Computer Vision
Perspective 3D

Ilan Guenet

November 12, 2021

Contents

1 Géométrie optique

Il existe deux matrices à calculer : * La matrice intrinsèque permet de passer du monde 3D caméra au monde des coordonnées 2D * La matrice extrinsèques permet de passer du monde 3D au monde 3D caméra.

1.1 Matrice intrinsèque (coordonnées 3D caméra à coordonnées 2D plan)

perspective projection to the image plane Supposons que la caméra est modélisée comme ci-dessous :

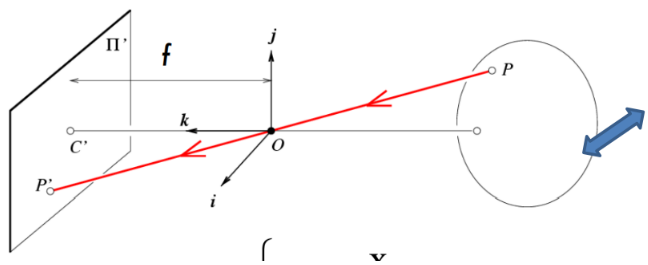

Figure 1: Projection des points 3D sur l'image 2D

- f: longueur focale
- o: centre de la caméra

On a $P = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ dans le monde caméra. Le but est de calculer $P' = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ dans le monde des coordonnées 2D (plan) tel que $P' = KP$ avec K la matrice de transformation appelée matrice intrinsèque.

1.1.1 Démonstration

Démontrons le calcul de la coordonnée x: On isole les coordonnées sur l'axe $\vec{i}$ et $\vec{k}$

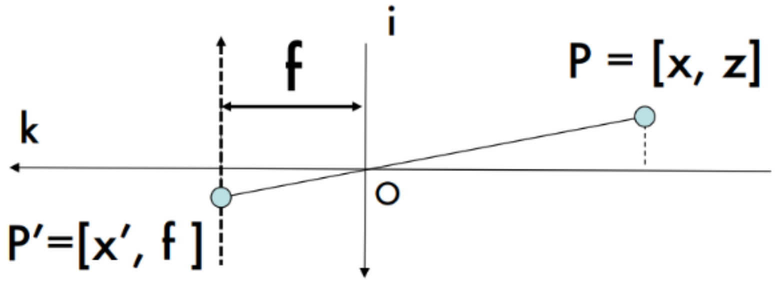

Figure 2: Visualisation sur l'axe x et k

La notation n'est pas la même sur le schéma.

En réalité, f est différent sur chaque axe Ainsi, à un signe négatif près $\frac{x}{f_x} = \frac{X}{Z} \Rightarrow x = f_x \frac{X}{Z}$

En généralisant :

$$\begin{cases} x = f_x \frac{X}{Z} \\ y = f_y \frac{Y}{Z} \end{cases}$$

Il faut maintenant passer du monde réel continue en mètres vers le monde discret de l'image en pixel, c'est-à-dire $P' = KP$ avec $P' = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$

$$\begin{cases} u = \frac{x}{d_x} \frac{X}{Z} = \frac{f_x}{d_x} \frac{X}{Z} = \alpha_x \frac{X}{Z} \\ v = \frac{y}{d_y} \frac{Y}{Z} = \frac{f_y}{d_y} \frac{Y}{Z} = \alpha_y \frac{Y}{Z} \end{cases}$$

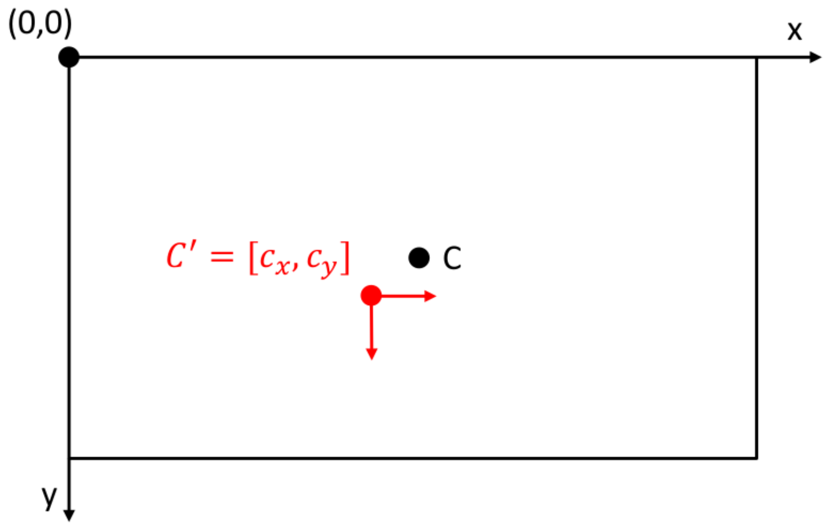

Figure 3: Différence centre optique et centre de l'image.

Enfin, il y a un décalage entre le centre optique et le centre de l'image. On applique une translation permettant de passer de centre optique au centre (0, 0) de l'image.

$$\begin{cases} u = \alpha_x \frac{X}{Z} + c_x \\ v = \alpha_y \frac{Y}{Z} + c_y \end{cases}$$

Matriciellement,

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{\begin{bmatrix} \alpha_x & 0 & c_x \\ 0 & \alpha_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}}{Z} \Rightarrow Z \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & 0 & c_x \\ 0 & \alpha_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \Rightarrow P' \approx KP$$

avec $K = \begin{bmatrix} \alpha_x & 0 & c_x \\ 0 & \alpha_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$ la matrice intrinsèque

1.2 Matrice extrinsèque (coordonnées 3D monde/objet à coordonnées 3D caméra)

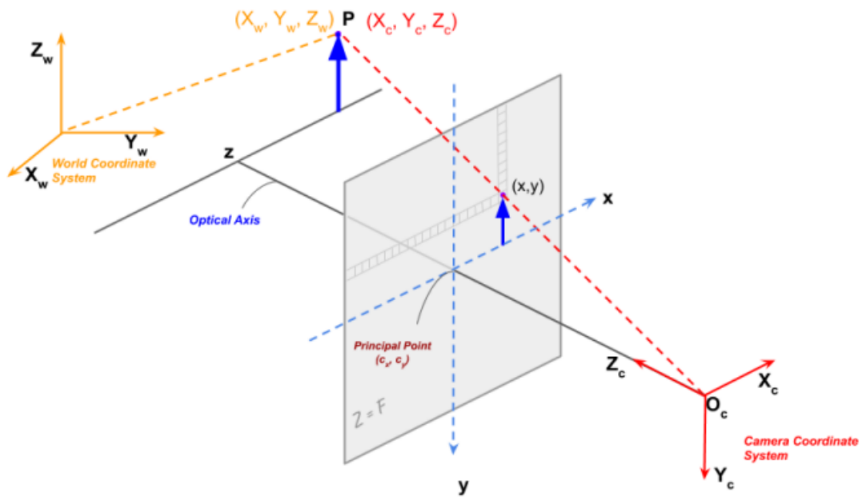

Figure 4: Les différents repères

Les coordonnées du monde ne sont pas nécessairement les mêmes que les coordonnées de la caméra. **Il faut convertir du système de coordonnées du monde vers le système de coordonnées de la caméra**

Notons $P_w = \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}$ les coordonnées 3D du système monde et $P_c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$ les coordonnées 3D du système caméra.

Pour convertir les coordonnées du système monde vers le système caméra, on a besoin d'une rotation et translation tel que

$$P_c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R. \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + t \text{ avec } \begin{cases} R: \text{ une matrice de rotation (3, 3)} \\ t: \text{ un vecteur de translation} \end{cases}$$

On peut réécrire avec une notation matricielle $\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \left[\begin{array}{c|c} R & t \end{array}\right] \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}$

On peut réécrire en coordonnées homogènes : $\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \left[\begin{array}{c|c} R & t \\ 0 & 1 \end{array}\right] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$

1.3 Matrice caméra

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \approx K. \left[\begin{array}{c|c} R & t \end{array} \right] . \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = C. \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

avec C matrice caméra de dimension (3,4)

Regardons avec les dimensions :

$$\underbrace{\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}}_{(3,1)} \approx \underbrace{\underbrace{\underbrace{K}_{(3,3)} . \underbrace{\left[\begin{array}{c|c} R & t \end{array} \right]}_{(3,4)}}_{\text{P } (3,4)} . \underbrace{\begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}}_{(4,\ 1)}}_{(3,1)}$$

L'égalité absolue est $\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = C. \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$

1.4 Estimation de la matrice caméra

On veut estimer la matrice C à partir de paires (P', P) connue.

On sait que : $\lambda P' = C.P$ avec $C = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m34 \end{bmatrix}$

donc $\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \approx \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} . \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$

Ainsi $\begin{bmatrix} \lambda.u \\ \lambda.v \\ \lambda \end{bmatrix} \approx \begin{bmatrix} m_{11}.X_w + m_{12}.Y_w + m_{13}.Z_w + m_{14} \\ m_{21}.X_w + m_{22}.Y_w + m_{23}.Z_w + m_{24} \\ m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34} \end{bmatrix}$

On divise tout par $\lambda = m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34}$ $\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{m_{11}.X_w + m_{12}.Y_w + m_{13}.Z_w + m_{14}}{m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34}} \\ \frac{m_{21}.X_w + m_{22}.Y_w + m_{23}.Z_w + m_{24}}{m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34}} \\ 1 \end{bmatrix}$

D'où $\begin{bmatrix} u.(m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34}) \\ v.(m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34}) \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}.X_w + m_{12}.Y_w + m_{13}.Z_w + m_{14} \\ m_{21}.X_w + m_{22}.Y_w + m_{23}.Z_w + m_{24} \\ 1 \end{bmatrix}$

On trouve deux équations :

$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} u.(m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34}) - m_{11}.X_w + m_{12}.Y_w + m_{13}.Z_w + m_{14} \\ v.(m_{31}.X_w + m_{32}.Y_w + m_{33}.Z_w + m_{34}) - m_{21}.X_w + m_{22}.Y_w + m_{23}.Z_w + m_{24} \end{bmatrix}$

Matriciellement,

$$\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix} u.\begin{bmatrix}m_{31} & m_{32} & m_{33} & m_{34}\end{bmatrix}\begin{bmatrix}X_w\\Y_w\\Z_w\\1\end{bmatrix} - \begin{bmatrix}m_{11} & m_{12} & m_{13} & m_{14}\end{bmatrix}\begin{bmatrix}X_w\\Y_w\\Z_w\\1\end{bmatrix} \\ v.\begin{bmatrix}m_{31} & m_{32} & m_{33} & m_{34}\end{bmatrix}\begin{bmatrix}X_w\\Y_w\\Z_w\\1\end{bmatrix} - \begin{bmatrix}m_{21} & m_{22} & m_{23} & m_{24}\end{bmatrix}\begin{bmatrix}X_w\\Y_w\\Z_w\\1\end{bmatrix}\end{bmatrix}$$

Finalement,

$$0 = \begin{bmatrix} -\begin{bmatrix}X_w & Y_w & Z_w & 1\end{bmatrix} & 0 & u.\begin{bmatrix}X_w & Y_w & Z_w & 1\end{bmatrix} \\ 0 & -\begin{bmatrix}X_w & Y_w & Z_w & 1\end{bmatrix} & v.\begin{bmatrix}X_w & Y_w & Z_w & 1\end{bmatrix}\end{bmatrix} \begin{bmatrix}\begin{bmatrix}m_{11} & m_{12} & m_{13} & m_{14}\end{bmatrix}^T \\ \begin{bmatrix}m_{21} & m_{22} & m_{23} & m_{24}\end{bmatrix}^T \\ \begin{bmatrix}m_{31} & m_{32} & m_{33} & m_{34}\end{bmatrix}^T\end{bmatrix}$$

$$0 = \begin{bmatrix} -X_w & -Y_w & -Z_w & -1 & 0 & 0 & 0 & 0 & u.X_w & u.Y_w \\ u.Z_w & u \\ 0 & 0 & 0 & 0 & -X_w & -Y_w & -Z_w & -1 & v.X_w & v.Y_w \\ v.Z_w & v \end{bmatrix} \begin{bmatrix}m_{11}\\m_{12}\\m_{13}\\m_{14}\\m_{21}\\m_{22}\\m_{23}\\m_{24}\\m_{31}\\m_{32}\\m_{33}\\m_{34}\end{bmatrix}$$

Dans le cas de N pairs (P', P)

$$0 = \underbrace{\begin{bmatrix} -X_1 & -Y_1 & -Z_1 & -1 & 0 & 0 & 0 & 0 & u_1.X_1 & u_1.Y_1 \\ u_1.Z_1 & u_1 \\ 0 & 0 & 0 & 0 & -X_1 & -Y_1 & -Z_1 & -1 & v_1.X_1 & v_1.Y_1 \\ v_1.Z_1 & v_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots \\ -X_2 & -Y_2 & -Z_2 & -2 & 0 & 0 & 0 & 0 & u_2.X_2 & u_2.Y_2 \\ u_2.Z_2 & u_2 \\ 0 & 0 & 0 & 0 & -X_2 & -Y_2 & -Z_2 & -2 & v_2.X_2 & v_2.Y_2 \\ v.Z_2 & v_2 \end{bmatrix}}_{\text{(2N, 12)}} \underbrace{\begin{bmatrix}m_{11}\\m_{12}\\m_{13}\\m_{14}\\m_{21}\\m_{22}\\m_{23}\\m_{24}\\m_{31}\\m_{32}\\m_{33}\\m_{34}\end{bmatrix}}_{\text{(12,1)}}$$

Notons $0 = A.s$ avec $s = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{21} & m_{22} & m_{23} & m_{24} & m_{31} & m_{32} \\ m_{33} & m_{34} \end{bmatrix}^T$

Pour résoudre l'estimation de la matrice caméra, on va utiliser un problème d'optimisation tel que :

$\hat{s} = \underset{s||s||^2=1}{\arg\min} ||A.s||^2$

2 méthodes de résolution : * SVD : $A = U\Sigma V^T$. La solution s est la colonne

de V correspondante à la plus petite valeur singulière * Vecteurs/valeurs propres : $A^T A$. La solution s est le vecteur propre correspondant à la plus petite valeur propre

Décomposition de la matrice caméra.

$$C = \left[\begin{array}{ccc|c} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{array}\right] = K \left[\begin{array}{c|c} R & t \end{array}\right] = K \left[\begin{array}{c|c} R & -Rc \end{array}\right]$$

$C = \left[\begin{array}{c|c} M & -Mc \end{array}\right]$ avec $M = KR$

Cherchons les inconnues :

- Trouver le centre de la caméra c. On pose $Pc = 0$ et on résout par SVD. c est le vecteur propre correspondant à la plus petite valeur propre.
- Trouver la matrice intrinsèque K et ;amatrice de rotation R. $M = \underbrace{K}_{\text{triangulaire right upper}} . \underbrace{R}_{\text{orthogonal}}$. On peut résoudre par décomposition RQ

1.5 Distorsion

Jusqu'ici nous avons supposé qu'il n'y avait de pas de distorsion. Définisson ce qu'est la distorsion.

Types de distorsion

- Distorsion radiale
  - Barrel distortion
  - Pincushion distortion
  - Mustache distortion

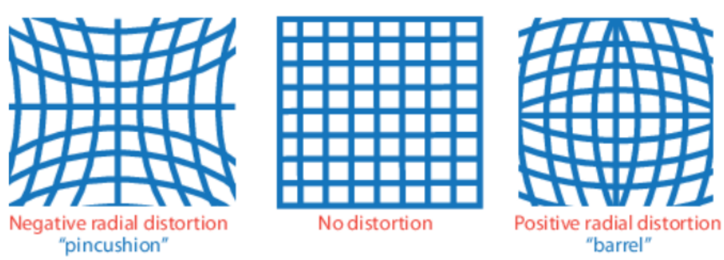

Figure 5: Types de distorsion radiale

- Distorsion tangentielle

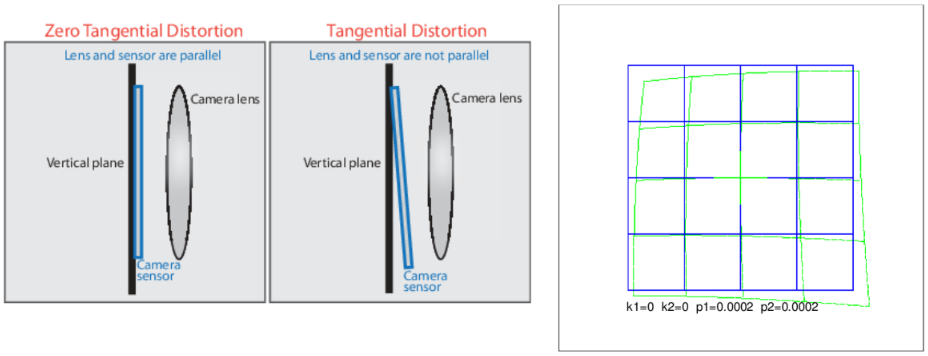

Figure 6: Distorsion tangentielle

La distorsion est plus importante pour les rayons qui passent près du bord de la lentille

Le modèle complet de la caméra avec distorsion utilise des coefficients de distorsions pour justement prendre en compte la distorsion. Un modèle de distorsion utilisé est celui de *Brown-Conrady* tel que

$$P' = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha_x . u_{distorted} + c_x \\ \alpha_y . v_{distorted} + c_y \\ 1 \end{bmatrix} \approx K . \left[\begin{array}{c|c} R & t \end{array} \right] . P_w$$

avec $\begin{cases} u_{distorted} = u . \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + 2 p_1 u v + p_2 (r^2 + 2u^2) \\ v_{distorted} = v . \frac{1 + k_1 r^2 + k_2 r^4 + k_3 r^6}{1 + k_4 r^2 + k_5 r^4 + k_6 r^6} + p_1 (r^2 + 2v^2) + 2 p_2 u v \end{cases}$

Attention, les équations ne sont plus linéaires.