Vision par Ordinateur

James L. Crowley

M2R IVR — Premier Bimestre 2005/2006

Séance 1 — 19 Octobre 2005

Coordonnées et Transformations Homogènes

Plan de la Séance :

Coordonnées Homogènes en Notation Tensorielle

- Vecteur et Matrice en Notation Tensorielle
- L'équation d'une droite
- Le produit Croisé
- Intersection de deux droites

Transformations en Coordonnées Homogènes.

- Transformations d'images
- Interpolation Linéaire
- Interpolation Bi-lineaire
- Translation en Coordonnées Homogènes.
- Rotation
- Translation et Rotation en Coordonnées Homogènes.
- Transformations d'échelle, rotation et translation

Homographie entre deux plans

- Rectification de l'Image d'un plan.
- Estimation de l'homographie

Coordonnées Homogènes en Notation Tensorielle

Les coordonnées homogènes repose sur une notation dans laquelle les vecteurs en N dimensions sont représentées par un vecteur en N+1 dimensions.

Les coordonnées homogènes sont un outil de base en vision, en robotique et en synthèse d'images.

Exemple : un point

Soit un plan Euclidienne en R^2 composé de points,

En notation classique, un point est un vecteur : $\vec{P} = \begin{pmatrix} x \\ y \end{pmatrix}$

En notation homogène, un point est un vecteur $\vec{P} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$

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

$(i, j, 1)^T = a\cdot(i, j, 1)^T = (ai, aj, a)^T$

On note que a, b : $a \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = b \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$

Vecteur et Matrice en Notation Tensorielle

En notation tensorielle, la signe " $\vec{}$ " est remplacé par un indice en super-scripte ou sous-scripte. Une super-scripte signifie un vecteur colonne. Par exemple, le point est indiqué par un vecteur p^i

$$P^i = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix}$$

une sous-scripte indique un vecteur ligne

Par exemple, la droite est indiquée par le vecteur l_i :

$$L_i = (l_1, l_2, l_3)$$

Une matrice est une ligne de vecteurs (ou une vecteurs de lignes).

$$M^i_i = \begin{pmatrix} m^1_1 & m^1_2 & m^1_3 \\ m^2_1 & m^2_2 & m^2_3 \\ m^3_1 & m^3_2 & m^3_3 \end{pmatrix}$$

Le produit d'un vecteur avec une indice en super scripte et une vecteur avec le même indice en sous-scripte signifie une produit scalaire. Ceci est fait par une sommation implicite des indices. (La Convention de Sommation d'Einstein).

Donc, le produit indique une annulation des souscriptes et superscriptes.

$$L_i P^i = l_1 p^1 + l_2 p^2 + l_3 p^3$$

Cette sommation et commutative

$$L_i P^i = P^i L_i$$

Pour la produit d'une matrice et une vecteur, ceci donne une nouvelle vecteur.

$$P^j = M^j_i P^i$$

Ceci représent une tranformation du repère "i" vers le repère "j".

L'équation d'une droite

Dans un plan Euclidienne en R^2, en notation "classique", une droite est définie par une équation

$$a x + b y + x = 0.$$

On peut exprimer cette équation comme la produit de deux vecteurs :

$$\vec{L} \cdot \vec{P} = 0$$

ou $\vec{L} = (a \;\; b \;\; c)$ et $\vec{P} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$

En notation tensorielle, cette equation est exprimé :

$$L_i P^i = 0 \qquad \text{avec } i=1, 2, 3.$$

La sommation des indices est implicites.

Le produit Croisé

Une droite est définie par deux points. Un point est défini par le croisement de deux droites. Il y a une dualité parfaite entre les points et les droites.

En notation classique, $ax + by + c = 0$
le manière de déterminer la droite pour deux points est :

où $a = (y_1 - y_2)$ $b = (x_2 - x_1)$

$$c = -(a\, x_1 + b\, y_1) = -x_1(y_1 - y_2) - y_1(x_2 - x_1)$$
$$= -x_1 y_1 + x_1 y_2 - y_1 x_2 + y_1 x_1$$
$$= x_1 y_2 - y_1 x_2$$

La droite est $x\,(y_2 - y_1) + y\,(x_1 - x_2) - x_1 y_2 + y_1 x_2 = 0$

Ceci peut être calculé par la déterminante, avec les variables libres dans le premier colonne :

$$\begin{vmatrix} x & x_1 & x_2 \\ y & y_1 & y_2 \\ 1 & 1 & 1 \end{vmatrix} = x(y_1 - y_2) + x_1(y_2 - y) + x_2(y - y_1) = 0$$

$$= x(y_1 - y_2) + y(x_2 - x_1) + x_1 y_2 - x_2 y_1 = 0$$

Il s'agit d'une méthode générale de déterminer les paramètres d’une équation linéaire à partir des contraintes. Ca marche aussi pour trouver la point d'intersection de deux lignes.

On peut, également, écrire la déterminante comme un produit croisé.

$$L^T = \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \vec{P}_1 \times \vec{P}_2 = \begin{pmatrix} 0 & -1 & y_1 \\ 1 & 0 & -x_1 \\ -y_1 & x_1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} y_1 - y_2 \\ x_2 - x_1 \\ x_1 y_2 - x_2 y_1 \end{pmatrix}$$

En notation tensorielle, le déterminant et fait par l’opérateur tensorielle E_{ijk} et E^{ijk}. Cette opérateur signifie une evaluation des indices pour une déterminant.

Exemplel La droite L_i est définit par les points P^j et Q^k:

$$L_i = E_{ijk} P^j Q^k$$

Pour

$$P^j = \begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix} = \begin{pmatrix} p^1/p^3 \\ p^2/p^3 \\ 1 \end{pmatrix} \quad \text{et} \quad Q^j = \begin{pmatrix} q^1 \\ q^2 \\ q^3 \end{pmatrix} = \begin{pmatrix} q^1/q^3 \\ q^2/q^3 \\ 1 \end{pmatrix}$$

$i = 1$, $ijk\text{-}ikj = 123\text{-}132$: $l_1 = p^2 q^3 - p^3 q^2$
$i = 2$, $ijk\text{-}ikj = 231\text{-}213$: $l_2 = p^3 q^1 - p^1 q^3$
$i = 3$, $ijk\text{-}ikj = 312\text{-}321$: $l_3 = p^1 q^2 - p^2 q^1$

et si $p^3 = 1$ et $q^3 = 1$ alors nous retrouvons notre forme :

$$P^j = \begin{pmatrix} p^1 \\ p^2 \\ 1 \end{pmatrix} \quad \text{et} \quad Q^j = \begin{pmatrix} q^1 \\ q^2 \\ 1 \end{pmatrix}$$

$i = 1$, $ijk\text{-}ikj = 123\text{-}132$: $l_1 = p^2 q^3 - p^3 q^2 = p^2 - q^2$
$i = 2$, $ijk\text{-}ikj = 231\text{-}213$: $l_2 = p^3 q^1 - p^1 q^3 = p^1 - q^1$
$i = 3$, $ijk\text{-}ikj = 312\text{-}321$: $l_3 = p^1 q^2 - p^2 q^1$

Intersection de deux droites

Pour le calcul d'un point d'intersection de deux droites.

soit deux droites : L: ax + by + c = 0 et M : dx + ey + f = 0.

En notation classique : Soit $\vec{L} = (a\ \ b\ \ c\ )$ et $\vec{M} = (d\ \ e\ \ f\ )$

Le point d'intersection est $x = \dfrac{bf-ce}{ae-bd}$ $y = \dfrac{cd-af}{ae-bd}$

Pour le demontrer :

$$\vec{P} = \vec{L} \times \vec{M} = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} bf-ce \\ cd-af \\ ae-bd \end{pmatrix} = \begin{pmatrix} \frac{bf-ce}{ae-bd} \\ \frac{cd-af}{ae-bd} \\ 1 \end{pmatrix}$$

Un point est l'intérsection d'une infini de droites. Soit deux droits
Soit deux droites (a, b, c) et (d, e, f). soit une droit "libre" u·x + v·y +w = 0.

$$\begin{vmatrix} u & v & w \\ a & b & c \\ d & e & f \end{vmatrix} = u(bf-ce) + v(cd-af) + w(ae-bd)= 0$$

ou bien $u \cdot \dfrac{bf-ce}{ae-bd} + v \cdot \dfrac{cd-af}{ae-bd} + w = 0 = u\cdot x + v\cdot y + w$

donc $x = \dfrac{bf-ce}{ae-bd}$ et $y = \dfrac{cd-af}{ae-bd}$ et 1=1.

En notation Tensorielle, nous avons l'opérateur tensorielle E^{ijk}

On a $P^i = E^{ijk} L_j M_k$

Pour

$$L^j = \begin{pmatrix} l^1 \\ l^2 \\ l^3 \end{pmatrix} = \text{ et } \quad M^k = \begin{pmatrix} m^1 \\ m^2 \\ m^3 \end{pmatrix}$$

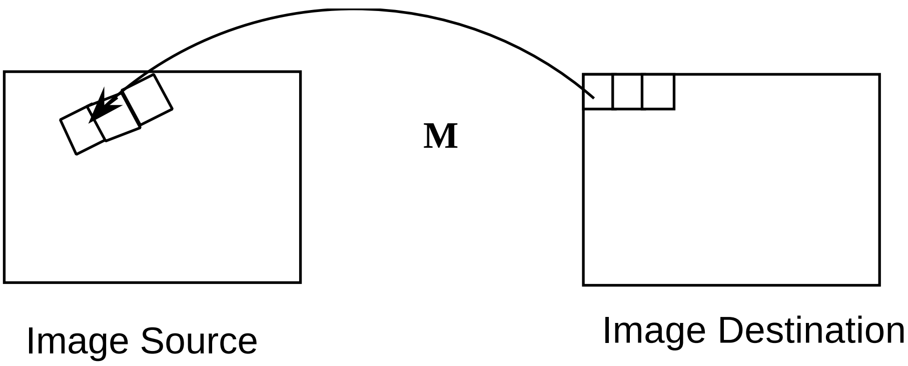
i =1, ijk-ikj = 123-132 : $p^1 = l_2\, m_3 - l_3\, m_2$
i =2, ijk-ikj = 231-213 : $p^2 = l_3\, m_1 - l_1\, m_3$
i =3, ijk-ikj = 312-321 : $p^3 = l_1\, m_2 - l_2\, m_1$

Transformations en Coordonnées Homogènes.

Les coordonnés homogènes fournissent une notation uniforme pour les transformations.

Par exemple, les transformations dans un plan sont décrites par une matrice homogène 3 x 3.

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.

$$P^s = \mathbf{M}^s_d \, P^d$$

ou
$$\begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix} = \begin{pmatrix} m^1_1 & m^1_2 & m^1_3 \\ m^2_1 & m^2_2 & m^2_3 \\ m^3_1 & m^3_2 & m^3_3 \end{pmatrix} \begin{pmatrix} x_d \\ y_d \\ 1 \end{pmatrix}$$

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

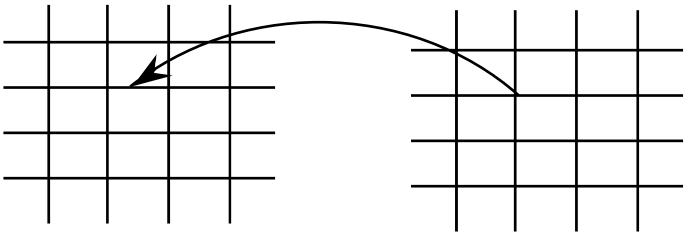
ou
$$\begin{pmatrix} x_s \\ y_s \\ 1 \end{pmatrix} = \begin{pmatrix} p^1/p^3 \\ p^2/p^3 \\ 1 \end{pmatrix}$$

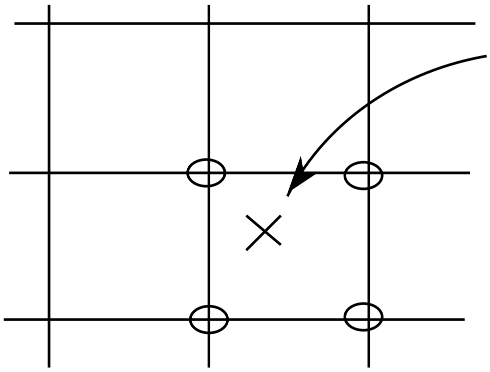
Ensuite la nouvelle valeur de
pixel destination, P^2, est calculé en fonction du voisinage de la position source.

MAIS, $P^s = \begin{pmatrix} x_s \\ y_s \\ 1 \end{pmatrix}$ 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.



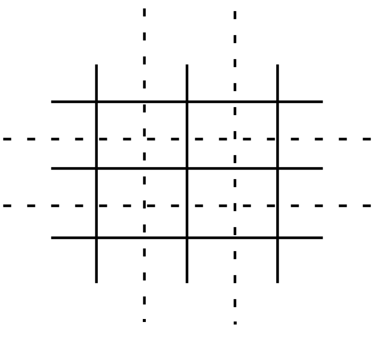


Il existe plusieurs fonctions d'interpolation.

Ordre zéro : Plus proche voisin.
Ordre unité : intérpolation linéaire et "bi-linéaire"
Ordre trois : Spline cubic.

Interpolation d'ordre zéro

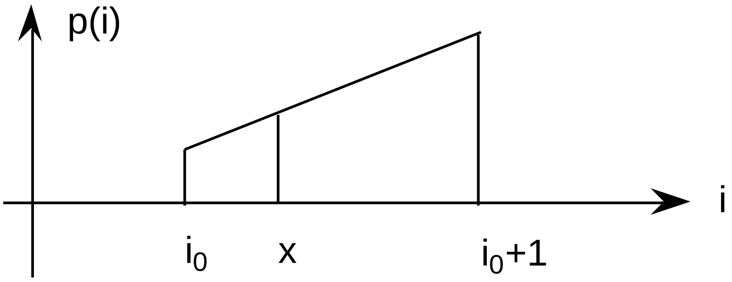
Pour les images Binaire, on peut fair que l'ordre zéro.
La valeur de $p(i_2, j_2)$ déterminer par arrondis de $p(i_1, j_1)$.



Surface de Décision - - - - - - -

Interpolation Linéaire

Interpolation Linéaire en 1-D. soit $i_o \leq x \leq i_o+1$



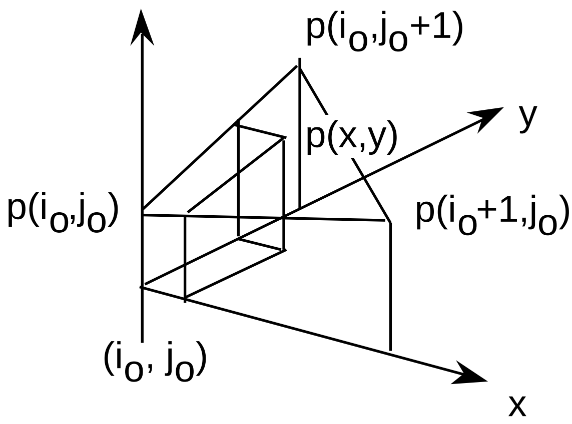
A partir de l'origine : $p(x) = p(0) + m_x\, x$

A partire de deux points i_o et i_0+1 :

pente : $m_x \equiv \frac{\Delta P}{\Delta x} = = p(i+1) - p(i)$

$$p(x) = (x - i_o)\, m_x + p(i_o)$$

Interpolation Linéaire en 2D

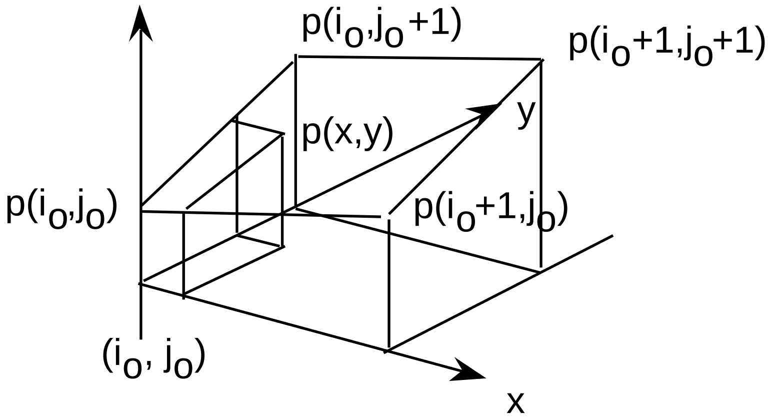


$$m_x \equiv \frac{\Delta P}{\Delta x} = p(i_o+1, j_o) - p(i_o, j_o)$$

$$m_y \equiv \frac{\Delta P}{\Delta y} = p(i_o, j_o+1) - p(i_o, j_o)$$

donc $p(x, y) = m_x \cdot (x - i_o) + m_y \cdot (y - j_o) + p(i_o, j_o)$

Interpolation Bi-lineaire



Forme Bilinéaire : Hyperbolic Paraboloïde

$$p(x, y) = a\,x + b\,y + c\,x\,y + d.$$

Une interpolation linéaire en "y" de deux interpolations linaire en "x".

Dérivation :

$$p(x, 0) = p(0, 0) + x \cdot (p(1, 0) - p(0, 0))$$
$$p(x, 1) = p(0, 1) + x \cdot (p(1, 1) - p(0, 1))$$

$$p(x, y) = p(x, 0) + y \cdot (p(x, 1) - p(x, 0))$$

$$\begin{aligned} &= p(0, 0) + x \cdot (p(1, 0) - p(0, 0)) \\ &\qquad + y \cdot (p(0, 1) + x \cdot (p(1, 1) - p(0, 1))) \\ &\qquad - y \cdot (p(0, 0) + x \cdot (p(1, 0) - p(0, 0))) \\ &= p(0, 0) + x \cdot (p(1, 0) - p(0, 0)) + y \cdot (p(0, 1) - p(0, 0)) \\ &\qquad + x \cdot y \cdot (p(1, 1) - p(0, 1) - p(1, 0) + p(0, 0)) \end{aligned}$$

Pour le point i_o, j_o, remplace : $0 \rightarrow i_o$, $1 \rightarrow i_o$, $x \rightarrow (x - i_o)$, $y \rightarrow (y - i_o)$

$$a \equiv m_x = \frac{\Delta P}{\Delta x} = p(i_o+1, j_o) - p(i_o, j_o)$$
$$b \equiv m_y = \frac{\Delta P}{\Delta y} = p(i_o, j_o+1) - p(i_o, j_o)$$
$$c \equiv m_{xy} = p(i_o+1, j_o) + p(i_o, j_o+1) - p(i_o, j_o) - p(i_o+1, j_o+1)$$
$$d = p(i_o, j_o)$$

$$p(x, y) = a \cdot (x - i_o) + b \cdot (y - j_o) + c \cdot (x - i_o) \cdot (y - j_o) + p(i_o, j_o)$$

Translation en Coordonnées Homogènes.

$x_2 = x_1 + t_x$,
$y_2 = y_1 + t_y$

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

En notation tensorielle :

$P^B = \mathbf{T}_A^B P^A$ pour A, B = 1, 2, 3.

Donc $\mathbf{T}_A^B$ est une transformation du repère A vers le repère B.

Les indices permet de noter les repères.

Rotation

(Repère main droite, rotation sens trigonométrique)

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

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

Quand le repère tourne dans le sanse "θ", le vecteur est tourner dans le sense –θ

Translation et Rotation en Coordonnées Homogènes.

$x_2 = \text{Cos}(\theta)\, x_1 + \text{Sin}(\theta)\, y_1 + t_x$
$x_2 = \text{Sin}(\theta)\, x_1 - \text{Cos}(\theta)\, y_1 + t_y$

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

Ceci modèlise la rotation, suivi de la translation.

En tensorielle, $P^B = \mathbf{R}_A^B P^A$

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 de le repère 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, θ, t_x, t_y) de

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

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

ou bien

$$x_2 = s_x\text{Cos}(\theta)\, x_1 + s_y \text{ Sin}(\theta) y_1 + t_x,$$
$$y_2 = s_y\text{Sin}(\theta)\, x_1 - s_x \text{ Cos}(\theta) y_1 + t_y$$

Homographie entre deux plans

La projection d'un plan vers un autre plan est une transformation projective entre deux plans. Cette transformation s'appele une "homographie.
L'homographie est bijective.

Elle est bijective est facile a estimate et de "recetifié.

$$\boxed{Q^B = \mathbf{H}_A^B \; P^A}$$

En notation "classique".

$$\begin{pmatrix} w\ x_B \\ w\ y_B \\ w \end{pmatrix} = \mathbf{H}_A^B \begin{pmatrix} x_A \\ y_A \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} \begin{pmatrix} x_A \\ y_A \\ 1 \end{pmatrix}$$

$$x_B = \frac{w\ x_B}{w} = \frac{m_{11}\ x_A + m_{12}\ y_A + m_{13}}{m_{31}\ x_A + m_{32}\ y_A + m_{33}}$$

$$y_B = \frac{w\ y_B}{w} = \frac{m_{21}\ x_A + m_{22}\ y_A + m_{23}}{m_{31}\ x_A + m_{32}\ y_A + m_{33}}$$

En notation "tensorielle"

$$\boxed{Q^B = \mathbf{H}_A^B \; P^A}$$

$$\begin{pmatrix} q^1 \\ q^2 \\ q^3 \end{pmatrix} = \mathbf{H}_A^B \begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix} = \begin{pmatrix} h_1^1 & h_2^1 & h_3^1 \\ h_1^2 & h_2^2 & h_3^2 \\ h_1^3 & h_2^3 & h_3^3 \end{pmatrix} \begin{pmatrix} p^1 \\ p^2 \\ p^3 \end{pmatrix}$$

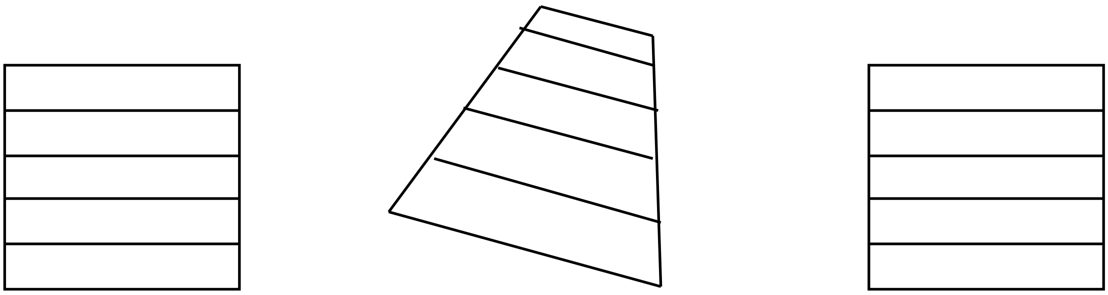
$$x_B = \frac{q^1}{q^3} \qquad y_B = \frac{q^2}{q^3}$$

Rectification de l'Image d'un plan.

La projection d'un plan 2D dans un scène 3D vers le retine 2D d'un caméra est une homographie. Elle est bijective est facile a "corriger".

Ceci a des nombreuse applications. Par exemple, l'image d'un tableau blanc peut être "rectifié" par l'homographie inverse.

$$\mathbf{H}_D^S = (\mathbf{H}_S^D)^{-1}$$



Dans le scène | Dans l'image (source) | L'image Rectifié (destination)

Pour chaque pixel de l'image destination, on calcul son position dans l'image source.

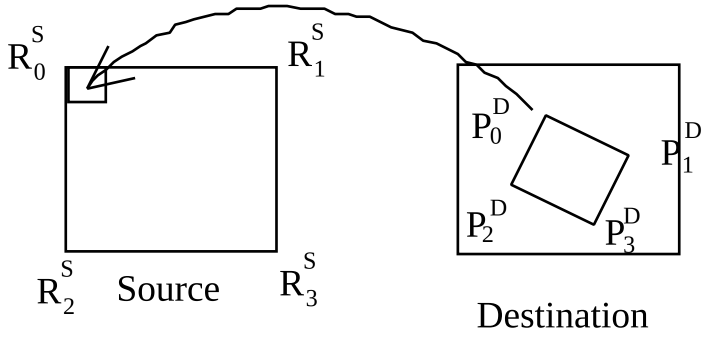
$$Q^S = \mathbf{H}_D^S \; P^D \quad \text{pour } S, D = 1, 2, 3.$$

Ensuite Pour chaque pixel P^D on détermine la valeur de la pixel au position

$$\begin{pmatrix} x_s \\ y_s \\ 1 \end{pmatrix} = \begin{pmatrix} q^1/q^3 \\ q^2/q^3 \\ 1 \end{pmatrix}$$

Estimation et Rectification par Homographie

La Homographie peut être déterminer par observation des 4 coins d'un carré.



Soit les qautre coins, k=1, 2, 3, 4 de l'image dans repere destination :

$$P_k^D = \begin{pmatrix} p_0^1 & p_1^1 & p_2^1 & p_3^1 \\ p_0^2 & p_1^2 & p_2^2 & p_3^2 \\ p_0^3 & p_1^3 & p_2^3 & p_3^3 \end{pmatrix}$$

On dit qu'ils corresponde au sommets de l'image source

$$R_k^S = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

On note $R_k^S = \mathbf{H}_D^S \; P_k^D$

Le 3 x 3 matrice inconnu, $\mathbf{H}_D^S$ a 9 coefficients.
Parce que $\mathbf{H}_D^S$ est en coordonées homogénes, on peut fixer $\mathbf{H}_3^3 = 1$.
Il nous reste 8 coéfficients à estimer.

Donc :

$$R_k^1 / R_k^3 = \mathbf{H}_D^1 \; P_k^D \; / \; \mathbf{H}_D^3 \; P_k^D$$

et
$$R_k^2 / R_k^3 = \mathbf{H}_D^2 \; P_k^D \; / \; \mathbf{H}_D^3 \; P_k^D$$

$$R_k^1 \; \mathbf{H}_D^3 \; P_k^D = R_k^3 \; \mathbf{H}_D^1 \; P_k^D$$
$$R_k^2 \; \mathbf{H}_D^3 \; P_k^D = R_k^3 \; \mathbf{H}_D^2 \; P_k^D$$

$$R_k^1 \; \mathbf{H}_D^3 \; P_k^D - R_k^3 \; \mathbf{H}_D^1 \; P_k^D = 0$$
$$R_k^2 \; \mathbf{H}_D^3 \; P_k^D - R_k^3 \; \mathbf{H}_D^2 \; P_k^D = 0$$

Ceci donne deux équations (S=1, 2) pour chaque coin (k=1,2,3,4) , donc 8 equations pour les 8 inconnues du $\mathbf{H}_D^S$. L'equation pour S =3 n'est pas indépendant de S =1 et S=2.

Ensuite

Pour chaque pixel, P^D :

SetPixel (Destination, P^D)= BiLinear_Interpolate(Source, ($\mathbf{H}_D^S$ P^D));