

Le traitement des images

tome 2

I. Bloch, Y. Gousseau, H. Maître, D. Matignon,

B. Pesquet-Popescu, F. Schmitt, M. Sigelle, F. Tupin

Polycopié du cours ANIM
Département TSI - Télécom-Paris

version 5.0

16 septembre 2004

Table des matières

1	Les prétraitements	1
1.1	Les traitements photométriques ou colorimétriques	1
1.1.1	Linéarité	1
1.1.2	Homogénéité	1
1.1.3	Binarisation, seuillage	3
1.1.4	Augmentation de contraste	3
1.2	Suppression des bruits	3
1.2.1	Filtrages linéaires	3
1.2.2	Filtrage de rang	5
1.2.3	Filtrages morphologiques	5
1.2.4	Filtrages par équations de diffusion	5
1.3	Les filtres adaptatifs	8
1.3.1	Les filtres à coefficients adaptatifs	8
1.3.2	Les filtres à fenêtres adaptatives	9
1.4	Le rééchantillonnage des images	10
1.4.1	Interpolation avec des polynômes	10
1.4.2	Interpolation par des B-splines	11
1.4.3	Interpolation adaptative	13
2	La détection des contours dans les images	15
2.1	Un modèle continu de contour.	16
2.2	Les approches classiques	17
2.2.1	Les détecteurs de gradient par filtrage	17
2.2.2	Les détecteurs de gradient par masques	18
2.2.3	Pré- et post traitements	19
2.2.4	Les détecteurs de passage par zéro du laplacien	20
2.2.5	Les détecteurs par masquage adapté	21
2.3	Les approches analytiques	22
2.3.1	Les critères de Canny	22
2.3.2	Les filtres de Deriche et Shen et Castan	23

2.3.3	L'extension à 2D	24
2.3.4	Variantes autour des filtres analytiques	25
2.4	Les contours actifs = les <i>snakes</i>	26
2.4.1	Une approche physique des contours	26
2.4.2	Mise en équation des contours actifs	27
2.4.3	Les ensembles de niveaux (<i>level sets</i>)	29
2.4.4	Les modèles géodésiques actifs	30
2.5	La poursuite et la fermeture des contours	30
2.5.1	Méthodes de recherche dans des graphes	30
2.5.2	Les automates	32
3	La segmentation par régions	35
3.1	Les méthodes sur histogramme	35
3.1.1	Avec apprentissage	35
3.1.2	Seuillage sans apprentissage	36
3.1.3	Méthodes de classification	37
3.1.4	Sélection sur l'histogramme et dans l'image	37
3.1.5	Sélection sur histogramme et régularisation	37
3.2	Les méthodes par transformation de régions	39
3.2.1	La croissance de région	40
3.2.2	Le partage de région	40
3.2.3	La réunion de région	41
3.2.4	Les pyramides	42
3.3	Les graphes d'adjacence	42
3.4	La méthode MDL = Minimum Description Length	43
3.5	L'approche de Mumford et Shah	45
3.5.1	Approche formelle	45
3.5.2	Les variantes autour de la formulation de Mumford-Shah	46
4	Les textures	49
4.1	Qu'est ce qu'une texture ?	49
4.1.1	Distribution aléatoire ou régulière ?	50
4.1.2	Les échelles des textures	51
4.1.3	Analyse ou synthèse	51
4.2	Modèles de texture	51
4.2.1	Un modèle biologiquement plausible	51
4.2.2	Modèles stochastiques	52
4.3	Analyse et reconnaissance de textures	52
4.3.1	Schéma général	52
4.3.2	Approches par mesures statistiques	53
4.3.3	La détection d'énergie dans des canaux fréquentiels	57

4.3.4	Les filtres optimisés	58
4.3.5	Les modélisations autorégressives	59
4.4	Les approches par champs markoviens	59
4.4.1	La méthode de Manjunath et Chellappa	60
4.4.2	La méthode de Kervrann et Heitz	61
4.5	Les méthodes structurales	62
4.6	Textures hétérogènes	62
5	Description de contours et de formes	63
5.1	Fonction caractéristique	63
5.2	Description de formes	65
5.2.1	Représentation par les moments	65
5.2.2	Représentation par les moments invariants	65
5.2.3	Boîtes englobantes et boîtes minimales	66
5.3	Polygones de Guzman	66
5.4	Chaînes de Freeman	67
5.4.1	Définition	68
5.4.2	Les propriétés des chaînes de Freeman	69
5.4.3	Reconnaissance des formes par des chaînes de Freeman	71
5.5	Descripteurs de Fourier	72
5.5.1	Descripteur par tangente	72
5.5.2	Représentation complexe	72
5.6	Approximations polynomiales	73
5.6.1	Approximation d'un nuage de points par une droite unique	74
5.6.2	Approximations polygonales, simplification de contours polygonaux	77
5.6.3	Approximation par des splines	78
5.7	Transformation de Hough	80
5.7.1	Définitions	80
5.7.2	Mise en œuvre de la TH	82
5.7.3	Détections par TH	84
5.8	Conclusion	85
6	Éléments de géométrie tridimensionnelle	87
6.1	Géométrie projective	87
6.1.1	Espace projectif et coordonnées homogènes	87
6.1.2	Espaces projectifs \mathcal{P}^1 , \mathcal{P}^2 et \mathcal{P}^3	88
6.1.3	Expression linéaire de problèmes géométriques	89
6.1.4	Stratification, groupes de transformations et invariants	90
6.2	Représentations des rotations de \mathbf{R}^3	92
6.2.1	Les angles d'Euler	92
6.2.2	Exponentielle d'une matrice anti-symétrique	93

6.2.3	Forme canonique et théorème de Rodriguez	94
6.2.4	Quaternions	94
6.3	Géométrie différentielle	97
6.3.1	Courbes, repère de Frenet	97
6.3.2	Courbes sur une surface, formes fondamentales	98
6.3.3	Courbures sur des surfaces	100
6.3.4	Une notation fréquente	101
6.3.5	Interprétations géométriques	101
6.3.6	Principales méthodes de calcul des courbures	103
6.3.7	Géométrie différentielle dans des volumes numériques	105
7	Formation de l'image et couleur	107
7.1	Formation des images	107
7.1.1	Imagerie cohérente et incohérente	107
7.1.2	Interaction matière rayonnement	109
7.1.3	Les sources de lumière	112
7.2	Définition des différents espaces couleur	112
7.2.1	Trivariance et primaires	112
7.2.2	Choix d'un espace de couleur	113
7.2.3	L'espace RVB de la CIE 1931	114
7.2.4	L'espace XYZ	116
7.2.5	L'espace Lab et le repère TLC	117
7.2.6	L'espace YC1C2	119
7.2.7	L'espace TLS	119
7.2.8	Le système visuel hypothétique	120
7.2.9	L'espace TLC	122
7.2.10	Les espaces de la télévision	122
7.3	Modélisation de la perception colorée des images	123
7.4	Les traitements des images en couleur	125
7.4.1	Un ordre pour l'espace R^3 ?	125
7.4.2	Qu'est ce qu'une couleur représentative d'un nuage ?	127
7.4.3	Comment calculer une palette ?	127
7.4.4	Comment filtrer 2 couleurs ?	128
7.4.5	La détection des contours des images en couleur	128
7.4.6	Représentation par quaternions	129
7.4.7	Couleur et quaternion	129
A	Décomposition en Valeurs Singulières : SVD	131
A.1	Définition	131
A.2	Lien avec le rang et interprétation de la décomposition	131
A.3	Normes	132

B Applications de la SVD à la résolution des systèmes linéaires	133
B.1 Conditionnement d'un système linéaire non singulier	133
B.2 Pseudo-inverse	133
B.3 Solutions robustes d'un système linéaire, moindres carrés	134
C Optimisation sous contraintes	137
C.1 Contraintes égalités	137
C.2 Contraintes inégalités	137

Chapitre 1

Les prétraitements

Chapitre rédigé par Henri MAÎTRE

On regroupe souvent sous le terme de prétraitement toutes les opérations qui sont appliquées aux images, indépendamment de leur usage futur, pour leur assurer une bonne qualité. Elles concernent donc essentiellement les corrections de contraste et la suppression du bruit. Cette pratique est souvent critiquée car l'expérience montre qu'il est toujours important d'adapter les traitements de plus bas niveau aux traitements plus élaborés et donc aux objectifs à long terme du projet, mais les prétraitements sont souvent appliqués dans des circonstances où l'on destine les images à de nombreuses applications différentes dont on ignore souvent les besoins exacts.

1.1 Les traitements photométriques ou colorimétriques

1.1.1 Linéarité

Le défaut de non-linéarité apparaît si la valeur du signal d'image en un pixel n'est pas proportionnelle à l'énergie qu'il reçoit. La linéarité des capteurs d'image est très souvent médiocre. C'est particulièrement vrai pour les films et pour les tubes vidéo. Les capteurs solides (CCD par exemple) leur sont sur ce point très supérieurs [Marion, 1997]. On sait pourtant que cette linéarité est indispensable pour un grand nombre de problèmes, en particulier pour la restauration puisque c'est l'un des premiers prérequis des approches, qu'elles soient analytiques ou algébriques (cf. chapitre ??). Corriger une non-linéarité est une opération simple lorsqu'il n'y a pas de saturation du capteur. Cela revient à inverser cette non-linéarité et cela se fait par des tables de transcodage après calibrage du capteur (aussi appelé LUT : *Look-up Tables*). Il n'est pas possible de corriger une saturation sans information supplémentaire sur le signal.

1.1.2 Homogénéité

Ce défaut apparaît lorsque le capteur n'a pas une réponse égale en tous les plans de son champ. Le vignettage [Perez et Heitz, 1991] est un exemple de défaut d'homogénéité frappant les systèmes optiques et affectant les points du champ éloignés du centre de l'image. Mais il peut également provenir d'un éclairage inégal de la surface à analyser.

Avec des capteurs en ligne (comme on en trouve par exemple en télédétection ou sur les scanners de documents), les défauts d'homogénéité apparaissent sous forme de trainées. Les corrections d'homogénéité sont alors pratiquées par apprentissage des éléments de calibrage pour chaque cellule sensible, puis correction en associant une table de transcodage à chaque cellule. Si le capteur est linéaire (au sens de la section 1.1.1), le calibrage revient

à apprendre pour chaque cellule i le gain α_i et l'offset β_i qui fait passer du signal d'entrée $e(i)$ au pixel image $f(i)$:

$$f(i) = \alpha_i e(i) + \beta_i$$

Pour les capteurs bidimensionnels (films, matrices) les corrections rigoureuses sont très difficiles car il n'est pas possible de généraliser l'approche précédente en raison de la dimension beaucoup trop grande du problème et de la très grande difficulté d'en apprendre les trop nombreux paramètres. On préfère alors des approches souvent approchées : méthodes par morceaux (subdivision du champ) ou corrections globales paramétriques (par exemple approximation par des plans ou par des quadriques).

Homogénéisation par subdivision du champ

Une méthode souvent employée, en particulier pour corriger les défauts d'éclairage, procède par subdivision du champ de façon à rechercher des statistiques constantes dans le champ de l'image. Cette méthode fait donc l'hypothèse que la scène observée possède une homogénéité que l'image a perdu.

Les corrections par moyenne constante remplacent l'image $f(x, y)$ par une image corrigée $f'(x, y)$ qui vérifie :

$$f'(x, y) = f(x, y) - m(x, y) + M$$

avec m une moyenne locale calculée dans un voisinage \mathcal{V}_{xy} autour de (x, y) :

$$m(x, y) = \frac{1}{s} \int_{\mathcal{V}_{xy}} f(x', y') dx' dy'$$

et M la moyenne d'ensemble sur l'image complète \mathcal{I} :

$$M = \frac{1}{S} \int_{\mathcal{I}} f(x, y) dx dy$$

La surface de \mathcal{V}_{xy} est dénotée par s et celle de \mathcal{I} par S .

Les corrections par contraste constant créent une image f'' telle que :

$$f''(x, y) = \frac{V}{v} \cdot f(x, y)$$

où v et V sont les variances de f sur \mathcal{V}_{xy} et \mathcal{I} .

On peut imposer à l'image d'avoir une moyenne et un contraste constants, mais la formule n'est plus alors directe.

Homogénéisation par approximation par une quadrique

Une autre façon de rendre homogène une image est d'en soustraire une composante qui serait inégale dans l'image, par exemple en raison d'un éclairage peu uniforme. Si le signal d'intérêt est faible autour de la valeur moyenne du signal, fortement variable, on peut par exemple approcher celle-ci à l'aide d'un polynôme dépendant des variables d'espace. Voici comment on s'y prend lorsque le polynôme est une quadrique $Q(x, y)$ de paramètres c_i inconnus :

$$Q(x, y) = c_1 + c_2 x + c_3 y + c_4 xy + c_5 x^2 + c_6 y^2$$

On recherche le meilleur polynôme en minimisant, par rapport aux c_i , l'erreur ϵ :

$$\epsilon = \int_{\mathcal{I}} [f(x, y) - Q(x, y)]^2 dx dy$$

On obtient un système de taille 6×6 avec pour inconnues les c_i . Il se met sous la forme :

$$AC = M$$

où C est le vecteur des coefficients inconnus c_i , A est la matrice des moments de l'image (matrice symétrique mesurée pour chaque image) et M le vecteur des monômes en x et y .

1.1.3 Binarisation, seuillage

Ce sont des cas extrêmes de prétraitements qui ramènent l'image à deux ou à quelques niveaux seulement. Nous ne les évoquons ici que par souci d'exhaustivité. Ils sont traités plus complètement au chapitre 3 et dans les références [Sahoo et al., 1988, Cham et al., 1998].

1.1.4 Augmentation de contraste

Une famille de traitements a pour objectif de donner à l'image un plus grand contraste. Ces méthodes procèdent de trois façons différentes [Pratt, 1978, Zamperoni, 1996] :

1. Par étirement d'histogramme : on cherche à exploiter au mieux la dynamique de l'histogramme, tout d'abord en utilisant toute l'échelle de gris ou de couleurs disponible (par une table de transcodage), mais aussi en modifiant l'histogramme de façon à le transformer en un histogramme de référence. Les plus utilisés sont les histogrammes plats et hyperboliques (nous avons vu une telle transformation au chapitre ??) [Bovik, 2000].
2. Par filtrage passe-haut de l'image : l'idée de base est ici de réduire l'importance du terme continu et des basses fréquences. Un point délicat de ces méthodes est de définir le gabarit du filtre utilisé.
3. Par des méthodes locales : ces méthodes modifient localement l'histogramme pour conserver toujours une bonne dynamique, même dans des zones de fort contraste (par exemple Dorst). Un exemple de réhaussement de contraste local par des méthodes morphologiques est présenté au chapitre ?? sur la figure ??.

1.2 Suppression des bruits

C'est l'un des sujets les plus délicats du traitement des images. Il a vu couler beaucoup d'encre et de nombreuses méthodes lui ont été consacrées, tout d'abord très intuitives, mais progressivement de plus en plus complexes. Nous verrons tout d'abord les approches linéaires, puis les méthodes non-linéaires.

1.2.1 Filtrages linéaires

Dans de nombreux cas on considère que l'image non dégradée $f(x, y)$ est affectée d'un bruit additif $n(x, y)$:

$$g(x, y) = f(x, y) + n(x, y)$$

Le cas d'un bruit multiplicatif est beaucoup plus difficile. Il est par exemple traité dans le cas de l'imagerie radar [Maître, 2001].

Dans les conditions les plus simples, on suppose le signal stationnaire, et on adopte donc souvent une approche par filtrage linéaire. Intuitivement, et en l'absence de toute autre information, on recherche des filtres passe-bas, éliminant les variations à très haute fréquence du signal.

Expérimentalement ces filtres ne sont pas très bons (cf. figure 1.1) car ils dégradent considérablement les contours et rendent l'image floue, ce qui est subjectivement désagréable. Une seule exception : si le signal est effectivement stationnaire (et donc sans contours) et si le bruit est gaussien (dans ce cas ce sont les meilleurs filtres au sens des moindres carrés et du maximum de vraisemblance). Les plus utilisés sont le filtre de moyenne (sur de petites fenêtres de 3×3 ou 5×5 pixels) et le filtre gaussien.

Complexité des filtres linéaires

La complexité des filtres linéaires est généralement en $\mathcal{O}(\nu^2 N)$, si N est le nombre de pixels dans l'image et ν le côté de la fenêtre carrée sur laquelle est faite le filtrage. Leur coût grandit donc très vite avec l'étendue du filtre et cela prohibe souvent les filtres larges.



FIG. 1.1 – **Filtres de moyenne** - A gauche, image bruitée par un bruit gaussien de variance 120, au centre, filtrage passe-bas sur une fenêtre 3×3 , à droite sur une fenêtre 7×7 .

Il existe cependant des exceptions notables :

- Le filtre de moyenne a une complexité indépendante de la taille de la fenêtre (si l'on garde en mémoire les traitements appliqués aux points et aux lignes précédentes).
- Le filtre gaussien est séparable, il peut être obtenu par 2 filtrages monodimensionnels. Sa complexité est donc $\mathcal{O}(\nu N)$.

On utilise aussi la propriété des filtres gaussiens d'être à noyau auto-reproducteur (on obtient une nouvelle gaussienne d'écart-type fort, par convolution de deux gaussiennes d'écart-type faible) en filtrant successivement par des gaussiennes étroites pour obtenir des gaussiennes larges. On a aussi beaucoup utilisé des approximations splines des fonctions gaussiennes, plus simples à calculer.

Le filtre de Wiener

Dans l'approche de Wiener (cf. chapitre ??, section ??), on recherche le filtre $h(x, y)$ qui minimise :

$$\left\langle \int_{\mathcal{I}} (\hat{f} - f)^2 dx dy \right\rangle \quad (1.10)$$

où l'espérance $\langle . \rangle$ est prise sur toutes les réalisations du bruit, le domaine d'intégration \mathcal{I} est l'image complète et $\hat{f} = f * h$.

Si on connaît le spectre de densité de puissance du bruit $P_n(u, v)$ et le spectre de densité de puissance de l'image $P_f(u, v)$, le filtrage optimal au sens des moindres carrés est le filtre adapté. C'est le filtre de Wiener avec un défaut égal à l'identité. Sa fonction de transfert (TF de sa réponse impulsionnelle) a pour expression :

$$H(u, v) = \left[1 + \frac{P_n(u, v)}{P_f(u, v)} \right]^{-1}$$

Lorsque l'on ne connaît pas ces spectres de densité de puissance, on adopte souvent, pour ces grandeurs, les expressions (cf. chapitre ??) :

$$P_f(u, v) = \frac{1}{1 + \alpha u^2} \frac{1}{1 + \alpha v^2} \quad \text{et} \quad P_n(u, v) = N^2$$

Malheureusement, compte tenu des difficultés à estimer les paramètres exacts du signal et du bruit, ce filtre conduit en pratique à des filtrages de bruit souvent médiocres.

1.2.2 Filtrage de rang

Ces approches reposent sur un raisonnement très différent. On considère que le signal vrai doit être estimé à partir d'échantillons bruités. Ces échantillons appartiennent à une fenêtre \mathcal{V}_{xy} de $n \times n$ pixels entourant le pixel à filtrer. Si la fenêtre est entièrement à l'intérieur d'une zone homogène, on est ramené au problème traité ci-dessus. Si la fenêtre est à cheval sur des champs de statistiques différentes, on doit trouver un **estimateur robuste** qui non seulement écartera le bruit, mais supprimera également l'influence du signal parasite (différence entre « bruit » et « outlyers »). Les filtres d'ordre ou de rang sont de tels filtres [Kotopoulos et al., 1997, Tagare et Figueiredo, 1985, Kim et Yaroslavski, 1986, Pitas et Venetsanopoulos, 1990, Arce et al., 2000].

On trie les valeurs des pixels dans la fenêtre \mathcal{V}_{xy} par amplitude croissante, on remplace le pixel central par un pixel choisi dans la liste ordonnée :

- si on choisit le plus grand : dilatation en morphologie mathématique (cf. chapitre ??),
- si on choisit le plus petit : érosion en morphologie mathématique,
- si on choisit celui du milieu : filtre médian.

Les avantages des filtrages de rang sont les suivants :

- ils ne sont pas sensibles aux valeurs extrêmes qui n'influencent donc pas le filtrage (cela ne concerne bien sûr pas les érosions et les dilatations) ;
- ils attribuent au signal filtré une valeur appartenant à l'ensemble des valeurs du signal bruité et donc ne créent pas de valeur nouvelle (à ce titre, ils permettent de filtrer des signaux portant une information plus *symbolique* qu'un simple niveau de gris, par exemple des images de classes obtenues après une étape de segmentation).

On a développé des filtres de rang plus complexes, par exemple le filtre d_α qui remplace la valeur courante par la valeur qui minimise :

$$d_\alpha = \left[\sum_{i \in \mathcal{V}} |\hat{f} - f_i|^\alpha \right]^{1/\alpha}$$

où $\hat{f} \in \{f_i\}$ et \mathcal{V} est la fenêtre d'estimation. Lorsque α est faible, on donne plus de poids aux valeurs f_i proches de \hat{f} tandis que les grandes valeurs de α donnent plus de poids aux écarts importants.

Ces filtres de rang donnent des résultats souvent satisfaisants dans les cas simples d'élimination du bruit. Ainsi, le filtre médian est pratiquement toujours préféré au filtre de moyenne ou au filtre gaussien pour améliorer les images bruitées.

1.2.3 Filtrages morphologiques

On a déjà vu que les érosions et les dilatations sont des filtres de rang (mais pas des filtres au sens morphologique), mais on leur préfère les ouvertures et les fermetures pour leur propriété d'idempotence (cf. chapitre ??). Pour des filtrages de taille croissante on prend des filtres alternés séquentiels (FAS) (cf. figure 1.2) [Schmitt et Mattioli, 1994, Serra, 1982].

Enfin, on fait également des filtres d'augmentation de contraste en répartissant judicieusement le niveau de sortie entre l'érodé, le dilaté ou la moyenne selon la distance du pixel à la valeur moyenne (on peut remplacer dans la phrase précédente *moyenne* par *médiane*).

1.2.4 Filtrages par équations de diffusion

Ces opérations ont fait l'objet du chapitre ?? où l'on a mis en évidence leur rôle de filtrage aussi bien pour l'analyse que pour la restauration. Ce sont ces propriétés qui nous intéresseront ici encore. Mais rappelons tout d'abord que les équations de diffusion constituent aussi l'approche duale de la minimisation d'une intégrale totale sur l'image. Dans cette approche, à une équation intégrale d'énergie comme l'équation 1.10, on associe une



FIG. 1.2 – **Filtre Alterné Séquentiel** - Une, deux ou trois itérations de la séquence ouverture puis fermeture, d'éléments structurants croissants.

équation d'Euler-Lagrange et des conditions aux limites (généralement les conditions de Neumann). On résout alors l'équation différentielle par des techniques de gradient (résolution itérative), à partir d'une solution initiale généralement égale à l'image bruitée. S'il y a convergence, c'est alors vers un minimum local.

Diffusion isotrope

On a vu au chapitre ?? que le filtrage par des équations de diffusion isotrope est, dans le cas continu, mathématiquement équivalent au filtrage gaussien [Marr et Hildreth, 1980, Babaud et al., 1986, Witkin, 1983, Koenderink, 1984]. Il est obtenu à partir de l'équation de la chaleur :

$\frac{\partial f}{\partial t} = \Delta f$ où l'on remplace le temps par des itérations successives. Ce filtrage, comme le filtrage gaussien, a le défaut de rendre les contours d'images très flous. Ce n'est donc pas un bon filtre d'amélioration d'images.

Diffusion anisotrope

On a montré au chapitre ?? que l'on préfère les équations de diffusion anisotrope [Perona et Malik, 1990, Alvarez, 1996]. Leur équation est :

$$\frac{\partial f}{\partial t} = \text{div}[g(f, t) \cdot \nabla f] \quad (1.14)$$

où $g(f, t)$ est une fonction généralement du module de la dérivée de f , maximale pour une dérivée nulle et décroissante lorsque le gradient croît : on choisit généralement les fonctions suivantes (cf. figure 1.3) :

$$\begin{aligned} - g(f, t) &= \exp\left(-\frac{|\nabla f|^2}{\kappa^2}\right), \\ - g(f, t) &= \left[1 + \left[\frac{|\nabla f|}{\kappa}\right]^{1+\alpha}\right]^{-1}. \end{aligned}$$

Le coefficient κ est proportionnel à un gradient. Il caractérise l'amplitude des gradients qui autoriseront une forte diffusion. Pour des valeurs de gradient très inférieures à κ , la diffusion sera très faible car la zone sera considérée comme déjà lisse. Pour des valeurs très supérieures, la diffusion sera interdite (on considérera qu'il y a présence d'un contour que le filtrage devra préserver). Cela se voit sur la figure 1.3 où l'on a porté l'évolution du produit $g(f, t) \cdot |\nabla f|$ en fonction de $|\nabla f|$.

La mise en œuvre des filtres de diffusion anisotrope se fait en calculant une série d'images f^n telles que :

$$f^n = f^{n-1} + \frac{\partial}{\partial x} \left[g(f, n) \cdot \frac{\partial f^{n-1}}{\partial x} \right] + \frac{\partial}{\partial y} \left[g(f, n) \cdot \frac{\partial f^{n-1}}{\partial y} \right]$$

Deux étapes de ce processus sont présentées sur la figure 1.4.

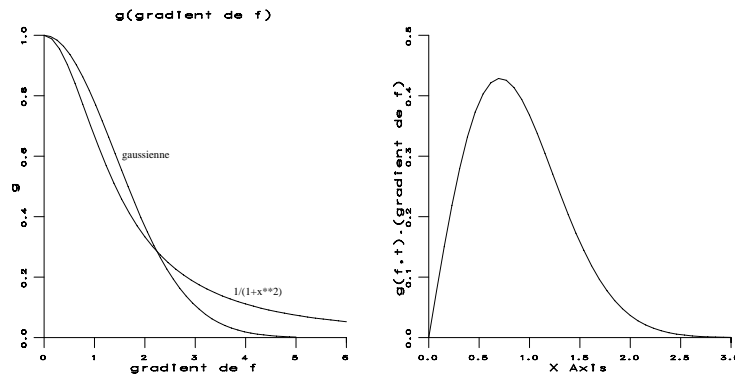


FIG. 1.3 – A gauche, terme contrôlant la diffusion dans les techniques de diffusion anisotropique pour deux exemples de filtres utilisés : la gaussienne $\exp(-\frac{|\nabla f|^2}{\kappa^2})$ ou la lorentzienne : $[1 + \frac{|\nabla f|^2}{\kappa^2}]^{-1}$. À droite, pour le cas de la gaussienne, variation du produit $g(f, t)|\nabla f|$ en fonction de $|\nabla f|$. Le maximum est proche de $|\nabla f| = \kappa$.



FIG. 1.4 – **Filtrage par Diffusion Anisotrope** : image originale bruitée (à gauche) et deux filtrages par 15 et 80 itérations avec $\kappa = 25$.

Filtres de courbure moyenne

On a également vu au chapitre ?? que plusieurs autres formes avaient été proposées pour les filtres par E.D.P. pour garantir la conservation de certaines propriétés des images. D'un intérêt particulier sont les filtres qui garantissent une invariance du traitement par changement de contraste. C'est le cas des filtres de courbure moyenne (ou *mean curvature equation*) [Kimia et al., 1992] :

$$\frac{\partial f}{\partial t} = \text{div}\left[\frac{\nabla f}{|\nabla f|}\right] \cdot \nabla f \quad (1.16)$$

Filtrage à Variation Totale

Dans la plupart des approches de filtrage, on recherche la meilleure solution au sens de l'erreur quadratique moyenne. Ce critère n'est pas très satisfaisant pour un observateur humain qui préfère souvent des optimisations sous d'autres normes. On préfère parfois la norme L^1 , mais celle ci conduit à des solutions souvent difficiles à mettre en œuvre. Cet argument a été poursuivi dans [Rudin et al., 1992] pour proposer un filtrage dans la classe des signaux à variation totale bornée (ou signaux **BV**), sous-classe des signaux **TV**, c'est-à-dire à contrainte sur la variation totale [Rudin et Osher, 1994].

Ainsi, sous ce formalisme, on minimise :

$$\int_{\Omega} \hat{f}_x^2 + \hat{f}_y^2 dx dy$$

où \hat{f}_x est la dérivée en x de l'estimée de l'image, sous les contraintes :

$$\int_{\Omega} \hat{f} dx dy = \int_{\Omega} f dx dy \quad \text{et} \quad \int_{\Omega} (\hat{f} - f)^2 dx dy = \sigma^2$$

Ces deux contraintes sont des hypothèses sur le bruit (valeur moyenne nulle et écart-type fixé).

Ces relations conduisent aux équations d'Euler-Lagrange :

$$\frac{\partial}{\partial x} \left(\frac{\hat{f}_x}{\sqrt{\hat{f}_x^2 + \hat{f}_y^2}} \right) + \frac{\partial}{\partial y} \left(\frac{\hat{f}_y}{\sqrt{\hat{f}_x^2 + \hat{f}_y^2}} \right) - \lambda_1 - \lambda_2(\hat{f} - f) = 0$$

avec, comme condition aux limites $\frac{\partial \hat{f}}{\partial n} = 0$. On la résout itérativement par une descente de gradient :

$$\frac{\partial f}{\partial t} = \frac{\partial}{\partial x} \left(\frac{\hat{f}_x}{\sqrt{\hat{f}_x^2 + \hat{f}_y^2}} \right) + \frac{\partial}{\partial y} \left(\frac{\hat{f}_y}{\sqrt{\hat{f}_x^2 + \hat{f}_y^2}} \right) - \lambda(\hat{f} - f) \quad (1.20)$$

avec une valeur de λ donnée par :

$$\lambda = -\frac{1}{2\sigma^2} \int \left[\sqrt{\hat{f}_x^2 + \hat{f}_y^2} - \left(\frac{f_x \hat{f}_x}{\sqrt{\hat{f}_x^2 + \hat{f}_y^2}} + \frac{f_y \hat{f}_y}{\sqrt{\hat{f}_x^2 + \hat{f}_y^2}} \right) \right] dx dy$$

Selon cette méthode, on filtre progressivement l'image d'origine, tout en respectant ses contours. Le schéma donne des résultats assez semblables à ceux du filtre de courbure moyenne. On s'en convaincra en notant l'identité des équations 1.16 et 1.20, dans le cas où $\lambda = 0$ et le membre de droite est multiplié par $|\nabla f|$.

1.3 Les filtres adaptatifs

1.3.1 Les filtres à coefficients adaptatifs

Ces filtres s'inspirent des filtres de moyenne, mais chaque terme de la moyenne est pondéré par un coefficient qui décroît avec la similarité entre le pixel considéré et le pixel central de la fenêtre. On peut également interpréter ces deux filtres comme des filtres de diffusion isotrope mais adaptative.

Filtre de gradient inverse

Il est appelé également filtre σ ; il calcule une moyenne des pixels de la fenêtre, pondérés par l'inverse de leur gradient, de façon que les points trop différents du point traités ne contribuent que peu :

$$f'(i, j) = \frac{1}{\alpha} \sum_k \sum_l \frac{f(i+k, j+l)}{|f(i, j) - f(i+k, j+l)| + 1}$$

C'est un filtre d'action très modérée que l'on itère généralement plusieurs fois.

Filtre de Saint Marc

Il repose sur un principe assez semblable, mais la pondération est exponentielle négative :

$$f^{k+1}(i, j) = \frac{1}{\alpha} \sum_l \sum_{l'} f^k(i+l, j+l') \exp(-\beta |\nabla f^k|)$$

Son action est plus nette que celle du précédent. Il conduit à de bons filtrages respectant bien les contours [Saint-Marc et al., 1989].

Toboggan de Fairfield

Il agit de façon très différente des précédents. Il creuse l'histogramme de l'image en attribuant à chaque pixel le niveau de gris du point de minimum de gradient dans son voisinage. On trace tout d'abord une carte du gradient. Puis, à partir d'un point non déjà parcouru, on descend sur cette carte comme sur un toboggan (d'où le nom) jusqu'à ce qu'on tombe dans un minimum local du gradient. On attribue alors la valeur du niveau de gris du point minimum local. Tous les points du chemin parcouru sont mis au même niveau, et tous ces points sont marqués afin de n'être plus traités. On répète ce traitement jusqu'à ce que tous les points aient été parcourus [Fairfield, 1990].

1.3.2 Les filtres à fenêtres adaptatives

Dans le cas de filtres à fenêtres adaptatives, on recherche, autour de chaque pixel, la fenêtre la plus adaptée au filtrage. Cela se fait de deux façons :

1. soit en sélectionnant parmi une famille de fenêtres celle qui convient le mieux (c'est le cas du filtre de Nagao),
2. soit en faisant croître une fenêtre et en contrôlant la croissance (cas du filtre de Wu).

Filtre de Nagao

Dans l'approche de Nagao [Natsuyama, 1979], on travaille sur une fenêtre de taille 5×5 entourant le pixel central (qui appartient à toutes ces fenêtres). On définit 9 fenêtres possibles (cf. figure 1.5), toutes de 9 pixels, identifiées par un indice k . Sur chaque fenêtre on mesure la moyenne m_k et la variance σ_k^2 des niveaux de gris. On choisit alors de remplacer le pixel central par la valeur moyenne de la fenêtre dont la variance est la plus faible.

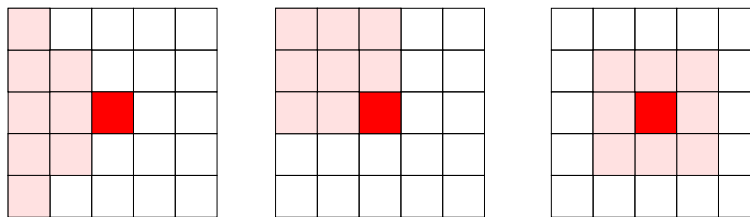


FIG. 1.5 – Le pixel central d'une fenêtre 5×5 appartient à 9 fenêtres de 9 pixels chacune : 4 se déduisent de la fenêtre représentée à gauche par rotations de $n \times 90^\circ$, 4 de la fenêtre au centre et la fenêtre de droite.

Filtre de Wu : filtrage par fenêtre maximale

Il procède par croissance de région [Wu et Maître, 1992]. Il utilise un prédicat d'homogénéité sur une région (cf. chapitre 3, section 3.2), c'est souvent une mesure de la variance (ce peut être le coefficient de variation pour

une image de radar). On part d'une zone de taille 3×3 , si elle est homogène on l'étend à une fenêtre 5×5 puis 7×7 en lui ajoutant des couronnes successives. Si le prédicat n'est plus vérifié, on recherche la cause de l'hétérogénéité. Si elle n'est présente que sur un côté de la fenêtre, on procède à une croissance sur une fenêtre rectangulaire en interdisant de croître sur le côté hétérogène. Si 2 côtés adjacents sont hétérogènes, on croît selon un triangle. Lorsque la fenêtre ne peut plus croître, on remplace le point initial par la moyenne sur la plus grande fenêtre obtenue (cf. figure 1.6).

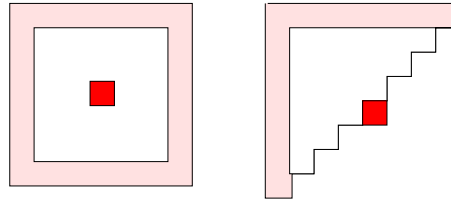


FIG. 1.6 – Méthode de Wu. A gauche, passage d'une fenêtre à la suivante dans le cas d'une croissance isotrope. A droite, s'il y a un contour dans le 1/2 plan inférieur droit, la croissance ne se fait plus que selon un triangle.

1.4 Le rééchantillonnage des images

Cette opération est nécessaire chaque fois que l'on doit transformer une image dans une géométrie prédéfinie (par exemple en télédétection pour se ramener dans une référence cartographique, ou en imagerie médicale pour aligner les images sur les repères anatomiques). Ces opérations de rééchantillonnage s'appuient sur les bases mathématiques de la théorie de l'échantillonnage de Shannon (cf. chapitre ??).

L'idée de principe est assez simple : considérant que l'image d'origine vérifie le théorème de Shannon, on reconstruit tout d'abord le signal continu dont cette image discrète est issue, puis cette image continue est filtrée de façon à donner une image dont le spectre est compatible avec le nouvel échantillonnage que l'on souhaite réaliser, enfin on échantillonne cette image aux points que l'on désire connaître.

Mais cette idée de principe se heurte à des contraintes de mise en œuvre matérielle très difficiles (temps de calcul, place mémoire, etc.) qui nécessitent que l'on trouve des solutions de remplacement.

Cela se fait en combinant reconstruction et filtrage en un seul filtrage. Cela se fait également très souvent en remplaçant les fonctions canoniques de reconstruction (généralement des sinus cardinaux) par des fonctions plus compactes, souvent polynomiales, permettant d'effectuer des sommations finies et non infinies.

1.4.1 Interpolation avec des polynômes

Dans ces approches, on convole la fonction initiale $f(x)$ ¹ par un noyau $h(x)$:

$$g(x) = f(x) * h(x) = \int_{-\infty}^{\infty} f(y)h(x-y)dy$$

ce qui s'exprime dans l'espace de Fourier par :

$$G(u) = F(u)H(u)$$

F et H étant les TF de f et h .

¹Nous la considérons ici monodimensionnelle pour simplifier, pour les fonctions bidimensionnelles il convient généralement de faire le produit d'un polynôme en x par un polynôme en y .

L'interpolation idéale est obtenue pour :

$$h(x) = \frac{\sin(2\pi x)}{x} = \text{sinc } \pi x \quad ; \quad H(u) = \text{rect}(u)$$

Mais d'autres interpolations, plus rapides, sont souvent utilisées :

- Interpolation au plus proche voisin :

$$h(x) = \text{rect}(x) \quad ; \quad H(u) = \text{sinc } \pi u.$$

- Interpolation linéaire :

$$h(x) = (1 - |x|)\text{rect}(2x) \quad ; \quad H(u) = \text{sinc}^2 \pi u.$$

- Interpolation quadratique :

Plusieurs formules ont été proposées. Celle de [Dodgson, 1997] est particulièrement utilisée :

$$\begin{aligned} h(x) &= -2x^2 + 1 & \text{si} & \quad |x| \leq 1/2 \\ h(x) &= x^2 - 5/2|x| + 3/2 & \text{si} & \quad 1/2 \leq |x| \leq 3/2 \\ h(x) &= 0 & \text{sinon} & \end{aligned}$$

dont la fonction de transfert est :

$$H(u) = \frac{6 \sin \pi u - 6 \sin 3\pi u - \cos \pi u + \cos 3\pi u}{(2\pi u)^2}$$

- Interpolation cubique :

Les solutions possibles sont encore plus nombreuses et ont fait l'objet de très nombreux travaux [Maeland, 1988]. On retiendra la famille paramétrée par la variable a :

$$\begin{aligned} h(x) &= (a + 2)|x|^3 - (a + 3)x^2 + 1 & \text{si} & \quad |x| \leq 1 \\ h(x) &= a(|x|^3 - 5x^2 + 8|x| - 4) & \text{si} & \quad 1 \leq |x| \leq 2 \\ h(x) &= 0 & \text{sinon} & \end{aligned}$$

La valeur $a = -1/2$ est souvent retenue. Elle conduit à une interpolation dont la fonction de transfert est :

$$H(u) = \frac{18 - 24 \cos 2\pi u + 6 \cos 4\pi u - \pi u(2 \sin 2\pi u - \sin 4\pi u)}{(2\pi u)^4}$$

1.4.2 Interpolation par des B-splines

Les B-splines forment également une famille de polynômes intéressants [Unser et al., 1993a, Unser, 1999, Chalmond, 2000]. Ils sont de degré k et se déduisent récursivement du polynôme de degré 0 par la formule :

$$\beta^k(x) = *^{k+1} \beta^0(x)$$

où $*^k$ exprime la $k^{\text{ème}}$ convolution. Ainsi :

$$\begin{aligned} \beta^1(x) &= \beta^0(x) * \beta^0(x) \\ \beta^2(x) &= \beta^0(x) * \beta^0(x) * \beta^0(x) \\ &\dots \text{etc.} \end{aligned}$$

La fonction d'ordre 0 est définie par la fonction porte : $\beta^0(x) = \text{rect } x$, et les diverses fonctions sont explicitées dans [Unser et al., 1993b] jusqu'à l'ordre 7. Leurs TF s'expriment simplement par :

$$TF[\beta^k] = (\text{sinc } u/2)^{k+1}$$

Une représentation par les splines d'ordre n prend la forme générale :

$$s(x) = \sum_{k \in Z} c(k) \beta^n(x - k) \quad (1.33)$$

Lorsqu'on interpole un signal dans l'approche par splines, on choisit généralement des splines **interpolantes**, qui passent exactement par tous les points de l'échantillonnage, et non des splines **approximantes** qui passent au plus près d'un ensemble de points. Calculer une telle approximation consiste donc à déterminer les coefficients inconnus c_k tels que en tout point x appartenant aux échantillons initiaux on ait :

$$\sum_{k \in Z} c(k) \beta^n(x - k)|_{x=k} = s(k) \quad (1.34)$$

les $s(k)$ étant les valeurs connues de l'échantillonnage.

On peut calculer les approximations par splines de deux façons :

1. par résolution du système d'équations 1.34 liant les coefficients inconnus des splines aux coordonnées et aux amplitudes connues des échantillons ; on est alors amené à inverser une matrice Toeplitz (cf. chapitre ??) (invariante par translation le long de x). Les coefficients de la matrice sont les valeurs des β^n aux nœuds du maillage, ils sont au nombre de $n + 1$. Les inconnues, les $c(k)$, sont plus nombreuses que les équations (on ne dispose pas d'équations complètes pour les $n/2 - 1$ premiers et derniers points de l'échantillonnage). On choisit donc généralement de répéter les premières et les dernières valeurs, ou on les reproduit selon une symétrie miroir.
2. par filtrage (c'est-à-dire par convolution) [Goshtasby et al., 1990] :
Pour cela, utilisons le noyau b_m^n obtenu en échantillonnant la B-spline de degré n dilatée d'un facteur m et associons lui sa transformée en z :

$$b_m^n(k) = \beta(x/m)|_{x=k} \quad \rightarrow \quad B_m^n(z) = \sum_{k \in Z} b_m^n(z)(k) z^{-k}$$

Nous voulons déterminer les coefficients $c(k)$ des splines qui garantissent l'égalité aux échantillons entiers :

$$\sum_{l \in Z} c(l) \beta^n(x - l)|_{x=k} = s(k)$$

ce que l'on peut réécrire comme une convolution :

$$s(k) = b_1^n * c$$

Si l'on définit par $\gamma(k)$ la fonction dont la transformée en Z (TZ) est l'inverse de $B_1^n(z)$:

$$\gamma(k) \quad \rightarrow \quad \frac{1}{B_1^n(z)}$$

on obtient une solution du problème par :

$$c(k) = \gamma(k) * s(k)$$

On montre que cette déconvolution est stable et peut être effectuée très rapidement [Unser et al., 1993a, Unser et al., 1993b].

À deux dimensions, l'interpolation spline se déduit du cas monodimensionnel par produit tensoriel de l'interpolation en x et en y (équation 1.33) :

$$s(x, y) = \sum_{k \in Z} \sum_{k' \in Z} c(k, k') \beta^n(x - k) \beta^n(y - k') \quad (1.40)$$

1.4.3 Interpolation adaptative

Des méthodes complètement différentes des précédentes effectuent des interpolations adaptatives, dans le but de préserver les objets présents dans la scène et leurs contours en particulier. Ces techniques contrôlent souvent l'interpolation par une détection de contours ou une segmentation de l'image. Un exemple de telles techniques est donné par [Karabassis et Spetsakis, 1995].

D'autres méthodes [Calle et Montanvert, 1998] s'appuient sur des décompositions fractales qui permettent de retrouver les propriétés fines - filtrées par l'échantillonnage original - mais présentes à travers la pyramide d'homothétie interne (cf. chapitre ??).

Chapitre 2

La détection des contours dans les images

Chapitre rédigé par Henri MAÎTRE

La détection de contours dans les images a débuté de façon extrêmement empirique par des opérateurs locaux qui, soit estimaient un gradient, soit convoluait l'image par des masques caractéristiques des contours [Haralick et Shapiro, 1985]. Dans les années 80, des approches plus systématiques ont été mises en place par Marr [Marr et Hildreth, 1980], puis Canny [Canny, 1986], pour obtenir des contours plus significatifs. Ces travaux ont abouti à une bonne compréhension de ce qu'il faut faire pour détecter les contours, mais la définition même des contours demeure très vague, ce qui rend ces techniques encore peu efficaces sur un problème concret. De plus, on a pu montrer que le problème de détection de contours est généralement mal posé (au sens de la résolution des systèmes) [Torre et Poggio, 1986]. Les seuls modèles de contours utilisables sont ceux de contours idéalisés, comme ceux présentés sur la figure 2.1 ; ils sont bien loin de la réalité. C'est pourquoi, même si de très gros progrès ont été accomplis dans ce domaine, les techniques empiriques d'estimation du gradient proposées dans les années 70-80 restent souvent encore employées en concurrence de techniques plus modernes. Une excellente référence à ce problème est l'ouvrage collectif [Cocquerez et Philipp, 1996]. Plusieurs articles existent sur la comparaison des performances de détecteurs de contours, en particulier [Palmer et al., 1996, Heath et al., 1997].

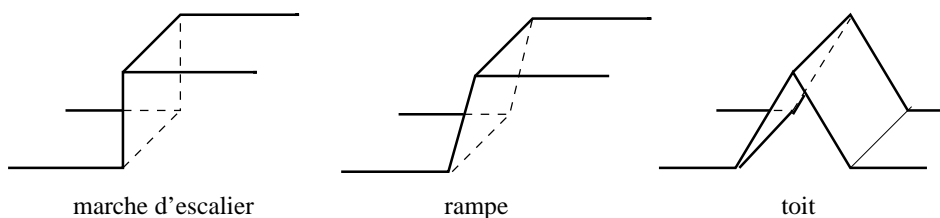


FIG. 2.1 – Quelques modèles de contours. Le plus utilisé est celui en marche d'escalier.

Nous présenterons tout d'abord une approche formelle (mais malheureusement stérile !) de la détection des contours, puis les méthodes empiriques encore utilisées seront vues dans la partie 2. Une approche plus analytique, proposée par Canny, sera ensuite présentée, avec ses dérivées qui sont aujourd'hui les plus employées en raison de leur efficacité. Enfin les techniques fondées sur le principe des **contours actifs** de Kass, Witkins et Terzopoulos seront présentées, ainsi que les approches par **ensembles de niveaux** (*level sets*). Nous terminerons par une brève présentation des méthodes de poursuite et de fermeture des contours.

2.1 Un modèle continu de contour.

Dans une image supposée continue $f(x, y)$, un contour apparaît comme une ligne où sont localisées les très fortes variations de f . Soit \vec{G} le gradient de f :

$$\vec{G} = \vec{\nabla} f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

On associe à f une image du module du gradient de f :

$$G = |\vec{\nabla} f| = \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right]^{1/2} \quad (2.1)$$

ainsi qu'une image des directions du gradient à partir de :

$$\vec{g} = \frac{\vec{\nabla} f}{|\vec{\nabla} f|}$$

et par application de $[-\pi, \pi[$ sur $\{0, 255\}$. Un contour est alors défini comme **le lieu des maximums du gradient dans la direction \vec{g} du gradient**. Un point de contour vérifie donc ;

$$\frac{\partial G}{\partial g} = 0 \quad \text{et} \quad \frac{\partial^2 G}{\partial g^2} \leq 0 \quad (2.2)$$

avec :

$$\frac{\partial}{\partial g} = \vec{g} \cdot \vec{\nabla} \quad (2.3)$$

L'équation obtenue est complexe et non linéaire :

$$\frac{\partial f}{\partial x} \cdot \frac{\partial}{\partial x} \sqrt{\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2} + \frac{\partial f}{\partial y} \cdot \frac{\partial}{\partial y} \sqrt{\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2} = 0$$

Elle n'a donc en général pas de solution explicite et ne peut être résolue que par tâtonnement. On procède alors en deux étapes : on calcule tout d'abord le gradient, puis on recherche les extrémums dans la direction du gradient.

Afin de garantir la double dérivation même en présence de discontinuités de type *marche d'escalier*, on prétraite l'image f par convolution avec une fonction au moins deux fois dérivable. On a beaucoup utilisé pour cela la gaussienne, mais on peut également prendre des polynômes de faible degré.

On peut également se placer dans les axes locaux définis par la tangente \vec{t} et la normale \vec{g} à la surface $f(x, y)$ (cf. figure 2.2). Ce repère est tourné d'un angle Φ par rapport au repère $\{x, y\}$:

$$\Phi = \text{Arctg} \left[\frac{\partial f}{\partial x} / \frac{\partial f}{\partial y} \right] \quad (2.4)$$

On a alors :

$$\frac{\partial f}{\partial g} = \frac{\partial f}{\partial x} \cos \Phi + \frac{\partial f}{\partial y} \sin \Phi$$

et les équations 2.2 donnent :

$$\frac{\partial^2 f}{\partial g^2} = \frac{\partial^2 f}{\partial x^2} \cos^2 \Phi + 2 \cdot \frac{\partial^2 f}{\partial x \partial y} \cos \Phi \sin \Phi + \frac{\partial^2 f}{\partial y^2} \sin^2 \Phi$$

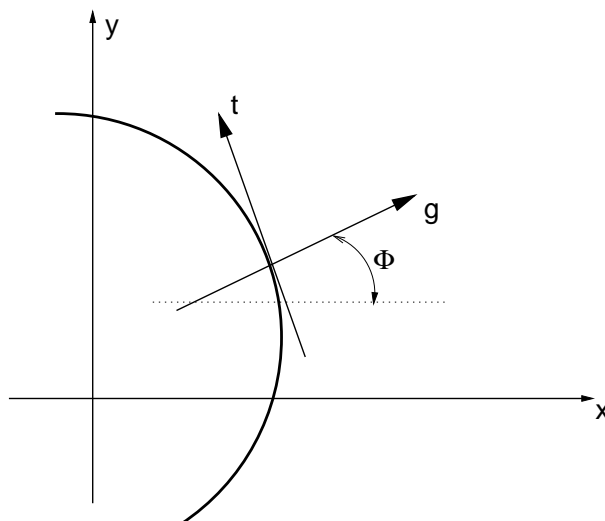


FIG. 2.2 – Repère local défini par le vecteur normal et le vecteur tangent à l’image.

Cette équation n’est pas plus simple à résoudre que la précédente. Cependant elle se relie assez aisément à l’équation du laplacien de f :

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

qui s’exprime également (par invariance du laplacien par changement de repère euclidien) dans le repère local $\{t, g\}$:

$$\Delta f = \frac{\partial^2 f}{\partial g^2} + \frac{\partial^2 f}{\partial t^2}$$

On voit que si le terme $\frac{\partial^2 f}{\partial t^2}$ est négligeable (c’est-à-dire si le contour a une courbure très faible), alors :

$$\Delta f = 0 \quad \approx \quad \frac{\partial^2 f}{\partial g^2} = 0 \quad (2.5)$$

Cette propriété sera utilisée dans les approches dites par passage par zéro du laplacien (cf détecteur de Marr).

2.2 Les approches classiques

Cette section présente un ensemble de méthodes qui ont eu historiquement une grande importance en traitement des images. Bien que certaines soient encore régulièrement employées, un lecteur intéressé par les techniques plus actuelles pourra se rendre directement à la section 2.3.

2.2.1 Les détecteurs de gradient par filtrage

Ces détecteurs reposent tous sur une recherche d’un extremum de la dérivée première (ou d’un passage par zéro d’une dérivée seconde), celle-ci étant calculée de diverses manières, mais généralement par un filtrage passe-haut précédé d’un léger filtrage passe-bas pour s’affranchir des bruits. Ces approches par **filtrage linéaire** sur des critères très simples de sélection en fréquence ont reçu une base théorique dans [Modestino et Fries, 1977, Shanmugam et Green, 1979]. La somme des connaissances sur ce sujet se trouve dans [Torre et Poggio, 1986].

Ainsi, Modestino et Fries ont proposé de détecter les contours par un filtrage de Wiener, le but de ce filtrage étant d'effectuer sur l'image non bruitée une opération de laplacien qui sera suivie par la suite d'une détection de zéros alliée à un seuillage (nous verrons ces étapes ultérieurement).

Le filtre de base de Modestino et Fries est un filtre récursif de la forme :

$$H_0(z_1, z_2) = \frac{\sum_i^n \sum_j^n b_{i,j} z_1^{-i} z_2^{-j}}{1 + \sum_i^n \sum_j^n a_{i,j} z_1^{-i} z_2^{-j}}$$

où les coefficients $a_{i,j}$ et $b_{i,j}$ sont déterminés à partir de 3 paramètres spécifiques de l'image à traiter : λ , ρ et ϵ . λ est le paramètre poissonnien de densité de contours (il exprime la distance moyenne entre 2 contours lors d'un balayage de l'image), ρ la constante de la loi exponentielle de l'auto-corrélation de l'image (cf. chapitre ??) et ϵ la puissance du bruit. Comme un tel filtre ne filtre que du coin supérieur gauche au coin inférieur droit, on lui associe les 3 autres filtres obtenus par rotation de $\pi/2$, pour obtenir un filtre final. Dans la pratique pour assurer des temps de calcul raisonnables, on choisit des filtres d'ordre 1 ($n = 1$).

2.2.2 Les détecteurs de gradient par masques

A côté de ces approches très inspirées du traitement du signal, des filtres de dérivation plus empiriques ont été proposés à partir d'estimateurs locaux de l'image f ou de ses dérivées $\frac{\partial f}{\partial x}$. Ces estimées sont obtenues à l'aide de masques (cf. figure 2.3) appliqués sur des fenêtres de 2×2 pixels ou 3×3 pixels (exceptionnellement, en cas d'images très bruitées, sur des fenêtres plus grandes). On note sans surprise que la somme des coefficients de ces filtres est nulle (fonction de transfert nulle à la fréquence 0), et que les coefficients sont anti-symétriques. Les filtres les plus utilisés sont, dans l'ordre décroissant :

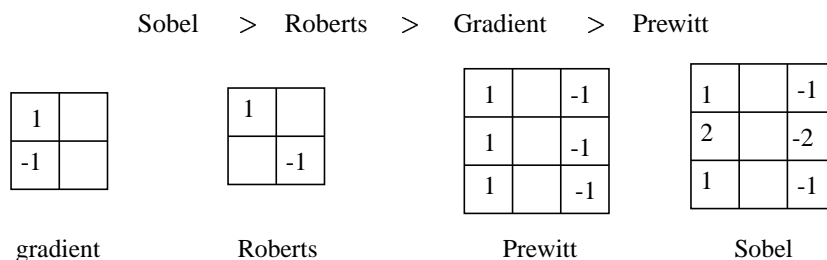


FIG. 2.3 – Quatre filtres de détection de contours par estimation du gradient. Les filtres représentés estiment une seule dérivée. Par rotation de $\pi/2$, on calcule la seconde dérivée.

Les filtres 3×3 sont un peu moins précis (c'est-à-dire que les contours qu'ils détectent sont moins bien localisés et souvent épais), mais les images ainsi obtenues sont généralement plus fiables et permettent des post-traitements plus poussés, ils sont également centrés sur un pixel et non entre des pixels (filtres à phase nulle).

La réponse de l'un quelconque de ces filtres s'obtient de la façon suivante : le filtre est centré en chaque pixel successivement (au coin supérieur gauche pour les filtres 2×2 , au point central pour les filtres 3×3). Le produit du masque par les valeurs des pixels correspondant étant fait, la valeur absolue de la somme est retenue. Puis le masque est tourné de 90° autour de son centre et la même mesure est répétée. Les deux mesures sont alors ajoutées et leur somme constitue la mesure du gradient en ce point selon la formule de l'équation 2.1, ou selon la formule approchée (en norme \mathcal{L}_1), un peu plus rapide à calculer :

$$\hat{G} = \left| \frac{\partial f}{\partial x} \right| + \left| \frac{\partial f}{\partial y} \right|.$$

On peut également mesurer l'orientation Φ du contour au point donné en faisant le rapport des réponses des deux filtres comme dans l'équation 2.4.

De tels filtres sont particulièrement simples de mise en œuvre, et rapides de calcul, permettent en particulier d'assurer des calculs en temps réel pour des applications industrielles.

2.2.3 Pré- et post traitements

La qualité de la détection est très liée à la qualité des contours dans l'image (cela sera vérifié pour toutes les méthodes, mais particulièrement pour ces méthodes peu élaborées). Ainsi, elles sont souvent bonnes pour les applications de robotique, en environnement artificiel, pour des images aériennes à forte résolution (en milieu urbain par exemple), elles se dégradent très vite pour les scènes d'extérieur, les images médicales, et deviennent totalement inapplicables en imagerie bruitée (thermographie, imagerie ultrasonore, radar SAR, etc.). On a alors intérêt à revenir à des démarches moins intuitives comme celles exposées dans [Modestino et Fries, 1977, Shanmugam et Green, 1979].

Ces détecteurs doivent toujours être suivis d'une étape de post-filtrage, et souvent précédés d'une étape de pré-filtrage.

Parmi les filtres de **pré-traitement**, on recherche ceux qui diminuent le bruit tout en préservant les discontinuités. Ils ont fait l'objet du chapitre ?? : filtre médian par exemple, mais aussi filtre toboggan [Fairfield, 1990] (cf. section 1.3.1), ou filtre de diffusion anisotrope de Perona et Malik [Perona et Malik, 1987] (cf. section 1.2.4). Ces filtres ont cependant le défaut de créer des contours artificiels qui peuvent être par la suite difficiles à éliminer.

Les **post-traitements** commencent le plus souvent par l'élimination des points de contour trop faibles par un seuillage (seuillage adaptatif si l'image est très hétérogène). On élimine ensuite les points qui ne sont pas des extrémums locaux dans la direction du gradient. Cela s'obtient à partir de la direction Φ mesurée précédemment et une comparaison simple des points rencontrés (ou interpolés si l'on souhaite une précision meilleure que le pixel). On peut opérer également des seuillages *par hystérésis*, afin de ne conserver que les composantes les plus importantes des contours. Pour cela, on procède à deux seuillages, le second étant très tolérant (c'est-à-dire laissant un très grand nombre de contours candidats) ; on ne garde des contours sélectionnés par ce second seuil que les points connexes d'un point de contour préservé par le premier (cela se fait aisément par un double étiquetage des images seuillées) cf. figure 2.4.

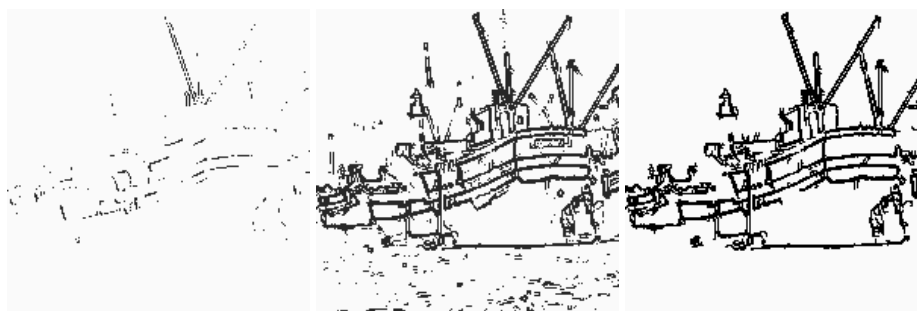


FIG. 2.4 – Filtrage de contours par hystérésis : à gauche, application d'un seuil sévère (70) au gradient de l'image bateau, seuls les contours les plus fiables sont détectés. Au centre, application d'un seuil tolérant (25), beaucoup de contours sont détectés, mais également beaucoup de bruit. A droite, résultat du filtre par hystérésis : seuls les contours du second filtrage connexes à ceux du premier sont conservés. Le détecteur de contours est un détecteur de Sobel. Aucun filtrage n'est appliqué pour réduire l'épaisseur des contours.

Enfin on procède à des étapes de **poursuite** et de **fermeture** des contours. La première opération a pour objectif de supprimer les petites disparitions de contours qui peuvent se produire par suite du bruit ou d'une perte de contraste, la seconde se propose de détecter des objets topologiquement clos, c'est-à-dire des zones fermées. Ces étapes seront vues plus en détail plus loin.

2.2.4 Les détecteurs de passage par zéro du laplacien

Ces méthodes ont été proposées en 1976 [Marr et Hildreth, 1980]. Elles ont eu une grande importance historique, étant considérées comme le prototype du détecteur de contour inspiré des systèmes biologiques (*primal sketch* de Marr). Elles utilisent le fait que le passage par zéro du laplacien permet de bien mettre en évidence les extrémums de la dérivée. Nous avons eu l'occasion de voir (cf. équation 2.5) qu'en l'absence de forte courbure, le passage par zéro du laplacien correspond en effet bien au maximum du gradient dans la direction du gradient. Ces méthodes tirent en outre profit du fait que les zéros de la dérivée seconde constituent un réseau de lignes fermées (évitant donc, en principe, les étapes de poursuite et de fermeture). Il en est de même du réseau des lignes de crête du gradient, mais le premier est plus aisément détecté à partir d'un simple étiquetage des zones positives et négatives. Mais l'estimation de la dérivée seconde étant très sensible aux bruits, il convient de filtrer très fortement l'image avant d'en mesurer le laplacien. Cela conduit au filtrage suivant :

$$\psi * \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \right) = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) [\psi * f].$$

où ψ est un filtre passe-bas ; ce qui se récrit symboliquement :

$$image\ des\ contours = passage\ par\ zéro \left(f * \left[\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right] \right).$$

cela exprime qu'une image de contours est obtenue par filtrage de l'image par la dérivée seconde d'un filtre passe-bas, puis détection des zéros de la fonction ainsi obtenue.

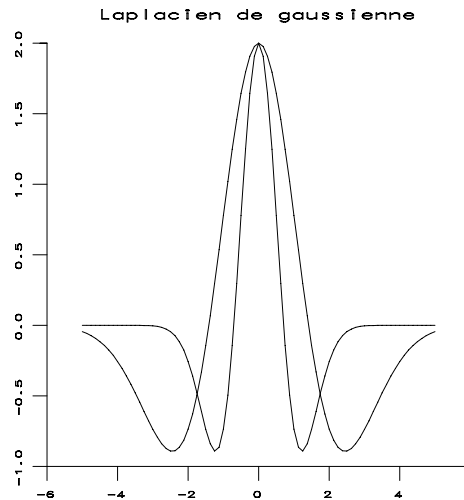


FIG. 2.5 – Filtre LOG pour deux valeurs différentes de σ .

Les filtres les plus utilisés pour ces filtrages passe-bas sont les filtres gaussiens [Marr et Hildreth, 1980]. Marr a montré en effet qu'avec de tels filtres on pouvait approcher de très près les effets donnés par le système visuel humain. Le filtre obtenu par convolution avec le laplacien d'une gaussienne est connu sous le nom de LOG (cf figure 2.5). Un autre filtre très utilisé et très proche du LOG est le DOG (*Difference of Gaussians*) (figure 2.6, qui procède par différence de deux gaussiennes dont les écarts-type sont dans un rapport 1,6 (pour simuler au mieux le système visuel). Les gaussiennes sont bien sûr affectées des facteurs adéquats pour que la somme des parties positives du filtre soit égale à la somme des parties négatives (valeur nulle à la fréquence 0 de la fonction

de transfert). Sa mise en œuvre peut tirer profit des diverses qualités de la gaussienne : séparabilité, noyau auto-reproducteur, limite de suites de polynômes, approximation par des splines, etc., ce qui permet d'en accélérer la réalisation.

Dans des approches très sommaires, le DOG est parfois remplacé par le DOB (*Difference of Box*), filtre très rapide même pour des grandes tailles (on tire profit alors du fait que les filtres de moyenne (*Box*) ont une complexité de calcul indépendante de la taille de la fenêtre d'intégration au prix d'un coup mémoire important). Les résultats de ces filtres est cependant généralement de piètre qualité.

Les filtres LOG et DOG, n'ont plus aujourd'hui qu'un intérêt historique. On leur préfère les filtres analytiques que nous verrons plus loin.

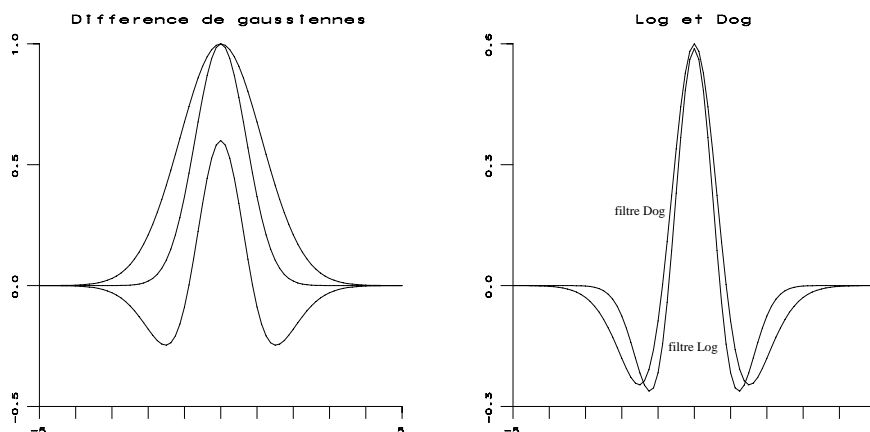


FIG. 2.6 – A gauche : filtre DOG et les deux gaussiennes dont il est la différence (leurs écarts-types sont dans un rapport 1,6). A droite filtre LOG et filtre DOG comparés.

2.2.5 Les détecteurs par masquage adapté

Dans ces techniques, on recherche en chaque point de l'image la présence d'une configuration conforme à un gabarit appartenant à un *dictionnaire de contours*. Pour cela, on définit une distance entre un contour type et une fenêtre de l'image. C'est par exemple l'inverse du produit scalaire des deux fenêtres : si l'image est dénotée $f(x, y)$, et le $k^{\text{ème}}$ gabarit $g_k(x, y)$, on calcule par exemple :

$$r_k(x, y) = \frac{\left(\sum_{x, y \in \text{fenêtre}} f^2(x, y) \cdot \sum_{x, y \in \text{fenêtre}} g_k^2(x, y) \right)^{1/2}}{\sum_{x, y \in \text{fenêtre}} f(x, y) g_k(x, y)}$$

pour toutes les valeurs de k . On ne conserve que la meilleure des valeurs et seulement si elle est suffisamment faible. L'approche par masques adaptés (*template matching*), permet de connaître aisément la direction du contour, elle permet également de soumettre, en parallèle, ces calculs à des architectures adaptées. Mais la qualité des contours, ainsi que les post-traitements qu'ils requièrent sont très semblables à ceux que l'on a vus pour les filtres de maximum du gradient.

Les filtres adaptés les plus utilisés sont le filtre de Kirsch (cf. tableau 2.1), le filtre de Nevatia et Babu (cf. tableau 2.2) et les boussoles directionnelles de Prewitt (cf. tableau 2.3).

Au-delà de ces approches souvent heuristiques, Hueckel a proposé une démarche beaucoup plus rationnelle, s'appuyant sur la décomposition du signal d'image sur une base de fonctions orthogonales en coordonnées polaires (polynômes d'Hermite utilisés également pour l'oscillateur harmonique), tronquée à ses 8 premiers termes

5	5	5
-3	0	-3
-3	-3	-3

-3	5	5
-3	0	5
-3	-3	-3

TAB. 2.1 – Masques de Kirsch : il y a 8 filtres issus de ces 2 fenêtres par rotation de $\pi/4$.

-100	-100	0	100	100
-100	-100	0	100	100
-100	-100	0	100	100
-100	-100	0	100	100
-100	-100	0	100	100

-100	32	100	100	100
-100	-78	92	100	100
-100	-100	0	100	100
-100	-100	-92	78	100
-100	-100	-100	32	100

TAB. 2.2 – Masques de Babu et Nevatia : il y a 12 filtres issus de ces 2 fenêtres par rotation, la première de $\pi/2$, la seconde de $\pi/4$.

[Hueckel, 1971]. Malheureusement les résultats n'ont pas été au niveau des investissements et cette voie est maintenant abandonnée.

2.3 Les approches analytiques

Nous allons voir maintenant une approche qui a permis une bien meilleure compréhension des conditions d'une bonne détection de contours et qui a ainsi conduit à des détecteurs de très bonne qualité. On les voit émerger dans les années 85, à partir des travaux : [Torre et Poggio, 1986, Shen et Castan, 1986, Canny, 1986].

2.3.1 Les critères de Canny

Canny, dans une approche originale, [Canny, 1986] a proposé un filtre déterminé analytiquement à partir de 3 critères :

1. garantir une bonne détection, c'est-à-dire une réponse forte même à de faibles contours,
2. garantir une bonne localisation,
3. assurer que pour un contour il n'y aura qu'une seule détection (éviter les effets de *rebonds* dus, par exemple, à la troncature des filtres).

Ces 3 critères s'expriment par l'optimisation conjointe de 3 fonctionnelles qui permettent de définir le **filtre linéaire optimal** pour la détection d'une **marche d'escalier** sous l'hypothèse d'un bruit additif indépendant du signal ¹.

Si l'on considère que le filtre a pour réponse impulsionnelle $\psi(x)$, ces fonctionnelles s'écrivent :

$$\text{bonne détection : } \Sigma = \frac{\int_0^{\infty} \psi(x) dx}{\sqrt{\int_{-\infty}^{\infty} \psi^2(x) dx}}$$

$$\text{bonne localisation : } \Lambda = \frac{|\psi'(0)|}{\sqrt{\int_{-\infty}^{\infty} \psi'^2(x) dx}}$$

¹Par rapport aux approches proposées précédemment, on note que le filtre reste linéaire, mais que le critère d'optimisation est plus complet.

1	1	1
1	-2	1
-1	-1	-1

1	1	1
-1	-2	1
-1	-1	1

TAB. 2.3 – Masques de boussole directionnelle : il y a 8 filtres issus de ces 2 fenêtres par rotation.

$$\text{réponse unique : } \frac{|\psi'(0)|}{\sqrt{\int_{-\infty}^{\infty} \psi''^2(x) dx}} = k \frac{\int_{-\infty}^0 \psi(x) dx}{\sqrt{\int_{-\infty}^{\infty} \psi^2(x) dx}}$$

La résolution du système est assez complexe (on maximise le produit $\Sigma\Lambda$ sous la contrainte du $3^{ème}$ terme constant). Par ailleurs, Canny souhaite obtenir un filtre monodimensionnel (dans la direction orthogonale au contour), à réponse impulsionnelle finie (RIF) ; celui-ci a alors une expression complexe, composée de 4 termes, chacun combinant des lignes exponentielles et trigonométriques :

$$\psi(x) = a_1 e^{x/\sigma} \sin \omega x + a_2 e^{x/\sigma} \cos \omega x + a_3 e^{-x/\sigma} \sin \omega x + a_4 e^{-x/\sigma} \cos \omega x$$

où les coefficients a_i et ω sont déterminés à partir de la taille du filtre. Le paramètre σ est un paramètre de grande importance que nous retrouverons dans tous les autres filtres dérivés de l'approche de Canny. C'est un paramètre d'échelle qui indique en-deça de quelle distance deux contours parallèles seront confondus en un seul. Canny montre que la dérivée d'une gaussienne est une bonne approximation de son filtre.

$$\psi(x) \approx -x \exp\left(-\frac{x^2}{2\sigma^2}\right)$$

C'est donc la réduction à 1 dimension du filtre proposé par Marr à 2 dimensions (le maximum de la dérivée de la gaussienne est obtenu pour le passage par zéro du laplacien de gaussienne). Ce filtre donne des résultats de bonne qualité. Le critère de Canny (critère $\Sigma\Lambda$) vaut 1,12 pour le filtre RIF optimal, 0,97 pour le filtre gaussien. Remarquons cependant que le filtre gaussien n'est pas RIF.

2.3.2 Les filtres de Deriche et Shen et Castan

Au filtre de Canny, on préfère souvent le détecteur de Deriche, qui répond exactement aux mêmes critères de qualité que celui de Canny, mais qui possède une réponse impulsionnelle infinie (filtre RII). Il a pu donc être synthétisé de façon récursive particulièrement efficace [Deriche, 1987]. Le filtre de Deriche a une expression générale de la forme :

$$\psi(x) = -cx \exp(-\alpha|x|),$$

avec :

$$c = \frac{[1 - \exp(-\alpha)]^2}{\exp(-\alpha)}.$$

avec les critères même de Canny il est supérieur au filtre de Canny ($\Sigma\Lambda = 2$). Le paramètre α de Deriche représente alors l'inverse de l'écart type σ de la gaussienne du filtrage de Canny ($\alpha = \frac{\sqrt{\pi}}{\sigma}$). Le filtre de Deriche s'écrit de façon récursive en fonction des valeurs $f(x)$ de l'image, par l'intermédiaire de 2 filtres l'un décrivant l'image de gauche à droite et l'autre de droite à gauche :

$$z^+(x) = c\epsilon + 2\epsilon z^+(x-1) - \epsilon^2 z^+(x-2)$$

$$z^-(x) = -c\epsilon + 2\epsilon z^-(x+1) - \epsilon^2 z^-(x+2)$$

$$z(x) = z^+(x) + z^-(x)$$

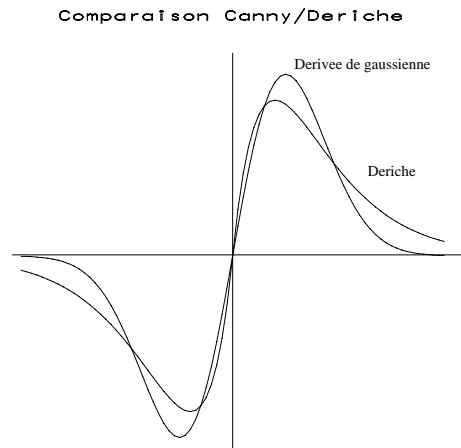


FIG. 2.7 – Le filtre de Deriche : $x \cdot \exp(-|x|)$ et la dérivée de la gaussienne : $\frac{x}{\sqrt{\pi}} \cdot \exp(-x^2/\pi)$.

avec :

$$\epsilon = \exp(-\alpha)$$

Le filtre de Shen et Castan [Shen et Castan, 1986] possède également de très bonnes qualités de détection, il s'apparente à la même famille des filtres exponentiels (mais comme il n'est pas dérivable à l'origine, il ne se prête pas à une comparaison avec le critère de Canny) :

$$f(x) = c \cdot \text{sgn}(x) \exp(-\alpha|x|)$$

Le filtre de Shen et Castan, en raison de sa forme à l'origine, donne une localisation très précise des contours, mais il est sensible aux bruits.

Dans [Demigny et Kamlé, 1997], Demigny a étendu les critères d'optimalité des filtres de Canny à des images discrètes. Il les a ensuite étendus à des profils en rampe et non plus en marche d'escalier.

2.3.3 L'extension à 2D

Tous ces filtres sont mono-dimensionnels, le filtrage 2D est donc obtenu par l'action de deux filtres croisés, l'un en x , l'autre en y .

Pour définir l'un de ces deux filtres, lorsque l'on a choisi le gabarit du filtre dans la direction perpendiculaire au contour, afin de mieux intégrer le signal, on utilise, dans la direction du contour, une **fonction de prolongement** qui permet de filtrer les bruits. C'est souvent une gaussienne de même écart-type que celle qui est dérivée dans le détecteur. Par exemple la dérivée en x pourra être estimée par convolution de l'image par :

$$\psi(x, y) = -x \exp(-(\alpha x^2)) \exp(-(\alpha y^2))$$

Ces filtres ne sont donc généralement pas isotropes. C'est-à-dire qu'ils ne répondent pas de façon identique à des contours identiques mais d'orientations différentes. Les seules fonctions séparables et isotropes sont les fonctions gaussiennes. Mais même si l'on utilise des filtrages gaussiens, la mise en œuvre discrète du filtre sur une maille carrée rend les filtres anisotropes. Des travaux importants ont été consacrés à la réduction de cet effet [Kundur et al., 1999] utilisant une combinaison de deux détecteurs à $\pi/4$ l'un de l'autre.

En conclusion, dans une approche moderne de la détection des contours, un filtre de détection se compose de deux estimateurs de dérivées, l'un selon Ox , l'autre selon Oy . L'un de ces détecteurs (choisissons celui selon Ox) se compose du produit de 2 fonctions :

- selon Oy c'est une fonction passe-bas, symétrique, (la fonction de prolongement) dont l'étendue est fonction de l'importance du bruit et de la distance que l'on souhaite respecter entre deux contours ;
- selon Ox c'est une fonction passe-haut (anti-symétrique), souvent égale à la dérivée de la précédente.

Les couples « fonctions de prolongement - dérivées » sont typiquement issus de la gaussienne (Canny), de l'exponentielle décroissante (Shen et Castan) ou du produit de l'exponentielle par x (Deriche) (cf. figure 2.8), mais on peut également les concevoir très librement pour une application spécifique.

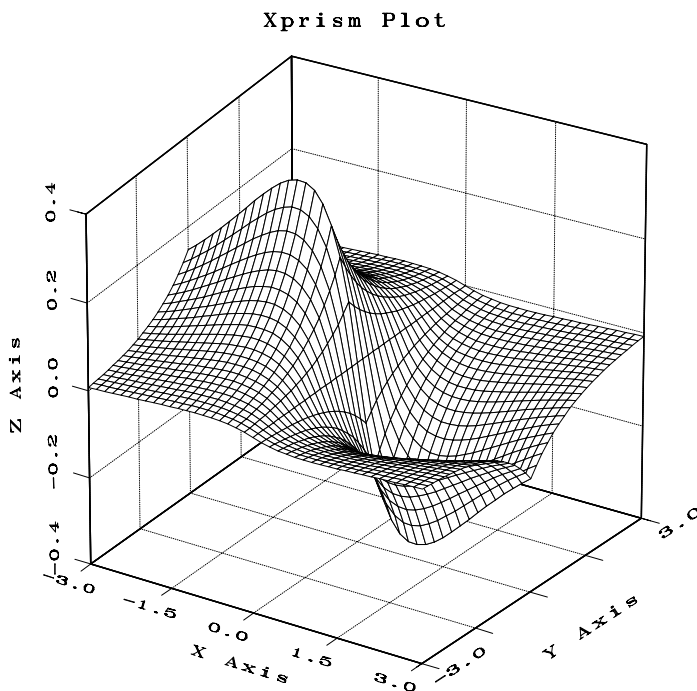


FIG. 2.8 – Profil général d'un filtre d'estimation de la dérivée dans la direction Ox . Selon Oy le filtre est un intégrateur (ici $\exp(-|y|)$), selon Ox , c'est un dérivateur (ici $x.\exp(-|x|)$).

2.3.4 Variantes autour des filtres analytiques

D'autres filtres aux performances comparables, mais résultats d'optimisations différentes, ont été proposés dans :

- [Spacek, 1986] : On modifie le critère de Canny, et on cherche pour filtre une spline cubique de forme $x^3 - 2x^2 + x$.
- [Petrou et Kittler, 1991] : On considère alors que le contour est également le résultat d'un flou, ce n'est donc plus une marche, mais une rampe. Les critères de sélection sont ceux de Spacek et le filtre est une combinaison non-linéaire de lignes trigonométriques et d'exponentielles décroissantes.
- [Sarkar et Boyer, 1991] : On modifie dans ce cas les critères de Canny, on adopte une approche variationnelle et on choisit la représentation récursive d'un filtre RII comme Deriche.

Une excellente comparaison et analyse des filtres de détection des contours est donnée dans [Cocquerez et Philipp, 1996].

Utilisant les résultats précédents ainsi que des méthodes éprouvées d'optimisation semi-globale développées dans les contours actifs (cf. section 2.4), des solutions variationnelles ont également été proposées qui sont fort efficaces et permettent d'ajouter des informations connues a priori sur l'allure des contours lors de leur détection. Voir par exemple [Fua et Leclerc, 1990].

2.4 Les contours actifs = les *snakes*

2.4.1 Une approche physique des contours

Une approche très différente des méthodes antérieures de détection de contours a été proposée en 1987 par Kass, Witkins et Terzopoulos [Kass et al., 1988] (voir aussi [Blake et Isard, 1998]), appelée contours actifs ou *snakes*. Il s'agit d'une méthode semi-interactive dans laquelle l'opérateur place dans l'image, au voisinage de la forme à détecter, une ligne initiale de contour. Cette ligne sera amenée à se déformer sous l'action de plusieurs forces :

- une énergie propre, assimilée à l'énergie mécanique de tension et de torsion d'une ligne matérielle,
- une énergie potentielle imposée par l'image qui vise à plaquer la courbe sur les contours,
- une énergie externe, introduite par l'utilisateur pour traduire les contraintes spécifiques du problème qu'il se pose.

Sous ces énergies, le contour *actif* va évoluer pour rechercher la position d'énergie minimale, qui sera ainsi un compromis entre les diverses contraintes du problème.

L'écriture formelle du problème passe par la définition paramétrique du contour, en fonction d'une variable s généralement l'abscisse curviligne :

$$v(s) = [x(s), y(s)]^t \quad s \in [0, 1]$$

$$E_{totale} = \int_0^1 [E_{interne}(v(s)) + E_{image}(v(s)) + E_{externe}(v(s))] ds \quad (2.6)$$

avec :

$$E_{interne} = \alpha(s) \left(\frac{dv}{ds} \right)^2 + \beta(s) \left(\frac{d^2v}{ds^2} \right)^2$$

où :

- la première dérivée prend en compte les variations de longueur de la courbe, (c'est donc un terme de tension (résistance à la rupture), qui est contrôlé par l'élasticité que l'on attribue au contour),
- tandis que la seconde exprime les variations de la courbure (c'est un terme de flexion contrôlé par la raideur du contour).

Ces deux termes agiront donc pour produire une courbe régulière (cf. figure 2.9).

Le second terme d'énergie : E_{image} caractérise les lignes que l'on souhaite suivre. Dans le cas de détection de contours, ce sont des lignes de fort gradient, il vaut donc généralement :

$$E_{image} = -\nabla f$$

mais il peut être adapté pour suivre les maximums des niveaux de gris (dans ce cas $E_{image} = f$), ou toute autre fonction définie à partir de ceux-ci.

Enfin le dernier terme d'énergie : $E_{externe}$ est choisi par l'utilisateur. Il peut avoir des formes très variées afin, par exemple, de contraindre le contour à ressembler à un gabarit donné, à s'approcher d'un contour déjà détecté sur une autre image (suivi de séquences, ou images à 3D)².

²On notera cependant que, dans le cas de volumes numériques tridimensionnels $f(x, y, z)$, il est possible d'étendre les contours actifs à des surfaces actives pour lesquelles l'énergie s'exprime comme pour des plaques minces.

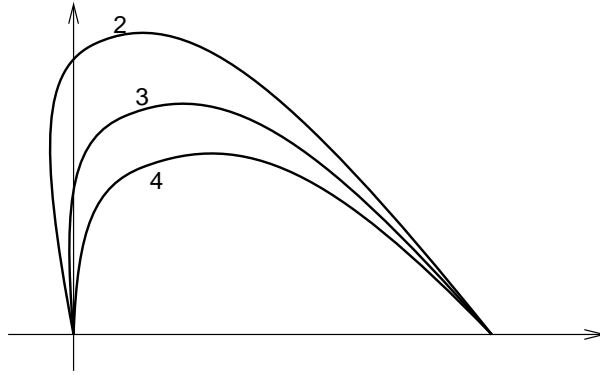


FIG. 2.9 – Déformation d'un contour actif à extrémités fixes pour des valeurs du rapport α/β (élasticité/raideur) égales à 2, 3 et 4, en l'absence d'une énergie liée à l'image.

En l'absence de ce dernier terme, et dans les nombreux cas où l'on recherche les lignes de fort gradient, l'équation 2.6 s'écrit alors :

$$E_{totale} = \int_0^1 \left[-\nabla f + \alpha(s) \left(\frac{dv}{ds} \right)^2 + \beta(s) \left(\frac{d^2v}{ds^2} \right)^2 \right] ds \quad (2.7)$$

2.4.2 Mise en équation des contours actifs

L'équation intégrale 2.7 est résolue généralement de façon variationnelle. On suppose que le contour évolue vers un minimum d'énergie, soit vers un zéro de $\frac{\partial E_{totale}}{\partial t}$. En désignant par v' et v'' les dérivées de v le long de la courbe, on obtient une équation différentielle vectorielle :

$$\gamma \frac{\partial v}{\partial t} + (\alpha v')' - (\beta v'')'' = \frac{\partial |\nabla f|^2}{\partial v}$$

où α , β et γ sont potentiellement variables le long de s . Il y a plusieurs façons de concevoir la discrétisation de la courbe :

1. selon les différences finies : les éléments de la courbe sont réduits en des points auxquels sont attachés les éléments mécaniques (masse, raideur, etc.) de la courbe considérée concentrée en ces points ;
2. selon les éléments finis on remplace chaque portion de courbe par le segment élémentaire et les éléments mécaniques sont calculés sur ces segments.

Nous ne décrivons ici que la première approche la plus fréquemment employée. Après discrétisation de la courbe en un nombre n de points et en posant :

$$V^t = [v_0^t, v_1^t, v_2^t, \dots, v_{n-1}^t]^t$$

et :

$$h_v = -|\nabla f(v)|^2$$

on obtient une équation matricielle de la forme :

$$(A + \gamma I)V^t = \gamma V^{t-1} - h_v(V^{t-1})$$

où A est une matrice pentadiagonale³ (figure 2.11) de taille $n \times n$ fonction de α et β , I est la matrice unité de taille $n \times n$ et γ exprime une *inertie* de la courbe aux déplacements. Sa résolution donne :

$$V^t = (A + \gamma I)^{-1}(\gamma V^{t-1} - h_v(V^{t-1}))$$

Si les paramètres α , β et γ sont invariants le long de la courbe, il suffit alors de calculer une seule fois $(A + \gamma I)^{-1}$ pour résoudre le système pour tous les temps t . Sinon, on inverse la matrice à chaque instant. Les choix des paramètres α , β et γ qui ne sont pas dictés par le problème demeurent souvent délicats pour garantir une convergence convenable, ainsi que le choix de l'estimateur de ∇f . L'utilisation d'un estimateur par filtre de Deriche est parfois préconisée.

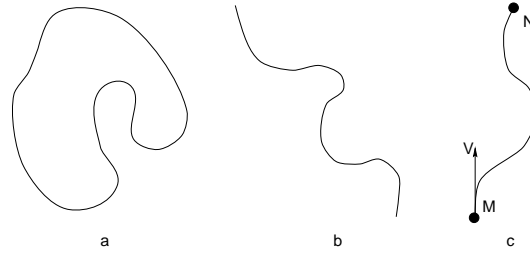


FIG. 2.10 – Contour actif fermé en a, ouvert à extrémités libres en b et ouvert à extrémités fixes en c (on impose en M une extrémité fixe et une orientation fixe donnée par le vecteur V).

$$\begin{bmatrix} 2\alpha+6\beta & -\alpha-4\beta & \beta & 0 & 0 & \dots & \dots \\ -\alpha-4\beta & 2\alpha+6\beta & -\alpha-4\beta & \beta & 0 & \dots & \dots \\ \beta & -\alpha-4\beta & 2\alpha+6\beta & -\alpha-4\beta & \beta & & \\ 0 & \beta & -\alpha-4\beta & 2\alpha+6\beta & -\alpha-4\beta & & \\ 0 & 0 & \beta & -\alpha-4\beta & 2\alpha+6\beta & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{bmatrix}$$

FIG. 2.11 – Un contour actif fermé est représenté par une matrice circulante pentadiagonale.

Malgré les difficultés de réglage de la convergence, les contours actifs apportent une solution heureuse à la détection de contour, intermédiaire entre les solutions purement locales (opérateurs de type Sobel) et les segmentations globales. Trois types de contours actifs différents sont utilisés (cf. figure 2.10) :

1. les contours actifs fermés (où $v_0^t = v_{n-1}^t$),
2. les contours actifs à extrémités libres,
3. les contours actifs à extrémités fixes (où les positions v_0 et v_{n-1} sont fixes dans le temps, ainsi que, potentiellement les premières dérivées en ces points).

Selon le type de contours actifs, les matrices A présentent la propriété d'être circulante (contours actifs fermés), Töplitz (contours actifs à extrémités libres) ou quelconques (figure 2.11).

³Dans le cas d'une discrétisation par éléments finis, la matrice est heptadiagonale

Laissé libre d'évoluer seul, en l'absence de forces d'attraction dues au gradient dans l'image, le contour actif a tendance à se réduire à un point s'il est fermé, à une droite s'il est à extrémités libres. Pour compenser cette tendance, on est parfois amené à introduire des forces internes de *gonflage*. On a alors affaire à des *ballons* [Cohen, 1991]. Comme les équations des forces de gonflage ne dérivent plus d'un potentiel, les écritures s'en trouvent un peu complexifiées.

Des solutions élégantes ont également été proposées pour rechercher la meilleure position du snake en en déplaçant ses nœuds sur toute la grille discrète de l'image par une technique de programmation dynamique [Amini et al., 1990] (cf. figure 2.12).

Par ailleurs, le couplage de techniques de contours actifs avec des méthodes s'appuyant sur des propriétés statistiques distinguant la forme de son fond, et optimisées par un maximum de vraisemblance au sens bayésien, a montré sa très grande efficacité, permettant de faire le lien entre techniques de contours et de régions [Chesnaud et al., 1999].

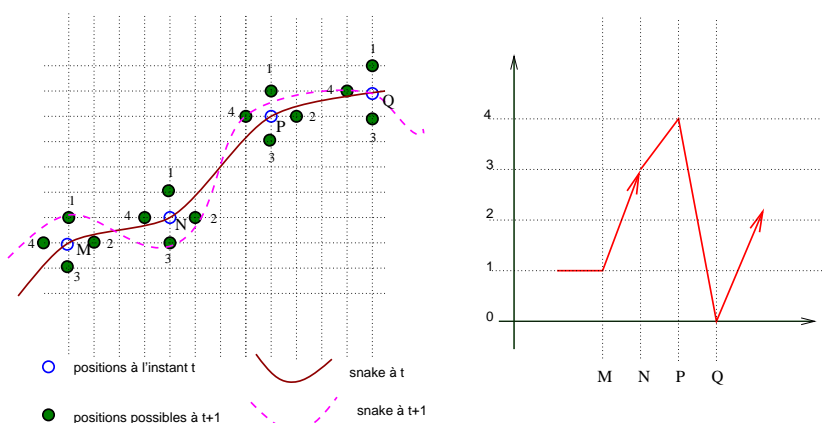


FIG. 2.12 – Dans la technique d'optimisation des contours actifs par programmation dynamique, chaque nœud M , N , P , Q peut se déplacer dans l'un des nœuds 1, 2, 3, ou 4, ou rester en place (position 0). L'optimisation se fait dans le graphe de droite par propagation de gauche à droite.

2.4.3 Les ensembles de niveaux (*level sets*)

Ce sont également des représentation variationnelles des contours qui conduisent donc à des solutions qui évoluent au cours du temps dans l'image, régies par un critère global. Ces modèles ont cependant le mérite remarquable de pouvoir changer de topologie si les contours l'imposent : par exemple un contour simple peut évoluer en deux contours séparés, ou, à l'inverse, deux contours séparés se réunir en un seul contour. Cela est rendu possible par l'utilisation de fonctions d'une dimension supérieure à celle des contours recherchés (si l'on cherche des lignes de contours, on introduira pour inconnue une surface de \mathbf{R}^3 , si on cherche des surfaces de \mathbf{R}^3 , on introduit des volumes de \mathbf{R}^4). Le contour est alors défini comme l'ensemble de niveau zéro (*level sets*) de cette fonction : $v(s)$ est une ligne de niveau $z = 0$ de la surface $z = f(x, y)$, souvent prise comme la distance au contour [Osher et Sethian, 1988, Sethian, 1996].

Sur le contour :

$$\frac{dz}{ds} = 0 = \frac{\partial z}{\partial x} \frac{dx(s)}{ds} + \frac{\partial z}{\partial y} \frac{dy(s)}{ds}$$

et la courbe de niveau évolue en fonction des itérations t selon l'équation fondamentale :

$$\frac{\partial z}{\partial t} = \|\text{grad}(z)\| \vec{N} \wedge \frac{d\vec{v}}{dt}$$

où \vec{N} représente la normale au contour v .

Des équations semblables existent à 3D [Zhao et al., 2000].

2.4.4 Les modèles géodésiques actifs

Les approches par contours actifs peuvent être reliées à la famille des problèmes d'optimisation géodésique dans des espaces dont la géométrie est fixée par la dynamique de l'image (cf. chapitre 3.5.1). Ces techniques, appelées modèles géodésiques actifs, ont été proposées dans [Caselles et al., 1997] et le lien avec les contours actifs y est discuté.

Le problème du snake, s'il n'y a pas d'élasticité, peut s'écrire comme la minimisation de l'énergie :

$$J_1(v) = \int_a^b |v'(s)| ds + \lambda \int_a^b \psi(|\nabla f(v(s))|)^2 ds$$

où ψ représente le détecteur de contour. Caselles a introduit une fonctionnelle du type :

$$J_2(v) = 2\sqrt{\lambda\alpha} \int_a^b |v'(s)| \psi(|\nabla f(v(s))|) ds$$

Minimiser J_2 revient à trouver le chemin de longueur minimale dans l'espace riemannien dont la métrique est induite par l'image. On a pu montrer un certain nombre d'équivalences entre les propriétés des minimums de J_1 et de J_2 dans [Aubert et Blanc-Féraud, 1999].

Les contours géodésiques utilisent des solutions proches de celles des ensembles de niveaux pour le codage et la représentation de la ligne de contours.

2.5 La poursuite et la fermeture des contours

Si l'on excepte les contours actifs et les passages par zéro des laplaciens, la plupart des détecteurs de contours fournissent des contours ouverts, c'est-à-dire qu'ils ne séparent pas les composantes de l'image dans des objets topologiquement distincts. Cela rend souvent plus difficile l'étape de reconnaissance des formes qui doit suivre. Pour cela, on a proposé plusieurs méthodes permettant d'obtenir des contours clos. C'est ce que l'on appelle la **fermeture de contours**.

De nombreuses solutions ont été proposées pour cette étape fondamentale, reposant sur des principes très différents (les contours actifs à extrémités fixes que nous venons de voir en sont un exemple). La qualité des résultats obtenus est généralement directement liée au coût informatique consenti, de très bonnes solutions existant, mais à des coûts prohibitifs pour la plupart des applications.

2.5.1 Méthodes de recherche dans des graphes

Ce sont les méthodes qui reposent sur les bases les plus solides et susceptibles de donner les meilleurs résultats, mais leurs coûts sont souvent très élevés, car elles se proposent d'explorer l'espace de toutes les solutions et de retenir la meilleure en fonction d'un critère que l'opérateur s'est donné. L'écueil de la recherche combinatoire exhaustive peut être évité par des recherches polynomiales assez complexes, surtout fondées sur la programmation dynamique.

La technique de base, présentée par Martelli [Martelli, 1972] et Montanari [Montanari, 1971], est celle d'une optimisation combinatoire minimisant la fonction de coût (ou maximisant une fonction de mérite). Les algorithmes de programmation dynamique (Viterbi, A^*), sont particulièrement bien adaptés à ce problème, mais, malgré leur

complexité polynomiale sont généralement limités par des profondeurs de recherche de quelques pixels si le graphe est la totalité de l'image.

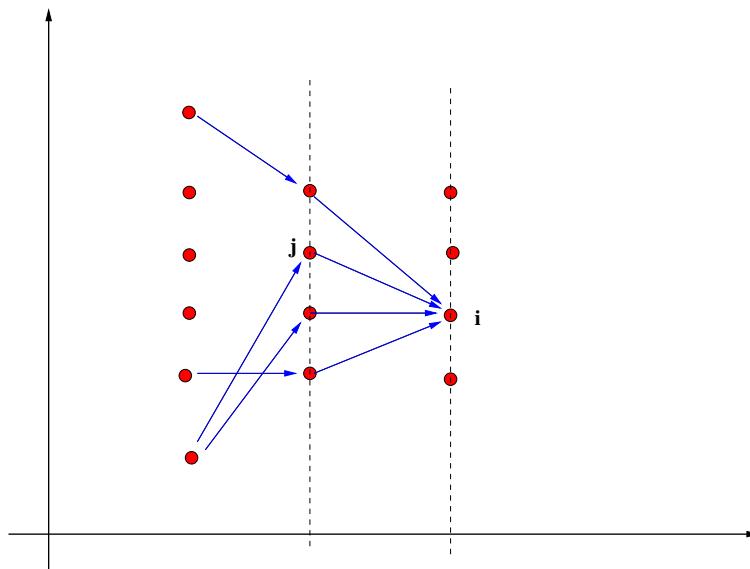


FIG. 2.13 – Résolution d'un problème de recherche de chemin optimal dans un graphe par programmation dynamique : la propagation se fait de la gauche vers la droite. On trouve le meilleur chemin arrivant à l'état i en optimisant la fonction de coût sur tous les états j qui ont été calculés à l'étape antérieure.

Dans une approche de type Viterbi, on considère l'évolution pas à pas d'un contour. À chaque pas le coût du contour est évalué pour toutes les provenances possibles du contour et n'est retenu pour tout point du parcours que le seul chemin qui minimise le coût pour arriver à ce point (cf. figure 2.13). On garde alors en chaque point ce coût minimum ainsi que l'adresse du *père* qui a permis d'y aboutir. Un nouveau pas est alors fait et toutes les destinations sont explorées que l'on peut atteindre en un pas à partir des points déjà atteints. La fonction de coût se calcule de la façon suivante :

$$\Gamma(i) = \gamma(i) + \min_{j \in \text{antécédents}} [\Gamma(j) + \delta(i, j)] = \gamma(i) + \epsilon(i)$$

où $\gamma(i)$ est inversement proportionnel à la qualité du point i comme point de contour et $\delta(i, j)$ exprime l'incompatibilité de i et j le long d'un contour. Le terme γ est un terme d'innovation le long du chemin tandis que ϵ est un terme de mémoire. La pondération de ces deux termes est un problème délicat en programmation dynamique.

Un contour est fermé lorsque le point i est confondu avec un point d'un autre contour. Le coût de la fermeture est alors l'intégrale le long de la chaîne des coûts de liaison δ et des coûts individuels γ . Pour éviter de trop pénaliser les chaînes longues, il est possible de calculer le *coût moyen* le long du parcours en pondérant le terme de mémoire ϵ par le chemin parcouru.

En limitant les domaines de recherche à des champs convexes, ou plus généralement en introduisant des limitations sur les types de chemin que l'on s'autorise à trouver, on peut étendre notablement la portée de ces fermetures (cf. figure 2.14). Des approches multi-résolutions permettent de travailler de façon hiérarchique, d'une fermeture grossière à une fermeture précise, en limitant progressivement le domaine de recherche à une bande étroite autour du contour détecté à l'étape précédente. Dans d'autres cas, on se limite à des horizons faibles pour entreprendre une recherche par programmation dynamique qui est validée, puis relancée à partir du résultat précédent.

Dans une recherche de type A^* , on dispose d'une estimée de la fonction de coût ($\hat{\Gamma} > \Gamma_{opt}$) et la fonction de coût doit être croissante le long d'une fermeture. On fait une recherche en largeur d'abord en comparant à chaque

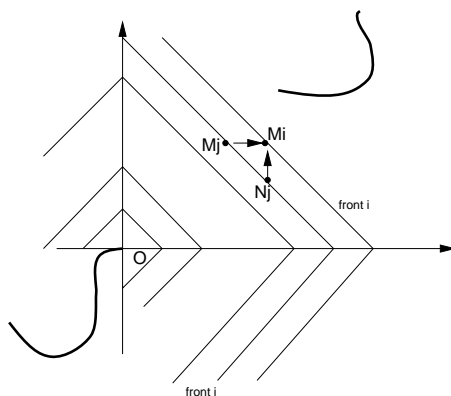


FIG. 2.14 – Exemple d’une fermeture de contours par recherche d’un chemin optimal dans un graphe. Le contour à fermer s’arrête en O . Les horizons de recherche à chaque étape de la programmation dynamique sont des carrés (fronts d’ordre i), centrés en O . Chaque point M_i peut avoir 2 antécédents (M_j et N_j). Une fermeture s’obtient lorsqu’un des points appartient à un autre contour. Pour éviter que le contour ne se ferme trop rapidement sur lui-même (près de O), il faut invalider les points du contour lui-même. Une telle propagation contraint les contours à une courbure toujours de même signe.

nœud la valeur locale du coût à cette estimée et on invalide le nœud si le coût dépasse l’estimée. On a tout intérêt à trouver donc une estimée très proche du coût optimal réel.

Les fonctions de coût sont les points déterminants de la qualité de la méthode. Cette fonction de coût doit être locale pour préserver le caractère markovien nécessaire à la programmation dynamique (le coût en un point ne dépend que de γ et du coût au point précédent), mais elle doit également incorporer toutes les informations que l’on connaît sur un *bon contour* : contraste entre les deux plages séparées par le contour bien sûr, mais aussi, longueur et courbure, voire écart à un contour modèle dans certains cas. C’est dans cette fonction de coût que l’on rajoute l’information qui a manqué au détecteur local placé en amont pour détecter le contour.

2.5.2 Les automates

Ces solutions, au contraire des précédentes, ne proposent que des solutions très sous-optimales, fondées sur de seuls critères locaux (il n’y a pas d’optimum global comme par programmation dynamique). Elles sont particulièrement rapides de mise en œuvre [Giraudon, 1987], mais peuvent parfois faire diverger la combinatoire si l’on gère mal la prolifération des chemins par ramification. Les résultats d’une détection par masquage adapté sont les candidats idéaux à ces techniques de fermeture puisqu’ils fournissent des directions de recherche.

Les diverses méthodes se distinguent par le choix des successeurs d’un point (parmi des configurations prédéfinies dans [Robinson, 1977, Cocquerez et Philipp, 1996, Nevatia et Babu, 1980], par prédiction de type Kalman) et le mode de balayage de l’image (par exemple 1 balayage dans [Nevatia et Babu, 1980], 2 balayages, l’un horizontal l’autre vertical dans [Cocquerez et Philipp, 1996], traitement de listes chaînées dans [Giraudon, 1987]). Il est également important de définir des fonctions de mérite pour valider les chaînes ainsi trouvées, car les automates, comme la programmation dynamique, trouvent toujours un contour, mais ne garantissent pas toujours qu’il correspond véritablement au contenu de l’image.

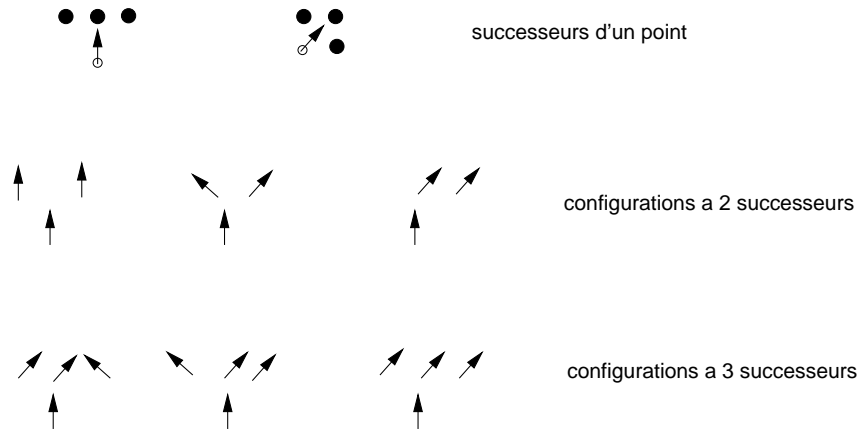


FIG. 2.15 – Configurations retenues par l’automate de Nevatia et Babu pour orienter un contour vers 3 successeurs au plus.

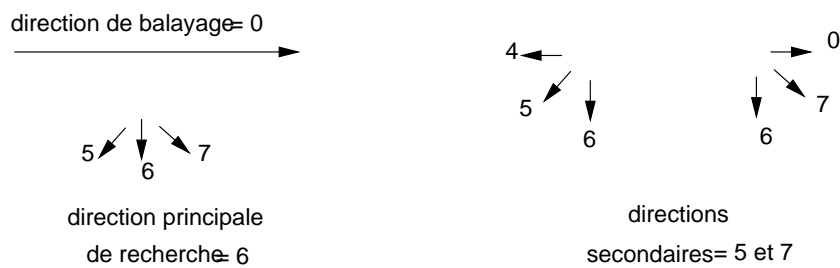


FIG. 2.16 – L’automate de Cocquerez lors d’un balayage horizontal utilise les directions 5, 6 et 7 de Freeman (cf. chapitre 5.4) comme directions principales. Lorsque l’une d’elle a été retenue, les directions adjacentes (par exemple le 4 et le 6 pour la direction principale 5) sont utilisées pour élargir la recherche. Cocquerez propose de faire 2 balayages successifs, l’un horizontal, l’autre vertical.

Chapitre 3

La segmentation par régions

Chapitre rédigé par Henri MAÎTRE

L'approche duale de la détection des contours pour la décomposition d'une image en ses formes élémentaires est l'approche par régions. Elle repose sur la recherche de zones possédant des attributs communs, soit de luminosité, soit, plus rarement, de textures (nous n'aborderons pratiquement pas ce second point dans ces lignes, mais au chapitre 4).

Nous allons voir ici, tout d'abord, les méthodes utilisant l'histogramme, qui sont souvent très proches des méthodes de classification conventionnelles, mais dans lesquelles serait renforcé l'aspect de cohérence de zone. Nous verrons ensuite les méthodes pyramidales, prototypes des méthodes de subdivision hiérarchique descendante, puis les méthodes de croissance de régions, qui s'inspirent des méthodes de conquête de l'optimisation combinatoire, et les méthodes de partage et réunion qui tirent profit simultanément des avantages des deux méthodes précédentes.

Enfin, nous donnerons très brièvement un aperçu de méthodes à base de règles qui se proposent de combiner approches par contours et approches par régions dans un seul vaste programme de segmentation.

3.1 Les méthodes sur histogramme

Ces méthodes sont de mise en œuvre assez simple et de performances souvent réduites car elles ne tirent pas profit de l'aspect spatial de l'information d'image. Elles sont recommandées dans les cas suivants :

1. lorsque les images présentent des classes évidentes : documents écrits ou schémas en noir et blanc ou en couleur, objets très contrastés (par exemple cellules d'une biopsie ou avion sur un ciel), etc.
2. lorsque les images sont définies sur de nombreux canaux (images multi- ou hyper-spectrales), ce qui enrichit l'information portée par l'histogramme.

L'idée générale de ces méthodes consiste à isoler des pics de l'histogramme. À une dimension on procède donc à des seuillages ou des multi-seuillages. À n-dimensions on procède à des classifications [Dubuisson, 1990].

3.1.1 Avec apprentissage

Seuillage bayésien

Si l'on dispose d'une connaissance sur les classes que l'on recherche (en particulier la probabilité d'occurrence d'un niveau de gris pour chaque classe, et la probabilité à priori des classes), alors on peut aisément se replacer

dans les conditions de la théorie bayésienne de la décision [Dubuisson, 1990, Duda et Hart, 1973] et choisir les seuils qui minimisent le coût des fausses erreurs.

Pour des histogrammes à 1 dimension et pour 2 populations X et Y , en dénotant par :

- $P(X)$ et $P(Y)$ les probabilités à priori des classes X et Y (sous la contrainte $P(X) + P(Y) = 1$),
- $P(n/X)$ et $P(n/Y)$ les probabilités conditionnelles qu'un pixel de la classe X ait un niveau de gris n (et idem pour Y),
- C_X et C_Y les coûts des mauvaises classifications de X et Y ,

en supposant ces grandeurs toutes connues, et sous l'hypothèse de stationarité de l'image (c'est-à-dire que les propriétés statistiques sont invariantes en tout point de l'image), le seuil optimal est défini comme celui minimisant le coût global :

$$\Gamma(s) = C_x \int_s^N P(X).P(n/X)dn + C_y \int_0^s P(Y).P(n/Y)dn \quad (3.1)$$

Ce seuil est obtenu en testant pour tout n le rapport de vraisemblance $\Lambda(n)$ face au seuil s :

$$\Lambda(n) = \frac{P(n/Y)}{P(n/X)} \quad \text{et} \quad s = \frac{C_y P(Y)}{C_x P(X)}$$

Seuillage de Neyman-Pearson

Une autre décision peut se faire selon le critère de Neyman-Pearson. Dans ce cas on s'intéresse en particulier à une classe (ici on choisit X). On définit la probabilité de fausse alarme pour cette classe :

$$P_f = \int_0^s P(Y)P(n/Y)dn$$

On maximise alors la probabilité de détection pour une probabilité de fausse alarme donnée. Dans ce cas on considère que la fausse alarme est l'erreur la plus grave. On fixe sa probabilité à une valeur acceptable : $P_f = \alpha$. La probabilité de détection de X est donnée par :

$$P_d = \int_0^s P(X)P(n/X)dn$$

Le seuil de Neyman-Pearson est donné par la méthode du lagrangien :

$$L = P_d - \lambda(P_f - \alpha) \quad (3.2)$$

$$= \lambda\alpha - \int_0^s [P(n/Y)P(Y) - P(n/X)P(X)]dn \quad (3.3)$$

et la décision se fait en comparant le rapport de vraisemblance $\Lambda(n)$ à α .

Cette décision est moins sensible aux probabilités à priori et conduit en particulier à des décisions plus proches de nos choix intuitifs que la décision bayésienne dans le cas où un événement est très rare.

D'autres critères existent bien sûr qui s'appuient sur d'autres choix statistiques (par exemple minimum d'entropie).

3.1.2 Seuillage sans apprentissage

Lorsque l'on dispose du seul histogramme pour extraire des classes, on recherche généralement des modes de cet histogramme qu'on isole par des seuillages au creux des vallées. Souvent les histogrammes sont trop irréguliers et il convient alors de les filtrer, soit par des fenêtres glissantes, soit par des équations de diffusion (éventuellement isotropes). De nombreuses règles empiriques ont été proposées pour choisir les seuils automatiquement qui ne sont guère généralisables [Kittler et al., 1985, Weszka, 1978].

3.1.3 Méthodes de classification

Disposant d'un histogramme, éventuellement multidimensionnel, la plupart des techniques de classification s'appliquent à sa segmentation. Les plus utilisées sont :

- les techniques de nuées dynamiques (*k-means*) qui procèdent alternativement en classifiant au plus proche voisin le nuage des points, selon une distance à des noyaux donnés, puis en estimant la position des meilleurs noyaux de ces classes ainsi obtenues. Il est important pour cette méthode de disposer du nombre de classes recherchées. Si l'on ne le connaît pas, on choisit souvent des critères entropiques (comme le critère d' Akaike ou le critère de Hannan et Quinn [Olivier et al., 1997]) qui permettent d'évaluer des classifications obtenues avec des nombres de classes différents¹.
- les réseaux neuromimétiques et en particulier les cartes de Hopfield.
- et, si l'on dispose d'un certain nombre d'échantillons déjà classés, les plans, ou les courbes, séparateurs, ainsi que les k-plus-proches voisins.

Dans les espaces de grande dimension (imagerie hyperspectrale par exemple), on peut avoir intérêt à réduire tout d'abord la dimension des espaces pour éviter d'avoir à estimer des probabilités très faibles. On peut le faire par ACP (analyse en composantes principales) ou par analyse de Karhunen-Loeve. Ces deux transformations sont identiques à une normalisation près des axes. Elles procèdent en projetant le nuage de points sur le sous-espace construit à partir d'un nombre limité des vecteurs propres de la matrice de covariance des données.

Dans les espaces de dimension p , les distances utilisées entre deux nuages de points caractérisés par leur moyenne (vectorielle) m_i et leur matrice de covariance Γ_i sont, par ordre de complexité croissante [Fukunaga, 1990, Zhang et Modestino, 1990] :

- la distance euclidienne qui pondère également toutes les variables :

$$(m_1 - m_2)^t (m_1 - m_2)$$

- la distance de Mahalanobis (par exemple du nuage 2 par rapport au nuage 1), qui rend compte de l'orientation des inerties des nuages dans l'espace :

$$(m_1 - m_2)^t \Gamma_2^{-1} (m_1 - m_2)$$

- la distance de Bhattacharyya, qui permet de distinguer des lois de mêmes moyennes, mais de variances différentes. :

$$1/4(m_1 - m_2)^t (\Gamma_1 + \Gamma_2)^{-1} (m_1 - m_2) + \frac{1}{2} \ln \frac{\frac{1}{2} |\Gamma_1 + \Gamma_2|}{\sqrt{|\Gamma_1| |\Gamma_2|}}$$

3.1.4 Sélection sur l'histogramme et dans l'image

Le défaut des approches par classification est de complètement négliger les relations spatiales entre pixels, pour ne s'attacher qu'à leurs propriétés de radiométrie. Pour pallier ce défaut, dans une approche proposée dans [Ohlander et al., 1978] et souvent reprise, on procède de façon itérative dans l'espace image et sur l'histogramme :

1. sur l'histogramme on sélectionne un mode isolé,
2. parmi les zones de l'images contribuant à ce mode on sélectionne la zone connexe la plus grande.

On itère ce processus, soit en subdivisant à nouveau l'histogramme de la zone retenue, soit en s'occupant d'un autre mode de l'histogramme original.

3.1.5 Sélection sur histogramme et régularisation

Mais si l'on souhaite améliorer les propriétés spatiales des zones obtenues par sélection de modes sur l'histogramme ou par classification, l'une des méthodes les plus appropriées est de modéliser le problème par un champ

¹On note cependant avec [Olivier et al., 1997] que ces modèles ont tendance à sur-évaluer le nombre de classes trouvées.

markovien (voir chapitre ??). On considère que les régions forment une partition sur l'image (cf. section 3.2). Chaque région est représentée par une fonction caractéristique et identifiée par une étiquette $\lambda_i \in E$. Le champ d'étiquettes est un champ caché à l'utilisateur qui n'observe que la réalisation bruitée de l'image $f(x, y)$ en chaque pixel, et qui va essayer d'estimer la meilleure distribution des étiquettes λ_i connaissant les n (généralement selon le critère du Maximum A Posteriori de Λ (MAP)).

Dans la formalisation markovienne on passe d'une représentation probabiliste à une représentation en énergie (généralement au moyen d'une formalisation bayésienne de la probabilité à postériori). Très souvent l'énergie (représentant la probabilité à postériori de la classe sachant l'observée) est constituée de deux termes : l'un traduit la similarité de la classification aux données que l'on observe, c'est le terme d'attache aux données (probabilité des observées conditionnellement aux classes), le second exprime les a priori que nous avons sur les classes (par exemple classes compactes, aux contours réguliers ou de forme prédéfinie, dans des relations de voisinage particulières, probabilité à priori de classe, etc.). Des hypothèses peuvent également être faites sur les distributions des niveaux de gris (par exemple lois gaussiennes sur les régions), des étiquettes (dépendance ou indépendance spatiale), ainsi que des distributions conditionnelles des niveaux de gris sachant les étiquettes.

Dans la démarche la plus simple on utilise alors souvent un terme d'attache aux données de la forme :

$$U_0 = (f(x, y) - \mu_i)^2$$

où $f(x, y)$ est le niveau de gris du pixel, μ_i le niveau de gris de la classe λ_i dans laquelle on a placé le pixel (x, y) . Un tel terme est équivalent à une probabilité gaussienne d'appartenance à la classe λ_i :

$$p(f(x, y)|\lambda_i) \propto \exp(-U_0) = \exp\left[-\frac{(f(x, y) - \mu_i)^2}{2\sigma_i^2}\right]$$

Le terme de régularisation est calculé sur les cliques $\{c\}$ constituant le voisinage que l'on souhaite donner à (x, y) (cf. figure ??). Il prend en compte la compatibilité $\Phi(\lambda_i, \lambda_k(x', y'))$ entre l'étiquette λ_i attribuée à (x, y) et celle $\lambda_k(x', y')$ attribuée à (x', y') :

$$E_1 = \sum_{(x', y') \in \{c\}} \Phi(\lambda_i, \lambda_k(x', y'))$$

Par exemple, dans le cas le plus simple, on peut attribuer une énergie négative si $\lambda_i = \lambda_k$ et positive dans le cas contraire (modèles de Potts et d'Ising, voir chapitre ??). Des modèles plus complexes peuvent prendre en compte des compatibilités plus subtiles (classes plus ou moins proches). On recherche alors un minimum de $U = U_0 + \theta U_1$ par toutes les techniques usuelles d'optimisation itératives (ICM, recuit simulé, etc.). Le coefficient θ permet de pondérer les effets de la régularisation par rapport à ceux de la classification. Pour un θ fort, on obtiendra des classes très compactes et grandes au détriment d'une bonne séparation des classes sur l'histogramme [Derin et Elliott, 1987, Lakshmanan et Derin, 1989, Heitz et al., 1994, Won et Derin, 1992].

Dans d'autres approches, on peut se dispenser d'introduire des étiquettes de classes en introduisant explicitement un processus de bord, (cf. section ?? et [Geman et Geman, 1984]), processus b discret binaire (0 ou 1) ou continu (entre 0 et 1) qui prend, généralement, ses valeurs entre les pixels. Dans le cas binaire, lorsqu'il vaut 1 le voisin concerné n'a plus d'effet sur le site, selon la formule :

$$U_1 = \sum_{(x', y') \in \{c\}} (1 - b) \cdot \Phi(g_i, g_k(x', y'))$$

C'est donc l'équivalent d'un contour introduit dans l'image. Mais on pénalise usuellement l'introduction d'un contour par un terme de la forme :

$$U'_1 = \sum_{(x', y') \in \{c\}} b$$

ou par un coût de type MDL (Minimum Description Length) (cf. section 3.4).

Enfin, on peut également supprimer le processus de bord en le « cachant » à l'intérieur du potentiel U_1 . Pour cela on choisit une fonction Φ qui sature pour des valeurs fortes des différences entre niveaux de gris de sites voisins. On perd alors, avec la convexité de la fonctionnelle d'énergie, des bonnes propriétés de convergence du champ markovien et l'optimisation devient plus complexe (cf. figure ?? et section ??) [Blake et Zisserman, 1987, Black et Rangarajan, 1996].

3.2 Les méthodes par transformation de régions

Les méthodes que l'on va examiner maintenant s'appuient toutes sur la notion de **prédicat** et sur celle de **partition**.

Un prédicat \mathcal{P} est une proposition logique dont la valeur dépend de son argument. Nous nous intéresserons aux prédicats sur les régions \mathcal{R}_i de l'image.

Le prédicat de base de la segmentation est : *la région \mathcal{R}_i est homogène*. Pour vérifier le prédicat, parmi les arguments les plus utilisés en segmentation d'images, nous retiendrons les suivants :

- le contraste sur la région : $\mathcal{P}(\mathcal{R}_i) = \text{vrai} \iff \max_{\mathcal{R}_i}[f(x, y)] - \min_{\mathcal{R}_i}[f(x, y)] < \sigma$
- l'écart-type de la région : $\mathcal{P}(\mathcal{R}_i) = \text{vrai} \iff \sqrt{\frac{1}{N} \sum_{\mathcal{R}_i} (f(x, y) - m)^2} < \sigma$ (avec $N = \text{Card}(\mathcal{R}_i)$ et $m = \frac{1}{N} \sum_{\mathcal{R}_i} n(x, y)$),
- la distance interquartile sur la région, c'est-à-dire la distance séparant les 25 % inférieurs des 25 % supérieurs de l'histogramme,
- les différences limitées : $\mathcal{P}(\mathcal{R}_i) = \text{vrai} \iff \forall (k, j) \text{ voisins}, |f(k) - f(j)| < \sigma$,
- l'entropie : $\mathcal{P}(\mathcal{R}_i) = \text{vrai} \iff - \sum_{\mathcal{R}_i} p(f) \log(p(f)) < \sigma$.
- etc.

Un autre prédicat également utilisé est : *la région \mathcal{R}_i est distincte de ses voisines*. On utilise des arguments qui mettent en jeu les différences de valeurs moyennes, les distances inter-médianes, les contrastes moyens ou minimums aux frontières, etc.

Une partition Π est un ensemble de sous-ensembles \mathcal{R}_i (appelées régions) de l'image, vérifiant :

$$\Pi \iff \begin{cases} \forall i, j \ \mathcal{R}_i \cap \mathcal{R}_j = \emptyset \\ \bigcup_i \mathcal{R}_i = \text{support}(\text{image}) \\ \forall i \ \mathcal{R}_i \neq \emptyset \end{cases}$$

Lorsque toutes les régions d'une partition vérifient le prédicat, on dit que la partition le vérifie. La partition triviale où tous les pixels sont dans des régions différentes de cardinal 1 vérifie tout prédicat qui peut se vérifier sur une image. Tout prédicat qui ne peut être vérifié par la partition triviale est impossible. On connaît donc au moins une partition qui vérifie tout prédicat non impossible.

Il existe bien sûr un très grand nombre de partitions d'une image et il existe également généralement un très grand nombre de partitions qui vérifient le prédicat (il suffit, partant d'une telle partition, de subdiviser n'importe quelle région pour obtenir une nouvelle partition vérifiant le prédicat). On ne sait pas trouver toutes les partitions vérifiant le prédicat (la combinatoire est généralement inaccessible) mais on ne sait non plus généralement pas choisir entre des partitions vérifiant un prédicat. Des critères empiriques peuvent être utiles :

- le cardinal de la partition (à minimiser),
- la taille de la plus petite région (à maximiser),
- une « distance entre régions » (par exemple somme des distances entre zones adjacentes) (à maximiser).

Dans de nombreux cas par exemple, on ne recherche que des partitions maximales, c'est-à-dire telles que deux régions adjacentes ne vérifient pas le prédicat :

$$\forall (i, j : i \text{ et } j \text{ adjacents}) \Rightarrow \mathcal{P}(\mathcal{R}_i \cup \mathcal{R}_j) = \text{faux}$$

En l'absence de stratégie pour trouver la meilleure partition vérifiant un prédicat donné, nous allons décrire des stratégies de base qui sont très souvent utilisées.

3.2.1 La croissance de région

Les méthodes de croissance de région (*region growing*) n'aboutissent pas à des partitions car la propriété $\bigcup_i \mathcal{R}_i = image$ n'est généralement pas satisfaite. Partant de **germes**² on applique successivement à l'image des prédicats plus sévères que le prédicat \mathcal{P} . Ainsi on commence à associer aux germes les seuls pixels qui sont en très bon accord avec le prédicat. On réduit cette sévérité progressivement, et on se rapproche ainsi du prédicat objectif. La décision d'associer un pixel à une région se fait alors le plus souvent sans ambiguïté à moins que ses « distances » à deux régions soient égales (et en ce cas un choix quelconque est peu décisif). Plus importante est la décision de regrouper 2 régions qui sont adjacentes et vérifient le prédicat. Il a été montré qu'il convient alors d'être assez sévère pour fusionner des régions et qu'il est souvent préférable de traiter ce point lors d'une étape ultérieure en acceptant donc une **sur-segmentation** de l'image, plutôt qu'une fusion abusive qui ne serait plus récupérable.

Les méthodes de croissance de région cessent souvent de créer de nouveaux germes bien avant que le prédicat ne vaille \mathcal{P} et rejettent les régions de cardinal 1. Ainsi tous les pixels ne se retrouvent pas dans des régions (cf. figure 3.1) [Adams et Bishof, 1994, Beaulieu et Goldberg, 1989, Chang et Li, 1994].

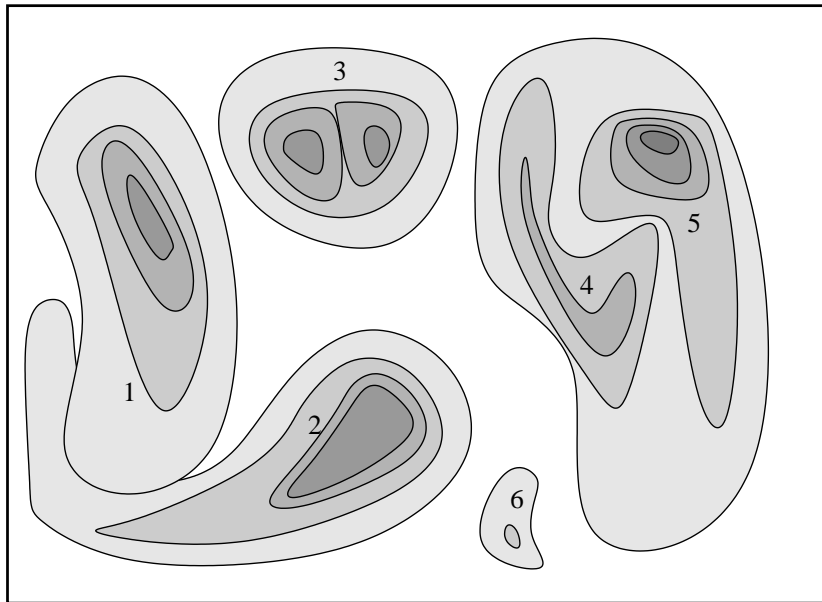


FIG. 3.1 – Croissance de région : 6 régions ont été obtenues par 4 prédicats successifs. La zone 6 n'a été créée qu'au second prédicat. La zone 3 regroupe 2 sous-régions qui ont été fusionnées car leur union vérifie le postulat et des critères annexes sur la forme résultante sont vérifiés. Ce n'est pas le cas de 1 et 2.

3.2.2 Le partage de région

Dans une démarche par partage de région (*region splitting*), on part de l'image entière (à laquelle généralement le prédicat \mathcal{P} ne s'applique pas). On appelle \mathcal{R} cette région. On lui applique plusieurs divisions δ produisant de nouvelles régions \mathcal{R}_i^δ . On teste \mathcal{P} sur chaque \mathcal{R}_i^δ , et on retient la meilleure subdivision δ , c'est-à-dire :

²les germes sont souvent des régions où le prédicat est trivialement vérifié, par exemple des zones où l'image est constante

- celle qui conduit à des sous-régions vérifiant toutes \mathcal{P} ,
- ou celle qui donne le plus de sous-régions vérifiant toutes \mathcal{P} ,
- ou, si aucune \mathcal{R}_i^δ ne vérifie \mathcal{P} , celle qui fournit la meilleure valeur à un critère \mathcal{E} dit critère d'échec.

Pour δ choisi, chaque sous-région \mathcal{R}_i^δ ne vérifiant pas \mathcal{P} devient alors une région \mathcal{R} passible du traitement ci-dessus.

Les stratégies de subdivision sont peu nombreuses, on utilise généralement les 2 bipartitions régulières horizontale ou verticale (cf. figure 3.2), parfois des tripartitions régulières.

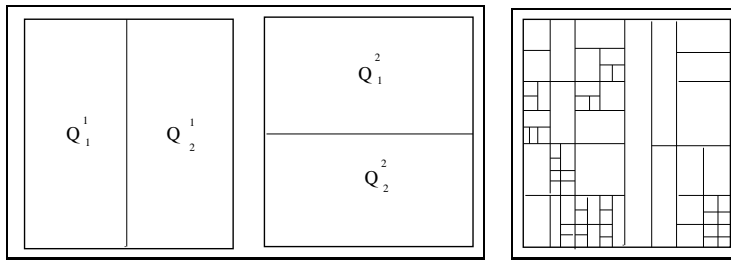


FIG. 3.2 – A gauche : partition d'une zone : on choisit entre la partition verticale ou la partition horizontale. A droite, à la fin du partage, l'image partitionnée.

Les critères d'échec \mathcal{E} sont souvent les mesures de variance ou de dynamique qui sont testées dans les prédicats \mathcal{P} . Lorsque toutes les régions ont été récursivement testées, il peut se trouver que des zones adjacentes le long d'une frontière d'ordre supérieur vérifient le prédicat. On peut alors les réunir avec les mêmes précautions que celles signalées au paragraphe 3.2.1.

Les algorithmes de partage sont mal adaptés à une mise en œuvre informatique sur machine séquentielle car les calculs effectués sur tous les pixels de la zone \mathcal{Q} ne sont généralement pas réutilisables au niveau inférieur \mathcal{Q}_i .

3.2.3 La réunion de région

Les techniques de réunion (*region merging*) prennent l'exact contre-pied des précédentes. Ce sont des méthodes ascendantes (*bottom-up*) lorsque les autres étaient descendantes (*top-down*). Les pixels sont visités systématiquement. Pour chaque carré de 2×2 pixels le postulat \mathcal{P} est testé, et s'il est accepté les pixels sont regroupés en une région. Après le parcours de toute l'image, les groupes de 2×2 régions se voient appliquer le même test et, éventuellement, les mêmes conséquences (réunion en une région de niveau 2) [Zhu et Yuille, 1996].

Les tests de réunion de région sont fréquemment faits sur des tests statistiques [Saporta, 1978]. On se place souvent dans l'hypothèse de bruits gaussiens sur des fonctions à valeur moyenne constante. Les tests les plus utilisés sont :

- le Chi2 (χ^2),
- le test de Wilcoxon, il travaille sur les pixels triés par ordre croissant de niveaux de gris des deux régions. Pour chaque pixel de la liste issue de la région 1 on compte combien de pixels de 2 sont plus sombres. Ces nombres sont additionnés pour donner une variable U de distribution asymptotiquement gaussienne de moyenne $n_1 n_2 / 2$ et de variance $n_1 n_2 (n_1 + n_2 + 1) / 12$. On teste donc cette valeur U face à sa distribution asymptotique.
- le test de Student d'égalité des espérances, on teste la variable de Student :

$$T(n_1 + n_2 - 2) = \frac{(m_1 - m_2)\sqrt{n_1 + n_2 - 2}}{\sqrt{(n_1\sigma_1^2 + n_2\sigma_2^2)(1/n_1 + 1/n_2)}}$$

– le test de Fisher-Snedecor d'égalité des moyennes et des variances, on teste la variable de Fisher :

$$F(n_1 - 1; n_2 - 1) = \frac{n_1 \sigma_1^2 (n_2 - 1)}{n_2 \sigma_2^2 (n_1 - 1)}$$

3.2.4 Les pyramides

Afin de bénéficier des avantages des deux méthodes précédentes, Horowitz a proposé une approche par pyramide [Horowitz et Pavlidis, 1976]. L'image est représentée sur une pyramide appelée *quad-tree*, constituée de N niveaux, l'image originale étant au niveau 0. Chaque pixel au niveau ν a 4 fils au niveau $\nu - 1$. Le pixel au niveau ν est la moyenne de ses 4 fils. Un schéma plus complet a été proposé par Burt [Burt, 1981] où les passages entre niveaux se font par filtrage gaussien. Une abondante littérature a été proposée depuis pour améliorer ces filtrages par une approche en ondelettes.

La segmentation procède directement à un niveau intermédiaire (par exemple $\nu = 2$), et tous les pixels fils sont testés avec le prédicat \mathcal{P} . Si le prédicat n'est pas vérifié, le pixel considéré est étiqueté comme candidat à un partage. Sinon il est candidat à une réunion avec ses voisins. C'est donc une technique de partage et réunion, où l'on profite du passage au niveau supérieur (ici $\nu = 3$) pour accélérer la procédure (cf. figure 3.3).

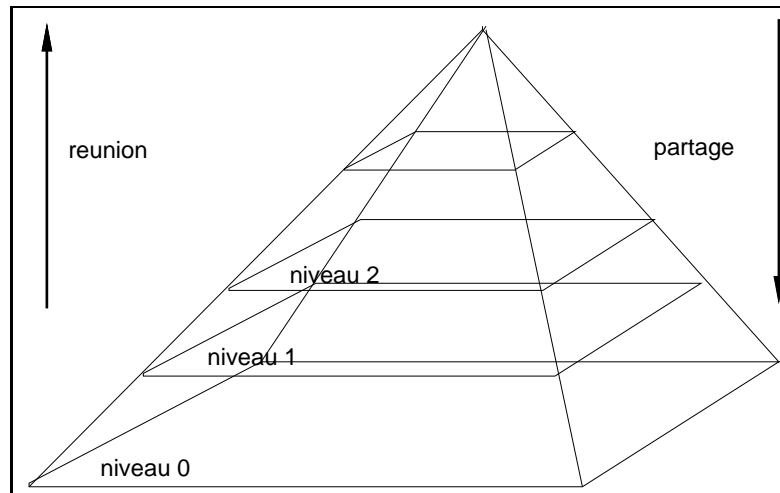


FIG. 3.3 – Méthode de pyramidage par quad-tree.

Mais la stratégie du quad-tree est trop rigide et conduit à des partitions trop régulières que les techniques de croissance ne permettent plus de rattraper. Des techniques inspirées de la pyramide ont été proposées qui donnent de bien meilleurs résultats. Ainsi, la méthode proposée dans [Suk et Chung, 1993], autorise beaucoup plus de fusion que les simples réunions de pixels 2×2 . Douze fenêtres différentes sont autorisées, avec une priorité aux plus grandes (cf figure 3.4) (voir aussi [Chen et al., 1991]).

3.3 Les graphes d'adjacence

Les techniques par graphes d'adjacence sont beaucoup utilisées à partir de sur-segmentations (c'est-à-dire de segmentations où les zones sont subdivisées trop finement). Ces sur-segmentations sont par exemple le résultat d'un prédicat trop sévère dans la phase de segmentation ou d'un algorithme très sensible aux variations locales comme la technique des lignes de partage des eaux (cf. chapitre ?? et [Schmitt et Mattioli, 1994]).

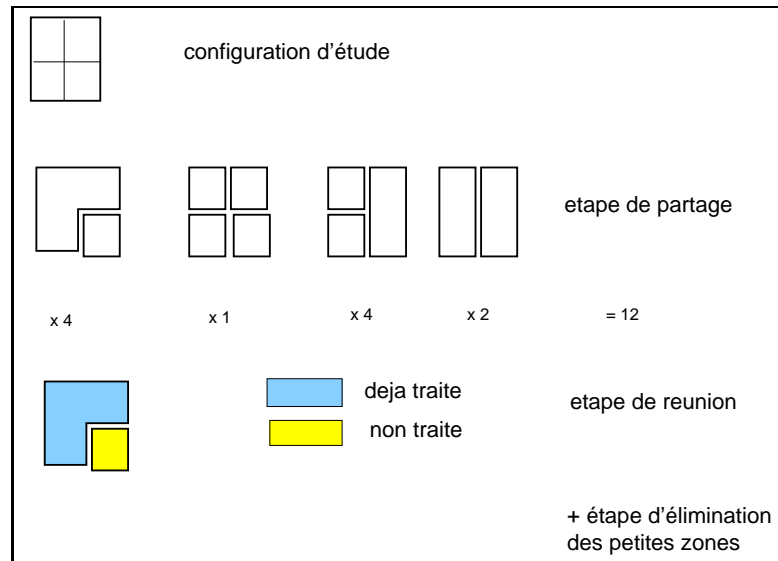


FIG. 3.4 – Les 3 étapes de la méthode de partage et réunion de Suk. Pour chaque étape il y a un prédicat différent. Celui du partage et de la réunion est généralement un prédicat de contraste, celui sur l'élimination des petites zones est un prédicat de taille. Les petites zones sont associées à la zone la plus proche en niveaux de gris.

L'idée de base consiste à plonger les régions obtenues dans une structure de graphe où une région est un nœud et un arc une relation d'adjacence. Puis on définit une fonction de similarité entre 2 nœuds. On trie tous les couples de nœuds adjacents dans une liste ordonnée. On regroupe les 2 meilleurs candidats. On remet à jour la liste et on itère.

La méthode proposée par Beaulieu et Goldberg [Beaulieu et Goldberg, 1989] est un bon exemple d'une telle technique. Le critère d'homogénéité $H(\mathcal{R}_i)$ d'une zone \mathcal{R}_i est le résidu de son approximation par un polynôme de degré d fixé :

$$H(\mathcal{R}_i) = \sum_{(x,y) \in \mathcal{R}_i} \left[f(x,y) - \sum_{p+q \leq d} a_{p,q}^i x^p y^q \right]^2$$

et plus particulièrement d'une approximation par une constante. En partant d'une partition très fine en régions très homogènes, et en acceptant progressivement des réunions de régions de moins en moins semblables, on s'approche de l'optimisation du critère global de minimum de variance sur les partitions.

D'autres schémas plus complexes sont mis en œuvre, par exemple dans [Wang, 1998] pour obtenir des segmentations meilleures.

3.4 La méthode MDL = Minimum Description Length

C'est une technique issue de la théorie stochastique de l'information [Rissanen, 1984, Rissanen, 1987, Rissanen, 1989] pour optimiser la représentation de données. Elle a été reprise de façon simplifiée dans divers domaines du traitement des images (en particulier en codage). L'idée du MDL consiste à exploiter l'analogie, au sens de la théorie de l'information, entre longueur minimale de description et quantité d'information. On cherche donc à optimiser le choix d'un modèle pour décrire les données conduisant à la plus courte description :

- d'une part par le choix d'un modèle qui s'adapte bien aux données,
- d'autre part par le choix d'un modèle simple qui nécessite peu de paramètres.

L'optimisation au sens du MDL procure un compromis entre ces deux termes. Dans le cas de la description d'une image, le modèle doit décrire les contours des régions et approcher au mieux l'image entre les contours.

On se place donc dans le domaine de la représentation paramétrique des données. On dispose d'un modèle dépendant de paramètres et on recherche le meilleur jeu de paramètres pour approcher les données. Soit $\phi = (\phi_1, \dots, \phi_k)$ le vecteur à k composantes des paramètres. Soit x l'observée (l'image). Dans le cadre de l'estimation au maximum de vraisemblance, on cherche le meilleur $\phi \in \mathbb{R}^k$ qui maximise $p(x/\phi)$. Cela revient à maximiser la log-vraisemblance :

$$\mathcal{L}(x/\phi) = -\log p(x/\phi) \quad (3.4)$$

Si l'on dispose de plusieurs modèles, chacun ayant des jeux différents de ϕ_i , la transmission de la représentation de x nécessite la transmission d'une part des paramètres décrivant la forme dans ce modèle, d'autre part les éléments permettant de reconstruire le modèle puisqu'il existe de nombreux modèles potentiellement disponibles. On sait que selon la théorie de l'information, il existe des codes optimaux pour transmettre le premier terme en utilisant un nombre de bits égal au logarithme de la probabilité $p(x/\phi)$, c'est à dire le terme de l'équation 3.4. On devra ensuite transmettre le second terme. Pour rechercher la meilleure description au sens de la théorie de l'information, on minimisera donc une expression plus complète que celle utilisée pour la représentation au maximum de vraisemblance (équation 3.4) :

$$\mathcal{L}'(x/\phi) = -\log p(x/\phi) + \mathcal{L}(\phi) = \mathcal{L}(x/\phi) + \mathcal{L}(\phi) \quad (3.5)$$

où $\mathcal{L}(\phi)$ représente la quantité d'information nécessaire pour transmettre le modèle. Dans le cadre général de la classification ou du traitement du signal, il est parfois possible de mener une optimisation complète et explicite du MDL.

Utilisé dans le cadre de la segmentation d'image, le MDL introduit dans la segmentation un terme $\mathcal{L}'(x/\phi)$ de pénalité lors d'une représentation en régions. Il vient limiter d'une part le nombre de contours (le premier terme $\mathcal{L}(x/\phi)$), d'autre part la complexité des contours par le choix si possible d'un petit nombre de paramètres très simples à coder (le second terme : $\mathcal{L}(\phi)$).

Dans la pratique, il n'est généralement pas possible en traitement des images d'obtenir une solution explicite de l'équation 3.5 et l'on recherche très souvent des solutions itératives (par modification des lignes de contour) sous le contrôle du terme 3.5. Le MDL est utilisé en segmentation d'images de différentes façons. On exprime généralement :

1. le coût de la représentation d'une région par une fonction (par exemple constante). Ce coût s'exprime par le nombre de bits nécessaires à coder la constante plus le nombre de bits nécessaires à coder l'erreur résiduelle,
2. le coût de la représentation des contours (coût de codage de la chaîne des contours). Une chaîne de contours peut être codée de façons très diverses : Freeman (cf. chapitre 5.4), approximations polygonales, splines, processus de bords, etc. et conduira alors à des codages très différents.

Le MDL tend à procurer un compromis entre ces deux termes, c'est-à-dire à donner peu de contours réguliers mais au bon endroit. Les techniques mises en œuvre en MDL sont à base de champs de Markov (optimisation par recuit simulé) [Lee, 1998, Leclerc, 1989, Zhu et Yuille, 1996], ou d'optimisation de courbes [Darell et al., 1990, Keren et al., 1990].

3.5 L'approche de Mumford et Shah

3.5.1 Approche formelle

C'est une façon formelle de poser le problème de segmentation [Mumford et Shah, 1989] qui n'a malheureusement pas de solution exacte. Soit $f(x, y)$ l'image³, de support \mathcal{I} , on la remplace sur des domaines \mathcal{R}_i par des fonctions régulières $g_i(x, y)$ qui se rejoignent le long de contours Γ_j (cf. figure 3.5). La fonctionnelle à minimiser se compose de :

$$U(\Gamma, g, f) = \mu^2 \iint_{\mathcal{I}} (f(x, y) - g_i(x, y))^2 dx dy + \iint_{\mathcal{R}_i} \|\nabla g_i(x, y)\|^2 dx dy + \nu \int_{\Gamma_i} dl \quad (3.6)$$

On montre aisément que les 3 termes sont simultanément nécessaires sous peine de convertir l'expression en une formule triviale. Mumford et Shah ont déduit des propriétés particulières des solutions Γ_j et g , que nous examinons ci-dessous.

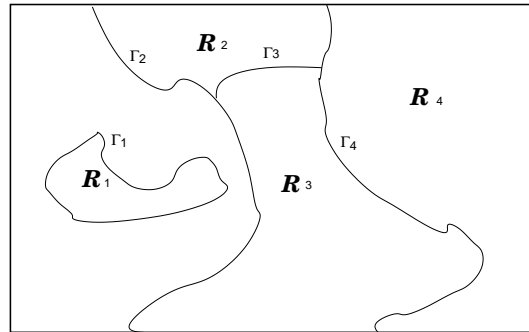


FIG. 3.5 – Approche de Mumford et Shah : on recouvre chaque domaine \mathcal{D}_i d'une fonction continue, on cherche à minimiser la distance entre la fonction de représentation et l'image, ainsi que la longueur des côtés.

On sait résoudre ce problème dans quelques cas simples (par exemple si g_i est une constante qui ne prend que deux valeurs (généralement choisies comme +1 et -1), on retrouve le problème d'Ising qui a une solution - en théorie exacte - par recuit simulé (cf. chapitre ??)⁴.

Mumford et Shah ont montré par ailleurs que l'équation générale 3.6 peut évoluer vers deux formes limites.

– Tout d'abord la minimisation de l'énergie U_o :

$$U_o(\Gamma, f) = \frac{1}{\mu^2} U(\Gamma, f, g) = \sum_i \iint_{\mathcal{R}_i} (f - g_i)^2 dx dy + \nu \int_{\Gamma_j} dl$$

³Attention dans le texte original de Mumford et Shah, [Mumford et Shah, 1989] les rôles de f et g sont interchangés, afin de garder la cohérence de nos notations nous avons préféré désigner par f l'image et par g son approximation.

⁴On se souvient cependant que les conditions de convergence vers le maximum absolu par recuit simulé sont liées à une descente en température infiniment lente.

que l'on obtient en imposant à g d'être constante par morceaux sur les \mathcal{R}_i : $g(x, y) = g_i$ ssi $(x, y) \in \mathcal{R}_i$. Il est alors aisé de montrer que $g_i = \frac{1}{s_i} \iint_{\mathcal{R}_i} f(x, y) dx dy$ où s_i représente la surface de \mathcal{R}_i (c'est donc la valeur moyenne sur la surface).

U est la limite de U_o quand μ tend vers 0. La minimisation de U_o est un problème bien posé. Le problème d'Ising évoqué plus haut en est un sous-problème réduit au cas à 2 classes, le problème de Potts l'étend à un plus grand nombre de classes.

– Puis la minimisation de l'énergie :

$$U_\infty(\Gamma) = \int_\Gamma \left[\nu_\infty - \left(\frac{\partial f}{\partial n} \right)^2 \right] dl$$

où ν_∞ est une constante et $\frac{\partial f}{\partial n}$ mesure la composante normale à Γ du gradient de f .

Ce problème peut se ramener à un problème de géodésique : on cherche simultanément à déterminer les trajectoires Γ_i de longueur minimale (terme en ν_∞) et à maximiser le gradient transverse de f le long de Γ_i ⁵. Ce problème est généralement mal posé si les Γ sont quelconques ; sur des contours de forme restreinte, le problème peut être bien posé.

On peut montrer que U_∞ est la limite de U quand μ tend vers l'infini.

On montre que si les Γ sont fixés (c'est-à-dire si l'on a fait par ailleurs une détection de contours), alors la minimisation de U est la recherche du minimum d'une fonction quadratique définie positive possédant un minimum unique. Elle est solution de :

$$\Delta g = \mu^2(g - f) \quad (3.7)$$

avec pour conditions aux limites :

$$\frac{\partial g}{\partial n} \Big|_{\Gamma_j} = 0$$

Si les contours Γ_j sont des courbes régulières (généralement on recherche des solutions parmi les courbes de classe \mathcal{C}^2)⁶, on déduit alors les propriétés suivantes :

- les contours Γ_j ne peuvent posséder que des points singuliers de deux types :
 - soit des points triples où 3 courbes se joignent à 120° ,
 - soit des points extrémités d'où démarre une courbe unique (démarrage d'un pli) ;
- les contours se rattachent aux bords de l'image selon des angles droits ;
- par ailleurs, les fonctions g ont une dérivée normale aux Γ_j horizontale le long des Γ_j .

3.5.2 Les variantes autour de la formulation de Mumford-Shah

Des variantes du problème de Mumford et Shah ont été proposées qui abordent le problème de nombreuses façons. Par exemple des méthodes déterministes issues de la mécanique et de la résistance des matériaux ont été proposées pour apporter des solutions en termes de *membranes* ou de *plaques minces* [Blake et Zisserman, 1987]. On cherche alors à résoudre l'équation 3.7. Ces solutions conduisent à des équations données par :

– dans le cas d'une membrane :

$$U(g, f) = \alpha_1 \cdot \sum_i \iint_{\mathcal{D}_i} (f(x, y) - g_i(x, y))^2 dx dy + \alpha_2 \cdot \sum_i \iint_{\mathcal{D}_i} (\|\nabla g_i(x, y)\|)^2 dx dy \quad (3.8)$$

⁵On a également abordé ce problème lorsqu'on a évoqué les modèles géodésiques actifs (cf. paragraphe 2.4.4)

⁶Une courbe de classe \mathcal{C}^n possède n dérivées continues.

– dans le cas d'une plaque mince :

$$\begin{aligned}
 U(g, f) &= \alpha_1 \cdot \sum_i \iint_{\mathcal{D}_i} (f(x, y) - g_i(x, y))^2 dx dy \\
 &+ \alpha_2 \cdot \sum_i \iint_{\mathcal{D}_i} (\Delta g_i(x, y))^2 dx dy
 \end{aligned} \tag{3.9}$$

Il est également possible de prendre en compte dans ce second cas un terme de torsion (dérivées croisées) dans le terme de régularisation.

Dans les 2 cas les discontinuités Γ sont fixées (par exemple par un détecteur de contour préalable).

La résolution de ces problèmes (connaissant la position des contours on recherche la valeur en tout point de g_i) se fait traditionnellement par des techniques convolutionnelles. À chaque fonctionnelle est attachée une réponse impulsionnelle - ou fonction de Green - G . C'est-à-dire que, sur chaque domaine, \mathcal{D}_i , $g(x, y)$ est de la forme :

$$g(x, y) = \int \int_{\mathcal{D}_i} G(x, y, u, v) \cdot f(u, v) du dv$$

Par exemple en l'absence de contours, la membrane (équation 3.8) a pour fonction de Green K_o , fonction de Bessel modifiée de seconde espèce :

$$G(\rho = \sqrt{(x - x')^2 + (y - y')^2}) = \frac{1}{2\pi\lambda^2} K_o\left(\frac{\rho}{\lambda}\right)$$

équivalente à $\log(1/\rho)$ près de Γ et équivalente à $\frac{1}{\rho} \exp(-\rho)$ pour ρ grand. C'est la solution de l'équation sans second membre issue de l'équation 3.7 :

$$g - \frac{\alpha_2}{\alpha_1} \cdot \nabla^2 g = \delta(x, y)$$

La plaque mince (équation 3.9) a pour fonction de Green :

$$G(\rho) = \frac{i}{4\pi\mu^2} \left[K_o\left(\frac{\rho\sqrt{i}}{\mu}\right) - K_o\left(\frac{\rho}{\mu\sqrt{i}}\right) \right]$$

et elle est solution de l'équation :

$$g + \mu^4 \nabla^4 g = \delta(x, y)$$

avec $\mu^4 = \frac{\alpha_2}{\alpha_1}$.

Dans le cas où il n'y a pas de discontinuités de g le long des Γ , le problème est assez simple mais peu satisfaisant en termes de segmentation d'image [Grimson, 1981, Terzopoulos, 1983]. La solution est obtenue par des techniques de relaxation [Grimson, 1981] ou de relaxation multi-grilles [Terzopoulos, 1983]⁷.

Dans le cas discret on étend les solutions continues par des approches par éléments finis permettant de calculer de façon discrète les dérivées (on est alors dans un schéma très semblable à celui adopté pour les contours actifs vus au paragraphe 2.4).

On définit un processus de bord b booléen (cf figure ??) et on calcule les divers termes de l'énergie en fonction des valeurs $b = 0$ (pas de bord) ou $b = 1$ (présence d'un bord) de ce processus.

Si l'on ne connaît pas la position exacte des contours, on met en œuvre des techniques itératives et (souvent) sous-optimales qui partent de contours estimés et les améliorent. La résolution se fait par des techniques

⁷Les techniques multi-grilles sont des techniques d'optimisation qui utilisent la régularité spatiale de la solution pour trouver tout d'abord une solution exacte sur une grille d'échantillonnage très grossière, puis par un raffinement itératif de la solution, de la calculer sur des grilles de plus en plus fines pour aboutir à la résolution ultime de l'image.

de Markov avec processus de bord explicites [Geman et Geman, 1984] - cf paragraphe 3.1.5 - ou implicites et des schémas d'optimisation adaptés (par exemple la méthode GNC) ou des techniques à base de réseaux neuro-mimétiques de Hopfield. La méthode GNC (*Graduated Non Convexity* de Zisserman [Blake et Zisserman, 1987]), est particulièrement adaptée à ce problème. Elle remplace les fonctions d'énergie non convexes à optimiser par une famille de fonctions convexes qui, successivement optimisées par une technique de gradient, permettent d'atteindre l'optimum global.

Dans ces schémas, on peut aussi introduire un terme de pénalité sur la longueur des contours comme dans l'équation 3.6. On trouve dans la littérature des formes variées de cette énergie :

1. la plus simple est celle de Mumford et Shah et s'exprime par :

$$U = \int_{\Gamma} dl$$

2. une minimisation de la courbure des contours :

$$U = \int_{\Gamma} \left(\frac{d\theta}{dl} \right) dl$$

3. une pénalisation pour les seuls forts changements de direction des contours.

Des techniques où l'on mélange des termes de dérivées premières et secondes (comme dans les contours actifs - cf. paragraphe 2.4) sont également possibles, mais réputées moins bonnes.

Chapitre 4

Les textures

Chapitre rédigé par Henri MAÎTRE

4.1 Qu'est ce qu'une texture ?

Dans le domaine du traitement de l'image et de la vision, il n'existe pas de définition satisfaisante de la texture. Les définitions mathématiques construites à partir de propriétés statistiques sont soit trop générales et imprécises soit trop restrictives pour s'adapter à la diversité des cas rencontrés. La définition que nous proposons n'est pas opérationnelle et prête sur ce point à critique. Elle s'appuie sur une constatation expérimentale : une texture est un champ de l'image qui apparaît comme un domaine cohérent et homogène, c'est-à-dire formant un tout pour un observateur. C'est cette propriété de cohérence de la texture placée dans son contexte d'être perçue comme un tout homogène par l'œil humain qui sera recherchée le plus souvent par le traiteur des images, dans le but d'isoler les textures, soit pour segmenter l'image, soit pour reconnaître des régions. La figure 4.1 illustre la diversité des textures et l'ouvrage de référence de Brodatz [Brodatz, 1966] offre une collection de textures naturelles qui constituent d'excellents exemples et sont souvent utilisées pour tester les algorithmes et les méthodes.

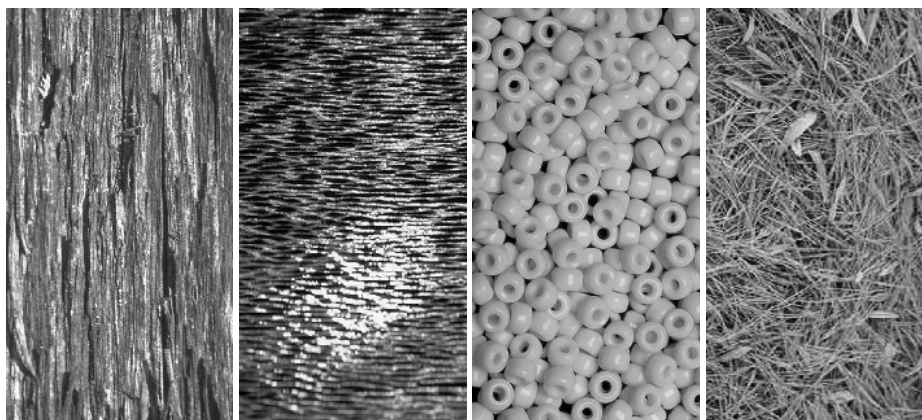


FIG. 4.1 – Quelques exemples de textures naturelles : écorce d'arbre, poil court, perles, tapis d'aiguilles de pin et de feuilles.

Les capacités du système visuel humain à discerner les textures différentes sont remarquables et très mal ex-

pliées. Nos capacités à mémoriser et discriminer des textures sont très variables. Elles deviennent excellentes dès lors que l'on a reconnu l'objet de la vie courante dont la texture est issue. Notre mémorisation est alors remarquable et invariante à de très nombreuses transformations (changement d'échelle, d'orientation, d'éclairage, de couleur, etc.). Au contraire, des textures non reconnues (champs aléatoires créés mathématiquement par exemple, ou textures placées hors de leur contexte) sont mémorisées de façon très fugitive et sont peu robustes aux transformations. Nos capacités à distinguer des textures différentes peuvent parfois s'expliquer par les capacités de filtrage du système visuel (bande passante des voies optiques, discrimination angulaire, sensibilité aux variations locales de luminosité [Hubel et Wiesel, 1969, Julesz, 1971]). Parfois elles relèvent de mécanismes de psycho-vision et peuvent trouver des explications par exemple dans la Gestalt Theorie qui propose des mécanismes d'association (groupement perceptuel) ou de discrimination, à partir des symétries, proximités, similarités des stimulus élémentaires composant la texture [Wertheimer, 1944, Grossberg et Mingolla, 1985].

L'une des remarquables capacités du système visuel est de reconnaître des textures déjà mémorisées même sous de très fortes distorsions géométriques (vues perspectives, recouvrement de surfaces gauches, etc.). Dans ces conditions, et même en l'absence d'un contexte favorable (par exemple présence de silhouettes d'objet ou d'alignements perspectifs), le système visuel interprète la déformation de la texture, non comme une perturbation des propriétés géométriques du champ texturé, mais comme une variation de la géométrie qui supporte la texture, celle-ci restant globalement invariante dans un repère imprécisé (cf. figure 4.2). Ainsi un tronc d'arbre est vu comme une texture homogène enroulée sur un cylindre et non comme une texture plane dont les fréquences spatiales s'accroissent en approchant des bords.

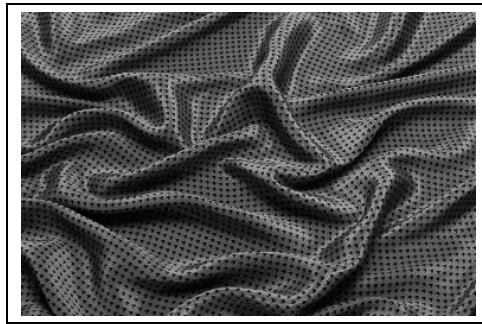


FIG. 4.2 – Lorsque la texture n'est pas homogène, on interprète naturellement les variations locales de ses propriétés statistiques comme des déformations de la surface sur laquelle elle est projetée.

Pour toutes ces raisons, (mais aussi pour simplifier considérablement la tâche du traiteur d'image), nous considérerons par la suite les textures comme spatialement invariantes, c'est-à-dire que nous négligerons généralement les effets de perspectives et les variations d'homogénéité du champ de distribution des intensités. Nous lèverons cette hypothèse au paragraphe 4.6.

4.1.1 Distribution aléatoire ou régulière ?

Une première constatation que l'on fait en examinant des textures naturelles est le rôle particulier que joue l'aléatoire dans la texture. On distingue assez naturellement deux modèles extrêmes de textures, entre lesquels se positionnent un peu toutes les textures :

1. les textures régulières, dans lesquelles la périodicité du motif est évidente : grilles, murs, tissus, etc.
2. les textures aléatoires pour lesquelles la distribution des intensités n'est l'objet d'aucune régularité apparente : sable, nuages, herbe, foule.

La première famille sera bien décrite par des approches fréquentielles ou des approches structurelles dans lesquelles on associera un motif et des règles de placement sur un pavage régulier.

La seconde approche sera plutôt décrite par des lois statistiques, des moments, une description spectrale en termes de densité de puissance, des propriétés de corrélation ou d'isotropie.

Mais une texture n'est jamais strictement périodique ni totalement aléatoire et les deux modèles se complètent naturellement. L'une des difficultés est de les mêler dans des modèles capables de s'adapter à la variabilité des textures étudiées.

4.1.2 Les échelles des textures

Une autre particularité très importante des textures est qu'elles présentent généralement plusieurs niveaux d'échelle auxquels on peut les étudier.

À petite échelle (donc pour des détails fins), on observe souvent un objet élémentaire qui constitue la base de la texture (poil, perle, aiguille de pin, sur la figure 4.1). Cette primitive peut être très régulière géométriquement (perle), ou photométriquement (aiguille de pin). Elle peut être au contraire relativement variable (personne dans une foule, caillou sur une plage, nuage). Elle possède une taille et des propriétés statistiques caractéristiques (qui joueront sur sa fonction de corrélation, son isotropie, ...). Cette composante élémentaire de la texture a tendance naturellement à disparaître lorsqu'on observe le champ global de la texture. Elle est intégrée à la texture.

À plus grande échelle (donc pour une analyse plus grossière), la texture apparaît comme une juxtaposition plus ou moins régulière des motifs ci-dessus. Cette juxtaposition possède ses propres lois d'isotropie, de périodicité, de régularité et induit donc d'autres lois statistiques (corrélation, densité de puissance) qui se combinent à celles du motif de base.

Une bonne analyse de texture donnera des informations sur ces deux composantes également.

4.1.3 Analyse ou synthèse

Un effort très important a été fait dans le domaine de la synthèse des images pour créer artificiellement ou copier des textures naturelles : forêts ou cultures pour des simulateurs de vols, bois, tissus, métaux pour des représentations d'intérieur, murs, toitures, eau, pour des jeux vidéo, etc. Ces travaux se distinguent notablement de ceux conduits en traitement d'image qui visent plutôt à extraire des paramètres discriminants et robustes permettant de séparer des textures différentes. Mais ces travaux se rejoignent en ce qu'ils contribuent simultanément à une meilleure connaissance et compréhension des textures à travers la boucle « analyse/synthèse » qui a fait ses preuves en reconnaissance des formes. Dans la suite, nous empruntons les méthodes aussi bien à l'analyse qu'à la synthèse des images en précisant lorsque c'est nécessaire les limites des méthodes dans l'une ou l'autre application.

4.2 Modèles de texture

4.2.1 Un modèle biologiquement plausible

Compte tenu du rôle important de la perception humaine dans la définition même de texture, des modèles de textures ont été proposés, s'inspirant de ce que l'on connaît aujourd'hui de cette perception. Ces travaux s'appuient sur les mécanismes d'adaptation du système visuel aux fréquences spatiales et aux orientations [Hubel et Wiesel, 1969, de Valois et al., 1982] ainsi que sur des expérimentations sur les effets de masquage psychovisuel [Phillips et Wilson, 1984].

On trouve par exemple dans [Bergen et Landy, 1991] un tel modèle. On décompose tout d'abord le flot optique en une pyramide (cf. paragraphe 3.2.4) par une succession de filtrages gaussiens et de sous-échantillonnages adaptés (semblables aux pyramides de Burt et Adelson [Burt, 1981]). Chaque niveau de la pyramide est alors filtré par 4 filtres qui produisent une dérivée seconde directionnelle selon l'horizontale, la verticale et les deux diagonales. On procède ensuite à une mesure d'énergie dans chaque image par une intégration locale des sorties

précédentes et l'on soustrait ces énergies, dans un même niveau de résolution, entre orientations voisines (pour exprimer le contraste relatif entre directions). On procède enfin à une étape de normalisation pour tenir compte d'une sorte de gain variable de contrôle de contraste vérifié expérimentalement.

Un tel schéma de fonctionnement du système visuel humain est très schématisé et controversé. Il existe d'autres schémas tout aussi intéressants [Mayhew et Frisby, 1978], mais il a le mérite de comporter la plupart des étapes que nous reverrons plus loin pour analyser les textures.

4.2.2 Modèles stochastiques

Ces modèles, au contraire, mettent l'accent sur la distribution statistique des pixels et sur leur dépendance spatiale. La texture est alors considérée comme la réalisation d'un processus aléatoire gouverné par ses lois. Nous avons vu dans les chapitres antérieurs deux modèles qui ont été abondamment utilisés pour modéliser les textures :

- le schéma booléen (cf. section ??), et plus généralement les ensembles fermés aléatoires, qui considèrent la texture comme une distribution poissonnienne de motifs et qui est particulièrement bien analysé par les outils développés en morphologie mathématique qui permettent de définir sa capacité de Choquet [Schmitt et Mattioli, 1994] ;
- les modèles autorégressifs et leurs dérivés, chaînes de Markov et champs de Markov, ces derniers étant aujourd'hui les plus universellement adoptés. Nous nous attarderons particulièrement sur ce point à la section 4.4 [Cross et Jain, 1983, Chellappa et Kashyap, 1985].

4.3 Analyse et reconnaissance de textures

Il existe de nombreuses revues des diverses approches de l'analyse de textures, et nous renvoyons le lecteur à trois textes qui couvrent assez bien l'évolution du domaine : [VanGool et al., 1985, Reed et du Buf, 1993, Randen et Husøy, 1999]

4.3.1 Schéma général

Le principe le plus général de l'analyse statistique des textures est le suivant.

1. On définit un voisinage \mathcal{V}_{ij} de tout pixel (i, j) , de taille et de forme appropriées.
2. Sur le voisinage \mathcal{V}_{ij} on mesure des propriétés particulières de l'image : soit π_k la mesure attachée à la configuration k parmi les N mesurées. Ces mesures seront attachées au pixel (i, j) .
3. On classe les pixels à partir du vecteur formé par les π_k par l'une des nombreuses méthodes de la reconnaissance des formes (rappelées en section 3.1.3).
4. Éventuellement, avant cette classification, et si le vecteur de mesures est trop grand, on réduit la dimension de l'espace par une sélection judicieuse des composantes les plus significatives ou par analyse en composantes principales.

La fenêtre d'analyse

La dimension du voisinage \mathcal{V}_{ij} est importante. C'est l'un des choix délicats de l'analyse des textures (cf. figure 4.3). Elle doit comprendre au moins un motif de base de la texture mesurée pour fournir des statistiques homogènes sur la texture, mais, si \mathcal{V}_{ij} est trop grand, la précision de localisation des frontières des textures détectées sera médiocre. La forme est généralement carrée. Taille et forme peuvent éventuellement s'adapter au signal à mesurer si l'on a quelque méthode pour le faire (pré-segmentation, masque de contours, etc.), mais il faut s'assurer alors que la technique de classification n'est pas perturbée par des mesures faites sur des échantillons variables.

La classification des données

Ce sont celles que nous avons vues au chapitre 3.1.3 : nuées dynamiques, plans séparateurs, réseaux neuro-mimétiques, C moyennes floues, etc. Dans l'idéal on aimerait pouvoir obtenir à partir de l'image $f(i, j)$, une nouvelle image $g(i, j)$ dont les plages seront très distinctes et pourront, idéalement, être segmentées par simple seuillage (cas de la mesure d'un seul paramètre discriminatoire) ou par des seuils multiples sur les diverses composantes. En pratique il est rare que l'on puisse séparer les nuages des mesures et il faut trouver des compromis lors des décisions. Les erreurs de classification peuvent bénéficier de techniques contextuelles de reclassification comme la relaxation ou les champs de Markov comme nous le verrons plus bas (section 4.4).

Lorsque l'on cherche non seulement à discriminer des plages de textures différentes, mais aussi à reconnaître des textures particulières, on utilise des techniques de classification supervisées : k-plus proches voisins, classifieurs bayésiens, réseaux neuro-mimétiques, algorithmes génétiques, préalablement entraînés sur des ensembles d'apprentissage [Vickers et Modestino, 1982].

La réduction des dimensions

Les diverses variables mesurées sont rarement indépendantes et le vecteur des π_k occupe rarement toutes les dimensions de N . On calcule alors la matrice Π de taille $N \times N$ et telle que $\Pi(k, l) = \langle \pi_k \pi_l \rangle$. On diagonalise Π et on ne retient que ses p plus grands vecteurs propres sur lesquels on projette toutes les mesures. La classification se fait alors dans l'espace de dimension p , ce qui réduit les calculs et permet de prendre de meilleures décisions car les composantes sont alors non corrélées.

4.3.2 Approches par mesures statistiques

Dans ces approches on mesure les propriétés statistiques autour du point (i, j) dans le voisinage \mathcal{V}_{ij} .

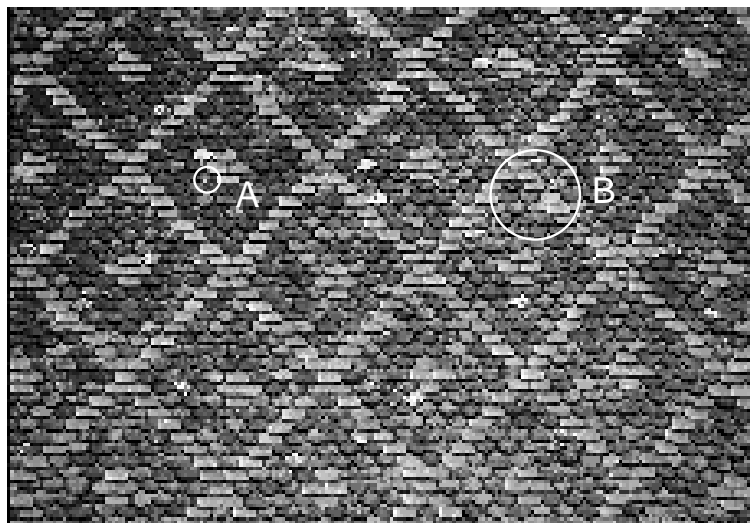


FIG. 4.3 – Le choix de la taille de la fenêtre d'analyse est important pour sélectionner les détails que l'on souhaite préserver. La fenêtre A permettra de déterminer des propriétés caractéristiques d'une tuile de ce toit et distinguera tuiles claires de tuiles sombres, la fenêtre B mesurera des propriétés moyennes sur tout le toit et ignorera les différences entre tuiles claires et sombres.

Statistiques du premier ordre

Ce sont celles qui ne prennent en compte qu'un pixel à la fois¹. Elles se déduisent de la probabilité empirique $p(n)$ du niveau de gris n dans la fenêtre ou de l'histogramme $h(n) \simeq \nu p(n)$ où ν est le nombre total de pixels dans l'image. Ce sont :

- les moments d'ordre k (non centrés) : $\mu_k = \sum_n n^k p(n)$,
- les moments centrés d'ordre k : $\tilde{\mu}_k = \sum_n (n - \mu_1)^k p(n)$, et en particulier :
 1. la moyenne : μ_1 ,
 2. la variance : $\sigma^2 = \tilde{\mu}_2$,
 3. le biais : $\gamma_1 = \frac{\tilde{\mu}_3}{\sigma^3}$,
 4. l'aplatissement (ou kurtosis) : $\gamma_2 = \frac{\tilde{\mu}_4}{\sigma^4} - 3$.
- l'énergie : $W = \sum_n |p(n)|^2$,
- l'entropie : $E = -\sum_n p(n) \log p(n)$,
- le contraste : $C = \frac{\max(n) - \min(n)}{\max(n) + \min(n)}$,
- la dynamique : $D = \max(n) - \min(n)$,
- le coefficient de variation (surtout pour les images cohérentes dont le bruit est multiplicatif : radar, images ultra-sonores) : $v = \frac{\mu}{\sigma}$,
- l'exposant de Holder, caractéristique de la dimension fractale : [Pentland, 1984].

Si les fenêtres de mesure sont petites, les statistiques mesurées sont souvent peu significatives (pour une fenêtre de 30×30 pixels et un histogramme de 256 niveaux de gris, l'occurrence moyenne des pixels est inférieure à 4 par niveau de gris). On choisit alors souvent de réduire la dynamique des images. On a montré que les propriétés de discrimination des textures se conservaient remarquablement même pour des quantifications très fortes des textures (jusqu'à 8 ou 4 niveaux de gris seulement), à la condition d'adapter la quantification. Une quantification brutale se fait en ne conservant que les bits de poids fort. Une quantification adaptée choisit de partager les niveaux de gris en fonction des statistiques de la texture :

- par moyenne et écart type par exemple (cela donne 4 classes non équiréparties),
- par médiane et distance interquartile², ce qui donne des classes équiréparties.

Les statistiques de points particuliers

On recherche dans ce cas, dans la fenêtre \mathcal{V}_{ij} , la densité moyenne de certains points d'intérêt. Les plus souvent utilisés sont :

- les maximums locaux de l'intensité,
- les points de contour (après application d'un détecteur et seuillage),

Les statistiques d'ordre élevé

Ce sont surtout les statistiques d'ordre 2 qui sont exploitées, c'est-à-dire celles qui mettent en jeu deux pixels simultanément. En effet, B. Julesz a émis une conjecture, fondée sur une très vaste étude expérimentale, que le système perceptif humain ne distingue pas les textures qui ont des statistiques similaires aux ordres 1 et 2, même si elles diffèrent aux ordres supérieurs [Julesz, 1971]. Cette conjecture a été démontrée fautive par A. Gagalowicz qui a produit des textures artificielles ne différant qu'à l'ordre 3 [Gagalowicz et Tournier-Lasserve, 1986], mais perceptivement différentes (cf. figure 4.4). La conjecture de Julesz demeure cependant un guide important pour les études sur les textures car elle est assez bien vérifiée pour les textures naturelles.

¹ Attention, on désigne également parfois par premier ordre les statistiques comme la moyenne, qui ne mettent en jeu que des moments du premier ordre.

² La distance interquartile est la distance qui, dans l'histogramme, sépare les niveaux de gris des 25% de pixels les plus sombres de ceux des 25% de pixels les plus clairs. C'est une mesure robuste de l'étalement d'une loi.

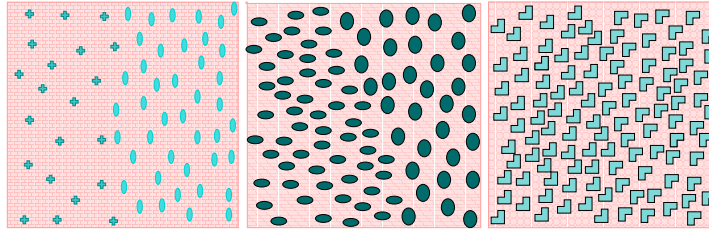


FIG. 4.4 – A gauche, les 2 textures diffèrent dès le premier ordre. Au centre, les 2 textures sont identiques au premier ordre mais diffèrent au second. A droite les 2 textures sont identiques à l'ordre 1 et à l'ordre 2, mais différentes à l'ordre 3. Ces deux textures sont distinguables visuellement, en contradiction avec la conjecture de Julesz.

Les statistiques d'ordre 2 sont bien appréhendées par la **fonction d'autocorrélation**³ de la texture, et plus particulièrement par la fonction d'autocorrélation normée centrée :

$$C_f(k, l; i, j) = \frac{\sum_{\nu_{ij}} (f(i, j) - \bar{f}) \cdot (f(i + k, j + l) - \bar{f})}{\sum_{\nu_{ij}} (f(i, j) - \bar{f})^2}$$

On déduit de nombreuses propriétés de cette fonction :

1. dans une direction donnée, et au voisinage de l'origine, elle peut être fréquemment approchée par une fonction régulière, par exemple :

$$C_f(k, l; i, j) \sim e^{-\alpha(i, j)|k|} e^{-\beta(i, j)|l|}$$

et les paramètres α et β sont très significatifs de la dépendance spatiale (en i et j) des niveaux de gris à l'intérieur de la texture ;

2. les deux directions d'inertie maximale $\lambda_0(i, j)$ et minimale $\lambda_1(i, j)$ extraites de $C_f(k, l; i, j)$ expriment bien les directions privilégiées de la texture et, si le rapport $\frac{\lambda_0}{\lambda_1}$ est fort, la texture est fortement anisotrope.
3. si la fonction $C_f(k, l; i, j)$ possède des maximums locaux différents de $k = 0, l = 0$, alors la texture est fortement périodique et les périodes de la fonction de corrélation permettent d'estimer de façon robuste la période de la texture.
4. enfin par transformation de Fourier, la fonction d'autocorrélation donne accès au spectre de densité de puissance (théorème de Wiener-Kinchine, cf. section ??), et à tous les traitements spectraux que nous verrons ci-dessous.

Plus encore que la fonction d'autocorrélation, les **matrices de cooccurrence** sont des outils adaptés à l'analyse des textures. Ces matrices sont définies pour un vecteur de translation V_i donné comme la probabilité jointe de l'amplitude d'un point M et d'un point $N = M + V_i$. Elles sont obtenues pratiquement par une estimation sur la texture, c'est-à-dire un décompte des occurrences des niveaux m en M et n en N . C'est donc un tableau généralement de taille 256×256 si la texture à 256 niveaux de gris :

$$\Gamma(m, n; V_i) = \text{proba} (f(M) = m, f(N = M + V_i) = n)$$

En raison de sa taille même, la matrice de cooccurrence est sous cette forme un outil ni très pratique, ni très fiable (les valeurs $\Gamma(m, n; V_i)$ sont des estimations très médiocres des probabilités réelles). On a donc intérêt à comprimer ces matrices. Cela se fait de plusieurs façons :

- en diminuant le nombre de niveaux de gris par une quantification régulière ou adaptative, par exemple 8 niveaux de gris déterminés de façon adaptative (par moyenne et écart-type ou par médiane et distance inter-quartile) donnent souvent une bonne qualité de représentation ;

³certains auteurs préfèrent étudier le **variogramme** qui s'exprime comme $1 - C_f(k, l)$.

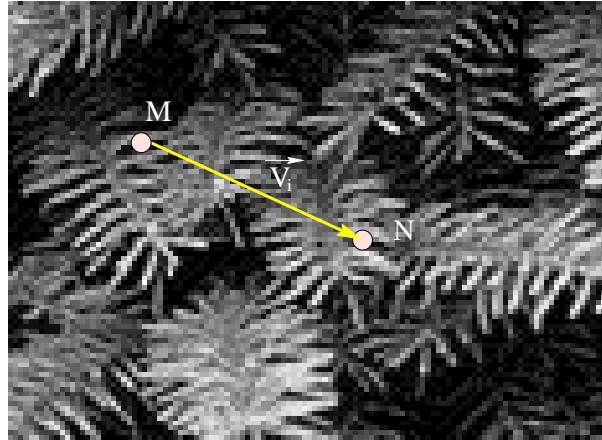


FIG. 4.5 – Construction d’une matrice de cooccurrence : on définit un vecteur V_i , puis on mesure l’occurrence des couples de niveaux de gris m en M et n en N . La matrice de cooccurrence ainsi calculée est $\Gamma(m, n, V_i)$.

- en tirant profit des symétries existant dans la texture pour regrouper des vecteurs V_i contribuant de façon identique à l’apparence de la texture (par exemple des vecteurs symétriques le long des axes, ou des vecteurs de même module si la texture est isotrope) ;
- en ne conservant que les seuls vecteurs significatifs ;
- enfin, on a pu montrer que quelques descripteurs des matrices de cooccurrences pouvaient très bien les représenter : position du centre de gravité, rapport des inerties, énergie, entropie, facteurs de symétrie, etc. [Chen et Pavlidis, 1979].

Les configurations particulières de pixels

On s’intéresse dans cette méthode à mesurer la probabilité d’apparition de certaines configurations particulières de pixels (cf. figure 4.6). Ce sont donc des techniques qui s’appliquent surtout à des textures binaires ou à très peu de niveaux de gris (ou rendues telles). Les mesures ainsi faites permettent de construire un vecteur d’attributs qui permettra de discriminer des textures différentes. Des configurations bien choisies pour un problème particulier peuvent être très discriminantes. Mais il est souvent difficile d’éviter que des mesures ne soient statistiquement dépendantes car les probabilités de certaines fenêtres sont évidemment dépendantes. Il convient donc souvent de réduire la dimension de l’espace des paramètres par analyse en composantes principales.

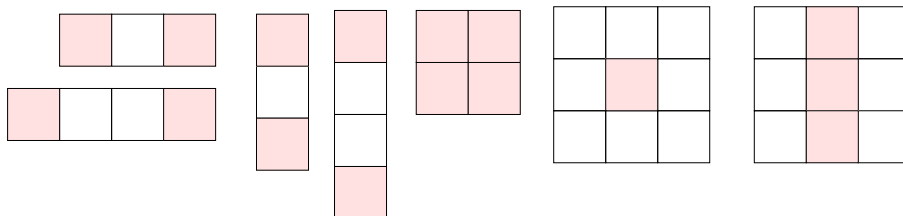


FIG. 4.6 – Quelques configurations particulières de pixels qui peuvent être mesurées pour fournir un descripteur statistique : par exemple on mesure le nombre d’occurrences de trois pixels alignés horizontalement, les 2 des bords étant noirs et celui du centre blanc (en haut à gauche).

4.3.3 La détection d'énergie dans des canaux fréquentiels

L'idée sous-jacente à toutes ces techniques est d'extraire l'énergie portée par le signal dans des bandes de fréquences diverses. Il n'est alors pas nécessaire d'utiliser de fenêtre d'estimation, car les analyses fréquentielles utilisées disposent généralement de leurs propres fonctions de fenêtrage, souvent gaussiennes. Il existe donc deux familles de techniques :

- celles qui analysent systématiquement tout l'espace fréquentiel (par exemple par une décomposition en ondelettes),
- celles qui sélectionnent un petit nombre de domaines fréquentiels significatifs (c'est par exemple l'analyse par filtres de Gabor).

Les filtres de Law

Ce sont des filtres très simples qui sont appliqués dans l'espace de l'image par des masques, dans l'esprit des filtres de détection de contour par masquage adapté. Les filtres sont au nombre de 25, obtenus par produit en x et y de 5 filtres de base (cf. figure 4.7). Leurs performances sont modestes, mais ils sont très rapides.

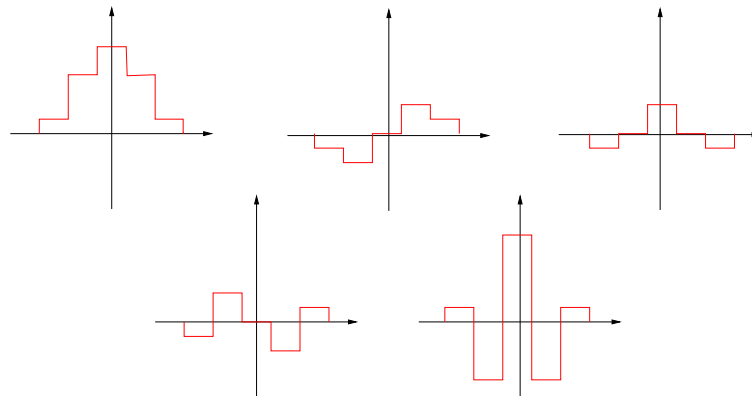


FIG. 4.7 – Les 5 filtres de base de Law. Un filtre de détection est le produit séparable de deux de ces filtres. Il est appliqué dans l'espace de l'image.

Les filtres en anneau et en coin

Ce sont des filtres définis dans l'espace de Fourier à 2 dimensions [Coggins et Jains, 1985]. Ils sont constitués du produit de deux familles élémentaires de filtres :

- 7 filtres en anneau, échelonnés de façon dyadique (c'est-à-dire centrés autour de fréquences suivant une progression géométrique de fréquences, le premier de fréquence $\nu_1 = \nu_0$, le second de fréquence $\nu_2 = 2\nu_0$, le troisième de fréquence $\nu_3 = 4\nu_0$, etc.), et de profil gaussien (cf. figure 4.8) :

$$\Phi_i(\nu = \sqrt{u^2 + v^2}) = \exp \left[-\frac{(\nu - \nu_i)^2}{2\alpha \nu_i^2} \right]$$

le recouvrement de ces filtres est contrôlé par le paramètre α .

- 4 filtres directionnels selon les axes et les diagonales, également de profils gaussiens.

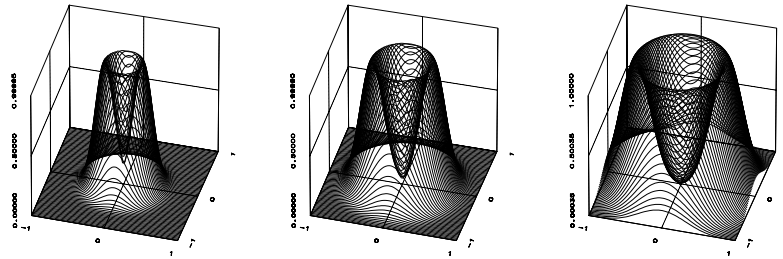


FIG. 4.8 – Les filtres en anneau dans le domaine spectral centrés sur les fréquences : 0,25, 0,40 et 0,60. Ce ne sont donc pas des filtres dyadiques.

Les filtres dyadiques de Gabor

Ce sont des filtres qui essaient de copier les fonctions de la vision des mammifères. Ils sont assez proches des précédents. Ce sont aussi des filtres à couverture dyadique qui se déduisent par rotation et homothétie du filtre de base qui a pour réponse impulsionnelle [Jain et Farrokhnia, 1991] :

$$\phi(x, y) = \exp \left[-1/2 \left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right] \right] \cos(2\pi\nu_0 x)$$

Dans [Jain et Farrokhnia, 1991], on conseille l'emploi de 5 fréquences radiales différentes et de 4 orientations. Dans [Randen et Husøy, 1999], on utilise aussi 4 orientations et les fréquences radiales suivantes :

$$\frac{\sqrt{2}}{2^6} \quad \frac{\sqrt{2}}{2^5} \quad \frac{\sqrt{2}}{2^4} \quad \frac{\sqrt{2}}{2^3} \quad \frac{\sqrt{2}}{2^2}$$

Les ondelettes et les QMF

La transformation en ondelettes discrètes permet de choisir une famille complète de filtres de décomposition en sous-bandes. Ces méthodes ont été présentées au chapitre ???. Les ondelettes de Mallat [Mallat, 1989] utilisent elles aussi une décomposition dyadique (mais deux directions d'analyse seulement). Des travaux suggèrent qu'une décomposition plus serrée que l'octave (c'est-à-dire des décompositions non-dyadiques) serait plus efficace pour l'analyse des textures [Chang et Kuo, 1993]. On propose ainsi l'usage de paquets d'ondelettes [Laine et Fan, 1993, Saito et Coifman, 1995]. Les ondelettes de Daubechies demeurent l'une des bases d'ondelettes les plus utilisées pour l'analyse de textures en raison de leur bonne efficacité de calcul, de leurs bonnes performances pour séparer les fréquences ainsi que de la grande variété des fonctions qu'elles autorisent [Daubechies, 1992, Unser, 1995].

Des modèles ont également été créés alliant les propriétés de sélection des ondelettes et les modèles statistiques [Portilla et Simoncelli, 2000].

4.3.4 Les filtres optimisés

Lorsque l'on cherche à distinguer 2 textures, ou un petit nombre de textures connues par avance, la recherche systématique de l'énergie dans de nombreuses décompositions en sous-bandes peut être très lourde et hasardeuse. Il est possible de mettre en place des techniques adaptées pour les seules textures que l'on recherche qui permettent de se concentrer sur quelques mesures seulement de l'espace fréquentiel.

Les filtres de Gabor, qui disposent d'un petit nombre de paramètres seulement se prêtent assez bien à ce type

d'optimisation. On cherche alors à optimiser un critère comme :

$$J_1 = \frac{\mu_{t_1}}{\mu_{t_2}} \quad J_2 = \frac{(\mu_{t_1} - \mu_{t_2})^2}{\mu_{t_1} \mu_{t_2}} \quad \text{ou} \quad J_3 = \frac{(\mu_{t_1} - \mu_{t_2})^2}{\sigma_{t_1}^2 + \sigma_{t_2}^2}$$

où μ_{t_i} est la valeur moyenne de la mesure donnée par la texture t_i et σ_{t_i} sa variance [Unser, 1986]. Il est parfois possible de déterminer des valeurs optimales analytiques des filtres par ces critères, dans les autres cas on optimise par essai-erreur.

4.3.5 Les modélisations autorégressives

De la même façon que les signaux monodimensionnels sont analysés par des modélisations autorégressives uni-dimensionnelles (AR ou ARMA) afin de mettre en évidence les périodicités qu'ils contiennent, les signaux bidimensionnels peuvent être analysés par des modèles autorégressifs bi-dimensionnels. Une étude très complète de ces méthodes de description est présentée dans [Garello, 2001].

Dans une approche autorégressive à moyenne ajustée (ARMA), l'image est décrite par la formule :

$$f(i, j) = \sum_{(k,l) \in \mathcal{D}_s} a_{k,l} f(i-k, j-l) + b_{0,0} B(i, j) + \sum_{(k,l) \in \mathcal{D}_e} b_{k,l} B(i-k, j-l) \quad (4.1)$$

où \mathcal{D}_s représente le domaine de prédiction lié à la sortie du filtre, \mathcal{D}_e , celui lié à l'entrée du filtre et B décrit un processus de bruit. Si les coefficients $b_{k,l}$ sont nuls pour tout (k, l) appartenant à \mathcal{D}_e , alors le processus est AR. Si les coefficients $a_{k,l}$ sont nuls pour tout (k, l) appartenant à \mathcal{D}_s , alors le processus est MA.

Le choix des domaines \mathcal{D}_s et \mathcal{D}_e incombe à l'utilisateur. Dans de nombreux cas on choisit $\mathcal{D}_s = \mathcal{D}_e = \mathcal{D}$. Afin de conserver aux modèles ARMA une bonne localité, il est souhaitable que le nombre de termes dans \mathcal{D} soit faible, mais pour bien représenter des signaux à longue périodicité, il vaut mieux qu'il soit grand. Ces modèles expriment naturellement une dépendance causale et suscitent donc des discussions comme celles que nous avons abordées au chapitre ??.

Les modèles ARMA permettent de représenter très bien n'importe quelle densité spectrale de puissance. Ceci se fait en utilisant les équations normales associées de Yule Walker [Alata et Cariou, 2001]. Les représentations par modélisation autorégressive des textures est particulièrement efficace pour des textures très périodiques et de large extension.

4.4 Les approches par champs markoviens

Les champs de Markov (cf. chapitre ??) peuvent être vus comme une sous-classe des processus ARMA, mais leur adaptation à traiter des images les distinguent. Ils se prêtent doublement à l'analyse des textures :

1. tout d'abord parce qu'ils comportent naturellement des descriptions des dépendances spatiales entre pixels par le choix des cliques et des potentiels d'interaction au sein des cliques,
2. mais aussi parce qu'ils ont, dans le terme d'énergie lié aux connaissances à priori, les éléments qui permettent de décrire les régions et leurs interactions.

Les champs de Markov sont particulièrement adaptés pour synthétiser et modéliser des textures, puisqu'il suffit de se donner les potentiels correspondant aux propriétés de dépendance, d'isotropie, de distance aux centres de classe, etc. que l'on souhaite voir représenter. Il y a donc une abondante littérature qui exploite le formalisme des champs markoviens pour segmenter des textures : [Hu et Fahmy, 1992, Derin et Elliott, 1987] [Won et Derin, 1992, Andrey et Tarroux, 1998]

très faciles d'emploi en synthèse des textures, ils sont cependant d'un usage plus difficile en analyse car il n'existe pas de formulation explicite permettant de déterminer les potentiels correspondant à une texture dont on dispose. Ce problème a été abordé au chapitre ??, à la section ??.

En règle générale, pour retrouver un champ de Markov à partir de la réalisation d'une texture, il faut définir les cliques significatives, puis les potentiels qui lient les sites dans ces cliques [Azencott et al., 1997].

1. La détermination des cliques utiles : on peut l'obtenir soit par l'analyse des fonctions de corrélation, soit par des techniques de croissance de voisinage, accompagnées de tests d'indépendance au fur et à mesure que l'on agrandit le voisinage.
2. La détermination des potentiels : elle passe souvent par la réduction des potentiels possibles à des classes étroites (par exemple les potentiels quadratiques ou les potentiels polynomiaux), elle peut alors se faire par des techniques de moindres carrés ou par des méthodes de filtrage (techniques de renormalisation).

4.4.1 La méthode de Manjunath et Chellappa

C'est une technique assez typique des approches markoviennes. Elle repose sur l'hypothèse de probabilités gaussiennes dans chaque classe (donc de potentiels quadratiques) [Manjunath et Chellappa, 1991]. Les énergies d'attache aux données pour une classe $L_s = l$ sont de la forme :

$$U_1(y_s|y_r, l) = \frac{1}{2\sigma_l^2} \left[y_s^2 - 2 \sum_r \theta_i^l y_s y_r \right] \quad (4.2)$$

dont les inconnues sont, pour chaque texture l , les θ_i , μ et σ^2 , où les θ_i correspondent aux cliques des directions S, SE, E, etc. du point s (cf. figure 4.9).

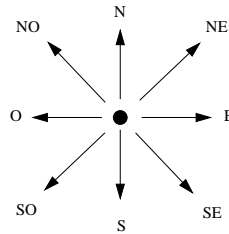


FIG. 4.9 – Le 8-voisinage utilisé par Manjunath et Chelappa. Seules les cliques d'ordre 2 sont utilisées.

Dans une première étape, il faut apprendre les paramètres des textures. L'image est arbitrairement subdivisée en petites fenêtres Ω . Prenons une seule fenêtre de taille $n \times n$ dont on va supposer la texture homogène. Sur cette fenêtre on estime $\hat{\Theta}$ aux moindres carrés par la formule⁴ :

$$\hat{\Theta} = \left[\sum_{\Omega} Q_s Q_s^t \right]^{-1} \left[\sum_{\Omega} Q_s y_s \right]$$

avec $Q_s = [y_{s+1}, y_{s-1}, y_{s+2}, \dots]^t$

et $\hat{\sigma}^2 = \frac{1}{n^2} \sum_{\Omega} [y_s - \hat{\Theta}^t Q_s]^2$.

Une tâche délicate est l'estimation du nombre de classes qui se fait soit de façon supervisée, soit par analyse du nuage des points obtenus sur toutes les fenêtres. Supposons que l'on recherche N classes. On détermine les

⁴C'est ici que l'hypothèse gaussienne simplifie le calcul : l'estimation peut se faire par moindres carrés de façon explicite.

N meilleures classes du nuage des points obtenus sur toutes les fenêtrés par exemple par un algorithme de nuées dynamiques. Cela nous fournit N vecteurs de paramètres $f_k = \theta_1^k, \theta_2^k, \theta_3^k, \theta_4^k, \mu^k, \sigma_2^k$.

On peut alors procéder à la phase de segmentation. On va désormais travailler au niveau des sites du champ de Markov. On introduit un champ de liaisons entre les pixels, décrit par la seule variable binaire (processus de bord) b qui vaut 0 si deux pixels voisins sont dans une même classe et 1 s'ils sont dans des classes différentes. On recherche la distribution des b et des $V_{s\lambda}$, variables caractéristiques de l'appartenance du pixel s à la classe λ ($V_{s\lambda} = 0$ ssi le pixel au site s appartient à la classe λ) qui minimise l'énergie :

$$U = \sum_s \sum_\lambda u(s, \lambda) V_{s\lambda} - (1 - b) \sum_s \sum_\lambda \sum_{t \in N_s} V_{t\lambda} V_{s\lambda}$$

où le premier terme exprime l'attache aux données et le second l'a priori sur les régions. Les u incluent toute l'information connue sur le site s appartenant à la classe λ : par exemple $u(s, \lambda) = w(\lambda) + u_1(Y_s, \lambda)$, u_1 étant comme dans la formule 4.2 et w étant un biais propre à la classe λ . L'optimisation d'une telle fonction d'énergie se fait soit par relaxation déterministe (par exemple l'ICM *Iterated Conditional Mode* qui remplace à chaque étape la classe λ par celle qui minimise l'énergie U au site considéré), soit par recuit simulé (relaxation stochastique) (cf. chapitre ??).

4.4.2 La méthode de Kervrann et Heitz

Dans cette méthode [Kervrann et Heitz, 1995], on calcule sur des fenêtrés recouvrant l'image, des matrices de cooccurrence (cf. section 4.5). En chaque fenêtré, on représente la texture par un vecteur d'attributs issus de ces matrices $O_s = \{o_s^1, o_s^2, \dots, o_s^k\}$. Le terme d'attache aux données s'écrit sous la forme :

$$U_1 = \sum_{s \in \Omega} V(O_s, O_\lambda)$$

où O_s est le vecteur d'attributs du site s et O_λ est le vecteur d'attributs de la région d'étiquette λ affectée au site s . Le potentiel V s'écrit à partir de la distance Δ de Smirnov-Kolmogorov des composantes o_s^i et o_λ^i des vecteurs :

$$V(O_s, O_\lambda) = \sum_{i=1}^k [2\Gamma(\Delta(o_s^i, o_\lambda^i) > c^i) - 1]$$

où la fonction Γ vaut 0 ou 1 selon que les valeurs de Δ sont supérieures ou inférieures à une valeur de seuil c^i issue des tables du test de Smirnov-Kolmogorov [Saporta, 1990].

Le champ évolue vers le critère du MAP (Maximum A Posteriori) par la technique de relaxation sous-optimale de l'ICM : pour chaque site on choisit la meilleure classe parmi les seules classes des voisins du site considéré (cf. chapitre ??).

La construction des classes o_λ se fait de la façon suivante. Initialement il n'y a qu'une classe, identifiée au vecteur de la première fenêtré rencontrée. Les autres vecteurs de texture sont associés à cette classe si leur distance est inférieure au seuil donné. Afin de créer de nouvelles classes, on crée une classe « fourre-tout » dans laquelle sont rangées toutes les textures trop éloignées de celle déjà trouvée. Après un passage sur l'image, on détermine les paramètres de la classe trouvée (en calculant un vecteur O_λ sur toutes les réalisations trouvées) et on accroît de 1 le nombre de classes. On reprend le balayage de l'image. On itère ce procédé jusqu'à ce qu'un nombre pré-établi de classes soit trouvé ou qu'il n'y ait plus de candidats dans la classe fourre-tout.

Comme dans le cas précédent, la taille des fenêtrés doit être :

1. assez grande pour permettre une caractérisation statistique réaliste des textures (et de leurs motifs) ;
2. assez petite pour conduire à des segmentations fines des détails.

Notons que cette méthode, au contraire de la précédente, ne permet pas de déterminer les potentiels et les cliques des textures retenues, autrement que par référence aux primitives de la matrice de cooccurrence.

4.5 Les méthodes structurales

Ces méthodes ne s'appliquent qu'aux textures construites sur une primitive bien identifiable qui se répète sur une maille régulière. L'analyse d'une texture homogène se fait en plusieurs étapes :

1. l'image est découpée en sous-fenêtres, chacune contenant plusieurs primitives (typiquement de l'ordre de 5×5 primitives),
2. sur chaque sous fenêtre on détermine les 2 vecteurs moyens de périodicité, par exemple par l'examen de la fonction d'auto-corrélation,
3. connaissant ces vecteurs, chaque sous-image est normalisée par un rééchantillonnage de façon à ce que les nouveaux vecteurs de périodicité soient tous dans une position de référence,
4. on calcule alors par TF le motif de la primitive en sommant dans l'espace de Fourier toutes les sous-fenêtres normalisées et en isolant le sous-domaine du fondamental,
5. on est alors en mesure de reconstruire une texture parfaite en répétant le motif moyen sur la maille moyenne, ou des textures plus ou moins parfaites en mélangeant des motifs réels ou moyens sur des mailles réelles ou moyennes.

Des méthodes plus structurales encore s'appuient sur des techniques de graphes ou de grammaires. Dans les approches par graphes, on représente les dépendances entre les diverses composantes de la texture par des graphes (arbres, cycles) mettant en évidence les dépendances entre pixels. Dans les approches par grammaires, on parcourt la texture ou des sous-ensembles de la texture par un balayage systématique de tous les sites, construisant ainsi une chaîne de descripteurs. Les relations entre les descripteurs successifs sont exprimées par des règles qui indiquent l'enchaînement des niveaux de gris le long du parcours. Les chaînes font alors l'objet d'un traitement syntaxique qui extrait ces dépendances [Lu et Fu, 1978, Lu et Fu, 1979].

4.6 Textures hétérogènes

Si l'on fait l'hypothèse que la texture sous-jacente est stationnaire, on peut déduire des variations de ses propriétés statistiques dans l'image, des éléments suffisants pour remonter à une connaissance de certains éléments de relief et d'orientation 3D des objets de la scène observée. Cela est naturellement fait par système visuel humain (cf. figure 4.2). Pour faire cela par traitement numérique, il convient de mesurer en chaque point de l'image (c'est-à-dire en des zones suffisamment petites autour de chaque point) des propriétés statistiques d'ordre 2 (prenant en compte les pixels 2 par 2, comme par exemple la décroissance de la corrélation à $1/2$), puis d'étudier les variations spatiales de ces propriétés. Ces variations sont alors reliées à l'orientation de la surface observée (généralement sous l'hypothèse que celle-ci fait un angle constant avec la direction d'observation).

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.

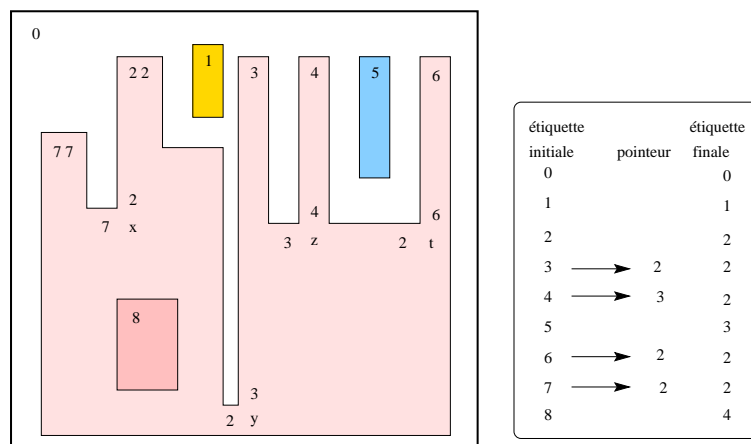


FIG. 5.1 – Etiquetage d’une image en 4 connexité. L’image de gauche est le résultat du premier passage jusqu’au point noté x. Les étiquettes ont été attribuées à chaque point soit en donnant une nouvelle étiquette si les 2 antécédents du point sont à 0, soit en donnant l’étiquette des antécédents si 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).

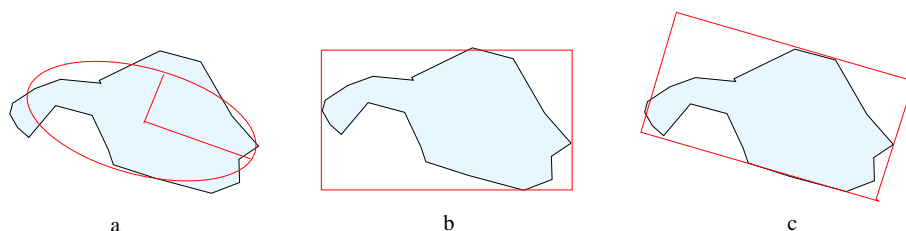


FIG. 5.2 – Description de formes par : a - ellipse et axes d'inertie, b - boîte englobante, c - boîte minimale.

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 ;

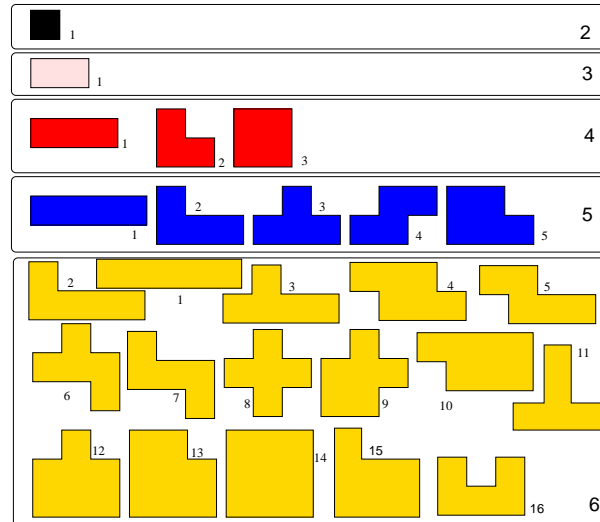


FIG. 5.3 – Les premières formes du dictionnaire de Guzman.

- 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).

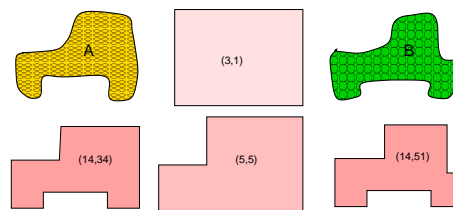


FIG. 5.4 – Représentation de Guzman. Les 2 formes A et B sont semblables aux niveaux de représentation 3 et 5, mais diffèrent au niveau 14, où l'une est représentée par le prototype 34 et l'autre par le prototype 51.

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-).

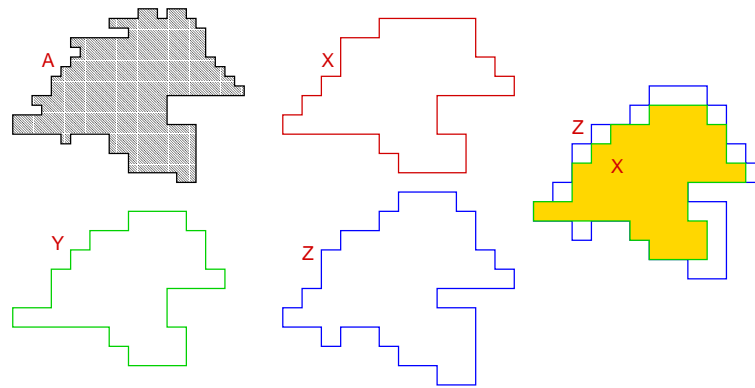


FIG. 5.5 – L'arbitraire de l'approximation des formes : la forme A est l'objet. Les 3 formes X, Y et Z en sont des approximations. La distance d'une représentation à la forme initiale peut se mesurer par la surface entre les 2 contours. Il y a de nombreuses formes à distance minimale. X est une forme 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.

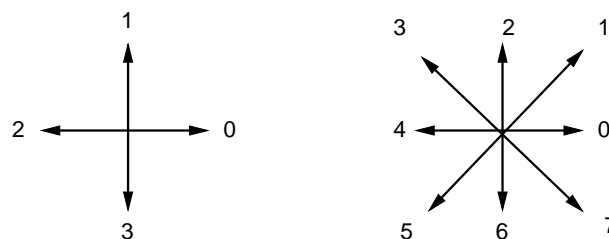


FIG. 5.6 – Les codes de Freeman en 4-connextité (à gauche) et en 8-connextité (à droite).

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].

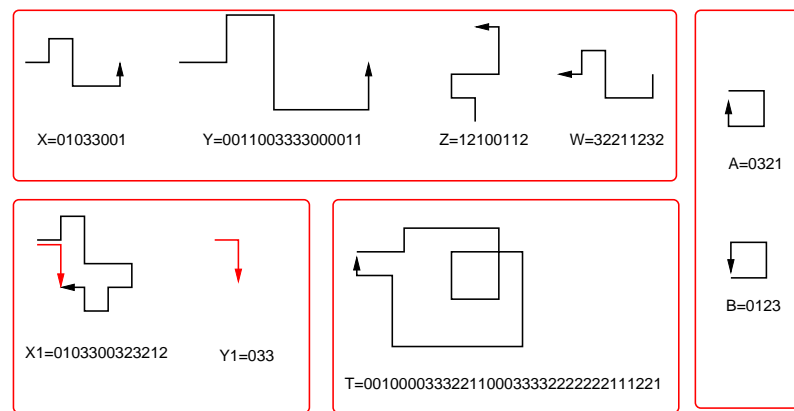


FIG. 5.9 – Quelques propriétés des chaînes de Freeman : X = chaîne originale, Y = contour double, Z = rotation de $\pi/2$, W = contour inverse. $X1$ = chaîne originale et $Y1$ = chaîne réduite. T = exemple de boucle dans une chaîne. A et B sont les deux contours élémentaires (inverse et direct) en 4-connexité.

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.

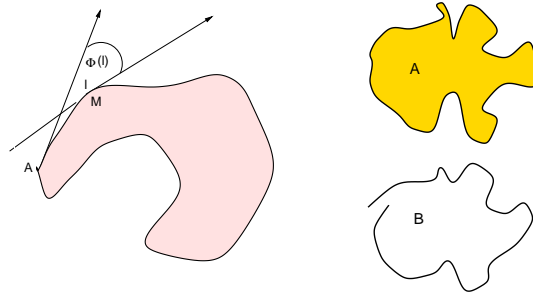


FIG. 5.10 – A gauche : descripteurs de Fourier par tangente. Le point de départ A est choisi arbitrairement. Son vecteur tangent sert de référence à la paramétrisation de la courbe. A droite : après troncature du développement de Fourier, la forme A, initialement fermée devient B, plus régulière, mais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

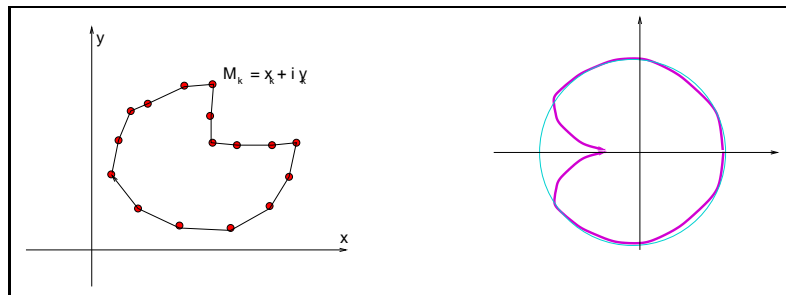


FIG. 5.11 – A gauche : descripteurs de Fourier par représentation complexe. Le point courant M_k est décrit par ses coordonnées complexes dans le plan image. A droite : le cercle résulte de la prise en compte du seul coefficient Z_1 , la courbe présentée résulte de l'adjonction d'un coefficient de pression (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 $\Delta : y = a_0 + a_1x$ qui minimise la distance (cf. Fig 5.12) :

$$d_1^2 = \sum_i^N [y_i - (a_0 + a_1x_i)]^2 \quad (5.1)$$

La solution est donnée par : $A = X^\#Y = (X^tX)^{-1}X^tY$, 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.

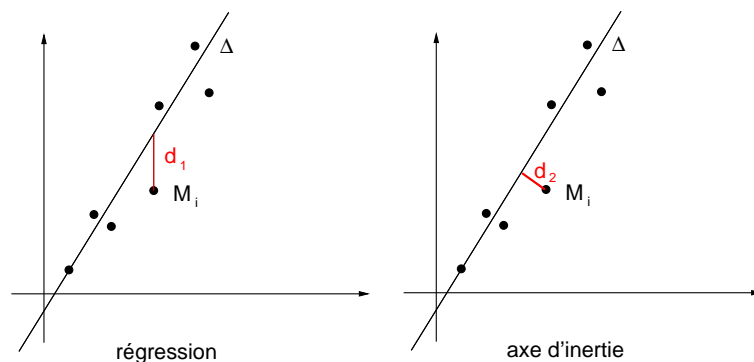


FIG. 5.12 – Deux approximations linéaires aux moindres carrés minimisant des distances différentes : à gauche par régression linéaire, à droite par axe d’inertie.

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 $50\% - \epsilon$ 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 indiquée 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 :

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

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 $n - 1$.

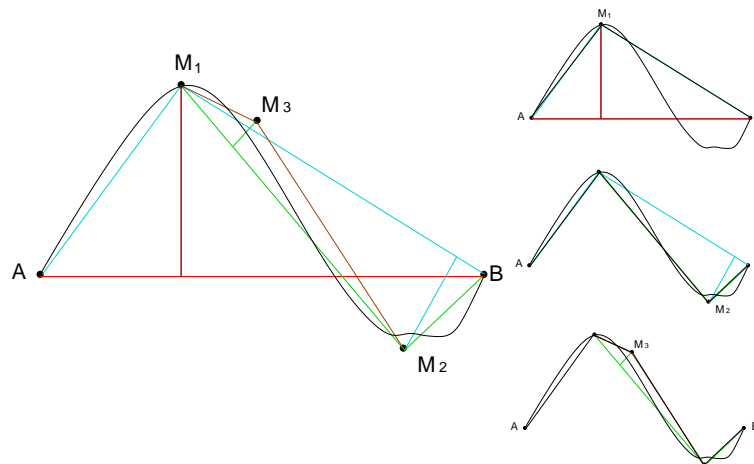


FIG. 5.13 – Algorithme de la corde. Les points M_1 , M_2 et M_3 sont successivement sélectionnés pour créer la ligne polygonale représentant la courbe AB . Le critère de sélection est la distance maximale à la corde polygonale précédemment obtenue.

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.

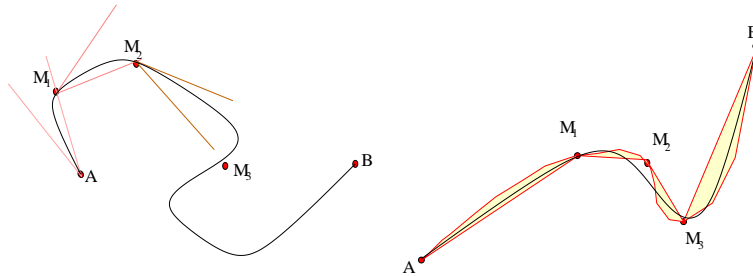


FIG. 5.14 – Algorithme de Dunham, à gauche, les points M_1 , M_2 et M_3 sont choisis lorsque la tolérance angulaire ϵ est insuffisante à approximer la courbe. A droite, algorithme de Wahl et Danielsson : il cherche un minimum des surfaces laissées entre la ligne polygonale et la courbe.

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 :

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

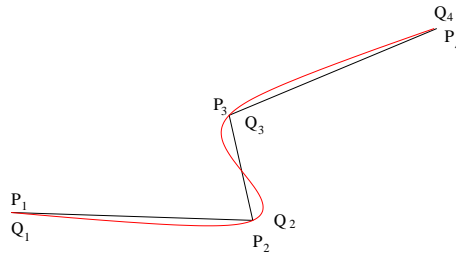


FIG. 5.15 – Dans le cas d'une interpolation spline, les jonctions entre les morceaux de splines sont les points de contrôle eux-mêmes.

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 (trouver des droites) en un problème mieux connu : retrouver des nuages de points. Ceci s'obtient en associant à l'espace de l'image (dénomé \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 l à m et la transformation de Hough de m à l .

5.7.1 Définitions

Transformation de l 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 l 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 l

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

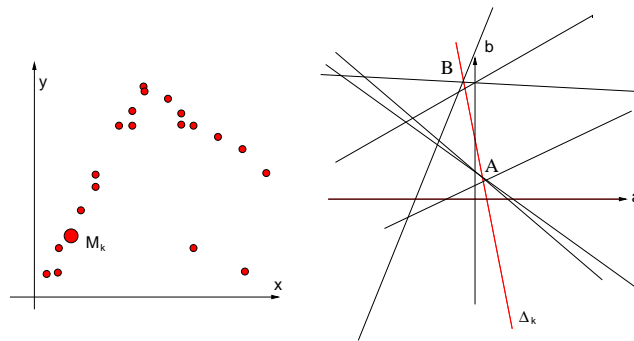


FIG. 5.16 – Transformation de Hough de 1 à m . Un point M_k est transformé en une droite Δ_k de l'espace des paramètres $\{a, b\}$. On a représenté quelques-unes seulement de toutes les droites existantes. Elles définiront finalement 2 points d'intersection, l'un en A, l'autre en B.

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.

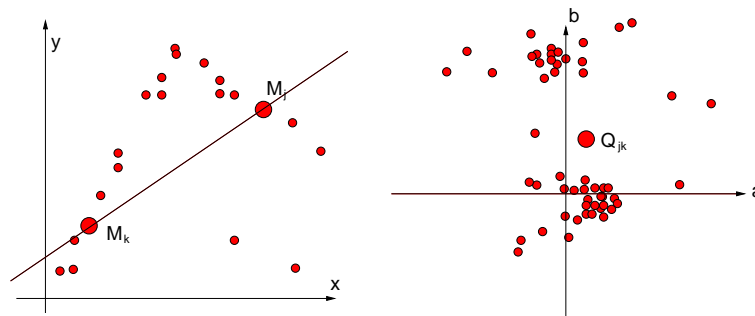


FIG. 5.17 – Transformation de Hough de m à 1 . Deux points M_j, M_k définissent une droite qui se transforme en 1 point Q_{jk} de l'espace des paramètres (a, b) .

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 :

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

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 :

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

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} .

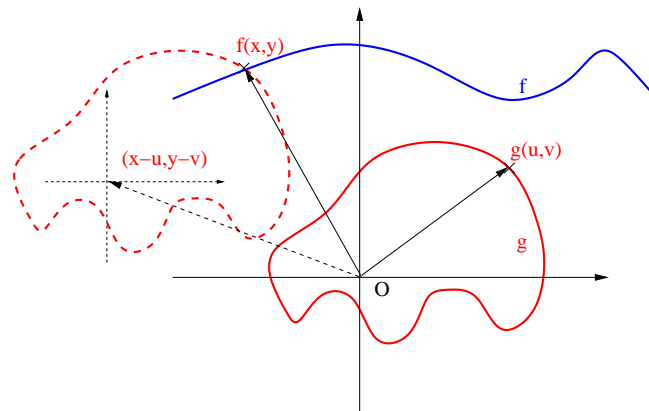


FIG. 5.18 – Transformation de Hough : le point de coordonnées x, y de f sera associé au point $g(u, v)$ de g si l'on fait subir à cette courbe g une translation de $(x - u, y - v)$.

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

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 l associée au doublet M_i, M_j le point de \mathcal{H} de coordonnées :

$$\rho_{ij} = \frac{|x_i y_j - x_j y_i|}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}$$

$$\theta_{ij} = -\text{Arctg} \left[\frac{x_j - x_i}{y_j - y_i} \right]$$

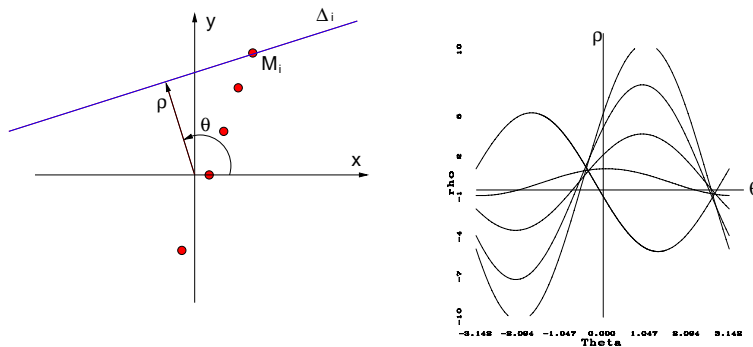


FIG. 5.19 – Transformation de Hough de l à m en coordonnées normales. Une droite Δ_i est repérée par l'angle θ et la distance au pied de sa normale ρ . Dans l'espace \mathcal{H} , chaque point donne une sinusoïde.

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.

Chapitre 6

Éléments de géométrie tridimensionnelle

Chapitre rédigé par Isabelle BLOCH et Henri MAÎTRE

Les images sont des représentations bi-dimensionnelles de l'espace tri-dimensionnel qui nous entoure. La structure des images est donc régie par les lois des transformations géométriques qui accompagnent la formation de l'image. Ces transformations ont fait l'objet d'études intensives ces dernières années [Faugeras, 1993, Mohr et al., 1994, Gallice, 2002], qui font intervenir des outils adaptés de représentation et de traitement de l'espace. C'est l'objet de ce chapitre de présenter ces outils, leur rôle dans la formation de l'image et les principaux résultats qui leurs sont attachés. Nous verrons tout d'abord la structuration de l'espace et des transformations de l'espace et introduirons l'espace projectif qui permet de traiter linéairement les transformations les plus utiles. Nous examinerons ensuite les diverses représentations possibles des rotations. Nous aborderons enfin la géométrie différentielle qui permet de décrire la surface des objets de l'espace.

6.1 Géométrie projective

Les principaux éléments de géométrie projective peuvent être trouvés dans [Faugeras, 1993].

6.1.1 Espace projectif et coordonnées homogènes

Un espace projectif de dimension n , noté \mathcal{P}^n est un espace dans lequel un point x est représenté par un vecteur de $n + 1$ coordonnées :

$$x = (x_1, \dots, x_{n+1})^t \quad (6.1)$$

dans lequel au moins une coordonnée x_i est non nulle. Ces coordonnées sont appelées coordonnées homogènes.

Deux vecteurs $(x_1, \dots, x_{n+1})^t$ et $(y_1, \dots, y_{n+1})^t$ représentent le même point si et seulement si il existe un réel non nul λ tel que $\forall i, x_i = \lambda y_i$.

Un point pour lequel $x_{n+1} = 0$ est appelé point à l'infini.

Les transformations linéaires d'un espace projectif \mathcal{P}^m dans un autre \mathcal{P}^n sont appelées colinéations. Une colinéation est définie par une matrice H de taille $(m + 1) \times (n + 1)$ de déterminant non nul, à un facteur non nul près (H et λH définissent la même transformation).

Une base projective est un ensemble de $n + 2$ points, tel que tous les $(n + 1)$ -uplets soient linéairement indépendants. La base projective standard est définie par :

$$\forall i \leq n + 1, e_i = (0, \dots, 1, \dots, 0)^t \text{ et } e_{n+2} = (1, \dots, 1)^t \quad (6.2)$$

où la i -ème coordonnée de e_i est égale à 1. Tout point x de \mathcal{P}^n s'écrit comme combinaison linéaire de n'importe quel $(n + 1)$ -uplet de la base, par exemple :

$$x = \sum_{i=1}^{n+1} x_i e_i. \quad (6.3)$$

Un changement de base est défini de manière unique par une colinéation, représentée par une matrice inversible de taille $(n + 1) \times (n + 1)$.

On peut montrer que l'espace projectif \mathcal{P}^n est topologiquement équivalent à la sphère unité S^n de \mathbf{R}^{n+1} repliée, c'est-à-dire dans laquelle on identifie les points aux antipodes. La source principale de ce résultat est qu'un point $x = (x_1, \dots, x_{n+1})^t$ est un point de S^n si $\sum_{i=1}^{n+1} x_i^2 = 1$. Un tel point est également un point de \mathcal{P}^n . Le point $-x$, aux antipodes de x , est également sur la sphère unité mais représente le même point de \mathcal{P}^n . Une des conséquences de ce résultat est que l'espace projectif \mathcal{P}^n est compact, ce qui peut sembler surprenant puisqu'intuitivement on manipule dans \mathcal{P}^n des points à l'infini. Les sphères S^1 et S^2 repliées (dans \mathbf{R}^2 et \mathbf{R}^3 respectivement) sont utilisées pour représenter des directions, et la sphère S^3 de \mathbf{R}^4 repliée pour représenter des rotations à l'aide de la notion de quaternions (nous y reviendrons plus loin). Le théorème d'équivalence permet de manipuler \mathcal{P}^n plutôt que S^n , ce qui est plus simple.

6.1.2 Espaces projectifs \mathcal{P}^1 , \mathcal{P}^2 et \mathcal{P}^3

L'espace \mathcal{P}^1 est appelé droite projective. Sa base standard est composée des points $e_1 = (1, 0)^t$, $e_2 = (0, 1)^t$, $e_3 = (1, 1)^t$. Ainsi tout point x de la droite projective peut s'écrire par exemple :

$$x = x_1 e_1 + x_2 e_2 \quad (6.4)$$

avec $x_1 \neq 0$ ou $x_2 \neq 0$. Si l'on se restreint aux points tels que $x_2 \neq 0$, ce qui revient à exclure le point e_1 , alors x s'écrit de manière équivalente comme :

$$x = \alpha e_1 + e_2 \quad (6.5)$$

avec $\alpha = \frac{x_1}{x_2}$. Le point e_1 est obtenu pour $\alpha = \pm\infty$ et est appelé point à l'infini.

L'espace \mathcal{P}^2 est appelé plan projectif. Il est très important pour la représentation d'une image. Un point x de \mathcal{P}^2 est défini par un triplet de coordonnées (x_1, x_2, x_3) non toutes nulles dans la base standard (e_1, \dots, e_4) . Une droite d de \mathcal{P}^2 est également définie par un triplet de nombres (d_1, d_2, d_3) non tous nuls. Un point x appartient à la droite d si et seulement si :

$$d^t x = 0 \Leftrightarrow \sum_{i=1}^3 d_i x_i = 0. \quad (6.6)$$

Cette équation de droite fait apparaître une dualité entre point et droite et il n'y a pas de différence formelle entre les deux. Cette équation représente également l'ensemble des droites qui passent par x .

Une droite passant par deux points x et y est représentée par le produit vectoriel $d = x \wedge y$. En effet, tout point z de la droite peut s'écrire $z = \alpha x + \beta y$ pour α et β quelconques, ce qui signifie que le déterminant de (z, x, y) est nul, soit encore $z^t(x \wedge y) = 0$. Par dualité, le point d'intersection de deux droites d et d' est représenté par le produit vectoriel $d \wedge d'$. On en déduit que deux droites s'intersectent toujours dans l'espace projectif \mathcal{P}^2 .

La droite obtenue pour $x_3 = 0$ est appelée droite à l'infini.

L'espace projectif \mathcal{P}^3 a pour base standard $(e_1, \dots, e_5)^t$ définie comme précédemment, et un point x de \mathcal{P}^3 est défini (toujours à un facteur non nul près) par 4 coordonnées (x_1, x_2, x_3, x_4) non toutes nulles. Un plan π de \mathcal{P}^3 est également défini par quatre nombres non tous nuls $(\pi_1, \pi_2, \pi_3, \pi_4)$ et l'équation du plan est :

$$\pi^t x = 0. \quad (6.7)$$

Cette équation fait apparaître une dualité entre point et plan cette fois, et représente également l'ensemble des plans qui passent par le point x .

Une droite de \mathcal{P}^3 peut être définie soit à partir de deux points, soit comme intersection de deux plans. Par dualité, une droite est donc transformée en une droite.

Le plan à l'infini est obtenu pour $x_4 = 0$. Tout plan intersecte le plan à l'infini selon une droite à l'infini.

6.1.3 Expression linéaire de problèmes géométriques

Un des intérêts de la géométrie projective est de permettre d'exprimer certains problèmes géométriques simplement par des équations d'algèbre linéaire. Nous en donnons ici quelques exemples.

Une conique¹ de \mathcal{P}^2 est définie par l'expression quadratique suivante :

$$x^t C x = 0 \quad (6.8)$$

où C est une matrice de taille 3×3 définie à un facteur non nul près (elle a donc 5 coefficients indépendants).

L'intersection d'une droite et d'une conique peut alors être exprimée facilement. Soit $x + \alpha y$ un point de la droite définie par les deux points x et y . Ce point appartient à la conique si et seulement si :

$$(x + \alpha y)^t C (x + \alpha y) = 0 \quad (6.9)$$

soit :

$$x^t C x + 2\alpha x^t C y + \alpha^2 y^t C y = 0 \quad (6.10)$$

et la résolution de cette équation en α donne les intersections de la droite et de la conique (deux solutions en général).

Dans le cas où le discriminant de l'équation est nul, les deux points d'intersection sont confondus et la droite est alors tangente à la conique. La tangente t à la conique au point x est alors donnée par :

$$t = C x. \quad (6.11)$$

Le rang de la conique (rang de la matrice C) permet de classer les coniques. Il est important de noter que ce rang est un invariant projectif, c'est-à-dire qu'il est indépendant de la base projective choisie. C'est donc bien une caractéristique intrinsèque de la conique. Les coniques non-dégénérées (de rang 3) sont les cercles, les ellipses, les paraboles et les hyperboles du plan. Les coniques dégénérées, de rang 2 ou 1, sont respectivement des paires de droites et des droites.

De manière similaire, une quadrique dans \mathcal{P}^3 est définie par une matrice Q de taille 4×4 (à un facteur non nul près) et l'équation suivante :

$$x^t Q x = 0. \quad (6.12)$$

Le plan π tangent à la quadrique au point x est donné par $\pi = Q x$.

Le rang de la matrice Q définit le rang de la quadrique et est un invariant projectif. Les quadriques non-dégénérées (de rang 4) sont les sphères, les ellipsoïdes, les paraboloides et les hyperboloïdes de l'espace. Une quadrique dégénérée de rang 3 est un cône, si le rang vaut deux c'est une paire de plans, et si le rang vaut 1 c'est un plan.

L'intersection d'un plan π et d'une quadrique est une conique de π . L'intersection d'une droite et d'une quadrique s'obtient en résolvant une équation du second degré similaire à celle obtenue pour une conique, en remplaçant C par Q .

¹Les coniques prendront une place importante en vision par ordinateur dans toutes les opérations de localisation et d'orientation des caméras à partir de points connus de la scène, par exemple lors de la recherche des formes fondamentales en section 6.3.2.

6.1.4 Stratification, groupes de transformations et invariants

Le monde est en général considéré comme un espace euclidien tridimensionnel. C'est l'espace le plus complet. Souvent nous n'en avons qu'une représentation partielle, par exemple sous forme d'une image, et la structure réduite plus simple de l'espace projectif permet alors de manipuler cette représentation. Les espaces métriques et affines sont intermédiaires, ce qui permet de construire une stratification des espaces, comme le montre le tableau 6.1. A chaque niveau sont associés un groupe de transformations linéaires et des invariants. L'espace projectif étant le moins structuré, c'est aussi celui qui a le plus grand groupe de transformations et le plus petit nombre d'invariants. D'un niveau à l'autre, le groupe des transformations est un sous-groupe du groupe des transformations du niveau précédent. Ainsi les transformations euclidiennes forment le plus petit groupe.

Espace	Degrés de liberté	Forme des transformations linéaires	Invariants
Projectif	15	$T_P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ p_{41} & p_{42} & p_{43} & p_{44} \end{pmatrix}$	bi-rapport
Affine	12	$T_A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix}$	distances relatives le long de directions, parallélisme
Métrique	7	$T_M = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & t_x \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & t_y \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$	distances relatives, angles
Euclidien	6	$T_E = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$	distances absolues

TAB. 6.1 – Stratification des espaces, du moins structuré au plus complet.

Dans l'espace projectif, le groupe des transformations est celui des colinéations, le groupe le plus général des transformations linéaires. Les seize coefficients d'une matrice 4×4 représentant une collinéation de \mathcal{P}^3 sont définies à un facteur non nul près, ce qui laisse 15 degrés de liberté.

Les invariants de \mathcal{P}^3 sont, outre les relations d'incidence, de colinéarité et de tangence, les bi-rapports, définis de la manière suivante. Soient quatre points colinéaires de paramètres λ_i . La droite portant ces points est par exemple définie par les points x et y et chacun des quatre points est représenté par $x + \lambda_i y$. Le bi-rapport des points est alors :

$$\frac{\frac{\lambda_1 - \lambda_3}{\lambda_1 - \lambda_4}}{\frac{\lambda_2 - \lambda_3}{\lambda_2 - \lambda_4}} \quad (6.13)$$

et est un invariant projectif : il ne dépend pas des choix de x et y , et est invariant par toute transformation projective T_P .

Ces bi-rapports peuvent également être calculés pour 4 droites s'intersectant en un point ou 4 plans s'intersectant en une droite. Ce sont également des invariants.

Le passage de l'espace projectif à l'espace affine se fait en éliminant les points à l'infini. C'est la transformation qui fait passer de la projection perspective (où les fuyantes ont un point de fuite en commun) à la perspective cavalière (où les fuyantes sont parallèles), voir figure 6.1.

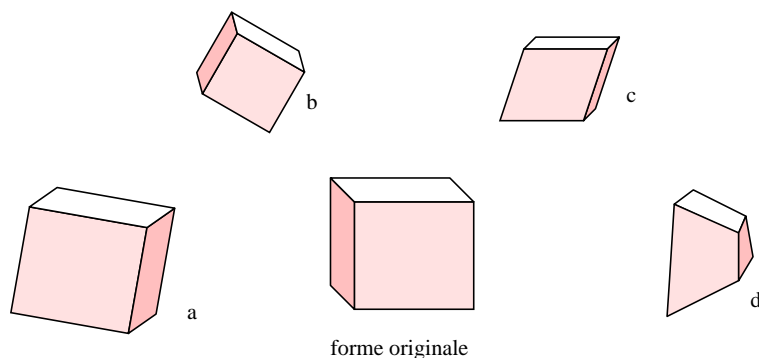


FIG. 6.1 – Dans une transformation euclidienne, le cube original se transforme : en a) par rotation et translation (transformation euclidienne); en b), par une transformation métrique; en c), par une affinité (sous forme d'un parallélogramme); en d), par une projection perspective.

Par exemple, le passage de la droite projective à la droite affine se fait par une bijection pour $\alpha \neq \pm\infty$. De même le passage entre l'espace affine 3D et \mathcal{P}^3 se fait par la bijection suivante :

$$(x, y, z)^t \mapsto (x, y, z, 1)^t. \quad (6.14)$$

Le plan à l'infini de \mathcal{P}^3 peut être considéré comme les points limites de norme infinie, donc représentés par les coordonnées homogènes $(x_\infty, y_\infty, z_\infty, 0)^t$, qui n'ont pas d'équivalent dans l'espace affine.

En utilisant cette bijection, les transformations affines sont les transformations linéaires T_A dont la forme générale est donnée dans le tableau 6.1. Elles ont 12 degrés de liberté et laissent le plan à l'infini invariant.

Les invariants projectifs sont a fortiori des invariants affines. De plus, le parallélisme est un invariant des transformations affines, ainsi que les rapports des longueurs dans une direction donnée (correspondant au birapport lorsque deux des points sont à l'infini).

Le groupe des transformations de l'espace métrique est le groupe des similarités (composées de rotations, translations et homothéties). La forme générale d'une transformation métrique T_M est montrée dans le tableau 6.1. En coordonnées cartésiennes, une telle transformation s'écrit comme la succession d'une rotation (matrice orthonormale 3×3), puis un produit par un facteur d'homothétie, puis une translation (vecteur de longueur 3). Les coordonnées homogènes permettent de regrouper ces trois termes dans une seule matrice.

Cette fois les distances relatives et les angles sont des invariants supplémentaires.

Dans l'espace euclidien, le facteur d'homothétie est fixé et les transformations euclidiennes sont donc les transformations rigides (6 degrés de liberté au lieu des 7 des transformations métriques). Les distances absolues sont donc cette fois des invariants.

L'ensemble des transformations linéaires de \mathcal{P}^3 forment donc un emboîtement de sous-groupes, appelé stratification, tel que représenté sur la figure 6.2.

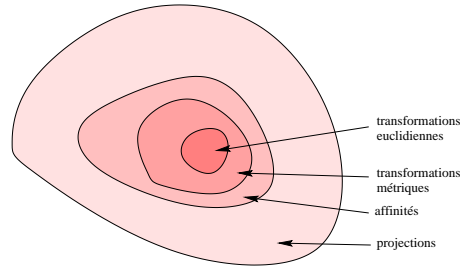


FIG. 6.2 – Les transformations d'espaces qui s'expriment linéairement en représentation projective forment un emboîtement appelé stratification.

6.2 Représentations des rotations de \mathbb{R}^3

6.2.1 Les angles d'Euler

Les rotations de \mathbb{R}^3 sont définies par des matrices R orthogonales (satisfaisant $RR^t = I$, où I est la matrice identité) et de déterminant égal à 1. Elles peuvent se décomposer sous la forme de trois rotations à 2 dimensions, autour de chacun des axes du repère. Soit α l'angle de la rotation autour de l'axe x (dans le plan yz), β l'angle de la rotation autour de y et γ l'angle de rotation autour de z . Ces angles sont appelés angles d'Euler de la rotation (figure 6.3).

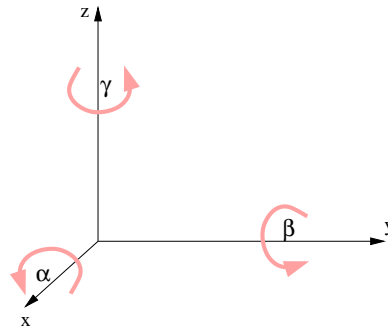


FIG. 6.3 – Représentation de la rotation par les angles d'Euler.

Si l'on note $R(\alpha, \beta, \gamma)$ une telle matrice, elle peut s'écrire :

$$R(\alpha, \beta, \gamma) = R(0, 0, \gamma) \circ R(0, \beta, 0) \circ R(\alpha, 0, 0)$$

où la loi de composition \circ s'effectue de la droite à la gauche. On note que l'ordre dans lequel s'effectue ces rotations est important pour le résultat final.

Finalement $R = R(\alpha, \beta, \gamma)$ s'écrit :

$$R = \begin{pmatrix} \cos \gamma \cos \beta & -\cos \gamma \sin \beta \sin \alpha - \sin \gamma \cos \alpha & -\cos \gamma \sin \beta \cos \alpha + \sin \gamma \sin \alpha \\ \sin \gamma \cos \beta & -\sin \gamma \sin \beta \sin \alpha + \cos \gamma \cos \alpha & -\sin \gamma \sin \beta \cos \alpha - \cos \gamma \sin \alpha \\ \sin \beta & \cos \beta \sin \alpha & \cos \beta \cos \alpha \end{pmatrix}. \quad (6.15)$$

Inversement, la connaissance de la matrice de rotation R permet en théorie de retrouver les angles d'Euler. Posons $u = \sqrt{R_{11}^2 + R_{21}^2}$. On a :

- si $u \neq 0$ alors : $\cos \alpha = R_{33}/u$, $\cos \beta = u$ et $\cos \gamma = R_{11}/u$,
- si $u = 0$ alors $\beta = \pm\pi/2$ et on ne peut connaître que $\alpha + \gamma$.

Mais cette déduction mathématique ignore totalement la présence possible de bruit sur les coefficients R_{ij} de la matrice R . La résolution doit en pratique prendre en compte la propriété de la matrice R d'être unitaire pour régulariser la solution.

Parallèlement à cette écriture classique, il existe d'autres formulations, souvent plus élégantes, qui sont souvent utilisées en traitement d'images et en vision par ordinateur [Faugeras, 1993], et que nous décrivons dans les paragraphes suivants.

6.2.2 Exponentielle d'une matrice anti-symétrique

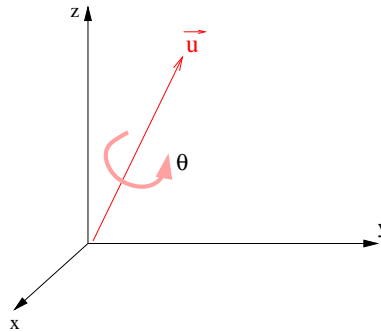


FIG. 6.4 – Représentation canonique de la rotation.

Une rotation de \mathbb{R}^3 laisse une droite invariante, appelée axe de la rotation. Elle peut alors être définie à partir de cet axe et d'un angle θ (figure 6.4). C'est la représentation dite canonique.

Etant donnée une rotation de \mathbb{R}^3 d'axe quelconque, il est possible de la représenter à partir d'une matrice anti-symétrique, sans revenir à la décomposition en trois rotations selon les trois axes du repère.

Plus précisément, soit la matrice anti-symétrique

$$H = \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}. \quad (6.16)$$

On définit l'exponentielle de H par la somme :

$$e^H = \sum_{n=0}^{\infty} \frac{H^n}{n!}. \quad (6.17)$$

On peut alors montrer que e^H est une matrice de rotation d'axe $\vec{u} = (a, b, c)^t$ et d'angle $\theta = \|\vec{u}\| [2\pi]$. La preuve de ce résultat peut être trouvée par exemple dans [Faugeras, 1993].

6.2.3 Forme canonique et théorème de Rodrigues

Le théorème de Rodrigues permet d'éviter le calcul de la somme infinie et de la ramener à la somme de trois termes uniquement. Ce théorème exprime la rotation R sous la forme suivante :

$$R = e^H = I + \frac{\sin \theta}{\theta} H + \frac{1 - \cos \theta}{\theta^2} H^2. \quad (6.18)$$

On a ainsi une expression simple d'une rotation d'axe et angle quelconques. En effet, le calcul des puissances successives de H donne :

$$H^2 = \begin{pmatrix} -c^2 - b^2 & ab & ac \\ ab & -c^2 - a^2 & bc \\ ac & bc & -a^2 - b^2 \end{pmatrix},$$

$$H^3 = -(a^2 + b^2 + c^2) \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix} = -\theta^2 H,$$

$$H^4 = -\theta^2 H^2,$$

$$H^5 = \theta^4 H,$$

$$H^{2n} = (-1)^{n-1} \theta^{2n-2} H^2 \quad \text{pour } n \geq 1,$$

$$H^{2n+1} = (-1)^n \theta^{2n} H \quad \text{pour } n \geq 0.$$

De plus, on a :

$$\sin \theta = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n+1}}{n!},$$

$$\cos \theta = \sum_{n=0}^{\infty} (-1)^n \frac{\theta^{2n}}{n!},$$

dont on déduit le théorème.

6.2.4 Quaternions

Les quaternions ont été introduits pour étendre les propriétés des nombres complexes (qui représentent élégamment les rotations du plan) à des espaces de dimension plus élevée. Ils sont définis comme des vecteurs de \mathbb{R}^4 , sur lesquels on définit une multiplication particulière, non commutative. Un quaternion $(q_1, q_2, q_3, q_4)^t$ peut être représenté de manière équivalente sous la forme (s, v) , où s est un réel et v un vecteur de \mathbb{R}^3 , avec $s = q_1$ et $v = (q_2, q_3, q_4)^t$. Par analogie avec les complexes, on appelle s la partie réelle du quaternion et v sa partie imaginaire. Un réel pur a alors la forme $(s, 0)$ et un imaginaire pur (ou quaternion vectoriel) a la forme $(0, v)$.

Le produit de deux quaternions $q = (s, v)$ et $q' = (s', v')$ est défini de la manière suivante :

$$q \times q' = (ss' - v \cdot v', sv' + s'v + v \wedge v') \quad (6.19)$$

où \cdot et \wedge désignent les produits scalaire et vectoriel classiques de \mathbb{R}^3 . En particulier, le produit de deux quaternions réels est le produit classique sur \mathbb{R} , alors que le produit de deux quaternions vectoriels est leur produit vectoriel.

Le conjugué d'un quaternion $q = (s, v)$ est le quaternion \bar{q} défini par :

$$\bar{q} = (s, -v) \quad (6.20)$$

La norme d'un quaternion est définie de manière naturelle par :

$$|q|^2 = \bar{q} \times q = q \times \bar{q} = (s^2 + \|v\|^2, 0) = (\|q\|^2, 0), \quad (6.21)$$

où $\|q\|$ désigne la norme euclidienne de q dans \mathbf{R}^4 . La norme d'un quaternion est un réel pur. Cette équation fait apparaître que la multiplication d'un quaternion avec son conjugué commute, mais ce n'est pas le cas du produit de deux quaternions quelconques.

La norme du produit de deux quaternions vaut simplement $|q \times q'| = |q||q'|$.

L'intérêt des quaternions ici réside dans leur lien avec les rotations de \mathbf{R}^3 . En effet, une rotation d'axe \vec{u} et d'angle θ peut être représentée par le quaternion (s, v) tel que :

$$s = \cos \frac{\theta}{2}, \quad (6.22)$$

et

$$v = \sin \frac{\theta}{2} \vec{u}. \quad (6.23)$$

Notons que le quaternion $(-s, -v)$ représente la même rotation. Ce n'est pas surprenant puisque une rotation d'axe \vec{u} et d'angle θ est équivalente à une rotation d'axe $-\vec{u}$ et d'angle $2\pi - \theta$.

Définissons la relation \mathcal{R} par $\mathcal{R}(q, q') \Leftrightarrow q = -q'$. Cette relation est une relation d'équivalence. Soit \mathcal{Q}_1 l'ensemble des quaternions de norme 1. Muni du produit défini plus haut, c'est un groupe (non abélien). L'ensemble des rotations de \mathbf{R}^3 muni de la composition (produit des matrices de rotations) est également un groupe \mathcal{R}^3 . On peut montrer que \mathcal{R}^3 est isomorphe au groupe quotient $\mathcal{Q}_1/\mathcal{R}$. Cet isomorphisme est établi par la relation entre (s, v) et (θ, \vec{u}) exprimée plus haut, et par la préservation de la structure de groupe : si deux rotations r_1 et r_2 sont représentées par les quaternions q_1 et q_2 (en choisissant une des deux représentations possibles pour chaque rotation), alors leur composition $r_1 \circ r_2$ est représentée par $q_1 \times q_2$ (et $-q_1 \times q_2$).

Si maintenant une rotation est représentée sous forme matricielle R et si q est un des deux quaternions correspondants, on a pour tout vecteur x de \mathbf{R}^3 :

$$(0, Rx) = q \times (0, x) \times \bar{q} \quad (6.24)$$

ou sous forme simplifiée en identifiant un vecteur et un quaternion :

$$Rx = q \times x \times \bar{q}. \quad (6.25)$$

On peut alors exprimer les coefficients de R à partir de ceux de q . Soit $q = (s, v)$, avec $v = (l, m, n)^t$ et $l^2 + m^2 + n^2 + s^2 = 1$. On a :

$$\begin{aligned} q \times x \times \bar{q} &= (s, v) \times (0, x) \times (s, -v) \\ &= (-v \cdot x, sx + v \wedge x) \times (s, -v) \\ &= (-sv \cdot x + v \cdot (sx + v \wedge x) + (sx + v \wedge x) \wedge (-v)) \\ &= (0, 2(v \cdot x)v + (s^2 - v^2)x + 2sv \wedge x). \end{aligned}$$

Le produit vectoriel peut s'écrire comme produit d'une matrice anti-symétrique et d'un vecteur. En effet, soit V la matrice définie à partir de v par :

$$V = \begin{pmatrix} 0 & -n & m \\ n & 0 & -l \\ -m & l & 0 \end{pmatrix}.$$

Alors on a $v \wedge x = Vx$. De même en posant $A = (lv, mv, nv)$ on peut écrire sous forme matricielle $(v \cdot x)v = Ax$. On a alors :

$$R = 2A + (s^2 - v^2)I + 2sV \quad (6.26)$$

$$= \begin{pmatrix} s^2 + l^2 - m^2 - n^2 & 2(lm - sn) & 2(ln + sm) \\ 2(lm + sn) & s^2 - l^2 + m^2 - n^2 & 2(mn - sl) \\ 2(ln - sm) & 2(mn + sl) & s^2 - l^2 - m^2 + n^2 \end{pmatrix}. \quad (6.27)$$

Faisons maintenant le lien avec l'espace projectif \mathcal{P}^3 . L'ensemble des quaternions de norme 1, \mathcal{Q}_1 , est de manière immédiate isomorphe à la sphère unité S^3 de \mathbb{R}^4 . On en déduit que le groupe quotient $\mathcal{Q}_1/\mathcal{R}$ est isomorphe à la sphère S^3 repliée, dans laquelle on identifie les points aux antipodes. On en déduit donc que le groupe des rotations \mathcal{R}^3 est isomorphe à l'espace projectif \mathcal{P}^3 .

En coordonnées homogènes, une rotation d'axe x et d'angle θ par exemple s'exprime dans \mathcal{P}^3 par :

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (6.28)$$

Une des applications typiques des quaternions est le recalage rigide, où l'on cherche à optimiser la transformation entre deux ensembles de points ou deux objets. Soient $X = \{x_i, i = 1..n\}$ et $Y = \{y_i, i = 1..n\}$ les deux ensembles de points à recaler, le point x_i correspondant au point y_i (on suppose ici que cette correspondance est connue). Soit (t, r) une transformation rigide formée d'une translation t et d'une rotation r . Le critère de recalage le plus souvent utilisé est celui des moindres carrés, et on cherche alors la transformation qui minimise :

$$E = \sum_{i=1}^n \|x_i - (r(y_i) + t)\|^2. \quad (6.29)$$

On peut montrer facilement que la translation optimale est celle qui fait coïncider les centres de gravité de X et Y , et on supposera donc dans la suite que X et Y ont même centre de gravité ($\sum x_i = \sum y_i$) et que la translation est nulle.

Soit q un des deux quaternions représentant la rotation r . L'expression de E devient alors :

$$E = \sum_{i=1}^n |x_i - q \times y_i \times \bar{q}|^2. \quad (6.30)$$

En utilisant les propriétés des quaternions de norme 1, on a :

$$\begin{aligned} E &= \sum_{i=1}^n |x_i - q \times y_i \times \bar{q}|^2 |q|^2 \\ &= \sum_{i=1}^n |x_i \times q - q \times y_i \times \bar{q} \times q|^2 \\ &= \sum_{i=1}^n |x_i \times q - q \times y_i|^2. \end{aligned}$$

D'après la définition du produit de quaternions, l'expression $x_i \times q - q \times y_i$ est une fonction linéaire des quatre composantes de q , et on peut donc écrire E sous la forme :

$$E = \sum_{i=1}^n q^t A_i^t A_i q$$

où les A_i sont des matrices 4×4 ne dépendant que de x_i et y_i . La rotation optimale s'obtient en calculant les valeurs propres de la matrice A de taille 4×4 suivante :

$$A = \sum_{i=1}^n A_i^t A_i.$$

La solution du problème de minimisation aux moindres carrés est alors le quaternion qui est le vecteur propre de norme 1 correspondant à la plus petite valeur propre de A .

6.3 Géométrie différentielle

Dans cette partie, nous rappelons quelques éléments de géométrie différentielle, utiles en particulier pour le traitement d'images et de données tridimensionnelles. L'ouvrage [Carmo, 1976] est une bonne référence dans ce domaine.

6.3.1 Courbes, repère de Frenet

Soit une courbe de l'espace paramétrée par u . Un point P de la courbe est repéré par rapport à l'origine de l'espace par le vecteur $r(u) = \vec{OP}$ et par rapport à un point de référence sur la courbe par l'abscisse curviligne s représentant la longueur de la courbe entre le point de référence et P :

$$s(u) = \int_{u_0}^u \|r'(u)\| du \quad (6.31)$$

où u_0 est le paramètre du point de référence sur la courbe.

La tangente à la courbe en P est définie par le vecteur :

$$T = \frac{dr}{ds}. \quad (6.32)$$

D'après la définition de s , on a :

$$\frac{ds}{du} = \|r'(u)\|$$

et donc

$$\frac{dr}{ds} = \frac{dr}{du} \frac{du}{ds} = \frac{r'(u)}{\|r'(u)\|}$$

dont on déduit que le vecteur tangent T est unitaire.

La dérivée de la tangente par rapport à l'abscisse curviligne définit un vecteur unitaire N appelé normale à la courbe en P et un réel k appelé courbure de la courbe au point P :

$$\frac{dT}{ds} = kN. \quad (6.33)$$

Le vecteur N est effectivement normal à T , puisqu'en dérivant l'expression $T \cdot T = 1$ (T est unitaire) on obtient $T \cdot \frac{dT}{ds} = 0$.

Le cercle osculateur, tangent à la courbe en P , a pour rayon $\frac{1}{k}$. Le plan défini par le point P et les vecteurs T et N est appelé plan osculateur à la courbe en P et contient le cercle osculateur.

La binormale est définie par :

$$B = T \wedge N \quad (6.34)$$

et est donc un vecteur unitaire orthogonal à T et à N .

Le triplet (T, N, B) définit un repère orthonormé, local en P , appelé repère de Frenet.

Les formules de Frenet-Serret donnent les dérivées des vecteurs introduits ci-dessus par rapport à l'abscisse curviligne dans le repère de Frenet :

$$\frac{dr}{ds} = T \quad (6.35)$$

$$\frac{dT}{ds} = kN \quad (6.36)$$

$$\frac{dN}{ds} = \tau B - kT \quad (6.37)$$

$$\frac{dB}{ds} = -\tau N \quad (6.38)$$

où τ est appelé la torsion de la courbe en P . Ces équations s'obtiennent facilement en dérivant les expressions suivantes :

$$T \cdot T = 1 \quad N \cdot N = 1 \quad B \cdot B = 1$$

et

$$T \cdot N = 0 \quad N \cdot B = 0 \quad B \cdot T = 0.$$

Ces expressions peuvent s'écrire sous forme condensée :

$$\frac{d}{ds} \begin{pmatrix} T \\ N \\ B \end{pmatrix} = \begin{pmatrix} -\tau \\ 0 \\ -k \end{pmatrix} \wedge \begin{pmatrix} T \\ N \\ B \end{pmatrix}. \quad (6.39)$$

En pratique, il est souvent plus facile de manipuler les expressions qui dépendent d'un paramètre u (qui peut être quelconque) plutôt que des expressions qui dépendent de l'abscisse curviligne (qui peut être difficile à calculer). On a alors :

$$\dot{r} = \frac{dr}{du} \quad (6.40)$$

$$T = \frac{\dot{r}}{\dot{s}} \quad (6.41)$$

$$\dot{s} = \frac{ds}{du} = |\dot{r}| \quad (6.42)$$

$$kB = \frac{\dot{r} \wedge \ddot{r}}{\dot{s}^3} \quad (6.43)$$

$$N = B \wedge T \quad (6.44)$$

$$\tau = \frac{\dot{r}(\ddot{r} \wedge \ddot{\ddot{r}})}{\dot{s}^6 k^2} \quad (6.45)$$

Dans ces équations, les points au-dessus des symboles désignent les dérivées par rapport à u .

6.3.2 Courbes sur une surface, formes fondamentales

Considérons maintenant une surface S , paramétrée par (u, v) . Un point P sur la surface est repéré par le vecteur $r(u, v) = \vec{OP}$. Le plan tangent à la surface en P est le plan $T_P(S)$ défini par les deux vecteurs suivants :

$$\frac{\partial r}{\partial u} = r_u, \quad (6.46)$$

$$\frac{\partial r}{\partial v} = r_v. \quad (6.47)$$

La normale à la surface en P est le vecteur unitaire orthogonal à $T_P(S)$:

$$n = \frac{r_u \wedge r_v}{|r_u \wedge r_v|}. \quad (6.48)$$

On appelle souvent sphère de Gauss la sphère unité de \mathbf{R}^3 sur laquelle un point représente un vecteur unitaire de \mathbf{R}^3 . L'application qui à tout point P de la surface associe le point correspondant à n sur la sphère de Gauss est appelée application de Gauss. Nous verrons par la suite l'intérêt de cette application et de sa différentielle.

Soit une courbe sur la surface, paramétrée par t , dont les points s'écrivent $U = (u(t), v(t))^t$. La dérivée de r par rapport à t , vecteur tangent à la courbe, vaut :

$$\dot{r} = \frac{dr}{dt} = \frac{\partial r}{\partial u} \frac{du}{dt} + \frac{\partial r}{\partial v} \frac{dv}{dt} = A\dot{U} \quad (6.49)$$

où \dot{U} est le vecteur des dérivées de U par rapport à t et A est la matrice 2×2 dont les colonnes sont les vecteurs représentant les dérivées partielles de r par rapport à u et v respectivement :

$$A = \left(\frac{\partial r}{\partial u} \mid \frac{\partial r}{\partial v} \right).$$

La longueur de ce vecteur tangent est :

$$\dot{s}^2 = |\dot{r}|^2 = \dot{U}^t A^t A \dot{U} = \dot{U}^t G \dot{U} \quad (6.50)$$

où s désigne toujours l'abscisse curviligne sur la courbe, et G est la matrice

$$G = A^t A = \begin{pmatrix} \frac{\partial r}{\partial u} \cdot \frac{\partial r}{\partial u} & \frac{\partial r}{\partial u} \cdot \frac{\partial r}{\partial v} \\ \frac{\partial r}{\partial v} \cdot \frac{\partial r}{\partial u} & \frac{\partial r}{\partial v} \cdot \frac{\partial r}{\partial v} \end{pmatrix} \quad (6.51)$$

appelée première forme fondamentale. Elle mesure le déplacement dr sur la surface au point de paramètres (u, v) pour des variations (du, dv) de ces paramètres.

Le vecteur tangent unitaire à la courbe peut alors s'écrire en fonction des matrices A et G sous la forme :

$$T = \frac{\dot{r}}{|\dot{r}|} = \frac{A\dot{U}}{(\dot{U}^t G \dot{U})^{1/2}}$$

et l'abscisse curviligne :

$$s = \int_{t_0}^t \frac{ds}{dt} dt = \int_{t_0}^t (\dot{U}^t G \dot{U})^{1/2} dt.$$

Examinons maintenant les dérivées secondes de r . Soient N la normale à la courbe (définie à partir de la dérivée de T comme nous l'avons vu plus haut), et n la normale à la surface. La dérivée de l'expression $\dot{r} = \dot{s}T$ (sur la courbe) peut s'exprimer par les deux formes équivalentes suivantes :

$$\ddot{r} = \dot{s}T + \dot{s}^2 k N \quad (6.52)$$

$$= \frac{\partial^2 r}{\partial u^2} \dot{u}^2 + 2 \frac{\partial^2 r}{\partial u \partial v} \dot{u} \dot{v} + \frac{\partial^2 r}{\partial v^2} \dot{v}^2 + \frac{\partial r}{\partial u} \ddot{u} + \frac{\partial r}{\partial v} \ddot{v}. \quad (6.53)$$

Le long de la normale à la surface, on a :

$$\ddot{r} \cdot n = \dot{s}^2 k N \cdot n = \dot{U}^t D \dot{U} \quad (6.54)$$

où k est la courbure de la courbe, et D est la matrice

$$D = \begin{pmatrix} n \cdot \frac{\partial^2 r}{\partial u^2} & n \cdot \frac{\partial^2 r}{\partial u \partial v} \\ n \cdot \frac{\partial^2 r}{\partial v \partial u} & n \cdot \frac{\partial^2 r}{\partial v^2} \end{pmatrix} \quad (6.55)$$

appelée deuxième forme fondamentale.

6.3.3 Courbures sur des surfaces

Introduisons maintenant les notions de courbures de surface. Contrairement au cas d'une courbe où l'on définit une courbure, plusieurs types de courbure peuvent être définis pour une surface.

Le vecteur tangent $\dot{r} = A\dot{U}$ et la normale à la surface n définissent un plan, dont l'intersection avec la surface est une courbe. On appelle courbure normale k_n de la surface dans la direction $A\dot{U}$ la courbure de cette courbe. C'est la longueur (signée) de la projection de kN sur n , et on a (figure 6.5) :

$$k_n = kN \cdot n \quad (6.56)$$

qui est donc positive si la courbe est tournée du côté positif de n et négative sinon.

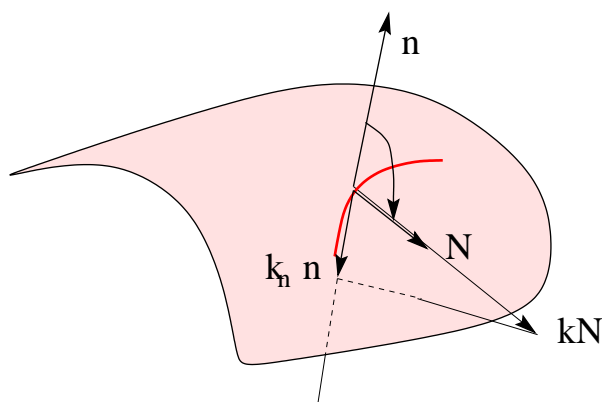


FIG. 6.5 – Courbe sur une surface, normale N à la courbe et normale n à la surface, courbure k de la courbe et courbure normale k_n de la surface dans la direction de la tangente à la courbe.

La courbure normale peut également être exprimée en fonction des formes fondamentales :

$$k_n = kN \cdot n = \frac{\ddot{r} \cdot n}{s^2} = \frac{\dot{U}^t D \dot{U}}{\dot{U}^t G \dot{U}}. \quad (6.57)$$

Les courbures principales k_1 et k_2 sont définies comme les valeurs minimale et maximale de k_n , et les directions associées sont appelées directions principales. On appelle point ombilic un point pour lequel $k_1 = k_2$ (c'est-à-dire que toutes les courbes sur la surface passant par ce point ont la même courbure).

Deux courbures particulières sont souvent utilisées, la courbure gaussienne définie par :

$$K = k_1 k_2 = \frac{|D|}{|G|}, \quad (6.58)$$

et la courbure moyenne :

$$H = \frac{k_1 + k_2}{2}. \quad (6.59)$$

La courbure gaussienne est le déterminant de la différentielle de l'application de Gauss, et la courbure moyenne est proportionnelle à sa trace. Ces deux valeurs sont utilisées pour caractériser localement des surfaces, en reconnaissance d'objets 3D, etc.

Ces utilisations sont possibles car les courbures gaussienne et moyenne sont des propriétés intrinsèques de la surface et sont invariantes par changement de paramétrisation. En effet, soit une autre paramétrisation (u', v') de

la surface, définie par la matrice de changement de paramétrisation P par rapport à (u, v) :

$$P = \begin{pmatrix} \frac{\partial u}{\partial u'} & \frac{\partial u}{\partial v'} \\ \frac{\partial v}{\partial u'} & \frac{\partial v}{\partial v'} \end{pmatrix}.$$

Les matrices A , G et D deviennent alors :

$$A' = AP \quad G' = P^t GP \quad D' = P^t DP$$

et donc n , les directions et courbures principales, les courbures gaussienne et moyenne sont invariantes par changement de paramétrisation.

6.3.4 Une notation fréquente

Une autre notation, fréquemment rencontrée, consiste à écrire la première forme fondamentale sous la forme :

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix} \quad (6.60)$$

et la deuxième forme fondamentale sous la forme :

$$\begin{pmatrix} L & M \\ M & N \end{pmatrix}. \quad (6.61)$$

Les courbures principales sont alors les valeurs propres de :

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}^{-1} \begin{pmatrix} L & M \\ M & N \end{pmatrix}$$

et les courbures gaussiennes et moyennes sont données par :

$$K = \frac{LN - M^2}{EG - F^2},$$

$$H = \frac{LG + EN - 2FM}{2(EG - F^2)}.$$

6.3.5 Interprétations géométriques

Plusieurs mesures géométriques peuvent s'exprimer de manière simple en fonction de r et de ses dérivées. Ainsi, la longueur d'une courbe $r(u)$ entre les points de paramètres u_0 et u_1 est donnée par :

$$L = \int_{u_0}^{u_1} \dot{s} du = \int_{u_0}^{u_1} |\dot{r}(u)| du. \quad (6.62)$$

Dans le cas d'une courbe plane, on peut calculer l'aire définie par la courbe entre deux points de paramètres u_0 et u_1 et les segments joignant chacun de ces points à l'origine O (figure 6.6) par :

$$A = \frac{1}{2} \int_{u_0}^{u_1} r(u) \wedge \dot{r}(u) du. \quad (6.63)$$

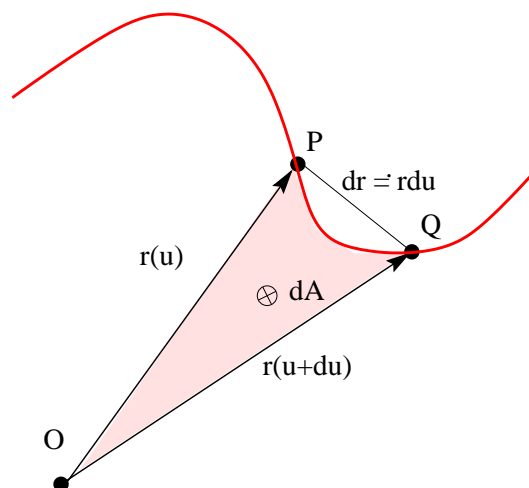


FIG. 6.6 – Le repère local et la courbure le long d'une courbe

Plaçons nous maintenant dans le cas d'une surface, et considérons le morceau de surface défini par le point A de paramètres (u, v) , et les points $B(u + du, v)$, $C(u, v + dv)$ et $D(u + du, v + dv)$. La surface de ce morceau élémentaire est :

$$\left| \frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} \right| du dv = |G|^{1/2} du dv, \quad (6.64)$$

et donc la surface totale de S (obtenue pour des paramètres variant dans un ensemble R) est :

$$\int \int_R |G|^{1/2} du dv. \quad (6.65)$$

Le volume de la portion d'espace limitée par le morceau de surface et les segments joignant chacun des quatre points à l'origine O vaut :

$$\frac{1}{3} r \cdot \left(\frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} \right) du dv \quad (6.66)$$

et donc le volume total défini par O et S vaut :

$$\frac{1}{3} \int \int_R r \cdot \left(\frac{\partial r}{\partial u} \wedge \frac{\partial r}{\partial v} \right) du dv. \quad (6.67)$$

L'application de Gauss permet d'interpréter la courbure gaussienne. Soit un élément dS de la surface. Si l'on calcule les normales à la surface en tout point de dS , l'application de Gauss transporte dS sur un élément de surface dA de la sphère unité. La courbure gaussienne est alors le rapport de ces deux surfaces (figure 6.7) :

$$K = \frac{dA}{dS}. \quad (6.68)$$

Le signe des courbures gaussienne et moyenne permet de classifier les surfaces (localement) [Besl et Jain, 1986]. On distingue ainsi :

- un surface elliptique, pour laquelle $K > 0$, le signe de H déterminant le sens dans lequel est tournée la surface par rapport à l'origine (figure 6.8) ;
- une surface hyperbolique, pour laquelle $K < 0$ (en forme de selle) (figure 6.9) ;

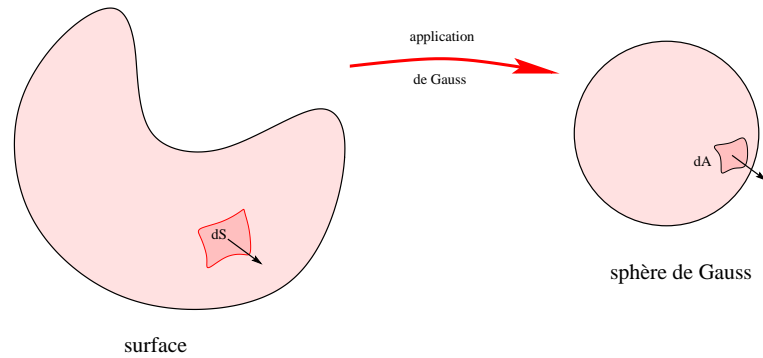


FIG. 6.7 – Application de Gauss : l’aire de surface est projetée sur la sphère unité. La courbure gaussienne de la surface est égale à $K = \frac{dA}{dS}$.

- une surface parabolique pour laquelle $K = 0$ (c’est-à-dire qu’une des deux courbures principales est nulle), et $H \neq 0$, le signe de H déterminant le sens dans lequel est tournée la surface (figure 6.10) ;
- une surface plane pour laquelle $K = H = 0$ (c’est-à-dire que les deux courbures principales sont nulles) (figure 6.11).

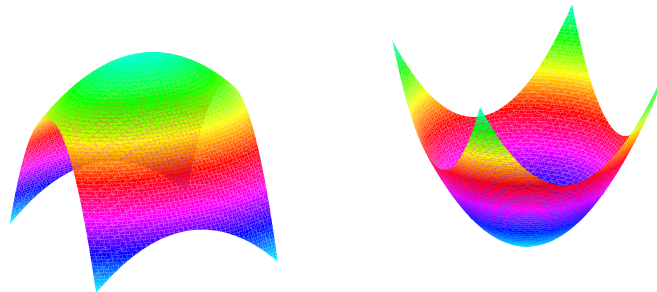


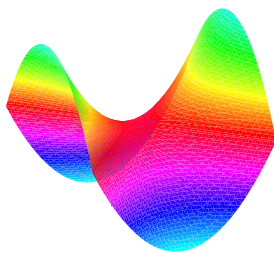
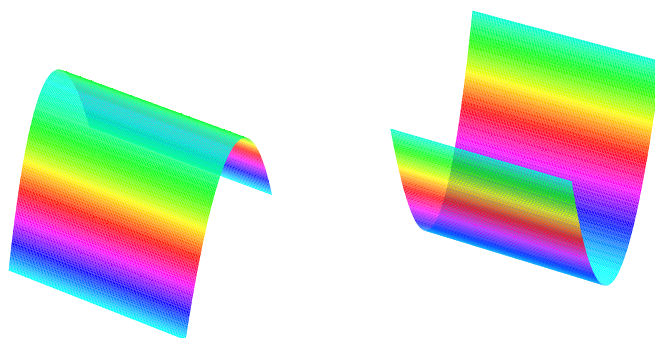
FIG. 6.8 – Surfaces elliptiques : $K > 0$; à gauche $H < 0$, à droite $H > 0$.

Considérons un cycle de points sur un morceau de surface elliptique ($K > 0$), auquel on donne un sens de parcours. Le cycle de points correspondants sur la sphère de Gauss, obtenu par l’application de Gauss, a le même sens. L’application de Gauss préserve donc l’orientation dans le cas d’une surface elliptique. Au contraire, dans le cas d’une surface hyperbolique ($K < 0$), le sens est inversé sur la sphère de Gauss.

6.3.6 Principales méthodes de calcul des courbures

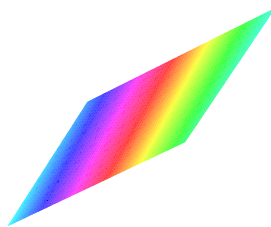
En général en traitement d’images et vision on ne connaît qu’un échantillonnage de la surface. Les principales méthodes de calcul des courbures se divisent en deux classes : les méthodes analytiques et les méthodes discrètes.

Les méthodes analytiques nécessitent de transposer les points échantillonnés sur la surface en une forme analytique, au moins localement. Le principe le plus courant consiste à approcher localement la surface par une surface paramétrée, une quadrique par exemple, dont on calcule les paramètres pour minimiser une fonction d’ajustement (moindres carrés typiquement). Le calcul des courbures peut alors se faire analytiquement, en appliquant directement les équations vues plus haut.

FIG. 6.9 – Surface hyperbolique : $K < 0$.FIG. 6.10 – Surfaces paraboliques : $K = 0$ et $H \neq 0$; à gauche $H < 0$, à droite $H > 0$.

Donnons maintenant quelques exemples de méthodes discrètes, qui utilisent directement les points échantillonnés sur la surface :

- méthode par dérivation numérique : les dérivées sont calculées par différences finies entre points voisins afin

FIG. 6.11 – Surface plane : $K = H = 0$.

de définir des équations approchées des équations analytiques ;

- méthode de dérivation de la normale : soient P et Q deux points proches sur la surface, et n_P et n_Q les normales à la surface en ces points, alors la courbure gaussienne est donnée par :

$$K = \sigma \frac{|n_P - n_Q|}{|PQ|} \quad (6.69)$$

avec $\sigma = 1$ si $|PQ| < |(P + n_P) - (Q + n_Q)|$, et $\sigma = -1$ sinon ;

- calcul par déficit angulaire : pour calculer la courbure gaussienne en un point P , on triangule la surface autour de P , en construisant des triangles formés de P et de deux voisins P_i et P_{i+1} de P , dont on calcule l'angle θ_i en P et la surface S_i ; on a alors :

$$K = \frac{2\pi - \sum_i \theta_i}{\frac{1}{3} \sum_i S_i}, \quad (6.70)$$

cette équation correspondant à l'interprétation de K en termes d'application de Gauss.

6.3.7 Géométrie différentielle dans des volumes numériques

Dans le cas d'images volumiques $I(x, y, z)$, on peut calculer directement les deux formes fondamentales (et donc les courbures principales, gaussienne et moyenne) pour toute isosurface $I(x, y, z) = I_0$ sans exprimer cette surface explicitement. Pour cela, la surface est paramétrée par $u = x$ et $v = y$. La coordonnée z est alors une fonction implicite de u et v , notée $\Phi(u, v)$ et on a :

$$r(u, v) = (u, v, \Phi(u, v))^t. \quad (6.71)$$

Le théorème des fonctions implicites permet de relier les dérivées de Φ par rapport à u et v aux dérivées de l'image I par rapport à x, y et z :

$$\frac{\partial \Phi}{\partial u} = -\frac{I'_x}{I'_z} \quad (6.72)$$

$$\frac{\partial \Phi}{\partial v} = -\frac{I'_y}{I'_z} \quad (6.73)$$

où I'_x, I'_y, I'_z désignent les dérivées partielles de I par rapport à x, y et z respectivement. On a alors :

$$\frac{\partial r}{\partial u} = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial \Phi}{\partial u} \end{pmatrix} \quad \frac{\partial r}{\partial v} = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial \Phi}{\partial v} \end{pmatrix} \quad (6.74)$$

et on peut alors calculer les coefficients des formes fondamentales :

$$E = \left\| \frac{\partial r}{\partial u} \right\|^2 = 1 + \left(\frac{\partial \Phi}{\partial u} \right)^2 = \frac{I_x'^2 + I_z'^2}{I_z'^2} \quad (6.75)$$

$$F = \frac{I'_x I'_y}{I_z'^2} \quad (6.76)$$

$$G = \frac{I_y'^2 + I_z'^2}{I_z'^2} \quad (6.77)$$

$$L = \frac{2I'_x I'_z I''_{xz} - I_x'^2 I''_{zz} - I_z'^2 I''_{xx}}{H I_z'^3} \quad (6.78)$$

$$H^2 = \frac{I_x'^2 + I_y'^2 + I_z'^2}{I_z'^2} \quad (6.79)$$

etc.

En pratique, les dérivées de l'image se calculent de manière discrète, par exemple :

$$I'_x = \frac{\partial I}{\partial x}(x, y, z) = \frac{I(x + 1, y, z) - I(x - 1, y, z)}{2},$$

ou par convolution avec les dérivées d'un filtre gaussien ou de Deriche (on calcule ainsi les dérivées de l'image filtrée), ce qui permet d'avoir des dérivées moins sensibles au bruit de l'image.

Chapitre 7

Formation de l'image et couleur

Chapitre rédigé par Henri MAÎTRE et Anne-Catherine CARRILERO

7.1 Formation des images

7.1.1 Imagerie cohérente et incohérente

Deux grandes classes d'images doivent être distinguées qui ont des propriétés très différentes (voir chapitre 1) : les images obtenues en imagerie cohérente et celles obtenues en imagerie incohérente. Pour compliquer les choses, des images sont faites en imagerie "partiellement cohérente", comme celles obtenues en microscopie par exemple ; elles présentent des propriétés intermédiaires entre celles des deux types principaux.

Imagerie cohérente

Une source est cohérente si elle présente deux types de propriétés [Pérez, 1991] :

1. elle est **temporellement cohérente**, c'est-à-dire monochromatique (une seule longueur d'onde), et les photons qu'elle émet présentent une relation de phase qui se maintient dans le temps au delà de la durée de vie des photons ;
2. elle possède une **cohérence spatiale**, c'est-à-dire que les divers points de la source émettent des rayonnements en phase.

Les sources naturelles dans le domaine visible ne sont pas cohérentes. On peut les rendre cohérentes en les filtrant par des filtres monochromatiques très étroits et en les diaphragmant.

Les principales sources cohérentes sont :

- dans le domaine visible : les lasers, les lampes spectrales (Hg, Cd, etc.). Dans la vie quotidienne certaines lampes utilisées en éclairage urbain ou des tubes en éclairage domestique ou professionnel peuvent se comporter, dans des conditions très particulières, comme des sources cohérentes ;
- dans les autres domaines des ondes électromagnétiques : le radar est le type-même de source cohérente ; il émet un rayonnement dans les gammes centimétriques (voir [Maître, 2001]), certains rayonnements utilisés en médecine ou en contrôle non destructif : PET (Imagerie par émission de Positons), rayons γ , etc.
- dans le domaine des ondes acoustiques : en imagerie ultrasonore, en prospection sismique, etc. Les sources sont alors des transducteurs piézo-électriques ou des générateurs électro-mécaniques.

Lors de la formation d'une image en éclairage cohérent, une source primaire est réfléchiée ou diffractée par la scène que l'on veut imager, chaque point de cette scène est une source secondaire de rayonnement caractérisée

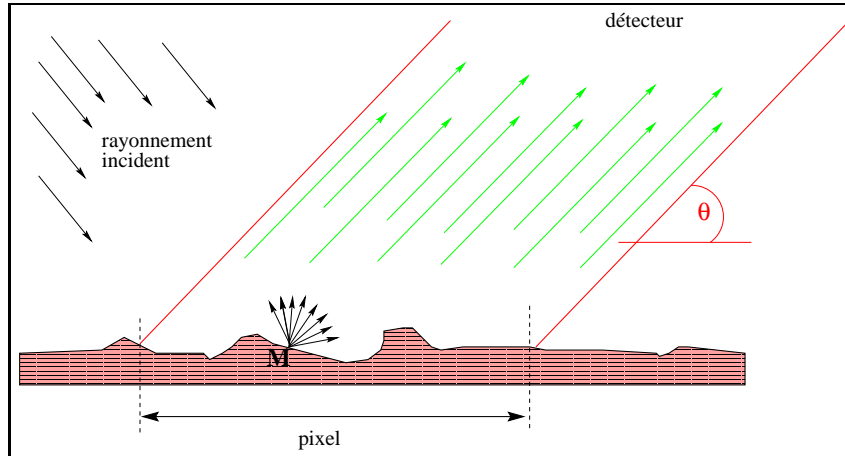


FIG. 7.1 – Formation d’une image : le rayonnement renvoyé par chaque point de la scène est collecté (par exemple par une sélection angulaire en fonction de θ) et sommé sur le capteur. En éclairage cohérent on somme des amplitudes complexes, en éclairage incohérent, on somme des densités spectrales d’énergie.

par sa phase et son amplitude dans chaque direction de l’espace (principe d’Huygens). Le capteur sélectionne les seules contributions provenant du pixel dont il fera l’image (sur des critères angulaires ou sur des critères de temps de propagation) et construira un signal $R(k, l)$ égal à (nous considérons ici une sélection angulaire) :

$$R(k, l) = \left[\left\| \int_{x-\delta x}^{x+\delta x} \int_{y-\delta y}^{y+\delta y} \rho(x, y) a(x, y) \exp(-2\pi\phi(x, y)) dx dy \right\| \right]^2 \quad (7.1)$$

Le terme $a(x, y)$ exprime les effets des absorptions et des atténuations dues à la réflexion sur le pixel ; il s’exprime en fonction de l’amplitude de l’onde incidente et des propriétés du matériau. Le terme de phase comporte d’une part les retards dus à la propagation (de la source au point de la scène, puis de la scène au capteur), ainsi que les déphasages introduits par l’interaction avec la scène (ce que l’on appelle fréquemment la phase propre du pixel). Le terme $\rho(x, y)$ traduit les atténuations dues à la propagation entre le point du sol et le capteur. La sommation se fait sur les amplitudes complexes. Le capteur mesure une énergie, c’est-à-dire le carré du module de l’intégrale ainsi obtenue.

En imagerie cohérente, tous les effets de phase sont très importants car ils régissent pour une grande part le bilan de l’équation 7.1. Ainsi la contribution d’un point du pixel peut-elle être positive ou négative selon que sa phase est en accord avec celle de ses voisins ou non. La variation très rapide de l’amplitude crée le phénomène de **granularité**¹. L’imagerie cohérente est considérée comme très bruitée par rapport à l’imagerie incohérente (voir [Nicolas, 1999, Maître, 2001]), mais son traitement bénéficie de la connaissance de très fortes propriétés statistiques qui n’existent pas en optique incohérente.

Imagerie incohérente

Le processus de formation des images en optique incohérente est finalement beaucoup plus simple, puisque le capteur intègre les énergies issues de chaque point source du pixel, et cela pour chacune des longueurs d’onde

¹Des termes différents sont employés dans des domaines différents : *granularité* est le terme employé en optique, *tavelure* en astronomie, *chatoiement* en radar. Le terme anglo-saxon consacré est *speckle*.

auquel il est sensible (avec une sensibilité $s(\lambda)$) :

$$R(k, l) = \int_{\lambda_0}^{\lambda_{max}} s(\lambda) \int_{x-\delta x}^{x+\delta x} \int_{y-\delta y}^{y+\delta y} A(x, y, \lambda) dx dy d\lambda \quad (7.2)$$

où $A(x, y, \lambda)$ représente la densité spectrale d'énergie réémise par le point M dans la direction du capteur, et $s(\lambda)$ la sensibilité spectrale du capteur.

En comparant les équations 7.1 et 7.2 on comprend que les effets géométriques interviennent beaucoup moins dans ce second cas, puisqu'ils n'interviennent que par la modulation du terme A (par exemple en traduisant la nature plus ou moins directionnelle des réflexions), cette modulation étant, toutes choses égales par ailleurs, aussi présente dans le terme a de l'équation 7.1.

7.1.2 Interaction matière rayonnement

Cette partie a été rédigée en partie à partir du livre de Patrick Callet [Callet, 1998].

C'est l'un des problèmes les plus complexes de la formation de l'image. Les processus mis en jeu sont très nombreux et opèrent fréquemment de façon simultanée rendant leurs contributions individuelles très difficiles à évaluer. Par ailleurs ces phénomènes peuvent s'expliquer par des modélisations à divers niveaux de la physique, ces niveaux n'étant pas compatibles. Ainsi des explications peuvent provenir d'une simple interprétation en optique géométrique, en optique ondulatoire, en électromagnétique classique, en électromagnétique quantique, en thermodynamique ou en chimie [Callet, 1998]. Dans de nombreux cas il est très difficile de mettre en évidence les phénomènes dominants et des interprétations à des niveaux différents peuvent conduire à des résultats très différents. Pour compliquer le tableau des connaissances actuelles, et devant l'ampleur et la complexité des bilans et des modèles qui cherchent à s'appuyer sur la physique, des modélisations complètement artificielles et purement mathématiques ont été proposées pour simuler les **apparences** dans des domaines comme la synthèse des images. Ces modèles ont été hybridés avec des modèles physiques conduisant à un tableau extrêmement confus.

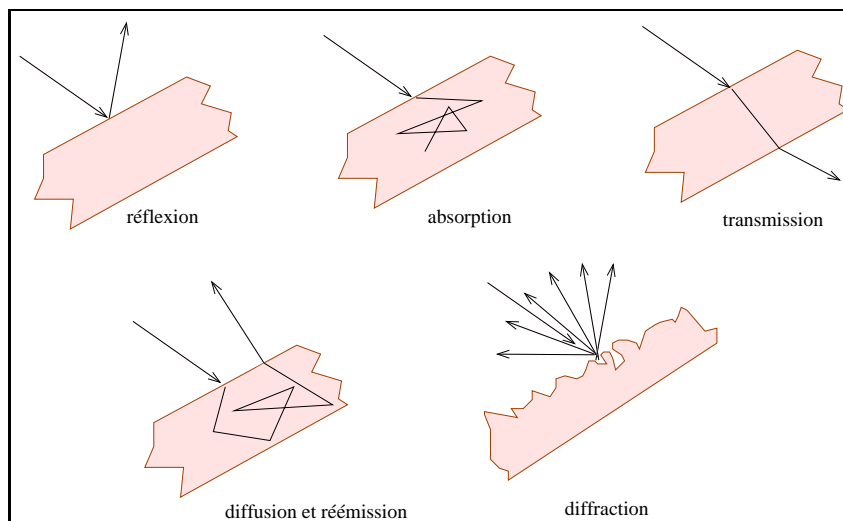


FIG. 7.2 – Les principaux mécanismes de propagation des rayons lumineux

Réflexion et réfraction

Elles résultent d'une modélisation en optique (géométrique : sans interférences, ou ondulatoire : avec interférences et diffraction). Les relations de Descartes ($n_1 \sin i_1 = n_2 \sin i_2$) régissent les réflexions spéculaires (miroirs ou dioptrés parfaits). Des relations empiriques comme la formule de Phong [Phong, 1975] s'en inspirent pour proposer des modèles de miroirs moins parfaits pour lesquels la décroissance de l'intensité autour de l'angle spéculaire est de la forme $I(\theta) \propto \cos^\nu(i_2 - \theta)$, ν étant une constante adaptée à chaque matériau.

Des approches photométriques suivant ces directions s'appuient sur une fonction dite de **réflexion bidirectionnelle** (FRB ou BDRF) qui relie l'énergie reçue dans une direction repérée par rapport à la normale à la surface en fonction de l'énergie émise par la source (elle même repérée par rapport à la normale). Cette FRB est normalement donc fonction de 4 angles [Nicodemus et al., 1977]. Des propriétés d'isotropie peuvent parfois être mises en avant pour ramener ces variables de 2 à 4.

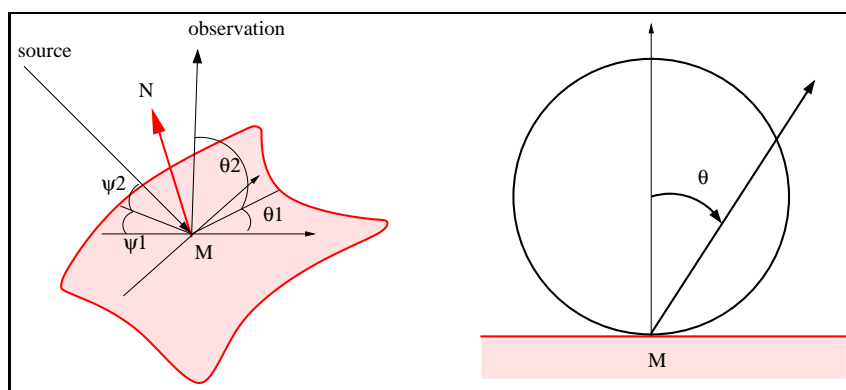


FIG. 7.3 – La fonction de réflectance bidirectionnelle exprime l'intensité réémise par la surface en fonction de l'intensité de la source et des 4 angles repérant source et récepteur par rapport à la normale. A droite, dans le cas de la diffusion lambertienne, l'intensité réémise est proportionnelle au cosinus de l'angle θ

Outre la réflexion spéculaire, la forme la plus usitée de FRB est la loi de Lambert qui correspond à une surface isotrope et parfaitement diffusante. Pour une surface vérifiant la loi de Lambert, la luminance est indépendante de la direction d'observation et ne dépend que de la direction ψ de la source. Pour une surface lambertienne l'énergie réémise par unité d'angle solide est proportionnelle au cosinus de l'angle fait par cette direction avec la normale.

La représentation vectorielle des champs électro-magnétiques conduit des équations de Descartes aux équations de Fresnel. Un modèle de ce type est par exemple utilisé dans les modèles de milieu rugueux de Torrance-Sparrow [Torrance et Sparrow, 1967] qui considère une distribution gaussienne de micro-facettes aléatoires. C'est un modèle de ce type qui a été exploité en synthèse d'images par Blinn [Foley et al., 1995], souvent en concurrence au modèle de Phong dont il ne diffère que peu.

Les lois de Descartes-Fresnel permettent en particulier de relier les champs absorbés et dispersés par le milieu (relations entre parties réelles et imaginaires de ϵ (relations de Kramers-Kronig)). En généralisant les relations de Fresnel aux milieux métalliques (d'indice ϵ complexe), on peut justifier un modèle très classiquement utilisé en synthèse d'images, celui de Cook et Torrance [Cook et Torrance, 1982] qui donne au matériau un indice complexe effectif du type :

$$\hat{n} = n(1 + i)$$

On construit également des modèles de milieux rugueux (rugosité de Beckmann ou de Beckmann et Spizzichino, formules de Goodman pour l'imagerie cohérente et la granularité totalement développée) [Beckmann et Spizzichino, 1963].

On peut également traiter les milieux périodiques ainsi que les réseaux (équations de Kirchoff, diffraction de Bragg).

Dans le cas de propagation dans des matériaux volumiques épais on retrouve là aussi la théorie des réseaux épais de Bragg à partir d'une modélisation en optique ondulatoire.

Dispersion dans les milieux homogènes

Cette modélisation considère la propagation des ondes électromagnétiques dans la matière. On cherche à exprimer les modifications du champ électro-magnétique incident sous l'effet des perturbations induites dans le milieu par ce champ. Les dipôles étant les structures sensibles au champ EM, les plus simples à modéliser, elle s'appuie donc sur la résolution des équations de Maxwell dans un milieu caractérisé par la présence de dipôles [Born et Wolf, 1975]. On établit ainsi des relations entre la fonction diélectrique complexe ϵ et la polarisabilité des dipôles. Une modélisation en est faite par la méthode de Lorentz qui conduit aux équations de Clausius-Mossotti (liant ϵ à la polarisabilité α des dipôles élémentaires). Pour la formation des images, ceci a conduit à la définition d'un **milieu effectif**, c'est-à-dire d'un milieu virtuel permettant de rendre compte des mêmes effets coloriques, défini uniquement par son indice de réfraction effectif. On explique ainsi l'apparence des vitraux et le rôle des ions métalliques comme Co^+ et Fe^+ (équations de Maxwell Garnett) [Berthier, 1993]. On peut également appliquer la théorie du champ moyen (méthode de Bruggeman) qui permet d'éviter de connaître les fonctions diélectriques des particules participant au milieu.

Diffusion et diffraction par des particules

Une autre famille de problèmes provient de la diffusion et de la diffraction de la lumière par des particules de dimension comparable ou inférieure à la longueur d'onde : poussières, molécules ou atomes.

Dans l'approche de Rayleigh, on traite des populations de particules incohérentes dans un milieu très dilué (cela entraîne qu'il n'y aura qu'une interaction au plus pour un rayon donné) de taille très petite devant λ . On montre alors que la puissance diffusée par unité de volume est proportionnelle à λ^{-4} , entraînant d'importants effets colorés en fonction de l'angle d'observation. L'intensité de l'onde traversant un espace de longueur l prend alors la forme :

$$I = I_0 \exp(-ld)$$

où d est la densité optique, proportionnelle à la concentration des particules, et à λ^{-4} .

La théorie de Rayleigh a été complétée par Mie qui a considéré non seulement les dipôles, mais aussi les multipôles. Sa théorie conduit à des résultats asymptotiquement équivalents à ceux de Rayleigh quand la taille des particules tend vers 0.

La diffusion par les particules explique de très nombreux phénomènes physiques : couleur du ciel ou des mers, fumées, etc.

Ces théories ne prenant en compte qu'une seule interaction sur un faisceau, ont été complétées par des approches plus complexes de diffusion multiple, mais celles-ci n'aboutissent pas à des expressions explicites et on leur préfère des démarches plus empiriques comme par exemple les équations du transfert radiatif [Kerker, 1969].

Parallèlement à ces travaux, des approches phénoménologiques ont été faites du phénomène de diffusion pour rendre compte des résultats sans en expliquer les causes exactes. C'est souvent des approches qui sont retenues pour résoudre des problèmes concrets. Une telle approche est par exemple celle de Melamed qui essaie de rendre compte de l'effet de poudres pulvérulentes (en ce cas on ramène le milieu à une seule couche faiblement absorbante de particules de taille grande devant λ) [Melamed, 1963]. Il rend compte ainsi de la neige, de la poussière, du sable, du café, etc. Le modèle de Kubelka-Munk² est particulièrement adapté à présenter les effets des matériaux

²On se reportera à l'article [Duntley, 1942] plutôt qu'à l'original de Kubelka-Munk en allemand.

synthétiques : pigments dans les peintures, plastiques, papier coloré, etc. Il découpe le milieu en couches successives et met en évidence le rôle important du rapport du coefficient d'absorption sur le coefficient de diffusion.

Approche chimique

Dans une telle démarche on s'intéresse essentiellement à la structure chimique des composés. Dans un composé chimique, il y a une étroite relation entre l'énergie de liaison et la position des bandes d'absorption des matériaux qui contrôlent l'aspect extérieur du matériau. Ceci est décrit par les équations de la mécanique quantique appliquées à la structure en question. Ainsi l'augmentation de l'intensité de la liaison de covalence favorise l'apparition de niveaux peu serrés et donc l'absorption se fera aux plus grandes longueurs d'ondes, plutôt vers le rouge que le bleu. Une bande d'absorption dans le violet donnera au matériau une apparence jaune ou rouge très vive.

Ces propriétés expliquent les couleurs vives de certains oxydes. D'autres apparences découlent de la position relative des anions et des cations au sein de la structure cristalline, de l'existence de transitions électroniques particulières, de la dimension de certaines structures, ou de la présence et de la concentration d'impuretés.

7.1.3 Les sources de lumière

1. Illuminant A : température $T = 2856\text{K}$,
2. Illuminant B : $T = 4800\text{K}$, il est proche d'une lumière solaire directe,
3. Illuminant C : $T = 6500\text{K}$, il est proche d'une lumière du jour moyenne,
4. Illuminants D50, D55, D65 et D75 : les températures respectives sont 5000K , 5500K , 6500K et 7500K , ils correspondent à quatre rayonnements diurnes.

7.2 Définition des différents espaces couleur

7.2.1 Trivariance et primaires

Maxwell a montré qu'une lumière blanche peut être reproduite "en apparence" par une combinaison linéaire de trois primaires. Grassman a étendu cette propriété à toute couleur et a donc établi la "trivariance" de l'espace coloré : **tout rayonnement peut être remplacé par une combinaison linéaire de 3 primaires qui aura la même apparence colorée pour un observateur.**

Dans la pratique, la plupart du temps, le choix des primaires se porte sur le rouge, le vert, et le bleu dans les synthèses additives (les seules que nous aborderons ici), et le jaune, le magenta et le cyan (leurs couleurs complémentaires) dans les synthèses soustractives. Dans la synthèse additive, on additionne les primaires pour obtenir la teinte cherchée ; dans la synthèse soustractive, partant d'une lumière considérée blanche, on soustrait à l'aide de filtres les primaires.

Deux rayonnements qui donnent une même impression colorée mais n'ont pas le même contenu spectral sont dits **métamères**. Le métamérisme forme une classe d'équivalence. Le métamérisme naturel exact est extrêmement rare.

Si l'on mélange les primaires rouge, vert et bleu, on obtient les couleurs secondaires qui sont également les primaires soustractives :

- rouge + vert = jaune,
- rouge + bleu = magenta (couleur mauve),
- vert + bleu = cyan (couleur bleu-turquoise),
- rouge + vert + bleu = blanc.

Les couleurs complémentaires (qui additionnées en bonne proportion donnent du blanc) sont donc :

- jaune et bleu,
- magenta et vert,
- cyan et rouge.

Les lois de la colorimétrie [Marion, 1997] sont complétées par :

- l'égalisation visuelle des couleurs demeure si l'on change également les quantités de chaque lumière du mélange. Cette égalisation ne disparaît qu'aux très faibles éclaircissements (perte de la sensation colorée) et aux forts éclaircissements (éblouissement).
- La luminance visuelle d'un mélange de couleurs est la somme des luminances visuelles des luminances de ses composantes.

Un espace colorimétrique est défini par un blanc de référence (c'est-à-dire une répartition $W(\lambda)$ donant une sensation visuelle de blanc) et 3 primaires P_1, P_2 et P_3 . Il existe une infinité d'espaces colorimétriques. On appelle composante trichromatique de la couleur C dans le système colorimétrique (P_1, P_2, P_3, W) , les 3 rapports :

$$T_i(C) = \frac{L_i(C)}{L_i(W)} \quad i = 1, 2, 3$$

où les $l_i(C)$ sont les luminances de la primaire nécessaires pour égaler visuellement la sensation de la lumière C (et idem pour W). La luminance, exprimée en candéla par m^2 , se calcule comme :

$$L = K \int_{400nm}^{800nm} S(\lambda)\sigma(\lambda)d\lambda \quad (7.3)$$

où $S(\lambda)$ est la densité spectrale de puissance de la lumière réfléchiée sur le capteur, K un facteur ne dépendant que de la géométrie de prise de vue et $\sigma(\lambda)$ la courbe de sensibilité spectrale de l'œil.

7.2.2 Choix d'un espace de couleur

Nous désirons créer l'image d'un objet éclairé par une source primaire de lumière (par exemple le Soleil). L'objet est alors considéré, du point de vue du capteur, comme une source secondaire de lumière émettant un rayonnement qui dépend du rayonnement incident auquel il est soumis. Le capteur, par exemple une caméra, est conçu pour recevoir l'information photométrique émise par l'objet et traduire ces données énergétiques en une image. On interpose des filtres entre l'objet et la caméra pour sélectionner le contenu spectral de la lumière mesurée.

Suivant le schéma de Maxwell, on décompose donc le rayonnement arrivant sur le capteur, suivant les trois primaires choisies, à l'aide de filtres appliqués successivement sur le capteur³. On obtient ainsi trois images, par exemple rouge, verte et bleue qui nous permettent de recréer l'image en couleur. Tout point de l'image est donc repéré dans un espace tridimensionnel par ses trois composantes chromatiques R, V et B (dénommées **composantes trichromatiques**), propres au capteur choisi, dont les expressions sont les suivantes :

$$R = \int_{-\infty}^{\infty} S(\lambda) T_r(\lambda) s_R(\lambda) d\lambda$$

$$V = \int_{-\infty}^{\infty} S(\lambda) T_v(\lambda) s_V(\lambda) d\lambda$$

$$B = \int_{-\infty}^{\infty} S(\lambda) T_b(\lambda) s_B(\lambda) d\lambda$$

où $S(\lambda)$ est la distribution spectrale de la source lumineuse secondaire,

³ou de trois capteurs masqués par un filtre chacun

$T_r(\lambda)$, $T_v(\lambda)$, $T_b(\lambda)$ sont les transmittances des filtres rouge, vert et bleu,
 $s_R(\lambda)$ est la sensibilité du capteur rouge, et de même pour le vert et le bleu.

L'espace RVB peut être représenté sous forme d'un cube dont trois des arêtes portent les trois primaires. La diagonale reliant les points caractérisés par les trois composantes $R = V = B = 0$ (noir), et $R = V = B = 1$ (blanc) est appelée axe achromatique. Cette représentation conventionnelle laisse croire que les composantes portées par les 3 axes orthogonaux sont indépendantes, mais on sait expérimentalement qu'il n'en est rien.

7.2.3 L'espace RVB de la CIE 1931

La CIE⁴ a défini un espace RVB de référence (CIERVB 1931) à partir des éléments suivants :

1. les 3 primaires ont pour puissance 1 Watt et sont monochromatiques, de longueurs d'ondes : 700 nm pour le rouge, 546 nm pour le vert et 436 nm pour le bleu ;
2. le blanc de référence est choisi comme W_0 de densité de puissance constante et égale à $5,3 \cdot 10^{-2}$ W/nm ; c'est le blanc d'égale énergie.

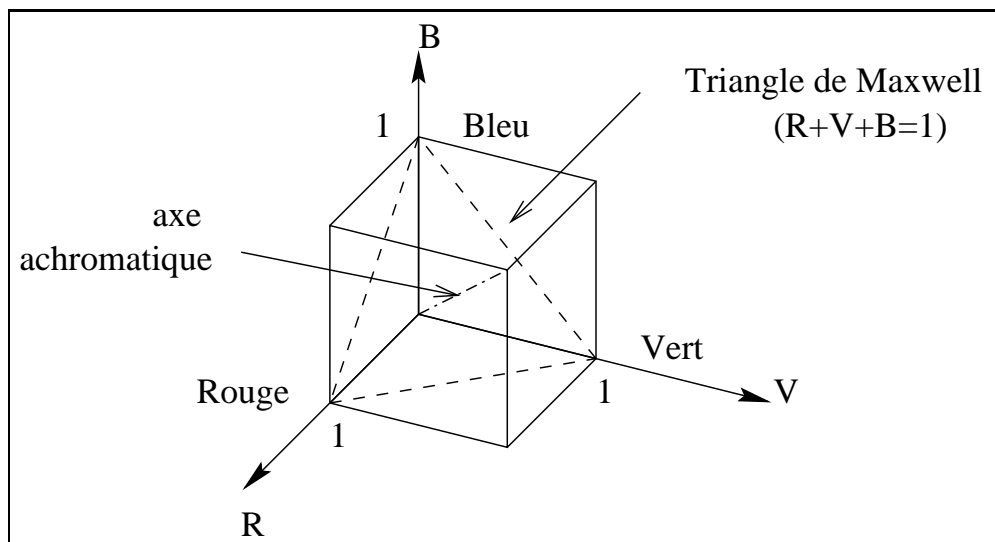


FIG. 7.4 – Le cube des couleurs et le triangle de Maxwell

Le triangle de Maxwell est le triangle défini, dans ce cube, par l'équation $R + V + B = 1$

Les fonctions colorimétriques de RVB CIE 1931

Les fonctions colorimétriques, définies pour le triplet rouge, vert, bleu, ont des valeurs nulles en dehors du domaine visible des longueurs d'onde $[400 \text{ nm}, 800 \text{ nm}]$. Elles sont égales aux **composantes trichromatiques** lorsque le stimulus à caractériser est de luminance constante et d'intégrale unité. Elles sont alors notées : $r(\lambda)$, $v(\lambda)$, $b(\lambda)$. On en déduit l'expression des composantes chromatiques lorsque le stimulus n'est pas de luminance unité, mais de luminance $L(\lambda)$:

$$R = \int L(\lambda) \overline{r(\lambda)} d\lambda$$

⁴CIE=Commission Internationale de l'Eclairage

$$V = \int L(\lambda) \overline{v(\lambda)} d\lambda$$

$$B = \int L(\lambda) \overline{b(\lambda)} d\lambda$$

Les fonctions colorimétriques sont également nommées courbes d'excitation ou fonctions de mélange. Elles sont représentées figure 7.5.

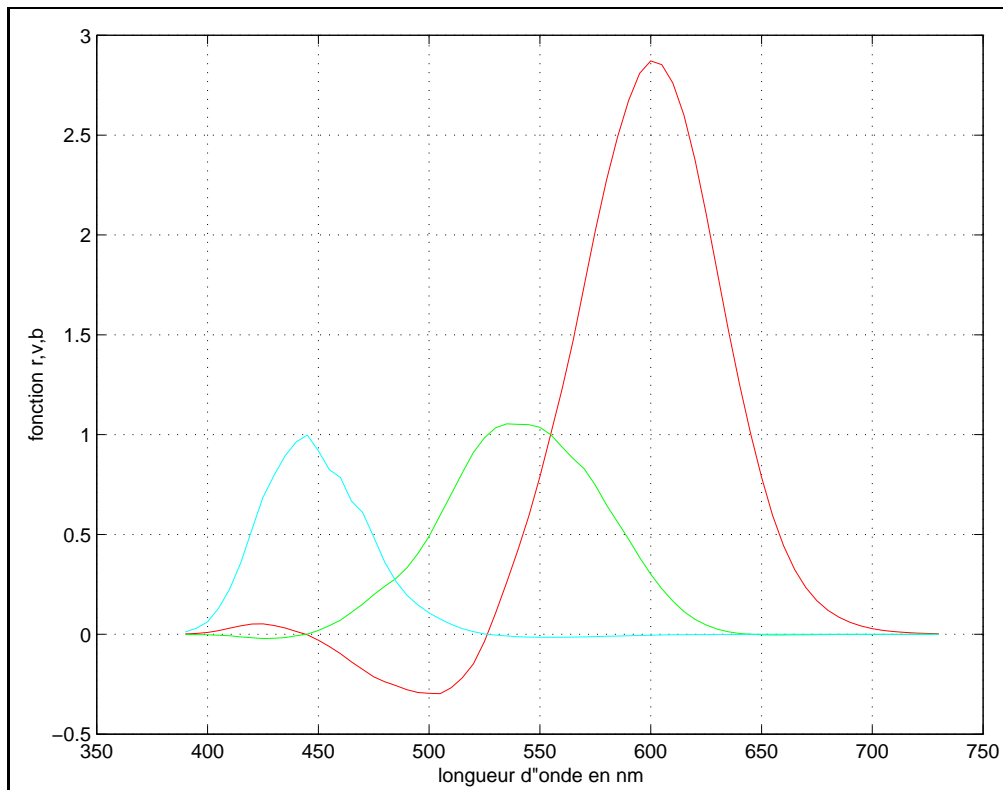


FIG. 7.5 – Fonctions colorimétriques. De gauche à droite : Bleue, Verte et Rouge. Notez les valeurs négatives (en particulier de Rouge), qui caractérisent une inhibition.

Limites de cet espace

1. L'espace RVB n'est pas le meilleur espace pour reproduire la perception visuelle humaine : il est par exemple difficile d'attacher simplement une couleur à un point de cet espace.
2. Les trois composantes RVB sont fortement corrélées (diminuer la composante V, fait apparaître la teinte plus rouge), et il est difficile de séparer la notion d'intensité de la notion de chromaticité.
3. une couleur pure a toujours au moins une composante négative (voir fig 7.5).

Pour éviter ces inconvénients, il a été proposé de passer à des espaces différents : tout d'abord l'espace XYZ, puis des espaces dont la représentation se rapproche de la perception humaine des couleurs, et dans lequel on peut décorréler l'intensité de la couleur (Lab, TLS, visuel hypothétique, et TLC).

Dans ces différents espaces, on définit mathématiquement les caractéristiques de la couleur : la Teinte, la Saturation et le Chroma. On peut également les définir littéralement indépendamment des espaces [Sève, 1996].

Définitions des grandeurs chromatiques

La Teinte est définie de la façon suivante : “Attribut de la sensation visuelle qui peut être décrit par des qualificatifs tels rouge, jaune...”. La teinte détermine donc la couleur d’une surface.

Le Chroma est “le niveau de coloration d’une surface, évalué relativement à la lumière qu’elle reçoit. Le chroma d’une surface donnée est un attribut perceptif indépendant du niveau d’éclairément. Pour une surface de chromaticité constante, le chroma augmente avec la clarté de la surface, contrairement à la saturation.”

La Saturation quant à elle peut être définie de la façon suivante : “Niveau de coloration d’une surface, évalué relativement à sa luminosité. Il permet d’estimer la proportion de couleur chromatique pure dans la sensation totale. La saturation d’une surface de chromaticité constante est un attribut perceptif indépendant de sa clarté.”

La Clarté est “un attribut d’une sensation visuelle, selon laquelle une surface paraît diffuser plus ou moins de lumière relativement à celle reçue. La clarté s’évalue relativement au diffuseur parfait éclairé dans les mêmes conditions.”

La Luminance est “une grandeur physique qui caractérise une surface émettant un rayonnement dans une direction donnée. C’est le quotient du flux émis par l’angle solide et l’aire apparente de la surface depuis la direction d’émission. ”

Les notions de chroma et de saturation sont souvent confondues, leurs expressions formelles seront définies par la suite et leurs notations seront respectivement C et S.

Dans les paragraphes suivants les différents espaces utilisés, faisant intervenir ces grandeurs, seront présentés. Ils nécessitent que l’on définisse tout d’abord l’espace intermédiaire XYZ.

7.2.4 L’espace XYZ

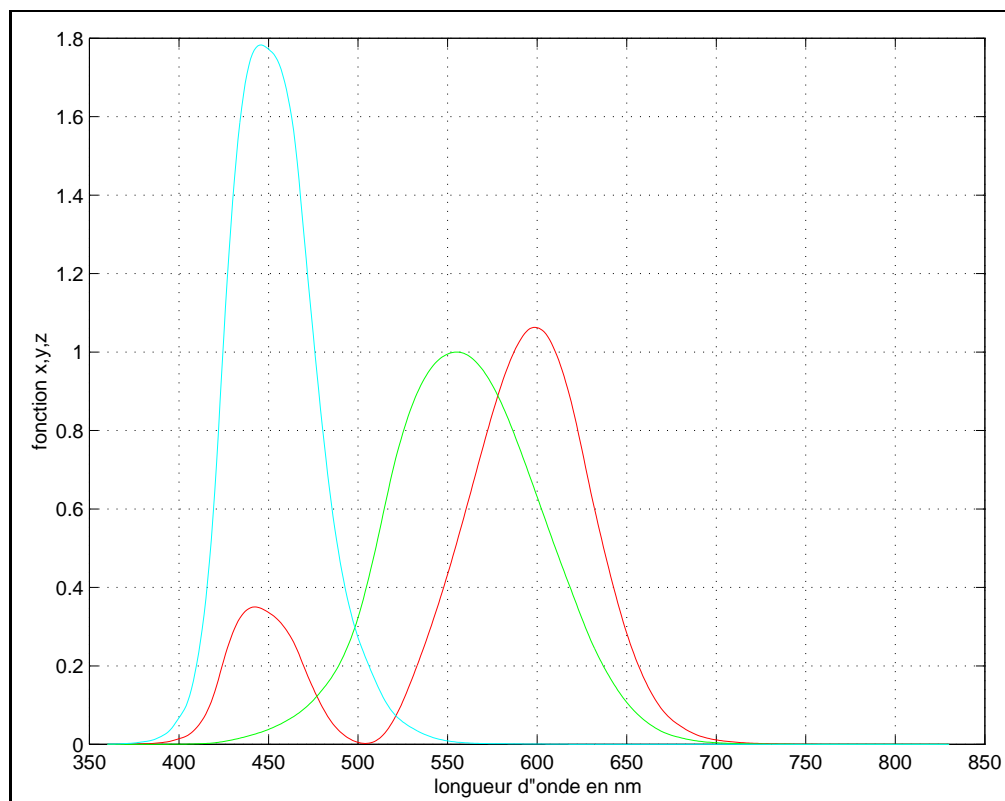
Comme nous pouvons le voir sur la figure 7.5, les fonctions colorimétriques prennent des valeurs négatives, qui correspondent donc à des inhibitions. La CIE proposa un changement de primaires afin d’avoir des fonctions colorimétriques à valeurs positives. Ces primaires furent nommées : $[X]$, $[Y]$, $[Z]$ et les fonctions colorimétriques associées \bar{x} , \bar{y} , \bar{z} . Les nouvelles fonctions colorimétriques sont représentées sur le schéma 7.6. Ces primaires ne sont pas physiques mais irréelles. On choisit en particulier \bar{y} égale à la sensibilité relative spectrale $\sigma(\lambda)$ de l’œil, ce qui permet de relier directement la luminance de la couleur à la valeur Y (voir équation 7.3) : $L(C) = KY$.

Le passage de l’espace RVB à l’espace XYZ se fait par une transformation linéaire. On définit donc une matrice de passage inversible nommée : RVB_XYZ (dont les valeurs dépendent du système physique d’acquisition et d’affichage de l’image si l’on ne se place pas dans RVB 1931). La transformation réciproque se fait grâce à la matrice inverse de RVB_XYZ nommée XYZ_RVB. Ces deux matrices sont, dans le cas du passage des espaces CIE 1931, définies de la façon suivante :

$$RVB_XYZ = \begin{bmatrix} 0.42 & -0.16 & -0.08 \\ -0.09 & 0.25 & 0.16 \\ 0.001 & -0.002 & 0.179 \end{bmatrix}$$

$$XYZ_RVB = RVB_XYZ^{-1} = \begin{bmatrix} 2.78 & 1.75 & 1.13 \\ 1 & 4.59 & 0.06 \\ 0 & 0.057 & 5.59 \end{bmatrix}$$

Les espaces RVB et XYZ n’ayant pas les mêmes blancs (ils sont en fait proportionnels), on définit un espace

FIG. 7.6 – Primaires \bar{x} , \bar{y} , \bar{z} de la CIE.

$X'Y'Z'$ qui se déduit de XYZ par changement de blanc et qui est relié à RVB par :

$$RVB_{X'Y'Z'} = \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.18 & 0.81 & 0.01 \\ 0.0 & 0.01 & 0.99 \end{bmatrix}$$

$$X'Y'Z'_{RVB} = RVB_{X'Y'Z'}^{-1} = \begin{bmatrix} 2.37 & -0.9 & -0.47 \\ -0.52 & 1.43 & 0.09 \\ 0.004 & -0.014 & 1.01 \end{bmatrix}$$

Des matrices semblables existent, permettant de passer par exemple des espaces NTSC, PAL ou SECAM aux espaces de la CIE et d'un type de télévision à l'autre (les deux systèmes européens ont les mêmes blancs et les mêmes luminophores).

A partir de cet espace nous pouvons définir l'espace Lab.

7.2.5 L'espace Lab et le repère TLC

L'espace Lab est un espace uniforme, ce qui signifie que les écarts de couleur dans cet espace sont égaux aux écarts de couleur perçus par un observateur (en première approximation). C'est un espace normalisé par le CIE. L'origine de cet espace vient de la propriété suivante : la luminance d'un rayonnement est indépendante de la chromaticité de ce dernier.

Il est important également de constater que dans le cas de l'espace RVB les composantes trichromatiques sont corrélées. Ainsi il est impossible d'obtenir l'information de luminance indépendamment de la chromaticité, en revanche, cela est possible dans le cas de l'espace Lab. On dispose de trois composantes pour un seul point de l'espace : L la clarté, a et b les coordonnées de chrominance. Celles-ci sont obtenues à partir des composantes X , Y , Z . On définit de plus un blanc de référence dans l'espace XYZ dont les coordonnées X_n, Y_n, Z_n sont déterminées en fonction de la nature de l'illuminant choisi [Sève, 1996].

- Illuminant A : $X_n = 109.85, Y_n = 100, Z_n = 35.58$.
- Illuminant C : $X_n = 98.07, Y_n = 100, Z_n = 118.23$.
- Illuminant D65 : $X_n = 95.04, Y_n = 100, Z_n = 108.88$.

L'espace Lab, tel qu'il a été adopté par le CIE en 1976, porte le nom de CIELab. Il est défini par les équations suivantes :

$$\begin{aligned} L &= 116 \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 \\ a &= 500 \left[\left(\frac{X}{X_n} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} \right] \\ b &= 200 \left[\left(\frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_n} \right)^{\frac{1}{3}} \right] \end{aligned}$$

Pour des valeurs de l'argument inférieures à 0.008856, la fonction puissance $x^{\frac{1}{3}}$ est remplacée par $7.787x + \frac{16}{116}$.

On peut ainsi voir que le plan orthogonal à l'axe des clartés est le plan chromatique contenant les deux axes a et b . On remarque que l'axe a peut être décomposé en deux couleurs : vert et rouge et l'axe b en bleu et jaune. On a donc :

- bleu pour $a = 0$ et $b < 0$ et jaune pour $a = 0$ et $b > 0$
- vert pour $a < 0$ et $b = 0$ et rouge pour $a > 0$ et $b = 0$

L'axe achromatique L est donc défini par $a = 0$ et $b = 0$.

Plutôt que de caractériser la couleur par les valeurs a et b , on préfère souvent utiliser les coordonnées cylindriques. Elles permettent d'introduire deux caractéristiques de la couleur : la teinte T et le chroma C dont les expressions sont les suivantes :

$$C = \sqrt{a^2 + b^2}$$

$$\begin{aligned} T &= a \cos \frac{a}{\sqrt{a^2 + b^2}} && \text{si } b > 0 \\ &= 2\pi - a \cos \frac{a}{\sqrt{a^2 + b^2}} && \text{sinon} \end{aligned}$$

La signification du chroma est la suivante : si dans l'espace Lab on effectue une coupe orthogonale à l'axe des luminances, à une valeur de L fixée, on définit un plan particulier de chromaticité. Le chroma est alors la distance d'un point de ce plan à l'axe achromatique. Les points d'isochroma se situent sur un cercle dont le centre est le point d'intersection entre l'axe achromatique et le plan de chrominance choisi.

La teinte est définie comme un angle, ses valeurs varient donc entre 0 et 2π ou entre $-\pi$ et π .

Remarque : la teinte est une notion délicate car une valeur de teinte à 0 semble très éloignée d'une valeur de teinte égale à 2π , alors que ce sont deux teintes identiques. Ces deux grandeurs sont représentées sur le schéma 7.7.

Caractéristiques de cet espace

1. L'espace Lab est un espace uniforme, donc par définition tout écart entre deux couleurs est égal à l'écart perçu par l'homme c'est donc un espace dont la métrique est proche de celle de l'espace visuel humain.

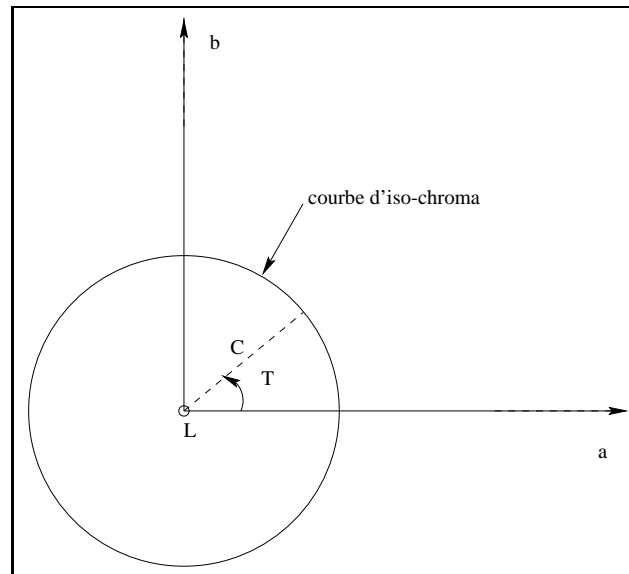


FIG. 7.7 – Coupe orthogonale à l'axe des luminances dans l'espace Lab : le repère TLC.

2. La composante L est décorrélée des composantes chromatiques, Malheureusement le chroma dépend de la clarté.
3. Le passage de RVB à Lab entraîne des pertes de précision numérique qui sont dues d'une part aux changements successifs d'espaces, et d'autre part à l'approximation causée par l'exposant $\frac{1}{3}$.
4. L'accès aux informations de teinte et de chroma est faussé par les différentes étapes nécessaires au calcul. C'est pourquoi il est préférable de passer par l'intermédiaire d'autres espaces pour définir ces grandeurs, notamment l'espace TLC.

7.2.6 L'espace YC1C2

C'est un espace générique. Il est obtenu à partir de l'espace RVB par transformation linéaire. La composante Y est définie comme étant la luminance et $C1$, $C2$ les valeurs de chrominance. On peut écrire cette transformation sous forme matricielle, on obtient alors :

$$\begin{bmatrix} Y \\ C2 \\ C1 \end{bmatrix} = [M] \begin{bmatrix} R \\ V \\ B \end{bmatrix}$$

La matrice M est définie suivant l'espace utilisé : TLS ou visuel hypothétique.

7.2.7 L'espace TLS

C'est un espace créé par les traiteurs d'images pour sa commodité d'emploi. Les composantes d'un point de cet espace sont : Teinte (T) Luminance (L) et Saturation (S). On l'appelle en anglais l'espace HIS (*Hue, Intensity, Saturation*). Les composantes sont définies, en passant à l'espace YC1C2, par la matrice M définie par :

$$M = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1 & -1/2 & -1/2 \\ 0 & \sqrt{3}/2 & -\sqrt{3}/2 \end{bmatrix}$$

On obtient alors les valeurs de T , L , S de la façon suivante :

$$L = Y$$

$$\begin{array}{lll} \text{si } R = V = B & \text{alors } S = 0 & \text{et } T = 0 \\ \text{sinon} & S = 1 - \frac{3\text{Min}(R,V,B)}{R+V+B} & \text{et si } C2 > 0 \text{ alors } T = \arccos\left(\frac{C1}{\sqrt{C1^2+C2^2}}\right) \\ & & \text{sinon } T = 2\pi - \arccos\left(\frac{C1}{\sqrt{C1^2+C2^2}}\right) \end{array}$$

La *saturation* est définie comme une distance séparant un point quelconque du point achromatique, on remarquera que ce n'est pas une distance euclidienne. Soit D' la droite passant par un point quelconque et par le point origine de l'axe achromatique. La teinte est définie comme l'angle compris entre la droite D' et l'axe de référence $C1$.

Cet espace présente un inconvénient : la fonction Min , qui définit la saturation, compromet la transformation réciproque passant de l'espace TLS à YC1C2. Pour éviter cela on peut utiliser le système visuel hypothétique [Carron, 1995].

7.2.8 Le système visuel hypothétique

Pour correspondre à la perception visuelle Garbay a défini ce système dont la sensibilité est indépendante de la longueur d'onde, du facteur de pureté et de l'intensité de la couleur considérée [Carron, 1995].

Dans ce modèle on définit une matrice M de passage à l'espace YC1C2 par :

$$M = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ \sqrt{3}/2 & -\sqrt{3}/2 & 0 \\ -1/2 & -1/2 & 1 \end{bmatrix}$$

Par la suite on définit les deux grandeurs suivantes : la *teinte* et le *chroma*.

Soit P un point de l'espace dont on désire connaître les propriétés chromatiques. Soit P' la projection orthogonale de P sur l'axe achromatique. Soit D la droite passant par P et P' .

La teinte est l'angle que font la droite D et l'axe de référence $C1$. Le chroma est la distance séparant le point considéré, de coordonnées $(C1, C2)$, de l'axe achromatique. D'où l'expression suivante pour le chroma : $C = \sqrt{C1^2 + C2^2}$

Remarques : Différences entre l'espace TLS et le système visuel hypothétique.

L'espace TLS est également appelé système de coordonnées triangulaires. Cela vient du fait que le plan de chrominance est défini par le triangle de Maxwell, l'axe achromatique est l'axe lui étant orthogonal, donc la diagonale principale du cube dans l'espace RVB. On remarquera qu'on différencie les termes de *Chroma* et de *Saturation*. En effet dans l'espace TLS, étant donné la définition de S ($S=f(\text{Min}(R,V,B))$) on peut dire que les courbes d'iso-saturation sont des triangles homothétiques au triangle de Maxwell et de taille inférieure à ce dernier, figure 7.8. Or dans le cas du système visuel hypothétique, le chroma n'est pas défini par une fonction Min , il est défini comme une distance euclidienne, donc les courbes d'iso-chroma sont des cercles, figure 7.8.

On peut noter également une différence dans les matrices de passage de RVB à YC1C2, la matrice M de l'espace visuel hypothétique introduit une rotation de 30° des axes $C1$ et $C2$ défini dans l'espace TLS, ce qui

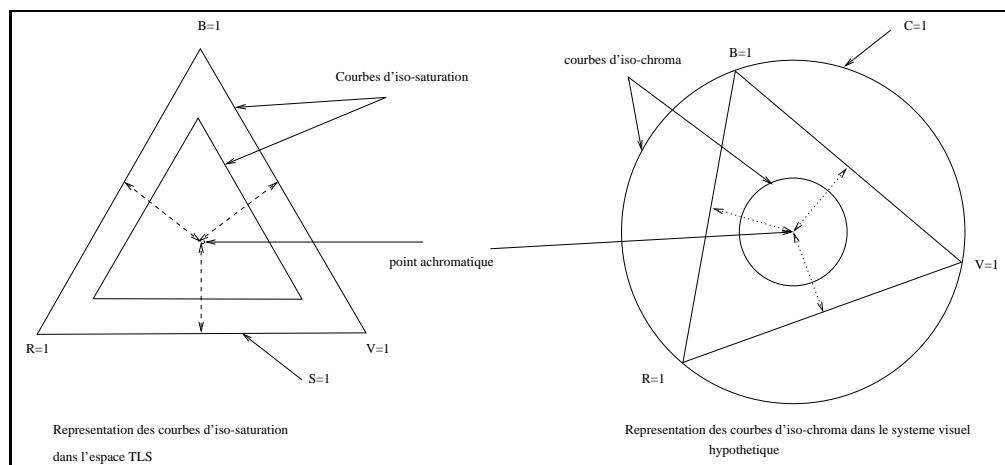


FIG. 7.8 – Les courbes d’iso-saturation dans l’espace TLS et d’iso-chroma dans l’espace visuel hypothétique.

implique que l’axe de référence C1 ne passe plus par le sommet rouge du triangle de Maxwell, mais correspond à l’axe Rouge-Vert. Si on reprend la schématisation de l’espace RVB (voir figure 7.4), on peut représenter les axes C1 et C2 des deux espaces, en considérant l’axe achromatique perpendiculaire au plan de la figure, voir schéma 7.9.

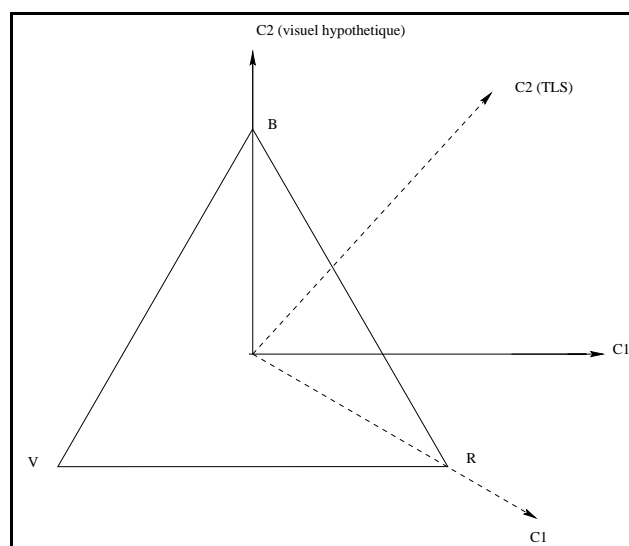


FIG. 7.9 – Les plans C1C2 de l’espace TLS et C1C2 du système visuel hypothétique

On voit qu’il existe des différences entre l’espace TLS et l’espace visuel hypothétique, cependant on constate que la formule de la saturation dans TLS ne permet pas la transformation inverse, pour cela il est préférable de choisir la formule du chroma définie dans l’espace visuel hypothétique. En revanche, dans ce système les axes C1 et C2 ne portent aucun sommet du triangle de Maxwell, ce qui privilégie le choix de la représentation par TLS. Mais plutôt que d’avoir à choisir entre ces deux espaces, on peut adopter la voie décrite par Carron [Carron, 1995] : faire un compromis entre les deux espaces, d’où l’apparition de l’espace TLC.

7.2.9 L'espace TLC

Cet espace est obtenu en effectuant un passage de l'espace RVB à l'espace YC1C2 par l'intermédiaire de la transformation linéaire exprimée à l'aide de la matrice M définie de la façon suivante :

$$M = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1 & -1/2 & -1/2 \\ 0 & -\sqrt{3}/2 & \sqrt{3}/2 \end{bmatrix}$$

La matrice M ainsi définie est très proche de celle définie pour l'espace TLS, on constate uniquement un changement d'orientation pour l'axe C2.

La matrice de passage de l'espace YC1C2 à l'espace RVB est définie par :

$$M^{-1} = \begin{bmatrix} 1 & 2/3 & 0 \\ 1 & -1/3 & \sqrt{3}/3 \\ 1 & -1/3 & -\sqrt{3}/3 \end{bmatrix}$$

Grâce au composantes Y, C1 et C2 nous pouvons définir les grandeurs suivantes : la luminance, la teinte et le chroma .

On obtient ainsi :

$$\begin{aligned} L &= Y \\ C &= \sqrt{C1^2 + C2^2} \\ \text{si } C1 > 0 \text{ alors } T &= \arccos\left(\frac{C1}{C}\right) \\ \text{sinon } T &= 2\pi - \arccos\left(\frac{C1}{C}\right) \end{aligned}$$

Ces formules sont inversibles ce qui permet un retour après traitement à l'espace YC1C2 puis à l'espace RVB.

De surcroît, on constate que le passage aux valeurs de teinte, luminance et chroma se fait en peu d'étapes et donc nécessite moins d'approximations que les expressions obtenues à partir de l'espace Lab. C'est pourquoi on préfère parfois utiliser l'espace TLC, afin d'éviter la perte de précision au cours des changements d'espaces.

7.2.10 Les espaces de la télévision

Afin de garantir une reproductivité universelle des couleurs, les systèmes de télévision se sont dotés de normes, aussi bien pour l'affichage que pour la transmission.

Système NTSC

Il propose un espace RVB relié à l'espace XYZ par les matrices :

$$RVB_XYZ_{NTSC} = \begin{bmatrix} 0.607 & 0.174 & 0.020 \\ 0.299 & 0.287 & 0.114 \\ 0 & 0.066 & 0.117 \end{bmatrix}$$

$$XYZ_RVB_{NTSC} = RVB_XYZ_{NTSC}^{-1} = \begin{bmatrix} 1.910 & -0.533 & -0.288 \\ -0.987 & 2.00 & -0.028 \\ 0.058 & -0.118 & 0.896 \end{bmatrix}$$

Systèmes PAL et SECAM

Ils proposent le même espace RVB relié à l'espace XYZ par les matrices :

$$RVB_XYZ_{PAL/SECAM} = \begin{bmatrix} 0.429 & 0.343 & 0.178 \\ 0.222 & 0.707 & 0.071 \\ 0.019 & 0.132 & 0.939 \end{bmatrix}$$

$$XYZ_RVB_{PAL/SECAM} = RVB_XYZ_{PAL/SECAM}^{-1} = \begin{bmatrix} 3.073 & -1.403 & -0.475 \\ -0.970 & 1.877 & 0.042 \\ 0.074 & -0.235 & 1.068 \end{bmatrix}$$

NTSC : les signaux en composantes

La transmission se fait par Y et I et Q : liées à RVB de NTSC par la matrice :

$$\begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.512 \end{bmatrix}$$

PAL : les signaux en composantes

On transmet Y et $U = 0.493(B - Y)$ ainsi que $V = 0.877(R - Y)$.

SECAM : les signaux en composantes

On transmet Y et $R - Y$ $B - Y$.

7.3 Modélisation de la perception colorée des images

Devant la complexité des expérimentations sur la couleur, un grand intérêt est apparu dans la modélisation de la perception de façon à évaluer automatiquement les performances d'un opérateur humain ou d'apprécier la qualité d'un document sans passer par une longue étape expérimentale.

Nous présentons ici un modèle de la vision - assez ancien - mais très représentatif de ceux que l'on a pu élaborer sur ce sujet. Ce modèle est cependant moins complet d'un point de vue colorimétrique que des modèles plus récents comme celui de Nayatani [Nayatani et al., 1990] ou celui de Guth [Guth, 1991], qui néanmoins ne permettent pas aisément de faire varier la distance d'observation.

Le modèle de Frei-Faugeras fait passer d'une image représentée par ses coordonnées RVB à une image dans un espace perceptuel $A C_1 C_2$, où A code la luminance et C_1 et C_2 les antagonismes rouge/vert et jaune/bleu. Un tel modèle comporte un étage de récepteurs non-linéaires, une séparation de l'information colorée en un 1 canal achromatique et 2 canaux chromatiques, un filtre spatial sur chaque canal (il est représenté sur la figure 7.10).

Le signal RVB , transformé en signal CIE 1931 XYZ , est tout d'abord converti en valeurs LMS . Celles-ci expriment assez bien l'absorption par ce qui semble être les trois familles de récepteurs des cônes des primates, L , M et S dénotant les longueurs d'ondes longues, moyennes et courtes. Cette absorption est suivie d'une étape de compression (traduite par un filtrage non-linéaire logarithmique) conduisant aux valeurs \tilde{L} , \tilde{M} , \tilde{S} . Un second étage donne naissance au canal achromatique A et aux valeurs d'antagonismes chromatiques C_1 et C_2 ($C_1 = \tilde{L} - \tilde{M}$, et $C_2 = \tilde{L} - \tilde{S}$). Enfin 3 filtres différents $H(f)$, $H_1(f)$ et $H_2(f)$ viennent rendre compte des interactions spatiales

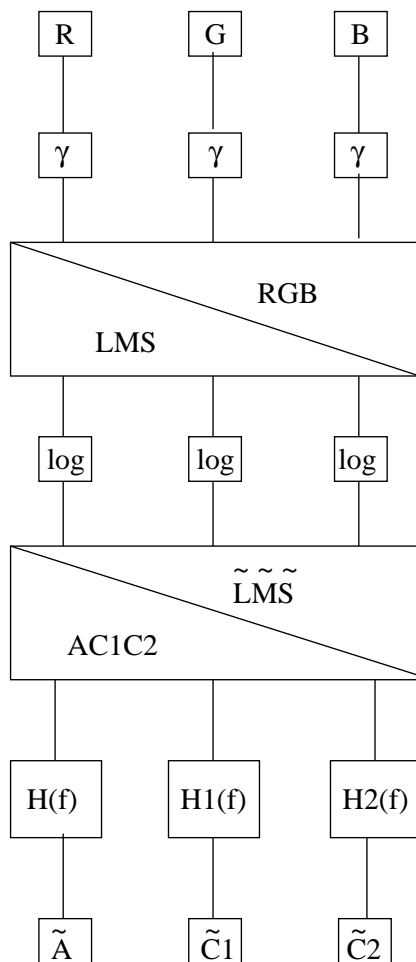


FIG. 7.10 – Le modèle de vision proposé par Frei et développé par Faugeras [Faugeras, 1976].

de chaque canal (respectivement A , C_1 et C_2) indépendamment. Les valeurs de ces divers filtres, ainsi que des matrices de composition et des non-linéarités du modèle sont proposées dans [Faugeras, 1976].

Les filtres spatiaux H_1 et H_2 agissant sur les canaux antagonistes, sont assez étroits, exprimant une nette sélection fréquentielle dont le rôle sera déterminant dans notre expérience (maximum de réponse respectivement à 4 cycles par degré et 2,6 cycles par degré).

Ces filtres ont été réalisés par différences de gaussiennes, conformément à l'approche de Marr [Marr et Hildreth, 1980] (voir chapitre sur la segmentation, paragraphe 2.5). Ils sont présentés sur la figure 7.11.

Le modèle développé par von Kreis reprend le modèle de Frei-Faugeras, mais tient compte du phénomène d'adaptation chromatique qui permet de conserver une perception quasi-permanente de la couleur même lorsque l'éclairage incident varie fortement en teinte. Pour cela von Kreis procède tout d'abord à une normalisation de la couleur sur l'ensemble de la scène observée et introduit des gains chromatiques pour rattraper un possible déséquilibre de l'éclairage.

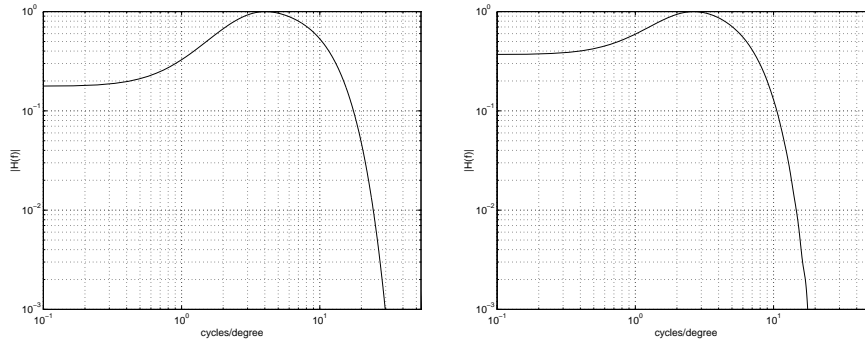


FIG. 7.11 – A gauche la fonction de transfert $H_1(f)$ du canal C_1 (à une dimension), calculée à partir d'une différence de deux gaussiennes. A droite : la fonction de transfert $H_2(f)$ du canal C_2 .

7.4 Les traitements des images en couleur

7.4.1 Un ordre pour l'espace R^3 ?

Un espace est **ordonné** s'il est muni d'une relation entre ses éléments (notée \leq) qui vérifie les 3 propriétés suivantes :

1. réflexivité : $x \leq x$,
2. transitivité : $x \leq y$ et $y \leq z$ implique $x \leq z$,
3. anti-symétrie : si $x \leq y$ et $y \leq x$ alors $x = y$.

Si seules les relations 1 et 2 sont vérifiées, alors la relation \leq est un **pré-ordre**.

Les entiers naturels sont ordonnés dans N . Les couples de N^2 ou les triplets de N^3 ne bénéficient pas d'ordre naturel.

On peut induire un ordre sur un espace non ordonné T' à partir d'un espace ordonné T si l'on fait correspondre à tout élément $x' \in T'$ un élément $x \in T$ par une relation : $x' = h(x)$ injective. Si la relation n'est pas injective alors elle définit un pré-ordre sur T' . On peut alors créer un ordre sur T' en remplaçant les x' par leur classe d'équivalence définie par : $X' = x' : h(x') = a$

On dit qu'il existe un **ordre total** sur un espace T si quelque soit $x, y \in T \times T$ soit $x \leq y$ soit $y \leq x$ (soit $x \leq y$ et $y \leq x$).

Si un ordre n'est pas total, il est dit **partiel**.

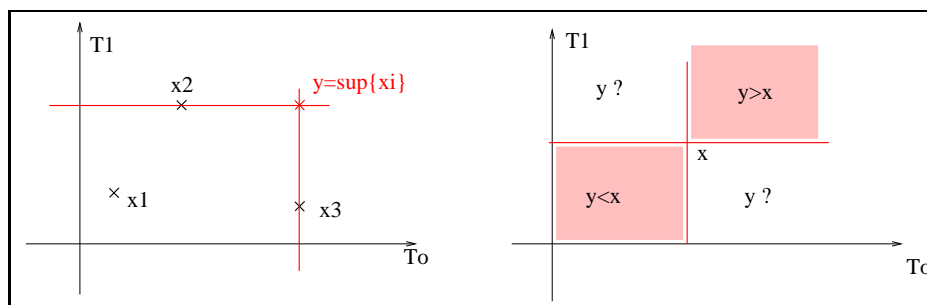


FIG. 7.12 – A gauche : y , le sup des x_i n'appartient pas à l'ensemble qui n'a donc pas de plus grand élément. A droite : l'ordre canonique définit les éléments supérieurs et inférieurs à x , mais ne classe pas les autres.

Si un espace $T' = \{X = \{x_0, x_1\}\}$ est espace produit de 2 espaces ordonnés $T_0 = x_0$ et $T_1 = x_1$, on obtient un ordre partiel dit **ordre canonique** par la relation conjonctive des deux relations d'ordre (voir figure 7.12). On ne peut donc définir au moyen de l'ordre canonique le dilaté (ou l'érodé) d'une image en couleur par cet ordre.

Si l'on veut bénéficier d'un certain nombre d'outils développés en traitement des images, il faut pouvoir induire un ordre total dans l'espace des couleurs. Avec un tel ordre, il est alors possible d'appliquer les algorithmes de la morphologie mathématique ou des filtres non-linéaires de rang , par exemple, (comme les filtres médians en particulier). L'impossibilité de le faire au moyen de l'ordre canonique est illustré sur la figure 7.12 pour la dilatation. On ne peut définir le dilaté (ou l'érodé) d'une image en couleur par cet ordre, puisque ni le min ni le max d'un ensemble de valeurs n'appartient forcément à cet ensemble.

Il a été proposé donc d'induire un ordre dans cet espace par un balayage judicieux des 3 dimensions (le choix d'une application h injective ci-dessus). Plusieurs ordres sont possibles, qui sont illustrés à 2 dimensions dans la figure 7.13. Aucun de ces balayages ne résoud complètement le problème car l'ordre induit dans l'espace transformé respecte mal la notion de voisinage que l'on avait dans l'espace \mathbb{R}^3 initial. Ceci est confirmé par le théorème de Netto :

Toute application injective de $[0, 1]^2$ dans $[0, 1]$ est discontinue.

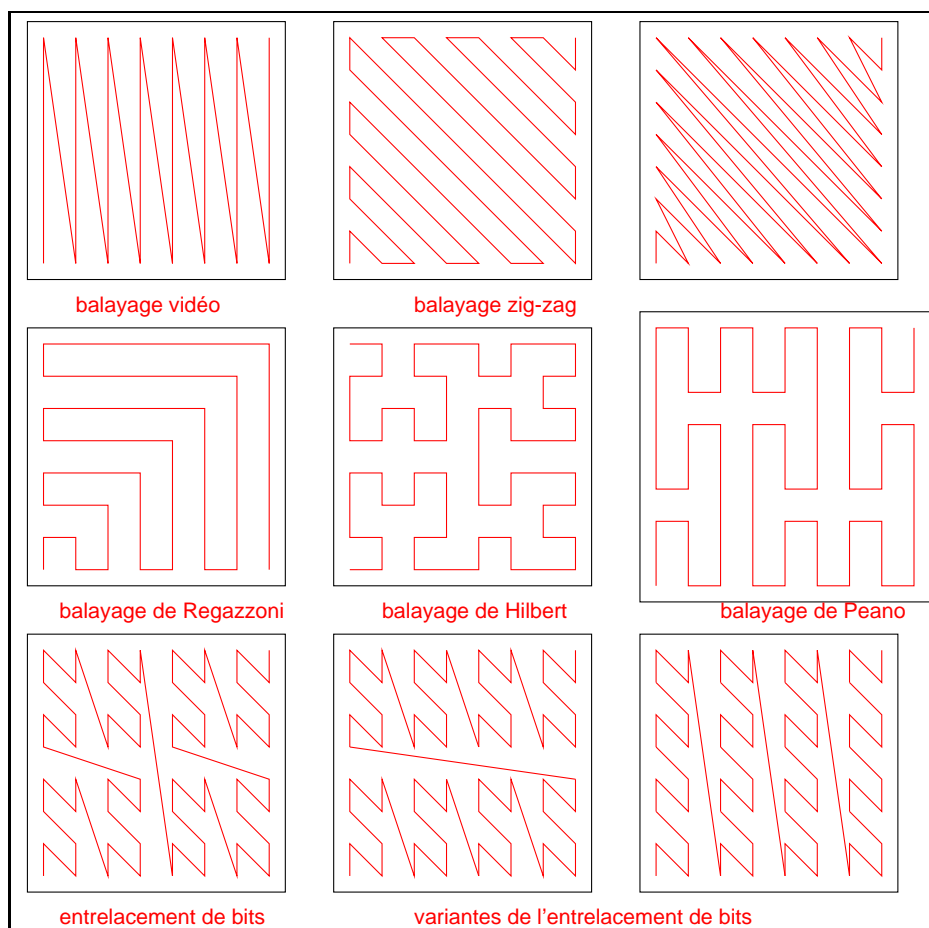


FIG. 7.13 – Les divers modes de balayage de l'espace R^2

Considérons la 8 connexité de l'espace \mathbb{R}^2 . Chaque point x a alors 8 voisins appartenant au voisinage V_x . Définissons une 8-connexité le long du balayage de l'image, qui associe les 4 points précédant x et les 4 points succédant à x le long de la courbe. Soit W_x ce voisinage. Un balayage préservera la topologie de \mathbb{R}^2 si les points de V_x se retrouvent dans W_x . Statistiquement, sur toute la surface $[0, 1]^2$, on trouve les résultats suivants :

balayage	Hilbert	Peano	Entrelacement	ZigZag	Regazzoni	TV
pourcentage de points du voisinage 1D appartenant au voisinage 2D	58 %	57 %	45-50 %	27 %	26 %	25 %

A ce titre, les balayages de Peano et de Hilbert (deux balayages à structure fractale puisqu'auto-similaires) et l'entrelacement de bits sont parmi les plus satisfaisants.

7.4.2 Qu'est ce qu'une couleur représentative d'un nuage ?

On suppose que l'on dispose de N points de couleur, on veut en extraire une valeur significative. Une fois choisi l'espace des couleurs, le choix d'une couleur moyenne est trivial puisqu'il revient à filtrer canal par canal et à constituer la valeur moyenne à partir des 3 moyennes des composantes.

On calcule parfois des moyennes "frustrées", c'est-à-dire calculée à partir de l'ensemble privé de ses éléments extrêmes. Pour cela on travaille en deux étapes :

1. à partir de l'ensemble des points on calcule une moyenne,
2. on écarte de l'ensemble tous les points à une distance trop grande de cette valeur moyenne,
3. une autre possibilité consiste à remplacer les points à distance trop grande par leur projection sur la sphère centrée en la moyenne et de rayon égal au seuil de distance,
4. pour les points retenus, on recalcule une moyenne. Cette valeur moyenne est la moyenne frustrée de l'ensemble.

Le choix d'une valeur médiane n'est pas trivial à cause des problèmes d'ordre exposés ci-dessus. Parmi les techniques choisies, on utilise souvent la technique du minimum de distance :

1. pour chaque point de l'ensemble on calcule la somme des distances à tous les autres,
2. on trie ces distances par ordre croissant,
3. on choisit pour point médian celui qui minimise la somme des distances.

On peut également choisir dans l'algorithme ci-dessus, non le point qui minimise la somme des distances, mais celui qui minimise la médiane des distances d'un point à tous les autres.

7.4.3 Comment calculer une palette ?

Les diverses techniques permettant de quantifier la couleur partent généralement d'un signal sur 2^{24} niveaux. Elles réduisent ce nombre par :

1. la quantification uniforme ; ex : palette 5,5,5 pour coder sur 15 bits (on réduit le nombre de couleurs à 36 000 environ, ce qui permet de construire des histogrammes) ou palette 4,2,2 pour coder sur 1 octet une image couleur.
2. la méthode LBG (Linde, Buzo, Gray) ou méthode des k-moyennes ; cette méthode de classification permet de minimiser la distorsion entre la représentation originale et la représentation quantifiée. Elle étend à N dimensions la solution classique de Lloyd et Max en choisissant [Linde et al., 1980] :
 - pour classe celle qui est la plus proche de l'échantillon considéré au sens de la norme adoptée pour mesurer la distorsion,

- pour représentant d'une classe le centre de gravité de cette classe.
- 3. la méthode des k-moyennes floues,
- 4. l'algorithme de population,
- 5. le *median-cut*,
- 6. l'*octree*
- 7. les réseaux de Kohonen

7.4.4 Comment filtrer 2 couleurs ?

Les expressions classiques du filtrage linéaire s'appliquent bien sûr aux images en couleur et l'extension à des signaux vectoriels ($f \in \mathbb{R}^3$ au lieu de $f \in \mathbb{R}$ comme dans les images à niveaux de gris), ne change pas fondamentalement le principe du filtrage. Il est cependant très important de bien choisir l'espace de représentation des signaux afin que la linéarité soit bien prise en compte dans l'espace souhaité. En effet, les équations fortement non linéaires de changement d'espaces (par exemple de XYZ à Lab) interdisent d'optimiser (par exemple aux moindres carrés) les filtres simultanément dans les divers espaces (figure 7.14).

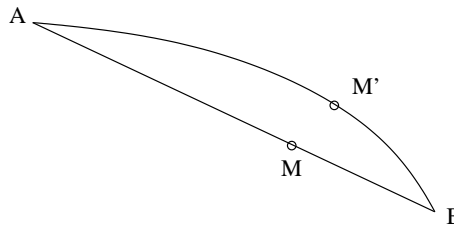


FIG. 7.14 – Lors de l'interpolation de deux couleurs, A et B, le passage par une transformation non-linéaire (par exemple l'exposant $\frac{1}{3}$ du passage de XYZ à Lab) conduit à des couleurs (ici M et M') très différentes.

7.4.5 La détection des contours des images en couleur

La détection des contours des images en couleur peut se faire tout d'abord en appliquant séparément une détection de contours scalaires (voir chapitre 2) sur chacune des composantes [diZenzo., 1986]. On dispose de 3 images de contours qu'il convient alors de composer pour obtenir une seule image de contours (on dit souvent « fusionner »). On choisit fréquemment une combinaison des 3 contours qui accepte tous les contours : par exemple au moyen d'un *max* des 3 images de contours.

$$\text{contour résultant} = \max \{ \text{contours des canaux} \}$$

Si l'on choisit une combinaison en modules :

$$\text{contour résultant} = \sqrt{\text{somme des contours}^2}$$

on se ramène alors souvent à une détection de contours sur le module de l'image (c'est-à-dire à peu près sur l'image de luminance) car beaucoup de détecteurs de contours sont linéaires (détecteur de Deriche, de Canny, de Shen). Attention, la recherche du maximum du gradient dans la direction du gradient n'est - elle - pas linéaire et il n'est pas équivalent de l'appliquer avant ou après la combinaison des canaux.

On peut tirer profit d'informations particulières sur le problème à traiter pour privilégier un type de représentation à d'autres. Par exemple on sait que les ombrages (aussi bien propres que portés) sont peu colorés. On peut alors ne segmenter que les composantes chromatiques des objets si l'on veut séparer des objets sans les fragmenter en fonction des ombres.

7.4.6 Représentation par quaternions

Cette technique particulière et originale peut s'appliquer à n'importe quel espace chromatique. Elle a été proposée dans [Sangwine, 1996]. Elle permet de manipuler la couleur dans une représentation analytique compacte autorisant la plupart des traitements accomplis sur les scalaires par une extension formelle.

Elle utilise la représentation par quaternions, de façon à représenter par un nombre unique un vecteur de \mathbb{R}^4 (et donc a fortiori de \mathbb{R}^3) de la même façon que la représentation par nombre complexe permet de manipuler des objets de \mathbb{R}^2 . Les quaternions forment un anneau qui peut être construit à partir de tout anneau commutatif. Dans le domaine du traitement de l'image on utilise les quaternions construits sur \mathbb{R} dans les applications comme la calibration des caméras où le changement d'espace de représentation ou le matching de points. Mais les quaternions sont également utilisés à partir de leur construction sur Z et sur en théorie des nombres ou en mécanique quantique.

Ils se construisent à partir d'une base constituée de $(1, i, j, k)$, vérifiant des propriétés de multilatativité semblables à celles des complexes :

$$\begin{aligned} 1.i &= i & 1.j &= j & 1.k &= k \\ ij &= -ji = k & jk &= -kj = i & ki &= -ik = j \\ i^2 &= j^2 = k^2 & &= -1 \end{aligned}$$

Un quaternion p se compose de 2 parties, l'une scalaire a , l'autre vectorielle q :

$$p = a + bi + cj + dk = a + q$$

Les quaternions vectoriels (c'est-à-dire tels que $a = 0$), également appelés quaternions purs, forment un espace vectoriel isomorphe à \mathbb{R}^3 .

Les opérations sur les quaternions permettent de définir le conjugué \bar{p} de $p = a + q$, ($\bar{p} = a - q$), la somme, le produit, le produit scalaire de 2 quaternions ($\langle p', p \rangle = 1/2(p'\bar{p} + p\bar{p}')$) et la norme d'un quaternion ($\|p\| = \sqrt{p\bar{p}}$). Le produit n'est pas commutatif, sauf si les 2 quaternions ont leurs parties vectorielles proportionnelles. En effet :

$$pp' = (aa' - qq', aq' + a'q + q \wedge q')$$

7.4.7 Couleur et quaternion

Représenter un point de couleur R, V, B par un quaternion pur, c'est tout d'abord choisir un espace de couleur (par exemple RVB), puis une origine de l'espace des couleurs (ce peut être le noir $(0,0,0)$, mais dans [Evans et al., 2000] on préfère le point de coordonnées $(127.5, 127.5, 127.5)$ qui se situe au centre du cube des couleurs. On représente ensuite un point de couleur par un quaternion pur. Par exemple comme :

$$q = i(R - O) + j(V - O) + k(B - O)$$

Les opérations conduites sur les quaternions (produits, convolutions, etc.) héritent des propriétés spécifiques du produit de quaternions. En particulier, pour des quaternions purs q et q' , on vérifie aisément que la partie scalaire $Scal[q.q']$ du produit de quaternions q et q' vaut :

$$Scal[qq'] = -q.q'$$

où \cdot dénote le produit scalaire de 2 vecteurs, et la partie vectorielle $Vect$:

$$Vect[qq'] = q \times q'$$

où \times représente le produit vectoriel. Ces propriétés permettent d'obtenir des résultats particuliers par des mises en équation très simples.

Par exemple dans [Evans et al., 2000], on montre que l'on peut filtrer une image en couleur par un filtre passe-bas qui n'agit que sur la composante chromatique parallèle à une couleur donnée quelconque C_1 (et pas sur les composantes orthogonales) en construisant le quaternion unitaire $\mu_{C_1} : \mu_{C_1} = \frac{C_1}{\|C_1\|}$, puis en convolant (à droite ou à gauche) l'image q par un noyau N (par exemple de 3×3 pixels) d'amplitude μ_{C_1} :

$$N = \frac{1}{9} \begin{bmatrix} \mu_{C_1} & \mu_{C_1} & \mu_{C_1} \\ \mu_{C_1} & \mu_{C_1} & \mu_{C_1} \\ \mu_{C_1} & \mu_{C_1} & \mu_{C_1} \end{bmatrix}$$

En effet, ces convolutions : $N * q = u$ et $q * N = v$ fournissent 2 quaternions u et v de même partie réelle et de parties vectorielles opposées. Les parties scalaires obtenues sont égales à la moyenne des composantes chromatiques de l'image parallèles à μ_{C_1} .

$$\text{Scal}[N * f](x, y) = \sum_{s, t} -\frac{1}{9} \mu_{C_1} f(x - s, y - t)$$

Les parties vectorielles sont \pm égales au produit vectoriel :

$$\text{Vect}[N * f](x, y) = -\text{Vect}[f * N](x, y) = \sum_{s, t} \frac{1}{9} \mu_{C_1} \times f(x - s, y - t)$$

Leur somme $u + v$ est un scalaire qui peut être orienté selon C_1 en le multipliant par μ_{C_1} . On aboutit à un filtrage hypercomplexe M qui vaut :

$$M = q - \frac{\mu_{C_1}}{2} [N' * q + q * N']$$

avec

$$N' = \frac{1}{9} \begin{bmatrix} \mu_{C_1} & \mu_{C_1} & \mu_{C_1} \\ \mu_{C_1} & -8\mu_{C_1} & \mu_{C_1} \\ \mu_{C_1} & \mu_{C_1} & \mu_{C_1} \end{bmatrix}$$

Cette représentation par quaternions permet de définir d'une façon semblable des détecteurs de contours [Sangwine, 1998], des filtres passe-bas [Evans et al., 2000], des TF de l'espace des couleurs ([Ell et Sangwine, 2000]).

Annexe A

Décomposition en Valeurs Singulières : SVD

Chapitre rédigé par Denis MATIGNON

A.1 Définition

On étudie la décomposition d'une matrice $A \in \mathbb{C}^{p \times m}$, éventuellement dans $\mathbb{R}^{p \times m}$; les résultats donnés pour les matrices réelles se particularisent en changeant les transposées-conjuguées (V^H) en transposées (V^T), et les matrices unitaires sont remplacées par des matrices orthogonales.

Théorème 1. Soit $A \in \mathbb{C}^{p \times m}$, il existe des matrices unitaires $U \in \mathbb{C}^{p \times p}$ et $V \in \mathbb{C}^{m \times m}$ telles que :

$$U^H A V = \Sigma \in \mathbb{R}^{p \times m} \quad \text{avec } \Sigma = \text{diag}(\sigma_1, \dots, \sigma_n) \quad \text{où } n = \min(p, m) \quad \text{et } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0.$$

La matrice rectangulaire Σ est constituée de 0 sauf sur la diagonale qui comporte n réels positifs, les σ_i ou *valeurs singulières* de A ; les u_i et les v_i (colonnes de U et V respectivement) constituent les *vecteurs singuliers* de A à gauche et à droite respectivement :

$$A v_i = \sigma_i u_i \quad \text{et} \quad A^H u_i = \sigma_i v_i$$

L'existence de la SVD provient de la diagonalisabilité en base orthonormée des matrices hermitiennes $A^H A$ et $A A^H$, et de leur positivité. On a l'unicité des valeurs singulières (comme on a celle du spectre de $A^H A$); en revanche, l'unicité de la SVD n'est assurée que si la matrice est carrée et possède n valeurs singulières deux à deux distinctes. Mais ceci n'est jamais vraiment gênant pour les propriétés structurelles que l'on peut étudier grâce à la SVD.

A.2 Lien avec le rang et interprétation de la décomposition

On appelle $\sigma_{\max}(A) = \sigma_1$ la plus grande valeur singulière de A et $\sigma_{\min}(A) = \sigma_n$ la plus petite valeur singulière de A (éventuellement nulle). Si le classement se précise comme suit :

$$\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_n = 0$$

alors r est le *rang* de A , et on peut caractériser le noyau et l'image de A par :

$$\ker A = \{v_{r+1}, \dots, v_m\} \quad \text{et} \quad \text{im } A = \{u_1, \dots, u_r\}$$

On peut alors donner une expression plus compacte de la décomposition de A , en notant $U_r = U(:, 1 : r)$, $V_r = V(:, 1 : r)$ les sous-matrices rectangulaires extraites, et $\Sigma_r = \Sigma(1 : r, 1 : r) = \text{diag}(\sigma_1, \dots, \sigma_r)$ la matrice carrée diagonale extraite ou partie *utile* de Σ :

$$A = U\Sigma V^H = U_r \Sigma_r V_r^H = \sum_{i=1}^r \sigma_i u_i v_i^H$$

Sur cette dernière décomposition, on voit très bien comment agit A sur chaque vecteur v_i de la base orthornormée de l'espace de départ \mathbb{C}^m : elle le transforme en σ_i fois le vecteur u_i de la base orthornormée de l'espace d'arrivée \mathbb{C}^p .

Pour A^H , on montre de même que $\ker A^H = \{u_{r+1}, \dots, u_p\}$ et $\text{im } A^H = \{v_1, \dots, v_r\}$; et l'on a :

$$A^H = V\Sigma^T U^H = V_r \Sigma_r U_r^H = \sum_{i=1}^r \sigma_i v_i u_i^H$$

A.3 Normes

La norme 2 et la norme Frobenius de A se calculent aisément à partir de la SVD :

$$\begin{aligned} \|A\|_2 &\triangleq \sup_{\|x\|=1} \|Ax\| = \sqrt{\max(\text{spec } A^H A)} = \sigma_1 \\ \|A\|_F &\triangleq \sqrt{\text{tr}(A^H A)} = \sqrt{\sum_{i,j} |a_{ij}|^2} = \sqrt{\sigma_1^2 + \dots + \sigma_r^2} \end{aligned}$$

La propriété suivante joue un très grand rôle dans les problèmes d'approximation :

Théorème 2. Soit $\{U, \Sigma, V\}$ une SVD de A , considérons pour $1 \leq k \leq r = \text{rang}(A)$ la suite de matrices $A_k = \sum_{i=1}^k \sigma_i u_i v_i^H$ de rang exactement k , alors :

$$\min_{B, \text{rang}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

En fait, les valeurs singulières de A mesurent la distance (en norme 2) de la matrice A par rapport aux matrices de rang inférieur.

Plus précisément, si l'on se donne $\varepsilon > 0$ comme tolérance numérique, et que l'on mesure le rang numérique de A correspondant par $r_\varepsilon = \text{rang}(A, \varepsilon) = \min_{B, \|A-B\|_2 \leq \varepsilon} \text{rang}(B)$, le théorème 2 permet d'écrire :

$$\sigma_1 \geq \dots \geq \sigma_{r_\varepsilon} > \varepsilon \geq \sigma_{r_\varepsilon+1} \geq \dots \geq \sigma_r$$

Annexe B

Applications de la SVD à la résolution des systèmes linéaires

Chapitre rédigé par Denis MATIGNON

B.1 Conditionnement d'un système linéaire non singulier

Considérons le système linéaire $Ax = b$ supposé régulier, c'est-à-dire A inversible. On cherche à évaluer les variations de la solution x relativement aux variations de A et de b (imprécisions sur le système et les données).

Le calcul différentiel permet une analyse directe dans le cas régulier : comme $x = A^{-1}b$, alors $dx = d(A^{-1}b) = d(A^{-1})b + A^{-1}db$; enfin $d(A^{-1}) = -A^{-1}dA A^{-1}$, d'où :

$$dx = -A^{-1}dA A^{-1}b + A^{-1}db = A^{-1}[-dA x + db], \text{ et } \|dx\| \leq \|A^{-1}\| [\|dA\| \|x\| + \|db\|]$$

En divisant le tout par $\|x\|$, et en utilisant que $\|b\| \leq \|A\| \|x\|$, on obtient la majoration suivante :

$$\frac{\|dx\|}{\|x\|} \leq \|A\| \|A^{-1}\| \left[\frac{\|dA\|}{\|A\|} + \frac{\|db\|}{\|b\|} \right]$$

Le nombre $C = \|A\| \|A^{-1}\|$ apparaît clairement comme une mesure de la sensibilité de l'opération d'inversion. Caractérisons la par les valeurs singulières de A .

En introduisant une SVD de $A : A = U\Sigma V^H = \sum_{i=1}^r \sigma_i u_i v_i^H$ (dans le cas régulier, $r = n = m = p$), on trouve immédiatement la décomposition correspondante de $A^{-1} : A^{-1} = V\Sigma^{-1}U^H = \sum_{i=1}^r \sigma_i^{-1} v_i u_i^H$. Il est alors évident que $C = \sigma_1/\sigma_n$; de plus, un raisonnement précis utilisant les vecteurs singuliers montre que la borne proposée est optimale pour certaines directions de A, dA, b, db .

Définition 1. $C = \|A\| \|A^{-1}\| = \sigma_{\max}(A)/\sigma_{\min}(A)$ est le nombre de condition ou conditionnement de la matrice A .

Remarquons que C est au minimum égal à 1 pour les matrices $A = \lambda U$ où U est unitaire. C peut être très grand, et ce même pour des matrices dont les valeurs propres sont raisonnables.

B.2 Pseudo-inverse

Plusieurs causes de singularité ($\sigma_{\min}(A) = 0$) existent dans les systèmes linéaires $y = Ax$:

- A carrée n'est plus inversible ; si $y \in \text{im } A$ alors il y a un sous-espace affine de solutions et il peut être intéressant de chercher celle de norme minimale ; si $y \notin \text{im } A$, il n'y a donc pas de solution exacte, en revanche il y a forcément une solution qui minimise la distance $\|y - Ax\|$.
- A est rectangulaire verticale avec $m < p$, autrement dit il y a plus d'équations que d'inconnues : on parle alors de système surdéterminé ; là encore, il peut n'y avoir aucune solution exacte, mais une solution qui minimise l'erreur $\|y - Ax\|$.
- A est rectangulaire horizontale avec $p < m$, autrement dit il y a plus d'inconnues que d'équations : on parle alors de système sous-déterminé ; bien souvent il y a un sous-espace affine de solutions, et il peut être intéressant de calculer la solution de norme minimale.

Nous donnons à présent la définition de la pseudo-inverse d'une matrice en ayant recours à la SVD puis nous la spécifions dans deux cas particuliers parmi ceux évoqués ci-dessus.

Définition 2. Soit $\{U, \Sigma, V\}$ une SVD de A de rang r , on appelle pseudo-inverse de A la matrice A^\dagger définie par :

$$A^\dagger = V_r \Sigma_r^{-1} U_r^H = \sum_{i=1}^r \sigma_i^{-1} v_i u_i^H$$

Elle vérifie :

$$AA^\dagger A = A \quad \text{et} \quad A^\dagger AA^\dagger = A^\dagger$$

Contrairement à ce que laisserait supposer la définition, A^\dagger est définie de manière unique et ne dépend pas du choix de SVD que l'on fait pour la construire. On a en particulier :

$$A^\dagger A = V_r V_r^H \text{ est la projection orthogonale sur } \text{im } A^H = (\ker A)^\perp = \{v_1, \dots, v_r\} \text{ dans } \mathbb{C}^m,$$

$$AA^\dagger = U_r U_r^H \text{ est la projection orthogonale sur } \text{im } A = (\ker A^H)^\perp = \{u_1, \dots, u_r\} \text{ dans } \mathbb{C}^p.$$

On donne maintenant une expression directe de A^\dagger dans les deux cas particuliers de rang plein, où $r = n$.

Cas de système surdéterminé avec $r = m < p$ dans ce cas $V_r = V$, et l'on a alors $A^H A = V_r \Sigma_r^2 V_r^H$ carrée $m \times m$ inversible, d'inverse $(A^H A)^{-1} = V_r \Sigma_r^{-2} V_r^H$, d'où :

$$(A^H A)^{-1} A^H = V_r \Sigma_r^{-2} V_r^H V_r \Sigma_r U^H = V_r \Sigma_r^{-1} U^H = A^\dagger$$

Cas de système sous-déterminé avec $r = p < m$ dans ce cas $U_r = U$, et l'on a alors $AA^H = U_r \Sigma_r^2 U_r^H$ carrée $p \times p$ inversible, d'inverse $(AA^H)^{-1} = U_r \Sigma_r^{-2} U_r^H$, d'où :

$$A^H (AA^H)^{-1} = V_r \Sigma_r U_r^H U_r \Sigma_r^{-2} U_r^H = V_r \Sigma_r^{-1} U_r^H = A^\dagger$$

Nous voyons maintenant l'utilité de la notion de pseudo-inverse pour calculer des solutions robustes d'un système linéaire général, c'est-à-dire potentiellement singulier ou proche de la singularité.

B.3 Solutions robustes d'un système linéaire, moindres carrés

On cherche à résoudre $y = Ax$, où $x \in \mathbb{C}^m, y \in \mathbb{C}^p$; au cas où ce système n'a pas de solution exacte on cherchera la solution qui minimise l'erreur $\|y - Ax\|$ (ou solution en *moindres carrés* puisque la norme est euclidienne) ; au cas où il y a plusieurs solutions, on prendra celle qui est de norme minimale, autrement dit qui minimise $\|x\|$.

Théorème 3. La solution de norme $\|x\|$ minimale dans \mathbb{C}^m qui minimise l'erreur $\|y - Ax\|$ dans \mathbb{C}^p est exactement donnée par :

$$x^* = A^\dagger y$$

Pour prouver ce théorème fondamental, qui donne tout son sens à la notion de pseudo-inverse, nous allons utiliser $\{U, \Sigma, V\}$ une SVD de A . Comme U, V sont des matrices unitaires, elles conservent la norme euclidienne dans $\mathbb{C}^p, \mathbb{C}^m$ respectivement; en conséquence :

$\delta = y - Ax$ et $U^H \delta$ ont même norme,
 x et $V^H x$ ont même norme,

En notant désormais $a = V^H x$ et $b = U^H y$, le problème se met sous la forme beaucoup plus simple : trouver la solution de norme $\|a\|$ minimale dans \mathbb{C}^m qui minimise l'erreur $\|b - \Sigma a\|$ dans \mathbb{C}^p . Or en décomposant Σ par blocs, il vient :

$$\|b - \Sigma a\|^2 = \sum_{i=1}^r |b_i - \sigma_i a_i|^2 + \sum_{i=r+1}^p |b_i|^2$$

On en déduit que le minimum vaut $\sum_{i=r+1}^p |b_i|^2$, et qu'il est atteint pour les a tels que $a_i = \sigma_i^{-1} b_i$ pour $1 \leq i \leq r$ et a_i quelconques pour $r + 1 \leq i \leq m$ (ce qui correspond exactement à la paramétrisation de $\ker A$); dans ce sous-espace affine de solutions, celle de norme minimale est celle pour laquelle $a_i = 0$ pour $r + 1 \leq i \leq m$: en réinterprétant cela matriciellement, il vient $a^* = \Sigma^\dagger b$. En revenant dans les bases initiales $A^\dagger = V \Sigma^\dagger U^H$, et l'on a alors prouvé le théorème.

Nous regardons plus précisément ce qui se passe dans les deux cas particuliers envisagés précédemment, où $r = n$.

Cas de système surdéterminé avec $r = m < p$ dans ce cas $\ker A = \{0\}$ et le minimum de $\|y - Ax\|$ est atteint en un point unique $x^* = A^\dagger y = (A^H A)^{-1} A^H y$; ce minimum est nul si et seulement si $y \in \text{im } A$.

Cas de système sous-déterminé avec $r = p < m$ dans ce cas $\text{im } A = \mathbb{C}^p$ tout entier et $y = Ax$ possède un sous-espace affine de solutions exactes $x^* + \ker A$, où $x^* = A^\dagger y = A^H (A A^H)^{-1} y$ est celle de norme minimale.

Annexe C

Optimisation sous contraintes

Chapitre rédigé par Denis MATIGNON

C.1 Contraintes égalités

Nous donnons le théorème relatif à la minimisation d'une forme quadratique positive sous contraintes égalités linéaires.

Théorème 4. Soit à trouver le minimum de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ avec $f(x) = \frac{1}{2}x^T Qx + L_0^T x$ sous m contraintes égalités $L_i^T x = y_i$. $Q \geq 0$ et les L_0, L_1, \dots, L_m sont des vecteurs.

On définit le lagrangien du problème :

$$\mathcal{L}(x, \lambda_1, \dots, \lambda_m) \triangleq f(x) - \sum_{i=1}^m \lambda_i (L_i^T x - y_i)$$

f admet un minimum en x^* si et seulement si il existe des multiplicateurs de Lagrange associés λ_i^* pour lesquels le point $(x^*, \lambda_1^*, \dots, \lambda_m^*)$ est un point critique du lagrangien, soit :

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial x}(x, \lambda_1, \dots, \lambda_m) = x^T Q + L_0^T - \sum_{i=1}^m \lambda_i L_i^T \\ 0 &= \frac{\partial \mathcal{L}}{\partial \lambda_i}(x, \lambda_1, \dots, \lambda_m) = L_i^T x - y_i \quad \text{pour } 1 \leq i \leq m \end{aligned}$$

Ce théorème est une version très particulière et fort utile dans la pratique de l'optimisation sous contraintes égalités de fonctions à valeurs réelles ; « condition de Lagrange » optimisation.

C.2 Contraintes inégalités

Nous donnons le théorème relatif à la minimisation d'une forme quadratique positive sous contraintes égalités et inégalités linéaires.

Théorème 5. Soit à trouver le minimum de $f : \mathbb{R}^n \rightarrow \mathbb{R}$ avec $f(x) = \frac{1}{2}x^T Qx + L_0^T x$ sous m contraintes égalités $L_i^T x = y_i$ et p contraintes inégalités $M_j^T x \geq z_j$. $Q \geq 0$ et les $L_0, L_1, \dots, L_m, M_1, \dots, M_p$ sont des vecteurs.

On définit le lagrangien du problème :

$$\mathcal{L}(x, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_p) \triangleq f(x) - \sum_{i=1}^m \lambda_i (L_i^T x - y_i) - \sum_{j=1}^p \mu_j (M_j^T x - z_j)$$

f admet un minimum en x^* si et seulement s'il existe des multiplicateurs de Lagrange associés λ_i^* et $\mu_j^* \geq 0$ pour lesquels le point $(x^*, \lambda_1^*, \dots, \lambda_m^*, \mu_1^*, \dots, \mu_p^*)$ est un point critique partiel du lagrangien, soit :

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial x}(x, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_p) = x^T Q + L_0^T - \sum_{i=1}^m \lambda_i L_i^T \\ 0 &= \frac{\partial \mathcal{L}}{\partial \lambda_i}(x, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_p) = L_i^T x - y_i \quad \text{pour } 1 \leq i \leq m \\ 0 &\leq \frac{\partial \mathcal{L}}{\partial \mu_j}(x, \lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_p) = M_j^T x - z_j \quad \text{pour } 1 \leq j \leq p \\ \text{si } 0 &\neq M_j^T x^* - z_j \quad \text{alors } \mu_j = 0 \quad \text{pour } 1 \leq j \leq p \end{aligned}$$

Voir des résultats plus généraux « condition de Kuhn et Tucker » en optimisation.

Bibliographie

- [Adams et Bishof, 1994] Adams, R. et Bishof, L. (1994). Seeded region growing. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 16 :641–647.
- [Alata et Cariou, 2001] Alata, O. et Cariou, C. (2001). Modélisation stochastique linéaire 2-D. In Garello, R., éditeur, *Analyse de signaux bidimensionnels*, chapitre 2, pages 67–116. Hermès - Collection IC2, Paris.
- [Alvarez, 1996] Alvarez, L. (1996). Images and pde's. In Berger, M., Deriche, R., Herlin, I., Jaffré, J., et Morel, J.-M., éditeurs, *Lecture Notes in Control and Information Sciences*, volume 219, pages 3–14.
- [Amini et al., 1990] Amini, A., Weymouth, T., et Jain, R. (1990). Using dynamic programming for solving variational problems in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Pattern Analysis and Machine Intelligence-12(9) :855–867.
- [Andrey et Tarroux, 1998] Andrey, P. et Tarroux, P. (1998). Unsupervised segmentation of MRF modeled textured images using selectionist relaxation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 20(3) :252–262.
- [Arce et al., 2000] Arce, G., Paredes, J. L., et Mullan, J. (2000). Non-linear filtering for image analysis and enhancement. In Bovik, A. C., éditeur, *Handbook of Image and Video Processing*, pages 81–100. Academic Press, San Diego (USA).
- [Aubert et Blanc-Féraud, 1999] Aubert, G. et Blanc-Féraud, L. (1999). Some remarks on the equivalence between 2D and 3D classical snakes and geodesic active contours. *Intern. J. Computer Vision*, 34(1) :19–28.
- [Azencott et al., 1997] Azencott, R., Wang, J., et Younès, L. (1997). Texture classification using windowed Fourier filters. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 19(2) :148–153.
- [Babaud et al., 1986] Babaud, J., Witkin, A., Baudin, M., et Duda, R. (1986). Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8 :26–33.
- [Ballard, 1978] Ballard, D. (1978). Generalizing the Hough transform to detect a arbitrary shapes. *Pattern Recognition*, 13(2) :111–122.
- [Beaulieu et Goldberg, 1989] Beaulieu, J. et Goldberg, M. (1989). Hierarchy in picture segmentation : A stepwise optimization. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 11(2) :150–163.
- [Beckmann et Spizzichino, 1963] Beckmann, P. et Spizzichino, A. (1963). *The scattering of electromagnetic waves from rough surfaces*. Artech House, N. Y.
- [Bergen et Landy, 1991] Bergen, J. et Landy, M. (1991). Computational modeling of visual texture segregation. In Landy, M. S. et Movshon, J. A., éditeurs, *Computational models of visual processing*, pages 253–271. The MIT Press.
- [Berthier, 1993] Berthier, S. (1993). *Optique des milieux continus*. Polytechnica, Paris.
- [Bertrand et al., 1982] Bertrand, G., Queval, R., et Maître, H. (1982). Shape Interpolation by Fourier Descriptors with Application to Animation Graphics. *Signal Processing*, 4 :53–58.
- [Besl et Jain, 1986] Besl, P. J. et Jain, R. C. (1986). Invariant Surface Characteristics for 3D Object Recognition. *Computer Vision, Graphics and Image Processing*, 33 :33–80.

- [Bezdek, 1981] Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithm*. Plenum Press, New York.
- [Black et Rangarajan, 1996] Black, M. et Rangarajan, A. (1996). On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *Int. J. Comp. Vision*, 19(1) :58–93.
- [Blake et Isard, 1998] Blake, A. et Isard (1998). *Active Contours*. Springer Verlag.
- [Blake et Zisserman, 1987] Blake, A. et Zisserman, A. (1987). *Visual reconstruction*. MIT Press.
- [Born et Wolf, 1975] Born, M. et Wolf, E. (1975). *Principles of Optics*. Pergamon Press, Oxford (UK).
- [Bovik, 2000] Bovik, A. C. (2000). Basic gray-level image processing. In Bovik, A. C., éditeur, *Handbook of Image and Video Processing*, pages 21–36. Academic Press, San Diego (USA).
- [Boyer et Moore, 1977] Boyer, R. et Moore, J. (1977). A fast string searching algorithm. *Comm. of Ass. Comp. Mach.*, 20(10).
- [Brodatz, 1966] Brodatz, P. (1966). *Textures*. Dover Pub. (New York).
- [Burt, 1981] Burt, P. (1981). Fast filter transforms for image processing. *Comp. Graphics and Image Proc.*, 16 :20–51.
- [Calle et Montanvert, 1998] Calle, D. et Montanvert, A. (1998). Super-résolution inducing of an image. In *IEEE ICIP'98 - Int. Conf. on Image Processing*, volume III, pages 232–236.
- [Callet, 1998] Callet, P. (1998). *Couleur-lumière couleur matière*. Diderot Ed. Arts et Sciences.
- [Canny, 1986] Canny, J. (1986). A computational approach to edge detection. *IEEE Pattern Analysis and Machine Intelligence*, 8(6) :679–698.
- [Carmo, 1976] Carmo, M. P. D. (1976). *Differential Geometry of Curves and Surfaces*. Prentice Hall, Englewood Cliffs, New-Jersey.
- [Carron, 1995] Carron, T. (1995). *Segmentation d'images couleur dans la base Teinte Luminance Saturation : approche numérique et symbolique*. Thèse de doctorat, Université de Savoie.
- [Caselles et al., 1997] Caselles, V., Kimmel, R., et Sapiro, G. (1997). On geodesic active contours. *Int. J. Computer Vision*, 21(1) :61–79.
- [Chalmond, 2000] Chalmond, B. (2000). *Eléments de modélisation pour l'Analyse d'Images*, volume 33. Springer Verlag, Mathématiques et Applications.
- [Cham et al., 1998] Cham, F., Lam, F., et Zhu, H. (1998). Adaptive thresholding by variational methods. *IEEE Trans on Image Proc.*, 7(3) :468–473.
- [Chang et Kuo, 1993] Chang, T. et Kuo, C. (1993). Texture analysis and classification with tree structured wavelet transform. *IEEE Trans. on Image Processing*, 2(4) :429–441.
- [Chang et Li, 1994] Chang, Y. et Li, X. (1994). Adaptive image region growing. *IEEE Trans on Image Proc.*, 3 :868–872.
- [Chellappa et Kashyap, 1985] Chellappa, R. et Kashyap, R. (1985). Texture synthesis using 2d non-causal AR models. *IEEE Trans on Acoustic Speech and Signal Proc.*, 33 :194–293.
- [Chen et Pavlidis, 1979] Chen, P. et Pavlidis, T. (1979). Segmentation by textures using a cocurrence matrix and a split and merge algorithm. *Comp. graph. and Image Proc.*, 10 :172–182.
- [Chen et al., 1991] Chen, S., Lin, W., et Chen, C. (1991). Split and merge image segmentation based on localized image features and statistical tests. *Comp. Vision Graphics and Image Proc : Graph. Models and Image Proc.*, 53 :457–475.
- [Chesnaud et al., 1999] Chesnaud, C., Réfrégier, P., et Boulet, V. (1999). Statistical region snake-based segmentation adapted to different physical noise models. *IEEE PAMI*, 21(11) :1145–1157.
- [Cocquerez et Philipp, 1996] Cocquerez, J. et Philipp, S. (1996). *Analyse d'Images : Filtrage et Segmentation*. Masson.

- [Coggins et Jains, 1985] Coggins, J. et Jains, A. (1985). A spatial filtering approach to texture analysis. *Pattern Recognition Letters*, 3(3) :195–203.
- [Cohen, 1991] Cohen, L. D. (1991). On active contour models and balloons. *Computer Vision, Graphics, and Image Processing : Image Understanding*, 53(2) :211–218.
- [Cook et Torrance, 1982] Cook, R. et Torrance, K. (1982). A reflectance model for computer graphics. *ACM Transactions on Graphics*, 1(1) :7–24.
- [Coster et Chermant, 1985] Coster, M. et Chermant, J. L. (1985). *Précis d'analyse d'images*. Editions du CNRS.
- [Cross et Jain, 1983] Cross, G. et Jain, A. (1983). Markov random fields texture models in image analysis. *J. Applied Statistics*, 5(1) :25–39.
- [Darell et al., 1990] Darell, T., Sclaroff, S., et Pentland, A. (1990). Segmentation by minimum description length. In *Int. Conf. on Computer Vision (Osaka)*, pages 112–116.
- [Daubechies, 1992] Daubechies, I. (1992). *Wavelets*. S.I.A.M., Philadelphia.
- [de Valois et al., 1982] de Valois, R. L., Yund, E., et Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22 :531–544.
- [Demigny et Kamlé, 1997] Demigny, D. et Kamlé, T. (1997). A discrete expression of Canny's criteria for step edge detector. Performances evaluation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 19(11) :1199–1211.
- [Deriche, 1987] Deriche, R. (1987). Using canny's criteria to derive recursively implemented optimal edge detector. *Int. J. Computer Vision*, 1(2) :167–187.
- [Derin et Elliott, 1987] Derin, H. et Elliott, H. (1987). Modeling and segmentation of noisy and textured images using gibbs random fields. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 9 :39–55.
- [diZenko., 1986] diZenko., S. (1986). A note on the gradient of multi-image. *Computer Vision, Graphics and Image Proc.*, 33 :116–125.
- [Dodgson, 1997] Dodgson, N. (1997). Quadratic interpolation for image resampling. *IEEE Trans. on Image Processing*, 6(9) :1322–1326.
- [Dubuisson, 1990] Dubuisson, B. (1990). *Diagnostic et reconnaissance des formes*. Hermès, Paris.
- [Duda et Hart, 1973] Duda, R. et Hart, P. (1973). *Pattern classification and scene analysis*. Interscience, New-York.
- [Duntley, 1942] Duntley, S. (1942). The optical properties of diffusing materials. *J. Optical Society of America*, 32(2).
- [Ell et Sangwine, 2000] Ell, T. et Sangwine, S. (2000). Hypercomplex Wiener-Khintchine theorem with application to color image correlation. In IEEE, éditeur, *ICIP 2000*, volume II, pages 792–796.
- [Evans et al., 2000] Evans, C., Sangwine, S. J., et Ell, T. A. (2000). Hypercomplex color-sensitive smoothing filters. In IEEE, éditeur, *ICIP 2000*, volume 1, pages 541–544, Vancouver.
- [Fairfield, 1990] Fairfield, J. (1990). Toboggan contrast enhancement for contrast segmentation. In *IAPR 10th Int. Conf. on Pattern Recognition*, pages 712–716, Atlantic City.
- [Faugeras, 1976] Faugeras, O. (1976). *Digital Color Image Processing and Psychophysics within the framework of a human visual model*. Thèse de doctorat, University of Utah.
- [Faugeras, 1993] Faugeras, O. (1993). *Three-Dimensional Computer Vision : A Geometric Viewpoint*. MIT Press, Cambridge, MA.
- [Fischler et Bolles, 1981] Fischler, M. et Bolles, R. (1981). Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography. *Communication of the ACM*, 24 :381–385.
- [Foley et al., 1995] Foley, J., van Damm, A., Feiner, S. K., Hughes, J. F., et Philips, R. L. (1995). *Introduction à l'infographie*. Addison-Wesley, Paris.

- [Freeman, 1961] Freeman, H. (1961). On the encoding of arbitrary geometric configurations. *IEEE Trans on Electr. Comput.*, 10 :260–268.
- [Freeman, 1977] Freeman, H. (1977). Computer processing of line drawing images. *ACM Computer Survey*, 6 :57–97.
- [Fua et Leclerc, 1990] Fua, P. et Leclerc, Y. G. (1990). Model driven edge detection. *Machine Vision and Applications*, 3(1) :45–56.
- [Fukunaga, 1990] Fukunaga, K. (1990). *Introduction to statistical pattern recognition*. Academic Press.
- [Gagalowics et Tournier-Lasserve, 1986] Gagalowics, A. et Tournier-Lasserve, C. (1986). Third order model for non homogeneous natural textures. In *Int. Conf. Pattern Recognition*, volume 86, pages 409–411, Vancouver.
- [Gallice, 2002] Gallice, J. (2002). *Images de profondeur*. Hermes, Paris.
- [Garello, 2001] Garello, R. (2001). *Analyse de signaux bidimensionnels*. Collection IC2. Hermès, Paris.
- [Geman et Geman, 1984] Geman, S. et Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 12(7) :721–741.
- [Giraudon, 1987] Giraudon, G. (1987). Chaînage efficace de contours. Rapport technique 605, INRIA.
- [Goshtasby et al., 1990] Goshtasby, A., Cheng, F., et Barsky, B. A. (1990). B-spline curves and surfaces viewed as digital filters. *Computer Vision Graphics and Image Processing*, 52 :264–275.
- [Grimson, 1981] Grimson, D. (1981). *From images to surface*. MIT Press, Cambridge USA.
- [Grossberg et Mingolla, 1985] Grossberg, S. et Mingolla, E. (1985). Neural dynamics of perceptual grouping : textures, boundaries and emergent segmentations. *perception Psychophys.*, 38 :141–171.
- [Guth, 1991] Guth, S. (1991). Model for color vision and light adaption. *J. Opt. Soc. Am. A*, 8(6).
- [Guzman, 1968] Guzman, A. (1968). Decomposition of visual scene into three-dimensional bodies. In *Proc J. computer Conf. AFIPS*, volume 33, pages 291–304.
- [Haralick et Shapiro, 1985] Haralick, R. M. et Shapiro, S. G. (1985). Survey : Image segmentation techniques. *Comp. Vision Graphics and Image Processing*, 29 :100–132.
- [Heath et al., 1997] Heath, M., Sarkar, S., Sanocki, T., et Bowyer, K. (1997). A robust visual method for assessing the relative performance of edge detection algorithms. *IEEE Pattern Analysis and Machine Intelligence*, 19(12) :1338–1359.
- [Heitz et al., 1994] Heitz, F., Pérez, P., et Boutheimy, P. (1994). Multiscale minimization of global energy function in some visual recovery problems. *Comp. Vision Graphics and Image Proc : Image Understanding*, 59(1).
- [Horowitz et Pavlidis, 1976] Horowitz, S. et Pavlidis, T. (1976). Picture segmentation by a tree traversal algorithm. *J. Ass. Comp. Mach.*, 23(2) :368–388.
- [Hu, 1962] Hu, M. K. (1962). Visual pattern recognition by moment invariants. *IRE Trans on Information Theory*, IT-8(2) :179–187.
- [Hu et Fahmy, 1992] Hu, R. et Fahmy, M. (1992). Texture segmentation based on a hierarchical Markov random field. *Signal Processing*, 26 :285–305.
- [Hubel et Wiesel, 1969] Hubel, D. et Wiesel, T. (1969). Anatomical demonstration of columns in the monkey striate cortex. *Nature*, 221 :747–750.
- [Huber, 1981] Huber, P. (1981). *Robust Statistics*. John Wiley and Sons, New York.
- [Hueckel, 1971] Hueckel, M. (1971). An operator which locates edges in digitized pictures. *J. Ass. Comp. Mach.*, 18(1) :113–125.
- [Jain et Farrokhnia, 1991] Jain, A. et Farrokhnia, F. (1991). Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, 24(12) :1167–1186.
- [Julesz, 1971] Julesz, B. (1971). *Foundations of Cyclopean Perception*. The University of Chicago Press, Chicago.

- [Karabassis et Spetsakis, 1995] Karabassis, E. et Spetsakis, M. (1995). An analysis of image interpolation, differentiation and reduction using local polynomial fits. *Graphical Models and Image Processing*, 57(3) :183–196.
- [Kass et al., 1988] Kass, M., Witkin, A., et Terzopoulos, D. (1988). Active contour models. *Int. J. Comp. Vision*, 1 :321–331.
- [Keren et al., 1990] Keren, D., Marcus, R., Werman, M., et Peleg, S. (1990). Segmentation by minimum length encoding. In *10th Int. Conf. on Pattern recogn. (Atlantic City)*, pages 681–683.
- [Kerker, 1969] Kerker, M. (1969). *The scattering of light and other electromagnetic radiation*. Academic Press, New York.
- [Kervrann et Heitz, 1995] Kervrann, C. et Heitz, F. (1995). A Markov random field model-based approach to unsupervised texture segmentation using local and global statistics. *IEEE Trans Image Processing*, 4(6) :856–862.
- [Kim et Yaroslavski, 1986] Kim, V. et Yaroslavski, L. (1986). Rank algorithms for picture processing. *Comp. Vision Graphics and Image Proc.*, 35 :234–258.
- [Kimia et al., 1992] Kimia, B. B., Tannenbaum, A., et Zucker, S. W. (1992). On the evolution of curves via a function of curvature. I. the classical case. *Journal of Mathematical Analysis and Applications*, 163(2) :438–458.
- [Kittler et al., 1985] Kittler, J., Illingworth, J., et Foglein, J. (1985). Threshold selection based on a simple image statistic. *Comp. Vision Graphics and Image Proc*, 30 :125–147.
- [Koenderink, 1984] Koenderink, J. (1984). The structure of images. *Biological Cybernetics*, 50 :363–370.
- [Kotropoulos et al., 1997] Kotropoulos, C., Gabrani, M., et Pitas, I. (1997). Adaptive multichannel L-filters with structural constraints. In *IEEE - Int. Work on Non Linear Signal and Image Processing, NSIP'97*, pages 181–183, Michigan.
- [Kunduri et al., 1999] Kunduri, S., Maître, H., Roux, M., Demigny, D., et Karabernou, M. (1999). An improved method for isotropic edge orientation estimation . In *SPIE, éditeur, Europto '99*, Florence.
- [Laine et Fan, 1993] Laine, A. et Fan, J. (1993). Texture classification by wavelet packet signature. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 15 :1186–1190.
- [Lakshmanan et Derin, 1989] Lakshmanan, S. et Derin, H. (1989). Simultaneous parameter estimation and segmentation of gibbs random fields using simulated annealing. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 11(8) :799–813.
- [Leclerc, 1989] Leclerc, Y. (1989). Constructing simple stable descriptions for image partitioning. *Int. J. of Computer Vision*, 3(1989) :73–102.
- [Lee et al., 1998] Lee, K.-M., Meer, P., et Park, R.-H. (1998). Robust adaptive segmentation of range image. *IEEE T. Pattern Analysis and Machine Intelligence*, 20(2) :200–205.
- [Lee, 1998] Lee, T. (1998). Segmenting images corrupted by correlated noise. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 20(5) :481–492.
- [Linde et al., 1980] Linde, Y., Buzo, A., et Gray, R. (1980). An algorithm for vector quantizer designs. *IEEE Trans on Communications*, C-28 :84–95.
- [Lowrance et Wagner, 1975] Lowrance, R. et Wagner, R. (1975). An extension of the string to string correction problem. *J. Ass. Comp. Mach.*, 22(2) :177–183.
- [Lu et Fu, 1978] Lu, S. et Fu, K. (1978). A syntactic approach to texture analysis. *Computer Graphics and Image Proc.*, 7(3) :303–330.
- [Lu et Fu, 1979] Lu, S. et Fu, K. (1979). Stochastic tree grammar inference for texture synthesis and discrimination. *Computer Graphics and Image Proc.*, 9 :234–245.
- [Maeland, 1988] Maeland, E. (1988). On the comparison of interpolation methods. *IEEE Trans. on Medical Imaging*, 7(3) :213–217.

- [Maître, 1985] Maître, H. (1985). Un panorama de la transformation de Hough. *Traitement du Signal*, 2(4) :305–317.
- [Maître, 2001] Maître, H. (2001). *Le traitement des images de Radar à Synthèse d’Ouverture (Ouvrage collectif)*. Hermes , collection IC2, Paris.
- [Mallat, 1989] Mallat, S. (1989). A theory of multiresolution signal decomposition : the wavelet representation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 11 :674–693.
- [Manjunath et Chellapa, 1991] Manjunath, B. et Chellapa, R. (1991). Unsupervised texture segmentation using Markov random field models. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 13(5) :478–482.
- [Marion, 1997] Marion, A. (1997). *Acquisition et Visualisation des Images*. Eyrolles, Paris.
- [Marr et Hildreth, 1980] Marr, D. et Hildreth, E. (1980). Theory of edge detection. In *Proc. Royal Soc. London*, volume 207 of *B*, pages 187–217.
- [Martelli, 1972] Martelli, A. (1972). Edge detection using heuristic search methods. *Comp. Graph. and Image Proc.*, 1 :169–182.
- [Mayhew et Frisby, 1978] Mayhew, J. et Frisby, J. P. (1978). Texture discrimination and Fourier analysis in the human visual system. *Nature*, 275 :438–439.
- [Meer et al., 1991] Meer, P., Mintz, D., Rosenfeld, A., et Kim, D. (1991). Robust regression methods for computer vision : A review. *Int. J. Computer Vision*, 6 :59–70.
- [Melamed, 1963] Melamed, N. (1963). Optical properties of powder. *J. of Applied Physics*, 34 :560–570.
- [Miclet, 1984] Miclet, L. (1984). *Méthodes structurelles pour la reconnaissance des formes*. Eyrolles, Paris.
- [Modestino et Fries, 1977] Modestino, J. et Fries, R. (1977). Edge detection in noisy images using recursive digital filters. *Comp. Graph. and Image Proc.*, pages 409–433.
- [Mohr et al., 1994] Mohr, R., Boufama, B., et Brand, P. (1994). Accurate projective reconstruction. In Mundy, J. L., Zisserman, A., et Forsyth, D., éditeurs, *Application of invariance in computer vision*, volume LNCS 825, pages 297–316. Springer-Verlag, Berlin.
- [Montanari, 1971] Montanari, U. (1971). On the optimal detection of curves in noisy images. *Com. ACM*, 14 :335–345.
- [Mumford et Shah, 1989] Mumford, D. et Shah, J. (1989). Optimal approximations by piecewise smoothed functions and associated variational problems. *Com. Pure and Applied Math.*, XLII :577–685.
- [Natsuyama, 1979] Natsuyama, M. N. T. (1979). Edge preserving smoothing. *Computer Graphics and Image Processing*, 9 :394–407.
- [Nayatani et al., 1990] Nayatani, Y., Takahama, K., Sobagaki, H., et Hashimoto, K. (1990). Color-appearance model and chromatic-adaption transform. *Color Research and Application*, 15(14) :210–221.
- [Nevatia et Babu, 1980] Nevatia, R. et Babu, K. (1980). Linear feature extraction and description. *Comp. Graphics and Image Proc.*, 13(3) :257–269.
- [Nicodemus et al., 1977] Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., et Limperis, T. (1977). *Geometrical considerations and nomenclature for reflectance*. Number 160. National Bureau of Standards, Washington.
- [Nicolas, 1999] Nicolas, J. M. (1999). Systèmes d’imagerie satellitale, cours de 3a., Rapport technique, ENST, Paris.
- [Ohlander et al., 1978] Ohlander, R., Price, K., et Reddy, D. (1978). Picture segmentation using a recursive region splitting method. *Comp. Graphics and Image Proc.*, 8 :313–333.
- [Olivier et al., 1997] Olivier, C., Jouzel, F., ElMatouat, A., et Courtellemont, P. (1997). Un nouveau critère pour la sélection de l’ordre d’un modèle. In *GRETSI’97*, pages 451–454, Grenoble.

- [Osher et Sethian, 1988] Osher, S. et Sethian, J. (1988). Front propagating with curvature dependant speed : Algorithm based on the Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79 :12–49.
- [Palmer et al., 1996] Palmer, P., Dabis, H., et Kittler, J. (1996). A performance measure for boundary detection algorithms. *Comp. Vis. and Image Understanding*, 63(3) :476–494.
- [Pentland, 1984] Pentland, A. (1984). Fractal based description of natural scenes. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 6 :661–674.
- [Perez et Heitz, 1991] Perez, P. et Heitz, F. (1991). Une approche multiéchelle à l’analyse d’images par champs markoviens. In *8ième Congres AFCET RFIA*, pages 1003–1015.
- [Perona et Malik, 1987] Perona, P. et Malik, J. (1987). Scale space and edge detection using anisotropic diffusion. In *IEEE Workshop on Computer Vision*, pages 16–22, Miami. IEEE.
- [Perona et Malik, 1990] Perona, P. et Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7) :629–639.
- [Petrou et Kittler, 1991] Petrou, M. et Kittler, J. (1991). Optimal edge detectors for ramp edges. *IEEE-Pattern Analysis and Machine Intelligence*, 13(5) :1483–1491.
- [Phillips et Wilson, 1984] Phillips, G. C. et Wilson, H. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *JOSA - A*, 1 :226–232.
- [Phong, 1975] Phong, B. T. (1975). Illumination for computer generated pictures. *Communications of the ACM*, 6(18) :311–317.
- [Pitas et Venetsanopoulos, 1990] Pitas, I. et Venetsanopoulos, A. (1990). *Nonlinear digital filters*. Kluwer Academic.
- [Portilla et Simoncelli, 2000] Portilla, J. et Simoncelli, E. (2000). A parametric texture model based on joint statistics of complex wavelets coefficients. *International Journal of Computer Vision*, 40(1) :49–71.
- [Pratt, 1978] Pratt, W. (1978). *Digital image processing*. Wiley Interscience, New-York.
- [Pérez, 1991] Pérez, J. P. (1991). *Optique géométrique, ondulatoire et polarisation*. Masson, Paris.
- [Princen et al., 1992] Princen, J., Illingworth, J., et Kittler, J. (1992). A formal definition of the Hough transform : properties and relationships. *J. Math. Image and Vision*, 1 :153–168.
- [Randen et Husøy, 1999] Randen, T. et Husøy, J. (1999). Filtering of texture classification : a comparative study. *IEEE Pattern Analysis and Machine Intelligence*, 21(4) :291–310.
- [Reed et du Buf, 1993] Reed, T. et du Buf, H. (1993). A review of recent texture segmentation and feature extraction techniques. *Comp. Vision Graphics and Image Proc : Image Undersrtand.*, 57(3) :359–372.
- [Rissanen, 1984] Rissanen, J. (1984). *IEEE Tr. Inform. Theory*, 30 :629–636.
- [Rissanen, 1987] Rissanen, J. (1987). Minimum description length principle. *Encyclopedia of Statistical Sciences*, 3 :523–527.
- [Rissanen, 1989] Rissanen, J. (1989). *Stochastic complexity in Statistical Inquiry*. World Scientific, Singapore.
- [Robinson, 1977] Robinson, G. (1977). Edge detection by compass gradient masks. *Computer Graphics and Image Processing*, 6 :492–501.
- [Rousseeuw et Leroy, 1987] Rousseeuw, P. et Leroy, A. (1987). *Robust regression and outlier detection*. John Wiley, New-York.
- [Rudin et Osher, 1994] Rudin, L. et Osher, S. (1994). Total variation based image restoration with free local constraints. In *IEEE ICIP’94*, volume I, pages 31–35, Dallas (TX).
- [Rudin et al., 1992] Rudin, L., Osher, S., et Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4) :259–268.
- [Sahoo et al., 1988] Sahoo, P., Soltani, S., Wong, A., et Chen, Y. (1988). A survey of thresholding techniques. *Comp. Vision Graphics and Image Proc.*, 41(1) :233–260.

- [Saint-Marc et al., 1989] Saint-Marc, P., Chen, J., et Medioni, G. (1989). Adaptive smoothing : A general tool for early vision,. In *CVPR'89*, pages 618–624.
- [Saito et Coifman, 1995] Saito, N. et Coifman, R. (1995). Local discriminant bases and their applications. *J. Mathematical Imaging and Vision*, 5(4) :337–358.
- [Sangwine, 1998] Sangwine, S. (1998). Colour image edge-detector based on quaternion convolution. *Electronics Letters*, 34(10) :969–971.
- [Sangwine, 1996] Sangwine, S. J. (1996). Fourier transforms of colour images using quaternions or hypercomplex numbers. *Electronics Letters*, 32(21) :1979–1980.
- [Saporta, 1978] Saporta, G. (1978). *Théorie et Méthodes Statistiques*. Princeton University Press.
- [Saporta, 1990] Saporta, G. (1990). *Probabilités, analyse de données et statistiques*. Editions Technip, Paris.
- [Sarkar et Boyer, 1991] Sarkar, S. et Boyer, K. (1991). Optimal infinite impulse response zero crossing based edge detectors. *Comp. Vision Graphics and Image Proc Image Understanding*, 54(2) :224–243.
- [Schmitt et Mattioli, 1994] Schmitt, M. et Mattioli, J. (1994). *Morphologie Mathématique*. Masson.
- [Serra, 1982] Serra, J. (1982). *Image analysis and mathematical morphology*. Academic Press, New York.
- [Sethian, 1996] Sethian, J. (1996). *Level-set methods*. Cambridge University Press.
- [Shanmugam et Green, 1979] Shanmugam, K. et Green, J. (1979). An optimal frequency domain filter for edge detection in digital pictures. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1 :37–49.
- [Shen et Castan, 1986] Shen, J. et Castan, S. (1986). An optimal linear operator for edge detection. In *Proceeding CVPR Miami*.
- [Sklansky, 1978] Sklansky, J. (1978). On the Hough technique for curve detection. *IEEE Trans. Comp.*, 27(100) :923–926.
- [Spacek, 1986] Spacek, L. (1986). Edge detection and motion detection. *Image and Vision Computing*, 4 :43–56.
- [Suk et Chung, 1993] Suk, M. et Chung, S. (1993). A new image segmentation technique based on partition mode test. *Pattern Recognition*, 16(5) :469–480.
- [Sève, 1996] Sève, R. (1996). *Physique de la couleur : de l'apparence colorée à la technique colorimétrique*. Masson - Physique fondamentale et appliquée, Paris.
- [Tagare et Figueiredo, 1985] Tagare, H. et Figueiredo, R. (1985). Order filters. *Proc of the IEEE*, 73(1) :163–165.
- [Terzopoulos, 1983] Terzopoulos, D. (1983). Multilevel computational process for visual surface reconstruction. *Comp. Vision Graph. and Image Proc.*, 24 :52–96.
- [Torrance et Sparrow, 1967] Torrance, K. et Sparrow, E. (1967). Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.*, 57(9) :1105–1114.
- [Torre et Poggio, 1986] Torre, V. et Poggio, T. (1986). On edge detection. *IEEE Pattern Analysis and Machine Intelligence*, 8 :147–153.
- [Unser, 1986] Unser, M. (1986). Local linear transforms for texture measurement. *Signal Processing*, 11(1) :61–79.
- [Unser, 1995] Unser, M. (1995). Texture classification and segmentation using wavelet frame. *IEEE Trans on Image Processing*, 4(11) :1549–1560.
- [Unser, 1999] Unser, M. (1999). Splines. a perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6) :22–38.
- [Unser et al., 1993a] Unser, M., Aldroubi, A., et Eden, M. (1993a). B-spline signal processing - Part i. *IEEE Trans. on Signal Processing*, 41(2) :821–833.
- [Unser et al., 1993b] Unser, M., Aldroubi, A., et Eden, M. (1993b). B-spline signal processing - Part II. *IEEE Trans. on Signal Processing*, 41(2) :834–848.

- [VanGool et al., 1985] VanGool, L., Dewaele, P., et Oosterlinck, A. (1985). Textural analysis anno 1983. *Comp. Graph. and Image Processing*, 29 :336–357.
- [Vickers et Modestino, 1982] Vickers, A. et Modestino, J. (1982). A maximum likelihood approach to texture classification. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 4 :61–68.
- [Wagner et Fisher, 1974] Wagner, R. et Fisher, M. (1974). The string to string correction problem. *J. Ass. Comp. Mach.*, 21(1) :168–173.
- [Wang, 1998] Wang, J.-P. (1998). Stochastic relaxation on partitions with connected components and its application to image segmentation. *IEEE Pattern Analysis and Machine Intelligence*, 20(6) :619–636.
- [Wertheimer, 1944] Wertheimer, M. (1944). Gestalt Theorie. *Social Research*, 11 :78–99.
- [Weszka, 1978] Weszka, J. (1978). A survey of threshold selection techniques. *Comp. Graphics and Image Proc.*, 7 :259–265.
- [Witkin, 1983] Witkin, A. (1983). Scale-space filtering. In *International Joint Conference on Artificial Intelligence*, pages 1019–1023.
- [Won et Derin, 1992] Won, C. et Derin, H. (1992). Unsupervised segmentation of noisy and textured images using Markov random fields. *Comp. Vision Graphics and Image Proc : Graphical Models and Image Processing*, 54(4) :308–328.
- [Wu et Maître, 1992] Wu, Y. et Maître, H. (1992). Smoothing speckled synthetic aperture radar images by using maximum homogeneous region filters. *Optical Engineering*, 31(8) :1785–1792.
- [Zamperoni, 1996] Zamperoni, P. (1996). Feature extraction. In Maître, H. et Zinn-Justin, J., éditeurs, *Progress in Picture Processing, Les Houches 92*, volume LVIII, pages 119–182. North-Holland.
- [Zhang et Modestino, 1990] Zhang, J. et Modestino, J. (1990). A model fitting approach to cluster validation with applications to stochastic model-based image segmentation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 4 :1009–1017.
- [Zhao et al., 2000] Zhao, H.-K., Osher, S., Merriman, B., et Kang, M. (2000). Implicit non-parametric shape reconstruction from unorganized data using variational level set method. *Comp. Vision Graphics and Image Proc*, 80(3) :295–314.
- [Zhu et Yuille, 1996] Zhu, S. et Yuille, A. (1996). Region competition : Unifying snakes, region growing and Bayes/MDL for multiband image segmentation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 18 :884–900.

Index

- équation
 - de diffusion, 6
 - de Perona et Malik, 6
- adaptation chromatique, 124
- algorithme
 - de la corde, 77
- analyse
 - en composantes principales, 37
- aplatissement, 54
- approximations polygonales, 74
- attache aux données, 38
- autorégressif, 59

- B-splines, 11
- balayage, 127
- BDRF, 110
- Bhattacharyya, 37
- binormale, 97
- birapport, 90
- blanc de référence, 117
- Blinn, 110
- boîte englobante, 66
- boussoles directionnelles, 21
- BV, 7

- C-moyennes floues, 76
- Canny, 22
- cercle osculateur, 97
- chaînes de Freeman, 68
- champ
 - de Markov, 38, 60
- chatoiement, 108
- chi², 42
- chroma, 116, 118
- chrominance, 117
- CIE, 114
- CIELab, 118
- clarté, 117
- coefficient de variation, 54
- colinéation, 87

- composantes trichromatiques, 113, 114
- conditionnement, 133
- conique, 89
- coniques, 77
- contour, 16
- contours actifs, 26
- cooccurrence, 55, 61
- coordonnées homogènes, 87
- courbes d'excitation, 115
- courbure, 97, 100
- courbure
 - gaussienne, 100
 - moyenne, 100
 - principale, 100
- critère
 - de Neyman Pearson, 36
- critère d'Akaike, 37
- croissance de région, 40
- cubique, 11

- décision bayésienne, 36
- décomposition en valeurs singulières, 131
- détection de contours, 15
- Deriche, 23
- descripteurs de Fourier, 72
- diffusion, 6
- diffusion anisotrope, 6
- distance
 - interquartile, 39, 54
- distance frustrée, 127
- DOG, 21
- droite
 - à l'infini, 88
 - projective, 88

- ensemble
 - de niveau, 29
- entrelacement, 127
- entropie, 39
- espace Lab, 117
- espace projectif, 87, 88, 96

- espace XYZ, 116
- estimation
 - robuste, 75, 76
- fausse alarme, 36
- fermeture , 19
 - de contours, 30
- filtrage par fenêtre maximale, 9
- filtre
 - adapté, 4
 - de diffusion anisotrope, 19
 - de Nagao, 9
 - de Perona et Malik, 6
 - de rang, 5
 - de Rudin, Osher et Fatemi, 8
 - de Saint Marc, 8
 - de Wiener, 4
 - de Wu, 9
 - LMS, 76
 - médian, 76
 - morphologique, 5
 - sigma, 8
 - Toboggan, 9, 19
- filtres
 - à phase nulle, 18
 - de Gabor, 57
 - miroirs, 58
- Fisher-Snedecor, 42
- fonction
 - d'autocorrélation, 55
- fonctions colorimétriques, 114
- fonctions de mélange, 115
- forme fondamentale, 99
- formes fondamentales, 98
- Freeman, 68
- Frenet, 97
- Frenet-Serret, 98
- Fuzzy-C means, 76

- géométrie différentielle, 97
- géométrie projective, 87
- Gabor, 57
- Gestalt Theorie, 50
- gradient , 18
- Graduated Non Convexity, 48
- granularité, 108, 110
- graphe
 - d'adjacence, 42
- groupement perceptuel, 50
- growing, 40

- Guzman, 66

- Hilbert, 127
- homogénéité, 1
- Hough, 80
- Hueckel, 21
- hystérésis, 19

- ICM, 61
- image d'étiquettes, 63
- indice d'allongement, 65
- interpolation, 11
- invariant, 90
- invariant
 - projectif, 90
- invariants
 - affines, 91

- Julesz, 54

- Karhunen Loeve, 37
- Kirsch, 21
- kurtosis, 54

- labels, 63
- lagrangien, 138
- Lambert, 110
- laplacien , 20
 - d'une gaussienne, 21
- LBG, 128
- level sets, 29
- lignes
 - de crête, 20
- linéarité, 1
- LOG, 20
- log-vraisemblance, 44
- loi de Lambert, 110
- look-up table, 1

- M-estimateurs, 76
- mélange, 76
- métamères, 112
- Mahalanobis, 37
- Marr, 21
- Martelli, 30
- masques adaptés, 21
- Maximum A Posteriori, 38
- maximums du gradient, 16
- mean curvature, 7
- membrane, 46
- merging, 41

- modèles
 - géodésiques actifs, 30
- moindres carrés, 134
- moment
 - invariant, 66
 - normé, 66
- Montanari, 30
- mouvement par courbure moyenne, 7
- moyenne frustrée, 127
- Mumford et Shah, 45

- Netto, 127
- Nevatia et Babu, 21
- non-linéarité, 1
- NTSC, 122
- nuées dynamiques, 76

- ondelettes, 58
- ordre, 125
- ordre canonique, 126
- ordre partiel, 125
- ordre total, 125
- orientation, 18

- PAL, 122
- partage de région, 40
- partition, 39
- Peano, 127
- plan à l'infini, 89
- plaques minces, 46
- point
 - ombilic, 100
- point à l'infini, 87
- Potts, 46
- poursuite, 19
- pré-ordre, 125
- prédicat, 39
- prétraitements, 1
- Prewitt, 21
- primal sketch, 20
- probabilité de détection, 36
- processus
 - ARMA, 59
 - de bords, 39, 61
- programmation dynamique, 30
- pseudo-inverse, 133
- pyramide, 42

- QMF, 58
- quad-tree, 42

- quadrique, 89
- quantification adaptée, 54
- quantification vectorielle, 128
- quaternions, 94, 130

- réflexion bidirectionnelle, 110
- régularisation, 38
- RanSac, 76
- rapport de vraisemblance, 36
- rapport iso-périmétrique, 65
- recalage rigide, 96
- recuit simulé, 61
- représentation canonique, 93
- Rodriguez, 94

- Saint Marc, 8
- saturation, 116
- SECAM, 122
- Shen et Castan, 24
- Smirnov-Kolmogorov, 61
- snakes, 26
- speckle, 108, 110
- sphère de Gauss, 99
- spline, 78
- splines, 11
- splitting, 40
- statistiques du premier ordre, 54
- stratification, 90, 91
- Student, 42
- Suk, 42
- sur-segmentations, 42
- SVD, 131

- table de transcodage, 1
- tavelure, 108
- teinte, 116, 118
- template matching, 21
- tests statistiques, 42
- texture, 49
- théorème
 - de Netto, 127
 - de Rodriguez, 94
- théorie bayésienne de la décision, 36
- Torrance-Sparrow, 110
- transformée
 - en Z, 12
- transformation
 - affine, 91
 - de Hough, 80
- tri planaire, 63

trièdre de Frenet, 97
triangle de Maxwell, 114
trivariance, 112
TV, 7

Wiener-Kinchine, 55
Wilcoxon, 42

Yule-Walker, 59

zéro du laplacien, 20
zigzag, 127