



Institut MONTEFIORE
Service de Télécommunications et d'Imagerie

ELEN016-0
Traitement numérique des images

Professeur Marc VAN DROOGENBROECK

Septembre 2007 (version 4.80)

Remerciements

Merci à J.-M. WAGNER pour son aide précieuse et pour son dévouement.

Je tiens également à remercier Frédéric NOO et Yves DAOUST pour leur exposé et leurs notes.

Prof. Marc Van Droogenbroeck

Modalités d'examen

Seule la matière couverte dans ces notes constitue la matière d'examen.

L'examen est écrit. Il se déroulera à livre ouvert ; toutes les notes (cours, exercices, encyclopédie, etc) sont autorisées.

Les notes de cours sont disponibles (sous forme *HTML*) en ligne à l'adresse suivante :

<http://www.ulg.ac.be/telecom>

En consultant cette version en ligne, vous aurez la possibilité de voir certaines images en couleur (notamment dans le chapitre introductif) mais vous profiterez également de tous les hyperliens créés dans ce document.

Notations

Cadre analytique:

\mathbb{N}	$\mathbb{N} = \{0, 1, 2, \dots\}$
\mathbb{Z}	$\mathbb{Z} = \{\dots - 2, -1, 0, 1, 2, \dots\}$
\mathbb{R}	Ensemble des réels
$\overline{\mathbb{R}}$	Ensemble des réels complété, $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$
\mathbb{Z}^n	$\mathbb{Z}^n = \mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z}$
\mathbb{R}^n	$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$
\mathcal{E}	Référentiel
$\mathcal{P}(\mathcal{E})$	Ensemble des parties de \mathcal{E}

Constantes:

j	Vecteur unité de l'axe imaginaire $j = \sqrt{-1}$
\emptyset	Ensemble vide

Variables et fonctions:

x	Variable de position horizontale
y	Variable de position verticale
z	Variable de position de profondeur
t	Variable de temps continue
f	Variable fréquence
ω	Variable de pulsation $\omega = 2\pi f$
Δx	Pas d'échantillonnage en x
Δy	Pas d'échantillonnage en y
$x(t)$	Fonction à valeurs continues, définie pour tout temps t
$f(x, y)$	Fonction représentant une image
$f(x, y, t)$	Fonction représentant une séquence vidéo
$f[x, y]$	Fonction échantillonnée spatialement
$\delta(\cdot)$	Fonction delta de DIRAC

Matrices:

\underline{A}	Matrice
-----------------	---------

Notations fonctionnelles:

$ a $	Valeur absolue
\vec{a}	Vecteur
$\ \vec{a}\ $	Norme de \vec{a}
$\vec{a} \cdot \vec{b}$	Produit scalaire de \vec{a} et \vec{b}
$\nabla^2 \varphi$	Laplacien de φ
$\nabla \varphi$	Gradient de φ
X^*	Complexe conjugué de X
$Re(a)$	Partie réelle de a
$Im(a)$	Partie imaginaire de a
$\mathcal{X}(f)$	Transformée de FOURIER du signal $x(t)$
\rightleftharpoons	Correspondance entre un signal et sa transformée
\otimes	Convolution
$hist(a)$	Histogramme de la fonction a

Fonctions stochastiques:

$p(A)$	Probabilité de A
μ_X	Espérance mathématique du signal X
σ_X^2	Variance de X

$R_{XX}(\tau)$	Fonction d'autocorrélation du processus aléatoire stationnaire $X(t)$
$\gamma_X(f)$	Densité spectrale de puissance de $X(t)$
<i>Cadre ensembliste:</i>	
$x \in X$	x appartient à X
X^c	Complémentaire de X
\check{X}	Transposé de X
X_b	Translaté de X par b
$X \subseteq Y$	X inclus ou égal à Y
$X \supseteq Y$	X comprend ou est égal à Y
$X \cup Y$	X union Y
$X \cap Y$	X inter Y
$X \setminus Y$	X moins Y
$\partial(X)$	Frontière de X
$co(X)$	Enveloppe convexe euclidienne de X
$\overline{co}(X)$	Enveloppe convexe euclidienne fermée de X
<i>Opérations morphologiques:</i>	
$X \ominus B$	Érosion morphologique de X par B
$X \oplus B$	Dilatation morphologique de X par B
$X \circ B$	Ouverture morphologique de X par B
$X \bullet B$	Fermeture morphologique de X par B
<i>Cadre algébrique:</i>	
\mathcal{L}	Treillis
$f \vee g$	Supremum de f et g
$f \wedge g$	Infimum de f et g
<i>Mesures:</i>	
$l(X)$	Longueur de $X \subseteq \mathbb{R}^n$
$A(X)$	Aire de $X \subseteq \mathbb{R}^n$
$V(X)$	Volume de $X \subseteq \mathbb{R}^n$
$\sharp(X)$	Cardinal de $X \subseteq \mathbb{Z}^n$

Contents

Remerciements	i
Modalités d'examen	ii
Notations	iii
1 Introduction	1
1.1 Généralités	1
1.2 Éléments de perception visuelle	2
1.2.1 Le système visuel humain	2
1.2.2 La lumière	2
Représentation fréquentielle des couleurs	4
Les espaces de couleurs	4
L'espace de couleurs additif RGB.	5
Vers d'autres systèmes de couleurs: le système XYZ.	5
L'espace de couleurs soustractifs CMY.	8
D'autres systèmes de couleurs.	8
Les systèmes de couleurs en pratique.	10
Les fausses couleurs.	10
La couleur des objets	11
1.2.3 Effets visuels	11
1.2.4 L'éclairage	15
1.3 Représentation des images	15
1.3.1 Objets à représenter	15
La vidéo	17
Standards de télévision analogiques	18
1.3.2 Structures de données	18
1.4 Traitement d'images industriel	21
1.4.1 Remarques préalables	21
Décomposition d'une image en plans binaires	21
Critères de qualité	24
Visualisation 3D d'une image 2D	24
1.4.2 Segmentation d'images	24
1.4.3 Reconnaissance de caractères	24
1.4.4 Autres opérations	26

2	Transformations unitaires	27
2.1	Introduction	27
2.2	Bases théoriques du calcul matriciel	27
2.3	La transformée de FOURIER discrète	28
2.3.1	Périodicité intrinsèque	29
2.3.2	Note sur l'implémentation	31
	Orthogonalité et vecteurs propres	31
	Algorithme rapide	31
2.4	La transformée de HADAMARD	32
2.5	La transformée en cosinus discrète	32
2.5.1	Note sur l'implémentation	33
2.6	Transformée de FOURIER	33
2.6.1	Définition	33
2.6.2	Propriétés	34
2.6.3	Illustrations	36
	La fonction Rectangle	36
	La fonction Disque	38
	Dual de la fonction Rectangle	39
	Dual de la fonction Disque	39
2.7	Échantillonnage	40
2.7.1	Impulsion de DIRAC bidimensionnelle	40
	Propriétés	40
	Transformée de FOURIER	40
	Liens avec quelques images typiques	41
2.7.2	Le processus d'échantillonnage	41
2.8	Reconstruction	44
2.8.1	Filtre passe-bas à fenêtre rectangulaire	44
2.8.2	Filtre passe-bas à fenêtre circulaire	45
	Conditions de reconstruction	46
2.8.3	Autres filtres de reconstruction et d'interpolation	47
3	Filtrage linéaire et déconvolution	49
3.1	Notion de filtre idéal	49
3.1.1	Catégories de filtres idéaux	50
3.2	Passage d'une image dans un système linéaire	52
3.2.1	Implémentation	54
3.2.2	Forme de la fenêtre	54
3.2.3	Filtres passe-bas	54
3.2.4	Filtres passe-haut	57
3.2.5	Filtres de GABOR	57
	Définition	57
	Incertitude	59
3.2.6	Implémentation	62
	Normalisation	62
	Implémentations d'un filtre de GABOR	62
3.3	Introduction à la déconvolution	62
3.3.1	Modèle	63
3.3.2	Forme itérative des opérateurs de déconvolution	64

3.3.3	Approche déterministe ou aléatoire?	66
3.3.4	Algorithmes de déconvolution sélective	67
3.3.5	Déconvolution sélective	68
3.3.6	Convergence et unicité de la déconvolution sélective	68
4	Morphologie mathématique	71
4.1	Introduction	71
4.1.1	Historique	71
4.2	Rappels sur les notions d'ensembles	72
4.2.1	Égalité d'ensembles	72
4.2.2	Inclusion	72
4.2.3	Intersection	72
4.2.4	Union	72
4.2.5	Différence	72
4.2.6	Complémentaire	73
4.2.7	Symétrique	73
4.2.8	Translaté	73
4.2.9	Addition et soustraction de MINKOWSKI	74
4.2.10	Propriétés de base des transformations morphologiques	74
4.3	Transformations morphologiques élémentaires	75
4.3.1	Érosion	75
4.3.2	Dilatation	76
	Propriétés des opérations d'érosion et de dilatation	78
4.3.3	Ouverture morphologique	78
4.3.4	Fermeture morphologique	79
	Propriétés de l'ouverture et de la fermeture	79
	Interprétation de l'ouverture et de la fermeture	80
4.3.5	Propriétés générales	80
	Remarques concernant l'implémentation des opérations morphologiques élé- mentaires	80
	Un problème pratique: la gestion du bord dans les réalisations	82
4.4	Les transformations de voisinage	83
4.4.1	Amincissement et épaissement	85
4.5	Géodésie et reconstruction	85
4.5.1	Dilatation géodésique	86
4.5.2	Érosion géodésique	87
4.5.3	Reconstruction morphologique	87
4.6	Opérations morphologiques sur des images en niveaux de gris	88
4.6.1	Notion de fonction	88
4.6.2	Définition de propriétés des opérateurs de fonctions	89
4.6.3	Érosion, dilatation et combinaisons	89
	En pratique	90
4.6.4	Ouverture et fermeture morphologiques	93
4.6.5	Propriétés des opérateurs morphologiques de fonctions	97
4.6.6	Reconstruction des images à niveaux de gris	97

5	Filtrage non-linéaire	101
5.1	Généralités	101
5.2	Filtres non-linéaires	101
5.2.1	Filtres de rang	101
	Filtre médian	102
	Implémentation	103
5.2.2	Filtres à masque tournant	103
5.3	Filtres algébriques et filtres morphologiques	104
5.3.1	Introduction	104
5.3.2	Comment construire un filtre?	105
5.3.3	Exemples de filtres	106
	Filtres alternés séquentiels	106
	Toggle mappings	108
6	Extraction de traits caractéristiques et détection des contours	109
6.1	Opérateurs linéaires	110
6.1.1	Opérateurs basés sur le calcul de la dérivée	110
6.1.2	Opérateur de dérivée première	112
6.1.3	Opérateur de dérivée seconde	113
6.1.4	Échantillonnage de la dérivée et calcul de l'erreur résultante	113
	Calcul de l'erreur	114
	Comportement fréquentiel des opérateurs échantillonnés	115
6.1.5	Expressions pratiques d'opérateurs de gradient et masques de convolution	116
	Formes de base pour la dérivée première	117
	Problème pratiques.	117
	Filtres gradient de PREWITT.	118
	Filtres gradient de SOBEL.	119
	Formes de base pour la dérivée seconde	119
6.2	Opérateurs non-linéaires	120
6.2.1	Gradients morphologiques	120
7	Texture	125
7.1	Introduction	125
7.2	Caractérisation statistique des textures	127
7.2.1	Moyenne locale	129
7.2.2	Écart-type local	129
7.2.3	Histogramme local	129
	Moyenne	131
	Écart-type	131
	Obliquité	131
	"Kurtosis"	131
	Énergie	131
	Entropie	131
7.2.4	Matrice de co-occurrence des niveaux de gris	132
	Probabilité maximale	133
	Moments	134
	Moments de la différence inverse	134
	Contraste	134

	Entropie	134
7.3	Caractérisation géométrique des textures	134
	7.3.1 Approche fréquentielle	134
	7.3.2 Textures et énergie	136
8	Rehaussement et restauration	139
8.1	Généralités	139
8.2	Rehaussement	141
	8.2.1 Définition	141
	8.2.2 Rehaussement par fonction croissante	141
	8.2.3 Rehaussement par accentuation des contours	142
	Implémentation	142
	8.2.4 Manipulation de l’histogramme	142
	Exemple: fonction parabolique	143
	Égalisation de l’histogramme	143
	8.2.5 Filtrage homomorphique	144
8.3	Restauration	146
	8.3.1 Filtrage de WIENER	148
	Définition d’un critère	148
	Erreur quadratique moyenne	149
	Minimisation de l’erreur quadratique moyenne	150
	Cas particulier : signal idéal et bruit indépendants	152
	8.3.2 Déconvolution de WIENER	152
	Application aux cas des images	154
9	Étude de la forme	155
9.1	Description de la forme	155
	9.1.1 Description linéique de la forme	155
	Le codage en chaîne	155
	Approximation polygonale	157
	Descripteur de FOURIER	157
	9.1.2 L’arbre comme descripteur par éléments de surface	159
	9.1.3 Le squelette morphologique comme descripteur par éléments de surface	159
	Définition intuitive du squelette	159
	Propriétés d’une transformée par squelette	160
	Préliminaires: la taille morphologique et les filtres multirésolution	161
	Taille morphologique.	161
	Homothétie d’un ensemble discret.	161
	Remarque pratique concernant l’implémentation.	161
	Filtres multirésolution.	162
	Définition formelle du squelette	162
	Autres formules.	164
	9.1.4 Autres descripteurs par éléments de surface	164
9.2	Mesures	164
	9.2.1 Mesures géométriques de base	166
	Périmètre	166
	Périmètre d’un polygone	166
	Aire	166

	Caractéristique d'EULER-POINCARRE	166
	Propriétés topologiques.	169
9.2.2	Indices de forme	169
9.2.3	Moments	170
9.2.4	Mesures morphologiques	170
	Granulométries	170
	Covariogramme géométrique	170
	Enveloppe convexe	171
10	Segmentation d'images	173
10.1	Présentation du problème	173
10.1.1	Définition	173
10.2	Segmentation par seuillage	174
10.3	Segmentation par détection de régions	175
10.3.1	Ligne de partage des eaux	175
10.3.2	Expression formelle d'un algorithme de segmentation par ligne de partage des eaux	177
	Analyse d'images binaires	177
	Algorithme de construction du squelette par zone d'influence.	178
	Images en niveaux de gris	178
	Notations:	178
	Initialisation	179
	Construction	179
	Marquage	180
11	Codage et compression	181
11.1	Introduction	181
12	Applications industrielles du traitement d'images	183
12.1	Introduction	183
12.1.1	Spécificités du traitement d'image industriel	183
12.1.2	Applications typiques	185
12.1.3	Tâches d'inspection	191
12.2	Moyens matériels	196
12.2.1	Chaîne d'acquisition et de traitement	196
	Éclairage	197
	Optique	199
	Caméra	200
	Carte d'interface	201
	Processeur	202
	Caméras intelligentes	202
12.2.2	Défauts introduits par la chaîne d'acquisition	202
	Éclairage	203
	Optique	203
	Caméra	204
12.3	Moyens logiciels	204
12.3.1	Organisation générale des traitements	205
12.3.2	Techniques de pré-traitement	206
12.3.3	Techniques de localisation	208

12.3.4	Techniques d'analyse et de mesure	209
12.3.5	Techniques de diagnostic	210
12.4	Un exemple concret: inspection de marquage	210
	Glossaire	217

Chapter 1

Introduction

1.1 Généralités

La description d'une image est une démarche faussement facile. S'il est vrai qu'une image consiste en un jeu de valeurs représentables, sur un écran par exemple, comprendre la nature d'une image s'avère délicat, d'autant que l'on confond souvent l'image, la visualisation et les traitements effectués sur l'image.

Pour aborder la question du traitement d'image, nous distinguons les éléments suivants:

- la *perception* d'une image. Elle s'articule autour des caractéristiques du système visuel humain. Ainsi, il apparaît que l'œil est sensible à certaines fréquences du spectre électromagnétique; ces fréquences représentent la lumière en général. Dans certains cas, le rayonnement électromagnétique à observer se situe en dehors de la plage des fréquences visibles –les rayons X par exemple. On doit alors recourir à un révélateur pour transposer le signal à observer dans une plage visible.
- la *représentation* d'une image. Il s'agit de représenter une entité physique sous une forme électrique ou une forme informatique. La représentation joue un rôle essentiel dans une chaîne de traitement car elle conditionne la capacité de stockage nécessaire ainsi que la mise en œuvre. En outre se pose régulièrement la question du changement de représentation. À titre d'illustration, une séquence vidéo au format PAL nécessite un débit de $199 [Mb/s]$ sans compression. Le changement de représentation que représente la compression ramène ce débit à des valeurs de l'ordre de 4 à 9 $[Mb/s]$ pour une qualité similaire.
- les *traitements d'image*. Les traitements sont des opérations qui interprètent ou affectent la représentation d'une image; nous considérerons exclusivement le traitement d'images échantillonnées.

Le traitement d'images est un domaine très vaste qui a connu, et qui connaît encore, un développement important depuis quelques dizaines d'années. Il a donc fallu faire un choix à l'heure de rédiger ce cours. Notre démarche a consisté à choisir une série de thèmes spécifiques et à les développer tant formellement que pratiquement. C'est ainsi que les opérateurs de traitement d'image seront étudiés suivant trois axes:

- la *définition analytique des opérateurs*.
- l'*étude des propriétés*. Elle complète et illustre chaque définition. Une compréhension fine d'un opérateur et de ses propriétés se révèle un excellent moyen de développer l'intuition pour aborder un problème concret.

- l'*implémentation*. Un opérateur peut avoir une description mathématique unique mais plusieurs réalisations présentant chacune des avantages et des inconvénients. En pratique, le choix d'un algorithme sera guidé par une série de contraintes comme l'utilisation de la mémoire, la vitesse, les possibilités de parallélisation, etc.

1.2 Éléments de perception visuelle

Tout document de traitement d'image qui se respecte débute par un résumé du fonctionnement du système visuel humain. Assez paradoxalement il s'agit d'une des tâches les plus complexes car le fonctionnement du système visuel humain fait intervenir la subjectivité de l'observateur et car, en pratique, il apparaît difficile d'inclure la plupart des résultats des études psychovisuelles dans un traitement d'image courant.

1.2.1 Le système visuel humain

L'œil est un système complexe (cf. figure 1.2). La lumière incidente est réfractée par la cornée et dirigée vers la pupille; la pupille est l'ouverture de l'iris par laquelle la lumière pénètre dans l'œil. La lumière est ensuite réfractée une seconde fois en direction du fond du globe oculaire où elle illumine la rétine. Cette dernière est composée d'une série de récepteurs, appelés photorécepteurs, reliés à des cellules qui transmettent des signaux au nerf optique.

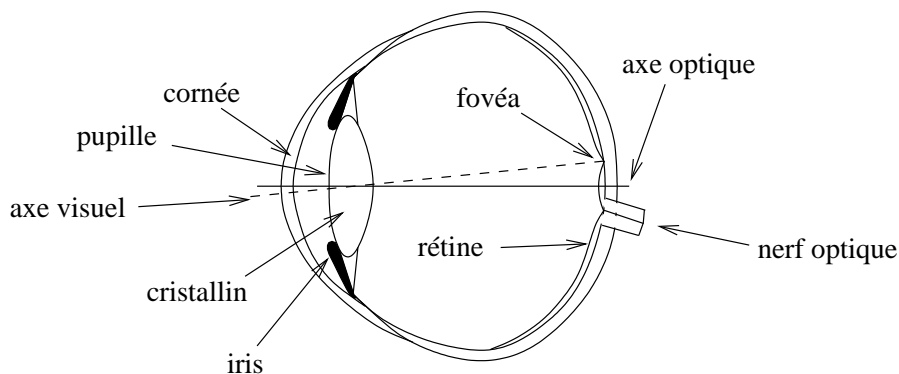


Figure 1.1: Coupe latérale simplifiée de l'œil.

1.2.2 La lumière

La lumière couvre une partie du spectre d'énergie électromagnétique. Un rayonnement électromagnétique est en général constitué d'un certain nombre de longueurs d'onde (ou fréquences) que les dispositifs dispersifs de séparer en un spectre. Le spectre est soit discret, soit continu.

Les longueurs d'onde du spectre visible s'étendent approximativement de 380 à 720 [nm]. Une source est caractérisée par (i) son rayonnement, mesurable dans un système de grandeur correspondant à l'action proprement visuelle, (ii) par le mélange des longueurs d'onde de cette énergie, mélange qui produit une sensation de *couleur*. La lumière est donc une distribution d'énergie émise à certaines fréquences ayant une certaine *intensité*.

Pour caractériser une couleur *monochromatique*, il suffit de connaître sa longueur d'onde λ et la *luminance* L , expression qualitative de la brillance énergétique. Dans ces conditions, l'œil se

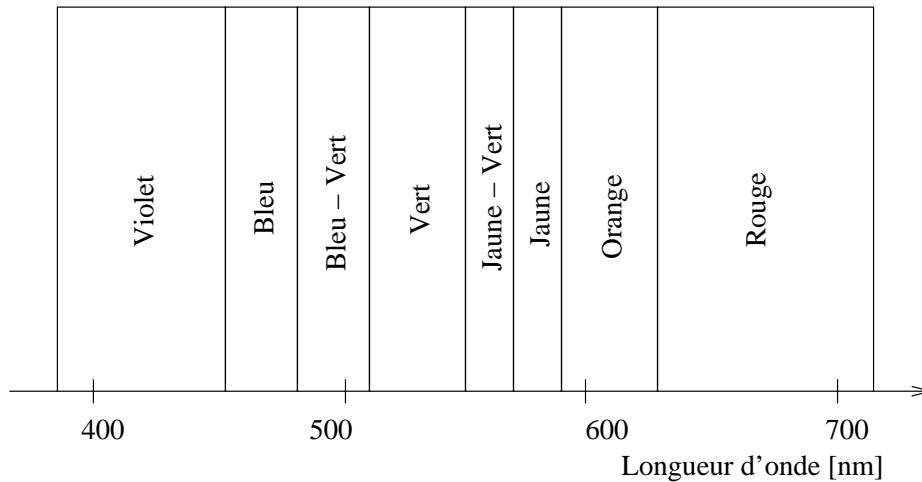


Figure 1.2: Les longueurs d'onde associées aux couleurs.

comporte donc comme un récepteur bivariant, la connaissance de λ et de celle de L étant nécessaires et suffisantes pour caractériser l'impression visuelle créée.

Supposons maintenant une source constituée d'un certain nombre de radiations simples formant un spectre de raies ou un spectre continu avec une répartition énergétique donnée. La question qui se pose est le problème de la résultante, pour le récepteur visuel, de l'addition de l'ensemble de ces radiations. Des études psychovisuelles ont révélé les lois suivantes (cf. le livre de BUSER et d'IMBERT [1] pour les détails):

- lorsqu'on additionne des radiations monochromatiques, l'œil ne sait pas distinguer les composantes de la couleur résultante, contrairement à l'oreille qui demeure capable de discerner les composantes fréquentielles d'une onde acoustique.
- avec des combinaisons adéquates de longueurs d'ondes et de brillance complètement différentes, il est possible de réaliser l'égalisation de deux demi-écrans, en luminance et en couleurs.
- l'addition de certains couples de couleurs peut produire une impression de *lumière blanche*. Les deux couleurs sont alors dites *complémentaires*. Pour toute radiation extérieure à l'intervalle $490 - 570 [nm]$, il existe une couleur complémentaire. En revanche, toute radiation inférieure à cet intervalle n'a pas de couleur complémentaire dans le spectre visible.
- lorsque les luminances L_1 et L_2 , de deux couleurs complémentaires, sont convenablement ajustées, on obtient du blanc.
- si les luminances L_1 et L_2 des couleurs λ_1 et λ_2 ne sont pas dans un rapport adéquat, on obtient une couleur "délavée" –on dira non *saturée*– au lieu du blanc.
- en généralisant, on peut démontrer que l'addition de deux couleurs quelconques non complémentaires produira une couleur nouvelle λ_r non saturée

$$L_1 + L_2 = L_b + L_r \quad (1.1)$$

où L_r correspond à la longueur d'onde λ_r .

- l'équation précédente se généralise du fait des lois de GRASSMANN qui énoncent l'additivité, la multiplicativité, l'associativité et la transitivité dans le mélange des couleurs. Finalement, on aboutit à ce qu'un nombre quelconque N de couleurs $\lambda_1, \lambda_2, \dots, \lambda_N$ équivaut à l'addition d'un flux L_b de blanc et d'un flux L_r d'une couleur λ_r résultante, ce qui revient à la détermination de 3 variables L_b, L_r et λ_r

$$\sum_{i=1}^N L_i = L_b + L_r \quad (1.2)$$

Représentation fréquentielle des couleurs

Une première possibilité de représenter une couleur consiste à estimer l'ensemble de ses composantes spectrales. En principe, il suffirait de filtrer le signal de couleur par une série de filtres à bande étroite et de mesurer l'intensité de chaque sortie filtrée. Le signal de couleur vaut alors

$$\int_{\lambda} L(\lambda) d\lambda \quad (1.3)$$

Le problème posé par cette approche est le grand nombre de capteurs nécessaires à la description de la couleur.

Les espaces de couleurs

Une autre possibilité de représentation consiste à utiliser un espace de couleurs à k dimensions. Des expériences psychovisuelles d'égalisation (voir KUNT *et al.* [18]) ont montré qu'en combinant trois stimuli de longueur d'ondes particulières, il est possible de synthétiser presque toutes les couleurs existantes. Cette expérience peut s'effectuer à l'aide d'un système de reproduction de couleurs de type additif ou de type soustractif.

Choisissons trois radiations monochromatiques que nous qualifierons de *primaires*. Le problème sera de déterminer dans quelle mesure il est possible, à partir de ces trois couleurs, de reproduire une radiation colorée quelconque et selon quelles règles quantitatives. Dans le cas d'un mélange additif, l'expérience d'égalisation des couleurs se déroule par exemple de la manière suivante: trois couleurs primaires A, B et C sont, après pondération de leur intensité respectivement par a, b et c , projetées sur écran noir avec un certain recouvrement comme l'illustre la figure 1.3. En jouant sur

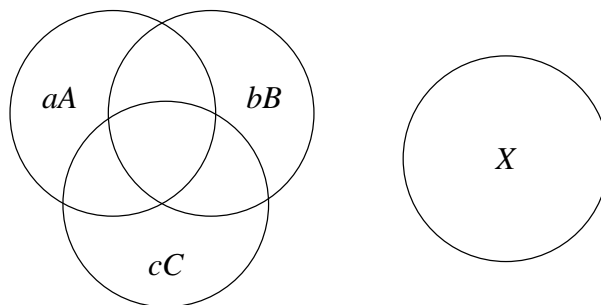


Figure 1.3: Expérience d'égalisation d'une couleur X au moyen de trois couleurs primaires A, B et C .

les paramètres a, b et c , l'observateur essaie d'égaliser la couleur à définir X avec le mélange additif $aA + bB + cC$. Au cas où l'égalisation est impossible, une ou plusieurs des couleurs primaires sont additionnées non pas au mélange mais à la couleur à définir X . On obtient alors, pour cette couleur, un ou plusieurs facteurs négatifs $-a, -b$ ou $-c$.

L'espace de couleurs additif RGB. Pour des raisons de standardisation, la commission internationale de l'éclairage (CIE) a réalisé en 1931 l'expérience d'égalisation de couleur par mélange additif. Un système qui s'imposa presque naturellement était fondé sur les trois couleurs monochromatiques, rouge R (700 [nm]), vert V (546,1 [nm]) et bleu B (435,8 [nm]), ces deux dernières étant émises par l'arc au mercure (d'où leur choix). Ce fut le système RGB. On a ainsi déterminé pour ce système:

- les coordonnées trichromatiques désignées dans le cas d'espèce r , g et b ;
- les unités lumineuses respectives des trois primaires en imposant qu'une certaine source blanche tombe au centre du diagramme;
- à partir de ci-dessus, les coefficients de distribution pour une couleur quelconque.

La figure 1.4 montre les courbes des coefficients de distribution des couleurs primaires.

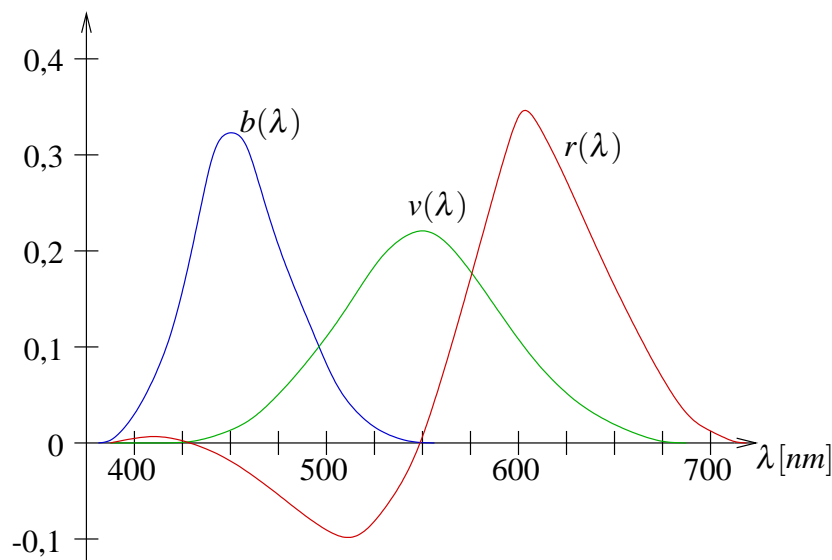


Figure 1.4: Courbes d'égalisation spectrale obtenues par égalisation des couleurs au moyen d'un mélange additif (d'après [14]).

La figure 1.5 représente le diagramme chromatique-type dans le système RGB : chaque couleur primaire est placée au sommet d'un triangle isocèle rectangle; l'ensemble des couleurs monochromatiques s'aligne sur le contour du diagramme, dit *lieu spectral*, tandis qu'une radiation complexe se situe à l'intérieur du contour.

Pour être plus précis, les trois couleurs RGB définissent un triangle significatif de la perception visuelle. À condition de respecter les proportions, il est possible de moduler les composantes d'une intensité minimale à une intensité maximale. L'ensemble des couleurs représentables se situe donc à l'intérieur d'une pyramide telle que représentée à la figure 1.6.

Vers d'autres systèmes de couleurs: le système XYZ. À l'usage, il s'est avéré que ce système RGB présentait un certain nombre d'inconvénients (cf. [1]) du fait, en particulier que: (i) en raison du choix des primaires, l'une des coordonnées trichromatiques était toujours négative pour les radiations monochromatiques, (ii) on risquait d'attribuer aux couleurs primaires une signification physiologique qu'elles n'ont pas (puisque leur choix est arbitraire).

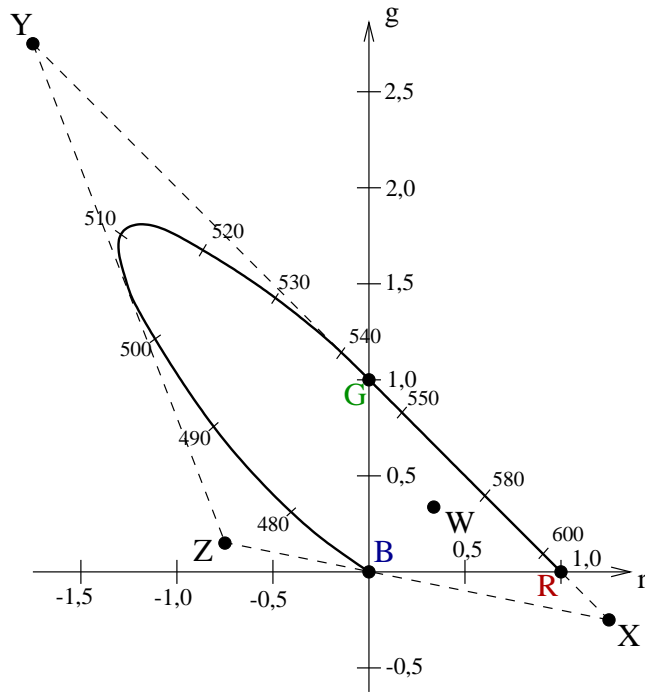


Figure 1.5: Diagramme chromatique RGB de la CIE.

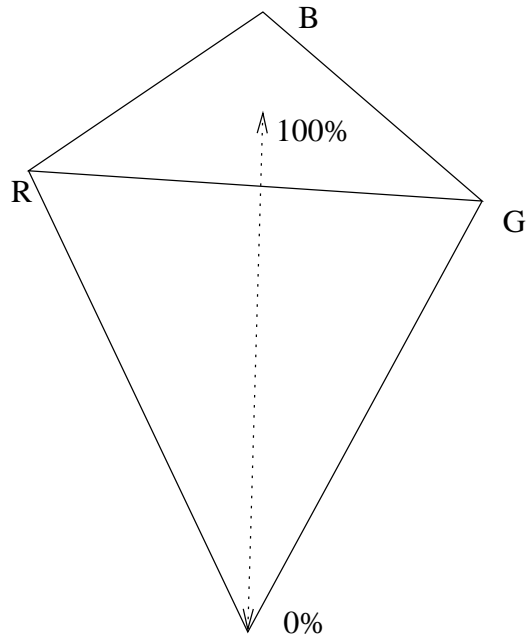


Figure 1.6: Pyramide des couleurs obtenues au moyen du tri-stimulus RGB.

Afin de parer à ces inconvénients, la *CIE* a introduit, également en 1931, un autre espace de couleurs appelé *XYZ*. Les composantes *X*, *Y* et *Z* qui décrivent un stimulus sont liées aux composantes *RGB* par la relation matricielle suivante [18]

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 2,769 & 1,7518 & 1,13 \\ 1 & 4,5907 & 0,0601 \\ 0 & 0,0565 & 5,5943 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1.4)$$

En normalisant le tri-stimulus *XYZ* par rapport à $X + Y + Z$, on obtient le système *xyz*. Une particularité intéressante de ce nouveau référentiel est que seules deux variables, par exemple *x* et *y*, sont indépendantes. Ces deux composantes sont appelées les composantes *chromatiques* du stimulus. On a donc les relations

$$x = \frac{X}{X+Y+Z} \quad (1.5)$$

$$y = \frac{Y}{X+Y+Z} \quad (1.6)$$

$$z = \frac{Z}{X+Y+Z} \quad (1.7)$$

$$x+y+z = 1 \quad (1.8)$$

Après transformation du stimulus original en deux composantes normalisées indépendantes *x* et *y*, il est possible de décrire l'ensemble des couleurs par les deux variables de chrominance *x* et *y*. Il en résulte un diagramme tel que celui représenté à la figure 1.7.

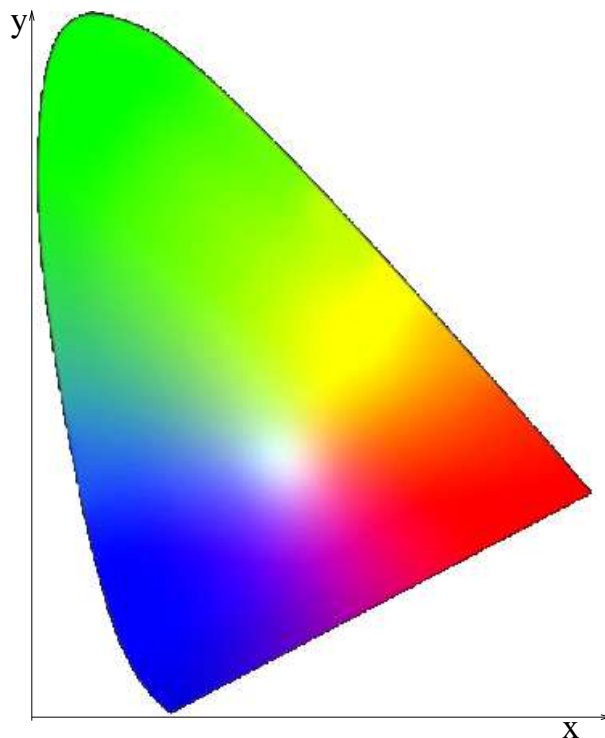


Figure 1.7: Diagramme chromatique (approché!) défini par les deux variables de chrominance *x* et *y*.

Les couleurs monochromatiques sont représentées sur la partie non rectiligne du contour (comme un U retourné). Les couleurs qui résultent d'un mélange d'un mélange additif de certaines couleurs de base sont représentées à l'intérieur de la forme en U. Il existe aussi certains point particuliers, comme

le *point d'égale énergie* qui correspond à une contribution égale des trois composantes de base X , Y et Z , ce qui correspond à $x = y = 1/3$. L'expérience d'égalisation des couleurs a montré qu'il suffisait de trois couleurs de base pour synthétiser la quasi-totalité des stimuli colorés. Comme le diagramme chromatique xy ne reprend que deux variables indépendantes, ce diagramme est une version réduite de l'espace à trois dimensions tel défini par le cube de la figure 1.8.

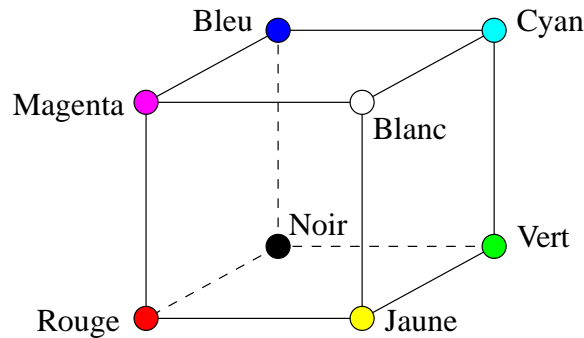


Figure 1.8: Espace tridimensionnel des stimuli produits par les composantes RGB .

C'est dans ce genre d'espace que nous avons représenté les couleurs du diagramme chromatique approché à la figure 1.7; le diagramme tridimensionnel correspondant est illustré à la figure 1.9.

La dimension non représentée sur ce diagramme est celle de la *luminance* qui permet d'éclaircir ou d'assombrir une couleur en modulant son degré de noir. Par exemple, en modulant du minimum au maximum la luminance du point d'égale énergie, on passe du noir au blanc en passant par toutes les teintes de gris. Chaque couleur possède une luminance maximale dont la représentation graphique est fournie à la figure 1.10.

Le diagramme chromatique xy est donc le lieu des couleurs de chrominance (x, y) de luminosité maximale. En effet, un changement des proportions entre x et y conduirait automatiquement à la perception d'une autre couleur. Les deux variables de chrominance x et y peuvent se transformer en variables possédant une interprétation physique plus intuitive à savoir la *teinte* et la *saturation*. La notion de teinte (*hue* en anglais) est associée à la longueur d'onde, c'est-à-dire à la position le long du U retourné. La saturation exprime quant à elle le degré de blanc; elle est définie comme la distance entre le point d'égale énergie et la couleur à définir.

La figure 1.11 illustre la décomposition d'une image dans les trois canaux RGB et HSI .

L'espace de couleurs soustractifs CMY. La détermination des composantes RGB d'une onde s'opère par addition sur fond noir. L'arrière-plan est donc supposé absorbant pour toutes les couleurs. Un tel système n'est pas adéquat pour traiter l'impression sur feuille blanche car cette dernière réfléchit l'ensemble des couleurs. Pour l'impression, on a donc défini d'autres systèmes de couleurs, complémentaires au système RGB . Il s'agit par exemple du système Cyan, Magenta et Yellow (CMY), complémentaire du système RGB et dont la somme produit le noir¹.

D'autres systèmes de couleurs. À côté du système RGB , couramment utilisé en infographie, il existe d'autres systèmes de couleurs tels les systèmes YIQ , YUV ou YC_bC_r . Ces systèmes de couleurs peuvent être calculés par transformation matricielle des composantes RGB auxquelles on a préalablement appliqué une correction γ ².

¹On préfère parfois ajouter d'emblée le noir au composante CMY , ce qui aboutit au système quadrichromatique $CMYK$.

²La *correction gamma* consiste à appliquer un fonction inverse de celle qui traduit, à l'intérieur du téléviseur, le signal électrique en une intensité des canons à électrons destinés à éclairer l'écran.

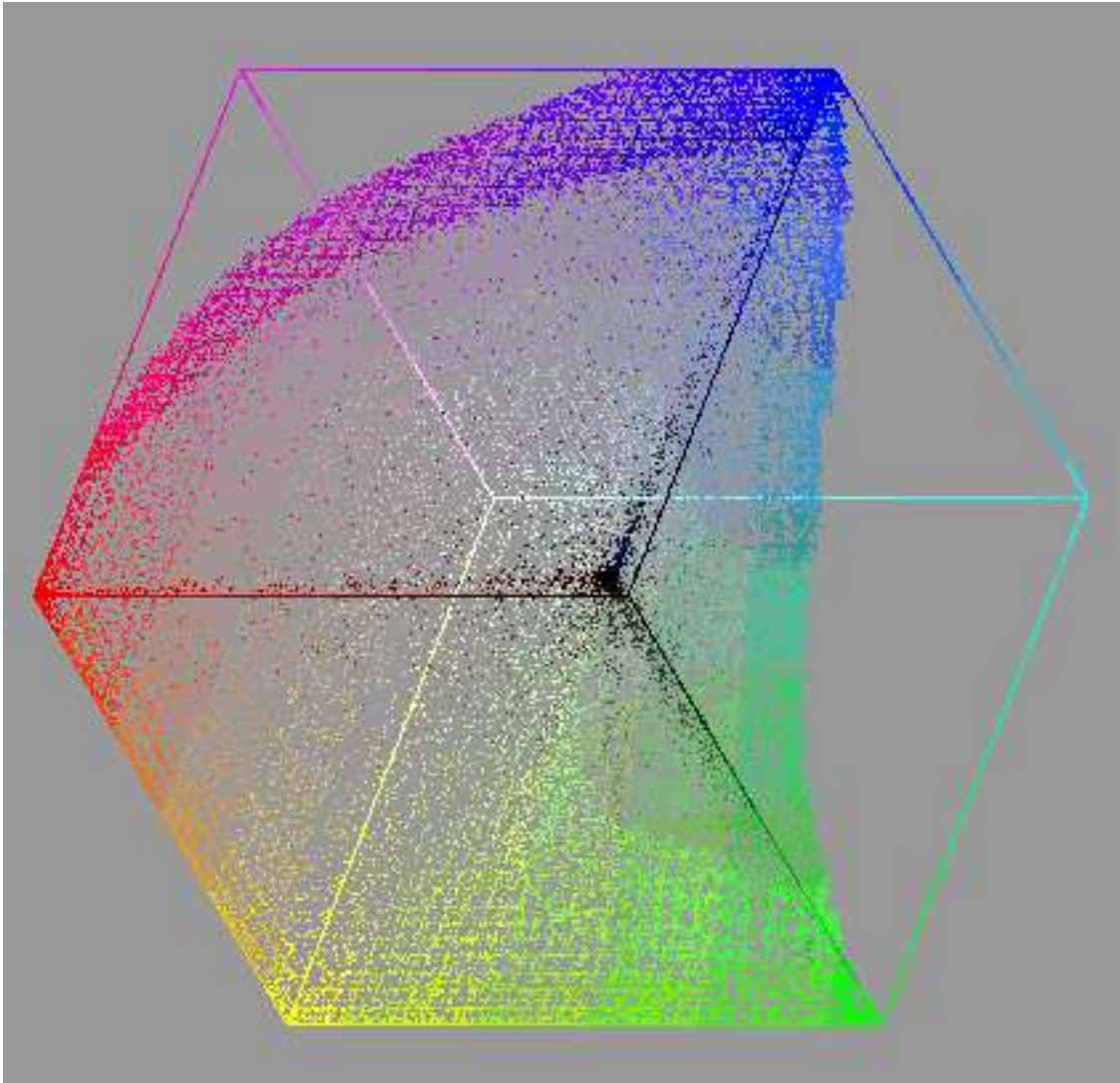


Figure 1.9: Espace tridimensionnel des couleurs du diagramme de chrominance approché.

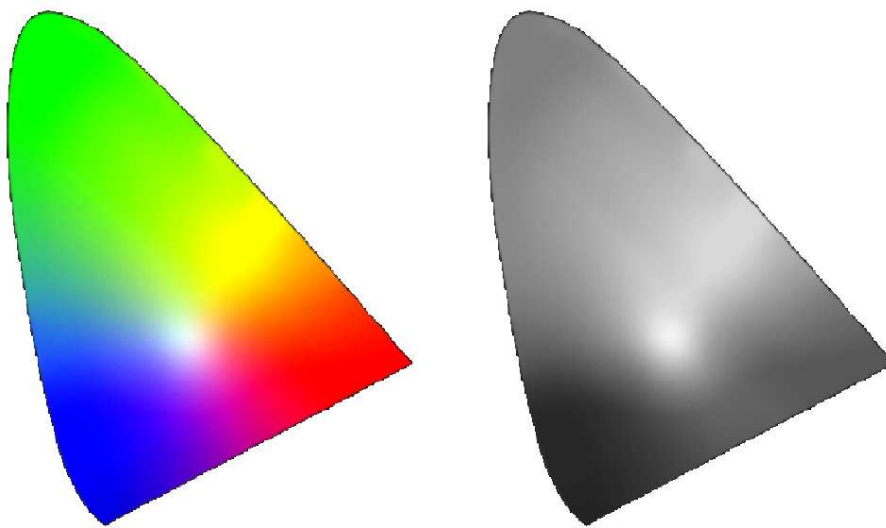


Figure 1.10: Diagramme chromatique xy et luminance maximale en chaque point.

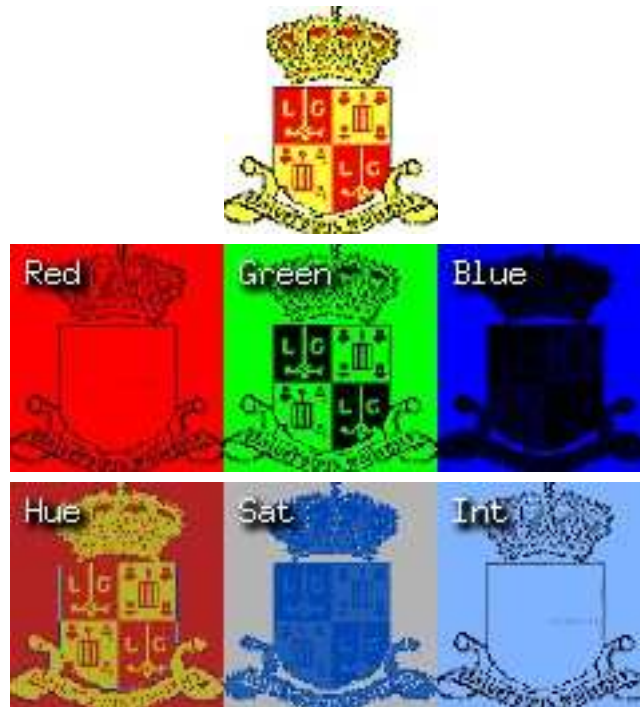


Figure 1.11: Décomposition d'une image en couleurs.

Les systèmes de couleurs en pratique. Les standards de télévision analogiques (PAL, NTSC) définissent chacun les positions de couleurs RGB dans le diagramme chromatique. À cet égard, il est intéressant de remarquer que, contrairement aux apparences, ces positions ne se situent pas sur le pourtour. Autrement dit, les couleurs de base ne sont pas monochromatiques! Comme les positions de ces systèmes ne coïncident pas, la totalité des couleurs d'un système n'est pas représentable dans un autre système. D'autre part, un téléviseur n'est pas réellement calibré à la fabrication et son usage produit une usure. Cela nous amène à considérer avec beaucoup de prudence la signification physique réelle d'une couleur affichée par un téléviseur. On peut exprimer des réserves similaires pour le cas des moniteurs utilisés dans le monde informatique. Pour désigner une couleur en infographie, on utilise généralement une quantification sur 8 bits par composante de couleurs. Les couleurs sont alors représentées par un triplet de valeurs exprimées en hexadécimal, en décimal (cf. table 1.1) ou en binaire.









Hexadécimal		R	G	B
00 00 00		0	0	0
00 00 FF		0	0	255
00 FF 00		0	255	0
00 FF FF		0	255	255
FF 00 00		255	0	0
FF 00 FF		255	0	255
FF FF 00		255	255	0
FF FF FF		255	255	255

Table 1.1: Table de correspondance de couleurs définies sur 8 bits.

Les fausses couleurs. La représentation des valeurs d'une image nécessite 8 bits dans le cas d'une image monochrome et $3 \times 8 = 24$ dans le cas d'une image RGB. Pour réduire le nombre de bits dans

le cas d'une image en couleurs, on réduit la taille de l'espace des couleurs en sous-échantillonnant l'espace RGB. Il est également possible de considérer les 256 valeurs d'une image comme les 256 indices d'un vecteur tridimensionnel reprenant les composantes RGB de couleurs. On parle alors de palette de couleurs ou de *Color Look Up Table* (CLUT). La figure 1.12 montre la palette de couleurs utilisées par le logiciel de navigation Netscape Navigator ainsi que la décomposition des couleurs dans les composantes RGB.

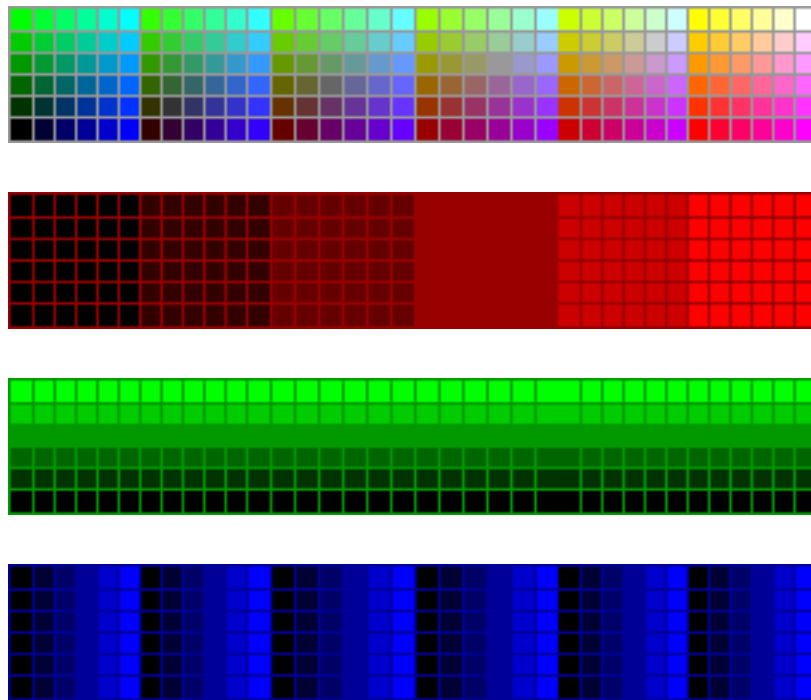


Figure 1.12: Palette de couleurs utilisée par les logiciels de navigation ainsi que la décomposition en R, G et B.

La figure 1.13 montre une image monochrome et une image obtenue en considérant une palette de couleurs aléatoires –on parle représentation en *fausses couleurs* car elles ne représentent pas des couleurs physiques.

La couleur des objets

La couleur d'un objet est significative de la réflectance de l'objet ainsi que de l'éclairage. Dès lors, à supposer qu'un onde lumineuse isochromatique (c'è-d. blanche) soit envoyée en direction de l'objet, sa surface absorbera certaines longueurs d'onde et en réfléchira d'autres. On ne pourra donc jamais définir la couleur que comme la couleur liée à la longueur d'onde dominante.

De plus, la réflexion peut être *spéculaire* (effet de miroir) ou être *diffuse*. L'infographie utilise abondamment ces techniques pour créer des effets d'ambiance. Un exemple, produit avec le logiciel BLENDER, est fourni à la figure 1.14.

1.2.3 Effets visuels

Parmi les nombreux effets de perception visuelle, le phénomène de contraste joue un rôle essentiel. Il est illustré à la figure 1.15. On dispose une pastille d'intensité $I + \Delta I$ sur un fond d'intensité fixée. Des expériences physiologiques montrent que la visibilité de la pastille disparaît sous très faible ou très forte intensité [1]. La courbe marquant cette transition est fournie à la figure 1.15.(a), partie droite.

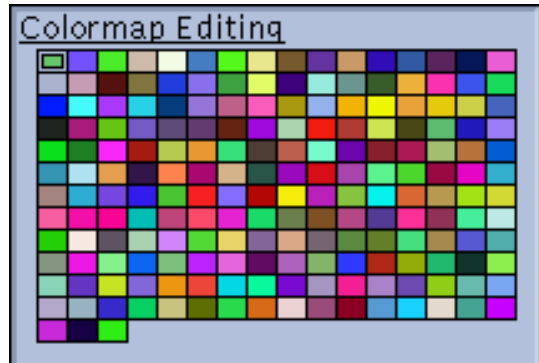
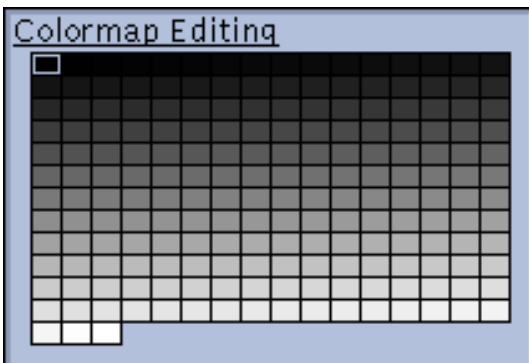


Figure 1.13: Une image en niveaux de gris, l'équivalent avec une palette de couleurs aléatoires et les palettes respectives.

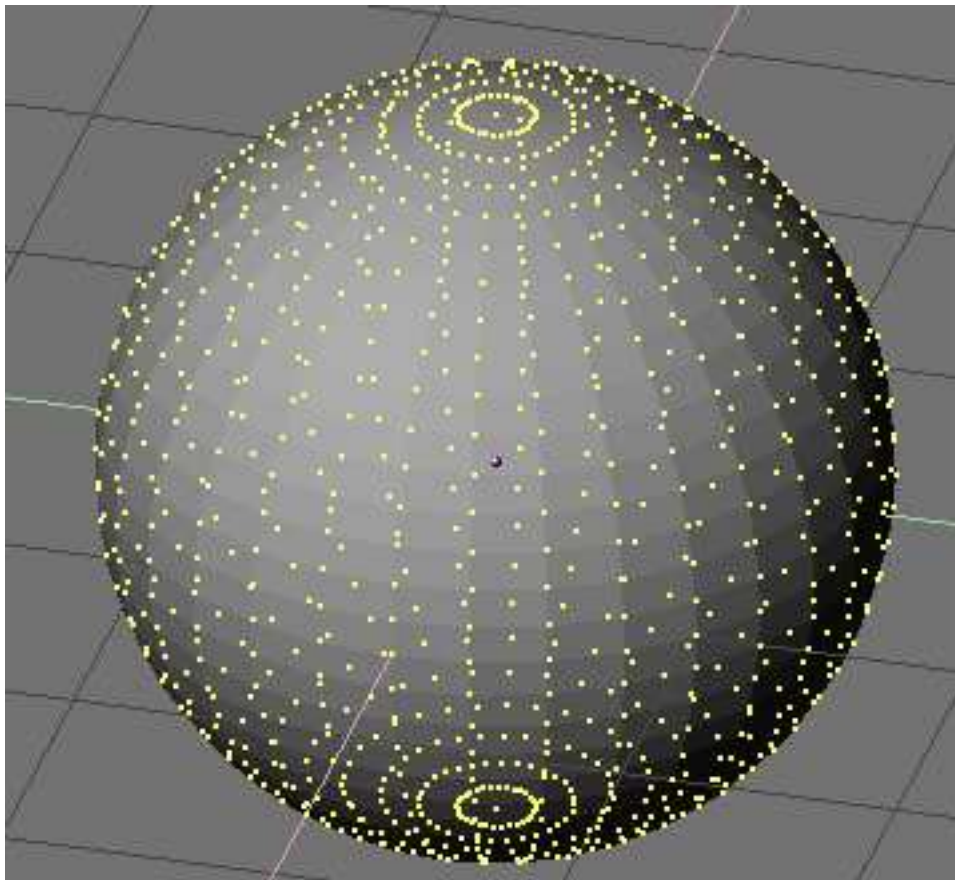


Figure 1.14: Une image de synthèse 3D.

On peut encore compliquer l'expérience en découpant la pastille centrale en deux parties. On obtient

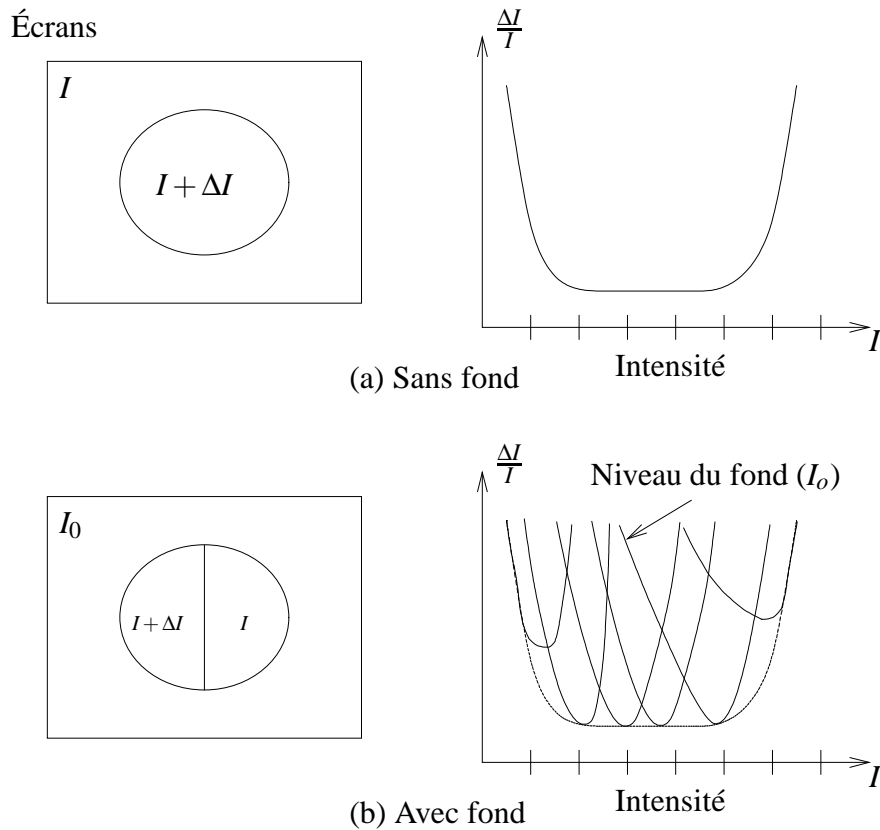


Figure 1.15: Différents arrangements expérimentaux permettant d'étudier les variations supraliminaire d'intensité, ΔI , à partir de l'intensité I .

alors des courbes de sensibilité qui rétrécissent en largeur, ce qui indique une perte de sensibilité accrue. Ce genre d'effets est également illustré à la figure 1.16 où l'œil perçoit différemment les trois rectangles centraux alors qu'ils sont de luminance égale.

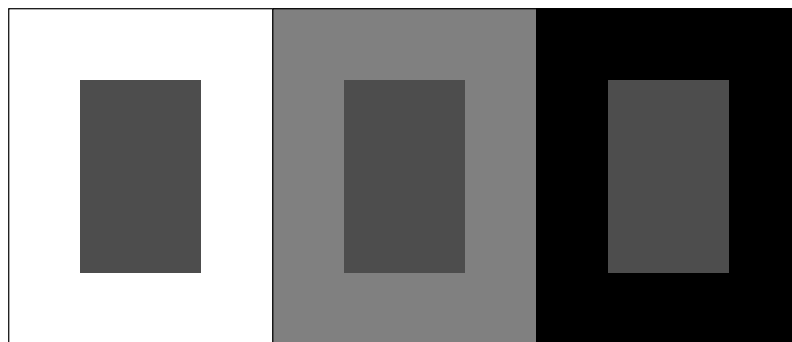


Figure 1.16: Illustration d'un effet perceptif.

La considération des aspects physiologiques dans le cadre d'applications industrielles reste délicate. En pratique, on se contente de certaines lois, comme la loi de WEBER ou le caractère passe-bas de l'œil; pire, on fait l'impasse sur les effets perceptifs.

1.2.4 L'éclairage

Dans le monde des applications d'imagerie industrielles, l'éclairage joue un rôle essentiel. En effet, une non-uniformité de couleurs entraîne des artefacts dans l'observation et, par conséquent, dans le traitement effectué. Il est dès lors primordial de démarrer la conception d'une chaîne de traitement d'images par la question de l'éclairage.

1.3 Représentation des images

1.3.1 Objets à représenter

Tout naturellement, la question de la représentation des images part de la volonté de trouver un moyen de définir un jeu de valeurs de luminances ou de couleurs. Pour répondre à cette question, il faut considérer plusieurs éléments:

- la nature des objets à représenter. Il peut s'agir de la luminance, des composantes RGB, de fausses couleurs ou encore des bits de transparence. La figure 1.17 montre l'usage de bits de transparence. L'usage de ces bits est simple. Soient $i(x,y)$ la valeur de l'image au point (x,y) , $t(x,y)$ la valeur de transparence et $o(x,y)$ la valeur de l'image finale. La valeur de transparence est typiquement définie sur 1 bit (elle vaut alors 0 ou 1) ou sur 8 bits. Prenons une transparence définie sur 8 bits; nous supposons que ces valeurs soient comprises dans l'intervalle $[0,255]$ tel que 0 soit la valeur d'intensité nulle ou d'opacité totale. Le principe consiste à appliquer la formule suivante

$$o(x,y) = \frac{t(x,y)}{255}i(x,y) \quad (1.9)$$

On peut bien entendu avoir une transparence par tout ou rien – 1 bits de transparence suffit

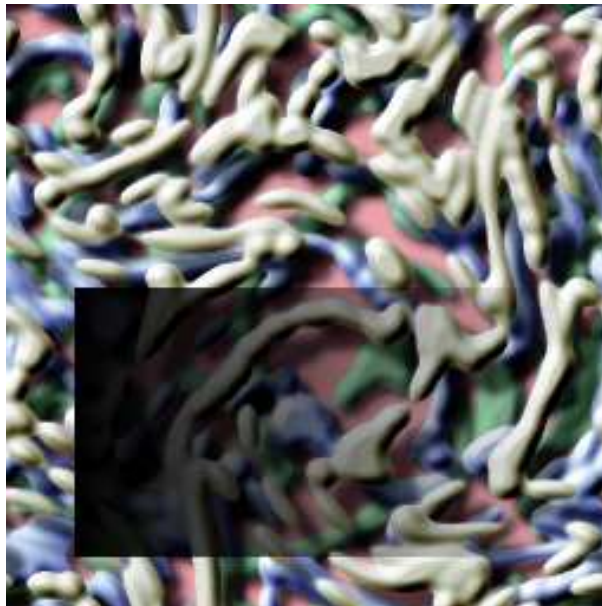


Figure 1.17: Utilisation de bits de transparence à l'intérieur du petit rectangle.

alors–, ou avoir une série ordonnée de plans de transparence.

- le cas échéant, la structure d'échantillonnage. L'échantillonnage aboutit à définir des images sur un domaine de définition discret, sous-ensemble de \mathbb{Z}^n , appelé trame. Chaque point est

appelé `pixel` (pour *picture element*).

Il existe deux sortes de trame suivant la grille d'échantillonnage adoptée. Le choix de la trame

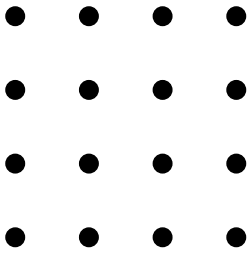
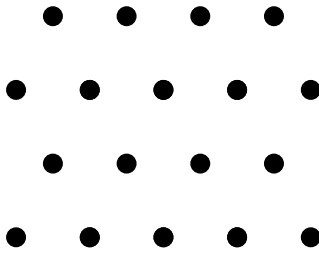
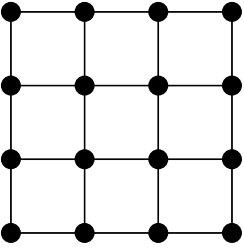
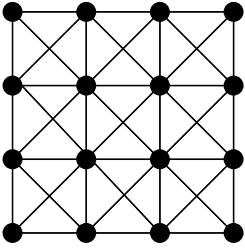
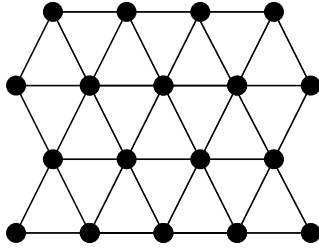
Trame carrée		Trame hexagonale	
			
4-connexité	8-connexité	6-connexité	
			

Table 1.2: Types de trames couramment utilisées et connexités associées.

hexagonale est judicieux pour parler de la notion de *connexité*. La connexité représente le nombre de voisins liés à un point. À ce titre, on remarquera que le complémentaire d'un objet 4-connexe est un objet 8-connexe. Par contre, le complémentaire d'un objet 6-connexe est également 6-connexe. La 6-connexité offre une plus grande symétrie (pour le calcul de la distance par exemple) mais son usage exige quelques précautions. On peut aisément passer d'une trame carrée à une trame hexagonale en décalant une ligne sur deux d'un demi-pas d'échantillonnage horizontal. Si cette solution est théoriquement insatisfaisante –il faut en toute rigueur recourir à une interpolation–, elle permet d'effectuer un traitement en trame hexagonale moyennant l'introduction d'une erreur bien souvent acceptable.

Nous adopterons la terminologie suivante pour désigner les objets à représenter:

- 2D.** Il s'agit d'une matrice bidimensionnelle de valeurs. Ces valeurs (luminance, couleurs, ...) sont par exemple issues de l'échantillonnage d'une image fixe.
- 3D.** Cette fois, le jeu de valeurs est défini sur base d'une grille d'échantillonnage tridimensionnelle. On obtient cette type d'images par scanner par exemple. La plupart des techniques de traitement d'image bidimensionnelles se généralisent sans peine à des images 3D.
- 2D+t.** Comme t représente le temps, les images $2D+t$ désignent une animation, une séquence vidéo ou les images d'un film.
- 3D+t.** Les images de type $3D+t$ sont des images tridimensionnelles animées. Il s'agit par exemple d'images de synthèse 3D animées.

La vidéo

D'un point de vue informatique, on représente une image $2D + t$ comme une succession d'images numériques $2D$. Dans le monde de la télévision analogique, on ne disposait pas de cette représentation et il a fallu trouver d'autres moyens de représenter une séquence vidéo.

Le procédé imaginé a consisté à fabriquer un signal électrique composite mélangeant luminance, couleurs, son et informations de service. Ce signal électrique balaie les lignes d'un écran à la manière de la figure 1.18.

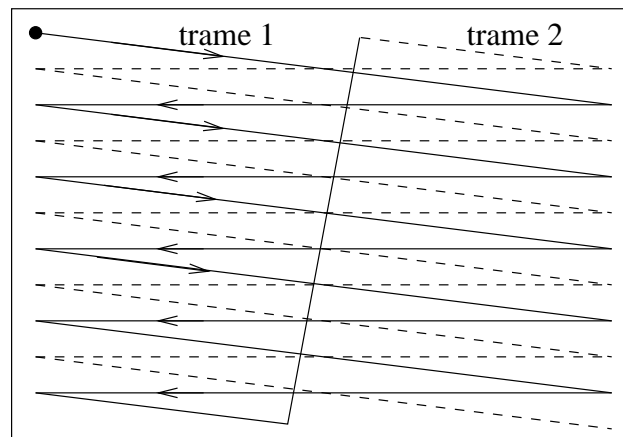


Figure 1.18: Description du format entrelacé.

L'image présentée sur un téléviseur est lue ligne après ligne, mais une ligne sur deux. Chaque ligne est lue de gauche à droite et les lignes sont parcourues de haut en bas. L'ensemble des lignes lues lors d'une première analyse constitue une première trame. Les lignes restantes constituent une seconde trame. Ces deux trames sont entrelacées; elles forment une image. Ce procédé, utilisé pour la télévision analogique, porte le nom d'entrelacement. L'autre manière de faire consistant à envoyer toutes les images d'une image est à la base du format *progressif*; le format progressif est utilisé pour les moniteurs de PC.

Qu'il s'agisse d'un format progressif ou entrelacé, la transmission d'un signal vidéo analogique s'effectue ligne par ligne. Le signal électrique d'un téléviseur est décrit à la figure 1.19.

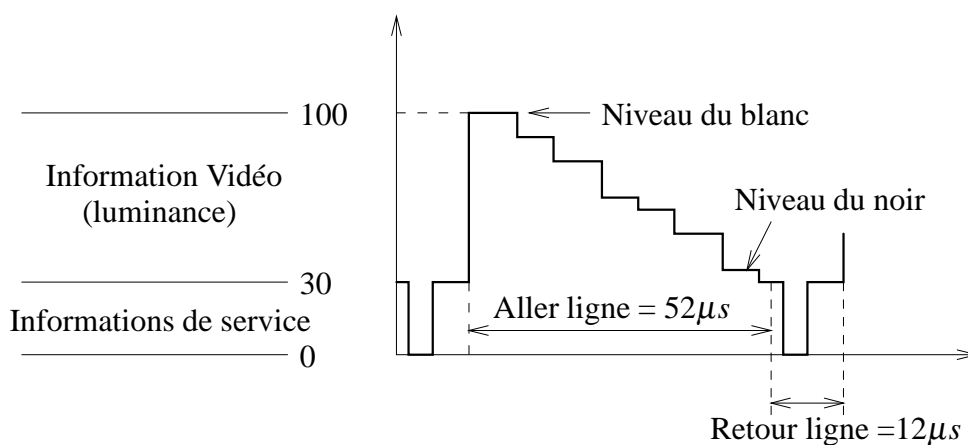


Figure 1.19: Une ligne du signal de luminance d'un téléviseur.

Le schéma plus général d'une chaîne acquisition-visualisation d'un signal vidéo est repris à la figure 1.20.

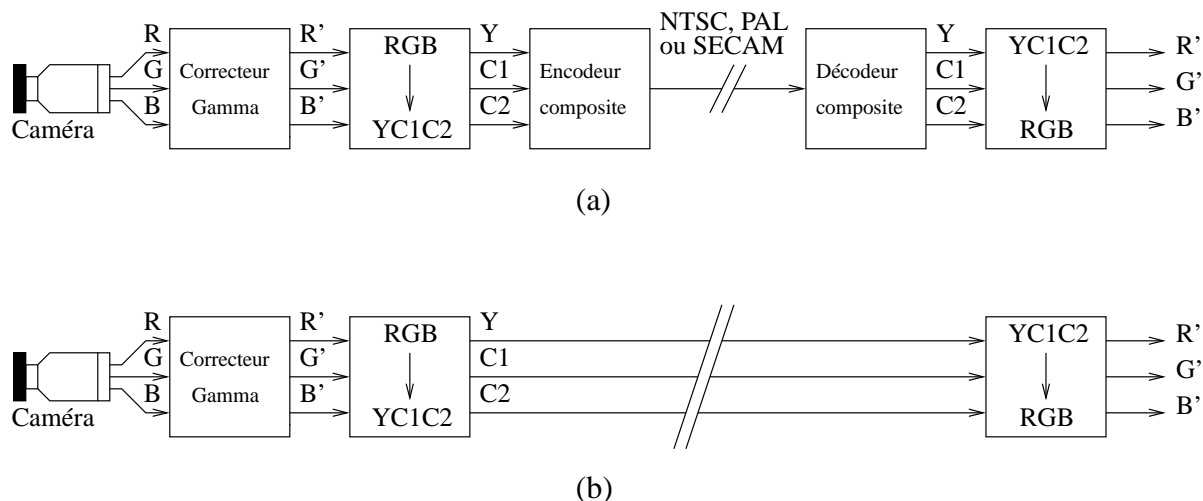


Figure 1.20: Chaîne de transmission des signaux de télévision composites.

Standards de télévision analogiques

Les États-Unis et l'Europe ont malheureusement défini des standards de télévision différents. Le système NTSC est utilisé aux États-Unis et le PAL en Europe (hormis la France et la Russie qui ont défini le système SÉCAM). Ces formats incompatibles sont à l'origine de bien des problèmes lors de la conversion de l'un vers l'autre.

Un signal PAL³ est un signal composite; il comporte l'information de luminance, les signaux de service (synchronisation de ligne et de trame, effacement) et deux signaux de chrominance.

Le signal de luminance Y est obtenu avec un fréquence maximale de $5 [MHz]$, ce qui représente la borne supérieure autorisée en norme G . Les information de couleur sont obtenus par différence. Il s'agit des signaux

- $R - Y$: signal de différence rouge, et
- $B - Y$: signal de différence bleu.

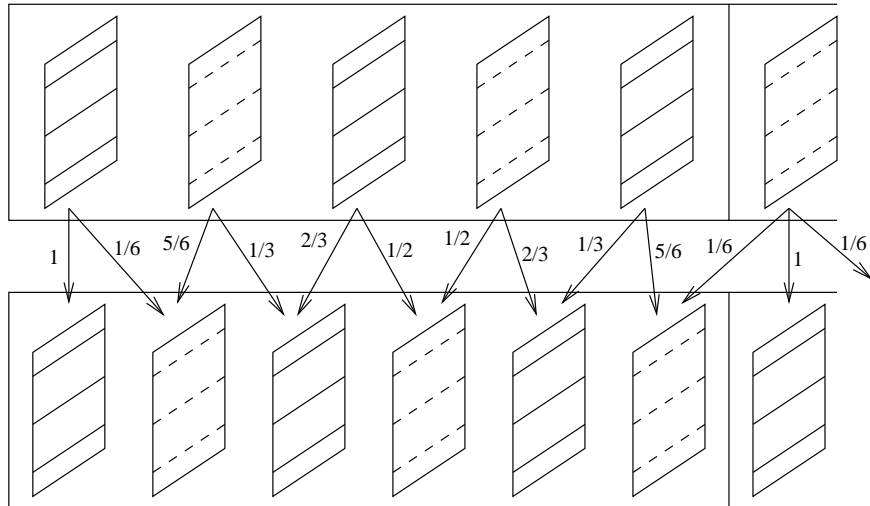
Les principales différences entre les standards de télévision analogiques sont reprises dans le tableau 1.3.

À titre d'exemple, la figure 1.21 montre comment s'effectue la conversion de la fréquence image entre un système $50 [Hz]$ et $60 [Hz]$. Seules les images obtenues à la cadence $10 [Hz]$ sont communes; les autres doivent être calculées par interpolation entre les images existantes.

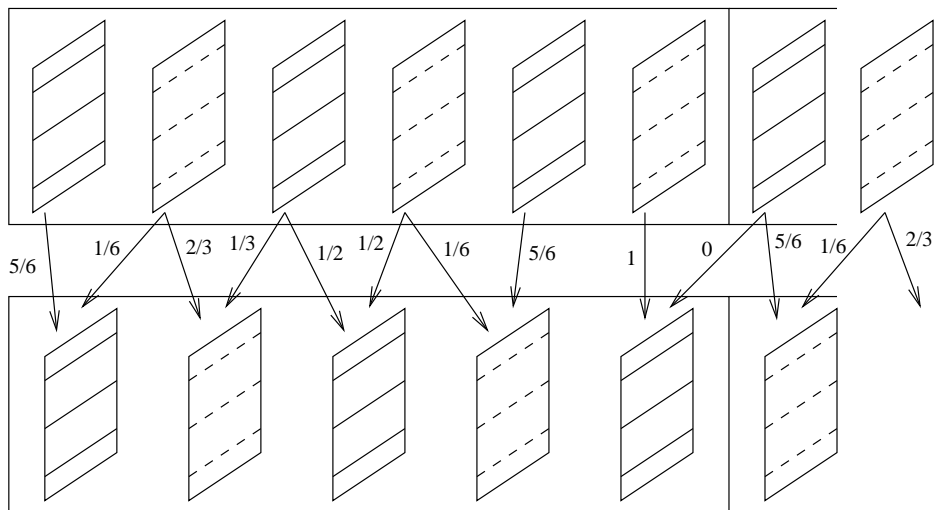
1.3.2 Structures de données

Les structures de données utilisées en traitement d'images sont multiples. On fait bien sûr un usage courant de la plupart des structures de données informatiques classiques: matrices (ou tableaux), vecteurs, arbres, listes, piles, ... Certaines structures de données ont été adaptées et enrichies pour le traitement d'images. Il s'agit par exemple de la structure en arbre quaternaire, appelé *quadtree*. Comme l'illustre la figure 1.22, cette structure attribue une valeur unique à un ensemble de pixels de valeurs homogènes. En l'absence d'homogénéité, la zone est à nouveau divisée en quatre parties égales.

³LAURENT [19] décrit en détail les principes de fonctionnement du PAL.



(a) Conversion 50Hz vers 60Hz



(b) Conversion 60Hz vers 50Hz

Figure 1.21: Conversion de la fréquence image.

Paramètres	PAL	NTSC	SÉCAM
Fréquence de trame [Hz]	50	59,94	50
Nombre de lignes par trame	625	525	625
Facteur de correction gamma γ	2,8	2,2	2,8
Porteuse audio [MHz]	QAM	4,5	FM
Porteuse couleur [MHz]	4,43	3,57	4,25 (+U) - 4,4 (-V)
Technique de modulation des signaux de couleur	QAM	QAM	FM
Largeur de bande de la luminance [MHz]	5,0 ou 5,5	4,2	6.0
Largeur de bande des chrominances [MHz]	1,3 (U et V)	1,3 (I) - 0,6 (Q)	>1,0 (U et V)

Table 1.3: Comparaison des 3 principaux formats de télévision analogiques.

00	01	1
02	03	
200	201	3
202	203	
22	23	

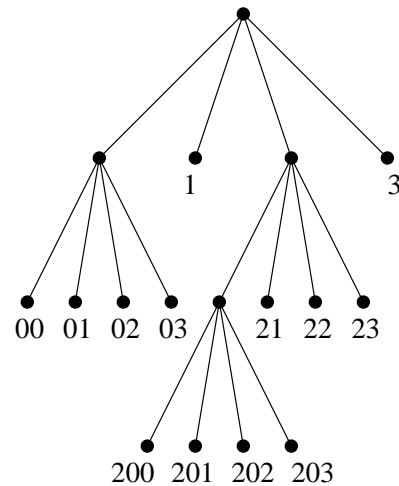
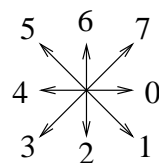
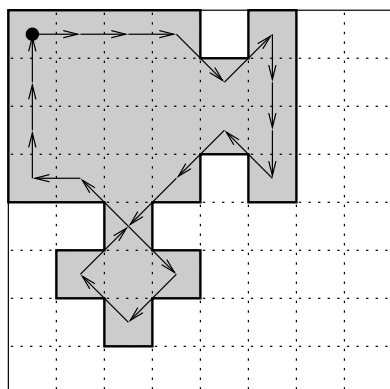


Figure 1.22: Découpage d'une image en arbre quaternaire.

S'agissant des contours, on peut adopter une description au moyen d'un *lacet*. Le lacet est une succession de directions qui indiquent à chaque fois la position d'un point voisin jusqu'à revenir au point de départ. En anglais, on parle de *chain code*.



Code : 00017222533135754666

Figure 1.23: Description d'une forme par suivi du contour.

Il existe également des structures de données *hiérarchiques* ou *multirésolution*, appelées généralement *pyramides*. Ces structures sont utilisées lorsqu'on désire coder une image à plusieurs niveaux de résolution différents.

Travailler à plusieurs niveaux de résolution peut s'avérer intéressant. En fait, une image 4 fois plus petite que l'originale sera traitée plus rapidement. On peut alors obtenir très rapidement une idée du résultat sur base de ce traitement rapide sans qu'il ne soit nécessaire de traiter l'image originale. La figure 1.4 montre une même image produite à plusieurs niveaux de résolution différents, ramenés à une même taille pour l'affichage.

Il n'y a que lorsque les détails sont seulement présents dans l'image à plus haute résolution qu'une telle approche n'est d'aucune utilité.

1.4 Traitement d'images industriel

Il est évidemment exclus de parcourir l'ensemble du traitement d'images industriel en quelques pages. Aussi nous contenterons-nous de décrire quelques opérations courantes et de commenter certains aspects du traitement industriel.

1.4.1 Remarques préalables

Décomposition d'une image en plans binaires

L'information contenue dans chaque bit de chaque pixel n'est pas la même. Pour s'en convaincre, il suffit de regarder le tableau 1.5. Ce tableau reprend l'image originale (sur 8 bits) et les différents plans de bit en partant du bit de poids fort (*Most Significant Bit* ou MSB) jusqu'au bit de poids faible (*Least Significant Bit* LSB). On peut voir que l'information locale, appelée *entropie* en théorie de l'information, n'est pas la même dans tous les plans. D'autre part, on devine tout de suite que les plans LSB contiennent moins d'information pertinente.

La différenciation des plans de bits se rattache à la question de la quantification. Une quantification sur 8 bits est usuelle mais on peut augmenter le nombre de niveaux de quantification jusqu'à 10 bits, valeurs que recommandent les intervenants du monde de l'édition. Il faut néanmoins garder à l'esprit que l'augmentation du nombre de pas de quantification accroît le besoin en mémoire et que l'information présente dans les derniers bits ne justifient pas toujours cet accroissement.

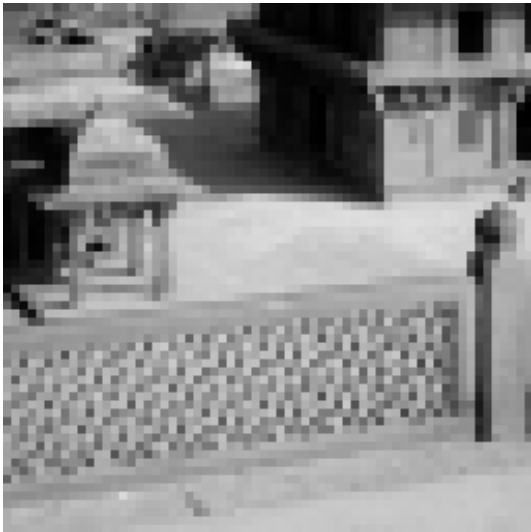


Table 1.4: Différents niveaux de résolution: chaque image est de résolution moitié de la précédente.

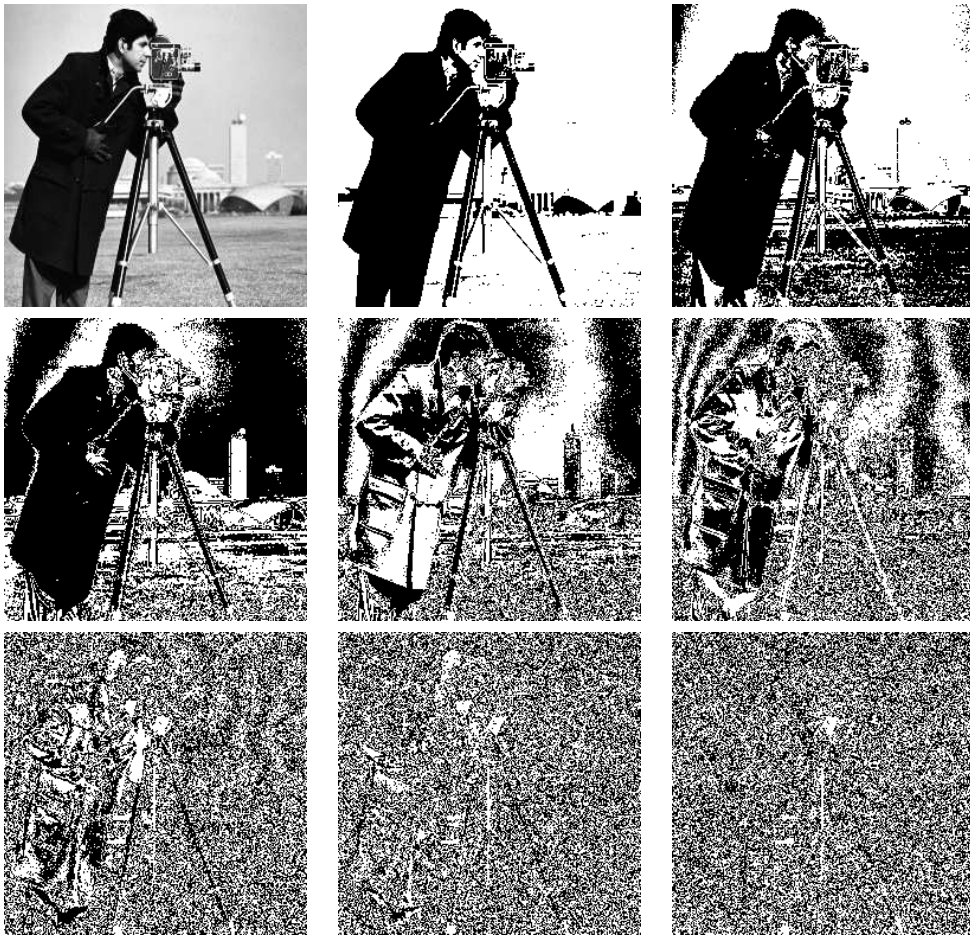


Table 1.5: Image originale et plans binaires en partant du MSB.

Critères de qualité

Un autre point important est la définition de mesures de distorsion permettant de comparer une image traitée de l'image originale. Soient une image originale f de taille $N \times N$ et \hat{f} l'image traitée.

On définit par exemple la métrique de distorsion d'erreur quadratique moyenne (MSE)

Définition 1 [Mean Square Error]

$$MSE = \frac{1}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} (f(j,k) - \hat{f}(j,k))^2 \quad (1.10)$$

On peut aussi définir des valeurs de type rapport signal à bruit.

Définition 2 [Signal to noise ratio]

$$SNR = \frac{\sum_{j=0}^{N-1} \sum_{k=0}^{N-1} (f(j,k))^2}{\sum_{j=0}^{N-1} \sum_{k=0}^{N-1} (f(j,k) - \hat{f}(j,k))^2} \quad (1.11)$$

Ou encore, en considérant une image de dynamique $[0, 255]$,

Définition 3 [Peak signal to noise ratio]

$$PSNR = \frac{N^2 \times 255}{\sum_{j=0}^{N-1} \sum_{k=0}^{N-1} (f(j,k) - \hat{f}(j,k))^2} \quad (1.12)$$

Il existe aussi des mesures de distorsion corrélatives ou basées sur l'évaluation de l'histogramme (pour une liste plus complète, voir par exemple [16, page 112])

Visualisation 3D d'une image 2D

Avant de traiter une image, il peut être intéressant de la visualiser sous la forme d'un volume 3D. On se choisit alors un axe de vue et on projette la profondeur de l'image suivant cet axe. La figure 1.24 illustre le principe⁴.

1.4.2 Segmentation d'images

La segmentation d'images consiste à décomposer une image en une série d'objets. Des exemples de segmentation sont illustrés aux figures 1.25 et 1.24. Dans le premier cas, il s'agit de segmenter les faces d'un parallélépipède rectangle. On peut voir que l'algorithme utilisé pour produire l'image segmentée ne traite pas correctement la zone d'ombre.

Le second cas est typique des séquences de vidéotéléphonie où l'on désire détecter les personnes de manière à mieux coder les zones qui leur correspondent.

1.4.3 Reconnaissance de caractères

Un processus complexe de segmentation est la reconnaissance de caractères. À partir d'un jeu de caractères connus, on cherche à identifier toute occurrence d'un caractère du jeu dans une image. La reconnaissance s'effectue habituellement en plusieurs étapes:

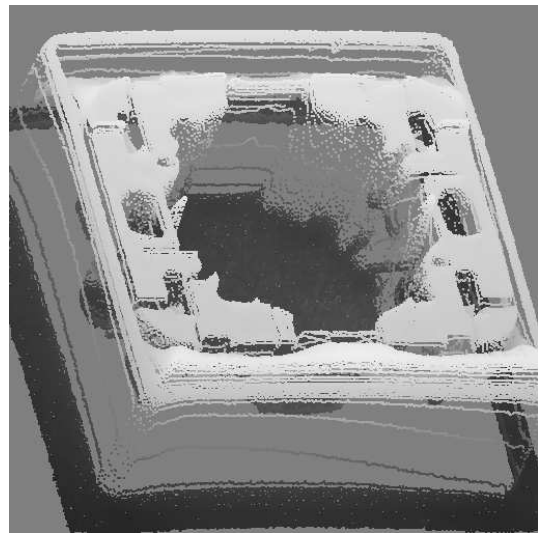
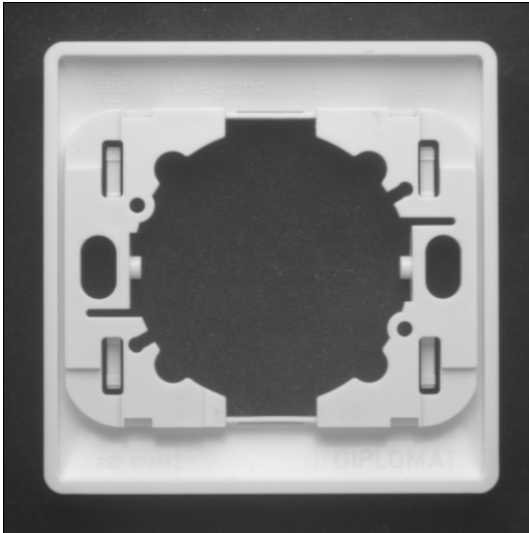


Figure 1.24: Une image et visualisation de la profondeur sous la forme d'une image 3D.

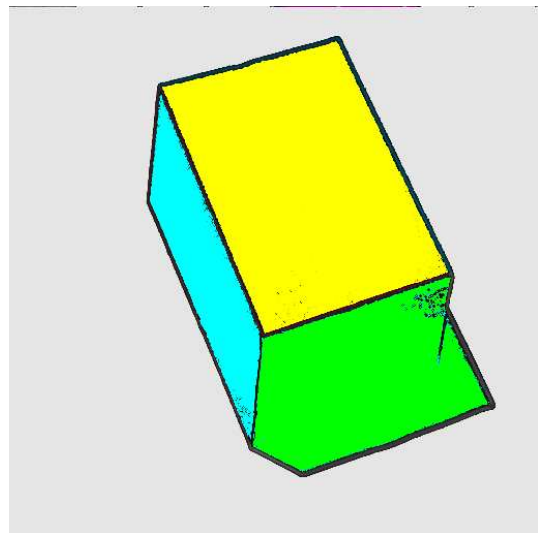
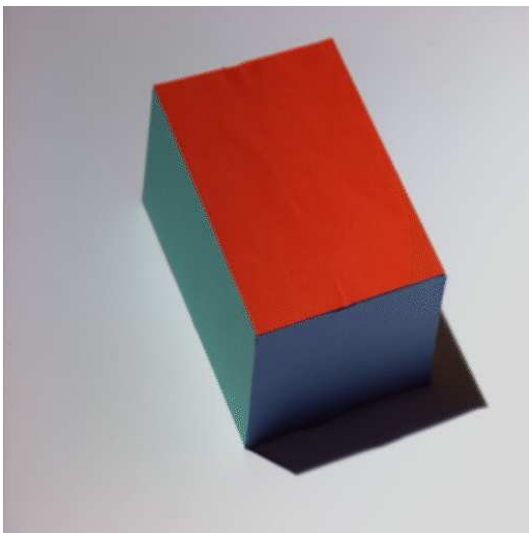


Figure 1.25: Segmentation d'un objet contenant une zone d'ombre.



Figure 1.26: Segmentation d'une image d'une séquence vidéo.

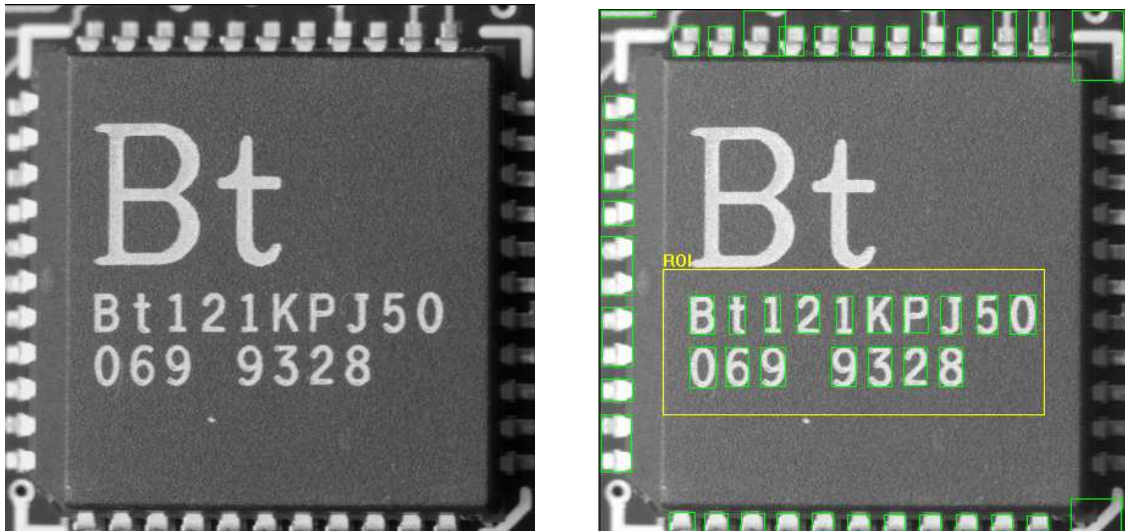


Figure 1.27: Reconnaissance de caractères par sélection d'une région d'intérêt.

- un opérateur ou un algorithme taillé sur mesure pour le type d'images à analyser spécifie une région à examiner; ce type de région, qui porte le nom de *Region Of Interest* (ROI), est destiné à sélectionner la partie significative de l'image (cf. figure 1.27).
- à l'intérieur de la région d'intérêt, un algorithme détecte les contours extérieurs de rectangle identifiées pour la reconnaissance de caractères.
- enfin, la reconnaissance de caractères s'effectue par mise en correspondance avec des gabarits connus. Cette étape porte le nom de *pattern matching*.

1.4.4 Autres opérations

D'autres opérations sont encore courantes dans la pratique industrielle, nous les citons ici:

- suppression d'une inhomogénéité d'éclairage,
- comptage de particules,
- contrôle de qualité par analyse des propriétés de l'image,
- analyse granulométrique, etc.

⁴Ces images, ainsi que d'autres encore, ont été produites à l'aide du logiciel EasyAccess de la société EURESYS.

Chapter 2

Transformations unitaires

2.1 Introduction

Les transformations discrètes sont couramment utilisées pour le traitement et l'analyse d'images. Ces transformées permettent, entre autres, des applications de filtrage, de compression d'images, de description d'images, ... tant pour des approches déterministes que stochastiques.

Elles furent activement étudiées dès la fin des années 1960. Aujourd'hui, on en connaît les caractéristiques des plus fondamentales (propriétés décorrélatives) jusqu'aux détails d'implémentation (effet d'échantillonnage, erreurs d'arrondi, etc). C'est évidemment une raison supplémentaire pour y recourir.

2.2 Bases théoriques du calcul matriciel

Considérons l'image échantillonnée f , représentée par une matrice de $M \times N$ points ou pixels. Elle est notée $\underline{f}(m,n)$ ou plus simplement \underline{f} avec $m = 0, \dots, M-1$ et $n = 0, \dots, N-1$:

$$\underline{f} = \begin{bmatrix} f(0,0) & \dots & f(0,N-1) \\ \vdots & & \vdots \\ f(M-1,0) & \dots & f(M-1,N-1) \end{bmatrix} \quad (2.1)$$

Soient \underline{P} et \underline{Q} deux matrices dites de transformation de dimensions respectives $M \times M$ et $N \times N$. Ces deux matrices permettent de transformer la matrice \underline{f} en la matrice \underline{F} de dimension $M \times N$ de telle sorte que

$$\underline{F} = \underline{P}\underline{f}\underline{Q} \quad (2.2)$$

ce qui peut également s'écrire

$$\underline{F}(u,v) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \underline{P}(u,m) \underline{f}(m,n) \underline{Q}(n,v) \quad (2.3)$$

tel que

$$u = 0, \dots, M-1 \quad v = 0, \dots, N-1 \quad (2.4)$$

Si les matrices \underline{P} et \underline{Q} sont non-singulières (c'est-à-dire qu'elles ont un déterminant non nul) les matrices inverses \underline{P}^{-1} et \underline{Q}^{-1} existent et la transformée inverse est donnée par

$$\underline{f} = \underline{P}^{-1}\underline{F}\underline{Q}^{-1} \quad (2.5)$$

Il est bon, avant de décrire quelques transformations particulières, de rappeler quelques propriétés concernant le calcul matriciel. La matrice transposée de la matrice \underline{M} est notée \underline{M}^T .

Propriétés

- \underline{M} est symétrique si $\underline{M} = \underline{M}^T$
- \underline{M} est orthogonale si $\underline{M}^T \underline{M} = \underline{I}$ où \underline{I} est la matrice identité.
- Si \underline{M} est symétrique, réelle et orthogonale, alors $\underline{M}^{-1} = \underline{M}$
- Une matrice complexe \underline{C} est hermitienne si $\underline{C}^{*T} = \underline{C}$, où \underline{C}^* est la matrice complexe conjuguée de \underline{C} .
- Une matrice complexe \underline{C} est unitaire si $\underline{C}^{*T} \underline{C} = \underline{I}$
- Si \underline{C} est une matrice complexe, carrée, hermitienne et unitaire, alors $\underline{C}^{-1} = \underline{C}$

Revenons à présent aux matrices de transformation \underline{P} et \underline{Q} . Si \underline{P} et \underline{Q} sont symétriques, réelles et orthogonales, alors

$$\underline{F} = \underline{P} \underline{f} \underline{Q} \quad \underline{f} = \underline{P} \underline{F} \underline{Q} \quad (2.6)$$

On parle alors de *transformations orthogonales* ou *unitaires*. Les transformations unitaires ont la propriété intéressante d'utiliser le même algorithme pour la transformation inverse que pour la transformation directe. Si les matrices \underline{P} et \underline{Q} sont complexes, la relation 2.6 tient toujours à condition que \underline{P} et \underline{Q} soient hermitiennes et unitaires.

2.3 La transformée de FOURIER discrète

Tout comme les séries de Fourier fournissent le point de départ pour les transformations et l'analyse de fonctions périodiques, la transformée de FOURIER discrète (*Discrete FOURIER Transform* ou DFT) est l'équivalent pour l'analyse de vecteurs. Elle effectue un passage du domaine spatial à un domaine spectral.

Un vecteur à N composantes est calculé comme une combinaison linéaire de N vecteurs de base. Les éléments de ces vecteurs de base sont tous des puissances de l'exponentielle imaginaire

$$W = e^{2\pi j/N} \quad (2.7)$$

ces vecteurs (lignes) valent

$$\left(1, \vec{W}^k, \vec{W}^{2k}, \dots, \vec{W}^{(N-1)k}\right), \quad k = 0, 1, \dots, N-1 \quad (2.8)$$

À partir de ces expressions, on peut définir la matrice de transformation $\underline{\Phi}_{JJ}$ de dimension $J \times J$:

$$\underline{\Phi}_{JJ}(k, l) = \frac{1}{J} \exp\left(-j \frac{2\pi}{J} kl\right) \quad k, l = 0, 1, \dots, J-1 \quad (2.9)$$

Définition 4 La transformée de FOURIER discrète est définie, étant données les deux matrices de transformation $\underline{P} = \underline{\Phi}_{MM}$ et $\underline{Q} = \underline{\Phi}_{NN}$, par

$$\underline{F} = \underline{\Phi}_{MM} \underline{f} \underline{\Phi}_{NN} \quad (2.10)$$

ce qui peut également se mettre sous la forme

$$\underline{F}(u, v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \exp\left[-2\pi j \left(\frac{mu}{M} + \frac{nv}{N}\right)\right] \quad (2.11)$$

avec

$$u = 0, 1, \dots, M-1 \quad v = 0, 1, \dots, N-1 \quad (2.12)$$

La matrice de transformation inverse $\underline{\Phi}_{JJ}^{-1}$ est donnée par

$$\underline{\Phi}_{JJ}^{-1}(k, l) = \exp\left(j\frac{2\pi}{J}kl\right) \quad (2.13)$$

et la transformée de FOURIER discrète inverse est donnée par

$$\underline{f}(m, n) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \underline{F}(u, v) \exp\left[2\pi j\left(\frac{mu}{M} + \frac{nv}{N}\right)\right] \quad (2.14)$$

avec

$$m = 0, 1, \dots, M-1 \quad n = 0, 1, \dots, N-1 \quad (2.15)$$

2.3.1 Périodicité intrinsèque

La périodicité est une propriété importante de la transformée de FOURIER discrète. La transformée est définie par une matrice d'éléments $\underline{F}(u, v)$ pour $u = 0, \dots, M-1$ et $v = 0, \dots, N-1$. Si on permet aux indices u et v de prendre d'autres valeurs, on obtient une transformée périodique et une image périodique:

$$\underline{F}(u, -v) = \underline{F}(u, N-v) \quad \underline{F}(-u, v) = \underline{F}(M-u, v) \quad (2.16)$$

$$\underline{f}(-m, n) = \underline{f}(M-m, n) \quad \underline{f}(m, -n) = \underline{f}(m, N-n) \quad (2.17)$$

Plus généralement,

$$\underline{F}(aM+u, bN+v) = \underline{F}(u, v) \quad \underline{f}(aM+m, bN+n) = \underline{f}(m, n) \quad (2.18)$$

où a et b sont des entiers. Ces propriétés ne doivent pas nous étonner car elles résultent de l'échantillonnage dans les deux plans.

La figure 2.1 montre une image ainsi que le module de sa transformée de FOURIER calculée par le logiciel *Matlab* sur base de la formule 2.10.



Figure 2.1: L'image *Lena* et le module de sa transformée de FOURIER discrète.

Une remarque toute particulière est à faire au sujet de la visualisation de la transformée de FOURIER discrète. La composante fréquentielle $\underline{F}(0,0)$, appelée *composante continue* ou *composante DC*, se trouve en haut et à gauche de l'image alors qu'on a l'habitude de voir cette composante située au milieu de l'image. En utilisant les propriétés de périodicité, il est possible d'observer le module de la transformée de FOURIER d'une manière plus conventionnelle (cf. figure 2.2). Le spectre modifié par décalage de l'image *Lena* est donné à la figure 2.3.

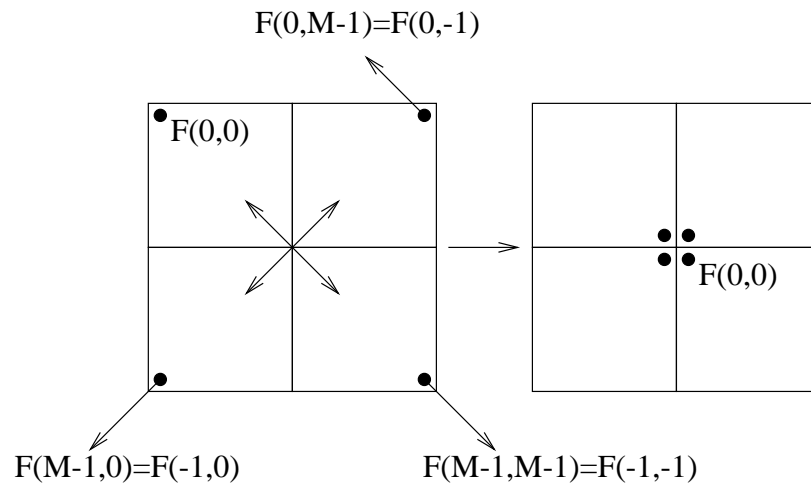


Figure 2.2: Décalage du spectre pour arriver à centrer l'origine.

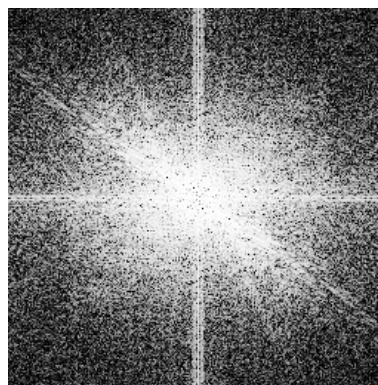


Figure 2.3: Visualisation du module du spectre de l'image *Lena* après centrage de l'origine.

2.3.2 Note sur l'implémentation

En majorité, la DFT dispose des mêmes propriétés que son homologue continu (linéarité, translation spatiale, multiplication, convolution, ...), mis à part la nature périodique de l'image et de sa transformée suite aux échantillonnages effectués dans les deux domaines. On peut donc à souhait profiter des propriétés connues pour faciliter le calcul. En plus de ces propriétés, la nature discrète de la DFT permet de bénéficier d'autres avantages.

Orthogonalité et vecteurs propres

Les vecteurs W^k qui composent la matrice de FOURIER sont orthogonaux entre eux. La démonstration de cette propriété est simple. En effet, le produit scalaire de deux vecteurs (distincts) quelconques donne

$$\vec{W}^k \cdot \vec{W}^l = \sum_{i=0}^{N-1} (W^k)^i (W^{l*})^i = \frac{(W^k W^{l*})^N - 1}{W^k W^{l*} - 1} \quad (2.19)$$

Le numérateur est toujours nul car $\forall k \neq l$,

$$(W^k W^{l*})^N = (e^{2\pi jk/N} e^{-2\pi jl/N})^N = e^{2\pi j(k-l)} = 1 \quad (2.20)$$

Quant au dénominateur, il est non nul car $k \neq l$.

Les vecteurs de base sont bien donc orthogonaux entre eux. De plus, ce sont les valeurs propres de la matrice de FOURIER (voir [36] pour la démonstration), ce qui signifie que l'application de la matrice de FOURIER diagonalise toute image périodique.

Algorithme rapide

La transformée de FOURIER discrète peut être implémentée de manière efficace malgré le fait que la matrice Φ_{NN} soit pleine. La raison pour laquelle il est possible de réduire le nombre d'opérations intervenant dans le calcul matriciel tient au fait que tous les vecteurs de base sont des multiples d'un même vecteur. Cela permet de factoriser plus efficacement la matrice en série de matrices creuses. Les algorithmes qui en découlent font partie de la famille des transformées de FOURIER rapides (notée FFT).

Détails pratiques. Remarquons que la formule 2.11 peut encore s'écrire sous la forme

$$\underline{F}(u, v) = \frac{1}{M} \sum_{m=0}^{M-1} \left[\frac{1}{N} \sum_{n=0}^{N-1} \underline{f}(m, n) \exp\left(-2\pi j \frac{nv}{N}\right) \right] \exp\left(-2\pi j \frac{mu}{M}\right) \quad (2.21)$$

Le terme entre crochets correspond à la transformée de FOURIER discrète à une dimension de la $m^{\text{ième}}$ ligne de la matrice \underline{f} . Chaque ligne de \underline{f} est remplacée par sa transformée de FOURIER discrète et la transformée de FOURIER discrète de chaque colonne est alors calculée. On peut tout aussi bien traiter d'abord les colonnes puis les lignes.

Le calcul d'une FFT est le plus efficace lorsque N est une puissance de 2: $N = 2^k$. Lorsque la taille de l'image n'est pas une puissance de 2, il est d'usage de compléter l'image en lui ajoutant des pixels de valeur nulle jusqu'à obtention de la taille adéquate. Cette démarche n'est pas sans entraîner certains inconvénients car elle modifie la résolution spectrale du résultat. On peut, pour éviter ce désagrément, recourir à des implémentations plus complexes de FFT qui particularisent le calcul pour des puissances autres que 2. Lorsqu'une telle implémentation n'est pas disponible, on pourra toujours recourir à l'application pure et simple de la définition, mais au détriment du temps de calcul!

Quelque soit l'implémentation choisie, le calcul d'une transformée de FOURIER manipule des nombres complexes et il fournit des valeurs complexes. La manipulation n'en est pas toujours comode. C'est pourquoi, certains chercheurs ont proposé des transformées qui reposent sur des simplifications de la transformée de FOURIER. La transformée de HADAMARD et la transformée en cosinus discrète font partie de ces familles de transformées à valeurs entières ou réelles.

2.4 La transformée de HADAMARD

Si les matrices de transformation \underline{P} et \underline{Q} sont des matrices de HADAMARD, alors $\underline{F} = \underline{P}\underline{f}\underline{Q}$ est appelée transformée de HADAMARD de l'image \underline{f} . Une matrice de HADAMARD \underline{H}_{JJ} est une matrice symétrique, de dimension $J \times J$, dont les éléments sont tous égaux à ± 1 et définie comme suit. La matrice d'HADAMARD du second ordre est

$$\underline{H}_{22} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (2.22)$$

La matrice de HADAMARD d'ordre 2^k peut être écrite sous la forme

$$\underline{H}_{2^k 2^k} = \begin{bmatrix} \underline{H}_{JJ} & \underline{H}_{JJ} \\ \underline{H}_{JJ} & -\underline{H}_{JJ} \end{bmatrix} \quad (2.23)$$

Des matrices de HADAMARD d'ordre autre que 2^k existent mais elles sont rarement utilisées en traitement d'images. L'inverse d'une matrice de Hadamard est donnée par

$$\underline{H}_{JJ}^{-1} = \frac{1}{J}\underline{H}_{JJ} \quad (2.24)$$

Définition 5 La transformée de HADAMARD et son inverse sont alors définies par

$$\underline{F} = \underline{H}_{MM}\underline{f}\underline{H}_{NN} \quad \underline{f} = \frac{1}{MN}\underline{H}_{MM}\underline{F}\underline{H}_{NN} \quad (2.25)$$

Étant donnée la composition des matrices de transformation, la transformée de HADAMARD et son inverse peuvent être calculée uniquement par des additions. La transformée de HADAMARD est parfois appelée transformée de WALSH-HADAMARD. Elle est occasionnellement utilisée dans des applications de codage d'images.

2.5 La transformée en cosinus discrète

Il existe 4 définitions de la transformée en cosinus discrète, parfois notées DCT-1, DCT-2, DCT-3, et DCT-4. Ces définitions diffèrent par les conditions limites qu'elles imposent aux bords¹.

La transformée la plus utilisée en traitement et compression d'images est la DCT-2 dont la définition suit. D'un point de vue théorique, cette variante présuppose que l'image a été miroirisée le long de ses bords car, dans ce cas, la transformée de FOURIER de la séquence miroirisée est rigoureusement égale à la DCT-2 de la séquence originale.

Nous faisons l'hypothèse que l'image est carrée de dimension $N \times N$. Soit la matrice de transformation $\underline{C}_{NN}(k, l)$ définie par

$$\underline{C}_{NN}(k, l) = \begin{cases} \frac{1}{\sqrt{N}} & l = 0 \\ \sqrt{\frac{2}{N}} \cos \left[\frac{(2k+1)l\pi}{2N} \right] & \text{sinon} \end{cases} \quad (2.26)$$

¹STRANG [36] fournit le détail de ces conditions.

Définition 6 La transformée en cosinus discrète *DCT* et son inverse sont données par

$$\underline{F} = \underline{C}_{NN} \underline{f} \underline{C}_{NN}^T \quad \underline{f} = \underline{C}_{NN}^T \underline{F} \underline{C}_{NN} \quad (2.27)$$

Une autre forme pour la transformée en cosinus discrète est donnée par

$$\underline{F}(u, v) = \frac{2c(u)c(v)}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \underline{f}(m, n) \cos\left(\frac{2m+1}{2N}u\pi\right) \cos\left(\frac{2n+1}{2N}v\pi\right) \quad (2.28)$$

$$u = 0, 1, \dots, N-1 \quad v = 0, 1, \dots, N-1 \quad (2.29)$$

où

$$c(k) = \begin{cases} \frac{1}{\sqrt{2}} & k = 0 \\ 1 & \text{sinon} \end{cases} \quad (2.30)$$

et la transformée inverse est donnée par

$$\underline{f}(m, n) = \frac{2}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} c(u)c(v) \underline{F}(u, v) \cos\left(\frac{2m+1}{2N}u\pi\right) \cos\left(\frac{2n+1}{2N}v\pi\right) \quad (2.31)$$

$$m = 0, 1, \dots, N-1 \quad n = 0, 1, \dots, N-1 \quad (2.32)$$

2.5.1 Note sur l'implémentation

L'implémentation de la transformée en cosinus discrète peut être basée sur la transformée de FOURIER. Pour une ligne (ou colonne), les N coefficients de la DCT peuvent être calculés en utilisant une FFT sur $2N$ points. On montre en effet aisément que la DCT d'une séquence est équivalente à la transformée de FOURIER de cette même séquence préalablement reproduite par symétrie autour de l'origine.

Cette miroirisation a au moins un effet négatif: elle induit une perte de discrimination pour les orientations diagonales. Cet effet n'est pas trop gênant pour le codage d'images car on sait qu'un pourcentage important de l'énergie d'une image se situe le long des axes principaux, mais il est préférable de recourir à une analyse de FOURIER pour détecter, par exemple, les orientations de textures.

La transformée en cosinus discrète constitue la pierre angulaire du format de compression d'image JPEG.

2.6 Transformée de FOURIER

2.6.1 Définition

La transformée de FOURIER d'une image $f(x, y)$ est définie par

$$\mathcal{F}(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-2\pi j(xu+yv)} dx dy \quad (2.33)$$

où u et v désignent les fréquences spatiales dans les directions x et y . À partir de la transformée de FOURIER, il est possible de reconstituer exactement l'image originale en prenant la transformée inverse

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{F}(u, v) e^{2\pi j(ux+vy)} du dv \quad (2.34)$$

Pour que la transformée d'une image existe, la condition suivante suffit

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(x,y)| dx dy < +\infty \quad (2.35)$$

La transformée de FOURIER nous fournit une interprétation intéressante puisqu'elle décompose l'image en composantes fréquentielles définie sur $[-\infty, +\infty] \times [-\infty, +\infty]$. $f(x,y)$ et $\mathcal{F}(u,v)$ forment une paire de transformée de FOURIER représentée par

$$f(x,y) \rightleftharpoons \mathcal{F}(u,v) \quad (2.36)$$

En général, $\mathcal{F}(u,v)$ est une fonction de u et de v à valeurs complexes. Nous pouvons donc l'exprimer sous la forme

$$\mathcal{F}(u,v) = \|\mathcal{F}(u,v)\| e^{j\theta(u,v)} \quad (2.37)$$

où $\|\mathcal{F}(u,v)\|$ est appelé *module* de $\mathcal{F}(u,v)$ ou *spectre fréquentiel* de l'image $f(x,y)$ et $\theta(u,v)$ est la *phase* de $\mathcal{F}(u,v)$ ou *spectre de phase* de $\mathcal{F}(u,v)$. Dans le cas particulier important où $f(x,y)$ est une fonction à valeurs réelles, nous avons

$$\mathcal{F}(-u,-v) = \mathcal{F}^*(u,v) \quad (2.38)$$

Dès lors,

$$\|\mathcal{F}(-u,-v)\| = \|\mathcal{F}(u,v)\| \quad (2.39)$$

$$\theta(-u,-v) = -\theta(u,v) \quad (2.40)$$

On peut en déduire deux propriétés importantes d'une image à valeurs réelles:

1. le spectre fréquentiel de l'image est symétrique par rapport à l'origine du système d'axes $u-v$. C'est-à-dire que la connaissance d'un demi-plan suffit. Ce résultat est satisfaisant car si, dans le plan spatial, on dispose de $M \times N$ variables indépendantes, la transformée de FOURIER fournit seulement $(M \times N)/2$ variables indépendantes mais ces variables sont complexes.
2. le spectre de phase de l'image est anti-symétrique par rapport à l'origine du système d'axes $u-v$.

2.6.2 Propriétés

1. Séparabilité

(a) En permutant l'ordre d'intégration dans (2.33), nous avons

$$\mathcal{F}(u,v) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(x,y) e^{-2\pi jxu} dx \right] e^{-2\pi jyv} dy \quad (2.41)$$

La transformée de FOURIER d'une image $f(x,y)$ peut se réaliser en deux étapes: (i) transformée de FOURIER unidimensionnelle de la fonction $f(x,y)$ pour tout y fixé, transformant la variable x en la variable u et (ii) transformée de FOURIER unidimensionnelle de la fonction obtenue pour tout u fixé, transformant la variable y en la variable v .

(b) Si l'image $f(x, y)$ est spatialement séparable², c'est-à-dire

$$f(x, y) = f_x(x) f_y(y) \quad (2.42)$$

alors

$$\mathcal{F}(u, v) = \mathcal{F}_x(u) \mathcal{F}_y(v) \quad (2.43)$$

Dans ce cas, la transformée de FOURIER de l'image $f(x, y)$ est égale au produit des transformées de FOURIER unidimensionnelles des fonctions $f_x(x)$ et $f_y(y)$.

2. Linéarité

Soient $f_1(x, y) \Leftrightarrow \mathcal{F}_1(u, v)$ et $f_2(x, y) \Leftrightarrow \mathcal{F}_2(u, v)$. Alors, pour toutes constantes c_1 et c_2 ,

$$c_1 f_1(x, y) + c_2 f_2(x, y) \Leftrightarrow c_1 \mathcal{F}_1(u, v) + c_2 \mathcal{F}_2(u, v) \quad (2.44)$$

3. Homothétie

Si $f(x, y) \Leftrightarrow \mathcal{F}(u, v)$, alors

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} \mathcal{F}\left(\frac{u}{a}, \frac{v}{b}\right) \quad (2.45)$$

Il est à noter que l'image $f(ax, by)$ correspond à l'image $f(x, y)$ compressée dans l'espace par un facteur a dans la direction x et par un facteur b dans la direction y . Une compression dans le domaine spatial équivaut donc à une extension dans le domaine fréquentiel et vice-versa.

4. Dualité

Si $f(x, y) \Leftrightarrow \mathcal{F}(u, v)$, alors

$$\mathcal{F}(x, y) \Leftrightarrow f(-u, -v) \quad (2.46)$$

5. Translation spatiale

Si $f(x, y) \Leftrightarrow \mathcal{F}(u, v)$, alors

$$f(x - x_0, y - y_0) \Leftrightarrow \mathcal{F}(u, v) e^{-2\pi j(x_0 u + y_0 v)} \quad (2.47)$$

6. Translation fréquentielle

Si $f(x, y) \Leftrightarrow \mathcal{F}(u, v)$, alors

$$f(x, y) e^{j2\pi(u_0 x + v_0 y)} \Leftrightarrow \mathcal{F}(u - u_0, v - v_0) \quad (2.48)$$

7. Aires

Si $f(x, y) \Leftrightarrow \mathcal{F}(u, v)$, alors

$$\mathcal{F}(0, 0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy \quad (2.49)$$

Le coefficient $\mathcal{F}(0, 0)$, appelé parfois *composante DC*, est la somme des pixels de l'image. Sa dynamique est donc très importante par rapport aux autres coefficients. Ceci pose des problèmes pratiques qui amène à traiter séparément ce coefficient.

$$f(0, 0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{F}(u, v) du dv \quad (2.50)$$

²Ce n'est généralement pas le cas pour des images naturelles. Ce cas de figure peut néanmoins se produire pour certaines images de synthèse.

8. Convolution

Le produit de convolution de deux fonctions $f(x, y)$ et $g(x, y)$ est défini par

$$(f \otimes g)(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) g(x - \alpha, y - \beta) d\alpha d\beta \quad (2.51)$$

Si $f(x, y) \rightleftharpoons \mathcal{F}(u, v)$ et $g(x, y) \rightleftharpoons \mathcal{G}(u, v)$, alors

$$(f \otimes g)(x, y) \rightleftharpoons \mathcal{F}(u, v) \mathcal{G}(u, v) \quad (2.52)$$

La convolution de deux images dans le domaine spatial est transformée dans le domaine fréquentiel en la multiplication de leur transformée de FOURIER respective.

9. Multiplication

Si $f(x, y) \rightleftharpoons \mathcal{F}(u, v)$ et $g(x, y) \rightleftharpoons \mathcal{G}(u, v)$, alors

$$f(x, y) g(x, y) \rightleftharpoons (\mathcal{F} \otimes \mathcal{G})(u, v) \quad (2.53)$$

La multiplication de deux images dans le domaine spatial est transformée dans le domaine fréquentiel en la convolution de leur transformée de FOURIER respective.

2.6.3 Illustrations

La fonction Rectangle

Considérons l'image $f(x, y)$ définie par

$$f(x, y) = A \text{rect}_{[a, b]}(x, y) \quad (2.54)$$

où

$$\text{rect}_{[a, b]}(x, y) = \begin{cases} 1 & |x| < \frac{a}{2}, |y| < \frac{b}{2} \\ 0 & \text{ailleurs} \end{cases} \quad (2.55)$$

Cette fonction est appelée *fonction Rectangle*. L'image $f(x, y)$ vaut donc A à l'intérieur du rectangle de longueur a et de largeur b dont les côtés sont parallèles aux axes x et y , et zéro partout ailleurs. La transformée de FOURIER de l'image $f(x, y)$ est donnée par

$$\mathcal{F}(u, v) = \int_{-a/2}^{+a/2} dx \int_{-b/2}^{+b/2} dy A e^{-2\pi j(xu + yv)} \quad (2.56)$$

$$= Aab \left(\frac{\sin(\pi au)}{\pi au} \right) \left(\frac{\sin(\pi bv)}{\pi bv} \right) \quad (2.57)$$

Il s'agit du produit de deux sinus cardinaux. Ce résultat aurait pu être obtenu en remarquant que l'image $f(x, y)$ est séparable spatialement et en utilisant la propriété 2.43 de séparabilité. Nous avons donc la paire de transformées de FOURIER

$$\text{rect}_{[a, b]}(x, y) \rightleftharpoons ab \left(\frac{\sin(\pi au)}{\pi au} \right) \left(\frac{\sin(\pi bv)}{\pi bv} \right) \quad (2.58)$$

Une représentation de l'image $f(x, y)$ est donné à la figure 2.4 tandis que le module de sa transformée de FOURIER est représenté à la figure 2.5.

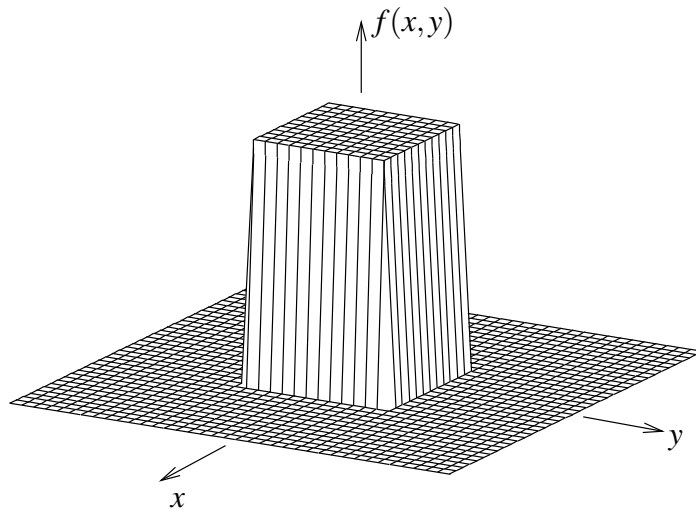


Figure 2.4: Illustration de la fonction Rectangle.

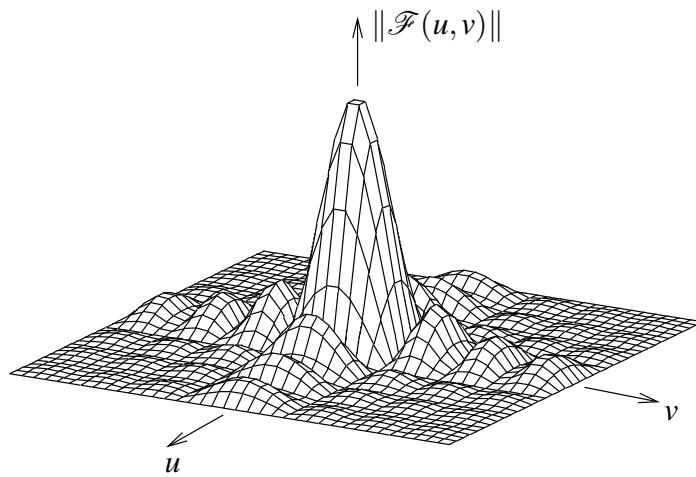


Figure 2.5: Module de la transformée de FOURIER de la fonction Rectangle.

La fonction Disque

Considérons l'image $f(x,y)$ définie par

$$f(x,y) = A \text{Disque}_R(x,y) \quad (2.59)$$

où

$$\text{Disque}_R(x,y) = \begin{cases} 1 & \sqrt{x^2+y^2} < R \\ 0 & \text{ailleurs} \end{cases} \quad (2.60)$$

est appelée *fonction disque*. L'image $f(x,y)$ vaut donc A sur le disque de rayon R et zéro partout ailleurs. La transformée de FOURIER de l'image $f(x,y)$ est donnée par

$$\mathcal{F}(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) e^{-2\pi j(xu+yv)} dx dy \quad (2.61)$$

$$= \int_0^R r dr \int_0^{2\pi} A e^{-2\pi jr(u \cos \theta + v \sin \theta)} d\theta \quad (2.62)$$

où nous avons effectué un changement de variables des coordonnées cartésiennes vers les coordonnées polaires. En se basant sur les propriétés des fonctions de BESSEL, nous obtenons le résultat suivant

$$\mathcal{F}(u,v) = AR \frac{J_1\left(2\pi R \sqrt{u^2+v^2}\right)}{\sqrt{u^2+v^2}} \quad (2.63)$$

où $J_1(r)$ est la fonction de BESSEL d'ordre 1. Il est à noter que la transformée de FOURIER de l'image $f(x,y)$ est purement réelle et à symétrie radiale, c'est-à-dire qu'elle ne dépend que de la distance $\sqrt{u^2+v^2}$ à l'origine. Nous avons donc la paire de transformées de FOURIER

$$\text{Disque}_R(x,y) \Leftrightarrow R \frac{J_1\left(2\pi R \sqrt{u^2+v^2}\right)}{\sqrt{u^2+v^2}} \quad (2.64)$$

L'image $f(x,y)$ est représentée à la figure 2.6 tandis que le module de sa transformée de FOURIER est représenté à la figure 2.7.

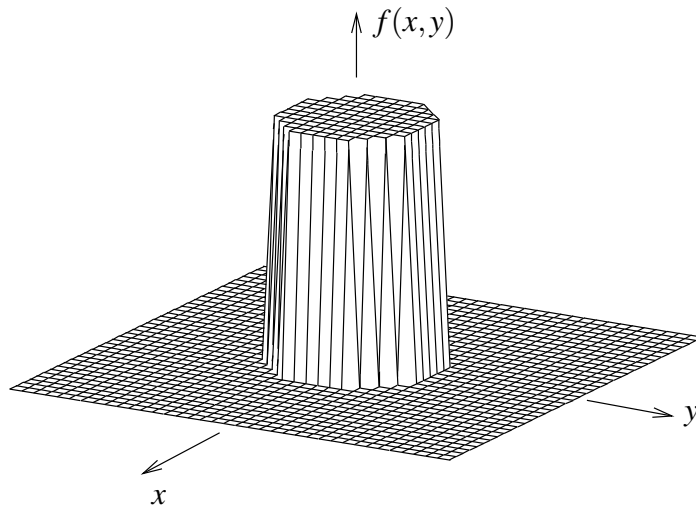


Figure 2.6: Illustration de la fonction Disque.

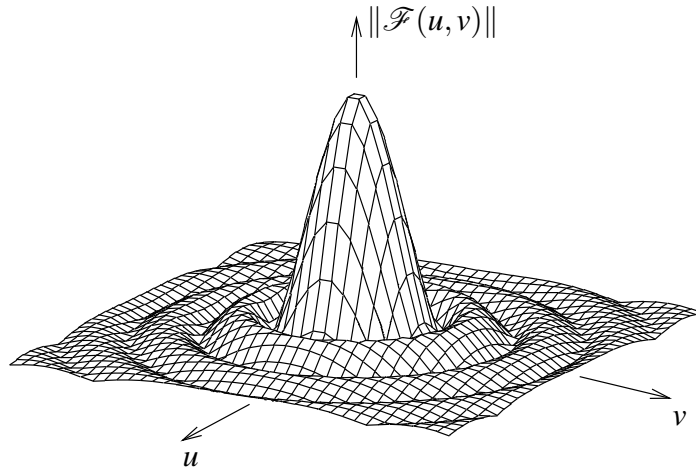


Figure 2.7: Module de la transformée de FOURIER de la fonction Disque.

Dual de la fonction Rectangle

Pour rappel, nous avons la paire de transformées de FOURIER suivante

$$\text{rect}_{[a,b]}(x,y) \Leftrightarrow ab \left(\frac{\sin(\pi au)}{\pi au} \right) \left(\frac{\sin(\pi bv)}{\pi bv} \right) \quad (2.65)$$

En utilisant la propriété de dualité (relation 2.46) de la transformée de FOURIER, nous obtenons

$$\left(\frac{\sin(\pi ax)}{\pi ax} \right) \left(\frac{\sin(\pi by)}{\pi by} \right) \Leftrightarrow \frac{1}{ab} \text{rect}_{[a,b]}(u,v) \quad (2.66)$$

La nouvelle image ainsi obtenue est dite à bande limitée car son spectre en fréquence est limité à une région finie du plan $u - v$. En l'occurrence, cette région est, dans ce cas-ci, un rectangle de longueur a et de largeur b centré à l'origine. D'après la définition de la fonction rectangle, le spectre est nul en dehors de ce rectangle.

Dual de la fonction Disque

Pour rappel, nous avons la paire de transformées de FOURIER suivante

$$\text{Disque}_R(x,y) \Leftrightarrow R \frac{J_1(2\pi R \sqrt{u^2 + v^2})}{\sqrt{u^2 + v^2}} \quad (2.67)$$

En utilisant la propriété de dualité (relation 2.46) de la transformée de FOURIER, nous obtenons

$$\frac{J_1(2\pi f_0 \sqrt{x^2 + y^2})}{\sqrt{x^2 + y^2}} \Leftrightarrow \frac{1}{f_0} \text{Disque}_{f_0}(u,v) \quad (2.68)$$

La nouvelle image ainsi obtenue est à bande limitée. Dans ce cas, le spectre est limité au disque de rayon f_0 centré à l'origine. f_0 est appelée *fréquence radiale* de coupure.

2.7 Échantillonnage

La question de l'échantillonnage bidimensionnel est singulièrement plus délicate que celle de l'échantillonnage unidimensionnel car rien n'impose que l'échantillonnage soit effectué suivant les axes horizontaux et verticaux.

Pour l'étude de l'échantillonnage, nous généralisons tout d'abord la notion d'impulsion de DIRAC. Cette fonction servira de base pour comprendre le processus d'échantillonnage.

2.7.1 Impulsion de DIRAC bidimensionnelle

Définition 7 La fonction delta ou impulsion de DIRAC bidimensionnelle, notée $\delta(x, y)$, est définie par

$$\delta(x, y) = 0 \quad \forall (x, y) \neq (0, 0) \quad (2.69)$$

et

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x, y) dx dy = 1 \quad (2.70)$$

Il est possible de donner une autre définition équivalente de l'impulsion de DIRAC qui englobe les équations 2.69 et 2.70

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x - \alpha, y - \beta) dx dy = f(\alpha, \beta) \quad (2.71)$$

où $f(x, y)$ est une fonction continue de x et y .

Propriétés

La fonction $\delta(x, y)$ est spatialement séparable en deux fonctions de DIRAC à une dimension

$$\delta(x, y) = \delta(x) \delta(y) \quad (2.72)$$

Étant donné que la fonction $\delta(x, y)$ est une fonction symétrique par rapport à l'origine, l'équation (2.71) peut encore s'écrire, en changeant les variables d'intégration, sous la forme

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) \delta(x - \alpha, y - \beta) d\alpha d\beta = f(x, y) \quad (2.73)$$

Le membre de gauche de l'équation 2.73 représente le produit de convolution de $f(x, y)$ par $\delta(x, y)$. On peut donc écrire

$$f(x, y) \otimes \delta(x, y) = f(x, y) \quad (2.74)$$

Dès lors, la convolution d'une fonction $f(x, y)$ avec l'impulsion de DIRAC laisse cette fonction inchangée.

Transformée de FOURIER

Par définition, la transformée de FOURIER de $\delta(x, y)$ est donnée par

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x, y) e^{-2\pi j(xu+yv)} dx dy \quad (2.75)$$

Étant donné que la fonction $e^{-2\pi j(xu+yv)}$ évaluée en l'origine $(0, 0)$ vaut 1, il vient finalement

$$\delta(x, y) \rightleftharpoons 1 \quad (2.76)$$

Donc, le spectre en fréquence de l'impulsion de DIRAC s'étend uniformément sur tout l'intervalle fréquentiel $[-\infty, +\infty] \times [-\infty, +\infty]$.

Liens avec quelques images typiques

1. Image continue

En utilisant la propriété de dualité (relation 2.46) de la transformée de FOURIER et étant donné que $\delta(x, y)$ est symétrique par rapport à l'origine, on peut écrire

$$1 \Leftrightarrow \delta(u, v) \quad (2.77)$$

Le spectre d'une image continue est donc discret et comporte *une seule raie* située à l'origine $(0, 0)$ du plan $u - v$.

2. Image complexe exponentielle

En appliquant la propriété de translation fréquentielle (relation 2.48) de la transformée de FOURIER à 2.77, il vient

$$e^{-2\pi j(u_0x + v_0y)} \Leftrightarrow \delta(u - u_0, v - v_0) \quad (2.78)$$

Le spectre d'une image complexe exponentielle de fréquence (u_0, v_0) se limite donc à *une raie* située en (u_0, v_0) dans le plan $u - v$.

3. Image sinusoidale

Soit l'image $\cos[2\pi(u_0x + v_0y)]$. Rappelons tout d'abord la formule bien connue

$$\cos[2\pi(u_0x + v_0y)] = \frac{e^{2\pi j(u_0x + v_0y)} + e^{-2\pi j(u_0x + v_0y)}}{2} \quad (2.79)$$

En utilisant la relation 2.78 et la propriété de linéarité de la transformée de FOURIER, il vient

$$\cos[2\pi(u_0x + v_0y)] \Leftrightarrow \frac{\delta(u - u_0, v - v_0) + \delta(u + u_0, v + v_0)}{2} \quad (2.80)$$

Le spectre d'une image cosinusoidale comporte donc 2 raies situées en (u_0, v_0) et $(-u_0, -v_0)$ dans le plan $u - v$. Il est également aisé de montrer que

$$\sin[2\pi(u_0x + v_0y)] \Leftrightarrow \frac{\delta(u - u_0, v - v_0) - \delta(u + u_0, v + v_0)}{2j} \quad (2.81)$$

2.7.2 Le processus d'échantillonnage

Considérons une fonction $f(x, y)$ connue pour tout x et y . Supposons que cette fonction soit échantillonnée uniformément de manière rectangulaire. Nous disposons ainsi d'une infinité d'échantillons de $f(x, y)$ espacés de Δx dans la direction de l'axe x et de Δy dans la direction de l'axe y . Δx et Δy sont appelées *pas d'échantillonnage* selon x et y respectivement. Les échantillons sont alors notés $f(i\Delta x, j\Delta y)$ où i et j sont des nombres entiers. $\frac{1}{\Delta x}$ et $\frac{1}{\Delta y}$ sont appelées *fréquences d'échantillonnage* selon x et y respectivement.

Introduisons à présent la *fonction d'échantillonnage idéale* définie par

$$s(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta(x - i\Delta x, y - j\Delta y) \quad (2.82)$$

et représentée à la figure 2.8.

Établissons à présent le lien entre les échantillons $f(i\Delta x, j\Delta y)$ et la fonction d'échantillonnage idéale.

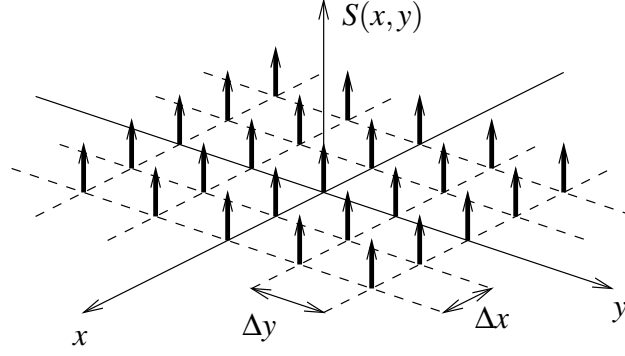


Figure 2.8: Représentation de la fonction $s(x,y)$.

Définition 8 Nous définissons ainsi la fonction échantillonnée idéale

$$f_s(x,y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(i\Delta x, j\Delta y) \delta(x - i\Delta x, y - j\Delta y) \quad (2.83)$$

La fonction échantillonnée idéale est donc un réseau d'impulsions de DIRAC pondérées par la valeur de la fonction $f(x,y)$ évaluée aux droits d'échantillonnage.

Déterminons à présent la relation existant entre les transformées de FOURIER respectives de $f_s(x,y)$ et $f(x,y)$. Pour cela, nous partons de la relation 2.83 et nous utilisons les propriétés de la fonction de DIRAC

$$f_s(x,y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(x,y) \delta(x - i\Delta x, y - j\Delta y) \quad (2.84)$$

$$= f(x,y) \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta(x - i\Delta x, y - j\Delta y) \quad (2.85)$$

$$= f(x,y) s(x,y) \quad (2.86)$$

La fonction échantillonnée idéale est donc égale au produit de la fonction $f(x,y)$ et de la fonction d'échantillonnage idéale dans le domaine spatial. En utilisant la propriété de multiplication (relation 2.53) de la transformée de FOURIER, il vient

$$f_s(x,y) \Leftrightarrow \mathcal{F}(u,v) \otimes \mathcal{S}(u,v) \quad (2.87)$$

où $\mathcal{S}(u,v)$ est la transformée de FOURIER de la fonction d'échantillonnage idéale. Étant donné que

$$\sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta(x - i\Delta x, y - j\Delta y) \Leftrightarrow \frac{1}{\Delta x \Delta y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta\left(u - \frac{i}{\Delta x}, v - \frac{j}{\Delta y}\right) \quad (2.88)$$

nous pouvons déterminer la transformée de FOURIER de $f_s(x,y)$

$$\mathcal{F}_s(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{S}(\alpha,\beta) \mathcal{F}(u-\alpha, v-\beta) d\alpha d\beta \quad (2.89)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left\{ \frac{1}{\Delta x \Delta y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta\left(\alpha - \frac{i}{\Delta x}, \beta - \frac{j}{\Delta y}\right) \right\} \mathcal{F}(u-\alpha, v-\beta) d\alpha d\beta$$

$$= \frac{1}{\Delta x \Delta y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta\left(\alpha - \frac{i}{\Delta x}, \beta - \frac{j}{\Delta y}\right) \mathcal{F}(u-\alpha, v-\beta) d\alpha d\beta$$

$$= \frac{1}{\Delta x \Delta y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \mathcal{F}\left(u - \frac{i}{\Delta x}, v - \frac{j}{\Delta y}\right) \quad (2.90)$$

Il vient finalement

$$f_s(x, y) \Leftrightarrow \frac{1}{\Delta x \Delta y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \mathcal{F} \left(u - \frac{i}{\Delta x}, v - \frac{j}{\Delta y} \right) \quad (2.91)$$

Ce résultat est particulièrement intéressant: mis à part le facteur $1/\Delta x \Delta y$, la transformée de FOURIER de la fonction échantillonnée $f_s(x, y)$ est obtenue par la répétition de celle de $f(x, y)$ à tous les multiples de la fréquence d'échantillonnage $(1/\Delta x, 1/\Delta y)$. Elle est donc périodique de période fondamentale $(1/\Delta x, 1/\Delta y)$.

Comme pour le cas des signaux unidimensionnels, il existe un théorème de SHANNON permettant d'éviter le recouvrement des copies du spectre occasionné par un sous-échantillonnage; ce phénomène porte le nom de *repli de spectre* ou d'*aliasing*. Afin d'illustrer ce théorème, considérons la fonction suivante

$$g(x, y) = \frac{J_1 \left(2\pi W \sqrt{x^2 + y^2} \right)}{\sqrt{x^2 + y^2}} \quad (2.92)$$

dont nous savons que la transformée de FOURIER est la fonction disque de rayon W .

Supposons que $g(x, y)$ soit échantillonnée à la fréquence $(1/\Delta x, 1/\Delta y)$. La transformée de FOURIER de la fonction échantillonnée correspondante est représentée à la figure 2.9. Sur cette figure, on peut remarquer qu'il y a recouvrement du spectre dans la direction v mais pas dans la direction u .

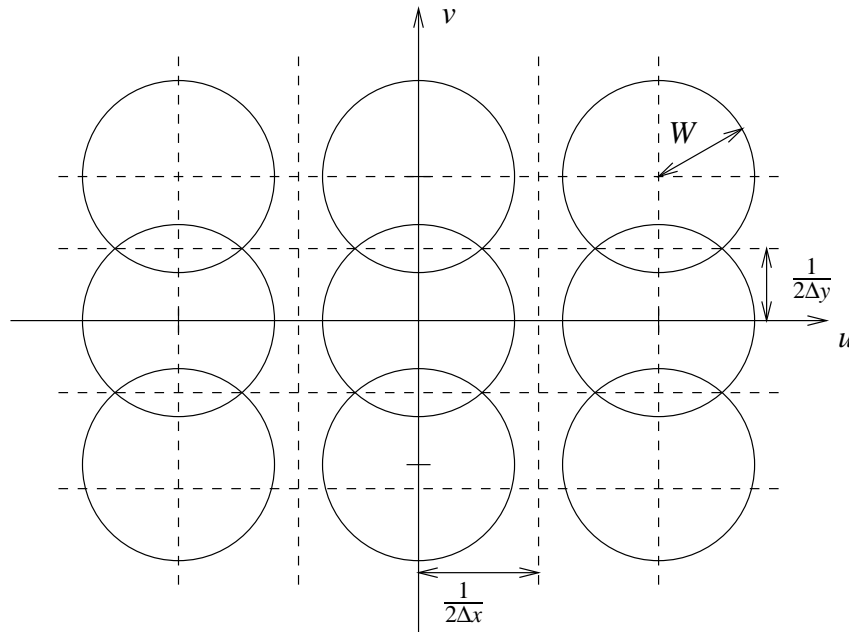


Figure 2.9: Transformée de FOURIER de la fonction échantillonnée de $g(x, y)$ ($1/\Delta x = 2,5W$ et $1/\Delta y = 1,5W$).

Nous pouvons à présent énoncer le théorème de SHANNON. Soit u_{max} et v_{max} les fréquences maximales d'une fonction $f(x, y)$ dans les directions x et y respectivement. On suppose donc que l'image $f(x, y)$ est à bande limitée.

Théorème 9 *Le théorème de SHANNON (aussi appelé critère de NYQUIST) dit que Δx et Δy doivent respecter les conditions suivantes*

$$\frac{1}{\Delta x} > 2u_{max} \quad \frac{1}{\Delta y} > 2v_{max} \quad (2.93)$$

pour éviter tout repli de spectre.

Cependant, les images rencontrées dans la pratique sont rarement à bande limitée. Il y aura donc toujours du recouvrement lors de l'échantillonnage. Dès lors, on réalise couramment un filtrage passe-bas de l'image avant échantillonnage, par exemple en défocalisant légèrement l'objectif de la caméra d'acquisition.

2.8 Reconstruction

Dans ce qui suit, nous ferons l'hypothèse simplificatrice que l'image $f(x, y)$ est à bande limitée et que les conditions d'échantillonnage de SHANNON sont respectées, cela pour séparer les problèmes dus au recouvrement des problèmes de reconstruction à partir de la fonction échantillonnée.

Le but ici est de reconstruire, à partir de la fonction échantillonnée $f_s(x, y)$, une fonction $f_r(x, y)$ qui se rapproche le plus possible de $f(x, y)$. Idéalement, on devrait avoir $f_r(x, y) = f(x, y)$.

Étant données les hypothèses qui ont été introduites, les répliques du spectre de $f(x, y)$ dans celui de la fonction échantillonnée ne se recouvrent pas et il est possible de reconstruire la fonction $f(x, y)$ exactement en choisissant judicieusement un filtre de reconstruction. Celui-ci sera simplement un filtre passe-bas permettant de conserver la réplique du spectre de $f(x, y)$ centrée à l'origine. En toute généralité, la réponse impulsionnelle du filtre de reconstruction, également appelé fonction d'*interpolation*, sera noté $r(x, y)$. La fonction reconstruite est alors obtenue par

$$f_r(x, y) = f_s(x, y) \otimes r(x, y) \quad (2.94)$$

$$= \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(i\Delta x, j\Delta y) r(x - i\Delta x, y - j\Delta y) \quad (2.95)$$

2.8.1 Filtre passe-bas à fenêtre rectangulaire

La fenêtre de filtrage a ici une forme rectangulaire. À la figure 2.10, on peut voir le spectre d'une fonction échantillonnée ainsi que la fenêtre rectangulaire de filtrage.

Le filtre est défini dans le domaine fréquentiel par la fonction de transfert suivante

$$\mathcal{R}(u, v) = \Delta x \Delta y \text{Rect}_{\frac{1}{\Delta x}, \frac{1}{\Delta y}}(u, v) \quad (2.96)$$

On voit aisément que l'application de cette fenêtre fréquentielle à la fonction échantillonnée permet de reconstruire exactement $f(x, y)$. La réponse impulsionnelle du filtre s'obtient par transformée de FOURIER inverse de 2.96. Il vient

$$r(x, y) = \frac{\sin\left(\frac{\pi x}{\Delta x}\right)}{\frac{\pi x}{\Delta x}} \frac{\sin\left(\frac{\pi y}{\Delta y}\right)}{\frac{\pi y}{\Delta y}} \quad (2.97)$$

En introduisant la notation de la fonction *sinus cardinal*, $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$, on peut encore écrire

$$r(x, y) = \text{sinc}\left(\frac{x}{\Delta x}\right) \text{sinc}\left(\frac{y}{\Delta y}\right) \quad (2.98)$$

Ce filtre conduit à la formule de reconstruction suivante

$$f_r(x, y) = f(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(i\Delta x, j\Delta y) \text{sinc}\left(\frac{x - i\Delta x}{\Delta x}\right) \text{sinc}\left(\frac{y - j\Delta y}{\Delta y}\right) \quad (2.99)$$

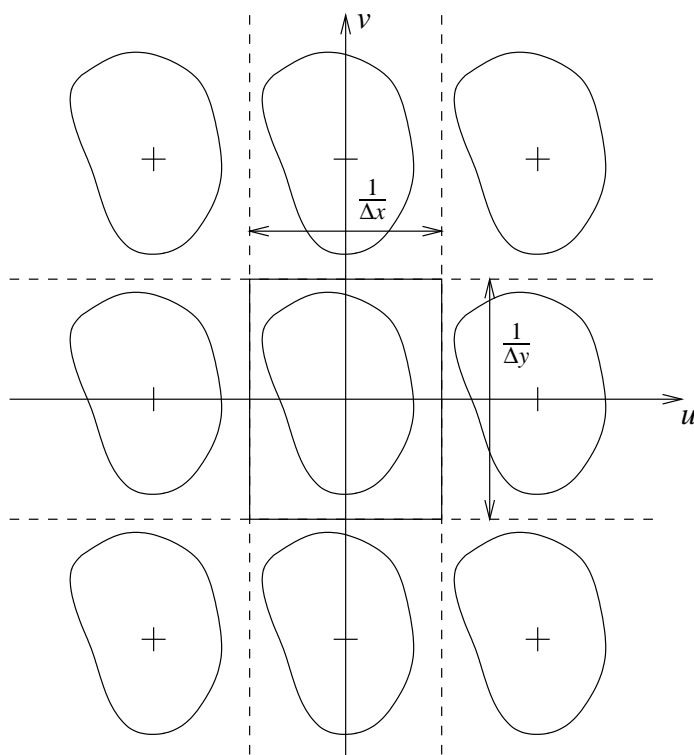


Figure 2.10: Filtre passe-bas à fenêtre rectangulaire

Cette dernière formule constitue une *formule d'interpolation* qui permet de reconstruire exactement $f(x, y)$ à partir de ses échantillons $f(i\Delta x, j\Delta y)$. Cependant, il faut bien garder à l'esprit que cela n'est valable que si les conditions d'échantillonnage de SHANNON sont vérifiées. L'expression 2.99 est l'expression de la formule d'interpolation de WHITTAKER pour le cas des signaux à une dimension.

2.8.2 Filtre passe-bas à fenêtre circulaire

Contrairement au cas unidimensionnel, le traitement d'image offre plus de liberté quant au choix du filtre passe-bas idéal. Il est ainsi possible de choisir une fenêtre circulaire comme le montre la figure 2.11.

Il est clair que, pour que la reconstruction soit exacte, la fenêtre circulaire doit contenir entièrement la réplique du spectre de $f(x, y)$ située à l'origine et elle ne peut intersecter les autres répliques du spectre. Ce filtre est alors défini par la fonction de transfert suivante

$$\mathcal{R}(u, v) = \Delta x \Delta y \text{Disque}_W(u, v) \quad (2.100)$$

La réponse impulsionnelle correspondante est donnée par

$$r(x, y) = W \Delta x \Delta y \frac{J_1\left(2\pi W \sqrt{x^2 + y^2}\right)}{\sqrt{x^2 + y^2}} \quad (2.101)$$

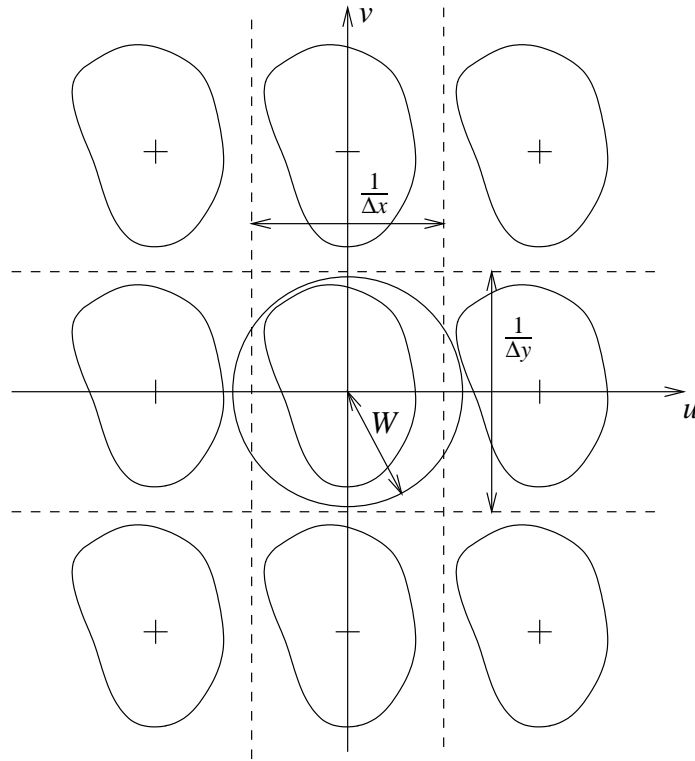


Figure 2.11: Filtre passe-bas à fenêtre circulaire.

La formule de reconstruction ou d'interpolation est finalement

$$\begin{aligned}
 f_r(x,y) &= f(x,y) & (2.102) \\
 &= W\Delta x\Delta y \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(i\Delta x, j\Delta y) \frac{J_1 \left(2\pi W \sqrt{(x-i\Delta x)^2 + (y-j\Delta y)^2} \right)}{\sqrt{(x-i\Delta x)^2 + (y-j\Delta y)^2}}
 \end{aligned}$$

Cette dernière formule constitue une *formule d'interpolation* faisant intervenir des fonctions de BESSEL du premier ordre.

Conditions de reconstruction

Les fenêtres rectangulaire et circulaire ne sont pas les seules à pouvoir être utilisées. Pour obtenir une reconstruction exacte de $f(x,y)$, la fenêtre fréquentielle de filtrage doit vérifier les conditions suivantes:

1. elle doit être symétrique par rapport à l'origine afin d'assurer que l'image reconstruite soit réelle;
2. elle doit contenir entièrement la réplique du spectre de $f(x,y)$ située à l'origine; et
3. elle ne peut intersecter les autres répliques du spectre de $f(x,y)$.

À nouveau, il est bon de rappeler que tout ceci n'est valable que si la fonction $f(x,y)$ est à bande limitée et que les conditions d'échantillonnage sont vérifiées. Dans le cas contraire, l'image reconstruite présentera ce qu'on appelle du *repli de spectre* ou *aliasing*.

2.8.3 Autres filtres de reconstruction et d'interpolation

Dans les deux sections précédentes, les filtres ont été définis dans le domaine fréquentiel de l'image. Il est également possible de travailler directement dans le domaine spatial, c'est-à-dire d'établir le filtre de reconstruction $r(x, y)$ et de voir son effet dans le domaine fréquentiel grâce à sa transformée de FOURIER $\mathcal{R}(u, v)$.

Considérons l'exemple suivant. Supposons que la fonction $f(x, y)$ ait été échantillonnée avec des pas Δx et Δy constants. Soit $f_r(x, y)$ l'image reconstruite définie par

$$f_r(x, y) = f(i\Delta x, j\Delta y) \quad (2.103)$$

où i et j sont tels que

$$\begin{cases} i\Delta x - \frac{\Delta x}{2} \leq x < i\Delta x + \frac{\Delta x}{2} \\ j\Delta y - \frac{\Delta y}{2} \leq y < j\Delta y + \frac{\Delta y}{2} \end{cases} \quad (2.104)$$

Cette reconstruction correspond à l'*interpolation du plus proche voisin* car la valeur de f_r en (x, y) correspond au plus proche échantillon connu. L'image reconstruite peut donc être vue comme une image en *escaliers*. Elle peut également s'écrire

$$f_r(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(i\Delta x, j\Delta y) \text{Rect}_{\Delta x, \Delta y}(x - i\Delta x, y - j\Delta y) \quad (2.105)$$

De cette expression, nous pouvons extraire la fonction d'interpolation (ou filtre de reconstruction) utilisée

$$r(x, y) = \text{Rect}_{\Delta x, \Delta y}(x, y) \quad (2.106)$$

Il est alors intéressant de regarder l'effet du filtre de reconstruction sur l'image échantillonnée. Pour cela, observons la fonction de transfert du filtre

$$\mathcal{R}(u, v) = \Delta x \Delta y \frac{\sin(\pi \Delta x u)}{\pi \Delta x u} \frac{\sin(\pi \Delta y v)}{\pi \Delta y v} \quad (2.107)$$

Deux remarques importantes peuvent être faites à propos de ce filtre de reconstruction:

1. dans le plan fréquentiel (u, v) , la fenêtre de filtrage $\mathcal{R}(u, v)$ n'est pas constante sur tout le rectangle de dimension $\frac{1}{\Delta x} \times \frac{1}{\Delta y}$ centré à l'origine, c'est-à-dire la région contenant la réplique du spectre de $f(x, y)$ centrée à l'origine. Dans ce rectangle, les composantes fréquentielles (u, v) de l'image sont d'autant plus atténuées que le couple (u, v) est éloigné de l'origine.
2. la fenêtre fréquentielle n'est pas nulle en dehors de la région d'intérêt rectangulaire de dimension $\frac{1}{\Delta x} \times \frac{1}{\Delta y}$ centrée à l'origine. L'image reconstruite tient non seulement compte de la réplique du spectre de $f(x, y)$ située à l'origine mais en plus d'autres répliques qui vont introduire du repli de spectre dans l'image reconstruite.

En toute généralité, l'utilisation de filtres de reconstruction non idéaux (comme par exemple l'*interpolation du plus proche voisin*) introduit les problèmes suivants:

- modification du contenu fréquentiel intrinsèque de l'image,
- introduction de hautes fréquences parasites dans l'image reconstruite.

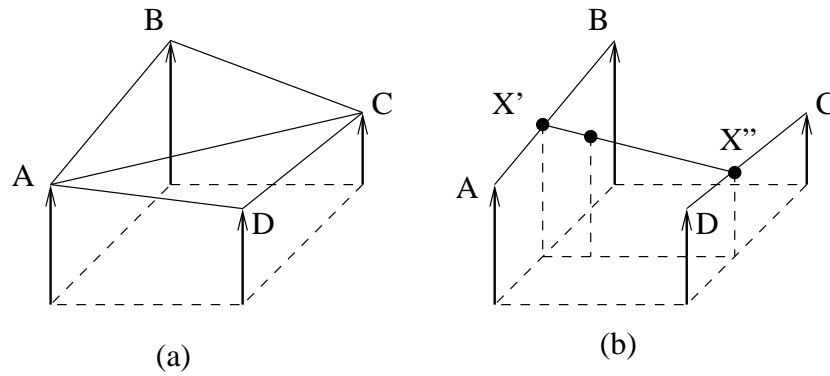


Figure 2.12: Exemples d'interpolateurs: (a) interpolation linéaire par morceaux, (b) interpolation bilinéaire.

Il existe évidemment d'autres filtres de reconstruction comme par exemple l'interpolation linéaire par morceaux (figure 2.12.a) ou l'interpolation bilinéaire (figure 2.12.b).

Insistons une fois encore sur le fait que l'utilisation d'un filtre de reconstruction idéal ne fournira pas nécessairement une reconstruction parfaite. Il faut pour cela que l'image $f(x,y)$ soit à bande limitée (ce qui extrêmement rare en pratique) et échantillonnée correctement (en concordance avec le théorème de SHANNON).

Chapter 3

Filtrage linéaire et déconvolution

De l'interprétation des transformées unitaires, il apparaît aisément qu'une modification des coefficients transformés aura pour effet de supprimer ou d'augmenter la présence de certaines composantes fréquentielles dans le domaine spatial. Il est donc tout naturel de chercher à filtrer le signal d'image par le truchement d'une modification des coefficients transformés. Ces techniques font partie du domaine du filtrage linéaire, même si leur implémentation s'effectue parfois par convolution dans le domaine spatial.

Avant d'aborder l'étude du filtrage linéaire, une précision s'impose. On pourrait croire que le filtrage linéaire est le seul moyen efficace de filtrer une image. En fait, il n'en est rien car on peut construire des filtres très intéressants en partant de propriétés algébriques à respecter. Cette démarche conduit à synthétiser des filtres non nécessairement linéaires.

Le filtrage d'une image est la généralisation à deux dimensions du filtrage de signaux temporels rencontrés dans la théorie des systèmes et en télécommunications. À ce niveau, la variable temporelle t est remplacée par deux variables spatiales x et y . La variable fréquentielle f est quant à elle remplacée par les variables u et v .

Dans certains cas, l'image a subi l'effet d'un filtre dont la transmittance est connue. Ainsi, la défocalisation de l'objectif d'une caméra produit un effet de flou que l'on cherche à éliminer. La suppression de cet effet porte le nom de *déconvolution*; c'est l'opération inverse du filtrage. L'étude de la déconvolution sera également abordée dans ce chapitre.

3.1 Notion de filtre idéal

On dira d'un filtre qu'il est *idéal*¹ s'il multiplie tous les coefficients transformés par 0 ou par 1.

Définition 10 [Filtre idéal] *Un filtre est idéal si sa transmittance est telle que*

$$\forall(u, v), \mathcal{H}(u, v) = 0 \text{ ou } 1 \quad (3.1)$$

La notion de filtre idéal est intimement liée à celle d'idempotence. L'idempotence doit être comprise ici telle que, pour toute image $f(x, y)$,

$$\mathcal{F}(u, v)\mathcal{H}(u, v) = \mathcal{F}(u, v)\mathcal{H}(u, v)\mathcal{H}(u, v) \quad (3.2)$$

Autrement dit, un signal préalablement filtré par un filtre idéal n'est pas modifié par application récursive de ce filtre.

¹La notion de filtre *idéal* est purement formelle. En effet, en pratique, un filtre idéal n'a rien d'idéal; il est irréalisable et ses propriétés le rendent inutilisable.

3.1.1 Catégories de filtres idéaux

Dans le cas d'un signal unidimensionnel, on distingue habituellement trois catégories de filtres idéaux en fonction de la position de leur gabarit le long de l'axe des fréquences (cf. figure 3.1).

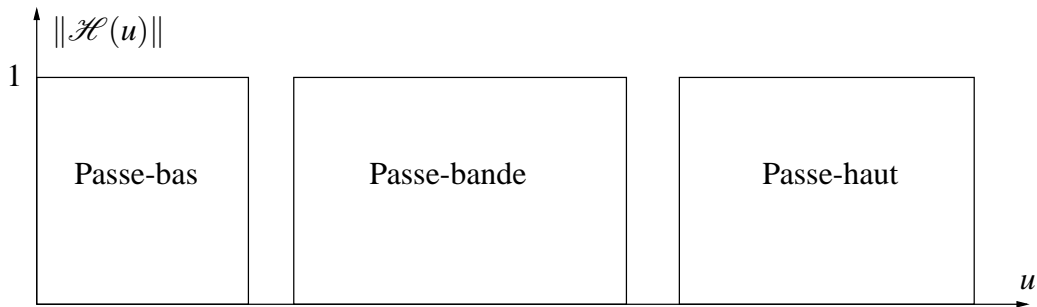


Figure 3.1: Filtres idéaux unidimensionnels.

L'extension à des signaux bidimensionnels est simple si l'on utilise des filtres à symétrie circulaire. Les trois types de filtres idéaux à symétrie circulaire sont:

- les *filtres passe-bas*, pour lesquels seuls les coefficients de la transmittance à proximité de l'origine sont non nuls. Un filtre passe-bas est un système linéaire ne modifiant pas ou peu les basses fréquences de l'image d'entrée. Par basses fréquences, on entend les couples (u, v) du plan de fréquence $u - v$ situé dans un voisinage de l'origine, la taille de ce voisinage caractérisant la *bande passante* du filtre. Les hautes fréquences de l'image, en l'occurrence celles situées en dehors du voisinage de l'origine sont atténuées, voir même annulées. Physiquement, le filtre passe-bas a pour effet d'atténuer les variations rapides d'intensité de l'image pouvant aller jusqu'à faire apparaître une certaine impression de flou dans l'image filtrée.

La figure 3.2 montre une image de taille 256×256 et l'image filtrée au moyen d'un filtre à symétrie circulaire ayant une fréquence de coupure $f_c = 30$ pixels. Ce filtre a un effet moyennneur sur le niveau de luminance. En raison de la suppression des composantes à haute fréquence, les transitions se retrouvent adoucies.



(a) Image originale



(b) Image filtrée

Figure 3.2: Filtrage passe-bas d'une image.

Le filtre passe-bas idéal circulaire est défini par une fonction de transfert de la forme

$$\mathcal{H}(u, v) = \begin{cases} 1 & \sqrt{u^2 + v^2} \leq R_0 \\ 0 & \sqrt{u^2 + v^2} > R_0 \end{cases} \quad (3.3)$$

Les composantes fréquentielles de l'image correspondant aux couples (u, v) situés à l'intérieur du disque de rayon R_0 , dites *basses fréquences*, ne subissent aucune modification tandis que les autres composantes fréquentielles, dites *hautes fréquences*, sont complètement supprimées. La réponse impulsionnelle du filtre est obtenue par transformée de FOURIER inverse de la fonction de transfert; elle fait intervenir une fonction de BESSEL et vaut

$$h(x, y) = R_0 \frac{J_1\left(2\pi R_0 \sqrt{x^2 + y^2}\right)}{\sqrt{x^2 + y^2}} \quad (3.4)$$

Ce filtre est dit idéal car il y a suppression totale des composantes fréquentielles ayant une fréquence radiale $\sqrt{u^2 + v^2}$ supérieure à R_0 .

- les *filtres passe-haut*. Ces filtres sélectionnent un contenu en hautes fréquences. Un filtre passe-haut linéaire est un système linéaire ne modifiant pas ou peu les hautes fréquences de l'image d'entrée. Les basses fréquences sont, quant à elles, atténuées, voire même complètement annulées. Visuellement, un filtre passe-haut a pour effet de supprimer la composante continue de l'image et de ne garder que les variations rapides d'intensité dans l'image filtrée.

Le filtre passe-haut idéal est défini par la fonction de transfert suivante

$$\mathcal{H}(u, v) = \begin{cases} 1 & \sqrt{u^2 + v^2} \geq R_0 \\ 0 & \sqrt{u^2 + v^2} < R_0 \end{cases} \quad (3.5)$$

Les basses fréquences dont la fréquence radiale $\sqrt{u^2 + v^2}$ est inférieure à R_0 sont complètement rejetées tandis que les hautes fréquences restent inchangées.

- les *filtres passe-bande*. Ils sont équivalents au complémentaire d'un filtre passe-bas et d'un filtre passe-haut. Un filtre passe-bande est un système linéaire qui atténue ou supprime les basses et les hautes fréquences. Seule une plage de fréquences n'est pas ou peu atténuée par le passage au travers de ce filtre.

Le filtre passe-bande idéal est défini par la fonction de transfert suivante

$$\mathcal{H}(u, v) = \begin{cases} 1 & R_0 \leq \sqrt{u^2 + v^2} \leq R_1 \\ 0 & \text{sinon} \end{cases} \quad (3.6)$$

Seules les composantes fréquentielles dont la fréquence radiale est comprise entre R_0 et R_1 sont conservées. Toutes les autres composantes fréquentielles sont rejetées. La fonction de transfert du filtre passe-bande idéal est représentée à la figure 3.3.

La figure 3.4 compare les effets d'un filtrage passe-bas, d'un filtrage passe-bande et d'un filtrage passe-haut ². En vertu de la linéarité de la transformée de FOURIER, l'addition des 3 spectres reconstitue le signal original. L'observation des images filtrées montre clairement des effets de filtrage complémentaires.

²Pour la visualisation des spectres, nous avons adopté la convention de *vidéo inverse*; elle consiste simplement à inverser l'échelle de luminance.

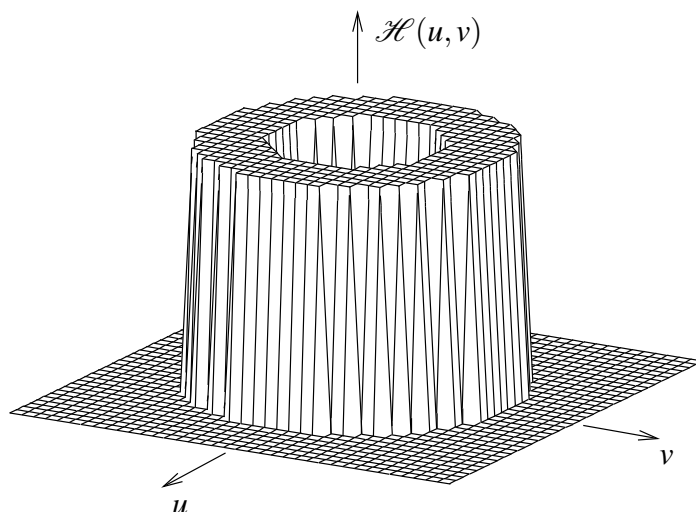


Figure 3.3: Fonction de transfert du filtre passe-bande idéal.

En pratique, on préfère souvent recourir à des filtres à support rectangulaire. Un tel choix se justifie par une plus grande facilité de réalisation car, en présupposant la séparabilité de l'image, le filtrage s'effectuera par mise en cascade d'un filtrage horizontal et d'un filtrage vertical. En contrepartie, on court le risque de ne pas traiter au mieux les motifs à orientation diagonale. Cela ne prête pas souvent à conséquence puisqu'une majorité d'images concentrent leur énergie suivant les axes principaux en raison de la présence massive d'objets à arrêtes horizontales ou verticales.

3.2 Passage d'une image dans un système linéaire

Comme dans le cas d'un système unidimensionnel, on peut interpréter l'action d'un opérateur bidimensionnel, linéaire et invariant en translation, comme un filtrage. Le problème général du filtrage bidimensionnel consiste à élaborer un filtre qui possède la réponse impulsionnelle voulue et qui soit réalisable efficacement.

Comme pour les signaux temporels, l'analyse de FOURIER sert d'outil pour l'étude des systèmes de traitement d'images. Tout d'abord, un système est linéaire si le principe de superposition est vérifié: si $g_1(x, y)$ et $g_2(x, y)$ sont les images de sortie du système relatives aux images d'entrée $f_1(x, y)$ et $f_2(x, y)$, toute combinaison linéaire $af_1(x, y) + bf_2(x, y)$ produit en sortie l'image $ag_1(x, y) + bg_2(x, y)$. Un filtre linéaire est un exemple important de système linéaire. Il est utilisé pour modifier le contenu spectral d'une image afin d'en retirer l'information recherchée.

Un filtre linéaire bidimensionnel est caractérisé par sa *réponse impulsionnelle* $h(x, y)$. L'image de sortie $g(x, y)$ résulte du produit de convolution de la réponse impulsionnelle du filtre par l'image d'entrée $f(x, y)$

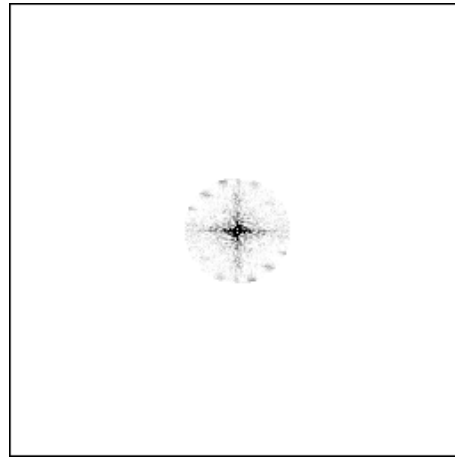
$$g(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha, \beta) h(x - \alpha, y - \beta) d\alpha d\beta \quad (3.7)$$

qui est encore noté

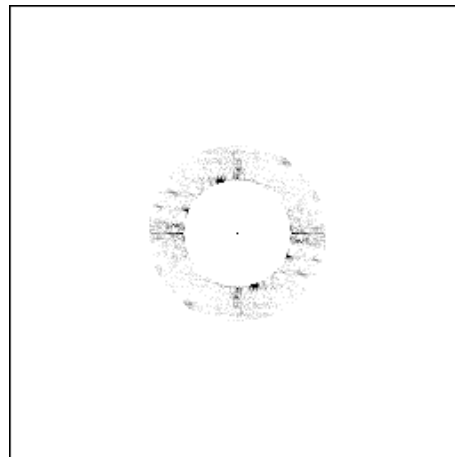
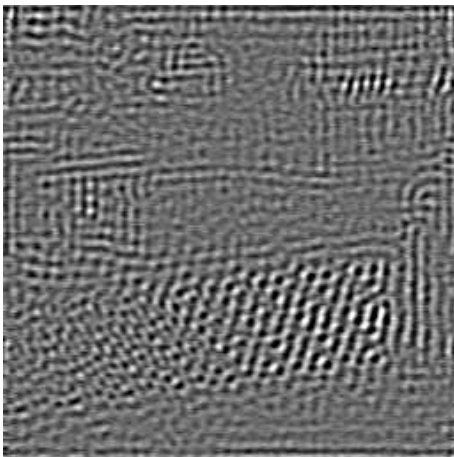
$$g(x, y) = (f \otimes h)(x, y) \quad (3.8)$$

Par la propriété de convolution de la transformée de FOURIER, nous pouvons encore écrire

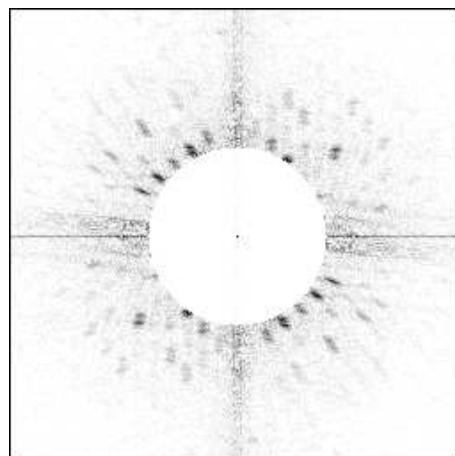
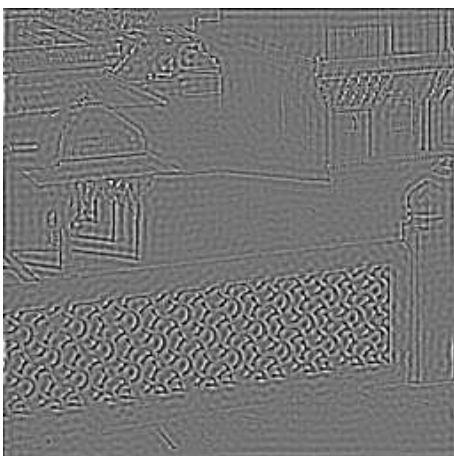
$$\mathcal{G}(u, v) = \mathcal{H}(u, v) \mathcal{F}(u, v) \quad (3.9)$$



Application d'un passe-bas circulaire de fréquence de coupure $f_c = 30$ pixels.



Application d'un passe-bande circulaire conservant les fréquences $[30, 50]$ pixels.



Application d'un passe-haut circulaire de fréquence de coupure $f_c = 50$ pixels.

Figure 3.4: Images filtrées au moyen de filtres idéaux circulaires et spectres correspondants.

La transformée de FOURIER de la réponse impulsionnelle du filtre, notée $\mathcal{H}(u, v)$, est appelée *fonction de transfert* ou encore *transmittance* du filtre. La fonction de transfert étant, en général, une fonction à valeurs complexes de u et v , nous pouvons la réécrire sous la forme

$$\mathcal{H}(u, v) = \|\mathcal{H}(u, v)\| e^{j\beta(u, v)} \quad (3.10)$$

Le module $\|\mathcal{H}(u, v)\|$ de la fonction de transfert est souvent appelée *réponse en amplitude* ou *spectre*³ tandis que $\beta(u, v)$ est souvent appelée *réponse en phase*.

3.2.1 Implémentation

Le filtrage linéaire d'une image peut donc s'envisager de deux manières. Le filtrage peut tout d'abord se réaliser dans le domaine spatial en effectuant un produit de convolution. Ou alors, le filtrage est réalisé dans le domaine fréquentiel en multipliant la transformée de FOURIER de l'image par la fonction de transfert du filtre. L'image filtrée est alors obtenue par transformée de FOURIER inverse. Enfin, on montre qu'un système linéaire invariant par translation peut également être caractérisé par une équation à coefficients [17, page 338]. C'est parfois la voie la plus facile pour réaliser le filtrage.

3.2.2 Forme de la fenêtre

Le problème majeur dans le choix d'un filtre est la forme de la fenêtre. On remarquera que les filtres idéaux ont une transmittance limitée spectralement mais une réponse impulsionnelle infinie. Cela signifie que, dans le plan spatial, une information locale va être fortement étalée. Une filtre dont la réponse impulsionnelle serait finie aurait lui une transmittance infinie. C'est alors l'information spectrale qui sera étalée dans l'image. En fait, il n'est pas possible d'échapper à ce paradoxe pour des raisons de conservation d'énergie entre le plan spatial et le plan spectral. La synthèse d'un filtre, comme ceux que nous allons étudier ci-après, résultera d'un compromis.

3.2.3 Filtres passe-bas

Un exemple commun de filtre passe-bas est le filtre passe-bas de BUTTERWORTH d'ordre n défini par la fonction de transfert suivante

$$\mathcal{H}(u, v) = \frac{1}{1 + \left(\frac{\sqrt{u^2+v^2}}{R_0}\right)^{2n}} \quad (3.11)$$

On remarque ici que toutes les composantes fréquentielles, hormis l'origine, subissent une atténuation d'autant plus grande que le couple (u, v) est éloigné de l'origine. Plus l'ordre n du filtre est élevé, plus l'atténuation des hautes fréquences est importante. La transmittance du filtre passe-bas de BUTTERWORTH pour $n = 1$ est représentée à la figure 3.5.

L'effet d'un filtrage passe-bas est généralement néfaste pour les contours. La figure 3.7 montre une image originale et trois versions filtrées par un filtre de plus en plus sélectif. On y observe une atténuation progressive du bruit ainsi qu'une dégradation croissante des contours.

³On appelle indifféremment *spectre* $\|\mathcal{H}(u, v)\|$ ou $\|\mathcal{H}(u, v)\|^2$. De fait, en unités logarithmiques, les deux valeurs sont équivalents à un facteur multiplicatif 2 près.

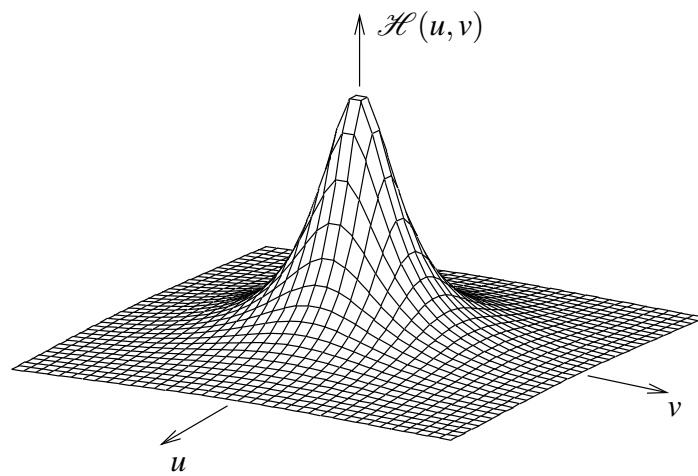
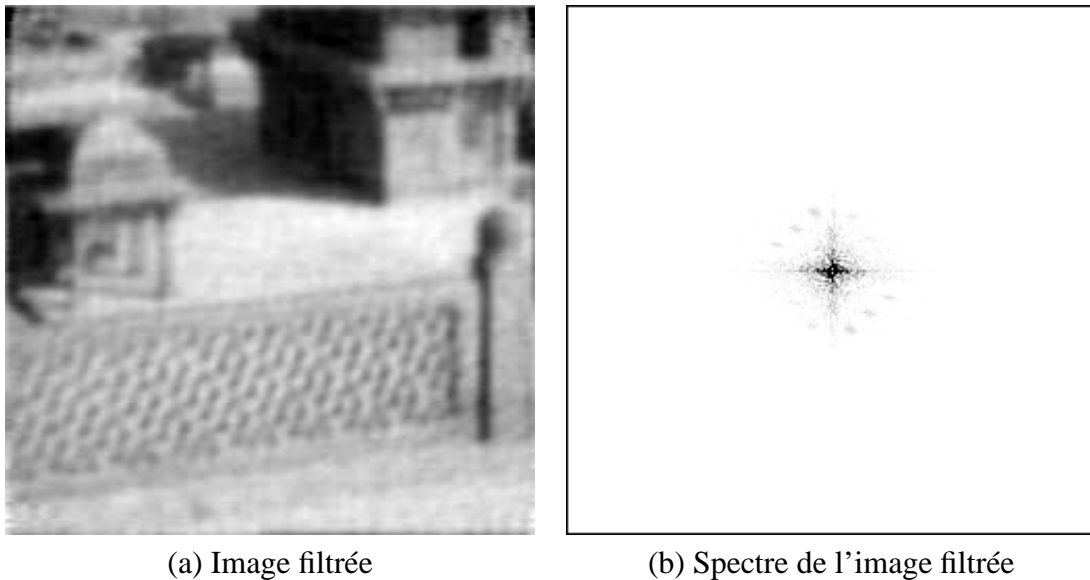


Figure 3.5: Transmittance du filtre passe-bas de BUTTERWORTH pour $n = 1$.



(a) Image filtrée

(b) Spectre de l'image filtrée

Figure 3.6: Image filtrée par un filtre passe-bas de BUTTERWORTH d'ordre 1 ($f_c = 30$).

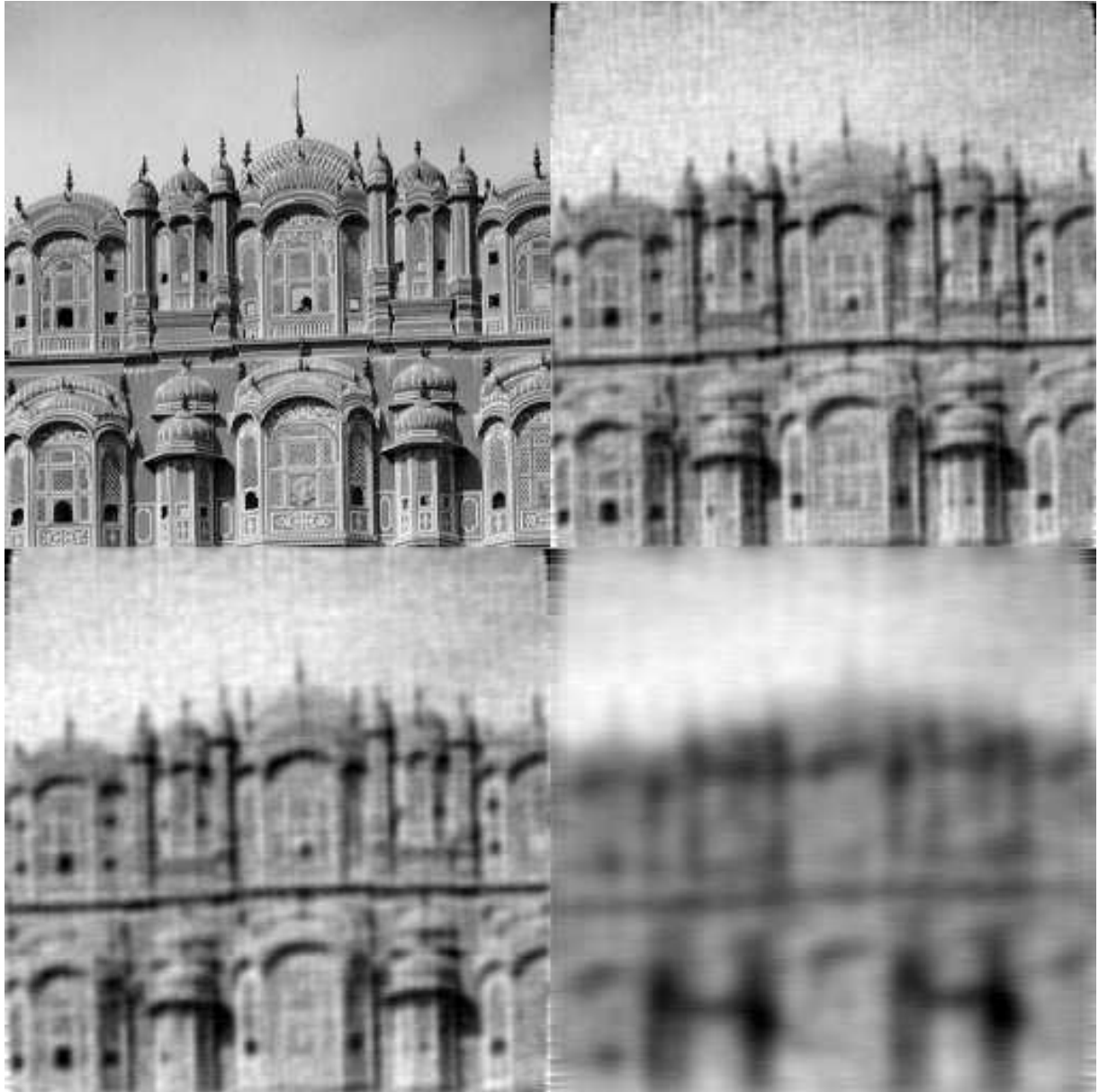


Figure 3.7: Effet d'un filtrage passe-bas sur une image.

3.2.4 Filtres passe-haut

Il existe évidemment d'autres formes de filtres passe-haut comme par exemple le filtre passe-haut de BUTTERWORTH d'ordre n dont la fonction de transfert est donnée par

$$\mathcal{H}(u, v) = \frac{1}{1 + \left(\frac{R_0}{\sqrt{u^2 + v^2}}\right)^{2n}} \quad (3.12)$$

Comme pour le cas du filtre passe-bas, toutes les fréquences sont atténuées et cela d'autant plus que $\sqrt{u^2 + v^2}$ est petit par rapport R_0 . De plus, n fixe la pente de la transition entre basses et hautes fréquences.

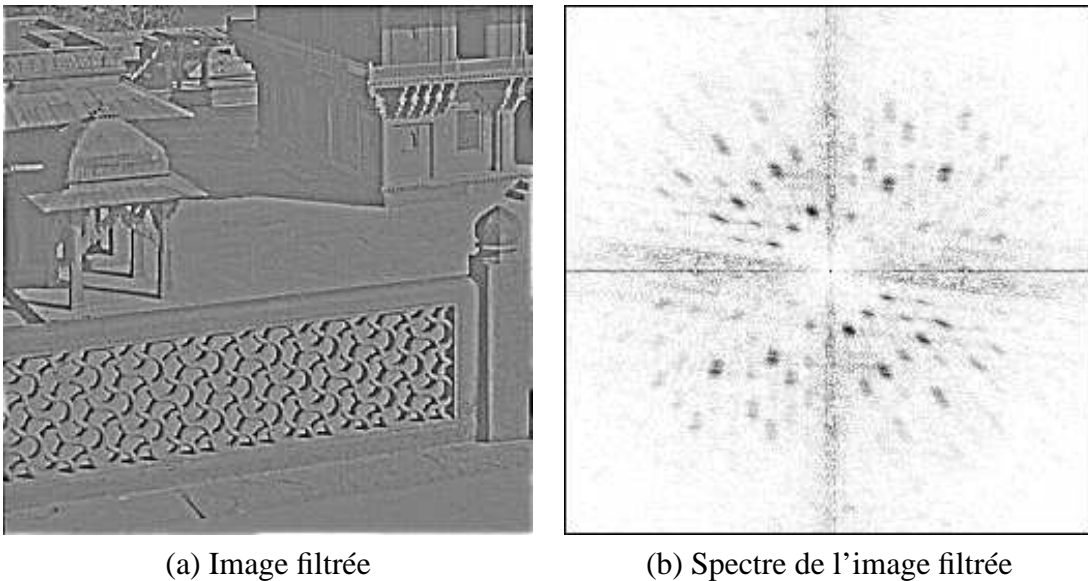


Figure 3.8: Image filtrée par un filtre passe-haut de BUTTERWORTH d'ordre 1 ($f_c = 50$).

Le filtrage passe-haut a un effet dérivateur. De fait, l'opérateur de dérivation a une transmittance de type $\mathcal{H}(u, v) = \alpha(ju)(jv) = \alpha'uv$; il amplifie donc les hautes fréquences et atténue les basses fréquences. Cet effet dérivateur peut être observé à la figure 3.9; les filtres sont de plus en plus sélectifs.

3.2.5 Filtres de GABOR

Définition

Les filtres de GABOR ou filtres *gaussiens* constituent une classe particulière des filtres linéaires; ce sont des filtres *orientés*. Ces filtres ont une réponse impulsionnelle de la forme:

$$h(x, y) = g(x', y') e^{2\pi j(Ux + Vy)} \quad (3.13)$$

où

- $(x', y') = (x \cos \phi + y \sin \phi - x \sin \phi + y \cos \phi)$, c'est-à-dire les coordonnées (x, y) tournées d'un angle ϕ , et
- $g(x', y') = \frac{1}{2\pi\sigma^2} e^{-(x'/\lambda)^2 + y'^2/2\sigma^2}$.

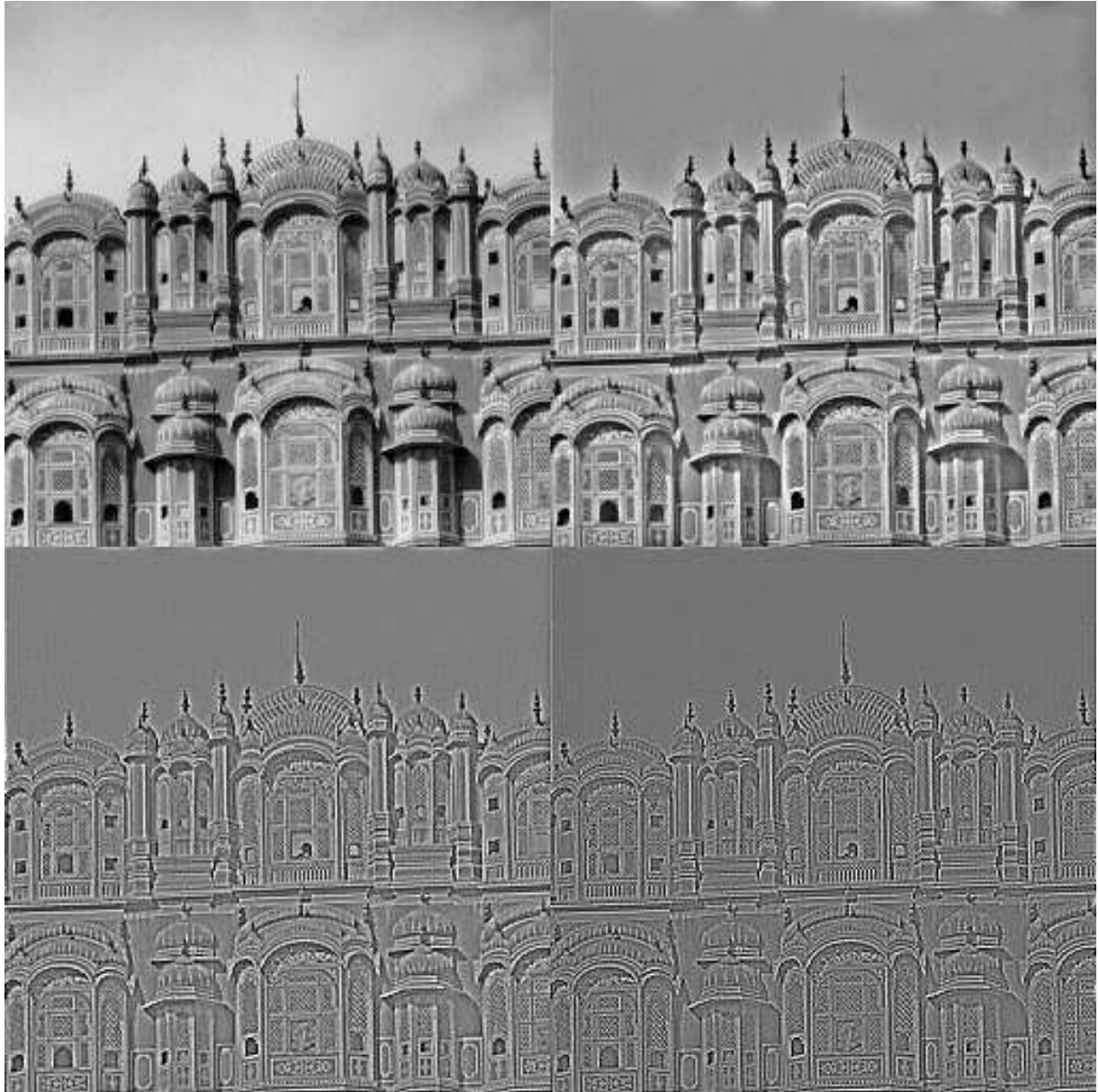


Figure 3.9: Effet d'un filtrage passe-haut sur une image.

La réponse impulsionnelle $h(x,y)$ est donc une fonction complexe sinusoïdale modulée par une gaussienne bidimensionnelle de rapport d'axes λ , de facteur de dilatation σ et où ϕ est l'orientation de l'axe x' par rapport à l'axe x .

La transformée de FOURIER correspondante est:

$$\mathcal{H}(u,v) = e^{-2\pi^2\sigma^2[(u'-U')^2\lambda^2+(v'-V')^2]} \quad (3.14)$$

où

- $(u',v') = (u\cos\phi + v\sin\phi - u\sin\phi + v\cos\phi)$, et
- (U',V') est obtenu par rotation du point (U,V) par le même angle ϕ .

La trace d'un filtre de Gabor orienté à 45° est illustrée à la figure 3.10.

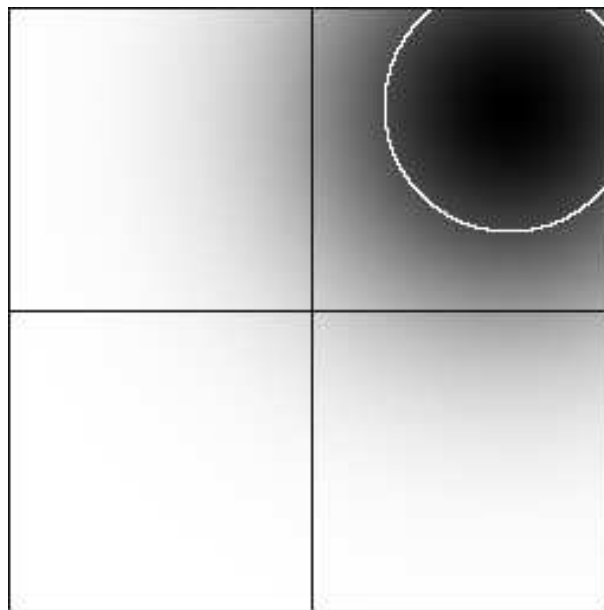


Figure 3.10: Trace d'un filtre de GABOR dans le plan spectral. Le cercle blanc représente la bande passante à 3 [dB].

$\mathcal{H}(u,v)$ est une gaussienne passe-bande de rapport d'axes $1/\lambda$, tournée d'un angle ϕ , centrée en $F = \sqrt{U^2 + V^2}$ et d'orientation $\theta = \tan^{-1}(V/U)$. On prend habituellement $\theta = \phi$.

Présentés ainsi, les filtres de GABOR jouissent de quatre degrés de liberté qui permettent d'extraire une partie de l'information spectrale de l'image. Il s'agit du rapport d'axes λ , du facteur d'excentricité σ des gaussiennes dans les deux plans, et de la translation imposée par (U,V) ou (F,ϕ) dans le domaine des fréquences.

Les filtres de GABOR sont une classe particulière des filtres linéaires, ce sont des filtres *orientés*. Ils permettent de mettre en évidence des textures ainsi que des zones homogènes d'une image. Grâce à la forme gaussienne des filtres, les enveloppes des images filtrées apportent une *information spectrale locale* en chaque pixel. De plus, elles renseignent sur le contenu énergétique de l'image dans la direction du filtre choisi.

Incertitude

L'usage des filtres gaussiens se justifie par le fait qu'ils minimisent un certain type d'incertitude. Considérons une fonction $f(x,y)$ et sa transformée de FOURIER $\mathcal{F}(u,v)$. Selon la propriété

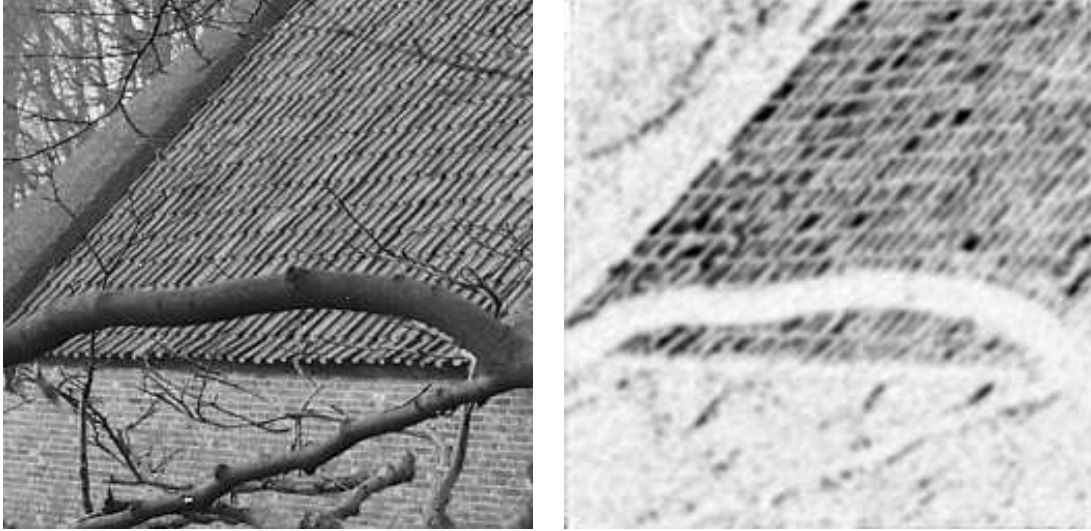


Figure 3.11: Une image originale et l'image filtré avec un filtre orienté à 135° .

d'homothétie de la transformée de FOURIER, nous avons

$$f(ax, by) \Leftrightarrow \frac{1}{|ab|} \mathcal{F}\left(\frac{u}{a}, \frac{v}{b}\right) \quad (3.15)$$

Plus une fonction est localisée dans le domaine spatial, moins elle l'est dans le domaine fréquentiel, et vice-versa. Cette propriété est similaire au principe d'incertitude d'HEISENBERG.

On peut définir les extensions effectives d'une fonction $f(x, y)$, centrée en (x_0, y_0) et dont la transformée de FOURIER est centrée en (u_0, v_0) , dans le domaine spatial Δx et Δy , et fréquentiel Δu et Δv respectivement, en considérant la racine carrée de la variance de la distribution d'énergie de cette fonction. Les variances valent

$$(\Delta x)^2 = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - x_0)^2 \|f(x, y)\|^2 dx dy}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|f(x, y)\|^2 dx dy} \quad (3.16)$$

$$(\Delta y)^2 = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (y - y_0)^2 \|f(x, y)\|^2 dx dy}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|f(x, y)\|^2 dx dy} \quad (3.17)$$

$$(\Delta u)^2 = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (u - u_0)^2 \|\mathcal{F}(u, v)\|^2 du dv}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|\mathcal{F}(u, v)\|^2 du dv} \quad (3.18)$$

$$(\Delta v)^2 = \frac{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (v - v_0)^2 \|\mathcal{F}(u, v)\|^2 du dv}{\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|\mathcal{F}(u, v)\|^2 du dv} \quad (3.19)$$

Le principe d'incertitude montre que (cf. DAUGMAN [3])

$$\Delta x \Delta y \Delta u \Delta v \geq \frac{1}{16\pi^2} \quad (3.20)$$

On peut montrer que ce produit d'incertitude est minimisé par les fonctions de GABOR qui sont des fonctions gaussiennes modulées

$$f(x, y) = e^{-\pi[a^2(x-x_0)^2 + b^2(y-y_0)^2]} e^{-2\pi j[u_0(x-x_0) + v_0(y-y_0)]} \quad (3.21)$$

$$\mathcal{F}(u, v) = e^{-\pi \left[\frac{(u-u_0)^2}{a^2} + \frac{(v-v_0)^2}{b^2} \right]} e^{-2\pi j [x_0(u-u_0) + y_0(v-v_0)]} \quad (3.22)$$

En utilisant les définitions des extensions effectives, on pourrait montrer que

$$\Delta x = \frac{1}{2a\sqrt{\pi}} \quad \Delta y = \frac{1}{2b\sqrt{\pi}} \quad (3.23)$$

$$\Delta u = \frac{a}{2\sqrt{\pi}} \quad \Delta v = \frac{1}{2\sqrt{\pi}} \quad (3.24)$$

Dès lors, quelles que soient les valeurs de a et de b , les fonctions de GABOR minimisent l'incertitude 3.20.

La fonction de GABOR (équation 3.21) est le produit d'une gaussienne elliptique de rapport b/a dont le centre est situé en (x_0, y_0) par une exponentielle complexe représentant une modulation par une fréquence radiale $\sqrt{u_0^2 + v_0^2}$ dont l'orientation est $\tan^{-1}(v_0/u_0)$. La représentation de la partie réelle d'une fonction de GABOR est donnée à la figure 3.12 tandis que sa transformée de FOURIER est représentée à la figure 3.13.

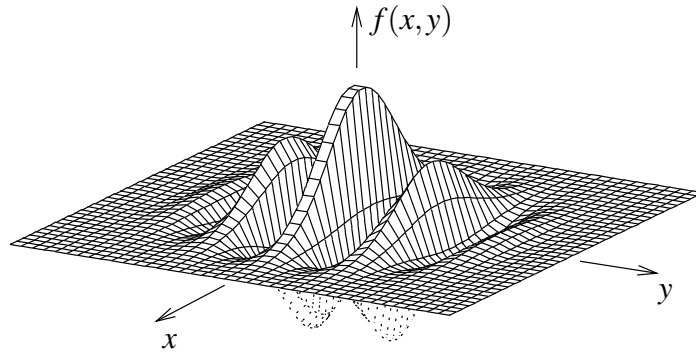


Figure 3.12: Représentation de la partie réelle d'un membre de la famille des fonctions de GABOR ($a = b, x_0 = y_0 = 0, u_0 = 0, v_0 = 2$).

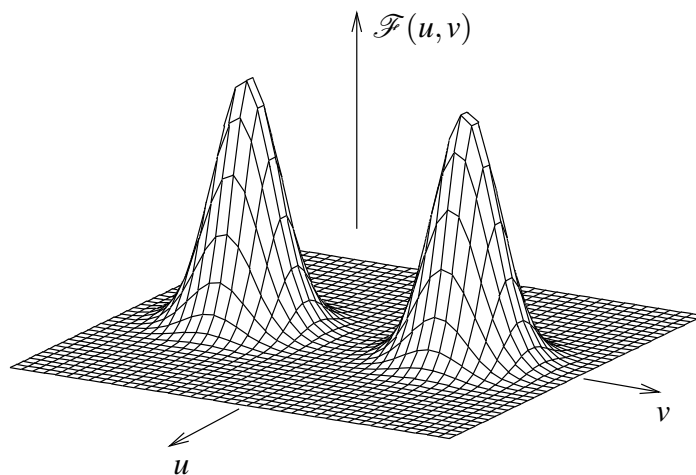


Figure 3.13: Module de la transformée de FOURIER de la fonction de GABOR représentée à la figure 3.12.

Comme le montre la figure 3.13, les fonctions de GABOR peuvent servir de filtres sélectifs, ne gardant que les composantes fréquentielles de l'image situées dans un voisinage centré en (u_0, v_0) , la

taille de ce voisinage étant fixée par les valeurs de a et de b . Il est également possible de faire subir une rotation aux fonctions de GABOR afin que les axes principaux des ellipses ne coïncident plus avec les axes $u - v$. Enfin, il est courant que l'on utilise un banc de filtres de GABOR afin de mettre en évidence différentes textures et zones homogènes qui sont présentes dans l'image.

3.2.6 Implémentation

Normalisation

L'implémentation des filtres linéaires se heurte à un problème d'échelle. En effet, la relation de PARSEVAL

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|f(x,y)\|^2 dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \|\mathcal{F}(u,v)\|^2 du dv \quad (3.25)$$

établit clairement que le filtrage de $\mathcal{F}(u,v)$ modifie l'énergie contenue dans l'image. On se retrouve à devoir gérer une éventuelle remise à niveau au cas par cas.

Implémentations d'un filtre de GABOR

Il y a principalement 4 moyens d'implémenter un filtre gaussien [42]:

1. Convolution avec un nombre restreint du noyau gaussien. Il est d'usage de choisir $N_0 = 3\sigma$ ou 5σ

$$g_{1D}[n] = \begin{cases} \frac{1}{\sqrt{2\sigma}} e^{-(n^2/2\sigma^2)} & |n| \leq N_0 \\ 0 & |n| > N_0 \end{cases} \quad (3.26)$$

2. Convolution itérative avec un noyau uniforme

$$g_{1D}[n] \simeq u[n] \otimes u[n] \otimes u[n] \quad (3.27)$$

avec

$$u[n] = \begin{cases} \frac{1}{(2N_0+1)} & |n| \leq N_0 \\ 0 & |n| > N_0 \end{cases} \quad (3.28)$$

C'est une application de la loi des grands nombres connue en statistique.

3. Multiplication dans le domaine fréquentiel.
4. Implémentation d'un filtre récursif. L'expression est fournie dans [42].

3.3 Introduction à la déconvolution

La déconvolution est un domaine vaste et bien connu. Pour la facilité de l'exposé, nous avons choisi d'illustrer la déconvolution par le biais d'une équivalence avec la notion intuitive d'extrapolation. Cette approche n'est pas générale mais elle illustre à souhait les principes de base.

3.3.1 Modèle

Pour la commodité du développement, nous utiliserons une notation unidimensionnelle. Le passage à un signal bidimensionnel ne présente pas de difficulté particulière. Les signaux de variable n sont des échantillons des signaux analogiques correspondants.

Les algorithmes d'extrapolation travaillant dans un plan transformé sélectionnent dans celui-ci des amplitudes et des positions. Malheureusement, pour le calcul du spectre, la région est étendue à un support rectangulaire si bien que le spectre d'une région de forme quelconque comprend l'information propre au motif à étendre mais aussi celle due au contour. Pour déjouer l'imbrication des deux informations, nous désirons estimer le spectre du seul motif. Au premier abord, le problème paraît insoluble car l'imbrication des informations introduit une dispersion spectrale subséquente (voir par exemple l'article de FRANKE [7], figure 2) qui interdit toute sélection efficace des fréquences. La modification spectrale due à l'imbrication est connue sous le nom d'*effet de fenêtre*. Dans une image, la fenêtre est une fonction bidimensionnelle w valant 1 à l'intérieur de l'objet et 0 à l'extérieur. La fenêtre est en quelque sorte l'*indicatrice* de la forme de l'objet traité. Pour supprimer l'influence du contour, une solution consiste à extrapoler la texture jusqu'à occupation complète d'un bloc rectangulaire. De cette façon, le réseau de données ne contient plus la moindre information concernant le contour mais uniquement l'aspect.

Soit $y(n)$, le signal observé. $y(n)$ s'obtient par '*fenêtrage*' ou '*apodisation*' de la fonction⁴ $f(n)$ à extrapoler:

$$y(n) = w(n)f(n) \quad n = 0, 1, \dots, N-1 \quad (3.29)$$

avec

$$w(n) = \begin{cases} 1 & \text{pour } n \in D_w \\ 0 & \text{si } n \notin D_w \end{cases} \quad (3.30)$$

D_w est la partie du domaine de la fonction $w(n)$ sur laquelle la valeur de la fenêtre est 1. Bien qu'inclus dans \mathbb{Z} , cet ensemble n'est pas nécessairement un intervalle. Cette précision compte pour le traitement d'image car, si le traitement s'effectue d'abord dans une direction puis dans la direction orthogonale à la première par séparation de variables, l'échantillonnage d'une ligne conduit à définir un ensemble D_w composé de plusieurs parties de \mathbb{Z} .

La détermination de la fonction $f(n)$ à l'extérieur de l'ensemble D_w requiert une modélisation plus fournie du problème sans quoi il n'offre aucune solution. Nous choisissons l'*hypothèse* qui considère que l'équation 3.29 résulte de l'action d'un système linéaire, invariant en translation. L'équation s'exprime donc aussi par une convolution dans le domaine spectral (cf. CANDY [2]). Par analogie avec le problème classique qui fait correspondre une convolution temporelle à une multiplication dans un domaine transformé et, comme les transformées directe et inverse sont des opérations duales, le problème d'extrapolation se ramène à la convolution de la transformée de $f(n)$ par celle de $w(n)$ ⁵

$$\mathcal{Y}(k) = \mathcal{W}(k) \otimes \mathcal{F}(k) \quad k = 0, 1, \dots, N-1 \quad (3.31)$$

Cette équation indique clairement que le spectre de la fonction à extrapoler est altéré par le spectre de la fenêtre $\mathcal{W}(k)$, effet dont il convient de se débarrasser au mieux. La figure 3.14 montre l'effet de fenêtrage et l'extrapolation résultant de la suppression du masque.

⁴Les fonctions dont il est question ici sont des applications de \mathbb{Z} dans \mathbb{R} . Si les niveaux de gris d'une image sont forcément des valeurs discrètes, il faut obligatoirement travailler dans \mathbb{R} durant la phase d'extrapolation, sous peine de perte de précision importante. La quantification aura lieu à la fin.

⁵Le facteur qui se place devant l'équation est régi par le type de transformée, nous évitons d'en placer un. C'est donc une relation valable à un facteur près.

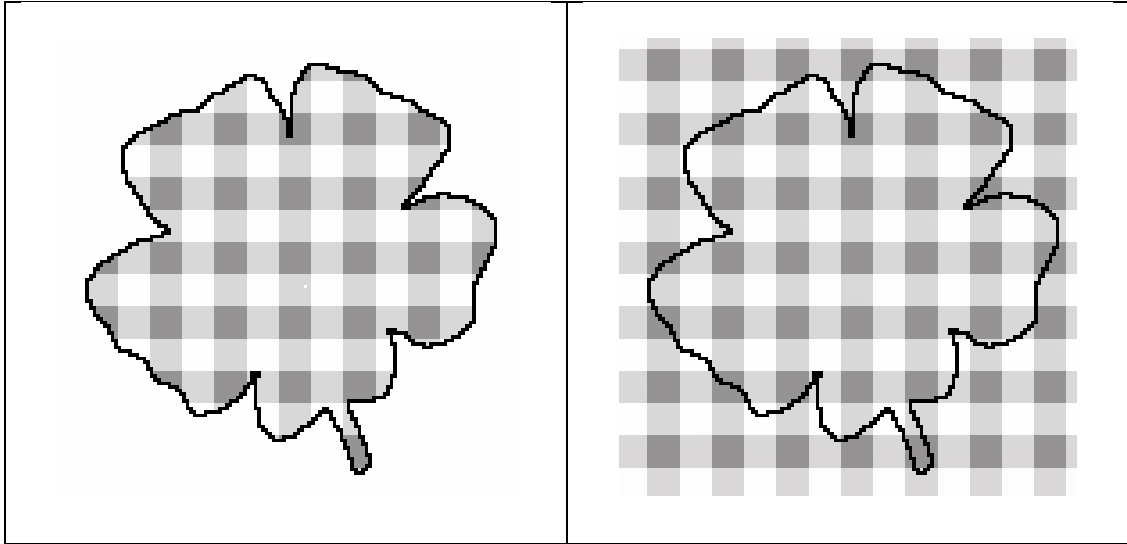


Figure 3.14: Exemple d'extrapolation d'un signal. L'image de droite est l'image extrapolée avec quelques 3% des coefficients spectraux seulement.

La démarche habituelle d'extrapolation consiste à construire d'abord une approximation ou une interpolation du signal sur son support à partir d'un modèle. Les fonctions de décomposition mises en œuvre lors de l'approximation sont ensuite étendues hors support en conservant les coefficients des fonctions exprimées à l'intérieur du domaine de départ. La littérature propose bien des techniques d'interpolation et d'approximation: surfaces-splines, méthode de SHÉPARD, techniques d'éléments finis, etc. La méthode d'extrapolation de texture que nous avons mise au point part des mêmes principes: elle évalue les paramètres significatifs de la texture sur le segment de départ et le reproduit tout autour bien que, à cette fois, le traitement ne se fasse plus dans le plan spatial mais dans le domaine spectral.

La figure 3.15 montre une ligne d'une texture⁶ $y(n)$ et le signal idéal $f(n)$ qui lui correspond dans les deux domaines. Pour bien distinguer la fenêtre $w(n)$ dans le domaine spatial, nous avons placé la valeur de la moyenne, c'est-à-dire 196, là où la fonction aurait dû s'annuler. Au vu des signaux, il est facile de comprendre que le passage du spectre $\mathcal{Y}(k)$ à celui du signal idéal est délicat, d'autant que le spectre d'une fonction $w(n)$ rectangulaire perd beaucoup d'énergie dans les lobes secondaires.

3.3.2 Forme itérative des opérateurs de déconvolution

Une simple inversion ne permet pas de résoudre le problème de la déconvolution. Et si certains auteurs affirment y parvenir (SABRI et STEENAART [29]), nous émettons de nettes réserves sur la validité des résultats rapportés pour deux raisons: les valeurs propres de la matrice d'inversion sont très proches de 1, compliquant à l'extrême les résolutions numériques, et puis, nous ne sommes pas parvenu à reproduire les résultats présentés ni à obtenir des résultats similaires sur d'autres signaux! Ceci explique sans doute pourquoi les auteurs ont eu recours à une résolution itérative, usage usuel pour les problèmes numériques. Le principe est simple: par le choix d'un opérateur O , on construit une suite de valeurs $\dots, f_i(n), f_{i+1}(n), \dots$ obtenues par

$$f_{i+1}(n) = O[f_i(n)] \quad (3.32)$$

Déterminer un opérateur itératif à partir d'une équation est assez direct. Encore faut-il que la convergence du procédé itératif soit assurée et que la solution soit unique! Ceci n'est pas garanti

⁶Il s'agit d'une ligne centrale de la texture de la figure 3.14.

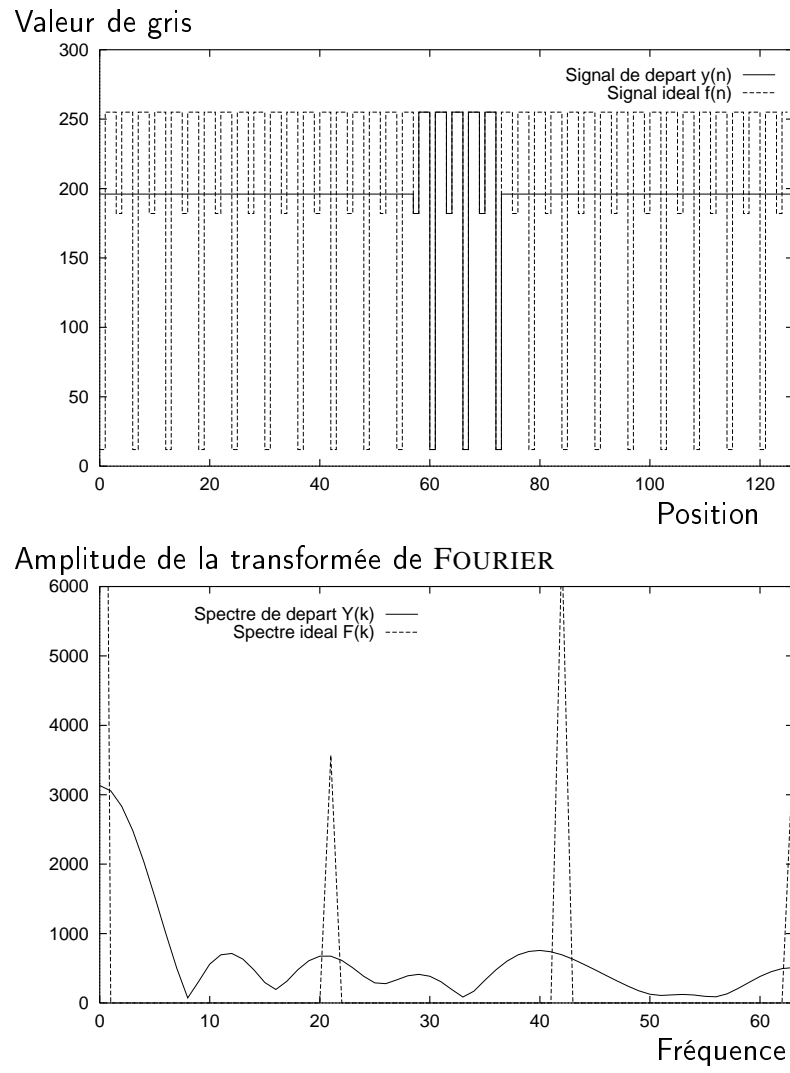


Figure 3.15: Un signal unidimensionnel dans le domaine spatial et dans le domaine spectral.

surtout que dans certains cas l'opérateur se modifie au fil des itérations. Le processus est alors du type: $f_{i+1}(n) = O_i[f_i(n)]$. Le procédé de déconvolution sélective développé plus loin est précisément un *opérateur adaptatif* de ce type.

Illustration: l'algorithme approximant une fonction par une somme de fonctions de base, croissante à chaque itération, génère un processus itératif. L'erreur d'approximation diminue à chaque étape ou reste à la rigueur constante (c'est ce qui se produit si une fonction de base nécessaire à la reconstruction est volontairement écartée).

3.3.3 Approche déterministe ou aléatoire?

En toute généralité les échantillons $f(n)$ peuvent provenir d'un signal déterministe ou aléatoire. La théorie du traitement du signal rattache quantité de méthodes et d'algorithmes à chacune de ces hypothèses. Pour y voir plus clair, nous avons dressé un tableau récapitulatif à la figure 3.16.

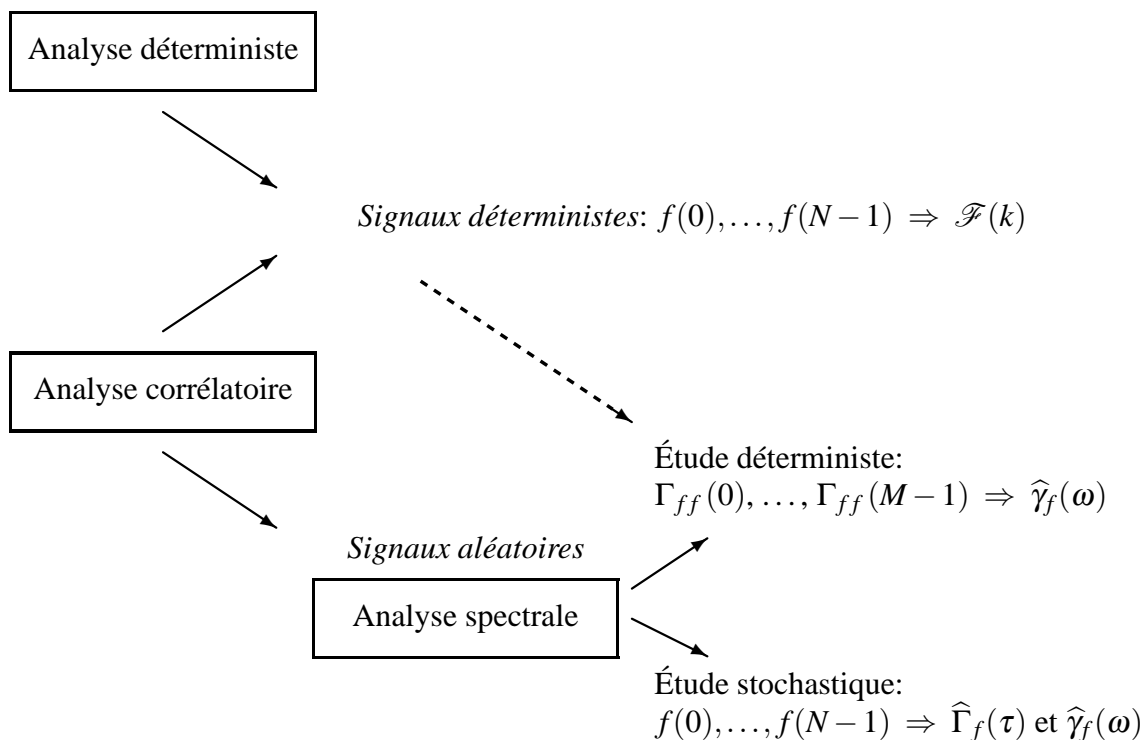


Figure 3.16: Récapitulatif des techniques de caractérisation des signaux.

L'*analyse déterministe* veille à tirer des renseignements de la séquence $f(n)$: étude de la moyenne, recherche d'une fréquence dominante dans la transformée de Fourier, conditionnement du signal pour éviter le repli de spectre, ...

L'*analyse corrélatoire*, terme emprunté à DUVAUT [6], introduit un outil de comparaison entre deux signaux décalés, la fonction d'autocorrélation $\Gamma_{ff}(\tau)$. Si ceux-ci sont déterministes, la forme de la fonction $\Gamma_{ff}(\tau)$ reste voisine des variations de $f(n)$; elle reproduit par exemple les mêmes périodicités. S'ils sont aléatoires, l'aspect de la fonction varie à chaque réalisation et dès lors ne constitue qu'une indication. Ce n'est que lorsque les signaux sont *stationnaires (faiblement au moins)*, que l'analyse corrélatoire prend son sens. À ce stade, l'*analyse spectrale* complète l'étude par une signature spectrale de la fonction d'autocorrélation; il s'agit de sa transformée de Fourier appelée *densité spectrale*, notée $\gamma_f(\omega)$, et son usage, assujéti à la même contrainte de stationnarité, indique la répartition fréquentielle de l'énergie.

Selon PAPOULIS [23], l'analyse spectrale, c'est-à-dire l'analyse corrélatoire appliquée à des signaux aléatoires stationnaires, se découpe en deux catégories dépendant de la connaissance de la fonction d'autocorrélation. Supposons que soient connus M échantillons $\Gamma_{ff}(0), \dots, \Gamma_{ff}(M-1)$. L'étude s'apparente à une étude déterministe où certains échantillons de $\Gamma_{ff}(\tau)$ sont manquants. Pour estimer la densité spectrale $\tilde{\gamma}_f(\omega)$, il faut tenir compte de cette apodisation soit en gérant l'influence de la fenêtre (méthodes non paramétriques), soit en identifiant le spectre à celui d'un modèle de données (méthodes paramétriques), ce qui conduit à une extrapolation de $\Gamma_{ff}(\tau)$ au-delà du segment connu. Souvent on utilise des modèles autorégressifs ou à moyenne mobile. Parfois le signal est supposé à bande spectrale limitée ou alors l'objectif consiste à maximiser l'entropie de la densité spectrale. Il arrive aussi qu'aucun échantillon de la fonction d'autocorrélation ne soit à disposition. Ce cas, repris sous le terme d'une *étude stochastique*, nécessite le calcul conjoint des estimées $\tilde{\Gamma}_{ff}(\tau)$ et $\tilde{\gamma}_f(\omega)$.

Ce survol des techniques spectrales suffit à comprendre la réelle difficulté de la sélection d'un modèle. Nous avons jugé préférable de restreindre le travail à une analyse déterministe. Pourquoi un tel choix?

1. La raison principale est la facilité de mise en œuvre: le recours à l'analyse spectrale nécessite un complément de modélisation, or aucune méthode ne satisfait tous les points de vue (qualité de l'estimation, réduction du bruit, minimisation des biais, etc). Il ne faudrait pas en conclure que nous perdons tout bénéfice des méthodes propres à des signaux aléatoires. En effet, les méthodes s'appuient sur des algorithmes qui eux s'appliquent à toutes les analyses. Par exemple, l'extrapolation itérative d'une séquence, à bande étroite, de signaux déterministes ou de la fonction d'autocorrélation, peut se réaliser de la même manière.
2. Un autre argument justifiant notre choix est la mise en doute de l'hypothèse de stationnarité qui sous-tend les développements de l'analyse spectrale. HUNT [13] met sa validité en cause dans le cas d'une image. Pour y remédier, il propose de stationnariser le signal, de le traiter et de revenir au signal de départ par inversion du processus de stationnarisation.
3. Enfin, si l'analyse spectrale achève une bonne discrimination de fréquences, elle remplit moins notre objectif de codage. De fait, les textures sont représentées par des échantillons de la transformée de Fourier du signal extrapolé et non de sa fonction d'autocorrélation. Une analyse corrélatoire devra donc se doubler d'un schéma pour convertir ces types d'information.

En conclusion, les développements théoriques repris dans ce chapitre dérivent d'une approche déterministe au problème d'extrapolation, quoique les algorithmes étudiés soient parfois repris par la littérature dans d'autres contextes, tel celui de l'analyse corrélatoire.

3.3.4 Algorithmes de déconvolution sélective

L'utilisation d'un procédé itératif d'extrapolation par déconvolution spectrale remonte à 1974. C'est GERCHBERG [8] qui, le premier, a appliqué la technique au traitement d'image. Quasi simultanément, PAPOULIS [22] proposa d'estimer le spectre d'un signal à bande limitée multipliée temporellement par une fonction fenêtre $w(n)$; la méthode correspond à un cas particulier de G , à savoir $G = DM^{-1}SM$. Ces auteurs partent d'un signal limité en fréquences et lui appliquent un filtre S ayant une fréquence de coupure σ_c , ce qui nécessite de connaître la valeur de la fréquence de coupure du signal non corrompu par le spectre de forme.

Le choix de la fréquence de coupure σ_c du filtre est critique: une bande passante de filtre trop petite contrarie la reconstitution du signal dans la fenêtre temporelle, et dans la situation inverse, l'algorithme perd de son efficacité et aboutit à une solution qui n'est pas unique et qui dépend de la valeur initiale ainsi que l'a montré FRANKE [7].

3.3.5 Déconvolution sélective

Résoudre l'extrapolation, quand aucune information relative au signal n'est disponible, doit s'accomplir autrement.

Nous avons suivi le principe d'extrapolation mis en œuvre par FRANKE [7] et qui est illustré à la figure 3.17; il consiste à choisir $G = DM^{-1}SM$, où S opère une sélection des coefficients transformés.

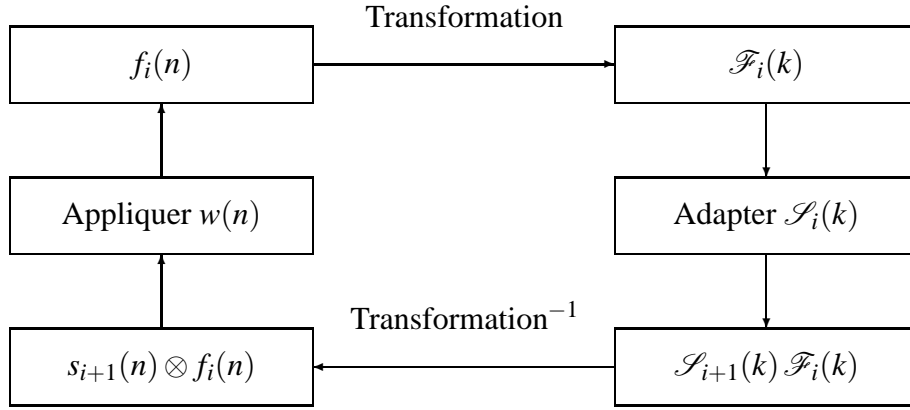


Figure 3.17: Bloc-diagramme de l'algorithme d'extrapolation par déconvolution sélective.

À la base de l'algorithme de déconvolution sélective apparaît une matrice S qui est modifiée à chaque itération. La matrice de sélection de l'étape $i + 1$ s'obtient en sélectionnant une fréquence supplémentaire non choisie dans S_i ; c'est-à-dire que, si k_s est la position du nouveau coefficient choisi, alors

$$S_{i+1}(k) = \begin{cases} 1 & \text{si } k = k_s \\ S_i(k) & \text{si } k \neq k_s \end{cases} \quad (3.33)$$

FRANKE suggère de sélectionner à chaque étape le plus grand coefficient non sélectionné auparavant. En remplissant progressivement la matrice S , le filtre est devenu *adaptatif*. Remarquons qu'il faut s'arranger pour ne conserver, à la fin du processus, qu'un faible nombre de coefficients spectraux, inférieur à la taille de la fenêtre. Dans le cas contraire, on dépasserait le nombre de raies strictement nécessaires pour décrire le signal.

DOCAMPO et FIGUEIRAS [5] agissent à l'opposé de FRANKE en éliminant les raies peu significatives à partir d'un algorithme du gradient. Si les résultats semblent acceptables, les premières étapes du processus itératif bafouent la condition d'unicité stricte et présentent de ce fait un risque non négligeable d'instabilité. Au contraire, la déconvolution sélective extrait en premier lieu les raies importantes d'où il résulte que, en conformité avec le théorème de Parseval, l'erreur diminue au fur et à mesure.

3.3.6 Convergence et unicité de la déconvolution sélective

Un théorème[38] établit la convergence et l'unicité d'un opérateur fixe. Or dans l'algorithme de déconvolution, l'opérateur se modifie durant les étapes et, en toute rigueur, le théorème ne s'y applique pas; il faudrait un nombre de boucles internes, autrement dit sans modification de S , infini avant de modifier l'opérateur. La pratique montre que le signal se stabilise après quelques boucles et que, en outre, leur nombre peut être réduit durant le processus de déconvolution. Pour assurer une démarche rigoureuse et pratique, nous avons adopté un algorithme itératif dégressif. Partant du signal original, nous sélectionnons une raie et effectuons K boucles internes; $K = 10$ dans nos essais. Au

moment d'ajouter une autre raie, K passe à $K - 1$ et ainsi de suite jusqu'au moment où $K = 1$. Cette démarche satisfait les conditions du théorème et conduit à de très bons résultats alors que prendre dès le départ $K = 1$ produit une interférence entre la stabilisation et la déconvolution, qui se répercute sur la qualité du résultat.

La figure 3.18 montre l'évolution du spectre de la figure 3.15 durant le processus de déconvolution. Le spectre de départ est représenté au temps $t = 0$. Puis l'algorithme sélectionne la composante DC; pour la clarté du dessin, nous avons écrêté l'amplitude de cette composante. L'effet de la déconvolution s'observe déjà près de l'origine. L'algorithme détecte ensuite les fréquences 42 et 21 au temps respectifs $t = 2$ et $t = 3$. Il en résulte un aplatissement important du spectre. Remarquons que l'amplitude des raies sélectionnées augmente au fil du temps avant de saturer, ce qui signifie qu'il ne faut pas s'arrêter trop tôt. D'autant que la fréquence 1, extraite au temps $t = 4$, ne représente plus qu'une faible partie du contenu énergétique du signal, fournissant un critère d'arrêt facile à mettre en œuvre.

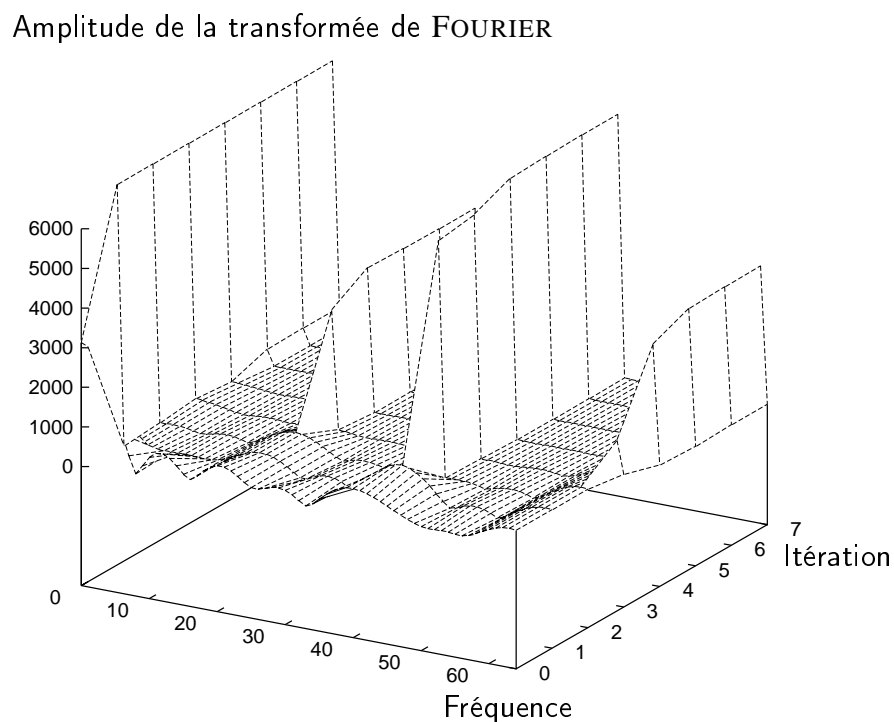


Figure 3.18: Évolution du spectre d'un signal unidimensionnel durant le processus de déconvolution sélective.

Les résultats de la déconvolution sélective d'une image sont présentés dans le tableau 3.1. Il convient de remarquer que les premières étapes de la déconvolution fournissent déjà un résultat très acceptable $-s_4(n) \otimes f_3(n)$ est obtenu avec seulement 4 raies spectrales !-, ce qui s'accompagne d'une réduction de l'entropie du spectre. La déconvolution atténue la dispersion du spectre autour des raies sélectionnées.

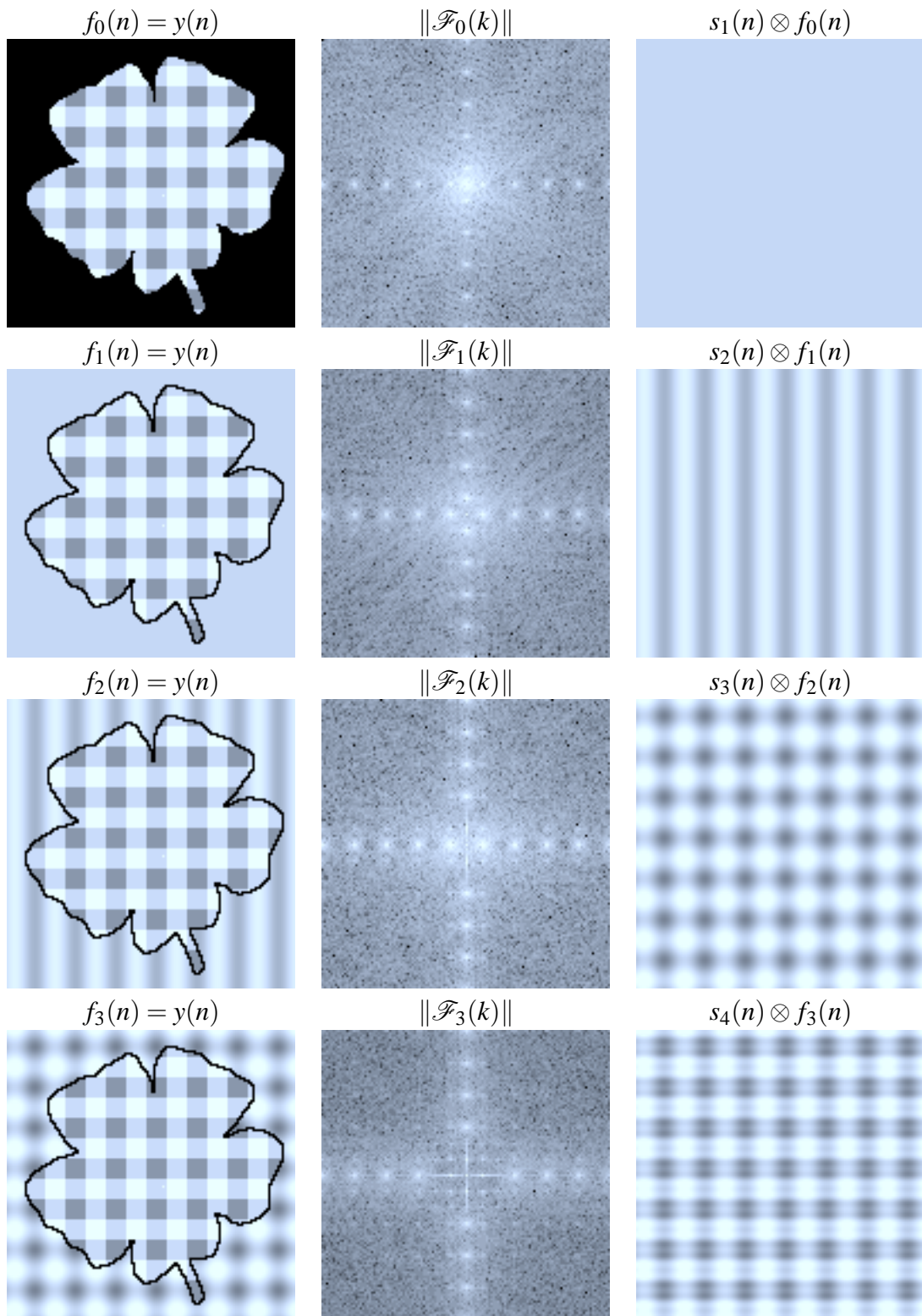


Table 3.1: Premières étapes de l'algorithme de déconvolution sélective.

Chapter 4

Morphologie mathématique

4.1 Introduction

L'analyse d'image a pour objectif la description aussi complète que possible des caractéristiques d'une image. Mais une image est constituée d'objets. Pour extraire des informations quantitatives de ces objets, plusieurs attitudes sont possibles: adopter une démarche relativement passive en déterminant *tous* les paramètres relatifs à cette image –cette démarche aboutit à un surplus informatif–, ou alors se contenter de quelques paramètres suffisants à caractériser l'image. L'approche morphologique s'inscrit dans ce second courant de pensée. De nature fondamentalement géométrique, elle consiste à *comparer les objets à analyser à un autre objet de forme connue*, appelé *élément structurant*. De la sorte, chaque élément structurant fait apparaître l'image sous un jour nouveau; tout l'art revient à choisir le bon éclairage.

4.1.1 Historique

La description morphologique repose sur des notions purement ensemblistes. C'est en 1901 que H. MINKOWSKI balbutie les premiers concepts d'une théorie, qui sous l'impulsion de H. HADWIGER, A. HAAS, G. MATHERON et J. SERRA, enfantera la morphologie mathématique.

Trois aspects la caractérisent: l'*aspect algébrique* traitant les opérations ensemblistes et plus généralement les transformations, l'*aspect probabiliste* qui décrit des modèles aléatoires applicables aux ensembles et celui de la *géométrie intégrale*, lié à l'étude de fonctionnelles d'images.

Actuellement la théorie morphologique jouit d'une place de choix à côté des théories que sont

- le *traitement de signal classique*, nous entendons par là celle qui fait usage des opérateurs linéaires et s'occupe de filtrage linéaire ou convolution, des transformées, ... ,
- les *méthodes statistiques* où l'accent est placé sur l'extraction de paramètres et non sur la transformation d'une image en une autre image, et
- les *méthodes syntaxiques*, qui extraient des primitives significatives de l'image, primitives souvent géométriques, et leur associent des règles grammaticales.

La théorie de la morphologie est largement diffusée dans la communauté internationale de nos jours comme l'indique la kyrielle de conférences, d'articles et d'applications industrielles qui l'utilisent. Par ailleurs, son champ d'application couvre quantité de domaines: imagerie médicale, sciences des matériaux, vision robotique, télédétection, analyse topographique, reconnaissance de caractères, ...

L'objectif de ce chapitre est de présenter les éléments théoriques essentiels, permettant de manipuler les concepts de la morphologie mathématique dans des applications. L'étude ne se veut en aucun cas exhaustive.

Nous commençons par un rappel de notions de la théorie des ensembles. La section suivante définit les opérations élémentaires de morphologie mathématique sur des ensembles et décrit leurs propriétés. Nous étudierons ensuite le cas des fonctions en niveaux de gris.

4.2 Rappels sur les notions d'ensembles

Les ensembles seront notés par des majuscules A, B, \dots et les éléments qu'ils comprennent par des minuscules a, b, \dots . Faisons un bref rappel sur des notions de la théorie des ensembles.

4.2.1 Égalité d'ensembles

Deux ensembles sont *égaux* s'ils sont formés des mêmes éléments: $X = Y \Leftrightarrow (x \in X \Rightarrow x \in Y \text{ et } x \in Y \Rightarrow x \in X)$. L'égalité entre ensembles est réflexive, symétrique et transitive. L'ensemble vide sera noté \emptyset .

4.2.2 Inclusion

X est un sous-ensemble de Y (X *inclus* dans Y) si tous les éléments de X appartiennent à Y : $X \subseteq Y \Leftrightarrow (x \in X \Rightarrow x \in Y)$. L'inclusion est réflexive, antisymétrique et transitive.

4.2.3 Intersection

L'*intersection* de deux ensembles X et Y est l'ensemble des éléments qui appartiennent aux deux: $X \cap Y = \{x \text{ tel que } x \in X \text{ et } x \in Y\}$. L'intersection est commutative, associative et *idempotente*, ce dernier terme traduit $X \cap X = X$.

4.2.4 Union

La *réunion* de deux ensembles est constituée des éléments appartenant à l'un ou à l'autre, c'est-à-dire $X \cup Y = \{x \text{ tel que } x \in X \text{ ou } x \in Y\}$. La réunion d'ensembles est commutative, associative et idempotente. Remarquons que contrairement à l'intersection, la réunion de deux ensembles dont l'un n'est pas vide, n'est jamais vide.

4.2.5 Différence

Étant donnés X et Y , la *différence* de X par Y , notée $X - Y$ ou $X \setminus Y$ est l'ensemble des éléments de X qui n'appartiennent pas à Y : $X - Y = \{x | x \in X \text{ et } x \notin Y\}$. La différence entre ensembles n'est en général pas commutative et ne possède pas de bonnes propriétés. Certains utilisent parfois une autre loi de composition, appelée *différence symétrique*, notée $X \Delta Y$, et définie par $X \Delta Y = X \cup Y - X \cap Y$. C'est donc l'ensemble des éléments qui appartiennent à l'un ou à l'autre mais jamais aux deux ensembles. La différence symétrique est une opération commutative et associative.

Après le rappel de ces opérations d'ensembles usuelles, nous introduisons des opérations qui seront utiles pour la définition d'opérateurs morphologiques.

4.2.6 Complémentaire

Soit un sous-ensemble X d'un ensemble \mathcal{E} servant de référentiel, le *complémentaire* de X dans \mathcal{E} est le sous-ensemble noté X^c , fourni par $X^c = \{x \text{ tel que } x \in \mathcal{E} \text{ et } x \notin X\}$. La figure 4.1 illustre la notion de complémentaire. Si le référentiel \mathcal{E} est modifié, le complémentaire diffère.

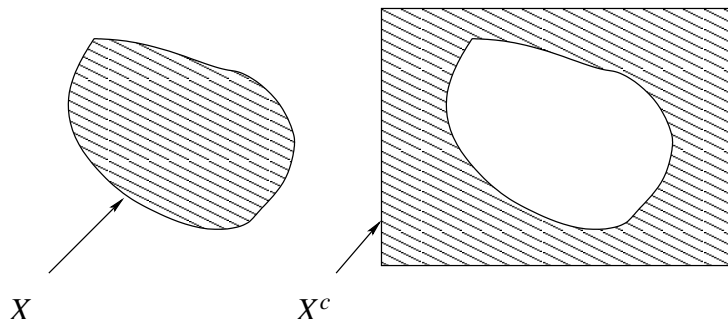


Figure 4.1: Complémentaire d'un ensemble.

Tout comme pour les opérateurs logiques, il existe des formules de MORGAN où intervient le complémentaire: le complémentaire d'une intersection est égal à l'union des complémentaires $(X \cap Y)^c = X^c \cup Y^c$ et vice versa.

La notion de complémentarité est importante. Elle permet de faire un lien entre un traitement effectué sur un objet (l'ensemble X) et l'arrière-plan. Quand on peut traiter indifféremment un objet ou son complémentaire, on parle d'*auto-dualité*. Remarquons que les traitements sont rarement auto-duaux.

4.2.7 Symétrie

Soit un référentiel \mathcal{E} sur lequel on a défini l'addition $+$; on dispose alors d'une structure mathématique plus riche (un *groupe* par exemple), permettant de définir de nouveaux opérateurs. L'opposé de l'addition est la soustraction.

Le symétrique \check{X} de l'ensemble X est défini par $\check{X} = \{-x | x \in X\}$.

4.2.8 Translaté

Le translaté de X par b vaut $\{z \in \mathcal{E} | z = x + b, x \in X\}$. Comme le montre la figure 4.2, la translation d'un ensemble consiste à déplacer l'ensemble dans le référentiel.

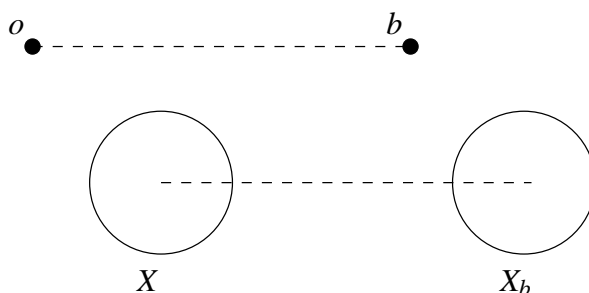


Figure 4.2: Translation d'un ensemble par un élément b .

4.2.9 Addition et soustraction de MINKOWSKI

Soient X et B deux ensembles quelconques de l'espace \mathcal{E} . À tout $x \in X$ et $b \in B$, il est possible de faire correspondre la somme algébrique $x + b$. Choisisant tous les x et b , on forme un nouvel ensemble appelé *addition de MINKOWSKI* et noté $X \oplus B$. Cet ensemble se définit par $X \oplus B = \{z \in \mathcal{E} \mid z = x + b, x \in X, b \in B\}$.

Du fait de la commutativité de l'addition, $X \oplus B = \bigcup_{x \in X} B_x = \bigcup_{b \in B} X_b = B \oplus X$. Par analogie, il est possible de définir la *soustraction* de deux ensembles notée $X \ominus B$, l'opération $X \ominus B = \bigcap_{b \in B} X_{-b}$.

4.2.10 Propriétés de base des transformations morphologiques

Toute opération morphologique est par construction le résultat d'une ou plusieurs transformations sur des ensembles, en abrégé $\psi(X)$ suivi éventuellement d'une mesure $P(\psi(X))$. Assez grossièrement, une transformation morphologique est vue comme la réalisation d'opérations ensemblistes (union, intersection, complémentarité, ...) faisant intervenir X et l'ensemble ou *élément structurant* B déplacé à travers tout l'espace \mathcal{E} , opérations dont le résultat s'interprète pour tout point de cet espace. Pour que l'outil de transformation ait une quelconque utilité, les opérations satisfont souvent quatre principes (les opérateurs morphologiques introduits par la suite ne rencontrent pas tous ces principes):

1. *Invariance en translation.*

La translation temporelle d'un signal introduit un facteur de phase dans la transformée de FOURIER, facteur qu'il est possible d'extraire et de neutraliser. Dans le même état d'esprit, il paraît logique d'imposer l'invariance par translation des transformations morphologiques, traduite mathématiquement par la relation $\psi(X_b) = (\psi(X))_b$. Il est clair qu'ainsi l'origine de l'espace ne joue plus le moindre rôle privilégié et que les seules transformations admises seront insensibles à la position des coordonnées.

2. *Compatibilité avec les homothéties.*

Les transformations dépendent parfois d'un facteur d'échelle positif λ . Soient λX un ensemble homothétique de X dans \mathbb{R}^n et ψ_λ une transformation de paramètre λ . La transformation est compatible avec les homothéties s'il est équivalent d'appliquer ψ_λ à l'ensemble X dilaté du facteur λ ou de dilater le résultat de la transformée de X pour $\lambda = 1$. Cela se traduit par la relation: $\psi_\lambda(\lambda X) = \lambda \psi(X)$ où $\psi = \psi_1$. À titre d'exemple, la transformation consistant à extraire d'un ensemble les cordes horizontales de longueur λ , est compatible avec les homothéties. Ce principe concerne toutes les transformations où opère un paramètre métrique.

3. *Connaissance locale.*

Le principe de connaissance locale intervient lorsque le champ de mesures ne contient qu'une partie de l'ensemble X : c'est le cas de structures observées au microscope. Grâce à cette propriété, la complexité en calcul est réduite et l'extérieur de l'image peut être adapté de manière à faciliter l'application, sans que cela n'ait de conséquence directe sur le contenu informatif de l'image; c'est la notion de *géodésie* sur laquelle nous reviendrons ultérieurement.

Soit M l'ensemble borné fermé (de volume fini) définissant le champ de mesures. L'ensemble X est connu à l'intérieur de ce champ, car on dispose de $X \cap M$. La transformation ψ vérifie le principe de connaissance locale si, quel que soit l'ensemble M borné, il existe N borné, ne dépendant pas de X , tel que $\psi(X \cap N) \cap M = \psi(X) \cap M$; ce qui signifie que le comportement de ψ à l'intérieur de M ne dépend que de ce qu'il y a dans N . La condition de connaissance locale élimine toutes les transformations qui concernent des maxima (longueur ou surface maximales, ...). Pour l'étude de signaux à plusieurs niveaux de gris, la notion de connaissance est étendue à ce que nous appellerons "connaissance locale simple", qui offre la possibilité de

reconsidérer les extrema de fonctions, notion indispensable pour quiconque désire étendre le cadre de la mathématique morphologique. Le principe de connaissance locale simple dit qu'il existe une taille du masque d'observation Z au-delà de laquelle la connaissance locale est assurée. Pour les ensembles N dont la taille est inférieure à celle de Z , la connaissance locale n'est pas satisfaite. C'est par l'intermédiaire des extrema qu'il est possible de montrer que toute convolution est un cas particulier de filtrage morphologique.

4. Continuité uniforme.

Considérons deux ensembles X et Y ; intuitivement, une transformation est dite *continue* lorsque, si en se rapprochant de plus en plus de X , la transformée de Y ($\psi(Y)$) se rapproche de plus en plus de la transformée de X ($\psi(X)$): $X \rightarrow Y \Rightarrow \psi(X) \rightarrow \psi(Y)$. Formellement la *continuité* est obtenue lorsque, pour tout voisinage V de $\psi(X)$, il existe un voisinage V' de X tel que pour $Y \in V'$, $\psi(Y) \in V$. Très clairement la continuité dépend de la notion de voisinage, c'-à-d. de la topologie.

SERRA parle aussi de semi-continuité.

Dans la suite, nous verrons apparaître des correspondances avec les propriétés énoncées ci-dessus, parfois pour constater qu'elles ne sont pas toutes respectées par certains opérateurs morphologiques!

4.3 Transformations morphologiques élémentaires

Les développements repris ci-après concernent des ensembles et non des fonctions. Les définitions seront étendues plus tard à des fonctions.

4.3.1 Érosion

La transformation par érosion est la première à avoir été proposée; c'est d'ailleurs la plus importante avec la dilatation. Pour définir intuitivement l'opération d'érosion, situons-nous dans l'espace euclidien \mathbb{R}^2 partiellement occupé par un ensemble X . Prenons un élément structurant B représentant une figure géométrique simple, par exemple un disque. Cet élément B_z est repéré par son centre et placé en z dans l'espace \mathbb{R}^2 . Il est ensuite déplacé de telle sorte que son centre occupe successivement toutes les positions de l'espace. Pour chacune de ces positions, la question suivante est posée: l'ensemble B est-il entièrement inclus dans l'ensemble X ($B_z \subseteq X$) pour cette position de z ? L'ensemble des z fournissant une réponse positive forme un nouvel ensemble appelé *érosion de X par B* et noté $X \ominus B$.

Définition 11 Érosion morphologique

$$X \ominus B = \{z \in \mathcal{E} \mid B_z \subseteq X\} \quad (4.1)$$

La notion d'érosion correspond à la notion de soustraction introduite précédemment; par ailleurs, on montre aisément que la définition suivante, de forme plus algébrique, fournit le même résultat

Définition 12

$$X \ominus B = \bigcap_{b \in B} X_{-b} \quad (4.2)$$

Ce qui signifie que la transformation par érosion d'un ensemble X par B s'obtient en translatant X par l'opposé de chaque élément de B et en ne conservant que les points appartenant à toutes les

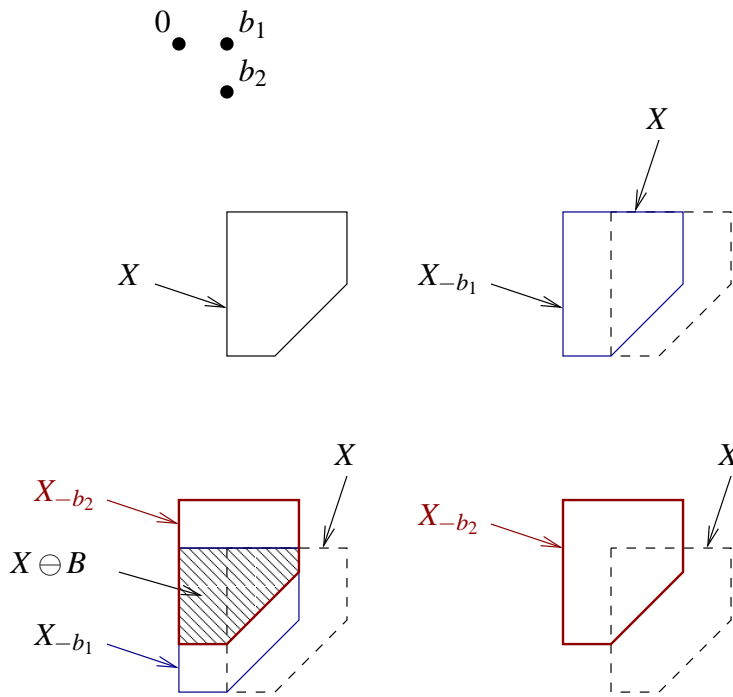


Figure 4.3: Interprétation algébrique de l'érosion.

translations de X . La figure illustre cette formulation algébrique. Avec cette interprétation, il est clair que si la taille de B dépasse celle de X , l'ensemble X érodé par B est vide.

Parfois l'opération d'érosion est définie comme la soustraction d'ensembles où B a préalablement subi une symétrie centrale autour de l'origine (cf. SERRA [31]), dans le but d'accroître le formalisme de dualité inhérent à la plupart des transformées morphologiques.

L'opération d'érosion est illustrée par la figure 4.4. On voit que la petite composante de X a disparu; en effet, elle ne peut contenir le disque B . Les effets habituels de l'érosion sont: la séparation des objets au droit des étranglements, le rétrécissement des objets de taille supérieure à B et la disparition des petites composantes.

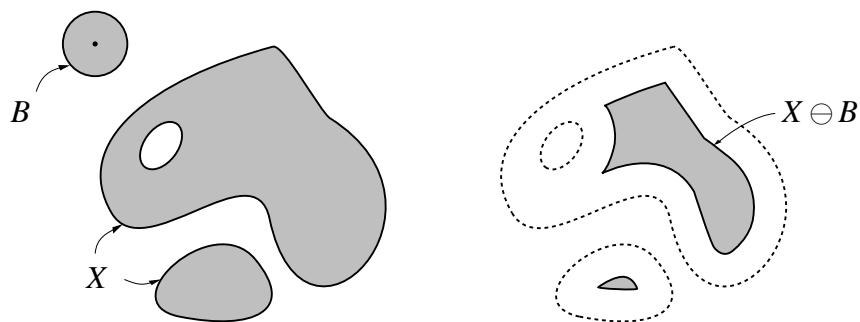


Figure 4.4: Érosion de X par un disque B . L'origine de l'élément structurant est représentée par un point noir.

4.3.2 Dilatation

L'opération de *dilatation*, appelée *dilation* en anglais, due à MINKOWSKI se définit de manière analogue. En prenant le même élément structurant B , pour chaque position z du centre de B la question

est: l'ensemble B_z touche-t-il X ($X \cap B_z \neq \emptyset$)? L'ensemble des points de \mathcal{E} correspondant à une réponse positive forme le nouvel ensemble $X \oplus \check{B}$ – l'ensemble \check{B} s'obtient par réflexion de l'ensemble B autour de l'origine – dont la frontière est le lieu des centres géométriques de B_z lorsque B_z touche X .

Définition 13 *En termes ensemblistes, la dilatation est l'union d'ensembles translatés:*

$$X \oplus B = \bigcup_{b \in B} X_b = \bigcup_{x \in X} B_x = \{x + b \mid x \in X, b \in B\} \quad (4.3)$$

Le principe de la dilatation, à savoir l'union des translatés de X par les éléments de B , est illustré à la figure 4.5.

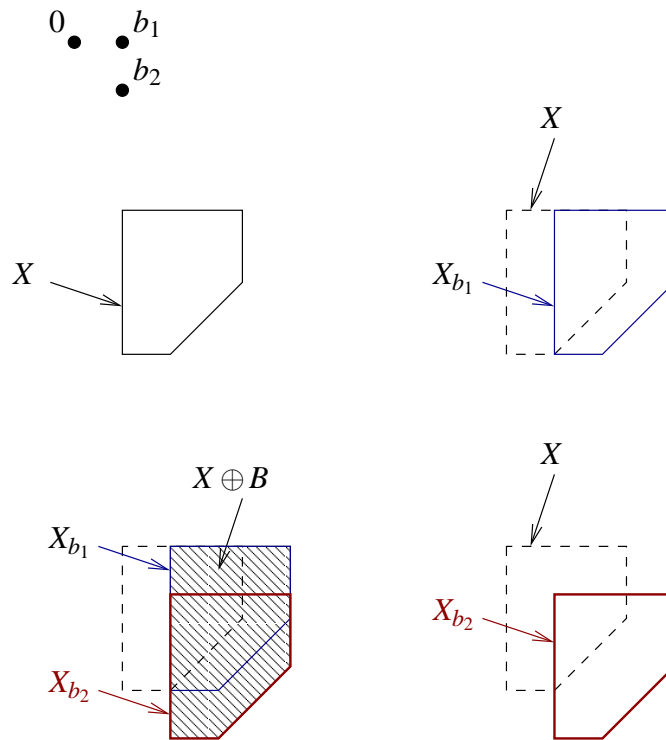


Figure 4.5: Illustration de la définition algébrique de la dilatation.

La figure 4.6 montre la dilatation par un disque du même ensemble X que celui traité lors de l'érosion. Dans cet exemple, les deux composantes connexes¹ sont réunies dans le dilaté; tout comme l'érosion, la dilatation ne préserve pas la connexité.

Propriété 14 [Dualité] *La transformation de dilatation n'est pas indépendante de l'opération d'érosion. Éroder X revient à dilater son complémentaire X^c . Les opérations de dilatation et d'érosion sont donc deux opérations duales par rapport à la complémentation:*

$$X \ominus \check{B} = (X^c \oplus B)^c \quad (4.4)$$

$$X \ominus B = (X^c \oplus \check{B})^c \quad (4.5)$$

¹Un ensemble X est *connexe* si $\forall x_1, x_2 \in X$, il existe un chemin reliant x_1 à x_2 totalement inclus dans X .

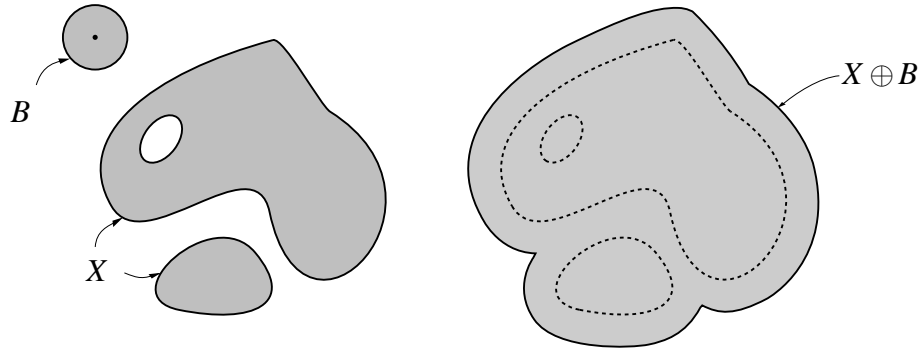


Figure 4.6: Dilatation de X par un disque B .

Avec nos définitions d'érosion et de dilatation, la présence du symétrique de B est inévitable. C'est pour l'éviter que certains auteurs adoptent parfois une autre définition de l'érosion.

L'exploitation de cette propriété permet bien souvent de transposer le travail fait pour une érosion à celui d'une dilatation et vice-versa. De plus, la pratique favorise l'emploi d'éléments symétriques, d'où découle une forme simplifiée du principe de dualité par complémentation.

Propriétés des opérations d'érosion et de dilatation

Les opérations d'érosion et de dilatation vérifient les principes des transformations morphologiques idéales:

1. l'érosion et la dilatation sont invariantes en translation: de fait, il revient au même de translater un ensemble X puis de l'éroder, que de l'éroder puis de translater l'image transformée: $X_z \ominus B = (X \ominus B)_z$. De même, $X_z \oplus B = (X \oplus B)_z$;
2. l'érosion et la dilatation sont compatibles avec les homothéties: $\lambda X \ominus \lambda B = \lambda(X \ominus B)$ et $\lambda X \oplus \lambda B = \lambda(X \oplus B)$;
3. l'érosion et la dilatation vérifient la condition de connaissance locale pour B borné;
4. nous admettrons sans démonstration que l'érosion et la dilatation sont des transformations continues. Pour une démonstration rigoureuse, le lecteur consultera par exemple l'ouvrage de SERRA [31]. Dans la pratique, la mathématique morphologique s'utilise sur des images discrétisées; la discrétisation escamote considérablement les problèmes de continuité.

Les deux opérations morphologiques jouissent aussi de propriétés algébriques:

- l'érosion et la dilatation sont des opérations *croissantes*: si $X \subseteq Y$, alors $(X \ominus B) \subseteq (Y \ominus B)$ et $(X \oplus B) \subseteq (Y \oplus B)$;
- si l'élément structurant contient l'origine, l'érosion est une opération *anti-extensive* alors que la dilatation est *extensive*, c'-à-d. $X \ominus B \subseteq X$ et $X \subseteq X \oplus B$.

4.3.3 Ouverture morphologique

Puisque l'érosion et la dilatation sont des opérations morphologiques, il vient à l'esprit de combiner ces opérations élémentaires.

Définition 15 L'opération obtenue par la succession d'une érosion et d'une dilatation est l'ouverture morphologique². Elle se note

$$X \circ B = (X \ominus B) \oplus B \quad (4.6)$$

En général, l'ensemble traité diffère de l'ensemble de départ: l'ensemble ouvert est plus régulier et moins riche en détails que l'ensemble initial. La transformation par ouverture adoucit donc les contours. L'ouverture joue le rôle d'un filtrage que nous retrouverons d'une façon formelle plus loin.

4.3.4 Fermeture morphologique

Définition 16 En inversant l'ordre des opérations utilisées pour définir l'ouverture, nous obtenons une nouvelle opération appelée fermeture; c'est la transformation

$$X \bullet B = (X \oplus B) \ominus B \quad (4.7)$$

Il s'agit bien de l'opération duale de l'ouverture pour la complémentation, car $(X \circ B)^c = X^c \bullet \check{B}$ et $(X \bullet B)^c = X^c \circ \check{B}$. La fermeture d'un ensemble effectue donc une opération de filtrage similaire à celle effectuée par l'ouverture mais sur le complémentaire de l'ensemble. L'ouverture et la fermeture sont illustrées par la figure 4.7.

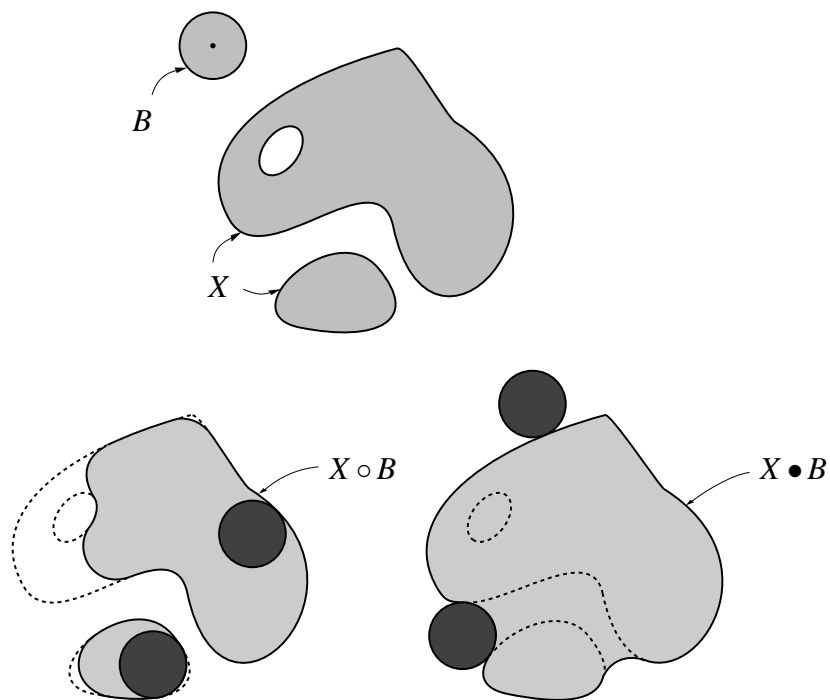


Figure 4.7: Ouverture et fermeture de X par un disque B .

Propriétés de l'ouverture et de la fermeture

Par construction, l'ouverture et la fermeture respectent les quatre principes de toute transformation morphologique. Concernant les propriétés algébriques, citons les propriétés suivantes, qui sont toutes duales pour $X \circ B$ et $X \bullet B$:

²Il faut encore distinguer l'ouverture morphologique de l'ouverture algébrique. Cette dernière n'est pas la résultante d'une érosion suivie d'une dilatation mais elle possède les mêmes propriétés algébriques.

- l'ouverture et la fermeture sont des transformations *croissantes*. Si $X \subseteq Y$, alors

$$(X \circ B) \subseteq (Y \circ B) \text{ et } (X \bullet B) \subseteq (Y \bullet B) \quad (4.8)$$

- l'ouverture est *anti-extensive* alors que la fermeture est extensive

$$X \circ B \subseteq X, \quad X \subseteq X \bullet B \quad (4.9)$$

- l'ouverture et la fermeture sont des opérateurs *idempotents* (ou opérateurs de projection), ce qui revient à dire que

$$(A \circ B) \circ B = A \circ B \text{ et } (A \bullet B) \bullet B = A \bullet B \quad (4.10)$$

Interprétation de l'ouverture et de la fermeture

L'interprétation la plus commode de l'opération d'ouverture est illustrée par la propriété que voici:

$$X \circ B = \bigcup \{B_z | z \in \mathcal{E} \text{ et } B_z \subseteq X\} \quad (4.11)$$

Ainsi, l'ouverture d'une figure par un élément structurant B est l'ensemble des points recouverts lors du déplacement de B à l'intérieur de la figure. Une propriété similaire vaut par dualité pour la fermeture; mais, à cette fois, l'élément structurant parcourt le complémentaire de la figure.

4.3.5 Propriétés générales

Pour décrire des réalisations logicielles ou matérielles, il est utile de mettre en lumière quelques propriétés supplémentaires (les démonstrations manquantes sont disponibles dans l'article d'HARALICK [10]):

Propriété 17 *La dilatation est commutative et associative*

$$X \oplus B = B \oplus X \quad (4.12)$$

$$(X \oplus Y) \oplus C = X \oplus (Y \oplus C) \quad (4.13)$$

Démonstration. [Associativité]

$z \in X \oplus (Y \oplus C)$ si et seulement si (ssi) il existe $x \in X, y \in Y$ et $c \in C$ tel que $z = x + (y + c)$. Mais comme l'addition dans \mathcal{E} est une opération associative $z = (x + y) + c$, ce qui suffit à démontrer la propriété. ■

Remarques concernant l'implémentation des opérations morphologiques élémentaires

Dans la pratique, X et B sont des objets très différents puisque B est l'*élément structurant*. La dilatation par un carré de taille 3×3 est une opération aisée à réaliser dans une architecture actuelle. La propriété d'associativité est fondamentale pour ce qui est du nombre d'opérations à réaliser lorsque X est analysé par $(B \oplus C)$. Soit N une dimension caractéristique de B et C (largeur ou hauteur), une application brutale de la dilatation de X par $B \oplus C$ nécessite N^2 opérations alors que la succession de la dilatation $X \oplus B$ suivie de la dilatation par C requiert $2N$ opérations. Cependant l'intérêt reste restreint pour des architectures qui effectuent la dilatation par accès aléatoire à tous les points du voisinage.

Des efforts de recherche constants sont entrepris pour réduire le nombre d'opérations nécessaires. Par exemple, VINCENT [41] a mis au point un algorithme souple et performant pour déterminer les

dilatations et érosions d'images binaires avec des éléments structurants quelconques. La première étape consiste en un suivi de contours de l'image et en leur codage sous forme de boucles. Dans un second temps, l'élément structurant représenté par une structure de données appropriée, est efficacement propagé le long de ces boucles. D'autres références [37] proposent un algorithme de complexité N et un autre algorithme indépendant de la taille de l'élément structurant sur base de la manipulation d'un histogramme local. Ces algorithmes permettent aussi d'effectuer une ouverture en une seule passe. ZHUANG [43] propose une méthode utilisable pour des éléments structurants convexes (c'est ce qu'on appelle des "STEINER") du type $B = (B_1 \oplus B_2)$. En fait, la plupart des propriétés citées ci-après induisent un algorithme particulier permettant une certaine optimisation des opérations morphologiques. Tantôt, l'élément structurant est décomposé par dilatation, tantôt par intersection ou encore par union.

Propriété 18 *La dilatation distribue l'union*

$$\left(\bigcup_j X_j\right) \oplus B = \bigcup_j (X_j \oplus B) \quad (4.14)$$

Démonstration.

Nous présentons la démonstration pour une union qui se limite à deux termes X et Y , à savoir $(X \cup Y) \oplus B = (X \oplus B) \cup (Y \oplus B)$.

$$(X \cup Y) \oplus B = \bigcup_{z \in X \cup Y} B_z = \left(\bigcup_{z \in X} B_z\right) \cup \left(\bigcup_{z \in Y} B_z\right) = (X \oplus B) \cup (Y \oplus B) \quad (4.15)$$

■

Par commutativité, $X \oplus (B \cup C) = (X \oplus B) \cup (X \oplus C)$. Cette égalité est importante; elle permet la décomposition d'un élément structurant en une union d'éléments structurants, offrant ainsi une alternative dans l'évaluation du coût d'une réalisation.

Propriété 19 *L'érosion est distributive par rapport à l'intersection*

$$\left(\bigcap_j X_j\right) \ominus B = \bigcap_j (X_j \ominus B) \quad (4.16)$$

Démonstration. Utilisons le formalisme algébrique:

$$\left(\bigcap_j X_j\right) \ominus B = \bigcap_{b \in B} \bigcap_j X_{j-b} = \bigcap_j \bigcap_{b \in B} X_{j-b} = \bigcap_j (X_j \ominus B) \quad (4.17)$$

■

La relation suivante est régulièrement utilisée pour réduire le temps de calcul d'une érosion par un grand élément structurant. C'est la *relation de chaîne*, appelée *chain rule*.

Propriété 20

$$X \ominus (B \oplus C) = (X \ominus B) \ominus C \quad (4.18)$$

Démonstration.

$$(X \ominus B) \ominus C = \bigcap_{c \in C} \left(\bigcap_{b \in B} X_{-b}\right)_{-c} = \bigcap_{c \in C} \left(\bigcap_{b \in B} X_{-(b+c)}\right) = \bigcap_{c \in C} \bigcap_{b \in B} X_{-(b+c)} \quad (4.19)$$

$$= X \ominus (B \oplus C) \quad (4.20)$$

■

Cette relation a son importance car elle permet d'effectuer des érosions d'éléments structurants de grande taille et décomposables en une dilatation de deux plus petits éléments –par exemple, un carré est équivalent à une dilatation d'un segment horizontal par un segment vertical– par deux érosions successives de taille réduite.

Propriété 21

$$X \oplus B = (X \oplus B) \circ B = (X \bullet B) \oplus B \tag{4.21}$$

Propriété 22 *L'ouverture et la fermeture ne dépendent pas de l'origine de l'élément structurant, comme le suggérait l'interprétation de ces opérations. Soit $z \in \mathcal{E}$*

$$X \circ B_z = X \circ B \tag{4.22}$$

$$X \bullet B_z = X \bullet B \tag{4.23}$$

Un problème pratique: la gestion du bord dans les réalisations

Si l'invariance en translation découle des définitions des opérations morphologiques, la pratique contredit sa véracité parce que le bord de champ limite la connaissance de l'objet à un certain domaine. Une translation de l'image entraîne une perte irréversible d'information. Deux attitudes, correspondant à des hypothèses physiques sur l'extérieur du domaine d'observation et illustrées à la figure 4.8, sont possibles: soit l'objet s'étend au-delà du masque (hypothèse 1), soit on suppose que, physiquement, il n'y a rien en dehors du masque (hypothèse 2).

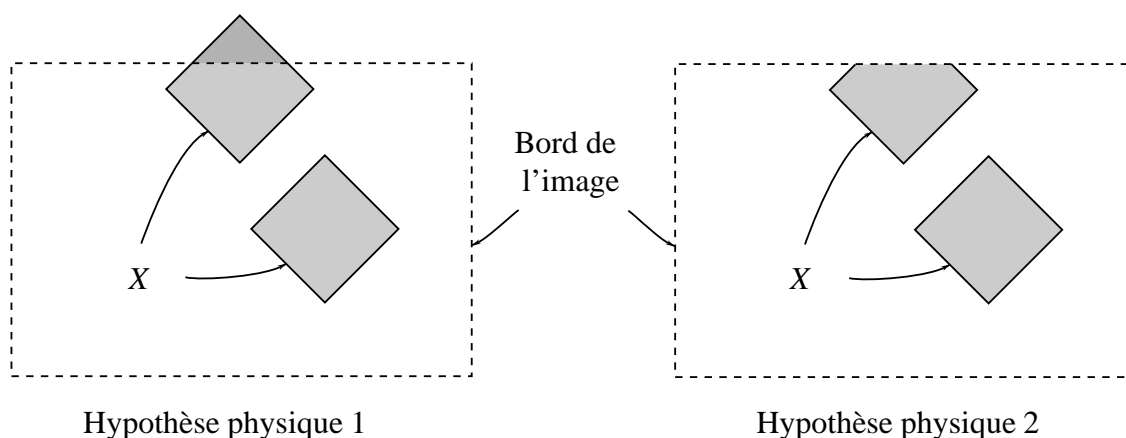


Figure 4.8: Deux hypothèses physiques pour traiter les objets qui touchent le bord de l'image.

Dans la seconde situation, les traitements morphologiques ne posent pas de problème car on dispose de toute l'information nécessaire. Il en va autrement pour la première hypothèse. Soit à effectuer, par exemple, une érosion. On a beau admettre que l'objet s'étende en dehors de l'image, cela ne signifie pas pour autant que l'on puisse délimiter cette zone d'extension. Le pis-aller habituel consiste à ajouter à X tous les points situés à l'extérieur de l'image lors du calcul de l'érodé. Cela modifie le résultat de l'érosion, comme le montrent les érosions par un disque centré, représentées à la figure 4.9. Par dualité, le calcul de la dilatation s'effectue en supposant que les points extérieurs n'appartiennent pas à l'ensemble.

Contrairement aux apparences, ceci n'est pas suffisant pour traiter correctement une ouverture dans le cas de l'hypothèse 1. Prenons l'exemple de l'ouverture de X par le disque B décentré, dessiné à la figure 4.10. Malgré la présence d'un bord lors du calcul de l'érosion, le rectangle du haut ne

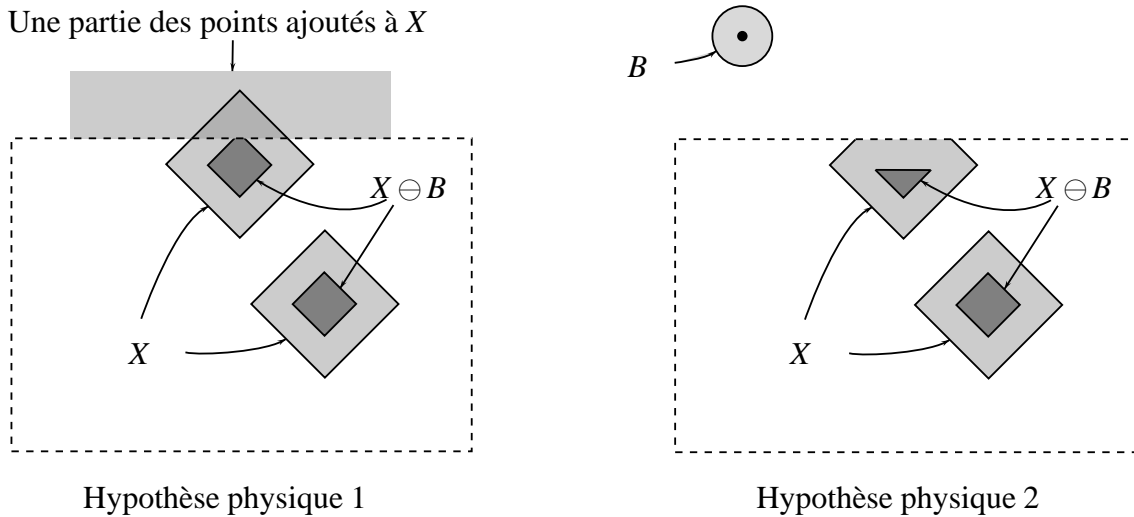


Figure 4.9: Comparaison de l'érosion de X , composé de deux losanges, pour les deux hypothèses physiques.

créé aucune contribution dans $X \ominus B$, et donc dans l'ouvert, alors qu'une même surface située à droite subsistera en partie dans l'ouvert. De plus, déplacer l'origine de B vers l'extrémité droite du disque inverse le résultat. Il n'est pourtant pas logique que le résultat dépende à ce point de la position de l'origine, surtout après avoir énoncé la propriété d'indépendance de cette position dans le résultat de l'ouverture.

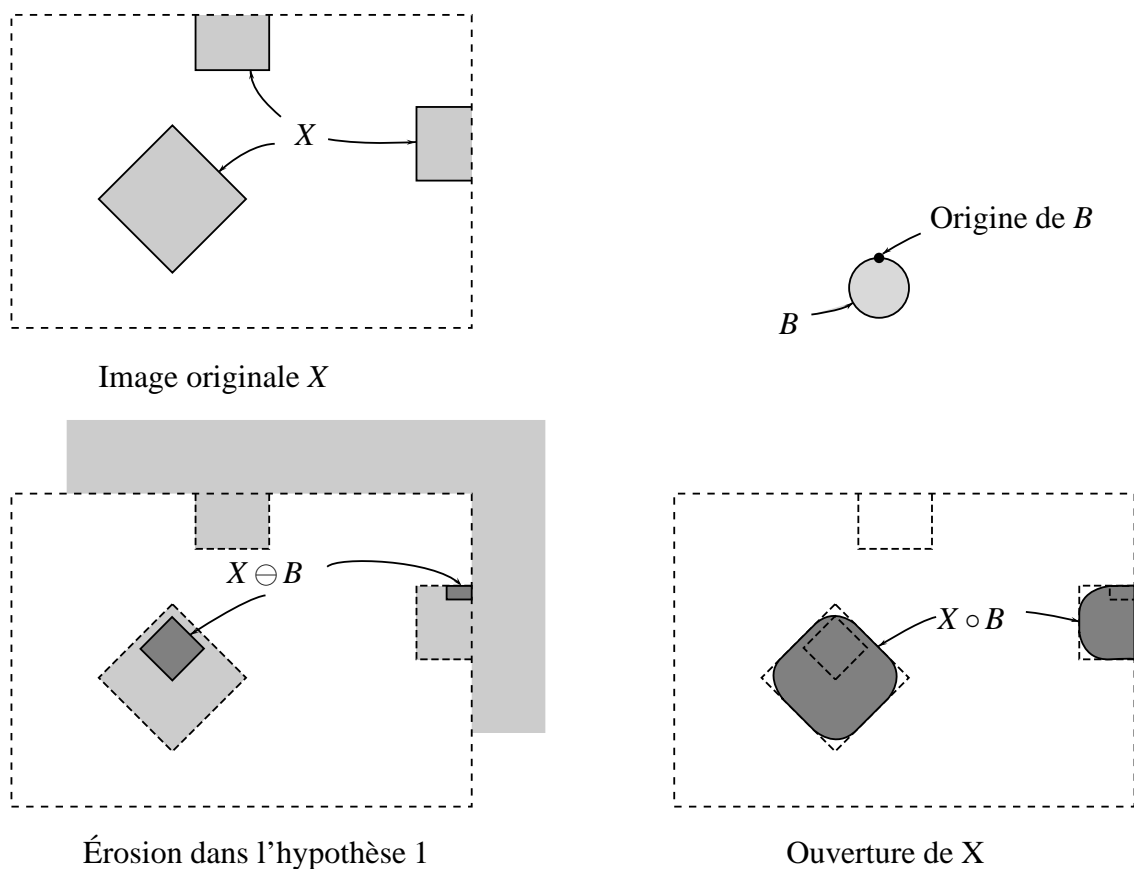


Figure 4.10: Pour illustrer que l'ouverture n'est pas unique malgré l'ajout d'un bord lors de l'érosion.

Soit M le masque d'observation, celui fixé par le bord de l'image. Comme remède, nous proposons une "définition pratique" de l'ouverture, à savoir

$$X \circ B = \bigcup_{p \in \mathbb{R}^2} ((X \cup M^c) \ominus B_p) \oplus B_p \cap M \quad (4.24)$$

Pour la plupart des éléments structurants, quelques positions p suffiront à obtenir la totalité de l'ouvert. Dans le pire des cas, il faudra les prendre en nombre égal au périmètre de B .

4.4 Les transformations de voisinage

Considérons un ensemble discret X . Le voisinage d'un point ne sera rien d'autre que le carré composé de 9 pixels, centré sur ce point. Si l'on attribue la valeur 1 ou 0 à chacun des points constituant le carré élémentaire, on forme une *configuration de voisinage*. Chacune de ces configurations constitue un élément structurant v^i , associé au réseau considéré. Dans le cas du réseau carré, il existe 2^9 configurations possibles. Dans la plupart des algorithmes utilisant des configurations de voisinage, elles sont groupées en famille $v = \{v^1, v^2, \dots, v^n\}$.

Appelant v_x la configuration de voisinage centrée en x , on définit une *transformation de voisinage* par la transformation qui attribue la valeur 1 (c'est la construction de la fonction *indicatrice*³) à chaque point ayant une configuration semblable à un élément de la famille de configurations v :

$$X \nabla v(x) = \begin{cases} 1 & \text{si } \exists v^i \in v \text{ tel que } v_x = v^i \\ 0 & \text{sinon} \end{cases} \quad (4.25)$$

Cette transformation satisfait les principes des transformations morphologiques énoncés précédemment.

L'érosion par un élément structurant n'est qu'un cas particulier de transformation de voisinage; elle revient à une transformation de voisinage où $v = \{v^i | B \subseteq v^i\}$.

Il existe aussi une transformation de voisinage appelée *transformée par tout ou rien* (*Hit or Miss transform* en anglais) telle que

$$X \uparrow (B, C) = \{x | B_x \subseteq X, C_x \subseteq X^c\} \quad (4.26)$$

Si $C = \emptyset$ la transformation dégénère en une opération d'érosion de X par B .

Les transformations de voisinage sont donc plus générales que les transformations morphologiques décrites précédemment, mais compte tenu de leur généralité, on ne peut leur associer aucune propriété fondamentale; elles dépendent de la configuration choisie.

4.4.1 Amincissement et épaissement

Les transformations de voisinage permettent d'introduire deux nouvelles catégories de transformation: l'amincissement et l'épaississement.

L'*amincissement* d'un ensemble X consiste à lui enlever des points qui correspondent à une configuration donnée, tandis que, par dualité, l'*épaississement* d'un ensemble revient à lui ajouter des points correspondant à une configuration donnée. Si l'on note par \ominus l'amincissement et par \odot l'épaississement, on aura:

$$X \odot v = X - (X \ominus v) \quad (4.27)$$

³La fonction *indicatrice* prend la valeur 1 si le point appartient à l'ensemble et 0 dans le cas contraire. Ce n'est, somme toute, qu'un autre moyen de définir un ensemble en extension.

$$X \odot v = X \cup (X \nabla v) \quad (4.28)$$

Les deux transformations que nous venons de définir possèdent un certain nombre de propriétés que nous allons brièvement énumérer.

Propriété 23 *Comme l'érosion et la dilatation, l'amincissement et l'épaississement sont deux transformations duales vis-à-vis de la complémentation*

$$(X \circ v)^c = X^c \odot v^c \quad (4.29)$$

Propriété 24 *Double inclusion*

$$X \circ v \subseteq X \subseteq X \odot v \quad (4.30)$$

4.5 Géodésie et reconstruction

Géographiquement, pour se rendre d'un endroit à un autre, il y a généralement plusieurs chemins possibles même si personne ne contestera que le chemin le plus court est la droite qui relie les deux points. De même, en imagerie, la *distance euclidienne* (c'est-à-dire $\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$) n'est pas toujours d'utilité pratique; c'est le cas sur une grille d'échantillonnage carrée où un chemin entre deux points doit emprunter un passage via des points intermédiaires. Sur une trame, le plus court chemin est appelé *chemin géodésique*; il peut y avoir plusieurs chemins géodésiques. Le principe de la géodésie est à la base de nombreuses techniques de traitement morphologique.

4.5.1 Dilatation géodésique

Une dilatation géodésique fait toujours intervenir deux images. La première image est le dilaté par un élément structurant élémentaire isotropique adapté à la trame; il s'agit d'un carré de trois pixels de large pour une trame carrée et d'un hexagone dans une trame hexagonale. Quant à la seconde image, elle limite l'extension de la dilatation de la première.

Définition 25 *La dilatation géodésique de taille 1 de l'ensemble X conditionnellement à Y , notée $D_Y^{(1)}(X)$, est définie comme l'intersection du dilaté de taille 1 et de Y :*

$$\forall X \subseteq Y, D_Y^{(1)}(X) = (X \oplus B) \cap Y \quad (4.31)$$

où B est l'élément le plus simple adapté à la trame.

Suite à l'apparition de l'intersection dans la définition, $D_Y^{(1)}(X)$ est toujours inclus ou égal à Y . On dit aussi que Y sert de *masque géodésique*. La figure 4.11 illustre le principe de la dilatation géodésique de taille 1 sur un ensemble dans une trame hexagonale.

Définition 26 *La dilatation géodésique de taille n de l'ensemble X conditionnellement à Y , notée $D_Y^{(n)}(X)$, est définie comme une succession du dilaté géodésique de taille 1*

$$\forall X \subseteq Y, D_Y^{(n)}(X) = \underbrace{D_Y^{(1)}(D_Y^{(1)}(\dots D_Y^{(1)}(X)))}_{n \text{ fois}} \quad (4.32)$$

où B est l'élément le plus simple adapté à la trame.

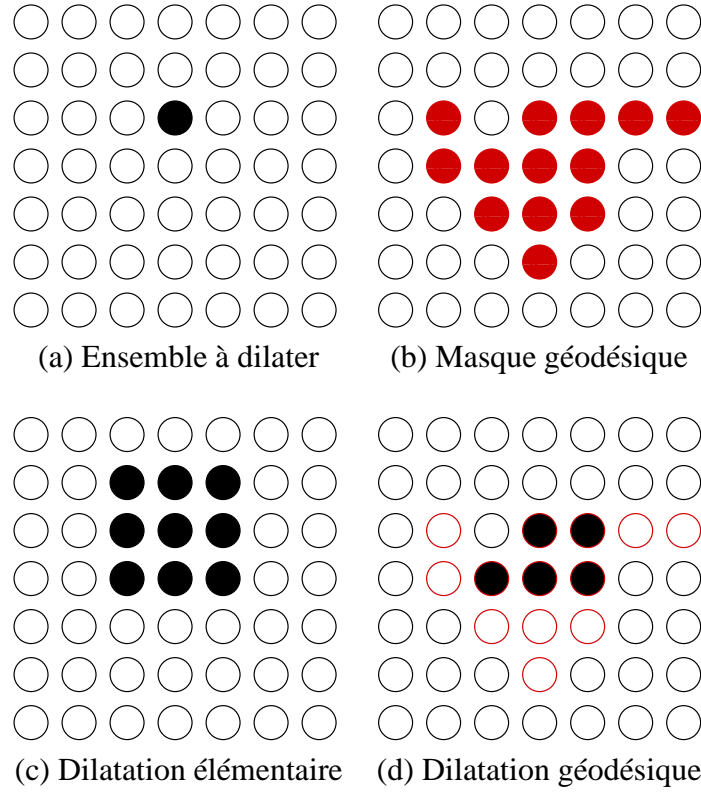


Figure 4.11: Dilatation géodésique de taille 1 d'un ensemble.

Il est important de remarquer que $D_Y^{(n)}(X)$ n'est pas toujours égal à l'ensemble obtenu en intersectant X dilaté par l'élément structurant B de taille n avec Y . Prenons le cas d'un ensemble Y formé de plusieurs composantes connexes et de X réduit à une composante de Y . Suite à la définition de la dilatation géodésique, $D_Y^{(n)}(X) = X$ alors que le résultat de l'intersection après plusieurs dilations pourra contenir des parties des autres composantes connexes de Y .

4.5.2 Érosion géodésique

La dualité inhérente aux opérations morphologiques élémentaires conduit à définir l'érosion géodésique de taille 1 et n .

Définition 27 L'érosion géodésique de taille n de l'ensemble X conditionnellement à Y , notée $(X \ominus B)^{(n)}$, est définie comme une succession de l'érodé géodésique de taille 1 et de Y

$$(X \ominus B)^{(n)} = \underbrace{(((X \ominus B)^{(1)} \ominus B)^{(1)} \dots)}_{n \text{ fois}} \quad (4.33)$$

où l'érodé géodésique de taille 1 est égal à

$$\forall X \supseteq Y, (X \ominus B)^{(1)} = (X \ominus B) \cup Y \quad (4.34)$$

Remarquons que cette fois, c'est l'union qui agit sur les deux ensembles. Au même endroit, une intersection aura pour conséquence de réduire l'ensemble jusqu'à en faire l'ensemble vide.

4.5.3 Reconstruction morphologique

Les opérations géodésiques sont rarement utilisées en tant que telles. C'est plutôt l'ensemble vers lequel convergent ces opérations lorsque n croît qui permet de résoudre certains problèmes pratiques. En fait, l'érosion et la dilatation géodésiques convergent après un certain nombre d'itérations. Le procédé de reconstruction morphologique tire profit de ce principe:

Définition 28 La reconstruction de X conditionnellement à Y est la dilatation géodésique de X jusqu'à idempotence. Soit i , la valeur à partir de laquelle l'idempotence est acquise, la reconstruction de X est définie par

$$R_Y(X) = D_Y^{(i)}(X) \text{ avec } D_Y^{(i+1)}(X) = D_Y^{(i)}(X) \quad (4.35)$$

Comme l'indique son nom, cette opération permet de reconstruire des détails perdus éventuellement lors d'une autre opération. La figure 4.12 représente les trois étapes mises en oeuvre pour extraire des particules choisies dans une image. L'image (a) est le signal de départ. Dans l'image (b), on sélectionne certaines particules en insérant quelques points dans celles-ci. La reconstruction conduit alors à l'image (c) où ne sont reconstituées que les particules sélectionnées; les autres ont été gommées.

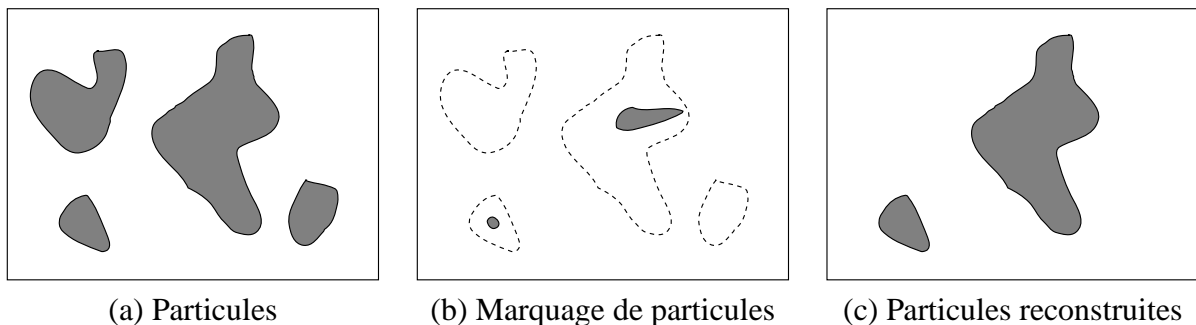


Figure 4.12: Extraction de particules pré-sélectionnées dans une image.

L'itération de la dilatation dans la phase de reconstruction pousse à penser que le temps de calcul de la reconstruction est élevé. En fait, il existe des algorithmes rapides et simples de réaliser la reconstruction; ils ont même été implantés dans un processeur spécialisé de traitement d'images par morphologie mathématique.

4.6 Opérations morphologiques sur des images en niveaux de gris

4.6.1 Notion de fonction

L'extension des outils de la morphologie mathématique à l'étude des images en niveaux de gris s'est réalisée longtemps après la définition des concepts de la théorie. MEYER et STERNBERG [35] furent les véritables initiateurs de techniques adaptées à des images en niveaux de gris.

Soit \mathcal{G} un ensemble de valeurs de niveaux de gris. Dans ce qui suit, nous supposons que \mathcal{G} est un espace complété ($\mathcal{G} = \mathbb{R} \cup \{-\infty, +\infty\}$ ou $\mathcal{G} = \mathbb{Z} \cup \{-\infty, +\infty\}$). Une image est représentée par une fonction $f : \mathcal{E} \rightarrow \mathcal{G}$, qui à tout point du plan fait correspondre une valeur. En pratique, l'image n'est pas définie sur la totalité du référentiel \mathcal{E} mais sur un support compact D qui constitue la partie visible de l'image. Néanmoins, à condition de bien comprendre les effets de bord, rien ne nous empêche de considérer un support infini par facilité.

Pour la morphologie mathématique en niveaux de gris, les opérations de base ne sont plus l'union et l'intersection mais le *supremum* \vee et l'*infimum* \wedge . Pour aborder le cadre algébrique, il faut d'abord définir la notion d'ordre entre fonctions, tout comme nous l'avons fait pour les ensembles en utilisant la notion d'inclusion.

Définition 29 [Ordre partiel entre fonctions] Soient deux fonctions f et g . La fonction f est inférieure à g ,

$$f \leq g \text{ si } f(x) \leq g(x), \forall x \in \mathcal{E} \quad (4.36)$$

La relation “est plus grand”, notée \geq , est définie de la même manière. L'ordre partiel permet de définir les bornes d'une famille d'opérateurs.

Insistons bien sur le fait que a fonction f est inférieure (\leq) à la fonction g si, en tout point x de \mathcal{E} , on a l'inégalité de valeur $f(x) \leq g(x)$. Comme toutes les valeurs d'une fonction ne sont pas nécessairement toutes inférieures ou supérieures à une autre fonction, il n'est pas toujours possible de déterminer l'ordre entre des fonctions; c'est la raison pour laquelle il s'agit d'un ordre partiel.

Définition 30 [Infimum et supremum] Soit une famille de fonctions $f_i, i \in I$. L'infimum (respectivement le supremum) de cette famille, noté $\bigwedge_{i \in I} f_i$ (resp. $\bigvee_{i \in I} f_i$) est la grande borne inférieure (resp. petite borne supérieure).

Dans le cas d'une famille I finie, le supremum et l'infimum se ramènent au maximum et au minimum respectivement. Dans ce cas,

$$\forall x \in \mathcal{E}, \begin{cases} (f \vee g)(x) = \max(f(x), g(x)) \\ (f \wedge g)(x) = \min(f(x), g(x)) \end{cases} \quad (4.37)$$

Nous conserverons l'expression formelle générale parce qu'elle est à la fois plus répandue dans la littérature scientifique et plus commode pour l'étude de propriétés délicates comme la convergence de certains opérateurs.

Tout comme pour des ensembles, il nous reste à définir le translaté d'une fonction.

Définition 31 La translation de la fonction f par b est la fonction f_b définie par

$$\forall x \in \mathcal{E}, f_b(x) = f(x - b) \quad (4.38)$$

Cette définition de translation agit sur le support d'une fonction. Nous aurions tout aussi bien pu définir une translation des valeurs de la fonction, mais ce concept ne nous servira pas directement.

4.6.2 Définition de propriétés des opérateurs de fonctions

Les opérateurs morphologiques de fonctions agissent comme des transformations d'ensembles. Dans ce paragraphe, nous étendons les concepts vus précédemment aux opérateurs de fonctions.

Définition 32 [Idempotence] Une transformation ψ est idempotente si, quelle que soit la fonction traitée, une nouvelle application de l'opérateur n'apporte aucun changement, c'est-à-dire si

$$\forall f, \psi(\psi(f)) = \psi(f) \quad (4.39)$$

Définition 33 [Extensivité] Un opérateur est extensif si la fonction transformée est plus grande que la fonction de départ

$$\forall f, f \leq \psi(f) \quad (4.40)$$

Définition 34 [Anti-extensivité] Une transformation de fonctions est anti-extensive si la fonction transformée est plus petite que la fonction de départ

$$\forall f, f \geq \psi(f) \quad (4.41)$$

La croissance est une autre propriété qui joue un rôle primordial dans les traitements morphologiques. Elle indique si un opérateur préserve la relation d'ordre.

Définition 35 [Croissance] Une transformation de fonctions est croissante si elle conserve l'ordre établi entre fonctions

$$\forall f, g, f \leq g \Rightarrow \psi(f) \leq \psi(g) \quad (4.42)$$

Par extension aussi, une transformation ψ_1 est inférieure à un transformation ψ_2 si, pour toute fonction f , $\psi_1(f)$ est inférieur à $\psi_2(f)$:

$$\psi_1 \leq \psi_2 \Leftrightarrow \forall f, \psi_1(f) \leq \psi_2(f) \quad (4.43)$$

4.6.3 Érosion, dilatation et combinaisons

Nous disposons à présent des éléments de théorie nécessaires à la définition des opérations morphologiques sur des images en niveaux de gris, étudiées ici sous l'angle des fonctions.

Définition 36 Si f et g sont deux fonctions, on peut définir l'addition (ou dilatation) $f \oplus g$ et la soustraction (ou érosion) $f \ominus g$ de la manière suivante: pour tout couple $(x, g(x))$, on prend pour $f \oplus g$ l'enveloppe supérieure des translatés du graphe de f par les $(x, g(x))$ et pour $f \ominus g$, l'enveloppe inférieure des translatés de f par les $(-x, -g(x))$. Algébriquement,

$$f \oplus g = \bigvee_{h \in \mathcal{E}} (f(x-h) + g(h)) \quad (4.44)$$

$$f \ominus g = \bigwedge_{h \in \mathcal{E}} (f(x+h) - g(h)) \quad (4.45)$$

(N.B. À nouveau, les travaux de SERRA et MATHERON couvrent une définition de $f \ominus g$ légèrement différente).

Il est d'usage de choisir une fonction structurante g prenant des valeurs finies sur un support compact (c'-à-d. borné et fermé) D et valant $-\infty$ ailleurs. La littérature fournit classiquement une interprétation géométrique de ces opérations en recourant à la notion d'ombre. L'ombre d'une fonction f est le sous-ensemble $U(f)$ de l'espace $\mathcal{E} \times \mathcal{G}$ tel que

$$U(f) = \{(x, t) \in \mathcal{E} \times \mathcal{G} | t \leq f(x)\} \quad (4.46)$$

En termes d'ombres, $f \oplus g$ et $f \ominus g$ sont les enveloppes supérieures respectives de $U(f) \oplus U(g)$ et de $U(f) \ominus U(g)$ où \oplus, \ominus sont les opérations ensemblistes d'addition et de soustraction originales.

En pratique

Telles quelles, les définitions ne se prêtent guère à une réalisation par logiciel ou par matériel; les algorithmes tirés des définitions sont de complexité proportionnelle au cube de la taille de la fonction structurante.

Dans la pratique, on utilise non pas une fonction g comme fonction structurante mais une fonction dégénérée binaire valant 0 à l'intérieur d'un petit masque et $-\infty$ ailleurs, auquel cas g se ramène à un ensemble B . Soit B , la partie du domaine de g où le signal est nul (ailleurs, il vaut $-\infty$), les définitions deviennent:

Définition 37

$$f \oplus B = \bigvee_{b \in B} f_b(x) \tag{4.47}$$

$$f \ominus B = \bigwedge_{b \in B} f_{-b}(x) \tag{4.48}$$

L'opération d'érosion est illustrée à la figure 4.13. Par commodité, nous assimilerons D à B sans faire

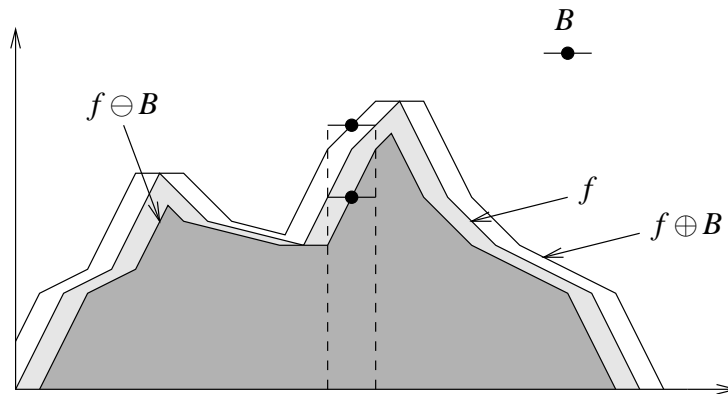


Figure 4.13: Érosion d'une fonction.

la distinction entre le support non infini négatif ou le domaine de définition.

Avec des éléments structurants plans, une dilataion se calcule comme l'enveloppe supérieure de tous les translatés de la fonction par les éléments de l'élément structurant. Soit à calculer par exemple le dilaté d'une fonction par un élément structurant $B = \{-1, 0, 1\}$. L'enveloppe supérieure se détermine à partir des trois fonctions $f_{-1}(x)$, $f(x)$ et $f_1(x)$, ce qui revient à chercher le maximum entre ces trois fonctions pour chaque point. C'est donc une traduction de la notion d'enveloppe supérieure sur un jeu de translatés par la recherche du maximum sur le voisinage du point en question. Quant à l'érosion, elle devient équivalente à la recherche du minimum, opération simple à réaliser.

4.6.4 Ouverture et fermeture morphologiques

L'ouverture morphologique $f \circ B$ et la fermeture morphologique $f \bullet B$ résultent de la mise en cascade de l'érosion et de la dilataion .

Définition 38 [Ouverture et fermeture morphologiques]

$$f \circ B = (f \ominus B) \oplus B \tag{4.49}$$

$$f \bullet B = (f \oplus B) \ominus B \tag{4.50}$$

Les figures 4.16 et 4.17 montrent le résultat d'une ouverture et d'une fermeture sur une fonction unidimensionnelle.

Par analogie avec le traitement d'ensembles, l'ouverture sur des images produit un effet de filtrage comme le montre la figure 4.20 où sont rassemblées plusieurs images dont l'image érodée, l'image dilatée et l'image ouverte, obtenues toutes avec un élément structurant carré. Puisqu'elle recherche un minimum, l'érosion assombrit l'image alors qu'à l'inverse, la dilataion éclaircit l'image. L'ouverture produit une image qui ne conserve que les parties claires ayant la taille et la forme de l'élément structurant; les autres sont proprement gommées, c'est-à-dire noircies.

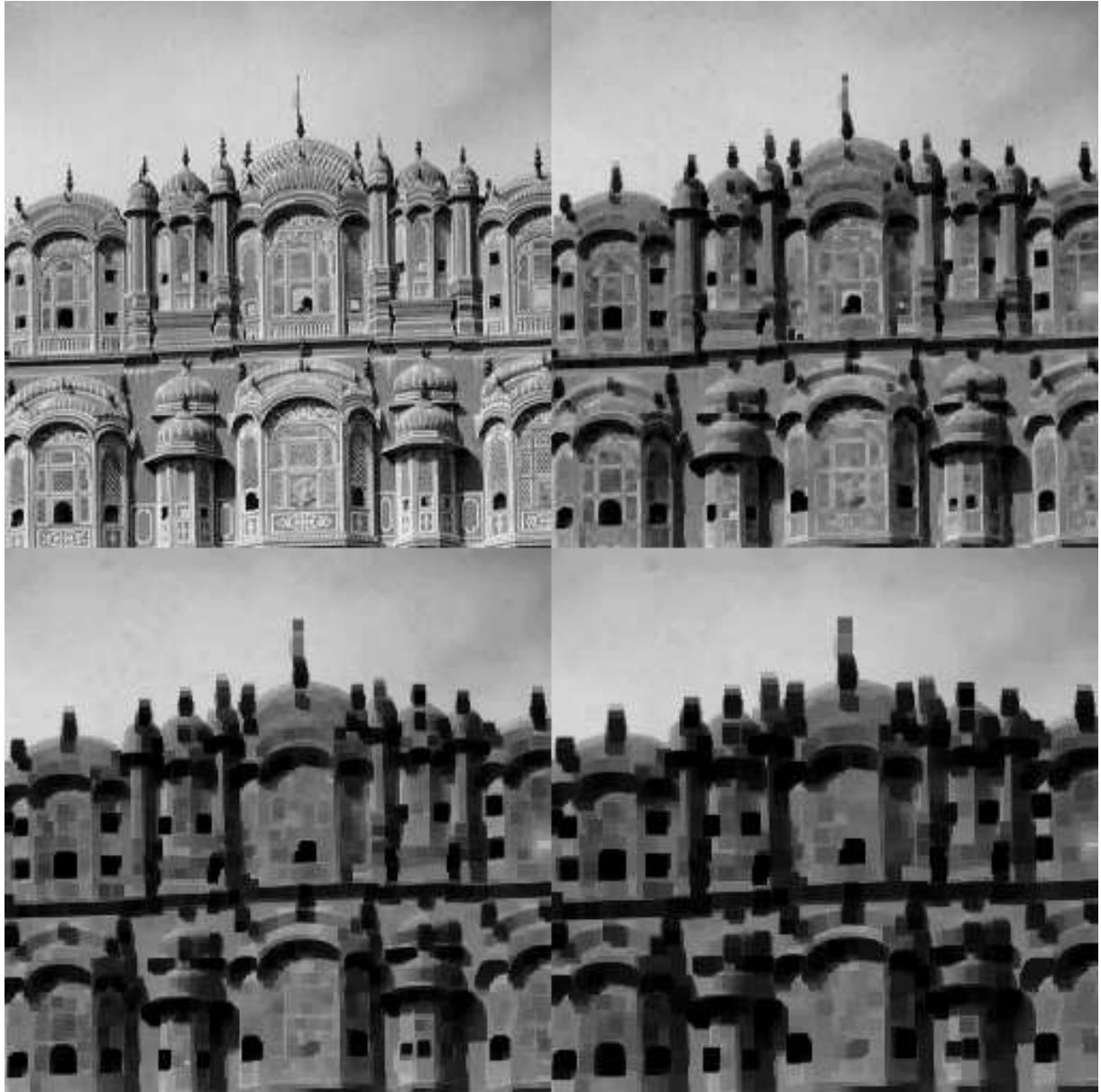


Figure 4.14: Érosions successives par des carrés de taille croissante.

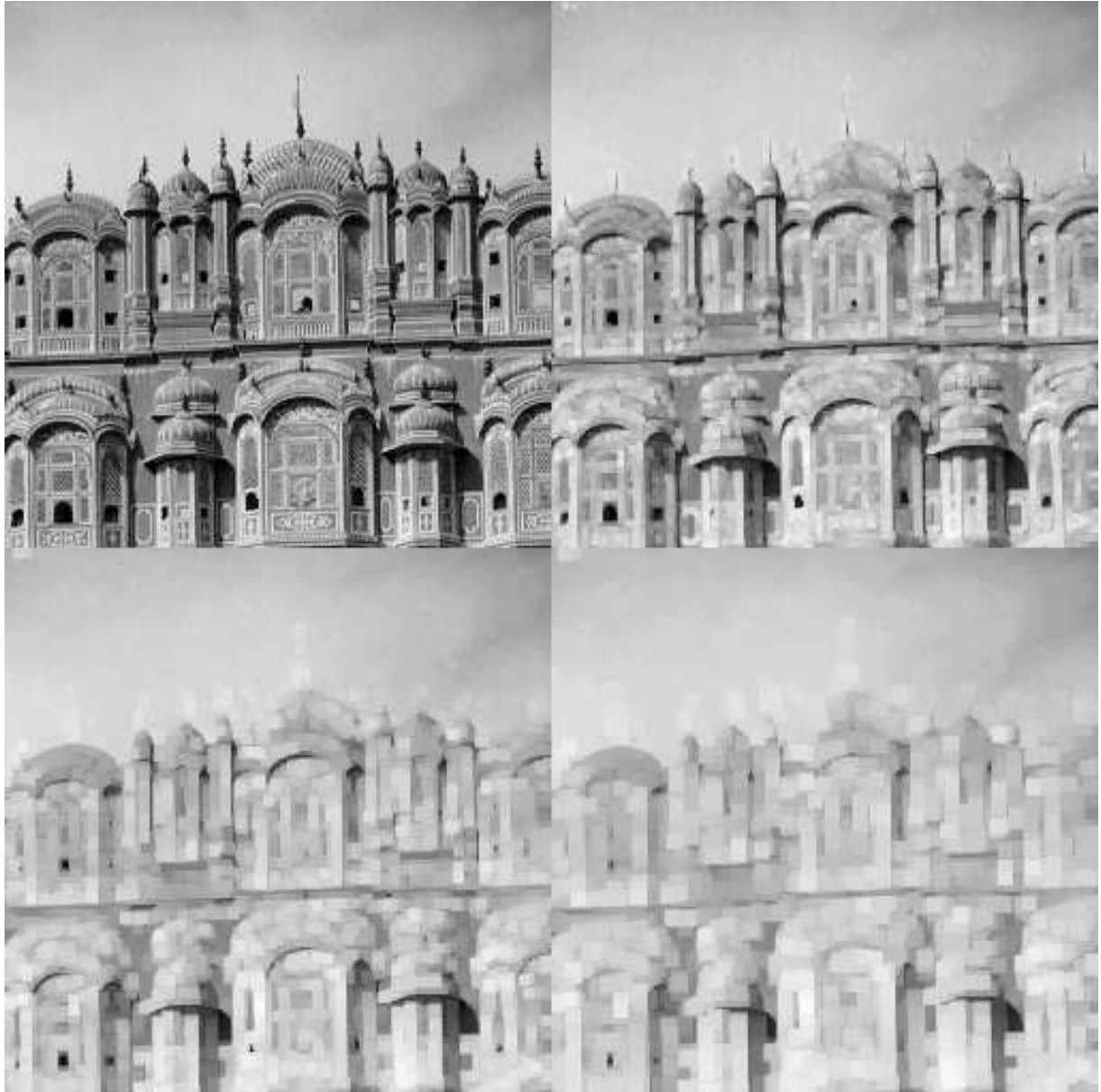


Figure 4.15: Dilatations successives par des carrés de taille croissante.

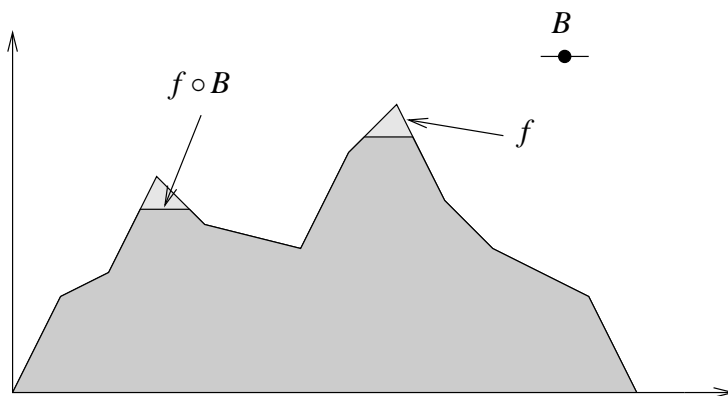


Figure 4.16: Ouverture d'une fonction.

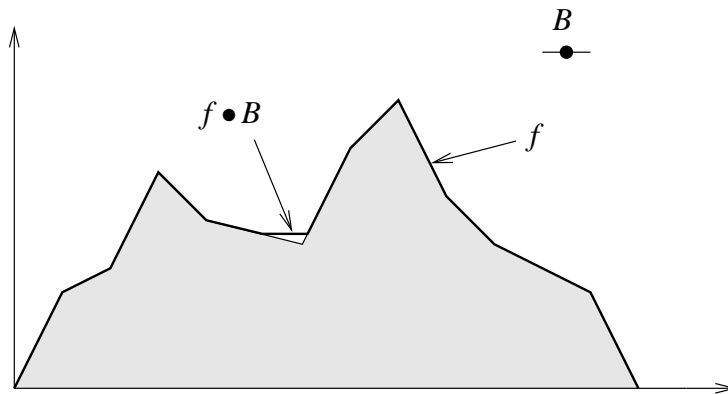


Figure 4.17: Fermeture d'une fonction.



Figure 4.18: Ouvertures successives par des carrés de taille croissante.

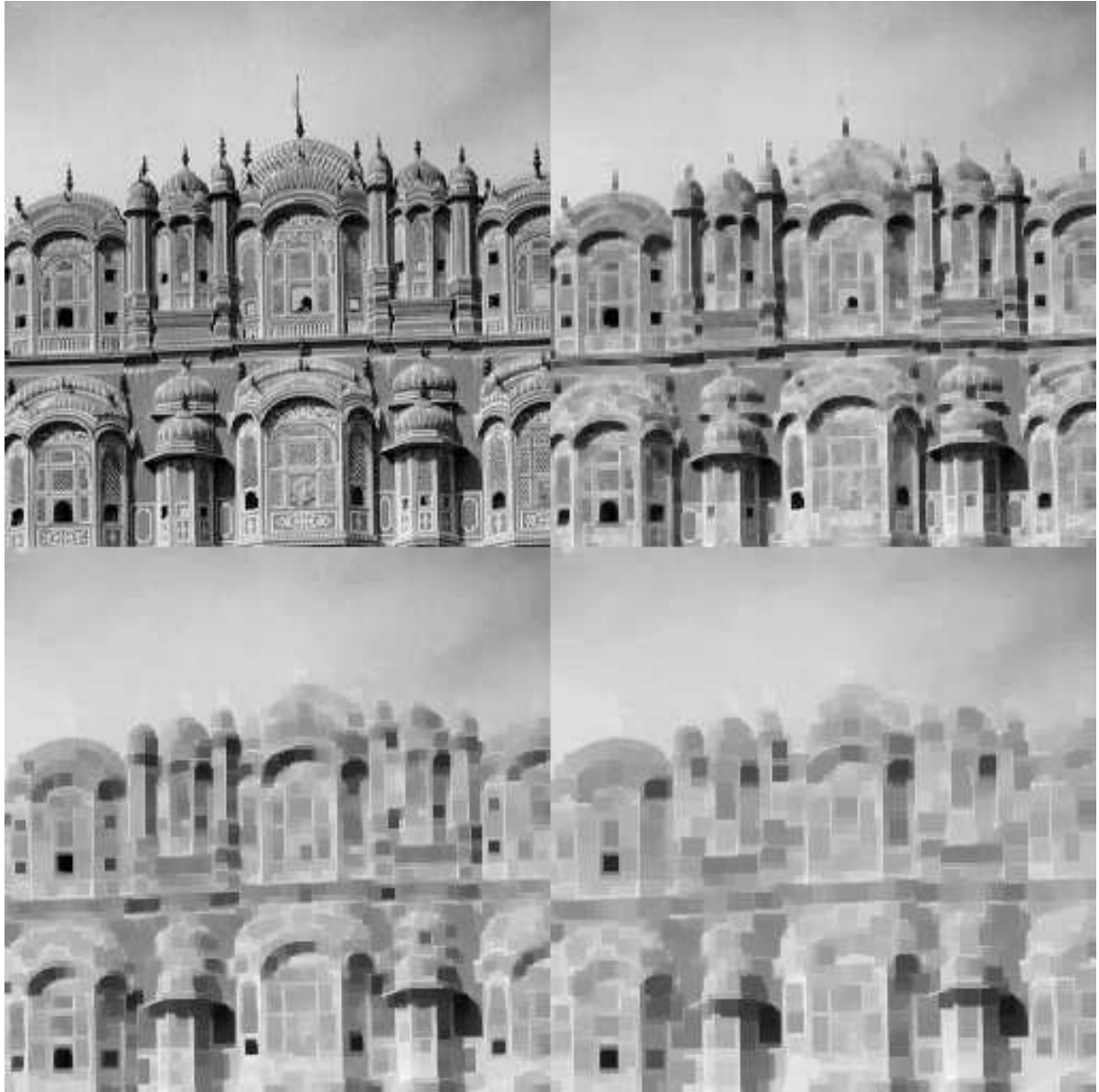


Figure 4.19: Fermetures successives par des carrés de taille croissante.



(a) Image originale f



(b) Érosion par un carré



(c) Dilatation par un carré



(d) Ouverture par un carré

Figure 4.20: Illustration des opérations élémentaires sur une image en niveaux de gris.

4.6.5 Propriétés des opérateurs morphologiques de fonctions

Toutes les propriétés relatives à des opérations morphologiques d'ensembles restent valables pour des fonctions à condition de remplacer les ensembles par des fonctions. Rappelons sous leur nouvelle forme quelques propriétés essentielles.

Propriété 39 *L'érosion et la dilatation sont des opérations croissantes*

$$f \leq g \Rightarrow \begin{cases} f \ominus B \leq g \ominus B \\ f \oplus B \leq g \oplus B \end{cases} \quad (4.51)$$

Propriété 40 *L'érosion distribue l'infimum et la dilatation distribue le supremum*

$$(f \wedge g) \ominus B = (f \ominus B) \wedge (g \ominus B) \quad (4.52)$$

$$(f \vee g) \oplus B = (f \oplus B) \vee (g \oplus B) \quad (4.53)$$

Propriété 41 *L'ouverture et la fermeture sont deux opérations idempotentes*

$$(f \circ B) \circ B = f \circ B \quad (4.54)$$

$$(f \bullet B) \bullet B = f \bullet B \quad (4.55)$$

Propriété 42 *L'ouverture et la fermeture sont deux transformations respectivement anti-extensive et extensive*

$$f \circ B \leq f \quad (4.56)$$

$$f \leq f \bullet B \quad (4.57)$$

4.6.6 Reconstruction des images à niveaux de gris

Tout comme en binaire, on peut définir la reconstruction morphologique en niveaux de gris.

Définition 43 *La reconstruction de la fonction f conditionnellement à la fonction g est la dilatation géodésique de f jusqu'à idempotence. Soit i , la valeur à partir de laquelle l'idempotence est acquise, la reconstruction de f est définie par*

$$R_g(f) = D_g^{(i)}(f) \text{ avec } D_g^{(i+1)}(f) = D_g^{(i)}(f) \quad (4.58)$$

La figure 4.21 montre une image originale, une image à reconstruire -il s'agit dans ce cas de l'image érodée- ainsi que des étapes de reconstruction.

Les figures 4.22 et 4.23 illustrent respectivement la reconstruction et la reconstruction duale qui s'obtient par itération de l'érosion géodésique.

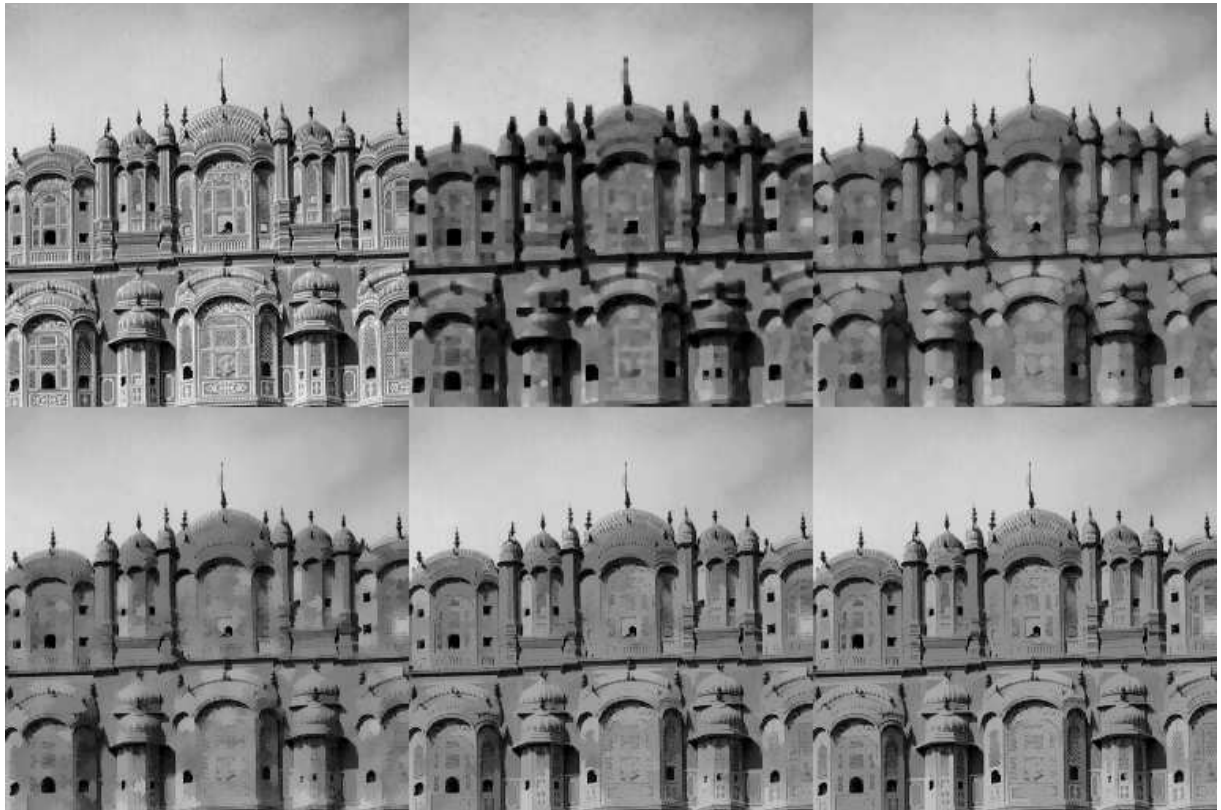


Figure 4.21: Image originale, image érodée et dilations géodésiques successives.

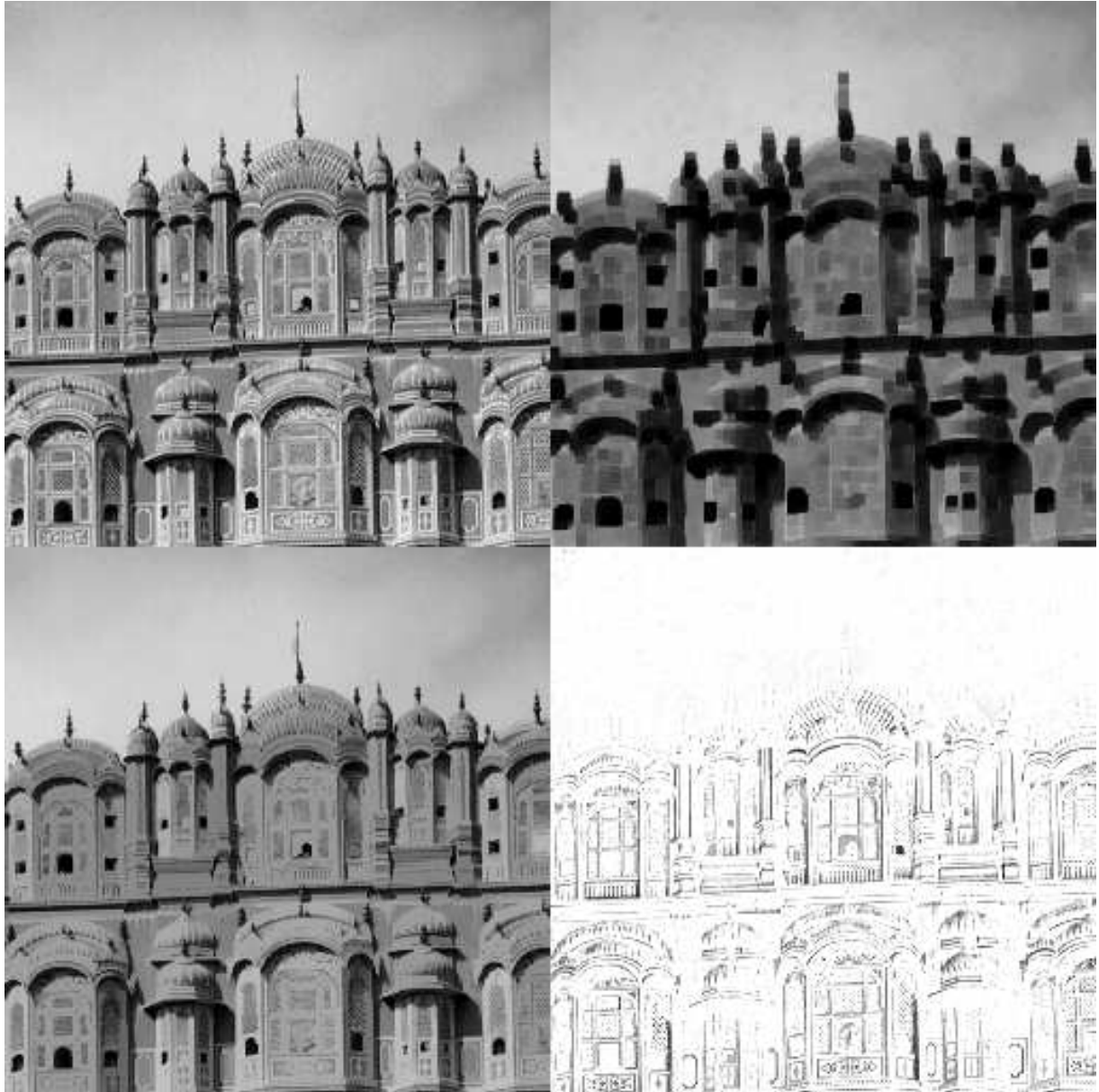


Figure 4.22: Image originale, image érodée, image reconstruite à partir de l'érodé et image différence (en vidéo inverse).

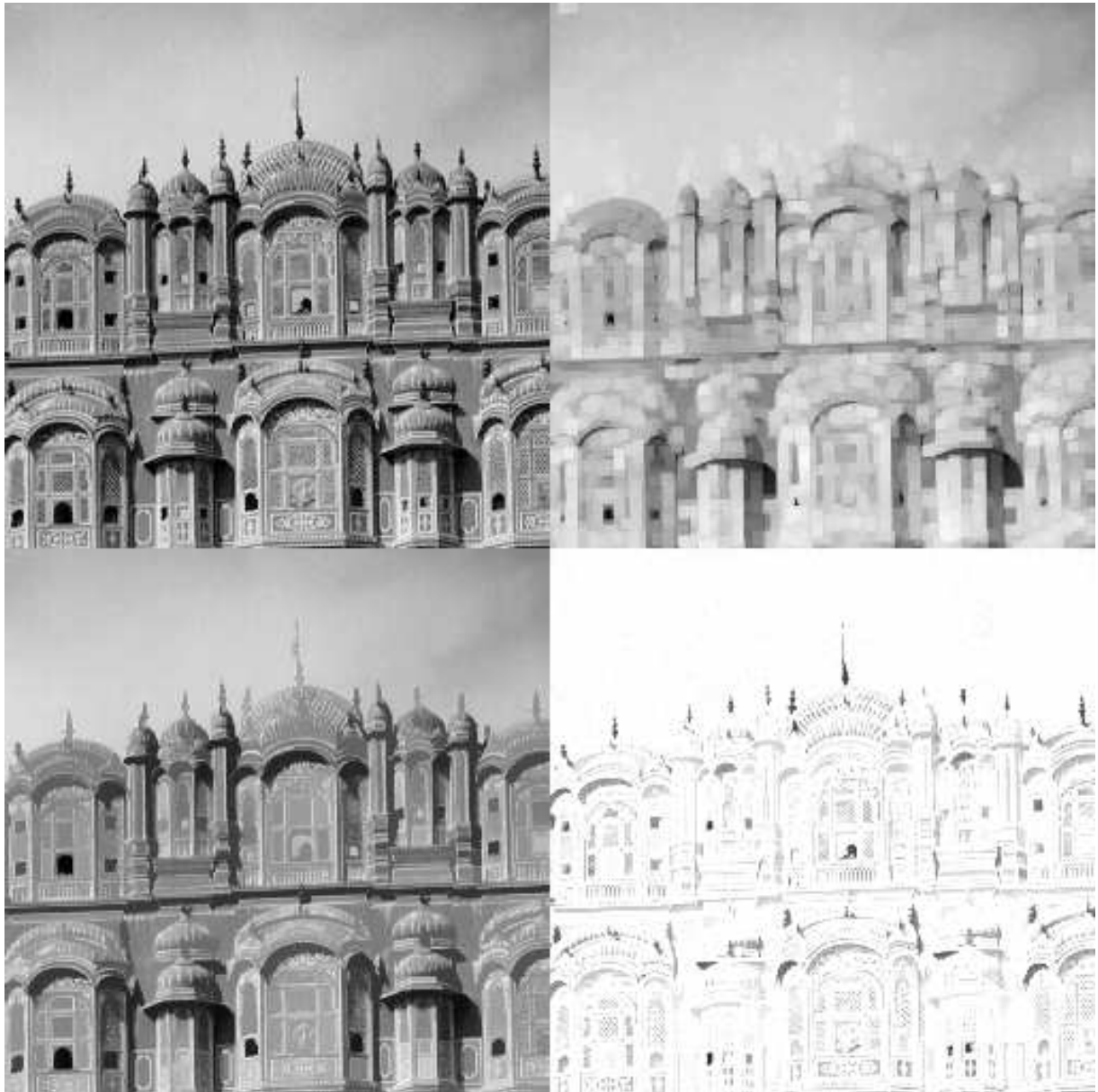


Figure 4.23: Image originale, image dilatée, image reconstruite à partir du dilaté (par reconstruction duale) et image différence (en vidéo inverse).

Chapter 5

Filtrage non-linéaire

5.1 Généralités

Le filtrage est une discipline à part entière du traitement d'images destinée à supprimer le bruit qui affecte l'image, à accroître les contrastes ou encore à favoriser une caractérisation de textures, ... Dans cette section, nous décrirons quelques fondements du filtrage non-linéaire ainsi que les familles de filtres les plus utilisés.

5.2 Filtres non-linéaires

5.2.1 Filtres de rang

L'érosion ou la dilatation correspondent à la sélection des valeurs extrêmes puisqu'il s'agit de déterminer l'infimum ou le supremum. Ce type d'opérations est donc très sensible au bruit, particulièrement à un bruit impulsionnel. On songe donc tout naturellement à considérer d'autres valeurs que ces valeurs extrêmes dans l'espoir de diminuer la sensibilité au bruit; c'est la notion de *filtre de rang*¹ qui s'en dégage.

Prenons un élément structurant B fini servant de domaine de définition. Soit $k \in \mathbb{N}$ la valeur d'un seuil.

Définition 44 [Filtre de rang] L'opérateur ou filtre de rang k , noté $\rho_{B,k}(f)(x)$, défini sur l'élément structurant B , est l'opérateur

$$\rho_{B,k}(f)(x) = \bigvee \{t \in \mathcal{G} \mid \sum_{b \in B} [f(x+b) \geq t] \geq k\} \quad (5.1)$$

L'interprétation la plus simple consiste à dire que $\rho_{B,k}(f)(x)$ est la k -ième valeur de la série obtenue en ordonnant toutes les valeurs $f(x+b)$ par ordre décroissant.

Les opérateurs de rang sont ordonnés entre eux. En effet, soit $\#(B)$ le cardinal de l'élément structurant B ,

$$\rho_{B,\#(B)}(f)(x) \leq \rho_{B,\#(B)-1}(f)(x) \leq \dots \leq \rho_{B,1}(f)(x) \quad (5.2)$$

¹On doit cette notion à RONSE.

Filtre médian

Si n est impair, le choix $k = \frac{1}{2}(\#(B) + 1)$ conduit à la définition d'un opérateur *auto-dual*, c'est-à-dire que l'opérateur appliqué à une image fournit le même résultat que l'opérateur appliqué à l'image complémentaire. Cet opérateur, noté med_B , est appelé *filtre médian*. On doit sa découverte à TUKEY.

Le filtrage médian est une technique de filtrage non-linéaire couramment utilisée en pratique. Il s'avère particulièrement efficace pour juguler les effets d'un bruit impulsionnel. Sa caractéristique essentielle est sa capacité à conserver des transitions fortes tout en supprimant une partie importante du bruit.

La figure 5.1 compare quelques opérations de filtrage. De même, la figure 5.2 montre l'effet d'un changement de la taille de la fenêtre sur la qualité de l'image filtrée.



(a) Image originale bruitée f



(b) Ouverture par un carré 5×5



(c) Passe-bas BUTTERWORTH ($f_c = 50$)



(d) Médian sur une fenêtre 5×5

Figure 5.1: Comparaison de différentes techniques de filtrage sur une image bruitée.

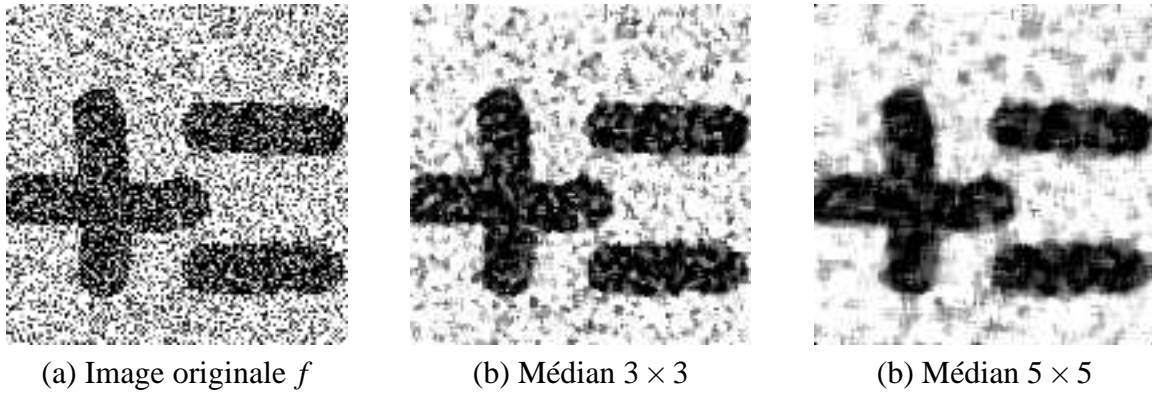


Figure 5.2: Effet de la taille de la fenêtre rectangulaire sur le filtrage.

Implémentation

Un moyen commode pour implémenter les filtres de rang consiste à utiliser un histogramme local glissant. Cet histogramme est mis à jour à chaque déplacement et une lecture de cet histogramme suffit à fournir le résultat.

Remarquons au passage que le filtre médian n'est pas idempotent et sa répétition peut entraîner des phénomènes d'oscillation (théoriquement sur une image de support infini). La figure 5.3 illustre le phénomène d'oscillations par applications successives de l'opérateur.

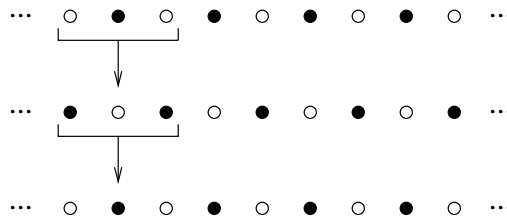


Figure 5.3: Applications répétées du filtre médian.

5.2.2 Filtres à masque tournant

SONKA *et al.* [34] détaillent l'usage d'autres filtres non-linéaires parmi lesquels on distingue les filtres à masque tournant. Ce type de filtres reposent sur l'utilisation d'un critère d'homogénéité et d'une famille d'éléments structurants B_j , appelés *masques*. Le critère d'homogénéité est par exemple une mesure de dispersion énergétique comme

$$\sigma_{B_j}^2 = \frac{1}{\#(B_j)} \sum_{b \in B_j} \left[f(x+b) - \frac{1}{\#(B_j)} \sum_{b' \in B_j} f(x+b') \right]^2 \quad (5.3)$$

$$= \frac{1}{\#(B_j)} \left[\sum_{b \in B_j} f^2(x+b) - \frac{1}{\#(B_j)} \left(\sum_{b' \in B_j} f(x+b') \right)^2 \right] \quad (5.4)$$

La dispersion est évaluée pour chaque masque. Le masque de plus petite dispersion sert ensuite de masque pour le calcul de la moyenne dont la valeur est affectée au pixel x considéré.

5.3 Filtres algébriques et filtres morphologiques

5.3.1 Introduction

Un filtre linéaire est constitué d'un opérateur, bien entendu linéaire et continu, mais aussi commutant avec les translations et ayant, dans certains cas, des propriétés d'idempotence (filtre passe-bande idéal, par exemple).

Dans une image, un objet placé au-devant de la scène occulte tous ceux qui se trouvent derrière lui dans l'axe d'observation. Pour exprimer cet aspect fondamental, on a nécessairement recours à la notion d'inclusion qui remplace, dans l'univers visuel, la structure d'additivité qui régit d'autres univers. Le filtrage morphologique constitue une alternative au filtrage linéaire. Rappelons que les filtres morphologiques ne possèdent aucune propriété de linéarité; supposons deux fonctions de gris $f(x)$ et $g(x)$ et une transformation morphologique θ , on aura généralement: $\theta(f(x) + g(x)) \neq \theta(f(x)) + \theta(g(x))$

Nous avons vu que les ouvertures morphologiques permettaient de lisser les objets d'une image binaire. D'où l'idée d'appliquer des filtrages sur des formes ou des images en niveaux de gris. Nous allons à présent enrichir la famille des filtres non-linéaires après avoir préalablement défini quelques propriétés algébriques.

Définition 45 *Par définition, un filtre algébrique est une opération croissante et idempotente:*

$$\psi \text{ est un filtre algébrique} \Leftrightarrow \forall f, g \begin{cases} f \leq g \Rightarrow \psi(f) \leq \psi(g) \\ \psi(\psi(f)) = \psi(f) \end{cases} \quad (5.5)$$

On peut dès lors définir une ouverture *algébrique*.

Définition 46 *Une ouverture algébrique est un opérateur qui possède les propriétés de croissance, d'idempotence et d'anti-extensivité. Formellement,*

$$\forall f, g, f \leq g \Rightarrow \psi(f) \leq \psi(g) \quad (5.6)$$

$$\forall f, \psi(\psi(f)) = \psi(f) \quad (5.7)$$

$$\forall f, \psi(f) \leq f \quad (5.8)$$

La *fermeture algébrique* se définit en remplaçant la propriété d'anti-extensivité par celle d'extensivité. Une ouverture morphologique est une ouverture algébrique mais l'inverse n'est pas vrai. Ainsi, il existe des ouvertures algébriques que l'on ne peut écrire sous la forme $(f \ominus B) \oplus B$. Les relations d'inclusion entre ouvertures sont illustrées à la figure 5.4.

La propriété de croissance dresse un fort parallélisme avec les filtres linéaires, surtout sous la forme suivante:

Corollaire 47

$$\psi \text{ est croissant} \Leftrightarrow \begin{cases} \psi(f \vee g) \geq \psi(f) \vee \psi(g) \\ \psi(f \wedge g) \leq \psi(f) \wedge \psi(g) \end{cases} \quad (5.9)$$

La propriété de croissance remplace l'additivité des filtres linéaires. Malheureusement, une transformation croissante produit une perte d'information; c'est pourquoi elle est irréversible. L'idempotence limite en quelque sorte la perte d'information parce qu'elle interrompt l'action simplificatrice à la première étape. Les filtres de base de la morphologie mathématique sont l'ouverture et la fermeture. Moyennant le choix judicieux des tailles et des formes d'éléments structurants, il sera possible de construire de nouveaux filtres suffisamment sélectifs pour éliminer des objets d'une certaine taille et qui auront les orientations indiquées par l'élément structurant. Par la suite, nous prendrons la liberté d'écrire "filtre" en lieu et place de "filtre algébrique".

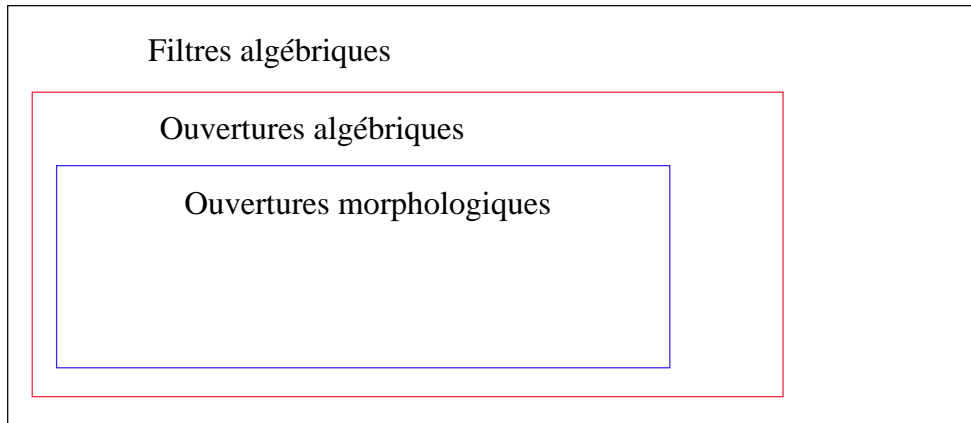


Figure 5.4: Relations d'inclusion entre familles de filtres.

5.3.2 Comment construire un filtre?

En combinant les filtres déjà connus! Mais toutes les combinaisons ne sont pas permises: par exemple, la composition de deux ouvertures n'est ni une ouverture ni un filtre, comme le montre la figure 5.5. Dans cet exemple, une nouvelle application donnera un ensemble vide et la propriété d'idempotence ne sera pas satisfaite.

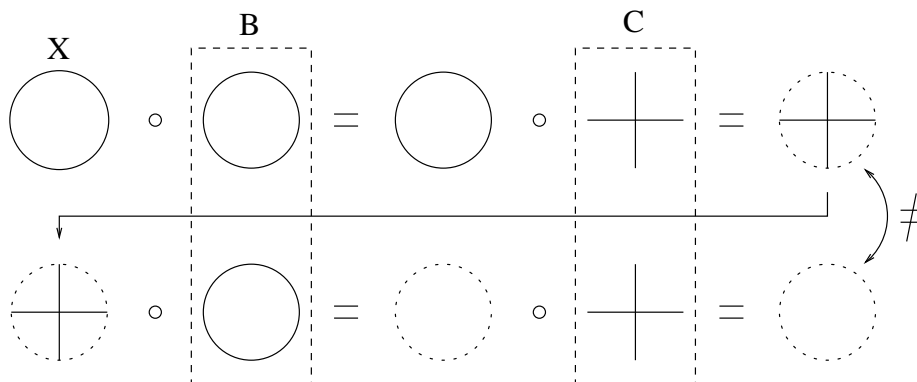


Figure 5.5: La composition de deux ouvertures n'est pas forcément une ouverture.

De nouveaux filtres peuvent être obtenus à partir des ouvertures, notées γ_i dans cette section, et des fermetures, notées ϕ_i , par les deux règles simples que voici:

1. le supremum d'ouvertures est une ouverture: $(\bigvee_i \gamma_i)$ est une ouverture;
2. l'infimum de deux fermetures est une fermeture: $(\bigwedge_i \phi_i)$ est une fermeture.

Certaines compositions de filtres permettent aussi d'obtenir des nouveaux filtres. Soient ψ_1 et ψ_2 deux filtres tels que $\psi_1 \geq I \geq \psi_2$ (par exemple, ψ_1 est une fermeture et ψ_2 une ouverture). Le théorème structurel définit quelles compositions sont des filtres et établit l'ordre des opérateurs.

Théorème 48 [Théorème structurel] Soient ψ_1 et ψ_2 deux filtres tels que $\psi_1 \geq I \geq \psi_2$

$$\psi_1 \geq \psi_1 \psi_2 \psi_1 \geq (\psi_2 \psi_1 \vee \psi_1 \psi_2) \geq (\psi_2 \psi_1 \wedge \psi_1 \psi_2) \geq \psi_2 \psi_1 \psi_2 \geq \psi_2 \quad (5.10)$$

$$\psi_1 \psi_2, \psi_2 \psi_1, \psi_1 \psi_2 \psi_1, \psi_2 \psi_1 \psi_2 \text{ sont des filtres} \quad (5.11)$$

Remarquons qu'il n'y a pas de relation d'ordre entre $\psi_1 \psi_2$ et $\psi_2 \psi_1$.

5.3.3 Exemples de filtres

Le nombre de filtres différents que l'on obtient en composant deux filtres est assez limité. Pour créer de nouveaux filtres, on peut composer non plus deux filtres mais deux *familles de filtres*. Cette idée a donné naissance aux filtres alternés séquentiels.

Filtres alternés séquentiels

Les ouvertures suppriment les détails blancs sans modifier les parties sombres de l'image tandis que les fermetures font l'inverse. Les filtres alternés séquentiels (FAS) adoptent un comportement plus symétrique par une composition alternée d'ouvertures et de fermetures. On commence par choisir un petit élément structurant qui est élargi au fur et à mesure jusqu'à une certaine taille.

Soient γ_i (ϕ_i) une ouverture (une fermeture) de taille i et I l'opérateur identité (i.e. $I(f) = f$). L'ordre entre ces opérations est le suivant:

$$\forall i, j \in \mathbb{N}, i \leq j, \quad \gamma_j \leq \gamma_i \leq I \leq \phi_i \leq \phi_j, \quad (5.12)$$

On définit ensuite les opérateurs suivants pour toute valeur de i :

$$\begin{aligned} m_i &= \gamma_i \phi_i, & r_i &= \phi_i \gamma_i \phi_i, \\ n_i &= \phi_i \gamma_i, & s_i &= \gamma_i \phi_i \gamma_i. \end{aligned}$$

En vertu du théorème structurel (cf. théorème 48), ces opérateurs sont bien des filtres. Ils entrent dans la composition des filtres alternés séquentiels.

Définition 49 [Filtres alternés séquentiels] *Pour tout $i \in \mathbb{N}$, les opérateurs suivants sont des filtres alternés séquentiels d'ordre i*

$$M_i = m_i m_{i-1} \dots m_2 m_1 \quad R_i = r_i r_{i-1} \dots r_2 r_1 \quad (5.13)$$

$$N_i = n_i n_{i-1} \dots n_2 n_1 \quad S_i = s_i s_{i-1} \dots s_2 s_1 \quad (5.14)$$

On peut montrer que M_i, N_i, R_i, S_i sont bien des filtres, c'est-à-dire croissants et idempotents. La loi d'absorption dit que si le filtrage le plus sévère a lieu après le filtrage le moins sévère, le résultat global du filtrage est égal à celui du filtrage sévère:

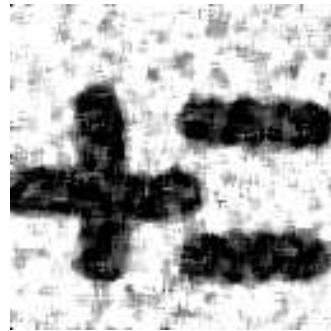
Propriété 50 [Loi d'absorption]

$$i \leq j \Rightarrow M_j M_i = M_j \text{ mais } M_i M_j \leq M_j \quad (5.15)$$

Les filtres alternés séquentiels sont très utiles pour l'étude des signaux bruités. On peut analyser le comportement du filtre sur la figure 5.6 dont l'image (a) de départ est fortement bruitée. Nous avons choisi un opérateur qui fait l'union de deux ouvertures obtenues avec des éléments structurants linéaires horizontaux et verticaux; la fermeture est l'intersection des fermetures par ces mêmes éléments structurants (cf. règles de construction de filtres à la page 105). Les images 5.6.(b) à 5.6.(d) montrent les étapes successives du filtre M_i et les images 5.6.(f) à 5.6.(h) celles de N_i . Pour la comparaison, nous avons également fourni l'image débruitée avec un filtre médian défini une fenêtre 5×5 .



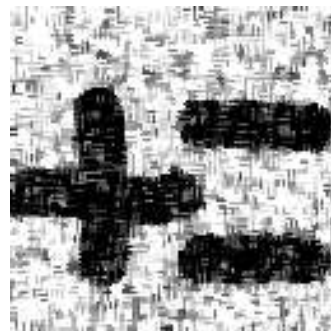
(a) Image originale f



(e) Médian 5×5



(b) $M_1(f)$



(f) $N_1(f)$



(c) $M_2(f)$



(g) $N_2(f)$



(d) $M_3(f)$



(h) $N_3(f)$

Figure 5.6: Utilisation de filtres alternés séquentiels pour supprimer le bruit.

Toggle mappings

Dans le but de créer d'autres filtres sans composition, considérons que nous disposons d'une famille de primitives et d'un critère tel que la sortie du filtre soit toujours l'une des valeurs de ces primitives. En général, la loi de choix de la primitive en question est fondé sur le supremum ou l'infimum, mais on peut trouver des critères plus complexes: c'est le *toggle mapping*.

Le centre morphologique est un exemple typique de toggle mapping.

Définition 51 [Centre morphologique] Soit ψ_i une famille d'opérateurs. Le centre morphologique β de la fonction f pour la famille ψ_i est défini en chaque x du domaine de définition de f de la manière suivante

$$\beta(f)(x) = (f(x) \vee (\bigwedge_i \psi_i(x))) \wedge (\bigvee_i \psi_i(x)) \quad (5.16)$$

Si toutes les fonctions de la famille indiquent un écart vers le haut ou vers le bas par rapport à la fonction de départ, le centre choisit de prendre la valeur la plus proche de f . Par contre si, pour certaines valeurs de x , les fonctions sont tantôt inférieures tantôt supérieures à f , le centre sélectionne l'original. En fait, dans le cas de trois fonctions, le centre choisit systématiquement la courbe centrale.

En général, le centre n'est pas idempotent mais si toutes les fonctions ψ_i sont croissantes, elles transmettent cette propriété au centre. De plus, si ψ_1 et ψ_2 sont des transformations duales, β est auto-dual, ce qui signifie que

$$\beta(f^c) = (\beta(f))^c \quad (5.17)$$

où $f^c(x) = \text{constante} - f(x)$.

L'auto-dualité assure que le blanc et le noir de l'image sont traités exactement de la même manière.

Prendre pour fonctions ψ_1 et ψ_2 une ouverture et une fermeture n'a aucun intérêt puisque le centre serait toujours égal à la fonction en vertu de leur propriété d'anti-extensivité et d'extensivité respective. On prendra plutôt l'opérateur $\gamma\phi\gamma$ et son dual $\phi\gamma\phi$.

L'action du centre morphologique est illustrée à la figure 5.7 pour un signal unidimensionnel. Si toutes les fonctions de la famille indiquent un écart vers le haut ou vers le bas par rapport à la fonction de départ, le centre choisit de prendre la valeur la plus proche de f . Par contre si, pour certaines valeurs de x , les fonctions sont tantôt inférieures tantôt supérieures à f , le centre sélectionne l'original.

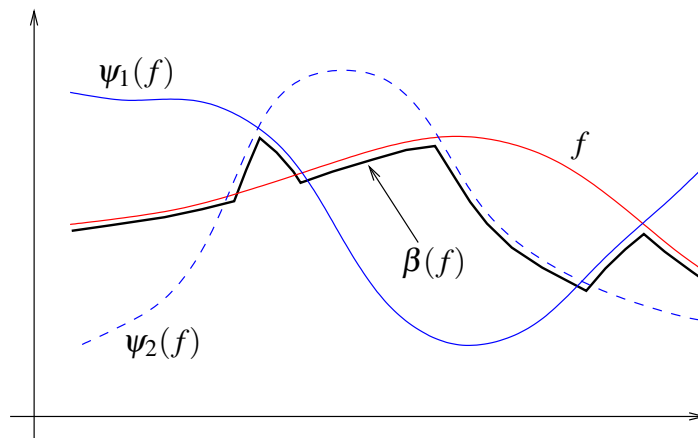


Figure 5.7: Centre morphologique d'un signal unidimensionnel.

Chapter 6

Extraction de traits caractéristiques et détection des contours

La détection de traits caractéristiques, comme une transition ou une orientation, est un domaine important du traitement d'image industriel. Quant à la détection de contours, elle est utilisée à différentes occasions, par exemple lors de la segmentation d'une image.

Par intuition, un contour est une transition marquée entre deux régions de luminosité distincte. En d'autres termes, un contour est une transition localisée dans le plan spatial et de contenu spectral à hautes fréquences mais également à large bande. Et c'est là tout l'origine du problème de la détection de contours car le bruit qui se superpose à l'image est bien souvent de contenu spectral similaire.

Plus fondamentale est la question de la définition d'un contour. Il se peut en effet fort bien qu'une image contienne des contours que l'œil ne perçoit pas. Cet effet est illustré à la figure 6.1 où l'on a considéré comme *contour* toute transition locale de luminance.

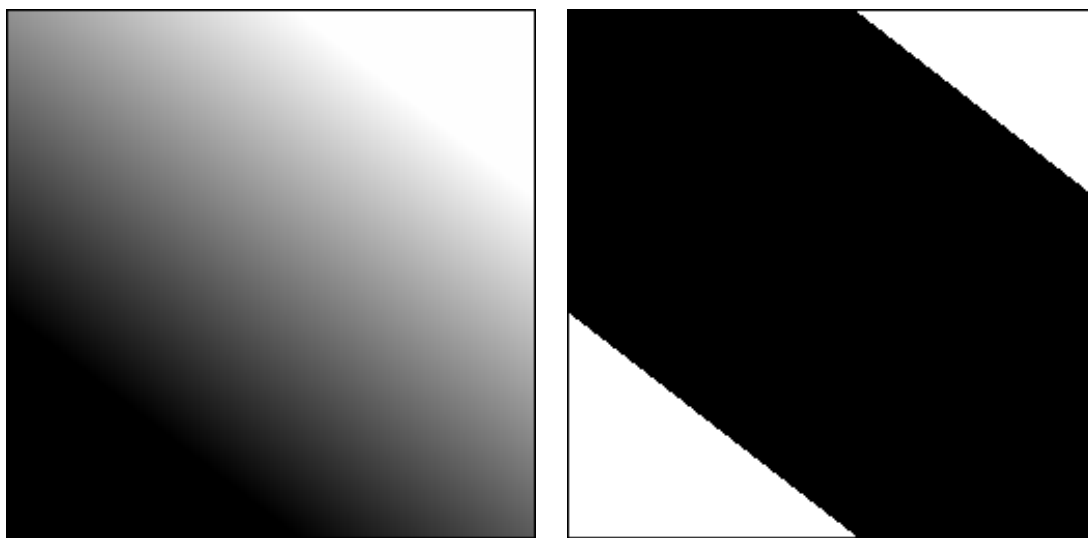


Figure 6.1: Une image en dégradé et ses contours (en noir).

Dans le cas de l'extraction de traits caractéristiques, c'est toute la configuration locale du contour qui importe (orientation, amplitude de la transition, signe de la dérivée, etc). La détection de contours se ramène assez souvent à ne considérer que l'amplitude de la dérivée. Remarquons aussi que la détection des contours peut fournir des contours non fermés.

Enfin, le positionnement d'un point de contour peut poser problème, notamment dans le cas d'une transition entre 2 zones plates (cf. figure 6.2): le point est-il situé à droite, à gauche ou milieu, étant

entendu que cette dernière solution ne permet pas une représentation simple du résultat?

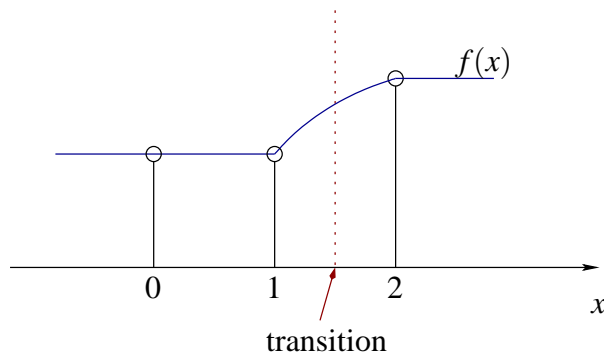


Figure 6.2: Problème de la position d'un point de contour.

6.1 Opérateurs linéaires

6.1.1 Opérateurs basés sur le calcul de la dérivée

Le but d'un opérateur d'extraction de contours est de fournir, en réponse à son application sur une image I_1 , une image I_2 à fortes variations aux différents endroits où I_1 présente des contours. Un opérateur répondant à ce critère est, sans conteste, l'opérateur différentiel ou dérivée. Deux choix usuels s'offrent à nous:

1. Si l'on se place dans le contexte d'une fonction continue $f(x)$ à une dimension, sa dérivée $f'(x)$ est maximale (en valeur absolue) aux différents endroits où $f(x)$ présente de fortes variations; cette valeur maximale étant d'autant plus grande que la variation de f est importante sur une courte distance.
2. La dérivée seconde $f''(x)$ présente, quant à elle, un passage par zéro aux différents endroits où $f(x)$ contient de fortes variations.

L'opérateur dérivée (première ou seconde) est linéaire et peut être vu comme un filtre. Nous verrons plus loin que le filtre dérivée est un filtre du type passe-haut. L'application de l'opérateur dérivée impliquera donc un très net effet accentuateur des hautes fréquences.

Le problème majeur que l'on rencontre en pratique est qu'une image n'est pas une fonction continue. Il n'est donc pas possible d'en calculer la dérivée exacte; on peut tout juste l'approximer par différentes formules. D'autre part, l'effet accentuateur des hautes fréquences, inhérent à l'opérateur de dérivée, entraîne une amplification du bruit.

On est donc face à un double problème:

- approximer au mieux la dérivée et
- éviter une amplification excessive du bruit; cette amplification est malheureusement intrinsèque à l'utilisation d'un opérateur dérivé. L'effet d'amplification du bruit est illustré à la figure 6.3.

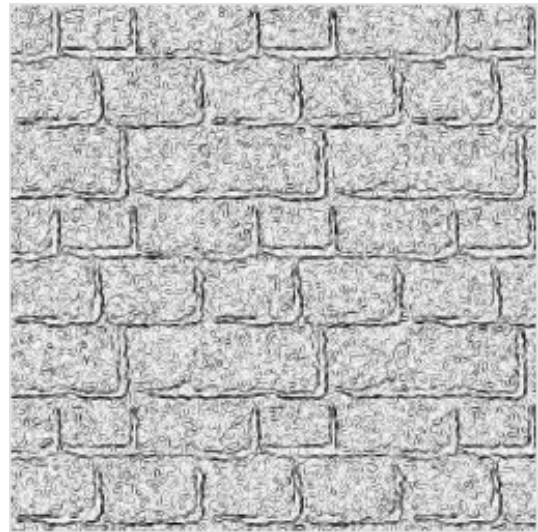
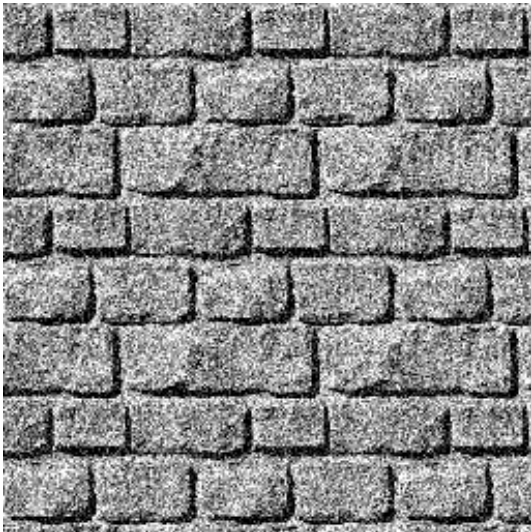
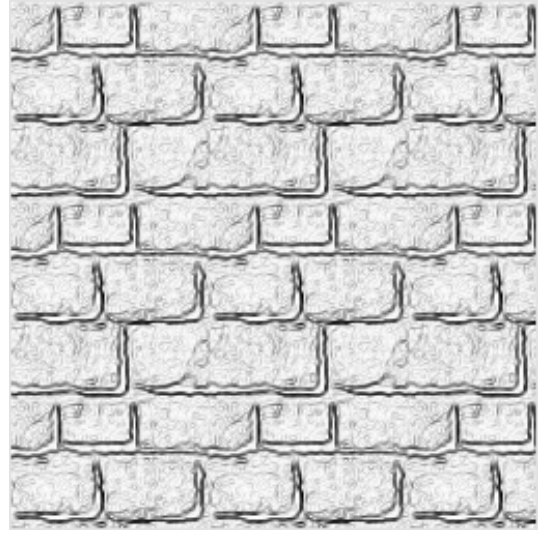
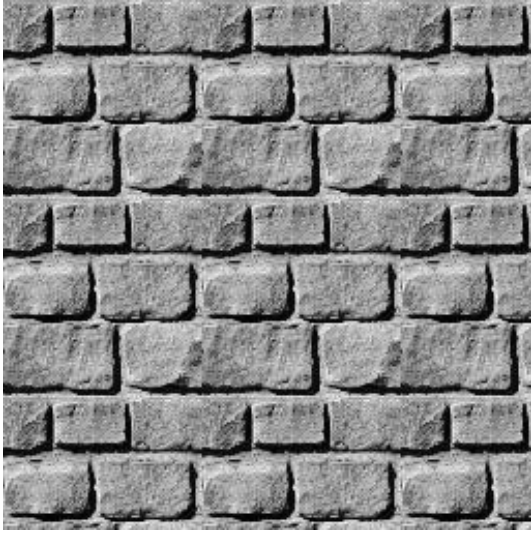


Figure 6.3: Images (à gauche) et gradients respectifs (à droite).

6.1.2 Opérateur de dérivée première

Par commodité, nous considérons une image caractérisée par une fonction $f(x, y)$ bidimensionnelle et continue. Il faut, dans ce cas, indiquer dans quelle direction on calcule la dérivée. Les principaux choix sont les directions horizontale \vec{e}_x ou verticale \vec{e}_y ; une dérivée dans une autre direction $\vec{\theta}$ peut se ramener à une combinaison des dérivées suivant les axes principaux –il s’agit alors de dérivée *directionnelle*.

Considérons la dérivée *partielle* de la fonction $f(x, y)$ par rapport à x . La transformée de FOURIER de cet opérateur vaut

$$\frac{\partial f}{\partial x}(x, y) \Leftrightarrow 2\pi ju \mathcal{F}(u, v) \quad (6.1)$$

Dérivée par rapport à x équivaut donc à multiplier la transformée de FOURIER de $f(x, y)$ par la fonction de transfert $\mathcal{H}_x(u, v) = 2\pi ju$, c’est-à-dire filtrer $f(x, y)$ avec le filtre dont la réponse impulsionnelle est donnée par

$$h_x(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (2\pi ju) e^{2\pi j(xu+yv)} dudv \quad (6.2)$$

De même, on peut définir la réponse impulsionnelle $h_y(x, y)$ correspondant au filtrée dérivée dans la direction \vec{e}_y . Le module de la fonction de transfert $\mathcal{H}_x(u, v)$ étant égal à $2\pi|u|$, on voit clairement qu’il y a un effet accentuateur des hautes fréquences, d’où le caractère passe-haut de l’opérateur dérivée. C’est pareil pour $h_y(x, y)$.

En adoptant une notation vectorielle de la dérivée, on définit le gradient ∇f de l’image f par

$$\nabla f = \frac{\partial f}{\partial x} \vec{e}_x + \frac{\partial f}{\partial y} \vec{e}_y = (h_x \otimes f) \vec{e}_x + (h_y \otimes f) \vec{e}_y \quad (6.3)$$

Le gradient d’une fonction $f(x, y)$ est une fonction vectorielle. Mais plutôt que de recourir aux composantes en x et y , on peut caractériser un gradient par son amplitude et sa direction.

Définition 52 [Amplitude du gradient]

$$|\nabla f| = \sqrt{(h_x \otimes f)^2 + (h_y \otimes f)^2} \quad (6.4)$$

Remarque pratique: approximation du gradient

On approxime parfois l’amplitude du gradient (ou d’une norme) par l’expression

$$|\nabla f| \simeq |h_x \otimes f| + |h_y \otimes f| \quad (6.5)$$

pour éviter le calcul de la racine carrée et les élévations au carré afin d’accélérer le calcul mais cette formule n’est qu’une approximation. Cette approximation est tout à fait raisonnable lorsqu’un terme dépasse largement l’autre. Par exemple, si $h_x \otimes f \ll h_y \otimes f$, $|\nabla f| \simeq |h_y \otimes f|$. Dans le cas où les deux termes sont d’amplitude égale, on a $|\nabla f| \simeq \sqrt{2} |h_x \otimes f|$. L’erreur peut donc atteindre 41%.

Définition 53 [Direction du gradient]

$$\varphi_{\nabla f} = \tan^{-1} \left(\frac{h_y \otimes f}{h_x \otimes f} \right) \quad (6.6)$$

Étant donnée la définition du gradient d’une fonction, il est maintenant possible de définir sa dérivée directionnelle par

Définition 54 [Dérivée directionnelle]

$$\frac{\partial f}{\partial \vec{\theta}} = \vec{\theta} \cdot \nabla f \quad (6.7)$$

Il s'agit du produit scalaire du gradient de f et d'un vecteur caractérisant la direction dans laquelle on désire calculer la dérivée. Si la direction $\vec{\theta}$ fait un angle θ avec l'axe x , c'est-à-dire $\vec{\theta} = \cos \theta \vec{e}_x + \sin \theta \vec{e}_y$, la dérivée directionnelle peut encore s'écrire

$$\frac{\partial f}{\partial \vec{\theta}} = \cos \theta \frac{\partial f}{\partial x} + \sin \theta \frac{\partial f}{\partial y} \quad (6.8)$$

Pour $\theta = 0$, nous retrouvons la dérivée directionnelle dans la direction \vec{e}_x . La dérivée directionnelle est utile pour mettre en évidence les contours d'une image qui ont une orientation particulière.

6.1.3 Opérateur de dérivée seconde

Comme pour la dérivée première, on peut définir les réponses impulsionnelles $h_{xx}(x,y)$ et $h_{yy}(x,y)$ correspondant aux filtres dérivée seconde dans les directions \vec{e}_x et \vec{e}_y . On définit le laplacien d'une fonction continue $f(x,y)$ par

Définition 55 [Laplacien]

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = (h_{xx} \otimes f) + (h_{yy} \otimes f) \quad (6.9)$$

On peut aisément tirer l'équivalence suivante dans le domaine de FOURIER

$$\nabla^2 f \rightleftharpoons -4\pi^2(u^2 + v^2)\mathcal{F}(u,v) \quad (6.10)$$

Ici encore, on remarque un effet accentuateur des hautes fréquences.

6.1.4 Échantillonnage de la dérivée et calcul de l'erreur résultante

En pratique, les images sur lesquelles nous devons calculer la dérivée sont échantillonnées. Il est alors nécessaire d'échantillonner l'opérateur de dérivée première et seconde en les approximant au mieux. Il va apparaître que l'échantillonnage entraîne une erreur dans le calcul de la dérivée. Afin de déterminer des expressions pratiques pour le calcul des opérateurs dérivés, nous adoptons la démarche suivante:

- synthèse d'approximations avec calcul de l'erreur induite,
- étude du comportement fréquentiel des approximations, et
- discussions de quelques formules pratiques fournies sous la forme de masques de convolution à étendue finie (cf. section 6.1.5).

Calcul de l'erreur

Considérons une fonction $f(x)$ à une dimension. Une première approximation de sa dérivée première est donnée par

$$f'_a(x) = \frac{f(x+h) - f(x-h)}{2h} \quad (6.11)$$

où h est le pas d'échantillonnage et l'indice a rappelle qu'il s'agit d'une *approximation*. Cette approximation correspond à l'utilisation du masque de convolution centré suivant

$$\frac{1}{2h} \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \quad (6.12)$$

Pour la dérivée seconde, nous avons l'approximation suivante

$$f''_a(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \quad (6.13)$$

Ces deux approximations sont dites *centrées* car elles utilisent des valeurs de la fonction de part et d'autre de x . Nous allons montrer qu'elles sont d'ordre $O(h^2)$, c'est-à-dire que l'erreur est composée de termes en h dont la puissance est supérieure ou égale à 2. Pour cela, nous utilisons les développements de TAYLOR suivant

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2}f''(x) + \dots + \frac{h^n}{n!}f^{(n)}(x) + \dots \quad (6.14)$$

$$f(x-h) = f(x) - hf'(x) + \frac{h^2}{2}f''(x) + \dots + (-1)^n \frac{h^n}{n!}f^{(n)}(x) + \dots \quad (6.15)$$

Dérivé première. Par soustraction membre à membre de ces deux égalités, nous obtenons

$$f(x+h) - f(x-h) = 2hf'(x) + \frac{2}{3!}h^3f^{(3)}(x) + \dots = 2hf'(x) + O(h^3) \quad (6.16)$$

Après réarrangement,

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2) \quad (6.17)$$

L'approximation $f'_a(x)$ est donc égale à la dérivée exacte $f'(x)$ à un terme d'erreur $O(h^2)$ près. Ce terme est d'autant plus petit que le pas d'échantillonnage est petit. Pour une erreur d'ordre $O(h^2)$, si on divise h par deux, le terme d'erreur est approximativement quatre fois plus petit. D'une manière générale, l'erreur de discrétisation $O(h^p)$ d'un opérateur est d'autant plus faible que l'ordre p de l'erreur est élevé.

Dérivée seconde. Similairement au cas de la dérivée première, par addition des deux égalités 6.14 et 6.15,

$$f(x+h) + f(x-h) = 2f(x) + h^2f''(x) + \frac{2}{4!}h^4f^{(4)}(x) + \dots \quad (6.18)$$

D'où

$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + O(h^2) \quad (6.19)$$

L'approximation $f''_a(x)$ est donc également d'ordre 2.

Synthèse d'opérateurs dérivés avec erreur préfixée. Supposons à présent que l'on désire construire une approximation centrée de la dérivée première dont l'erreur est d'ordre $O(h^{2p})$ où $p \geq 1$. Choisissons l'approximation $f'_a(x)$ de la manière suivante

$$f'_a(x) = \frac{1}{h} \sum_{i=-p}^p T_i f(x+ih) \quad (6.20)$$

où les $(2p + 1)$ coefficients T_i ($i = -p, \dots, 0, \dots, p$) doivent être déterminés. Cette expression garantit l'obtention d'une approximation *centrée*. Introduisons à présent le développement de TAYLOR suivant

$$f(x + ih) = \sum_{n=0}^{+\infty} \frac{(ih)^n}{n!} f^{(n)}(x) \quad (6.21)$$

dans l'expression précédente:

$$f'_a(x) = \frac{1}{h} \sum_{i=-p}^p T_i \left(\sum_{n=0}^{+\infty} \frac{(ih)^n}{n!} f^{(n)}(x) \right) \quad (6.22)$$

$$= \frac{1}{h} \sum_{n=0}^{+\infty} \left(f^{(n)}(x) \sum_{i=-p}^p \frac{i^n}{n!} T_i \right) h^n \quad (6.23)$$

$$= f'(x) + O(h^{2p}) \quad (6.24)$$

Cette dernière égalité est vérifiée si le système d'équations suivant est satisfait par les $(2p + 1)$ coefficients T_i ; c'est-à-dire pour $0 \leq n \leq 2p$, il faut que

$$\sum_{i=-p}^p \frac{i^n}{n!} T_i = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n \neq 1 \end{cases} \quad (6.25)$$

Cela nous donne un système de $(2p + 1)$ équations à $(2p + 1)$ inconnues. Pour $p = 1$, c'est-à-dire l'approximation d'ordre 2, le système nous donne $T = (T_{-1}, T_0, T_1) = \frac{1}{2}(-1, 0, 1)$, ce qui correspond à l'approximation donnée plus haut. Pour $p = 2$, la solution du système est donnée par $T = (T_{-2}, T_{-1}, T_0, T_1, T_2) = \frac{1}{8}(1, -8, 0, 8, -1)$, ce qui correspond à l'approximation d'ordre $O(h^4)$ suivante

$$f'_a(x) = \frac{-f(x + 2h) + 8f(x + h) - 8f(x - h) + f(x - 2h)}{12h} \quad (6.26)$$

Cette formule approxime mieux la dérivée mais elle nécessite plus d'efforts de calcul vu le nombre accru de termes utilisés pour son évaluation.

Comportement fréquentiel des opérateurs échantillonnés

Nous venons de voir que le fait d'approximer l'opérateur de dérivée provoque l'apparition d'erreurs d'échantillonnage. À présent, nous allons étudier l'effet de l'échantillonnage dans le domaine spectral.

Considérons une fonction continue $f(x)$ à une dimension et l'approximation de l'opérateur de dérivée première suivante

$$f'_a(x) = \frac{f(x + h) - f(x - h)}{2h} \quad (6.27)$$

Sa transformée de FOURIER vaut

$$\frac{f(x + h) - f(x - h)}{2h} \Leftrightarrow \frac{e^{2\pi juh} - e^{-2\pi juh}}{2h} \mathcal{F}(u) \quad (6.28)$$

ce qui peut encore s'écrire sous la forme

$$\frac{f(x + h) - f(x - h)}{2h} \Leftrightarrow (2\pi ju) \frac{\sin(2\pi hu)}{2\pi hu} \mathcal{F}(u) \quad (6.29)$$

Le facteur $(2\pi ju)$ correspond à l'opérateur idéal (ou continu) de la dérivée première. Dès lors, l'approximation de la dérivée première se ramène à l'opérateur idéal multiplié, dans le domaine

spectral, par une fonction fenêtre du type sinus cardinal. Étant donné le pas d'échantillonnage h , la fréquence de NYQUIST vaut $1/2h$.

Pour $u \ll 1/2h$, le sinus cardinal vaut approximativement 1 et l'approximation de la dérivée est bonne. Par contre, plus u tend vers la fréquence de NYQUIST (où le sinus cardinal s'annule), plus l'approximation perd en qualité. Au-delà de la fréquence de NYQUIST, l'approximation est médiocre.

Le même raisonnement peut être fait pour l'opérateur approximant la dérivée seconde suivant

$$f_a''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \quad (6.30)$$

Sa transformée de FOURIER est donnée par

$$\frac{f(x+h) - 2f(x) + f(x-h)}{h^2} \Leftrightarrow (-4\pi^2 u^2) \left(\frac{\sin(\pi hu)}{\pi hu} \right)^2 \mathcal{F}(u) \quad (6.31)$$

Le facteur $(-4\pi^2 u^2)$ correspond à l'opérateur idéal de dérivée seconde. Les mêmes remarques que pour la dérivée première peuvent être faites ici à la différence près que la fenêtre spectrale est du type sinus cardinal au carré.

La figure 6.4 montre le module de la transmittance des approximations précitées (trait plein) par rapport aux opérateurs idéaux (trait discontinu).

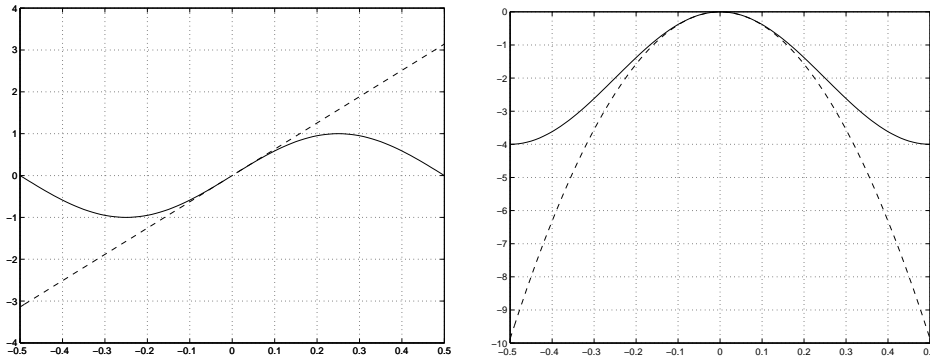


Figure 6.4: Comportement fréquentiel des approximations ($h = 1$). À gauche, la dérivée première et à droite, la dérivée seconde.

6.1.5 Expressions pratiques d'opérateurs de gradient et masques de convolution

À partir des différents opérateurs discrets introduits plus haut, nous allons voir comment définir différentes formes de base de filtre ainsi que les masques de convolution que l'on pourra appliquer directement aux images échantillonnées.

Le passage le plus simple des expressions analytiques valables pour des images continues à des images échantillonnées s'effectue en considérant un pas d'échantillonnage h identique en x et en y . Par convention, h est choisi égal à 1; dès lors, $f(x+h, y)$ représente alors le pixel situé juste à droite (ou à gauche selon l'orientation de l'axe x) du pixel $f(x, y)$.

Formes de base pour la dérivée première

L'approximation la plus sommaire de la dérivée première dans la direction x consiste à prendre la différence entre deux pixels voisins en utilisant le *masque de convolution* suivant¹

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \quad (6.32)$$

ce qui correspond en fait à l'approximation non-centrée suivante

$$\frac{f(x+h,y) - f(x,y)}{h} \quad (6.33)$$

Ce masque présente un inconvénient majeur: le résultat est décalé d'un demi pixel par rapport aux points utilisés pour le calcul. On lui préfère généralement le masque étendu suivant

$$\begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \quad (6.34)$$

qui correspond à l'approximation centrée décrite en détail précédemment. Ce petit masque s'applique directement sur la grille des pixels de l'image que l'on désire traiter, en plaçant le 0 sur le pixel en cours de traitement. Une approximation de la dérivée première dans la direction y est donnée par

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad (6.35)$$

Dans la foulée, on peut imaginer l'utilisation du masque suivant

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad (6.36)$$

pour procéder à une dérivation directionnelle dans la direction diagonale 135° . La figure 6.5 montre le résultat de l'application de divers opérateurs de dérivée. L'image originale est définie sur 8 bits de luminance; les images gradients ont été décalées de 128.

Problème pratiques. L'utilisation des masques de convolution pose quelques problèmes pratiques:

- **Effets de bord.** L'application d'un masque est délicate près d'un bord puisqu'il faudrait inclure des valeurs qui ne se situent hors image. Ce problème récurrent pour l'application d'opérateurs locaux est résolu de différentes manières, selon le type de l'opérateur:
 - (i) placement d'une *valeur* hors image *par défaut*; cette valeur dépendra de l'opérateur appliqué,
 - (ii) *miroirisation*, qui consiste à miroiriser les valeurs de l'intérieur vers l'extérieur de l'image,
 - (iii) *périodisation* de l'image –on retrouve alors les pixels de la gauche contre les pixels du bord droit–, qui se justifie par la périodisation spatiale subséquente à l'échantillonnage du spectre,

¹En toute rigueur, tous les masques sont précédés de facteurs de normalisation. Nous négligerons parfois ces facteurs car, en pratique, on renormalise bien souvent le résultat.

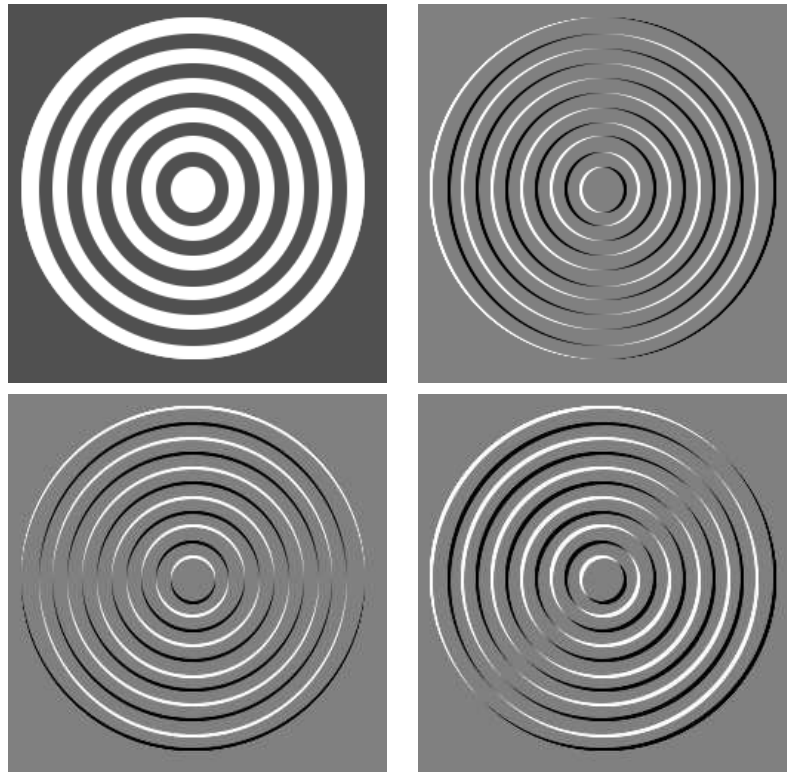


Figure 6.5: Illustrations de l'application de masques de dérivée première: (a) image originale, (b) application d'un masque horizontal, (c) application d'un masque vertical et (d) application d'un masque à 135° .

- (iv) *recopiage* des valeurs du bord de l'image. Cette opération consiste à recopier les valeurs des lignes et colonnes qui forment le bord intérieur de l'image.

- **Modifications de la dynamique de valeurs.** Le résultat de l'opérateur n'est pas nécessairement compris dans la dynamique de départ. En effet, considérons une image définie sur l'intervalle $[0, 255]$. L'application du masque $\begin{bmatrix} -1 & 1 \end{bmatrix}$ fournit un résultat compris dans l'intervalle $[-255, 255]$. Sans renormalisation, il n'est pas possible de représenter correctement l'ensemble de ces valeurs avec 8 bits. On procède alors à une normalisation de l'échelle de manière à ramener les valeurs dans l'intervalle de départ. Dans l'exemple précédent, on appliquera une division par 2 et suivie d'une addition de 128. On pourrait aussi ne considérer que le module du gradient, ce qui ramène automatiquement le résultat dans la dynamique initiale.
- **Facteurs de normalisation.** Pour des raisons connexes à celles évoquées au point précédent, il n'est pas toujours utile de considérer le facteur de normalisation correct de l'opérateur dérivé; on préfère en général procéder à la renormalisation des valeurs.

Filtres gradient de PREWITT. Les formes de base effectuent un gradient ligne par ligne ou colonne par colonne. Les directions horizontales et verticales sont dès lors privilégiées. Il existe des formes qui permettent de fournir un gradient aux caractéristiques plus isotropes. Pour y parvenir, ces formes s'étendent dans les deux directions.

Les filtres de gradient de PREWITT appartiennent à cette classe d'opérateurs gradient. Ils valent

$$[h_x] = \frac{1}{3} \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \otimes [-1 \ 0 \ 1] \quad (6.37)$$

$$[h_y] = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \otimes [1 \ 1 \ 1] \quad (6.38)$$

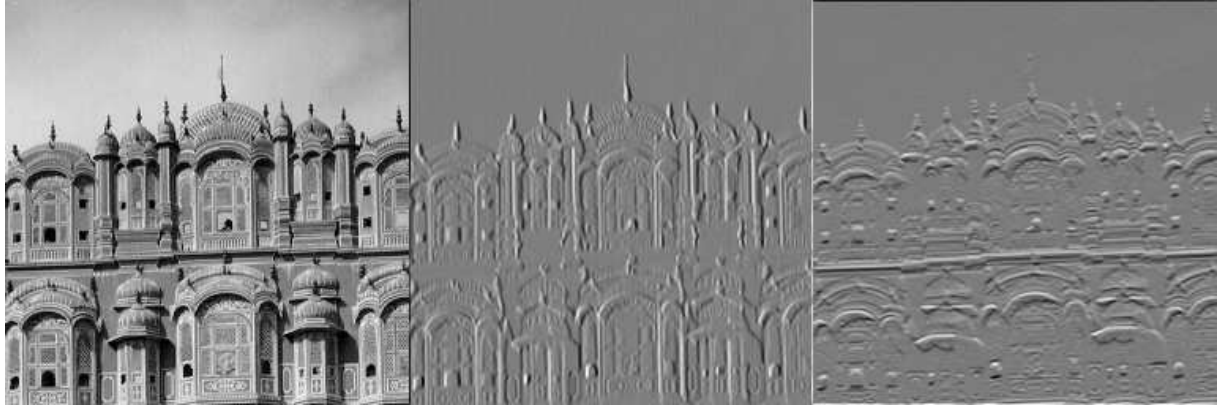


Figure 6.6: Image originale, images traitées respectivement par le filtre horizontal et le filtre vertical de PREWITT.

La figure montre l'effet des filtres de PREWITT. Le calcul pratique s'effectue en appliquant successivement une convolution par la colonne verticale puis par la ligne horizontale, plutôt qu'en traitant l'image par la matrice carrée. En effet, on sait que $f(x,y) \otimes (h_{xv} \otimes h_{xh}) = (f(x,y) \otimes h_{xv}) \otimes h_{xh}$. Cette astuce algorithmique réduit quelque peu le nombre d'opérations à effectuer.

Filtres gradient de SOBEL. Les filtres de SOBEL constituent une alternative aux filtres de PREWITT. Ils s'expriment comme suit

$$[h_x] = \frac{1}{4} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \otimes [-1 \ 0 \ 1] \quad (6.39)$$

$$[h_y] = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \otimes [1 \ 2 \ 1] \quad (6.40)$$

La figure 6.7 montre l'effet des filtres de SOBEL.

La comparaison des figures 6.6 et 6.7 montre que les différences entre les images filtrées sont minimes.

Formes de base pour la dérivée seconde

Les formes de base comprennent les filtres horizontaux et verticaux suivants

$$[1 \ -2 \ 1] \quad (6.41)$$

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad (6.42)$$

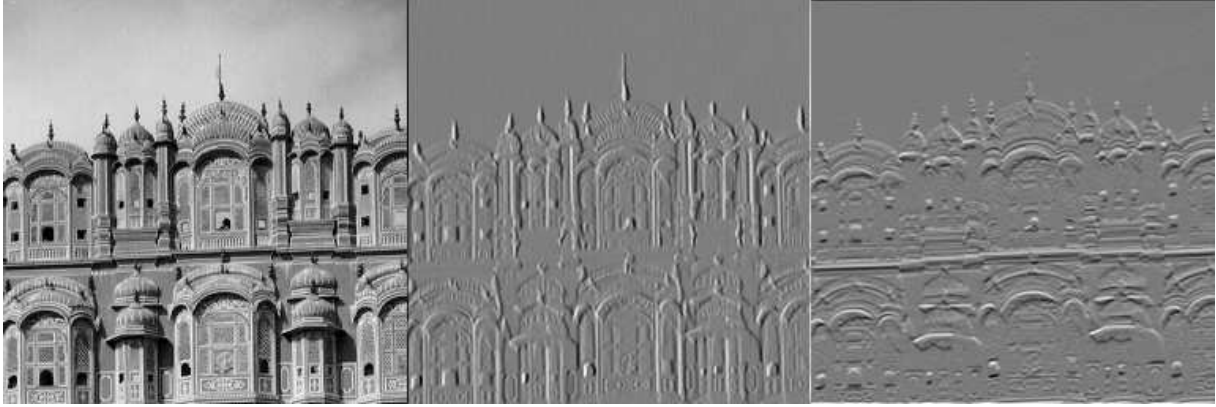


Figure 6.7: Une image originale et les images traitées par application d'un masque de SOBEL (respectivement horizontal et vertical).

provenant directement de l'approximation de la dérivée seconde donnée précédemment.

On peut obtenir un filtre bidimensionnel en combinant les deux expressions précédentes ou lui préférer un masque de type

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (6.43)$$

qui a un caractère plus isotrope.

La figure 6.8 illustre l'application de ces différents masques; on y a représenté respectivement l'image originale, le laplacien horizontal, le laplacien vertical, le laplacien obtenu avec le masque 6.43 et enfin l'image résultant de l'application du masque suivant

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (6.44)$$

6.2 Opérateurs non-linéaires

6.2.1 Gradients morphologiques

Si B est un élément structurant de petite dimension, la différence ensembliste $X - (X \ominus B)$ fournit les contours intérieurs de l'image binaire X , alors que la différence $(X \oplus B) - X$ produit les contours extérieurs de l'image. On peut réaliser des opérations similaires sur des images en niveaux de gris. L'opérateur

$$GE(f) = f - (f \ominus B) \quad (6.45)$$

appelé *gradient d'érosion*, constitue un moyen commode d'accentuer les transitions d'une image en niveaux de gris f . Un opérateur qui remplit la même fonction, est le *gradient de dilatation*

$$GD(f) = (f \oplus B) - f \quad (6.46)$$

Par combinaison de ces deux opérateurs, il est possible de synthétiser une kyrielle de nouveaux opérateurs, tous destinés à extraire les contours de l'image. L'objectif d'une utilisation des deux

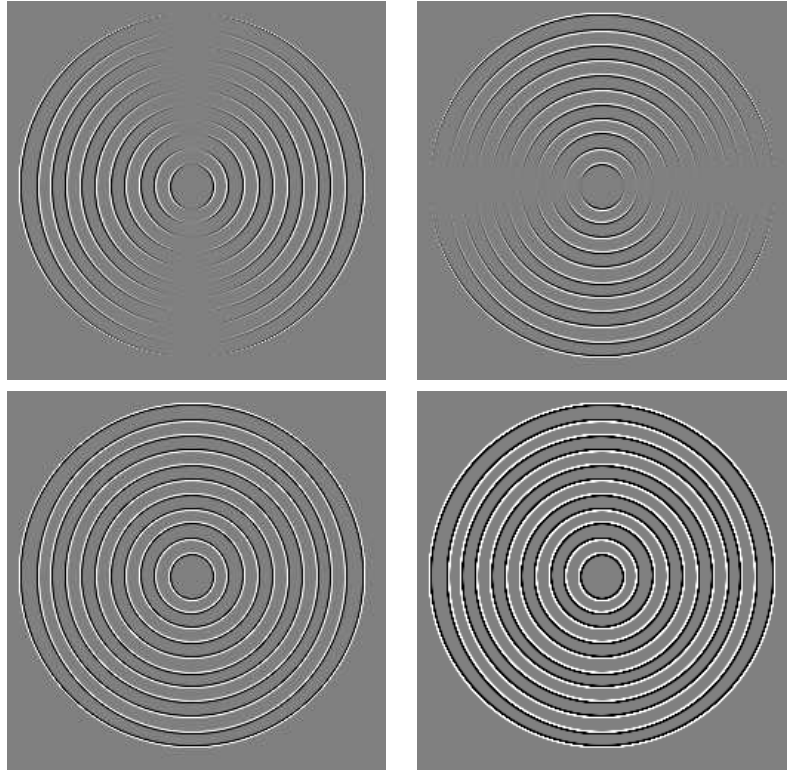


Figure 6.8: Illustrations de l'application de masques de dérivée seconde.

opérateurs est la symétrisation du traitement d'une image et de son arrière-fond. Les exemples sont nombreux (cf. figures 6.9 et 6.10):

- le *gradient morphologique* de BEUCHER: $GE(f) + GD(f)$. Il est illustré à la figure 6.9;
- le *chapeau haut-de-forme* (*top-hat* en anglais). On en définit 2:
 - le *white top-hat*, égal à l'image moins l'ouvert: $f - (f \circ B)$
 - le *black top-hat*, égal au fermé moins l'image: $(f \bullet B) - f$
- les opérateurs de détection de contours marqués: $\min(GE(f), GD(f))$, $\max(GE(f), GD(f))$
- le laplacien non-linéaire: $GD(f) - GE(f)$.

Tous ces opérateurs morphologiques sont rendus plus robustes si un filtrage usuel précède leur emploi. Les schémas de filtrage hybrides, formés par mélange d'opérateurs dits classiques et de gradients morphologiques, ont des performances comparables voire supérieures à ceux des schémas conventionnels basés sur la combinaison de filtres et de gradients linéaires [39]. En outre, les opérateurs morphologiques nécessitent moins d'opérations.



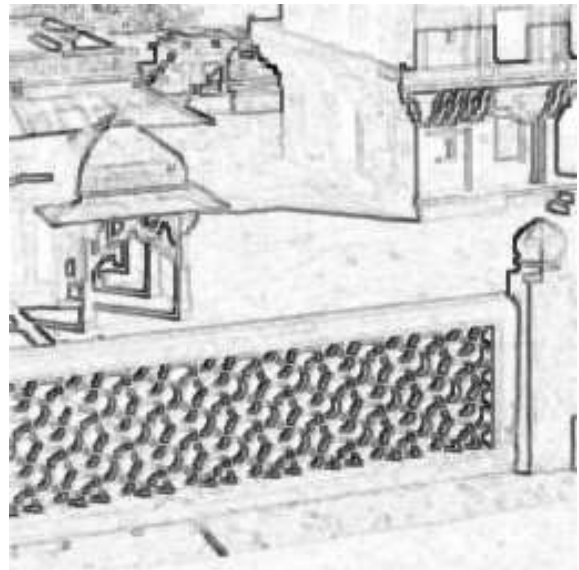
(a) Image originale f



(b) $f \oplus B$

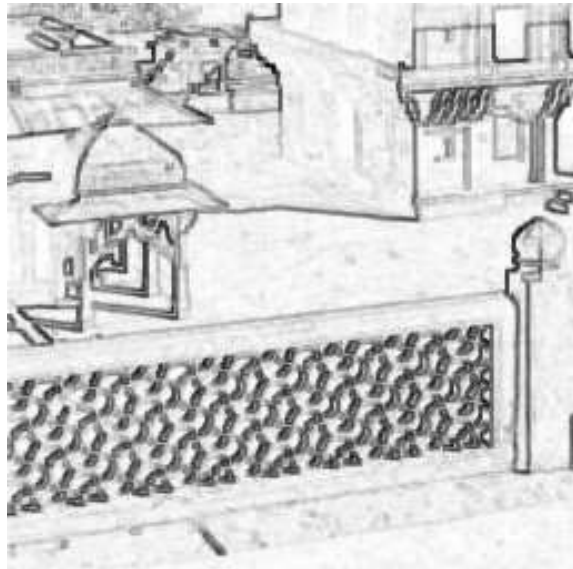


(c) $f \ominus B$

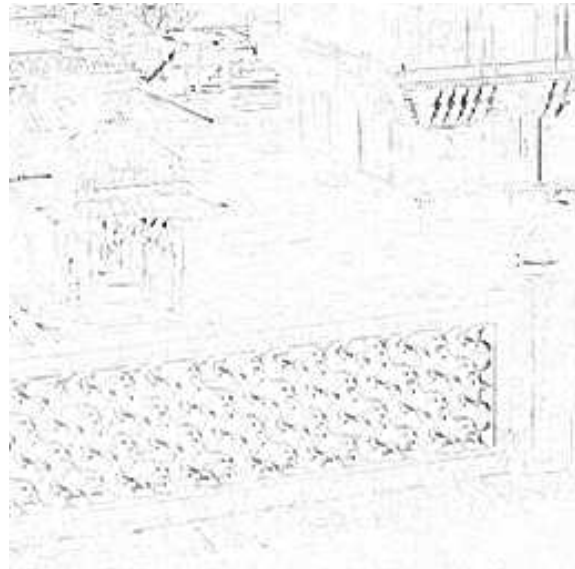


(d) $(f \oplus B) - (f \ominus B)$ (vidéo inverse)

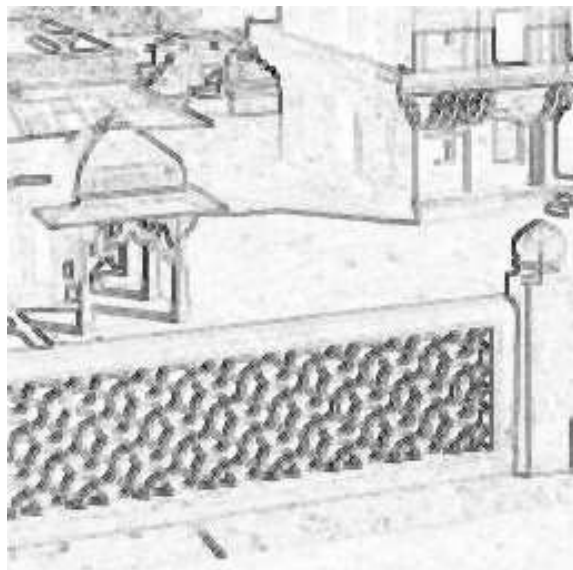
Figure 6.9: Gradient de BEUCHER.



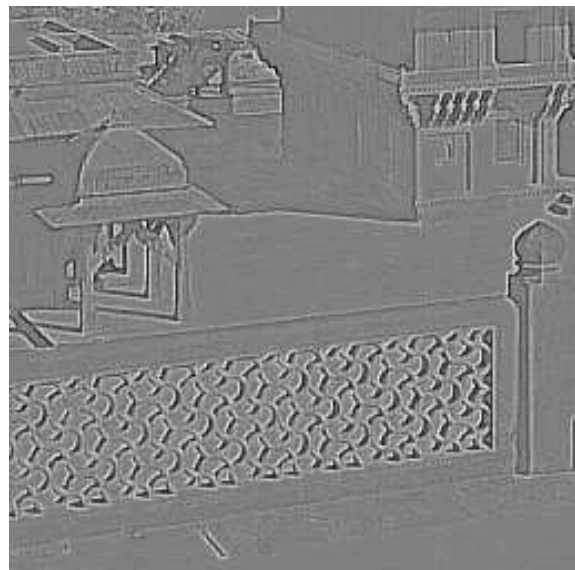
(a) $(f \oplus B) - (f \ominus B)$



(b) $f - f \circ B$ (white top-hat)



(c) $\max(GE(f), GD(f))$



(d) $GD(f) - GE(f)$

Figure 6.10: Différents détecteurs de contours non-linéaires (vidéo inverse).

Chapter 7

Texture

7.1 Introduction

Si caractériser la forme est relativement aisé, définir la notion de texture relève presque du défi. La texture fait référence aux propriétés de surface ou de structures d'un objet. Une texture est constituée de primitives ou d'éléments de textures parfois appelés *texels*. La difficulté majeure consiste à définir des constantes à plusieurs échelles différentes. En fait, la texture est fortement dépendante de l'échelle.

Les principales questions relatives aux textures sont

- l'*analyse des textures*, c'est-à-dire la caractérisation des textures par l'introduction de différents paramètres appelés *descripteurs* de texture. On décrit généralement des textures par des notions comme la finesse, la rugosité, etc.
- la *reconnaissance de textures*, c'est-à-dire la possibilité de mettre en évidence différentes textures présentes dans une image de test grâce, par exemple, à une base de donnée contenant tout un ensemble de textures déjà étudiées.
- la *segmentation d'images* qui, via la reconnaissance des textures, permet de découper l'image en différentes sous-régions présentant les mêmes caractéristiques.

La définition de texture fournie ci-après met en lumière l'origine des difficultés que l'on peut rencontrer lors de leur étude.

Définition 56 *Une texture est un signal qui peut être étendu naturellement au-delà de son domaine de définition.*

Ainsi donc, la définition d'une texture exige la présence d'une certaine régularité dans le signal, régularité qui ne signifie pas nécessairement périodicité, et elle s'appuie sur une extrapolation cohérente. Le signal de la figure 7.1 est un exemple de texture. En voyant ce signal, on comprend que l'extrapolation d'une texture n'est pas toujours unique; elle dépend de la manière d'exprimer la régularité intrinsèque.

Vu la complexité des textures naturelles (figure 7.2 pour quelques échantillons de textures naturelles), un procédé de segmentation s'avère incapable de les détecter toutes avec précision.

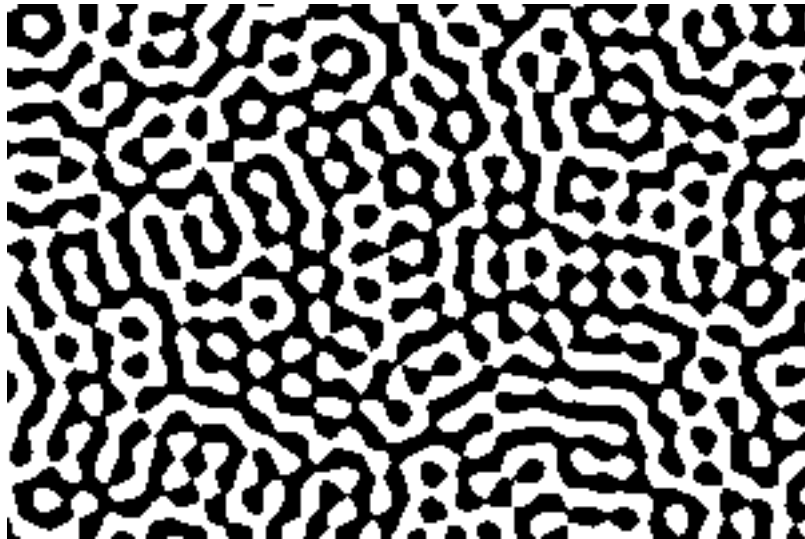


Figure 7.1: Un exemple de texture (d'après LANTUÉJOUL).

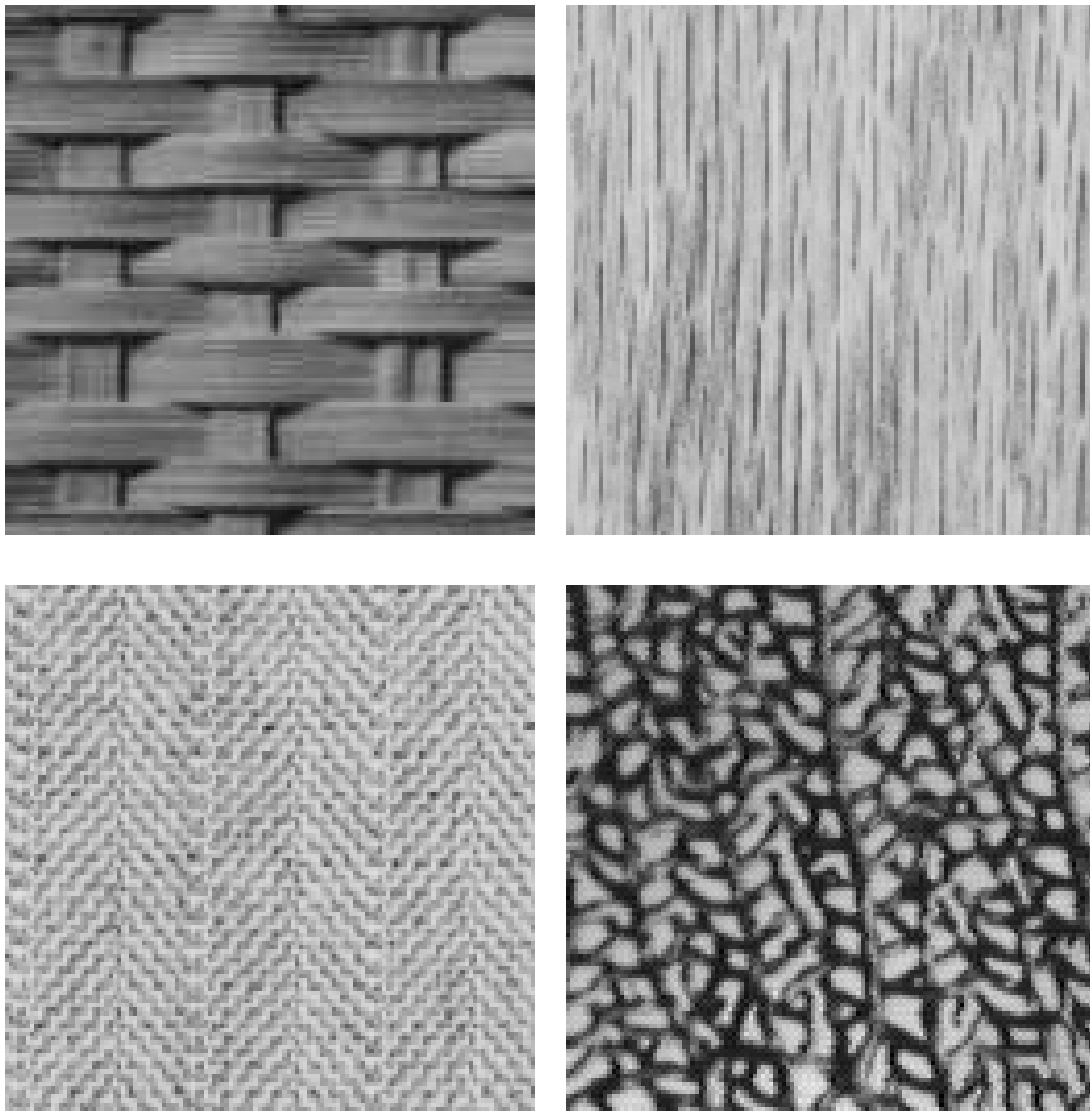


Figure 7.2: Exemples de textures naturelles.

Analyse simple d'une image en niveaux de gris

Par un exemple, nous allons introduire quelques paramètres permettant de caractériser au mieux une texture. Considérons l'image de la figure 7.3. Pour un observateur, il est bien évident que cette image comporte deux textures distinctes. La définition de descripteurs de texture a pour but de fournir des grandeurs liées à chaque texture et permettant ainsi leur identification ou leur reconnaissance.



Figure 7.3: Exemple d'image comportant deux textures.

Une étude sur l'image entière, ou étude globale de l'image, ne permettra pas de mettre en évidence les différentes textures présentes dans l'image. Par exemple, si nous calculons la transformée de FOURIER de l'image, il est très difficile de mettre en évidence les composantes fréquentielles relatives à telle ou telle texture de l'image. Il est donc préférable de découper l'image en sous-régions que l'on appelle généralement *fenêtres*. On effectuera donc un traitement local de l'image. Chaque fenêtre est étudiée séparément et peut être de forme et de taille quelconque. On peut, par exemple, prendre des blocs carrés de 13×13 pixels.

7.2 Caractérisation statistique des textures

Une première approche dans l'étude d'une texture est de calculer la moyenne des niveaux de gris des pixels appartenant à une fenêtre B . Ce n'est évidemment pas le seul moyen. Un autre paramètre qui peut nous renseigner sur une texture est le calcul de l'écart-type sur cette même fenêtre. La figure 7.4 montre des exemples de fenêtres de forme circulaire. Dans cette figure, l'intérieur des fenêtres a été remplacé, soit par la moyenne de l'image sur cette fenêtre (figure 7.4a), soit par la déviation standard de l'image sur la fenêtre (figure 7.4b).

On remarque que les valeurs moyennes des deux textures sont assez différentes, ce qui permet de les distinguer mais non de les caractériser complètement. La moyenne est le *descripteur statistique* de texture le plus simple que l'on puisse introduire. La variance est également un descripteur de texture. Dans l'exemple donné, les écart-type respectifs des deux textures sont égaux à 32 et à 66, ce qui ne permet pas de les distinguer facilement.

En général, la description des textures par la *moyenne* et l'*écart-type* ne suffit pas. En particulier, des textures différentes peuvent avoir même moyenne et même variance, comme le montre la figure 7.5. Il est donc nécessaire d'introduire d'autres descripteurs qui pourront mettre en évidence les propriétés des textures, comme la directivité, la rugosité, etc.

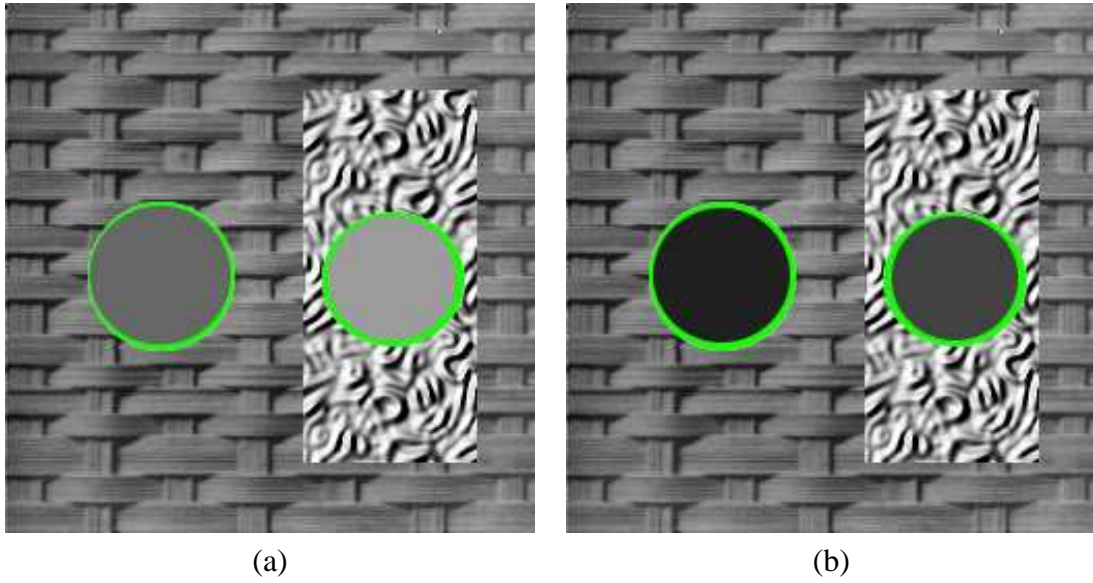


Figure 7.4: Illustration de fenêtres. (a) calcul de la moyenne des niveaux de gris (103 et 156 respectivement) (b) calcul de l'écart-type (32 et 66 respectivement).

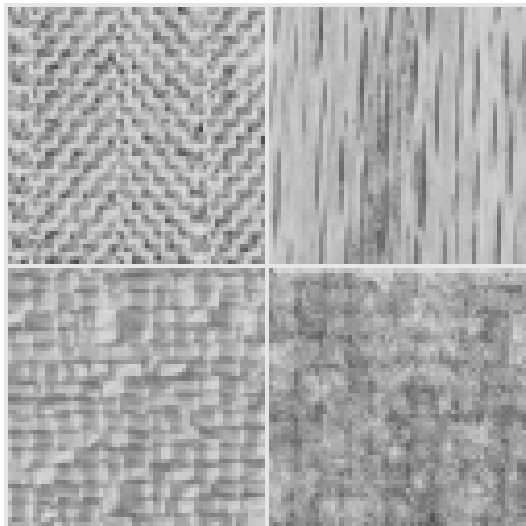


Figure 7.5: Textures ayant même moyenne et même variance.

Nous donnons à présent une définition plus précise de la notion de moyenne et d'écart-type sur un voisinage (ou fenêtre). Ensuite, nous parlerons d'une famille de descripteurs basés sur l'utilisation d'un histogramme local.

Considérons une image f contenant au moins une texture. Soit une fenêtre B complètement incluse dans cette texture; on évite ainsi des fenêtres qui seraient à cheval sur deux textures. Nous noterons $\#(B)$ le cardinal de B . L'étude de la texture se fait via l'étude de l'image f sur la fenêtre B de manière à fournir une information locale pertinente.

7.2.1 Moyenne locale

Définition 57 La moyenne locale sur la fenêtre B est définie par

$$\mu_f = \frac{1}{\#(B)} \sum_{(x,y) \in B} f(x,y) \quad (7.1)$$

Il convient de remarquer que, d'un point de vue statistique, la définition précédente n'est pas rigoureuse. En effet, la moyenne dépend de l'ensemble des valeurs possibles et de leur distribution à l'intérieur de la fenêtre. À ce titre, la définition précédente n'est pas réellement une moyenne mais l'estimation de cette moyenne sur base des échantillons. Pour la suite, nous parlerons de moyenne ou d'autres grandeurs statistiques sans faire cette distinction importante car, en pratique et en l'absence de modèle, les grandeurs statistiques sont toujours estimées par des expressions similaires à celle fournie ci-avant.

7.2.2 Écart-type local

Définition 58 L'écart-type sur la fenêtre B est défini par

$$\sigma_f = \sqrt{\frac{\sum_{(x,y) \in B} [f(x,y) - \mu_f]^2}{\#(B)}} \quad (7.2)$$

L'écart-type mesure l'énergie locale d'une texture.

7.2.3 Histogramme local

Commençons par définir la notion d'histogramme.

Définition 59 L'histogramme d'une image est la courbe représentant la fréquence des occurrences des niveaux de gris présents dans l'image.

Considérons l'image suivante:

```

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

```

comportant $L = 4$ niveaux de gris ($l = 0, 1, 2, 3$). Son histogramme est représenté à la figure 7.6. Pour le construire, il suffit de compter le nombre de pixels ayant le niveau de gris 0, 1, ... et de reporter cette valeur sur le graphique. On peut également diviser cette valeur par le nombre total de pixels de l'image pour obtenir la probabilité d'occurrence d'un niveau de gris. Un exemple d'histogramme pour une image en 256 niveaux de gris est également donné à la figure 7.7.

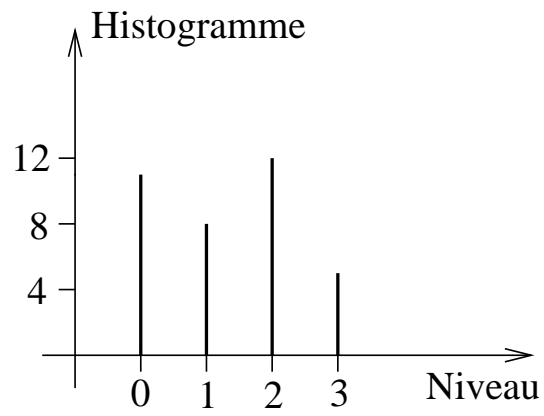


Figure 7.6: Histogramme (non normalisé) d'une image.

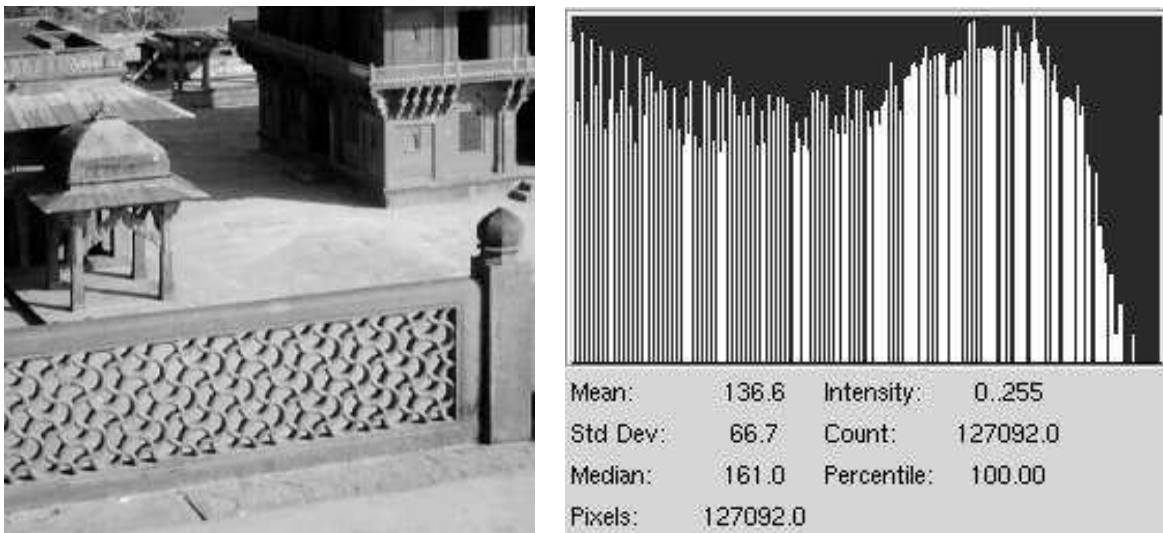


Figure 7.7: Une image et son histogramme (global car ici B correspond à l'image entière)

Définition 60 À partir d'une fenêtre B dans l'image, il est possible de définir un histogramme local normalisé $p(l)$ comme suit

$$p(l) = \frac{\#\{(x,y) \in B \mid f(x,y) = l\}}{\#(B)} \quad (7.3)$$

L'histogramme $p(l)$ représente donc la probabilité d'occurrence du niveau de gris l parmi les $\#(B)$ pixels de la fenêtre B . L'histogramme local ainsi défini peut directement servir de descripteur de texture mais on lui préfère d'autres mesures plus classiques permettant de le caractériser:

Moyenne

$$\mu_L = \sum_{l=0}^{L-1} l p(l) \quad (7.4)$$

où L est le nombre de niveaux de gris présent dans la fenêtre B . Cette formule fournit en fait un moyen simple de calculer la moyenne locale μ_f à partir de l'histogramme local. En effet, on remarquera que $\mu_f = \mu_L$.

Écart-type

$$\sigma_L = \sqrt{\sum_{l=0}^{L-1} (l - \mu_L)^2 p(l)} \quad (7.5)$$

Obliquité

$$S_s = \frac{1}{\sigma_L^3} \sum_{l=0}^{L-1} (l - \mu_L)^3 p(l) \quad (7.6)$$

“Kurtosis”

$$S_k = \frac{1}{\sigma_L^4} \sum_{l=0}^{L-1} (l - \mu_L)^4 p(l) - 3 \quad (7.7)$$

Énergie

$$S_E = \sum_{l=0}^{L-1} p(l)^2 \quad (7.8)$$

Entropie

$$S_H = - \sum_{l=0}^{L-1} p(l) \log_2 p(l) \quad (7.9)$$

Tous les paramètres qui viennent d'être introduits peuvent être utilisés comme descripteurs de la texture. Cependant, il n'existe pas de descripteur idéal. En raison du faible pouvoir discriminant des ces paramètres, il est nécessaire de combiner plusieurs descripteurs.

7.2.4 Matrice de co-occurrence des niveaux de gris

Les descripteurs statistiques décrits précédemment ne fournissent pas d'information concernant la nature périodique qu'une texture peut présenter. Ainsi, la texture de fond d'images (petits carrés obliques) de la figure 7.3 présente un caractère périodique qu'aucun descripteur introduit précédemment ne pourra mettre en lumière.

La matrice de co-occurrence introduite ci-après va nous offrir un premier moyen de caractériser la périodicité et la directivité des textures.

L'idée consiste à étudier la texture au moyen d'une fenêtre B d'observation de la texture en comptabilisant le nombre de paires de pixels distants de d qui présentent une différence Δz en niveaux de gris. En plus de la distance, on tient compte de la direction définie par toute paire de pixels.

Une matrice de co-occurrence est définie au moyen d'une relation géométrique R entre deux pixels (x_1, y_1) et (x_2, y_2) . Un exemple de relation géométrique simple est

$$x_2 = x_1 + 1 \quad (7.10)$$

$$y_2 = y_1 \quad (7.11)$$

pour lequel le pixel (x_2, y_2) est le pixel voisin de droite de (x_1, y_1) .

La matrice de co-occurrence $C_R(i, j)$ est carrée et de dimension $L \times L$, où L est le nombre de niveaux de gris présents dans B . Les indices de la matrice de co-occurrence sont donc les niveaux de gris de la texture étudiée. La matrice $C_R(i, j)$ se construit comme suit :

1. Initialisation de la matrice : $\forall i, j \in [0, L[: C_R(i, j) = 0$.
2. Remplissage de la matrice. Si la relation R entre deux pixels (x_1, y_1) et (x_2, y_2) est respectée:

$$C_R(f(x_1, y_1), f(x_2, y_2)) = C_R(f(x_1, y_1), f(x_2, y_2)) + 1$$

Dès lors, $C_R(i, j)$ contient le nombre de fois que l'on a rencontré dans B deux pixels (x_1, y_1) , (x_2, y_2) vérifiant la relation géométrique R et tel que $f(x_1, y_1) = i$ et $f(x_2, y_2) = j$. En pratique, on utilise des relations géométriques assez simples comme celle citée comme exemple. Ainsi, on définit une matrices orientées

$$P_{90^\circ, d}(i, j) = \# \{ (x_1, y_1), (x_2, y_2) \in B \mid x_1 = x_2, |y_2 - y_1| = d, f(x_1, y_1) = i \text{ et } f(x_2, y_2) = j \} \quad (7.12)$$

où 90° correspond à une direction verticale et d détermine la distance entre les deux pixels en question. De même, on définit

$$P_{0^\circ, d}(i, j) = \# \{ (x_1, y_1), (x_2, y_2) \in B \mid y_1 = y_2, |x_2 - x_1| = d, f(x_1, y_1) = i \text{ et } f(x_2, y_2) = j \} \quad (7.13)$$

Cette définition fournit une matrice symétrique –ce qui permet de ne stocker qu'une partie de la matrice.

Il est également possible de définir d'autres matrices correspondant aux directions 45° et 135° .

Exemple. Soit l'image suivante comportant $L = 4$ niveaux de gris ($l = 0, 1, 2, 3$) :

$$f(x, y) = \begin{matrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 2 & 2 & 3 \\ 2 & 2 & 3 & 3 \end{matrix} \quad (7.14)$$

Les matrices $P_{0^\circ,1}$ et $P_{90^\circ,1}$ sont donc de dimension 4 et valent

$$P_{0^\circ,1} = \begin{bmatrix} 6 & 2 & 1 & 0 \\ 2 & 2 & 0 & 0 \\ 1 & 0 & 4 & 2 \\ 0 & 0 & 2 & 2 \end{bmatrix} \quad P_{90^\circ,1} = \begin{bmatrix} 6 & 1 & 2 & 0 \\ 1 & 2 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \end{bmatrix} \quad (7.15)$$

On peut remarquer que les matrices de co-occurrence sont bien symétriques. D'autre part, la somme des valeurs des occurrences vaut 24 tant pour $P_{0^\circ,1}$ que pour $P_{90^\circ,1}$.

L'analyse de textures sur base de la matrice de co-occurrence est fondée sur l'occurrence répétée d'une configuration typique dans la texture; cette configuration varie rapidement en fonction de la distance pour des textures fines et lentement pour des textures plus grossière. Les matrices de co-occurrence sont généralement utilisées par des algorithmes automatiques, à des fins d'identification par exemple. Seuls des cas simples, comme celui décrit ci-après, permettent une interprétation intuitive du contenu d'une matrice de co-occurrence.

Considérons la texture suivante comportant deux niveaux de gris 0 et 1:

$$\begin{array}{cccccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{array} \quad (7.16)$$

composée de lignes noires (ou blanches !) horizontales et calculons les matrices de co-occurrence pour les directions horizontales et verticales ainsi que pour d valant 1 et 2 :

$$P_{0^\circ,1} = \begin{bmatrix} 30 & 0 \\ 0 & 30 \end{bmatrix} \quad P_{90^\circ,1} = \begin{bmatrix} 0 & 30 \\ 30 & 0 \end{bmatrix} \quad (7.17)$$

$$P_{0^\circ,2} = \begin{bmatrix} 24 & 0 \\ 0 & 24 \end{bmatrix} \quad P_{90^\circ,2} = \begin{bmatrix} 24 & 0 \\ 0 & 24 \end{bmatrix} \quad (7.18)$$

On peut donner une interprétation à ces résultats. Dans la direction horizontale (0°), on ne peut trouver que des transitions noir-blanc, les matrices $P_{0^\circ,d}$ ne peuvent donc être que diagonales. Pour les matrices $P_{90^\circ,d}$, on voit que l'on peut caractériser la taille de la texture grâce à la valeur de d .

Les matrices de co-occurrence permettent donc de caractériser la répétition de motifs dans une texture. Néanmoins, il faut remarquer que la taille d'une matrice de co-occurrence est proportionnelle au nombre de niveaux de gris de la texture. Pour $L = 256$, il faudrait donc des matrices carrées de taille 256×256 . De plus, il faut considérer une matrice par couple (direction, d), ce qui entraîne un nombre assez élevé de descripteurs de texture.

À partir des matrices de co-occurrence, on a créé toute une série de nouveaux descripteurs (en nombre restreint cette fois-ci) permettant de mettre en évidence les paramètres les plus importants de ces matrices de co-occurrence.

Probabilité maximale

C'est simplement la valeur maximale normalisée d'une des matrices ou de toutes les matrices de co-occurrence considérées.

Moments

Le moment d'ordre k d'une matrice de co-occurrence $C_R(i, j)$ est défini par

$$(m_R)_k = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-j) C_R(i, j) \quad (7.19)$$

Ce descripteur a de faibles valeurs lorsque les éléments diagonaux de C_R sont grands. L'effet inverse peut être obtenu grâce aux moments de la différence inverse.

Moments de la différence inverse

$$(m_R^{-1})_k = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{C_R(i, j)}{(i-j)^k}, i \neq j \quad (7.20)$$

Contraste

$$C(k, n) = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} |i-j|^k C_R(i, j)^n \quad (7.21)$$

Entropie

$$H = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} C_R(i, j) \log_2 C_R(i, j) \quad (7.22)$$

7.3 Caractérisation géométrique des textures

Dans notre étude des textures, nous n'avons étudié jusqu'ici que des outils statistiques, mettant en évidence des descripteurs du type moyenne, variance, etc. Nous allons voir que certaines textures présentant un réel caractère périodique peuvent être décrites de manière précise grâce, entre autres, à la transformée de FOURIER ou à des mesures d'énergie.

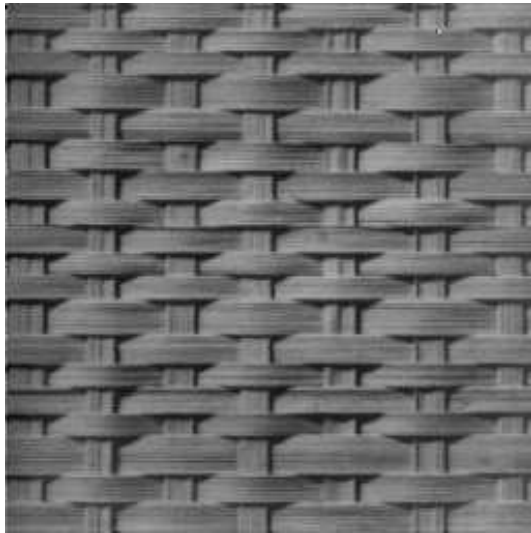
7.3.1 Approche fréquentielle

Pour mettre en évidence les caractéristiques fréquentielles d'une texture, il est nécessaire de calculer la transformée de FOURIER sur une image ne contenant que la texture, sous peine de voir apparaître dans le spectre des composantes fréquentielles qui n'ont rien à voir avec la texture étudiée. Un exemple de l'utilisation de la transformée de FOURIER est donnée à la figure 7.8.

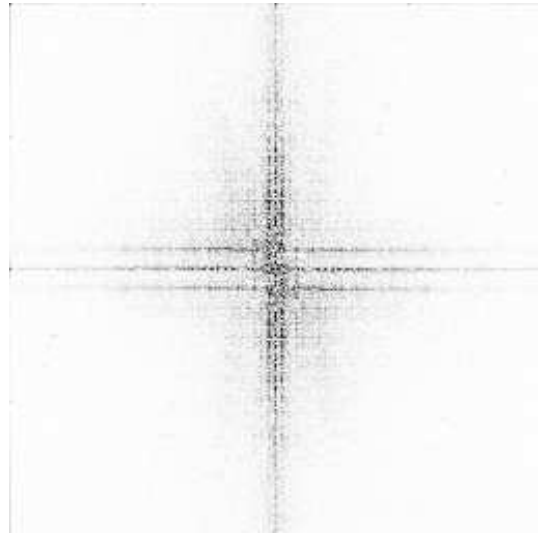
On voit clairement sur cette figure 7.8(b) l'apparition d'une structure dans le module de la transformée de FOURIER qui peut ainsi présenter différents pics. Ces pics nous informent sur l'orientation de la texture ainsi que de son importance (amplitude). Par contre, si on observe le spectre de la figure 7.8(d), on remarquera qu'il est assez aléatoire et ne permet de caractériser une véritable structure dans la texture.

Divers descripteurs peuvent être définis à partir de la transformée de FOURIER de la texture:

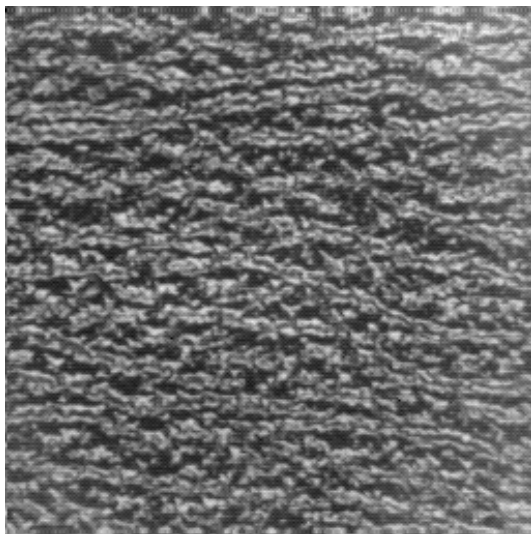
- le pic dominant dans le spectre donne la direction (ou orientation) principale de la texture,



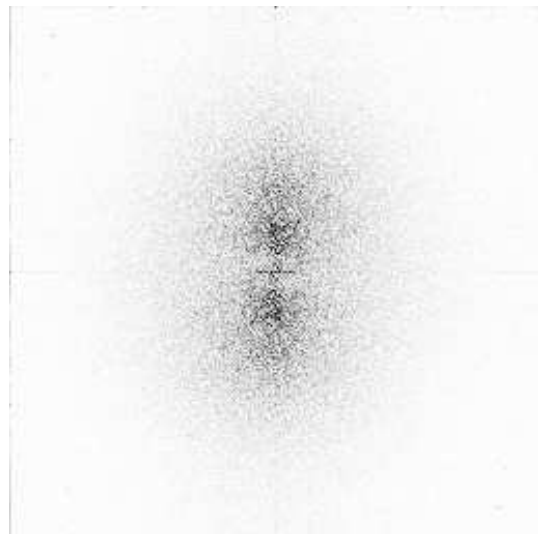
(a)



(b)



(c)



(d)

Figure 7.8: Caractéristiques fréquentielles d'une texture. Sur la gauche, les images texturées et sur la droite, le module de leur transformée de FOURIER respective.

- la localisation des pics permet de déterminer la période spatiale fondamentale de la texture

Il est également possible d'obtenir d'autres descripteurs. Ainsi, soit $\mathcal{F}(u, v)$ la transformée de FOURIER de la texture. À partir de cette fonction, nous pouvons déterminer le module de la transformée de FOURIER exprimé en coordonnées polaires: $\mathcal{S}(r, \theta)$. Pour θ fixé, $\mathcal{S}(r, \theta)$ peut être vue comme une fonction à une dimension et qui renseigne sur la texture pour une direction particulière. De même, pour r fixé, $\mathcal{S}(r, \theta)$ peut être vue comme une fonction à une dimension. Une description plus globale peut être obtenue en intégrant $\mathcal{S}(r, \theta)$ soit par rapport à r , soit par rapport à θ , ce qui nous permet de définir les deux fonctions suivantes :

$$\mathcal{S}_r(r) = \int_0^{2\pi} \mathcal{S}(r, \theta) d\theta \quad (7.23)$$

$$\mathcal{S}_\theta(\theta) = \int_0^{+\infty} \mathcal{S}(r, \theta) dr \quad (7.24)$$

Ces deux fonctions constituent une description énergétique spectrale de la texture. La fonction $\mathcal{S}_r(r)$ présente des maxima pour certaines fréquences caractéristiques correspondant à des distances (ou périodes) caractéristiques de la texture (sans tenir compte de l'orientation). Par contre, la fonction $\mathcal{S}_\theta(\theta)$ nous renseigne sur l'orientation de la texture en présentant des maxima aux angles (ou orientations) caractéristiques de la texture.

7.3.2 Textures et énergie

Dans l'étude des textures, LAWS a introduit différentes mesures permettant de caractériser l'énergie d'une texture.

Ces mesures déterminent les propriétés d'une texture à partir d'une estimation du profil de niveaux de gris en termes de *contours*, de *tacheté (spot)*, d'*ondulation* et de *vague*. Les mesures sont dérivées de trois vecteurs simples: $L_3 = (1, 2, 1)$ qui effectue la moyenne, $E_3 = (-1, 0, 1)$ qui détecte les contours et $S_3 = (-1, 2, -1)$ qui correspond à la dérivée seconde. Par convolution de ces vecteurs (symétriques) par eux-mêmes et entre eux, LAWS a introduit 9 masques de convolution:

$$\begin{array}{ccc} \frac{1}{36} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} & \frac{1}{12} \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} & \frac{1}{12} \begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix} \\ \text{Laws 1} & \text{Laws 2} & \text{Laws 3} \\ \\ \frac{1}{12} \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} & \frac{1}{4} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} & \frac{1}{4} \begin{bmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ 1 & -2 & 1 \end{bmatrix} \\ \text{Laws 4} & \text{Laws 5} & \text{Laws 6} \\ \\ \frac{1}{12} \begin{bmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{bmatrix} & \frac{1}{4} \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & -2 \\ -1 & 0 & 1 \end{bmatrix} & \frac{1}{4} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \\ \text{Laws 7} & \text{Laws 8} & \text{Laws 9} \end{array}$$

De même, il est possible de générer des vecteurs de taille 5:

$$L_5 = (1, 4, 6, 4, 1) \quad (7.25)$$

$$E_5 = (-1, -2, 0, 2, 1) \quad (7.26)$$

$$S_5 = (-1, 0, 2, 0, -1) \quad (7.27)$$

$$R_5 = (1, -4, 6, -4, 1) \quad (7.28)$$

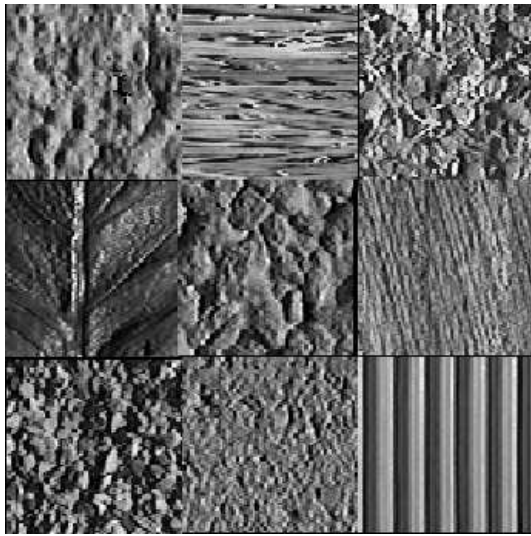
$$W_5 = (-1, 2, 0, -2, -1) \quad (7.29)$$

On peut aussi les combiner suivant les directions horizontales et verticales

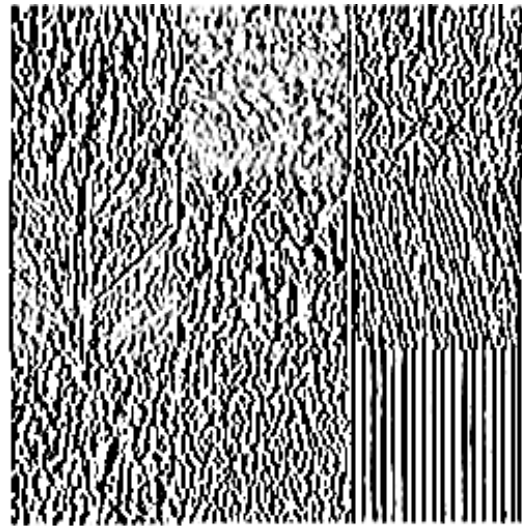
$$L_5^T \times S_5 = \begin{bmatrix} -1 & 0 & 2 & 0 & -1 \\ -4 & 0 & 8 & 0 & -4 \\ -6 & 0 & 12 & 0 & -6 \\ -4 & 0 & 8 & 0 & -4 \\ -1 & 0 & 2 & 0 & -1 \end{bmatrix} \quad (7.30)$$

La convolution d'une texture au moyen des masques de LAWS suivie d'un calcul d'une statistique d'énergie permet de dériver une information intéressante pour la description ou la discrimination de textures. La figure 7.9 montre des exemples de textures traitées par les masques de LAWS.

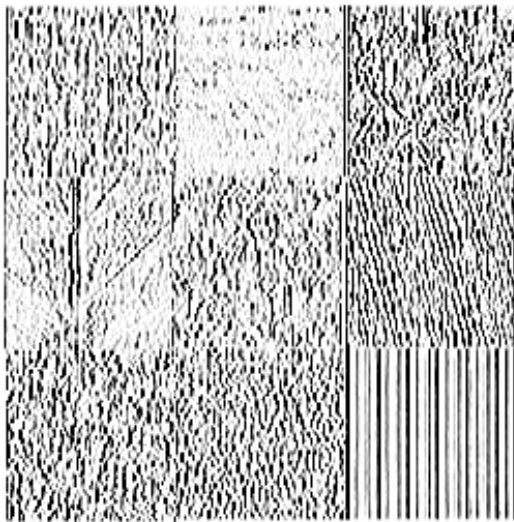
On remarque que le pouvoir discriminant des masques n'est pas identique. Par contre, la combinaison des résultats est à même de fournir une information suffisamment riche pour discriminer et identifier les textures. La combinaison des critères est bien souvent le seul moyen efficace d'aboutir. Il convient alors d'attirer l'attention sur le fait que les mesures n'ont pas toutes les mêmes dimensions physiques, certains sont des moyennes, d'autres des énergies ou des mesures adimensionnelles; leur combinaison est alors hasardeuse.



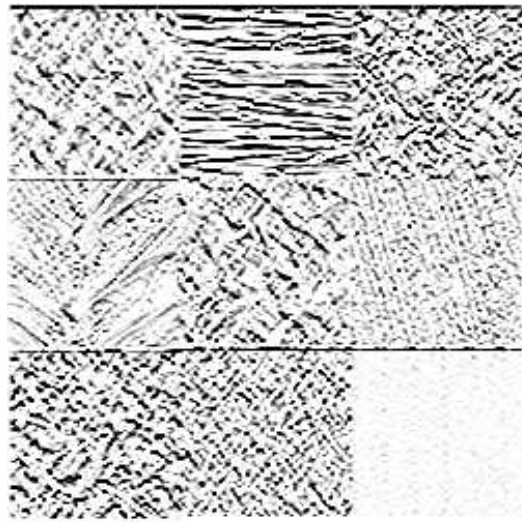
Textures



matrice 5×5 donnée en exemple



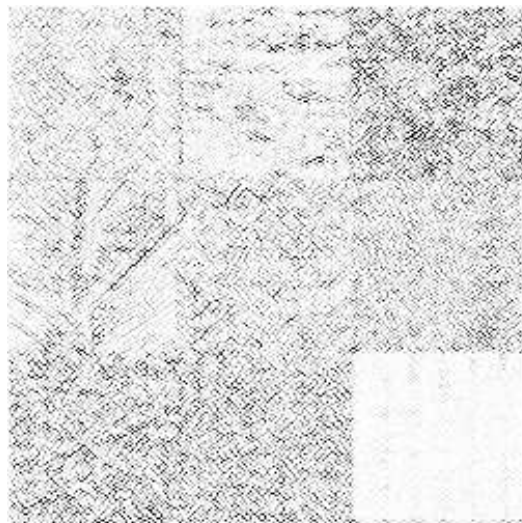
Laws 3



Laws 4



Laws 5



Laws 9

Figure 7.9: Illustration des mesures de LAWS (vidéo inverse)

Chapter 8

Rehaussement et restauration

8.1 Généralités

L'acquisition d'une image s'accompagne souvent d'une distorsion ou d'une certaine dégradation. On peut songer par exemple à un éclairage non uniforme, à des poussières sur une lentille ou encore à un flou créé par un mouvement lors de l'acquisition. Il n'y a dès lors pas d'autre solution que de chercher à corriger les défauts par un procédé algorithmique.

On distingue deux grandes familles de procédés de correction des défauts:

- le *rehaussement*, qui consiste à donner à l'image un aspect visuellement correct, et
- la *restauration*, qui cherche à rétablir la valeur exacte des pixels.

Dans les deux cas, on part à la recherche d'une estimation $\hat{f}(x,y)$ de l'image $f(x,y)$ qui soit visuellement ou analytiquement intéressante, c'est-à-dire qu'elle minimise une fonction d'erreur $e(x,y) = e(\hat{f}, f)$. Comme il n'existe pas de fonction d'erreur capable de modéliser la perception visuelle, le rehaussement est principalement de nature empirique. Ce n'est pas le cas de la restauration qui s'appuie sur une définition précise de la fonction d'erreur. Il s'agit par exemple de minimiser l'erreur quadratique moyenne

$$e(\hat{f}, f) = \frac{1}{MN} \sum_x \sum_y (\hat{f}(x,y) - f(x,y))^2 \quad (8.1)$$

Cette fonction est un choix commun car

- la fonction se calcule aisément,
- elle est différentiable, et
- elle correspond à l'énergie de l'erreur. Par la relation de PARSEVAL, l'énergie correspond également à l'énergie de la transformée de FOURIER.

La figure 8.1 illustre la marge de manœuvre dont on dispose pour du rehaussement par rapport à la restauration. L'image 8.1(b) a été obtenue en divisant, par division entière, l'image originale par 5 puis en multipliant toutes les valeurs par 5; c'est en fait un arrondi au plus petit multiple de 5 inférieur à chaque valeur. Les images (c) et (d) résultent d'approximations plus grossières encore.

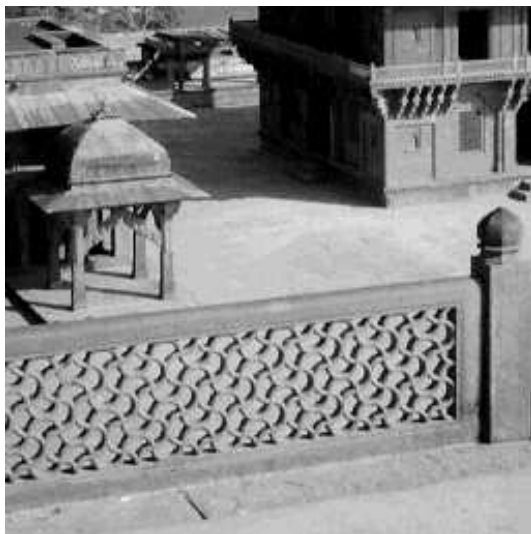
On voit clairement qu'un arrondi à 10 n'est pas gênant, ce qui revient à dire que les trois bits de poids faible sont peu visibles. Au-delà, l'erreur de quantification devient perceptible.



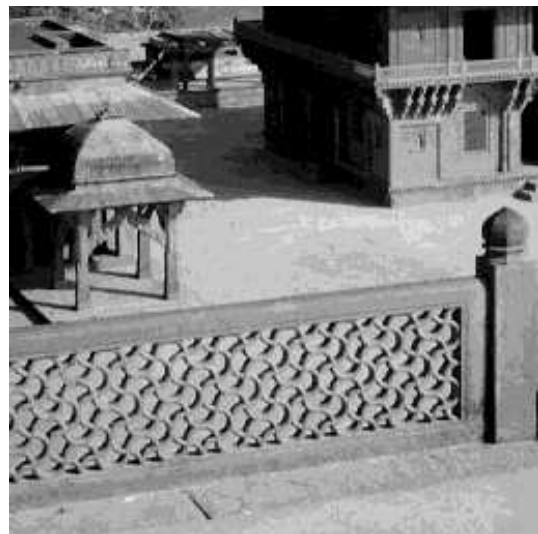
(a) Image originale f



(b) Arrondi à 5



(c) Arrondi à 10



(d) Arrondi à 20

Figure 8.1: Seuil de visibilité pour une opération d'arrondi.

8.2 Rehaussement

8.2.1 Définition

Soit L la valeur maximale de l'image $f(x,y)$. De plus, prenons une image positive, c'-à-d. que $\forall x,y, f(x,y) \geq 0$. En définitive, $f(x,y) \in [0,L[$.

Définition 61 Le rehaussement consiste à appliquer une fonction croissante $g()$ à toutes les valeurs l de l'image telle que $g(0) = 0$ et $g(L) = L$.

Le type de rehaussement dépend essentiellement de la pente de $g()$: les plages des valeurs où la pente est forte seront rehaussées. On impose que la fonction $g()$ soit croissante, de sorte que toute relation d'ordre entre pixels traduite par leur valeur soit maintenue, mais elle peut être de forme quelconque.

8.2.2 Rehaussement par fonction croissante

On distingue généralement trois zones à rehausser; elles correspondent aux faibles valeurs, aux valeurs moyennes et aux fortes valeurs. Si l'on souhaite rehausser les valeurs de la plage inférieure, on peut utiliser une fonction parabolique de la forme suivante

$$g(l) = l + \alpha l(L-l) \quad (8.2)$$

La figure 8.2(a) montre cette courbe pour $\alpha = 0,003$. Clairement, les valeurs des deux extrémités sont peu affectées par le rehaussement. Tout l'effet se concentre dans la partie centrale.

Un autre type de courbes consiste à appliquer une forme de S dans la partie centrale

$$g(l) = \frac{L}{2} \left(1 + \frac{\sin \left[\alpha \pi \left(\frac{l}{L} - \frac{1}{2} \right) \right]}{\sin \left[\alpha \frac{\pi}{2} \right]} \right), \quad \alpha \in]0, 1[\quad (8.3)$$

comme représenté à la figure 8.2(b) pour $\alpha = 0,95$. Cette dernière rehausse les valeurs centrales.

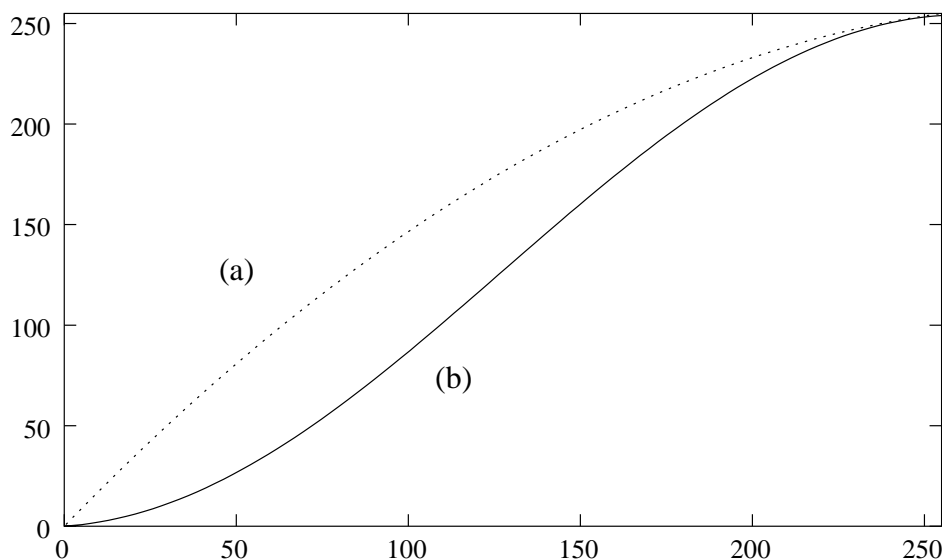


Figure 8.2: Différentes courbes de rehaussement.

8.2.3 Rehaussement par accentuation des contours

Une technique couramment utilisée en photographie consiste à accentuer les contours. Il faut donc d'abord localiser les contours pour adapter l'image par la suite. Par exemple, le laplacien convient parfaitement pour localiser les contours. Pour accentuer les contours, il suffit de soustraire les valeurs du laplacien, avec un éventuel facteur correctif. On a donc

$$\widehat{f}(x,y) = f(x,y) - \alpha \nabla^2 f(x,y) \quad (8.4)$$

La figure 8.3 montre le résultat de l'opérateur d'estimation pour $\alpha = \frac{1}{3}$. Très clairement, les contours ont été accentués.

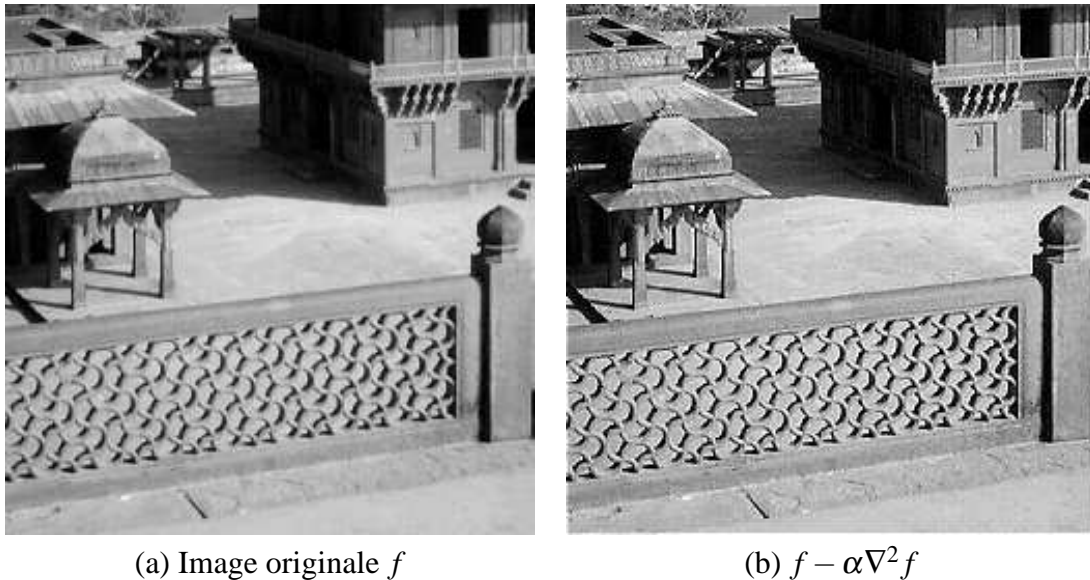


Figure 8.3: Rehaussement des contours.

Implémentation

Pour $\alpha = 1$, il suffit d'implémenter le filtre suivant (qui revient à combiner les masques respectifs):

$$\begin{array}{ccc} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{array} \quad (8.5)$$

8.2.4 Manipulation de l'histogramme

Soit $o(x,y)$ l'image de sortie. On cherche à déterminer le lien entre l'histogramme $hist_o(l)$ de cette image avec celui de l'image d'entrée f . D'une manière générale, pour une valeur l_o de l'image de sortie, il existe une fonction ϕ telle que

$$l_o = \phi(l_f) \quad (8.6)$$

Si cette fonction admet un inverse, ce qui est le cas si l'on impose la croissance et l'existence d'une dérivée non nulle,

$$l_f = \phi^{-1}(l_o) \quad (8.7)$$

Si l'on interprète l'histogramme comme une densité de probabilité, on arrive à une égalité entre les "masses" que sont

$$hist_o(l_o)dl_o = hist_f(l_f)dl_f \quad (8.8)$$

Dès lors,

$$hist_o(l_o) = \frac{hist_f(l_f)}{dl_o/dl_f} \quad (8.9)$$

$$= \frac{hist_f(l_f)}{d\phi(l_f)/dl_f} \Big|_{l_f=\phi^{-1}(l_o)} \quad (8.10)$$

et donc

$$hist_o(l) = \frac{hist_f(\phi^{-1}(l))}{\frac{d\phi}{dl}(\phi^{-1}(l))} \quad (8.11)$$

Exemple: fonction parabolique

Considérons la fonction ϕ suivante

$$l_o = \phi(l_f) = \frac{1}{L}(l_f)^2 \quad (8.12)$$

d'où également $l_f = \sqrt{Ll_o}$. Par application de la formule 8.11,

$$hist_o(l) = \frac{hist_f(\sqrt{Ll})}{\frac{2}{L}\sqrt{Ll}} = \frac{hist_f(\sqrt{Ll})}{2\sqrt{l/L}} \quad (8.13)$$

Égalisation de l'histogramme

Dans une image réelle, tous les niveaux ne sont pas présents avec une même occurrence. Il en résulte donc des disparités dans l'histogramme.

L'égalisation de l'histogramme d'une image est une méthode courante de rehaussement. Elle vise à assurer une distribution homogène des valeurs dans la totalité de la plage dynamique des valeurs possibles. Il s'agit donc d'une distorsion de l'échelle des valeurs. Formellement, pour toute valeur l de l'histogramme résultat, le nombre de pixels devrait valoir $\frac{A(D_f)}{L}$, où $A(D_f)$ représente l'aire du support de définition de la fonction f . L'équation 8.11 devient donc

$$\frac{d\phi}{dl}(l) = \frac{L}{A(D_f)} hist_f(l) \quad (8.14)$$

Par intégration des membres de cette dernière relation,

$$\phi(l) = \frac{L}{A(D_f)} \int_0^l hist_f(k)dk \quad (8.15)$$

Comme on peut estimer la probabilité d'un niveau par le rapport de la fréquence à l'aire du domaine de définition de la fonction

$$p(l) = \frac{hist_f(l)}{A(D_f)} \quad (8.16)$$

De plus, la fonction de répartition F est l'intégrale cumulée des probabilités

$$F(l) = \int_0^l p(k)dk = \frac{1}{A(D_f)} \int_0^l hist_f(k)dk \quad (8.17)$$

Dès lors,

$$\phi(l) = LF(l) \quad (8.18)$$

Autrement dit, l'image $o(x,y)$ traitée par la fonction qui effectue une égalisation de l'histogramme vaut

$$o(x,y) = LF(f(x,y)) \quad (8.19)$$

En raison des valeurs quantifiées, l'égalisation de l'histogramme peut fournir un histogramme non uniforme. Il ne faut pas s'en étonner. Ce phénomène est illustré à la figure 8.4.

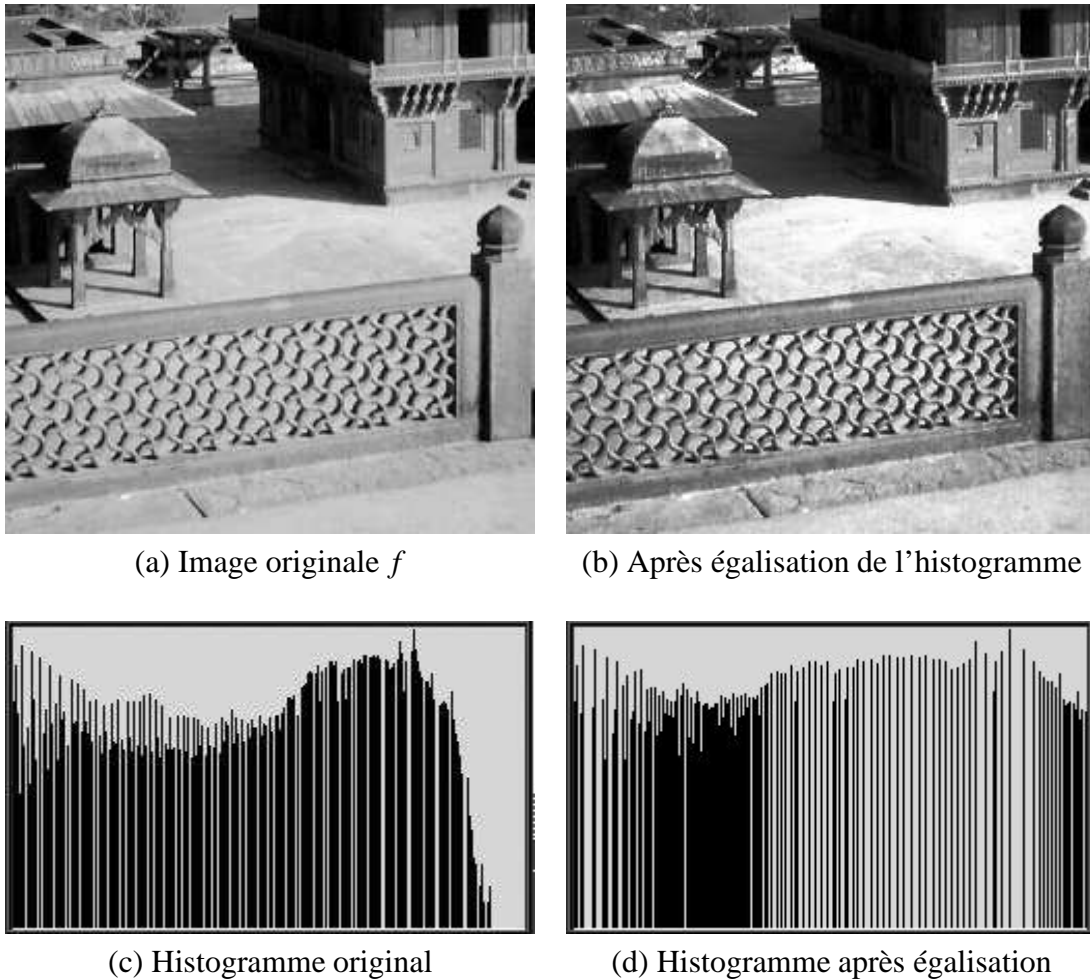


Figure 8.4: Égalisation d'histogramme.

8.2.5 Filtrage homomorphique

Une image peut être vue comme une fonction à deux dimensions $f(x,y)$ dont la valeur en un point représente l'intensité de l'image en ce point. Les images que l'on perçoit dans la vie courante sont obtenues par la lumière réfléchiée par les différents objets inclus dans la scène observée. La nature de $f(x,y)$ peut donc être caractérisée par deux composantes:

1. l'*illumination* $i(x,y)$ qui représente la quantité de lumière incidente qui illumine la scène en question.
2. la *réflectance* $r(x,y)$ qui représente la quantité de lumière réfléchiée par les objets de la scène.

L'illumination et la réflectance permettent de modéliser l'image $f(x,y)$ par le produit suivant:

$$f(x,y) = i(x,y)r(x,y) \quad (8.20)$$

où

$$0 < i(x,y) < +\infty$$

$$0 < r(x,y) < 1$$

La nature de $i(x,y)$ est déterminée par les sources lumineuses tandis que celle de $r(x,y)$ est déterminée par les caractéristiques des objets de la scène.

Le modèle qui vient d'être introduit peut être utilisé pour améliorer l'apparence d'une image en réalisant simultanément une compression de la dynamique de luminosité et une augmentation du contraste. L'expression 8.20 ne permet pas de réaliser un traitement séparé sur l'illumination et la réflectance. En effet, la transformée de FOURIER de f correspondant à la convolution des transformées de FOURIER respectives de i et de r , un filtrage linéaire simple ne peut pas être réalisé.

Afin de contourner ce problème, considérons la fonction

$$z(x,y) = \ln(f(x,y)) = \ln(i(x,y)) + \ln(r(x,y)) \quad (8.21)$$

La fonction $z(x,y)$ étant constituée d'une somme de deux fonctions, il est maintenant intéressant d'utiliser la propriété de linéarité de la transformée de FOURIER:

$$\mathcal{Z}(u,v) = \mathcal{F}\{\ln(i(x,y))\} + \mathcal{F}\{\ln(r(x,y))\} \quad (8.22)$$

$$= \mathcal{I}(u,v) + \mathcal{R}(u,v) \quad (8.23)$$

où $\mathcal{I}(u,v)$ et $\mathcal{R}(u,v)$ sont respectivement les transformées de FOURIER de $\ln(i(x,y))$ et $\ln(r(x,y))$. Il est possible de filtrer la fonction $\mathcal{Z}(u,v)$ en la multipliant par une fonction de transfert $\mathcal{H}(u,v)$, ce qui donne

$$\mathcal{S}(u,v) = \mathcal{H}(u,v)\mathcal{Z}(u,v) \quad (8.24)$$

$$= \mathcal{H}(u,v)\mathcal{I}(u,v) + \mathcal{H}(u,v)\mathcal{R}(u,v) \quad (8.25)$$

Si nous introduisons les deux fonctions

$$i'(x,y) = \mathcal{F}^{-1}\{\mathcal{H}(u,v)\mathcal{I}(u,v)\} \quad (8.26)$$

et

$$r'(x,y) = \mathcal{F}^{-1}\{\mathcal{H}(u,v)\mathcal{R}(u,v)\} \quad (8.27)$$

le résultat du filtrage de $z(x,y)$ sera noté $s(x,y)$ et sera égal à

$$s(x,y) = \mathcal{F}^{-1}\{\mathcal{S}(u,v)\} \quad (8.28)$$

$$= i'(x,y) + r'(x,y) \quad (8.29)$$

Vu que $z(x,y)$ avait été obtenue en calculant le logarithme népérien de $f(x,y)$, l'opération inverse va nous fournir l'image rehaussée $g(x,y)$ suivante:

$$g(x,y) = e^{s(x,y)} \quad (8.30)$$

$$= e^{i'(x,y)} e^{r'(x,y)} \quad (8.31)$$

$$= i''(x,y)r''(x,y) \quad (8.32)$$

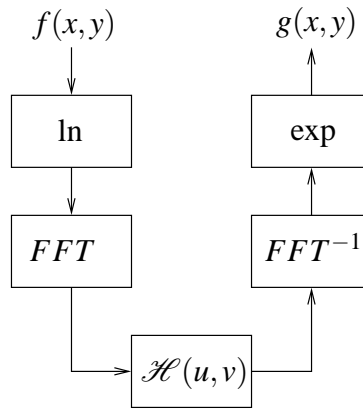


Figure 8.5: Schéma du filtrage homomorphique.

moyennant les notations $e^{i'(x,y)} = i''(x,y)$ et $e^{r'(x,y)} = r''(x,y)$ qui sont respectivement l'illumination et la réflectance de l'image rehaussée.

La méthode de rehaussement décrite ici est représentée à la figure 8.5. Cette méthode est un cas particulier d'une classe de systèmes connus sous le nom de système *homomorphique*.

L'illumination d'une image est en général une fonction qui présente de faibles variations tandis que la réflectance est plutôt une fonction qui varie de manière abrupte, particulièrement à la jonction d'objets de caractéristiques différentes. Ces observations tendent à associer les basses fréquences à la luminance et les hautes fréquences à la réflectance. Bien que ces associations ne soient qu'approximatives, elles peuvent être utilisées pour rehausser l'image.

Les spécifications du filtre homomorphique $\mathcal{H}(u,v)$ vont être établies de telle sorte qu'il affecte différemment les hautes et les basses fréquences. La figure 8.6 montre la coupe radiale d'un filtre homomorphique qui peut être utilisé. L'atténuation des basses fréquences va faire diminuer la dynamique de la luminosité tandis que l'amplification des hautes fréquences va augmenter le contraste.

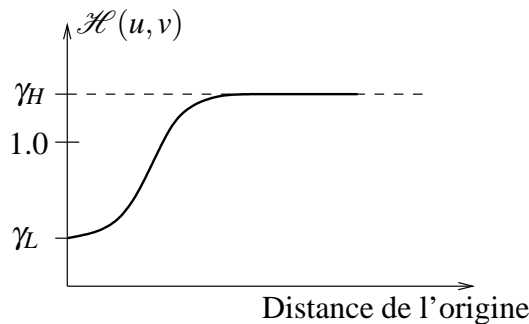


Figure 8.6: Coupe radiale d'un filtre homomorphique.

Un exemple d'application de filtre homomorphique est montré à la figure 8.7.

8.3 Restauration

Le problème de la restauration d'images est de corriger les défauts d'une image dégradée par l'ajout de bruit ou par une convolution (due par exemple à une défocalisation de la caméra d'acquisition). Le schéma de la figure 8.9 donne un modèle de système de restauration d'images.

Sur cette figure, $f(x,y)$ représente l'image sans défaut. Les dégradations sont dues à la convolution par la réponse impulsionnelle $h(x,y)$ et l'ajout d'un bruit $n(x,y)$. La fonction $w(x,y)$ représente



Figure 8.7: Illustration du filtre homomorphique. (a) image à rehausser (b) image rehaussée par le filtre homomorphique dont la coupe radiale est montrée à la figure 8.8.

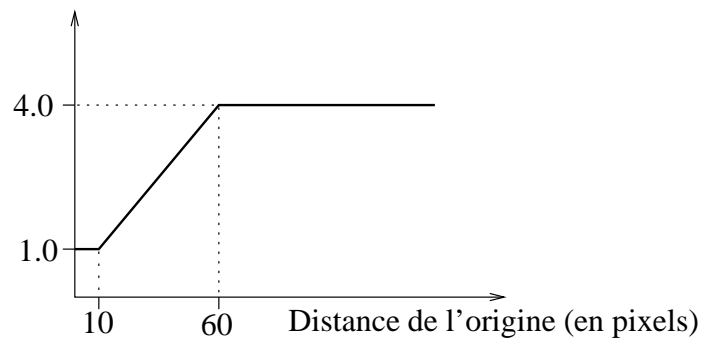


Figure 8.8: Exemple de coupe radiale d'un filtre homomorphique.

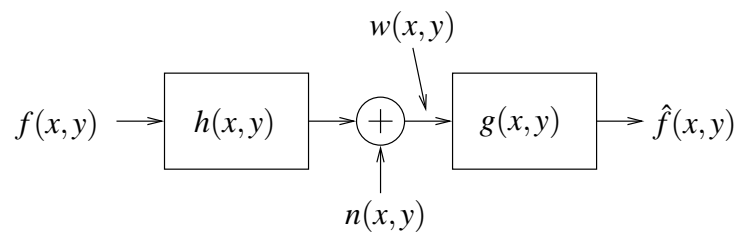


Figure 8.9: Modèle pour la restauration d'une image continue.

donc une version dégradée de l'image $f(x,y)$. Nous allons tenter de restaurer l'image $w(x,y)$ par un procédé de filtrage linéaire, via le filtre de réponse impulsionnelle $g(x,y)$. Tout le problème ici est de déterminer $g(x,y)$ de telle sorte que l'image restaurée $\hat{f}(x,y)$ soit la plus proche possible de $f(x,y)$.

8.3.1 Filtrage de WIENER

Le filtre de WIENER est un filtre linéaire souvent utilisé pour la réduction de bruit. Dès lors, dans cette section, nous considérons que l'image originale n'a subi comme dégradation que l'ajout d'un bruit aléatoire. Afin de simplifier les développements mathématiques, nous nous limiterons au cas d'une fonction à une dimension. Ce résultat sera alors aisément généralisé à deux dimensions.

La figure 8.10 illustre le modèle que nous allons utiliser. $w(t)$ représente la fonction dégradée par l'ajout du bruit $n(t)$ tandis que $\hat{f}(t)$ représente la fonction restaurée au départ de $w(t)$ par une convolution avec le filtre de réponse impulsionnelle $h(t)$. Idéalement, nous aimerions obtenir $\hat{f}(t) = f(t)$ mais en général, un filtre linéaire n'est pas assez puissant pour retrouver le signal de départ $f(t)$ de manière exacte. Le but est alors de déterminer la fonction $h(t)$ qui rende $\hat{f}(t)$ le plus proche de $f(t)$. $\hat{f}(t)$ est appelée *estimation* de $f(t)$.

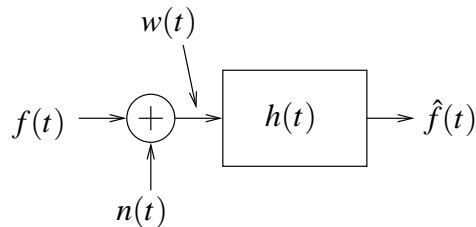


Figure 8.10: Modèle utilisé pour le filtre de WIENER.

Dans cette étude, nous ferons l'hypothèse que les signaux $f(t)$ et $n(t)$ sont des processus aléatoires ergodiques et de densité spectrale de puissance $\gamma_f(f)$ et $\gamma_n(f)$ connues.

Définition d'un critère

Avant de déterminer le filtre optimal, que l'on notera $h_o(t)$, il est nécessaire de définir un critère d'optimalité. Vu que $\hat{f}(t) = f(t)$ n'est pas réalisable, nous allons utiliser comme critère l'erreur quadratique moyenne.

Quelque soit l'expression de $h(t)$, nous obtiendrons en sortie un signal $\hat{f}(t)$ en réponse à $f(t)$. Le signal d'erreur à la sortie du filtre est défini par

$$e(t) = f(t) - \hat{f}(t) \quad (8.33)$$

qui mesure l'écart instantané entre la fonction restaurée et la fonction idéale. Comme mesure de l'erreur moyenne, nous utiliserons l'erreur quadratique moyenne définie par

$$E \{e^2(t)\} = \int_{-\infty}^{+\infty} e^2(t) dt \quad (8.34)$$

Ici, l'espérance mathématique peut se calculer comme une intégrale sur le temps vu que les processus sont supposés ergodiques. On remarquera que l'erreur quadratique moyenne pénalise plus les grands écarts entre $f(t)$ et $\hat{f}(t)$ comparés aux faibles écarts, ce qui constitue un critère tout à fait satisfaisant.

Erreur quadratique moyenne

Le problème se ramène à présent à déterminer une fonction $h(t)$ qui minimise l'erreur quadratique moyenne $E \{e^2(t)\}$, qui est un nombre. Il s'agit en fait de la minimisation d'une fonctionnelle qui relève du calcul des variations. Nous allons suivre la démarche suivante:

1. exprimer l'erreur quadratique moyenne en terme de la fonction $h(t)$,
2. déterminer l'expression optimale $h_o(t)$ en terme des densités spectrales de puissance connues, et finalement
3. calculer la valeur de l'erreur quadratique moyenne lorsque le filtre optimal $h_o(t)$ est utilisé.

Nous commençons par développer l'erreur quadratique moyenne de la manière suivante

$$E \{e^2(t)\} = E \{f^2(t) - 2f(t)\hat{f}(t) + \hat{f}^2(t)\} \quad (8.35)$$

$$= E \{f^2(t)\} - 2E \{f(t)\hat{f}(t)\} + E \{\hat{f}^2(t)\} \quad (8.36)$$

$$= T_1 + T_2 + T_3 \quad (8.37)$$

où T_1 , T_2 et T_3 sont trois termes que l'on va traiter séparément. T_1 se ramène à

$$T_1 = E \{f^2(t)\} = \int_{-\infty}^{+\infty} f^2(t)dt = \Gamma_{ff}(0) \quad (8.38)$$

où $\Gamma_{ff}(0)$ est la fonction d'autocorrélation de $f(t)$ évalué en 0. T_1 correspond donc à la puissance du signal $f(t)$. Le second membre T_2 peut s'écrire

$$T_2 = -2E \left\{ f(t) \int_{-\infty}^{+\infty} h(\tau)w(t-\tau)d\tau \right\} \quad (8.39)$$

où nous avons remplacé $\hat{f}(t)$ par le produit de convolution entre $w(t)$ et $h(t)$. Vu que l'espérance mathématique est une intégrale sur le temps, il est possible d'inverser l'ordre d'intégration:

$$T_2 = -2 \int_{-\infty}^{+\infty} h(\tau)E \{f(t)w(t-\tau)\}d\tau \quad (8.40)$$

$$= -2 \int_{-\infty}^{+\infty} h(\tau)\Gamma_{fw}(\tau)d\tau \quad (8.41)$$

où $\Gamma_{fw}(\tau)$ est la fonction d'intercorrélacion de $f(t)$ et $w(t)$. Le troisième terme T_3 peut être traité d'une manière similaire

$$T_3 = E \left\{ \int_{-\infty}^{+\infty} h(\tau)w(t-\tau)d\tau \int_{-\infty}^{+\infty} h(u)w(t-u)du \right\} \quad (8.42)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau)h(u)E \{w(t-\tau)w(t-u)\}d\tau du \quad (8.43)$$

Si l'on écrit le changement de variables $v = t - u$ dans l'espérance, nous obtenons

$$E \{w(t-\tau)w(t-u)\} = E \{w(v+u-\tau)w(v)\} = \Gamma_{ww}(u-\tau) \quad (8.44)$$

et donc

$$T_3 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau)h(u)\Gamma_{ww}(u-\tau) d\tau du \quad (8.45)$$

Dès lors, l'erreur quadratique moyenne peut maintenant s'écrire

$$E \{e^2(t)\} = \Gamma_{ff}(0) - 2 \int_{-\infty}^{+\infty} h(\tau)\Gamma_{fw}(\tau) d\tau + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\tau)h(u)\Gamma_{ww}(u-\tau) d\tau du \quad (8.46)$$

Étant donné que les fonctions d'autocorrélation et d'intercorrélacion sont connues, on remarque que l'erreur quadratique moyenne est bien une fonction de $h(t)$.

Minimisation de l'erreur quadratique moyenne

Considérons le filtre optimal $h_o(t)$. Tout autre filtre peut s'écrire comme

$$h(t) = h_o(t) + g(t) \quad (8.47)$$

où $h(t)$ est une réponse impulsionnelle arbitraire sous-optimale et $g(t)$ est définie comme l'écart instantané entre le filtre optimal $h_o(t)$ et le filtre sous-optimale $h(t)$. Si nous introduisons cette dernière expression dans l'expression de l'erreur quadratique moyenne, nous obtenons

$$E \{e^2(t)\} = \Gamma_{ff}(0) - 2 \int_{-\infty}^{+\infty} (h_o(\tau) + g(\tau))\Gamma_{fw}(\tau) d\tau \quad (8.48)$$

$$+ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (h_o(\tau) + g(\tau))(h_o(u) + g(u))\Gamma_{ww}(u-\tau) d\tau du \quad (8.49)$$

ce qui peut s'écrire sous la forme de sept termes:

$$E \{e^2(t)\} = \Gamma_{ff}(0) - 2 \int_{-\infty}^{+\infty} h_o(\tau)\Gamma_{fw}(\tau) d\tau \quad (8.50)$$

$$+ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h_o(\tau)h_o(u)\Gamma_{ww}(u-\tau) d\tau du \quad (8.51)$$

$$+ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h_o(\tau)g(u)\Gamma_{ww}(u-\tau) d\tau du \quad (8.52)$$

$$+ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h_o(u)g(\tau)\Gamma_{ww}(u-\tau) d\tau du \quad (8.53)$$

$$- 2 \int_{-\infty}^{+\infty} g(\tau)\Gamma_{fw}(\tau) d\tau + \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\tau)g(u)\Gamma_{ww}(u-\tau) d\tau du \quad (8.54)$$

Les trois premiers termes correspondent à l'erreur quadratique moyenne dans le cas du filtre optimal $h_o(t)$. On notera cette valeur E_o . Vu que la fonction d'autocorrélation $\Gamma_{ww}(\tau)$ est paire, les troisième et quatrième termes sont égaux. Nous pouvons alors les combiner avec le sixième terme pour obtenir

$$E \{e^2(t)\} = E_o + 2 \int_{-\infty}^{+\infty} g(u) \left[\int_{-\infty}^{+\infty} h_o(\tau)\Gamma_{ww}(u-\tau) d\tau - \Gamma_{fw}(u) \right] du \quad (8.55)$$

$$+ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\tau)g(u)\Gamma_{ww}(u-\tau) d\tau du \quad (8.56)$$

$$= E_o + T_4 + T_5 \quad (8.57)$$

où T_4 et T_5 sont à nouveau deux termes que nous allons traiter séparément. Montrons tout d'abord que le terme T_5 est toujours positif. En exprimant la fonction d'autocorrélation par une intégrale, nous obtenons

$$T_5 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\tau)g(u) \int_{-\infty}^{+\infty} w(t-\tau)w(t-u)dt d\tau du \quad (8.58)$$

$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(t-u)g(u)du \int_{-\infty}^{+\infty} g(\tau)w(t-\tau)d\tau dt \quad (8.59)$$

Si l'on définit $z(t) = g(t) \otimes w(t)$, ce terme vaut

$$T_5 = \int_{-\infty}^{+\infty} z^2(t)dt \geq 0 \quad (8.60)$$

qui ne peut jamais être négatif. Remarquons que cette valeur est indépendante de l'expression du filtre idéal $h_o(t)$.

Retournons à présent à l'expression de l'erreur quadratique moyenne:

$$E \{e^2(t)\} = E_o + 2 \int_{-\infty}^{+\infty} g(u) \left[\int_{-\infty}^{+\infty} h_o(\tau)\Gamma_{ww}(u-\tau) d\tau - \Gamma_{fw}(\tau) \right] du + T_5 \quad (8.61)$$

Nous allons à présent établir la condition qui assure que E_o soit la plus petite erreur quadratique moyenne que l'on puisse obtenir. Un moyen est de faire en sorte que le terme entre crochets soit nul pour toute valeur de u . Il faut alors assurer qu'il s'agit d'une condition nécessaire et suffisante pour optimiser le filtre. La condition est tout d'abord nécessaire. En effet, si le terme entre crochets était non nul pour certaines valeurs de u , le terme T_4 pourrait prendre des valeurs fortement négatives vu que $g(u)$ est une fonction totalement arbitraire. Il en découlerait que $E \{e^2(t)\}$ pourrait être inférieure à E_o , ce qui violerait notre définition. Nous obtenons donc une condition nécessaire pour que le filtre $h_o(t)$ soit optimal:

$$\Gamma_{fw}(\tau) = \int_{-\infty}^{+\infty} h_o(\tau)\Gamma_{ww}(u-\tau) d\tau \quad (8.62)$$

Cette condition est de plus suffisante. En effet, il suffit qu'elle soit vérifiée pour que le terme T_4 disparaisse et dès lors $E \{e^2(t)\}$ sera toujours supérieure ou égale à E_o , vu que $T_5 \geq 0$.

Pour les systèmes linéaires, il est aisé de montrer que

$$\Gamma_{\hat{f}w}(\tau) = h(\tau) \otimes \Gamma_{ww}(\tau) \quad (8.63)$$

où $\hat{f}(t)$ et $w(t)$ sont respectivement la sortie et l'entrée du système caractérisé par $h(t)$. Étant donnée la condition d'optimalité sur $h_o(t)$, nous pouvons écrire

$$\Gamma_{fw}(\tau) = h_o(\tau) \otimes \Gamma_{ww}(\tau) = \Gamma_{\hat{f}w}(\tau)$$

Cette seconde égalité montre que le filtre de WIENER rend l'intercorrélacion entre le signal idéal et le signal dégradé égal à l'intercorrélacion entre le signal dégradé et le signal restauré. En prenant la transformée de FOURIER de chacun des membres de l'égalité, nous obtenons

$$\gamma_{fw}(f) = \mathcal{H}_o(f)\gamma_{ww}(f) \quad (8.64)$$

et dès lors,

$$\mathcal{H}_o(f) = \frac{\gamma_{fw}(f)}{\gamma_{ww}(f)} \quad (8.65)$$

qui correspond donc à l'expression fréquentielle du filtre optimal en fonction des données du problème.

Cas particulier : signal idéal et bruit indépendants

Tel quel, on éprouve du mal à interpréter l'expression fréquentielle du filtre optimal. Ce n'est plus le cas lorsque l'on fait l'hypothèse supplémentaire que les signaux $f(t)$ et $n(t)$ sont indépendants. Dans ce cas, nous pouvons écrire

$$E \{f(t)n(t)\} = E \{f(t)\} E \{n(t)\} \quad (8.66)$$

Nous pouvons alors transformer le numérateur de $\mathcal{H}_o(f)$ de la manière suivante

$$\Gamma_{fw}(\tau) = E \{w(t)f(t+\tau)\} = E \{[f(t)+n(t)]f(t+\tau)\} \quad (8.67)$$

ou encore

$$\Gamma_{fw}(\tau) = \Gamma_{ff}(\tau) + E \{n(t)\} E \{f(t+z)\} = \Gamma_{ff}(\tau) + \int_{-\infty}^{+\infty} n(t) dt \int_{-\infty}^{+\infty} f(t+\tau) dt \quad (8.68)$$

et finalement

$$\Gamma_{fw}(\tau) = \Gamma_{ff}(\tau) + \mathcal{N}(0) \mathcal{F}(0) \quad (8.69)$$

où $\mathcal{N}(0)$ et $\mathcal{F}(0)$ sont respectivement les transformées de FOURIER de $n(t)$ et $f(t)$ évaluée en 0. Un raisonnement similaire peut être appliqué au dénominateur de $\mathcal{H}_o(f)$:

$$\Gamma_{ww}(\tau) = \Gamma_{ff}(\tau) + \Gamma_{nn}(\tau) + 2\mathcal{F}(0) \mathcal{N}(0) \quad (8.70)$$

On peut à présent réécrire l'expression fréquentielle du filtre optimal:

$$\mathcal{H}_o(f) = \frac{\gamma_f(f)}{\gamma_f(f) + \gamma_n(f)} \quad f \neq 0 \quad (8.71)$$

où nous n'avons pas considéré la composante continue. Remarquons qu'en l'absence de bruit, la fonction de transfert du filtre optimal est égale à 1. Cette dernière expression donne le filtre optimal en termes des différentes densités spectrales de puissance connues. La figure 8.11 montre des exemples de filtres de WIENER dans des cas typiques.

8.3.2 Déconvolution de WIENER

Ainsi que l'on vient de le voir, le filtrage de WIENER ne permet que de restaurer des images qui ont été dégradées par l'ajout d'un bruit aléatoire. Nous allons voir qu'il est possible de combiner le filtrage de WIENER avec un procédé de déconvolution.

Considérons le schéma de la figure 8.12. Comme pour le filtrage de WIENER, nous allons réaliser le raisonnement pour le cas à une dimension. La fonction idéale est ici dégradée par convolution avec un filtre dont la fonction de transfert est $\mathcal{K}(f)$ et par l'ajout d'un bruit $n(t)$. La fonction dégradée est toujours notée $w(t)$.

Avant d'appliquer la technique du filtrage de WIENER, il est nécessaire de réaliser la déconvolution par un filtre dont la fonction de transfert est égal à l'inverse de la fonction de transfert $\mathcal{K}(f)$. Le filtre global de restauration est obtenu par combinaison du filtre de déconvolution et du filtre de WIENER. Vu que le filtre de déconvolution est connu, il reste simplement à déterminer la fonction de transfert $\mathcal{H}_o(f)$ du filtre de WIENER avant de combiner les deux filtres.

Calculons la transformée de FOURIER du signal dégradé

$$\mathcal{W}(f) = \mathcal{K}(f)\mathcal{F}(f) + \mathcal{N}(f) \quad (8.72)$$

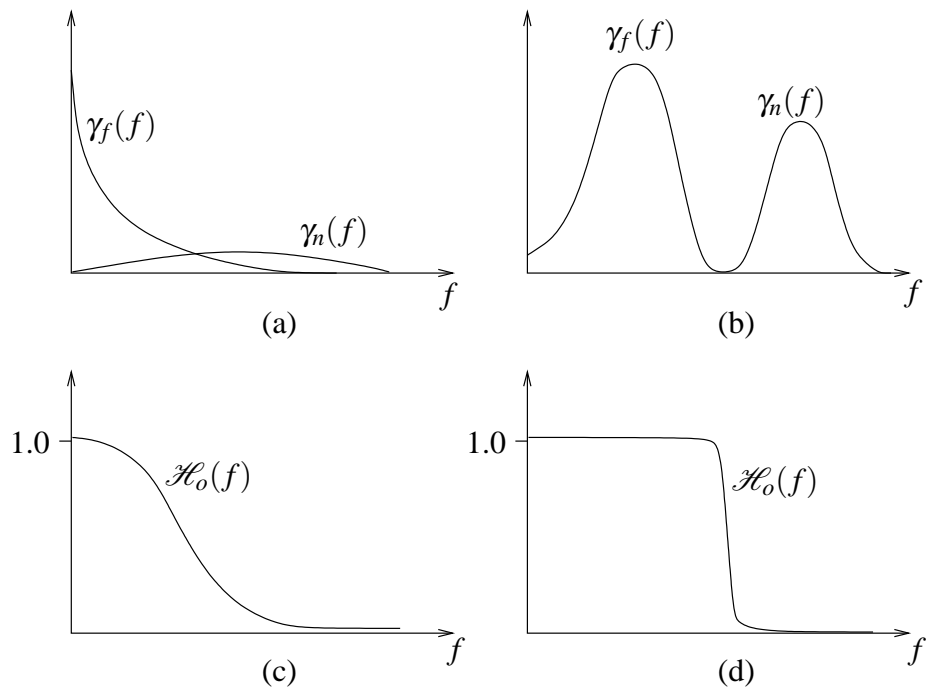


Figure 8.11: Illustration du filtre de WIENER. (a) et (c): signal utile passe-bas et bruit s'étalant sur toute la gamme de fréquences. (b) et (d): signal utile et bruit du type passe-bande.

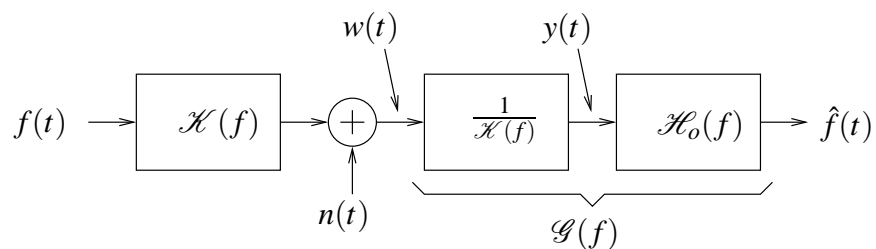


Figure 8.12: Schéma de déconvolution de WIENER.

ainsi que celle du signal $y(t)$ après déconvolution

$$\mathcal{Y}(f) = \mathcal{F}(f) + \frac{\mathcal{N}(f)}{\mathcal{K}(f)} \quad (8.73)$$

$$= \mathcal{F}(f) + \mathcal{N}'(f) \quad (8.74)$$

Tout se passe donc comme si $y(t)$ était le résultat d'une dégradation de la fonction $f(t)$ par un bruit aléatoire $n'(t)$. La détermination du filtre de WIENER devient alors évidente:

$$\mathcal{H}_o(f) = \frac{\gamma_f(f)}{\gamma_f(f) + \gamma_{n'}(f)} \quad (8.75)$$

dans le cas de signaux non-corrélés. Finalement, la fonction de transfert du filtre global de restauration (ou de déconvolution de WIENER) est donné par

$$\mathcal{G}(f) = \frac{\mathcal{H}_o(f)}{\mathcal{K}(f)} \quad (8.76)$$

$$= \frac{\mathcal{K}^*(f) \gamma_f(f)}{\|\mathcal{K}(f)\|^2 \gamma_f(f) + \|\mathcal{K}(f)\|^2 \gamma_{n'}(f)} \quad (8.77)$$

Si l'on se souvient à présent que le passage d'un processus aléatoire au travers d'un filtre revient à multiplier sa densité spectrale de puissance par le module de la fonction de transfert du filtre au carré, on peut finalement écrire

$$\mathcal{G}(f) = \frac{\mathcal{K}^*(f) \gamma_f(f)}{\|\mathcal{K}(f)\|^2 \gamma_f(f) + \gamma_n(f)} \quad (8.78)$$

Appliquer ce filtre dans le but de restaurer une image porte le nom de déconvolution de WIENER. Elle nécessite la connaissance des densités spectrales de puissance du signal original et du bruit ainsi que de la fonction de transfert $\mathcal{K}(f)$ du filtre ayant dégradé l'image.

Application aux cas des images

La déconvolution de WIENER peut aisément être adaptée aux cas des images à deux dimensions. Si l'on note $\gamma_f(u, v)$ la densité spectrale de puissance de l'image $f(x, y)$, $\gamma_n(u, v)$ celle du bruit aléatoire additionnel et $\mathcal{K}(u, v)$ la fonction de transfert du filtre ayant dégradé l'image originale, la déconvolution de WIENER se caractérise par le filtre de déconvolution suivant

$$\mathcal{G}(u, v) = \frac{\mathcal{K}^*(u, v) \gamma_f(u, v)}{\|\mathcal{K}(u, v)\|^2 \gamma_f(u, v) + \gamma_n(u, v)} \quad (8.79)$$

$$= \frac{\mathcal{K}^*(u, v)}{\|\mathcal{K}(u, v)\|^2 + \frac{\gamma_n(u, v)}{\gamma_f(u, v)}} \quad (8.80)$$

qui ne constitue qu'une généralisation de ce qui a été décrit à une dimension.

Chapter 9

Étude de la forme

Ce chapitre étudie plusieurs aspects relatifs à la forme des objets. Dans un premier temps, nous aborderons la question de la représentation de la forme. Nous étudierons ensuite quelques caractéristiques ou mesures des formes.

Traiter un contour n'est pas une tâche aussi aisée qu'il n'y paraît. Dans un premier temps, il faut décider si le contour appartient à l'objet ou non. Ensuite, en raison de l'échantillonnage, ce n'est pas le contour vrai que l'on observe mais une version approchée, souvent sujette à repli de spectre. Enfin, on peut jouer sur différents types de connexité sur une même grille d'échantillonnage, ce qui aboutit à définir des contours différents, pour un même objet!

9.1 Description de la forme

Il existe fondamentalement deux catégories de description de la forme:

- les descripteurs *linéiques*. Ils décrivent la forme pas à pas en suivant son contour.
- les descripteurs *surfaiques*. Ils opèrent par éléments de surface, ce qui en fait des opérateurs globaux. La *structure d'arbre* et le *squelette* sont les opérateurs de description centraux de toute une série d'opérateurs surfaiques.

9.1.1 Description linéique de la forme

Le codage en chaîne

Le *codage en chaîne* (run-length coding en anglais) est une description fort utilisée en codage des contours. Elle consiste à définir un jeu de directions (cf. figure 9.1) et à suivre les points de contours de proche en proche jusqu'à revenir au point de départ. Un contour se définit alors comme une structure contenant les coordonnées du point de départ et une liste de directions à suivre.

La figure 9.2 montre des exemples de codage en chaîne en 4 et 8-connexité. Le codage en chaîne en 8-connexité fournit (x_1, y_1) , 070675424452212. Le codage en 8-connexité est toujours plus court que le codage en 4-connexité mais il fait intervenir plus de directions distinctes.

Un problème typique du codage en chaîne intervient lors de la présence de points de rebroussement. On remarque alors qu'un même point de contour est décrit plusieurs fois. Ce phénomène est illustré à la figure 9.3.

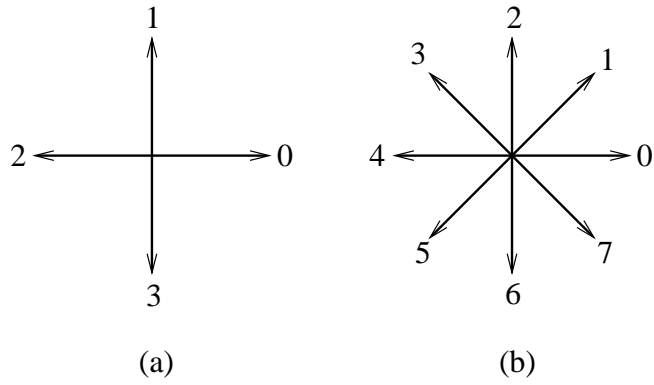


Figure 9.1: Directions en 4-connectivité (a) et en 8-connectivité (b).

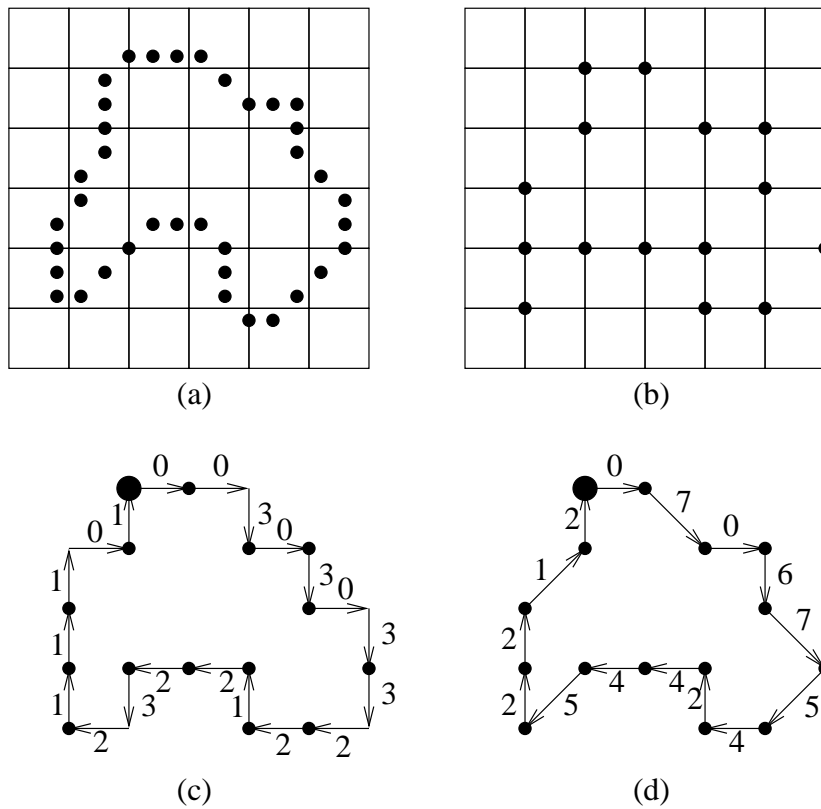


Figure 9.2: (a) Un contour, (b) contour après échantillonnage sur une grille, (c) codage en chaîne en 4-connectivité et (d) codage en chaîne en 8-connectivité.

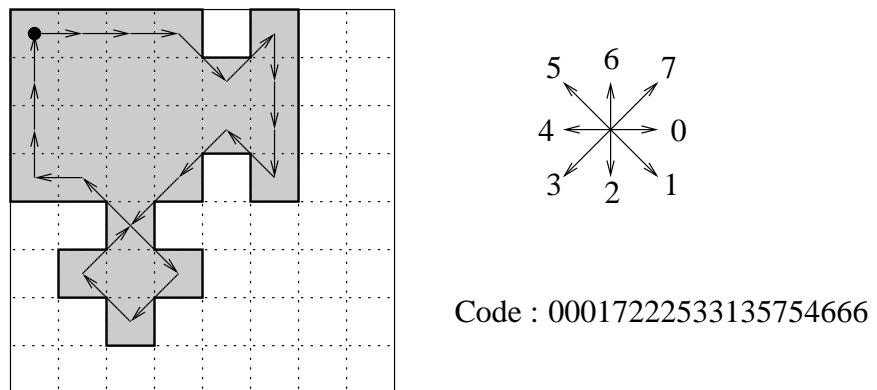


Figure 9.3: Contour avec des points de rebroussement.

Approximation polygonale

En général, on cherche à approcher au mieux la forme d'un contour avec un minimum de côtés. Le problème d'*approximation polygonale* d'un contour n'est pas trivial en tant que tel et l'on se heurterait rapidement à des temps de calcul excessif si on procède itérativement. Néanmoins, il existe des techniques satisfaisante pour le traitement d'images.

La figure 9.4 montre les étapes d'un procédé d'approximation polygonale d'un contour. On détermine tout d'abord la plus longue corde qu'il est possible de placer à l'intérieur de l'objet; les deux extrémités de cette corde correspondent à deux sommets du quadrilatère d'approximation. Pour trouver les autres sommets, on abaisse les plus longues perpendiculaires de part et d'autre de la corde. L'approximation peut être améliorée en ajoutant des sommets en abaissant les plus longues perpendiculaires sur chaque côté. La technique d'approximation consiste à ajouter des côtés jusqu'au moment où un critère d'arrêt est respecté, par exemple un écart surfacique entre la forme originale et son approximation inférieur à un certain pourcentage.

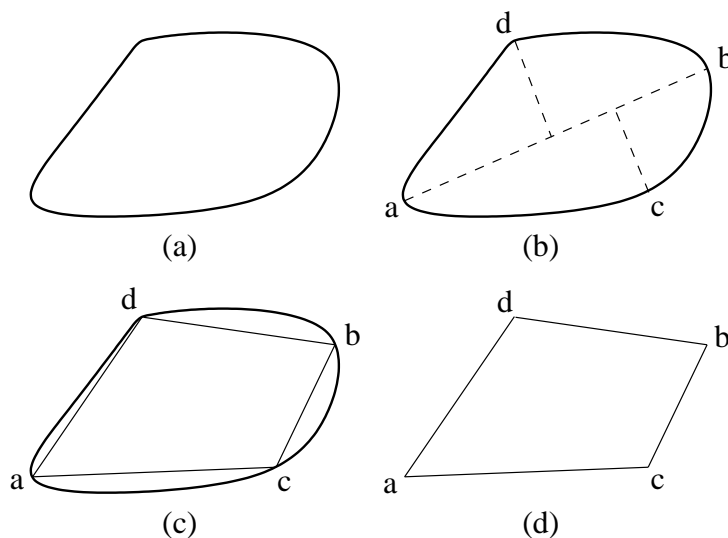


Figure 9.4: Approximation polygonale d'un contour.

Descripteur de FOURIER

Un contour est constitué d'une série de coordonnées de type (x_i, y_i) mais une description curviligne conviendrait puisqu'il ne s'agit jamais que d'un objet unidimensionnel parcourant un espace de dimension plus élevée. Partant de ce principe, on peut considérer une représentation sous forme de nombres complexes.

Soit une série de N points de contours dont les coordonnées sont représentées sous forme complexe, tels que représentés à la figure 9.5.

$$s[n] = x[n] + jy[n] \quad (9.1)$$

La transformée de FOURIER discrète de ce signal est

$$\mathcal{S}[u] = \frac{1}{N} \sum_{n=0}^{N-1} s[n] e^{-2\pi j u n / N} \quad (9.2)$$

Les coefficients $\mathcal{S}[u]$ sont appelés *descripteurs de FOURIER* du contour.

On peut approximer un contour en sélectionnant un certain nombre de descripteurs de FOURIER. La figure 9.6 illustre les approximations successives qui en découlent.

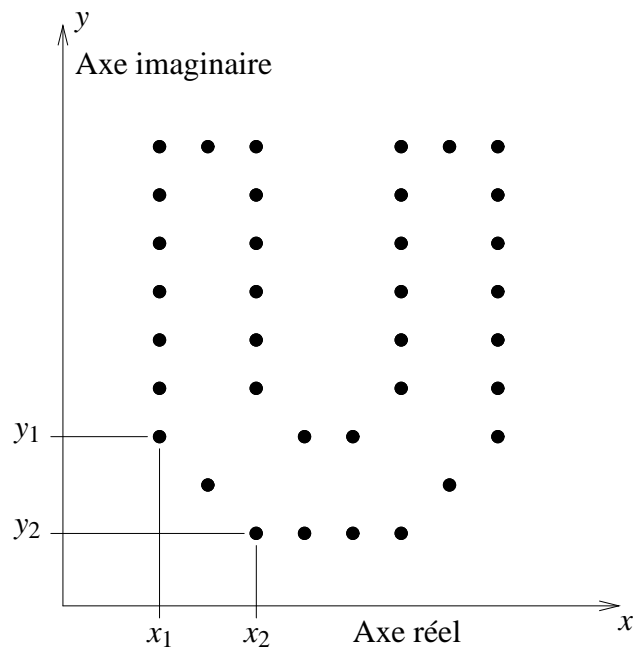


Figure 9.5: Représentation d'un contour comme une séquence de points complexes.

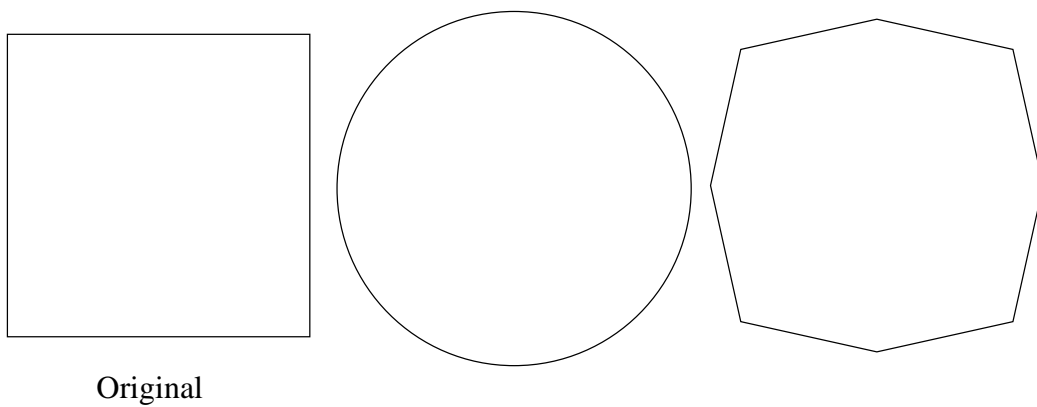


Figure 9.6: Forme originale et approximations du contour au moyen d'un nombre restreint de descripteurs de FOURIER.

9.1.2 L'arbre comme descripteur par éléments de surface

L'arbre quaternaire, appelé *quadtree* en anglais, est un outil de description de forme commode. Cet arbre est constitué de feuilles dont l'union reconstitue un ensemble et son complémentaire.

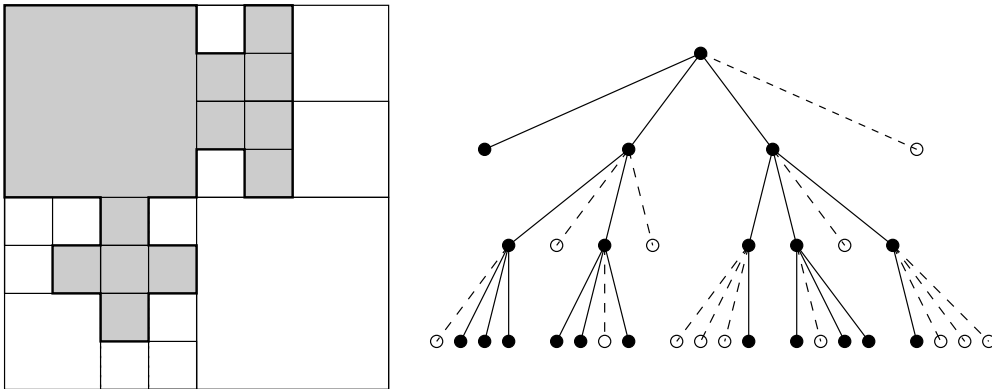


Figure 9.7: Descripteur de forme par arbre quaternaire.

Cette structure convient bien pour certains calculs. Par exemple, le calcul de l'aire s'obtient aisément en comptant le nombre de feuilles appartenant à l'objet en pondérant l'aire par la profondeur de la feuille dans l'arbre.

9.1.3 Le squelette morphologique comme descripteur par éléments de surface

Nous avons étudié précédemment des opérateurs morphologiques qui transforment un ensemble en un autre ensemble. Dans la mesure où la morphologie interprète une image en termes d'objets, il est licite de se demander s'il n'existe pas une opération qui est capable de décrire un objet par une série d'ensembles.

Une description en termes d'objets a plusieurs objectifs: interprétation automatique d'images, modélisation de matériau (séparation de phase), reconnaissance de formes, etc. Une description valable d'une scène intègre nécessairement une approche hiérarchique ou multirésolution des objets mais avant de travailler valablement avec des objets de résolution variable, il faudra définir le concept de *taille morphologique*. Ainsi tant pour le codage que pour le traitement d'objets, il est préférable de traiter en premier lieu les grands ensembles puis de passer successivement aux détails fins. La squelettisation offre une approche intéressante à plusieurs degrés de précision.

Définition intuitive du squelette

Considérons un ensemble continu X et sa frontière $\partial(X)$; un point x de X appartiendra au *squelette* de X , noté $S(X)$, s'il existe un disque centré en x inclus dans X et touchant $\partial(X)$ en deux points au moins. Plus intuitivement le squelette peut être défini comme le lieu des centres des boules maximales B contenues dans X (voir figure 9.8 pour quelques exemples de squelettes). Ces deux définitions ne sont pas tout à fait équivalentes mais elles ont la même fermeture topologique dans \mathbb{R}^2 .

L'opérateur issu de cette approche intuitive est défini par les propriétés suivantes:

- le résultat est large d'un pixel,
- le squelette passe par le milieu des objets, et
- l'opérateur de squelettisation préserve la topologie de l'objet.

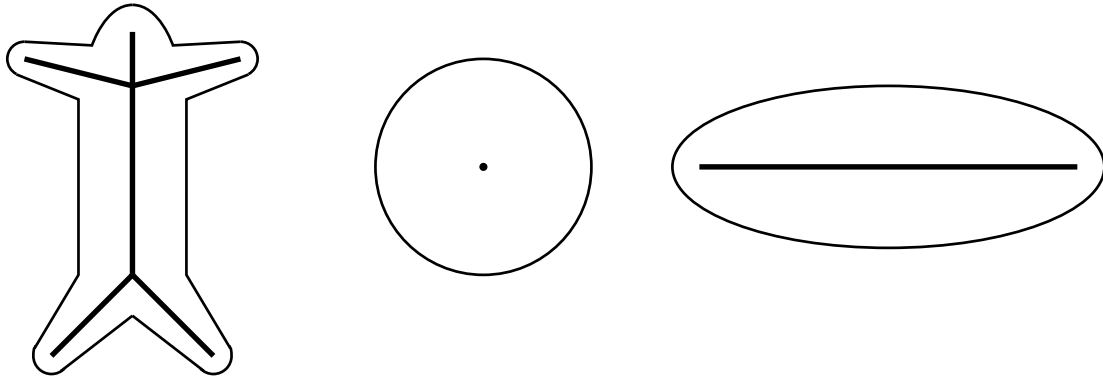


Figure 9.8: Exemple de squelettes de plusieurs ensembles.

Nous fournissons plus loin une définition mieux adaptée au cas d'ensembles discrétisés ($\subseteq \mathbb{Z}^2$) –la difficulté apparente est comment définir une boule maximale sur une grille d'échantillonnage–, faisant intervenir des érosions et des ouvertures. En fait, on cherche à définir le squelette d'un ensemble discret. Cette définition n'est pas évidente car la boule maximale n'a pas d'équivalent sur une grille d'échantillonnage. Il y a donc plusieurs approches, comme le laisse entrevoir l'illustration de la figure 9.9.

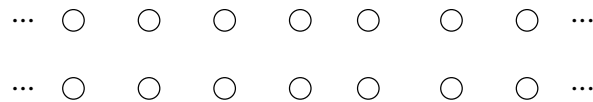


Figure 9.9: Que vaut le squelette de X ? Idéalement, il se situe au milieu. En pratique, en haut ou en bas.

Plus généralement, la définition pratique du squelette dans le cas d'objets définis sur une grille d'échantillonnage pose plusieurs problèmes:

- comment définir une boule sur une grille d'échantillonnage?
- comment associer une notion pratique de taille à une boule?
- faut-il garantir la préservation du nombre de connexité lors de la squelettisation?
- l'information fournie par le squelette (c'est-à-dire positions et rayons des boules) est-elle suffisante ou surabondante?

Propriétés d'une transformée par squelette

Si la squelettisation (opérateur de $X \rightarrow S(X)$) respecte les deux premiers principes de toute transformation morphologique, le respect de la continuité n'est pas nécessairement assuré. Par exemple, le squelette d'un polygone régulier est constitué de segments de droite joignant son centre aux sommets; lorsque, par un passage à la limite, on fait tendre le nombre de côtés vers l'infini (cercle), le squelette devient pratiquement équivalent au polygone alors que le squelette d'un cercle se réduit à son centre. La connaissance locale exige aussi quelques précautions.

Citons les propriétés algébriques suivantes.

Propriété 62 *La squelettisation est une opération ni croissante ni décroissante. En effet, $X \subseteq Y$ n'entraîne généralement pas que $S(X) \subseteq S(Y)$ ou $S(Y) \subseteq S(X)$.*

Propriété 63 La squelettisation est une opération anti-extensive: $S(X) \subseteq X$.

Propriété 64 La squelettisation est une transformation idempotente: $S(S(X)) = S(X)$.

À cette réflexion, on peut déjà deviner que le squelette est lié à l'ouverture morphologique par des éléments structurants de plusieurs tailles différentes.

Mais avant de fournir l'expression analytique du squelette, nous introduisons le concept de taille morphologique. C'est aussi l'occasion de présenter les filtres multirésolution.

Préliminaires: la taille morphologique et les filtres multirésolution

Taille morphologique.

Définition 65 Soit B un élément structurant dont la taille est fixée arbitrairement à r . On définit un espace continu d'ensembles par

$$rB = \{rb \mid b \in B\} \quad r \geq 0 \quad (9.3)$$

où r est un paramètre continu. L'objet rB a la même forme que B à condition de prendre un ensemble B connexe et borné.

Homothétie d'un ensemble discret. La définition précédente ne convient pas du tout pour le cas d'un ensemble discrétisé¹. En effet, si B est un carré de 2×2 pixels, alors $2B$ est composé uniquement des sommets d'un carré de taille 3×3 . En d'autres termes, le nombre d'éléments contenus dans l'ensemble $2B$ n'a pas été modifié alors que la taille l'était. Ceci nous conduit à proposer une autre définition basée sur l'opération de dilatation.

Définition 66 Soit B , un ensemble discret, fini de \mathbb{Z}^2 et de taille unitaire. On crée une famille d'objets binaires générée par B et caractérisée par le paramètre $n = 0, 1, \dots$ par

$$nB = \underbrace{B \oplus \dots \oplus B}_{n-1 \text{ dilatations}} \quad (9.4)$$

Dans cette expression, nB est obtenu par $(n - 1)$ dilatations successives. Si $n = 0$, $0B = \{(0, 0)\}$ par convention. Notons au passage que $nB \oplus mB = (n + m) \oplus B$ pour tout n, m . Si B est convexe, alors nB a exactement la géométrie de B mais est de taille n . Remarquons que cette dernière définition est aussi valable pour des ensembles B convexes et continus.

Remarque pratique concernant l'implémentation. Le théorème suivant permet d'arriver à une simplification de la formule 9.4.

Théorème 67 [38, page 132] Soit $\partial(B)$ la frontière du fermé B de \mathbb{R}^2 . Alors

$$\partial(B) \oplus B = B \oplus B \quad (9.5)$$

Dès lors,

$$nB = \underbrace{B \oplus \dots \oplus B}_{n-1 \text{ dilatations}} = B \oplus \underbrace{\partial(B) \oplus \dots \oplus \partial(B)}_{n-2 \text{ dilatations}} \quad (9.6)$$

Remarquons que cette formule donne un sorte de borne maximale. En effet, comme le montre le théorème suivant, il n'est pas nécessaire de prendre tous les points de la frontière lorsque la *condition de réduction* est respectée.

¹Le terme "discret" est à prendre au sens de "défini sur une grille d'échantillonnage", tel qu'on le retrouve dans la littérature scientifique.

Théorème 68 [Condition de réduction] [38, page 130] *Si $A \subseteq B$ alors*

$$A \oplus B \supseteq B \oplus B \Leftrightarrow A \oplus B = B \oplus B \quad (9.7)$$

Pour s'en convaincre, prenons le cas d'un ensemble $B = H \oplus V$. Alors $B \oplus B = B \oplus \partial(H) \oplus \partial(V)$; la condition de réduction est respectée si $A = \partial(H) \oplus \partial(V)$. Il suffit donc de dilater un rectangle B par ses quatre extrémités pour obtenir $B \oplus B$.

La figure 9.10 illustre un autre cas.

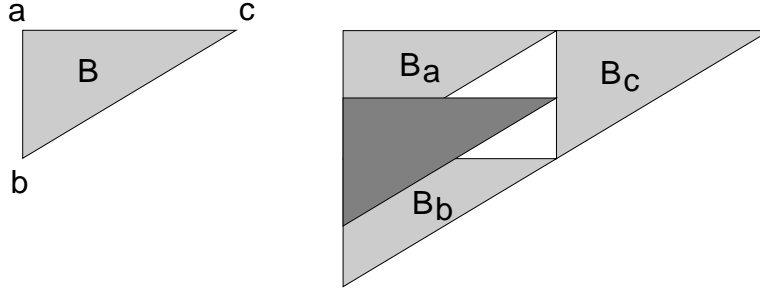


Figure 9.10: Ici, $\partial(B) \oplus B = ([a, b] \cup \{c\}) \oplus B$.

Filtres multirésolution.

Définition 69 *On définit l'ouverture et la fermeture à multirésolution par*

$$X \circ nB = (X \ominus nB) \oplus nB \quad (9.8)$$

$$X \bullet nB = (X \oplus nB) \ominus nB \quad (9.9)$$

Comme $(X \ominus B) \ominus C = X \ominus (B \oplus C)$ et $(X \oplus B) \oplus C = X \oplus (B \oplus C)$, l'ouverture s'exprime aussi par

$$X \circ nB = \underbrace{[(X \ominus B) \ominus B \dots \ominus B]}_{n \text{ érosions}} \oplus \underbrace{B \oplus B \dots \oplus B}_{n \text{ dilatations}} \quad (9.10)$$

Une forme similaire vaut pour la fermeture $X \bullet nB$. La complexité de cette dernière expression est directement proportionnelle au nombre de points de B alors que celle de $(X \ominus nB) \oplus nB$ est de complexité quadratique. Si B a une forme suffisamment régulière, $X \circ nB$ et $X \bullet nB$ fournissent des versions filtrées non linéaires et multirésolution de la frontière de X . Par définition,

$$X \circ nB = \bigcup_{(nB)_z \subseteq X} (nB)_z \quad (9.11)$$

Ainsi, $X \circ nB$ supprime toute la surface intérieure qui ne peut inclure un translaté de nB . C'est pour cette raison que l'on associe le concept de taille n à celle de l'objet analysé.

Définition formelle du squelette

Si en chaque point du squelette, on dessine le disque de taille maximale correspondant, l'union de ces cercles reconstitue l'objet de départ et l'enveloppe des disques formera le bord de l'objet. Soit $S_r(X)$, $r > 0$, le sous-ensemble du squelette $S(X)$ des centres des disques dont le rayon est égal à r ; les sous-ensembles $S_r(X)$ s'obtiennent par des opérations morphologiques d'érosion et d'ouverture. On peut démontrer qu'à l'exception des figures non bornées ou de celles incluant des points isolés,

voire des lignes qui sont leur propre squelette, le squelette existe et s'exprime de manière équivalente par

$$S(X) = \bigcup_{r>0} S_r(X). \quad (9.12)$$

Bien que le concept de squelette soit difficile à étendre au cas d'images discrétisées, cela reste possible. SERRA propose une définition de la squelettisation qui s'applique à tout ensemble discret, borné et non vide. Nous passons à une définition plus pratique, due à LANTUÉJOUL, faisant intervenir les concepts de taille morphologique et de filtre multirésolution.

Définition 70 *Le squelette d'un ensemble X s'obtient par l'union de sous-ensembles S_n , chacun de ces sous-ensembles étant l'ensemble des points de référence de copies de nB incluses dans X mais non incluses dans une copie de $(n+1)B$ contenue dans X*

$$S_n(X) = (X \ominus nB) \setminus ((X \ominus nB) \circ B), \quad n = 0, 1, \dots, N \quad (9.13)$$

$$S(X) = \bigcup_{n=0}^N S_n(X) \quad (9.14)$$

où B est un carré de 2×2 pixels en trame carrée et un hexagone en trame hexagonale.

Les images représentées à la figure 9.11 ont été obtenues par le biais de cette formule.

Remarquons que la relation de LANTUÉJOUL fournit un squelette qui ne préserve pas la connexité.

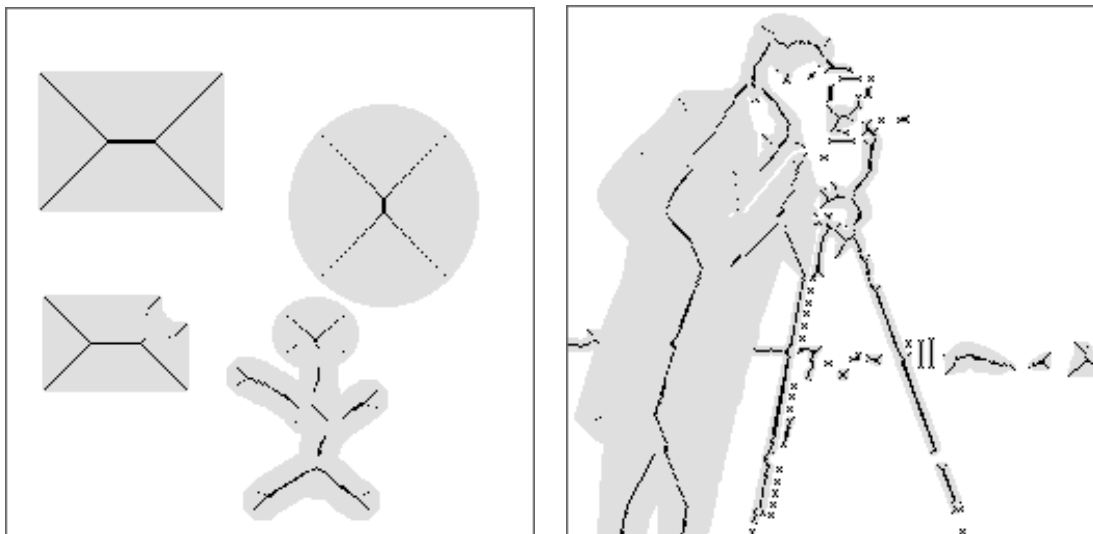


Figure 9.11: Exemples de squelette tel que calculé par la formule de LANTUÉJOUL.

Le sous-ensemble de taille n , $S_n(X)$, s'obtient en érodant X par nB , puis en ne conservant du résultat $X \ominus nB$ que les parties anguleuses et les lignes d'épaisseur nulle. Le squelette de taille n est la différence entre $X \ominus nB$ et son ouverture par B , $(X \ominus nB) \circ B$. Si l'objet est entièrement contenu dans un masque de taille $M \times M$, alors $n \leq M$. De ces propos, il découle que la famille des couples $(S_n(X), n)$ forme une description $\rho(X)$ de l'image. Pour obtenir une approximation $\pi(X)$ de l'image, il faut encore se munir d'une recombinaison valable $\tau(X)$.

La squelettisation est une des rares transformations morphologiques à avoir un inverse. L'image numérisée X est reconstruite avec exactitude par la réunion de ses $N+1$ sous-ensembles $S_n(X)$ dilatés

par l'élément de structure de taille adéquate:

$$\pi(X) = X = \bigcup_{n=0}^N [S_n(X) \oplus nB] \quad (9.15)$$

Il apparaît clairement que plus l'indice n est élevé, plus les détails de reconstruction sont grossiers. Si on tient compte de l'associativité de la dilatation et de la distributivité de la dilatation par rapport à l'union, il est possible de trouver une forme équivalente de reconstruction de X

$$\pi(X) = [[[S_N(X) \oplus B] \cup S_{N-1}(X)] \oplus B \cup S_{N-2}(X) \dots] \oplus B \cup S_0(X) \quad (9.16)$$

Nous avons vu que le squelette de X est plus mince que X par l'intermédiaire de la propriété d'anti-extensivité. De plus, les sous-ensembles $S_n(X)$ forment une partition du squelette.

Autres formules. Il existe d'autres formules de calcul de l'opération de squelettisation. VINCENT a proposé une méthode basée sur la fonction distance [40].

Définition 71 L'opérateur de fonction distance $\phi(X)$ d'un ensemble $X \subseteq \mathcal{E}$ est la fonction définie par

$$[\phi(X)](h) = d(X^c, h) \quad (9.17)$$

avec la convention que $d(\emptyset, h) = +\infty$ pour $h \in \mathcal{E}$.

Par exemple, si on définit une image de marqueurs à partir des maxima locaux de la fonction distance, on obtient le squelette par boules maximales. Le principe de cet algorithme est illustré à la figure 9.12.

La figure 9.13 montre quant à elle des squelettes calculés par l'algorithme de VINCENT. On voit que ce type d'algorithme préserve la connexité de la figure analysée.

9.1.4 Autres descripteurs par éléments de surface

Le choix d'une famille d'éléments structurants homothétiques d'un même élément peut sembler arbitraire, d'autant qu'il n'est nullement obligatoire. On peut tout aussi bien choisir une famille d'éléments structurants quelconques et décrire un objet comme l'union de ces éléments placés aux bons endroits.

La figure 9.14 montre un objet obtenu par la superposition de trois éléments structurants. À l'instar du squelette, la description comprend une série de coordonnées avec un indice qui, dans le cas présent, référence un élément structurant contenu dans la famille de départ et non une taille; c'est donc une sorte de généralisation de la notion de squelette.

Cette généralisation amène son lot de difficultés puisqu'il faut parfois choisir entre plusieurs possibilités pour recouvrir une forme. Il faut alors se doter d'un critère automatique d'optimisation.

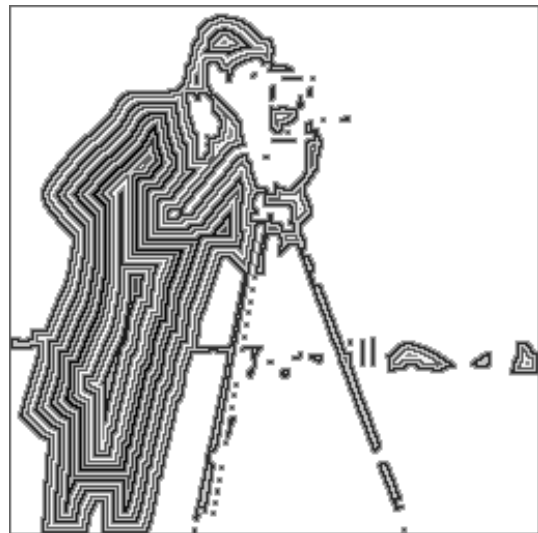
La figure 9.15 compare différents descripteurs sur une forme plus complexe.

9.2 Mesures

Les mesures effectuées sur des images sont multiples et variées, qu'il s'agisse d'étudier les formes ou les textures. On peut ainsi caractériser certains aspects géométriques comme la taille mais aussi des propriétés topologiques, le nombre de trous par exemple, ou encore analyser un objet en considérant plusieurs niveaux de résolution.



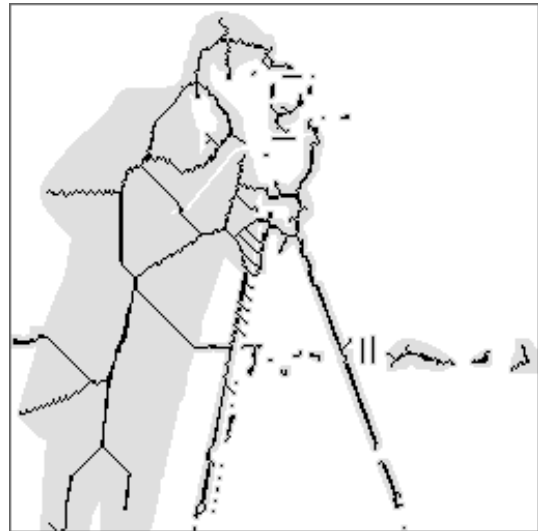
(a) Image binaire originale



(b) Courbes de niveau de la fonction distance



(c) Position des maxima locaux de (b)



(d) Squelette par boules maximales

Figure 9.12: Squelette par boules maximales.

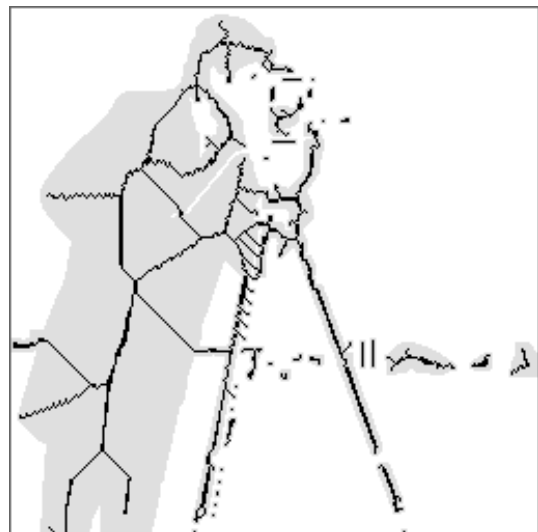
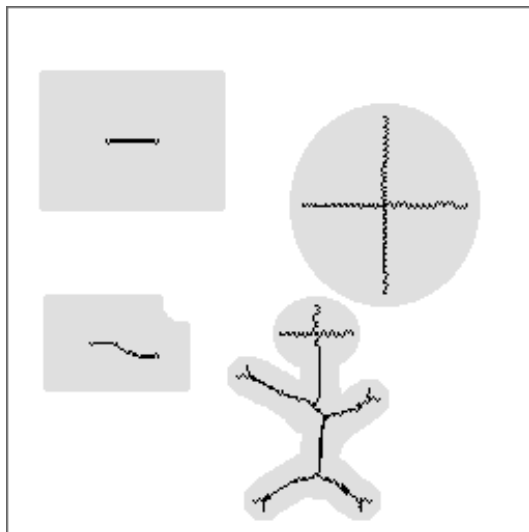


Figure 9.13: Exemples de squelettes tels que calculés par l'algorithme de VINCENT.

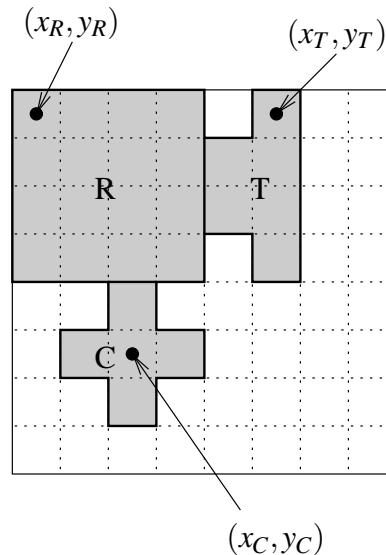


Figure 9.14: Description d'une forme comme l'union d'un rectangle, d'un triangle et d'un losange.

9.2.1 Mesures géométriques de base

Fondamentalement, les trois mesures de base caractérisant une forme sont la *périmètre*, la *surface* et le *nombre d'EULER-POINCARRÉ*. Par ailleurs, HADWIGER [9] a montré que toute mesure de la forme continue, additive et invariante par rotation et translation est nécessairement une combinaison linéaire de ces trois mesures de base.

Périmètre

Le *périmètre* d'une figure se calcule aisément à partir d'une description du contour par un codage en chaîne. En effet, il suffit de compter le nombre de points de ce codage en chaîne en considérant un facteur multiplicatif $\sqrt{2}$ pour les directions diagonales.

Périmètre d'un polygone Soient N_p et N_i le nombre d'éléments respectivement pairs et impairs contenus dans le codage en chaîne en 8-connexité d'un contour C .

$$P(C) = N_p + \sqrt{2}N_i \quad (9.18)$$

Aire

L'*aire* ou *surface* d'un objet s'obtient simplement en comptant le nombre de pixels qui le composent. Parfois, on distingue l'aire d'un objet de celle calculée en remplissant les trous présents dans l'objet. D'un point de vue pratique, un balayage direct de l'image suffit au calcul de l'aire.

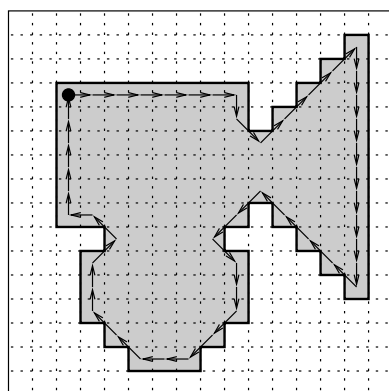
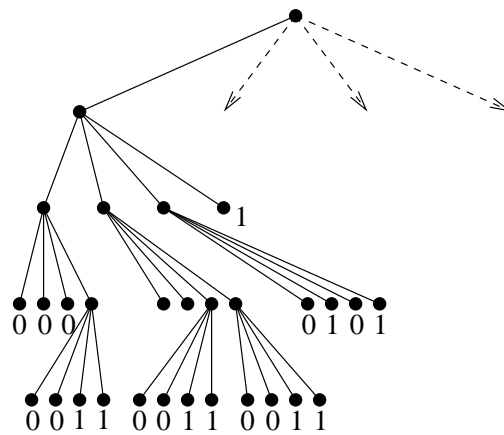
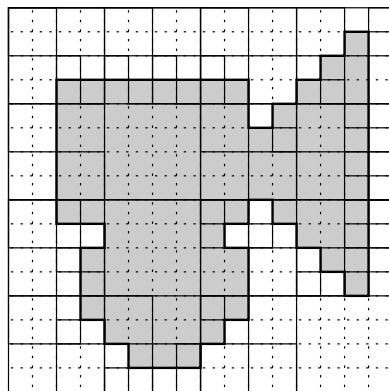
Caractéristique d'EULER-POINCARRÉ

Il existe une relation fondamentale entre le nombre de composantes connexes C et le nombre de trous T .

Définition 72 [Nombre d'EULER] On définit le nombre d'EULER E par

$$E = C - T \quad (9.19)$$

Les figures 9.16 et 9.17 illustrent le calcul du nombre d'EULER.



Code : 0000000217777222...

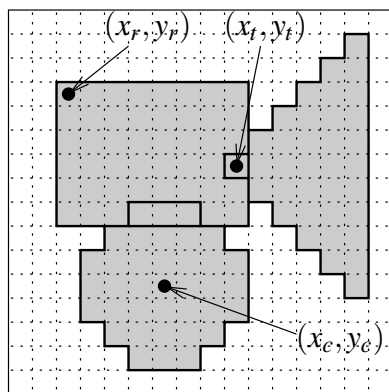


Figure 9.15: Comparaison de plusieurs descripteurs de formes.

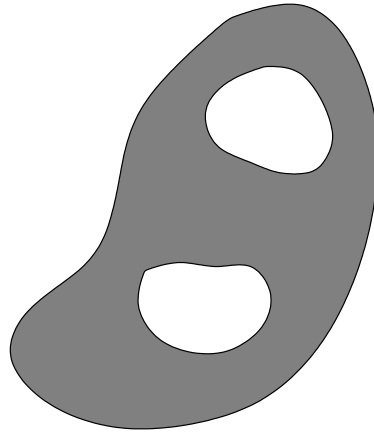


Figure 9.16: Nombre d'EULER: $E = 1 - 2 = -1$.



Figure 9.17: Deux lettres de nombre d'EULER différent.

Propriétés topologiques. Une *propriété topologique* est une propriété préservée lors de l'étirement d'une forme. Ainsi, la distance métrique n'est clairement pas une propriété topologique, ainsi que tous les opérateurs qui en découlent.

Le nombre d'EULER est une propriété topologique car C et T sont des propriétés topologiques. Le *nombre de connexité*, c'est-à-dire le nombre de particules, est également une propriété topologique.

9.2.2 Indices de forme

Un premier moyen, très utilisé en pratique, pour caractériser une forme est d'inscrire l'objet dans un rectangle, un cercle ou toute autre forme de référence; le principe en est illustré à la figure 9.18. Les paramètres de ces formes de référence caractérisent alors grossièrement l'objet analysé.

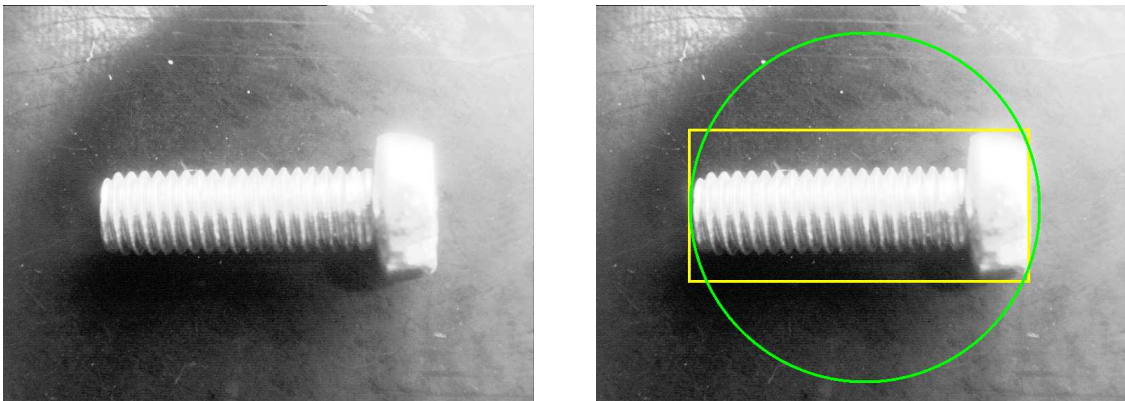


Figure 9.18: Un objet et différentes figures de référence exinscrites.

Ainsi, PARKER [24] définit la *rectangularité* d'un objet comme le rapport entre l'aire de l'objet et l'aire du plus petit rectangle exinscrit. La rectangularité est toujours inférieure à 1; elle vaut 1 pour un rectangle.

Plus généralement, un *indice de forme* est une description chiffrée de la forme d'un objet binaire. RUSS [28] définit une série d'indices de forme. En voici quelques-uns (cf. tableau 9.1). Soient $A(X)$, $P(X)$, D_{min} et D_{max} respectivement l'aire de X , le périmètre de X , le diamètre minimum de l'ellipse exinscrite à X et le diamètre maximum.

Nom	Expression
Facteur de forme	$\frac{4\pi A(X)}{P(X)^2}$
Circularité	$\frac{4A(X)}{\pi D_{max}^2}$
Rapport d'aspect	$\frac{D_{max}}{D_{min}}$

Table 9.1: Quelques indices de forme (d'après [28]).

On peut aussi considérer des indices qui mesurent l'écart, encore appelé *déficit*, entre la forme d'un objet et une forme de référence. Ainsi, on définit l'*indice d'écart au cercle inscrit* (de rayon R) comme

$$IR(X) = 1 - \frac{A(X)}{\pi R^2} \quad (9.20)$$

9.2.3 Moments

Les *moments* d'une image sont des descripteurs courants, notamment en reconnaissance de caractères. La définition fournie ci-après considère une image $f(x,y)$; le passage à un objet se fait en prenant une valeur 1 à l'intérieur de l'objet et 0 en dehors.

Définition 73 Le moment d'ordre $p + q$ de la fonction $f(x,y)$ est

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \quad (9.21)$$

On peut centrer ces moments en déduisant la moyenne, ce qui fournit les moments centrés

$$\mu_{pq} = \sum_x \sum_y (x - m_{10})^p (y - m_{01})^q f(x,y) \quad (9.22)$$

ou encore normaliser leur valeurs (cf. [34])

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{p+q}{2}}} \quad (9.23)$$

HU [12] a montré l'usage qu'on pouvait faire de tels moments centrés normalisés. Ils conduisent à des caractéristiques qui possèdent des propriétés d'invariance en translation et en rotation (notamment) intéressantes. Leur détermination est néanmoins coûteuse en raison des nombreuses opérations de calcul.

9.2.4 Mesures morphologiques

Les mesures morphologiques les plus courantes sont des *granulométries* qui représentent l'évolution d'une caractéristique en fonction de la taille de l'élément structurant considéré pour l'analyse.

Granulométries

Proposition 74 [?, page 93] Si B_i , $i > 0$, est une collection d'éléments structurants tels que B_i est égal à son ouvert au moyen de tout élément B_j si $i \geq j$, alors les ouvertures $\gamma_i = \gamma_{B_i}$ satisfont la propriété de semi-groupe

$$\gamma_i \gamma_j = \gamma_j \gamma_i = \gamma_i, \quad i \geq j \quad (9.24)$$

Une collection de telles ouvertures γ_i , $i > 0$ est appelée *granulométrie*. La mesure de LEBESGUE des ensembles ouverts –l'aire en pratique– $\gamma_i(X)$ définit une fonction décroissante de $i > 0$, appelée *courbe granulométrique*.

Covariogramme géométrique

Comme autre mesure morphologique, on peut définir le covariogramme qui effectue une mesure corrélatoire spatiale de l'objet.

Définition 75 [Covariogramme]

$$C_X(h) = A(X \cap X_h) \text{ pour } h \in \mathbb{R}^2 \quad (9.25)$$

L'emploi d'une telle mesure se justifie en présence de plusieurs objets non connexes et de forme fort proche, car le covariogramme fait ressortir l'écart entre les objets.

Enveloppe convexe

La notion de convexité et d'enveloppe convexe intervient souvent en traitement d'images. La convexité se définit de la manière suivante:

Définition 76 [Convexité] *Un ensemble $X \subseteq \mathcal{E}$ est convexe si $rx + (1 - r)y \in X$ pour tout $x, y \in X$ avec $r \in [0, 1]$.*

Autrement dit, toute droite reliant deux points quelconques de X doit être entièrement contenue dans X . La figure 9.19 montre quelques objets.

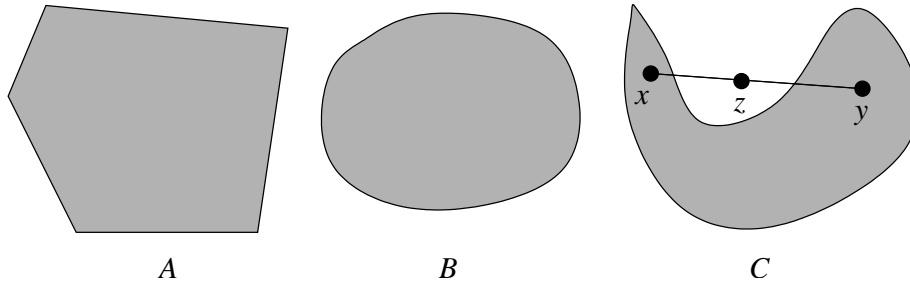


Figure 9.19: A, B sont convexes; C ne l'est pas.

La notion de convexité permet de définir celle d'enveloppe convexe.

Définition 77 *L'enveloppe convexe d'un ensemble $X \subseteq \mathcal{E}$, notée $co(X)$, est l'intersection de tous les ensembles convexes contenant X .*

La figure 9.20 montre un ensemble et son enveloppe convexe. Il est à remarquer que l'enveloppe convexe ne contiendra aucun trou, ce qui signifie que tous les trous de l'ensemble X seront bouchés.

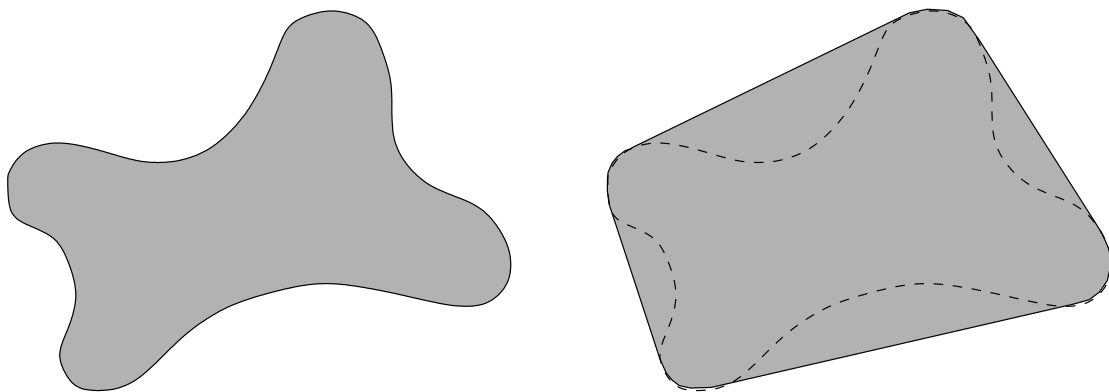


Figure 9.20: Un ensemble X et son enveloppe convexe.

La convexité est une notion géométrique riche qui a donné naissance à des développements mathématiques intéressants. Le résultat suivant est une conséquence directe d'un théorème de CARATHEODORY.

Proposition 78 [?, page 274] *Soient $X, Y \subseteq \mathcal{E}$*

$$co(X \oplus Y) = co(X) \oplus co(Y) \quad (9.26)$$

Bien entendu, on peut définir des indices de forme basé sur l'enveloppe convexe. Ainsi, l'*indice de convexité* vaut

$$\frac{P(\text{co}(X))}{P(X)} \quad (9.27)$$

Chapter 10

Segmentation d'images

10.1 Présentation du problème

Dans les approches classiques de compression ou de rehaussement, la sortie est une approximation ou une amélioration de l'image d'entrée. Une autre branche du traitement d'image s'occupe de l'*analyse d'image* ou *analyse de scène*. Dans ce domaine, l'entrée est toujours une image mais la sortie est une *description* de l'image. La plupart des descriptions nécessitent une *détection* préalable des formes présentes dans l'image, étape appelée *segmentation*. Ce chapitre est consacré à l'*étude de techniques de segmentation d'image*.

Pour situer le problème de la segmentation d'image, nous étudierons une série d'outils de segmentation en précisant leurs avantages et inconvénients.

10.1.1 Définition

Définition 79 En général, on considère que l'objectif de la segmentation revient à construire une série de régions R_1, \dots, R_n telles

$$\mathcal{E} = \bigcup_{i=1}^n R_n \text{ et } \forall i \neq j, R_i \cap R_j = \emptyset \quad (10.1)$$

Définition 80 Formellement, la segmentation s'apparente à un opérateur ϕ agissant sur l'image I et fournissant, par exemple, une image binaire $\phi(I)$ qui différencie les points des contours des objets.

Comme alternative, on pourrait fort bien considérer que la segmentation a pour objectif de définir une série de régions d'intérêt. C'est par ailleurs la manière usuelle de l'aborder dans le secteur du traitement d'images industriel.

La tâche n'est généralement pas simple et le lecteur attentif remarquera que les procédés de segmentation sont fort nombreux. D'une manière générale, l'utilisateur se voit contraint de construire un procédé de segmentation le plus adapté au problème qu'il traite. Il faut qu'il garde à l'esprit qu'une même image peut subir plusieurs segmentations toutes aussi valables les unes que les autres; c'est l'application en ligne de mire qui permet de déterminer si le résultat de la segmentation est adéquat ou non.

La segmentation, comme la plupart des techniques de traitement d'images, peut être abordée sous l'angle *local* ou *global*. Le premier considère que la segmentation ne requiert pas la connaissance du contenu global de l'image mais seulement le voisinage direct de chaque pixel. À l'inverse, dans une approche globale, c'est la totalité du contenu de l'image qui constitue l'unique guide. Ainsi, lorsqu'on désire segmenter un nombre précis d'objets (typiquement une centaine) s'agit-il d'un procédé global.

La différence essentielle entre l'approche locale ou globale tient du fait que le moindre changement dans la valeur d'un pixel peut, respectivement, se limiter à des perturbations au voisinage local de la valeur affectée ou modifier la totalité du résultat de la segmentation.

10.2 Segmentation par seuillage

Le *seuillage* d'une image à plusieurs niveaux est le procédé de segmentation le plus simple. L'hypothèse qui sou-tend l'approche consiste à dire que tout objet se différencie de l'arrière-fond. Elle aboutit à fournir une série d'objets extraits de l'arrière-plan.

La figure 10.1 illustre la difficulté à segmenter par simple seuillage; l'origine du problème est l'éclairage non uniforme de la scène, problème courant en pratique.

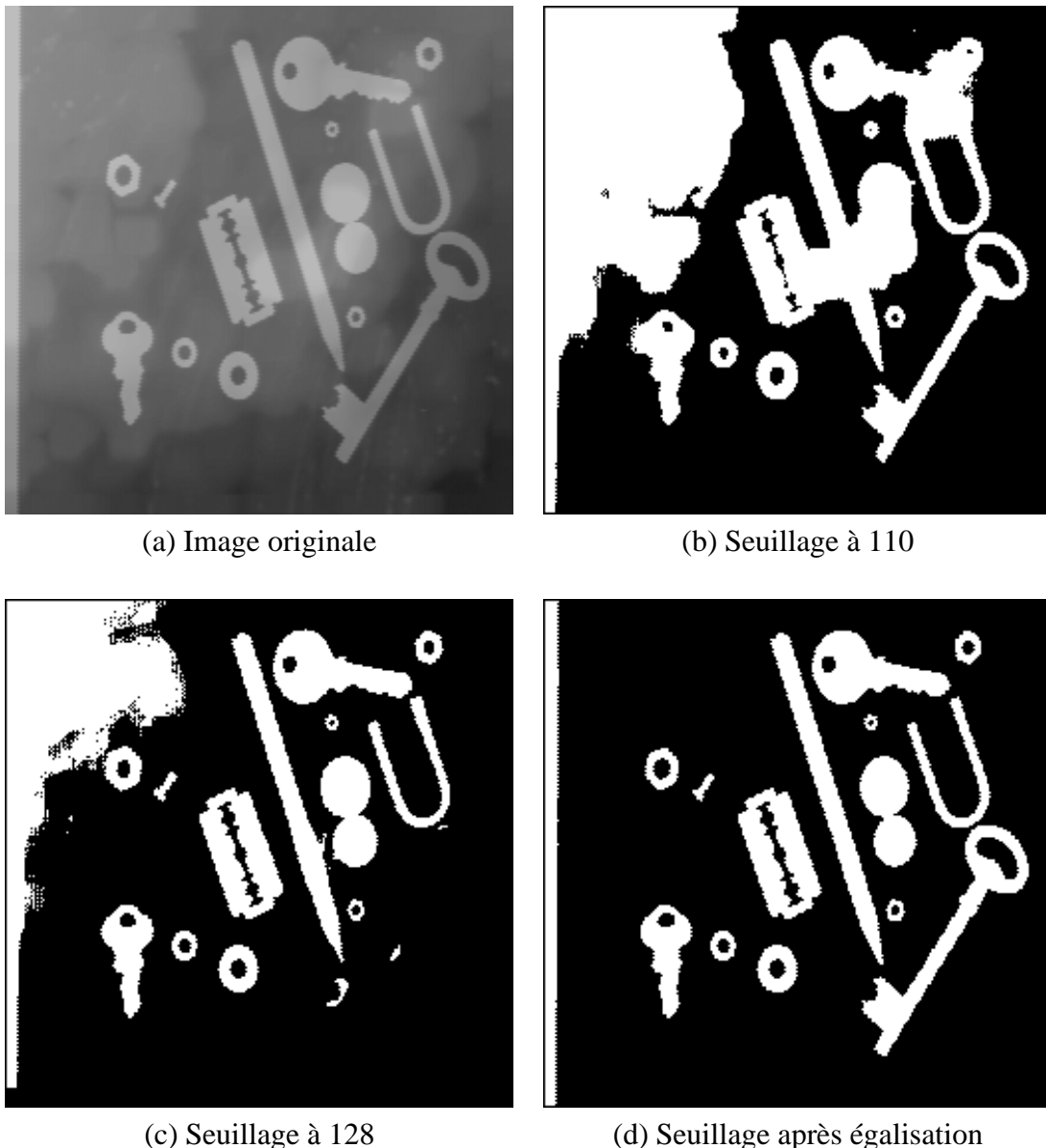


Figure 10.1: Seuillages.

Comme l'outil de segmentation par seuillage est conceptuellement simple, certains auteurs ont amélioré la segmentation en imaginant des techniques de détermination optimale du seuil. Cela aboutit par exemple à la détermination de seuils comme illustré à la figure 10.2.

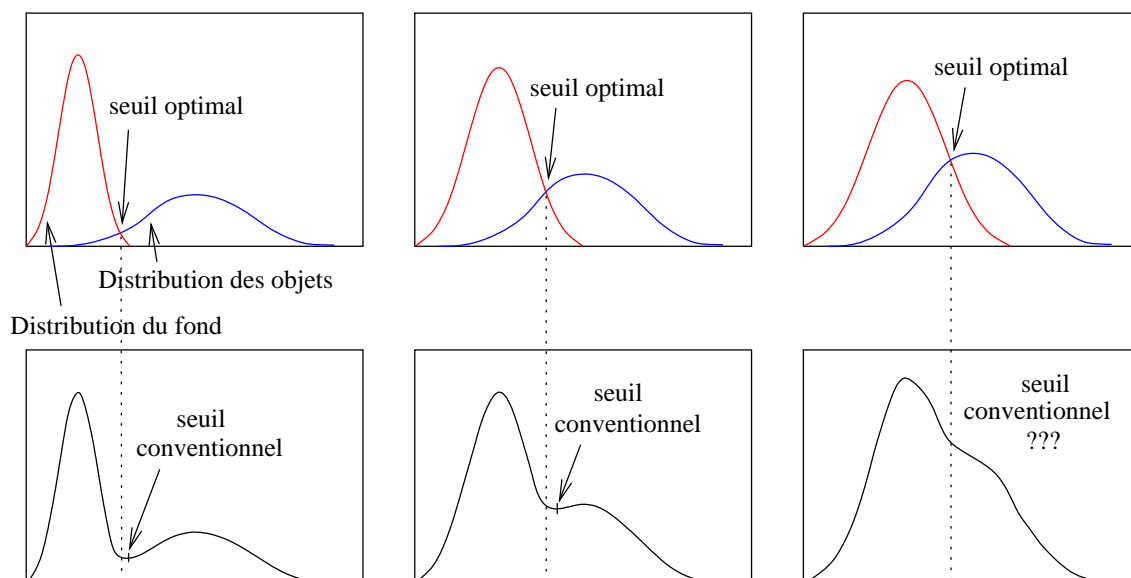


Figure 10.2: Seuillage optimal (d'après [34]).

10.3 Segmentation par détection de régions

Dans le domaine de l'analyse d'images et plus particulièrement en morphologie mathématique, on considère que les fonctions sont des surfaces où la valeur de luminance est synonyme d'élévation. Cette représentation est pratique pour l'étude de techniques de segmentation –technique destinée à détecter des objets dans une image– utilisant des outils morphologiques. Avec quelques variantes, les techniques de segmentation morphologique se façonnent à partir de deux outils que sont la *transformation de ligne de partage des eaux* et la *modification homotopique* qui, elle, résout les problèmes de sur-segmentation (elle ne sera pas abordée). Outre leur aptitude à détecter des contours, les opérateurs morphologiques sont aussi capables de remplir des tâches annexes mais souvent difficiles comme la fermeture de contours restés ouverts après segmentation.

10.3.1 Ligne de partage des eaux

Si dans une interprétation topographique, la valeur de luminance représente l'élévation, cela ne signifie pas que le profil puisse être découpé correctement par un simple effet de seuillage. Un seuil élevé détectera bien les pics; un seuil faible extraira les vallées; mais ces fonctions ne sont pas compatibles.

Puisque les valeurs de luminance ne semblent pas convenir directement, on est tenté de recourir à la norme des gradients morphologiques. L'inconvénient majeur de ceci est que le gradient est lui-même une fonction et, que comble de malheur, il faut à nouveau essayer des seuils pour segmenter. Bien sûr, a priori, l'information du gradient est plus pertinente que la simple valeur de luminance mais le problème reste entier. On peut encore remplacer le gradient par un outil qui analyse les transitions.

Trouver la procédure optimale pour placer des contours sur la fonction gradient n'est pas simple du tout. Par contre, identifier les régions où il n'y a pas de contours est plus simple: il s'agit de régions dans lesquelles le gradient est minimum. La surface de la fonction gradient décrit en quelque sorte des bassins séparés par des crêtes. Il paraît assez logique d'identifier ces crêtes à des contours idéaux. En termes imagés, on associe un bassin $\mathcal{C}(M)$ à chaque minimum M . Si une goutte d'eau tombe en un point x de la surface, elle descendra jusqu'à se retrouver dans le fond d'un bassin –c'est un minimum de la fonction. Ainsi chaque M a sa propre zone d'attraction constituée des points à partir desquels

le minimum sera atteint. Deux bassins voisins possèdent une frontière commune, frontière à partir de laquelle une goutte d'eau pourra glisser au choix dans l'un ou l'autre des bassins. L'ensemble de ces frontières forme la *ligne de partage des eaux (LPE)*. On peut comprendre ces notions sur la figure 10.3.

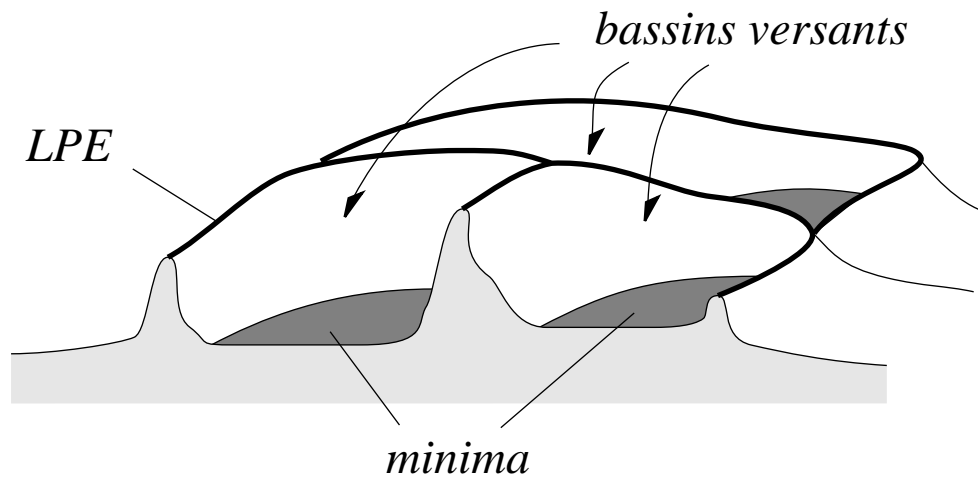


Figure 10.3: Minima, bassins versants et ligne de partage des eaux.

Une seconde interprétation est envisageable. Cette fois, le météorologiste n'annonce plus une chute de pluie mais une inondation. L'eau vient par le bas; des sources jaillissent dans le fond de la surface. Les bassins d'inondation se rejoignent et forment des frontières qui, elles aussi, sont des contours. Avant de passer à une description plus mathématique de toutes les idées précitées, nous parlerons encore d'un remède à la sur-segmentation. Une application pure et simple du schéma de détection de contours décrit conduit inéluctablement à une sur-segmentation parce que l'opérateur gradient est très sensible au bruit, ce qui a comme conséquence désastreuse d'introduire des minima artificiels et donc des bassins indésirables. Pour éviter ce désagrément, il nous faut une information supplémentaire: les centres des véritables vallées et sommets. L'introduction d'une distinction entre vrais et faux extrema modifie légèrement l'algorithme de segmentation; c'est l'étape de *marquage*.

Lors d'une montée des eaux, on construit un barrage entre deux vrais bassins; par contre, le faux bassin sera immergé par le flux d'eau en provenance de son voisin sans qu'un quelconque barrage ne soit construit. Après avoir placé le profil complètement sous eau, seuls les barrages feront office de vrais contours. N'avons-nous pas simplement déformé le problème de la sur-segmentation? La valeur du gradient n'est plus suffisante; il faut aussi des *marqueurs* différenciant les vraies et les fausses sources. En fait, le problème est souvent plus simple parce que l'analyste dispose d'une information a priori. Illustrons cela par quelques exemples:

- dans un mode interactif d'interprétation d'images où l'intéressé a la possibilité de donner des indications, la détermination des véritables extrema se reporte sur ce que fournit l'opérateur. Il n'est évidemment plus question d'interprétation automatique d'images;
- si l'image à traiter se compose d'objets de luminance fort distincte, un algorithme un tant soit peu intelligent pourra focaliser la recherche des extrema sur les centres des véritables régions;
- en recourant à une décomposition pyramidale, il est possible de franchir un premier pas dans la direction d'une automatisation de l'élimination des faux contours. Ainsi, une étape de la pyramide fournit une version filtrée et sous-échantillonnée de l'image de départ. L'image interpolée reconstruite ne contiendra plus les objets dont la taille est inférieure à celle de l'élément

structurant de filtrage. Plus généralement, il peut être souhaitable de supprimer par filtrage des régions non désirées, par exemple des régions étroites orientées dans une certaine direction;

- les ouvertures et les fermetures offrent un moyen commode de détection des pics et des vallées et elles effectuent un filtrage intéressant. L'opérateur de chapeau haut-de-forme $P(f)$ obtenu par $f - (f \circ B)$ est un opérateur toujours positif qui repère bien les pics présents dans le signal f . Par contre, les vallées seront localisées par l'opérateur suivant $V(f) = (f \bullet B) - f$, signal lui aussi toujours positif. Avec ces deux filtres, il est possible de détecter les quelques pics ou vallées accentués.

Toutes ces démarches se réduisent à exiger une information a priori concernant l'image analysée. Si la pratique s'accorde mieux à un marquage des vraies régions par un humain ou par un algorithme semi-automatique, le problème d'une détermination automatique, aveugle, universelle et intelligente est loin d'être résolu. Voyons à présent comment cet algorithme se traduit en langage morphologique.

10.3.2 Expression formelle d'un algorithme de segmentation par ligne de partage des eaux

Le cas des images binaires est introduit en premier lieu. Il sera plus aisé de généraliser ensuite le principe à des fonctions.

Analyse d'images binaires

Soit X une image binaire. Cette image contient une série de zones d'attraction centrées autour d'une famille de petits ensembles connexes Z_i . Le bassin associé au petit ensemble Z_i est le sous-ensemble de \mathcal{E} dont la distance géodésique à Z_i est plus petite que celle à Z_j . Pour parler de distance géodésique, nous introduisons quelques définitions et notations:

Définition 81 [Chemin géodésique] *Un chemin géodésique de longueur l entre deux points s et t est une famille de $l + 1$ pixels $x_0 = s, x_1, \dots, x_l = t$ telle que*

$$\forall i \in [0, l], x_i \in X \quad \text{et} \quad \forall i \in [0, l], x_{i-1}, x_i \text{ sont voisins} \quad (10.2)$$

Parmi tous les chemins, le plus court détermine la distance géodésique. La figure 10.4 fournit un exemple de chemin géodésique.

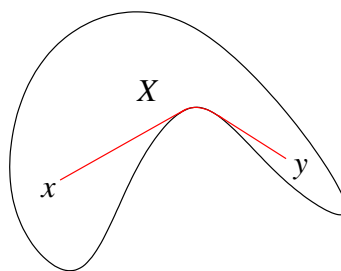


Figure 10.4: Le plus court chemin géodésique entre x et y .

Définition 82 [Distance géodésique] *La distance géodésique entre les points s et t est la longueur du plus court chemin géodésique reliant s à t ; elle est infinie si un tel chemin n'existe pas.*

Notations

La zone d'influence de l'ensemble Z_i est notée $ZI(\text{domaine} = X, \text{centre} = Z_i)$ et sa frontière $FR(\text{domaine} = X, \text{centre} = Z_i)$.

Algorithme de construction du squelette par zone d'influence. Le *squelette par zone d'influence* (SZI) s'obtient aisément par l'algorithme suivant:

- dans un premier temps, on délimite les zones Z_i des régions de base;
- pour les autres pixels, il suffit d'itérer la démarche suivante jusqu'à stabilisation du processus: si un pixel a un voisin indicé i , alors ce pixel prend le même indice; les points qui n'ont pas de voisin ou ceux qui en ont deux sont laissés inchangés;
- au terme de ce processus, les pixels sont attribués à l'une ou l'autre zone de départ Z_i .

Cet algorithme simple revient à faire croître des régions jusqu'à ce qu'elles se touchent; de la sorte, les contours obtenus sont fermés. La figure 10.5 montre un exemple de squelette géodésique.

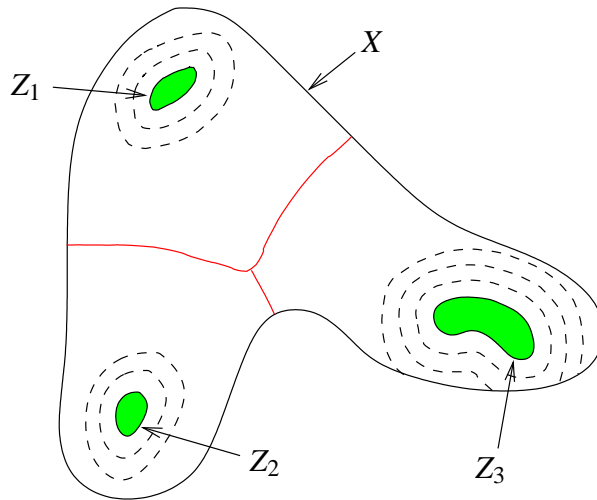


Figure 10.5: Squelette géodésique.

Images en niveaux de gris

Supposons que le relief soit progressivement immergé, l'incrément étant d'une unité de luminance par étape. C'est la clef de synthèse d'un algorithme de détermination de ligne de séparation. Le dessin de la figure 10.6 montre une étape intermédiaire du processus d'immersion.

Notations:

- f est la fonction étudiée.
- h_{min} et h_{max} sont les valeurs extrêmes de f sur son domaine de définition.
- $T_h(f) = \{x \in dom f : f(x) \leq h\}$ est l'ensemble constitué par un seuil h .
- M_i sont les minima et $\mathcal{C}(M_i)$ les bassins correspondants.

Les bassins seront construits étape par étape suite à une augmentation progressive du niveau de l'eau. Soit $\mathcal{C}_h(M_i)$ la partie du bassin M_i rempli au temps algorithmique h . Alors

$$\mathcal{C}_h(M_i) = \mathcal{C}(M_i) \cap T_h(f) \quad (10.3)$$

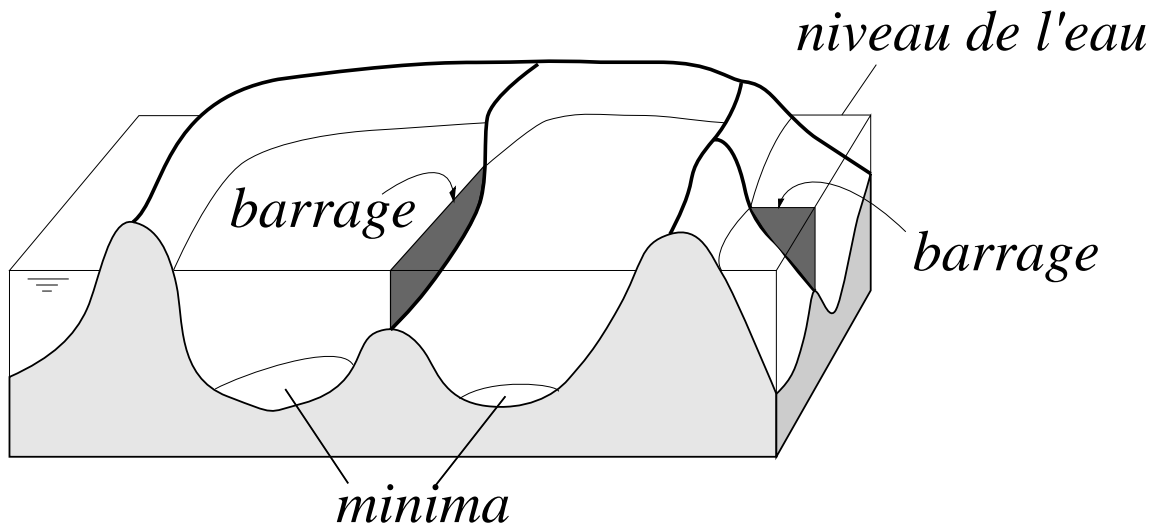


Figure 10.6: Construction d'un barrage au raccord de deux bassins versants.

Dans cette expression, $\mathcal{C}(M_i)$ est l'inconnue. Si l'altitude du minimum M_i est j , alors $\mathcal{C}_j(M_i) = M_i$ et $\mathcal{C}_k(M_i) = \emptyset$ pour $k < j$. L'union de tous les bassins est notée $\mathcal{C}(M)$ et sa partie immergée au temps h est $\mathcal{C}_h(M)$. Dans ces conditions, $\mathcal{C}_{h_{max}}(M) = \mathcal{C}(M)$. Il reste à initialiser l'algorithme et à donner un schéma de construction. Attention, il n'est pas question ici de supprimer des faux bassins.

Initialisation

- $\mathcal{C}_{h_{min}}(M) = T_{h_{min}}(f)$; l'initialisation considère que tous les points minima sont des réservoirs potentiels.

Construction

$$\forall h \in [h_{min} + 1, h_{max}] : \mathcal{C}_h(M) = ZI_h \cup Min_h \quad (10.4)$$

avec

- ZI_h = zone d'influence (de domaine $T_h(f)$) composée des centres contenus dans $\mathcal{C}_{h-1}(M)$ – c'est durant cette étape que seront déterminés les barrages traduits en contours;
- Min_h est l'ensemble des points de $T_h(f)$ qui, après le processus de recherche de zones d'influence n'ont toujours pas de voisin. Il s'agit de nouveaux minima, centres de nouvelles zones d'influence.

Cet algorithme peut subir quelques modifications destinées à supprimer les faux contours; des développements de ce type sont abondamment décrits par MEYER (le véritable initiateur de ces techniques) et BEUCHER [20]. MEYER a également développé une recherche de lignes de séparation entre bassins par analogie avec l'opération de squelettisation légèrement modifiée; de telles considérations nous mèneraient malheureusement trop loin. Remarquons simplement que l'algorithme développé est relativement complexe et que, dès lors, il n'est guère étonnant que certains auteurs (cf. [33] par exemple) aient développé des expressions plus agréables.

Marquage

Le remède communément adopté pour combattre la sur-segmentation est le *marquage*. Dans le cas présent, le marquage consiste à sélectionner certains minima seulement. On part de la constatation que le nombre de régions est généralement proportionnel au nombre de minima. Le fait de ne sélectionner qu'une partie des minima entraîne la disparition de certaines régions. En effet, lors du processus d'inondation, ces régions sont remplies par l'eau provenant d'un bassin voisin en passant par le point de crête ayant la plus faible altitude. Les bassins inondés sont alors englobés dans un bassin voisin.

Les avantages de la ligne de partage des eaux par rapport à d'autres techniques de segmentation se résument en définitive à

- la possibilité de travailler sur base d'une image de gradient quelconque,
- la possibilité d'effectuer une pré-sélection par un marquage adéquat.

Chapter 11

Codage et compression

Une description complète de ce chapitre n'est pas fournie dans ces notes.

11.1 Introduction

La compression dépend de plusieurs facteurs:

- la représentation de l'image. Il serait déjà intéressant de montrer une image et une autre image avec un autre contraste et le taux de compression sans perte des deux images.
- la perception visuelle

Définition 83 *Le taux de compression est le rapport entre le nombre de bits avant compression et le nombre de bits après compression*

$$\frac{\text{nombre de bits avant compression}}{\text{nombre de bits après compression}} \quad (11.1)$$

On peut aussi utiliser le gain de codage comme mesure de performance. Ainsi, le gain de codage [15] est l'amélioration du rapport signal à bruit due à une méthode par rapport à un codage PCM pour un même nombre de bits utilisés.

Chapter 12

Applications industrielles du traitement d'images

Auteur: YVES DAOUST

12.1 Introduction

Le traitement d'images industriel, appelé *Machine Vision* par les Anglo-Saxons, est une discipline appliquée de la vision artificielle qui emploie le traitement numérique d'images pour automatiser des tâches. Elle utilise des moyens électroniques et informatiques pour doter des machines et robots de capacités sensorielles analogues à la vision humaine. Autrement dit, elle permet d'utiliser des caméras comme capteurs pour détecter des défauts sur des objets manufacturés, inspecter des pièces en cours de fabrication, les compter, les trier, les classer, les mesurer, ... à partir de leur apparence visuelle.

Par exemple, des systèmes de vision sont intégrés sur des robots de soudure pour vérifier la continuité des cordons, sur des machines à coudre le cuir pour suivre les contours des pièces à assembler, sur des presses de moulage de CD-Roms pour vérifier l'absence de griffes, à la sortie d'imprimantes à gravure laser pour vérifier la lisibilité des numéros de série sur des composants électroniques, ... La gamme d'applications possibles est très étendue.

Il faut néanmoins se garder d'imaginer une analogie trop profonde entre les performances de la vision humaine, capable d'explorer la structure spatiale de son environnement et d'y reconnaître des objets complexes et variés, connus ou inconnus, en deux dimensions comme en trois, et les capacités bien plus modestes des systèmes de vision industriels: ces derniers ont essentiellement pour mission de mesurer des propriétés simples sur des objets et scènes connus à l'avance, pour en déduire la conformité par rapport à des modèles de référence. Bien que dans certains cas, des calculs élémentaires suffisent à résoudre les problèmes posés, tout l'arsenal des méthodes du traitement d'image vient bien à point pour parvenir à extraire l'information pertinente.

12.1.1 Spécificités du traitement d'image industriel

Le traitement d'images industriel présente une série d'exigences bien particulières:

- Il s'adresse en général à des *tâches répétitives* telles que la recherche de défauts sur de grandes surfaces de matériaux défilants, le contrôle dimensionnel de pièces identiques, le comptage de colonies bactériennes sur des boîtes de culture, ... Ces tâches sont simples et bien définies, et l'on connaît à l'avance les caractéristiques des scènes à observer et l'information à en extraire.

On peut souvent procéder par comparaison avec des modèles tels que des images d'objets sans défauts. Le degré d'intelligence et de "compréhension" des images requis est faible.

- Les approches utilisées doivent fonctionner de façon *autonome, non supervisée*. Après une phase de démarrage éventuellement réalisée par un opérateur en début de production, les systèmes doivent fonctionner sans intervention pendant de longues périodes et à une cadence rapide. Ils doivent pour cela être capables de s'adapter dans une certaine mesure et résister aux changements des conditions de fonctionnement telles que l'éclairage ambiant ou la disposition des pièces.
- Les méthodes de mesure, qui peuvent être *qualitatives* (présent/absent, nombre, bon/mauvais...) ou *quantitatives* (longueur, diamètre, couleur moyenne, rugosité, ...), doivent être fiables, précises et reproductibles. Cela implique une prise en compte correcte des variations du contenu de l'image inhérentes au processus observé et une exploitation judicieuse des détails pertinents dans les images.
- Les traitements doivent suivre la *cadence du temps réel* imposé par la chaîne de fabrication. Étant données les hautes vitesses de défilement souhaitées (souvent comptées en pièces ou dizaines de pièces par secondes) et le gros volume de données à traiter (de l'ordre du mégabyte par image), des processeurs puissants combinés à des algorithmes simples et efficaces sont requis. Si nécessaire, on peut avoir recours à du hardware spécialisé pour accélérer les calculs.
- Dans la plupart des cas, on se ramène à des problèmes à deux dimensions en choisissant judicieusement le point de vue et la disposition des objets pour se ramener à des cas plans. Le travail en trois dimensions nécessite en effet une reconstruction de l'information de profondeur absente des images, qui fait appel à des techniques de triangulation spatiale telles que la stéréoscopie (emploi de deux images) ou l'éclairage structuré (motif projeté sur les objets observés). Ces techniques sont relativement délicates à mettre en oeuvre.
- On se contente également la plupart du temps d'utiliser des images en niveaux de gris. Celles-ci permettent en effet souvent de voir les caractéristiques pertinentes des objets et suffisent pour les besoins courants. D'autre part, les images en couleur sont plus volumineuses (trois fois plus en général) et les valeurs associées aux pixels doivent être vues comme des vecteurs plutôt que des scalaires. Le surcroît de complexité ainsi introduit n'est pas toujours justifiable. L'emploi de la couleur s'impose uniquement lorsqu'elle permet de révéler des contrastes invisibles en niveaux de gris. Dans la pratique, les niveaux de gris sont quantifiés sur 8 bits (256 niveaux distincts); exceptionnellement, 10 ou 12 bits sont utilisés.
- On a souvent la liberté d'agir sur l'environnement pour faciliter la tâche du système d'inspection. On peut par exemple disposer judicieusement les sources d'éclairage pour éviter les ombres portées ou au contraire les renforcer, on peut améliorer le contraste en choisissant la couleur du fond, on peut placer des repères faciles à localiser pour mesurer les positions, ...
- Enfin, les industriels apprécient la facilité d'utilisation, d'intégration dans des automatismes existants. Les méthodes et outils proposés doivent rester simples d'emploi et ne pas nécessiter les compétences de spécialistes en vision.

Les avantages des systèmes de vision par rapport à d'autres types de capteurs sont nombreux:

- Ils sont souples d'emploi et permettent une très grande variété d'utilisations. Par simple reprogrammation, un capteur de vision peut changer radicalement de type d'inspection et peut s'adapter instantanément à d'autres pièces.
- Ils permettent de réaliser simultanément plusieurs mesures, éventuellement de nature différente.
- Ils permettent des mesures sans contact et non destructives, avec une vitesse et une précision appréciables.
- Ils permettent d'objectiver certaines mesures et de garantir ainsi la répétabilité de l'apparence des objets (constance de la couleur, de la forme, ...).
- Plus que tous dispositifs automatiques, ils permettent de remplacer des opérateurs humains dans des tâches fastidieuses qui requièrent attention et concentration.

12.1.2 Applications typiques

De nombreux secteurs industriels sont concernés par la vision artificielle.

On peut citer notamment les procédés où défilent à grande vitesse des matériaux en feuille et en rouleaux qui doivent être exempts de défauts tels que taches, griffes, trous, plis, reliefs et autres hétérogénéités. Ce type d'application se rencontre dans l'industrie papetière, dans la sidérurgie, la fabrication du verre plat, l'industrie textile, ... Dans les cas les plus simples, il s'agit de surfaces parfaitement lisses et uniformes où il suffit de détecter des changements de couleur ponctuels; dans d'autres, on est confronté à des textures périodiques ou non, et il est plus délicat de séparer les variations normales et accidentelles.



Figure 12.1: Industrie du papier: détection de taches, plis et déchirures.

L'inspection au défilé de particules de matériaux pulvérulents, de poudres, concassés et granulats permet leur caractérisation granulométrique et l'élimination d'éventuelles impuretés et le tri. Ceci concerne les producteurs de matières premières tels que les minéraliers, les cimentiers, les recycleurs de déchets, ...



Figure 12.2: Industrie métallurgique: inspection des tôles et caractérisation des alliages.

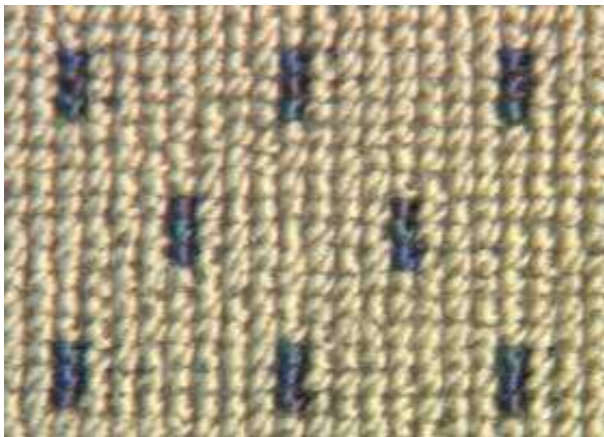


Figure 12.3: Industrie textile: contrôle des couleurs, détection des défauts de tissage.

Dans d'autres cas, le matériau défilant peut présenter un motif particulier dont on veut vérifier la conformité. C'est le cas par exemple en imprimerie où l'on veut contrôler la qualité des couleurs et l'opacité des encres, l'absence de bavures et autres défauts d'encre, ou le désalignement des matrices cyan, magenta, jaune et noire et plus généralement la conformité du motif ou du texte imprimé.



Figure 12.4: Imprimerie: mesure des couleurs, alignement des composantes, défauts d'encre, griffes des matrices.

On est également amené à inspecter des pièces et des assemblages de pièces. Dans l'industrie verrière, on contrôle des flacons et bouteilles pour vérifier leurs dimensions et l'absence de défauts qui compromettent leur esthétique (bulles, inclusions et impuretés) ou leur résistance (fissures). Dans l'industrie de la fabrication des composants électroniques, on guide les outils de façonnage au moyen de repères optiques (marques "fiduciaires"); on surveille les différentes phases de l'encapsulation dans un boîtier, de la pose du "chip" sur une embase à la mise en forme des pattes, en passant par l'apposition du logo du fabricant. Dans l'industrie du montage de circuits imprimés, on utilise la vision pour permettre aux robots de saisir les composants en amenant leurs pattes rigoureusement en regard des trous de fixation prévus à cet effet.



Figure 12.5: Industrie du verre: contrôle de flacons.

Dans l'industrie alimentaire, on utilise des caméras pour juger du degré de maturité de fruits et légumes et les classer en catégories selon la quantité de taches, coups, tavelures et autres irrégularités

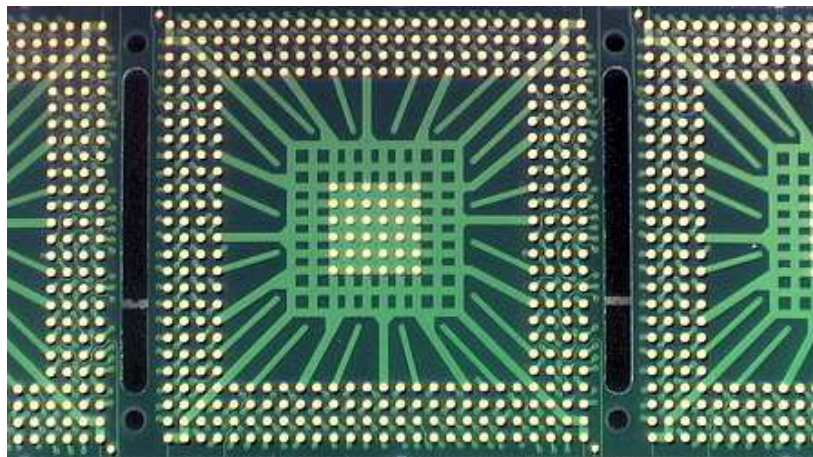
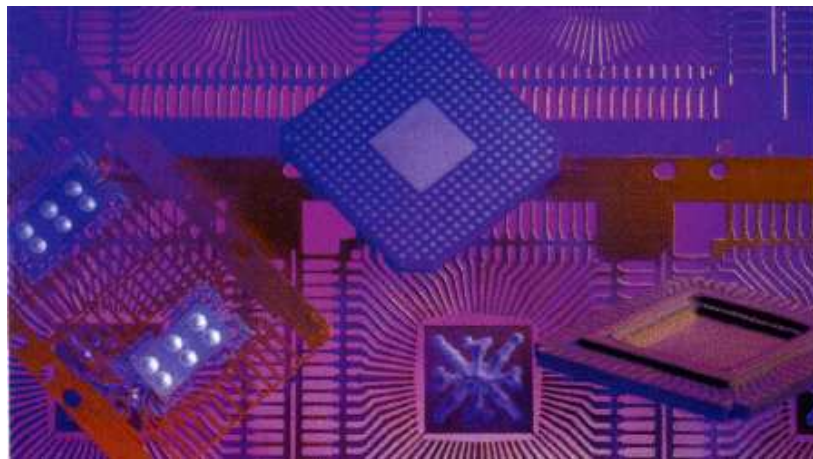


Figure 12.6: Fabrication de composants électroniques: moulage, découpage des pattes, formation des pattes, inspection des billes.

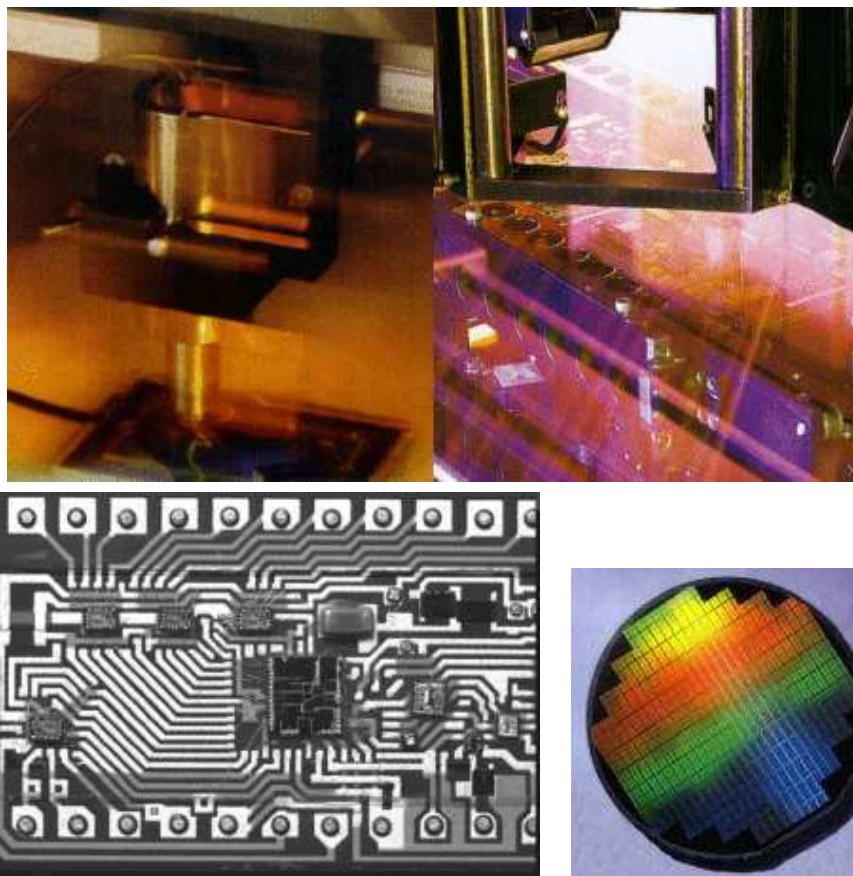


Figure 12.7: Fabrication de composants électroniques: découpage du *wafer*, placement des *dies* et soudage des contacts.



Figure 12.8: Montage de circuits imprimés (*pick & place*).

inesthétiques. On observe des pains à la sortie des fours pour contrôler leur taille, leur forme et l'absence de brûlures ou traces indésirables, ... Bref, les applications sont innombrables.

D'une façon générale, on peut distinguer les cas des matériaux naturels (en particulier issus du milieu vivant) et artificiels (résultant d'un processus de fabrication). Les premiers montrent en général des aspects beaucoup plus variés et changeants, et des formes moins régulières que les seconds, ce qui complique leur inspection; on recourt pour eux à des méthodes plus qualitatives que quantitatives. Il faut également avouer que les méthodes appliquées aux matériaux naturels doivent plus souvent être supervisées (intervention d'un opérateur) et sont réservées à l'utilisation en laboratoire plutôt qu'à l'atelier.



Figure 12.9: Industrie automobile: contrôle des soudures.

12.1.3 Tâches d'inspection

De cette variété d'applications, on peut néanmoins dégager un ensemble de procédés utilisés de façon récurrente et combinés pour résoudre complètement les problèmes de vision posés.

Dans une grande catégorie de cas, on ne connaît pas à l'avance le contenu de l'image, ou seulement qualitativement, et on se propose de le quantifier: lorsqu'il s'agit d'objets en vrac ou de défauts imprévisibles, on est amené à les compter. Cette tâche est triviale pour des objets simples et bien contrastés. Elle peut se compliquer si les objets présentent une structure interne, qu'ils se touchent ou se chevauchent et se distinguent mal du fond et si leur aspect est très variable.

L'extraction des objets dans une image est une tâche difficile dans le cas général. On s'arrange, par des artifices d'éclairage ou autres, pour créer un contraste important entre les objets à observer et le fond. De cette façon, une simple opération de seuillage permet de les séparer. Sinon, des techniques plus élaborées doivent être employées.

Lorsqu'on est capable d'isoler les objets de leur contexte, on peut décrire leur forme, leur taille, leur couleur, ... au moyen de mesurandes bien choisis. C'est notamment le cas de la recherche de défauts de surface, qu'ils soient contrastés, colorés ou en relief, ... On peut également faire des études statistiques de la répartition des valeurs associées. On caractérise ainsi des populations entières.

Par comparaison des paramètres mesurés avec des valeurs de référence, on peut ainsi classer les objets individuels ou populations d'échantillons en catégories pré-définies telles que "normal" et



112 Postgroot 1000 Karlsruhe 1.05.91 *Stegwe*

WEBER MARTIN

134228 8 0026 68050101

SPARKASSE FREIBURG

22 12.50

RIETE OSIBI

WEYICH, GEORGES, DR. *7800 FREIBURG

282404754

30226324000267 02824047544

BSL 204

Sparkasse Überall *Für BSE - MFL.*

ID *Weyich*

Muster *DM*

SPECIMEN

Muster *Stegwe*

Muster *Stegwe*

000000000003254 12354678 114

Figure 12.10: Tri postal automatisé.

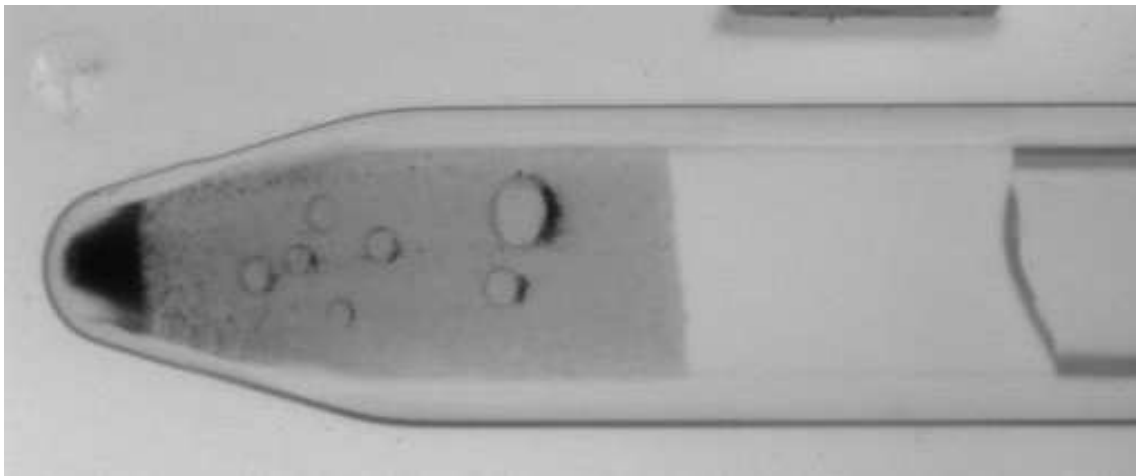


Figure 12.11: Imagerie biomédicale: analyse automatisée, comptage de cellules.

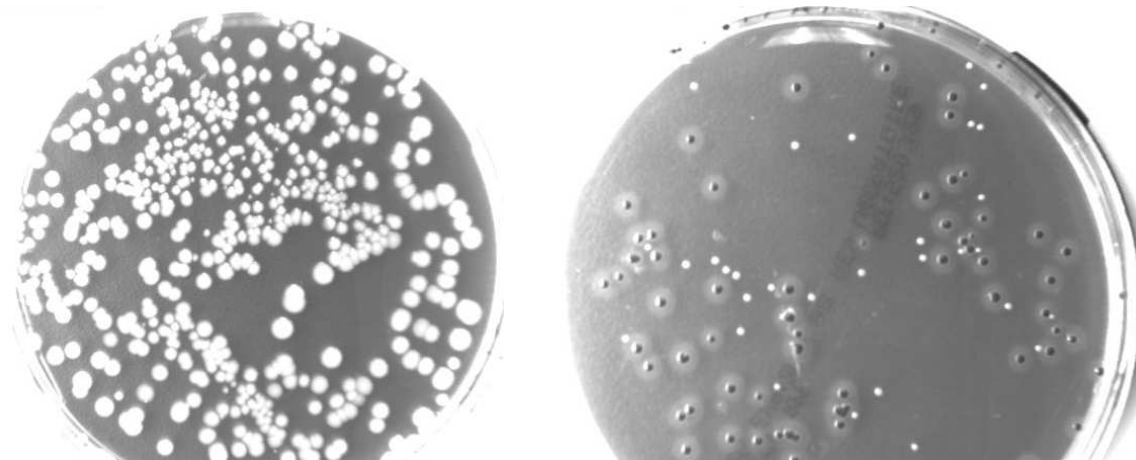


Figure 12.12: Comptage.

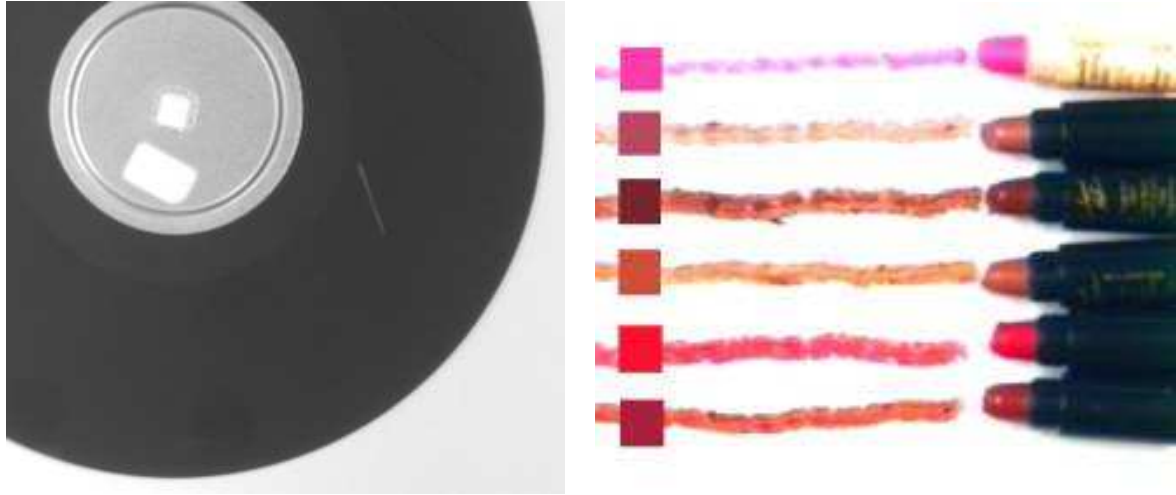


Figure 12.13: Inspection de surfaces: détection de défauts, colorimétrie.

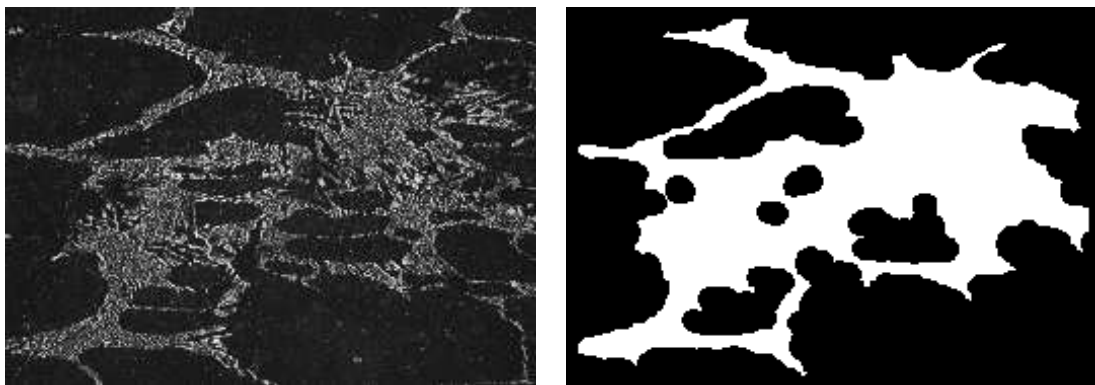


Figure 12.14: Contrôle de la forme et morphométrie.

“pathologique”, “fin”, “mi-fin” et “grossier”

Lorsqu’il s’agit de pièces de forme et dimensions bien connues à l’avance, les mesures peuvent être plus fines: un contrôle détaillé des cotes et de la forme des contours devient possible. Ces informations peuvent être comparées à des gabarits ou intervalles d’acceptation qui précisent les limites de variations tolérées.

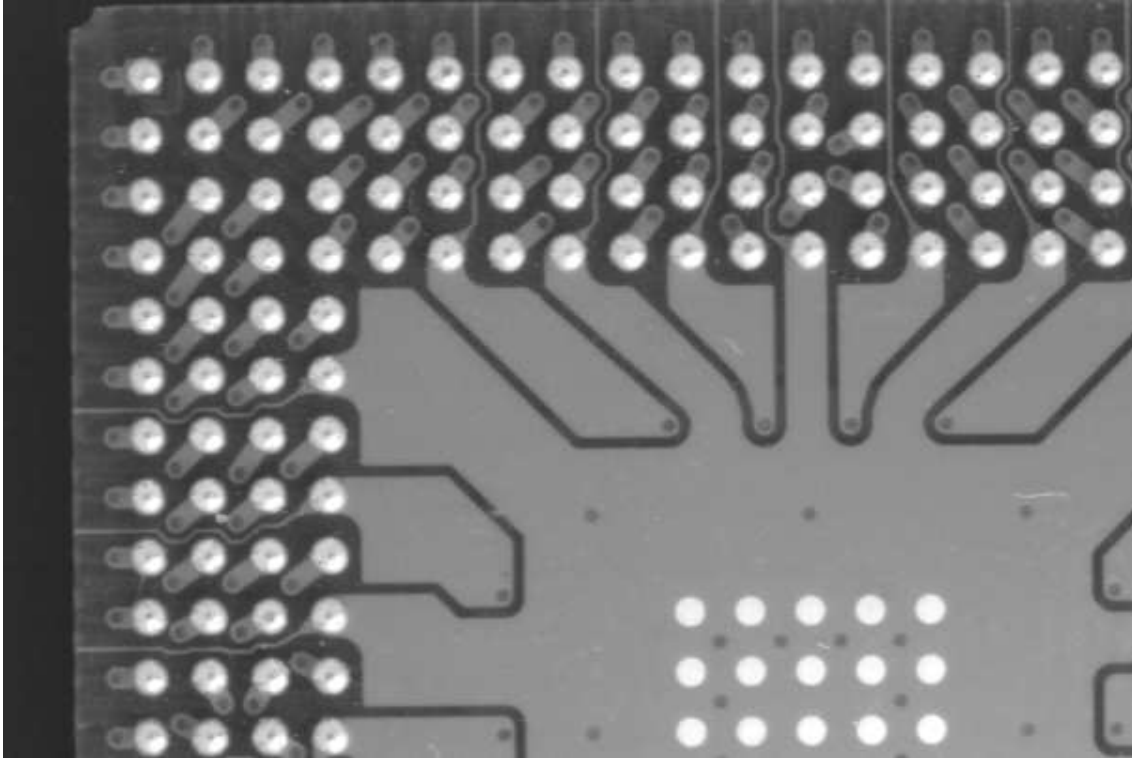


Figure 12.15: Contrôle de la taille de pièces.

Le contrôle dimensionnel suppose que les éléments à mesurer sont placés dans une position connue. Des techniques particulières permettent de localiser avec précision des objets connus dans une scène complexe, pour autant qu’on ait présenté au système un exemplaire représentatif. De tels outils permettent de réaliser le comptage ou la localisation précise en vue d’une inspection plus approfondie.

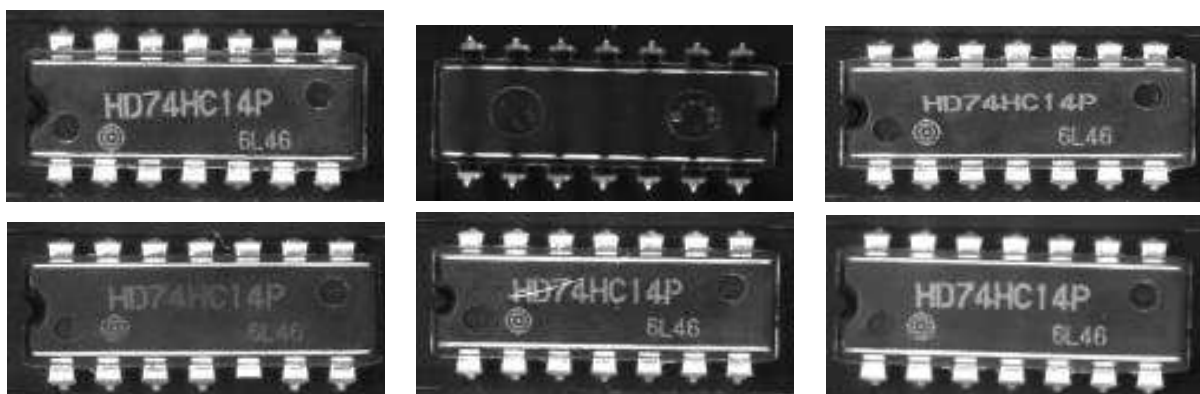


Figure 12.16: Inspection de marquage: lecture de caractères, contrôle de qualité.

Dans le cas d’assemblages, on peut de même observer les emplacements prévus pour les différentes pièces du montage et contrôler leur présence, leur bon positionnement et éventuellement le bon choix du modèle.

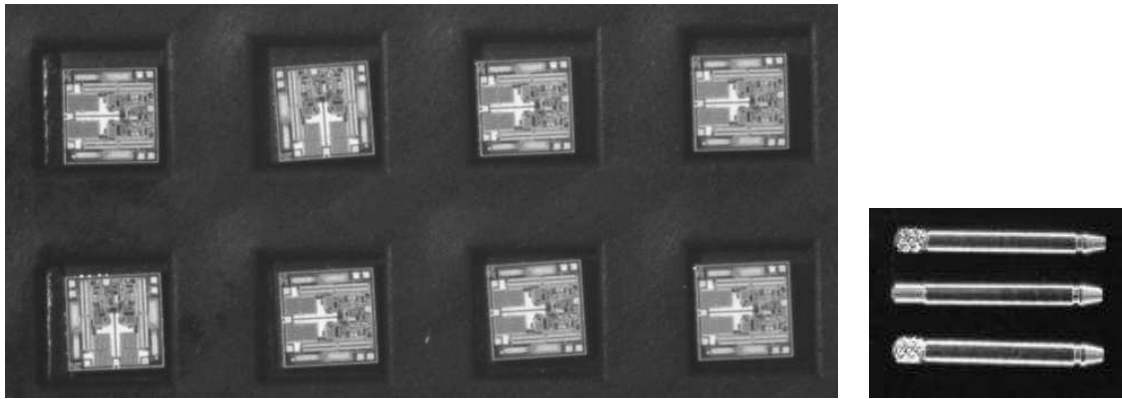


Figure 12.17: Contrôle d'assemblage: positionnement, présence/absence.

La plupart de ces tâches peuvent être menées à bien à partir d'images en noir et blanc. Elles nécessitent cependant que le contraste, en terme de luminosité, entre les objets à observer soit suffisant. La caractérisation de la couleur des surfaces sur base d'une échelle de gris (du foncé au clair) est relativement grossière. L'utilisation de plusieurs bandes spectrales, et en particulier la traditionnelle décomposition en canaux rouge, vert et bleu, enrichit de façon notable les possibilités de mesure. En particulier, on peut se livrer à une classification sur base d'une description qualitative ou quantitative de la couleur (défauts plutôt rouges, plutôt jaunes, ...).



Figure 12.18: Lectures de codes: codes matriciels, codes à barres et codes de couleurs.

Le traitement en couleur est plus exigeant à de nombreux points de vue. Il requiert en particulier que l'on se préoccupe de la stabilité des couleurs et de la fidélité de leur mesure. Par exemple, dans l'industrie de l'imprimerie, il est utile de pouvoir comparer les couleurs imprimées aux couleurs affichées à l'écran du système de composition de pages. On utilise pour ce faire des systèmes de représentation des couleurs indépendants du matériel.

12.2 Moyens matériels

12.2.1 Chaîne d'acquisition et de traitement

Les systèmes de vision industrielle sont constitués d'un certain nombre d'éléments matériels et logiciels dont on doit tenir compte et choisir judicieusement au moment de la conception, en fonction des exigences et contraintes de l'application visée.

Considérons ces éléments dans leur ordre d'apparition le long de la chaîne de formation de l'image (cf. figures 12.19 et 12.20).

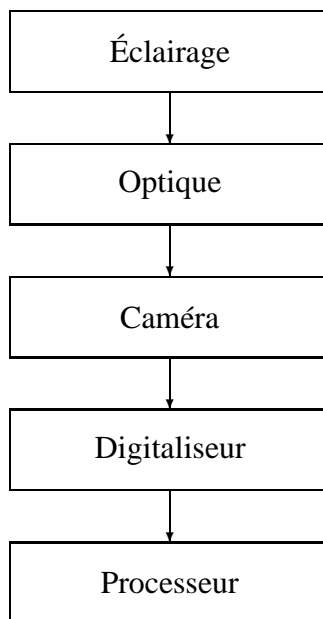


Figure 12.19: Chaîne de formation de l'image.

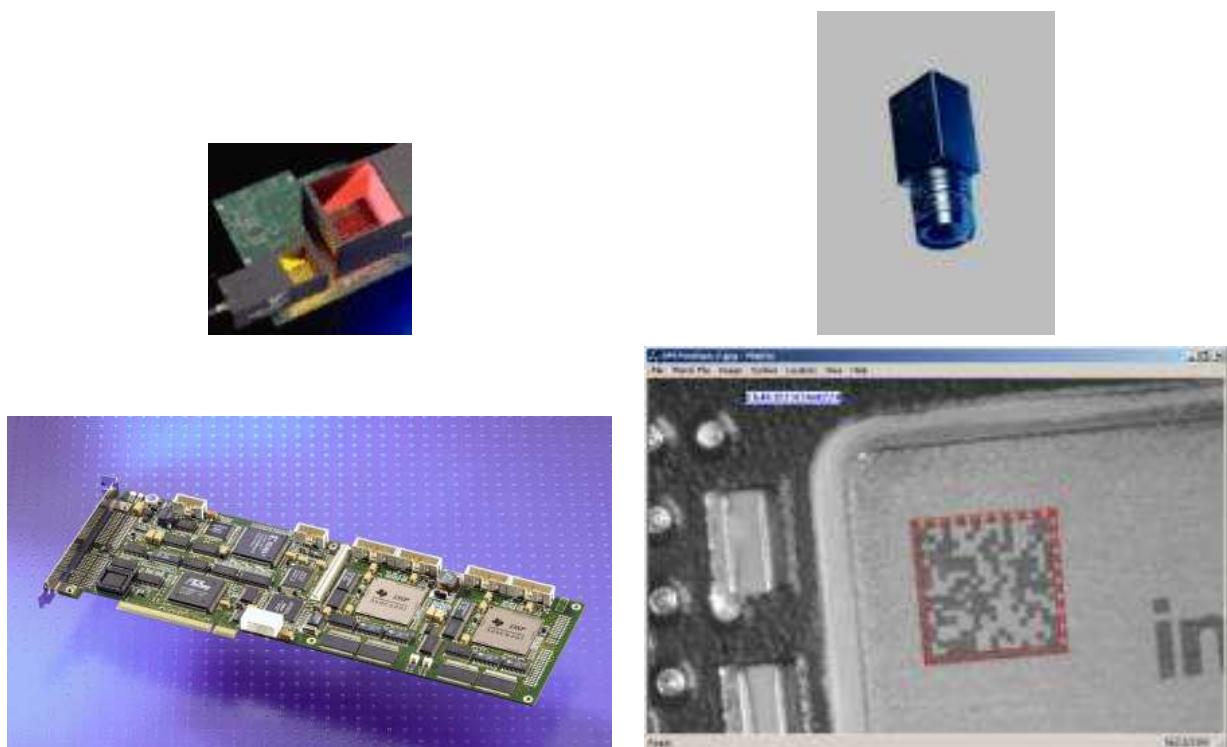


Figure 12.20: Éléments d'une chaîne d'acquisition: éclairage, caméra, digitaliseurs, logiciel.

Éclairage

En premier lieu, il est important de garder en mémoire qu'une image résulte de l'interaction entre la lumière émise par une ou plusieurs sources d'éclairage, et la surface des objets à inspecter et

de leur environnement, qui réfléchissent cette lumière vers l'observateur. L'effet résultant est donc bien la combinaison de la distribution de la lumière incidente et des propriétés de réflectance des surfaces. De façon schématique, on peut dire qu'une image permet de voir les sources lumineuses qui se réfléchissent sur la scène observée.

Cette remarque permet de comprendre combien le choix des sources lumineuses et de leur disposition est un élément essentiel du succès d'une solution. Si l'éclairage est bien choisi, les caractéristiques utiles seront faciles à mesurer; si l'éclairage est déficient, l'application échouera. En général, on ne peut pas agir sur les propriétés des objets à inspecter. On peut par contre ajuster les sources lumineuses à volonté.

À l'heure actuelle, plusieurs types de sources d'éclairage ont la faveur (cf. figure 12.21):

- Les **ampoules à incandescence** sont celles utilisées pour l'éclairage domestique. Elles sont bon marché et fournissent une lumière diffuse et blanche. Leur rendement énergétique n'est cependant pas excellent et leur vieillissement est sensible.
- Les **tubes fluorescents** ont un meilleur rendement, mais délivrent une lumière dont la température de couleur est moins bonne.
- Les **lampes halogènes** permettent d'atteindre de fortes luminosités à coût raisonnable.
- Les **diodes électroluminescentes (LEDs)** présentent une excellente durée de vie et peuvent être conçues pour produire de la lumière colorée ou blanche.
- Enfin, les **lampes à quartz** sont utilisées pour fabriquer des flashes.



Figure 12.21: Sources d'éclairage: lampe à incandescence, diodes électroluminescentes, flash, tubes fluorescents.

Dans la plupart des cas, on fabrique des sources lumineuses de formes variées (plaques, barrettes, anneaux, cylindres, sphères, cloches, ...) soit par assemblage de plusieurs ampoules, soit en utilisant des fibres optiques comme guides de lumière. Ces différentes dispositions permettent de distribuer la lumière de la façon la plus adéquate.

On distingue l'éclairage directionnel où la direction d'incidence de la lumière est bien définie, de façon telle que des ombres portées apparaissent nettement, et l'éclairage diffus, souvent obtenu par réflexion ou transmission sur des surfaces mates, qui atténue les ombres et adoucit le rendu.

La position de l'éclairage par rapport à l'observateur permet d'obtenir des effets variés. Lorsque les objets sont éclairés dans une direction proche de l'observateur, les surfaces sont vues avec leur couleur habituelle et le contraste qu'elles créent est lié à leur nature. On parle alors d'*éclairage en champ clair*. Le cas extrême est obtenu au moyen d'éclairages coaxiaux (la lumière est amenée

sur l'axe optique au moyen d'un miroir semi-réfléchissant), qui ont la propriété d'éviter toute zone d'ombre dans l'image.

Lorsque les objets sont éclairés selon une incidence rasante, les ombres portées par les accidents et la microstructure des surfaces donnent une apparence très différente qui fait ressortir le relief et la texture des objets plutôt que leur teinte. On parle alors de *champ sombre*.

Lorsque les objets sont éclairés par l'arrière, seule leur silhouette est perceptible et elle se détache avec un fort contraste sur un fond clair. On parle dans ce cas d'*ombroscopie*. Seule la forme des objets apparaît.

Enfin, certaines applications spéciales, telles que le relevé tridimensionnel de surfaces, nécessitent des procédés spéciaux, dits d'*éclairage structuré*, qui consistent à éclairer sélectivement en projetant des motifs géométriques judicieusement choisis. Par exemple, l'emploi d'un "plan laser" permet de tracer des lignes sur les objets et d'en déduire la forme spatiale en les observant de biais.

Il paraît naturel d'éclairer les objets en lumière blanche, puisque c'est ainsi que nous observons le monde qui nous entoure. Lorsque l'on utilise des caméras en noir et blanc, il est cependant fréquent d'éclairer en lumière rouge, car cette gamme de longueur d'onde correspond à une plage où la sensibilité du silicium à la lumière est meilleure. D'autres couleurs sont parfois utilisées pour maximiser le contraste ou pour d'autres raisons. On utilise également le rayonnement infrarouge qui a l'avantage d'être invisible à l'œil humain.

L'expérience montre que l'apparence des objets vus par un système d'imagerie peut s'écarter fortement de notre expérience de tous les jours, en particulier en raison des procédés d'éclairage, et il est bon d'en connaître toutes les possibilités.

Optique

Le choix correct de l'optique dont la caméra est équipée est également crucial. Il détermine pour une large part la qualité de l'image en termes de netteté, contraste et absence de déformations. Il est soumis à des contraintes telles que l'encombrement maximum permis, qui tient compte de la taille de l'objectif et de la distance d'observation. On retiendra que les optiques courantes sont loin d'égaliser les performances de l'œil humain qui dispose notamment d'une large faculté d'auto-adaptation (mise au point sur une large gamme de distances, auto-focalisation, auto-ajustement d'ouverture, ...).

Les caméras industrielles sont munies d'objectifs similaires à ceux utilisés en photographie, à une ou plusieurs lentilles. Les paramètres essentiels qui guident leur choix sont le *grossissement*, le *pouvoir séparateur* (mesuré en nombre de lignes résolues par unité de longueur) qui fixe la finesse des détails mesurables, la distance d'observation qui détermine l'encombrement, et la *profondeur de champ* (gamme d'éloignements entre lesquels l'image est nette). Des caractéristiques secondaires tels que l'absence de *distorsion géométrique* ou *chromatique* font la différence entre optiques de mauvaise et bonne qualité.

Les objectifs peuvent avoir des distances *focales fixes* ou *ajustables* (on parle alors de *zooms*). Les focales courantes s'étendent de courtes distances (quelques millimètres) dans le cas d'objectifs macrophoto ou de grand angles, à de longues distances (quelques décimètres) dans le cas des téléobjectifs. A l'extrême, on trouve les *optiques télécentriques*, correspondant à une focale infinie, qui donnent une image de la même taille que l'objet. Les petites focales correspondent à un plus large champ visuel mais donnent lieu à plus d'aberrations et distorsions optiques.

Les objectifs sont normalement munis d'un *obturateur* mécanique, qui permet d'ajuster la quantité de lumière entrante pour régler la luminosité de l'image. Plus l'obturateur est fermé, meilleure est la profondeur de champ, et vice-versa. Comme nous le verrons, des moyens électroniques intégrés à la caméra permettent également d'ajuster la luminosité et le contraste.

Parfois, des éléments optiques additionnels sont utilisés: miroirs et prismes pour modifier le

chemin optique pour réduire l'encombrement, filtres polarisants pour éliminer les reflets, écrans pour éliminer la lumière parasite, ...



Figure 12.22: Optiques de caméra et obturateurs.

Caméra

Les caméras industrielles actuelles sont constituées par des boîtiers compacts sur lesquels se visent les objectifs et qui contiennent un capteur photosensible logé derrière un verre protecteur. Le rôle du capteur est de transformer la quantité de lumière incidente en une charge électrique par effet photo-électrique, et de convertir cette charge en une grandeur électrique mesurable. Les capteurs sont décomposés en petites cellules jointives indépendantes, qui reçoivent chacune une petite partie de l'image, c'est-à-dire un pixel.

Dans la grande majorité des cas, les capteurs sont réalisés au moyen de la technologie *CCD* (*Charge Coupled Device*), qui permet de véhiculer les charges accumulées dans les éléments sensibles au moyen de registres à décalage. Le courant correspondant est converti en une tension qui est mise en forme et combinée avec des impulsions de contrôle pour constituer un *signal vidéo*. Une autre technologie entre en concurrence avec la précédente: il s'agit de l'intégration *CMOS* (*Complementary Metal-Oxyde semiconductors*), qui permet de réaliser des capteurs avec une technologie très proche de celle des mémoires RAM (accès aléatoire aux pixels) tout en profitant du faible coût de fabrication de ce type de circuits.

On distingue deux grandes familles de caméras, selon la disposition des éléments sensibles:

1. Les **caméras matricielles** comportent un réseau d'éléments sensibles à deux dimensions, qui couvre la totalité du champ visuel. Ces matrices de pixels sont balayées ligne par ligne et colonne par colonne pour transmettre les pixels en série. Les résolutions couramment utilisées vont de 640×480 ou 768×576 pixels pour les besoins standard à 1280×1024 pour la haute résolution, voire même 2048×2048 pixels et plus dans le très haut de gamme.
2. L'autre famille est celle des **caméras linéaires**, où les éléments sensibles sont répartis en ligne uniquement. De telles caméras nécessitent un déplacement relatif régulier entre la caméra et la scène, souvent naturellement présent dans les chaînes de fabrication automatisées, mais permettent de travailler à des résolutions élevées (1024, 2048, 4098 pixels et plus). Elles nécessitent cependant une parfaite synchronisation entre le mouvement et l'acquisition des lignes et demandent une beaucoup plus grande quantité de lumière. En effet, les éléments sensibles ne sont exposés que le temps nécessaire au balayage d'un interligne, alors que dans le cas matriciel, la période d'exposition de l'image complète est disponible.

Traditionnellement, les caméras industrielles délivraient un signal vidéo sous forme *analogique*, calqué sur le mode de transmission des images télévisées. Ce mode de codage présente l'énorme

avantage de ne nécessiter qu'une paire de fils pour la transmission, ce qui simplifie le câblage à l'extrême. En contrepartie, il nécessite un échantillonnage suivi d'une conversion analogique vers numérique susceptibles de dégrader le signal et d'introduire des erreurs de phase (*jitter*).

Les caméras plus récentes sont dites *numériques* car elles délivrent directement un signal numérisé véhiculé sur une liaison série ou parallèle. Dans le cas de la liaison parallèle, le nombre de bits à transmettre (bits de données, signaux de contrôle et de synchronisation) peuvent requérir un nombre de fils prohibitif. Dans le cas de liaisons en série, la fréquence de transmission doit par contre être très élevée.

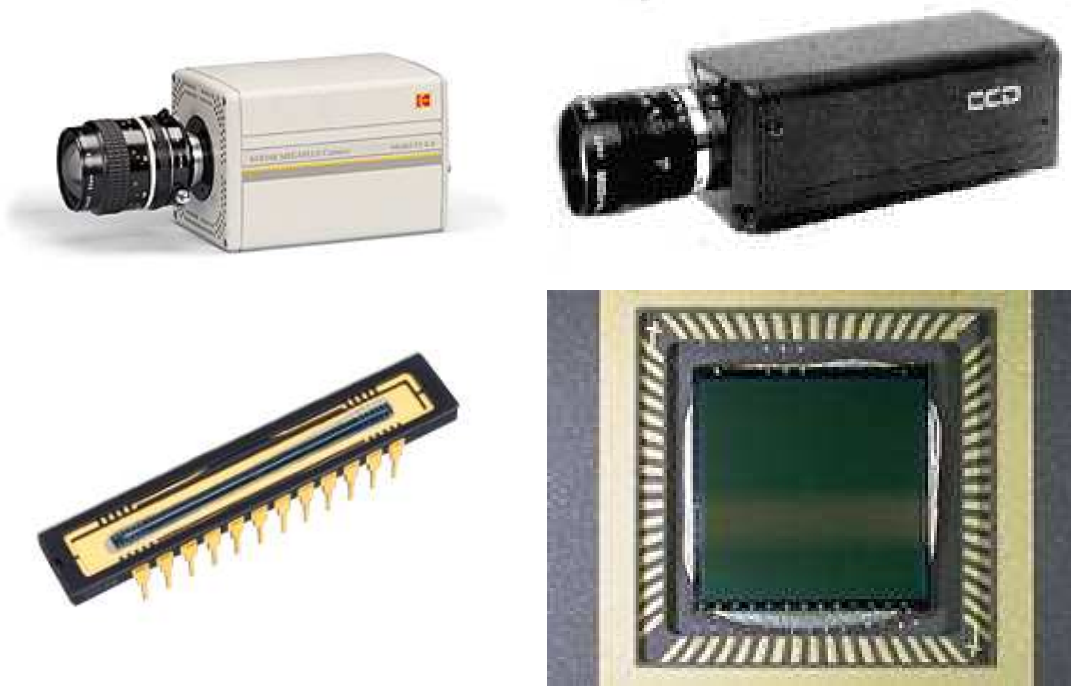


Figure 12.23: Caméras: capteurs CCD et CMOS.

Carte d'interface

Que les caméras délivrent un signal analogique ou numérique, le contenu de l'image doit être transféré dans la mémoire de travail d'une unité de traitement. Ce transfert est réalisé par un hardware d'interface spécialisé, appelé *frame grabber*, qui joue plusieurs rôles. Il récolte les signaux en provenance de la caméra, essentiellement liés au contenu de l'image et à la synchronisation, et les met en forme pour le processeur. S'il s'agit d'un signal analogique, celui-ci est échantillonné et numérisé; s'il s'agit d'un signal numérique en série, il est "*désérialisé*"; s'il s'agit d'un signal numérique en parallèle, il est pratiquement transmis tel quel.

Après éventuelle numérisation, les pixels sont stockés dans des *mémoires tampons* intermédiaires. Enfin, le contenu de ces mémoires est transféré vers la mémoire principale par blocs, pour perturber le moins possible le fonctionnement du processeur en monopolisant le bus d'accès mémoire.

Éventuellement, le frame grabber est capable de réaliser des traitements élémentaires tels que des filtres spatiaux, l'étirement de la dynamique (gain et offset), la séparation des couleurs, ... Ces traitements sont la plupart du temps appliqués au signal analogique.

Enfin, les modèles haut de gamme peuvent être munis de processeurs dédiés qui peuvent appliquer des pré-traitements "au vol" avant de délivrer l'image numérique. On peut citer par exemple la

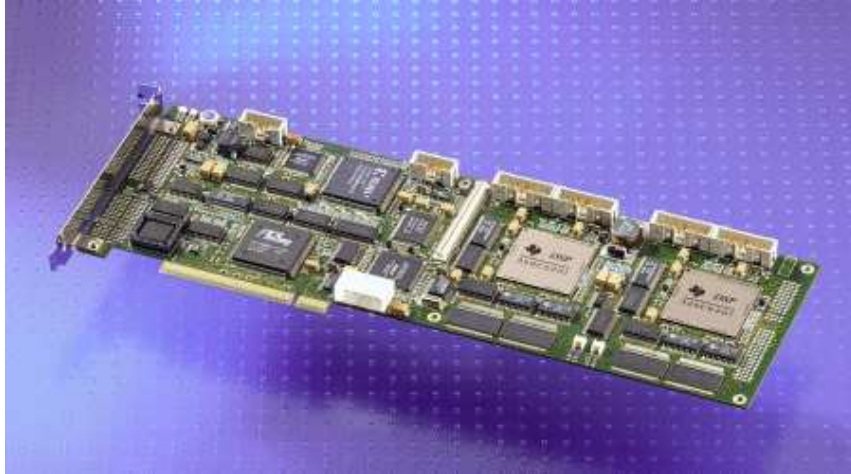


Figure 12.24: Carte d'acquisition vidéo.

compression, la correction de la balance des blancs, la correction de non-uniformité de réponse du capteur, ...

Processeur

Par le passé, les systèmes informatiques à *microprocesseur* ne disposaient pas de suffisamment de puissance de calcul pour être capables de servir à eux seuls les besoins des applications de traitement d'image. Il suffit de se rappeler qu'une image de taille standard comporte 640×480 ou 768×576 pixels codés sur un ou trois bytes, et que les images sont produites à la cadence de 25 par seconde. Cela représente un flux de données de $7,7 [Mbytes/s]$ à $33 [Mbytes/s]!$

Pour cette raison, de nombreux systèmes ont été construits sur base de hardware spécialisé, combinant des processeurs à usage général, des processeurs de traitement du signal (*DSPs*) et des unités arithmétiques sur mesure. L'évolution spectaculaire de la puissance de traitement des ordinateurs personnels (PCs) les a rendus tout à fait aptes à remplir ce rôle.

A l'heure actuelle, la majorité des systèmes industriels sont donc conçus sur base de *PCs industriels* ou non, qui abritent des frame grabbers. Certains constructeurs fournissent des systèmes propriétaires "tous faits", souvent construits autour d'un PC simplifié.

Caméras intelligentes

Profitant des progrès de la miniaturisation, certains fabricants sont parvenus à combiner dans un seul boîtier compact les fonctions de caméra, carte d'acquisition et unité de traitement. Les *caméras intelligentes* ainsi réalisées sont programmables et ne délivrent pas d'image: elles fournissent directement les résultats du traitement, comme par exemple le contenu d'un code à barres ou un diagnostic d'inspection.

12.2.2 Défauts introduits par la chaîne d'acquisition

Chacun des éléments qui viennent d'être cités introduit des dégradations et déperditions qui peuvent être dommageables. Étudions les au cas par cas et voyons quelles mesures correctives peuvent être prises.

Éclairage

L'image qui arrive au capteur de la caméra résulte de l'interaction de la lumière avec la scène observée. Lorsque la lumière est distribuée de façon parfaitement uniforme et homogène, son effet ne se fait pas sentir. Malheureusement, il est très difficile en pratique d'assurer cette uniformité, et la constance de l'intensité au cours du temps. De plus, des effets tels que les ombres portées, qui occultent une partie de la scène, ou les reflets, qui éblouissent le capteur, détruisent par endroit le signal.

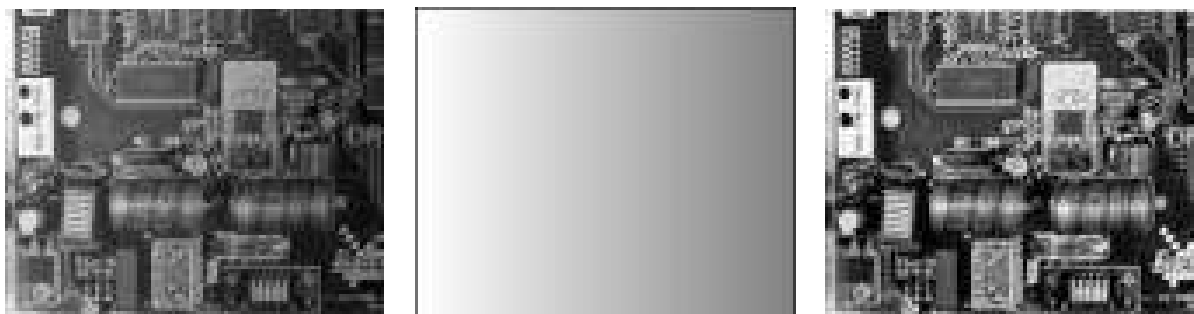


Figure 12.25: Défaut d'éclairage non-uniforme.

Les ombres peuvent être combattues par la disposition judicieuse des sources, à moins que leur présence ne soit désirée; l'éclairage diffus est une solution possible; les reflets peuvent être atténués par le recours à la lumière polarisée; la non-uniformité de l'éclairage peut être compensée, en mesurant à l'avance le champ lumineux sur un écran blanc et en recalibrant l'échelle des gris point à point. Si cela n'est pas possible ou si la non-uniformité est variable, un filtrage passe-haut peut améliorer la situation. La variation de l'intensité lumineuse peut être corrigée par des techniques d'analyse d'histogramme telles que, par exemple, la normalisation des niveaux de gris par la moyenne et l'écart-type.

Optique

Les imperfections et mauvais réglages de l'optique introduisent des distorsions dues aux aberrations et à l'effet de perspective, du flou dû à la défocalisation ou à la diffraction, de la non-uniformité de luminosité transmise, ...

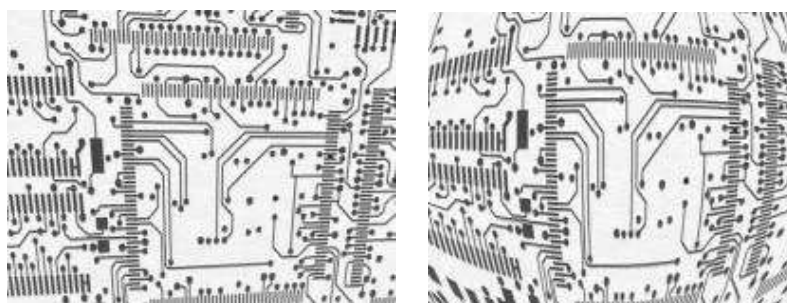


Figure 12.26: Défauts d'optique: perspective et distorsion.

Les distorsions, lorsqu'elles sont permanentes, peuvent être éliminées par calibration géométrique du champ visuel: en observant l'image d'une mire de forme connue, on peut mesurer la distorsion et la compenser au moyen de la transformation inverse. Dans certains cas, c'est l'objet mesuré lui-même qui peut servir de mire.

Le flou peut être compensé par des techniques de filtrage inverse. On distingue à ce propos les techniques de restauration standard, basées sur la mesure de la fonction de dispersion (PSF) qui mesure la réponse impulsionnelle du système optique, et les techniques de restauration aveugle, qui ne disposent d'aucune information a priori.

Caméra

Les caméras et leurs capteurs sont susceptibles d'introduire un grand nombre de défauts. Diverses sources de bruit, de nature thermique et électronique, polluent le signal. Pour les atténuer, on est parfois amené à refroidir la caméra. Les éléments sensibles du capteur peuvent présenter des sensibilités différentes, ce qui a pour conséquence l'apparition d'une non-uniformité de réponse permanente. Cette dernière peut être compensée par une calibration appropriée.

Lorsque l'espacement des pixels diffère horizontalement et verticalement, on dit que l'on a affaire à des pixels non carrés. Cet effet entraîne une déformation de l'image, sous forme d'un étirement. Le même étirement apparaît avec les caméras linéaires, lorsque la vitesse de défilement n'est pas ajustée par rapport à la fréquence de prise d'image. De telles déformations nécessitent un ré-échantillonnage.

Enfin, l'effet d'entrelacement des trames (lignes paires et lignes impaires balayées séparément) fait que des décalages se manifestent lorsqu'on observe des objets en mouvement. De plus, si les objets se déplacent rapidement, un flou de bougé apparaît. Pour l'éviter, on est amené à réduire la durée d'illumination et à faire appel à des lampes flash.

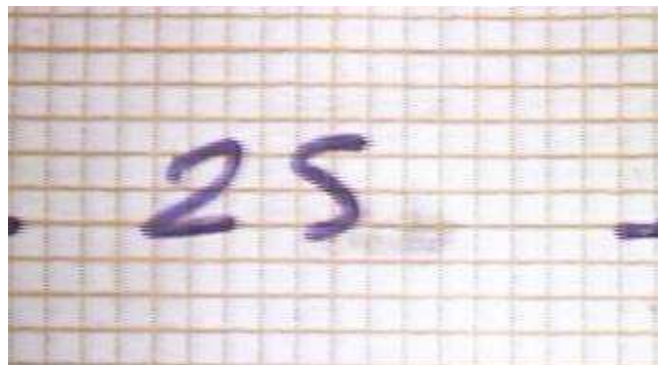


Figure 12.27: Défaut dû à l'entrelacement du capteur.

12.3 Moyens logiciels

Les logiciels de traitement d'image industriel, qui sont à l'heure actuelle très répandus sur le marché et qui représentent une technologie mûre, sont conçus pour s'intégrer sur des machines.

En général, on peut distinguer une mode de fonctionnement hors-ligne, durant lequel un opérateur prépare la tâche d'inspection en spécifiant le type de pièces à traiter, en ajustant des paramètres de fonctionnement et en présentant éventuellement des images-échantillons de référence (ce que l'on appelle l'apprentissage).

Ensuite, le fonctionnement en-ligne, à la cadence réelle, sans intervention de l'opérateur, peut prendre place. L'acquisition des images est synchronisée avec le déplacement des pièces, et dès qu'une anomalie est détectée, une action est commandée automatiquement: éjection de la pièce en cause, arrêt de la machine, signal d'alarme, ...

12.3.1 Organisation générale des traitements

Malgré la diversité des applications rencontrées, il est possible de dégager un schéma général pour la résolution des problèmes de vision et de définir une méthodologie appropriée. On peut décomposer les traitements en une séquence d'opérations de nature différente telle que représentées à la figure 12.28.

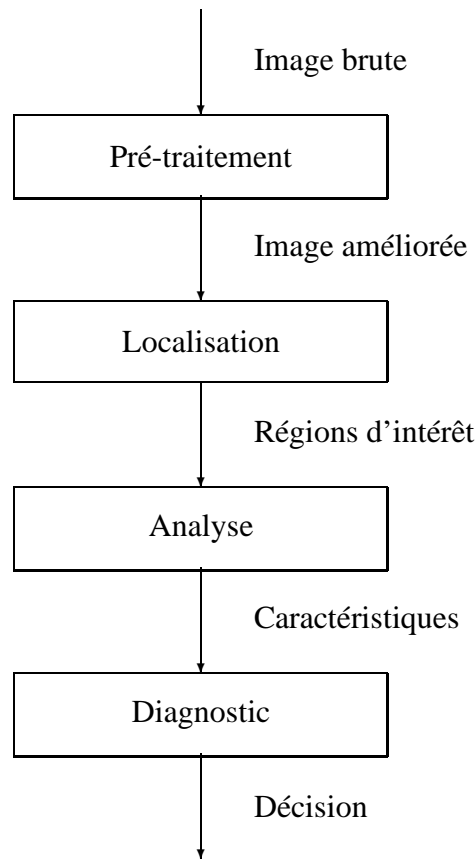


Figure 12.28: Organisation générale des opérations de traitement d'image.

Le *pré-traitement* consiste à préparer l'image brute pour faciliter les opérations qui suivent. Il s'agit soit de compenser divers défauts dont l'image pourrait être entachée, tels que bruit, manque de contraste, déformations géométriques, flou de bougé, non-uniformité d'éclairage, ... soit de mettre en évidence les contrastes intéressants: rehaussement des contours, filtrage des détails ayant une taille donnée, ou au contraire de réduire les contrastes indésirables. Les pré-traitements s'appliquent à des images et fournissent normalement de nouvelles images au contenu modifié. Parfois, on extrait une information plus condensée telle que des vecteurs (*images* à une dimension) constitués par des profils, projections ou autres.

La *localisation* est une partie essentielle des opérations. Il s'agit de trouver dans les images les éléments intéressants et de les délimiter. Comme on l'a déjà dit, ces éléments peuvent être

- inconnus à l'avance, auquel cas on doit procéder à une segmentation suivie d'une sélection des segments utiles, ou
- connus à l'avance par un modèle, soit extrait d'une image de référence, soit donné sous une forme synthétique.

Selon les cas, on se satisfera d'une localisation grossière ou au contraire la localisation sera l'objet principal de l'inspection. A l'issue de la phase de localisation, on connaîtra les emplacements des

éléments à inspecter et on aura précisé les pixels qui seront impliqués dans les mesures. Les ensembles de pixels ainsi délimités sont appelés *régions d'intérêt*. Les régions d'intérêt peuvent être des figures géométriques simples (fenêtres rectangulaires, circulaires ou polygonales, et combinaisons de celles-ci), ou correspondre à des formes extraites de l'image elle-même (segments).

L'*analyse* exploite les résultats de la localisation et a pour but de réduire à une petite quantité de paramètres l'information à traiter. Par exemple, on sera amené à décrire les objets par un petit nombre de paramètres de taille ou de forme en vue de les reconnaître; on pourra mesurer l'intensité lumineuse moyenne à l'intérieur d'une petite zone pour déterminer l'absence ou la présence de matière à cet endroit, ou la distance de bord à bord d'une pièce cylindrique pour déterminer son diamètre, ...

Le résultat de l'analyse est un vecteur de nombres résultant d'un calcul sur les diverses régions d'intérêt explorées.

Le *diagnostic* est la conclusion du traitement. Il exploite les valeurs des paramètres analysés, également appelés indicateurs de qualité, pour émettre un verdict tel que: la pièce est présente ou absente, exempte de défauts ou non, présente la bonne forme, appartient à une catégorie plutôt qu'une autre, ... Ces prises de décision sont intimement liées à la notion de classification. Le jeu de paramètres caractéristiques mesurés définissent un espace des états, lequel peut être partitionné en régions d'acceptation et régions de rejet, ou plus généralement en régions d'appartenance à des classes pré-définies.

Dans les cas les plus simples, le diagnostic se fait sur base d'un seul paramètre et il suffit de vérifier qu'il prend sa valeur dans un intervalle de variations acceptées. Les régions d'acceptation peuvent être déterminées par calcul, mais également être définies par comparaison avec des échantillons de référence, réputés bons.

En principe les différentes phases évoquées se succèdent au cours du traitement, mais il arrive que des retours en arrière aient lieu.

Lors de la résolution d'un cas d'application, le schéma global de traitement est défini au moment de la programmation de la solution et devient normalement immuable. Cependant, on laisse souvent accessible des paramètres réglables qui influencent le fonctionnement des algorithmes et qui permettent à un opérateur de réagir à des changements de conditions ambiantes ou des dérives de la production. Plus généralement, le passage d'un type de pièce à un autre demande de prendre en compte des caractéristiques qui changent, telles que les valeurs nominales des cotes à mesurer ou le texte du marquage à apposer.

L'entrée des paramètres par un opérateur, qui doit se faire à l'arrêt (avant de lancer l'inspection en-ligne), s'appelle l'*apprentissage*. L'apprentissage, quand il est nécessaire, peut se faire sur base d'images de référence ou de modèles synthétiques (représentations par C.A.O.). Lorsqu'un système est capable de s'adapter à un grand nombre de pièces de types différents, on lui adjoint un catalogue de modèles descriptifs prêts à l'emploi.

12.3.2 Techniques de pré-traitement

La préparation des images peut bénéficier une vaste gamme de méthodes aptes à résoudre divers problèmes. En voici une liste largement non exhaustive:

Amélioration du contraste: lorsque la quantité de lumière incidente est faible ou que les surfaces observées présentent des teintes voisines (en couleur comme en noir et blanc), le contraste peut être insuffisant. L'étirement d'histogramme (transformation linéaire ou non-linéaire) et son égalisation peuvent améliorer grandement la lisibilité de l'image.

Amélioration de la netteté: le flou peut être réduit très simplement par utilisation de filtres passe-haut, qui combinent d'une certaine façon l'image donnée et ses dérivées (rehaussement des

bords), de façon à accentuer les variations au détriment des zones uniformes.

Réduction de bruit par moyenne spatiale (filtres passe-bas) ou temporelle (intégration multi-images):

les pixels d'une image sont normalement fortement corrélés spatialement (d'un pixel à ses voisins) et temporellement (d'un instant à l'autre). Cette corrélation peut être exploitée pour éliminer les pixels "déviant".

Compensation de non-uniformité d'éclairage et de réponse du capteur: les non-uniformités d'éclairage (plus de lumière à certains endroits qu'à d'autres) sont en général lentes, c'est-à-dire qu'elles concernent essentiellement les fréquences spatiales basses. Un filtre passe-haut peut les atténuer. Une autre approche consiste à calibrer l'éclairage en conservant une image d'une mire uniforme qui sert de référence pour l'intensité des pixels de même intensité en chaque point. Le même procédé permet de tenir compte de la non-uniformité de la sensibilité du capteur (Une seule mire permet de corriger l'intensité. Deux mires, l'une claire et l'autre sombre, permettent de corriger à la fois l'intensité et le contraste).

Compensation des variations de contraste: en raison du vieillissement des sources d'éclairage, des conditions de lumière ambiante ou d'interventions sur les réglages du système, la quantité totale de lumière incidente peut varier. Cela peut influencer le résultat des traitements. Une façon de compenser de telles variations est la normalisation: on mesure la luminosité moyenne dans les conditions d'apprentissage, puis lors des inspections suivantes, on mesure à nouveau la luminosité moyenne de l'image et on applique un coefficient correcteur approprié. Cette technique n'est efficace que si le contenu de l'image ne change pas d'une image à l'autre.

Compensation de non-linéarité de réponse du capteur: une réponse linéaire peut être décrite par deux points de mesure (une relation linéaire en x , de la forme $ax + b$ comporte deux paramètres inconnus). Dans le cas général, par contre, la relation entre l'intensité lumineuse reçue et la tension mesurée peut être non-linéaire. Si nécessaire, la relation en question peut être étalonnée au moyen d'instruments de précision et être tabulée pour permettre la correction.

Réalignement (correction de position et direction): pour diverses raisons, en particulier à cause du jeu mécanique de la fixation des pièces, ou un déplacement involontaire de la caméra ou des pièces elles-mêmes, ... les objets observés n'occupent pas toujours exactement la même position dans le champ de vue. Si on est capable, par des méthodes exposées ci-après, de déterminer la position effective des pièces, on peut ramener celles-ci en position nominale par réalignement (translation et rotation, éventuellement remise à l'échelle). Cela permet de ramener l'objet inspecté dans une position telle que les régions d'intérêt utiles peuvent rester fixes.

Compensation des distorsions optique, perspective, des pixels non carrés: comme on l'a vu, la chaîne de formation de l'image peut introduire différents types de déformations qui donnent des objets observés une image faussée. Si on est capable de mesurer les déformations correspondantes, notamment par l'emploi de mires de formes connues –grilles, matrices de points, damiers– on est également capable de compenser ces déformations et de restituer la géométrie d'origine de la scène.

Ré-échantillonnage géométrique: parfois, le contenu brut de l'image n'est pas exploitable et on s'intéresse à des sous-parties d'images de forme particulière: les profils correspondent aux valeurs des pixels prises le long d'un segment de droite ou d'une courbe (arc de cercle ou autre). Ceux-ci sont particulièrement utiles pour analyser des formes longilignes. Les projections résultent d'un moyennage des valeurs des pixels autour d'une ligne de référence et donnent une vision unidimensionnelle de l'image (par exemple, une projection dans le sens horizontal

permet de séparer les lignes d'un texte). Le *déroulement* ou le *dégauchissement* (plus généralement appelés *anamorphoses*) consistent à compenser la déformation de la chose à observer, telle qu'un marquage annuaire au centre d'un CD-Rom ou l'étiquette d'une bouteille, pour pouvoir l'observer à plat.

Comparaison à une référence: dans de nombreux cas, les imperfections introduites par la chaîne d'acquisition peuvent être acceptées si l'on travaille par comparaison. Si les imperfections correspondent à des erreurs systématiques, c'est-à-dire se présentent toujours avec la même intensité et le même signe (distorsion optique permanente, non-uniformité stable dans le temps, désalignement constant, ...), il suffit de conserver une image de référence d'un objet sans défaut et de comparer cette image aux images à inspecter pour observer les défauts de la pièce sans être gêné par les défauts de la prise de vue. Souvent, une simple soustraction d'image, ou la comparaison à une tolérance de variation près suffit.

Tous les moyens présentés ci-avant permettent dans une certaine de réduire ou éliminer des causes d'erreurs accidentelles ou systématiques. Elles le font souvent au détriment d'autres qualités de l'image. Par exemple, la réduction de bruit s'accompagne toujours d'une perte de résolution qui se manifeste par une disparition des détails fins (à partir d'un certain niveau, et sans connaissance a priori, un algorithme ne peut pas faire la différence entre du bruit et un détail fin). Bien que le traitement numérique d'images procure une large gamme d'outils propres à réduire divers défauts, il est bien plus recommandable de les éliminer à la source, autant pour réduire la charge de calcul que pour éviter d'introduire d'autres défauts. Répétons-le également, les techniques d'éclairage sont de précieux auxiliaires pour mettre en évidence l'information utile et la séparer de l'information parasite.

12.3.3 Techniques de localisation

Après éventuel pré-traitement, c'est-à-dire lorsque la qualité d'image ne peut plus être améliorée sans une certaine "compréhension" du contenu, il est essentiel de pouvoir localiser dans l'image les régions intéressantes. Dans les cas les plus favorables, la localisation peut se faire a priori, en plaçant des régions d'intérêt fixes: lorsque la répétabilité du placement des pièces est bonne, on sait à l'avance où il faut inspecter. On peut donc placer des sondes de mesures statiques aux endroits appropriés.

Cependant, deux situations arrivent fréquemment: soit les positions des pièces fluctuent, d'une façon telle que des régions d'intérêt statiques ne conviennent pas, soit les positions des pièces sont totalement inconnues (voir même, on ne sait pas à l'avance quelles pièces sont présentées).

On peut distinguer au moins trois familles de méthodes qui s'appliquent plus ou moins bien à l'un ou l'autre cas de figure:

l'appariement de modèle (*template matching*): en premier lieu, on présente au système une instance de l'objet à localiser, sous la forme d'une image. C'est la phase d'apprentissage. Ensuite, lorsqu'une nouvelle image est présentée, une comparaison est faite par superposition entre l'image de référence et l'image inspectée, en faisant varier la position de la première. Le résultat de la recherche correspond à la meilleure correspondance trouvée. Deux grandes familles de méthodes sont utilisées: dans les méthodes surfaciques, on compare deux à deux tous les pixels du modèle et de la cible; les variantes les plus sophistiquées –basées sur la corrélation normalisée– permettent de compenser des variations globales de contraste. Dans les méthodes dites géométriques, on compare des représentations par descripteurs de formes plutôt que les pixels eux-mêmes¹.

¹Il ne faut pas confondre les expressions *pattern matching* (reconnaissance de formes) et *template matching* (ici rendu

la détection de contours: pour localiser des objets, il faut s'appuyer sur leur forme. Celle-ci est délimitée par leurs contours, c'est-à-dire les endroits où les variations de niveaux de gris sont les plus rapides: c'est ce que l'on appelle les bords (*edge detection*). Lorsque l'on franchit un bord, le profil des niveaux de gris suit une courbe en S, et sa dérivée première présente un extremum correspondant à un point d'inflexion. La position de ce point d'inflexion peut être déterminée avec grande précision par des techniques d'intégration ou d'ajustement de modèle, fournissant ainsi une précision sous-pixel (typiquement, de l'ordre du dixième de pixel). Lorsque la géométrie des bords d'un objet est connue à l'avance, la localisation des bords peut ainsi servir à les localiser. De plus, la localisation au sous-pixel des points de transition permet d'atteindre des précisions et répétabilités insoupçonnées.

la segmentation: lorsque le contenu de l'image est très mal connu à l'avance (par exemple dans le cas de la reconnaissance de caractères, on sait que l'on va trouver des caractères, mais on ignore lesquels), on doit avoir recours à des méthodes générales pour isoler des objets. La segmentation consiste à délimiter dans l'image des régions dont le contenu est homogène, telles que des à-plats uniformes, des zones texturées de la même façon, l'image d'arrière-plan plus sombre, ... De telles régions correspondent à des objets complets dans les cas simples (caractères imprimés en une seule couleur sur un fond contrasté), et à des parties d'objets dans le cas général.

La technique de segmentation la plus répandue est le seuillage suivi d'analyse des composantes connexes: on classe les pixels par comparaison à un niveau de seuil (foncé appartient au fond, clair appartient à l'objet), puis on regroupe les pixels de même type qui se touchent. On forme ainsi des *taches* dont on peut estimer la position par détermination du centre de gravité ou autre. Cette approche a le mérite de la simplicité, mais elle tombe en défaut lorsque le critère de classification par intervalles de niveaux de gris n'est pas vérifié. On parle alors d'*images non seuillables*. Un contournement possible est l'utilisation de seuils non uniformes, aussi appelés adaptatifs, tels que la valeur de seuil est ajustée pixel par pixel. On obtient des seuils adaptatifs par filtrage passe-bas ou à partir d'une image de référence.

Des techniques alternatives de segmentation sont plus sophistiquées et se basent non seulement sur la valeur de gris des pixels individuels, mais sur les relations de voisinage entre ceux-ci: lignes de partage des eaux, croissance de régions, fusion/division, ...

12.3.4 Techniques d'analyse et de mesure

Les moyens de localisation permettent de trouver dans l'image les objets ciblés. La connaissance de leur géométrie permet alors de définir des zones à mesurer. On s'efforcera en particulier de quantifier:

la couleur, en terme de niveaux de gris ou de composantes colorées. Pour cela, des statistiques sur des ensembles de pixels conviennent: valeur moyenne, minimum, diamètre du nuage de points, histogramme, ...

la taille et la forme: de nombreux descripteurs de forme sont disponibles. On peut en particulier songer à des statistiques sur la position des points de l'intérieur, telles que le centre de gravité, la boîte englobante, l'ellipse d'inertie équivalente, ... ou des statistiques sur les points des contours: longueur du périmètre, nombre de trous, sinuosité, élongation, ... en particulier, les descripteurs de FOURIER fournissent une description multi-échelle des formes.

par appariement de modèle). La première couvre les larges domaines de la mesure de la ressemblance et de la classification, et constitue une branche de l'intelligence artificielle, alors que la seconde a le sens précis de localisation en termes de position spatiale, d'un objet connu.

la position: les techniques de localisation telles que l'appariement de modèle ou la détection de contours peuvent servir non seulement à une localisation grossière des objets, mais aussi à une détermination rigoureuse de leurs dimensions telles que distances de bord à bords, angle entre arêtes, surface de trous. De tels contrôles dimensionnels ont l'avantage de se faire rapidement et sans contact, avec une précision importante. L'orientation peut également être mesurés avec une bonne précision²

De telles mesures sont cependant sensibles à des défauts tels que la distorsion optique qui modifie localement les facteurs d'échelle (correspondance entre les dimensions d'un pixel et la portion du champ visuel qu'il couvre effectivement sur l'objet observé). Ils nécessitent la mise en oeuvre d'une calibration géométrique, qui permet précisément de convertir les mesures faites en pixels en mesures en unités de longueur [*mm*]. La calibration est mise en oeuvre au moyen de mires ou de tout objet dont les dimensions sont connues avec précision et qui peuvent être mesurés facilement par le système d'imagerie.

12.3.5 Techniques de diagnostic

Le diagnostic, déclenchement d'alarme pour un objet hors tolérances, classification des pièces en bonnes ou mauvaises, reconnaissance d'une lettre dans un l'alphabet, ... ne fait pas à proprement parler partie du domaine du traitement d'images. Au contraire, il fait plutôt appel à des techniques statistiques et de reconnaissance de formes. Il en est cependant une partie indissociable.

D'une façon générale, il s'agit d'utiliser les résultats de mesures fournis par les phases d'analyse, qui ont pour but de mesurer quantitativement certaines caractéristiques des objets observés, pour vérifier leur conformité par rapport à un modèle idéal, compte tenu de tolérances qui englobent les variations inhérentes aux objets inspectés.

Les paramètres mesurés définissent un *espace des états* à N dimensions, dans lequel on découpe des zones d'acceptation et des zones de rejet, et plus généralement une partition en zones attribuées aux différentes classes rencontrées. Une façon simple de définir ces régions est de choisir un centre pour chaque classe au moyen d'un échantillon représentatif (exemplaire de référence), et de fixer des marges d'incertitudes autour de ce centre, soit par le raisonnement, soit par l'expérience.

Par exemple, on peut inspecter la surface de CD-Roms vierges et décider que les griffes plus courtes qu'un millimètre ne prêtent pas à conséquence.

Plus généralement, lorsqu'on dispose d'un échantillonnage suffisant, on peut répartir les pièces en *exemptes de défauts* et *griffées, trop cuites, malformées*, ... et introduire des paramètres d'aspect et de forme dans un classifieur à usage général (réseau de neurones ou autre). Celui-ci permettra d'attribuer une classe à tout nouvel échantillon présenté, en généralisant à partir des échantillons connus.

12.4 Un exemple concret: inspection de marquage

L'inspection de marquage, désignée par l'expression anglaise *Optical Character Verification* (OCV), est utilisée pour contrôler l'efficacité de dispositifs de marquage tels qu'imprimante à aiguille, graveur laser, tampon d'encrage, ... utilisés pour identifier des produits et y apposer des informations personnalisées: identité du fabriquant, numéro de série, date limite de vente, ... Il peut être extrêmement important de contrôler la lisibilité et la validité du contenu.

²À la différence des opérations de localisation évoquées ci-avant, qui ne requièrent pas une grande précision et ont pour seul but de délimiter les zones à analyser, les mesures dimensionnelles constituent le résultat proprement dit de l'inspection et requièrent la précision maximum.

Les systèmes OCV disposent normalement à l'avance d'un modèle des caractères à contrôler et procèdent typiquement comme suit:

- On présente à la caméra les composants marqués, un par un ou par petits groupes.
- Une phase d'apprentissage consiste à présenter un marquage exempt de défauts. Cet apprentissage permettra de définir la disposition et la forme des caractères, leur position relative au composant, le contraste, ... et fixera des tolérances de variation de ces paramètres.
- Au moment de l'inspection, les phases de traitement évoquées ci-avant interviendront:
 - **Pré-traitement:** on mesure le contraste global de l'image pour se faire une idée du niveau d'éclairage et s'adapter à la netteté des caractères par rapport au fond.
 - **Localisation:** on repère les bords de la boîte ou du composant qui comporte le marquage. A partir de ceux-ci, on détermine la région qui est susceptible de contenir le marquage. Par seuillage, on isole les caractères du fond. Par un critère de taille, on élimine les taches qui correspondent à du bruit ou à d'autres parties du marquage.
 - **Analyse:** on mesure sur chaque caractère des dimensions pertinentes, telles que dimensions hors-tout, surface, épaisseur des traits, "noirceur", lissage des contours, ...
 - **Diagnostic:** on compare ces mesurandes avec les valeurs acceptables établies par l'expérience. Éventuellement, on peut reconnaître les caractères par comparaison avec une police pré-enregistrée, sur base de leur forme, pour lire le numéro de série et assurer la traçabilité du composant. Cette dernière opération est appelée *Optical Character Recognition* (OCR).
- En fonction des résultats de l'inspection, des ajustements de position, de débit d'encre ou autres pourront avoir lieu, ainsi qu'un contrôle statistique permettant notamment d'appliquer une maintenance préventive.

Index

A

addition de Minkowski, 74
aire, 166
aliasing, 43, 47
amincissement, 85
anamorphose, 208
anti-extensivité, 78
appariement, 208
apprentissage, 206
approximation
 polygonale, 157
approximation polygonale, 157
arbre, 18
arbre quaternaire, 159
auto-dualité, 73

B

bande passante, 50
bord, 82
 effets, 117
boule maximale, 160
Butterworth, 54

C

caméra
 intelligente, 202
CCD, 200
centre morphologique, 108
chaîne, 81, 155
chain code, 21
chain rule, 81
champ
 clair, 198
 sombre, 199
chapeau haut-de-forme, 121
chemin géodésique, 85, 177
circularité, 169
CMOS, 200
CMY, 8
co-occurrence, 132
codage en chaîne, 155
coding

 run-length, 155
complémentaire, 73
composante
 continue, 29
 DC, 29, 35
compression
 taux, 181
condition de réduction, 162
connexité, 16, 77
 nombre, 169
contraste, 134
convexité, 171
convolution, 117
couleur, 2
courbe granulométrique, 170
covariogramme, 170
croissance, 78
 d'opérateurs de fonctions, 89

D

déconvolution, 49, 152
dégauchissement, 208
dérivée, 112
 directionnelle, 112
 partielle, 112
déroulement, 208
détection des contours, 109
DC, 29, 35
DCT, 32
descripteur de Fourier, 157
différence, 72
dilatation, 76
 de fonctions, 90
dilatation géodésique, 86
Dirac, 40
distance
 d'échantillonnage, 41
 euclidienne, 85
 fonction, 164
 géodésique, 177
distance focale, 199

dualité, 73, 77
dynamique, 118

E

échantillonnage, 41
éclairage, 197, 203
 en champ clair, 198
 structuré, 199
effet de bord, 82
égalisation de l'histogramme, 143
entrelacement
 défaut, 204
entropie, 131, 134
enveloppe convexe, 171
épaississement, 85
érosion, 75
 de fonctions, 90
 géodésique, 87
Euler, 166
extensivité, 78
 d'opérateurs de fonctions, 89

F

facteur de forme, 169
fermeture
 morphologique, 93
fermeture à multirésolution, 162
filtrage
 non-linéaire, 101
filtre
 algébrique, 104
 alterné séquentiel, 106
 de rang, 101
 de Wiener, 148
 Gabor, 57
 gaussien, 57
 homomorphique, 144
 idéel, 49
 linéaire, 49
 médian, 102
 multirésolution, 162
 passe-bande, 51
 passe-bas, 50
 passe-haut, 51
fonction
 d'échantillonnage, 41
 disque, 38
 Rectangle, 36
fonction distance, 164

forme
 indice, 169

Fourier

 descripteur, 157
 transformée discrète, 28, 33

fréquence

 d'échantillonnage, 41

G

géodésie, 85, 177
Gabor, 57
gamma, 8
gradient, 112
 approximation, 112
 d'érosion, 120
 de dilatation, 120
 morphologique, 121
granulométrie, 170
grossissement, 199

H

Hadamard, 32
histogramme, 129, 142
 égalisation, 143
Hit of Miss transform, 85
HMT, 85
homomorphique, 144
hue, 8

I

idempotence, 49, 72
 d'opérateurs de fonctions, 89
impulsion de Dirac, 40
incertitude, 59
inclusion, 72
indicatrice, 83
indice de forme, 169
infimum, 88
inspection, 191
intensité, 2
interpolation, 44–46
 bilinéaire, 48
 linéaire, 48
 plus proche voisin, 47
intersection, 72

K

kurtosis, 131

L

lacet, 21

lampe
 à quartz, 198
 halogène, 198
lampe à incandescence, 198
Lantuéjoul, 163
laplacien, 113
Laws, 136
LED, 198
ligne de partage des eaux, 175
localisation, 205
loi d'absorption, 106
LSB, 21
luminance, 2, 8

M
médian, 102
marquage, 176, 180
marqueur, 176
masque, 117
matrice
 orthogonale, 28
 symétrique, 28
 transposée, 27
 unitaire, 28
matrice de co-occurrence, 132
mesure, 164
Minkowski, 74
module, 34
moment, 134, 170
morphologie mathématique, 71
moyenne locale, 129
MSB, 21
multirésolution, 162

N
nombre
 d'Euler, 166
 de connexité, 169
normalisation, 118
norme
 approximation, 112
Nyquist, 43

O
obliquité, 131
ombre, 90
ombroscopie, 199
ondulation, 136
optique, 199

ordre partiel, 88
ouverture
 algébrique, 79, 104
 morphologique, 78, 93
ouverture à multirésolution, 162

P
périmètre, 166
pattern matching, 26, 208
phase, 34
pixel, 16
pouvoir séparateur, 199
Prewitt, 118
profondeur de champ, 199
propriété topologique, 169
pyramide, 21

Q
quadtree, 18, 159
quaternaire
 arbre, 159

R
région d'intérêt, 206
reconstruction, 44, 85
 morphologique, 87, 97
rectangle, 169
rectangularité, 169
rehaussement, 139, 141
repli de spectre, 43, 47
restauration, 139, 146
ROI, 26
run-length coding, 155

S
saturation, 8
segmentation, 173
seuillage, 174
Shannon, 43
Sobel, 119
soustraction de Minkowski, 74
spectre, 34, 54
spot, 136
squelette, 159, 163
 par zone d'influence, 178
supremum, 88
surface, 166

T
taille

- morphologique, 161
- taux de compression, 181
- teinte, 8
- template matching, 208
- texel, 125
- texture, 125
- théorème
 - Shannon, 43
 - structurel, 105
- toggle mapping, 108
- top-hat, 121
- trame, 15, 17
- transformée
 - DCT, 32
 - de Fourier, 28, 33
 - de Hadamard, 32
 - en cosinus discrète, 32
 - Hadamard, 32
- transformée par tout ou rien, 85
- transformation
 - orthogonale, 28
 - unitaire, 27, 28
- transformation de voisinage, 83
- translaté, 73
- transparence, 15
- tube fluorescent, 198

U

- union, 72

V

- vague, 136
- voisinage, 83

W

- Walsh, 32
- Whittaker, 45
- Wiener
 - déconvolution, 152
 - filtre, 148

Y

- YCbCr, 8
- YIQ, 8
- YUV, 8

Glossaire

- CIE *Commission internationale de l'éclairage*. Organisme qui a, entre autres, défini un système de représentation à partir de trois couleurs primaires. 5
- CLUT *Color Look Up Table*. Liste de couleurs, ainsi que leur numéros d'index associés, utilisée dans le rendu d'images 8 bits. 11
- CMYK *Cyan-Magenta-Yellow-Black*. Système de quatre couleurs de base utilisé par les imprimantes. Non approprié pour concevoir des pages web graphiques. (RGB est le système de couleurs utilisé pour les pages web graphiques). 8
- DCT *Discrete Cosine Transform*. La transformée en cosinus discrète est une opération d'analyse fréquentielle proche de la transformée de FOURIER. On l'utilise principalement pour le codage d'images. 33
- DFT *Discrete FOURIER Transform*. La transformée de FOURIER discrète s'obtient par échantillonnage de la transformée continue. 28
- entrelacement Processus qui consiste à combiner deux images ne contenant respectivement que les lignes paires ou impaires pour donner l'illusion du contenu. Ce processus est utilisé à l'intérieur des formats de télévision analogique PAL, NTSC et Sécam. 17
- FFT *Fast FOURIER Transform*. Transformée de FOURIER rapide. Il s'agit d'un algorithme de calcul rapide pour l'obtention de la transformée de FOURIER discrète (DFT). 31
- gamma Se réfère à la luminosité de l'écran d'un ordinateur. En termes techniques, il s'agit de l'ajustement de la relation non-linéaire entre une tension électrique et l'intensité lumineuse. 8
- JPEG *Joint Photographic Experts Group*. Format d'images présentant de la compression avec perte. Efficace pour les images du type photographique. 33
- LSB *Least Significant Bit*. Terme désignant le bit de poids faible dans un byte. 21
- MSB *Most Significant Bit*. Terme désignant le bit de poids fort dans un byte. 21
- NTSC *National Television Systems Committee*. Comité qui a normalisé le système de couleur utilisé pour la télévision analogique aux États-Unis. Il a donné son nom au dit format. 10, 18
- OCR *Optical Character Recognition*. Terme qui désigne les techniques optiques de reconnaissance de caractères. 211
- OCV *Optical Character Verification*. Techniques d'inspection de marquage utilisée pour contrôler l'efficacité de marquage en imagerie. 210
- PAL *Phase Alternate Line*. Format de télévision analogique utilisé dans la majorité des pays européens (hormis la France). 1, 10, 18
- palette Table indexant les couleurs des pixels d'une image. Ce principe est utilisé pour le format GIF. 11
- pixel *Picture Element*. Il s'agit d'un point d'une image échantillonnée. On parle aussi de PEL. 16
- RGB *Red Green Blue*. Système de représentation de couleurs au moyen de trois couleurs de base. Les téléviseurs et moniteurs représentent toutes les couleurs au moyen de 3 luminophores correspondants aux couleurs RGB. Il convient de remarquer que la plupart des écrans ne sont pas calibrés et qu'il n'y a pas de réel standard industriel unique pour la définition de ces composantes RGB, hormis celui défini par la CIE en 1931. Enfin, les composantes RGB des systèmes PAL et NTSC ne sont monochromatiques. 5

- ROI *Region Of Interest*. Région d'intérêt. Terme utilisé en traitement d'images pour désigner la zone de l'image utile pour des traitements ultérieurs. 26
- trame En traitement d'images, la trame est la grille d'échantillonnage. On considère généralement la trame carrée mais la trame peut aussi être rectangulaire ou hexagonale. Dans le cas du format entrelacé, la trame désigne une image ne contenant que les lignes paires ou impaires de l'image. En télécommunications, trame désigne un ensemble d'informations numériques temporelles constituant un tout. 15, 17

Bibliography

- [1] P. Buser et M. Imbert. *Neurophysiologie fonctionnelle IV*. Hermann, 1987. 3, 5, 11
- [2] J. Candy. *Signal processing: The modern approach*. McGraw-Hill, 1988. 63
- [3] J. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 2(7):1160–1169, July 1985. 60
- [4] C. Demant, B. Streicher-Abel, and P. Waszkewitz. *Industrial image processing: visual quality control in manufacturing*. Springer, 1999.
- [5] D. Docampo-Amoedo and A. Figueiras-Vidal. A deconvolution approach to harmonic signal extrapolation. In *Int. Conf. on Acoustics, Speech and Signal Processing*, pages 2345–2348. IEEE, May 1989. 68
- [6] P. Duvaut. *Traitement du signal: concepts et applications*. Hermès, Paris, 1991. 66
- [7] U. Franke. Selective deconvolution: A new approach to extrapolation and spectral analysis of discrete signals. In *Int. Conf. on Acoustics, Speech and Signal Processing*, pages 30.3.1–30.3.4. IEEE, May 1987. 63, 67, 68
- [8] R. Gerchberg. Super-resolution through error energy reduction. *Optica Acta*, 21(9):709–720, 1974. 67
- [9] H. Hadwiger. *Vorlesungen über inhalt, oberfläche and isoperimetric*. Springer Verlag, 1957. 166
- [10] R. Haralick, S. Sternberg, and X. Zhuang. Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):532–550, July 1987. 80
- [11] B. Haskell, A. Puri, and A. Netravali. *Digital video : an introduction to MPEG-2*. Chapman & Hall, 1997.
- [12] M. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 8:179–187, 1962. 170
- [13] B. Hunt. Nonstationary statistical image models (and their application to image data compression). *Computer Graphics and Image Processing*, 12:173–186, 1980. 67
- [14] R. Hunt. *The reproduction of colour*. Fountain Press, fifth edition, 1995. 5
- [15] N. Jayant and P. Noll. *Digital coding of waveforms*. Englewood Cliffs, N. J.: Prentice-Hall, 1984. 181

- [16] S. Katzenbeisser and F. Petitcolas. *Information hiding techniques for steganography and digital watermarking*. Artech House, 2000. 24
- [17] M. Kunt. *Traitement numérique du signal*. Dunod, 1981. 54
- [18] M. Kunt, G. Grunland, et M. Kocher. *Traitement de l'information: traitement numérique des images*, Volume 2. Presses polytechniques et universitaires romandes, 1993. 4, 7
- [19] G. Laurent. *Cours de télévision, tome 1: réception, norme, gestion, traitement vidéo*. Dunod, 1996. 18
- [20] F. Meyer and S. Beucher. Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1):21–46, September 1990. 179
- [21] J. O'Rourke. *Computational geometry in C*. Cambridge University Press, second edition, 1998.
- [22] A. Papoulis. A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Transactions on Circuits and Systems*, 22(9):735–742, September 1975. 67
- [23] A. Papoulis. *Probability, random variables, and stochastic processes*. McGraw-Hill, 1984. 67
- [24] J. Parker. *Practical computer vision using C*. John Wiley & Sons, 1994. 169
- [25] W. Pratt. *Digital image processing*. John Wiley & Sons, second edition, 1991.
- [26] U. Reimers. *Digital Video Broadcasting (DVB): the international standard for digital television*. Springer, 2001.
- [27] G. Roelofs. *PNG: the definitive guide*. O'Reilly & Associates, 1999.
- [28] J. Russ. *The image processing handbook*. CRC Press, second edition, 1991. 169
- [29] M. Sabri and W. Steenaart. An approach to band-limited signal extrapolation: the extrapolation matrix. *IEEE Transactions on Circuits and Systems*, 25(2):74–78, February 1978. 64
- [30] M. Schmitt et J. Mattioli. *Morphologie mathématique*. Masson, Paris, 1993.
- [31] J. Serra. *Image analysis and mathematical morphology*. Academic Press, New York, 1982. 76, 78
- [32] P. Soille and J.-F. Rivest. Principles and applications of morphological image analysis. Tutorial, Den Haag, September 1992.
- [33] P. Soille and L. Vincent. Determining watersheds in digital pictures via flooding simulations. In *Visual Communications and Image Processing, Vol. 1360*, pages 240–250, Lausanne, October 1990. SPIE. 179
- [34] M. Sonka, V. Hlavac, and R. Boyle. *Image processing, analysis, and machine vision*. Brooks/Cole Publishing Company, second edition, 1999. 103, 170, 175
- [35] S. Sternberg. Grayscale morphology. *Computer Vision, Graphics, and Image Processing*, 35:333–355, 1986. 88
- [36] G. Strang. The discrete cosine transform. *SIAM Review*, 41(1):135–147, 1999. 31, 32

- [37] M. Van Droogenbroeck. Algorithmes d'ouverture morphologique. Document interne au Centre de Morphologie Mathématique, N-17/93/MM, 22 pages, Fontainebleau, Mars 1993. 81
- [38] M. Van Droogenbroeck. *Traitement d'images numériques au moyen d'algorithmes utilisant la morphologie mathématique et la notion d'objet : application au codage*. Thèse de doctorat, Université catholique de Louvain, Mai 1994. 68, 161, 162
- [39] L. van Vliet, I. Young, and G. Beekers. A nonlinear Laplace operator as edge operator in noisy images. *Computer Vision, Graphics, and Image Processing*, 45:167–195, 1989. 121
- [40] L. Vincent. *Algorithmes morphologiques à base de files d'attente et de lacets, extension aux graphes*. Thèse de doctorat, École nationale supérieure des mines de Paris, Mai 1990. 164
- [41] L. Vincent. Morphological transformations of binary images with arbitrary structuring elements. *Signal Processing*, 22(1):3–23, January 1991. 80
- [42] I. Young, J. Gerbrands, and L. van vliet. *Image processing fundamentals*, chapter 51, pages 1–81. CRC Press, 1998. 62
- [43] X. Zhuang and R. Haralick. Morphological structuring element decomposition. *Computer Vision, Graphics, and Image Processing*, 35(35):370–382, 1986. 81