L'espace réel est de nature tridimensionnelle, alors que les outils usuels de formation d'une image sont bidimensionnels.

La projection est la fonction de passage du 3D au 2D . L'image 2D est la résultante de d'une combinaison des 3 dimensions de l'objet. Cette combinaison donne un aspect géométrique spécifique lorsque la géométrie de l'objet existe (droite, plan, polygone, polyèdre..) .

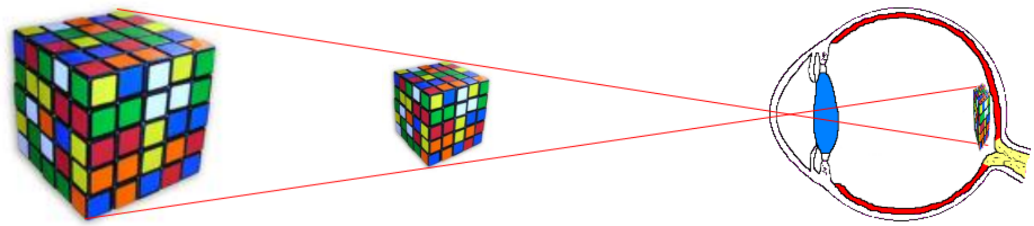


fig 1 - Projection

Le contenu de l'image 2D peut permettre de retrouver certains éléments caractéristiques de la scène initiale, sous réserve d'apporter des informations complémentaires (*connaissance à priori*).

Cependant, le seul moyen pour retrouver les caractéristiques 3D est d'augmenter la quantité d'information mesurée; l'une des solutions est la stéréoscopie qui combine deux points de vue différents pour reconstruire la scène 3D .

Projection

La projection est la loi géométrique selon laquelle les points de l'espace réel 3D se projettent sur la surface du capteur.

On suppose que le repère de l'espace des objets (environnement extérieur) correspond avec celui de la caméra.

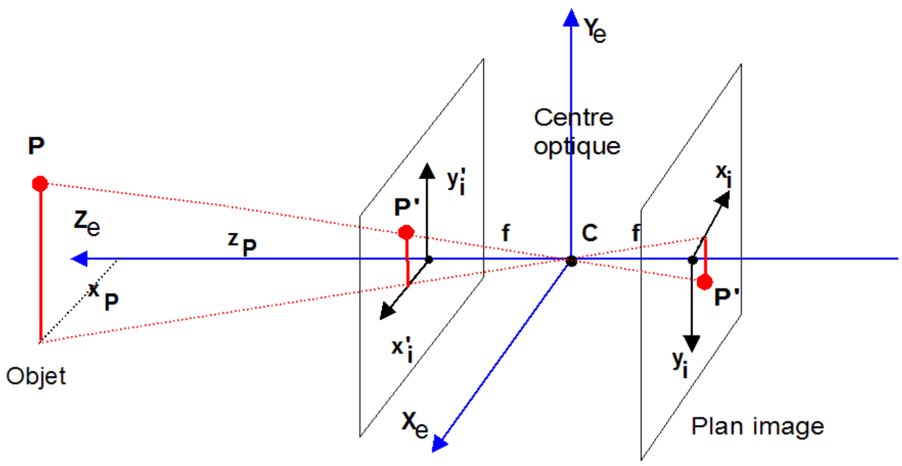


fig 2 - Projection géométrique

Les conditions géométriques sont les suivantes (fig 2):

- l'image se forme dans un plan (surface sensible du capteur) perpendiculaire à l'axe optique du système
- l'origine du repère environnement extérieur (X_e, Y_e, Z_e) est supposée être en C, centre optique de la lentille (point nodal du plan principal objet pour un système optique)

- l'axe Z est confondu avec l'axe principal du système optique
- les axes x et y du plan image sont parallèles aux axes X et Y du repère objet et leur origine est sur l'axe Z.
- le système optique est du type *sténopée* (pin hole)

Soit un objet P de coordonnées (x_e, y_e, z_e) dans le repère (X_e, Y_e, Z_e) et (x_i, y_i) les coordonnées de son image dans le plan du capteur. L'image se formant avec une *inversion,* on adopte souvent comme repère caméra le repère symétrique par rapport au centre optique (x_i', y_i') .

On remarque la relation par triangles semblables:

$$\frac{x_i}{x_e} = \frac{f}{z_e} \qquad \text{de même} \qquad \frac{y_i}{y_e} = \frac{f}{z_e}$$

On en déduit : $$x_i = f\frac{x_e}{z_e} \quad \text{et} \quad y_i = f\frac{y_e}{z_e}$$

On constate que les coordonnées de la projection P' de P dépendent de la profondeur z_P et que sa *localisation* (x_P, y_P) ne peut se faire à partir de $(x_{P'}, y_{P'})$ que si z_P est connu (cas où les objets sont dans plan parallèle au plan du capteur par exemple).

L'ensemble des points de coordonnées (ax_P, ay_P, az_P) appartenant à la droite PC se projette sur la même image; c'est pourquoi la profondeur z ne peut s'évaluer à partir d'une simple prise de vue !

Perspective

Le résultat de la projection provoque la notion de *perspective*. L'exemple le plus visible est celui de l'image d'un rectangle face à la caméra, ayant subi une rotation autour de l'axe vertical d'un angle θ /

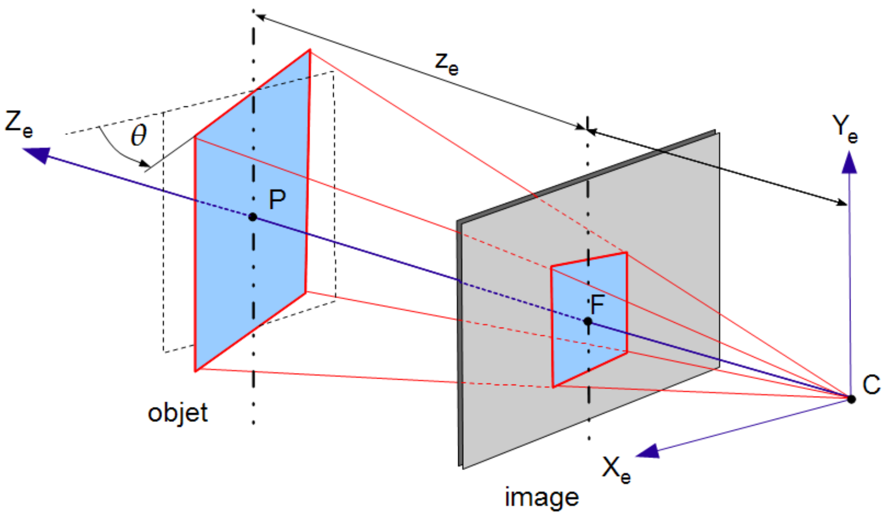


fig 2 - Perspective d'un carré

Les bords verticaux à distance constante sont projetées parallèlement à leur propre direction et restent parallèles entre eux. Les bords supérieur et inférieur sont projetés selon des droites non parallèle. Leur point d'intersection est appelé *point de fuite* .

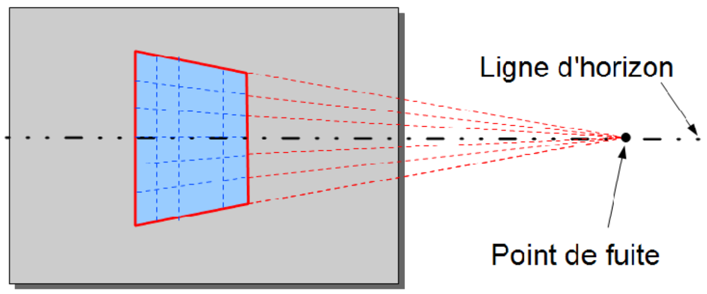


fig 3 - Ligne d'horizon et point de fuite en perspective

Lorsque la caméra n'est centrée sur le repère de l'environnement, la ligne d'horizon est modifiée.

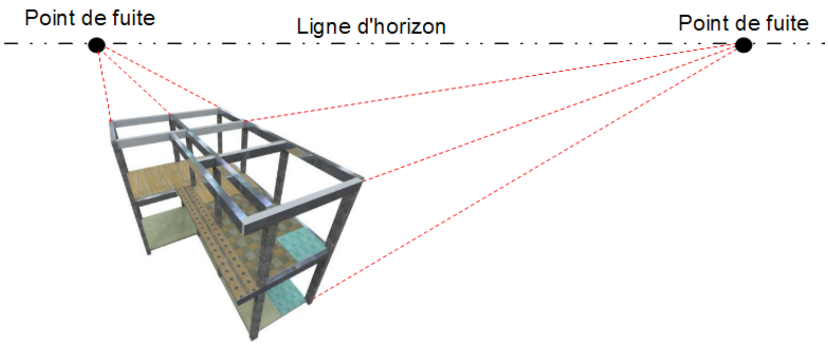


fig 4 - Points de fuite d'un objet tridimensionnel (3ème non tracé)

Paramètres extrinsèques de la caméra

Dans un montage réel, il est impossible de garantir que le repère réel soit aligné avec celui de la caméra.

La position de la caméra dans l'espace réel est la composition de :

- une translation du centre C par rapport à l'origine du repère (X_e, Y_e, Z_e)

- trois rotations des axes caméras par rapport au repère de l'environnement

Les paramètres caractérisant la relation entre la caméra et le référentiel sont dits *extrinsèques* .

Le passage des coordonnées environnement aux coordonnées caméra est :

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix} \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix} \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

L'ensemble de ces paramètres est représenté en coordonnées homogènes :

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{x/x} & r_{x/y} & r_{x/z} & t_x \\ r_{y/x} & r_{y/y} & r_{y/z} & t_y \\ r_{z/x} & r_{z/y} & r_{z/z} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_e \\ y_e \\ z_e \\ 1 \end{bmatrix}$$

Connaissant f ,la projection perspective sur la caméra peut s'écrire de façon matricielle:

$$\begin{bmatrix} x'_c . w \\ y'_c . w \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \quad \text{avec homogénéisation par} \quad w = z_e$$

Paramètres intrinsèques de la caméra

L'image se forme sur le capteur positionné dans le boitier caméra. Le capteur possède des propriétés géométriques spécifiques de positionnement dans le boitier car il n'est pas un simple plan de projection parfait. Ces paramètres sont dits paramètres *intrinsèques* de la caméra car ils ne changent pas si on déplace la caméra.

- la distance focale f (souvent mal connue)

- l'origine géométrique (O_x, O_y) du capteur qui n'est généralement pas l'axe optique,

- des facteurs d'échelle S_x et S_y permettant de traduire les facteurs associés à des pixels non-carrés et de faire un repérage évalué en pixel; ces coefficients s'expriment en pixel/m.

- un angle entre les axes de repère de la matrice (souvent négligé)

Nous obtenons en coordonnées homogènes:

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{S_x} & 0 & O_x \\ 0 & \frac{1}{S_y} & O_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x'_c \\ y'_c \\ 1 \end{bmatrix}$$

Le modèle final est donc :

$$\begin{bmatrix} x_i . w \\ y_i . w \\ w \end{bmatrix} = \begin{bmatrix} \textit{Transformation} \\ \textit{paramètres} \\ \textit{intrinsèques} \end{bmatrix} \left[\textit{Projection} \right] \begin{bmatrix} \textit{Transformation} \\ \textit{paramètres} \\ \textit{extrinsèques} \end{bmatrix} \begin{bmatrix} x_e \\ y_e \\ z_e \\ 1 \end{bmatrix}$$

Prise en compte des distorsions

Les optiques introduisent souvent une distorsion liée à la distance du point image au centre optique. Ces distorsions sont appelées distorsion en *coussin* ou *tonneau* ;

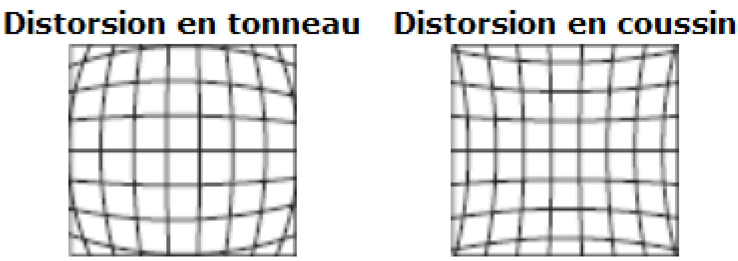


fig 5 - Distorsion d'une image

Une telle distorsion dite *radiale* peut être modélisée par :

$$\begin{aligned} x''_c &= (1 + k_2 r^2 + k_4 r^4 + k_6 r^6) x'_c \\ y''_c &= (1 + k_2 r^2 + k_4 r^4 + k_6 r^6) y'_c \end{aligned} \quad \text{avec} \quad r2 = (x'^2_c + y'^2_c)$$

La distorsion tangentielle a pour forme :

$$\begin{aligned} x'''_c &= 2\mathrm{p}_1 x''_c y''_c + p_2 (r^2 + 2x''^2) \\ y'''_c &= p_1 (r^2 + 2x''^2) + 2\mathrm{p}_2 x''_c y''_c \end{aligned}$$

Ces modèles de distorsion viennent compléter le modèle des transformations intrinsèques de la caméra. Ils ne se mettent pas sous forme matricielle.

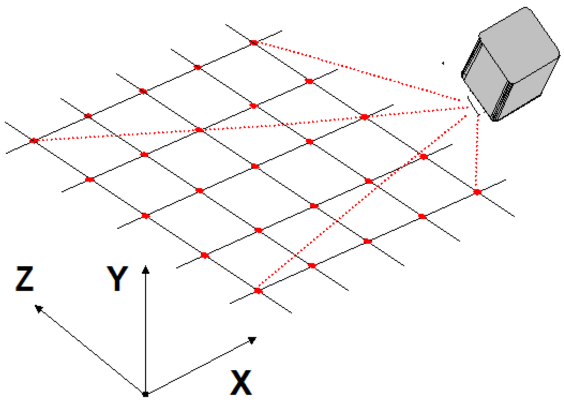
Étalonnage de la caméra

Si on néglige les distorsions , la projection peut être vue comme une transformation globale de la forme:

$$\begin{bmatrix} x_i . w \\ y_i . w \\ w \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x_e \\ y_e \\ z_e \\ 1 \end{bmatrix}$$

Cette transformation est est plus générale (12 inconnues alors que la projection a fait apparaître 11 paramètres - 3 translations, 3 rotations, la distance focale , 2 coordonnées de l'origine du capteur et 2 facteurs d'échelle -); il faut lui fixer une contrainte, par exemple $m_{34} = 1$.

La détermination des 11 inconnues en disposant d'au moins 6 couples de mesure soit 6 points non coplanaires.



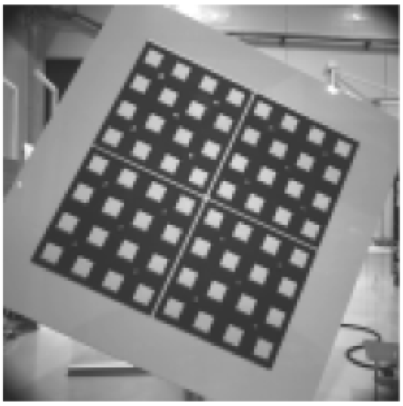


fig 6 - Etalonnage par une grille

Pour améliorer la précision, l'étalonnage se fait souvent en plaçant devant la caméra une grille de points de référence. La méthode des moindres carrés sera employée résoudre le système.

Il existe de nombreuses méthodes/programmes pour déterminer les paramètres de la projection, y compris ceux de distorsions. Nous pouvons citer :

- Camera calibration Toolbox for Matlab http://www.vision.caltech.edu/bouguetj

- Camera Calibration and 3D reconstruction http://opencv.willowgarage.com

- Direct Linear Transformation Method http://www.kwon3d.com