Représentation et analyse de formes

Yann GOUSSEAU
Télécom Paris - IP Paris

- Quelques expériences de perception des formes
- Représentations globales, moments invariants
- Frontière et rôle de la courbure
- Formes paramétriques : Hough, Ransac

Au prochain cours :

- Points clé et descripteurs locaux
- Détection d'objets par réseaux de neurones

Expériences de perception des formes et principes gestaltistes

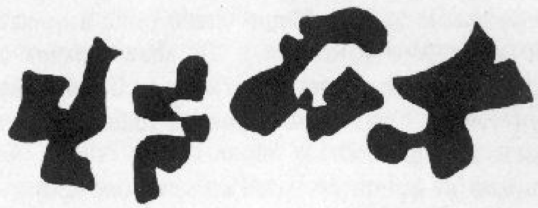


FIG. 1.2. Oggetti visivi sconosciuti, senza significato, ma perfettamente visibili e stabili per forma, colore, grandezza, rapporti spaziali.

G. Kanizsa, grammatica del vedere

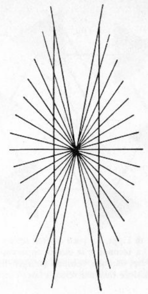


FIG. 4.29. L'illusione di Hering. Anche in questo caso l'incurvatura delle due parallele sarebbe una conseguenza della tendenza degli angoli verso l'angolo retto.

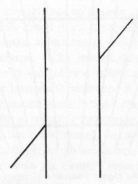


FIG. 4.30. L'illusione di Poggendorff. Ancora un presunto effetto secondario della ortogonalizzazione degli angoli: perdita della collinearità.

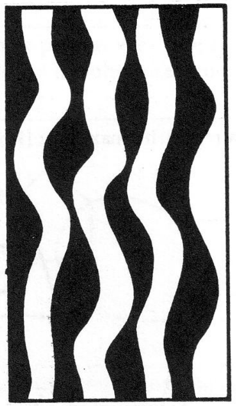


Fig. 1.45 : Action de la constante de largeur [Morinaga, 1941].

Expérience de **discrimination fond-forme**.

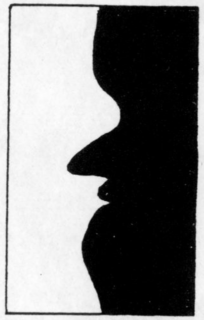
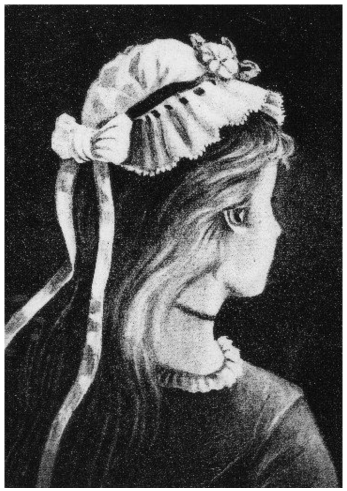
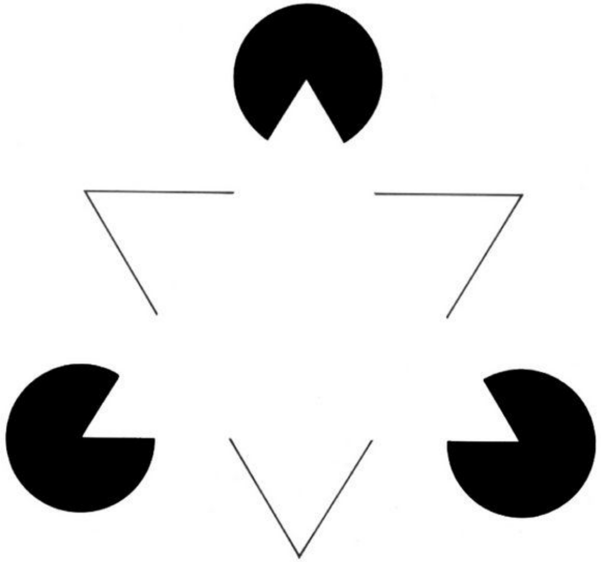


FIG. 5.21. Un'altra configurazione bistabile: il profilo nero e il profilo bianco si alternano nel ruolo di figura [da Rubin 1921].

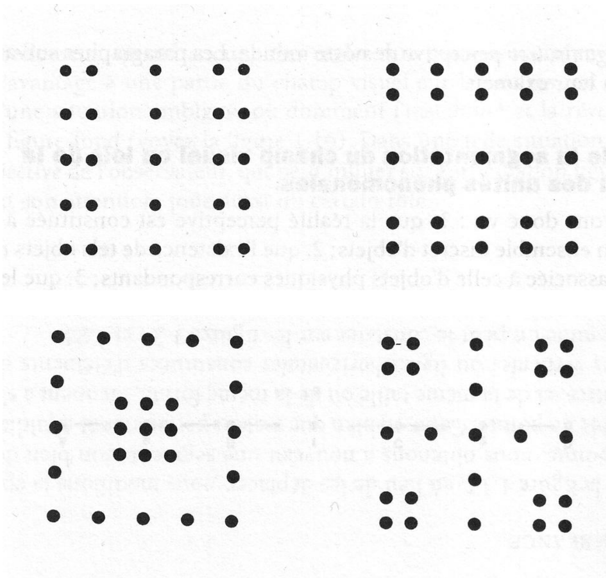




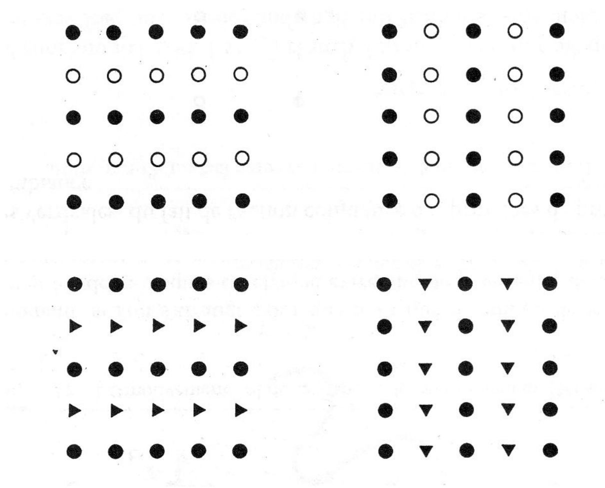
Lois de groupement

- Ecole gestaltiste (Wertheimer 1920, Metzger 1975, Kanizsa[1])
- Comment passe-t-on d'une "innombrable quantité d'éléments singuliers isolés les uns des autres" à la formation des objets ?
- Lois de groupement, de constitution des objets visuels, dont :
  - proximité, similarité
  - bonne continuation
  - fermeture, convexité, symétrie
  - constance de largeur
  - complétion amodale
  - Etc.

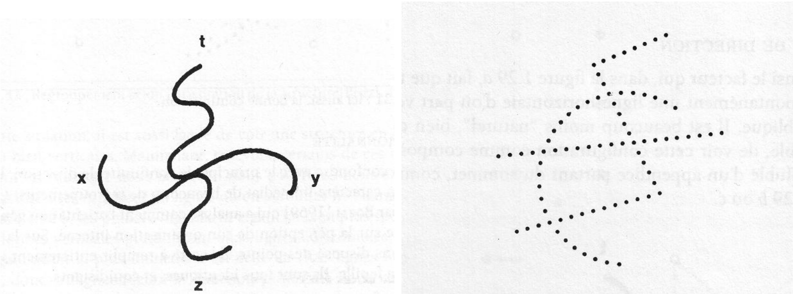
1. Pour une introduction, voir *La grammaire du voir*, Kanizsa, ed. Diderot, 1999



Regroupement selon la proximité.



Regroupement selon la similarité.

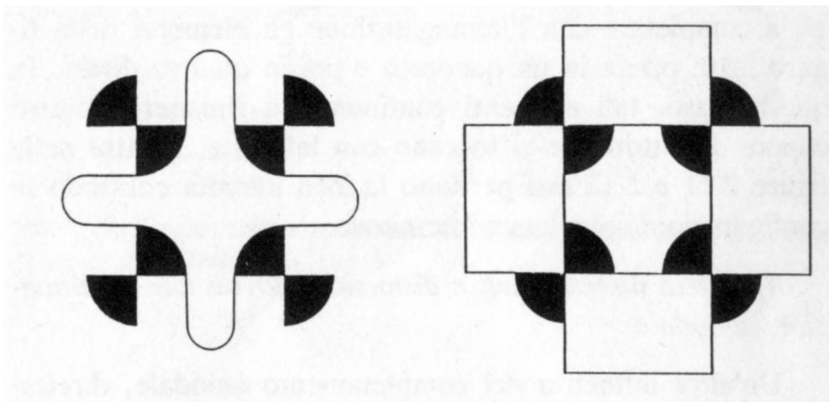


Principe de bonne continuation.

Tendance à la symétrie.

Fig. 1.45 : Action de la constante de largeur [Morinaga, 1941].

Constance de largeur - discrimination fond-forme.



Complétion amodale

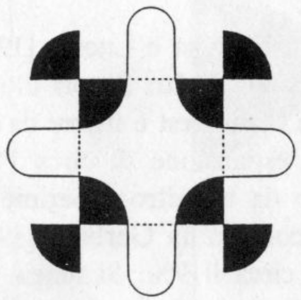


FIG. 2.13. Così si completa la figura 2.11.

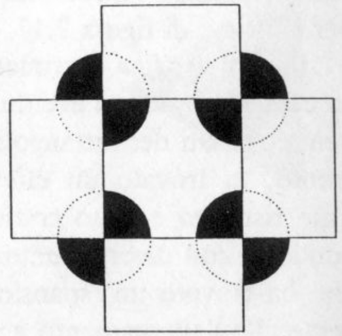
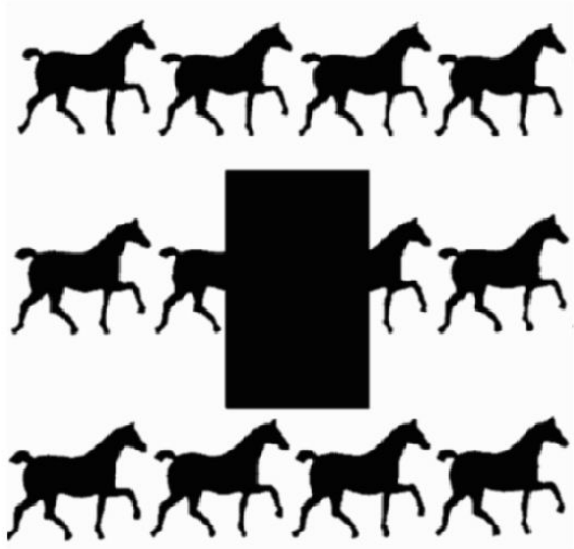


FIG. 2.14. Così si completa la figura 2.12.

Complétion amodale
Les courbes sont interpolées **régulièrement** entre les **jonctions en T**



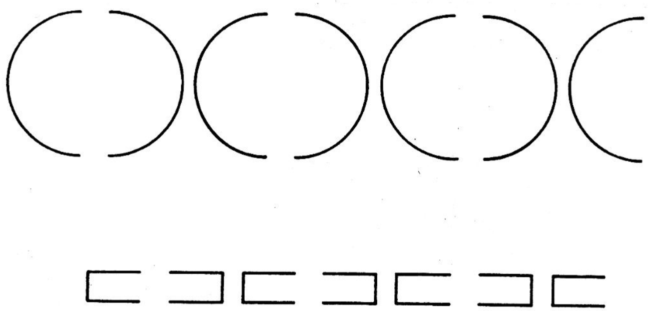
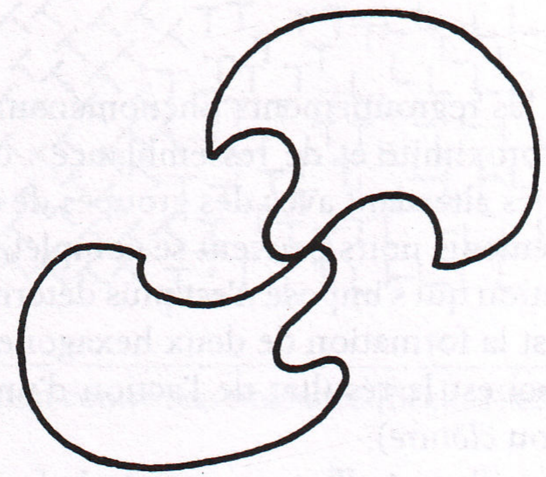
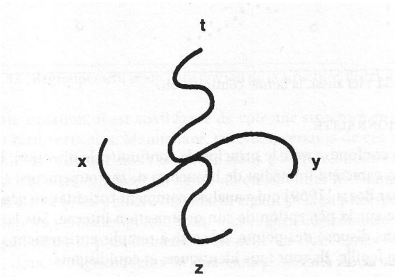


Fig. 1.38 : Fermeture contre proximité.





Principe de bonne continuation seul : l'interprétation est différente.

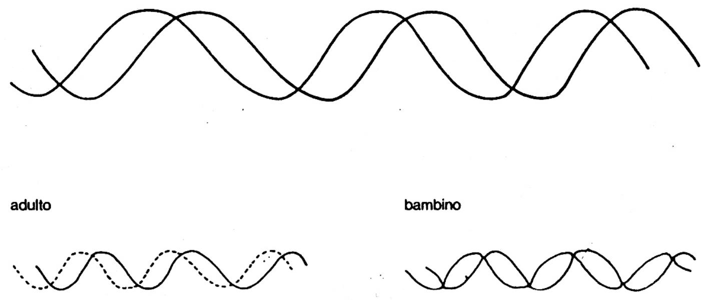


Fig. 1.65 : Regroupements différents chez l'adulte et chez l'enfant.

Descripteurs de formes

Descripteurs globaux

- Pour une forme $A \subset \mathbb{R}^2$, on s'intéresse à des descripteurs numériques :

$$\Phi : \mathcal{P}(\mathbb{R}^2) \to \mathbb{R}^p$$

  invariants à certaines transformations géométriques
  $\to$ prise en compte des variabilités de pose lors de l'acquisition de l'image
- Φ est invariant à T si

$$\Phi(T.A) = \Phi(A)$$

- On considère typiquement : translations, rotation, zoom, transformations affines.
- Une transformation affine est une application du plan dans lui-même définit par

$$X \to M.X + P$$

  avec M une matrice inversible et P un vecteur
  C'est une bonne approximation d'un changement de point de vue léger sur une scène plane.

Descripteurs globaux

Exemples de descripteurs globaux élémentaires :

- Périmètre $P(A)$, surface $S(A)$, rapport isopérimétrique
$$R = \frac{4\pi S(A)}{P(A)^2}$$
- Diamètre
$$\sup_{x,y \in A} d(x,y)$$
- Boîtes englobantes (axes fixes ou adaptatifs)
- Symétrie
$$\frac{S(A)}{S(A \cup \check{A})}, \text{ avec } \check{A} = \{-x : x \in A\}$$

Moments

$$m_{p,q} = \int\int_A (x - x_c)^p (y - y_c)^q dxdy,$$

avec (x_c, y_c) coordonnées du centre d'inertie de la forme :

$$x_c = \frac{1}{S(A)} \int\int_A xdxdy$$

$$y_c = \frac{1}{S(A)} \int\int_A ydxdy$$

En discret $\rightarrow$ somme sur les pixels

Matrice d'inertie

Utilisation la plus standard :

$$\Sigma = \begin{pmatrix} m_{2,0} & m_{1,1} \\ m_{1,1} & m_{0,2} \end{pmatrix}$$

- Valeurs propres $\lambda_1 \geq \lambda_2$: information sur la taille et l'allongement
- Angle du vecteur propre associé à λ_1 : orientation
- Représentation par une ellipse

Pour être plus discriminant : augmentation de $p, q \rightarrow$ problème de robustesse

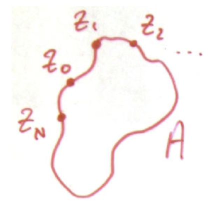
Moments invariants

- Normalisation pour une invariance au zoom (changement d'échelle) :
$$n_{p,q} = \frac{m_{p,q}}{m_{0,0}^{\alpha}}, \text{ avec } \alpha = \frac{1}{2}(p+q) + 1.$$
En effet : si $A \rightarrow kA$ alors $m_{p,q} \rightarrow m_{p,q} k^{p+q+2}$.
- Normalisation supplémentaire pour les rotations :
$$h_1 = n_{2,0} + n_{0,2} \text{ et } h_2 = (n_{2,0} - n_{0,2})^2 + 4n_{1,1}^2,$$
ou autre combinaison trace-déterminant de
$$S = \begin{pmatrix} n_{2,0} & n_{1,1} \\ n_{1,1} & n_{0,2} \end{pmatrix}$$
+ compliqué pour les ordres supérieurs
$$h_3 = (n_{3,0} - n_{0,3})^2 + (n_{0,3} - 3n_{2,1})^2,$$
etc.

Moments invariants (suite)

- Alternative pour le second ordre :
  Soient $\alpha_1 > \alpha_2$ les v.p. de S
  Elongation $e = \alpha_2/\alpha_1$:
  Compacité $\kappa = \frac{1}{4\pi\sqrt{\alpha_1\alpha_2}}$
- Propriétés :
  - $0 \leq e, \kappa \leq 1$
  - e invariant aux rotations
  - κ invariant aux transformations affines.

Descripteurs de Fourier



- La forme A est codée par son contour $Z_v = x_v + iy_v$, pour $v = 0, \ldots, N-1$ dont la transformée de Fourier discrète (TFD) est

$$\hat{Z}_k = \frac{1}{N}\sum_{v=0}^{N-1} Z_v \exp(-2i\pi kv/N),$$

  pour $k = -\frac{N}{2}+1, \ldots, \frac{N}{2}$
- On utilise comme descripteurs : $(\hat{Z}_{-N/2+1}, \ldots, \hat{Z}_{N/2})$, ou un sous-ensemble de ces valeurs.
- Remarque : il existe de nombreux autres descripteurs de Fourier reposant sur des principes similaires

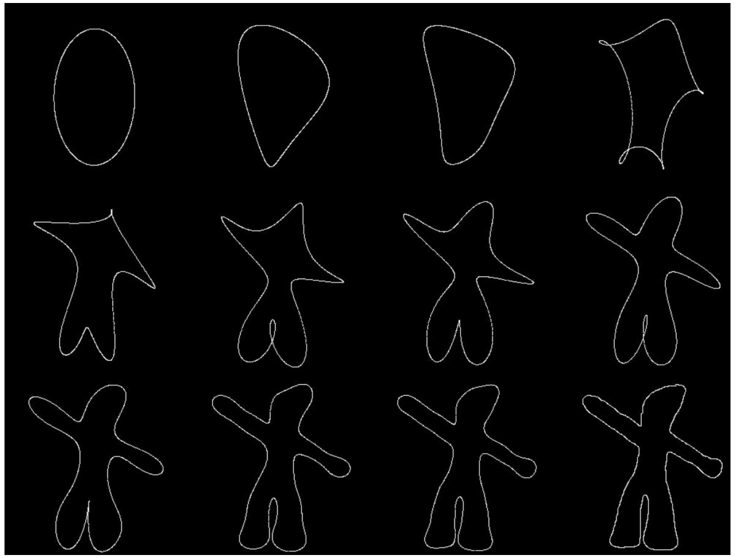
$$\hat{Z}_k = \frac{1}{N}\sum_{v=0} Z_v \exp(-2i\pi kv/N),$$

Propriétés

- $\hat{Z}_0$ est le barycentre de A
- pour $k \neq 0$, $\hat{Z}_k$ est invariant par translation
  en effet :
  si $A \to A + C$, alors pour tout v, $Z_v \to Z_v + C$, donc

$$\hat{Z}_k \to \hat{Z}_k + \sum\sum_{v=0} C\exp(-2i\pi kv/N) = \hat{Z}_k$$

- les $|\hat{Z}_k|$ sont invariants par rotation
- on obtient une invariance par zoom en divisant les descripteurs par $|\hat{Z}_1|$



Haut : 1,2,3,4 coeff. Milieu : 5,6,7,8, coeff. Bas : 10, 20, 30, 1257 coeff.
Extrait de Ruye Wang, *Introduction to orthogonal transforms*, 2010

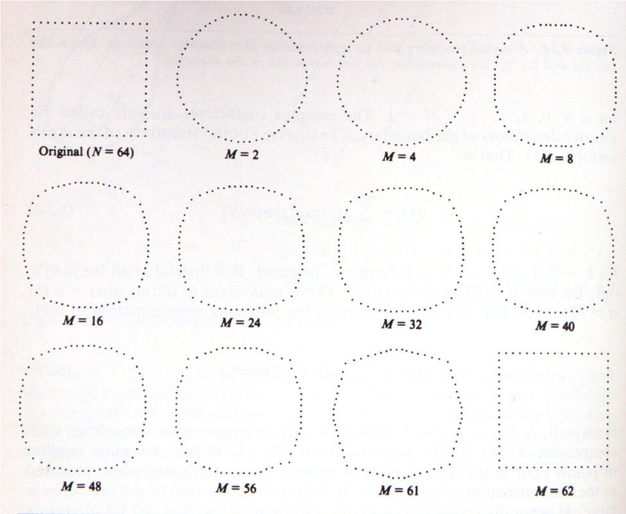
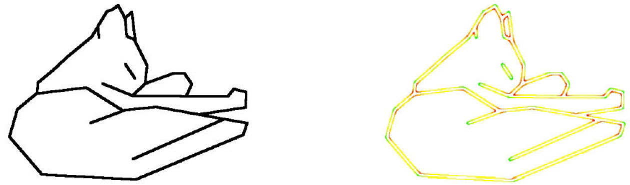


Figure 8.15 *Examples of reconstructions from Fourier descriptors for various values of M.*

Extrait de Gonzalez et Woods *Digital Image Processing*
... les points de forte courbure sont difficiles à reproduire.

Importance de la courbure



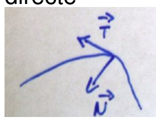
(a). Attneave's cat (b). Curvature map.

Expérience de F. Attneave (1954) : 38 points de courbure maximale (extrait de Ciomaga et al. 2011)

Rappels mathématiques sur les courbes planes

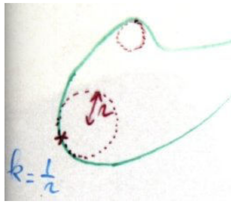
- $C : [a,b] \to \mathbb{R}^2$ est une **courbe de Jordan**
  si $C(p_1) \neq C(p_2)$ pour $p_1 \neq p_2$.

- le choix de la fonction C pour représenter une courbe n'est pas unique (différentes **paramétrisations**)
- $L(a,p)$: longueur de C entre a et p
  la paramétrisation est dite **euclidienne** si $\frac{dL}{dp} = 1$
- Si C est 2 fois différentiable et $C'(p) \neq 0$
  le vecteur tangent est $\overrightarrow{T} = \frac{C'(p)}{|C'(p)|}$
  le vecteur normal $\overrightarrow{N}$ est tel que $(\overrightarrow{T}, \overrightarrow{N})$ soit une base orthonormée directe

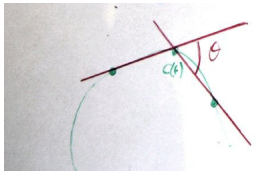


- **courbure**
  $\exists k$ tel que
  $$\frac{1}{|C'|}\frac{d\overrightarrow{T}}{dp} = k\overrightarrow{N}$$
  et k est indépendant de la paramétrisation
  $k\overrightarrow{N}$ s'appelle le vecteur courbure
- cas d'une paramétrisation euclidienne :
  $$\overrightarrow{T} = C'(s),$$
  $$\frac{d\overrightarrow{T}}{ds} = k\overrightarrow{N} = C''(s)$$
  (car $L(a,p) = \int_a^p |C'(u)|du$, donc $|C'(s)| = 1$.
- en notant $C(p) = (x(p), y(p))$, on a
  $$k = \frac{y''x' - x''y'}{(x'^2 + y'^2)^{3/2}}$$

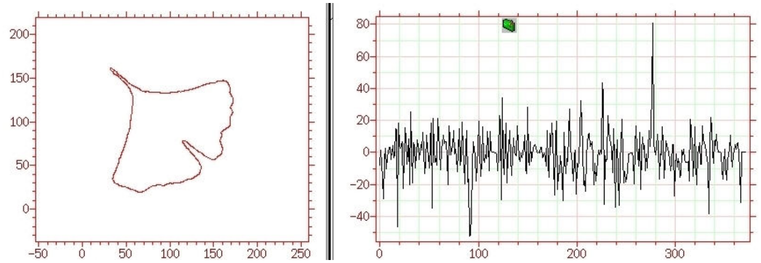
- la courbure vérifie $k(p) = r(p)^{-1}$, où $r(p)$ est le rayon du cercle qui approche le mieux la courbe en $C(p)$ (cercle osculateur)



- En pratique on peut approcher la courbure par la différence des directions des tangentes en deux points successifs (évite le calcul des dérivées secondes, peu robustes)



$$k \approx \frac{\theta}{\Delta s}$$

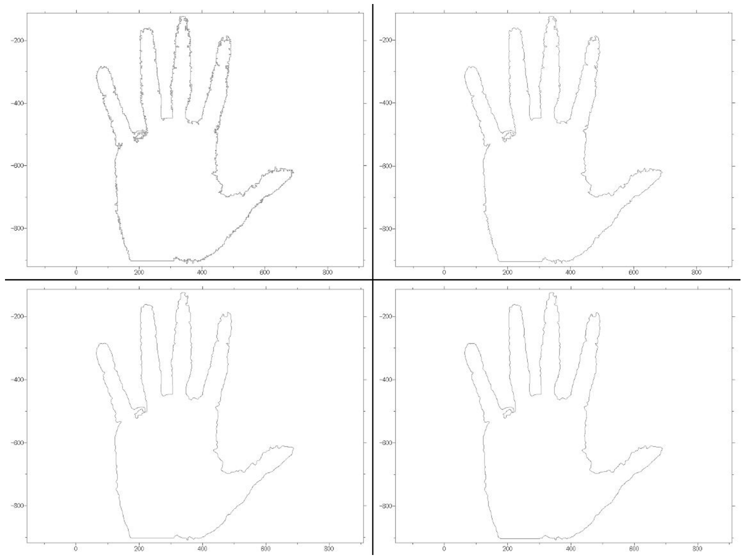


Régularisation de courbes

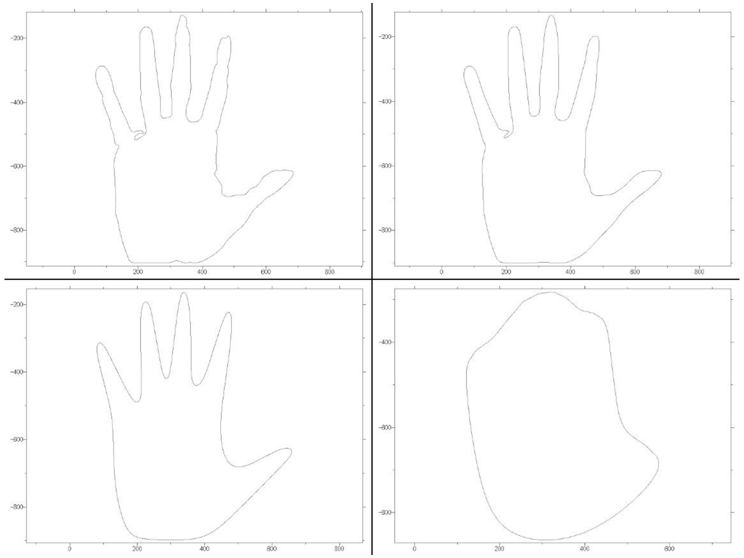
- On veut construire, à partir d'une courbe initiale $C_0 = \begin{pmatrix} x \\ y \end{pmatrix}$, une famille de courbes C_σ de plus en plus régulières (qui comporte de moins en moins de détails).
- Soit $G_\sigma(x) = \frac{1}{\sqrt{(2\pi)}\sigma} exp(-x^2/2\sigma^2)$ un noyau Gaussien
- On définit

$$C_\sigma = \begin{pmatrix} x * G_\sigma \\ y * G_\sigma \end{pmatrix}$$

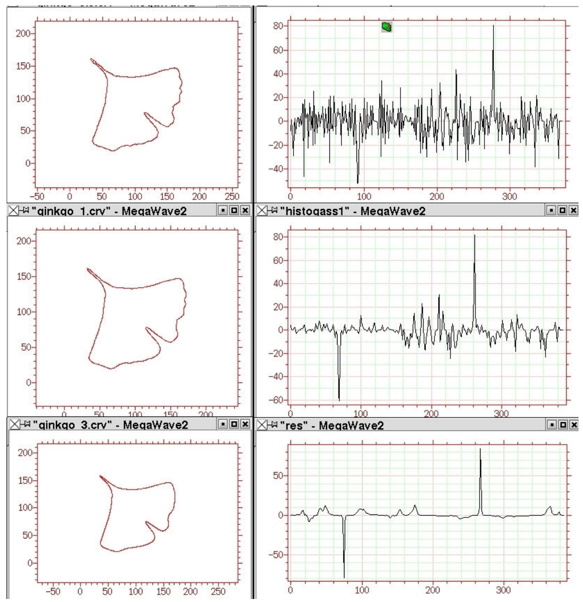
- $\rightarrow$ espace échelle linéaire (on y reviendra pour le cours sur les points d'intérêt).
- Considérer une courbe C_σ pour un σ suffisamment grand permet d'appliquer des opérateurs différentiels (en particulier la courbure) à la forme.



$\sigma = 1, 2, 3, 4$



$\sigma = 5, 10, 30, 100$



La courbure en fonction de l'abscisse curviligne, pour différentes échelles

Espace-échelle de courbure (Mokhtarian et al. 1992)

- L'ensemble des représentation d'une forme à différentes échelles $\{C_\sigma\}_{\sigma \geq 0}$ permet un codage efficace de l'information visuelle
  - Construction du descripteur :
  - Pour chaque valeur de σ on calcule les positions s (abscisse curviligne normalisée entre 0 et 1) de la courbure
  - Chacun de ces passages par zéro correspond à une position $(s, \sigma) \rightarrow$
  - On retient comme descripteurs l'ensemble des (s, σ) pour lesquels σ est maximum
- Pour comparer deux formes, on compare leurs descripteurs après les avoir alignés en s (une translation en s correspond à une rotation de la forme)
- Procédure robuste au bruit / invariante aux translations, rotations, zoom

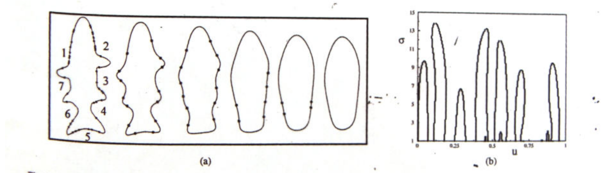
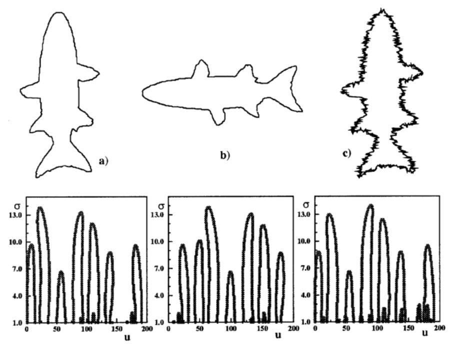
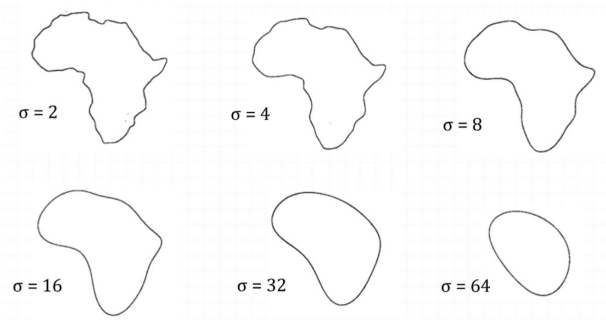
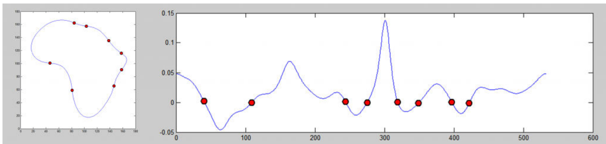


Fig. 1. (a) Shrinkage and smoothing of the curve and decreasing of the number of curvature zero crossings during the evolution, from left: σ = 1, 4, 7, 10, 12, 14. (b) The CSS image of the shape.



Curvature scale space - CSS (extrait de Mokhtarian et al. 1992 & 1999)





Curvature scale space - CSS (illustration opencv)

Extraction de formes paramétriques : Hough et RANSAC

Transformée de Hough

- Partant de $\{p_1, \ldots, p_n\}$ ensemble de points, on cherche des formes paramétrées dans
$$\{F_\theta, \theta \in \Theta\}.$$
- Transformée de Hough :
$$p_i \to \mathcal{C}_i \subset \Theta,$$
avec
$$\mathcal{C}_i = \{\theta : p_i \in F_\theta\}.$$
- Puis on s'intéresse à
$$\sum_{i=1}^{n} \mathbf{1}(\mathcal{C}_i),$$
votes des points pour les formes paramétrées.

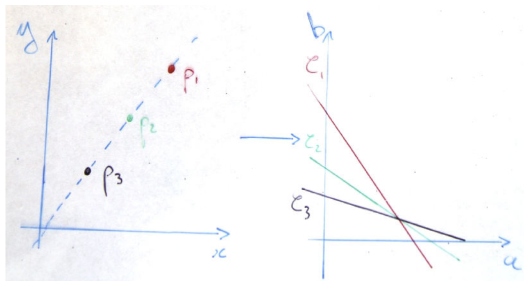
Exemple des droites

$$F_{a,b} : y = ax + b$$

alors si $p_i = (x_i, y_i)$,

$$\mathcal{C}_i = \{(a, b) : b = -x_i a + y_i\}$$

donc point $\to$ droite dans Θ.



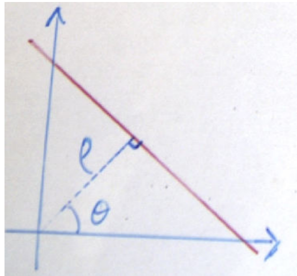
Choix de la paramétrisation (cas des droites)

$y = ax + b$ est une mauvaise paramétrisation

- $a, b \in (-\infty, \infty)$
- les variations sur a et b ne correspondent pas à des variations uniformes dans l'espace image

On choisit
$x\cos(\theta) + y\sin(\theta) = \rho$

- θ borné , ρ borné pour une image bornée
- $d\theta \to$ rotation dans le plan image
- $d\rho \to$ translation dans le plan image

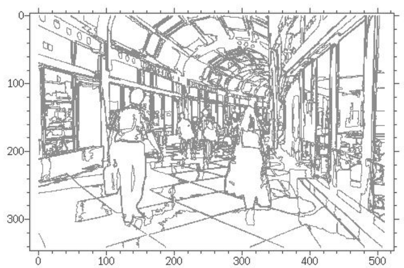


Transformée de Hough

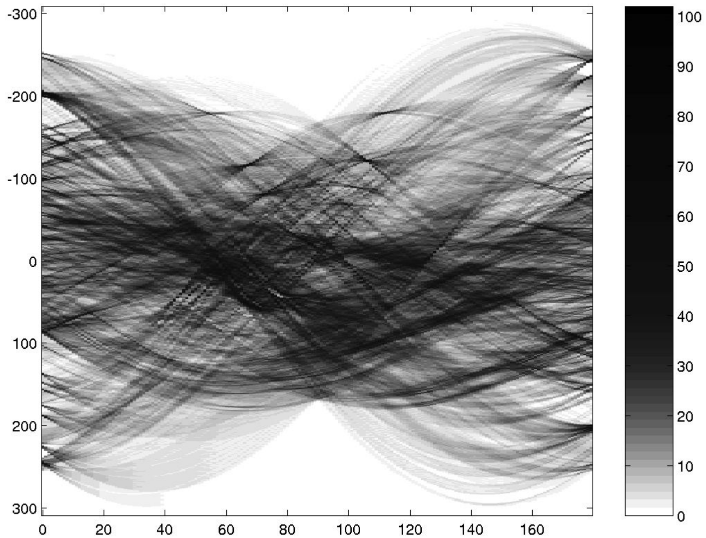
En pratique :

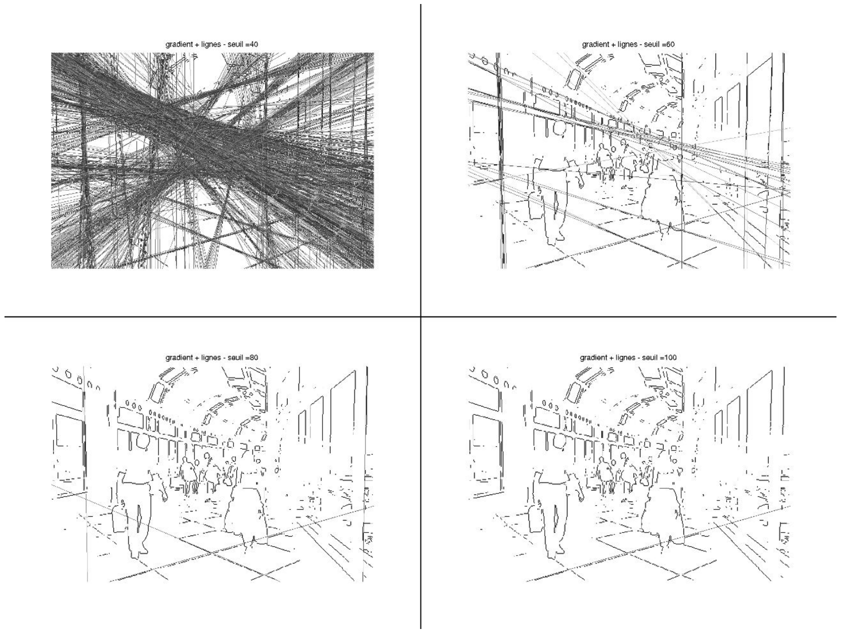
- Les p_i sont des pixels (typiquement des points de contour)
- L'espace Θ est partagé en cellules
- Pour chaque cellule, on compte les courbes l'intersectant $\to$ problème du choix de la taille des cellules
  - petites : bonne précision sur les paramètres, mais peu de votes
  - grandes : plus de votes, moins de précision

Segmentation initiale



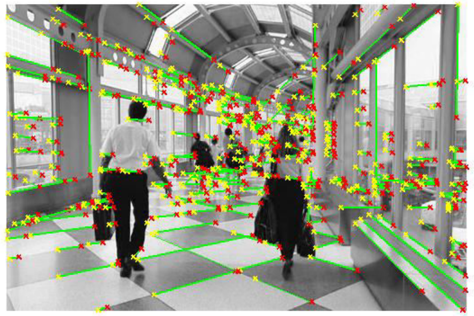
Espace des paramètres





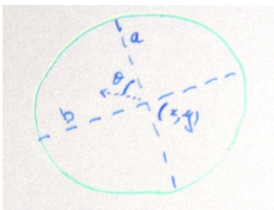
Après post-traitement

Post-traitement : on se restreint aux segments inclus dans l'image des contours et correspondant aux pics de la transformée de Hough



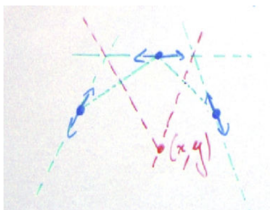
Pics $> 0.3 * \max(Hough)$;
longueur minimum 7 ;
deux segments sont groupés si à une distance inférieure à 5

Exemple des ellipses



5 paramètres : (x, y, a, b, θ)

- Θ est très densément occupé par les $\mathcal{C}_i$
  $\rightarrow$ on cherche les centres d'abord,
  puis les (a, b, θ) correspondant
- Pour trouver les centres, on utilise 3 points et des directions (e.g. calculées par le gradient)



RANSAC

RAndom SAmple Consensus (Fischler - Bolles 1981)

Le cas des droites

Données : $p_i = (x_i, y_i)$

Initialisation $\Delta = \emptyset, N = 0$

On itère

- p_i, p_j tirés au hasard $\rightarrow$ droite $\Delta_{i,j}$
- $N_{i,j} = \{p_k : d(p_k, \Delta_{i,j}) \leq \epsilon\}$
- si $N_{i,j} > N : \Delta = \Delta_{i,j}, N = N_{i,j}$

Utilisation classique : détermination de transformations entre images à partir de correspondances entre descripteurs locaux (cf deuxième partie du cours)