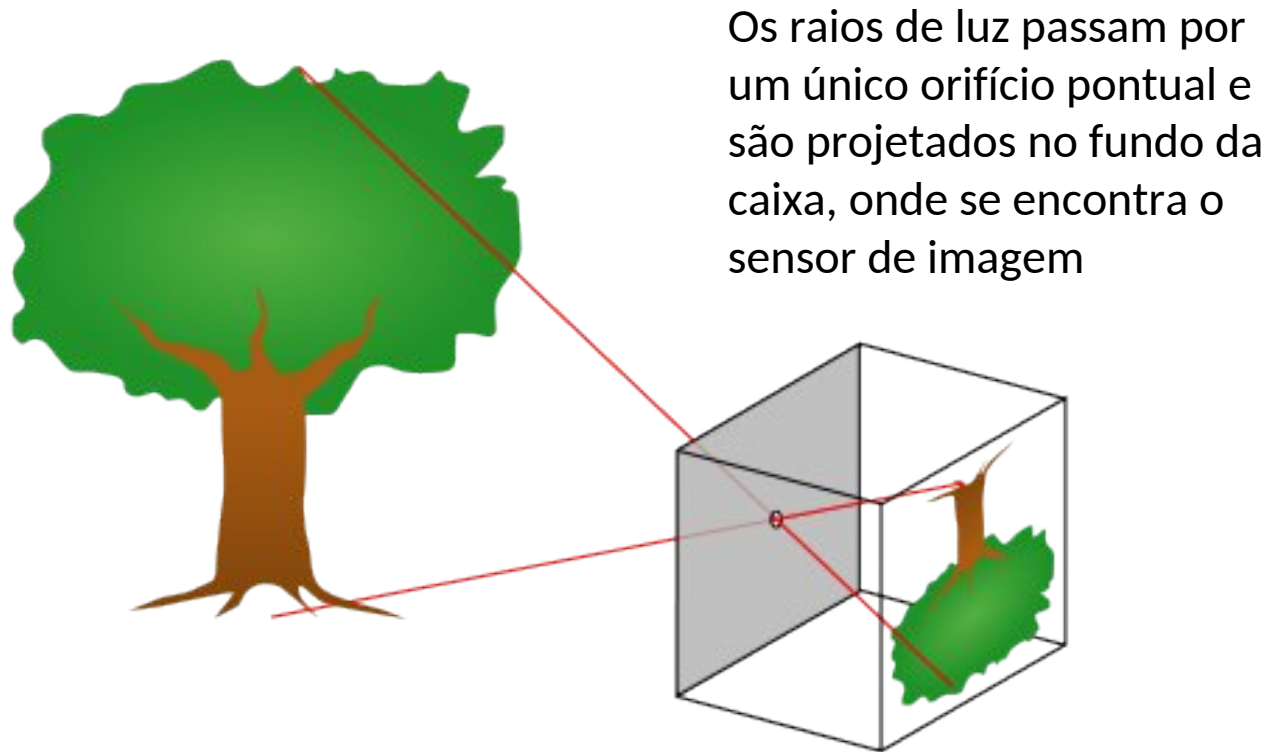


Robótica Computacional

Captura de Imagens

Modelo de câmera Pinhole



Extraído de http://pt.wikipedia.org/wiki/Câmera_pinhole

Projeção perspectiva

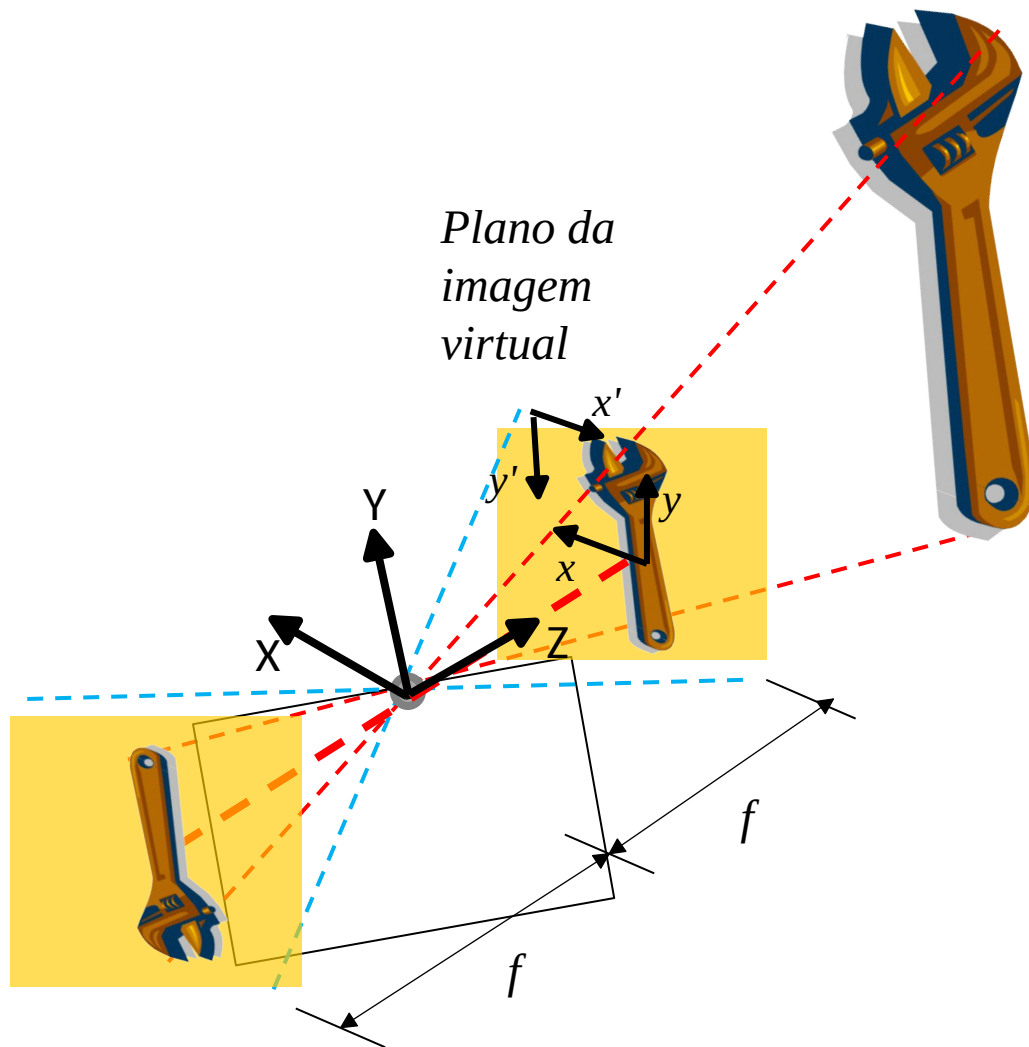
Nas câmeras de vídeo, são as lentes que fazem a projeção da imagem dos objetos no plano do sensor.

Matematicamente, essa projeção é expressa por: $(x,y)=M(X,Y,Z)$, onde (x,y) são as coordenadas do sensor e (X,Y,Z) são as coordenadas no mundo real.

É possível descrever essa projeção de diversas formas, de acordo com o nível de detalhamento dos fenômenos óticos.

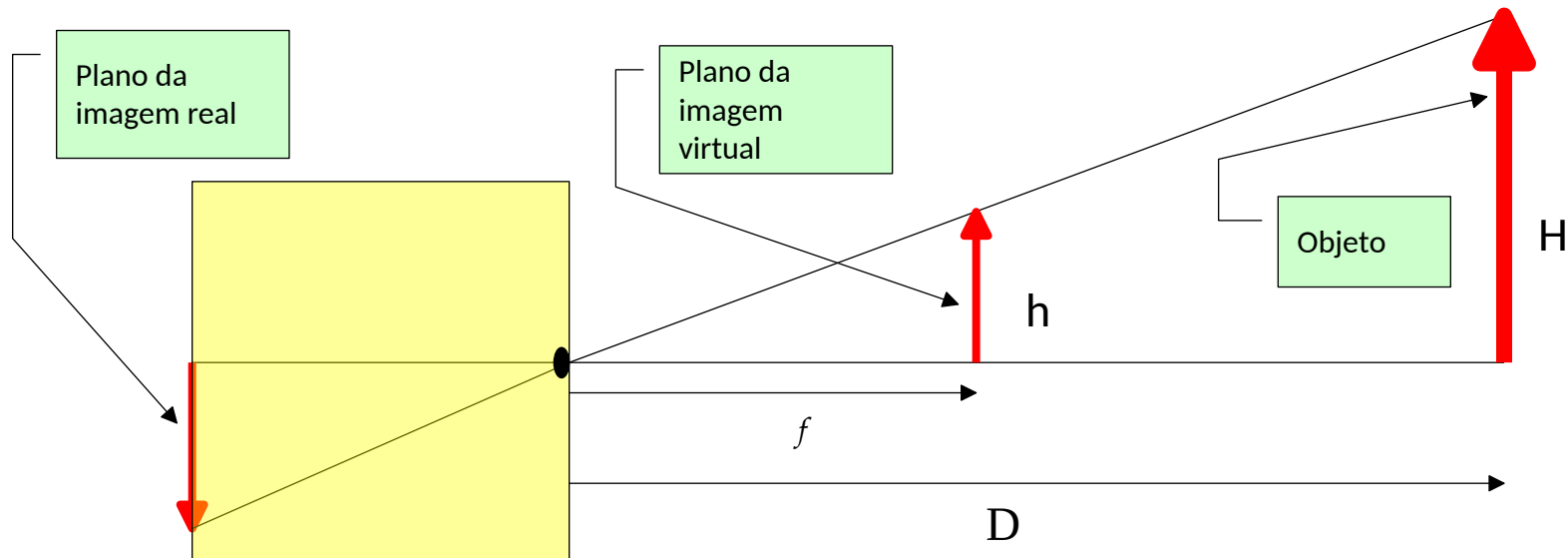
O modelo mais comumente usado é o *pinhole*, ou buraco de alfinete, que descreve como a luz é projetada através de um pequeno orifício dentro de uma caixa fechada.

Sistemas de coordenadas do modelo



- X, Y, Z : coordenadas da câmera
- x, y : coordenadas do plano da imagem
- x', y' : coordenadas da imagem em pixels
- f é a distância focal
- A resolução do sensor é c , podendo ser expressa em pixels por polegada ou pixels por mm

Projeção perspectiva simplificada



Semelhança de triângulos!

- Se tenho 3 entre f , h , H e D consigo encontrar os outros!
- Calibração (controlar o ambiente para descobrir parâmetros “desconhecidos”)

$$h = \frac{f}{D} H$$

Parâmetros do modelo pinhole

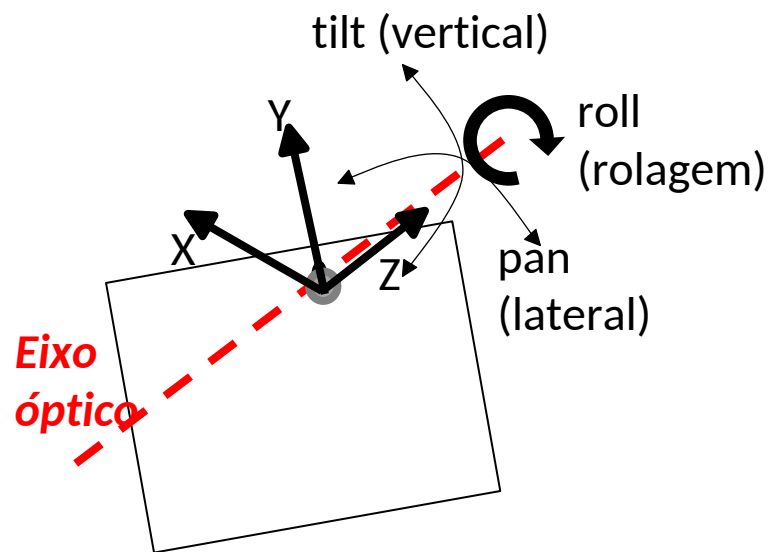
Segundo o modelo pinhole, a câmera possui parâmetros intrínsecos e extrínsecos:

Parâmetros intrínsecos: dependem exclusivamente do modelo de câmera utilizada

- Distância focal (f): distância entre o centro óptico (o ponto de maior convergência da luz) até a superfície de formação da imagem
- Tamanho e resolução do sensor: definem o tamanho da imagem, em pixels
- Ponto principal: posição na imagem do ponto onde o eixo óptico atravessa o sensor de imagem. Idealmente é posição do pixel central, mas em geral o sensor apresenta um leve deslocamento

Parâmetros extrínsecos: representados por um vetor de translação e uma matriz de rotação 3D, indicam o posicionamento da câmera com relação aos objetos sendo imageados

Relação entre as projeções dos pontos dos objetos na imagem

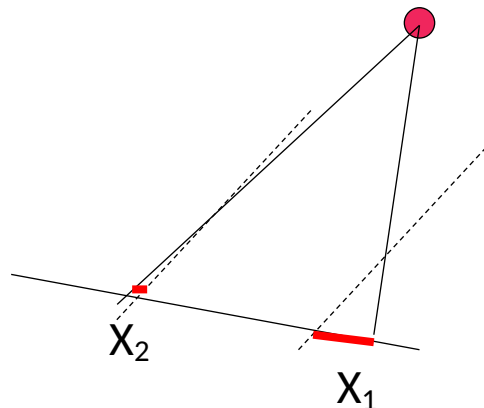
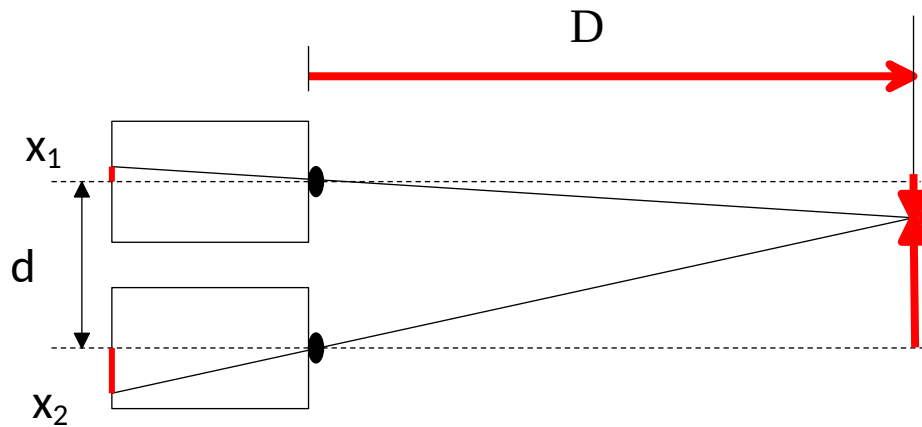


Estereoscopia

- Na formação da imagem, projetamos os objetos 3D e 2D, perdendo a dimensão da profundidade (Z)
- Pode ser de interesse descobrir a profundidade dos objetos na imagem. Para isso, precisamos de duas imagens do mesmo objeto, tirado por duas câmeras diferentes
- Cada ponto do cenário aparecerá em ambas as imagens em posições distintas.
- Às diferenças nas posições são chamadas *disparidades*.
- A partir da disparidade, é possível calcular a profundidade desse ponto

Câmera estéreo

- No caso de termos duas câmeras alinhadas horizontalmente, e apontando na mesma direção, os pontos aparecerão na mesma altura nas imagens de ambas as câmeras, mas em posições diferentes



Vamos calcular a profundidade?

$$x_1 = \frac{f}{Z_1} X_1$$

$$x_2 = \frac{f}{Z_2} X_2$$

$$X_2 = X_1 + d$$

$$Z_1 = Z_2 = D$$

$$D = \frac{fd}{x_2 - x_1}$$

Cálculo da profundidade



Observem que quanto maior a disparidade ($x_1' - x_2'$) menor a distância do objeto. Se a disparidade for zero, a distância é infinita.

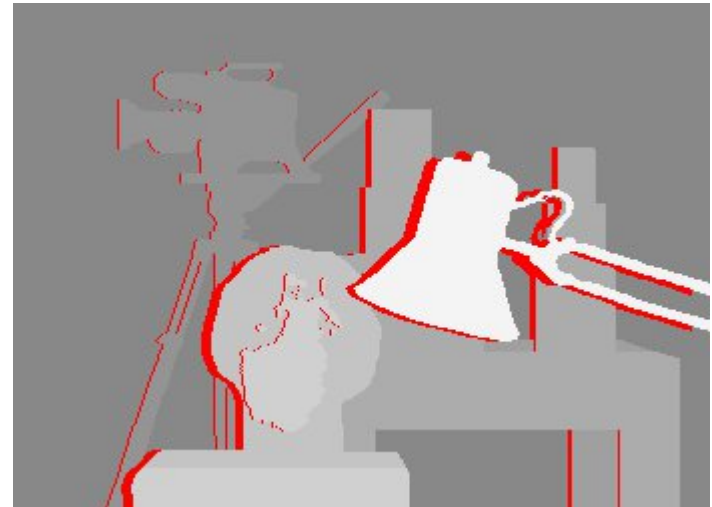


A maior dificuldade é saber que ponto na imagem de uma câmera corresponde a que ponto na imagem da outra.

Exemplo

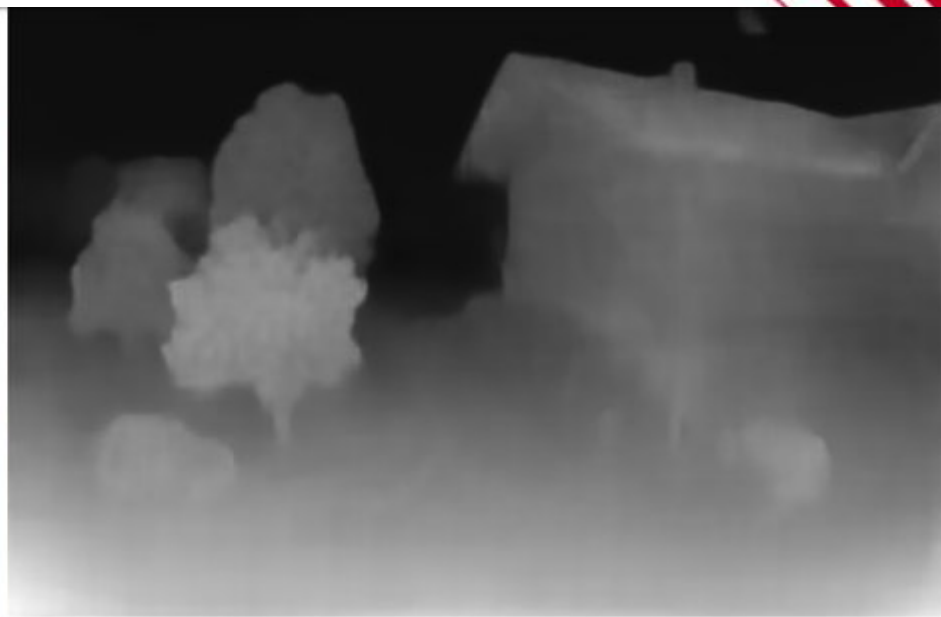


Uma das imagens tiradas por uma par de câmeras apontando na mesma direção



Profundidade dois objetos na cena. As regiões vermelhas representam as disparidades encontradas entre duas imagens

Profundidade sem estereoscopia



Atividade para entrega

- Calibrar uma câmera para encontrar distâncias entre objetos no mundo real
- Segmentar por cores e encontrar objetos nas máscaras criadas

Dada uma distância em pixels, encontrar a distância equivalente no “mundo real”