

Projet imagerie 3D : Reconstruction d'une scène à partir d'images 2D



telecom
saint-étienne
école d'ingénieurs
nouvelles technologies

DUTEYRAT ANTOINE, SÈVE LÉO

27 mai 2025

Table des matières

1	Objectif	2
2	Calibration de la caméra	3
2.1	Obtention de la matrice intrinsèque K	3
2.2	Lien matrices K et E, utilisation et projection	4
2.3	Application de la méthode Zhang pour estimer K et E	6
2.3.1	Calcul de la matrice d'homographie H	6
2.3.2	Calcul de la matrice intrinsèque K	6
2.3.3	Calcul de la matrice extrinsèque E	6
2.3.4	Optimisation par minimisation d'une fonction de perte	6
3	Où trouver notre travail ?	7

1 Objectif

L'objectif de ce projet est de reconstruire une scène 3D à partir d'images 2D prises d'une caméra arbitraire. Pour cela, plusieurs étapes sont nécessaires :

- Calibrer la caméra (calibration intrinsèque) par la méthode de Zhang (mire plane).
- Prendre plusieurs images d'une scène 3D avec la caméra.
- Triangulation des points 3D à partir des images 2D.

2 Calibration de la caméra

2.1 Obtention de la matrice intrinsèque K

La calibration de la caméra est une étape cruciale pour obtenir des images 3D précises. Elle permet de déterminer la matrice intrinsèque K, qui est utilisée pour projeter les points 3D du monde réel (repère caméra) sur l'image 2D capturée par la caméra.

De cette illustration, on tire les égalités suivantes :

$$\frac{y'}{y} = \frac{f}{z} \quad \text{et} \quad \frac{x'}{x} = \frac{f}{z} \quad (1)$$

où :

- x' et y' sont les coordonnées de l'image (en mm),
- x , y et z sont les coordonnées du repère caméra (en mm),
- f est la distance focale de la caméra (en mm).

D'où :

$$y' = \frac{f \times y}{z} \quad \text{et} \quad x' = \frac{f \times x}{z} \quad (2)$$

L'objectif est maintenant de passer des coordonnées en mm du repère $x'; y'$ au repère $v; u$ de l'image. Pour cela, on introduit m_u et m_v les facteurs d'échelle (en pixels/mm), et on ajoute le décalage v_0 et u_0 (en pixels) pour obtenir les coordonnées en pixels dans l'image.

$$v = m_v \times y' + v_0 = \frac{m_v \times f \times y}{z} + v_0 \quad \text{et} \quad u = m_u \times x' + u_0 = \frac{m_u \times f \times x}{z} + u_0 \quad (3)$$

Et en prenant $z = s$:

$$v = \frac{m_v \times f \times y}{s} + v_0 \quad \text{et} \quad u = \frac{m_u \times f \times x}{s} + u_0 \quad (4)$$

En multipliant par s et en réorganisant, on obtient :

$$\begin{pmatrix} s \times u \\ s \times v \\ s \end{pmatrix} = \begin{pmatrix} f \times m_u & 0 & u_0 \\ 0 & f \times m_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (5)$$

En notant :

Projet imagerie 3D : Reconstruction d'une scène à partir d'images 2D

- $f \times m_u = \alpha_u$: distance focale en pixels (en pixels),
- $f \times m_v = \alpha_v$: distance focale en pixels (en pixels),
- $s = z$: coordonnée de la caméra (en mm)

On obtient alors la matrice intrinsèque K de la caméra :

$$\begin{pmatrix} sv \\ su \\ s \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (6)$$

Avec la matrice K au centre de l'opération (nous ne considérons pas le cisaillement γ). Cette matrice est utilisée pour projeter les points 3D du repère caméra sur l'image 2D capturée par la caméra.

2.2 Lien matrices K et E, utilisation et projection

L'objectif de la calibration est de déterminer les valeurs u_0 et v_0 de la matrice intrinsèque K de la caméra de la forme suivante [1][2] (équation 7) :

$$\begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7)$$

où :

- α_u et α_v représentent la distance focale f en pixels dans les directions verticale et horizontale respectivement,
- u_0 et v_0 sont les coordonnées du centre optique en pixels,
- La dernière ligne est utilisée pour homogénéiser les coordonnées.

Cela implique d'avoir les informations suivantes sur la caméra :

- f : distance focale de la caméra,
- $w \times h$: dimensions du capteur de la caméra (en pixels),
- $W \times H$: dimensions du capteur de la caméra (en m) ou les dimensions d'un pixel (en m)

Projet imagerie 3D : Reconstruction d'une scène à partir d'images 2D

Cette matrice est ensuite utilisée dans la conversion des coordonnées 3D du monde réel en coordonnées 2D de l'image, en modèle sténopé sans défaut d'orthogonalité (équation 8).

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \quad (8)$$

avec la matrice d'extrinsèque E de la caméra de la forme suivante (équation 9) :

$$E = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix} \quad (9)$$

Et on note H la matrice d'homographie de la forme suivante (équation 10) :

$$H = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{pmatrix} \quad (10)$$

Pour estimer les paramètres de la matrice d'homographie, nous utiliserons la méthode de Zhang. La scène qui nous permettra de calibrer la caméra comportera une mire damier (figure 1).

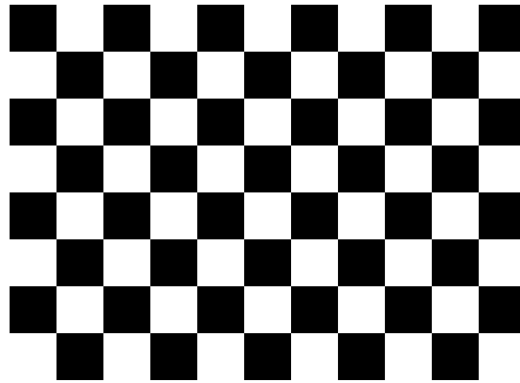


FIGURE 1 – Exemple d'une mire damier

2.3 Application de la méthode Zhang pour estimer K et E

2.3.1 Calcul de la matrice d'homographie H

2.3.2 Calcul de la matrice intrinsèque K

2.3.3 Calcul de la matrice extrinsèque E

2.3.4 Optimisation par minimisation d'une fonction de perte

3 Où trouver notre travail ?

Tout le travail dont il est question dans ce rapport est disponible sur github.

Bibliography

- [1] FrenchiDrone. Calibration de caméra. <https://www.frenchidrone.com/calibration-de-camera/>. Consulté le 20 mai 2025.
- [2] Wikipédia. Calibration de caméra. https://fr.wikipedia.org/wiki/Calibration_de_cam%C3%A9ra. Consulté le 20 mai 2025.