

Vision par ordinateur

Introduction au domaine

Applications

Frédéric Devernay

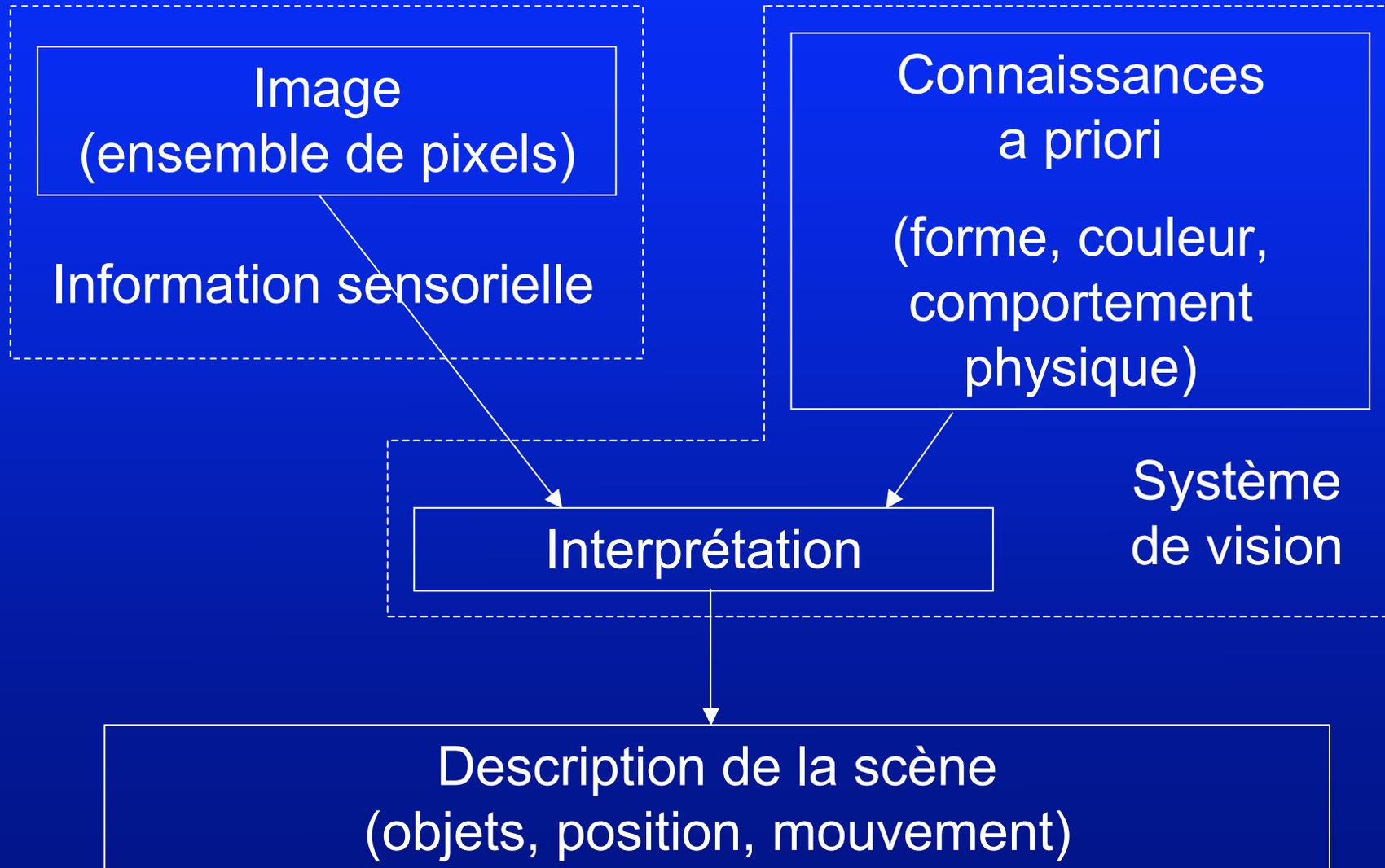
Qu'est-ce que la vision ?

- Le monde a une structure 3D et est composé d'objets
- L'être humain sait parfaitement décrire et interpréter ce monde
- Pourtant, l'information disponible sur la rétine n'est qu'un ensemble de points (environ un million de *picture elements* ou *pixels*)...

Qu'est-ce que la vision (2)

- Chaque pixel contient des informations sur la lumière (quantité et contenu spectral/couleur) reçue en ce point de la rétine
- Les objets (téléphone, voiture...) n'existent pas sur la rétine, et pourtant on les voit : leur interprétation est le résultat du processus visuel.

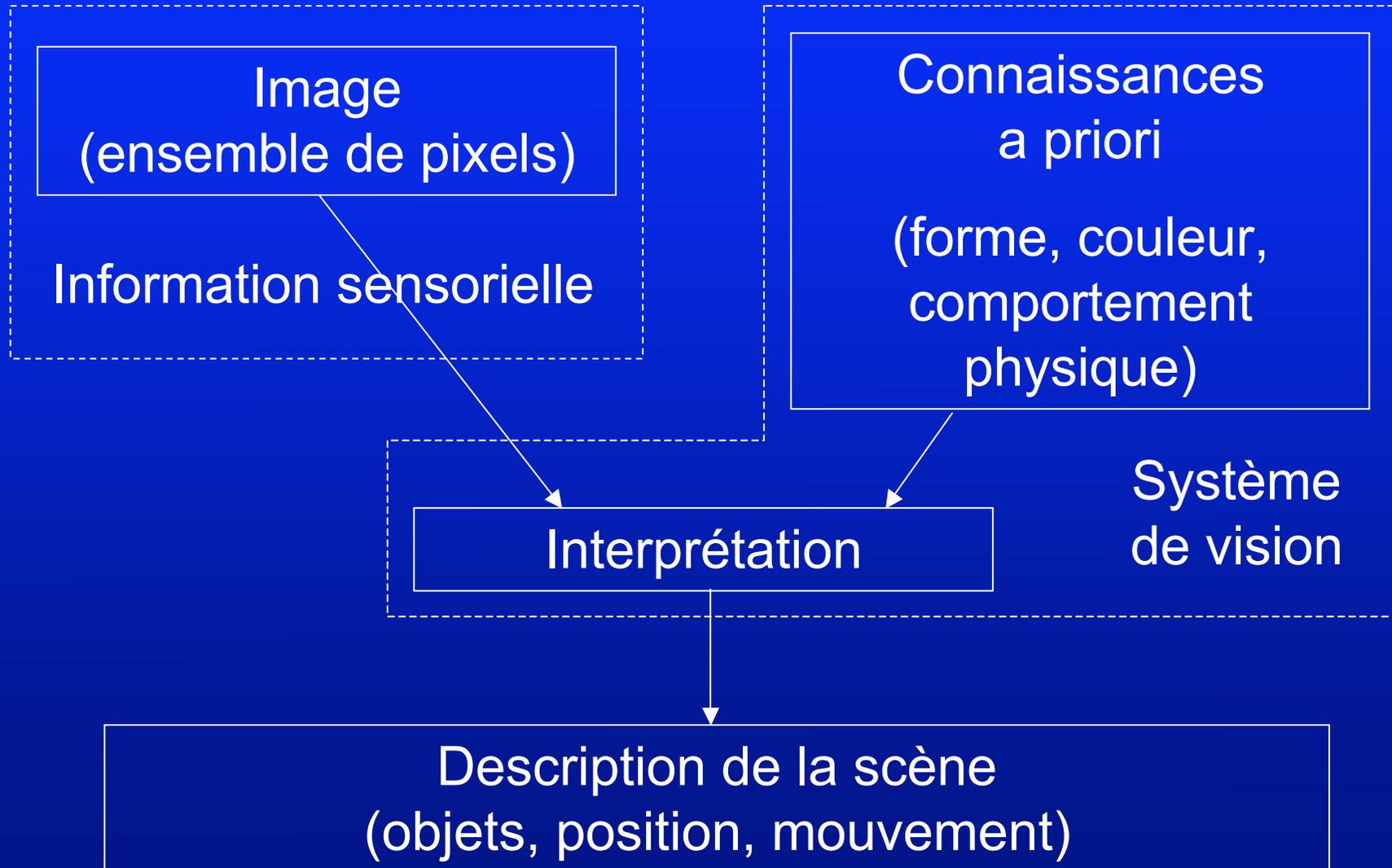
Qu'est-ce que la vision (3)



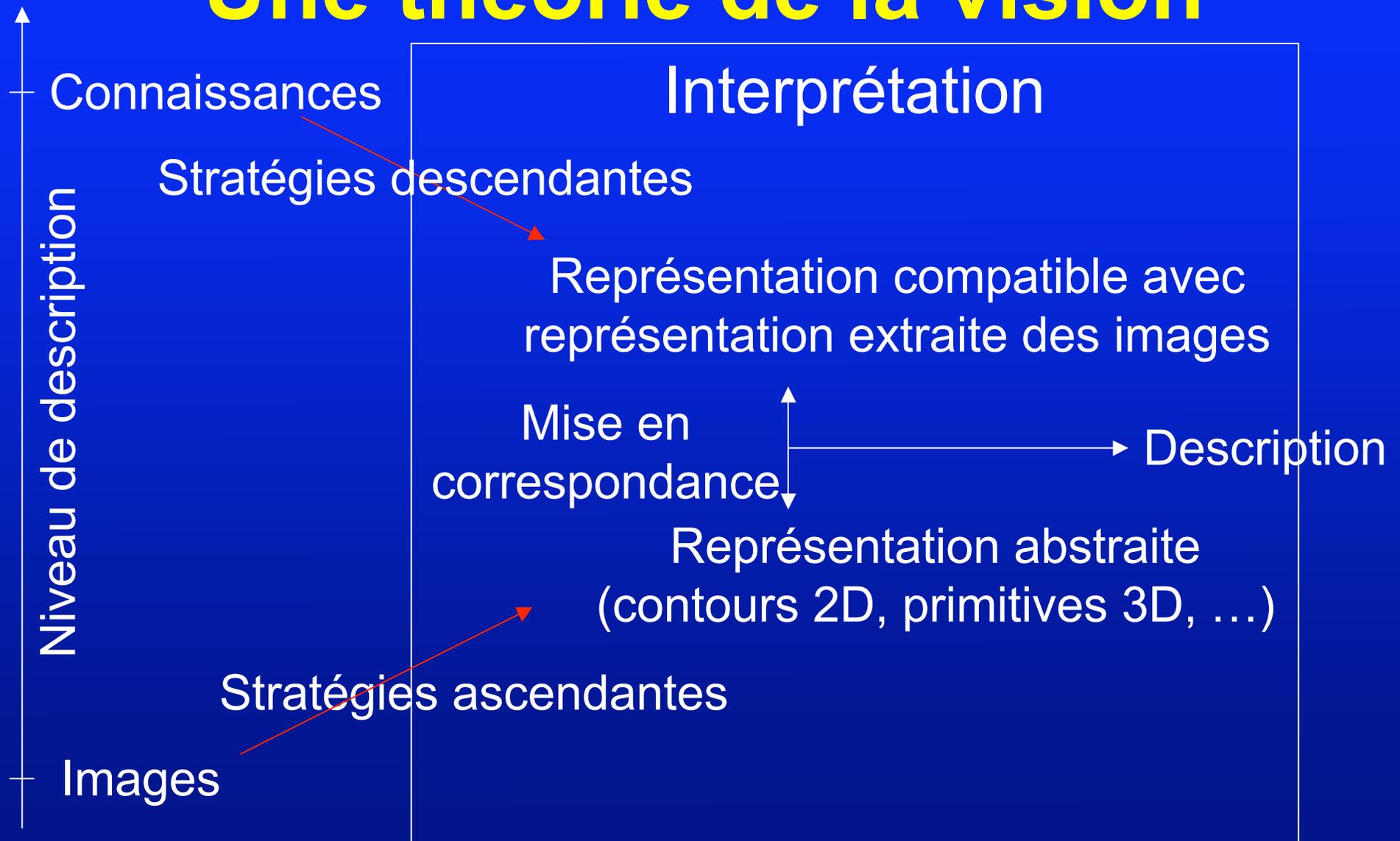
Résoudre le problème de la vision ?

- La vision humaine est *extrêmement complexe* (neurosciences)
- La vision par ordinateur ne cherche pas à comprendre ou à reproduire la vision humaine, mais à construire un *modèle algorithmique* qui, vu de l'extérieur, possède des propriétés semblables.
- De plus, un problème de vision par ordinateur correspond souvent à un *sous-ensemble* du système de vision humain.

Une théorie de la vision



Une théorie de la vision



Une théorie de la vision

Connaissances a priori

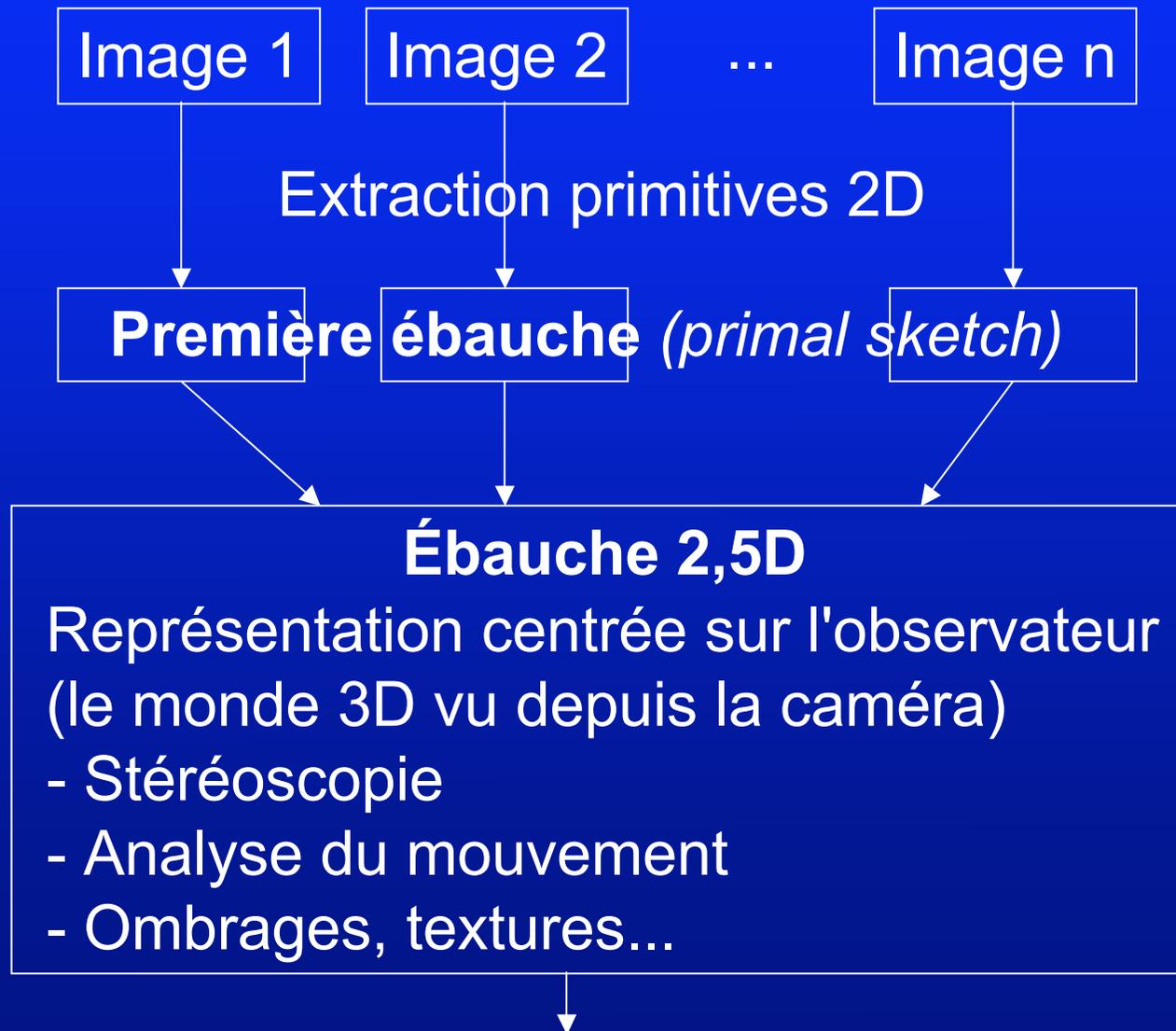
- Physiques (optique, forces/interactions)
- Géométriques (forme des objets, formation de l'image)
- Sémantiques (fonction des objets, comportement)

Exemple : que sait-on de *mesurable* sur une voiture qui permette de l'identifier dans une image ?

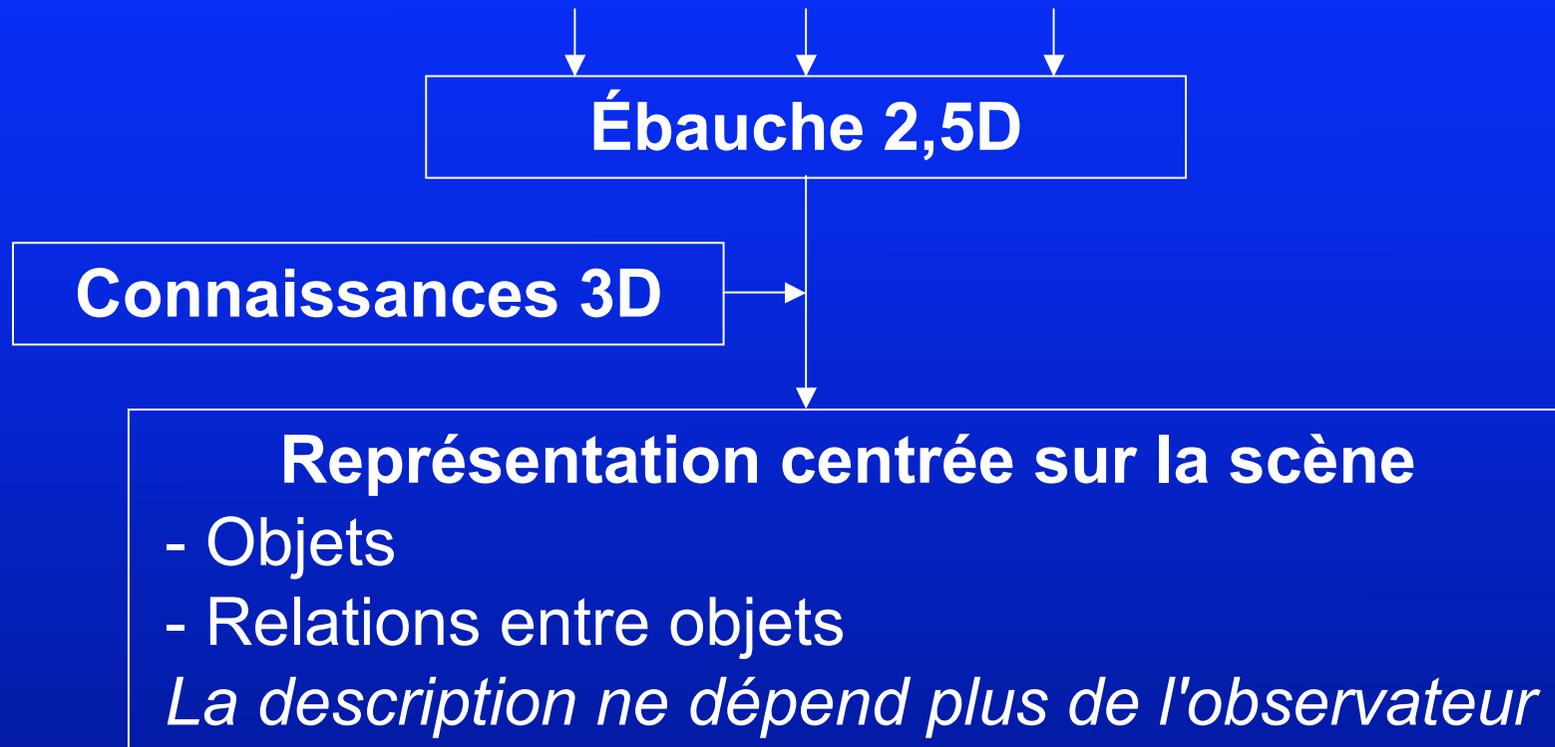
Une théorie de la vision

- Les **algorithmes** de vision sont des expressions des stratégies mises en œuvre :
 - extraction de primitives à partir des images
 - représentation des connaissances
 - mise en correspondance images/connaissances : *reconnaissance*
- **Un système de vision doit être fiable et rapide**

Le paradigme de David Marr



Le paradigme de David Marr



Paradigme datant de la fin des années 1970,
mais toujours d'actualité !

Segmentation, reconstruction, reconnaissance

- Traduction en algorithmes des étapes de Marr
- Tout part des images :
 - optiques (2D) : photo, satellite, document
 - médicales (Rayons X, IRM, 2D/3D...)
 - autres (radar, laser, sonar...)
 - images fixes ou séquences temporelles

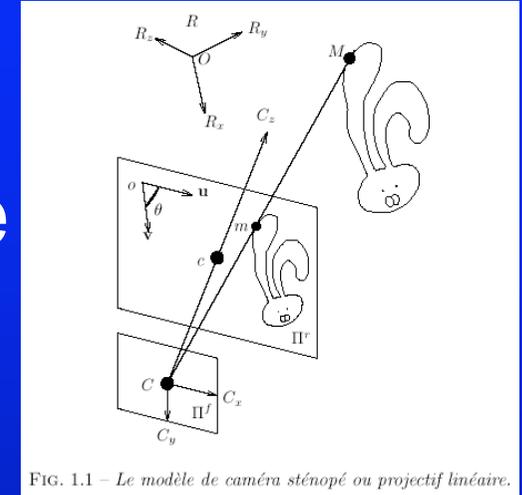
Segmentation

- Détection de points d'intérêt (coins...)
- Détection de contours
 - correspondent à des discontinuités des niveaux de gris
 - approximation par des représentations analytiques (polygones, coniques...)
- Extraction de régions (zones homogènes)

Les approches "contours" et "régions" sont duales : frontières de régions \Leftrightarrow contours

Reconstruction

- Commence par le calibrage de l'imageur
 - détermination de la fonction : point 3D & caractéristiques physiques → point image & intensité/couleur
 - revient à déterminer les paramètres d'un modèle de caméra/capteur (position, orientation, focale...)

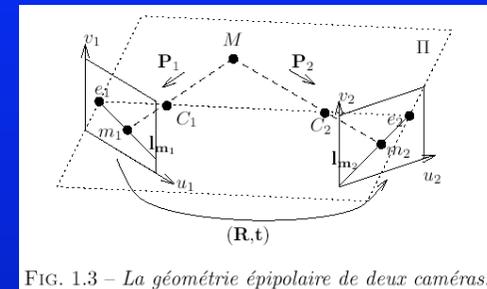


Reconstruction

- Des méthodes variées (*shape from...*) :

- Stéréoscopie

- procède par appariement de primitives, puis triangulation



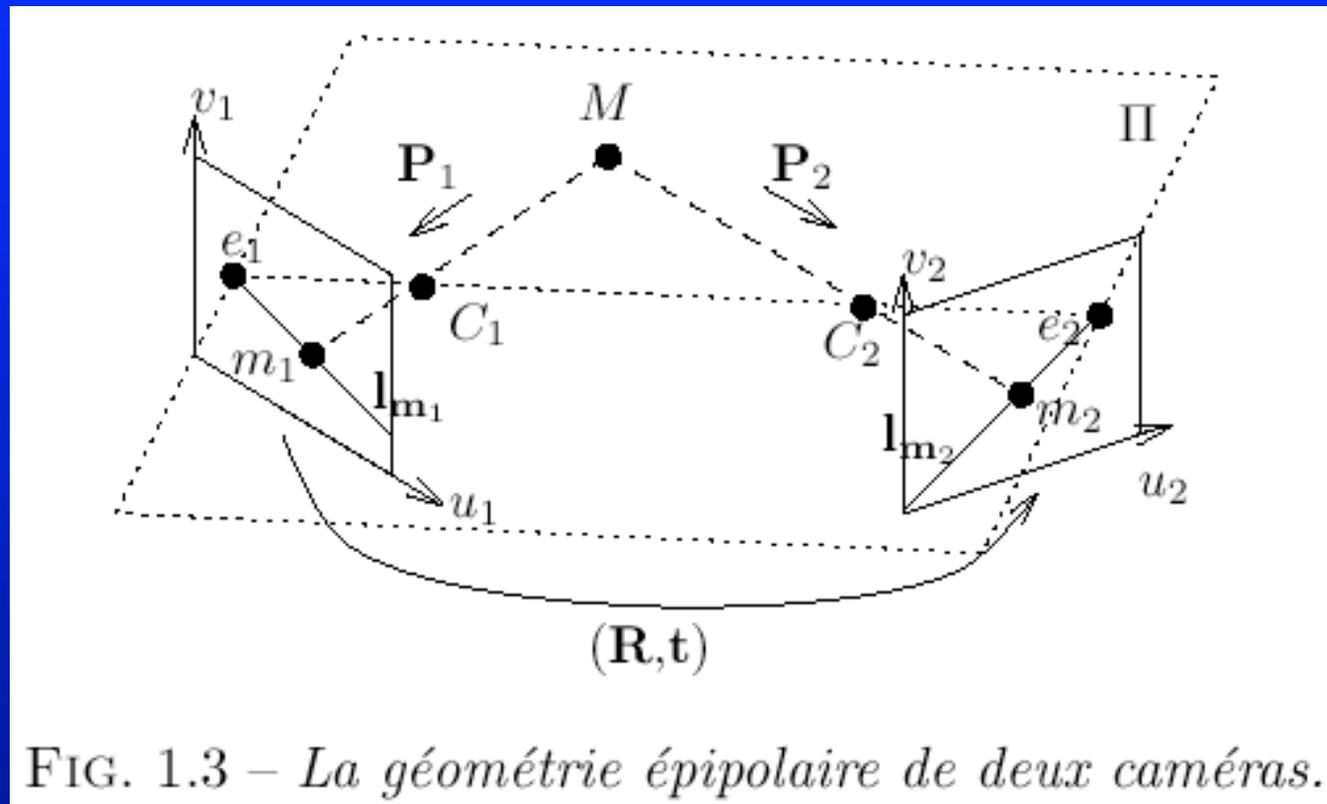
- Mouvement (structure from motion)

- on ne peut pas forcément déduire la structure 3D de la scène d'un mouvement de caméra (un panoramique ne contient pas d'infos 3D)

- Ombrage (shape from shading)

- seulement sur objets uniformes et/ou avec éclairage contrôlé

Reconstruction



- Stéréoscopie

Quelques exemples concrets

- Reconnaissance d'écriture (OCR)
- Contrôle visuel sur chaîne de fabrication
- Asservissement visuel de robots
 - une caméra observe l'évolution du robot et le renseigne sur sa position par rapport à la cible

Quelques exemples concrets

- Robotique mobile
 - évitement d'obstacles
 - découverte d'un environnement inconnu et construction d'une carte (exploration)
 - conduite automatique / assistée
- Exemples :
 - mission Mars Pathfinder
 - conduite "sans les mains" (A. Broggi, U Parma)

Quelques exemples concrets

- Modélisation à partir d'images
- Rendu à partir d'images
- Post-production cinématographique (objets virtuels dans scènes réelles et vice-versa)
 - [PhotoModeller](#)
 - [Résultats de Marc Pollefeys](#)
 - [système EyeVision au Superbowl'01 \(CMU\)](#)
 - Panoramas, mosaïques d'images