

Chapitre 5

Description de contours et de formes

Chapitre rédigé par Henri MAÎTRE

Dans de nombreuses applications de traitement des images, une fois l'image segmentée (on peut le faire, de la façon la plus simple, en seuillant l'image [Sahoo et al., 1988, Cham et al., 1998] ou à partir des méthodes décrites aux chapitres 2 et 3), on essaie de reconnaître les divers objets qui la composent à partir de leur seule silhouette. C'est particulièrement vrai dans les applications de tri dans lesquelles on est intéressé à saisir ou trier des objets arrivant sur un convoyeur¹, dans les applications de reconnaissance de caractères, ainsi que pour des applications de surveillance ou de guidage, par exemple en imagerie militaire.

Il est utile alors que l'utilisateur dispose d'une représentation de la forme vérifiant plusieurs propriétés :

1. une bonne fidélité à la forme initiale,
2. une bonne discrimination de formes différentes,
3. une bonne adaptation aux opérations de reconnaissance des formes, et en particulier une insensibilité aux déformations qui sont susceptibles d'entacher l'objet,
4. une certaine compacité pour permettre l'archivage de nombreuses formes et potentiellement de ces mêmes formes sous divers aspects.

De nombreuses représentations des formes ont été développées concurremment, chacune pour répondre à un problème parfois assez spécifique et donc mettant l'accent sur l'une ou l'autre des propriétés ci-dessus.

5.1 Fonction caractéristique

La représentation naturelle d'une forme dans une image est une représentation par fonction caractéristique, c'est à dire sous forme d'une image binaire, les seuls pixels non-nulles étant ceux couverts par la forme :

$$\begin{aligned} f(i, j) &= 0 && \text{si } (i, j) \notin \text{objet} \\ &= 1 && \text{si } (i, j) \in \text{objet} \end{aligned}$$

Dans une image où de nombreux objets existent, on utilise une image d'étiquettes (ou de *labels*) qui généralise la notion de fonction caractéristique :

$$f(i, j) = k \quad \text{si } (i, j) \in \text{objet } k$$

¹Le tri se partage en tri planaire (où les objets sont préalablement mis à plat, généralement sur un tapis roulant), et en tri en vrac, où ils sont présentés dans un désordre total.

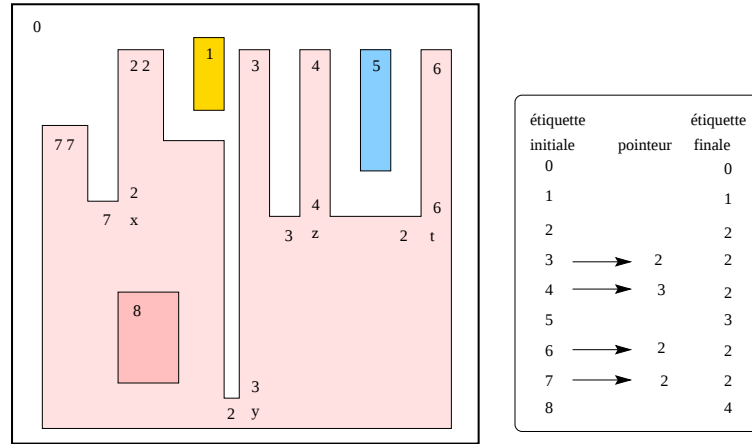


FIG. 5.1 – Etiquetage d’une image en 4 connexité. L’image de gauche est le résultat du premier passage jusqu’au point noté x. Les étiquettes ont été attribuées à chaque point soit en donnant une nouvelle étiquette si les 2 antécédents du point sont à 0, soit en donnant l’étiquette des antécédents si ces antécédents ont une étiquette semblable différente de 0. En x les étiquettes sont de valeur différente. Le pixel x se voit attribué l’étiquette 2 car $2 = \min(2, 7)$, le pixel y l’étiquette 2, le pixel z l’étiquette 3 et le pixel t l’étiquette 2. Le point x crée le pointeur $7 \sim 2$, y le pointeur $3 \sim 2$, z $4 \sim 3$ et t $6 \sim 2$. Après un second passage sur l’image, on détermine la liste finale des étiquettes de 0 à 4. Pour recréer une image d’étiquettes en conformité avec cette liste, il faut donc faire un deuxième passage sur l’image.

Une image d’étiquettes résulte d’une procédure d’étiquetage. Elle se fait par le choix d’une connexité qui permet de définir les propriétés de topologie des divers objets (4- ou 8- connexité sur les trames carrées, 6-connexité en trame hexagonale). L’étiquetage vient attribuer une étiquette semblable à tous les points connexes au sens de la segmentation. Par exemple en 4-connexité un algorithme d’étiquetage fonctionne de la façon suivante (cf. figure 5.1).

1. une liste d’étiquettes est initialisée à 0 ;
2. le balayage (par exemple vidéo) de l’image est initialisé ; au premier point est attribuée l’étiquette 0 ;
3. un point nouveau est pris dans l’ordre du balayage ; on examine sa connexité à son voisin supérieur et à son voisin de gauche (ses antécédents) :
 - si le point courant est connexe à ses deux antécédents (au sens de la segmentation), et si ses deux antécédents ont la même étiquette, on lui attribue l’étiquette de ces antécédents ;
 - si le point courant est est connexe à ses deux antécédents, mais que ses deux antécédents n’ont pas la même étiquette, on donne au point l’étiquette la plus petite parmi les deux, puis dans la liste des étiquettes, on crée un pointeur qui renvoie de l’étiquette la plus grande vers l’étiquette la plus petite ;
 - si le point courant est connexe à l’un seulement de ses antécédents, on lui attribue l’étiquette de cet antécédent ;
 - si le point courant n’est pas connexe à ses deux antécédents, on incrémente la liste des étiquettes, puis on attribue au point courant cette nouvelle étiquette ;
4. on retourne en 3 jusqu’à atteindre le dernier point de l’image.

Lorsque l’ensemble de l’image a été ainsi traité (après donc un passage complet sur l’image), on réorganise l’image des étiquettes. Pour cela on parcourt cette image et chaque valeur d’étiquette est remplacée par l’étiquette terminale rencontrée en remontant la liste des pointeurs. La liste des étiquettes est alors filtrée de toutes les étiquettes qui possèdent au moins un pointeur. On constitue ainsi une liste d’étiquettes qui occupe tous les entiers entre 0 et k_{max} s’il y a k_{max} objets se découpant sur un fond.

5.2 Description de formes

Les descripteurs (ou paramètres) de forme sont des nombres qui représentent chaque forme et permettent de les classer (par exemple par nuées dynamiques ou plans séparateurs). Ces paramètres ont été abondamment proposés dans la littérature [Coster et Chermant, 1985], beaucoup étant adaptés à des formes particulières. Citons les plus utilisés, par exemple :

- le rapport iso-périmétrique, proportionnel au rapport du carré du périmètre de l'objet à sa surface (dans des images continues, il est maximum pour le cercle);
- l'indice d'allongement [Schmitt et Mattioli, 1994], proportionnel au rapport du carré du diamètre géodésique à la surface de l'objet, (minimal pour les disques au sens du voisinage choisi).

Les paramètres de forme sont bien adaptés pour des objets de taille assez grande car ils ont souvent été définis pour des formes continues. Pour des objets de petite taille (quelques pixels), la discrétisation du maillage induit des comportements souvent peu satisfaisants.

5.2.1 Représentation par les moments

Connaissant l'objet par sa fonction caractéristique $f(x, y)$, une représentation classique de sa forme consiste à en mesurer les divers moments :

$$M_{mn} = \iint x^m y^n f(x, y) dx dy$$

En particulier les moments centrés (rapportés au centre de gravité (X_g, Y_g) de la forme) sont invariants par translation :

$$\bar{M}_{mn} = \iint (x - X_g)^m (y - Y_g)^n f(x, y) dx dy$$

Sur une image discrète, ces moments s'écrivent :

$$\bar{M}_{mn} = \frac{1}{J^m K^n} \sum_{j=1}^J \sum_{k=1}^K (j - X_g)^m (k - Y_g)^n f(j, k)$$

Les moments d'inertie (valeurs propres de la matrice d'inertie, matrice 2×2 de terme courant $\bar{M}_{mn}$, $m+n=2$) sont invariants par rotation. Les moments d'inertie normés par la plus grande valeur propre sont invariants par similitude (rotation et facteur d'échelle). Les moments d'inertie décrivent bien l'allongement de formes régulières comme des ellipses ou des distributions gaussiennes. Ils sont plus ambigus sur des formes complexes (cf. figure 5.2).

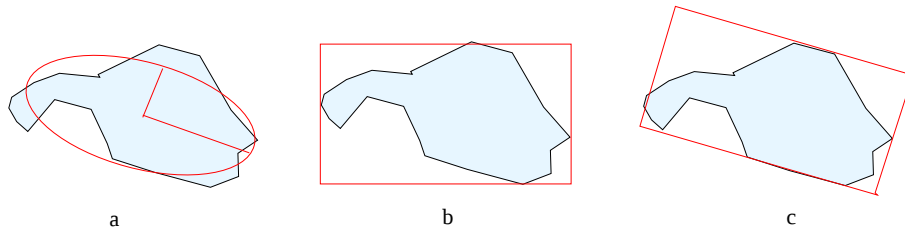


FIG. 5.2 – Description de formes par : a - ellipse et axes d'inertie, b - boîte englobante, c - boîte minimale.

5.2.2 Représentation par les moments invariants

Des moments invariants ont été proposés (nommés moments de Hilbert), invariants par translation, rotation et changement d'échelle (dans la limite de l'approximation par une maille discrète) [Hu, 1962]. Ils se construisent à

partir du moment normé :

$$N_{nm} = \frac{J^m K^n}{\bar{M}_{00}^\alpha} \bar{M}_{mn}$$

où $\alpha = \frac{1}{2}(m + n) + 1$, par les formules :

$$\begin{aligned} h_1 &= N_{20} + N_{02} \\ h_2 &= (N_{20} - N_{02})^2 + 4N_{11}^2 \\ h_3 &= (N_{30} - 3N_{12})^2 + (N_{03} - 3N_{21})^2 \\ h_4 &= (N_{30} + N_{12})^2 + (N_{03} + N_{21})^2 \\ h_5 &= (N_{30} - 3N_{12})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{03} + N_{21})^2] \\ &\quad + (3N_{21} - N_{03})(N_{03} + N_{21})[3(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ h_6 &= (N_{20} - N_{02})[(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ &\quad + 4N_{11}(N_{30} + N_{12})(N_{03} + N_{21}) \\ h_7 &= (3N_{12} - N_{30})(N_{03} + N_{21})[3(N_{30} + N_{12})^2 - (N_{03} + N_{21})^2] \\ &\quad + (3N_{21} - N_{03})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{03} + N_{21})^2] \end{aligned}$$

5.2.3 Boîtes englobantes et boîtes minimales

Lorsque les formes sont plus irrégulières, on préfère souvent des descriptions de l'allongement par la forme de la boîte englobante :

- soit la boîte alignée sur les axes et donc simplement définie par ses dimensions ($x_{max} - x_{min}$) et ($y_{max} - y_{min}$),
- soit la boîte orientée selon le diamètre minimal de l'objet (de calcul un peu plus complexe) (cf. figure 5.2).

5.3 Polygones de Guzman

C'est l'une des méthodes les plus anciennes [Guzman, 1968]. L'approche de Guzman consiste à envelopper l'objet à reconstruire dans des « boîtes » de formes de plus en plus précisément adaptées. Elle peut donc s'appliquer soit à partir d'une image d'étiquettes de la forme à analyser, soit à partir d'une liste des pixels de contour de la forme. Les formes prototypes auxquelles seront comparées les objets sont construites sur un maillage carré de façon systématique et se classent en niveaux en fonction de la longueur de leur périmètre. Cette longueur (exprimée en pixels) étant obligatoirement paire, on décrit chaque niveau par le demi-cardinal du nombre de pixels constituant la frontière. Dans chaque niveau, les divers prototypes sont repérés par un indice qui les identifie uniquement. Ainsi, la forme la plus simple (le carré unitaire) a une longueur 4 (et donc appartient au niveau 2). C'est la forme $\Phi(2, 1)$. Afin de tenir compte des symétries et des rotations de $\pi/2$, toutes les formes identiques par rotation modulo $\pi/2$ et toutes les formes identiques par symétrie droite (par rapport à un axe horizontal ou vertical) sont rapportées à une même prototype du dictionnaire des formes (cf. figure 5.3). Un objet quelconque est donc décrit par une succession d'indices décrivant l'objet à divers niveaux de résolution. Deux objets sont généralement identiques jusqu'à un niveau n , puis différents à partir du niveau $n + 1$ (cf. figure 5.4).

Cette approche rencontre plusieurs limites :

- pour les ordres grands, les formes deviennent très nombreuses et le dictionnaire trop grand. La recherche du représentant se faisant par comparaison au prototype devient très longue ;
- les distances d'une forme à un prototype ne s'imposent pas de façon unique (cf. figure 5.5), laissant place à des classements différents selon le critère adopté ;

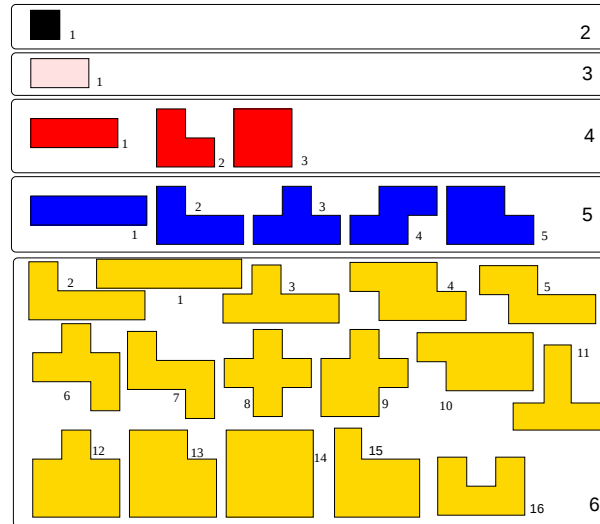


FIG. 5.3 – Les premières formes du dictionnaire de Guzman.

- les relations « verticales » entre niveaux sont complexes et ne permettent pas d’accélérer le parcours de l’arbre de façon commode ;
- certaines formes se retrouvent identiques dans des niveaux différents, à un facteur d’échelle près (par exemple le carré élémentaire se retrouve dans tous les niveaux impairs).

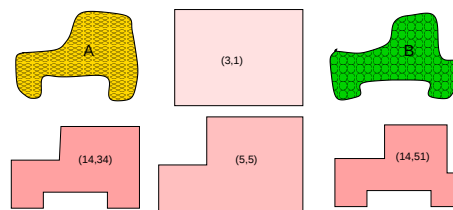


FIG. 5.4 – Représentation de Guzman. Les 2 formes A et B sont semblables aux niveaux de représentation 3 et 5, mais différent au niveau 14, où l’une est représentée par le prototype 34 et l’autre par le prototype 51.

L’approche de Guzman a été abandonnée aujourd’hui, mais elle apporte des concepts intéressants à la description des formes :

- la notion de raffinement de la description,
- la notion de similarité entre tous les objets « vus de très loin »,
- la notion d’invariance par rotation, symétrie et changement d’échelle que l’on pourrait étendre éventuellement à d’autres transformations.

5.4 Chaînes de Freeman

C’est la méthode la plus ancienne de description des contours dans les images et aussi la plus utilisée encore aujourd’hui [Freeman, 1961, Freeman, 1977]. Si elle est moins utilisée en reconnaissance des formes qu’elle le fut, elle est très fréquemment utilisée dans des applications récentes comme la transmission des images par zones (par exemple dans MPEG-).

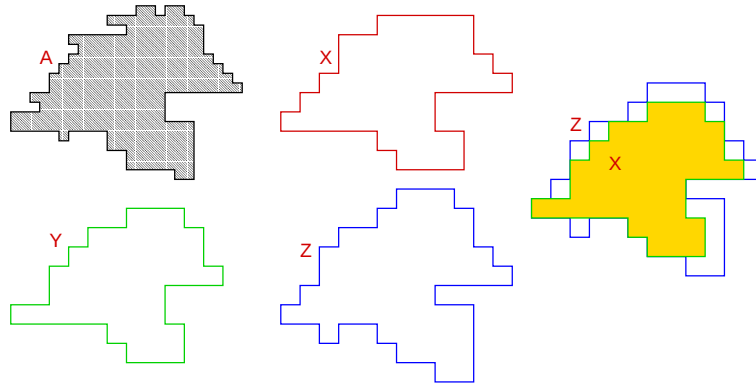


FIG. 5.5 – L'arbitraire de l'approximation des formes : la forme A est l'objet. Les 3 formes X, Y et Z en sont des approximations. La distance d'une représentation à la forme initiale peut se mesurer par la surface entre les 2 contours. Il y a de nombreuses formes à distance minimale. X est une forme la plus proche en terme de surface et « la plus simple » au sens du nombre d'angles droits lorsqu'il y a plusieurs solutions de même distance. Y est la forme la plus proche par défaut (entièrement contenue dans A), Z est celle par excès. Z et Y sont superposées à droite.

5.4.1 Définition

C'est une technique de représentation des directions du contour (on code la direction le long du contour dans un repère absolu lors du parcours du contour à partir d'une origine donnée). Les directions peuvent se représenter en 4-connextité (codage sur 2 bits) ou en 8-connextité (codage sur 3 bits)².

Le codage d'un contour se fait donc de la façon suivante :

1. transmission des coordonnées absolues du point de départ,
2. transmission de la liste des codes de déplacement d'un point du contour au suivant sur le maillage.

Les codes des contours sont donnés par la figure 5.6.

Dans d'autres techniques, on code de façon différentielle le changement de direction d'un point au suivant. Cela peut se justifier en codage sur 3 bits si une direction est très dominante par rapport aux autres.

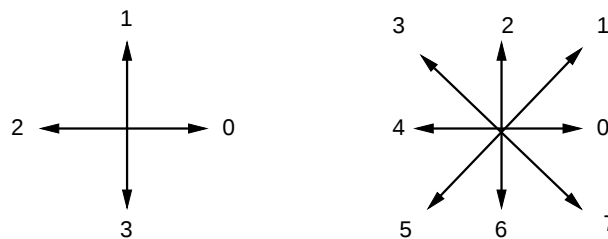


FIG. 5.6 – Les codes de Freeman en 4-connextité (à gauche) et en 8-connextité (à droite).

Deux exemples de codages par chaînes de Freeman sont présentés sur la figure 5.7 (le codage du positionnement absolu a été omis). Dans le cas de la 4-connextité, le codage de la suite de contours occupe 124 bits (62×2), tandis qu'en 8 connextité elle occupe 126 bits (42×3). Il n'y a pas de règle générale sur l'efficacité des représentations en termes de compression. Si la courbe est très complexe (beaucoup de changements de direction), le codage en

²des essais ont été faits pour coder également les chaînes de Freeman sur 4 bits en ajoutant les directions reliant le point central à la seconde couronne des voisins qui ne sont pas dans la direction des voisins de la première couronne. Mais cette approche n'a pas eu un très grand succès.

6. Simplification d'un chemin : c'est un chemin dont on a supprimé des détails sans changer globalement la forme. Cela s'obtient en remplaçant des séquences de p descripteurs consécutifs par des descripteurs équivalents reliant les mêmes points (cf. figure 5.8) :
exemple en 4-connexité : $\{012\} \rightarrow \{1\}$,
en 8-connexité : $\{03\} \rightarrow \{2\}$.
7. Réduction d'un chemin : c'est l'un des chemins de longueur minimale reliant les 2 extrémités de la courbe initiale. On associe 2 par 2 des descripteurs inverses de la chaîne et on les supprime.
exemple en 4-connexité : $X = \{00132122\} \rightarrow \tilde{X} = \{21\}$. On obtient tous les chemins réduits en changeant l'ordre des associations.
en 8-connexité la réduction est un peu plus complexe car il faut aussi regrouper des ensembles de 3 ou 4 descripteurs qui s'annulent : exemple $\{025\}$ ou $\{7225\}$.
8. Fermeture d'un contour : on teste la fermeture d'un contour en vérifiant que la chaîne réduite est nulle.
On ferme un contour en lui ajoutant le chemin inverse d'un de ses chemins réduits. Il y a beaucoup d'autres fermetures possibles que par addition de l'inverse d'un réduit, mais les fermetures obtenues ainsi sont de longueur minimale. Il y a d'autres fermetures minimales que celles obtenues par addition de l'inverse d'un réduit.
9. Courbe qui s'intersecte : pour savoir si une courbe décrite par sa chaîne se recoupe, on procède à une réduction de chemin systématique en partant de son origine et en testant si chaque nouveau descripteur possède un inverse dans la chaîne déjà parcourue. Si à un instant la chaîne déjà parcourue se réduit à une chaîne nulle, on a trouvé un point double.
10. Changement d'origine : le changement de l'origine d'une chaîne de longueur L revient à une permutations circulaires des descripteurs modulo L .
11. Sens de parcours d'un contour : le contour fermé d'une forme simplement connexe peut être décrit dans le sens direct ou dans le sens inverse. Pour connaître le sens d'un contour fermé, on réduit ce contour jusqu'à n'avoir que 4 descripteurs. En 4-connexité, il n'existe que 2 chaînes possibles (et celles qui s'en déduisent par changement d'origine) :
– le carré direct : $X_D = 0123$,
– le carré inverse : $X_I = 0321$.
En 8-connexité les configurations sont un peu plus nombreuses.
12. Surface d'une région simplement connexe : la technique est assez complexe mais ne nécessite pas la reconstruction de la région [Freeman, 1977].

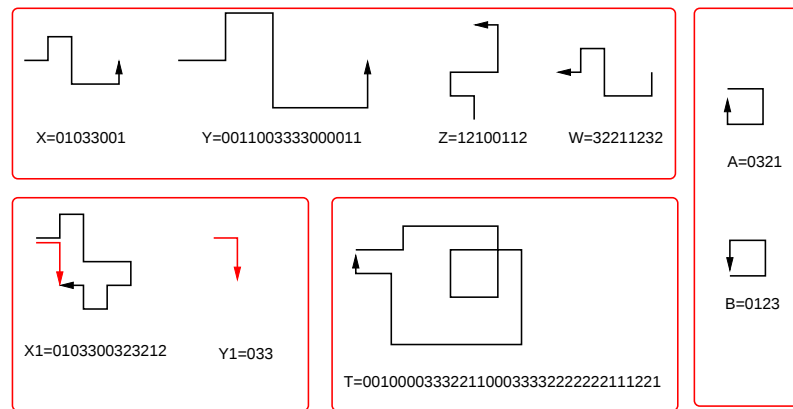


FIG. 5.9 – Quelques propriétés des chaînes de Freeman : X = chaîne originale, Y = contour double, Z = rotation de $\pi/2$, W = contour inverse. X1 = chaîne originale et Y1 = chaîne réduite. T = exemple de boucle dans une chaîne. A et B sont les deux contours élémentaires (inverse et direct) en 4-connexité.

5.4.3 Reconnaissance des formes par des chaînes de Freeman

Pour que 2 formes X et Y soient identiques, il faut tout d'abord qu'elles aient le même nombre de descripteurs. Puis, comme elles peuvent différer par le sens de description, il faut comparer X à Y d'une part et à $\bar{Y}$ d'autre part. Enfin, comme elles peuvent avoir des origines différentes, il convient de tester les deux chaînes descripteur à descripteur sous toutes les hypothèses de permutation circulaires. Des techniques rapides ont été développées pour faire ces recherches (d'une complexité sous-linéaire en L), [Boyer et Moore, 1977, Miclet, 1984]³.

Dans le cas où l'on ne recherche pas une similarité exacte mais seulement approchée (par exemple pour prendre en compte la présence du bruit lors de la détection des contours), on est amené à utiliser des techniques de **distance d'édition** pour lesquels on chiffre les coûts des erreurs possibles : omission de contours, adjonction de contours, permutation de descripteurs, etc. On utilise alors généralement des techniques de programmation dynamique (algorithme de Wagner et Fisher [Wagner et Fisher, 1974, Miclet, 1984], élargi pour des substitutions de plusieurs descripteurs consécutifs dans [Lowrance et Wagner, 1975]) ou de relaxation (algorithme de Davis) pour retrouver la distance minimale entre X et Y .

³Ce sont des automates finis semblables à ceux qui sont utilisés en particulier pour les compilateurs et les éditeurs de texte.

5.5 Descripteurs de Fourier

5.5.1 Descripteur par tangente

Dans cette approche on considère le contour comme une courbe continue qui peut être décrite par son abscisse curviligne s à partir d'une origine A choisie (cf. Fig 5.10). On paramètre la courbe par l'angle fait par le vecteur tangent en chaque point et celui au point origine : $\phi(s)$ et on crée la variable réduite t qui prend ses valeurs entre 0 et 2π : $t = 2\pi s/L$, où L est la longueur complète du contour. On construit alors la fonction $\Phi(t)$:

$$\Phi(t) = \phi \left[\frac{2\pi \cdot s}{L} \right] - \frac{2\pi s}{L}$$

le terme correctif prenant en compte l'enroulement de 2π de la tangente pour un tour de contour. La fonction $\Phi(t)$ est une fonction périodique sur $[0, 2\pi[$ qui admet donc une série de Fourier : $\Phi(t) = \sum_{k=0}^{\infty} a_k \exp(-ikt)$. On appelle descripteurs de Fourier l'ensemble des modules des a_k : $\{|a_k|\}$.

Ils bénéficient des propriétés suivantes :

1. ils sont invariants par translation de la forme,
2. ils sont invariants par changement d'échelle (puisque t est normalisé),
3. ils sont invariants par rotation, puisque l'on a choisi la différence d'angle entre 2 tangentes,
4. ils sont invariants par changement d'origine, car passer d'une origine A à une origine A' revient à :
 - retrancher $\Phi(t_{A'})$ à toutes les valeurs $\Phi(t)$,
 - changer t en $t - t_{A'}$.

Donc : $\Phi_{A'}(t) = \Phi_A(t) * \delta(t - t_{A'}) - \Phi(t_{A'})$ et : $a_k \rightarrow a_k \exp(-kt_{A'})$ si $k \neq 0$

Pour comparer des formes on compare leurs descripteurs par ordre croissant. Si de plus on veut simplifier le contour, il suffit de supprimer les ordres k élevés dans le développement. malheureusement, dans cette représentation, si un contour est fermé, le contour obtenu en filtrant les hautes fréquences ne l'est généralement plus. C'est pourquoi on préfère souvent les descripteurs de Fourier par représentation complexe qui n'ont pas ce défaut.

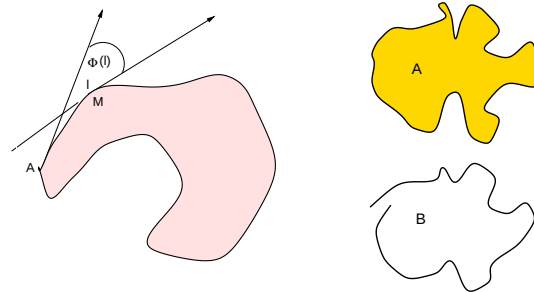


FIG. 5.10 – A gauche : descripteurs de Fourier par tangente. Le point de départ A est choisi arbitrairement. Son vecteur tangent sert de référence à la paramétrisation de la courbe. A droite : après troncature du développement de Fourier, la forme A , initialement fermée devient B , plus régulière, mais non fermée.

5.5.2 Représentation complexe

Dans cette représentation, on décrit la forme par un ensemble $\{M_j\}$ de points de contours, et on représente la forme dans le plan complexe. On attache donc à chaque M_j un nombre complexe $z_j = x_j + iy_j$. On appelle alors

descripteurs de Fourier, les coefficients de la TF Z de z :

$$Z_k = \sum_{j=1}^N z_j \exp(-2\pi i j k)$$

Les coefficients Z_k , pour $k \in [-N/2 + 1, N/2]$, jouissent d'intéressantes propriétés [Bertrand et al., 1982].

1. Pour $k = 0$, Z_0 est le centre de gravité de la forme. Si l'on l'omet, la description est invariante par translation.
2. Si tous les Z_k sont nuls sauf pour $k = 1$, la forme est un cercle de rayon Z_1 (ou un polygone régulier à N côtés), donc Z_1 joue le rôle de facteur d'échelle. La normalisation par Z_1 rend la forme invariante par homothétie.
3. Les coefficients $Z_{|k|}$ et $Z_{|1-k|}$ (pour $k \neq 1$ et $k \neq 0$) jouent des rôles symétriques (mais opposés) de la façon suivante :
 - l'ordre k indique le nombre d'actions sur le cercle unité (entre 0 et 2π) : 1 action pour $k = 2$ et $k = -1$, 2 actions pour $k = 3$ et $k = -2$, 3 actions pour $k = 4$ et $k = -3$, etc. Ces actions sont réparties régulièrement autour du cercle unité,
 - les valeurs de $k > 0$ indiquent des actions de traction sur la courbe, pour la déformer vers l'extérieur du cercle unité, les valeurs de $k < 0$ indiquent des actions de pression sur la courbe, pour creuser la courbe vers son centre,
 - la phase du nombre complexe Z_k : ϕ_k exprime le lieu, sur le cercle unité, où s'exerce l'action.
4. Plus les coefficients sont nombreux, plus la forme est complexe. Et plus les coefficients sont élevés plus les détails sont fins sur la courbe.

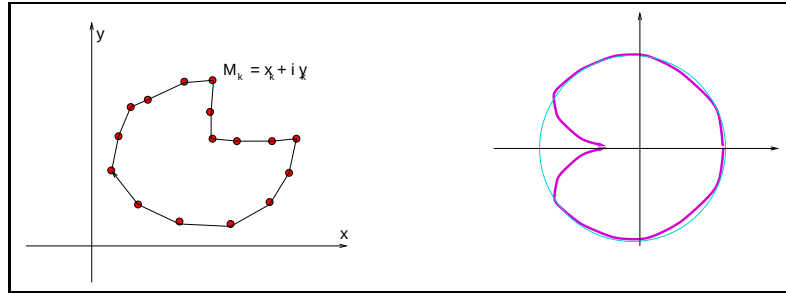


FIG. 5.11 – A gauche : descripteurs de Fourier par représentation complexe. Le point courant M_k est décrit par ses coordonnées complexes dans le plan image. A droite : le cercle résulte de la prise en compte du seul coefficient Z_1 , la courbe présentée résulte de l'adjonction d'un coefficient de pression (donc à $k < 0$), en 1 seul point (donc $k = 2$ ou $k = -1$, ici $k = -1$), de déphasage nul (puisque situé en π).

On voit que les descripteurs de Fourier par représentation complexe ont le même type de comportement que les descripteurs par tangente. Ils sont mieux adaptés aux formes discrètes puisqu'ils garantissent toujours que la forme demeure fermée après troncature du développement de Fourier. Ils peuvent être invariants par rotation si l'on s'intéresse aux seuls modules des coefficients Z_k .

Parmi leurs inconvénients, il faut noter que l'on ne peut pas aisément garantir qu'un contour de forme simplement connexe ne donnera pas, après troncature, un contour qui s'auto-intersectera.

5.6 Approximations polynomiales

Il y a essentiellement 3 types de telles approximations :

1. les approximations analytiques par ajustements de nuages de points par des polynômes du premier ordre (conduisant donc à des polygones);
2. les approximations polygonales par des critères géométriques obtenues en parcourant la courbe des points ordonnés selon une abscisse curviligne;
3. les approximations par des polynômes de degré > 1 et en particulier les approximations par des fonctions splines.

Les approximations polygonales des formes décrites par des points de contour sont les représentations les plus employées. Dans les cas les plus simples, on se contente de relier 2 à 2 des points de contour dans un ordre déterminé au préalable (donc selon une abscisse curviligne croissante). Deux problèmes se posent alors :

- peut-on éliminer d'éventuels points parasites qui n'appartiennent pas au contour ?
- peut-on faire l'économie de certains sommets de la ligne polygonale pour obtenir des formes plus simples ?

Ces problèmes seront abordés dans le paragraphe 5.6.2.

Mais avant d'en arriver à cette situation simple, se posent les problèmes difficiles de déterminer quels points appartiennent à un même contour et dans quel ordre les relier lorsque ces points sont issus d'un détecteur de contour par nature local. C'est ce que nous abordons tout d'abord.

5.6.1 Approximation d'un nuage de points par une droite unique

C'est un problème bien classique mais qui mérite pourtant un peu d'attention. Soit $M_i = (x_i, y_i)$ les points, en nombre N , que l'on cherche à approcher par une droite. Il y a 2 façons de prendre le problème.

Approximation par régression linéaire

C'est une approche aux moindres carrés : on recherche la droite $\Delta : y = a_0 + a_1 x$ qui minimise la distance (cf. Fig 5.12) :

$$d_1^2 = \sum_i^N [y_i - (a_0 + a_1 x_i)]^2 \quad (5.1)$$

La solution est donnée par : $A = X^\# Y = (X^t X)^{-1} X^t Y$, où $X^\#$ dénote la matrice pseudo-inverse de la matrice X , où X , Y et A sont donnés par :

$$X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_N \end{bmatrix}$$

$$Y = [y_1, y_2, \dots, y_N]^t$$

$$A = [a_0, a_1]^t$$

Ces formules s'étendent très aisément aux espaces de dimensions supérieures, ainsi qu'aux approximations par des polynômes d'ordre plus élevé. la distance minimisée est celle mesurée selon le seul axe y . C'est donc une mesure généralement mal adaptée en traitement d'image, puisque x et y jouent habituellement un rôle équivalent. On lui préfère donc les méthodes par axe d'inertie.

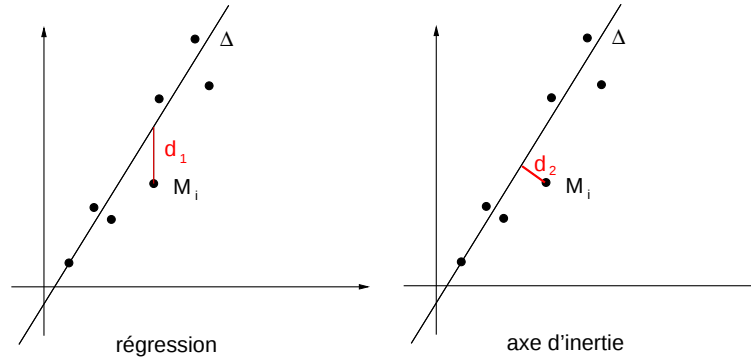


FIG. 5.12 – Deux approximations linéaires aux moindres carrés minimisant des distances différentes : à gauche par régression linéaire, à droite par axe d'inertie.

Approximation par axe principal d'inertie

C'est aussi une approche aux moindres carrés, mais on minimise dans ce cas la somme des distances de tous les points à la droite Δ :

$$d_2^2 = \sum_{i=1}^N \frac{[(a_0 + a_1 x_i - y_i)]^2}{a_1^2 + 1} \quad (5.2)$$

C'est l'équation de l'axe d'inertie des points, qui passe par leur centre de gravité X_g, Y_g et qui est donné comme vecteur propre de plus grande valeur propre de la matrice de forme quadratique :

$$S = \sum_N V_i V_i^t = \begin{bmatrix} \sum x_i x_i & \sum x_i y_i \\ \sum x_i y_i & \sum y_i y_i \end{bmatrix}$$

Ces équations s'écrivent sans problème en dimensions supérieures (pour estimer des variétés d'ordre variable). Elles se transcrivent beaucoup plus difficilement à des polynômes d'ordre plus grand car on ne sait pas, en règle générale, exprimer la distance d'un point courant à une telle fonction.

Estimations robustes

Si l'on repose le problème précédent dans le cadre plus général de l'estimation robuste, on est amené à considérer ce problème de façon un peu plus complète [Meer et al., 1991, Rousseeuw et Leroy, 1987]. On cherche ici la meilleure droite représentant au mieux l'ensemble des points M_i sous l'hypothèse d'un bruit entachant la position des points.

Lorsque le bruit est gaussien, les estimations aux moindres carrés nous assurent d'une bonne qualité de l'estimée. Mais il faut souvent tenir compte de la présence, parmi les points, de détections erronées dont les écarts à la droite ne sont pas gaussiens (ce sont des *outliers*⁴).

Une première façon de prendre en compte ces points consiste à faire une première estimation aux moindres carrés, puis à éliminer les points trop éloignés de la droite (par exemple à l'aide d'un test statistique tel celui de Cramer-von Mises⁵). Ces méthodes ont reçu une base théorique dans l'approche dite par **moindres carrés tronqués** (LTS : *Least Trimmed Squares*) ainsi que toutes ses formes dérivées : moindres quantiles carrés (LQS), moindre k^e carré (LKS) [Lee et al., 1998].

⁴on appelle *outlier* un point qui n'est pas régi par la même distribution statistique que les points dont on recherche une approximation.

⁵Le test de Cramer-von Mises calcule une distance entre la loi gaussienne théorique et la répartition expérimentale des erreurs. Un point est rejeté si sa distance à la distribution des autres points est trop improbable sous l'hypothèse gaussienne [Saporta, 1990].

Mais cette approche est réputée dangereuse (on peut éliminer successivement les points appartenant à la droite si un *outlier* est trop éloigné de celle-ci). On lui préfère des techniques d'estimation par médiane ou par médianes itérées, plus coûteuses en temps de calcul, mais capables d'éliminer beaucoup plus de points erronés.

Dans une estimation par médiane, on choisit 2 points quelconques M_i et M_j de l'ensemble. On leur associe une droite qui fournit des coefficients a_{ij} et b_{ij} . Par combinatoire sur l'ensemble des points, on obtient deux ensembles de coefficients : $a = \{a_{ij}\}$ et $b = \{b_{ij}\}$. Les ensembles a et b sont triés séparément et l'on choisit les valeurs médianes $\bar{a}$ et $\bar{b}$ comme estimateurs des paramètres de la droite. Le filtrage médian peut tolérer jusqu'à 50% - ϵ d'*outliers*. Dans le cas de l'estimation simultanée de deux variables, l'estimation par médiane permet de tolérer jusqu'à 25% d'*outliers* ($0, 5^2$).

De nombreux estimateurs robustes ont été calculés à partir de la notion de médiane. Ainsi, le minimum des médianes des distances quadratiques (LMS *Least Median of Squares*) détermine le centre d'un nuage de points en calculant pour chaque point la distance à tous les autres, puis en caractérisant la distance de ce point au nuage par la valeur moyenne des distances et enfin en choisissant comme centre du nuage celui dont la médiane est minimale [Rousseeuw et Leroy, 1987].

D'autres approches robustes reviennent à une estimation par filtrage (comme le font la moyenne ou la régression), mais en pondérant chaque point i par un coefficient $\rho(x_i)$ qui tend vers 0 pour des points loin de la solution. Une autre approche consiste à intégrer cette pondération des points lointains directement dans la distance (c'est le cas des M-estimateurs [Huber, 1981]). Dans ce cas, les « distances »⁶ croissent moins vite que la loi quadratique ou décroissent même pour tendre vers 0 pour les grandes distances. On les appelle des M-estimateurs.

Une autre famille d'estimateurs robustes s'appuie sur des recherches systématiques. c'est par exemple le cas de la méthode RanSac (*Random Sample Consensus*) [Fischler et Bolles, 1981]. Cette méthode consiste à choisir 2 points et à adopter la droite qui les joint comme approximation. On mesure alors le nombre de points qui sont en accord avec cette hypothèse (par exemple à l'aide d'un critère de distance du point à la droite). On choisit finalement le couple de points qui conduit au plus fort consensus. On montre qu'il n'est pas nécessaire de tester tous les couples mais qu'un petit nombre choisis aléatoirement (de l'ordre de quelques pourcents) est usuellement suffisant.

Estimation d'un mélange de droites

Si l'on sait que n droites existent dans l'ensemble des points, il est possible de faire une classification au sens des nuées dynamiques par exemple. Pour cela on initialise le processus en choisissant n droites (issues par exemple de l'ensemble des points par tirage aléatoire) représentées par leurs paramètres $\{a_0^i, a_1^i, i = 1, \dots, n\}$. On attribue chaque point à la droite qui minimise la distance 5.1 ou mieux 5.2. Après classification de tous les points on estime pour chaque droite ses paramètres $\{a_0^i, a_1^i\}$ par l'une des méthodes vues ci-dessus. On itère le processus enchaînant classification et estimation. On converge vers une solution qui dépend assez fortement de l'initialisation.

Une autre famille de méthodes utilise la version floue des nuées dynamiques (les C-moyennes-floues ou *Fuzzy-C-means*), dans laquelle chaque point k « appartient » à l'une des droites i avec une appartenance $\mu_i(k)$ fonction de sa distance à la droite. Pour tout point k , on vérifie : $\sum_{i=1}^n \mu_i(k) = 1$ [Bezdek, 1981].

Extension de l'estimation à des coniques

Dans le cas où l'on recherche des formes représentées par des cercles ou des ellipses, on peut choisir une paramétrisation matricielle de la forme quadratique les représentant. Chaque conique indexée par i est alors décrite par son centre q_i , son rayon r_i et sa matrice d'ellipticité A_i :

$$[(\mathbf{x} - q_i)^t A_i (\mathbf{x} - q_i)] = r_i^2$$

⁶Les distances ainsi définies ne vérifient en fait plus les propriétés d'une distance.

où $\mathbf{x}$ est le point courant du plan. La distance d'un point $\mathbf{x}_k$ à cette conique s'exprime par :

$$d_i^2(k) = \left[[(\mathbf{x}_k - \mathbf{q}_i)^t \mathbf{A}_i (\mathbf{x}_k - \mathbf{q}_i)]^{1/2} - r_i \right]^2$$

L'estimation des paramètres se fait de la même façon que précédemment à partir d'une conique initiale souvent circulaire.

5.6.2 Approximations polygonales, simplification de contours polygonaux

Lorsque l'on dispose d'une courbe continue ou finement échantillonnée et que l'on souhaite la réduire à une ligne polygonale, on dispose de très nombreux algorithmes qui proposent soit des critères d'approximation soit des mises en œuvre différents. Le plus connu est l'algorithme de la corde.

Algorithme de la corde (ou de Ramer)

C'est un processus de subdivision qui peut être entrepris soit de façon récursive (chaque segment créé fait l'objet d'une nouvelle subdivision), soit de façon itérative (la courbe Γ est considérée globalement à chaque étape). C'est cette dernière version que nous examinons sur l'exemple de la figure 5.13. Les sommets du polygone sont choisis successivement comme les points de Γ les plus éloignés des cordes précédemment tirées. Le processus s'arrête lorsque la nouvelle distance candidate est inférieure à un seuil ϵ fixé. Très employé dans de nombreuses applications par la simplicité de sa mise en œuvre, l'algorithme de la corde n'est pas très rapide. Il ne garantit pas non plus une convergence uniforme vers la courbe finale car, dans son implémentation itérative, il se peut que la distance à l'étape n soit supérieure à celle à l'étape $n - 1$.

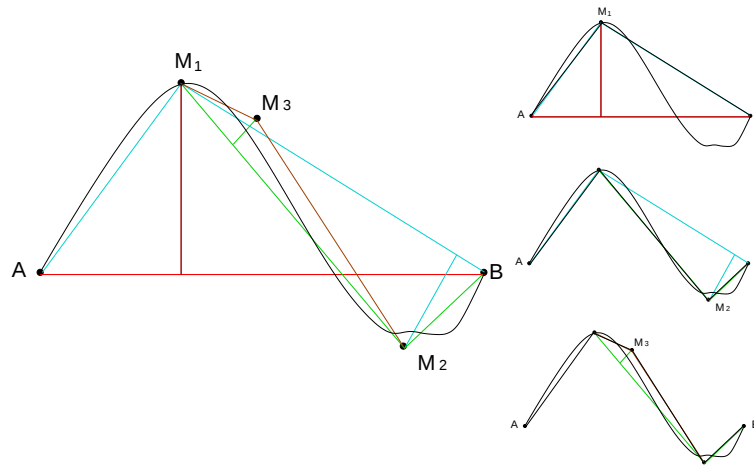


FIG. 5.13 – Algorithme de la corde. Les points M_1 , M_2 et M_3 sont successivement sélectionnés pour créer la ligne polygonale représentant la courbe AB . Le critère de sélection est la distance maximale à la corde polygonale précédemment obtenue.

Algorithme de Dunham

Dans cette méthode on cherche à supprimer les cordes successives qui sont presque alignées. On se fixe donc une tolérance angulaire ϵ . Partant de A , on couvrira par une même corde tous les points de la courbe qui se

trouvent dans la tolérance et l'on choisira le premier point du polygone comme le dernier point de cet ensemble. Le processus est répété en ce point.

Algorithme de Wahl et Danielsson

Son objectif est de minimiser l'aire laissée entre la courbe et la ligne polygonale.

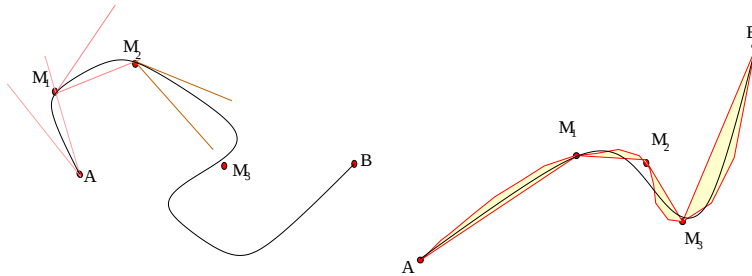


FIG. 5.14 – Algorithme de Dunham, à gauche, les points M_1 , M_2 et M_3 sont choisis lorsque la tolérance angulaire ϵ est insuffisante à approximer la courbe. A droite, algorithme de Wahl et Danielsson : il cherche un minimum des surfaces laissées entre la ligne polygonale et la courbe.

Algorithme progressif

Enfin, un algorithme progressif utilise le critère de distance de tout point de la courbe au segment d'approximation, mais dans un schéma progressif. Dans ce schéma, partant d'une extrémité, la courbe est parcourue, et le point courant est retenu comme extrémité d'un segment d'approximation s'il est le dernier point visité permettant une approximation de tous les points déjà visités, à une tolérance ϵ donnée.

5.6.3 Approximation par des splines

L'approximation d'un ensemble de points de contours par des polynômes d'ordre plus élevé que les droites est une bonne façon de représenter ce contour [Chalmond, 2000]. Le choix du degré du polynôme se fait généralement de façon à assurer des propriétés de continuité au contour obtenu :

- les segments de droite (degré 1) assurent la continuité du contour,
- les approximations par polynômes d'ordre 2 permettent d'avoir des dérivées continues,
- les polynômes d'ordre 3 permettent d'avoir des courbures continues.

Les fonctions splines sont de bons candidats pour ce type d'approximation. Nous en avons déjà vu l'usage à diverses occasions : au chapitre 1.4.2 lorsque nous avons discuté du rééchantillonnage des images, au chapitre 3.5.2 lorsque nous avons estimé des fonctions régulières lors de segmentation par régions. Deux types de fonctions peuvent être utilisées [Unser et al., 1993a, Unser, 1999] :

1. les **fonctions interpolantes** qui passent exactement par tous les points de contours qui sont utilisés pour calculer la spline ;
2. les **fonctions approximantes** qui ne passent pas nécessairement par les points mais qui s'en approchent de façon contrôlée.

Approximation

Dans le cas des fonctions approximantes, on définit la spline d'ordre k par morceaux en fonction d'une variable continue u , et des m points de contrôle P_i qui la déterminent. Q_i et P_i sont des fonctions de même dimension : des vecteurs de $\mathbb{R}^2$ si les points P_i sont définis par leurs coordonnées du plan ou des vecteurs de $\mathbb{R}^3$ dans l'espace.

Le morceau i est défini par :

$$Q_i(u) = \sum_{r=0}^{k-1} P_{i+r} b_r^k(u) \quad i = 0, \dots, m - k + 1 \quad (5.3)$$

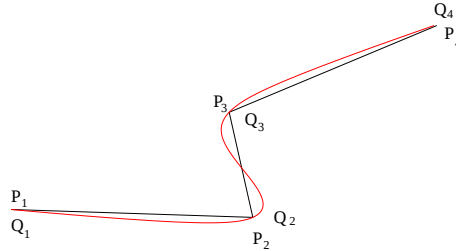


FIG. 5.15 – Dans le cas d'une interpolation spline, les jonctions entre les morceaux de splines sont les points de contrôle eux-mêmes.

Les fonctions $b_r^k(u)$ sont définies à partir des fonctions $\beta^k(u)$ qui ont été introduites au chapitre 1.4.2 équations 1.32 et suivantes :

$$b_r^k(u) = \beta^k \left[u - \frac{r-1}{2} \right]$$

Les positions des points de raccord entre splines sont données par $u = 0$ dans l'équation 5.3 :

$$Q_i(0) = \sum_{r=0}^{k-1} P_{i+r} b_r^k(0) \quad (5.4)$$

Il y a de nombreuses façons de définir les variables u . La plus naturelle consiste à considérer que les échantillons sont régulièrement distribués en fonction de u (attention cela conduit à une paramétrisation où u n'est pas l'abscisse curviligne le long du contour), alors $b_0^k(0) = 0$ et $b_{k-1}^k(1) = 0$, on peut réécrire cette équation [Goshtasby et al., 1990] sous la forme :

$$Q_i = \sum_{r=0}^{k-2} P_{i+r} b_r^k(0)$$

qui prend la forme d'une convolution : $Q = P * H_k$. Ceci montre que les Q sont des points filtrés issus des P .

Interpolation

Dans ce cas, les coefficients des splines de l'équation 5.3 sont inconnus et l'équation devient :

$$Q_i(u) = \sum_{r=0}^{k-1} c_{i+r} b_r^k(u) \quad i = 0, \dots, m - k + 1$$

et les coefficients inconnus c_i sont déterminés en contraignant les splines à passer par les points de contrôle P_i :

$$P_i(0) = \sum_{r=0}^{k-1} c_{i+r} b_r^k(0)$$

dont on a vu au chapitre 1.4.2 les divers modes de résolution soit par inversion de matrice soit par filtrage (cf. figure 5.15).

5.7 Transformation de Hough

La transformation de Hough est un outil de traitement des images dont l'usage déborde largement l'application pour laquelle nous la présentons maintenant [Ballard, 1978, Sklansky, 1978, Maître, 1985]. Ces applications seront vues plus tard. Nous nous intéressons actuellement au sous-problème suivant :

Connaissant un ensemble de pixels P (potentiellement bruité) appartenant à une frontière que l'on sait polygonale, comment déterminer le nombre de segments impliqués dans cette ligne polygonale et leur position ?

Exprimé dans les termes que nous avons vus précédemment, on se retrouve devant un difficile problème de **mélange**. Les solutions combinatoires (où l'on recherche systématiquement les classes possibles) sont généralement inacceptables. La transformation de Hough nous propose au contraire une solution élégante.

Elle consiste à transformer un problème inconnu (retrouver des droites) en un problème mieux connu : retrouver des nuages de points. Ceci s'obtient en associant à l'espace de l'image (dénnoté $\mathcal{I}$ et défini par ses variables d'espaces $\{x, y\}$), un espace de paramètres, dénoté par $\mathcal{H}$ (dans le cas où l'on recherche des droites représentées par l'équation $y = ax + b$, on a $\mathcal{H} = \{a, b\}$). Pour faire cette transformation il y a 2 façons : la transformation de Hough de 1 à m et la transformation de Hough de m à 1 .

5.7.1 Définitions

Transformation de 1 à m

Dans cette définition, on associe à tout point $M_i = (x_i, y_i)$ de P toutes les droites du plan. Elles sont définies dans $\mathcal{H}$ par : $b = -x_i a + y_i$, ce qui est l'équation d'une droite. A un point de $\mathcal{I}$ est associée une droite de $\mathcal{H}$ (d'où le nom de 1 à m : à 1 point sont associés m points, *many* en anglais) et à un point de $\mathcal{H}$ une droite de $\mathcal{I}$.

Lorsque l'on transforme tous les points de P par la transformation, on associe à $\mathcal{I}$ un ensemble de droites qui, idéalement, se coupent en des points H_k . Le nombre de transformations à faire est $\nu = N = \text{Card}(P)$. Les transformés des H_k sont les droites cherchées dans $\mathcal{I}$. Ces points H_k s'obtiennent aisément en ne conservant que les intersections des droites de $\mathcal{H}$ et en recherchant les nuages d'intersections (figure 5.16).

Transformation de m à 1

Dans cette définition, on choisit d'associer à tout bi-point $\{M_i, M_j\}$, une droite Δ_{ij} de $\mathcal{I}$ qui se transforme en un unique point⁷ Q_{ij} de $\mathcal{H}$. Les coordonnées a_{ij} et b_{ij} du point Q_{ij} se déduisent immédiatement de l'équation de Δ_{ij} . En combinant tous les couples de points de P , on obtient tous les points Q de $\mathcal{H}$. Le nombre de transformations à accomplir est $\nu' = \frac{1}{2}N(N-1)$. L'espace de Hough est alors constitué de ν' points que l'on identifie immédiatement aux intersections des ν droites obtenues par la transformation de 1 à m .

On se ramène donc, comme précédemment, à la recherche des nuages de points les plus denses du plan $\mathcal{H}$.

⁷la dénomination « m à 1 » se traduirait alors précisément par « 2 à 1 », attention, m n'a pas la même signification dans les 2 définitions.

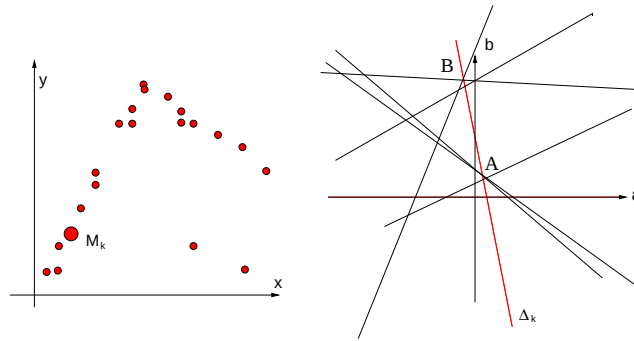


FIG. 5.16 – Transformation de Hough de 1 à m . Un point M_k est transformé en une droite Δ_k de l'espace des paramètres $\{a, b\}$. On a représenté quelques-unes seulement de toutes les droites existantes. Elles définiront finalement 2 points d'intersection, l'un en A, l'autre en B.

Passage d'une TH à l'autre

Il est assez simple de montrer que l'accumulateur de la TH de m à 1 se déduit de celui de la TH de 1 à m par un seuillage au niveau 1. Ces deux transformations sont donc tout à fait équivalentes dans leurs résultats.

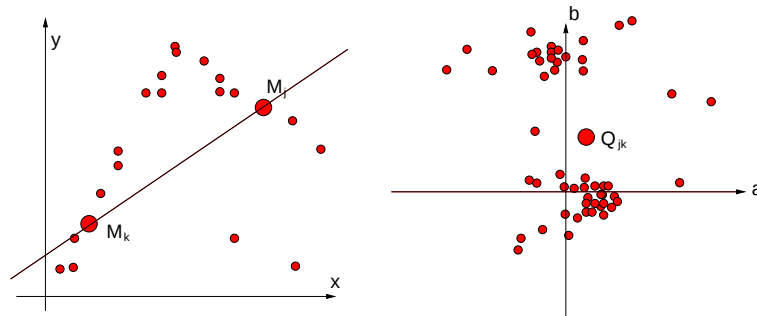


FIG. 5.17 – Transformation de Hough de m à 1. Deux points M_j, M_k définissent une droite qui se transforme en 1 point Q_{jk} de l'espace des paramètres (a, b) .

Généralisation

On appellera transformation de Hough associée à une forme paramétrée par les variables $\{a_i\}$ la transformation qui fait passer de l'espace image $\mathcal{I}$ à l'espace des paramètres $\mathcal{H}$.

La transformation est de 1 à m si on associe à un point de $\mathcal{I}$ la variété de $\mathcal{H}$ décrivant toutes les formes passant par ce point. La transformation est de m à 1 si l'on associe à toutes les combinaisons m points de $\mathcal{I}$ une seule forme (donc un point unique de $\mathcal{H}$).

Il existe généralement de nombreuses transformations de n à m ($n < m$), où l'on se place dans une situation intermédiaire : en associant n points de $\mathcal{I}$ on limite le sous-espace d'arrivée à une partie de la variété de la transformation de 1 à m . Par exemple, on peut utiliser simultanément la position d'un point de contour et la direction de la tangente en ce point. Ces problèmes n'étant pas directement liés à la recherche de contours, des exemples seront vus plus loin.

TH et filtrage adapté

Sklansky [Sklansky, 1978] a démontré qu'il y a équivalence entre la TH et la détection par filtrage adapté (par corrélation) dans une situation particulière mais fréquente. Si les seules inconnues d'un problème de détection sont les deux translations du plan (δx et δy), c'est à dire si la forme ne subit ni changement d'échelle ni rotation, ni transformation perspective. Soit $f(x, y)$ la forme caractéristique de l'objet disponible ($f(x, y) = 1$ si (x, y) appartient au contour de l'objet, et 0 sinon). Soit $g(x, y)$ la forme recherchée. Le filtrage adapté conduit à une détection optimale au sens de l'erreur quadratique moyenne, en recherchant les maximums de la fonction de corrélation :

$$C_{fg}(\delta x, \delta y) = \iint f(x, y)g(x - \delta x, y - \delta y)dxdy = f(x, y) * g(-x, -y)$$

Cette démonstration est assez simple si l'on voit que chaque point M de coordonnées (x, y) de l'image f est candidat à être associé à un point quelconque de g de coordonnées (u, v) si l'on décale g du vecteur $x - u, y - v$ (cf. figure 5.18). Cette contribution à l'accumulateur de position $\delta x = x - u, \delta y = y - v$ vaut exactement : $f(x, y)g(u, v)$. Sommé sur toute la courbe f on obtient :

$$\begin{aligned} \gamma(u, v) = \gamma(-\delta x, -\delta y) &= \iint f(x, y)g(x - \delta x, y - \delta y)dxdy \\ &= f(x, y) * g(-x, -y) \end{aligned}$$

Ce qui est bien le résultat cherché. En fait on montre que la TH agit en pratique en inversant l'ordre des intégrales par rapport au filtrage adapté. Celui-ci procède usuellement en calculant l'intégrale ci-dessus pour un décalage u, v donné, en sommant sur tous les points du plan tandis que la TH choisit un point du plan et incrémente toutes les intégrales (les compteurs) qui lui sont associées par un déplacement u, v quelconque. Cette inversion des opérations peut permettre de tirer profit d'a priori importants pour accélérer les calculs.

Grâce à cette propriété la TH peut hériter des nombreux résultats obtenus en filtrage adapté : pour prendre en compte le rôle du bruit si celui-ci à une densité de probabilité connue, pour déterminer la taille des accumulateurs lorsque la PSF est connue, etc. [Maître, 1985, Princen et al., 1992].

5.7.2 Mise en œuvre de la TH

Discretisation de $\mathcal{H}$

Dans la pratique le plan $\mathcal{H}$ est discrétisé. Chaque cellule de position a_i, b_i de largeur $\delta a, \delta b$ est appelée accumulateur. Elle contiendra un nombre d'autant plus important que la droite $y = a_i x + b_i$ sera plus probable dans $\mathcal{I}$.

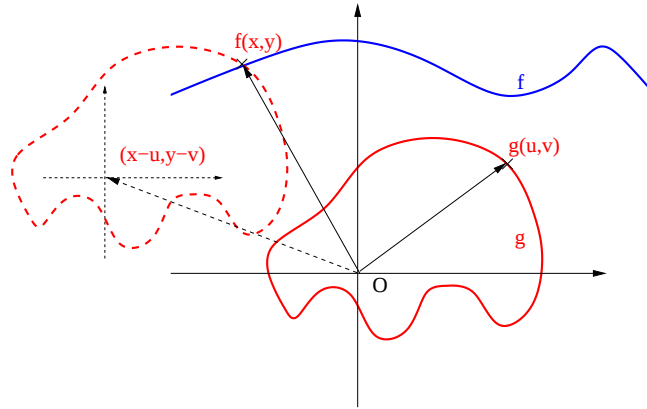


FIG. 5.18 – Transformation de Hough : le point de coordonnées x, y de f sera associé au point $g(u, v)$ de g si l'on fait subir à cette courbe g une translation de $(x - u, y - v)$.

Au début de la transformation, tous les accumulateurs sont à zéro. Les points sont alors visités successivement (individuellement dans la TH de 1 à m , combinatoirement pour les autres). Chaque hypothèse donne naissance à des votes pour des cellules particulières. Pour chaque vote, l'accumulateur de la cellule est incrémenté de 1.

Lorsque tous les candidats ont voté, on recherche les accumulateurs de $\mathcal{H}$ de comptes localement maximaux. Le choix de la taille des accumulateurs est donc un problème délicat.

1. Des accumulateurs trop petits recevront très peu de votes et donneront des statistiques peu significatives. On ne saura décider de la position d'un maximum entre plusieurs cellules voisines de comptes faibles et proches.
2. Des accumulateurs trop grands donneront une mauvaise précision dans la détection de la droite.

Par ailleurs, la mise en œuvre de la TH nécessite un espace mémoire de taille $\frac{\Delta a \Delta b}{\delta a \delta b}$ où Δa et Δb représentent les domaines de variation de a et b .

Des solutions ont donc été proposées pour adapter localement la taille des accumulateurs (très précis près des lieux où les votes sont nombreux). Cela peut se faire dynamiquement par des approches de subdivision récursive des cellules (selon le principe des **quad-tree** cf. 3.2.4), ou par des passages successifs sur l'ensemble des points de P , le premier passage permettant de dimensionner les cellules.

Ce problème est particulièrement important dans le cas où l'on recherche des formes possédant de nombreux paramètres, puisque l'espace $\mathcal{H}$ est l'espace produit de tous ces paramètres. Il est alors de grande dimension.

Paramétrisation des formes

Le choix d'une bonne paramétrisation des formes recherchées apparaît immédiatement à la lecture du paragraphe ci-dessus si l'on souhaite appliquer la démarche proposée à la recherche de droites. Dans le cas d'une paramétrisation cartésienne ($y = ax + b$), les domaines de variation de a et b sont potentiellement de moins l'infini à plus l'infini, conduisant à un insoluble problème de résolution de cellule. Par ailleurs, si l'on examine la variable a par exemple, on voit que, statistiquement, elle décrit 25% des droites du plan pour $a \in [0, 1[$ et 25% pour $a \in]1, \infty[$, conduisant donc à des cellules qui seront très inégalement remplies si leur taille est uniforme.

On peut cependant choisir une paramétrisation plus heureuse pour la droite, par exemple la paramétrisation normale (cf. figure 5.19). Dans ce cas la droite $y = ax + b$ est repérée par la distance ρ et l'angle θ .

La transformation de 1 à m s'écrit : $\rho = x_i \cos \theta + y_i \sin \theta$, et la forme associée à un point est donc une sinusoïde dans l'espace $\rho = f(\theta)$, d'équation :

$$\rho = \sqrt{x_i^2 + y_i^2} \cos(\theta + \phi)$$

avec : $\cos(\phi) = x_i / \sqrt{x_i^2 + y_i^2}$.

Cette paramétrisation est tout d'abord homogène puisque les angles θ sont a priori équiprobables, ainsi que les distances ρ . Elle fournit également des domaines de variation finis : $\Delta\theta = [-\pi, +\pi[$ et $\Delta\rho = [0, \sqrt{2}L]$ si L est le côté de l'image, puisqu'une droite dont ρ serait supérieur à la longueur de la diagonale de l'image ne serait pas visible dans l'image.

La transformation de m à 1 associée au doublet M_i, M_j le point de $\mathcal{H}$ de coordonnées :

$$\begin{aligned} \rho_{ij} &= \frac{|x_i y_j - x_j y_i|}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}} \\ \theta_{ij} &= -\text{Arctg} \left[\frac{x_j - x_i}{y_j - y_i} \right] \end{aligned}$$

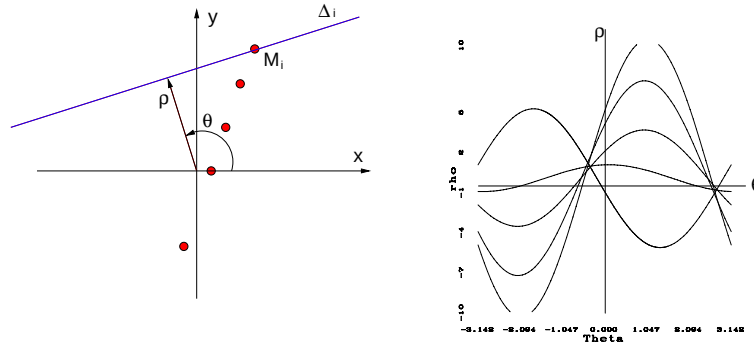


FIG. 5.19 – Transformation de Hough de 1 à m en coordonnées normales. Une droite Δ_i est repérée par l'angle θ et la distance au pied de sa normale ρ . Dans l'espace $\mathcal{H}$, chaque point donne une sinusoïde.

5.7.3 Détections par TH

La TH est très souvent utilisée, comme nous l'avons vu précédemment, pour détecter des droites. Les autres courbes paramétrées détectées par la TH sont :

1. les paraboles (3 paramètres si l'on connaît la direction de leur axe, 5 dans le cas général). Par exemple pour détecter des côtes dans les radiographies du corps humain);

2. les cercles (3 paramètres) par exemple pour détecter des cuves ou des citernes en imagerie aérienne. On paramètre alors le cercle par son centre et son rayon et la forme associée dans cet espace est un cône, centré au centre du cercle.
3. les ellipses (5 paramètres : les 3 du cercle + la direction de l'axe et l'aplatissement ou l'ellipticité), également utilisée pour détecter des formes circulaires vues en perspectives.
4. des traces sinusoïdales (5 paramètres), forme fréquemment recherchée en expérimentation en physique ou en géologie.

La TH est également utilisée pour suivre des formes quelconques sous des hypothèses de transformations particulières. Les paramètres sont alors ceux qui caractérisent la déformation : changement d'échelle, rotation à 2 ou 3 D, projection perspective, etc.

5.8 Conclusion

Il n'est pas possible d'épuiser toutes les variétés des représentations et des modélisations qui ont été adoptées pour décrire les formes des objets en traitement des images. en effet, beaucoup de ces représentations tirent profit de spécificités attachées à la famille particulière des objets à traiter pour une application donnée dans un contexte précis. Nous avons cependant vu que toutes les représentations cherchent à marier la simplicité de la description et la fidélité à la forme. En cela, elles relèvent déjà toutes de la reconnaissance des formes, puisque ces deux objectifs sont aussi ceux que la reconnaissance des formes se donne.

