HAL
open science

Ondelettes pour la détection de caractéristiques en traitement d'images. Application à la détection de région d'intérêt.

Christophe Damerval

▶ To cite this version:

Christophe Damerval. Ondelettes pour la détection de caractéristiques en traitement d'images. Application à la détection de région d'intérêt.. Mathématiques [math]. Université Joseph-Fourier - Grenoble I, 2008. Français. NNT : . tel-00287008

HAL Id: tel-00287008
https://theses.hal.science/tel-00287008v1

Submitted on 10 Jun 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE JOSEPH FOURIER (Grenoble I)
Ecole Doctorale "Mathématiques, Sciences et
Technologies de l'Information, Informatique"

THESE

pour l'obtention du grade de
DOCTEUR DE L'UNIVERSITÉ JOSEPH FOURIER
Spécialité : Mathématiques Appliquées
préparée au Laboratoire Jean Kuntzmann (LJK) de Grenoble
présentée et soutenue publiquement par

CHRISTOPHE DAMERVAL

le 7 mai 2008

Ondelettes pour la détection de caractéristiques en traitement d'images

Applications à la détection de régions d'intérêt

Composition du Jury

M. ANESTIS ANTONIADIS	PROFESSEUR, UNIVERSITÉ JOSEPH FOURIER	PRÉSIDENT
M. ANDRÉS ALMANSA	CHERCHEUR CNRS (HDR), TELECOM PARISTECH	EXAMINATEUR
M. FRANCESC ARANDIGA	PROFESSEUR, UNIVERSITÉ DE VALENCIA (ESPAGNE)	RAPPORTEUR
M. PATRICK FLANDRIN	DIRECTEUR DE RECHERCHE, ENS LYON	RAPPORTEUR
M. SYLVAIN MEIGNEN	MAÎTRE DE CONFÉRENCES, INP GRENOBLE	CO-DIRECTEUR
MME CORDELIA SCHMID	DIRECTEUR DE RECHERCHE, INRIA RHÔNE-ALPES	DIRECTEUR

Table des matières

Introduction **v**
Notations ix

1 Décompositions en ondelettes (1D) **1**
1.1 Généralités sur les ondelettes monodimensionnelles 1
1.2 Transformées en ondelettes 3
1.2.1 Analyse multirésolution (AMR) 3
1.2.2 Transformée en ondelettes dyadique 6
1.2.3 Transformée en ondelettes continue 7
1.2.4 Choix de la transformée et de l'ondelette 9
1.3 Propriétés des coefficients d'ondelettes 11
1.3.1 EDP vérifiée par les coefficients 11
1.3.2 Evolution des coefficients en fonction de l'échelle 14
1.3.3 La TOC vue comme un opérateur différentiel 15
1.3.4 Coefficients d'ondelette et régularité d'un signal 16
1.4 Modules maxima de la transformée en ondelette 18
1.4.1 Notion de module maximum 18
1.4.2 Lien avec les singularités présentes dans le signal 18
1.4.3 Reconstruction d'un signal à l'aide des modules maxima 18

2 Extrema et Maxima Lines d'ondelettes (1D) **21**
2.1 Lignes d'extrema et Maxima Lines 21
2.1.1 Extrema de la TOC et lignes d'extrema 21
2.1.2 Modules Maxima et Maxima Lines 22
2.2 Etude des extrema locaux de la TOC 23
2.2.1 Cas général 23
2.2.2 Illustration dans le cas d'un signal type "rampe" 24
2.3 Etude numérique des Maxima Lines 25
2.3.1 Construction des ML (en 1D) 25
2.3.2 Exemple de disparition d'une ligne de maxima 27
2.3.3 Interaction entre singularités d'importances différentes 28
2.4 Applications 29
2.4.1 Détection de singularités et calcul des régularités associées 29
2.4.2 Mise en évidence de certaines structures 29
2.4.3 Etude dans un cadre bruité 32

3 Décompositions 2D et lignes de maxima **37**
3.1 Introduction . . . 37
3.2 Décompositions multiéchelles . . . 38
3.2.1 Ondelettes bidimensionnelles . . . 38
3.2.2 Coefficients d'ondelettes associés à une image . . . 39
3.2.3 Transformées adaptées à la géométrie . . . 41
3.3 Modules maxima et Maxima Lines . . . 44
3.3.1 Définitions . . . 44
3.3.2 Equation de la chaleur et principe du maximum . . . 44
3.3.3 Existence de Maxima Lines pour l'ondelette ΔG . . . 45
3.4 Maxima au sens de Canny . . . 46
3.4.1 Définitions . . . 46
3.4.2 Formulation multiéchelle utilisant une ondelette gradient . . . 47
3.4.3 Notion de ligne de Canny . . . 48

4 Théorie du Scale-Space **49**
4.1 Eléments de la théorie du Scale-Space . . . 49
4.1.1 Représentation espace-échelle . . . 50
4.1.2 Notion de dérivée γ-normalisée . . . 51
4.1.3 Application de l'invariance à l'échelle : spectre de puissance . . . 53
4.1.4 Normalisation du gradient et de la Hessienne . . . 53
4.2 Approche classique pour la détection de caractéristiques locales . . . 54
4.2.1 Opérateurs en Vision par Ordinateur . . . 54
4.2.2 Principe de sélection automatique d'échelle . . . 57
4.3 Prise en compte de déformations géométriques . . . 58
4.3.1 Transformations considérées . . . 58
4.3.2 Invariance par transformation affine du Scale-Space . . . 58
4.4 Estimation de déformations géométriques . . . 59
4.4.1 Matrice des moments d'ordre deux (*Second Moment Matrix*) . . . 59
4.4.2 Applications . . . 61
4.5 Généralisations par des fonctions échelles 2D . . . 62
4.5.1 Nouvelle représentation espace-échelle . . . 62
4.5.2 MM2 définie par une fonction échelle . . . 63
4.6 Liens entre Scale-Space et ondelettes . . . 64
4.6.1 Cadre 1D . . . 64
4.6.2 Opérateur Laplacien et Maxima Lines en 2D . . . 66
4.6.3 Applications . . . 67
4.7 Conclusion . . . 67

5 Construction des lignes de maxima **69**
5.1 Motivation . . . 69
5.2 Construction numérique des lignes de maxima . . . 70
5.2.1 Construction des Maxima Lines (ML) . . . 71
5.2.2 Construction des lignes de Canny (LC) . . . 73
5.3 Complexité algorithmique (ML et LC) . . . 76

5.4 Synthèse sur les lignes de maxima en 2D . . . 77
5.4.1 Quelques propriétés des Maxima Lines . . . 78
5.4.2 Quelques propriétés des lignes de Canny . . . 79
5.4.3 Maxima Lines et lignes de Canny . . . 80
5.5 Synthèse : décompositions, maxima et singularités . . . 83

6 Détection de structures monodimensionnelles en 2D 85
6.1 Singularités dans une image . . . 85
6.2 Caractérisation de structures monodimensionnelles . . . 86
6.2.1 Calcul des régularités . . . 86
6.2.2 Comparaison des méthodes d'estimation de la régularité . . . 87
6.3 Robustesse de la régularité Lipschitzienne . . . 91
6.3.1 Influence d'une déformation sur la régularité . . . 91
6.3.2 Influence du bruit . . . 91
6.3.3 Influence d'une rotation . . . 92
6.3.4 Influence d'une déformation affine . . . 93
6.4 Questions d'invariance et de robustesse . . . 96
6.4.1 Régularité Lipschitzienne . . . 96
6.4.2 Maxima Lines et lignes de Canny . . . 96
6.5 Conclusion . . . 98

7 Détection de structures bidimensionnelles en 2D 99
7.1 Détection de points d'intérêt en vision par ordinateur . . . 100
7.1.1 Notion de région d'intérêt . . . 100
7.1.2 Régions associées à des maxima dans le *Scale-Space* . . . 101
7.1.3 Application des détecteurs de régions d'intérêt . . . 102
7.1.4 Principe de détection de points d'intérêt dans le Scale-Space . . . 103
7.1.5 Exemples de détecteurs utilisés . . . 104
7.2 Approche par ML . . . 110
7.2.1 Mécanisme de sélection le long de ML . . . 110
7.2.2 Comparaison avec la sélection classique . . . 110
7.3 Détection de point d'intérêt par ML . . . 112
7.3.1 Quelques notions associées aux ML . . . 112
7.3.2 Algorithme de détection par ML . . . 113
7.3.3 Sélection de Maxima Lines pertinentes . . . 115
7.3.4 Détermination des points d'intérêts – Jonction de ML . . . 117
7.3.5 Calcul de l'échelle caractéristique . . . 117
7.4 Validation de l'approche par ML . . . 118
7.4.1 Sensibilité au bruit de l'échelle caractéristique . . . 118
7.4.2 Exemples pour une image naturelle . . . 119
7.4.3 Comparaison avec l'approche classique . . . 119
7.4.4 Estimation géométrique . . . 123
7.5 Conclusions et perspectives . . . 127

8 Conclusion générale **129**
8.1 Contributions principales 129
8.2 Perspectives 131

A Tests de répétabilité **133**
A.1 Performance d'un détecteur 133
A.2 Evaluation d'un détecteur – Test de répétabilité 135
A.3 Application au détecteur par ML 136

B Estimation des régularités par curvelets **145**
B.1 Contours et curvelets 145
B.2 Analyse de singularités par curvelets 147

C Quelques aspects de l'EMD **151**
C.1 Présentation de la méthode EMD en dimension 1 151
C.1.1 Principe de la décomposition 151
C.1.2 Algorithme EMD 153
C.1.3 Degrés de liberté dans la méthode 154
C.2 Problématique de l'EMD en 2D 155
C.3 Calcul des enveloppes 156
C.3.1 Utilisation de fonctions à base radiale 156
C.3.2 Interpolation basée sur une triangulation 157
C.3.3 Conditions aux bords 158
C.3.4 Calculs d'enveloppes 159
C.4 Raffinement des enveloppes – *Sifting Process* (SP) 161
C.5 Analyse de signaux bidimensionnels par EMD 163
C.6 Conclusions et perspectives 163

Bibliographie **167**

Introduction

Cette thèse en traitement d'images aborde le problème de la mise en évidence de certaines structures remarquables, comme des objets que nous percevons visuellement. Nous nous intéressons à la détection de telles structures, ainsi qu'à l'extraction de grandeurs caractéristiques de celles-ci. Nous montrons que ces tâches peuvent être réalisées par des méthodes d'ondelettes. En particulier, nous verrons comment certaines lignes de maxima permettent de mettre en évidence des régions d'intérêt.

Mots clés : ondelettes, *Scale-Space*, représentations multiéchelles, lignes de maxima, invariance ou robustesse à des transformations de l'image, détection de régions d'intérêt, calcul de grandeurs caractéristiques.

Approches issues de la vision par ordinateur

La détection de régions d'intérêt constitue un problème central en vision par ordinateur, de par ses diverses applications, comme la reconnaissance d'objets, l'appariement d'images, le suivi de mouvement ou le rehaussement de certains éléments particuliers [2, 27, 66, 67, 107, 110, 111]. Cette étape de détection est essentielle, dans la mesure où elle constitue une première étape avant d'autres plus complexes, comme la détermination de déformations géométriques ou l'extraction de caractéristiques invariantes [68, 77, 79, 99]. Sa spécificité réside dans le caractère local des régions extraites : il s'agit d'identifier les structures bidimensionnelles les plus remarquables dans l'image. Ceci contraste avec des approches qui recherchent des structures monodimensionnelles ou des singularités ponctuelles. Ainsi, il est crucial de déterminer les échelles auxquelles apparaissent ces objets.

Dans cette optique, des approches multiéchelles ont été introduites, et plus particulièrement celles impliquant une certaine redondance, de manière à ce que les éléments les plus remarquables soient présents à différentes échelles. Parmi les différentes théories permettant de formaliser de telles approches, la théorie du *Scale-Space*, introduite par Witkin [125] puis approfondie par Lindeberg [75], définit des méthodes de détection automatique, qui prennent en compte certaines transformations de l'image, notamment le changement d'échelle et les déformations affines. Ceci conduit alors au concept d'invariance à une certaine transformation (d'où le terme d'invariance à l'échelle par exemple) : soit que les régions extraites soient invariantes à celle-ci, soit qu'elles la prennent en compte (l'invariance étant atteinte par la suite).

Dans ces méthodes classiques de détection, l'image est d'abord lissée, typiquement par un noyau Gaussien, à différentes échelles, avant d'appliquer un opérateur adapté [7]. Une fois cette opération effectuée, la sélection de certains maxima permet de mettre en évidence différentes structures plus ou moins complexes présentes dans l'image. Naturellement, la diversité et la complexité de celles-ci rendent d'autant plus difficile la formulation d'un détecteur efficace de régions d'intérêt.

De nombreuses méthodes ont été proposées afin d'extraire les régions les plus pertinentes d'une image (algorithmes de détection) puis de calculer des grandeurs caractérisant le contenu de ces régions (calcul de descripteurs) [53, 83, 99]. Lorsqu'il s'agit d'extraire de telles caractéristiques d'une image, il importe qu'elles soient invariantes (ou au moins robustes) à certaines transformations de l'image, ces dernières pouvant être géométriques (rotations, déformations affines) ou autres (changement de luminosité ou de contraste, image bruitée). Une application majeure de la théorie du *Scale-Space* a été de rendre ces détecteurs invariants à certaines transformations (comme les déformations affines), et un important travail d'amélioration de ces algorithmes a permis d'obtenir des performances accrues. Notons néanmoins que la nature des objets détectés n'est pas toujours claire, et en particulier il est difficile de caractériser les régions détectées.

Ondelettes en traitement d'images

Les ondelettes sont des outils en traitement du signal et de l'image, intégrant simultanément des paramètres de localisation et d'échelle [34, 85, 93]. Le développement d'algorithmes rapides adaptés aux signaux digitaux a permis de nombreuses applications en traitement d'images, où les ondelettes peuvent être utilisées soit en tant que systèmes de représentation (compression d'images dans une base d'ondelettes), soit comme outils d'analyse (détection de contours multiéchelles) [42, 86, 87]. Dans une optique de détection de structures remarquables dans une image, c'est ce dernier aspect qui attire notre attention.

Un aspect intéressant des ondelettes est qu'elles permettent de prendre en compte à la fois la notion d'échelle et celle de singularité. En effet, les transformées en ondelettes reviennent à convoluer l'image avec des fonctions oscillantes, tout en intégrant un aspect multiéchelles (ou multirésolution). Par ailleurs, à partir de celles-ci, il est possible de définir certains maxima (associés à une échelle particulière) qui correspondent à des singularités présentes dans l'image. De surcroît, il est possible de définir des lignes de maxima, qui sont des chaînes de maxima traversant les échelles à travers les échelles [88].

Ces lignes de apparaissent comme intéressantes sur plusieurs points. D'abord, elles peuvent être associées à certaines singularités, et permettent de suivre l'interaction entre différentes singularités à mesure que l'échelle croît ; en particulier, il apparaît que certaines lignes fusionnent en certains points du *Scale-Space* (notion de jonction de lignes). Ensuite, leurs propriétés vis à vis des certaines transformations de l'image les rend pertinentes pour la détection robuste de certaines structures : monodimensionnelles, comme des contours, ou bidimensionnelles,

ce qui correspond à des objets plus complexes [30, 31]. Enfin, elles permettent l'extraction de grandeurs caractéristiques, telles que la régularité Lipschitzienne associée à une singularité ponctuelle, ou une échelle caractéristique associée à une structure 2D. Ces grandeurs sont dites caractéristiques dans la mesure où elles identifient le type d'une singularité ou la taille d'un objet présent dans l'image.

Ainsi, il apparaît comme intéressant de tisser des liens entre les ondelettes et la théorie du *Scale-Space*, dans l'optique de formuler de méthodes de détection originales. Idéalement, ces méthodes seront à la fois robustes à différentes transformations de l'image (déformations géométriques, ajout de bruit) et capables de caractériser les objets détectés.

Approche proposée

La contribution essentielle de cette thèse est l'utilisation de lignes de maxima afin de détecter des objets présents dans l'image. Du point de vue théorique, nous montrons des liens entre le *Scale-Space* et les ondelettes. Ceci justifie alors l'utilisation des ondelettes, de manière similaire aux méthodes de détection de régions d'intérêt. L'apport de cette approche originale est qu'elle permet de caractériser les objets mis en évidence. Ceci est d'autant plus pertinent que dans les nombreuses applications de ces détecteurs issus de la vision par ordinateur, il importe de connaître le type d'objet extrait (points de contour par exemple), ou dans le cas plus complexe d'un objet 2D ayant une certaine cohérence, de pouvoir identifier des singularités associées.

Nous montrons que ces lignes de maxima permettent de détecter certains éléments caractéristiques de la partie géométrique de l'image, comme des contours ou des structures bidimensionnelles. Ces méthodes de détection peuvent alors être comparées avec les approches issues de la vision par ordinateur précédemment évoquées. A cet égard, notons que ces dernières comportent souvent deux étapes, de détection puis de description de régions d'intérêt [83, 99, 100]. Ici, nous nous focalisons sur la détection, et les grandeurs extraites permettent avant tout d'identifier le type d'objet détecté (par exemple, un point de contour pourra être identifié comme appartenant à une frontière entre deux régions relativement homogènes).

Plan de la thèse

Nous nous plaçons d'abord dans un cadre 1D, où nous présentons des décompositions en ondelettes et expliquons comment ces approches multiéchelles sont utiles en traitement du signal (chap. 1). Puis, nous insistons sur la détection et le suivi à travers les échelles de singularités présentes dans un signal (chap. 2) ; nous montrons notamment comment ces singularités évoluent à mesure que l'échelle augmente, grâce à la notion de *Maxima Line* (ML). Dans ce cadre, nous verrons que ces ML permettent plusieurs choses : détecter des singularités ponctuelles, les caractériser (par la régularité Lipschitzienne), et également mettre en évidence certaines structures non-ponctuelles (grâce à la jonction de ML).

Ensuite, dans le cadre bidimensionnel du traitement d'images, nous présentons des approches multiéchelles utilisant les ondelettes (chap. 3), en particulier nous définissons les principaux outils de cette thèse : les lignes de Canny (LC) et les Maxima Lines (ML). Afin de montrer la pertinence de ces lignes pour la détection d'objets, nous rappelons comment la théorie du *Scale-Space*, étroitement liée à des problématiques issues de la vision par ordinateur, permet de sélectionner des points et des échelles remarquables associés à une image (chap. 4). Ces points et ces échelles sont définis comme les maxima dans l'espace-échelle d'un certain opérateur. L'intérêt de ceux-ci est qu'ils correspondent à certains objets présents dans l'image, et également qu'ils possèdent des propriétés intéressantes vis à vis de transformations géométriques. Ceci conduit notamment à la notion de caractéristique invariante (invariance au changement d'échelle par exemple). Nous montrons alors des liens entre les ondelettes et le *Scale-Space*, du point de vue théorique.

Du point de vue pratique, nous détaillons la construction algorithmique des lignes de maxima (chap. 5), et effectuons une synthèse sur leurs différentes propriétés. Nous abordons alors la détection de structures remarquables dans l'image, en les distinguant suivant leur caractère monodimensionnel ou bidimensionnel. Au chapitre 6, nous effectuons d'abord une synthèse sur les différentes caractéristiques extractibles d'une image, où nous comparons les ML et les LC. Nous détaillons ensuite comment des structures 1D peuvent être détectées et caractérisées, notamment grâce aux LC et à la régularité Lipschitzienne. Au chapitre 7, nous décrivons comment il est possible, grâce aux ML, d'identifier des structures 2D. Nous présentons un nouveau mécanisme de sélection d'échelle, et un détecteur de points d'intérêt fondé sur les ML. Un objet ainsi détecté peut alors être associé à un certain nombre de ML, qui le caractérisent. En outre, ces ML permettent de donner une estimation de la forme de cet objet.

Enfin, nous discutons de l'intérêt de l'approche fondée sur les ML, ainsi que de ses limites. Nous décrirons également les perspectives de ce travail de thèse.

Travaux annexes

Nous présentons en annexe des résultats complémentaires à cette thèse. En annexe A, nous rappelons d'abord la notion de répétabilité, qui permet de mesurer la performance d'un détecteur de régions d'intérêt [101] ; nous présentons alors les résultats de tests de répétabilité sur le détecteur ML (vu au chap. 7). Par ailleurs, en annexe B, nous expliquons comment la régularité Lipschitzienne (étudiée au chap. 6) peut être estimée grâce aux curvelets.

En marge de cette thèse (annexe C), nous présentons quelques aspects d'une méthode adaptative de décomposition de signaux, appelée EMD (Empirical Mode Decomposition, ou décomposition modale empirique). En particulier, nous détaillons un algorithme rapide que nous avons proposé [33].

Notations et abréviations

L^p	Espace des fonctions $f : \mathbb{R} \longrightarrow \mathbb{C}$ telles que $\|\|f\|\|_p = \left(\int_{\mathbb{R}} \|f(t)\|^p dt\right)^{1/p} < +\infty$
$\|\|f\|\|_p$	Norme L^p d'une fonction f
$< f, g >$	Produit scalaire de deux fonctions dans L^2
$\widehat{f}$	Transformée de Fourier de $f \in L^1$
FFT	Transformée de Fourier rapide (*Fast Fourier Transform*)
$E \bigoplus F$	Somme directe de deux espaces vectoriels
$E \overset{\perp}{\bigoplus} F$	Somme directe de deux espaces vectoriels orthogonaux
$E \bigotimes F$	Produit tensoriel de deux espaces vectoriels
∂_x	Dérivée partielle par rapport à la variable x
∂_ξ	Dérivée γ-normalisée
∇	Opérateur gradient
$\mathcal{H}$	Matrice Hessienne
Δ	Opérateur Laplacien
$\|x\|_2$	Norme euclidienne d'un vecteur $x \in \mathbb{R}^n$
$M_2(\mathbb{R})$	Matrices réelles 2×2
GL_2	Matrices de $M_2(\mathbb{R})$ inversibles
SP_2	Matrices de $M_2(\mathbb{R})$ symétriques positives
SDP_2	Matrices de $M_2(\mathbb{R})$ symétriques définies positives
det	Déterminant d'une matrice
trace	Trace d'une matrice
TOC	Transformée en ondelette continue
MM	Modules maxima
ML	Maxima Lines
MC	Maxima au sens de Canny
LC	Lignes de Canny
α	Régularité Lipschitzienne
s^*	Echelle caractéristique

Définitions

Transformée de Fourier

Définition 1. *Pour $f : \mathbb{R} \to \mathbb{R}, f \in L^1$, sa transformée de Fourier (TF) est définie comme*

$$\mathcal{F}f(\omega) = \widehat{f}(\omega) = \int_{\mathbb{R}} f(t) \exp(-2i\pi\omega t)dt \tag{1}$$

et sa transformée conjuguée est donnée par

$$\overline{\mathcal{F}}f(\omega) = \int_{\mathbb{R}} f(t) \exp(2i\pi\omega t)dt \tag{2}$$

Définition 2. *Pour $f : \mathbb{R}^2 \to \mathbb{R}, f \in L^1$, sa transformée de Fourier est définie comme*

$$\mathcal{F}f(\omega_1, \omega_2) = \widehat{f}(\omega_1, \omega_2) = \int_{\mathbb{R}^2} f(x, y) \exp(-2i\pi(\omega_1 x + \omega_2 y))dxdy \tag{3}$$

et sa transformée conjuguée est donnée par

$$\overline{\mathcal{F}}f(\omega_1, \omega_2) = \int_{\mathbb{R}^2} f(x, y) \exp(2i\pi(\omega_1 x + \omega_2 y))dxdy \tag{4}$$

Bases et Frames

Nous considérons ici une famille de fonctions $\{\phi_n\}_{n\in\mathbb{N}}$ appartenant à un espace de Hilbert $\mathcal{H}$.

Définition 3. *(Frame) La famille $\{\phi_n\}_{n\in\mathbb{N}}$ est un frame de $\mathcal{H}$ s'il existe deux constantes A et B $(A, B > 0)$ telles que :*

$$\forall f \in \mathcal{H}, \quad A||f||^2 \leq \sum_{n\in\mathbb{N}} | < f, \phi_n > |^2 \leq B||f||^2 \tag{5}$$

Définition 4. *(Base de Riesz) La famille $\{\phi_n\}_{n\in\mathbb{N}}$ est une base de Riesz si elle est un frame et si pour tout $f \in \mathcal{H}$,*

$$\exists!(c_n)_{n\in\mathbb{N}} \in \mathbb{C}^{\mathbb{N}}, \lim_{n\to+\infty} ||f - \sum_{k=0}^{n} c_k\phi_k|| = 0 \tag{6}$$

où $\forall t \in \mathbb{R}, \phi_k(t) = \phi(t - k)$.

Note : un frame est dit étroit si $A = B$ dans la définition 3 ; si la famille $\{\phi_n\}_{n\in\mathbb{N}}$ est un frame composée de vecteurs linéairement indépendants, c'est une base de Riesz ; par ailleurs, toute base Hilbertienne est une base de Riesz. Dans le cadre d'espaces fonctionnels, un frame donne une représentation complète d'une fonction, mais à la différence d'une base, cette représentation est redondante.

Chapitre 1

Décompositions en ondelettes (1D)

1.1 Généralités sur les ondelettes monodimensionnelles

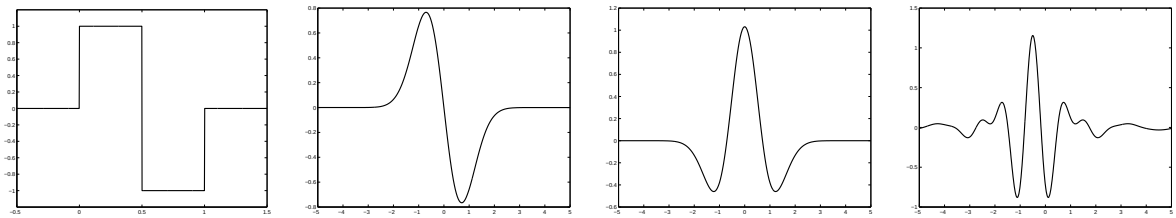
Nous présentons ici des généralités sur la théorie des ondelettes dans un cadre monodimensionnel, notamment dans l'optique d'analyse de signaux et de détection de singularités. Nous insistons particulièrement sur les propriétés des coefficients d'ondelettes et des modules maxima de la transformée en ondelette continue [86–88]. L'intérêt de ceux-ci est que les coefficients d'ondelettes donnent une représentation du signal, tandis que les modules maxima permettent d'identifier des singularités présentes dans celui-ci.

Notion d'ondelette

Les ondelettes sont des fonctions à partir desquelles il est possible d'analyser un signal (extraire des éléments remarquables, caractéristiques) et de représenter un signal (décomposition dans une base).

Définition 5. *Une ondelette est une fonction $\psi : \mathbb{R} \longrightarrow \mathbb{R}$, $\psi \in L^2$, vérifiant la condition d'admissibilité :*

$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(\omega)|^2}{\omega} d\omega < \infty \tag{1.1}$$



(A) (B) (C) (D)

Fig. 1.1 – (A) Ondelette de Haar ; (B) Ondelette dérivée première de Gaussienne ; (C) Ondelette *Sombrero*, égale à $-\Delta G$ (dérivée seconde de Gaussienne) ; (D) Ondelette de Meyer.

L'analyse en ondelettes permet une localisation à la fois spatiale (ou temporelle) et fréquentielle. Alors que l'analyse de Fourier est adaptée à l'étude de signaux stationnaires, l'analyse par ondelettes est adaptée aux signaux non-stationnaires. Citons quelques exemples d'ondelettes :

– ondelette de Haar, utilisée en compression (Fig. 1.1 (A))

$$\psi(t) = \begin{cases} 1 & \text{si } t \in [0, 1/2] \\ -1 & \text{si } t \in]1/2, 1] \\ 0 & \text{sinon} \end{cases}$$

– dérivée première de Gaussienne, utilisée pour la détection de contours (Fig. 1.1 (B))

$$\psi(t) = -tG(t) \text{ où } G(t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2}) \tag{1.2}$$

– dérivée seconde de Gaussienne, dont nous verrons plus loin l'utilisation pour la détection d'éléments caractéristiques :

$$\psi(t) = (-1 + t^2)G(t) \tag{1.3}$$

Note : nous appelons cette ondelette ΔG, et l'ondelette *Sombrero* est définie comme $-\psi$ (Fig. 1.1 (C)).

– ondelette de Meyer, utile pour la localisation fréquentielle (Fig. 1.1 (D)).

Les propriétés de ces ondelettes sont détaillées dans [35, 87, 93–96].

Notion de filtre

Le filtrage est une opération fondamentale en traitement du signal, qui, à partir d'un signal à analyser f, fournit un signal plus simple à interpréter Lf, certaines fréquences étant généralement privilégiées. Dans le domaine discret, cette opération revient à effectuer une convolution entre le signal à analyser f (donné par $(f[n])_{n\in\mathbb{Z}}$) et un signal connu h (donné par $(h[n])_{n\in\mathbb{Z}}$), appelé filtre

$$Lf[n] = \sum_{p\in\mathbb{Z}} f[p]h[n-p] = (f * h)[n] \tag{1.4}$$

L'utilisation de filtres permet notamment d'isoler certaines bandes de fréquences (Fig. 1.2). Etant donné un filtre $h \in l^1(\mathbb{Z})$ ($\sum_{k\in\mathbb{Z}} |h(k)| < +\infty$), sa fonction de transfert est définie comme : $\widehat{h}(\omega) = \sum_{p\in\mathbb{Z}} h[p]e^{-i\omega p}$. Cette fonction de transfert caractérise ce filtre, et donne également les fréquences sélectionnées, vu que sous certaines hypothèses [50], nous avons

$$\widehat{Lf}(\omega) = \widehat{f}(\omega) \cdot \widehat{h}(\omega) \tag{1.5}$$

Application : certaines ondelettes étant définies par un filtre [35], dans un certain cadre, une transformée en ondelette peut s'exprimer comme un banc de filtres ; un algorithme rapide permet alors de décomposer un signal discret dans une base d'ondelettes [87].

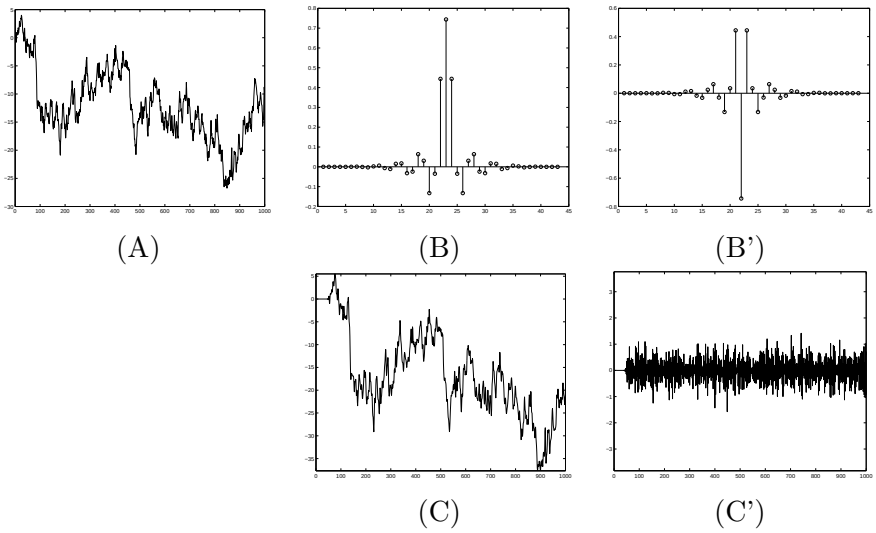


FIG. 1.2 – (A) Signal à analyser (issu de la simulation d'un mouvement brownien) ; (B) Filtre passe-bas h ; (C) Signal filtré correspondant ; (B') Filtre passe-haut g ; (C') Signal filtré correspondant. Remarque : les deux filtres h et g sont associés à l'ondelette de Meyer.

1.2 Transformées en ondelettes

Nous définissons maintenant des transformées en ondelettes associées à un signal : d'abord la transformée en ondelette dyadique (TOD), qui s'avère intéressante dans le cadre d'une analyse multirésolution de L^2 (AMR) ; puis la transformée en ondelette continue (TOC), qui sera l'outil privilégié de cette thèse.

1.2.1 Analyse multirésolution (AMR)

Définition 6. *(AMR [94]) Une analyse multirésolution de $L^2(\mathbb{R})$ est une suite de sous-espaces fermés $(V_j)_{j\in\mathbb{Z}}$ (espaces d'approximation) vérifiant les propriétés suivantes :*

1. $\forall j \in \mathbb{Z},\ V_{j+1} \subset V_j$
2. $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$
3. $\bigcup_{j\in\mathbb{Z}} V_j$ *dense dans* $L^2(\mathbb{R})$
4. $\forall (j,k) \in \mathbb{Z}^2,\quad f \in V_j \Leftrightarrow f(\cdot - 2^j k) \in V_j$
5. $f \in V_j \Leftrightarrow f(\frac{\cdot}{2}) \in V_{j+1}$
6. *Il existe $\varphi \in V_0$ telle que la famille $\{\varphi(\cdot - k), k \in \mathbb{Z}\}$ forme une base de Riesz de V_0 (cette fonction φ est appelée fonction d'échelle de l'AMR).*

L'intérêt d'une AMR de L^2 est qu'elle permet une représentation multiéchelles, le passage d'une échelle à une autre étant donné par des relations simples [94].

Dans le cadre de l'analyse d'un signal f par ondelettes, une certaine fonction échelle conduit à différentes approximations tandis que les ondelettes donnent les détails entre deux approximations successives. Les ondelettes sont alors définies comme les détails intermédiaires entre deux espaces d'approximation. Pour tout $j \in \mathbb{Z}$, vu que $V_j \subset V_{j-1}$, il existe un espace supplémentaire de V_j dans V_{j-1}, noté W_j (celui-ci n'étant pas *a priori* unique). De plus, comme les espaces V_j sont emboîtés, nous pouvons écrire :

$$\begin{aligned} V_{j-1} &= V_j \bigoplus W_j \\ V_J &= V_L \bigoplus_{j=J+1}^{L} W_j \quad \text{pour } L > J \\ L^2 &= V_0 \bigoplus_{j\leq 0} W_j = \bigoplus_{j\in\mathbb{Z}} W_j \end{aligned}$$

Cadre orthogonal

Pour une AMR donnée, les espaces V_j sont définis de manière unique, mais il peut exister un degré de liberté dans les espaces W_j, puisque l'équation $V_{j-1} = V_j \bigoplus W_j$ est une somme directe, non nécessairement orthogonale. Maintenant, si nous nous plaçons dans un cadre orthogonal, l'équation $V_{j-1} = V_j \overset{\perp}{\bigoplus} W_j$ définit les espaces d'ondelettes W_j de manière unique. Dans ce cadre, étant donnée une fonction échelle φ, il existe une ondelette ψ telle que

$$W_0 = Vect\,(\psi(\cdot - k), k \in \mathbb{Z}) \tag{1.6}$$

En notant

$$\begin{cases} \psi_{jk} &= 2^{-j/2}\psi(2^{-j}\cdot - k) \\ \varphi_k &= \varphi(\cdot - k) \\ \varphi_{Jk} &= 2^{-J/2}\varphi(2^{-J}\cdot - k) \end{cases}$$

alors des fonctions $f \in L^2$ et $f_J \in V_J$ s'écrivent de manière unique comme

$$f = \sum_{k\in\mathbb{Z}} < f, \varphi_k > \varphi_k + \sum_{j\leq 0}\sum_{k\in\mathbb{Z}} < f, \psi_{jk} > \psi_{jk} \tag{1.7}$$

$$f_J = \sum_{k=-\infty}^{+\infty} < f, \varphi_{Jk} > \varphi_{Jk} \tag{1.8}$$

ce qui permet de décomposer des signaux sur des bases d'ondelettes orthogonales [35].

Relations à deux échelles

La donnée d'une AMR permet de décrire des relations d'une échelle à une autre. D'une part, vu que $\frac{1}{\sqrt{2}}\phi(\cdot/2) \in V_1 \subset V_0$ (dont nous connaissons une base), il existe une suite de coefficients réels $h = (h_k)_{k\in\mathbb{Z}}$ telle que

$$\frac{1}{\sqrt{2}}\varphi(\frac{x}{2}) = \sum_{k\in\mathbb{Z}} h_k\varphi(x-k) \tag{1.9}$$

Par transformée de Fourier, il vient alors

$$\hat{\varphi}(2\xi) = m_0(\xi)\hat{\varphi}(\xi) \text{ avec } m_0(\xi) = \frac{1}{\sqrt{2}}\sum_{k\in\mathbb{Z}} h_k e^{-2i\pi k\xi} \tag{1.10}$$

Dans la suite, nous nous plaçons dans le cadre où il existe une ondelette ψ telle que la famille $\left(2^{-j/2}\psi(2^{-j}t-n)\right)$ soit une base de Riesz de W_j (ce qui est le cas dans le cadre orthogonal). Comme $\psi \in W_0$, $\frac{1}{\sqrt{2}}\phi(\cdot/2) \in W_1 \subset V_0$, et ainsi il existe une suite $g = (g_k)_{k\in\mathbb{Z}}$ telle que

$$\frac{1}{\sqrt{2}}\psi(\frac{x}{2}) = \sum_{k\in\mathbb{Z}} g_k\varphi(x-k) \tag{1.11}$$

Une transformée de Fourier conduit alors à

$$\hat{\psi}(2\xi) = n_0(\xi)\hat{\varphi}(\xi) \text{ avec } n_0(\xi) = \frac{1}{\sqrt{2}}\sum_{k\in\mathbb{Z}} g_k e^{-2i\pi k\xi} \tag{1.12}$$

Ces relations permettent de passer d'un espace d'approximation à un autre. m_0 est la fonction de transfert associée au filtre h et n_0 est celle associée au filtre g. La suite $h = (h_k)_{k\in\mathbb{Z}}$ est un filtre passe-bas, associé à la fonction échelle ϕ ; la suite $g = (g_k)_{k\in\mathbb{Z}}$ est un filtre passe-haut associé à la fonction ondelette ψ.

Ondelettes et filtres

En itérant l'éq. 1.10 (et vu l'éq. 1.12), nous obtenons

$$\hat{\varphi}(\xi) = \hat{\varphi}(0)\prod_{j=1}^{+\infty} m_0\left(\frac{\xi}{2^j}\right) \tag{1.13}$$

$$\hat{\psi}(\xi) = n_0(\xi/2)\hat{\varphi}(\xi/2) \tag{1.14}$$

Dans le cadre orthogonal – lorsque $V_{j-1} = V_j \overset{\perp}{\bigoplus} W_j$ – la fonction d'échelle φ et la fonction ondelette ψ sont totalement définies par la donnée des filtres h et g (à condition que $\varphi, \psi \in L^2$). Nous les représentons sur la figure 1.3, pour l'ondelette de Daubechies à 2 moments nuls et pour l'ondelette de Meyer [87]. Les valeurs de φ et ψ sont calculées par l'algorithme en cascade de Daubechies [35].

La décomposition orthogonale permet d'obtenir une représentation dans une base (décomposition non redondante). A cet égard, les propriétés des coefficients d'ondelettes assurent que pour une grande classe de signaux, la représentation dans une base d'ondelettes est creuse : un nombre restreint de coefficients d'ondelettes permet de représenter le signal tandis que nombre de coefficients sont négligeables [41]. Ceci a motivé l'utilisation de telles représentations en compression de signaux [42] et en compression d'images [70, 92].

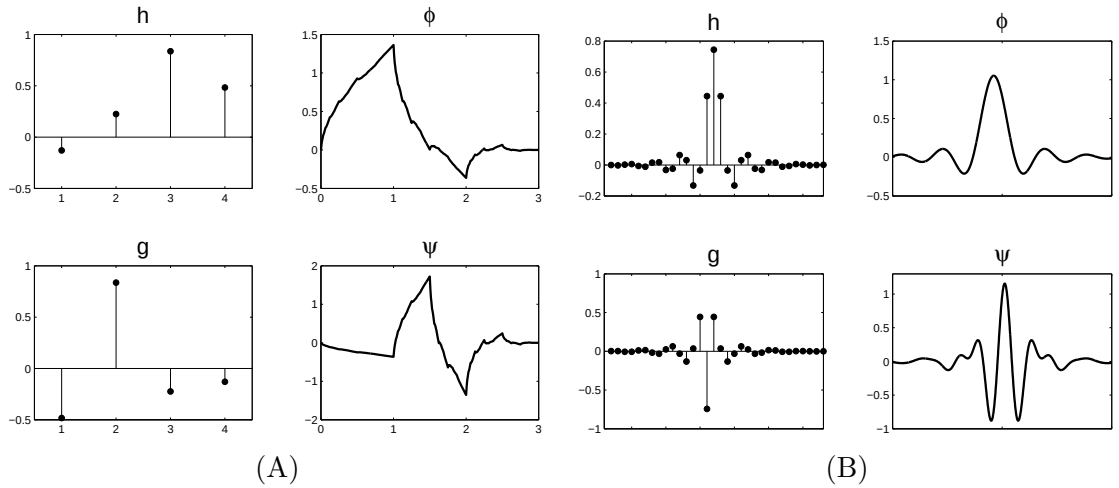


FIG. 1.3 – Filtres, fonction échelle et ondelette ψ : (A) Ondelette de Daubechies à 2 moments nuls (B) Ondelette de Meyer.

1.2.2 Transformée en ondelettes dyadique

Définition 7. *(Transformée en Ondelettes Dyadique, TOD)*
Etant donné $f \in L^2(\mathbb{R})$, la transformée en ondelettes dyadique à la position u et à l'échelle 2^j est définie comme

$$Wf(u, 2^j) = \frac{1}{\sqrt{2^j}} \int_{\mathbb{R}} f(t)\psi\left(\frac{t-u}{2^j}\right) dt = \left(f * \overline{\psi}_{2^j}\right)(u) \tag{1.15}$$

$$\text{où} \quad \psi_{2^j}(t) = \frac{1}{\sqrt{2^j}}\psi\left(\frac{t-u}{2^j}\right) \quad \text{et} \quad \overline{\psi}_{2^j}(t) = \psi_{2^j}(-t)$$

Algorithme à trous de Mallat

L'algorithme de Mallat permet de calculer rapidement les coefficients d'ondelettes de la TOD, grâce à des convolutions discrètes. Nous définissons d'abord les quantités suivantes pour $f \in L^2$ ($< .,. >$ désigne le produit scalaire dans L^2, ϕ une fonction échelle et ψ une ondelette) :

– coefficients d'approximation

$$a_j[n] =< f, \frac{1}{\sqrt{2^j}}\phi\left(\frac{\cdot - 2^j n}{2^j}\right) > \tag{1.16}$$

– coefficients de détails

$$d_j[n] =< f, \frac{1}{\sqrt{2^j}}\psi\left(\frac{\cdot - 2^j n}{2^j}\right) > \tag{1.17}$$

Proposition 1. *Etant donné un filtre h, nous notons h_j le filtre obtenu par insertion de $2^j - 1$ zéros entre deux coefficients de h, et $\overline{h}_j[n] = h_j[-n]$. A partir des coefficients a_0, les coefficients $(a_j)_{j\geq 0}$ et $(d_j)_{j\geq 0}$ définis ci-dessus peuvent être obtenus récursivement d'après les formules de décomposition suivantes : $\forall j \geq 0$,*

$$a_{j+1} = a_j * \overline{h}_j \tag{1.18}$$

$$d_{j+1} = a_j * \overline{g}_j \tag{1.19}$$

L'application de ces formules se représente par un banc de filtres (Fig. 1.4).

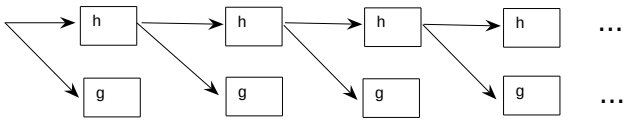


FIG. 1.4 – Banc de filtres impliquant un filtre passe-bas h, correspondant à la fonction échelle ϕ et un filtre passe-haut g, correspondant à l'ondelette ψ.

Application aux décompositions multiéchelles

Nous considérons un signal discret de longueur 2^J. L'hypothèse fondamentale est que les données discrètes $(f_n)_{1\leq n\leq N}$ représentant le signal sont des coefficients d'approximation $(a_J[n])_{1\leq n\leq N}$ (où J est fixé) :

$$f = a_J[n] \tag{1.20}$$

La proposition précédente permet de calculer rapidement l'ensemble $\{a_j[n], d_j[n], J \geq j \geq L\}$ (L entier, $L > J$) ; ceci constitue une décomposition multiéchelles de f, appelée transformée en ondelette dyadique non décimée.

1.2.3 Transformée en ondelettes continue

La transformée en ondelettes continue (TOC) est la décomposition qui sera la plus utilisée dans cette thèse. Ce choix est motivé par la liberté dans le choix des échelles et des propriétés que nous détaillons plus loin. Dans la suite, nous considérons une ondelette $\psi : \mathbb{R} \longrightarrow \mathbb{R}$, $\psi \in L^2 \cap L^p$.

Définition 8. *Etant donné un signal, la transformée en ondelettes continue (TOC) normalisée L^p ($p > 0$) utilisant l'ondelette ψ est définie comme*

$$\forall (u,s) \in \mathbb{R} \times \mathbb{R}_+^*, \quad Wf(u,s) = \frac{1}{s^{1/p}} \int_{\mathbb{R}} f(t)\psi\left(\frac{t-u}{s}\right) dt \tag{1.21}$$

Remarques :

- u est un paramètre de position, s un paramètre d'échelle, l'ensemble des points $(u,s) \in \mathbb{R} \times \mathbb{R}_+^*$ est appelé plan espace-échelle,
- pour (u,s) donné, $Wf(u,s)$ est appelé coefficient d'ondelette,
- à s fixé, $Wf(.,s)$ est un produit de convolution. En effet, si nous posons $\psi_s(x) = \frac{1}{s^{1/p}}\psi_s(x/s)$ et en notant $\overline{\psi}_s(x) = \psi(-x)$, la TOC s'écrit :

$$Wf(u,s) = \left(f * \overline{\psi}_s\right)(u) \tag{1.22}$$

- la normalisation L^p signifie que $||\psi_s||_p$ est indépendante de s :

$$\int_{\mathbb{R}} |\psi_s(x)|^p dx = \int_{\mathbb{R}} |\frac{1}{s^{1/p}}\psi(\frac{x}{s})|^p dx = \int_{\mathbb{R}} \frac{1}{s}|\psi(\frac{x}{s})|^p dx = \int_{\mathbb{R}} |\psi(y)|^p dy$$

  et ainsi $||\psi_s||_p = ||\psi||_p$
- à une échelle 2^j, la TOC normalisée L^2 correspond à la TOD (cf. définition 7).

Calcul effectif

Nous pouvons être amenés à calculer numériquement $Wf(u,s)$ pour $u \in \mathbb{R}$ et pour des échelles $s > 0$ quelconques [34]. Dans le cas où une expression analytique de $\hat{\psi}$ est connue, ce qui sera le cas par la suite, il est possible de calculer la TOC à une échelle quelconque (non nécessairement dyadique), et ce par un algorithme rapide (utilisation d'une FFT).

Rappelons que si $f, g \in L^2$, alors

$$f * g = \overline{\mathcal{F}}\left(\hat{f} \cdot \hat{g}\right) \tag{1.23}$$

Ainsi, si f et ψ appartiennent à L^2, nous pouvons alors écrire :

$$\forall (u,s) \in \mathbb{R} \times \mathbb{R}_+^*, Wf(u,s) = \overline{\mathcal{F}}\left(\widehat{f} \cdot \widehat{\overline{\psi}_s}\right)(u) \tag{1.24}$$

Des propriétés classiques de la transformée de Fourier assurent que

$$\widehat{\overline{\psi}_s}(\omega) = \overline{\widehat{\psi_s}(\omega)} = \frac{1}{s^{1/p}}\overline{s\widehat{\psi}(s\omega)} = \frac{1}{s^{\frac{1}{p}-1}}\overline{\widehat{\psi}(s\omega)} \tag{1.25}$$

et ainsi une expression analytique de $\widehat{\psi}$ donne une expression analytique de $\widehat{\overline{\psi}_s}$.

Par ailleurs, en échantillonnant la fonction f par pas de $\frac{1}{a}, a > 0$, nous avons la relation :

$$\widehat{f}(\frac{n}{a}) \approx \frac{a}{N}\sum_{k=0}^{N-1} f(\frac{ka}{N})e^{-2i\pi\frac{kn}{N}} \tag{1.26}$$

Ainsi les valeurs de $\widehat{f}$ aux points $\{\frac{n}{a}, 0 \leq n \leq N-1\}$ peuvent être approchées par la transformée de Fourier discrète de $\left(\frac{a}{N}f(\frac{ka}{N})\right)_{0 \leq k \leq N-1}$, ce qui peut être calculé par une transformée de Fourier rapide (FFT).

Finalement, le calcul de la TOC de f peut être effectué d'après les étapes suivantes :

1. Calcul approché de $\widehat{f}$ en $\left(\frac{n}{a}\right)_{0\leq n\leq N-1}$, par FFT (eq. 1.26),
2. Calcul exact de $\widehat{\overline{\psi}_s}$, en $\left(\frac{n}{a}\right)_{0\leq n\leq N-1}$ (eq. 1.25),
3. Calcul approché de Wf par transformée de Fourier inverse (inverse-FFT) du produit des deux quantités précédentes (eq. 1.24).

Remarque : afin d'éviter des effets de bords, étant donné un signal de taille N, nous le symétrisons d'abord (par rapport à la dernière valeur) avant d'appliquer ces opérations au signal symétrisé (de taille $2N$). Alors la restriction aux N premières valeurs donne la TOC voulue. Dans toute la suite, les TOC seront calculées d'après cette technique.

1.2.4 Choix de la transformée et de l'ondelette

Chacune des transformées (TOC et TOD) possède des particularités liées à sa formulation, et également aux algorithmes associés. Nous comparons ici certains de leurs aspects. En ce qui concerne les échelles choisies, la TOC n'impose rien *a priori*. Notons qu'avec la TOD, le choix des échelles dyadiques fait que ces échelles croissent très rapidement, ce qui est clairement un problème si nous souhaitons obtenir une localisation fine suivant les échelles. Dans ce but, il est possible de décomposer chaque octave en un certain nombre de voix (ce nombre étant fixé), ce qui permet de rajouter des échelles supplémentaires entre les échelles dyadiques, et qui conduit à $s = 2^j/l, j \in \mathbb{N}, l = 1, ..., L$. En ce qui concerne le choix de l'ondelette, certaines sont plus adaptées à la TOD (ondelettes orthogonales de Daubechies, ondelettes biorthogonales) et d'autres plutôt à la TOC (ondelettes dérivées n-ième de Gaussienne, ondelettes dont la transformée de Fourier admet une expression analytique). A cet égard, notons que la plupart des ondelettes orthogonales ont rarement une expression analytique, ce qui soulève des difficultés pour le calcul de la TOC. Relevons enfin qu'une transformée en ondelettes dyadique non décimée utilisant des ondelettes splines (par les filtres associés [120, 121]) conduit à une décomposition similaire à celle d'une transformée en ondelettes continue utilisant une dérivée n-ème de Gaussienne.

Propriétés des ondelettes

Les ondelettes sont des fonctions qui peuvent être caractérisées par certaines propriétés remarquables ; ces dernières pouvant être incompatibles entre elles, cela implique des choix à faire suivant l'application souhaitée. Nous donnons des propriétés usuelles ainsi que des exemples d'ondelettes.

- **Régularité** : certaines peuvent être de classe C^∞ (comme les dérivées de Gaussienne), d'autres de classe $C^k, k \in \mathbb{N}$ (ondelettes splines), alors que certaines peuvent être discontinues (ondelette de Haar) ou présenter une régularité Lipschitzienne inférieure à 1 (comme certaines ondelettes de Daubechies).
- **Support compact** : $\overline{\{x \in \mathbb{R}/\psi(x) \neq 0\}}$ borné. Les ondelettes utilisées sont localisées dans le plan espace-fréquence (contrairement aux fonctions $(t \mapsto e^{ikt})_{k\in\mathbb{Z}}$, lesquelles ne

sont pas localisées en espace mais uniquement en fréquence). Remarque : une fonction $g : \mathbb{R} \longrightarrow \mathbb{R}$ ($g \in L^1$ ou L^2) ne peut pas vérifier simultanément g et $\widehat{g}$ à support compact ; ceci impose ainsi un compromis entre la localisation spatiale et fréquentielle.

- **Parité** : $\forall x \in \mathbb{R}, \psi(-x) = \psi(x)$.
- **Décroissance rapide** : une ondelette sera dite à décroissance rapide si

$$\forall m \in \mathbb{N}, \exists C_m > 0 / \forall t \in \mathbb{R}, |\psi(t)| \leq \frac{C_m}{1 + |t|^m}$$

- **Nombre de moments nuls** : pour $n \in \mathbb{N}$, une ondelette admet n moments nuls si

$$\forall k = 0...n-1, \int_{-\infty}^{+\infty} t^k \psi(t) dt = 0$$

- **Orthogonalité** : une ondelette ψ est dite orthogonale si pour tout $(j, j', n, n') \in \mathbb{Z}^4$:

$$< \psi_{jn}, \psi_{j'n'} >= \delta_{jj'} \delta_{nn'} \quad (\delta_{jj'} = 0 \text{ si } j \neq j', \delta_{jj} = 1)$$

Exemples

Définition 9. ***Ondelettes dérivées n-ièmes de Gaussienne***
Etant donné un entier $n \geq 1$, l'ondelette dérivée n-ième de Gaussienne est définie à partir de la Gaussienne G, laquelle est donnée par $G(t) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{t^2}{2})$, comme :

$$\psi_n(t) = (-1)^n \frac{d^n}{dt^n} G(t)$$

Remarque : ψ_n est à décroissance rapide, et admet n moment nuls.

Ondelettes dont la transformée de Fourier est connue

Nous avons vu précédemment que le calcul effectif de la TOC d'un signal est simplifié par la connaissance d'une expression analytique de $\widehat{\psi}$. Nous donnons ici des exemples de telles ondelettes. Les ondelettes dérivées n-ièmes de Gaussienne possèdent cette propriété, puisque

$$\widehat{\psi}_n(\omega) = (2\pi\omega)^n \widehat{G}(\omega) \tag{1.27}$$

Par ailleurs, certaines ondelettes sont définies explicitement dans le domaine de Fourier, comme l'ondelette de Shannon ou celle de Meyer [87].

Ondelettes complexes

Un exemple classique d'ondelette complexe est l'ondelette de Morlet : en posant $k_\sigma = e^{-\sigma^2/2}$ et $c_\sigma = \sqrt{1 + e^{-\sigma^2} - 2e^{3\sigma^2/4}}$ (typiquement, $\sigma = 5$), celle-ci est définie explicitement comme

$$\psi(t) = c_\sigma \pi^{-1/4} e^{-t^2/2} (e^{i\sigma t} - k_\sigma)$$

Vu que $\psi \in L^1 \cap L^2$ et que $\widehat{\psi}$ admet une expression analytique [87], la TOC associée est calculable rapidement.

Ondelettes données par un filtre

Tandis que certaines ondelettes sont données explicitement par une expression analytique de ψ ou de $\widehat{\psi}$, certaines ondelettes sont définies par un filtre. Dans un cadre orthogonal, il est possible de calculer des valeurs de ψ, grâce à l'algorithme en cascade proposé par I. Daubechies [35]. Ces filtres jouent un rôle important dans la mesure où ils permettent de calculer rapidement des coefficients d'ondelettes, que ce soit pour calculer une représentation multiéchelles redondante ou pour décomposer un signal dans une base (analyse multirésolution, algorithme de décomposition et de reconstruction).

Bilan

Dans la suite, nous utilisons une TOC. Ceci permettra de suivre précisément l'évolution des singularités. De plus, les ondelettes ψ utilisées ici sont telles qu'il existe une expression analytique de $\widehat{\psi}$, de sorte que le calcul de la TOC est rapide. Enfin, nous privilégions les ondelettes dérivées n-ièmes de Gaussienne, dans la mesure où elles permettent d'obtenir des opérateurs différentiels (utilisés notamment en vision par ordinateur), et où elles assurent l'existence de certaines lignes de maxima, ce que nous verrons au chapitre 2.

1.3 Propriétés des coefficients d'ondelettes

Nous nous intéressons maintenant aux propriétés des coefficients d'ondelettes, donnés par $Wf(u,s)$, $(u,s) \in \mathbb{R} \times \mathbb{R}_+^*$. D'une part, il est possible de mettre en relation l'évolution du module des coefficients avec l'échelle, et la régularité de la fonction f.

1.3.1 EDP vérifiée par les coefficients

Les coefficients d'ondelette $Wf(u,s)$ dépendent du signal f, de l'ondelette ψ, des variables d'espace u et d'échelle s. Il est intéressant de remarquer que la fonction de deux variables $(u,s) \mapsto Wf(u,s)$ vérifie une équation aux dérivées partielles (EDP), qui décrit la loi d'évolution des coefficients. Nous écrivons cette loi dans un cadre L^p, ce qui généralise le cas L^2, vu dans [15].

Pour des raisons de lisibilité, les équations qui suivent sont implicitement valables pour tout $u \in \mathbb{R}$ et pour tout $s \in \mathbb{R}_+^*$, et nous adoptons les notations suivantes :

$$\begin{aligned}
\Theta(u,s) &= \frac{1}{s^{1/p}}\psi(-\frac{u}{s}) = \overline{\psi_s}(u) \\
w(u,s) &= (f * \Theta)(u,s) = Wf(u,s) \\
v(u,s) &= Wh(u,s), \text{ où } h(t) = tf(t)
\end{aligned}$$

Cas général

Proposition 2. *(EDP vérifiée par les coefficients d'ondelette) Si ψ est dérivable, alors la TOC w vérifie l'EDP suivante :*

$$\partial_s w + \frac{u}{s}\partial_u w + \frac{1}{p}\frac{1}{s}w = \frac{1}{s}\partial_u v \tag{1.28}$$

Démonstration

Si nous dérivons Θ par rapport à chacune des variables u et s, nous obtenons

$$\begin{aligned}
\partial_u\Theta(u,s) &= -\frac{1}{s}\frac{1}{s^{1/p}}\psi'(-\frac{u}{s}) \\
\partial_s\Theta(u,s) &= -\frac{1}{p}\frac{1}{s}\Theta(u,s) + \frac{u}{s^2}\frac{1}{s^{1/p}}\psi'(-\frac{u}{s})
\end{aligned}$$

ce qui permet de déduire :

$$\begin{aligned}
\partial_s\Theta(u,s) + \frac{u}{s}\partial_u\Theta(u,s) + \frac{1}{p}\frac{1}{s}\Theta(u,s) &= 0 \\
\partial_s\Theta(u,s) + \partial_u\left(\frac{u}{s}\Theta(u,s)\right) + (\frac{1}{p}-1)\frac{1}{s}\Theta(u,s) &= 0
\end{aligned}$$

La convolution avec f étant une opération linéaire, nous pouvons écrire :

$$f * \partial_s\Theta(u,s) + f * \left(\partial_u\frac{u}{s}\Theta(u,s)\right) + f * \left((\frac{1}{p}-1)\frac{1}{s}\Theta(u,s)\right) = 0$$

ce qui nous ramène à une équation impliquant w : le passage des dérivées de Θ à celles de w se fait d'après les propriétés $\partial_u(f*\Theta) = f*(\partial_u\Theta)$ et $\partial_s(f*\Theta) = f*(\partial_s\Theta)$, d'où :

$$\begin{aligned}
f * \partial_s\Theta(u,s) &= \partial_s w(u,s) \\
f * \left(\partial_u\frac{u}{s}\Theta(u,s)\right) &= \partial_u\left(f * \frac{u}{s}\Theta(u,s)\right) \\
f * \left((\frac{1}{p}-1)\frac{1}{s}\Theta(u,s)\right) &= (\frac{1}{p}-1)\frac{1}{s}w(u,s)
\end{aligned}$$

Il est alors possible de détailler le 2^e^terme apparaissant dans la 2^e^équation ci-dessus :

$$\begin{aligned}
f * \frac{u}{s}\Theta(u,s) &= \int_{\mathbb{R}} f(t)\frac{u-t}{s}\frac{1}{s^{1/p}}\psi(\frac{t-u}{s})dt \\
&= \frac{u}{s}\int_{\mathbb{R}} f(t)\frac{1}{s^{1/p}}\psi(\frac{t-u}{s})dt - \frac{1}{s}\int_{\mathbb{R}} tf(t)\frac{1}{s^{1/p}}\psi(\frac{t-u}{s})dt \\
&= \frac{u}{s}w(u,s) - \frac{1}{s}v(u,s) \\
\partial_u\left(f * \frac{u}{s}\Theta\right) &= \frac{1}{s}w + \frac{u}{s}\partial_u w - \frac{1}{s}\partial_u v
\end{aligned}$$

En combinant ces différents résultats, nous obtenons alors le résultat énoncé.

Simplification dans le cas du signal "rampe"

Pour des signaux simples, l'EDP précédente peut parfois se simplifier, comme pour un signal de type "rampe" :

$$f(t) = \begin{cases} b_0 & \text{pour } t < t_0 \\ a_0(t-t_0) + b_0 & \text{pour } t \geq t_0 \end{cases} \tag{1.29}$$

Dans un premier temps nous pouvons écrire explicitement $w(u,s)$:

$$\begin{aligned} w(u,s) &= \frac{1}{s^{1/p}} \int_{\mathbb{R}} b_0 \psi(\frac{t-u}{s}) dt + \int_{t_0}^{+\infty} a_0(t-t_0)\psi(\frac{t-u}{s}) dt \\ &= 0 + \frac{1}{s^{1/p}} \int_{(t_0-u)/s}^{+\infty} a_0(u+sy-t_0)\psi(y) s dy \end{aligned}$$

$$w(u,s) = \frac{1}{s^{1/p}} a_0 s \int_{(t_0-u)/s}^{+\infty} (u+sy-t_0)\psi(y) dy$$

ainsi que sa dérivée partielle $\partial_u w(u,s)$:

$$\partial_u w(u,s) = \frac{1}{s^{1/p}} a_0 s \left((u + s(t_0-u)/s - t_0)\psi((t_0-u)/s) + \int_{(t_0-u)/s}^{+\infty} \psi(y) dy \right)$$

$$\partial_u w(u,s) = \frac{1}{s^{1/p}} a_0 s \int_{(t_0-u)/s}^{+\infty} \psi(y) dy$$

Dans un deuxième temps, nous pouvons écrire $v(u,s)$:

$$v(u,s) = Wh(u,s) = \frac{1}{s^{1/p}} \int_{\mathbb{R}} t f(t) \psi(\frac{t-u}{s}) dt$$

– En ce qui concerne le terme constant en b_0 :

$$\begin{aligned} \int_{\mathbb{R}} t b_0 \psi(\frac{t-u}{s}) dt &= \int_{\mathbb{R}} (sy+u) b_0 \psi(y) s dy \\ \text{d'où } \partial_u v(u,s) &= b_0 s \int_{\mathbb{R}} \psi(y) dy = 0 \end{aligned}$$

– En ce qui concerne le terme linéaire en $a_0(t-t_0)$:

$$\begin{aligned} \int_{t_0}^{+\infty} t a_0 (t-t_0) \psi(\frac{t-u}{s}) dt &= a_0 \int_{(t_0-u)/s}^{+\infty} (sy+u)(sy+u-t_0)\psi(y) s dy \\ &= a_0 s \int_{(t_0-u)/s}^{+\infty} \left(u^2 + u(2sy-t_0) - ((sy)^2 - t_0)\right) \psi(y) dy \\ \text{d'où } \partial_u (...) &= a_0 s \int_{(t_0-u)/s}^{+\infty} (2u + 2sy - t_0)\, \psi(y) dy \end{aligned}$$

– La somme des deux contributions conduit alors à :

$$\partial_u v(u,s) = a_0 s \int_{(t_0-u)/s}^{+\infty} (2u + 2sy - t_0)\, \psi(y) dy$$

Vu que nous connaissons $\partial_u v$ et w, nous pouvons alors écrire :

$$\begin{aligned} \partial_u v(u,s) - 2w(u,s) &= \frac{1}{s^{1/p}} a_0 s \int_{(t_0-u)/s}^{+\infty} t_0 \psi(y) dy \\ &= t_0 \partial_u w(u,s) \end{aligned}$$

Ainsi nous avons les deux relations suivantes :

$$\begin{aligned} \frac{1}{s}\partial_u v &= \partial_s w + \frac{u}{s}\partial_u w + \frac{1}{p}\frac{1}{s} w \\ \partial_u v &= 2w + t_0 \partial_u w \end{aligned}$$

Ceci conduit alors l'EDP

$$\partial_s w + \frac{u - t_0}{s}\partial_u w + (\frac{1}{p} - 2)\frac{1}{s} w = 0 \tag{1.30}$$

.

1.3.2 Evolution des coefficients en fonction de l'échelle

L'évolution des coefficients d'ondelettes aux échelles fines est liée à la régularité locale du signal. Aux plus grandes échelles, cette évolution est liée aux propriétés globales du signal : à mesure que l'échelle augmente, la transformée en ondelettes a pour effet de lisser les différents détails, au profit des plus grandes tendances. Afin de préciser cela, rappelons que

$$Wf = f * \overline{\psi}_s, \text{ avec } \psi_s(x,y) = \frac{1}{s^{1/p}} \psi\left(\frac{x}{s}\right)$$

et qu'à $s > 0$ fixé, $\psi_s \in L^1 \cap L^2 \cap L^\infty$. Lorsque $f \in L^1$, $f \in L^2$ ou $f \in L^\infty$, les inégalités de Young pour la convolution [52, 102] s'expriment comme

$$\begin{aligned} ||f * \overline{\psi}_s||_\infty &\leq ||f||_1 ||\overline{\psi}_s||_\infty \\ ||f * \overline{\psi}_s||_\infty &\leq ||f||_2 ||\overline{\psi}_s||_2 \\ ||f * \overline{\psi}_s||_\infty &\leq ||f||_\infty ||\overline{\psi}_s||_1 \end{aligned}$$

Relevons alors que

$$\begin{aligned}
||\overline{\psi}_s||_\infty &= \frac{1}{s^{1/p}}||\psi(-\frac{\cdot}{s})||_\infty = \frac{1}{s^{1/p}}||\psi||_\infty \\
||\overline{\psi}_s||_2 &= \frac{1}{s^{1/p}}||\psi(-\frac{\cdot}{s})||_2 = \frac{1}{s^{1/p}}\sqrt{\int_{\mathbb{R}} \psi^2\left(\frac{-x}{s}\right)dx} = \frac{1}{s^{1/p}}\sqrt{\int_{\mathbb{R}} \psi^2(y)sdy} \\
||\overline{\psi}_s||_2 &= \frac{1}{s^{(1/p)-1/2}}||\psi||_2 \\
||\overline{\psi}_s||_1 &= \frac{1}{s^{1/p}}||\psi(-\frac{\cdot}{s})||_1 = \frac{1}{s^{1/p}}\int_{\mathbb{R}} |\psi\left(\frac{-x}{s}\right)|dx| = \frac{1}{s^{1/p}}\int_{\mathbb{R}} |\psi\left(u,v\right))\,|sdy| \\
||\overline{\psi}_s||_1 &= \frac{1}{s^{(1/p)-1}}||\psi||_1
\end{aligned}$$

Ainsi, lorsque $s \longrightarrow +\infty$:

$$\begin{cases}
||\overline{\psi}_s||_\infty & \longrightarrow \quad 0 \\
||\overline{\psi}_s||_2 & \longrightarrow \quad 0 \text{ pourvu que } p < 2 \\
||\overline{\psi}_s||_1 & \longrightarrow \quad 0 \text{ pourvu que } p < 1
\end{cases}$$

Ainsi, asymptotiquement (lorsque $s \to +\infty$), nous sommes assurés que $||Wf||_\infty$ tend vers zéro dans tous les cas si $f \in L^1$, pour $p < 2$ si $f \in L^2$ et pour $p < 1$ si $f \in L^\infty$.

1.3.3 La TOC vue comme un opérateur différentiel

Proposition 3. *[87] Une ondelette ψ à décroissance rapide possède n moments nuls ($n \in \mathbb{N}$) si et seulement si il existe une fonction θ à décroissance rapide telle que*

$$\psi(t) = (-1)^n \frac{d^n\theta}{dt^n}$$

et si $\int_{\mathbb{R}} \theta \neq 0$, alors ψ a exactement n moments nuls.

Proposition 4. *Si ψ est à décroissance rapide et possède n moments nuls, la TOC de $f \in L^1$ peut s'écrire comme : $\forall (u,s) \in \mathbb{R} \times \mathbb{R}_+^*$,*

$$Wf(u,s) = s^n \frac{d^n}{du^n}(f * \overline{\theta}_s)(u) \text{ où } \overline{\theta}_s(t) = \frac{1}{s^{1/p}}\theta(-t/s). \tag{1.31}$$

Cette proposition permet alors de voir que dans le cas où l'ondelette admet un certain nombre de moments nuls, la TOC se comporte comme un opérateur différentiel multiéchelles. Alors les maxima du module de la TOC à une certaine échelle s peuvent être interprétés comme des points de variation brutale pour l'ondelette dérivée première de Gaussienne, et des maxima locaux de la courbure pour l'ondelette ΔG (dérivée seconde de Gaussienne).

1.3.4 Coefficients d'ondelette et régularité d'un signal

Le signal analysé peut être plus ou moins régulier (de classe $C^k, k \in \mathbb{N}$), et présenter des discontinuités en certains points. Afin de préciser la notion de régularité d'une fonction, nous utilisons la notion de régularité Lipschitzienne :

Définition 10. *(Régularité Lipschitzienne en 1D) Etant donné un réel $\alpha > 0$, une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ est ponctuellement Lipschitz-α en t_0 s'il existe une constante $K > 0$ et un polynôme p de degré $m = E[\alpha]$ tels que*

$$\forall t \in \mathbb{R}, |f(t) - p(t)| \leq K|t - t_0|^\alpha$$

Remarques :

- une fonction sera uniformément Lipschitz-α sur un intervalle I si l'inégalité ci-dessus est valable pour tout $t_0 \in I$ (la constante K étant indépendante de t_0) ;
- cette notion de régularité s'étend pour des valeurs de $\alpha \in \mathbb{R}$, et nous verrons que certaines singularités conduisent à des régularités locales $\alpha \leq 0$;
- la régularité Lipschitzienne de f en t_0 (appelée également régularité ponctuelle) est définie par $\alpha = \sup\{\alpha_0 \in \mathbb{R}, f \text{ Lipschitz} - \alpha_0 \text{ en } t_0\}$. L'intérêt de cette grandeur apparaîtra notamment aux chapitres 2 et 6.

Exemples :

1. la fonction $f : \mathbb{R}_+ \to \mathbb{R}$ définie par $f(t) = \sqrt{t}$ est uniformément Lipschitz-$\frac{1}{2}$ sur $\mathbb{R}_+$;
2. la fonction créneau de Heavyside est définie par $f : \mathbb{R}_+ \to \mathbb{R}$, $f(t) = 1$ si $t \geq 0, f(t) = 0$ sinon ; elle est discontinue en 0 et la régularité ponctuelle associée est $\alpha = 0$.

Définition 11. *(Cône d'influence) Dans le cadre où ψ est à support compact, le cône d'influence de u_0 dans le plan espace échelle est l'ensemble des points (u, s) tels que u_0 est inclus dans le support de la fonction $\psi_{u,s}$, définie par $\psi_{u,s}(t) = \frac{1}{s}\psi(\frac{t-u}{s})$.*

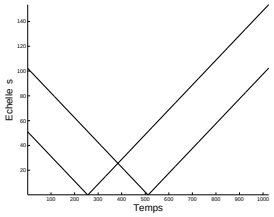


FIG. 1.5 – Cônes d'influences de deux singularités dans le plan espace-échelle (représenté en échelle linéaire).

Remarques :

- étant donné un signal comportant une unique singularité, le cône d'influence correspond à l'ensemble des points $(u, s) \in \mathbb{R} \times \mathbb{R}_+^*$ tels que les coefficients d'ondelettes associés sont non nuls ;
- dans le cas où ψ est symétrique et à support compact (égal à $[-C, C], C > 0$), le cône d'influence est symétrique ;
- dans le cas où ψ n'est pas à support compact, il est possible de définir un pseudo-cône, dans la mesure où l'ondelette est négligeable pour $|t| > C$ (où C est une constante, par exemple $C = 5$ pour l'ondelette ΔG) ;
- si nous considérons un signal dont les singularités sont isolées, lorsque l'échelle augmente, les cônes d'influence sont d'abord disjoints (échelles fines) puis s'intersectent : il y a alors interférence entre les deux singularités (Fig. 1.5).

Les coefficients d'ondelettes $Wf(u, s)$ permettent d'obtenir des informations sur la fonction analysée f. En particulier, ces coefficients sont forts (en module) là où la fonction est irrégulière, tandis qu'ils sont faibles là où la fonction est régulière. Plus précisément, le résultat suivant a été démontré (que nous donnons pour une TOC normalisée L^1) :

Proposition 5. *(Jaffard [57]) Soit une ondelette ψ possédant n moments nuls, de classe $\mathcal{C}^n$ et telle que ses dérivées soient à décroissance rapide. Si $f \in L^2$ est α-Lipschitz ($\alpha \leq n$) en v, alors il existe une constante A telle que*

$$\forall (u, s) \in \mathbb{R} \times \mathbb{R}^+, |Wf(u, s)| \leq As^\alpha \left(1 + \left|\frac{u - v}{s}\right|^\alpha\right)$$

Réciproquement, si $\alpha < n$ est non-entier, et s'il existe une constante A et $\alpha' < \alpha$ tels que :

$$\forall (u, s) \in \mathbb{R} \times \mathbb{R}^+, |Wf(u, s)| \leq As^\alpha \left(1 + \left|\frac{u - v}{s}\right|^{\alpha'}\right)$$

alors f est α-Lipschitz- en v.

Une application de ce résultat est que nous pouvons estimer la régularité Lipschitzienne ponctuelle $\alpha \in \mathbb{R}$, grâce à l'inégalité suivante (vraie pour (u, s) dans le cône d'influence de v) :

$$\log(|Wf(u, s)|) \leq \alpha \log s \tag{1.32}$$

Nous disposons de deux méthodes d'estimation de α, suivant les cas suivants :

1. en un point u_0 correspondant à une singularité, nous considérons la loi d'évolution du maximum dans le cône d'influence de u_0. L'inégalité (1.32) devient une quasi-égalité aux fines échelles, et en effectuant une régression sur les échelles les plus fines, nous obtenons alors une estimation de α [88] ;
2. en un point quelconque (u, s) nous définissons $M(u, s)$ comme la moyenne des coefficients d'ondelettes dans le cône d'influence de u à l'échelle s. En remplaçant $Wf(u, s)$ par $M(u, s)$ dans la relation (1.32), il est possible d'obtenir une estimation de α par une régression similaire.

1.4 Modules maxima de la transformée en ondelette

1.4.1 Notion de module maximum

Définition 12. *(Module Maximum, MM, [88]) Un point* $(u_0, s_0) \in \mathbb{R} \times \mathbb{R}_+^*$ *est un module maximum si la fonction* $u \mapsto |Wf(u, s_0)|$ *admet un maximum local en* $u = u_0$ *(ce maximum étant strict au moins à gauche ou à droite).*

Les MM sont les maxima locaux en espace, à une échelle fixée, du module de la transformée en ondelette de f. Notons que si ψ est dérivable, alors $Wf(., s_0)$ l'est aussi, et s'il y a un MM en (u_0, s_0), nous pouvons écrire :

$$\frac{\partial Wf}{\partial u}(u_0, s_0) = 0 \tag{1.33}$$

Remarque : cette propriété est vraie pour tous les extrema, les MM étant un sous-ensemble. Relevons que les modules maxima se démarquent des extrema de la TOC : si un extremum correspond à un minimum négatif ou un maximum positif, il est également un module maximum ; cependant, un minimum positif ou un maximum négatif n'est pas un module maximum.

1.4.2 Lien avec les singularités présentes dans le signal

L'étude des MM est intéressante dans la mesure où ils sont reliés aux singularités présentes dans le signal : à une échelle $s > 0$ donnée, les coefficients d'ondelette $Wf(u, s)$ sont importants (en module) là où f est irrégulière et faibles là où f est régulière. Plus précisément, nous avons le résultat suivant :

Proposition 6. *(Hwang, Mallat [88]) Soit* ψ *de classe* C^n*, à support compact, et pouvant s'écrire*

$$\psi = (-1)^n \theta^{(n)} \text{ avec } \int_{\mathbb{R}} \theta \neq 0 \text{ (ce qui est notamment le cas si } \psi \text{ possède } n \text{ moments nuls).}$$

Pour $f \in L^1[a, b]$*, s'il existe* $s_0 > 0$ *tel que* $|Wf(u, s)|$ *n'ait aucun maximum local pour* $u \in [a, b]$ *et* $s < s_0$*, alors* f *est uniformément Lipschitz* n *sur tout intervalle* $[a + \epsilon, b - \epsilon]$ *(pour tout* $\epsilon > 0$*).*

Etant donné un MM en (u_0, s_0), $|Wf(u_0, s_0)|$ est d'autant plus important que la singularité est marquée, tandis que le signe de $Wf(u_0, s_0)$ donne une information supplémentaire. En outre, il importe de noter qu'une singularité conduit à un certain nombre de MM, d'autant plus grand que l'ondelette ψ est oscillante (ce qui est relié au nombre maxima locaux de $|\psi|$).

1.4.3 Reconstruction d'un signal à l'aide des modules maxima

Grâce aux propriétés des ondelettes (existence bases d'ondelettes, propriétés de *frame*), il est possible de reconstruire un signal à l'aide de ses coefficients d'ondelettes. En particulier, l'ensemble des coefficients de la TOC ou de la TOD constitue une information complète

(redondante) sur le signal. Par ailleurs, comme les modules maxima sont reliés aux singularités présentes dans un signal, ils constituent une certaine information sur celui-ci. Il est alors intéressant de savoir si cette information est complète, et également s'il est possible de reconstruire le signal original à partir des MM.

Du point de vue théorique, il a été prouvé qu'il existe des signaux différents ayant les mêmes MM de la TOD [10]. Néanmoins, du point de vue pratique, des algorithmes de reconstruction efficaces ont été proposés [61, 89]. Typiquement, un tel algorithme cherche un signal de norme minimale ayant les mêmes coefficients d'ondelettes là où sont situés les maxima. Dans ce type de problème consistant à reconstruire un signal à partir de peu d'information, les approches proposées minimisent souvent une norme L^1 [23, 24]. Ces résultats montrent que les MM constituent une information privilégiée sur le signal.

Conclusion

Dans une optique d'analyse de signaux, nous privilégions l'utilisation de TOC, construite à l'aide d'ondelettes régulières, en particulier les dérivées n-ièmes de Gaussienne. Ceci permettra d'exploiter à la fois les propriétés des coefficients d'ondelettes et celles des modules maxima (MM). En particulier, nous insistons sur les méthodes de détection et de caractérisation de certaines singularités présentes dans le signal.

Chapitre 2

Extrema et Maxima Lines d'ondelettes (1D)

Introduction

Le propos de ce chapitre est de détailler comment la décomposition en ondelettes d'un signal permet de mettre en évidence certaines caractéristiques de celui-ci. Au chapitre précédent, nous avons vu des propriétés relatives aux coefficients d'ondelettes ainsi qu'aux MM, qui sont des extrema de la TOC. D'abord, nous nous intéressons à l'existence de lignes d'extrema de la TOC, ainsi qu'à celle de lignes de MM, appelées Maxima Lines (ML). Ensuite, nous étudions ces lignes d'extrema pour des ondelettes suffisamment régulières, puis ces ML dans le cadre d'ondelettes dérivées n-ièmes de Gaussienne. Nous insistons sur le comportement de celles-ci dans le plan espace-échelle : trajectoire de la ML (notion de dérive spatiale) et évolution de la réponse associée en fonction de l'échelle. Comme application, nous verrons ensuite qu'à partir des ML, il est possible d'extraire différentes caractéristiques du signal : des singularités ponctuelles, dont nous estimons la régularité Lipschitzienne, ou bien des structures non-ponctuelles.

2.1 Lignes d'extrema et Maxima Lines

2.1.1 Extrema de la TOC et lignes d'extrema

Rappelons qu'à une échelle s fixée, les extrema de la TOC sont donnés par les maxima et minima locaux de $Wf(.,s)$. Dans le cas où ψ est dérivable, ils sont donnés par

$$\{(u,s) \in \mathbb{R} \times \mathbb{R}_+^*, F(u,s) = 0 \text{ où } F(u,s) = \partial_u Wf(u,s)\} \tag{2.1}$$

Proposition 7. *(Théorème des fonctions implicites) Soit $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$, différentiable, et soit $(u_0, s_0) \in \mathbb{R}^2$ tel que $F(u_0, s_0) = 0$. Sous l'hypothèse $\partial_u F(u_0, s_0) \neq 0$, il existe une unique fonction g continue telle que $g(s_0) = u_0$ et un intervalle ouvert $I \ni s_0$ tel que*

$$\forall s \in I, F(g(s), s) = 0 \tag{2.2}$$

De plus, g est dérivable et $\forall s \in I, g'(s)\partial_u F(g(s), s) + \partial_s F(g(s), s) = 0$.

Interprétation : s'il existe un extrema en (u_0, s_0), et si $\partial_u F(u_0, s_0) \neq 0$, alors cet extrema fait partie d'une ligne d'extrema (il n'est pas isolé). Cette propriété assure également l'unicité de la ligne d'extrema. A partir de cette équation implicite, nous pourrons déduire des propriétés sur ces lignes d'extrema.

2.1.2 Modules Maxima et Maxima Lines

La notion de Maxima Line (ML) est un outil intéressant, d'une part aux échelles les plus fines, pour détecter et caractériser des singularités, et d'autre part à de plus grandes échelles, où certaines jonctions de ML peuvent être mises en correspondance avec des structures présentes dans le signal.

Définition 13. *Une Maxima Line (ML) est une courbe connexe du plan espace-échelle* (u, s) *formée de MM. Une telle ML sera donnée par*

$$(u(s), s, Wf(u(s), s)), s \in I \text{ intervalle de } \mathbb{R}_+. \tag{2.3}$$

Proposition 8. *(Hummel, Poggio, Yuille [126]) Propagation des MM vers les échelles fines : dans le cadre des ondelettes dérivées n-ième de Gaussienne (i.e,* $\psi = (-1)^n \theta^{(n)}$*, où* θ *est une Gaussienne), et pour* $f \in L^2$*, les MM appartiennent à des courbes connexes qui ne s'interrompent jamais lorsque l'échelle diminue.*

Dans le cas d'ondelettes dérivées n-ièmes de Gaussienne, ce qui sera le cas en 2.4 et 2.3, la prop. 8 assure l'existence de ces Maxima Lines. L'intérêt des ML vient du fait qu'elles permettent de suivre des singularités, ce qui conduit à la notion de dérive de ML.

Définition 14. *(Dérive spatiale d'une ML) Etant donné une ML* $(u(s), s, Wf(u(s), s)), s > 0$*, sa dérive spatiale est définie comme* $s \longmapsto u(s)$ *(pour* $s > 0$*).*

Notons qu'un MM présent à une certaine échelle est associé à au moins une singularité présente dans le signal, et les singularités auxquelles nous nous intéressons sont définies comme les extrémités des ML. Dans la plupart des cas rencontrés dans la pratique, ces singularités sont isolées : étant donné un point v, il existe une échelle $s_0 > 0$ telle que pour tout $s < s_0$, tous les modules maxima qui convergent vers v se trouvent dans le cône d'influence de v. Cependant, il peut exister des singularités oscillantes, telles que celles présentes en 0 pour le signal $f(t) = \sin \frac{1}{t}$ (sujet que nous n'abordons pas ici [5]). Nous verrons que les ML construites numériquement permettent de caractériser la singularité présente en un tel point v, mais il importe de noter que ceci n'est valable que sous l'hypothèse que la singularité associée est isolée – si tel n'est pas le cas, l'approche par ML conduira à une estimation biaisée de la régularité Lipschitzienne. En revanche, à des échelles plus grossières, étant donné que des ML fusionnent, elles sont associées à plusieurs singularités.

Etant donné que les ML sont composées de MM, qui sont des extrema, les ML sont des lignes d'extrema particulières. Ainsi les ML vérifient les propriétés des lignes d'extrema, ce qui complète la proposition 8.

2.2 Etude des extrema locaux de la TOC

Nous nous intéressons aux extrema locaux de la TOC, d'abord dans le cas général où $f \in L^2$, puis dans le cas d'un signal simple (fonction "rampe"). Dans la suite, nous supposons que ψ est dérivable deux fois, et que ψ' et ψ'' sont des ondelettes. Nous notons

$$\begin{aligned} w(u,s) &= \frac{1}{s^{1/p}} \int_{\mathbb{R}} f(t)\psi(\frac{t-u}{s})dt \\ w_1(u,s) &= \frac{1}{s^{1/p}} \int_{\mathbb{R}} f(t)\psi'(\frac{t-u}{s})dt \\ w_2(u,s) &= \frac{1}{s^{1/p}} \int_{\mathbb{R}} f(t)\psi''(\frac{t-u}{s})dt \end{aligned}$$

2.2.1 Cas général

Les dérivées partielles de $w(u,s)$ s'écrivent comme :

$$\begin{aligned} \partial_u w &= -\frac{1}{s} w_1 \\ \partial_s w &= -\frac{1}{p}\frac{1}{s} w + \frac{u}{s^2} w_1 \\ \partial_u \partial_s w &= (\frac{1}{p}+1)\frac{1}{s^2} w_1 - \frac{1}{s^3} w_2 = \partial_s \partial_u w \end{aligned}$$

et cela permet d'exprimer :

$$\begin{aligned} F(u,s) &= \partial_u w(u,s) = -\frac{1}{s} w_1 \\ \partial_u F(u,s) &= \partial_u^2 w(u,s) = \frac{1}{s^2} w_2 \end{aligned}$$

Nous pouvons alors distinguer deux cas. Dans le premier cas, si $(\partial_u F)(u_0, s_0) \neq 0$, il s'agit d'extrema simples : $w_1(u_0, s_0) = 0$. Les extrema sont donnés par (u, s), où $u = g(s)$, et localement $g'(s)$ s'écrit comme :

$$g'(s) = -\frac{\partial_s F}{\partial_u F} = -\frac{\partial_s \partial_u w}{\partial_u^2 w}$$

Dans le deuxième cas, si $(\partial_u F)(u_0, s_0) = 0$, il s'agit de points pour lesquels nous avons $w_1(u_0, s_0) = 0$ et $w_2(u_0, s_0) = 0$. Rappelons alors l'équation 1.28 vérifiée par les coefficients d'ondelette dans le cas général :

$$\partial_s w + \frac{u}{s}\partial_u w + \frac{1}{p}\frac{1}{s} w = \frac{1}{s}\partial_u v$$

En dérivant par rapport à u, nous obtenons :

$$\partial_u \partial_s w + \frac{1}{s}\partial_u w + \frac{u}{s}\partial_u^2 w + \frac{1}{p}\frac{1}{s}\partial_u w = \frac{1}{s}\partial_u^2 v$$

ce qui se simplifie en (u_0, s_0) (où $\partial_u w = 0, \partial_u^2 w = 0$) :

$$\partial_u \partial_s w(u_0, s_0) = \frac{1}{s}\partial_u^2 v(u_0, s_0) \tag{2.4}$$

Remarque : pour la fonction rampe, le 2^{e}terme est nul et ainsi $\partial_u \partial_s w(u_0, s_0) = 0$ (ainsi, $w(.,.)$ admet un point selle en (u_0, s_0)).

2.2.2 Illustration dans le cas d'un signal type "rampe"

Pour le signal de type "rampe" (vu en 1.3.1, formule 1.29), rappelons que la TOC w vérifie l'EDP (éq. 1.30)

$$\partial_s w + \frac{u - t_0}{s}\partial_u w + (\frac{1}{p} - 2)\frac{1}{s}w = 0 \tag{2.5}$$

Nous nous intéressons alors aux extrema en espace de w. Un tel extremum à l'échelle s, présent en u, vérifie $\partial_u w(u, s) = 0$. De plus, nous supposons que $\partial_u^2 w \neq 0$, et ainsi (u, s) appartient à une ligne d'extrema, laquelle est donnée par $u = g(s)$ (prop. 7). Nous nous intéressons maintenant à la trajectoire et à la réponse associée à cette ligne d'extrema.

Trajectoires des extrema dans l'espace-échelle

D'une part, en dérivant l'équation précédente par rapport à u, puis en se plaçant sur un module maximum, où $\partial_u w(u, s) = 0$, nous obtenons successivement :

$$\partial_u \partial_s w + \frac{1}{s}\partial_u w + \frac{u - t_0}{s}\partial_u^2 w + \left(\frac{1}{p} - 2\right)\frac{1}{s}\partial_u w = 0$$

$$\partial_u \partial_s w + \frac{g(s) - t_0}{s}\partial_u^2 w = 0 \tag{2.6}$$

D'autre part, la propriété $\partial_u w(u, s) = 0$ pour tout $u = g(s)$ permet d'affirmer successivement :

$$\tfrac{d}{ds}\left((\partial_u w)(g(s), s)\right) = 0 \quad \text{et} \quad g'(s)\partial_u^2 w + \partial_s \partial_u w = 0$$

En combinant ces deux résultats, nous obtenons :

$$\left(g'(s) - \frac{g(s) - t_0}{s}\right)\partial_u^2 w = 0 \tag{2.7}$$

Vu que $\partial_u^2 w \neq 0$, nous obtenons alors une équation différentielle ordinaire (EDO) donnant la trajectoire d'un extremum dans l'espace-échelle :

$$g'(s) = \frac{1}{s}(g(s) - t_0) \tag{2.8}$$

En fixant comme condition initiale $g(0) = t_0$, cette EDO admet une unique solution. Notons de plus que :

- si ψ admet un extremum en zéro, l'équation 2.8 admet pour solution $g(s) = t_0 \quad (\forall s > 0)$ (la trajectoire associée à cet extremum est constante) ;
- sinon, pour un extremum non centré, cette EDO admet une solution de la forme $g(s) = t_0 + C_0 s \quad \forall s > 0$ (où C_0 est une constante).

Evolution de la réponse associée à une ligne d'extrema dans l'espace-échelle
Aux points où il y a un extremum, vu la propriété $\partial_u w = 0$, l'équation 2.5 devient :

$$\partial_s w = \lambda \frac{1}{s} w, \text{ avec } \lambda = 2 - \frac{1}{p} \tag{2.9}$$

Si nous nous plaçons le long d'une ligne d'extrema, le coefficient d'ondelette $w(u, s)$ s'écrit comme une fonction d'une variable : $h(s) = w(g(s), s)$. Vu que $\partial_u w(g(s), s) = 0$, nous avons

$$h'(s) = \lambda \frac{1}{s} h(s), \text{ avec } \lambda = 2 - \frac{1}{p} \tag{2.10}$$

D'où il vient : $h(s) = K \exp\left(\int \frac{\lambda}{s} ds\right) = K \exp\left(\lambda \log |s|\right) = K s^{\lambda}$. Ainsi le coefficient d'ondelette le long d'une ligne de maxima évolue comme

$$w(g(s), s) = K s^{2 - \frac{1}{p}} \tag{2.11}$$

Pour une normalisation L^2, nous obtenons $h(s) = K s^{3/2}$. Notons qu'il s'agit du comportement des coefficients le long d'une ligne isolée (pas de fusion avec d'autres lignes), dans le cadre du signal type "rampe".

Remarque : l'équation (2.11) étant valable pour les extrema de Wf, nous avons $|Wf| \leq K s^{2 - \frac{1}{p}}$, de sorte que pour $p < 1/2$, $||Wf||_\infty \longrightarrow 0$ lorsque $s \longrightarrow +\infty$. Cette propriété, valable pour une fonction "rampe", complète les résultats vus en 1.3.2.

2.3 Etude numérique des Maxima Lines

2.3.1 Construction des ML (en 1D)

Pour déterminer les ML associées à un signal (utilisant une certaine ondelette), nous calculons d'abord une TOC associée à différentes échelles $s_1, ..., s_{\max}$ (section 1.2.3). Puis, nous déterminons les cartes de MM associées, (un MM étant supérieur à ses voisins, au sens strict à droite ou à gauche). Partant d'un MM (localisé en (x', y'), à l'échelle la plus fine s_1), nous construisons une ML d'après les règles suivantes :

- Initialisation : $L \longleftarrow (u, s_1, Wf(u, s_1))$
- Chaînage de s_i à s_{i+1} :
  - s'il existe un MM de même signe à l'échelle suivante, localisé en un point u' voisin de u ($u' = u \pm 1$), la ML se prolonge comme $L \longleftarrow (L, (u', s_{i+1}, Wf(u', s_{i+1}))$;
  - sinon, la ligne s'interrompt.

Nous présentons sur la figure 2.1 quelques exemples de ces squelettes pour des signaux simples. En particulier, nous pouvons observer la différence, suivant l'ondelette utilisée, entre l'ondelette dérivée première de Gaussienne et l'ondelette dérivée seconde de Gaussienne (plus généralement, nous observons des résultats similaires entre une ondelette impaire et une ondelette paire). Ceci permet de voir comment se comporte globalement l'ensemble de ces ML, appelé squelette des ML. Notons en particulier que lorsque l'échelle croît, certaines ML peuvent fusionner tandis que d'autres peuvent disparaître. Ceci conduit à poser la définition suivante :

Définition 15. *(Jonction de ML) Un point* $(u^*, s^*) \in \mathbb{R} \times \mathbb{R}_+^*$ *du plan espace échelle est appelé jonction de ML s'il existe deux ML,* $L_i = (u_i(s), s, Wf(u(s), s)$*, (*$i = 1, 2$*), telles que*

$$\forall s < s^*, u_1(s) \neq u_2(s) \text{ et } \forall s \geq s^*, u_1(s) = u_2(s)$$

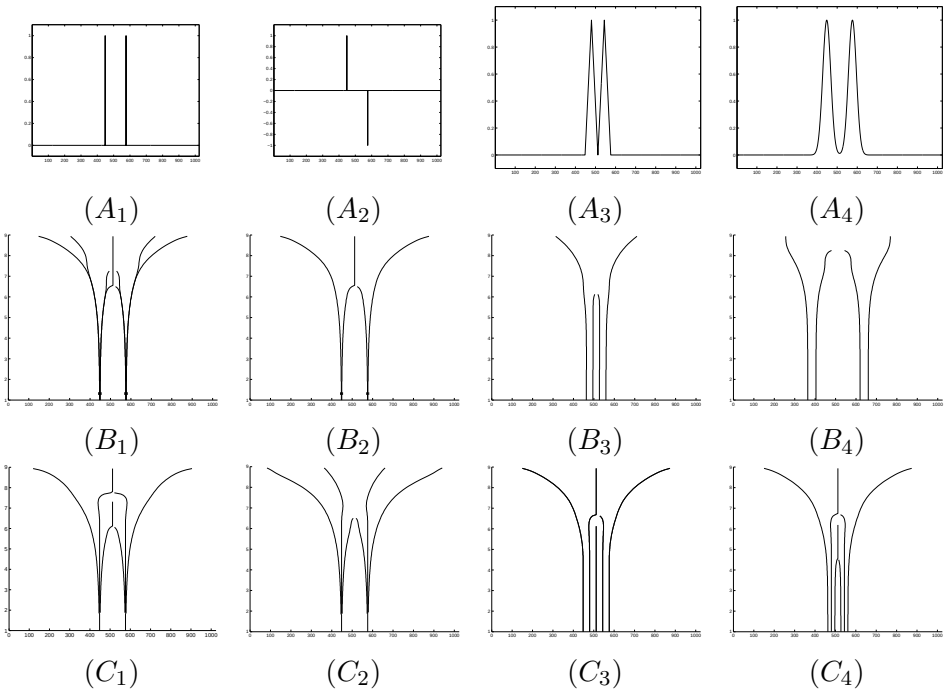


FIG. 2.1 – Pour tout $i \in \{1, 2, 3, 4\}$: (A_i) Signal ; (B_i) Squelette des ML (ondelette dérivée de Gaussienne) ; (C_i) Squelette des ML (ondelette ΔG) ; $i = 1$, deux impulsions de Dirac de même signe ; $i = 2$, deux impulsions de Dirac de signes opposés ; $i = 3$, signal linéaire par morceaux ; $i = 4$, somme de deux Gaussiennes.

2.3.2 Exemple de disparition d'une ligne de maxima

Notons que le cadre des MM est assez différent de celui des extrema de la TOC. Afin d'illustrer cela, nous considérons un signal simple, composé de trois masses de Dirac (Fig. 2.2 (A)), et calculons le squelette des ML (Fig. 2.2 (B)). Sur cet exemple, où les singularités sont isolées, les cônes d'influence disjoints aux échelles fines, et ainsi nous voyons apparaître la TOC comme la superposition de fonctions ondelettes (Fig. 2.2 (C)). Etant donné que l'ondelette présente des oscillations autour de zéro, c'est plusieurs MM qui apparaissent à partir d'une seule singularité, et leur nombre sera d'autant plus grand que l'ondelette est oscillante (Fig. 2.2 (D)). Par ailleurs, il est intéressant de mesurer comment évoluent Wf et $|Wf|$ à mesure que l'échelle croît. Notons d'abord que si l'ondelette est suffisamment régulière, la TOC tend à être de plus en plus lisse. Ainsi ses extrema fusionnent peu à peu. Quant à elle, l'évolution de $|Wf|$ est plus complexe. Dans l'exemple considéré ici, nous pouvons mettre en évidence la disparition d'un MM (et donc l'interruption d'une ML) autour d'une certaine échelle critique, notée s_c (Fig. 2.2 (A'–C'). Enfin, nous nous intéressons à l'évolution des ML issues de la singularité centrale. Sous l'influence des deux autres, dans un premier temps, les MM issus de cette singularité subsistent ; puis, dans un deuxième temps, ces MM sont atténués, jusqu'à disparaître. Plus précisément, nous voyons que :

- pour $s < s_c$, un MM positif associé à l'atome central subsiste (Fig. 2.2 (A')) ;
- pour $s = s_c$, ce MM disparaît (interruption d'une ML) (Fig. 2.2 (B')) ;
- pour $s > s_c$, seuls sont présents les MM associés aux autres singularités (Fig. 2.2 (C')).

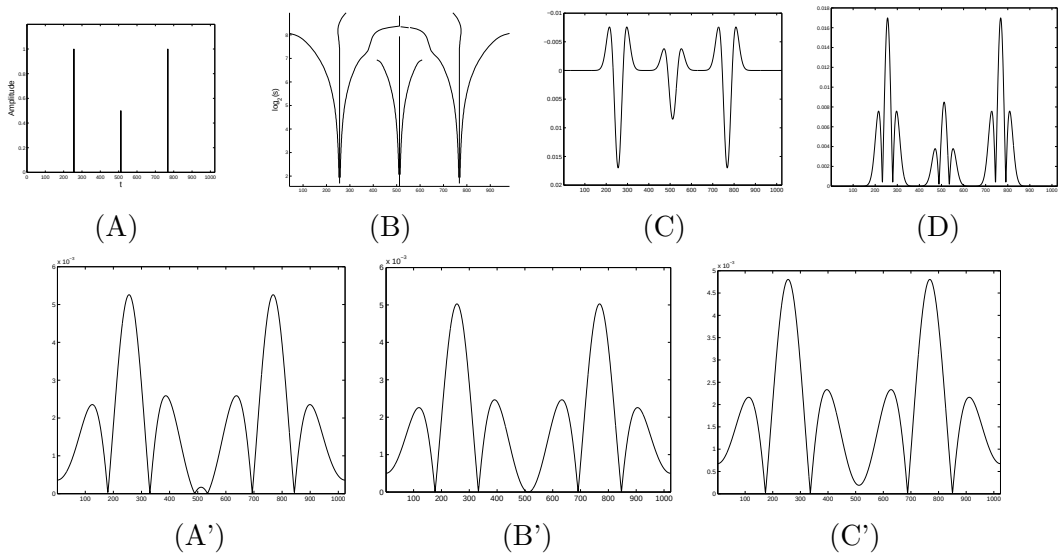


FIG. 2.2 – (A) Somme de trois masses de Dirac ; (B) Squelette des ML (ondelette ΔG) ; (C) TOC associée à une échelle intermédiaire s ; (D) Module de la TOC (échelle s). (A'–C') Evolution de $|Wf|$ autour d'une échelle critique s_c, à laquelle une ligne de maxima s'interrompt (lorsque cette échelle croît) : (A') $s < s_c$; (B') $s = s_c$; (C') $s > s_c$.

2.3.3 Interaction entre singularités d'importances différentes

Nous considérons le signal $f = \lambda\delta_{a_1} + \delta_{a_2}$ où $\lambda \in \mathbb{R}$ et nous observons l'effet de la variation de λ sur le squelette des ML (Fig. 2.3) :

- pour $\lambda = 1$, plusieurs ML fusionnent : les deux singularités interfèrent de manière équivalente l'une sur l'autre (Fig. 2.3 (A)) ;
- pour λ légèrement différent de 1, une ML issue de la plus petite singularité subsiste, tandis que les deux autres sont interrompues du fait de la présence d'une plus grande singularité. Notons aussi que la plus petite est nettement influencée, ce qui rend les ML associées instables (Fig. 2.3 (B)) ;
- pour λ très différent de 1, une des singularités prédomine complètement, et les ML associées à l'autre atome s'interrompent rapidement (Fig. 2.3 (C)).

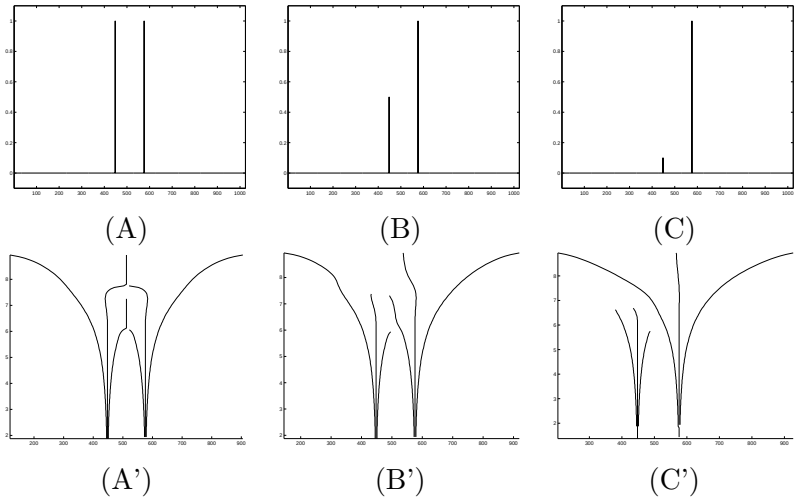


FIG. 2.3 – Singularités d'importance différentes – Etude du signal $\delta_{a_1} + \lambda\delta_{a_2}$ en fonction de λ (TOC, ondelette ΔG) : (A) $\lambda = 1$; (B) $\lambda = 0.5$ (C) $\lambda = 0.1$ (A', B', C') Squelettes associés.

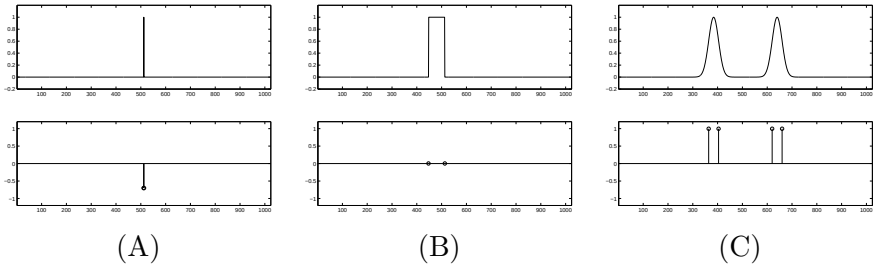


FIG. 2.4 – Etude des régularités, basée sur les ML de la TOC (ondelette dérivée première de Gaussienne) : (A) Impulsion type masse de Dirac ; (B) Créneau ; (C) Deux bosses continues.

2.4 Applications

2.4.1 Détection de singularités et calcul des régularités associées

Nous présentons ici des exemples de détection de singularités présentes dans un signal. A cet effet, nous considérons différents signaux (impulsion type masse de Dirac, créneau, somme de deux Gaussiennes), à partir desquels nous pouvons calculer la TOC associée à l'ondelette dérivée de Gaussienne, normalisée L^1. Rappelons que, dans ce cadre, cela implique que les MM correspondent à des points d'inflexion. Ceci explique alors la localisation des MM. Une fois les ML construites, il est possible de se focaliser sur les plus petites échelles, et d'estimer la régularité Lipschitzienne associée (Fig. 2.4). Nous constatons que les régularités estimées d'après les ML sont cohérentes avec les valeurs théoriques (Tab. 2.1), l'estimation étant légèrement biaisée.

Tab. 2.1 – Valeurs numériques des régularités estimées le long des ML (A, B, C : *idem* fig. précédente). Pour chaque singularité, nous obtenons une valeur de α.

	α	Valeurs estimées $\widehat{\alpha}$			
(A)	-1	-0.995			
(B)	0	0.0026	0.0024		
(C)	1	0.997	0.997	0.997	0.997

2.4.2 Mise en évidence de certaines structures

Nous nous plaçons maintenant dans le cadre d'une TOC normalisée L^1 utilisant l'ondelette ΔG. Nous avons vu que les ML aux échelles fines permettent de détecter et de caractériser certaines singularités ponctuelles. Il est intéressant de voir qu'à des échelles plus grandes, elles permettent de mettre en évidence certaines structures locales. Cela provient des deux éléments : d'une part la dérive spatiale des ML ($s \mapsto (u(s))$) et d'autre part l'évolution de la réponse en fonction de l'échelle ($s \mapsto |Wf(u(s), s)|$). En premier lieu, notons qu'aux échelles les plus fines, les détails fins sont présents tandis qu'aux échelles les plus grandes, tous les détails sont atténués et seule une tendance globale subsiste (l'atténuation de la TOC d'une fonction $f \in L^\infty$ a été vue en 1.3.2). Relevons que cette évolution est progressive, en particulier nous observons qu'à des échelles intermédiaires (non connues *a priori*), il existe des points de jonctions où certaines ML fusionnent.

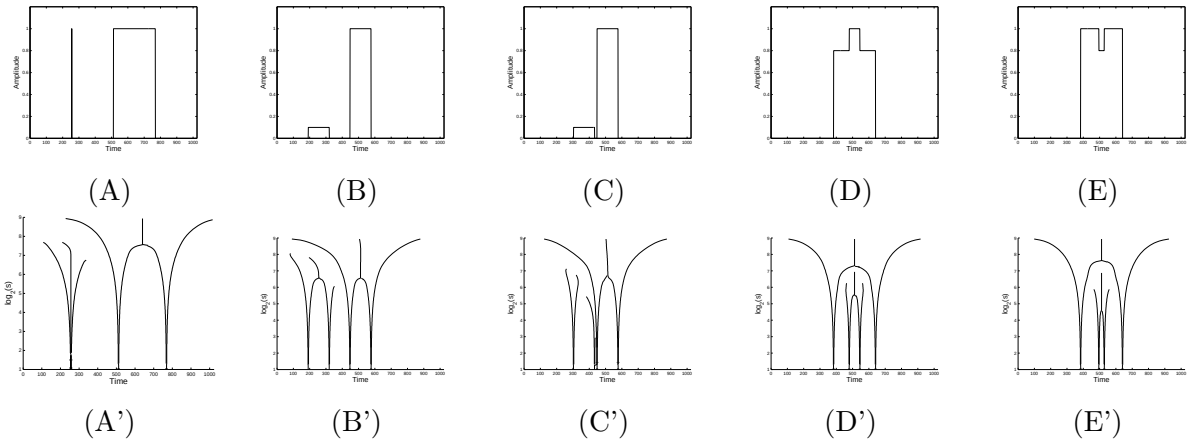


FIG. 2.5 – Quelques exemples utilisant l'ondelette ΔG : (A) Impulsion type Dirac et créneau ; (B) Deux créneaux relativement éloignés ; (C) Deux créneaux proches l'un de l'autre ; (D) Deux créneaux emboîtés (de même signe) ; (E) Deux créneaux emboîtés (de signes opposés) ; (A'– E') Squelettes des ML associés aux signaux précédents.

Ces jonctions peuvent alors être mises en correspondance avec des structures présentes dans le signal. Nous observons que dans le cas de l'ondelette ΔG, cela conduit à mettre en évidence des structures de type créneau, et plus généralement des structures délimitées à droite et à gauche par des singularités dont les MM associés sont de même signe et associées à des réponses similaires (Fig. 2.5). Les ML associées à ces MM dérivent dans l'espace-échelle et fusionnent en un point (u^*, s^*). Ce point, correspondant à une jonction de ML (définition 15) est caractéristique de la structure, en ce sens que u^* correspond à une position correspondant approximativement à son centre (ou exactement si la structure est symétrique) et que s^* est proportionnel à sa taille. Plus précisément, nous relevons sur la figure 2.5 les éléments suivants :

- En (A–A'), le créneau (comprenant deux singularités type 'step') conduit à une jonction tandis que la masse de Dirac (singularité ponctuelle) conduit à des ML s'interrompant ;
- En (B–B'), les deux créneaux conduisent à une jonction, et les ML issues du créneau le plus important sont dominantes par la suite ;
- En (C–C'), le plus petit créneau ne se distingue pas suffisamment de l'autre : les ML issues d'une de ses singularités (celle située à droite) s'interrompent prématurément du fait de la proximité avec la structure dominante ;
- En (D–D') et (E–E'), des structures imbriquées donnent lieu à des squelettes imbriqués.

Remarque : cette notion de jonction de ML sera vue ultérieurement dans un cadre bidimensionnel, en particulier pour la détection de structures 2D présentes dans une image.

Nous nous intéressons maintenant à l'évolution de la réponse en fonction de l'échelle, plus particulièrement nous reprenons l'exemple du créneau (Fig. 2.6 (A)). Nous considérons cette

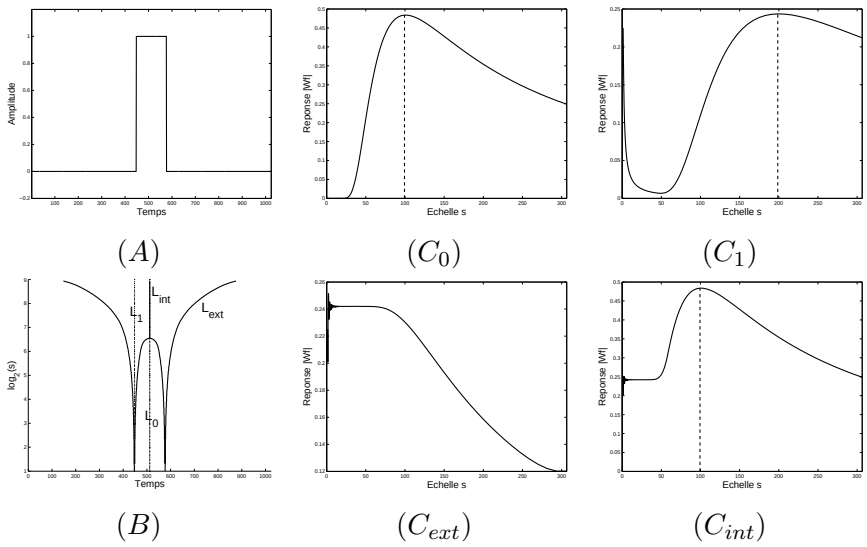


FIG. 2.6 – (A) Créneau ; (B) Squelette des ML – L_0 est un chemin localisé au centre du créneau (indépendante de s), L_1 est un chemin localisé au niveau d'une singularité aux échelles fines, L_{ext} est une ML extérieure (par rapport à la structure "créneau"), L_{int} est une ML intérieure ; (C_0) Evolution de la réponse suivant L_0 ; (C_1) Evolution de la réponse suivant L_1 ; (C_{ext}) Evolution de la réponse suivant L_{ext} ; (C_{int}) Evolution de la réponse suivant L_{int}.

évolution le long des quatre chemins dans l'espace échelle , représentés sur la figure 2.6 (B)) :

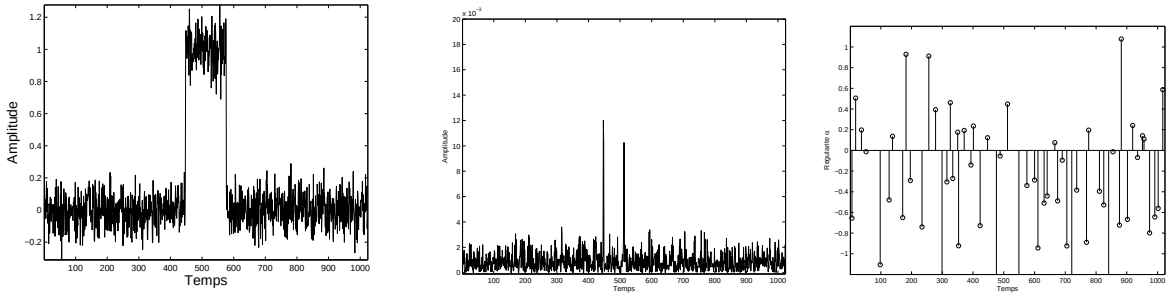
- L_0 – chemin localisé au centre du créneau (indépendant de s). Pour de faibles valeurs de s, la réponse est nulle, puis elle augmente progressivement (le cône d'influence contient le centre du créneau) ; elle atteint alors un maximum en s^* avant de décroître.
- L_1 – chemin localisé au niveau d'une singularité aux échelles fines. Aux fines échelles, comme L_1 se situe dans le cône d'influence de la singularité, l'évolution de la réponse dépend de la régularité Lipschitzienne associée (ici $\alpha = 0$) ; ensuite, aux échelles intermédiaires, vu que la structure créneau apparaît, cela influence la réponse dans tout le voisinage, de sorte que la réponse suivant L_1 augmente, jusqu'à atteindre un maximum (relevons que ce maximum n'est pas atteint en s^* mais à une échelle supérieure) ;
- L_{ext} – ML extérieure (par rapport à la structure "créneau"). La réponse est stable aux fines échelles (réponse non nulle car associée à une singularité) puis décroît ; ceci provient du fait qu'aux fines échelles, cette ML est assez proche de la structure et qu'elle s'en éloigne à mesure que l'échelle croît ;
- L_{int} – ML intérieure. La réponse est stable aux fines échelles (comme L_{ext}) augmente jusqu'à atteindre un maximum (en s^*, vu queL_{int} rejoint L_0), et décroît par la suite (aux grandes échelles, les deux singularités se comportent comme une seule).

2.4.3 Etude dans un cadre bruité

Le SNR (*Signal-to-Noise Ratio*) est une grandeur mesurant le niveau de bruit. Plus il y a de bruit, plus cette valeur est faible.

Définition 16. *Etant donné un signal auquel un bruit est ajouté, en notant A_{signal} et A_{bruit} leurs amplitudes respectives, le SNR (mesuré en dB) est défini comme*

$$SNR = 20 \log \frac{A_{signal}}{A_{bruit}} \tag{2.12}$$



(A) (B) (C)

FIG. 2.7 – Estimation de régularités avec l'ondelette dérivée de Gaussienne : (A) Créneau bruité (SNR= $20dB$) ; (B) Modules maxima à l'échelle la plus fine ; (C) Régularités estimées le long de ML associées.

Nous considérons un signal de type créneau, auquel un bruit blanc Gaussien a été ajouté (SNR=20 dB , Fig. 2.7 (A)). En utilisant l'ondelette dérivée de Gaussienne, nous pouvons calculer les MM et les régularités Lipschitziennes associées (Fig. 2.7 (B–C)). En utilisant différentes ondelettes, nous représentons les TOC (normalisées L^1) associées au créneau bruité (Fig. 2.8 (A–C)), ainsi que les squelettes des ML correspondants (Fig. 2.8 (A'–C')).

En ce qui concerne les TOC, nous observons à la fois les singularités marquées et les fluctuations liées au bruit. A mesure que l'échelle croît, les singularités prédominent par rapport au bruit du fait de l'élargissement des cônes d'influence (et que les singularités correspondent à une réponse plus forte). Par ailleurs, en observant la structure hétérogène du squelette des ML, il se dégage deux types de ML : celles liées au fluctuations liées au bruit – nombreuses et de faible réponse – et celles liées à des singularités remarquables (ici les bords du créneau). Notons également que lorsque leurs cônes d'influence se rencontrent, il y a une interaction entre les singularités, de sorte que les ML résultantes sont alors associées à plusieurs singularités.

Distinguons alors suivant l'ondelette utilisée :

– ondelette dérivée (première) de Gaussienne (Fig. 2.8 (A–A')) : les ML dérivent vers l'extérieur du créneau ;

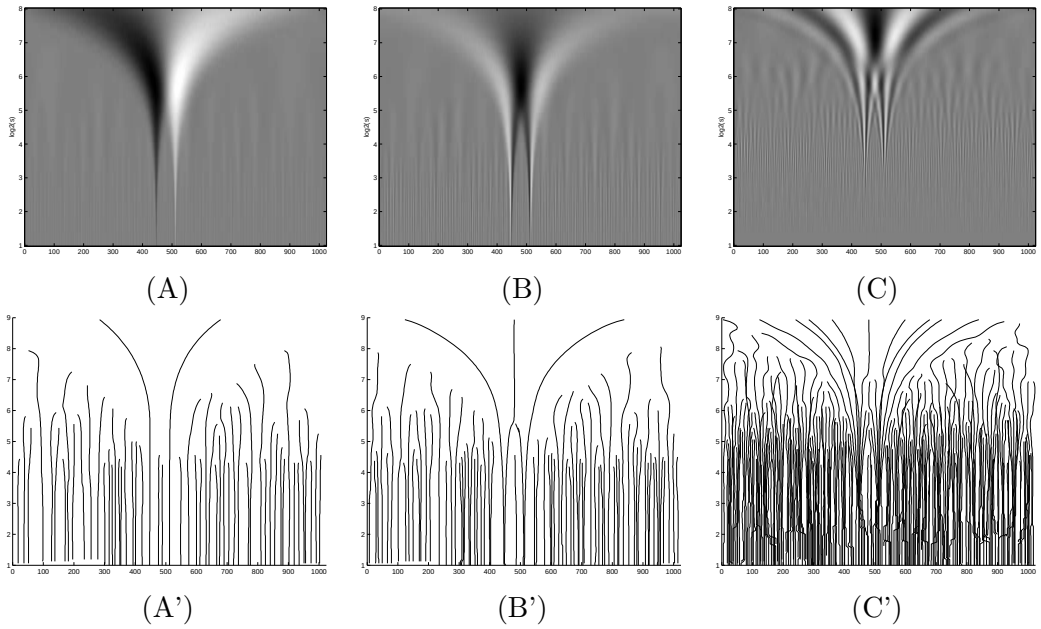


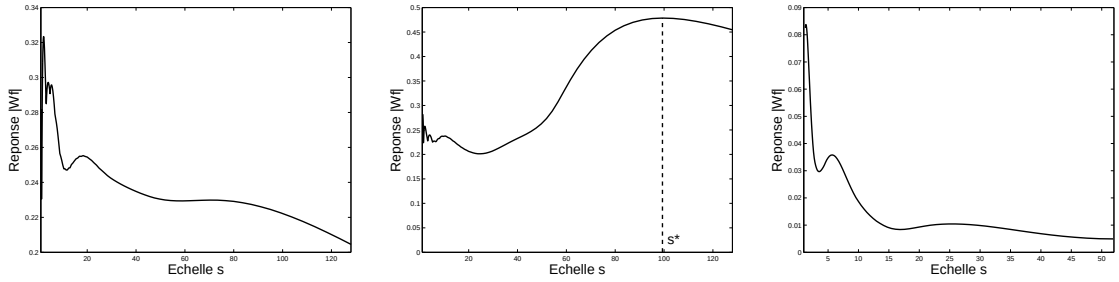
FIG. 2.8 – Influence de l'ondelette utilisée (exemple du créneau bruité) – Transformée en ondelette continue (A–C) et squelettes des Maxima Lines associés (A'–C'). (A, A') Ondelette dérivée de Gaussienne ; (B, B') Ondelette ΔG ; (C, C') Ondelette de Morlet (cette dernière étant complexe, nous considérons le module de la TOC).

- ondelette ΔG, dérivée seconde de Gaussienne (Fig. 2.8 (B–B')) : certaines ML dérivent vers l'extérieur du créneau, tandis que celles qui dérivent vers l'intérieur, et fusionnent à une certaine échelle (jonction de ML). Cette échelle est dite caractéristique du créneau, en ce sens que les singularités qui le délimitent correspondent aux ML qui fusionnent ;
- ondelette de Morlet (Fig. 2.8 (C–C')) la structure est plus complexe, tout en présentant certaines similitudes avec le cas précédent : le module de l'ondelette de Morlet est une fonction paire, d'où la similitude avec l'ondelette ΔG (rappelons néanmoins qu' ici l'existence des lignes n'est pas assurée théoriquement).

Nous nous plaçons maintenant dans le cadre de l'ondelette ΔG. En complément de l'aspect spatial (dérive des ML), il est alors intéressant de comparer les réponses associées aux différentes ML. Notations :

L_{ext}, L_{int}	:	lignes associées à des singularités marquées,
L_{ext}	:	dérivant vers l'extérieur du créneau,
L_{int}	:	dérivant vers l'intérieur.
L_b	:	ligne associée à du bruit

Notons d'abord que L_b s'interrompt rapidement et que l'évolution de la réponse est fluctuante (Fig. 2.9 (C)), dans la mesure où elle interagit rapidement avec d'autres lignes également associées à du bruit, lesquelles peuvent autant la renforcer que l'atténuer.



(A) (B) (C)

FIG. 2.9 – Evolution de la réponse le long de différentes ML (ondelette ΔG) : (A) ML associée à une singularité, dérivant vers l'extérieur du créneau ; (B) ML associée à une singularité, dérivant vers l'intérieur du créneau ; (C) ML correspondant à du bruit.

Ensuite, en comparant L_{ext} et L_{int}, nous relevons les éléments suivants :

- aux échelles fines, l'évolution est conditionnée par la régularité Lipschitzienne (et la normalisation de la TOC, ici L^1) ; ainsi, si L_{ext} (ou L_{int}) correspond à une singularité type Heavyside ($\alpha = 0$), la réponse associée est croissante ;
- à des échelles plus grandes, la réponse suivant L_{ext} s'atténue rapidement (ligne isolée) tandis que celle suivant L_{int} augmente jusqu'à atteindre un pic (interaction avec une autre ligne qui la renforce). Ce pic de ML est atteint à une échelle s^* dite caractéristique, et nous relevons que s^* correspond approximativement à la taille du créneau (nous verrons cela en détail au chapitre 4).
- lorsque $s \longrightarrow +\infty$, les réponses (suivant L_{ext} et L_{int}) tendent vers zéro (conformément aux résultats vus en 1.3.2, pour une TOC normalisée L^1 et $f \in L^\infty$).

En conclusion, l'approche par ML permet de distinguer les singularités les plus marquées dans le signal de celles liées aux bruit, et nous pouvons relever différents éléments :

- Amplitude de la réponse : L_{ext} et L_{int} sont associées à des réponses plus fortes que L_b ;
- Dérive spatiale : L_{ext} et L_{int} seront plus stables spatialement (ceci est lié à l'amplitude des réponses associées) alors que L_b tend à être influencée par d'autres lignes ;
- Persistance des lignes : aux grandes échelles, L_{ext} et L_{int} perdurent alors que L_b disparaît ;
- Evolution de la réponse en fonction de l'échelle s : tandis que le comportement suivant L_b n'est pas très significatif, les comportements se distinguent suivant L_{ext} (Fig. 2.9 (A)) ou L_{int} (Fig. 2.9 (B)).
- Interaction avec d'autres lignes : globalement, des lignes correspondant à du bruit tendent à se neutraliser alors que des lignes associées à des singularités peuvent se renforcer mutuellement dans certains cas, et aboutir à une certaine échelle caractéristique, correspondant à la taille d'un objet présent dans le signal, ici le créneau (nous verrons ultérieurement comment cela se généralise en 2D).

Conclusion

Dans un cadre monodimensionnel, nous avons vu que les décompositions en ondelettes permettent d'extraire certains éléments remarquables du signal grâce aux ML. Ceux-ci peuvent être des singularités ponctuelles, ou des objets plus complexes délimités par des singularités, comme la structure associée au créneau dans l'exemple précédent. Nous soulignons que l'approche multiéchelles est ici pleinement utilisée, autant pour la dérive spatiale d'une ML, qu'en ce qui concerne l'évolution de la réponse suivant cette ML en fonction de l'échelle. L'approche par ML permet de mesurer la dérive spatiale des ML à mesure que l'échelle croît, et également de voir si certaines ML perdurent ou disparaissent. Un élément intéressant de cette approche est qu'une fois les ML construites, il est possible de parcourir cette ML vers les échelles les plus fines, et d'identifier la singularité associée. Nous verrons plus loin l'intérêt de ces propriétés, dans un cadre bidimensionnel, où nous tenterons de mettre en correspondance des objets présents dans l'image et des singularités associées à certaines ML (chap. 6 et 7). Notons enfin que les caractéristiques mises en évidence ne sont pas nécessairement ponctuelles, puisqu'elles peuvent correspondre à certaines structures présentes dans le signal.

Chapitre 3

Décompositions multiéchelles d'une image et lignes de maxima

3.1 Introduction

La question de trouver un modèle satisfaisant pour décrire une image quelconque est un problème difficile. Des résultats en théorie de l'approximation, dans des espaces fonctionnels tels que L^2, montrent que certaines familles de fonctions sont bien adaptées au traitement d'images. Cependant, la complexité des images naturelles est telle qu'il n'y a pas d'espace canonique pour étudier celles-ci. Différents espaces ont été proposés, tels que les espaces BV (fonction à variations bornées), les espaces de Besov et les espaces de fonctions C^α hors de contours C^α [65, 6, 21, 122]. Schématiquement, la plupart des images naturelles peuvent être décomposées en deux parties : une géométrique et l'autre texturée [48, 62, 64]. Notons que la frontière entre géométrie et texture est floue : par exemple, un motif géométrique qui se répète peut être apparenté à de la texture. Nous nous intéressons plus particulièrement à la partie géométrique. Celle-ci correspond à des régions où le niveau de gris est homogène (ou plus généralement associé à une fonction régulière), lesquelles sont séparées par des contours. Ces derniers correspondent à de fortes énergies, et ils concentrent une grande partie de l'information de l'image. Une bonne représentation des images est essentielle, que ce soit pour des problèmes de compression d'images, de caractérisation de singularités (chap. 6), ou de détection de régions d'intérêt (chap. 7). Nous présentons d'abord différentes décompositions multiéchelles, comme les transformées en ondelettes [3] et des décompositions spécifiques, telles que la transformée en curvelets [21]. A partir de celles-ci, il est possible de définir des maxima multiéchelles, ainsi que des lignes de maxima associées, qui constituent l'outil central de cette thèse.

Régularité Lipschitzienne en 2D

La notion de régularité Lipschitzienne (vue en 1D) peut être étendue en 2D. L'intérêt de cette notion de régularité est qu'en un point présentant une discontinuité, la valeur de la régularité Lipschitzienne $\alpha \in \mathbb{R}$ donne une information sur le type de singularité.

Définition 17. *(Régularité Lipschitzienne en 2D) Une fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ sera dite α-Lipschitz (pour $\alpha \in]0,1[$) en un point $v \in \mathbb{R}^2$ s'il existe une constante $K > 0$ telle que*

$$\forall u \in \mathbb{R}^2, |f(u) - f(v)| \leq K|u - v|_2^\alpha \tag{3.1}$$

Remarques :

1. La régularité locale en $v \in \mathbb{R}^2$ est définie comme : $\alpha = \sup\{\alpha_0 / f \ \ \alpha_0-\text{Lipchitz au point } v\}$
2. Cette notion peut être étendue à $\alpha \in \mathbb{R}$ [58] ; le cas $\alpha < 1$ correspond à f non-différentiable (singularités) ;
3. f sera dite uniformément Lipschitz sur un ouvert Ω s'il existe une constante $K > 0$ telle que f soit α-Lipschitz en tout $v \in \Omega$;
4. Cette notion de régularité ne privilégie pas de direction particulière. Cependant, dans le cadre 2D, il existe souvent des directions privilégiées. Par exemple, suivant des contours, il est intéressant de remarquer que l'intensité est régulière dans la direction tangentielle au contour et irrégulière dans la direction normale (cette dernière étant donnée par la direction du gradient). Ceci justifie alors l'emploi de méthodes qui cherchent d'abord certaines directions particulières, puis estiment la régularité suivant cette direction (problème monodimensionnel).

3.2 Décompositions multiéchelles

3.2.1 Ondelettes bidimensionnelles

Définition 18. *Une ondelette bidimensionnelle est une fonction $\Psi : \mathbb{R}^2 \to \mathbb{R}, \Psi \in L^2$ vérifiant la condition d'admissibilité*

$$\int_{(\mathbb{R}_+)^2} \frac{\left|\widehat{\Psi}(\omega)\right|^2}{|\omega|_2^2} d\omega < +\infty$$

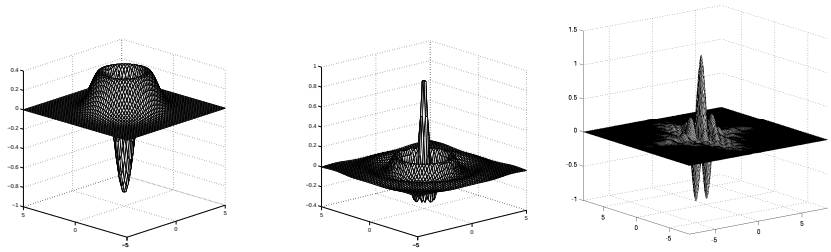


FIG. 3.1 – (A) Ondelette ΔG ; (B) Ondelette isotrope issue de l'ondelette de Meyer ; (C) Ondelette obtenue par produit tensoriel (de type $\psi(x)\phi(y)$, ψ ondelette, ϕ fonction échelle).

Afin de se donner une ondelette bidimensionnelle Ψ, il existe différentes options. En premier lieu, celle-ci peut être donnée explicitement par une fonction 2D $\Psi(x, y)$, telle que l'ondelette isotrope ΔG (Fig. 3.1 (A))

$$\Delta G(x, y) = \partial_{xx}G + \partial_{yy}G = \frac{1}{2\pi}\left(-2 + x^2 + y^2\right)\exp\left(-\frac{x^2 + y^2}{2}\right) \tag{3.2}$$

ou exprimée comme une fonction radiale $\psi(r)$, où ψ est l'ondelette de Meyer par exemple (Fig. 3.1 (B)) ; une alternative consiste à se placer dans le cadre d'une analyse multirésolution 2D, donnée par des produits tensoriels d'AMR 1D $(V_j \otimes V_j)_{j\in\mathbb{Z}}$. La propriété $V_{j+1} = V_j \oplus W_j$ permet d'écrire

$$V_{j+1} \otimes V_{j+1} = (W_j \otimes W_j) \oplus (V_{j+1} \otimes W_j) \oplus (W_j \otimes V_j) \oplus (V_j \otimes V_j)$$

Ceci conduit alors à des ondelettes 2D Ψ^H, Ψ^V et Ψ^D, ainsi qu'à une fonction échelle 2D Φ, exprimées comme produits tensoriels de fonctions ondelettes ψ et de fonctions échelles ϕ (Fig. 3.1 (C)) ;

$$\begin{cases} \Psi^H(x, y) = \psi(x)\phi(y) & < f, \Psi^H >\text{: coefficients de détail } c_H \text{ (bords verticaux)} \\ \Psi^V(x, y) = \phi(x)\psi(y) & < f, \Psi^V >\text{: coefficients de détail } c_V \text{ (bords horizontaux)} \\ \Psi^D(x, y) = \psi(x)\psi(y) & < f, \Psi^D >\text{: coefficients de détail } c_D \text{ (coins)} \\ \Phi(x, y) = \phi(x)\phi(y) & < f, \Phi >\text{: coefficients d'approximation} \end{cases}$$

Relevons que le cadre anisotrope, où certaines directions (plus ou moins nombreuses) sont privilégiées, peut se révéler intéressant : soit pour généraliser des résultats connus dans un cadre isotrope (chap.4), soit pour compléter une analyse isotrope (algorithme de *Shape Adaptation* [79]).

3.2.2 Coefficients d'ondelettes associés à une image

Définition 19. *(Transformée en Ondelette Continue 2D, TOC2D)*

Etant donné un signal $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ et une ondelette 2D $\Psi : \mathbb{R}^2 \longrightarrow \mathbb{R}$, la transformée en ondelette continue de f (normalisée L^p) est définie comme :

$$Wf(x, y, s) = \frac{1}{s^{2/p}} \int_{\mathbb{R}^2} f(u, v)\Psi\left(\frac{u - x}{s}, \frac{v - y}{s}\right) dudv \tag{3.3}$$

ce qui, en notant

$$\Theta_s(u, v) = \frac{1}{s^{2/p}}\Psi\left(-\frac{u}{s}, -\frac{v}{s}\right) \tag{3.4}$$

s'écrit également comme

$$Wf(x, y, s) = (f * \Theta_s)(x, y) \tag{3.5}$$

Remarque : dans la suite, nous utilisons une TOC normalisée L^1 $(p = 1)$.

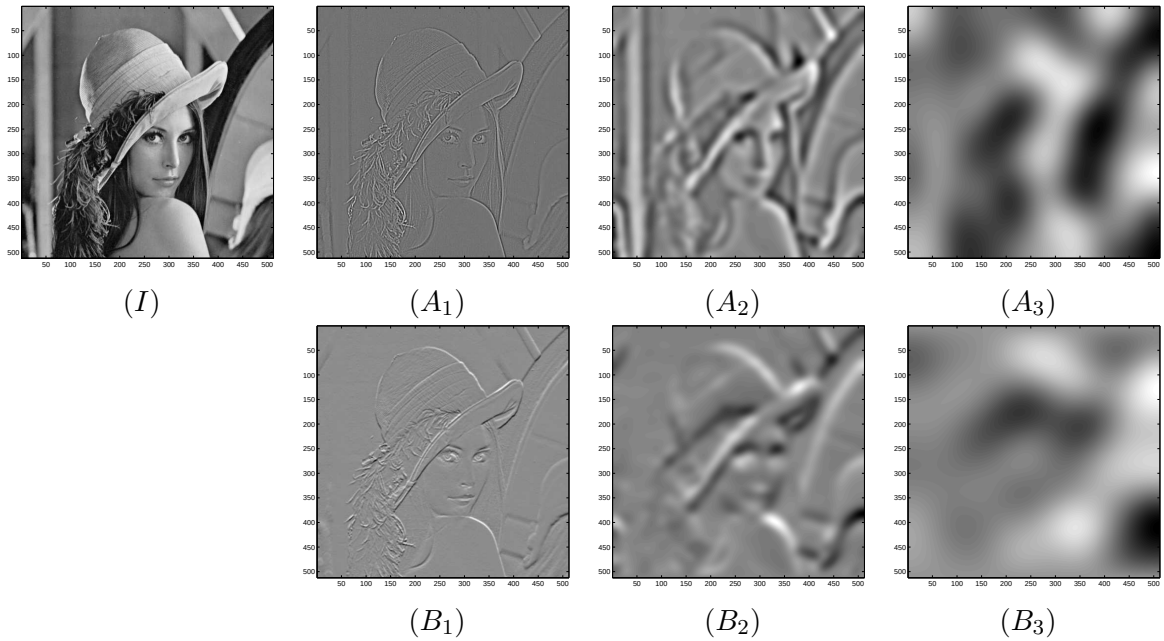


FIG. 3.2 – (I) Image originale, à partir de laquelle une TOC est calculée à une échelle s $(i = 1 : s = 1; i = 2 : s = 10; i = 3 : s = 50)$; (A_i) Ondelette ΔG (isotrope) ; (B_i) Ondelette anisotrope (mettant en valeur les bords horizontaux).

Calcul effectif

Nous étendons en 2D la méthode 1D vue en 1.2.3. Dans le cadre où f et Θ sont dans L^2, nous pouvons écrire la relation suivante :

$$\mathcal{W}f(\xi_1, \xi_2, s) = \overline{\mathcal{F}}\left(\widehat{f}(\xi_1, \xi_2) \cdot \widehat{\Theta_s}(\xi_1, \xi_2)\right) \tag{3.6}$$

où ξ_1 et ξ_2 les variables duales de u et v.

Notons alors que si la transformée de Fourier de l'ondelette Ψ possède une expression analytique, alors il est possible de calculer directement les valeurs de $\widehat{\Theta}$. Par ailleurs, étant donné que la transformée de Fourier discrète donne une approximation de la transformée de Fourier, des valeurs approchées de $\widehat{f}$ peuvent être obtenues. Ainsi, par produit de ces deux termes puis en effectuant une transformée de Fourier inverse, une approximation de la TOC de f est calculable pour toute échelle $s > 0$ (non nécessairement dyadique ou entière). Des exemples d'application d'une telle méthode sont donnés sur la figure 3.2. En notant N la taille des données, le coût algorithmique de la TOC à une échelle s donnée revient à $O(N \log N)$, la FFT étant l'opération la plus coûteuse (ce coût est indépendant de l'échelle s).

En résumé, nous pouvons calculer $Wf(.,.,s)$ de la manière suivante (comme en 1.2.3) :

1. Calcul exact de $\widehat{\Theta_s}$;
2. Calcul approché de $\widehat{f}$ par transformée de Fourier discrète (FFT) ;
3. Calcul de $Wf(.,.,s)$ par transformée de Fourier discrète inverse (Inverse-FFT) du produit des deux quantités précédentes.

Conditions au bord

L'image considérée I (d'intensité f) étant donnée sur un domaine borné, les conditions de bord sont primordiales. En particulier, l'utilisation d'une FFT dans le calcul de la TOC implique une périodisation implicite, et en calculant la TOC directement sur l'image I, cela conduit à des discontinuités au bord. Afin d'éviter ceci, nous utilisons une condition miroir : l'image originale I (de taille $n \times n$) est symétrisée suivant les deux directions, ce qui conduit à une image I' (de taille $2n \times 2n$). Ainsi nous calculons d'abord la TOC de l'image I', puis la restriction de celle-ci au domaine initial donne la TOC de l'image originale. En pratique, pour la construction de lignes de maxima (décrite plus loin dans ce chapitre), le calcul de la TOC n'est pas l'étape la plus coûteuse, et ainsi le surcoût en calcul est limité.

Remarque sur les images non carrées

Cette méthode permet de calculer le produit de convolution avec un noyau isotrope pour une image carrée (de taille $n \times n$). Dans le cas d'une image de taille $n_1 \times n_2$, si nous souhaitons conserver ce caractère isotrope, il convient de l'étendre à une image de taille $\max(n_1, n_2) \times \max(n_1, n_2)$ (en symétrisant une partie de l'image), puis d'appliquer la méthode précédente.

3.2.3 Transformées adaptées à la géométrie

L'information directionnelle est une caractéristique propre aux signaux multidimensionnels, et apparaît comme importante dans nombre d'applications en traitement d'images. La transformée en ondelettes bidimensionnelle séparable permet de sélectionner des informations suivant des directions privilégiées (les variables x et y sont traitées indépendamment, par produit tensoriel). Les directions associées sont l'horizontale, la verticale et les diagonales (ces dernières étant mélangées). Diverses approches ont été proposées, afin de mieux capter cette information directionnelle [1, 16, 40].

Rappelons que les données numériques associées à une image I sont des valeurs correspondant à des points d'une grille régulière fixée. Si nous effectuons une rotation sur I (d'un angle de rotation non multiple de $\pi/2$), en notant I' l'image transformée, il est important de noter que les valeurs associées aux différents pixels des images I et I' ne peuvent pas être directement mises en correspondance (il y a une interpolation sous-jacente). De manière plus générale, une transformation géométrique altère le contenu de l'image, tout en conservant

certaines propriétés. Il importe alors de déterminer des caractéristiques issues de l'image qui soient invariantes par une transformation donnée (cette notion d'invariance sera étudiée aux chapitres 4 et 6). Deux approches sont alors envisageables :

- soit utiliser une transformée invariante par le type de transformation considéré ; par exemple, l'ondelette ΔG est invariante par rotation ;
- soit estimer une déformation, et la prendre en compte par la suite (chap. 7).

Transformée en ondelette directionnelle – Invariance par rotation

Définition 20. *Soit Ψ une ondelette bidimensionnelle et $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ telle que $f \in L^2$. La transformée en ondelette directionnelle est définie comme :* $\forall s > 0, \forall b \in \mathbb{R}^2, \forall \theta \in [0, 2\pi[$,

$$\mathcal{T}f(s,b,\theta) = \int_{\mathbb{R}^2} \frac{1}{s^2}\Psi\left(R_{-\theta}\left(\frac{x-b}{s}\right)\right) f(x)dx \tag{3.7}$$

La transformée en ondelette directionnelle d'un signal bidimensionnel f reprend les paramètres d'espace et d'échelle présents dans la TOC, et ajoute un paramètre d'angle. Cette transformée a été appliquée avec succès à des problèmes d'imagerie médicale [105], où le signal à analyser est donné directement en fonction du paramètre d'angle (données tomographiques provenant d'un scanner). Cependant, lorsque ce n'est pas le cas, une implémentation directe est délicate [103]. Il est néanmoins possible de définir des bancs de filtres directionnels, utilisant certains angles particuliers [40, 106].

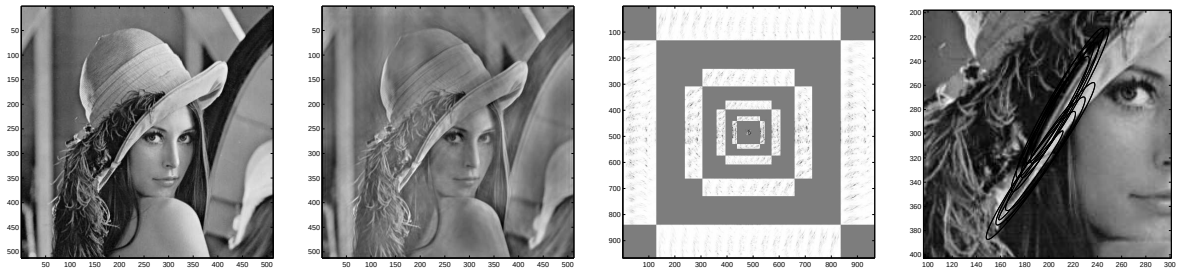
Par ailleurs, étant donné deux images L et R reliées par $f_L(x) = f_R(R_\alpha x)$, où R_α est une rotation d'angle α, notons que

$$\begin{aligned}
\mathcal{T}f_L(s,b,\theta) &= \int_{\mathbb{R}^2} \frac{1}{s^2}\Psi\left(R_{-\theta}\left(\frac{x-b}{s}\right)\right) f_R(R_\alpha x)dx \\
&= \int_{\mathbb{R}^2} \frac{1}{s^2}\Psi\left(R_{-\theta}\left(\frac{R_{-\alpha}y-b}{s}\right)\right) f_R(y)dy (y = R_\alpha x, dy = dx) \\
&= \int_{\mathbb{R}^2} \frac{1}{s^2}\Psi\left(R_{-(\theta+\alpha)}\left(\frac{y-R_\alpha b}{s}\right)\right) f_R(y)dy \\
&= \mathcal{T}f_R(s, R_\alpha b, \theta+\alpha)
\end{aligned}$$

d'où il vient

$$\mathcal{T}f_L(s,b,\theta) = \mathcal{T}f_R(s, R_\alpha b, \theta+\alpha) \tag{3.8}$$

Ainsi, dans le cas où f_L et f_R sont reliées par une rotation, l'équation (3.8) permet de relier $\mathcal{T}f_L$ et $\mathcal{T}f_R$. Ce type de propriété – décrivant l'influence d'une certaine transformation géométrique – sera développée au chapitre 4 et constitue un point important dans la détection de régions d'intérêt, en particulier pour la recherche de caractéristiques invariantes à certaines transformations de l'image [75] (chap. 4–7).



(A) (B) (C) (D)

FIG. 3.3 – Exemple de décomposition en curvelets : (A) Image classique *Lena* (512×512) ; (B) Image compressée, obtenue en conservant 5% des plus forts coefficients de curvelets ; (C) Décomposition obtenue par l'algorithme FDCT (appproximativement $2 \cdot 10^6$ coefficients) ; (D) Pseudo-supports associés à quelques curvelets dont les coefficients sont les plus forts en module.

Curvelets – Transformée en curvelets

Introduites par E. Candès et D. Donoho [22, 25], les curvelets sont des fonctions anisotropes, multiéchelles et multiorientations. Elles vérifient une propriété de *frame* (def. 3, chap. 1) et permettent d'obtenir des résultats d'approximations meilleurs que ceux connus pour les ondelettes [18, 21]. De plus, il existe un algorithme rapide de calcul des coefficients de curvelets associés à une image [19] (*Fast Curvelet Transform*, http://www.curvelab.org). Cette décomposition en curvelets (Fig. 3.3 (C)) se démarque des décompositions pyramidales classiques [11, 16, 115], à la fois par le caractère anisotrope des fonctions (largeur $w = 2^{-j}$, longueur $l = 2^{-j/2}$, $w = l^2$) et par la résolution angulaire qui augmente avec l'échelle. Schématiquement, la construction des curvelets, détaillée dans [22, 25], consiste à définir

- $\varphi(x), x = (x_1, x_2) \in \mathbb{R}^2$: fonction oscillante suivant x_1, noyau de lissage suivant x_2 ;
- $D_j = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}$ $(j \geq 0)$: matrice de dilatation ;
- $R_{\theta_{jl}}, \theta_{jl} = 2\pi \cdot l 2^{-\lfloor (j/2) \rfloor}$ ($\lfloor . \rfloor$: partie entière) : matrice de rotation

puis à appliquer successivement à $\varphi(x)$: une dilatation, $2^{3j/4}\varphi(D_j x)$; une rotation, $2^{3j/4}\varphi(D_j R_{\theta_{jl}})$; une translation, $2^{3j/4}\varphi(D_j R_{\theta_{jl}})$. Les applications des curvelets se trouvent notamment en compression, débruitage et restauration d'image [117, 116] (Fig. 3.3 (B)). Le succès des curvelets vient notamment du fait que celles associées aux plus forts coefficients sont localisées sur les contours (Fig. 3.3 (D)). Nous verrons également qu'elles sont un bon outil pour l'analyse de singularités, grâce à l'estimation de la régularité Lipschitzienne ponctuelle (annexe B).

Maxima multiéchelles associés à une image

Le cadre 2D conduit à différents types de maxima associés à une image, en particulier les modules maxima et les maxima au sens de Canny que nous présentons ici. Les décompositions

en ondelettes et leurs extensions permettent d'établir des cartes de maxima multiéchelles : ainsi, en posant telle ou telle définition d'un maximum, à chaque échelle est associée une certaine carte, indiquant s'il y a ou non un maximum, associée à un certain coefficient réel ou complexe (coefficient d'ondelette par exemple). Après le rappel de ces notions de maxima, nous présentons quelques résultats sur l'existence de lignes de maxima associées. Ces résultats utilisent notamment le principe du maximum pour l'équation de la chaleur.

3.3 Modules maxima et Maxima Lines

3.3.1 Définitions

Les notions de MM et de ML, vues en 1D (définitions 12 et 13) s'étendent en 2D :

Définition 21. *(Module Maximum, MM [89]) Etant donné un signal bidimensionnel f auquel est associée une décomposition en ondelette $Wf(.,.,s)_{s>0}$, un point (x_0, y_0) est un module maximum à l'échelle s si $|Wf(.,.,s)|$ admet un maximum local en (x_0, y_0) (ce maximum étant strict au moins dans une direction).*

Définition 22. *(Maxima Line, ML) Une Maxima Line (ML) est une courbe connexe du plan espace-échelle (x, y, s) formée de MM. Une telle ML sera donnée par*

$$(x(s), y(s), s, Wf(u(s), s)), s \in I \text{ intervalle de } \mathbb{R}_+. \tag{3.9}$$

3.3.2 Equation de la chaleur et principe du maximum

L'équation de la chaleur est l'exemple fondamental d'une diffusion. Sur un domaine Ω (ouvert de $\mathbb{R}^n$), elle s'écrit comme

$$\partial_t u = \Delta u \qquad \forall x \in \Omega, \forall t \in \mathbb{R}_+^* \tag{3.10}$$

$$u(x, 0) = u_0(x) \qquad \forall x \in \Omega \tag{3.11}$$

Pour $\Omega = \mathbb{R}^n$ et $u_0 \in L^2$, cette équation peut être résolue explicitement – par résolution dans le domaine de Fourier. La solution $u(x, t)$ s'exprime sous la forme :

$$u(x, t) = \int_{\mathbb{R}^n} G(x - y, t) u_0(y) dy \tag{3.12}$$

où G est la fonction définie par

$$G(z, t) = \frac{1}{(4\pi t)^{n/2}} \exp\left(-\frac{|z|^2}{4t}\right) \tag{3.13}$$

et qui vérifie $\forall z \in \mathbb{R}^n, \forall t > 0, G(z, t) > 0$, $\int_{\mathbb{R}^n} G(z, t) dz = 1$. Ainsi, la solution de l'équation de la chaleur peut s'écrire comme la convolution entre la solution initiale et un noyau strictement positif de moyenne égale à 1 (voir [108] pour les propriétés de telles convolutions).

Proposition 9. *(Principe du maximum) Etant donné une condition initiale $u_0 \in L^2(\Omega)$ et u la solution de l'équation de la chaleur donnée par 3.10, nous avons l'inégalité suivante :*

$$Min\{0, \inf_{\Omega} u_0\} \leq u(x,t) \leq Max\{0, \sup_{\Omega} u_0\} \qquad \forall (x,t) \in \Omega \times]0, +\infty[\tag{3.14}$$

Ce résultat assure que la solution est bornée par la donnée initiale, et cela induit certaines propriétés sur les extrema de la solution. Il est intéressant de noter qu'il existe un résultat de principe du maximum pour certaines équations d'évolution plus complexes [108].

Equation de la chaleur semi-linéaire

L'équation de la chaleur semi-linéaire (en dimension $n = 1$ ou $n = 2$) est définie par

$$\begin{array}{rcll} \frac{\partial u}{\partial t}(x,t) & = & \Delta u(x,t) + f(u(x,t)) & x \in \Omega, t > 0 \\ u(x,t) & = & 0 & x \in \partial\Omega, t > 0 \\ u(x,0) & = & u_0(x) & x \in \Omega \end{array} \tag{3.15}$$

où Ω est un ouvert borné et $f : \mathbb{R} \longrightarrow \mathbb{R}$.

Suivant les hypothèses faites sur f, il existe des résultats similaires à ceux existants pour l'équation de la chaleur classique 3.10 (existence d'une solution, principe du maximum) [60, 108]. Ainsi, l'équation de la chaleur présente une certaine stabilité, au sens où si elle est légèrement perturbée la solution conserve des propriétés de l'équation classique.

3.3.3 Existence de Maxima Lines pour l'ondelette ΔG

Nous avons généralisé la démonstration de l'existence des ML (vue en 1D, pour les ondelettes dérivées n-ièmes de Gaussienne) au cadre 2D, dans le cadre de l'ondelette ΔG (eq. 3.2).

Ondelette ΔG et équation de la chaleur

Etant donné $t > 0$, nous définissons $G_t : \mathbb{R}^2 \to \mathbb{R}$ comme

$$G_t(x,y) = \frac{1}{2\pi t} \exp\left(-\frac{x^2+y^2}{2t}\right) \tag{3.16}$$

Notons alors que $u(x,\tau) = u_0 * G_t$ est solution de l'équation de la chaleur :

$$\partial_\tau u(x,\tau) = \frac{1}{2}\Delta u(x,\tau) \quad \text{avec } u(x,0) = u_0 \tag{3.17}$$

Par ailleurs, d'après les propriétés de la convolution et de l'ondelette ΔG, nous avons

$$t\Delta f * G_t = f * (\Delta G)_t = Wf(\cdot, \cdot, \sqrt{t}) \tag{3.18}$$

Existence de ML

Proposition 10. *Soit $f \in L^2$ et Wf la TOC (normalisée L^1) utilisant l'ondelette Δg. Alors les MM de Wf forment des courbes ininterrompues vers les échelles fines.*

Démonstration

1. Dans le cadre de l'ondelette ΔG, la TOC $Wf(x, y, s)$ est proportionnelle à la solution de l'équation de la chaleur (en 2D), avec pour condition initiale Δf (cette dérivée étant prise au sens des distributions).
2. En notant g cette solution, et par application du principe du maximum à l'équation de la chaleur, un maximum global de $|g(x, y, s)|$ pour $(x, y, s) \in [a, b[\times[c, d[\times[s_0, s_1[$ ne peut être atteint que sur la frontière du domaine (x égal à a ou b, y égal à c ou d). ou à l'instant initial (en $s = s_0$) [14, 108, 126].
3. Raisonnement par l'absurde : supposons qu'une courbe de MM s'interrompe en un point $(x_1, y_1, s_1), s_1 > 0$ (cette interruption ayant lieu lorsque l'échelle diminue).
   Alors il existe $\varepsilon > 0$ tel que $|g(x, y, s)|$ soit un maximum global dans le voisinage $[x_1-\varepsilon, x_1+\varepsilon]\times[y_1-\varepsilon, y_1+\varepsilon]\times[s_1-\varepsilon, s_1+\varepsilon]$: d'une part $|g(., ., s_1)|$ est localement maximum en espace (MM) et d'autre part il n'y a pas de MM au voisinage de (x_1, y_1), aux échelles immédiatement inférieures (interruption de la courbe de MM). Ceci contredit le principe du maximum.
4. Conclusion : toutes les courbes de maxima se prolongent jusqu'aux échelles les plus fines.

3.4 Maxima au sens de Canny

3.4.1 Définitions

Les maxima au sens de Canny sont définis à partir du gradient de l'image [26], et plus précisément par son module et son orientation, qui s'expriment comme

$$\begin{cases} \text{Gradient de l'image} & \nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} f_x \\ f_y \end{pmatrix} \\ \text{Module du gradient} & |\nabla f| = \sqrt{f_x^2 + f_y^2} \\ \text{Direction du gradient} & \text{donnée par l'angle } (0x, \nabla f) \end{cases}$$

Si f n'est pas suffisamment régulière, le gradient de l'image ∇f peut être néanmoins défini à l'aide de la théorie des distributions [113]. En pratique, l'image est d'abord convoluée avec un noyau régularisant, le gradient étant alors bien défini.

Définition 23. *(Maxima au sens de Canny, MC [26]) Un point (x, y) est appelé maximum au sens de Canny si, dans la direction du gradient, le module du gradient est localement maximum.*

L'intérêt de ces maxima est que les plus significatifs d'entre eux sont localisés sur des contours de l'image (d'autres MC pouvant correspondre à du bruit). Une fois ces maxima détectés, il est possible d'estimer les contours en les regroupant de manière judicieuse [69]. Un détecteur efficace doit localiser précisément ces contours tout en étant robuste au bruit. A cet effet, l'algorithme suivant a été proposé [26, 39] :

1. Lissage de l'image par un noyau Gaussien :

$$L(x,y,s) = (f * G_s)(x,y) \quad \text{où } G_s(x,y) = \frac{1}{2\pi s^2} \exp\left(-\frac{x^2+y^2}{2s^2}\right)$$

2. Calcul du gradient de l'image lissée ∇L, conduisant à son module $M = |\nabla L|_2$ et sa direction $A = (Ox, \nabla L)$;
3. Calcul des maxima de Canny à partir de M et A ;
4. Seuillage des maxima les plus significatifs, d'après le module M.

3.4.2 Formulation multiéchelle utilisant une ondelette gradient

Définition 24. *(Ondelette gradient) Etant donné un noyau de lissage $\Lambda : \mathbb{R}^2 \longrightarrow \mathbb{R}$, une ondelette gradient en 2D est définie comme*

$$\Psi = \nabla\Lambda = \begin{pmatrix} \frac{\partial\Lambda}{\partial x} \\ \frac{\partial\Lambda}{\partial y} \end{pmatrix} = \begin{pmatrix} \psi_x \\ \psi_y \end{pmatrix} \tag{3.19}$$

Alors la transformée en ondelette (associée à Ψ, appliquée à f) s'écrit comme

$$Wf = \begin{pmatrix} W_x f \\ W_y f \end{pmatrix} \tag{3.20}$$

Dans le cas où le noyau régularisant est une Gaussienne 2D, il s'exprime comme le produit tensoriel de deux Gaussiennes 1D, ce qui simplifie le calcul de $W_x f$ et $W_y f$. En particulier, comme $\widehat{\psi_x}$ et $\widehat{\psi_y}$ admettent une expression analytique connue, nous pouvons utiliser la méthode de calcul présentée dans la section 3.2.2. En posant $\Lambda_s = \frac{1}{s^2}\Lambda(-\frac{\cdot}{s}, -\frac{\cdot}{s})$, nous avons

$$\begin{aligned} W_x f &= -s\partial_x(f * \Lambda_s) \\ W_y f &= -s\partial_y(f * \Lambda_s) \end{aligned}$$

d'où il ressort

$$Wf = -s\nabla L \tag{3.21}$$

3.4.3 Notion de ligne de Canny

Des résultats pratiques montrent que certains maxima de Canny persistent d'une échelle à une autre (nous verrons cela au chapitre 5). Cette notion de persistance conduit alors au problème d'existence de lignes de maxima, au sens où certaines lignes (telles que les ML vues précédemment) ne s'interrompent pas lorsque l'échelle diminue.

Définition 25. *(Ligne de Canny, LC) Une ligne de Canny est une courbe connexe du plan espace-échelle formée de MC* (x, y, s)*. Une telle LC sera donnée par*

$$(x(s), y(s), s, Mf(x(s), y(s), s), Af(x(s), y(s), s)), s \in I \text{ intervalle de } \mathbb{R}_+. \tag{3.22}$$

A notre connaissance, il n'existe pas de résultats théoriques sur l'existence de lignes associées aux maxima au sens Canny dans un cadre général. Notons cependant que le problème se simplifie dans le cadre où la direction du gradient est constante (indépendante de s) : en notant $r = \nabla(f * \Lambda_s)$, et en se plaçant dans le repère $(r, r^{\perp})$, nous sommes alors ramenés à un cadre monodimensionnel. De plus, dans le cas où le noyau Λ est Gaussien, nous pouvons faire l'analogie entre un maxima au sens de Canny et un MM associé à une TOC 1D appliquée dans la direction du gradient (constante) utilisant une dérivée première de Gaussienne 1D : d'une part, un MC correspond à un maximum du module d'un gradient de Gaussienne (direction fixée) et d'autre part, un MM correspond à un maximum du module d'une dérivée de Gaussienne (l'utilisation d'ondelettes dérivées n-ièmes de Gaussienne assurant l'existence de ML).

Conclusion

Dans cette thèse, nous nous intéressons plus particulièrement aux lignes de maxima présentées ici (ML et LC), dont la construction algorithmique sera détaillée au chapitre 5. Au chapitre suivant nous présentons des liens entre les ondelettes et la théorie du *Scale-Space*. Celle-ci met l'accent sur l'association entre certains maxima dans l'espace-échelle et des structures présentes dans l'image (objets apparaissant visuellement). Ces liens justifient alors l'utilisation des ondelettes pour la détection de telles structures, ce qui sera l'objet des chapitres 6 et 7.

Chapitre 4

Théorie du Scale-Space
Sélection de points et d'échelles remarquables

Introduction

Nous présentons dans ce chapitre des éléments de la théorie du Scale-Space, issue de la vision par ordinateur. En particulier, nous expliquons comment sélectionner des points et des échelles remarquables, par utilisation d'opérateurs adaptés et de maxima dans le *Scale-Space* (approche classique [75]). L'intérêt de ces maxima est qu'il peuvent être associés à des régions de l'image jugées particulièrement importantes, ce qui conduit à des détecteurs de telles régions (ceci sera détaillé au chap. 7).

Nous aborderons également l'effet de déformations géométriques sur le *Scale-Space*, comme un changement d'échelle ou une déformation affine. La prise en compte de ces déformations est cruciale dans les applications, où souvent celle-ci n'est pas connue *a priori*. Par ailleurs, nous montrerons d'une part les généralisations que permettent la théorie des ondelettes, et, d'autre part, le lien entre celles-ci et le *Scale-Space*, en particulier du point de vue des Maxima Lines.

4.1 Eléments de la théorie du Scale-Space

La théorie du Scale-Space permet de donner une représentation multiéchelles d'une image f, notée $L(x,t)$, $x \in \mathbb{R}^d$ étant un paramètre spatial et $t > 0$ un paramètre d'échelle. Un problème important consiste alors à savoir comment L est modifié lorsque lorsque f subit une modification telle qu'un changement d'échelle, une déformation géométrique, un changement d'intensité ou de contraste. Dans cette section, nous donnons des éléments de cette théorie, proposée par Lindeberg [75–77].

4.1.1 Représentation espace-échelle

Cadre isotrope

Définition 26. *(Représentation espace-échelle linéaire [75]) Etant donné un signal $f : \mathbb{R}^d \longrightarrow \mathbb{R}$, sa représentation espace-échelle linéaire est définie par : $\forall x \in \mathbb{R}^d, \forall t > 0$,*

$$L(x,t) = \int_{\xi \in \mathbb{R}^d} f(x-\xi) G(\xi,t) d\xi \tag{4.1}$$

où G est le noyau Gaussien donné par

$$\begin{aligned} G &: \mathbb{R}^d \times \mathbb{R}_+^* \to \mathbb{R} \\ G(x,t) &= \left(\frac{1}{\sqrt{2\pi t}}\right)^d \exp\left(-\frac{|x|_2^2}{2t}\right) \end{aligned} \tag{4.2}$$

$$\tag{4.3}$$

L'ensemble $\{(x,t) \in \mathbb{R}^d \times \mathbb{R}_+^*\}$ est appelé espace-échelle ou *Scale-Space*.

Cadre anisotrope (2D)

Définition 27. *(Représentation espace-échelle affine) Nous considérons une famille $(\Sigma_t)_{t>0}$ de matrices symétriques définies positives, et le noyau 2D défini par : $\forall x \in \mathbb{R}^2, \forall t > 0$,*

$$G(x, \Sigma_t) = \frac{1}{2\pi\sqrt{\det(\Sigma_t)}} \exp(-\frac{1}{2} x^T \Sigma_t^{-1} x) \quad \textit{(noyau Gaussien affine)} \tag{4.4}$$

Alors la représentation espace-échelle affine de $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ est donnée par

$$L(x, \Sigma_t) = (f * G(\cdot, \Sigma_t))(x) \tag{4.5}$$

Notons que si la famille $(\Sigma_t)_{t>0}$ est quelconque, il est difficile de décrire les propriétés de L. Néanmoins, dans le cas où $\Sigma_t = t\Sigma_0$, où Σ_0 est une matrice symétrique définie positive, L est solution du problème [79] :

$$\partial_t L = \frac{1}{2} \mathrm{div}(\Sigma_0 \nabla L) \text{ avec } L(x,0) = f(x) \tag{4.6}$$

Ce problème généralise celui de l'équation de la chaleur (cas où $\Sigma_0 = Id$), et représente un processus de diffusion anisotrope. De plus, la matrice Σ_0 étant symétrique définie positive, il s'agit d'un problème de type parabolique, ce qui induit notamment que le principe du maximum est vérifié (prop. 9, section 3.3.2). Enfin, notons que Σ_0 est indépendant de L, l'opérateur de diffusion est global.

4.1.2 Notion de dérivée γ-normalisée

Définition 28. *[77] Pour $\gamma > 0$, la dérivée γ-normalisée par rapport à une variable x est définie comme :*

$$\partial_\xi = t^{\gamma/2}\partial_x, \text{ où } t \text{ est le paramètre d'échelle} \tag{4.7}$$

Ainsi un opérateur impliquant des dérivées (en dimension finie quelconque) peut être normalisé en remplaçant les dérivées partielles classiques par cette nouvelle dérivée. Par exemple, une dérivée ∂_x^m sera remplacée par une dérivée $\partial_\xi^m = t^{m\gamma/2}\partial_x^m$.

Plus généralement, dans $\mathbb{R}^d$, nous posons

- α_i l'ordre de la dérivée dans la direction i $(1 \leq i \leq d)$,
- $\alpha = (\alpha_1, \alpha_2, ..., \alpha_d)$, $|\alpha| = \sum_{i=1}^{d} \alpha_i$,
- $L_{x^\alpha} = \partial_\alpha L = \partial_{\alpha_1}\partial_{\alpha_2}...\partial_{\alpha_d} L$,

et nous considérons l'opérateur

$$\mathcal{D} = \sum_{k=1}^{K} c_k L_{x^{\alpha_k}} \tag{4.8}$$

Alors, dans le cas où les ordres de dérivation vérifient $\forall k \in \{1, ..., K\}, |\alpha_k| = m$ (constante indépendante de k), l'opérateur normalisé s'écrit comme :

$$\mathcal{D}_{norm} = t^{m\gamma/2}\mathcal{D} \tag{4.9}$$

Ce terme multiplicateur $t^{\gamma/2}$ correspond au changement de variable $\xi = \frac{x}{t^{\gamma/2}}$, et la notion de dérivée γ-normalisée revient à une normalisation L^p, avec :

$$p = \frac{1}{1 + \frac{m}{d}(1-\gamma)} \quad (d: \text{ dimension de l'espace })$$

En particulier, pour $\gamma = 1$, nous obtenons une normalisation L^1, et ce indépendamment de la dimension d et de l'ordre des dérivées considérées [77].

Intérêt de la γ-normalisation : étude du cas sinusoïdal pur [77]

Afin d'étudier comment se comporte la fréquence d'un signal dans le Scale-Space, nous considérons la fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$ définie par $f(x) = \sin(\omega x)$.

Nous adoptons les notations suivantes :

- $L_{\max}$: amplitude de L
- L_{ξ^m} : dérivée m-ième de L (γ-normalisée)

$$L_{\xi^m} = \partial_\xi^m L = t^{m\gamma/2}\partial_x^m L \tag{4.10}$$

- $L_{\xi^m,\max}$: amplitude de L_{ξ^m}

Considérons la représentation espace-échelle linéaire $L = g_t * f$. Celle-ci vérifie l'équation de la chaleur

$$\frac{\partial L}{\partial t} = \Delta L = g_t * f'' = g_t * (-\omega^2 f) = -\omega^2 L$$

D'où il vient :

$$\begin{aligned} L(x,t) &= e^{-\omega^2 t/2} \sin(\omega x) \\ L_{\max}(t) &= e^{-\omega^2 t/2} \\ L_{\xi^m,\max}(t) &= t^{m\gamma/2} \omega^m e^{-\omega^2 t/2} \end{aligned}$$

Notons que cette quantité positive $L_{\xi^m,\max}$ croît puis décroît, atteignant un unique maximum en $t^* = \frac{\gamma m}{\omega^2}$. En notant $s = \sqrt{t}$ l'échelle (associée à la représentation L) et $\lambda = \frac{2\pi}{\omega}$ la longueur d'onde (associée au signal sinusoïdale), il est intéressant de remarquer que

$$s^* = \frac{\sqrt{\gamma m}}{2\pi} \lambda \tag{4.11}$$

ce qui signifie que l'échelle à laquelle l'amplitude de la dérivée γ-normalisée atteint son maximum est proportionnelle à la longueur d'onde du signal considéré. De plus, la valeur maximale de $L_{\xi^m,\max}$ atteinte en t^* vaut

$$L_{\xi^m,\max}(t^*) = \frac{(\gamma m)^{\gamma m/2}}{e^{\gamma m/2}} \omega^{(1-\gamma)m} \tag{4.12}$$

et cette amplitude varie différemment en fonction de m, ω et γ (Fig. 4.1.2). Relevons que pour $\gamma = 1$, cette valeur est indépendante de la fréquence du signal. Ce point constitue la base de l'invariance à l'échelle. Nous verrons ultérieurement qu'étant donné un opérateur détectant un certain type d'objets, il est possible de le normaliser de manière à ce que l'amplitude de ses maxima soit caractéristique de la taille des objets.

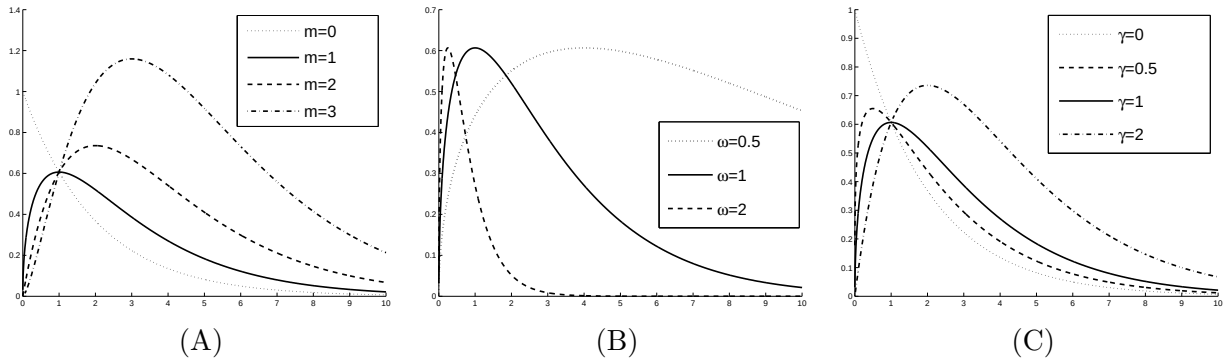


FIG. 4.1 – Evolution de $L_{\xi^m,\max}$: (A) en fonction de m ($\omega = 1$, $\gamma = 1$) ; (B) en fonction de ω ($m = 1$, $\gamma = 1$) ; (C) en fonction de γ ($m = 1$, $\omega = 1$).

4.1.3 Application de l'invariance à l'échelle : spectre de puissance

Etant donné une image $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dont la représentation échelle-espace est donnée par $L : \mathbb{R}^2 \times \mathbb{R} \longrightarrow \mathbb{R}$, il est possible de définir différentes mesures d'énergie (pour $t \in \mathbb{R}_+^*$) [77] :

$$\begin{cases} E_0 &= \int_{x\in\mathbb{R}^2} L(x,t)^2 dx \\ E_1 &= \int_{x\in\mathbb{R}^2} t^\gamma L_{x_1}^2 L_{x_2}^2 dx \\ E_2 &= \int_{x\in\mathbb{R}^2} t^{2\gamma} \left(L_{x_1x_1}^2 + 2L_{x_1x_2}^2 + L_{x_2x_2}^2\right) dx \end{cases}$$

et plus généralement, en dimension d, et pour $m \in \mathbb{N}^*$:

$$E_m = \int_{x\in\mathbb{R}^d} \sum_{|\alpha|=m} t^{m\gamma} L_{x^\alpha}^2 dx \tag{4.13}$$

Considérons maintenant un signal d-dimensionnel, dont le spectre de puissance est donné par

$$S_f(\omega) = (\hat{f}\hat{f}^*)(\omega) \simeq |\omega|^{-2\beta} \quad \text{appelé signal en } 1/f$$

Etant donné $\beta > 0$, il est possible de simuler des bruits Gaussiens correspondant à un tel spectre. Pour des images naturelles ($d = 2$), des études ont montré qu'en moyenne, ce modèle de puissance spectrale était valide, la valeur de β étant proche de 1 [17]. Il est alors intéressant de noter que, pour m fixé, le terme d'énergie E_m varie en fonction de t comme [77]

$$E_m \sim t^{\beta - d/2 - m(1-\gamma)} \tag{4.14}$$

de sorte que cette expression ne dépend plus de l'échelle si et seulement si $\beta = d/2 - m(1-\gamma)$ (en particulier, dans le cas $d = 2$ et $\beta = 1$, nous retrouvons la valeur $\gamma = 1$). Ainsi, pour des signaux dont le spectre de puissance est de la forme $S_f(\omega) = |\omega|^{-(d+2m(1-\gamma))}$, une fois E_m correctement normalisé, il ne dépend plus de l'échelle.

4.1.4 Normalisation du gradient et de la Hessienne

Dans le cadre 2D, le vecteur gradient ∇f et la matrice Hessienne $\mathcal{H}$ sont deux primitives permettant de formuler des opérateurs intéressants en vision par ordinateur. Rappelons que si $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ est suffisamment régulière en un point $x \in \mathbb{R}^2$, nous pouvons écrire le développement de Taylor (pour $\delta x \in \mathbb{R}^2$)

$$f(x + \delta x) = f(x) + \delta x^T \nabla f(x) + \frac{1}{2}\delta x^T \mathcal{H}(x)\delta x + o(|\delta x|^2)$$

ce qui justifie leur utilisation. De plus, afin d'obtenir des propriétés d'invariance à l'échelle, il est intéressant d'utiliser leurs versions γ-normalisées, comme nous le verrons par la suite.

Définition 29. *(Gradient γ-normalisé, Hessienne γ-normalisée [76]) Etant donné une représentation L, son gradient normalisé est défini comme*

$$\nabla L_{norm} = t^{\gamma/2} \nabla L \tag{4.15}$$

et sa Hessienne normalisée comme

$$\mathcal{H}_{norm} = t^{\gamma} \begin{pmatrix} L_{xx} & L_{xy} \\ L_{xy} & L_{yy} \end{pmatrix} \tag{4.16}$$

4.2 Approche classique pour la détection de caractéristiques locales

Nous donnons ici des exemples de détecteurs de caractéristiques locales proposés par Lindeberg [77]. La démarche adoptée consiste généralement à poser d'abord un opérateur adapté à la caractéristique, puis à écrire sa version γ-normalisée ; enfin un critère détermine si telle caractéristique est présente en tel point du Scale-Space.

4.2.1 Opérateurs en Vision par Ordinateur

Nous avons présenté auparavant (chap. 3) comment détecter les maxima au sens de Canny, qui permettent de mettre en évidence des points de contours. Ce qui est présenté ici correspond à un raffinement, de manière à extraire des caractéristiques plus précises (bord, ligne, coin) [76, 77]. Nous verrons ultérieurement (chap. 6) que ceci contraste avec l'approche multiéchelles de Canny (formulée par ondelettes), qui permet de détecter simultanément ces caractéristiques, et également d'identifier précisément chaque type de caractéristique (calcul de la régularité Lipschitzienne).

Dans ce qui suit, nous utilisons parfois comme bases locales :

- (u, v), où v correspond à la direction du gradient ;
- (p, q), correspondant aux directions associées aux vecteurs propres de la matrice Hessienne (p correspondant à la plus grande valeur propre).

Notons que (u, v) correspond à des dérivées du premier ordre (gradient) tandis que (p, q) à des dérivées du second ordre (courbure). Dans ces bases, les dérivées considérées sont prises par rapport à ces coordonnées locales, ce qui conduit à des expressions telles que $L_u = \partial_u L$.

Détection de bords (*edges*) [77]

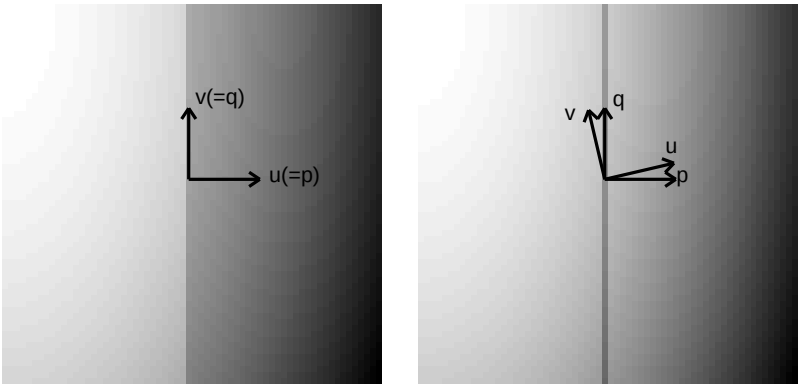
En un point correspondant à un bord (Fig. 4.2 (B)), la magnitude du gradient est localement maximale dans sa direction. Ceci conduit à l'opérateur

$$\mathcal{E}L = L_v^2 \tag{4.17}$$

dont la version γ-normalisée est : $\mathcal{E}_{norm}L = t^{\gamma}\mathcal{E}$.

Un critère permettant de déterminer si un point du Scale-Space correspond à un bord est alors donné par :

$$\begin{cases} \partial_t(\mathcal{E}_{norm}L)(x,y,t) = 0 \text{ et } \partial_{tt}(\mathcal{E}_{norm}L)(x,y,t) < 0 & \text{Maximum local en échelle} \\ L_{vv}(x,y,t) = 0 \text{ et } L_{vvv}(x,y,t) < 0 & \text{Maximum du gradient dans la direction du gradient} \end{cases}$$



(A) (B)

FIG. 4.2 – Exemples de bases locales (u, v) (associée au gradient) et (p, q) (associée à la matrice Hessienne) sur une image dont l'intensité est la somme d'un paraboloïde et d'un bord pour (A), d'une ligne pour (B) (légèrement lissée pour assurer l'existence des dérivées).

Détection de lignes (*ridges*) [77]

En un point appartenant à une ligne d'une image (ligne tracée dans une région où le niveau de gris est une fonction régulière, Fig. 4.2 (B)), l'intensité admet un maximum local suivant la direction de courbure principale et un minimum local suivant la direction orthogonale. Ceci conduit à l'opérateur suivant :

$$\mathcal{R}L = (L_{pp} - L_{qq})^2 \tag{4.18}$$

et l'opérateur γ-normalisé associé est : $\mathcal{R}_{norm} = t^{2\gamma}\mathcal{R}$.

Afin de caractériser dans le Scale-Space, un tel point de ligne, il est alors possible d'utiliser le critère suivant :

$$\begin{cases} \partial_t(\mathcal{R}_{norm}L)(x,y,t) = 0 \text{ et } \partial_{tt}(\mathcal{R}_{norm}L)(x,y,t) < 0 & \text{Maximum local en échelle} \\ L_p(x,y,t) = 0 \text{ et } L_{pp}(x,y,t) < 0 & \text{Maximum dans la direction définie par } p \end{cases}$$

Détection de coins (*corners*) [77]

Afin d'identifier les coins, il a été proposé d'utiliser le produit de la courbure κ multiplié par la magnitude du gradient L_v élevé à une certaine puissance (3 par exemple) [63]. Le premier terme κ correspond véritablement aux coins tandis que le deuxième évite des coins peu marqués (il ne change pas la localisation des points, tout en améliorant la qualité du détecteur, puisqu'il évite des divisions, sources d'instabilités numériques). Ceci conduit alors à l'opérateur suivant

$$\widetilde{\kappa}L = L_v^3\kappa = L_v^2 L_{uu} \tag{4.19}$$

Alors l'opérateur γ-normalisé correspondant est : $\widetilde{\kappa}_{norm}L = t^{2\gamma}L_v^2 L_{uu}$, et le critère associé est $(\widetilde{\kappa}_{norm}L)^2$ localement maximum en espace et en échelle.

Détection de structures type *blob* [77]

Les structures de type "blob" sont définies comme les lieux du Scale-Space où le module du Laplacien normalisé $t^\gamma L$ est localement maximum (en échelle et en espace). Cet opérateur s'écrit également comme

$$\text{trace}(\mathcal{H}_{norm}) \quad = \quad t^\gamma(L_{xx} + L_{yy}) = t^\gamma \nabla^2 L = t^\gamma \Delta L \tag{4.20}$$

Notons que ce n'est qu'*a posteriori* que le lien peut être fait entre un maximum dans le Scale-Space (x^*, y^*, s^*) et la région d'intérêt associée – disque centré en (x^*, y^*), de rayon (s^*). Ainsi, le terme *blob* est utilisé comme dénominateur pour un ensemble de structures telles qu'une surface Gaussienne (non nécessairement isotrope), une surface associée à la fonction indicatrice d'une ellipse, ou plus généralement un objet délimité par certaines singularités et ne présentant pas de contours internes. Nous verrons plus loin que les Maxima Lines permettent de mieux appréhender la forme de cette structure [30].

Remarque : pour la détection de tels *blobs*, il est possible d'utiliser également l'opérateur $\det(\mathcal{H}_{norm}) = t^{2\gamma}(L_{xx}L_{yy} - L_{xy}^2)$. Cela se justifie de la manière suivante : vu que $\mathcal{H}_{norm}$ est une matrice symétrique définie positive, elle est diagonalisable dans une base orthonormée, et ainsi :

$$\mathcal{H}_{norm} = P^{-1}AP = P^T AP \text{ , où } A = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \quad (\alpha, \beta > 0)$$

de sorte que $\text{trace}(\mathcal{H}_{norm}) = \alpha + \beta$ et $\det(\mathcal{H}_{norm}) = \alpha\beta$. Notons que si α et β sont forts en module, $|\text{trace}(\mathcal{H}_{norm})|$ et $|\det(\mathcal{H}_{norm})|$ le sont simultanément (néanmoins, ce n'est pas toujours le cas, par exemple si α est fort et β faible).

Ainsi, les structures de type *blob* sont détectées par sélection des maxima de $\text{trace}(\mathcal{H}_{norm})$ ou $\det(\mathcal{H}_{norm})$, en espace et en échelle. De plus, pour $\gamma = 1$, il y a invariance à l'échelle, au sens où si deux structures (type *blob*, au contenu identique) apparaissent à des échelles différentes, elles conduiront à la même réponse de l'opérateur.

Facteurs de γ-normalisation conduisant à l'invariance à l'échelle

Les opérateurs précédemment définis existent pour tout $\gamma > 0$. Si maintenant nous souhaitons obtenir l'invariance à l'échelle, i.e., que deux caractéristiques – au contenu identique, apparaissant à des échelles différentes – soient traitées de la même manière (même amplitude de la réponse de l'opérateur), il importe de fixer une valeur appropriée à γ, laquelle est variable suivant l'opérateur utilisé (cf. tableau 4.1 [76]). Cette valeur est calculée d'après le modèle de structure à détecter (bord, ligne, blob, ...), comme nous l'avons vu pour la fonction $\sin(\omega\cdot)$ (section 4.1.2).

TAB. 4.1 – Caractéristiques à extraire, exemple d'opérateur adapté, et valeur de γ pour laquelle il y a invariance à l'échelle.

Caractéristique détectée	Entité différentielle	γ-valeur
Bord	$t^{\gamma/2}L_v$	1/2
Ligne	$t^{2\gamma}(L_{pp} - L_{qq})^2$	3/4
Coin	$t^{2\gamma}L_v^2 L_{uu}$	1
Blob	$t^{\gamma}\left(L_{xx} + L_{yy}\right)$	1

4.2.2 Principe de sélection automatique d'échelle

L'intérêt d'utiliser des opérateurs multiéchelles est qu'ils permettent de mettre en valeur certaines échelles particulières, et suivant l'approche utilisée, celles-ci peuvent être caractéristiques de certaines structures de l'image. Nous présentons ici l'approche classique de sélection d'échelle, fondée sur des maxima dans le Scale-Space [77, 99].

Etant donné un opérateur Γ correctement normalisé (au sens de l'invariance à l'échelle) et une représentation espace-échelle L, nous considérons les maxima locaux de $|\Gamma L|$, à la fois en espace et en échelle [72, 77] :

Définition 30. *(Maxima dans le Scale-Space)*

$$MaxI = \{(x^*, y^*, s^*)/|\Gamma L| : \mathbb{R}^3 \longrightarrow \mathbb{R} \textit{ localement maximum.}\} \tag{4.21}$$

Dans un cadre 2D discret, l'espace-échelle est de dimension 3 et cela revient à comparer chaque point avec ses 26 voisins. L'intérêt de ces maxima est qu'ils correspondent à des structures importantes dans l'image, et qu'ils sont robustes ou invariants par certaines transformations sur l'image. Un maximum dans le Scale-Space est donné par $(x^*, y^*, s^*, |\Gamma L(x^*, y^*, s^*)|)$:

- (x^*, y^*) : position,
- s^* : échelle caractéristique,
- $|\Gamma L(x^*, y^*, s^*)|$: réponse associée.

Un tel maximum peut être associé à une région d'intérêt définie comme le disque centré en (x^*, y^*) de rayon s^*. Cette notion de région d'intérêt sera détaillée au chapitre 7, où nous verrons en particulier que le lien entre l'existence d'un tel maximum et un objet (structure présente dans l'image) soulève certaines difficultés.

4.3 Prise en compte de déformations géométriques

Nous nous intéressons maintenant aux conséquences d'une déformation géométrique de l'image : d'une part sur la représentation dans le *Scale-Space*, et d'autre part sur la matrice des moments d'ordre deux (MM2), qui est un outil permettant d'estimer en pratique de telles déformations. Nous présentons ici l'approche proposée par Lindeberg [79].

4.3.1 Transformations considérées

En premier lieu, nous nous donnons une application inversible $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$. Si nous considérons une image I, dont l'intensité est donnée par une fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, une nouvelle image I_1 (image déformée) dont l'intensité f_1 est définie comme

$$f_1(T(x, y)) = f(x, y).$$

Notons que cette transformation est d'ordre géométrique – ceci inclut les rotations, les changements d'échelle, les transformations affines locales. Nous cherchons alors à mettre en relation la représentation espace-échelle de f et celle de f_1. Dans le cas simple où I et I_1 sont deux images reliées par une translation, les deux représentations linéaires L et L_1 sont translatées l'une par rapport à l'autre. En revanche, il importe de noter que pour une transformation quelconque de l'image, cette relation est délicate à établir. Ici, nous nous intéressons plus particulièrement au cadre des transformations affines, où l'application T s'écrit comme

$$\begin{array}{rcccccl} T & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 & B \in M_2(\mathbb{R}) & : \text{ terme linéaire} \\ & & X & \mapsto & BX + C & C \in \mathbb{R}^2 & : \text{ translation} \end{array}$$

Remarque : la translation ne soulève pas de difficulté particulière (dans la suite nous supposons que $C = 0$). En revanche, la déformation géométrique liée à la matrice B constitue un point essentiel à traiter.

4.3.2 Invariance par transformation affine du Scale-Space

Nous nous intéressons ici aux représentations espace-échelle affines linéaires [79]. Une telle représentation est donnée par :

$$L(., \Sigma_t) = G(., \Sigma_t) \star f(.) \text{ où } \Sigma_t \text{ est symétrique définie positive.}$$

Si nous considérons maintenant deux images f_L et f_R reliées par une transformation linéaire, définie par $B \in GL_2$, la relation entre leurs intensités s'écrit :

$$\forall x \in \mathbb{R}^2, f_L(x) = f_R(Bx) \tag{4.22}$$

Les représentations affines de L et R s'écrivent alors comme

$$\begin{aligned} L(., \Sigma_t^L) &= G(., \Sigma_t^L) * f_L(.) \\ R(., \Sigma_t^R) &= G(., \Sigma_t^R) * f_R(.) \end{aligned}$$

Proposition 11. *(Lindeberg [79]) Pour $\Sigma_t^R = B\Sigma_t^L B^T$, nous avons (pour tout $x \in \mathbb{R}^2$)*

$$L(x, \Sigma_t^L) = R(y, \Sigma_t^R) \text{ avec } y = Bx, \tag{4.23}$$

Cette propriété montre que la représentation espace-échelle affine est bien adaptée aux transformations affines, dans le cas où cette transformation est globale (B constante). Notons que dans les applications, ce n'est que localement que les déformations peuvent être considérées comme affines.

4.4 Estimation de déformations géométriques

Précédemment, nous avons supposé que la matrice inversible B était connue. Dans l'optique de détection de caractéristiques invariantes, il est intéressant de pouvoir estimer des déformations géométriques. Ainsi, nous cherchons maintenant à calculer B, en faisant appel à un descripteur de l'image : la matrice des moments d'ordre deux (MM2) [2, 79].

4.4.1 Matrice des moments d'ordre deux (*Second Moment Matrix*)

La matrice des moments d'ordre deux – *second moment matrix* – est une primitive utilisée à la fois pour la détection de caractéristiques locales telles que des points d'intérêt, et également pour la description de structures locales présentes dans l'image. Elle possède différentes formulations plus ou moins complexes.

Définition 31. *(MM2 [79]) Etant donné une représentation espace-échelle affine L, la matrice des moments d'ordre deux associée est définie comme*

$$\mu_L(q) = \int_{\mathbb{R}^2} [\nabla L(x)][\nabla L(x)]^T w(q-x)dx \tag{4.24}$$

- *∇L est le gradient de l'image lissée,*
- *$w : \mathbb{R}^2 \longrightarrow \mathbb{R}$ est une fenêtre symétrique et normalisée (moyenne égale à 1),*
- *$\mu_L(q) \in SP_2$ (matrice 2×2 symétrique positive).*

Notons que la matrice $M(x) = (\nabla L(x))(\nabla L(x))^T$ est une matrice de rang 1. C'est en intégrant $M(x)$ sur une certaine fenêtre que nous obtenons une matrice μ_L qui sera inversible dans la plupart des cas ; les cas de non-inversibilité concernent les zones où $M(x)$ est constante sur le support de la fenêtre, ce qui survient rarement dans la pratique. Enfin, pour le choix de la fenêtre w, il est courant d'utiliser des noyaux Gaussiens, mais ce n'est pas le seul choix possible.

Effet d'un changement de coordonnées sur la MM2 [79]

Nous considérons deux images d'intensités f_L et f_R, reliées par $f_L(x) = f_R(Bx)$, où B est inversible. Leurs représentations espace-échelle linéaires L et R sont alors reliées par $\nabla L(x) = B^T \nabla R(Bx)$.

Ainsi

$$\begin{aligned}
\mu_L(q) &= \int_{\mathbb{R}^2} [\nabla L(x)][\nabla L(x)]^T w(q-x)dx \\
&= \int_{\mathbb{R}^2} [B^T \nabla R(Bx)][B^T \nabla R(Bx)]^T w(q-x)dx \\
&= B^T \left(\int_{\mathbb{R}^2} \nabla R(Bx) \nabla R(Bx)^T w(B^{-1}(Bq - Bx))dx \right) B
\end{aligned}$$

En effectuant alors le changement de variable

$$\begin{cases} y &= Bx \\ dy &= |det(B)|dx \end{cases}$$

nous obtenons alors

$$\begin{aligned}
\mu_L(q) &= B^T \int_{\mathbb{R}^2} \nabla R(y) \nabla R(y)^T \left(\frac{1}{|det(B)|} w(B^{-1}(Bq - y)) \right) dyB \\
\mu_L(q) &= B^T \mu'_R(p) B \text{ avec } p = Bq
\end{aligned}$$

où μ'_R est la MM2 calculée avec une fenêtre w', dépendant de w et B, et généralement différente de w. Ceci conduit Lindeberg à introduire une MM2 plus générale [79].

Formulations plus générales de la MM2

Dans la définition précédente de la MM2, la représentation L étant calculée d'après un noyau Gaussien isotrope. Ceci peut être généralisé par des noyaux anisotropes et une représentation L affine [75].

MM2 basées sur des noyaux Gaussiens affines [79]

Définition 32. *Etant donné deux matrices symétriques définies positives Σ_s et Σ_t, la MM2 affine est définie comme $\mu_L : \mathbb{R}^2 \times SDP_2 \times SDP_2 \longrightarrow SP_2$:*

$$\mu_L(\cdot, \Sigma_t, \Sigma_s) = g(\cdot, \Sigma_s) * \left([\nabla L(\cdot, \Sigma_t)][\nabla L(\cdot, \Sigma_t)]^T \right) \tag{4.25}$$

- *Σ_s dépend d'une échelle d'intégration s (sélection d'échelle),*
- *Σ_t dépend d'une échelle locale t (calcul du gradient),*
- *$g(., \Sigma_s)$ et $L(\cdot, \Sigma_t)$ ont été définis en 4.1.1 (eq. (4.4) et (4.5)).*

Proposition 12. *Etant donné deux images reliées par $f_L(x) = f_R(Bx)$ où $B \in GL_2$, leurs MM2 affines μ_L et μ_R sont reliées par*

$$\mu_L(q, \Sigma_t, \Sigma_s) = B^T \ \mu_R(Bq, B\Sigma_t B^T, B\Sigma_s B^T) \ B \tag{4.26}$$

4.4.2 Applications

Propriété d'invariance [79]

Dans le cas particulier où $\Sigma_t = t\Sigma_1$ et $\Sigma_s = s\Sigma_2$, les propositions 12 et 14 deviennent

$$\mu_L(q, t\Sigma_1, s\Sigma_2) = B^T \mu_R(Bq, tB\Sigma_1 B^T, sB\Sigma_2 B^T)B \tag{4.27}$$

$$\widetilde{\mu}_L(q, t\Sigma_1, s\Sigma_2) = B^T \widetilde{\mu}_R(Bq, tB\Sigma_1 B^T, sB\Sigma_2 B^T)B \tag{4.28}$$

Par ailleurs, il est possible d'effectuer une normalisation à l'échelle (cette normalisation est relative à $L(., \Sigma_t)$), ce qui conduit à

$$\mu_{L,norm} = t\Sigma_1 \mu_L(q, t\Sigma_1, s\Sigma_2) \tag{4.29}$$

Partant des équations (4.27) et (4.29), nous déduisons

$$\begin{aligned}
\det(\mu_L(q, t\Sigma_1, s\Sigma_2)) &= \det(B^T)\det(\mu_R(Bq, tB\Sigma_1 B^T, sB\Sigma_2 B^T))\det(B) \\
\det(\mu_{L,norm}) &= \det(t\Sigma_1 \mu_L) \\
&= \det(t\Sigma_1)\det(B^T)\det(\mu_R)\det(B) \\
&= \det(tB^T\Sigma_1 B)\det(\mu_R) \\
&= \det(\mu_{R,norm})
\end{aligned}$$

Les équations ci-dessus étant valables pour μ_L et $\widetilde{\mu}_L$, nous pouvons écrire

$$\det(\mu_{L,norm}(q, t\Sigma_1, s\Sigma_2)) = \det(\mu_{R,norm}(Bq, tB\Sigma_1 B^T, sB\Sigma_2 B^T) \tag{4.30}$$

$$\det(\widetilde{\mu}_{L,norm}(q, t\Sigma_1, s\Sigma_2)) = \det(\widetilde{\mu}_{R,norm}(Bq, tB\Sigma_1 B^T, sB\Sigma_2 B^T) \tag{4.31}$$

Comme conséquence, nous voyons que les maxima en échelle du déterminant normalisé d'une MM2 sont conservés par déformation affine. Remarquons que cette propriété, valable pour le déterminant $\det(\mu_{L,norm})$ n'est pas nécessairement valable pour la trace $\text{trace}(\mu_{L,norm})$.

Estimation de la déformation – Calcul effectif de B [79]

Nous cherchons ici à estimer la matrice de déformation B Nous nous plaçons dans le cas particulier où $\Sigma_t = t\Sigma_1$ et $\Sigma_s = s\Sigma_2$, de sorte que

$$\mu_L(q, t\Sigma_1, s\Sigma_2) = B^T \mu_R(Bq, tB\Sigma_1 B^T, sB\Sigma_2 B^T)B \tag{4.32}$$

ce qui permet d'écrire :

$$B = \mu_L^{-1/2} R_\alpha \mu_R^{1/2} \text{ où } R_\alpha \text{ est une matrice de rotation} \tag{4.33}$$

Ainsi, dans le cas où nous connaissons μ_L et μ_R, la matrice B est déterminée à une rotation près. Relevons que l'inconnue $B \in GL_2(\mathbb{R})$ dépend a priori de 4 paramètres, et que les données μ_L et μ_R dépendent de 3 paramètres (matrices symétriques, reliées par l'équation 4.32), ce qui explique qu'il reste un degré de liberté – l'angle de rotation. En pratique, cet angle peut être retrouvé par d'autres moyens. Relevons enfin qu'il existe un algorithme, appelé *Shape Adaptation*, qui permet d'estimer la matrice de déformation B [79] (cet algorithme est fondé sur une recherche de point fixe).

4.5 Généralisations par des fonctions échelles 2D

Nous avons généralisé les outils vus précédemment (représentation espace-échelle affine, matrice des moments d'ordre deux). Cette formulation originale utilise une fonction échelle 2D, $\Phi : \mathbb{R}^2 \longrightarrow \mathbb{R}$.

4.5.1 Nouvelle représentation espace-échelle

Dans la représentation espace-échelle donnée par l'éq. 4.5, nous remplaçons le noyau Gaussien affine par Φ. Etant donné $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, nous définissons $\mathcal{A}f$ pour tout $b \in \mathbb{R}^2$ et pour toute matrice Σ symétrique définie positive comme

$$\mathcal{A}f(b, \Sigma) = \int_{\mathbb{R}^2} \frac{1}{\sqrt{det(\Sigma)}} \Phi\left(\Sigma^{-1/2}(x-b)\right) f(x)dx \tag{4.34}$$

où $\Sigma^{-1/2}$ est une matrice symétrique définie positive telle que $\Sigma^{-1/2}\Sigma^{-1/2} = \Sigma$.

Proposition 13. *Etant donné deux images f_L et f_R reliées par $f_L(x) = f_R(Bx)$ (où B est une matrice inversible), nous avons*

$$\mathcal{A}f_L(b, \Sigma_L) = \mathcal{A}f_R(Bb, B\Sigma_L B^T) \tag{4.35}$$

Démonstration

Comme B est inversible, nous avons $det(B) \neq 0$, $|det(B)| = \sqrt{det(B)det(B^T)}$ et nous pouvons écrire :

$$\begin{aligned}
\mathcal{A}f_L(b, \Sigma_L) &= \int_{\mathbb{R}^2} \frac{1}{\sqrt{det(\Sigma_L)}} \Phi\left(\Sigma_L^{-1/2}(x-b)\right) f_R(Bx)dx \\
&= \int_{\mathbb{R}^2} \frac{1}{\sqrt{det(B\Sigma_L B^T)}} \Phi\left((B\Sigma_L^{1/2})^{-1}(Bx - Bb)\right) f_R(Bx)|det(B)|dx
\end{aligned}$$

En effectuant le changement de variable

$$\begin{cases} y &= Bx \\ dy &= |det(B)|dx \end{cases}$$

nous obtenons alors le résultat énoncé :

$$\begin{aligned}
\mathcal{A}f_L(b, \Sigma_L) &= \int_{\mathbb{R}^2} \frac{1}{\sqrt{det(B\Sigma_L B^T)}} \Phi\left((B\Sigma_L^{1/2})^{-1}(y - Bb)\right) f_R(y)dy \\
&= \mathcal{A}f_R(Bb, B\Sigma_L B^T)
\end{aligned}$$

4.5.2 MM2 définie par une fonction échelle

Nous généralisons également la définition 32 ainsi que la prop. 12.

Définition 33. *(MM2 généralisée $\widetilde{\mu}_L$)*

$$\begin{array}{rcccl}
\widetilde{\mu}_L & : & \mathbb{R}^2 \times SDP_2 \times SDP_2 & \longrightarrow & SP_2 \\
& & (q, \Sigma_t, \Sigma_s) & \longmapsto & \widetilde{\mu}_L
\end{array}$$

$$\text{où } \widetilde{\mu}_L(q, \Sigma_t, \Sigma_s) = \int_{\mathbb{R}^2} [\nabla_b \mathcal{A}f(b, \Sigma_t)][\nabla_b \mathcal{A}f(b, \Sigma_t)]^T \frac{1}{\sqrt{det(\Sigma_s)}} \Phi(\Sigma_s^{-1/2}(q-b))db \tag{4.36}$$

Proposition 14. *Etant donné deux images reliées par $f_L(x) = f_R(Bx)$ où $B \in GL_2$, leurs MM2 en ondelettes $\widetilde{\mu}_L$ et $\widetilde{\mu}_R$ sont reliées par*

$$\widetilde{\mu}_L(q, \Sigma_t, \Sigma_s) = B^T \widetilde{\mu}_R(Bq, B\Sigma_t B^T, B\Sigma_s B^T)B \tag{4.37}$$

Démonstration

Nous rappelons l'équation (4.35), que nous différencions (gradient par rapport à $b \in \mathbb{R}^2$), d'où

$$\begin{aligned}
\mathcal{A}f_L(b, \Sigma_L) &= \mathcal{A}f_R(Bb, B\Sigma_L B^T) \\
\nabla_b \mathcal{A}f_L(b, \Sigma_t) &= B^T \nabla_b \mathcal{A}f_R(Bb, B\Sigma_L B^T) \\
[\nabla_b \mathcal{A}f_L(b, \Sigma_x t)]\,[\nabla_b \mathcal{A}f_L(b, \Sigma_t)]^T &= B^T[\nabla_b \mathcal{A}f_R(Bb, B\Sigma_L B^T)][\nabla_b \mathcal{A}f_R(Bb, B\Sigma_L B^T)]^T B
\end{aligned}$$

Ceci permet alors d'écrire :

$$\begin{aligned}\widetilde{\mu}_L(q,\Sigma_t,\Sigma_s) &= B^T\left(\int_{\mathbb{R}^2}[\nabla_b\mathcal{A}f_R(Bb,B\Sigma_LB^T)][\nabla_b\mathcal{A}f_R(Bb,B\Sigma_LB^T)]^T\frac{\Phi(\Sigma_s^{-1/2}(q-b))}{\sqrt{\det(\Sigma_s)}}db\right)B\\ &= B^T\widetilde{\mu}_R(Bq,B\Sigma_tB^T,B\Sigma_sB^T)B\end{aligned}$$

4.6 Liens entre Scale-Space et ondelettes

Nous présentons ici quelques liens que nous avons établis entre le Scale-Space et la théorie des ondelettes [30, 31], lesquels ont été appliqués à la détection de *blobs* (chap. 7).

4.6.1 Cadre 1D

Dans un cadre 1D, les expressions de la Gaussienne $G(x)$, du noyau Gaussien $G(x,t)$, et de la représentation linéaire espace-échelle L s'écrivent

$$\begin{array}{rclcl} G(x) &=& \frac{1}{\sqrt{2\pi}}\exp(-\frac{x^2}{2}) & & \\ G(x,t) &=& \frac{1}{\sqrt{t}}G(\frac{x}{\sqrt{t}}) &=& \frac{1}{\sqrt{2\pi t}}\exp(-\frac{x^2}{2t}) \\ L(x,t) &=& (G(.,t)*f)(x) &=& \int_{\mathbb{R}}\frac{1}{\sqrt{t}}G(\frac{x-u}{\sqrt{t}})f(u)du \end{array}$$

Si nous considérons une dérivée m-ième de L, $\partial_x^m L$, la normalisation proposée par Lindeberg conduit à étudier la quantité $t^{m\gamma/2}\partial_x^m L$. Celle-ci s'écrit alors comme [31] :

$$\begin{aligned} t^{m\gamma/2}\partial_x^m L(x,t) &= t^{m\gamma/2}\int_{\mathbb{R}}\frac{1}{\sqrt{t}}\partial_x^m\left(G(\frac{x-u}{\sqrt{t}})\right)f(u)du\\ &= t^{m\gamma/2}\int_{\mathbb{R}}(\frac{1}{\sqrt{t}})^{m+1}G^{(m)}(\frac{x-u}{\sqrt{t}})f(u)du\\ &= t^{m/2(\gamma-1)}\int_{\mathbb{R}}\frac{1}{\sqrt{t}}G^{(m)}(\frac{x-u}{\sqrt{t}})f(u)du \end{aligned}$$

et dans le cas particulier de $\gamma=1$ (invariance à l'échelle), nous obtenons

$$t^{m/2}\partial_x^m L(x,t)=\int_{\mathbb{R}}\frac{1}{\sqrt{t}}G^{(m)}(\frac{x-z}{\sqrt{t}})f(z)dz \tag{4.38}$$

Etant donné que $G^{(m)}$ est une ondelette (dérivée m-ème de Gaussienne), nous pouvons effectuer un parallèle entre l'expression ci-dessus et un coefficient d'ondelette (Tab. 4.2). Rappelons qu'étant donné une ondelette Ψ, un tel coefficient est donné par :

$$Wf(u,s)=\int_{\mathbb{R}}\frac{1}{s}\Psi(\frac{z-u}{s})f(z)dz \tag{4.39}$$

Dans le cas où $\Psi = (-1)^m G^{(m)}$, nous avons

$$t^{m/2}\partial_x^m L(x,t) = Wf(x,\sqrt{t}) \tag{4.40}$$

Par ailleurs, notons que

$$Wf(x,\sqrt{t}) = \begin{cases} \int_{\mathbb{R}} \frac{1}{\sqrt{t}}\Psi\left(\frac{x-z}{\sqrt{t}}\right) f(z)dz & \text{si } \Psi \text{ est symétrique} \\ -\int_{\mathbb{R}} \frac{1}{\sqrt{t}}\Psi\left(\frac{x-z}{\sqrt{t}}\right) f(z)dz & \text{si } \Psi \text{ est antisymétrique} \end{cases} \tag{4.41}$$

En outre, étant donné $f : \mathbb{R} \longrightarrow \mathbb{R}$, et $a > 0$, en posant $f_1(az) = f(z)$, nous avons :

$$\begin{aligned} Wf_1(u,s) &= \int_{\mathbb{R}} \frac{1}{s}\Psi(\frac{z-u}{s}) f_1(az)dz \\ &= \int_{\mathbb{R}} \frac{1}{s}\Psi(\frac{z-u}{s}) f(z/a)dz \\ &= \int_{\mathbb{R}} \frac{1}{s}\Psi(\frac{az-u}{s}) f(z)adz \\ &= \int_{\mathbb{R}} \frac{1}{s/a}\Psi(\frac{z-u/a}{s/a}) f(z)dz \end{aligned}$$

d'où il ressort

$$Wf_1(u,s) = Wf(u/a, s/a) \tag{4.42}$$

Tab. 4.2 – Correspondance entre *Scale-Space* et ondelettes

	Scale-Space	Ondelettes
Espace	x	u
Echelle	$\sqrt{t}$	s
Opérateur	Fonction des dérivées de $L(x,t)$	Transformée en ondelette $Wf(u,s)$

Conséquences sur les ML en 1D [79]

Nous nous intéressons à l'effet d'une dilatation d'un signal sur les ML. A cet effet, nous considérons deux signaux f et f_1 reliés par $f_1(u) = f(\lambda u)$, où $\lambda > 0$; la représentation dans le *Scale-Space* de f_1 s'écrit (après le changement de variable $v = \lambda u$)

$$L_1(x,t) = \int_{\mathbb{R}} \frac{1}{\sqrt{t}} G(\frac{x-u}{\sqrt{t}}) f(\lambda u)du = \int_{\mathbb{R}} \frac{1}{\sqrt{\lambda^2 t}} G(\frac{\lambda x - v}{\sqrt{\lambda^2 t}}) f(v)dv = L(\lambda x, \lambda^2 t)$$

de sorte que pour tout $m \in \mathbb{N}$:

$$\partial_\xi^m L_1(x,t) = \lambda^m \partial_\xi^m L(\lambda x, \lambda^2 t) \tag{4.43}$$

Ainsi, dans le cas d'une déformation globale (λ constant), dès lors que les ML associées à f sont connues, il est possible de déduire celles associées à f_1. Notons que le cadre de déformations quelconques est plus délicat, car l'équation (4.43) n'est plus forcément vérifiée.

4.6.2 Opérateur Laplacien et Maxima Lines en 2D

Des travaux ont montré que le Laplacien de la représentation espace-échelle, ΔL, est un opérateur classique pour la détection de structures bidimensionnelles dans une image [77, 99]. Sa version γ-normalisée s'écrit alors $t^\gamma \Delta L$. En notant Δf le Laplacien de f (pris au sens des distributions [113]), il est intéressant de remarquer que, pour $\gamma = 1$ (invariance à l'échelle)

$$t^\gamma \Delta L = t\Delta L = t\Delta f * G_t = f * (\Delta G)_t \tag{4.44}$$

Par ailleurs, l'ondelette $\Psi = \Delta G$ (définie en 3.2.1, eq. (3.2)) présente une symétrie de révolution ($\Psi(x, y) = \phi(|(x, y)|_2)$). Ceci permet de voir que la transformée en ondelette continue de f, normalisée L^1, utilisant l'ondelette ΔG, calculée en $(x, y, \sqrt{t})$, est égale à $t\Delta L$, calculée en (x, y, t).

Remarque : $\text{trace}(\mathcal{H}_{norm}) = t\Delta L$ solution de l'équation de la chaleur au temps t, avec la condition initiale Δf.

Conséquence sur les Maxima Lines en 2D

Nous étudions ici l'influence d'une transformation géométrique sur les ML, dans le cas d'une déformation affine constante sur toute l'image : nous considérons deux images f et f_1 reliées par $f(x) = f_1(Bx), B \in GL_2$. Nous avons vu (section 4.3.2, prop. 11) que pour toute matrice $\Sigma_t \in SDP_2$,

$$L(x, \Sigma_t) = g(\cdot, \Sigma_t) * f = g(\cdot, B\Sigma_t B^T) * f_1 \tag{4.45}$$

Ainsi, dans le cas où $\Sigma_t = t\Sigma_0$, $L(x, \Sigma_t)$ est solution de chacune des EDP suivantes :

$$\partial_t L = \tfrac{1}{2}\text{div}(\Sigma_0 \nabla L) \quad \text{avec } L(x, 0) = f(x) \tag{4.46}$$

$$\partial_t L = \tfrac{1}{2}\text{div}(B\Sigma_0 B^T \nabla L) \quad \text{avec } L(x, 0) = f_1(x) \tag{4.47}$$

Rappelons que ces deux EDP sont paraboliques, et que pour $f, f_1 \in L^2$, elles admettent une unique solution et le principe du maximum est vérifié [14, 108]. En particulier les modules maxima de $L(x, \Sigma_t)$ se propagent vers les échelles fines [126].

Maintenant, nous notons $\mathcal{H}_{norm}^{\Sigma_t}$ la matrice Hessienne associée à $L(x, \Sigma_t)$. Relevons alors que $\text{trace}(\mathcal{H}_{norm}^{\Sigma_t})$ est solution de l'EDP (4.46) (resp. (4.47)), en utilisant comme condition initiale Δf (resp. Δf_1). Vu l'unicité de la solution et le principe du maximum, ces EDP conduisent alors à des lignes de maxima identiques, associées à $\text{trace}(\mathcal{H}_{norm}^{\Sigma_t})$.

4.6.3 Applications

Dans une optique de détection de structures remarquables dans un signal (comme le créneau dans un cadre bruité vu en 2.4.3), nous nous intéressons aux conséquences d'un changement d'échelle, dont le facteur est constant. Celui-ci déplace la position des coefficients dans l'espace-échelle ; néanmoins, la propriété d'extremum local – en 2D, suivant u et s – est conservée. Ceci est intéressant pour la détection de points d'intérêt, où, après avoir détectés des extrema dans le Scale-Space, ceux-ci sont seuillés indépendamment de l'échelle. Si toutes les échelles n'étaient pas également traitées, ce seuillage induirait un biais non souhaité. Ainsi, lorsqu'il y a invariance d'échelle (par exemple $\gamma = 1$ pour la détection de *blobs*, Tab. 4.1), la représentation de Lindeberg peut être vue comme une transformée en ondelettes. Plus précisément, dans le cas où Ψ est symétrique ou antisymétrique, une transformée en ondelette continue normalisée L^1 constitue une représentation espace-échelle plus générale, normalisée au sens de l'invariance à l'échelle. Ceci justifie alors l'utilisation des ML de l'ondelette ΔG pour la détection de points d'intérêt et, plus généralement, l'utilisation de transformées en ondelettes pour la détection de structures remarquables et le calcul de grandeurs caractéristiques.

4.7 Conclusion

La théorie du *Scale-Space* permet de définir des opérateurs normalisés, dont les maxima (en espace et en échelle) possèdent des propriétés d'invariance vis à vis de certaines transformations géométriques. Ceci permet alors de formuler des méthodes pertinentes de détection automatique de certaines structures, dont la taille est proportionnelle à une certaine échelle caractéristique. Le calcul de telles grandeurs – caractéristiques des entités détectées – apparaît comme essentiel dans les applications, où les déformations sont souvent inconnues *a priori*.

Dans ce chapitre, nous avons montré des liens existants entre la théorie du *Scale-Space* et les ondelettes, et les extensions que celles-ci permettent. Aux chapitres suivants, nous insistons sur deux points : d'une part la détection de certaines entités, et d'autre part l'extraction de caractéristiques associées à celles-ci (échelle proportionnelle à sa taille, estimation de sa forme). Nous verrons notamment qu'au lieu d'utiliser des opérateurs adaptés spécifiquement aux bords, lignes et coins, il peut être pertinent de détecter de telles structures 1D par ondelettes (chap. 6). En outre, afin de détecter des structures 2D, nous utiliserons un mécanisme original de sélection d'échelle par maxima le long de ML (chap. 7). Dans les deux cas, nous montrerons l'intérêt d'utiliser les ML tant pour la détection que la caractérisation des objets détectés.

Chapitre 5

Construction des lignes de maxima

5.1 Motivation

Dans l'optique de la détection de caractéristiques locales, nous souhaitons formuler des approches alternatives à celles issues de la théorie du *Scale-Space*, utilisant les ondelettes. L'originalité de ces approches est qu'elles sont fondées sur des lignes de maxima – Maxima Lines (ML) et lignes de Canny (LC) – dont la construction numérique est détaillée ici.

Détection de contours multiéchelles

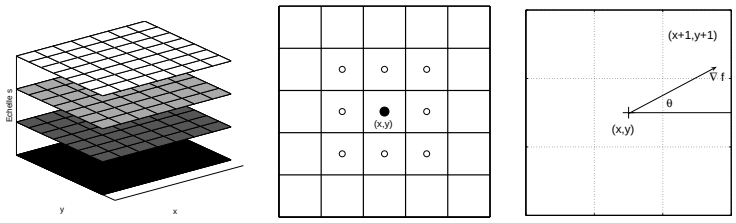
Nous avons vu précédemment une méthode de détection de contours issue du *Scale-Space*, où l'accent est mis sur l'échelle comme moyen de sélection. Par ailleurs, nous avons vu que les ondelettes permettaient de détecter des contours multiéchelles (chap. 3). Nous verrons au chapitre 6 une approche fondée sur les LC, permettant de détecter et caractériser des points de contour. Celle-ci part du principe que les contours apparaissent tous à l'échelle la plus fine, et le nombre d'échelles traversées par une ligne de Canny (LC) indique l'importance du point de contour associé. En outre, nous verrons comment calculer la régularité Lipschitzienne et l'échelle caractéristique associée à un contour (structure monodimensionnelle).

Détection de *blob*

Nous avons vu que ΔL équivaut à une transformée en ondelette continue utilisant l'ondelette ΔG. Ceci est intéressant dans la mesure où il existe des méthodes spécifiques aux ondelettes (bancs de filtres, méthodes spectrales) qui se démarquent des opérations couramment utilisées en informatique (lissage par un filtre Gaussien discret, dérivation discrète). Nous verrons au chapitre 7 que les ML permettent de formuler un nouveau mécanisme de sélection d'échelle – utilisant la jonction de certaines ML dans le *Scale-Space*. Ceci conduira alors à un détecteur d'objets de type *blob*. De plus, nous verrons que les propriétés des ML permettent de mieux décrire la structure de l'objet associé à une jonction de ML.

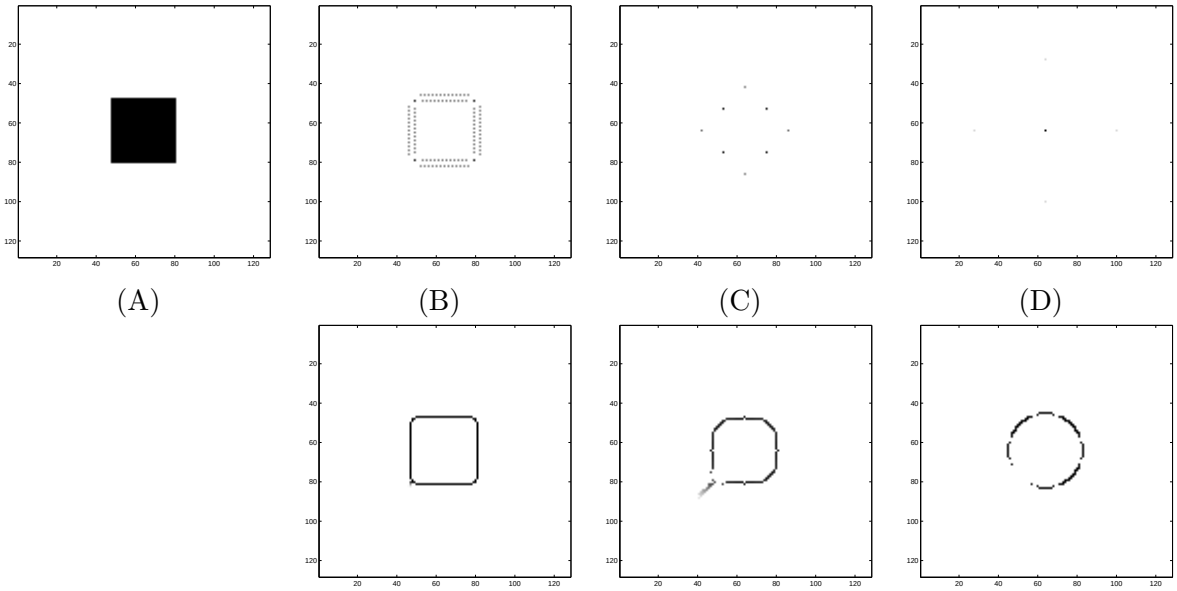
5.2 Construction numérique des lignes de maxima

Schématiquement, la construction de lignes de maxima procède en deux étapes : d'abord, nous calculons un ensemble de cartes de maxima associées à différentes échelles (Fig. 5.1-5.2) ; puis, des règles de chaînage permettent de définir des lignes de maxima (reliant des maxima d'une échelle à une autre). Relevons que ces règles sont différentes pour les ML (associées aux MM de l'ondelette ΔG) et les LC (associés aux MC).



(A) (B) (C)

FIG. 5.1 – (A) Exemple de cartes multiéchelles ; (B) Module maximum : point où le module de la TOC admet un maximum local (maximum dans toutes les directions) ; (C) Maximum au sens de Canny : point où le module du gradient admet un maximum suivant la direction du gradient.



(A) (B) (C) (D)

(B') (C') (D')

FIG. 5.2 – (A) Image représentant un carré ; (B–D) Modules maxima (MM) de l'ondelette ΔG, (B'–D') Maxima de Canny (MC) multiéchelles ; ceux-ci sont représentés à différentes échelles : (A–A') la plus fine, (B–B') intermédiaire, (C–C') grossière.

5.2.1 Construction des Maxima Lines (ML)

Nous détaillons ici la construction des ML, que nous avons proposée et implémentée : le calcul des cartes de MM et les règles permettant de chaîner ces MM d'une échelle à une autre. Rappelons que chaque MM considéré est un maximum en espace, les cartes étant établies de manière indépendante les unes des autres. Du point de vue théorique, nous sommes assurés que certains maxima perdurent vers les échelles fines, ce qui aboutit à la notion de Maxima Line (ML). Nous proposons alors des règles permettant la construction effective de lignes de MM, de manière cohérente avec le comportement théorique des ML. En particulier, celles-ci peuvent fusionner à une certaine échelle, et nous verrons que cette notion de jonction de ML est cruciale pour la détection d'objets (chap. 7).

Notons que lorsque l'échelle croît, le nombre de MM tend à diminuer. De plus, certains MM dérivent vers l'extérieur du carré et d'autres vers l'intérieur. En outre, des résultats pratiques permettent de remarquer que ces MM correspondent à des singularités différentes :

- aux échelles fines, à des singularités bien localisées (coins, bords, points de contours) ;
- à de plus grandes échelles, à des entités présentant une certaine cohérence, ce que nous préciserons plus loin.

Calcul des cartes de MM

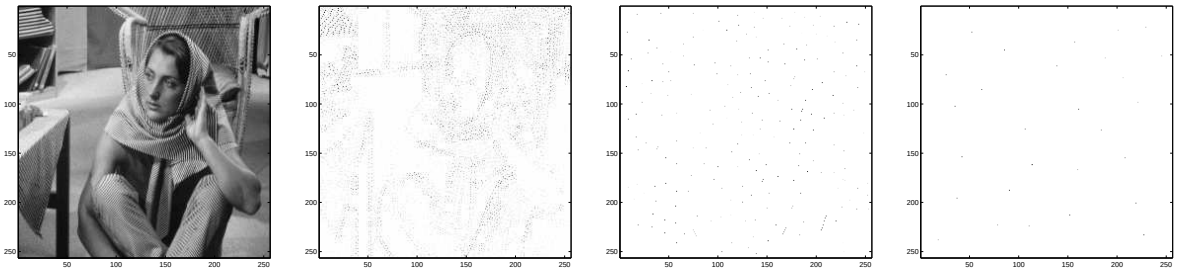
Si un module maxima est présent en (x, y) à une échelle s, celui-ci vérifie $|Wf(x, y, s)| > 0$ et donc la valeur associée $Wf(x, y, s)$ est non nulle. Ainsi, une carte associée à une certaine échelle est donnée par un tableau T de taille $n \times n$, où en tout point $(x, y) \in \{1, 2, ..., n\}$:

$$T(x, y) = \begin{cases} Wf(x, y, s) & \text{s'il y a un MM en } (x, y) \text{ à l'échelle } s \\ 0 & \text{sinon} \end{cases}$$

Afin d'obtenir de telles cartes pour une certaine gamme d'échelles, $s \in \{1, 2, 3...s_{\max}\}$ par exemple, nous calculons d'abord, pour chaque échelle s, la TOC utilisant l'ondelette ΔG. Nous disposons alors à chaque échelle d'une carte de MM, et nous conservons Wf comme information supplémentaire ; ainsi un MM est associé à un module et à un signe (positif ou négatif). Un MM sera donné par $(x(s), y(s), s, Wf(x(s), y(s), s))$, et l'ensemble des MM peut alors être représenté comme une série de cartes indexées par l'échelle et de même taille que l'image (Fig. 5.3).

Règles de construction des ML

Relevons que pour les application visées, toutes les échelles sont importantes : les plus petites permettront de caractériser des singularités (chap. 6) tandis que de plus grandes seront utiles pour la détection de *blobs* (chap. 7). Ainsi, la construction des ML présentée ici est fondée sur des cartes correspondant à des échelles entières (nous écartons les échelles dyadiques dans la mesure où elles croissent trop rapidement).



(A) (B) (C) (D)

FIG. 5.3 – Localisation des modules maxima (MM) de l'ondelette ΔG sur une image naturelle : (A) Image classique *Barbara* ; (B) MM à l'échelle la plus fine ; (C) MM à une échelle intermédiaire ; (D) MM à une échelle grossière.

Dans l'approche proposée ici, le point de départ consiste à considérer un MM à l'échelle la plus fine $s = 1$, présent en $(x_0, y_0) = (x(1), y(1))$, ce qui définit l'origine d'une ML, et associé à la valeur $Wf(x(1), y(1), 1)$. Nous cherchons alors à le connecter à un MM de même signe, de manière à ce que la ML construite permette de suivre une singularité donnée à travers les échelles :

- aux plus petites échelles ($s = 1, 2, 3, ...$), où il y a généralement un grand nombre de MM, parmi ceux de même signe présents dans un voisinage 3x3 à l'échelle $s + 1$, nous le connectons à celui dont le module est le plus proche (s'il n'y en a aucun, la ML s'interrompt) ;
- à des échelles plus grandes, là où les MM sont plus épars, nous connectons un MM à l'échelle s à celui qui est le plus proche en distance euclidienne à l'échelle $s + 1$, à condition que :
  - il n'y ait pas de MM de signe opposé dans le rectangle défini par les coins $(x(s), y(s))$ et $(x(s + 1), y(s + 1))$ (pas de croisement de ML),
  - les modules associés $Wf(x(.), y(.), .)$ soient suffisamment proches (de sorte que la ML suit bien la même singularité).

Remarque : un autre choix possible consiste en une discrétisation fine – comme $s = n\delta, n \in \mathbb{N}^*$ avec $\delta \in]0, 1[$ – ce qui permet de mieux suivre la dérive spatiale des ML. En particulier, lorsque deux (ou plusieurs) ML fusionnent, la dérive spatiale peut être importante entre deux échelle successives ; ce phénomène vu en 1D se retrouve en 2D, où des ML fusionnent à des échelles inconnues *a priori*. Une discrétisation fine assurerait alors qu'entre deux échelles successives, la ML ne dérive pas de plus d'un pixel et ainsi la règle de recherche dans un voisinage 3×3 serait valable pour toutes les échelles. Cependant, en pratique, cela amène à choisir une résolution très fine en échelle (typiquement $\delta \leq 0.01$) et conduit à un temps de calcul très élevé. Dans l'optique de construire des ML, une solution envisageable consisterait à utiliser d'abord des échelles entières, puis à raffiner de manière adaptative. Notons enfin,

dans un cadre plus général, qu'il est possible de considérer un ensemble quelconque de cartes dépendant de l'échelle (comme par exemple les MM de l'opérateur $det(\mathcal{H}_{norm})$, vu au chap. 4) ; les règles précédentes permettent alors de construire empiriquement des chemins (ou pseudo lignes de maxima) dans l'espace-échelle (x, y, s), sans qu'une continuité de ces chemins soit établie théoriquement. Pour les MM de l'ondelette ΔG, le choix d'échelles entières couplé à l'application des règles que nous avons proposées permet d'obtenir des lignes satisfaisantes en un temps de calcul raisonnable.

5.2.2 Construction des lignes de Canny (LC)

Calcul des cartes de MC

Nous présentons d'abord un algorithme permettant de calculer des cartes de maxima au sens de Canny dans un cadre multiéchelles. Initialement, le détecteur de Canny considérait une échelle s fixée ; ici nous utilisons la version multiéchelles (vue au chap. 3) formulée à l'aide d'une ondelette gradient [89]. Dans le cas où le noyau est Gaussien, le module du gradient s'exprime comme

$$Mf = \sqrt{|W_x f|^2 + |W_y f|^2} \tag{5.1}$$

En outre, l'orientation du gradient s'exprime comme $Af = Af(x, y, s)$:

$$Af = \begin{cases} \alpha(u) & \text{si } W_x f \geq 0 \\ \pi - \alpha(u) & \text{si } W_x f < 0 \end{cases} \quad \text{avec } \alpha(u) = \arctan\left(\frac{W_y f}{W_x f}\right) \tag{5.2}$$

Une fois Mf et Af calculés en tout (x, y, s), les maxima de Canny à une échelle s sont déterminés d'après les règles résumées dans le tableau 5.1. Les maxima multiéchelles de Canny sont alors de la forme $(x(s), y(s), s, Mf, Af)$ (Fig. 5.4).

Tab. 5.1 – Règles permettant de déterminer si (x, y) est un maxima au sens de Canny. La direction du gradient $\theta = Af$ détermine un point voisin PV ; si le module du gradient Mf en (x, y) est plus grand qu'en PV_θ et $\text{PV}_{-\theta}$ (strictement dans au moins un des deux cas), alors (x, y) est un MC.

Valeurs	de	θ		(mod.2π)	PV_θ
$-\pi/8$	$\leq$	θ	$<$	$\pi/8$	$(x+1, y)$
$\pi/8$	$\leq$	θ	$<$	$3\pi/8$	$(x+1, y)$
$3\pi/8$	$\leq$	θ	$<$	$5\pi/8$	(x, y)
$5\pi/8$	$\leq$	θ	$<$	$7\pi/8$	$(x-1, y)$
$7\pi/8$	$\leq$	θ	$<$	$-7\pi/8$	$(x-1, y)$
$-7\pi/8$	$\leq$	θ	$<$	$-5\pi/8$	$(x-1, y)$
$-5\pi/8$	$\leq$	θ	$<$	$-3\pi/8$	(x, y)
$-3\pi/8$	$\leq$	θ	$<$	$-\pi/8$	$(x+1, y)$

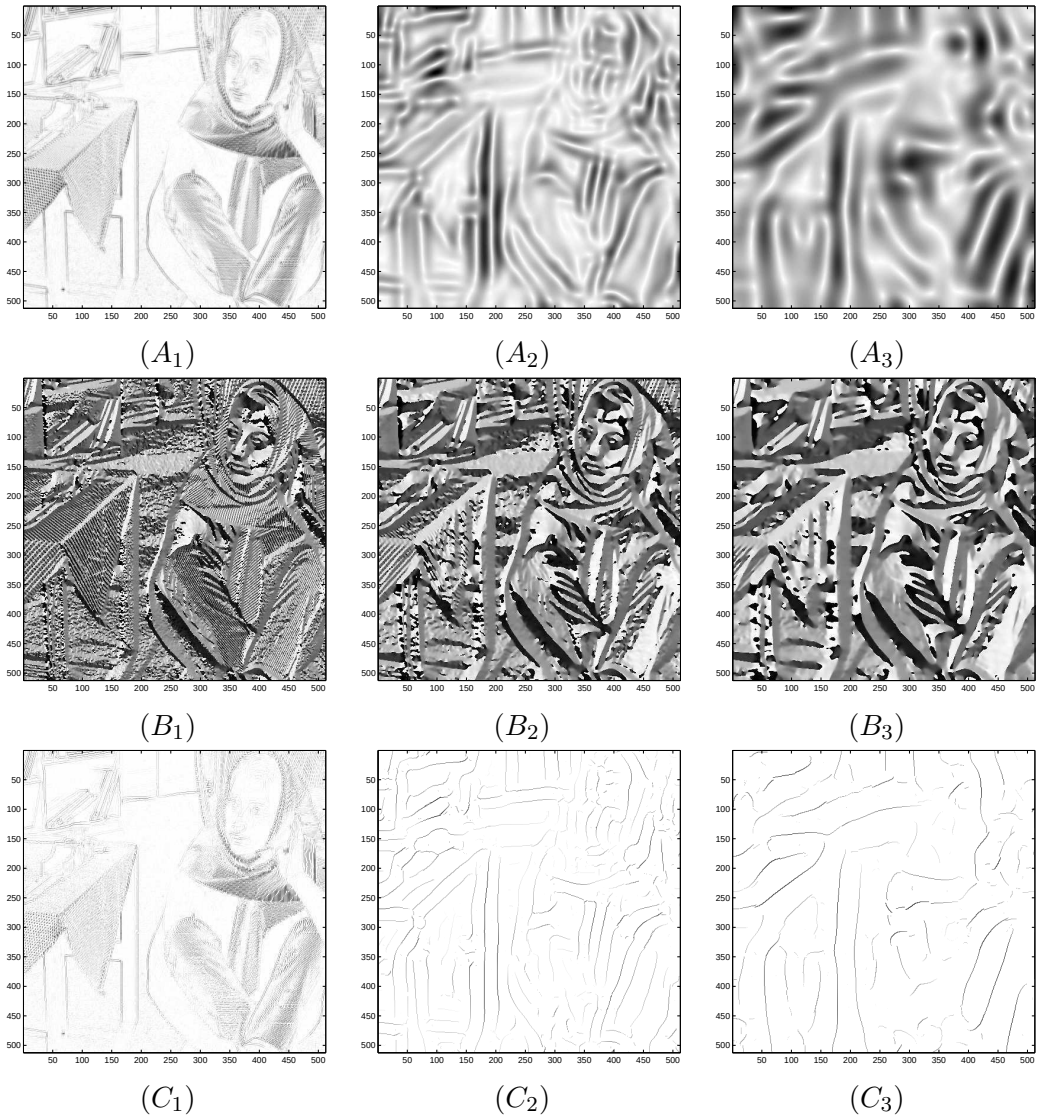


FIG. 5.4 – Calcul des maxima au sens de Canny, à des échelles s_i ($s_1 = 1, s_2 = 10, s_3 = 20$), sur l'image *Barbara* : (A_i) Module du gradient ; (B_i) Orientation du gradient ; (C_i) Maxima au sens de Canny.

Règles de construction des LC

Nous présentons ici des règles permettant de définir empiriquement des chemins à travers les échelles de maxima au sens de Canny, appelées lignes de Canny (LC). L'implémentation que nous avons réalisée utilise des règles similaires à celles présentées dans [69]. Celles-ci permettent de construire des lignes non triviales (i.e., qui s'interrompent pas trop rapidement) et cohérentes avec le contenu de l'image. Ainsi, une ligne doit suivre un contour à travers les échelles : à mesure que l'échelle croît, le lissage sous-jacent est de plus en plus important et seul les contours les plus significatifs sont conservés. La règle de chaînage est la suivante : étant donné deux maxima au sens de Canny à des échelles successives, notés $(x_i, y_i, s_i, Mf_i, Af_i)$ ($i = 1$ ou 2), ils seront chaînables à condition que :

- (x_2, y_2) soit dans le voisinage 3×3 de (x_1, y_1) ;
- les angles Af_1 et Af_2 soient assez proches ($|Af_1 - Af_2| < \pi/4$).

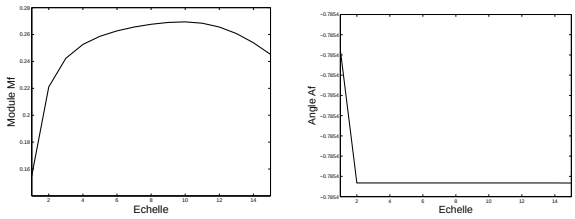
Alors, pour la construction des LC, nous choisissons les maxima de Canny tels que l'orientation du gradient soit la plus stable possible le long de ces lignes.

Par ailleurs, notons que deux choix sont possibles pour le chaînage : depuis les petites vers les grandes échelles ou en sens inverse. Dans le sens des échelles croissantes (en partant des échelles fines), il y aura un grand nombre de singularités détectées ; cependant, parmi celles-ci, nombreuses seront celles qui ne traverseront pas un nombre d'échelles suffisant, et qui ne seront pas exploitables. Notons qu'en parcourant les échelles dans le sens croissant, il est possible de décider qu'un maximum ne perdure pas de s à $s + 1$, ce qui permet de gérer l'interruption de certaines lignes. A l'inverse, dans le sens des échelles décroissantes (en partant des échelles grossières), nous garantissons que les lignes de maxima correspondront à des singularités significatives. Apparaît alors l'importance du choix de l'échelle $s_{\max}$, qui doit être ajustée de manière à ce que suffisamment de maxima soient présents ($s_{\max} = 10$ constitue une valeur satisfaisante). Si nous supposons que les lignes ne s'interrompent pas lorsque l'échelle diminue, cela implique que tout maximum à l'échelle s doit être chaîné à un autre maximum à l'échelle $s - 1$.

Ainsi suivant les échelles considérées, l'un ou l'autre sens de parcours pourra être utilisé : dans le sens décroissant si $s_{\max}$ est petit et dans le sens croissant si $s_{\max}$ est grand ; dans ce dernier cas, il est important de réussir à chaîner jusqu'aux plus petites échelles, afin d'assurer que ces lignes soient associés à des singularités de l'image (visibles aux fines échelles). Notons enfin que dans le cas du chaînage dans le sens des échelles décroissantes, nous excluons des lignes de Canny qui s'interrompent prématurément : ceci est justifié par le fait que seuls les contours les plus significatifs nous intéressent.

Remarque sur les lignes de Canny

En dépit de l'absence de théorie des LC, leur construction soulève rarement des difficultés pratiques. Nous avons remarqué que la stabilité de la direction du gradient Af joue un rôle important. En pratique, cette direction du gradient n'est pas constante : nous observons globalement qu'elle est instable aux échelles fines, et stable pour des échelles plus grandes (à partir de $s = 3$, Fig. 5.5, voir également les figures 5.4 ($B_1 - B_3$)). Cette stabilité s'explique par le fait qu'il ne reste alors que les contours les plus significatifs, alors les maxima aux fines échelles sont sujets à des instabilités. Notons enfin que pour des points de contour localisés près d'un bord, la direction du gradient Af est relativement stable, tandis que près d'un coin, celle-ci peut être plus facilement influencée par son voisinage dans l'image.



(A) (B)

FIG. 5.5 – Evolution en fonction de l'échelle, le long d'une ligne de Canny associée à un point de contour : (A) du module Mf du gradient de l'image ; (B) de l'orientation Af de celui-ci.

5.3 Complexité algorithmique (ML et LC)

Nous étudions ici le coût algorithmique de la construction des ML et des LC, en théorie et en pratique. Nous notons N la taille des données (pour une image carrée $n \times n$, $N = n^2$), et $s_{\max}$ la plus grande échelle considérée ; de plus, nous supposons que $s_{\max} << N$ (par exemple, pour $N = 256 \times 256$ et $s_{\max} \leq 50, s_{\max}/N \leq 0.01$).

En ce qui concerne la construction des ML, notons que le calcul de la TOC à une échelle donnée nécessite $O(N \log N)$ opérations (en utilisant une FFT), et que la détermination des MM aux échelles considérées est fait en $O(N)$ opérations. En notant $N_1 < N$ le nombre de MM à l'échelle la plus fine $s = 1$, la construction des ML est réalisée en $O(N_1)$ opérations. Ainsi, le coût théorique de la construction des ML est $O(N \log N)$.

En ce qui concerne la construction des LC, le coût du calcul des cartes est $O(N \log N)$. En notant N_1 (resp. $N_{\max}$) le nombre de maxima de Canny à l'échelle $s = 1$ (resp. $s = s_{\max}$), le coût associé à la construction des lignes de Canny est $O(N_1)$ dans le sens des échelles croissantes et $O(N_{max})$ dans le sens des échelles décroissantes. Vu que $N_{max} < N_1 < N$, la complexité de la construction des lignes de Canny est $O(N \log N)$.

A titre d'exemple nous mentionnons des temps de calculs obtenus sous *matlab* (Tab. 5.2).

Ces résultats sont en accord avec la complexité $O(N \log N)$ de la construction des ML ou des LC (Fig. 5.6). Relevons également qu'une fois les maxima déterminés (MM ou maxima de Canny), l'étape de construction des lignes de maxima conduit à un certain surcoût en temps de calcul effectif, du fait de la complexité de construction d'une ligne de maxima (recherche de maxima à l'échelle supérieure, différents tests, etc.).

TAB. 5.2 – Temps de calcul de la construction des LC et des ML, (en secondes, pour des images de taille $n \times n$, $1 \leq s \leq s_{\max} = 10$).

	LC				ML			
n	64	128	256	512	64	128	256	512
Carré	2.06	7.35	32.1	144	0.23	0.85	5.51	57.5
Carré bruité	3.73	25.1	346	5911	0.41	3.75	60.8	939
Barbara	X	X	168	22291	X	X	24.5	433

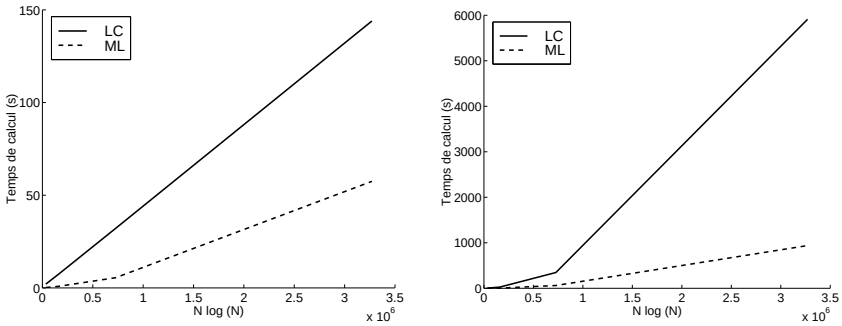


FIG. 5.6 – Temps de calcul effectif en fonction de $N \log N$, pour un carré en (A), et pour un carré bruité en (B) (les images étant de taille $n \times n$, la taille des données vaut $N = n^2$).

5.4 Synthèse sur les lignes de maxima en 2D

Nous étudions ici les propriétés des maxima multiéchelles au sens de Canny (MC) et des modules maxima (MM) de l'ondelette ΔG (chap. 3). Etant donné que ces maxima sont associés à des singularités présentes dans l'image, les lignes de maxima correspondantes permettent de suivre ces singularités dans l'espace-échelle. Deux choses apparaissent alors comme essentielles : d'une part la dérive spatiale des ML, et d'autre part l'évolution de la réponse en fonction de l'échelle. En outre, ces lignes permettent de calculer certaines grandeurs caractéristiques, telles que la régularité Lipschitzienne associée à une singularité, ou l'échelle caractéristique associée à un certain objet.

5.4.1 Quelques propriétés des Maxima Lines

Dérive spatiale des ML

La dérive spatiale d'une ML en 2D est donnée par $s \mapsto (x(s), y(s))$ pour $s > 0$. Pour des ML traversant suffisamment d'échelles, cette notion apparaît comme intéressante. Relevons qu'aux échelles fines, les MM sont localisés au voisinage des singularités, tandis qu'à des échelles plus grandes, les MM proviennent de l'interaction entre différentes singularités. Ainsi la dérive spatiale des ML s'explique par la somme des influences des différentes singularités présentes dans l'image sur la TOC. Cette interaction est complexe, mais nous retrouvons les résultats vus en 1D, à savoir que pour un objet délimité par des frontières, nous observons certaines ML dérivant vers l'extérieur, et d'autres ML dérivant vers l'intérieur, jusqu'à fusionner à une certaine échelle.

Evolution de la réponse le long des ML [87]

Aux échelles les plus fines, une ML permet d'identifier la singularité présente dans l'image, dont la localisation est donnée par l'origine de la ML (x_0, y_0) (MM à l'échelle $s = 1$). Il est possible d'estimer ponctuellement la régularité Lipschitzienne α en (x_0, y_0) d'après la relation

$$|Wf(x(s), y(s), s)| \leq Cs^{\alpha} \tag{5.3}$$

($|Wf(x(s), y(s), s)|$: module du coefficient d'ondelette, normalisé L^1, le long d'une ML).

Alors qu'aux échelles fines, les ML proviennent souvent d'une unique singularité (dont le cône d'influence contient la ML), la situation est différente à des échelles plus grandes. Plus précisément, nous constatons un phénomène analogue à celui observé pour la TOC d'un créneau 1D utilisant l'ondelette (1D) ΔG. Deux ML – associées à deux singularités différentes – fusionnaient à une certaine échelle. Autour de cette échelle, nous constatons qu'en plus de fusionner, elles admettent un pic : la réponse associée le long de ces ML admet un maximum local (la valeur du maximum est identique, puisqu'il y a fusion). De surcroît, la ML issue de la fusion s'avère stable spatialement pour des échelles immédiatement supérieures (pour $s^* \leq s \leq s^* + 2$).

Du point de vue théorique, il serait intéressant de montrer les liens existants entre la jonction de ML et le pic de ML impliquées dans cette jonction. Nous constatons empiriquement qu'aux échelles intermédiaires, il existe effectivement des points dans l'espace-échelle qui correspondent simultanément à des jonctions de ML et à des pics de ML. Nous verrons ultérieurement des applications de ces jonctions de ML admettant un pic en détection de points d'intérêt (chap. 7, [30–32]).

Synthèse

Il est important de noter que la dérive spatiale et la réponse associée à une ML sont reliées. Schématiquement, dans le cas où deux ML se rapprochent, soit elles renforcent leur

réponses respectives jusqu'à fusionner, soit l'une d'elles domine nettement, auquel cas l'autre s'interrompt (la première subissant une légère dérive et une légère atténuation). Suivant les échelles considérées, la réponse $s \mapsto Wf(.,.,s)$ le long d'une ML a un comportement différent :

- aux petites échelles, la dérive spatiale est liée au support de l'ondelette qui s'élargit, et la réponse évolue en fonction de la régularité Lipschitzienne ;
- à des échelles intermédiaires, la dérive spatiale est plus importante, et la réponse subit diverses fluctuations liées à l'interaction entre différentes ML ;
- aux plus grandes échelles, tout se passe comme s'il n'y avait qu'une seule singularité ; la dérive spatiale est similaire à celle aux petites échelles, et la réponse tend vers zéro.

5.4.2 Quelques propriétés des lignes de Canny

Dérive spatiale

Une première chose que nous constatons en pratique est que les LC dérivent relativement peu, ce qui facilite leur construction en pratique. Si deux LC voisines sont associées au même type de contour (vertical par exemple) elles tendent à se renforcer puis à fusionner, comme si à partir d'une certaine échelle, il n'y avait qu'un seul contour présent. Si ce sont deux LC associées à des contours différents (un horizontal et un vertical, ou deux verticaux de signes opposés), l'une des deux lignes tend à s'interrompre tandis que l'autre perdure. Relevons également que les lignes associées à des bords sont plus stables à travers les échelles que celles associées à des coins.

Evolution de la réponse le long des lignes de Canny [87]

Aux échelles fines, le comportement de Mf le long d'une LC peut être relié à la régularité Lipschitzienne ponctuelle. En effet, les singularités associées aux MC sont des points de contours. Or un contour présente deux directions particulières, l'intensité f étant irrégulière dans la direction du gradient et régulière dans la direction orthogonale. De plus, nous disposons de la relation

$$Mf(x(s), y(s), s) \leq Cs^{\alpha} \quad \text{où} \quad Mf = \sqrt{W_x f^2 + W_y f^2} \tag{5.4}$$

qui est une quasi-égalité aux échelles fines, vu que les LC suivent ces contours à travers les échelles. Ceci permet alors de donner une estimation de la régularité Lipschitzienne en tout point de contour. Remarquons que si les relations (5.3) et (5.4) diffèrent (approches ML et de Canny), elles conduisent néanmoins toutes les deux à des estimations correctes de α, comme nous le verrons en 6.2 et 6.3.

Afin d'étudier le comportement global de la réponse en fonction de l'échelle, nous considérons deux images géométriques (carré et disque de taille 128×128, Fig. 5.7 (A–B)) auxquelles un bruit blanc Gaussien a été ajouté. Le détecteur multiéchelles de Canny permet de mettre en évidence certains points de contours plus ou moins marqués, (Fig. 5.7 (A–B')). Les points

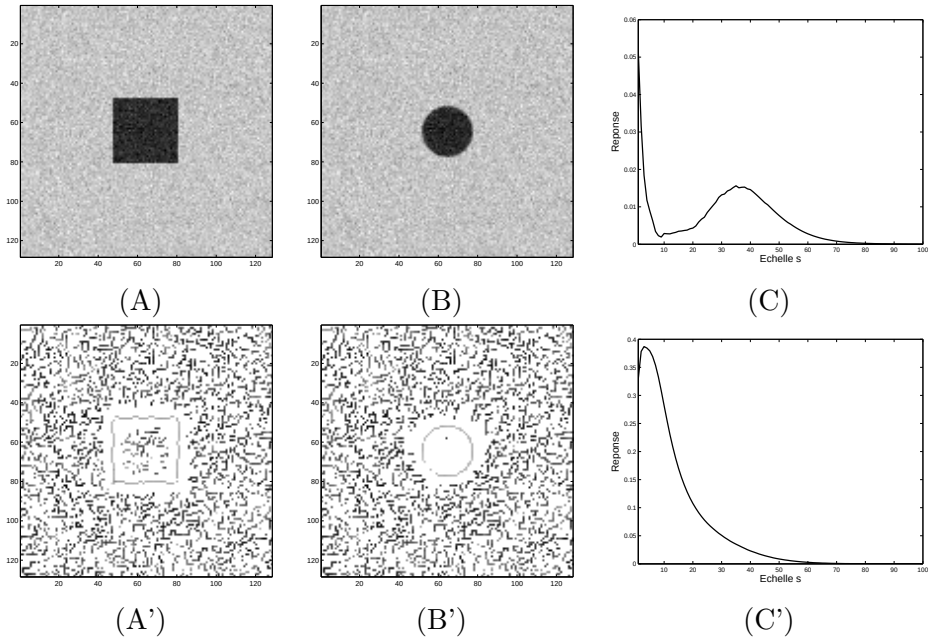


FIG. 5.7 – Approche de Canny dans un cadre bruité : (A) Carré bruité (SNR= $20dB$) ; (B) Disque bruité (SNR= $20dB$) ; (A', B') MC associés. Etude de la réponse en fonction de l'échelle, le long de LC relatives au disque bruité : (C) Ligne associée à du bruit ; (C') Ligne associée à un contour significatif.

de contours liés au bruit peuvent être éliminés sur la base du nombre d'échelles traversées, de la régularité Lipschitzienne ponctuelle associée ou de l'amplitude de la réponse à l'échelle la plus fine. Concernant les lignes associées à du bruit, soit elles s'interrompent à une certaine échelle, soit elles se joignent à d'autres lignes associées à de plus fortes réponses (Fig. 5.7 (C)). En ce qui concerne les lignes associées à des points de contour, nous observons que la réponse croît aux fines échelles (comportement lié à la régularité Lipschitzienne), atteint un pic (maximum local de la réponse) à une certaine échelle, puis tend vers zéro. Relevons que ce pic est relatif à l'importance du point de contour par rapport au voisinage. Le lien entre des objets présents dans l'image et des échelles pouvant leur être associées permet de qualifier celles-ci de *caractéristiques* (voir section suivante).

5.4.3 Maxima Lines et lignes de Canny

D'abord, notons que si la continuité des ML de l'ondelette ΔG est assurée théoriquement (comme conséquence du principe du maximum), le résultat est moins évident pour les LC. La stabilité de la direction du gradient joue un rôle clé : nous avons remarqué auparavant qu'à mesure que l'échelle augmente, le gradient ne reste pas nécessairement dans la même direction ; mais tend à se stabiliser assez rapidement.

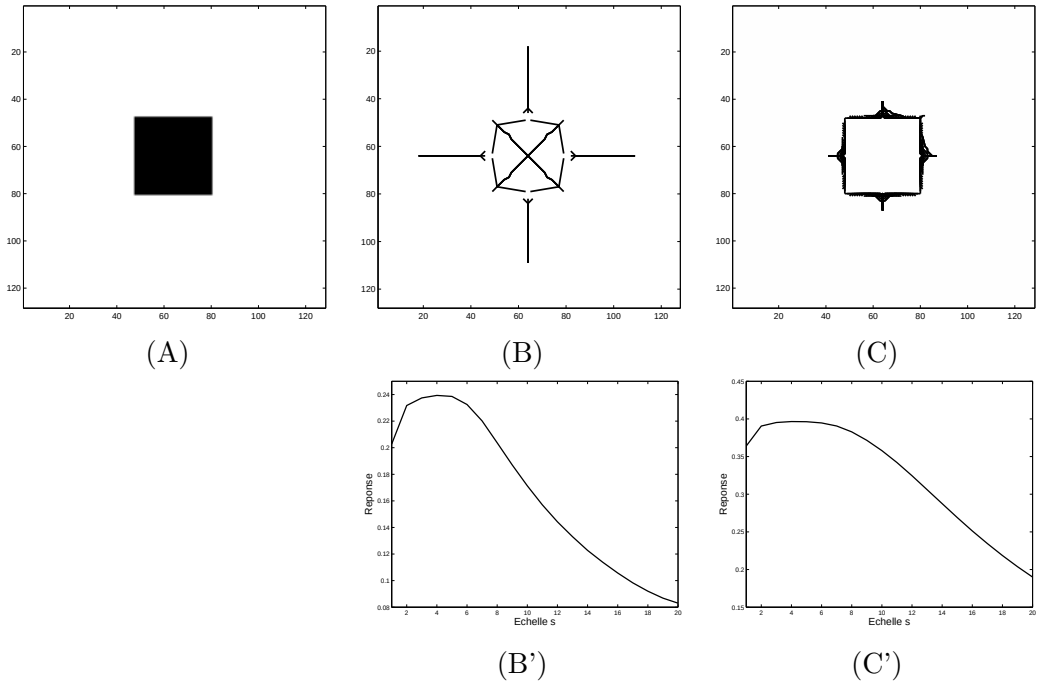


FIG. 5.8 – (A) Image représentant un carré ; (B) ML associées (les ML partent toutes des bords du carré, certaines dérivent vers l'intérieur, d'autres vers l'extérieur) ; (C) LC associées ; (B') Evolution de la réponse le long d'une ML ; (C') Evolution de la réponse le long d'une LC.

Du point de vue pratique, pour la construction des LC, il est aisé de chaîner les MC d'une échelle à une autre ; pour les ML, le chaînage des MM d'une échelle à une autre est parfois plus délicat, étant donné qu'il peut exister une grande dérive spatiale entre deux échelles successives. Par ailleurs, il est intéressant de remarquer que dans le cas des ML, l'opérateur est basé sur les dérivées secondes (ondelette ΔG) tandis que dans le cas des LC, il est basé sur des dérivées premières (ondelette gradient). Ceci explique notamment que les ML et les LC ont des comportements assez différents à mesure que l'échelle augmente.

Afin de mieux cerner cette différence, nous comparons ces deux types de lignes sur l'exemple simple du carré (Fig. 5.8 (A)). Le fait que les LC soient stables spatialement s'explique par le fait que les contours les plus significatifs sont visibles à toutes les échelles. Quant aux ML, leur dérive peut être exploitée pour mettre en évidence la structure globale du carré. Plus précisément, aux fines échelles, les ML sont reliées aux bords du carré, puis dérivent (vers l'extérieur ou l'intérieur du carré). Nous remarquons alors qu'à de plus grandes échelles, les ML dérivant vers l'intérieur sont reliées au carré : il y a jonction de ML. En outre, nous pouvons comparer l'évolution de la réponse associée à ces lignes en fonction de l'échelle s ($|Wf|$ pour les ML, Mf pour les LC).

En ce qui concerne les ML, les réponses associées aux ML intérieures correspondent toutes à l'évolution représentée sur la figure 5.8 (B'), en particulier la réponse admet le même maxi-

mum local. Nous relevons alors que la localisation associée correspond au centre du carré, tandis que l'échelle associée est proportionnelle à la taille de l'objet (ce qui est conforme à la théorie de Scale-Space, étant donné que nous utilisons une TOC normalisée L^1, cf. chap. 4).

En ce qui concerne les LC associées à l'exemple du carré, la réponse le long d'une LC admet également un maximum local, et l'échelle associée est alors identique à celle issue des ML (Fig. 5.8 (C')). Ainsi, les LC se focalisent sur les contours de l'objet (bords du carré), tandis que les ML, par la jonction de ML intérieures, mettent en évidence l'objet entier (le carré). De plus, certaines échelles apparaissent comme remarquables : celles pour lesquelles la réponse atteint un maximum local le long d'une ligne de maxima. Nous précisons cela maintenant.

Objets présents dans l'image – Caractéristiques associées

Distinguons deux applications des lignes de maxima : d'une part la détection de certaines structures (1D ou 2D), et d'autre part leur caractérisation (calcul de grandeurs caractéristiques des structures). Nous avons vu en 5.4.3 que les lignes de maxima permettent de détecter de telles entités : des structures 1D (contours) par les LC (Fig. 5.8 (B)) et des structures 2D par les ML (Fig. 5.8 (C)). D'une part, les LC permettent de détecter des points de contours, le nombre d'échelles traversées (persistance de la ligne) indiquant le caractère plus ou moins marqué du contour dans l'image. Une localisation précise de ceux-ci est donnée par les extrémités des lignes (aux échelles fines), et les LC donnent une estimation de la régularité Lipschitzienne en ces points de contours. D'autre part, les ML permettent de détecter des objets (critère de jonction) et d'extraire certaines caractéristiques de l'objet : une localisation et une échelle qui correspondent à sa taille (notion d'échelle caractéristique), ainsi que la régularité Lipschitzienne associée aux extrémités des ML, lesquelles sont situées près des frontières de l'objet.

Il importe de remarquer que si une structure 2D peut être délimitée par des contours (comme pour le carré), ce n'est pas toujours le cas, comme par exemple pour une surface Gaussienne bidimensionnelle. Par ailleurs, étant donné une structure 1D, il n'est pas évident de définir si elle appartient à tel ou tel objet 2D, comme par exemple pour une ligne isolée. Cependant, dans le cas où les structures 2D sont délimitées (au moins partiellement) par des singularités, la notion de jonction de ML permet d'associer ces singularités à des objets. De surcroît, ces singularités peuvent être également détectées comme des points de contours par les LC. Une fois cette association effectuée entre singularités, objets et points de contours, nous pouvons alors comparer les grandeurs caractéristiques associées à ces lignes.

Les échelles correspondant à un maximum local de la réponse le long d'une ligne de maxima sont remarquables, au sens où elles correspondent à la taille d'une certaine région. Dans le cas des ML, une telle échelle est proportionnelle à la taille d'un objet, tandis que dans le cas des LC, celle-ci correspond à une certaine zone d'influence, où le contour associé prédomine dans l'image. Par exemple, dans le cas du carré, en notant l la longueur d'une arête, s_{ML} une échelle

remarquable le long d'une ML intérieure, s_{LC} une échelle remarquable le long d'une LC, nous obtenons $s_{LC} \approx l/2 \approx s_{ML}$. La relation $s_{LC} \approx l/2$ s'explique par le fait que l'influence d'un bord est limitée par celle du bord opposé, et nous étudierons en détail la relation $s_{ML} \approx l/2$ au chap. 7. Quant à elle, la régularité Lipschitzienne permet d'identifier le type de singularité associée à l'origine de la ligne de maxima, ce qui est précisé dans la suite. En particulier, nous verrons que l'estimation d'après les LC est souvent plus pertinente.

A ce stade, nous avons vu que les ML possèdent une structure globale liée au contenu de l'image. En particulier, il existe des points du *Scale-Space* où certaines ML fusionnent et admettent simultanément un pic. Ceci aboutit au critère de jonction, qui permet alors de mettre en évidence certains objets 2D. Les ML associées à une telle jonction permettent alors de décrire sommairement cet objet. Par ailleurs, les LC correspondent à des points de contours, et l'association entre ceux-ci et des structures de l'image n'est pas immédiate, surtout avec des structures 2D : les LC ne permettent pas de mettre en évidence de tels objets. Néanmoins, le lien avec des structures 1D est possible, grâce à des algorithmes de reconstruction des contours par des segments [45, 69].

5.5 Synthèse : décompositions, maxima et singularités

Schématiquement, les approches présentées s'appuient sur trois points (l'exemple pour l'approche par ML est mis entre parenthèses) :

1. Définir un ensemble de cartes associées à différentes échelles (TOC),
2. Définir des points remarquables, nommés par le terme générique de maxima (MM),
3. Construire des lignes reliant ces points d'une échelle à une autre (ML).

L'intérêt de construire ces lignes est qu'elles permettent de suivre des singularités à travers les échelles. En effet, les coefficients des décompositions utilisées sont faibles là où l'image est régulière, et forts là où elle est irrégulière. Ainsi, les maxima sélectionnés correspondent à certaines singularités qui pourront être caractérisées par la suite (chap. 6 et 7).

Notons enfin qu'il est envisageable de généraliser cette approche, en particulier à l'aide de processus de diffusion. De telles méthodes pourraient alors se révéler intéressantes en termes d'applications : des lignes de maxima plus générales permettraient de formuler des détecteurs de structures caractéristiques dans l'image, à partir de processus de diffusion anisotropes ou invariants à l'affine.

Chapitre 6

Détection de structures monodimensionnelles en 2D

6.1 Singularités dans une image

Définir quels sont les éléments les plus remarquables dans une image donnée n'est pas une chose immédiate : si nous le demandons à différents observateurs, nous n'obtiendrons pas de consensus. Cependant, il apparaît certaines structures importantes, telles que des coins (singularités ponctuelles), des lignes de contours plus ou moins nets (singularités curvilignes), un disque, une surface Gaussienne (structures isotropes, pouvant être légèrement déformées), des contours flous ou des objets déformés présentant une forte anisotropie. Relativement à ces singularités, il se dégage globalement deux types de structures :

- monodimensionnelles, telles que des lignes de contours,
- bidimensionnelles, correspondant à des objets que nous percevons visuellement.

Notons d'emblée que l'analyse des contours ne permet pas de caractériser simplement les objets qu'ils délimitent, ce qui motive le développement de méthodes plus sophistiquées. Une idée centrale développée dans cette thèse est que de telles structures peuvent être identifiées de manière automatique, par détection de points et d'échelles remarquables (telles que celles vues au chap. 4). En outre, celles-ci pourront être associées à des grandeurs caractéristiques de la géométrie ou du type de singularité.

Dans ce chapitre, nous nous focalisons sur les singularités associées à des structures monodimensionnelles (l'étude des structures bidimensionnelles étant l'objet du chapitre 7). A cet effet, nous rappelons d'abord le principe de détection de telles structures, puis nous expliquons comment les différentes lignes de maxima (ML et LC, détaillées au chapitre 5) permettent de détecter et de caractériser de telles structures monodimensionnelles. En particulier, nous montrerons que la régularité Lipschitzienne (chap 3, def. 17) constitue une grandeur caractéristique robuste, que ce soit au bruit ou à des transformations géométriques. Enfin, nous effectuerons une synthèse sur le caractère invariant ou robuste des primitives extraites, chose importante tant pour les structures monodimensionnelles que bidimensionnelles.

Détection de structures monodimensionnelles

Les outils que nous avons présentés au chapitre 3 (ML et LC) permettent de détecter différentes entités (structures présentes dans l'image). Nous privilégions ici celles qui sont monodimensionnelles, et afin de les détecter, nous cherchons à identifier les points de l'image qui leur appartiennent. Au chapitre 4, nous avons présenté différents opérateurs permettant de détecter des singularités spécifiques : bords, lignes, coins. Quant à elles, les approches utilisant des lignes de maxima – ML ou LC – détectent des singularités associées à différents maxima (MM, MC). Par la suite, le type de singularité (bord, ligne ou coin) pourra être identifié à l'aide de la régularité Lipschitzienne.

Pour détecter de tels points appartenant à des structures monodimensionnelles, il importe de définir un critère adéquat. Une première possibilité consiste à choisir les MM d'une TOC, ce qui conduit typiquement à des coins, des points de contours où la courbure est importante, ou des points isolés remarquables. Une autre option consiste à considérer les MC à l'échelle la plus fine, ce qui conduit à des points appartenant aux contours de l'image. De plus, les lignes de maxima associées mesurent leur importance : étant donné que ce sont des données stables dans l'image, ces lignes permettent de préciser l'importance d'un maximum associé à leur origine tout en conservant une localisation précise : par exemple, si un maximum localisé en (x, y) est associé à une ligne qui perdure jusqu'à l'échelle $s = 5$, cela indique l'importance de ce maximum par rapport au voisinage. Un (léger) seuillage sur la réponse dans ce voisinage élimine enfin les points peu significatifs. Finalement, le critère de sélection s'exprime comme :

- (x, y) est un maximum (MM, MC), appartenant à une certaine ligne de maxima L ;
- L traverse suffisamment d'échelles (typiquement 5) ;
- $|Wf(x, y)|$ (ou $Mf(x, y)$) dépasse un certain seuil.

6.2 Caractérisation de structures monodimensionnelles

6.2.1 Calcul des régularités

Les régularités Lipschitziennes α ponctuelles associées à des singularités de l'image correspondent souvent à des valeurs comprises entre -2 et 1, et dépendent du type de contour :

$$\begin{cases} \text{Frontière} & \alpha = 0 \\ \text{Ligne} & \alpha = -1 \\ \text{Point isolé} & \alpha = -2 \end{cases}$$

Ici, nous nous intéressons aux régularités ponctuelles, ce qui contraste avec des approches globales d'estimation des régularités [4]. Diverses applications, telles que le recalage de signaux [12], motivent l'étude plus approfondie de la régularité Lipschitzienne. Concernant l'estimation de α, les méthodes d'ondelettes permettent de détecter certains points singuliers, et également d'estimer la régularité Lipschitzienne en ces points. Ces derniers peuvent être, par exemple,

des MM à l'échelle la plus fine d'une TOC, ou des points de contours. Dans ces deux cas, ces maxima peuvent être associés à des chemins traversant les échelles (ML, LC) qui donnent alors une information supplémentaire sur la singularité sous-jacente.

En notant M_s la réponse de l'opérateur le long d'un tel chemin, nous avons une relation du type (cf. eq. (5.3) et (5.4))

$$\log M_s \leq \alpha \log s + C \tag{6.1}$$

où C est une constante ($C = 0$ pour la normalisation L^1 que nous utilisons ici). Cette inégalité est une quasi-égalité aux échelles fines (et d'autant meilleure que la singularité est bien suivie), ce qui permet alors une estimation de α (Fig. 6.1). Enfin, notons que des estimations plus générales de α sont possibles [59]. Ici, nous nous focalisons sur les méthodes par ML et LC, pour lesquelles les régularités ponctuelles sont finement estimées en des points particuliers.

Tant avec les ML que les LC, les valeurs obtenues concordent avec celles attendues sur une image géométrique (Fig. 6.1 (A–A")) et permettent d'estimer des valeurs de la régularité Lipschitzienne en un grand nombre de points sur une image naturelle (Fig. 6.1 (B–B")). Relevons au passage que la distribution des régularités est bien différente entre ces deux images (Fig. 6.2). Maintenant, nous montrons que l'estimation de la régularité α fondée sur les LC apparaît comme la plus adaptée, au sens où des régularités sont correctement estimées, et où le nombre de points obtenus sur une image naturelle est important. Par la suite, nous montrerons également que cette estimation est robuste (section 6.3).

6.2.2 Comparaison des méthodes d'estimation de la régularité

Afin de comparer les approches ML et LC, nous considérons l'image géométrique vue précédemment, à laquelle nous appliquons une déformation affine ou une rotation, ce qui conduit à trois images déformées contenant les mêmes régularités (Fig. 6.3). Nous calculons alors les régularités par ML et LC, ce qui permet de comparer les valeurs obtenues en certains points singuliers, et observons que celles-ci s'avèrent globalement meilleures pour les LC (Tab. 6.1). Par ailleurs, du point de vue quantitatif, l'approche par LC permet de mettre en évidence un plus grand nombre de points singuliers que celle par ML. En considérant différentes images (un bruit blanc Gaussien, une image naturelle, une géométrique), nous représentons comment se comportent le nombre de MC ainsi que le nombre de MM en fonction de l'échelle (Tab. 6.2, Fig. 6.4). Nous remarquons alors que les MC sont bien plus nombreux que les MM, et qu'à mesure que l'échelle augmente, les MC persistent mieux que les MM. Ainsi, l'estimation par les LC apparaît comme mieux adaptée que par les ML.

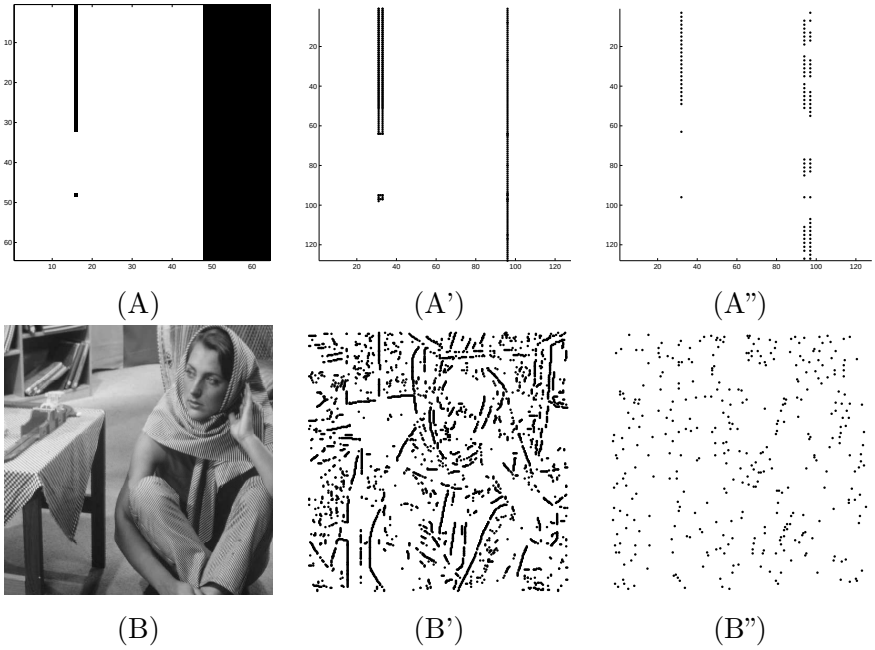


FIG. 6.1 – (A) Image géométrique contenant des régularités connues *a priori* ; (B) Image naturelle dont les régularités sont inconnues *a priori* ; (A') et (B') Régularités Lipschitziennes calculées d'après les LC ; (A") et (B") Régularités Lipschitziennes calculées d'après les ML.

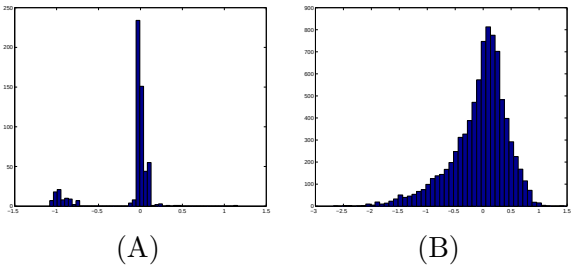
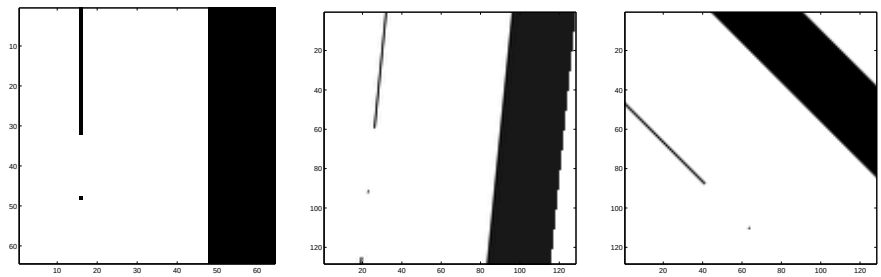


FIG. 6.2 – Histogrammes des régularités Lipschitziennes α (estimées d'après les LC) : (A) Image géométrique (cf. Fig. 6.1 (A)) (B) Image *Barbara* (cf. Fig. 6.1 (B)).

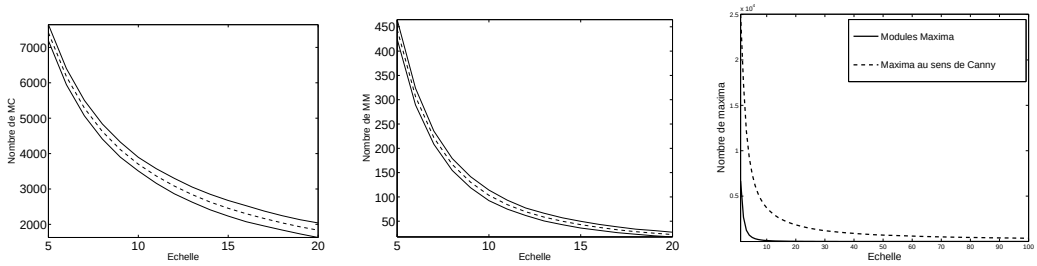


(A_1) (A_2) (A_3)

FIG. 6.3 – Images géométriques : (A_1) originale ; (A_2) ayant subi une déformation affine ; (A_3) ayant subi une rotation.

TAB. 6.1 – Comparaison des estimations de α d'après les ML et les LC, sur des images géométriques.

Image	(A_1)	(A_2)	(A_3)
Frontière ($\alpha = 0$)			
α_{ML}	0.15	-0.04	0.09
α_{LC}	0.08	0.08	0.04
Ligne ($\alpha = -1$)			
α_{ML}	-0.98	-0.90	-0.74
α_{LC}	-1.00	-0.97	-0.87
Point Isolé ($\alpha = -2$)			
α_{ML}	-1.98	-1.74	-1.87
α_{LC}	-2.00	-1.91	-1.78



(A) (B) (C)

FIG. 6.4 – Evolution des nombres de MC et de MM en fonction de l'échelle – étude réalisée en moyenne sur 100 signaux de taille 256x256, sur des bruits blanc Gaussiens d'écart-type $\sigma = 100$; notations : s désigne l'échelle, m_{mc} et σ_{mc} la moyenne et l'écart-type relatifs aux MC m_{mm} et σ_{mm} la moyenne et l'écart-type relatifs aux MM. (A) $m_{mc} = f(s)$ et $m_{mc} \pm 2\sigma_{mc} = f(s)$ (MC) ; (B) $m_{mm} = f(s)$ et $m_{mm} \pm 2\sigma_{mc} = f(s)$ (MM) ; (C) $m_{mc} = f(s)$ et $m_{mm} = f(s)$.

TAB. 6.2 – Nombre de modules maxima (MM) et de maxima au sens de Canny (MC) pour différentes images (taille 256x256) : un bruit blanc Gaussien, l'image naturelle *Barbara*, et une image géométrique représentant un carré.

Echelle	Bruit blanc		Barbara		Carré	
s	MM	MC	MM	MC	MM	MC
1	6658	25154	5128	17372	232	810
2	2682	17640	1470	9567	152	761
3	1258	12128	628	6878	33	333

6.3 Robustesse de la régularité Lipschitzienne

6.3.1 Influence d'une déformation sur la régularité

Proposition 15. *Etant donné une image f et une matrice $B \in GL_2$ nous considérons l'image déformée g définie par :*

$$\forall X \in \mathbb{R}^2, f(X) = g(BX) \tag{6.2}$$

Supposons que f soit α-Lipschitz, avec $\alpha \in [0, 1[$. En notant $||.||$ une norme sur $\mathbb{R}^2$ et $||B||$ la norme de la matrice B, il existe une constante $K > 0$ telle que :

$$\begin{aligned} |g(X) - g(Y)| = |f(B^{-1}X) - f(B^{-1}Y)| &\leq K|B^{-1}X - B^{-1}Y|^\alpha \\ &\leq \left(K||B^{-1}||^\alpha\right)|X - Y|^\alpha \end{aligned}$$

Ainsi, g est également α-Lipschitz, de sorte que la régularité Lipschitzienne est conservée par déformation affine.

Interprétation : tant qu'une déformation ne change pas la nature d'une singularité (comme une ligne légèrement déformée par exemple), la régularité Lipschitzienne associée n'est pas modifiée. Par ailleurs, vu que l'estimation de α est basée sur des lignes de maxima, il importe que ces lignes ne soient pas détruites par la transformation (problème abordé à la section précédente).

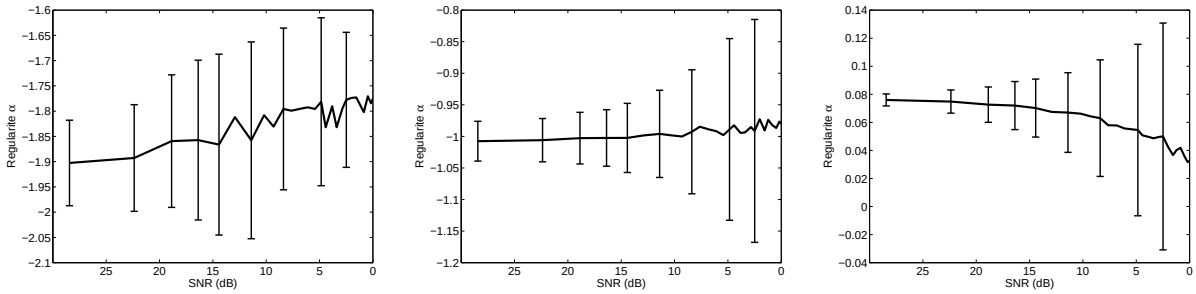
Du point de vue pratique, il est alors intéressant de quantifier la robustesse de la régularité Lipschitzienne, estimée d'après les LC. Nous évaluons cette robustesse d'abord vis à vis du bruit, puis vis à vis de transformations géométriques, comme des rotations et des déformations affines. Dans ce qui suit, afin de synthétiser l'information associée à une image, nous traitons les régularités obtenues comme des variables aléatoires, ce qui permet d'utiliser le vocabulaire statistique (moyenne, écart-type, estimation, biais).

Relevons que dans les exemples qui suivent, où des singularités sont présentes, nous obtenons souvent des valeurs négatives de α. Nous observons alors qu'empiriquement, la proposition 15 reste valable pour $\alpha < 0$ (son extension théorique est en cours d'étude).

6.3.2 Influence du bruit

Nous considérons une image 64×64 contenant différentes régularités : $\alpha = -2$ pour le point isolé, $\alpha = -1$ pour la ligne et $\alpha = 0$ pour la frontière (Fig. 6.5 (A)) ; un bruit blanc Gaussien est alors ajouté (pour chaque niveau de bruit, 100 simulations sont effectuées). Nous calculons les régularités d'après les LC, et isolons (d'après un critère géométrique) celles qui correspondent au point isolé, à la ligne et à la frontière. En l'absence de bruit, les régularités calculées valent -2.003 pour le point isolé, -1.004 pour la ligne, 0.078 pour la frontière.

Nous pouvons alors représenter l'évolution des régularités associées à ces trois types de singularités (Fig. 6.5 (B–C)). Pour le point isolé, la régularité est légèrement instable ; notons l'écart-type non-négligeable, présent dès un faible niveau de bruit. En ce qui concerne la ligne, l'estimation de la régularité est correcte (très proche de -1) et l'estimation stable à mesure que le niveau de bruit augmente. Quant à la frontière, l'estimation est légèrement biaisée (comme dans le cas non-bruité), et s'avère très stable. Globalement, retenons que les régularités sont robustes au bruit.



(A) (B) (C)

FIG. 6.5 – (A–C) Evolution en fonction du niveau de bruit (moyenne et écart-type) des régularités estimées d'après les LC sur une image géométrique. Ces régularités sont calculées : (A) au point isolé, (B) à la ligne et (C) à la frontière.

6.3.3 Influence d'une rotation

Nous considérons deux images, l'une représentant une ligne et l'autre un carré, auxquelles nous appliquons une rotation d'un angle variant de 0 à 90 degrés (Fig. 6.6 (A–B, A'–B')). Nous calculons alors les régularités associées à chaque angle, ainsi que la moyenne et l'écart-type des régularités obtenues. Nous observons que celles-ci ce concentrent autour de 0 pour le carré, et autour de -1 pour la ligne, ce qui est conforme à la théorie. Nous représentons alors la régularité Lipschitzienne en fonction de cet angle de rotation (Fig. 6.6 (A"–B"). Dans le cas de la ligne, l'estimation de α présente un certain biais (de 0.15 environ) pour tous les angles sauf 0 et 90 degrés, pour lesquels l'estimation donne exactement la valeur attendue ($\alpha = -1$). Dans le cas du carré, l'estimation de α est très proche de la valeur attendue, avec un biais compris entre 0.04 et 0.06. Notons par ailleurs que pour une image donnée, il existe une légère hétérogénéité des régularités : les extrémités de la ligne et les coins du carré diffèrent des autres régularités. Ceci explique qu'il existe un certain écart-type, mais celui-ci reste inférieur à 0.1 dans tous les cas. Ainsi, la régularité Lipschitzienne s'avère robuste à la rotation.

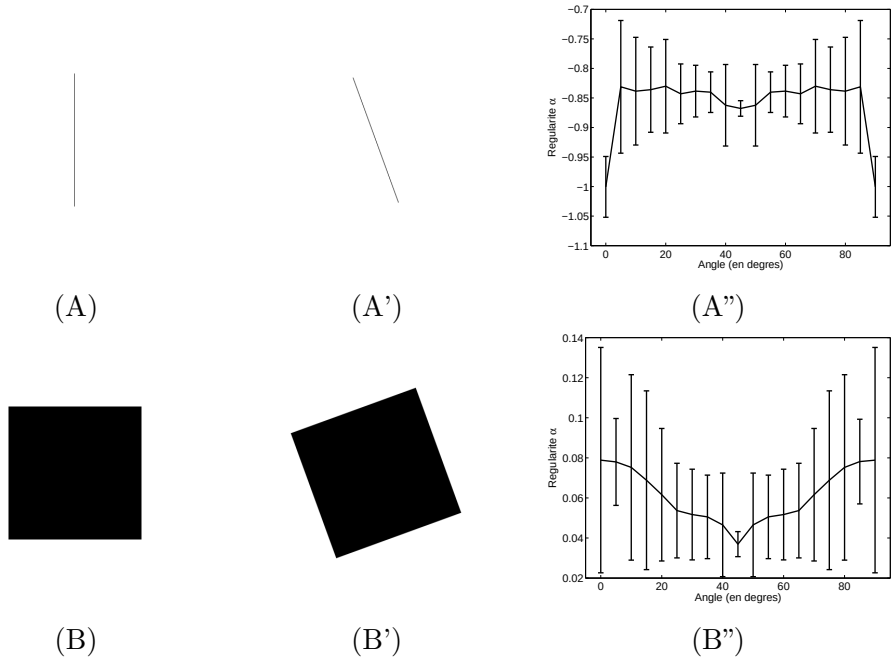


FIG. 6.6 – Influence d'une rotation sur la régularité Lipschitzienne : (A–A') Images originale et déformée par une rotation ($\alpha = -1$) ; (B–B') *idem* ($\alpha = 0$) ; (A") Evolution des régularités Lipschitziennes (moyenne et écart-type) suivant l'angle de rotation appliqué à l'image (A) ; (B") *idem*, l'image originale étant (B).

6.3.4 Influence d'une déformation affine

Nous considérons maintenant deux images de taille 256×256 – une image géométrique et *Barbara* (Fig. 6.7) – auxquelles nous appliquons une déformation affine connue. Puis, nous calculons les régularités Lipschitziennes associées, d'après les LC. Nous apparions alors certains pixels deux à deux, et comparons alors les valeurs des régularités.

Comme nous connaissons la matrice de déformation, l'appariement est aisé : étant donné deux points $p_O^i = (x_O^i, y_O^i)$ (issu de l'image originale) et $p_D^j = (x_D^j, y_D^j)$ (issu de l'image déformée), nous les apparions à condition que la distance Euclidienne entre ces deux points, en exprimant leurs coordonnées dans l'image *originale*, soit suffisamment faible. En pratique, pour la plupart des points appariés, cette distance est inférieure à 1 pixel. Une fois cet appariement effectué, nous nous intéressons à l'écart $d_\alpha = |\alpha_O^i - \alpha_D^j|$ (Tab. 6.3). Nous remarquons alors que les valeurs des régularités sont très proches dans les images originale et déformée.

De plus, nous relevons que la proportion de points appariés pour lesquels $d_\alpha < 0.2$ est élevée (97% et 80%) ; ainsi, dans le cas d'une déformation affine, la régularité Lipschitzienne permet d'identifier correctement le type de singularité.

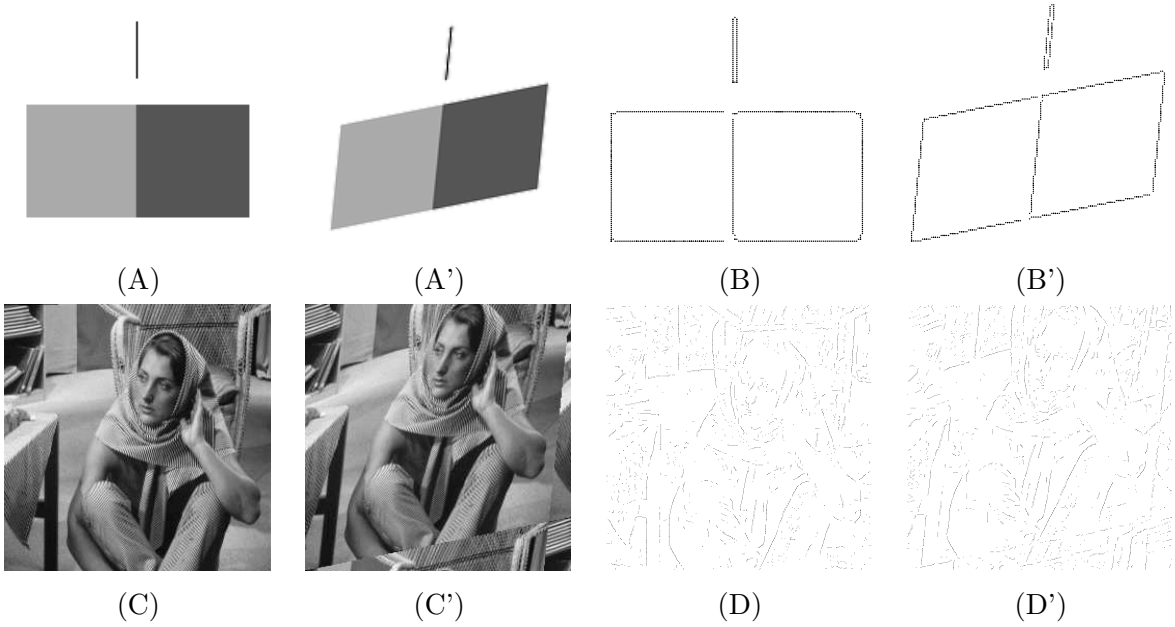


FIG. 6.7 – (A–A') Image géométrique, originale et déformée ; (B–B') Régularités associées. (C–C') Image *Barbara*, originale et déformée ; (D–D') Régularités associées.

TAB. 6.3 – Influence d'une déformation affine sur le calcul des régularités Lipschitziennes. Celles-ci sont calculées en un certain nombre de points de régularité, définis comme l'extrémité d'une LC.

Image	Géométrique	*Barbara*
Nombre de points de régularité		
- dans l'image originale	599	8474
- dans l'image déformée	657	8517
Nombre de de points appariés	599	6929
Ecart d_α		
Valeur moyenne	0.077	0.15
Ecart-type	0.06	0.23
Proportion de points pour lesquels $d_\alpha < 0.2$	97%	80%

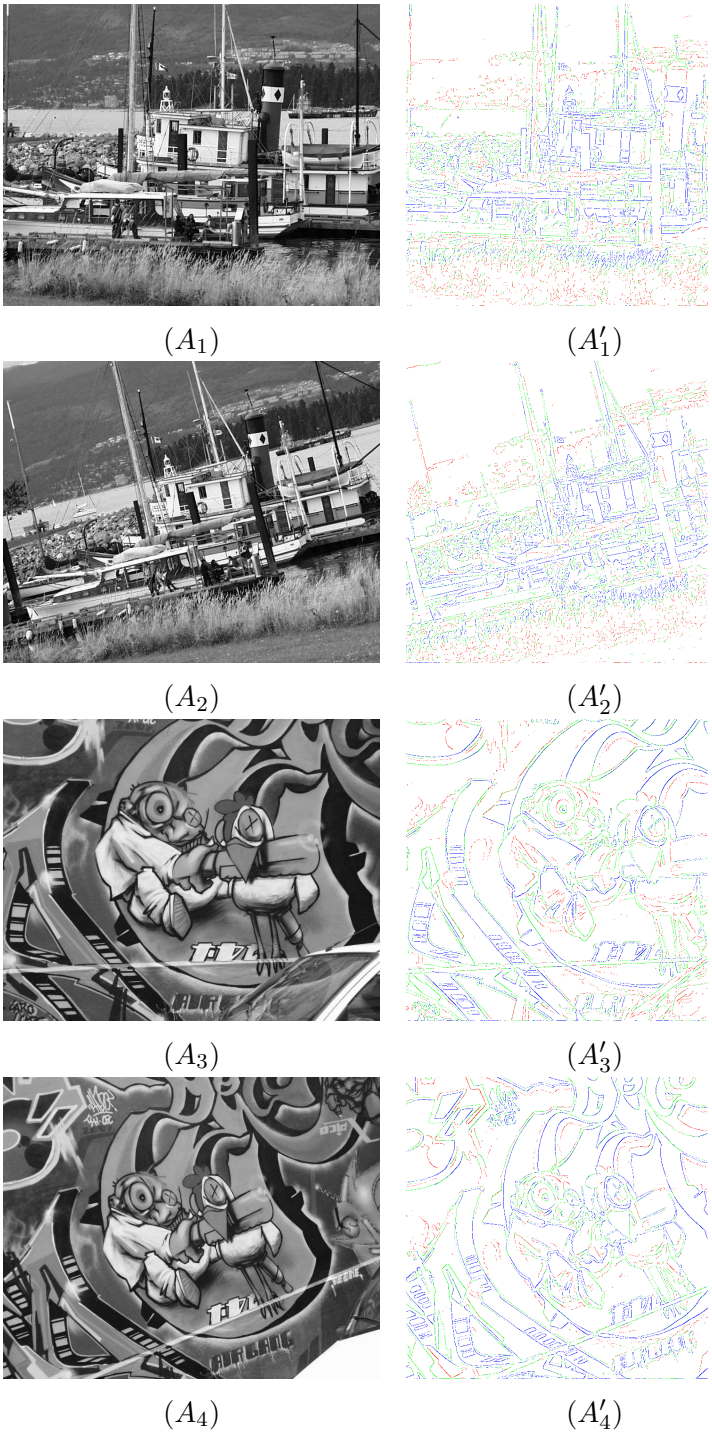


FIG. 6.8 – Estimation de la régularité Lipschitzienne, sur des images ayant subi des transformations géométriques. $(A_1 - A_2)$ Changement d'échelle et rotation ; $(A_3 - A_4)$ Changement de point de vue ; $(A'_1 - A'_4)$ Régularités Lipschitziennes associées (en bleu : frontières ($\alpha \approx 0$), en vert : lignes ($\alpha \approx -1$), en rouge : points isolés ($\alpha \approx -2$)).

6.4 Questions d'invariance et de robustesse

Dans cette section, nous expliquons pourquoi certaines caractéristiques constituent des primitives invariantes ou robustes à certaines transformations de l'image. Deux types d'arguments entrent en ligne de compte :

- d'une part, les données ponctuelles considérées sont robustes (typiquement, la réponse associée est suffisamment forte) ;
- d'autre part, l'opérateur utilisé est invariant par certaines transformations (ainsi un noyau isotrope est invariant par rotation), ou prend en compte certaines d'entre elles (approches multiéchelles pour le changement d'échelle).

6.4.1 Régularité Lipschitzienne

Nous avons montré en 6.3 que la régularité Lipschitzienne est une grandeur robuste au bruit, à la rotation et aux (légères) déformations affines. Maintenant, nous considérons des images photographiques (disponibles à l'adresse http://www.robots.ox.ac.uk/vgg/research/affine) ayant été prises sous des conditions différentes – changement d'échelle et rotation pour l'une, changement de point de vue pour l'autre (Fig. 6.8 ($A_1 - A_4$)). Nous représentons alors les régularités Lipschitziennes ponctuelles, calculées d'après les LC (Fig. 6.8 ($B_1 - B_4$)). Au vu des résultats vus en 6.3, et conformément à la proposition 15, la déformation translate les points où sont calculées les régularités, tandis que les valeurs des régularités sont très proches entre les images originale et déformée. En particulier, nous avons vu que les valeurs proches de -1 et de 0 étaient robustes, ce qui permet d'identifier correctement des points appartenant à des lignes ou à des frontières (ce qui constitue une alternative aux méthodes issues de la vision par ordinateur, vues en 4.2.1).

6.4.2 Maxima Lines et lignes de Canny

Du point de vue théorique, nous avons vu auparavant que dans le cadre d'une transformation affine de l'image, il existe toujours des ML – pourvu que la matrice B de changement de coordonnées soit inversible et constante pour toute l'image (chap. 4, section 4.6.2). Dans un cadre plus général où la transformation affine n'est plus constante, nous ne pouvons pas écrire explicitement la manière dont les ML sont transformées. Cependant, si la transformation peut s'écrire (approximativement) comme une transformation affine dans un certain voisinage, et que ce voisinage inclut la trajectoire de certaines ML, alors les propriétés des ML restent valables. En outre, il est intéressant de voir quel est l'effet de différentes déformations. Rappelons que les MM sont des caractéristiques ponctuelles, et ce à toutes les échelles, puisque la TOC a la même résolution que l'image : l'image est lissée à mesure que l'échelle croît, mais les MM sont toujours bien localisés. Ainsi, en fonction de la régularité de l'opérateur utilisé et de l'amplitude des déformations appliquées à l'image, les singularités peuvent être conservées,

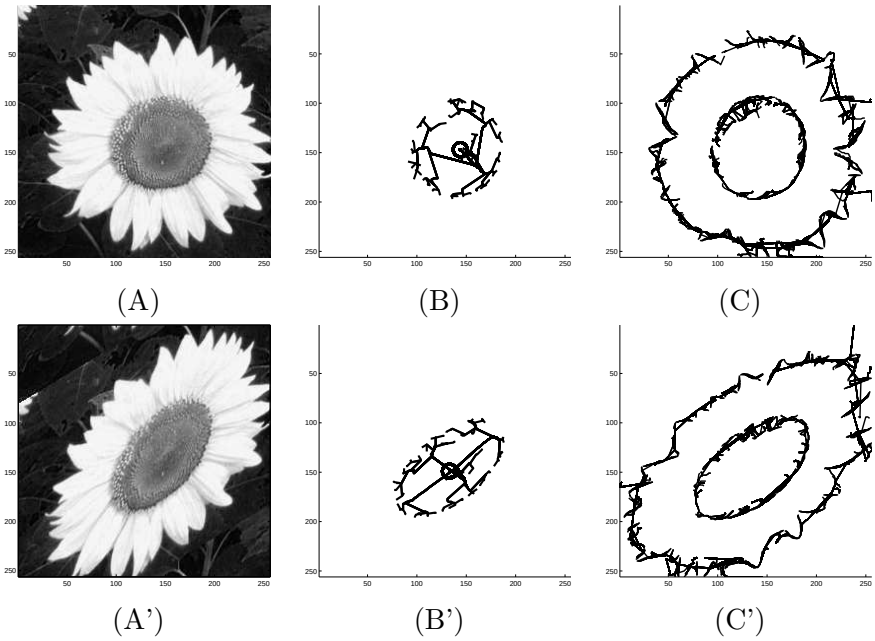


FIG. 6.9 – Effet d'une déformation sur une image naturelle (A–A') Images originale et déformée ; (B–B') Trajectoires des ML associées à l'objet central (les autres ML n'étant pas représentées ici) ; (C–C') Trajectoires des LC associées aux principaux contours.

ce qui conduit alors à des MM multiéchelles associés à ces singularités.

En ce qui concerne les LC, l'existence des lignes de maxima pose autant un problème pour l'image originale que pour l'image déformée. Nous avons néanmoins vu que leur construction empirique ne soulevait pas de difficulté particulière. A cet égard, rappelons que les singularités associées aux LC sont associés à des points de contours, lesquels sont conservés pour de légères déformations (i.e., tant que la nature du contour n'est pas altérée).

Du point de vue pratique, nous considérons une image, à laquelle nous appliquons une déformation affine (Fig. 6.9 (A–A')). Nous construisons alors les ML et les LC, sur les images originale et déformée. En ce qui concerne les ML (Fig. 6.9 (B–B')), nous relevons que les extrémités (localisation des ML aux échelles fines) ainsi que les points de jonction (fusion de ML) sont conservés ; cependant, les trajectoires des ML s'avèrent instables aux échelles intermédiaires. En ce qui concerne les LC (Fig. 6.9 (C–C')), celles-ci sont conservées, ce qui s'explique par leur faible dérive spatiale et le lien avec les contours, lesquels ne sont pas altérés par une légère déformation affine. Globalement, les ML et les LC apparaissent comme des caractéristiques robustes aux déformations affines.

6.5 Conclusion

Dans ce chapitre, nous avons d'abord rappelé le principe de détection de certaines singularités – des points de contour – et leur caractérisation par la régularité Lipschitzienne. Puis, nous avons montré que la régularité Lipschitzienne était une caractéristique robuste au bruit ainsi qu'à des déformations géométriques.

En comparant les propriétés des ML et des LC, nous avons montré que les LC étaient les plus adaptées au calcul des régularités. Ensuite, nous avons vu que lorsque l'image subit une déformation géométrique, les ML et les LC sont déformées (certaines pouvant être détruites). Ainsi, nous pouvons considérer que les ML, les LC ainsi que la régularité Lipschitzienne, de part leurs propriétés d'invariance à différentes transformations de l'image, sont pertinentes pour des applications telles que la vision par ordinateur.

Un problème important en traitement d'image consiste à formuler des outils permettant à la fois de détecter certaines entités dans l'image (contours, coins, *blobs*) et d'extraire des caractéristiques de ceux-ci. Afin que les caractéristiques soient pertinentes, il importe que de telles entités puissent être correctement identifiées, quelle que soit la transformation appliquée à l'image (ceci sera précisé par la notion de *répétabilité*, annexe A). Par exemple, l'analyse de deux images correspondant à une même scène, prises sous des angles de vue différents, doit mettre en évidence des entités pouvant être mises en correspondance (au moins en partie) [100, 101, 112]. Au chapitre suivant, nous détaillons des détecteurs de telles entités.

Chapitre 7

Détection de structures bidimensionnelles en 2D

Nous nous intéressons ici à la détection de structures bidimensionnelles présentes dans l'image. Nous rappelons d'abord la notion de région d'intérêt, le principe de sélection de telles régions dans le *Scale Space* (chap. 4), ainsi que quelques opérateurs classiques. Nous présentons ensuite un détecteur de régions d'intérêt fondé sur les ML (chap. 3), contribution essentielle de cette thèse [30–32]. Les objets détectés correspondent à la fois à la jonction de ML et un maximum local de la réponse de l'opérateur. Un tel objet peut alors être caractérisé grâce à l'association à différentes ML ; en particulier, nous définissons une nouvelle échelle caractéristique, et donnons une estimation de sa forme.

Cette approche se distingue dans la mesure où le mécanisme de sélection est basé sur les propriétés des ML, ce qui contraste avec des méthodes basées uniquement sur un seuillage. Nous discuterons également des régions détectées, et une comparaison avec d'autres détecteurs pourra être réalisée grâce aux tests de répétabilité que nous avons réalisés (annexe A).

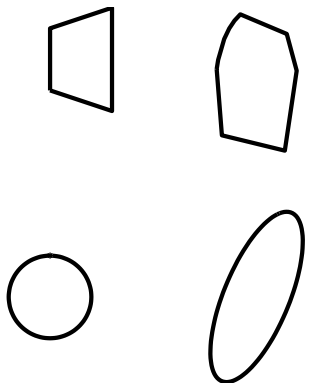


Fig. 7.1 – Exemples de régions d'intérêt

7.1 Détection de points d'intérêt en vision par ordinateur

7.1.1 Notion de région d'intérêt

Si nous lui présentons une image, un observateur humain est capable d'identifier quelles sont les régions qui lui paraissent les plus pertinentes ou les plus remarquables. Le domaine de la vision par ordinateur propose des méthodes permettant de mettre en évidence de telles régions, appelées *régions d'intérêt*. La détection de régions d'intérêt constitue souvent une première étape permettant diverses applications, dont nous donnerons des exemples plus loin. En premier lieu, il est nécessaire qu'une telle région vérifie certaines propriétés :

- correspondre à un domaine bien défini de l'image, généralement convexe (Fig. 7.1) ;
- être invariante ou robuste vis à vis de certaines transformations de l'image (chap. 6) ;
- contenir une information pertinente et distinctive ; par exemple, la région d'intérêt correspond à un objet défini (coin, bord, structure plus complexe).

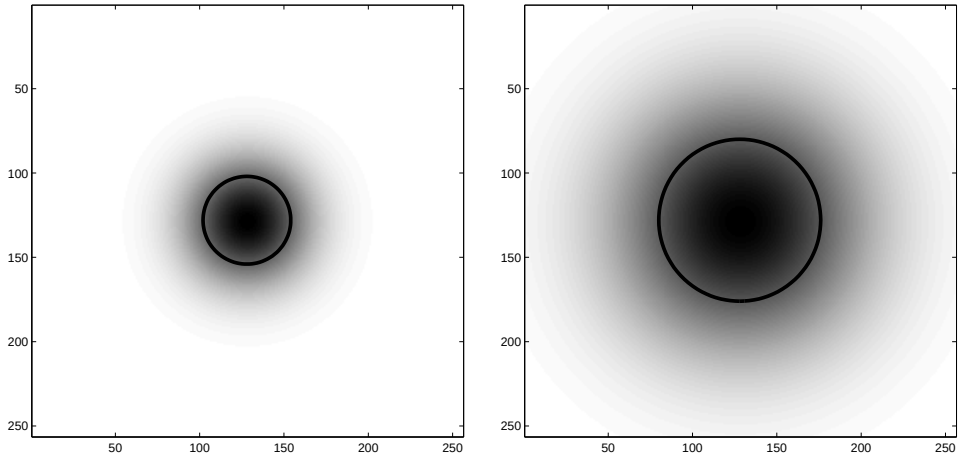
En vision par ordinateur, il est intéressant de mettre en évidence un certain nombre de régions – généralement assez grand – afin d'exploiter au maximum le contenu de l'image, quitte à ce qu'il y ait une grande redondance. Ainsi, dans une problématique de mise en correspondance de deux images, plutôt que de tenter de calculer la transformation entre les deux, il est plus judicieux d'extraire des éléments particuliers, puis d'en mettre certains en correspondance.

Pour une région donnée, la frontière doit être définie précisément : un nombre fini de point pour un polygone convexe, tous les pixels associés à la frontière dans un cadre général. Dans le cas où la région est définie à partir d'un cercle ou d'une ellipse, nous parlerons de point d'intérêt ; pour un cercle, ce point est défini comme $(x^*, y^*, s^*) \in \mathbb{R}^2 \times \mathbb{R}_+^*$ et la région associée est donnée par $(x-x^*)^2+(y-y^*)^2 \leq (s^*)^2$; pour une ellipse, ce point est défini par un couple $(X^*, \Sigma^*) \in \mathbb{R}^2 \times SDP_2$ et la région associée est donnée par $(X^*)^T \Sigma\, X^* \leq 1$.

Notion d'invariance

L'invariance comme la robustesse vis à vis de certaines transformations sont des conditions nécessaires pour traiter les applications visées. Par exemple, en ce qui concerne le problème de mise en correspondance entre deux images, la transformation permettant de passer de l'une à l'autre est *a priori* inconnue. Les transformations de l'image que nous considérons s'entendent en un sens assez général :

- déformations géométriques : la géométrie est modifiée, le contenu est préservé. Il peut s'agir de rotations, de changements d'échelle, de transformations affines, de changement de perspective ;
- modification de l'intensité : changement d'illumination ou de contraste ;
- altération de l'image : effet d'une compression avec perte (JPEG), occlusion d'une partie de l'image, perturbation par un bruit.



(A) (B)

FIG. 7.2 – (A–B) Deux surfaces Gaussiennes bidimensionnelles reliées par un changement d'échelle. La sélection de maxima dans le Scale-Space met en évidence deux maxima respectifs (x_A^*, y_A^*, s_A^*) et (x_B^*, y_B^*, s_B^*). Ceux-ci sont représentés par des cercles centrés en (x_A^*, y_A^*) et (x_B^*, y_B^*), de rayons s_A^* et s_B^* . Le rapport s_B^*/s_A^* est égal au facteur du changement d'échelle.

Afin de prendre en compte ces transformations, deux approches sont envisageables :

- soit calculer une primitive invariante (les régions issues d'une méthode basée sur un noyau isotrope pourront être invariantes à la rotation, celles issues d'une approche multiéchelles pourront être invariantes au changement d'échelle) ;
- soit estimer la transformation, et la prendre en compte par la suite (calcul d'échelle caractéristique, estimation de déformation par l'algorithme de *Shape Adaptation*).

7.1.2 Régions associées à des maxima dans le *Scale-Space*

La détection d'entités locales (objets, structures bidimensionnelles) constitue une approche intéressante dans le domaine de la vision par ordinateur. En partant du principe que certaines régions de l'image apparaissent comme particulièrement distinctives, cela conduit à formuler des détecteurs automatiques de telles régions [74, 75, 78, 83, 111, 112]. Nous nous intéressons ici à ce type de détecteurs, basés sur la sélection de maxima dans le *Scale-Space*. Nous avons vu que deux structures identiques, apparaissant à des échelles différentes, conduisaient à des maxima dans le *Scale-Space* de même amplitude, pourvu que l'opérateur soit correctement normalisé (chap. 4, [75]). Ces maxima peuvent être associés à certaines régions de l'image, dites *régions d'intérêt*, qui prennent en compte la dimension caractéristique de l'objet sous-jacent (Fig. 7.2).

La pertinence d'une région d'intérêt est difficile à définir. Afin de la cerner, deux idées rentrent en ligne de compte ; d'une part, le caractère saillant d'une région : contraste entre

la région d'intérêt et son voisinage immédiat ; d'autre part, le caractère informatif de celle-ci. Globalement, il n'est pas évident de déterminer quels sont les éléments les plus importants d'une image. Visuellement, certaines structures apparaissent comme importantes, mais il est difficile de les associer à un modèle d'objet particulier ; ceci contraste avec les structures 1D (coins, bords, frontières) qui sont plus aisément identifiables (chap. 6).

L'approche par ML présentée ici constitue un premier pas vers le développement de modèles de structures 2D et de méthodes de caractérisations de celles-ci. Dans cette optique, un détecteur de régions d'intérêt doit extraire différents objets (blobs, ...), tout en étant à même de pouvoir identifier leur type (comme par la régularité Lipschitzienne pour les structures 1D, chap. 6). Les détecteurs existants insistent sur la notion de *répétabilité* (annexe A), de sorte que c'est le caractère stable ou robuste qui est privilégié au détriment des propriétés de la région.

Notion de détecteur invariant

Nous précisons ici la notion de détecteur invariant, que ce soit au changement d'échelle, à la rotation ou aux déformations affines (nous indiquons un exemple entre parenthèses). Formellement, nous considérons des régions R_i (ellipses) extraite d'une image I, correspondant à certaines entités $E_i \subset I$ (surfaces Gaussiennes anisotropes) à l'aide d'un détecteur (tel que ceux présentés ci-après). Une question importante dans les applications est de savoir pour quelles déformations le détecteur sera à même de mettre en évidence ces entités dans l'image déformée. Etant donné une déformation affine, l'image I est déformée en une image I', et les entités $E_i \subset I$ sont déformées en des entités $E'_i \subset I'$ (surfaces Gaussiennes déformées). Parallèlement, des régions R'_i sont extraites de I'. Alors, si pour tout i, R'_i correspond à E'_i (en termes de recouvrement), le détecteur sera dit invariant à cette déformation. La différence entre cet idéal et les résultats concrets sera quantifié par la notion de répétabilité (annexe A). Si un détecteur est invariant à une déformation, les régions seront dites également invariantes. Ainsi des régions affines désignent des régions mises en évidence par un détecteur invariant aux déformations affines.

7.1.3 Application des détecteurs de régions d'intérêt

Les détecteurs de points d'intérêt extraient d'une image un certain nombre de régions, généralement assez grand (100 à 2000). L'intérêt de cette approche est qu'elle se fonde sur des structures locales de l'image : le grand nombre de caractéristiques assure une certaine robustesse, tandis que leur caractère local assure leur répétabilité pour des séquences d'images ayant approximativement le même contenu. Nous présentons ici quelques applications, dans leurs grandes lignes.

Mise en correspondance – *Matching*

Problématique : étant donné deux images I_1 et I_2, nous cherchons à établir si elles correspondent à la même scène.

Méthodologie [8, 9, 118, 107, 110]

1. Application d'un détecteur à I_1 et I_2, ce qui permet d'extraire n_1 et n_2 régions d'intérêt, notées $(P_1^i)_{1\leq i\leq n_1}$ et $(P_2^j)_{1\leq j\leq n_2}$.
2. Calcul de descripteurs : à chaque région P_1^i (resp. P_2^j) est associée une donnée numérique, appelée descripteur [100], notée D_1^i (resp. D_2^j). Un tel descripteur peut être par exemple un vecteur composé de 128 valeurs réelles, qui synthétise l'information contenue dans la région, tout en prenant en compte des transformations géométriques.
3. Une fois ces descripteurs calculés, certains peuvent être alors appariés deux à deux : par exemple si D_1^i et D_2^j sont suffisamment proches (au sens de la distance euclidienne par exemple), ce qui conduit alors à une mise en correspondance de la forme $P_1^i \longleftrightarrow P_2^j$.
4. Enfin, si suffisamment de correspondances de ce type peuvent être établies, alors il est raisonnable de penser que les deux images I_1 et I_2 correspondent à la même scène.

Evaluation des détecteurs de points d'intérêt : supposons que les images I_1 et I_2 soient reliées par une déformation connue. Il est alors possible de quantifier dans quelle mesure les régions d'intérêt se recouvrent d'une image à l'autre. Plus ces régions se recouvrent, plus le détecteur est performant.

Reconnaissance d'objets – *Object Recognition*

Problématique : nous souhaitons savoir si une image contient telle ou telle entité, et si tel est le cas, pouvoir la localiser. Par exemple, un visage contient des éléments caractéristiques tels que les yeux ou la bouche, ou un manuscrit comporte des lettres alphabétiques.

Méthodologie [43, 98, 124, 119] : un détecteur met en évidence des régions pertinentes. A partir de leur contenu, une règle permet de décider si la région correspond aux caractéristiques recherchées. Alors, si suffisamment de régions sont identifiées comme telles, il est vraisemblable que l'image corresponde effectivement à l'entité recherchée.

7.1.4 Principe de détection de points d'intérêt dans le Scale-Space

Partant d'une image f de taille $n_1 \times n_2$, nous nous donnons un ensemble d'échelles $\{s_1, ..., s_{max}\}$. Le choix des échelles utilisées dépend de la taille des structures recherchées ; dans la plupart des applications, la gamme d'échelles est assez fine (comme par exemple $s = n\delta, 0.1 \leq \delta \leq 2$), ce qui permet d'être suffisamment robuste aux changements d'échelle. La détection de points d'intérêt peut alors être effectuée par une recherche d'extrema parmi les points du Scale-Space discret, donné par

$$S_{3d} = \left\{(x, y, s) \in \{1, ..., n\}^2 \times \{s_1, ..., s_{max}\}\right\} \tag{7.1}$$

Cette recherche peut être effectuée par un algorithme dont les grandes étapes sont les suivantes [72, 83] :

1. Calcul d'un ensemble de cartes multiéchelles discrètes $T(x, y, s), (x, y, s) \in S_{3d}$;
2. Recherche des extrema : points du Scale-Space discret (x^*, y^*, s^*) tels que $|T(., ., .)|$ soit localement maximum dans un voisinage $3 \times 3 \times 3$ (chaque point de S_{3d} est comparé à ses 26 voisins) ;
3. Détermination des points d'intérêt :
   - dans le cas de l'invariance au changement d'échelle (régions d'intérêt : cercles)

$$\{(x^*, y^*, s^*, T(x^*, y^*, s^*)), |T(x^*, y^*, s^*)|\}$$

   - dans le cas de l'invariance à l'affine (régions d'intérêt : ellipses)

$$\{(x^*, y^*, s^*, \Sigma^*, T(x^*, y^*, s^*)), |T(x^*, y^*, s^*)|\}$$

   où $\Sigma^* \in SDP_2$ est issue de l'algorithme de *Shape Adaptation* [79], ou de la méthode d'estimation de forme par ML, voir section 7.4.4).

Primitives extraites d'une région d'intérêt

Les détecteurs de points d'intérêt permettent de mettre en évidence certaines régions de l'image jugées pertinentes. Notons qu'à ce stade, le contenu de ces régions n'est pas exploité. Décrire le contenu de ces régions est l'objet d'une opération supplémentaire, le calcul des descripteurs. Cette étape comporte principalement deux volets :

- d'une part, extraire des éléments synthétiques, tels que l'orientation dominante du gradient ce qui permet de compléter des informations sur la forme de l'objet sous-jacent ;
- d'autre part, agréger l'information contenue dans une région donnée, ce qui conduit typiquement à un vecteur de grande dimension appelé *descripteur*.

7.1.5 Exemples de détecteurs utilisés

Nous présentons ici trois détecteurs classiques en vision par ordinateur : d'abord, le Laplacien de Gaussienne, correspondant à la détection de structures de type blob [75] ; ensuite, la différence de Gaussienne, similaire au précédent et permettant un calcul rapide [83] ; enfin le détecteur de Harris-Laplace, qui est un détecteur de coins [53].

Dans le cadre 2D, rappelons les expressions du noyau Gaussien G et de la représentation espace-échelle linéaire L d'une image f :

$$\begin{aligned} G(x, y, s) &= \frac{1}{2\pi s^2} \exp\left(-\frac{x^2 + y^2}{2s^2}\right) \\ L(x, y, s) &= f(., .) * G(., ., s) \end{aligned}$$

Remarque : la variable t utilisée par Lindeberg vaut ici $t = s^2$ et G vérifie l'EDP suivante :

$$\frac{\partial G}{\partial s} = s\Delta G \tag{7.2}$$

Laplacien de Gaussienne (LoG)

Nous avons vu que l'opérateur basé sur le Laplacien d'une représentation échelle-espace est adapté à la détection de caractéristiques dans une image, pourvu qu'une normalisation adéquate soit adoptée. Ceci conduit poser l'opérateur *LoG* (*Laplacian of Gaussian*) :

$$LoG(x, y, s) = s^2\Delta L = s^2\left(L_{xx} + L_{yy}\right) \tag{7.3}$$

Numériquement, l'image est d'abord lissée par un filtre Gaussien, puis les dérivées sont calculées sur l'image lissée. La sélection d'échelle est basée sur les maxima locaux de la quantité $|LoG|$ dans S_{3d}.

Différence de Gaussienne (DoG)

L'opérateur DoG (*Difference of Gaussian*) est défini comme :

$$DoG(x, y, \sigma) = L(x, y, ks) - L(x, y, s) \text{ où } k \text{ est une constante} \tag{7.4}$$

Relevons que, pour k proche de 1 :

$$\begin{aligned} L(x, y, ks) - L(x, y, s) &= f * (G(x, y, ks) - G(x, y, s)) \\ &\approx f * ((k-1)s\partial_\sigma G(x, y, s)) \\ &\approx (k-1)s^2\left(f * \Delta G\right) \\ &\approx (k-1)s^2\Delta L \end{aligned}$$

ce qui correspond à l'opérateur *LoG* vu précédemment, à un facteur $(k-1)$ près ; or ce facteur est constant avec l'échelle. Ainsi, dans la mesure de l'approximation ci-dessus, les opérateurs *LoG* et *DoG* conduisent à des détecteurs similaires. L'intérêt de cet opérateur réside principalement dans la rapidité du calcul des cartes multiéchelles.

Détecteur de Harris-Laplace (HL)

Détecteur de Harris

Le détecteur de Harris utilise un opérateur très différent des précédents, basé sur l'autocorrélation des gradients. En un point (x, y), étant donné un voisinage $V = V(x, y)$, et $L = f * G(., ., \sigma)$ (σ fixé, relativement faible) nous considérons les matrices suivantes :

$$B(x, y) = \begin{pmatrix} L_x^2(x, y) & L_xL_y(x, y) \\ L_xL_y(x, y) & L_y^2(x, y) \end{pmatrix} \tag{7.5}$$

$$C(x,y) = \frac{1}{|V|} \sum_{(x,y)\in V} B(x,y) \tag{7.6}$$

Notons que $C(x,y)$ est un cas particulier de la MM2 vue au chap. 4, vu que $B(x,y) = [\nabla L][\nabla L]^T$ (la fenêtre utilisée ici revient à une moyenne dans le voisinage). Par ailleurs, B est de rang au plus égal à 1, mais pour la plupart des images, $\det(C(x,y))$ est souvent non nul. En notant $\lambda_1 > \lambda_2$ les valeurs propres de $C(x,y)$, relevons que :

- dans des zones uniformes, λ_1 et λ_2 sont faibles ;
- sur une arête, λ_1 est forte par rapport à λ_2 ;
- sur un coin, λ_1 et λ_2 sont fortes.

La sélection de points pour lesquels λ_1 et λ_2 sont fortes (seuillage) conduit alors à détecter des coins, qui sont des données stables. Relevons que l'échelle d'analyse étant fixée, nous ne disposons à ce stade que de la localisation des points d'intérêt (x^*, y^*). Afin d'obtenir un détecteur de points d'intérêt invariant à l'échelle, il importe de calculer une échelle caractéristique en ces points [111].

Calcul de l'échelle caractéristique par un Laplacien

Nous avons vu auparavant la notion d'échelle caractéristique associée à un maximum dans le *Scale-Space* (chap. 4). La méthode décrite ici permet de calculer une certaine échelle caractéristique en un point quelconque de l'image (x,y) (mentionnons au passage l'existence d'approches globales, calculant une échelle caractéristique moyenne associée à une image [84]). En particulier, à partir d'un point (x^*, y^*) détecté précédemment, une échelle caractéristique s^* peut être calculée, ce qui conduit à un point d'intérêt (x^*, y^*, s^*). Nous reprenons l'opérateur Laplacien normalisé et introduisons la fonction Γ, définie en un point $(x,y) \in \mathbb{R}^2$ par :

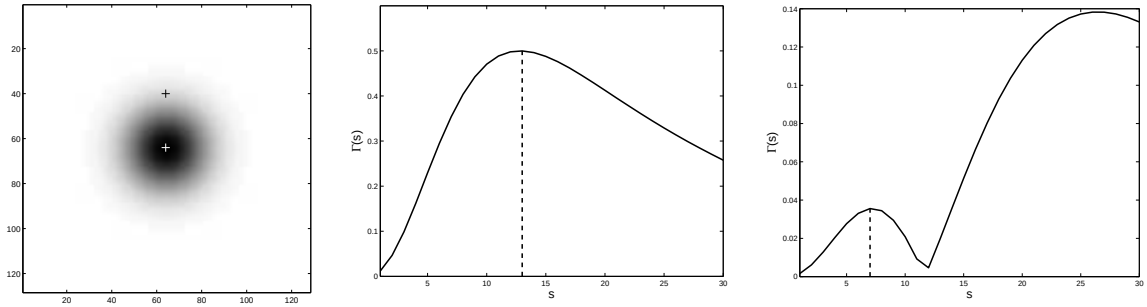
$$\begin{array}{rccl} \Gamma(x,y): & \mathbb{R}_+^* & \longrightarrow & \mathbb{R}_+ \\ & s & \longmapsto & s^2|\Delta L(x,y,s)| \end{array} \tag{7.7}$$

L'échelle caractéristique associée au point (x,y) est alors donnée par le premier maximum local de $\Gamma(x,y)$, sa valeur Γ_{max} étant atteinte en $s^* = \text{Argmax}\ \Gamma(x,y)(s)$. Remarquons que lorsque (x^*, y^*, s^*) est un extremum dans le Scale-Space, $\Gamma(x^*, y^*)$ admet un maximum local en s^*. Cependant, si (x', y') est un extremum de $\Delta L(.,.,1)$ (échelle fine), et si $\Gamma(x', y', .)$ admet un maximum local en s', (x', y', s') n'est pas nécessairement un extremum dans le Scale-Space (chap. 4). Relevons aussi que cette manière de calculer l'échelle caractéristique a été vue en 1D (chap. 2, section 2.4, Fig. 2.6). Notons également que le calcul effectif des échelles caractéristiques peut être effectué à l'aide d'une transformée en ondelettes (avec l'ondelette Δg par exemple), et qu'il peut être rendu très précis par un raffinement des échelles.
En considérant deux points différents d'une surface Gaussienne bidimensionnelle (Fig. 7.3 (A)), nous pouvons tracer l'évolution de Γ en fonction de l'échelle s. Pour le point central (x^*, y^*), qui est associé à un extremum dans le Scale-Space (x^*, y^*, s^*), nous observons un unique maximum, atteint en s^* (Fig. 7.3 (B)). Pour un point excentré, en notant s' l'échelle associée

au premier maximum (Fig. 7.3 (C)), celle-ci est appelé caractéristique dans la mesure où, si l'image subit un changement d'échelle d'un facteur λ, l'échelle caractéristique associée à ce même point sera égale à $\lambda \cdot s'$ (invariance au changement d'échelle).

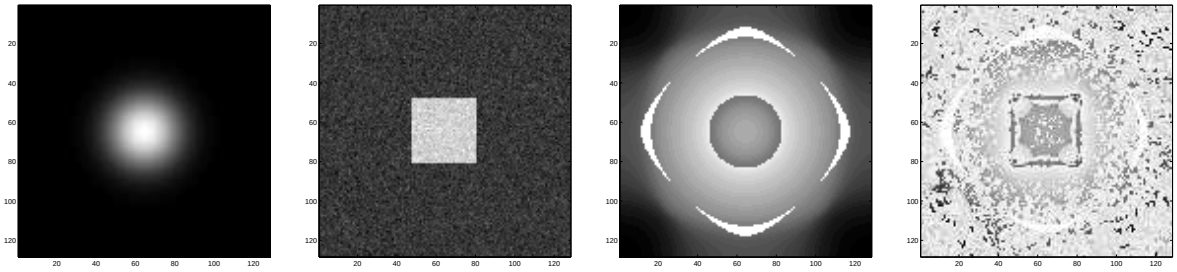
Pour des structures de type blob, l'échelle caractéristique est relative à la taille de la structure sous-jacente ; ici, cette échelle correspond à la zone d'influence d'un certain point (souvent un coin), de sorte que la région d'intérêt concerne à la fois le point lui-même et son voisinage. Ainsi, le sens donné à l'échelle caractéristique est différent :

- pour une structure de type blob, cette échelle est proportionnelle à la taille de la structure (laquelle est souvent délimitée par des frontières) ;
- pour un coin, cette échelle est d'autant plus grande qu'il n'y a pas d'autres structures importantes à proximité.



(A) (B) (C)

FIG. 7.3 – Calcul de l'échelle caractéristique en un point quelconque : (A) Surface Gaussienne bidimensionnelle, à partir de laquelle $\Gamma(x, y)$ est calculée ; (B) Evolution de Γ en fonction de s au centre (croix blanche de (A)) ; (C) Evolution de Γ en fonction de s en un point excentré (croix noire de (A)).



(A) (B) (A') (B')

FIG. 7.4 – Calcul de l'échelle caractéristique en tout point – le niveau de gris indique l'échelle caractéristique (A) Surface Gaussienne isotrope ; (B) Carré bruité (SNR : 20 dB) ; (A', B') Echelles caractéristiques calculées en tout point.

Nous considérons maintenant deux images – une représentant une surface Gaussienne, l'autre un carré bruité – et calculons l'échelle caractéristique en tout point (représentée en niveau de gris, Fig. 7.4). Notons qu'au voisinage de points associés à des extrema dans le Scale-Space (centre de la structure, coins), Γ_{max} est proche de la valeur de s^* associée aux extrema, ce qui montre bien le caractère stable de ces extrema (Fig. 7.4 (A') et (B')). Par ailleurs, en d'autres points, nous observons des valeurs de Argmax Γ élevées : ces valeurs élevées correspondent à la symétrisation de l'image (utilisée dans le calcul de la TOC), ce qui revient dans ce cas particulier à une périodisation de l'image : d'une part la Gaussienne implique une structure de type blob, et d'autre part, du fait de la périodisation, une autre structure apparaît entre la Gaussienne originale et celle qui provient de la périodisation (le phénomène est similaire pour le carré bruité).

Exemples

L'application de ces détecteurs sur l'image classique *Barbara* conduit à différentes régions d'intérêt représentées sur la figure 7.5. Notons que les opérateurs présentés ici sont basés sur un noyau Gaussien isotrope, ainsi les détecteurs sont invariants au changement d'échelle et à la rotation et *a priori* non-invariants à l'affine. En appliquant l'algorithme de *Shape Adaptation*, l'anisotropie est prise en compte, et il est alors possible d'obtenir des détecteurs invariants à l'affine (Fig. 7.5). Note : ces détecteurs sont disponibles auprès de l'équipe LEAR de l'Inria Rhône-Alpes (http://lear.inrialpes.fr/software). Relevons enfin que le nombre de régions est assez important (Tab. 7.1). Dans les applications, le fait que de nombreuses régions soient détectées permet d'atteindre une certaine robustesse, comme dans le cas de l'occlusion d'une partie de l'image.

Nous considérons d'abord les points d'intérêt représentés sur les figures 7.5 (A–C). En ce qui concerne *LoG* et *DoG*, les maxima sont détectées dans le *Scale-Space* 3D, ce qui fait que de nombreuses régions d'intérêt qui s'intersectent. Si certains points d'intérêt peuvent être identifiés à certaines entités (parties du visages en (A–B)) il reste que pour une majorité de régions, le lien entre la région détectée et son contenu est délicat à établir. En ce qui concerne HL, les régions sont reliées à des coins (points où la courbure est élevée), et certaines régions semblent s'accumuler autour de certains points, de sorte qu'une singularité donne lieu à plusieurs régions d'intérêt. En outre, en considérant les régions affines associées (Fig. 7.5 (A'–C')), nous voyons que certaines zones se prêtent bien à l'adaptation affine (comme le pied de table en (A') ou la jambe droite de *Barbara* en (B')), tandis que pour d'autres, l'algorithme de *Shape Adaptation* conduit à des régions dont le contenu n'apparaît pas clairement comme un élément remarquable.

Globalement, ces méthodes ont montré leur efficacité [101]. Cependant, le lien entre ces régions et des singularités présentes dans l'image n'est pas évident. Nous détaillons dans la suite une méthode basée sur les Maxima Lines permettant de détecter des régions d'intérêt et de les caractériser précisément.

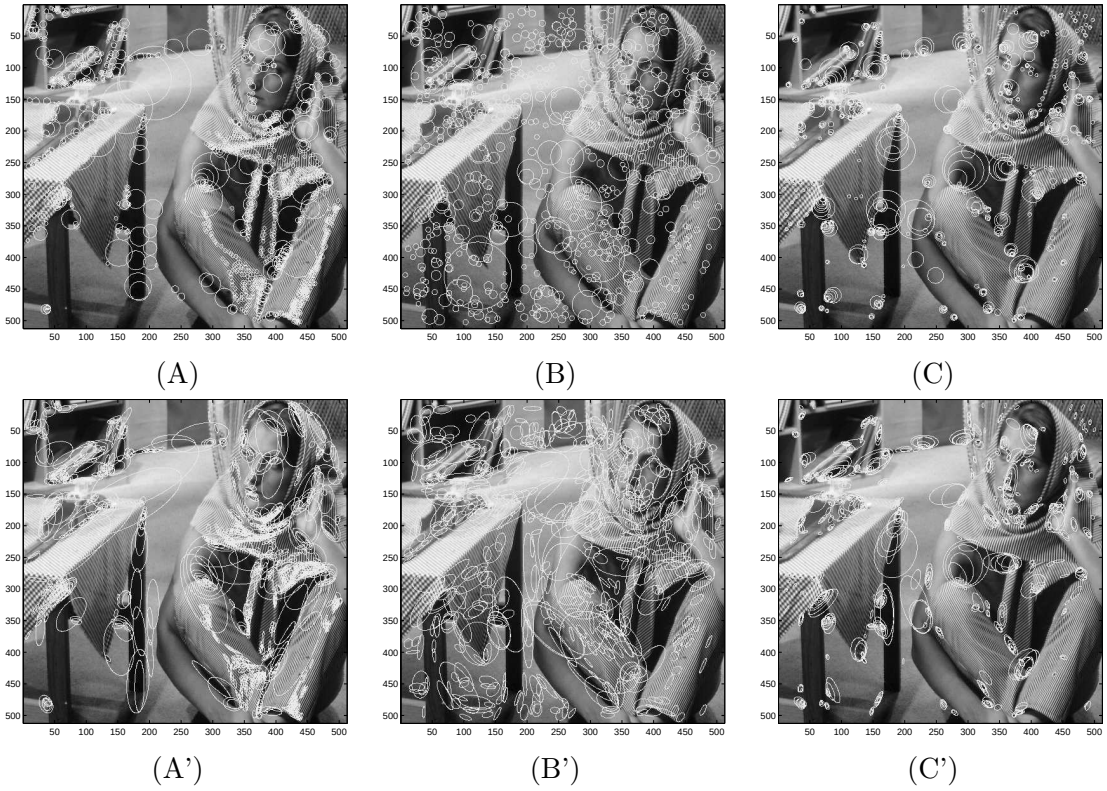


FIG. 7.5 – Application de détecteurs classiques en vision par ordinateur (A, A') Laplacien de Gaussienne (LoG) ; (B, B') Différence de Gaussienne (DoG) ; (C, C') Harris-Laplace (HL) ; (A, B, C) Points d'intérêts (x^*, y^*, s^*) ; (A', B', C') Régions affines $(x^*, y^*, s^*, \Sigma^*)$.

TAB. 7.1 – Nombre de points obtenus pour la détection de l'image *Barbara* 512x512, suivant l'opérateur utilisé, et suivant que l'on ait estimé la matrice des moments d'ordre deux (auquel cas les régions d'intérêt sont des ellipses).

Opérateur utilisé	Points d'intérêt	Régions affines
LoG	1353	907
DoG	724	585
HL	721	657

7.2 Approche par ML

Nous présentons maintenant une approche originale pour mettre en évidence des structures bidimensionnelles dans une image, fondée sur les ML de l'ondelette ΔG [30–32]. Celle-ci est basée sur un nouveau mécanisme de sélection d'échelle, que nous comparons à la sélection classique. La sélection de ML pertinentes conduit alors à un détecteur de régions d'intérêt de type *blob*.

7.2.1 Mécanisme de sélection le long de ML

Nous nous plaçons dans le cadre de l'ondelette ΔG, où il existe des ML. Une telle ligne est donnée par :

$$\begin{aligned} &L = (x(s), y(s), s, Wf(x(s), y(s), s))_{s\in]0,s_{int}[}, s_{int} \in \overline{\mathbb{R}_+^*} \\ &\text{où pour tout } s,\ (x(s), y(s)) \text{ est un MM de } Wf(.,.,s). \end{aligned} \tag{7.8}$$

Définition 34. *Maxima le long de lignes de maxima*

$$\begin{aligned} &MaxML = \{(x^*, y^*, s^*) \textit{ appartenant à une ML } L, \\ &\textit{telle que la réponse associée à } L \textit{ admet un maximum local en } s = s^*.\} \end{aligned} \tag{7.9}$$

Cette définition constitue un nouveau mécanisme de sélection. Rappelons que le mécanisme classique est fondé sur $MaxI$ (définition 30, section 4.2.2) :

$$MaxI = \{(x^*, y^*, s^*)/|\Gamma L| : \mathbb{R}^3 \longrightarrow \mathbb{R} \text{ localement maximum.}\}$$

Tandis que l'ensemble $MaxI$ est constitué de maxima locaux dans toutes les directions du Scale-Space, l'ensemble $MaxML$ est constitué de maxima en échelle, suivant des trajectoires particulières de maxima en espace. L'intérêt principal de l'approche fondée sur $MaxML$ est qu'elle permet de donner un supplément d'information sur la structure sous-jacente, ce qui n'est pas le cas dans l'approche classique.

7.2.2 Comparaison avec la sélection classique

Nous comparons ici les deux méthodes de sélection d'extrema dans le Scale-Space, en nous plaçant dans cadre 1D discret, la démonstration en 2D étant similaire.

Sens $MaxI \Rightarrow MaxML$

Soit $(u^*, s^*) \in MaxI$. Nous considérons le voisinage de (u^*, s^*) suivant :

$$V(u^*, s^*) = \{(u, s) \in \mathbb{Z}^2/u^* - 1 \leq u \leq u^* + 1, s^* - 1 \leq s \leq s^* + 1\}$$

Alors :

$$\forall (u, s) \in V(u^*, s^*), |Wf(u^*, s^*)| \geq |Wf(u, s)|$$

En particulier $(u^*, s^*) \in MaxI$ implique que (u^*, s^*) correspond à un MM à l'échelle s^*, et ainsi (u^*, s^*) appartient à une ML (rappelons que tous les MM appartiennent à une ML). De plus, sous l'hypothèse que la discrétisation soit suffisamment fine, la continuité des lignes de maxima induit qu'il existe (au moins) un module maximum à l'échelle $s^* - 1$ localisé dans $\{u^* - 1, u^*, u^* + 1\}$, noté (u^-, s^-). Distinguons alors deux cas.

- Cas où la ligne s'interrompt : dans ce cas, on a $|Wf(u^*, s^*)| > |Wf(u^-, s^-)|$, ce qui donne un maximum local (en un sens assez large, puisqu'il n'y a pas de module maximum à l'échelle $s + 1$)
- Cas où la ligne de maxima perdure à l'échelle $s + 1$: il existe (au moins) un module maximum à l'échelle $s+1$ localisé dans $\{u^* - 1, u^*, u^* + 1\}$, noté (u^+, s^+). Etant donné que $(u^+, s^+), (u^-, s^-) \in V(u^*, s^*)$, nous avons

$$\begin{cases} |Wf(u^*, s^*)| & > & |Wf(u^+, s^+)| \\ |Wf(u^*, s^*)| & > & |Wf(u^-, s^-)| \end{cases}$$

  En conclusion, $(u^*, s^*) \in MaxI \Rightarrow (u^*, s^*) \in MaxML$.

Sens $MaxML \Rightarrow MaxI$

Réciproquement, soit $(u^*, s^*) \in MaxML$. Par hypothèse, il est situé sur une ligne de maxima ininterrompue de l'échelle la plus fine à l'échelle $s^* + 1$ (au moins). Différentes possibilités se présentent :

- la ligne de maxima est stable spatialement autour de l'échelle s^* : alors elle passe par $(u^*, s^* - 1), (u^*, s^*), (u^*, s^* + 1)$. Comme (u^*, s^*) constitue un maximum (local) le long de la ligne de maxima à laquelle il appartient, nous avons :

$$\begin{array}{rcl} |Wf(u^*, s^*)| & \geq & |Wf(u^*, s^* - 1)| \\ |Wf(u^*, s^*)| & \geq & |Wf(u^*, s^* + 1)| \end{array}$$

  Par ailleurs, les points $(u^*, s^*-1), (u^*, s^*), (u^*, s^*+1)$ correspondant à des MM (maxima locaux en espace), ils vient

$$\begin{array}{rcl} |Wf(u^*, s^* - 1)| & \geq & |Wf(u^* \pm 1, s^* - 1)| \\ |Wf(u^*, s^*)| & \geq & |Wf(u^* \pm 1, s^*)| \\ |Wf(u^*, s^* + 1)| & \geq & |Wf(u^* \pm 1, s^* + 1)| \end{array}$$

$$\text{et ainsi } \forall (u, s) \in V(u^*, s^*), |Wf(u^*, s^*)| \geq |Wf(u, s)|$$

- la ligne de maxima dérive : dans ce cas, il n'est pas certain qu'il s'agisse d'un maximum isolé ; considérons en effet la configuration suivante, où $(u^*, s^*) \in MaxML$ et $(u^*, s^*) \notin MaxI$:

$$\begin{cases} \text{Une ligne de maxima passant par } (u^* - 1, s^* - 1), (u^*, s^*), (u^* + 1, s^* + 1) \\ (u^*, s^*) \text{ maximum local le long de cette ligne de maxima,} \\ \text{Il existe un module maximum en } (u^* + 1, s^* + 1), \\ |Wf(u^* + 1, s^* + 1)| > |Wf(u^*, s^*|. \end{cases}$$

Synthèse

Nous avons vu que $MaxI \Rightarrow MaxML$ et que $MaxML \Rightarrow MaxI$ seulement pour des lignes stables spatialement. Relevons par ailleurs que le nombre de MM décroît très rapidement à mesure que l'échelle augmente. Ainsi, pour s suffisamment grand, ceux-ci seront relativement isolés spatialement, de sorte que la plupart des ML sont elles-mêmes isolées les unes des autres. Ainsi, le cas de deux lignes disjointes et proches est marginal : soit les ML fusionnent (ce qui notamment le cas dans le cas dans la détection de *blobs*), soit l'une disparaît au profit de l'autre. L'idée de rechercher des maxima le long de certains chemins (*extremum paths*) avait déjà été émise [73], sans être complètement exploitée. Nous verrons ultérieurement que certaines entités (structures locales) peuvent être associées à plusieurs ML ayant une double propriété : d'une part elles fusionnent à une certaine échelle s^*, et d'autre part la réponse le long de celles-ci admet un maximum local en s^*. Les ML associées à une entité donnée permettent alors d'en donner une description sommaire.

7.3 Détection de point d'intérêt par ML

Nous rappelons d'abord quelques notions associées aux ML, puis décrivons notre algorithme permettant d'extraire des points d'intérêt d'une image.

7.3.1 Quelques notions associées aux ML

Pic le long d'une ML – Echelle caractéristique d'une ML

Définition 35. *(Pic de ML) Etant donné une ML* $(x(s), y(s), s, Wf(x(s), y(s), s))_{s>0}$*, nous dirons qu'elle admet un pic en* $(x^*, y^*, s^*) \in \mathbb{R}^2 \times \mathbb{R}_+^*$ *si* (x^*, y^*, s^*) *appartient à cette ML et si la réponse* $s \mapsto |Wf(x(s), y(s), s)|$ *le long de cette ML admet un maximum local en* $s = s^*$*.*

Etant donné une ML, une échelle s^* sera dite caractéristique de celle-ci si :
- s^* est supérieure au niveau de bruit,
- la réponse le long de la ML admet un maximum local en s^*.

Caractère stable d'une ML

Nous considérons une ML telle que la réponse le long de celle-ci atteigne un maximum local à l'échelle s^*, localisé en (x^*, y^*). Cette ML sera dite *stable* si elle ne dérive pas spatialement pour des échelles localement supérieures (i.e., de $s = s^*$ à $s = s^* + \delta$, $\delta > 0$). En d'autres termes, les MM associés à cette ML sont localisés en (x^*, y^*) pour ces échelles-là. En pratique, étant donné une ML correspondant à la séquence de points (x_s, y_s, s), nous considérons qu'elle est stable si :
- elle admet un pic pour une certaine échelle s^*, en (x^*, y^*),
- $\forall s \in [s^*, s^* + 2], \max_s |(x_s, y_s) - (x^*, y^*)|_2 \leq 1$.

Dans le cadre de l'étude d'une structure de type blob en 1D, nous observons que deux ML donnent simultanément une réponse maximale pour une certaine échelle s^*, se rejoignent à cette même échelle, et sont stables spatialement pour des échelles localement supérieures. Cette propriété de stabilité est intéressante pour des applications en vision par ordinateur dans la mesure où nous cherchons justement à extraire d'une image des caractéristiques robustes.

Jonction de plusieurs ML

La construction proposée permet de mettre en évidence un certain nombre de ML, n'ayant au départ aucun lien entre elles. Il est alors intéressant de pouvoir localiser là où certaines ML fusionnent (jonctions de deux ou plusieurs ML) de manière à regrouper celles-ci ; par la suite ce groupe de ML pourra être associé à un certain objet présent dans l'image. Plusieurs cas peuvent se présenter :

- une ML ne perdure pas ; par exemple, une ML de signe opposé devient prédominante ;
- des ML d'importance équivalente se rencontrent, et elles admettent simultanément un pic ;
- des ML fusionnent sans que cela corresponde nécessairement à un pic (ML de même signe, d'importance nettement différentes).

Dans le cadre de la détection de points d'intérêt, nous sélectionnons les ML qui fusionnent (critère de jonction) et le long desquelles la réponse admet un maximum local (pic de ML). Ceci constitue un mécanisme de sélection d'échelles et de points remarquables, comme nous l'avons vu au chapitre 4 (où nous avons comparé ce mécanisme et la sélection d'échelle classique). En pratique, de manière à mettre en évidence ces jonctions, nous calculons pour chaque ML L_i un point caractéristique (x_i^*, y_i^*, s_i^*) (s_i^* : échelle caractéristique) et nous apparions ensuite les ML pour lesquelles (x_i^*, y_i^*, s_i^*) sont identiques (en utilisant une certaine marge de tolérance).

7.3.2 Algorithme de détection par ML

L'approche de détection d'objets par ML que nous proposons peut être formulée par l'algorithme suivant :

1. Calculer une transformée en ondelette continue 2D (TOC), utilisant l'ondelette Δg, sur des échelles entières $s \in \{1, 2, 3, ..., s_{max}\}$;
2. Calculer les cartes de MM – maxima locaux du module de la TOC à une échelle fixée – et construire les ML associées (chaînes de MM à travers les échelles) ;
3. Sélectionner les ML les plus pertinentes ;
4. Associer des ML d'après un critère de jonction, ce qui conduit à mettre en évidence les points d'intérêts du détecteur ;
5. Pour chaque point d'intérêt, à partir des ML associées, extraire des informations sur l'objet sous-jacent, telles que sa localisation, une échelle caractéristique, et une estimation de sa forme géométrique.

L'opérateur utilisé revient à utiliser un Laplacien, permettant de mettre en évidence des structures dans l'image dont le nom générique est *blob*. Les détecteurs classiques considèrent des extrema en espace et en échelle tandis que le détecteur présenté considère des maxima le long de ML. Nous avons vu auparavant que cela conduit à sélectionner quasiment les mêmes points dans le Scale-Space 3D. Une différence notable est la manière dont est effectuée la sélection : dans un cas, un seuillage sur la réponse est effectué, tandis que dans l'autre, nous imposons des conditions particulières sur les ML afin de sélectionner celles qui sont les plus pertinentes ; cette sélection sur les propriétés des ML apparaît plus intéressante qu'un seuillage direct. En outre, cette approche apporte un supplément d'information, qui renseigne sur la nature de la structure du blob. En effet, un objet détecté peut être associé à un certain nombre de ML, qui donnent alors une information géométrique sur ce même objet. Le calcul de la TOC et la construction des ML ayant été détaillés auparavant, nous nous focalisons maintenant sur le mécanisme de sélection des ML, la mise en évidence de points d'intérêts, le nouveau calcul d'échelle caractéristique par ML et l'estimation de la forme de l'objet, qui constitue une alternative à l'algorithme de *Shape Adaptation* [79].

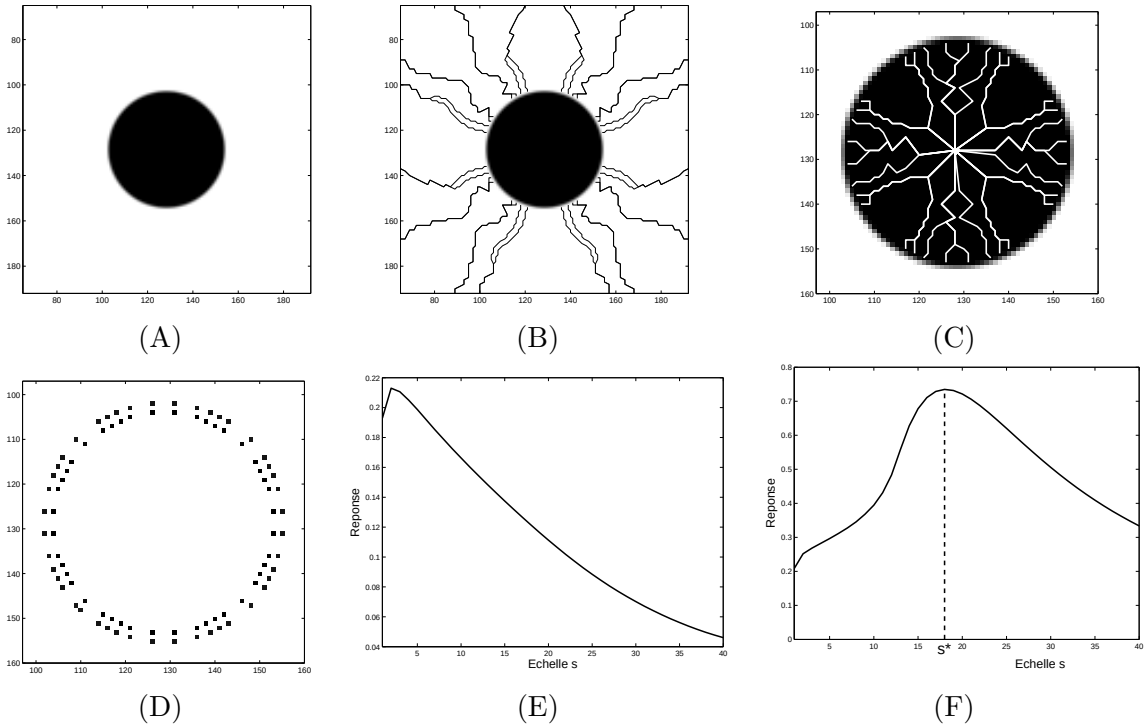


FIG. 7.6 – (A) Image représentant un disque ; (B) Trajectoire des ML extérieures ; (C) Trajectoire des ML intérieures ; (D) MM à l'échelle la plus fine ; (E) Réponse le long d'une ML dérivant vers l'extérieur ; (F) Réponse le long d'une ML dérivant vers l'intérieur.

7.3.3 Sélection de Maxima Lines pertinentes

Dans le cadre du problème de détection de structures (ou objets de type *blob*), nous cherchons à mettre en évidence les ML pertinentes parmi toutes celles qui ont été construites. Les échelles considérées sont $s = 1, 2, 3, ..., s_{max}$, où $s_{max} \in \mathbb{N}$ a été choisi en fonction de la taille des structures que l'on souhaite détecter (typiquement $s_{max} = 40$ pour une image naturelle).

Nous considérons d'abord l'exemple d'une image représentant un disque (Fig. 7.6 (A)) dont l'intensité est donnée par :

$$f(x, y) = \begin{cases} 1 \text{ si } (x, y) \text{ appartient au disque,} \\ 0 \text{ sinon.} \end{cases}$$

Notons que les MM à l'échelle la plus fine sont situés de part et d'autre de la frontière du disque (Fig. 7.6 (D)). Nous pouvons alors distinguer différentes ML, d'après leur dérive spatiale :

- celles qui dérivent vers l'extérieur (Fig. 7.6 (E)) ;
- celles qui dérivent vers l'intérieur et fusionnent au centre (Fig. 7.6 (F)) ;

Suivant la situation, la réponse le long d'une ML est différente (Fig. 7.6 (E–F)). Dans cet exemple, l'évolution de la réponse est identique pour toutes les ML intérieures et toutes les ML extérieures. Ainsi l'évolution de la réponse permet de caractériser les ML. Notons également que les coefficients d'ondelettes le long de celles-ci sont opposés (par exemple, les ML intérieures correspondent à des MM positifs tandis que les ML extérieures correspondent à des MM négatifs). Enfin, aux petites échelles, deux MM assez proches de signes opposés ont tendance à se renforcer dans un premier temps ; puis les ML intérieures se renforcent mutuellement (somme d'atomes de même signe) alors que les ML extérieures s'atténuent (les atomes sont de plus en plus éloignés). Ceci explique que dans les deux cas (ML intérieures et extérieures), l'évolution de la réponse aux petites échelles est similaire.

Dans une optique de sélection d'échelle, nous sommes amenés à sélectionner une échelle s^* pour laquelle la réponse admet un maximum local en s^*, et supérieure à s_{min} le choix de $s_{min} = 4$ par exemple permet d'éliminer les ML extérieures (Fig. 7.6 (E–E)). Cette échelle est alors caractéristique de la ML. Par ailleurs, pour un objet *a priori* inconnu, la trajectoire d'une ML ne permet pas de mettre en évidence celui-ci (la distinction intérieur-extérieur n'a pas de sens dans ce cas). Néanmoins, une fois que des ML sont associées à un objet (critère de jonction), elles permettent de le cerner approximativement : les extrémités des ML (aux échelles fines) donnent des points proches de ses bords, tandis que le lieu où la réponse est maximum est proche de son centre.

TAB. 7.2 – Exemples numériques de l'approche par ML : sur une image géométrique (disque au centre de l'image) et sur une scène naturelle (*Barbara*), dans un cadre bruité ou non – dans le cas bruité, un bruit blanc Gaussien a été ajouté à l'image originale ($SNR = 20dB$).

Image 512x512	Disque	Disque bruité	Barbara	Barbara bruitée
ML construites	166	9769	4907	6592
ML associées à du bruit	0	9694	3150	5253
ML pertinentes	166	75	1736	1324
Objets détectés (jonction de ML)	1	1	136	129

Prise en compte du bruit

Pour de nombreuses applications, en particulier la détection de points d'intérêt, les méthodes développées doivent présenter une certaine robustesse au bruit. En présence de bruit, le squelette des ML devient plus complexe, en particulier les ML correspondant à du bruit sont bien plus nombreuses que celles associées aux objets présents dans l'image. Néanmoins, il est possible de dissocier ces différentes ML, et de sélectionner uniquement les ML associés à des objets structurés.

Dans la mesure où ceux-ci sont associés à des échelles relativement grandes, il est naturel de ne considérer que les ML traversant un nombre suffisant d'échelles. Par ailleurs, il a été montré que pour un bruit blanc Gaussien, en moyenne, le nombre de MM diminue d'un facteur 2 lorsque l'échelle augmente d'un même facteur 2 [88]. Ainsi nous éliminons toutes les ML s'interrompant avant l'échelle $s = 5$. Rappelons également que le long d'une ML associée à des fluctuations liées au bruit, la réponse décroît à mesure que l'échelle croît aux fines échelles [88]. Ceci conduit alors à éliminer ce type de ML, identifiées comme du bruit. Relevons enfin que si s_{max} est faible par rapport à la taille de l'objet, les ML associées à celui-ci s'interrompront prématurément et la réponse associée à une ML intérieure sera strictement croissante (échelle caractéristique de l'objet supérieure à s_{max}). Ceci conduit à éliminer les ML dont la réponse est une fonction croissante (sur l'intervalle $[1, s_{max}]$).

Finalement, cette sélection des ML permet de se ramener à un nombre restreint de ML, dites ML pertinentes, dont le comportement s'apparente au cas simple étudié précédemment. Cette étape de sélection conduit à ne retenir typiquement que 1% des ML pour une image géométrique et 20% des ML pour une image naturelle. Enfin, en comparant deux images, l'une étant bruitée et l'autre non, nous voyons que la présence de bruit conduit à un plus grand nombre de ML, et qu'un certain nombre de ML pertinentes disparaissent lorsqu'un bruit est ajouté : le bruit détruit un nombre non négligeable de ML (Tab. 7.2).

7.3.4 Détermination des points d'intérêts – Jonction de ML

Pour déterminer un ensemble de points d'intérêt au sein d'une image donnée, après une étape de construction des ML, nous effectuons l'étape de sélection précédemment décrite, ce qui conduit à un ensemble restreint de ML appelée ML pertinentes, et chacune vérifie les propriétés suivantes :

1. elle perdure au moins jusqu'à l'échelle $s = 5$,
2. la réponse associée ne décroît pas aux échelles fines, et elle n'est pas une fonction croissante sur $[1, s_{max}]$.

Un élément important à relever est la jonction de ML dans le Scale-Space. Rappelons qu'une jonction de ML est un point $(x^*, y^*, s^*) \in \mathbb{R}^2 \times \mathbb{R}_+^*$ tel qu'il existe au moins deux ML passant par ce point. Ces jonctions permettent de mettre en évidence des objets présents dans l'image. Etant donné une jonction (x^*, y^*, s^*), nous notons $(L_i)_{i \in I}$ les ML passant par celle-ci. Ces ML sont associées à un certain objet présent dans l'image, et elles permettront de le décrire plus précisément. Enfin, nous observons que le long de chacune de ces ML $(L_i)_{i \in I}$, la réponse atteint un maximum local en $s = s^*$. De manière plus détaillée, nous voyons que :

- avant s^*, l'évolution est propre à la ML ;
- juste avant s^*, la dérive spatiale est importante ;
- en $s = s^*$, il y a fusion de ML au point de jonction (x^*, y^*, s^*), et la réponse admet un maximum local ;
- juste après s^*, l'unique ML résultant de la fusion est stable spatialement et sa réponse décroît légèrement avec l'échelle ;
- après s^*, l'évolution dépend du contexte :
  - si la ML résultante est influencée par d'autres ML, l'interaction est complexe (comme nous l'avons vu en 1D avec les atomes d'ondelette),
  - si la ML résultante est isolée, elle reste stable spatialement et la réponse décroît avec l'échelle.

A partir d'un ensemble pertinent de ML (cf. sélection de ML), le critère de jonction permet de déterminer un certain nombre de points d'intérêt, lesquels sont d'autant plus robustes que le nombre de ML impliquées dans la jonction est important. Ce critère est valable à la fois pour des images géométrique et pour des images naturelles, que l'image soit bruitée ou non (Tab. 7.2). Ainsi le critère de jonction de ML permet de mettre en évidence un certain nombre de structures bidimensionnelles présentes dans l'image.

7.3.5 Calcul de l'échelle caractéristique

L'approche par ML permet de définir une nouvelle échelle caractéristique, à la fois pour une ML et pour une jonction de ML. Cette grandeur, analogue à l'échelle caractéristique classique, est fondée sur la dérive spatiale des ML.

Définition 36. *(Echelle caractéristique associée à une ML) Etant donné une ML admettant un pic en (x^*, y^*, s^*) et dont l'extrémité est localisée en (x_0, y_0), l'échelle caractéristique associée à cette ML est définie par*

$$S^* = \sqrt{(x^* - x_0)^2 + (y^* - y_0)^2} \tag{7.10}$$

Relevons que dans le cadre de la détection d'un objet délimités par certains bords (comme un disque), (x_0, y_0) correspond à une singularité près du bord, tandis que (x^*, y^*) correspond approximativement au centre de l'objet. Alors cette grandeur S^* représente la distance euclidienne entre le centre et le bord de l'objet ; ainsi elle peut être considérée comme une échelle caractéristique puisque S^* est intrinsèque à la géométrie de l'objet.

Définition 37. *(Echelle caractéristique associée à une jonction de ML) Etant donné une jonction de plusieurs ML $(L_i)_{i\in I}$ en (x^*, y^*, s^*) (admettant toutes un pic en (x^*, y^*, s^*)), l'échelle caractéristique de la jonction (et de l'objet sous-jacent) est alors définie comme :*

$$\mathcal{S}^* = \textit{médiane}\{S_i^*\} \tag{7.11}$$

où $(S_i^*)_{i\in I}$ sont les échelles caractéristiques des ML $(L_i)_{i\in I}$.

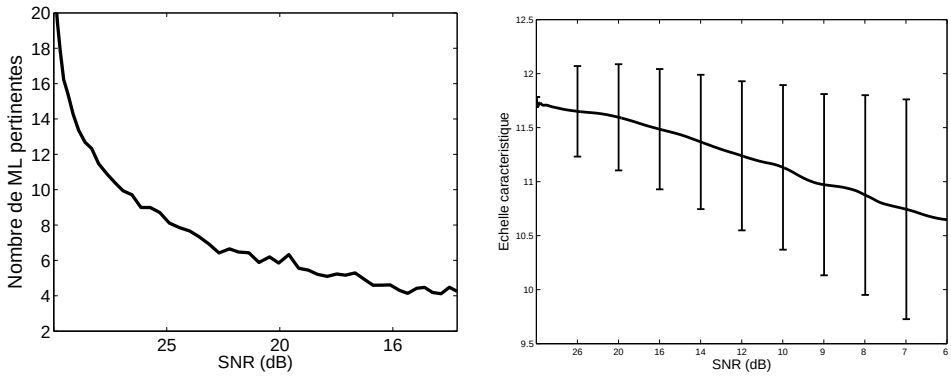
Remarque : ce choix de la médiane (plutôt que de la moyenne) est motivé par deux choses : d'une part, certaines ML ont une grande dérive spatiale, et dans certains cas cela conduit à des valeurs trop élevées de S_i^* (par rapport à la taille de l'objet) ; d'autre part, le nombre de ML impliquées dans la jonction peut varier fortement d'un cas à un autre – un grand nombre sur une image géométrique contenant un seul objet, quelques unes seulement pour un objet dans une image naturelle.

Maintenant, nous montrons la pertinence de cette approche par ML, et la comparons à l'approche classique, qui consiste à proposer (x^*, y^*, s^*) comme un point d'intérêt, (x^*, y^*) correspondant à sa localisation, et s^* à son échelle caractéristique.

7.4 Validation de l'approche par ML

7.4.1 Sensibilité au bruit de l'échelle caractéristique

Afin de tester la robustesse de cette grandeur vis à vis du bruit, nous considérons une image 128×128 représentant un disque, auquel nous ajoutons un bruit blanc Gaussien. Nous étudions alors le comportement de cette nouvelle échelle caractéristique en fonction du niveau de bruit (donné par le SNR). Pour chaque niveau de bruit, 100 simulations sont effectuées, et nous calculons alors la moyenne et la variance des valeurs obtenues de $\mathcal{S}^*$ (Fig. 7.7). A mesure que le niveau de bruit augmente, le nombre de ML permettant de calculer $\mathcal{S}^*$ diminue, et ainsi $\mathcal{S}^*$ décroît en moyenne. Sur l'exemple considéré, $\mathcal{S}^* = 11.78$ en l'absence de bruit. Notons que la variation de $\mathcal{S}^*$ n'excède pas 1 pixel, et ce même pour des niveaux de bruit élevés. Ainsi, cette échelle caractéristique basée sur la règle de la médiane s'avère résistante au bruit et varie linéairement en fonction de la taille de l'objet.



(A) (B)

FIG. 7.7 – Influence du bruit sur les ML, sur une image 128×128 représentant un disque (100 simulations par niveau de bruit). (A) Evolution du nombre de ML (en moyenne) ; (B) Comportement de l'échelle caractéristique $\mathcal{S}^*$ moyenne (courbe tracée) et écart-type (barres d'erreur).

7.4.2 Exemples pour une image naturelle

L'approche par ML permet de mettre en évidence un certain nombre d'objets (structures bidimensionnelles). Dans le cadre de la détection d'un seul objet, après que la TOC a été calculée (Fig. 7.8 (A–C)), il est possible de sélectionner un certain nombre de ML associées à celui-ci d'après le critère de jonction (Fig. 7.8 (A')). Chaque ML impliquée dans la jonction, notée L_i, conduit alors à une échelle caractéristique S_i^* (Fig. 7.8 (B')), ce qui mène alors à l'échelle caractéristique $\mathcal{S}^*$ de l'objet (Fig. 7.8 (C')). Dans un cadre bruité, cette approche par ML est également valable (Fig. 7.9), même si le bruit détruit un certain nombre de ML. Enfin, relevons que l'approche par ML permet de détecter différents objets ayant une structure bidimensionnelle, plus ou moins déformée (Fig. 7.9).

7.4.3 Comparaison avec l'approche classique

Nous comparons ici deux approches pour la détection de points d'intérêt : celle classique (extrema dans le Scale-Space 3D) et celle par ML (maxima le long de ML pertinentes). Ceci complète la comparaison théorique vue entre $MaxI$ et $MaxML$ (section 7.2.2). Il est important de noter que le mécanisme de sélection est différent, puisque dans un cas nous effectuons un seuillage sur la réponse tandis que dans l'autre nous exprimons des conditions sur les ML (dans l'approche ML, un seuillage est effectivement utilisé, mais celui-ci a une moindre importance). Par ailleurs, nous avons effectué des tests de répétabilité pour l'approche par ML (annexe A), ce qui permet alors de la comparer avec d'autres détecteurs de régions d'intérêt.

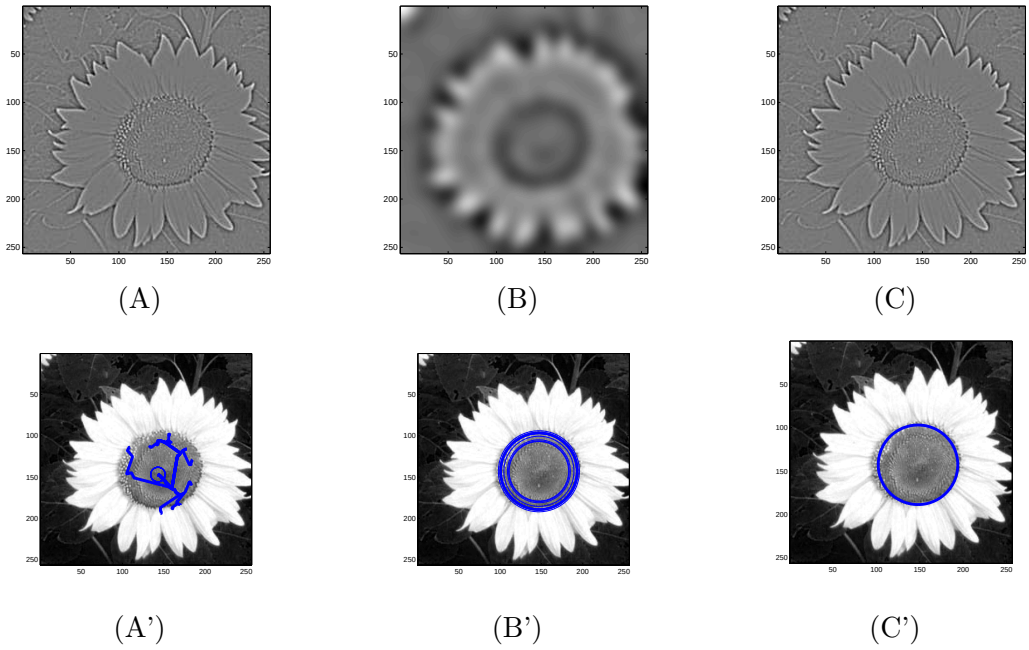


FIG. 7.8 – Approche par ML sur une image naturelle. (A) Transformée en ondelette continue (TOC) à l'échelle la plus fine $s = 1$; (B) TOC à une échelle intermédiaire ; (C) TOC à une échelle grossière ; (A') Ensemble des ML L_i correspondant à la structure centrale ; (B') Echelle caractéristique $\mathcal{S}_i^*$ (calculée d'après la distance Euclidienne) ; (C') Echelle caractéristique $\mathcal{S}^*$ (calculée d'après la règle de la médiane).

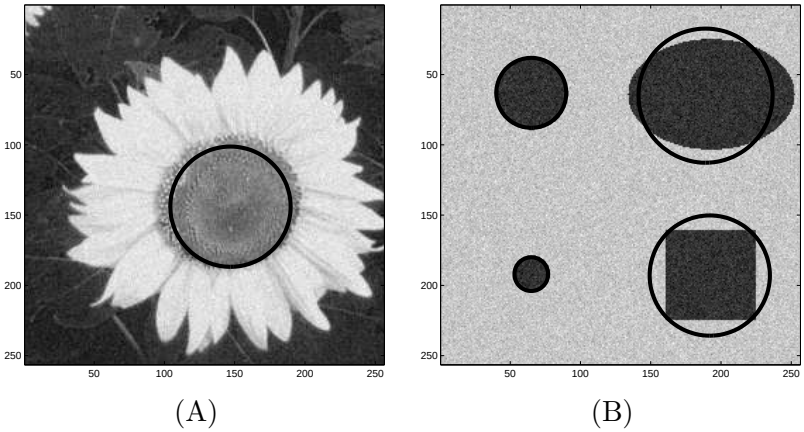


FIG. 7.9 – Exemple de régions d'intérêt sur une image géométrique et sur une image naturelle, dans un cadre bruité ($SNR = 20dB$). L'approche par ML permet la détection de points d'intérêt (x^*, y^*, s^*) et le calcul de l'échelle caractéristique $\mathcal{S}^*$, les régions d'intérêt sont des cercles centrés en (x^*, y^*) et de rayon $\mathcal{S}^*$.

Calculs d'échelles caractéristiques

Notons que l'échelle classique s^* correspond à un modèle de surface Gaussienne bidimensionnelle, tandis que la grandeur $\mathcal{S}^*$ que nous avons introduite correspond à une structure bidimensionnelle délimitée par certaines singularités, et sans contours internes. Les limites de la pertinence de $\mathcal{S}^*$ correspondent à des objets sans frontières nettes : ainsi, la Gaussienne 2D conduit à $\mathcal{S}^* = 0$; cependant la Gaussienne 2D bruitée conduit à une bonne estimation. Globalement, si l'objet donne lieu à très peu de MM, l'estimation de $\mathcal{S}^*$ est peu fiable. Relevons également que, tant pour des images tant géométriques que naturelles, les objets présents dans l'image s'apparentent plus souvent à ce modèle de structure bidimensionnelle – objets sans frontières nettes – qu'à des surfaces Gaussiennes.

Tab. 7.3 – Comparaison des échelles caractéristiques s^* et $\mathcal{S}^*$ sur deux images simples (carré d'arête $2 \cdot l$, disque de rayon l, 512×512).

	Carré	Disque
(Approche classique) s^*/l	0.796	0.698
(Approche par ML) $\mathcal{S}^*/l$	1.05	0.976
Moyenne des S_i^*/l	1.078	0.976
Ecart-type	0.092	0.008

Enfin, nous comparons ces deux échelles sur deux images l'une représentant un carré, d'arête de longueur l l'autre un disque, de diamètre l (Tab. 7.3). Les régions d'intérêt associées correspondent approximativement à l'objet détecté. Nous observons empiriquement que $\mathcal{S}^*$ et s^* sont des grandeurs caractéristiques de l'objet détecté, sans pour autant que ces valeurs soient égales. Ceci vient du fait que l'échelle caractéristique classique s^* est liée à la discrétisation des échelles de la décomposition, tandis que la nouvelle échelle caractéristique $\mathcal{S}^*$ est basée sur des grandeurs géométriques. En particulier, notons que la distribution des S_i^* (associées aux ML L_i) peut varier d'un objet à un autre (elle est plus homogène dans le cas du disque, Tab. 7.3).

Enfin, pour le calcul ultérieur de descripteurs, il est courant de considérer tous les pixels autour de (x^*, y^*), dans un rayon de $2s^*$, de manière à bien recouvrir l'objet [100]. Ceci montre bien l'importance de ces points situés au bord de l'objet, qui permettent d'estimer leur géométrie [32].

Régions détectées

A partir d'une image géométrique représentant des Gaussiennes de différentes tailles (avec ou sans bruit blanc Gaussien) et d'une image naturelle, nous obtenons les résultats représentés sur la figure 7.10. Dans le premier cas (Fig. 7.10 (A–A')), nous cherchons à détecter une seule région d'intérêt (associée à de nombreuses ML) ; les deux approches conduisent à détecter la

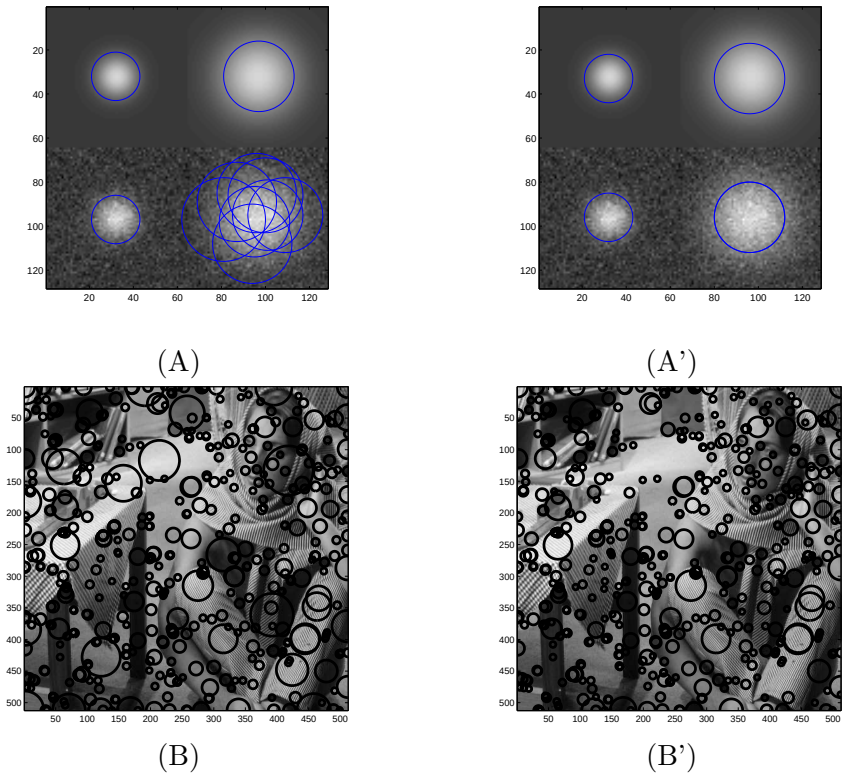


FIG. 7.10 – Détection de points d'intérêt sur une image géométrique (A, A') et sur une image naturelle (B, B'). (A, B) Approche classique ; (A', B') Approche par ML.

structure attendue, et celle par ML s'avère plus stable dans un cadre bruité. Dans le deuxième cas (Fig. 7.10 (B–B')), nous cherchons à détecter un certain nombre de régions d'intérêt, auxquelles sont associées un nombre restreint de ML. Sur l'image géométrique, nous obtenons les mêmes régions, sauf pour l'image fortement bruitée (Fig. 7.10 (A–A'), en bas à droite) : dans ce cas, les multiples régions d'intérêt extraites par l'approche classique correspondent à des maxima associés à quelques ML, tandis que le mécanisme de sélection dans l'approche par ML – jonctions d'un nombre suffisant de ML – ne conduit qu'à une seule région d'intérêt.

Dans le cas de l'image *Barbara* (Fig. 7.10 (A–B)), relevons que les objets apparaissent à une résolution moindre ; ainsi, à un objet donné, il peut correspondre un nombre restreint de ML. En particulier, lorsque les jonctions de ML ne sont pas clairement identifiées, l'approche par ML ne permet pas de de mettre en évidence certaines régions. Par ailleurs, il existe des structures correspondant à la jonction de plusieurs ML et à une réponse relativement faible, de sorte que celles-ci apparaissent sur la figure 7.10 (B) mais pas sur la figure 7.10 (B'). Ainsi, en plus du critère portant sur les jonctions de ML, il peut être utile d'appliquer un léger seuillage sur la réponse associée aux extrema dans S_{3d}.

7.4.4 Estimation géométrique

L'approche par ML permet de détecter un certain nombre d'objets. Chacun de ces objets est associé à un ensemble de ML $(L_i)_{i\in I}$ se rejoignant en un certain point de jonction (x^*, y^*, s^*). Une première estimation de la forme de l'objet est donnée par la région d'intérêt (circulaire) centrée en (x^*, y^*), et d'un certain rayon. Celui-ci peut être choisi proportionnel à l'échelle caractéristique classique s^*, mais nous préférons le choisir comme égal à $\mathcal{S}^*$: ceci se justifie par le fait $\mathcal{S}^*$ est une grandeur géométrique.

Il est alors possible de préciser cette forme, notamment grâce aux extrémités des ML. Si de nombreuses ML sont impliquées dans la jonction et que leurs extrémités sont réparties relativement uniformément sur le bord de l'objet, alors l'enveloppe convexe de ces points donne une bonne approximation de la forme de l'objet ; généralement, celle-ci est située à l'intérieur de l'objet, vu que les extrémités des ML sont localisées légèrement à l'intérieur de l'objet, près du bord. Notons au passage que l'enveloppe convexe des extrémités $(x_i^0, y_i^0)_{i\in I}$ est donnée par le polygone passant par certains points $(x_k^0, y_k^0)_{k\in K}, K \subset I$ (en nombre fini). Une forme plus lisse peut ensuite être obtenue en considérant les points $(x_k^0, y_k^0)_{k\in K}$ comme les nœuds d'une spline cubique par exemple [123]. Remarquons enfin que si certains nœuds sont répartis inégalement, ceci peut entraîner dans certains cas des formes inappropriées (par exemple, c'est le cas si les $(x_k^0, y_k^0)_{k\in K}$ sont concentrés sur une seule partie de l'objet).

Un cas particulier de cette estimation de la forme, lié à la détection de régions d'intérêt invariantes à l'affine, consiste à chercher une ellipse cohérente avec les points $(x_k^0, y_k^0)_{k\in K}$. L'estimation de cette ellipse peut s'écrire comme un problème d'optimisation [28] : étant donné $\alpha = (\alpha_i)_{1\leq p\leq 6} \in \mathbb{R}^6$, nous définissons $F_\alpha : \mathbb{R}^2 \longrightarrow \mathbb{R}$ par :

$$\begin{array}{rccl} F_\alpha & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & (x,y) & \longmapsto & F_\alpha(x,y) = \alpha_1 x^2 + \alpha_2 y^2 + \alpha_3 xy + \alpha_4 x + \alpha_5 y + \alpha_6 \end{array} \tag{7.12}$$

et par une régression aux moindres carrés, nous résolvons alors

$$\min_\alpha \sum_{k\in K} F_\alpha(x_k^0, y_k^0)^2 \tag{7.13}$$

ce qui donne une solution $\hat{\alpha} \in \mathbb{R}^6$ et ainsi l'ellipse cherchée est donnée par l'équation $F_{\hat{\alpha}}(x) = 0$. Ce problème peut encore être simplifié en fixant le centre de l'ellipse (x_c, y_c), par exemple à (x^*, y^*). ou le barycentre des $(x_k^0, y_k^0)_{k\in K}$. L'équation de l'ellipse recherchée s'écrit alors comme $Z^T A Z = 1$, où Z est le vecteur de coordonnées centrées et A une matrice semi-définie positive :

$$Z = \begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix} \text{ et } A = \begin{pmatrix} a & c \\ c & b \end{pmatrix}$$

Ce choix est motivé d'une part par la réduction du nombre de paramètres (3 au lieu de 6) et d'autre part par le fait que (x^*, y^*) est une donnée stable proche du centre.

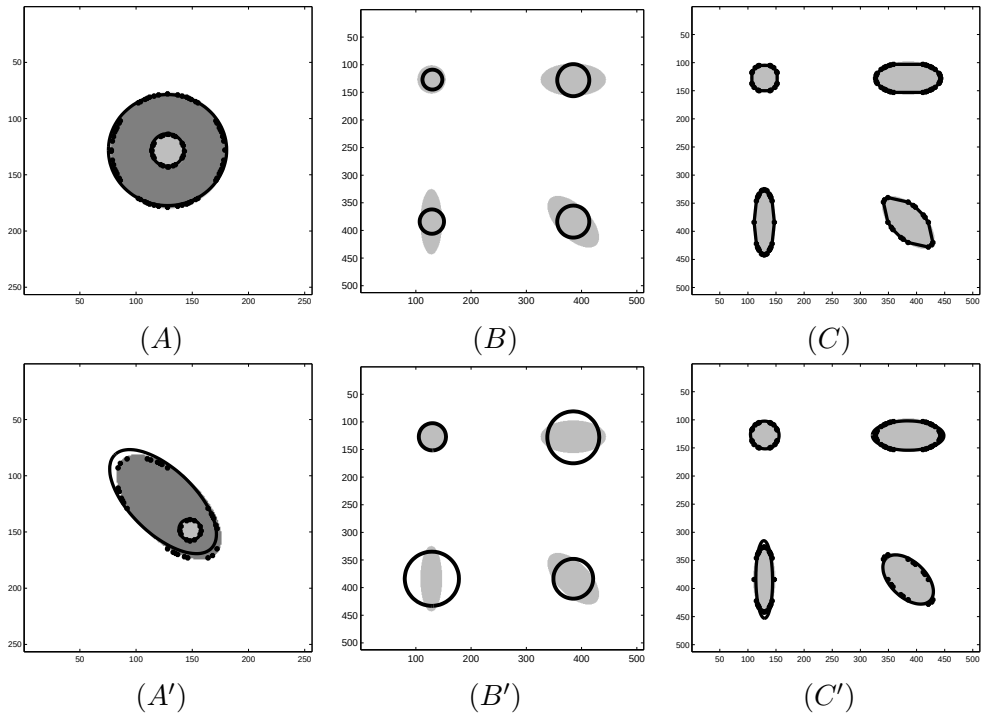
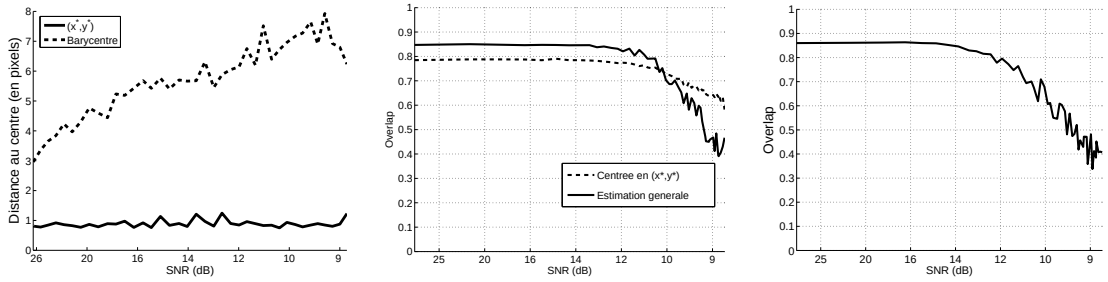


FIG. 7.11 – Exemples simples d'estimation de forme par l'approche ML. (A, A') Deux structures emboîtées, ellipse centrée en (x^*, y^*) ; (B) Régions d'intérêts données par (x^*, y^*, s^*) ; (B') Régions d'intérêts données par (x^*, y^*, S^*) (S^* étant obtenue par la règle de la médiane) ; (C) Enveloppe convexe des extrémités des ML (x_i^0, y_i^0) ; (C') Ellipse estimée de manière générale, à partir des points $(x_k^0, y_k^0)_{k \in K}$ (K correspondant aux points situés sur l'enveloppe convexe).

Afin de valider cette méthode d'estimation, nous considérons d'abord des images géométriques dont les données sont connues *a priori* (Fig. 7.11). Dans le cas de deux structures emboîtées toutes deux centrées en c, le centre obtenu $c^* = (x^*, y^*)$ est identique à c (Fig. 7.11 (A)). En revanche, dans le cas de deux structures emboîtées et décentrées, en notant c_1 et c_2 leurs centres respectifs ($c_1 > c_2$), nous remarquons que le centre de la plus grande structure c_1 est décalé par rapport au véritable centre (Fig. 7.11 (A')) ; cependant, l'écart entre ces points reste faible ($||c_1^* - c_1|| = 6.4$ pixels pour cette image 256x256). De manière plus générale, sur une image naturelle par exemple, la localisation du point (x^*, y^*) dépend de tout ce qui est présent dans la structure, et pas seulement des extrémités des ML. En particulier, un bord présentant une certaine courbure donnera lieu à plusieurs ML qui fusionneront à une certaine échelle s^*, en un point (x^*, y^*), mais cela ne correspondra pas nécessairement au centre d'un objet visuellement identifiable dans l'image – il correspondra plutôt au centre de courbure correspondant à la portion du bord considéré.



(A) (B) (C)

FIG. 7.12 – Estimation de la forme d'un objet par ML, en faisant varier le niveau de bruit. (A) Distance au véritable centre (disque 256x256) ; (B) Overlap entre l'objet (disque 256x256) et le cercle $(x^*, y^*, \mathcal{S}^*)$ ainsi qu'entre l'objet et l'ellipse estimée d'après les extrémités des ML (estimation sans *a priori* sur le centre) ; (C) Overlap entre l'objet (ellipse 256x256, Fig. 7.13) et l'ellipse estimée sans *a priori* sur le centre.

Afin de comparer les différentes méthodes d'estimation de cette ellipse, nous considérons deux objets – images de taille 128×128 représentant un disque et un disque déformé (Fig. 7.13) – auquels sont ajoutés un bruit blanc Gaussien plus ou moins important (100 simulations sont effectuées pour chaque niveau de bruit). D'abord, nous considérons le disque, et en faisant varier le niveau de bruit, il est possible de comparer la position du véritable centre avec celle :

- du point $c^* = (x^*, y^*)$
- du barycentre des extrémités des ML : $c_b = (x_b, y_b) = \sum_{k \in K} w_k (x_k^0, y_k^0)$
  (K : enveloppe convexe, $w_k = \frac{1}{|K|}$)

Le point c^* apparaît ici comme plus stable que le barycentre c_b, (Fig. 7.12 (A)). Ceci provient du fait qu'à partir d'un certain niveau de bruit, il ne reste qu'un petit nombre de ML pertinentes, et si de plus les extrémités de ces ML sont concentrées au même endroit, cette estimation est complètement faussée. D'autre part, comme les lignes correspondant au bruit traversent peu d'échelles, le point c^* reste stable tant qu'il reste des ML correspondant à la structure (i.e., que la structure est dominante par rapport au bruit). Enfin, à mesure que le niveau de bruit augmente, le nombre de ML pertinentes se réduit (jusqu'à 4 ML) ce qui rend l'estimation de l'ellipse encore possible, même si celle-ci est moins stable. Ainsi, pour fixer le centre de l'ellipse, c^* apparaît comme un choix judicieux.

Maintenant, nous nous intéressons à la stabilité de l'ellipse. A cet effet, nous introduisons la notion d'*overlap*, définie comme le rapport :

$$\frac{\text{Aire}(E_e \cap E_o)}{\text{Aire}(E_e \cup E_o)} \in [0, 1] \tag{7.14}$$

où E_e correspond au domaine de l'ellipse estimée et E_o à celui du disque (original ou déformé). Ceci permet alors de comparer les différentes méthodes d'estimation d'ellipse (sans fixer le centre *a priori*, et en fixant le centre à c^*).

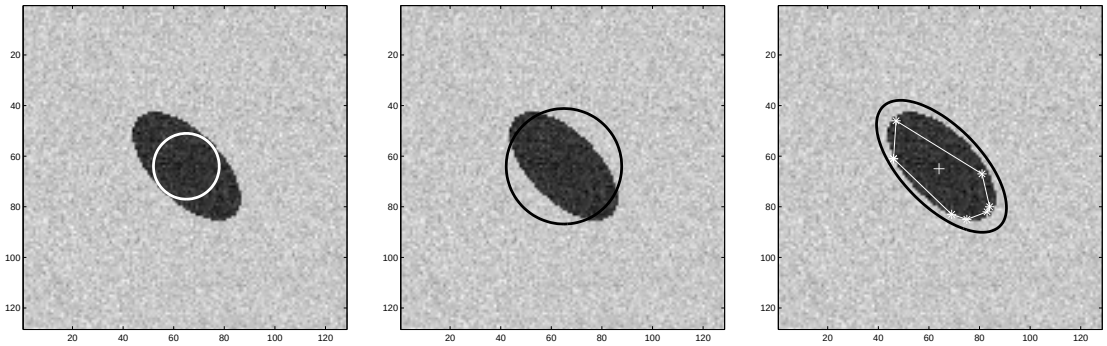
En ce qui concerne le disque (Fig. 7.12 (B), notons que pour des niveaux de bruit modérés, la méthode sans fixer le centre apparaît comme la plus performante (les ML sont en nombre suffisant) ; en revanche, pour des niveaux de bruit élevés, lorsqu'il ne reste qu'un nombre limité de ML, la méthode fixant le centre s'avère plus robuste que l'estimation générale (3 paramètres à estimer au lieu de 6 dans l'estimation générale).

En ce qui concerne le disque déformé, nous raffinons la méthode d'estimation de l'ellipse :

- calcul précédent de l'ellipse sans fixer le centre *a priori*, d'après les extrémités des ML ; cette ellipse est définie par un centre (x_Q, y_Q) et une matrice $M_Q \in SP_2$;
- normalisation de M_Q de manière à ce que l'aire de l'ellipse soit égale à $\pi \cdot (\mathcal{S}^*)^2$

Nous constatons le bénéfice de cette normalisation en comparant les figures 7.13 (B) et (C). Nous observons alors la dégradation de cette estimation de l'ellipse à mesure que le niveau de bruit augmente (Fig. 7.12) ; en particulier, nous relevons que l'overlap reste supérieur à 0.8 tant que le SNR reste inférieur à 12 (soit $A_{signal}/A_{bruit} \leq 4$).

En outre, il est intéressant de comparer qualitativement la différence entre le cercle (x^*, y^*, s^*) (échelle caractéristique classique, Fig. 7.13 (A)), le cercle $(x^*, y^*, \mathcal{S}^*)$ (nouvelle échelle caractéristique, Fig. 7.13 (B)), et les estimations basées uniquement sur les extrémités des ML (Fig. 7.13 (C)). Nous voyons que $\mathcal{S}^*$ est nettement supérieure à s^*. Ceci s'explique par le fait que les extrémités des ML les plus résistantes au bruit sont celles situées aux points de forte courbure (aux extrémités du grand axe de l'ellipse) ; ainsi, le calcul de S^* conduit à un cercle enveloppant le disque déformé tandis que s^* conduit à un cercle contenu dans celui-ci. D'autre part, nous observons que l'estimation à partir des ML est plus fine, et réalise une adaptation satisfaisante à la forme de l'objet. En conclusion, pour un objet détecté par ML, l'estimation générale est la plus pertinente s'il y suffisamment de ML associées (au moins 4). Néanmoins, dans le cas où il a peu de ML (comme cela est le cas des images naturelles), il peut être intéressant de considérer une estimation de l'ellipse fixant le centre en (x^*, y^*).



(A) (B) (C)

FIG. 7.13 – Estimation de forme sur une ellipse 128×128 bruitée ($SNR = 20dB$) : (A) Cercle centré en (x^*, y^*), de rayon s^* ; (B) Cercle centré en (x^*, y^*), de rayon $\mathcal{S}^*$; (C) Enveloppe convexe des extrémités des ML (associées à l'objet) et ellipse estimée d'après celles-ci.

Nous avons mentionné que Lindeberg a proposé un algorithme de *Shape Adaptation* permettant d'obtenir une ellipse adaptée à l'anisotropie de la structure sous-jacente à une région d'intérêt (chap. 4, [79]). Celui-ci estime une matrice représentant la forme quadratique associée à une ellipse, ce qui constitue une région d'intérêt invariante à l'affine. Cependant, il n'y a pas à notre connaissance de résultats théoriques sur la convergence de cet algorithme, et en pratique cette méthode converge dans seulement 50% des cas [97]. En revanche, l'approche par ML donne des points situés sur le bord de l'objet, et le problème de l'estimation de la forme est formulé uniquement à partir des extrémités des ML (x_i^0, y_i^0) et de la localisation de la jonction en (x^*, y^*) [31], ce qui constitue un avantage.

7.5 Conclusions et perspectives

Le point fondamental du détecteur par ML est que les régions extraites peuvent être reliées à des singularités dans l'image, ce qui contraste avec d'autres détecteurs pour lesquels le lien entre les régions d'intérêt et la structure de l'image est parfois difficile à appréhender. Etant donné un objet, une fois que les ML associées ont été déterminées, les extrémités de ces ML peuvent être directement mises en relation avec l'image, vu que les MM aux échelles fines sont associées à des singularités.

Un autre élément à relever est que les structures ne sont pas redondantes : les régions ne s'intersectent pas, même si elles peuvent être imbriquées. Or, pour de nombreux détecteurs, les régions d'intérêts s'intersectent, et en particulier nous observons de nombreux recouvrements, de sorte que plusieurs régions sont associées à un seul objet. Dans ce cas, l'approche par ML permet d'obtenir une seule région, associée à un certain nombre de ML.

En outre, par construction, la taille des régions est comprise entre l'échelle $s_{\min}$ où l'on considère que l'on dépassé le niveau de bruit et l'échelle maximale considérée $s_{\max}$. Ainsi l'approche par ML conduit à des régions de tailles intermédiaires ; ceci exclut les plus petites structures, qui sont généralement peu exploitables étant donné la résolution de l'image ; la valeur de $s_{\max}$ permet de s'adapter à une taille quelconque, sachant que dans l'optique de détection caractéristiques locales, il est souhaitable que cette valeur soit raisonnable (typiquement $s_{\max} \leq 40$).

Ces différents éléments conduisent à une densité de points d'intérêts (nombre de régions pour une image donnée) relativement faible, comparativement à des détecteurs donnant des régions fortement redondantes. Le fait qu'un détecteur donne des régions redondantes assure une certaine stabilité du détecteur : si l'image subi une certaine transformation, même si certaines régions disparaissent, il en existe d'autres qui jouent un rôle de remplacement. Cependant, cela se fait au détriment de la cohérence entre les régions détectées et le contenu de l'image. En conclusion, le détecteur par ML met en évidence des régions qui ne donnent pas de représentation complète de l'image, mais qui sont cohérentes avec des singularités visuellement identifiables.

Perspectives

La question de la pertinence des régions extraites est délicate, car les éléments les plus distinctifs pour un observateur humain sont très hétérogènes (objets, contours, ...). Ainsi, le modèle de Gaussienne apparaît comme limité pour la détection de structures pertinentes (notamment pour des applications telles que la reconnaissance ou la mise en correspondance). Plus généralement, les régions issues des détecteurs actuels sont liées à un modèle d'objet (surface Gaussienne, ...). Nous autorisons ces objets à être plus ou moins déformés, mais en tout cas cela exclut certaines structures, comme des données ponctuelles.

Dans la perspective d'un nouveau type de détecteur, il serait intéressant d'effectuer une classification des structures bidimensionnelles, comme cela a été fait pour les structures monodimensionnelles (où la régularité Lipschitzienne caractérise le type de contour). Plutôt que de détecter des *blobs* sans connaître la structure sous-jacente (forme géométrique, singularités), il serait intéressant de détecter de tels objets soit séparément (différents opérateurs) soit simultanément (à condition de pouvoir les caractériser *a posteriori*). De plus, différentes grandeurs permettraient de décrire sommairement l'objet détecté (échelle caractéristique, régularité Lipschitzienne, caractéristiques géométriques de l'objet). Une telle approche permettrait de détecter différents objets tout en les caractérisant, et ainsi d'obtenir une diversité de régions d'intérêt tout en conservant une richesse de leur contenu.

Chapitre 8

Conclusion générale

Nous avons montré dans cette thèse que les ondelettes constituent des outils pertinents pour la détection de structures présentes dans l'image, qu'elles soient monodimensionnelles (comme des contours) ou bidimensionnelles (comme des entités de type *blob*).

8.1 Contributions principales

Propriétés des ML (1D)

Dans un cadre monodimensionnel, nous avons d'abord établi, pour une TOC normalisée L^p utilisant une ondelette suffisamment régulière, que les coefficients d'ondelettes sont solution d'une EDP, ce qui nous a permis d'étudier le comportement des extrema de la TOC en fonction de l'échelle. Du point de vue théorique, nous avons vu que la détermination des trajectoires associées à des lignes d'extrema n'était pas immédiate, même dans le cas d'un signal simple. Par ailleurs, nous avons étudié empiriquement sur le comportement des ML, lignes d'extrema particulières, associées à certaines singularités. Nous avons vu en particulier qu'il existe des points du *Scale-Space*, où, simultanément, il y a jonction de ML et maximum local de la réponse le long de ces ML. Ces points peuvent être alors associés à des structures locales (non-ponctuelles) présentes dans le signal. Soulignons que ces structures n'apparaissent pas aux fines échelles, où seules des singularités les délimitant sont connues.

Lignes de maxima (2D)

Après avoir rappelé les notions de ML et de LC, nous avons précisé des règles de construction empirique de celles-ci, et étudié le problème d'existence de ces lignes. Nous avons vu en outre l'effet d'une transformation affine sur les ML. En comparant les comportements des ML et des LC en fonction de l'échelle, nous avons vu que tandis que la dérive des ML est prononcée (depuis une singularité vers une jonction de ML), celle des LC est limitée (localisation sur les contours).

Régularité Lipschitzienne

Après avoir rappelé comment calculer des régularités ponctuelles par des approches multiéchelles, nous avons montré que ces régularités constituent des grandeurs caractéristiques des singularités associées aux lignes de maxima, et qu'elles présentent une certaine robustesse à des déformations locales de l'image (la rédaction d'un article de synthèse est en cours).

Sélection de points et d'échelles remarquables

Nous avons montré des liens entre ondelettes et *Scale-Space*, et proposé quelques généralisations que permettent les ondelettes. Nous avons également présenté un nouveau mécanisme de sélection d'échelle, fondé sur les ML, que nous avons comparé à la sélection classique. Ceci justifie l'emploi des ondelettes pour formuler un détecteur de régions d'intérêt.

Détecteur par ML

La contribution essentielle de cette thèse est la formulation d'une méthode originale de détection de régions d'intérêt, fondée sur les Maxima Lines (ML). Cette approche se distingue d'autres méthodes dans la mesure où le critère de sélection est basé sur les propriétés des ML plutôt que sur un seuillage. en particulier, nous avons associé différentes structures bidimensionnelles à un certain nombre de ML (ML admettant un maximum local et fusionnant). Nous avons également défini une nouvelle échelle caractéristique, dont nous avons montré la robustesse et que nous avons comparé avec l'échelle caractéristique classique.

Notons que dans l'approche classique, les détecteurs existants se limitent à donner des régions d'intérêt, l'estimation de déformations locales et l'extraction de l'information contenue dans celles-ci étant des tâches distinctes. En ce qui concerne l'estimation des déformations, en remarquant que les extrémités des ML correspondent à des singularités proches des frontières de l'objet, nous avons formulé une nouvelle méthode d'estimation de sa forme, ce qui constitue une alternative à l'algorithme de *Shape Adaptation* proposé par Lindeberg (qui en pratique, ne converge que dans 50% des cas). De surcroît, relativement à l'étape de description de la région, l'intérêt des ML est qu'elles permettent d'anticiper cette étape, au sens où elles permettent de décrire sommairement l'objet. En effet, pour un objet donné, les ML associées permettent de dire si la région contient telle ou telle singularité, grâce au calcul des régularités Lipschitziennes.

Performance du détecteur

Nous avons montré que la formulation d'un détecteur de points d'intérêt par ML était réalisable, et apportait de nouveaux éléments, comme les extrémités des ML ou les régularités associées. En réalisant des tests de répétabilité, nous obtenons des résultats moins performants que ceux des détecteurs classiques. Un premier élément d'explication est la diversité

des conditions sur les ML, qui fait que la performance du détecteur dépend de nombreux paramètres délicats à ajuster. Par ailleurs, notons que pour une image naturelle, les objets recherchés apparaissent à une résolution moyenne, de sorte que le nombre de ML associées à tel ou tel objet peuvent être peu nombreuses. Ainsi, les objets associés à peu de ML sont parfois instables, tandis que celles associés à plus de 5 ML s'avèrent très stables.

Plus généralement, notons que les détecteurs classiques mettent en évidence différents *blobs*, et les objets associés n'apparaissent pas toujours clairement. Nous relevons que certains d'entre eux correspondent effectivement à des structures 2D associées à des jonctions de ML. Cependant, il existe une grande diversité de ces structures 2D, et l'approche par ML n'en détecte qu'une partie. L'intérêt de celle-ci est que les régions extraites peuvent être mises en correspondance avec certaines singularités de l'image. Ceci constitue ainsi un premier pas vers la caractérisation des structures 2D détectées, comme nous avons vu que la régularité Lipschitzienne permet de caractériser une structure 1D (détectée par LC).

8.2 Perspectives

Propriétés des ML

Du point de vue théorique, il sera intéressant de faire le lien – dans un cadre monodimensionnel puis bidimensionnel – entre les points de jonction de ML dans le *Scale-Space* et les maxima locaux de la réponse le long de ML. Une autre perspective est l'étude du comportement des ML dans un cadre plus général : leur existence pour des diffusions plus générales que celle de la chaleur, et leurs propriétés lorsqu'une transformation géométrique quelconque est appliquée à l'image. Ceci précisera les résultats empiriques vu ici.

Extension de l'approche par ML

Relevons que l'approche par ML peut être étendue de différentes manières. Une première possibilité consiste à utiliser un autre opérateur que l'ondelette ΔG, la construction des ML étant identique. Une seconde possibilité consiste à considérer d'autres maxima que les MM (dont la définition est assez stricte), ce qui conduit à de nouvelles lignes. A cet égard, relevons que les ML et les LC présentent des propriétés complémentaires : tandis que l'intérêt des LC réside dans le lien avec les contours, et le calcul de la régularité Lipschitzienne en un grand nombre de points, l'élément le plus remarquable des ML est l'existence de points de jonction dans le Scale-Space, associés à une réponse localement maximale. Ainsi, il sera pertinent de définir de nouvelles lignes de maxima, associées à de nombreuses singularités de l'image, et possédant également cette propriété de jonction–maximum local. Un détecteur fondé sur de telles lignes permettra alors d'obtenir des performances accrues, tout en étant à même de caractériser les objets détectés.

Caractérisation de structures 2D

Nous avons vu que les structures 2D sont plus complexes que les structures 1D, que ce soit du point de vue de leur détection ou de leur caractérisation (identification du type de singularité). En particulier, le modèle de *blob* classique, la Gaussienne 2D, apparaît comme limité. Ainsi, il apparaît comme nécessaire d'élaborer un modèle plus complet d'entités 2D, puis de formuler un détecteur adapté ainsi que des moyens de caractérisation. Une telle démarche conduirait alors à des triplets du type (entité, détecteur, éléments de caractérisation) tels que celui que nous avons décrit dans cette thèse :

$$\text{blob, détecteur ML,} \begin{cases} \text{ML impliquées dans une jonction} \\ \text{régularités Lipschitziennes associées aux extémités des ML} \\ \text{échelle caractéristique } \mathcal{S}^* \\ \text{estimation de la forme par ML} \end{cases}$$

Calcul de descripteurs

Dans cette thèse, nous nous sommes intéressés principalement à un aspect de la vision par ordinateur : la détection de régions d'intérêt. Celle-ci constitue une étape préliminaire avant le calcul de descripteurs associés aux régions détectées, qui sont alors utilisés dans les applications pratiques, comme la reconnaissance d'objets ou la mise en correspondance entre deux images. Ces descripteurs sont souvent fondés sur la distribution des gradients. Ainsi, de nouveaux descripteurs pourront être formulés, en intégrant les régularités Lipschitziennes ponctuelles (dont l'estimation par les LC est robuste). Enfin, plus généralement, il sera intéressant d'utiliser les coefficients d'ondelettes pour calculer un descripteur invariant ou robuste à différentes transformations de l'image.

Scale-Space et ondelettes

Les liens que nous avons remarqué entre le *Scale-Space* et les ondelettes ne semblent pas limités aux détecteurs de type *blob*, et il apparaît comme prometteur de tisser des liens entre les ondelettes et des opérateurs utilisés en vision par ordinateur (souvent non-linéaires). Cela permettra de mieux cerner le type de caractéristique extraites par ces derniers, soit par le calcul de grandeurs caractéristiques, soit par la mise en relation entre un objet et des singularités de l'image. En outre, nous avons vu que les ondelettes permettent de généraliser certains outils du *Scale-Space*, par l'utilisation de fonctions échelles 2D au lieu de noyaux Gaussiens. Ceci ouvre ainsi la voie à des généralisations plus sophistiquées des outils de détection issus de la vision par ordinateur.

Annexe A

Tests de répétabilité

A.1 Performance d'un détecteur

Problématique

Les différents détecteurs de points d'intérêt existant en vision par ordinateur sont formulés de différentes manières, et leurs auteurs insistent surtout sur leur cohérence interne : d'abord, un modèle de structure à détecter (coins, bords) est proposé, puis un algorithme de détection est détaillé, ce qui conduit à des résultats sur une image (géométrique, naturelle, texturée) où les régions détectées sont plus ou moins cohérentes avec ce que nous pouvons percevoir visuellement. Il est alors intéressant de pouvoir comparer différents détecteurs, d'après les régions qu'elles donnent sur un ensemble d'images. Naturellement, cet ensemble doit présenter une certaine diversité, à la fois en termes de contenu (textures, structures monodimensionnelles et bidimensionnelles) ainsi qu'en termes de situation (changement d'échelle, perspectives différentes). Idéalement, si nous considérons une scène donnée représentée de manières différentes (l'angle de vue, l'échelle, la luminosité ou le contraste pouvant avoir varié), les régions extraites devraient correspondre au même contenu. La notion de répétabilité quantifie dans quelle mesure nous nous rapprochons de cela, et détermine le degré de stabilité de la méthode utilisée. Etant donné que de multiples transformations sont considérées, les approches proposées ne doivent pas être systématiquement invariantes, il suffit souvent qu'elles soient robustes.

Mesure d'erreur de recouvrement

Nous considérons ici deux images (images originale et déformée) reliées par une homographie H connue. Nous supposons également que les régions s'expriment comme des ellipses, celles-ci étant de la forme $z^T M z = 1$, où $M \in M_3(\mathbb{R})$ et $z \in \mathbb{R}^3$ (la dimension 3 permet de décrire à la fois les translations, les homographies ainsi que les déformations affines). Une telle région recouvre un domaine du plan 2D noté D_M.

Définition 38. *(Erreur de recouvrement – Overlap error)*
Etant donné deux régions R_A et R_B (issues de A et B) associées aux matrices M_A et M_B, l'erreur de recouvrement est définie comme :

$$\epsilon = 1 - \frac{D_{M_A} \cap D_{H^T M_B H}}{D_{M_A} \cup D_{H^T M_B H}} \tag{A.1}$$

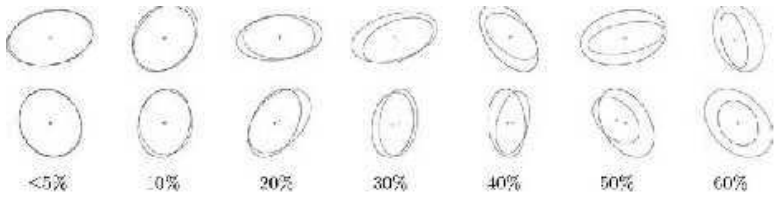


FIG. A.1 – Erreur de recouvrement entre deux régions

Correspondances

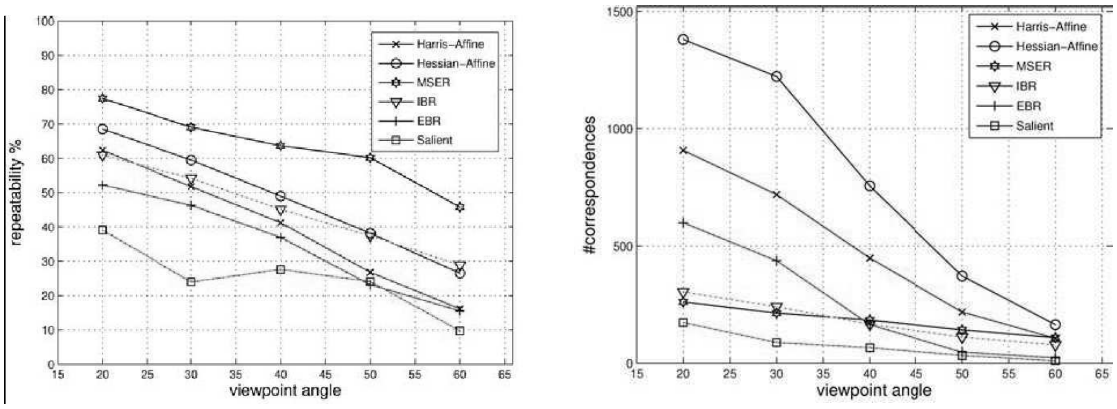
Nous décidons que deux régions se correspondent si l'erreur de recouvrement est suffisamment faible ; à cet égard, relevons qu'une erreur de 20% correspond à une différence de 10% entre les rayons des deux régions, de sorte que tant que l'erreur reste inférieure à 50%, il est possible de mettre en correspondance les deux régions à l'aide d'un descripteur robuste (Fig. A.1). Une fois que les erreurs de recouvrement ont été calculées entre toutes les régions de A et toutes celles de B, cela conduit à définir un certain nombre de régions qui peuvent être potentiellement appariées, ce qui définit $n_{corresp}$ régions correspondantes. Remarque : vu qu'en pratique, les images sont définies sur un domaine fini, certaines régions ne peuvent pas être mises en correspondance : si une image a subi une déformation comme un changement d'échelle, certaines parties de la scène présentes dans l'image originale ne le seront pas dans l'image déformée ; ces régions ne sont alors pas prises en compte.

Notion de répétabilité

Définition 39. *(Score de répétabilité – Repeatability score)*
Etant donné deux images A et B, à partir desquelles un détecteur a mis en évidence n_A et n_B régions d'intérêt, et où $n_{corresp}$ correspondances ont été déterminées entre les régions issues de A et celle issues de B, le score de répétabilité entre A et B est défini par le ratio

$$RS(A, B) = \frac{n_{corresp}}{min(n_A, n_B)} \tag{A.2}$$

Idéalement, chaque région de l'image déformée correspond à une autre région de l'image originale, le contenu visuel étant le même, et le score de répétabilité d'un détecteur parfait est égal à 1.



(A) (B)

FIG. A.2 – Evaluation de différents détecteurs, pour une séquence d'images correspondant à un changement de point de vue : (A) Score de répétabilité ; (B) Nombre de correspondances.

A.2 Evaluation d'un détecteur – Test de répétabilité

Afin d'attribuer une mesure de performance (ou score) à un détecteur donné, une approche proposée dans [101] consiste à se donner différentes séquences d'images $(A_i)_{0 \leq i \leq K}$, ici 8 séquences de 6 images (deux images de chaque séquence sont données, Fig. A.3 et A.4). Par hypothèse, A_0 est l'image originale, et les images $(A_i)_{i>1}$ peuvent se déduire de celle-ci, par une opération d'un certain type (changement d'échelle, rotation, compression JPEG). Nous supposons également que la transformation est quantifiée par une fonction $i \mapsto param(i)$ (taux de compression JPEG, angle de la rotation, facteur d'un changement d'échelle).

1. Pour chaque image, déterminer les régions d'intérêts associées ;
2. Mesurer le recouvrement (en termes d'aire) entre les régions de A_0 et celles de $(A_i)_{i>0}$, puis calculer le nombre de nombre de correspondances entre A_0 et A_i pour tout $i > 0$;
3. Calculer alors le score de répétabilité $RS(A_0, A_i)$ pour tout $i > 0$;
4. La courbe de répétabilité associée à la transformation étudiée est alors donnée par

$$(param(i), RS(A_0, A_i)), 1 \leq i \leq K \tag{A.3}$$

Ainsi, pour chaque type de transformation, nous mesurons le degré d'invariance ou de robustesse du détecteur. L'ensemble des courbes de répétabilité pour différentes transformations permet alors de rendre compte de la performance générale du détecteur de régions d'intérêt considéré. Afin de réaliser un test de répétabilité, des séquences d'images ainsi qu'une procédure *ad hoc* sont disponibles à l'url `www.robots.ox.ac.uk/vgg/research/affine`. Ceci permet alors de comparer les performances de différents détecteurs, par le tracé du nombre de correspondances et du score de répétabilité (Fig. A.2).

Dans tels graphes, l'abscisse correspond à une image de la séquence, et à un certain degré de la transformation (changement de point de vue de plus en plus important par exemple) ; quant à l'ordonnée, elle représente :

- soit le nombre de correspondances (établies d'après la paramètre ϵ, eq. (A.1)),
- soit le score de répétabilité, représenté ici en pourcentage (eq. (A.2)).

Plus les valeurs sont élevées, plus le détecteur est performant.

A.3 Application au détecteur par ML

Méthodologie

Nous avons appliqué le détecteur par ML aux 48 images de test, organisées en 8 séquences de 6 images, chaque image étant traitée de manière distincte. Ceci conduit à différents points d'intérêt (Tab. A.1). Nous pouvons alors effectuer des tests de répétabilité pour les 8 séquences, les régions d'intérêt étant :

- soit des régions circulaires, associées à $(x^*, y^*, \mathcal{S}^*)$, (x^*, y^*) correspondant à la jonction de ML et $\mathcal{S}^*$ à échelle caractéristique calculée d'après les ML (vu en 7.3.2).
- soit des régions elliptiques, associées à (x^*, y^*, Σ_Q), correspondant à l'estimation de la forme d'un objet – l'ellipse est estimée d'après les extrémités des ML, son centre étant fixé à (x^*, y^*) (vu en 7.4.4).

En outre, pour quelques images de test – pour chaque séquence, il s'agit des images I_1 et I_6 du tableau A.1 – nous avons représenté ces régions circulaires (Fig. A.5) ainsi que ces régions elliptiques (Fig. A.6).

TAB. A.1 – Nombre de régions d'intérêt détectées pour les 6 images de chaque séquence (I_1 et I_6 sont représentées sur les figures A.3 et A.4).

Séquence	I_1	I_2	I_3	I_4	I_5	I_6
Blur1	398	560	539	531	484	511
Blur2	347	314	446	609	484	448
Compression	179	186	195	210	234	211
Light	251	233	241	238	222	252
ZoomRotation1	317	351	333	251	273	245
ZoomRotation2	294	285	251	295	306	240
Viewpoint1	370	370	398	353	320	327
Viewpoint2	234	260	262	257	273	272

Grille de lecture des courbes de répétabilité

Nous représentons les résultats de répétabilité concernant les régions circulaires d'une part (Fig. A.7 et A.8) et elliptique d'autre part (Fig. A.9 et A.10). Relevons que sur la figure A.2, les courbes tracées représentent la performance de différents détecteur pour ε fixé ; sur les figures A.7 à A.10, les courbes tracées correspondent au détecteur par ML, pour différentes valeurs de ε (resp. 0.4, 0.5 et 0.6). Sur les figures A.7 à A.10, nous représentons à gauche le score de répétabilité (exprimé en pourcentage) et à droite le nombre de correspondances. Le score (tout comme le nombre de correspondances) est issu de la comparaison entre l'image originale I_1 et l'image transformée $(I_i)_{2\leq i\leq 6}$, et représenté par un marqueur (un rond pour $\varepsilon = 0.6$, un carré pour $\varepsilon = 0.5$, une croix pour $\varepsilon = 0.4$).

Analyse des résultats

En ce qui concerne les séquences Blur1, Blur2, Compression et Light, le score de répétabilité est satisfaisant et s'avère stable (Fig. A.7), ce qui est cohérent avec la stabilité de régions associées à plusieurs ML.

En ce qui concerne les séquences ZoomRotation1 et ZoomRotation2, le score est plus faible que pour les séquences précédentes, ce qui s'explique par le fait que certains éléments (comme le ciel nuageux dans ZoomRotation2) ne sont pas présents dans toute la séquence. Le score diminue avec le degré de la transformation (angle de rotation, facteur du changement d'échelle), mais cette diminution est limitée lorsque ce degré est faible (facteur de changement d'échelle inférieur à 2.5).

Quant aux séquences Viewpoint1 et Viewpoint2, notons qu'il s'agit de transformations de l'image plus générales que les déformations affines. Dans ce cadre, les performances du détecteur par ML sont plus faibles que pour les séquences précédentes, en particulier, le score diminue rapidement à mesure que la transformation est de plus en plus radicale.

Enfin, notons qu'en intégrant les déformations affines, le score gagne en stabilité et perd en performance (Fig. A.9), ou n'apporte pas de gain (Fig. A.10). Nous avons montré que l'estimation d'une ellipse était pertinente dans un cadre simple (chap. 7), et nous voyons apparaître ici ses limites. Une solution envisageable consisterait à n'utiliser cette estimation d'ellipse seulement dans le cas d'un objet associé à nombreuses ML.

Globalement, retenons que pour une première implémentation, les performances de ce détecteur fondé sur les ML s'avèrent correctes, comparées à celles des versions initiales des détecteurs issus de la vision par ordinateur, qui ont été améliorées par la suite. Enfin, la performance limitée du détecteur par ML est liée au fait que nous nous sommes focalisés sur un certain type d'objet (ceux associés à des jonctions de ML), ce qui constitue une partie seulement de l'ensemble des structures bidimensionnelles présentes dans l'image (entités que nous pouvons identifier visuellement).

FIG. A.3 – Séquences correspondant à des modifications de l'image Blur1 et Blur2 : l'image est lissée ; Light : changement d'illumination ; Compression : compression JPEG (note : nous montrons ici les images I_1 et I_6 de chaque séquence).

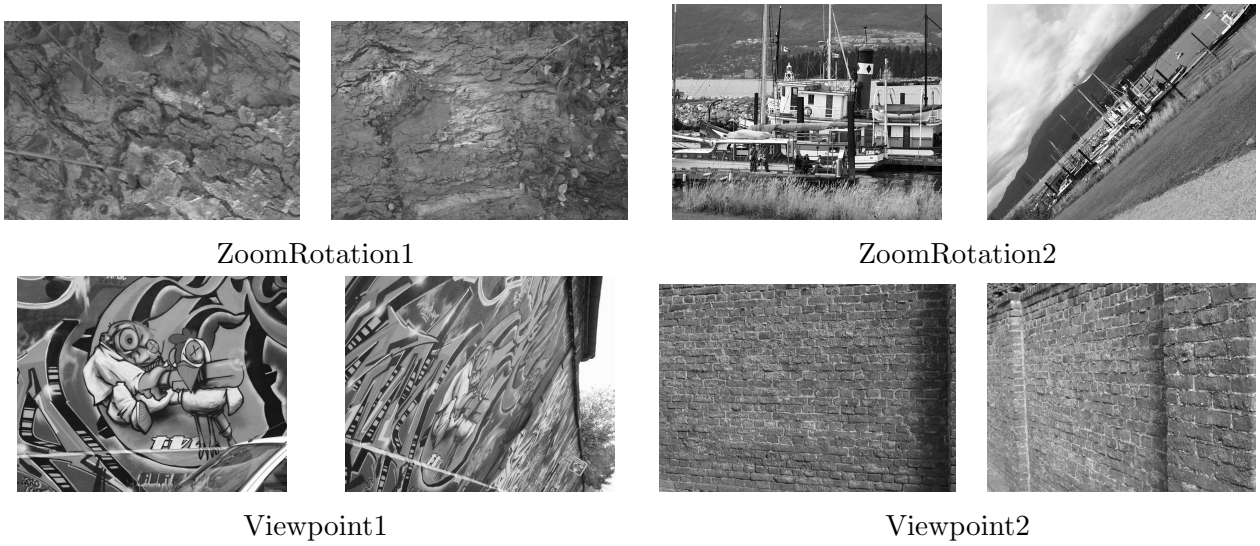


FIG. A.4 – Séquences correspondant à des déformations de l'image : ZoomRotation1, ZoomRotation2 : changement d'échelle et rotation ; Viewpoint1, Viewpoint2 : changement de point de vue

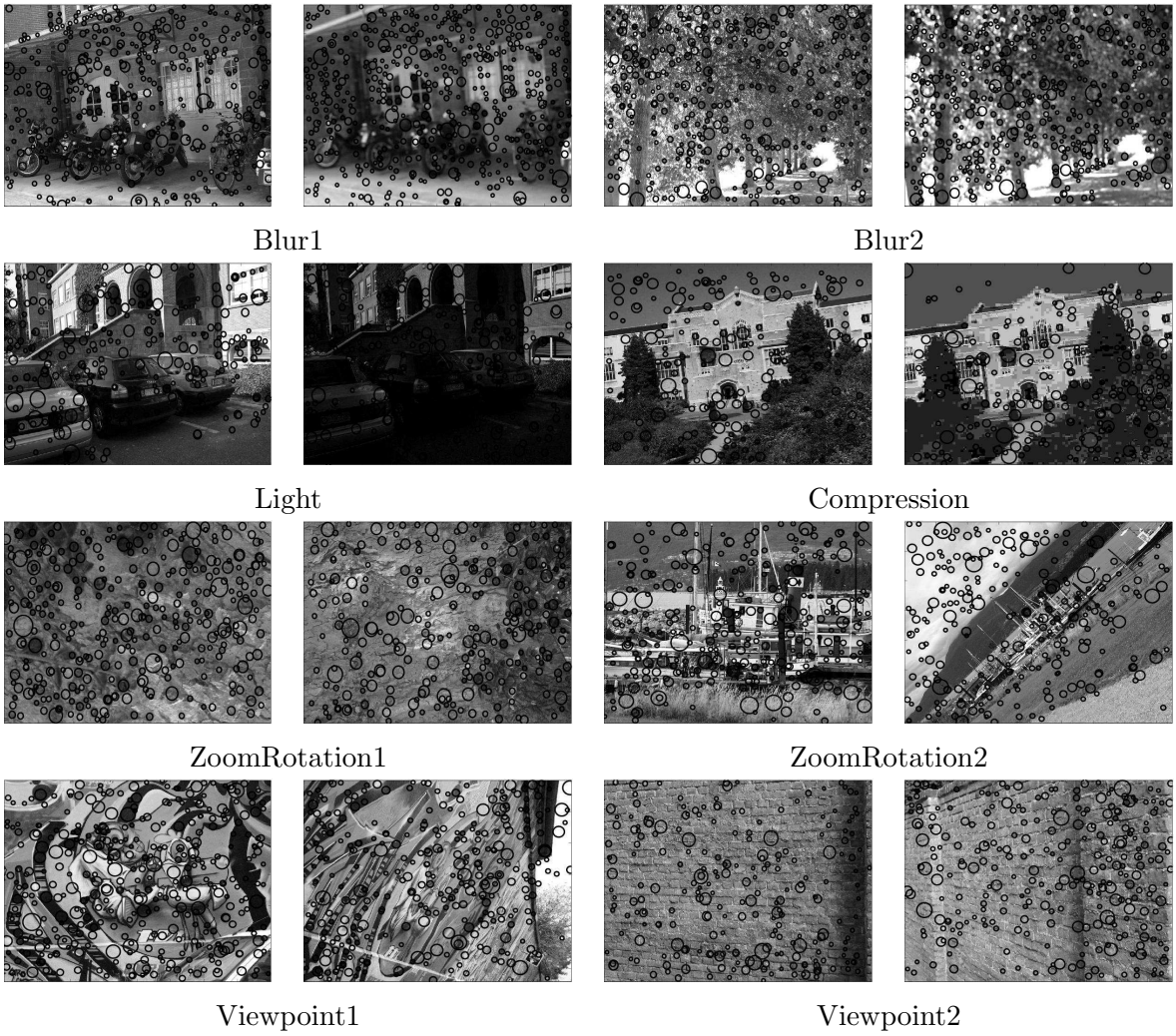


FIG. A.5 – Détection par ML de point d'intérêt $(x^*, y^*, \mathcal{S}^*)$ sur différentes images de test

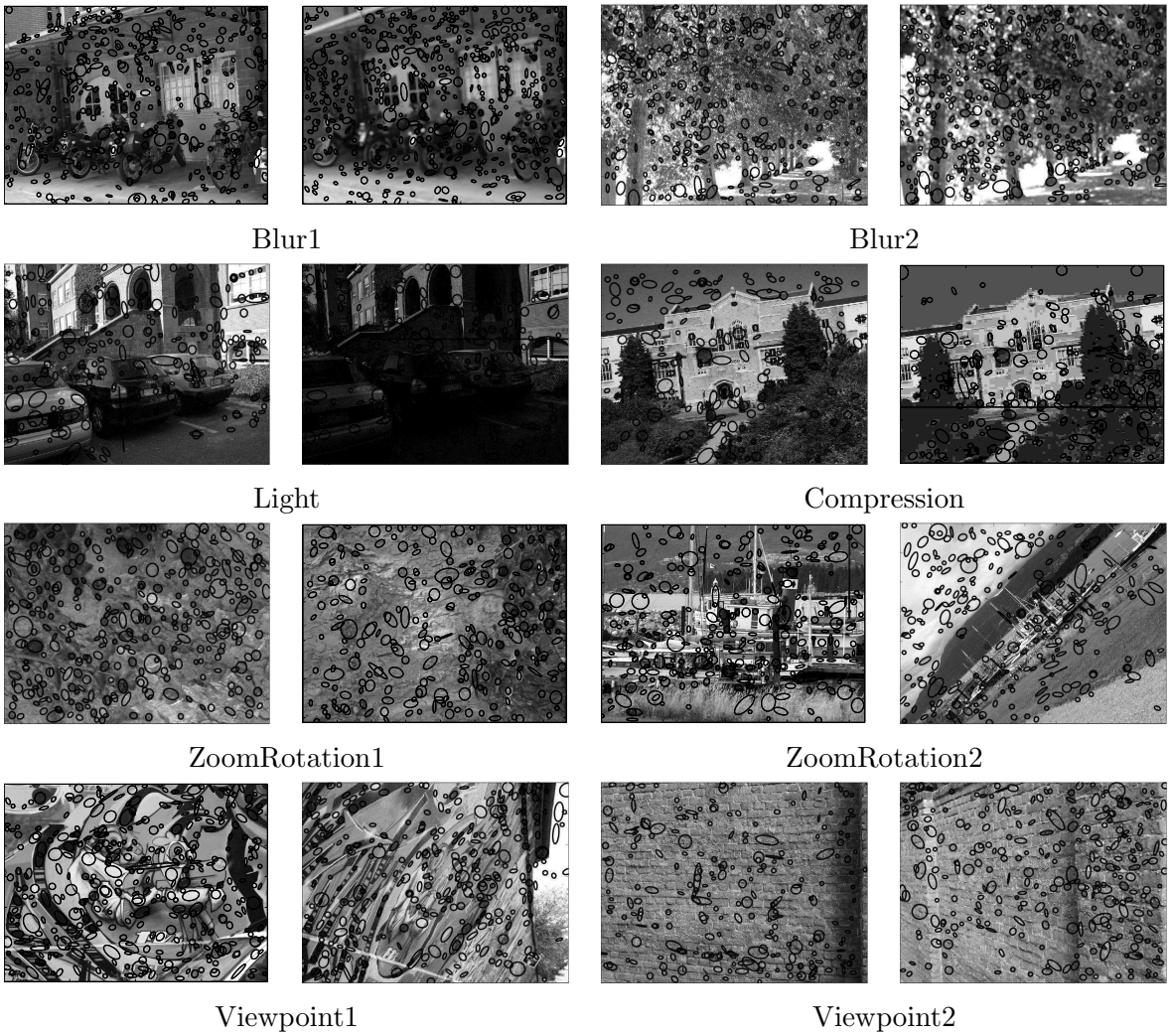


FIG. A.6 – Détection par ML de régions d'intérêt (x_C, y_C, Σ_Q) sur différentes images de test.

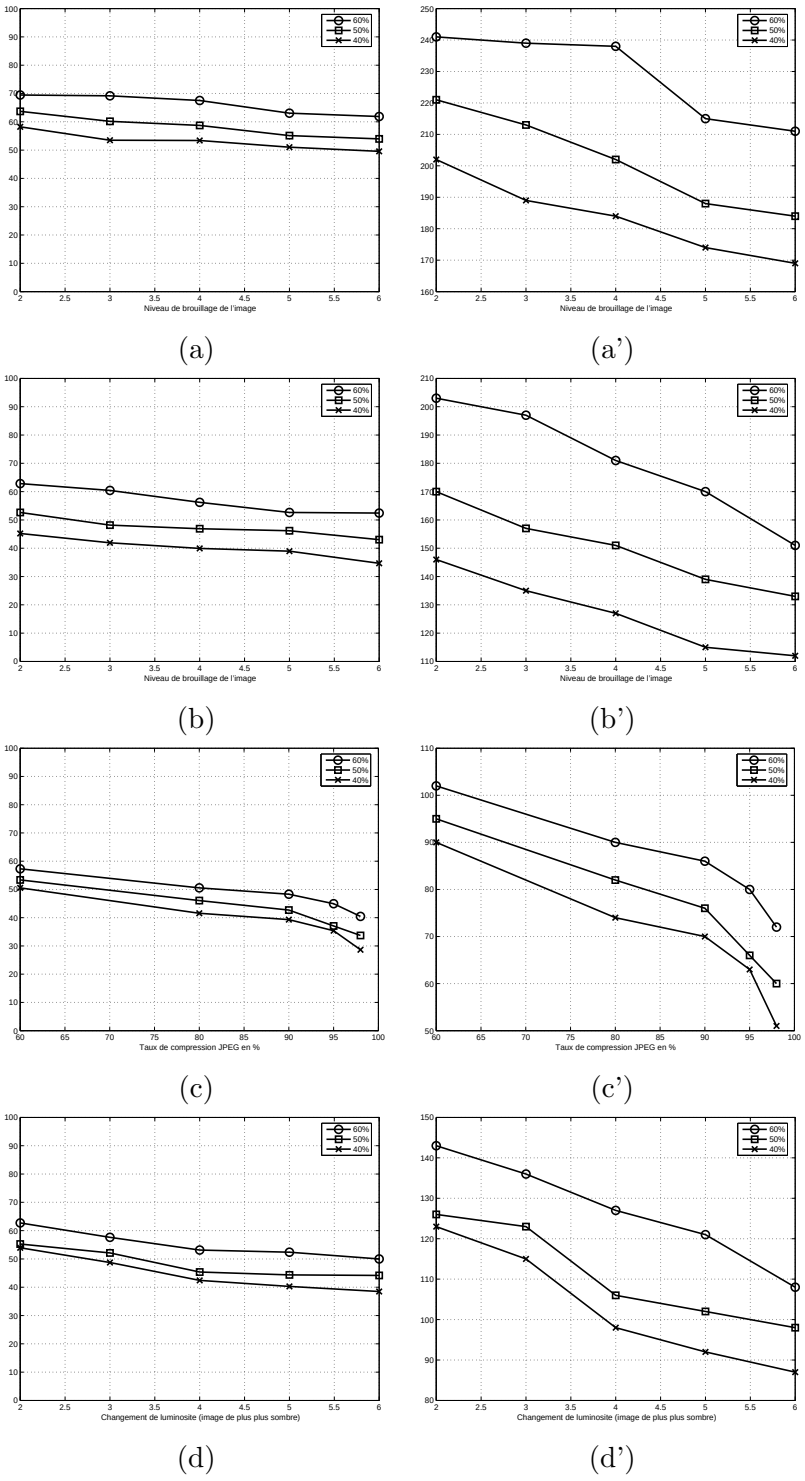


FIG. A.7 – Performance du détecteur ML (points d'intérêt : cercles $(x^*, y^*, \mathcal{S}^*)$). Score de répétabilité (a–d) et nombre de correspondances (a'–d') pour les séquences : (a) Blur1, (b) Blur2, (c) Compression, (d) Light.

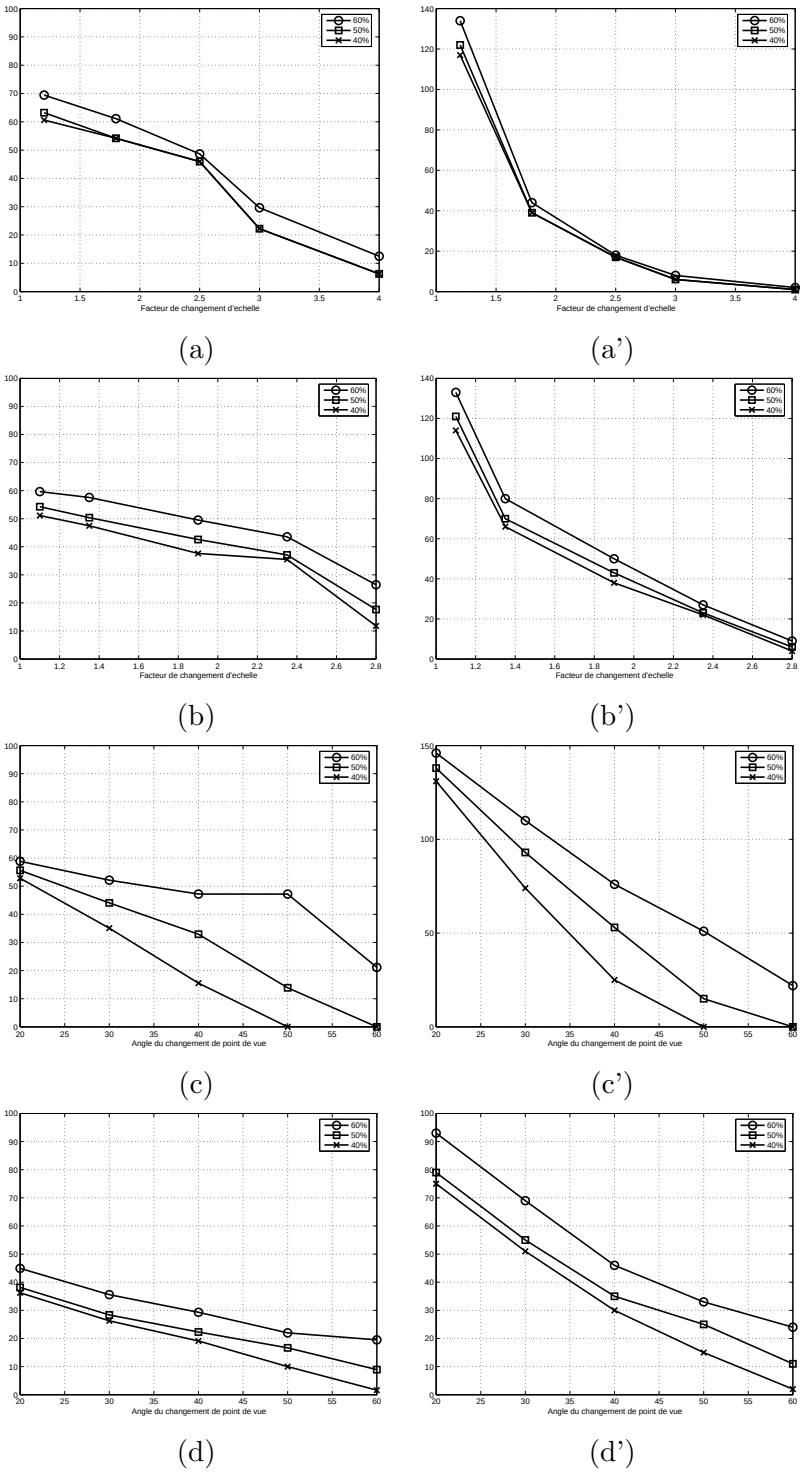


FIG. A.8 – Performance du détecteur ML (points d'intérêt : cercles $(x^*, y^*, \mathcal{S}^*)$). Score de répétabilité et nombre de correspondances pour les séquences : (a) ZoomRotation1, (b) ZoomRotation2, (c) Viewpoint1, (d) Viewpoint2.

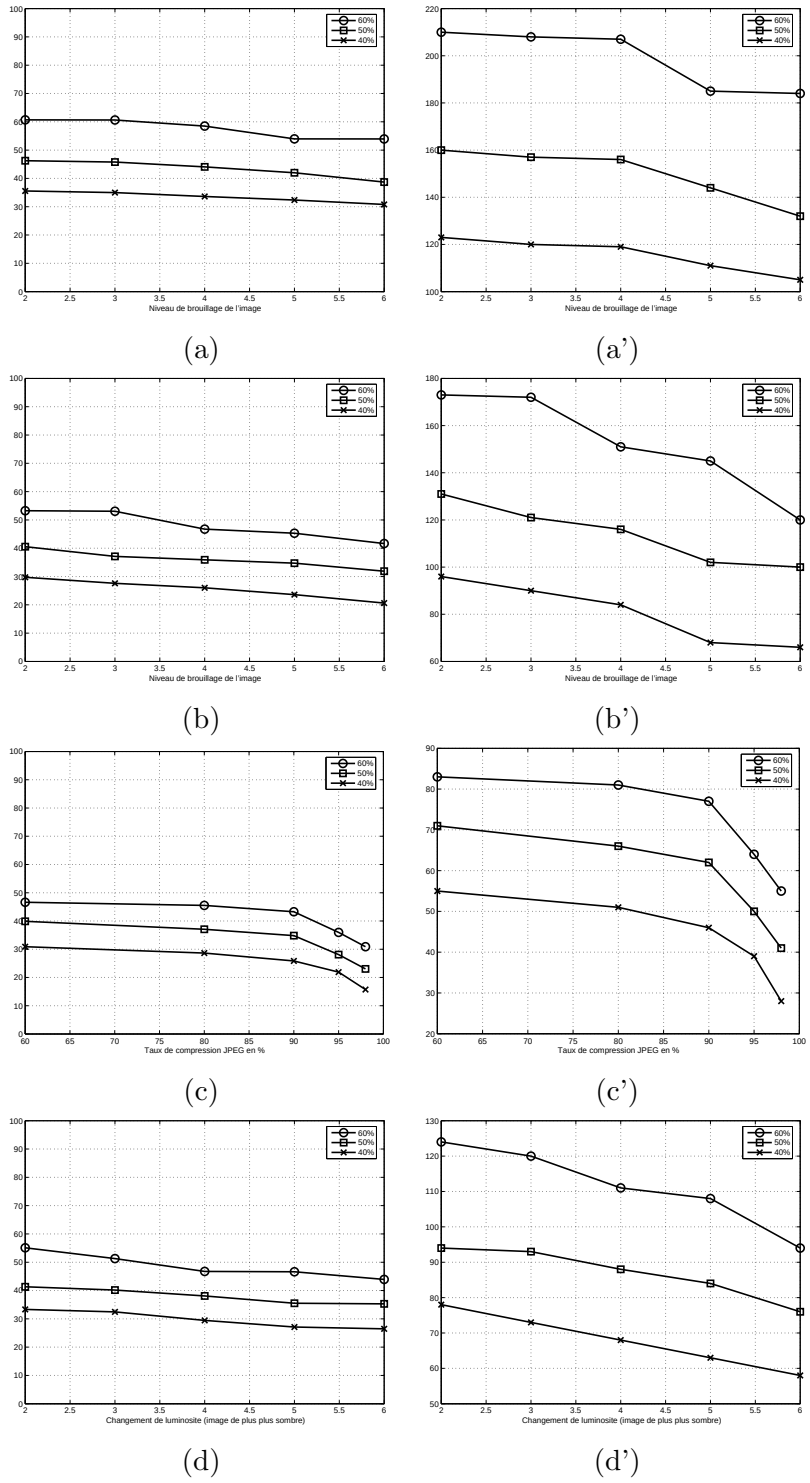


FIG. A.9 – Performance du détecteur ML, prenant en compte les déformations affines (points d'intérêt : ellipses (x_C, y_C, Σ_Q)). Score de répétabilité (a–d) et nombre de correspondances (a'–d') pour les séquences : (a) Blur1, (b) Blur2, (c) Compression, (d) Light.

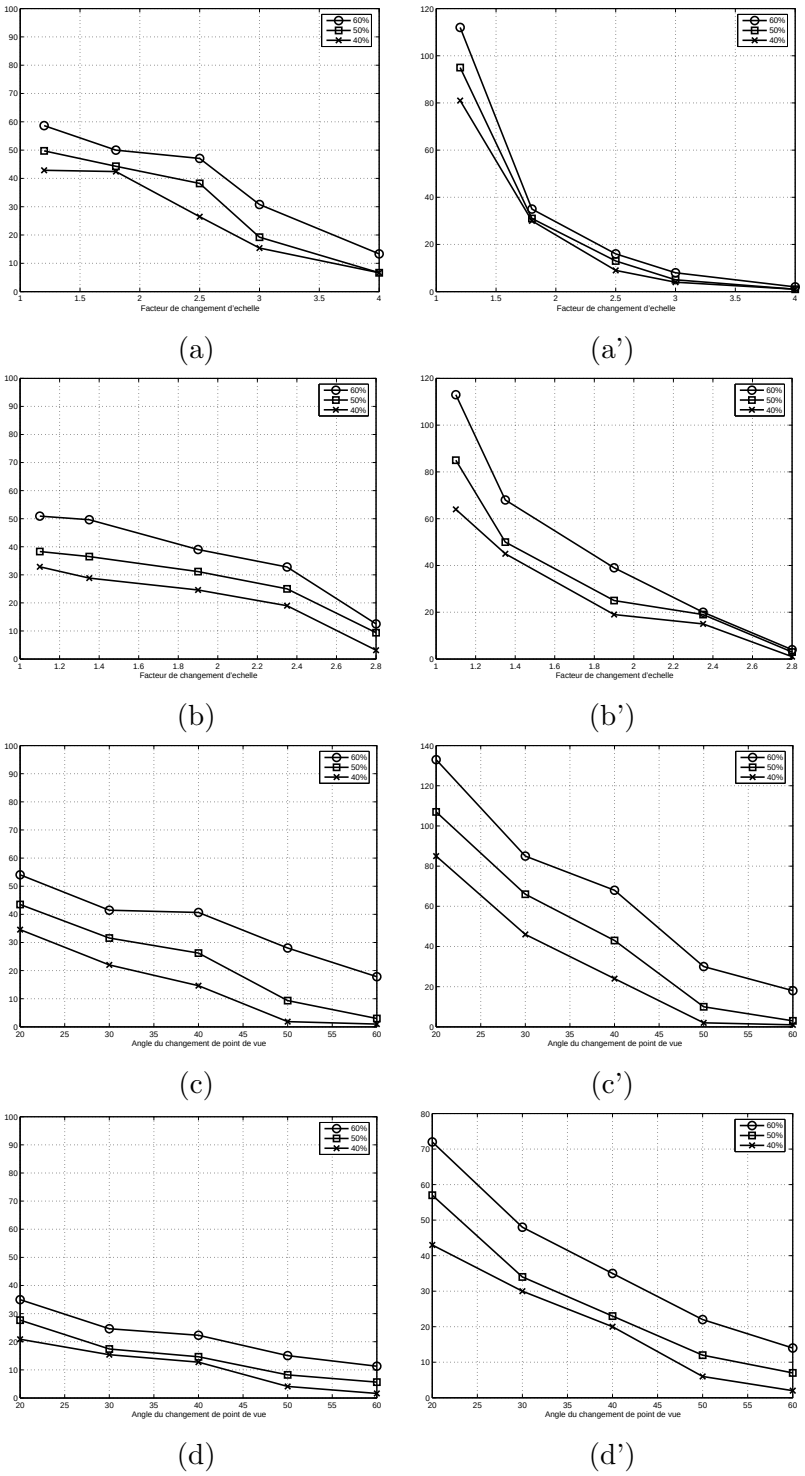


FIG. A.10 – Performance du détecteur ML, prenant en compte les déformations affines (points d'intérêt : ellipses (x_C, y_C, Σ_Q)). Score de répétabilité et nombre de correspondances pour les séquences : (a) ZoomRotation1, (b) ZoomRotation2, (c) Viewpoint1, (d) Viewpoint2.

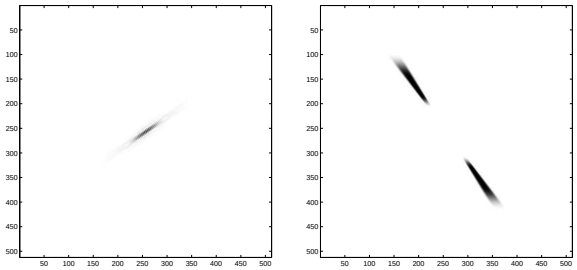
Annexe B

Estimation des régularités par curvelets

Nous décrivons ici une méthode d'estimation de la régularité Lipschitzienne, fondée sur les curvelets. Elle utilise en particulier leur lien avec les contours (cette méthode a été présentée au congrès Wave2006, Lausannne, 10-14 juillet 2006).

B.1 Contours et curvelets

Nous avons rappelé une règle due à Canny [26] permettant de déterminer si un point (pixel de l'image) appartient à un contour (chap. 3). En nous plaçant dans le voisinage d'un tel point, nous pouvons distinguer deux directions orthogonales : l'une normale au contour, qui présente le plus d'irrégularité (sauts et fluctuations associées au contour), et l'autre tangente au contour, présentant une certaine régularité.



(A) (B)

Fig. B.1 – Exemple d'une curvelet φ : (A) représentation dans le domaine spatial, $\varphi(x, y)$; (B) représentation dans le domaine fréquentiel, $\widehat{\varphi}(\omega_1, \omega_2)$.

Les singularités qui nous intéressent sont celles pour lesquelles les régularités sont les plus faibles (souvent négatives). Ainsi, il est pertinent de rechercher la direction de plus forte irrégularité, afin d'appliquer des résultats en 1D suivant cette direction. Dans cette optique, l'estimation de cette direction peut être obtenue en considérant la direction du gradient, celle-ci pouvant être estimée par des méthodes d'ondelettes [89, 88]. Ces méthodes supposent que la direction du gradient est stable lorsque l'échelle augmente, ce qui en pratique est vérifiée

approximativement (à condition que l'échelle ne soit pas trop petite).

L'approche proposée ici, basée sur les curvelets, autorise un certain degré de liberté dans la direction. Rappelons qu'une fonction *curvelet* à l'échelle s peut être vue comme une forme d'onde orientée dans une certaine direction θ, dont le pseudo-support est en forme d'aiguille (sa largeur w et sa longueur l obéissant à une relation $l \approx w^2$). Formellement, une telle fonction peut s'écrire, dans un système de coordonnées locales (x_1, x_2) comme

$$\varphi_s(x_1, x_2) = \frac{1}{s^{3/4}}\psi(\frac{x_1}{s})\phi(\frac{x_2}{\sqrt{s}}) \tag{B.1}$$

où ψ est une fonction oscillante et ϕ un passe-bas. La transformée FDCT – *Fast Discrete Curvelet Transform* – permet de calculer numériquement les coefficients de curvelets associés à une image. De sorte qu'un coefficient de curvelet peut être vu approximativement comme le produit scalaire

$$< f, \varphi_s > . \tag{B.2}$$

En supposant que la fonction f peut s'écrire (localement) comme $f(x_1, x_2) = f_1(x_1) \cdot f_2(x_2)$, nous avons

$$\begin{aligned} | < f, \varphi_s > | &= \left| \int_{\mathbb{R}^2} f(x_1, x_2)\varphi_s(x_1, x_2)dx_1dx_2 \right| \\ &= \left| \int_{\mathbb{R}^2} f_1(x_1)f_2(x_2)\frac{1}{s^{3/4}}\psi(\frac{x_1}{s})\phi(\frac{x_2}{\sqrt{s}})dx_1dx_2 \right| \\ &= \left| \left(\int_{\mathbb{R}} f_1(x_1)\frac{1}{s^{3/4}}\psi(\frac{x_1}{s})dx_1 \right) \left(\int_{\mathbb{R}} f_2(x_2)\phi(\frac{x_2}{\sqrt{s}})dx_2 \right) \right| \end{aligned}$$

Le premier terme correspond à une TOC W_pf_1 (avec une normalisation $L^p, p = 4/3$). Si de plus f_1 est de régularité α, alors nous avons (pour (u, s) dans un certain domaine) une inégalité du type :

$$|W_pf_1(u, s)| \leq As^{\alpha+\frac{1}{p}}.$$

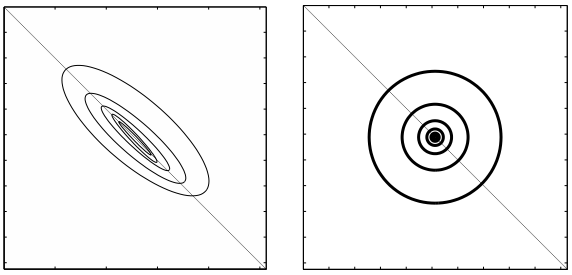
Par ailleurs, si f_2 est bornée, ϕ étant un passe-bas, le deuxième terme est également borné par une constante B. Ceci conduit alors à

$$| < f, \varphi_s > | \quad \leq \quad (As^{\alpha+\frac{3}{4}})B$$

et ainsi

$$| < f, \varphi_s > | \leq Cs^{\alpha+\frac{3}{4}}. \tag{B.3}$$

Les curvelets étant associées à des singularités curvilignes, il peut être intéressant de tenter de suivre ces singularités. Il n'existe pas à notre connaissance de théorie équivalente à celle des Maxima Lines qui permettrait de suivre des singularités curvilignes à travers les échelles. Nous présentons maintenant une méthode empirique permettant de suivre une singularité.



(A) (B)

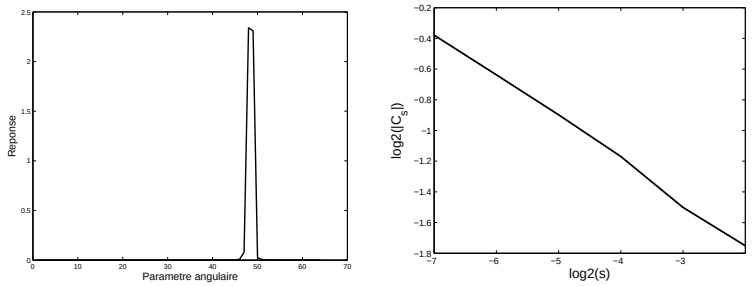
FIG. B.2 – Image 512×512 d'intensité f contenant une ligne diagonale ($f(x, y) = 1$ si (x, y) est sur la diagonale, 0 sinon). La régularité Lipschitzienne associée à cette singularité est $\alpha = -1$, et nous considérons le point de coordonnées $(x_0, y_0) = (256, 256)$. Sur (A), les ellipses représentent le pseudo-support de quelques curvelets centrées autour du point (x_0, y_0) ; ces curvelets sont celles qui donnent la réponse maximale à une échelle donnée suivant tous les angles possibles (à chaque échelle 2^j est associée un certain nombre d'angles). Sur (B), chaque cercle représente le pseudo-support d'une ondelette isotrope centrée en (x_0, y_0), dilatée à une certaine échelle 2^j.

B.2 Analyse de singularités par curvelets

Nous nous donnons un point (x, y) de l'image, préalablement identifié comme une singularité. Nous proposons d'analyser la régularité associée d'après la méthode suivante :

1. Trouver un certain nombre de curvelets localisées autour du point (x, y), i.e., à une certaine distance de ce point ; plus précisément, les curvelets n'étant pas à support compact, nous considérons toutes les curvelets centrées autour du point (x, y) ;
2. A chaque échelle s ($s = 2^j, j \in \mathbb{N}$), déterminer la curvelet la plus significative sur la base des coefficients de curvelets, i.e., le coefficient C_s maximal suivant différents angles (Fig. B.3 (B)). Ceci permet d'obtenir un ensemble de curvelets qui sont les plus significatives à chaque échelle (Fig. B.2 (A)) ;
3. Sur la base de la relation (empirique)
$$\log(|C_s|) \approx \left(\alpha + \frac{3}{4}\right) \log s + C \tag{B.4}$$
estimer α en utilisant une régression linéaire.

Notons que la première étape consiste à se donner un ensemble de curvelets qui pourraient correspondre aux mieux à la singularité présente (au sens de l'approximation des fonctions) : sont sélectionnées les curvelets dont le pseudo-support contient (x, y). Quant à elle, la deuxième étape sélectionne la curvelet la plus pertinente ; ceci se fait selon une direction appropriée (pouvant varier d'une échelle à une autre) qui capture au mieux la singularité. Enfin, nous constatons empiriquement que l'inégalité (B.4) (correspondant au logarithme de l'équation (B.3)) est presque une égalité aux échelles fines pourvu que les curvelets sélectionnées soient adaptées à la singularité.



(A) (B)

FIG. B.3 – (A) Evolution du module des coefficients de curvelet en fonction de l'angle, à une échelle fixée (la valeur du coefficient associé à la réponse maximale est noté C_s) ; (B) Evolution du module de C_s en fonction de l'échelle s ; une régression aux fines échelles donne une estimation $\hat{\alpha} = -1.01$, très proche de la valeur théorique $\alpha = -1$.

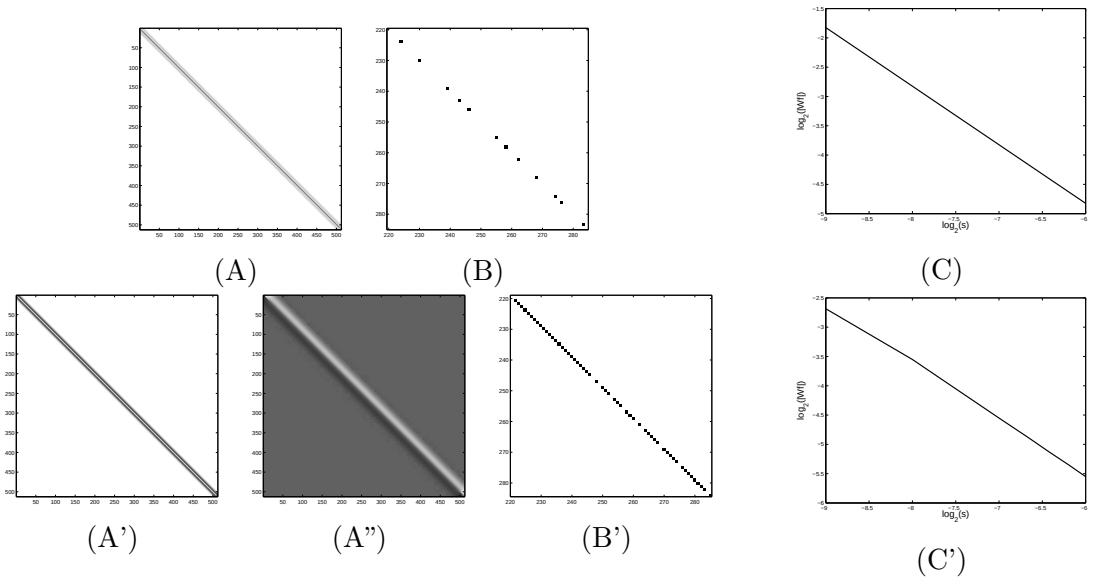


FIG. B.4 – Estimation de la régularités en $(x, y) = (256, 256)$, par des méthodes d'ondelettes ; (A) TOC utilisant l'ondelette Δg $(s = 3)$; (A') Module du gradient Mf $(s = 3)$; (A") Orientation du gradient Af $(s = 3)$; (B) Extrémités des ML $(s = 1)$; (B') Extrémités des LC $(s = 1)$; (C) Evolution de $|Wf|$ en fonction de s (le long d'une ML dont l'extrémité est (x, y)) ; (C') Evolution de Mf en fonction de s (le long d'une LC dont l'extrémité est (x, y)).

Cette approche par curvelets calcule la régularité α en un point (x, y), ce point pouvant être quelconque ou bien issu d'un détecteur particulier. Elle permet ainsi de caractériser des contours, par des fonctions qui sont adaptées aux contours [20, 21], alors que l'approche utilisant les LC se focalise sur des points appartenant aux contours de l'image. Relevons également que pour une image quelconque, suivant le point considéré, l'évolution de la réponse suivant l'angle est plus complexe que dans l'exemple étudié (Fig. B.3 (A)).

Enfin, à titre de comparaison, nous avons estimé la régularité au point $(x, y) = (256, 256)$ par des méthodes d'ondelettes (Fig. B.4). L'estimation par ML conduit à $\widehat{\alpha}_{ML} = -1.001$ (Fig. B.4 (C)) et celle par LC donne $\widehat{\alpha}_{LC} = -0.86$ (Fig. B.4 (C')), tandis que l'approche par curvelets conduit à $\widehat{\alpha} = -1.01$ (Fig. B.3 (B)).

Conclusion

L'approche par curvelets permet de fournir une régularité en un point quelconque de l'image, grâce à une sélection de fonctions adaptées aux contours. Ceci contraste avec l'approche de Canny, où les points de régularités sont fixés, et qui est basée sur l'orientation du gradient. Rappelons que dans [89], cette direction est supposée constante. Globalement, l'approche de Canny fournit une très bonne estimation dans le cas où les bords sont soit verticaux soit horizontaux, mais il existe un certain biais pour les bords orientés suivant des angles différents, comme nous l'avons vu au chap. 6. Ici, l'estimation peut être réalisée en un point quelconque, et surtout, les fonctions curvelets choisies peuvent être adaptées aux contours. Une perspective intéressante consisterait alors à effectuer une comparaison plus complète (sur une séquence d'images variées) de ces différentes méthodes d'estimation (ML, LC, curvelets).

Annexe C

Quelques aspects de l'EMD

Introduite par Huang et al. [55], la décomposition modale empirique – Empirical Mode Decomposition (EMD) – est une méthode adaptative de représentation de signaux non-stationnaires. Dans un premier temps, nous la présentons dans un cadre monodimensionnel. Nous verrons ensuite qu'elle peut être généralisée en 2D et détaillerons un algorithme rapide de décomposition EMD [33], lequel s'applique tant à des signaux bidimensionnels qu'à des images.

C.1 Présentation de la méthode EMD en dimension 1

Le propos de l'EMD est de décomposer un signal en une somme restreinte de composantes, vérifiant certaines propriétés : chaque composante doit être une fonction oscillante de moyenne locale nulle, modulée en amplitude et en fréquence. De par son caractère adaptatif, cette méthode s'applique dans des domaines tels que l'océanographie, la sismologie [55], la biologie (étude de la pression artérielle [54, 56]) et plus généralement à des signaux non-stationnaires.

C.1.1 Principe de la décomposition

L'idée fondamentale dans l'EMD est de considérer les extrema d'un signal monodimensionnel. Ces extrema permettent alors de déterminer ses oscillations locales. A titre d'exemple, supposons que t_1 et t_2 soient deux minima consécutifs d'un signal $x(t)$ (continu, non constant) ; alors il existe un maximum entre t_1 et t_2. Les variations du signal $x(t)$ entre ces deux minima peuvent être vues comme la somme d'une oscillation et d'une tendance (Fig. C.1).

Formellement, en notant $d(t)$ une composante localement haute-fréquence extraite de $x(t)$, et en posant $m(t) = x(t) - d(t)$ (tendance localement basse-fréquence), nous obtenons une décomposition de la forme :

$$\underbrace{x(t)}_{\text{signal}} = \underbrace{m(t)}_{\text{moyenne ou tendance}} + \underbrace{d(t)}_{\text{détails ou oscillations}} \tag{C.1}$$

Dans l'optique d'obtenir une décomposition du signal en plusieurs composantes, il est possible de décomposer à nouveau l'une ou l'autre des deux composantes précédemment obtenues.

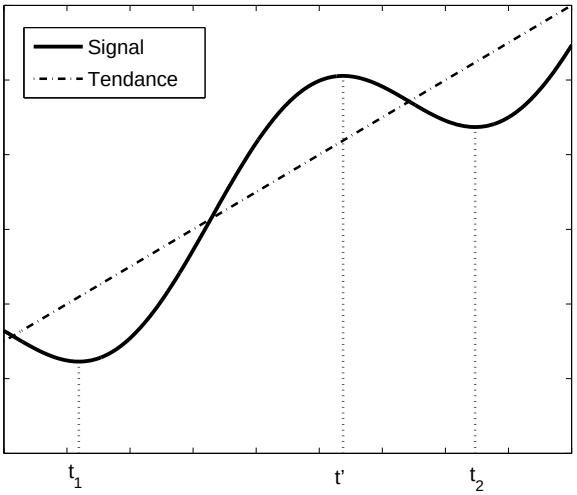


FIG. C.1 – Principe de l'EMD : décomposer un signal oscillant à partir de ses extrema, suivant une tendance et une oscillation.

En particulier, si $d(t)$ isole les plus hautes fréquences, en décomposant itérativement $m(t)$, la décomposition obtenue correspond à l'application d'un banc de filtres adaptatif. Ainsi, les grandes lignes de la méthode EMD sont les suivantes :

1. Identifier tous les extrema locaux de $x(t)$,
2. Interpoler les minima (resp. les maxima) de manière à construire une certaine enveloppe $EnvMin$ (resp. $EnvMax$),
3. Calculer la moyenne $m(t) = \frac{1}{2}(EnvMin(t) + EnvMax(t))$,
4. Extraire le détail $d(t) = x(t) - m(t)$,
5. Itérer sur le résidu $m(t)$ jusqu'à ce qu'il ne soit plus possible de construire d'enveloppe (nous verrons que le nombre de d'extrema tend à diminuer rapidement).

Ainsi, nous obtenons une décomposition du signal sous la forme :

$$x(t) = \sum_{k=1}^{K} d_k(t) + r(t) \ , K \in \mathbb{N}^* \tag{C.2}$$

Dans une telle décomposition, nous cherchons à ce que les composantes $d_k, (k \in \mathbb{N})$ obtenues soient des IMFs.

Définition 40. *(Intrinsic Mode Function, IMF) Une IMF est une fonction $d : \mathbb{R} \to \mathbb{R}$ oscillante, de moyenne locale nulle, modulée en amplitude et en fréquence, et dont les nombres de maxima et de zéros diffèrent au plus de 1 (ce qui signifie qu'une IMF passe par zéro entre un minimum et un maximum).*

La fonction $r(t)$ est appelé résidu de la décomposition. Elle représente la tendance globale du signal original $x(t)$ ($r(t)$ n'est pas nécessairement une quantité négligeable). Par abus de langage, le résidu pourra parfois être qualifié d'IMF, en tant que composante obtenue par l'algorithme EMD.

Notons que l'approche EMD contraste avec celle de la théorie de Fourier, ce qui est résumé dans le tableau suivant :

Fourier	EMD
$x(t) = \sum_{k=-\infty}^{+\infty} c_k e^{ikt}$	$x(t) = \sum_{k=1}^{K} d_k(t) + r(t)$
Grand nombre de composantes	Nombre restreint de composantes
Fonctions d'une base (e^{ikt})	Fonctions adaptées
Signaux stationnaires	Signaux non-stationnaires

C.1.2 Algorithme EMD

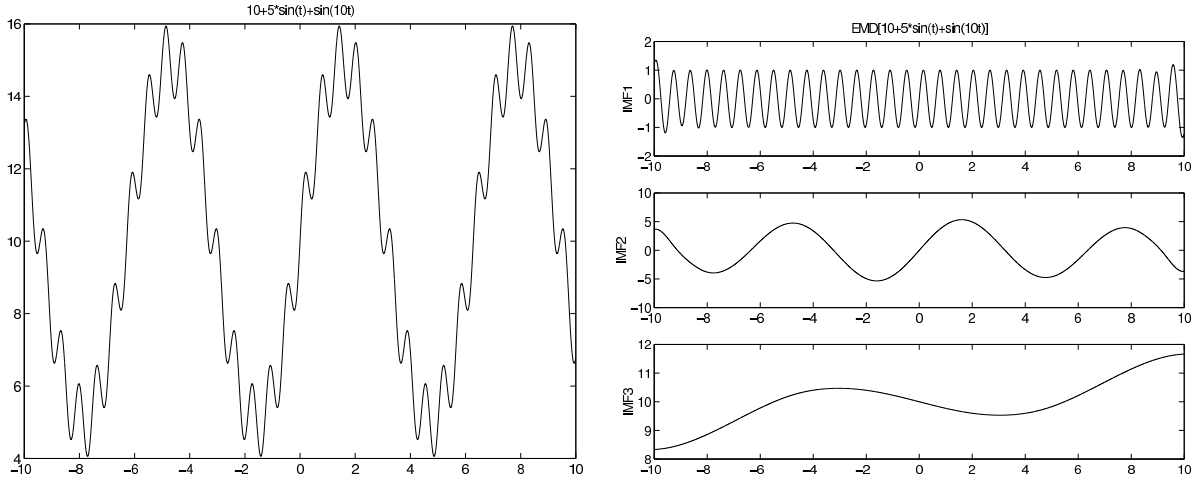
Afin d'obtenir une décomposition EMD (eq. (C.2)), l'algorithme suivant a été proposé [55] :

1. Initialisation : poser $r_0(t) = x(t)$, $k = 1$
2. Extraction de la k^{e}IMF, notée $d_k(t)$ [étape appelée *Sifting Process* (SP)]
   (a) Initialisation : $h_0(t) = r_{k-1}(t), j = 1$
   (b) Extraire les minima et les maxima locaux de $h_{j-1}(t)$
   (c) Interpoler les minima locaux (resp. les maxima) pour obtenir une enveloppe $\text{EnvMin}_{j-1}(t)$ (resp. $\text{EnvMax}_{j-1}(t)$)
   (d) Calculer la moyenne des deux enveloppes :
   $$m_{j-1}(t) = \frac{1}{2}(\text{EnvMin}_{j-1}(t) + \text{EnvMax}_{j-1}(t)) \quad (= \text{EnvMoy}_{j-1}(t))$$
   (e) $h_j(t) = h_{j-1}(t) - m_{j-1}(t)$
   (f) Evaluation du critère d'arrêt
   – s'il est satisfait : $d_k(t) = h_j(t)$
   – sinon retourner en (b) avec $j = j + 1$
3. $r_k(t) = r_{k-1}(t) - d_k(t)$
4. – si $r_k(t)$ a encore au moins deux extrema, retourner en (2) avec $k = k + 1$,
   – sinon la décomposition est terminée ; $r_k(t)$ constitue le résidu $r(t)$ de cette décomposition.

A titre d'illustration, nous appliquons cet algorithme à $x(t) = 10 + 5\sin(t) + \sin(10t)$ (Fig. C.2). Note : un programme implémentant cet algorithme, réalisé par P. Flandrin, est disponible sur `http ://perso.ens-lyon.fr/patrick.flandrin/emd.html`.

La particularité de cet algorithme EMD est qu'il permet de séparer automatiquement certains modes, au sens où les composantes obtenues correspondent à certaines bandes de fréquences contenues dans le signal [47]. Dans cet exemple, l'algorithme conduit à trois composantes d_1, d_2 et d_3, et nous constatons qu'elles correspondent approximativement aux modes présents dans $x(t)$: d_1 à $\sin(10t)$, $d_2(t)$ à $5\sin(t)$, tandis d_3 (le résidu) est proche de 10 (le caractère approximatif s'explique notamment par des défauts d'interpolation et des effets de bord). Ainsi, les modes présents dans le signal sont effectivement bien séparés, tant en amplitude qu'en fréquence.

Le résultat obtenu sur cet exemple simple montre le principe de l'EMD : décomposer automatiquement un signal oscillant (plus ou moins complexe) en différentes composantes, couvrant son spectre depuis les hautes vers les basses fréquences, et ce de manière adaptative (aucune gamme de fréquences particulière n'étant spécifiée initialement).



(A) (B)

FIG. C.2 – (A) Signal $x(t) = 10+5\sin(t)+sin(10t)$; (B) Décomposition EMD correspondante (2 composantes, 1 résidu).

C.1.3 Degrés de liberté dans la méthode

L'algorithme décrit ici nécessite quelques précisions, en particulier le type d'interpolation utilisée pour le calcul des enveloppes et le critère d'arrêt du *Sifting Process*.

Choix d'une méthode d'interpolation

Le type d'interpolation utilisée pour calculer les différentes enveloppes à partir des minima et des maxima est importante, dans la mesure où elle conditionne la forme des IMFs et certaines de leurs propriétés (par exemple, l'utilisation d'une interpolation linéaire par morceaux conduirait à une décomposition en un trop grand nombre de composantes). Dans la littérature,

l'interpolation par des splines cubiques est souvent utilisée, grâce au compromis réalisé entre la qualité de l'interpolation et la simplicité du calcul. Par ailleurs, retenons également une interpolation de type *pchip* (Piecewise Cubic Hermite Interpolation Polynomial), qui préserve la monotonie [49] : si aux nœuds d'interpolation sont associés des valeurs croissantes, la fonction polynômiale par morceaux interpolante sera croissante également ; ceci évite notamment la création d'oscillations artificielles dans les composantes obtenues.

Sifting Process

L'algorithme décrit précédemment comporte deux boucles imbriquées l'une dans l'autre ; celle indicée par j est appelée *Sifting Process* (SP). Ce SP correspond à un certain nombre d'itérations supplémentaires, effectués sur les détails extraits. L'intérêt du SP est qu'il permet de faire en sorte que les composantes obtenues puissent être considérées comme des IMFs (de manière satisfaisante). La terminaison du SP est liée à un certain critère d'arrêt, qui repose sur l'idée que $h_j(t)$ soit de moyenne locale nulle ; ceci peut être formulé par l'écart-type [55] ou un paramètre local plus sophistiqué [109]. Nous verrons également en 2D pourquoi il peut être pertinent d'effectuer un nombre fixé d'itérations [33].

Synthèse

L'EMD a fait l'objet de différentes études empiriques [46, 109]. Il apparaît que les composantes obtenues sont approximativement orthogonales, et surtout, l'étude du spectre de celles-ci montre que l'EMD se comporte comme un banc de filtres adaptatif [47], aucune gamme de fréquence n'ayant été spécifiée à l'avance. Ainsi, par ses différents aspects, l'EMD se distingue par rapport à d'autres méthodes de traitement du signal (analyse de Fourier, ondelettes). Relevons les particularités suivantes :

- méthode définie algorithmiquement,
- analyse locale du signal, adaptée au cadre non-stationnaire,
- décomposition adaptative en modes oscillants quelconques.

C.2 Problématique de l'EMD en 2D

Nous présentons maintenant des éléments sur l'EMD en 2D. En premier lieu, notons que formellement, l'algorithme EMD vu en C.1.2 s'applique également en 2D, à condition de préciser :

- la notion d'extrema ;
- le calcul des enveloppes (interpolation, gestion des effets de bord) ;
- le critère d'arrêt lié au *Sifting Process* (SP).

La BEMD (*Bidimensional* EMD) associée à cet algorithme généralisé conduit à écrire un signal bidimensionnel $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ sous la forme (cf. eq. (C.2))

$$f(x,y) = \sum_{k=1}^{K} d_k(x,y) + r(x,y) \ , K \in \mathbb{N}^* \tag{C.3}$$

Rappelons que la décomposition EMD dépend fortement des degrés de liberté de la méthode, et ceci est plus marqué dans le cadre bidimensionnel. Ainsi les différentes versions formulées conduisent à des résultats assez hétérogènes [80, 82, 104]. Dans un cadre discret, les extrema de f sont définis comme les points où f est supérieure à ses 8 voisins (dont au moins un au sens strict). Relevons que la localisation des extrema est *a priori* quelconque, de sorte que le problème de calcul des enveloppes est un problème d'interpolation de données dispersées (*scattered data*). Nous décrivons maintenant des méthodes permettant de calculer efficacement des enveloppes (interpolant des maxima ou des minima). Nous montrerons également qu'il peut être intéressant d'effectuer un nombre restreint d'itérations dans le SP.

C.3 Calcul des enveloppes

Interpolation de données dispersées

Rappelons que dans le cadre où des valeurs $(z_i)_{1\leq i\leq n}$ à interpoler sont associées à une grille régulière $(x_i, y_j)_{1\leq i,j\leq n}$, il est courant d'utiliser une interpolation en produit tensoriel. Ainsi, le problème de dimension deux se ramène à la dimension un, de sorte que le calcul correspondant peut être effectué rapidement. Dans le cadre de l'EMD, les extrema du signal bidimensionnel sont répartis de manière quelconque, et ainsi le calcul des enveloppes rentre dans le cadre de l'interpolation de données dispersées (*scattered data interpolation*). Ceci conduit au problème d'interpolation suivant :

$$(\mathcal{P}) \text{ Trouver } f : \mathbb{R}^2 \longrightarrow \mathbb{R} \text{ telle que } \forall k \in \{1...N\}, f(x_k, y_k) = z_k.$$

C.3.1 Utilisation de fonctions à base radiale

Les fonctions à base radiale constituent un outil intéressant pour le problème d'interpolation de données dispersées [13, 44, 91, 123]. Nous recherchons une fonction f sous la forme :

$$f(x,y) = \sum_{i=1}^{N} w_i \ \varphi\left(|(x,y) - (x_i, y_i)|_2\right) \tag{C.4}$$

où $(w_i)_{1\leq i\leq N}$ sont les poids (inconnues du problème), et φ est une fonction d'une variable appelée fonction à base radiale, *Radial Basis Function* (RBF). Différents choix sont possibles pour la fonction φ, comme par exemple $\varphi(r) = \exp(-cr^2)$, $\varphi(r) = \sqrt{r^2 + c^2}, c > 0$, ou celle que nous utilisons ici :

$$\varphi(r) = r^2 \ln(r) \qquad \textit{Thin-plate spline (TPS)} \tag{C.5}$$

qui est l'équivalent en dimension deux de la spline cubique. L'intérêt d'utiliser une telle RBF est qu'elle minimise une certaine énergie, définie comme

$$E(f) = \int_{R^2} \left(\partial_x^2 f\right)^2 + 2\left(\partial_x \partial_y f\right)^2 + \left(\partial_y^2 f\right)^2 \quad dxdy \tag{C.6}$$

Ce choix implique que la fonction f donnée par (C.4) est aussi aplatie que possible tout en respectant la contrainte d'interpolation. Alors la résolution du problème $(\mathcal{P})$ s'effectue par inversion d'un système linéaire, qui permet d'obtenir les poids $(w_i)_{1 \leq i \leq N}$. La solution f peut alors être représentée comme une surface bidimensionnelle $z = f(x, y)$ (Fig. C.3). Notons enfin que dans cette approche par RBF, il est possible de relaxer la contrainte d'interpolation, ce qui conduit à rechercher une fonction f minimisant :

$$(1 - \lambda) \sum_{i=1}^{N} \left(y_i - f(x_i)\right)^2 + \lambda E(f) \tag{C.7}$$

où $\lambda \in [0, 1]$ quantifie l'influence de ce terme d'énergie ($\lambda = 0$: régression aux moindres carrés, $\lambda = 1$: interpolation par TPS).

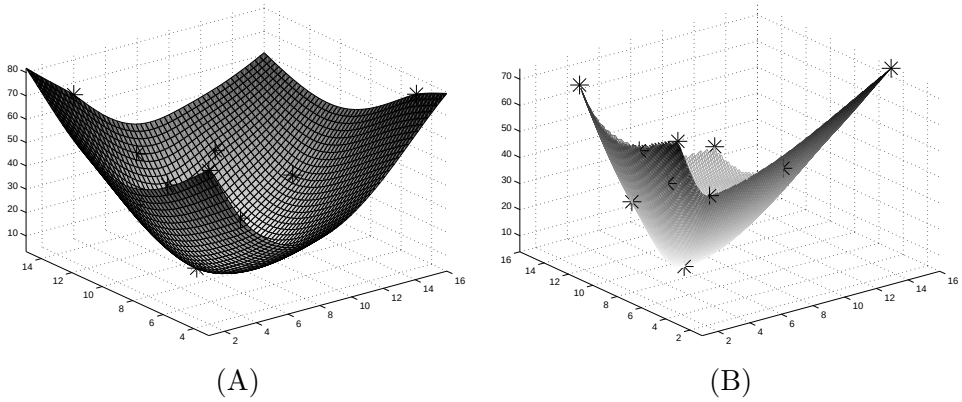


FIG. C.3 – Surfaces obtenues par des méthodes d'interpolation de données dispersées : (A) *Thin-Plate Spline*, (B) à partir d'une triangulation.

C.3.2 Interpolation basée sur une triangulation

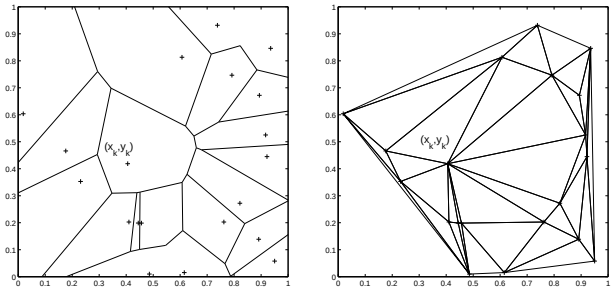
Afin de simplifier le problème d'interpolation, il est possible de le traiter en deux étapes :

1. Triangulation des points $(x_k, y_k)_{1 \leq k \leq N}$,
2. Interpolation sur chacun des triangles.

En ce qui concerne la triangulation, un choix usuel est celle de Delaunay, qui permet de décomposer un domaine en un ensemble de triangles. Rappelons que cette triangulation est unique, et que la construction de cette triangulation est réalisable par la construction de son dual, le diagramme de Voronoï (Fig. C.4). D'autres types de triangulations ont été proposées

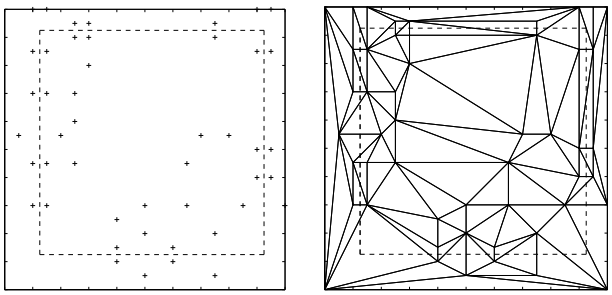
afin de prendre en compte la géométrie de l'image [37, 38, 51, 71], mais nous privilégions ici celle de Delaunay, laquelle peut être calculée d'après un algorithme rapide. Dans le cadre de l'EMD 2D, ceci est intéressant vu le nombre d'enveloppes à calculer.

En ce qui concerne la méthode d'interpolation, les fonctions s'appuyant sur les triangles précédents sont généralement simples. L'interpolation utilisée ici est similaire à un *pchip* en 1D : la surface obtenue est C^2 à l'intérieur des triangles, les raccords soient C^1 sur chaque arête de la triangulation (Fig. C.3).



(A) (B)

FIG. C.4 – (A) Diagramme de Voronoï ; (B) Triangulation de Delaunay.



(A) (B)

FIG. C.5 – (A) Symétrisation des points à interpoler ; (B) Triangulation de Delaunay correspondante (le tracé en pointillés correspond au domaine initial).

C.3.3 Conditions aux bords

Dans l'optique de définir des enveloppes satisfaisantes, il importe de gérer les conditions aux bords de l'image. Ceci est valable autant dans le cadre d'une interpolation par RBF que pour une surface s'appuyant basée sur une triangulation. Si aucune contrainte n'est imposée, cela conduit à des enveloppes explosant au bord. Inversement, si les contraintes sont trop fortes, cela crée artificiellement des structures au bord qui, du fait des propriétés de la méthode d'interpolation (raccords C^1 ou poids associés au RBF), se répercutent sur toute l'image. Ceci

est valable si, par exemple, nous imposons certaines valeurs prédéterminées aux quatre coins de l'image ou le long du bord. Une première option consiste à symétriser les données à interpoler, par rapport à chacun des bords et des coins : pour une image initiale de taille $n_1 \times n_2$, ceci conduit à un domaine de taille $9n_1 \times n_2$, d'où un surcoût algorithmique élevé. Une deuxième option que nous proposons consiste à ne symétriser qu'une partie du domaine (Fig. C.5).

En notant $\mathcal{D}_{extr}$ l'ensemble des points à interpoler et $\mathcal{B}$ le bord de l'image, nous définissons l'ensemble de points à symétriser comme

$$\mathcal{S}_{extr} = \left\{ A_i \in \mathcal{D}_{extr}, \min_{B_i \in \mathcal{B}} ||A_i - B_i||_2 \leq \sqrt{\frac{1}{|\mathcal{D}_{extr}|}} \right\} \tag{C.8}$$

puis nous considérons le domaine $\mathcal{D}_{extr} \cup T(\mathcal{S}_{extr})$, où T est l'opérateur de symétrie par rapport aux bord de l'image. Notons alors que l'enveloppe convexe de ce domaine (qui sert pour la triangulation de celui-ci) ne contient pas nécessairement le domaine de l'image $\mathcal{D}$. Afin d'obtenir cette condition, nous ajoutons les quatre coins du domaine étendu, (auxquels est associé la valeur du point le plus proche). Ainsi le domaine à interpoler est de taille $(n_1 + 2m) \times (n_2 + 2m)$, où m est une marge qui correspond à l'ajout d'une bande pour chaque bord de l'image. La largeur de cette bande est inversement proportionnelle à la densité de points à interpoler, celle-ci étant donnée par $\frac{N^2}{|\mathcal{D}_{extr}|}$.

Dans le cadre de l'EMD, ce traitement des conditions de bord se révèle adapté. En effet, la densité du nombre d'extrema diminue fortement au fur et à mesure du calcul des différentes IMFs ; ainsi la bande est étroite pour les premières IMFs et large pour les IMFs suivantes. De sorte que le surcoût lié à la gestion de problèmes de bord est faible pour les premières IMF, qui sont les plus coûteuses en temps de calcul. Globalement, cette technique permet de gérer les conditions aux bords tout en nécessitant peu de calculs supplémentaires.

C.3.4 Calculs d'enveloppes

Nous considérons un bruit blanc Gaussien bidimensionnel ($\sigma = 1$), auquel nous appliquons une itération de l'EMD (calcul de $EnvMin$ et $EnvMax$). Ceci permet de comparer, suivant la méthode utilisée, l'enveloppe moyenne $m(x, y) = \frac{1}{2}(EnvMin + EnvMax)$ ainsi que les détails $d(x, y) = f(x, y) - m(x, y)$ (Fig. C.6). Dans l'EMD, l'extraction de la première IMF est fondamentale car elle conditionne les IMFs suivantes. En particulier, il importe que l'enveloppe moyenne (la tendance) présente encore des oscillations, ce qui est vérifié pour l'approche basée sur une triangulation (Fig. C.6 (A')) mais qui n'est pas très marqué dans le cas d'utilisation de TPS. En outre, pour le même type de signal, nous comparons les deux méthodes en termes de coût algorithmique (Tab. C.3.4). Il apparaît que celle utilisant des TPS est bien plus coûteuse que celle basée sur une triangulation, en particulier pour des signaux présentant de nombreux extrema, comme c'est le cas de bruits 2D. Ainsi l'approche basée sur une triangulation, proposée dans [33], apparaît comme plus pertinente que d'autres formulations utilisant des TPS [81].

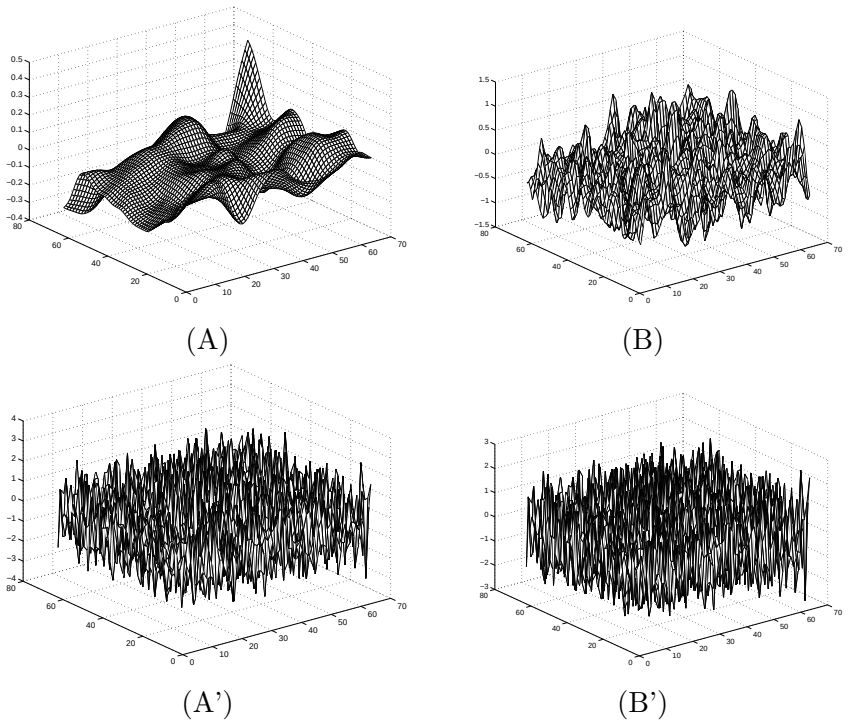


FIG. C.6 – (A-B) Enveloppes moyennes ; (A'-B') Détails extraits ; (A–A') par *Thin-Plate Spline* ; (B–B') à partir d'une triangulation.

TAB. C.1 – Coût algorithmique de l'EMD en 2D, en fonction de n (les données étant de taille $n \times n$) suivant que la méthode d'interpolation utilisée (PC : Piecewise Cubic, basée sur une triangulation, TPS : Thin-Plate spline) et suivant le nombre d'itérations (3 et 11) au cours du SP.

	3 itérations		11 itérations	
n	PC	TPS	PC	TPS
16	1	2	3	7
32	2	6	5	22
64	4	57	14	1675
128	10	X	44	X
256	43	X	187	X
512	185	X	871	X

C.4 Raffinement des enveloppes – *Sifting Process* (SP)

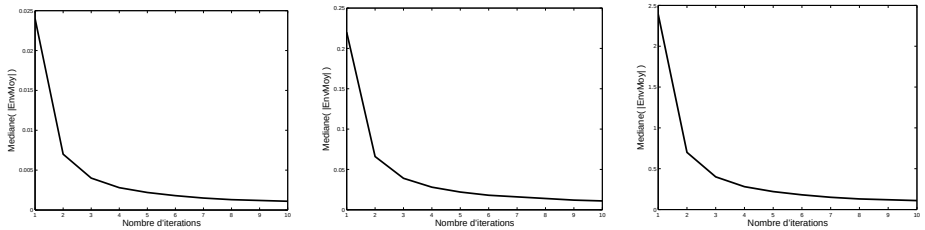
Rappelons qu'une fois que la première décomposition du signal f a été calculée (sous la forme $f = d + m$), le SP consiste à itérer sur les détails d. L'intérêt de cette opération est que la composante haute-fréquence corresponde à des fréquences suffisamment élevées. Ainsi, le SP permet de raffiner les enveloppes, au sens où plus il y a d'itérations, plus les détails extraits seront concentrés sur les hautes fréquences (étant donné que nous ôtons des composantes basse-fréquence). Il importe de noter que cette opération est déterminée par le nombre d'itérations effectuées. Ainsi le critère d'arrêt détermine ce nombre d'itérations à effectuer, afin que globalement (et dans une certaine mesure), les composantes extraites par l'algorithme EMD soient des IMFs.

Différents critères d'arrêt ont été proposés en 1D [55, 109] et leur étude a montré qu'un trop grand nombre d'itérations conduisait souvent à sur-décomposer le signal et à créer des artefacts liés plus à la méthode d'interpolation utilisée qu'au signal analysé. Par ailleurs, relevons que cette opération est particulièrement déterminante pour le calcul de la première IMF. L'approche proposée ici recherche empiriquement un nombre fixé d'itérations à effectuer pour que les composantes obtenues soient satisfaisantes. A cet effet, nous étudions l'impact du nombre d'itérations effectué dans le SP, sur l'enveloppe moyenne correspondant à la 1^{re}IMF. Plus précisément, nous considérons 100 réalisations d'un bruit blanc Gaussien bidimensionnel, sur lesquelles nous appliquons l'algorithme EMD (tel que défini précédemment), en variant le nombre d'itérations correspondant au SP. En notant N_r le nombre de réalisations (ici $N_r = 100$), et pour j itérations effectuées, nous définissons

$$g(j) = \frac{1}{N_r} \sum_{i=1}^{N_r} \text{mediane}(|m_j^i(.,.)|)$$

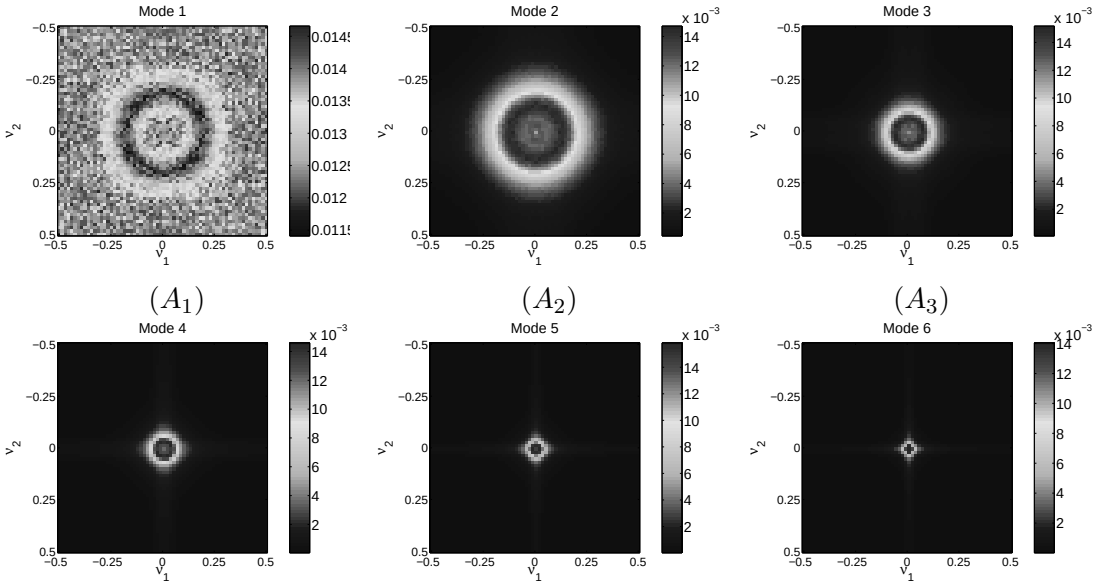
ce qui donne l'évolution de la médiane de $|EnvMoy|$ en fonction du nombre d'itérations j, en moyenne, sur les 100 réalisations effectuées (Fig. C.7 (A_i)).

Nous observons que cette grandeur diminue rapidement avec le nombre d'itérations effectuées. De plus, la courbure de la courbe $(j, g(j))$ admet un minimum pour $j = 3$. Ainsi, dans le cadre d'un bruit blanc bidimensionnel, il est pertinent d'effectuer uniquement 3 itérations pour calculer la première IMF (alors que d'autres études suggéraient d'utiliser bien plus d'itérations [55]). Nous obtenons des résultats similaires pour une image naturelle, tant dans le cas bruité que non-bruité. Enfin, étant donné que le nombre d'extrema diminue très rapidement au cours de l'extraction successive des IMFs, les procédures de SP correspondantes ne nécessiteront pas plus de 3 itérations.



(A_0) (A_1) (A_2)

FIG. C.7 – Evolution de la médiane de $|EnvMoy(x, y)|$, en fonction du nombre d'itérations au cours du SP (évolution en moyenne sur 100 réalisations d'un bruit blanc Gaussien 64×64) : $(A_1)\sigma = 0.1, (A_2)\sigma = 1, (A_3)\sigma = 10$.



(A_1) (A_2) (A_3)

(A_4) (A_5) (A_6)

FIG. C.8 – Module de la transformée de Fourier des IMFs, calculé en moyenne à partir d'un échantillon composé de 1000 réalisations de bruit blanc Gaussien bidimensionnel 64×64 de variance $\sigma^2 = 1$.

C.5 Analyse de signaux bidimensionnels par EMD

L'algorithme EMD que nous proposons utilise le même principe que l'algorithme original de Huang et al. [55]. Il construit des enveloppes régulières basées sur une triangulation de Delaunay des extrema, et utilise 3 itérations pour le SP [33].

D'abord, nous analysons un bruit blanc Gaussien bidimensionnel (en moyenne, sur 1000 réalisations de taille 64×64, $\sigma = 1$), et représentons le module de la transformée de Fourier de chaque IMF obtenue (Fig. C.8). Tandis que la 1^{re}IMF se concentre sur les hautes fréquences, les spectres des IMFs suivantes correspondent à des bandes de fréquences relativement séparées, des plus hautes vers les plus basses. Ainsi, cela a été vu en 1D [47], l'EMD se comporte comme en banc de filtres auto-adaptatif.

Ensuite, nous effectuons des décompositions EMD sur un bruit blanc Gaussien et sur une image naturelle, dont l'intensité f (niveau de gris) est un signal 2D. Pour le bruit (Fig. C.9), les composantes obtenues s´avèrent de moins en moins oscillantes, ce qui est conforme aux résultats précédents. Quant à l'image naturelle (Fig. C.10), les composantes obtenues correspondent à des détails associés à des échelles de plus en plus grossières. Sur ces exemples, il apparaît que la première IMF contient beaucoup plus d'information que les suivantes. Ceci constitue un problème lié à la fois au caractère adaptatif de l'EMD (nombre limité de composantes) et au fait que l'extraction s'effectue des plus hautes fréquences vers les plus basses.

Enfin, pour un bruit blanc Gaussien 512×512, pour l'image *Lena* 512×512 et pour la somme des deux, nous présentons des données numériques associées aux différentes IMFs extraites (Tab. C.6). En ce qui concerne le bruit, les composantes obtenues respectent bien la propriété d'IMF, que ce soit la moyenne nulle ou le caractère oscillant autour de zéro (quasiment tous les maxima sont positifs, et presque tous les minima sont négatifs). En ce qui concerne l'image *Lena* (bruitée ou non), les conditions précédentes sont moins respectées, ce qui s'explique par la complexité de la structure des images naturelles.

C.6 Conclusions et perspectives

L'intérêt principal de la méthode EMD est son caractère adaptatif, relié à sa formulation algorithmique utilisant les extrema du signal. Initialement formulée dans un cadre 1D, elle s'étend en 2D, tout en présentant de nettes différences. En particulier, nous avons vu les problèmes liés à l'interpolation de données dispersées et au SP. De par ce caractère algorithmique, l'étude théorique de cette méthode soulève des difficultés, même si les propriétés des IMFs sont mieux cernées [114]. Dans le cadre monodimensionnel, de nouvelles approches ont été proposées afin de reformuler l'algorithme, l'enveloppe moyenne étant directement calculée à partir des extrema [36, 90] (ce qui évite le calcul de deux enveloppes interpolant resp. les maxima et les minima). L'extension de ces approches au cadre bidimensionnel est en construction.

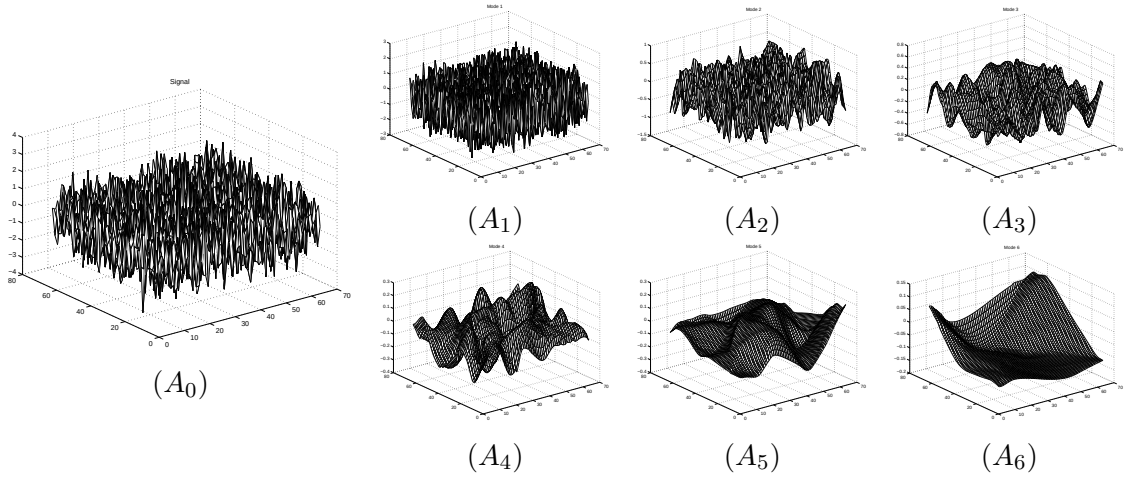


FIG. C.9 – Décomposition EMD sur un bruit blanc Gaussien bidimensionnel : (A_0) Signal analysé ; $(A_i)_{1\leq i\leq 6}$ 6 premières IMFs correspondantes.

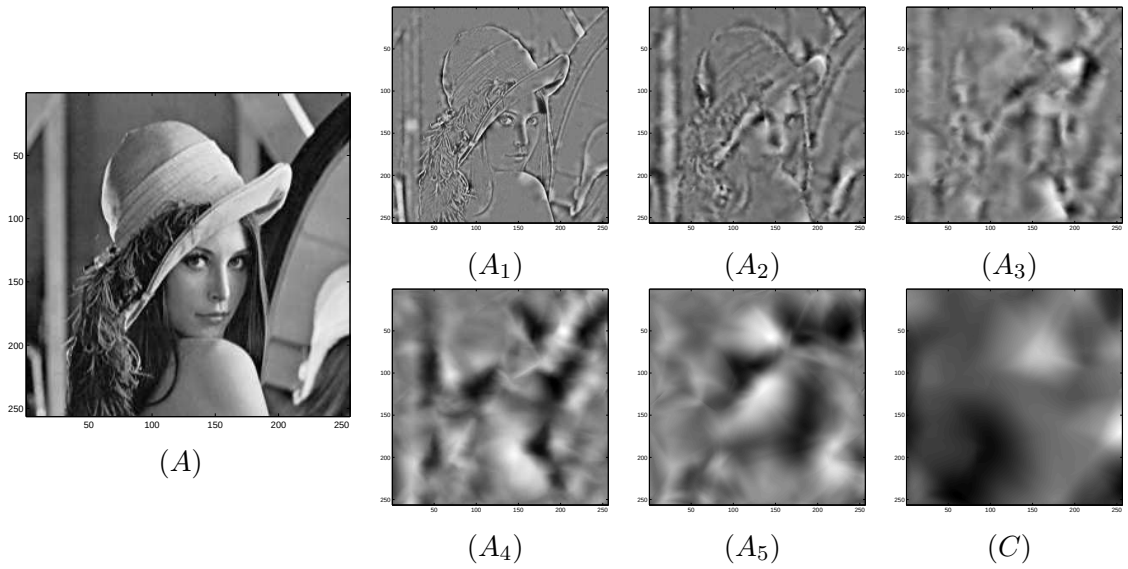


FIG. C.10 – Décomposition EMD sur l'image *Lena* : (A) Image originale ; (B_i) IMFs obtenues ; (C) Résidu.

TAB. C.2 – Exemples numériques de la décomposition EMD sur des images de taille 512×512 en utilisant 3 itérations par mode : un bruit blanc Gaussien bidimensionnel, l'image classique *Lena*, et la somme des deux ($SNR = 20dB$). Pour chaque IMF (et pour le résidu), nous donnons les nombres de maxima, de minima, de maxima positifs et de minima négatifs, ainsi que la moyenne et l'écart-type de l'IMF.

		nb max	nb min	nb max> 0	nb min < 0	Moyenne	Ecart-type
IMF_1	Bruit	30645	30692	30645	30692	9.10^{-4}	0.98
	Lena	21776	21294	21776	21296	-4.10^{-2}	11
	Lena+Bruit	22846	22357	22846	22357	-0.07	10.59
IMF_2	Noise	8889	8910	8889	8910	-6.10^{-4}	0.36
	Lena	4236	4206	4236	4206	0.28	11.6
	Lena+Bruit	4529	4600	4529	4600	0.18	11.4
IMF_3	Bruit	3153	3198	3153	3198	-3.10^{-4}	0.20
	Lena	1420	1424	1420	1424	-0.26	12.2
	Lena+Bruit	1533	1517	1533	1517	0.27	12.1
IMF_4	Bruit	1486	1480	1486	1480	10^{-3}	0.12
	Lena	928	980	928	980	0.25	11.4
	Lena+Bruit	961	931	961	931	0.22	11.9
IMF_5	Bruit	836	896	836	896	-2.10^{-5}	0.07
	Lena	873	812	873	812	0.35	10.3
	Lena+Bruit	741	765	741	765	0.55	11.4
IMF_6	Bruit	656	652	656	652	10^{-4}	0.05
	Lena	676	712	676	712	-1.6	12
	Lena+Bruit	706	680	706	680	1.02	10.4
IMF_7	Bruit	576	562	576	562	5.10^{-3}	0.03
	Lena	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$
	Lena+Bruit	706	718	706	718	-0.6	11.7
Résidu	Bruit	177	211	$\times$	$\times$	$\times$	$\times$
	Lena	169	161	$\times$	$\times$	$\times$	$\times$
	Lena+Bruit	119	131	$\times$	$\times$	$\times$	$\times$

Bibliographie

[1] E. Adelson and W. Freeman. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(9) :891–906, 1991.

[2] A. Almansa and T. Lindeberg. *Enhancement of fingerprint images using shape-adaptation of scale-space operators*, volume 8, pages 21–29. Kluwer, 1997.

[3] J. P. Antoine. The continuous wavelet transform in image processing. *CWI Quarterly*, 11(4) :323–345, 1998.

[4] A. Arneodo, E. Bacry, S. Jaffard, and J. F. Muzy. Singularity spectrum of multifractal functions involving oscillating singularities. *J. Fourier Anal. Appl.*, 4(2) :159–174, 1998.

[5] A. Arneodo, E. Bacry, S. Jaffard, and J. F. Muzy. Oscillating singularities and fractal functions. In *Spline functions and the theory of wavelets (Montreal, PQ, 1996)*, volume 18 of *CRM Proc. Lecture Notes*, pages 315–329. Amer. Math. Soc., Providence, RI, 1999.

[6] J. F. Aujol and A. Chambolle. Dual norms and image decomposition models. *International Journal of Computer Vision*, 63(1) :85–104, 2005.

[7] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda. Uniqueness of the gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1) :26–33, 1986.

[8] A. Baumberg. Reliable feature matching across widely separated views. In *Proceedings of CVPR*, pages 774–781, 2000.

[9] S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2(4) :509–522, 2002.

[10] Z. Berman and J. Baras. Properties of multiscale maxima and zero-crossing representations. *IEEE Transactions on Image Processing*, 41(12) :3216–3231, 1993.

[11] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms. *Communications on Pure and Applied Mathematics*, XLIV :141–183, 1991.

[12] J. Bigot. A scale-space approach with wavelets to singularity estimation. *ESAIM*, 9 :143–164, 2005.

[13] F. L. Bookstein. Principal warps : Thin plate splines and the decomposition of deformations. *IEEE Transactions on PAMI*, 11 :567–585, 1989.

[14] H. Brezis. *Analyse fonctionnelle, théorie et applications.* Paris, Masson, 1983.

[15] V. Bruni and D. Vitulano. Wavelet based signal denoising via simple singularities approximation. *Signal Processing*, 86 :859–876, 2006.

[16] P. J. Burt and E. H. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 4 :532–540, 1983.

[17] G. J. Burton and I. R. Moorhead. Color and spatial structure in natural scenes. *Journal of Applied Optics*, 26 :157–170, 1987.

[18] E. Candes and L. Demanet. The curvelet representation of wave propagators is optimally sparse. *Comm. Pure Appl. Math*, 58 :1472–1528, 2004.

[19] E. Candes, L. Demanet, D. L. Donoho, and L. Ying. Fast discrete curvelet transforms. *Multiscale Model. Simul*, 5 :861–899, 2005.

[20] E. Candes and D. L. Donoho. Curvelets and curvilinear integrals. *J. Approx. Theory*, 113 :59–90, 2000.

[21] E. Candes and D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise-c2 singularities. *Comm. Pure Appl. Math*, 57 :219–266, 2000.

[22] E. Candes and D. L. Donoho. Continuous curvelet transform : I. resolution of the wavefront set. *Appl. Comput. Harmon. Anal.*, 19 :198–222, 2002.

[23] E. Candes, J. Romberg, and T. Tao. Robust uncertainty principles : Exact recovery from highly incomplete fourier information. *IEEE Transactions on Information Theory*, 2006.

[24] E. Candes and T. Tao. Near-optimal signal recovery from random projections and universal encoding strategies. *IEEE Transactions on Information Theory*, 2004.

[25] E. J. Candes and D. L. Donoho. Continuous curvelet transform : Ii. discretization and frames. *Appl. Comput. Harmon. Anal.*, 19 :198–222, 2000.

[26] J. Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence*, 8(6) :679–698, 1986.

[27] R. Collins. Mean-shift blob tracking through scale space. In *Computer Vision and Pattern Recognition (CVPR'03)*. IEEE, 2003.

[28] Y. Cui, J. Weng, and H. Reynolds. Estimation of ellipse parameters using optimal minimum variance estimator. *Pattern recognition letters*, 17(3) :309–316, 1996.

[29] C. Damerval. Notes sur la régularité lispchitzienne. Technical Report, HAL Inria, no. 188421, 2007.

[30] C. Damerval and S. Meignen. Computation of blob-like structures characteristic scale with wavelet maxima lines. In T. Lyche P. Chenin, J. L. Merrien, editor, *Proceedings of the 6th Conference on Curves and Surfaces*, pages 111–120. Nashboro Press, June 29 - July 5 2006.

[31] C. Damerval and S. Meignen. Blob detection with wavelet maxima lines. *IEEE Signal Processing Letters*, 14(1) :39–42, 2007.

[32] C. Damerval and S. Meignen. Interest point detection with wavelet maxima lines. Technical Report, HAL Inria, no. 171678, 2007.

[33] C. Damerval, S. Meignen, and V. Perrier. A fast algorithm for bidimensional emd. *IEEE Signal Processing Letters*, 12(10) :701–704, 2005.

[34] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Transactions on Information Theory*, 36 :961–1005, 1990.

[35] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, 1992.

[36] E. Deléchelle, J. Lemoine, and O. Niang. Empirical mode decomposition : An analytical approach for sifting process. *IEEE Signal Processing Letters*, 12(11) :764–767, 2005.

[37] L. Demaret, N. Dyn, and A. Iske. Image compression by linear splines over adaptive triangulations. *IEEE Trans. on Signal Processing*, 86(7) :1604–1616, 2006.

[38] L. Demaret and A. Iske. Adaptive image approximation by linear splines over locally optimal delaunay triangulations. *IEEE Signal Processing Letters*, 13(5) :281–284, 2006.

[39] R. Deriche. Using canny's criteria to derive an optimal edge detector recursively implemented. *International Journal on Computer Vision*, 2 :15–20, 1987.

[40] M. N. Do and M. Vetterli. The contourlet transform : an efficient directional multiresolution image representation. *IEEE Transactions Image on Processing*, 14(12) :2091–2106, 2005.

[41] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard. Density estimation by wavelet thresholding. *Annals of Statistics*, 24(2) :508–539, 1996.

[42] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies. Data compression and harmonic analysis. *IEEE Transactions on Information Theory*, 44(6) :2435–2476, 1998. Information theory : 1948–1998.

[43] G. Dorko and C. Schmid. Selection of scale-invariant parts for object class recognition. In *Proceedings of the 9th ICCV*, pages 634–640, 2003.

[44] J. Duchon. Splines minimizing rotation invariant seminorms. *Sobolev spaces, constructive theory of functions of several variables*, 1 :85–100, 1976.

[45] V. Ferrari, L. Fevrier, F. Jurie, and C. Schmid. Groups of adjacent contour segments for object detection, 2006. Technical report, INRIA, no 5980.

[46] P. Flandrin and P. Goncalves. Empirical mode decompositions as a data-driven wavelet-like expansions. *International Journal of Wavelets, Multiresolution and Information Processing*, 2(4) :477–496, 2004.

[47] P. Flandrin, P. Goncalves, and G. Rilling. Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters*, 11(2) :112–114, 2004.

[48] L. Florack. *Image Structure.* Kluwer, 1997.

[49] F.N. Fritsch and R. E. Carlson. Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17 :238–246, 1980.

[50] C. Gasquet and P. Witomski. *Analyse de Fourier et applications. Filtrage, calcul numérique et ondelettes.* Dunod, 2000.

[51] L. Guibas and D. Russel. An empirical comparison of techniques for updating delaunay triangulations. *ACM Symposium on Computational Geometry*, pages 170–179, 2004.

[52] G. H. Hardy, J. E. Littlewood, and G. Pólya. *A Theorem of W. H. Young.* Cambridge University Press, 1988.

[53] C. Harris and M. J. Stephens. A combined corner and edge detector. In *Alvey Vision Conference*, pages 147–152, 1988.

[54] N. E. Huang and Z. Wu. A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society of London*, 460 :1597–1611, 2004.

[55] N.E. Huang and al. The empirical mode decomposition and the hilbert spectrum for non-linear and non-stationary time series analysis. *Proceedings of the Royal Society A*, 454 :903–995, 1998.

[56] N.E. Huang and al. A confidence limit for the empirical mode decomposition and the hilbert spectral analysis. *Proceedings of the Royal Society of London*, 459 :2317–2345, 2003.

[57] S. Jaffard and Y. Meyer. Wavelet methods for pointwise regularity and local oscillations of functions. *American Mathematical Society*, 1996.

[58] S. Jaffard and Y. Meyer. On the pointwise regularity of functions in critical Besov spaces. *J. Funct. Anal.*, 175(2) :415–434, 2000.

[59] L. M. Kaplan and R. Murenzi. Texture segmentation using multiscale hurst features. In *ICIP97*, pages 205–208, 1997.

[60] T. Kato. *Perturbation Theory for Linear Operators.* Springer, New York, 1966.

[61] C. J. Kicey and C. J. Lennard. Unique reconstruction of band-limited signals by a mallat-zhong wavelet transform algorithm. *Journal of Fourier Analysis and Applications*, 3(1) :63–82, 1997.

[62] R. Kimmel. *The Numerical Geometry of Images.* Springer, New York, 2004.

[63] J. Koenderink. The structure of images. *Biological Cybernetics*, 50 :363–370, 1984.

[64] J. Koenderink. *Solid Shape.* Artificial Intelligence Series. MIT Press, 1990.

[65] P. Kornprobst, R. Deriche, and G. Aubert. Image sequence analysis via partial differential equations. *Journal of Mathematial Imaging and Vision*, 11(1) :5–26, 1999.

[66] I. Laptev. On space-time interest points. *International Journal of Computer Vision*, 64(2/3), 2005.

[67] I. Laptev, H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner. Automatic extraction of roads from aerial images based on scale-space and snakes. *Machine Vision and Applications*, 12(1) :23–31, 2000.

[68] S. Lazebnik, Cordelia Schmid, and Jean Ponce. A sparse texture representation using local affine regions. *IEEE Transactions on PAMI*, 27(8) :1265–1278, 2005.

[69] O. LeCadet. Méthodes d'ondelettes pour la segmentation d'images – applications à l'imagerie médicale et au tatouage d'images, 2004. Thèse de doctorat de l'INPG.

[70] E. LePennec and S. Mallat. Image compression with geometrical wavelets, 2000.

[71] S. Lettrattanapanich and N. K. Bose. High resolution image formation from low resolution frames using delaunay triangulation. *IEEE Transactions on Image Processing*, 11(12), 2002.

[72] T. Lindeberg. Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3) :234–254, 1990.

[73] T. Lindeberg. Scale-space behaviour of local extrema and blobs. *Journal of Mathematical Imaging and Vision*, 1(1) :65–99, 1992.

[74] T. Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space primal sketch : A method for focus-of-attention. *International Journal of Computer Vision*, 11 :283–318, 1993.

[75] T. Lindeberg. *Scale-Space Theory in Computer Vision.* Kluwer, 1994.

[76] T. Lindeberg. Scale-space theory : A framework for handling image structures at multiple scale. In *Proc. CERN School of Computing, The Netherlands, Sept. 8–21*, 1996.

[77] T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision*, 30(2) :77–116, 1998.

[78] T. Lindeberg and D. Fagerstroem. Scale-space with causal time direction. *Lecture Notes in Computer Science*, 1064 :229–243, 1996.

[79] T. Lindeberg and J. Garding. Shape-adapted smoothing in estimation of 3-d shape cues from affine deformations of local 2-d brightness structure. *Image and Vision Computing*, 15(6) :415–434, 1997.

[80] A. Linderhed. 2d empirical mode decomposition in the spirit of image compression. *Wavelet and Indendant components analysis applications, SPIE Proceedings*, 4738(1–8), 2002.

[81] A. Linderhed. 2d variable sampling of the empirical mode decomposition of two-dimensional signals. *International Journal of Wavelets, Multiresolution and Information Processing*, 3(3), 2005. Special Issue on Frame Theory and Sampling Problems in Time-Frequency Analysis and Wavelet Theory (Part II).

[82] Z. Liu and S. Peng. Boundary processing of bemd using texture synthesis. *IEEE Signal Processing Letters*, 12(1), 2005.

[83] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2) :91–110, 2004.

[84] B. Luo, J. F. Aujol, Y. Gousseau, S. Ladjal, and H. Maître. Resolution independent characteristic scale dedicated to satellite images. *IEEE Transactions on Image Processing*, 16(10) :2503–2417, 2007.

[85] S. Mallat. Multiresolution approximations and wavelet orthonormal bases in $\mathbf{L}^2(\mathbf{R})$. *Trans. Amer. Math. Soc.*, 315 :69–87, 1989.

[86] S. Mallat. A theory for multiresolution signal decomposition : The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11 :674–693, 1989.

[87] S. Mallat. *A wavelet tour of signal processing.* Academic Press, 1998.

[88] S. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2) :617–643, 1992.

[89] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Anal. and Mach. Intell.*, 14(7) :710–732, 1992.

[90] S. Meignen and V. Perrier. A new formulation for empirical mode decomposition based on constrained optimization, 2007.

[91] J. Meinguet. Multivariate interpolation at arbitrary points made simple. *Journal of Applied Mathematics and Physics*, 30 :292–304, 1979.

[92] F. Meyer, A. Averbuch, and R. Coifman. Multi-layered image representation : application to image compression. *IEEE Transactions on Image Processing*, 11(9) :1072–1080, 2002.

[93] Y. Meyer. *Wavelets.* Springer Verlag, 1989.

[94] Y. Meyer. *Ondelettes et opérateurs I : ondelettes.* Paris, Hermann, 1990.

[95] Y. Meyer. *Ondelettes et opérateurs II : opérateurs de Calderon-Zygmund.* Paris, Hermann, 1990.

[96] Y. Meyer. *Ondelettes et opérateurs III : opérateurs multilinéaires.* Paris, Herman, 1990.

[97] K. Mikolajczyk. Detection of local features invariant to affine transformations, ph.d thesis, inpg grenoble, 2002.

[98] K. Mikolajczyk and C. Schmid. Indexing based on scale invariant interest points. In *Proceedings of the 8th IEEE International Conference on Computer Vision*, volume 1, pages 525–531, 2001.

[99] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *International Journal of Computer Vision*, 60(1), 2004.

[100] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10) :1615–1630, 2005.

[101] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 62(1) :43–72, 2005.

[102] D. S. Mitrinovic. *Analytic Inequalities.* New York, Springer-Verlag, 1970.

[103] F. Natterer. *The Mathematics of Computerized Tomography.* SIAM, Philadelphia, 2001.

[104] J.C. Nunes, S. Guyot, and E. Deléchelle. Texture analysis based on local analysis of the bidimensional empirical mode decomposition. *Journal of Machine Vision and Applications*, 16(3) :177–188, 2005.

[105] V. Perrier, A. Bilgot, O. LeCadet, and L. Desbat. Transformée en ondelettes continue directionnelle : applications en imagerie médicale. *ESAIM, 18, Ed. Gerbeau et Labbé*, pages 216–228, 2007.

[106] D. Po and M. N. Do. Directional multiscale modeling of images using the contourlet transform. *IEEE Transactions on Image Processing*, 15(6) :1610–1620, 2006.

[107] P. Pritchett and A. Zisserman. Wide baseline stereo matching. In *Proceedings ICCV*, pages 754–760, 1998.

[108] M. Protter and H. Weinberger. *Maximum Principles in differential equations.* Englewood Cliffs, Nj Prentice hall, 1967.

[109] G. Rilling, P. Flandrin, and P. Goncalves. On empirical mode decomposition and its algorithms. *IEEE Signal Processing Letters*, 2003.

[110] F. Schaffalitzky and A. Zisserman. Viewpoint invariant texture matching. In *Proceedings ICCV*, 2001.

[111] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5) :530–535, 1997.

[112] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2) :151–172, 2000.

[113] L. Schwartz. *Théorie des Distributions.* Hermann, Paris, 1966.

[114] R.C. Sharpley and V. Vatchev. Analysis of the intrinsic mode functions. *Constructive Approximation*, 24(1) :17–47, 2006.

[115] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D J Heeger. Shiftable multi-scale transforms. *IEEE Trans. Information Theory*, 38(2) :587–607, 1992.

[116] J. L. Starck, E. J. Candès, and D. L. Donoho. Very high quality image restoration by combining wavelets and curvelets. *Wavelet Applications in Signal and Image Processing IX*, SPIE 4478, 2001.

[117] J. L. Starck, E. J. Candès, and D. L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11 :670–684, 2002.

[118] D. Tell and S. Carlsson. Wide baseline point matching using affine invariants computed from intensity profiles. In *Proceedings ECCV*, 2000.

[119] T. Tuytelaars and L. Van Gool. Content-based image retrieval based on local affinely invariant regions. In *Int. Conf. on Visual Information Systems*, pages 493–500, 1999.

[120] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing – part i – theory. *IEEE Transactions on Signal Processing*, 41 :821–833, 1993.

[121] M. Unser, A. Aldroubi, and M. Eden. B-spline signal processing – part ii – efficient design. *IEEE Transactions on Signal Processing*, 41 :834–848, 1993.

[122] L. A. Vese and S. J. Osher. Image denoising and decomposition with total variation minimization and oscillatory functions. *Journal of Mathematical Imaging and Vision*, 20(1-2) :7–18, 2004.

[123] G. Wahba. *Spline models for observational data*. Society for Industrial and Applied Mathematics, Philadelphia, 1990.

[124] M. Weber, M. Welling, and P. Perona. Unsupervised learning of models for recognition. In *Proceedings of ECCV*, pages 18–32, 2000.

[125] A. P. Witkin. Scale space filtering. *Proc. Int. Joint Conf. on Artificial Intelligence*, pages 1019–1022, 1983.

[126] A. L. Yuille and T. A. Poggio. Scaling theorems for zero-crossing. *IEEE Transactions on Pattern Analysis and Machine Intelligence,*, 8 :15–25, 1986.

Remerciements

Je remercie en premier lieu Cordelia Schmid et Sylvain Meignen pour m'avoir encadré pendant cette thèse, où j'ai beaucoup appris d'un point de vue scientifique. Je souligne également les nombreux conseils Valérie Perrier, qui en dehors de son niveau d'expertise, témoigne autant d'énergie que de patience.

Francesc Arandiga et Patrick Flandrin m'ont fait l'honneur d'être les rapporteurs de cette thèse, et de lire avec attention mon manuscrit. Je tiens également à remercier les membres du jury pour leur travail critique vis à vis de ce manuscrit, et également pour leur attention lors de la soutenance.

En outre, je suis reconnaissant vis à vis de toutes les personnes avec qui j'ai pu discuter des problématiques de cette thèse, que ce soit avec des membres du laboratoire LJK, ou des personnes recontrées lors de séminaires et à l'occasion de différentes conférences (Wave 2006, Curves and Surfaces 2006, congrès SMAI 2007).

Je souhaite enfin remercier toutes les personnes du LJK qui contribuent à la bonne ambiance qui y règne, Jean-Baptiste Durand, Boris Thibert, Claire, Carine, Irène, Elie, Thomas, Alexandre, Erwan, Morgan, Ibrahim, Adrien, Yann, Azmi, sans oublier Mathieu, pour de nombreuses et intéressantes conversations.

Ondelettes pour la détection de caractéristiques en traitement d'images
Applications à la détection de régions d'intérêt

Cette thèse en traitement d'images aborde le problème de la mise en évidence de certaines structures remarquables, comme des objets que nous percevons visuellement. Celles-ci peuvent être autant monodimensionnelles, comme des contours, que bidimensionnelles, ce qui correspond à des objets plus complexes. Un problème important issu de la vision par ordinateur est de détecter de telles structures, ainsi que d'extraire des grandeurs caractéristiques de celles-ci. Dans diverses applications, comme la reconnaissance d'objets, l'appariement d'images, le suivi de mouvement ou le rehaussement de certains éléments particuliers, il s'agit d'une première étape avant d'autres opérations de plus haut niveau. Ainsi, la formulation de détecteurs performants apparaît comme essentielle. Nous montrons que cela peut être réalisé grâce des décompositions en ondelettes ; en particulier, il est possible de définir certaines lignes de maxima, qui s'avèrent pertinentes vis à vis de ce problème : d'une part, pour détecter des objets (par des régions d'intérêt), et, d'autre part, afin de les caractériser (calculs de régularité Lipschitzienne et d'échelle caractéristique). Cette approche originale de détection fondée sur des lignes de maxima peut alors être comparée aux approches classiques.

Mots-clés : ondelettes, *Scale-Space*, représentations multiéchelles, lignes de maxima, invariance ou robustesse à des transformations de l'image, détection de régions d'intérêt, calcul de grandeurs caractéristiques.

Wavelets for detection of features in image processing
Applications to the detection of regions of interest

This thesis in image processing address the problem of the highlight of some remarquable structures, such as objects we perceive visually. These can be monodimensional, like contours, as well as bidimensional, corresponding to more complex objects. An important problem in computer vision consists on detecting such structures, and also extracting characteristic features from them. In many applications, such as object recognition, image matching, motion tracking or the enhancement of some particular elements, it is a first step before other high-level operations. Thereby, the formulation of performant detectors appears as essential. We show that this can be carried out using wavelet decompositions ; in particular, it is possible to define some maxima lines, which turn out as relevant to this problem : one the one hand, so as to detect objects (given by some regions of interest), and, on the other hand, in order to characterize them (computations of Lipschitz regularity and of characteristic scale). This original approach for detection, based on maxima lines, can thus be compared to classical approches.

Keywords : wavelets, *Scale-Space*, multiscale representations, maxima lines, invariance of robustness to image transformations, detection of interest regions, computation of characteristic values.