

UE Vision, Réalité Augmentée et Applications

Matière : Vision

Partie I (Cours)

Sylvie CHAMBON
schambon@toulouse-inp.fr

24 septembre 2025

Plan

Avant propos

Introduction

Modèle géométrique d'une caméra

Présentation du modèle

Transformation repère scène \rightarrow repère caméra

Transformation repère caméra \rightarrow repère image

Calibrage d'une caméra

Capteur stéréoscopique

Plan

Avant propos

Introduction

Modèle géométrique d'une caméra

Présentation du modèle

Transformation repère scène \rightarrow repère caméra

Transformation repère caméra \rightarrow repère image

Calibrage d'une caméra

Capteur stéréoscopique

Organisation de ce cours

- Pensez à lire/relire les objectifs
 1. Apprendre les notions géométriques fondamentales pour aborder un système de vision
 2. Découvrir et apprendre les différents éléments liés à la détection de primitives d'intérêt
 3. Découvrir et apprendre différentes techniques de suivi/mise en correspondance
 4. Mettre en œuvre une chaîne de traitement permettant de répondre à un besoin spécifique/une application particulière (APP)
- Des indications sur les temps de lecture
- Aujourd'hui on doit aborder environ 30% du cours
- Pour la classe inversée :

Les questions à poser concernent la partie à lire (+ les précédentes)

Plan

Avant propos

Introduction

Modèle géométrique d'une caméra

Présentation du modèle

Transformation repère scène \rightarrow repère caméra

Transformation repère caméra \rightarrow repère image

Calibrage d'une caméra

Capteur stéréoscopique

Plan

Avant propos

Introduction

Modèle géométrique d'une caméra

Présentation du modèle

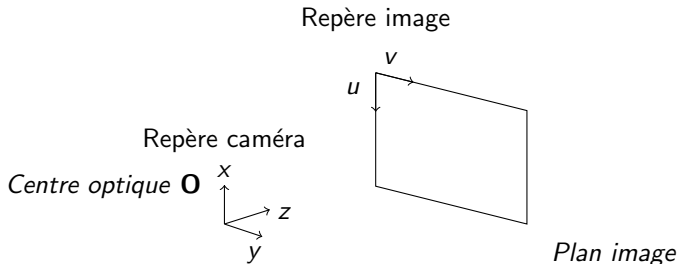
Transformation repère scène \rightarrow repère caméra

Transformation repère caméra \rightarrow repère image

Calibrage d'une caméra

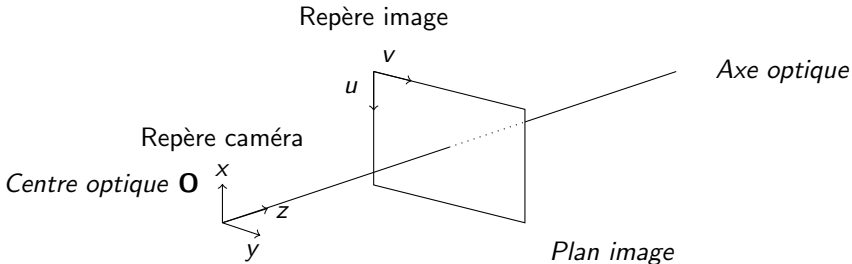
Capteur stéréoscopique

Modèle sténopé



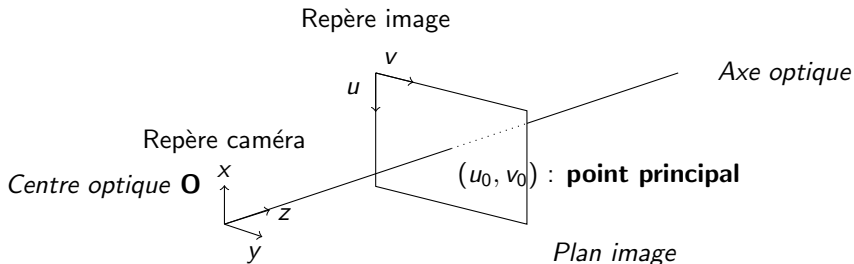
O, centre optique de la caméra, avec son repère associé et le plan image
 Plan image = plan de projection

Modèle sténopé



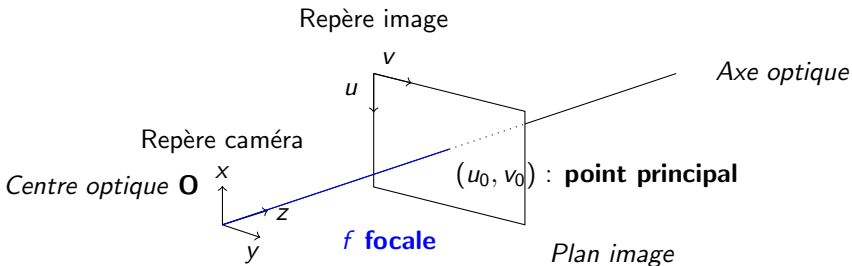
Droite passant par \mathbf{O} et perpendiculaire au plan image = *axe optique*

Modèle sténopé



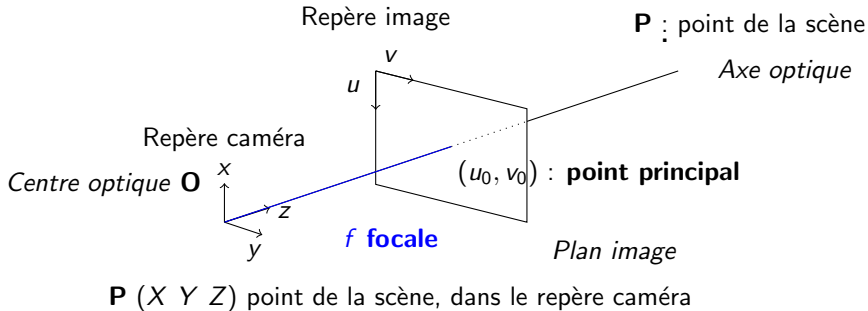
Intersection *axe optique* et plan image, notée $(u_0, v_0) =$ **point principal**

Modèle sténopé

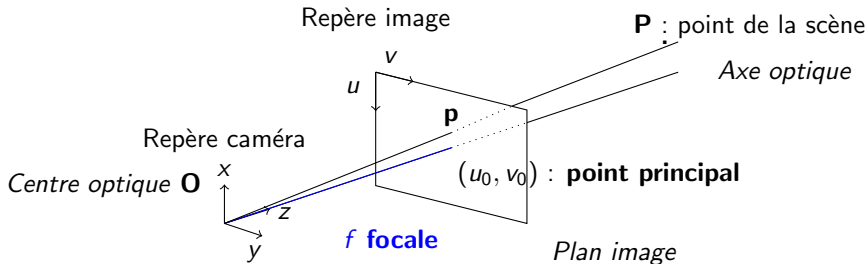


Centre optique \mathbf{O} , à une distance f du plan image = **focale**

Modèle sténopé

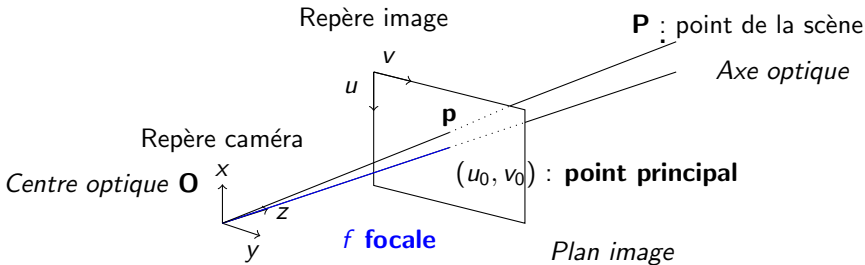


Modèle sténopé



Projection sur le plan image en \mathbf{p} (u v), dans le repère image =
Intersection de (\mathbf{OP}) avec le plan image

Modèle sténopé



\mathbf{M} = matrice de projection perspective des paramètres du modèle

$$\lambda \begin{pmatrix} u & v & 1 \end{pmatrix}^T = \mathbf{M} \begin{pmatrix} X & Y & Z & 1 \end{pmatrix}^T, \lambda \in \mathbb{R}^*$$

Calibrage d'une caméra = estimer la matrice \mathbf{M}

Notions de coordonnées homogènes (1)

- **Notations**

- $\mathbf{p} (x \ y)$ dans \mathbb{R}^2
- Si \mathbb{R}^2 , un espace vectoriel alors $(x \ y)$: un vecteur
- un point dans le plan représenté par $(x \ y)^T$

- **Représentation homogène d'une droite**

- $ax + by + c = 0$: $(a \ b \ c)^T$
- $ax + by + c = 0$ et $(ka)x + (kb)y + (kc) = 0$: la même droite, pour tout $k \neq 0$

Ainsi $(a \ b \ c)^T$ et $k(a \ b \ c)^T$: strictement la même droite donc vecteurs strictement équivalents

et classe d'équivalence de vecteurs = vecteur homogène

Conclusion : Ensemble des vecteurs de cette classe d'équivalence dans $\mathbb{R}^3 - (0 \ 0 \ 0)^T =$ espace projectif

Notions de coordonnées homogènes (2)

- **Représentation homogène d'un point**

- $\mathbf{p} = (x \ y)^T$ appartient à une droite $\mathbf{l} = (a \ b \ c)^T$ si et seulement si $ax + by + c = 0$

Ainsi $(x \ y \ 1)(a \ b \ c)^T = (x \ y \ 1)\mathbf{l} = 0$

donc \mathbf{p} peut être représenté par un vecteur de taille 3 dont la dernière coordonnée est 1

- $\mathbf{p} = (x \ y \ 1)^T =$ **coordonnées homogènes** de \mathbf{p}

De plus $\forall k \neq 0, (kx \ ky \ k)\mathbf{l} = 0$ si et seulement si $(x \ y \ 1)\mathbf{l} = 0$

Ainsi $\{(kx \ ky \ k)^T\} =$ représentation de \mathbf{p} dans \mathbb{R}^2

- Les points sont représentés par des vecteurs homogènes

Étapes pour estimer M

1. Repère scène \rightarrow repère caméra
= Matrice des paramètres extrinsèques
2. Repère caméra \rightarrow repère image
= Projection perspective
+ Matrice des paramètres intrinsèques

Repère scène → repère caméra

Soit $(x \ y \ z)$ les coordonnées de \mathbf{P} , dans le repère caméra

$$(x \ y \ z)^T = \mathbf{R}(X \ Y \ Z)^T + \mathbf{t} \quad (1)$$

Nous notons :

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \quad \text{et} \quad \mathbf{t} = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}$$

Matrice des paramètres extrinsèques

$$(x \ y \ z \ 1)^T = \mathbf{A}(X \ Y \ Z \ 1)^T \quad (2)$$

avec :

$$\mathbf{A} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

- **A** = matrice des paramètres extrinsèques
- 12 paramètres à estimer
- **R** = 9 éléments mais sachant que $\mathbf{R} \mathbf{R}^T = \mathbf{1} \rightarrow$ 3 éléments indépendants : 3 angles polaires
- **t** = 3 éléments

Donc 6 paramètres indépendants

= 6 correspondances de points

- En pratique : une centaine de points, voire plusieurs centaines !

Repère caméra → repère image

- Passage de (x, y, z) à (u, v)
- Deux transformations
 1. Projection perspective
 2. Transformation avec changement d'unités**= matrice des paramètres intrinsèques**

Projection perspective

- $(x_p \ y_p \ z_p)$ les coordonnées du point \mathbf{p} , du plan image, dans le repère caméra
- Avec f la focale + le théorème de Thalès

$$\begin{cases} x_p &= f \frac{X}{z} \\ y_p &= f \frac{y}{z} \\ z_p &= f \end{cases} \quad (4)$$

Matrice des paramètres extrinsèques

- Pour obtenir (u, v) **en pixels** (avec (u_0, v_0) point principal) :

$$\begin{cases} u &= -x_p + u_0 \\ v &= y_p + v_0 \end{cases} \quad (5)$$

- D'où mise à l'échelle avec k_u , facteur vertical, et k_v , facteur horizontal (*pixels/mm*) :

$$\begin{cases} u &= -k_u x_p + u_0 \\ v &= k_v y_p + v_0 \end{cases} \quad (6)$$

- En reprenant la projection perspective :

$$\begin{cases} u &= -k_u f \frac{x}{z} + u_0 \\ v &= k_v f \frac{y}{z} + v_0 \end{cases} \quad (7)$$

Matrice des paramètres extrinsèques

- Sous forme matricielle, en coordonnées homogènes :

$$\begin{pmatrix} \lambda u \\ \lambda v \\ \lambda \end{pmatrix} = \begin{pmatrix} -k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (8)$$

- Coordonnées cartésiennes (u, v) obtenues en divisant le résultat obtenu par λ (la troisième coordonnée) :

$$\lambda(u \ v \ 1)^T = \mathbf{K} (x \ y \ z)^T \quad \text{avec} \quad (9)$$
$$\mathbf{K} = \begin{pmatrix} -k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

- **K = matrice des paramètres intrinsèques**
- 4 paramètres : $\alpha_u = -k_u f$, $\alpha_v = -k_v f$, u_0 et v_0

Expression de la matrice de projection perspective **M**

$$\begin{cases} u &= -k_u f \frac{r_{11}X + r_{12}Y + r_{13}Z + t_x}{r_{31}X + r_{32}Y + r_{33}Z + t_z} + u_0 \\ v &= k_v f \frac{r_{21}X + r_{22}Y + r_{23}Z + t_y}{r_{31}X + r_{32}Y + r_{33}Z + t_z} + v_0 \end{cases} \quad (10)$$

Expression de la matrice de projection perspective \mathbf{M}

- Réduction au même dénominateur commun :

$$\begin{cases} u &= \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \\ v &= \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \end{cases} \quad (11)$$

- On retrouve bien :

$$\lambda (u \quad v \quad 1)^T = \mathbf{M} (X \quad Y \quad Z \quad 1)^T, \quad \lambda \in \mathbb{R}^*. \quad (12)$$

- Définition à un facteur multiplicatif près
- Si \mathbf{M} connue = possibilité d'estimer les paramètres intrinsèques \mathbf{K} et extrinsèques \mathbf{A}
- Inversement :

$$\mathbf{M} = (\mathbf{K} \mathbf{0}_{3 \times 1}) \mathbf{A}. \quad (13)$$

Expression des paramètres à partir de **M**

$$\mathbf{A} = \begin{pmatrix} \mathbf{r}_1 & t_x \\ \mathbf{r}_2 & t_y \\ \mathbf{r}_3 & t_z \\ \mathbf{0} & 1 \end{pmatrix} \text{ avec } \mathbf{r}_i = (r_{i1} \ r_{i2} \ r_{i3}). \quad (14)$$

$$\mathbf{M} = \begin{pmatrix} \mathbf{m}_1 & m_{14} \\ \mathbf{m}_2 & m_{24} \\ \mathbf{m}_3 & m_{34} \end{pmatrix} \text{ avec } \mathbf{m}_i = (m_{i1} \ m_{i2} \ m_{i3}). \quad (15)$$

Expression des paramètres à partir de \mathbf{M}

- En utilisant $\mathbf{M} = (\mathbf{K} \mathbf{0}_{3 \times 1}) \mathbf{A}$
- les propriétés d'orthogonalité de la matrice de rotation
- $\alpha_u < 0$

alors

$$\begin{aligned}
 \mathbf{r}_3 &= \mathbf{m}_3 \\
 u_0 &= \mathbf{m}_1 \cdot \mathbf{m}_3 \\
 v_0 &= \mathbf{m}_2 \cdot \mathbf{m}_3 \\
 \alpha_u &= -\|\mathbf{m}_1 \wedge \mathbf{m}_3\| \\
 \alpha_v &= \|\mathbf{m}_2 \wedge \mathbf{m}_3\| \\
 \mathbf{r}_1 &= \frac{\mathbf{m}_1 - u_0 \mathbf{m}_3}{\alpha_u} \\
 \mathbf{r}_2 &= \frac{\mathbf{m}_2 - v_0 \mathbf{m}_3}{\alpha_v} \\
 t_x &= \frac{m_{14} - u_0 m_{34}}{\alpha_u} \\
 t_y &= \frac{m_{14} - v_0 m_{34}}{\alpha_v} \\
 t_z &= m_{34}
 \end{aligned} \tag{16}$$

Plan

Avant propos

Introduction

Modèle géométrique d'une caméra

Présentation du modèle

Transformation repère scène \rightarrow repère caméra

Transformation repère caméra \rightarrow repère image

Calibrage d'une caméra

Capteur stéréoscopique

Calibrage d'une caméra

- Deux éléments :
 1. Estimer la matrice de projection \mathbf{M} = un problème de resection ou *resection problem*
 2. Estimer, éventuellement, si nécessaire, les paramètres extrinsèques et intrinsèques de la caméra : \mathbf{A} et \mathbf{K}
- En général, méthodes de calibrage =
 1. Estimer la matrice de projection
 2. À partir de cette matrice estimer les paramètres intrinsèques et extrinsèques
- Nombreuses solutions : sans ou **avec points connus dans la scène**

Calibrage d'une caméra : le problème

- **Hypothèse** : N correspondances de points connues
 $\mathbf{P}_i = (X_i \ Y_i \ Z_i)$ de la scène vers $\mathbf{p}_i = (u_i \ v_i \ 1)$ de l'image

$$\lambda_i \mathbf{p}_i = \mathbf{M} \mathbf{P}_i. \quad (17)$$

- **Inconnues** : 11 paramètres de \mathbf{M} et les $N\lambda_i$
- Il y a $3N$ équations
- Il faut $3N \geq 11 + N$, soit $N \geq 6$

Calibrage d'une caméra : la solution simple

- Transformation linéaire directe ou *Direct Linear Transformation* (DLT)

$$\mathbf{M} = \begin{pmatrix} \mathbf{m}_1^T \\ \mathbf{m}_2^T \\ \mathbf{m}_3^T \end{pmatrix}$$

- Avec $\mathbf{0}_{1 \times 4} = \mathbf{0}$

$$\begin{pmatrix} \mathbf{P}_i & \mathbf{0} & \mathbf{0} & -u_i \\ \mathbf{0} & \mathbf{P}_i & \mathbf{0} & -v_i \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_i & -1 \end{pmatrix} \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \\ \lambda_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (18)$$

Calibrage d'une caméra : en pratique

- **Nombreuses approches** : avec points connus de la scène ou sans
- Utilisation de photos acquises en présence d'une mire : en TP, en réalité augmenté, un damier (plein d'autres possibilités)
- Détection de points d'intérêt de cette mire : par exemple, les coins du damier bien contrastés
- À partir des correspondances de points, estimation de la matrice **M**
- À partir de **M**, calcul de **A** et **K**

Plan

Avant propos

Introduction

Modèle géométrique d'une caméra

Présentation du modèle

Transformation repère scène \rightarrow repère caméra

Transformation repère caméra \rightarrow repère image

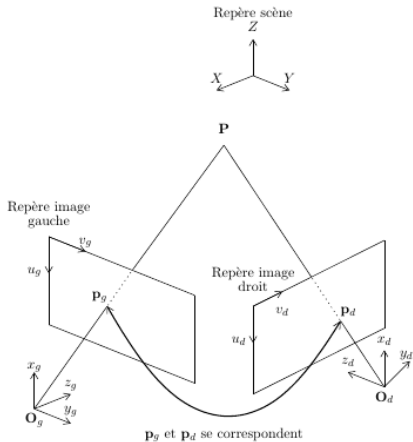
Calibrage d'une caméra

Capteur stéréoscopique

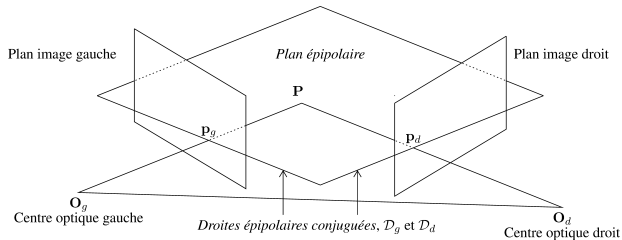
Modèle stéréoscopique

- Deux images acquises de deux points de vue différents
 - Deux caméras
 - ou un déplacement
- Position relative de telle sorte qu'une grande partie de la scène soit visible dans les deux points de vue
- Pas de rotation autour du centre optique de la caméra sinon reconstruction impossible
- Deux matrices de projection perspective

Modèle stéréoscopique



Géométrie épipolaire



Contrainte et rectification épipolaire

- Contrainte épipolaire = relation entre les points de l'image 1 et les points de l'image 2 qui ne dépend que des paramètres des caméras et des coordonnées image et qui est indépendante des coordonnées du point de la scène
- Matrice fondamentale, cf. cours de vision par ordinateur avec Jean-Denis Durou
- Rectification épipolaire = transformation géométrique à appliquer aux deux images
- Difficulté = estimation des paramètres de cette transformation