

# Calibration de caméra et Vision Stéréo

Fabien Baldacci

Enseirb-Matmeca, IPB

# Ce cours est très largement inspiré de

Informatique visuelle / vision par ordinateur

Un peu de géométrie

Elise Arnaud

[elise.arnaud@imag.fr](mailto:elise.arnaud@imag.fr)

cours inspiré par A. Boucher, E. Boyer, E. Marchand



# Plan

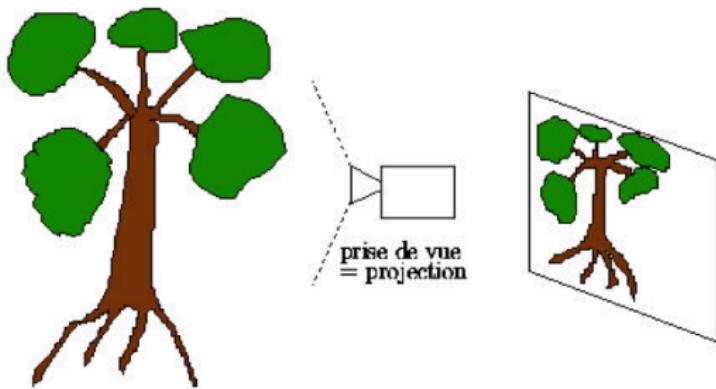
Projection perspective

Calibration de caméra

Vision stéréoscopique

Disparité

- Le monde réel : 3D, une image : 2D



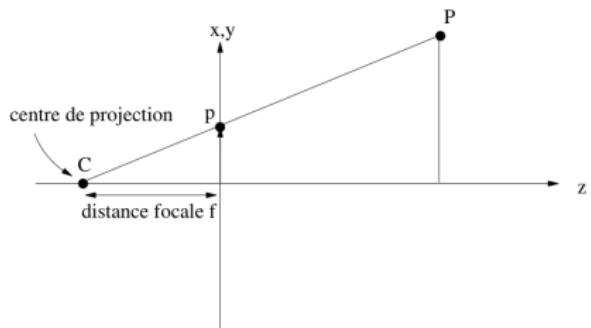
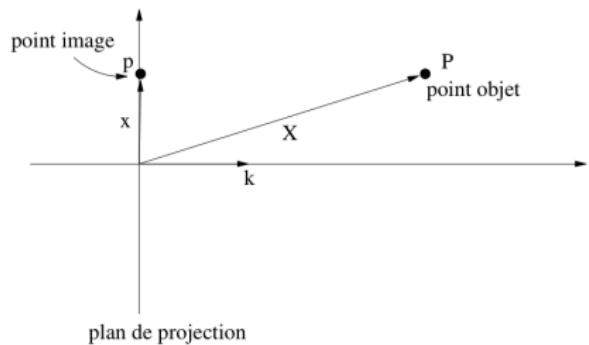
scène 3D

scène 2D

source : E. Marchand, Irisa

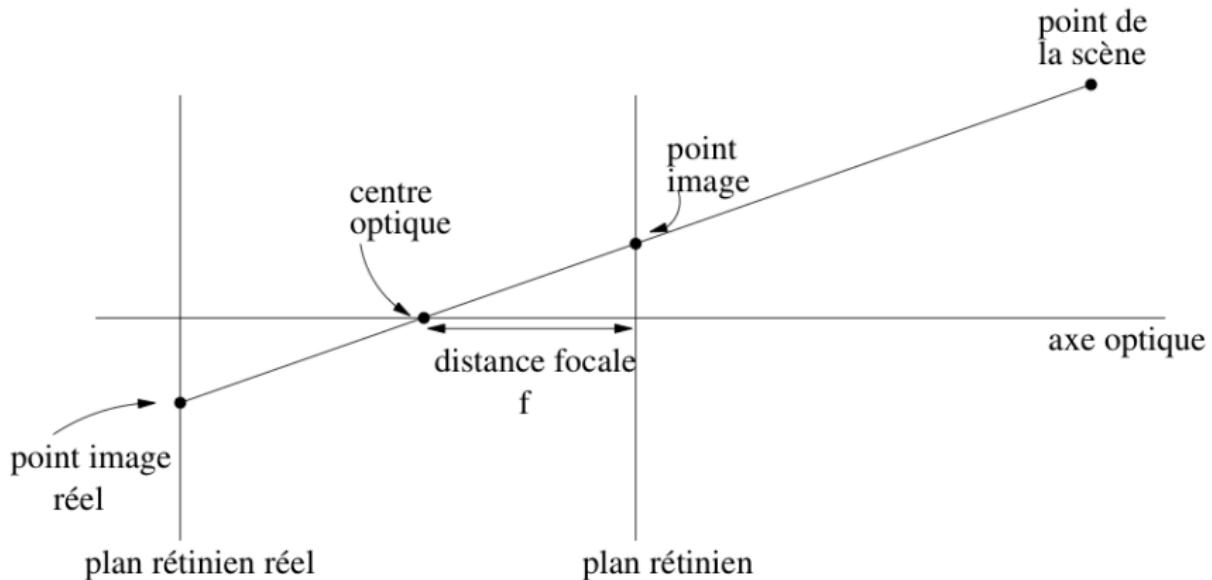
# Notions de projection

- ▶ projection orthographique
- ▶ projection perspective



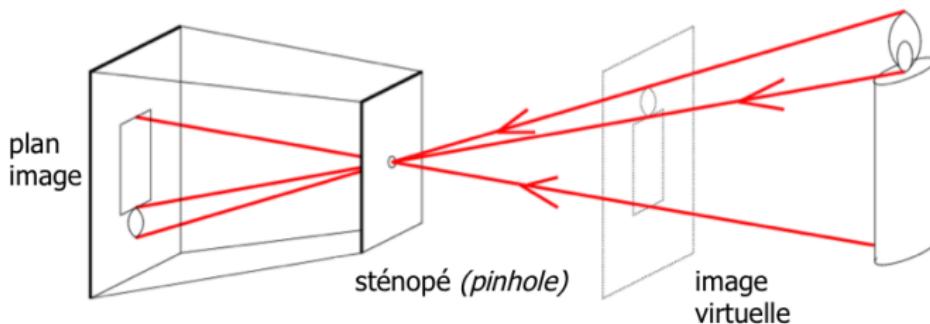
source : E. Boyer, Inria

# Projection perspective - modèle de sténopé



source : E. Boyer, Inria

# Projection perspective - modèle de sténopé

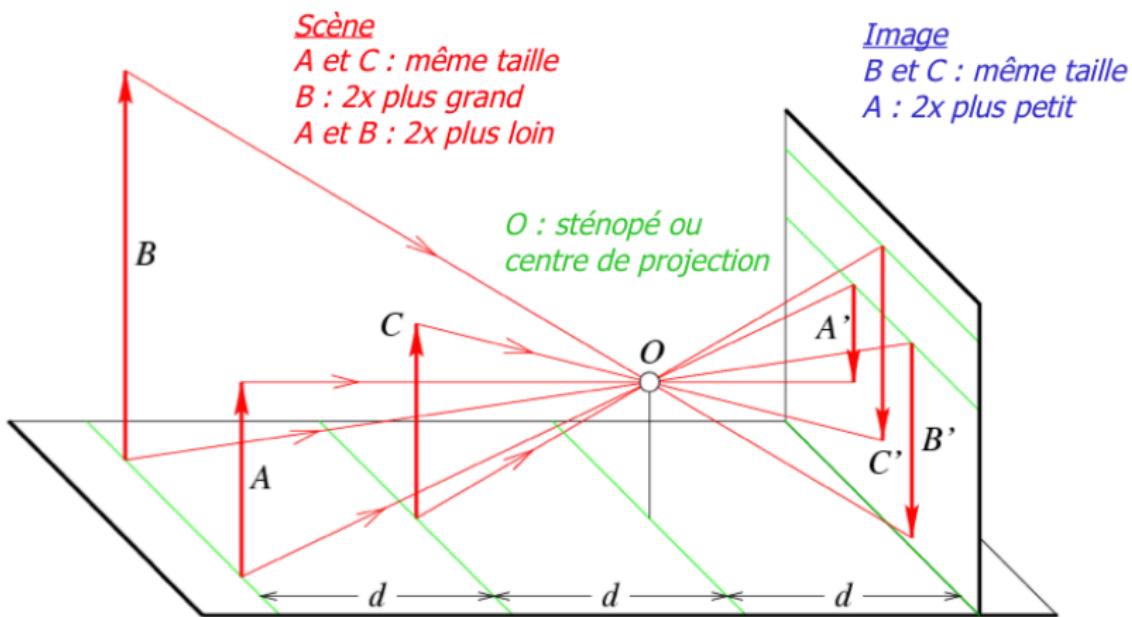


source : A. Boucher, IFI

- ▶ Modèle sans lentille qui date du 15ème siècle
- ▶ Les objets apparaissent inversés sur le plan image

# Projection perspective - modèle de sténopé

- ▶ effets de la perspective



3

source : A.

Boucher, IFI

# Projection perspective - modèle de sténopé

$P_c = (x_c, y_c, z_c)$  point du monde 3D (dans le repère caméra)

$p = (x, y)$  projection dans le plan image

- ▶ projection perspective

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{x_c \cdot f}{z_c} \\ \frac{y_c \cdot f}{z_c} \end{bmatrix}$$

- ▶ en coordonnées homogènes

$$\begin{bmatrix} x.w \\ y.w \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/f \end{bmatrix} \cdot \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$

# Projection perspective - modèle de sténopé

$p = (x, y)$  projection dans le plan image

$P_i = (x_i, y_i)$  coordonnées des pixels

- ▶ prise en compte de l'échantillonnage des pixels, d'une éventuelle distortion radiale, etc.

- ▶ ici,  $(u_o, v_o)$  : centre de l'image en pixels
- ▶  $(k_x, k_y)$  : facteurs d'échelles en pixels / mm

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} k_x \cdot x \\ k_y \cdot y \end{bmatrix} + \begin{bmatrix} u_o \\ v_o \end{bmatrix}$$

- ▶ en coordonnées homogènes

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_o \\ 0 & k_y & v_o \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

# Plan

Projection perspective

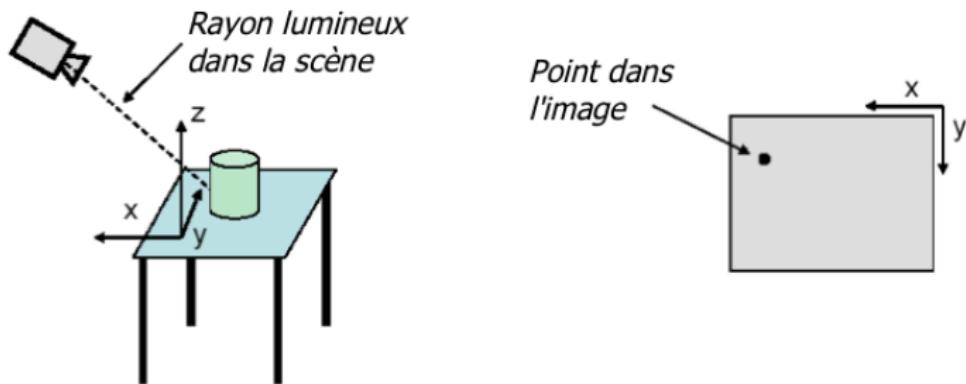
Calibration de caméra

Vision stéréoscopique

Disparité

# Calibration de la caméra

consiste à déterminer les paramètres entre les coordonnées du monde et les coordonnées image



source : A. Boucher, IFI

# Calibration de la caméra

- ▶ Transformation du système de coordonnées de la scène vers celui de l'image :

scène → caméra → image

## paramètres extrinsèques

- ▶ relation entre le référentiel environnement et le référentiel caméra
- ▶ position et orientation de la caméra par rapport à la scène

## paramètres intrinsèques

- ▶ propriétés optiques et géométriques de la caméra
- ▶ ne changent pas si on déplace la caméra

# Systèmes de coordonnées

Trois systèmes de coordonnées distincts

la scène  $R_e$

- ▶ système de référence connu. Tout (caméra, objet, etc. ) est défini par rapport à ce système de coordonnées

la caméra  $R_c$

- ▶ système dans lequel s'effectue la projection des points 3D vers 2D

l'image  $R_i$

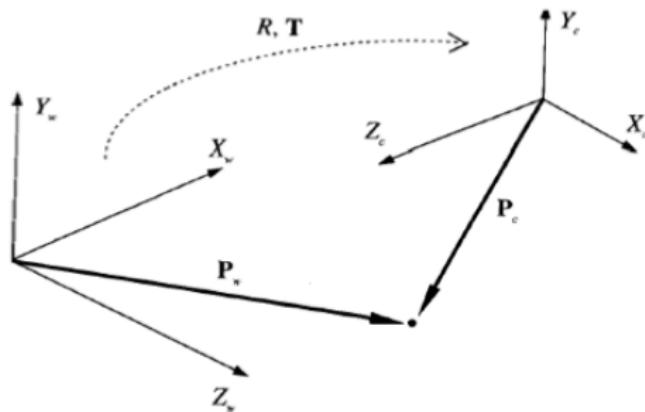
- ▶ système établissant la référence aux pixels

$$P_e(x_e, y_e, z_e) \rightarrow P_c(x_c, y_c, z_c) \rightarrow P_i(x_i, y_i)$$

# Paramètres extrinsèques

La position et l'orientation composent une transformation rigide 3D

- ▶ La translation  $T$  pour faire correspondre les origines des systèmes de coordonnées
- ▶ la rotation  $R$  pour faire correspondre les axes des systèmes



source : A. Boucher, IFI

# Paramètres extrinsèques

Transformation rigide décrite par les paramètres extrinsèques

$$P_c = R P_e + T$$

- écriture matricielle

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} R \end{bmatrix} \cdot \begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix} + \begin{bmatrix} T \end{bmatrix}$$

- coordonnées homogènes

$$\begin{bmatrix} x_c \cdot w \\ y_c \cdot w \\ z_c \cdot w \\ w \end{bmatrix} = \begin{bmatrix} R & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_e \\ y_e \\ z_e \\ 1 \end{bmatrix}$$

# Du monde à l'image



$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_o \\ 0 & k_y & v_o \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$



$$\begin{bmatrix} x_{i.w} \\ y_{i.w} \\ w \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_o \\ 0 & k_y & v_o \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x.w \\ y.w \\ w \end{bmatrix}$$



$$\begin{bmatrix} x_{i.w} \\ y_{i.w} \\ w \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_o \\ 0 & k_y & v_o \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/f \end{bmatrix} \cdot \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$



$$\begin{bmatrix} x_{i.w} \\ y_{i.w} \\ w \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_o \\ 0 & k_y & v_o \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \cdot \begin{bmatrix} R & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_e \\ y_e \\ z_e \\ 1 \end{bmatrix}$$

# Du monde à l'image

$$\begin{bmatrix} x_i \cdot w \\ y_i \cdot w \\ w \end{bmatrix} = \begin{bmatrix} k_x & 0 & u_o \\ 0 & k_y & v_o \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \cdot \begin{bmatrix} R & & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_e \\ y_e \\ z_e \\ 1 \end{bmatrix}$$

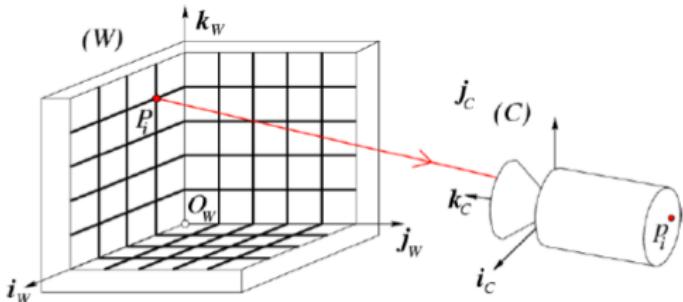
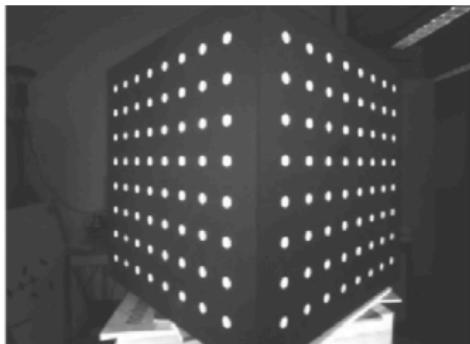


$$P_i = M \cdot P_e = K \cdot [R | T] \cdot P_e$$

- ▶  $M$  : matrice  $3 \times 4$ , contenant les paramètres intrinsèques et extrinsèques
- ▶ tous les paramètres sont définis à un facteur d'échelle près (on ne connaît pas la profondeur des points)
- ▶ 11 inconnues

# Estimation de M

- Utilisation d'une mire pour calibrer les caméras
- La mire (ou cible) doit contenir des formes faciles à reconnaître et dont la position est très bien connue selon le référentiel environnement
- On mesure le centre des points blancs qui sont connus à l'avance et facile à identifier
- La cible doit comporter des éléments non-coplanaires



Sources : [cis.poly.edu/cs664](http://cis.poly.edu/cs664) + Justus Piater, Computer vision, Université de Liège (Belgique)

source : A. Boucher, IFI

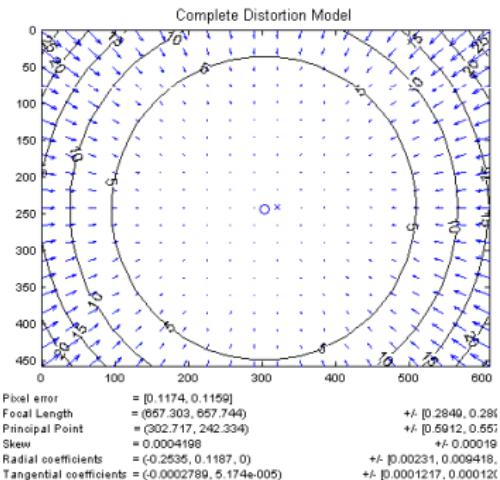
# Distortion

Image rectifiée  $I$ Image distordue  $I_d$ 

**Modèle générique de distorsion:**

$$\mathbf{m}_d = \text{distord}(\mathbf{m})$$

avec *distord* fonction non linéaire  
représentant le champs de distorsion

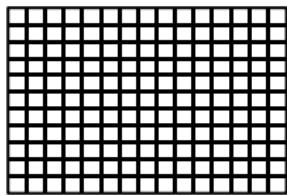


Exemple de champ de distorsion estimé  
avec la Matlab Calibration Toolbox

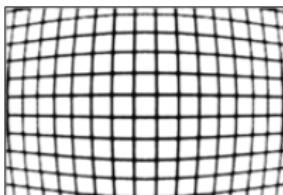
source : A. Boucher, IFI

# Distortion

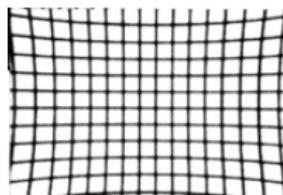
- Effet principal d'une optique non idéale
  - Notamment sur optiques grand-angle



Vue rectifiée d'une grille



Distorsion en barillet



Distorsion en croissant

source : A. Boucher, IFI

# Modèle de distortion radiale

Position rectifiée

$$\mathbf{m} \propto \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\mathbf{m}_d = \text{distord}(\mathbf{m}; \mathbf{k})$$

$$\mathbf{m} = \text{undistord}(\mathbf{m}_d; \mathbf{k})$$

Position distordue

$$\mathbf{m}_d \propto \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix}$$

## Modèle polynomial de distortion radiale

$$\mathbf{r}_d = (1 + k_1 r^2 + k_2 r^4 + \dots) \mathbf{r} \quad \text{en notant} \quad \mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix} \quad r = \|\mathbf{r}\|$$

$$\mathbf{r}_d = \begin{bmatrix} x_d \\ y_d \end{bmatrix}$$

Les paramètres de distorsion  $\mathbf{k}=(k_1, k_2, \dots)$  sont estimés lors de la calibration intrinsèque de la caméra (voir Matlab Calibration Toolbox)

source : A. Boucher, IFI

# Correction de la distortion

Estimé en même temps que les paramètres intrinsèques et extrinsèques ; avec une mire



Compensation de la distorsion

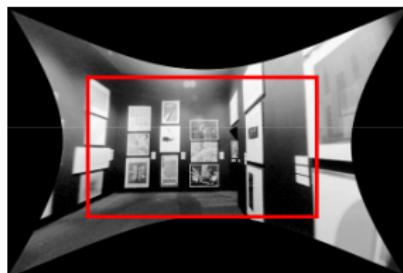


Image capturée par une caméra  
avec objectif très grand angle  
(présence de distorsions : les droites  
3D n'apparaissent pas rectilignes)



Image rectifiée  
(les droites 3D apparaissent rectilignes)

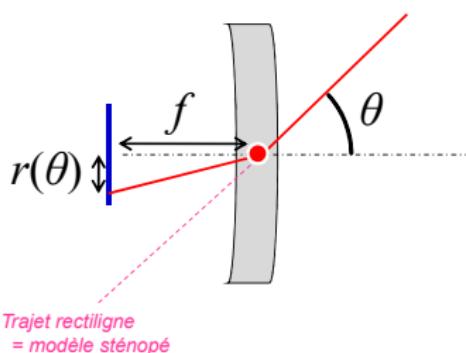


Image rectifiée et recadrée

# Objectif très grand angle : “fish-eye”



Angle de vue approchant les 180°



<http://cmp.felk.cvut.cz/demos/Omni/fisheye/>

Modèle	Fonction $r(\theta)$
Sténopé	$r = f \tan(\theta)$
Angles équisolides (fish-eye le plus courant)	$r = 2f \sin(\theta/2)$
Equidistant	$r = f\theta$
Projection orthogonale	$r = f \sin(\theta)$

# Plan

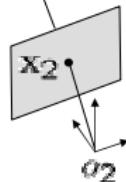
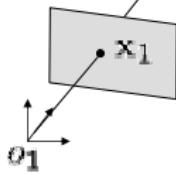
Projection perspective

Calibration de caméra

Vision stéréoscopique

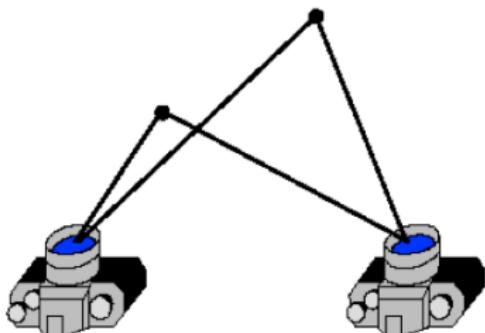
Disparité

## General Formulation

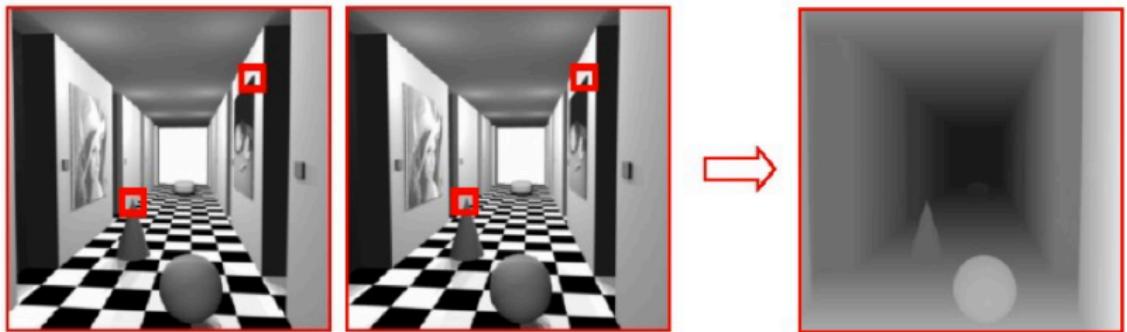


Given two views of the scene  
recover the unknown camera  
displacement and 3D scene  
structure

# Vision stéréoscopique



Nous allons évaluer la position d'un point dans l'espace à partir de plusieurs caméras

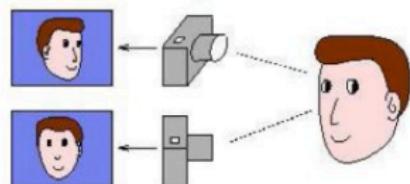


Carte des profondeurs

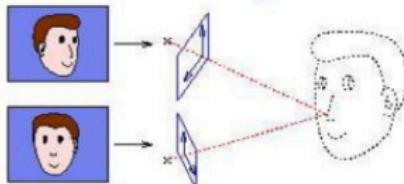
source : cis.poly.edu/cs64

# Problème de la vision stéréoscopique

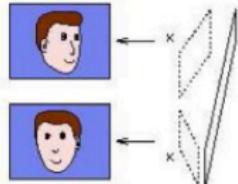
1. acquisition



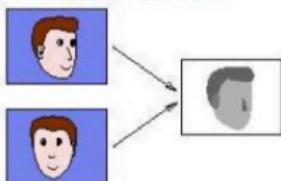
2. calibrage



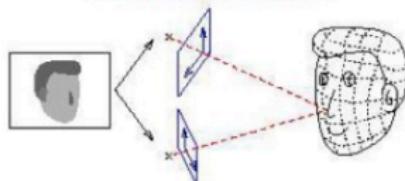
3. rectification



4. corrélation

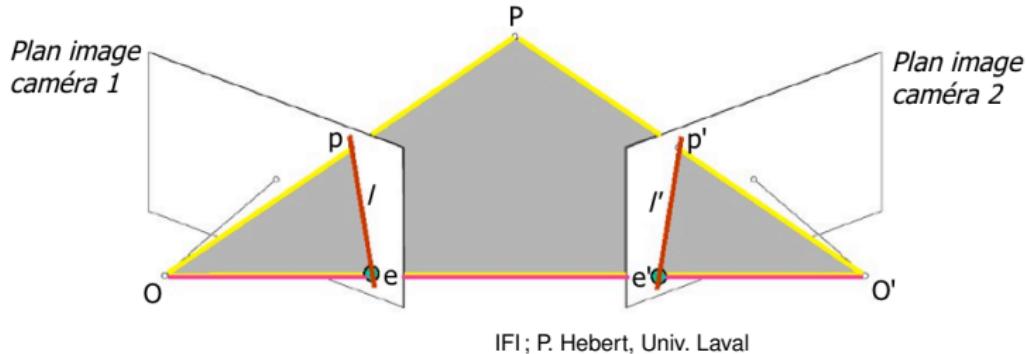


5. reconstruction...



source : F. Devernay, Inria

# Géométrie épipolaire

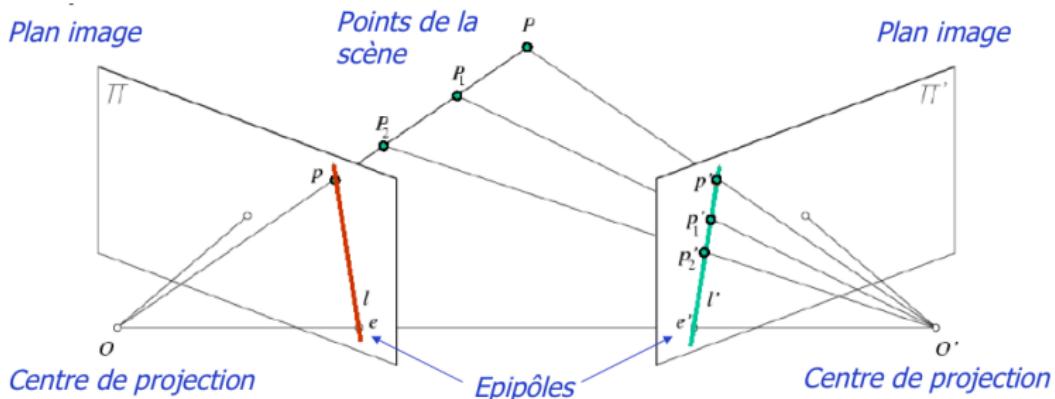


source : A. Boucher,

- ▶ 5 points dans le même plan :  $P, p, p', O, O'$
- ▶ les droites  $l$  et  $l'$  sont les droites épipolaires
- ▶ les points  $e$  et  $e'$  sont respectivement les épipôles gauche et droit
- ▶  $e$  est en fait  $O'$  vu dans l'image gauche
- ▶  $e'$  est  $O$  vu dans l'image droite

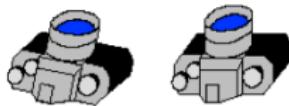
# Géométrie épipolaire

- Un point dans l'image gauche se situe sur la droite épipolaire correspondante dans l'image droite



source : A. Boucher, IFI

# Exemples de droites épipolaires



source : A. boucher. IFI

# Vision stéréoscopique

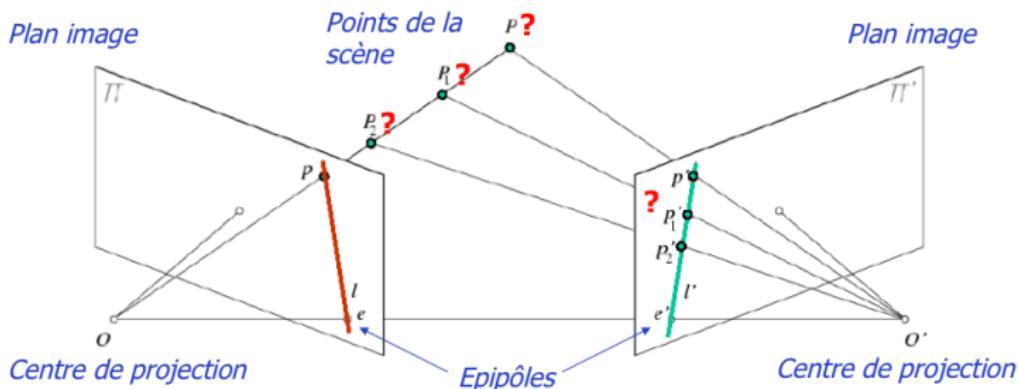
- ▶ Première étape : calibration du système à deux caméras :
  - ▶ Calibrer chaque caméra
  - ▶ Connaitre la transformations qui permet de passer d'une caméra vers l'autre caméra
  - ▶ Calcul des droites épipolaires
- ▶ Deuxième étape (éventuelle) : rectification

# Vision stéréoscopique



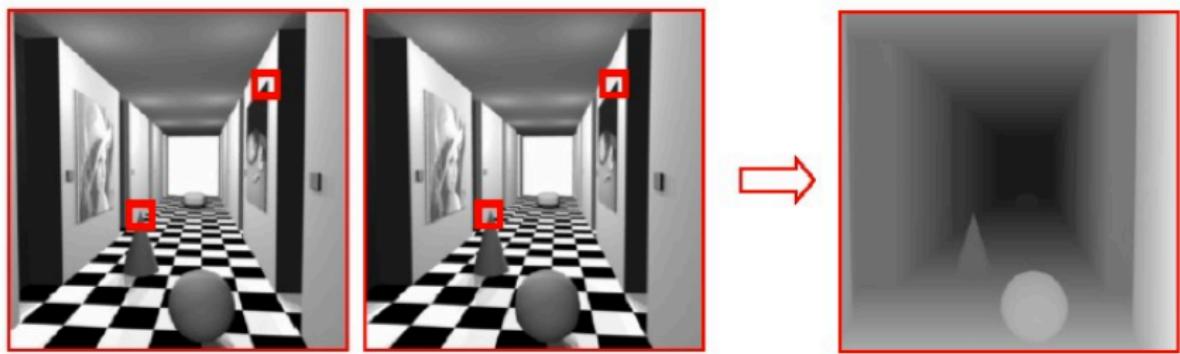
# Vision stéréoscopique

- ▶ Troisième étape : mise en correspondance (appariement)
- ▶ A partir d'un point dans l'image gauche, et de la géométrie épipolaire, quel est le bon point correspondant dans l'image droite ?



# Vision stéréoscopique

- ▶ Si on retrouve l'ensemble des correspondances, on peut calculer une carte de profondeur de la scène
- ▶ en fait, il s'agit d'une carte de disparités entre les correspondances (distances ( $p_1, p_2$ ))

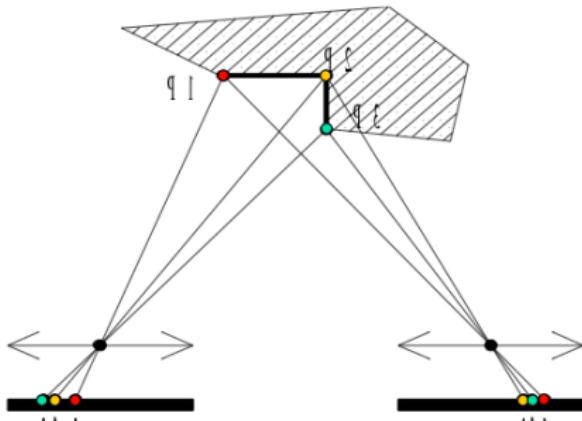


*Carte des profondeurs*

# Vision stéréoscopique

Difficultés de la mise en correspondance

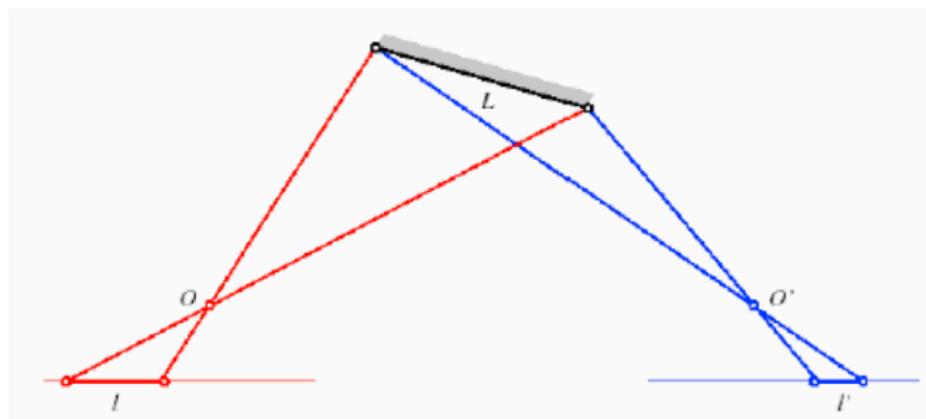
*Les points ne sont  
pas forcément dans  
le même ordre dans  
les deux images*



# Vision stéréoscopique

Difficultés de la mise en correspondance

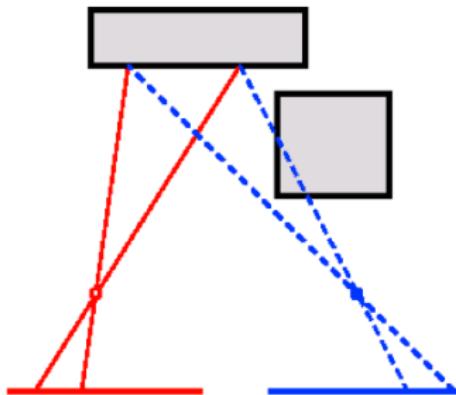
- ▶ Les tailles et les distances ne sont pas les mêmes d'une image à l'autre



# Vision stéréoscopique

## Difficultés de la mise en correspondance

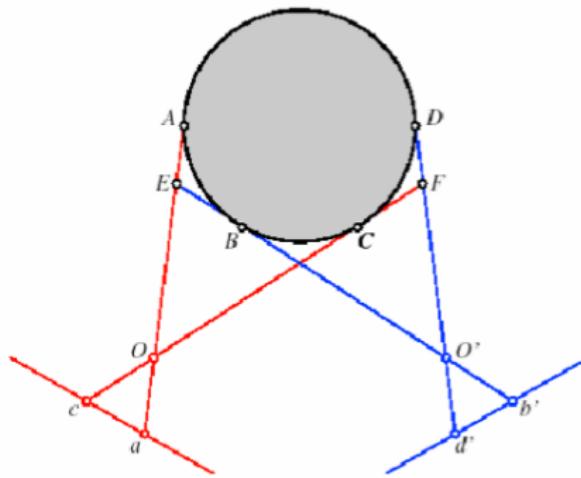
- ▶ Occultations : des objets ou parties d'objets peuvent être cachés.  
La correspondance n'existe pas dans ce cas.



# Vision stéréoscopique

## Difficultés de la mise en correspondance

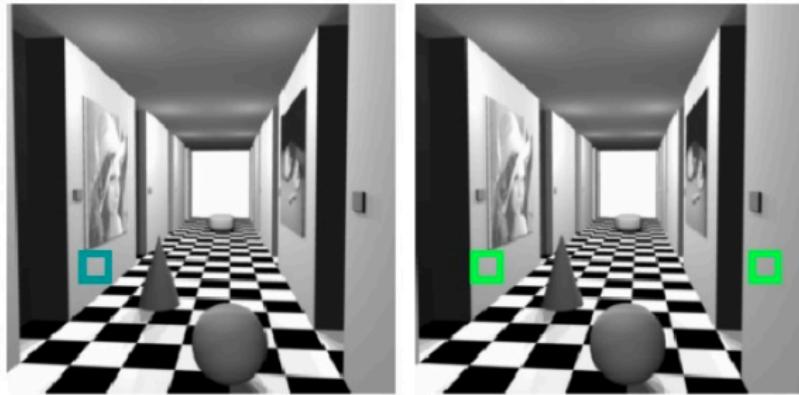
- Dans le cas des objets sphériques ou curvilignes, les arêtes perçues ne correspondent pas d'une image à l'autre



# Vision stéréoscopique

## Difficultés de la mise en correspondance

- ▶ Correspondance difficile (voire impossible) dans les zones faiblement texturées.



# Plan

Projection perspective

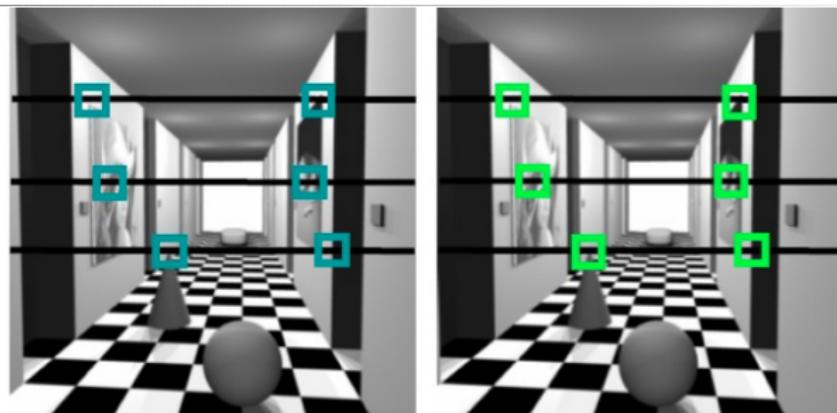
Calibration de caméra

Vision stéréoscopique

Disparité

# Vision stéréoscopique - mise en correspondance

- ▶ Si les images sont rectifiées, alors la recherche des correspondances est beaucoup plus facile
- ▶ Il suffit de rechercher les zones similaires sur la même ligne
- ▶ la distance ( $x_2 - x_1$ ) entre les deux correspondances est la disparité : indication de la profondeur du point dans la scène



# Vision stéréoscopique - mise en correspondance

## Hypothèses

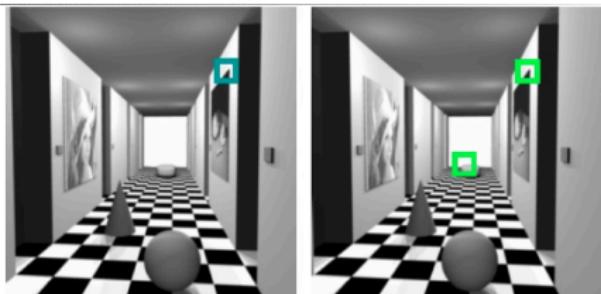
- ▶ Les points de la scène sont visibles, en général, des deux points de vues
- ▶ les points correspondants sont similaires dans les images

## Approche basée sur la corrélation

- ▶ On sélectionne un voisinage dans l'image gauche
- ▶ On cherche le voisinage correspondant dans l'image de droite
  - ▶ Si on connaît la géométrie épipolaire alors recherche le long de la droite épipolaire sinon recherche dans toute l'image
  - ▶ Concentrer d'abord la recherche dans les environs du point observé

# Vision stéréoscopique - mise en correspondance

Approche basée sur la corrélation



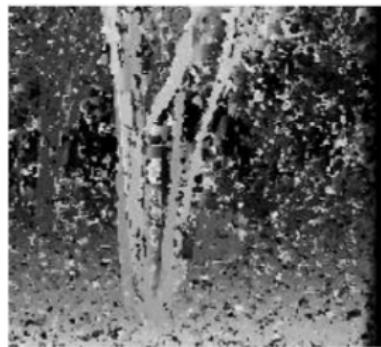
$$\sum ( \text{[Red Box Image]} - \text{[Green Box Image]} )^2 = ssd$$

$$\sum ( \text{[Red Box Image]} - \text{[Green Box Image]} )^2 = ssd$$

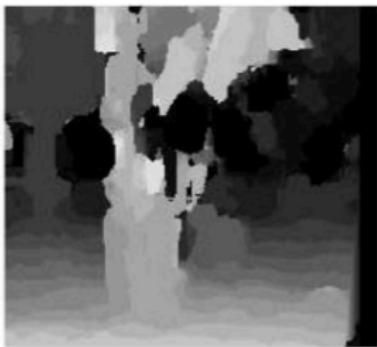
# Vision stéréoscopique - mise en correspondance

Approche basée sur la corrélation

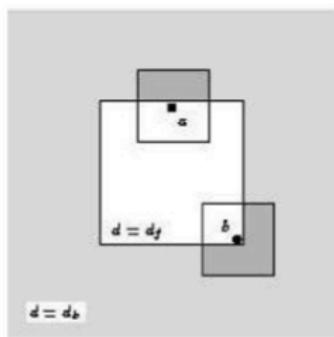
- ▶ importance de la taille de la fenêtre de corrélation



$w = 3$



$w = 20$



# Vision stéréoscopique - mise en correspondance

Approche basée sur la corrélation

image  $I(x,y)$



Carte des disparités  $D(x,y)$

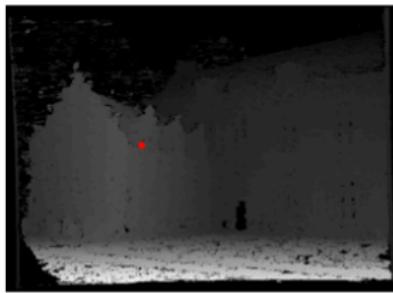


image  $I'(x',y')$



$$(x',y') = (x + D(x,y), y)$$