

Computer Vision

Raphaël Viards – Magellium – Imagerie et Applications

raphael.viards@magellium.fr

- I. Introduction
- II. Systèmes d'acquisition 3D
- III. Modélisation de la camera
- IV. Stéréovision

Partie I : Perception 3D

I. Introduction

Domaines d'activité



Observation de la terre



Géo-information



Imagerie & applications

2003
création

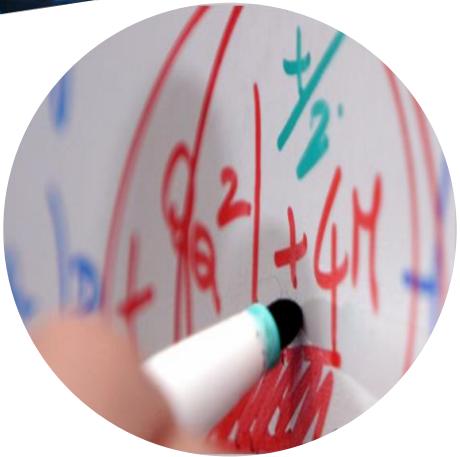
150+
employés

2 sites
Paris
Toulouse*

15* M€
CA 2017

Intégrée au groupe ARTAL depuis le
1^{er} Septembre 2016

Offres & marchés



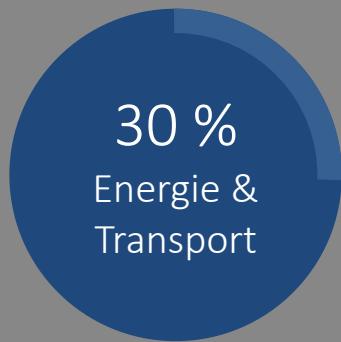
Etudes scientifiques
& techniques



Développement logiciel
& intégration de systèmes



Consulting
& assistance technique



Imagerie & Applications

Vision par ordinateur



Robotique



Lidar & 3D

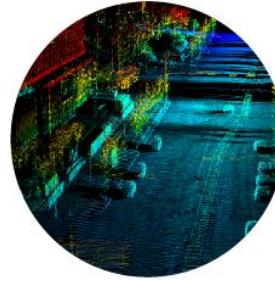


Image & vidéo



Etudes scientifiques



Conception systèmes de vision



Développement logiciel



Défense & Sécurité



Exploration spatiale



Transports



Acquisition & traitement d'image Plateforme IT

Propriété Magellium Alté Group



Adaptation métier

- Mon métier
 - Ingénieur d'études et développement, responsable technique
 - Thématiques de robotique spatiale:
 - Principalement robotique d'exploration planétaire
 - Robotique orbitale



Data Fusion for Space Robotics



AUTONOMOUS DECISION MAKING IN VERY LONG TRAVERSES (ADE)

A H2020 project funded by the European Commission
within the Strategic Research Cluster on Space Robotics
Technologies Programme

<http://www.h2020-ade.eu>

KEY CONCEPTS OF ADE

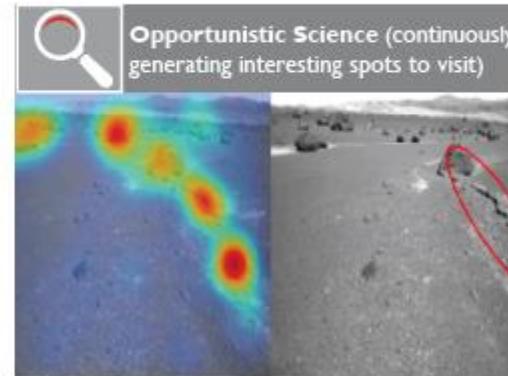
The ADE system encompasses the contemporaneous presence of the following **autonomous functionalities** for a planetary mission.



Decision making in presence of conflicting objectives

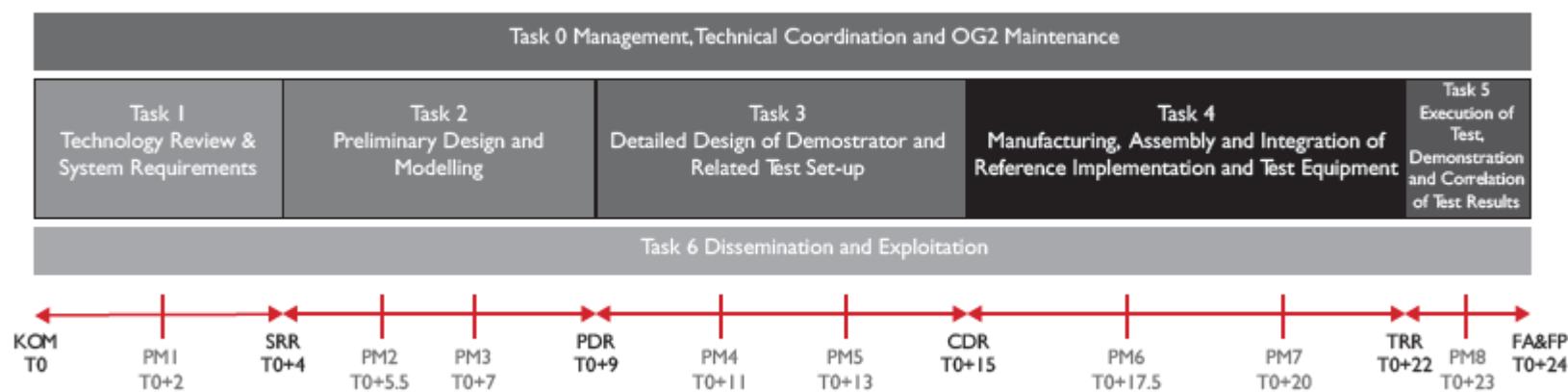


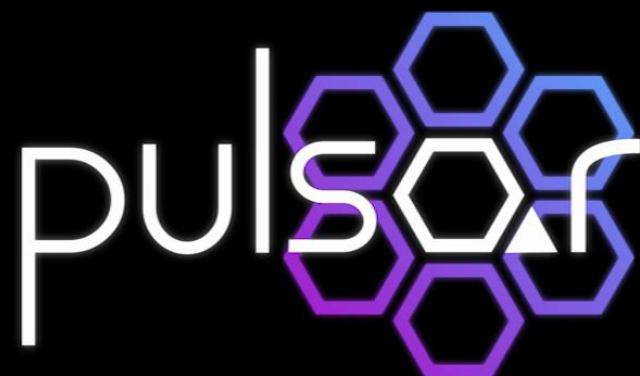
Long-term path planning and execution.
Adaptive execution of activities



Identification of specific features in the landscape,
assessment of their scientific value and planning of
scientific tasks

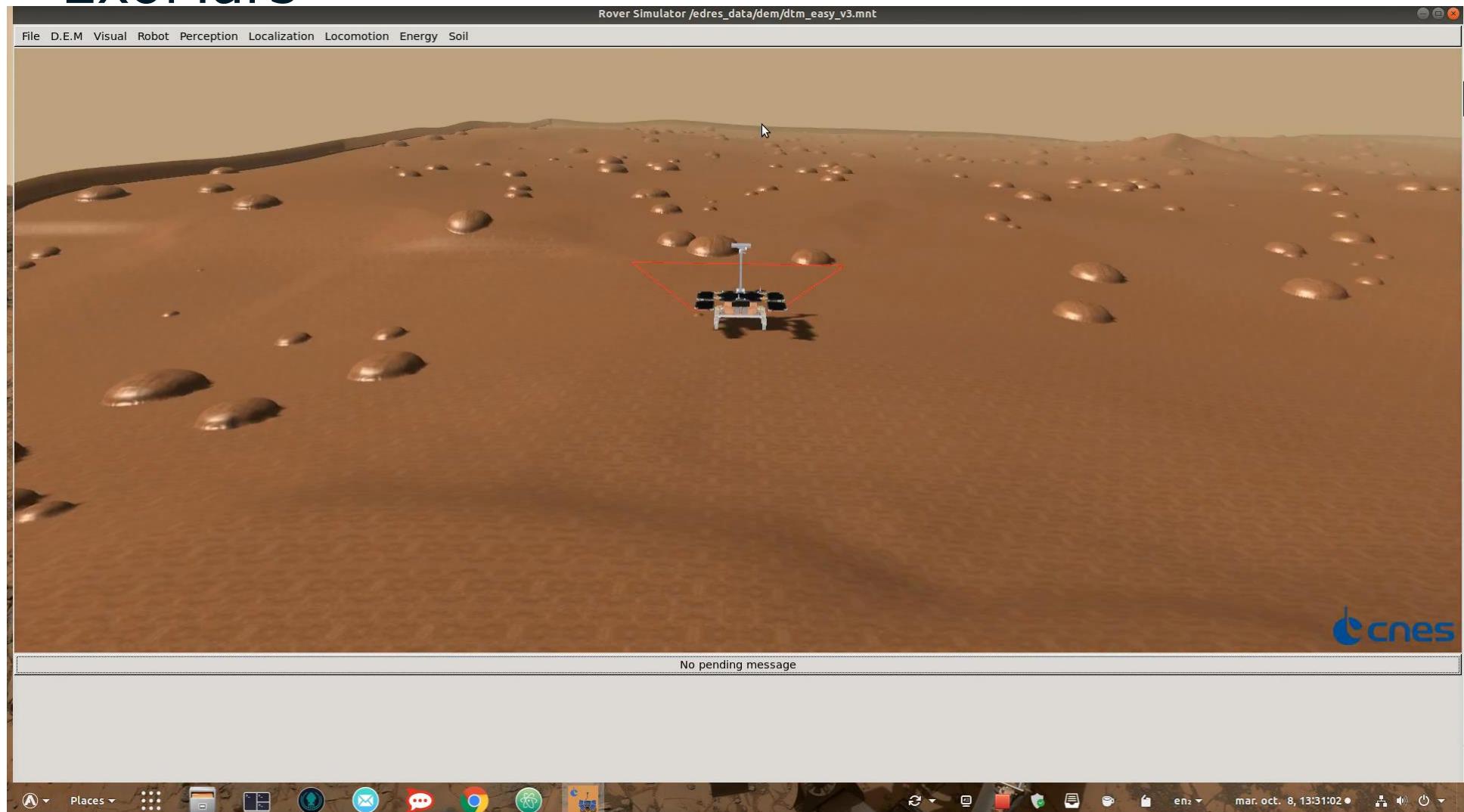
TIMELINE OF THE PROJECT





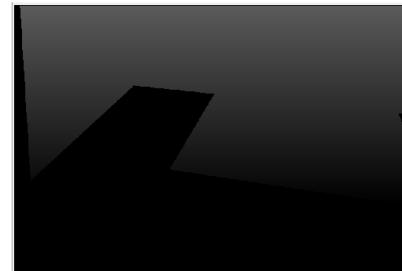
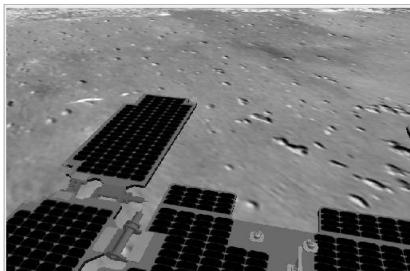
Prototype for an Ultra Large Structure Assembly Robot

- ExoMars



- ExoMars

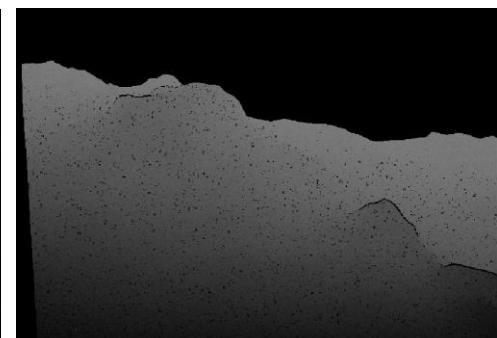
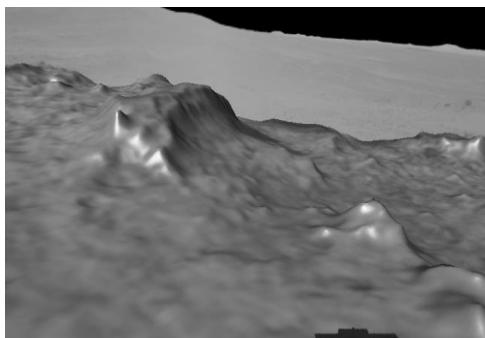
Ajout du modèle de rover ExoMars dans le simulateur



IHM pour le paramétrage des modèles de bruit

Altitude Perturbations		Localization Perturbations	
Pitch Temp. Drift (dg/s)	0.0	X Temp. Drift (mm/s)	0.0
Roll Temp. Drift (dg/s)	0.0	Y Temp. Drift (mm/s)	0.0
Azim Temp. Drift (dg/s)	0.0	X Dist. Drift (percent)	0.0
Pitch Dist. Drift (dg/100m)	0.0	Y Dist. Drift (percent)	0.0
Roll Dist. Drift (dg/100m)	0.0	X Rand. Noise std (mm)	0.0
Azim Dist. Drift (dg/100m)	0.0	Y Rand. Noise std (mm)	0.0
Pitch Const. Bias (dg)	0.0	X Const. Bias (mm)	0.0
Roll Const. Bias (dg)	0.0	Y Const. Bias (mm)	0.0
Azim Const. Bias (dg)	0.0		
Pitch Rand. Noise std (dg)	0.0		
Roll Rand. Noise std (dg)	0.0		
Azim Rand. Noise std (dg)	0.0		

Visualisation du bruit de perception

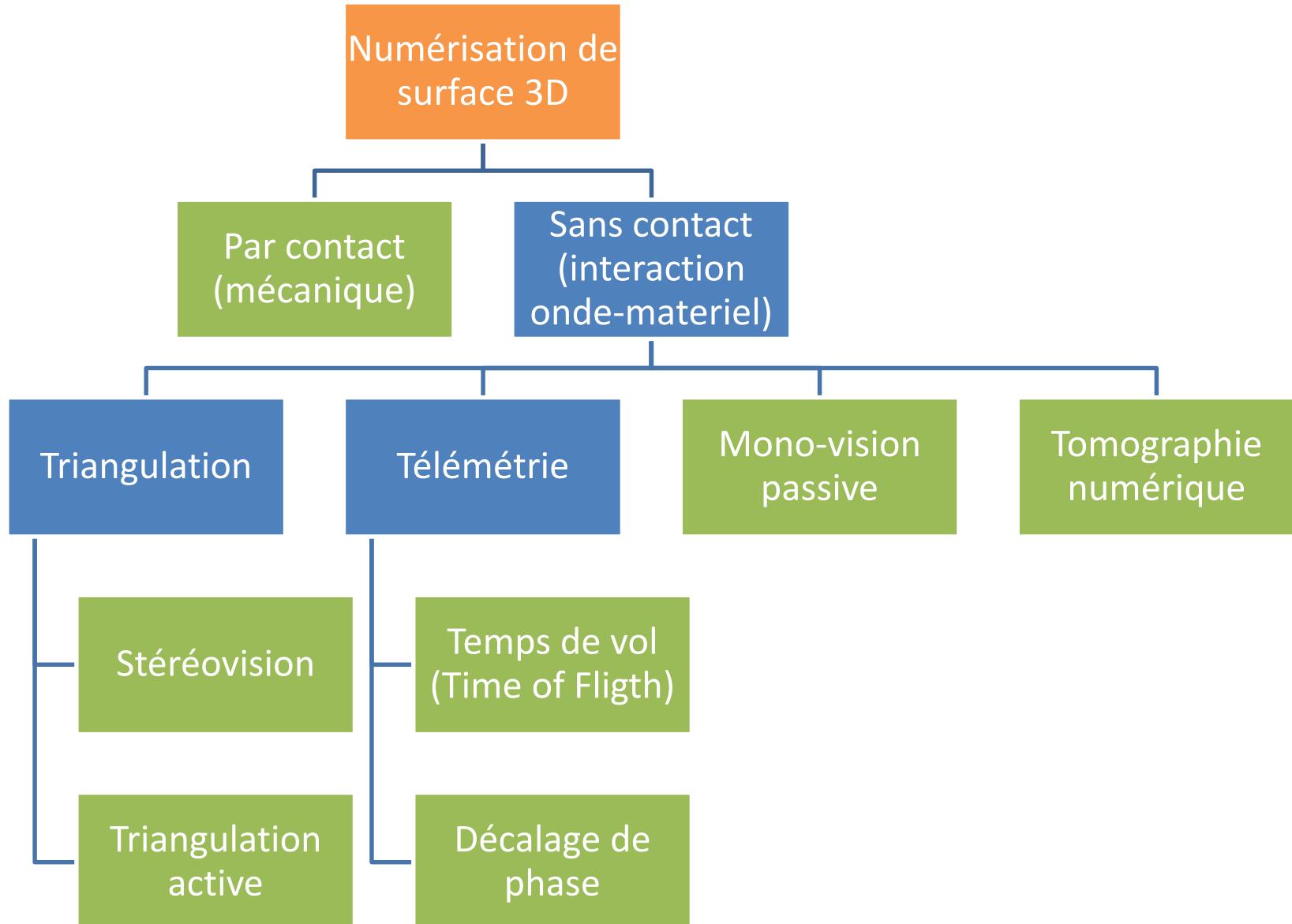


- Qu'est-ce que le domaine de la computer vision?
 - Computer vision is an interdisciplinary field that deals with how computers can be made to gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to automate tasks that the human visual system can do.(source Wikipedia)

Partie I : Perception 3D

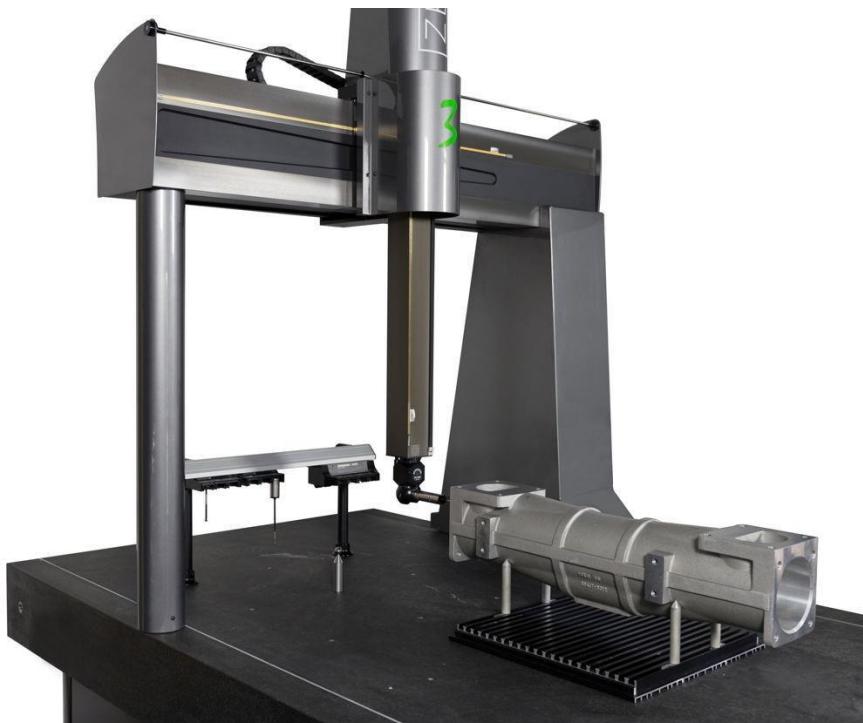
II. Systèmes d'acquisition 3D

Systèmes d'acquisition 3D



- Numérisation par contact

Propriété Magellium Artal Group

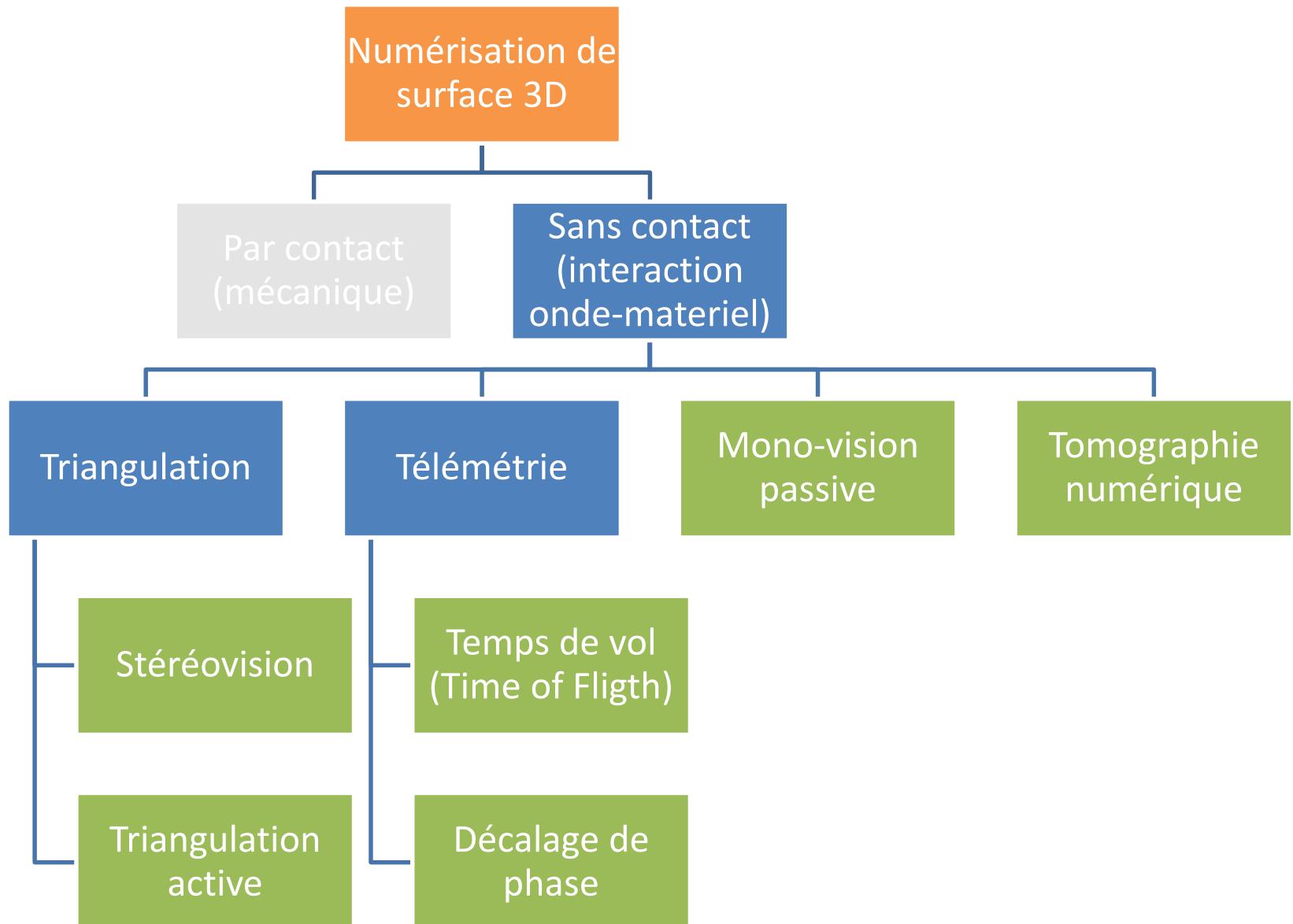


Machine à mesurer tridimensionnelle

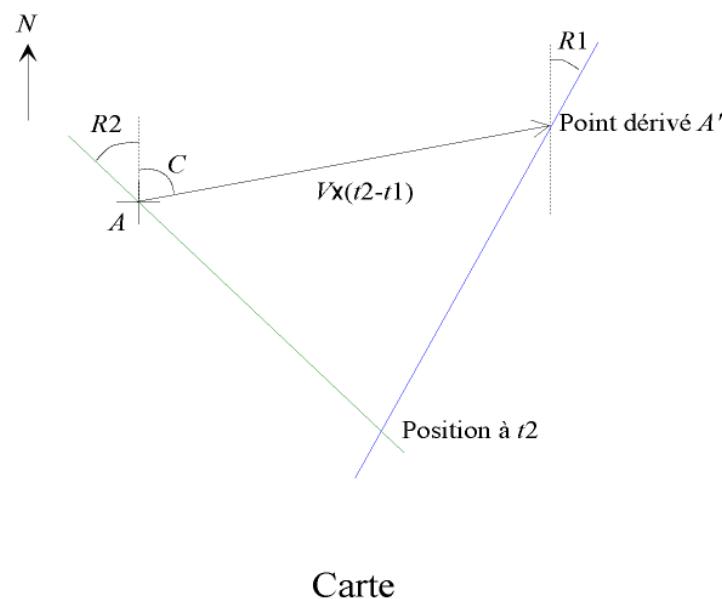
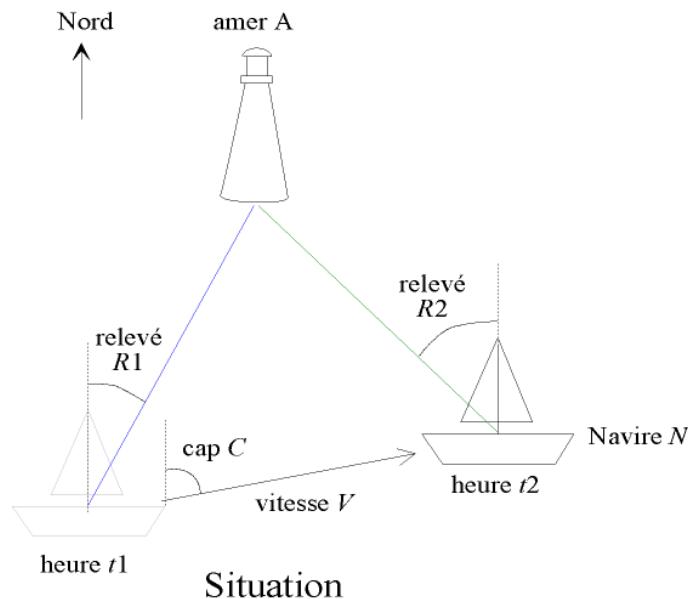


Bras articulé

Systèmes d'acquisition 3D

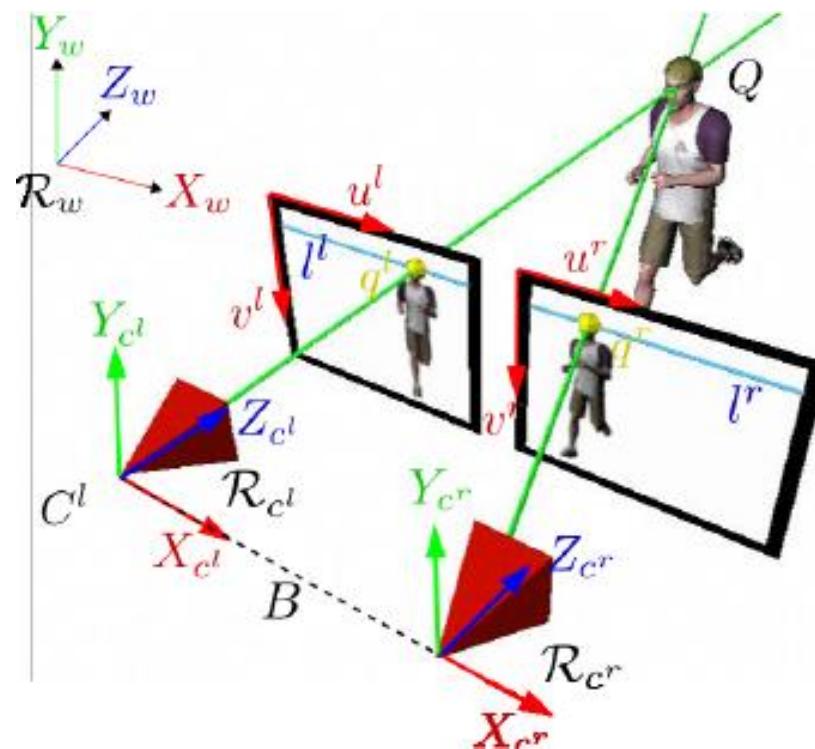


- Principe de la triangulation
 - Principe utilisé par les géomètres, les marins...



- Stéréovision

- Vision d'une même scène de deux endroits légèrement décalés l'un par rapport à l'autre
- Principe de la perception 3D chez l'Homme

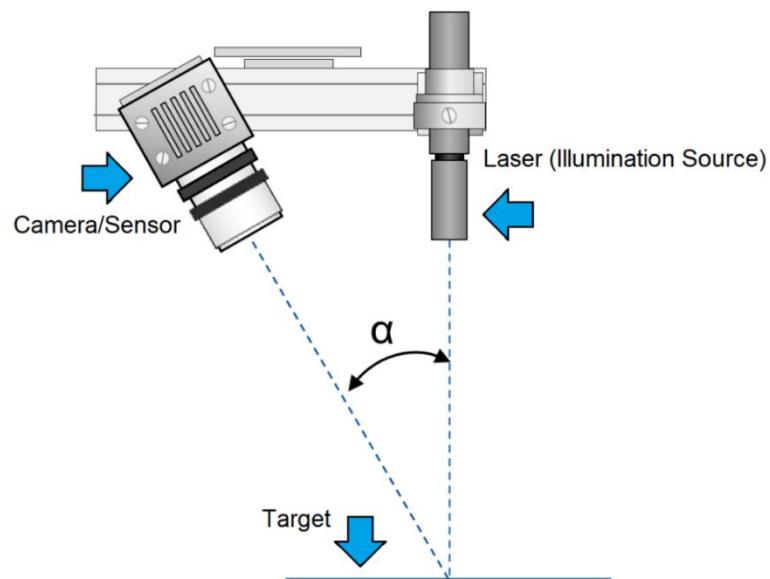


- Intérêts et limitations de la stéréovision
 - Avantages:
 - Coordonnées 3D obtenues sans éclairage spécifique (technique de vision passive)
 - Capteurs peu onéreux et robustes
 - Inconvénients:
 - Difficulté de l'appariement des points
 - Coûteux en ressources calculatoires

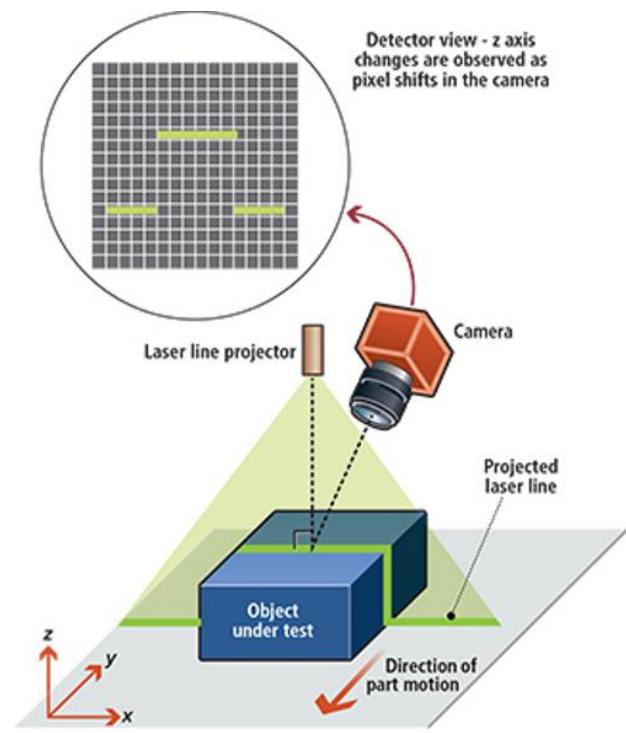
- Triangulation active

- Pointeur laser

- Un rayon laser est envoyé vers l'objet à mesurer
 - La lumière diffusée est observée par une caméra
 - On peut en déduire la profondeur du point



- Triangulation active
 - Laser ligne, profilométrie
 - Une image saisie donne une ligne de points
 - Un unique balayage suffit pour assurer la couverture de la surface

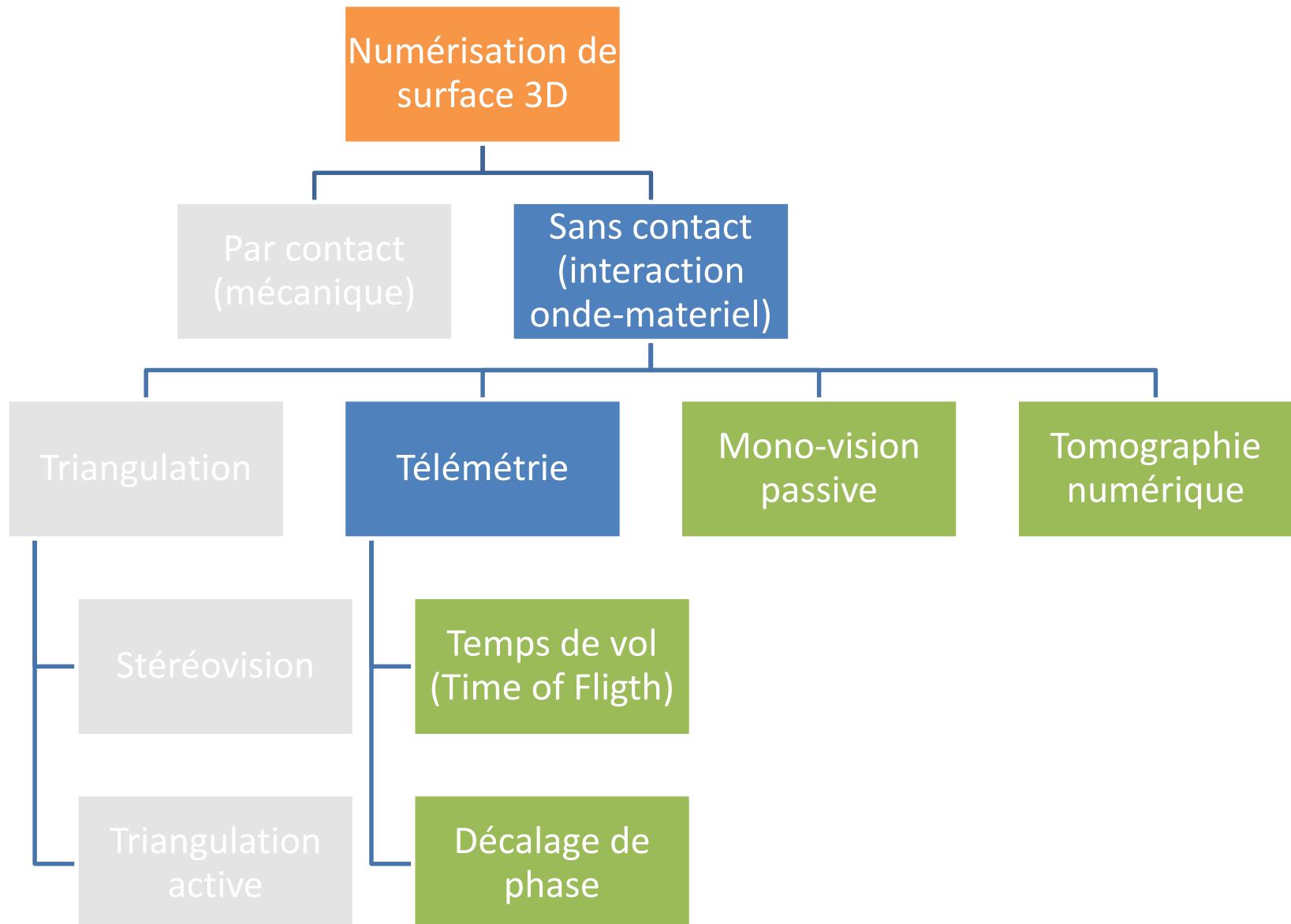


- Triangulation active
 - Lumière structurée
 - Projection de motifs aléatoires
 - Matching pixels/motif ou caméra-caméra



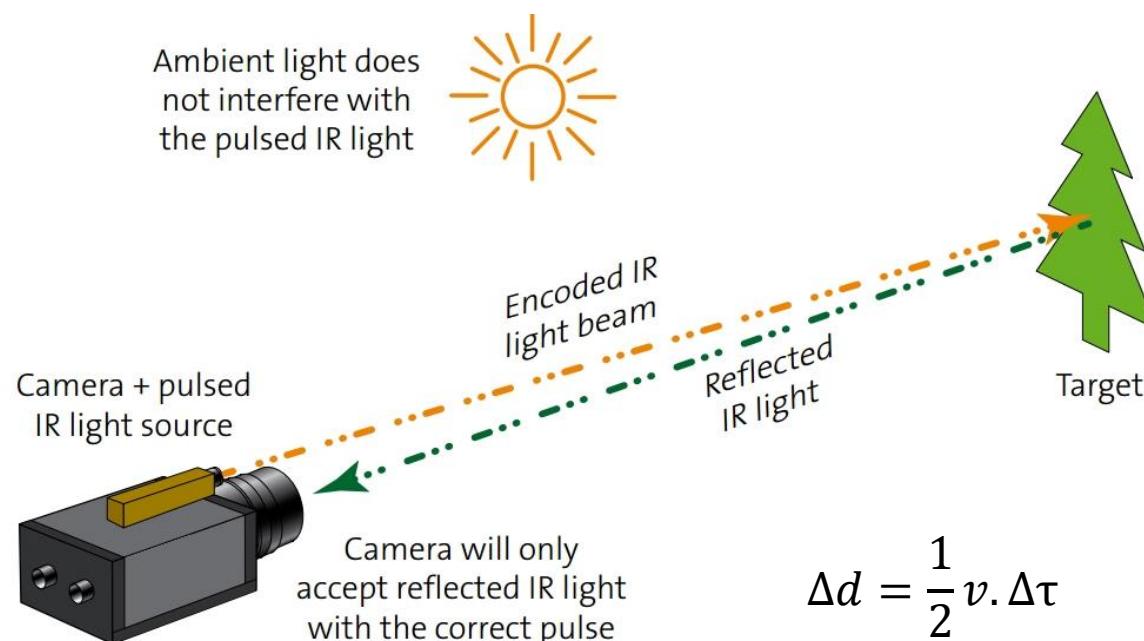
- Démo RealSense
 - Lumière structurée

Systèmes d'acquisition 3D



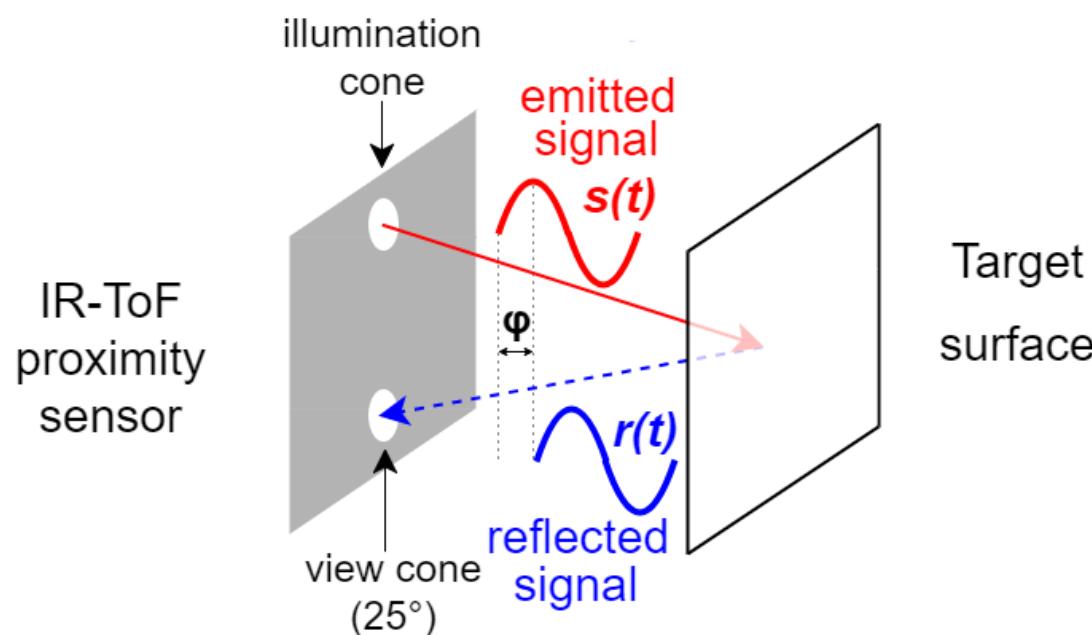
- Télémétrie par temps de vol (ToF)

- Envoi d'une impulsion
 - Lumineuse (laser)
 - Ultrasonore (SONAR)
 - Ondes radio (RADAR)

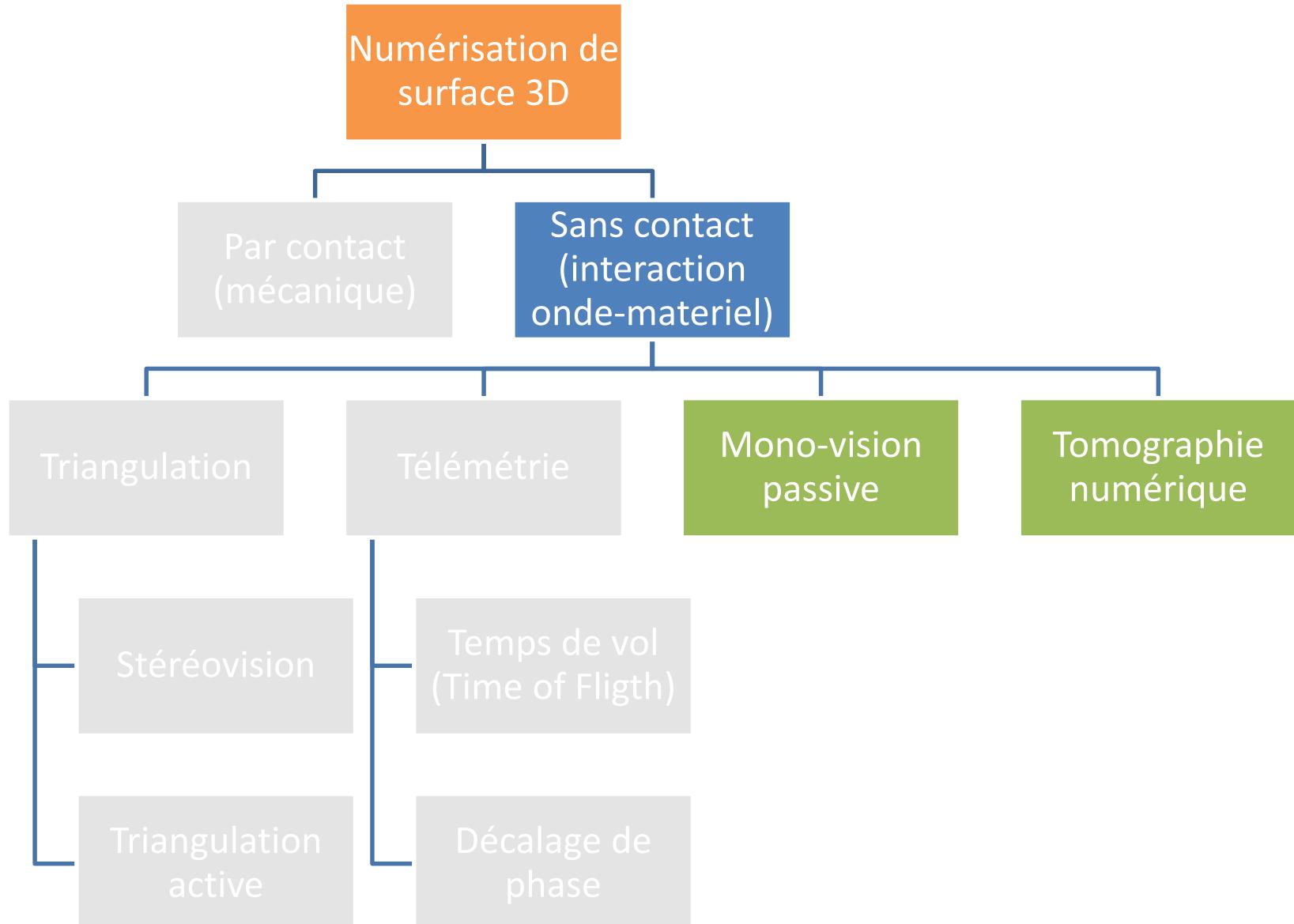


- Télémétrie par décalage de phase

- Envoi d'une impulsion
 - Lumineuse (laser)
 - Ultrasonore (SONAR)
 - Ondes radio (RADAR)

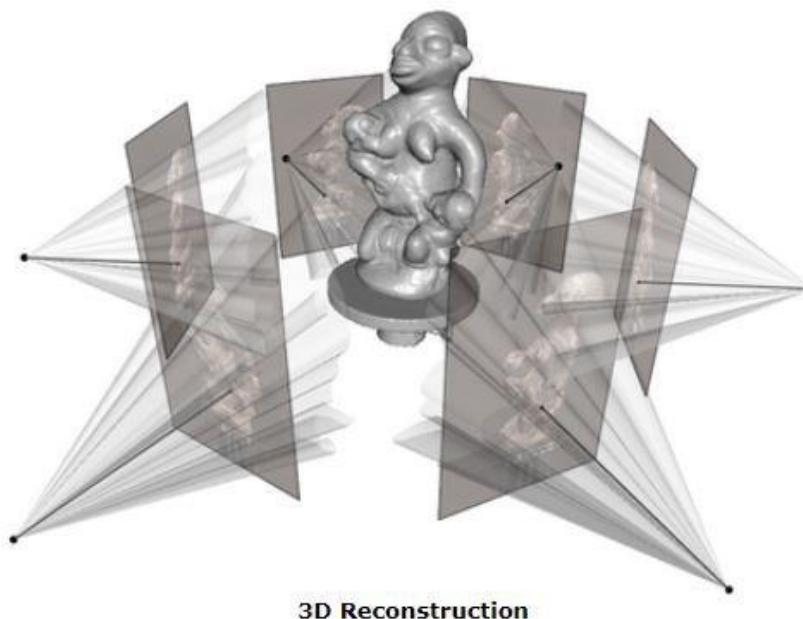


Systèmes d'acquisition 3D

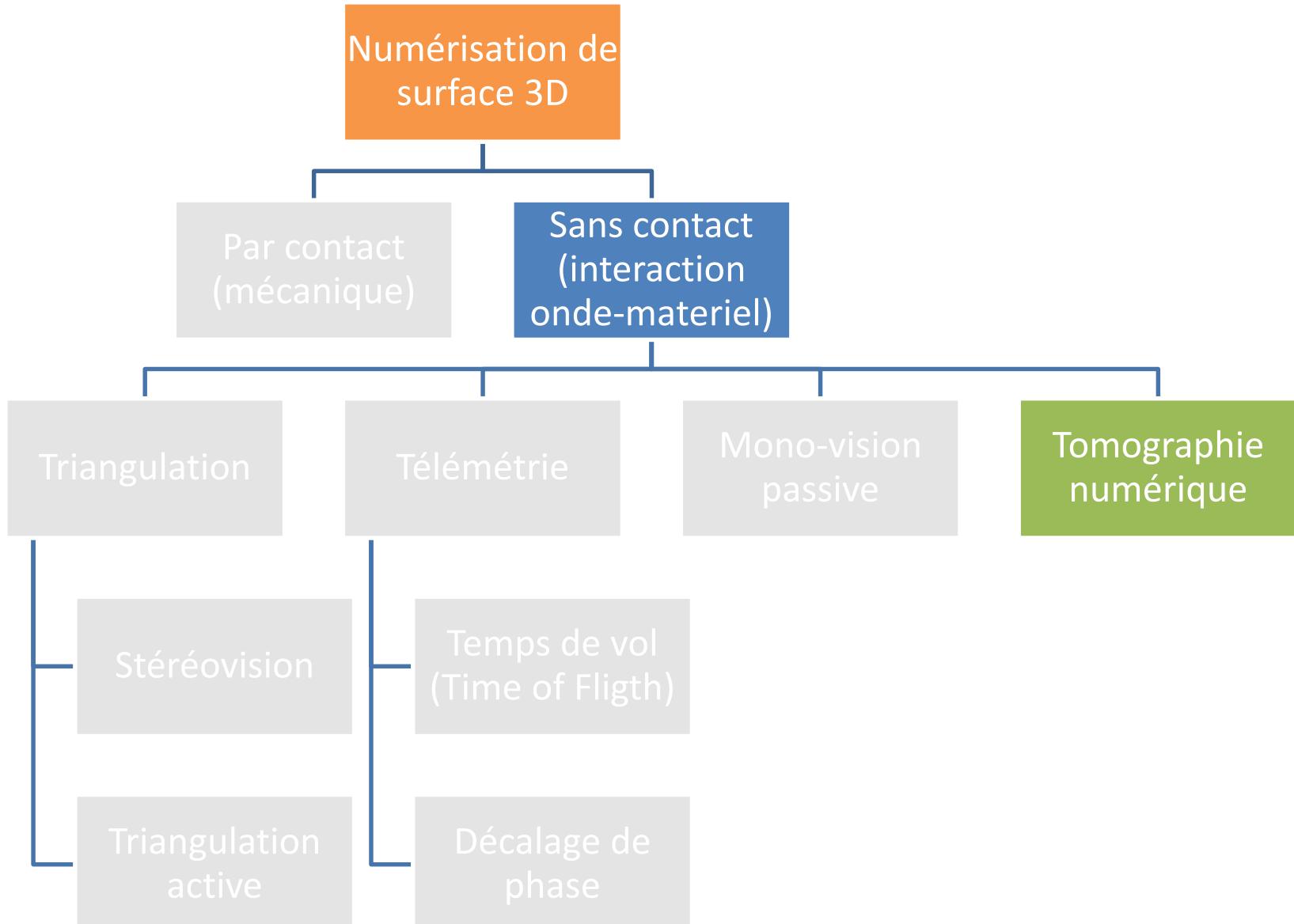


- **Monovision passive**

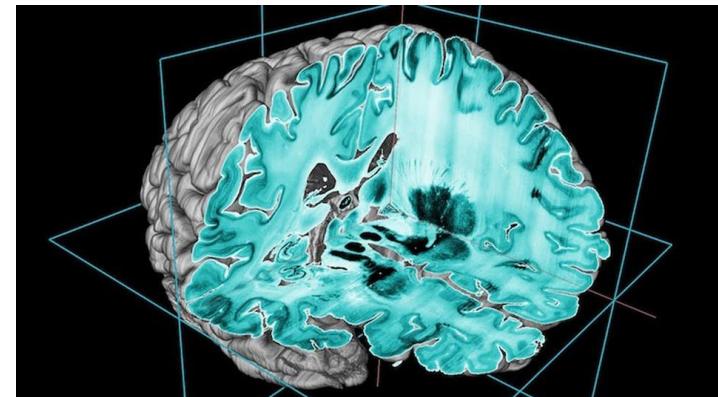
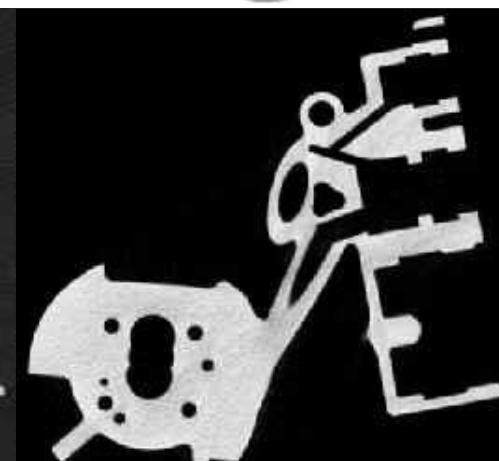
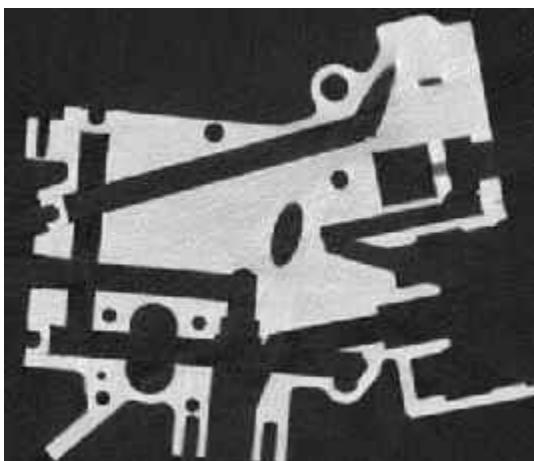
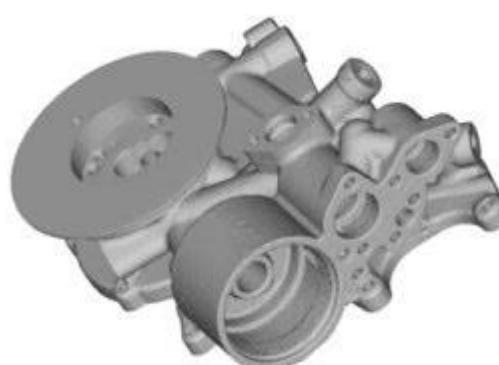
- Utilisation de simples images 2D pour obtenir des informations 3D
- Obtention d'angles de vue différents par mouvement de l'objet ou de la caméra



Systèmes d'acquisition 3D



- Tomographie numérique (Rayons X)
 - Principe de l'IRM
 - Superposition de coupes 2D successives

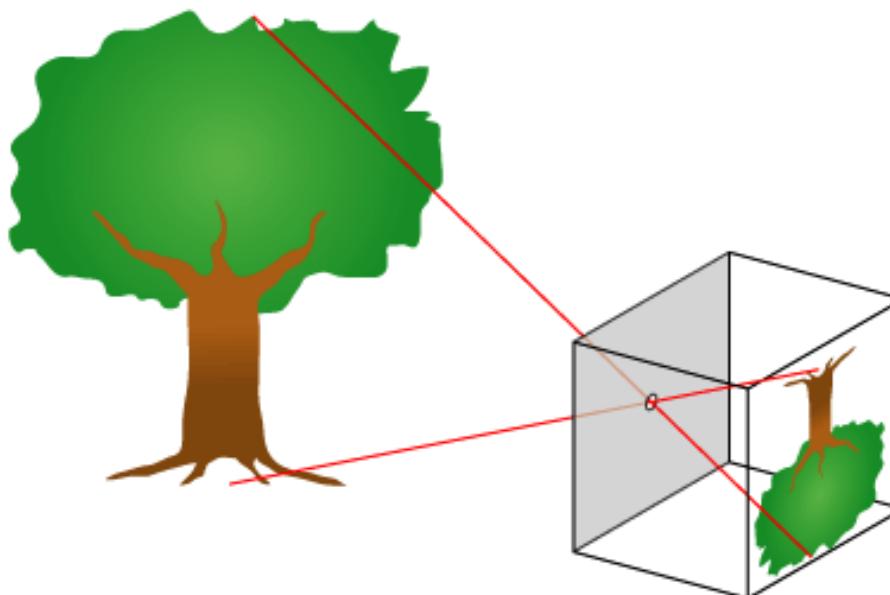


- Précision et étalonnage
 - Les points 3D sont des mesures géométriques obtenues par des principes physiques (lumière, contact, etc.) et mécaniques
 - Les erreurs systématiques de mesure peuvent être améliorées par calibrage/étalonnage

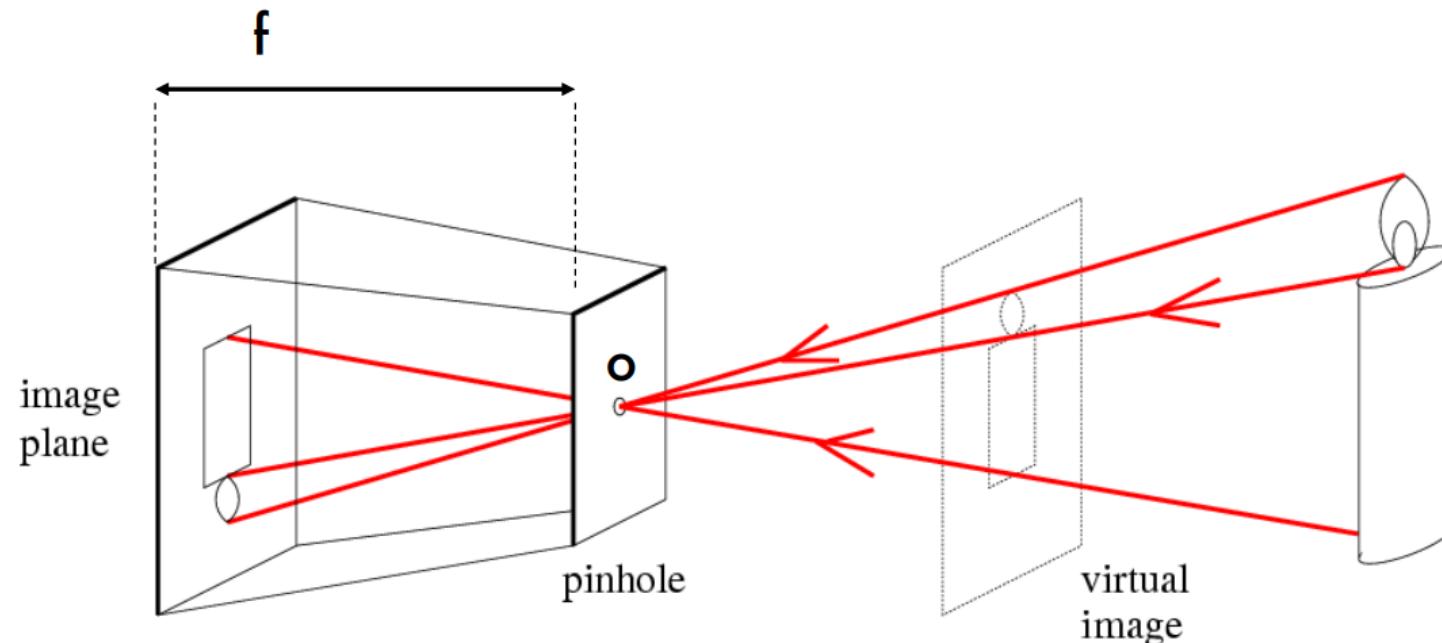
Partie I : Perception 3D

III. Modélisation de la camera

- Modèle sténopé ou pinhole



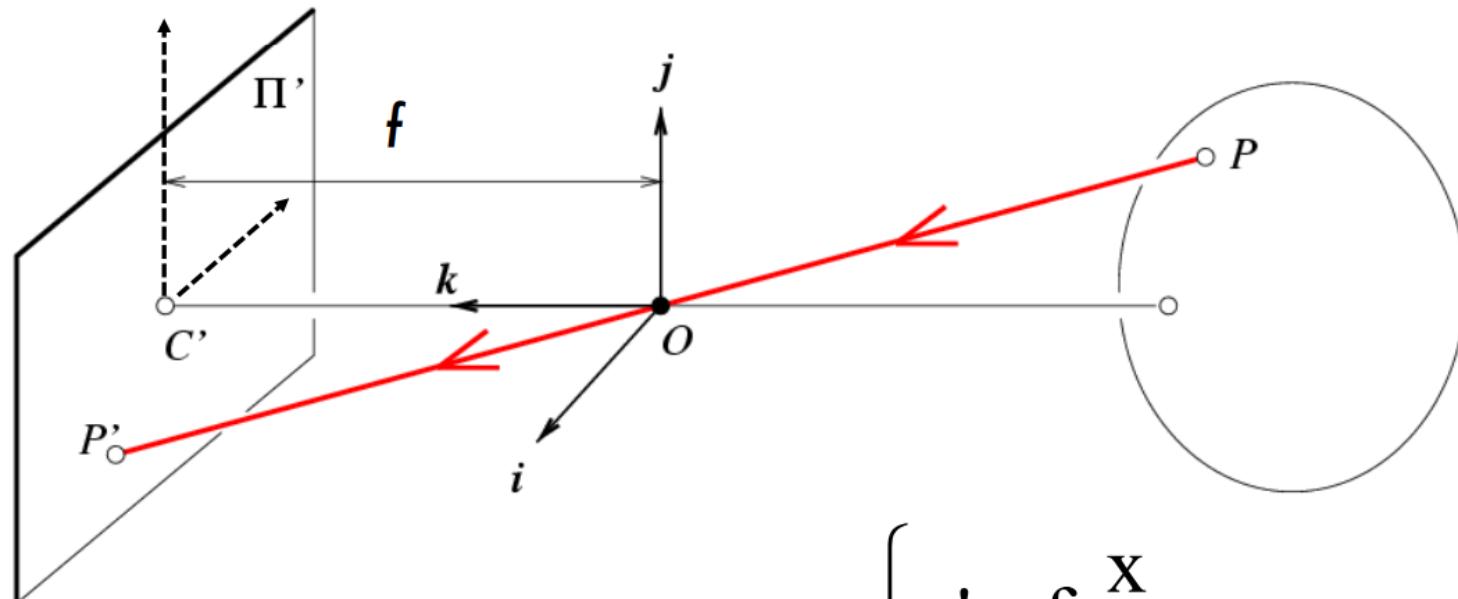
- Modèle sténopé ou pinhole



f = focal length

o = aperture = pinhole = center of the camera

- Modèle sténopé ou pinhole

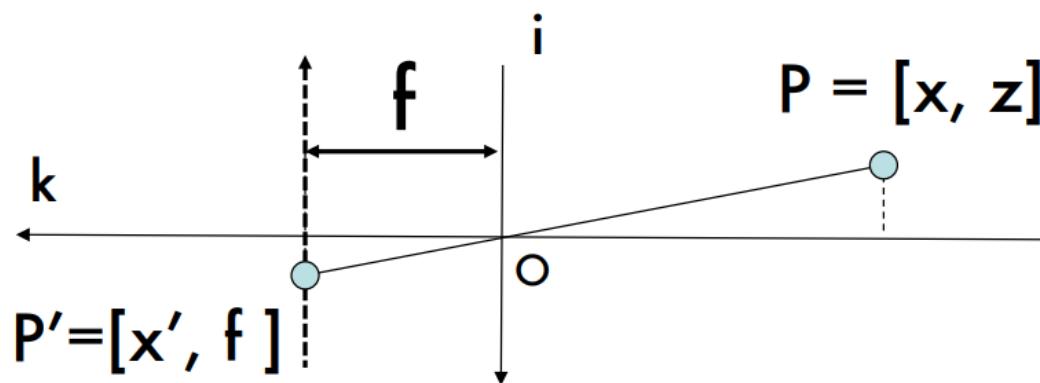
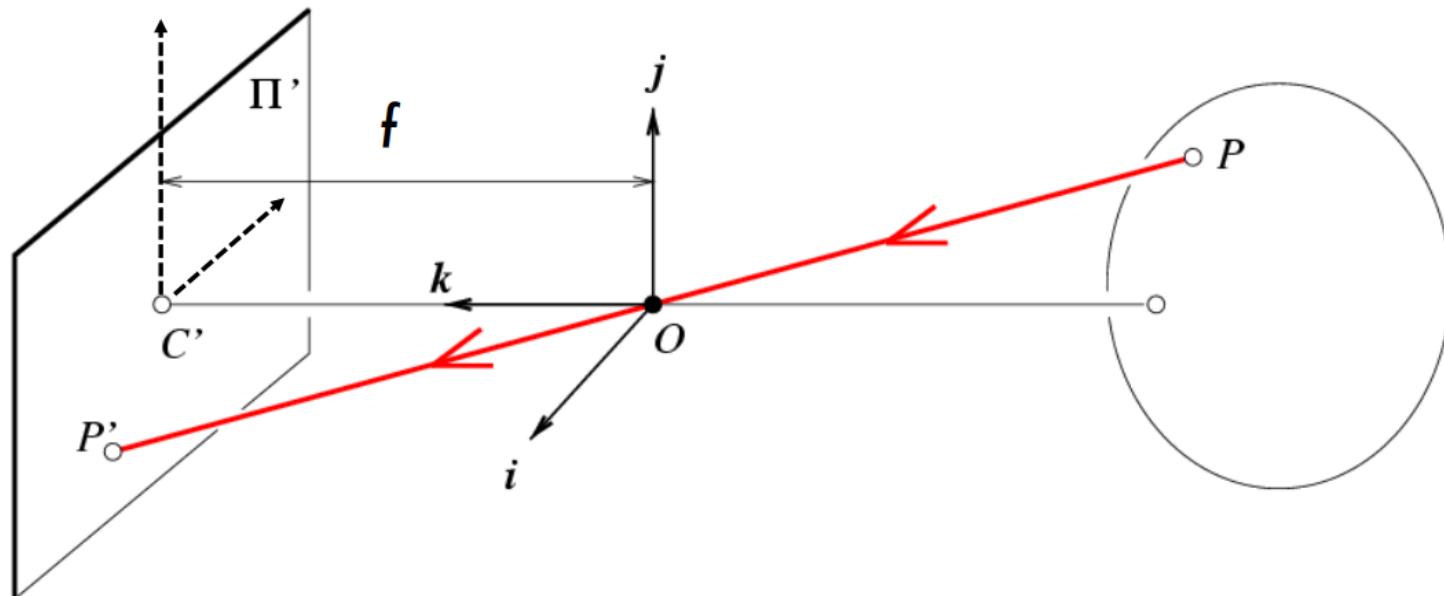


$$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow P' = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

$$\left\{ \begin{array}{l} x' = f \frac{x}{z} \\ y' = f \frac{y}{z} \end{array} \right.$$

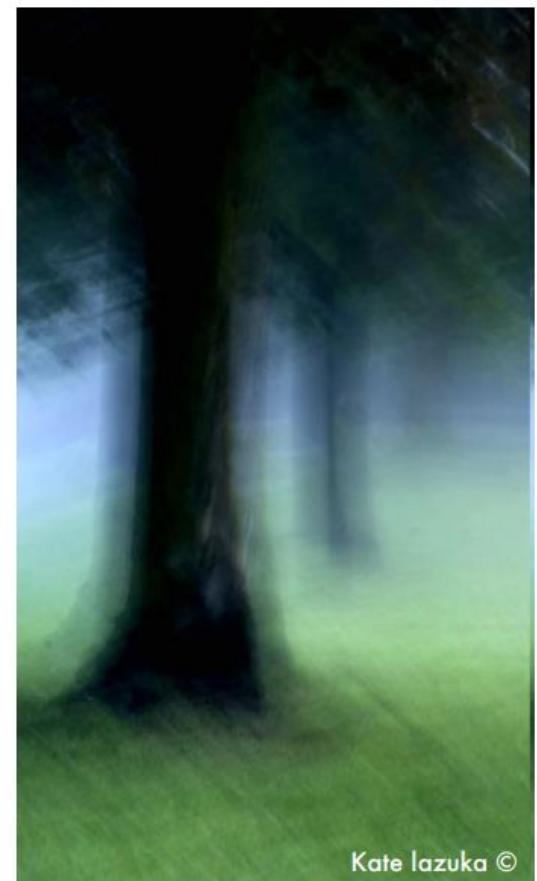
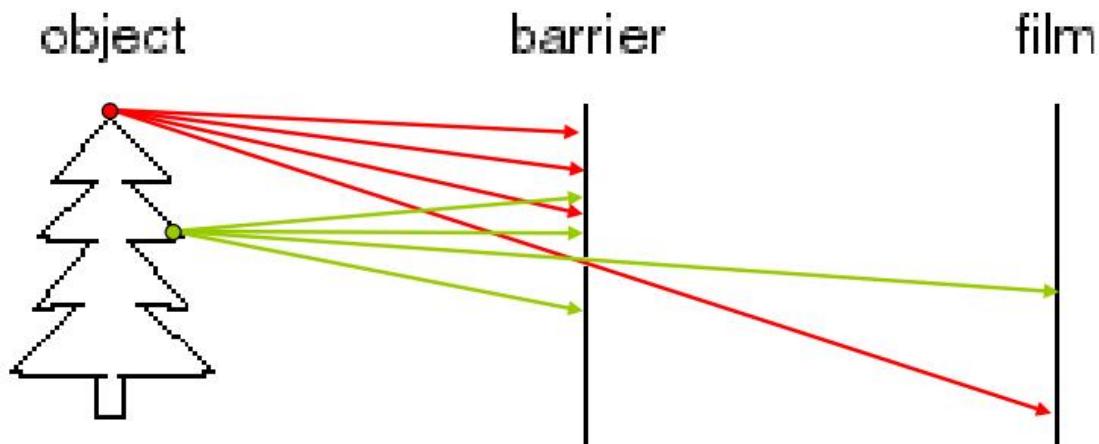
Derived using similar triangles

- Modèle sténopé ou pinhole



$$\frac{x'}{f} = \frac{x}{z}$$

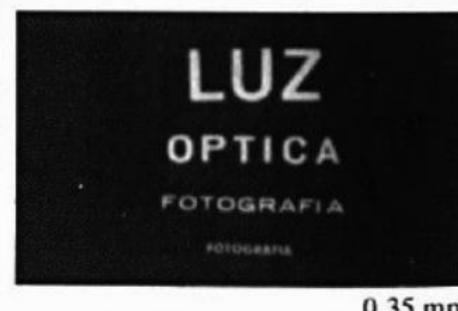
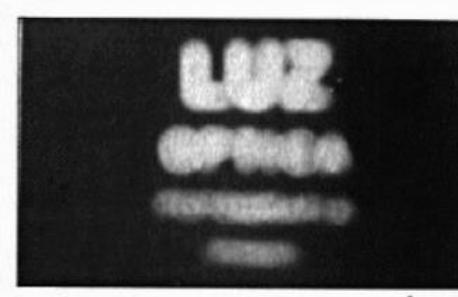
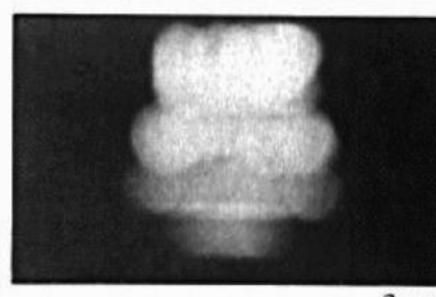
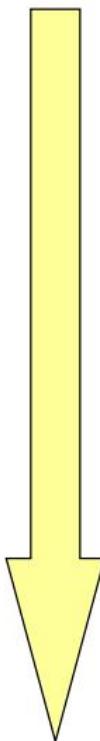
- Taille de l'ouverture



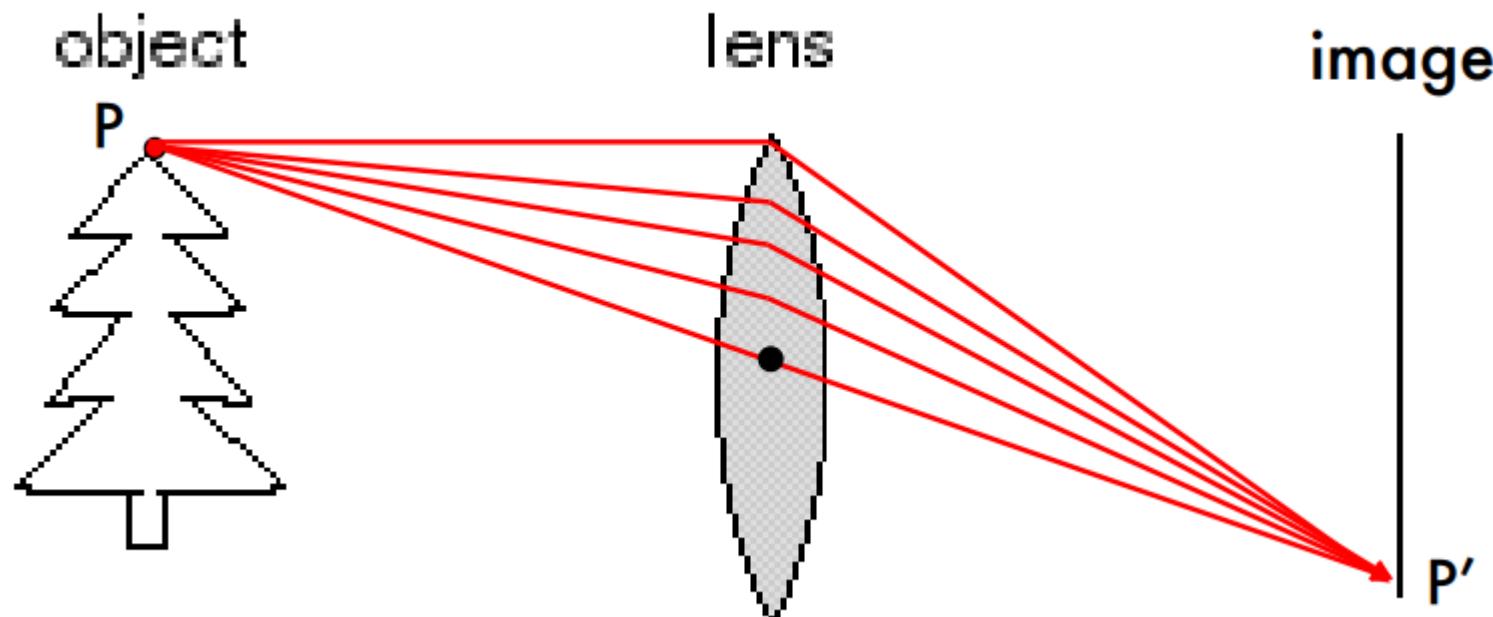
Kate lazuka ©

- Réduction de la taille de l'ouverture
 - Amélioration de la netteté
 - Réduction de la luminosité

Shrinking
aperture
size

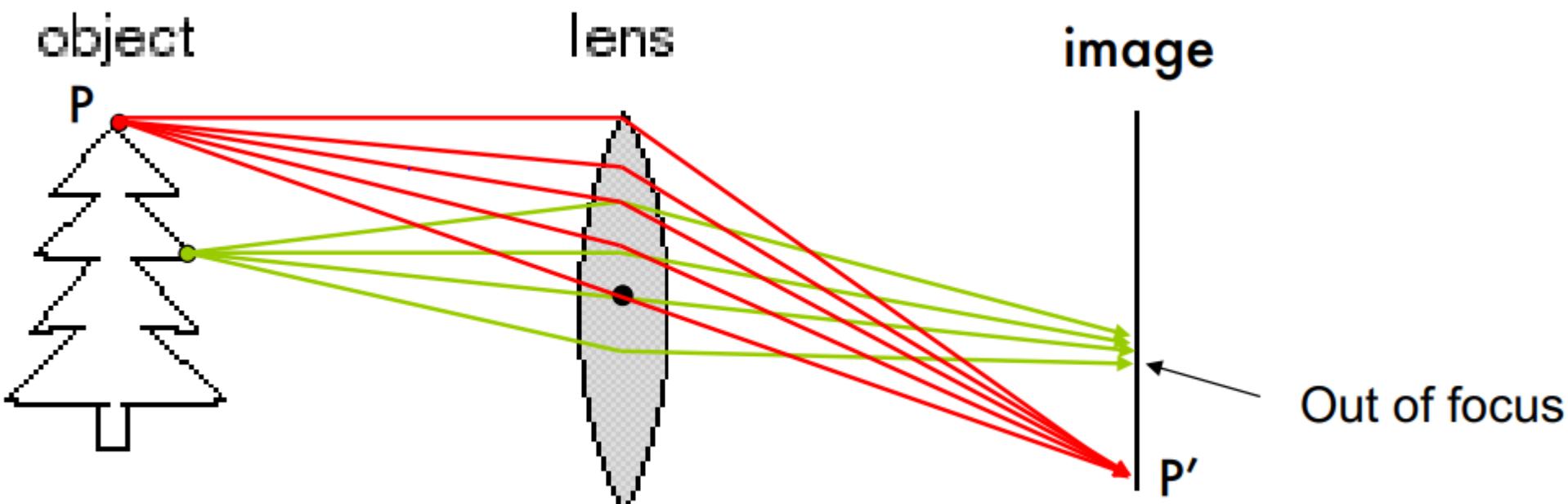


- Ajout d'une lentille pour concentrer la lumière sur le plan image
 - Amélioration de la luminosité



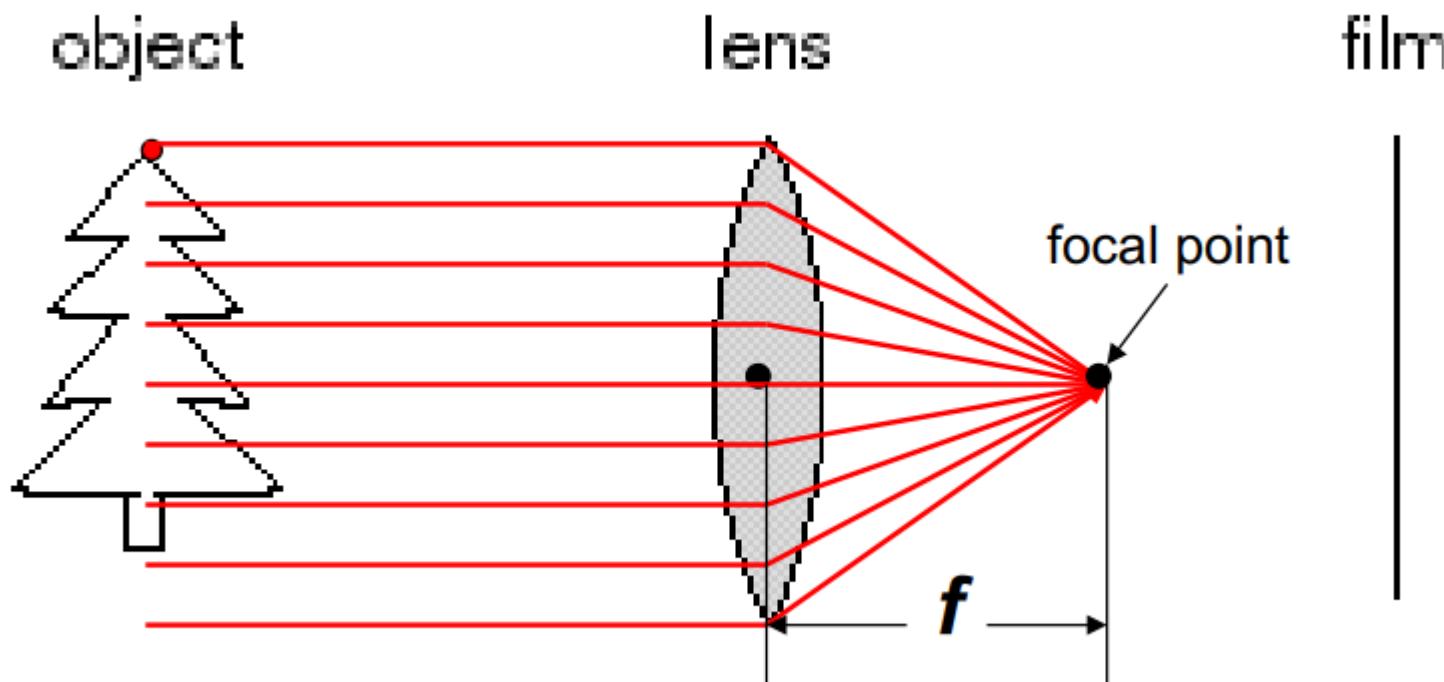
- **Distance focale**

- Les objets sont correctement projetés sur le plan image uniquement lorsqu'ils sont à une certaine distance de la lentille

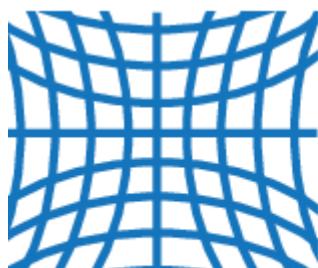


- **Distance focale**

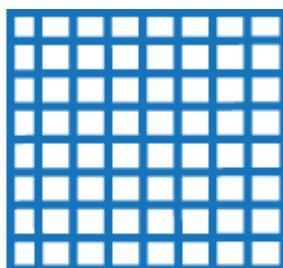
- Tous les rayons parallèles à l'axe optique convergent en un point unique, le point focal, sur un plan situé à la distance focale du centre de la lentille



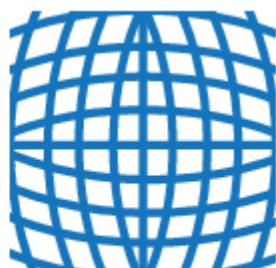
- Problème des lentilles
 - Distorsion radiale
 - La distorsion est plus importante pour les rayons qui passent près du bord de la lentille



Negative radial distortion
"pincushion"



No distortion

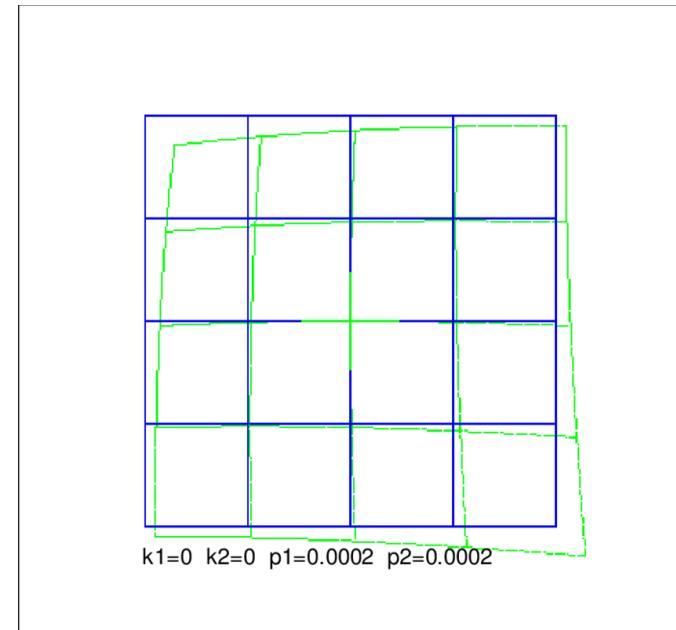
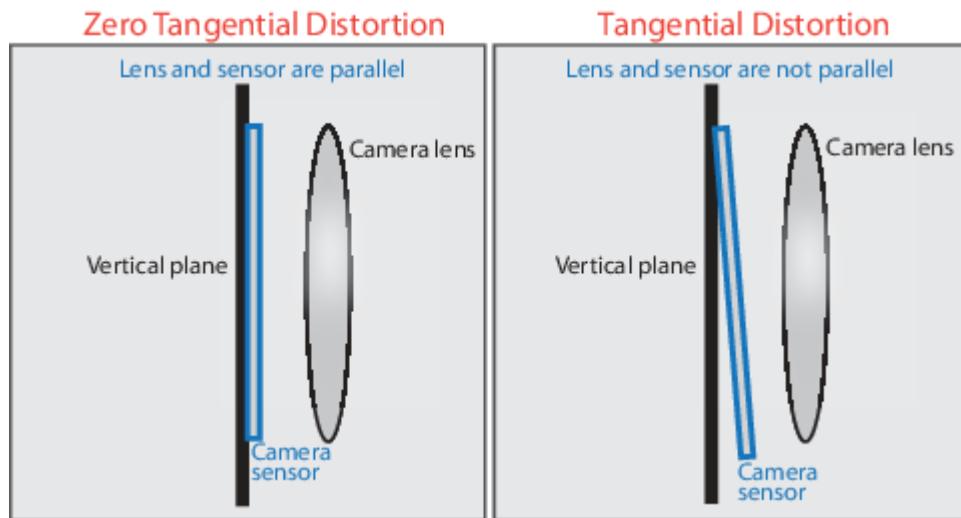


Positive radial distortion
"barrel"

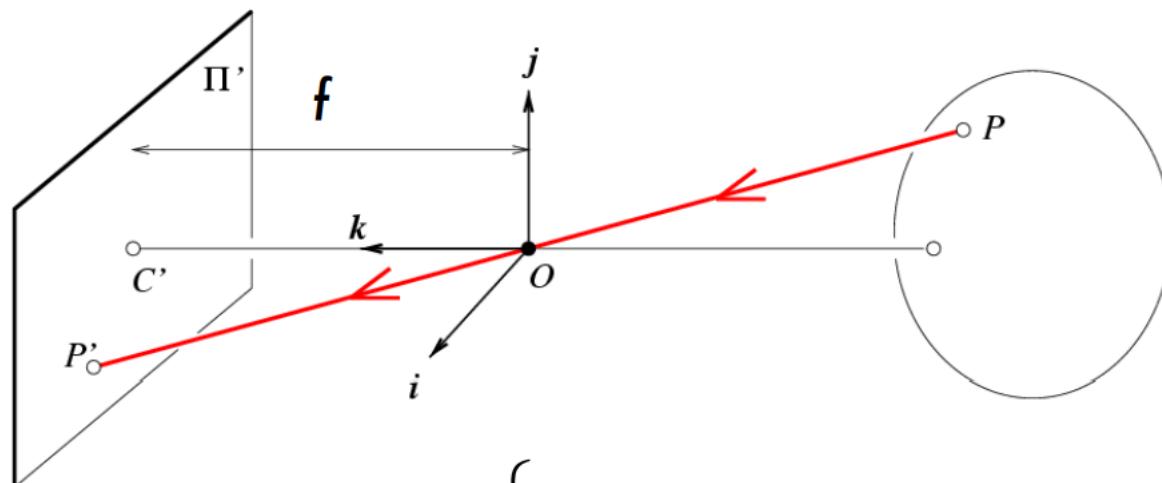


Image magnification decreases
with distance from the optical axis

- Problème des lentilles
 - Distorsion tangentielle
 - La distorsion est plus importante pour les rayons qui passent près du bord de la lentille



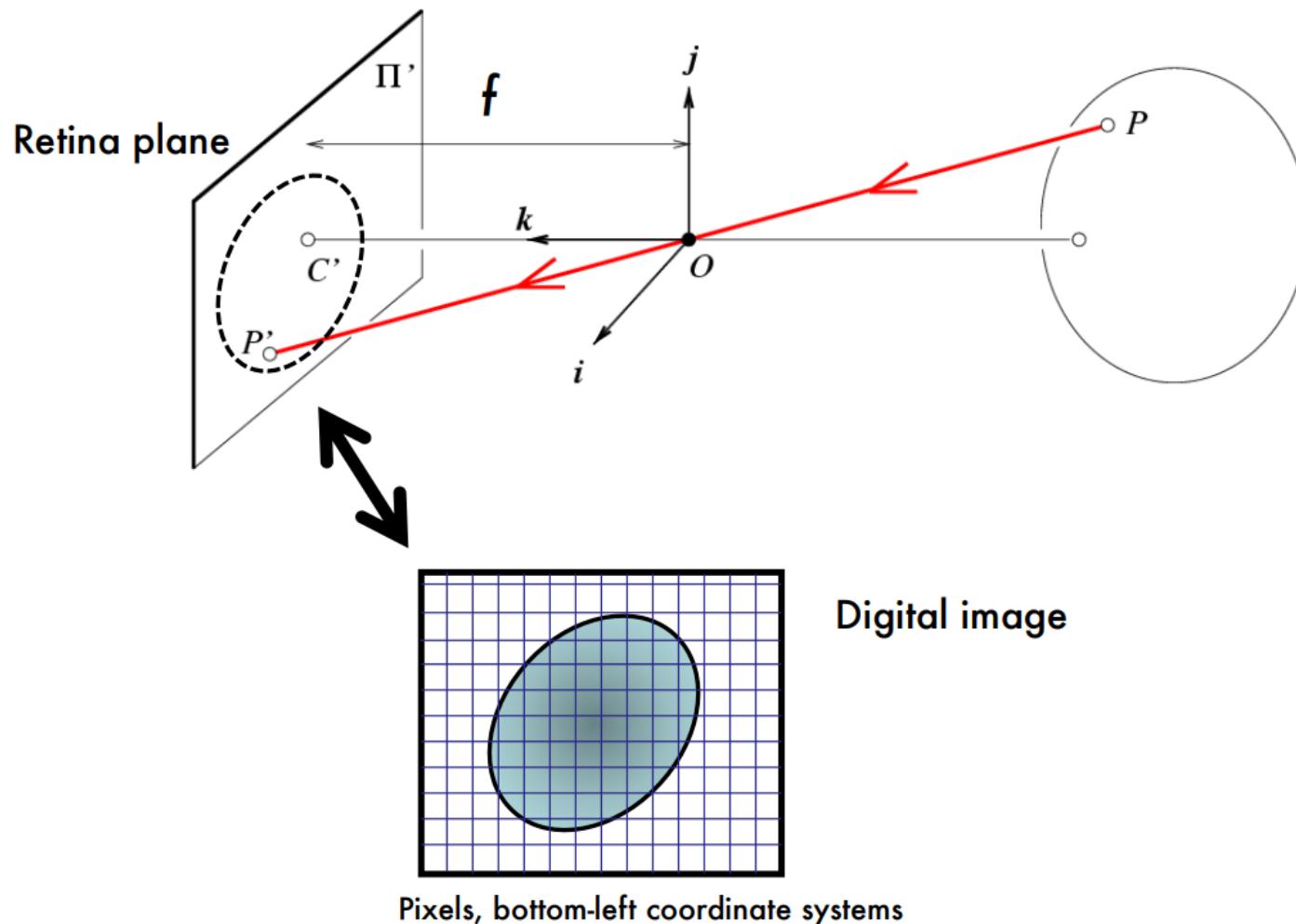
- Géométrie optique



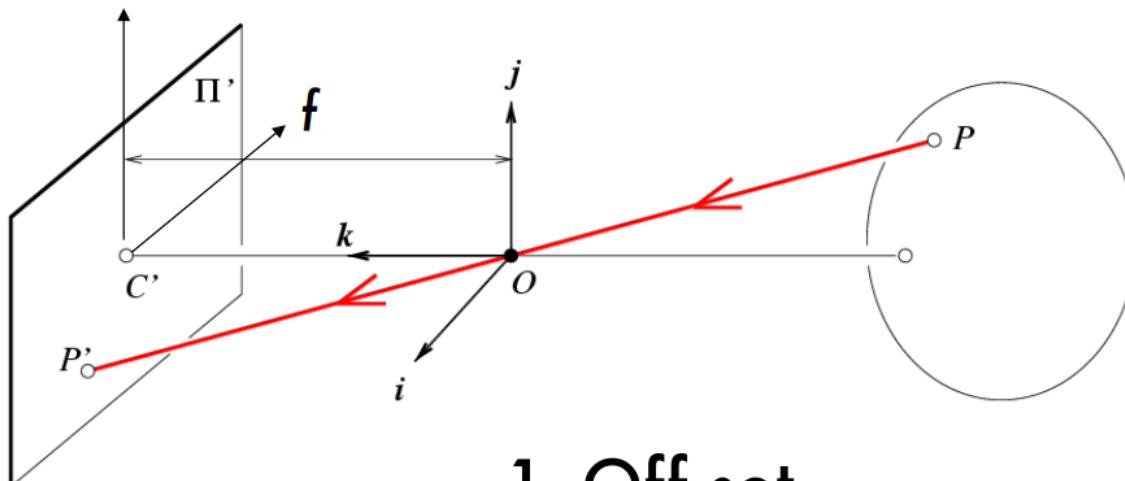
$$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \rightarrow P' = \begin{bmatrix} x' \\ y' \end{bmatrix} \quad \left\{ \begin{array}{l} x' = f \frac{x}{z} \\ y' = f \frac{y}{z} \end{array} \right. \quad \mathcal{R}^3 \xrightarrow{E} \mathcal{R}^2$$

f = focal length
o = center of the camera

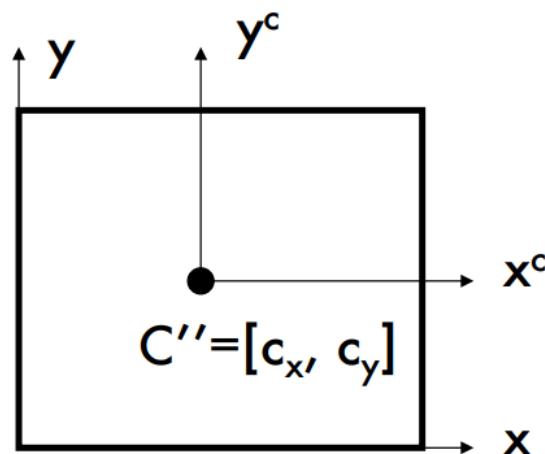
- Géométrie optique



- Centre optique

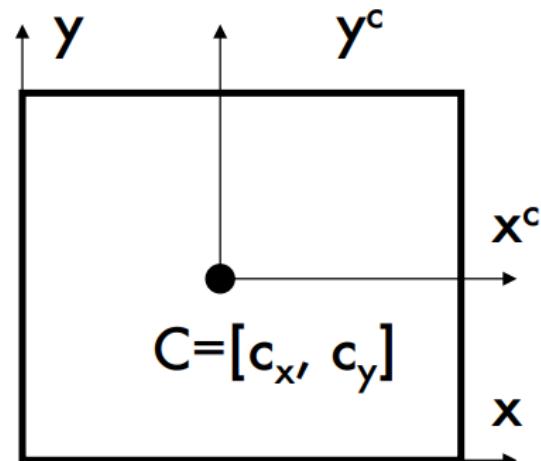
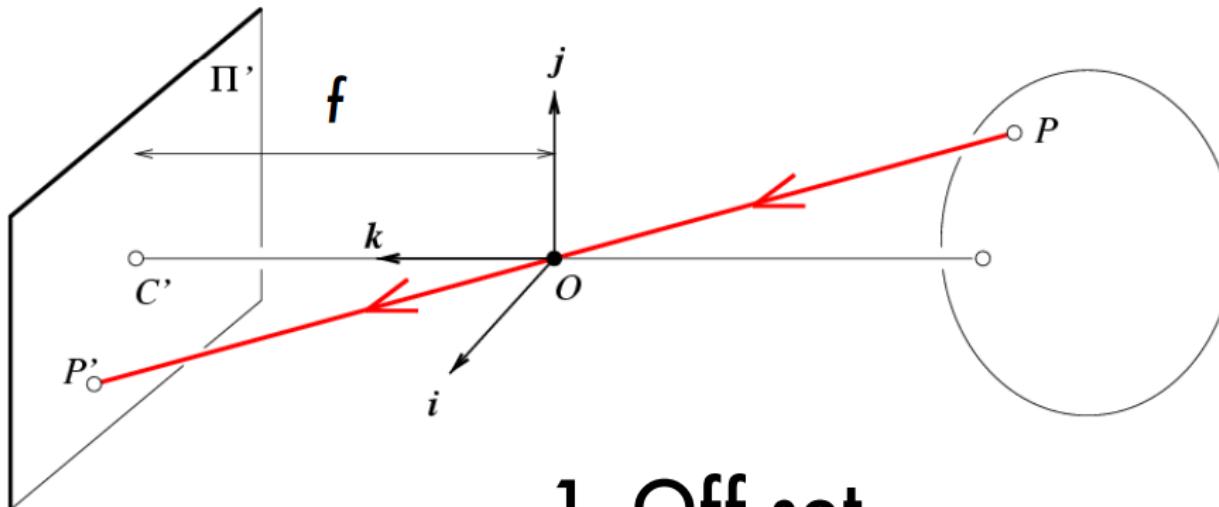


1. Off set



$$(x, y, z) \rightarrow \left(f \frac{x}{z} + c_x, f \frac{y}{z} + c_y \right)$$

- Conversion en pixels



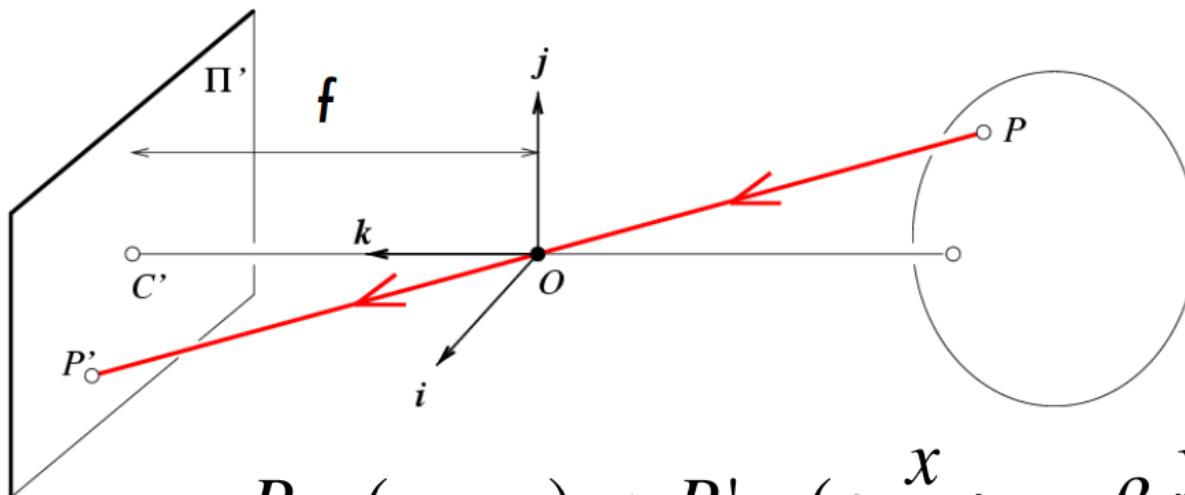
1. Off set
2. From metric to pixels

$$(x, y, z) \rightarrow \left(f \frac{k}{z} + c_x, f \frac{l}{z} + c_y \right)$$

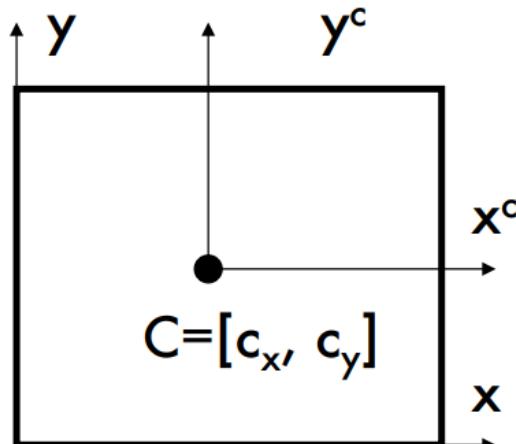
Units: k, l : pixel/m
 f : m

Non-square pixels
 α, β : pixel

- Linéarité



$$P = (x, y, z) \rightarrow P' = (\alpha \frac{x}{z} + c_x, \beta \frac{y}{z} + c_y)$$



- Is this a linear transformation?
No – division by z is nonlinear
- Can we express it in a matrix form?

- Coordonnées homogènes

E → H

$$(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

homogeneous image coordinates

$$(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

homogeneous scene coordinates

H → E

$$\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)$$

- Transformation projective

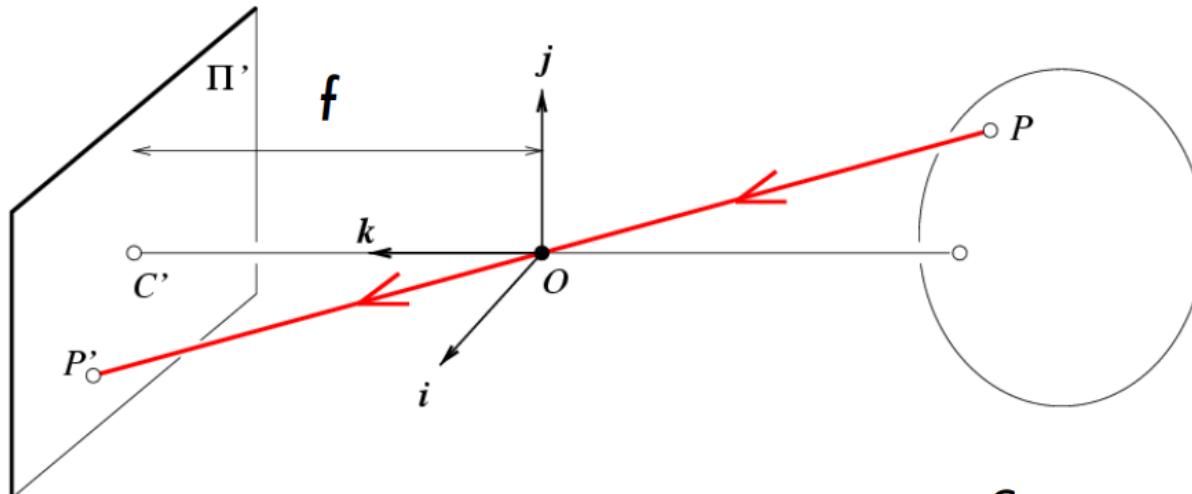
$$P_h' = \begin{bmatrix} \alpha x + c_x z \\ \beta y + c_y z \\ z \end{bmatrix} = \begin{bmatrix} \alpha & 0 & c_x & 0 \\ 0 & \beta & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad P_h$$

Homogenous **Euclidian**

$\overbrace{P_h'}^{} \rightarrow P' = (\alpha \frac{x}{z} + c_x, \beta \frac{y}{z} + c_y)$

$M = \begin{bmatrix} \alpha & 0 & c_x & 0 \\ 0 & \beta & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

- Matrice intrinsèque ou matrice caméra

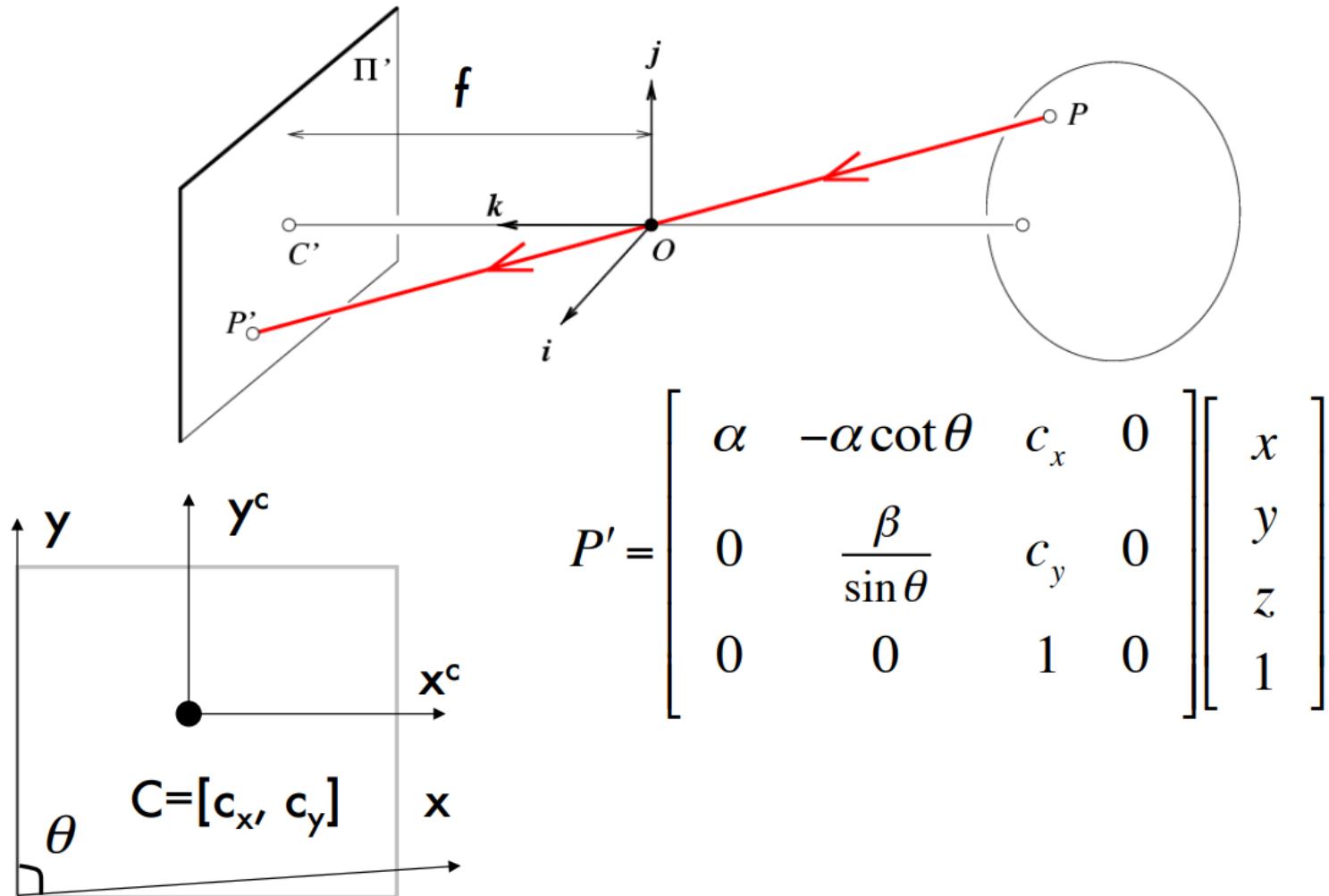


$$\begin{aligned}
 P' &= M P \\
 &= K \begin{bmatrix} I & 0 \end{bmatrix} P
 \end{aligned}$$

Camera matrix K

$$P' = \begin{bmatrix} \alpha & 0 & c_x \\ 0 & \beta & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

- Le coefficient 'skew'



- Transformation projective canonique

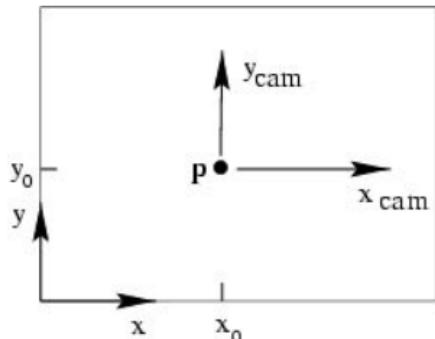
$$P' = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_M \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$P' = M P$$

$$\mathcal{R}^4 \xrightarrow{H} \mathcal{R}^3$$

$$P'_i = \begin{bmatrix} \frac{x}{z} \\ \frac{y}{z} \end{bmatrix}$$

- Résumé



$$(X, Y, Z)^T \mapsto (fX/Z + p_x, fY/Z + p_y)^T$$

$(p_x, p_y)^T$ principal point

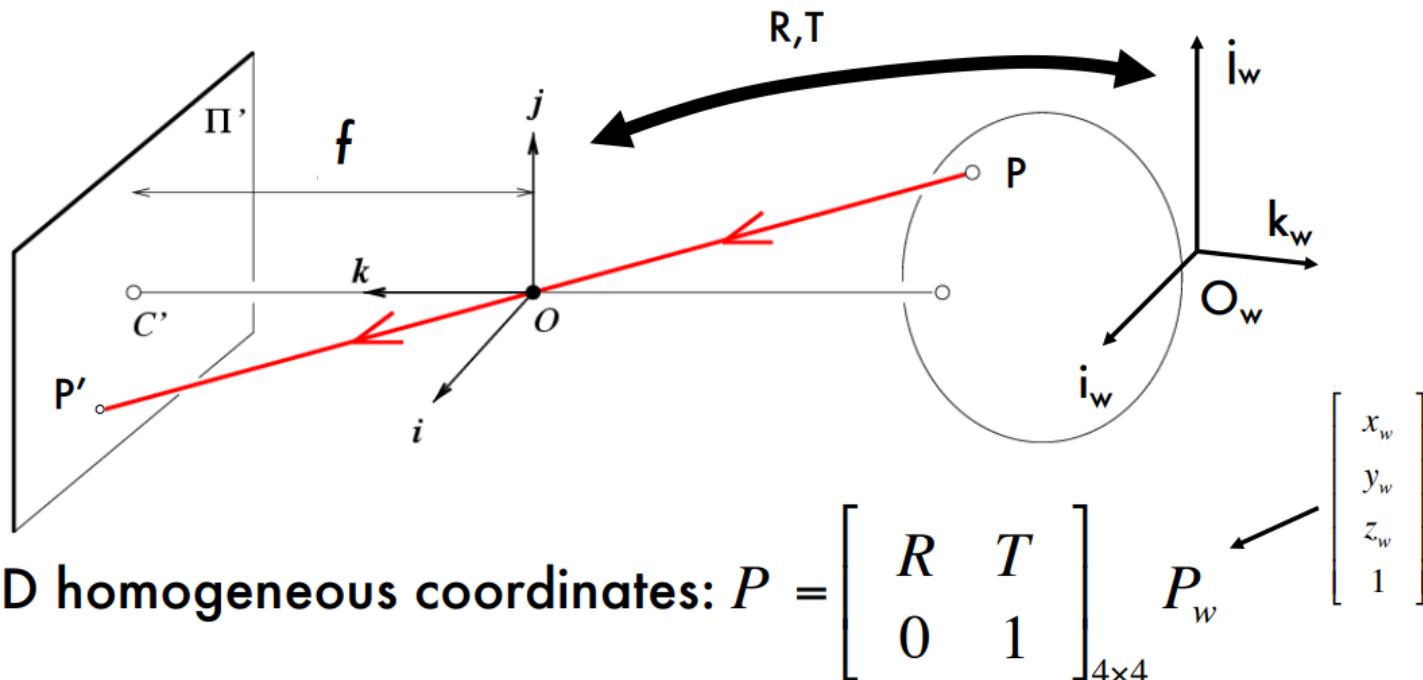
$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & p_x & 0 \\ f & p_y & 0 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$x = K[I|0]X_{cam}$$

$$K = \begin{bmatrix} f & p_x \\ f & p_y \\ 1 & 0 \end{bmatrix}$$

Matrice de calibrage

- Paramètres extrinsèques



$$P' = K \begin{bmatrix} I & 0 \end{bmatrix} P = K \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}_{4 \times 4} P_w = K \boxed{\begin{bmatrix} R & T \end{bmatrix}} P_w$$

Internal parameters External parameters

M

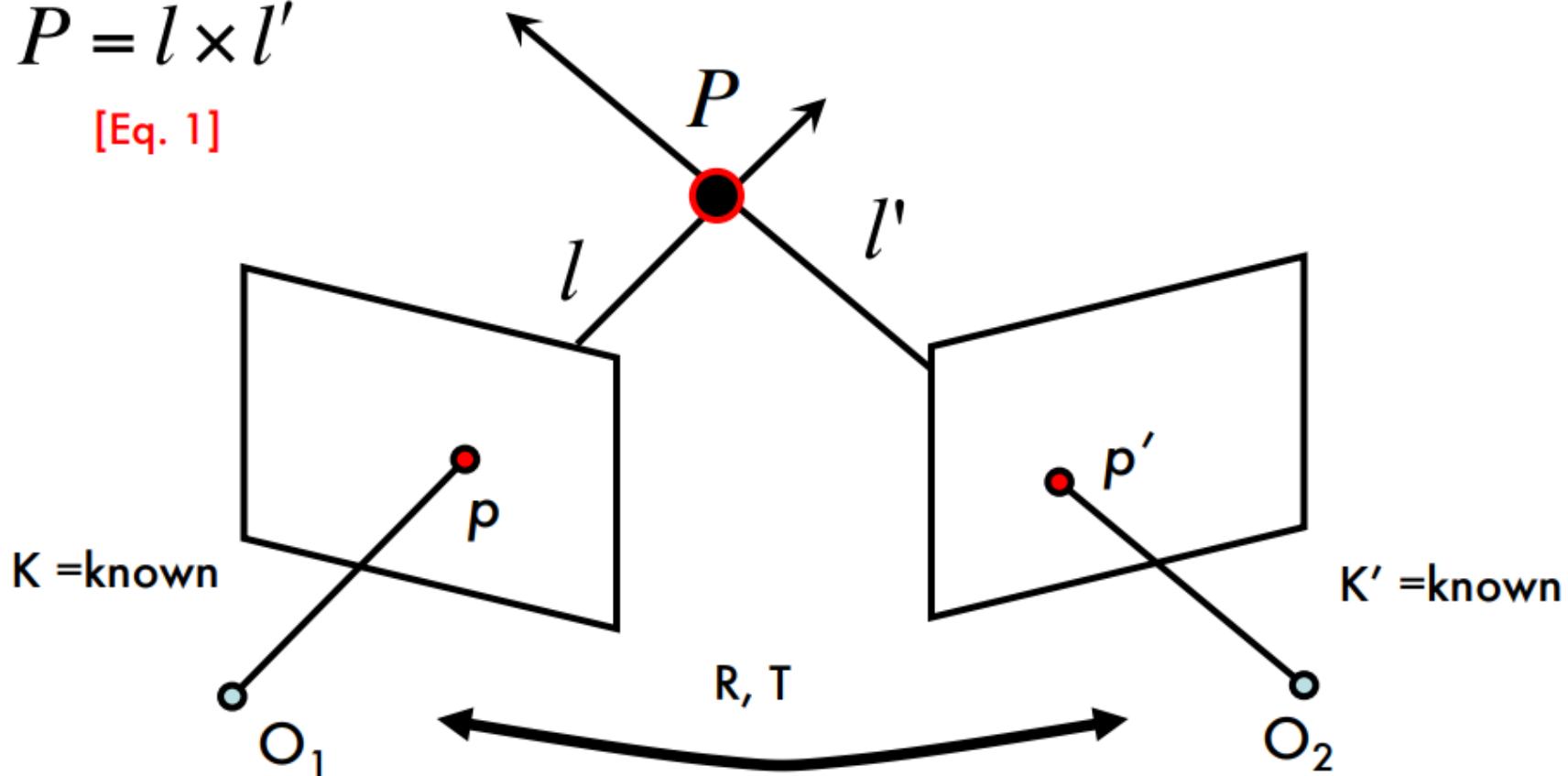
Partie I : Perception 3D

IV. Stéréovision

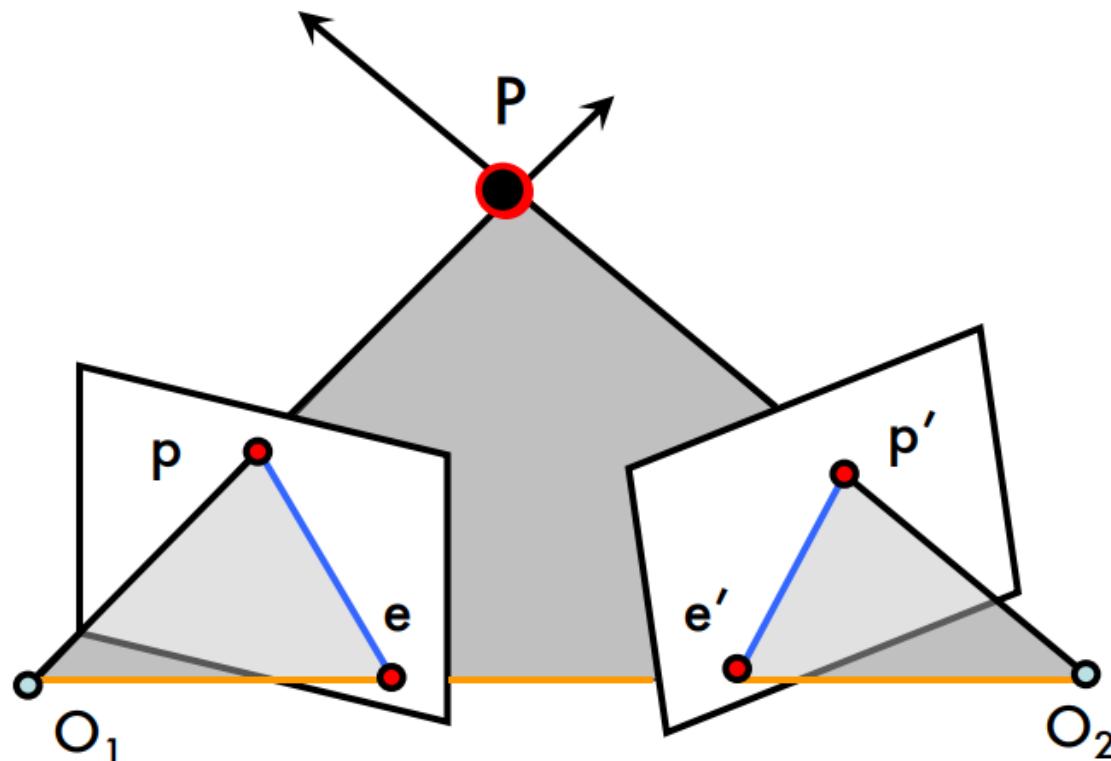
- Triangulation

$$P = l \times l'$$

[Eq. 1]

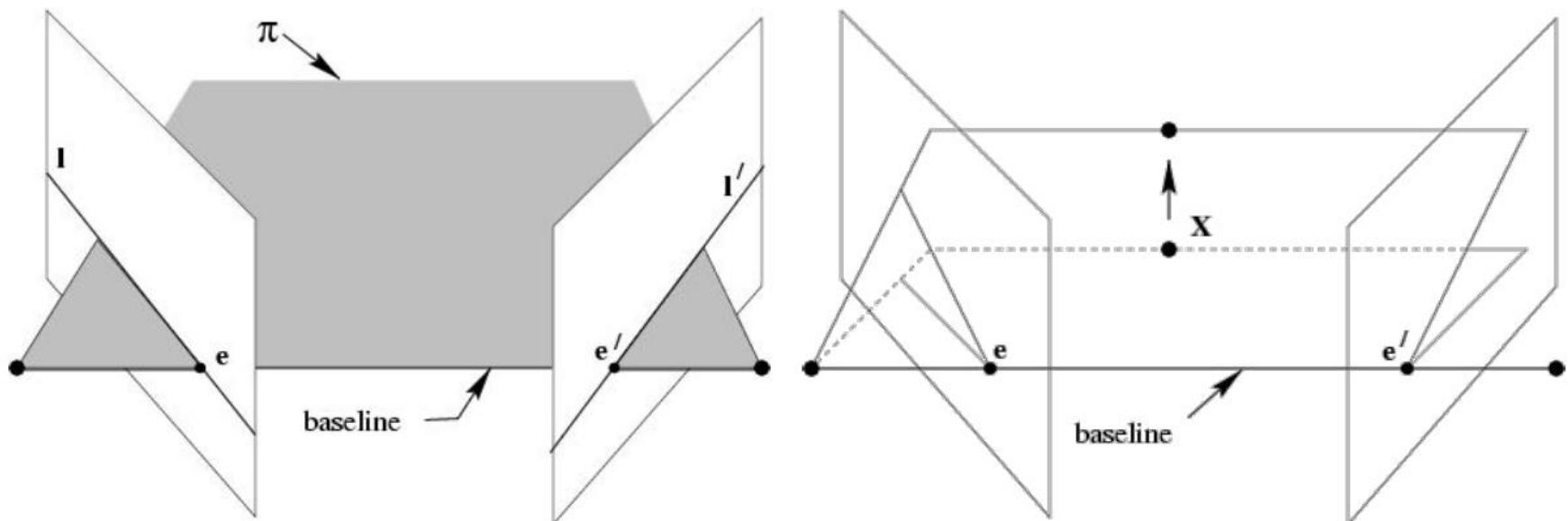


- Géométrie épipolaire



- Epipolar Plane
- Baseline
- Epipolar Lines
- Epipoles e, e'
 - = intersections of baseline with image planes
 - = projections of the other camera center

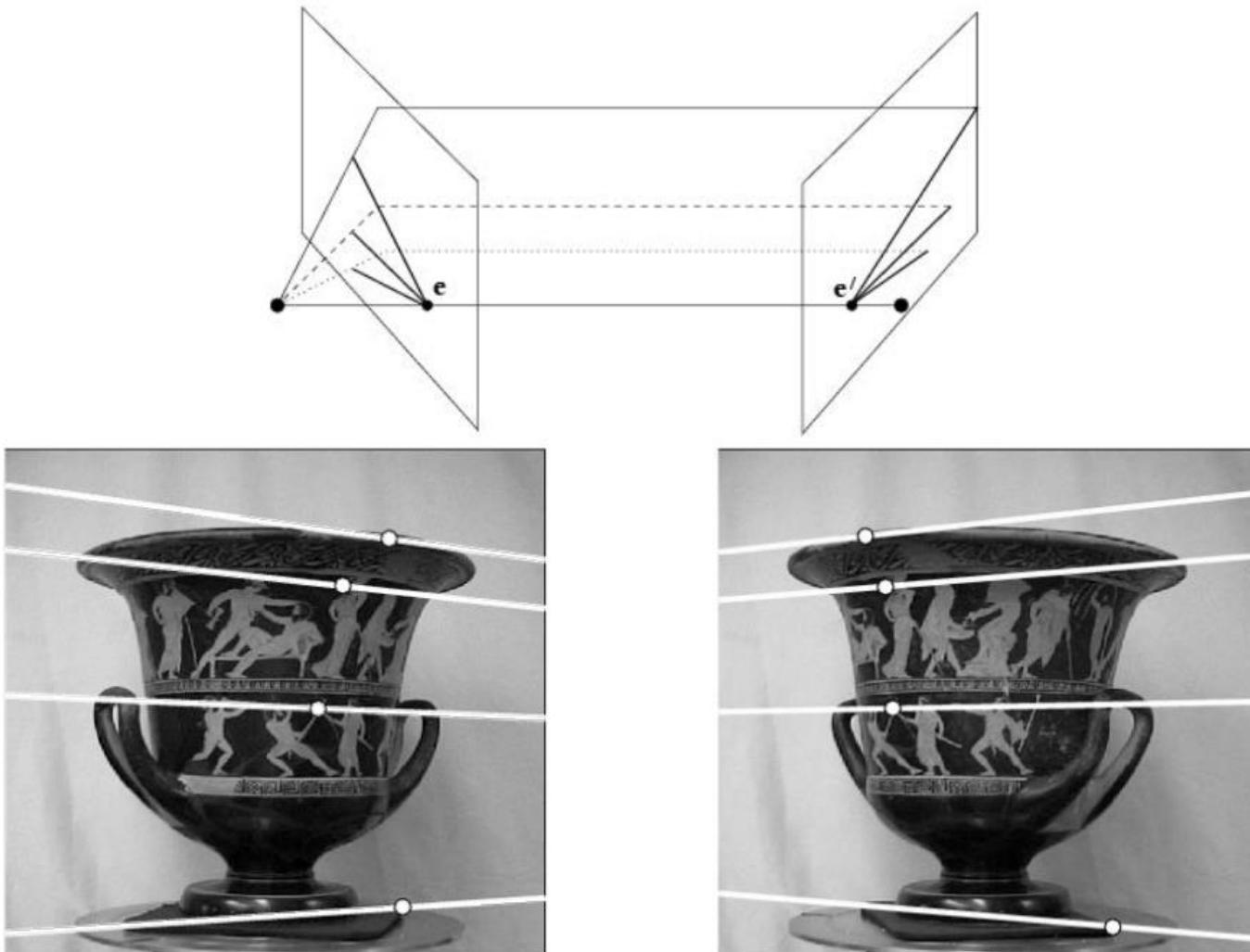
- Géométrie épipolaire



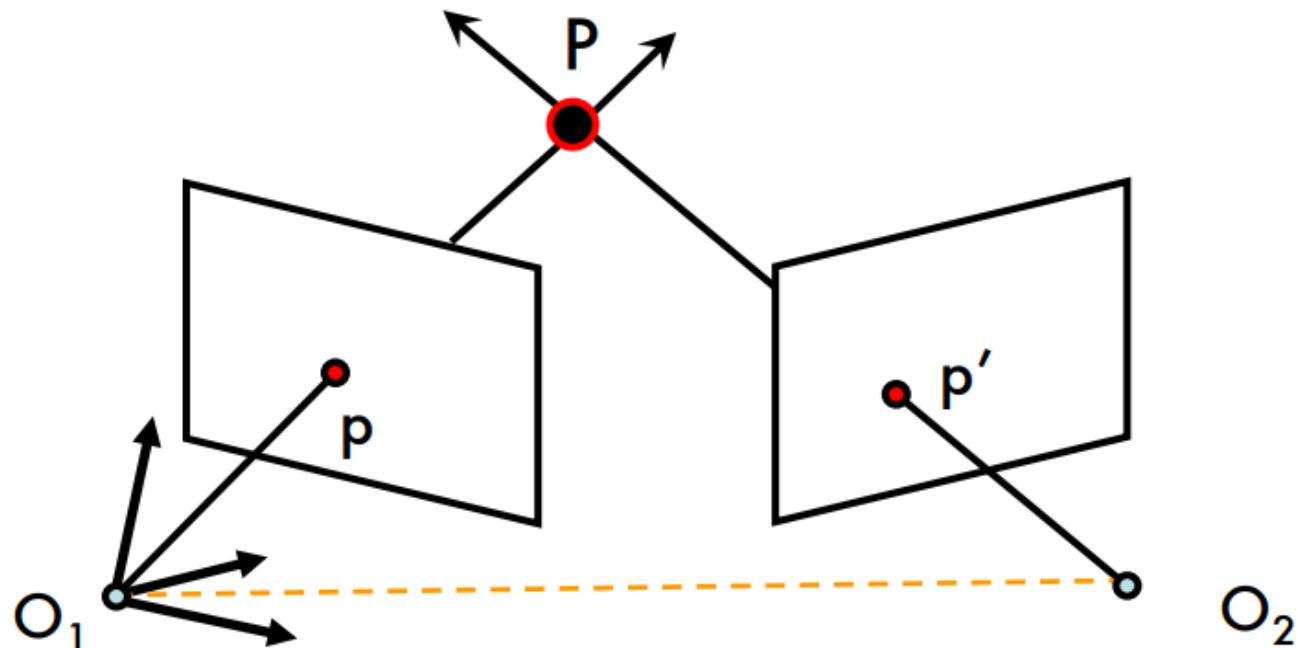
Tous les points sur π se projettent
sur l et l'

Famille de plans π et lignes l et l'
Se coupent en e et e'

- Lignes épipolaire



- Algèbre: Matrice Essentielle



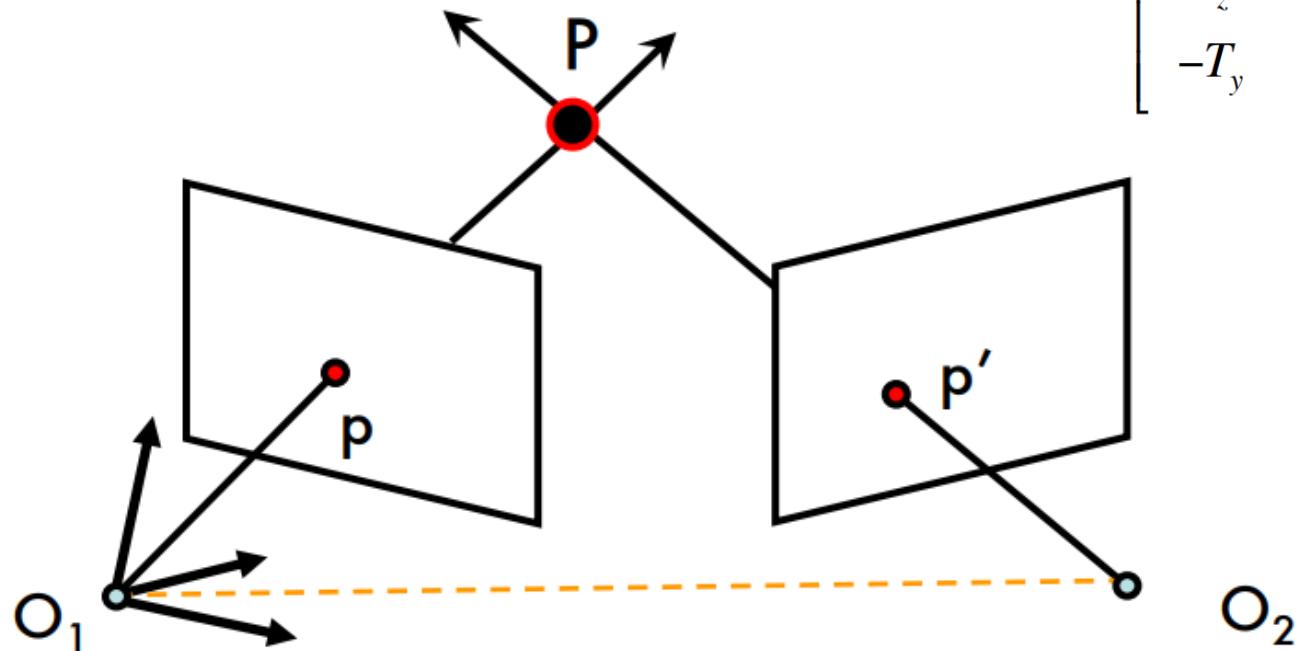
$$p^T E p' = 0$$

E = Essential Matrix

$$E = [T_x] \cdot R$$

- Algèbre: Matrice essentielle E

$$E = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} R$$

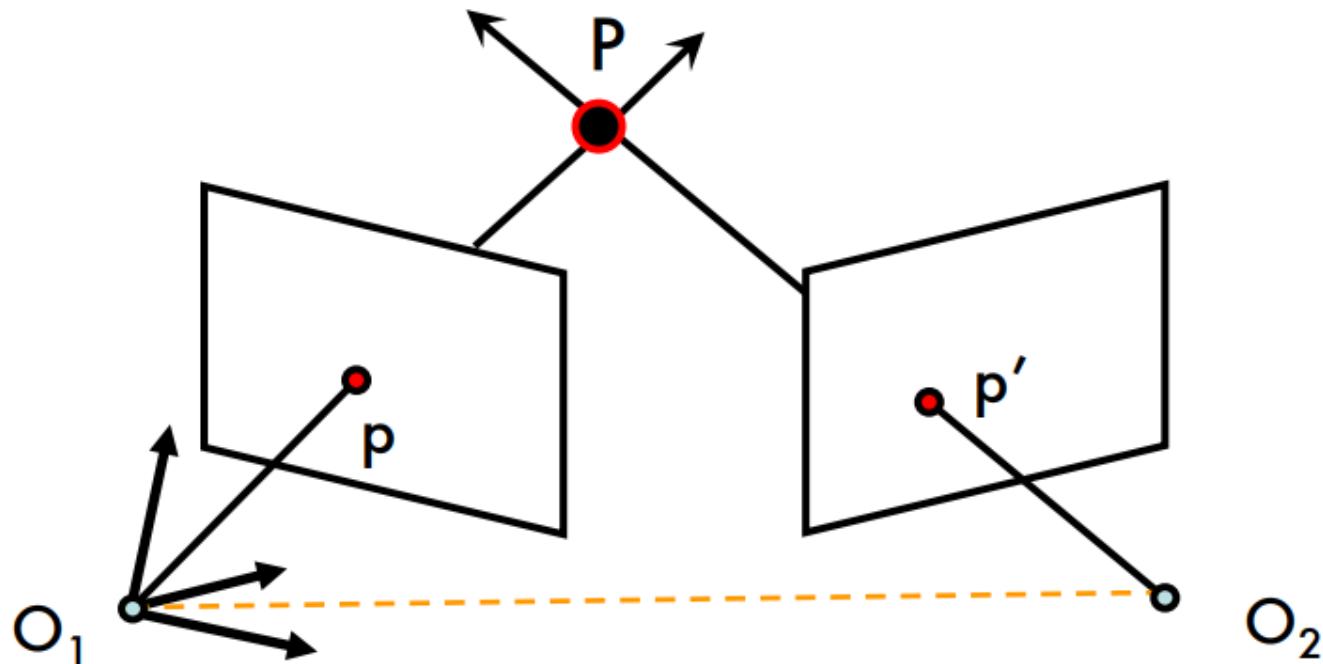


$$p^T E p' = 0$$

E = Essential Matrix

$$E = [T_x] \cdot R$$

- Algèbre: Matrice fondamentale F

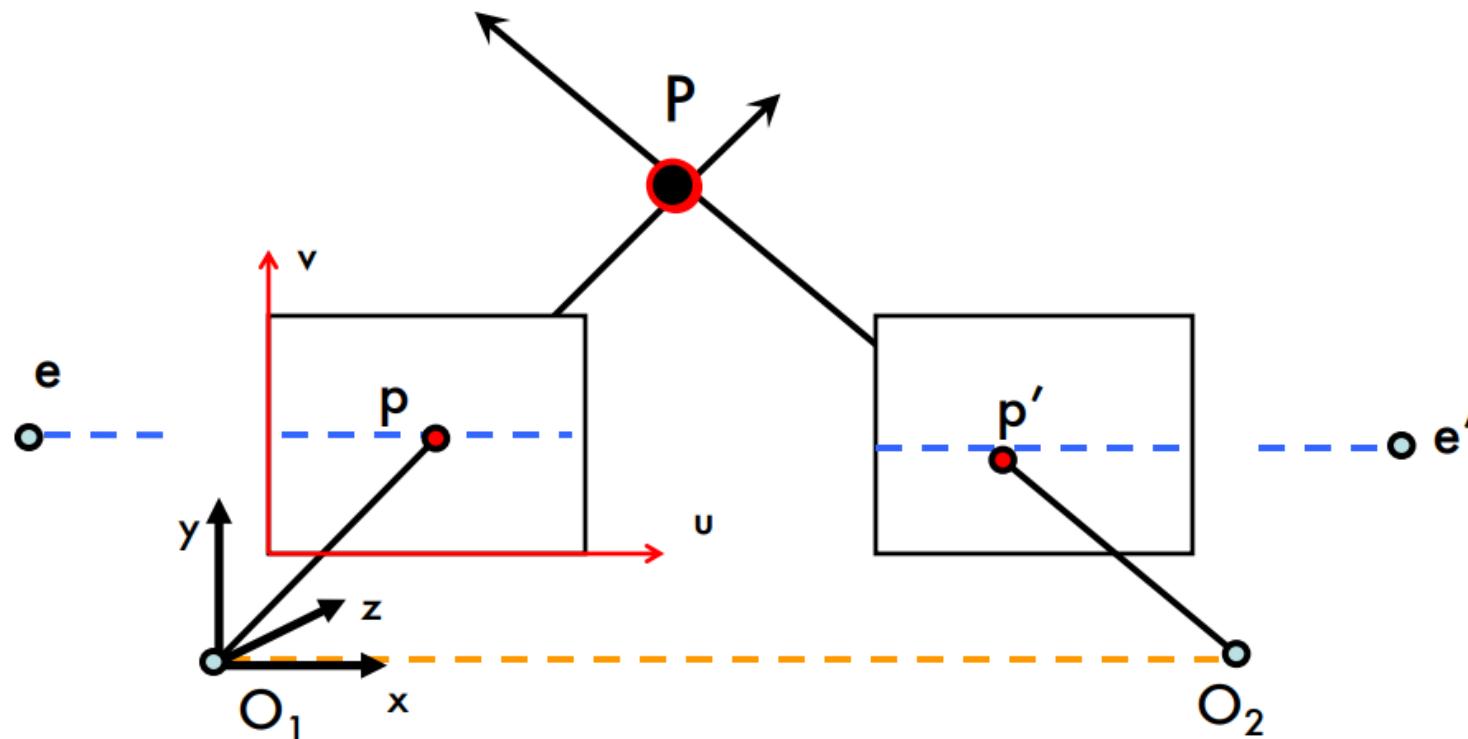


$$p^T F p' = 0$$

$$F = K^{-T} \cdot [T_x] \cdot R \cdot K'^{-1}$$

F = Fundamental Matrix

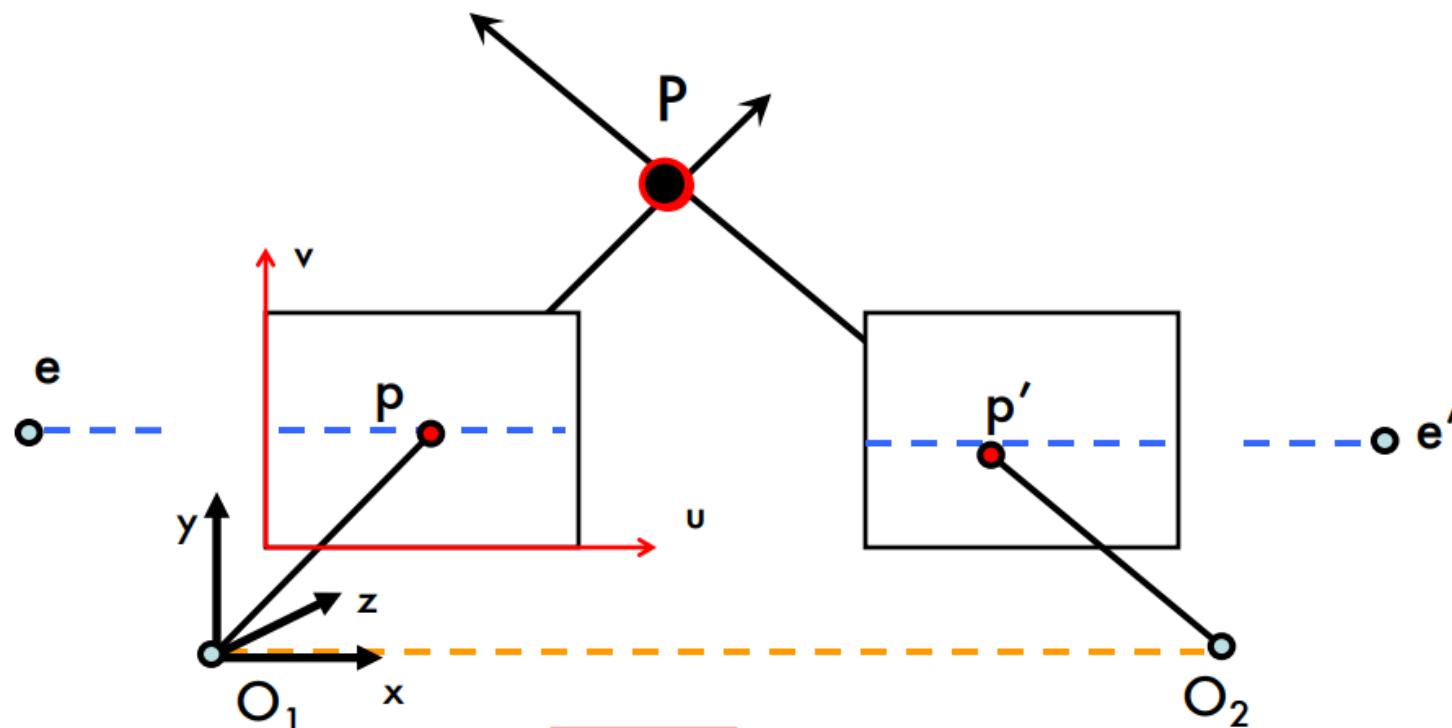
- Plans image parallèles



- Epipolar lines are horizontal
- Epipoles go to infinity
- v -coordinates are equal

$$p = \begin{bmatrix} p_u \\ p_v \\ 1 \end{bmatrix} \quad p' = \begin{bmatrix} p'_u \\ p'_v \\ 1 \end{bmatrix}$$

- Plans image parallèles



$K_1 = K_2 = \text{known}$
 x parallel to O_1O_2

$$E = ?$$

Hint :
 $R = I$ $T = (T, 0, 0)$

- Plans image parallèles

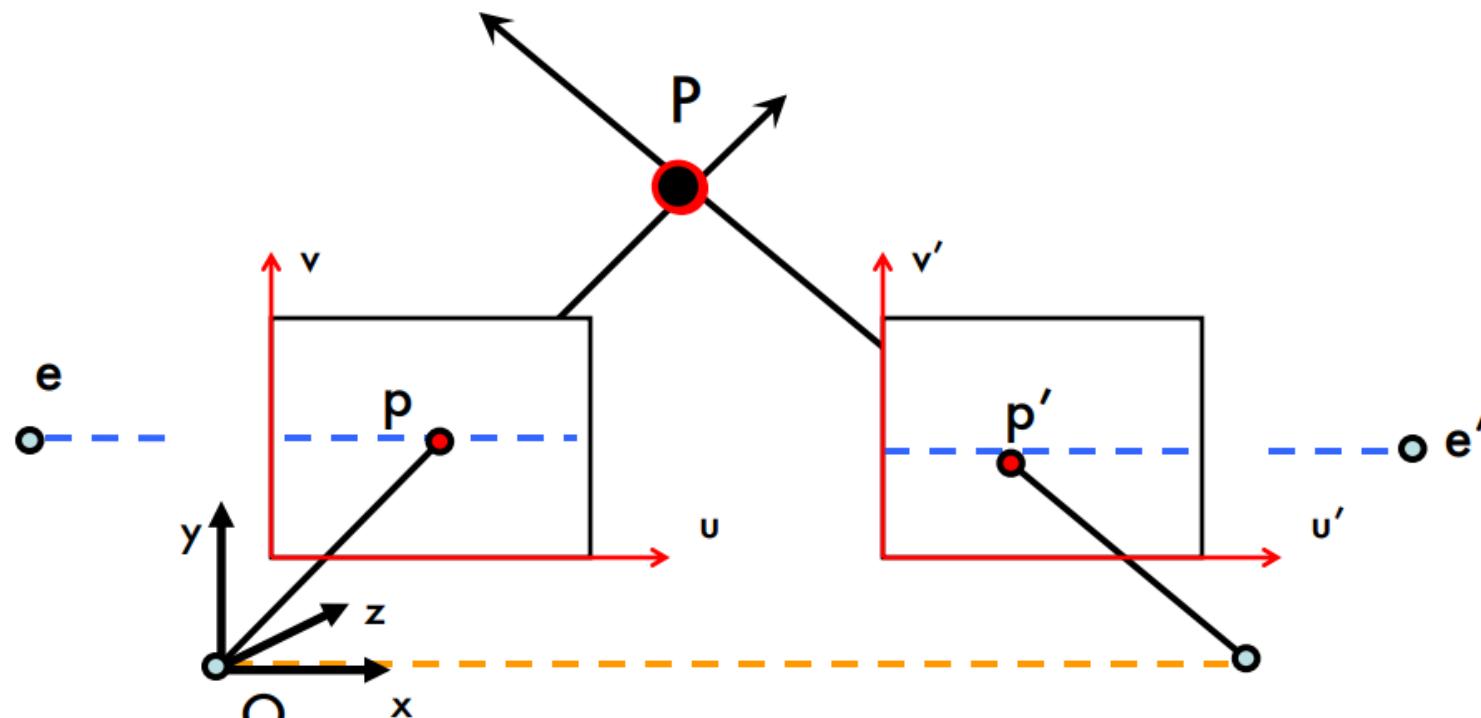
$$\mathbf{E} = [\mathbf{T}_x] \cdot \mathbf{R}$$

$$\mathbf{E} = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix}$$

$$\mathbf{T} = [T \ 0 \ 0]$$

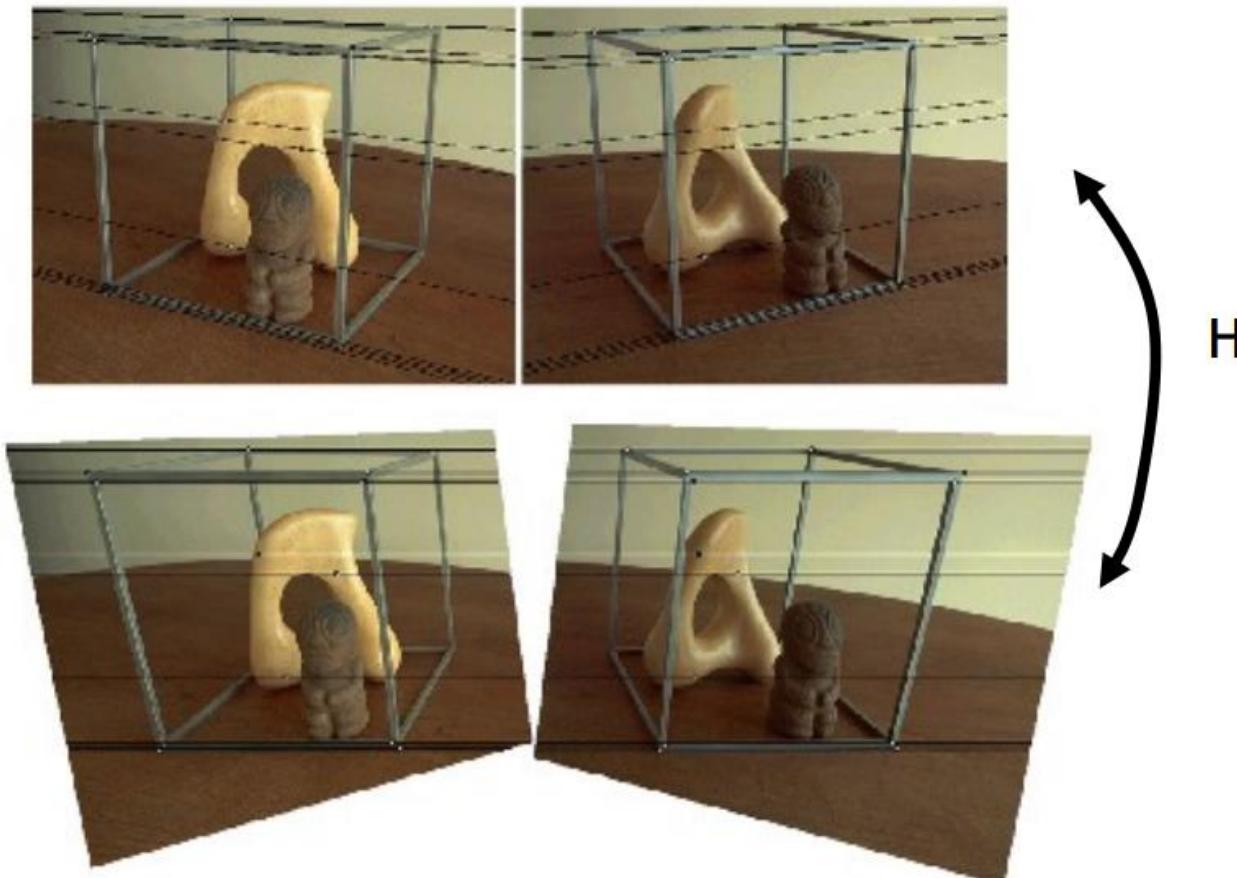
$$\mathbf{R} = \mathbf{I}$$

- Plans image parallèles: relation entre p et p'

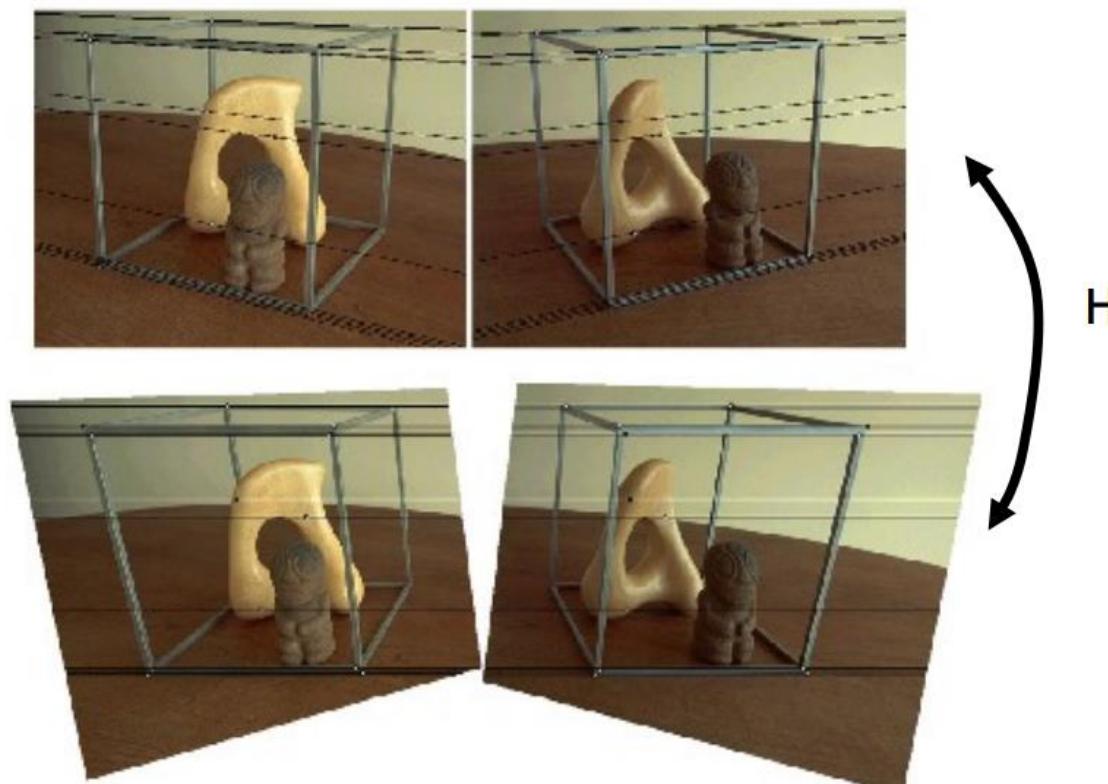


$$p^T \cdot E p' = 0 \Rightarrow (u \ v \ 1) \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -T \\ 0 & T & 0 \end{bmatrix} \begin{pmatrix} u' \\ v' \\ 1 \end{pmatrix} = 0 \Rightarrow (u \ v \ 1) \begin{pmatrix} 0 \\ -T \\ Tv' \end{pmatrix} = 0 \Rightarrow Tv = Tv' \Rightarrow v = v'$$

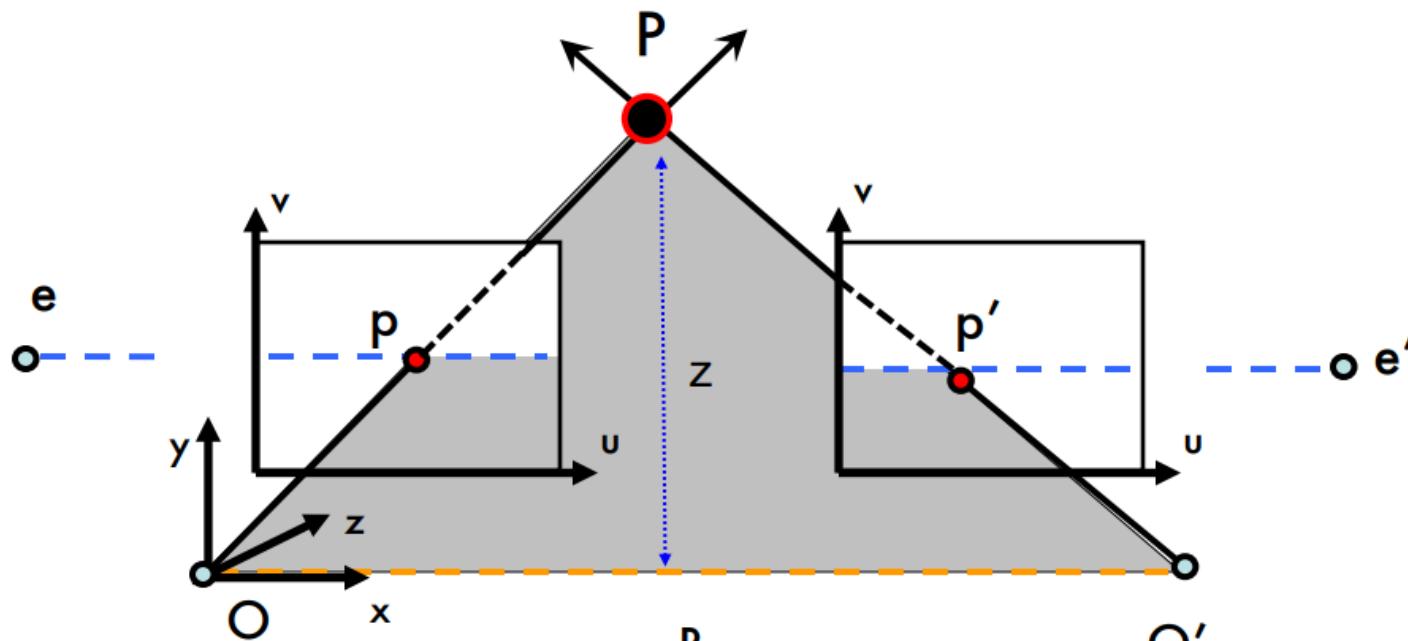
- Rectification:
 - Le fait de rendre deux images parallèles



- Rectification:
 - Le fait de rendre deux images parallèles
 - Rends la triangulation facile
 - Facilite l'appariement



- Disparité

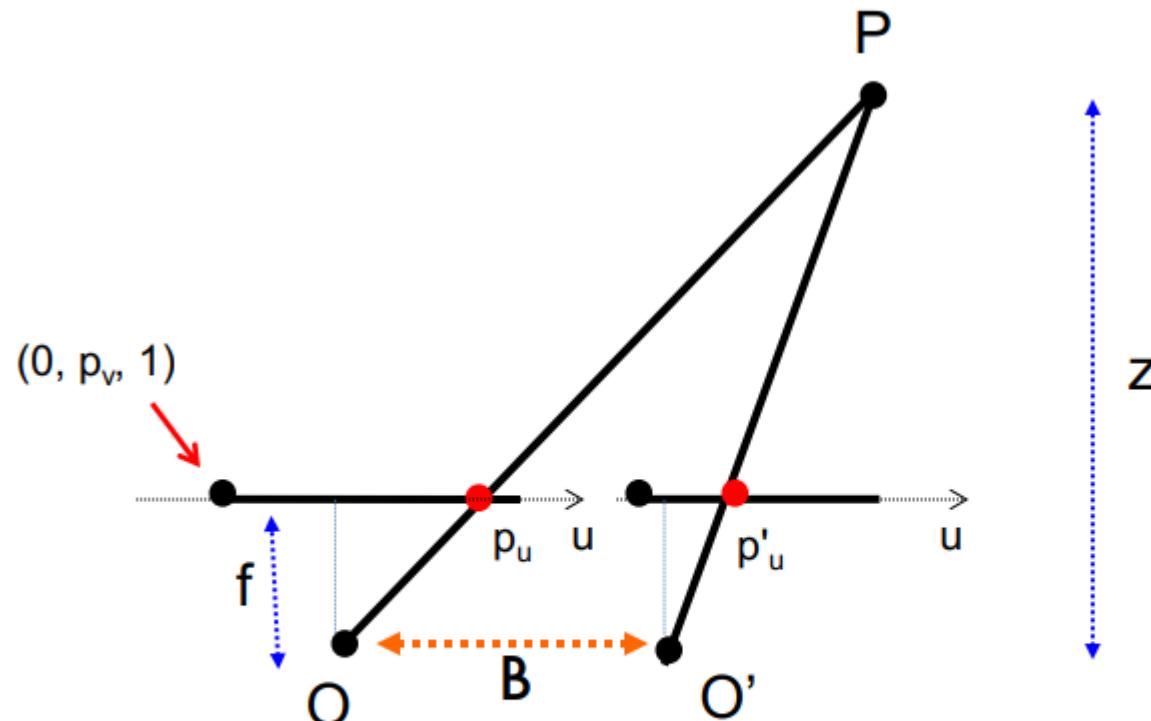


$$p = \begin{bmatrix} p_u \\ p_v \\ 1 \end{bmatrix} \quad p' = \begin{bmatrix} p'_u \\ p'_v \\ 1 \end{bmatrix}$$

$$\text{disparity} = p_u - p'_u \propto \frac{B \cdot f}{z}$$

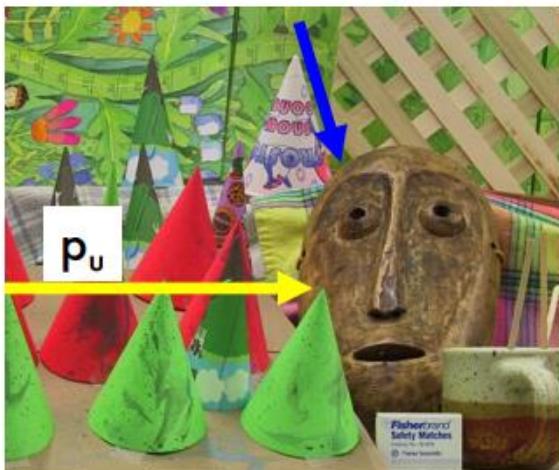
Disparity is inversely proportional to depth z !

- Calcul de la profondeur



$$\text{disparity} = p_u - p'_u \propto \frac{B \cdot f}{z}$$

- Image de disparité/carte de profondeur:



$$p_u - p'_u \propto \frac{B \cdot f}{z}$$

[Eq. 1]

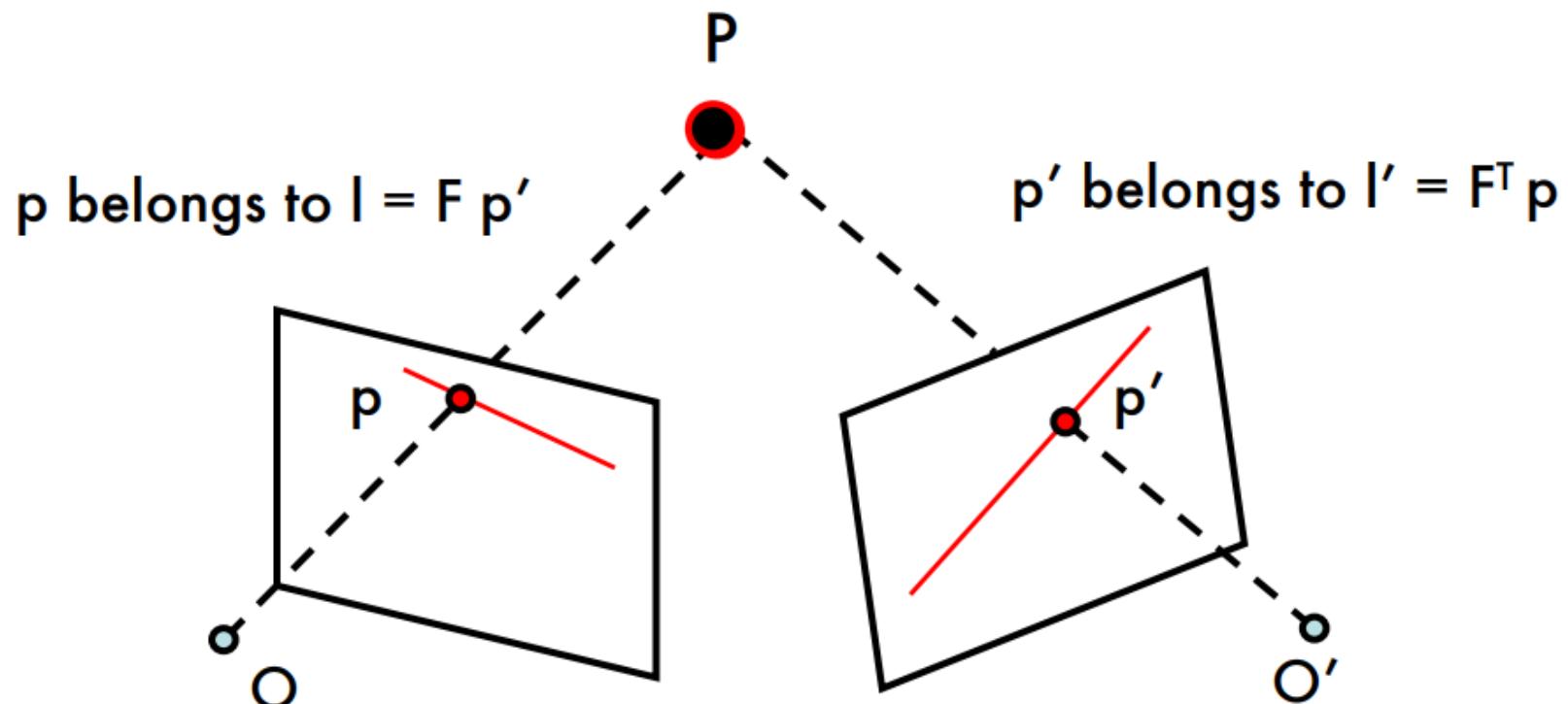
Stereo pair



Disparity map / depth map

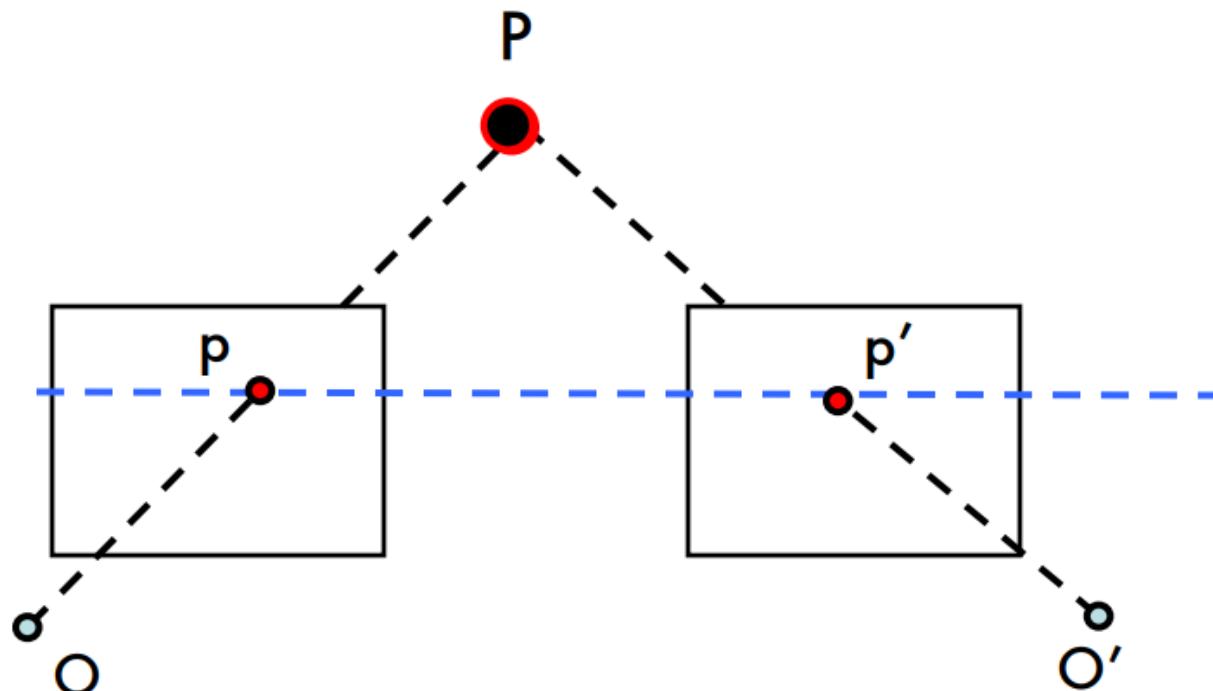
- Appariement de points:

- Trouver les coordonnées en pixel d'un point 3D dans le plan image des deux caméras



- Appariement de points:

- Trouver les coordonnées en pixel d'un point 3D dans le plan image des deux caméras
- Les points images sont sur la même ligne lorsque les images sont rectifiées



- Méthodes de corrélation:

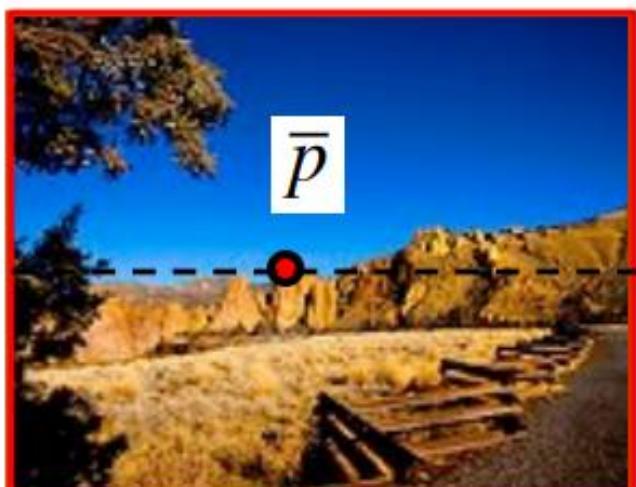


image 1

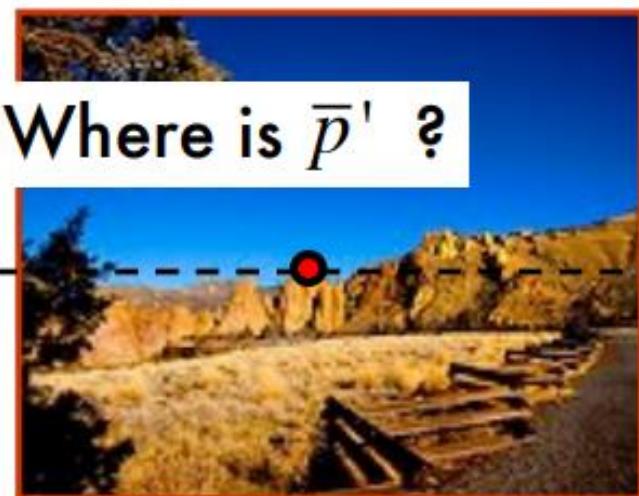
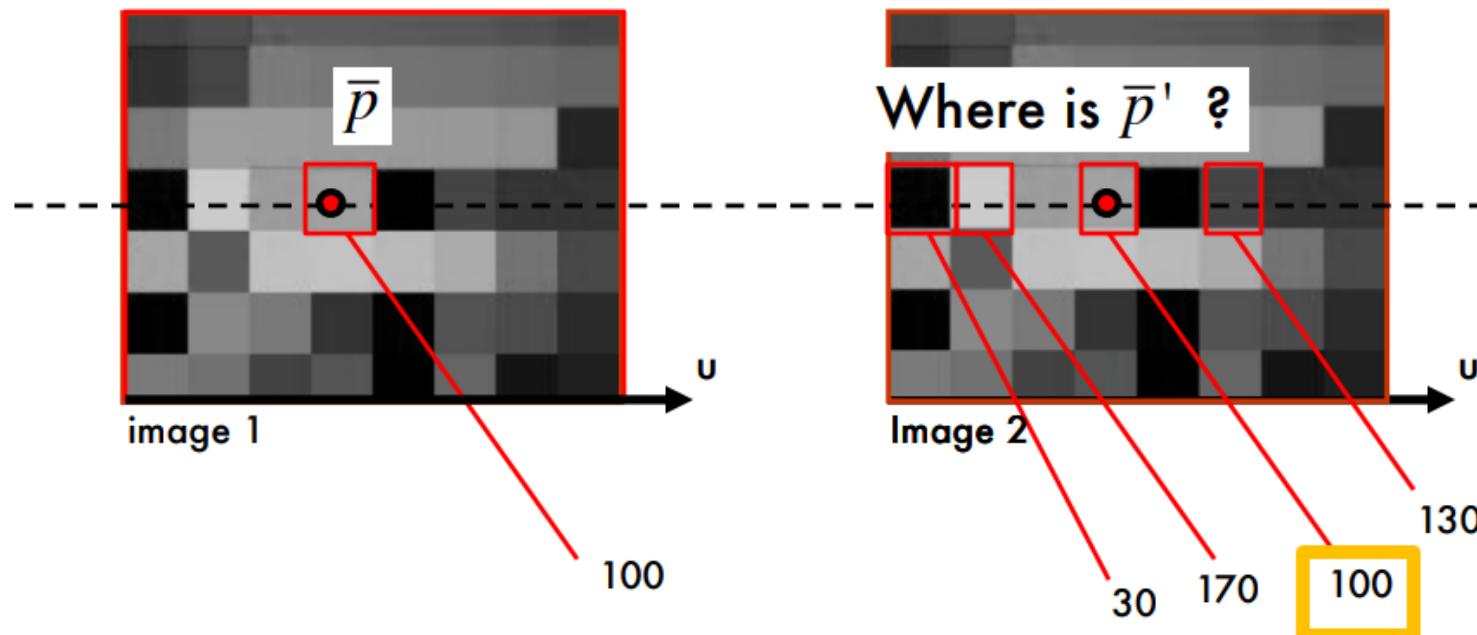
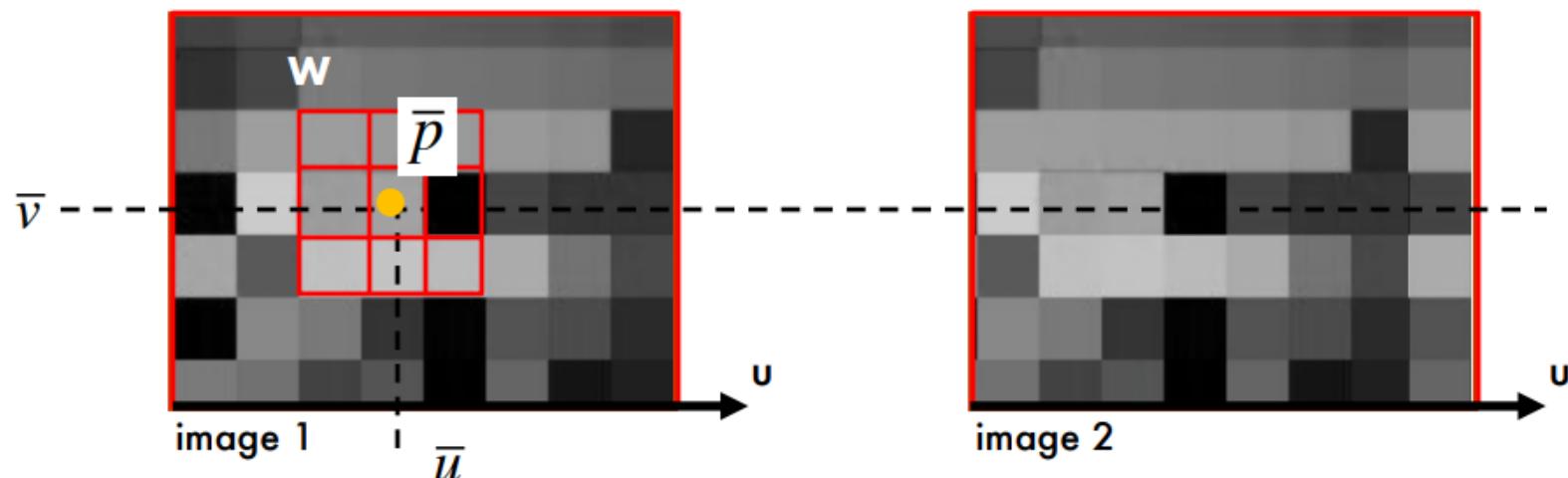


Image 2

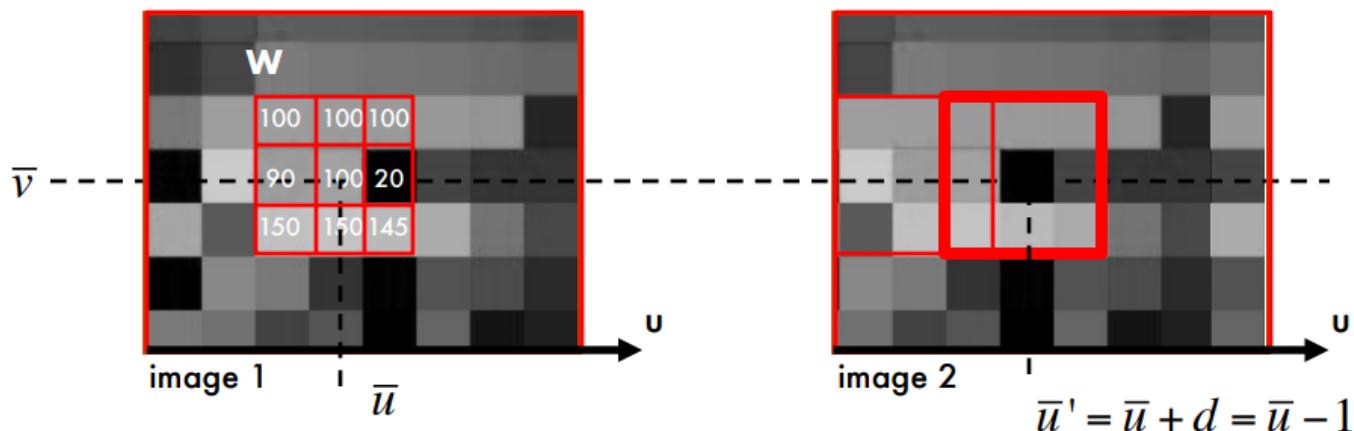
- Méthodes de corrélation:
 - Recherche de valeur identique



- Méthodes de corrélation:
 - Sélection d'une fenêtre



- Méthodes de corrélation:
 - Sélection d'une fenêtre



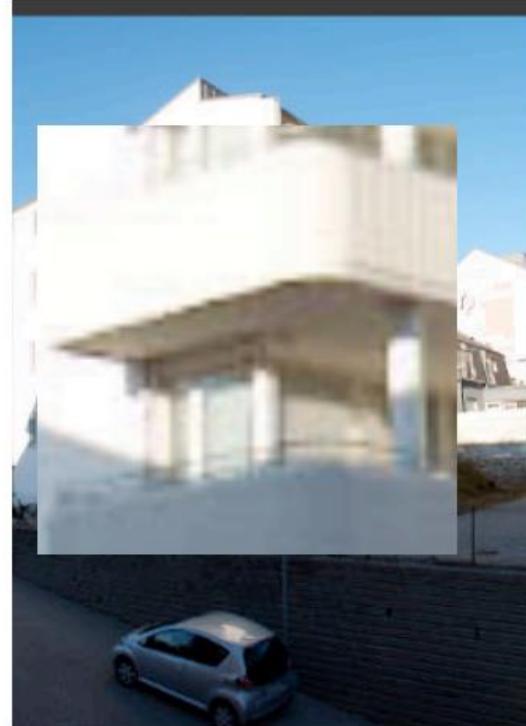
Example: **W** is a 3×3 window in red

w is a 9×1 vector

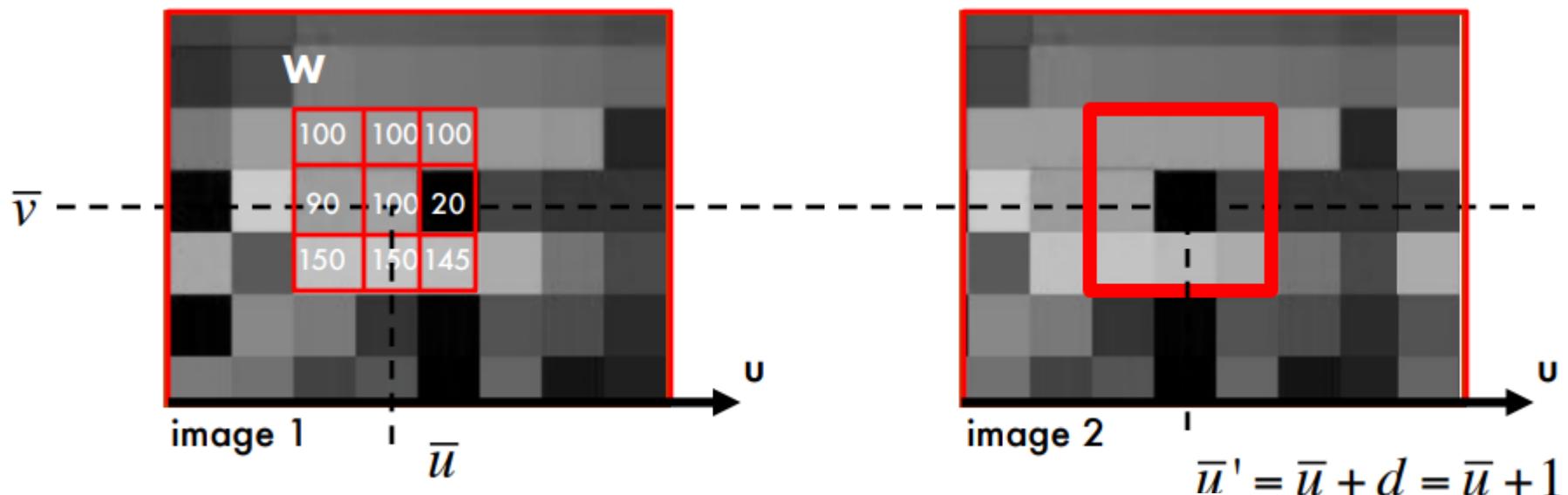
$$\mathbf{w} = [100, 100, 100, 90, 100, 20, 150, 150, 145]^T$$

- Pick up a window **W** around $\bar{p} = (\bar{u}, \bar{v})$
- Build vector **w**
- Slide the window **W** along $v = \bar{V}$ in image 2 and compute $\mathbf{w}'(u)$ for each u
- Compute the dot product $\mathbf{w}^T \mathbf{w}'(u)$ for each u and retain the max value

- Méthodes de corrélation:
 - Sélection d'une fenêtre
 - Sensible aux différences d'exposition



- Méthodes de corrélation:
 - Normalized cross-correlation



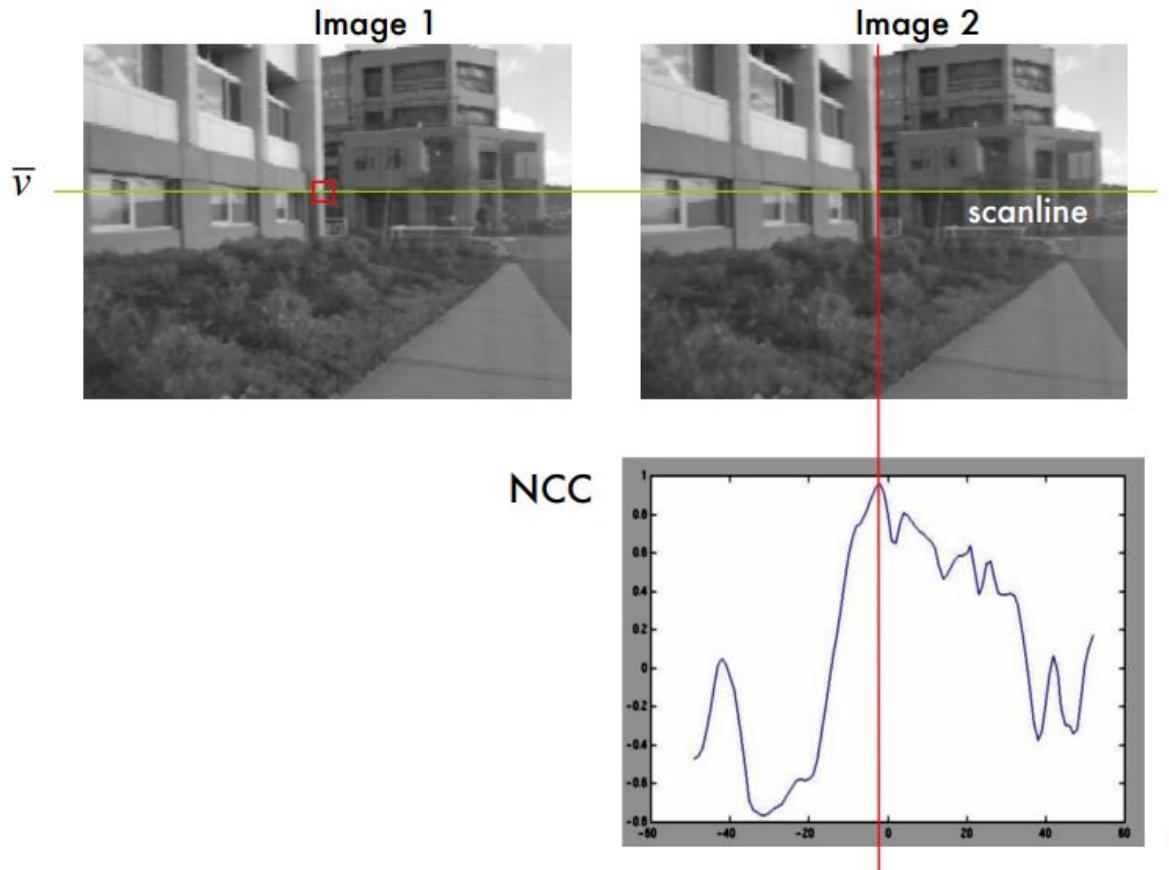
Find u that maximizes:

$$\frac{(w - \bar{w})^T (w'(u) - \bar{w}')}{\|(w - \bar{w})\| \| (w'(u) - \bar{w}') \|}$$

\bar{w} = mean value within **W**
located at u^* in image 1

$\bar{w}'(u)$ = mean value within **W**
located at u in image 2

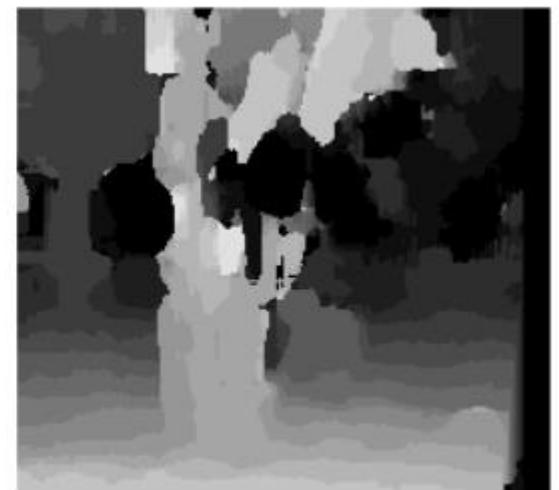
- Méthodes de corrélation:
 - Normalized cross-correlation



- NCC taille de la fenêtre de corrélation:
 - Petite fenêtre:
 - Plus de détail, plus de bruit
 - Grande fenêtre:
 - Peu de bruit, lissage de la disparité



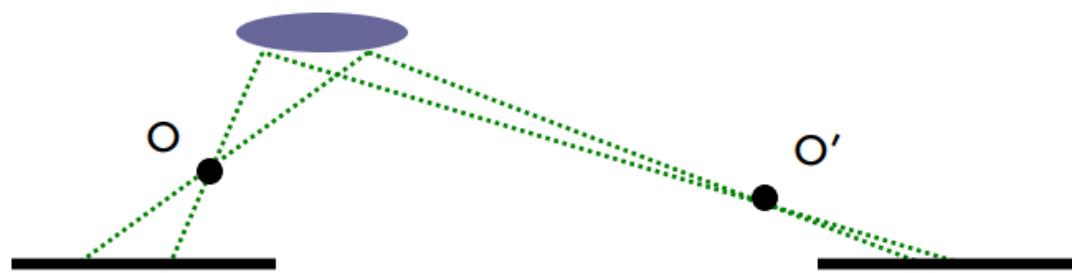
Window size = 3



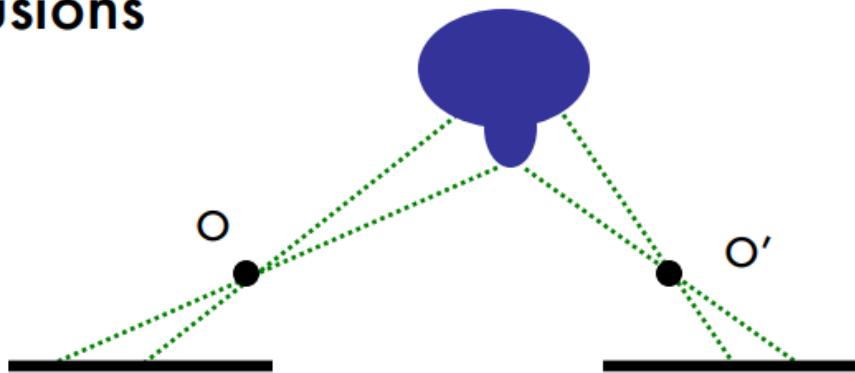
Window size = 20

- NCC problèmes:

- Fore shortening effect



- Occlusions



- NCC problèmes:
 - Régions homogènes



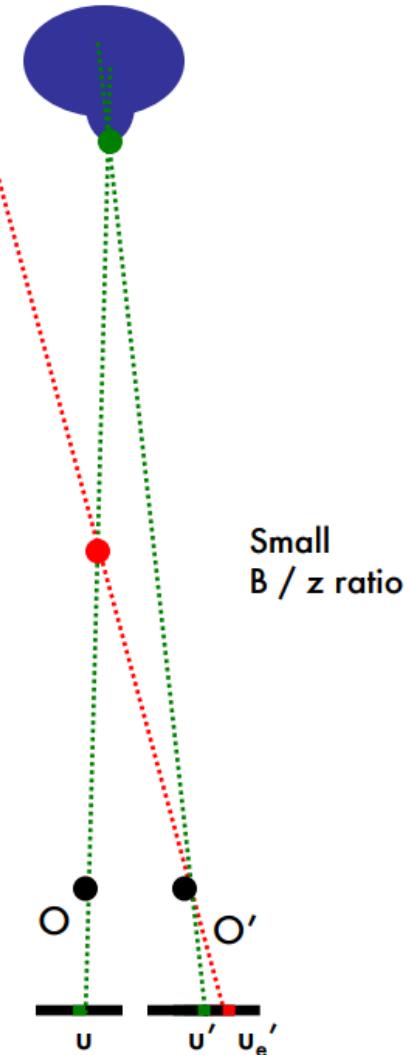
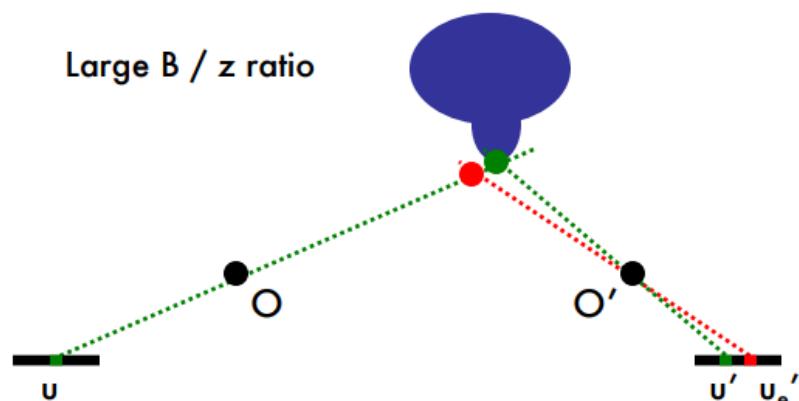
mismatch

- NCC problèmes:
 - Patterns répétitifs

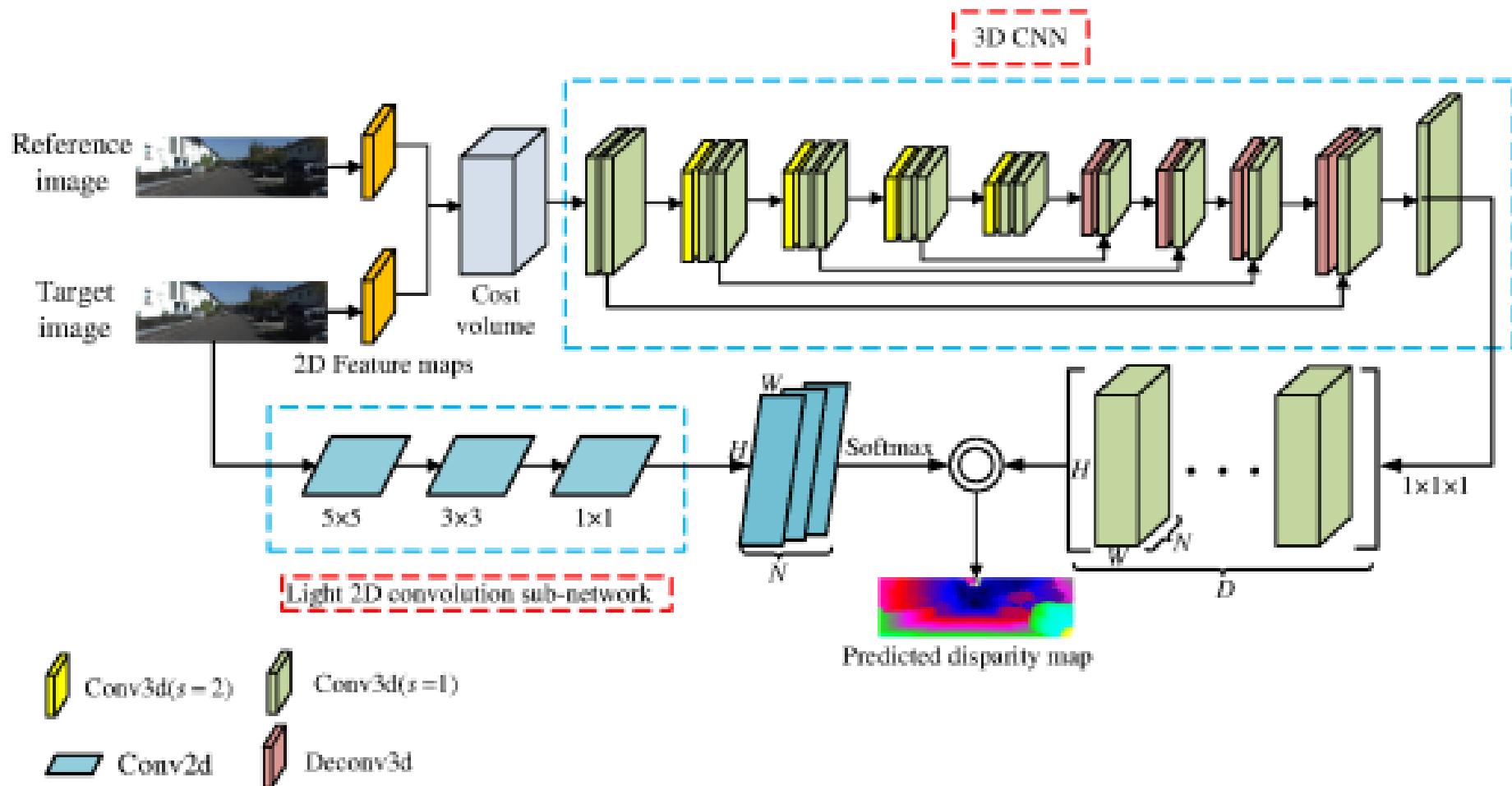


- NCC Influence de la Baseline

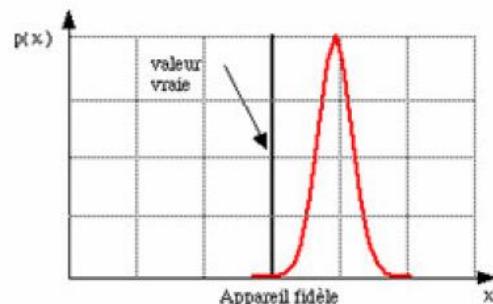
- b/z petit
 - réduction des effets de rapprochement
 - Augmentation des erreurs



- Réseaux de neurones convolutifs

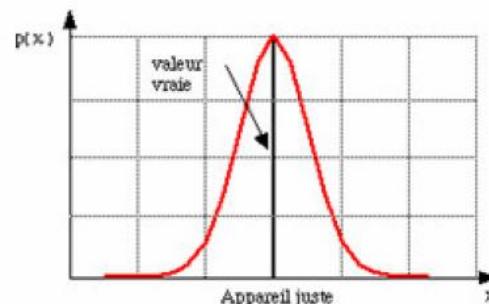


- Etat de l'art
 - <http://vision.middlebury.edu/stereo/>
- Evaluation



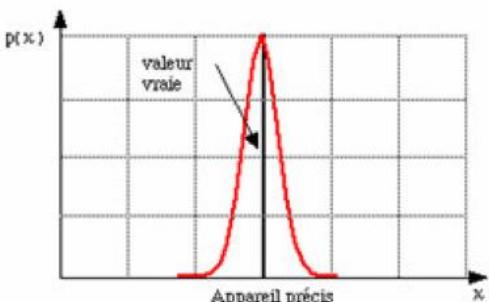
Fidélité : analysée par l'écart-type

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$



Justesse : analysée par la moyenne

$$\bar{x} = E(X) = \frac{1}{n} \sum_{i=1}^n x_i$$



Précision : analysée par la valeur RMS

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - x_{vrai})^2}$$

Merci pour votre attention

Questions?