

MASTER  
Vision et Machines Intelligentes



MATHÉMATIQUES ET INFORMATIQUE

**Sciences**

**Université Paris Cité**

Nicolas Loménie, Ass. Prof., HDR, Ing.

# Analyse de Séquences Vidéo

<https://perceiving-systems.blog/en/post/my-first-iccv>

## 2 cadres d'étude :

**Hardware** : robotique, caméras, microscopie, bio-mimétique etc. (<https://www.prophesee.ai/>)



- **Software** : Modèles, Cadres algorithmiques  
-> **objet de ce cours pour l'essentiel**

Ce qui est **brevetable** : l'algorithme associé à un capteur (en général donc un mix<sup>2</sup> des deux)

## 3 points de vue / cadres applicatifs :

- **Vision par ordinateur / Jeux :**  
**aspect perception** de l'environnement réel et agent intelligent orienté 3D
- **Imagerie Médicale de type Screening :**  
**aspect tracking** rapide et massif orienté 2D
- **Télécommunications / Multimédia :**  
**aspect traitement du signal** (compression, représentation, indexation)

# Analyse de Séquences Vidéo

Nicolas Loménie

## INTRODUCTION A L'ASPECT VISION PAR ORDINATEUR

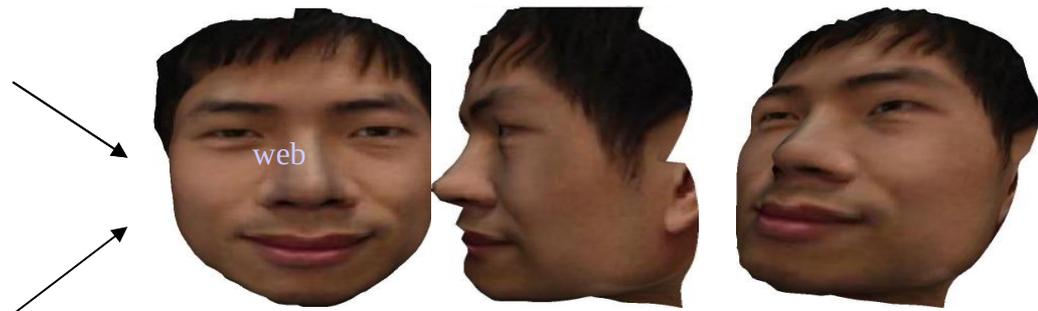
# Objectif de la VpO :

Calculer des propriétés du monde réel (**3D en mouvement**) à partir d'images numériques dans un but de **réalité mixte**.

Ces propriétés sont de type :

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

Dans cette optique, les séquences issues de vidéo sont une modalité d'acquisition particulière associée à des algorithmes adaptés.



Vues Multiples contre Vidéo

# Disciplines :

- Image Analysis
- Scene Analysis
- Image Understanding
  
- Image Processing
- Pattern Recognition
- Photogrammetry

# 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...

# Champs de Recherches :

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



# Ressources :

## Conférences : Journaux : Internet :

- ICCV
- CVPR
- ECCV
- ICPR ...

- IJCV
- PAMI
- CVIU
- MVA
- PR
- PRL

- Computer Vision Home Page :  
<http://www-2.cs.cmu.edu/~cil/vision.html>
- <http://peipa.essex.ac.uk>
- <http://iris.usc.edu/Vision-Notes/bibliography/contents.html>  
*Et le catalogue de la BU de Université de Paris*  
<https://u-paris.fr/bibliotheques/>

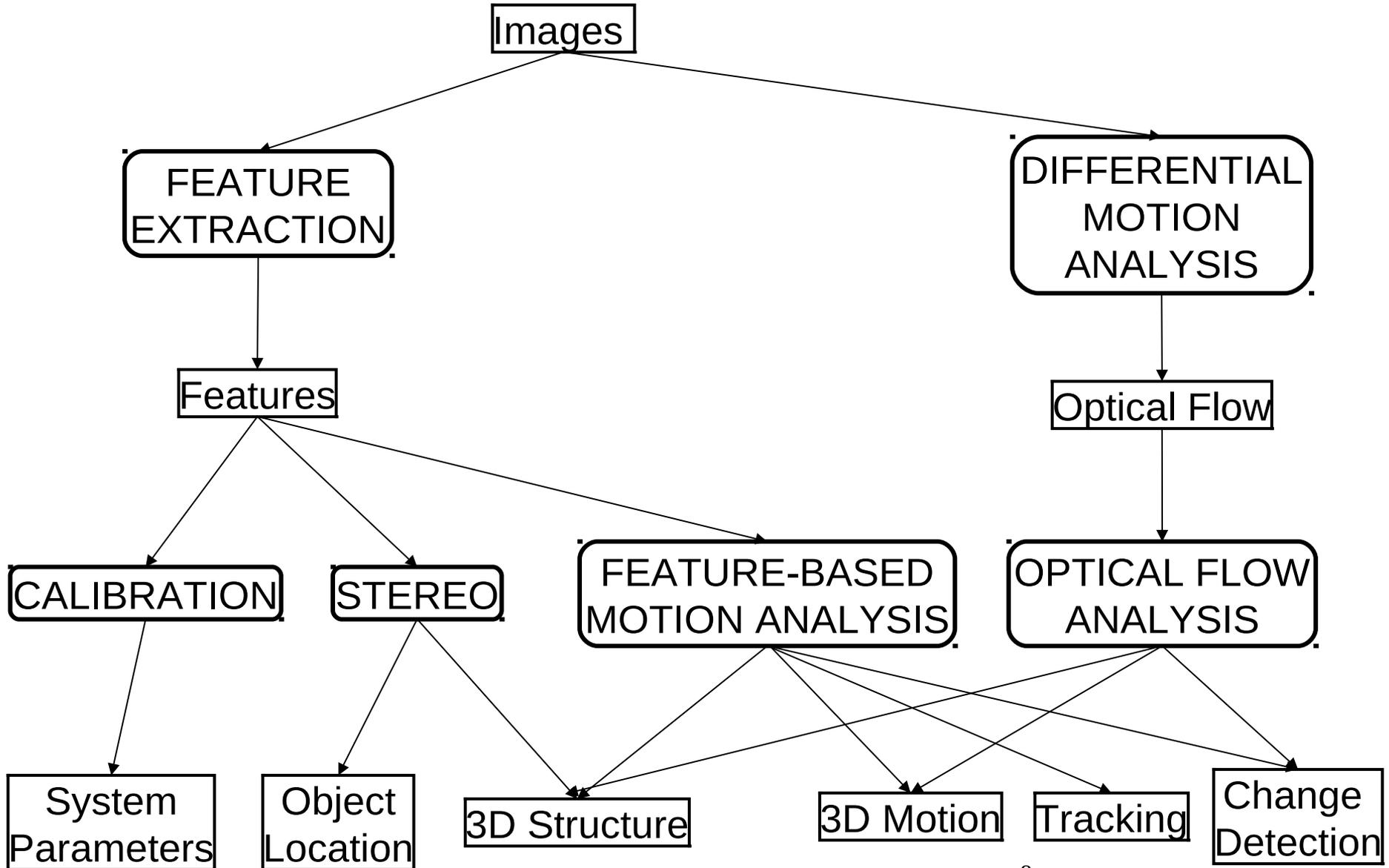
### Exemples de périodiques en ligne accessibles depuis Université de Paris (to be checked again) :

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>

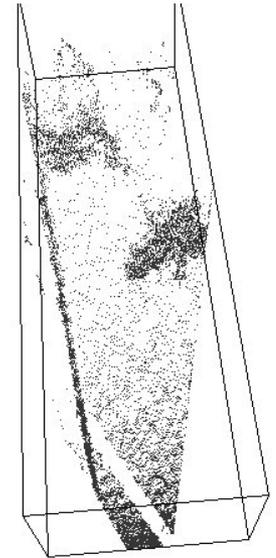
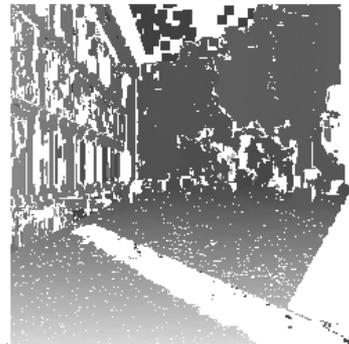
### Sont également accessibles de nombreux autres périodiques, publiés par :

Elsevier - Science Direct  
Kluwer Online  
Wiley Interscience  
Springer

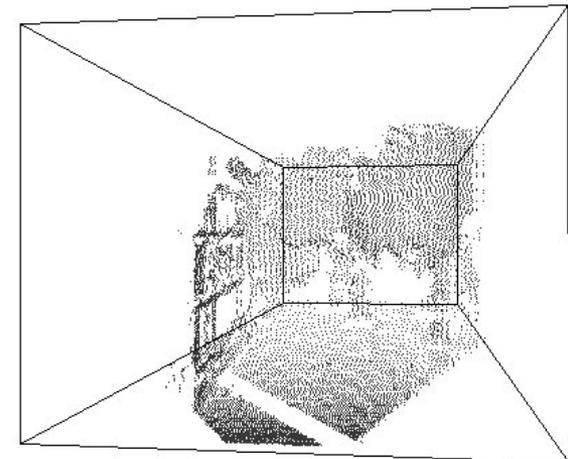
Tiré de "Introductory techniques for 3D computer vision", Emanuele Trucco et Alessandro Verri, Prentice Hall, 1998  
A vous en fin de formation de la modifier, compléter, réarranger et de le remplir des compétences acquises durant l'ensemble des UE de cette formation.



# Intensity Image vs. Range Image Passive Optical vs. Active Device



web



Dans ce cours, on ne considère que les données optiques issus de capteur optiques traditionnels, c-a-d grand public -> [large marché](#)

Exemple industriel de CyCab : [cycab\\_plaquette\\_inria.pdf](#) de Robosoft Company

# Analyse de Séquences Vidéo

Nicolas Loménie

## Capteurs et paramètres Le problème de la calibration

# Objectifs :

- Analyse du phénomène : Comprendre comment est imagé le monde 3D sur une rétine 2D ( capteur optique dit passif = oeil humain)
  - Géométriquement -> dans un but de modélisation du réel
  - Algébriquement -> dans un but algorithmique
- Apprendre la notion de modèle de caméra pinhole et de matrice de projection
- Définir les paramètres extrinsèques et intrinsèques d'un capteur passif
- Synthèse : Montrer une méthode de calibration de caméra classique (au passage, apprendre une technique de résolution d'équations par décomposition SVD) et comprendre que cela facilite la rétroprojection directe du monde 2D vers le monde physique 3D
- Première illustration de la méthodologie générale en VpO optique :  
**Analyse 3D vers 2D -> Modélisation -> Synthèse 2D vers 3D**
- Sans oublier qu'une multitude de capteurs nouveaux existent ou sont en cours de conception pour passer directement du monde physique 3D à une représentation et un **Display 3D**. En ce sens, ce chapitre est à la limite d'un cours sur l'Analyse de Séquence Vidéo et intéresse principalement le cours d'Imagerie 3D mais nous verrons que les notions de Flot Optique utilisées en Analyse du Mouvement utilisent cette modélisation du capteur et que la problématique de calibration est générale à la VpO notamment pour la réalité mixte ou **augmentée**.

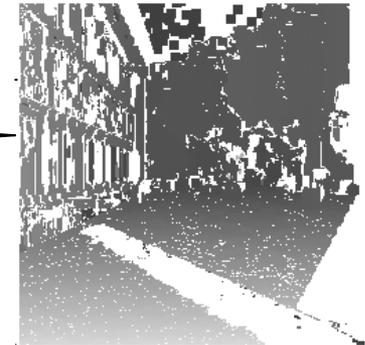
Intensity image :

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

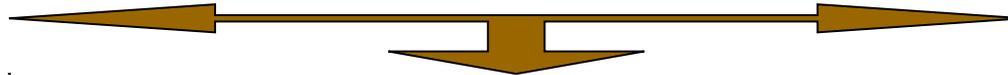


Range image :

images codant la forme et la distance, acquises par des capteurs spécifiques comme le scanner laser. On parle d'images de profondeur ou topographiques ((voir cours Imagerie 3D)



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	143	172	219	214	223	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	224	229	210
137	132	138	137	166	221	219	224	229	210
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	210
149	140	144	130	169	223	225	223	227	223
142	139	143	144	172	224	224	225	229	214
147	144	145	147	171	227	220	228	213	225
142	141	140	147	173	226	227	227	213	228

118	122	118	118	121	119	118	118	116	116
119	117	118	118	116	117	115	117	119	112
119	116	116	116	117	118	117	115	119	113
120	121	120	120	116	117	115	117	117	116
117	122	116	118	120	116	114	115	114	120
119	118	117	117	121	118	118	116	116	117
118	118	119	116	118	119	119	119	116	113
118	121	117	119	117	115	118	112	118	116
117	119	115	117	120	117	115	118	116	118
118	119	116	119	116	116	117	118	114	115
118	117	119	119	118	119	118	119	118	118
117	120	117	119	117	117	119	115	119	120
120	117	119	119	115	119	119	119	119	116
119	118	118	121	121	118	116	120	119	119
120	122	118	123	119	121	121	124	118	120
118	118	118	120	120	119	120	119	119	122

Algorithmes d'extraction d'information

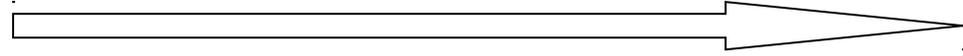


# Formation des Image d'intensités

## Monde réel 3D vers Monde 2D numérisé

### 1. Paramètres optiques :

- Type de lentille
- Focale
- Ouverture angulaire



### 2. Paramètres photométriques :

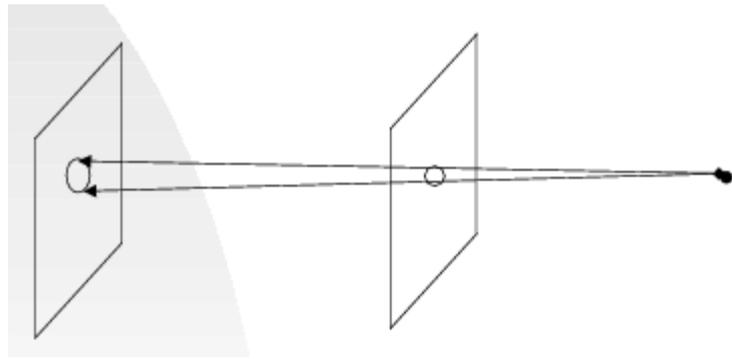
- Type, intensité et direction de l'illumination
- Propriétés de réflectance des surfaces observées

### 3. Paramètres géométriques :

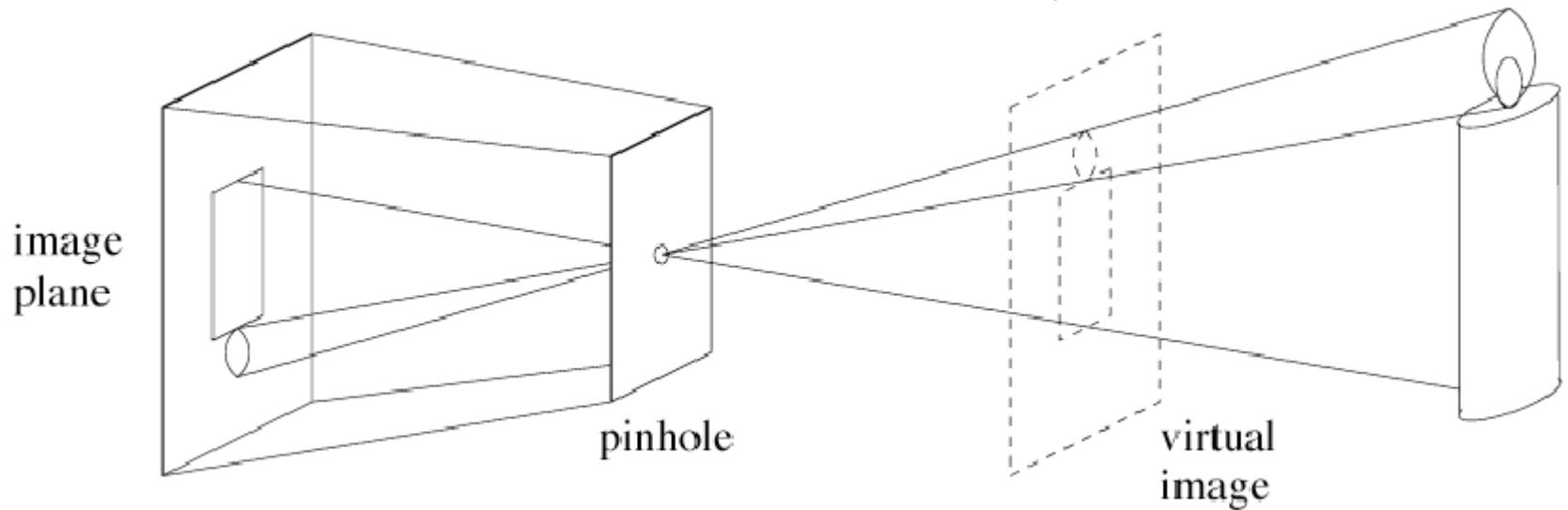
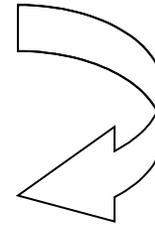
- Type de projection
- Position et orientation de la caméra
- Distortions de perspective

4. Nature discrète des photorécepteurs
5. Quantification de l'échelle d'intensité

# Pinhole Camera (ou modèle dit trou d'aiguille)



Nécessité de focaliser l'image ...



Ouverture faible donc temps d'exposition long

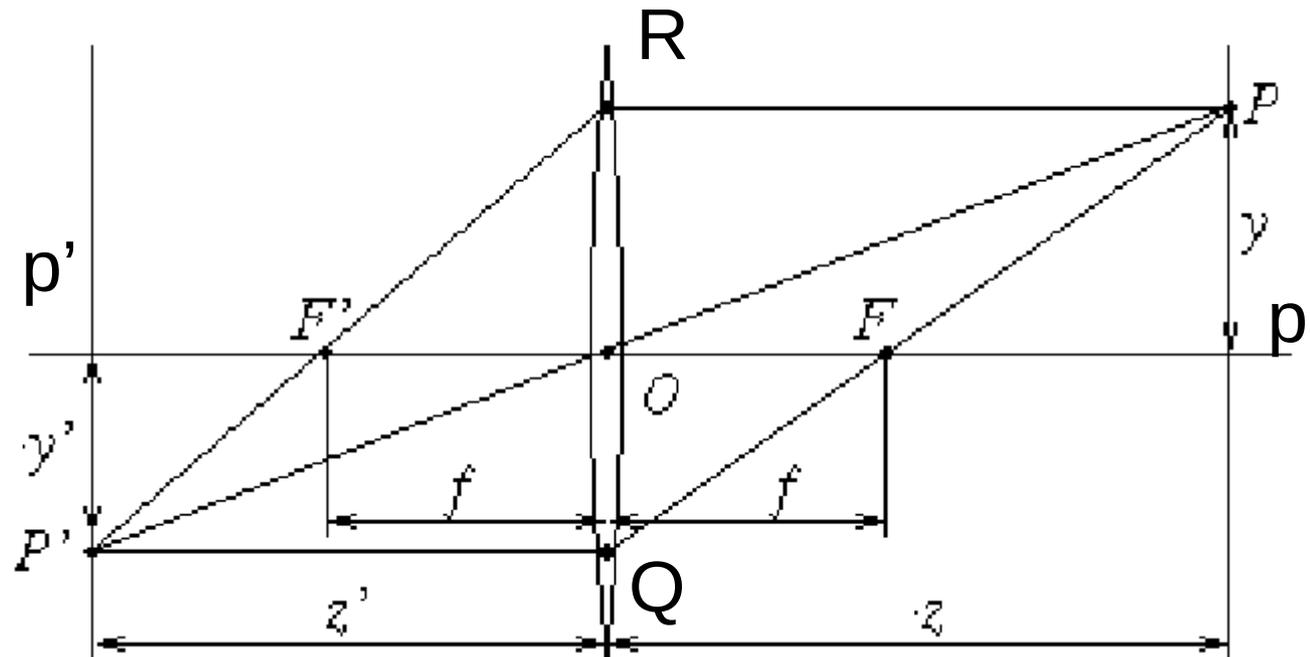


# Systèmes optiques réels

## Lentille mince

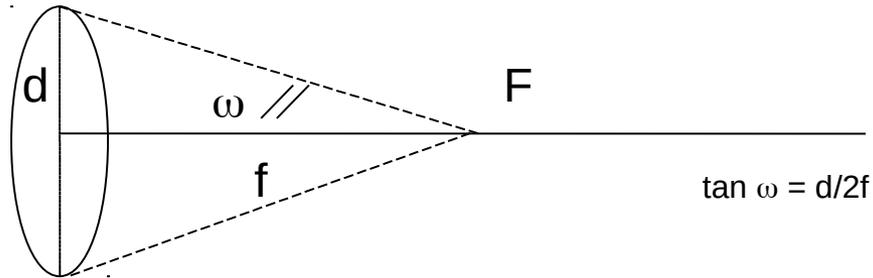
$$\frac{1}{z'} + \frac{1}{z} = \frac{1}{f}$$

Focale



## Champ visuel :

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



## Problèmes bien connus en photographie:

- 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 :

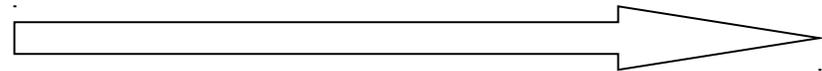
- Type de lentille
- Focale
- Ouverture angulaire

## 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



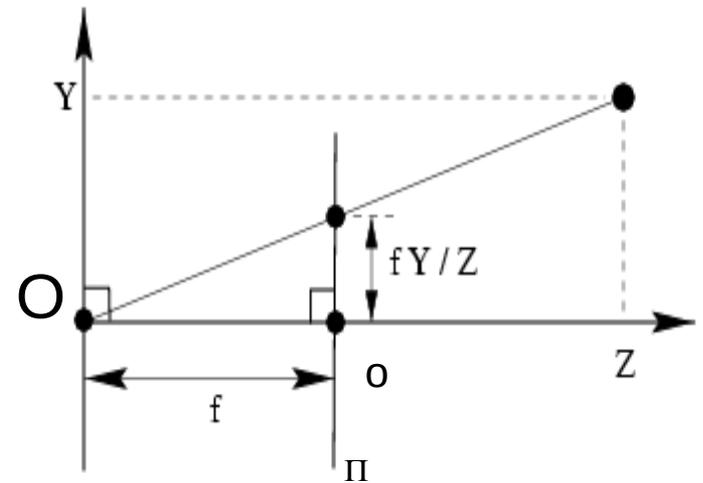
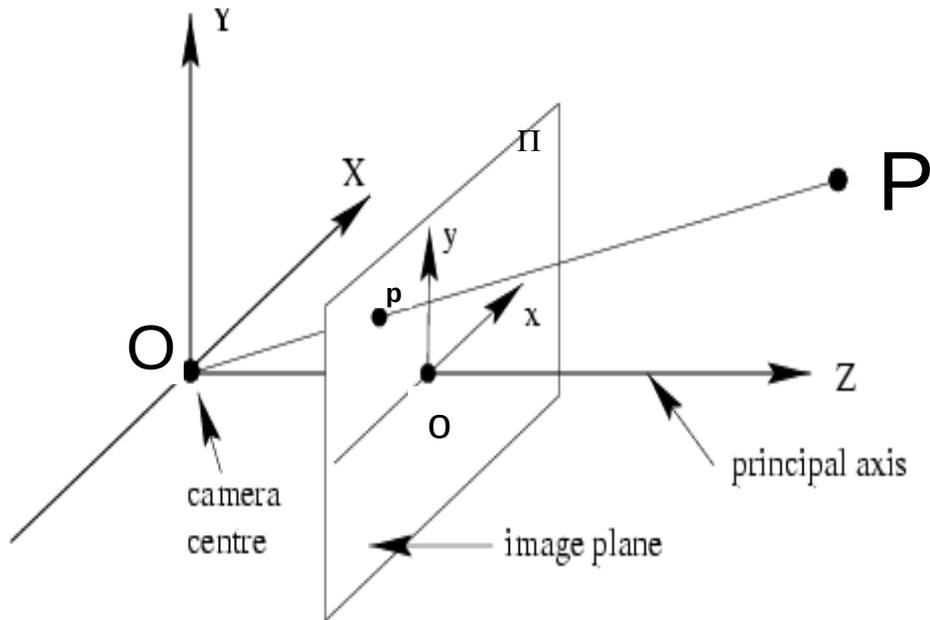
## 4. Nature discrète des photorécepteurs

## 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

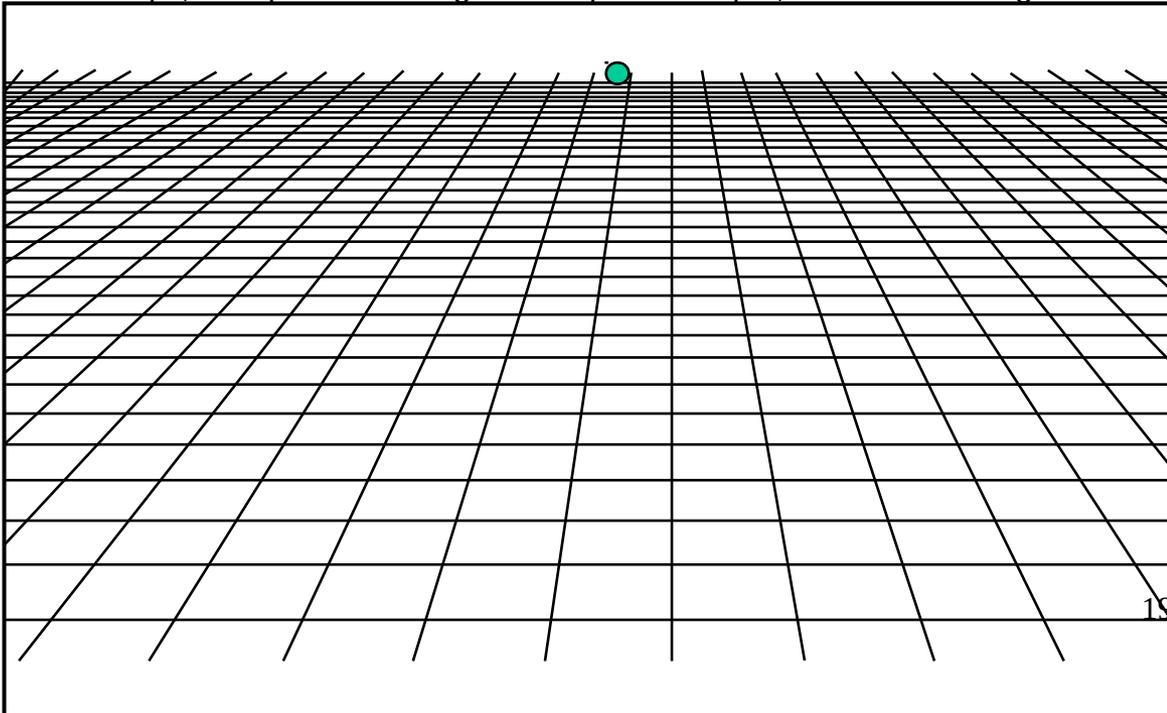
Dans le repère caméra 3D, on a :

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

Voir coordonnées homogènes

Math

- Dans le repère image 2D, on écrira  $p = [x,y]^T$
- On verra qu'en géométrie projective, un point de l'espace projectif 2D possède 3 coordonnées dites coordonnées homogènes  $p=[fX,fY,Z]^T$
- Equations non linéaires qui, d'un point de vue géométrique classique, conservent l'alignement et c'est tout



# Géométrie projective et coordonnées homogènes : simplifier l'écriture des algorithmes en linéarisant

- Coordonnées homogènes 2D :  $\bar{u}$

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

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

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

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

<https://mobidev.biz/blog/augmented-reality-trends-future-ar-technologies>

## Définitions :

Un point de l'espace projectif à  $n$  dimensions,  $\mathbb{P}^n$ , 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

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

-> 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

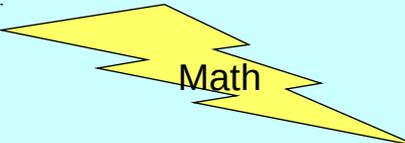
$$v = R_{\theta}u + \begin{bmatrix} 2 \\ 3 \end{bmatrix} \Leftrightarrow \bar{v} = M \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = M\bar{u}$$

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

- Identique en 3D

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

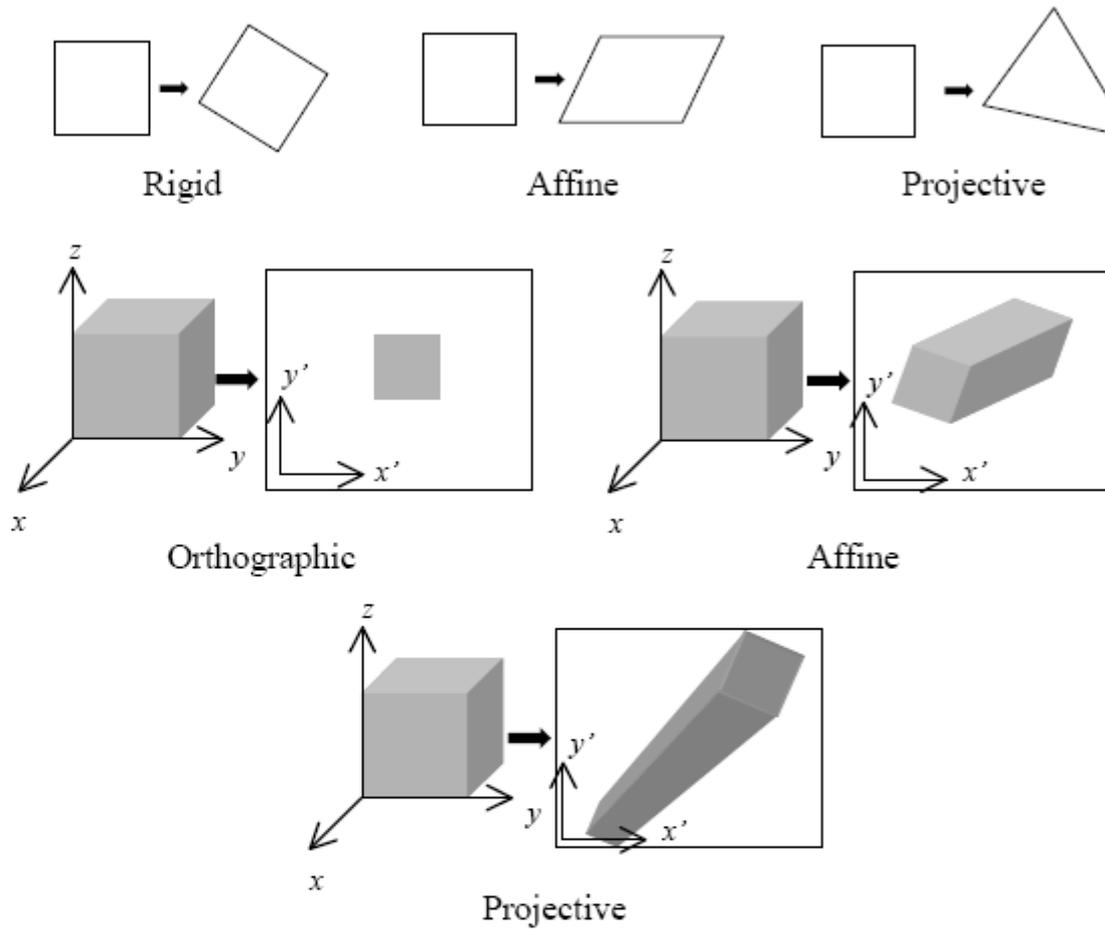
Transformation	Vector Coordinates	Homogeneous Coordinates	Degrees of Freedom	Invariants
Translation	$\mathbf{y} = \mathbf{x} + \mathbf{t}$	$\begin{bmatrix} \mathbf{I} & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{bmatrix}$	3	lengths, angles
Rotation	$\mathbf{y} = \mathbf{R}\mathbf{x}$ $\mathbf{R}^T \mathbf{R} =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



Math

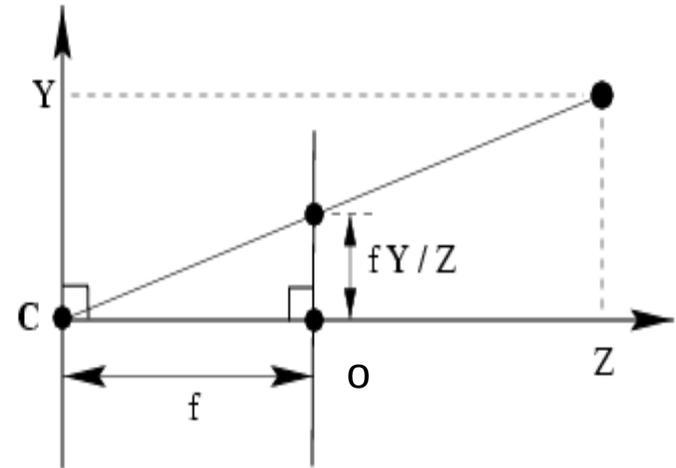
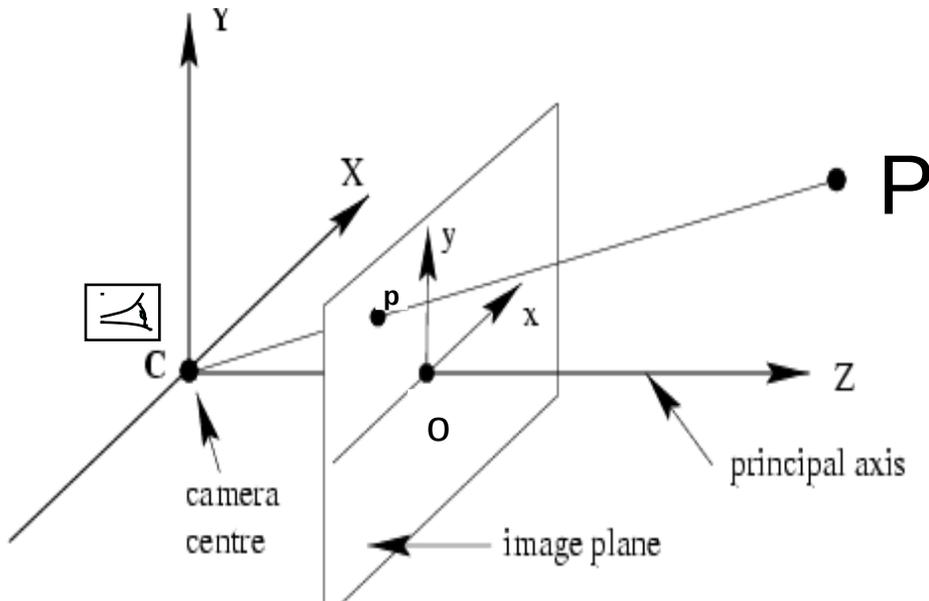
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = M \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad x' = \frac{u}{w} \text{ et } y' = \frac{v}{w}$$

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



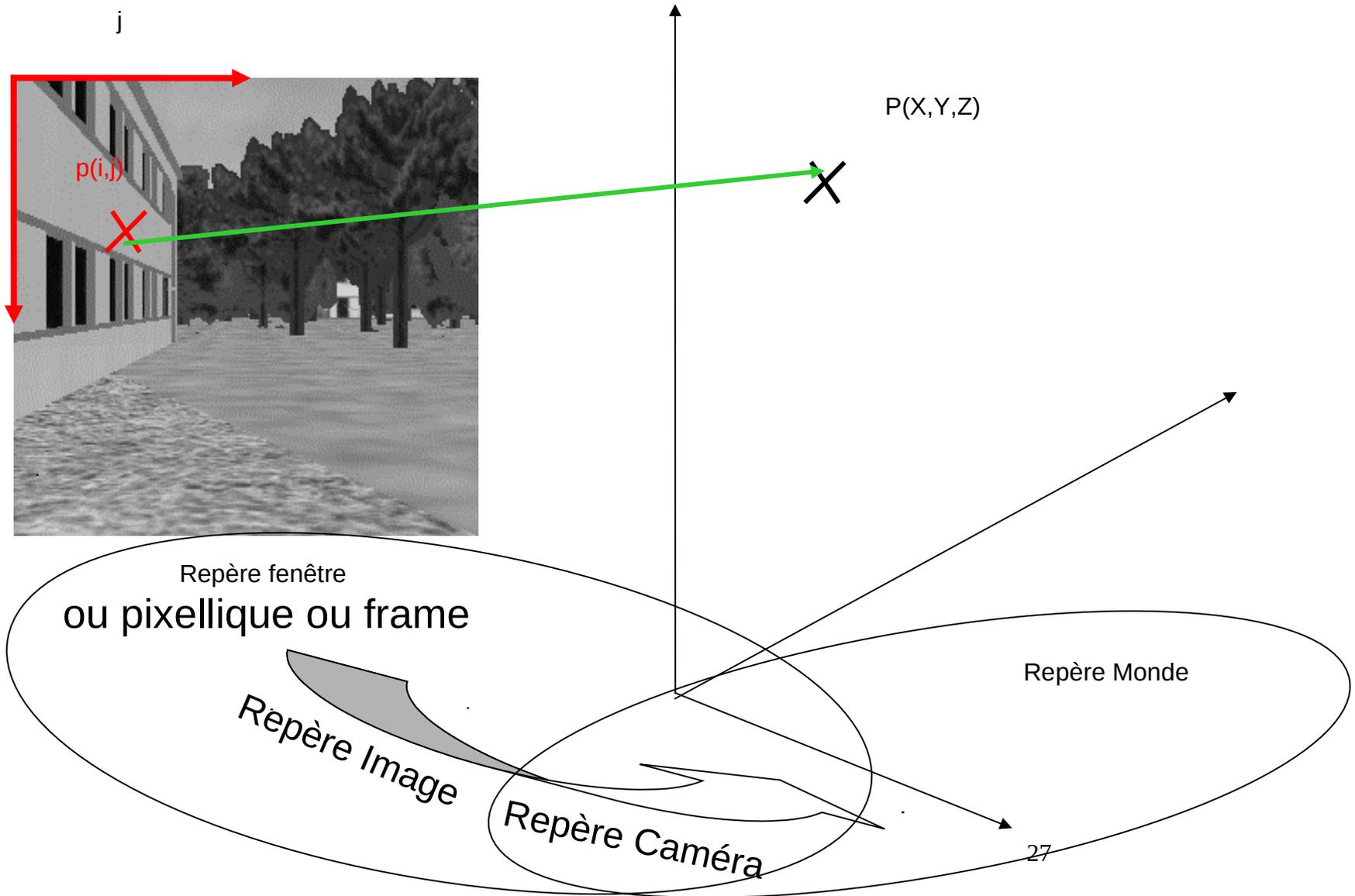
**Figure 1.** Examples of 2-D transformations (top) and 3-D to 2-D transformations (bottom).

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



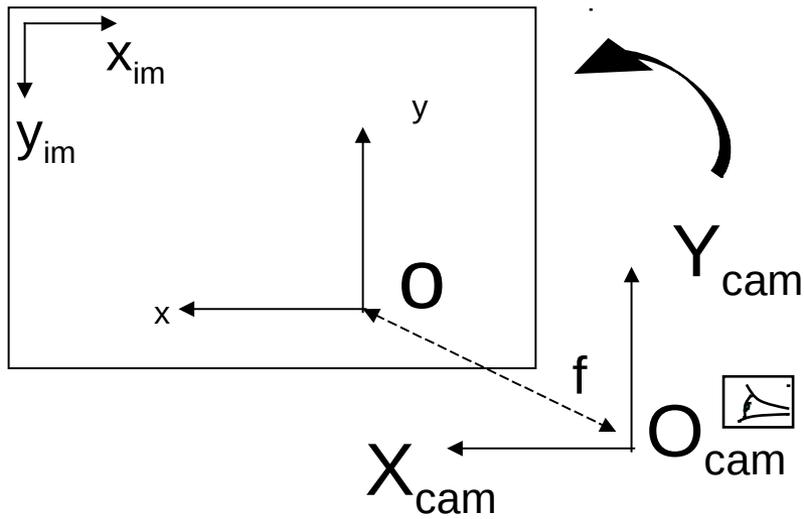
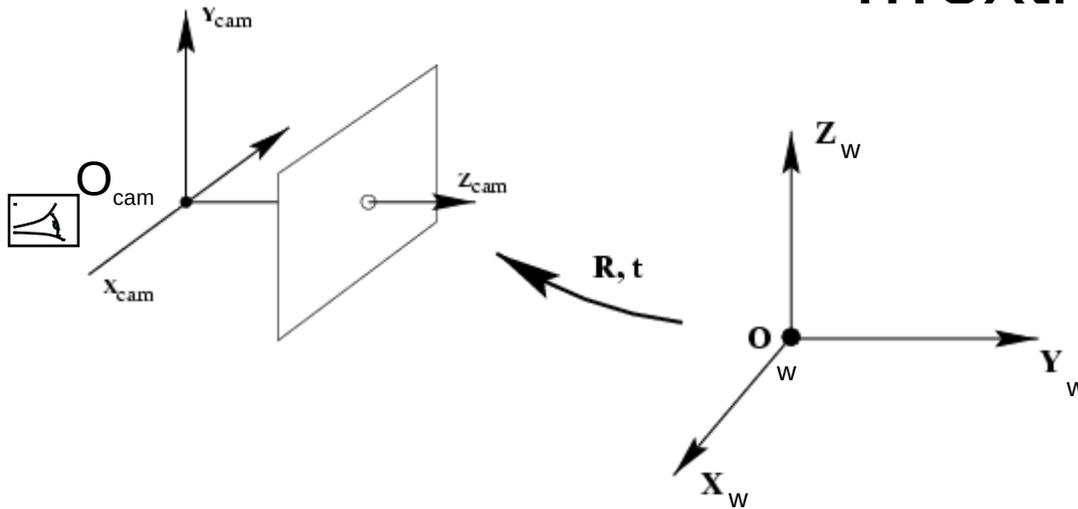
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad x = \frac{u}{w} \text{ et } y = \frac{v}{w}$$

# Paramètres de caméra...



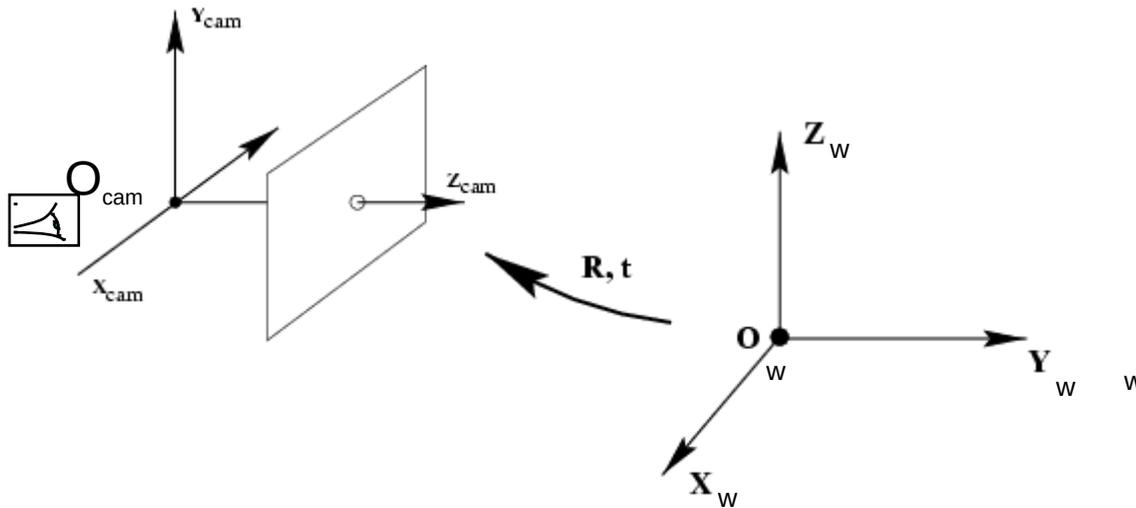
# Paramètres de caméra...

...extrinsèques



...intrinsèques

# ...extrinsèques



$$P_c = R(P_w - t)$$

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

Avec  $T = -Rt$

## Paramètres extrinsèques :

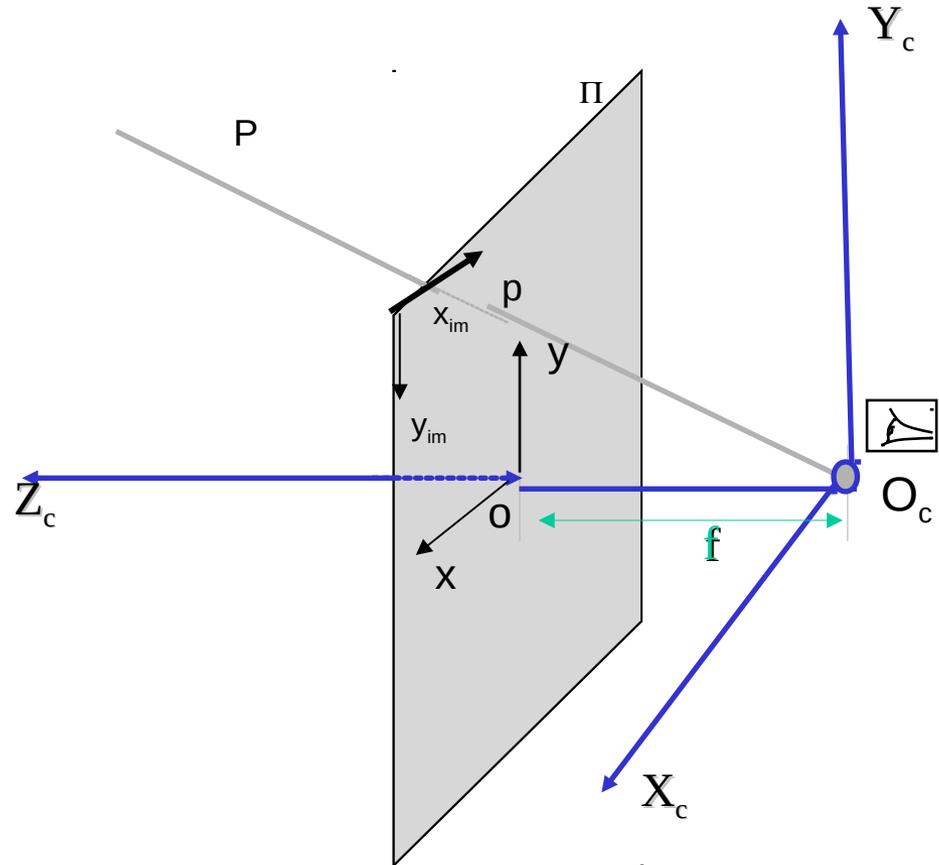
vecteur translation,  $T$ , et matrice de rotation,  $R$  (ou mieux ses paramètres libres), qui spécifient la transformation entre la caméra et le repère monde

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

## ...intrinsèques

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

- Image Plane or Focal Plane  $\Pi$
- 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 ( $x_{im}, y_{im}$ )



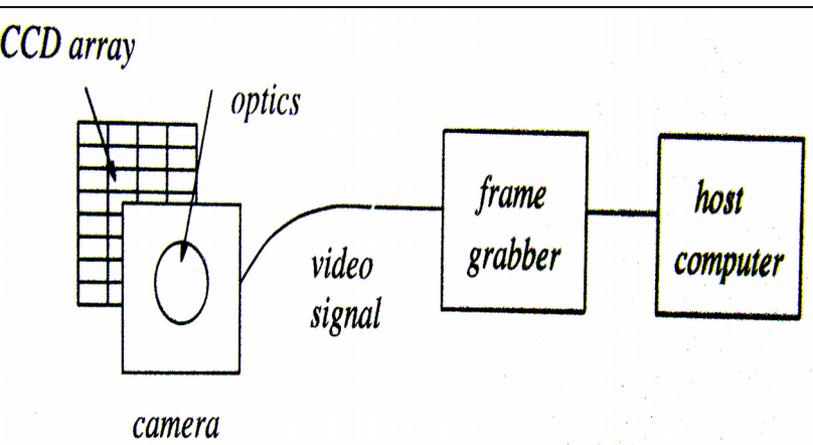
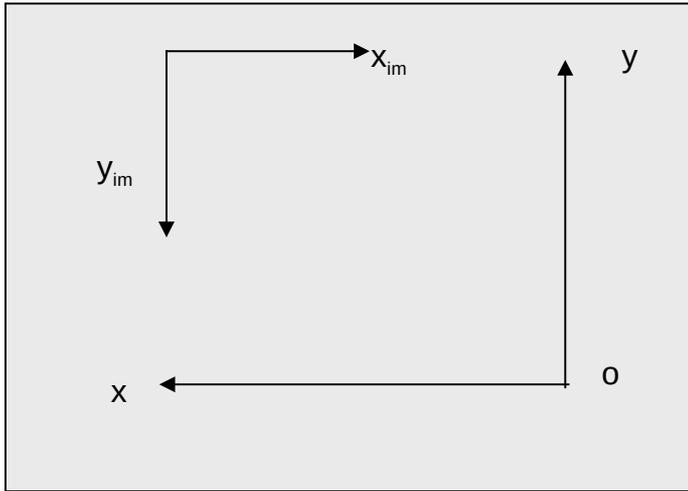
Pour un modèle de caméra perspectif, 3 ensembles de paramètres intrinsèques spécifiant :

- La projection perspective caractérisant le passage repère caméra 3D  $(X, Y, Z)$  -> repère image projectif  $(x, y)$  : paramètre  $f$
- La numérisation caractérisant le passage repère image  $(x, y)$  -> repère fenêtre CCD  $(x_{im}, y_{im})$  :  
paramètres  $o_x, o_y, s_x, s_y$
- 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  $k_1, k_2$

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

avec  $r^2 = x_d^2 + y_d^2$

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

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



## 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 :

$$1. f_x = f / s_x$$

$$2. f_y = f / s_y$$

$$3. o_x$$

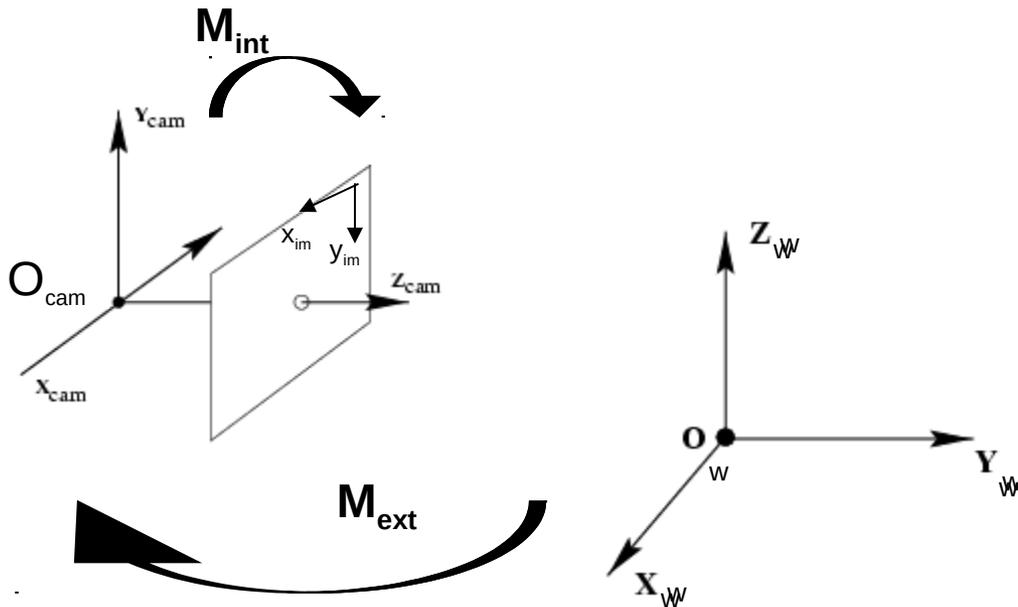
$$4. o_y$$

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

L'équation linéaire matricielle de la projection perspective

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

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



avec

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

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

<https://www.youtube.com/watch?v=E0QLVj9FJ0A>

## Paramètres extrinsèques :

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

**-> utilisation pour localiser une caméra mobile ou un objet mobile observé par une caméra fixe**

## Paramètres intrinsèques

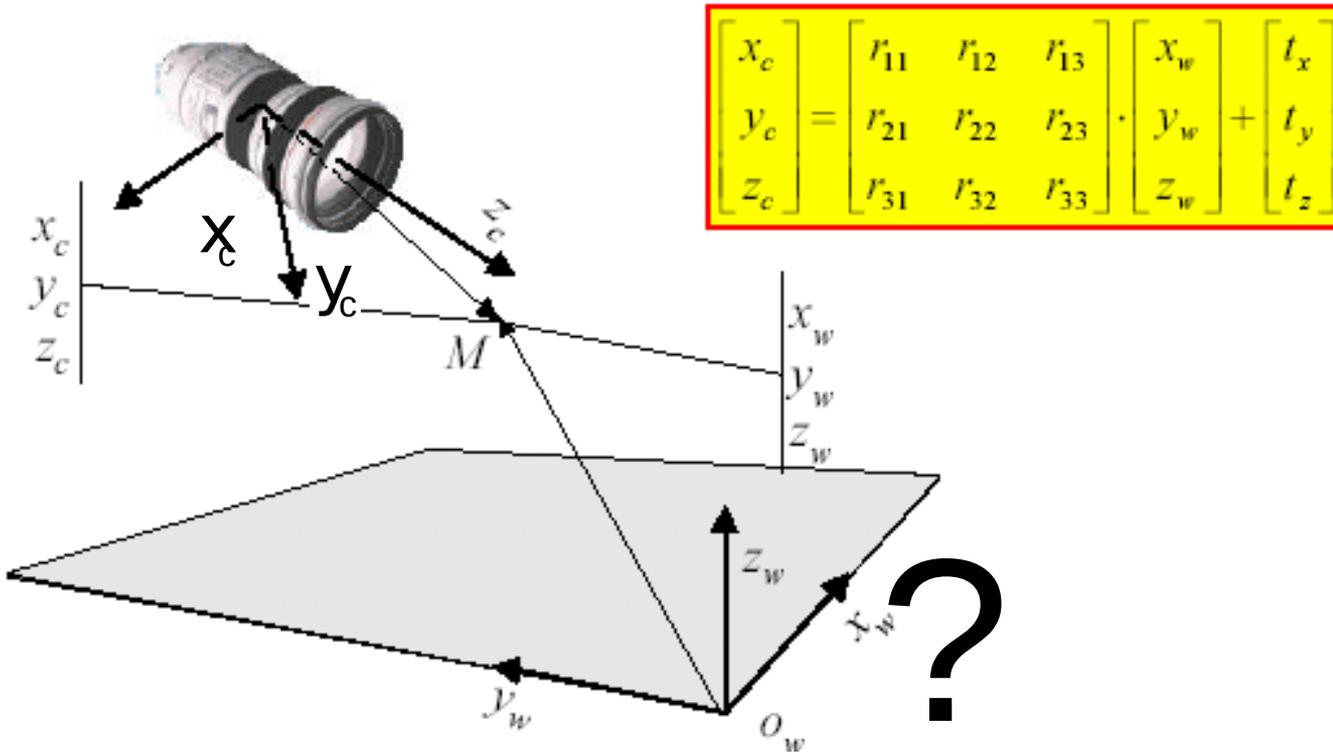
paramètres nécessaires pour lier les coordonnées pixelliques d'un point image aux coordonnées correspondantes dans le repère caméra

**-> utilisation pour dériver la vitesse réelle 3D d'observations de déplacement pixelliques dans une vidéo 2D**

Estimation de ces ou d'une partie de ces paramètres = **calibration** de caméra

# Rappel !!!

- Extrinsic Parameters



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

+ intrinsic parameters ->

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} -f_x & 0 & 0 \\ 0 & -f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} R_{3 \times 3} \\ 0_{1 \times 3} \end{bmatrix}^T \begin{bmatrix} X^{(w)} \\ Y^{(w)} \\ Z^{(w)} \\ 1 \end{bmatrix}$$

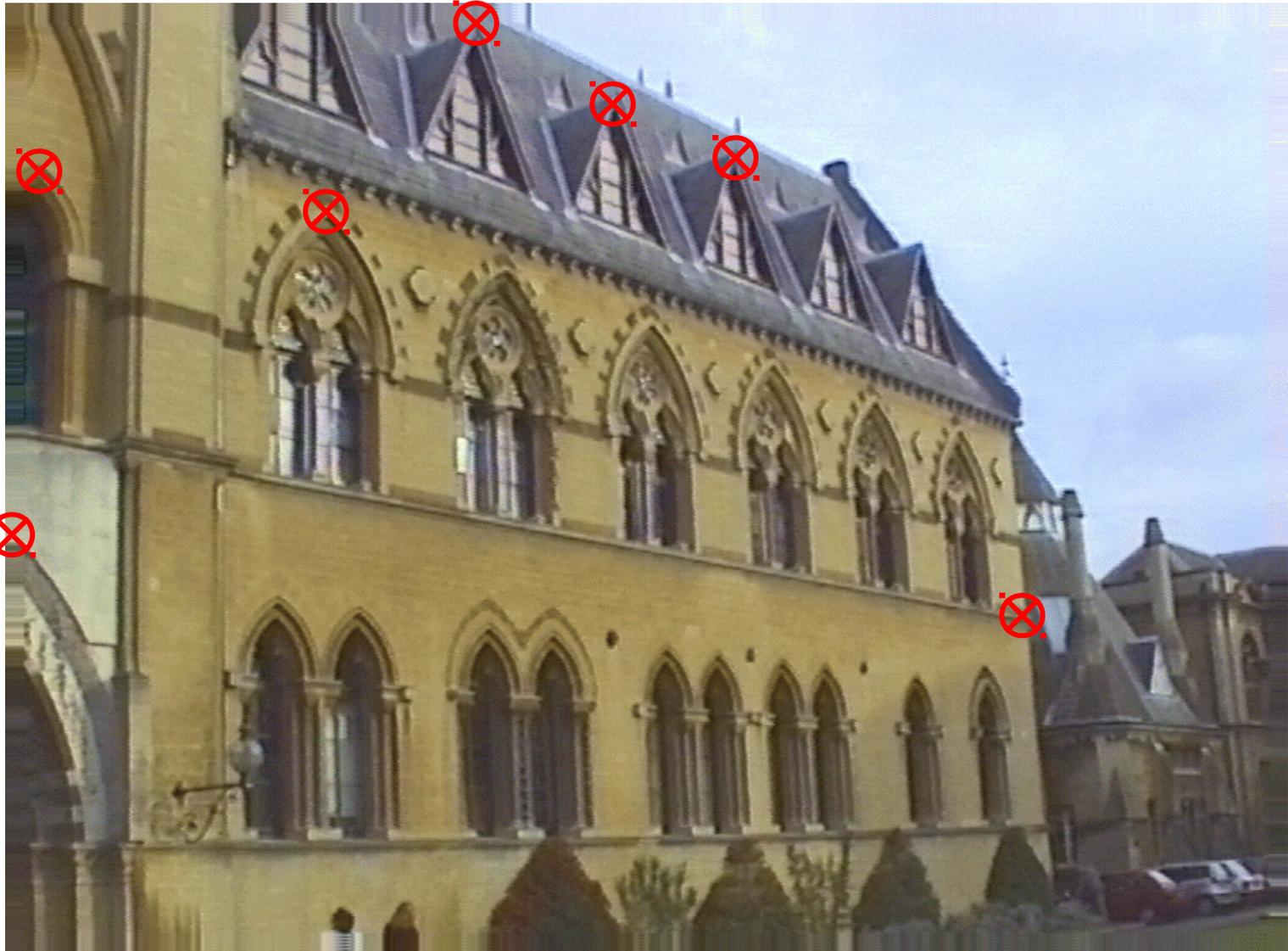
## Problème de la calibration

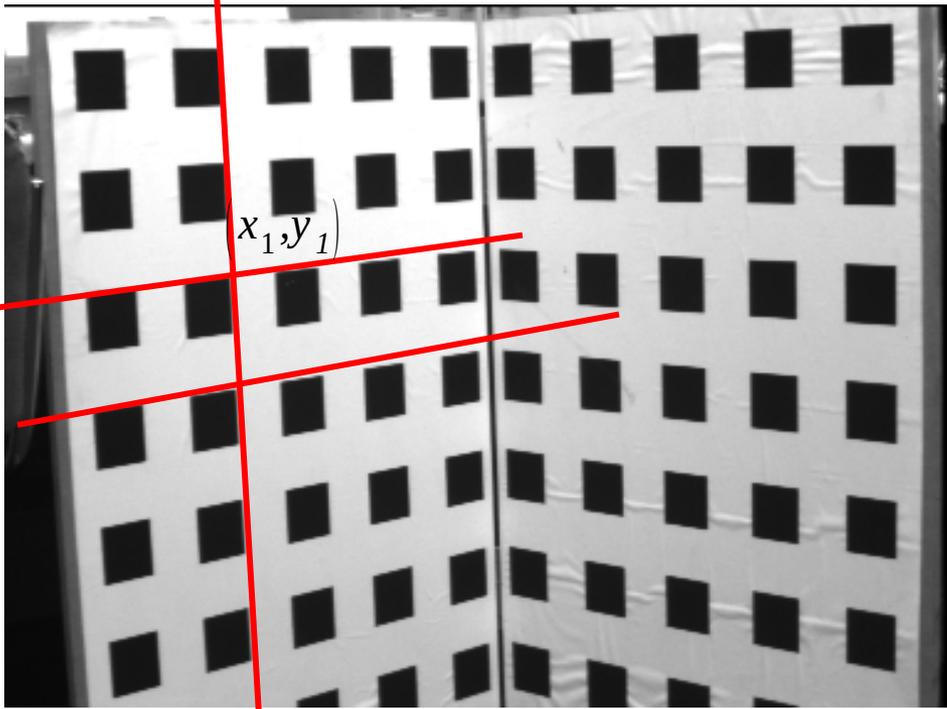
Étant donné une ou plusieurs images d'un motif de calibration, estimer :

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

-> la caméra est utilisée comme instrument de mesure

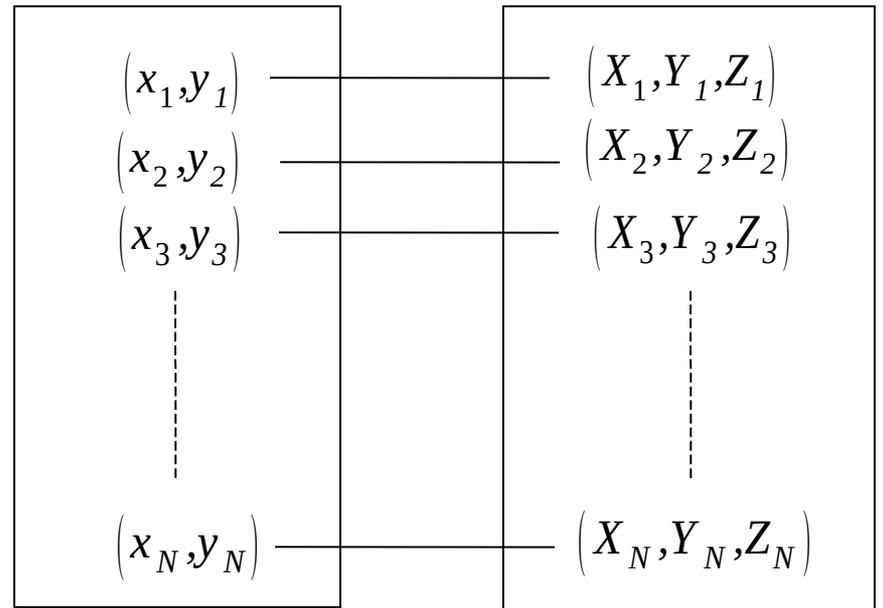
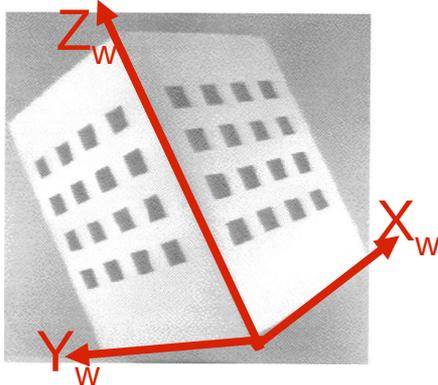
-> pas indispensable mais ouvre la voie à des algorithmes de reconstruction et reconnaissance 3D performants





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

Motif de calibration



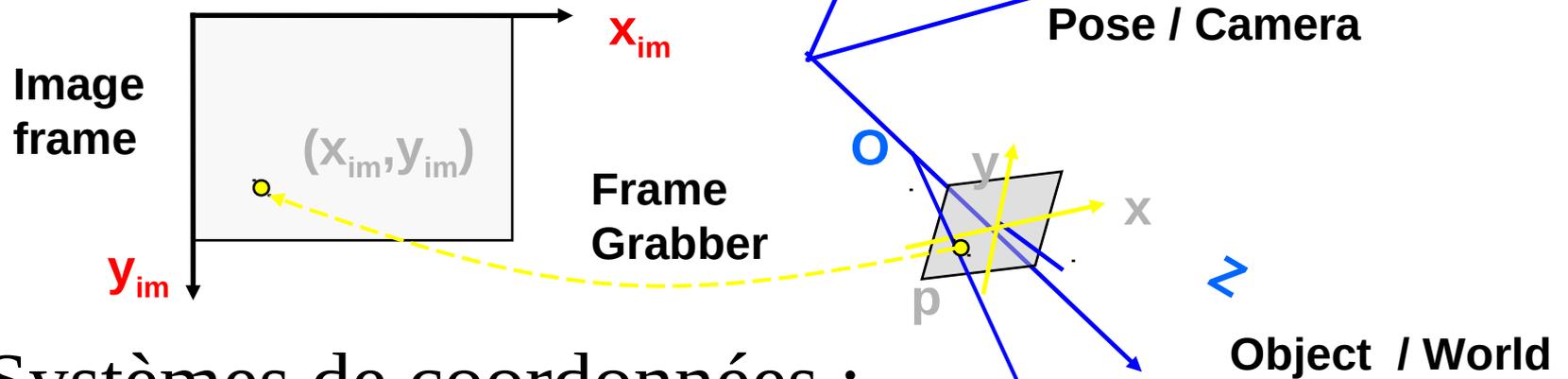
## 2 méthodes reconnues :

### 1. par estimation directe des paramètres

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

### 1. par matrice de projection

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



- Systèmes de coordonnées :

- Coordonnées fenêtre  $(x_{im}, y_{im})$  pixels
- Coordonnées image  $(x, y)$  in mm
- Coordonnées caméra  $(X, Y, Z)$
- Coordonnées monde  $(X_w, Y_w, Z_w)$

- Paramètres de caméra :

- Paramètres intrinsèques (de la camera et du "frame grabber") : relie les "**frame coordinates**" d'un point image point à ses "**camera coordinates**" correspondantes
- Paramètres extrinsèques : définissent la localisation et l'orientation du "**camera coordinate system**" par rapport au "**world coordinate system**"

- **World to Camera**

- Camera:  $P = (X, Y, Z)^T$
- World:  $P_w = (X_w, Y_w, Z_w)^T$
- Transformations : R, T

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

- **Camera to Image**

- Camera:  $P = (X, Y, Z)^T$
- Image:  $p = (x, y)^T$
- 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 :
  - On suppose connues  $(o_x, o_y)$ , le centre de l'image, et on estime tous les autres paramètres 
  - Puis on calcule  $(o_x, o_y)$ .

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

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

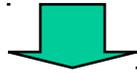
- Paramètres Extrinsèques :
  - **R**, matrice de rotation 3x3
    - Trois angles  $\alpha, \beta, \gamma$
  - **T**, vecteur translation 3-D
  
- Paramètres Intrinsèques :
  - $\alpha = f_x/f_y = s_y/s_x$ , et  $f_x$
  - $(o_x, o_y)$ : on suppose le **centre de l'image connu** ->  $(x', y')$  connues
  - $k_1$ , coefficient de distorsion radiale : **négligé dans l'algorithme de base**

## Étape 1

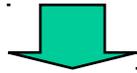
Supposant que la localisation du centre de l'image  $(o_x, o_y)$  est connu, et que la distorsion radiale peut être négligée, on doit estimer  $f_x$ ,  $\alpha$ ,  $R$  et  $T$  à partir des points images  $(x_i, y_i)$ ,  $i = 1$  à  $N$ , projections des  $N$  points connus  $(X_i, Y_i, Z_i)$  dans 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  $\mathbf{v} = (v_1, \dots, v_8)^T$

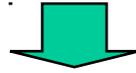
$$x_i X_i v_1 + x_i Y_i v_2 + x_i Z_i v_3 + x_i v_4 - y_i X_i v_5 - y_i Y_i v_6 - y_i Z_i v_7 - y_i v_8 = 0$$

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

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

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

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

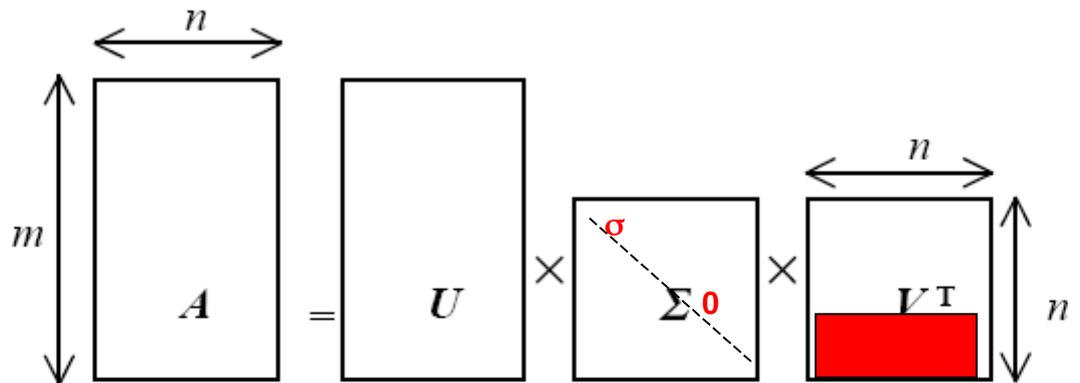


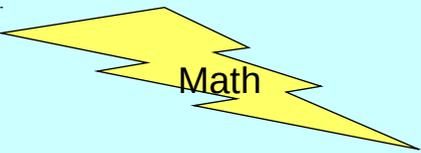
$$\mathbf{A}\mathbf{v} = \mathbf{0}$$

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

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<sup>ème</sup> 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

Cette décomposition s'applique à une matrice rectangulaire.

Elle permet 3 applications importantes :

- Résoudre des systèmes d'équations linéaires non homogènes,
- Résoudre 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).

## Théorème :

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

- $U \in M_{m \times m}$  et  $V \in M_{n \times n}$  et leurs colonnes correspondent à des vecteurs unités mutuellement orthogonaux;
- $D \in M_{m \times n}$  est une matrice diagonale dont les éléments diagonaux,  $\sigma_i$ , 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

## 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.

## Propriété 2 :

Si  $A$  est une matrice rectangulaire, le nombre de  $\sigma_i$  non nuls égale le rang de  $A$ .

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

## Propriété 3 :

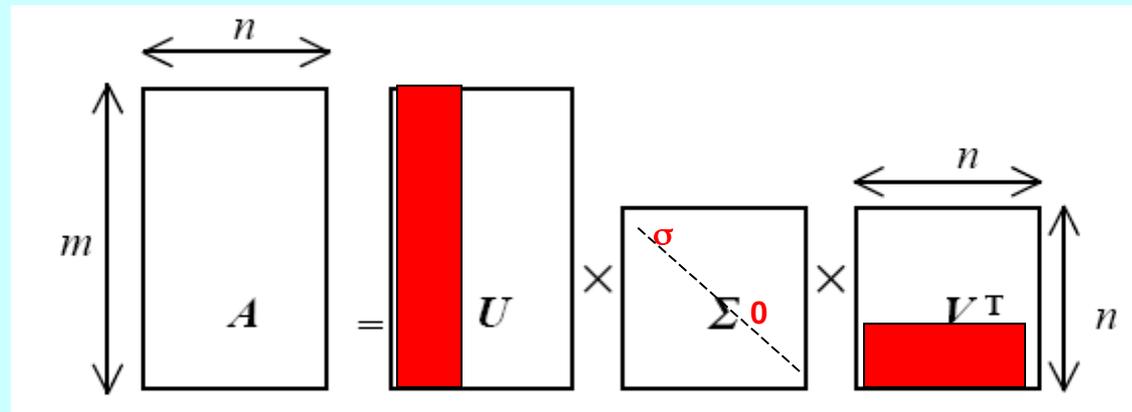
Si  $A$  est une matrice carrée, non singulière, son inverse peut s'écrire  $A^{-1} = V D^{-1} U^T$

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  $A^+ = A^{-1}$

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  $A$ .

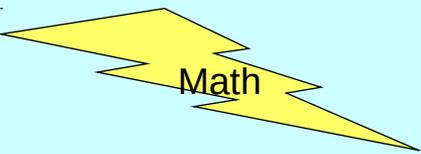


Propriété 5 :

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  $\sigma_k$ .



$$\begin{matrix} \boxed{A} \\ \boxed{x} \end{matrix} = \begin{matrix} \boxed{b} \end{matrix}$$

**S**ingular  
**V**alue  
**D**ecomposition

Première Application :

Résolution de systèmes d'équations linéaires non homogènes

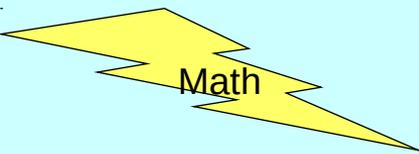
Soit le système de  $m$  équations linéaires à  $n$  inconnues  $Ax=b$

- $A \in M_{m \times n}$  la matrice du système linéaire non homogène
- $x \in \mathcal{R}_n$  le vecteur des inconnues
- $b \in \mathcal{R}_m$  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  $(A^T A)^{-1}$ .



$$\begin{matrix} \boxed{A} \\ \times \\ \boxed{x} \end{matrix} = \boxed{0}$$

**S**ingular  
**V**alue  
**D**ecomposition

### 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  $A$ ).

$$\begin{matrix} m \\ \updownarrow \\ \boxed{A} \\ \updownarrow \\ n \end{matrix} = \begin{matrix} \boxed{U} \\ \times \\ \begin{matrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_n \\ \text{---} \\ 0 \end{matrix} \\ \times \\ \begin{matrix} \boxed{V^T} \\ \updownarrow \\ n \end{matrix} \end{matrix}$$

## Dernière Application :

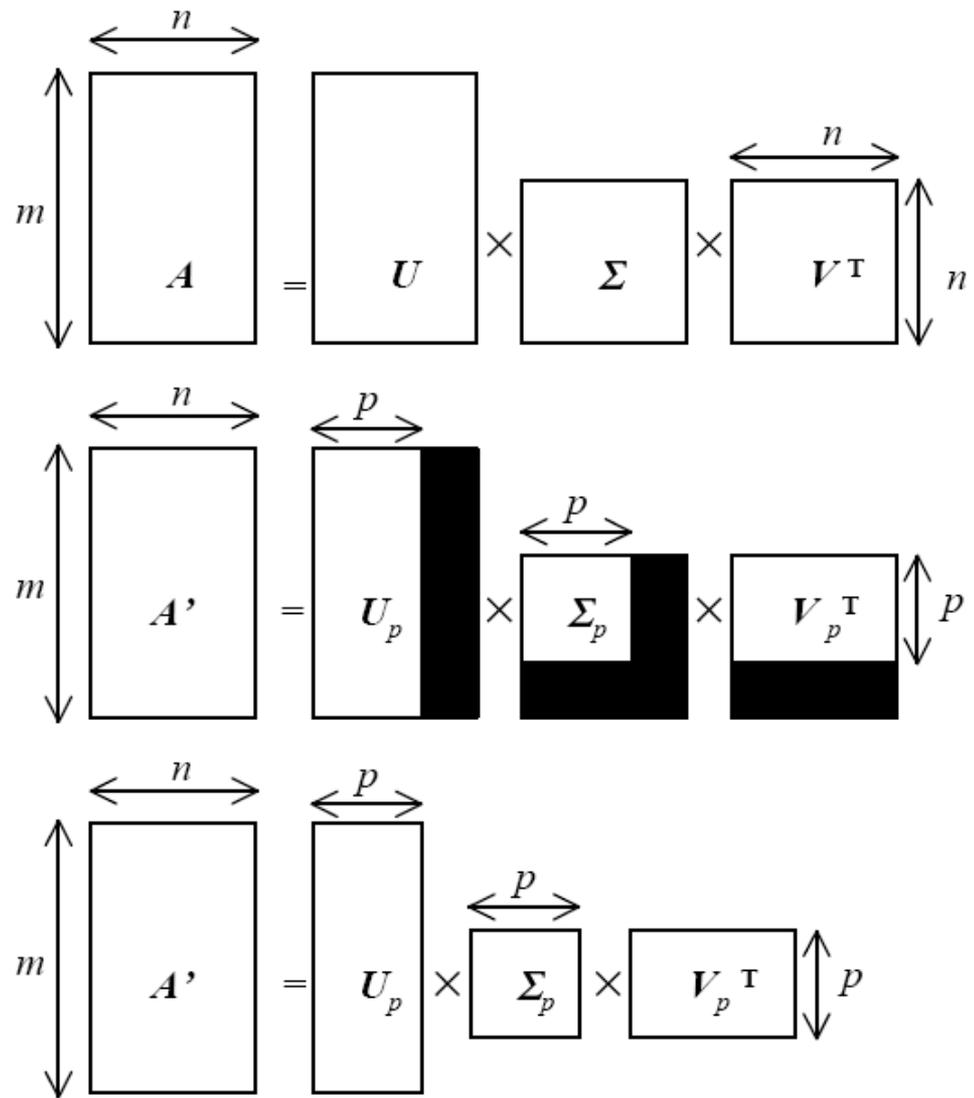
### Assurer des contraintes

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  $\gamma$  and aspect ratio  $\alpha$

- Knowledge:  $\bar{\mathbf{v}}$  is an orthogonal matrix  $\bar{\mathbf{v}} = \gamma (r_{21}, r_{22}, r_{23}, T_y, \alpha r_{11}, \alpha r_{12}, \alpha r_{13}, \alpha T_x)$

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

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

- $$\mathbf{R}_i^T \mathbf{R}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$$\mathbf{R} = (r_{ij})_{3 \times 3} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix}$$

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

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

# Rotation R and Translation T

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

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

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

$$sR_1^T, sR_2^T, sT_x, sT_y$$

- The third row of the rotation matrix by vector product

$$R_3^T = R_1^T \times R_2^T = sR_1^T \times sR_2^T$$

- Remaining Questions :

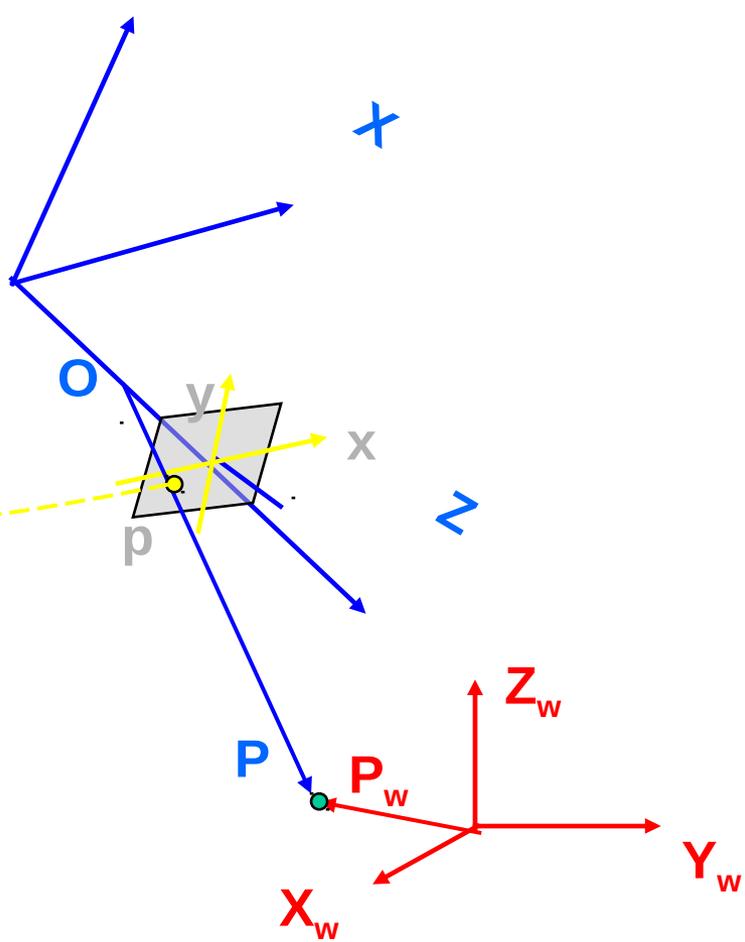
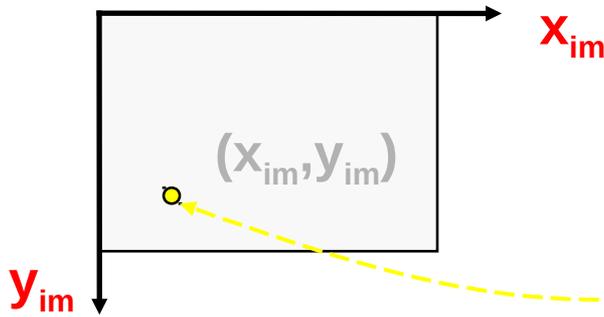
– How to find the sign  $s$ ?

– Is R orthogonal?

– How to find  $T_z$  and  $f_x$ ?

$$R = (r_{ij})_{3 \times 3} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix}$$

# Find the sign s



Facts:

$$-f_x > 0$$

$$-Z_c > 0$$

- x known

-  $X_w, Y_w, Z_w$  known

Solution

⇒ Check the sign of  $X_c$

⇒ 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

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

$$R = (r_{ij})_{3 \times 3} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} R_1^T \\ R_2^T \\ R_3^T \end{bmatrix}$$

$$R_3^T = R_1^T \times R_2^T = sR_1^T \times sR_2^T$$

- Solution:
  - Use SVD of estimate R

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



Replace the diagonal matrix D with the 3x3 identity matrix



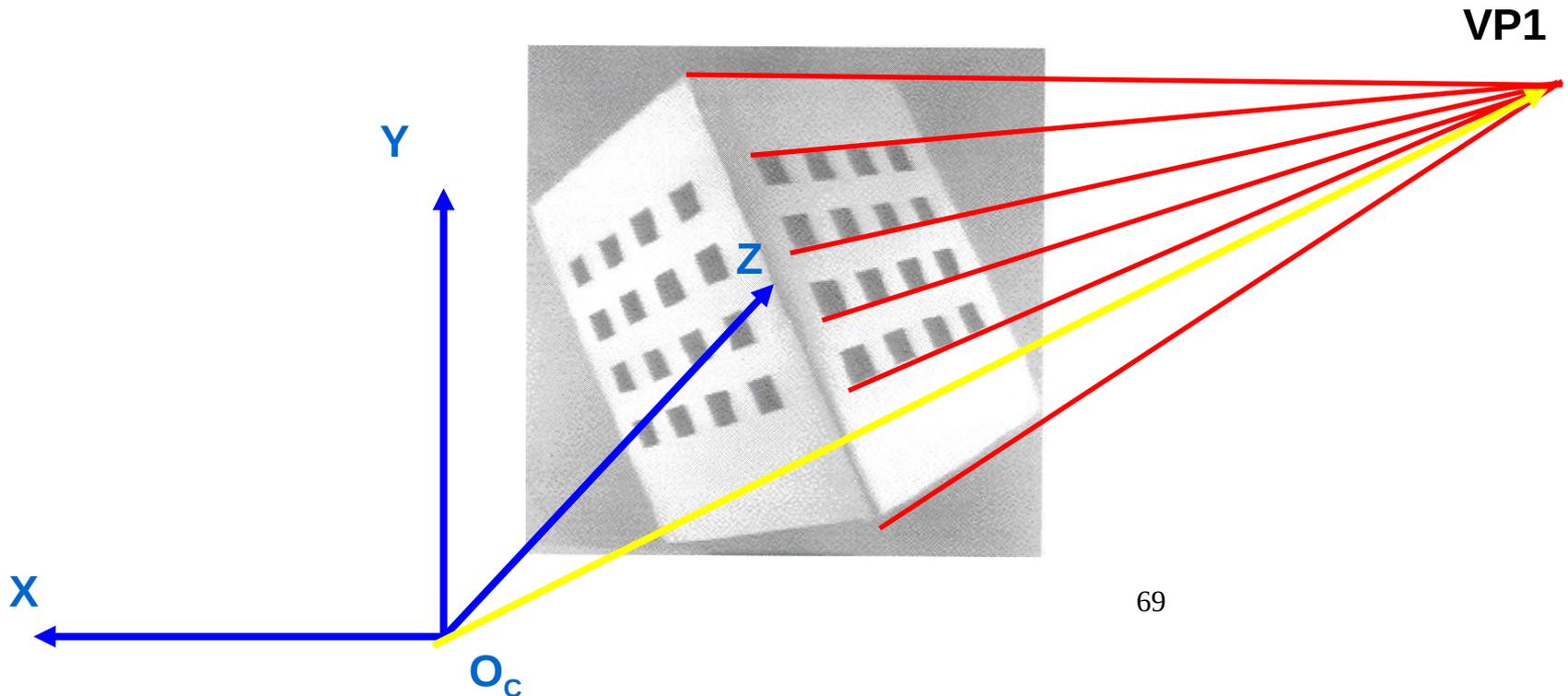
## Étape 2

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

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

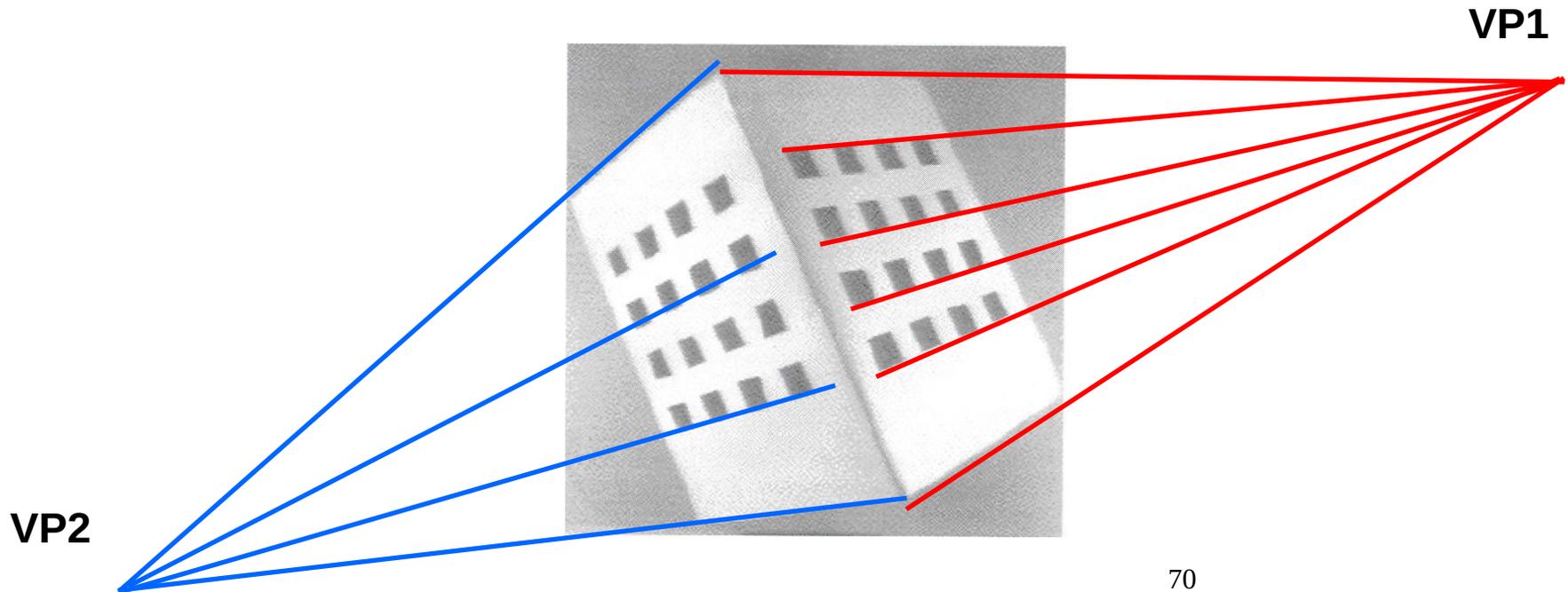
## Définition : Point de fuite

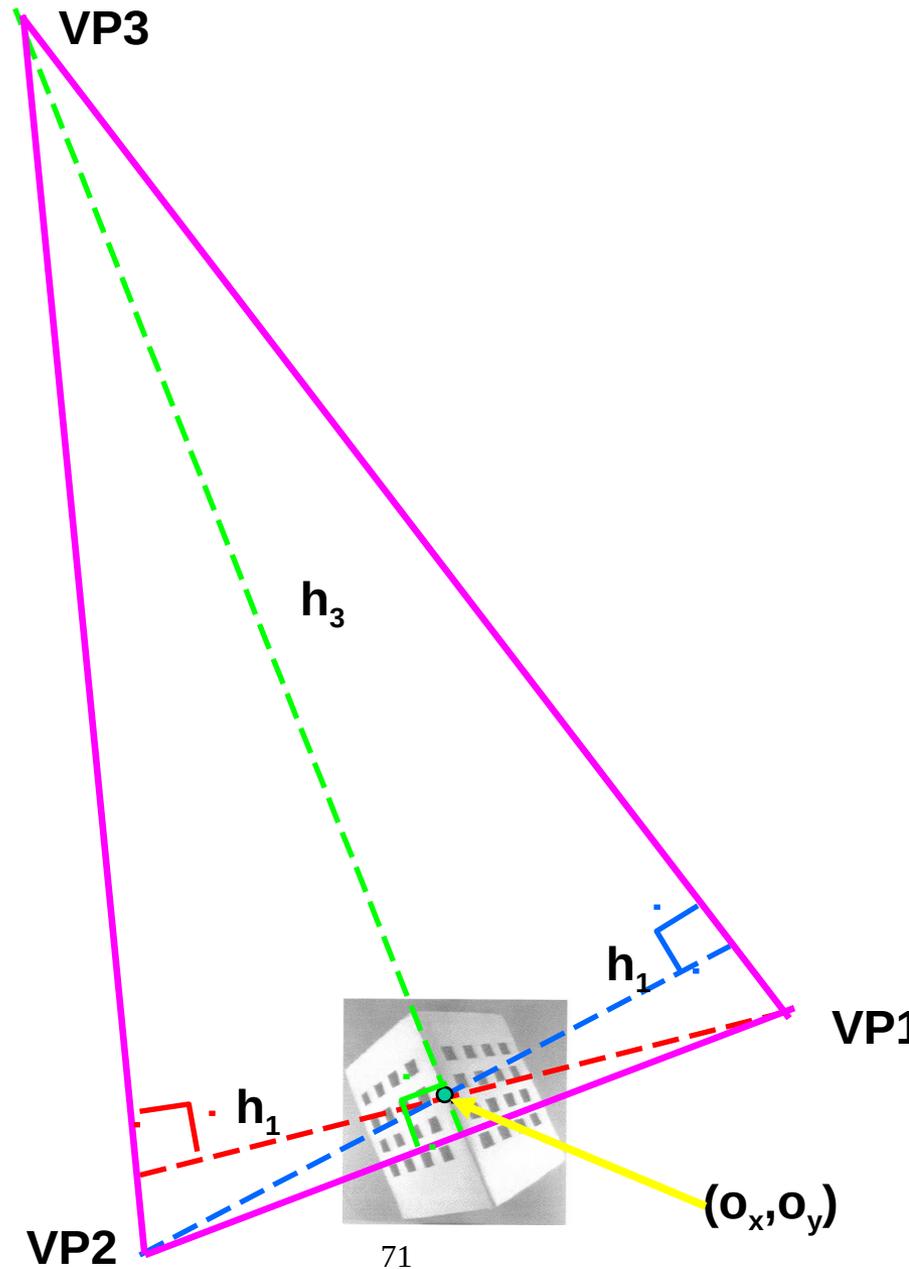
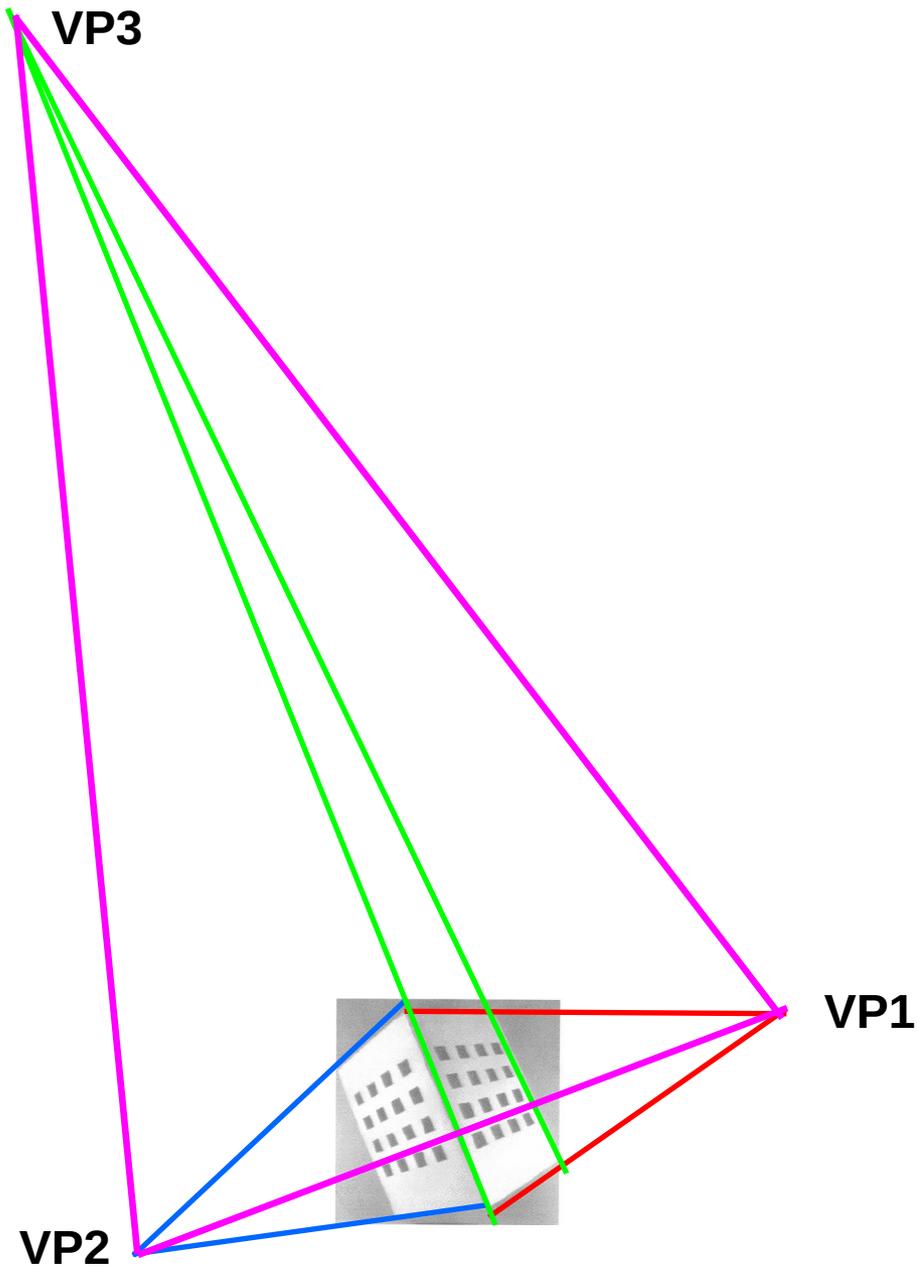
Soient  $L_i$ ,  $i=1, \dots, N$  droites parallèles dans l'espace 3D, et  $l_i$  les droites images correspondantes. En raison de la projection perspective, les droites  $L_i$  apparaissent comme se rencontrant en un point  $p$ , appelé point de fuite, défini comme l'intersection commune de toutes les 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  $T$ .



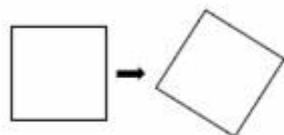


Piste pour la preuve du théorème :

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

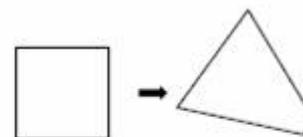
	Euclidean	similarity	affine	projective
<b>Transformations</b>				
rotation	X	X	X	X
translation	X	X	X	X
uniform scaling		X	X	X
nonuniform scaling			X	X
shear			X	X
perspective projection				X
composition of projections				X
<b>Invariants</b>				
length	X			
angle	X	X		
ratio of lengths	X	X		
parallelism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



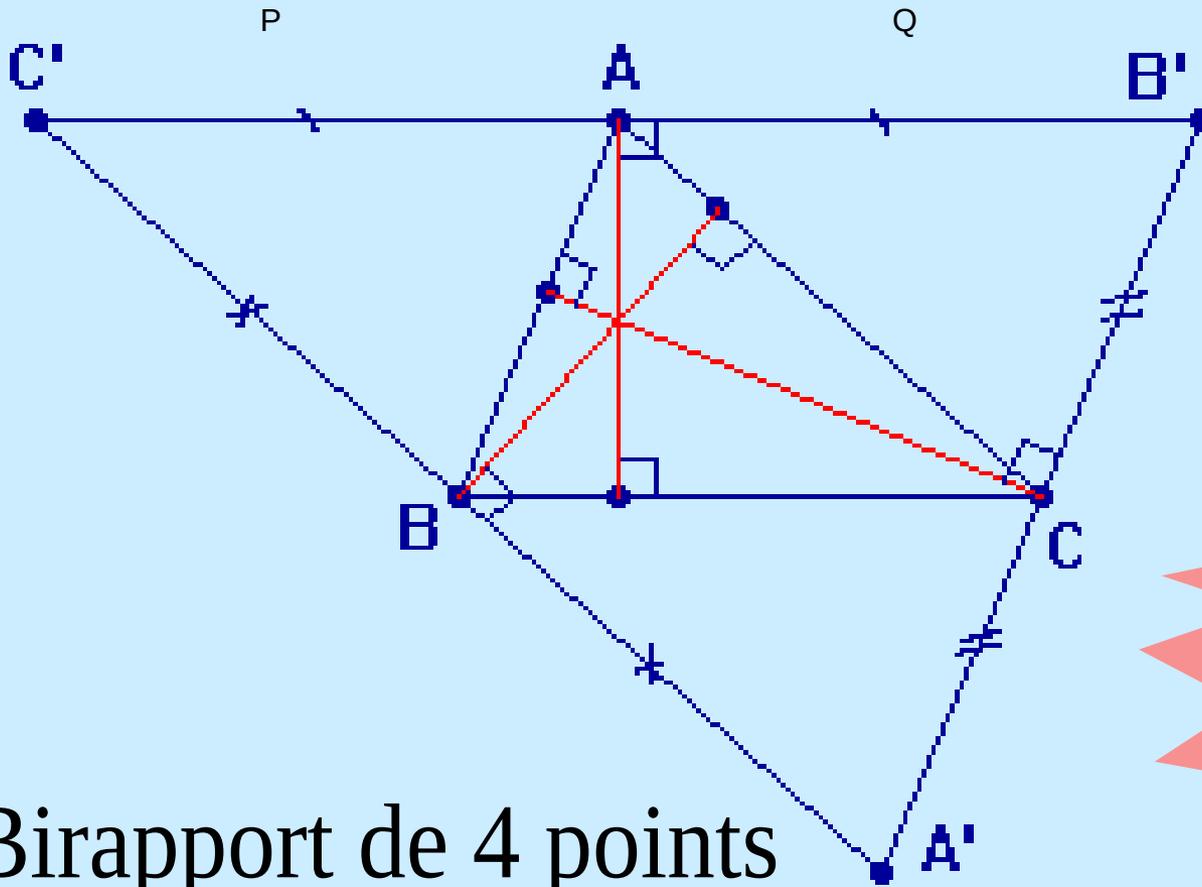
Rigid



Affine



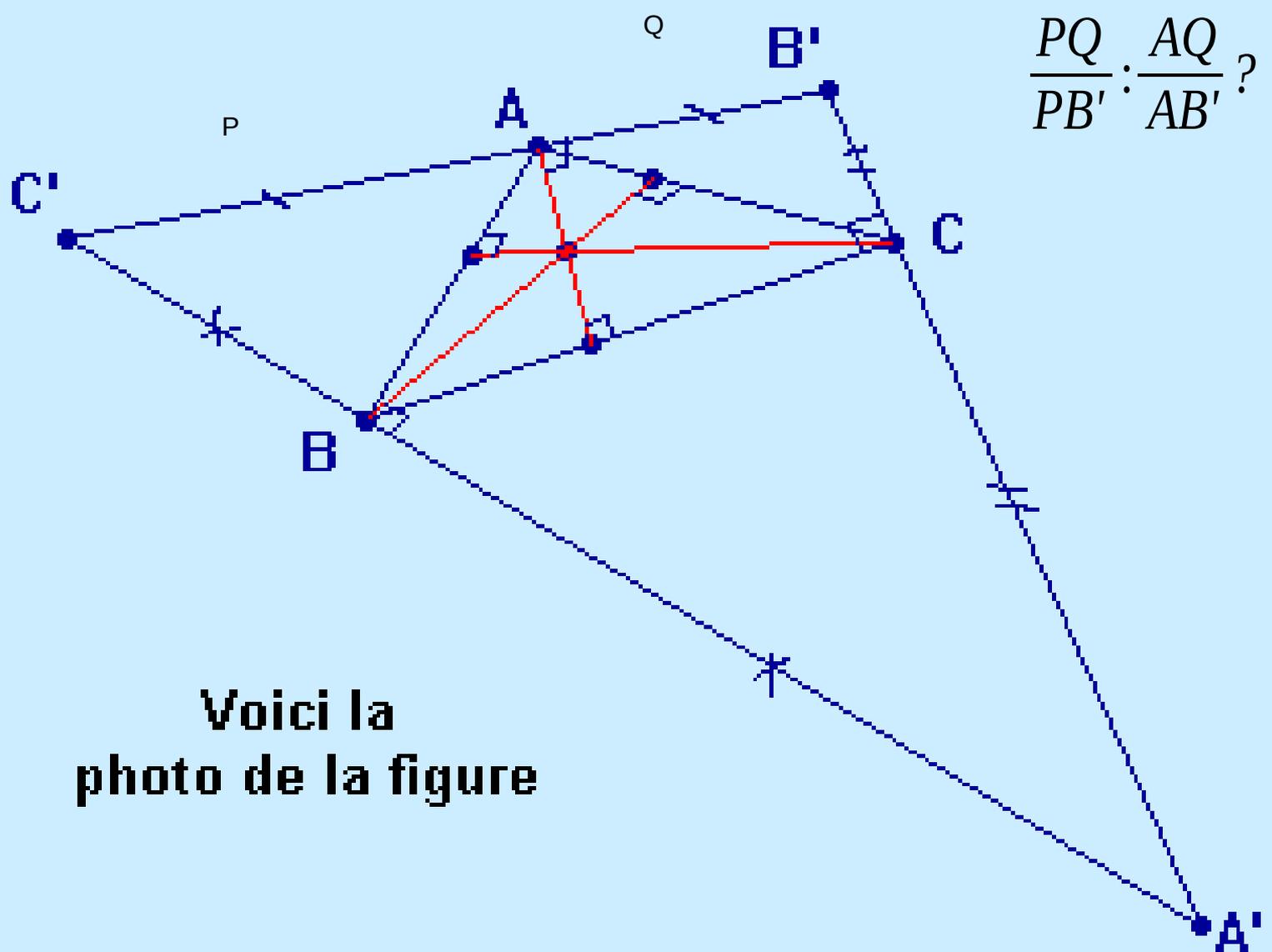
Projective



Birapport de 4 points

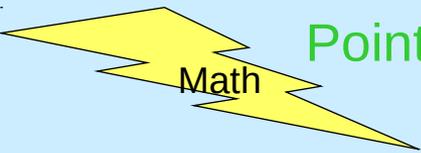
$$(P, A; Q, B') = \frac{PQ}{PB'} \cdot \frac{AQ}{AB'}$$



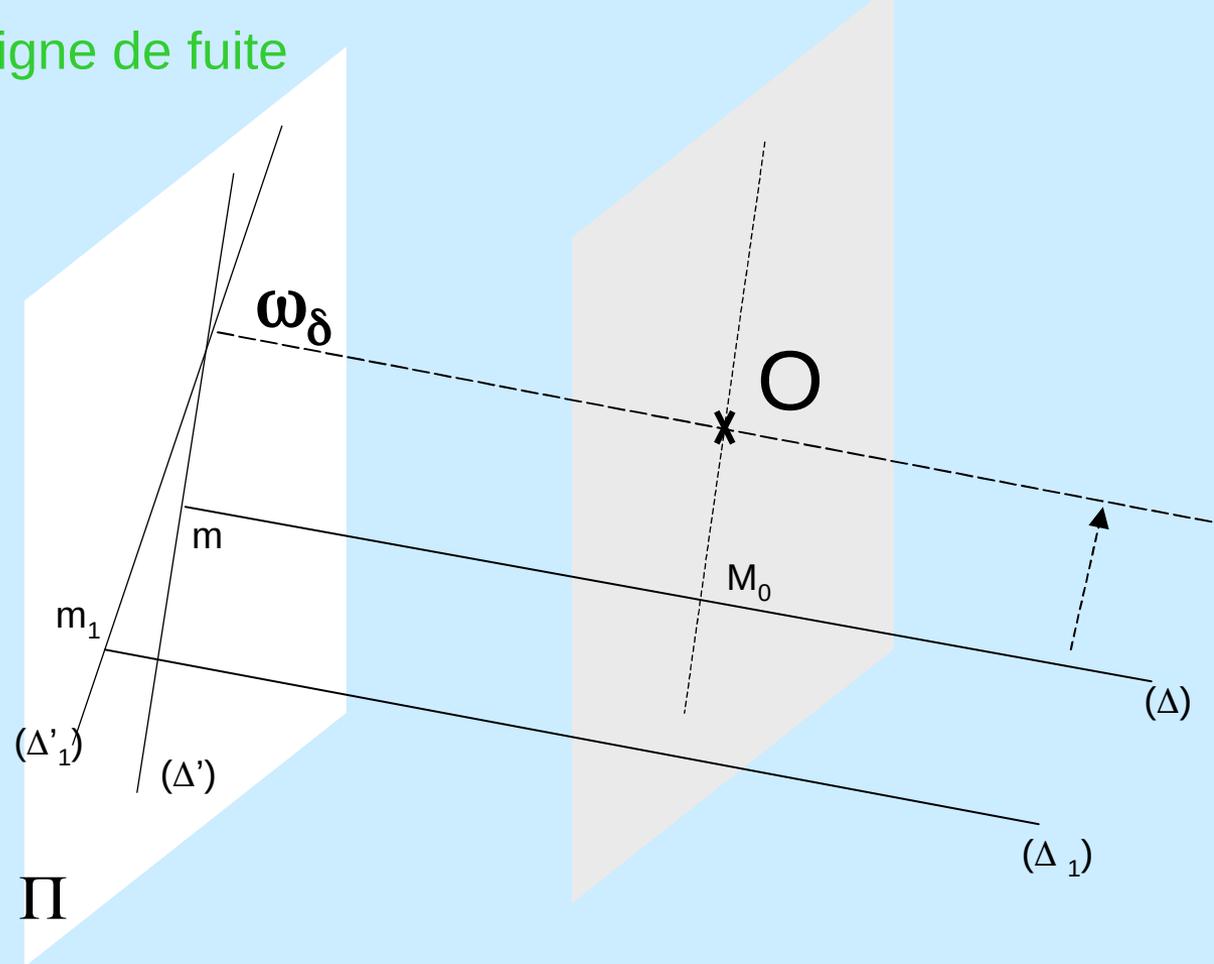


$$\frac{PQ}{PB'} \cdot \frac{AQ}{AB'} ?$$

**Voici la  
photo de la figure**



## Point de fuite – Ligne de fuite



### Théorème 1

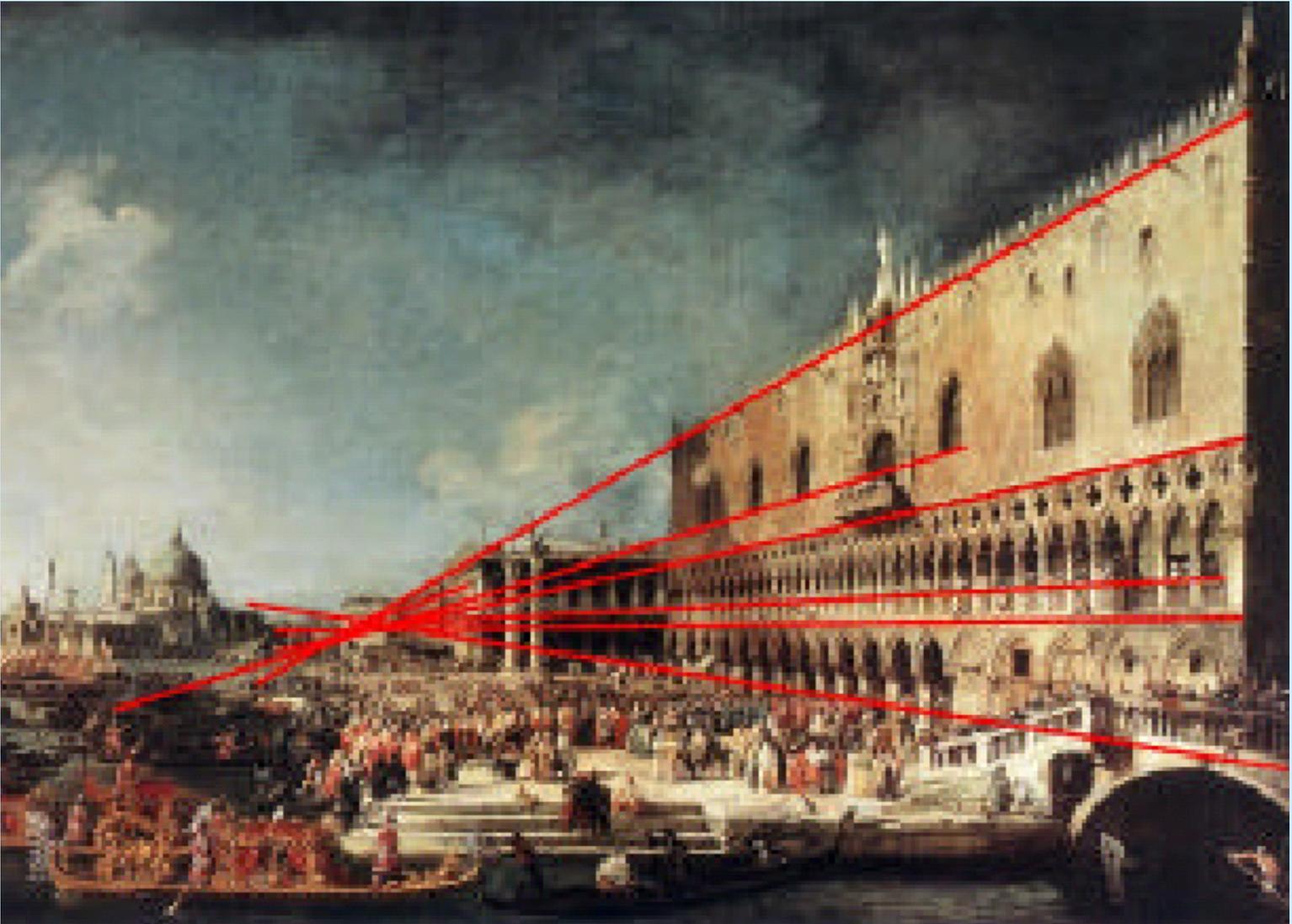
La projection conique de centre  $O$  sur le plan  $\Pi$  est une bijection de  $\Delta \setminus \{M_0\}$  sur  $\Delta' \setminus \{\omega_\delta\}$ , si  $\Delta$  n'est pas parallèle à  $\Pi$ .

### Théorème 2

Une famille de droite toutes parallèles à une même direction  $\delta$  a pour image une famille de droites dans  $\Pi$  qui concourent toutes en le même point  $\omega_\delta$  lorsque  $\delta$  n'est pas une direction de  $\Pi$ .



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<sup>e</sup> 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<sup>e</sup> siècle. Elle est aujourd'hui largement utilisée par les systèmes de visualisation sur ordinateur [GL](#) et [OpenGL](#) grâce notamment à O. Faugeras (INRIA).



## 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  $\Delta$  de  $E$  qu'on appellera son point à l'infini, noté  $\infty_{\Delta}$ .

On définit ce point à l'infini comme la direction  $\delta$  de  $\Delta$ .

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

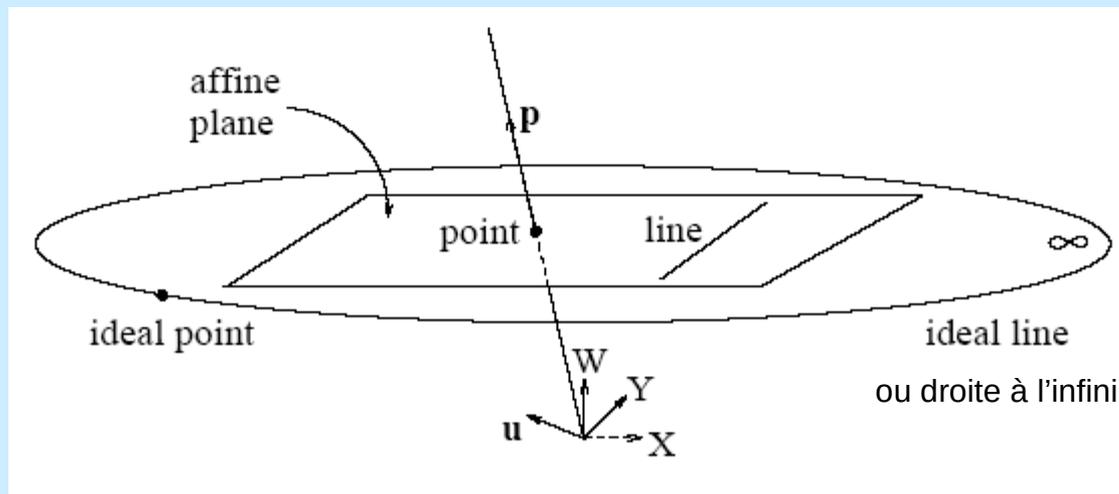
De même,  $\hat{P} = P \bullet \{\infty_P\}$  où  $\infty_P$  est l'ensemble des points à l'infini des droites de  $P$  et  $\infty_P$  s'appelle la droite à l'infini de  $P$ ,

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

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

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

$\{\infty_E\} \bullet \{\hat{P}\} =$  ensemble des plans projectifs de  $E$  <sup>86</sup>



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,
- de plus,

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

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

## Prolongement des projections

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

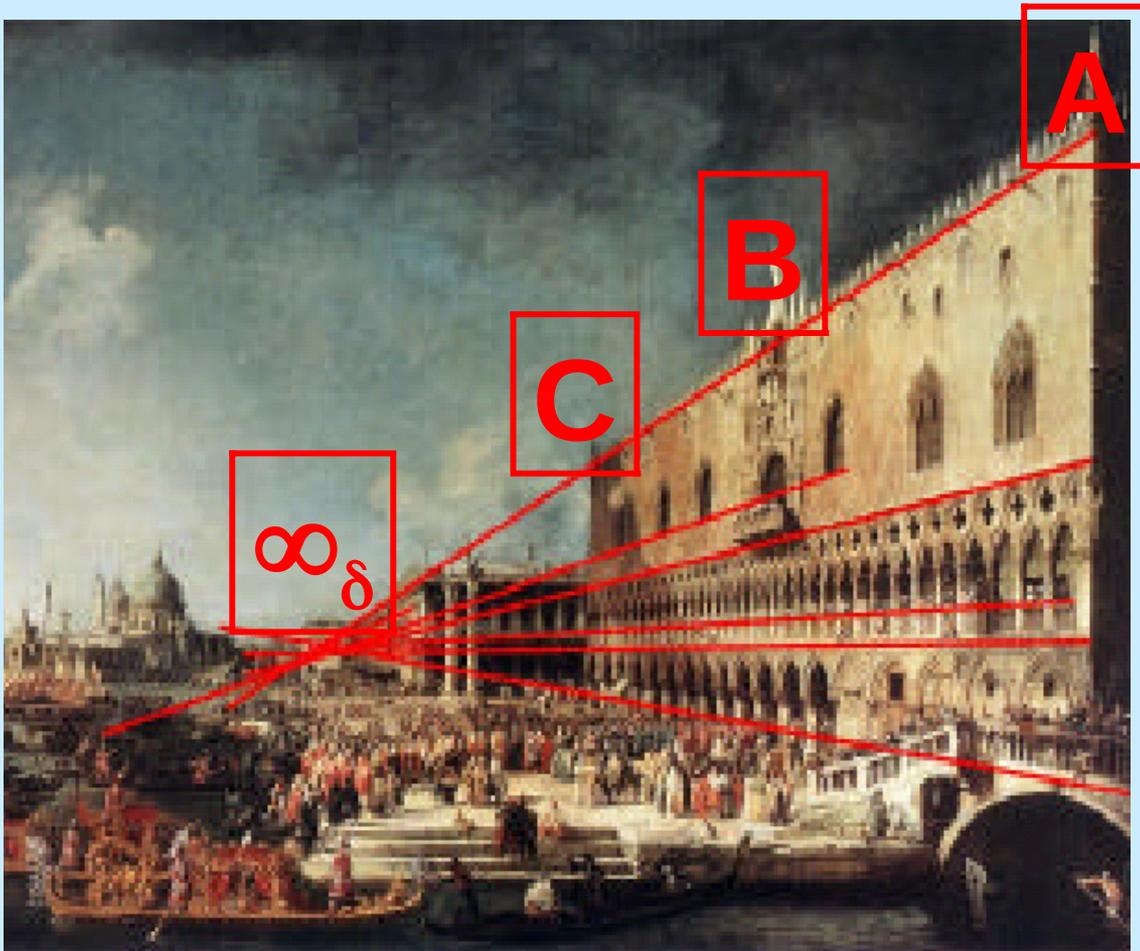
$$p: \hat{E} \setminus \{O\} \rightarrow \hat{\Pi}$$

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 :

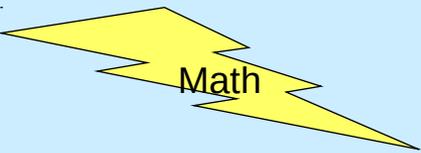
$$(A, B; C, \infty_{\Delta}) = \frac{AC}{BC} = (p(A), p(B), p(C), \omega_{\delta})$$



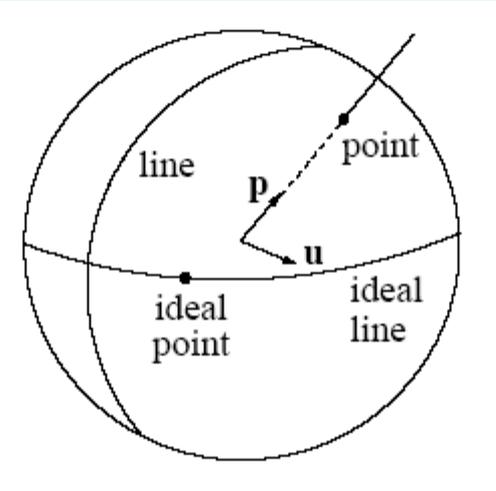
$$(A,B;C, \infty_\delta) = \frac{AC}{A \infty_\delta} : \frac{BC}{B \infty_\delta} = 2$$

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

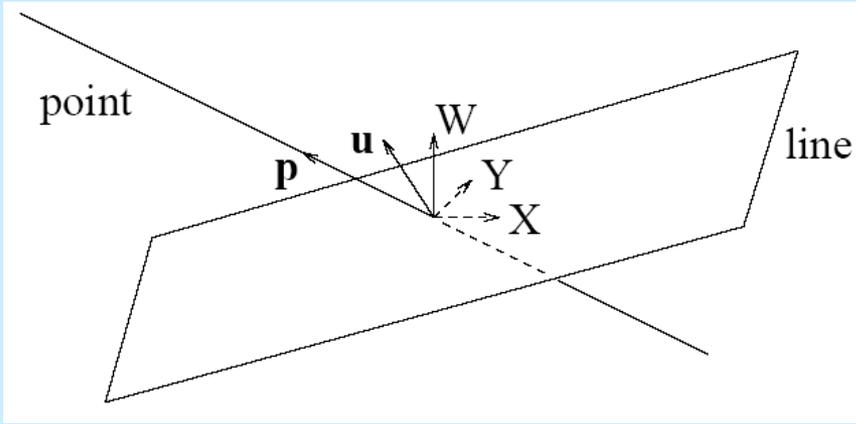
à  $\frac{A \infty_\delta}{B \infty_\delta}$  la valeur du même rapport en 3D



# Visualiser le plan projectif : ses points et ses droites



Sphère unité



Ray Space

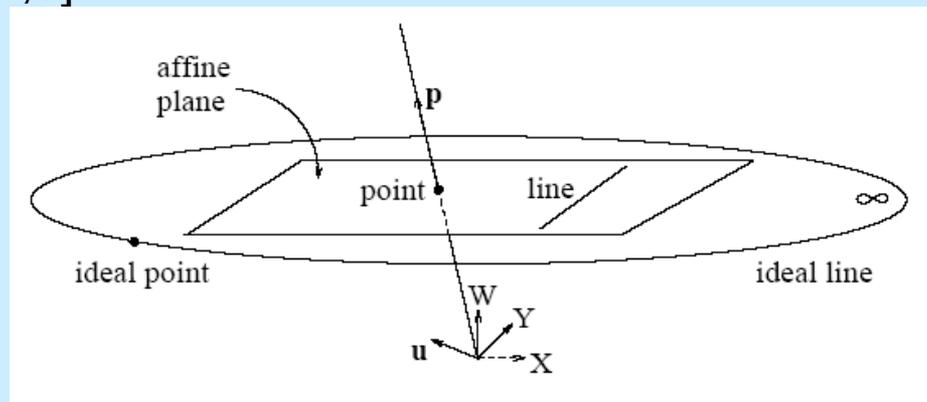
# La géométrie projective et le plan projectif

- 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  $[x, y]^T$

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

- 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  $\Delta_0$  est une droite de  $\Pi_0$  passant par  $O$  et  $\Delta$  est une droite affine de  $\Pi$  parallèle à  $\Delta_0$ , on pose  $p(\Delta_0) = \infty_{\Delta} = \infty_{\Delta_0}$ .

La projection conique  $p$  associe au point  $M$  de  $E \setminus \Pi_0$  de coordonnées  $(X, Y, Z)$  ( $Z \neq 0$ ) le point  $m$  de  $\hat{\Pi}$  de coordonnées  $(x = X/Z, y = Y/Z)$ .

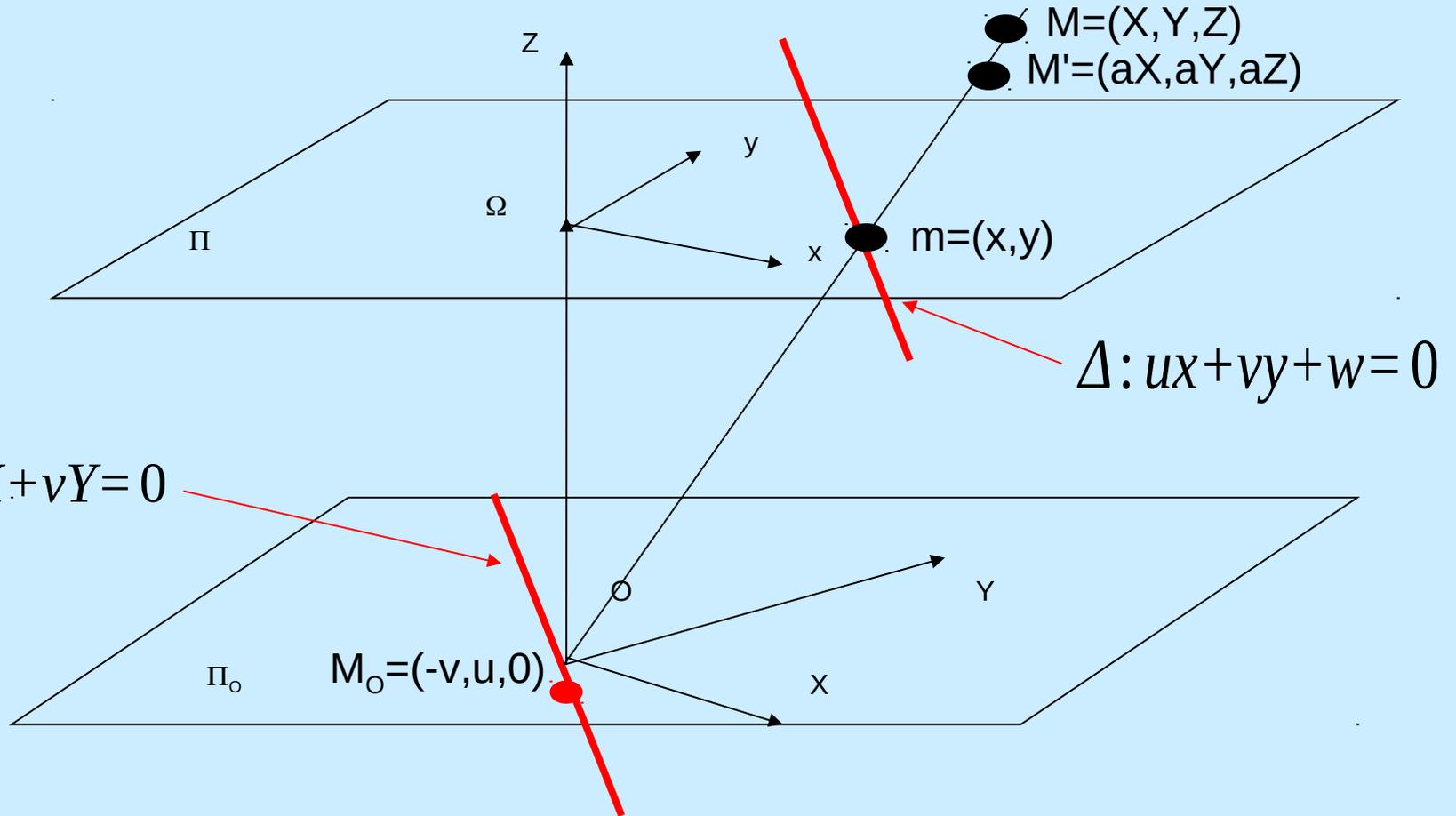
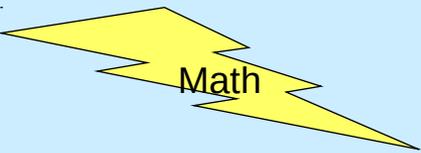
On dit que  $(X, Y, Z)$  est un système de coordonnées homogènes du point  $m$  et l'on note  $m = [X, Y, Z]$ .

Dans le plan affine, le repère affine  $(\Omega, i, j)$  donne :

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

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

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



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

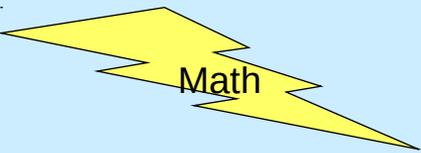
Or  $p(\Delta_o) = \infty_{\Delta}$  et si  $ux+vy+w=0$  est l'équation affine dans le repère  $(\Omega, i, j)$  de  $(\Delta)$ , l'équation affine de  $\Delta_o$  dans  $(O, i, j, k)$  est  $uX+vY=0$ , or tout point de  $\Delta_o$  définit un système de coordonnées homogènes de  $\infty_{\Delta}$ ,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

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

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  $\infty_{\Pi}$  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.



• Equation d'une droite :

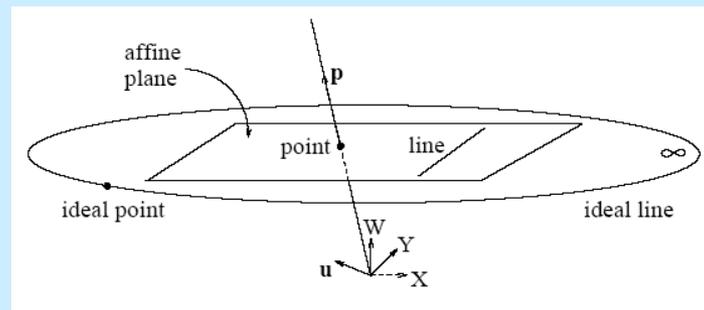
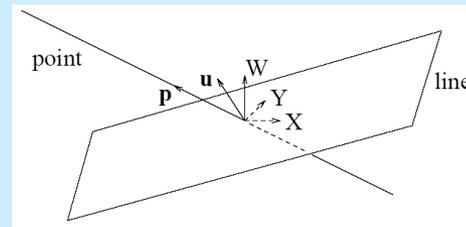
• Euclidienne (a,b,c):

$$ax+by+c=0$$

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

• 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 \ \mathbf{p}_2 \ \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 \ \mathbf{u}_2 \ \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 au points

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

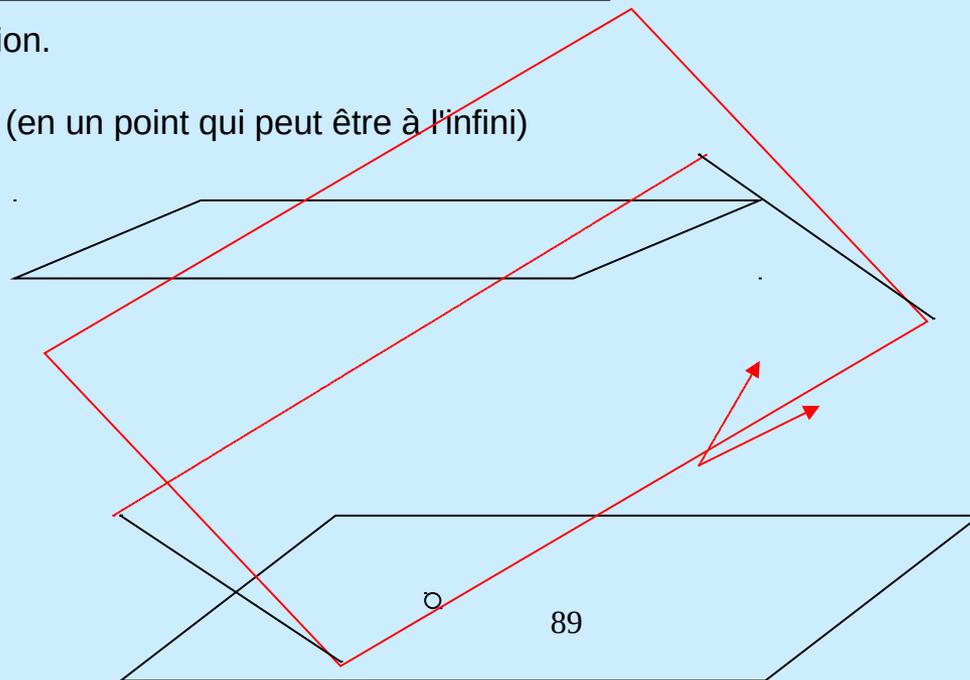
Tout théorème du plan projectif peut être reformulé en substituant les points aux droites et inversement

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)

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

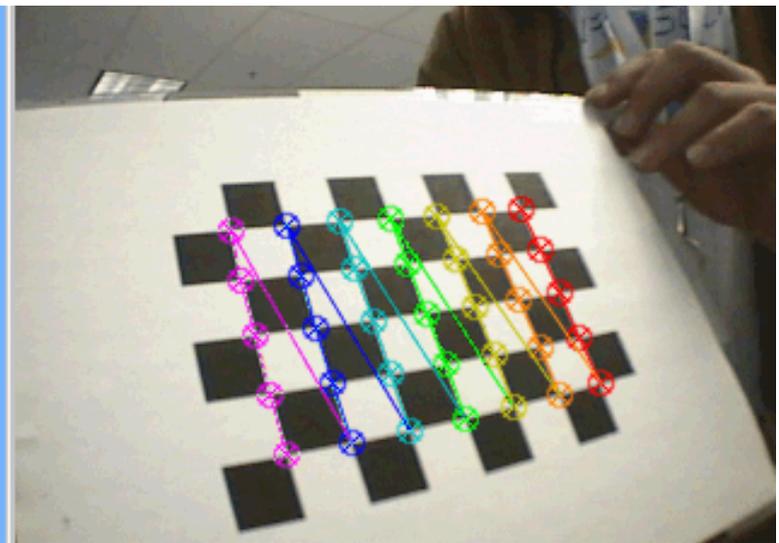
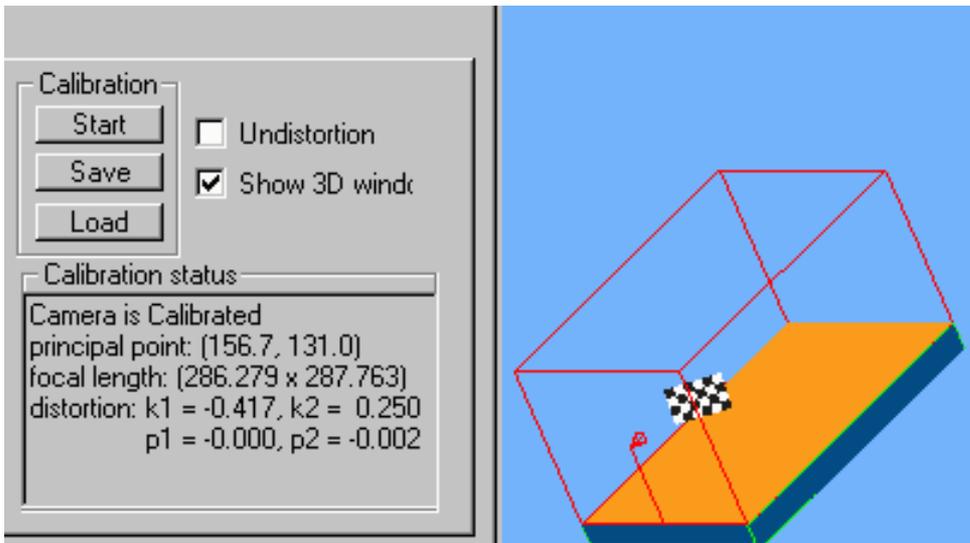


# Procédure PARAM\_CAL

**INPUT** : Une image du motif de calibration.

1. Mesurer les coordonnées 3D de chaque sommet des  $n$  carrés sur le motif de calibration dans le référentiel Monde. Soit  $N = 4n$
2. 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
3. Calculer les trois points de fuite  $p_1, p_2, p_3$  déterminés par les trois faisceaux de droites obtenus.
4. Calculer l'orthocentre  $O$  du triangle  $p_1p_2p_3$ .  $O$  est le centre de l'image.
5. établir les  $N$  correspondances entre les points image et monde, calculer  $SVD(A)$ . La solution est la colonne de  $V$  correspondant à la plus petite valeur singulière de  $A$ .
6. Déterminer  $\alpha$  and  $|\gamma|$
7. Déterminer les deux premières lignes de  $R$  et deux premières composantes de  $T$ .
8. 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.
9. 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  $T$ .
10. Réinitialise  $A$  et  $b$  pour estimer  $T_z$  et  $f_x$ .

**OUTPUT** :  $o_x, o_y, \alpha, f_x$  et les paramètres extrinsèques de la caméra.



<https://thesmartphonephotographer.com/smartphone-camera-focal-length>

« When we talk about crop-factors, what we really are talking about is the change in field of view by a smaller sensor. The reference is a full frame sensor, which is 36\*24mm. Therefore, a 10mm lens on an APS-C camera is still a 10mm lens - it just happens that it offers the same FOV as a 15mm lens would on a full frame camera.

### EXIF Format

The same of course is true for smartphones - unfortunately, the smartphone market is constantly and rapidly changing, and manufacturers usually do not state focal length or sensor size in their marketing texts, so research is needed on a per-phone-basis (or so I think).»

<https://photo.stackexchange.com/questions/38288/what-is-the-focal-length-on-your-typical-cell-phone-camera>

[https://en.wikipedia.org/wiki/Crop\\_factor](https://en.wikipedia.org/wiki/Crop_factor)

[https://fr.wikipedia.org/wiki/Format\\_35\\_mm](https://fr.wikipedia.org/wiki/Format_35_mm)

<https://www.bhphotovideo.com/explora/photography/tips-and-solutions/scanning-without-a-scanner-digitizing-your-film-with-a-dslr>

[https://fr.wikipedia.org/wiki/Appareil\\_photographique\\_reflex\\_num%C3%A9rique](https://fr.wikipedia.org/wiki/Appareil_photographique_reflex_num%C3%A9rique)

Global Positioning System	
GPS Altitude	31.9 m
GPS Latitude	6deg 14' 7.620"
GPS Longitude	106deg 49' 30.210"
Image Information	
Date and Time	2018:08:24 15:47:27
Manufacturer	Apple
Model	iPhone 6s
Photograph Information	
Aperture	F2.2
Exposure Bias	0 EV
Exposure Mode	Auto
Exposure Program	Auto
Exposure Time	1/874 s
Flash	No, auto
FNumber	F2.2
Focal Length	4.2 mm
ISO Speed Ratings	25
Metering Mode	Multi-segment
Shutter speed	1/874 s
White Balance	Auto

C'est un problème inverse intéressant :

étant donné une photographie prise par une caméra inconnue 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 :

- par estimation directe des paramètres

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

### 1. 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

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

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

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

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),

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

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

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

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



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



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

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

Avec

$$\mathbf{A} = \begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -x_1 X_1 & -x_1 Y_1 & -x_1 Z_1 & -x_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -y_1 X_1 & -y_1 Y_1 & -y_1 Z_1 & -y_1 \end{bmatrix}$$

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

# Etape 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.

$$\hat{m} = \alpha \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{21} & m_{22} & m_{23} & m_{24} & m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix}^T$$

# Etape 2 : Calculer les paramètres de caméra

$$\hat{M} = \begin{bmatrix} q_1 & q_{41} \\ q_2 & q_{42} \\ q_3 & q_{43} \end{bmatrix}$$
$$M = \begin{bmatrix} -f_x r_{11} + o_x r_{31} & -f_x r_{12} + o_x r_{32} & -f_x r_{13} + o_x r_{33} & -f_x T_x + o_x T_z \\ -f_y r_{21} + o_y r_{31} & -f_y r_{22} + o_y r_{32} & -f_y r_{23} + o_y r_{33} & -f_y T_y + o_y T_z \\ r_{31} & r_{32} & r_{33} & T_z \end{bmatrix}$$

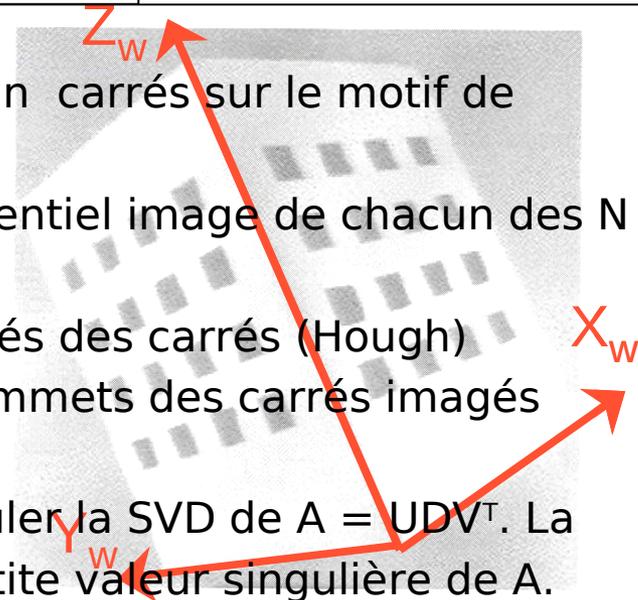
$$\hat{M} = \gamma M$$

# Procédure PROJ\_MAT\_CAL

**INPUT** : Une image du motif de calibration.

1. Mesurer les coordonnées 3D de chaque sommet des  $n$  carrés sur le motif de calibration dans le référentiel Monde. Soit  $N = 4n$
2. 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
- Etant donné  $N$  correspondances monde-image, calculer la SVD de  $A = UDV^T$ . La solution  $\hat{m} = \alpha m$  est la colonne de  $V$  corresp. à la plus petite valeur singulière de  $A$ .
- Trouver  $|\gamma|$  en utilisant  $\sqrt{\hat{m}_{31}^2 + \hat{m}_{32}^2 + \hat{m}_{33}^2} = |\gamma| \sqrt{r_{31}^2 + r_{32}^2 + r_{33}^2} = |\gamma|$
- $\hat{M} \leftarrow 1/|\gamma| \hat{M}$  ainsi  $M$  est connu à un signe  $\sigma$  près
- Calculer  $T_z$  par  $T_z = \sigma$
- Calculer  $R_3^T$  par  $r_{3i} = \sigma$
- $o_x = q_1^T q_3$  et  $o_y = q_2^T q_3$
- On peut retrouver  $f_x$  et  $f_y$  par  $f_x = \sqrt{q_1^T q_1 - o_x^2}$  et  $f_y = \sqrt{q_2^T q_2 - o_y^2}$
- Trouver le signe  $\sigma$  avec  $T_z > 0$  si  $O^w$  est devant  $O^c$  et orthogonaliser l'estimée de  $R$  trouvée par la technique SVD



**OUTPUT** :  $o_x, o_y, \alpha, f_x$  et les paramètres extrinsèques de la caméra.<sup>97</sup>

De façon similaire, on résoud d'abord un système linéaire puis on procède à une décomposition en paramètres physiques. Les résultats devraient être identiques. Mais :

■ 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
- Des méthodes spécifiques pour l’estimation de pose online (une fois la caméra calibrée intrinsèquement offline):

```
@article{DAV04,  
  author = "P. David, D. DeMenthon,R. Duraiswami and  
  H. Samet",  
  title = "SoftPOSIT : Simultaneous Pose and  
  Correspondence Determination",  
  journal = "Intl. Journal of Computer Vision",  
  ijournal = "Intl. Journal of Computer Vision",  
  volume = "59",  
  year = "2004",  
  number = "3",  
  pages = "259-284"}
```

```
@article{QUA99,  
  author = "L. Quan and Z. Lan",  
  title = "Linear N-Point Camera Pose  
  Determination",  
  journal = "IEEE Transactions on Pattern Analysis  
  and Machine Intelligence",  
  ijournal = "IEEE Transactions on Pattern Analysis  
  and Machine Intelligence",  
  volume = "21",  
  year = "1999",  
  number = "7",  
  pages = "774-780"}
```