Institut MONTEFIORE
Service de Télécommunications et d'Imagerie

Acquisition et traitement de l'image

Professeur M. VAN DROOGENBROECK

Septembre 2001 (version 4.14)

Table des matières

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 . . . 3
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. . . . 9
D'autres systèmes de couleurs. . . . 12
Les systèmes de couleurs en pratique. . . . 12
Les fausses couleurs. . . . 12
La couleur des objets . . . 13
1.2.3 Effets visuels . . . 13
1.2.4 L'éclairage . . . 16
1.3 Représentation des images . . . 17
1.3.1 Objets à représenter . . . 17
La vidéo . . . 19
Standards de télévision analogiques . . . 20
1.3.2 Structures de données . . . 21
1.4 Traitement d'images industriel . . . 25
1.4.1 Remarques préalables . . . 25
Décomposition d'une image en plans binaires . . . 25
Critères de qualité . . . 26
Visualisation 3D d'une image 2D . . . 26
1.4.2 Segmentation d'images . . . 27
1.4.3 Reconnaissance de caractères . . . 28
1.4.4 Autres opérations . . . 29

2 Transformations unitaires **31**
2.1 Introduction . . . 31
2.2 Transformée de FOURIER . . . 31

2.2.1 Définition . . . 31
2.2.2 Propriétés . . . 32
2.2.3 Illustrations . . . 34
La fonction Rectangle . . . 34

3 Filtrage linéaire **37**
3.1 Notion de filtre idéal . . . 37
3.1.1 Catégories de filtres idéaux . . . 38
3.2 Passage d'une image dans un système linéaire . . . 40
3.2.1 Forme de la fenêtre . . . 42
3.2.2 Filtres passe-bas . . . 42
3.2.3 Filtres passe-haut . . . 43

4 Rehaussement et restauration **45**
4.1 Généralités . . . 45
4.2 Rehaussement . . . 47
4.2.1 Définition . . . 47
4.2.2 Rehaussement par fonction croissante . . . 47
4.2.3 Rehaussement par accentuation des contours . . . 47
4.2.4 Manipulation de l'histogramme . . . 49
Exemple : fonction parabolique . . . 49
Égalisation de l'histogramme . . . 49

5 Segmentation d'images **53**
5.1 Présentation du problème . . . 53
5.1.1 Définition . . . 53
5.2 Segmentation par seuillage . . . 54
5.3 Segmentation par détection de régions . . . 54
5.3.1 Ligne de partage des eaux . . . 54
5.3.2 Expression formelle d'un algorithme de segmentation par ligne de partage des eaux . . . 58
Analyse d'images binaires . . . 58
Algorithme de construction du squelette par zone d'influence. . . 59
Images en niveaux de gris . . . 59
Notations : . . . 59
Initialisation . . . 60
Construction . . . 60

6 Tomographie par rayons X **63**
6.1 Introduction . . . 63
6.2 Mesures . . . 65
6.3 Reconstruction en géométrie parallèle . . . 67
6.3.1 Notations et définitions . . . 67

6.3.2 Lien entre une image et sa transformée de Radon 69
6.4 Algorithmes de reconstruction 72
6.4.1 Méthode directe de FOURIER 72
6.4.2 Méthode du filtrage de la rétroprojection 75
6.4.3 Méthode de rétroprojection des projections filtrées 76

Glossaire **81**

Chapitre 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 une 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.

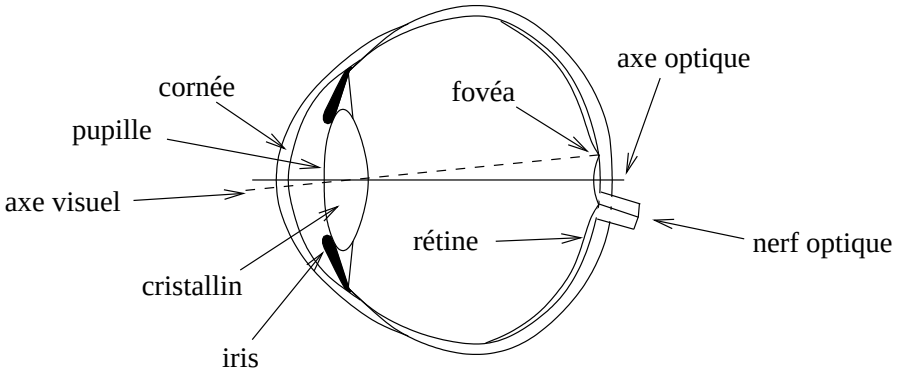


FIG. 1.1 – Coupe latérale simplifiée de l'œil.

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*.

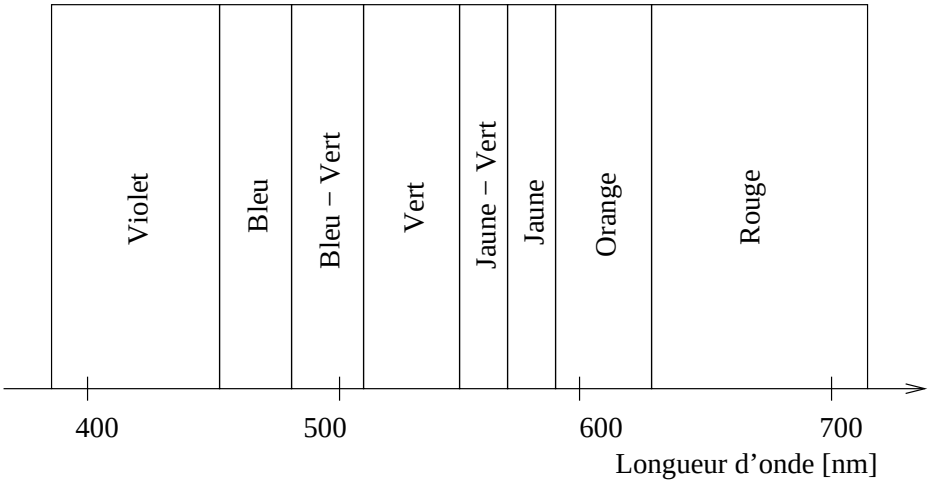


FIG. 1.2 – Les longueurs d'onde associées aux couleurs.

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 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'in-

tervalle $490 - 570\,[nm]$, 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

$$L_1 + L_2 = L_b + L_r \tag{1.1}$$

  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 $\lambda_1,\ \lambda_1,\ \ldots,\ \lambda_N$ 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

$$\sum_{i=1}^{N} L_i = L_b + L_r \tag{1.2}$$

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

$$\int_{\lambda} L(\lambda)\, d\lambda \tag{1.3}$$

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*. [4]) 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 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 les paramètres a, b et c, l'observateur essaie d'égaliser la couleur à

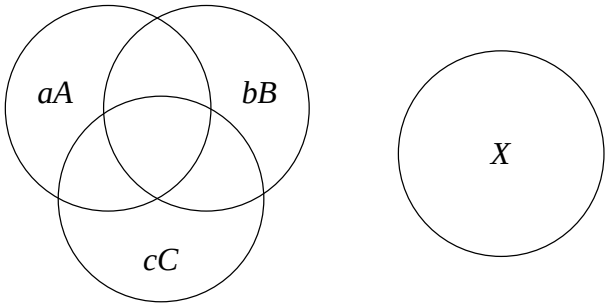


FIG. 1.3 – Expérience d'égalisation d'une couleur X au moyen de trois couleurs primaires A, B et C.

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.

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'un 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

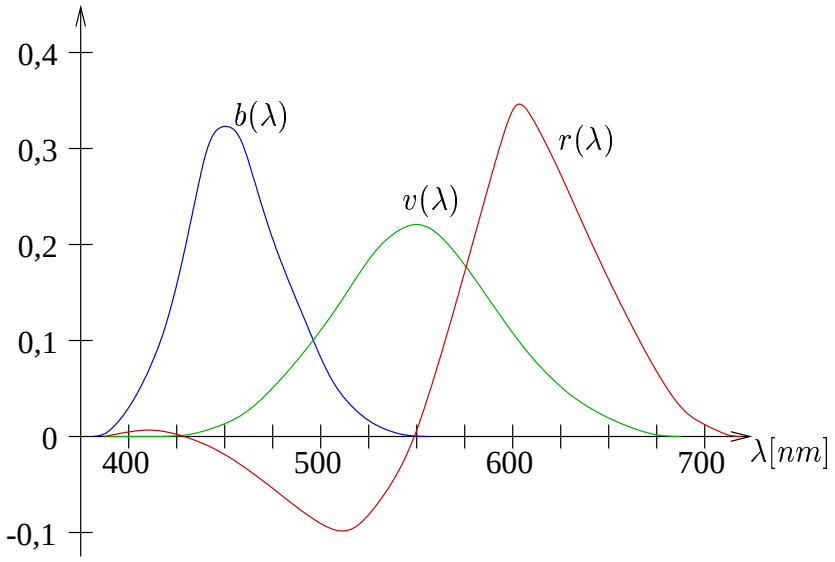


FIG. 1.4 – Courbes d'égalisation spectrale obtenues par égalisation des couleurs au moyen d'un mélange additif (d'après [2]).

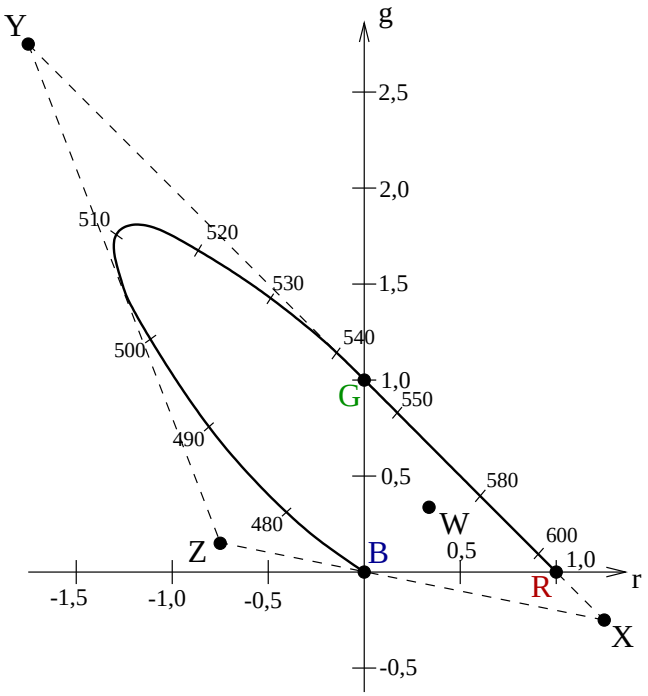


FIG. 1.5 – Diagramme chromatique RGB de la CIE.

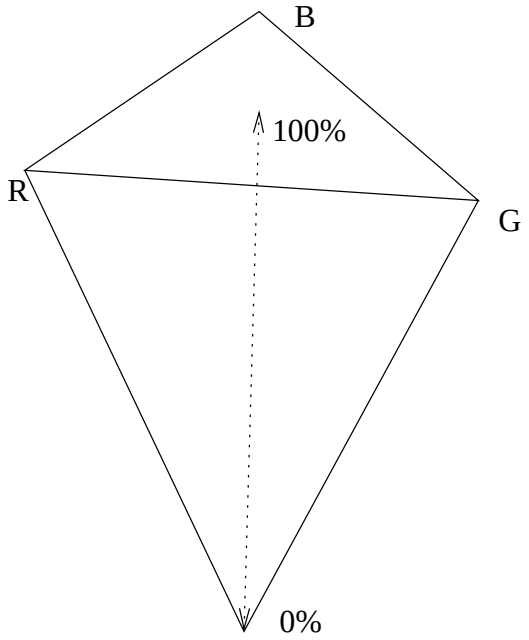


FIG. 1.6 – Pyramide des couleurs obtenues au moyen du tri-stimulus RGB.

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).

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 [4]

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

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

$$x = \frac{X}{X+Y+Z} \tag{1.5}$$

$$y = \frac{Y}{X+Y+Z} \tag{1.6}$$

$$z = \frac{Z}{X+Y+Z} \tag{1.7}$$

$$x+y+z = 1 \tag{1.8}$$

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.

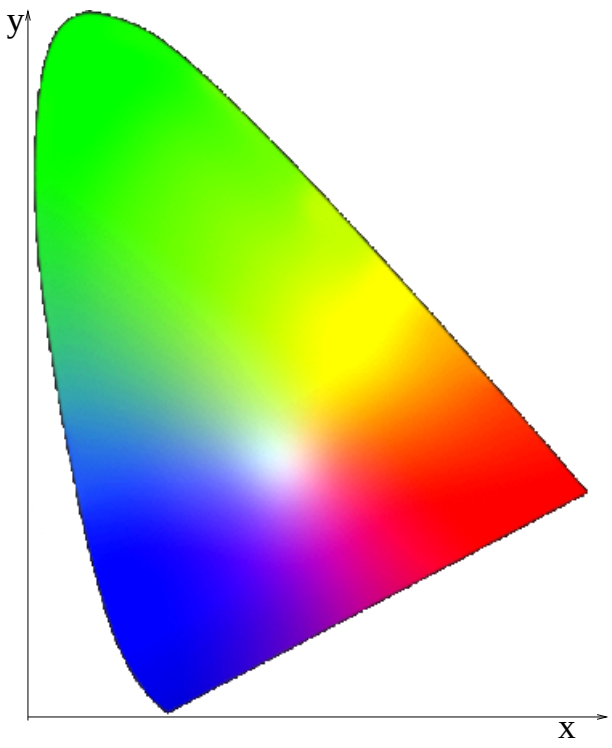


FIG. 1.7 – Diagramme chromatique (approché !) défini par les deux variables de chrominance x et y.

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.

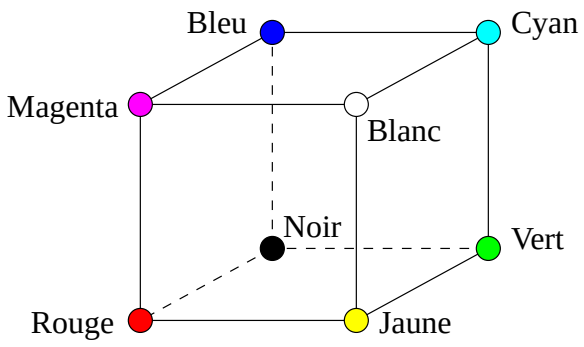


FIG. 1.8 – Espace tridimensionnel des stimuli produits par les composantes RGB.

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

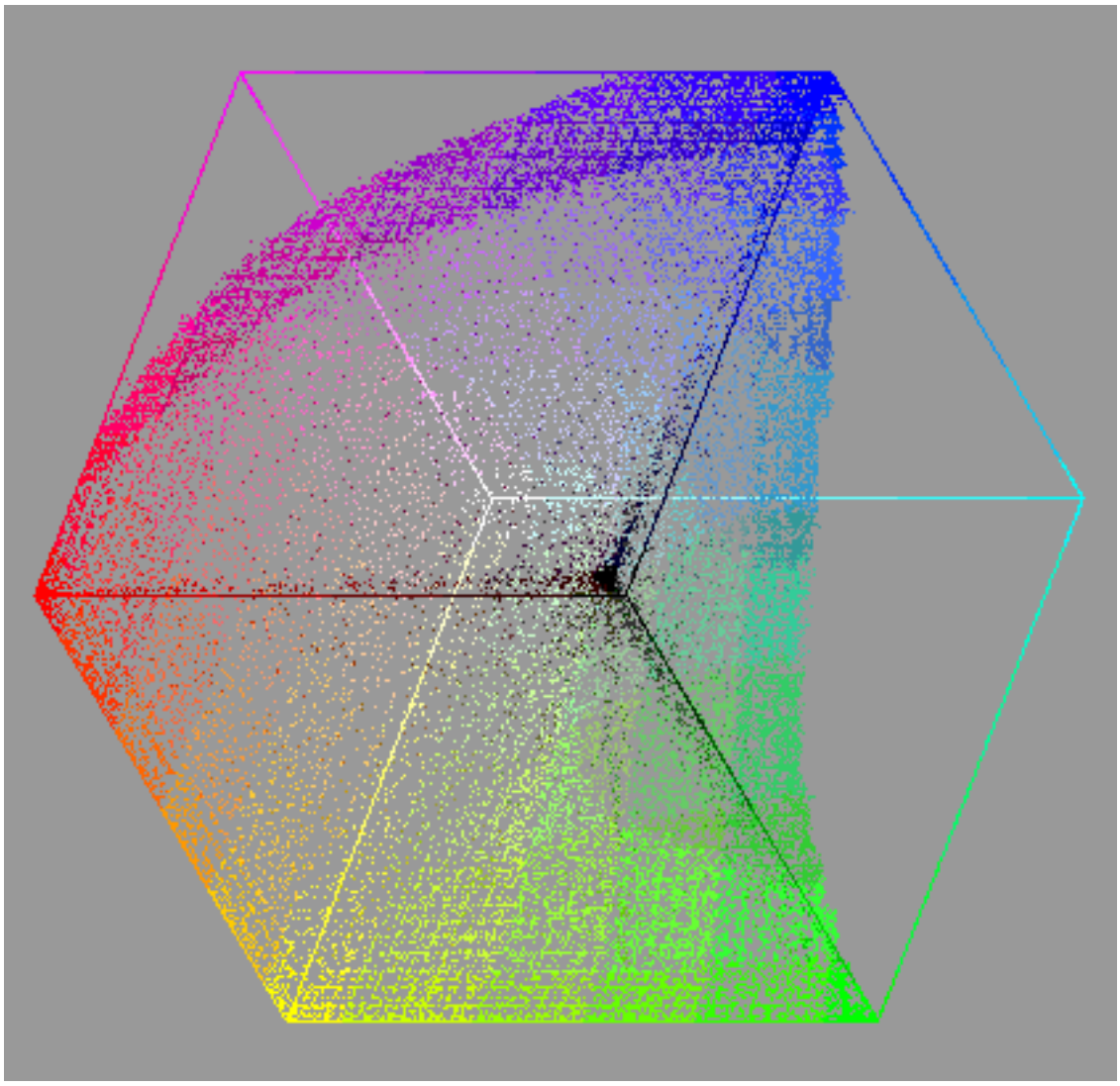


FIG. 1.9 – Espace tridimensionnel des couleurs du diagramme de chrominance approché.

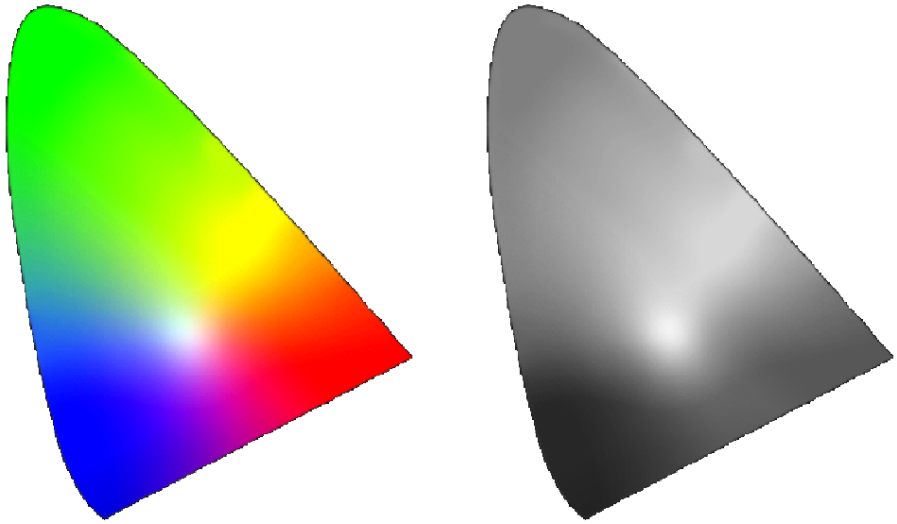


FIG. 1.10 – Diagramme chromatique xy et luminance maximale en chaque point.

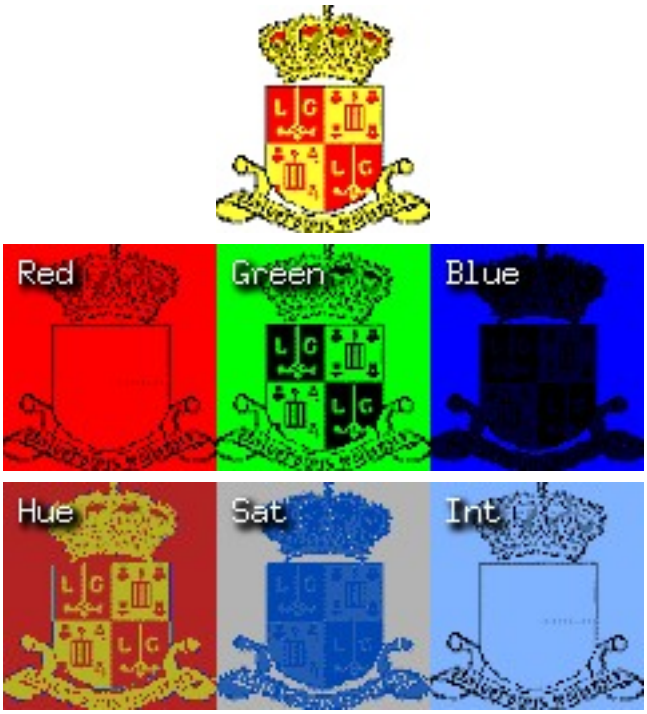


FIG. 1.11 – Décomposition d'une image en couleurs.

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 noir[1].

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 `gamma`[2].

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

TAB. 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

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

[2] La *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.

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.

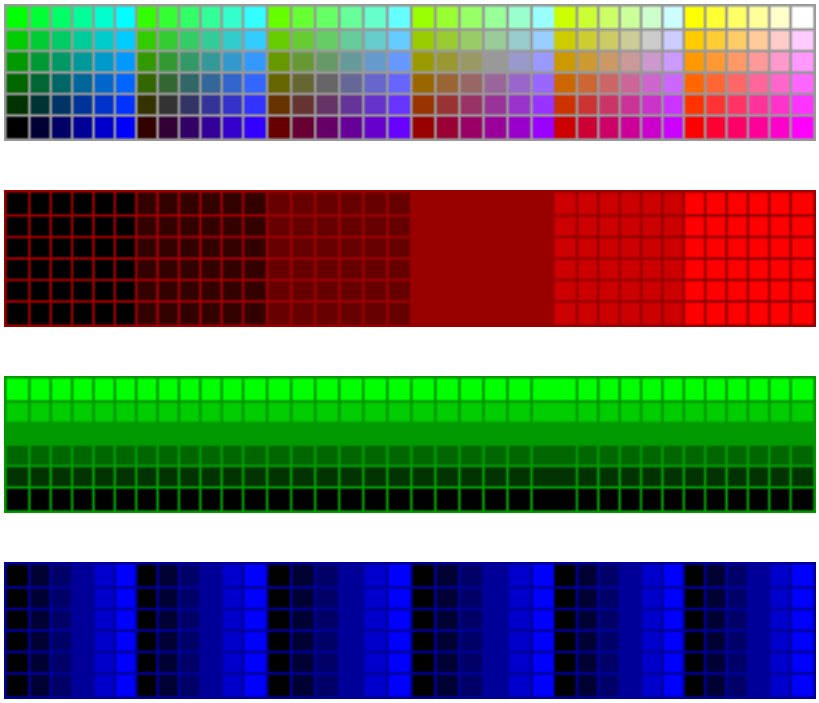


FIG. 1.12 – Palette de couleurs utilisée par les logiciels de navigation ainsi que la décomposition en R, G et B.

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'à-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 $I + \triangle I$ sur un fond

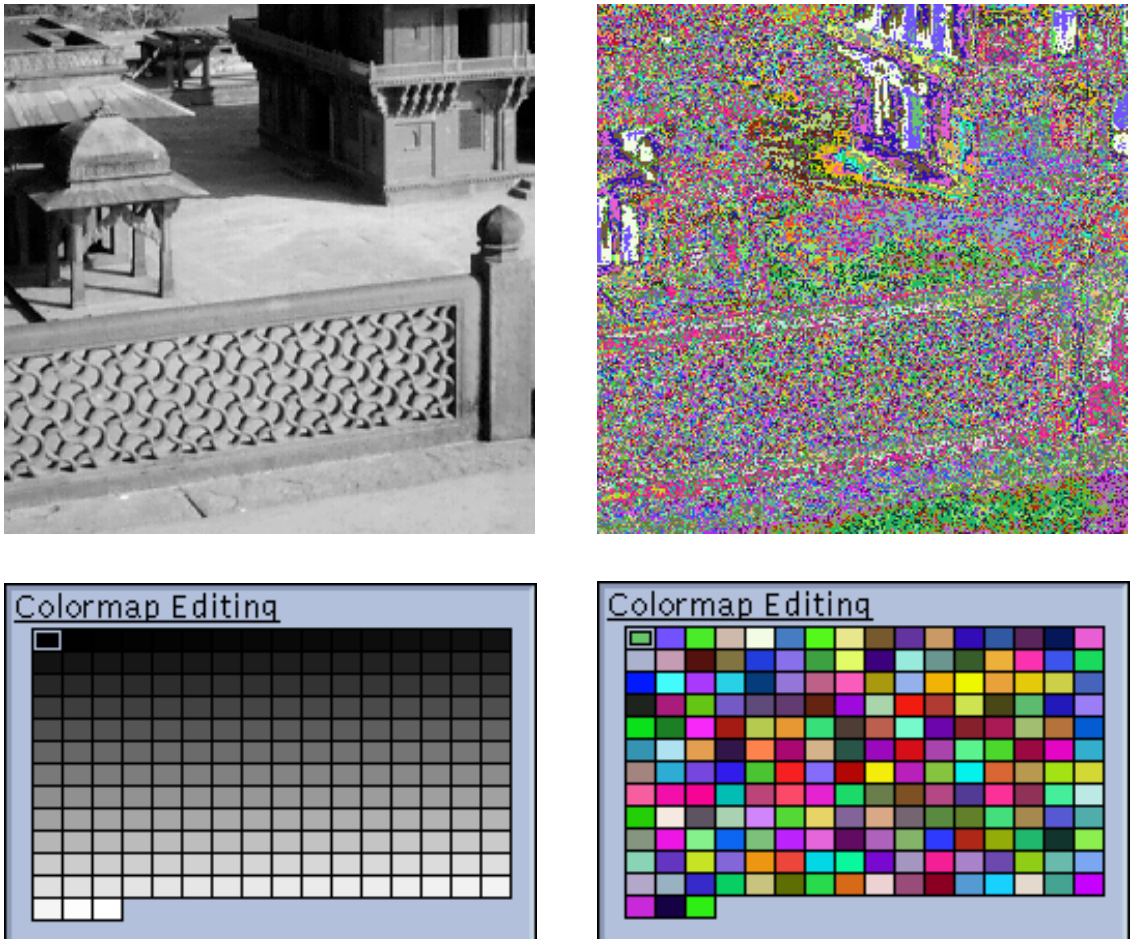


FIG. 1.13 – Une image en niveaux de gris, l'équivalent avec une palette de couleurs aléatoires et les palettes respectives.

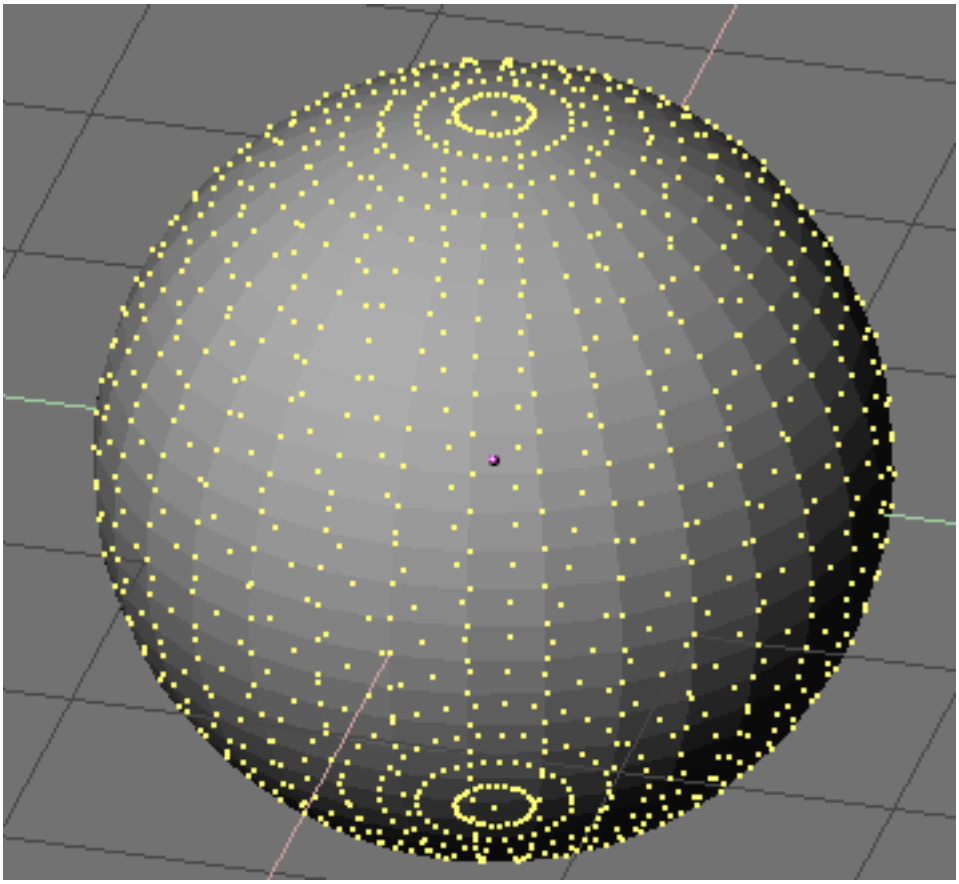


FIG. 1.14 – Une image de synthèse 3D.

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. On peut encore compliquer l'expérience en découpant la pastille

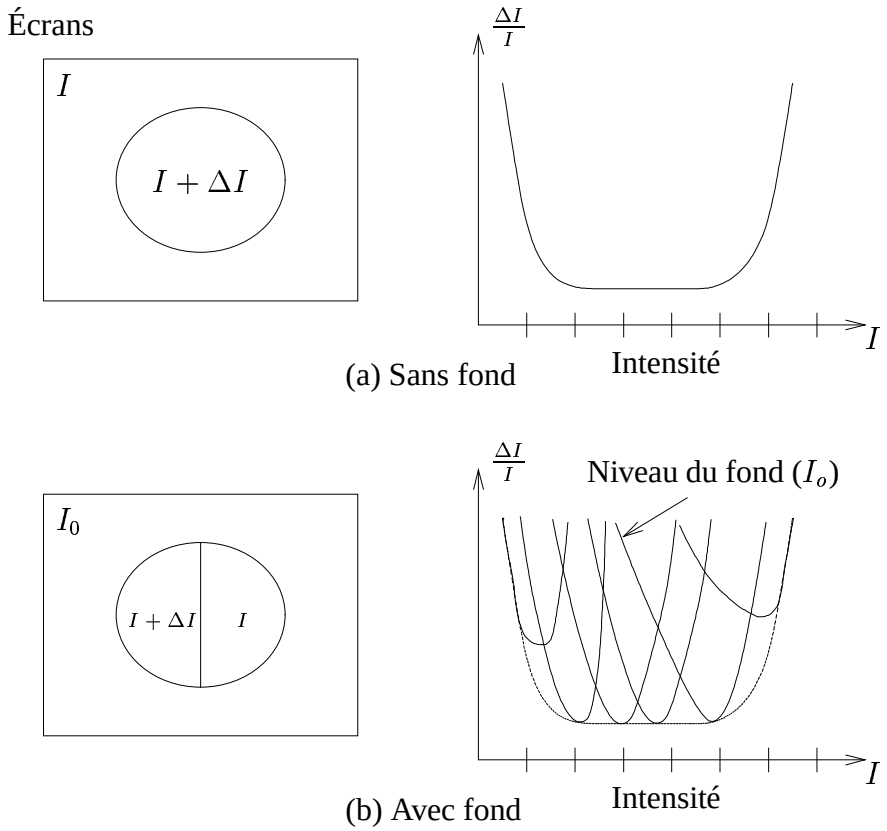


FIG. 1.15 – Différents arrangements expérimentaux permettant d'étudier les variations supraliminaires d'intensité, $\triangle I$, à partir de l'intensité I.

centrale en deux parties. On obtient 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.

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.

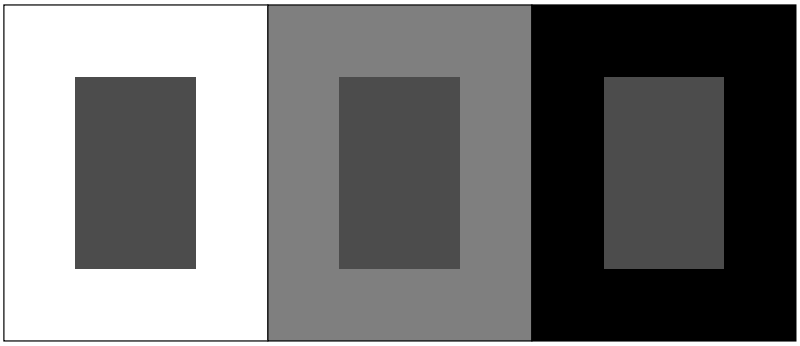


FIG. 1.16 – Illustration d'un effet perceptif.

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

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

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

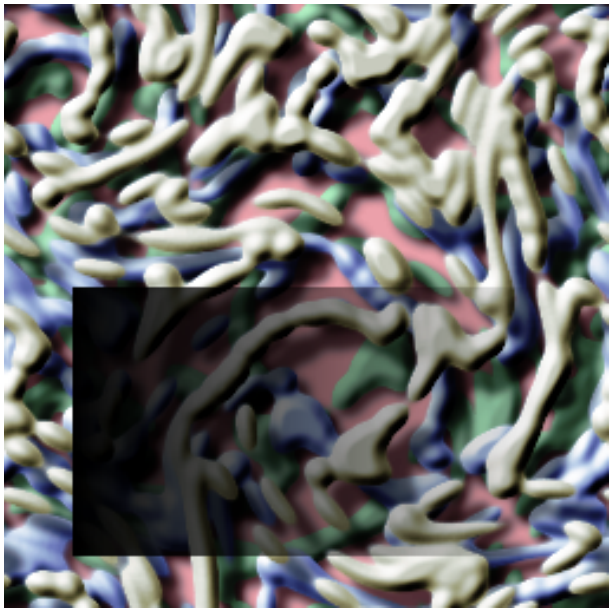
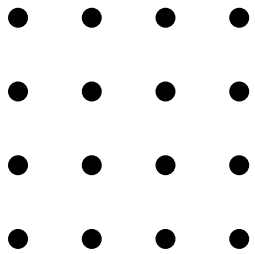
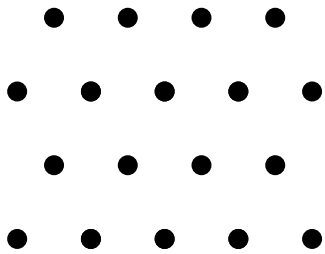
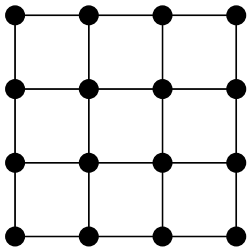
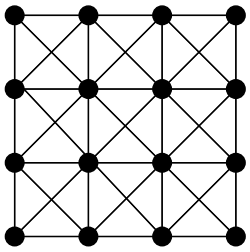
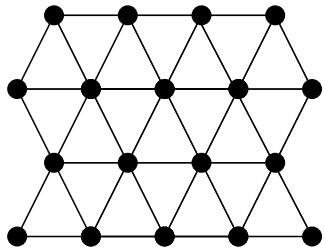


FIG. 1.17 – Utilisation de bits de transparence à l'intérieur du petit rectangle.

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

TAB. 1.2 – Types de trames couramment utilisées et connexités associées.

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.

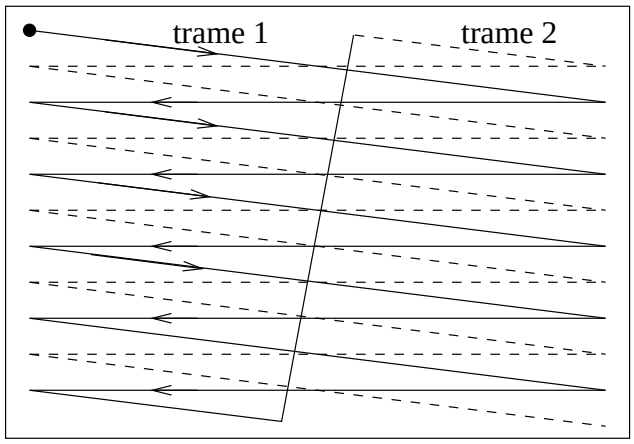


FIG. 1.18 – Description du format entrelacé.

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.

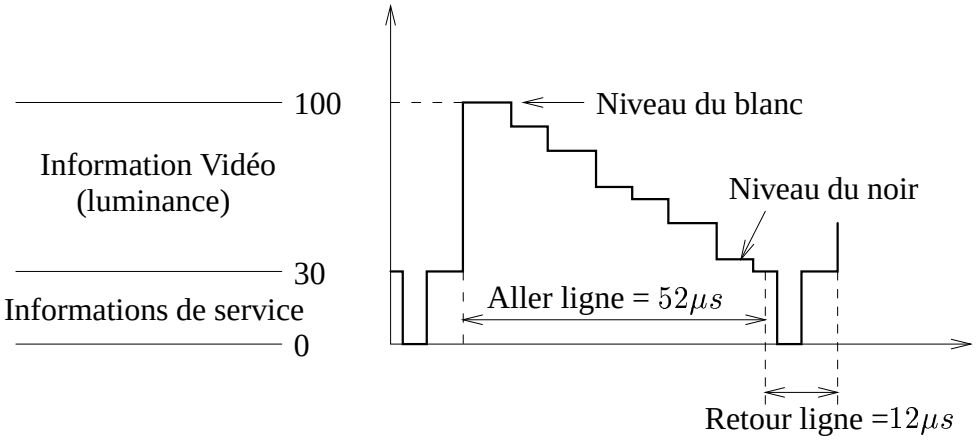


FIG. 1.19 – Une ligne du signal de luminance d'un téléviseur.

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

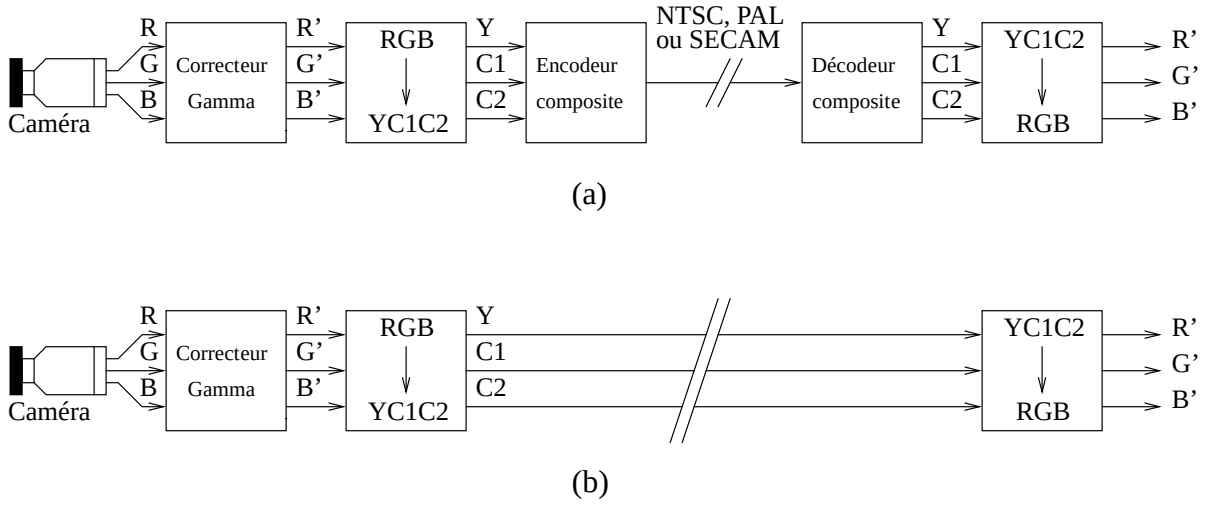


FIG. 1.20 – Chaîne de transmission des signaux de télévision composites.

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 PAL[3] 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.

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)

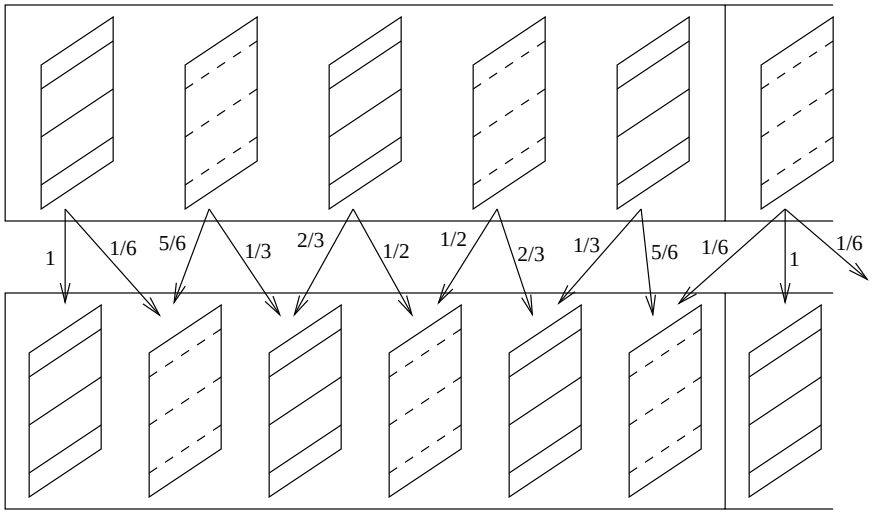
TAB. 1.3 – Comparaison des 3 principaux formats de télévision analogiques.

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.

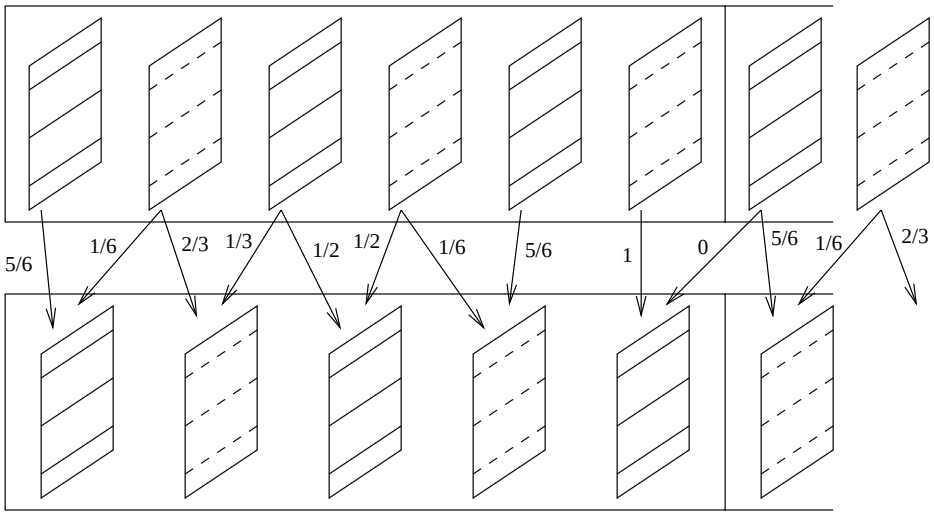
1.3.2 Structures de données

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.

[3] LAURENT [5] décrit en détail les principes de fonctionnement du PAL.



(a) Conversion 50Hz vers 60Hz



(b) Conversion 60Hz vers 50Hz

FIG. 1.21 – Conversion de la fréquence image.

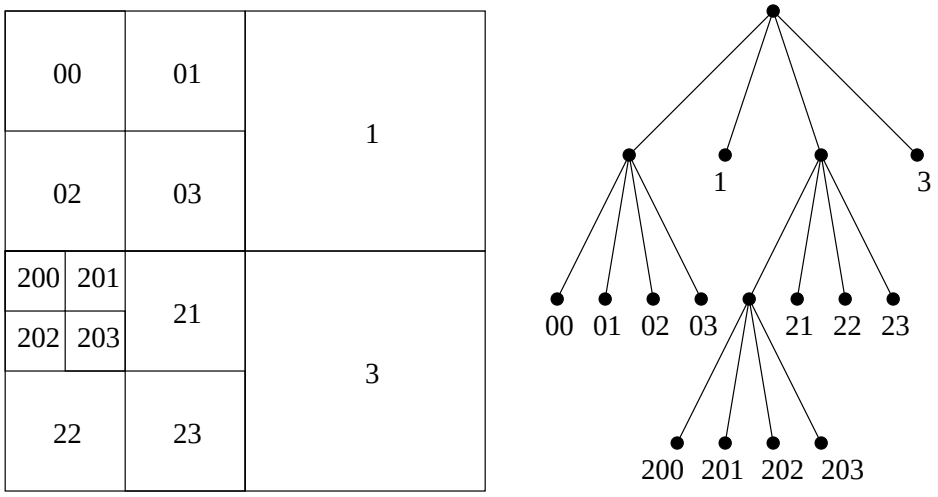
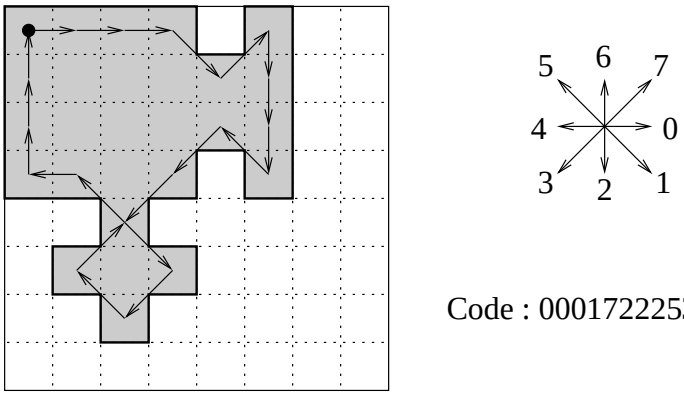


FIG. 1.22 – Découpage d'une image en arbre quaternaire.

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*.



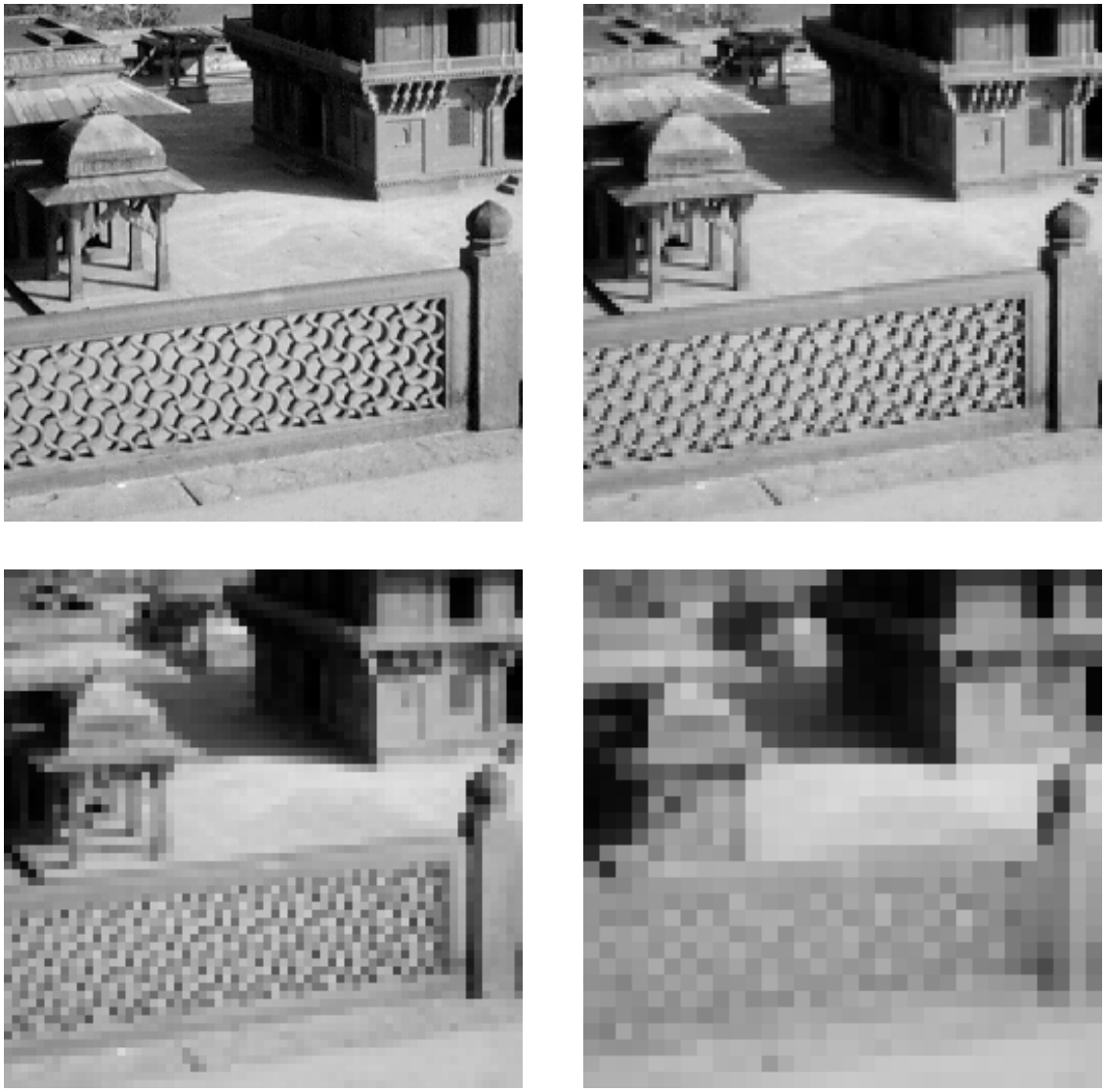
Code : 00017222533135754666

FIG. 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.



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

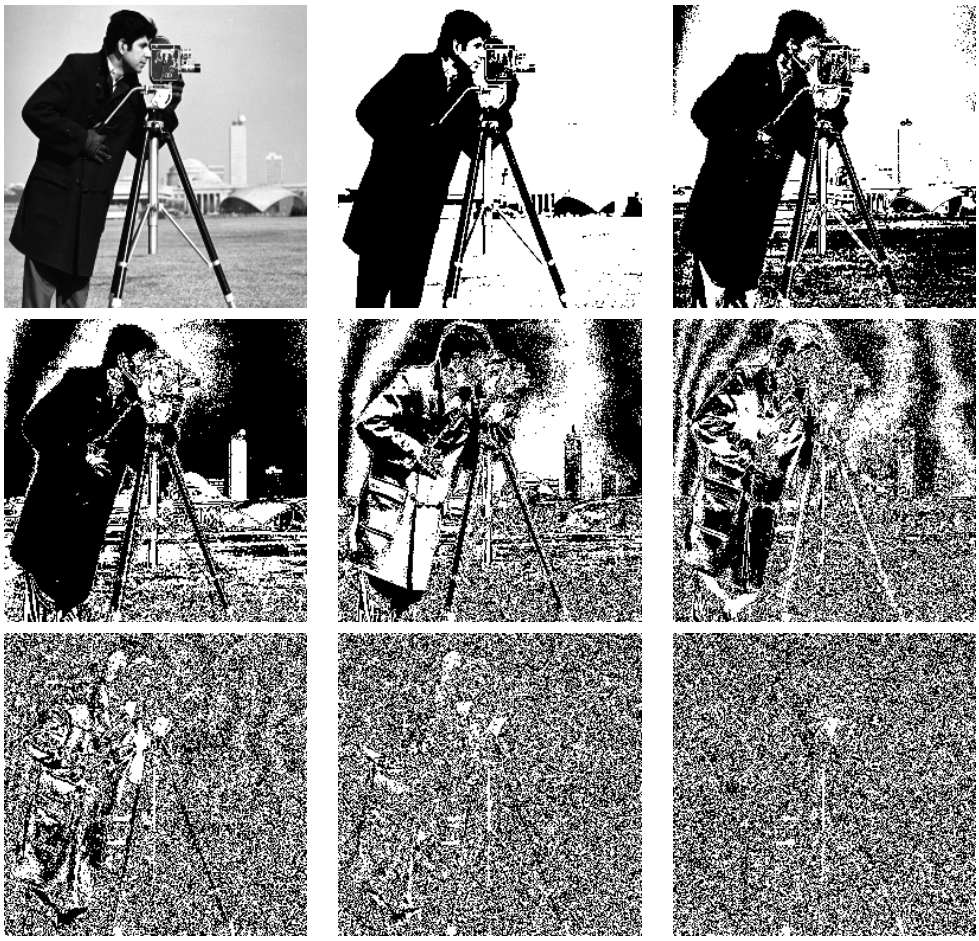
1.4 Traitement d'images industriel

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.



TAB. 1.5 – Image originale et plans binaires en partant du MSB.

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.

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]

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

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

Définition 2 [Signal to noise ratio]

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

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

Définition 3 [Peak signal to noise ratio]

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

Il existe aussi des mesures de distorsion corrélatives ou basées sur l'évaluation de l'histogramme (pour une liste plus complète, voir par exemple [3, 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 principe[4].

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

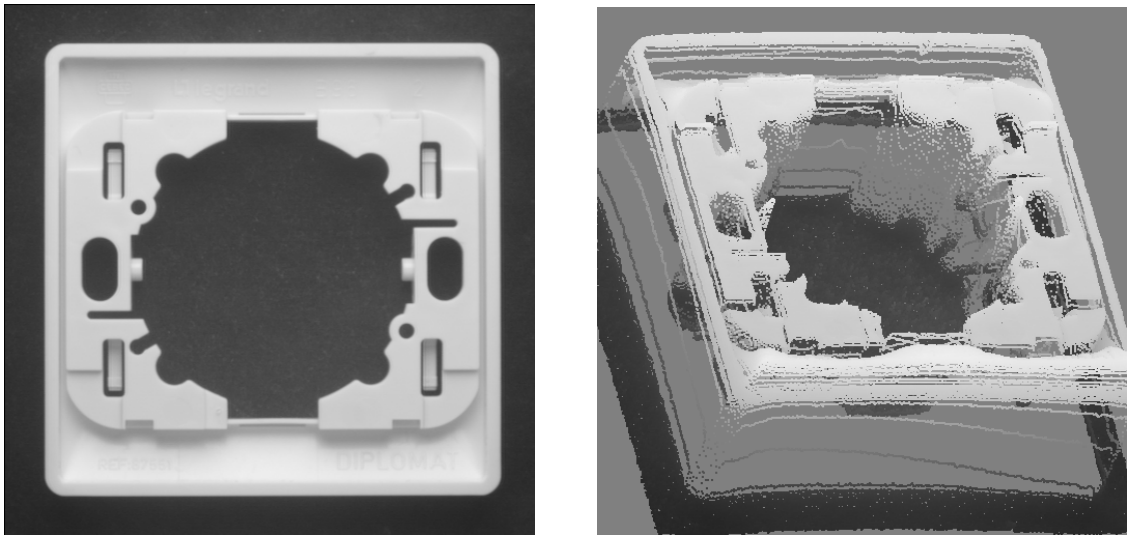


FIG. 1.24 – Une image et visualisation de la profondeur sous la forme d'une image 3D.

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 illustrées 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.

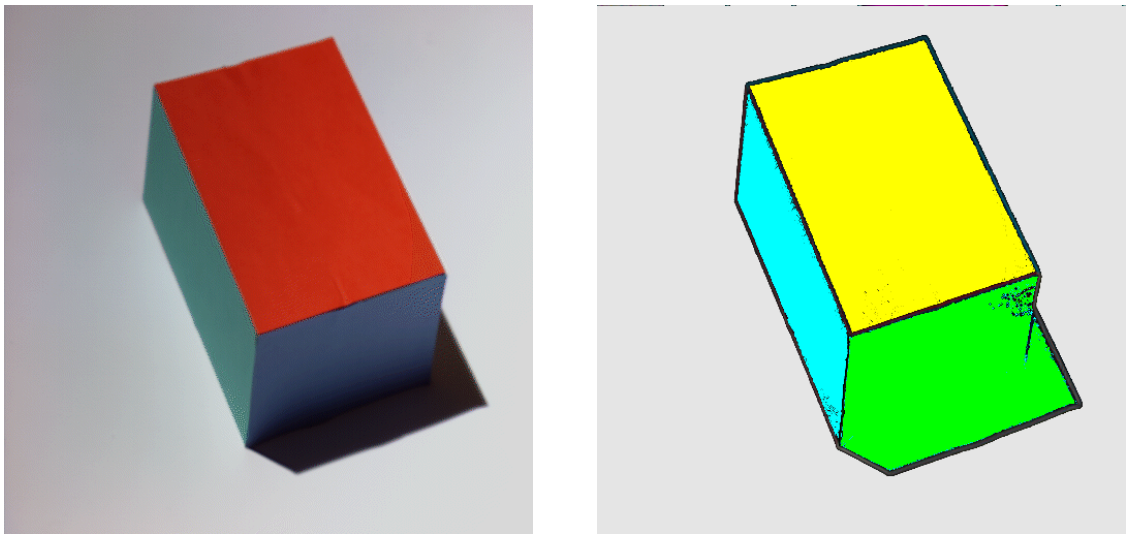


FIG. 1.25 – Segmentation d'un objet contenant une zone d'ombre.

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.

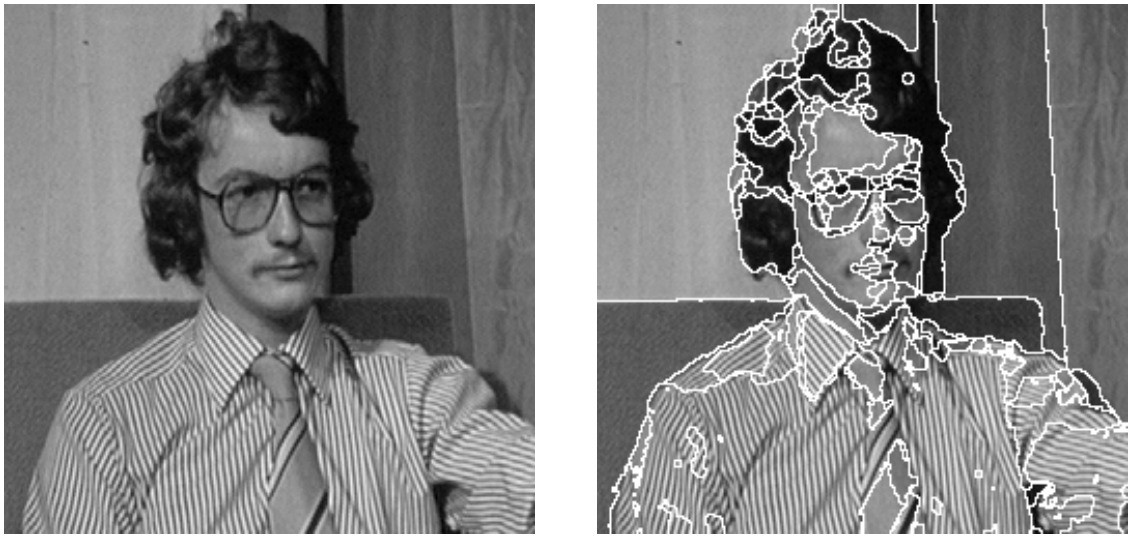


FIG. 1.26 – Segmentation d'une image d'une séquence vidéo.

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 :

– 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).

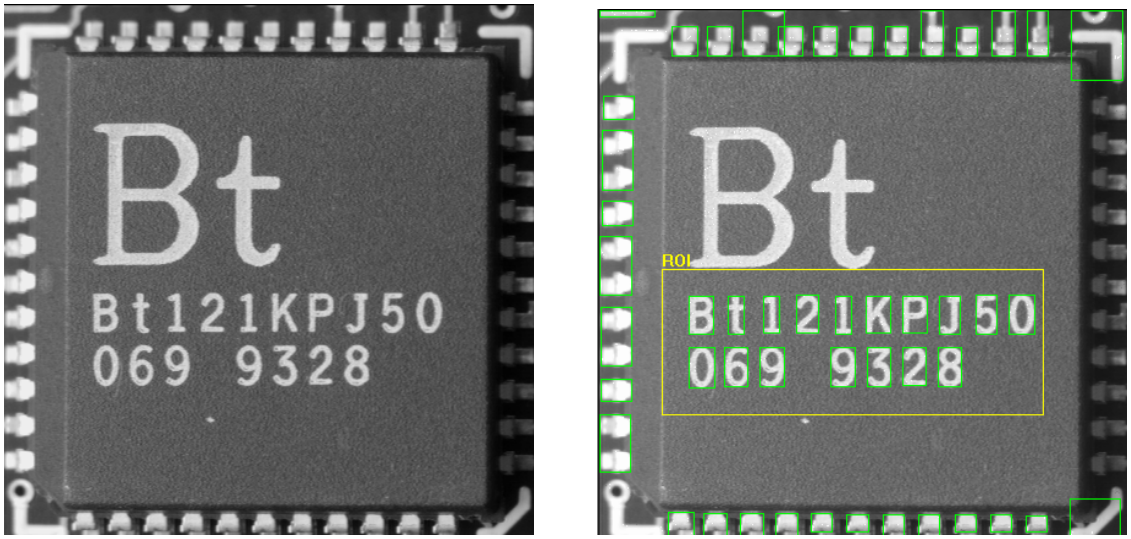


FIG. 1.27 – Reconnaissance de caractères par sélection d'une région d'intérêt.

– 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.

Chapitre 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 Transformée de FOURIER

2.2.1 Définition

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

$$\mathcal{F}(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, e^{-2\pi j(xu+yv)} dxdy \tag{2.1}$$

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

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{F}(u, v)\, e^{2\pi j(ux+vy)} dudv \tag{2.2}$$

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

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(x, y)|\, dxdy < +\infty \tag{2.3}$$

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

$$f(x, y) \rightleftharpoons \mathcal{F}(u, v) \tag{2.4}$$

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

$$\mathcal{F}(u, v) = \|\mathcal{F}(u, v)\| e^{j\theta(u,v)} \tag{2.5}$$

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

$$\mathcal{F}(-u, -v) = \mathcal{F}^*(u, v) \tag{2.6}$$

Dès lors,

$$\|\mathcal{F}(-u, -v)\| = \|\mathcal{F}(u, v)\| \tag{2.7}$$

$$\theta(-u, -v) = -\theta(u, v) \tag{2.8}$$

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 $u - v$.

2.2.2 Propriétés

1. **Séparabilité**

   (a) En permutant l'ordre d'intégration dans (2.1), nous avons

   $$\mathcal{F}(u, v) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f(x, y) e^{-2\pi jxu} dx \right] e^{-2\pi jyv} dy \tag{2.9}$$

   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 séparable[1], c'est-à-dire

$$f(x,y) = f_x(x)\, f_y(y) \tag{2.10}$$

alors

$$\mathcal{F}(u,v) = \mathcal{F}_x(u)\, \mathcal{F}_y(v) \tag{2.11}$$

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

$$c_1 f_1(x,y) + c_2 f_2(x,y) \rightleftharpoons c_1 \mathcal{F}_1(u,v) + c_2 \mathcal{F}_2(u,v) \tag{2.12}$$

3. **Homothétie**
Si $f(x,y) \rightleftharpoons \mathcal{F}(u,v)$, alors

$$f(ax,by) \rightleftharpoons \frac{1}{|ab|}\mathcal{F}\left(\frac{u}{a},\frac{v}{b}\right) \tag{2.13}$$

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

$$\mathcal{F}(x,y) \rightleftharpoons f(-u,-v) \tag{2.14}$$

5. **Translation spatiale**
Si $f(x,y) \rightleftharpoons \mathcal{F}(u,v)$, alors

$$f(x-x_0, y-y_0) \rightleftharpoons \mathcal{F}(u,v)\, e^{-2\pi j(x_0 u + y_0 v)} \tag{2.15}$$

6. **Translation fréquentielle**
Si $f(x,y) \rightleftharpoons \mathcal{F}(u,v)$, alors

$$f(x,y)\, e^{j2\pi(u_0 x + v_0 y)} \rightleftharpoons \mathcal{F}(u-u_0, v-v_0) \tag{2.16}$$

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

7. **Aires**
   Si $f(x,y) \rightleftharpoons \mathcal{F}(u,v)$, alors

$$\mathcal{F}(0,0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\, dxdy \tag{2.17}$$

   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.

$$f(0,0) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{F}(u,v)\, dudv \tag{2.18}$$

2.2.3 Illustrations

La fonction Rectangle

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

$$f(x,y) = A\text{rect}_{[a,b]}(x,y) \tag{2.19}$$

où

$$\text{rect}_{[a,b]}(x,y) = \begin{cases} 1 & |x| < \frac{a}{2}, |y| < \frac{b}{2} \\ 0 & \text{ailleurs} \end{cases} \tag{2.20}$$

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

$$\mathcal{F}(u,v) = \int_{-a/2}^{+a/2} dx \int_{-b/2}^{+b/2} dy A e^{-2\pi j(xu+yv)} \tag{2.21}$$

$$= Aab \left(\frac{\sin(\pi au)}{\pi au}\right)\left(\frac{\sin(\pi bv)}{\pi bv}\right) \tag{2.22}$$

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 2.11 de séparabilité. Nous avons donc la paire de transformées de FOURIER

$$\text{rect}_{[a,b]}(x,y) \rightleftharpoons ab \left(\frac{\sin(\pi au)}{\pi au}\right)\left(\frac{\sin(\pi bv)}{\pi bv}\right) \tag{2.23}$$

Une représentation de l'image $f(x,y)$ est donné à la figure 2.1 tandis que le module de sa transformée de FOURIER est représenté à la figure 2.2.

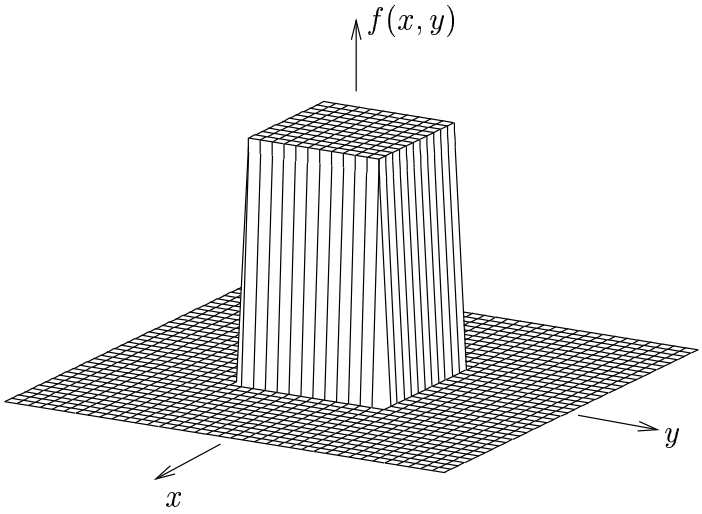


FIG. 2.1 – Illustration de la fonction Rectangle.

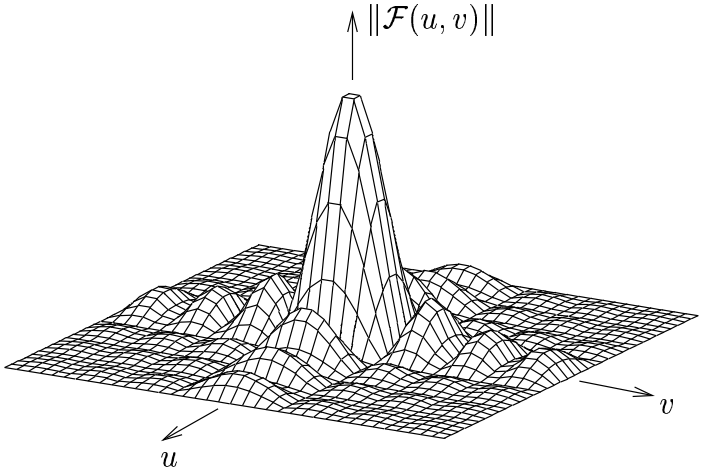


FIG. 2.2 – Module de la transformée de FOURIER de la fonction Rectangle.

FIG. 2.3 – Visualisation du module du spectre de l'image *Lena* après centrage de l'origine.

Chapitre 3

Filtrage linéaire

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.

3.1 Notion de filtre idéal

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

Définition 4 [Filtre idéal] *Un filtre est* idéal *si sa transmittance est telle que*

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

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

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

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

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).

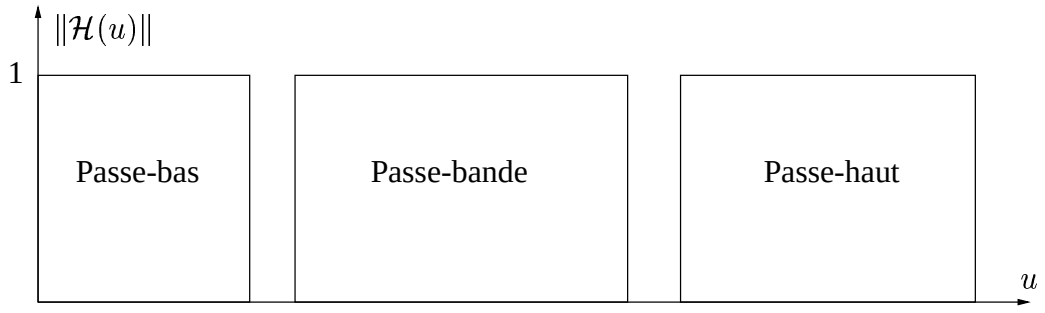


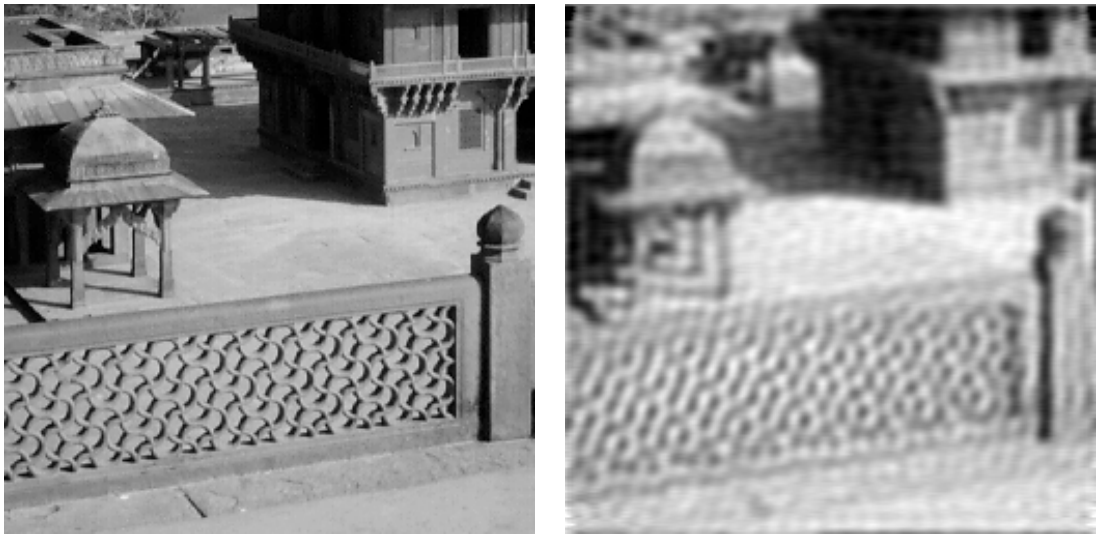
FIG. 3.1 – Filtres idéaux unidimensionnels.

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 moyenneur sur le niveau de luminance. En raison de la suppression des composantes à haute fréquence, les transitions se retrouvent adoucies.
  Le filtre passe-bas idéal circulaire est défini par une fonction de transfert de la forme

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

  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. 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.



(a) Image originale (b) Image filtrée

FIG. 3.2 – Filtrage passe-bas d'une image.

– 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

$$\mathcal{H}(u,v) = \begin{cases} 1 & \sqrt{u^2+v^2} \geq R_0 \\ 0 & \sqrt{u^2+v^2} < R_0 \end{cases} \tag{3.4}$$

  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

$$\mathcal{H}(u,v) = \begin{cases} 1 & R_0 \leq \sqrt{u^2+v^2} \leq R_1 \\ 0 & \text{sinon} \end{cases} \tag{3.5}$$

  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.

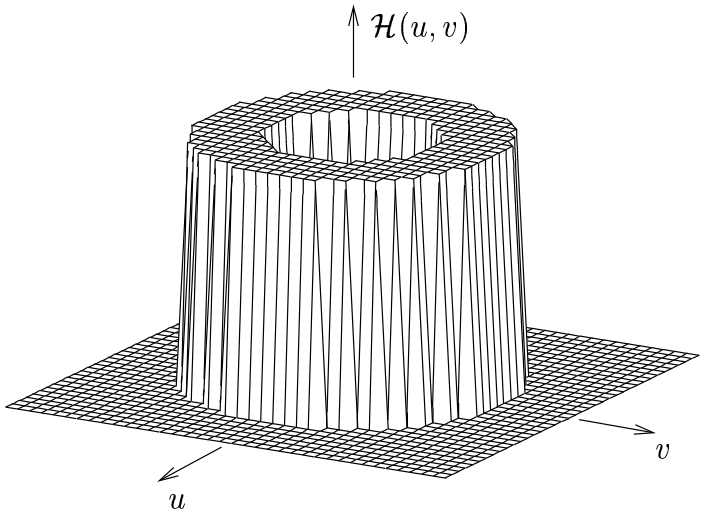


FIG. 3.3 – Fonction de transfert du filtre passe-bande idéal.

La figure 3.4 compare les effets d'un filtrage passe-bas, d'un filtrage passe-bande et d'un filtrage passe-haut [1]. 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.

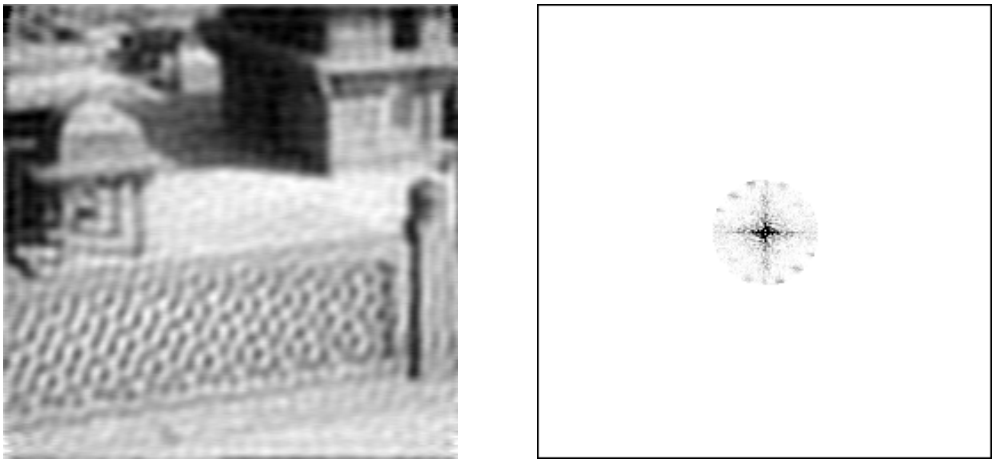
En pratique, on préfère souvent recourir à des filtres à support rectangulaire. Un tel choix se justifie par une plus grande facilité de réalisation car, en présupposant la séparabilité de l'image, le filtrage s'effectuera par mise en cascade d'un filtrage horizontal et d'un filtrage vertical. En contrepartie, 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

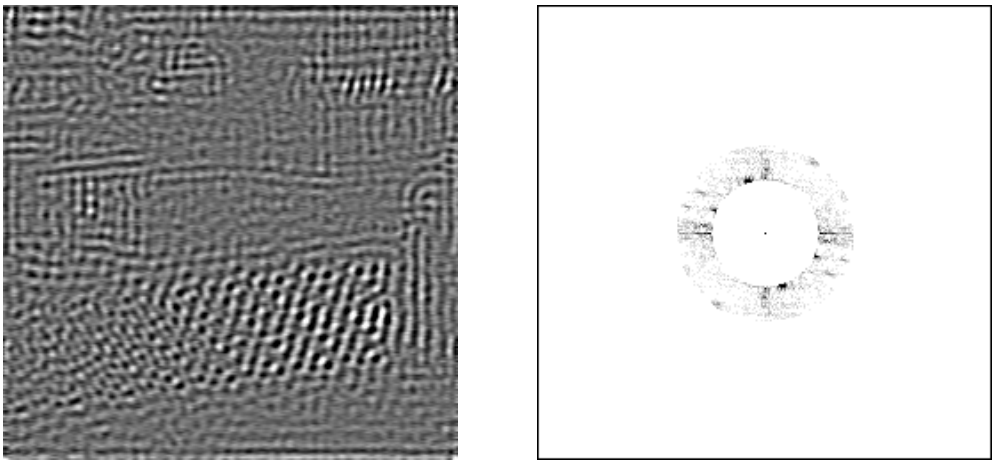
Comme dans le cas d'un système unidimensionnel, on peut interpréter l'action d'un opérateur bidimensionnel, linéaire et invariant en translation, comme un filtrage. Le problème général du filtrage bidimensionnel consiste à élaborer un filtre qui possède la réponse impulsionnelle voulue et qui soit réalisable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

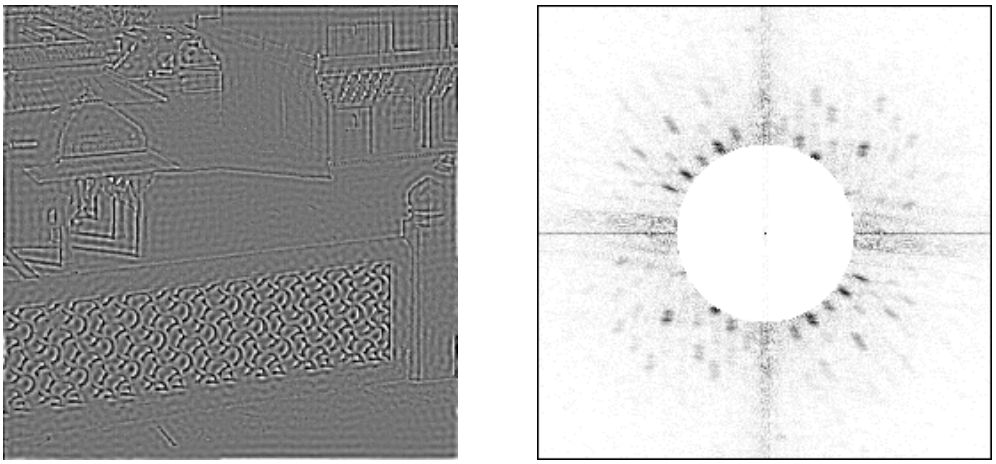
[1] Pour la visualisation des spectres, nous avons adopté la convention de *vidéo inverse* ; elle consiste simplement à inverser l'échelle de luminance.



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.

FIG. 3.4 – Images filtrées au moyen de filtres idéaux circulaires et spectres correspondants.

est utilisé pour modifier le contenu spectral d'une image afin d'en retirer l'information recherchée.

3.2.1 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.

3.2.2 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

$$\mathcal{H}\left(u,v\right) = \frac{1}{1+\left(\frac{\sqrt{u^2+v^2}}{R_0}\right)^{2n}} \tag{3.6}$$

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.

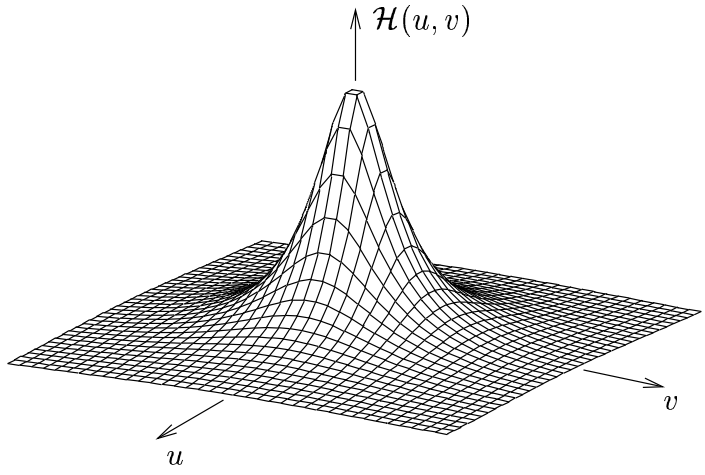
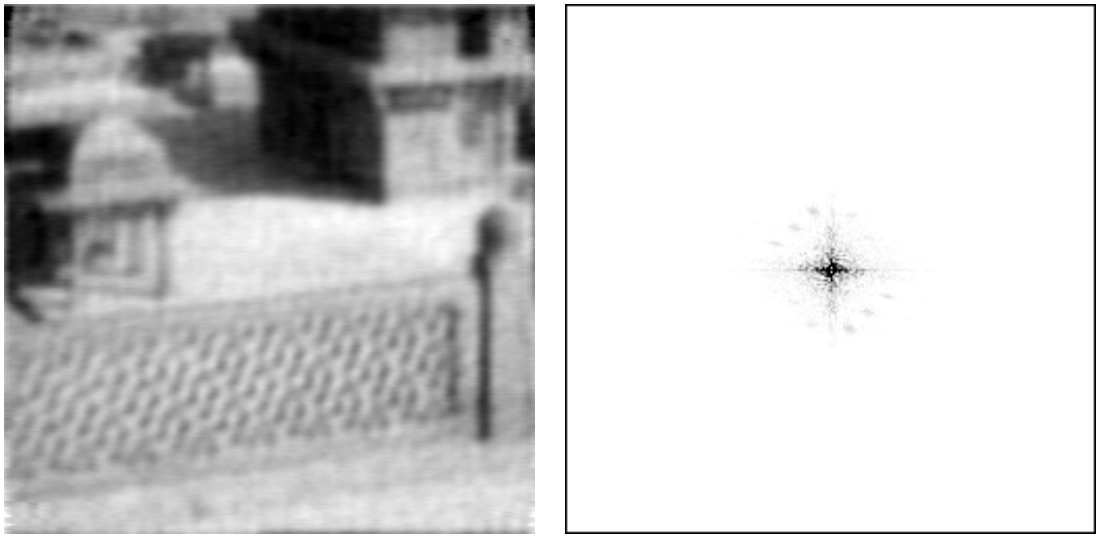


FIG. 3.5 – Transmittance du filtre passe-bas de BUTTERWORTH pour $n = 1$.



(a) Image filtrée (b) Spectre de l'image filtrée

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

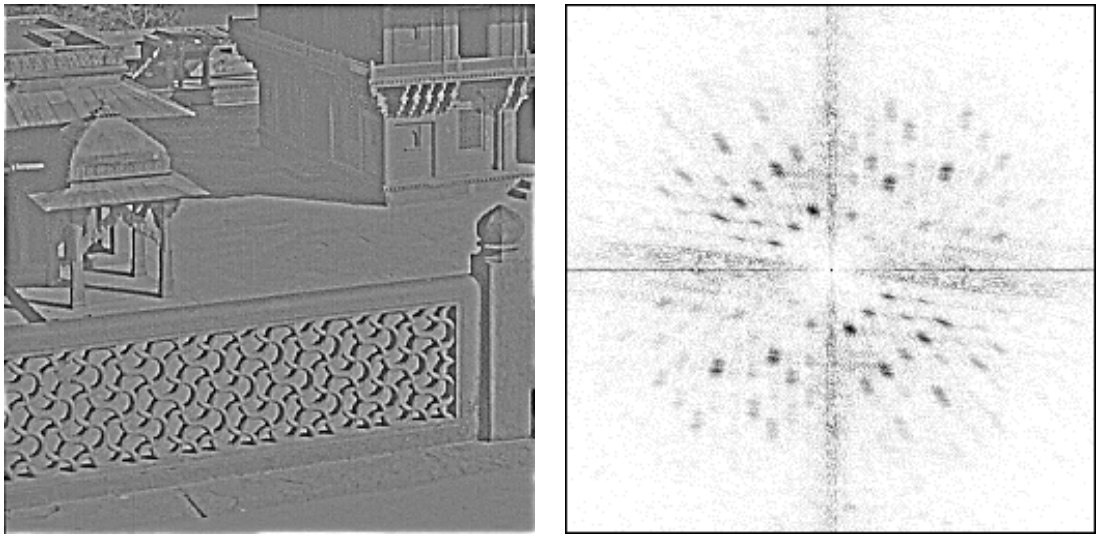
3.2.3 Filtres passe-haut

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

$$\mathcal{H}(u,v) = \frac{1}{1 + \left(\frac{R_0}{\sqrt{u^2+v^2}}\right)^{2n}} \tag{3.7}$$

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.

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.



(a) Image filtrée (b) Spectre de l'image filtrée

FIG. 3.7 – Image filtrée par un filtre passe-haut de BUTTERWORTH d'ordre 1 $(f_c = 50)$.

Chapitre 4

Rehaussement et restauration

4.1 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.

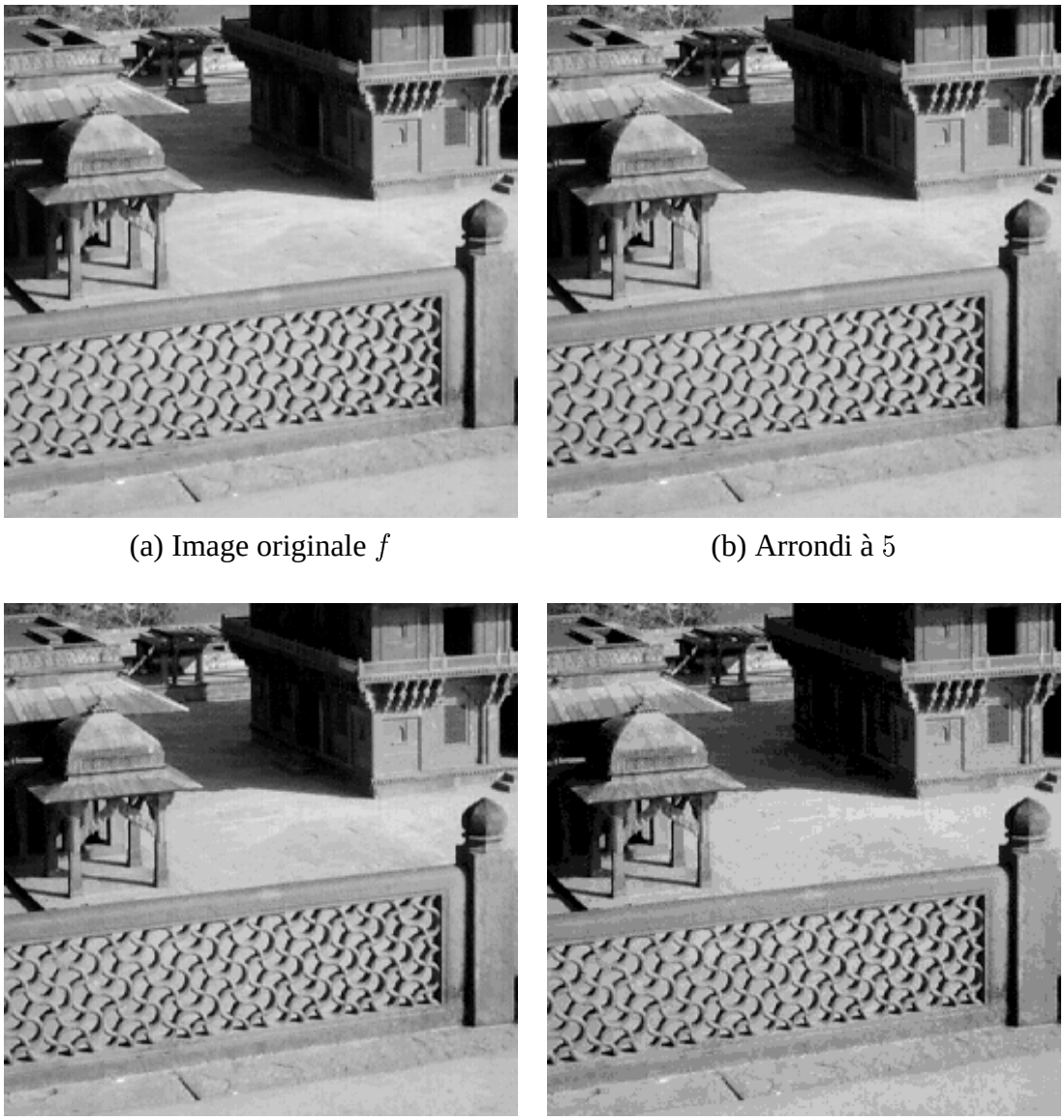
Dans les deux cas, on part à la recherche d'une estimation $\widehat{f}(x, y)$ de l'image $f(x, y)$ qui soit visuellement ou analytiquement intéressante, c'est-à-dire qu'elle minimise une fonction d'erreur $e(x, y) = e(\widehat{f}, f)$. Comme il n'existe pas de fonction d'erreur capable de modéliser 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

$$e(\widehat{f}, f) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (\widehat{f}(x, y) - f(x, y))^2 \tag{4.1}$$

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 4.1 illustre la marge de manœuvredont on dispose pour du rehaussement par rapport à la restauration. L'image 4.1(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.



(a) Image originale f

(b) Arrondi à 5

(c) Arrondi à 10

(d) Arrondi à 20

FIG. 4.1 – Seuil de visibilité pour une opération d'arrondi.

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.

4.2 Rehaussement

4.2.1 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 5 *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.

4.2.2 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

$$g(l) = l + \alpha l(L - l) \tag{4.2}$$

La figure 4.2(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

$$g(l) = \frac{L}{2}\left(1 + \frac{\sin\left[\alpha\pi(\frac{l}{L} - \frac{1}{2})\right]}{\sin\left[\alpha\frac{\pi}{2}\right]}\right), \quad \alpha \in]0, 1[\tag{4.3}$$

comme représenté à la figure 4.2(b) pour $\alpha = 0,95$. Cette dernière rehausse les valeurs centrales.

4.2.3 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

$$\widehat{f}(x, y) = f(x, y) - \alpha\nabla^2 f(x, y) \tag{4.4}$$

La figure 4.3 montre le résultat de l'opérateur d'estimation pour $\alpha = \frac{1}{3}$. Très clairement, les contours ont été accentués.

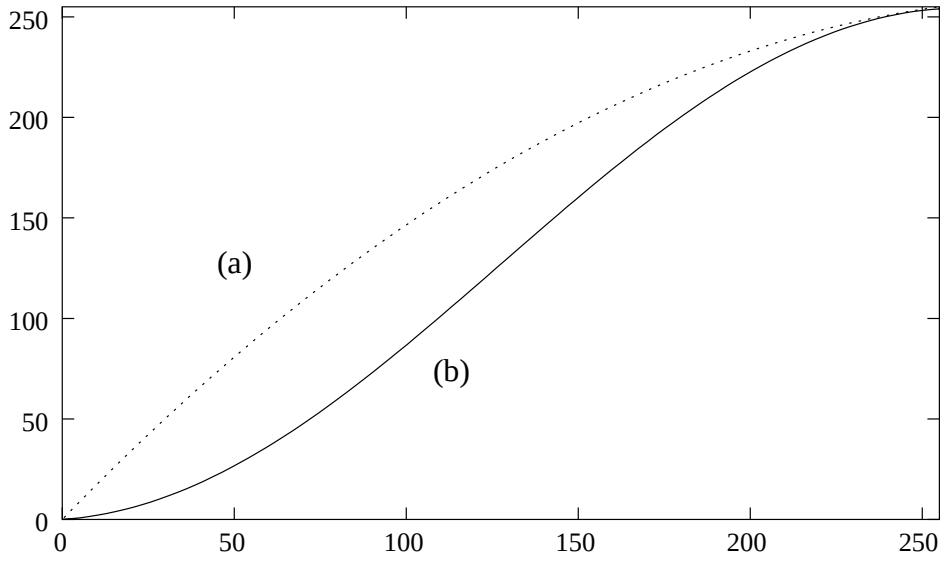
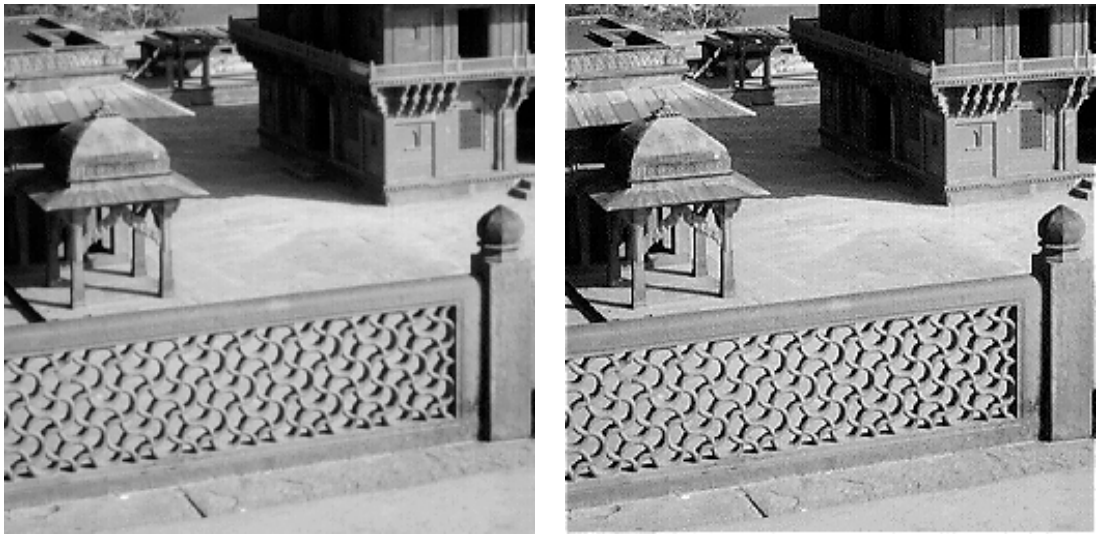


FIG. 4.2 – Différentes courbes de rehaussement.



(a) Image originale f (b) $f - \alpha\nabla^2 f$

FIG. 4.3 – Rehaussement des contours.

4.2.4 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

$$l_o = \phi(l_f) \tag{4.5}$$

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,

$$l_f = \phi^{-1}(l_o) \tag{4.6}$$

Si l'on interprète l'histogramme comme une densité de probabilité, on arrive à une égalité entre les "masses" que sont

$$hist_o(l_o)dl_o = hist_f(l_f)dl_f \tag{4.7}$$

Dès lors,

$$hist_o(l_o) = \frac{hist_f(l_f)}{dl_o/dl_f} \tag{4.8}$$

$$= \left.\frac{hist_f(l_f)}{d\phi(l_f)/dl_f}\right|_{l_f=\phi^{-1}(l_o)} \tag{4.9}$$

et donc

$$hist_o(l) = \frac{hist_f(\phi^{-1}(l))}{\frac{d\phi}{dl}(\phi^{-1}(l))} \tag{4.10}$$

Exemple : fonction parabolique

Considérons la fonction ϕ suivante

$$l_o = \phi(l_f) = \frac{1}{L}(l_f)^2 \tag{4.11}$$

Par application de la formule 4.10,

$$hist_o(l) = \frac{hist_f(L\sqrt{l})}{\frac{2}{L}L\sqrt{l}} = \frac{hist_f(L\sqrt{l})}{2\sqrt{l}} \tag{4.12}$$

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 4.10 devient donc

$$\frac{d\phi}{dl}(l) = \frac{L}{A(D(f))} hist_f(l) \tag{4.13}$$

Par intégration des membres de cette dernière relation,

$$\phi(l) = \frac{L}{A(D(f))} \int_0^l hist_f(k)dk \tag{4.14}$$

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

$$p(l) = \frac{hist_f(l)}{A(D(f))} \tag{4.15}$$

De plus, la fonction de répartition F est l'intégrale cumulée des probabilités

$$F(l) = \int_0^l p(k)dk = \frac{1}{A(D(f))} \int_0^l hist_f(k)dk \tag{4.16}$$

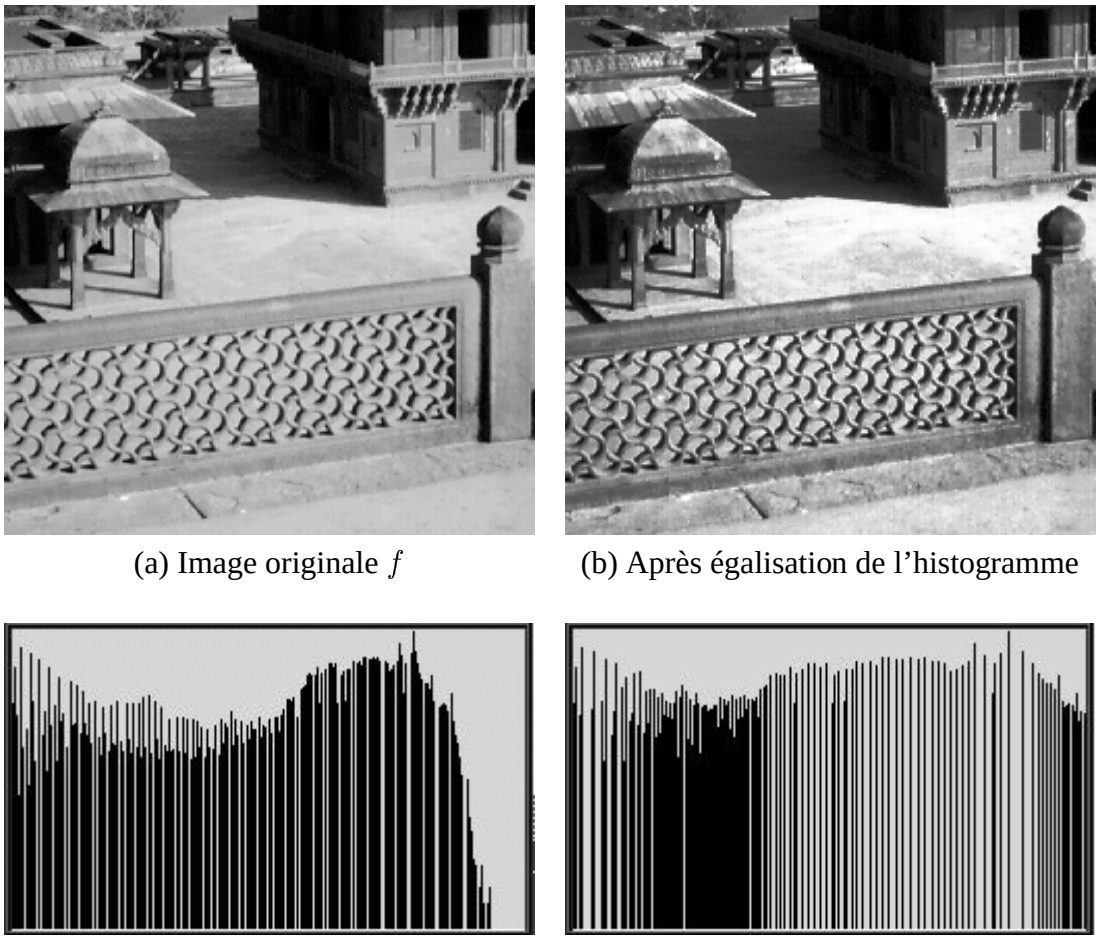
Dès lors,

$$\phi(l) = LF(l) \tag{4.17}$$

Autrement dit, l'image $o(x, y)$ traitée par la fonction qui effectue une égalisation de l'histogramme vaut

$$o(x, y) = LF(f(x, y)) \tag{4.18}$$

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 4.4.



(a) Image originale f (b) Après égalisation de l'histogramme

(c) Histogramme original (d) Histogramme après égalisation

FIG. 4.4 – Égalisation d'histogramme.

Chapitre 5

Segmentation d'images

5.1 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.

5.1.1 Définition

Définition 6 *En général, on considère que l'objectif de la* segmentation *revient à construire une série de régions $R_1, \ldots, R_n$ telles*

$$\mathcal{E} = \bigcup_{i=1}^{n} R_n \quad \textit{et} \quad \forall i \neq j,\ R_i \cap R_j = \emptyset \tag{5.1}$$

Définition 7 *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.*

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 s'agit-il d'un procédé global.

5.2 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 5.1 illustre la difficulté à segmenter par simple seuillage ; l'origine du problème est l'éclairage non uniforme de la scène, problème courant en pratique.

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 5.2.

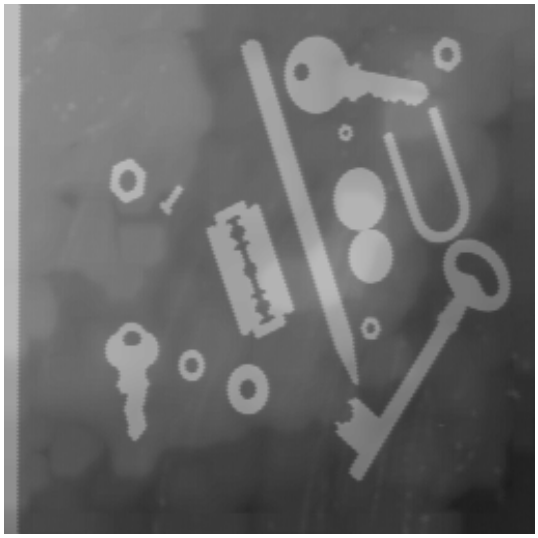
5.3 Segmentation par détection de régions

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.

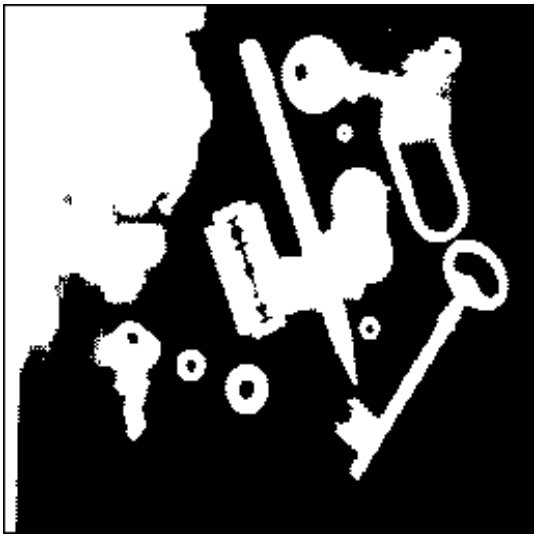
5.3.1 Ligne de partage des eaux

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.

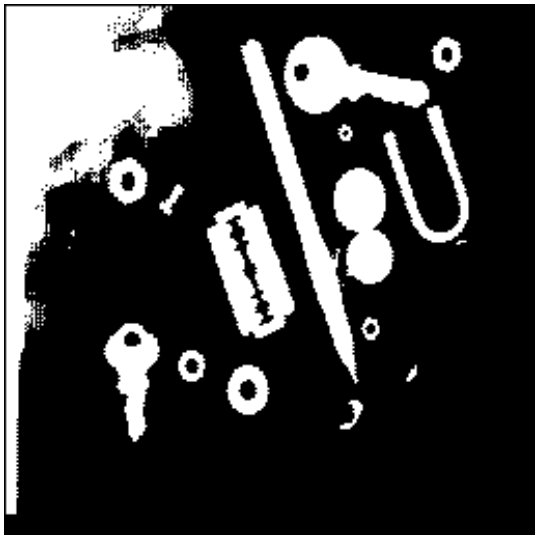
Puisque les valeurs de luminance ne semblent pas convenir directement, on est tenté de recourir à la norme des gradients morphologiques. L'inconvénient majeur de ceci est que le gradient est lui-même une fonction et, que comble de malheur, il faut à nouveau essayer des seuils pour segmenter. Bien sûr, 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.



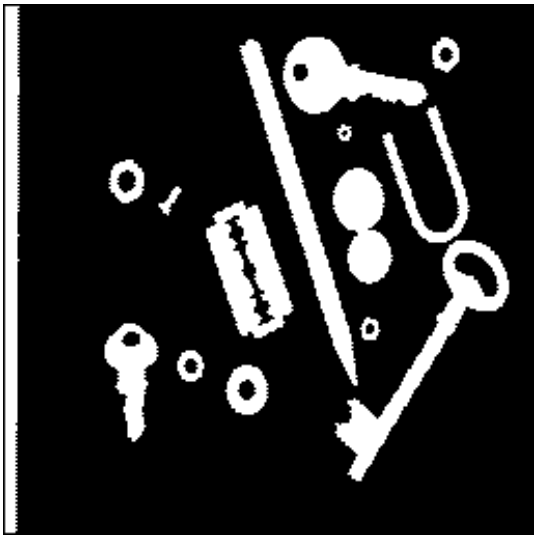
(a) Image originale



(b) Seuillage à 110



(c) Seuillage à 128



(d) Seuillage après égalisation

FIG. 5.1 – Seuillages.

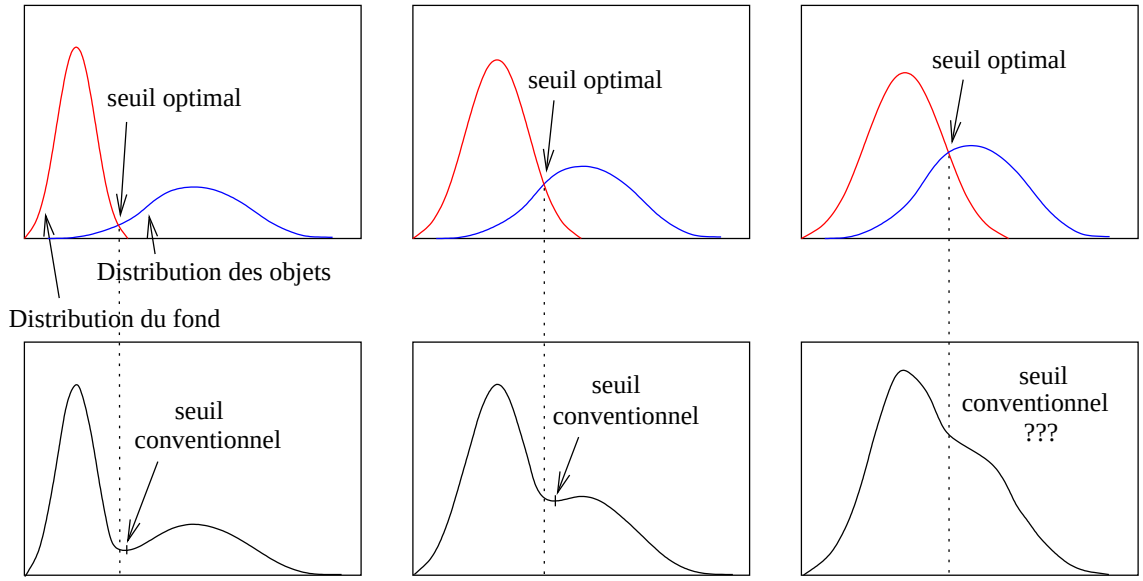


FIG. 5.2 – Seuillage optimal (d'après [8]).

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 5.3.

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 quel-

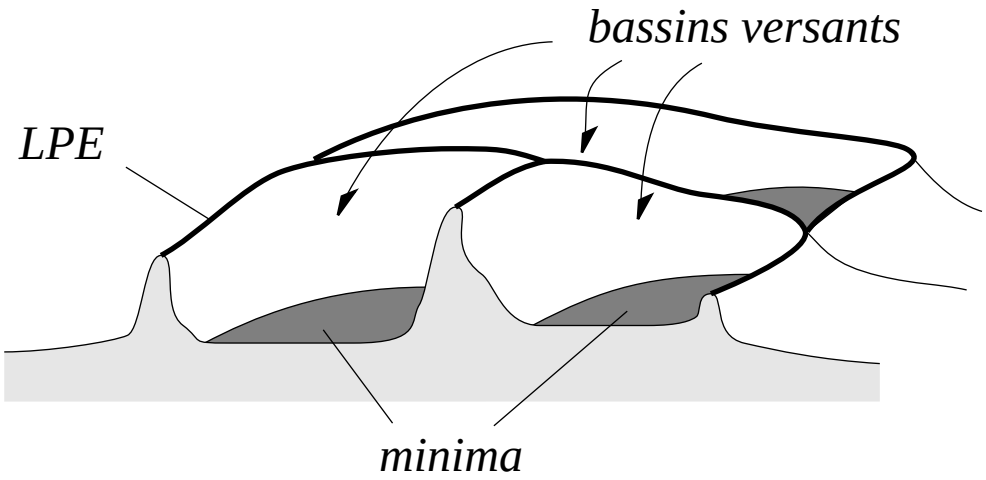


FIG. 5.3 – Minima, bassins versants et ligne de partage des eaux.

conque 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 fausse 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.

5.3.2 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 8 [Chemin géodésique] *Un* chemin géodésique de longueur l *entre deux points* s *et* t *est une famille de* $l+1$ *pixels* $x_0 = s,\ x_1,\ \ldots,\ x_l = t$ *telle que*

$$\forall i \in [0, l],\ x_i \in X \quad \textit{et} \quad \forall i \in [0, l],\ x_{i-1},\ x_i \textit{ sont voisins} \tag{5.2}$$

Parmi tous les chemis, le plus court détermine la distance géodésique. La figure 5.4 fournit un exemple de chemin géodésique.

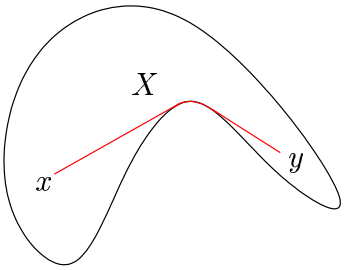


FIG. 5.4 – Le plus court chemin géodésique entre x et y.

Définition 9 [Distance géodésique] *La* distance géodésique *entre les points* s *et* t *est la longueur du plus court chemin 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 $ZI(domaine =$X, $centre = Z_i)$ et sa frontière $FR(domaine =$X, $centre = Z_i)$.

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 5.5 montre un exemple de squelette géodésique.

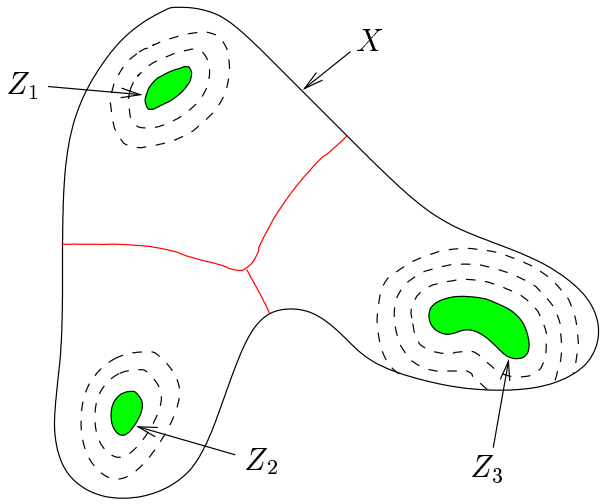


FIG. 5.5 – Squelette géodésique.

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. Mais avant de passer à une description formelle de la construction de la ligne de partage des eaux, nous proposons le dessin de la figure 5.6 pour illustrer la démarche.

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

$$\mathcal{C}_h(M_i) = \mathcal{C}(M_i) \cap T_h(f) \tag{5.3}$$

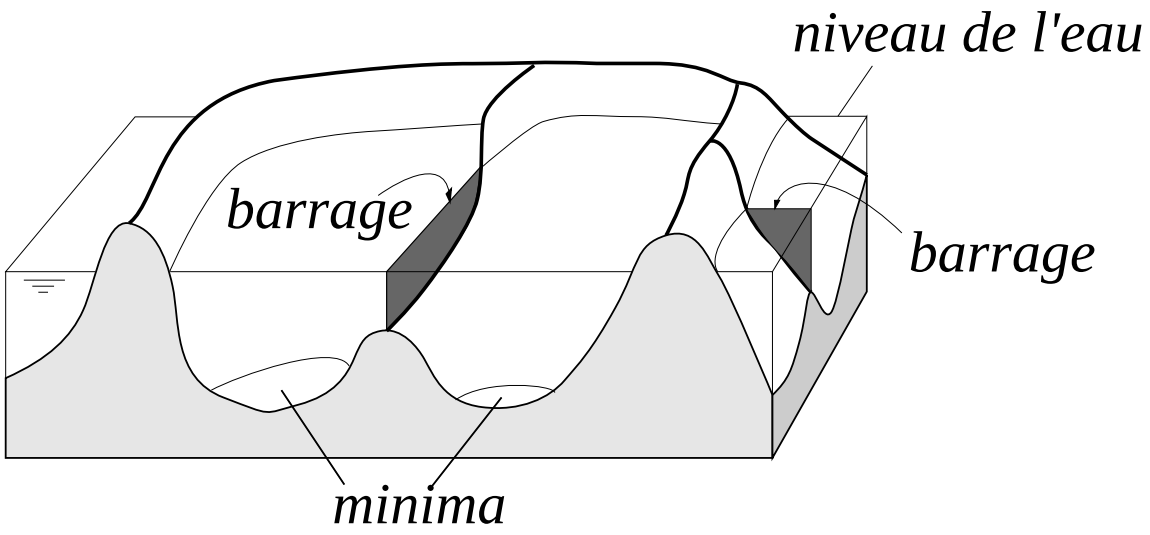


FIG. 5.6 – Construction d'un barrage au raccord de deux bassins versants.

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 $\mathcal{C}_k(M_i) = \phi$ 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

$$\forall h \in [h_{min} + 1, h_{max}] : \mathcal{C}_h(M) = ZI_h \cup Min_h \tag{5.4}$$

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 [6]. 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 [7] cherchent des expressions plus agréables.

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.

Chapitre 6

Tomographie par rayons X

Auteurs : FRÉDÉRIC NOO et JEAN-MARC WAGNER

Le mot *tomographie* signifie textuellement *visualisation (graphe) de coupes (tranches) au sein d'un corps 3D*. Techniquement parlant, il s'agit de la méthode utilisée dans les scanners pour visualiser l'intérieur d'un corps de façon non-invasive. De nos jours, les scanners sont très utilisés dans les hôpitaux comme moyen de diagnostic. Toutefois, ils servent aussi dans le monde industriel afin, par exemple, de détecter des fissures dans des pièces mécaniques.

Sur le marché, on peut trouver différents types de scanners : le scanner à rayons X, le scanner à émission de photons, et aussi le scanner à résonance magnétique nucléaire. Dans ce chapitre, nous ne présenterons que les principes de l'imagerie par rayons X.

6.1 Introduction

La tomographie par rayons X utilise les propriétés d'absorption des rayons X par la matière pour visualiser l'intérieur d'un corps 3D. Elle repose sur l'observation suivante : tout faisceau de rayons X passant au travers d'un matériau subit une perte de photons qui a la particularité d'être différente suivant le type de matériau traversé. Cette perte s'exprime sous la forme suivante : $\Delta N = -f\, N\, \Delta l$ (voir la figure 6.1), où f est un paramètre, appelé *facteur d'atténuation*, qui est propre au matériau considéré. On observe que différents matériaux correspondent à différentes valeurs de f et que donc f peut être utilisé pour classer les matériaux. Notons que f s'exprime en $[mm^{-1}]$ et prend la même valeur quelque soit l'angle d'incidence du faisceau de rayons X.

Considérons une coupe dans un corps donné, comme illustré à la figure 6.2, et utilisons des coordonnées cartésiennes (x, y) pour situer un pixel au sein de cette coupe. Le but de la tomographie par rayons X est d'estimer la fonction $f\,(x, y)$ qui attribue à chaque pixel de la coupe la valeur moyenne du facteur d'atténuation f sur la surface du pixel. Lorsque la fonction $f\,(x, y)$ est connue, on obtient aisément une visualisation des structures au sein de la coupe en représentant $f\,(x, y)$ comme une image sur un moniteur, avec différents tons de gris pour différentes valeurs de $f\,(x, y)$.

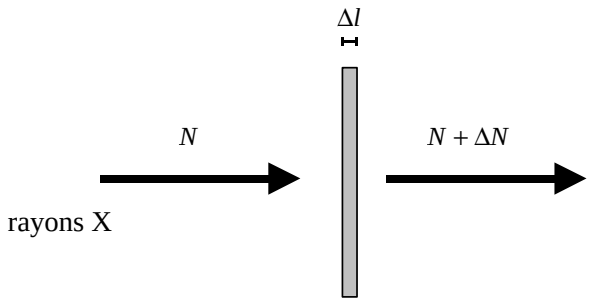


FIG. 6.1 – Perte de photons au travers d'un matériau ($\Delta N < 0$).

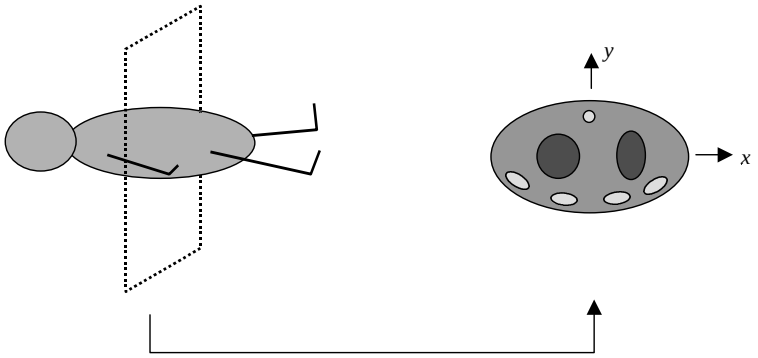


FIG. 6.2 – Visualisation d'une coupe.

6.2 Mesures

Comme nous venons de le voir, le but de la tomographie par rayons X est de trouver la fonction d'atténuation $f\,(x, y)$. Pour atteindre cet objectif, on effectue tout d'abord des mesures qui fournissent des équations mathématiques faisant intervenir $f\,(x, y)$. Ensuite, on résout ces équations pour obtenir $f\,(x, y)$.

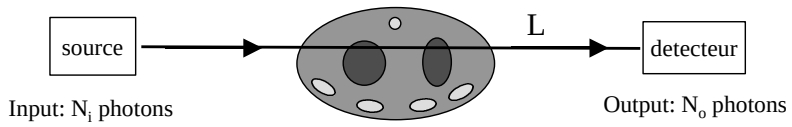


FIG. 6.3 – Principe d'acquisition des mesures.

Dans les scanners à rayons X, l'information accessible sur $f\,(x, y)$ se présente sous la forme d'intégrales de $f\,(x, y)$ sur différentes droites du plan (x, y). On parle d'intégrales de lignes. La figure 6.3 illustre la situation. Conceptuellement, on utilise une source de rayons X et un détecteur (compteur) de photons. La source envoie un faisceau de N_i photons dans la direction du détecteur au travers de la coupe, on détermine le nombre N_o de photons qui passent, puis, par le biais de la relation $\Delta N = -f\,N\,\Delta l$, on trouve

$$\int_{input}^{output} \frac{dN}{N} = -\int_{input}^{output} f\,(x, y)\;dl \quad \rightarrow \quad \log\left(\frac{N_i}{N_o}\right) = \int_{droite\;L} f\,(x, y)\;dl \tag{6.1}$$

La quantité $\log\,(N_i/N_o)$ constitue une mesure. Il s'agit de l'intégrale de $f\,(x, y)$ sur la droite qui relie la source au détecteur.

Pour calculer (reconstruire) $f\,(x, y)$ de façon exacte, nous verrons que les intégrales de lignes de $f\,(x, y)$ doivent être connues pour toutes les droites qui passent au travers de la coupe étudiée.

Il y a deux façons différentes de collecter les mesures sur l'ensemble des droites passant au travers de la coupe : en géométrie parallèle (voir la figure 6.4) et en géométrie en éventail (voir la figure 6.5).

En géométrie parallèle, les mesures $\log\,(N_i/N_o)$ sont réalisées sur des ensembles de droites parallèles. On déplace la source et le détecteur dans une direction donnée $\vec{\theta}$ pour obtenir les mesures sur toutes les droites perpendiculaires au vecteur $\vec{\theta}$. Ensuite, on applique une rotation à l'ensemble source-détecteur afin d'obtenir des nouvelles mesures en répétant le mouvement de translation. Le mouvement de rotation s'effectue sur un intervalle angulaire de longueur π.

En géométrie en éventail, on utilise tout un banc de détecteurs. La source envoie simultanément ses rayons en direction de tous les détecteurs de ce banc de telle sorte que l'on obtient pour une position angulaire donnée tout un jeu de mesures correspondant à des droites formant

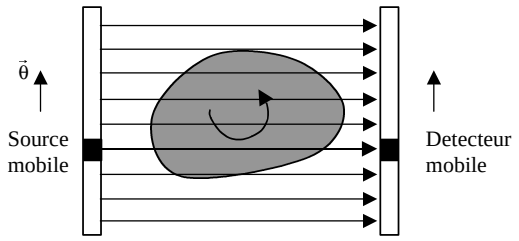


FIG. 6.4 – Acquisition des mesures en géométrie parallèle.

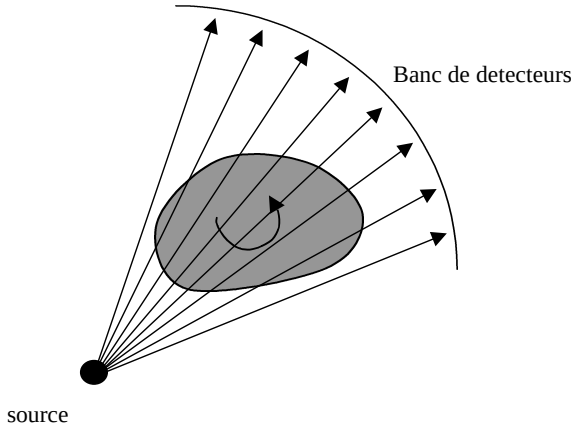


FIG. 6.5 – Acquisition des mesures en géométrie en éventail.

un éventail. Pour couvrir toutes les droites du plan, on considère généralement un mouvement de rotation de l'ensemble source-banc sur un intervalle de longueur 2π, ce qui, notons-le, conduit à une redondance d'ordre 2 dans les mesures. Par rapport à la géométrie parallèle, on observe que la géométrie en éventail permet une acquisition plus rapide des données, grâce à la suppression du mouvement de translation.

Conceptuellement, l'appareillage requis pour la géométrie en éventail est plus coûteux que celui requis pour la géométrie parallèle. Toutefois, la géométrie en éventail est préférée à la géométrie parallèle car elle permet d'acquérir les mesures plus rapidement. La vitesse d'acquisition des mesures est de fait un facteur très important en tomographie. Dans la discussion ci-dessus, nous avons fait l'hypothèse implicite que la fonction $f(x, y)$ ne varie pas au cours du temps. Lorsque le patient bouge durant l'acquisition des mesures (ne serait-ce qu'en respirant), l'expression de $f(x, y)$ peut changer d'une intégrale de ligne à une autre et la qualité des images obtenues est alors médiocre. Les scanners actuels utilisent la géométrie en éventail et peuvent collecter les mesures pour une coupe donnée en moins de 750 $[ms]$, de telle sorte que le patient ne bouge pratiquement pas.

Pour la formulation mathématique du problème de reconstruction de $f(x, y)$ à partir de ses intégrales de lignes, notons qu'il est plus facile de travailler en géométrie parallèle, comme effectué ci-après.

6.3 Reconstruction en géométrie parallèle

6.3.1 Notations et définitions

Comme illustré à la figure 6.6, nous décrivons les droites dans le plan (x, y) à l'aide de deux coordonnées : $s \in]-\infty, +\infty[$ et $\theta \in [0, 2\pi[$. La droite définie par s et θ est notée $L(\theta, s)$ tandis que la mesure le long de la droite $L(\theta, s)$ est notée $g(\theta, s)$. Notons que θ et s ne sont pas des coordonnées polaires : lorsque θ_1 et θ_2 sont tels que $\theta_2 \neq \theta_1 + k\pi$, on a $L(\theta_1, s = 0) \neq L(\theta_2, s = 0)$. Notons aussi que $L(\theta + \pi, -s) \equiv L(\theta, s)$ et donc $g(\theta + \pi, -s) = g(\theta, s)$.

Mathématiquement parlant, on dit que $g(\theta, s)$ est un échantillon de la transformée de RADON de $f(x, y)$. (La transformée de RADON de $f(x, y)$ est la fonction qui associe à chaque droite du plan (x, y), l'intégrale de $f(x, y)$ sur cette droite.) On note cette transformée sous la forme

$$g(\theta, s) = (\mathcal{R}f)(\theta, s) = \int_{-\infty}^{+\infty} dt\, f(s\vec{\theta} + t\vec{\theta}^{\perp}) \tag{6.2}$$

avec $\vec{\theta} = (\cos\theta, \sin\theta)$, $\vec{\theta}^{\perp} = (-\sin\theta, \cos\theta)$ et $f(x, y) \equiv f(\vec{x})$ pour $\vec{x} = (x, y)$.

En pratique, l'ensemble des mesures $g(\theta, s)$ obtenues pour une valeur fixée de θ avec $s \in]-\infty, +\infty[$ est appelé *projection de* $f(x, y)$, tandis que l'ensemble des mesures $g(\theta, s)$ obtenues pour $s \in]-\infty, +\infty[$ et $\theta \in [0, 2\pi[$ est appelé *sinogramme*. On parle de sinogramme parce que les données associées au point objet $f(x, y) = \delta(x - x_0)\,\delta(y - y_0)$ sont uniquement non-nulles le long de la sinusoïde $s = x_0 \cos\theta + y_0 \sin\theta$ dans le domaine (θ, s). L'ensemble des mesures obtenues pour une fonction quelconque $f(x, y)$ peut donc s'exprimer comme une somme de sinusoïdes en coordonnées (θ, s). Les figures 6.7, 6.8 et 6.9 illustrent les sinogrammes que l'on

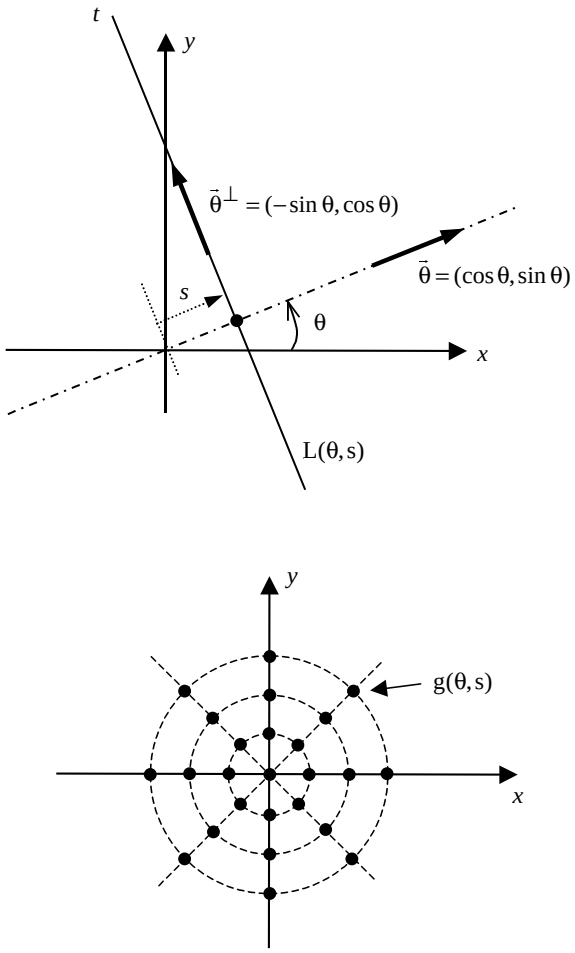
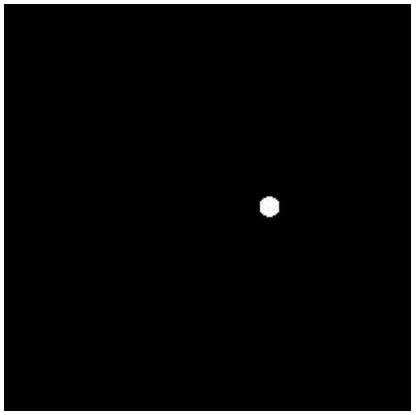


FIG. 6.6 – Représentation de droites dans l'espace (x, y).

peut obtenir pour un point objet, trois points objet, puis le fantôme de SHEPP et LOGAN qui est largement utilisé en pratique pour tester la stabilité numérique des algorithmes de reconstruction.



$f(x, y)$

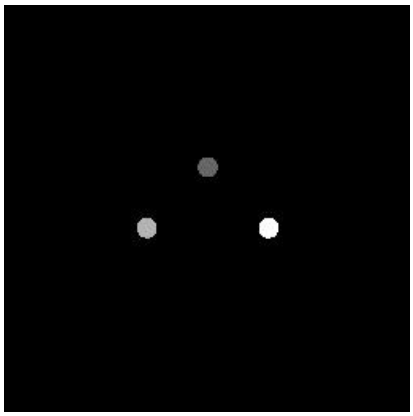
$g(\theta, s)$

FIG. 6.7 – Sinogramme pour un point objet.

6.3.2 Lien entre une image et sa transformée de Radon

En 1917, Radon établit le lien existant entre une image $f(x, y)$ et ses mesures $g(\theta, s)$ telles que décrites ci-avant. Cette relation est connue sous l'appellation *théorème du profil central* :

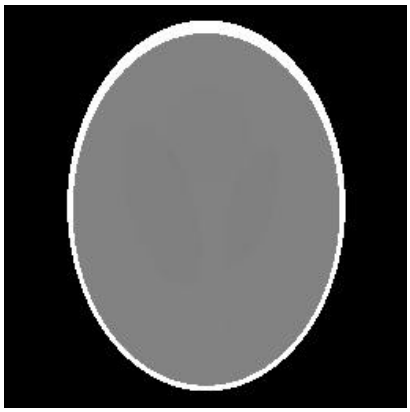
$$(\mathcal{F}_1 g)(\theta, \sigma) = (\mathcal{F}_2 f)(\sigma \cos\theta, \sigma \sin\theta) \tag{6.3}$$



$f(x, y)$

$g(\theta, s)$

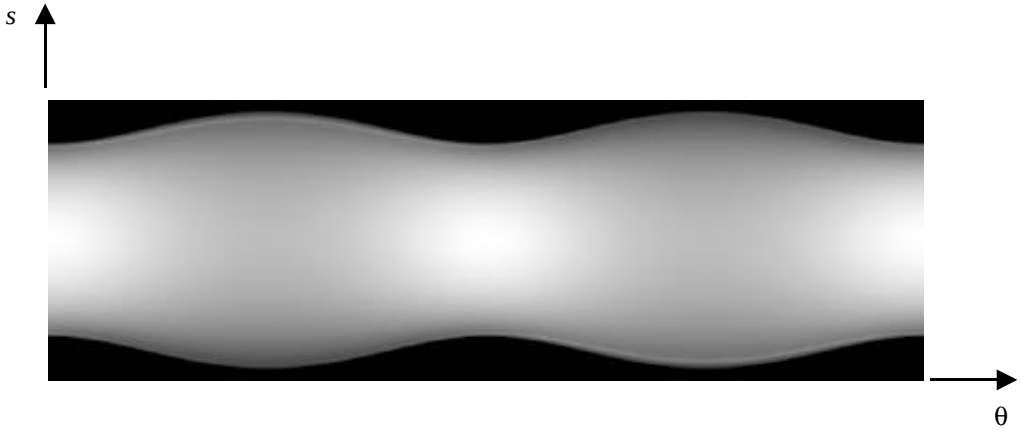
FIG. 6.8 – Sinogramme pour 3 points objet.



Echelle de gris [0,2]

Echelle de gris [1.0,1.04]

$f(x, y)$



$g(\theta, s)$

FIG. 6.9 – Sinogramme du fantôme de SHEPP-LOGAN.

ou encore

$$(\mathcal{F}_1 g)(\theta, \sigma) = (\mathcal{F}_2 f)(\sigma\vec{\theta}) \tag{6.4}$$

si l'on utilise des notations vectorielles. Dans cette formule, $(\mathcal{F}_1 g)(\theta, \sigma)$ représente la transformée de FOURIER 1D (une dimension) de $g(\theta, s)$ par rapport à s, i.e.

$$(\mathcal{F}_1 g)(\theta, \sigma) = \int_{-\infty}^{+\infty} ds\, e^{-2\pi j s\sigma}\, g(\theta, s) \tag{6.5}$$

tandis que $(\mathcal{F}_2 f)(\sigma\vec{\theta})$ est la transformée de FOURIER 2D de $f\,(x, y)$ évaluée en $\vec{X} = \sigma\vec{\theta}$

$$(\mathcal{F}_2 f)(X, Y) = \int_{-\infty}^{+\infty} dx \int_{-\infty}^{+\infty} dy\, e^{-2\pi j x X}\, e^{-2\pi j y Y}\, f(x, y) = \int_{\mathbb{R}^2} d\overrightarrow{x}\, e^{-2\pi j\, \overrightarrow{x}\,\overrightarrow{X}}\, f(\overrightarrow{X}) = (\mathcal{F}_2 f)(\overrightarrow{X})$$

avec $\vec{X} = (X, Y)$.

Le théorème du profil central peut se lire comme suit.

Théorème 10 [Théorème du profil central] *La transformée de* FOURIER *de la projection de $f\,(x, y)$ à l'angle θ est identique à la transformée de* FOURIER *2D de $f\,(x, y)$ le long de la droite de direction $\vec{\theta}$ qui passe par l'origine dans le domaine fréquentiel de $f\,(x, y)$.*

La figure 6.10 illustre la situation.

6.4 Algorithmes de reconstruction

Il existe trois méthodes analytiques pour reconstruire $f\,(x, y)$ à partir de ses intégrales de lignes : (i) la méthode directe de FOURIER, (ii) la méthode du filtrage de la rétroprojection, et (iii) la méthode de rétroprojection des projections filtrées. D'un point de vue pratique, la plus intéressante des méthodes est la troisième car elle est numériquement plus stable.

6.4.1 Méthode directe de FOURIER

La méthode directe de Fourier correspond à une implémentation directe du théorème du profil central. Celui-ci nous apprend que la reconstruction de $f\,(x, y)$ à partir de ses intégrales de lignes est uniquement possible lorsque les projections de $f\,(x, y)$ sont connues pour tout angle θ situé dans un intervalle angulaire de longueur π. Lorsque cette condition n'est pas satisfaite, il est impossible d'obtenir $(\mathcal{F}_2 f)(\vec{X})$ en tout point $\vec{X}$, et donc impossible de reconstruire $f\,(x, y)$, vu que la relation entre $f\,(x, y)$ et sa transformée de FOURIER est biunivoque.

Lorsque les projections de $f\,(x, y)$ sont connues sur un intervalle angulaire de longueur π, il est possible de reconstruire $f\,(x, y)$ comme expliqué à la figure 6.11.

Il s'agit de la méthode directe de FOURIER qui peut se résumer comme suit :

1. Calcul de la transformée de FOURIER 1D des projections par FFT (algorithme de transformée de FOURIER rapide) pour obtenir les échantillons de la transformée de FOURIER 2D de $f\,(x, y)$ sur une grille circulaire.

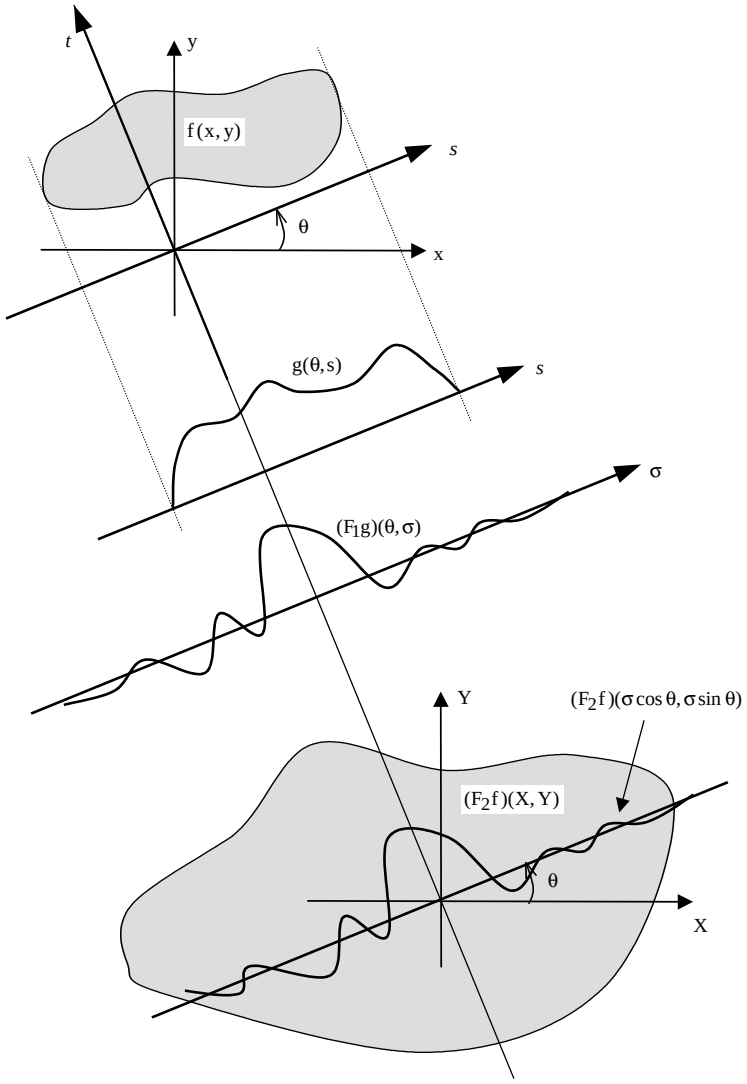


FIG. 6.10 – Illustration du théorème du profil central.

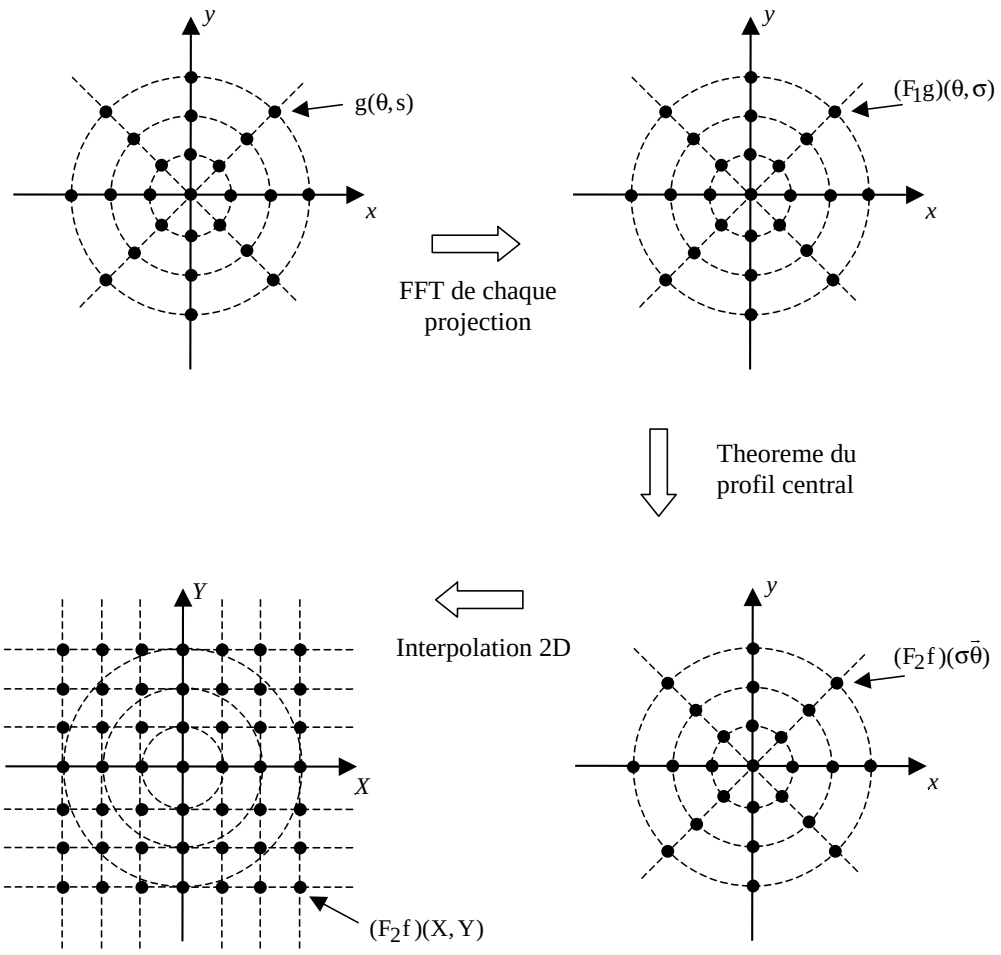


FIG. 6.11 – Méthode directe de FOURIER.

2. Interpolation 2D pour obtenir des échantillons de la transformée de FOURIER 2D de $f(x, y)$ sur une grille rectangulaire.
3. Calcul de $f(x, y)$ par inversion de FOURIER 2D rapide (FFT 2D).

Concrètement, la méthode directe de FOURIER est très difficile à implémenter de façon numériquement stable pour différentes raisons que nous n'aborderons pas dans cette introduction. Cependant, la méthode directe de FOURIER présente par rapport aux autres méthodes analytiques l'avantage d'être potentiellement plus rapide.

6.4.2 Méthode du filtrage de la rétroprojection

Dans cette méthode, la reconstruction de l'image $f(x, y)$ se fait en deux étapes : (i) rétroprojection des projections $g(\theta, s)$, ce qui donne l'image $f_B(x, y)$ (différente de $f(x, y)$ malheureusement ...) et (ii) filtrage de $f_B(x, y)$ afin d'obtenir l'image $f(x, y)$ recherchée.

L'image $f_B(x, y) = f_B(\vec{x})$ est obtenue par rétroprojection des mesures $g(\theta, s)$:

$$f_B(x, y) = \int_0^{\pi} d\theta \, g(\theta, s = x \, \cos\theta + y \, \sin\theta) \tag{6.6}$$

Géométriquement, pour θ fixé, $s = x \, \cos\theta + y \, \sin\theta$ correspond à la distance séparant l'origine de la projection orthogonale du point $\overrightarrow{x} = (x, y)$ sur la projection $g(\theta, s)$ (voir figure 6.12). La rétroprojection $f_B(x, y)$ est alors égale à la somme des contributions de chaque projection.

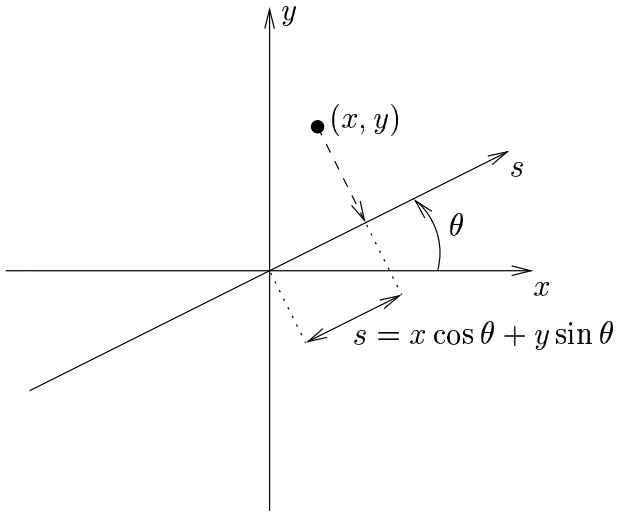


FIG. 6.12 – Illustration de la rétroprojection.

On observe que $f_B(x, y)$ est une version floue de $f(x, y)$ avec valeurs non nulles, en dehors du support de $f(x, y)$. On cherche alors un filtre 2D qui permet de retrouver $f(x, y)$ à partir de $f_B(x, y)$. La figure 6.13 montre le résultat $f_B(x, y)$ correspondant à l'image $f(x, y)$ définie à la figure 6.8.

FIG. 6.13 – Résultat de la rétroprojection pour un fantôme composé de 3 points objet.

Il existe une relation directe entre les transformées de FOURIER respectives de $f(x, y)$ et $f_B(x, y)$:

$$(\mathcal{F}_2 f)(\vec{X}) = |\vec{X}|\,(\mathcal{F}_2 f_B)(\vec{X}). \tag{6.7}$$

De là, on décrit l'algorithme du filtrage de la rétroprojection comme suit :

1. Rétroprojection des projections pour obtenir $f_B(\vec{x})$.
2. Calcul de la transformée de FOURIER de $f_B(\vec{x})$, application du filtre $|\vec{X}|$ et calcul de la transformée de FOURIER inverse du résultat.

L'étape de filtrage est très difficile à implémenter en pratique parce que $f_B(\vec{x})$ n'est pas à support compact (son support peut sortir de la grille de reconstruction). En pratique, pour obtenir des images de qualité satisfaisante en utilisant la méthode du filtrage de la rétroprojection, il faut calculer l'image $f_B(\vec{x})$ sur une grille de pixels trois à quatre fois plus grande que celle utilisée pour $f(\vec{x})$. On se retrouve alors avec un algorithme beaucoup plus lent que la méthode dite de rétroprojection des projections filtrées.

6.4.3 Méthode de rétroprojection des projections filtrées

À l'inverse de la méthode du filtrage de la rétroprojection, l'opération de filtrage est réalisée avant la rétroprojection. On utilise l'expression mathématique

$$f(x, y) = \int_0^{\pi} d\theta\, g^F(\theta, s = x\,\cos\theta + y\,\sin\theta) \tag{6.8}$$

où $g^F(\theta, s)$ sont les projections filtrées. Chaque projection $g(\theta, s)$ est filtrée indépendemment des autres grâce à la formule

$$(F_1 g^F)(\theta, \sigma) = |\sigma|\,(F_1 g)(\theta, \sigma) \tag{6.9}$$

Le filtre $|\sigma|$ est bien connu en tomographie, il porte le nom de filtre rampe. Il est représenté à la figure 6.14.

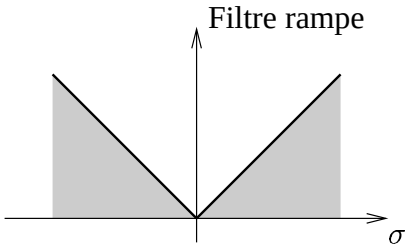


FIG. 6.14 – Filtre rampe.

De là, on décrit l'algorithme de rétroprojection des projections filtrées (*filtered backprojection* en anglais ou FBP) :

1. Transformée de FOURIER 1D de chaque projection $g(\theta, s)$, application du filtre rampe $|\sigma|$ et transformée de FOURIER 1D inverse pour obtenir la projection filtrée $g^F(\theta, s)$.
2. Rétroprojection des projections filtrées pour obtenir $f(x, y)$.

L'avantage de cette dernière méthode est tout d'abord de simplifier l'étape de filtrage, vu que celle-ci ne se fait plus qu'à une dimension. Un autre avantage est que, dès que la première projection a été acquise, il est possible d'obtenir une première approximation de l'image avant même que l'acquisition globale des mesures ne soit terminée. Dès qu'une nouvelle projection est mesurée, elle est filtrée et rétro-projetée ; et ainsi de suite jusque la dernière projection. L'algorithme de rétroprojection des projections filtrées est couramment utilisé dans les scanners actuels.

La figure 6.15 montre la reconstruction obtenue pour le fantôme de SHEPP-LOGAN sur une grille de 256×256 pixels avec un sinogramme de taille 400×256.

FIG. 6.15 – Reconstruction du fantôme de SHEPP-LOGAN.

Index

A
arbre, 20

B
bande passante, 36
Butterworth, 40

C
chain code, 22
chemin géodésique, 56
CMY, 10
composante
    DC, 32
connexité, 18
couleur, 3

D
déconvolution, 35
DC, 32
distance
    géodésique, 56

E
égalisation de l'histogramme, 47
éventail, 63

F
facteur d'atténuation, 61
fantôme, 67
filtre
    idéal, 35
    linéaire, 35
    passe-bande, 37
    passe-bas, 36
    passe-haut, 37
fonction
    Rectangle, 32
Fourier
    transformée discrète, 29

G
géodésie, 56
géométrie
    en éventail, 63
    parallèle, 63
gamma, 11

H
histogramme, 47
    égalisation, 47
hue, 10

I
idempotence, 35
intensité, 3

L
lacet, 22
ligne de partage des eaux, 52
LSB, 24
luminance, 3, 8

M
marquage, 54, 59
marqueur, 55
module, 30
MSB, 24

P
pattern matching, 28
phase, 30
pixel, 16
projection, 65
pyramide, 22

Q
quadtree, 22

R
rétroprojection, 70, 73
Radon, 65
reconstruction, 65
rehaussement, 43, 45
restauration, 43
ROI, 27

S
saturation, 10
segmentation, 51
seuillage, 52
sinogramme, 65, 67
spectre, 30
squelette
    par zone d'influence, 57

T
teinte, 10
théorème du profil central, 70
tomographie, 61
trame, 16, 19
transformée
    de Fourier, 29
transformation
    unitaire, 29
transparence, 16

Y
YCbCr, 10
YIQ, 10
YUV, 10

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. 13

`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). 12

`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. 37

`DFT` *Discrete* FOURIER *Transform.* La transformée de FOURIER discrète s'obtient par échantillonnage de la transformée continue. 32

`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. 20

`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). 35

`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. 12

`JPEG` *Joint Photographic Experts Group.* Format d'images présentant de la compression avec perte. Efficace pour les images du type photographique. 38

`LSB` *Least Significant Bit.* Terme désignant le bit de poids faible dans un byte. 25

`MSB` *Most Significant Bit.* Terme désignant le bit de poids fort dans un byte. 25

`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. 12, 20

`OCR` *Optical Character Recognition.* Terme qui désigne les techniques optiques de reconnaissance de caractères. 251

`OCV` *Optical Character Verification.* Techniques d'inspection de marquage utilisée pour contrôler l'efficacité de marquage en imagerie. 250

`PAL` *Phase Alternate Line.* Format de télévision analogique utilisé dans la majorité des pays européens (hormis la France). 1, 12, 20

`palette` Table indexant les couleurs des pixels d'une image. Ce principe est utilisé pour le format GIF. 13

`pixel` *Picture Element.* Il s'agit d'un point d'une image échantillonnée. On parle aussi de PEL. 17

`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. 28

`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. 17, 19

Bibliographie

[1] P. Buser et M. Imbert. *Neurophysiologie fonctionnelle IV*. Hermann, 1987.

[2] R. Hunt. *The reproduction of colour*. Fountain Press, fifth edition, 1995.

[3] S. Katzenbeisser and F. Petitcolas. *Information hiding techniques for steganography and digital watermarking*. Artech House, 2000.

[4] M. Kunt, G. Grunland, et M. Kocher. *Traitement de l'information : traitement numérique des images*, Volume 2. Presses polytechniques et universitaires romandes, 1993.

[5] G. Laurent. *Cours de télévision, tome 1 : réception, norme, gestion, traitement vidéo*. Dunod, 1996.

[6] F. Meyer and S. Beucher. Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1) :21–46, September 1990.

[7] 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.

[8] M. Sonka, V. Hlavac, and R. Boyle. *Image processing, analysis, and machine vision*. Brooks/Cole Publishing Company, second edition, 1999.