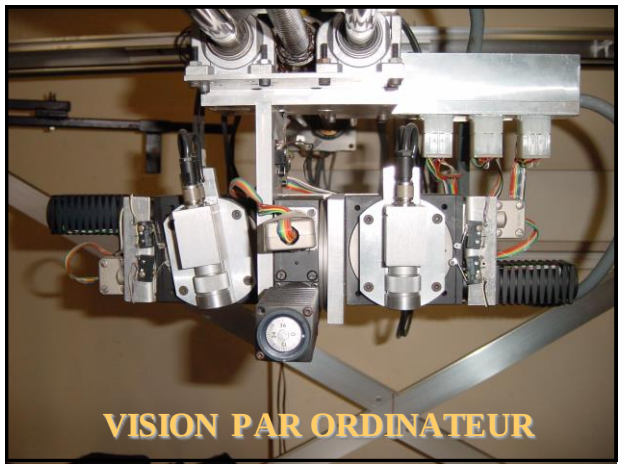


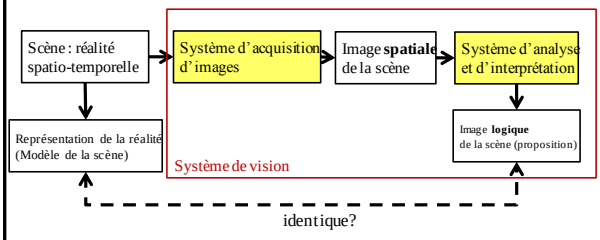
Introduction

Image

- **Image:** Représentation de la **réalité**
  - Il y a plusieurs types d'images, chacun ayant une forme de représentation particulière
    - image spatiale 2D ou 3D
    - séquence temporelle d'images spatiales
    - image acoustique (signal acoustique) …
    - image logique (sa forme de représentation est une proposition)
  - Chaque image représente la réalité dont elle a la forme
    - une image spatiale représente les propriétés et les relations spatiales des objets de la scène qui est imagée
  - Toute image (à l'exclusion des images logiques) doit être « interprétée », pour acquérir un « sens », i.e. traduite en une image logique (une proposition) décrivant un « fait », l'existence d'une combinaison d'« objets » ayant certaines propriétés.

Système de Vision

- Qu'est-ce donc qu'un système de vision ?



Vérité ou fausseté d'une image logique

- D'un point de vue *philosophique*:
  - « L'image (*interprétée grâce au système de vision, l'image logique donc*) s'accorde ou non avec la réalité; elle est correcte ou incorrecte, vraie ou fausse » (Wittgenstein, Tractatus Logico-Philosophicus, 2.21, 1922)
  - «C'est dans l'accord ou le désaccord de son sens avec la réalité que consiste sa vérité ou sa fausseté » (Wittgenstein, TLP, 2.222)
- Traduction *informatique*:
  - L'information sur la réalité (le sens), extraite de l'image par le système de vision, peut être **vraie** ou **fausse**.
  - Lorsque l'information est fausse, on parle d' **illusion visuelle**

Vision

image spatiale colorée de la scène

Vision →

Un lustre éteint de style empire
possédant 4 ampoules
est suspendu à un plafond et projette
son ombre

image logique de la scène
(proposition)

- **Vision :** Saisie de la réalité à partir de systèmes d'acquisition d'images (yeux, caméras). **La vision donne un sens aux images spatiales.**
  - + particulièrement, description des phénomènes spatio-temporels à partir de leurs images
  - +particulièrement, description de la géométrie 3D des objets d'une scène et de leurs relations spatiales à partir de leurs images (vision géométrique statique)

Image (spatiale) numérique: objet informatique

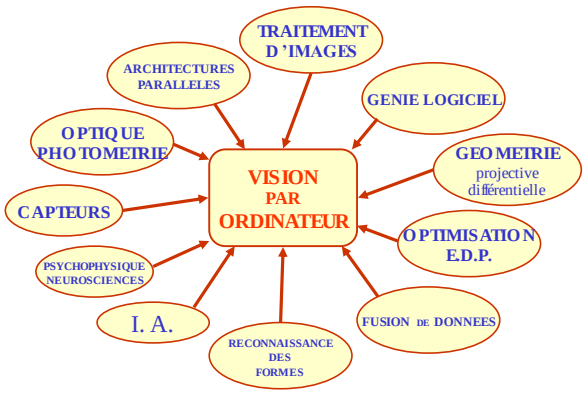
- Image (spatiale) numérique : tableau de « pixels » (éléments d'image) contenant une **information** relative à la scène réelle qui est imagée.
  - Dans le cas d'une image optique, chaque pixel contient un nombre entier qui est une mesure quantifiée de l'éclairement de ce pixel (le nombre de photons reçus par ce pixel durant une période de temps donné dans une certaine bande de fréquences)
- Cette image devra être « traitée » et « interprétée » par le système de vision pour reconstruire la réalité dont elle est l'image
- Un système de vision est donc un système de traitement de l'information

SYSTEMES DE VISION
ACTIF vs PASSIF

- S.V. **ACTIF** : possède sa propre source de lumière
  Exemple : radar, satellite ERS

- S.V. **PASSIF** : utilise une source de lumière externe
  Exemple : système visuel humain, satellite SPOT

PLURIDISCIPLINARITE DE LA V.O.



Le système visuel humain

- un système **passif** (ne disposant pas d'une source d'éclairage)
  - Certains philosophes de l'antiquité grecque pensaient que le svh était un système actif : les yeux émettaient des corpuscules (en présence d'une source de lumière) qui venaient « toucher » les objets (cette approche a été invalidée par l'avènement de la photographie : lorsqu'on observe une photo (2D) on perçoit le relief.
- un système comportant
  - un sous-système d'acquisition d'images : 2 yeux
  - un sous-système d'analyse et d'interprétation des images rétiniennes situé essentiellement dans le cortex visuel
- un **système de traitement de l'information** (David Marr, 1980) où l'information est contenue dans les signaux lumineux captés par les rétines des deux yeux, les « images rétiniennes »
- Le système visuel humain est très performant, à la fois rapide et fiable…d'où l'intérêt de comprendre son fonctionnement…dans le but de construire des systèmes de vision artificielle !

Modèle mathématique

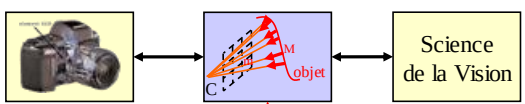
Modèle mathématique: description en termes logiques et/ou mathématique de phénomènes ou d'objets réels (physiques)

La Science construit des modèles à partir des lois de la Nature (physique, biologique, économique, …).
Il faut distinguer un modèle de la réalité qu'il est censé représenter !!
Un modèle est une interface entre la Réalité et la Science : la Science atteint la réalité grâce à ses modèles …

Modèle mathématique

Modèle mathématique: description en termes logiques et/ou mathématique de phénomènes ou d'objets réels (physiques)

• La Science construit des modèles mathématiques à partir des lois de la Nature (physique, biologique, économique, …).
• Il faut distinguer un modèle de la réalité qu'il est censé représenter !!
• Un modèle est une interface entre la Réalité et la Science : la Science atteint la réalité grâce à ses modèles …



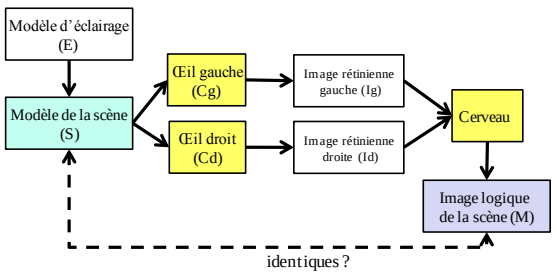
attention ! Il s'agit ici du modèle **fonctionnel** et non pas du modèle physique (géométrique, radiométrique) de l'objet caméra !

Le système visuel humain

- Comment le système visuel humain peut-il calculer une image logique de la scène qui soit faible ?
  - Notre cerveau reçoit des informations du monde : ce sont les images qui se forment sur la rétine (« rétine » vient de rete, filet en latin). A partir de ces images rétiniennes, il doit calculer une image logique (un modèle) de la scène.
  - Pour que notre cerveau puisse calculer un modèle de la scène qui soit fiable, il doit connaître le processus (modèle) de formation des images rétiniennes. Il devra en être de même pour tous système naturel ou artificiel.
  - La modélisation de la formation des images rétiniennes comprend la modélisation des différents composants du système de vision savoir :
    - la modélisation de la scène
    - la modélisation de l'éclairage
    - la modélisation du système d'acquisition de l'image
    - la modélisation de l'image spatiale

    Ainsi que les relations (interactions) existant entre ces différents modèles

Modèle du système visuel humain



- Il ne faut pas confondre:
  - Le modèle de la scène S : représentation de la réalité
  - L'image fournie par le système d'acquisition d'images I (ici l'image rétinienne Ig ou Id)
  - L'image logique de la scène M calculée par le cerveau

Méthodologie pour un système de vision

- F : modèle de la formation des images spatiales
- E : modèle d'éclairage
- S : modèle de la scène
- C = {C1, C2,...} : modèles fonctionnels des systèmes d'acquisition d'images spatiales
- I = {I1,I2,...} : modèles des images spatiales
- M : image logique (proposition) de la scène calculée par le système de vision
- Equation fondamentale : I = F(S,E,C) (eq. 0)
- Equation que doit résoudre le système de vision : trouver M* tel que

  $$I = F(M^*,E,C) \text{ (eq. 1)}$$

  Soit M* une solution de l'eq. 1 alors d'après l'eq. 0 :

  $$F(M^*,E,C) = F(S,E,C) \text{ (eq. 2)}$$

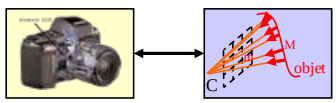
  Cependant l'eq. 2 n'implique pas que M* = S !! (possibilité de solutions multiples)

Méthodologie pour un système de vision

- *Définition* : on appelle « modèle possible de la scène » tout modèle M* tel que I = F(M*,E,C) (eq. 1)
- *Théorème* : Tout système de vision doit calculer un modèle possible de la scène
- *Remarques*
  - Dans un système de vision, tout modèle est composé d'un modèle géométrique et d'un modèle radiométrique
  - Dans un système de vision tout modèle comporte des paramètres dont certains sont connus et d'autres non. Le but d'un système de vision est de découvrir les paramètres inconnus du modèle de la scène

Modèlisation et Calibrage

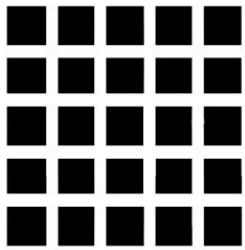
Un modèle dépend toujours d'un certain nombre de **paramètres**.
L'opération de **calibrage** ou d'**étalonnage** consiste à déterminer ces paramètres
Exemple : les paramètres du modèle fonctionnel d'une caméra sont : la focale, le nombre de pixels, les dimensions des pixels, le temps de pose, …



ILLUSIONS VISUELLES

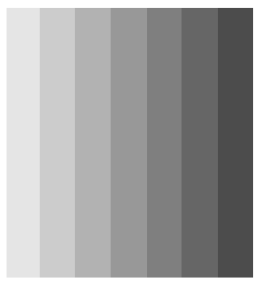
- Le but poursuivi est d'étudier les illusions visuelles pour comprendre le système visuel humain

GRILLE D 'HERMANN

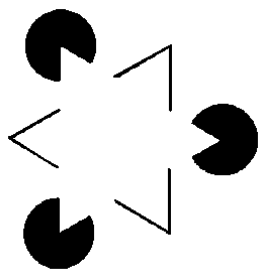


BANDES DE MACH

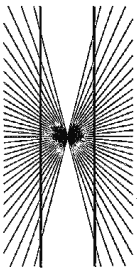
ILLUSION DE RENFORCEMENT DU CONTRASTE



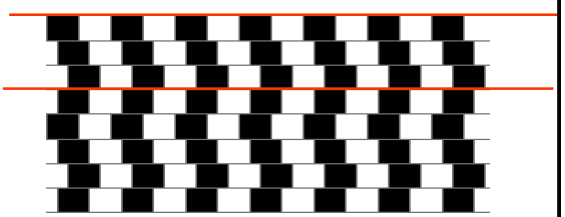
CONTOURS SUBJECTIFS



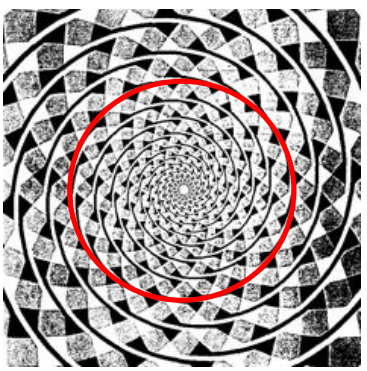
ILLUSION D ’HERING



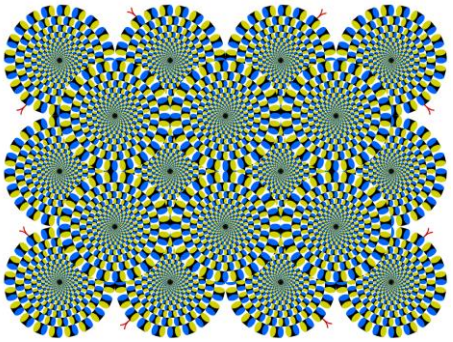
ILLUSION DU MUR



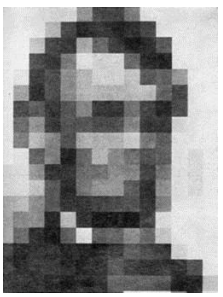
SPIRALE DE FRASER



SPIRALE DE FRASER

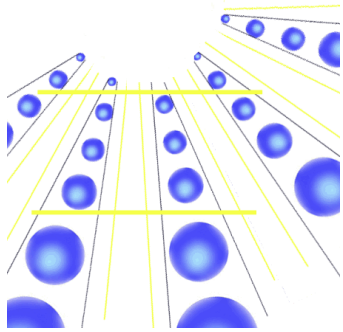


MULTIRESOLUTION

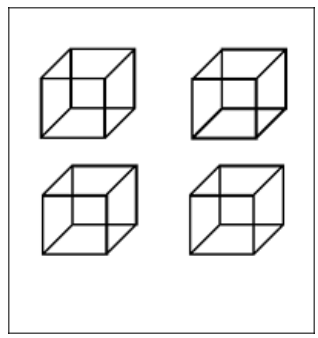


VISION 3D MONOCULAIRE

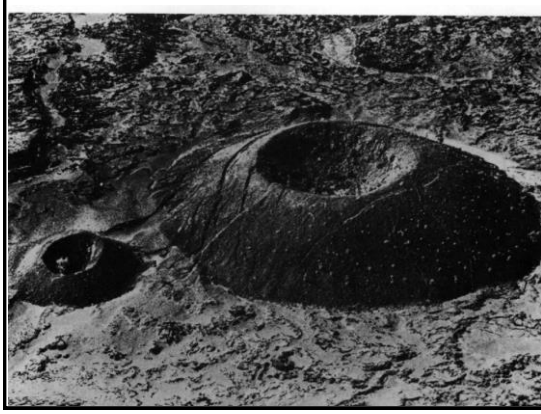
VISION 3D MONOCULAIRE



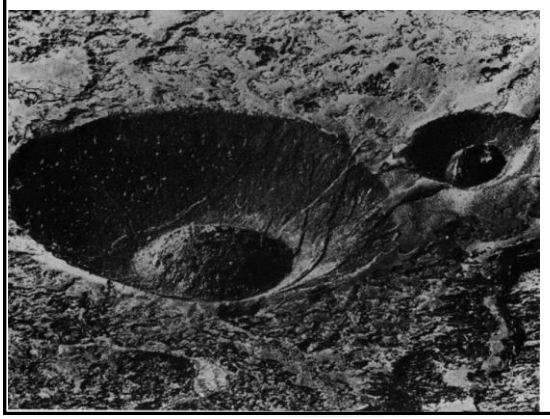
CUBES DE NECKER



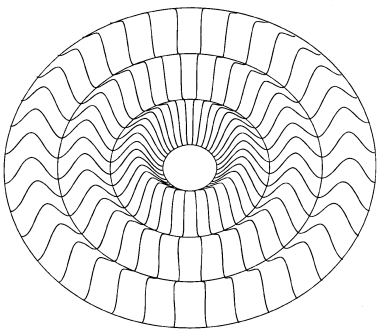
ILLUSION DU CRATERE



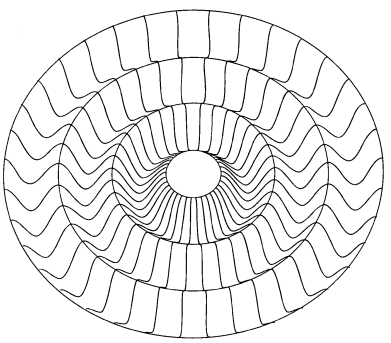
ILLUSION DU CRATERE



SURFACE AMBIGUË



SURFACE AMBIGUË

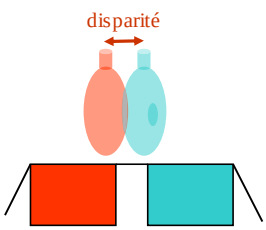


INVERSION DE RELIEF

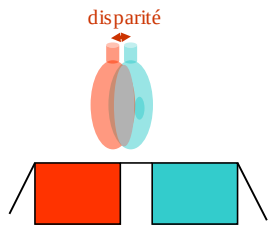
ANAGLYPHE

- Couple de photographies stéréoscopiques en deux couleurs complémentaires

PRINCIPE DE L 'ANAGLYPHE

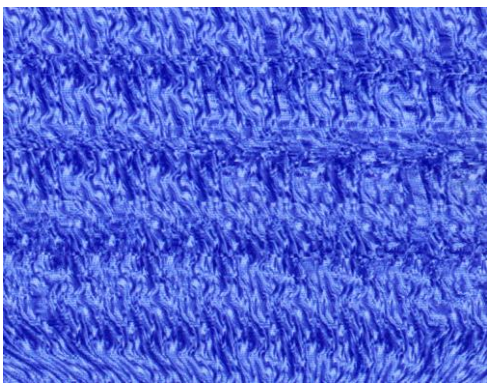


Objet ELOIGNE



Objet RAPPROCHE

ANAGLYPHE



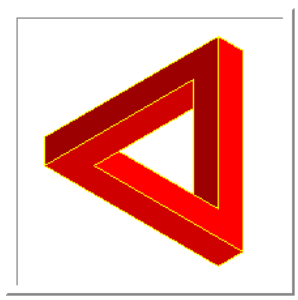
CAMOUFLAGE

FIGURE AMBIGUË

FIGURE AMBIGUË

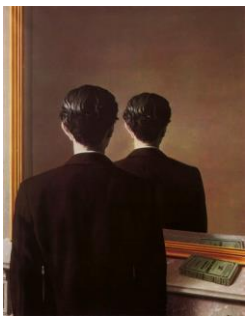
VASE DE RUBIN

TRIANGLE IMPOSSIBLE



Triangle impossible

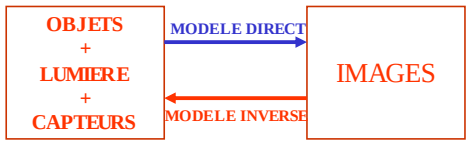
Un tableau de Magritte



Applications de la Vision par ordinateur

- **Robotique**
  - Chaîne d'assemblage, guidage, détection d'obstacles …
- **Métrologie**
  - Contrôle dimensionnel, inspection, …
- **Télédétection**
  - Construction de Modèles Numériques de Terrain, analyse de scènes, …
- **Imagerie médicale**
- **Vidéo-surveillance**
- **Réalité augmentée**
- …

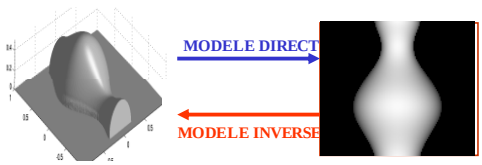
VISION vs SYNTHESE D'IMAGES



•Appliquer le **modèle direct** (de formation d'une image) revient à calculer l'image numérique du modèle d'une scène, en utilisant les relations physiques existant entre les modèles des objets de la scène, de l'éclairage et de la caméra: **SYNTHESE D'IMAGES**

• Appliquer le **modèle inverse** revient à calculer le modèle d'une scène à partir d'un modèle d'éclairage, d'un modèle de caméra et d'un modèle de l'image (ou de modèles de plusieurs images) en utilisant le modèle direct : **VISION PAR ORDINATEUR**

VISION vs SYNTHESE D'IMAGES



•Appliquer le **modèle direct** (de formation d'une image) revient à calculer l'image numérique du modèle d'une scène, en utilisant les relations physiques existant entre les modèles des objets de la scène, de l'éclairage, de la caméra et de l'image numérique: **SYNTHESE D'IMAGES**

• Appliquer le **modèle inverse** revient à calculer le modèle d'une scène à partir d'un modèle d'éclairage, d'un modèle de caméra et d'un modèle de l'image (ou de modèles de plusieurs images) en utilisant le modèle direct : **VISION PAR ORDINATEUR**

Image réelle et Image de synthèse

Image réelle

Image de synthèse

Réalité augmentée

ou comment rendre visible l'invisible !!

Image réelle

Vision + Synthèse d'images

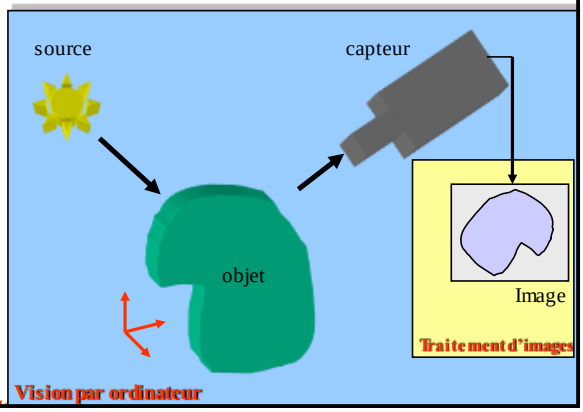
Image de synthèse

Exemple d'application : **système Karma** : cf . **http://www.loria.fr/~gsimon/ra/**

Des **éléments graphiques** ainsi que du **texte** apparaissent par-dessus les éléments réels d'une photocopieuse, permettant d'indiquer à l'utilisateur la suite des opérations à suivre.
Le même principe a été adopté pour l'assemblage de systèmes de fermeture de portes de voitures au Fraunhofer-Institut für Graphische Datenverarbeitung (IGD) ou de fils électriques chez Boeing.

Méthodologie

Formation de l'image



Méthodologie de la vision par ordinateur

1. **Apprentissage (off-line)**
   1. **Modélisation**
      1. Détermination des modèles des composants du système de vision: éclairage, objets et images de la scène, caméra(s)
      2. Détermination des relations entre ces modèles aboutissant au **modèle de la formation d'une image**
   2. **Calibrage** (ou étalonnage)
      1. Détermination de certains paramètres des modèles précédents
2. **Résolution des équations (on-line)**

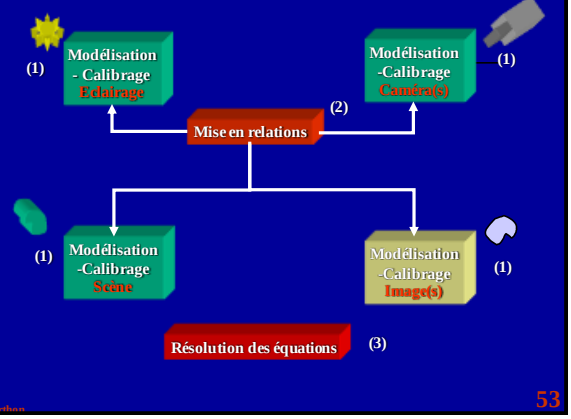
   Détermination des paramètres des modèles des objets et de la scène par ajustage de l'image de la scène à l'image réelle

2 méthodes sont possibles:
- algorithmique : trouver l'algorithme de résolution
- systémique : trouver le réseau de « neurones » résolvant les équations

Remarque: chaque modèle est composé d'un modèle géométrique et d'un modèle radiométrique.

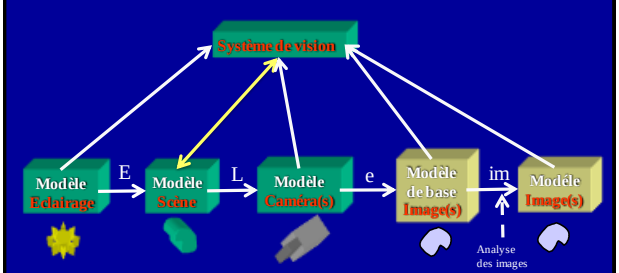
2009 Ph. Marthon 52

Méthodologie de la V.O.



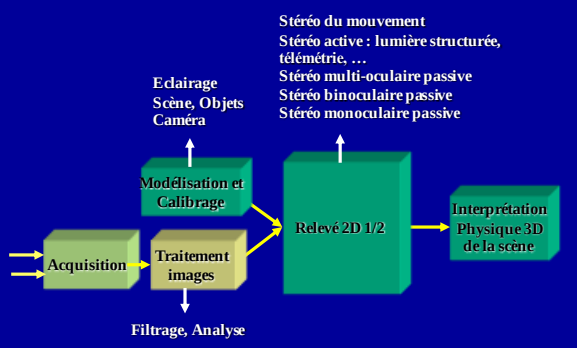
2009 Ph. Marthon 53

Représentation systémique de la vision



2009 Ph. Marthon 54

« Une » chaîne type de vision



Plan du cours

1. **Apprentissage**
   Modélisation et calibrage
   1. Système d'acquisition d'images
   2. Scène
   3. Eclairage
   4. Images
   5. Formation des images
   6. Relations entre images d'une même scène
2. **Résolution des équations**
   Calcul des paramètres des modèles d'objets et de la scène
   1. Stéréovision monoculaire
   2. Stéréovision binoculaire
   3. Stéréovision du mouvement
   4. Stéréovision active
   5. Reconnaissance et interprétation

Etape 1
APPRENTISSAGE

APPRENTISSAGE

SYSTEME D'ACQUISITION D'IMAGES

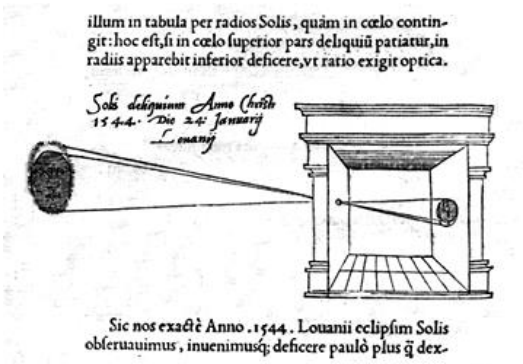
APPRENTISSAGE

MODELISATION GEOMETRIQUE

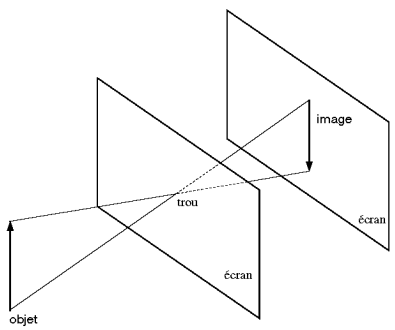
Modèle, calibrage, étalonnage géométrique de caméra

Relation entre pixels et rayons ?

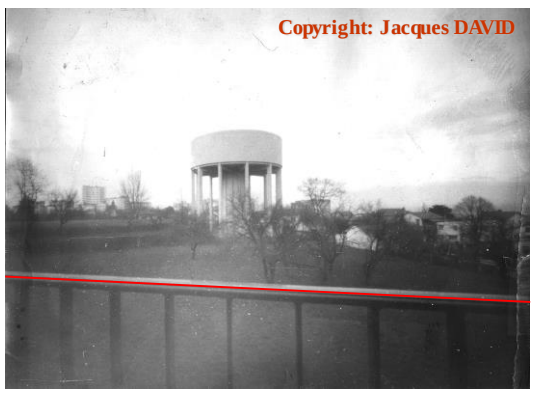
CAMERA EN TROU D 'EPINGLE (STENOPE)



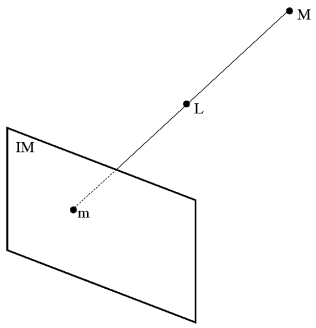
CAMERA EN TROU D 'EPINGLE (STENOPE)



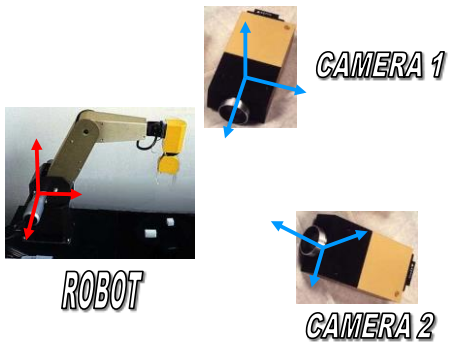
STENOPE



CAMERA EN TROU D 'EPINGLE

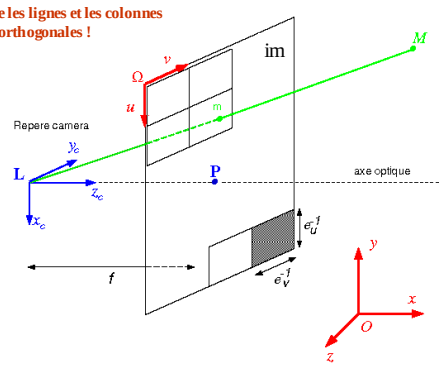


REPERES



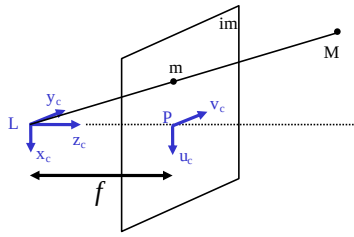
CAMERA EN TROU D 'EPINGLE
MODELISATION

On suppose que les lignes et les colonnes des pixels sont orthogonales !



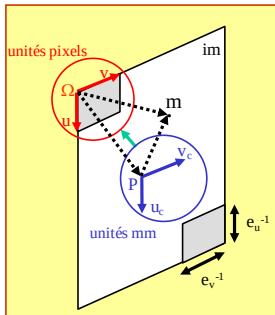
CAMERA EN TROU D 'EPINGLE
MODELISATION

$$\exists \lambda > 0 \,\big|\, L\vec{m} = \lambda\, L\vec{M}$$



$$u_c = f\,\frac{x_c}{z_c} \qquad v_c = f\,\frac{y_c}{z_c} \quad \Longleftrightarrow \quad \begin{bmatrix} wu_c \\ wv_c \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix}$$

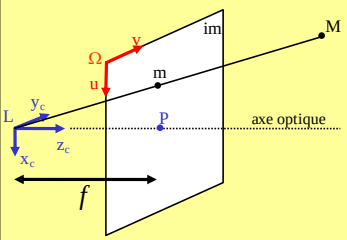
CAMERA EN TROU D 'EPINGLE
CHANGEMENT DE REPERE DANS LE PLAN IMAGE



$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = \begin{bmatrix} e_u & 0 & p_u \\ 0 & e_v & p_v \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} wu_c \\ wv_c \\ w \end{bmatrix}$$

P (p_u , p_v) dans Ωuv

CAMERA EN TROU D 'EPINGLE
MODELISATION INTERNE



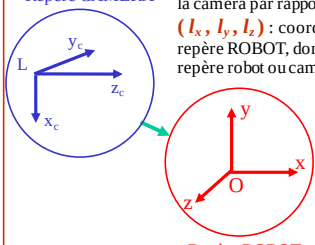
$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = K \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix}$$

$$K = \begin{bmatrix} fe_u & 0 & p_u \\ 0 & fe_v & p_v \\ 0 & 0 & 1 \end{bmatrix}$$

4 paramètres internes
- p_u , p_v : coordonnées du point principal dans (Ωuv)
- $\alpha_u = fe_u$, $\alpha_v = fe_v$: distances focales en unités pixels

CAMERA EN TROU D'EPINGLE
CHANGEMENT DE REPERE SPATIAL

Repère CAMERA



Repère ROBOT

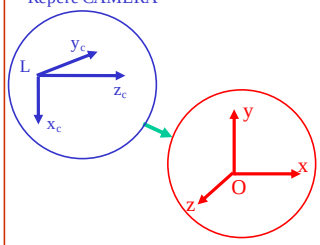
$R^T = [\vec{r}_1, \vec{r}_2, \vec{r}_3]$ matrice de rotation donnant l'orientation de la caméra par rapport au repère du robot ou de la scène

(l_x, l_y, l_z) : coordonnées du centre optique L dans le repère ROBOT, donne la position de la caméra dans le repère robot ou caméra

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = R^T \begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} + \begin{pmatrix} l_x \\ l_y \\ l_z \end{pmatrix}$$

CAMERA EN TROU D'EPINGLE
CHANGEMENT DE REPERE SPATIAL

Repère CAMERA



Repère du monde (ou du robot)

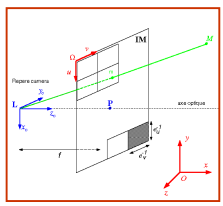
$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = [R\ t] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$t = -R \begin{pmatrix} l_x \\ l_y \\ l_z \end{pmatrix}$$

Les lignes de R donnent l'orientation de la caméra dans le repère du monde

L = (l_x , l_y , l_z)^T : coordonnées du centre optique L dans le repère du monde

CAMERA EN TROU D'EPINGLE
MODELISATION INTERNE ET EXTERNE



$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = \tilde{C} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{avec } \tilde{C} = KR[I_3 | -L]$$

$$\begin{bmatrix} wu \\ wv \\ w \end{bmatrix} = \begin{bmatrix} \alpha_u \vec{r}_1^T + p_u \vec{r}_3^T & \alpha_u t_x + p_u t_z \\ \alpha_v \vec{r}_2^T + p_v \vec{r}_3^T & \alpha_v t_y + p_v t_z \\ \vec{r}_3^T & t_z \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

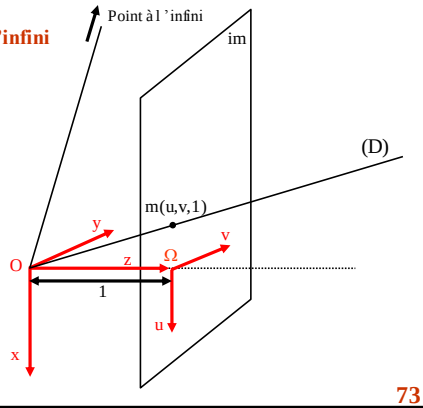
Forme développée

Modèle à **10** paramètres
- **4** paramètres **internes**
- **6** paramètres **externes**

PLAN PROJECTIF

$P^2 = E^2$ + points à l 'infini

m <------> (D)



PLAN PROJECTIF

$P^2 = E^2$ + points à l 'infini

m de E^2 <------> (D) de R^3

$(D) \equiv \tilde{m}$ si (D) n'est pas parallèle à im

$(D) \equiv \tilde{\omega}$ (point à l'infini) si (D) est parallèle à im

Si $m \in E^2$ a comme coordonnées (u, v),

$\tilde{m} \in P^2$ a pour coordonnées (u, v,1)

Si une droite (d) de E^2 (et de P^2) a pour vecteur directeur (d_u, d_v) dans E^2,

son point à l'infini ω_d de P^2 a pour coordonnées $(d_u, d_v, 0)$

x

ESPACE PROJECTIF

- Dans l 'ensemble R^{n+1}-{(0,0,…,0} la relation
  $(y_1, y_2, \ldots, y_{n+1}) R (x_1, x_2, \ldots, x_{n+1}) \Leftrightarrow \exists \lambda \in R^* \mid (y_1, y_2, \ldots, y_{n+1}) = \lambda (x_1, x_2, \ldots, x_{n+1})$
  définit une relation d 'équivalence dite de **colinéarité**
- On appelle **espace projectif P^n** issu de l 'espace vectoriel R^{n+1}, l 'ensemble quotient de R^{n+1}-{(0,0,…,0} par la relation de colinéarité
- On appelle **coordonnées homogènes** d 'un point M de P^n par rapport à une base B de R^{n+1}, les coordonnées dans cette base de l 'un quelconque $\vec{M}$ des représentants (dits « homogènes ») de M

Remarque : les coordonnées homogènes sont définies à un facteur multiplicatif non nul près

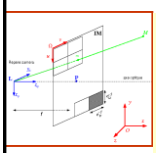
ESPACES AFFINE ET PROJECTIF

- Soit une base de R^{n+1}, $(x_1, x_2, \ldots, x_{n+1})$ un vecteur de R^{n+1}, représentant un point projectif
  - **si x_{n+1} non nul**, on associe au point projectif M, le point $M(x_1/x_{n+1}, x_2/x_{n+1}, \ldots, x_n/x_{n+1})$ de l'espace affine E^n.

    Inversement, au point affine $M(x_1, x_2, \ldots, x_n)$, on associe le point projectif $M(x_1, x_2, \ldots, x_n, 1)$
  - **si x_{n+1} nul**, on associe au point **« à l'infini »** Ω, le point à l'infini de la droite dont la direction est $(x_1, x_2, \ldots, x_n)$

$P^n = E^n$ + points à l'infini

INTERPRETATION PROJECTIVE



- Si (X,Y,Z,T) sont les coordonnées homogènes de M considéré comme un point de l'espace projectif P^3
- Si (U,V,W) sont les coordonnées homogènes de m considéré comme un point de l'espace projectif P^2

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \tilde{C} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix} = \begin{bmatrix} \alpha_u \vec{r}_1^T + p_u \vec{r}_3^T & \alpha_u t_x + p_u t_z \\ \alpha_v \vec{r}_2^T + p_v \vec{r}_3^T & \alpha_v t_y + p_v t_z \\ \vec{r}_3^T & t_z \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ T \end{bmatrix}$$

La caméra réalise une transformation projective LINEAIRE de P^3 dans P^2

CENTRE OPTIQUE

Soit $\tilde{C} = [C \;\; \tilde{c}_4]$ $C(3,3)$ inversible, $\tilde{c}_4(3,1)$

$$\tilde{C}\tilde{L} = \tilde{C} \begin{pmatrix} l_x \\ l_y \\ l_z \\ 1 \end{pmatrix} = 0$$

$$\Rightarrow \boxed{\begin{pmatrix} l_x \\ l_y \\ l_z \end{pmatrix} = -C^{-1}\tilde{c}_4}$$

MODELE INVERSE

Soit $\tilde{C} = [C \;\; \tilde{c}_4]$ $C(3,3)$ inversible , $\tilde{c}_4(3,1)$

Le rayon de projection associé à m est la droite qui passe par L et m.

Cette droite contient aussi son point à l'infini $\tilde{\Omega} = \begin{bmatrix} \Omega \\ 0 \end{bmatrix}$

Ce point a pour image m donc :

$\tilde{m} = \tilde{C}\tilde{\Omega} = C\Omega$ et $\Omega = C^{-1}\tilde{m}$

De plus, $O\tilde{\tilde{\Omega}}$ est parallèle à $L\vec{m}$

Enfin, un point courant du rayon de projection est :

$M = L + \lambda L\vec{m} \quad \lambda > 0$

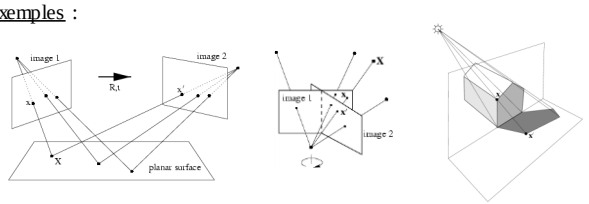
$$\Rightarrow \boxed{M = C^{-1}\left(-\tilde{c}_4 + \lambda \tilde{m}\right) \qquad \lambda > 0 \text{ ou } < 0}$$

Transformations projectives du plan

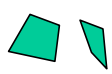
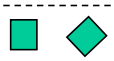
Définition : une transformation projective ou homographie de P^2 dans P^2 est une application inversible transformant toute droite en une droite

Théorème : une application h de P^2 dans P^2 est une transformation projective ssi il existe une matrice inversible H (3,3) telle que pour chaque point de P^2 représenté par son vecteur de coordonnées homogènes x, il est vrai que h(x) = H x

Exemples :



Hiérarchie des transformations projectives du plan

		Carrés transformés	Invariants
Homographie 8ddl	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$		Convergence, colinéarité, ordre de contact (intersection, tangence, inflexion, etc.), bi-rapport
Affine 6ddl	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Parallélisme, rapport des aires, rapport de longueur sur droites colinéaires ou parallèles (e.g points milieu), combinaisons linéaires de vecteurs (centres de gravité). **Droite à l'infini L.**
Similitude 4ddl	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Rapports de longueur, angles. **Points circulaires I,J**
Euclidienne 3ddl	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		longueurs, aires

$R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$: matrice de rotation

Droite à l'infini

Soit l une droite du plan, H une homographie alors la transformée de cette droite par H est la droite $l' = H^{-T}l$

$$x \in l \Leftrightarrow Hx \in l' \quad car \quad 0 = l^T Hx = \left(H^{-T} l\right)^T Hx = l^T H^{-1} Hx = l^T x = 0$$

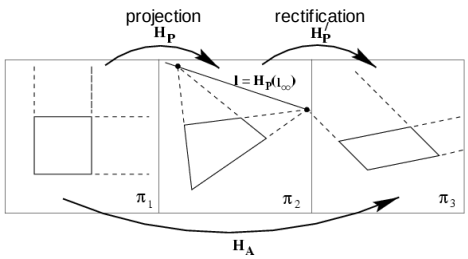
La droite à l'infini l_∞ est une droite invariante sous une transformation projective H *ssi* H est une transformation **affine**

Si H_A affine $H_A = \begin{bmatrix} A & t \\ 0 & 1 \end{bmatrix}$ et $l' = H_A(l_\infty) = H_A^{-T} l_\infty = \begin{bmatrix} A^{-T} & 0 \\ -t^T A^{-T} & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = l_\infty$

La réciproque est vraie : le montrer

Rque: l_∞ est invariante globalement mais pas ponctuellement

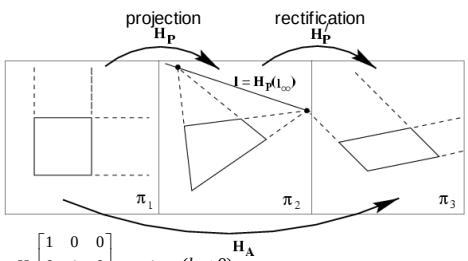
Rectification affine



Il suffit de trouver une homographie H'_P qui transforme l'image de la droite à l'infini $l=(l_1, l_2, l_3)^T = H_P(l_{infini})$ en la droite à l'infini. En effet, la transformation globale est alors $H'_P H_P$ et l_{infini} reste invariante dans cette transformation qui est donc affine. Si H_A est une transformation affine:

$$H'_P = H_A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} \text{ convient } \quad (l_3 \neq 0)$$

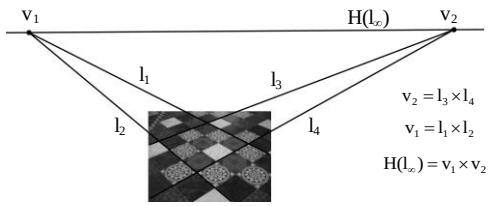
Rectification affine



$$H'_P = H_A \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} \text{ convient } (l_3 \neq 0)$$

$$l' = H'_P(l) = H_P'^{-T} l = H_A^{-T} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix}^{-T} l = H_A^{-T} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

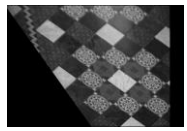
Rectification affine



$v_2 = l_3 \times l_4$

$v_1 = l_1 \times l_2$

$H(l_\infty) = v_1 \times v_2$



Coniques et coniques duales / 1

- Equation d'une **conique** (ponctuelle) **C** dans P^2 :

$a X_1^2 + b X_1 X_2 + c X_2^2 + d X_1 X_3 + e X_2 X_3 + f X_3^2 = 0$

Forme matricielle : $\boldsymbol{x}^T \boldsymbol{C} \boldsymbol{x} = \boldsymbol{0}$ avec

$$C = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}$$

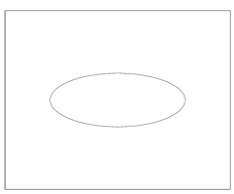
Remarquer que C est symétrique !

- **5 points définissent une conique**
- **La droite l tangente à C au point x s'écrit : l = C x**
- La conique duale **C*** à la conique (ponctuelle) C est formée par l'ensemble des droites tangentes à C :

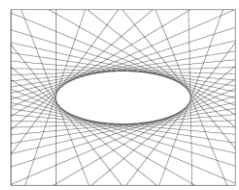
$l \in C^* \Leftrightarrow l^T C^* l = 0$ où C^* est la matrice ajointe de C

$C^* = C^{-1}$ si C est inversible

Coniques et coniques duales / 2



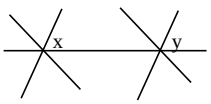
Conique (ponctuelle)



Conique duale C*
C est l'enveloppe des droites l appartenant à C*

Coniques et coniques duales / 3

- La conique C (ou C^*) est **dégénérée** si rg(C ou C^*) = 1 ou 2
- Exemple : $\mathbf{C^* = x\,y^T + y\,x^T}$ où x et y sont 2 points de P^2
  - rg (C^*) = 2
  - C^* est formé des 2 faisceaux de droites passant par x ou par y



- Soit C une conique ponctuelle, H une homographie, alors la transformée de cette conique par H est la conique : $\mathbf{C' = H^T\,C\,H^{-1}}$
- Soit C^* une conique duale, H une homographie, alors la transformée de cette conique duale par H est la conique : $\mathbf{C^{*\prime} = H C^*\,H^T}$

Points circulaires

- On complexifie le plan projectif.
- Il y a 2 **points « circulaires »** sur la droite à l'infini :

  $I = (1,i,0)^T$ et $J=(1,-i,0)^T$
- Les points circulaires et leur conique duale $C^* = I\,J^T + J\,I^T$ sont invariants sous une transformation homographique H *ssi* si H est une **similitude**
- Chaque cercle coupe la droite à l'infini en I et J (d'où le nom de points circulaires)
  - dans P^2 un cercle s'écrit : $X_1^2 + X_2^2 + d\,X_1X_3 + e\,X_2X_3 + f\,X_3^2 = 0$

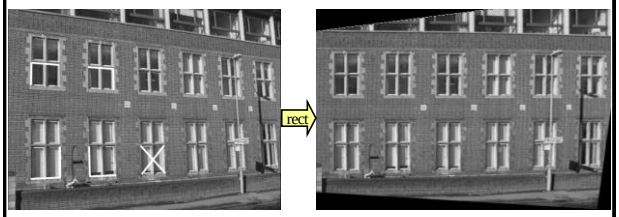
    Faisant $X_3 = 0$, on obtient : $X_1^2 + X_2^2 = 0$ de solutions I et J

Rectification métrique

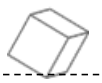
1. Identifier 5 couples de droites $(l_i\,,\,m_i)$ $i=1\ldots5$ dans le plan de l'image qui sont orthogonales dans le plan de l'objet
2. Identifier $C'^* = H_P\,(I\,J^T + J\,I^T)$ en résolvant le système $l'^T_i C'^* m'_i = 0$, $i=1\ldots5$
3. Calcul de la SVD de C'^*
4. Application à l'image de l'homographie H_{Rect} issue de la SVD de C'^* et rendant invariante la conique duale C^*

L'application $H_{Rect} \circ H_P$ est une similitude

Rectification métrique



Hierarchie des transformations projectives de l'espace

		Transformées d'un cube	Invariants
Homographie 15ddl	$\begin{bmatrix} A & t \\ v^T & v \end{bmatrix}$		Intersection et tangence
Affine 12ddl	$\begin{bmatrix} A & t \\ 0^T & 1 \end{bmatrix}$		Parallélisme de plans, Rapports de volume, centre de gravité, **Plan à l'infini π_∞**
Similitude 7ddl	$\begin{bmatrix} sR & t \\ 0^T & 1 \end{bmatrix}$		Angles, rapports de longueur **La conique absolue Ω_∞**
Euclidienne 6ddl	$\begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}$		Volume

R est une matrice (3,3) de rotation, t un 3-vecteur de translation

Plan à l'infini

Le plan à l'infini π_∞ est un plan invariant sous une transformation homographique H ssi H est une transformation affine

$$\pi'_\infty = H_A^{-T}\pi_\infty = \begin{bmatrix} A^{-T} & 0 \\ -t^T A^{-T} & 1 \end{bmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \pi_\infty$$

1. Position canonique $\pi_\infty = (0,0,0,1)^T$
2. Contient les directions $D = (X_1, X_2, X_3, 0)^T$
3. 2 plans sont parallèles $\Leftrightarrow$ droite d'intersection dans π_∞
4. droite // droite (ou plan) $\Leftrightarrow$ point d'intersection dans π_∞

Conique absolue (ombilic)

La conique absolue (ou ombilic) Ω_∞ est un cercle imaginaire de centre 0 et de rayon i situé dans le plan à l'infini π_∞.

dans un repère métrique: $\left.\begin{array}{r} X_1^2 + X_2^2 + X_3^2 \\ X_4 \end{array}\right\} = 0$

La conique absolue Ω_∞ est une conique invariante sous la transformation projective H ssi H est une **similitude**

1. Ω_∞ est simplement invariant en tant qu'ensemble
2. Un cercle coupe Ω_∞ en deux points circulaires
3. Les sphères coupent π_∞ en Ω_∞

CALIBRAGE ou ETALONNAGE de CAMERA

- Calibrer ou étalonner (géométriquement ou radiométriquement) une caméra, consiste à déterminer les paramètres de son modèle (géométrique ou radiométrique)
  - Pour une caméra « pinhole », l'étalonnage géométrique revient à déterminer 4 (modèle interne) ou 10 (modèle externe) paramètres
  - Pour étalonner géométriquement une caméra, on utilise une mire d'étalonnage, objet dont la géométrie est connue avec précision et on repère des points d'étalonnage (ou amers) sur cette mire ainsi que leurs images

APPRENTISSAGE

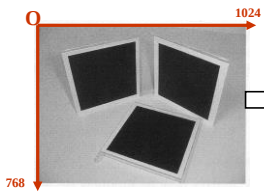
CALIBRAGE GEOMETRIQUE

CAMERA HOMOGRAPHIQUE

- CAMERA HOMOGRAPHIQUE : caméra dont l'image d'une droite est une droite (donc sans aberration géométrique)
- Une caméra en « trou d'épingle » est un cas particulier d'une caméra homographique

Technique de l'ombilic

- L'image de l'ombilic par une caméra homographique est une conique dont les paramètres internes sont reliés à ceux de la caméra.
- Même si l'ombilic est toujours présent dans la nature, il est néanmoins nécessaire de disposer d'une mire d'étalonnage car il faut être en mesure d'identifier l'image de l'ombilic.



$$K = \begin{bmatrix} 1097{,}8 & -9{,}8 & 395{,}9 \\ 0 & 1108{,}3 & 525{,}8 \\ 0 & 0 & 1 \end{bmatrix}$$

L'image de ces 3 carrés suffit à étalonner géométriquement la caméra homographique !

Calibrage du modèle interne d'une caméra homographique
Techniquede l'ombilic

- L'image ω de l'ombilic Ω par une caméra homographique (sans aberration) est la conique $\omega = (KK^T)^{-1}$: elle ne dépend donc que des paramètres internes de la caméra. Ainsi, en déterminant ω puis en appliquant la factorisation de Cholesky, il est possible de retrouver K et les paramètres internes de la caméra
- Chaque plan possède 2 points circulaires appartenant à l'ombilic. Leurs images appartiennent donc à ω.
- En déterminant 3 plans et leurs images puis en déterminant les homographies plan vers image du plan, on détermine 6 points appartenant à ω et donc ω
- Une figure plane commode est un carré et d'après ce qui précède l'image de 3 carrés non coplanaires permettra de déterminer les paramètres internes de la caméra homographique.

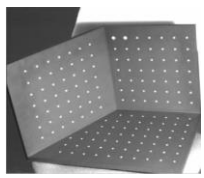
$$K = \begin{bmatrix} 1097{,}8 & -9{,}8 & 395{,}9 \\ 0 & 1108{,}3 & 525{,}8 \\ 0 & 0 & 1 \end{bmatrix}$$

APPRENTISSAGE

CALIBRAGE GEOMETRIQUE externe

Mires d'étalonnage

Trièdre



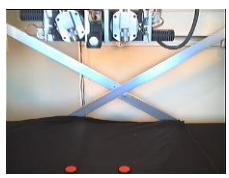
Points d'étalonnage:
Centres des cercles

Dièdre
(pour caméra en trou d'épingle uniquement)

Points d'étalonnage:
Sommets des carrés

Un seul plan de calibrage ne suffit jamais !

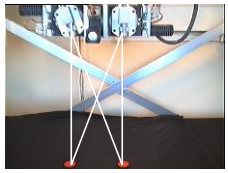
Etalonnage géométrique caméras NOISIV



2D
?
3D

Points de calibrage : 2 centres de disque

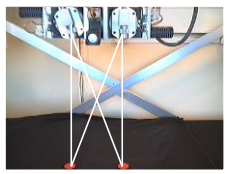
Etalonnage géométrique caméras NOISIV



2D

3D

Etalonnage géométrique caméras NOISIV



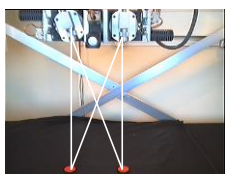
2D

3D

Les capteurs se déplacent !

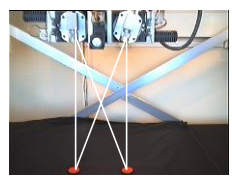
⇨ Création artificielle de 2 nouveaux points de calibrage

Etalonnage géométrique caméras NOISIV



2D

3D



Multi-cibles et Multi-vues

Ajustement de faisceaux : 1er TP

METHODE LINEAIRE DE CALIBRAGE EXTERNE

- Soit M_i de coordonnées (x_i, y_i, z_i) N points de calibrage
- Soit m_i leurs images de coordonnées (u_i, v_i)

Soit $\tilde{C} = \begin{bmatrix} \vec{c}_1^T & c_{14} \\ \vec{c}_2^T & c_{24} \\ \vec{c}_3^T & c_{34} \end{bmatrix}$ la matrice (3,4) cherchée

alors :

$$\forall i = 1 \ldots N \quad \tilde{C}\tilde{M}_i = \tilde{m}_i \Leftrightarrow \begin{cases} \vec{c}_1^T M_i - u_i \vec{c}_3^T M_i + c_{14} - u_i c_{34} = 0 \\ \vec{c}_2^T M_i - v_i \vec{c}_3^T M_i + c_{24} - v_i c_{34} = 0 \end{cases}$$

qui sont 2 équations linéaires homogènes par rapportsaux inconnues c_{ij}

On obtient donc 2N équations linéaires homogènes :

$$\boxed{A\vec{c} = 0 \text{ avec } A(2N,12) \text{ et } \vec{c} = \left[\vec{c}_1^T, c_{14}, \vec{c}_2^T, c_{24}, \vec{c}_3^T, c_{34}\right]^T}$$

$\vec{c}$ est donc défini à un facteur d'échelle près

METHODE LINEAIRE DE CALIBRAGE
pour caméra pinhole

- En général, si N>= 6 et si les points de calibrage ne sont pas coplanaires, les 2N équations $Ac = 0$ admettent une solution unique non nulle (à un facteur d 'échelle près).
- On a aussi 2 autres contraintes...

METHODE LINEAIRE DE CALIBRAGE

- THEOREME

Soit $\tilde{C} = \begin{bmatrix} \vec{c}_1^T & c_{14} \\ \vec{c}_2^T & c_{24} \\ \vec{c}_3^T & c_{34} \end{bmatrix}$ une matrice (3,4) de rang 3

Alors $\tilde{C}$ modélise une caméra en trou d'épingle si et seulement si

$$\|\vec{c}_3\| = 1$$
$$\left(\vec{c}_1 \wedge \vec{c}_3 | \vec{c}_2 \wedge \vec{c}_3\right) = 0$$

METHODE DE RESOLUTION

- Elle consiste à résoudre l 'un des 2 problèmes d 'optimisation avec contraintes suivant:

$$(P_1)\begin{cases} \mathrm{Min}\|A\vec{c}\|^2 \\ \|\vec{c}_3\|^2 = 1 \end{cases} \qquad (P_2)\begin{cases} \mathrm{Min}\|A\vec{c}\|^2 \\ \|\vec{c}_3\|^2 = 1 \\ (\vec{c}_1 \wedge \vec{c}_3 | \vec{c}_2 \wedge \vec{c}_3) = 0 \end{cases}$$

SOLUTION EXPLICITE

Soit: $(P)\begin{cases} \mathrm{Min}\|Ax\|^2 \\ \|Cx\|^2 = 1 \end{cases}$

où A est une matrice (N,M) et C est une matrice (P,N) qui échantillonne certaines coordonnées du vecteur $\boldsymbol{x}$

Soit $\boldsymbol{z} = C\boldsymbol{x}$ et $\boldsymbol{y}$ le (M-P) vecteur composé des coordonnées restantes de $\boldsymbol{x}$.

Soit B(N,M-P) et D(N,P) les sous-matrices de A correspondant à y et z : $Ax = By + Dz$

alors

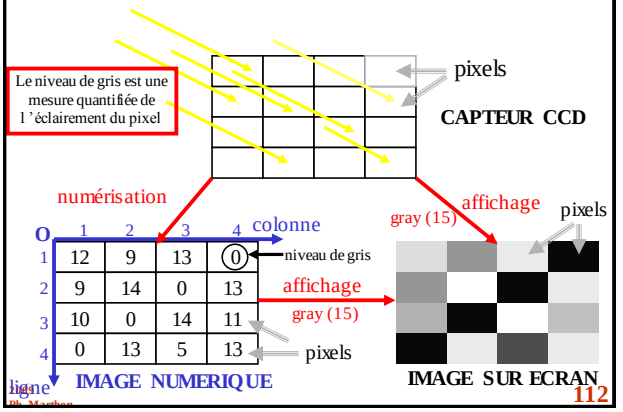
posant $E = D^TD - D^TB(B^TB)^{-1}B^TD$, E est une matrice semi-définie, et toute solution de (P) $\bar{x}$ est telle que:

- $\bar{z}$ est un vecteur propre unitaire associé à la plus petite valeur propre de E
- $\bar{y} = -(B^TB)^{-1}B^TDz$

APPRENTISSAGE

MODELISATION RADIOMETRIQUE

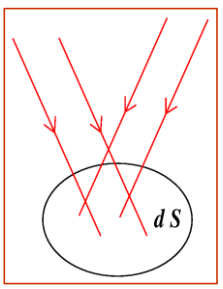
IMAGE NUMERIQUE MONOCHROME



ECLAIREMENT

- L'éclairement énergétique ou irradiance E en un point d'une surface est égal au rapport de la puissance à l'aire réceptrice :

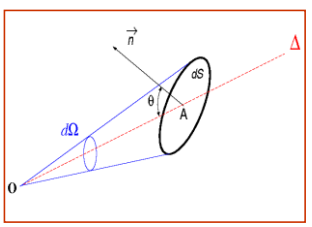
$$E = \frac{dP}{dS}$$



INTENSITE

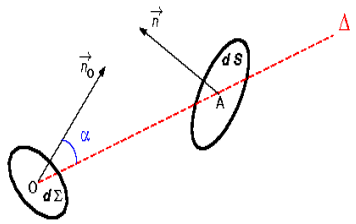
- L'intensité lumineuse I d'une source ponctuelle O dans une direction Δ est égale au rapport de la puissance dP émise par la source à l'intérieur d'un cône infiniment petit de sommet O contenant Δ, à l'angle solide dΩ de ce cône

$$I = \frac{dP}{d\Omega}$$

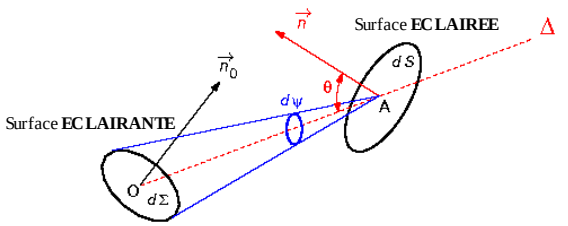


LUMINANCE

$$L = \frac{dI}{d\Sigma \cos\alpha}$$

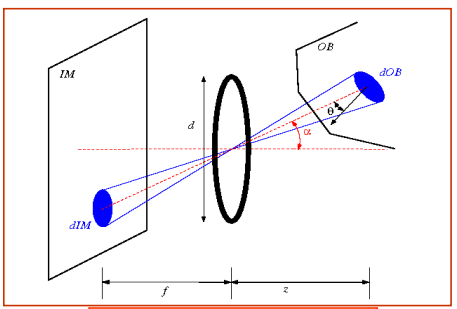


LUMINANCE ET ECLAIREMENT



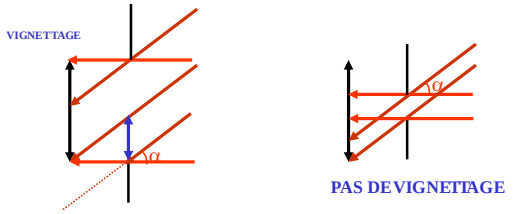
$$dE = L\cos\theta\, d\psi$$

PHOTOMETRIE DE LA CAMERA



$$E = L\ \pi\left(\frac{d}{2f}\right)^2 \cos^4\alpha$$

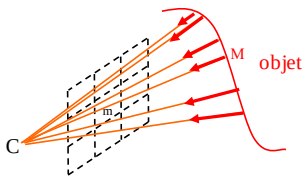
EFFET DE VIGNETTAGE



- L'effet de vignettage se manifeste par une luminosité plus forte du centre de l'image lorsqu'on ouvre le diaphragme
- Il est d'autant plus important que le diaphragme est ouvert et que l'angle α est grand
- Il se superpose à l'affaiblissement en $\cos^4 \alpha$

Calibrage des caméras optiques

- Le point **M** de l'espace physique a pour image le point **m**.
- Nous avons vu 2 types de calibrage:
  - **Géométrique** : à chaque point image m, on associe le rayon de projection contenant le point M
  - **Radiométrique** : à chaque point image m, on associe la luminance en provenance de M



APPRENTISSAGE

SCENE ET OBJETS

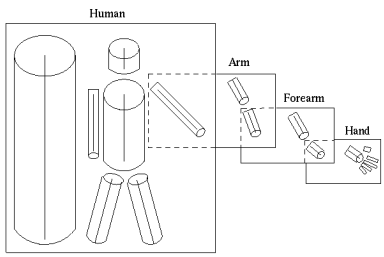
1.2. Modélisation des objets

- Les phénomènes, les objets ainsi que leurs images sont **localisés** dans l'espace et dans le temps.
- On peut :
  - **Soit modéliser l'objet 3D lui-même**
  - **Soit modéliser les images 2D de cet objet**
- Classes de modèles
  - Modèles **syntaxiques**
  - Modèles **géométriques** **(statiques et dynamiques)**
    - Faces-arêtes-sommets (polyèdres)
    - Modèle à base de Cylindres Généralisés
    - Modèle géométrique dynamique d'aspect (AAM)
  - Modèles **radiométriques**
    - Réflectance
      - surfaces lambertienne, spéculaire, …
    - Luminance
      - Modèle radiométrique dynamique d'aspect (AAM)
      - Statistique : Champs de Markov, modèles autorégressifs AR …

APPRENTISSAGE

MODELISATION GEOMETRIQUE

Modélisation géométrique



Représentation 3D des objets: composition de **cylindres généralisés** (modèle proposé par D. Marr comme représentation utilisée par le système visuel humain)

Modèles d'aspect (Active Appearance Models)

- AAM : Modèle (paramétré) dynamique (« active ») d'un certain phénomène visuel ou des images de ce phénomène visuel
- Un AAM se composed'un
  - Modèle géométrique
  - Modèle radiométrique
- Applications :
  - suivi d'objets et estimation de pose (exemple: visage)
  - analyse de déformations
  - reconnaissance d'objets
- Mise en œuvre d'une application
  1. Apprentissage des modèles AAM
  2. Ajustage (fitting) du modèle

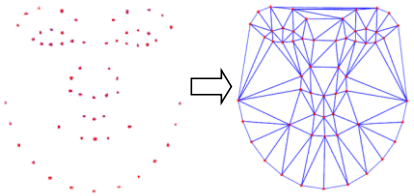
Suivi de visages par modélisation AAM 2D

Triangulation de Delaunay

- En géométrie, une **triangulation** est une partition d'un objet en un ensemble de simplexes (de triangles si l'objet est plan).
- Une triangulation d'un ensemble de points P du plan est une triangulation de l'enveloppe convexe de P, tous les points de P formant alors les sommets de cette triangulation
- La **triangulation de Delaunay** d'un ensemble de points P du plan est une triangulation telle qu'aucun point de P n'est à l'intérieur du cercle circonscrit d'un des triangles de cette triangulation

Triangulation de Delaunay

Triangulation de Delaunay



Modèle de forme d'un AAM 2D

- Définition
  - Maille de base s_0 construite à partir d'une triangulation de Delaunay de n sommets
  - Combinaison linéaire de k vecteurs codant les déformations de la maille

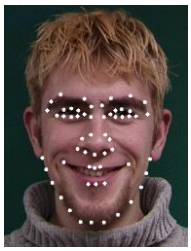
$$s = (x_1, y_1, x_2, y_2, \ldots, x_n, y_n)^T \ : \text{sommets de la maille}$$

$$s = s_0 + \sum_{i=1}^{k} p_i s_i \ : \ p_i \equiv \text{paramètres de forme}$$

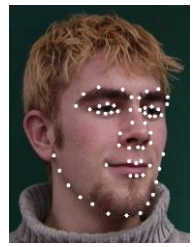
Modèle de forme d'un AAM

- Apprentissage
  - Relevé manuel des sommets sur des « exemples » (images d'apprentissage)
  - Normalisation : alignement des exemples par analyse de Procrustes itérative (élimination des effets de translation, rotation et changements d'échelle)
  - Calcul des n vecteurs s_i codant les déformations locales et non-rigides de la forme moyenne s_0 : analyse en composantes principales des exemples, les s_i sont les k vecteurs propres associés aux k plus grandes valeurs propres (les autres vecteurs s_j j > k sont non-significatifs)
  - Triangulation de Delaunay de s_0

Apprentissage du modèle de forme



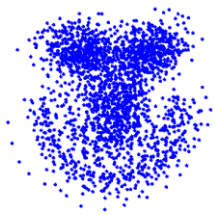
Relevé des sommets sur exemple 1



Relevé des sommets sur exemple 2

Apprentissage du modèle de forme

Analyse de Procrustes

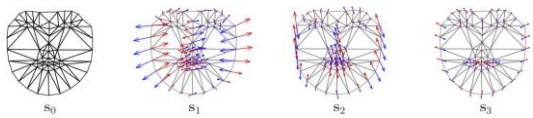


Formes non alignées

Formes alignées

Apprentissage du modèle géométrique

Analyse en composantes principales et triangulation de Delaunay de la forme moyenne s_0



APPRENTISSAGE

MODELISATION RADIOMETRIQUE

Modèle d'aspect d'un AAM

- Définition
  - Image numérique (couleur ou d'intensité) moyenne $A_0(x)$
  - Combinaison linéaire des m images d'aspect $A_i(x)$ codant les variations de couleur ou d'intensité des pixels x.

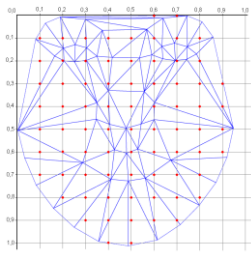
$$A(x) = A_0(x) + \sum_{i=1}^{m} \lambda_i A_i(x) \quad x \in s_0 \quad \lambda_i \text{ paramètres d'aspect}$$

s_0 désigne maintenant l'ensemble des pixels $x = (x,y)^T$ contenus dans la maille de base s_0

Modèle d'aspect d'un AAM

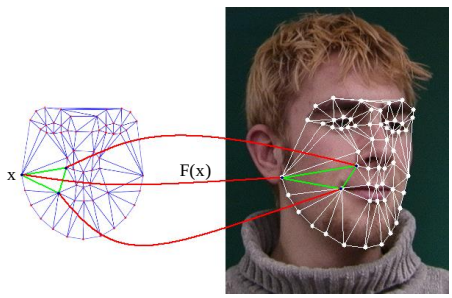
- Apprentissage
  - Choix du nombre de pixels décrivant le modèle d'aspect, autrement dit, densité de la grille des pixels à l'intérieur de la maille de base
  - Pour chaque exemple I_l, déterminer la déformation linéaire affine par morceaux F faisant correspondre chaque triangle de s à son triangle homologue dans l'exemple I_l
  - Calculer : $\forall x \in s_0 \quad I_l^0(x) = I_l(F(x))$
  - $A_0(x)$ est l'image moyenne des $I_l^0(x)$
  - Effectuer une analyse en composantes principales des $I_l^0(x)$: les $A_i(x)$ sont les m images propres associées aux m plus grandes valeurs propres

Modèle d'aspect d'un AAM



Choix du nombre de pixels décrivant le modèle d'aspect, autrement dit, densité de la grille des pixels à l'intérieur de la maille de base

Modèle d'aspect d'un AAM



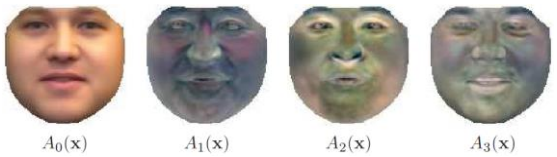
l'exemple I_l

Modèle d'aspect d'un AAM

Exemple d'apprentissage

Exemple d'apprentissage ajusté à la forme moyenne s_0

Modèle d'aspect d'un AAM

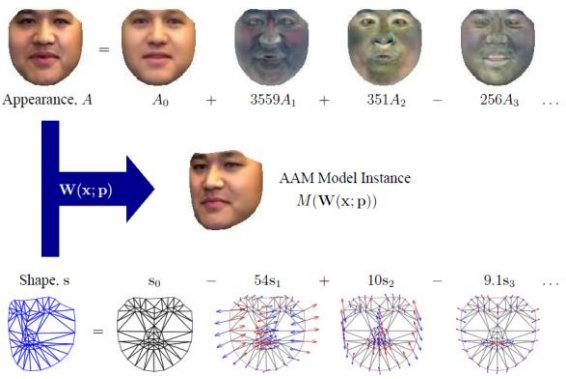


Modèle Actif d'Aspect (AAM)

- *Définition* : ***M(p,λ)*** est un modèle représentant l'image d'un objet qui se déforme.
- *Construction* : on définit la déformation (« warp ») **W(x,p)** , x pixel de s_0, comme suit : x appartient nécessairement à un triangle t de la triangulation de Delaunay de s_0 qui se transforme en un autre triangle t'(p). Sur ce triangle t, *W(.,p)* est alors l'unique transformation affine transformant le triangle t en triangle t'.

$$M(W(x;p)) = A(x) = A_0(x) + \sum_{i=1}^{m} \lambda_i A_i(x) \quad x \in s_0$$

Modèle AAM complet



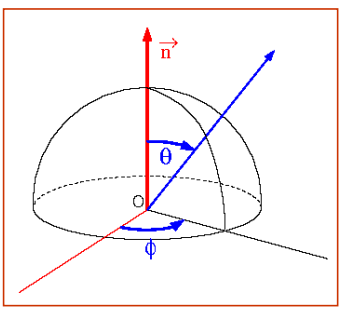
Ajustage (fitting) d'un AAM

- *On minimise sur (p,λ)*

$$\sum_{x\in s_0}\left[M(W(x;p))-I(W(x;p))\right]^2=\sum_{x\in s_0}\left[A_0(x)+\sum_{i=1}^{m}\lambda_i A_i(x)-I(W(x;p))\right]^2$$

REFLECTANCE

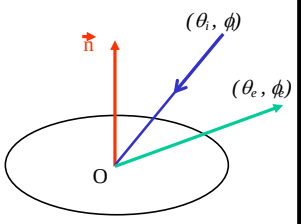
REPERE LOCAL



FONCTION DE
REFLECTANCE BIDIRECTIONNELLE

- $dE(\theta_i, \phi_i)$ l'éclairement d'un élément de surface en provenance d'une direction (θ_i, ϕ_i)
- $dL(\theta_e, \phi_e)$ la luminance de ce même élément de surface dans la direction (θ_e, ϕ_e)

$$f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{dL(\theta_e, \phi_e)}{dE(\theta_i, \phi_i)}$$



CONDITION DE RECIPROCITE D'HELMHOLTZ

$$f(\theta_e, \phi_e, \theta_i, \phi_i) = f(\theta_i, \phi_i, \theta_e, \phi_e)$$

- **En effet, lorsque 2 surfaces sont en équilibre thermique, le rayonnement de l'une sur l'autre est contrebalancé par le rayonnement se propageant en sens opposé.**
- **Si ce n'était pas le cas, l'une se réchaufferait tandis que l'autre se refroidirait, ce qui violerait la seconde loi de la thermodynamique.**

REFLECTANCE : caractérisation

- La réflectance d'une surface se caractérise par 2 paramètres:
  - Son albédo
  - Ses effets directionnels (reflets)

ALBEDO

- L 'ALBEDO ρ en un point d 'une surface est égal à la fraction de l 'énergie réfléchie par cet élément.
- Une surface est dite **BLANCHE** si elle réfléchit toute la lumière qu 'elle reçoit.

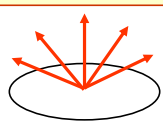
  Son albedo est donc égal à 1.
- Une surface est dite **NOIRE** si elle absorbe toute la lumière qu 'elle reçoit.

  Son albedo est donc égal à 0.
- Une surface **GRISE** est une surface ni blanche ni noire.

  Son albedo ρ tel que $0 < \rho < 1$

SURFACE DE LAMBERT

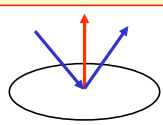
- Une **surface de Lambert** (ou surface mate) est une surface dont la luminance reste la même dans toutes les directions à éclairement constant.
- Sa fonction de réflectance f est constante et s 'écrit: $f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{\rho}{\pi}$



SURFACE SPECULAIRE

- Une **surface spéculaire** (ou miroir) est une surface qui reflète toute la lumière venant de la direction (θ_i, ϕ_i) dans la direction symétrique $(\theta_i, \phi_i + \pi)$
- Sa fonction de réflectance f s 'écrit:

$$f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{\delta(\theta_e - \theta_i)\delta(\phi_e - \phi_i - \pi)}{\sin\theta_i \cos\theta_i}$$



APPRENTISSAGE

ECLAIRAGE

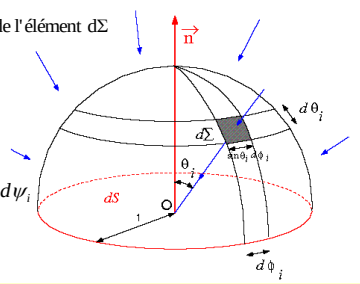
1.3. Modélisation de l'éclairage

- **Eclairage cohérent :** fluctuations lumineuses corrélées, phénomènes d'interférences (ex: radar)
- **Eclairage incohérent :** fluctuations lumineuses décorrélées, pas de phénomènes d'interférences (ex: lumière solaire)
- **Lumière polarisée** (ex: radar) ou **non polarisée**

ECLAIREMENT D 'UNE SURFACE PAR UNE SOURCE ETENDUE INCOHERENTE

$E(\theta_i, \phi_i)$ luminance de l'élément dΣ

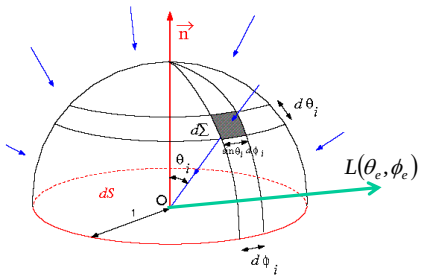
$dE = E(\theta_i, \phi_i)\cos\theta_i d\psi_i$



$$E = \int dE = \int_{-\pi}^{+\pi}\int_{0}^{\pi/2} E(\theta_i, \phi_i)\sin\theta_i \cos\theta_i \, d\theta_i d\phi_i$$

LUMINANCE D'UNE SURFACE ECLAIREE PAR UNE SOURCE ETENDUE

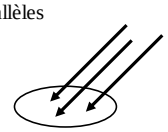
$dL = f\, dE$



$$L(\theta_e,\phi_e) = \int dL = \int_{-\pi}^{+\pi}\int_{0}^{\pi/2} f(\theta_i,\phi_i,\theta_e,\phi_e)E(\theta_i,\phi_i)\sin\theta_i\cos\theta_i\, d\theta_i d\phi_i$$

ECLAIRAGE D'UNE SURFACE PAR UNE SOURCE PONCTUELLE A L'INFINI

- On appelle **source ponctuelle à l'infini** une source de lumière suffisamment éloignée de la scène pour que l'on puisse admettre que tous les rayons qu'elle reçoit soient parallèles



(θ_s,ϕ_s) : direction de la source
I : intensité de la source dans cette direction

$$E(\theta_i,\phi_i) = K\delta(\theta_i-\theta_s)\delta(\phi_i-\phi_s)$$

$$I = \int dI = \int_{-\pi}^{\pi}\int_{0}^{\pi/2} E(\theta_i,\phi_i)\sin\theta_i\, d\theta_i d\phi_i = K\sin\theta_s$$

$$\Rightarrow E(\theta_i,\phi_i) = \frac{I}{\sin\theta_s}\delta(\theta_i-\theta_s)\delta(\phi_i-\phi_s)$$

$$E = \int_{-\pi}^{\pi}\int_{0}^{\pi/2} \frac{I}{\sin\theta_s}\delta(\theta_i-\theta_s)\delta(\phi_i-\phi_s)\sin\theta_i\cos\theta_i\, d\theta_i d\phi_i = I\cos\theta_s$$

LUMINANCE D'UNE SURFACE DE LAMBERT

- **Source ponctuelle à l'infini** de direction (θ_s, ϕ_s) d'intensité I :

$$L(\theta_e,\phi_e) = \frac{\rho}{\pi} I\cos\theta_s$$

La luminance est indépendante de la direction d'observation mais elle dépend de l'orientation de la surface : condition favorable pour Voir en relief !

- **Ciel uniforme** de luminance E :

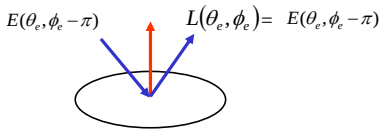
$$L(\theta_e,\phi_e) = \rho E$$

Difficulté pour voir le relief car l'orientation de la surface n'intervient pas (cf champs de neige)

LUMINANCE D'UNE SURFACE SPECULAIRE

$$L(\theta_e,\phi_e)=\int dL=\int_{-\pi}^{+\pi}\int_{0}^{\pi/2}\frac{\delta(\theta_e-\theta_i)\delta(\phi_e-\phi_i-\pi)}{\sin\theta_i\cos\theta_i}E(\theta_i,\phi_i)\sin\theta_i\cos\theta_i\,d\theta_i d\phi_i$$
$$=E(\theta_e,\phi_e-\pi)$$

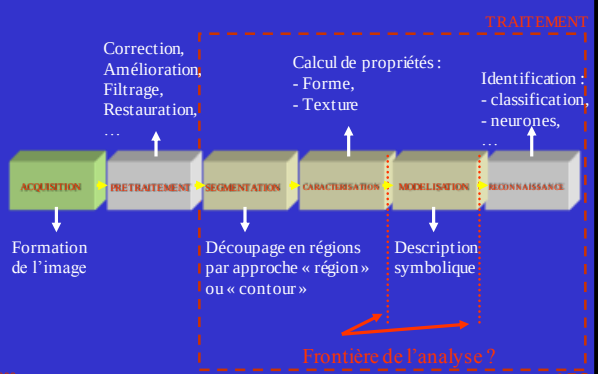
Résultat qui était attendu !



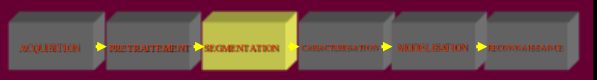
APPRENTISSAGE

IMAGES

« Une » chaîne type d'analyse d'images et d'interprétation



Segmentation



Scène - Objet - Objet élémentaire

A. Une **scène** est composée d'**objets** qui sont localisés dans l'espace et dans le temps. Il en est de même des images de ces objets, appelées des « **régions** ».

B. Un objet se compose d'un ou de plusieurs **objets élémentaires.**

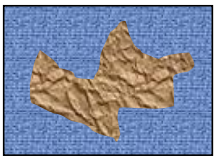
C. La surface S d'un **objet élémentaire** possède un albédo ρ (M) caractéristique. Il en est de même de l'image d'un objet élémentaire, appelée « **région élémentaire** », qui possède une **texture** caractéristique.

Segmentation

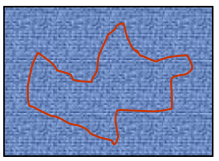
- **Segmentation** : partition d'une image en régions (ou segments)
  - Chaque région est connexe (**sauf occultation**)(propriété A)
  - Chaque région est une réunion de régions élémentaires (prop.B)
  - Chaque région élémentaire possède une texture unique (prop.C)
- Pour segmenter une image, on procède en deux temps:
  - On détecte les régions élémentaires
  - On détecte les régions en appliquant des **lois de regroupement** aux régions élémentaires

Texture et Contour

- Pour détecter une région élémentaire, on peut
  - Soit détecter sa **texture**
  - Soit détecter son **contour** (i.e. la frontière de cette région élémentaire)



Texture



Contour

Segmentation

Image avec les contours des régions superposés

Image d'étiquettes, chaque pixel porte le numéro de la région à laquelle, il appartient

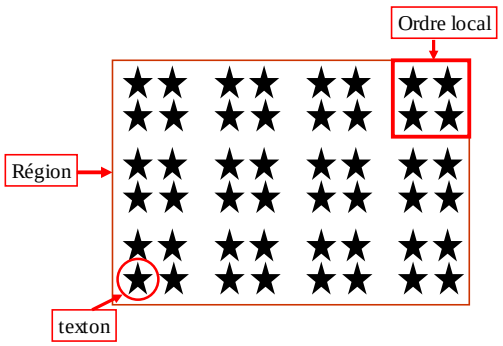
TEXTURE/

- <u>Définition</u> : distribution spatiale des niveaux de gris d'une région élémentaire
- Plusieurs modélisations des textures ont été proposées:
  - à partir de **textons** : 3 macro-paramètres : texton, ordre local, ordre global
  - à partir d'**indices** de texture extraits de matrices de co-occurrence
  - à partir de **champs de Markov**

Modèle de TEXTURE

- Modèle: une **texture** a 3 composantes:
  - un ordre local est répété sur une région grande par rapport à cet ordre local
  - cet ordre local consiste en un arrangement déterminé d 'objets élémentaires
  - ces objets élémentaires – des « **textons** »- sont des entités en gros uniformes ayant à peu près partout dans la région des dimensions identiques

TEXTURE



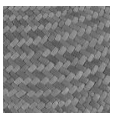
TEXTURES NATURELLES

Ecorce

Pierre

Sable

Tissu



Eau

Bois

Feuille

Bois, Eau, Ecorce, Feuille, Pierre, Sable, Tissu **???**

MATRICE DE CO-OCCURRENCE

Pour un vecteur de translation t = (d,θ), l'élément MC_t (a,b) d'une matrice de co-occurrence est égal au nombre de couples de pixels (p,p+t) de la région analysée, séparés par le vecteur t, le pixel p ayant le niveau de gris a et le pixel p+t le niveau de gris b.

0	0	2	2	0
1	1	2	2	1
0	3	3	3	0
0	0	1	2	0

IMAGE

	0	1	2	3
0	2	1	1	1
1	0	1	2	0
2	2	1	2	0
3	1	0	0	2

MATRICE DE CO-OCCURRENCE
d=1 θ = 0

Indices de texture

Homogénéité

$$H = \frac{1}{N^2} \sum_a \sum_b \left(MC_t(a,b)\right)^2$$

Contraste

$$C = \frac{1}{N(L-1)^2} \sum_{k=0}^{L-1} k^2 \sum_{|a-b|=k} MC_t(a,b)$$

Entropie

$$E = 1 - \frac{1}{N\,Ln(N)} \sum_a \sum_b MC_t(a,b)\,Ln\left(MC_t(a,b)\right)$$

N : nombre de couples (p, p+t) de la région R
L : nombre de niveaux de gris possible

Indices de texture

- Homogénéité : d'autant plus élevé que l'on retrouve souvent le même couple de pixels
  - ce qui est le cas quand le niveau de gris est uniforme ou quand il y a périodicité dans le sens de la translation
- Contraste : chaque terme de la matrice est pondéré par sa distance à la diagonale; élevé quand les niveaux de gris sont très différents
- Entropie : faible si on a souvent le même couple de pixels

CHAMP DE MARKOV

- Le champ image est **markovien** si le niveau de gris de chaque pixel ne dépend statistiquement que des niveaux de gris de ses voisins
- Soit **S** l'ensemble des sites ou des pixels
- **V** un système de « voisinages », formé de parties v_s de S tel que:

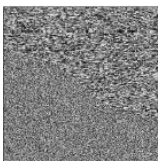
$$\begin{cases} s \notin v_s \\ s \in v_t \Leftrightarrow t \in v_s \end{cases}$$

- t est dit **voisin** de s si $t \in v_s$

im est un champ markovien ssi :

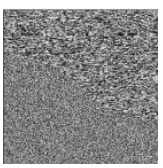
$$P(im(s) = a_s | im(r) = a_r \; r \neq s) = P(im(s) = a_s | im(r) = a_r \; r \in v_s) \quad \forall s$$

Segmentation de textures

im=  ? *=IM*

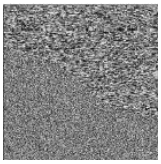
Une approche possible : les **champs de Markov**

Segmentation de textures

im=  ? *=IM*

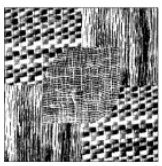
Une approche possible : les **champs de Markov**

Segmentation de textures

im=  ? *=IM*

Une approche possible : les **champs de Markov**

Segmentation de textures

im=  ? *=IM*

Une approche possible : les **champs de Markov**

CHAMP DE MARKOV

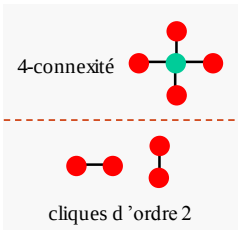
$$P(im(s)=a_s|im(r)=a_r\ r\neq s)=P(im(s)=a_s|im(r)=a_r\ r\in v_s)\quad \forall s$$

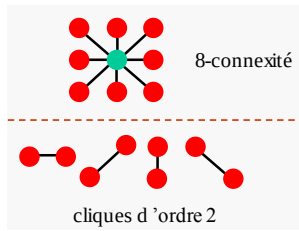
CHAMP DE MARKOV

$$P(im(s)=a_s|im(r)=a_r\ r\neq s)=P(im(s)=a_s|im(r)=a_r\ r\in v_s)\quad \forall s$$

CLIQUE

- Un sous-ensemble c de **S** est appelé une **clique** relative au système de voisinages **V** si :
  - c est un singleton
  - ou 2 pixels quelconques de c sont voisins (au sens de V)





CHAMP DE MARKOV

- **Théorème d 'Hammersley-Clifford**
  - Un champ image ***im*** est un **champ de Markov** associé au système de voisinages **V** si et seulement si sa distribution de probabilité *P(im = x)* est une **mesure de Gibbs** définie par:

$$\forall x\in\Omega^Q\quad P(im=x)=\frac{e^{-U(x)}}{Z}$$

où ***U*** est une fonction appelée **énergie** définie par:

$$U(x)=\sum_{c\in C}V_c(x/c)\qquad C\text{: ensemble des cliques}$$

Les fonctions V_c sont appelées des fonctions **potentiel** et ne dépendent que des éléments *x* qui appartiennent à *c*, la notation *x/c* désignant la restriction de *x* à *c*

Z appelé **fonction de partition** est la constante de normalisation :

$$Z=\sum_{x\in\Omega^Q}e^{-U(x)}$$

Champ de Markov

- Démo condition suffisante

$$P(im(s)=a_s|im(r)=a_r\ r\neq s)=\frac{P(im(s)=a_s,im(r)=a_r\ r\neq s)}{\sum_{x\in\Omega}P(im(s)=x,im(r)=a_r\ r\neq s)}$$

$$=\frac{e^{-\sum_{c\in sc}V_c(a/c)}\ \cancel{e^{-\sum_{c\in sc}V_c(a/c)}}}{\sum_{x\in\Omega}e^{-\sum_{c\in sc}V_c(a/c,\cancel{im(s)=x})}\ \cancel{e^{-\sum_{c\in sc}V_c(a/c)}}}$$

se factorise !

$$=\frac{e^{-\sum_{c\in sc}V_c(a/c)}}{\sum_{x\in\Omega}e^{-\sum_{c\in sc}V_c(a/c,im(s)=x)}}=P(im(s)=a_s|im(r)=a_r,r\in v_s)$$

Modèle d'Ising d'ordre 2

- Les seules cliques c ayant un potentiel non nul sont les 4 cliques à 2 éléments (pour la 8-connexité)

soit : $C_1=$ $C_2=$ $C_3=$ $C_4=$

- Le potentiel de ces cliques s'écrit:

$$V_c(x/c)=\beta_c\,\Delta(x/c)\quad (c=(r,s))$$
$$\Delta(x_r,x_s)=-1\quad si\quad x_s=x_r$$
$$\Delta(x_r,x_s)=+1\quad si\quad x_s\neq x_r$$

- Soit $b=(\beta_1,\beta_2,\beta_3,\beta_4)$, le vecteur de paramètres de ce modèle, alors l'énergie $U(x)=b.K(x)$

avec $K(x)=(k_1(x),k_2(x),k_3(x),k_4(x))$ et $k_i(x)=\sum_{c\in C_i}\Delta_c(x)$

Approche markovienne de la segmentation

- *Hypothèses de modélisation*:
  - L'image numérique observée a, est composée d'un certain nombre de régions élémentaires, chaque région élémentaire possédant une texture unique.
  - Chaque texture est la réalisation d'un champ de Markov
  - L'image segmentée λ^* est la réalisation d'un champ de Markov IM
- *Formulation du problème:*
  - La segmentation consiste à estimer λ^* connaissant l'image a.

Approche bayésienne par estimation du maximum a posteriori (MAP)

- **Critère MAP** : trouver la segmentation λ^{**} qui maximise la probabilité conditionnelle: $P(IM=\lambda / im=a)$
- En utilisant le **théorème de Bayes**:

$$P(IM=\lambda | im=a) = \frac{P(IM=\lambda, im=a)}{P(im=a)} = \frac{P(im=a | IM=\lambda) P(IM=\lambda)}{P(im=a)}$$

- **Critère MAP (reformulation) :**

trouver λ^{**} qui maximise $P(im=a \mid IM=\lambda)\, P(IM=\lambda)$

Approche markovienne

- Chaque texture étant modélisée comme un champ markovien: pour la texture l :

$$P(a_R | l) = \frac{1}{Z^{(l)}} e^{-U^{(l)}(a_R)}$$

R étant une région de l'image possédant la texture l

En général, les textures sont spatialement indépendantes.

$$P(im=a | IM=\lambda) = \prod_{R \in im} \frac{1}{Z^{l(R)}} e^{-U^{l(R)}(a_R)} = \frac{e^{-U_1(a,\lambda)}}{Z_1(\lambda)} \qquad Z_1(\lambda) = \sum_{a \in \Omega_1^Q} e^{-U_1(a,\lambda)}$$

- IM étant aussi un champ markovien: $P(IM=\lambda) = \frac{e^{-U_2(\lambda)}}{Z_2}$

$$P(im=a | IM=\lambda)\, P(IM=\lambda) \div \frac{e^{-U_1(a,\lambda)-U_2(\lambda)}}{Z_1(\lambda)}$$

Algorithme ICM

- ICM = Iterated Conditional Mode
- Choisir une image initiale IM_0 aussi proche que possible de la solution optimale IM^*
- Balayer l'ensemble des sites (pixels) S. La nouvelle texture (région élémentaire) $\lambda_s^{(n+1)}$ réalise le maximum de la probabilité conditionnelle :

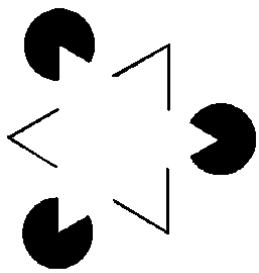
$$P\{IM(s)=\lambda | IM(r)=\lambda_r^{(n)}\ r \neq s\ ,\ im=a\}$$

- Itérer jusqu'à ce que le test d'arrêt soit satisfait

Contour

- **Définition d'un Contour** : frontière d'un phénomène ou d'un objet
- *Comment se caractérise un contour ?*
  - Différence de luminance
  - Différence de texture
  - Courbe régulière (à cause du camouflage)

Contours « subjectifs »



Extraction de contours

- Les **contours** des objets sont souvent (mais pas toujours !!) les lignes le long desquelles les niveaux de gris varient rapidement.
- D'où une extraction en 2 étapes:
  - Calcul du module du gradient de l'image
  - Calcul des contours comme maxima locaux du gradient dans la direction du gradient ou comme passage par zéro de la dérivée seconde dans la direction du gradient

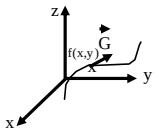
Opérateurs de contour

Cas continu

- Opérateur **gradient** (1er ordre)

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x,y) - f(x+h,y)}{h}$$

$$\vec{G} = \begin{pmatrix} G_x = \dfrac{\partial f}{\partial x} \\ G_y = \dfrac{\partial f}{\partial y} \end{pmatrix}$$



Norme : $\left(G_x^2 + G_y^2\right)^{1/2}$ **Direction** : arctg [G_y / G_x]

- Opérateur **laplacien** (2ème ordre)

$$\Delta = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Cas discret

Approximation de ces opérateurs par des masques

Ex : gradient G_x

(h=1)

f(x,y)	f(x+1,y)
f(x,y+1)	f(x+1,y+1)

→

+1	- 1
0	0

2009
Ph. Marthon

190

Opérateurs de contour

Opérateurs gradient

On définit 2 masques H1 et H2

K1=H1 ⊗ Image
K2=H2 ⊗ Image

$G=(K1^2+K2^2)^{1/2}$
ou
$G=|K1|+|K2|$

ROBERTS (croisé)

H1= $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ H2= $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

SOBEL

H1= $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}$ H2= $\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$

PREWITT

H1= $\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}$ H2= $\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

Opérateurs Laplacien

$\begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix}$ $\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$ $\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}$

2009
Ph. Marthon

191

Opérateurs de contour

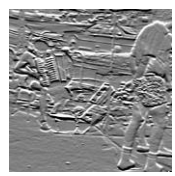
IMAGE

ROBERTS

SOBEL

PREWITT

LAPLACIEN



COMPASS (N)

2009
Ph. Marthon

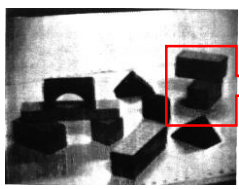
192

DETECTION DE CONTOURS
seuillage du module du gradient SOBEL

Image originale

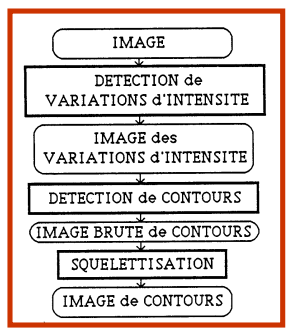
Image binaire de contours
seuil = 26,15

S.I.R.O.P.



Modules des gradients de l 'image obtenus par le filtre de Sobel

S.I.R.O.P.



Les 2 parallélépipèdes

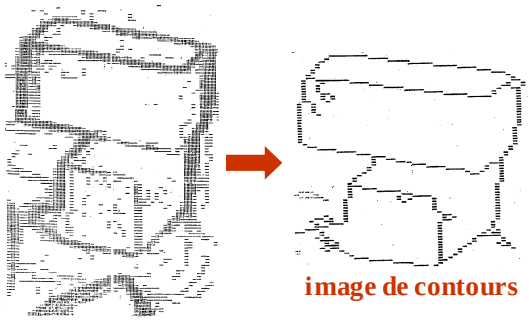


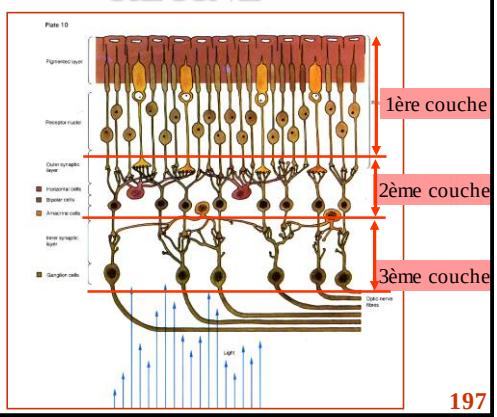
image des modules de gradients

image de contours

2009
Ph. Marthon

196

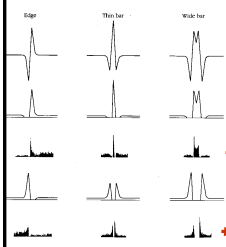
RETINE



2009
Ph. Marthon

197

RETINE ET LAPLACIEN DE GAUSSIENNE



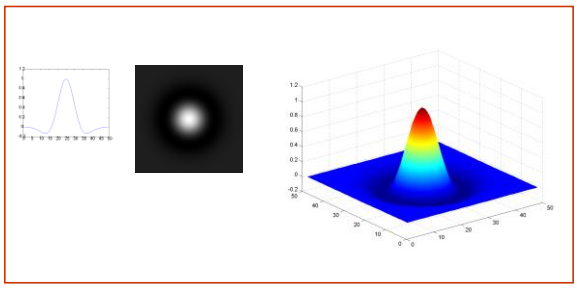
- Réponses du modèle laplacien de gaussienne
- Réponses des cellules ganglionnaires X à centre ON
- Réponses des cellules ganglionnaires X à centre OFF

2009
Ph. Marthon

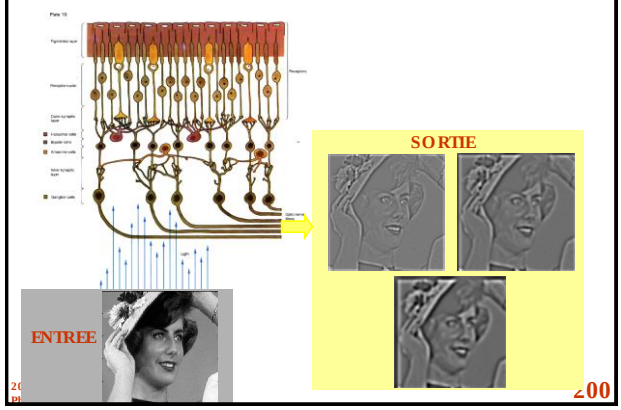
198

LAPLACIEN DE GAUSSIENNE

$$\Delta G_2(x,y) = \left(1 - \frac{x^2+y^2}{2\sigma^2}\right)\, e^{-\frac{x^2+y^2}{2\sigma^2}}$$



RETINE ET LAPLACIEN DE GAUSSIENNE



CONVOLUTIONS AVEC LAPLACIENS DE GAUSSIENNE

DETECTION DE CONTOURS

Passage par zéro du laplacien de gaussienne

Image originale

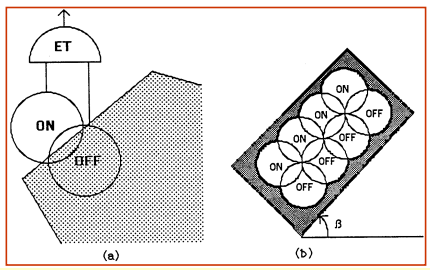
σ= 2

σ = 4

σ= 3

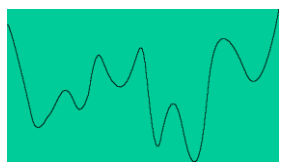
SYSTEME VISUEL HUMAIN

- **Modèle d'une cellule simple du cortex strié comme détecteur de segments de passages par zéro**



Chaque cercle représente le champ récepteur du lobe central d'une cellule ganglionnaire X

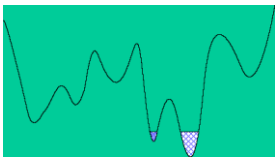
Algorithme de ligne de partage des eaux (LPE)



Profil du module du gradient ou de la puissance du contour

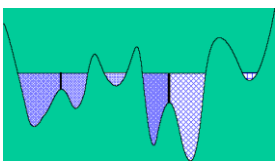
Algorithme de ligne de partage des eaux (LPE)

- L'image de puissance de contour est considérée comme une surface
- Simulation d'immersion



Algorithme de ligne de partage des eaux (LPE)

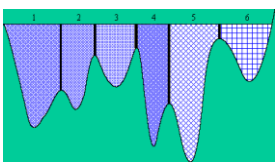
- L'image de puissance de contour est considérée comme une surface
- Simulation d'immersion



- Construction de barrages aux points de rencontre

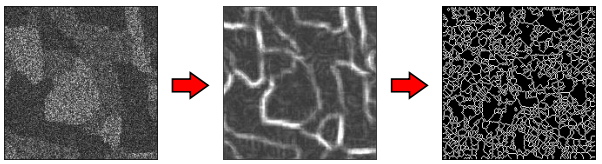
Algorithme de ligne de partage des eaux (LPE)

- L'image de puissance de contour est considérée comme une surface
- Simulation d'immersion



- Construction de barrages aux points de rencontre
- Les barrages constituent à la fin les contours

Algorithme de ligne de partage des eaux (LPE)

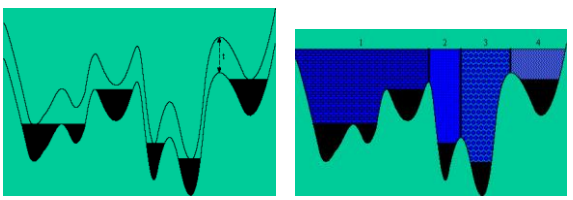


- Contours fermés et squelettisés
- Problème : sur-segmentation

- Versions plus robustes au bruit

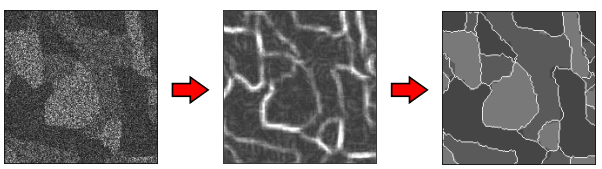
Algorithme de ligne de partage des eaux (LPE)

- Correspond à une érosion géodésique suivie par l'algorithme LPE



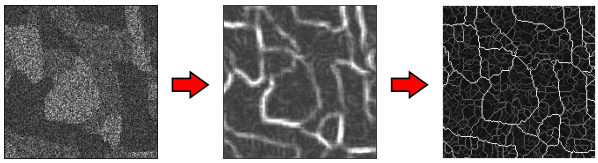
- Elimination de certains bassins

Seuillage des dynamiques de bassin



- Permet une forte réduction de la sur-segmentation par rapport à l'algorithme LPE, tout en préservant les contours significatifs.

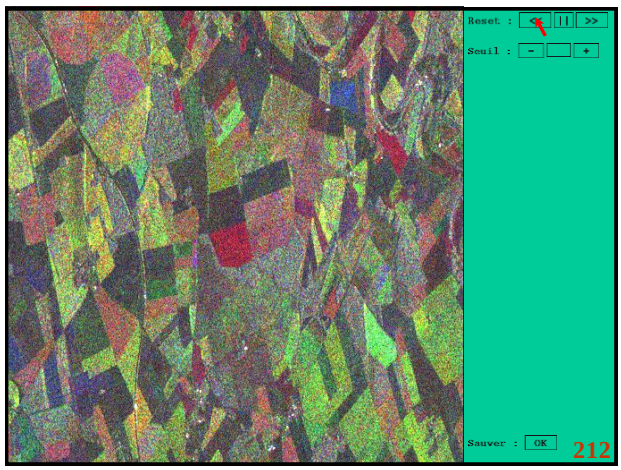
Seuillage des dynamiques de contour



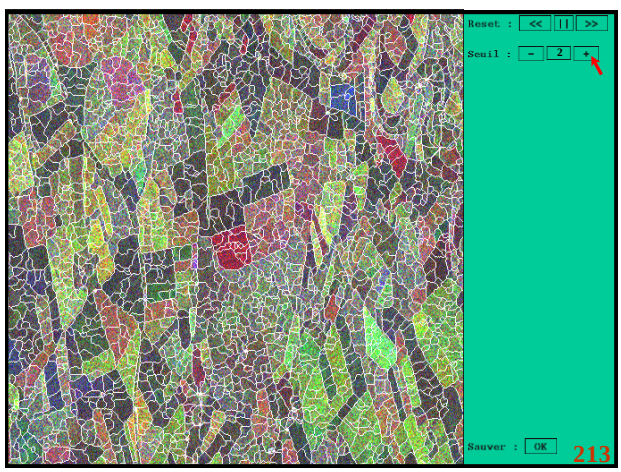
- L'attribution des dynamiques de bassin aux arcs de contour permet de créer une représentation hiérarchique des segmentations correspondant aux différents seuils.
- Possibilité de réglage interactif du seuil

2009
Ph. Marthon

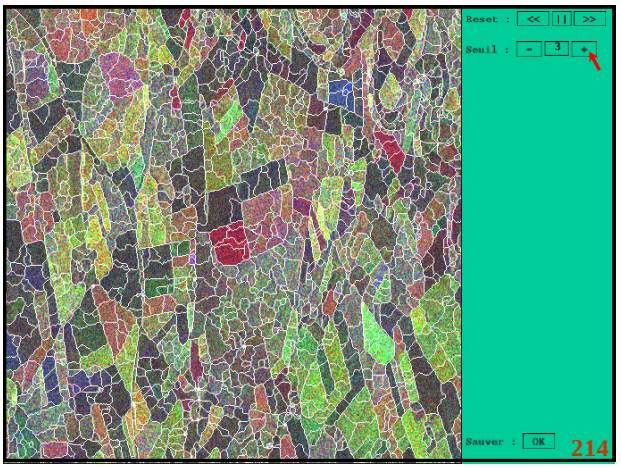
211

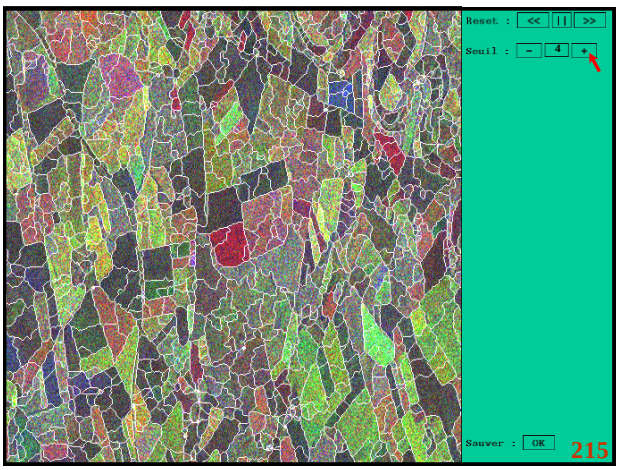


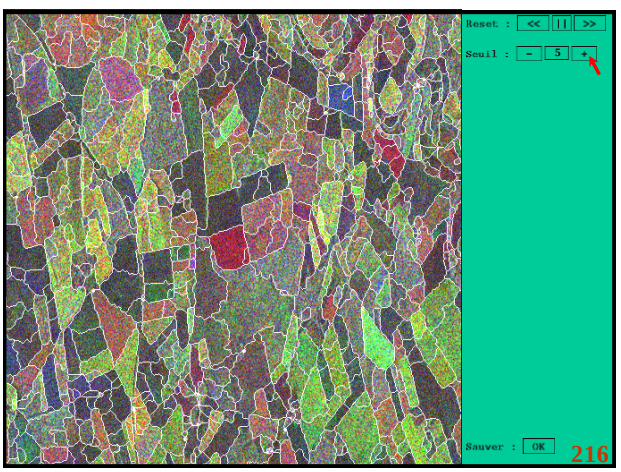
212

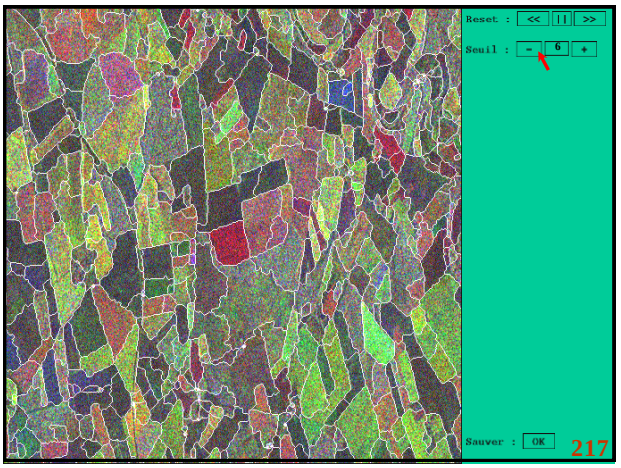


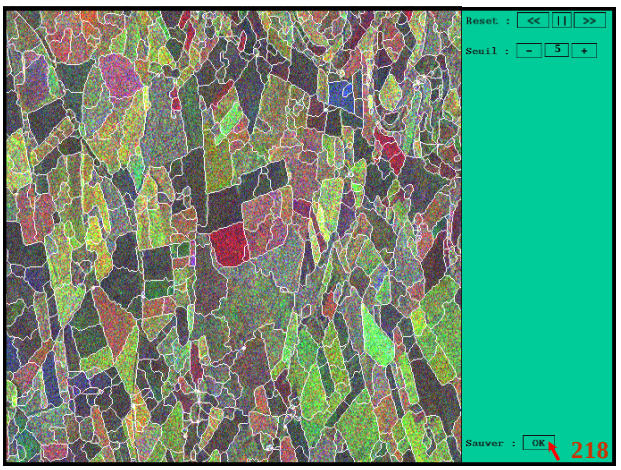
213

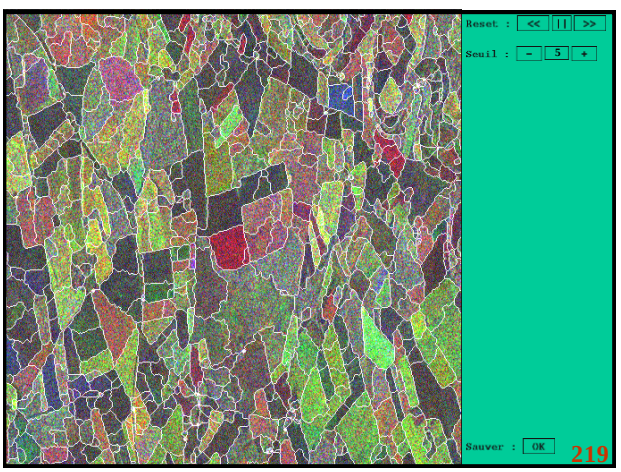

Reset : << || >>
Seuil : - 3 +
Sauver : OK
214

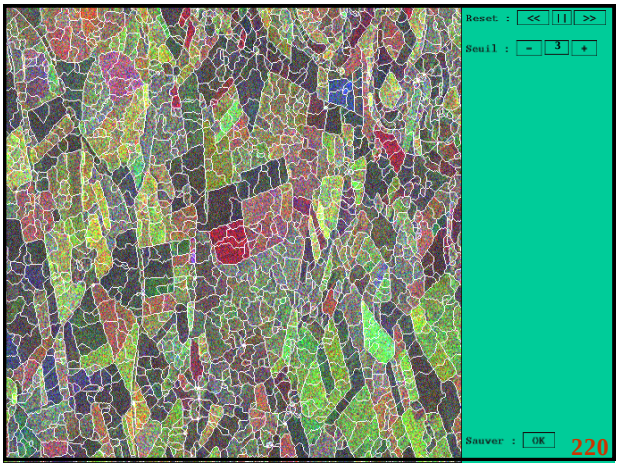

Reset : << || >>
Seuil : - 4 +
Sauver : OK
215


Reset : << || >>
Seuil : - 5 +
Sauver : OK
216


Reset : << || >>
Seuil : - 6 +
Sauver : OK
217


Reset : << || >>
Seuil : - 5 +
Sauver : OK
218


Reset : << || >>
Seuil : - 5 +
Sauver : OK
219



220

Contour actif

- Un contour actif C(t) est une courbe, fermée ou non, qui évolue au cours du temps t, de sa position initiale vers les bords de l'objet à extraire
- On cherche à minimiser une énergie :

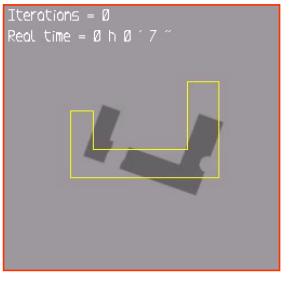
$$E(C(t))=E_1(C(t)) + E_2(C(t))$$

  - E_1 : énergie interne, régulatrice, permettant de réduire les déformations du contour dues au bruit
  - E_2 : terme d'attache aux données, faisant interagir le contour actif avec des caractéristiques de l'image

2009
Ph. Marthon

221

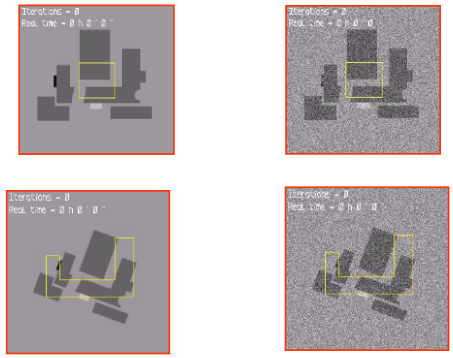
Contours actifs



2009
Ph. Marthon

222

Contours actifs



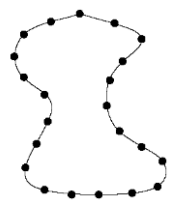
2009
Ph. Marthon

223

Représentation explicite d'un contour actif

$$C = \{C(p,t) \in R^2 \mid p \in [a,b] \subset R,\ t > 0\}$$

Après discrétisation de la variable p, le contour actif se résume à une liste de nœuds dont les positions sont mises à jour au fil de l'évolution de C

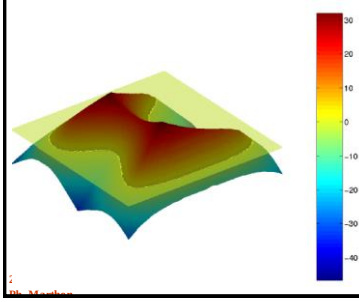


2009
Ph. Marthon

224

Représentation implicite d'un contour actif

- Ensemble de niveaux (« level sets »)
- Un contour actif fermé C(t) est représenté comme le zéro d'une fonction $\phi(x,t)$ représentant la distance signée entre le point et le contour.

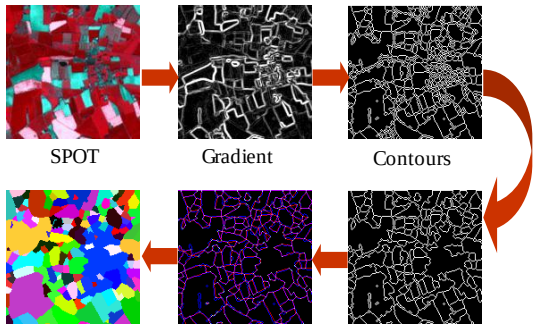


$$\phi(x,t) = 0 \quad si\ x \in C(t)$$
$$\phi(x,t) = d(x, C(t)) \quad si\ x \in \mathrm{int}(C(t))$$
$$\phi(x,t) = -d(x, C(t)) \quad si\ x \in \mathrm{ext}(C(t))$$

Ph. Marthon

225

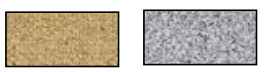
Traitement d 'images SPOT



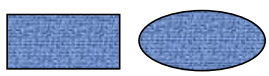
Caractérisation

Forme et texture

- Une région élémentaire se caractérise par sa forme et sa texture



Même forme
Textures distinctes



Même texture
Formes distinctes

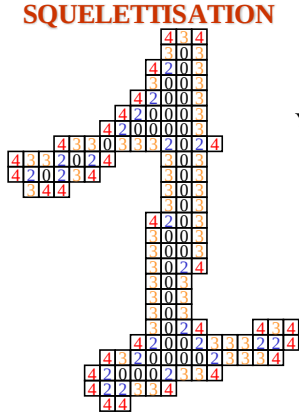
SQUELETTE

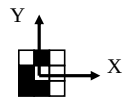
F = forme
B = fond
S = squelette

$$S = \{x \in F \mid \exists y, z \in B, y \neq z, d(x,B) = d(x,y) = d(x,z)\}$$

La squelettisation constitue un prétraitement utile pour extraire la forme d'une région

SQUELETTISATION





$|X| + |Y| = 3$

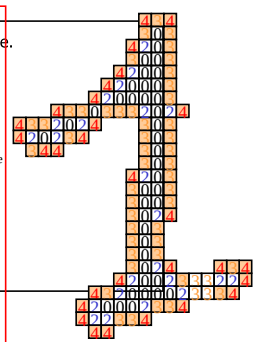
Sommes des distances locales $|X|+|Y|$

SQUELETTISATION

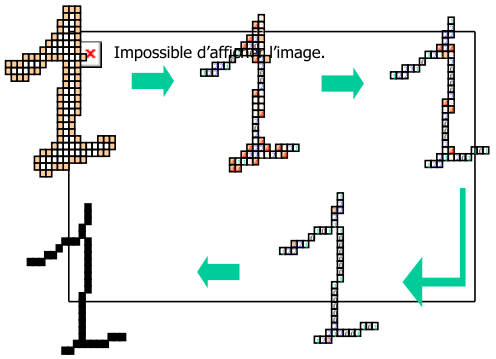
Premier balayage :
pour chaque pixel p faire
calculer $|x|, |y|, nvois$
si $|x|+|y| \neq 3$ alors $p \leftarrow |x|+|y|$
sinon
si $nvois = 2$ ou $nvois = 6$ ou ($nvois = 3$ et $\max(|x|,|y|) = 3$)
ou ($nvois = 5$ et $\max(|x|,|y|) = 3$) alors
si configuration ligne double ou colonne double
alors $p \leftarrow 3$
sinon $p \leftarrow 5$
fsi
sinon $p \leftarrow 3$
fsi
fsi
fpour

Deuxième balayage :
Eliminer tous les pixels p tels que $p \geq 4$

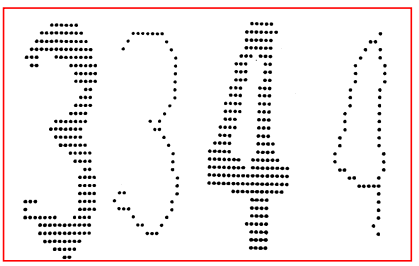
Impossible d'afficher l'image.



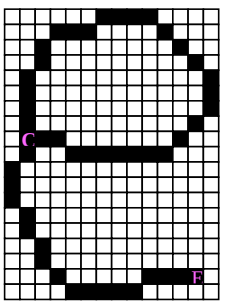
SQUELETTISATION

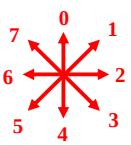


SQUELETTISATION



CHAÎNE DE FREEMAN

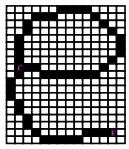




E : extrémité
C : carrefour

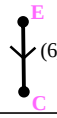
E66656666770707001000001012212223333445556666667 66C

GRAPHE DE BOUCLES



E666566667707070010000010122122233334455566666766C

6 7 0 1 2 3 4 5 6



(6, -360°)

Le « e minuscule » est formé d'une seule boucle d'amplitude -360°, partant d'une extrémité, avec une direction « ouest » et aboutissant à un carrefour.

Descripteurs géométriques

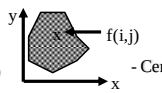
Descripteurs géométriques

- Surface, périmètre
- Compacité : Périmètre^2 / Surface (min : cercle)

MOMENTS

$M_{pq} = \sum_i \sum_j x_i^p . y_j^q . f(i,j)$

M_{00}=surface (image binaire)



- Centre de gravité : $X_{cg}=M_{10}/M_{00}$ $Y_{cg}=M_{01}/M_{00}$

MOMENTS CENTRES

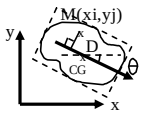
$$\mu_{pq} = \sum_i \sum_j (xi-X_{cg})^p . (yj-Y_{cg})^q . f(i,j)$$

Propriété : invariance en translation

ORIENTATION

$I(\theta) = \sum_i \sum_j D^2(i,j)$

min pour : $\frac{\delta I}{\delta \theta} = 0$

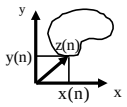


$$\theta = 1/2 * \mathrm{arctg}\left[\frac{2 * \mu_{11}}{\mu_{20} - \mu_0}\right]$$

-> Rectangle minimum
-> Ellipse équivalente
-> Exentricité : Grand axe / Petit axe

Descripteurs de Fourier

Contour d'un objet = courbe fermée



z(n)=x(n)+i y(n)
n : abscisse curviligne

$$z(n) = \frac{1}{N}\sum_k Z(k) e^{\frac{2\pi j k n}{N}}$$

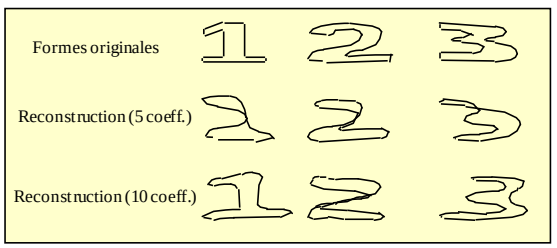
La fonction z(n) est périodique donc développable en série de Fourier

$$\boxed{Z(k) = \sum_n z(n) e^{\frac{-2\pi j k n}{N}}}$$

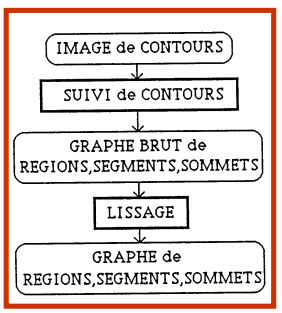
descripteurs

5 étapes :
- Suivi de contour -> liste des N points ordonnés
- Interpolation pour avoir un nombre en puissance de 2
- Transformée de Fourier rapide (FFT)
- Filtrage (suppression de coefficients)
- Transformée inverse (IFFT) : reconstruction contour

Descripteurs de Fourier



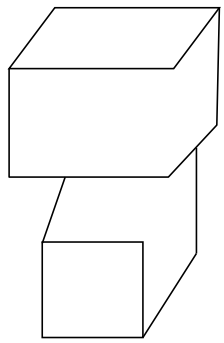
Suivi de contours



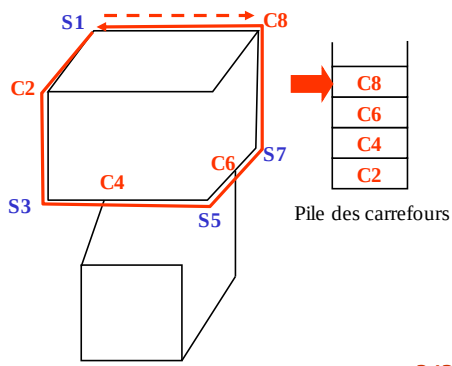
SUIVI ET APPROXIMATION DES CONTOURS

- **SUIVI D 'UNE COMPOSANTE CONNEXE**
  - Passer d 'un point de contour à un pixel de contour voisin
  - Eviter de repasser sur les mêmes pixels : chaque pixel visité doit être effacé
  - Pour tout visiter, on mémorise dans une pile, tous les pixels de contour qui sont des **CARREFOURS** (au moins 3 voisins)
- **APPROXIMATION**
  - Chaque portion de contour entre 2 carrefours est segmenté en segments de droite ou en portions de courbe.

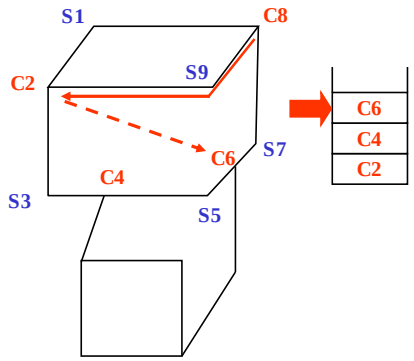
SUIVI



SUIVI



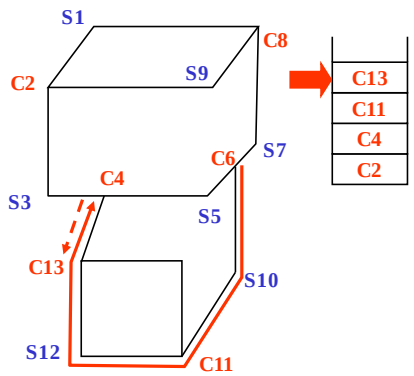
SUIVI



2009
Ph. Marthon

244

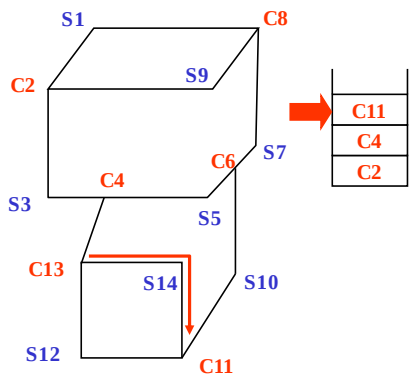
SUIVI



2009
Ph. Marthon

245

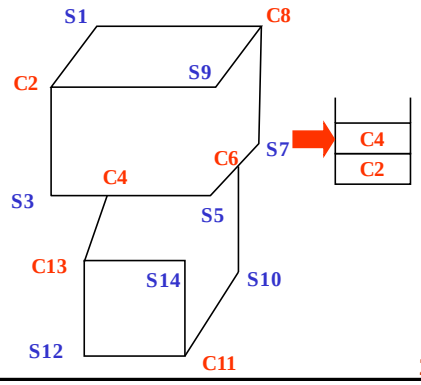
SUIVI



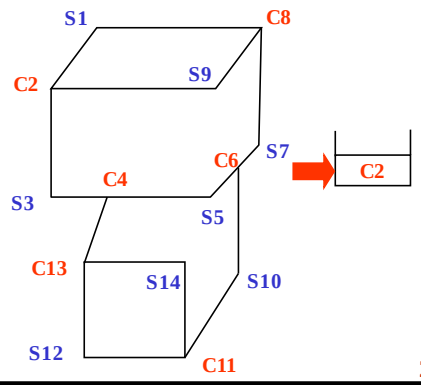
2009
Ph. Marthon

246

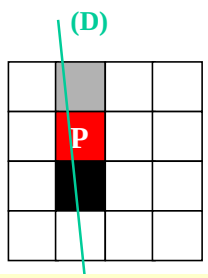
SUIVI



SUIVI



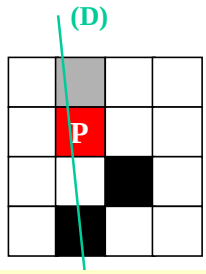
APPROXIMATION POLYGONALE



P NON SOMMET

- (D) est la droite d'approximation courante du segment de contour; c'est l'**axe principal d'inertie**, soit la 1ère composante principale du « nuage » de pixels de contour visités.

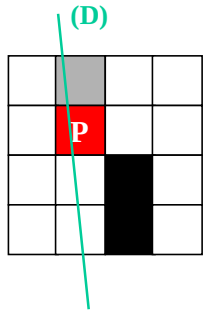
APPROXIMATION POLYGONALE



P NON SOMMET

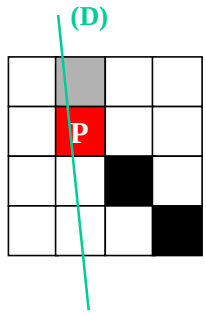
- (D) est la droite d'approximation courante du segment de contour; c'est l'**axe principal d'inertie**, soit la 1ère composante principale du « nuage » de pixels de contour visités.

APPROXIMATION POLYGONALE



P NON SOMMET

APPROXIMATION POLYGONALE



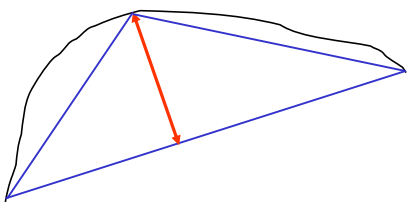
P SOMMET

APPROXIMATION POLYGONALE

- **P** : pixel courant, **(D)** : droite d'approximation courante du segment
- **Etape 1** : on cherche un pixel voisin de P, non encore visité, dont la distance à (D) soit inférieure à **0,5**.
- **Etape 2** :
  - si ce voisin existe, appelons-le **Q**
    - Q devient le pixel courant : (P <-- Q)
    - Q est ajouté à la liste des pixels servant à l'estimation de (D)
    - Retour à l'étape 1
  - si ce voisin n'existe pas
    - on détermine **Q'**, le pixel le plus proche de (D) et DISTMIN sa distance à (D) (=> DISTMIN > 0,5)
    - on cherche un pixel voisin de Q', non encore visité, dont la distance à (D) soit inférieure à DISTMIN
      - si ce voisin existe
        Q' devient le pixel courant : (P <-- Q')
        Q' est ajouté à la liste des pixels servant à l'estimation de (D)
        Retour à l'étape 1
      - si ce voisin n'existe pas
        P est un **sommet**

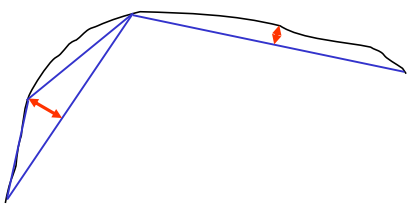
DECOUPAGE RECURSIF

Ecart maximal autorisé



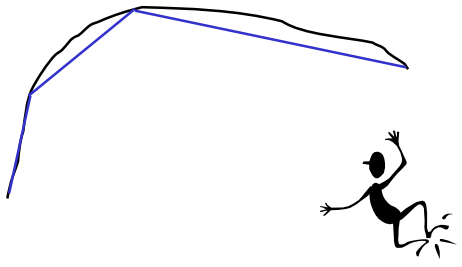
DECOUPAGE RECURSIF

Ecart maximal autorisé

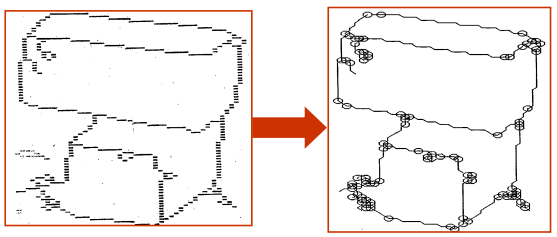


DECOUPAGE RECURSIF

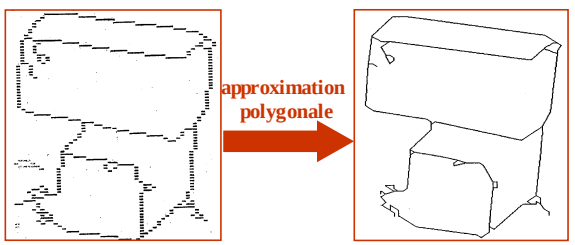
Ecart maximal autorisé



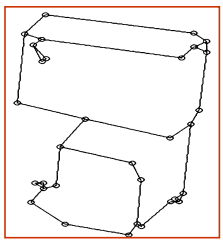
DETECTION DES SOMMETS

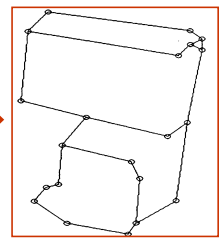


APPROXIMATION POLYGONALE



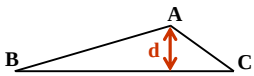
LISSAGE





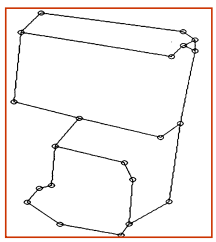
32 sommets

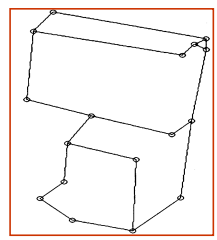
21 sommets



HAUTEUR=1,1*HAUTEUR

LISSAGE

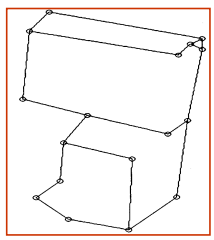


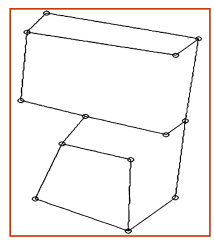


21 sommets

15 sommets

LISSAGE

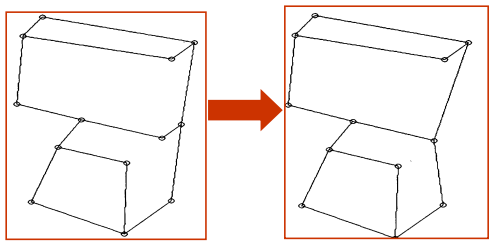




15 sommets

13 sommets

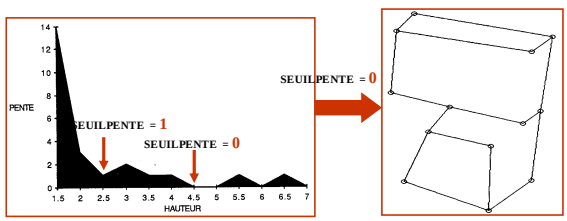
LISSAGE



13 sommets

12 sommets

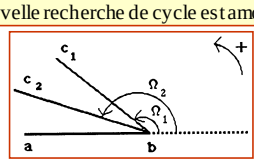
CRITERE D'ARRET DU LISSAGE



- L'algorithme s'arrête quand le nombre de sommets éliminés lors d'un accroissement de **HAUTEUR** de **0,5** pixel, noté **PENTE**, est inférieur à un seuil (noté **SEUILPENTE**), inférieur ou égal à **1** en pratique.

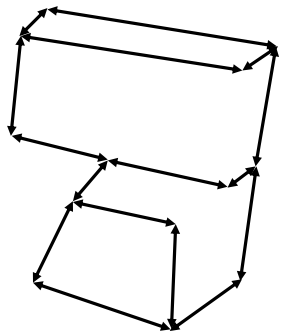
EXTRACTION DES REGIONS

- On oriente le graphe de sommets en remplaçant chaque arête par 2 arcs opposés (graphe G). On choisit un **arc (a,b)** de G. Cet arc va déterminer un **cycle minimal** ou la **silhouette** (une région): **on cherche un sommet voisin de b et différent de a, c, tel que l'angle orienté [(a,b),(b,c)] soit le plus grand possible.**
- On répète ensuite la même opération mais en faisant jouer à b le rôle de a et à c celui de b. On arrête dès qu'un cycle est trouvé. Ce cycle est retiré de G. Si G n'est pas vide, une nouvelle recherche de cycle est amorcé.

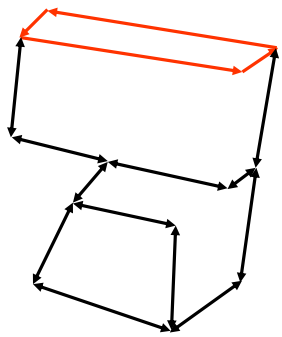


Venant de a, c_2 est préféré à c_1

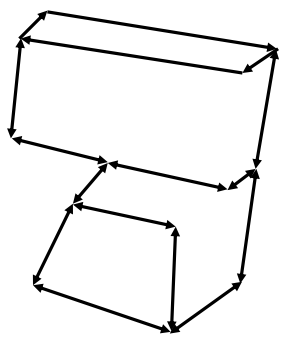
EXTRACTION DES REGIONS



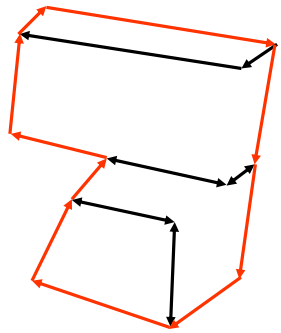
EXTRACTION DES REGIONS



EXTRACTION DES REGIONS

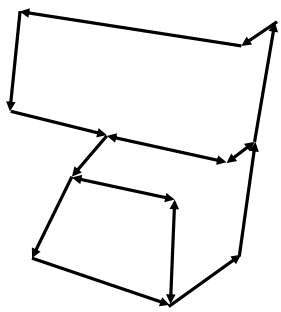


EXTRACTION DES REGIONS



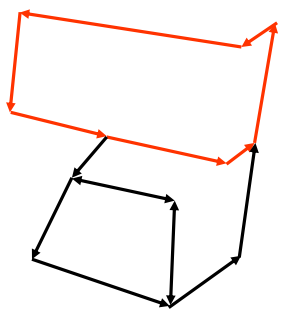
2009
Ph. Marthon
268

EXTRACTION DES REGIONS



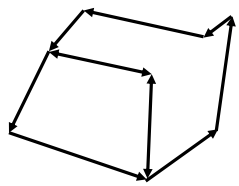
2009
Ph. Marthon
269

EXTRACTION DES REGIONS

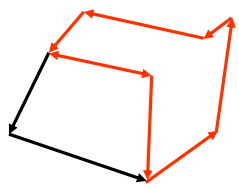


2009
Ph. Marthon
270

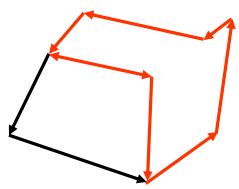
EXTRACTION DES REGIONS



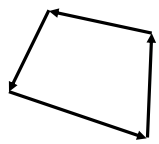
EXTRACTION DES REGIONS



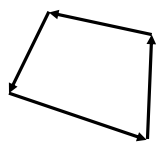
EXTRACTION DES REGIONS



EXTRACTION DES REGIONS



EXTRACTION DES REGIONS



Arbre de concavité

Z objet digital

PCE(Z) : polygone convexe exinscrit de Z

EC(Z) : enveloppe convexe de Z =ensemble des points intérieurs au PCE(Z)

DEC(Z) : déficience d'enveloppe convexe DEC(Z) = EC(Z)-Z

La DEC(Z) comprend 2 parties : les trous et les baies

Objet *Enveloppe convexe* *Décomposition*

Chaque concavité peut être assimilée à un objet et peut être décomposée à son tour.

L'arbre de concavité comprend plusieurs niveaux :

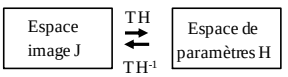
- *niveau 1* : objet et ses descripteurs
- *niveau 2* : concavités de l'objet et descripteurs
- *niveau 3* : concavités des concavités et descripteurs
- *etc....*

Transformation de Hough

- Outil classique de l'analyse d'images utilisé pour détecter la présence de courbes ayant une forme paramétrique (droite, conique, ...)
- Principe: à chaque « point caractéristique », on associe le faisceau de courbes passant par ce point.

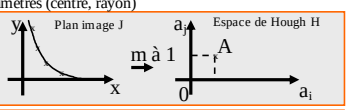
Transformation de Hough

PRINCIPE:

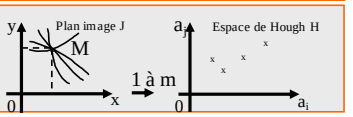


Espace image défini par $\{M_i\ (x_i, y_i)\}$
Un ensemble de points définit une courbe décrite par une équation paramétrique
$f(a_1, a_2, ..a_n, x, y)=0$
ex : une droite est définie par 2 paramètres (y=ax+b)
un cercle est défini par 3 paramètres (centre, rayon)

Transformation de m à 1
A tout m-uple de l'espace image $M_i(m)$ est associée une courbe de paramètre $\{a_i\}$



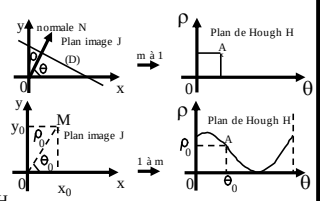
Transformation de 1 à m
Par un point M_i passent m courbes de la famille $\{a_i\}$



Transformation de Hough

DETECTION DE DROITES

Equation caractéristique d'une droite :
$\rho = x\cos(\theta) + y\sin(\theta)$



PROPRIETES
- A un point de H correspond une droite de J
- A un point de J correspond une sinusoïde de H
- Des points de la même droite de J donnent des courbes passant par un point commun de H
- Des points de la même courbe de H donnent des droites passant par un point commun de J

APPLICATION : recherche d'une structure d'image
a) TH des points caractéristiques de J
b) classification dans H
c) transformation inverse sur les points de H retenus

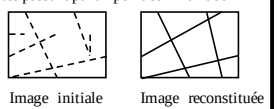


Image initiale Image reconstituée

TRANSFORMEE DE HOUGH

- Détection
  - Principe: accumulation d'évidences
  - Pratique (cas de droites)
    - On construit un histogramme 2D HIST(ρ,θ)
    - Pour chaque pixel significatif (x_i , y_i) faire:
      - HIST(ρ_m,θ_m) = HIST(ρ_m,θ_m) + 1 où $\rho_m = x_i \cos\theta_m + y_i \sin\theta_m$
    - La (ou les) meilleure(s) droite(s) sera (seront) associé(es) aux maxima présents dans l'histogramme

REGROUPEMENT

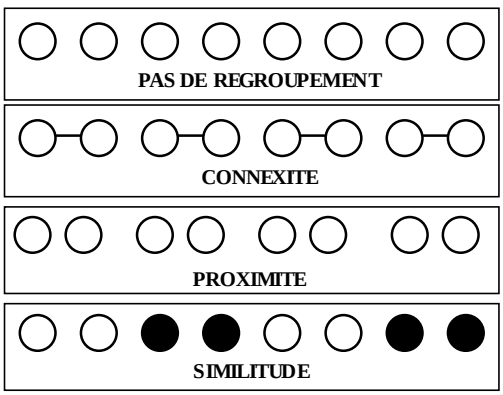
REGROUPEMENT

- En général, les différentes régions extraites par détection de contour et/ou de texture ne correspondent pas aux images des objets de la scène mais seulement à des sous-images de ces objets
- Il est donc nécessaire d'opérer des regroupements avant d'aborder l'étape de reconnaissance
- Les lois de regroupement utilisés par le système visuel humain ont été essentiellement découvertes par des psychophysiciens regroupés dans l'école de la Gestalt (1930)

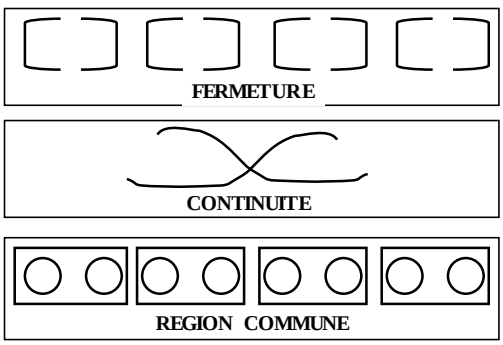
LOIS DE LA GESTALT

- **Loi générale de PREGNANCE**
  - Parmi toutes les structures possibles, il y en a une qui est **PREDOMINANTE**, qui s'impose de préférence aux autres
- **Loi de la BONNE FORME**
  - La forme « PREGNANTE » est toujours celle qui présente le maximum d'unité, de simplicité, de régularité
- **Loi générale la SEGREGATION DES UNITES**
  - Dans l'ensemble du champ, certaines unités perceptives se constituent spontanément selon des lois :
    - de PROXIMITE des éléments
    - de SIMILITUDE entre les éléments
    - de DIRECTION COMMUNE des éléments

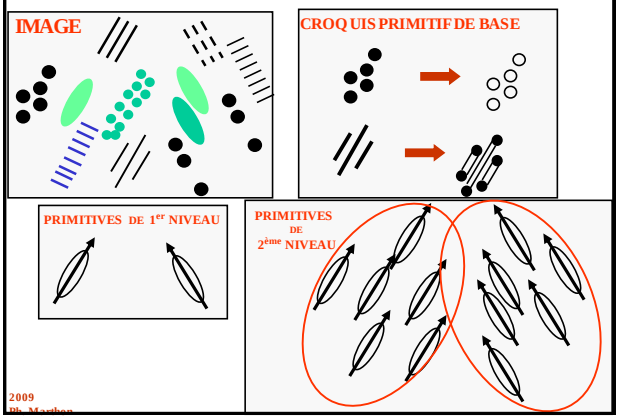
LOIS DE REGROUPEMENT



LOIS DE REGROUPEMENT



CROQUIS PRIMITIF



2009
Ph. Marthon

FIGURE et FOND

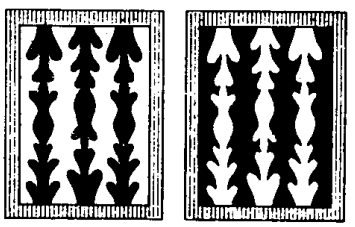
- **Loi générale la FIGURE et du FOND**

  Apparaît généralement comme **figure**:
  - ce qui a un **contour** déterminé par opposition au fond qui n'en a pas
  - ce qui est **articulé, différencié** par opposition au fond qui est plus homogène, plus uniforme
  - ce qui est **enveloppé** par le fond
  - ce qui correspond à certaines **directions privilégiées** de l'espace
  - ce qui est **symétrique** par opposition au fond asymétrique

2009
Ph. Marthon

287

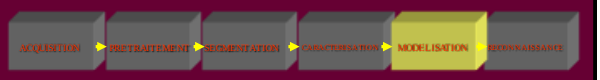
SYMETRIE



2009
Ph. Marthon

288

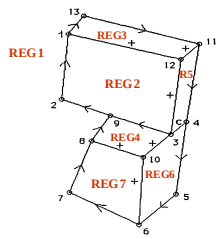
Modélisation



Modélisation d'une image

- Graphe (hiérarchique) de régions adjacentes
- Régions
  - paramètres géométriques
    - périmètre
    - aire
    - moments géométriques
    - descripteurs de Fourier …
  - paramètres radiométriques
    - histogrammes de couleur …
  - paramètres de texture
    - paramètres des champs de Gibbs-Markov
    - primitives extraites des filtres de Gabor
    - primitives calculées à partir des matrices de co-occurrence …

Objets polyédriques



est la suivante:
LISTEREGIONS = '(REG1 REG2 REG3 REG4 REG5 REG6 REG7)
REG1 = \(SEG17 SEG18 SEG22 SEG23 SEG4 SEG3 SEG35 SEG31 SEG12 SEG13)

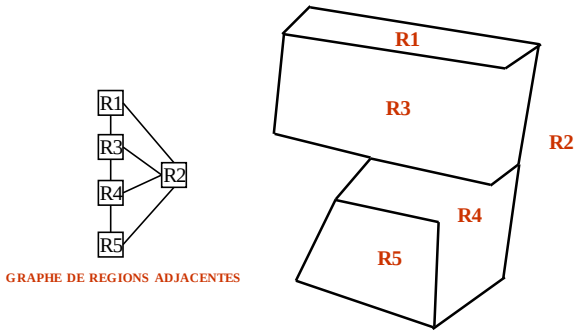
REG2 = \(SEG1 SEG5 SEG25 SEG9 SEG32)
REG3 = \(SEG2 SEG33 SEG29 SEG36)
REG4 = \(SEG6 SEG24 SEG21 SEG28)
REG5 = \(SEG7 SEG10 SEG30 SEG34)
REG6 = \(SEG8 SEG26 SEG15 SEG14 SEG11)
REG7 = \(SEG16 SEG27 SEG20 SEG19)

SEG1 = \(SOM1 SOM2) ; SEG2 = \(SOM1 SOM12) ; SEG3 = \(SOM1 SOM13)
SEG4 = \(SOM2 SOM1) ; SEG5 = \(SOM2 SOM9)
SEG6 = \(SOM3 SOM9) ; SEG7 = \(SOM3 SOM4) ; SEG8 = \(SOM3 SOM10);SEG9 = \(SOM3 SOM1
SEG10 = \(SOM4 SOM11) ; SEG11 = \(SOM4 SOM3) ; SEG12 = \(SOM4 SOM2)
SEG13 = \(SOM5 SOM6) ; SEG14 = \(SOM5 SOM4)
SEG15 = \(SOM6 SOM5) ; SEG16 = \(SOM6 SOM10) ; SEG17 = \(SOM6 SOM7)
SEG18 = \(SOM7 SOM8) ; SEG19 = \(SOM7 SOM6)
SEG20 = \(SOM8 SOM7) ; SEG21 = \(SOM8 SOM10) ; SEG22 = \(SOM8 SOM9)
SEG23 = \(SOM9 SOM2) ; SEG24 = \(SOM9 SOM8) ; SEG25 = \(SOM9 SOM5)
SEG26 = \(SOM10 SOM6) ; SEG27 = \(SOM10 SOM8) ; SEG28 = \(SOM10 SOM3)
SEG29 = \(SOM11 SOM13) ; SEG30 = \(SOM11 SOM12) ; SEG31 = \(SOM11 SOM4)
SEG32 = \(SOM12 SOM1) ; SEG33 = \(SOM12 SOM11) ; SEG34 = \(SOM12 SOM3)
SEG35 = \(SOM13 SOM11) ; SEG36 = \(SOM13 SOM1)

LISTEINT3D = (< + > > < > C + + < C > > < < + > > < < + > > < < + + + < + > + + + > <

SOM1 = '(15 5) ; Type = \(FLECHE SEG1 SEG2 SEG3)
SOM2 = '(38,4) ; Type = \(L SEG4 SEG5)
SOM3 = '(51 40) ; Type = \(K SEG6 SEG7 SEG8 SEG9)
SOM4 = '(48 45) ; Type = \(TE SEG31 SEG11 SEG12)
SOM5 = '(74 43) ; Type = \(L SEG13 SEG14)
SOM6 = '(86 10) ; Type = \(FLECHE SEG15 SEG16 SEG17)
SOM7 = '(82 12) ; Type = \(L SEG18 SEG19)
SOM8 = '(52 15) ; Type = \(FLECHE SEG20 SEG21 SEG22)
SOM9 = '(45 19) ; Type = \(TE SEG5 SEG24 SEG25)
SOM10 = '(60 31) ; Type = \(FOURCHE SEG26 SEG27 SEG28)
SOM11 = '(16 48) ; Type = \(FLECHE SEG29 SEG30 SEG31)
SOM12 = '(22 43) ; Type = \(FOURCHE SEG32 SEG33 SEG34)
SOM13 = '(6 12) ; Type = \(L SEG35 SEG36).

SEGMENTATION



MISE EN RELATIONS

Apprentissage

FORMATION DES IMAGES

1.4. Formation des images

- Etablir les relations géométriques et radiométriques entre la scène, la lumière, les capteurs et les images
- Différence géométrique entre optique et radar
  - Imagerie optique : séparation angulaire de l'espace
  - Imagerie radar : séparation de l'espace en distance

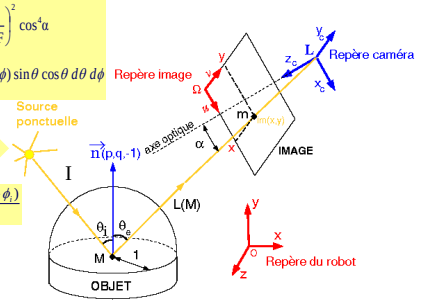
FORMATION DE L 'IMAGE

- **Formules générales**

$$im(m) = im(x,y) = \mathrm{L(M)}\ \pi \left(\frac{d}{2F}\right)^2 \cos^4\alpha$$

$$\mathrm{L(M)} = \int_{-\pi}^{\pi}\int_{0}^{\pi/2} f(\theta,\phi,\theta_i,\phi_i)\, E(\theta,\phi)\sin\theta\cos\theta\, d\theta\, d\phi$$

$$E(\theta,\phi) = \mathrm{I}\,\frac{\delta(\theta-\theta_i)\delta(\phi-\phi_i)}{\sin\theta_i}$$



VECTEUR GRADIENT

- On considère le repère de la caméra ($Lx_cy_cz_c$).
- Soit $z_c = F(x_c, y_c)$ l 'équation donnant la profondeur de la surface - supposée différentiable - d 'un objet visible de cette caméra
- Le vecteur

$$\vec{n} = \left(\frac{\partial F}{\partial x_c}(x_c,y_c), \frac{\partial F}{\partial y_c}(x_c,y_c), -1\right)$$

est un vecteur normal à la surface au point $(x_c, y_c, F(x_c, y_c))$ et dirigé vers l 'extérieur de l 'objet

- On définit le « **vecteur gradient** » $(p,q)^T$ en M :

$$p = \frac{\partial F}{\partial x_c}(x_c,y_c)$$
$$q = \frac{\partial F}{\partial y_c}(x_c,y_c)$$

CARTE DE REFLECTANCE

- **Elle relie l'orientation du surfel à sa luminance dans la direction de la caméra**

$$L(x_c, y_c, z_c) = \frac{im(x,y)}{\pi (d/2f)^2 \cos^4 \alpha}$$

alors on appelle carte de réflectance l'application R :

$$(p,q) \rightarrow R(p,q) = L(x_c, y_c, z_c)$$

où le point $m(x,y)$ est l'image du point $M(x_c, y_c, z_c) \Rightarrow x = f\frac{x_c}{z_c} \quad y = f\frac{y_c}{z_c}$

$$R(p,q) = L\left(\frac{fx_c}{z_c}, \frac{fy_c}{z_c}\right)$$

$$p = \frac{\partial F}{\partial x_c}(x_c, y_c) \qquad q = \frac{\partial F}{\partial y_c}(x_c, y_c)$$

CARTE DE REFLECTANCE

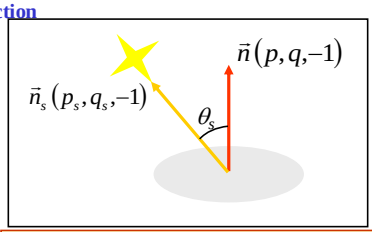
Si la surface est mate, la source ponctuelle située sur l'axe optique et la projection orthogonale:

$$\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 = \left(\frac{L_{\max}}{L(x,y)}\right)^2 - 1$$

Équation de l'eikonale

EXERCICE

- Calcul de la **carte de réflectance** dans le cas d'une surface de **Lambert** éclairée par une source **ponctuelle à l'infini, de direction (p_s, q_s,-1) dans le repère caméra et d'intensité I dans cette direction**



$$R(p,q) = \frac{\rho I}{\pi} \frac{p_s p + q_s q + 1}{\sqrt{p_s^2 + q_s^2 + 1}\sqrt{p^2 + q^2 + 1}}$$

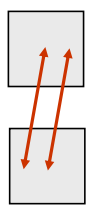
RELATION GEOMETRIQUE ENTRE IMAGES

POINTS HOMOLOGUES

- Deux **points** image sont **homologues** s 'ils sont les images d 'un même point 3D de la scène observée

PIXELS HOMOLOGUES

- Deux **pixels** sont **homologues** si l 'homologue du point central de l 'un appartient à l 'autre et vice-versa



CONTRAINTE D 'UNICITE

- Soit m^g un **point** de l 'image gauche alors:
  - soit il possède un homologue unique sur l 'image droite
  - soit il ne possède pas d 'homologue car
    - il est l 'image d 'un point M qui n 'appartient pas au champ visuel de la caméra droite (L^dM ne coupe pas le capteur)
    - le point M est occulté depuis le centre optique droit

➡ **Tout point image possède au plus un point homologue**

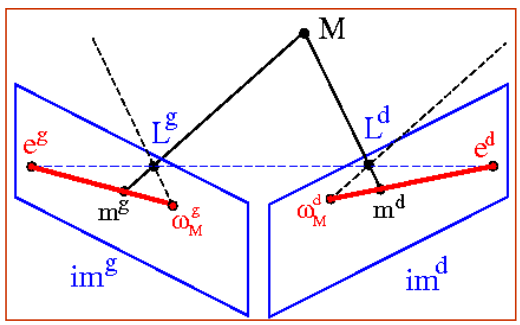
CONTRAINTE D 'UNICITE

- Les résultats précédents restent vrais pour un **pixel** au lieu d 'un **point**.
- On doit ajouter qu 'un pixel peut ne pas posséder d 'homologue lorsque la surface de l 'objet est très **inclinée** par rapport à la ligne de base L^gL^d :

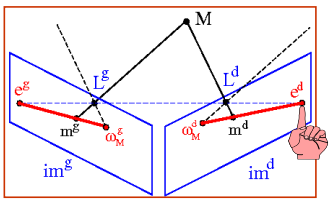
  le nombre de pixels de l 'image gauche de cette surface diffère du nombre de pixels de l 'image droite de cette surface (phénomène de **raccourcissement**)
- On peut encore énoncer:

tout pixel possède au plus un pixel homologue

GEOMETRIE EPIPOLAIRE

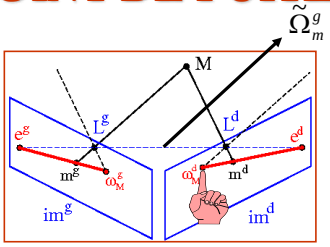


EPIPOLE



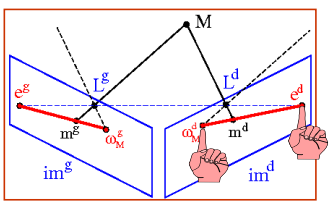
$$\tilde{e}^d = \tilde{C}^d \tilde{L}^g = \left[C^d \ \tilde{c}_4^d\right] \begin{bmatrix} -\left(C^g\right)^{-1} \tilde{c}_4^g \\ 1 \end{bmatrix} = \tilde{c}_4^d - C^d \left(C^g\right)^{-1} \tilde{c}_4^g$$

POINT DE FUITE



$$\tilde{\omega}_M^d = \tilde{C}^d \tilde{\Omega}_m^g = \left[C^d \ \tilde{c}_4^d\right] \begin{bmatrix} \left(C^g\right)^{-1} \tilde{m}^g \\ 0 \end{bmatrix} = C^d \left(C^g\right)^{-1} \tilde{m}^g$$

DROITE EPIPOLAIRE



$$\left(D^d\right) = \left\{\tilde{m} \in P^2 \middle| \exists \alpha, \beta \in \mathrm{R} \middle| \tilde{m} = \alpha \tilde{e}^d + \beta \tilde{\omega}_M^d\right\}$$
$$= \left\{\tilde{m} \in P^2 \middle| \tilde{m}^T \left(\tilde{e}^d \wedge \tilde{\omega}_M^d\right) = 0\right\}$$

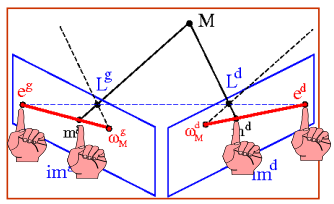
MATRICE FONDAMENTALE

Soit $\tilde{E}^d \,\Big|\, \tilde{E}^d \tilde{x} = \tilde{e}^d \wedge \tilde{x} \quad \mathrm{rg}\left(\tilde{E}^d\right) = 2$

Soit $\tilde{F} = \tilde{E}^d C^d \left(C^g\right)^{-1}$ matrice fondamentale $\mathrm{rg}\left(\tilde{F}\right) = \mathrm{rg}\left(\tilde{E}^d\right) = 2$

$$\left(D^d\right) = \left\{\tilde{m} \in P^2 \,\middle|\, \tilde{m}^T \tilde{F} \tilde{m}^g = 0\right\}$$

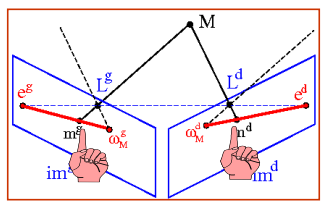
DROITES EPIPOLAIRES CONJUGUEES



$$\left(D^d\right) = \left\{\tilde{m} \in P^2 \,\middle|\, \tilde{m}^T \tilde{F} \tilde{m}^g = 0\right\}$$

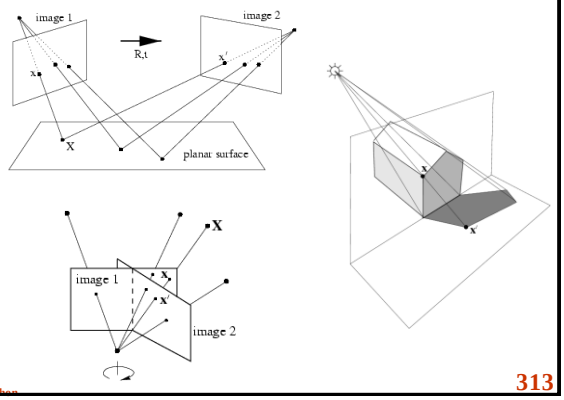
$$\left(D^g\right) = \left\{\tilde{m} \in P^2 \,\middle|\, \tilde{m}^T \tilde{F}^T \tilde{m}^d = 0\right\}$$

POINTS HOMOLOGUES



$$\tilde{m}^{d\,T} \tilde{F} \tilde{m}^g = 0$$

Appariement plan



Appariement plan

- La transformation entre 2 images d'un même plan est une homographie: y = Hx où

$$H = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}$$

- 4 couples de points en correspondance permettent de calculer H
- En pratique, on peut calculer H comme vecteur singulier correspondant à la plus petite valeur singulière de la SVD de la matrice A(2n,9) du système linéaire homogène résultant de la mise en correspondance de n >> 4 couples de points (algorithme DLT)

Suivi de visages par modélisation AAM 2D

SUIVI

RELATION RADIOMETRIQUE ENTRE IMAGES

Relation radiométrique entre deux images

- Si la surface est mate, les luminances des pixels homologues sont les mêmes (pas d'effet directionnel)

$$L^g(x,y) = L^d(x + d_l(x,y), y + d_c(x,y))$$

d : disparité

Attention : il n'y a pas égalité des niveaux de gris de deux pixels homologues

Etape 2
RESOLUTION

Résolution

- **Ontologique**
  - Reconnaissance des objets, interprétation de la scène
- **Géométrique** (**stéréovision, photogrammétrie**)
  - Calcul des formes 3D des objets, de leurs positions et orientations (absolues et relatives)
- **Radiométrique**
  - Calcul des fonctions de réflectance
- **Temporelle**
- **Cinématique** (stéréovision du mouvement)

Ajustage (fitting) d'un AAM

- *On minimise sur* (p,λ)

$$\sum_{x\in s_0}\left[M(W(x;p))-I(W(x;p))\right]^2=\sum_{x\in s_0}\left[A_0(x)+\sum_{i=1}^{m}\lambda_i A_i(x)-I(W(x;p))\right]^2$$

RESOLUTION

STEREOVISION

Stéréovision

- Toute méthode de stéréovision possède 2 caractéristiques essentielles :
  - Elle est **monoculaire** ou **multi-oculaire**
  - Elle utilise ou non du **traitement d'image(s)**

RESOLUTION

STEREOVISION OPTIQUE

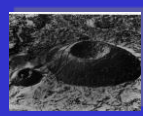
Stéréovision monoculaire

- Stéréovision **SANS** traitement d'images
  - Stéréophotométrie
  - Analyse des variations de luminance (photoclinométrie, shape from shading)
- Stéréovision **AVEC** traitement d'images
  - Analyse des contours et des lignes
  - Analyse des variations de texture
  - Calcul des points de fuite

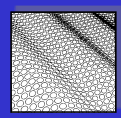
Stéréovision monoculaire passive

- Photoclinométrie

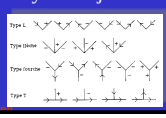
$$\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 = \left(\frac{L_{max}}{L(x, y)}\right)^2 - 1$$

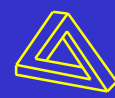


- Texturoclinométrie



- Analyse des jonctions





RESOLUTION

STEREOPHOTOMETRIE

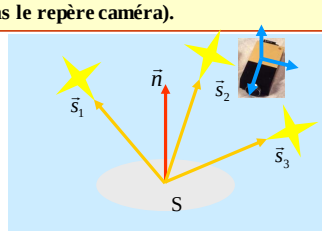
STEREOPHOTOMETRIE

- On produit plusieurs (n) éclairages de la scène.
- Scène et caméra restent fixes.
- On obtient n cartes de réflectance pour chaque pixel

$$R_i(p,q) = L_i(x_c, y_c, z_c) \quad i = 1, \ldots, n$$

STEREOPHOTOMETRIE : EXERCICE

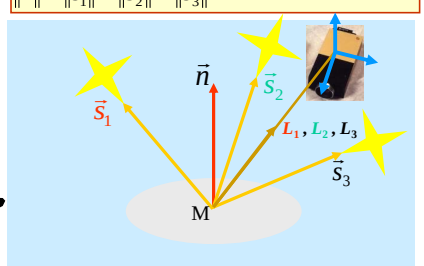
- On éclaire successivement une **surface de Lambert** par une source ponctuelle éloignée d'intensité **I = 1**, dont la position varie successivement. L'orientation de la source est toujours connue (dans le repère caméra).
- Montrer qu'en général **3** éclairages sont nécessaires et suffisants pour trouver l'**albedo et l'orientation de la surface (dans le repère caméra).**



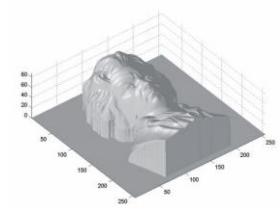
STEREOPHOTOMETRIE : EXERCICE

$$\frac{\rho}{\pi}(\vec{n}|\vec{s}_1) = L_1 \quad \frac{\rho}{\pi}(\vec{n}|\vec{s}_2) = L_2 \quad \frac{\rho}{\pi}(\vec{n}|\vec{s}_3) = L_3$$

$$\|\vec{n}\| = \|\vec{s}_1\| = \|\vec{s}_2\| = \|\vec{s}_3\| = 1$$



STEREOPHOTOMETRIE



2010
Ph. Marthon

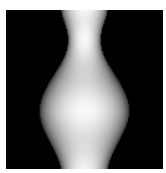
331

STEREOVISION DES LUMINANCES

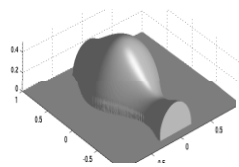
2009
Ph. Marthon

332

Shape From Shading (Photoclinométrie) / 1



?

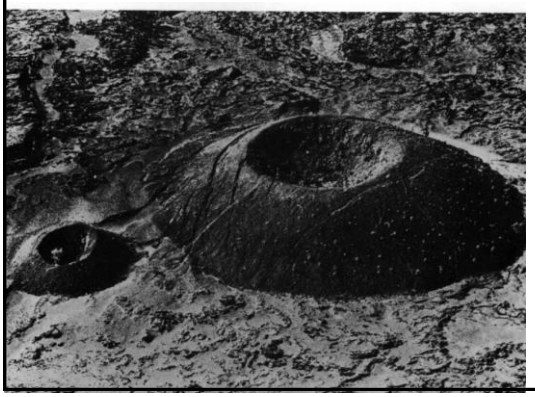


Comment retrouver la forme d'un objet (3D) à partir d'une image (2D) de celui-ci ?

2009
Ph. Marthon

333

Shape From Shading / 2



334

Shape From Shading / 3

335

Shape from shading (optique)

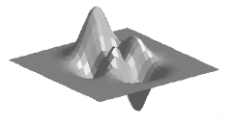
- Résolution de l'équation de l'eikonale, équation aux dérivées partielles du premier ordre, non linéaire …

$$\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2 = \left(\frac{L_{\max}}{L(x,y)}\right)^2 - 1$$

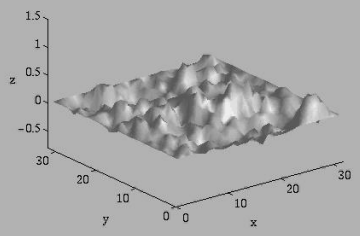
Shape from shading (optique)

- Exemple : On cherche le relief correspondant à l'image de synthèse suivante (la donnée) :

- Le relief ayant servi à calculer cette image est le suivant (l'inconnue) :



Shape from shading (optique)



Réalisation : **Jean-Denis DUROU**
Algorithme : **Recuit simulé**

RECUIT SIMULE

- Calculer une solution (réalisable) initiale X
- $X_{meilleur} \leftarrow X$
- $f_{min} \leftarrow f(X_{meilleur})$
- Initialiser la température t
- Tant que « condition 1 » est vraie, faire:
  - Tant que « condition 2 » est vraie, faire :
    - choisir X' dans V(X) et calculer $\Delta f \leftarrow f(X') - f(X)$
    - si $\Delta f < 0$ alors
      - $X \leftarrow X'$
      - si $f(X) < f_{min}$ alors
        - » $f_{min} \leftarrow f(X)$
        - » $X_{meilleur} \leftarrow X$
    - sinon
      - tirer p dans [0,1] avec une distribution uniforme
      - si $p <= e^{-\Delta f / t}$ faire $X \leftarrow X'$
  - $t \leftarrow g(t)$ (g est décroissante)
- Retourner $X_{meilleur}$

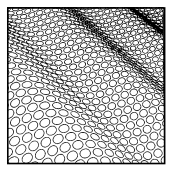
STEREOVISION DES TEXTURES

Shape from texture

- Filtrer les effets dus à la géométrie de l'éclairage pour retrouver la « texture intrinsèque », c'est-à-dire l'albédo ρ(M) de la surface

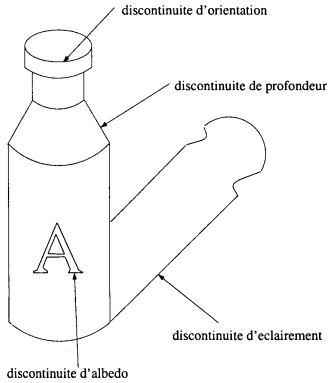
$$L_M(\theta_e, \phi_e) = \frac{I}{\pi}\,\rho(M)\cos(\theta_s(M))$$

- Calculer les variations géométriques de texture pour inférer la forme de la surface: « shape from texture »



STEREOVISION DES CONTOURS

ORIGINE DES VARIATIONS DE LUMINANCE



STEREO DES CONTOURS MODELISATION

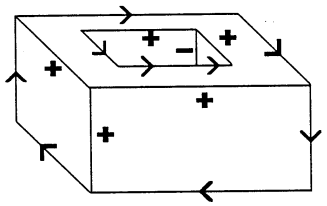
Modélisation

- On suppose que les objets sont polyédriques
- Les parties des surfaces de ces objets sont des facettes planes
- Les frontières des facettes planes sont des arêtes (segments de droite 3D)
- Les images (par projection centrale) des facettes planes sont des surfaces dont les frontières sont des polygônes (suite de segments de droite).

CLASSIFICATION DES ARETES

4 types d 'arêtes

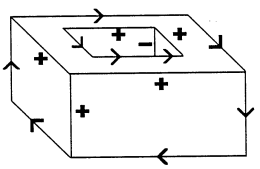
- **les arêtes saillantes (+)**
- **les arêtes rentrantes (-)**
- **les arêtes frontalières (→ et ←)**



CLASSIFICATION DES JONCTIONS

4 types de jonctions autour des images de sommets trièdre en supposant un point de vue non accidentel

- **les jonctions à 2 branches**
  - **L**
- **les jonctions à 3 branches**
  - **fourches**
  - **flèches**
  - **T**



Modélisation

- On suppose que les objets sont polyédriques
- Les frontières des facettes planes sont des arêtes (segments de droite 3D)
  - chaque arête appartient à une et à une seule des 4 classes suivantes :
    - saillante
    - rentrante
    - frontalière droite
    - frontalière gauche
- Les images (par projection centrale) des facettes planes sont des surfaces dont les frontières sont des polygônes (suite de segments de droite).
  - chaque segment-frontière appartient à l'une des 11 classes suivantes:
    - segment gauche ou droit d'une jonction en L
    - segment gauche ou intermédiaire ou droit d'une jonction en flèche
    - segment gauche ou intermédiaire ou droit d'une jonction en T
    - segment gauche ou intermédiaire ou droit d'une jonction en fourche

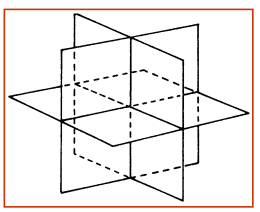
STEREO DES CONTOURS
MISE EN RELATIONS

Mise en relations

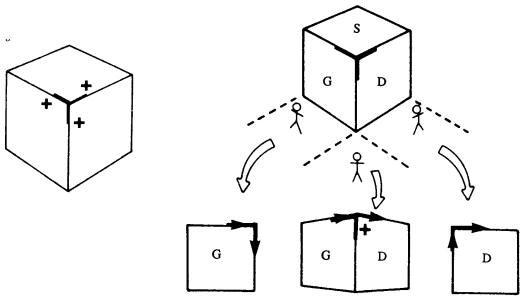
- Quelles relation existe-t-il entre les 4 classes d' arêtes et les 11 classes de segments (images d'arêtes) ?
- La réponse est donnée par le « dictionnaire d'Huffman-Clowes »

INTERPRETATION DES ARETES D 'UNE JONCTION

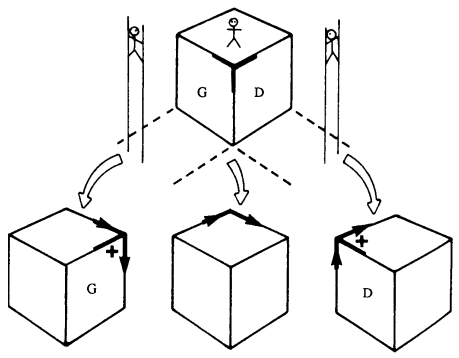
- **On examine les sommets trièdre depuis toutes les directions possibles, afin d 'établir la liste complète des configurations des classes d'arêtes autour des jonctions qui soient géométriquement possibles.**
- **On remarque que 3 faces définissant un trièdre divisent l 'espace en 8 octants**



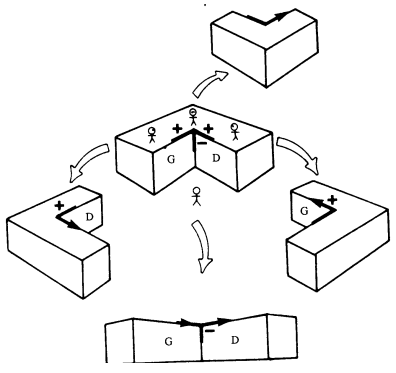
1 OCTANT REMPLI



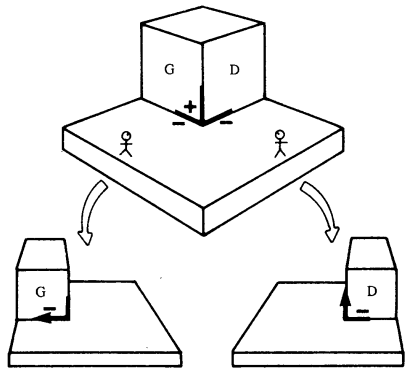
1 OCTANT REMPLI



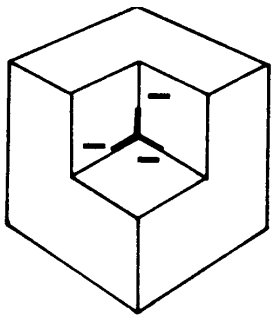
3 OCTANTS REMPLIS



5 OCTANTS REMPLIS

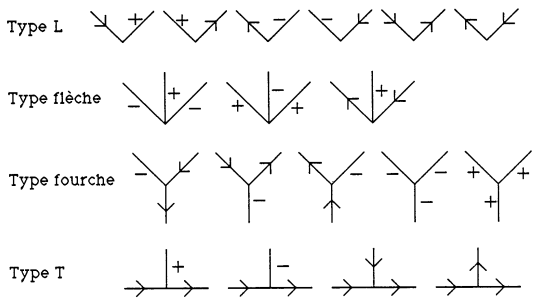


7 OCTANTS REMPLIS

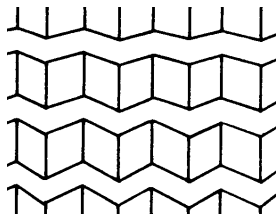


INTERPRETATION DES ARETES

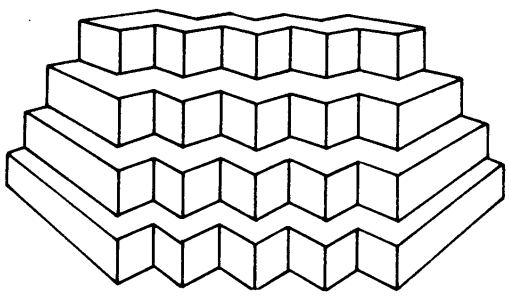
DICTIONNAIRE D'HUFFMAN-CLOWES



CONTRAINTE DU BORD

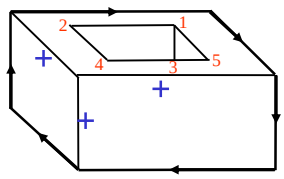


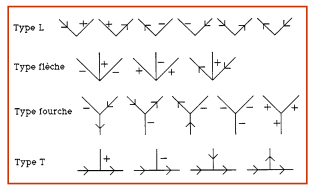
CONTRAINTE DU BORD



STEREO DES CONTOURS
RESOLUTION

EXECUTION D'UNE RECHERCHE

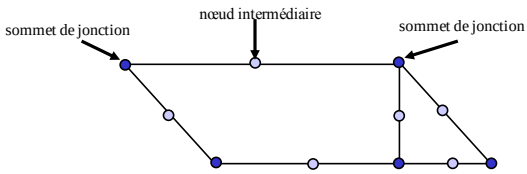




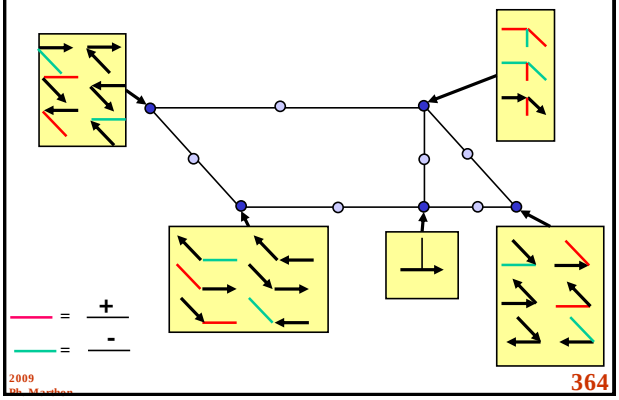
Résolution systémique

- L'approche systémique consiste à construire un système (graphe + règles) qui résout le problème.
- Pour construire le graphe du système on recopie le dessin et l'on insère sur chaque segment un nouveau nœud.
- On détermine le type de chaque jonction (flèche, fourche, T ou L)
- Grâce au dictionnaire d'Huffman-Clowes, on détermine les différentes interprétations 3D possibles des arêtes autour des jonctions. L'état initial de chaque composant (jonction) du système est l'ensemble de ces interprétations a priori possible.
- Le système évolue en propageant ces contraintes d'interprétation. S'il n'y a qu'une seule interprétation possible, le système converge vers cette interprétation

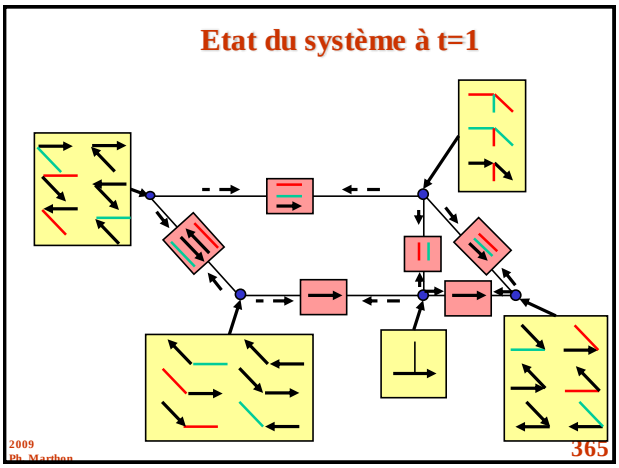
Architecture du système



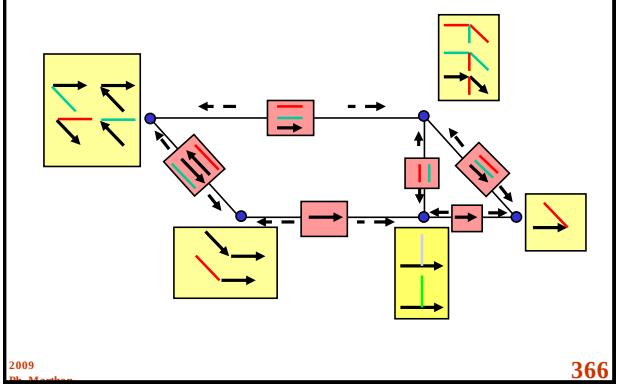
Initialisation du système (t=0)



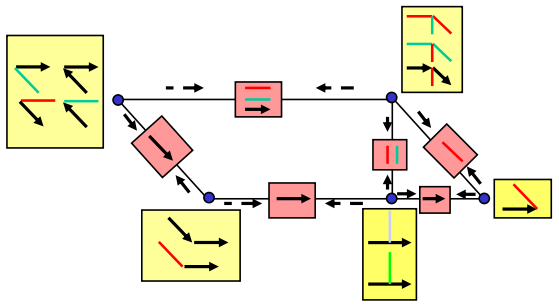
Etat du système à t=1



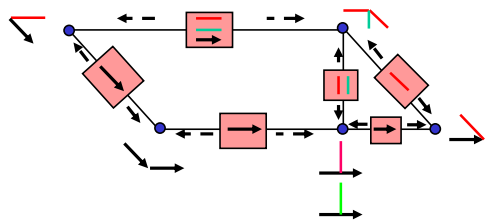
Etat du système à t=2



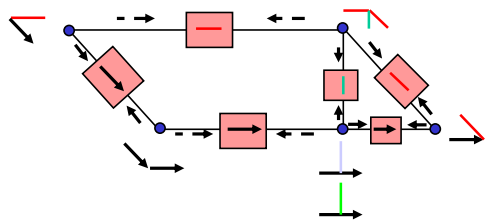
Etat du système à t=3



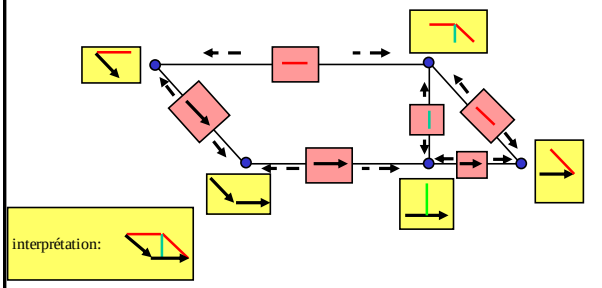
Etat du système à t=4



Etat du système à t=5



Etat du système à t=6 : convergence



Exercice : écrire les règles d'évolution de ce système !

HC-INTERPRETATION

- Une figure composée de segments de droite est dite **HC-INTERPRETABLE** si chacun de ses segments peut être classé en + , - , > ou < , de telle façon que toutes les jonctions ainsi interprétées appartiennent au dictionnaire de Huffman-Clowes

- L'image d'un objet polyédrique dont chaque sommet est à l'intersection de 3 plans est **HC-interprétable**
- La réciproque est fausse : une figure HC-interprétable peut représenter un objet géométriquement impossible

- L'image d'un objet polyédrique dont un sommet au moins est à l'intersection de plus de 3 plans n'est pas forcément HC-interprétable

OBJET IMPOSSIBLE en tant que polyèdre

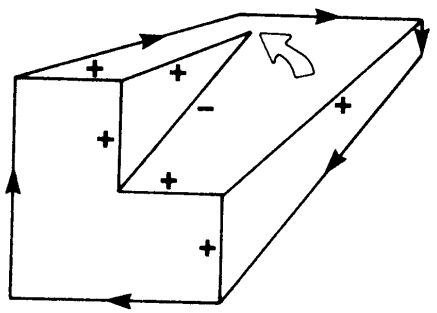
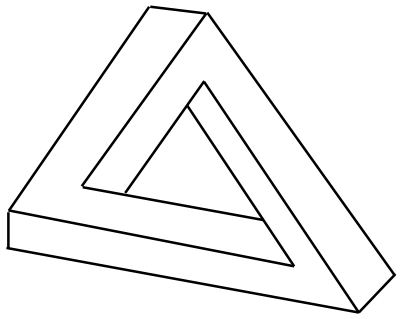
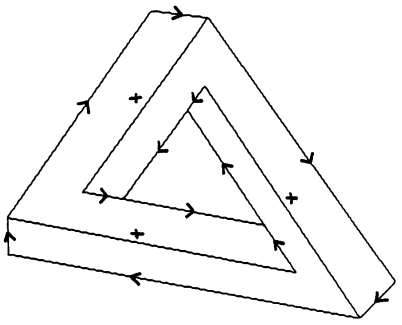


Figure NON HC_interprétable

TRIANGLE IMPOSSIBLE

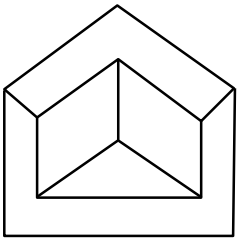


TRIANGLE IMPOSSIBLE

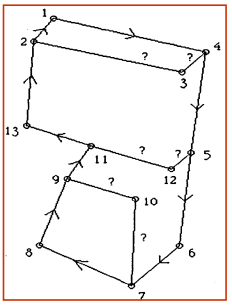


EXERCICE

- Cette figure est-elle HC-interprétable ?
- Représente-elle un objet géométriquement réalisable ?

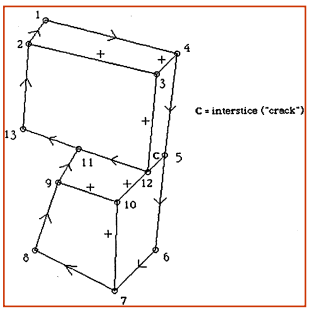


S.I.R.O.P.

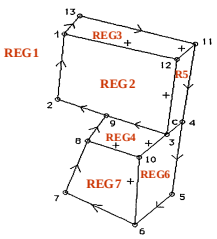


Objet impossible ! Donc il manque des arêtes

S.I.R.O.P.



S.I.R.O.P.



est la suivante:

LISTEREGIONS = \REG1 REG2 REG3 REG4 REG5 REG6 REG7)

REG1 = \SEG17 SEG18 SEG22 SEG23 SEG4 SEG3 SEG35 SEG31 SEG12 SEG13)

REG2 = \SEG1 SEG5 SEG25 SEG9 SEG32)

SEG3 = \SEG2 SEG33 SEG29 SEG36)

SEG4 = \SEG6 SEG24 SEG21 SEG28)

SEG5 = \SEG7 SEG10 SEG30 SEG34)

SEG6 = \SEG8 SEG26 SEG15 SEG14 SEG11)

SEG7 = \SEG16 SEG27 SEG20 SEG19)

SEG1 = \SOM1 SOM2) ; SEG2 = \SOM1 SOM12) ; SEG3 = \SOM1 SOM13)

SEG4 = \SOM2 SOM1) ; SEG5 = \SOM2 SOM9)

SEG6 = \SOM3 SOM9) ; SEG7 = \SOM3 SOM4) ; SEG8 = \SOM3 SOM10);SEG9 = \SOM3 SOM12)

SEG10 = \SOM4 SOM11) ; SEG11 = \SOM4 SOM5) ; SEG12 = \SOM4 SOM3)

SEG13 = \SOM5 SOM6) ; SEG14 = \SOM5 SOM4)

SEG15 = \SOM6 SOM5) ; SEG16 = \SOM6 SOM10) ; SEG17 = \SOM6 SOM7)

SEG18 = \SOM7 SOM8) ; SEG19 = \SOM7 SOM6)

SEG20 = \SOM8 SOM7) ; SEG21 = \SOM8 SOM10) ; SEG22 = \SOM8 SOM9)

SEG23 = \SOM9 SOM2) ; SEG24 = \SOM9 SOM8) ; SEG25 = \SOM9 SOM3)

SEG26 = \SOM10 SOM6) ; SEG27 = \SOM10 SOM8) ; SEG28 = \SOM10 SOM3)

SEG29 = \SOM11 SOM13) ; SEG30 = \SOM11 SOM12) ; SEG31 = \SOM11 SOM4)

SEG32 = \SOM12 SOM1) ; SEG33 = \SOM12 SOM11) ; SEG34 = \SOM12 SOM3)

SEG35 = \SOM13 SOM11) ; SEG36 = \SOM13 SOM1)

LISTEINT3D = (< + > > < > C + + < C > > < < + > > < < + > > < < + + + < + > + + + > <)

SOM1 = \(15 5) ; Type = \FLECHE SEG1 SEG2 SEG3)

SOM2 = \(36,4) ; Type = \L SEG4 SEG5)

SOM3 = \(51 40) ; Type = \K SEG6 SEG7 SEG8 SEG9)

SOM4 = \(48 45) ; Type = \TE SEG31 SEG11 SEG12)

SOM5 = \(74 43) ; Type = \L SEG13 SEG14)

SOM6 = \(86 30) ; Type = \FLECHE SEG15 SEG16 SEG17)

SOM7 = \(82 12) ; Type = \L SEG18 SEG19)

SOM8 = \(52 15) ; Type = \FLECHE SEG20 SEG21 SEG22)

SOM9 = \(45 19) ; Type = \TE SEG5 SEG24 SEG25)

SOM10 = \(60 31) ; Type = \FOURCHE SEG26 SEG27 SEG28)

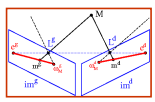
SOM11 = \(16 48) ; Type = \FLECHE SEG29 SEG30 SEG31)

SOM12 = \(22 43) ; Type = \FOURCHE SEG32 SEG33 SEG34)

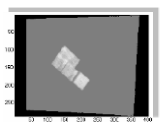
SOM13 = \(6 12) ; Type = \L SEG35 SEG36).

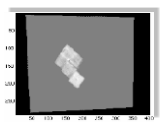
STEREOVISION BINOCULAIRE

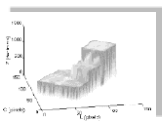
Stéréovision binoculaire passive

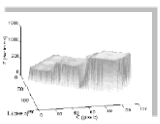


$$\tilde{m}^{d^T} \tilde{F} \tilde{m}^g = 0$$



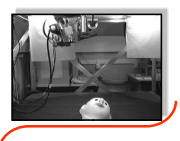


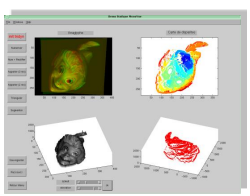


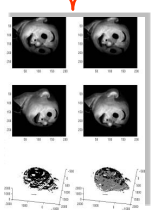


Stéréovision binoculaire incrémentale

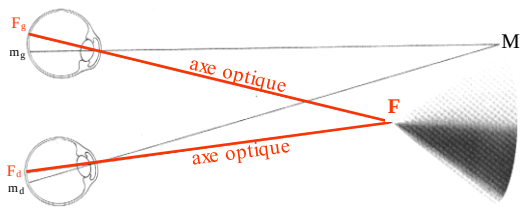
Banc NOISIV (LIMA-IRIT)





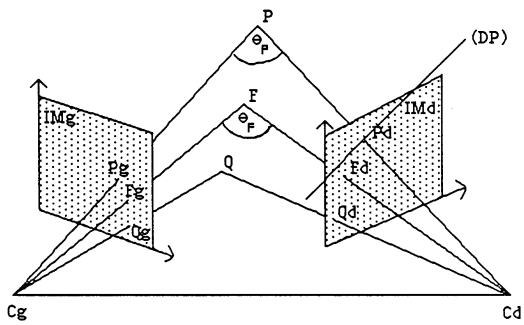


SYSTEME VISUEL HUMAIN
DISPARITE BINOCULAIRE



$$\delta = \overline{F_g m_g} - \overline{F_d m_d}$$

GEOMETRIE BINOCULAIRE



ALGORITHME DE RECONSTRUCTION

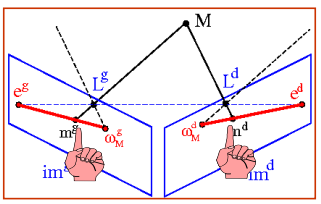
- Déterminer les matrices de perspective $\tilde{C}^g$ et $\tilde{C}^d$
- Pour chaque pixel (de centre m^g) de l 'image gauche
  - déterminer son pixel homologue (de centre m^d) s 'il existe
  - calculer les deux rayons de projection:

$$\boxed{\begin{aligned} M &= \left(C^g\right)^{-1}\left(-\tilde{c}_4^g + \lambda\,\tilde{m}^g\right) \\ M &= \left(C^d\right)^{-1}\left(-\tilde{c}_4^d + \lambda\,\tilde{m}^d\right) \end{aligned}}$$

  - trianguler pour trouver M (calculer l 'intersection des deux droites)

MISE EN CORRESPONDANCE

POINTS HOMOLOGUES



$$\tilde{m}^{d^T} \tilde{F} \tilde{m}^g = 0$$

MISE EN CORRESPONDANCE

- 3 paramètres doivent être définis :
  - –les **primitives**
  - –les **critères d'appariement**
  - –la **stratégie**

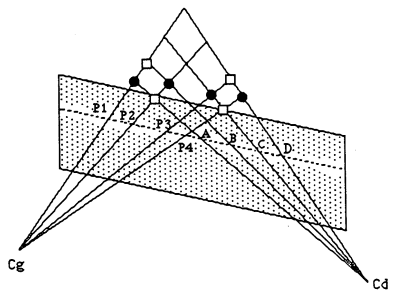
PRIMITIVES

- **Pixels (ou points) et leurs attributs:**
  - **Luminance (ou niveaux de gris)**
  - **Couleur**
  - **Contour**
  - **Sommet**
  - **Texture**
- **Segments de contour (et attributs)**
- **Régions élémentaires (et attributs)**
- **Régions (et attributs)**

CRITERE

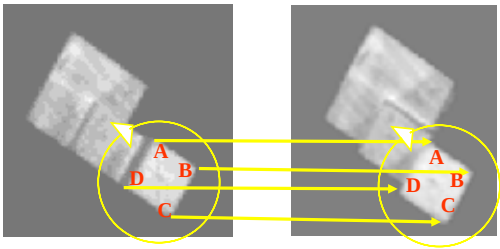
- Besoin de 2 critères, l'un **local** l'autre **global**
  - **critère local** : mesure de **ressemblance** des attributs des primitives, corrélation
    - ex : si la surface est **mate**, les luminances (pas les éclairements !) des pixels homologues sont les mêmes
    - dans le cas général, ils sont « voisins »
    - pour des segments de contour homologues, les signes des contrastes sont les mêmes, les orientations « voisines » (écart < 60°) en général
  - **critère global** : mesure de **continuité** (gradient de disparité, contrainte d'ordre); en effet, les surfaces des objets étant régulières et les variations brutales de profondeur rares, les disparités varient de manière continue à de rares exceptions près
    - cas de scènes **polyédriques** : l'ordre des des sommets sur un cycle de contour est le même que celui de leurs sommets homologues

AMBIGUITE



L'appariement qui respecte la contrainte d'ordre (ronds noirs) est le plus probable

APPARIEMENT DE SOMMETS

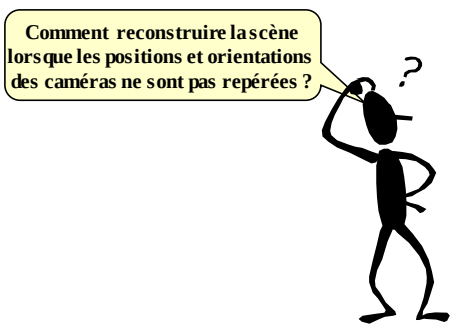


Respect de la contrainte d'ordre

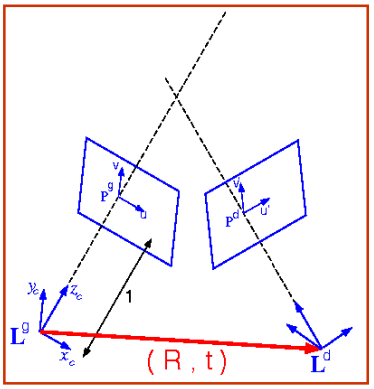
STRATEGIE

- **stratégie MULTIRESOLUTION** : la mise en correspondance des 2 images à un certain niveau de résolution guidera celle du niveau supérieur
- **Hypothèse de MARR** : « Si un segment de passages par zéro apparaît au « même » endroit et selon la « même » orientation dans des images filtrées par des LOG ou DOG de tailles consécutives, alors cet ensemble de segments révèle la présence d'une variation d'intensité dont l'origine est due à un seul phénomène physique (variation de la réflectance, de l'éclairement, de l'orientation d'une surface) »
  - Remarques :
    - les images de basse résolution sont plus faciles à apparier
    - cette stratégie est présente dans le système visuel humain
- **optimisation combinatoire** : on optimise un critère f(x) (local + global), x étant un appariement global qcq.
  - Méthodes : **recuit simulé**, **recherche TABOU**, **algorithmes génétiques**, ...

ALGORITHME DE RECONSTRUCTION



RECONSTRUCTION 3D SANS REPERAGE

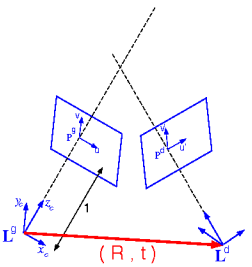


RECONSTRUCTION 3D SANS REPERAGE

- Les paramètres internes étant supposés connus, on repère les points image dans les repères centrés normalisés (Pu'v')

$$u' = \frac{u - p_u}{\alpha_u}$$

$$v' = \frac{v - p_v}{\alpha_v}$$



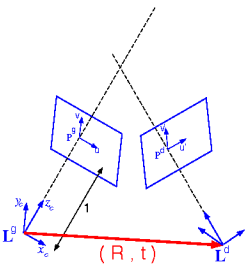
RECONSTRUCTION 3D SANS REPERAGE

- Le repère 3D de référence étant celui de la caméra gauche ($L^g x_c y_c z_c$), les matrices de perspective s'écrivent:

$$\tilde{C}^g = \begin{bmatrix} I_3 & 0 \end{bmatrix}$$

$$\tilde{C}^d = \begin{bmatrix} I_3 & 0 \end{bmatrix} \begin{bmatrix} R^T & -R^T t \\ 0_3^T & 1 \end{bmatrix}$$

$$= \begin{bmatrix} R^T & -R^T t \end{bmatrix}$$

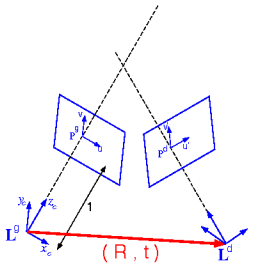


RECONSTRUCTION 3D SANS REPERAGE

- Calcul des épipoles

$$\tilde{e}^g = \tilde{C}^g \tilde{L}^d = \begin{bmatrix} I_3 & 0 \end{bmatrix} \begin{bmatrix} t \\ 1 \end{bmatrix} = t$$

$$\tilde{e}^d = \tilde{C}^d \tilde{L}^g = \begin{bmatrix} R^T & -R^T t \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -R^T t$$



DROITES EPIPOLAIRES

$$\tilde{C}^g = \begin{bmatrix} I_3 & 0 \end{bmatrix} \qquad \tilde{C}^d = \begin{bmatrix} R^T & -R^T t \end{bmatrix}$$

$$\tilde{e}^g = t \qquad \tilde{e}^d = -R^T t$$

$$\Downarrow$$

$$\tilde{e}^d \wedge C^d \left(C^g\right)^{-1} \tilde{m}^g = -R^T t \wedge R^T \tilde{m}^g = R^T \left(-t \wedge \tilde{m}^g\right)$$

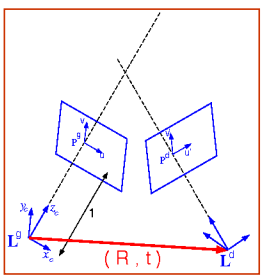
$$\tilde{e}^g \wedge C^g \left(C^d\right)^{-1} \tilde{m}^d = t \wedge R\tilde{m}^d$$

RECONSTRUCTION 3D SANS REPERAGE

- Droites épipolaires

associé à m^g : $R^T\left(t \wedge \tilde{m}^g\right)$

associé à m^d : $t \wedge R\tilde{m}^d$



m^g et m^d homologues $\Longrightarrow$ $\left(\tilde{m}^g \middle| t \wedge R\tilde{m}^d\right) = 0$

MATRICE ESSENTIELLE

Soit T la matrice : $Tx = t \wedge x$

Soit $\tilde{E} = T\,R$ la matrice essentielle $\quad \mathrm{rg}(\tilde{E}) = 2$

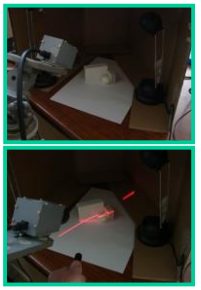
m^g et m^d homologues $\Rightarrow$ $\left(\tilde{\mathrm{m}}^g \middle| \tilde{E}\,\tilde{\mathrm{m}}^d\right) = 0$

ALGORITHME DE LONGUET-HIGGINS

- **1ère étape** : calcul de la **matrice essentielle** en déterminant au moins **8** couples de points homologues
- **2ème étape** : on en déduit la **translation** ***t*** à un facteur d'échelle près
- **3ème étape** : on en déduit la matrice de **rotation** ***R***; le déplacement entre les 2 caméras est maintenant connu
- **4ème étape** :**appariement et triangulation** : pour chaque couple de points homologues, on calcule les coordonnées de M à un facteur d'échelle près

STEREOVISION ACTIVE

Stéréovision active



Projection d'une ombre

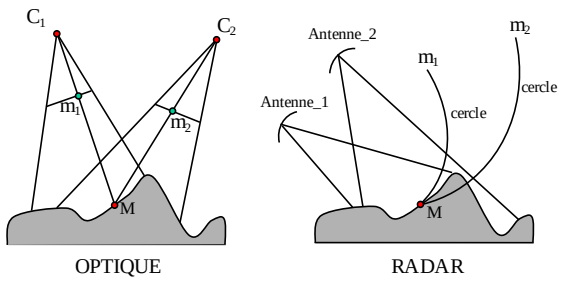
Projection d'un plan laser

STEREOVISION RADAR

2.2. La résolution géométrique : stéréovision radar

1. Stéréovision monoculaire
   - Radarclinométrie (analogue à la photoclinométrie)
2. Stéréovision binoculaire
   1. Radargrammétrie
   2. Interférométrie

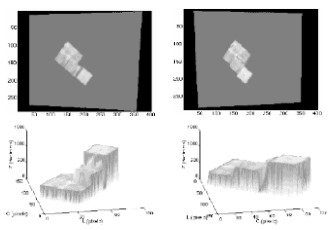
Stéréovision binoculaire vs Radargrammétrie



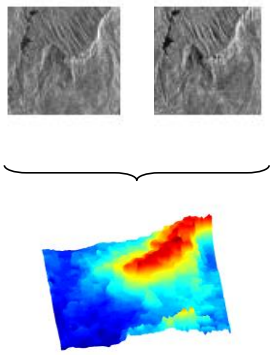
UNE METHODE

Objectif	appariement dense de pixels (après rectification) triangulation
Modélisation	problème discret d'optimisation combinatoire intégrant une mesure de ressemblance locale
Résolution	Recherche Tabou, Algorithmes Génétiques, Recherche Dispersée
Précision	au pixel près ! bonne détection des parties cachées

Stéréo binoculaire optique



Radargrammétrie



2009
Ph. Mordon
409

VISION DU MOUVEMENT

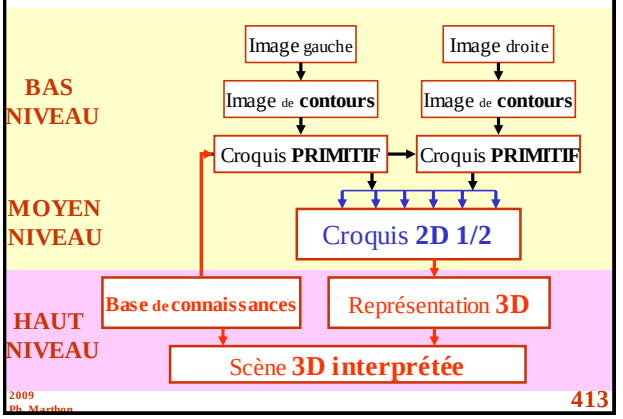
2009
Ph. Mordon
410

SUIVI

2009
Ph. Mordon
411

RECONNAISSANCE ET INTERPRETATION

MODELE DU SYSTEME VISUEL HUMAIN



TRAITEMENTS DE HAUT NIVEAU

- Organiser les surfaces en objets
- Décrire intrinsèquement les objets (description statique)
- Décrire intrinsèquement le mouvement des objets (description dynamique)
- Reconnaître les objets
- Localiser les objets
- Décrire et reconnaître la scène

OBJETS POLYEDRIQUES

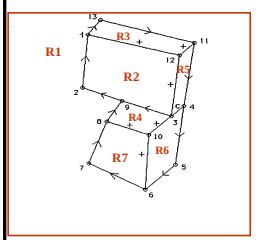
VISIBILITE / OCCULTATION

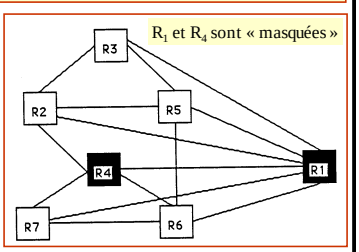
- F : face d'un polyèdre
- R : région polygonale image de F
- $(S_1, S_2, ..., S_n)$: suite de segments de droite formant la frontière de R et ordonnés dans le sens trigonométrique

$$F \text{ entièrement visible} \Leftrightarrow \begin{cases} R \neq \phi \\ \forall i = 1 \ldots n \;\; \text{etiquette}(S_i) \neq '>' \end{cases}$$

VISIBILITE / OCCULTATION

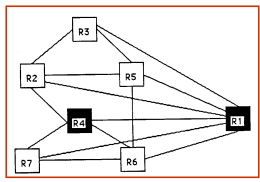
$$F \text{ entièrement visible} \Leftrightarrow \begin{cases} R \neq \phi \\ \forall i = 1 \ldots n \;\; \text{etiquette}(S_i) \neq '>' \end{cases}$$

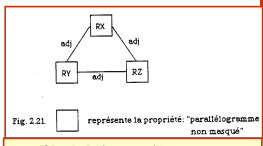




GRAPHE D'ADJACENCE DES REGIONS

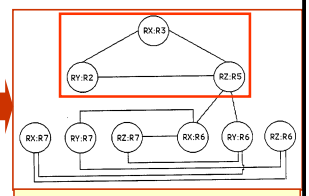
RECONNAISSANCE





Modèle de l 'image d 'un parallélipipède entièrement visible

Présence d 'une clique
➡ présence d 'un parallélipipède



GRAPHE D 'ASSOCIATION

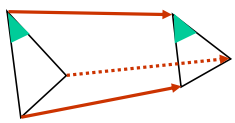
NOMBRE D 'ARETES

F : face d 'un polyèdre
R : région polygonale image de F
NARE = nombre de ses arêtes
NSEG = nombre de segments propres $(\text{etiquette}(S_i) \neq '>')$
de la frontière de R

- F entièrement visible ➡ NARE = NSEG
- F partiellement masquée ➡ NARE $\geq$ NSEG

ANGLE CONVEXE/CONCAVE

- L 'image d 'un angle convexe (saillant) est convexe
- L 'image d 'un angle concave (rentrant) est concave



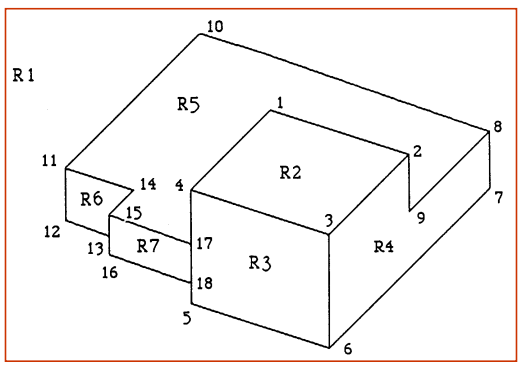
MISE EN CORRESPONDANCE IMAGE 2D - MODELE 3D

- On suppose connues les descriptions (« modèles ») de tous les objets polyédriques susceptibles d'être présents dans la scène observée
- On cherche une description « compatible » avec l'image
- A cet effet, on utilise les contraintes 2D-3D suivantes :
  - interprétation 3D des segments (Huffman-Clowes)
  - nombre de segments « propres » de chaque région
  - type (convexe/concave) de chaque angle visible

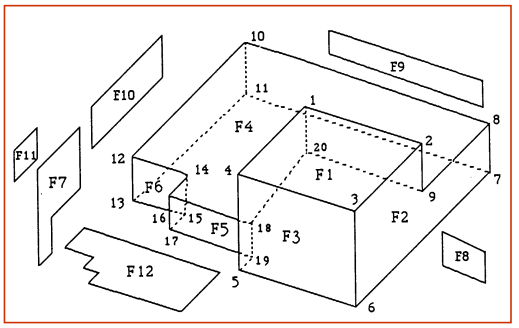
ALGORITHME

- 1ère étape : association régions-faces
  - une association n'est possible que si les contraintes 2D-3D sont vérifiées
- 2ème étape : vérification du contexte
  - on vérifie que régions et faces adjacentes sont compatibles
- 3ème étape : association
  sommets de l'image - sommets du modèle

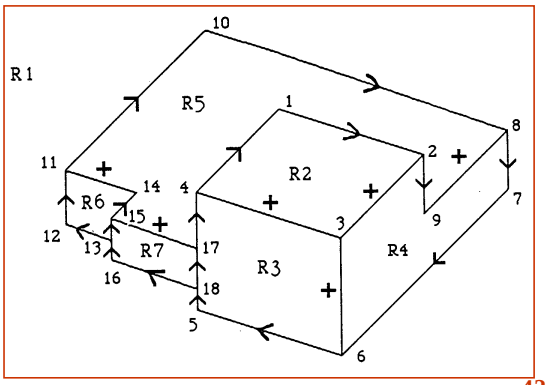
IMAGE SEGMENTEE



MODELE

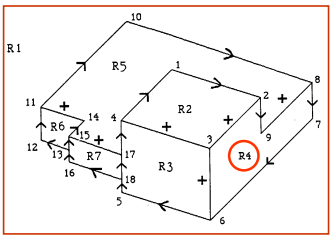


INTERPRETATION DES SEGMENTS



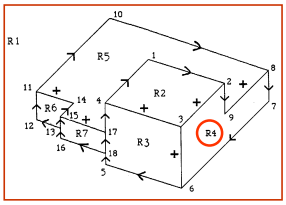
REGION INITIALE

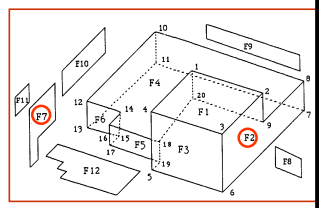
- **Choix d'une région initiale**
  - entièrement visible
  - ayant le plus grand nombre d'angles concaves



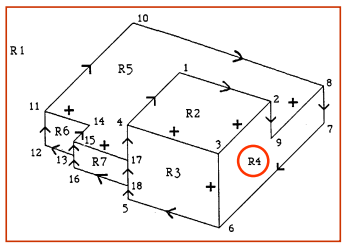
FACES COMPATIBLES

- **Une face compatible F avec la région initiale R** (entièrement visible)
  - possède le même nombre de côtés que R
  - possède le même nombre d'angles concaves que R





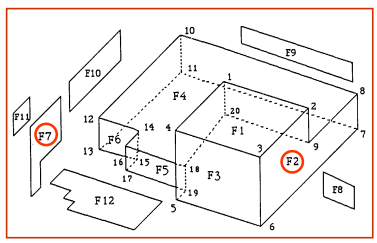
REGIONS ADJACENTES



R4 : (R2 R3 < < R5 <)

R5 : (R4 < < R6 < R7 > >)

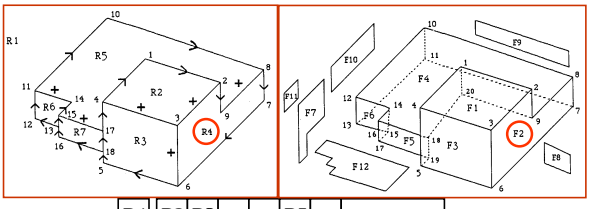
FACES ADJACENTES



F2 : (F1 F3 F12 F9 F4 F8)

F7 : (F1 F8 F4 F5 F12 F3)

TEST DE COMPATIBILITE

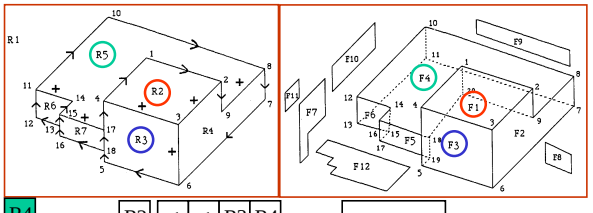


R4	R2	R3	<	<	R5	<	COMPATIBILITE ?
F2	F1	F3	F12	F9	F4	F8	OUI
	F3	F12	F9	F4	F8	F1	NON
	F12	F9	F4	F8	F1	F3	NON
	F9	F4	F8	F1	F3	F12	NON
	F4	F8	F1	F3	F12	F9	NON
	F8	F1	F3	F12	F9	F4	NON

2009
Ph. Marthon

430

TEST DE COMPATIBILITE

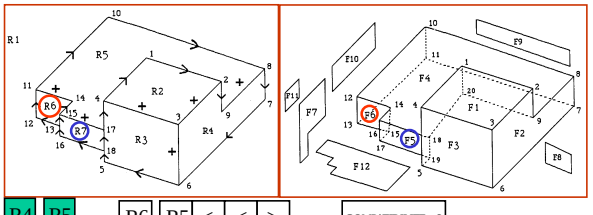


R4	R2	<	<	R3	R4		COMPATIBILITE ?
F2	F1	F8	F7	F3	F2		OUI
	R3	R2	<	<	R4		COMPATIBILITE ?
	F3	F1	F7	F12	F2		OUI
R5	R4	<	<	R6	<	R7	> > COMPATIBILITE ?
F4	F2	F9	F10	F6	F11	F5	F7 F8 OUI

Ph. Marthon

431

TEST DE COMPATIBILITE

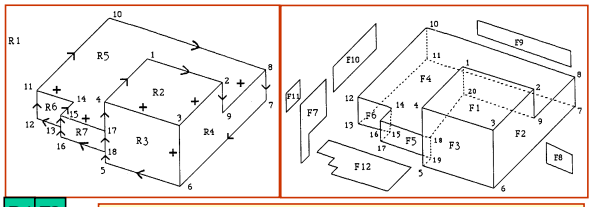


R4	R5	R6	R5	<	<	>	COMPATIBILITE ?
F2	F4	F6	F4	F10	F12	F11	OUI
		R7	R5	<	<	>	COMPATIBILITE ?
		F5	F4	F11	F12	F7	OUI

2009
Ph. Marthon

432

SOMMETS

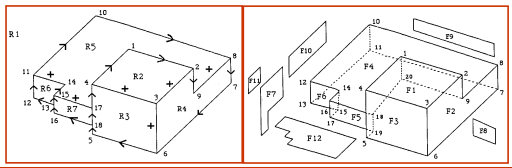


R4	F2
R2	F1
R3	F3
R5	F4
R6	F6
R7	F5

- Trouver région R contenant le sommet-image S considéré
- Trouver les 2 segments SEG1 et SEG2 de R dont S est l'intersection
- Le sommet du modèle SM est à l'intersection de la face F associée à R et des 2 faces adjacentes à F et contenant les 2 arêtes dont les images sont SEG1 et SEG2

$$S9 \leftrightarrow F2 \cap F4 \cap F8 = SM9$$

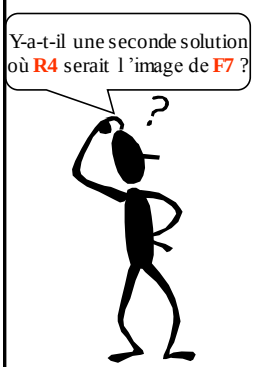
COMPATIBILITE SOMMETS

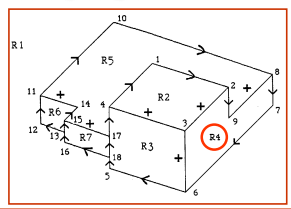


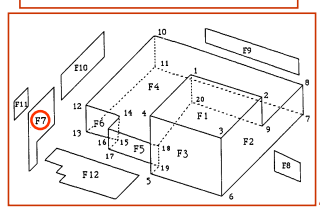
Chaque sommet d'un angle visible et concave doit être associé à un sommet du modèle qui est lui-aussi à l'extrémité d'un angle concave de l'une des 3 faces qui le définissent.

S1	SM1	S4	SM4	S7	SM7	S10	SM10	S14	SM14
S2	SM2	S5	SM5	S8	SM8	S11	SM12	S15	SM16
S3	SM3	S6	SM6	S9	SM9	S12	SM13	S16	SM17

EXERCICE

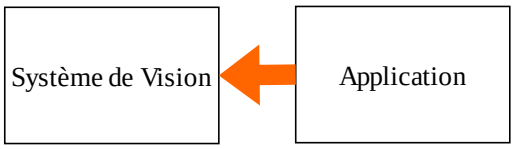






SYSTEME DE VISION

Système et application



Système de vision

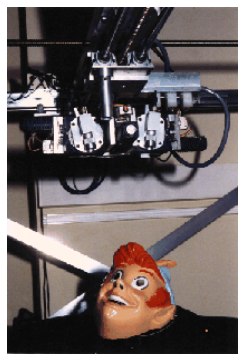
- Un système de vision comprend 4 sous-systèmes:
  - Un système de transport et de visée
  - Un système d'éclairage
  - Un système d'acquisition d'images numériques
  - Un système de traitement de l'information

Système de transport et de visée

2009
Ph. Marthon

439

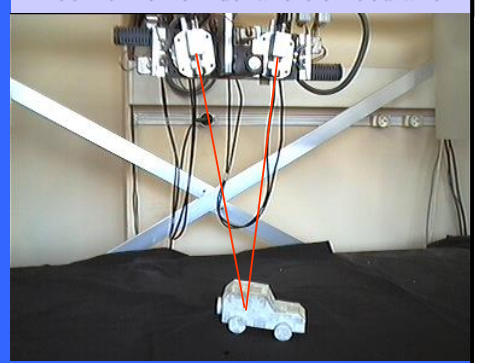
STEREOVISION BINOCULAIRE EN PRATIQUE SUR NOISIV



2009
Ph. Marthon

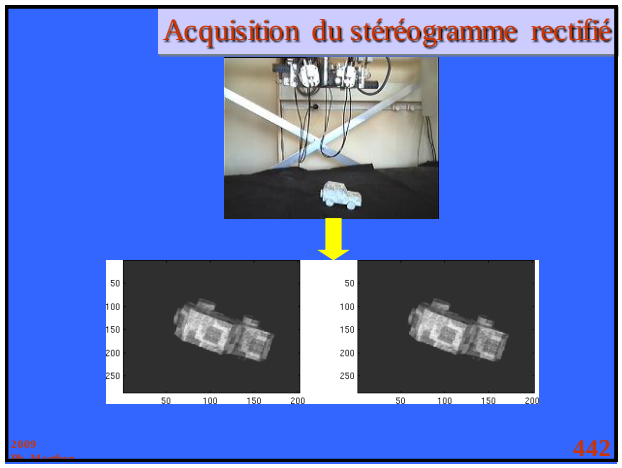
440

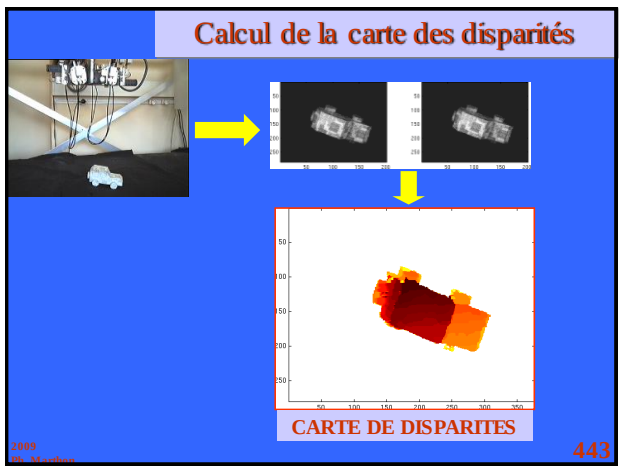
Positionnement de la tête binoculaire

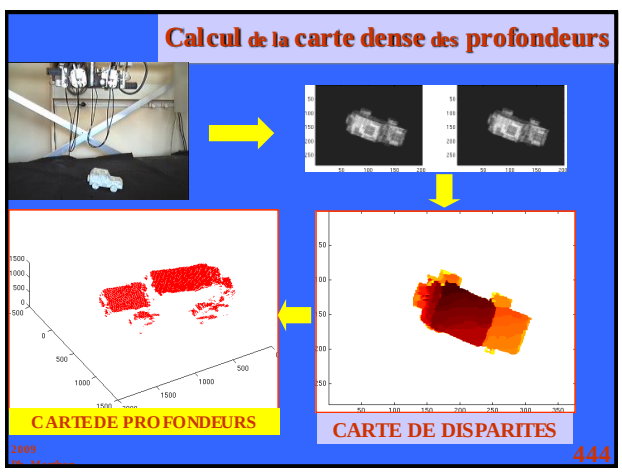


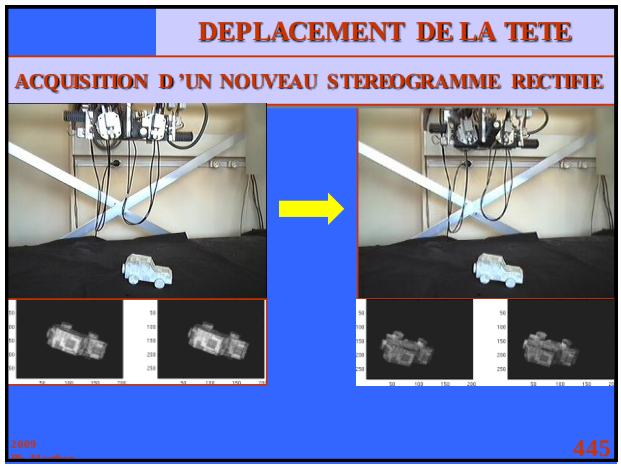
2009
Ph. Marthon

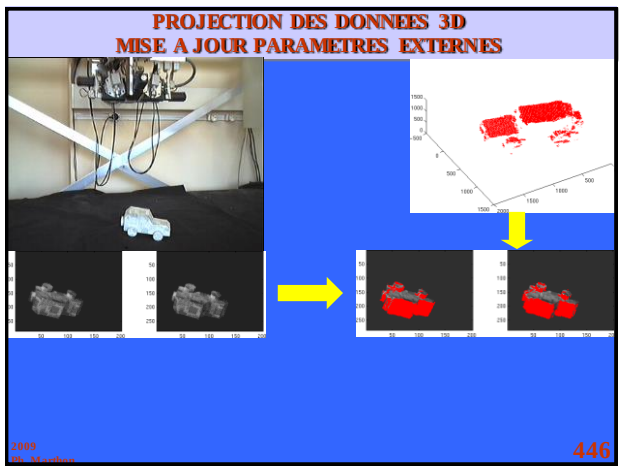
441

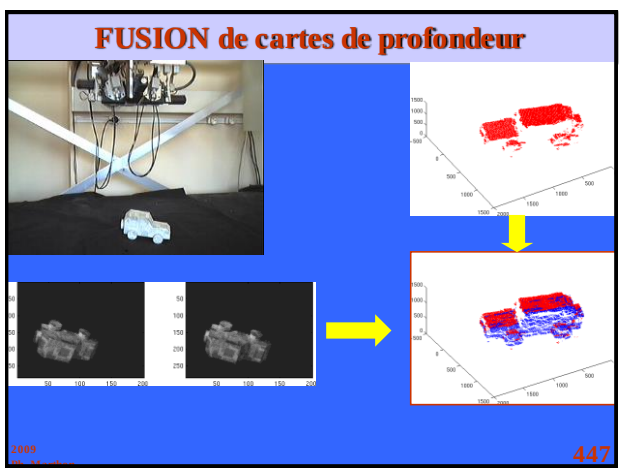

Acquisition du stéréogramme rectifié
2009
442


Calcul de la carte des disparités
CARTE DE DISPARITES
2009
443

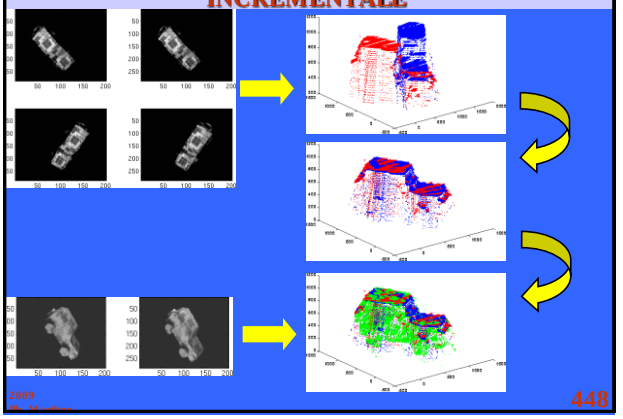

Calcul de la carte dense des profondeurs
CARTE DE PROFONDEURS
CARTE DE DISPARITES
2009
444


DEPLACEMENT DE LA TETE
ACQUISITION D'UN NOUVEAU STEREOGRAMME RECTIFIE
2009
445


PROJECTION DES DONNEES 3D
MISE A JOUR PARAMETRES EXTERNES
2009
446

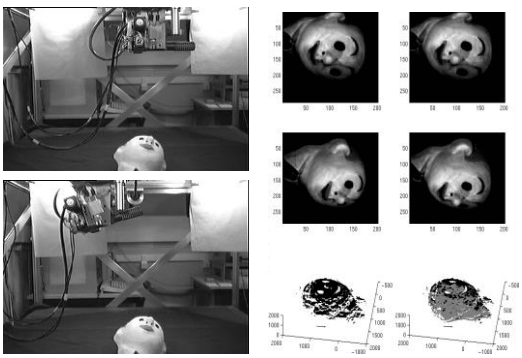

FUSION de cartes de profondeur
2009
447

STEREOVISION BINOCULAIRE INCREMENTALE



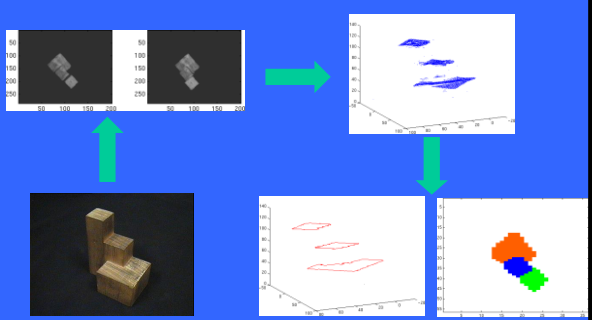
2009
Ph. Marthon
448

STEREOVISION BINOCULAIRE INCREMENTALE



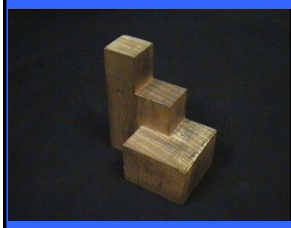
2009
Ph. Marthon
449

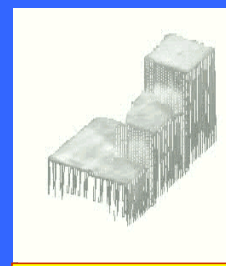
Segmentation 3D d 'une carte de profondeur en facettes planes

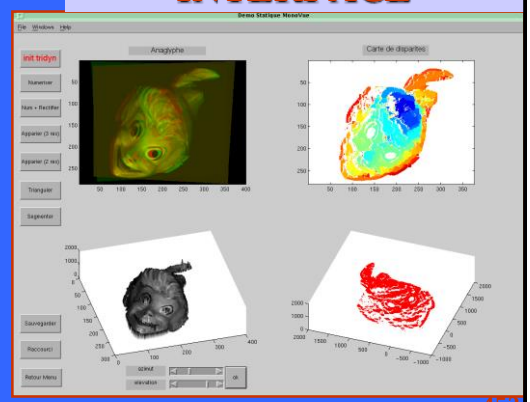


2009
Ph. Marthon
450

RECONSTRUCTION 3D





RECONSTRUCTION 3D

INTERFACE