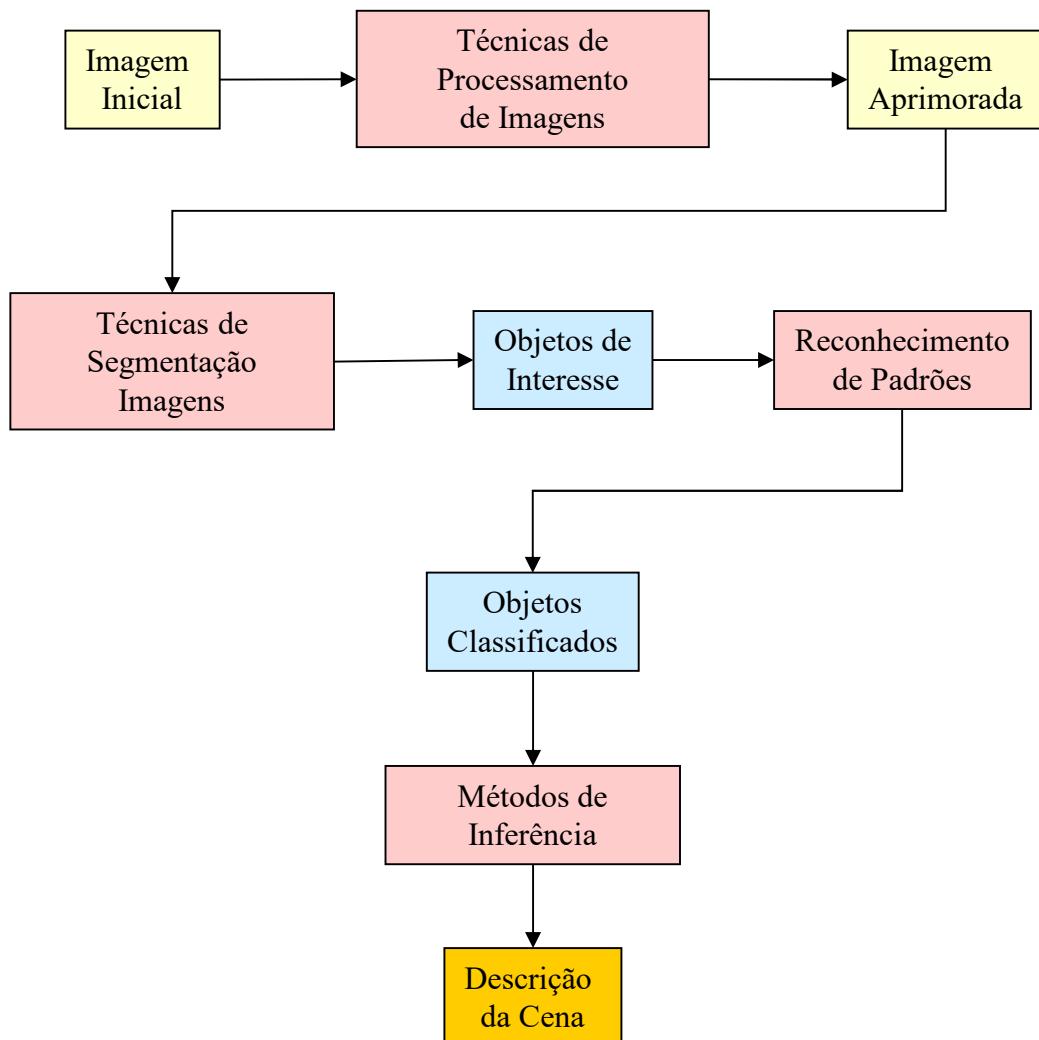
Constata-se que

- O padrão  $C$  é o único para o qual  $\text{referência} = \text{teste}$ .
- O padrão  $B$  é tal que  $\text{referência}$  e  $\text{teste}$  diferem em um único pixel.
- Para o padrão  $I$ ,  $\text{referência}$  e  $\text{teste}$  são iguais a menos de uma translação; em uma comparação pixel a pixel seriam considerados distintos.
- Há um padrão que poderia ser classificado como  $I$  ou como  $T$ .

# Visão Computacional

## Disciplinas envolvidas

- **Processamento de imagens:** produz, a partir de uma imagem, novas imagens com qualidades visuais e propriedades mais adequadas à observação e análise humanas.
- **Geometria Projetiva:** possibilita determinar as propriedades afim e de similaridade das imagens, bem como reconstruir modelos 3D a partir de pares estereoscópicos de imagens de uma mesma cena.
- **Estatística:** oferece inúmeros métodos úteis à identificação e classificação dos objetos de interesse da imagem processada.
- **Inteligência Artificial:** oferece métodos de inferência adequados à identificação de relações entre objetos da imagem, possibilitando a extração de conceitos e regras que possibilitam compreender a cena.



# Visão Computacional

## Processamento de imagens × visão computacional

- Processamento de imagens: Gera novas imagens a partir das imagens originais. Dentre seus métodos, podem-se citar:
  - a) Filtragem.
  - b) Restauração e correção de defeitos em imagens degradadas.
  - c) Ênfase e suavização de características das imagens.
  - d) Alteração do histograma de tons de cinza das imagens.
- Visão computacional: Identifica **objetos característicos** nas imagens e realiza inferências de modo a extrair da cena significados e descrições. Dentre suas técnicas características, destacam-se:
  - a) Métodos de segmentação de imagens
  - b) Métodos de identificação de padrões.
  - c) Métodos de Geometria Projetiva.
  - d) Métodos de estimativa de campos de velocidades a partir da análise de seqüências de imagens.

## Computação gráfica × visão computacional

- **Computação Gráfica:** Disciplina eminentemente algorítmica. Produz uma representação gráfica de um objeto a partir de um modelo matemático representativo de suas características geométricas e fotométricas, em conformidade com um dado observador.
- **Visão Computacional:** Por operar sobre sistemas complexos, utilizando informação incompleta, necessita de heurísticas e conhecimento específico do domínio para identificar características geométricas e relações a partir de conjuntos de imagens.

# Visão Computacional

## Tópicos de visão computacional

- **Geração de imagens**
  - a) Formação de imagens;
  - b) Colorimetria;
  - c) Hardware de captura de imagens;
  - d) Representação, armazenamento e compressão.
- **Digitalização de imagens**
  - a) Transformada de Fourier;
  - b) Teorema de Shanon;
  - c) *Aliasing*;
  - d) Reconstrução.
- Aprimoramento de imagens
  - a) Transformações de ponto;
  - b) Filtros espaciais;
  - c) Filtros no domínio das freqüências.
- **Restauração de imagens**
  - a) Modelo de degradação.
  - b) Modelos de ruído.
  - c) Filtros espaciais.
  - d) Filtros no domínio das freqüências: filtro inverso; filtro de Wiener.
  - e) Distorção geométrica.
- **Detecção de bordas, vértices, retas e curvas**
  - a) Operadores baseados em gradiente.
  - b) Operadores baseados em momentos.
  - c) Transformada de Hough.
  - d) Rastreamento de bordas.

# **Visão Computacional**

- **Métodos de segmentação de imagens**
  - a) Métodos globais de ‘limiarização’.
  - b) Métodos locais de ‘limiarização’.
  - c) Rotulação.
  - d) Crescimento de regiões.
  - e) Técnicas de relaxação.
- **Processamento de imagens binárias**
  - a) Morfologia matemática.
  - b) Métricas e propriedades geométricas características.
- **Métodos de classificação de padrões**
  - a) Classificadores estatísticos,
  - b) ‘Cluster analysis’,
  - c) Métodos sintáticos.
- **Redes neurais**
  - a) Percéptrons;
  - b) MLP.
  - c) Kohonen.
  - d) ART.
- **Algoritmos indutivos**
  - a) ID3.
  - b) C5-2.
- **Estimação do movimento**
  - a) Fluxo óptico.
  - b) Registro de imagens.
  - c) Correlação cruzada.
- **Reconstrução de formas tridimensionais**
  - a) Calibração de câmeras.
  - b) Estereoscopia.
  - c) Fotogrametria.
  - d) Uso de iluminação estruturada.
  - e) Tomografia.

# **Visão Computacional**

## **Campos de aplicação**

- Reconhecimento de caracteres impressos e manuscritos.
- Interpretação de imagens de diagramas e desenhos de engenharia.
- Interpretação de imagens de partituras musicais.
- Interpretação de imagens de formulários.
- Reconhecimento de assinaturas.
- Restauração e pós-produção de filmes.
- Reconhecimento de impressões digitais.
- Reconhecimento de íris.
- Reconhecimento de faces humanas.
- Caracterização do movimento de indivíduos.
- Contagem de população em aglomerados humanos.
- Inspeção de placas de circuito impresso.
- Inspeção de juntas de solda.
- Inspeção de tecidos.
- Inspeção de frutas e alimentos industrializados.
- Visão robótica.
- Engenharia reversa baseada em contornos do objeto.
- Medição e metrologia dimensional.
- Reconstrução de cenas para produção cinematográfica.
- Caracterização morfológica de células.
- Análise de cromossomos.
- Estereologia de células e tecidos de seres vivos.
- Detecção de células cancerígenas.

# **Visão Computacional**

## **Campos de aplicação**

- Sistemas de apoio visual à cirurgia.
- Reconstrução de imagens de ultra-som.
- Identificação de aviões e mísseis.
- Reconhecimento e rastreamento de alvos.
- Sistemas de navegação.
- Reconhecimento de pista de pouso.
- Identificação e rastreamento de nuvens.
- Reconhecimento de rodovias em imagens de satélite.
- Mapeamento de terrenos.
- Caracterização de campos de produção agrícola.
- Sistemas de informação geográfica.
- Interpretação de imagens rádio-astronômicas.
- Vigilância de estacionamentos.
- Detecção e rastreamento de intrusos.
- Detecção de comportamentos anômalos.
- Análise de cenas em jogos esportivos.
- Monitoramento de tráfego rodoviário.
- Reconhecimento de placas de veículos automotores.

# **Visão Computacional**

## **Inspeção visual automática**

Tais sistemas têm dois objetivos principais:

- Tomar decisões sobre a similaridade ou não do produto com o assim chamado “golden template” (por exemplo: arquivos CAD).
- Coletar dados sobre as características de desempenho de um produto, através da identificação e contagem de defeitos.
- Classificar os defeitos encontrados.

## **Inspeção de placas de circuito impresso**

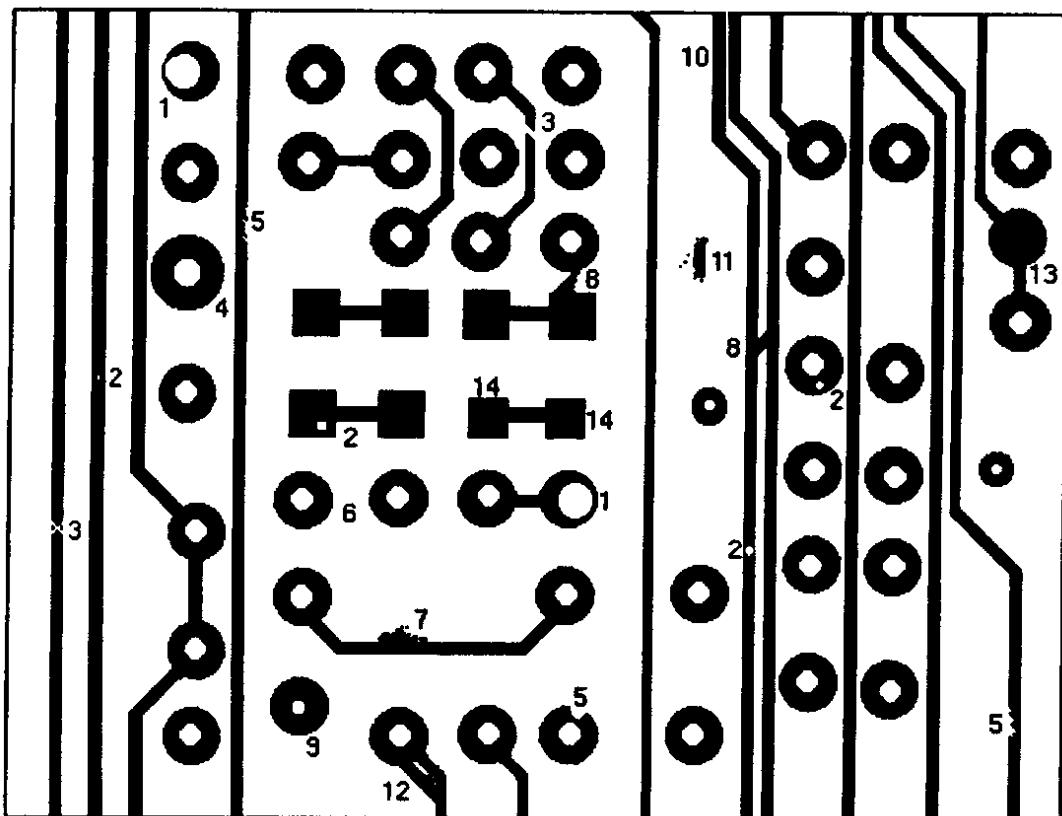
Nessa aplicação, a inspeção visual automática é essencial, pois o método tradicional de inspeção, além de tedioso

- requer o monitoramento visual de um grande número de etapas do processo (mais de 50).
- baseia-se em avaliações humanas, que são muito subjetivas.
- apresenta taxas de falha consideráveis: 10%, para um inspetor usando lente de aumento.
- é invasivo, capaz de introduzir novos defeitos no produto.
- é inaplicável em placas multicamada.
- não satisfaz às exigências de qualidade e produtividade requeridas pela indústria.
- é inadequado para a verificação dos atuais níveis de tolerância dimensional.

# Visão Computacional

**Defeitos típicos em placas de circuito impresso.**

Na figura abaixo ilustram-se alguns desses defeitos.



- 1.expansão
- 3.interrupção;
- 5.“mordida de rato”
- 7.“raiz”;
- 9.diâmetro errado
- 11.metal espúrio
- 13.furo ausente

- 2 .“furinho”
- 4 .sub-corrosão
- 6 .trilha ausente
- 8 .curto-circuito
- 10 .trilhas próximas
- 12 .curto-circuito
- 14 .super-corrosão

# **Visão Computacional**

## **Monitoração de tráfego**

Sistemas de inspeção visual automática são bastante indicados para realizarem as tediosas tarefas de:

- vigilância de ambientes reservados;
- monitoração de tráfego.

Em ambos os casos, opera-se com seqüências de imagens que devem ser analisadas, detalhadamente, apenas quando se detectam anormalidades.

Essa classe de aplicações requer a resolução dos seguintes problemas:

- captura e análise de imagens em tempo real
- detecção de intrusos na cena
- atualização do ‘background’
- reconhecimento de objetos de interesse nas imagens onde existe intrusão.

Nas figuras a seguir apresentam-se alguns exemplos de tarefas que devem ser executadas para o cumprimento das missões de um sistema de leitura automática de placas de veículos automotivos.

# Visão Computacional

Para analisar a intrusão na cena, é preciso investigar as variações de propriedades do fluxo óptico através de faixas de controle que emulam a função dos tradicionais ‘loops’ indutivos.

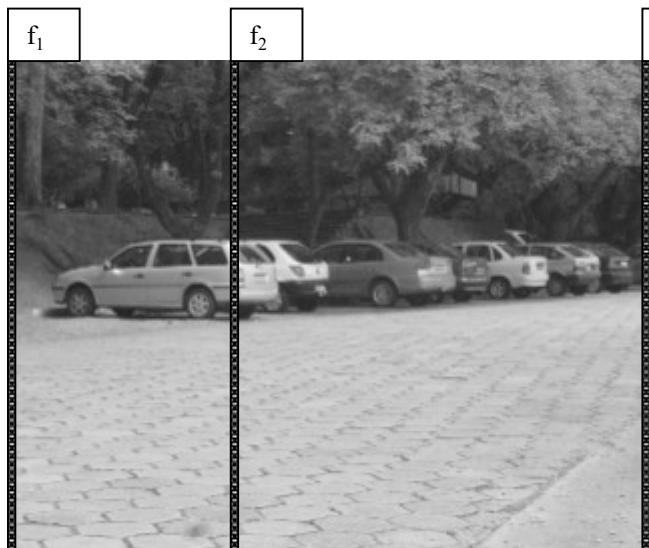


Imagen do *background*



Cena com intrusão nas faixas de controle  $f_1$  e  $f_2$

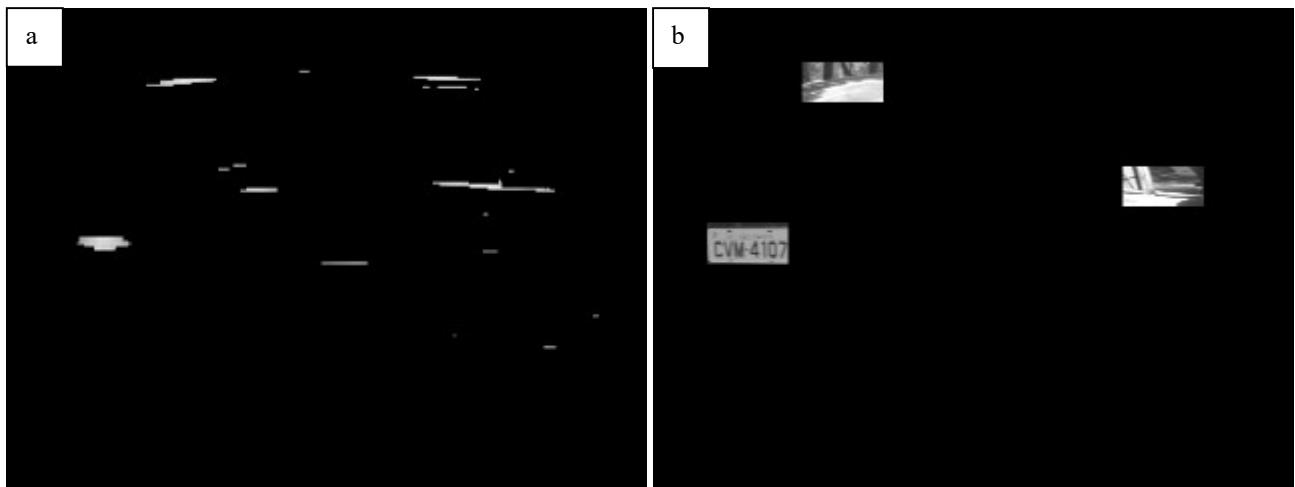
Diferentes métodos de estimação podem ser utilizados para recuperar o ‘background’ a partir de uma seqüência de intrusões.



Estimação do *background*: (a) via filtro de Kalman; (b) via filtro de mediana

# Visão Computacional

Para identificar a provável região da imagem onde se localiza a placa, é preciso investigar características da **textura** da imagem.



Análise das texturas: (a) Padrões locais de textura; (b) sub-imagens correspondentes

Aplicando o mesmo processo a uma seqüência de imagens, obtém-se um conjunto de sub-imagens da placa que podem ser **fundidas** em uma única imagem de **maior resolução**, visando-se, com isso, facilitar o subsequente processo de reconhecimento de caracteres.



Essas sub-imagens, fundidas, produzirão uma imagem de maior resolução.

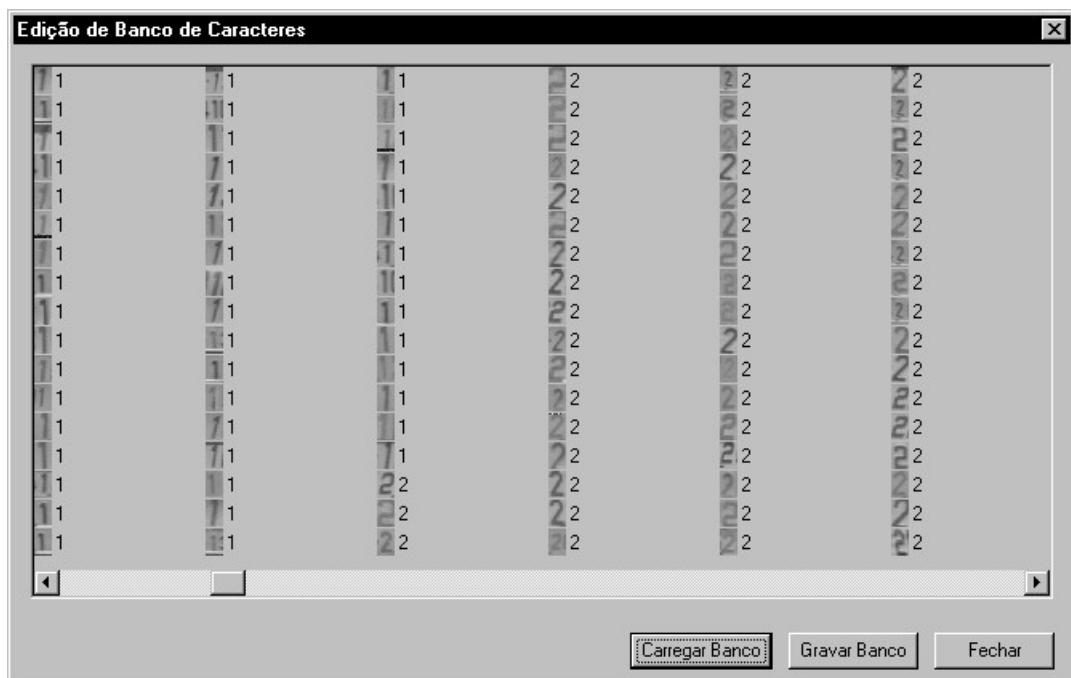
# Visão Computacional

**Métodos de segmentação** de imagens são aplicados à imagem de maior resolução, com o intuito de isolar os prováveis caracteres.

Sub-imagens dos mesmos são, finalmente, submetidas a uma rede neural treinada que realiza a identificação final.



Identificação das sub-imagens de caracteres, após processo de segmentação.



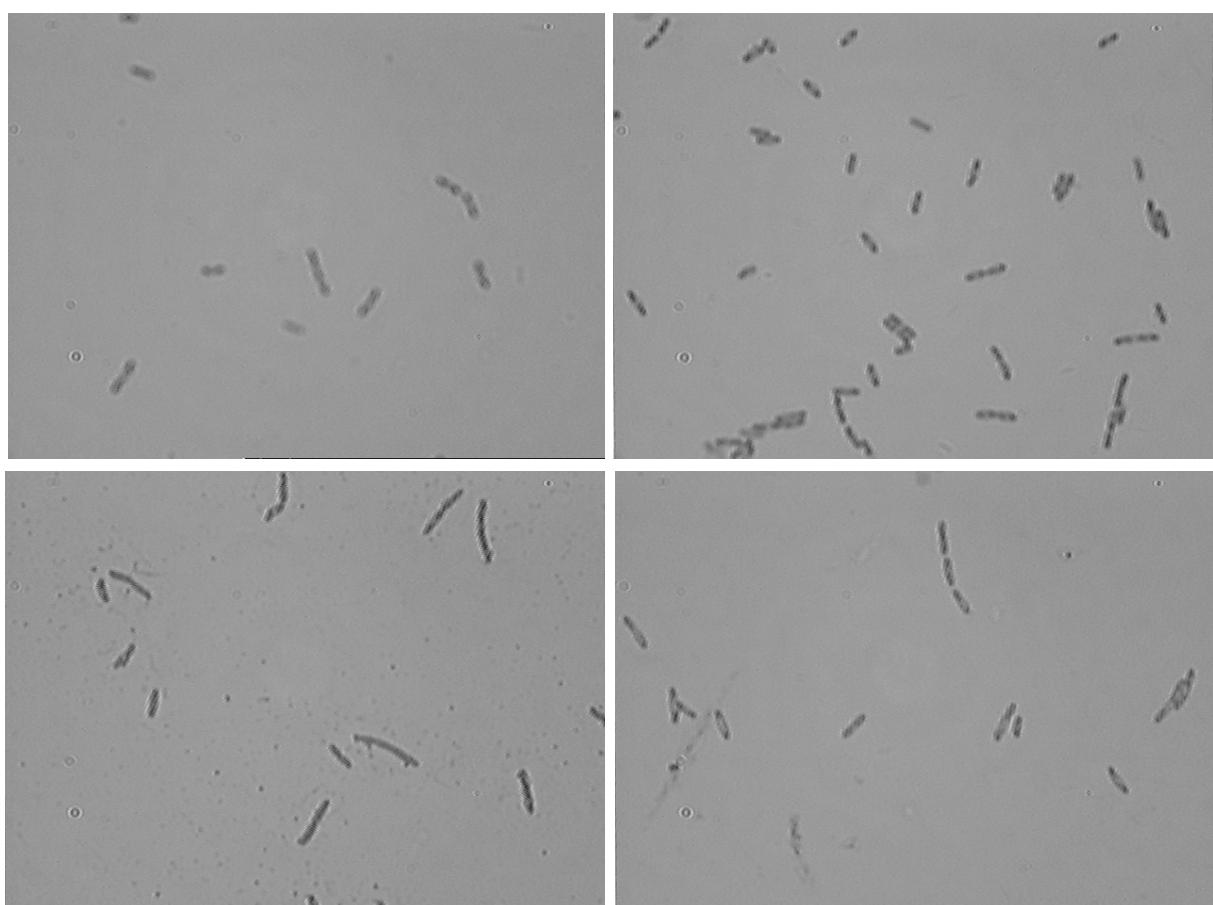
*Training set* utilizado na rede neural de reconhecimento de caracteres.

# Visão Computacional

## Caracterização de micro-organismos

A modelagem de um processo fermentativo requer o estudo da evolução temporal da concentração de micro-organismos na solução, bem como de certas características morfológicas dos mesmos.

Visão Computacional é bastante indicada para a automação desse processo, mas é preciso lidar com os problemas de estimativa de profundidade utilizando-se as técnicas ‘depth from focus’ e estereologia.



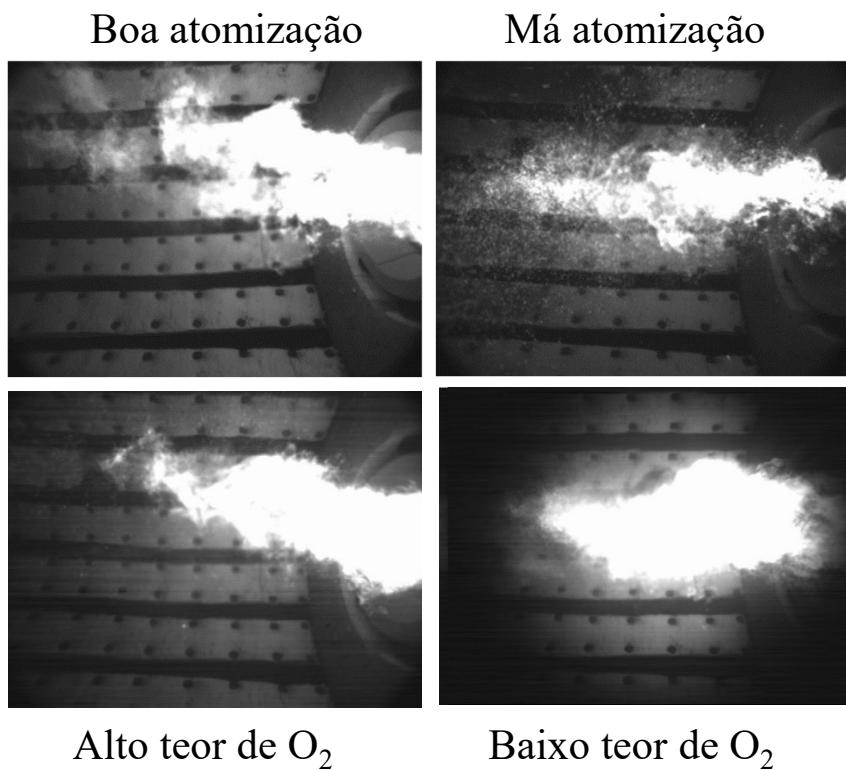
Imagens de uma cultura de micro-organismos produtores de *PHB*, ao longo de várias fases do processo de fermentação

# Visão Computacional

## Defeitos de processos de combustão

A inspeção visual de chamas em fornos de refinarias de petróleo é um trabalho extenuante, mas de importância crucial, pois, se operados em condições inadequadas, podem sofrer sérias avarias, além de causarem danos ao meio ambiente e prejuízos à produção.

Utilizando-se os métodos da visão computacional, pode-se construir um sistema de diagnóstico de defeitos de chamas, capaz de identificar um grande número de condições anormais de operação mediante a comparação dos vetores de características das imagens teste e referência.

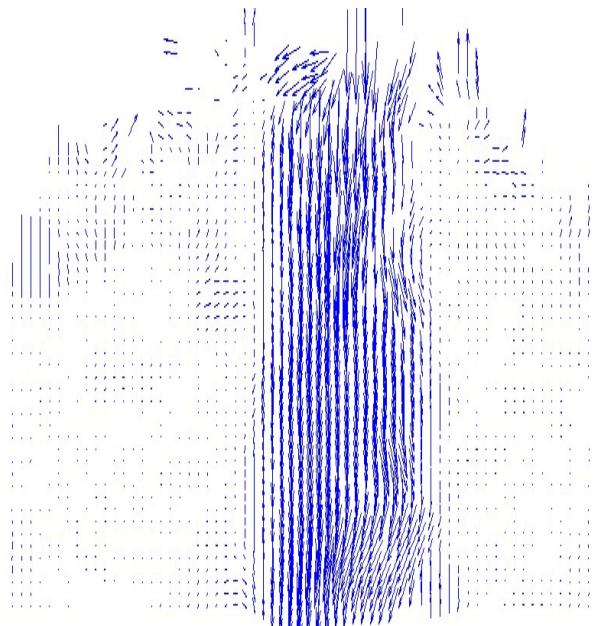


# Visão Computacional

## Particle Image Velocimetry

*PIV* ('Particle Image Velocimetry') é um método para estimação do campo de velocidades de um escoamento, baseado na análise de seqüências de imagens da luz espalhada por partículas microscópicas adicionadas ao fluido em um ponto a montante da região observada, iluminada por uma lâmina de luz pulsada.

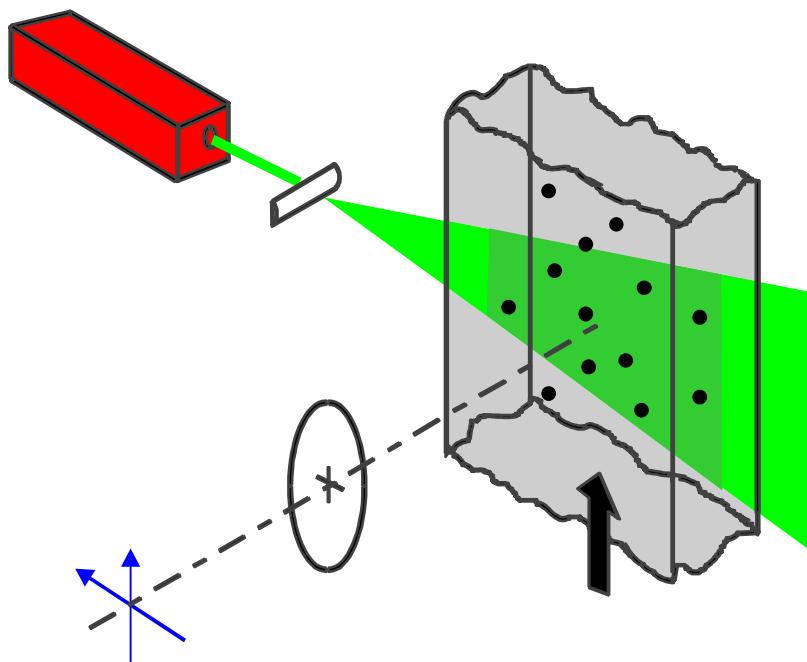
Através da análise da correlação de pares de imagens consecutivas, estima-se, para cada ponto da grade, o vetor velocidade, e, com o auxílio de métodos numéricos apropriados, determinam-se as linhas de corrente, caracterizando-se, assim, o campo de velocidade do escoamento.



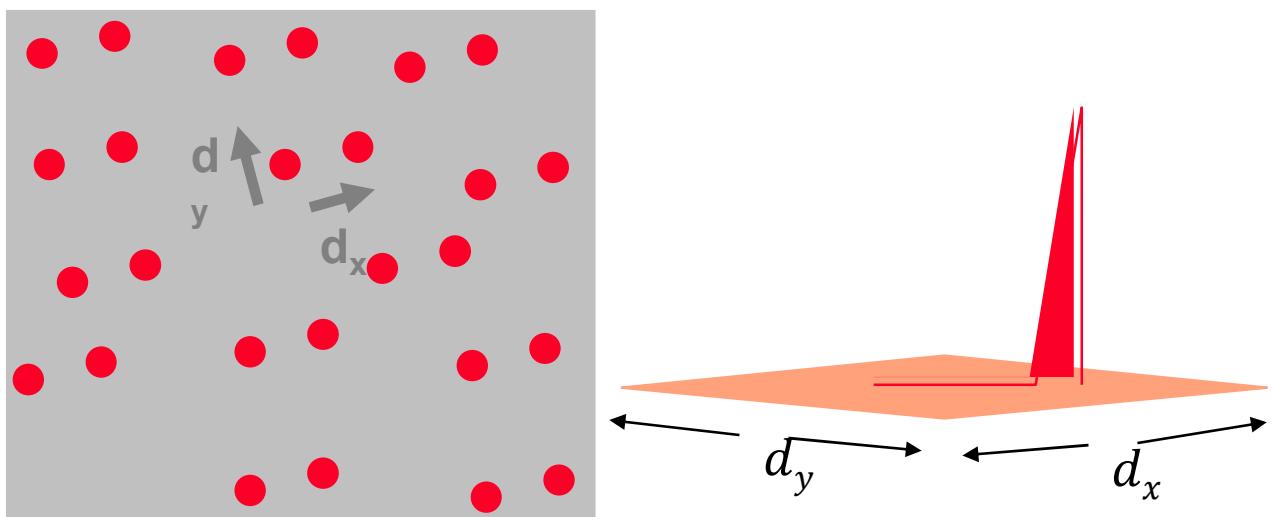
Campo de velocidades estimado pelo método *PIV*

# Visão Computacional

A instrumentação típica de um sistema *PIV* compõe-se de uma fonte de laser pulsado, lente cilíndrica para geração de uma lâmina de luz e equipamento de aquisição de imagens.



Instrumentação típica de um sistema *PIV*



Estimação da correlação de duas imagens consecutivas

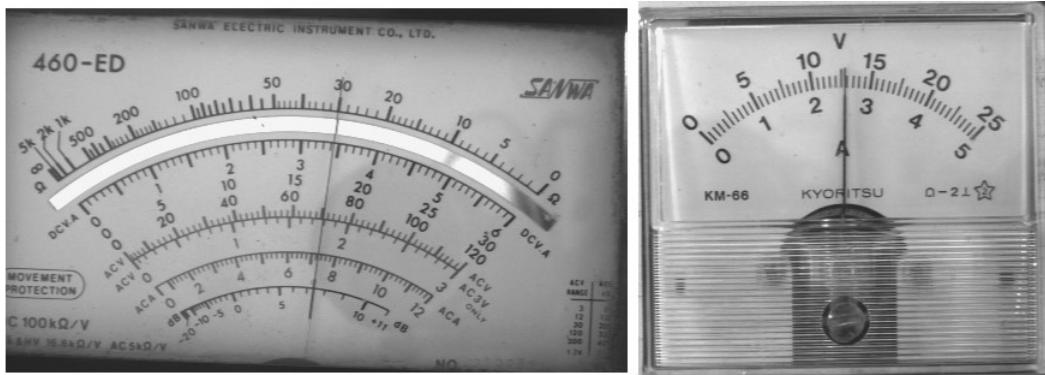
# Visão Computacional

## Leitura de painéis de instrumentos analógicos

Dada a impossibilidade de se modificar os circuitos de controle dos instrumentos analógicos, a alternativa para se automatizar suas leituras consiste na construção de uma interface de visão computacional que, aplicada a seqüências de imagens capturadas dos mostradores, emula as funções de observação, monitoração e anotação desempenhadas por um técnico.

Neste problema há vários desafios a serem vencidos, a saber:

- construção de um sistema especialista para análise das características geométricas do mostrador.
- implementação de algoritmos eficientes e precisos para localização de retas.



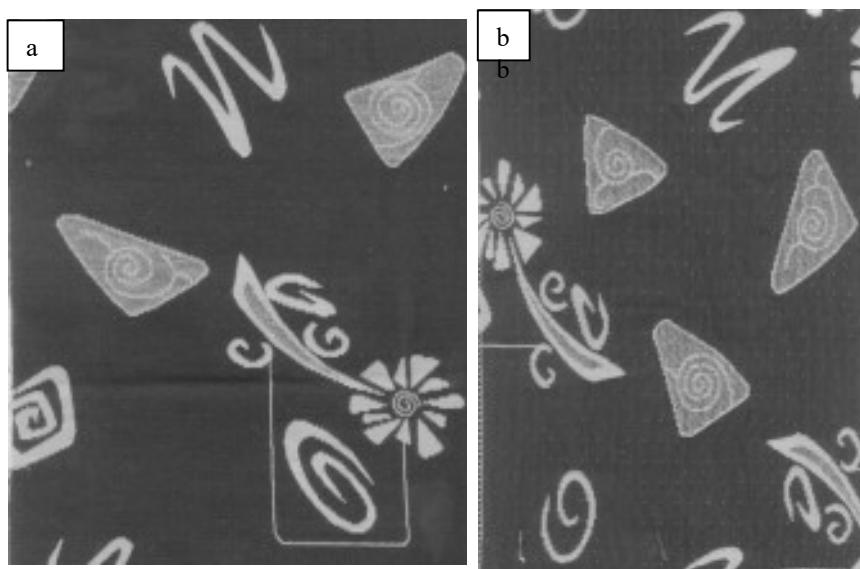
Imagens de mostradores de instrumentos analógicos

# Visão Computacional

## Identificação de plágio em ‘design’ de tecidos

A identificação de plágios em estampas de tecidos, fato que tem causado sérios prejuízos à indústria têxtil, é um tema bastante indicado para a aplicação de técnicas de Visão Computacional.

Apesar de, em geral, a fraude ser baseada na utilização de transformações geométricas simples, facilmente notadas por um observador humano, a resolução desse problema demanda a utilização de técnicas de caracterização e medição de formas de objetos de interesse.



(a) Padrão original; (b) Plágio

# Visão Computacional

## Imagen médica de referência

Propósito: construir uma imagem de referência de um tórax suíno.

Imagens originais

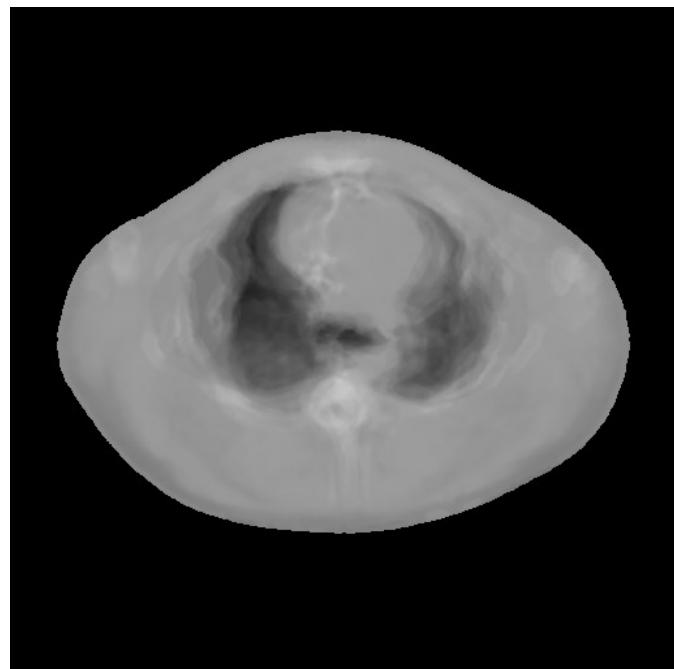
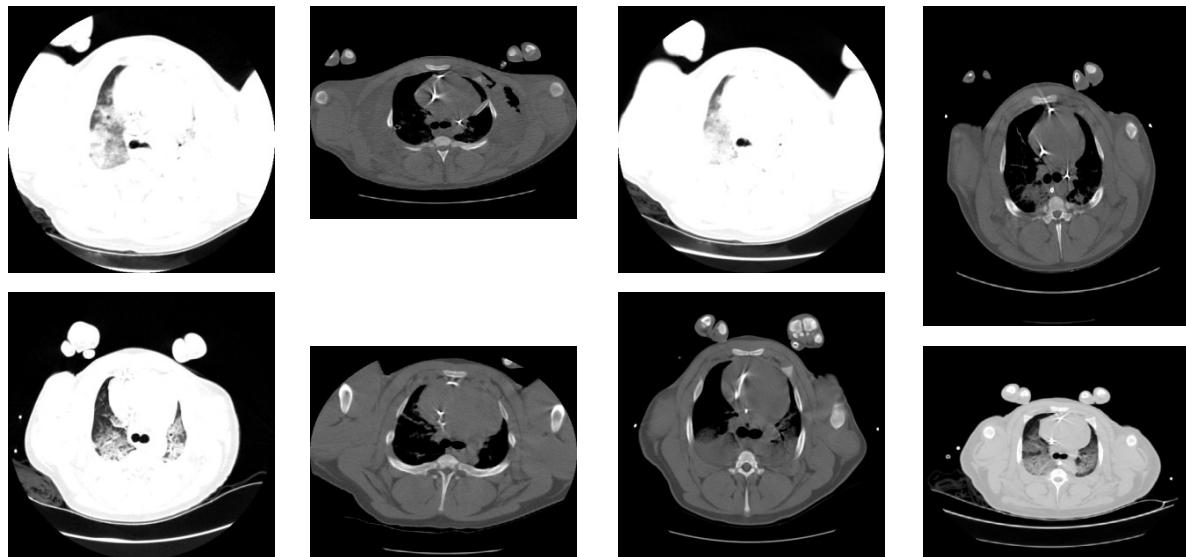


Imagen média referência