

Formation et Analyse d'Images

James L. Crowley

ENSIMAG 3 - Option IRV

Premier Bimestre 2006/2007

Séance 2

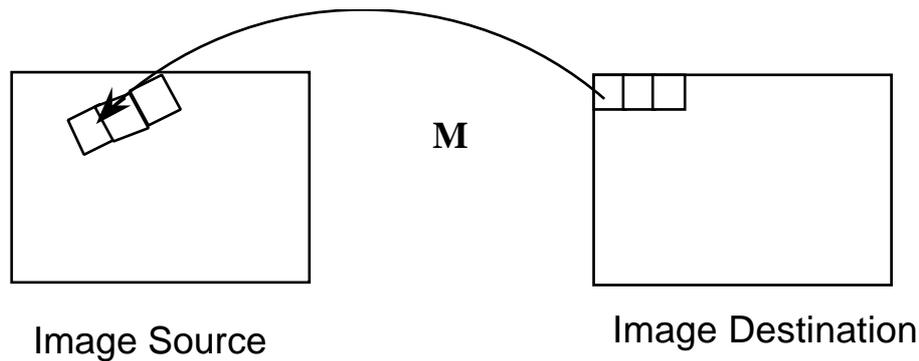
29 septembre 2006

Plan de la Séance :

Modèle de la Caméra

Transformations d'images	2
Translation	3
Rotation.....	3
Translation et Rotation.....	3
Transformations d'échelle, rotation et translation.....	4
Modèle de la Caméra	5
Les Repères	5
Transformations entres reperes.....	6
La Transformation Scène - Caméra.....	6
La Projection Caméra - Rétine (Le modèle Stenope)	8
La transformation Rétine - Image.....	10
Les paramètres Intrinsèques de la caméra.....	10
Echantillonnage et Numérisation.....	10
La Composition de la Projection Scène - Image	12

Transformations d'images



Pour chaque pixel de l'image de destination (x_d, y_d) , on calcule une position (x_s, y_s) dans l'image de la source.

$$P^s = M_d^s P^d$$

$$\text{ou } \begin{matrix} p^1 \\ p^2 \\ p^3 \end{matrix} = \begin{matrix} m_1^1 & m_2^1 & m_3^1 \\ m_1^2 & m_2^2 & m_3^2 \\ m_1^3 & m_2^3 & m_3^3 \end{matrix} \begin{matrix} x_d \\ y_d \\ 1 \end{matrix}$$

$$x_s = \frac{p^1}{p^3} \quad y_s = \frac{p^2}{p^3}$$

$$\text{ou } \begin{matrix} x_s \\ y_s \\ 1 \end{matrix} = \begin{matrix} p^1/p^3 \\ p^2/p^3 \\ 1 \end{matrix}$$

Ensuite la nouvelle valeur de pixel destination, P^d , est calculé en fonction du voisinage de la position source.

MAIS, $P^s = \begin{matrix} x_s \\ y_s \\ 1 \end{matrix}$ n'est pas des entier!

Quelle valeur faut-il prendre pour les pixels?

Pour chaque pixel du destination, (x_d, y_d) on calcul le position du source, (x_s, y_s) . Ensuite, on détermine une valeur par interpolation avec les pixels voisins.

Translation

$$\begin{aligned}x_2 &= x_1 + t_x, \\y_2 &= y_1 + t_y\end{aligned}$$

$$\begin{array}{rcccl}x_2 & & 0 & 0 & t_x & x_1 \\y_2 & = & 0 & 0 & t_y & y_1 \\1 & & 0 & 0 & 1 & 1\end{array}$$

En notation tensorielle :

$$P^B = T_A^B P^A \quad \text{pour } A, B = 1, 2, 3.$$

Donc T_A^B est une transformation du repère A vers le repère B.

Les indices permet de noter les repères.

Rotation

(Repère main droite, rotation sens trigonométrique)

$$\begin{aligned}x_2 &= \text{Cos}(\theta) x_1 + \text{Sin}(\theta) y_1, \\y_2 &= \text{Sin}(\theta) x_1 - \text{Cos}(\theta) y_1\end{aligned}$$

$$\begin{array}{rcccl}x_2 & & \text{Cos}(\theta) & \text{Sin}(\theta) & 0 & x_1 \\y_2 & = & -\text{Sin}(\theta) & \text{Cos}(\theta) & 0 & y_1 \\1 & & 0 & 0 & 1 & 1\end{array}$$

Quand le repère tourne dans le sens " ", le vecteur est tourner dans le sens –

Translation et Rotation

$$\begin{aligned}x_2 &= \text{Cos}(\theta) x_1 + \text{Sin}(\theta) y_1 + t_x \\y_2 &= \text{Sin}(\theta) x_1 - \text{Cos}(\theta) y_1 + t_y\end{aligned}$$

$$\begin{array}{rcccl}x_2 & & \text{Cos}(\theta) & \text{Sin}(\theta) & t_x & x_1 \\y_2 & = & -\text{Sin}(\theta) & \text{Cos}(\theta) & t_y & y_1 \\1 & & 0 & 0 & 1 & 1\end{array}$$

En tensorielle, $P^B = R_A^B P^A$

Transformations d'échelle, rotation et translation

Une transformation de similitude d'une image est définie par une rotation, une translation, et un changement de taille des axes.

Soit un changement d'échelle s_x et s_y des axes x_1 et y_1 (repère source) suivi d'une rotation d'angle θ dans le plan de l'image source, suivi d'une translation t_x, t_y s'exprimes dans le repère de la destination.

Ces paramètres donne une transformation $(s_x, s_y, \theta, t_x, t_y)$ de

- 1) Un changement d'échelle des axes, puis
- 2) Une rotation des axes, puis
- 3) Une translation.

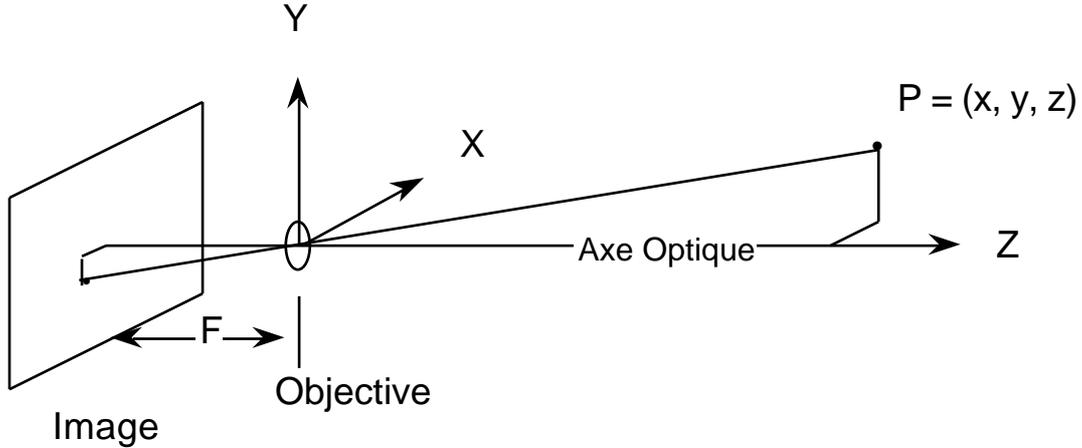
$$\begin{array}{rcccl} x_2 & = & s_x \cos(\theta) & s_y \sin(\theta) & t_x & x_1 \\ y_2 & = & -s_x \sin(\theta) & s_y \cos(\theta) & t_y & y_1 \\ 1 & & 0 & 0 & 1 & 1 \end{array}$$

ou bien

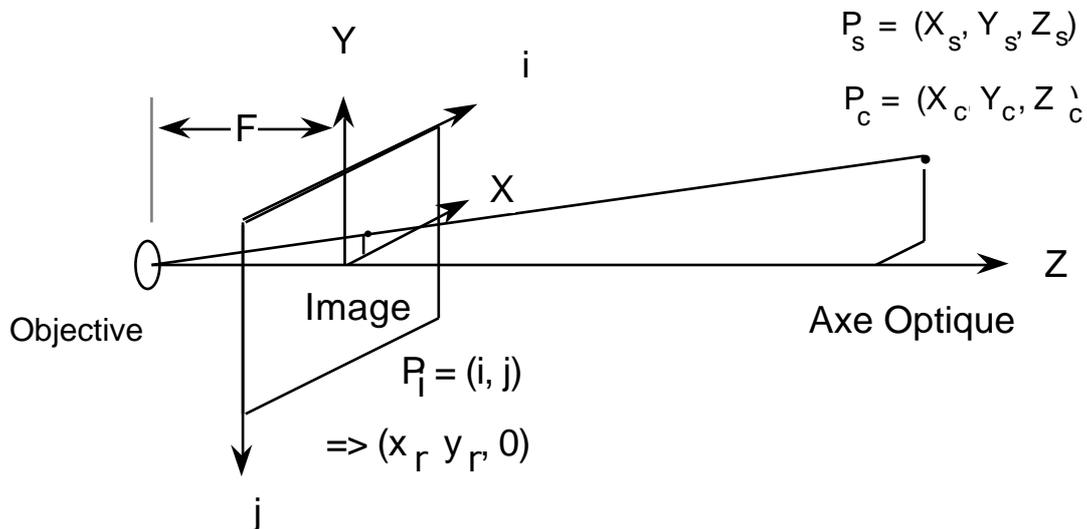
$$\begin{aligned} x_2 &= s_x \cos(\theta) x_1 + s_y \sin(\theta) y_1 + t_x, \\ y_2 &= s_y \sin(\theta) x_1 - s_x \cos(\theta) y_1 + t_y \end{aligned}$$

Modèle de la Caméra

Modèle de la Caméra Géométrie Physique



Modèle Mathématique: Projection Centrale



Les Repères

Coordonnées de la Scène :

Point Scène : $P^s = (x_s, y_s, z_s, 1)^T$

Coordonnées de la Caméra :

Point Caméra : $P^c = (x_c, y_c, z_c, 1)^T$

Point Image : $P^r = (x_r, y_r, 1)^T$

Coordonnées de l'Image :

Point Image : $P^i = (i, j, 1)^T$

Nota : En coordonnées homogènes, les points sont invariant au multiplication par un constant.

$$(i, j, 1)^T = A \cdot (i, j, 1)^T = (Ai, Aj, A)^T$$

Transformations entres reperes

La modèle de la caméra est composé d'une composition de transformations.

Transformation entre repère Caméra et repère Scène :

$$P^c = T_s^c P^s$$

Matrice de Projection ${}^r P$ du repère Caméra vers le repère Rétine

$$P^r = P_c^r P^c$$

Transformation entre repère image et repère caméra

$$P^i = C_r^i \circ P^r$$

Composition : $P^i = C_r^i \circ P_c^r \circ T_s^c \circ P^s = M_s^i \circ P^s$

La Transformation Scène - Caméra

Dans les coordonnées homogènes, les translations, rotations, transformées affines et perspectives sont représentées par les produits des matrices.

La transformation T_s^c a la forme :

$$T_s^c = \begin{pmatrix} \mathbf{R}_s^c & x_s \\ 0 & y_s \\ 0 & z_s \\ 0 & 1 \end{pmatrix}$$

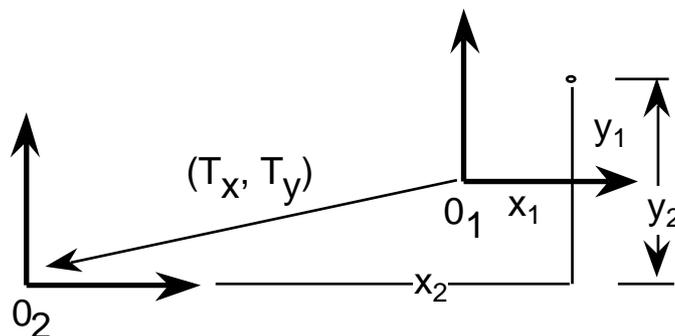
ou (x_s, y_s, z_s) est la position du repère scène dans le repere caméra.

et \mathbf{R}_s^c est l'orientation du repère scène dans le repère caméra.

Translation

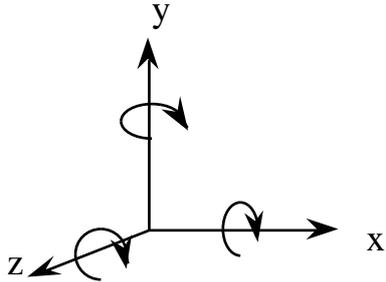
Addition en espace cartésien

Multiplications en Coordonnées Homogènes



T est un vecteur (vecteur de translation) de P_1 vers P_2

En 3D



Au tour de l'axe X :

$$\mathbf{R}_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & \sin(\alpha) & 0 \\ 0 & -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

l'axe Y :

$$\mathbf{R}_y(\alpha) = \begin{pmatrix} \cos(\alpha) & 0 & -\sin(\alpha) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

l'axe Z :

$$\mathbf{R}_z(\alpha) = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) & 0 & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

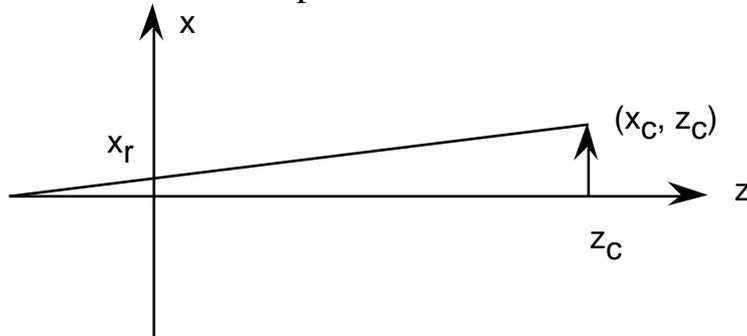
En Générale :

$$\mathbf{T} = \begin{pmatrix} \mathbf{R} & \begin{matrix} T_x \\ T_y \\ T_z \end{matrix} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{ou } \mathbf{R} = \mathbf{R}_z(\alpha) \mathbf{R}_y(\beta) \mathbf{R}_x(\gamma)$$

La Projection Caméra - Rétine (Le modèle Stenope)

La projection de la caméra vers la rétine est une projection perspective.

Considère une caméra 1-D dans un plan 2-D.



Dans le repère de la caméra :

$(x_c, y_c, z_c, 1)$: Point dans la scène en repère caméra

$(x_r, y_r, 1)$: Point dans la rétine en repère rétine.

Par triangles semblables :

$$\frac{x_r}{F} = \frac{x_c}{z_c} \quad x_r = x_c \frac{F}{z_c} \quad x_r \frac{z_c}{F} = x_c$$

$$\frac{y_r}{F} = \frac{y_c}{z_c} \quad y_r = y_c \frac{F}{z_c} \quad y_r \frac{z_c}{F} = y_c$$

soit : $w = \frac{z_c}{F}$

alors : $w x_r = x_c$

$w y_r = y_c$

$w = \frac{z_c}{F}$

En matrice :

$$\begin{array}{l} w x_r \\ w y_r \\ w \end{array} = \begin{array}{ccccc} 1 & 0 & 0 & 0 & x_c \\ 0 & 1 & 0 & 0 & y_c \\ 0 & 0 & \frac{1}{F} & 0 & z_c \\ & & & & 1 \end{array}$$

La transformation du repère Caméra vers le repère Rétine est :

$$\mathbf{P}^r = \begin{array}{l} w x_r \\ w y_r \\ w \end{array} = \mathbf{P}_c^r \mathbf{P}^c \quad \text{tel que} \quad \begin{array}{l} x_r \\ y_r \\ 1 \end{array} = \begin{array}{l} p^1/p^3 \\ p^2/p^3 \\ 1 \end{array}$$

$$\text{et donc : } \mathbf{P}_c^r = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{F} & 0 \end{pmatrix}$$

Noter que \mathbf{P}_c^r n'est pas inversible .

Si l'origine est dans la rétine, par triangles semblables :

$$\frac{x_r}{F} = \frac{x_c}{(F + z_c)} \quad \Rightarrow \quad x_r = \frac{x_c F}{(F + z_c)}$$

Equations de perspective:

$$x_r = \frac{x_c F}{(F+z_c)} \qquad y_r = \frac{y_c F}{(F+z_c)}$$

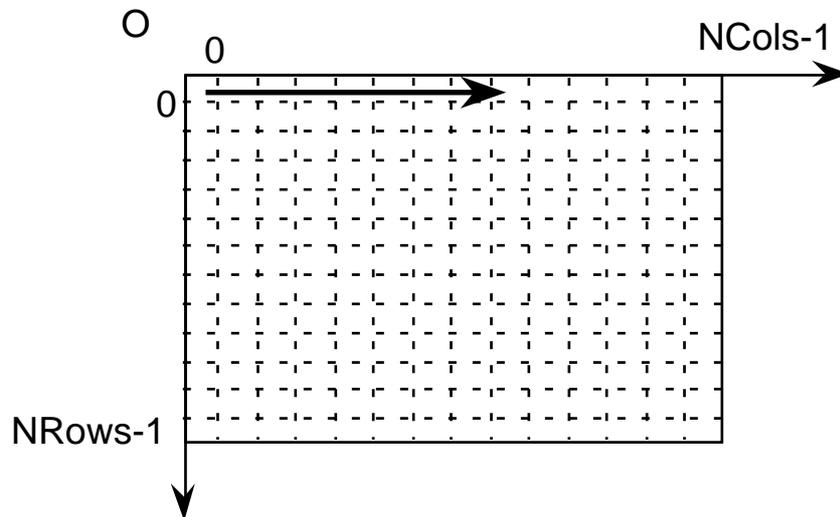
$$z_r = 0$$

$$\text{et puis : } \mathbf{P}_c^r = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{F} & 1 \end{pmatrix}$$

La transformation Rétine - Image

Les paramètres Intrinsèques de la caméra

La Trame : L'image est composée des "pixels" (picture elements)



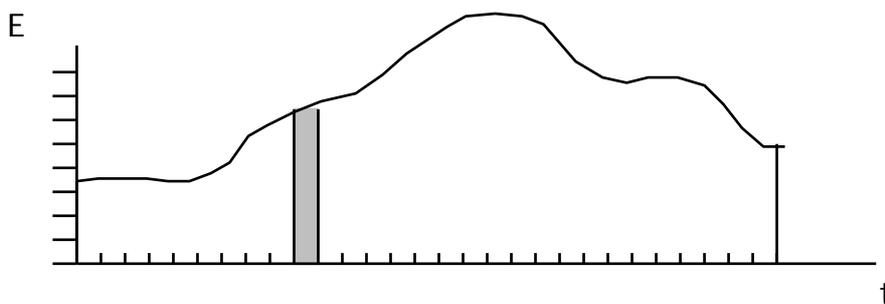
Question : Est-ce-que les pixels sont "carrés"?

Réponse : non.

Le nombre de pixels de l'image dépend de la matérielle.

Par exemple : VGA : 640 x 480

Echantillonnage et Numérisation



Les Paramètre Intrinsèque de la Caméra

F : Distance focale
 C_i, C_j : Centre Optique de l'Image (en pixels)
 D_i, D_j : Taille Physique d'un Pixel dans la Rétine (pixel/mm)

$$i = x_r D_i \text{ (mm} \cdot \text{pixel/mm)} + C_i \text{ (pixel)}$$

$$j = y_r D_j \text{ (mm} \cdot \text{pixel/mm)} + C_j \text{ (pixel)}$$

Transformation entre repère de l'image et repère de la rétine :

$$P^i = C_r^i P^r$$

$$\begin{array}{l} i \\ j \\ 1 \end{array} = \begin{array}{cccc} D_i & 0 & C_i & x_r \\ 0 & D_j & C_j & y_r \\ 0 & 0 & 1 & 1 \end{array}$$

ou bien:

$$\begin{array}{l} w_i \\ w_j \\ w \end{array} = \begin{array}{cccc} D_i & 0 & C_i & w x_r \\ 0 & D_j & C_j & w y_r \\ 0 & 0 & 1 & w \end{array}$$

La Composition de la Projection Scène - Image

$$\mathbf{P}^i = \mathbf{C}_r^i \mathbf{P}_c^r \mathbf{T}_s^c \mathbf{P}^s = \mathbf{M}_s^i \mathbf{P}^s$$

$$\begin{pmatrix} w \ i \\ w \ j \\ w \end{pmatrix} = \mathbf{M}_s^i \begin{pmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{pmatrix}$$

et donc

$$i = \frac{w \ i}{w} = \frac{M_{s1}^1 \cdot P^s}{M_{s3}^1 \cdot P^s} \qquad j = \frac{w \ j}{w} = \frac{M_{s1}^2 \cdot P^s}{M_{s3}^2 \cdot P^s}$$

ou bien

$$i = \frac{w \ i}{w} = \frac{M_{11} X_s + M_{12} Y_s + M_{13} Z_s + M_{14}}{M_{31} X_s + M_{32} Y_s + M_{33} Z_s + M_{34}}$$

$$j = \frac{w \ j}{w} = \frac{M_{21} X_s + M_{22} Y_s + M_{23} Z_s + M_{24}}{M_{31} X_s + M_{32} Y_s + M_{33} Z_s + M_{34}}$$