

IFT 6145
Vision tridimensionnelle
Modèles de caméra II

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

hiver 2004

En bref...

- Paramètres de caméra : suite et fin
 - Paramètres externes
 - Paramètres internes
 - Modèles de caméra
- TP0 : Modèle de Caméra et Intro à Mathematica

Systemes de coordonnees

Trois systemes de coordonnees distincts :

Le monde (ou la scene)

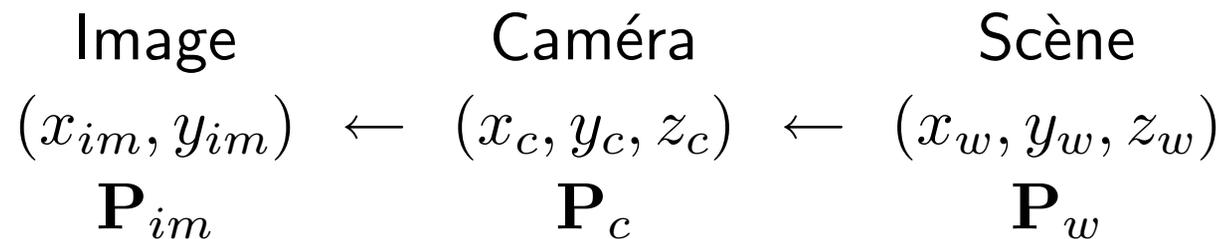
Systeme de reference connu. Tout (camera, objet, etc...) est defini par rapport a ce systeme de coordonnees.

La camera

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

L'image

Systeme etablissant le concept de pixel.



Modèle de caméra comme transformation projective

On va relier les différents système par une transformation linéaire dans l'espace projectif.

$$\mathbf{P}'_{im} = \mathbf{M} \mathbf{P}'_w$$

Pourquoi ?

- Système linéaire \Rightarrow simple à résoudre
- Capable de bien représenter la réalité
- Simple à implanter en hardware...

Paramètres de caméra

Les paramètres du modèle de caméra décrivent la transformation entre les systèmes de coordonnées.



On sépare les paramètres en deux grandes catégories :

Paramètres externes

Position et orientation de la caméra par rapport à la scène.
Relation entre le système *Caméra* et le système *Scène*.

Paramètres internes

Caractéristiques de la caméra elle-même.
Ces paramètres ne changent pas si on déplace la caméra.
Relation entre le système *Image* et le système *Caméra*.

Paramètres de caméra

- Paramètres externes

$$\begin{aligned}\mathbf{P}'_c &= \mathbf{M}_{ext} \mathbf{P}'_w \\ &= \mathbf{R} \mathbf{T} \mathbf{P}'_w\end{aligned}$$

avec rotation \mathbf{R} et translation \mathbf{T} .

- Paramètres internes

$$\begin{aligned}\mathbf{P}'_{im} &= \mathbf{J} \mathbf{M}_{int} \mathbf{P}'_c \\ &= \mathbf{J} \mathbf{A} \mathbf{P}'_c\end{aligned}$$

avec transformation d'image \mathbf{A} et projection \mathbf{J} .

- Tous les paramètres combinés

$$\mathbf{P}'_{im} = \mathbf{J} \mathbf{M} \mathbf{P}'_w = \mathbf{J} \mathbf{M}_{int} \mathbf{M}_{ext} \mathbf{P}'_w$$

Rotation \mathbf{R}

Rotation autour de θ degrés autour d'un axe \mathbf{n}

$$\mathbf{R} = \begin{pmatrix} \mathbf{r} & \mathbf{0} \\ \mathbf{0}^\top & 1 \end{pmatrix}$$

avec

$$\mathbf{r} = \mathbf{I} \cos \theta + \mathbf{n} \mathbf{n}^\top (1 - \cos \theta) + [\mathbf{n}]_\times \sin \theta$$

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$[(n_1, n_2, n_3)]_\times = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix}$$

Rotation \mathbf{R}

Comment comparer deux rotation \mathbf{R}_0 et \mathbf{R}_1 ?

$$\|\log(\mathbf{R}_0 \mathbf{R}_1^{-1})\|$$

ou encore

$$3 - \text{Tr}((\mathbf{R}_0 \mathbf{R}_1^{-1}))$$

ou encore

...

Translation \mathbf{T}

Pour une translation $\mathbf{t} = (t_x, t_y, t_z)$, on a une matrice de translation

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Comment comparer t_0 et t_1 ?

$$\|t_0 - t_1\|$$

Transformation d'image \mathbf{A}

\mathbf{A} représente une transformation du plan image.

Paramètres : Centre de l'image (o_x, o_y) , taille d'un pixel (s_x, s_y) et distance focale f .

$$\mathbf{A} = \begin{pmatrix} \frac{f}{s_x} & 0 & o_x & 0 \\ 0 & \frac{f}{s_y} & o_y & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Plus généralement, simple transformation affine de l'image :

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 & a_3 & 0 \\ a_4 & a_5 & a_6 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Transformation d'image \mathbf{A}

- \mathbf{A} est en général exprimée par une 3×3 *après* la projection, plutôt qu'une 4×4 *avant* la projection.
- Si a_1 et a_5 sont positifs, on a un plan de projection *en avant* de la caméra (négatif \Rightarrow en arrière de la caméra).
(Trucco & Verri utilise le négatif)

Image

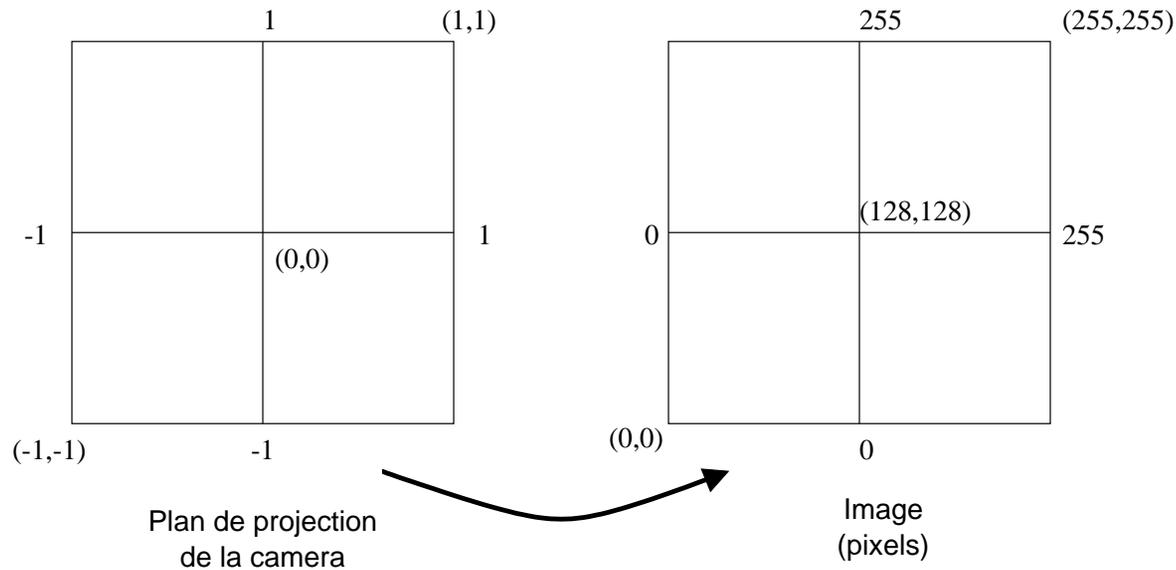


Image normalisée

Sur le plan de projection de la caméra, l'origine est au milieu de l'image et les unités sont celles du monde.

Image (pixels)

L'image exprime ce même plan de projection en pixels.

Image

L'intervale des coordonnées normalisées détermine l'angle de vue.

Si l'intervale est de $\pm k$, alors l'angle de vue est

$$\theta = 2 \tan^{-1}(k)$$

Exemples

- l'intervale ± 1 donne 90 degrés
- l'intervale ± 0.5 donne 53 degrés
- l'intervale ± 1.5 donne 112 degrés

Projection J

Dans l'espace projectif, la projection consiste à éliminer la dernière dimension.

$$\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Modèle Perspective

$$\begin{bmatrix} u \\ v \\ h \end{bmatrix}' = \mathbf{JART} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}'$$

avec

$$\mathbf{M}_{int} = \mathbf{A}$$

$$\mathbf{M}_{ext} = \mathbf{RT}$$

Lorsque $\mathbf{A} = \mathbf{R} = \mathbf{T} = \mathbf{I}$, on a

$$x = \frac{u}{h} = \frac{X}{Z} \quad y = \frac{v}{h} = \frac{Y}{Z}$$

Modèle Perspective faible

$$\begin{bmatrix} u \\ v \\ h \end{bmatrix}' = \mathbf{JA} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \bar{Z} \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{RT} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}'$$

avec $\bar{Z} = \frac{1}{n} \sum Z_i$, la profondeur moyenne de la scène.

Lorsque $\mathbf{A} = \mathbf{R} = \mathbf{T} = \mathbf{I}$, on a

$$x = \frac{u}{h} = \frac{1}{\bar{Z}} X \quad y = \frac{v}{h} = \frac{1}{\bar{Z}} Y$$

Modèle Orthographique

$$\begin{bmatrix} u \\ v \\ h \end{bmatrix}' = \mathbf{JA} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{RT} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}'$$

Lorsque $\mathbf{A} = \mathbf{R} = \mathbf{T} = \mathbf{I}$, on a

$$x = \frac{u}{h} = X \quad y = \frac{v}{h} = Y$$

Modèle Affine

$$\begin{bmatrix} u \\ v \\ h \end{bmatrix}' = \mathbf{J} \begin{pmatrix} m_1 & m_2 & m_3 & m_4 \\ m_5 & m_6 & m_7 & m_8 \\ m_9 & m_{10} & m_{11} & m_{12} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}'$$

- Modèle de projection perspective généralisée.
- Mathématiquement élégant, numériquement simple.
- Pas très réaliste...
- Aucune structure comparable à **ART**
- Moins de structure \Rightarrow moins d'hypothèses
 - Plan de projection perpendiculaire à l'axe optique
 - Image rectangulaire
 - ...

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)

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

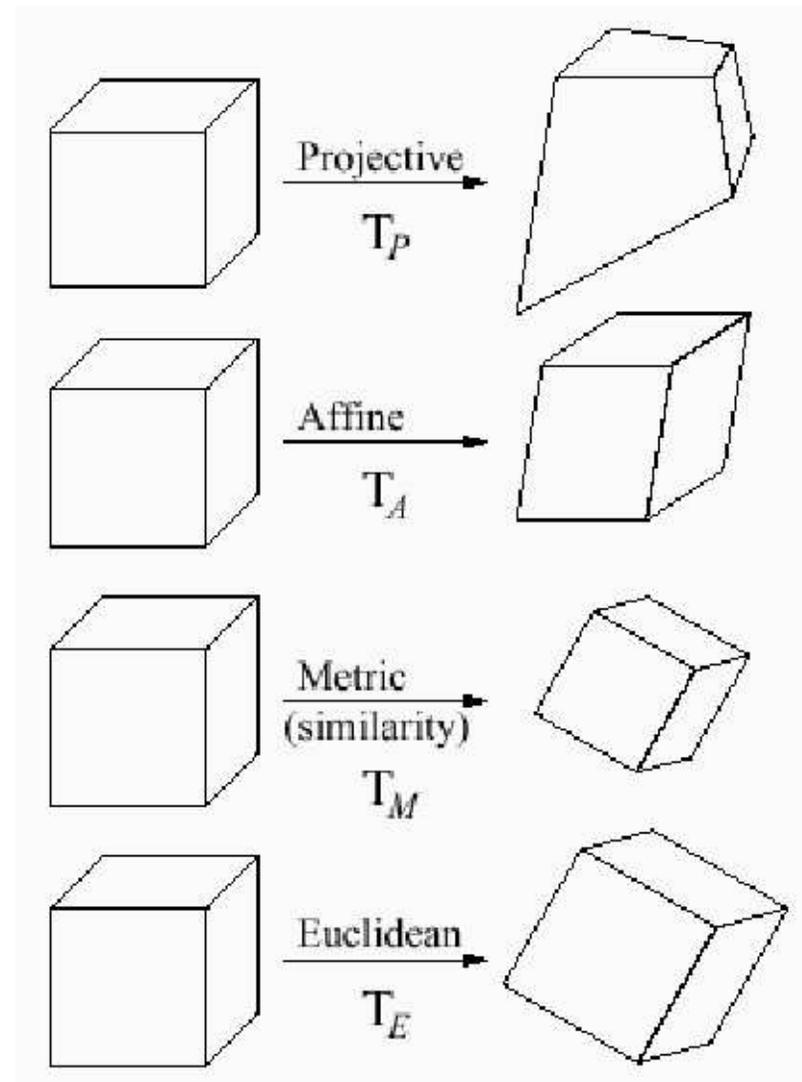
métrique (7 inconnues)

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

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



En pratique...

TP0

Travail pratique # 0 Introduction à Mathematica Modèles de caméra

Problème à résoudre

Soit une matrice \mathbf{M} représentant un modèle de caméra.

Illustrer graphiquement cette caméra, en incluant

- les axes x, y, z de la scène
- quelques points 3D
- la projection de ces points sur le plan de projection (l'image)
- un polygone représentant l'image

Illustrer plusieurs exemples de géométrie de caméra et plusieurs modèles de caméra.