

VISÃO COMPUTACIONAL

AULA 3

IMAGENS DIGITAIS

Parâmetros de Câmeras, Dados e
Sensores de Profundidade

IMAGENS DIGITAIS

PARÂMETROS DE CÂMERAS

- Definições

Assume-se que

- O sistema de referência da câmera pode ser localizado com respeito a algum outro sistema conhecido, p.ex. o *sistema de coordenadas externo*.
- As coordenadas dos pontos da imagem no sistema de referência da câmera podem ser obtidas das *coordenadas de píxeis*, os únicos diretamente disponíveis da imagem.

Os parâmetros das câmeras são divididos em

- Parâmetros extrínsecos

- Definem a localização e orientação do sistema de referência da câmera com respeito a um sistema externo conhecido

Sistemas de Coordenadas: **Externo (mm) \Longleftrightarrow Câmera (mm)**

- Parâmetros intrínsecos

- Necessários para relacionar as coordenadas de píxeis de uma imagem com as coordenadas correspondentes no sistema de referência da câmera.

Sistemas de Coordenadas: **Imagem (pixel) \Longleftrightarrow Câmera (mm)**

- Parâmetros Extrínsecos

- Qualquer conjunto de parâmetros geométricos que identificam univocamente as transformações entre o sistema de referência da câmera desconhecido e um sistema conhecido, chamado sistema de coordenadas externo.

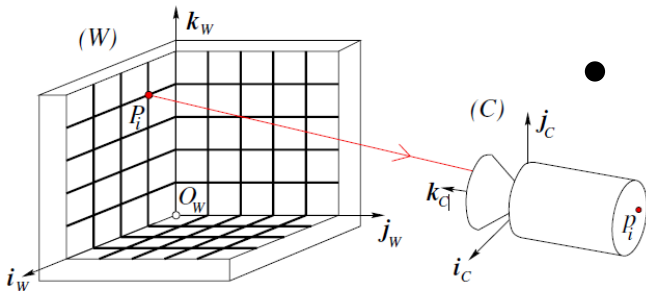
- Uma escolha típica é

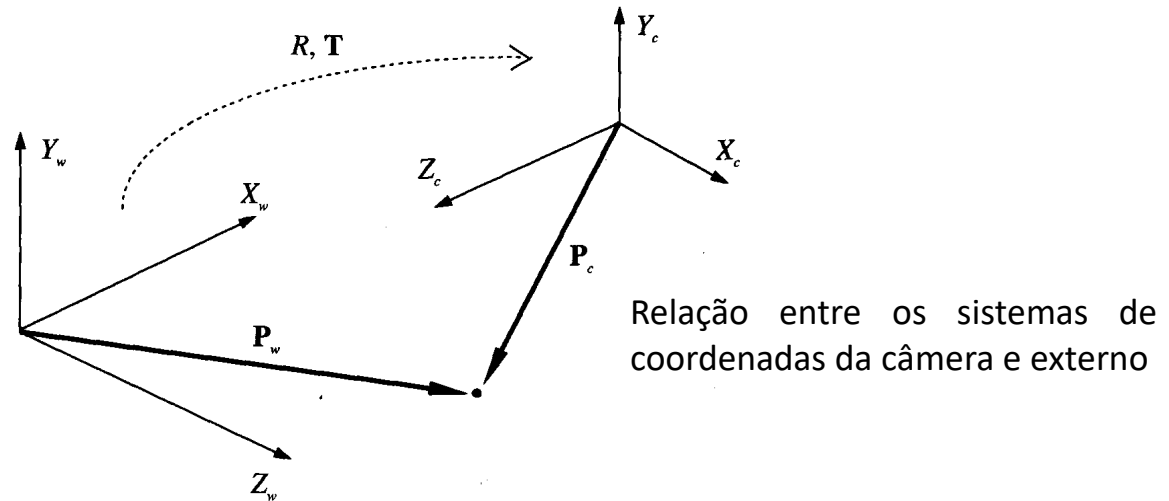
- Um vetor de translação 3-D, \mathbf{T}

- Posições relativas entre a origem de dois sistemas

- Uma matriz de rotação 3x3, ortogonal, \mathbf{R}

- Alinha os eixos de dois sistemas





Um ponto \mathbf{P} pode ser representado no sistema externo e da câmera como \mathbf{P}_w e \mathbf{P}_c respectivamente e

$$\mathbf{P}_c = \mathbf{R} \cdot (\mathbf{P}_w - \mathbf{T})$$

e

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

- Por definição os parâmetros extrínsecos são o vetor de translação, \mathbf{T} , e a matriz de rotação, \mathbf{R} , que especificam a transformação entre o sistema de coordenadas externo e o sistema de referência da câmera.

- Parâmetros Intrínsecos

- Caracterizam as características ópticas, geométricas e digitais da câmera. Em um modelo *pinhole* os p.i. são

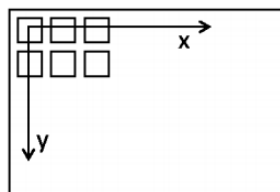
- A projeção de perspectiva, unicamente com f
- Transformação entre os sistemas da câmera e imagem

Assume-se que não haja distorções geométricas ópticas, e que o sensor CCD é uma grade retangular de elementos fotossensíveis

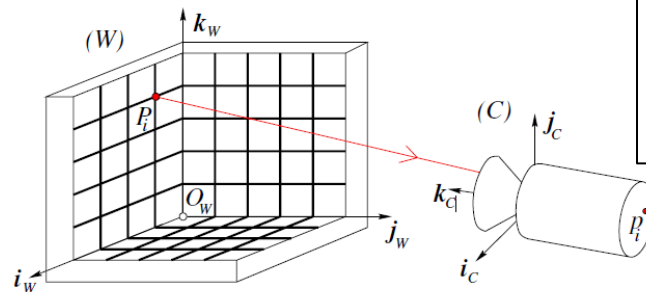
$$x = -(x_{im} - o_x) \cdot s_x \quad \text{e} \quad y = -(y_{im} - o_y) \cdot s_y$$

(o_x, o_y) = coord. do centro da imagem em píxeis (ponto principal)

(s_x, s_y) = tamanhos efetivos dos píxeis (em milímetros) nas direções horizontais e verticais respectivamente.

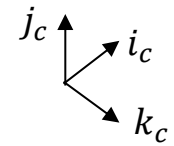


(x_{im}, y_{im})



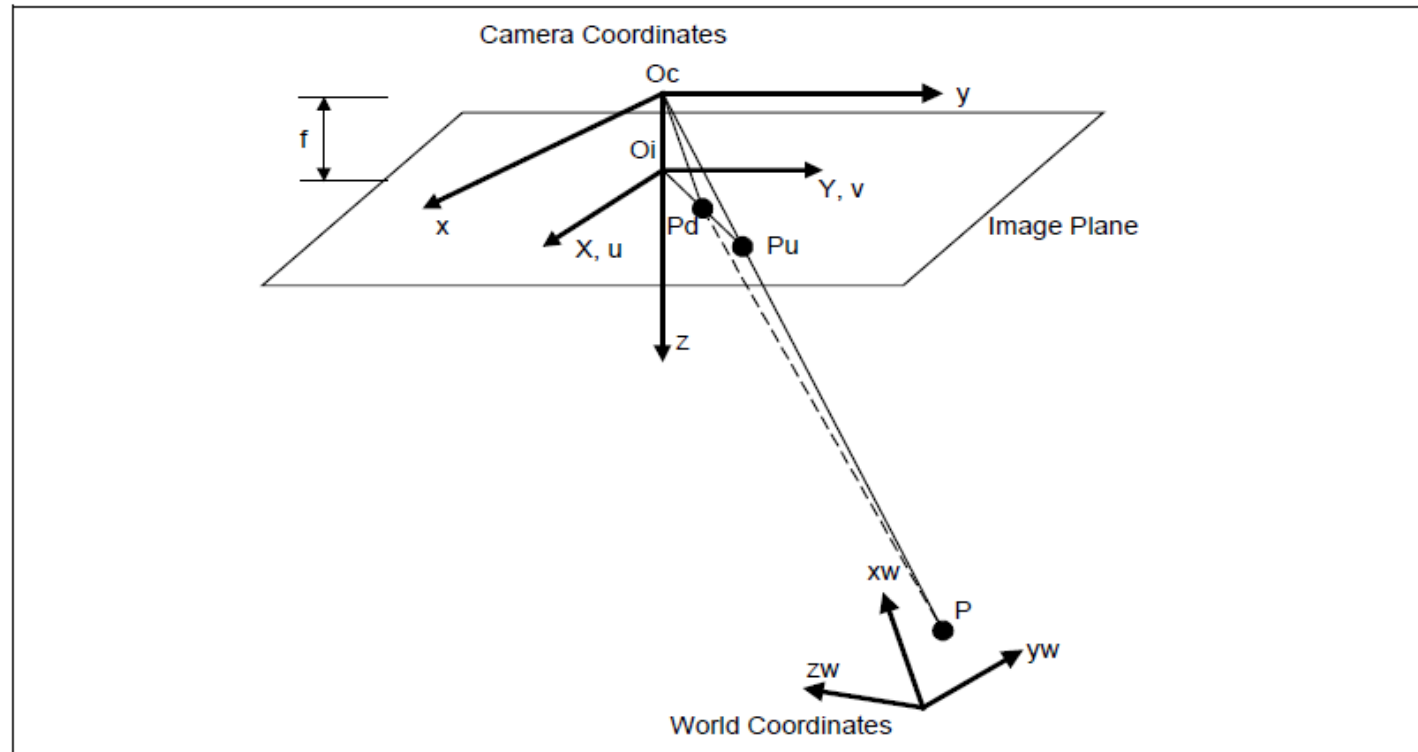
O sentido de x_c e z_c pode também ser oposto:

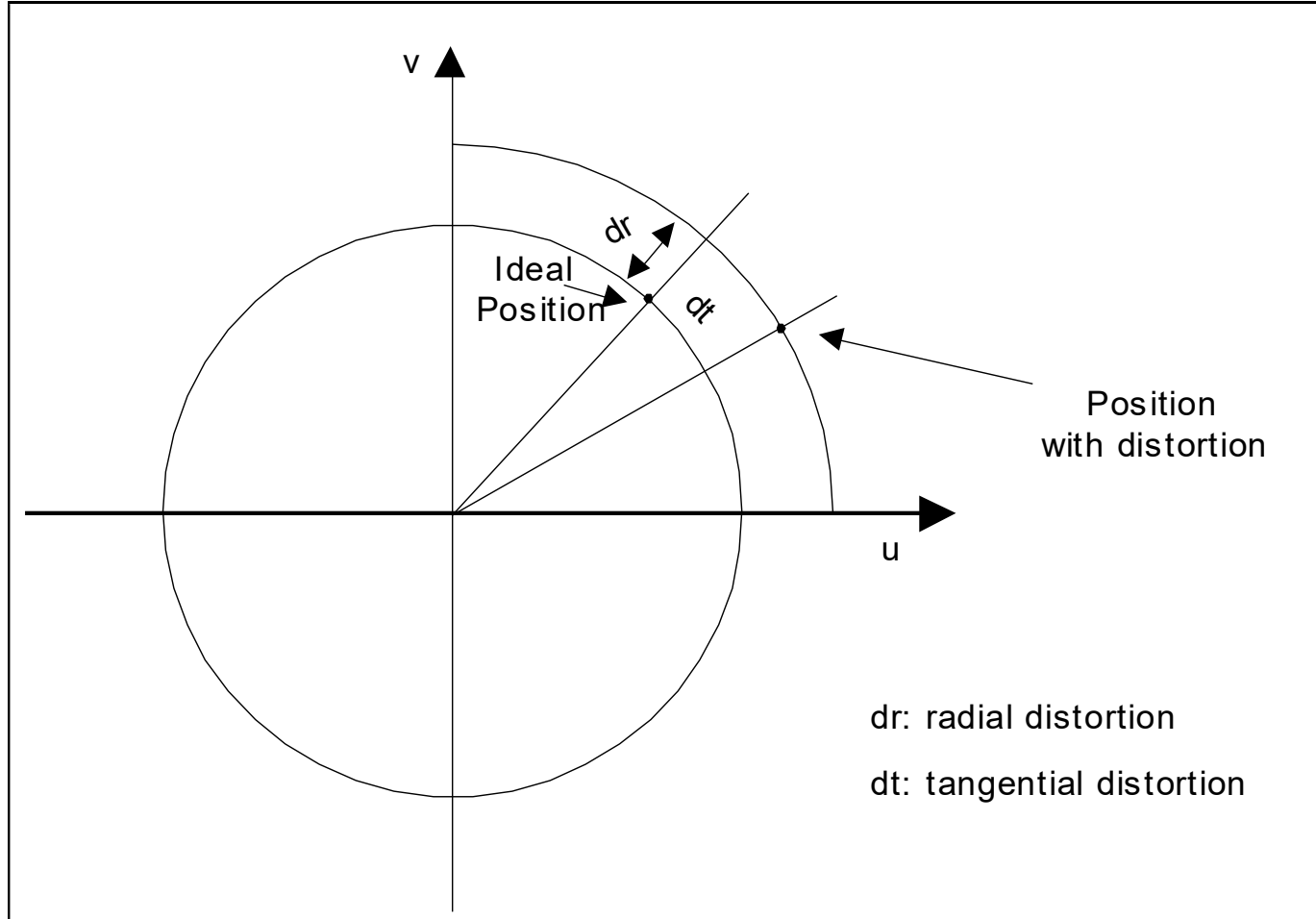
$$x = (x_{im} - o_x) \cdot s_x$$

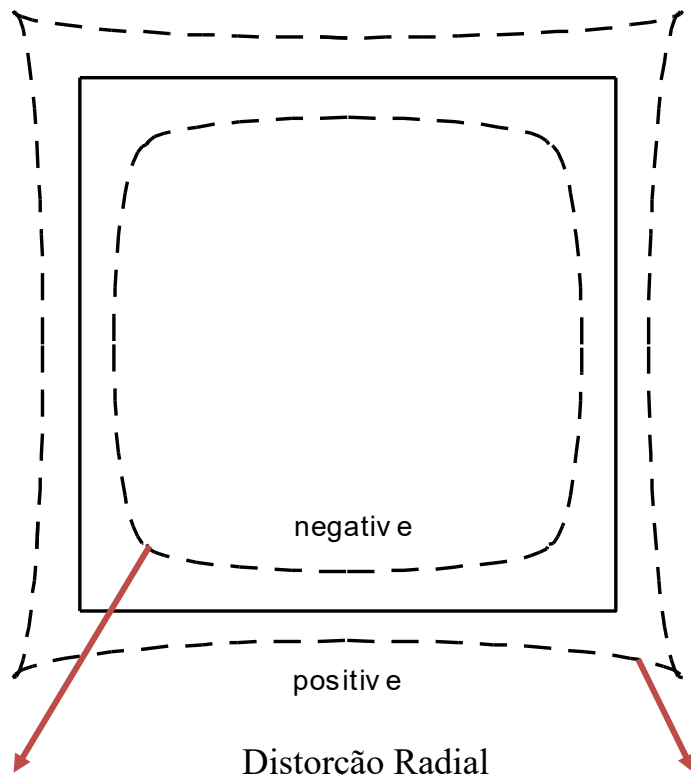


(x_c, y_c)

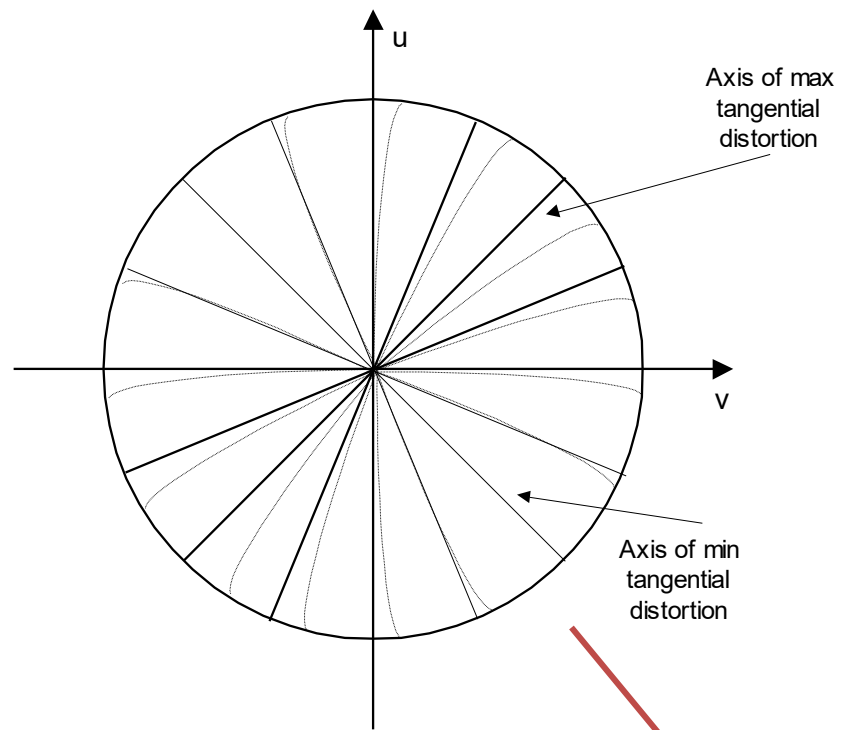
- Distorções geométricas introduzidas pelo sistema óptico.



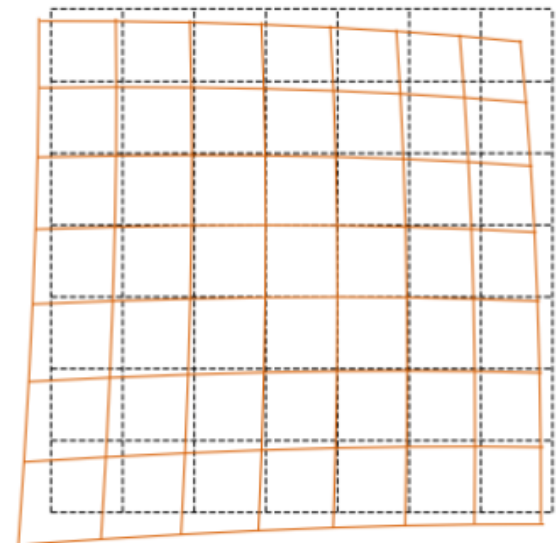




Distorção Radial



Distorção Tangencial



- A distorção radial é dominante, e aumenta com um maior campo de visão

$$x = x_d \cdot (1 + k_1 \cdot r^2 + k_2 \cdot r^4)$$

$$y = y_d \cdot (1 + k_1 \cdot r^2 + k_2 \cdot r^4)$$

(x_d, y_d) = coordenadas do ponto distorcido, observável

(x, y) = coordenadas do ponto corrigido

$$r^2 = x_d^2 + y_d^2$$

Parâmetros Intrínsecos

f = distância focal

o_x, o_y = Centro da Imagem (em coord. pixels)

s_x, s_y = fatores de escala

k_1 = coef. de distorção radial

- **Modelos de Câmeras**

Pode-se então, agora, relacionar o sistema externo diretamente às coordenadas de imagem, sem passar pelo sistema da câmera como na transformação de perspectiva.

- Versão Linear das Equações de Projeção de Perspectiva

Juntando

$$P_c = R.(P_w - T) , \quad x = -(x_{im} - o_x).s_x \text{ e } y = -(y_{im} - o_y).s_y \quad ; \quad (P_c = [X,Y,Z]^T)$$

e as equações fundamentais da projeção perspectiva

$$x = f.\frac{X}{Z} \quad ; \quad y = f.\frac{Y}{Z}$$

tem-se

$$-(x_{im} - o_x).s_x = f.\frac{R_1^T.(P_w - T)}{R_3^T.(P_w - T)}$$

$$-(y_{im} - o_y).s_y = f.\frac{R_2^T.(P_w - T)}{R_3^T.(P_w - T)}$$

Eq. 1

onde R_i , $i = 1, 2, 3$ é um vetor 3-D formado pela i -ésima linha de R .

As equações anteriores podem ser reescritas como uma multiplicação matricial simples. Definindo 2 matrizes M_{int} e M_{ext} como

$$M_{\text{int}} = \begin{pmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{pmatrix} \quad M_{\text{ext}} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & -R_1^T T \\ r_{21} & r_{22} & r_{23} & -R_2^T T \\ r_{31} & r_{32} & r_{33} & -R_3^T T \end{pmatrix}$$

M_{int} depende apenas dos parâmetros intrínsecos
 M_{ext} depende apenas dos parâmetros extrínsecos

Equação Matricial Linear de Projeções de Perspectiva

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = M_{\text{int}} \cdot M_{\text{ext}} \cdot \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

em que

$$x_1/x_3 = x_{\text{im}} \quad \text{e} \quad x_2/x_3 = y_{\text{im}}$$

Portanto

- M_{ext} realiza a transformação entre o sistema de coordenadas externo e o sistema de referência da câmera.
- M_{int} realiza a transformação entre o sistema de referencia da câmera e o sistema de coordenadas da imagem.

○ O Modelo de Câmera Perspectiva

Assumindo $o_x = o_y = 0$ e $s_x = s_y = 1$, $M = M_{\text{int}} \cdot M_{\text{ext}}$ é

$$M = \begin{pmatrix} -f.r_{11} & -f.r_{12} & -f.r_{13} & f.R_1^T T \\ -f.r_{21} & -f.r_{22} & -f.r_{23} & f.R_2^T T \\ r_{31} & r_{32} & r_{33} & -R_3^T T \end{pmatrix}$$

M descreve o modelo de câmera de perspectiva completa (full-perspective), e é chamada ***Matriz de Projeção***

- Não se preservam distâncias entre pontos, ou ângulos entre linhas
- Mapeiam linhas em linhas

○ O Modelo de Câmera Perspectiva Fraca

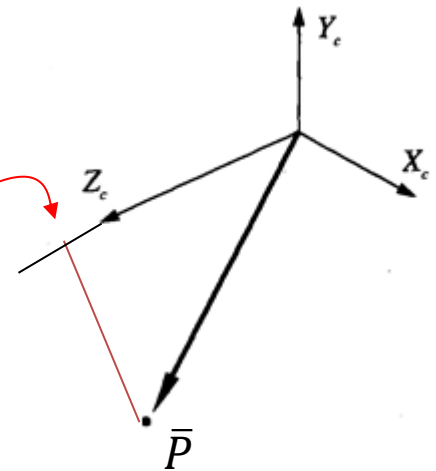
De acordo com a equação Equação Matricial Linear de Projeções de Perspectiva, a imagem \mathbf{p} de um ponto \mathbf{P}_w é

$$\mathbf{p} = \mathbf{M} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \mathbf{M} \cdot \begin{bmatrix} f \cdot \mathbf{R}_1^T \cdot (\mathbf{T} - \mathbf{P}_w) \\ f \cdot \mathbf{R}_2^T \cdot (\mathbf{T} - \mathbf{P}_w) \\ \mathbf{R}_3^T \cdot (\mathbf{T} - \mathbf{P}_w) \end{bmatrix}$$

Eq. 1

Mas $\|\mathbf{R}_3^T \cdot (\mathbf{P}_w - \mathbf{T})\|$ é a distância de \mathbf{P} do centro de projeção ao longo do eixo óptico; portanto, a restrição básica para a aproximação de perspectiva fraca é

$$\left| \frac{\mathbf{R}_3^T \cdot (\mathbf{P}_i - \bar{\mathbf{P}})}{\mathbf{R}_3^T \cdot (\bar{\mathbf{P}} - \mathbf{T})} \right| \ll 1$$



onde $\mathbf{P}_1, \mathbf{P}_2$ são dois pontos no espaço 3-D, e $\bar{\mathbf{P}}$ o centróide de \mathbf{P}_1 e \mathbf{P}_2 .

Fazendo a substituição ($P = P_i$)

$$p_i \approx \begin{bmatrix} f.R_1^T.(T-P_{wi}) \\ f.R_2^T.(T-P_{wi}) \\ R_3^T.(\overline{P_w}-T) \end{bmatrix}$$

e a matriz de projeção M se torna

$$M_{wp} = \begin{bmatrix} -f.r_{11} & -f.r_{12} & -f.r_{13} & f.R_1^T.T \\ -f.r_{21} & -f.r_{22} & -f.r_{23} & f.R_2^T.T \\ 0 & 0 & 0 & R_3^T.(\overline{P_w}-T) \end{bmatrix}$$

- O Modelo de Câmera Afim (Affine)

$$M_{\text{aff}} = \begin{pmatrix} -f.r_{11} & -f.r_{12} & -f.r_{13} & f.R_1^T T \\ -f.r_{21} & -f.r_{22} & -f.r_{23} & f.R_2^T T \\ 0 & 0 & 0 & -R_3^T.T \end{pmatrix}$$

M perspect. completa com
última linha da projeção = 0

- Não preserva ângulos, mas preserva paralelismo
- Em comparação com o modelo de persp. fraca, apenas as razões entre distâncias medidas ao longo de direções paralelas são preservadas.

Modelo Câmera Perspectiva

$$\mathbf{M} = \begin{pmatrix} -f.r_{11} & -f.r_{12} & -f.r_{13} & f.R_1^T T \\ -f.r_{21} & -f.r_{22} & -f.r_{23} & f.R_2^T T \\ r_{31} & r_{32} & r_{33} & -R_3^T T \end{pmatrix}$$

Modelo Câmera Perspectiva Fraca

$$\mathbf{M}_{wp} = \begin{pmatrix} -f.r_{11} & -f.r_{12} & -f.r_{13} & f.R_1^T T \\ -f.r_{21} & -f.r_{22} & -f.r_{23} & f.R_2^T T \\ 0 & 0 & 0 & R_3^T . (\bar{P} - T) \end{pmatrix}$$

Modelo Câmera Affine

$$\mathbf{M}_{aff} = \begin{pmatrix} -f.r_{11} & -f.r_{12} & -f.r_{13} & f.R_1^T T \\ -f.r_{21} & -f.r_{22} & -f.r_{23} & f.R_2^T T \\ 0 & 0 & 0 & -R_3^T . T \end{pmatrix}$$

DADOS E SENSORES DE PROFUNDIDADE

- Nas imagens de profundidade cada pixel expressa a *distância entre um sistema de referência conhecido e um ponto visível na cena.*
- Reproduz a *estrutura 3-D* de uma cena
- Pode ser vista como uma *superfície amostrada*
- Também denominadas mapa de profundidade, ou mapas xyz, ou perfil de superfície, ou imagens 2.5-D

- **Representação de Imagens de Profundidade**

- *Forma xyz ou nuvem de pontos*

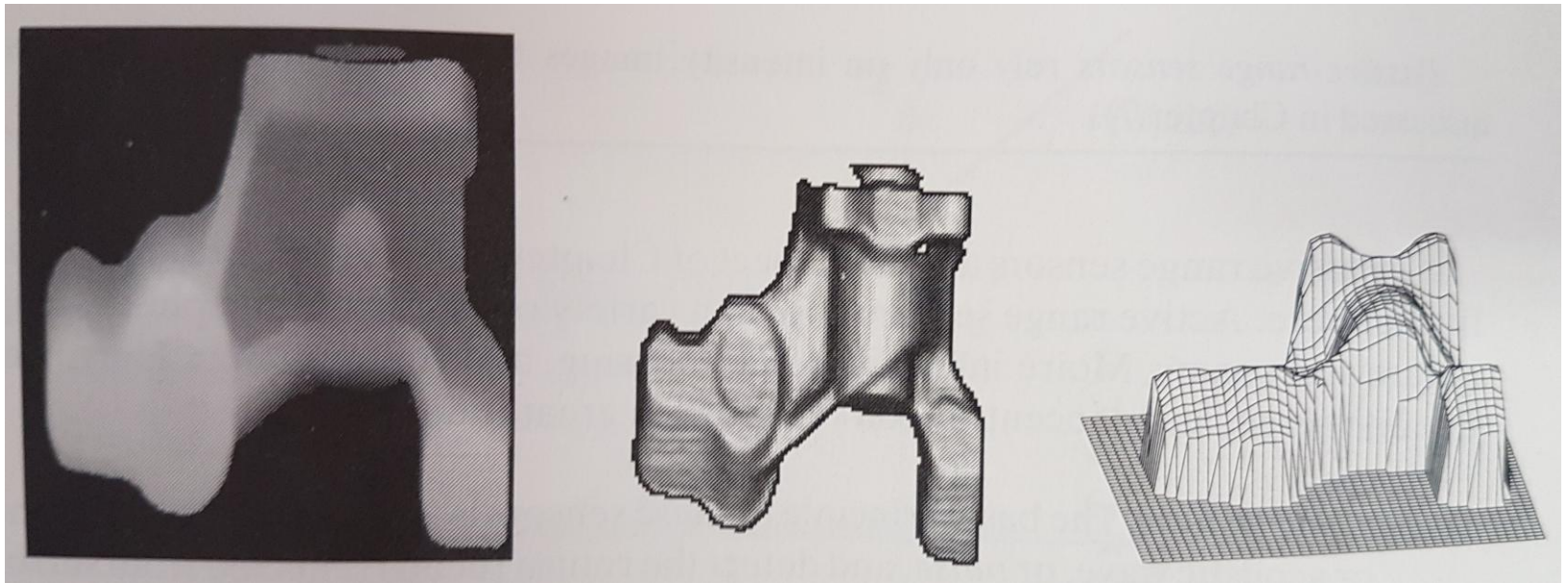
- Lista de coordenadas 3-D
 - Não há ordem específica

- *Forma r_i*

- Matrizes de valores de profundidade de pontos ao longo das direções x,y da imagem
 - Fornece informação espacial explícita

- *Sombreamento cossenoidal*

- Valor de intensidade do pixel é proporcional à norma do gradiente nas superfícies de profundidade



Imagens de Profundidade de um componente mecânico mostradas como imagem de intensidade. (mais claro mais próximo (esquerda), sombreamento cossenoidal (meio), e superfície 3-D (direita).

- Sensores de Profundidade

- Medem profundidade em apenas um ponto, ou
- Distância e forma de perfil de superfícies ou de superfícies completas

1. Ativos

- Projetam energia (p. ex. padrão de luz, pulsos de sonar) na cena e detectam sua posição para realizar a medida, ou
- Exploram o efeito de mudanças controladas de parâmetros, p. ex. foco, ou calculam tempo de percurso da luz (TOF)

2. Passivos

- Baseiam-se unicamente em imagens de intensidade para reconstruir a profundidade, p. ex. estereoscopia

- Ativos

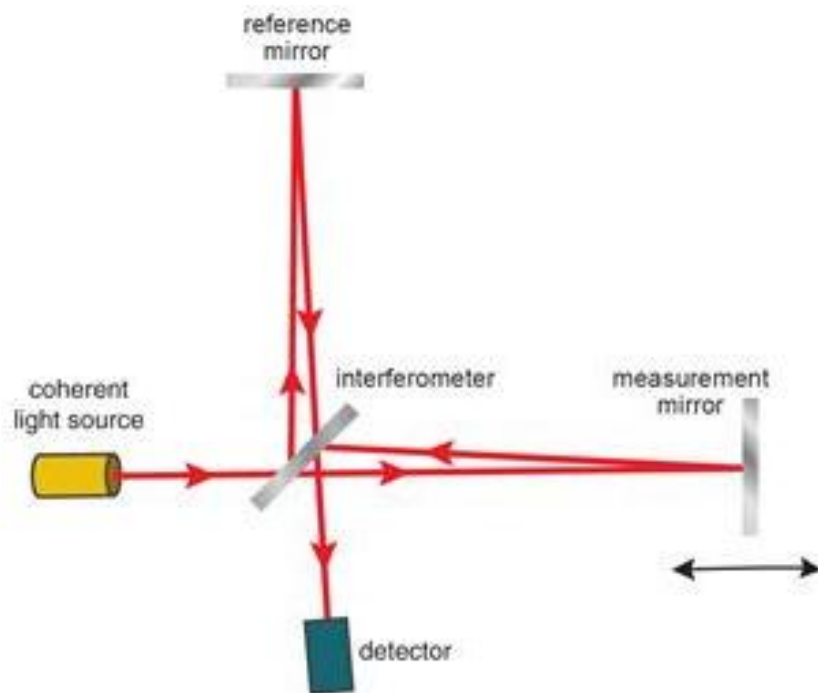
- Radares e Sonares (RADAR = Radio Detection and Ranging)

- Baseado no cálculo do tempo de resposta de um sinal eletromagnético ou sonoro (time of flight).
 - Usado para varrer uma superfície.

- Detectores a laser (LADAR ou LIDAR – Laser (Light) Detection and Ranging) emitem sinal com amplitude modulada e medem a diferença de fase.

- Interferometria

- Duas malhas padronizadas (p. ex. linhas) são projetadas em uma superfície e a diferença de fase ou distância entre os dois padrões fornece a distância 3D.
- Mede apenas distância relativa



- Focalização/Desfocalização Ativa

- Variação da distância focal de uma lente motorizada permite calcular a distância 3D

Podem ser por:

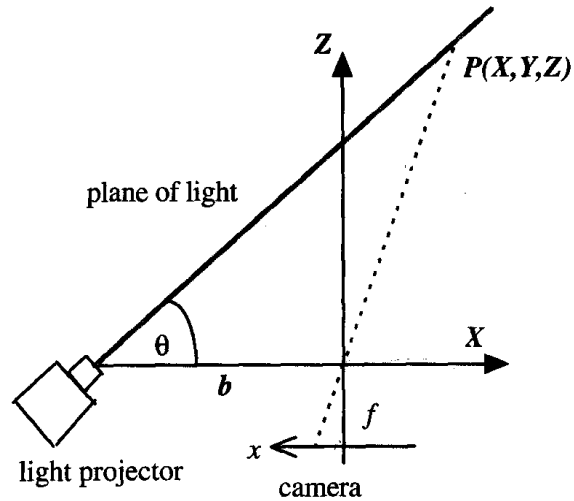
- *Detecção de contraste*

O sistema analisa um grupo de pixels na imagem e move a lente para frente e para trás em pequenos incrementos

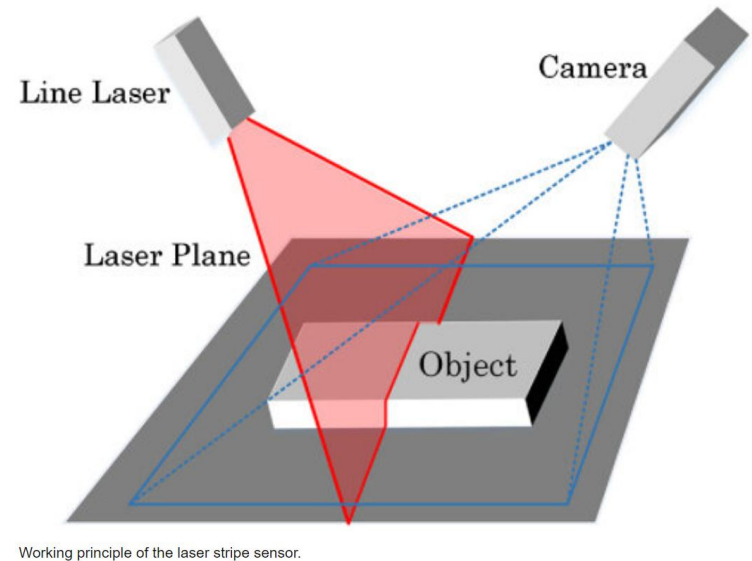
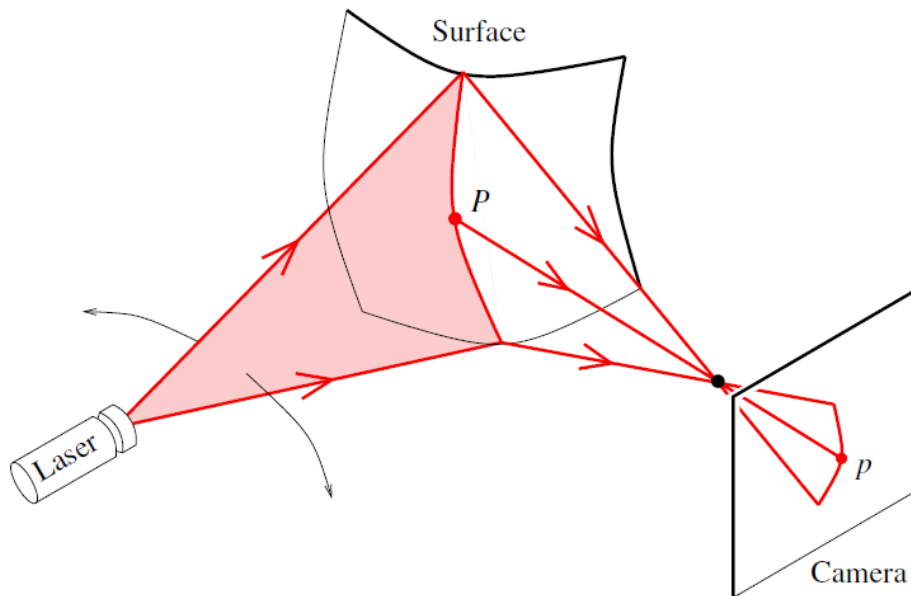
- *Detecção de fase*

O sistema divide a luz que entra em duas imagens separadas e as compara.

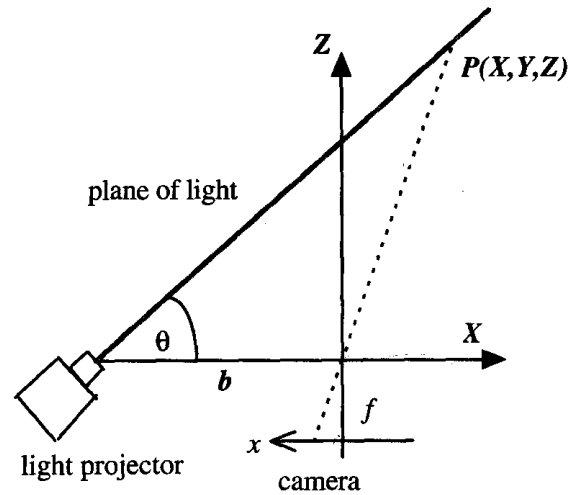
- **Triangulação Ativa**



Geometria da triangulação óptica ativa (vista plana XZ). Os eixos Y e y são perpendiculares ao plano da figura.



- **Triangulação Ativa**



Geometria da triangulação óptica ativa (vista plano XZ). Os eixos Y e y são perpendiculares ao plano da figura.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{b}{f \cdot \cot \theta - x} \cdot \begin{pmatrix} x \\ y \\ f \end{pmatrix}$$

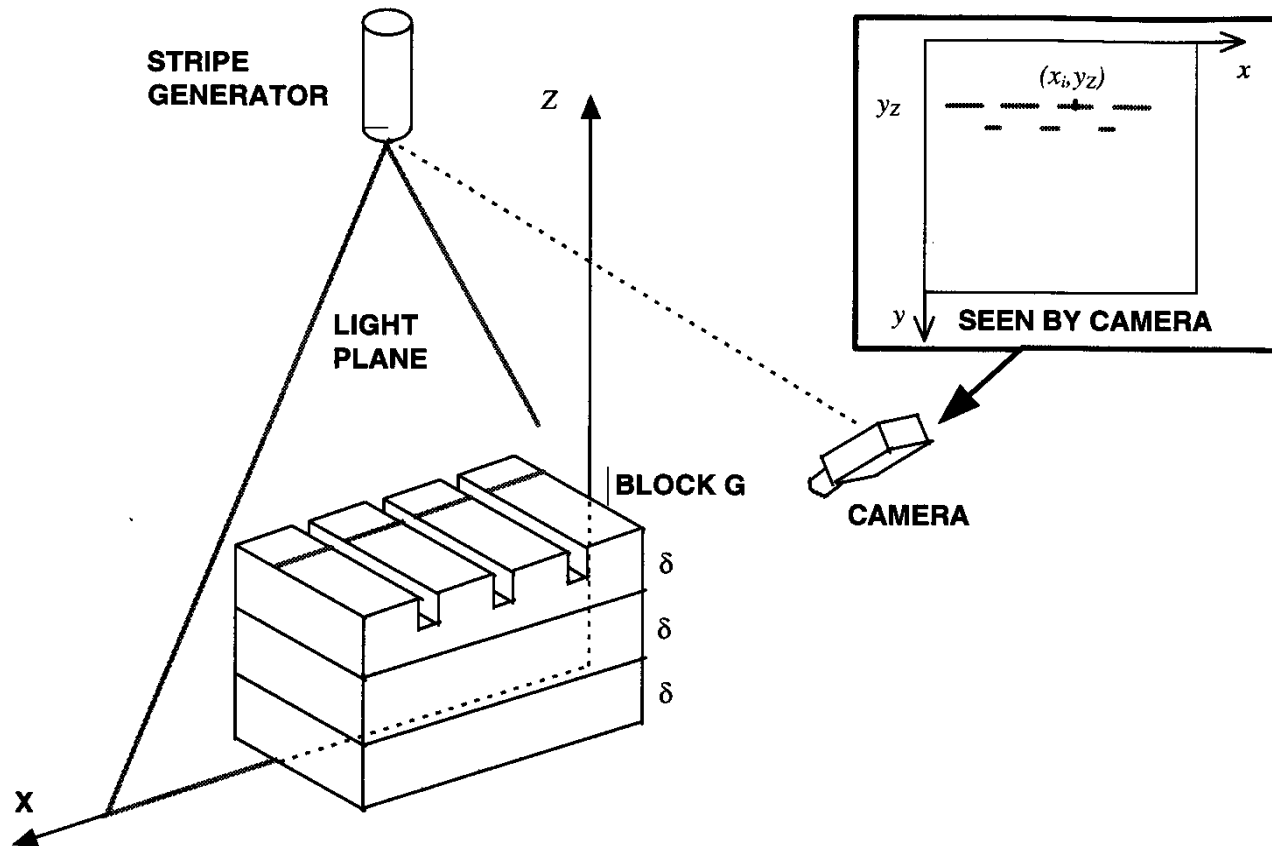
b = linha base conhecida

f = distância focal conhecida

θ = ângulo monitorado

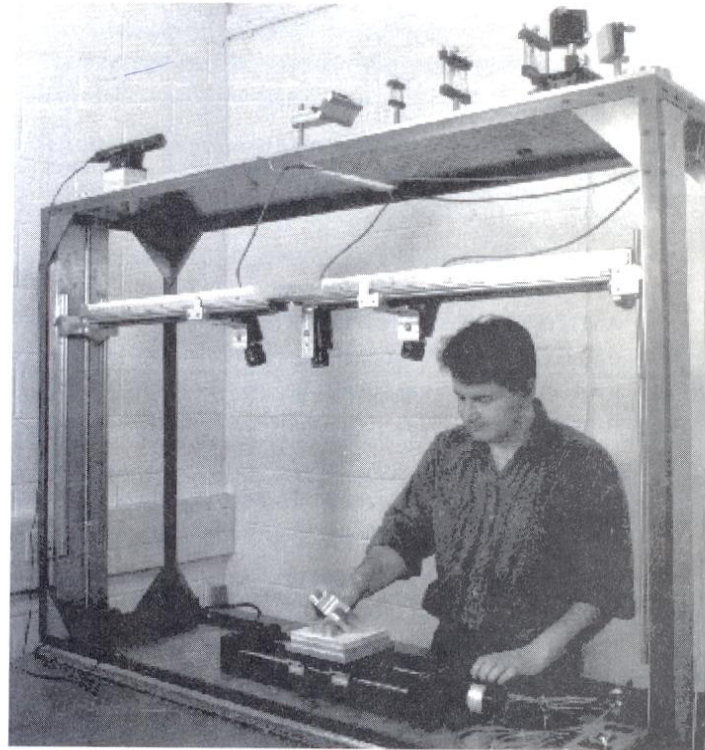
- Um Sensor Simples

Método para calibrar f , b e θ sem o uso da equação



Calibração direta de um sensor de profundidade de perfil simples

- A equação é aplicada a todos os pontos visíveis na linha de luz
- A varredura do objeto, movendo-se a camera ou o objeto, permite a determinação do perfil da superfície
- A linha projetada deve ter contraste com o fundo (objeto), e concavidades ou pontos brilhantes podem confundir a detecção da linha.
- Mais de uma câmera pode ser utilizada para evitar oclusões



Sistema de triangulação 3D. Diodo laser e sistema óptico projetam plano de luz laser sobre componente e câmeras capturam imagem com luz estruturada. Componente se move sobre plataforma motorizada.

Algoritmo RANGE_CAL

1. Colocar o bloco G debaixo da linha de luz, com rasgos perpends. ao plano da luz. Assegure que a linha apareça paralela a x (y cte.).
2. Adquira uma imagem da linha sobre G. Ache as coords. dos pontos da linha sobre a superfície mais alta do bloco, varrendo as colunas da imagem.
3. Calcule as coords. $[x_i, y_Z]^T$, $i = 1, \dots, n$ dos centros dos segmentos da linha sobre a parte alta do bloco G, localizando o centro dos segmentos na linha varrida $y = y_Z$. Entre com cada ponto de imagem $[x_i, y_Z]^T$ e seus correspondentes pontos 3-D $[X, Z]^T$ (conhecidos) na tabela T.
4. Ponha um outro bloco sob G, elevando a sua superfície de δ . Assegure-se de que as condições do passo 1 ainda se aplicam. Seja cuidadoso para não mover o sistema de referência.

5. Repita os passos 2, 3 e 4 até a superfície do topo do bloco G ser visualizada próxima a $y = 0$.
6. Converta T em uma tabela de consulta L, indexada por coords. de imagem $[x, y]^T$, com x entre 0 e $(x_{\max} - 1)$, e y entre 0 e $(y_{\max} - 1)$, e retornando $[X, Z]^T$. Para associar valores de pixels não medidos diretamente, interpole linearmente usando os quatro pontos vizinhos mais próximos.

A saída será uma Tabela de Consulta (LUT) associando coords. de pontos de imagem a coords. de pontos da cena.

Algoritmo RANGE_ACQ

Para usar a Tabela de Consulta para obter um perfil de profundidade de um objeto

1. Ponha um objeto debaixo da linha e adquira uma imagem da mesma sobre G
 2. Calcule as coords. de imagem $[x, y]^T$ dos pontos da linha varrendo cada coluna da imagem.
 3. Indexe L usando as coords. de imagem (x,y) do ponto da linha para obter os pontos de profundidade $[X, Z]^T$.
-