Sistemas de Visão e Percepção Industrial

Geometria e Imagem

Parte 2 - Geometria Projetiva

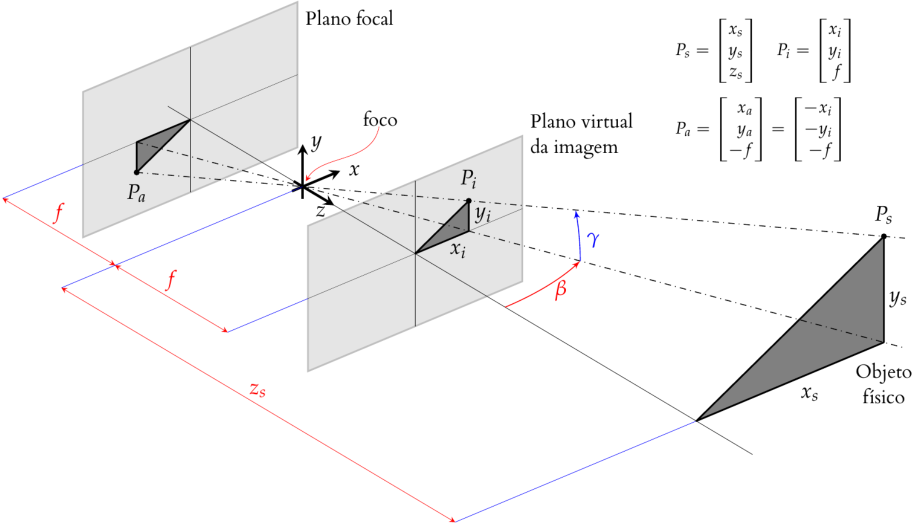
Universidade de Aveiro
Vitor Santos, 2022-2023

Sumário

Formação de imagem

Formação de imagem

- Modelo “pin-hole”



$$P_s = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} \qquad P_i = \begin{bmatrix} x_i \\ y_i \\ f \end{bmatrix}$$

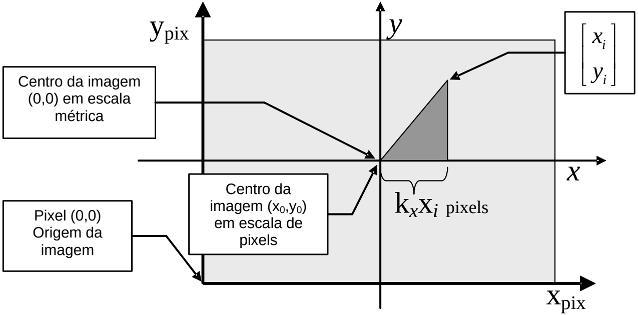
$$P_a = \begin{bmatrix} x_a \\ y_a \\ -f \end{bmatrix} = \begin{bmatrix} -x_i \\ -y_i \\ -f \end{bmatrix}$$

Relações entre coordenadas na imagem

- Pixels vs. "metros"

$$\begin{cases} \dfrac{x_s}{z_s} = \dfrac{x_i}{f} \\ \dfrac{y_s}{z_s} = \dfrac{y_i}{f} \end{cases}$$

$$\begin{cases} x_{pix} = x_0 + k_x x_i \\ y_{pix} = y_0 + k_y y_i \end{cases}$$



$$\begin{cases} x_{pix} = x_0 + k_x f \dfrac{x_s}{z_s} \\ y_{pix} = y_0 + k_y f \dfrac{y_s}{z_s} \end{cases}, \quad \begin{cases} x_{pix} = x_0 + \alpha_x \dfrac{x_s}{z_s} \\ y_{pix} = y_0 + \alpha_y \dfrac{y_s}{z_s} \end{cases}, \quad \begin{bmatrix} x_{pix} \\ y_{pix} \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \end{bmatrix} \begin{bmatrix} \dfrac{x_s}{z_s} \\ \dfrac{y_s}{z_s} \\ 1 \end{bmatrix}$$

Parâmetros intrínsecos

Matriz intrínseca

- Na forma matricial virá:

$$z_s \begin{bmatrix} x_{pix} \\ y_{pix} \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$$

$$x_{pix} = \frac{u}{z_s} \qquad y_{pix} = \frac{v}{z_s}$$

Matriz intrínseca em formato homogéneo

- Em coordenadas homogéneas será:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} , \; P_{i_h} = \mathbf{K} P_S \; , \; \mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$P_{i_h} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \rightarrow \text{coordenadas na imagem em formato homogéneo}$$

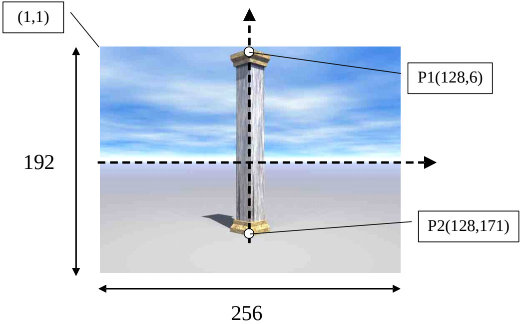
$$P_S = \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} \rightarrow \text{coordenadas métricas (mundo) em formato homogéneo}$$

- onde:

$$w = z_s, \quad x_{pix} = \frac{u}{w}, \quad y_{pix} = \frac{v}{w}$$

Exemplo de cálculo geométrico da imagem

- Numa imagem de 256 X 192 pixels
  - CCD de 150 pixels/mm (*dot-pitch*)
  - distância focal $f = 6$ mm
- Determinar a distância a que está uma coluna que está paralela ao plano de imagem (da câmara) e tem 2 metros de altura.
  - Os pontos relevantes P1 e P2 estão indicados na figura.



Passos do procedimento:

1. Obter matriz intrínseca
2. Obter coordenadas (x_{pix}, y_{pix}) dos pixels em coordenadas na câmara
3. Escrever as equações para os pontos conhecidos P1 e P2
4. Estabelecer a relação conhecida dos pontos correspondentes no ambiente real
5. Resolver as equações resultantes e obter a distância a que está a coluna

Solução: aprox. 10.9 m

Resolução do problema anterior

$$z_s \begin{bmatrix} x_{pix} \\ y_{pix} \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}, \alpha_x = \alpha_y = 6 \times 150 = 900\, pixels \; ,$$

$$x_0 = \frac{256}{2} = 128, \quad y_0 = \frac{192}{2} = 96$$

- Observar que os pixels estão indicados no enunciado num sistema diferente do apresentado, i.e. a coordenada y está a contar do topo. Assim, as coordenadas corretas dos pixels para os cálculos são:
  - $y_{pix_1} = 192 - 6 = 186$
  - $y_{pix_2} = 192 - 171 = 21$

$$\mathbf{K} = \begin{bmatrix} 900 & 0 & 128 \\ 0 & 900 & 96 \end{bmatrix}, P_1 = \begin{bmatrix} x_{pix1} \\ y_{pix1} \end{bmatrix} = \begin{bmatrix} 128 \\ 186 \end{bmatrix}, P_2 = \begin{bmatrix} x_{pix2} \\ y_{pix2} \end{bmatrix} = \begin{bmatrix} 128 \\ 21 \end{bmatrix}$$

- Obtém-se assim as seguintes equações matriciais a ser resolvidas:

$$z_s \begin{bmatrix} 128 \\ 186 \end{bmatrix} = \begin{bmatrix} 900 & 0 & 128 \\ 0 & 900 & 96 \end{bmatrix} \begin{bmatrix} x_{1s} \\ y_{1s} \\ z_s \end{bmatrix}, \quad z_s \begin{bmatrix} 128 \\ 21 \end{bmatrix} = \begin{bmatrix} 900 & 0 & 128 \\ 0 & 900 & 96 \end{bmatrix} \begin{bmatrix} x_{2s} \\ y_{2s} \\ z_s \end{bmatrix}$$

Resolução do problema anterior (cont.)

- Expandindo as matrizes e operando as equações:

$$\begin{cases} 128z_s = 900x_{1s} + 128z_s \\ 186z_s = 900y_{1s} + 96z_s \end{cases} \Leftrightarrow \begin{cases} x_{1s} = 0 \\ 900y_{1s} = 90z_s \end{cases}$$

$$\begin{cases} 128z_s = 900x_{2s} + 128z_s \\ 21z_s = 900y_{2s} + 96z_s \end{cases} \Leftrightarrow \begin{cases} x_{2s} = 0 \\ 900y_{2s} = -75z_s \end{cases}$$

$$(90 - (-75))\, z_s = 900\, (y_{1s} - y_{2s})$$

- Sabendo que a altura real da coluna são 2 m (2000 mm):

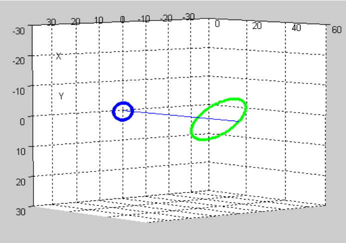
$$y_{1s} - y_{2s} = 2000$$

- Obtem-se o resultado final pedido:

$$z_s = \frac{900 \times 2000}{165} = \frac{1800000}{165} \approx 10909\, mm \approx 10.9\, m$$

Exemplo em Matlab de cálculo dos pixels por projeção

- Calcular a imagem de pontos de uma circunferência com 10 unidades de raio, centrada em (0,0,50) e inicialmente num plano paralelo ao da imagem mas depois rodada em torno do eixo x de α rad
- Elementos do sistema:
  - $f = 2$ unidades de comprimento (por ex. mm)
  - $K_x = K_y = 10$ pixels/unidade de comprimento
  - Centro da imagem em pixels $(x_0 = 0, y_0 = 0) \rightarrow$ considera-se que a numeração dos pixels é (0,0) no centro da imagem (o que, aliás, raramente é usado)



Aula4a.m

Código - 1

- Criação dos pontos no espaço em coordenadas homogéneas

```
%% define pontos de circunferencia
NN=20;
t=linspace(0,2*pi,NN);
x=10*cos(t);
y=10*sin(t);
z=zeros(1,NN);
%% torna as coordenadas homogeneas
CC=[ x; y; z; ones(1,NN)];
```

Código - 2

- Recolocação dos pontos reais noutra posição para melhor ilustrar o efeito da projeção (uma rotação e uma translação)

```
%%define a transformacao geometrica a aplicar
a=-pi/4;
trotx=[ 1 0        0       0
        0 cos(a) -sin(a) 0
        0 sin(a)  cos(a) 50
        0 0        0       1
      ];
DD=trotx*CC; %%Reposiciona a circunferencia em 3D
plot3( DD(1,:) ,DD(2,:) , DD(3,:) , 'g.')
```

Código - 3

- Estabelecimento da matriz intrínseca com parâmetros sugeridos

```
%% define a matriz intrinseca
f=2; Kx=10; x0=0; y0=0;
Ky=Kx; ax=f*Kx; ay=f*Ky;
K = [ax 0  x0 0
     0  ay y0 0
     0  0   1 0
     ];

%Plot do foco
plot3(0,0,0, '*k');
```

Código - 4

- Calcular as coordenadas por projeção e converter coordenadas homogénas em oordenadas de pixels:

```
% calcula as coordenadas na imagem em formato homogeneo

EE= K*DD;
% obtem as coordenadas na escala da imagem
% dividindo as coordenadas pelo termo homogeneo

Xpix=EE(1,:)./EE(3,:); % xpix=u/w
Ypix=EE(2,:)./EE(3,:); % ypix=v/w
% apenas para representar no Matlab a 3D
Zpix=ax*ones(size(EE(1,:)));
plot3( Xpix , Ypix , Zpix , 'b.')
% ax - distancia que coincide com o plano focal
```

Matriz intrínseca tradicional

- Na forma matricial, em coordenadas homogéneas, virá a seguinte representação:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix}$$

$$P_{i_h} = \mathbf{K} P_S$$

$$\mathbf{K} = \begin{bmatrix} \alpha_x & 0 & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$w = z_s, \quad x_{pix} = \frac{u}{w}, \quad y_{pix} = \frac{v}{w}$$

Variante da forma da matriz intrínseca

- Em certos casos pode haver um termo s de distorção (*skewing*) entre os eixos x e y.
- E, por outro lado, descartando a coluna de zeros (descartando a representação homogénea), também se pode decompor a matriz intrínseca em três componentes, com os seguintes significados:

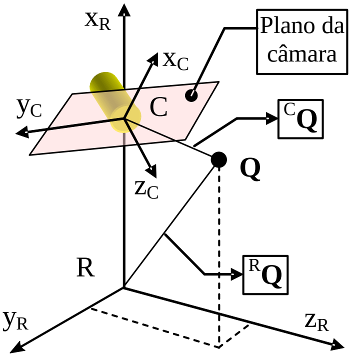
$$\mathbf{K} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Translação 2D}} \times \underbrace{\begin{bmatrix} 1 & s/\alpha_y & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Cisalhamento 2D}} \times \underbrace{\begin{bmatrix} \alpha_x & 0 & 0 \\ 0 & \alpha_y & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\text{Fator de escala 2D}}$$

Os parâmetros intrínsecos

- Os parâmetros intrínsecos da câmara são:
  - x_0, y_0
    - ponto principal (em geral são as coordenadas do centro da imagem, ou próximo)
  - α_x, α_y
    - distâncias focais em pixels ($\alpha_x = k_i f$, sendo f a distância focal em unidades métricas e k_i a densidade linear dos pixels)
  - s
    - fator de *skewing* entre eixos (zero, na maioria dos casos).

Variação de posição da câmara

- Quando os sistemas de coordenadas da câmara e do mundo real são diferentes (acontece na maioria dos casos reais), é necessária uma adequada transformação geométrica da câmara antes aplicar a matriz intrínseca:



- Simples relação de transformações...
  - $^{R}Q = {}^{R}\mathbf{T}_{C} \cdot {}^{C}Q$
  - $^{C}Q = \left({}^{R}\mathbf{T}_{C}\right)^{-1} \cdot {}^{R}Q = {}^{C}\mathbf{T}_{R} \cdot {}^{R}Q$
- ... e a transformação perspetiva:
  - $Q_{i_h} = \mathbf{K} \cdot {}^{C}Q = \mathbf{K} \cdot \left({}^{R}\mathbf{T}_{C}\right)^{-1} \cdot {}^{R}Q = \mathbf{K} \cdot {}^{C}\mathbf{T}_{R} \cdot {}^{R}Q$

A relação entre a câmara e o referencial é a Matriz Extrínseca (posição do referencial do mundo em relação ao referencial da câmara ${}^{C}\mathbf{T}_{R}$).

Parâmetros extrínsecos

Parâmetros Extrínsecos

- Na relação anterior, a transformação geométrica da "postura" da câmara, que se define à custa dos parâmetros extrínsecos, pode não ser conhecida, e só se tem acesso ao produto da matriz intrínseca com a extrínseca, e que se designa a matriz da câmara $(\mathbf{P})$:

$$Q_{i_h} = \mathbf{P} \cdot {}^{R}Q$$

$$\text{sendo } \mathbf{P} = \mathbf{K} \cdot {}^{C}\mathbf{T}_{R}$$

- Esta matriz $\mathbf{P}$ pode ser obtida por um processo experimental de calibração.

Ainda os parâmetros extrínsecos

- A matrix extrínseca $\mathbf{T}$ pode ser decomposta numa rotação seguida de uma translação na seguinte representação:

$$\left[\begin{array}{c|c} \mathbf{R} & \mathbf{t} \\ \hline \mathbf{0} & 1 \end{array}\right] = \left[\begin{array}{c|c} \mathbf{I} & \mathbf{t} \\ \hline \mathbf{0} & 1 \end{array}\right] \times \left[\begin{array}{c|c} \mathbf{R} & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array}\right] = \left[\begin{array}{ccc|c} 1 & 0 & 0 & t_1 \\ 0 & 1 & 0 & t_2 \\ 0 & 0 & 1 & t_3 \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \times \left[\begin{array}{ccc|c} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

- O vetor $\mathbf{t}$ é a posição do referencial do mundo visto da câmara, e a matriz $\mathbf{R}$ traduz as orientações do referencial do mundo visto da câmara. Isto é confirmado se se recordar que, de facto, a definição original é: ${}^{C}\mathbf{T}_{R}$

Decomposição da matriz da câmara

- O objetivo principal da calibração é saber $\mathbf{K}$ para depois a aplicar no cálculo da projeção e vice-versa. Assim, o processo de calibração tem de ser capaz de decompor a matriz $\mathbf{P}$ na matriz intrínsica $\mathbf{K}$ e na extrínseca $\mathbf{T}$.
- $\mathbf{P}$ é uma matriz de 3x4 que resulta do seguinte:

$$\mathbf{P} = \mathbf{K} \cdot \mathbf{T} = \begin{bmatrix} \alpha_x & s & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \left[\begin{array}{c|c} \mathbf{R}_{3\times 3} & \mathbf{t}_{3\times 1} \\ \hline 0_{1\times 3} & 1 \end{array}\right]$$

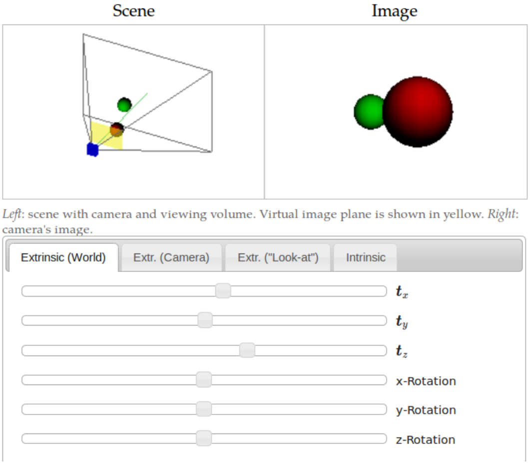
$$\mathbf{P} =$$

$$\begin{bmatrix} \alpha_x r_{11} + r_{21} s + r_{31} x_0 & \alpha_x r_{12} + r_{22} s + r_{32} x_0 & \alpha_x r_{13} + r_{23} s + r_{33} x_0 & t_x \alpha_x + t_y s + t_z x_0 \\ \alpha_y r_{21} + r_{31} y_0 & \alpha_y r_{22} + r_{32} y_0 & \alpha_y r_{23} + r_{33} y_0 & t_y \alpha_y + t_z y_0 \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}$$

- Não obstante a expressão anterior parecer complexa, $\mathbf{P}$ envolve "apenas" até 11 variáveis desconhecidas; portanto no mínimo são precisas 11 equações, mas em geral usam-se mais para se melhorar estatisticamente os resultados devido a limitações na obtenção de pontos nas imagens de calibração.

Ilustração interativa dos parâmetros

- `http://ksimek.github.io/2012/08/22/extrinsic/`



Calibração de câmara

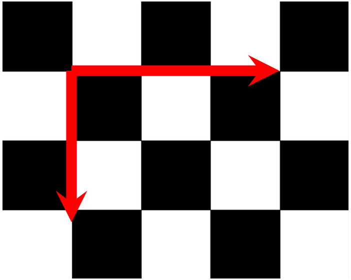
Calibração de câmara - 1

- Calibrar uma câmara é o processo de obtenção dos seguintes parâmetros:
  - Parâmetros Intrínsecos
    - Para uma dada câmara, basta fazer uma vez se a lente e o CCD não variarem.
  - Parâmetros de Distorção
    - Obtenção dos parâmetros de distorção ótica da lente (radial, tangencial, etc.). Em geral, calculados uma única vez, tal como os intrínsecos.
  - Parâmetros Extrínsecos
    - Obtenção da transformação (translação e orientações) entre a câmara e um referencial externo convencionado com o sistema de referência.

Calibração de câmara - 2

Metodologia para a calibração

- Obter imagens onde se saiba a posição real de pontos conhecidos (coordenadas reais e os pixels que lhe correspondem)
- Com essa informação, estabelecer um processo (iterativo, em geral) para determinar os 11 parâmetros.
- Recorre-se em geral a múltiplas imagens de um tabuleiro em xadrez de medidas conhecidas, o que permite detetar bem os cantos, e depois fazer os cálculos.
- Existem packages de software (em Matlab, OpenCV, etc.) que automatizam este processo de calibração de câmara.



Algumas *packages* de calibração consideram o vértice superior esquerdo do tabuleiro de xadrez como a origem do sistema de coordenadas de referência para definir os extrínsecos da câmara.