

# Reconstruction 3D géométrie épipolaire et stéréovision

François Cabestaing

Master ASE et Informatique, spécialité "Image Vision Interaction"

# reconstruction 3D

Objectif : tenter de reconstituer la troisième dimension qui a été perdue du fait de la projection sur le plan image.

<i>Mono-caméra</i> Shape from shading, from focus Stéréo-photométrie	<i>Vision active</i> Lumière structurée Laser
<i>Multi-caméra et Vision passive</i> Stéréovision et reconstruction stéréoscopique Space carving, cones généralisés et silhouettes	

## vue unique : shape from focus (from defocus)

Utilisation de l'effet de flou de défocalisation pour estimer la distance ou la forme des objets.

- les propriétés du système optique d'acquisition d'image sont connues : ouverture de l'objectif, profondeur de champ
- si la profondeur de champ est finie, le flou dépend de la distance entre l'objet et la caméra



image réelle



image avec flou simulé

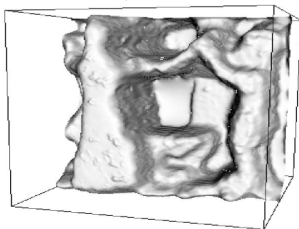
# vue unique : shape from focus (from defocus)



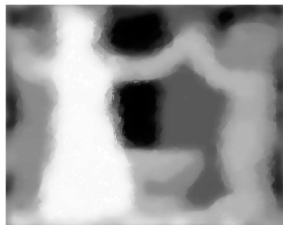
mise au point distante



mise au point proche



reconstruction 3D



carte des distances

[http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/FAVARO1/dfdutorial.html](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/FAVARO1/dfdutorial.html)

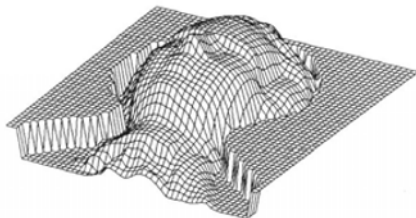
# vue unique : shape from shading

Utilisation des variations d'intensité dans l'image pour estimer la distance ou la forme des objets.

- les conditions d'éclairage de la scène sont parfaitement connues
- les surfaces sont lambertiennes : l'intensité de la lumière réfléchie dépend de son incidence



image initiale



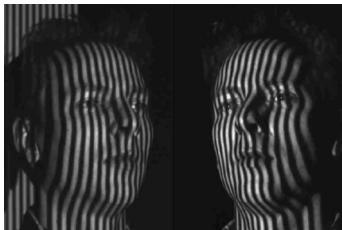
carte de profondeur

E. Trucco, A. Verri : Introductory Techniques for 3-D Computer Vision. Prentice Hall, 1998

## vue unique : lumière structurée

On contrôle la *structure* de la source lumineuse pour faire apparaître le relief des objets.

- nécessite une source de lumière dont les propriétés sont parfaitement connues (ex : faisceau laser)
- l'analyse des lignes de transition noir/blanc permet de reconstituer la forme



visage



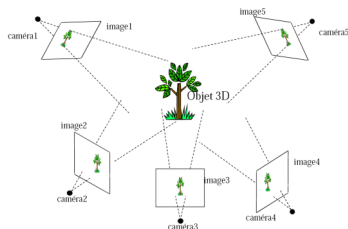
objets divers

[http://en.wikipedia.org/wiki/Structured-light\\_3D\\_scanner](http://en.wikipedia.org/wiki/Structured-light_3D_scanner)

# reconstruction à partir de vues multiples

On exploite les différences existant entre les images acquises depuis différents points de vue.

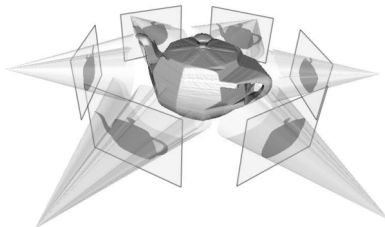
- objets mobiles : on doit acquérir les images à partir de plusieurs caméras donc les paramètres sont connus
- objets fixes : on peut acquérir les images au moyen d'une caméra qui se déplace dans la scène



# reconstruction par intersection de cones

On calcule l'intersection entre les cones généralisés reconstruits à partir des silhouettes des objets.

- ne permet de reconstruire que l'enveloppe convexe de l'objet
- la reconstruction est d'autant plus précise qu'on dispose de nombreux points de vue



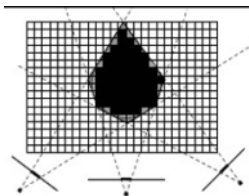
reconstruction à partir de 6 vues



# reconstruction par space carving

On reconstruit la forme 3D en éliminant des voxels qui ne se projettent pas sur la silhouette de l'objet.

- possibilité d'associer aux voxels des propriétés comme la couleur
- inconvénient : la qualité du modèle reconstitué est directement liée au pas de discrétisation 3D



principe du space carving



exemple d'objet reconstitué

# stéréovision

Reconstruction 3D de certains points de la scène à partir de deux images obtenues à partir de points de vue différents.

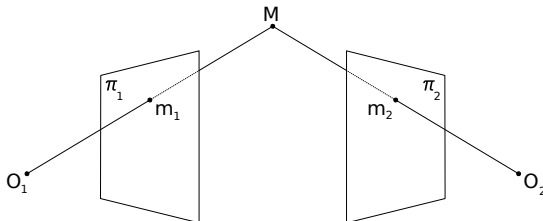
- recherche des *points homologues* dans les deux images par *mise en correspondance*
- reconstruction 3D à partir des propriétés des points homologues

Suite du cours :

- géométrie épipolaire : étude des propriétés des transformations projectives impliquant deux projections perspectives
- techniques de mise en correspondance stéréoscopique

## deux projections perspectives

Un point  $\mathbf{M}$  de  $\mathbb{P}^3$  est transféré sur deux plans projectifs  $\pi_1$  et  $\pi_2$  :



Connaissant les matrices des projections  $\mathbf{P}_1$  et  $\mathbf{P}_2$ , on peut calculer les coordonnées des points  $\mathbf{m}_1$  et  $\mathbf{m}_2$  à partir de celles de  $\mathbf{M}$  :

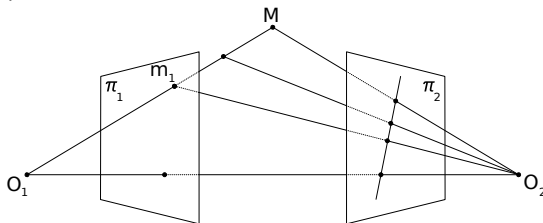
$$\mathbf{m}_1 \equiv \mathbf{P}_1 \mathbf{M} = \mathbf{A}_1(\mathcal{R}_1 \mid \mathbf{t}_1) \mathbf{M}$$

$$\mathbf{m}_2 \equiv \mathbf{P}_2 \mathbf{M} = \mathbf{A}_2(\mathcal{R}_2 \mid \mathbf{t}_2) \mathbf{M}$$

où  $\mathbf{A}_i$  et  $(\mathcal{R}_i \mid \mathbf{t}_i)$  désigneront les matrices intrinsèque et extrinsèque de la caméra  $i$  quand la formation de l'image est modélisée par une transformation projective.

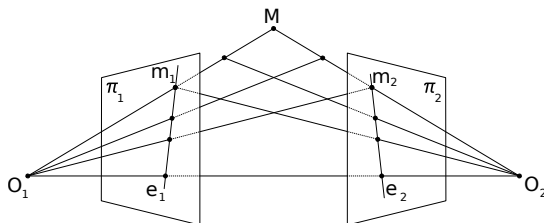
## droites épipolaires

Connaissant un point projeté sur l'un des plans, que peut-on déduire du point projeté dans l'autre ?



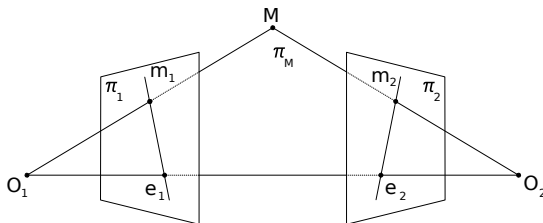
- le point  $m_1$  peut être la projection de n'importe quel point de l'espace situé sur la droite  $O_1 m_1$ .
- la projection de cette droite  $O_1 m_1$  dans le plan projectif  $\pi_2$  est également une droite.
- cette droite est appelée *droite épipolaire* de  $\pi_2$  associée à  $m_1$ .

# épipoles ou points épipolaires



- le centre de projection  $O_1$  appartient à toutes les droites joignant un point  $M$  de l'espace à sa projection  $m_1$  dans  $\pi_1$ .
- la projection  $e_2$  de  $O_1$  dans  $\pi_2$  appartient donc à toutes les droites épipolaires d'un point de  $\pi_1$ .
- les points  $e_1$  et  $e_2$  sont appelés épipoles ou points épipolaires.
- la droite joignant les centres optiques (donc passant par les épipoles) est appelée *droite de base*.

# plan épipolaire et faisceau épipolaire



- le plan épipolaire  $\pi_M$  associé à  $M$  est le plan de l'espace projectif défini par  $M$  et les deux centres de projection.
- les deux droites épipolaires associées aux deux projections de  $M$  sont les intersections de  $\pi_M$  avec les plans de projection.
- le *faisceau épipolaire* est l'ensemble de tous les plans épipolaires.

## produit vectoriel, notation matricielle

Rappel : pour calculer le point d'intersection de deux droites, ou la droite passant par deux points, on utilise un *produit vectoriel* des coordonnées homogènes.

Le produit vectoriel peut être considéré comme une homographie :

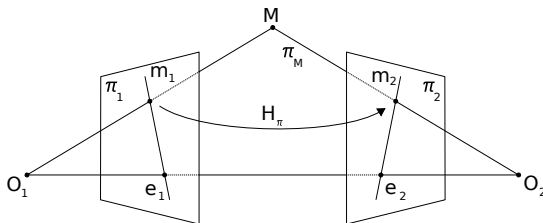
$$\mathbf{p} \times \mathbf{q} = \mathbf{p}^\times \cdot \mathbf{q}, \text{ avec :}$$

$$\mathbf{p}^\times = \begin{pmatrix} 0 & -p_z & p_y \\ p_z & 0 & -p_x \\ -p_y & p_x & 0 \end{pmatrix}$$

- la matrice de l'homographie  $\mathbf{p}^\times$  est de rang 2 (l'un de ses vecteurs propres est nul).
- $\mathbf{p}$  est le vecteur propre de  $\mathbf{p}^\times$  associé à cette valeur propre nulle car :

$$\mathbf{p}^\times \cdot \mathbf{p} = \mathbf{p} \times \mathbf{p} = \mathbf{0}$$

## matrice fondamentale



Il existe une homographie  $\mathcal{H}_\pi$  du plan  $\pi_M$  qui transforme  $\mathbf{m}_1$  en  $\mathbf{m}_2$ .

La droite épipolaire associée à  $\mathbf{m}_1$  est déterminée par le produit vectoriel de  $\mathbf{e}_2$  et de  $\mathbf{m}_2$ , qui est donné par :

$$\mathbf{e}_2 \times \mathbf{m}_2 = \mathbf{e}_2^\times \cdot \mathbf{m}_2 = \mathbf{e}_2^\times \cdot \mathcal{H}_\pi \mathbf{m}_1 = (\mathbf{e}_2^\times \mathcal{H}_\pi) \cdot \mathbf{m}_1$$

La matrice  $\mathbf{F} = \mathbf{e}_2^\times \mathcal{H}_\pi$  est appelée *matrice fondamentale*.



## propriétés de la matrice fondamentale

$\mathbf{F}$  est une matrice homogène de rang 2 définie par 7 degrés de liberté.

Quand  $\mathbf{m}_1$  et  $\mathbf{m}_2$  sont les projections d'un même point  $\mathbf{M}$  de l'espace sur  $\pi_1$  et  $\pi_2$  alors  $\mathbf{m}_2^T \mathbf{F} \mathbf{m}_1 = 0$ .

Droites épipolaires :

- la droite épipolaire de  $\pi_2$  associée à  $\mathbf{m}_1$  est  $\mathbf{d}_2 = \mathbf{F} \mathbf{m}_1$
- la droite épipolaire de  $\pi_1$  associée à  $\mathbf{m}_2$  est  $\mathbf{d}_1 = \mathbf{F}^T \mathbf{m}_2$

Les épipoles vérifient :  $\mathbf{F} \mathbf{e}_1 = 0$  et  $\mathbf{F}^T \mathbf{e}_2 = 0$ .

Calcul à partir des matrices de projection  $\mathbf{P}_1$  et  $\mathbf{P}_2$

$\mathcal{H}_\pi = \mathbf{P}_2 \mathbf{P}_1^+$ , où  $\mathbf{P}^+$  désigne la pseudo-inverse d'une matrice  $\mathbf{P}$ , donc :

$\mathbf{F} = (\mathbf{P}_2 \mathbf{O}_1)^\times \mathbf{P}_2 \mathbf{P}_1^+$ , où  $\mathbf{P}_2 \mathbf{O}_1$  est la projection de  $\mathbf{O}_1$  sur  $\pi_2$

# formation des images stéréoscopiques

Dans le cas de la géométrie projective, une *infinité* de points de l'espace se trouvent projetés sur un unique point du plan.

Dans le cas de la formation d'une image, en général chaque point d'une image correspond à la projection *d'un seul* point de la scène visible par la caméra.

Dans le cas de la stéréovision, un point de la scène peut être :

- visible par les deux caméras
- masqué pour les deux caméras
- visible pour une seule caméra et masqué pour l'autre

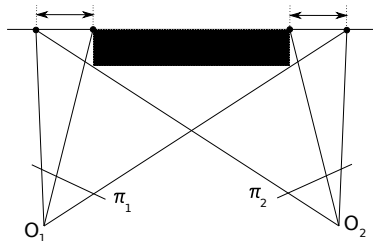
Dans ce dernier cas, on dit que le point est *occulté* pour l'une des deux caméras.

Attention : on parle d'*occultation* en français, le terme *occlusion* est un anglicisme incorrect.

# points homologues

Deux points image  $\mathbf{m}_1$  et  $\mathbf{m}_2$  sont dits *homologues* s'ils correspondent aux deux projections d'un même point  $\mathbf{M}$  de la scène sur les deux images.

Un point d'une image correspondant à un point de la scène qui est visible par sa caméra, mais qui est invisible pour l'autre caméra, ne possède pas d'homologue.



# reconstruction stéréoscopique

La reconstruction stéréoscopique est le procédé qui permet, en déterminant des paires de points homologues dans les deux images, d'estimer par triangulation la troisième dimension.

Étapes de la reconstruction stéréoscopique :

- 1 identification de points ou de primitives d'intérêt (*features*)
- 2 recherche de correspondance entre les primitives
- 3 reconstruction 3D par triangulation

On distingue les méthodes de reconstruction :

- *éparses*, dans lesquelles uniquement quelques points sont utilisés pour la mise en correspondance
- *denses*, dans lesquelles on recherche l'homologue de tous les points des images et on identifie les zones occultées

## exemple de méthode éparse (TP)

### Détection de points d'intérêt

- détection de points particuliers dans les images : les *coins*
- détecteur de *Shi & Tomasi* qui recherche les pixels pour lesquels la fonction image a une forte courbure locale.

### Mise en correspondance

- estimation d'une *distance* généralisée entre chaque paire possible de points (*gauche, droite*)
- la distance généralisée comprend :
  - un terme de distance euclidienne entre ce point et la droite épipolaire de son correspondant potentiel
  - un terme lié à la similarité des niveaux de gris dans les deux voisinages des images centrés sur ces deux points
- recherche de correspondance par traitement du tableau des distances généralisées

# bibliographie

## Éléments de cours

- Séance "tutorial" de Richard Hartley and Andrew Zisserman lors de la conférence CVPR'99.

<http://users.cecs.anu.edu.au/~hartley/Papers/CVPR99-tutorial/tutorial.pdf>

- Frédéric Devernay, INRIA Grenoble, cours "Vision par ordinateur 3-D".

<http://devernay.free.fr/cours/vision/>

- Cours de Sébastien Roy sur la vision 3D, Université de Montréal.

[http://www.iro.umontreal.ca/~roys/fr\\_ift6145H07.shtml](http://www.iro.umontreal.ca/~roys/fr_ift6145H07.shtml)

## Articles de chercheurs

- C. Harris and M. Stephens, *A combined corner and edge detector*. Proceedings of the 4th Alvey Vision Conference, 1988.
- J. Shi and C. Tomasi, *Good Features to Track*. 9th IEEE Conference on Computer Vision and Pattern Recognition, 1994.