

# Modélisation de l'environnement par vision en robotique mobile

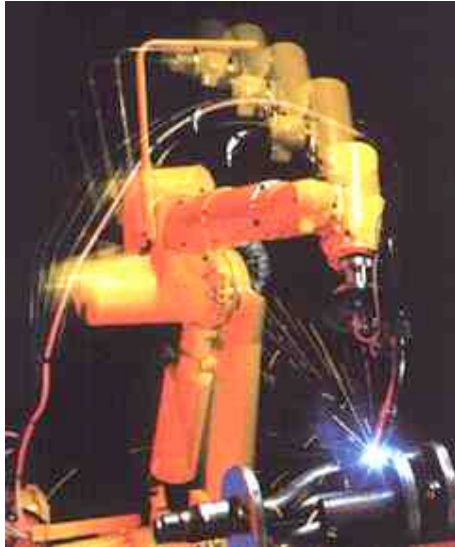
Simon Lacroix

Robotics and Interactions group  
LAAS/CNRS, Toulouse

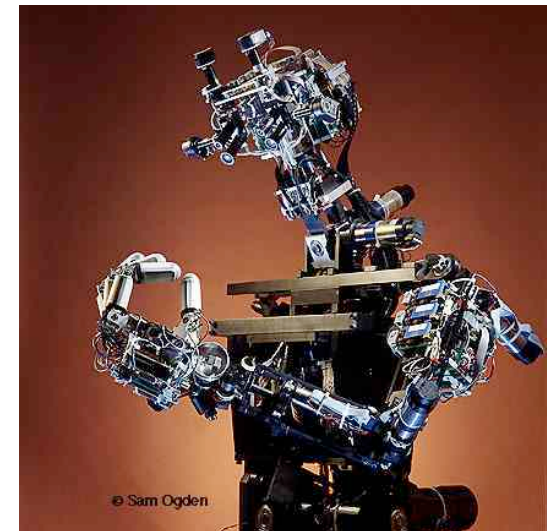
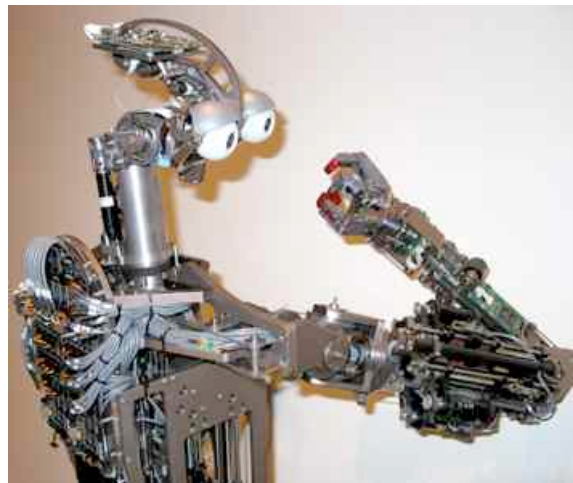
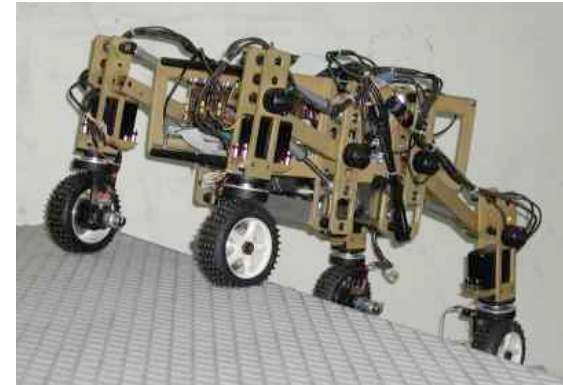


# Robots...

## Industriel



## Dans les laboratoires



...robots...





...et robots.



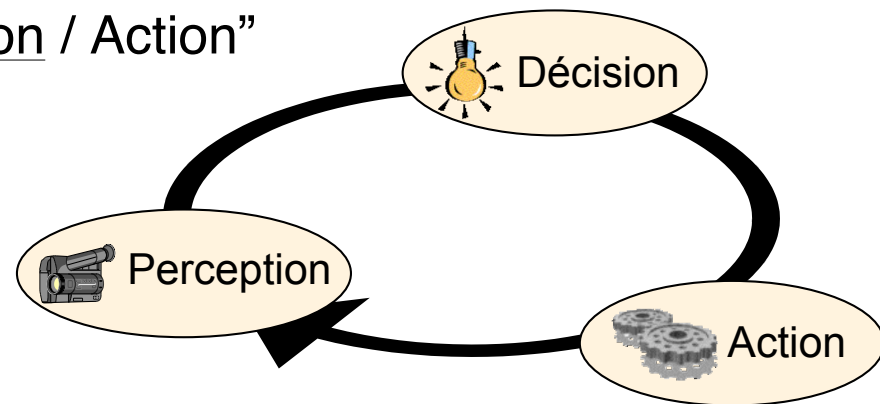
# De l'automatique à l'autonomie

- Automatique :
  - Tâche bien définie (*“réguler variable”, “suivre trajectoire”...*)
  - Lien “direct” entre perception et action
- Autonomie :
  - Tâche plus générale (*“atteindre position”, “surveiller zone ”...*)
  - Mise en oeuvre de *processus décisionnels*

! Boucle “perception / Décision / Action”

Plus :

- *Intégration* des fonctions
- Apprentissage
- Interaction avec opérateurs
- Interactions avec autres robots
- ...



# Un exemple : robotique d'exploration planétaire

Contraintes sur les communications  
Informations limitées au sol  
! autonomie indispensable



Tâche fondamentale : navigation autonome : *"Atteindre position"*

Première approche : «"réactive"»

Seconde approche : «"délibérative"»

# Perception pour la robotique

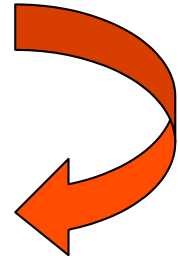
Perception :

« Acquisition et représentation des informations sur l'environnement et le robot lui-même »

Besoins :

- Planification d'actions (déplacements)
- Planification de tâches (et notamment tâches de perception)
- Localisation du robot
- Exécution d'action (de déplacements)
- Interaction avec l'opérateur, reconnaissance *d'objets*

Aujourd'hui : vision pour



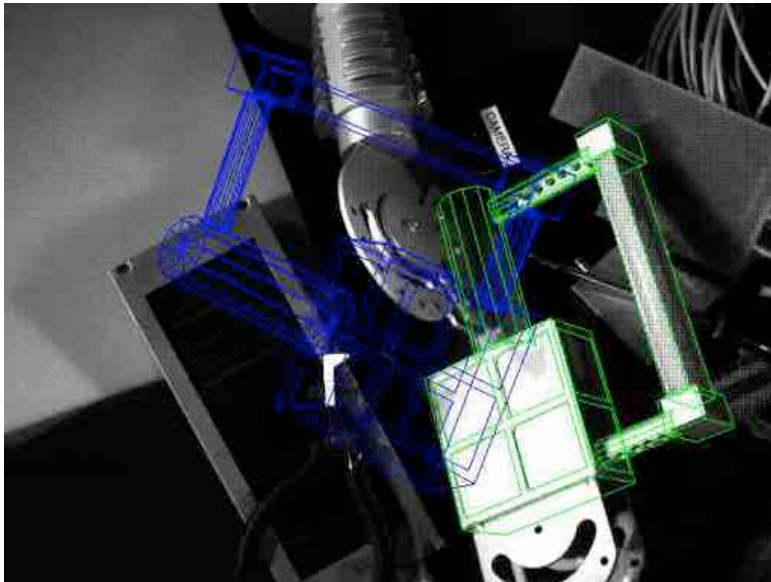
# Perception pour la robotique

Perception :

« Acquisition et représentation des informations sur l'environnement et le robot lui-même »

Besoins :

- Planification d'actions (déplacements)
- Planification de tâches (et notamment tâches de perception)
- Localisation du robot
- Exécution d'action (de déplacements)
- Interaction avec l'opérateur, reconnaissance *d'objets*





# Fil conducteur

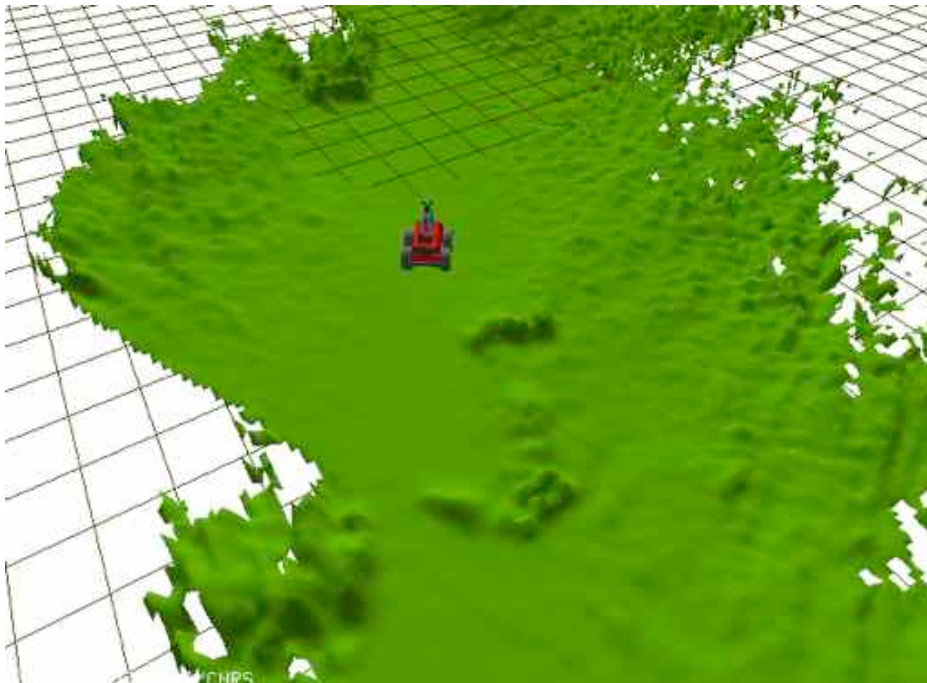
- Introduction
- Modélisation d'environnement par vision
- Localisation par vision
- Modélisation et localisation simultanées

# Fil conducteur

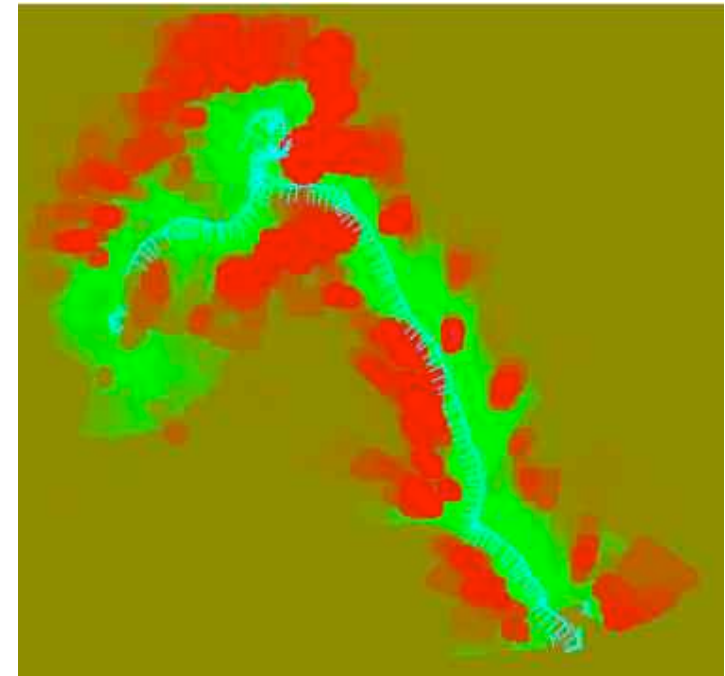
- Introduction
- Modélisation d'environnement par vision
  - Quels modèles sont nécessaires ?
  - Pourquoi la vision ?
  - Stéréovision
  - Exemples d'algorithmes de modélisation
- Localisation par vision
- Modélisation et localisation simultanées

# Modélisation d'environnement

- Pour planifier des mouvements et déplacements  
! évaluation/quantification des actions de déplacement



Modèle numérique de terrain  
 $z = f(x,y)$  sur grille cartésienne



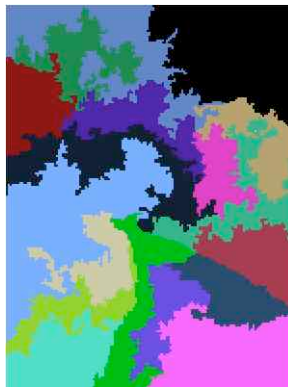
Modèle probabiliste d'obstacles  
 $P(Obst)$  sur grille cartésienne  
(modèle *robocentrique*)

# Modélisation d'environnement

- Pour réaliser des mouvements
  - ! Détection d'éléments sur lesquels s'appuyer (plus ou moins complexes)



Données brutes

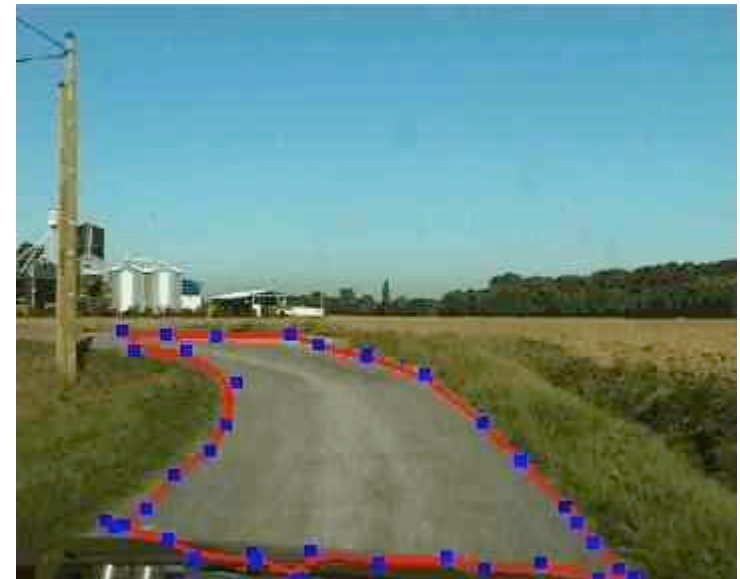


Segmentées



Interprétées

Par classification de couleur/texture



Exploitation pour l'asservissement  
de déplacements

# Modélisation d'environnement

- Pour réaliser des mouvements
  - ! Détection d'éléments sur lesquels s'asservir (plus ou moins complexes)



Données brutes



Segmentées



Interprétées

Par classification de couleur/texture



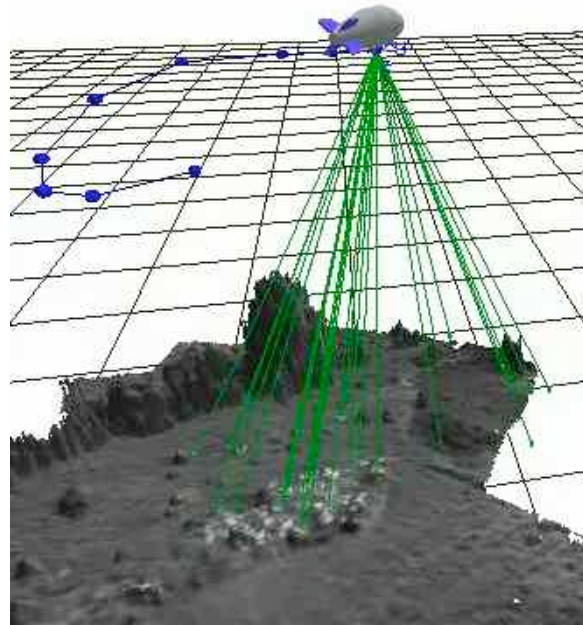
Exploitation pour l'asservissement  
de déplacements



# Modélisation d'environnement

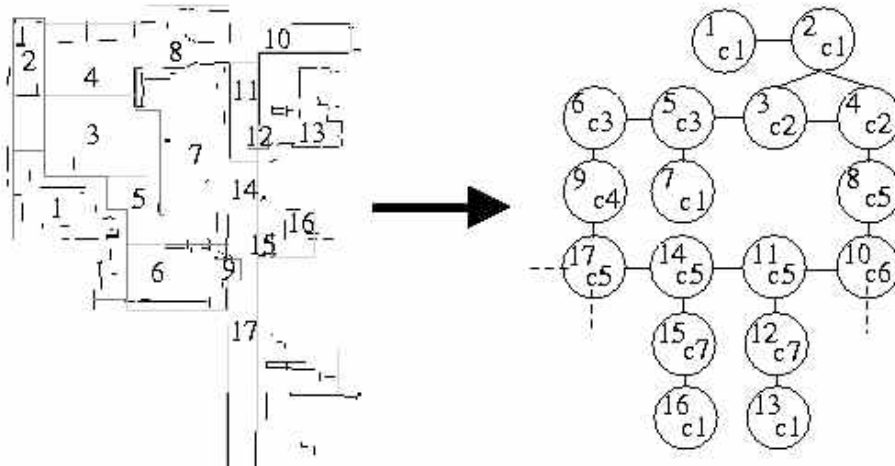
- Pour se localiser
  - ! Missions définies en terme de localisation ("aller à", "explorer/surveiller zone"...)
  - ! Pour assurer la cohérence spatiale des modèles
  - ! Pour assurer l'exécution des déplacements planifiés

Notion *d'amer*



# Modélisation d'environnement

- Au delà de l'autonomie du mouvement, pour planifier / exécuter des *tâches* :
  - ! De perception, d'exploration, de surveillance, dans un contexte multi-robots...



Graphe topologique  
(avec éventuellement  
informations sémantiques)

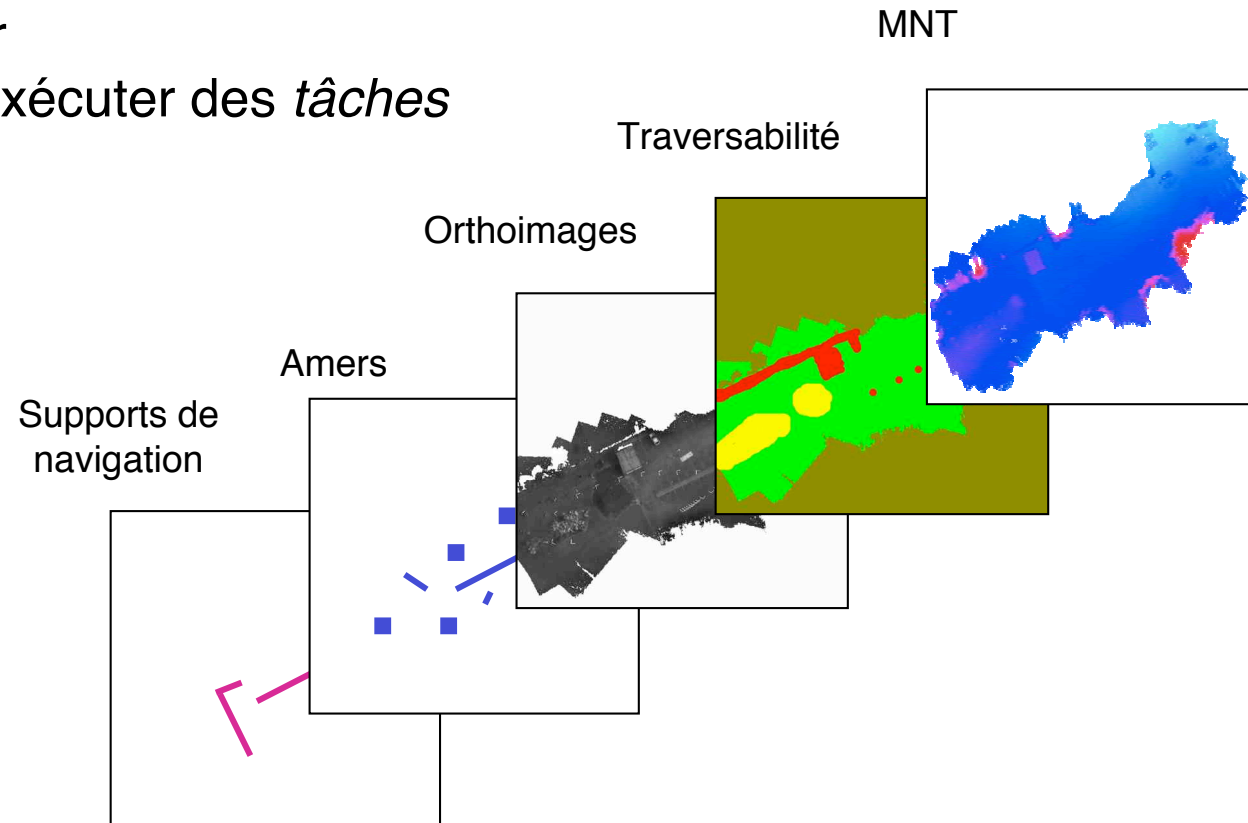


Représentation volumique  
(pour calculs de visibilité)

# Modélisation d'environnement

Bilan :

- Pour planifier des mouvements et déplacements
- Pour réaliser des mouvements
- Pour se localiser
- Pour planifier / exécuter des *tâches*



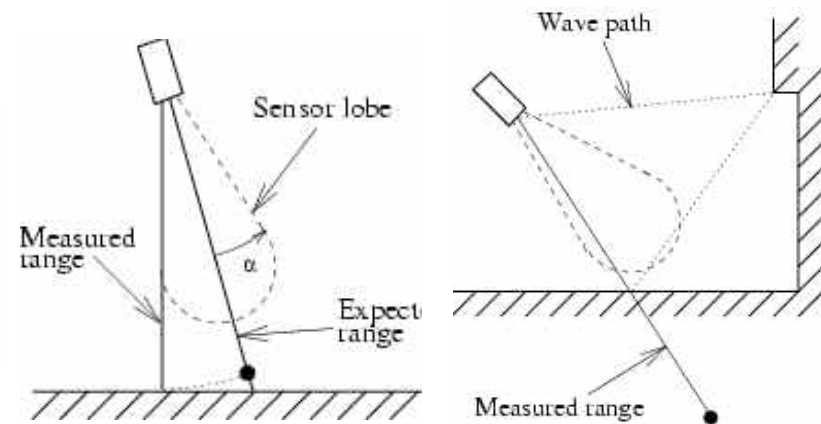
! Différents modèles organisés en «"couches"»

# Pourquoi la vision ?

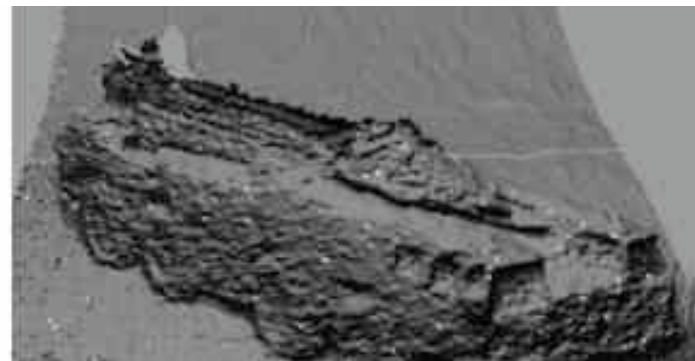
Perception :

« Acquisition et représentation des informations sur l'environnement »

- Télémètre ultrasons



- Imageur radar

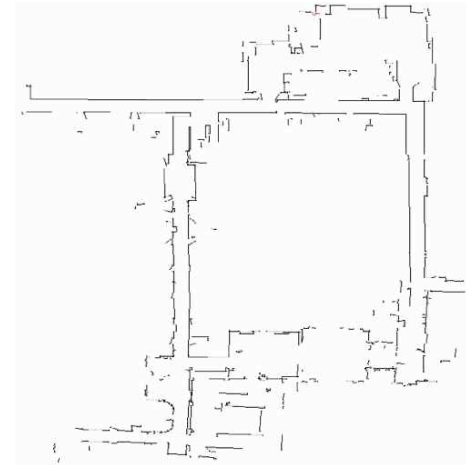


# Pourquoi la vision ?

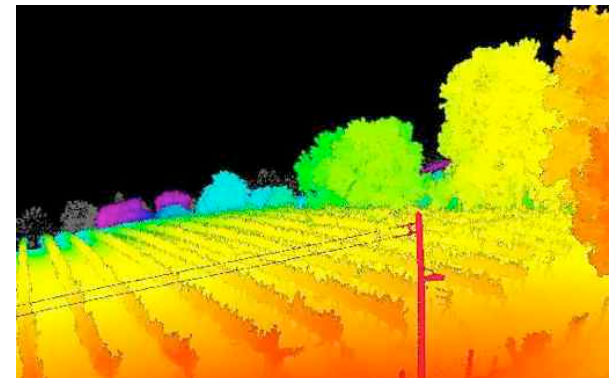
Perception :

« Acquisition et représentation des informations sur l'environnement »

- Télémètre laser («"nappe plane"»)



- Télémètre laser («"3D"»)



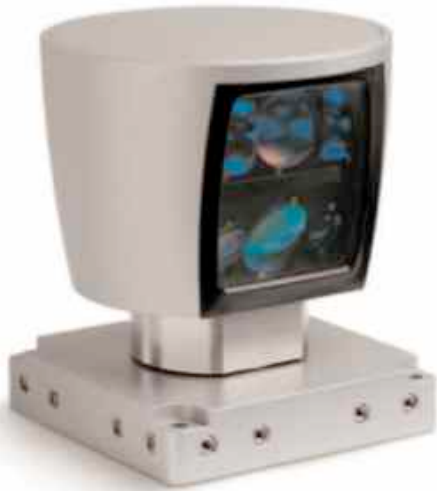


# Pourquoi la vision ?

Perception :

« Acquisition et représentation des informations sur l'environnement »

- Télémètre laser (« "3D" »)

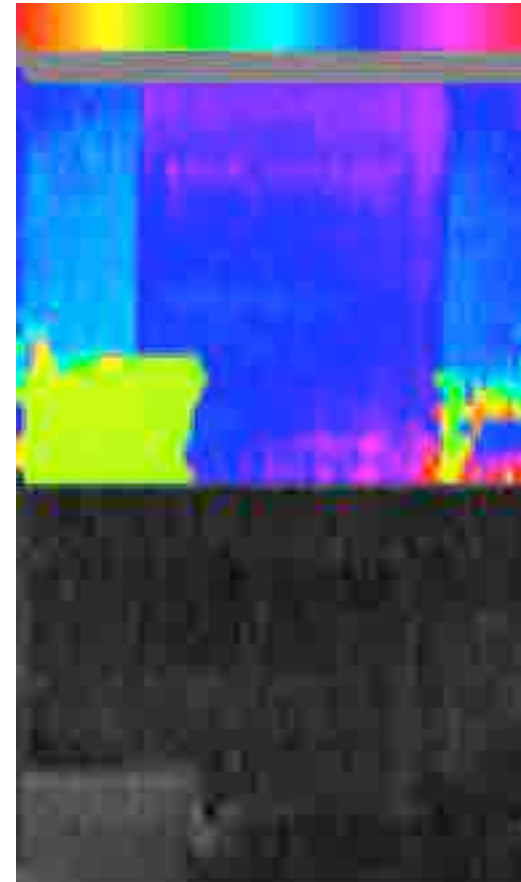
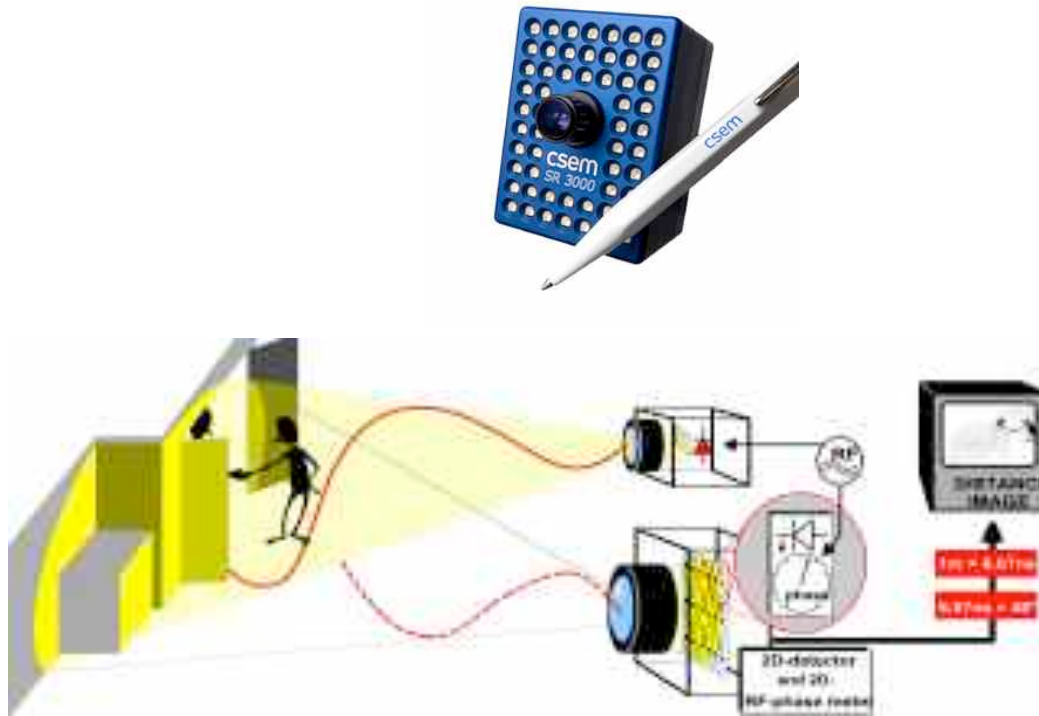


# Pourquoi la vision ?

Perception :

« Acquisition et représentation des informations sur l'environnement »

- Télémètre laser (« "3D scannerless" »)



# Pourquoi la vision ?

Perception :

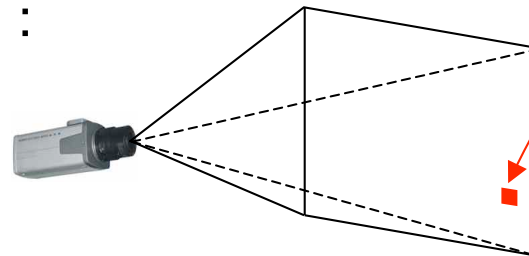
« Acquisition et représentation des informations sur l'environnement »

## Micro UAVs

- Cameras : bas coût, légères et peu gourmandes en énergie



- Perçoivent des données :
  - Dans un volume
  - Très loin
  - Très précisément



1024 x 1024 pixels  
60° x 60° FOV

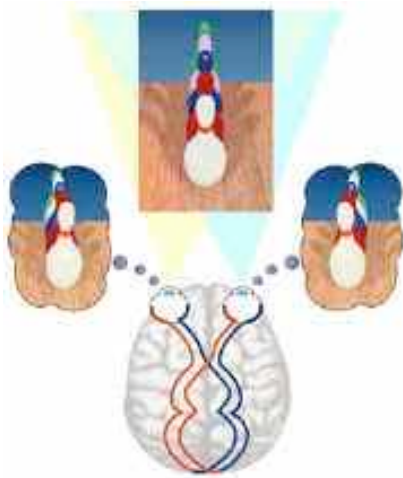
0.06 ° pixel resolution  
(1.0 cm at 10.0 m)

- La perte de profondeur n'est (presque) plus une difficulté
- Les images contiennent un énorme volume d'informations
- Un corpus très conséquent a été développé en vision par ordinateur

# Stéréovision

Du grec «"stéréo"», *solide, espace* : «"Perception de l'espace par vision"»

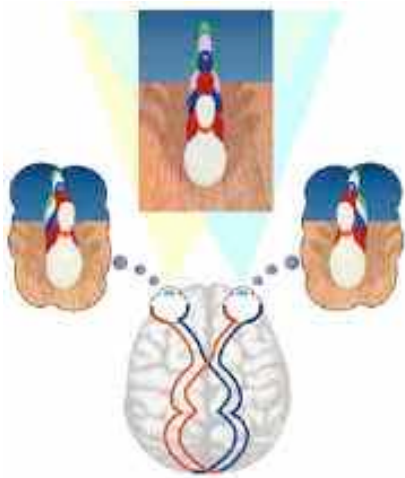
Premier moyen de mesure des distances à distance



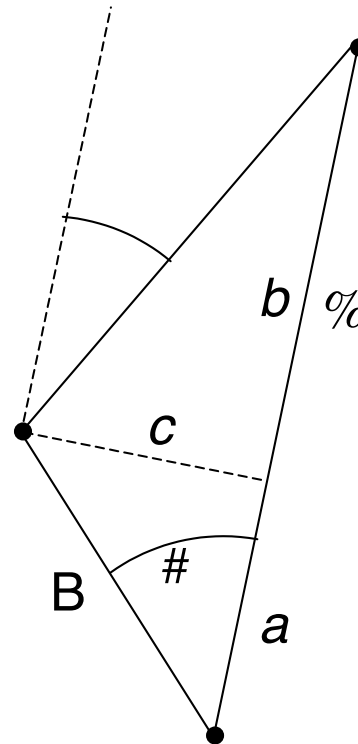
# Stéréovision

Du grec «"stéréo"», *solide, espace* : «"Perception de l'espace par vision"»

Premier moyen de mesure des distances à distance



Stéréovision = triangulation



$$\% = B \sin \# \left( \frac{1}{\tan \#} + \frac{1}{\tan B} \right)$$



# Stéréovision

Deux moyens différents de faire de la stéréovision :

## 1. Stéréovision binoculaire

- Deux caméras
- Base de triangulation fixe et connue («"calibrée"»)



Nombreuses variantes...



Dispositifs commerciaux

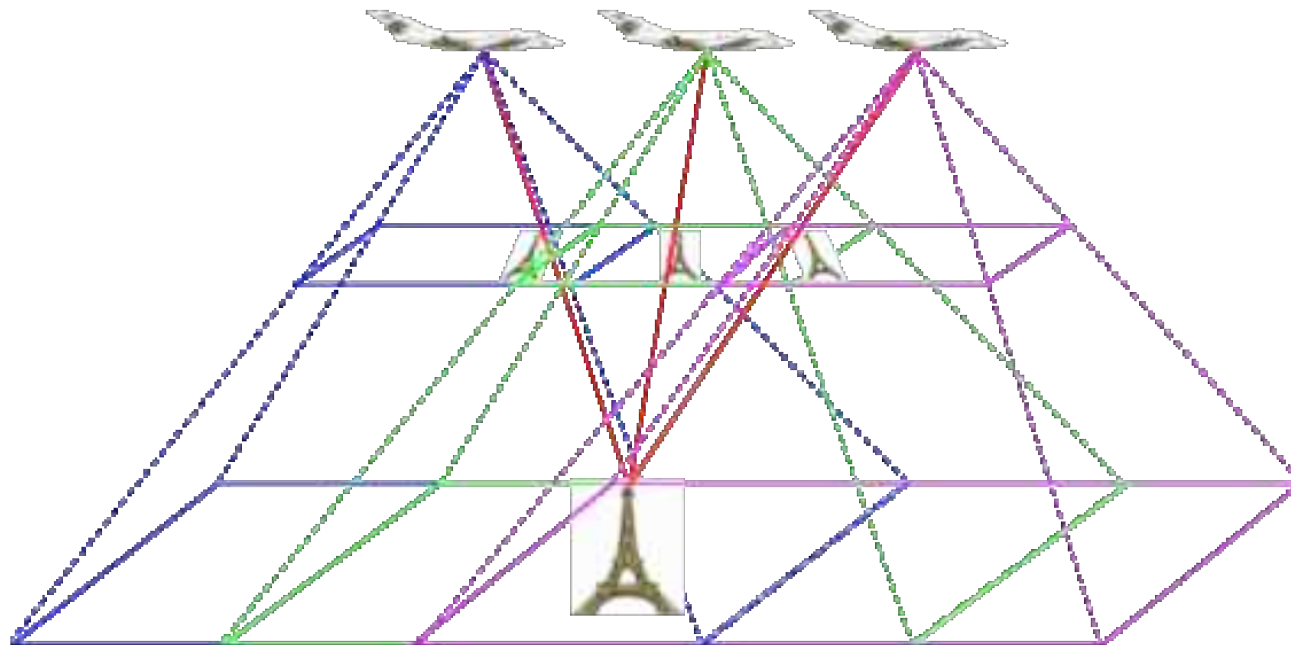


# Stéréovision

Deux moyens différents de faire de la stéréovision :

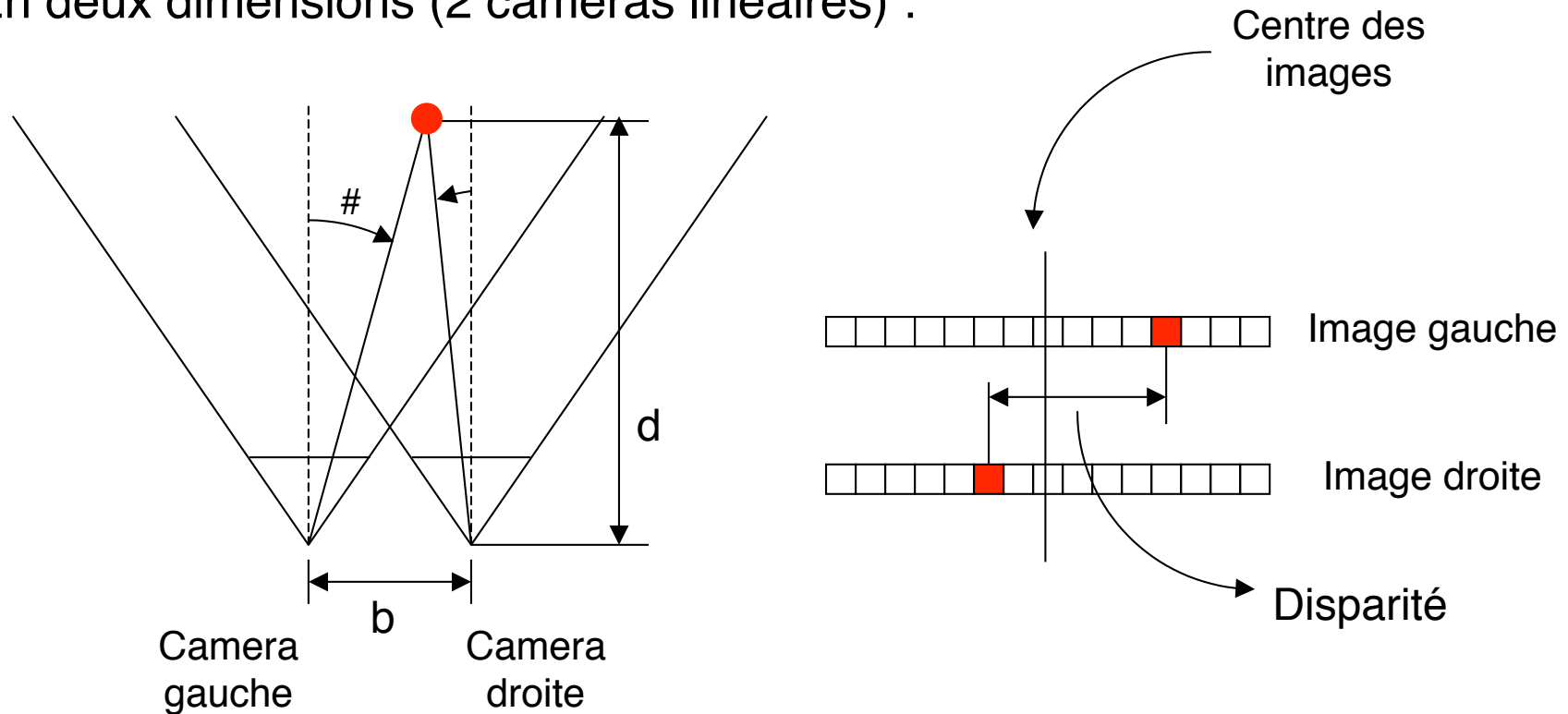
## 2. Stéréovision en mouvement

- Une seule caméra en déplacement
- Base de triangulation variable, à estimer



# Stéréovision binoculaire

En deux dimensions (2 caméras linéaires) :



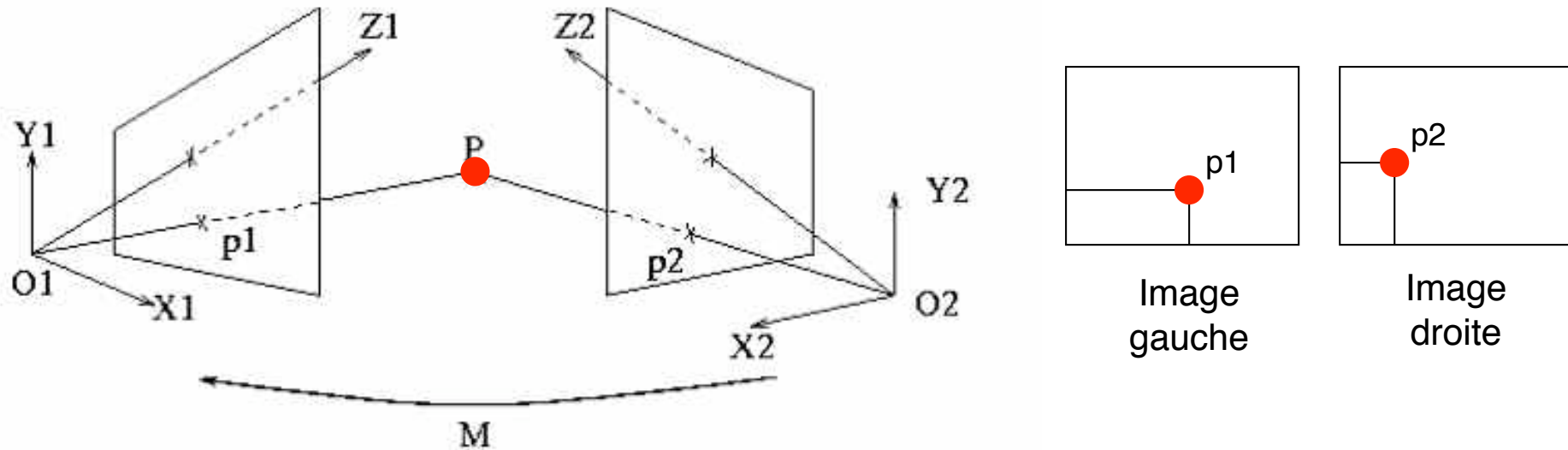
2 angles, une distance : 
$$d = \frac{b}{\tan(\alpha) + \tan(\beta)}$$

! Deux problèmes à résoudre :

1. Estimation de la *géométrie* du système
2. *Mise en correspondance* d'éléments entre les deux images

# Stéréovision binoculaire

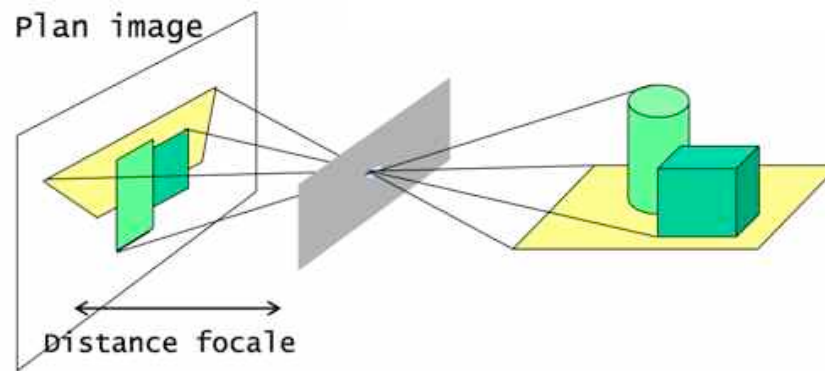
En trois dimensions (2 caméras matricielles) :



4 angles, une transformation 3D (3 angles et 3 distances)

# Modèle géométrique d'une caméra

- Modèle «"sténopé"» («"œil étroit"»)

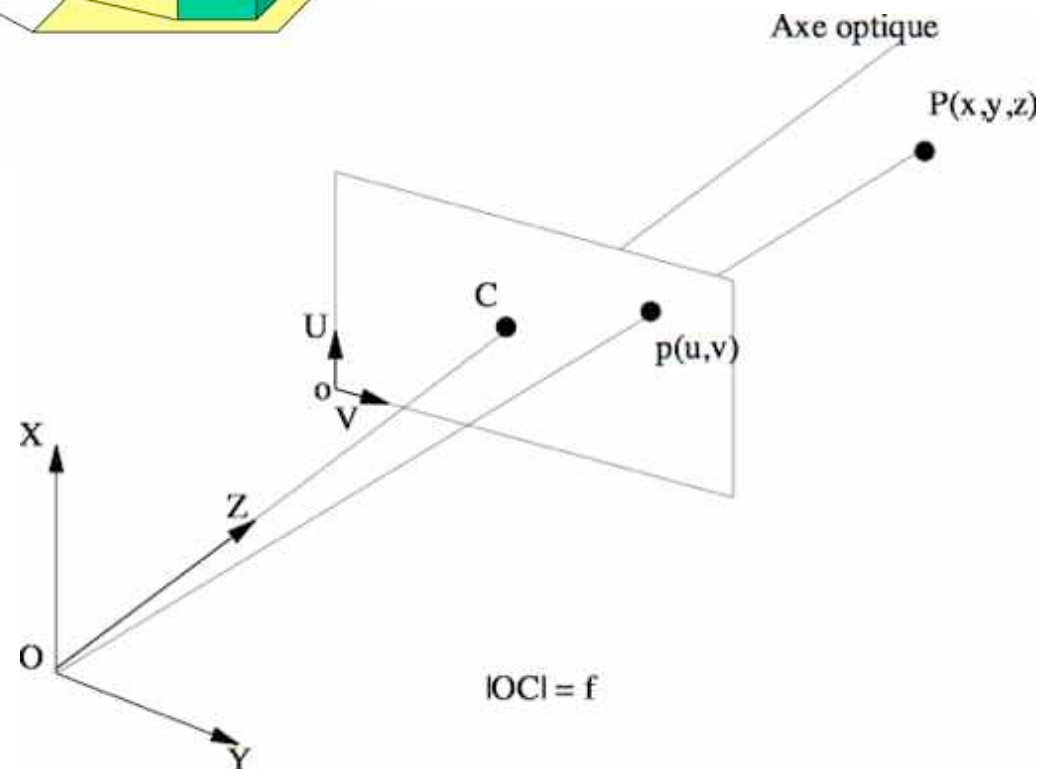


Projection perspective  
Modèle géométrique :

$$(u, v) = P(x, y, z)$$

$$u = k_u f x / z + u_0$$

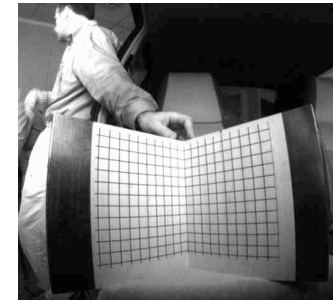
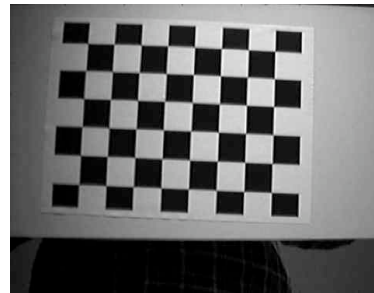
$$v = k_v f y / z + v_0$$



# Modèle géométrique d'une caméra

- Détermination des paramètres du modèle : calibrage
- Principe
  1. Acquisition d'une série d'images d'une mire de calibrage connue
  2. Extraction de points particuliers dans les images
  3. Appariements entre points image et points de la mire
  4. Application d'une technique de minimisation

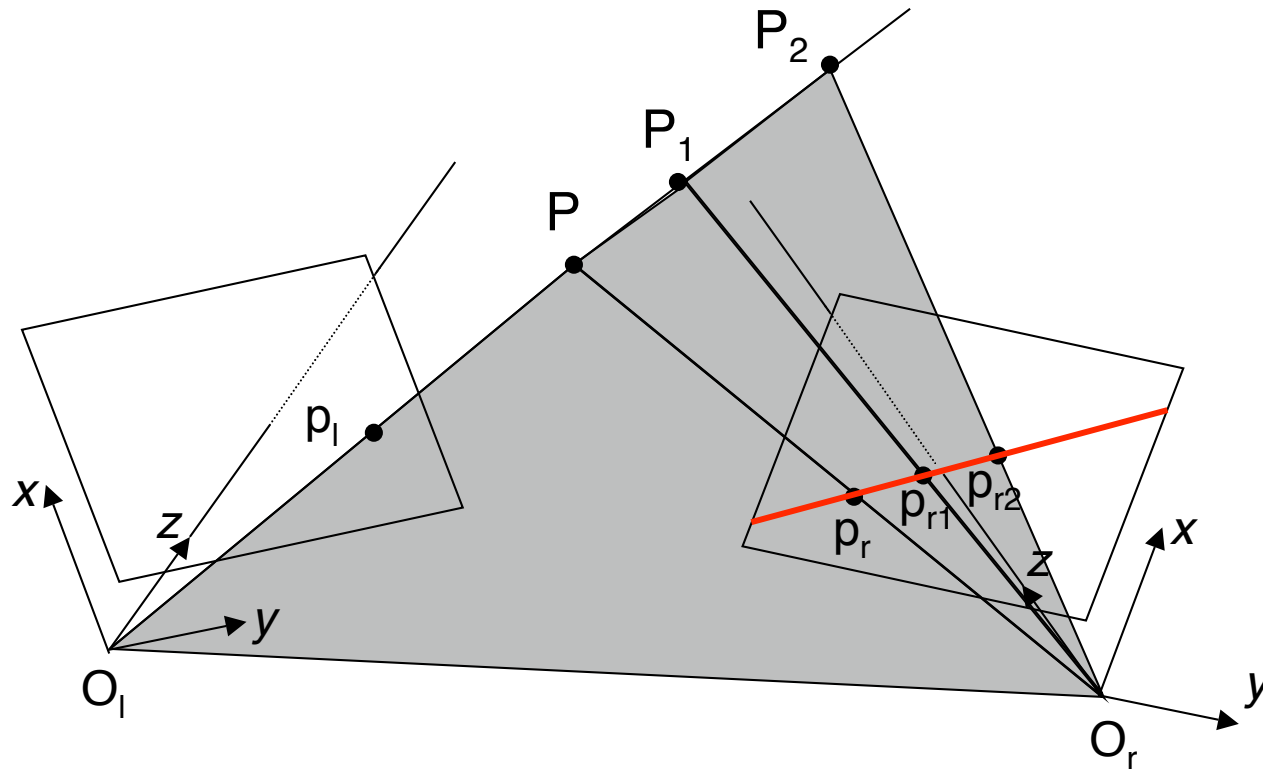
Exemples de mires



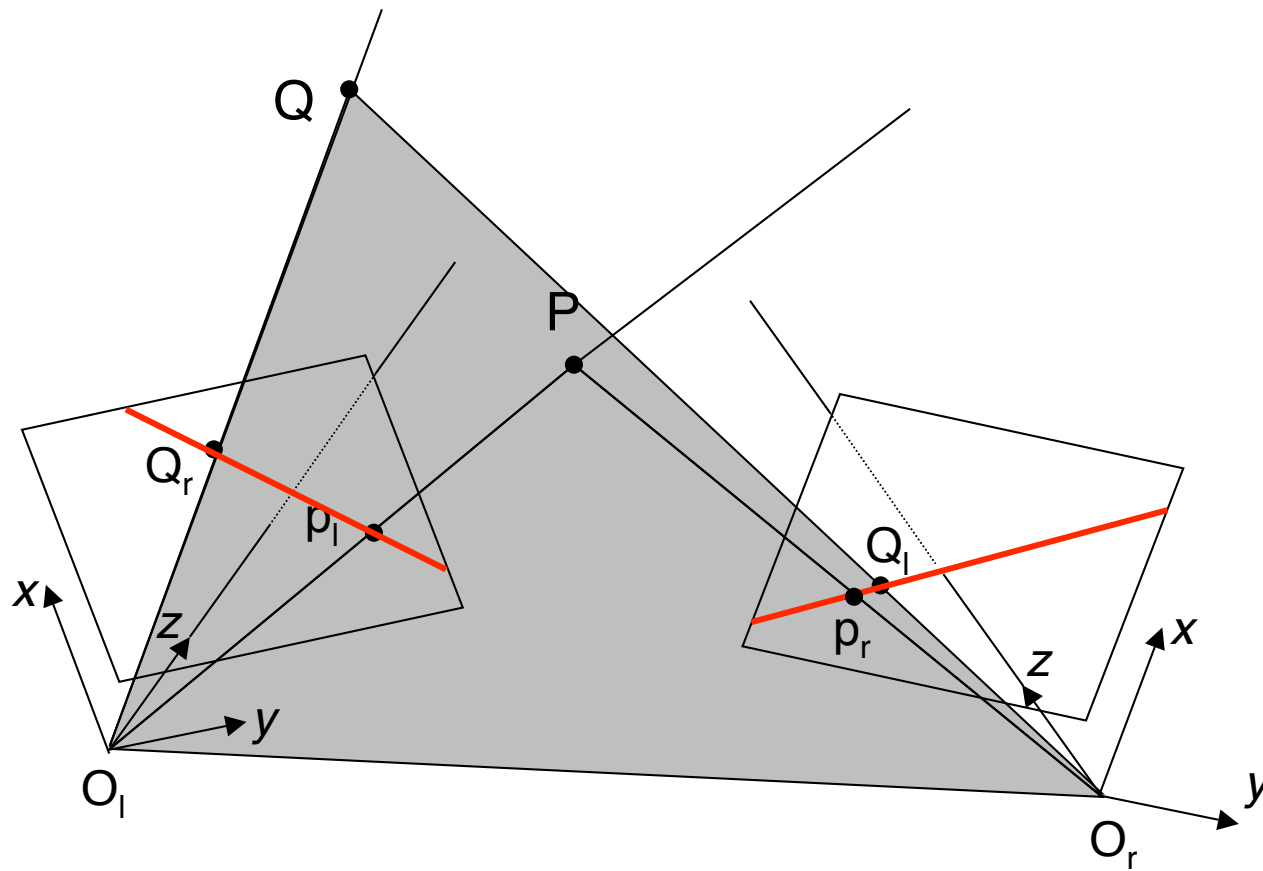
Problème “maîtrisé”, quelques outils disponibles en ligne



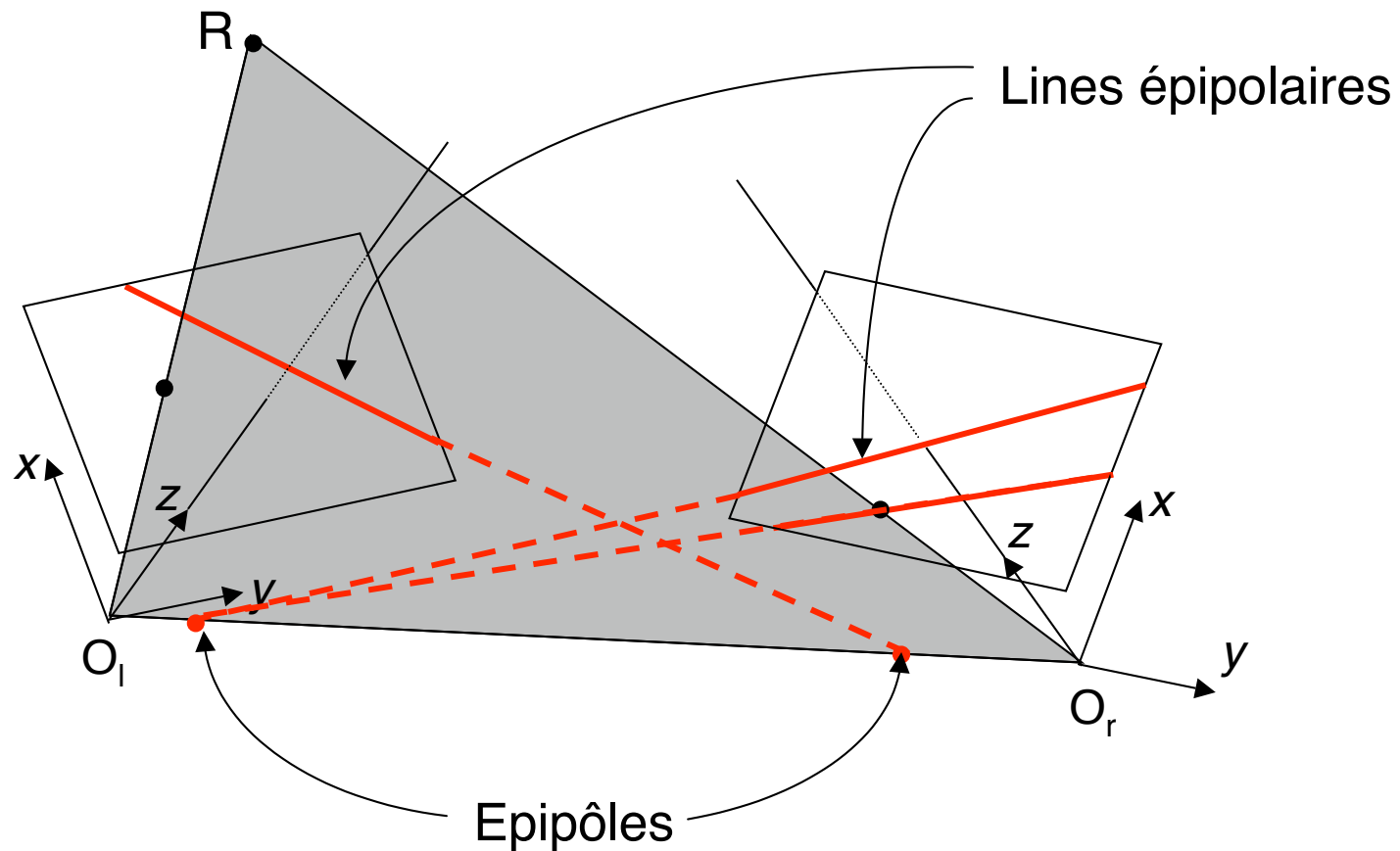
# Géométrie de la stéréovision



# Géométrie de la stéréovision



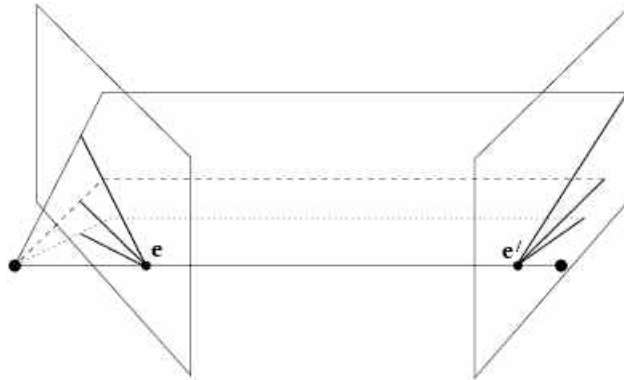
# Géométrie de la stéréovision



! Notion de géométrie épipolaire

# Géométrie de la stéréovision

Caméras convergentes :

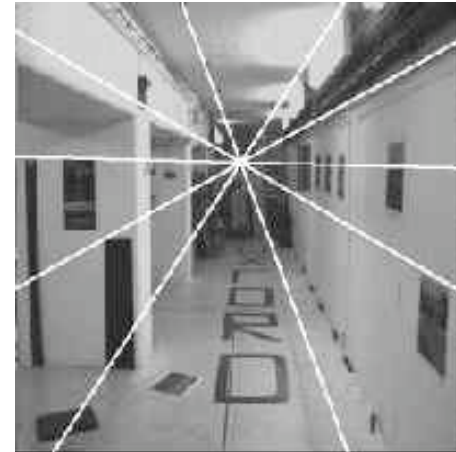
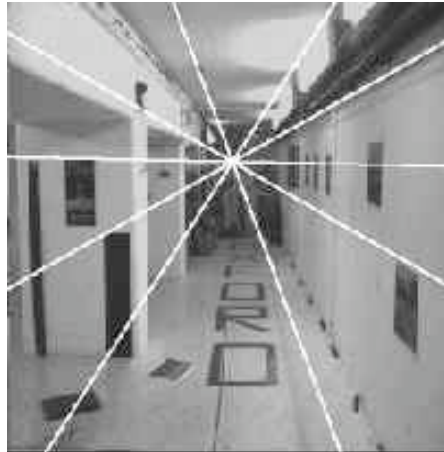
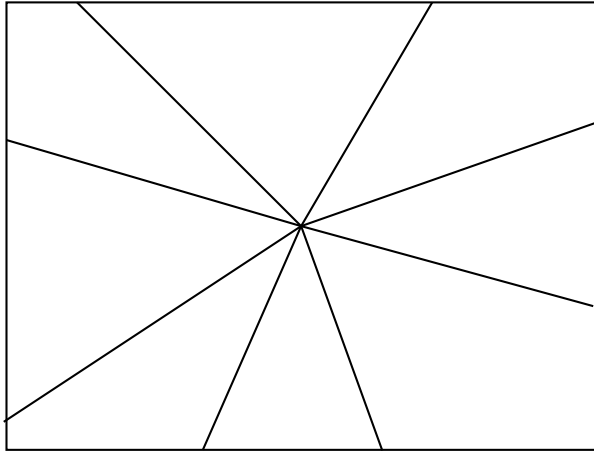


Mouvement parallèle  
au plan image



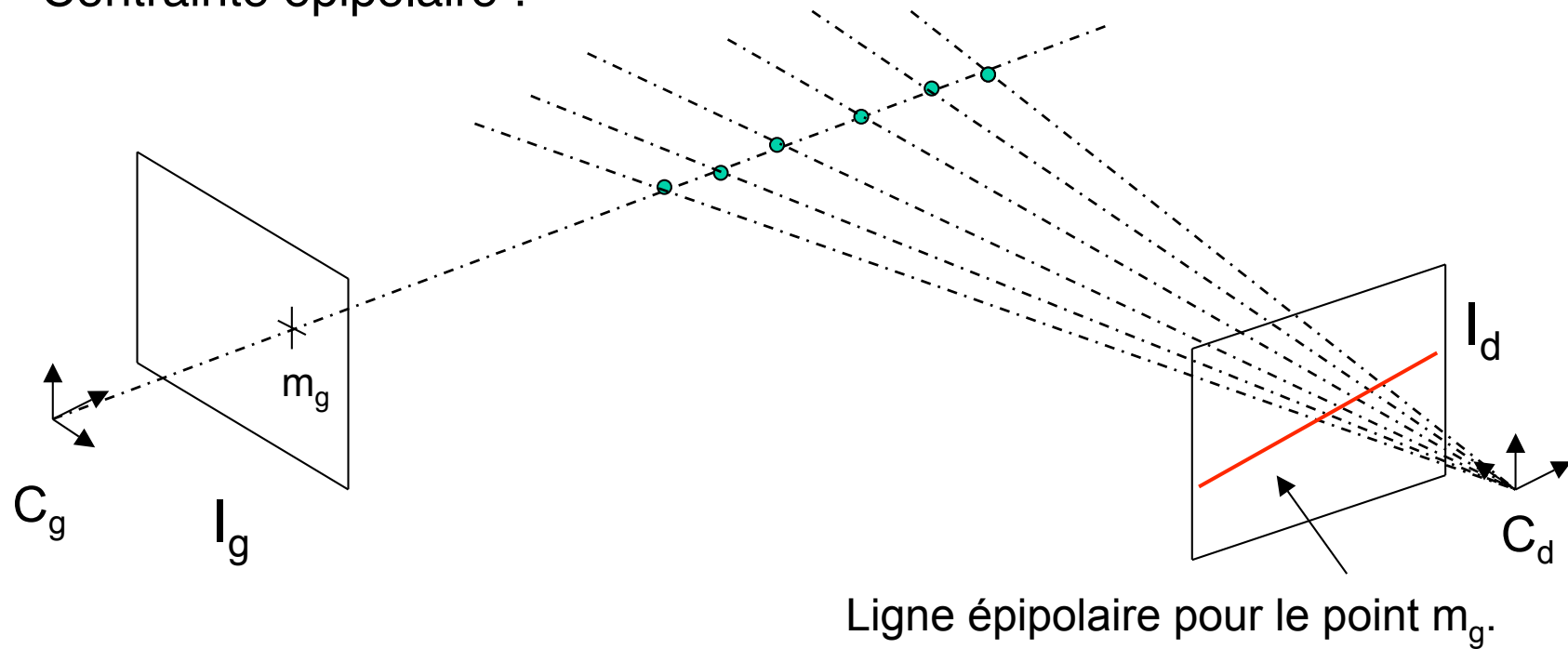
# Géométrie de la stéréovision

Caméra en rapprochement :



# Géométrie de la stéréovision

Contrainte épipolaire :



Contrainte épipolaire : le correspondant d'un pixel de  $I_g$  se trouve dans  $I_d$  sur la ligne épipolaire définie par  $C_d$ ,  $C_g$  et  $m_g$

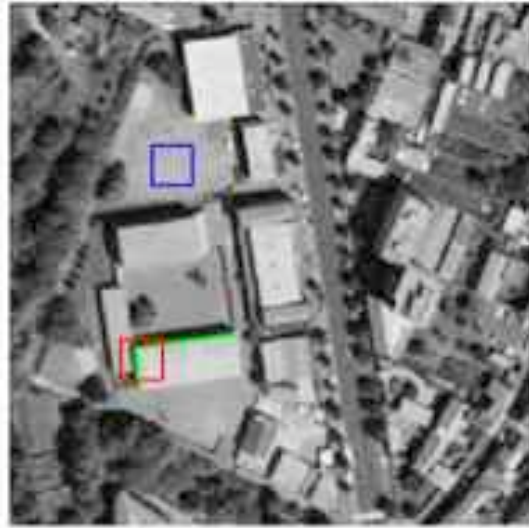
Si  $\mathbf{x}_g$  et  $\mathbf{x}_d$  sont images d'un même point 3D alors :  $\mathbf{x}_d^t \mathbf{F} \mathbf{x}_g = 0$

$\mathbf{F}$  est la matrice fondamentale



# Mise en correspondance

- Le problème :



Détail rouge



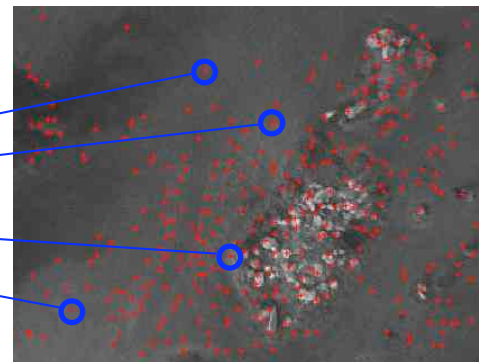
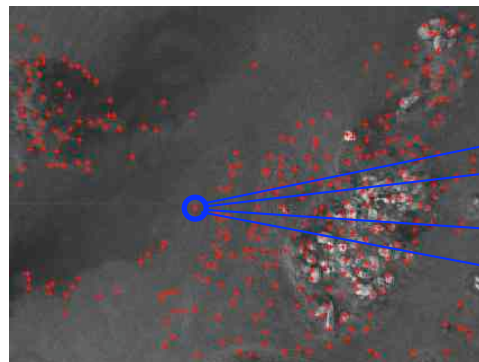
A



B



C



?

# Mise en correspondance

- Problème «"mal posé"» :
  - Bruit, ambiguïtés, imprécisions (variations radiométriques)
  - Difficulté très variable selon :
    - Le «"contenu"» des images (zones homogènes, textures, contours...)
    - Le type des images (noir&blanc, couleur, IR...)
    - Le type d'environnement (indoor, outdoor, structuré ou non, terrestre, aérien, ...)
    - Les différences de point de vue
    - Les occultations
  - Problème de grande dimension

# Mise en correspondance

- Que mettre en correspondance ?

Régions

Segments

Contours

Pixels



image gauche



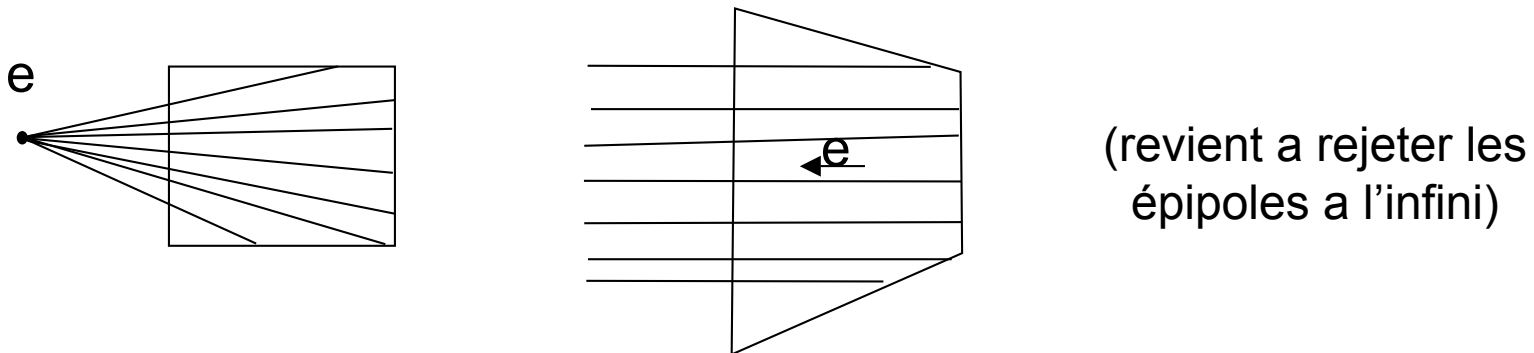
image droite

Lien entre mise en correspondance et géométrie :

- Si géométrie connue, mise en correspondance facilitée
- Si primitives mises en correspondances, possibilité de déterminer la géométrie

# Stéréovision binoculaire dense

- Objectif : mettre en correspondance l'ensemble des pixels d'une paire stéréoscopique
- Première étape : *rectification* des images  
Application d'une transformation au couple d'images afin que les **lignes épipolaires soient parallèles aux lignes des images**.



Pourquoi rectifier les images ?

Faciliter les traitements de mise en correspondance

Faciliter la triangulation

# Stéréovision binoculaire dense

- Première étape : *rectification* des images

Avant



Après

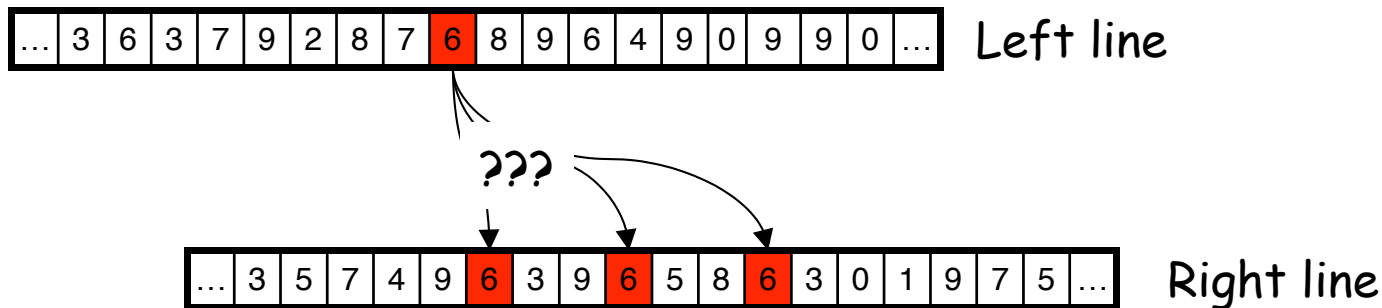


# Stéréovision binoculaire dense

- Deuxième étape : mettre en correspondance les pixels

Principe : pour chaque pixel de  $I_g$ , trouver le pixel de  $I_d$  qui lui correspond

Après rectification : recherche du pixel sur la même ligne



Simple comparaison de pixels un à un : trop ambiguë

! comparaison de voisinages

|   |   |   |          |   |   |   |
|---|---|---|----------|---|---|---|
| 1 | 3 | 5 | 4        | 7 | 6 | 4 |
| 8 | 7 | 2 | 6        | 8 | 9 | 6 |
| 0 | 8 | 6 | 7        | 4 | 6 | 3 |
| 6 | 7 | 8 | <b>9</b> | 0 | 7 | 5 |
| 6 | 6 | 5 | 3        | 2 | 6 | 7 |
| 9 | 8 | 8 | 9        | 0 | 9 | 8 |
| 6 | 7 | 3 | 6        | 5 | 4 | 3 |

←  $2M + 1$  →

?

|   |   |   |          |   |   |   |
|---|---|---|----------|---|---|---|
| 1 | 3 | 5 | 7        | 7 | 5 | 4 |
| 8 | 7 | 3 | 6        | 9 | 7 | 6 |
| 0 | 7 | 6 | 8        | 4 | 5 | 2 |
| 6 | 7 | 7 | <b>9</b> | 1 | 7 | 5 |
| 6 | 7 | 5 | 2        | 3 | 6 | 7 |
| 8 | 8 | 7 | 9        | 0 | 8 | 8 |
| 6 | 6 | 5 | 6        | 7 | 4 | 9 |

# Stéréovision binoculaire dense

- Deuxième étape : mettre en correspondance les pixels

Comparaison directes des radiométries :

$$score(p_1, p_2) = \sum_{x=-M}^{+M} \sum_{y=-M}^{+M} \|I(u_1 - x, v_1 - y) - I(u_2 - x, v_2 - y)\| \quad \text{SAD}$$

$$score(p_1, p_2) = \sum_{x=-M}^{+M} \sum_{y=-M}^{+M} (I(u_1 - x, v_1 - y) - I(u_2 - x, v_2 - y))^2 \quad \text{SSD}$$

**ZNCC :**

$$score(p_1, p_2) = \frac{\frac{1}{N} \sum_{x,y} I(u_1 - x, v_1 - y) I(u_2 - x, v_2 - y)}{\sqrt{\left( \frac{1}{N} \sum_{x,y} I(u_1 - x, v_1 - y)^2 \right) \left( \frac{1}{N} \sum_{x,y} I(u_2 - x, v_2 - y)^2 \right)}}$$

$$score(p_1, p_2) = \frac{\frac{1}{N} \sum_{x,y} I(u_1 - x, v_1 - y) I(u_2 - x, v_2 - y) - \mu_1 \mu_2}{\sigma_1 \sigma_2}$$



# Stéréovision binoculaire dense

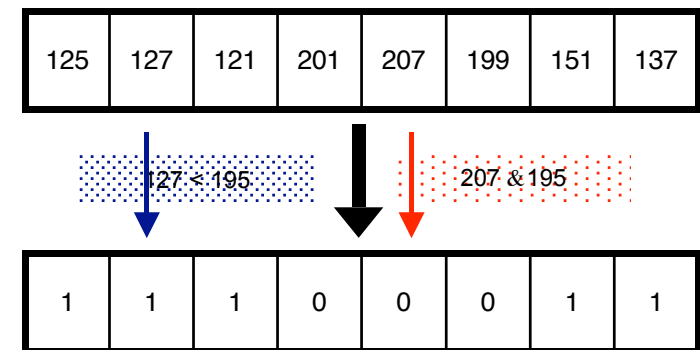
- Deuxième étape : mettre en correspondance les pixels

Un tout autre moyen (non-paramétrique) de mesurer la ressemblance :

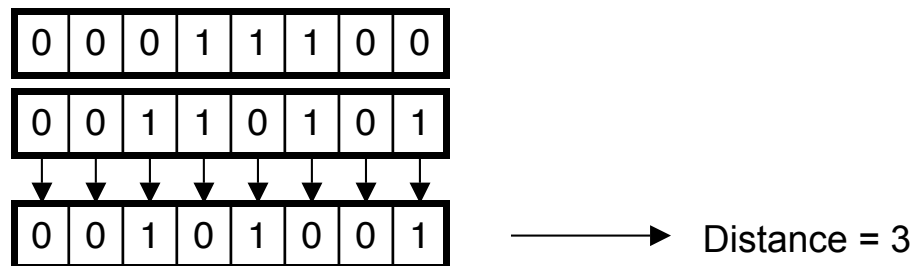
## 1. Transformée «"Census"» :

Générer un mot de bits par comparaisons successives de la luminance du pixel central avec celles de pixels voisins

|     |     |     |
|-----|-----|-----|
| 125 | 127 | 121 |
| 137 | 195 | 201 |
| 151 | 199 | 207 |



## 2. Comparaison de deux pixels individuels : distance de Hamming

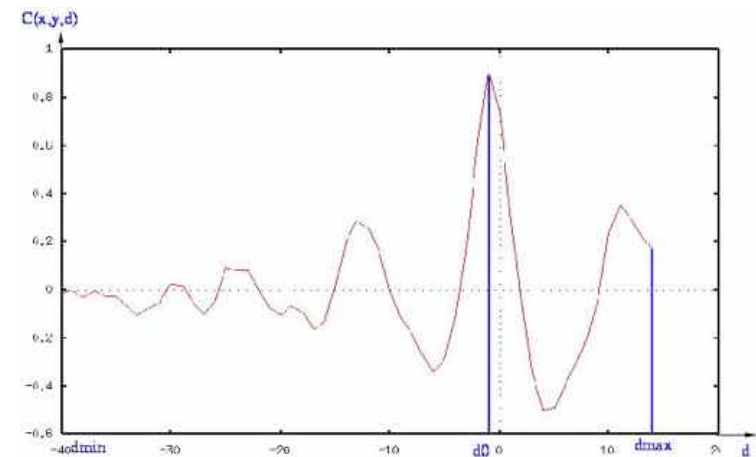
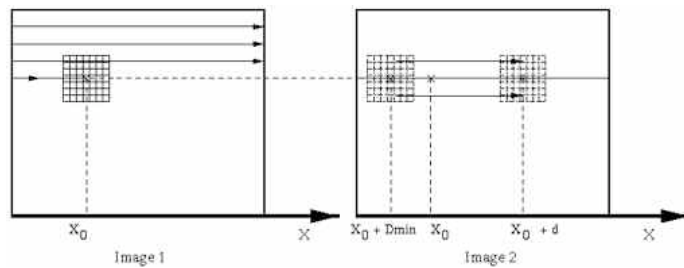


## 3. Calcul des scores sur une fenêtre (idem scores paramétriques)

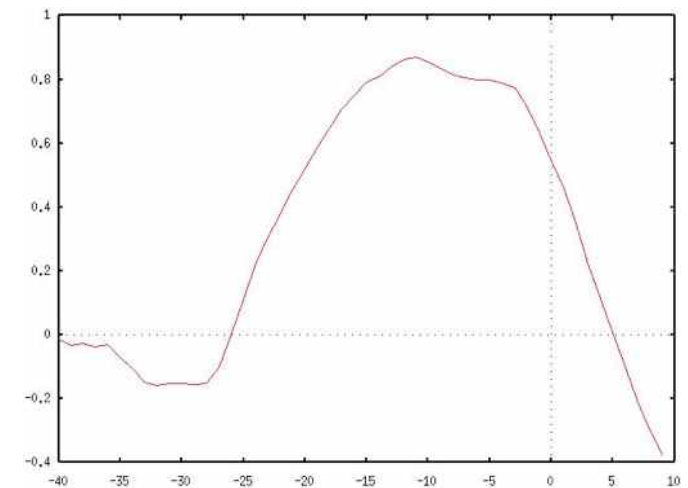
# Stéréovision binoculaire dense

- Deuxième étape : mettre en correspondance les pixels

Exemples de courbes de scores



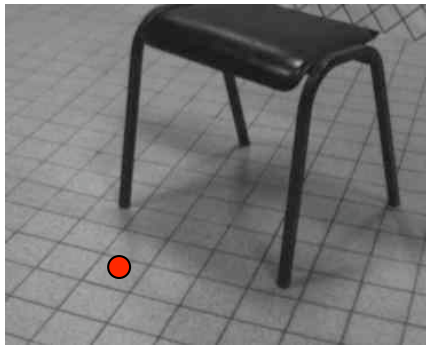
Cas de faibles textures



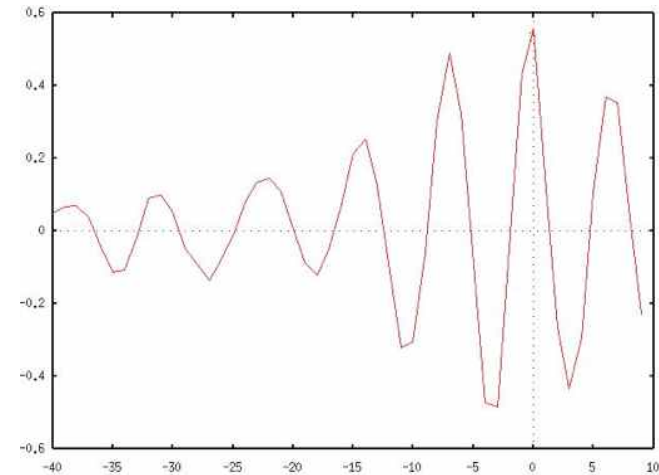
# Stéréovision binoculaire dense

- Deuxième étape : mettre en correspondance les pixels

Motifs répétitifs



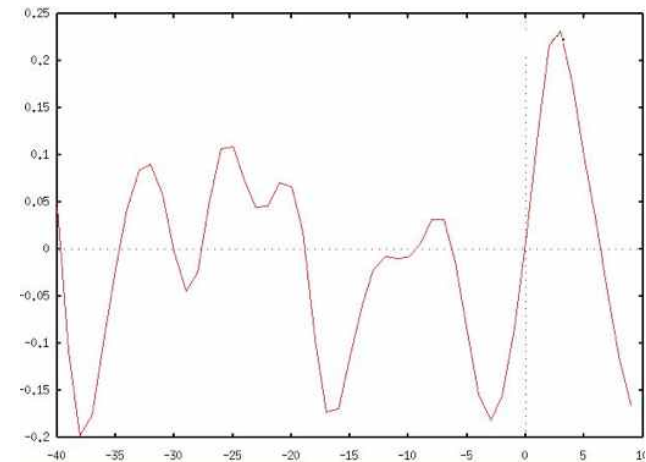
!



Ocultations



!



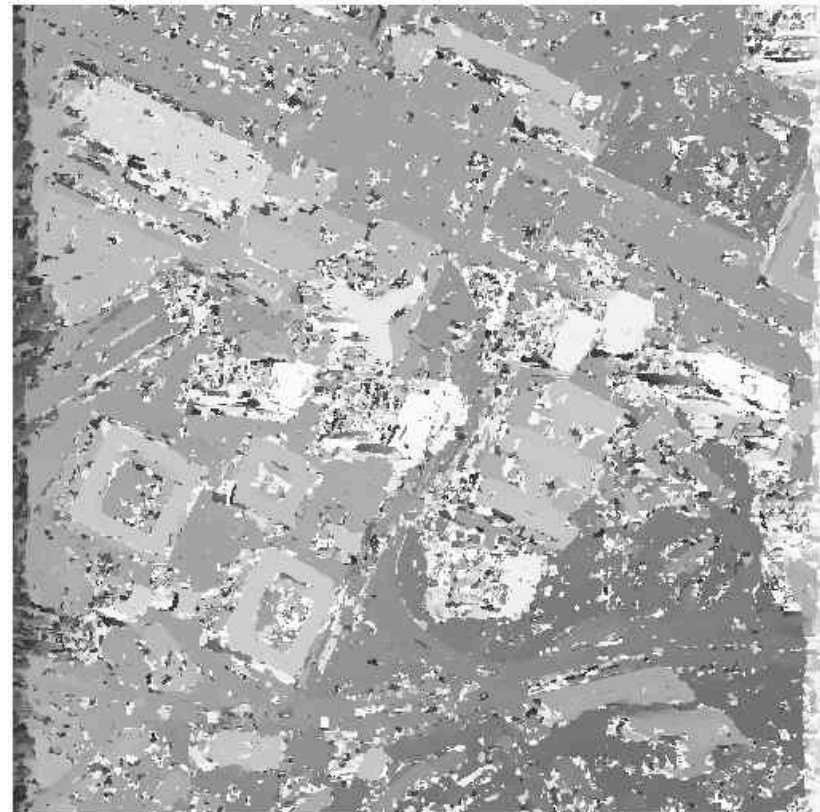
# Stéréovision binoculaire dense

- Deuxième étape : mettre en correspondance les pixels

Trop de seuils (paramètres), difficilement gérable

! Application de la contrainte d'unicité

$$d_{gd}(x_g, y_g) = u \text{ ssi } d_{gd}(x_g, y_g) = -d_{dg}(x_g + u, y_g)$$





# Stéréovision binoculaire dense

- Deuxième étape : mettre en correspondance les pixels

Trop de seuils (paramètres), difficilement gérable

! Application de la contrainte d'unicité

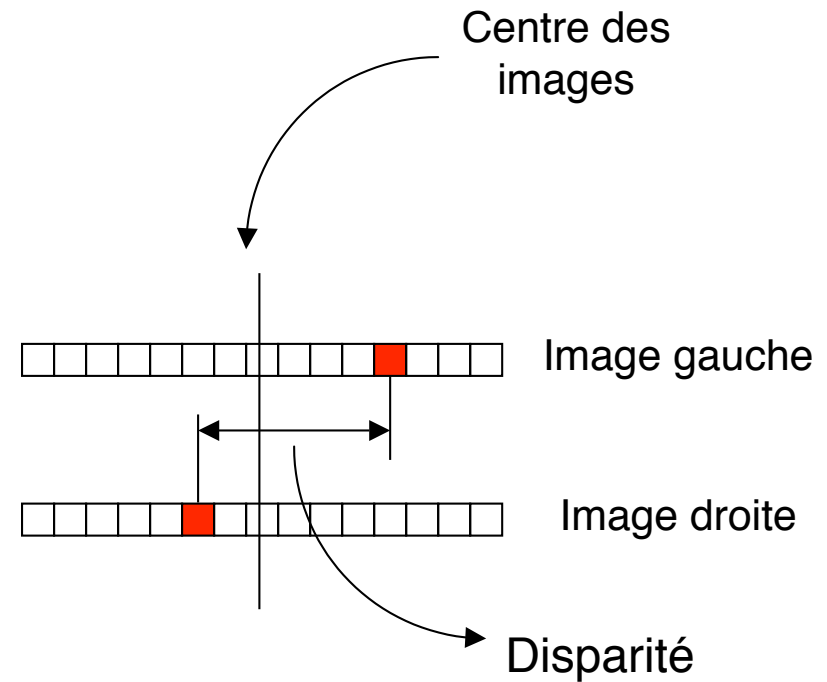
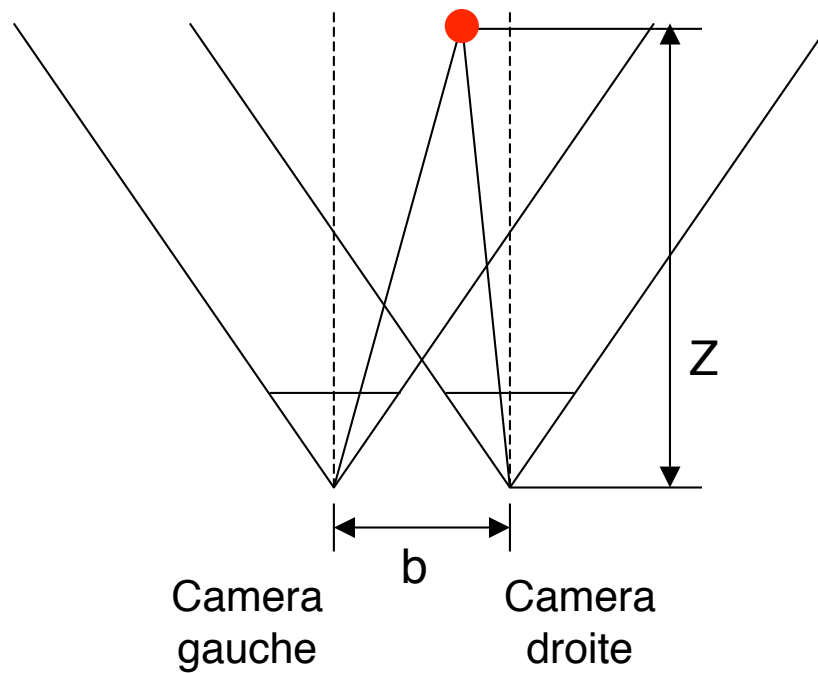
$$d_{gd}(x_g, y_g) = u \text{ ssi } d_{gd}(x_g, y_g) = - d_{dg}(x_g + u, y_g)$$



# Stéréovision binoculaire dense

- Troisième étape : triangulation

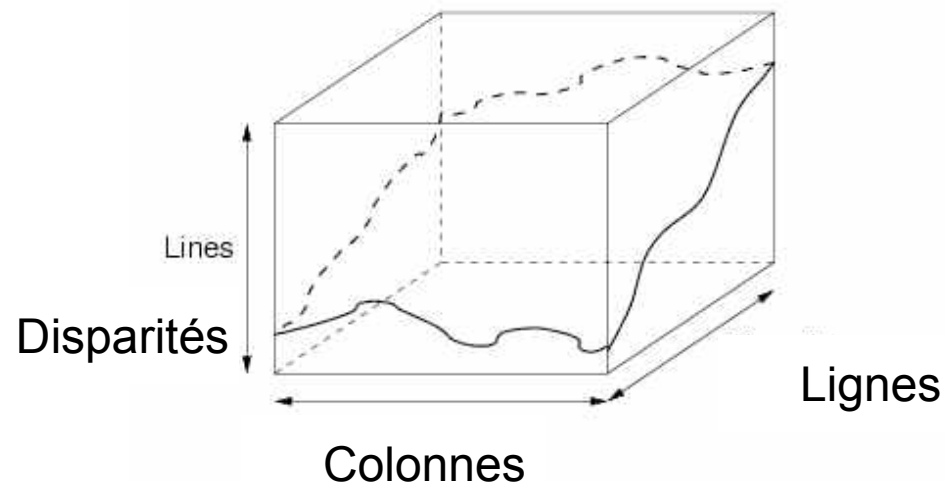
Images rectifiées : on se ramène à une triangulation dans le plan



$$Z = \frac{b}{d}$$

# Stéréovision binoculaire dense

- Complexité algorithmique : cubique !



|     | ZNCC (ms) | Census (ms) | Rectif (ms) | image size (pixel) |
|-----|-----------|-------------|-------------|--------------------|
| 1:1 | 59150     | 40840       | 300         | 572x768            |
| 1:2 | 5320      | 4600        | 60          | 286x384            |
| 1:4 | 560       | 470         | 10          | 142x192            |
| 1:8 | 70        | 70          | 1           | 70x96              |

! Développement d'une approche pyramidale : focalisation du domaine de recherche des disparités



# Stéréovision binoculaire dense

- Difficultés

1. grossissement des obstacles

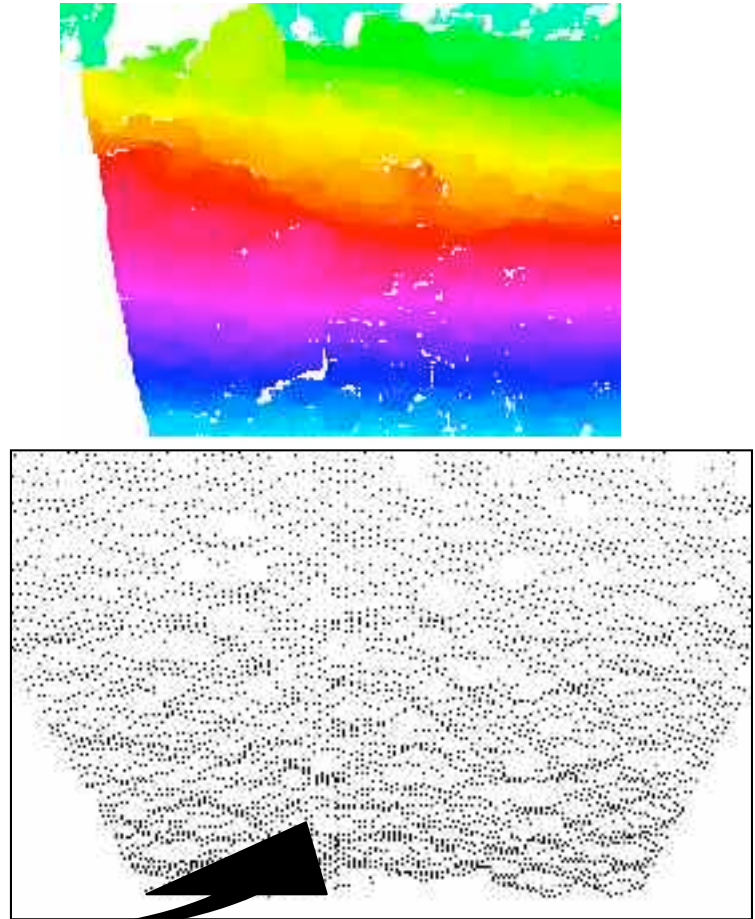
2. “Vaguelettes”



Image gauche



Points 3D



3. Lissage des discontinuités

4. Absence de données

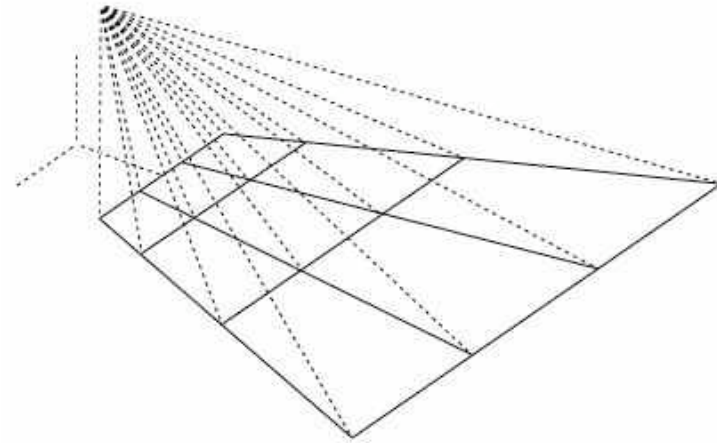
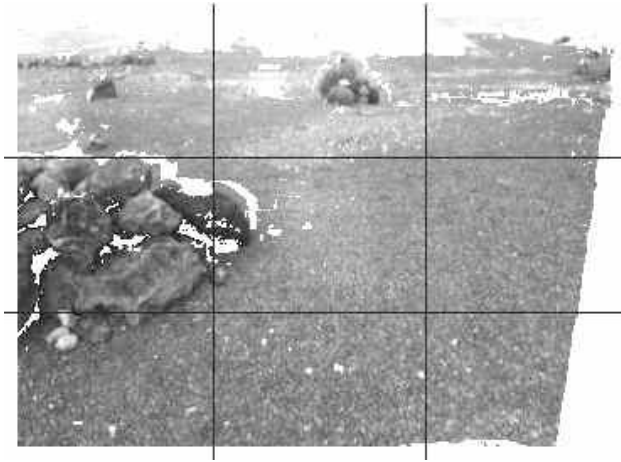
5. Calibration

# Fil conducteur

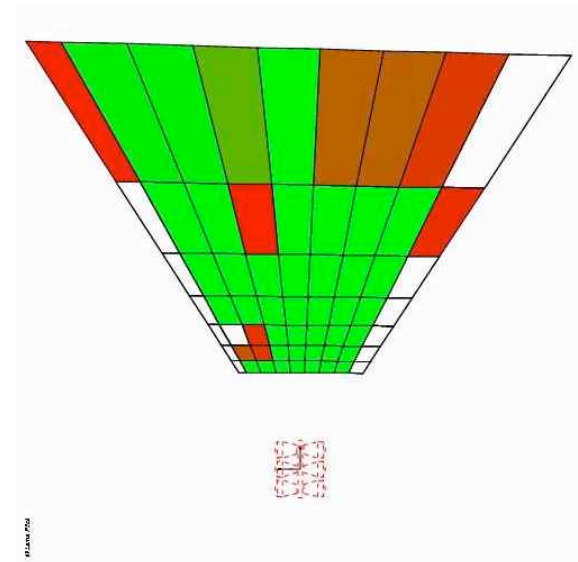
- Introduction
- Modélisation d'environnement par vision
  - Quels modèles sont nécessaires ?
  - Pourquoi la vision ?
  - Stéréovision
  - Exemples d'algorithmes de modélisation

# Modèle de traversabilité

## 1. Discrétisation de la zone perçue

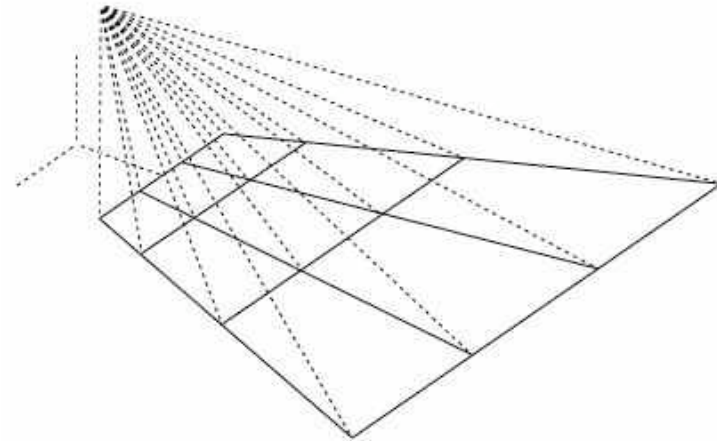
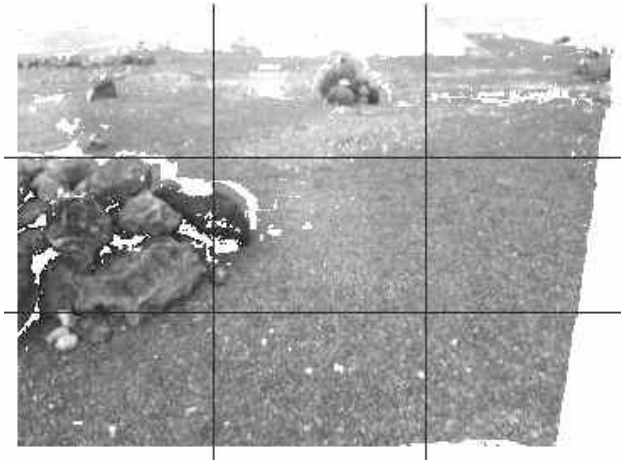


## 2. Étiquetage probabiliste



# Modèle de traversabilité

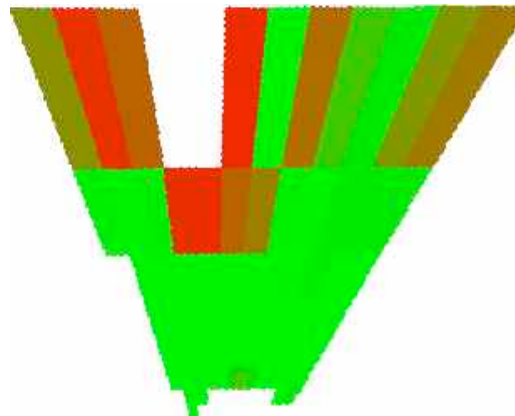
## 1. Discrétisation de la zone perçue



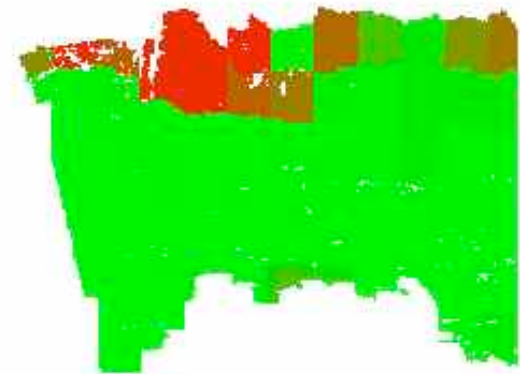
## 2. Étiquetage probabiliste



Pixels corrélés



Classes déterminées  
(vue du dessus)

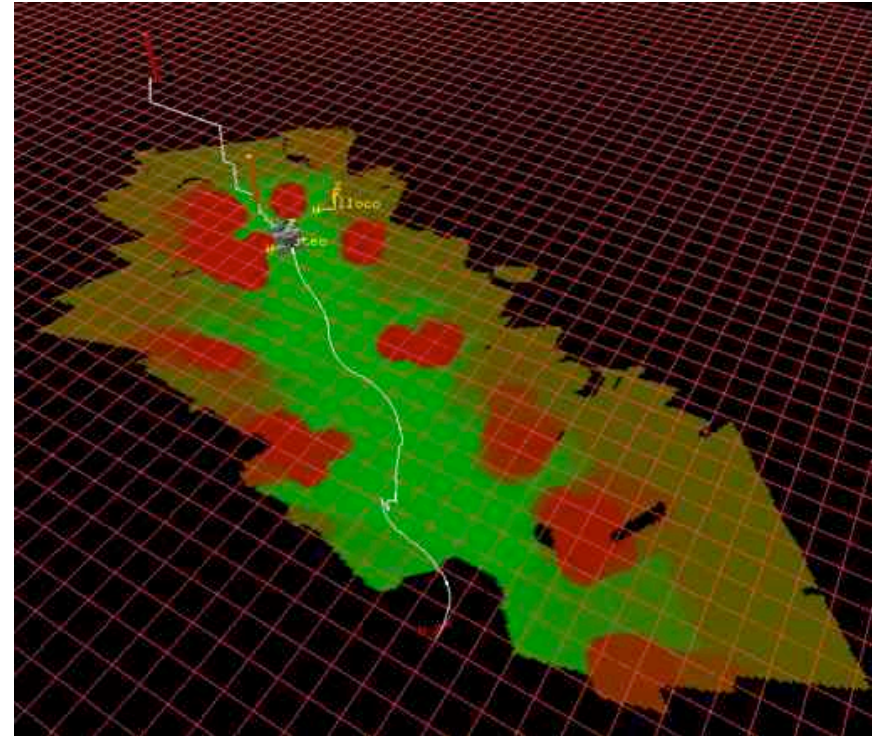
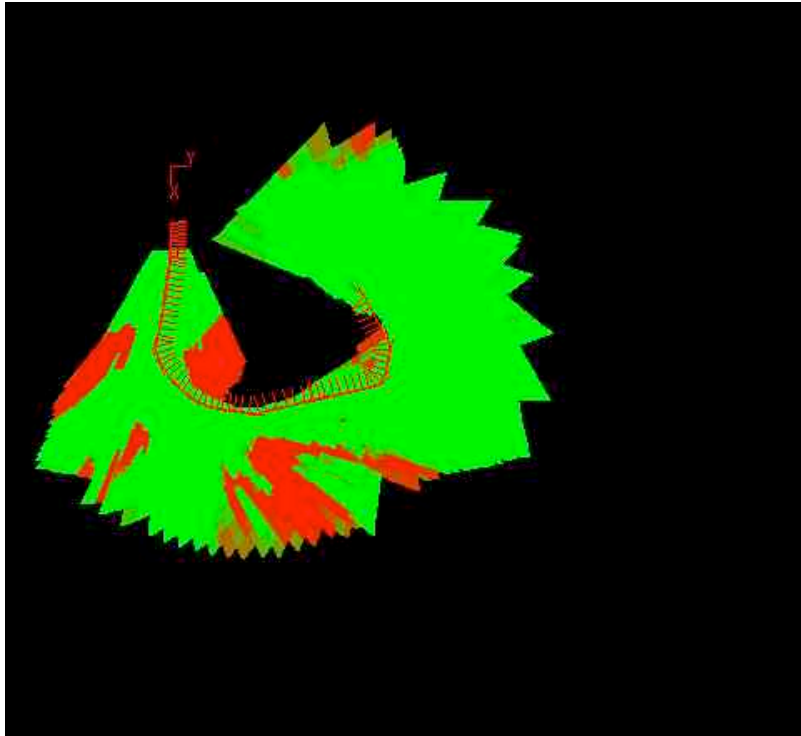


Classes déterminées  
(reprojection dans l'image)



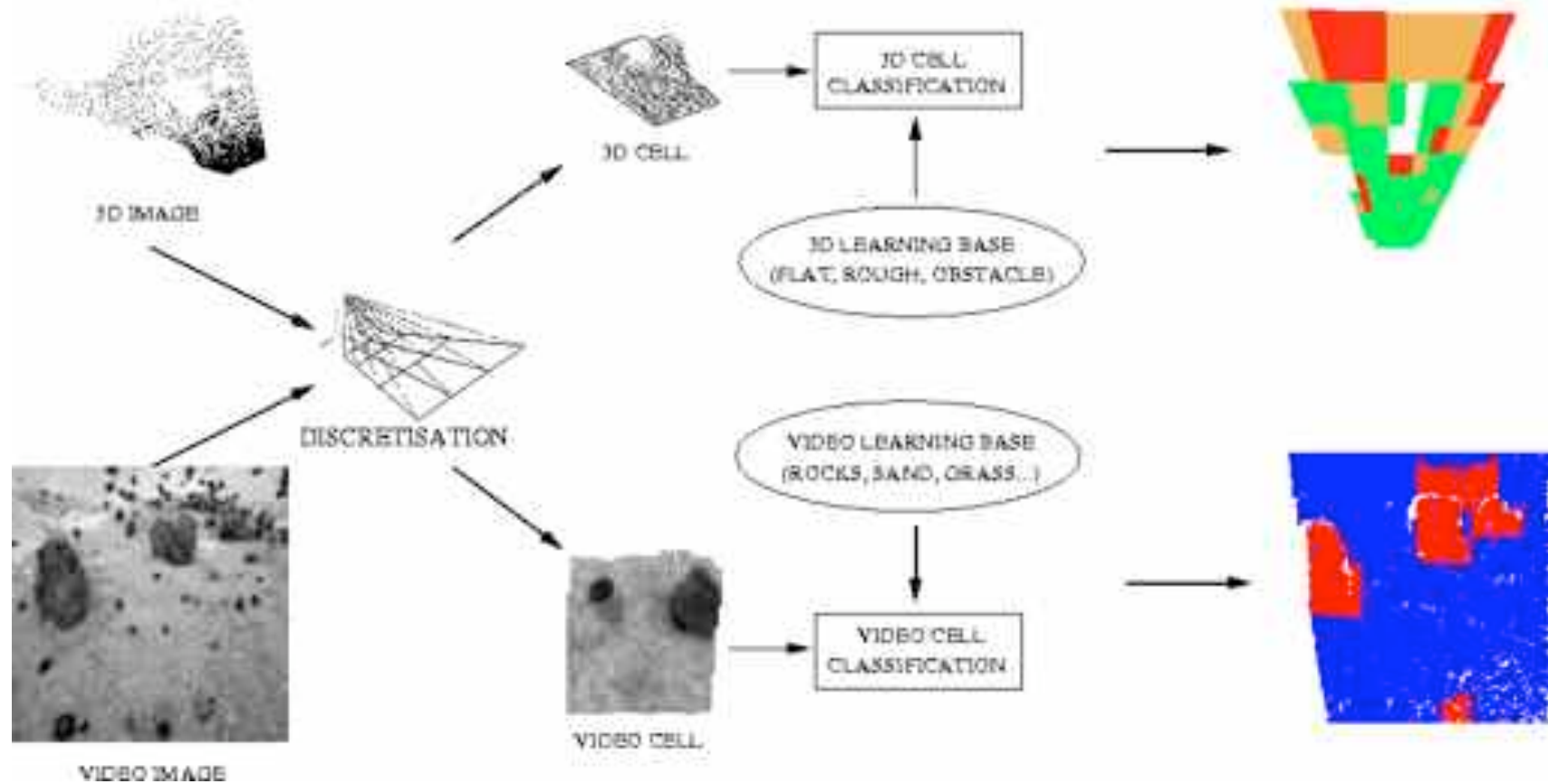
# Modèle de traversabilité

## 3. Fusion des probabilités lors des déplacements



# Modèle de traversabilité

## Introduction de la texture



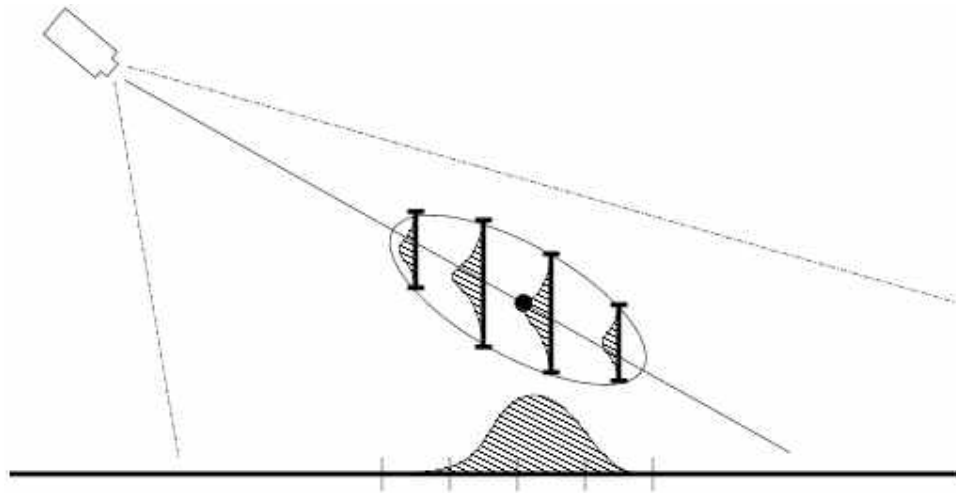


# Modèle numérique de terrain

DEM:  $z = f(x, y)$  on a regular Cartesian grid

Ground rover case:

- Varying resolution
- Imprecision on the data  
uncertainties in the values



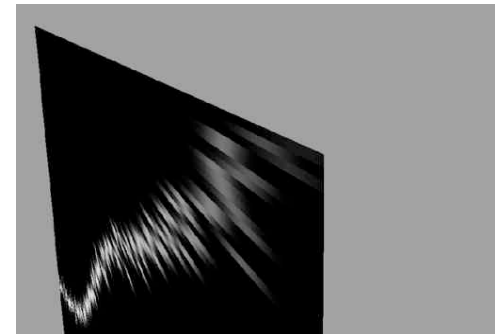
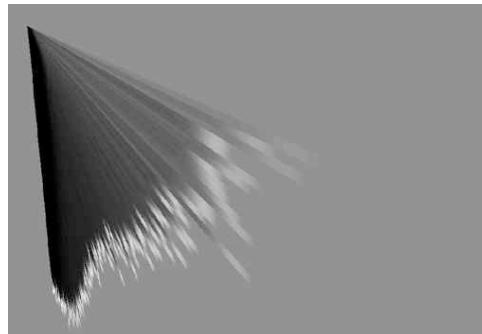
# Modèle numérique de terrain

Data uncertainties, vertical features and overheads:

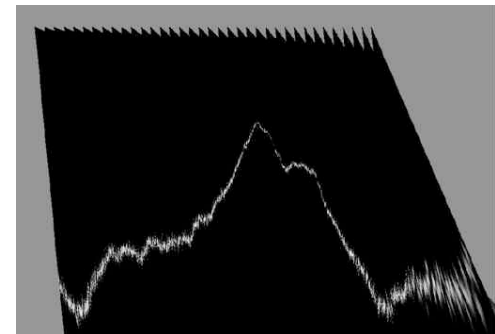
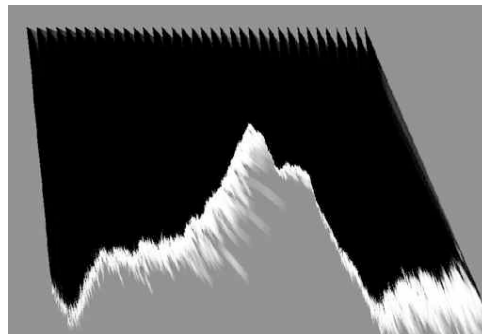
└─ occupancy 3D grid

Simulation results on a 2D profile:

One acquisition



Fusion of 36  
acquisitions



Update strategy

Bayes

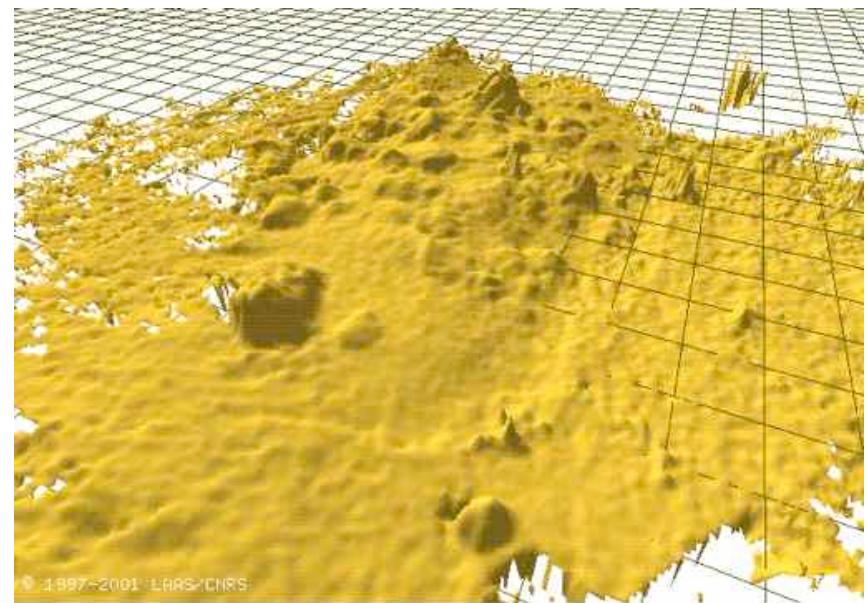
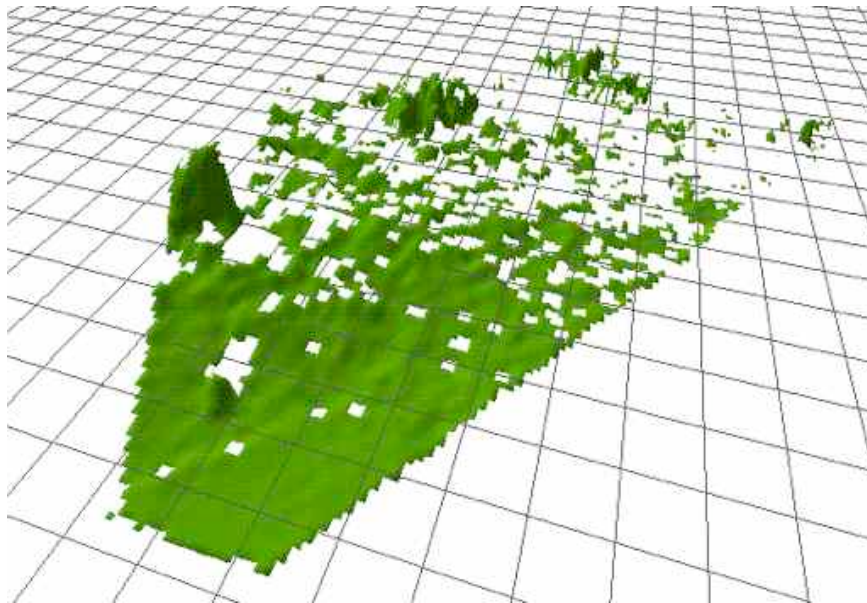
Dempster-Shafer

# Modèle numérique de terrain

3D grids require huge memory

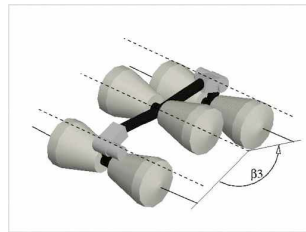
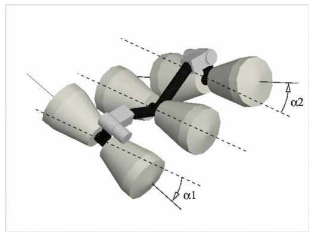
↳ current solutions: 1. Simple averaging  
2. Weighted averaging

Elevations are fused with a weight proportional to their uncertainty



# Modèle de traversabilité : autre approche

Sur la base d'un MNT, évaluation des positions acceptables



# Modèle de traversabilité : autre approche

Sur la base d'un MNT, évaluation des positions acceptables



# Fil conducteur

- Introduction
- Modélisation d'environnement par vision
  - Quels modèles sont nécessaires ?
  - Pourquoi la vision ?
  - Stéréovision
  - Exemples d'algorithmes de modélisation
- Localisation par vision
- Modélisation et localisation simultanées



# Fil conducteur

- Introduction
- Modélisation d'environnement par vision
- Localisation par vision
  - Le problème de la localisation
  - Localisation en robotique mobile
  - Odométrie visuelle
- Modélisation et localisation simultanées