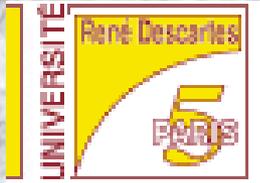


MASTER IPCC



Nicolas Loménie

Vision par Ordinateur 3D en Mouvement

MASTER IPCC



Nicolas Loménie

Vision par Ordinateur

INTRODUCTION

INTRODUCTION

But :

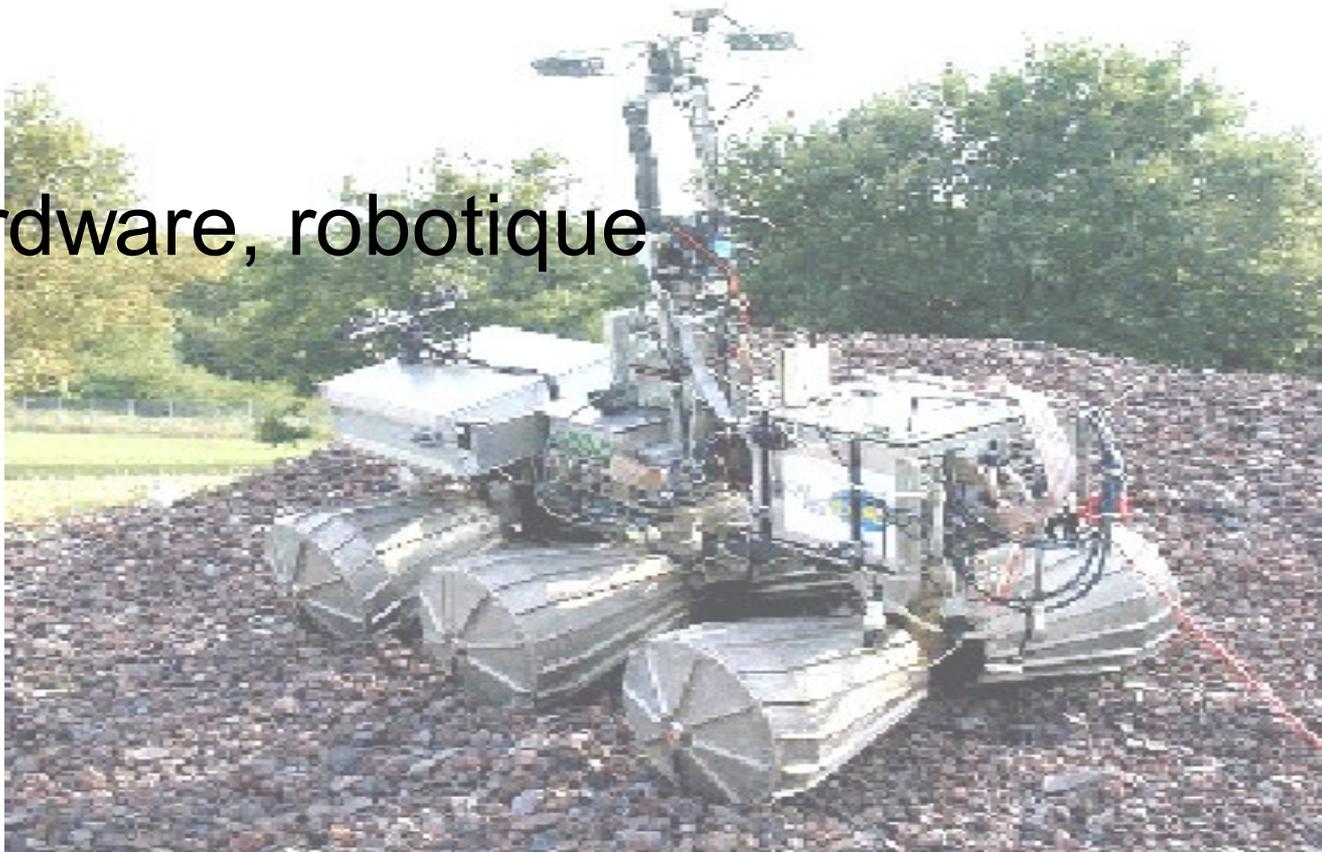
Calculer des propriétés du monde 3D
à partir d'une ou de plusieurs images numériques,

... propriétés :

- **Géométriques** : forme, position ...
- **Dynamiques** : vitesse ...

Outils :

- Hardware, robotique



- Algorithmes
-> **objet de ce cours pour l'essentiel**

Disciplines :

- Image Analysis
- Scene Analysis
- Image Understanding

- Image Processing
- Pattern Recognition
- Photogrammetry

Champs de Recherches :

- Image Feature Detection
- Contour Representation
- Feature-based segmentation
- Range Image Analysis
- Shape Modeling and Representation
- Shape Reconstruction from Single-image Cues (Shape from X)
- Stereo Vision
- Motion Analysis
- Color Vision
- Image Understanding...

Applications :

- Industrial Inspection and Quality Control
- Surveillance and Security
- Road monitoring and Autonomous vehicles
- Hand-eye robotics systems
- Space, Medical, Virtual reality, telepresence and telerobotics...

Réalité virtuelle



Ressources :

Conférences :

- ICCV
- CVPR
- ECCV
- ICIO
- ICPR

Journaux :

- Int. J. of Comp. Vision
- IEEE Trans. PAMI
- CVIU
- MVA
- PR
- PRL

Internet :

- Computer Vision Home Page :
<http://www-2.cs.cmu.edu/~cil/vision.html>
- *peipa.essex.ac.uk*
- *iris.usc.edu/Vision-Notes/bibliography/contents.html*
- *<http://www.math-info.univ-paris5.fr/sip-lab/>*

Quelques périodiques en ligne accessibles depuis Paris 5 :

Computational Geometry	<i>(depuis Mars 1995).</i>
Image and Vision Computing	<i>(depuis Février 1995).</i>
International journal of computer vision	<i>(depuis 1995)</i>
International Journal of Imaging Systems and Technology	<i>(depuis 1996).</i>
Pattern Recognition Letters	<i>(depuis Janvier 1995).</i>

Sont également accessibles de nombreux autres périodiques, publiés par :

Elsevier - Science Direct
Kluwer Online
Wiley Interscience

Site web et catalogue de quelques bibliothèques :

Bibliothèque Universitaire de Paris 5
BIU Scientifique Jussieu (via telnet) : login=mlibrary et passwd=jussieu.
Bibliothèque Informatique-Recherche (Jussieu - site du Capitaine Scott)

Problématiques abordées :

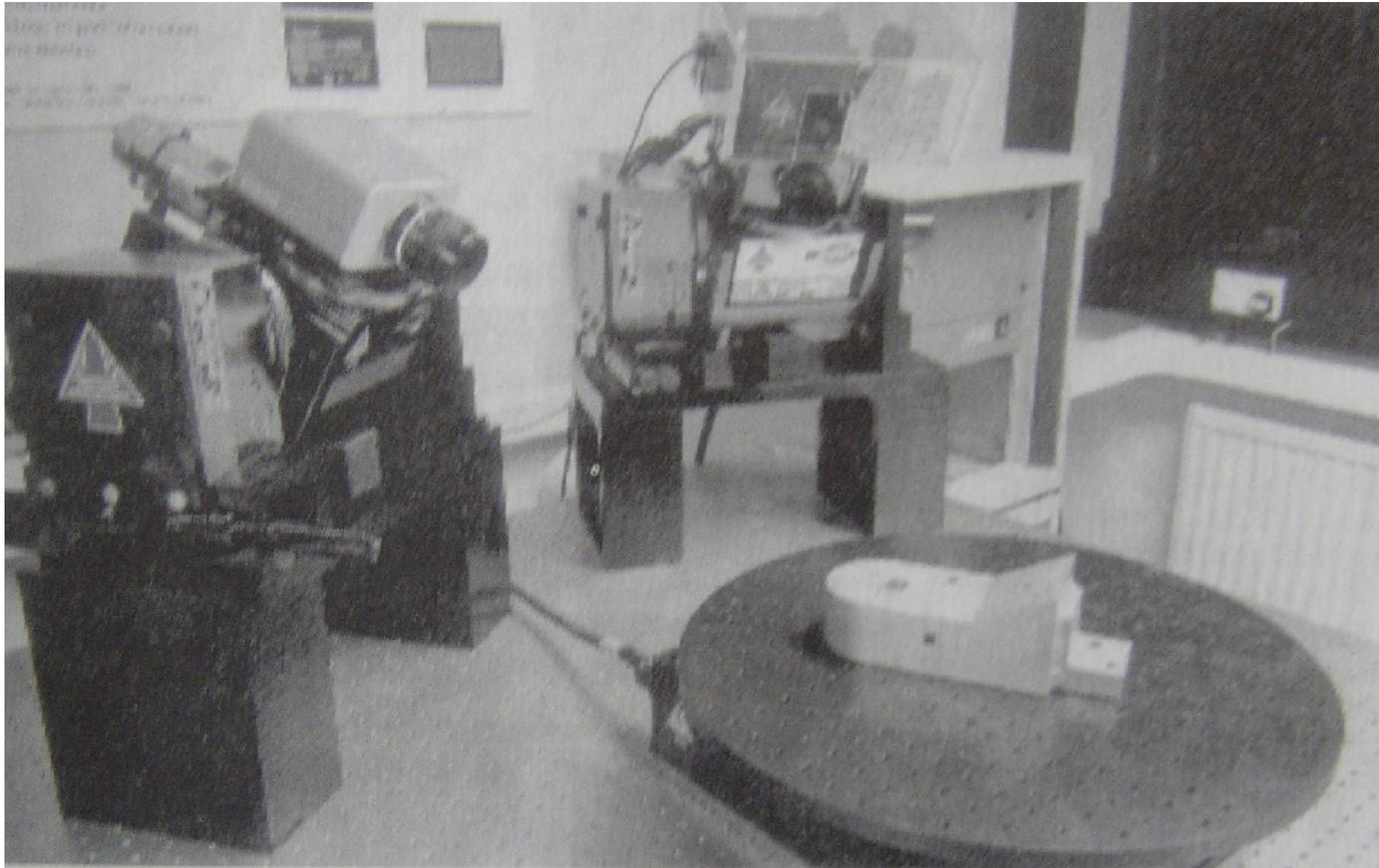
Methodes utilisant les “image features” :

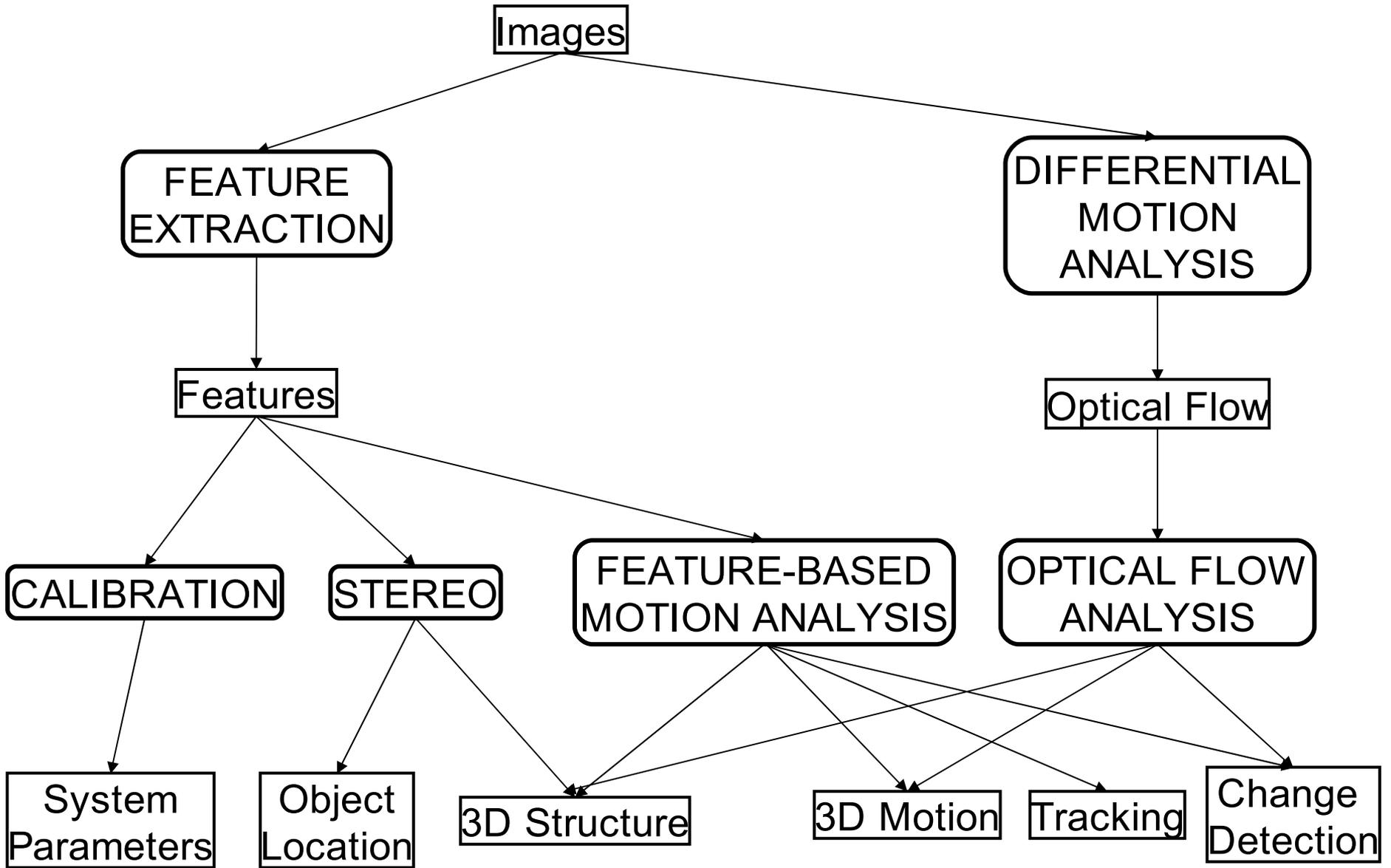
- Calibration
- Stereo Analysis
- Feature-based motion analysis

Methodes utilisant les images directement :

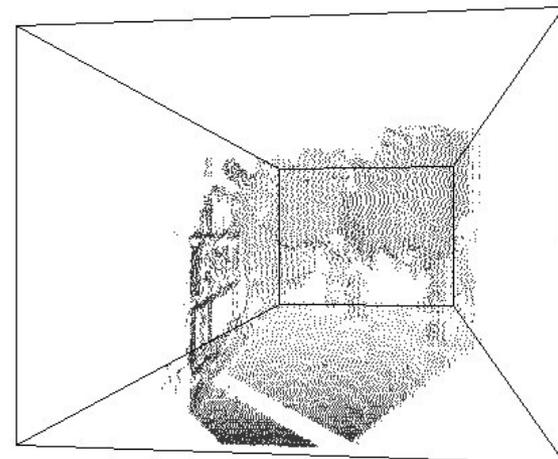
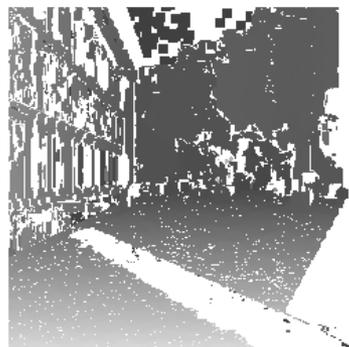
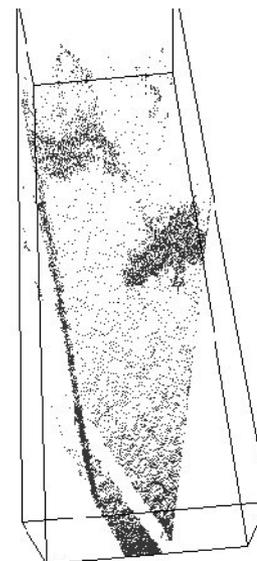
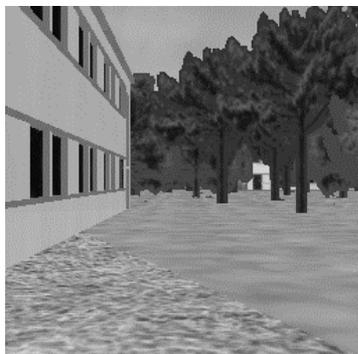
- Shape from single image
- Optical Flow methods
- Stereo Analysis

Inspection 3D : scanner laser + système à base de franges de Moiré

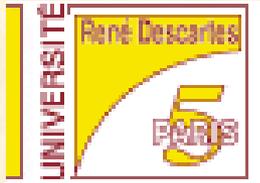




Intensity Image vs. Range Image



MASTER IPCC



Vision par Ordinateur

Nicolas Loménie

CAPTEURS ET PARAMETRES

Capteurs et paramètres

Objectifs :

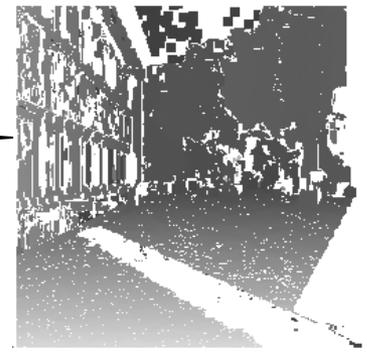
- Comprendre comment est imagé le monde 3D sur une rétine 2D (capteur optique dit passif = oeil humain)
 - Géométriquement
 - Algébriquement
- Apprendre la notion de modèle de caméra pinhole et de matrice de projection
- Définir les paramètres extrinsèques et intrinsèques d'un capteur passif
- Illustrer un système d'acquisition 3D actif

Intensity image :

images traditionnelles encodant les intensités lumineuses acquises par des caméras type TV (*voir cours traitement d'images*)

Range image :

images codant la forme et la distance, acquises par des capteurs spécifiques comme le scanner laser. On parle d'images de profondeur ou topographiques.



Processus d'acquisition

139	140	136	140	172	221	217	222	219	217
136	137	137	143	169	219	212	218	222	209
139	140	139	138	171	218	217	219	222	213
141	145	145	145	172	219	214	222	217	205
138	139	147	147	172	217	218	219	224	203
143	146	143	143	170	219	220	224	225	201
150	144	144	146	172	217	222	221	222	210
137	132	138	137	166	221	219	224	229	240
138	143	141	146	171	222	221	225	223	196
146	145	141	142	169	221	222	223	218	189
142	141	141	147	172	217	223	225	223	215
150	143	145	147	170	225	225	222	226	220
149	140	144	139	169	223	225	223	227	223
142	139	143	144	172	224	224	225	229	224
147	144	145	147	171	227	229	228	233	225
142	141	140	147	173	226	227	227	233	228

318	322	318	318	321	319	318	318	316	316
319	317	318	318	316	317	315	317	319	312
319	316	319	316	317	318	317	315	319	313
320	321	320	320	316	317	315	317	317	316
317	322	316	318	320	316	314	315	314	320
319	318	317	317	321	318	318	316	316	317
318	318	319	316	318	319	319	315	316	313
318	321	317	319	317	315	318	312	318	316
317	319	315	317	320	317	315	318	316	318
318	319	316	319	316	316	317	318	314	315
318	317	319	319	318	319	318	319	318	318
317	320	317	319	317	317	319	315	319	320
320	317	319	319	315	319	319	319	319	316
319	318	318	321	321	318	316	320	319	319
320	322	318	323	319	321	321	324	318	320
318	318	318	320	320	319	320	319	319	322

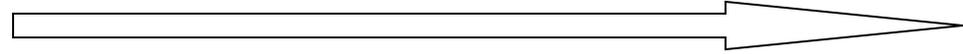
Algorithmes d'extraction d'information

Objet de ce cours : l'Analyse du monde 3D en mouvement !!

Le processus inverse, la Synthèse 3D, utilise les mêmes principes théoriques. Computer Graphics is not so far !!

Formation des Image d'intensités

1. Paramètres optiques :



- Type de lentille
- Focale
- Ouverture angulaire

2. Paramètres photométriques :

- Type, intensité et direction de l'illumination
- Propriétés de réflectance des surfaces observées

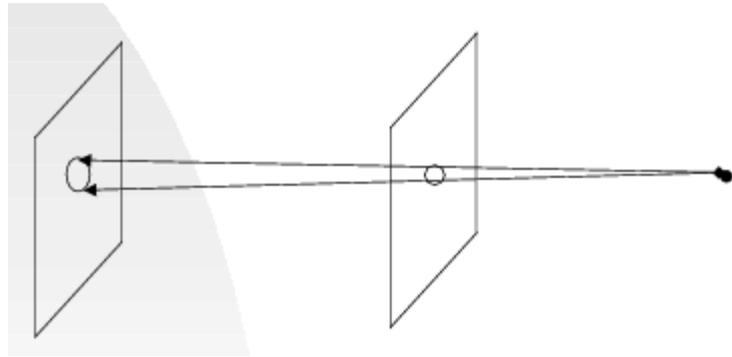
3. Paramètres géométriques :

- Type de projection
- Position et orientation de la caméra
- Distorsions de perspective

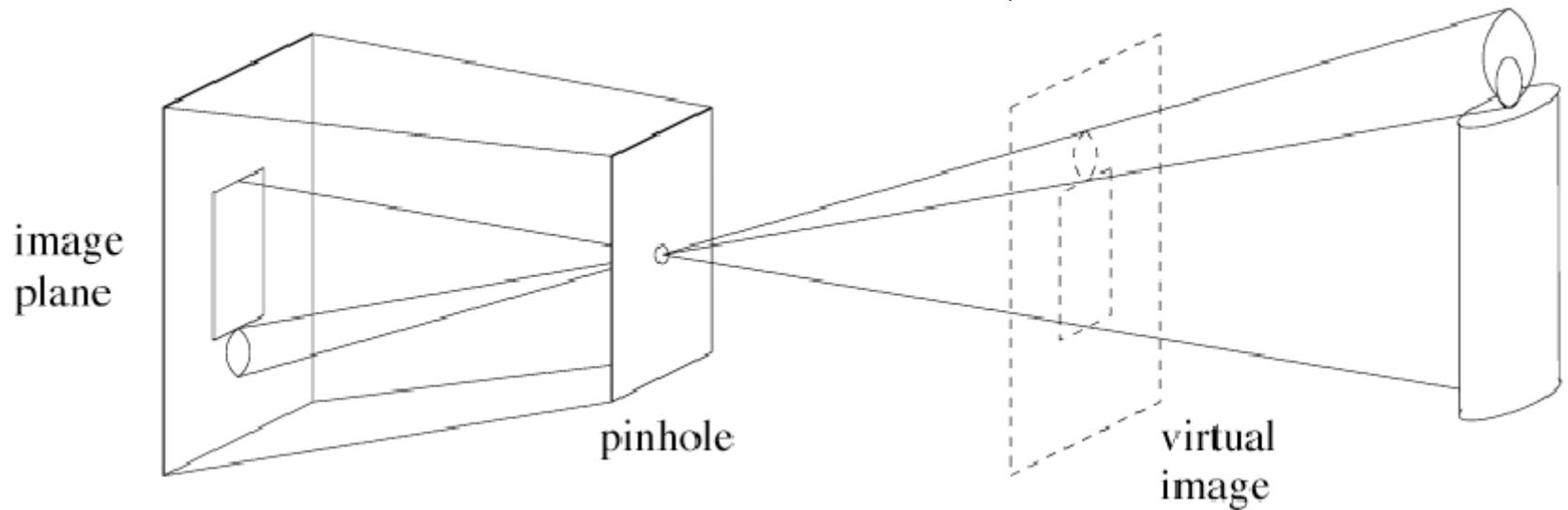
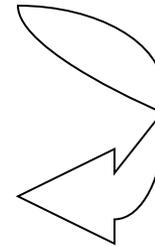
4. Nature discrète des photorécepteurs

5. Quantification de l'échelle d'intensité

Pinhole Camera



Nécessité de focaliser l'image ...



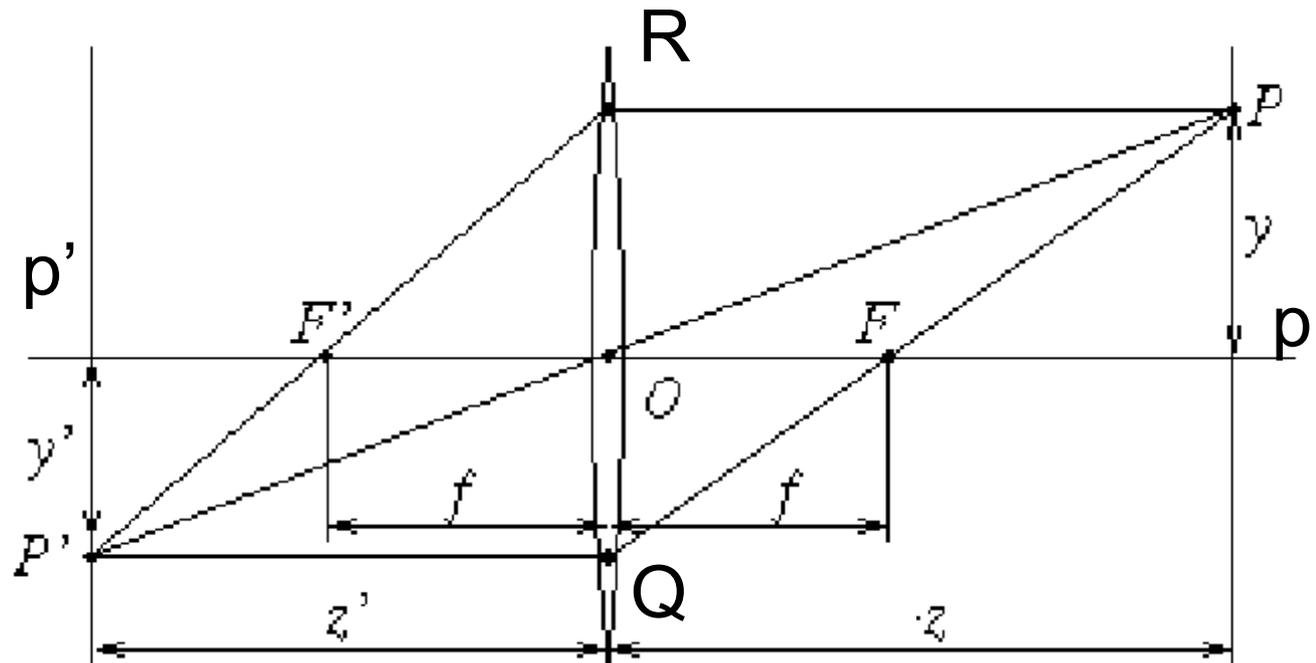
Ouverture faible donc temps d'exposition long



Systemes optiques réels

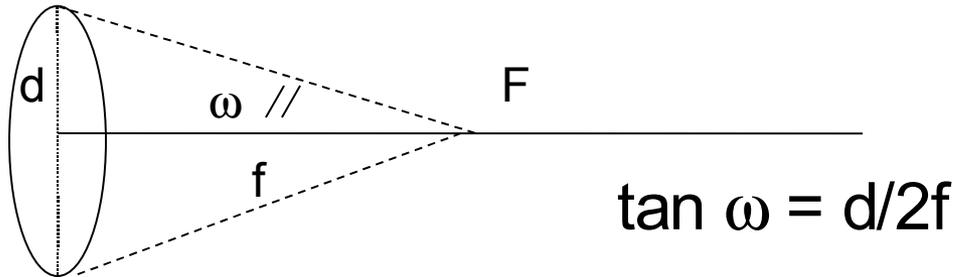
Lentille mince

$$\frac{1}{z'} + \frac{1}{z} = \frac{1}{f}$$



Champ visuel :

Mesure de la portion d'espace 3D véritablement observée par la caméra



Problèmes :

- objets focalisés à des distances à la caméra différentes : notion de **champ de profondeur** de la caméra (le modèle pinhole implique un champ de profondeur infini)

Intensity image :

1. Paramètres optiques :

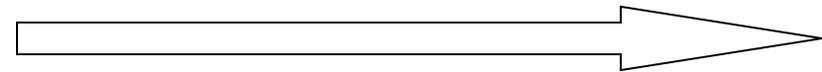
- Type de lentille
- Focale
- Ouverture angulaire

2. Paramètres photométriques :

- Type, intensité et direction de l'illumination
- Propriétés de réflectance des surfaces observées

3. Paramètres géométriques :

- Type de projection
- Position et orientation de la caméra
- Distorsions de perspective



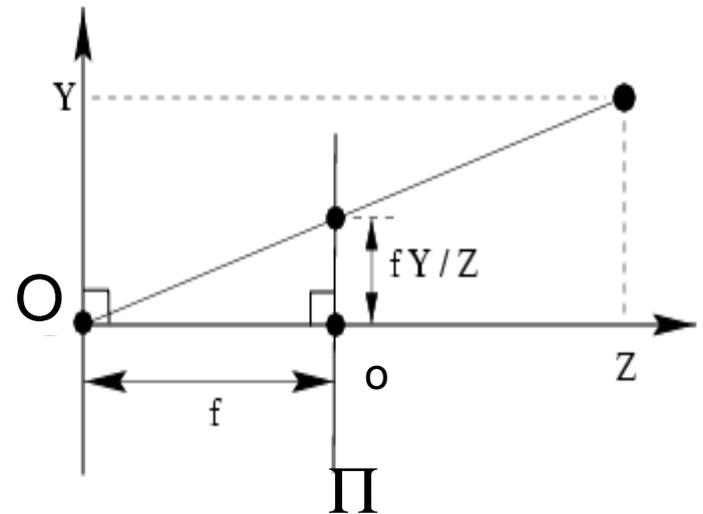
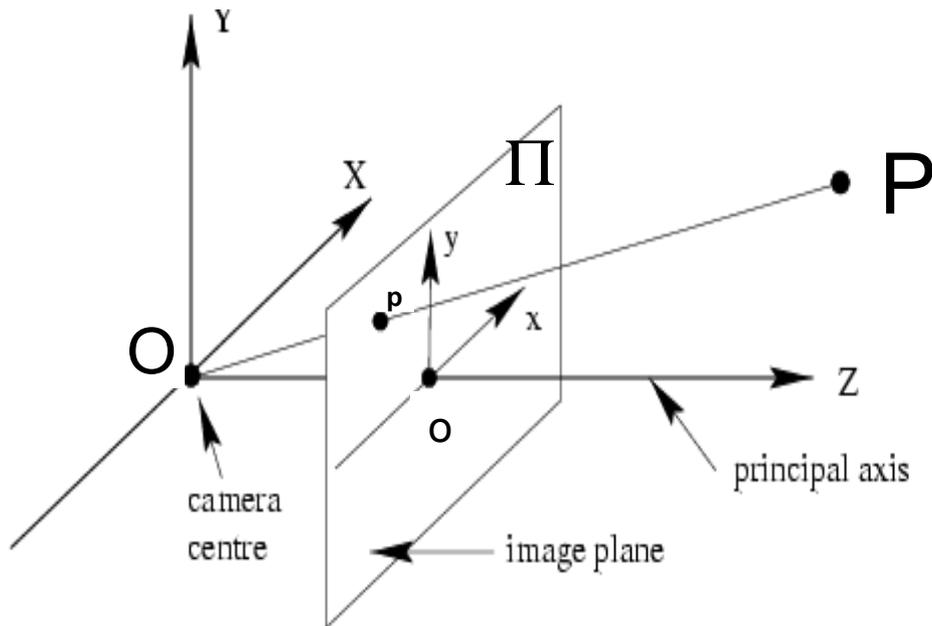
4. Nature discrète des photorécepteurs

5. Quantification de l'échelle d'intensité

Modèle de caméra perspective

Modèle de la formation géométrique de l'image =
Modèle de la projection géométrique opérée par le capteur

Communément, modèle de projection perspective pour un
modèle de caméra pinhole

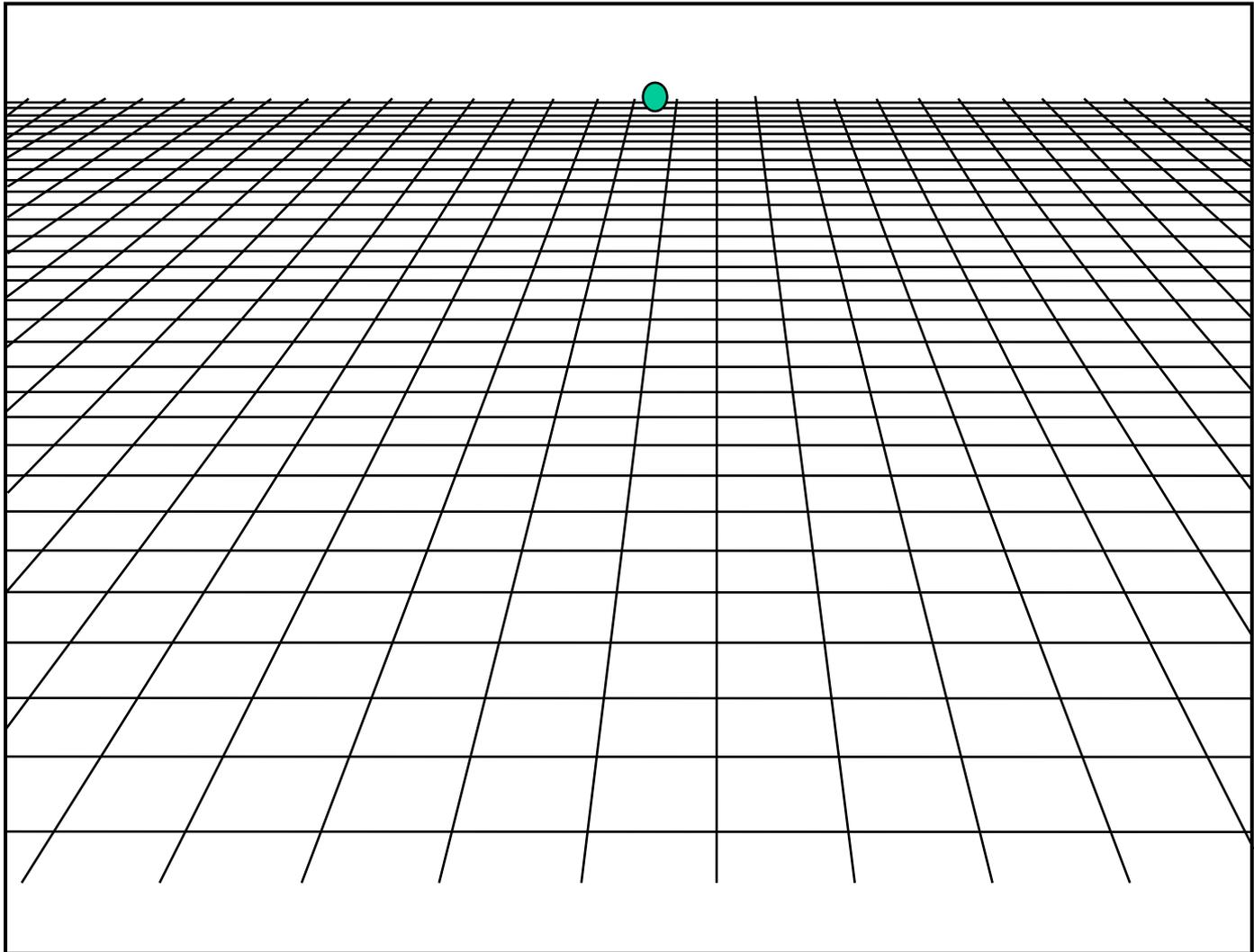


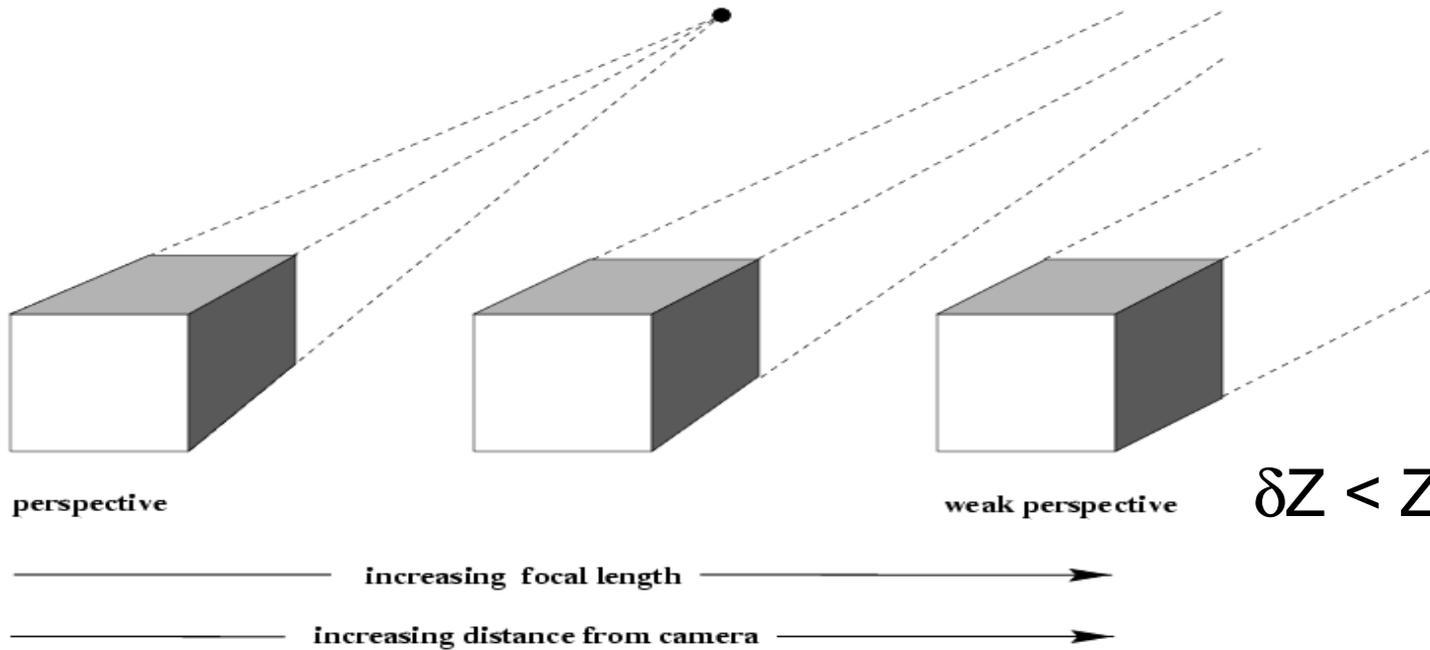
Les équations fondamentales de la caméra perspective

Dans le repère caméra, on a :

$$\begin{cases} x = f X/Z \\ y = f Y/Z \\ z = f \end{cases}$$

- On écrira $p = [x,y]^T \longrightarrow$ Voir coordonnées homogènes
- Equations non linéaires
- Conservation de l'alignement et c'est tout

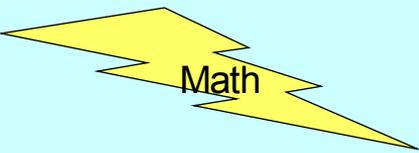




$$\delta Z < Z_{\text{moyen}}/20$$



Modèle orthographique : $f \rightarrow \infty$ et $Z_{\text{moyen}} \rightarrow \infty$



Géométrie projective et coordonnées homogènes

- Coordonnées homogènes 2D : \bar{u}

Si $u = [x \ y]^T$ en coordonnées cartésiennes

$$\left\{ \begin{array}{l} v = Au + t \Leftrightarrow v = M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = M\bar{u} = M \begin{bmatrix} \alpha x \\ \alpha y \\ \alpha \end{bmatrix} \\ \text{avec } M = [A|t] \end{array} \right.$$

u



- Une transformation affine générique $v=Au+t$ devient linéaire en coordonnées homogènes

$$\bar{v} = M\bar{u}$$

$$M_{3 \times 3} = \begin{bmatrix} A & t \\ 0 \ 0 & 1 \end{bmatrix}$$

Définitions :

Un point de l'espace projectif à n dimensions, \mathcal{P}^n , est représenté par un vecteur de dimension $n+1$ de coordonnées $x = [x_1, \dots, x_{n+1}]^T$, avec au moins une des coordonnées x_i non nulle

x_i est appelée coordonnée homogène ou projective

x est appelé vecteur coordonnées

Deux vecteurs $x = [x_1, \dots, x_{n+1}]^T$ et $y = [y_1, \dots, y_{n+1}]^T$ représentent le même

$$\exists \lambda \neq 0 / x_i = \lambda y_i \quad \forall i \in [1, n+1]$$

point ssi

-> Il n'y a pas de correspondance biunivoque entre points et vecteurs de coordonnées ce qui rend l'application de l'algèbre linéaire à la géométrie projective un peu plus ambiguë qu'en géométrie euclidienne

- Transformation Rigide 2D

$$v = R_\theta u + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \Leftrightarrow \bar{v} = M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = M\bar{u}$$

u

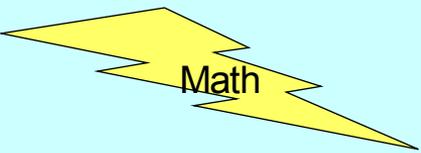
$$M = \begin{bmatrix} \cos\theta & -\sin\theta & 2 \\ \sin\theta & \cos\theta & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

- Identique en 3D

$$M_{4 \times 4} = \begin{bmatrix} A & t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Classification des transformations 3D

Transformation	Vector Coordinates	Homogeneous Coordinates	Degrees of Freedom	Invariants
Translation	$\mathbf{y} = \mathbf{x} + \mathbf{t}$	$\begin{bmatrix} \mathbf{I} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}$	3	lengths, angles
Rotation	$\mathbf{y} = \mathbf{R}\mathbf{x}$ $\mathbf{R}^T \mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}$	$\begin{bmatrix} \mathbf{R} & \mathbf{0} \\ 0 & 0 & 0 & 1 \end{bmatrix}$	3	lengths, angles
Rigid	$\mathbf{y} = \mathbf{R}\mathbf{x} + \mathbf{t}$	$\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}$	6	lengths, angles
Affine	$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{t}$	$\begin{bmatrix} \mathbf{A} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}$	12	ratios of lengths, parallelism
Projective		4×4 matrix \mathbf{M}	15	colinearity, incidence



Classification des transformations 3D ->2D d'un point de coordonnées (x,y,z) vers un point (x',y')

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = M \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$x' = \frac{u}{w} \quad \text{et} \quad y' = \frac{v}{w}$$

Transformation	Homogeneous Coordinates	Degrees of Freedom
Orthographic	$\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}$ $\mathbf{R}\mathbf{R}^T = \mathbf{I} \text{ (row-orthonormal)}$	5
Affine	$\begin{bmatrix} \mathbf{A} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}$	8
Projective	3 × 4 matrix \mathbf{M}	11

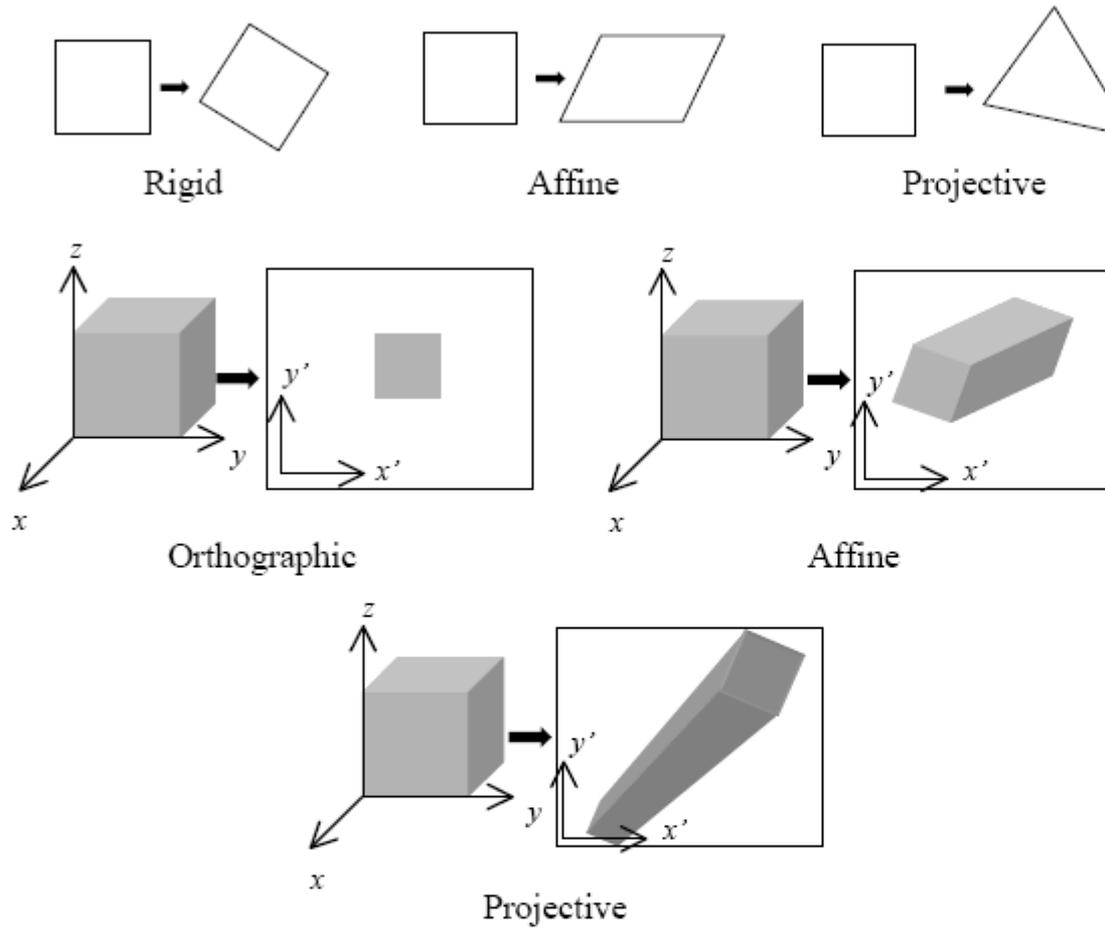
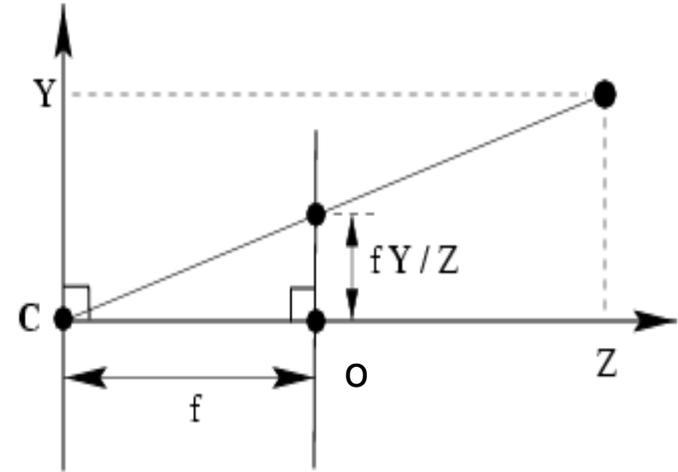
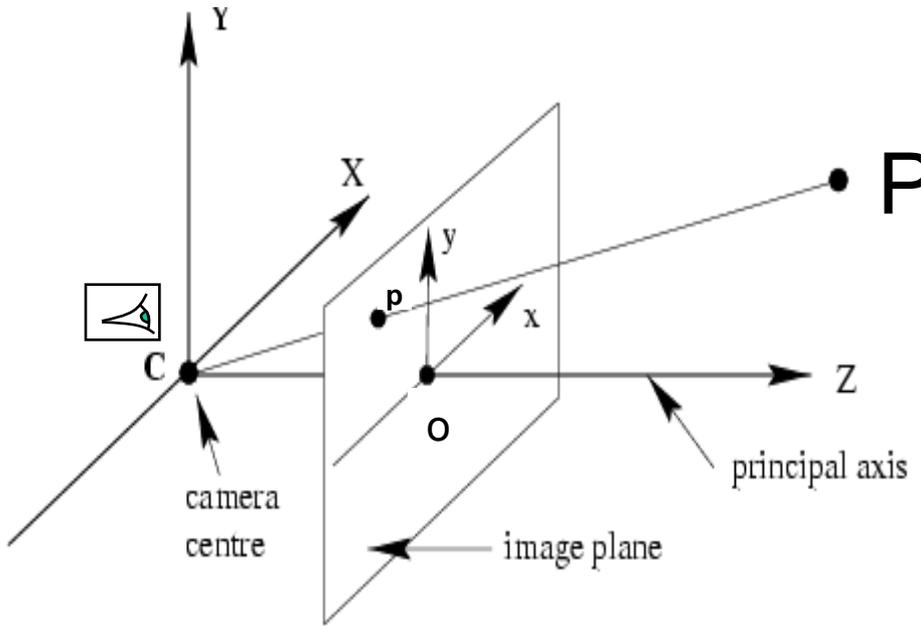


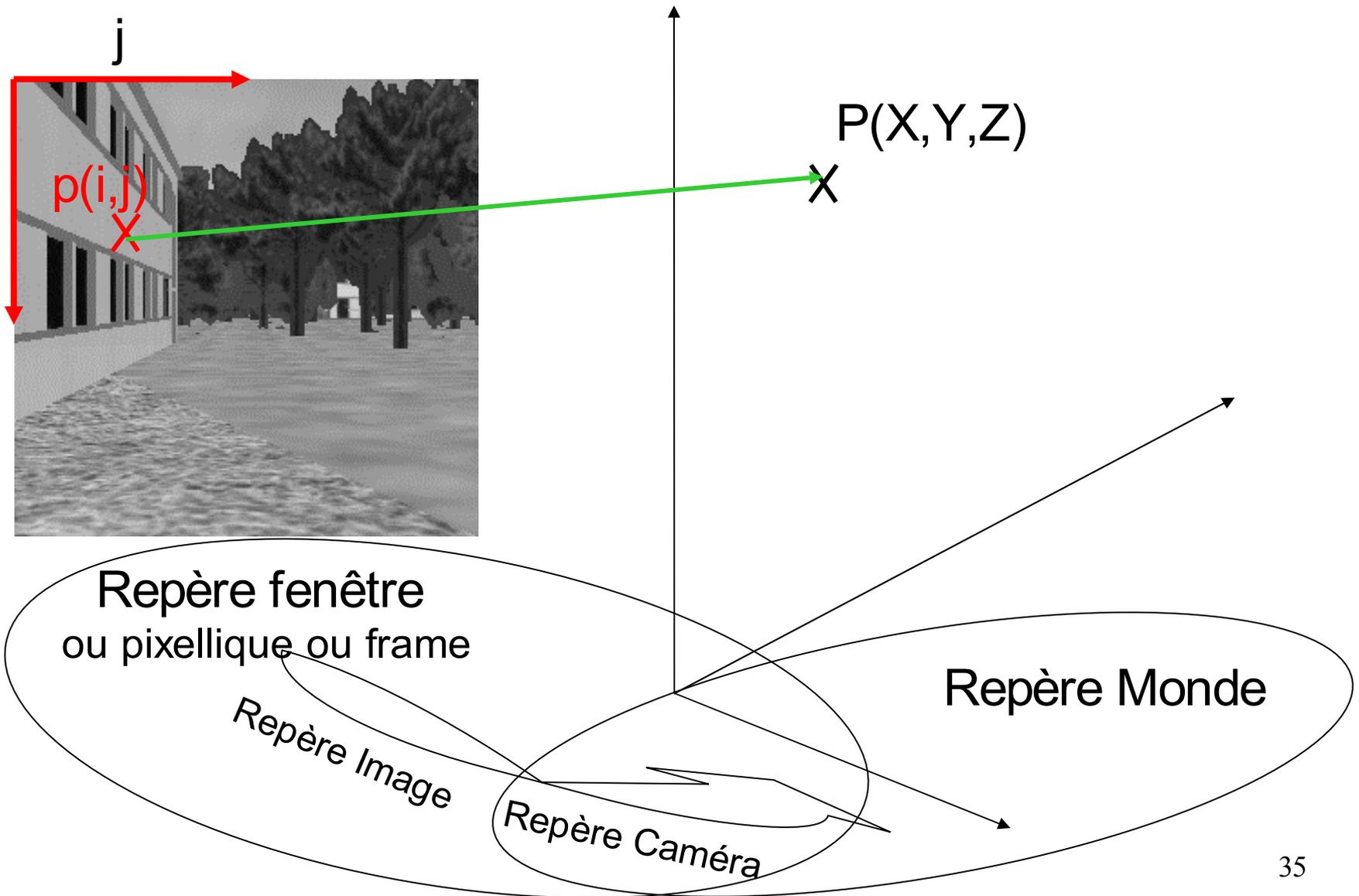
Figure 1. Examples of 2-D transformations (top) and 3-D to 2-D transformations (bottom).

Modèle de caméra perspective, version linéarisée dans le plan projectif !



$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad x = \frac{u}{w} \quad \text{et} \quad y = \frac{v}{w}$$

Paramètres de caméra...



Paramètres extrinsèques :

paramètres qui définissent l'orientation et la localisation du repère caméra par rapport à un repère monde connu

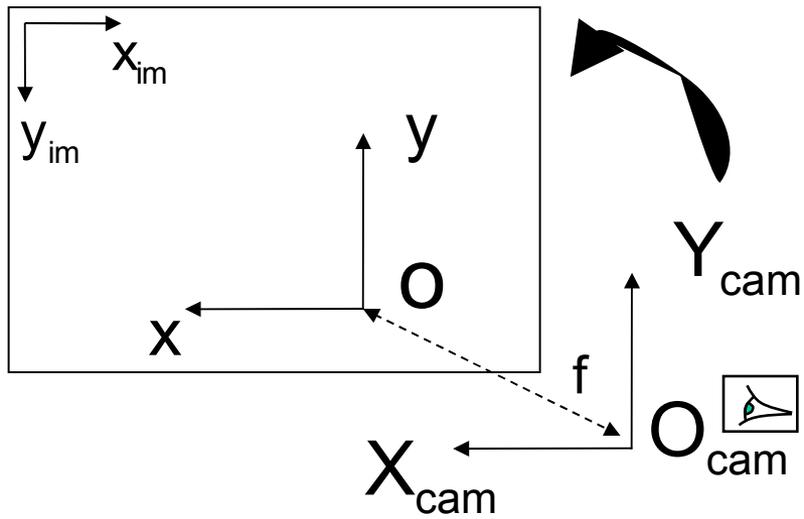
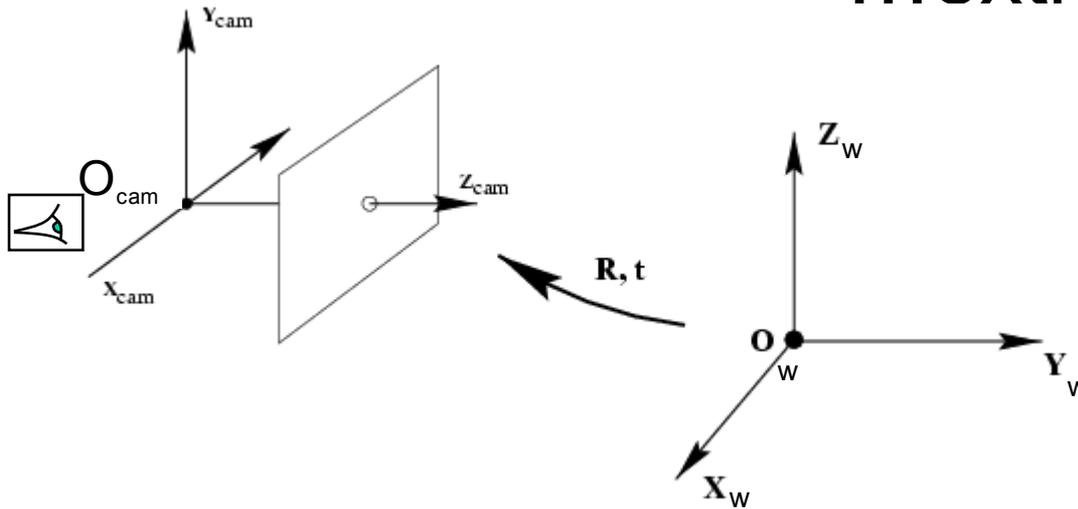
Paramètres intrinsèques

paramètres nécessaires pour lier les coordonnées pixelliques d'un point image aux coordonnées correspondantes dans le repère caméra

Estimation de ces paramètres = **calibration** de caméra

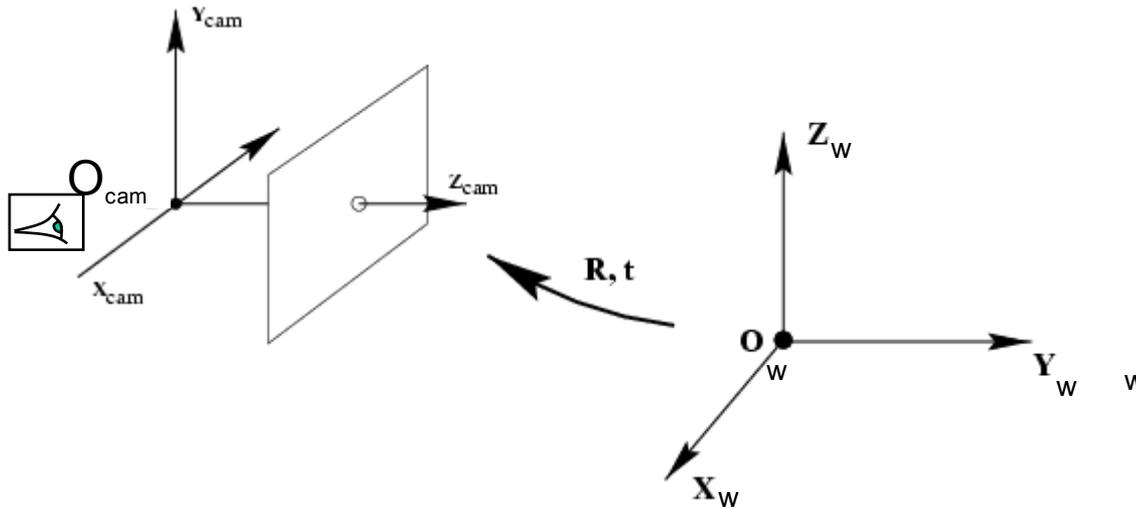
Paramètres de caméra...

...extrinsèques



...intrinsèques

...extrinsèques



$$P_C = R(P_W - t)$$

$$\begin{bmatrix} X^{(C)} \\ Y^{(C)} \\ Z^{(C)} \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

Avec $T = -Rt$

Paramètres extrinsèques :

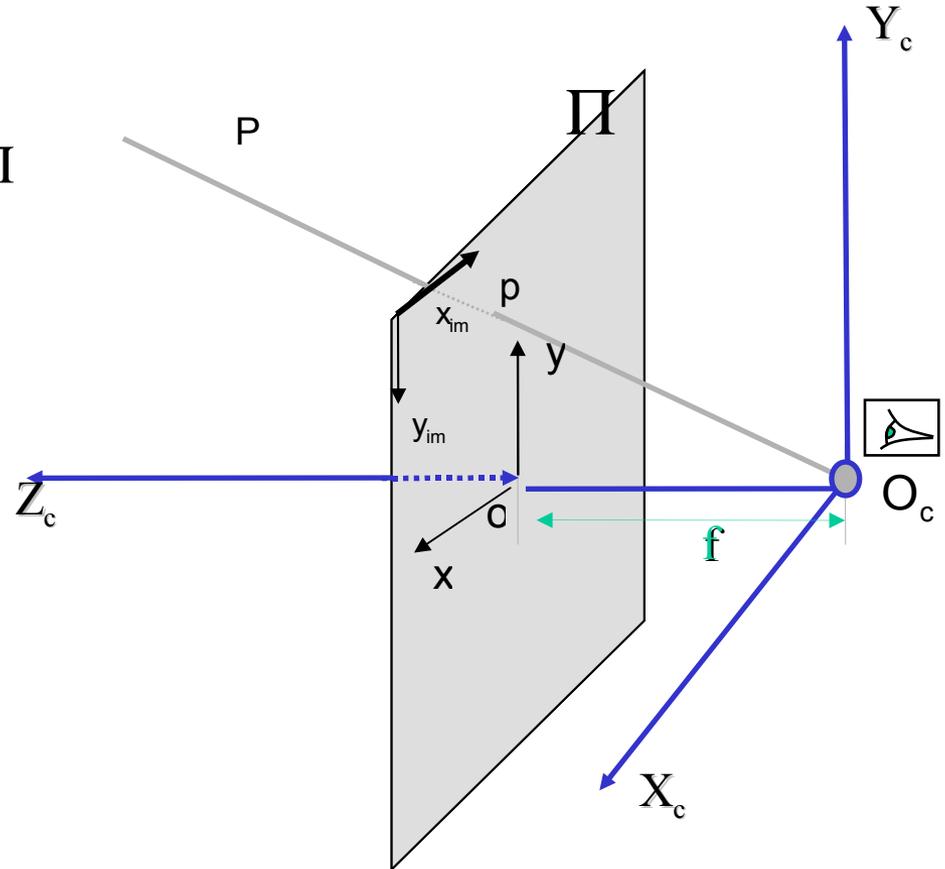
vecteur translation, T , et matrice de rotation, R (ou mieux ses paramètres libres), qui spécifient la transformation entre la caméra et le repère monde

$$\text{Let } R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \text{ and } T = \begin{bmatrix} T_X \\ T_Y \\ T_Z \end{bmatrix}$$

...intrinsèques

Caractérisation des propriétés géométriques, optiques et numériques de la caméra :

- Image Plane or Focal Plane Π
- Focal Distance f
- Camera Center O_c
- Principal Point o
- Principal Axis ($O_c Z_c$)
- Principal Plane (O_c, X_c, Y_c)
- Camera Coords (X_c, Y_c, Z_c)
- Image Coords (x_{im}, y_{im})



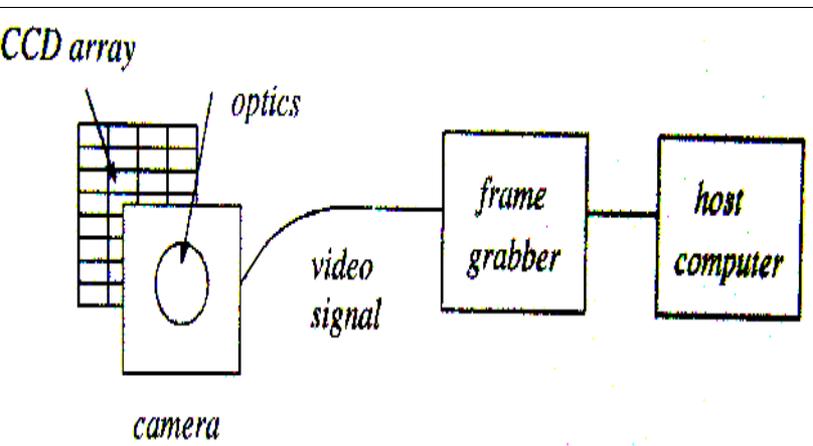
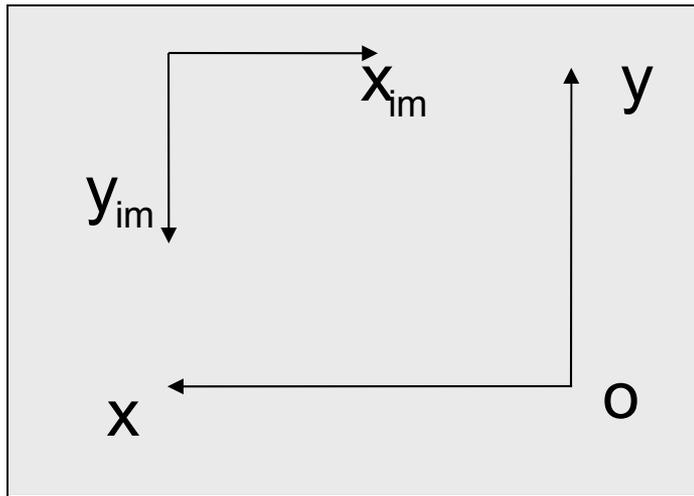
Pour un modèle de caméra perspectif, 3 ensembles de paramètres intrinsèques spécifiant :

3. La projection perspective caractérisant le passage repère caméra 3D (X, Y, Z) \rightarrow repère image projectif (x, y) :
paramètre f
5. La numérisation caractérisant le passage repère image (x, y) \rightarrow repère fenêtre CCD (x_{im}, y_{im}) :
paramètres O_x, O_y, S_x, S_y
7. La distorsion géométrique introduite par l'optique (en périphérie ou si large champ de profondeur), de nature radiale :
paramètres k_1 et k_2 .

1. Passage du repère caméra 3D (X,Y,Z) au repère image projectif (x,y) : paramètres f

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$x = \frac{u}{w} \quad \text{et} \quad y = \frac{v}{w}$$



2. Passage du repère caméra (x, y) au repère fenêtre CCD (x_{im}, y_{im}) :
paramètres o_x, o_y, s_x, s_y

En négligeant toute distorsion géométrique introduite éventuellement par l'optique et sous l'hypothèse que le tableau CCD est constitué de grilles rectangulaires d'éléments photosensibles, on a :

$$\begin{cases} x = - (x_{im} - o_x) s_x \\ y = - (y_{im} - o_y) s_y \end{cases}$$

avec (o_x, o_y) les coordonnées en pixels du centre de l'image (le point principal) et (s_x, s_y) la taille effective des pixels (en mm) dans les directions horizontales et verticales respectivement.

1. Correction de la distorsion géométrique :
paramètres k_1, k_2

$$\begin{cases} x = x_d(1+k_1r^2+k_2r^4) \\ y = y_d(1+k_1r^2+k_2r^4) \end{cases}$$

$$\text{avec } r^2 = x_d^2 + y_d^2$$

Le plus souvent, on néglige cette distorsion et sinon, on prend $k_2=0$

CCD de 500x500, lentille de
qualité moyenne, distorsion au
coin de l'image ≈ 5 pixels

Paramètres intrinsèques :

- la longueur focale f ,
- la localisation du centre de l'image en coordonnées pixelliques (o_x, o_y) ,
- la taille pixellique effective dans les directions verticales et horizontales (s_x, s_y)
- et si nécessaire le coefficient de distorsion radiale k_1 .

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X^{(c)} \\ Y^{(c)} \\ Z^{(c)} \\ 1 \end{bmatrix} \quad \text{avec}$$

$$x = u / w$$

$$y = v / w$$

or

$$\begin{bmatrix} x_{im} \\ y_{im} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{-1}{s_x} & 0 & o_x \\ 0 & \frac{-1}{s_y} & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

d'où

$$\begin{bmatrix} u^{new} \\ v^{new} \\ w \end{bmatrix} = \begin{bmatrix} \frac{-1}{s_x} & 0 & o_x \\ 0 & \frac{-1}{s_y} & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

avec

$$x = u / w \quad \text{et} \quad y = v / w$$

$$x_{im} = u^{new} / w \quad \text{et} \quad y_{im} = v^{new} / w$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -f_x & 0 & o_x & 0 \\ 0 & -f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X^{(c)} \\ Y^{(c)} \\ Z^{(c)} \\ 1 \end{bmatrix}$$

$$x_{im} = x_1 / x_3$$

$$y_{im} = x_2 / x_3$$

Intrinsic Parameters :

1. $f_x = f / s_x$

2. $f_y = f / s_y$

3. o_x

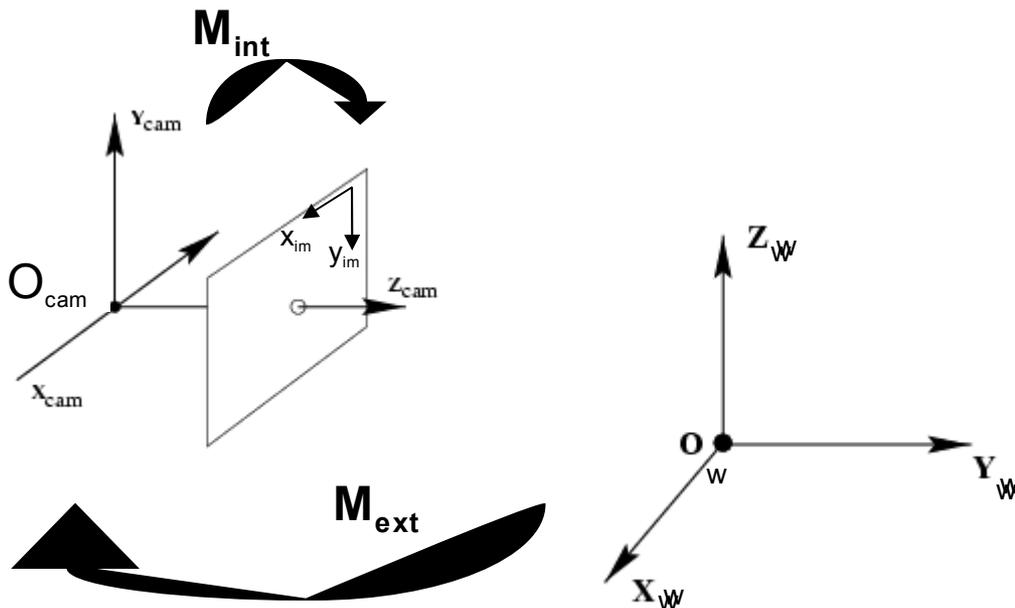
4. o_y

!! L'origine du plan image
peut être translatée par
rapport au centre de la
rétine

L'équation linéaire matricielle de la projection perspective

$$p^{(I)} = M_{\text{int}} P^{(C)}, P^{(C)} = M_{\text{ext}} P^{(W)} \Rightarrow p^{(I)} = M_{\text{int}} M_{\text{ext}} P^{(W)}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -f_x & 0 & o_x & 0 \\ 0 & -f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ 0_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$



avec

$$x_{im} = x_1 / x_3$$

$$y_{im} = x_2 / x_3$$

Images de profondeurs

Intérêts :

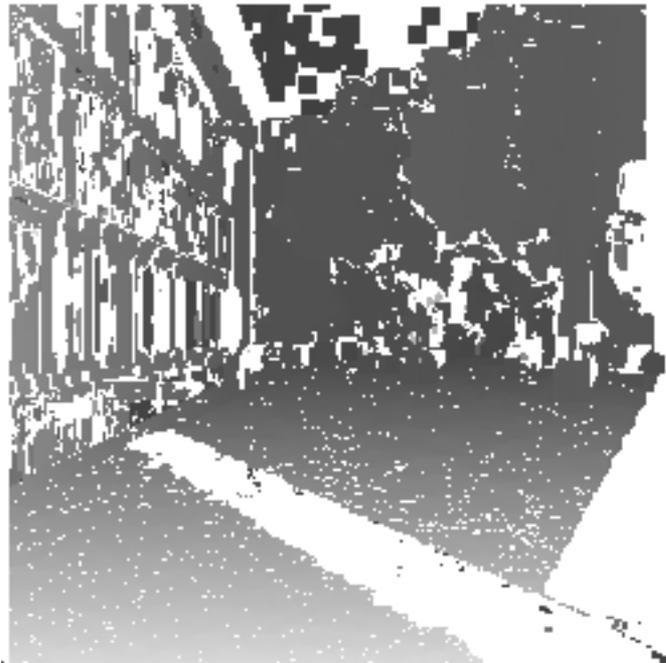
- Éviter les obstacles
- Estimer la forme des surfaces
- Inspecter des objets manufacturés

Difficile à faire partir d'une image d'intensité
-> voir suite du cours

Acquérir directement la forme : les "range sensors"
produisant des "range data"

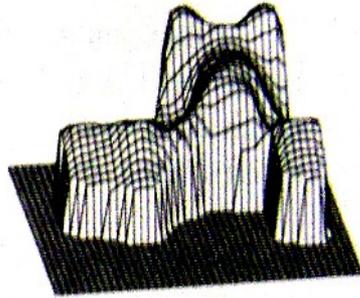
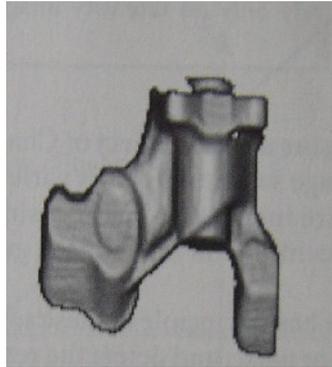
Range images :

Ce sont des cas spéciaux d'images numériques. Chaque pixel d'une image de profondeur exprime la distance entre un repère référence connu et un point visible de la scène. Ainsi, une image de profondeur reproduit la structure 3D d'une scène et est visualisable comme une surface échantillonnée.

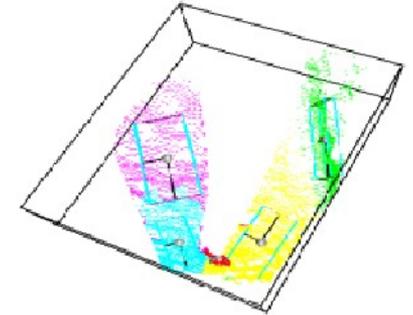
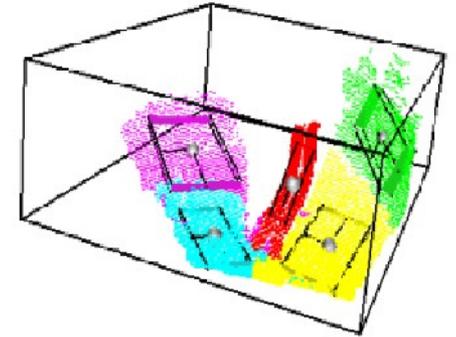


Représentation des « range data »

Sous la forme r_{ij}



Sous la forme xyz



Range image = Depth image = xyz images = surface profile = 2.5D images

Active vs. Passive range sensors

Les capteurs de profondeurs **passifs** reposent seulement sur des images d'intensités pour reconstruire la profondeur (ex. : stéréoscopie). -> *voir cours plus loin*

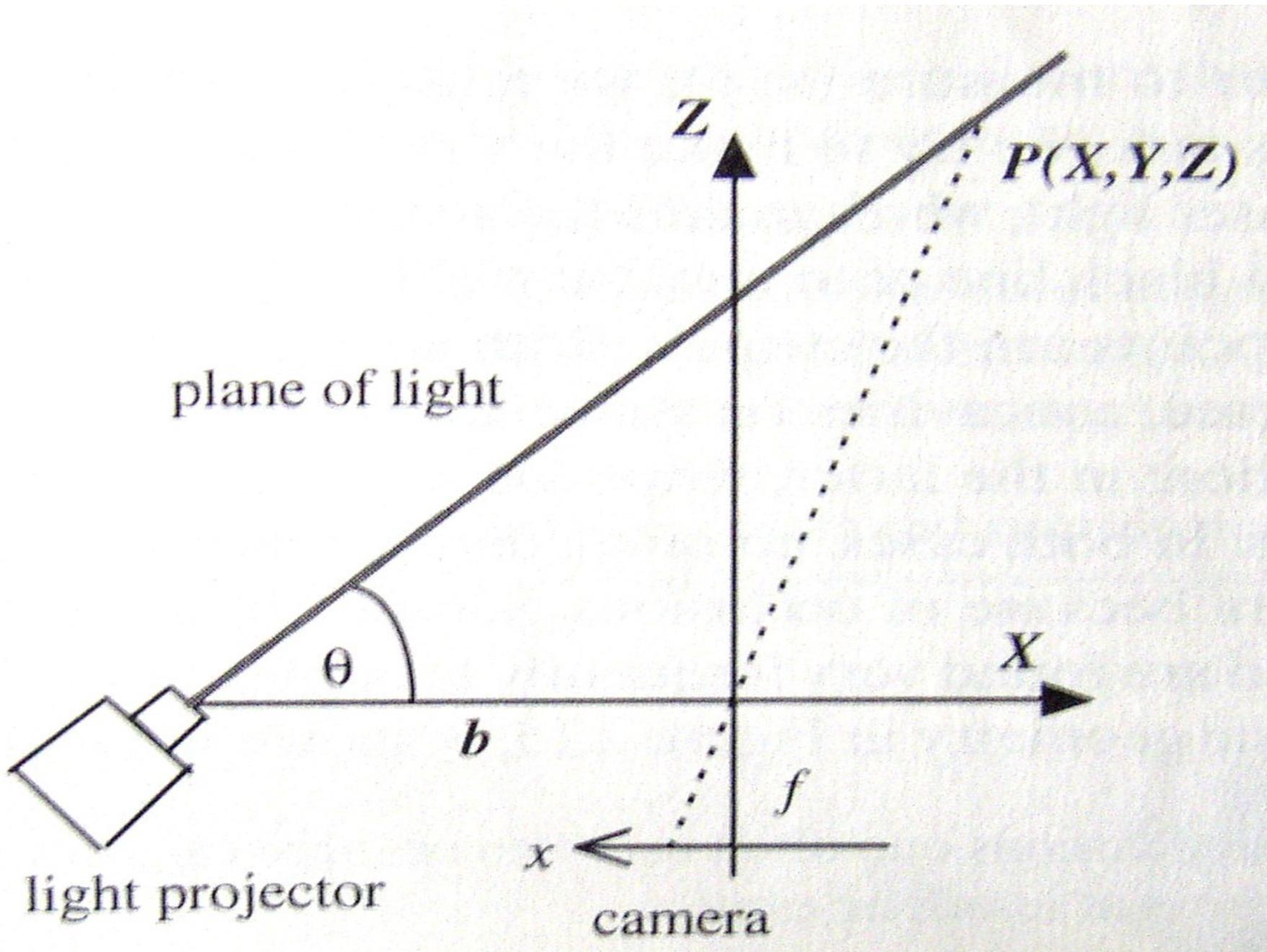
Les capteurs de profondeurs **actifs** projettent de l'énergie (ex. : un motif de lumière, des impulsions sonars) sur la scène et détecte sa position pour réaliser la mesure ou exploitent l'effet de changements contrôlés de certains paramètres du capteur (ex. : focus).

- Radar et sonar
- Interférométrie de Moiré
- Focalisation/Défocalisation active
- Triangulation active**



Systemes d'acquisition par Triangulation active

- basés sur caméras d'intensité
- OUTPUT : carte dense de coordonnées 3D précises
- faciles à comprendre et à construire
 - Projecteur
 - Caméra vidéo
 - Repère monde = Repère caméra
 - Plan lumière \perp (XZ)
 - Angle (Plan lumière, (XY)) = θ : paramètre de balayage
 - Plan lumière \cap Surface Objet = Stripe (courbe planaire) observée par la caméra



Pour chaque point visible de la bande,

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{b}{f \cos \theta - x} \begin{bmatrix} x \\ y \\ f \end{bmatrix}$$

-> Profile 3D des points de la surface ou cross-section

En faisant varier θ ,

-> ensemble de cross-sections 3D

-> range image complète de l'objet

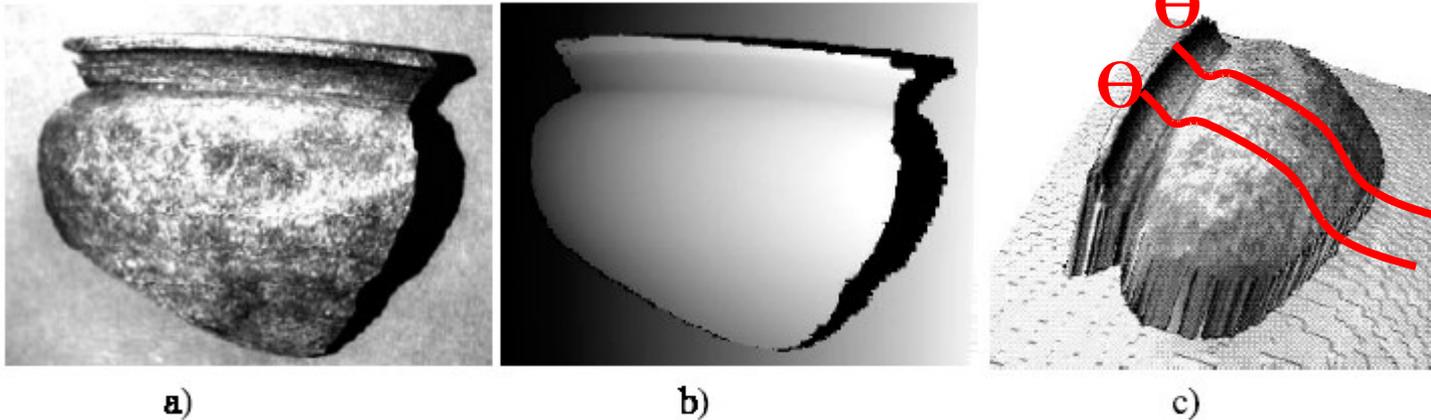


Figure 3 Fragment: a) intensity image, b) range image, c) object model

Comment rendre très visible la bande dans l'image :

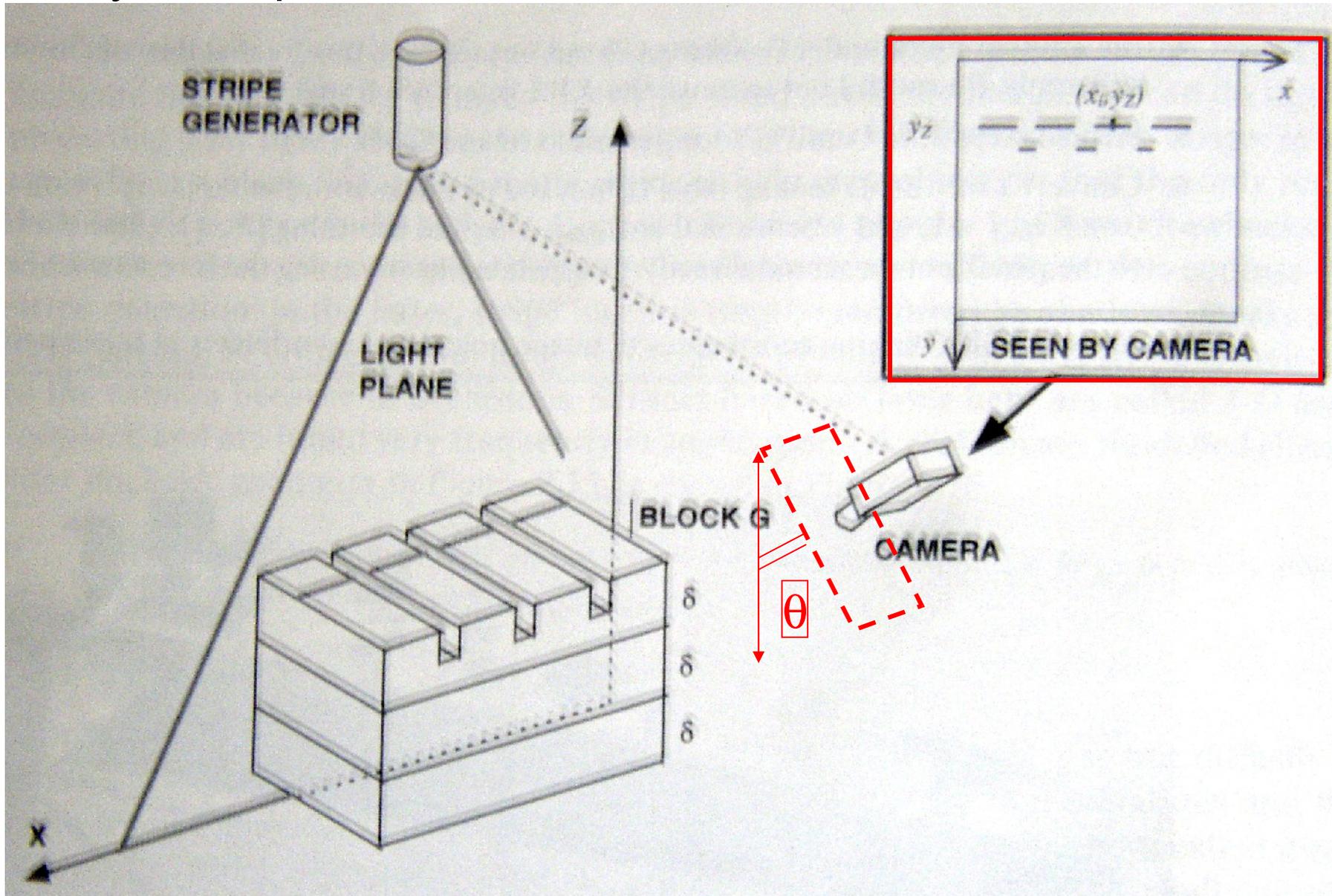
- ligne claire : laser light : 3D laser scanner
- ligne noire sur objet clair

Comment éviter les occlusions ?

Nécessiter de calibrer le système en f , b , et θ



Un système par calibration directe ?



Soit on procède à une calibration de f, b et θ comme précédemment,
Soit on calibre directement, simplement mais manuellement, sans équation

Construction d'une LUT (look-up table) liant une image et ses coordonnées 3D, en mesurant les coordonnées image d'une grille de points 3D connus, et en stockant à la fois les coordonnées image et monde pour chaque point.

Les valeurs de profondeurs de tous les autres points sont obtenues par interpolation.

La procédure utilise quelques blocs rectangulaires de hauteur connue δ . Un bloc G doit avoir un nombre n de rainures parallèles de forme rectangulaires. On suppose enfin que la taille de l'image en pixels est $x_{max} \times y_{max}$

Procédure RANGE_CAL

INPUT : Installer le système comme décrit. Sans objet dans la scène, la raie de lumière verticale tombe sur l'arrière-plan $Z=0$ et devrait être imagée près de $y=y_{\max}-1$

3. Placer le bloc G sous la raie, avec la rainure perpendiculaire au plan de lumière. S'assurer que la raie apparaît parallèle à l'axe x (à y constant).
4. Acquérir une image de la raie tombant sur G. Trouver les coordonnées en y des points de la raie tombant sur la plus haute surface de G (cad pas dans la rainure) en scannant les colonnes de l'image.
5. Calculer les coordonnées $[x_i, y_z]^T$, $i=1, \dots, n$, des centres des segments de la raie sur la surface supérieure de G, en prenant les centres des segments sur la ligne de balayage $y=y_z$. Entrer chaque point image $[x_i, y_z]^T$ et ses points 3D $[X, Z]^T$ correspondants (connus) dans une table T.
6. Mettre un autre bloc sous G, élevant la surface supérieure de G de δ . S'assurer que les conditions de l'étape 1 sont toujours valides. Faire attention de ne pas bouger le plan de référence XYZ.
7. Répéter les étapes 2,3, 4 jusqu'à ce que la surface supérieure de G soit imagée près de $y=0$.
8. Convertir T en une "lookup table" 2D L, indexée par les coordonnées image $[x, y]^T$, avec x entre 0 et $x_{\max}-1$, et y entre 0 et $y_{\max}-1$, et retournant $[X, Z]^T$. Pour associer des valeurs aux pixels non mesurés directement, interpoler linéairement en utilisant les quatre plus proches voisins.

OUTPUT : une LUT reliant les coordonnées des points image et les coordonnées des points de la scène.

Algorithme RANGE_ACQ

INPUT : LUT L construite par RANGE_CAL

3. Placer un objet G sous la raie et acquérir une image de la raie tombant sur G.
4. Calculer les coordonnées image $[x,y]^T$ des points de la raie en balayant chaque colonne image.
5. Indexer L en utilisant les coordonnées images $[x,y]^T$ de point de la raie pour obtenir sa profondeur $[X,Z]^T$

OUTPUT : ensemble de coordonnées 3D correspondants aux points de la raie imagée = "range profile"

Remarque : quand un bloc est ajouté à la scène de calibration, la raie doit s'élever d'au moins 1 ou 2 pixels, sinon, la calibration ne discriminera pas entre les niveaux Z.

Bibliographie

- “Basic Math for 16-720”, Martial Hebert, (*mathprimer.pdf et mathfigures.pdf*)
- “Trinocular Active Range Sensing”, A. Blake et al., IEEE PAMI, vol. 15, pp. 477-483, 1993 (Bibliothèque Informatique-Recherche -Jussieu - site du Capitaine Scott)
- “Shape from Focus”, S.K. Nayar et al., IEEE PAMI, vol. 16, pp. 824-831, 1994 (Bibliothèque Informatique-Recherche -Jussieu - site du Capitaine Scott)
- “3-D Digitizers”, T. Wohlers, Computer Graphics World, pp. 73-77, 1992

MASTER IPCC



Vision par Ordinateur

Nicolas Loménie

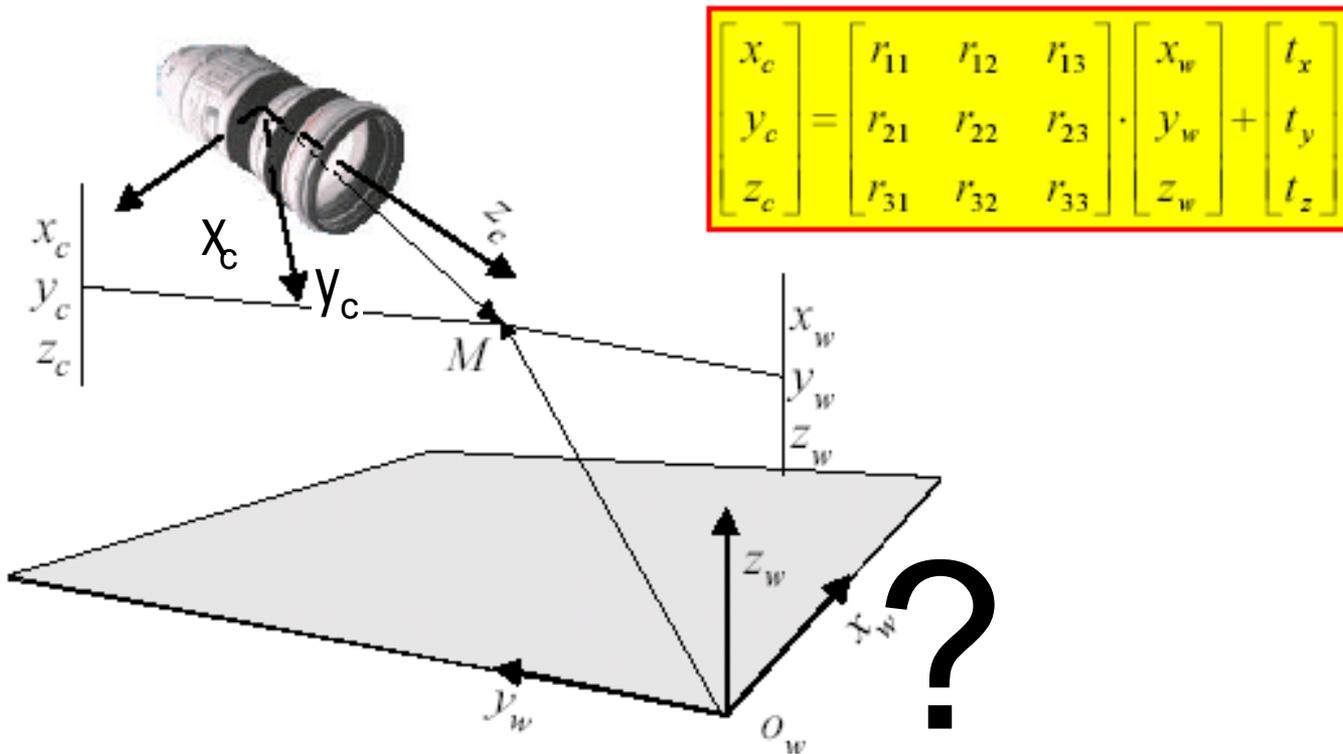
CALIBRATION

CALIBRATION

Objectifs :

- Redéfinir les paramètres extrinsèques et intrinsèques d'un capteur passif
- Estimer ces paramètres un par un en utilisant la caméra comme instrument de mesure
- Apprendre la résolution d'équations par décomposition SVD
- Estimer la matrice de projection directement

- Extrinsic Parameters



$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \cdot \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

+ intrinsic parameters \rightarrow

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} \\ 0_{1 \times 3} \end{bmatrix} \begin{bmatrix} T_{3 \times 1} \\ 1 \end{bmatrix} \begin{bmatrix} X^{(w)} \\ Y^{(w)} \\ Z^{(w)} \\ 1 \end{bmatrix}$$

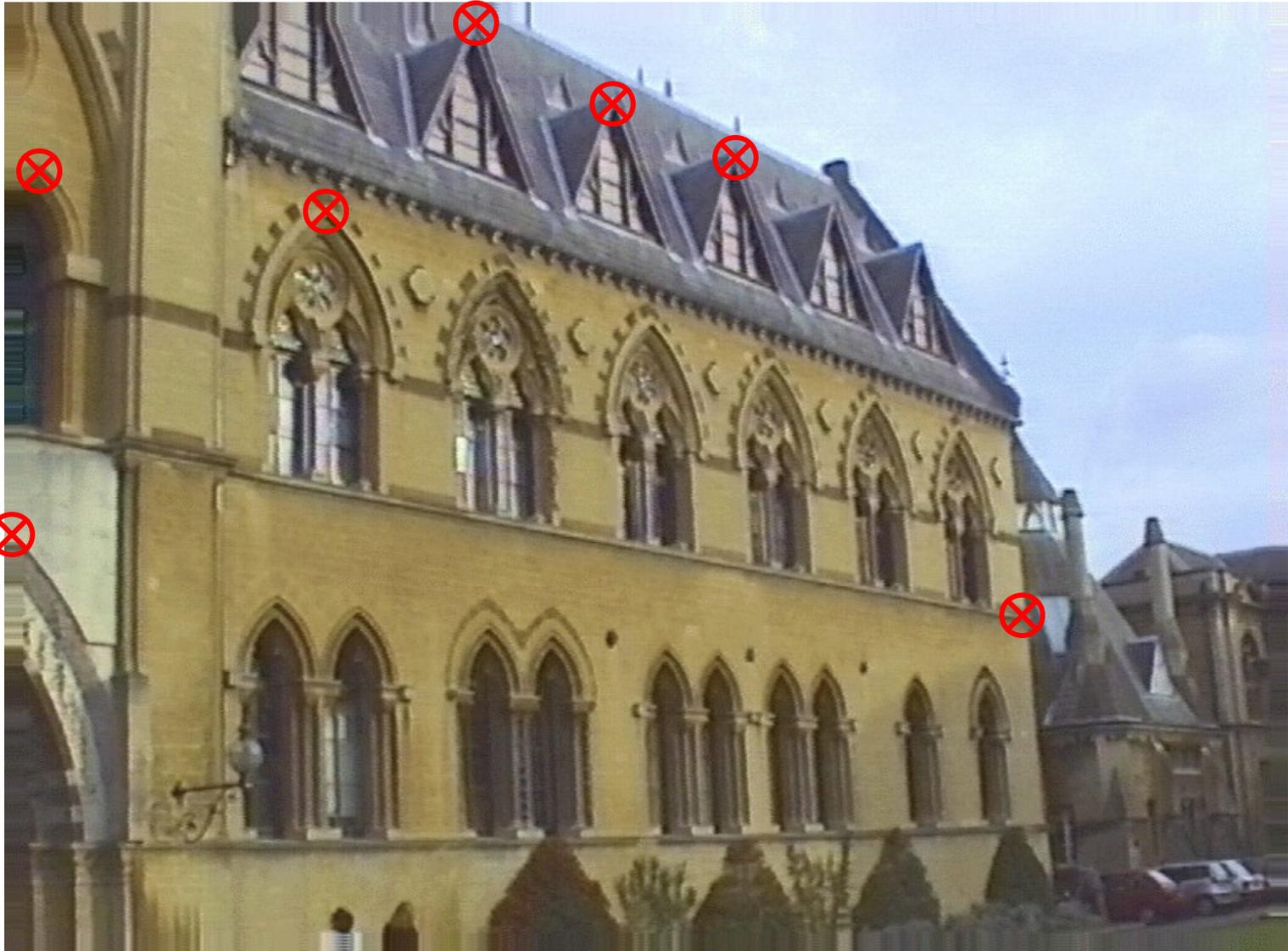
Problème de la calibration

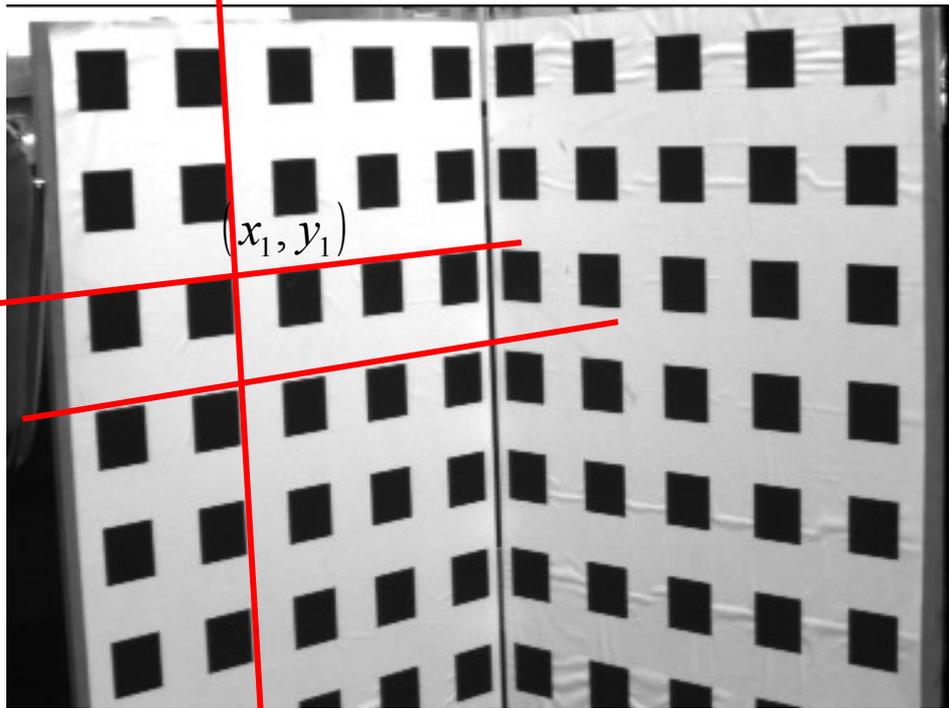
Étant donné une ou plusieurs images d'un motif de calibration, estimer :

1. Les paramètres intrinsèques
2. Les paramètres extrinsèques
3. Les deux

-> la caméra est utilisée comme instrument de mesure

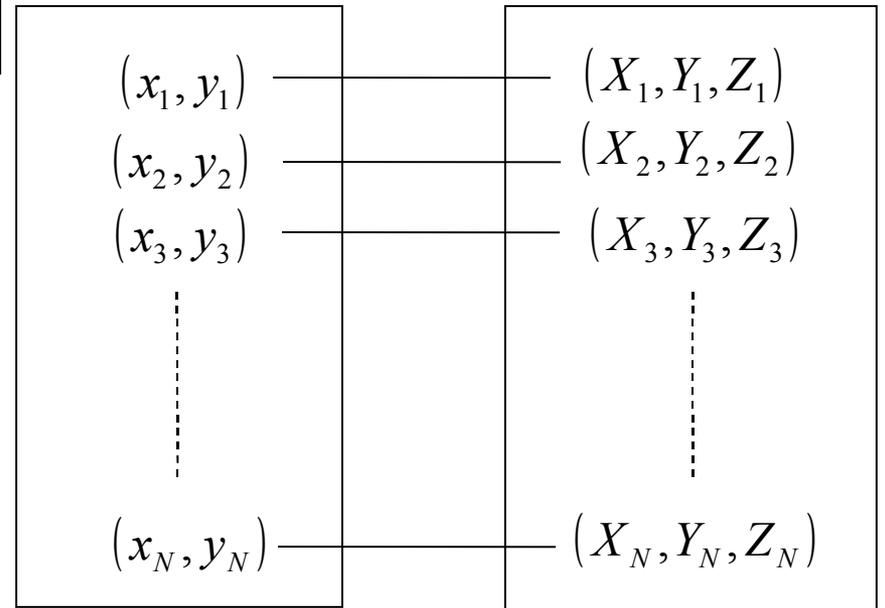
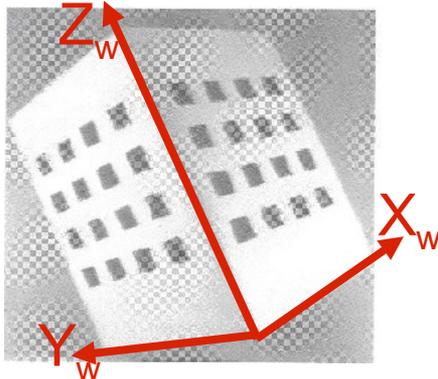
-> pas indispensable mais ouvre la voie à des algorithmes de reconstruction et reconnaissance 3D performants





$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Motif de calibration



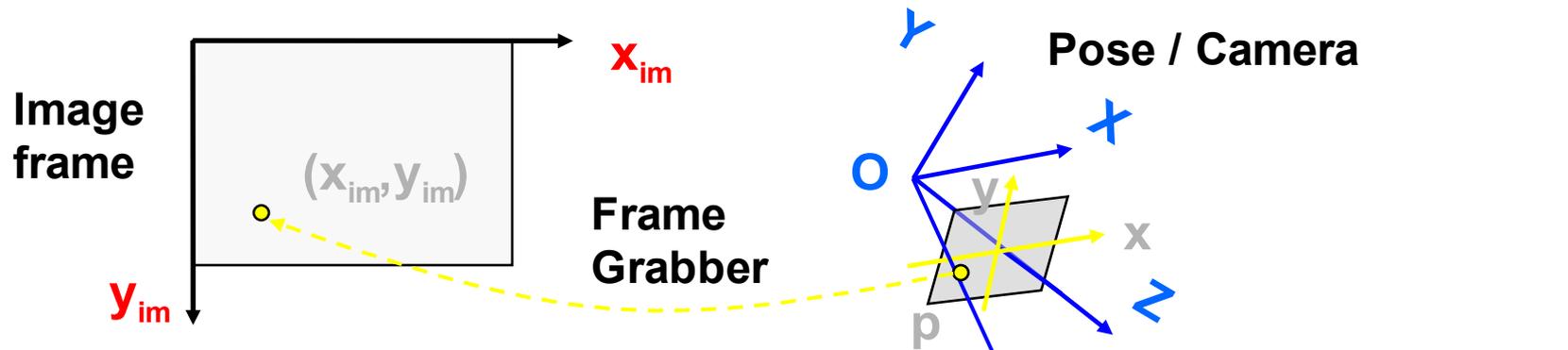
2 méthodes reconnues :

1. par estimation directe des paramètres

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -f_x & 0 & o_x & 0 \\ 0 & -f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

8. par matrice de projection

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$



- **Systemes de coordonnees :**

- Coordonnees fenetre (x_{im}, y_{im}) pixels
- Coordonnees image (x, y) in mm
- Coordonnees camera (X, Y, Z)
- Coordonnees monde (X_w, Y_w, Z_w)

- **Parametres de camera :**

- Parametres intrinseques (de la camera et du "frame grabber") : relie les "frame coordinates" d'un point image point a ses "camera coordinates" correspondantes
- Parametres extrinseques : definissent la localisation et l'orientation du "camera coordinate system" par rapport au "world coordinate system"

- World to Camera

- Camera: $\mathbf{P} = (X, Y, Z)^T$
- World: $\mathbf{P}_w = (X_w, Y_w, Z_w)^T$
- Transformations : R, T

$$\mathbf{P} = \mathbf{R}\mathbf{P}_w + \mathbf{T} = \begin{pmatrix} r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x \\ r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y \\ r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z \end{pmatrix} = \begin{bmatrix} \mathbf{R}_1^T \mathbf{P}_w + T_x \\ \mathbf{R}_2^T \mathbf{P}_w + T_y \\ \mathbf{R}_3^T \mathbf{P}_w + T_z \end{bmatrix}$$

- Camera to Image

- Camera: $\mathbf{P} = (X, Y, Z)^T$
- Image: $\mathbf{p} = (x, y)^T$
- Equations non linéaires

$$(x, y) = \left(f \frac{X}{Z}, f \frac{Y}{Z} \right)$$

- Image to Frame

- Distorsion négligée
- Frame $(x_{im}, y_{im})^T$

$$x = -(x_{im} - o_x)s_x$$

$$y = -(y_{im} - o_y)s_y$$

- World to Frame

- $(X_w, Y_w, Z_w)^T \rightarrow (x_{im}, y_{im})^T$
- Focales effectives
 - $f_x = f/s_x, f_y = f/s_y$

$$x_{im} - o_x = -f_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y_{im} - o_y = -f_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$x_{im} - o_x = -f_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y_{im} - o_y = -f_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

- Ce qui est inaccessible disparaît : le repère caméra
- On se fonde sur ce qui est mesurable : x , y , X , Y et Z
- On compte sur un nombre minimum de mesures pour résoudre un **système d'équations linéaire** ... 
- 2 étapes :
 1. On suppose connues **(o_x, o_y)** , le centre de l'image, et on estime tous les autres paramètres
 2. Puis on calcule **(o_x, o_y)** .

$$x' = x_{im} - o_x = -f_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y' = y_{im} - o_y = -f_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

- Paramètres Extrinsèques :
 - \mathbf{R} , matrice de rotation 3x3
 - Trois angles α, β, γ
 - \mathbf{T} , vecteur translation 3-D

- Paramètres Intrinsèques :
 - $\alpha = f_x/f_y = s_y/s_x$, et f_x
 - (o_x, o_y) : on suppose le **centre de l'image connu** -> (x', y') connues
 - k_1 , coefficient de distorsion radiale : **négligé dans l'algorithme de base**

Étape 1

Supposant que la localisation du centre de l'image (o_x, o_y) est connu, et que la distorsion radiale peut être négligée, on doit estimer f_x , α , R et T à partir des points images (x_i, y_i) , $i = 1$ à N , projections des N points connus (X_i, Y_i, Z_i) dans le référentiel monde.

Même dénominateur dans les deux équations

$$f_y(r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y) / y' = f_x(r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x) / x'$$



$$x' f_y(r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y) = y' f_x(r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x)$$



$$x'(r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y) = y' \alpha (r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x)$$

Pour chaque paire de points $\{(X_i, Y_i, Z_i), (x_i, y_i)\}$



$$x_i X_i r_{21} + x_i Y_i r_{22} + x_i Z_i r_{23} + x_i T_y - y_i X_i (\alpha r_{11}) - y_i Y_i (\alpha r_{12}) - y_i Z_i (\alpha r_{13}) - y_i (\alpha T_x) = 0$$



Équation linéaire à 8 inconnus $\mathbf{v} = (v_1, \dots, v_8)^T$

$$x_i X_i v_1 + x_i Y_i v_2 + x_i Z_i v_3 + x_i v_4 - y_i X_i v_5 - y_i Y_i v_6 - y_i Z_i v_7 - y_i v_8 = 0$$

$$(v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8)$$

$$= (r_{21}, r_{22}, r_{23}, T_y, \alpha r_{11}, \alpha r_{12}, \alpha r_{13}, \alpha T_x)$$

Pour N points, on obtient un système linéaire homogènes à 8 inconnus et N équations.

$$x_i X_i v_1 + x_i Y_i v_2 + x_i Z_i v_3 + x_i v_4 - y_i X_i v_5 - y_i Y_i v_6 - y_i Z_i v_7 - y_i v_8 = 0$$

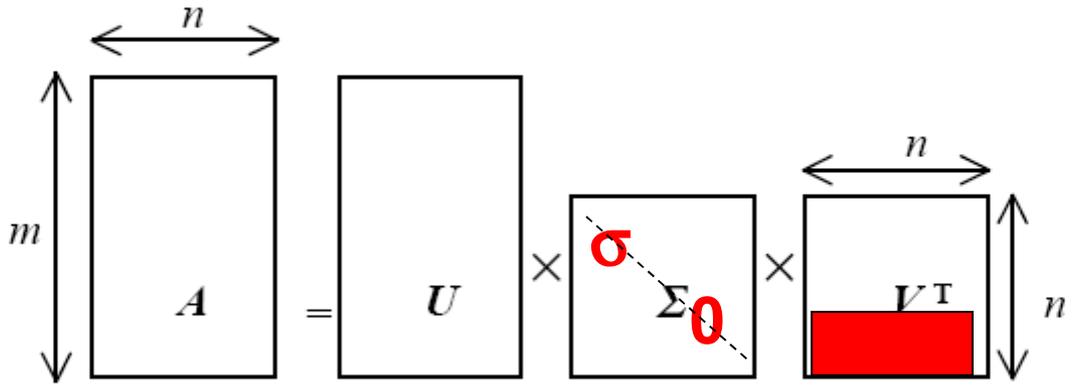


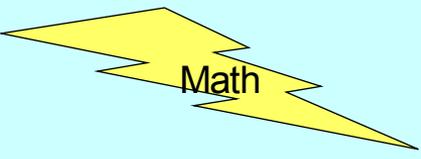
$$\mathbf{A} \mathbf{v} = \mathbf{0}$$

$$\mathbf{A} = \begin{bmatrix} x_1 X_1 & x_1 Y_1 & x_1 Z_1 & x_1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 \\ x_2 X_2 & x_2 Y_2 & x_2 Z_2 & x_2 & -y_2 X_2 & -y_2 Y_2 & -y_2 Z_2 & -y_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ x_N X_N & x_N Y_N & x_N Z_N & x_N & -y_N X_N & -y_N Y_N & -y_N Z_N & -y_N \end{bmatrix}$$

Le système admet une solution non triviale (à un facteur d'échelle près) :

- Si $N \geq 7$ et les N points sont non coplanaires $\Rightarrow \text{Rank}(\mathbf{A}) = 7$
 - preuve dans le livre de Faugeras
- Que l'on trouve grâce la décomposition SVD de $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ par exemple
 - \mathbf{v} est le vecteur colonne de \mathbf{V} correspondant à la seule valeur propre nulle singulière le long de la diagonale de \mathbf{D}
 - Matlab, Scilab
 - Lignes de \mathbf{V}^T : vecteurs propres $\{e_i\}$ de $\mathbf{A}^T\mathbf{A}$
 - Solution: $\mathbf{v} = \mathbf{e}_8$, la 8^{ème} ligne correspondant à la seule valeur singulière nulle $\sigma_8 = 0$
 - Considération pratique : les erreurs de mesures peuvent rendre $\text{rang}(\mathbf{A}) = 8$, on considère alors la seule valeur propre singulière la plus faible le long de la diagonale de $\mathbf{\Sigma}$





Décomposition en Valeurs Singulières

Singular **V**alue **D**ecomposition

Cette décomposition s'applique à une matrice rectangulaire.

Elle permet 3 applications importantes :

- Résoudre des systèmes d'équations linéaires non homogènes,
- Résoudre des des systèmes d'équations linéaires homogènes de rang déficient,
- Garantir que les entrées d'une matrice extraite numériquement satisfont des contraintes données (orthogonalité par exemple).

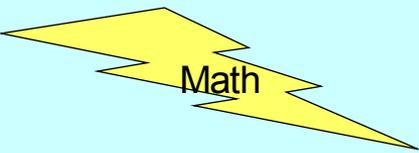
Singular Value Decomposition

Théorème :

Toute matrice $A \in M_{m \times n}$ peut s'écrire comme le produit de 3 matrices : $A = UDV^T$ telles que :

- $U \in M_{m \times n}$ et $V \in M_{n \times n}$ et leurs colonnes correspondent à des vecteurs unités mutuellement orthogonaux;
- $D \in M_{n \times n}$ est une matrice diagonale dont les éléments diagonaux, σ_i , appelés valeurs singulières, sont tels que $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

Remarque : alors que U et D ne sont pas uniques, les valeurs singulières sont complètement déterminées par A



Singular Value Decomposition

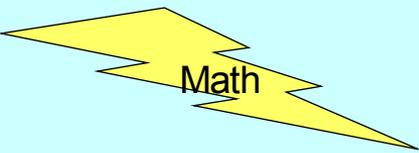
Propriété 1 :

Une matrice carrée A est non singulière ssi toutes les valeurs singulières sont différentes de 0.

Définition : conditionnement

Le ratio $c = \sigma_1 / \sigma_n$ appelé conditionnement mesure le degré de singularité de A .

Remarque : quand $1/c$ est comparable à la précision arithmétique de votre machine, la matrice A est **mal conditionnée** et de façon pratique peut être considérée comme singulière et donc non inversible.



Singular Value Decomposition

Propriété 2 :

Si A est une matrice rectangulaire, le nombre de σ_i non nuls égale le rang de A .

Remarque : en pratique on utilisera une tolérance $\varepsilon \approx 10^{-6}$

Propriété 3 :

Si A est une matrice carrée, non singulière, son inverse peut s'écrire $A^{-1} = V D^{-1} U^T$

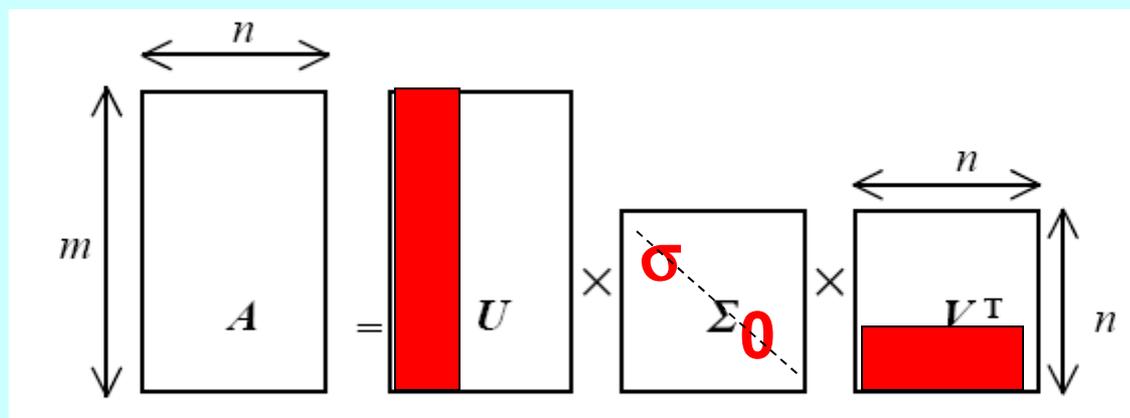
Remarque : que A soit singulière ou non, la pseudo-inverse de A , A^+ peut s'écrire $A^+ = V D_0^{-1} U^T$ avec D_0^{-1} égale à D^{-1} pour toutes valeurs singulières non nulles et zéro sinon.

Si A est non singulière, $D_0^{-1} = D^{-1}$ et $A^+ = A^{-1}$

Singular Value Decomposition

Propriété 4 :

Les colonnes de U correspondant aux valeurs singulières non nulles décrivent l'image de A , les colonnes de V correspondant aux valeurs singulières nulles décrivent le noyau de A .

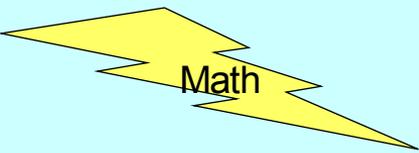


Propriété 5 :

Les carrés des valeurs singulières non nulles sont les valeurs propres non nulles de $A^T A \in M_{n \times n}$ et de $A A^T \in M_{m \times m}$.

Les colonnes de U sont les vecteurs propres de $A A^T$ et celles de V ceux de $A^T A$.

De plus, $A u_k = \sigma_k v_k$ et $A^T v_k = \sigma_k u_k$ avec u_k et v_k les colonnes de U et V correspondantes à σ_k .



$$\boxed{A} \quad \boxed{x} = \boxed{b}$$

Singular Value Decomposition

Première Application : Résolution de systèmes d'équations linéaires non homogènes

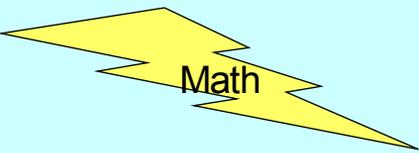
Soit le système de m équations linéaires à n inconnues $Ax=b$

- $A \in M_{m \times n}$ la matrice du système linéaire non homogène
- $x \in \mathfrak{R}_n$ le vecteur des inconnues
- $b \in \mathfrak{R}_m$ le vecteur de données

$$\text{Si } b \neq 0, \quad Ax = b \Leftrightarrow A^T Ax = A^T b$$

$$Ax = b \Leftrightarrow x = \left(A^T A \right)^+ A^T b$$

Cette solution est optimale aux sens des moindres carrés. De plus $A^T A$ est carrée. On calcule $(A^T A)^+$ grâce à la SVD. Dans le cas de plus d'équations que d'inconnues, la pseudo-inverse $(A^T A)^+$ est en effet plus apte à coïncider avec $(A^T A)^{-1}$.



$$A \cdot x = 0$$

Singular Value Decomposition

Deuxième Application :
Résolution de systèmes d'équations linéaires homogènes de rang déficient

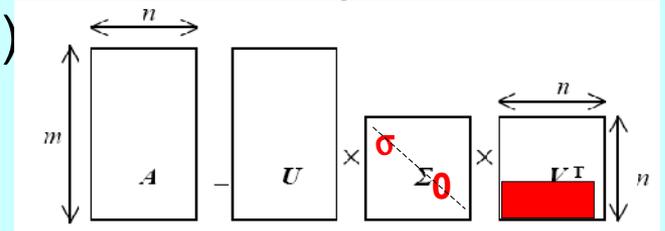
Soit le système de m équations linéaires à n inconnues $Ax=0$

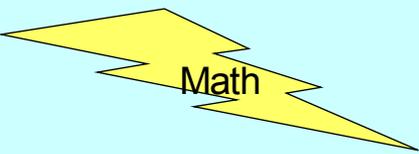
- $A \in M_{m \times n} / \text{rang}(A) = n-1$ et $m \geq n-1$
- $x \in \mathcal{R}_n$ le vecteur des inconnues

Si on ignore la solution triviale, une solution unique à un facteur d'échelle près peut être trouvée grâce à la SVD.

Cette solution est simplement proportionnelle au vecteur propre correspondant à la seule valeur propre nulle de $A^T A$ (toutes les autres valeurs propres étant >0 car $\text{rg}(A)=n-1$)

Propriété 4 et 5 \rightarrow cette solution est la colonne de V correspondant à la seule valeur propre nulle de A (le noyau de A)



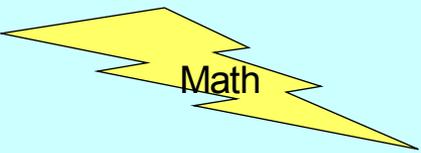


Singular Value Decomposition

Dernière Application :
Assurer des contraintes

On génère souvent des estimées d'une matrice A , dont les entrées ne sont pas toutes indépendantes, mais satisfont des contraintes algébriques (voire les matrices orthogonales).

Or, les erreurs produites par le bruit et les altérations dues aux calculs numériques abîment la matrice \hat{A} , si bien que ses entrées ne satisfont plus les contraintes données \rightarrow problème par la suite
La SVD permet de trouver la matrice la plus proche de A au sens de la norme de Frobenius et qui satisfasse les contraintes exactement.



Singular Value Decomposition

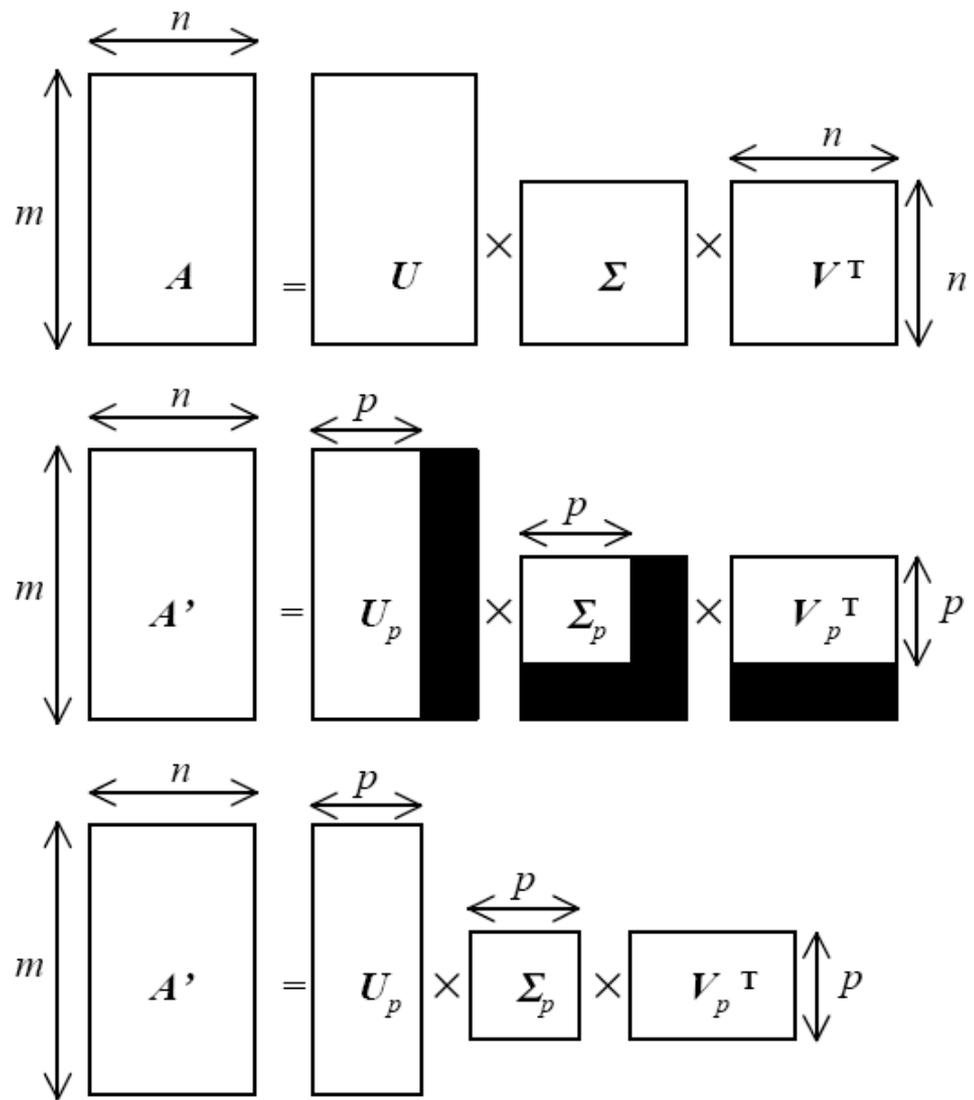


Figure 2. A simple way to construct a rank p matrix: Compute the SVD (top); Set the smallest $n-p$ singular values to zero (middle); Remove the corresponding entries from the matrices (bottom).

Scale Factor and Aspect Ratio

- Equations for scale factor γ and aspect ratio α

$$\bar{\mathbf{v}} = \gamma (r_{21}, r_{22}, r_{23}, T_y, \alpha r_{11}, \alpha r_{12}, \alpha r_{13}, \alpha T_x)$$

$$\mathbf{v}_1 \quad \mathbf{v}_2 \quad \mathbf{v}_3 \quad \mathbf{v}_4 \quad \mathbf{v}_5 \quad \mathbf{v}_6 \quad \mathbf{v}_7 \quad \mathbf{v}_8$$

- Knowledge: \mathbf{R} is an orthogonal matrix

$$\mathbf{R}_i^T \mathbf{R}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad \mathbf{R} = (r_{ij})_{3 \times 3} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1^T \\ \mathbf{R}_2^T \\ \mathbf{R}_3^T \end{bmatrix}$$

- Second row ($i=j=2$):

$$r_{21}^2 + r_{22}^2 + r_{23}^2 = 1 \quad \Rightarrow \quad |\gamma| = \sqrt{\bar{v}_1^2 + \bar{v}_2^2 + \bar{v}_3^2} \quad \Rightarrow \quad |\gamma|$$

- First row ($i=j=1$)

$$r_{11}^2 + r_{12}^2 + r_{13}^2 = 1 \quad \Rightarrow \quad \alpha |\gamma| = \sqrt{\bar{v}_5^2 + \bar{v}_6^2 + \bar{v}_7^2}$$

} α

Rotation R and Translation T

- Equations for first 2 rows of R and T given α and $|\gamma|$

$$\bar{\mathbf{V}} = s |\gamma| (r_{21}, r_{22}, r_{23}, T_y, \alpha r_{11}, \alpha r_{12}, \alpha r_{13}, \alpha T_x)$$

- First 2 rows of R and T can be found up to a common sign **s (+ or -)**

$$s\mathbf{R}_1^T, s\mathbf{R}_2^T, sT_x, sT_y$$

- The third row of the rotation matrix by vector product

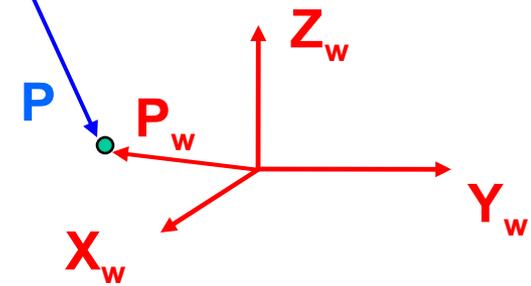
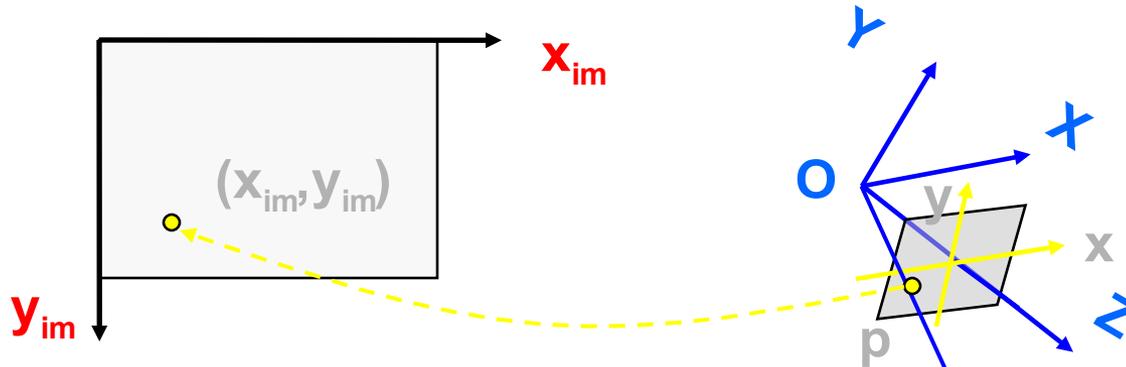
$$\mathbf{R}_3^T = \mathbf{R}_1^T \times \mathbf{R}_2^T = s\mathbf{R}_1^T \times s\mathbf{R}_2^T$$

- Remaining Questions :

$$\mathbf{R} = (r_{ij})_{3 \times 3} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1^T \\ \mathbf{R}_2^T \\ \mathbf{R}_3^T \end{bmatrix}$$

- How to find the sign **s**?
- Is **R** orthogonal?
- How to find **T_z** and **f_x**?

Find the sign s



- Facts:

- $f_x > 0$
- $Z_c > 0$
- x known
- X_w, Y_w, Z_w known

- Solution

- \Rightarrow Check the sign of X_c
- \Rightarrow Should be opposite to x

$$x = -f_x \frac{X_c}{Z_c} = -f_x \frac{r_{11}X_w + r_{12}Y_w + r_{13}Z_w + T_x}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

$$y = -f_y \frac{Y_c}{Z_c} = -f_y \frac{r_{21}X_w + r_{22}Y_w + r_{23}Z_w + T_y}{r_{31}X_w + r_{32}Y_w + r_{33}Z_w + T_z}$$

Rotation R : orthogonality

- Question:
 - First 2 rows of R are calculated without using the mutual orthogonal constraint

$$\mathbf{R} = (r_{ij})_{3 \times 3} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1^T \\ \mathbf{R}_2^T \\ \mathbf{R}_3^T \end{bmatrix}$$

$$\hat{\mathbf{R}}^T \hat{\mathbf{R}} = \mathbf{I}?$$

$$\mathbf{R}_3^T = \mathbf{R}_1^T \times \mathbf{R}_2^T = s\mathbf{R}_1^T \times s\mathbf{R}_2^T$$

- Solution:
 - Use SVD of estimate R

$$\hat{\mathbf{R}} = \mathbf{U}\mathbf{D}\mathbf{V}^T \quad \longrightarrow \quad \mathbf{R} = \mathbf{U}\mathbf{I}\mathbf{V}^T$$



Replace the diagonal matrix D with the 3x3 identity matrix

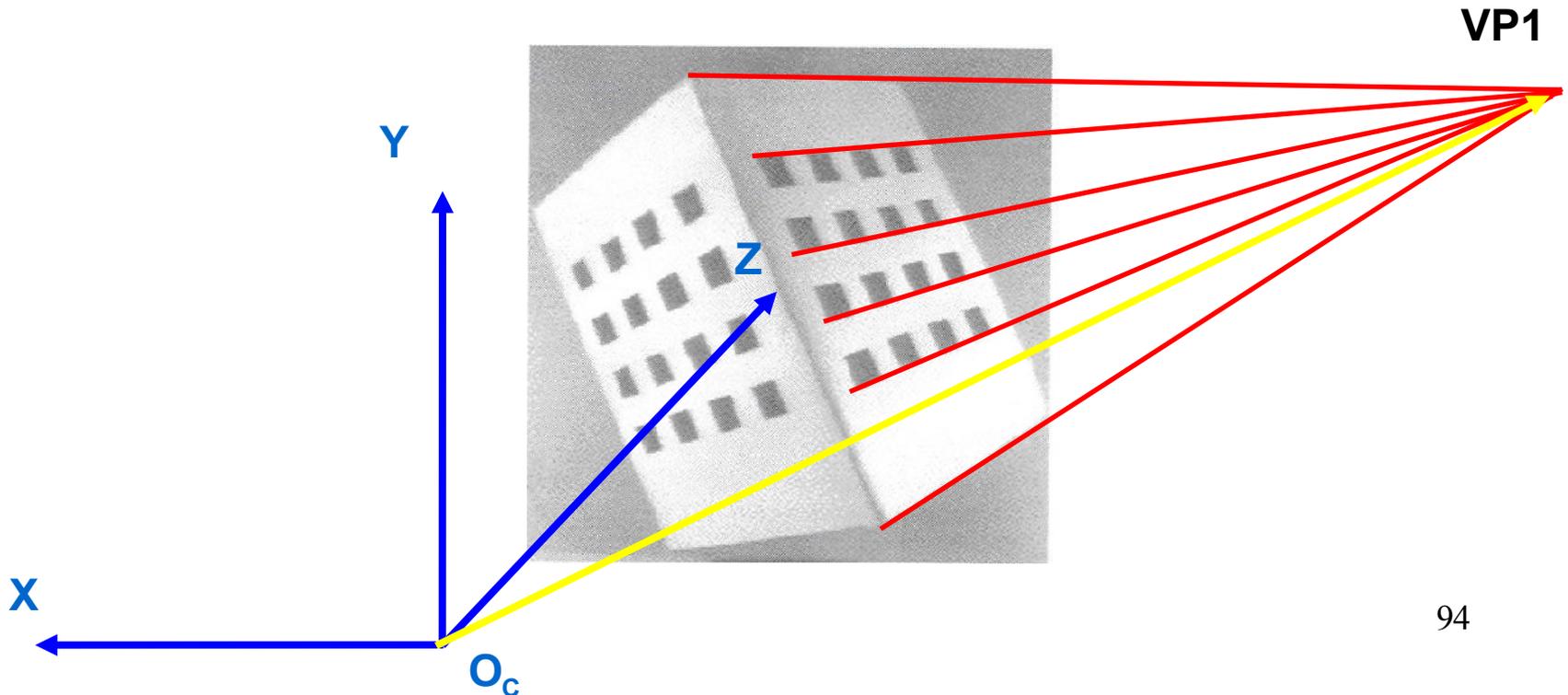
Étape 2

Estimation du centre de l'image (o_x, o_y)

Remarque : la précision de l'estimation du centre o de l'image n'influe pas sur les autres paramètres mais est importante pour d'autres applications (telle que la modélisation du mouvement)

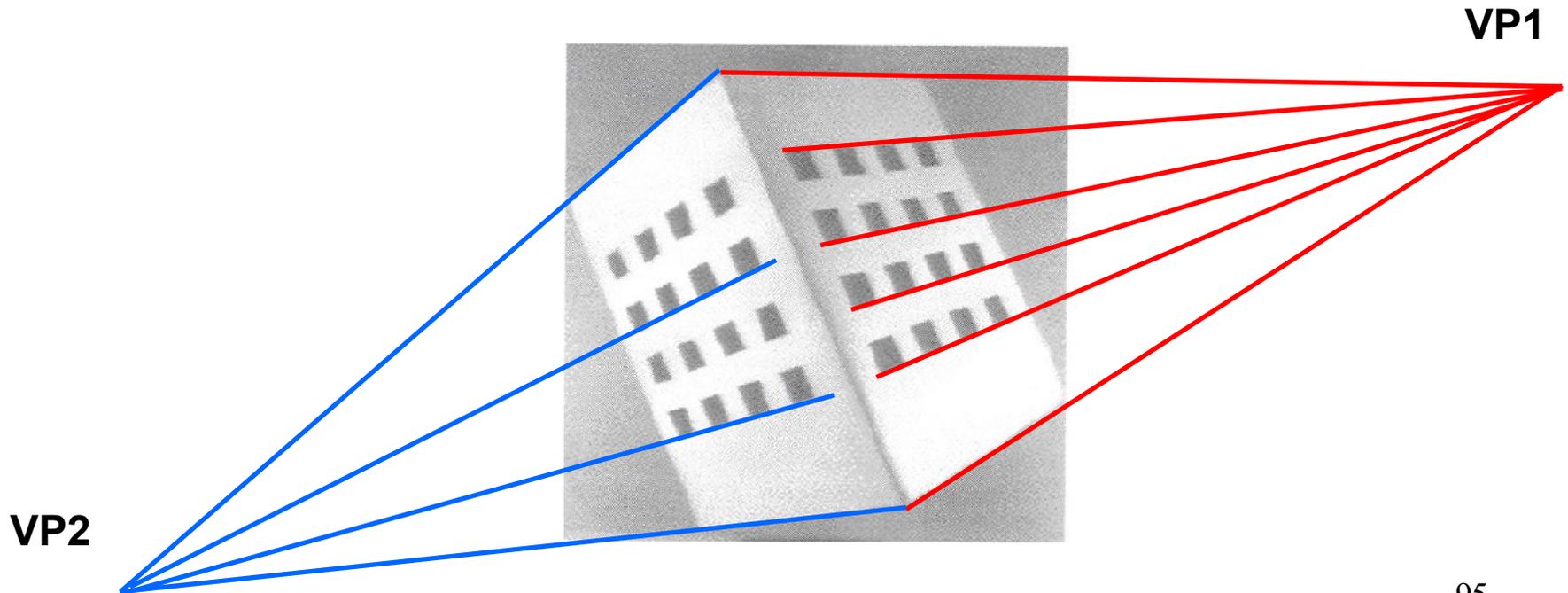
Définition : Point de fuite

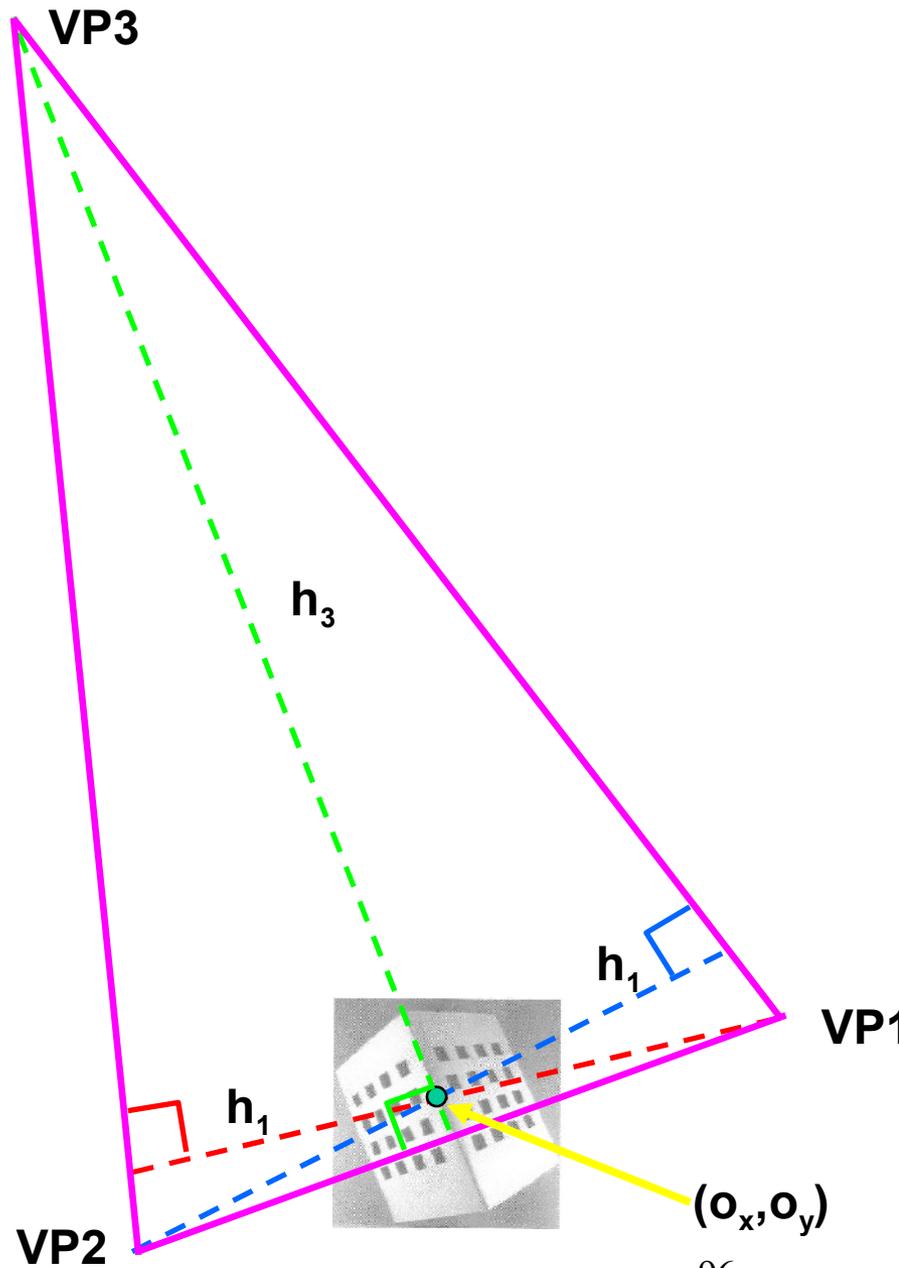
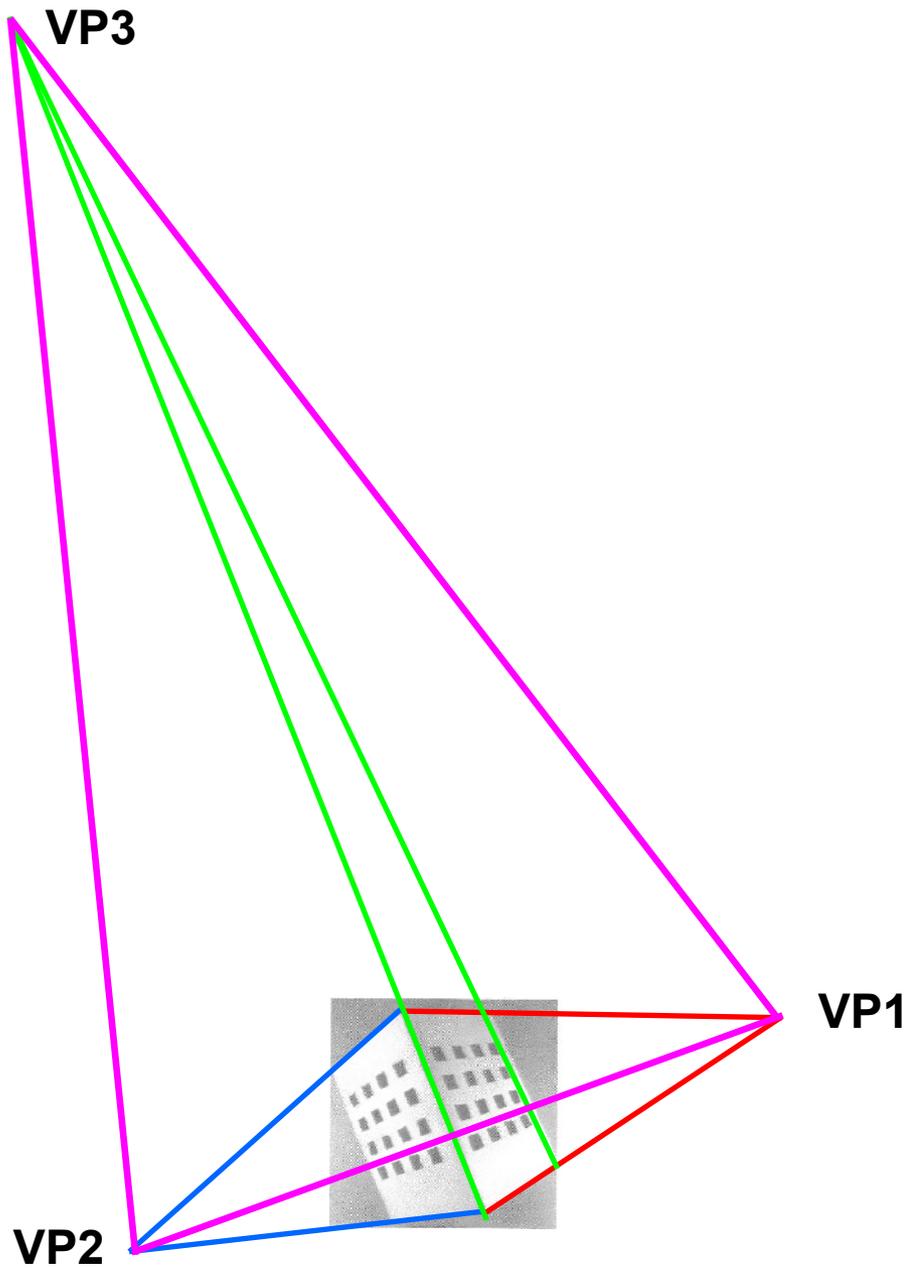
Soient L_i , $i=1, \dots, N$ droites parallèles dans l'espace 3D, et l_i les droites images correspondantes. En raison de la projection perspective, les droites l_i apparaissent comme se rencontrant en un point p , appelé point de fuite, défini comme l'intersection commune de toutes les droites images l_i .



Théorème : Point de l'orthocentre Centre de l'image à partir des points de fuite

Soit T le triangle du plan image défini par les trois points de fuite de 3 ensembles mutuellement orthogonaux de droites parallèles de l'espace. Le centre de l'image est l'orthocentre de T .



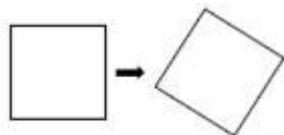


Piste pour la preuve du théorème :

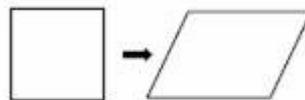
Soit h la hauteur reliant le sommet (et point de fuite) v au côté s , et O le centre de projection. Comme s est à la fois orthogonal à h et vO , le plan passant par h et vO est orthogonal à s et donc au plan image...

De la géométrie euclidienne à la géométrie projective

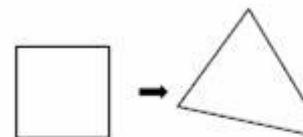
	Euclidean	similarity	affine	projective
Transformations				
rotation	X	X	X	X
translation	X	X	X	X
uniform scaling		X	X	X
nonuniform scaling			X	X
shear			X	X
perspective projection				X
composition of projections				X
Invariants				
length	X			
angle	X	X		
ratio of lengths	X	X		
parallelism	X	X	X	
incidence	X	X	X	X
cross ratio	X	X	X	X



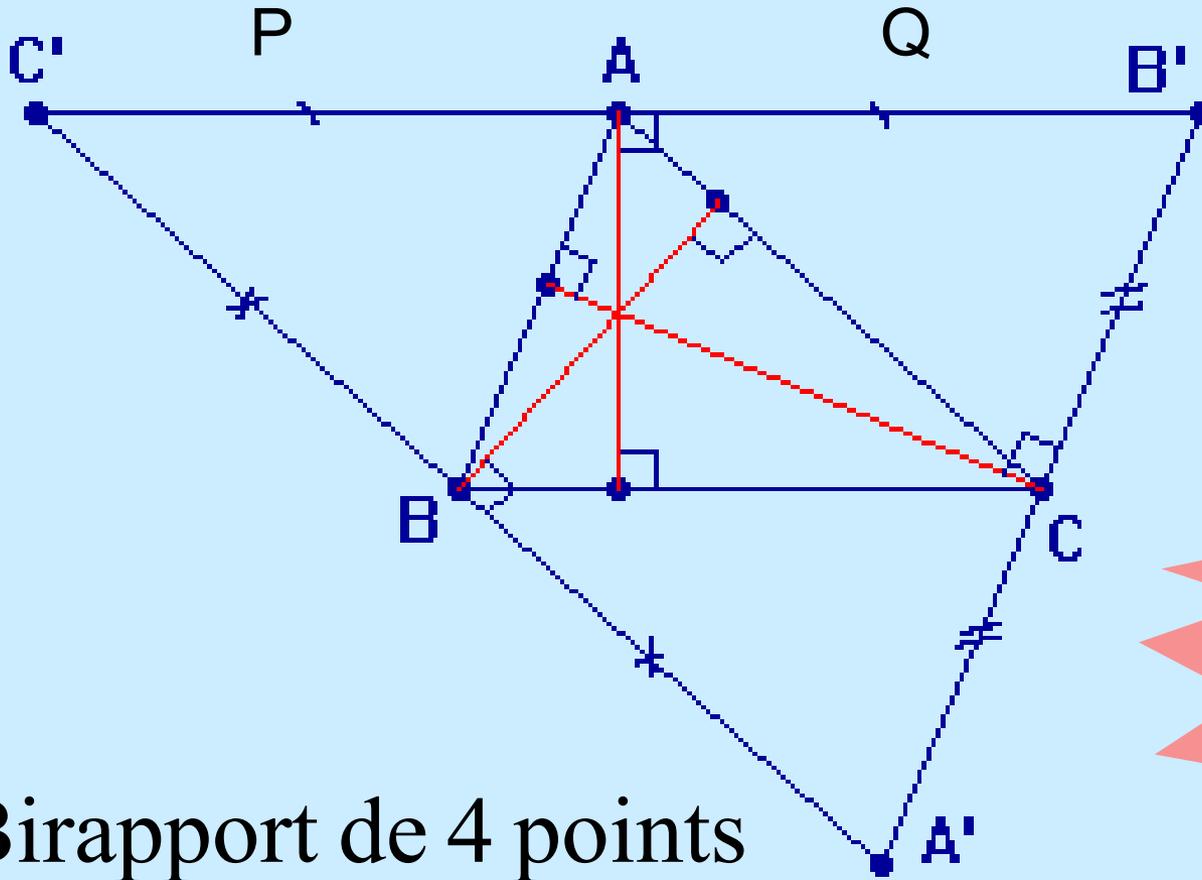
Rigid



Affine



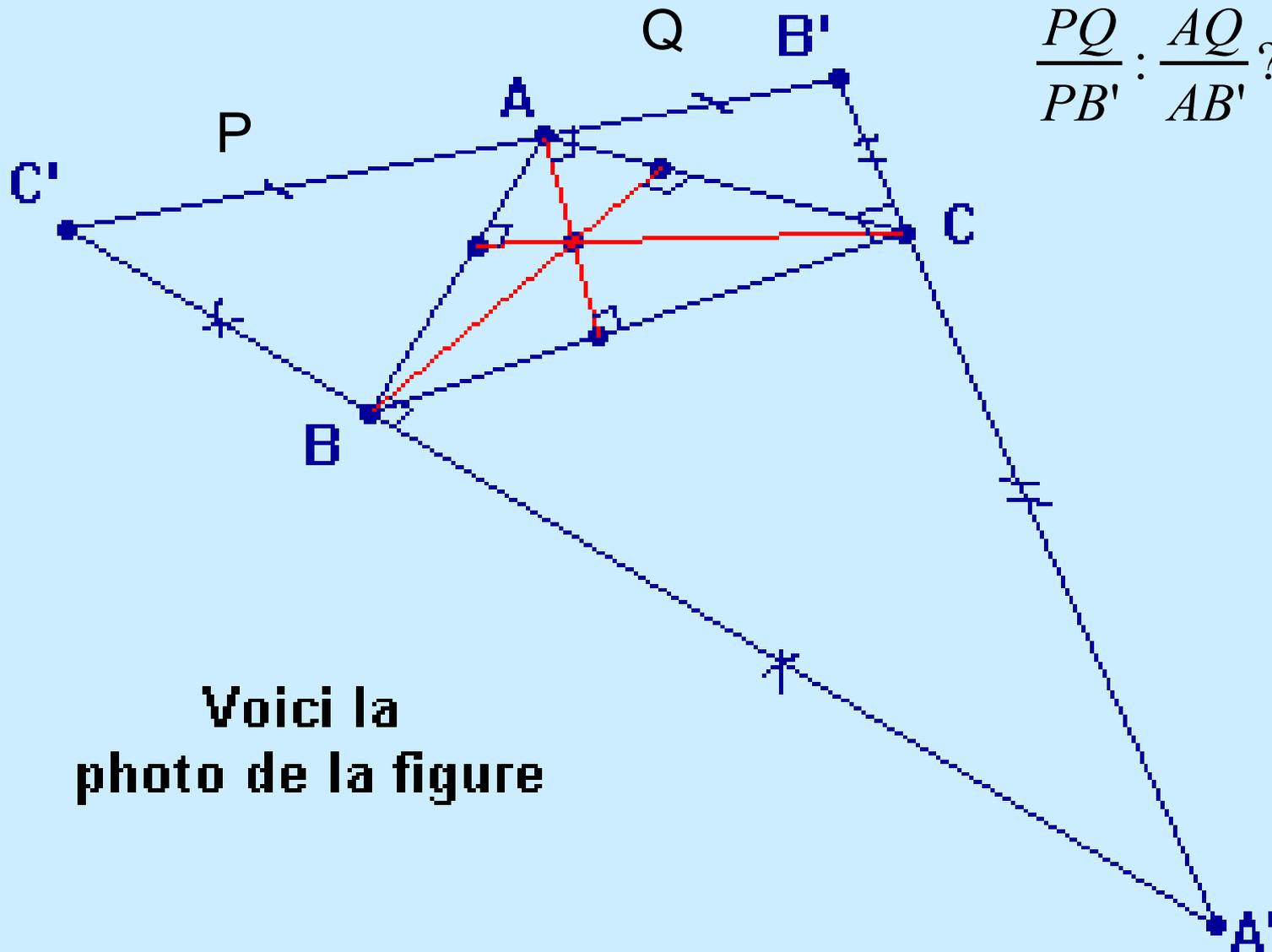
Projective



Birapport de 4 points

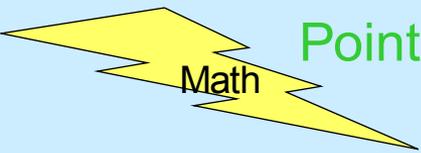
$$(P, A; Q, B') = \frac{PQ}{PB'} \cdot \frac{AQ}{AB'}$$



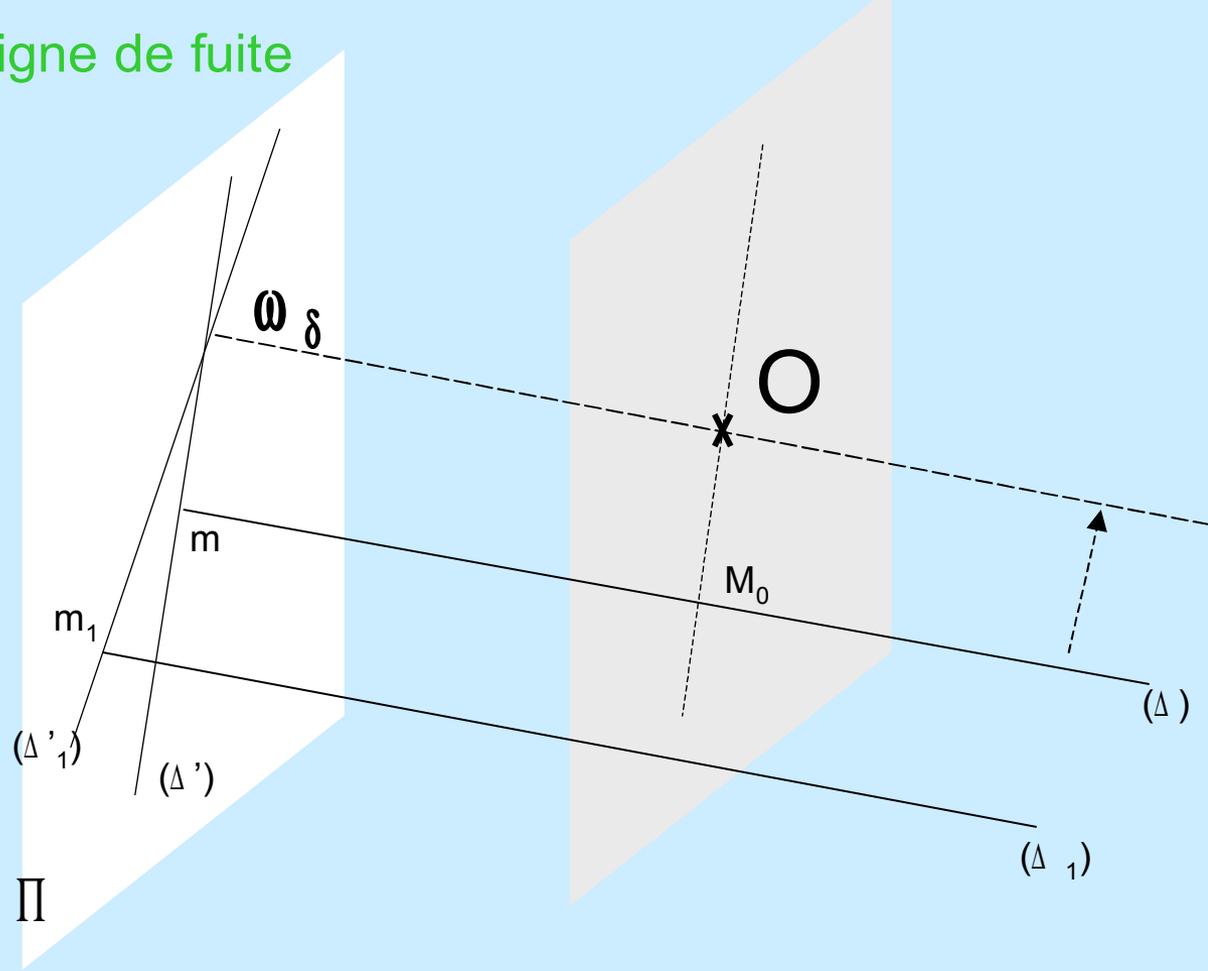


$$\frac{PQ}{PB'} : \frac{AQ}{AB'} ?$$

**Voici la
photo de la figure**



Point de fuite – Ligne de fuite

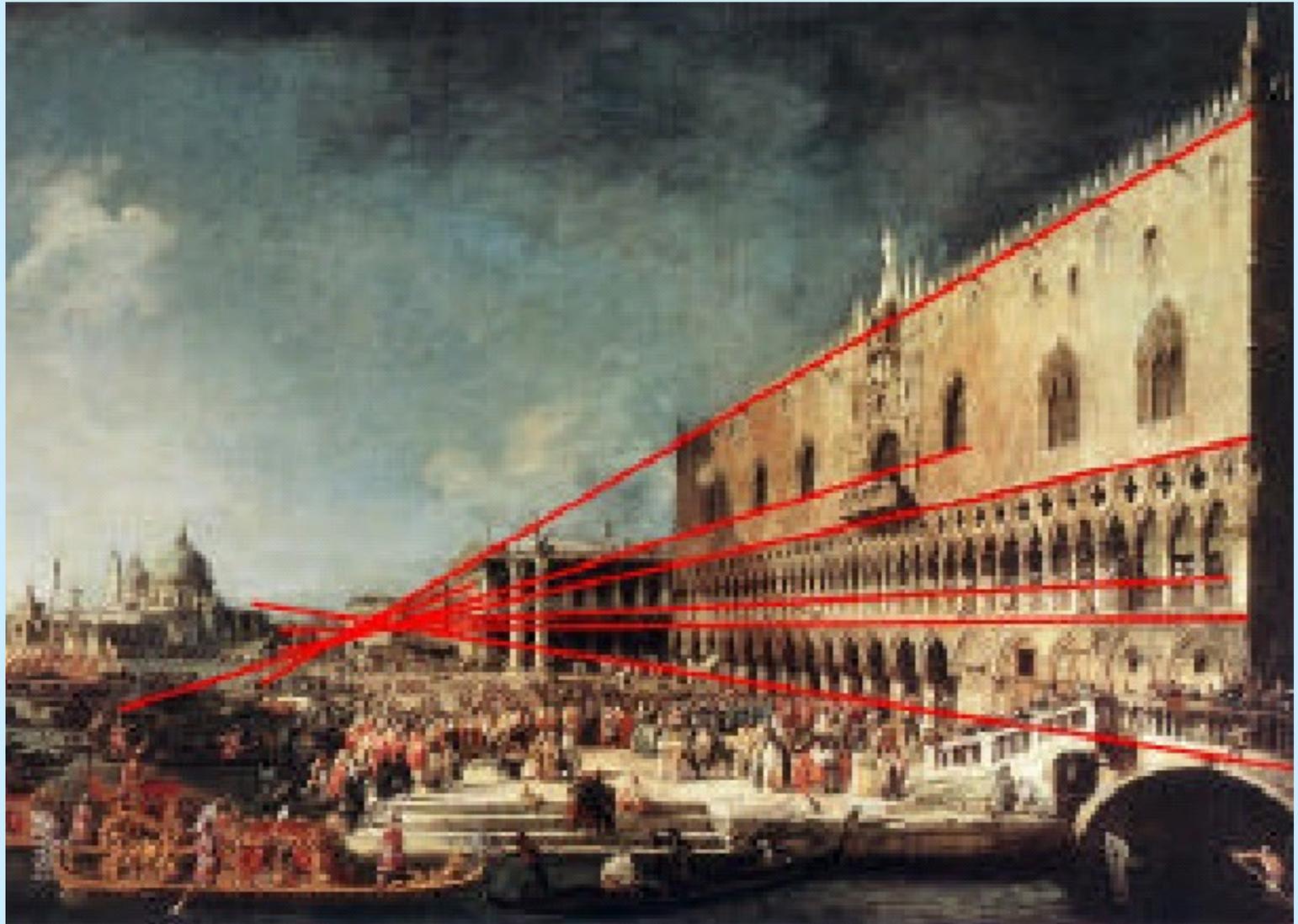


Théorème 1

La projection conique de centre O sur le plan Π est une bijection de $\Delta \setminus \{M_0\}$ sur $\Delta' \setminus \{\omega_\delta\}$, si Δ n'est pas parallèle à Π .

Théorème 2

Une famille de droite toutes parallèles à une même direction δ a pour image une famille de droites dans Π qui concourent toutes en le même point ω_δ lorsque δ n'est pas une direction de Π .



Adjonction des points à l'infini

Deux droites parallèles sont des droites qui se coupent à l'infini, en ajoutant un point à toute droite Δ de E qu'on appellera son point à l'infini, noté ∞_{Δ} .

On définit ce point à l'infini comme la direction δ de Δ .

$\hat{\Delta} = \Delta \sqcup \{\infty_{\Delta}\}$ complété projectif de Δ , droite projective de E

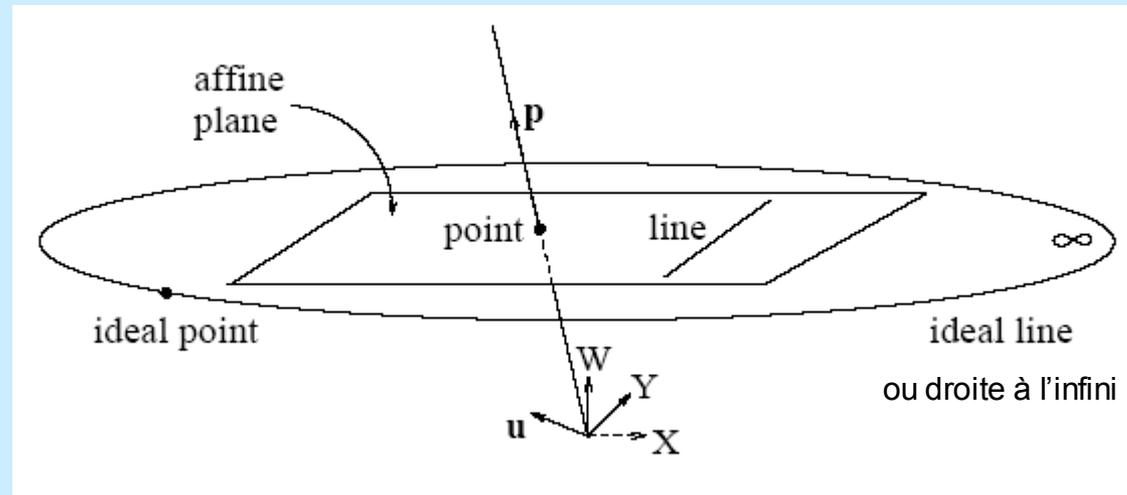
De même, $\hat{P} = P \sqcup \{\infty_P\}$ où ∞_P est l'ensemble des points à l'infini des droites de P et ∞_P s'appelle la droite à l'infini de P ,

et \hat{P} un plan projectif de E .

De même, $\hat{E} = E \sqcup \{\infty_E\}$, appelé espace projectif.

$\{\infty_P\} \sqcup \{\hat{\Delta}\} =$ ensemble des droites projectives de E

$\{\infty_E\} \sqcup \{\hat{P}\} =$ ensemble des plans projectifs de E



Plan affine augmenté

Avec ces définitions, deux droites projectives coplanaires distinctes de l'espace projectif se coupent toujours en un point de cet espace et un seul.

Remarque :

- 2 droites D et D' parallèles ont le même point à l'infini,

$$\infty_D = \infty_{D'}$$

$$(A, B; C, \infty_{\Delta}) = \frac{AC}{BC}$$

- de plus,

Prolongement des projections

La projection conique $p: E \setminus \Pi_O \rightarrow \Pi$ de point de vue O se prolonge naturellement en une application :

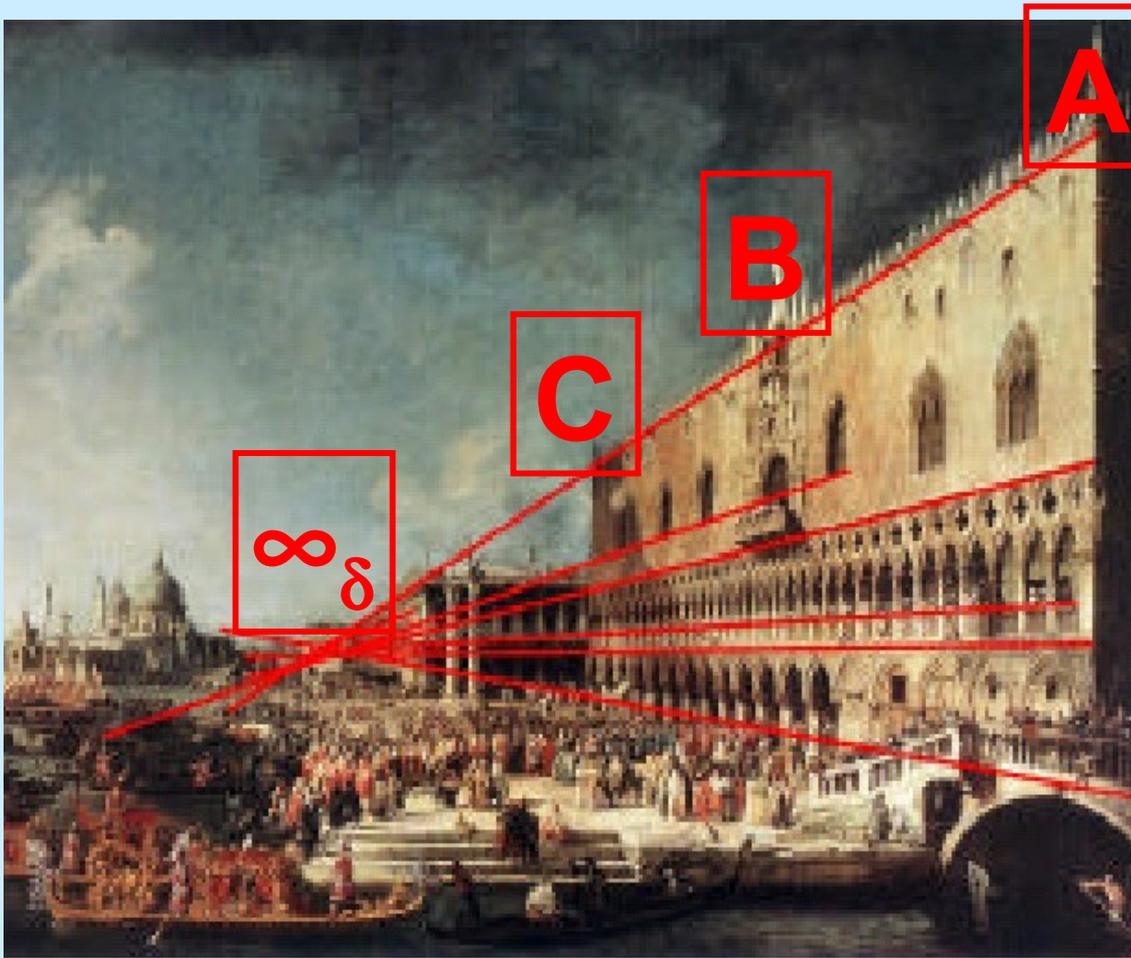
$$p: \hat{E} \setminus \{O\} \rightarrow \hat{\Pi}$$

En particulier, elle établit une bijection entre $\hat{\Delta}$ et $\hat{\Delta}'$

Ainsi, $p(M_O) = \infty_{\Delta'}$ et $p(\infty_{\Delta}) = \infty_{\delta}$

et la projection conique préserve le birapport de 4 points distincts quelconques de $\hat{E} \setminus \{O\}$, pourvus qu'ils soient alignés sur une droite de \hat{E} ne passant par O . Cas intéressant d'un milieu de segment :

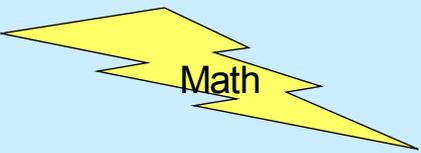
$$(A, B; C, \infty_{\Delta'}) = \frac{AC}{BC} = (p(A), p(B), p(C), \omega_{\delta})$$



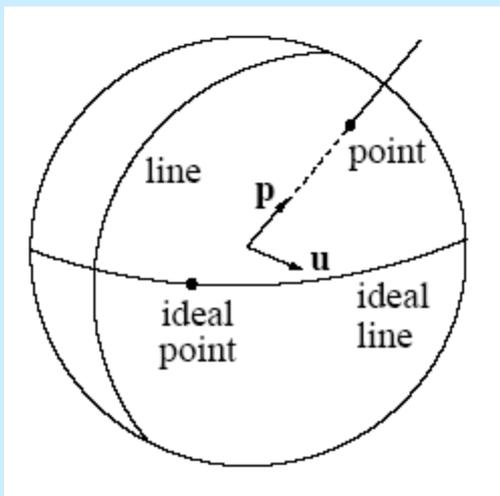
$$(A, B; C, \infty_{\delta}) = \frac{AC}{A\infty_{\delta}} : \frac{BC}{B\infty_{\delta}} = 2$$

La valeur $\frac{AC}{BC}$ une fois projetée en 2D est égale

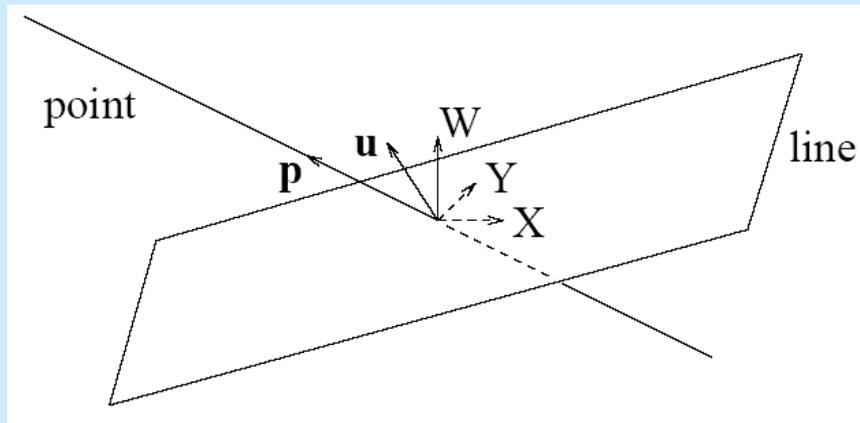
à $\frac{A\infty_{\delta}}{B\infty_{\delta}}$ la valeur du même rapport en 3D



Visualiser le plan projectif : ses points et ses droites



Sphère unité



Ray Space

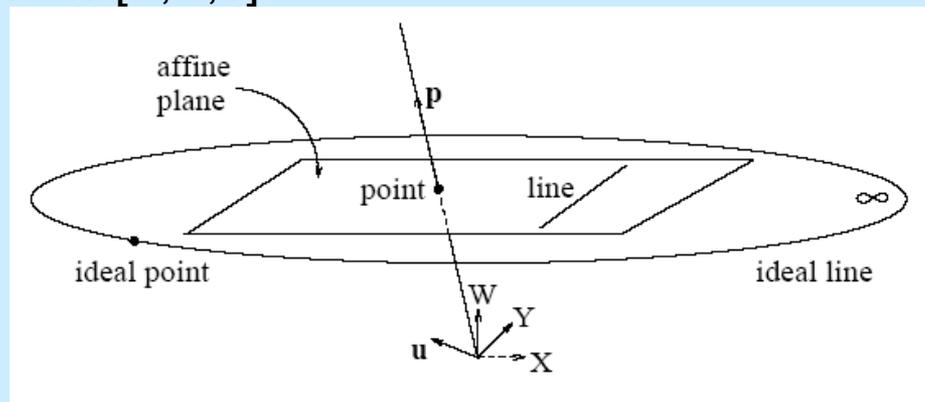
La géométrie projective et le plan projectif

- Une image plane est modélisée par le plan projectif \mathbb{P}^2
- Transformation d'un point du plan projectif $[X, Y, W]^T$ vers le plan euclidien $[x, y]^T$

$$x = \frac{X}{W} \quad \text{et} \quad y = \frac{Y}{W}$$

- Nécessité de points à l'infini ou points idéaux (cas où $W=0$) dans chaque direction : ensemble des points $[X, Y, 0]^T$

- Droite idéale représentée par $(0, 0, 1)$



Coordonnées Homogènes

La projection p est une bijection entre l'ensemble de toutes les droites de E passant par O , $P(E)$, et $\hat{\Pi}$.

En effet, si Δ_0 est une droite de Π_0 passant par O et Δ est une droite affine de Π parallèle à Δ_0 , on pose $p(\Delta_0) = \infty_{\Delta} = \infty_{\Delta_0}$.

Coordonnées Homogènes

La projection conique p associée au point M de $E \setminus \Pi_O$ de coordonnées (X, Y, Z) ($Z \neq 0$) le point m de $\hat{\Pi}$ de coordonnées $(x = X/Z, y = Y/Z)$.

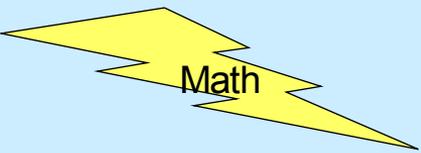
On dit que (X, Y, Z) est un système de coordonnées homogènes du point m et l'on note $m = [X, Y, Z]$.

Dans le plan affine, le repère affine (Ω, i, j) donne :

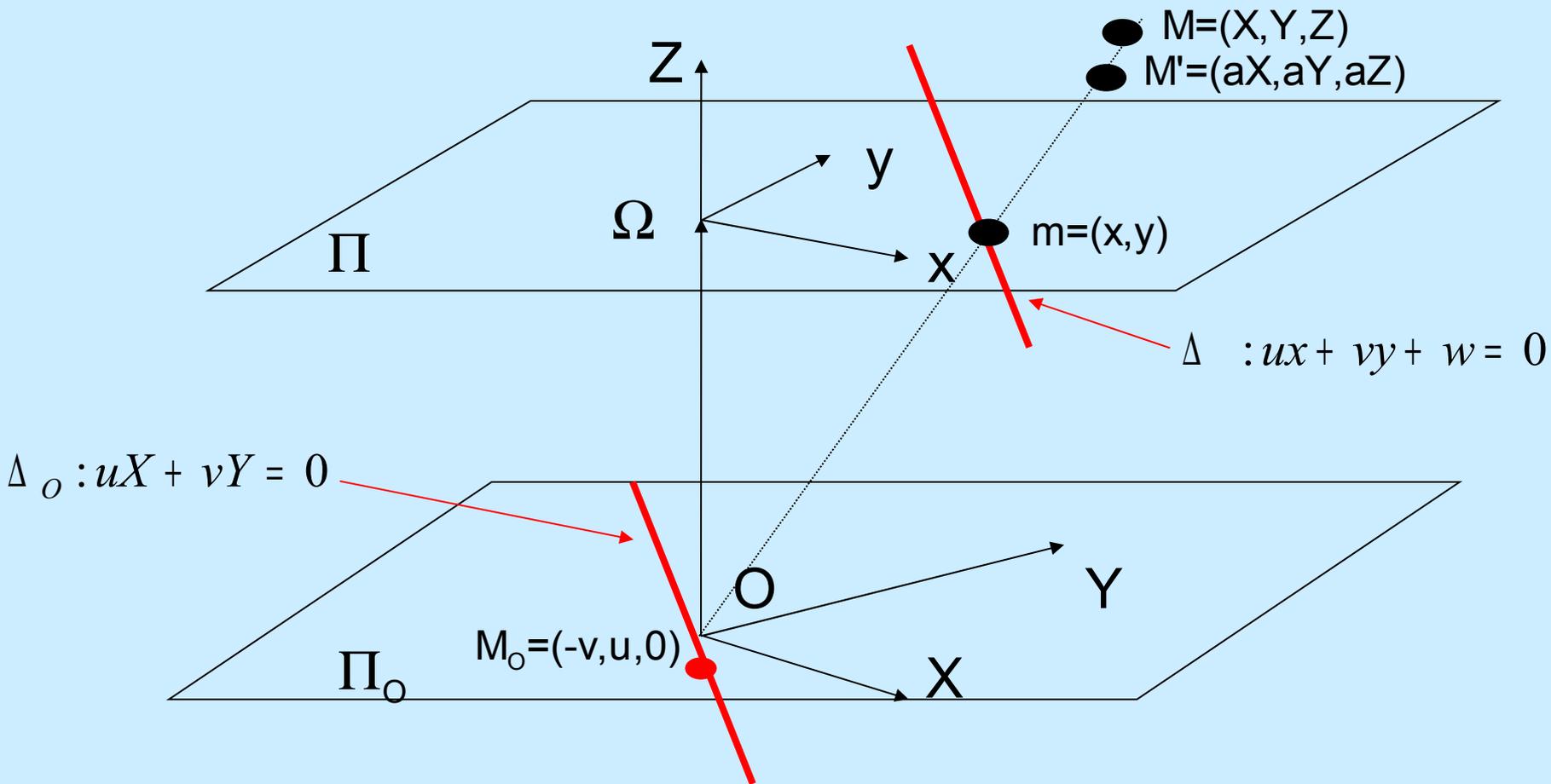
$$\overrightarrow{\Omega m} = xi + yj \quad \rightarrow \rightarrow \rightarrow$$

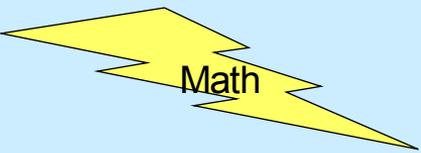
Si $k = \overrightarrow{O\Omega}$, dans le repère affine (O, i, j, k) de E ,

$$\overrightarrow{OM} = Xi + Yj + Zk$$



Coordonnées Homogènes





Coordonnées Homogènes

$p((OM))=m$ et tout point de (OM) définit un système de coordonnées homogènes de m $(\lambda X, \lambda Y, \lambda Z)$ ($Z \neq 0$)

Or $p(\Delta_O) = \infty_{\Delta}$ et si $ux+vy+w=0$ est l'équation affine dans le repère (Ω, i, j) de (Δ) , l'équation affine de Δ_O dans (O, i, j, k) est $uX+vY=0$, or tout point de Δ_O définit un système de coordonnées homogènes de ∞_{Δ} , donc $\infty_{\Delta} = [-v, u, 0]$ est un système de coordonnées homogènes de ∞_{Δ} .

$\hat{\Pi}$

Désormais tout point $[X, Y, Z]$ définit un élément de $\hat{\Pi}$ pourvu que X, Y, Z ne soient pas tous nuls simultanément.

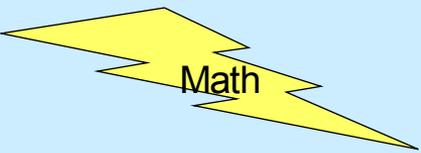
$\hat{\Delta}$

Coordonnées Homogènes

Avec ces conventions, le point $[X, Y, Z]$ appartient à $\hat{\Delta}$ ssi $uX+vY+wZ=0$.

C'est l'équation de $\hat{\Delta}$ en coordonnées homogènes.

De même, la droite à l'infini du plan ∞_{Π} a pour équation $Z=0$.
Ainsi toutes les droites projectives de $\hat{\Pi}$ sont définies en coordonnées homogènes par une équation de la forme $uX+vY+wZ=0$ avec u, v, w non tous nuls.



Principe de dualité ou d'équivalence des points et des droites

• Equation d'une droite :

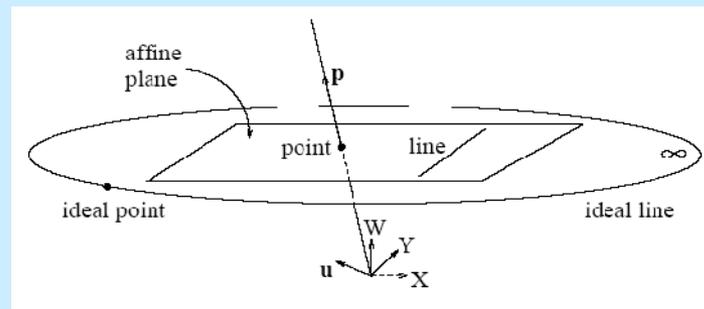
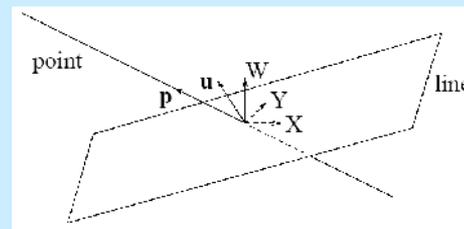
• Euclidienne (a,b,c):

$$ax+by+c=0$$

$$aX+bY+cW=0 \text{ avec } x=X/W \text{ et } y=Y/W$$

• Projective u:

$$u^T p = p^T u = 0$$



point	$\mathbf{p} = (X, Y, W)$
incidence	$\mathbf{p}^T \mathbf{u} = 0$
collinearity	$ \mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3 = 0$
join of 2 points	$\mathbf{u} = \mathbf{p}_1 \times \mathbf{p}_2$
ideal points	$(X, Y, 0)$

(a)

line	$\mathbf{u} = (a, b, c)$
incidence	$\mathbf{p}^T \mathbf{u} = 0$
concurrency	$ \mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3 = 0$
intersection of 2 lines	$\mathbf{p} = \mathbf{u}_1 \times \mathbf{u}_2$
ideal line	$(0, 0, c)$

(b)

Coordonnées Homogènes pour des (a) points (b) droites

S'en convaincre en visualisant les plans 3D sous-jacents aux droites ou aux points¹¹⁵

Principe de dualité ou d'équivalence des points et des droites

Nous dirons qu'un résultat, un concept, ou un problème de géométrie plane relève de la géométrie projective s'il peut s'exprimer en termes d'alignements de points, de concours de droites ou de plans et de birapports sans que l'on ait à distinguer les points à l'infini des autres.

Les concepts de conique est un concept de la géométrie projective. Le parallélisme appartient à la géométrie affine.

Principe d'équivalence des points et des droites

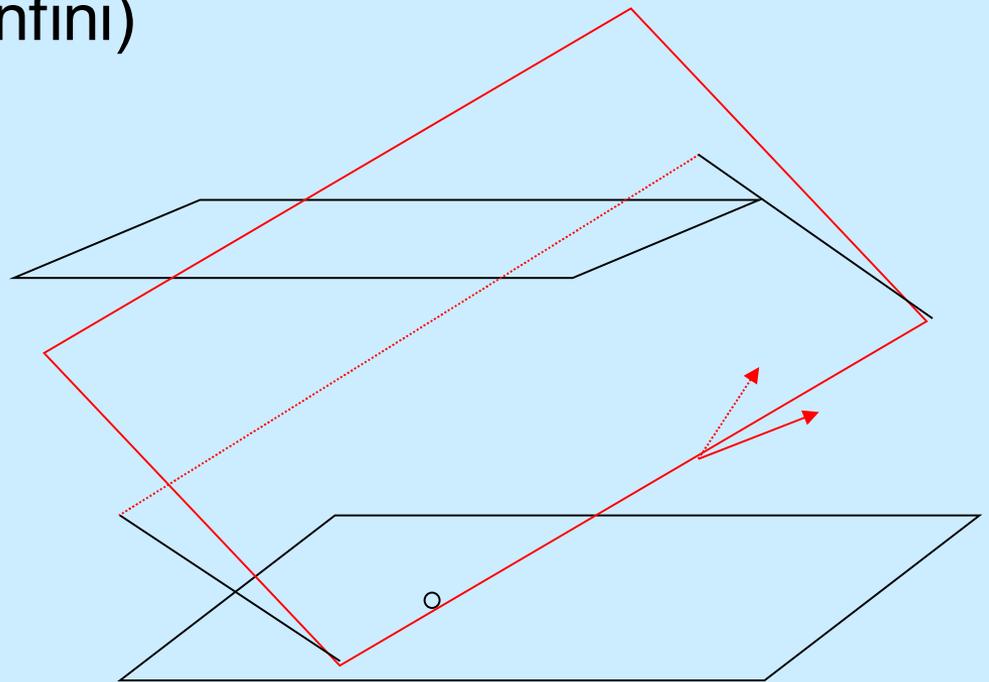
Tout théorème du plan projectif peut être reformulé en substituant les points aux droites et inversement

2 points du plan projectif sont alignés par définition.

Donc,

2 droites du plan projectif s'intersectent toujours (en un point qui peut être à l'infini)

$$\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \times \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix}$$

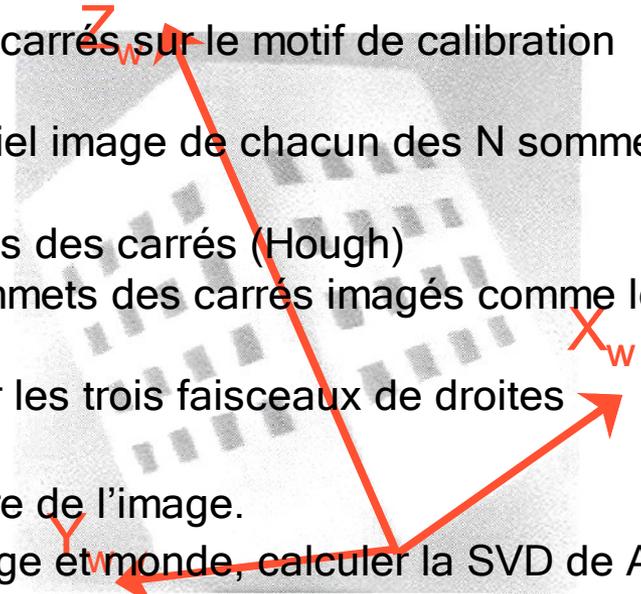


La géométrie descriptive (projective) est le domaine des mathématiques qui modélise les notions intuitives de *perspective* et d'*horizon*. Elle a d'abord été imaginée au XVII^e siècle par des mathématiciens comme Pascal ou Desargues, avant de tomber dans l'oubli jusqu'à sa redécouverte par Poncelet, et sa remise au goût du jour par Félix Klein à la fin du XIX^e siècle. Elle est aujourd'hui largement utilisée par les systèmes de visualisation sur ordinateur GL et OpenGL grâce notamment à O. Faugeras (INRIA).

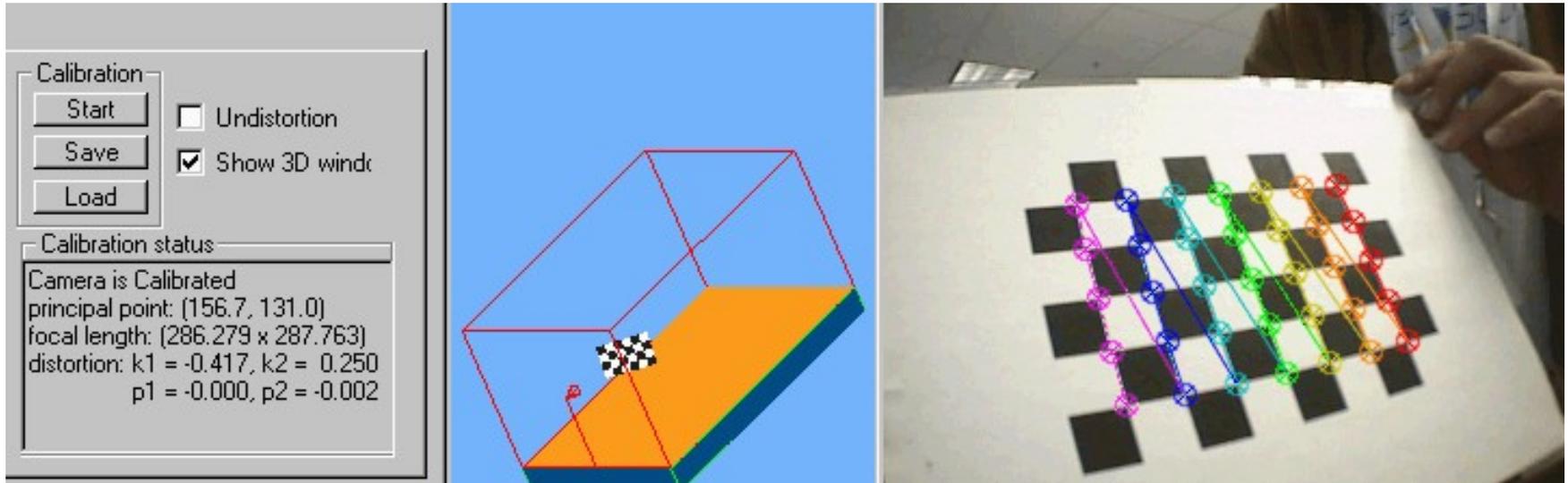
Procédure PARAM_CAL

INPUT : Une image du motif de calibration.

- Mesurer les coordonnées 3D de chaque sommet des n carrés sur le motif de calibration dans le référentiel Monde. Soit $N = 4n$
- Dans le but de trouver les coordonnées dans le référentiel image de chacun des N sommets :
 - Localiser les droites images définies par les côtés des carrés (Hough)
 - Estimer les coordonnées images de tous les sommets des carrés imagés comme les intersections des droites extraites
- Calculer les trois points de fuite p_1, p_2, p_3 déterminés par les trois faisceaux de droites obtenus.
- Calculer l'orthocentre O du triangle $p_1p_2p_3$. O est le centre de l'image.
- Ayant établi les N correspondances entre les points image et monde, calculer la SVD de A . La solution est la colonne de V correspondant à la plus petite valeur singulière de A .
- Déterminer α and $|\gamma|$
- Déterminer les deux premières lignes de R et deux premières composantes de T .
- Calculer la troisième ligne de R comme le produit vectoriel des deux premières lignes estimées en 7, et forcer la contrainte d'orthogonalité en estimant R par sa décomposition SVD.
- Choisir un point pour lequel $(x-o_x)$ est notablement différent de 0. Si l'inégalité $(x-o_x)(r_{11}X^w+r_{12}Y^w+r_{13}Z^w+T_x) > 0$ est vérifiée, inverser les signes des deux première lignes de R et des deux premières composantes de T .
- Réinitialise A et b pour estimer T_z et f_x .



OUTPUT : o_x, o_y, α, f_x et les paramètres extrinsèques de la caméra.



http://w3.impa.br/~pcezar/3dp/original/CVL_html/appPage/calib_doc/index.html

C'est un problème inverse intéressant :

étant donné une photographie prise par une caméra inconnu à partir d'une position inconnue et qui a été rétrécie ou élargie, comment déterminer la position et l'orientation de la caméra mais aussi jusqu'à quel point l'image a été rétrécie ou élargie ?

Les applications sont nombreuses :

En navigation autonome, un missile de croisière peut obtenir la matrice de transformation caméra à partir d'un modèle de terrain embarqué et ainsi calculer les paramètres de la caméra qui définissent la localisation du véhicule.

Ou bien encore, une caméra stationnaire observant la zone de travail d'un bras robotisé peut déterminer la position et l'orientation du bras articulé par marquage de l'instrument de manipulation.

2 méthodes reconnues :

1. par estimation directe des paramètres

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -f_x & 0 & o_x & 0 \\ 0 & -f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X^{(W)} \\ Y^{(W)} \\ Z^{(W)} \\ 1 \end{bmatrix}$$

8. par estimation directe de la Matrice de Projection

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Rappels

$$\begin{pmatrix} x_{im} \\ y_{im} \end{pmatrix} = \begin{pmatrix} u/w \\ v/w \end{pmatrix} \quad \leftarrow \quad \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \mathbf{M}_{\text{int}} \mathbf{M}_{\text{ext}} \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix}$$

$$\mathbf{M}_{\text{ext}} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & T_x \\ r_{21} & r_{22} & r_{23} & T_y \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1^T & T_x \\ \mathbf{R}_2^T & T_y \\ \mathbf{R}_3^T & T_z \end{bmatrix}$$

$$\mathbf{M}_{\text{int}} = \begin{bmatrix} -f_x & 0 & o_x \\ 0 & -f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

Soit, en laissant tomber les indices $_{im}$ et $_w$, pour un point P_i et son image p_i dans le référentiel fenêtre (qu'on appellera image),

$$\begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix} = \mathbf{M} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}$$

Avec, \mathbf{M} à 10 paramètres (4 intrinsèques + 6 extrinsèques)

$$\mathbf{M} = \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

D'où pour N paires $\{(x_i, y_i), (X_i, Y_i, Z_i)\}$ liant les référentiels image et monde,

$$\begin{pmatrix} u_i \\ v_i \\ w_i \end{pmatrix} = \mathbf{M} \begin{pmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{pmatrix}$$



$$\mathbf{A}\mathbf{m} = \mathbf{0}$$



$$x_i = \frac{u_i}{w_i} = \frac{m_{11}X_i + m_{12}Y_i + m_{13}Z_i + m_{14}}{m_{31}X_i + m_{32}Y_i + m_{33}Z_i + m_{34}}$$

$$y_i = \frac{v_i}{w_i} = \frac{m_{21}X_i + m_{22}Y_i + m_{23}Z_i + m_{24}}{m_{31}X_i + m_{32}Y_i + m_{33}Z_i + m_{34}}$$

Avec

$$\mathbf{A} = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1X_1 & -x_1Y_1 & -x_1Z_1 & -x_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1X_1 & -y_1Y_1 & -y_1Z_1 & -y_1 \\ \cdot & \cdot \end{bmatrix}$$

$$\mathbf{m} = [m_{11} \quad m_{12} \quad m_{13} \quad m_{14} \quad m_{21} \quad m_{22} \quad m_{23} \quad m_{24} \quad m_{31} \quad m_{32} \quad m_{33} \quad m_{34}]^T$$

Etape 1 : Estimation de la matrice de projection M

On a affaire à un système d'équations linéaires en m à $2N$ équations

et 11 variables indépendantes :

-> Si $N \geq 6$, la SVD indique que m sera connu à un facteur d'échelle près.

$$\hat{\mathbf{m}} = \alpha \left[m_{11} \quad m_{12} \quad m_{13} \quad m_{14} \quad m_{21} \quad m_{22} \quad m_{23} \quad m_{24} \quad m_{31} \quad m_{32} \quad m_{33} \quad m_{34} \right]^T$$

Etape 2 : Calculer les paramètres de caméra

M est une matrice de projection 3x4

- Contient à la fois les informations intrinsèques et extrinsèques

$$\hat{\mathbf{M}} = \begin{bmatrix} \mathbf{q}_1 & q_{41} \\ \mathbf{q}_2 & q_{42} \\ \mathbf{q}_3 & q_{43} \end{bmatrix}$$

$$\mathbf{M} = \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

$$\hat{\mathbf{M}} = \gamma \mathbf{M}$$

Procédure PROJ_MAT_CAL

INPUT : Une image du motif de calibration.

3. Mesurer les coordonnées 3D de chaque sommet des n carrés sur le motif de calibration dans le référentiel Monde. Soit $N = 4n$

4. Dans le but de trouver les coordonnées dans le référentiel image de chacun des N sommets :

- Localiser les droites images définies par les côtés des carrés (Hough)
- Estimer les coordonnées images de tous les sommets des carrés imagés comme les intersections des droites extraites

5. Etant donné N correspondances monde-image, calculer la SVD de $A = UDV^T$. La solution est la colonne de V correspondant à la plus petite valeur singulière de A .

6. Trouver $|\gamma|$ en utilisant $\sqrt{\hat{m}_{31}^2 + \hat{m}_{32}^2 + \hat{m}_{33}^2} = |\gamma| \sqrt{r_{31}^2 + r_{32}^2 + r_{33}^2} = |\gamma|$

7. $\hat{M} \leftarrow \frac{1}{|\gamma|} \hat{M}$ ainsi M est connu à un signe σ près

8. Calculer T_z par $T_z = \sigma \hat{m}_{34}$

9. Calculer R_3^T par $r_{3i} = \sigma \hat{m}_{3i}$

10. $o_x = q_1^T q_3$ et $o_y = q_2^T q_3$

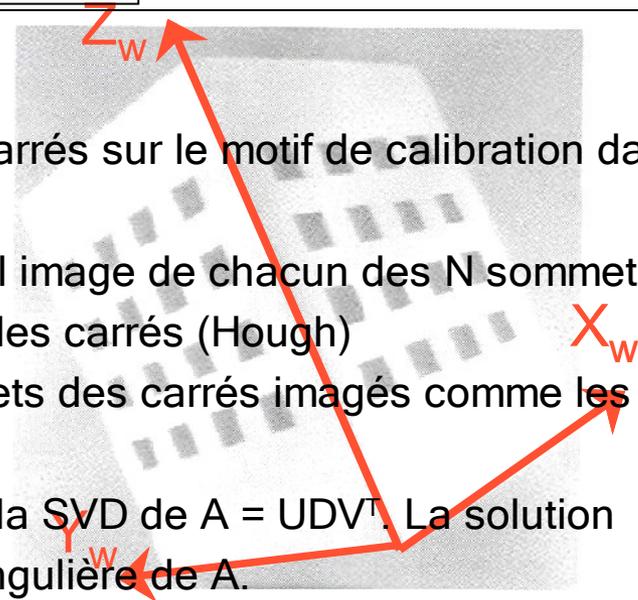
11. On peut retrouver f_x et f_y par $f_x = \sqrt{q_1^T q_1 - o_x^2}$ et $f_y = \sqrt{q_2^T q_2 - o_y^2}$

$r_{1i} = \sigma (o_x \hat{m}_{3i} - \hat{m}_{1i}) / f_x$, $r_{2i} = \sigma (o_y \hat{m}_{3i} - \hat{m}_{2i}) / f_y$

12. $T_x = \sigma (o_x T_z - \hat{m}_{14}) / f_x$, $T_y = \sigma (o_y T_z - \hat{m}_{24}) / f_y$

13. Trouver le signe σ avec $T_z > 0$ si O^w est devant O^c et orthogonaliser l'estimée de R trouvée par la technique SVD

OUTPUT : o_x, o_y, α, f_x et les paramètres extrinsèques de la caméra.



■ Propriétés en commun :

- On résoud D'abord un système linéaire puis on procède à une décomposition en paramètres physiques;
- Les résultats devraient être identiques.

■ Différences :

- Le nombre de variables dans les systèmes homogènes considéré :
 - Méthode matricielle directe : tous les paramètres d'un coup, puis 2N Equations de 12 variables
 - Méthodes paramétriques directe : N Equations de 8 variables, puis N équations de 2 Variables, puis trouver le Centre de l'Image – peut-être plus stable
- Les hypothèses :
 - Méthode matricielle directe : plus simple et plus général; parfois la matrice de projection est suffisante et donc on n'a pas besoin de la décomposition en paramètres physiques réels.
 - Méthodes paramétriques directe : elle suppose que le centre de l'image est connue au départ puis que le rapport d'aspect est connu pour l'estimation du centre de l'image

Bibliographie

- “Three-dimensional computer vision : A geometric Viewpoint”, O.D. Faugeras, MIT Press, Cambridge (MA), 1993
- “Introductory techniques for 3D computer vision”, E. Trucco et Alessandro Verri, Prentice Hall, 1998