

PMR2560 – Visão Computacional

Formação de imagens

Prof. Eduardo L. L. Cabral



Objetivos

- Visão computacional:
 - Sensores de visão;
 - Imagens digitais;
 - Noções de transformação de coordenadas;
 - Formação da imagem;
 - Lentes.

Sensores de imagem

- Efeito fotoelétrico \Rightarrow um raio de luz (que na mecânica quântica pode ser visto como uma partícula chamada fóton) incide em um material e se choca com um elétron de um átomo, provocando a saída desse elétron do átomo. Esse elétron que escapa fica “livre” para formar uma corrente elétrica.
- O fluxo incidente de fótons em um material semi-condutor gera pares elétron-lacuna.
- Um material semi-condutor aumenta o efeito fotoelétrico \Rightarrow para cada elétron livre de carga negativa se forma também uma lacuna de carga positiva.

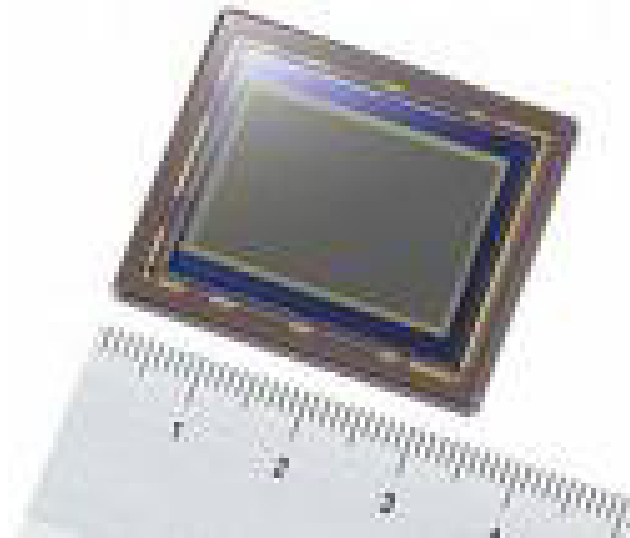
Sensores de imagem

- Eficiência do sensor de imagem (eficiência quântica), $q(\lambda)$:
 - Relação entre o fluxo de fótons e o fluxo de elétrons;
 - Depende da energia do fóton (λ) e do material do sensor;
 - Tubos de vácuo $\Rightarrow q(\lambda)$ baixo;
 - CCDs $\Rightarrow q(\lambda)$ elevado;
 - CMOS $\Rightarrow q(\lambda)$ médio;
 - Fotografia $\Rightarrow q(\lambda)$ baixo.

Sensores de imagem



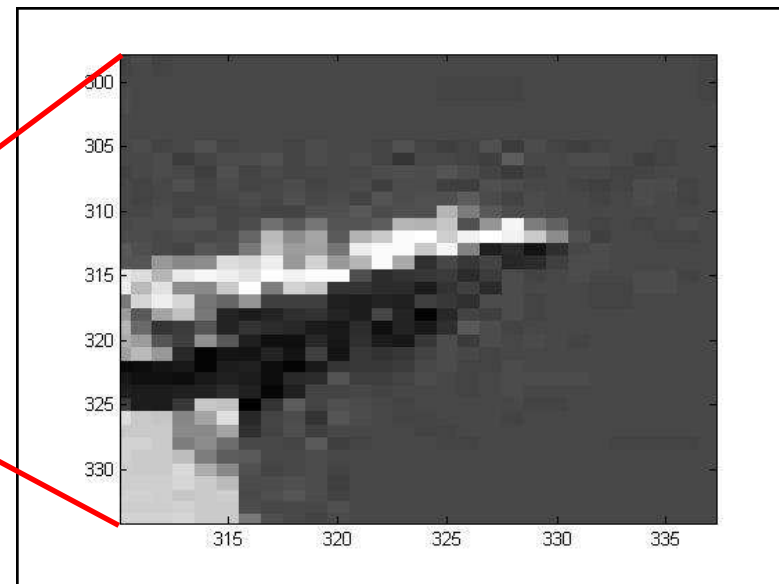
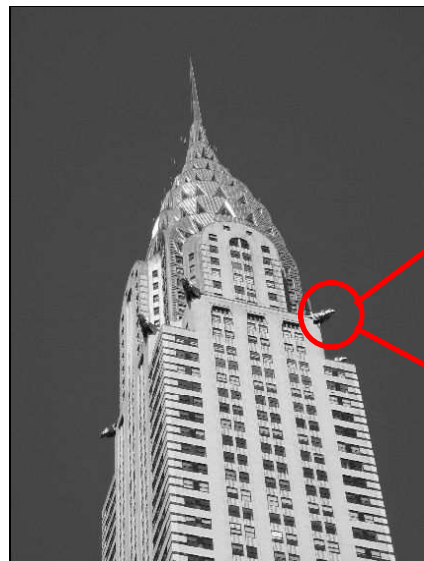
D40 CCD Sensor



*12.4-megapixel
CMOS Sensor*

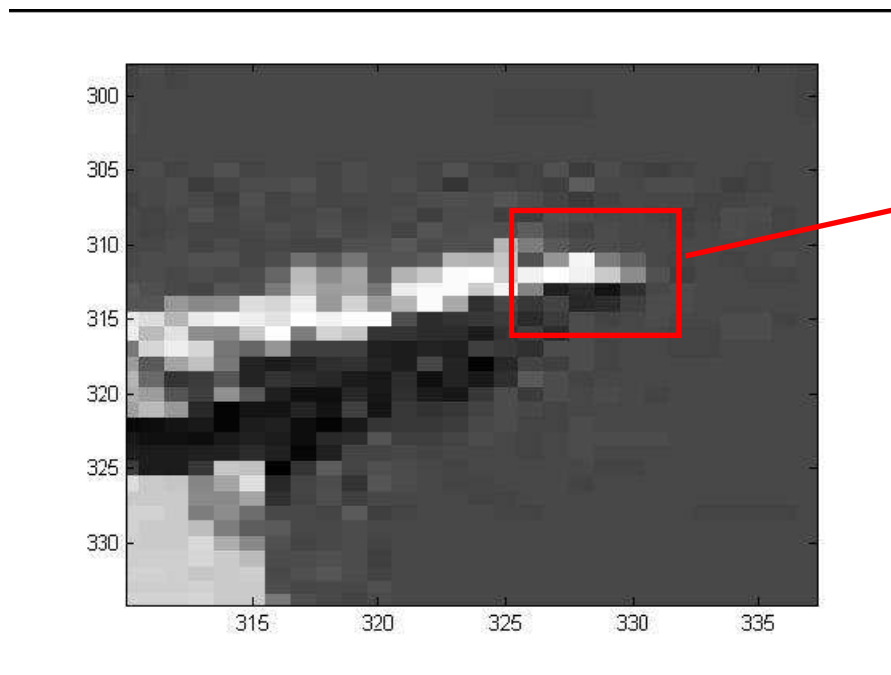
Imagens digitais

- Imagens digitais são funções discretas que correspondem à intensidade luminosa média recebida pela câmera durante um intervalo finito de tempo.
 - São discretas espacialmente;
 - Valores de intensidade luminosa quantizados.



Imagens digitais

- A imagem é somente um conjunto de números na memória do computador.
- É possível manipular esses números para transformá-los em informações úteis.



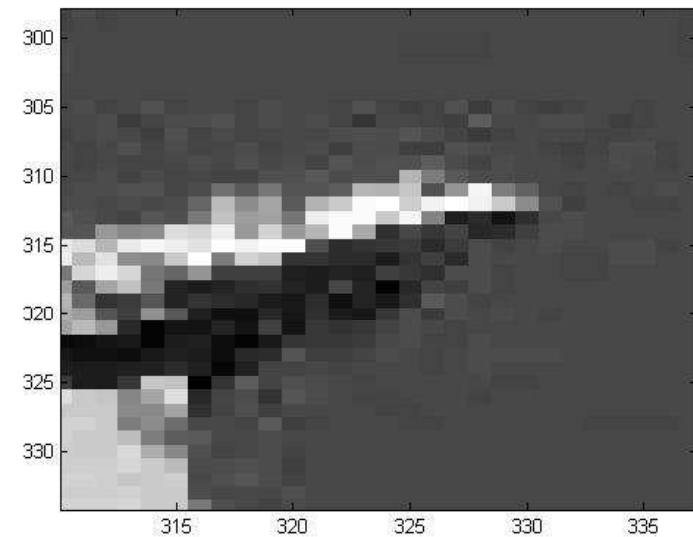
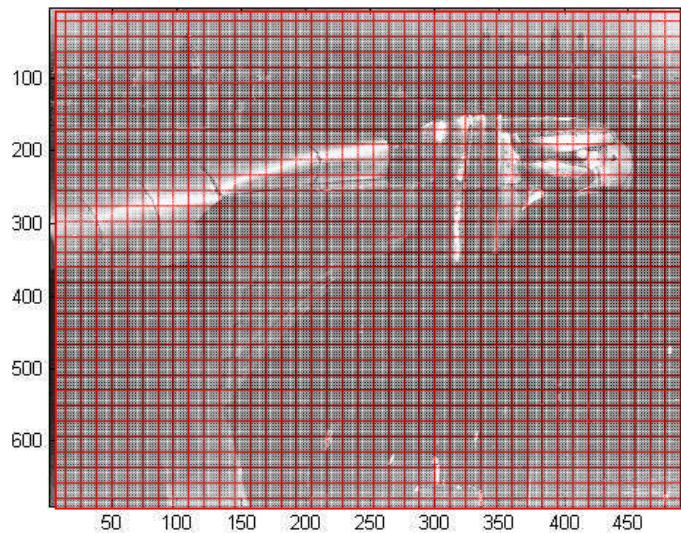
<u>61</u>	<u>29</u>	<u>29</u>	<u>57</u>	<u>199</u>	<u>192</u>
<u>222</u>	<u>200</u>	<u>197</u>	<u>135</u>	<u>167</u>	<u>222</u>
<u>203</u>	<u>203</u>	<u>203</u>	<u>137</u>	<u>137</u>	<u>165</u>
<u>208</u>	<u>208</u>	<u>201</u>	<u>124</u>	<u>142</u>	<u>111</u>
<u>208</u>	<u>203</u>	<u>200</u>	<u>190</u>	<u>127</u>	<u>92</u>
<u>204</u>	<u>201</u>	<u>200</u>	<u>218</u>	<u>173</u>	<u>139</u>

Imagens digitais

- Diversas propriedades definem o formato da imagem digital:
 - Número de pixels, ou resolução espacial (exemplo 640x480 pixels);
 - Número de bits por pixel (exemplos: 8-bit sem sinal, 16-bits com sinal etc.);
 - Taxa de aquisição (exemplo 30 Hz);
 - Número de planos \Rightarrow 1 para imagens em tons de cinza, 3 para imagens coloridas;
 - Forma de representar cor (RGB, HSI etc.);
 - Formato do pixel (plano versus empacotado).

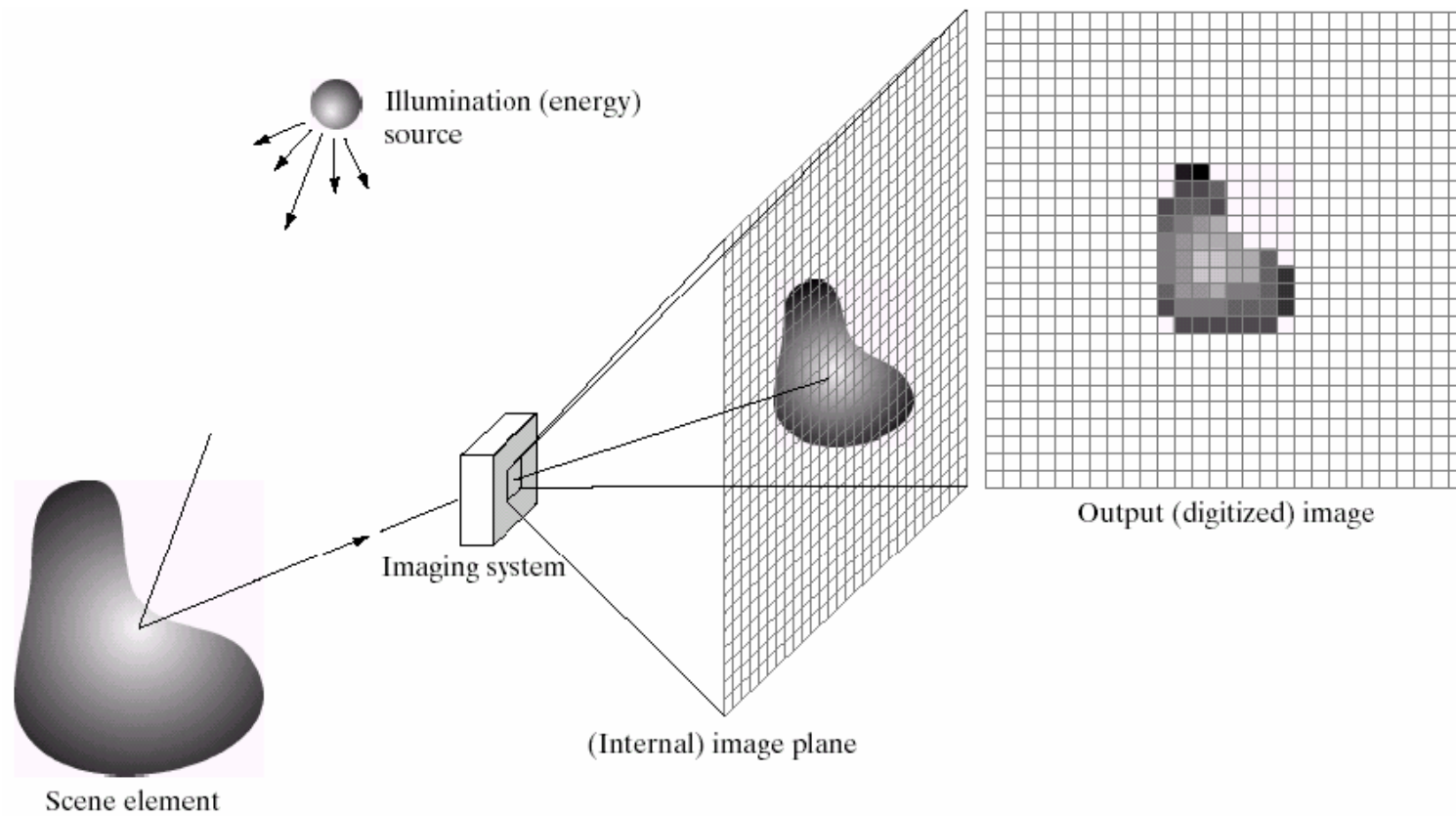
Imagens digitais

- Como chegamos da imagem analógica para a digital?



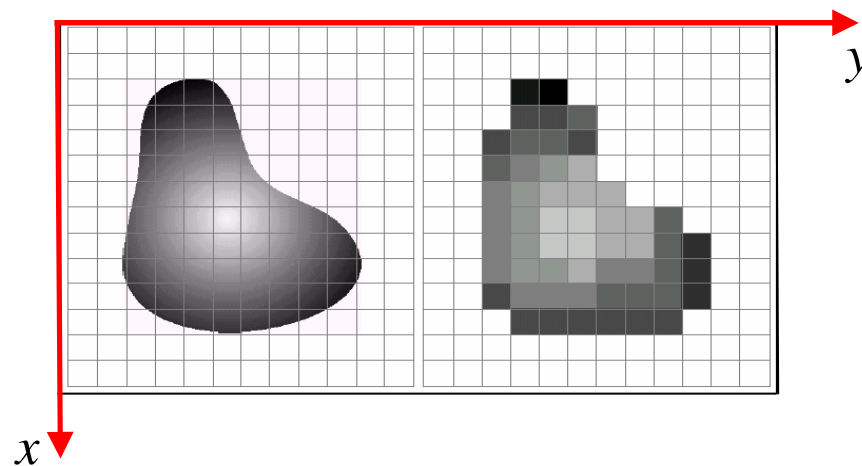
Amostragem e quantização.

Imagens digitais



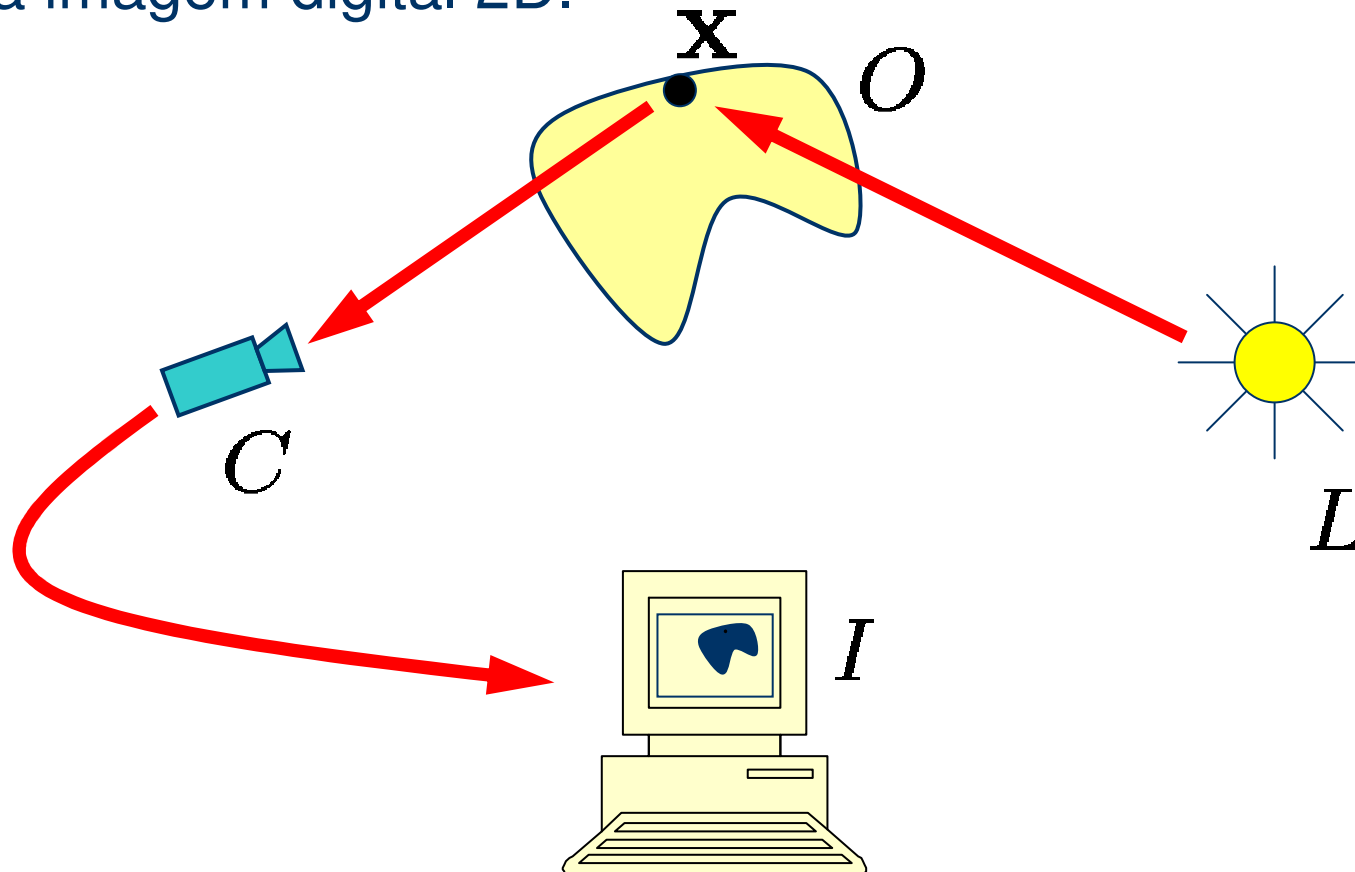
Imagens digitais

- Pode-se pensar que as imagens são funções discretas da intensidade luminosa em cada posição da imagem $I(x,y)$.
- Valores de $I(x,y)$ são definidos para cada ponto com a resolução definida pelo sensor de visão (CCD ou CMOS).
- A quantização da intensidade luminosa de cada ponto da imagem é definida pelo número de bits por pixel e pelo espaço de cor utilizado.



Formação da imagem

- Compreender como o mundo tridimensional se projeta na imagem digital 2D.



Formação da imagem

- O que determina a localização de um ponto da imagem? (geometria)
 - Transformação de sistemas de coordenadas;
 - Projeção perspectiva;
 - Lentes.
- Como se determina o brilho/cor que uma superfície/objeto apresenta numa imagem? (Radiometria)
 - Emissão de luz;
 - Interação da luz com as superfícies.
- Conversão da imagem analógica para digital:
 - Amostragem espacial;
 - Intervalo dinâmico;
 - Integração temporal.

Transformação de coordenadas

- Transformação de coordenadas do sistema 0 para o sistema 1:

$$\mathbf{p}_0 = \mathbf{x}_0 + \mathbf{R}_0^1 \mathbf{p}_1$$

onde:

- $\mathbf{x}_0 \Rightarrow$ vetor de translação;
- $\mathbf{R}_0^1 \Rightarrow$ matriz de rotação.

- Somente translação:

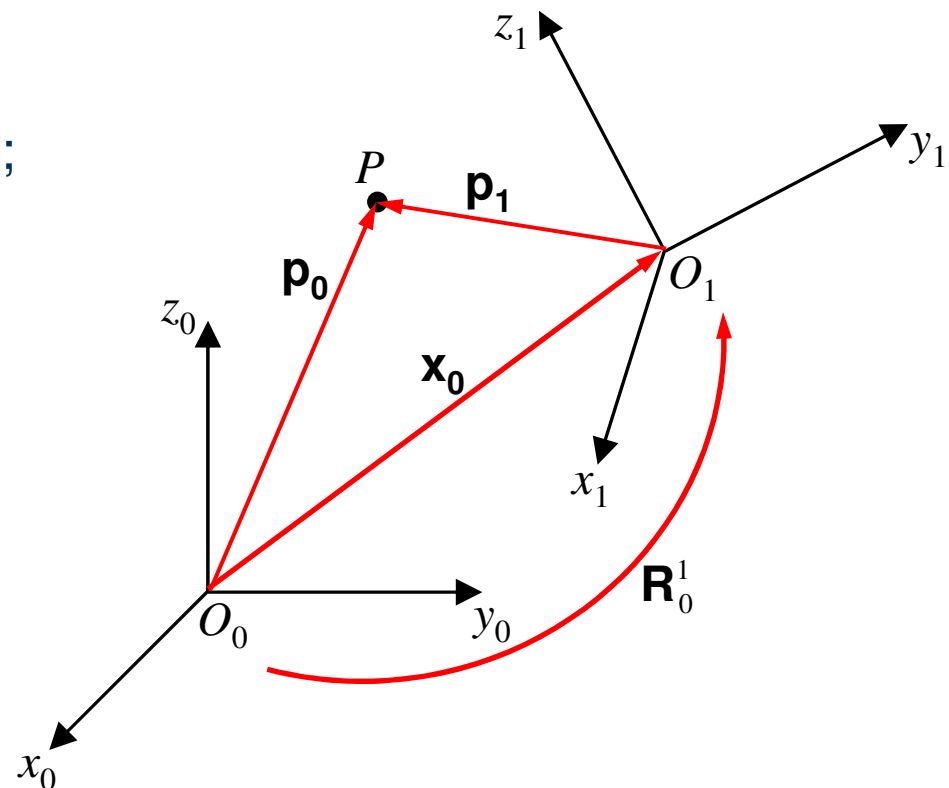
$$\mathbf{p}_0 = \mathbf{x}_0 + \mathbf{I} \mathbf{p}_1;$$

- Vetor de translação \mathbf{x}_0 ;
- $\mathbf{R}_0^1 = \mathbf{I}$ (identidade 3x3).

- Somente rotação:

$$\mathbf{p}_0 = \mathbf{R}_0^1 \mathbf{p}_1;$$

- $\mathbf{x}_0 = \mathbf{0}$.



Transformação de coordenadas

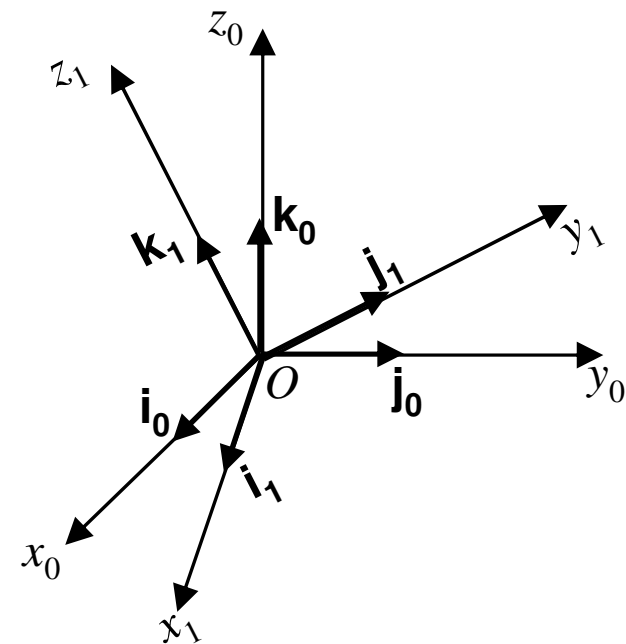
- Matriz de rotação:

$$\mathbf{R}_0^1 = \begin{bmatrix} i_1 \cdot i & j_1 \cdot i & k_1 \cdot i \\ i_1 \cdot j & j_1 \cdot j & k_1 \cdot j \\ i_1 \cdot k & j_1 \cdot k & k_1 \cdot k \end{bmatrix}$$

- Propriedades:

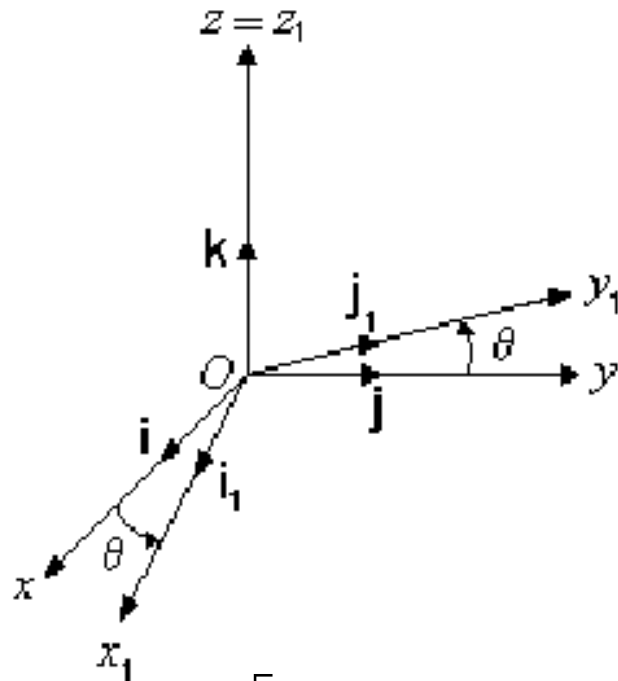
- Linhas e colunas tem módulo unitário;
- Linhas e colunas são perpendiculares entre si \Rightarrow produto escalar entre elas é zero;
- $\text{Det}(\mathbf{R}) = 1$ (se a regra da mão direita é seguida);
- Inversa da matriz de rotação:

$$\mathbf{R}_1^0 = (\mathbf{R}_0^1)^{-1} = \mathbf{R}_0^1{}^t$$



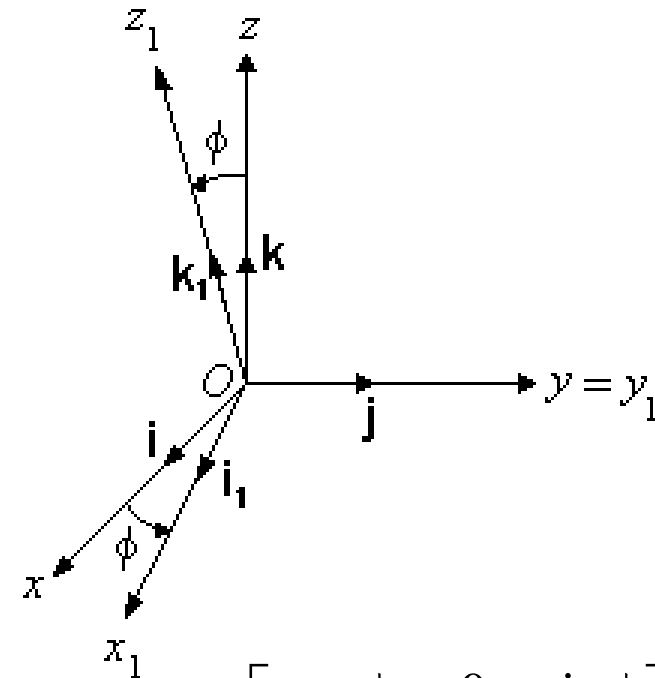
Transformação de coordenadas

- Rotação em relação a z:



$$\mathbf{R}_{z,\theta} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- Rotação em relação a y:



$$\mathbf{R}_{y,\phi} = \begin{bmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{bmatrix}$$

Transformação homogênea

- Define-se os vetores homogêneos \mathbf{P}_0 e \mathbf{P}_1 de dimensão 4x1:

$$\mathbf{P}_0 = (\lambda p_{x0}, \lambda p_{y0}, \lambda p_{z0}, \lambda)^t = \begin{bmatrix} \lambda \mathbf{p}_0 \\ \lambda \end{bmatrix} \quad \mathbf{P}_1 = (\lambda p_{x1}, \lambda p_{y1}, \lambda p_{z1}, \lambda)^t = \begin{bmatrix} \lambda \mathbf{p}_1 \\ \lambda \end{bmatrix}$$

onde p_x , p_y e p_z são as coordenadas de um ponto P no espaço e λ é um fator de escala.

- Define-se a matriz homogênea, \mathbf{H}_0^1 , de dimensão 4x4:

$$\mathbf{H}_0^1 = \begin{bmatrix} \mathbf{R}_0^1 & \mathbf{x}_0 \\ \mathbf{0} & 1 \end{bmatrix}$$

- matriz \mathbf{H}_0^1 representa a posição e a orientação do sistema $O_1-x_1y_1z_1$ em relação ao sistema $O_0-x_0y_0z_0$.

Transformação homogênea

- Uma transformação de coordenadas ($\mathbf{p}_0 = \mathbf{x}_0 + \mathbf{R}_0^1 \mathbf{p}_1$) em termos dos vetores e matriz homogêneos fica:

$$\mathbf{P}_0 = \mathbf{A}_0^1 \mathbf{P}_1 \Rightarrow \text{Transformação homogênea}$$

ou,

$$\begin{bmatrix} \mathbf{p}_0 \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_0^1 & \mathbf{x}_0 \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ 1 \end{bmatrix}$$

- **Vantagem da transformação homogênea \Rightarrow forma compacta de representação (conveniente para representar transformações consecutivas).**

Transformação homogênea

- A transformação homogênea inversa pode ser obtida da transformação inversa:

$$\mathbf{p}_1 = -(\mathbf{R}_0^1)^t \mathbf{x}_0 + (\mathbf{R}_0^1)^t \mathbf{p}_0$$

Escrevendo de forma matricial e usando os vetores e matriz homogêneos \mathbf{P}_0 , \mathbf{P}_1 e \mathbf{H}_0^1 fica:

$$\mathbf{P}_1 = (\mathbf{H}_0^1)^{-1} \mathbf{P}_0 = \mathbf{H}_1^0 \mathbf{P}_0$$

onde,

$$\mathbf{H}_1^0 = (\mathbf{H}_0^1)^{-1} = \begin{bmatrix} (\mathbf{R}_0^1)^t & -(\mathbf{R}_0^1)^t \mathbf{x}_0 \\ \mathbf{0} & 1 \end{bmatrix}$$



essa é a forma de calcular a inversa da matriz homogênea

Transformações consecutivas

- Transformação homogênea \Rightarrow forma compacta de representar diversas transformações de coordenadas consecutivas por um único termo.
- Transformação homogênea do sistema $i-1$ para o sistema i é descrita por:

$$\mathbf{P}_{i-1} = \mathbf{H}_{i-1}^i \mathbf{P}_i;$$

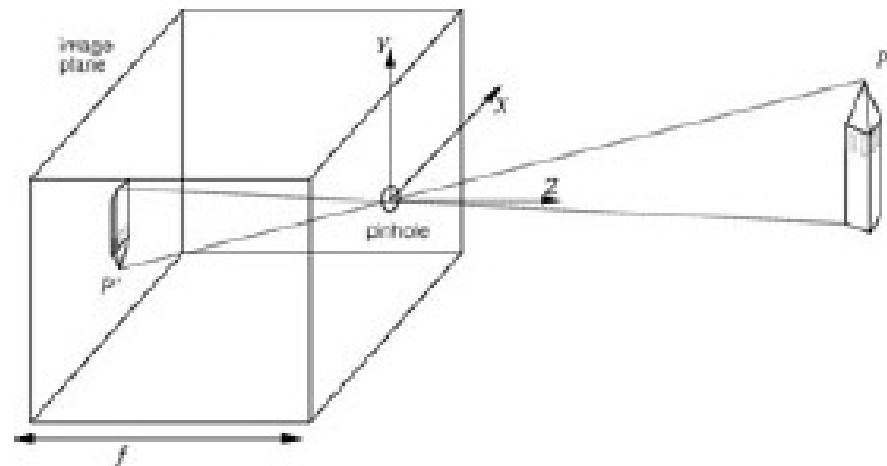
- Transformação do sistema 0 para o sistema n :

$$\mathbf{P}_0 = \mathbf{H}_0^1 \mathbf{H}_1^2 \dots \mathbf{H}_{n-1}^n \mathbf{P}_n$$

- \mathbf{P}_0 posição do ponto P no sistema 0;
- \mathbf{P}_n posição do ponto P no sistema n .

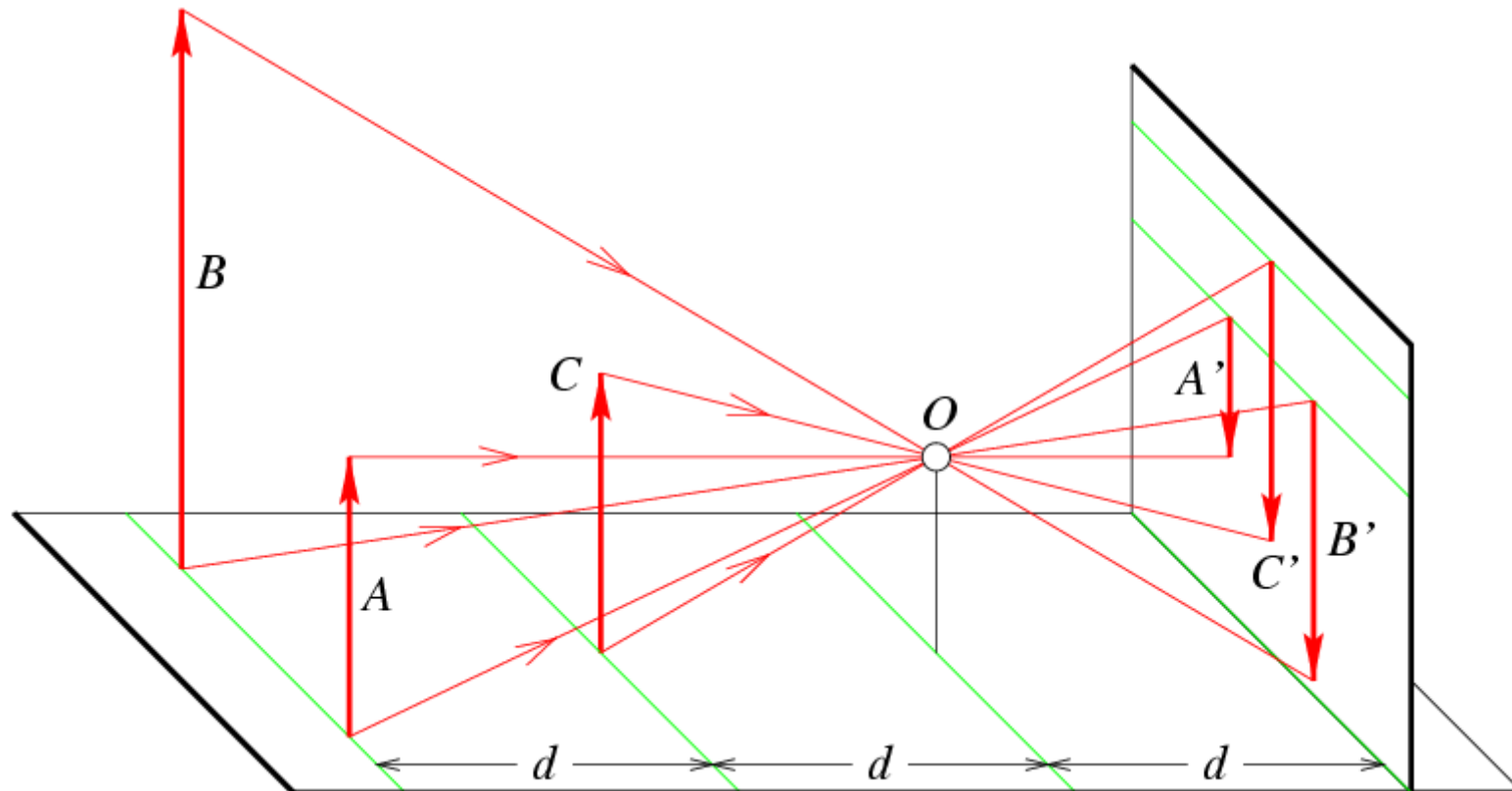
Formação da imagem

- Situação ideal:
 - Modelo de câmara pontual (*pinhole*);
 - Só a luz vinda do “furo” atinge o plano de imagem;
 - Cada ponto na imagem corresponde a um único ponto 3D.
- Representação alternativa.



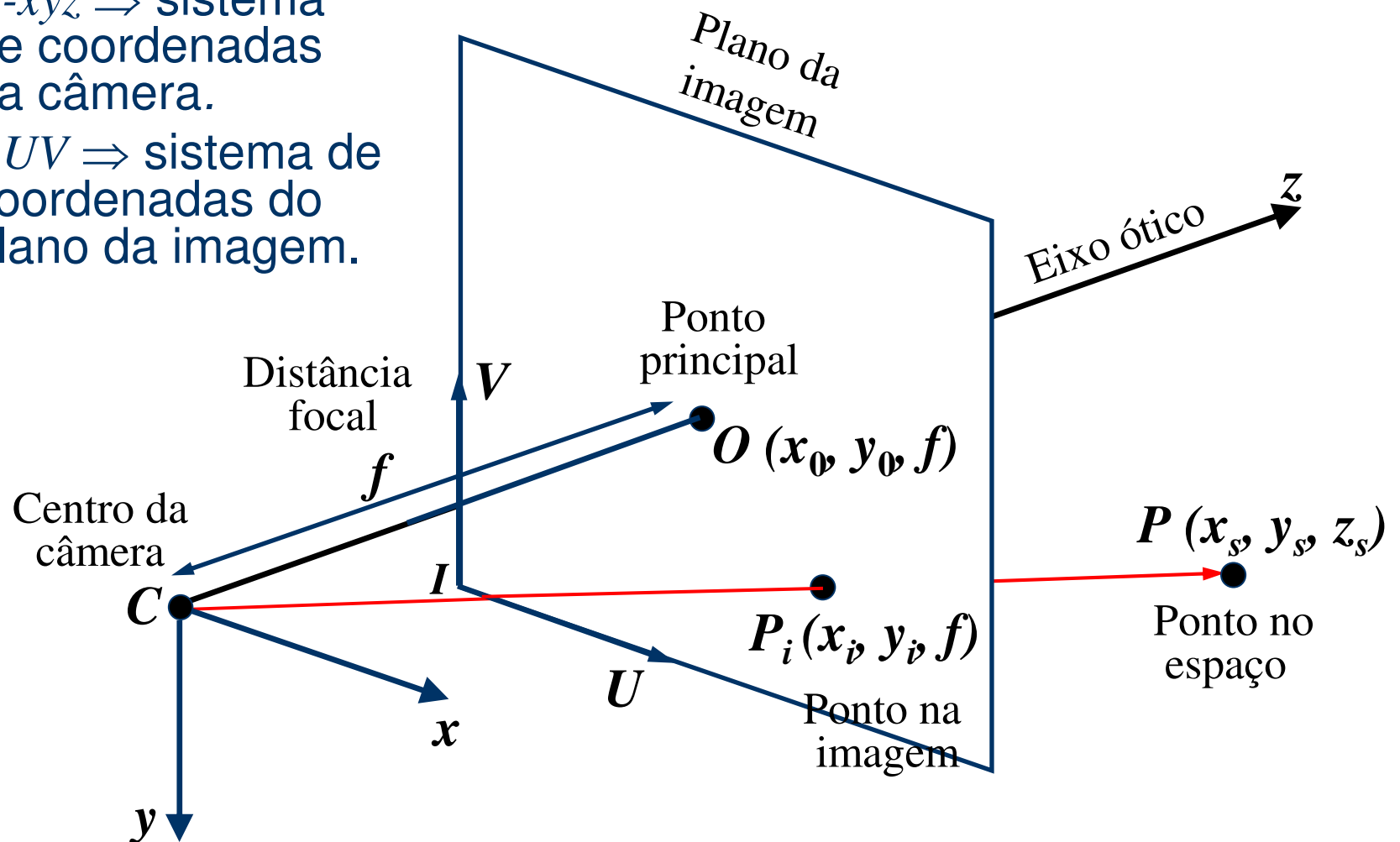
Formação da imagem

- Projeção perspectiva.



Formação da imagem

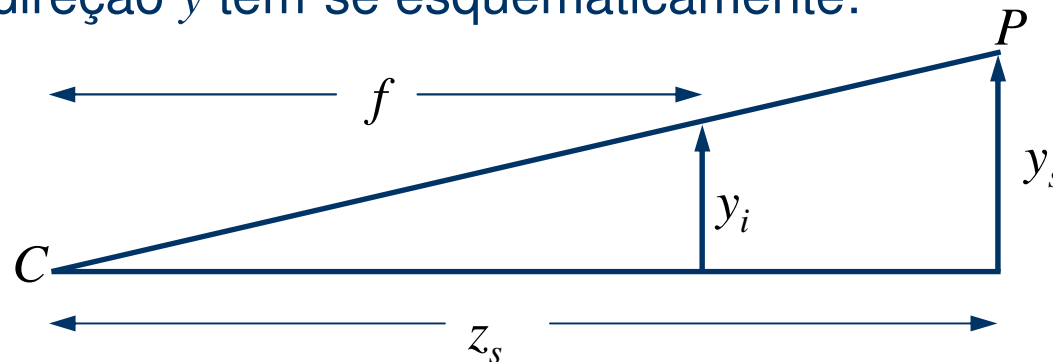
- C - $xyz \Rightarrow$ sistema de coordenadas da câmera.
- I - $UV \Rightarrow$ sistema de coordenadas do plano da imagem.



Formação da imagem

➤ Projeção perspectiva:

- Ponto P no espaço descrito no sistema de coordenadas da câmera $\Rightarrow (x_s, y_s, z_s)$;
- Ponto P no plano da imagem descrito no sistema de coordenadas da câmera $\Rightarrow (x_i, y_i, z_i)$;
- Na direção y tem-se esquematicamente:



- Por similaridade de triângulos:

$$x_i = f \frac{x_s}{z_s}; \quad y_i = f \frac{y_s}{z_s}; \quad z_i = f.$$

Formação da imagem

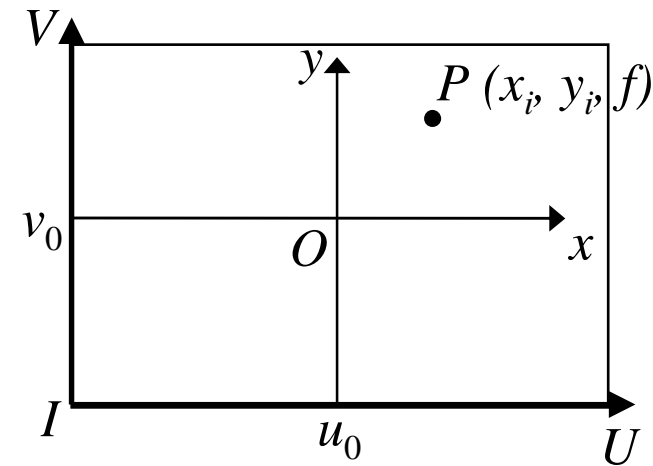
- Transformação do sistema de coordenadas da imagem (3D - métrica) para o sistema da imagem (2D - pixel):

$$u - u_0 = k_x x_i \Rightarrow u = k_x f \frac{x_s}{z_s} + u_0$$

$$v - v_0 = k_y y_i \Rightarrow v = k_y f \frac{y_s}{z_s} + v_0$$

onde:

- x_i e y_i coordenadas métricas;
- u e v em pixel;
- k_x e k_y são fatores de escala;
- u_0 e v_0 são coordenadas do centro da imagem em pixel.



Formação da imagem

- Utilizando matrizes homogêneas:

$$\begin{cases} uw = \alpha_x x_s + u_0 z_s \Rightarrow u = \alpha_x \frac{x_s}{z_s} + u_0 \\ vw = \alpha_y y_s + v_0 z_s \Rightarrow v = \alpha_y \frac{y_s}{z_s} + v_0 \\ w = z_s \end{cases}$$

– α_x e $\alpha_y \Rightarrow$ distâncias focais em pixels;

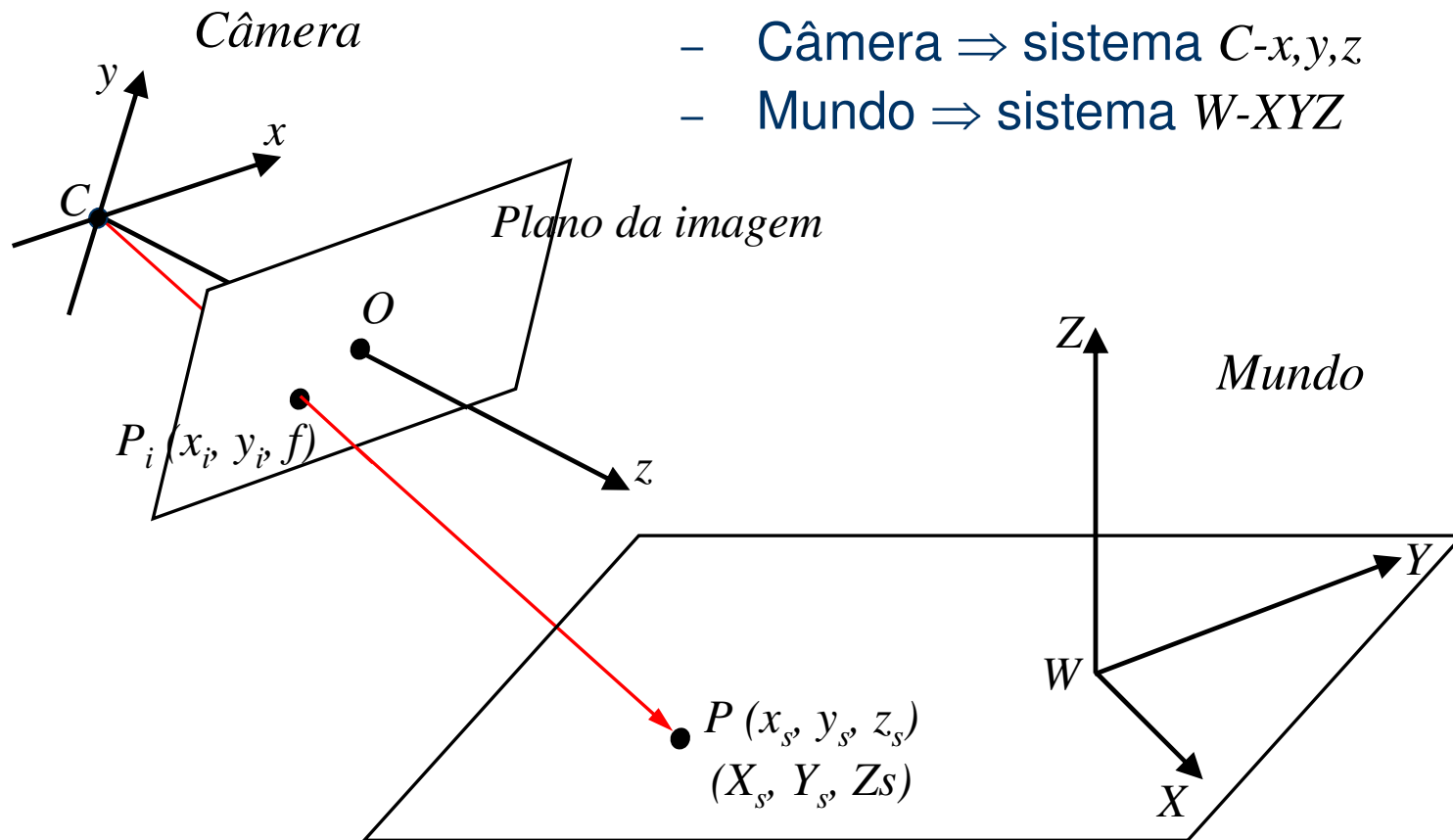
$$\alpha_x = k_x f$$

$$\alpha_y = k_y f$$

$$\begin{bmatrix} uw \\ vw \\ w \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} uw \\ vw \\ w \end{bmatrix} = \mathbf{K} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix}$$

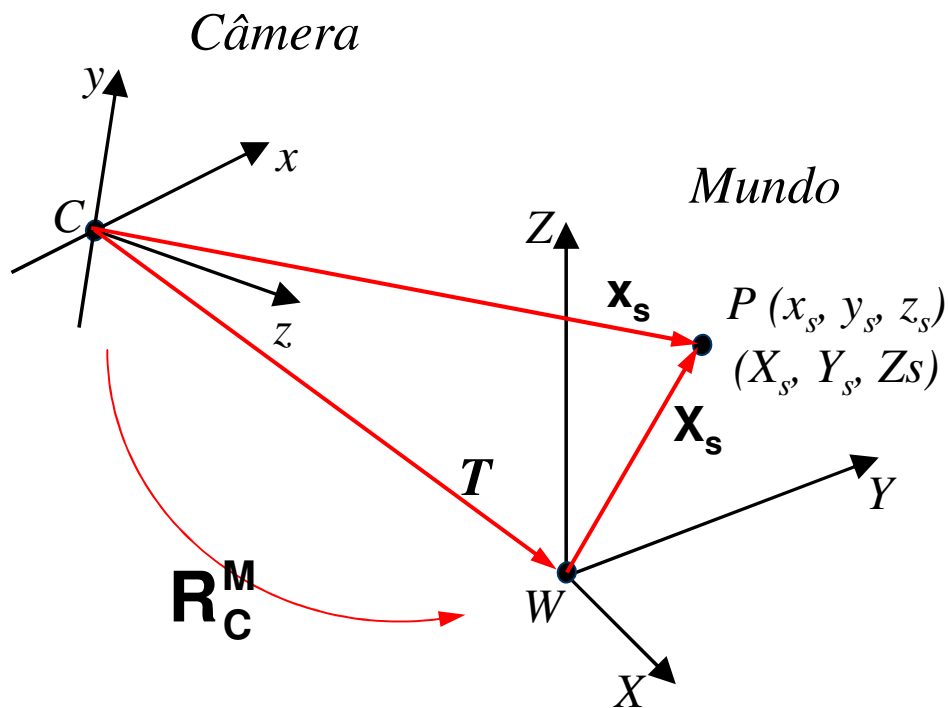
Formação da imagem

- Transformação do sistema da câmera para o sistema do mundo.



Formação da imagem

- Transformação de coordenadas do sistema da câmera para o mundo.



$$\begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_C^M & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{bmatrix}$$

$$\mathbf{x}_s = \begin{bmatrix} \mathbf{R}_C^M & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{X}_s$$

Formação da imagem

- Transformação de coordenadas do sistema da imagem (2D em pixels) para o mundo (3D métrico):

- Da imagem para a câmera:

$$\begin{bmatrix} uw \\ vw \\ w \end{bmatrix} = \mathbf{K} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix}$$

- Da câmera para o mundo:

$$\begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_C^M & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} uw \\ vw \\ w \end{bmatrix} = \mathbf{kH}_C^M \begin{bmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{bmatrix} = \mathbf{P} \begin{bmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{bmatrix}$$

Formação da imagem

- Propriedades da matriz **P**:

$$\mathbf{P} = \mathbf{kH}_C^M = \begin{bmatrix} \alpha_x & 0 & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_C^M & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix}$$

- Matriz de rotação $\mathbf{R}_C^M \Rightarrow 3$ parâmetros independentes;
 - Vetor de translação $\mathbf{T} \Rightarrow 3$ parâmetros;
 - Matriz $\mathbf{K} \Rightarrow 4$ parâmetros.
- **Lente** \Rightarrow como entra a lente na formação da imagem?
 - Como obter esses 10 parâmetros?

Lentes

- Problema com as câmeras *pinholes* \Rightarrow tamanho do furo é um compromisso entre nitidez e luminosidade:
 - O diâmetro do furo tem que ser “muito pequeno” para obter uma imagem nítida;
 - Um furo pequeno não produz luminosidade suficiente;
 - Se o furo é da mesma ordem de grandeza do comprimento de onda da luz \Rightarrow difração embaralha a imagem.
 - Uma imagem nítida com a câmera *pinhole* é obtida quando o diâmetro do furo é dado por:

$$d = 2 \sqrt{f' \lambda}$$

- Exemplo: se $f = 50mm$, para a cor vermelha $\lambda = 600nm \Rightarrow d = 0,6mm$.

➤ **Solução \Rightarrow uso de lentes.**

Lentes

- Exemplos de imagens de câmeras *pinhole*.

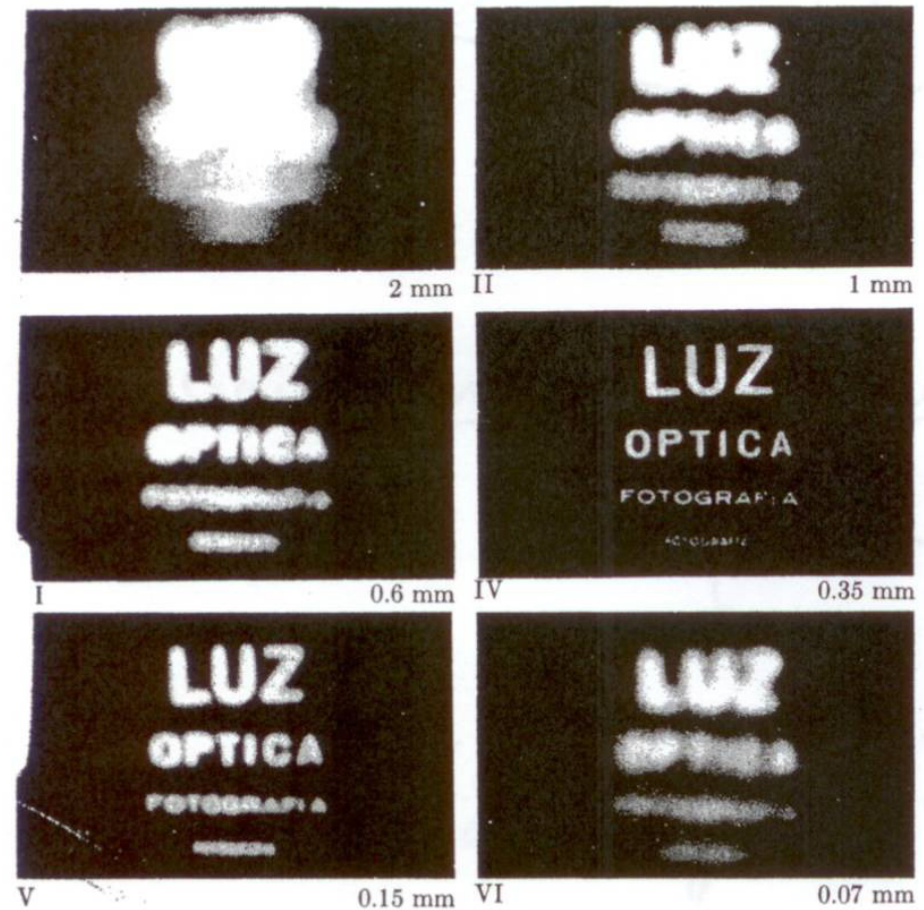


Fig. 5.96 The pinhole camera. Note the variation in image clarity as the hole diameter decreases. [Photos courtesy Dr. N. Joel, UNESCO.]

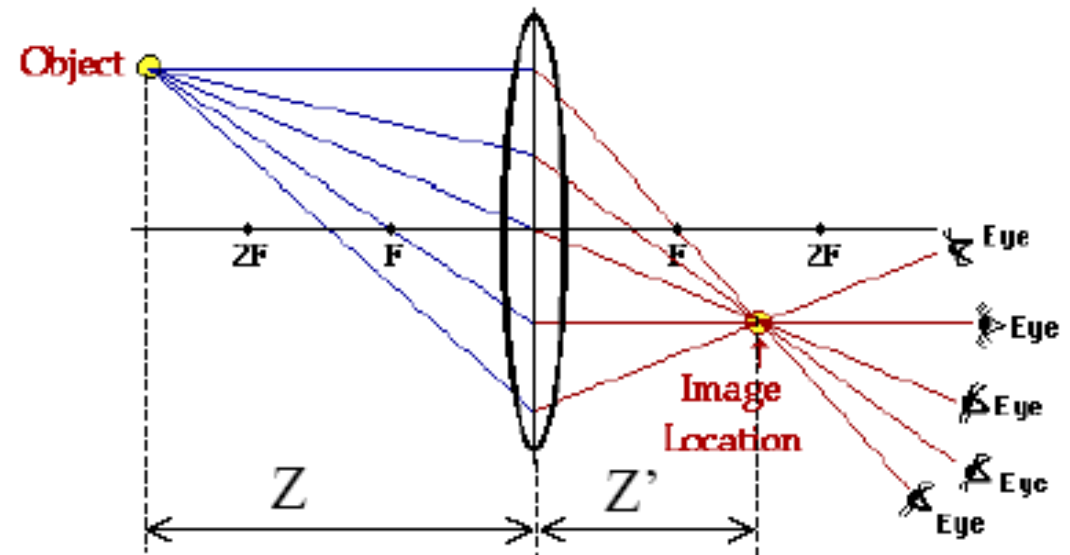
Lentes

- Benefícios das lentes \Rightarrow aumenta a intensidade de luz na imagem pela focalização de raios de luz da cena.
- Complicações das lentes:
 - Profundidade limitada;
 - Distorção radial e tangencial \Rightarrow linhas retas aparecem curvas;
 - Imagem parece mais escura nas bordas;
 - Aberração cromática \Rightarrow distância focal é função do comprimento de onda.
- Profundidade de campo:
 - Intervalo de distâncias bem focadas \Rightarrow depende do sensor;
 - Quanto maior a abertura da lente menor a profundidade de campo

Lentes

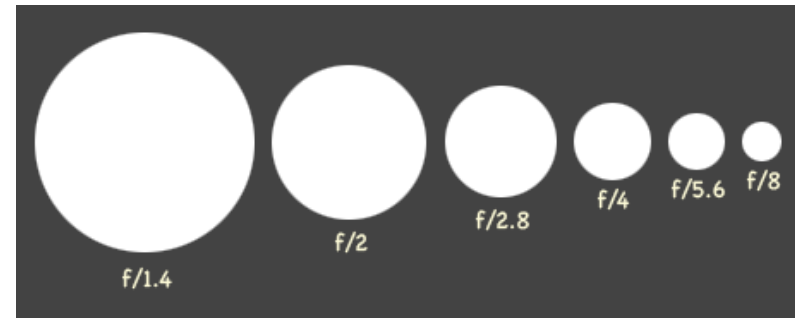
- Lente ideal:
 - O raio que passa pelo centro da lente, não é defletido.
 - Os restantes raios intersectam-se num ponto único juntamente com o raio central.
- Equação da lente:

$$\frac{1}{f} = \frac{1}{Z'} + \frac{1}{Z}$$



Lentes

- **Abertura** \Rightarrow diâmetro D da lente que é exposta à luz.



- **F-Number ($f/\#$):**

$$f/\# = \frac{f}{D}$$

- Por exemplo \Rightarrow se f for 16 vezes o diâmetro da pupila, então $f/\# = f/16$.
- Quanto maior o F-Number menos luz por unidade de área atinge o plano da imagem.

Lentes

- **Profundidade de campo**
⇒ intervalo de distâncias para as quais a imagem é bem focalizada.
- Controle da profundidade de campo ⇒ aumento da abertura da lente diminui a profundidade de campo.



Cenas próximas e distantes
ficam embaralhadas

Lentes



Aumento da abertura diminui a profundidade de visão

Lentes

- Modelo de distorção radial da lente:

- A distorção é função da distância do pixel ao centro da imagem:

$$\begin{bmatrix} u_d \\ v_d \end{bmatrix} = L(r) \begin{bmatrix} u \\ v \end{bmatrix}$$

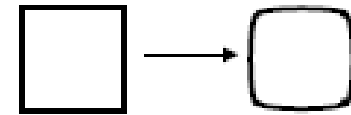
onde,

$$r = \sqrt{(u - u_0)^2 + (v - v_0)^2}$$

- Pode ser aproximada por um polinômio:

$$L(r) = 1 + k_1 r + k_2 r^2 + \dots$$

- Distorção tangencial em geral é pequena e pode ser desconsiderada.



Exemplo do que a distorção radial provoca na imagem

Lentes

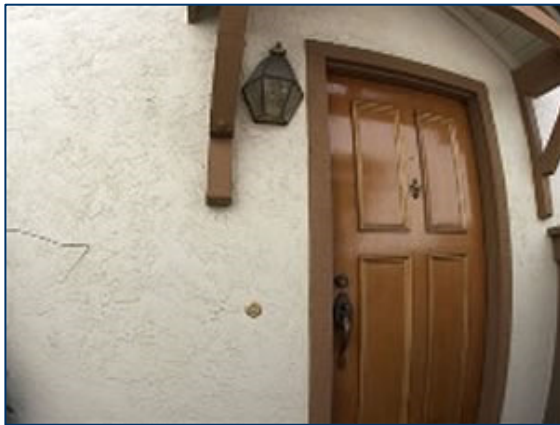
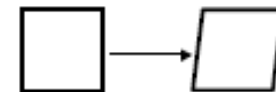
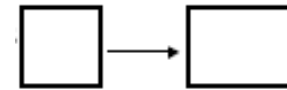


Imagem distorcida

Imagem corrigida

Formação da imagem

- Outras formas de distorção da imagem:
 - Fatores de escala diferentes para as colunas e linhas de pixels \Rightarrow pixels da câmera não são necessariamente quadrados.
 - Efeito considerado pelos fatores de escala k_x e k_y .
 - Fator de inclinação, s (*Skew factor*):
 - Causado pelo tempo finito de amostragem da imagem e pela disposição incorreta das linhas da imagem.



Formação da imagem

- Transformação de coordenadas do sistema da imagem (em pixels) para o mundo considerando a distorção da lente:

$$\begin{bmatrix} u_d \\ v_d \end{bmatrix} = L(r) \begin{bmatrix} u \\ v \end{bmatrix} \Rightarrow \begin{bmatrix} u_d w \\ v_d w \\ w \end{bmatrix} = L(r) \begin{bmatrix} \alpha_x & s & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{H}_C^M \begin{bmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{bmatrix}$$

- Em geral considera-se somente o termo linear da distorção radial:

$$L(r) \cong 1 + k_1 r$$

Formação da imagem

- No modelo de formação da imagem tem-se diversos parâmetros:
 - Parâmetros intrínsecos \Rightarrow internos à câmera:
 - $\alpha_x, \alpha_y, u_0, v_0, s, k_1$ (6 parâmetros).
 - Parâmetros extrínsecos \Rightarrow externos à câmera:
 - \mathbf{R}_c^M e \mathbf{T} (6 parâmetros).
 - Na maioria das aplicações de visão computacional é preciso conhecer esses parâmetros.
- Cálculo dos parâmetros da câmera \Rightarrow **calibração da câmera.**

Sumário

- Sensores de visão:
 - Visão computacional \Rightarrow CCD ou CMOS.
- Imagens digitais:
 - Representam luminosidade média de um ponto no espaço;
 - Resolução espacial;
 - Quantização da intensidade luminosa.
- Formação da imagem \Rightarrow modelo de câmera *pinhole*:
 - Transformação perspectiva e métrica para pixel.
 - Lente *pinhole* \Rightarrow imagem escura e sem nitidez;
 - Uso de lentes resolve problemas da lente *pinhole* \Rightarrow mas causa distorção;
 - Transformação do sistema da câmera para o do mundo;
 - Lentes reais.

Exercícios

1. Dados parâmetros da câmera, da lente e a transformação de coordenadas da câmera para o mundo calcular a matriz **P**.
2. Mapeamento de pontos no espaço na imagem digital.

Solução do exercício 1

- Câmera:
 - Imagem de 640x480 pixels;
 - Tamanho do \Rightarrow CCD 5x3mm;
 - Distância focal \Rightarrow 20mm;
- Transformação da câmera para mundo em mm:

$$\mathbf{T} = \begin{bmatrix} 0 \\ 0 \\ 2000 \end{bmatrix} \quad \mathbf{R}_C^M = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

- Centro da imagem:
 - $u_0 = 310$ pixels;
 - $v_0 = 230$ pixels;

Solução do exercício 1

- Fatores de escala:

$$k_x = \frac{640}{5} = 128 \text{ pixels/mm} \quad k_y = \frac{480}{3} = 160 \text{ pixels/mm}$$

- Matriz \mathbf{P} de formação da imagem:

$$\mathbf{P} = \mathbf{K}\mathbf{H}_c^M = \begin{bmatrix} \alpha_x & 0 & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_c^M & \mathbf{T} \\ \mathbf{0} & 1 \end{bmatrix}$$
$$\mathbf{P} = \begin{bmatrix} 128x20 & 0 & 310 & 0 \\ 0 & 160x20 & 230 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2000 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2560 & 310 & 620x10^3 \\ -3200 & 0 & 230 & 460x10^3 \\ 0 & 0 & 1 & 2000 \end{bmatrix}$$

Solução do exercício 2

- Dada a matriz de formação de imagem calculada no exercício 1, calcule a posição dos seguintes pontos no espaço no plano da imagem em pixel.

$$\mathbf{P}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} mm \quad \mathbf{P}_2 = \begin{bmatrix} 0 \\ 0 \\ 1000 \end{bmatrix} mm \quad \mathbf{P}_3 = \begin{bmatrix} 200 \\ -10 \\ 500 \end{bmatrix} mm$$