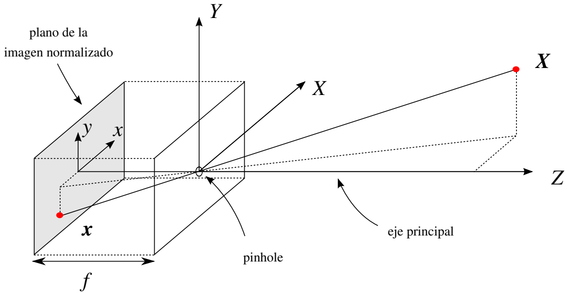
Modelo de cámara

Araguás, Gastón   Redolfi, Javier

17 de abril del 2019

Modelo de cámara - pinhole

Proyección central - Proyección perspectiva



Por triángulos semejantes

$$-\frac{x}{f} = \frac{X}{Z} \qquad -\frac{y}{f} = \frac{Y}{Z}$$

Modelo de cámara - pinhole

Proyección central - Proyección perspectiva

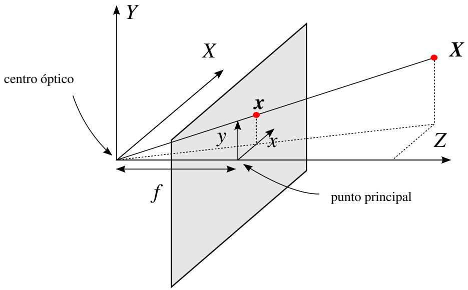


Imagen virtual en el plano $z = f$

$$\frac{x}{f} = \frac{X}{Z} \qquad \frac{y}{f} = \frac{Y}{Z}$$

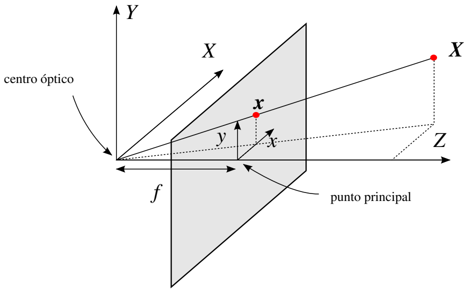
En coord. homogeneas

$$\boldsymbol{X} = [X, Y, Z, 1]^T \Rightarrow \boldsymbol{x} = [f\frac{X}{Z}, f\frac{Y}{Z}, 1]^T = [fX, fY, Z]^T$$

Mapea $\mathbb{P}^3 \to \mathbb{P}^2$.

Modelo de cámara - pinhole

Proyección central - Proyección perspectiva



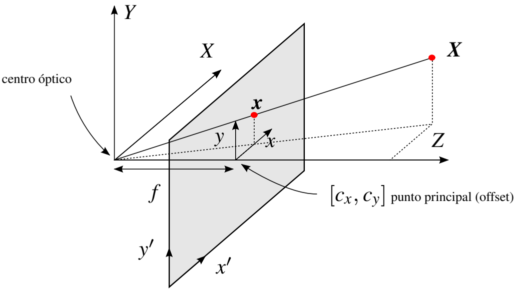
Matricialmente

$$\boldsymbol{x} = \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix} = \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = P\boldsymbol{X}$$

con P una matriz de proyección de 3×4 llamada *matriz de la cámara*

Modelo de cámara - pinhole

Proyección central - Proyección perspectiva



Offset en el *punto principal*

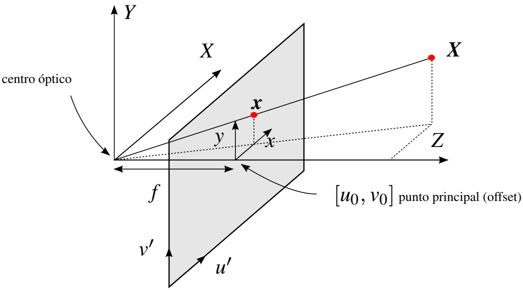
$$\boldsymbol{x}' = \left[f\frac{X}{Z} + c_x, f\frac{Y}{Z} + c_y, 1 \right]^T = \begin{bmatrix} f & & c_x & 0 \\ & f & c_y & 0 \\ & & 1 & 0 \end{bmatrix} \boldsymbol{X}$$

$$\boldsymbol{x}' = K[I|\mathbf{0}]\boldsymbol{X}$$

K es la matriz de calibración y (c_x, c_y) se llama punto principal.

Modelo de cámara - pinhole

Proyección central - Proyección perspectiva



Sistema de coordenadas en píxeles $\quad m\boldsymbol{x}' = [xm_x, ym_y, 1]^T \rightarrow \boldsymbol{u}' = [u', v', 1]^T$

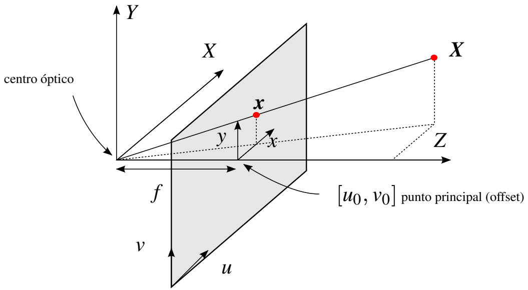
$$\boldsymbol{u}' = \left[f\frac{X}{Z}m_x + c_x m_x, f\frac{Y}{Z}m_y + c_y m_y, 1\right]^T = \begin{bmatrix} \alpha_x & 0 & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} [I|\mathbf{0}]\boldsymbol{X}$$

$$\boldsymbol{u}' = K[I|\mathbf{0}]\boldsymbol{X}$$

con $\alpha_x = fm_x$, $\quad \alpha_y = fm_y$, $\quad u_0 = c_x m_x \quad$ y $\quad v_0 = c_y m_y \quad$ en píxeles.

Modelo de cámara - pinhole

Proyección central - Proyección perspectiva



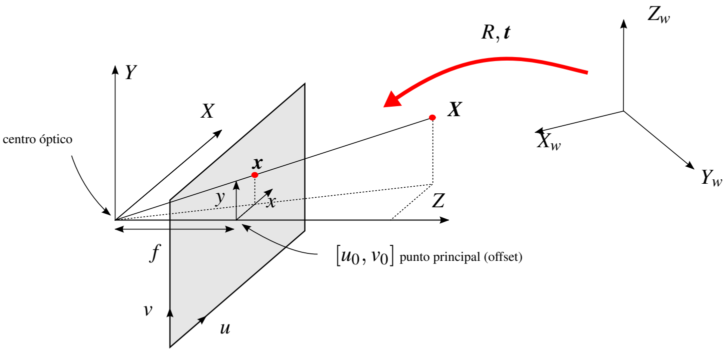
Ejes de la cámara no perpendiculares $\boldsymbol{u}' \rightarrow \boldsymbol{u} = [u' + sv', v', 1]^T$

$$\boldsymbol{u} = \left[\alpha_x \frac{X}{Z} + u_0 + s\left(\alpha_y \frac{Y}{Z} + v_0\right), \alpha_y \frac{Y}{Z} + v_0, 1\right]^T = \begin{bmatrix} \alpha_x & s & u_0 \\ 0 & \alpha_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} [I|\mathbf{0}]\boldsymbol{X}$$

$$\boldsymbol{u} = K[I|\mathbf{0}]\boldsymbol{X} = P\boldsymbol{X}$$

Modelo de cámara - pinhole

Proyección central - Proyección perspectiva



Rotación y traslación

$$\boldsymbol{X} = \begin{bmatrix} R & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \boldsymbol{X}_w$$

$$\boldsymbol{u} = K[I|\boldsymbol{0}]\boldsymbol{X} = K[R|\boldsymbol{t}]\boldsymbol{X}_w = P\boldsymbol{X}_w$$

P tiene 11 GDL, 5 de la matriz K, 3 de la rotación (3D) y 3 de la traslación.

Sean $\boldsymbol{x}$ y $\hat{\boldsymbol{x}}$ los puntos ideal y con distorsión

Distorsión radial

$$\hat{x} = x + x\left[k_1\left(x^2+y^2\right) + k_2\left(x^2+y^2\right)^2\right]$$

$$\hat{y} = y + y\left[k_1\left(x^2+y^2\right) + k_2\left(x^2+y^2\right)^2\right]$$

donde k_1 y k_2 son los coeficientes de distorsión radial; suponiendo para simplificar $s = 0$ y operando

$$\hat{u} = u + (u-u_0)\left[k_1\left(x^2+y^2\right) + k_2\left(x^2+y^2\right)^2\right]$$

$$\hat{v} = v + (v-v_0)\left[k_1\left(x^2+y^2\right) + k_2\left(x^2+y^2\right)^2\right]$$

Modelos más complejos contemplan k_3, k_4 y k_5, y p_1, p_2 (coeficientes de distorsión tangencial).

Homografía entre plano e imagen

Sea $\boldsymbol{X}^w$ un punto 3D sobre el plano $Z = 0$ en el sistema de coordenadas del mundo, entonces $\boldsymbol{X}^w = [X, Y, 0, 1]^T$

$$\boldsymbol{x} = K[\boldsymbol{r}_1\, \boldsymbol{r}_2\, \boldsymbol{r}_3\, \boldsymbol{t}]\boldsymbol{X}^w = K[\boldsymbol{r}_1\, \boldsymbol{r}_2\, \boldsymbol{t}]\boldsymbol{x}^w = H\boldsymbol{x}^w$$

donde $\boldsymbol{x}^w = [X, Y, 1]^T$, por lo tanto

$$H = K[\boldsymbol{r}_1\, \boldsymbol{r}_2\, \boldsymbol{t}]$$

es una homografía (matríz de 3×3)

[1] Z. Zhang. A flexible new tehcnique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 2000.

Algoritmo

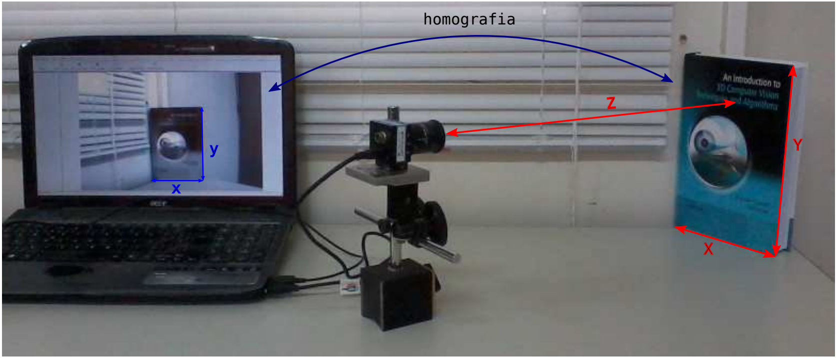
- Estimar H usando Direct Linear Transform (DLT)
- Usando varias imágenes se trata de minimizar:

$$\sum_{i}^{N}\sum_{j}^{P} d(\boldsymbol{x}_{ij}, \hat{\boldsymbol{x}}(K, k_1, k_2, R_i, \boldsymbol{t}_i, \boldsymbol{X}_j^w))$$

  donde $\hat{\boldsymbol{x}}(K, k_1, k_2, R_i, \boldsymbol{t}_i, \boldsymbol{X}_j^w)$ es la proyección del punto 3D $\boldsymbol{X}_j^w$, en la $i-$ésima imagen

Cada imagen aporta 8 GDL, de los cuales 3 son de la rotación y 3 de la traslación, los dos restantes ajustan los parámetros intrínsecos.

Asumiendo $s = 0$ y sin considerar la distorsión de la lente…



$$\alpha_x = \frac{x}{X} Z \qquad \alpha_y = \frac{y}{Y} Z \qquad u_0 = w/2 \qquad v_0 = h/2$$

Calibración simple

$$K = \begin{bmatrix} 600 & 0 & 372 \\ 0 & 600 & 240 \\ 0 & 0 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Calibración completa (8 imágenes)

$$K = \begin{bmatrix} 611{,}47746477 & 0 & 432{,}3113038 \\ 0. & 613{,}15654266 & 208{,}46520386 \\ 0 & 0 & 1 \end{bmatrix}$$

$$D = \begin{bmatrix} -2{,}57e-01 & 1{,}44e-01 & -1{,}70e-02 & 4{,}34e-03 & 3{,}24e-04 \end{bmatrix}$$