

IFT 6145
Vision tridimensionnelle
Modèles de caméra I

Sébastien Roy
Département d'informatique et de recherche opérationnelle
Université de Montréal

hiver 2004

Au programme...

- Formation de l'image (2.2)
- Paramètres de caméra (2.4)
 - Paramètres externes (2.4.2)
 - Paramètres internes (2.4.3)
 - Modèles de caméra (2.4.4)
- Représentation projective

Formation de l'image

D'où provient l'image ?

Système optique

Paramètres associés à la caméra.

(lentille, distance focale, angle de vue, ...)

Photométrie

Paramètres associés à la lumière.

(illumination, réflectance des surfaces, sensibilité du senseur)

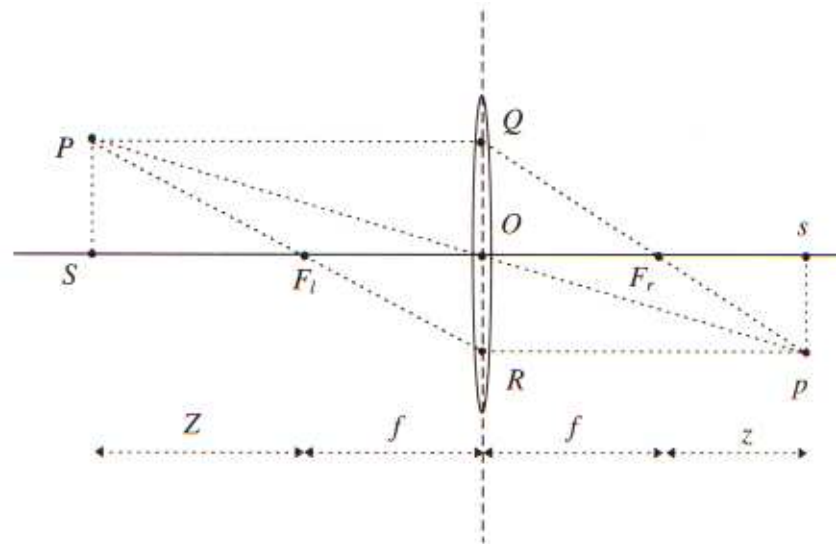
Géométrie

Paramètres de la relation entre le monde et l'image.

(type de projection, position et orientation de la caméra, distorsion, ...)

Un peu de photographie

Les lentilles....



On apprend par ses erreurs...

Source : <http://www.vanwalree.com>

Les défauts de l'image

Aberrations de Siedel

- Aberrations sphériques
- Coma
- Astigmatisme
- Courbure du plan focal
- Distorsion

Couleur

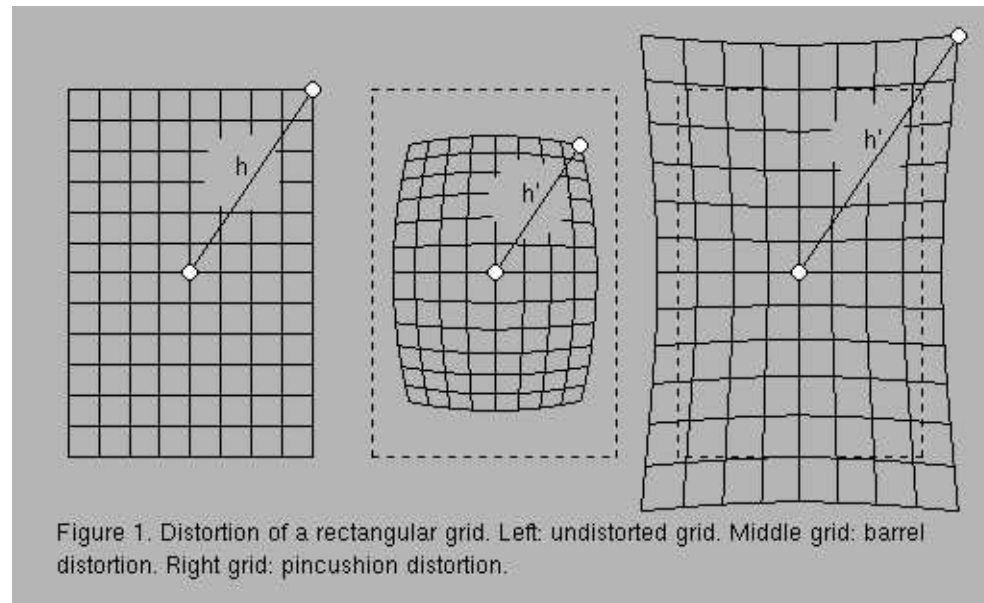
- Aberrations chromatiques

Autres effets

- Vignettage
- Reflets (Flare)
- Diffraction
- Profondeur de champs

Distorsion

Distorsion qui rend courbes les lignes droites.



Distorsion

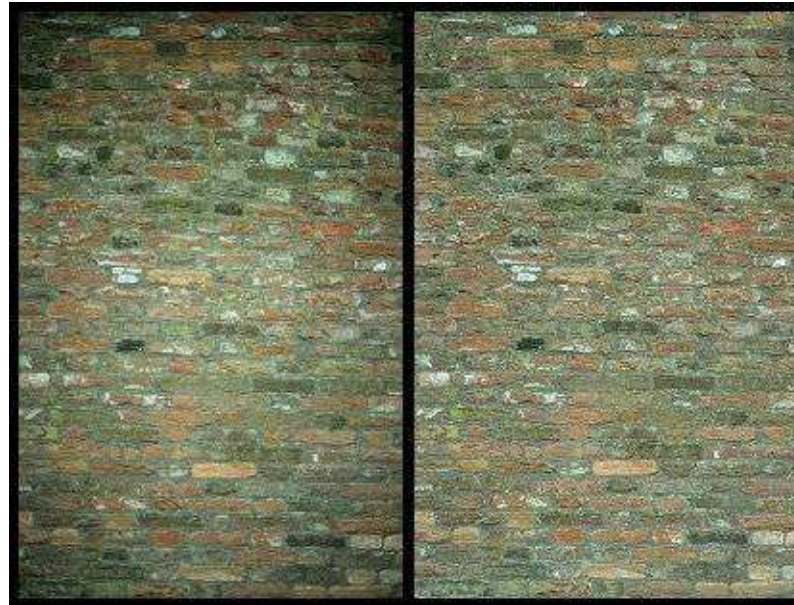
Voici un exemple de distorsion (à gauche). Ne pas confondre avec la perspective !!! (à droite).



La distorsion perspective dépend du point de vue seulement et de la profondeur des objets.

Vignettage

Diminution graduelle de l'intensité lumineuse vers les bords de l'image. Causes : naturelles ou mécaniques.



Vignettage mécanique

La lentille cause une obstruction...

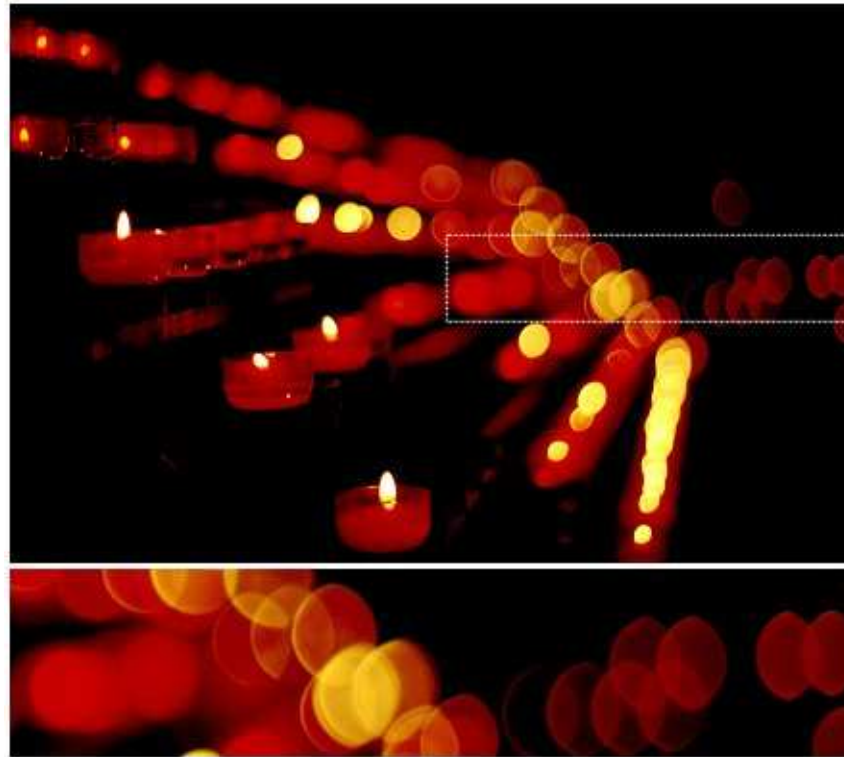


Vignettage mécanique

Lorsqu'on est sur le bord de l'images, la profondeur de la lentille affecte la lumière qui passe....

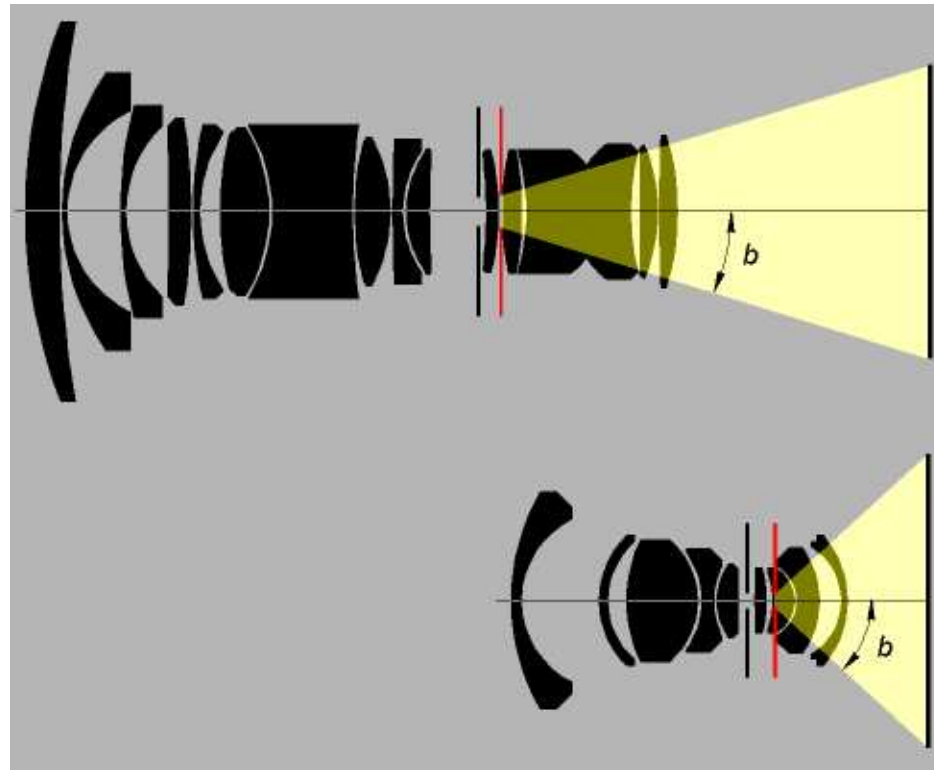


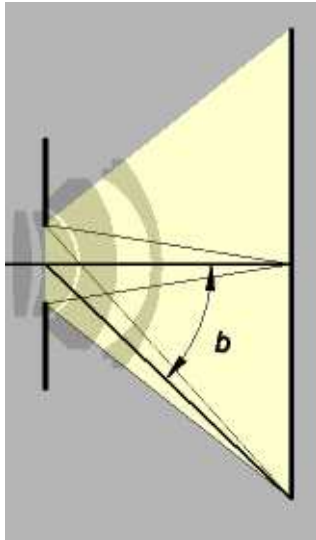
Il en résulte un effet d'oeil de chat sur les cotés de l'image.



Vignettage naturel

C'est la loi du \cos^4 .

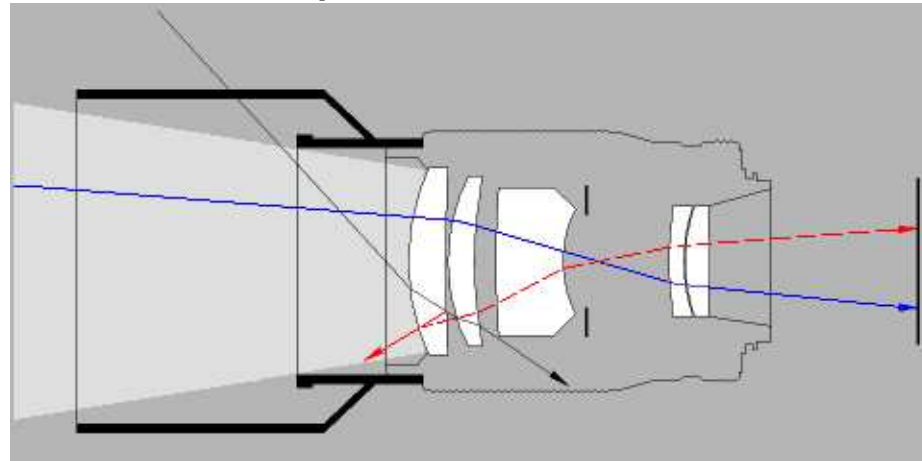


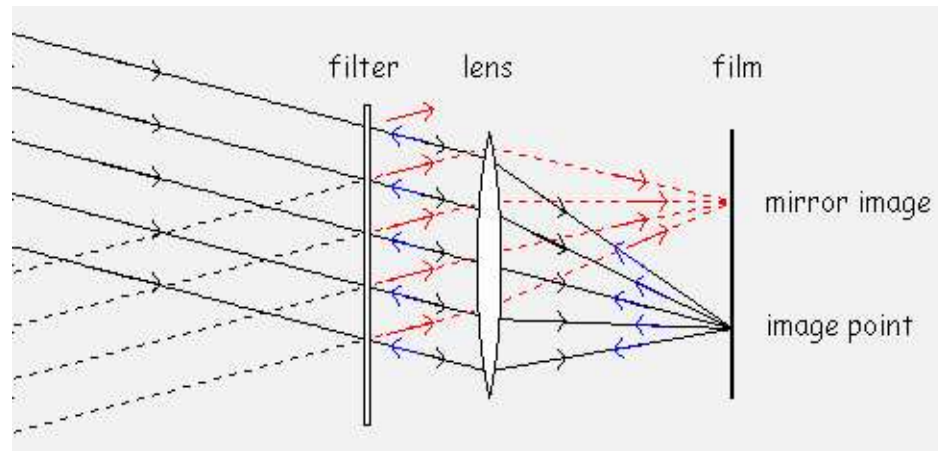


- $\cos^2(b)$: La loi de l'inverse du carré de la distance.
- $\cos(b)$: La pupille devient elliptique lorsque vue de côté.
- $\cos(b)$: La lumière frappe l'image à angle b (de côté).

Reflets

Les reflets sont causés par de lumière indésirables....





Profondeur de champs

La netteté de l'image est très importante.

Définition

Intervalle de profondeur sur lequel les objets sont projetés sur l'image avec une netteté suffisante.

Critère de netteté...



Formation de l'image numérique

Pour une image numérique :

- Propriétés physiques du senseur
- Nature discrète des des récepteurs
- Quantification et courbe de réponse de l'intensité

Radiométrie

Irradiance d'une image

Puissance totale de la lumière captée (par unité de surface) à un chaque point p d'une image

Radiance d'une scène

Puissance (par unité de surface) de la lumière émise par un point P d'une surface dans la direction d .

Réflectance

Réflectance d'une surface

Un modèle de réflectance détermine comment la lumière reflète sur une surface.

Exemple : Modèle Lambertien

Une surface est lambertienne si la lumière reflétée est indépendante de la direction d de l'observateur.

$$L = \rho I^\top n$$

ρ albedo

I direction de l'illumination

n normale de la surface

Réflectance

Note sur la surface lambertienne...

Si l'on change de point de vue vers une surface plus inclinée...

- L'observateur voit une plus grande partie de surface
- La surface reflète moins de lumière à cause de l'angle de vue

Les deux effets s'annulent → invariance au point de vue.

Géométrie

Associer un point 2D de l'image à un point 3D de la scène.

$$(x, y) = g(X, Y, Z)$$

1. Perspective

Modèle le plus réaliste pour les caméras photographiques

2. Perspective-Faible

(Weak-Perspective, Scaled-Orthographic)

Hypothèse : perspective avec les objets loin de la caméra

3. Orthographique

Modèle réaliste pour certaines caméras (rayon-X)

4. Affine

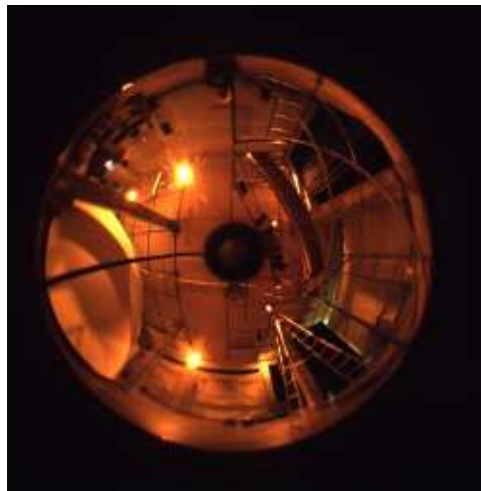
Modèle mathématiquement élégant mais physiquement irréaliste

Géométrie

$$(x, y) = g(X, Y, Z)$$

La fonction g peut être non-linéaire.

- Distorsion radiale
- Caméra catadioptriques (lentilles + miroir courbé)



Pourquoi tant de modèles ?

Chaque hypothèse simplificatrice diminue le réalisme du modèle mais simplifie le traitement mathématique.

Ne pas oublier...

En vision, on veut récupérer les paramètres du modèle de caméra.

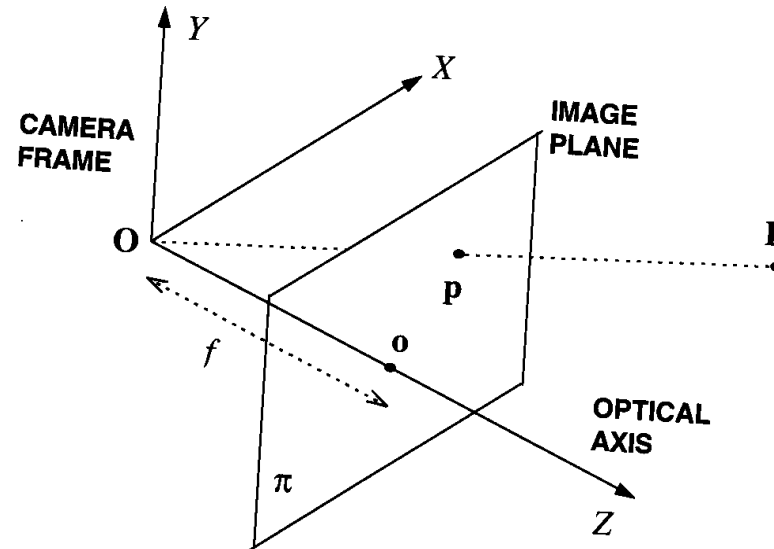
On doit donc trouver un compromis :

modèle très précis impossible à résoudre

VS

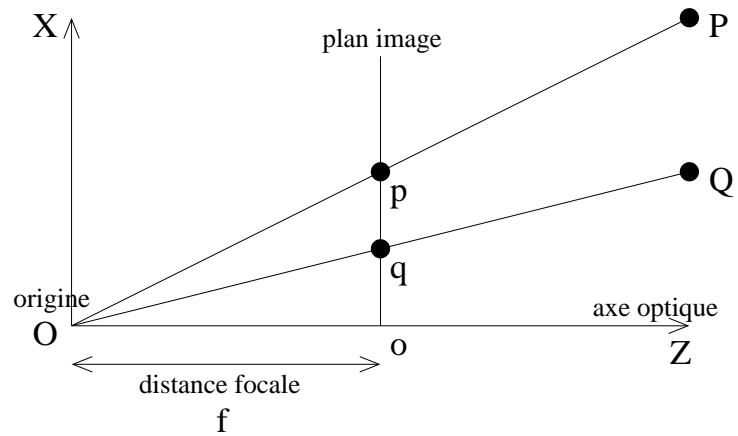
modèle facile à résoudre mais trop simpliste

Caméra Perspective



Plan de projection π , l'origine O . La distance entre π et O est la distance focale f . Le point P de la scène se projète au point p de l'image.

En équation...



$$x = f \frac{X}{Z} \quad y = f \frac{Y}{Z}$$

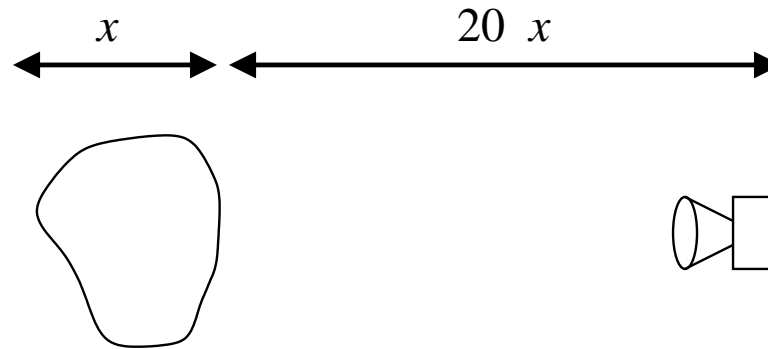
avec $\mathbf{p} = [x \quad y \quad f]^\top$ et $\mathbf{P} = [X \quad Y \quad Z]^\top$.

Caméra Perspective-Faible

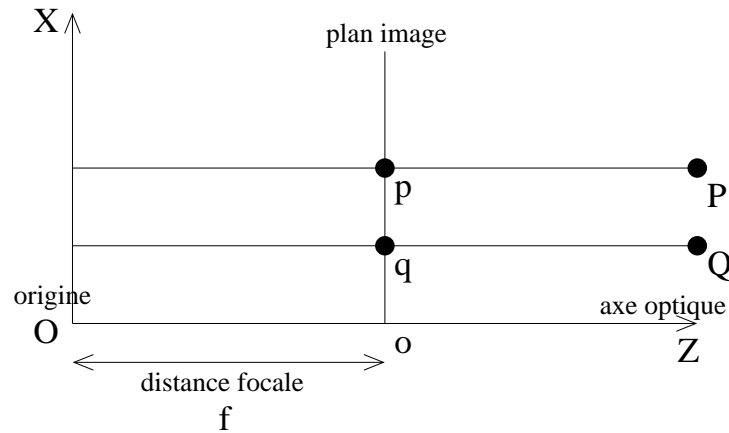
$$x = f \frac{X}{\bar{Z}} = \frac{f}{\bar{Z}} X \quad y = f \frac{Y}{\bar{Z}} = \frac{f}{\bar{Z}} Y$$

avec $\bar{Z} = \sum Z_i$, la profondeur moyenne de la scène.

- Cette transformation équivaut à une projection orthographique suivie d'un changement d'échelle d'un facteur f/\bar{Z} .



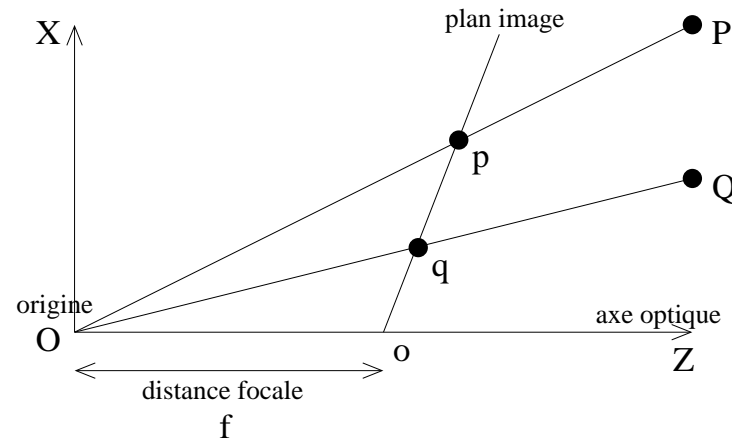
Caméra Orthographique



$$x = X \quad y = Y$$

- très simple
- pas toujours très réaliste

Caméra Affine



Modèle de projection perspective généralisée.

Intuitivement

Prenez une caméra normale, changez la position (et la forme!) du film arbitrairement (enfin presque).

- Mathématiquement élégant
- Numériquement simple
- Physiquement irréaliste

Numérisation

Le plan image doit être *numérisé*.

- CCD, senseurs CMOS, convertisseur D/A (Digital/Analogue), etc...
- Échantillonnage, filtrage, AGC (automatic gain control), AWB (automatic white balance), Anti-Blooming, ...

Fréquences dans l'image

Il y a une limite aux fréquences qu'on peut représenter dans une image.

Limite de diffraction et limite de résolution (Théorème d'échantillonnage).

Bruit dans l'image

Phénomène très **important**.

Le bruit affecte tous les algorithmes de vision.

Comment s'en sortir? → Un pixel est une variable aléatoire!!!

Comment estimer le bruit ?

Prendre de nombreuses images de la même scène, puis calculer la variance de chaque point.

Mais il y aura d'autres façons...

Systemes de coordonnees

Trois systemes de coordonnees distincts :

Le monde (ou la scene)

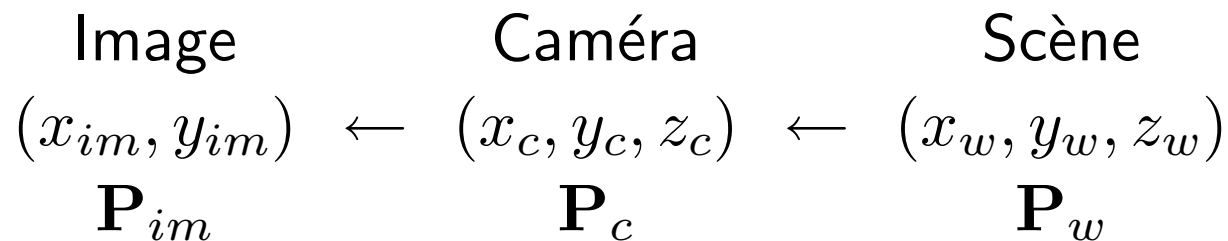
Systeme de reference connu. Tout (camera, objet, etc...) est defini par rapport a ce systeme de coordonnees.

La camera

Systeme dans lequel s'effectue la projection des points 3D vers 2D.

L'image

Systeme etablissant la reference aux pixels.



Paramètres de caméra

Les paramètres du modèle de caméra décrivent la transformation du système de coordonnées de la scène vers celui de l'image.



On sépare les paramètres en deux grandes catégories :

Paramètres externes (*extrinsèque, extrinsic*)

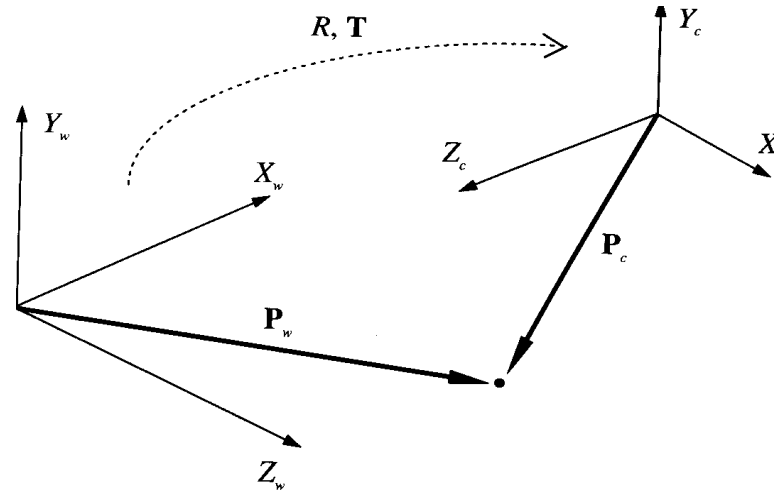
Position et orientation de la caméra par rapport à la scène.

Paramètres internes (*intrinsèque, intrinsic*)

Caractéristiques de la caméra elle-même.

Ces paramètres ne changent pas si on déplace la caméra.

Paramètres externes



La position et l'orientation sont une **transformation rigide** 3D :

- Translation \mathbf{T} pour faire correspondre les origines des systèmes de coordonnées.
- Rotation \mathbf{R} pour faire correspondre les axes des systèmes.

$$\mathbf{P}_c = \mathbf{R}(\mathbf{P}_w - \mathbf{T})$$

$$\mathbf{R} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{pmatrix} \quad \mathbf{T} = \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix}$$

\mathbf{R} est une matrice de rotation, donc orthonormale

$$\mathbf{R}^{-1} = \mathbf{R}^\top \quad \det(\mathbf{R}) = 1$$

(

Coordonnées homogènes

$$\mathbf{P}_c = \mathbf{R}(\mathbf{P}_w - \mathbf{T})$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} - \mathbf{RT}$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \left(\mathbf{R}_{(3 \times 3)} \mid -\mathbf{RT}_{(3 \times 1)} \right)_{(3 \times 4)} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$\mathbf{P}_c = (\mathbf{R} \mid -\mathbf{RT})\mathbf{P}'_w$$

$$\mathbf{P}_c = \mathbf{M}_{ext}\mathbf{P}'_w$$

Espaces projectifs

Le point \mathbf{P} euclidien est associé au point \mathbf{P}' dans l'espace projectif :

$$\mathbf{P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \mathbf{P}' = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}'$$

Par définition on a l'équivalence

$$\begin{aligned} \mathbf{P}' &\equiv h\mathbf{P}' && (h \neq 0) \\ \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}' &\equiv \begin{bmatrix} hx \\ hy \\ hz \\ h \end{bmatrix}' \end{aligned}$$

Espaces projectifs

Pour passer d'un à l'autre...

$$\text{euclidien} \rightarrow \text{projectif} : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow \begin{bmatrix} hx \\ hy \\ hz \\ h \end{bmatrix}' \quad (h \neq 0)$$

$$\text{projectif} \rightarrow \text{euclidien} : \begin{bmatrix} x \\ y \\ z \\ h \end{bmatrix}' \rightarrow \begin{bmatrix} x/h \\ y/h \\ z/h \end{bmatrix}$$

Dimensions d'un espace projectif

Espace projectif à deux dimensions :

$$\begin{bmatrix} x \\ y \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}' \quad \begin{bmatrix} x \\ y \\ h \end{bmatrix}' \rightarrow \begin{bmatrix} x/h \\ y/h \end{bmatrix}$$

Espace projectif à trois dimensions :

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}' \quad \begin{bmatrix} x \\ y \\ z \\ h \end{bmatrix}' \rightarrow \begin{bmatrix} x/h \\ y/h \\ z/h \end{bmatrix}$$

Addition projective

Addition de deux points projectifs

$$\begin{bmatrix} x_1 \\ y_1 \\ h_1 \end{bmatrix}' + \begin{bmatrix} x_2 \\ y_2 \\ h_2 \end{bmatrix}' = \begin{bmatrix} h_2x_1 + h_1x_2 \\ h_2y_1 + h_1y_2 \\ h_1h_2 \end{bmatrix}'$$

$$\begin{bmatrix} x_1 \\ y_1 \\ h \end{bmatrix}' + \begin{bmatrix} x_2 \\ y_2 \\ h \end{bmatrix}' = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ h \end{bmatrix}'$$

Équivalent euclidien

$$\begin{bmatrix} x_1/h_1 \\ y_1/h_1 \end{bmatrix} + \begin{bmatrix} x_2/h_2 \\ y_2/h_2 \end{bmatrix} = \begin{bmatrix} \frac{h_2x_1 + h_1x_2}{h_1h_2} \\ \frac{h_2y_1 + h_1y_2}{h_1h_2} \end{bmatrix}$$

Transformations

Transformation affine d'un point euclidien 2D

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Équivalent projectif

$$\begin{bmatrix} hu \\ hv \\ h \end{bmatrix}' = \begin{pmatrix} m_{11} & m_{12} & 0 \\ m_{21} & m_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} hx \\ hy \\ h \end{bmatrix}'$$

Addition euclidienne

Addition de deux points euclidiens 2D

$$\begin{bmatrix} x + a \\ y + b \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix}$$

Équivalent projectif

$$\begin{bmatrix} h(x + a) \\ h(y + b) \\ h \end{bmatrix}' = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} hx \\ hy \\ h \end{bmatrix}'$$

Projection

Projection du point \mathbf{P} de la scène 3D vers le point \mathbf{p} du plan image.

$$\mathbf{P} = [X \ Y \ Z]^\top \quad \mathbf{p} = [x \ y]^\top$$

Euclidien :

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} X/Z \\ Y/Z \end{bmatrix}$$

Projectif :

$$\mathbf{p}' = \mathbf{JP}'$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}'$$

Espaces projectifs

Avantage incroyable

La projection, opération non-linéaire dans l'espace euclidien, est une opération linéaire dans l'espace projectif.

Et ce n'est pas tout...

On peut même représenter l'infini.

L'infini dans l'espace projectif

Un point à l'infini suivant l'axe des x

$$\begin{bmatrix} \infty \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}'$$

On peut donc représenter l'infini *dirigé*.

La ligne à l'infini...

Il y a un point projectif qui est invariant sous une transformation affine : La ligne à l'infini \mathbf{L}_∞ .

$$\mathbf{L}_\infty = \mathbf{M}\mathbf{L}_\infty = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \mathbf{L}_\infty$$

Que vaut \mathbf{L}_∞ ?

$$\mathbf{L}_\infty = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

)

Paramètres internes

Relation entre la caméra et l'image :

Centre de l'image (o_x, o_y) , ratio entre les unités du monde et les pixels $(f/s_x, f/s_y)$,

$$\begin{bmatrix} x_{im} \\ y_{im} \end{bmatrix} = \begin{bmatrix} x_c \frac{f}{s_x} \\ y_c \frac{f}{s_y} \end{bmatrix} + \begin{bmatrix} o_x \\ o_y \end{bmatrix}$$
$$\begin{bmatrix} x_{im} \\ y_{im} \\ 1 \end{bmatrix} = \begin{pmatrix} \frac{f}{s_x} & 0 & o_x \\ 0 & \frac{f}{s_y} & o_y \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix}$$
$$\mathbf{P}_{im} = \mathbf{M}_{int} \mathbf{P}_c$$

Modèle Perspective

$$\begin{bmatrix} u \\ v \\ h \end{bmatrix}' = \mathbf{M}_{int} \mathbf{M}_{ext} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}'$$

avec

$$\mathbf{M}_{int} = \begin{pmatrix} \frac{f}{s_x} & 0 & o_x \\ 0 & \frac{f}{s_y} & o_y \\ 0 & 0 & 1 \end{pmatrix} \quad \mathbf{M}_{ext} = (\mathbf{R} \mid \mathbf{RT})$$

- Modèle Perspective faible ?
- Modèle Orthographique ?
- Modèle Affine ?

Résumé des transformations projectives

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{bmatrix}$$

projective (15 inconnues)

$$\begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

affine (12 inconnues)

$$\begin{bmatrix} \sigma r_{11} & \sigma r_{12} & \sigma r_{13} & t_1 \\ \sigma r_{21} & \sigma r_{22} & \sigma r_{23} & t_2 \\ \sigma r_{31} & \sigma r_{32} & \sigma r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

métrique (7 inconnues)

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

euclidienne (6 inconnues)

- $r_{11} \dots r_{33}$ représente une matrice de rotation 3×3 .
- (t_1, t_2, t_3) est un vecteur représentant une translation.
- σ est un facteur d'échelle

Types de transformations projectives

