

Géométrie

Contenu

1	Transformations élémentaires	2
2	Les projections	4
2.1	Les projections parallèles	4
2.2	Les projections perspectives	6
2.3	Points de fuite	7
3	La géométrie d'une image ou la formation d'images	10
3.1	Projection : le modèle à sténopé	11
3.2	Capteurs CCD	12
3.3	Transformation globale	13
3.4	Un modèle de caméra	14
3.5	Applications	16
3.5.1	Projection de plans	16
4	La géométrie de deux images	19
4.1	La contrainte épipolaire	19
4.2	La matrice fondamentale	20
4.3	La matrice essentielle	21
5	Reconstruction	23
5.1	Les différents types de reconstruction	23
6	Estimer la géométrie des caméras	25

Dans cette partie, nous allons aborder les problèmes géométriques en vision par ordinateur, en particulier les mécanismes de visualisation de scènes 3D ainsi que ceux de la formation d'une image.

Coordonnées homogènes

Nous nous plaçons dans l'espace projectif $\mathcal{P}^3$. Un point P de $\mathcal{P}^3$ est représenté par un vecteur 4x1 de coordonnées $[x, y, z, w]$, dont une au moins est non nulle. x, y, z et w sont les coordonnées homogènes de P .

- ☞ coordonnées homogènes d'un plan : (a, b, c, d) ,
- ☞ dualité point-plan,
- ☞ les points à l'infini : $(x, y, z, 0)$ définissent le plan à l'infini

1 Transformations élémentaires

Translations :

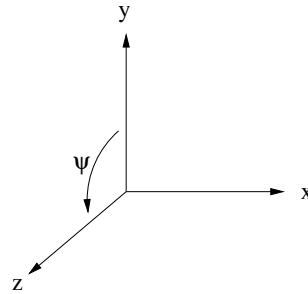
$$T = \begin{bmatrix} 1 & 0 & 0 & T_x \\ 0 & 1 & 0 & T_y \\ 0 & 0 & 1 & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotations : Représentation par les angles d'Euler :

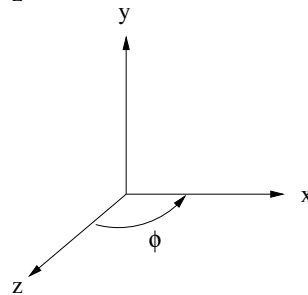
$$R = R_z \cdot R_y \cdot R_x,$$

avec :

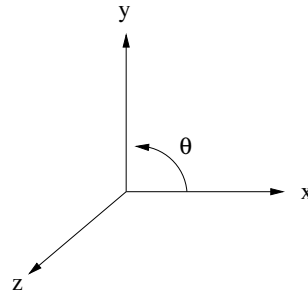
$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi & 0 \\ 0 & \sin \psi & \cos \psi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



$$R_y = \begin{bmatrix} \cos \phi & 0 & \sin \phi & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \phi & 0 & \cos \phi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



$$R_z = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Propriétés des matrices de rotation :

- $\det R = 1$,
- $R^{-1} = R^t$.
- groupe des rotations non commutatif.

Changements d'échelles :

$$T = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

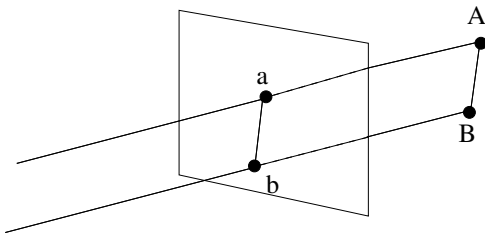
2 Les projections

Utilité :

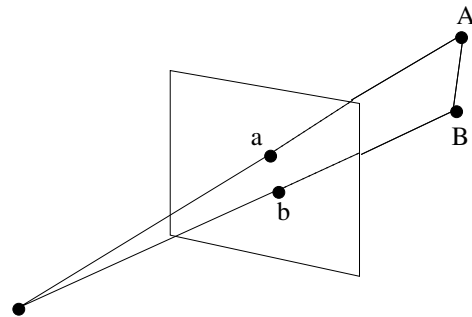
- Visualisation 3D (synthèse).
- Modélisation des caméras (vision).

Deux familles de projections :

1. Projections parallèles.
2. Projections perspectives.



Projection parallèle



Projection perspective

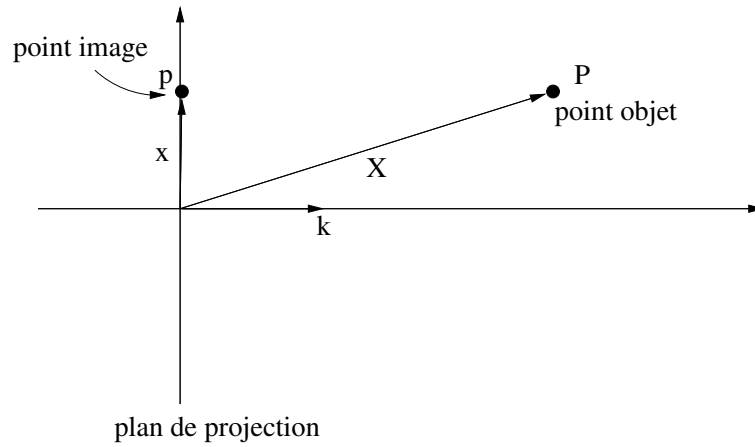
2.1 Les projections parallèles

La projection se fait dans le plan de projection suivant une direction. Il existe deux types de projections parallèles dépendant du fait que la direction de projection soit perpendiculaire au plan de projection :

1. Projections orthographiques (perpendiculaires).
2. Projections obliques (peu utilisées en analyse et synthèse d'images).

Les projections orthographiques

Le point P se projette au point p suivant la direction d'observation k .



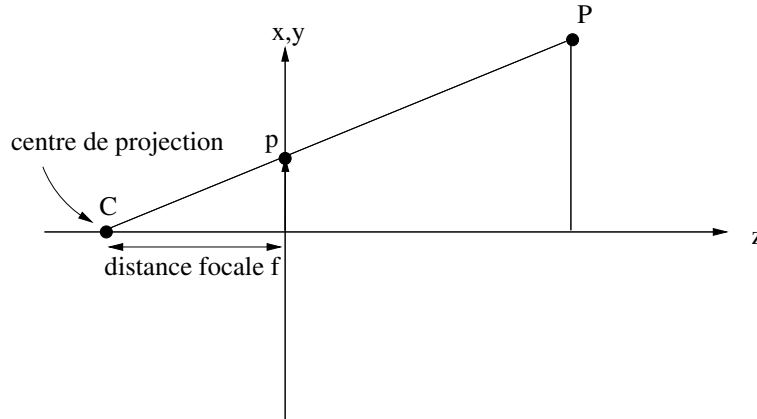
Soit X et x les vecteurs positions de P et p alors :

$$x = X - (X.k)k.$$

On choisit $k = (0, 0, 1)$ et on place le plan de projection en $z = 0$. Alors la projection orthographique s'écrit sous la forme matricielle suivante, et à l'aide des coordonnées homogènes :

$$x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot X$$

2.2 Les projections perspectives



La projection se fait suivant un centre de projection. On place l'origine du repère au centre de projection et on l'oriente de sorte que le plan de projection ait pour équation $z = f$ où f est la *distance focale*.

Soient X_P, Y_P, Z_P les coordonnées cartésiennes du point P et x_p, y_p, z_p celle du point image p . On a alors les relations :

$$\begin{cases} X_p = \frac{x_P f}{z_P}, \\ Y_p = \frac{y_P f}{z_P}, \\ Z_p = f. \end{cases}$$

d'où l'équation matricielle :

$$\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \cdot X$$

où $\mathbf{x}, X$ sont les vecteurs de coordonnées homogènes de p et P .

Autre représentation, le plan de projection est à $z = 0$ et le centre de projection à $z = -f$:

$$\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} \cdot X$$

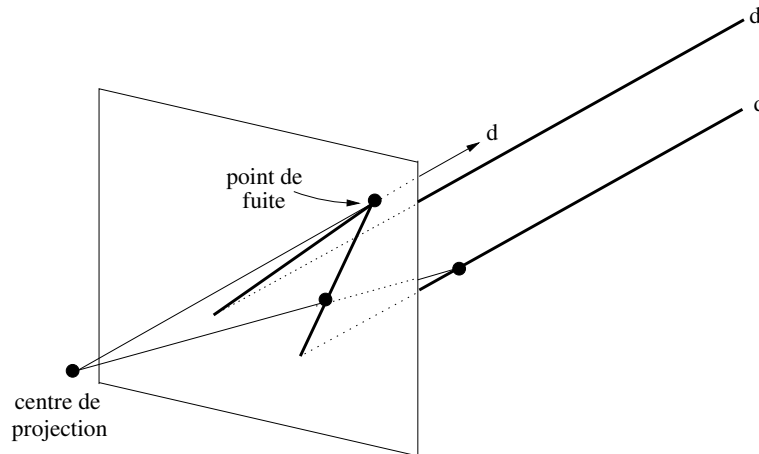
Exemple de projection : un cercle

Soit un cercle dans l'espace de rayon r , situé dans un plan parallèle au plan de projection à une distance Z de ce dernier.

1. On suppose que le cercle est centré en $(0, 0)$, montrer que sa projection est un cercle de rayon $fr/(Z + f)$.
2. On déplace le cercle une direction du plan de projection, que devient sa projection ?
3. si le plan du cercle n'est pas parallèle au plan de projection ?

2.3 Points de fuite

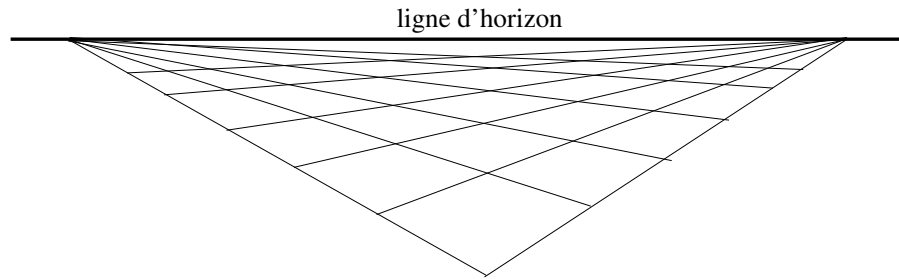
Les droites parallèles qui ne sont pas dans une direction du plan de projection définissent un *point de fuite*. Les projections de ces droites s'intersectent en effet dans le plan de projection, cette intersection étant le point de fuite.



Chaque paire de droites parallèles définit un point de fuite différent. Pour un ensemble de droites appartenant à un plan, les points de fuite parcourent une droite appelée ligne d'horizon.

De manière similaire, deux plans parallèles définissent une ligne d'horizon correspondant à la projection de l'intersection de ces deux plans. Les points à l'infini correspondent à des directions dans l'espace (cf. cours précédent).

Exercices :



1. Déterminez la position image du point de fuite associée à une droite de l'espace ?
2. Dans quel cas le point de fuite est rejeté à l'infini ?
3. Déterminez l'équation de la ligne d'horizon associée à un groupe de plans parallèles.
4. Dans quel cas la ligne d'horizon est rejetée à l'infini ?

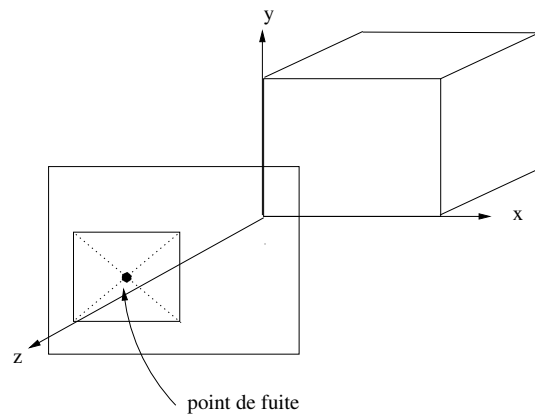


Figure 1: *Projection à un point de fuite.*

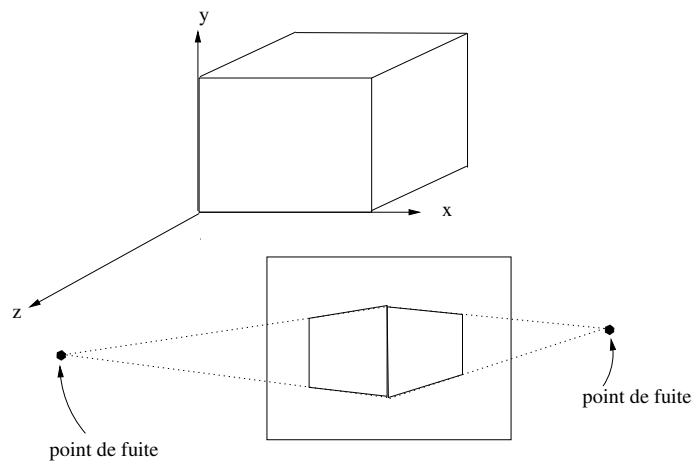


Figure 2: *Projection à deux points de fuite.*

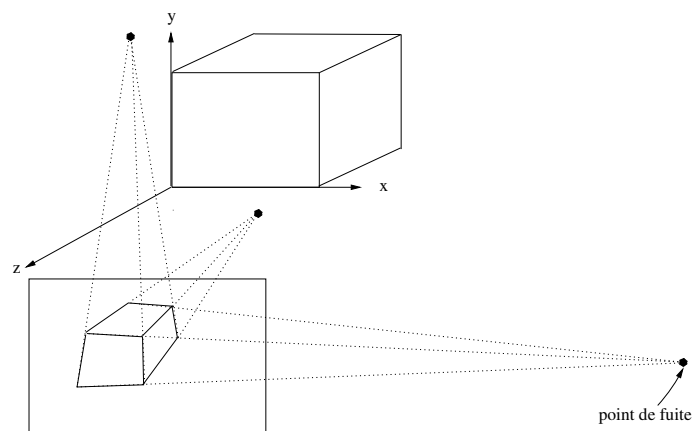


Figure 3: *Projection à trois points de fuite.*

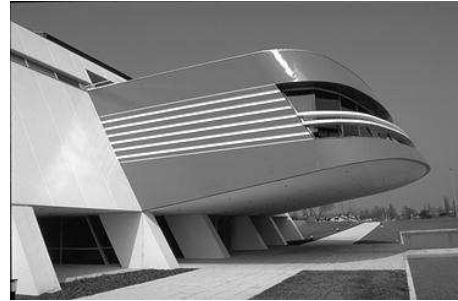
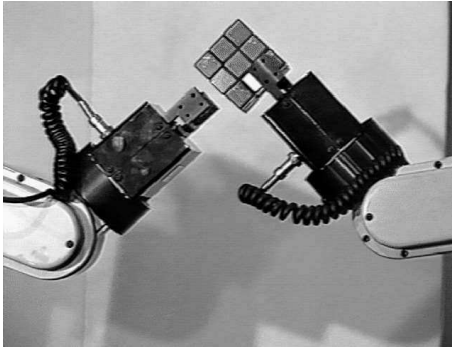
3 La géométrie d'une image ou la formation d'images

On se place dans le cas d'images réelles, prises avec un appareil numérique (CCD), et on cherche à modéliser la transformation scène $\leftrightarrow$ image. Celle-ci est composée de différentes transformations :

- projection,
- déplacement rigide,
- changement d'échelles.

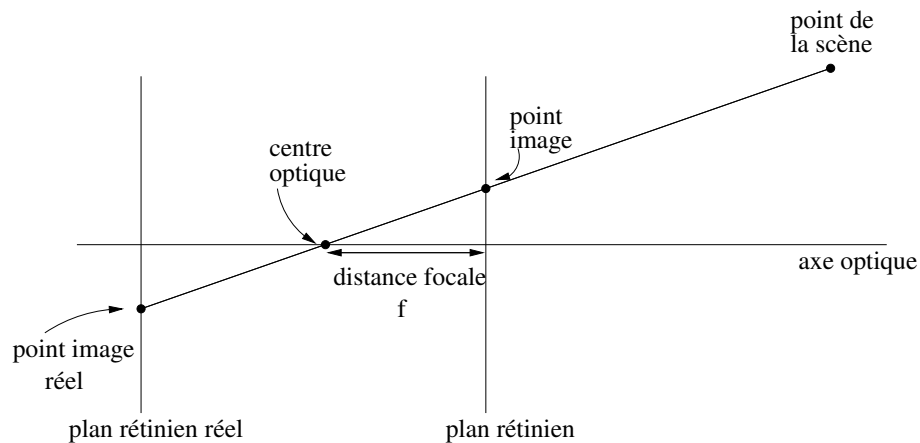
et peut donc être représentée sous forme matricielle.

3.1 Projection : le modèle à sténopé



projections parallèles ? perspectives ?

Le modèle de caméra le plus utilisé, *modèle à sténopé (pinhole)* :



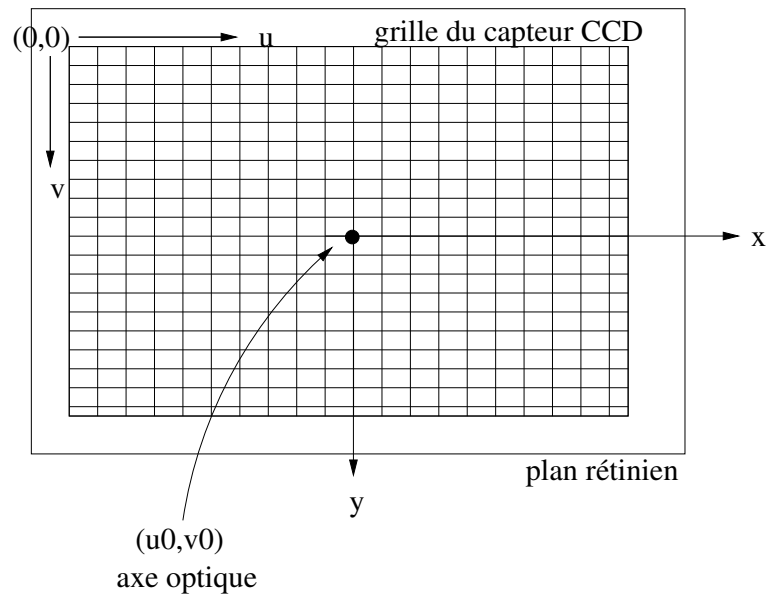
La matrice de projection correspondante :

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix}$$

qui représente donc la matrice de transformation du repère de la scène vers le repère de la rétine.

3.2 Capteurs CCD

Pour modéliser les capteurs CCD, on définit les coordonnées pixels (u, v) .

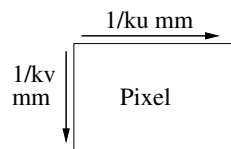


Transformation du repère de la rétine vers le repère image :

$$\begin{pmatrix} wu \\ wv \\ wz \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

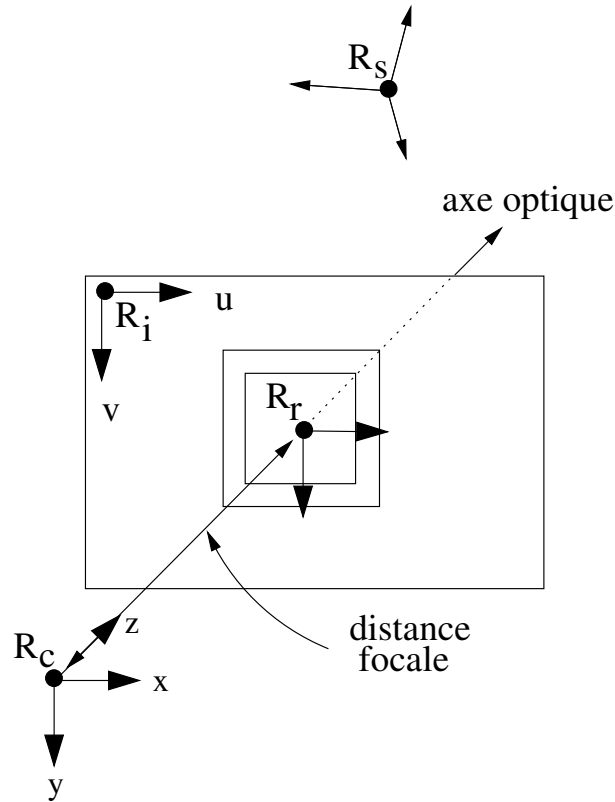
où :

1. k_u et k_v sont les facteurs d'échelles en pixels/mm.



2. (u_0, v_0) sont les coordonnées en pixels de l'intersection de l'axe optique avec le plan rétinien.

3.3 Transformation globale



La transformation du repère de la scène vers le repère image s'écrit :

$$\begin{pmatrix} wu \\ wv \\ wf \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \cdot \begin{pmatrix} R & T \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

En pratique, on utilise l'expression suivante dans laquelle la troisième coordonnée du point image n'apparaît plus (cette coordonnée est constante) :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \cdot \begin{pmatrix} R & T \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = {}^i_r C \cdot {}^r_c P \cdot {}^c_s T \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

soit encore :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} R & T \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = {}^i_r C \cdot {}^c_s T \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

☞ Paramètres intrinsèques et extrinsèques de la caméra.

☞ Une caméra est définie par 10 paramètres dans le cas où les pixels sont carrés, 11 dans le cas général.

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u f & \alpha & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} R & T \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

☞ Une matrice projective 3x4 correspond à 11 degrés de liberté.

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

3.4 Un modèle de caméra

Une caméra est généralement modélisée sous la forme projective suivante :

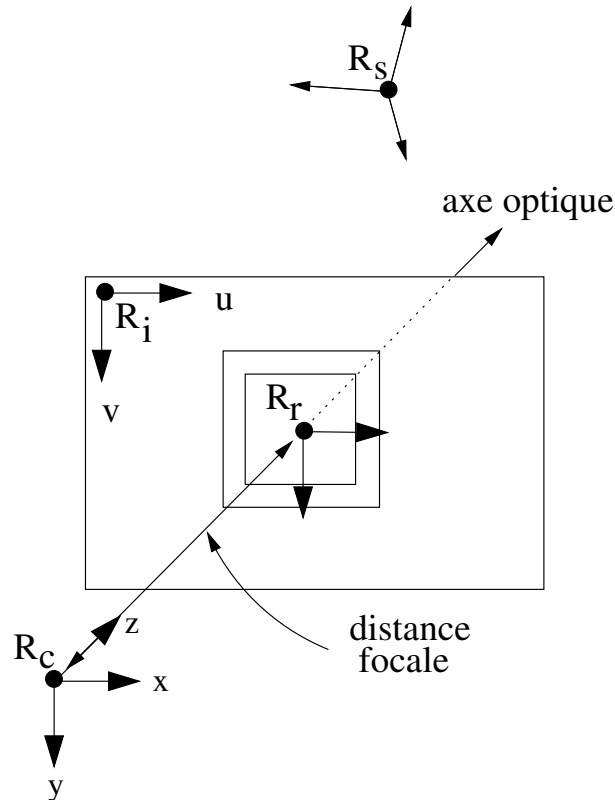
$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u & c & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \cdot \begin{pmatrix} R & T \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u f & c & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} R & T \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix},$$

où :

- k_u et k_v sont les facteurs d'échelles en pixels/mm,
- f est la distance focale. On note en général : $f k_u = \alpha_u$ et $f k_v = \alpha_v$,
- (u_0, v_0) sont les coordonnées en pixels de l'intersection de l'axe optique avec le plan rétinien,
- c est un coefficient qui est nul lorsque les axes de l'images sont orthogonaux.

Ces paramètres sont dits intrinsèques à la caméra.



Soit M la matrice de projection perspective dans une image. Si l'on note K la matrice des paramètres intrinsèques alors :

$$M = \begin{pmatrix} \alpha_u & c & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} R & T \end{pmatrix} = K \cdot \begin{pmatrix} R & T \end{pmatrix},$$

$$M = (K \cdot R \quad K \cdot T).$$

☞ M a onze degrés de liberté, il faut donc six correspondances entre points de la scène et points images pour déterminer cette matrice.

☞ On note par la suite $\overline{M}$ la matrice 3×3 $\overline{M} = K \cdot R$ et m le vecteur 3×1 $m = K \cdot T$. Soit $M = (\overline{M} \ m)$.

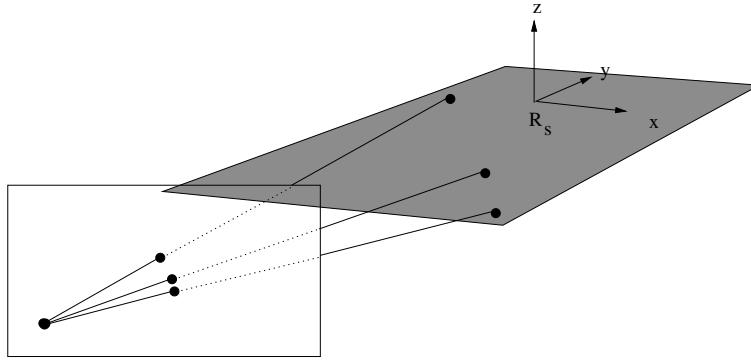
☞ Soit $C = (x_c \ y_c \ z_c)^t$ le vecteur de coordonnées du centre de projection de la caméra dans le repère de la scène, alors :

$$C = -\overline{M}^{-1} \cdot m.$$

3.5 Applications

3.5.1 Projection de plans

Nous allons illustrer le principe des projections dans le cas de points coplanaires. C'est à dire lorsque l'on observe des points appartenant à un même plan.



Le cas perspectif Nous supposons, sans perte de généralité, que les points appartiennent au plan d'équations $z = 0$ dans la scène. Nous pouvons réécrire l'équation de projection sous la forme :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{pmatrix} \cdot \begin{pmatrix} R_{11} & R_{12} & T_x \\ R_{21} & R_{22} & T_y \\ R_{31} & R_{32} & T_z \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

soit :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

La transformation entre (wu, wv, w) et $(x, y, 1)$ est une projectivité plan à plan, on parle d'*homographie*.

Exercice :

1. Que peut-on en déduire sur la transformation existant entre 2 projections perspectives d'une image ?
2. Dans le cas d'une projection orthographique que devient cette transformation ?

Le cas orthographique Nous allons tout d'abord établir l'équation de projection dans le cas orthographique. Reprenons l'équation de projection précédemment établie, dans le cas orthographique cette équation peut s'écrire (le plan de projection étant en $z = 0$ dans le repère de la caméra) :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{pmatrix} \cdot \begin{pmatrix} R & T \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

si l'on fait tendre f vers l'infini :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} R & T \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Maintenant si nous supposons que les points observés appartiennent au plan d'équation $z = 0$ alors l'équation de projection devient :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} R_{11} & R_{12} & T_x \\ R_{21} & R_{22} & T_y \\ R_{31} & R_{32} & T_z \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

soit :

$$\begin{pmatrix} wu \\ wv \\ w \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & m_{33} \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

☞ Dans le cas d'une projection orthographique, la transformation entre 2 projections d'une image est une transformation affine.

☞ Application dans le cas de texture mapping.

Exercice :

- Traitez le cas de points appartenant à une droite.

4 La géométrie de deux images

Considérons deux images provenant de deux caméras perspectives observant la même scène. Soient M_1 et M_2 les matrices de projections correspondant à ces deux images. Un point P de la scène se projette en $p_1 = M_1 \cdot P$ et $p_2 = M_2 \cdot P$. Nous allons nous intéresser ici aux liens qui existent entre p_1 et p_2 . Ces liens sont caractérisés par la *géométrie épipolaire* qui découle de la contrainte épipolaire.

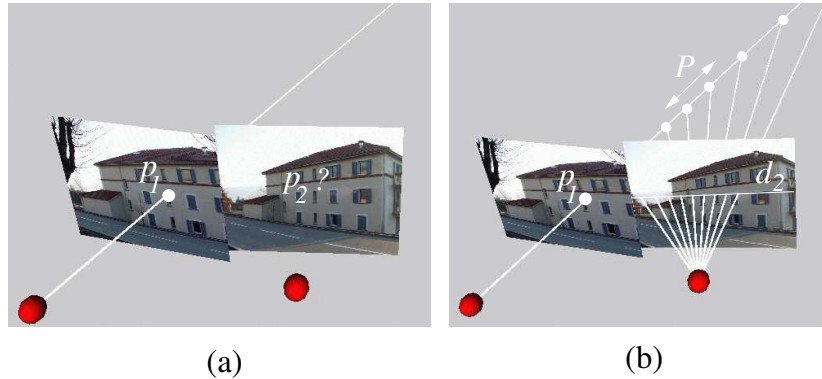


Figure 4: La géométrie épipolaire caractérise le fait que les correspondants potentiels dans l'image 2 d'un point p_1 de l'image 1 sont situés sur une droite d_2 appelée *droite épipolaire*.

4.1 La contrainte épipolaire

La contrainte épipolaire caractérise le fait le correspondant p_2 d'un point p_1 (i.e., p_1 et p_2 sont les projections images du même point P) se situe sur une droite dans l'image 2 : l_{p_1} . En effet, le point p_2 appartient nécessairement au plan défini par p_1, C_1 et C_2 . La droite l_{p_1} est appelée *droite épipolaire* du point p_1 dans l'image 2.

- ☞ La contrainte épipolaire est symétrique.
- ☞ Les droites épipolaires dans une image s'intersectent toutes en un point appelé *épipole*. Ce point correspond à la projection du centre de projection de l'autre image considérée :

$$e_1 = M_1 \cdot C_2,$$

$$e_2 = M_2 \cdot C_1.$$

- ☞ les épipoles jouent un rôle fondamental dans la vision stéréoscopique.

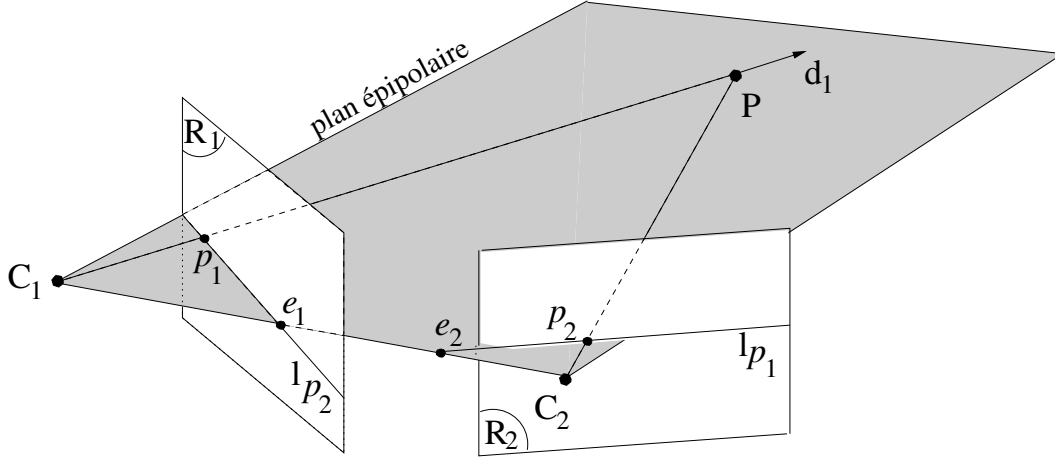


Figure 5: La contrainte épipolaire.

4.2 La matrice fondamentale

Nous allons exprimer ici, sous une forme algébrique, la contrainte épipolaire. Soit d_1 la direction de la ligne de vue de p_1 . Alors si P et C_1 sont exprimés en coordonnées homogènes :

$$P = C_1 + \lambda_1 \begin{pmatrix} d_1 \\ 0 \end{pmatrix},$$

où λ_1 est la profondeur du point P dans l'image 1 et $(d_1 \ 0)^t$ correspond au point à l'infini de la ligne de vue de p_1 . En projetant P :

$$p_1 = M_1 \cdot P = (\overline{M}_1 \ m_1) \cdot P,$$

$$p_1 = \overline{M}_1 \cdot d_1.$$

Soit :

$$d_1 = \overline{M}_1^{-1} \cdot p_1,$$

Tous les points de la ligne de vue de p_1 se projette sur la droite épipolaire l_{p_1} de p_1 dans l'image 2, en particulier le point à l'infini $(d_1 \ 0)^t$. La droite épipolaire l_{p_1} passe donc par l'épipoles e_2 et la projection dans l'image 2 de $(d_1 \ 0)^t$, soit en coordonnées homogènes :

$$l_{p_1} = e_2 \wedge M_2 \cdot \begin{pmatrix} d_1 \\ 0 \end{pmatrix},$$

et :

$$l_{p_1} = e_2 \wedge (\overline{M}_2 \cdot \overline{M}_1^{-1} \cdot p_1)$$

e_2 étant la projection de C_1 dans l'image 2 :

$$l_{p_1} = (M_2 \cdot C_1) \wedge (\overline{M_2} \cdot \overline{M_1}^{-1} \cdot p_1),$$

et :

$$l_{p_1} = (M_2 \cdot -\overline{M_1}^{-1} \cdot m_1) \wedge (\overline{M_2} \cdot \overline{M_1}^{-1} \cdot p_1).$$

La relation précédente est linéaire et peut s'écrire sous la forme :

$$l_{p_1} = F \cdot p_1,$$

ou F est une matrice 3x3. Cette matrice est appelée *matrice fondamentale*. La contrainte épipolaire s'écrit donc :

$$p_2^t \cdot F \cdot p_1 = 0.$$

- ☞ La matrice F est de rang deux. En effet : $F \cdot e_1 = 0$.
- ☞ La matrice F^t joue le rôle symétrique de F dans l'image 1, c'est à dire que pour tout point p_2 de l'image 2 le point correspondant de l'image 1 appartient à la droite de coordonnées homogènes l_{p_2} :

$$l_{p_2} = F^t \cdot p_2.$$

- ☞ La matrice fondamentale caractérise toute la géométrie entre deux images hors paramètres intrinsèques.
- ☞ La matrice fondamentale peut être estimée à partir d'au moins huit correspondances de points entre l'image 1 et l'image 2.

4.3 La matrice essentielle

Lorsque les paramètres intrinsèques des caméras sont connus, il est possible d'exprimer la contrainte épipolaire dans les repères rétinien. C'est à dire, pour un point image p_1 de l'image 1, d'exprimer la contrainte correspondant à :

$$x_1 = K_1^{-1} \cdot p_1,$$

où K_1 est la matrice des paramètres intrinsèques de l'image 1. Par simple analogie avec ce qui a été vu précédemment, la contrainte épipolaire devient :

$$x_2^t \cdot K_2^t \cdot F \cdot K_1 \cdot x_1 = 0,$$

soit :

$$x_2^t \cdot E \cdot x_1 = 0.$$

avec :

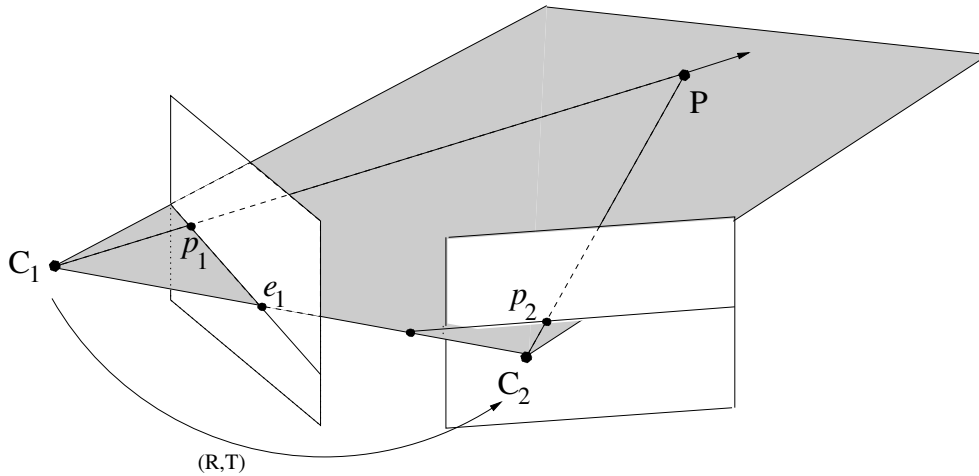
$$E = K_2^t \cdot F \cdot K_1.$$

La matrice E de l'expression précédente est appelée *matrice essentielle*. Cette matrice dépend uniquement du déplacement entre les deux positions des caméras. En effet, posons :

$$M_1 = K_1 \cdot (I \ 0), \quad M_2 = K_2 \cdot (R \ T),$$

où R et T sont respectivement la rotation et la translation entre les deux positions de la caméra. La contrainte de coplanarité s'écrit alors :

$$p_2^t \cdot K_2^{-t} (T \wedge R \cdot K_1^{-1} \cdot p_1),$$



Soit :

$$x_2^t \cdot (T \wedge R \cdot x_1).$$

- ☞ La translation T est ici définie à un facteur d'échelle près.
- ☞ La matrice essentielle est de rang deux, de plus les deux valeurs propres non nulles de E ont même valeurs.
- ☞ La matrice essentielle peut être estimée, théoriquement, à partir de cinq correspondances.
- ☞ Les épipoles ont pour coordonnées homogènes dans le plan rétinien :

$$\begin{aligned} K_1^{-1} \cdot e_1 &= -R^t \cdot T, \\ K_2^{-1} \cdot e_2 &= T. \end{aligned}$$

5 Reconstruction

La reconstruction consiste à déterminer les positions dans l'espace des primitives de la scène à partir des images de cette dernière. Nous allons nous intéresser ici aux reconstructions à partir de deux images.

Soit P un point de la scène se projetant en p_1 et p_2 . Reconstruire P consiste à déterminer sa position dans l'espace, soit à déterminer l'intersection des lignes de vues en p_1 et p_2 (triangulation). Soient M_1 et M_2 les deux matrices de projections des images 1 et 2. Alors :

$$p_1 = M_1 \cdot P, \quad p_2 = M_2 \cdot P.$$

Connaissant M_1 et M_2 la résolution des deux équations précédentes donne la position de P . Par contre, si M_1 et M_2 ne sont pas connus, le point P ne constitue pas la seule solution, en effet, soit H une homographie de $\mathcal{P}^3$ alors :

$$p_1 = M_1 \cdot H^{-1} \cdot H \cdot P, \quad p_2 = M_2 \cdot H^{-1} \cdot H \cdot P.$$

Le point $H \cdot P$ est donc aussi solution des équations de projections. La reconstruction de la scène s'effectue donc à la transformation H près. Cette transformation étant fonction des informations dont on dispose sur la scène.

☞ Un aspect fondamental de la reconstruction est que pour un point p_1 de l'image 1, il faut avant tout déterminer le point p_2 correspondant dans l'image 2. Ce problème *d'appariement* est primordial en vision par ordinateur.

5.1 Les différents types de reconstruction

Soient M_1 et M_2 les matrices de projection des images 1 et 2 et $\tilde{M}_1$ et $\tilde{M}_2$ les matrices estimées. Soit H la transformation telle que :

$$M_1 = \tilde{M}_1 \cdot H, \quad M_2 = \tilde{M}_2 \cdot H.$$

Reconstructions Euclidiennes

Considérons tout d'abord une scène dans laquelle des points de référence sont connus. Si les positions d'au moins six points de références sont connues, il est alors possible de déterminer pour chaque image les matrices :

$$\tilde{M}_1 = K_1 \cdot (R_1 \ T_1), \quad \tilde{M}_2 = K_2 \cdot (R_2 \ T_2),$$

par simple résolution des équations de projections aux points de référence. La reconstruction s'effectuera donc ici dans le repère où sont exprimés les points de

référence. H est donc ici une transformation Euclidienne.

Supposons maintenant qu'aucun point de référence n'est connu dans la scène mais que par contre les paramètres intrinsèques des caméras sont eux connus (K_1 et K_2). Il est alors possible de déterminer le déplacement entre des deux images à partir de correspondances de points. Par exemple :

1. Mise en correspondance de points,
2. Calcul de la matrice fondamentale,
3. Calcul de la matrice essentielle et du déplacement (R, T) entre les images 1 et 2.

On obtient alors les deux matrices de projection estimées suivantes :

$$\tilde{M}_1 = K_1 \cdot (I \ 0), \quad \tilde{M}_2 = K_2 \cdot (R \ \lambda T),$$

où : λ est un réel strictement positif. Ces deux matrices permettent donc d'effectuer une reconstruction de la scène à un facteur près. En effet, les points solutions des équations de projections sont de la forme : $P = (x \ y \ z \ 1/\lambda)$, où λ est un réel strictement positif. Donc H est ici une similitude.

Reconstructions affines

Les paramètres intrinsèques des caméras ne sont pas connus, par contre l'homographie du plan à l'infini H_∞ entre les deux points vue est connue. Soient $M_1 = K_1 \cdot (I \ 0)$ et $M_2 = K_2 \cdot (R \ T)$ les matrices de projection des images 1 et 2. L'homographie du plan à l'infini caractérise la transformation dans l'espace projectif, des directions de l'espace :

$$H_\infty = K_2 \cdot R \cdot K_1^{-1}.$$

Cette homographie peut être déterminée à partir de correspondances de points de fuites par exemple.

Lorsque H_∞ est connue, on peut alors définir les deux matrices de projections :

$$\tilde{M}_1 = (I \ 0), \quad \tilde{M}_2 = (H_\infty \ e_2),$$

et effectuer une reconstruction à partir de ces matrices. La transformation H est donc ici une affinité :

$$H = \begin{pmatrix} K_1 & 0 \\ 0 & 1 \end{pmatrix}$$

Reconstructions projectives

C'est le cas le plus général, aucune information *a priori* sur la scène n'est connue. La forme générale des matrices de projections est :

$$\tilde{M}_1 = (I \ 0), \quad \tilde{M}_2 = (H_i \ ce_2),$$

où H_i est l'homographie d'un plan de la scène ne contenant pas les centres de projection, et c un scalaire. Le passage d'une reconstruction projective à une reconstruction affine puis Euclidienne se fait à l'aide des relations :

$$\tilde{M}_1 = (I \ 0) = (K_1 \ 0) \cdot \begin{pmatrix} K_1^{-1} & 0 \\ 0 & 1 \end{pmatrix} = (K_1 \ 0) \cdot \begin{pmatrix} K_1^{-1} & 0 \\ a^t & c \end{pmatrix}$$

$$\tilde{M}_2 = (H_i \ ce_2) = (K_2 \cdot R \ K_2 \cdot T) \cdot \begin{pmatrix} K_1^{-1} & 0 \\ a^t & c \end{pmatrix}$$

soit :

$$\tilde{M}_1 = M_1 \cdot \begin{pmatrix} K_1^{-1} & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} I & 0 \\ a^t & c \end{pmatrix}$$

$$\tilde{M}_2 = M_2 \cdot \begin{pmatrix} K_1^{-1} & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} I & 0 \\ a^t & c \end{pmatrix}$$

Donc la transformation H s'écrit :

$$H = \begin{pmatrix} K_1 & 0 \\ -1/c (K_1^t \cdot a)^t & 1/c \end{pmatrix}$$

6 Estimer la géométrie des caméras

Nous passons brièvement en revue les principaux moyens d'estimer la géométrie des caméras.

Le moyen le plus simple, le *calibrage*, consiste à utiliser un objet dont la géométrie est connue. A partir d'une seule image de cet objet, la géométrie complète de la caméra peut alors être estimée : sa géométrie interne et son positionnement relatif par rapport à l'objet. Si la modélisation de la scène est effectuée à l'aide de plusieurs caméras *statiques*, la géométrie complète de ce système peut être estimée de la même manière en plaçant un objet connu dans le champ de vue commun des caméras. Des logiciels du domaine public¹ permettent, en particulier, de calibrer à l'aide d'une grille (voir la figure 1). Les caméras peuvent donc être



Figure 6: Un objet de calibration très largement utilisé : le damier. Les sommets constituent en effet des primitives qui sont facilement détectables dans les images, et dont les positions dans la scène sont connues.

calibrées et leurs positions relatives peuvent être déduites des positions relatives à l'objet de calibration.

Ce scénario n'est pas toujours applicable, en particulier si, au cours de l'acquisition d'images, les caméras se déplacent, ou si le zoom ou la mise au point sont modifiés (et, par voie de conséquence, la géométrie interne), ou bien si l'on souhaite éviter le processus de calibration pour des raisons de temps ou d'équipement. Dans ce cas, les outils de la géométrie multi-images doivent être mis en œuvre. En effet, comme nous l'avons vu dans la partie précédente, des correspondances entre images, peuvent permettre d'estimer la géométrie des caméras, même si la structure de la scène n'est pas connue. Dans le cas le plus général, les géométries interne *et* externe des caméras doivent être estimées à partir de ces correspondances. On parle alors d'*auto-calibration*. Il a été prouvé, au début des années 90, que l'auto-calibration était effectivement possible. Plusieurs approches ont été proposées depuis. La mise en œuvre pratique de l'auto-calibration reste néanmoins une opération délicate.

Une solution intermédiaire consiste à pré-calibrer les caractéristiques intrinsèques des caméras (la géométrie interne) que l'on suppose fixes, et à déterminer ensuite la géométrie externe uniquement. Dans ce cas, un auto-calibration complet n'est pas nécessaire, la géométrie interne des caméras étant connue. On parle alors d'*estimation du mouvement*. Il s'agit d'un problème plus simple, mais qui interdit l'usage du zoom (car il modifie la géométrie interne) et qui requiert toujours une étape préliminaire de calibration. Le choix de l'approche pour l'estimation de la géométrie des caméras, dépend du contexte de l'application envisagée et des contraintes qui en découlent : temps, précision, nombre de caméras, *etc.* Le tableau 1

¹<http://intel.com/research/mrl/research/opencv/>

récapitule les principales opérations géométriques existantes, et ce qui doit être connu et ce qui est déterminé pour chaque opération.

Opération	Géométrie de la scène	Géométrie externe des caméras	Géométrie interne des caméras
Calibrage	connue	déterminée	déterminée
Calcul de pose	connue	déterminée	connue
Triangulation	déterminée	connue	connue
Estimation du mouvement	inconnue	déterminée	connue
Reconstruction non-métrique	déterminée	inconnue	inconnue
Auto-calibrage	inconnue	inconnue	déterminée

Table 1: Les différentes opérations géométriques et les contextes correspondants.