



**HAL**  
open science

# Prise en compte de l'environnement marin dans le processus de reconnaissance automatique de cibles sous-marines

Laurent Picard

► **To cite this version:**

Laurent Picard. Prise en compte de l'environnement marin dans le processus de reconnaissance automatique de cibles sous-marines. Traitement du signal et de l'image [eess.SP]. Université de Bretagne occidentale - Brest, 2017. Français. NNT : 2017BRES0038 . tel-01616196

**HAL Id: tel-01616196**

**<https://theses.hal.science/tel-01616196v1>**

Submitted on 13 Oct 2017

**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.

**THÈSE / UNIVERSITÉ DE BRETAGNE OCCIDENTALE**

*sous le sceau de l'Université Bretagne Loire*

pour obtenir le titre de  
**DOCTEUR DE L'UNIVERSITÉ DE BRETAGNE OCCIDENTALE**

*Mention : Traitement du signal et des images*

présentée par

**Laurent PICARD**

Préparée à l'ENSTA Bretagne

**École Doctorale SICMA**

Prise en compte de  
l'environnement marin dans  
le processus de reconnaissance  
automatique de cibles  
sous-marines.

**Thèse soutenue le 18 mai 2017**

devant le jury composé de :

**Philippe BLONDEL**

Senior Lecturer, Université de Bath / Rapporteur

**Jean-Marie NICOLAS**

Professeur des Universités, Télécom Paris Tech / Rapporteur

**Christophe COLLET**

Professeur des Universités, Télécom Strasbourg

**Gilles BUREL**

Professeur des Universités, Université de Bretagne Occidentale

**Alexandre BAUSSARD**

Maître de Conférences, ENSTA Bretagne

**Isabelle QUIDU**

Maître de Conférences, ENSTA Bretagne

**Gilles LE CHENADEC**

Enseignant Chercheur, ENSTA Bretagne

**Dominique FATTACCIOLI**

Expert senior en acoustique sous-marine,  
DGA Techniques Navales

**Zamirddine MARI**

DGA Techniques Navales



# Résumé

Au cours des dernières décennies, les avancées en termes de technologies robotiques sous-marines ont permis de réaliser des levés sur les fonds marins à l'aide de véhicules sous-marins autonomes (AUV). Ainsi, équiper un AUV avec un sonar latéral permet de scanner une vaste zone de manière rapide. Naturellement, les forces armées se sont intéressées à de tels dispositifs pour effectuer des missions de chasses aux mines rapides et sécurisées pour le facteur humain. Néanmoins, analyser des images sonar par un ordinateur plutôt que par un opérateur, reste très complexe. En effet, les chaînes de reconnaissance automatique de cibles (ATR) doivent faire face à la variabilité de l'environnement marin et il a été démontré qu'une forte relation existe entre la texture d'une image et la difficulté d'y détecter des mines. Effectivement, sur des fonds fortement texturés, voire encombrés, les performances d'une chaîne ATR peuvent être très dégradées. Ainsi, intégrer des informations environnementales dans le processus apparaît comme une piste crédible pour améliorer ses performances. Ces travaux de thèse proposent d'étudier la manière de décrire cet environnement marin et comment l'intégrer dans un processus ATR.

Pour répondre à ces défis, nous proposons tout d'abord une nouvelle représentation des images sonar basée sur l'utilisation du signal monogène. Ce dernier permet d'extraire des informations énergétiques, géométriques et structurelles sur la texture locale d'une image. La nature multi-échelle de cet outil permet de tenir compte de la variabilité en taille des structures sous-marines. Ensuite, le concept de dimension intrinsèque est introduit pour décrire une image sonar en termes d'homogénéité, d'anisotropie et de complexité. Ces trois descripteurs sont directement reliés à la difficulté de détection des mines sous-marines dans un fond texturé et permettent de réaliser une classification très précise des images sonar en fonds homogènes, anisotropes et complexes. De notre point de vue, la chasse aux mines sous-marines ne peut pas être réalisée de la même manière sur ces trois types de fond. En effet, leurs natures et caractéristiques propres mènent à des challenges variés pour le processus ATR. Pour le démontrer, nous proposons de réaliser un premier algorithme de détection spécifique, appliqué aux zones anisotropes, qui prend en considération les caractéristiques environnementales de ces régions.

**Mots-clés :** Images sonar latéral, processus ATR, description des fonds marins, signal monogène, dimension intrinsèque, analyse multi-échelle.



# Table des matières

Résumé	i
Table des matières	iii
Table des figures	ix
Liste des abréviations	xxi
<b>1 Introduction</b>	<b>1</b>
1.1 Introduction générale	1
1.2 Contexte	3
1.3 Objectifs	4
1.4 Organisation	5
<b>2 Principes de l'imagerie sonar</b>	<b>7</b>
2.1 Introduction	8
2.2 Principes généraux des sonars de cartographie	9
2.2.1 Émission et acquisition des signaux	10
2.2.2 Étude et mise en forme des signaux en réception	15
2.3 Imagerie acoustique des fonds marins	19
2.3.1 Origine du speckle et modélisations statistiques associées	20
2.3.2 Images texturées et complexes	24
2.3.3 Apparition d'objets	26
2.4 Présentation des bases de données utilisées dans la thèse	28
2.5 Conclusion	30

<b>3</b>	<b>Détection et classification de cibles en imagerie sonar</b>	<b>31</b>
3.1	Introduction . . . . .	32
3.2	Les processus ATR . . . . .	32
3.2.1	Phase de détection . . . . .	33
3.2.2	Phase d'extraction d'attributs . . . . .	36
3.2.3	Phase de Classification . . . . .	37
3.2.4	Limites et perspectives . . . . .	39
3.3	Nécessité d'une caractérisation environnementale . . . . .	40
3.3.1	Trois grandes catégories de fonds marins pour l'ATR . . . . .	40
3.3.2	Descripteurs d'anisotropie et de complexité des fonds marins . . . . .	41
3.3.3	Exploitation des descripteurs environnementaux dans l'ATR . . . . .	52
3.4	Approche proposée dans cette thèse . . . . .	57
3.4.1	Position par rapport à la littérature . . . . .	57
<b>4</b>	<b>Représentation monogène d'une image</b>	<b>63</b>
4.1	Introduction . . . . .	65
4.2	Analyse complexe des signaux 1D . . . . .	66
4.2.1	La transformée de Hilbert . . . . .	66
4.2.2	Le signal analytique . . . . .	67
4.2.3	L'analyse complexe . . . . .	68
4.2.4	Discussion . . . . .	70
4.3	Les extensions 2D du signal analytique . . . . .	71
4.3.1	La transformée de Hilbert 2D . . . . .	71
4.3.2	Au delà des nombres complexes . . . . .	72
4.3.3	L'algèbre géométrique . . . . .	72
4.4	Définition du signal monogène . . . . .	73
4.4.1	La transformée de Riesz . . . . .	73
4.4.2	Le signal monogène . . . . .	78
4.4.3	Définition générale du signal monogène . . . . .	81
4.5	Signal monogène et Analyse multi-échelle . . . . .	83

---

4.5.1	Théorie de l'espace d'échelle . . . . .	83
4.5.2	Espace d'échelle de Poisson . . . . .	86
4.5.3	Analyse multi-échelle . . . . .	91
4.5.4	Représentation monogène multi-échelle . . . . .	100
4.6	Application à l'image sonar . . . . .	103
4.6.1	Paramétrage de l'analyse multi-échelle . . . . .	103
4.6.2	Représentation monogène d'une image sonar latéral . . . . .	107
4.6.3	Application à la classification des fonds marins . . . . .	111
4.7	Résumé du Chapitre . . . . .	115
<b>5</b>	<b>La dimension intrinsèque des images</b> . . . . .	<b>117</b>
5.1	Introduction . . . . .	119
5.2	La dimension intrinsèque discrète . . . . .	120
5.2.1	Introduction . . . . .	120
5.2.2	La dimension intrinsèque en traitement d'image . . . . .	120
5.2.3	Estimation de la dimension intrinsèque discrète . . . . .	122
5.2.4	Potentialités pour l'étude de la complexité des images sonar . . . . .	122
5.3	La dimension intrinsèque continue . . . . .	124
5.3.1	Topologie triangulaire . . . . .	124
5.3.2	Topologie conique . . . . .	129
5.3.3	Dimensions intrinsèques continues multi-échelles . . . . .	136
5.4	Application à l'imagerie sonar . . . . .	138
5.4.1	Paramétrage spécifique aux images sonar . . . . .	138
5.4.2	Analyse des fonds marins homogènes . . . . .	139
5.4.3	Analyse des fonds marins anisotropes . . . . .	141
5.4.4	Analyse des fonds marins complexes . . . . .	143
5.5	Résumé du chapitre . . . . .	144

<b>6</b>	<b>Approche proposée pour la description des fonds marins</b>	<b>147</b>
6.1	Introduction . . . . .	149
6.2	Description continue des fonds marins . . . . .	150
6.2.1	Formulation des objectifs . . . . .	150
6.2.2	Homogénéité des fonds . . . . .	150
6.2.3	Anisotropie des fonds . . . . .	156
6.2.4	Complexité des fonds . . . . .	162
6.2.5	Résultats . . . . .	163
6.2.6	Comparaison avec une méthode existante . . . . .	166
6.2.7	Liens avec l'ATR . . . . .	173
6.3	Classification non-supervisée des fonds marins . . . . .	175
6.3.1	Introduction . . . . .	175
6.3.2	Classification des fonds homogènes . . . . .	175
6.3.3	Classification des fonds anisotropes et complexes . . . . .	179
6.3.4	Résultats . . . . .	180
6.4	Conclusion . . . . .	183
<b>7</b>	<b>Exploitation de la description environnementale en ATR</b>	<b>185</b>
7.1	Introduction . . . . .	187
7.2	Application à la caractérisation des fonds marins . . . . .	187
7.2.1	Motivations . . . . .	187
7.2.2	Caractérisation des fonds homogènes . . . . .	188
7.2.3	Caractérisation des fonds anisotropes . . . . .	198
7.3	Architecture ATR adaptée au type de fond . . . . .	205
7.3.1	Prise en compte de la caractérisation environnementale . . . . .	205
7.3.2	Stratégie proposée . . . . .	206
7.4	Détection de cibles dans les zones anisotropes . . . . .	207
7.4.1	Analyse des zones anisotropes avec et sans objets . . . . .	207
7.4.2	Détection d'échos . . . . .	209
7.4.3	Détection d'ombres . . . . .	215

---

7.4.4	Résultats finaux de la détection . . . . .	216
7.5	Conclusion et perspectives . . . . .	220
<b>8</b>	<b>Conclusion générale</b>	<b>221</b>
<b>A</b>	<b>Les transformées de Hilbert 2D classiques</b>	<b>227</b>
A.1	La transformée de Hilbert partielle . . . . .	227
A.2	La transformée de Hilbert totale . . . . .	228
A.3	La transformée de Hilbert 2D par approche "Single Orthant" . . . . .	229
A.4	Le signal analytique hypercomplexe . . . . .	230
<b>B</b>	<b>Algèbre géométrique</b>	<b>233</b>
B.1	Origines de l'algèbre géométrique . . . . .	233
B.1.1	Premières définitions . . . . .	233
B.1.2	Cas de l'espace Euclidien 2D . . . . .	235
B.1.3	Cas de l'espace Euclidien 3D . . . . .	235
B.1.4	Les spineurs . . . . .	236
B.2	Algèbre géométrique et traitement du signal . . . . .	237
<b>C</b>	<b>Généralisation isotrope du signal analytique</b>	<b>241</b>
C.1	Généralisation de la transformée de Hilbert . . . . .	241
C.1.1	Équations de Cauchy-Riemann . . . . .	241
C.1.2	Équations de Cauchy-Riemann généralisées . . . . .	242
C.2	Résolution des équations de Cauchy-Riemann pour $z = 0$ . . . . .	243
C.3	Résolution des équations de Cauchy-Riemann pour $z > 0$ . . . . .	245
	<b>Bibliographie</b>	<b>247</b>



# Table des figures

1.1	Exemple de moyens de contrôle de la surface océanique : (a) avion de type ATL2, (b) navire de patrouille maritime (extraits du site internet du ministère de la défense). . . . .	2
1.2	(a) chasseur de mine "La Croix du Sud" de la marine nationale française (extrait du site du ministère de la défense), (b) AUV proche de la surface (extrait du site internet de l'IFREMER). . . . .	3
2.1	Exemple d'un sonar latéral Klein 5000 monté sur un poisson (documentation Klein). . . . .	9
2.2	Géométrie d'acquisition d'un sonar latéral positionné sur un poisson tracté par un bateau. . . . .	10
2.3	Géométrie d'une antenne considérée comme ponctuelle. . . . .	11
2.4	Exemple de diagramme de directivité en azimuth en coordonnées polaires. . . . .	12
2.5	Cellule de résolution d'un sonar. . . . .	13
2.6	Géométrie d'acquisition d'une voie de sonar latéral et le signal reçu correspondant. . . . .	15
2.7	Géométrie d'acquisition d'un sonar latéral positionné sur un poisson tracté par un bateau. L'image sonar résultante est représentée juste en dessous. . . . .	17
2.8	Exemple d'une image acoustique de sonar latéral avec certaines de ses caractéristiques. . . . .	18
2.9	Mosaïque d'images acoustiques provenant de différents systèmes sonars. En haut, un sonar latéral Marine Sonic Technology Ltd de fréquence 900 kHz. En bas, un sonar latéral Klein 5000 de fréquence 455 kHz. . . . .	19
2.10	Illustration du speckle pour deux blocs rectangulaires d'une image sonar Klein 5000. . . . .	21
2.11	Représentation de la loi de Rayleigh pour différentes valeurs de paramètres. . . . .	23

2.12	Mosaïque représentant différents types de fonds marins. Sur ces différentes images, nous pouvons apercevoir différentes structures sous-marines comme des zones de sable fin, des rides de sables, des roches, des traces de dragage, ou encore des objets posés sur le fonds. . . . .	25
2.13	Représentation schématique de l'insonification d'un objet en champ lointain (inspirée de [1]). . . . .	26
2.14	Mosaïque de différents objets apparaissant dans des image sonars. . . . .	27
2.15	Exemple de la difficulté de détection d'une cible posée sur un fond non-homogène (rides de sable ici). Le rectangle vert indique la position d'une mine sous-marine. En (a), paire écho-ombre de l'objet visible. En (b), paire écho-ombre non-visible. . . . .	28
2.16	Exemples d'images sonar pour chacune des trois bases de données, (a) base de données DSTO, (b) base de données DGA, et (c) base de données CMRE. . . . .	29
3.1	Séquence de traitements d'un processus ATR. . . . .	33
3.2	Exemple d'un fond homogène (a), anisotrope (b), et complexe (c). . . . .	40
3.3	Trois exemples de fonds marins : homogène (a), anisotrope (b), complexe (c), et leurs transformées de Fourier (TF) respectives (d), (e), et (f). L'origine du plan fréquentiel est considéré au centre de l'image. . . . .	43
3.4	Illustration d'un filtre de Gabor en 3D (a) et de ses versions orientées (b). . . . .	45
3.5	Emprunte des filtres de Gabor dans le domaine fréquentiel pour 6 orientations (une par couleur) et 3 échelles. . . . .	46
3.6	(a) Exemple d'une image sonar réelle issue d'un sonar latéral Marine Sonic sur laquelle sont extraits deux descripteurs à l'aide des filtres de Gabor : (b) l'anisotropie et (c) la complexité. . . . .	46
3.7	Différents filtres de Haar orientés selon 4 directions distinctes $0^\circ$ , $90^\circ$ , $45^\circ$ et $135^\circ$ . . . . .	48
3.8	Résultats des descripteurs d'anisotropie et de complexité extraits de l'étude de E. Fakiris [2], basée sur l'utilisation des filtres de Haar. Ces résultats sont proposés pour 4 et 12 orientations distinctes de ces filtres. . . . .	49
3.9	Orientation des ondelettes réelles (en haut), et imaginaires (en bas). . . . .	50
3.10	Exemple d'une image sonar réelle issue d'un sonar latéral Marine Sonic (a) sur laquelle sont extraits deux descripteurs à l'aide de la DTCWT : l'anisotropie (b) et la complexité (c). . . . .	51
3.11	Schématisation de la stratégie d'intégration mise en place par D.P. Williams pour intégrer des descripteurs environnementaux dans le processus ATR. . . . .	53

3.12 Exemples de trois environnements marins distincts dans lesquels une mine sous-marine est positionnée : (a) sur un fond homogène, (b) sur une frontière entre un fond homogène et un fond de rides de sable, et (c) sur un fond de rides de sable. La distribution des coefficients d'anisotropie et de complexité autour de ces trois cibles est représentée en (d) (résultats extraits de [3]). . . . .	54
3.13 Schématisation de la stratégie mise en place par O. Daniell pour intégrer des descripteurs environnementaux dans le processus ATR. . . . .	56
3.14 Exemple d'une image sonar où les résultats d'une chaîne ATR sont illustrés. L'ensemble des rectangles correspond aux détections fournies par cette dernière. Les rectangles bleus définissent les vérités terrains, les rectangles rouges correspondent aux détections filtrées par l'algorithme de O. Daniell, et les blancs sont les détections gardées par ce même algorithme (Résultats extraits de [4]). . . . .	57
3.15 Schématisation de la stratégie mise en place par O. Daniell pour intégrer des descripteurs environnementaux dans le processus ATR. . . . .	60
3.16 Zoom sur l'architecture de la description de l'environnement proposée. . . . .	61
4.1 Position de la représentation monogène dans le processus. . . . .	65
4.2 (a) Visualisation du signal $s(t) = A \cos(2\pi f_0 t)$ (en trait plein) et de sa transformée de Hilbert $s_{\mathcal{H}}(t)$ (en trait discontinu). (b) Illustration de l'évolution de l'amplitude instantanée $A_s(t)$ et de la phase instantanée $\phi_s(t)$ du signal $s(t)$ . . . . .	69
4.3 Valeurs caractéristiques de la phase instantanée et formes associées du signal. . . . .	69
4.4 Représentation du signal analytique dans le plan complexe. . . . .	71
4.5 Représentation 3D des deux noyaux de Riesz $r_1$ (à gauche) et $r_2$ (à droite). . . . .	74
4.6 Exemple d'une sinusoïde 2D de fréquence et d'orientation fixe (à gauche) pour laquelle la norme $\mathcal{N}_R$ et l'argument $\theta_R$ de la transformée de Riesz sont illustrées (au milieu et à droite respectivement). . . . .	77
4.7 Illustration d'une sinusoïde 2D isotrope (a) sur laquelle les informations d'amplitude (b), de phase (c) et d'orientation locale (d) sont extraites. . . . .	79
4.8 Illustration d'une sinusoïde 2D isotrope dont l'amplitude décroît de manière radiale à partir du centre de l'image (a) sur laquelle les informations d'amplitude (b), de phase (c) et d'orientation locale (d) sont extraites. . . . .	80
4.9 Illustration schématique de la représentation en espace d'échelle d'un signal 1D (extraite de [5]). . . . .	84
4.10 Résultats de l'application d'un filtrage gaussien sur l'image Lena (image la plus à gauche) pour des valeurs de $\sigma$ croissantes. . . . .	85

4.11	Représentation spatiale et fréquentielle des filtres de Poisson 2D (coupe horizontale selon la direction $x$ et $u_1$ respectivement) pour un ensemble de paramètres distincts. . . . .	86
4.12	Illustration d'un noyau gaussien (à gauche) et d'un noyau de Poisson (à droite). . . . .	87
4.13	Comparaison entre le noyau gaussien (en rouge) et de Poisson (en bleu) dans le domaine spatial et fréquentiel, de gauche à droite respectivement. . . . .	88
4.14	Illustration de plusieurs isophotes dans le cas 1D : en abscisse la variable de position $x$ et en ordonnée la variable d'échelle $s$ . (a) Isophote dont l'extremum local possède une convexité vers les échelles grossières. (b) Isophote dont l'extremum local possède une concavité vers les échelles grossières. (c) Isophote dont les deux extrémités sont reliées à l'image d'origine ( $s = 0$ ). . . . .	89
4.15	Résultats de l'application d'un noyau de Poisson sur l'image Lena (image la plus à gauche) pour des valeurs de $s$ croissantes. . . . .	90
4.16	A gauche : Exemple d'une image contenant 3 disques de diamètres respectifs 10, 20 et 40 pixels. Cette image est lissées pour 4 valeurs croissantes du paramètre d'échelle $s$ (de gauche à droite). . . . .	91
4.17	Processus de construction de l'analyse multi-échelle. . . . .	92
4.18	Représentation 3D spatiale (a) et fréquentielle (b) d'une différence de gaussienne pour $\sigma = 1$ et $k = \sqrt{2}$ . Une coupe 1D de ces deux noyaux est proposée en (c) et (d) respectivement. . . . .	93
4.19	Schématisation d'une découpe en octaves d'un axe fréquentiel. . . . .	94
4.20	Construction successive des filtres DoG. . . . .	95
4.21	Discrétisation en fréquence des différences de gaussiennes en fonction du paramètre $n$ . En trait plein, une seule décomposition par octave est réalisée (un octave par couleur). En pointillé, deux décompositions par octave est cette fois réalisée (même couleur pour les décompositions appartenant au même octave). . . . .	96
4.22	Représentation 3D spatiale (a) et fréquentielle (b) d'une différence de Poisson pour $s = 1$ et $k = \sqrt{2}$ . Une coupe 1D de ces deux noyaux est proposée en (c) et (d) respectivement. . . . .	97
4.23	Discrétisation en fréquence des différences de Poisson en fonction du paramètre $n$ . En trait plein, une seule décomposition par octave est réalisée (un octave par couleur). En pointillé, deux décompositions par octave est cette fois réalisée (même couleur pour les décompositions appartenant au même octave). . . . .	98

4.24	A gauche : Exemple d'une image contenant 3 disques de diamètres respectifs 5, 10 et 20 pixels. A droite : Résultats du filtrage de cette image par un filtre DoP pour 3 niveaux de décomposition. . . . .	99
4.25	Processus de construction de la représentation monogène multi-échelle. . . . .	101
4.26	Illustration d'une représentation monogène multi-échelle à quatre niveaux de détail pour une image synthétique. . . . .	102
4.27	Exemple de trois mines sous-marines : (a) une cylindrique, (b) une Manta et (c) une Rockan. . . . .	103
4.28	Dimensions d'une structure de l'image à la plus forte réponse du DoP en fonction des paramètres $s$ et $k$ . . . . .	104
4.29	Représentation des 4 filtres passe-bande DoP pour chacun des 3 octaves dans le domaine fréquentiel en (a), (b) et (c), puis dans le domaine spatial en (d), (e) et (f). . . . .	105
4.30	Sur la première ligne, exemple d'image sonar contenant un objet de petite dimension et ses réponses aux filtres DoP pour différents niveaux de détails. Sur la seconde ligne, les mêmes informations sont représentées mais pour une structure plus importante (on peut ici reconnaître un voilier). . . . .	106
4.31	Représentation monogène multi-échelle d'un objet posé sur un fond homogène (Données DSTO). . . . .	107
4.32	Représentation monogène multi-échelle d'un champ de rides de sable traversant l'image (Données DSTO). . . . .	109
4.33	Représentation monogène multi-échelle d'un amas de roches (Données DSTO). . . . .	110
4.34	Illustration de la base d'apprentissage : (de gauche à droite) posidonie, rides de sable, roche, sable, et vase. . . . .	112
4.35	Deux exemples d'image sonar et leur classification respective en 5 classes (a) et (b). Les couleurs correspondant aux différentes classes sont les suivantes, jaune = vase, vert = sable, bleu clair = posidonie, bleu foncé = rides de sable et rouge = roche. . . . .	114
4.36	Deux exemples d'image sonar et leur classification respective en 3 classes (a) et (b). Les couleurs correspondant aux différentes classes sont les suivantes, vert = homogène, bleu = anisotrope, et rouge = complexe. . . . .	115
5.1	Position de l'extraction d'informations dans le processus. . . . .	119

5.2	Exemple de 6 images aux dimensions intrinsèques différentes : (a)-(b) images de signaux homogènes i0D, (c)-(d) images de signaux unidirectionnel i1D, et (e)-(f) images de signaux ni constant ni unidirectionnel i2D. . . . .	121
5.3	Trois exemples de fonds marins : homogène (a), anisotrope (b), complexe (c), et leurs transformées de Fourier (TF) respectives (d), (e), et (f). L'origine du plan fréquentiel est considéré au centre de l'image. . . . .	123
5.4	Topologie de la dimension intrinsèque. (a) Topologie carrée initiale, (b) topologie triangulaire et (c) paramétrisation de la topologie triangulaire. . . . .	126
5.5	Positions dans le triangle de différents blocs i1D, A et B, orientés selon des directions opposées. La position de leur dimension intrinsèque moyenne C ainsi que celle de l'union des deux blocs D sont également présentées. . . . .	129
5.6	Topologie conique de la dimension intrinsèque. . . . .	130
5.7	Positions dans le cône de différents blocs i1D, A et B, orientés selon des directions opposées. La position de leur dimension intrinsèque moyenne C ainsi que celle de l'union des deux blocs D sont également présentées. . . . .	130
5.8	Illustration des indices de confiance $c_{i0D}$ , $c_{i1D}$ et $c_{i2D}$ pour quatre images synthétiques. . . . .	134
5.9	Illustration des indices de confiance $c_{i0D}$ , $c_{i1D}$ et $c_{i2D}$ pour une image réelle. . . . .	135
5.10	Illustration de la dimension intrinsèque multi-échelle appliquée à l'image Lena. . . . .	137
5.11	Illustration de la dimension intrinsèque continue multi-échelle appliquée à un fond marin homogène. . . . .	141
5.12	Illustration de la dimension intrinsèque continue multi-échelle appliquée à un fond marin anisotrope. . . . .	142
5.13	Illustration de la dimension intrinsèque continue multi-échelle appliquée à un fond marin complexe. . . . .	143
6.1	Position de la phase de description dans le processus. . . . .	149
6.2	Synoptique de la première étape. . . . .	150
6.3	Exemple de deux cartes d'homogénéité pour une image sonar latéral Klein 5000 (a). En (b), une carte basée sur la moyenne, et en (c) une carte basée sur la médiane. . . . .	152
6.4	Exemple de deux cartes d'homogénéité pour une image sonar latéral Marine Sonic (a). En (b), une carte basée sur la moyenne, et en (c) une carte basée sur la médiane. . . . .	153

6.5	Exemple de deux images sonar (a) (b), ainsi que leur carte d'homogénéité (c) (d) et d'inhomogénéité (e)(f) respectives. . . . .	155
6.6	Exemple de deux cartes d'anisotropie (a) et (b) correspondantes aux images sonar présentées à la Figure 6.5. . . . .	156
6.7	Schéma des différentes formes de rides possibles dans une image sonar (extrait et adapté de [6]). . . . .	158
6.8	Répartitions des longueurs d'onde et des hauteurs pour les structures directives du fond marin (extraite et adaptée de [7]). . . . .	158
6.9	Comparaison entre un objet présent dans un champ de rides de sable et dans une zone de dunes sous-marines. . . . .	159
6.10	Exemple d'une image sonar (a) et des deux cartes d'anisotropie $A_1$ (b) et $A_2$ (c) correspondantes. . . . .	160
6.11	Exemple d'une image sonar (a) et des deux cartes d'anisotropie $A_2$ (b) et $A$ (c) correspondantes. . . . .	162
6.12	Exemple d'une image sonar issue de la base de données DGA (a) pour laquelle les cartes d'homogénéité (b), d'anisotropie (c) et de complexité sont extraites. . . . .	163
6.13	Exemple d'une image sonar issue de la base de données DSTO (a) pour laquelle les cartes d'homogénéité (b), d'anisotropie (c) et de complexité sont extraites. . . . .	164
6.14	Exemple d'une image sonar issue de la base de données CMRE (a) pour laquelle les cartes d'homogénéité (b), d'anisotropie (c) et de complexité sont extraites. . . . .	165
6.15	Exemple d'une image sonar (a) issues de la base DSTO. Les cartes d'anisotropie et de complexité sont illustrées pour l'algorithme de O. Daniell sur la première ligne, en (b) et (c) respectivement, et pour notre approche en (d) et (e). . . . .	168
6.16	Exemple d'une image sonar (a) issues de la base DGA. Les cartes d'anisotropie et de complexité issues l'algorithme de O. Daniell sont illustrées sur la première ligne, en (b) et (c) respectivement, et celles issues de notre approche en (d) et (e). . . . .	170
6.17	A gauche, une image sonar issues de la base de donnée DGA. A droite, sa segmentation manuelle en termes de zones homogènes (en vert), anisotropes (en bleu) et complexes (en rouge). Les grandes zones d'ombre ont également été segmentées et apparaissent sans couleur. . . . .	171
6.18	Histogrammes 2D représentant la répartition des pixels en fonction des descripteurs d'anisotropie et de complexité calculés par O. Daniell dans les zones homogènes (a), anisotropes (b) et complexes (c) de l'image sonar présentée à la Figure 6.17. . . . .	172

6.19	Histogrammes 2D représentant la répartition des pixels en fonction des descripteurs d'anisotropie et de complexité que nous proposons dans ces travaux pour des zones considérées comme homogènes (a), anisotropes (b) et complexes (c) de l'image sonar présentée à la Figure 6.17. . . . .	172
6.20	Illustration du voisinage autour d'une détection donnée (en rouge) et de sa fenêtre garde (en vert). . . . .	174
6.21	Répartition des fausses alarmes (a) et des manques (b) du détecteur de D.P. Williams en fonction des coefficients d'anisotropie et de complexité. . . . .	174
6.22	Synoptique de l'étape de classification. . . . .	176
6.23	Effets de l'augmentation du seuil $T$ sur l'extraction des zones homogènes (en transparence verte) d'une image sonar (a) issue de la base de données DSTO. De (b) à (e), $T = 0, 0.25, 0.5, 0.75$ respectivement. . . . .	177
6.24	Pourcentage des cibles considérées comme n'appartenant pas à des zones homogènes de l'image en fonction du seuil $T$ . . . . .	178
6.25	Pourcentage de la surface totale des images sonar de la base de données dus CMRE considérée comme non-homogènes ( <i>i.e.</i> restantes à étudier) en fonction du seuil $T$ . . . . .	179
6.26	Illustration de résultats de classification pour deux images sonar issues de la base de données DGA. . . . .	181
6.27	Illustration de résultats de classification pour deux images sonar issues de la base de données CMRE. . . . .	182
6.28	Illustration de résultats de classification pour deux images sonar issues de la base de données DSTO. . . . .	183
7.1	Illustration de résultats de classification pour deux images sonar issues de la base de données DGA. . . . .	189
7.2	En (a) et (b) : illustration de l'histogramme en niveaux de gris pour une même image sonar avant et après rehaussement du contraste. En (c) : autre exemple d'image sonar et de son histogramme rehaussé. . . . .	190
7.3	Histogrammes des zones homogènes (en transparence verte) de deux images sonar de la base de données DGA. . . . .	191
7.4	Segmentation de deux histogrammes par la méthode FTC. . . . .	193
7.5	Résultats d'extraction des zones d'ombre par segmentation d'histogramme : (a) image sonar initiale, (b) zones d'ombre, (c) prise en compte dans l'image classée, où les zones sombres correspondent aux ombres détectées. . . . .	194

7.6	Résultats d'extraction des zones d'ombre par utilisation des quantiles pour une image sonar de la base de données du DSTO (a) pour trois valeurs distinctes de percentiles, à savoir pour le 1 <sup>er</sup> (b), 5 <sup>ème</sup> (c) et le 10 <sup>ème</sup> (d). . . . .	195
7.7	Synoptique de la chaîne de détection des ombres. . . . .	196
7.8	Résultats finaux de classification pour deux images sonar (a) et (c) de la base de données DSTO : la transparence verte indique les zones classées comme homogènes, la bleue les zones classées comme anisotropes, la rouge les zones classées comme complexes et la noire les zones classées comme de l'ombre. . . . .	197
7.9	Résultats finaux de classification pour deux images sonar (a) et (c) de la base de données DGA. . . . .	197
7.10	illustration des réponses énergétiques des filtres passe-bande DoP pour quatre niveaux de détail. . . . .	199
7.11	illustration des indices $c_{i1D}$ pour quatre niveaux de détail. . . . .	199
7.12	Image sonar issues de la base DSTO (a) et du CMRE (c) ainsi leurs cartes des échelles $S$ respectives (c) et (d). . . . .	200
7.13	Illustration des orientations monogènes pour quatre niveaux de détail. . . . .	201
7.14	Carte d'orientation des rides de sables pour les deux images sonar de la Figure 7.12. . . . .	202
7.15	Schéma d'une possibilité de replanification dans les zones anisotropes. . . . .	202
7.16	Illustration des phases monogènes pour 4 échelles parmi les 12 totales. . . . .	203
7.17	Représentation d'une imagerie issues de la base de données DGA (a) et de sa carte $P$ des phases monogènes. . . . .	203
7.18	Coupes 1D horizontales de l'imagerie (a) (en bleu) et de sa carte $P$ (b) (en rouge) de la Figure 7.17. . . . .	204
7.19	Classification d'une imagerie sonar en termes de crêtes (en bleu) et de vallées (en rouge). . . . .	205
7.20	Architecture de la stratégie proposée. . . . .	206
7.21	Représentation monogène de l'imagerie d'un fond marin anisotrope. . . . .	208
7.22	Représentation monogène de l'imagerie d'un fond marin anisotrope où une mine sous-marine se situe. . . . .	209
7.23	Image sonar SAS issues de la base de données CMRE (a) et sa carte des échos $E$ associées (b). . . . .	210
7.24	Cartes $E^2$ (a) et $E^3$ (b) ainsi que leurs distributions associées en (c) et (d). . . . .	211

7.25	Configuration de la fenêtre TFAC. . . . .	212
7.26	Courbe ROC du détecteur d'écho proposé. L'axe des abscisses représente le nombre de fausses alarmes par $\text{km}^2$ de rides de sable, alors que l'axe des ordonnées représente la probabilité de détection. . . . .	213
7.27	Exemples de la difficulté de détection d'un écho. En vert la vérité terrain, en rouge le résultat de détection de l'algorithme proposé. . . . .	214
7.28	Résultats de la détection d'échos pour deux images sonar de la base de données CMRE. La vérité terrain est entourée en vert alors que les résultats de la détection d'échos sont pointés en rouge. . . . .	214
7.29	Classification en termes de crêtes (en bleu) et de vallées (en rouge) pour deux imageries sonar (a) et (d), où une mine sous-marine est présente à l'intérieur du champ de rides de sable. En (b) et (e), l'information de phase monogène est également représentée. . . . .	215
7.30	Image sonar issues de la base de données CMRE et sa carte des vallées associée à la zone de rides de sable. . . . .	216
7.31	Configuration pour l'estimation de la hauteur d'un objet. . . . .	217
7.32	Exemple de détections finales sur des images sonar issues de la base de données CMRE : En rouge les détections de l'algorithme, en vert, les vérités terrain. . . . .	217
7.33	Courbe ROC du détecteur proposé (en rouge) et celle du détecteur d'écho seul (en bleue). L'axe des abscisses représente le nombre de fausses alarmes par $\text{km}^2$ de rides de sable alors que l'axe des ordonnées représente la probabilité de détection. . . . .	218
7.34	Exemples de la difficulté de détection d'une ombre. En vert la vérité terrain, en rouge le résultat de la détection d'échos, et en jaune le résultat de la détection d'échos et d'ombres. . . . .	218
7.35	Comparaison de la probabilité de détection et du nombre de fausses alarmes par $\text{km}^2$ entre le détecteur d'écho seul, le détecteur d'écho et d'ombre que nous proposons dans cette étude et le détecteur de D. Williams [8]. . . . .	219
A.1	Les quatre orthants du plan fréquentiel. . . . .	229
B.1	(a) Espace euclidien 2D composé des vecteurs de base $\mathbf{e}_1$ et $\mathbf{e}_2$ ainsi que du pseudoscalaire $\mathbf{I}_2$ . (b) Espace Euclidien 3D composé des vecteurs de base $\mathbf{e}_1$ , $\mathbf{e}_2$ , $\mathbf{e}_3$ , de la base des bivecteurs $\mathbf{e}_{12}$ , $\mathbf{e}_{13}$ , $\mathbf{e}_{23}$ , et du pseudoscalaire $\mathbf{I}_3 = \mathbf{e}_{123}$ (extrait de [9]). . . . .	235

---

C.1 Représentation 3D du vecteur signal monogène dans une base de l'algèbre géométrique. . . . .	244
--	-----



# Liste des abréviations

ACP	Analyse en Composantes Principales
AGC	Automatic Gain Control
ATR	Automatic Target Recognition
AUV	Autonomous Underwater Vehicle
ciD	continuous intrinsic Dimension
CMRE	Centre of Maritime Research and Experimentation
DGA	Direction Générale de l'Armement
DoG	Difference of Gaussians
DoP	Difference of Poisson
DSTO	Defence Science and Technology Organisation
FFI	Forsvarets forskningsinstitut
FTC	Fine To Coarse
GLCM	Gray-Level Co-Occurrence Matrix
iD	intrinsic Dimension
LoG	Laplacian of Gaussians
MCM	Mine Counter Measure
MLO	Mine Like Object
MLT	Mine Like Texture
ROI	Region Of Interest
SAS	Synthetic Aperture Sonar
SIFT	Shift Invariant Feature Transform
SONAR	SOund Navigation And Ranging
TFAC	Taux de Fausses Alarmes Constant
TVG	Time Varying Gain



# Chapitre 1

## Introduction

### Sommaire

---

<b>1.1</b>	<b>Introduction générale</b>	<b>1</b>
<b>1.2</b>	<b>Contexte</b>	<b>3</b>
<b>1.3</b>	<b>Objectifs</b>	<b>4</b>
<b>1.4</b>	<b>Organisation</b>	<b>5</b>

---

### 1.1 Introduction générale

Les océans recouvrent la très grande majorité du globe terrestre et sont étudiés avec la plus grande attention pour de nombreuses raisons, allant de l'intérêt scientifique, écologique, ou économique, en passant naturellement par les intérêts stratégiques militaires. En particulier, les forces navales à travers le monde sont confrontées aux problématiques de sécurité des voies de navigation, mais également à la sûreté des espaces maritimes nationaux. La surveillance de ces étendues, qui représentent près de 80% de la surface terrestre, est un enjeu crucial qui mobilise de nombreuses ressources. Ces dernières peuvent se scinder en deux catégories, à savoir les ressources liées à l'observation de la surface océanique et les ressources liées à l'observation du milieu sous-marin. Comme son nom l'indique, la première catégorie permet de contrôler la surface de l'océan comme par exemple la présence, la trajectoire ou le comportement des navires. Dans ce cas, les systèmes utilisés peuvent être de plusieurs natures. En premier lieu, il y a l'observation humaine par des opérateurs dans des sémaphores sur le littoral, depuis des bâtiments de surface en mer ou à bord d'avions de patrouille maritime dans les airs (cf. Figure 1.1). Cependant, leurs faibles taux de couverture et leurs dépendances aux conditions climatiques (brouillard, nuages, mer agitée, vent, etc.) ont poussé l'utilisation généralisée des systèmes dits de **téledétection**. Ces derniers permettent de scanner à longue distance et rapidement une surface très vaste par le biais d'instruments de mesure variés tels que les radars



(a)



(b)

FIGURE 1.1: Exemple de moyens de contrôle de la surface océanique : (a) avion de type ATL2, (b) navire de patrouille maritime (extraits du site internet du ministère de la défense).

(de l'anglais *radio detection and ranging*), les lidars (*light detection and ranging*) ou encore les capteurs optiques. Leurs plateformes peuvent être installées directement sur le littoral, sur des aéronefs ou sur des satellites. Les données qu'ils fournissent sont ensuite analysées soit directement par un opérateur humain, soit par des systèmes automatiques spécialement conçus pour des tâches particulières (comme par exemple la détection de navires ou de polluants). Néanmoins, bien que de tels systèmes soient aujourd'hui opérationnels, ils ne permettent pas une observation complète des océans. En particulier, les ondes d'origine électromagnétique ou optique utilisées par les technologies mentionnées précédemment, ne pénètrent pas dans l'eau ou très peu et ne sont donc pas adaptées à l'observation du milieu sous-marin. Pourtant, les menaces stratégiques que représentent les sous-marins et la dangerosité des mines sous-marines dans les zones de navigation, font de l'analyse de ce milieu une problématique réelle et prioritaire pour les forces armées. Or, l'exploration de ce milieu directement par un être humain ne peut être envisagé du fait des profondeurs associées, des coûts en matériels adaptés mais surtout en termes de temps. En conséquence, il est nécessaire de s'équiper de systèmes propres à ce milieu. Les systèmes sonars (acronyme issu du terme anglais *sound navigation and ranging*) ont alors vu le jour et se basent sur l'étude de la propagation des ondes acoustiques dans le but d'acquérir des informations sur l'environnement sous-marin. Une large gamme de systèmes sonars existent aujourd'hui allant des sonars passifs, qui se contentent d'écouter les bruits ambiants du milieu, aux systèmes actifs proposant d'étudier les phénomènes de réflexion d'ondes sonores émises. Alors que la discrétion des systèmes passifs est adaptée à la lutte anti-sous-marine, les opérations de chasses aux mines utilisent quant à elles des systèmes actifs qui permettent de repérer, localiser voire d'identifier un objet immergé.

Parmi l'ensemble de ces missions d'observation et de contrôle des océans, ces travaux de thèse se focalisent exclusivement sur les problématiques liées à la chasse aux mines sous-marines. Dans ce contexte, les forces navales à travers le monde se sont équipées de navires particuliers, à faible signature électromagnétique, capables de détecter, d'identifier et de neutraliser ces engins explosifs (cf. Figure 1.2(a)). Pour permettre de telles missions, ces bâtiments de surface sont dotés de systèmes sonars spécifiques, dits "latéraux" permettant de couvrir une vaste

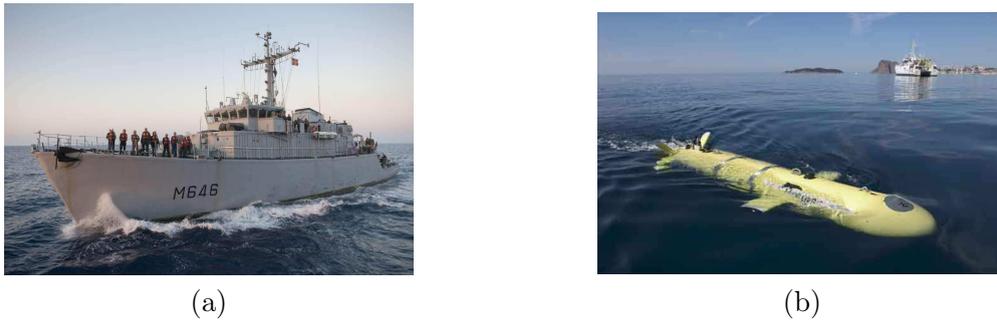


FIGURE 1.2: (a) chasseur de mine "La Croix du Sud" de la marine nationale française (extrait du site du ministère de la défense), (b) AUV proche de la surface (extrait du site internet de l'IFREMER).

zone de fonds marins rapidement et d'en construire une image détaillée. A bord du navire, un opérateur humain est spécialement entraîné pour analyser ces images et y détecter les objets potentiellement dangereux. Cependant, les progrès en matière de sonars font que l'opérateur doit aujourd'hui analyser des images sonar à haute résolution du fond marin. Ainsi, le temps nécessaire à l'étude de ces données augmente inévitablement et cet opérateur doit être au minimum assisté par ordinateur, voire remplacé.

Par ailleurs, les avancées en terme de robotique sous-marine permettent aujourd'hui de limiter encore plus l'intervention humaine dans le processus de chasse aux mines, en contrôlant à distance des véhicules non-habités, voire d'utiliser des robots autonomes. Ces derniers, que l'on appelle AUV (de l'anglais *Autonomous Underwater Vehicles*) sont des véhicules équipés de nombreux capteurs qui permettent leur propre navigation mais également la collecte de données concernant le milieu dans lequel ils évoluent (cf. Figure 1.2(b)). En particulier, dans notre contexte de chasse aux mines, ces AUV peuvent être équipés d'un sonar latéral pour scanner rapidement le fond marin et y détecter automatiquement les cibles via des algorithmes automatiques implémentés directement à bord.

Néanmoins, que le sonar latéral soit tracté par un navire de surface ou porté par un AUV, les problématiques liées au remplacement d'un opérateur humain par un ordinateur pour analyser les images restent les mêmes. De ce fait, dans ces travaux de thèse, nous nous intéressons à l'analyse automatique des images sonar sans *a priori* sur la manière dont elles ont été collectées.

## 1.2 Contexte

Rendre automatique l'analyse et la détection d'objets sur des images n'est pas un problème nouveau, ni spécifique au domaine de la chasse aux mines. En effet, depuis maintenant plusieurs décennies, différents domaines se sont attachés à cette problématique et particulièrement à la manière de détecter des objets ou des cibles dans des images issues de capteurs

variés. Nous pouvons citer par exemple la détection de visages et de piétons sur des images issues de caméras de surveillance, ou encore la détection de navires et d'avions sur des images radars ou optiques. Dans le domaine particulier de la défense, les algorithmes automatiques liés à la détection de cibles se regroupent généralement sous l'acronyme de chaînes ATR (de l'anglais *Automatic Target Recognition*). Ces algorithmes cherchent à détecter des cibles spécifiques selon le type de données traitées (radar, optique, laser ou sonar) telles que des bâtiments, des chars, des avions, des navires ou, dans ce qui nous intéresse particulièrement ici, des mines sous-marines. Dans ces travaux, nous allons alors étudier le concept de chaîne ATR dans le cas spécifique d'images issues de systèmes d'imagerie sonar.

Depuis maintenant plusieurs décennies, les chaînes ATR sont l'objet d'attentions particulières dans la communauté de la chasse aux mines sous-marines. L'enjeu derrière ces processus automatiques est de taille et peut être critique pour la sécurité. Or, actuellement, les performances de ces traitements n'atteignent pas encore des niveaux suffisants pour espérer être totalement opérationnels. Le principal frein reconnu par la communauté est lié à la variabilité de l'environnement marin proche d'une cible. Ce dernier, qui apparaît généralement sur l'image sonar du fait du relief et de la nature du fond marin, impacte directement l'onde acoustique rétro-diffusée mais également la signature des cibles posées sur le fond. Ainsi, une cible peut potentiellement se confondre avec son environnement proche, menant généralement à un échec pouvant être critique pour la phase de détection. De plus, cet environnement peut également se révéler propice aux fausses alarmes, dégradant fortement les performances d'un processus ATR. Néanmoins, ces phénomènes ne sont que rarement pris en compte dans le traitement. Parfois, l'environnement est tout de même caractérisé pour pouvoir déclarer une zone chassable ou non, voire pour anticiper certains défauts connus du processus de reconnaissance. L'approche classique consiste alors le plus souvent à le gommer ou à s'en affranchir par différents outils de traitement d'image. Parmi ces traitements, on trouve le filtrage ou la segmentation dans le but de ne conserver uniquement l'information dite utile pour la détection et la reconnaissance de la cible (à savoir les zones d'ombres et d'échos, elles-mêmes post-traitées parfois). Après des années d'efforts consacrés à la recherche d'une chaîne de traitements capable de s'affranchir de tous les artéfacts dus à l'environnement d'une cible, force est de constater qu'une chaîne universelle n'existe pas et que le problème doit être abordé autrement.

### 1.3 Objectifs

L'état de l'art sur les chaînes de reconnaissance automatique de cibles montre que leurs performances dépendent fortement des conditions environnementales. En effet, sur des fonds homogènes comme le sable, où une mine apparaîtra clairement comme un motif constitué d'un fort écho et d'une zone d'ombre facilement identifiable, la majorité des processus ATR actuels fourniront des taux de performance très élevés. Cependant, si le type de fond marin change,

alors ces performances se dégradent. En effet, les fonds marins qui présentent des rides de sable, une forte densité de cailloux ou de roches, ou de la végétation dense, entraînent une diminution significative de la probabilité de détection des cibles, ainsi qu'une augmentation de la probabilité de fausses alarmes.

Pour répondre à cette dégradation des performances, nous proposons de ne plus considérer l'environnement marin comme une contrainte qu'il faut éliminer, mais plutôt comme une source d'informations à part entière, qui doit être intégrée dans les processus ATR pour le rendre plus robuste et plus efficace. Dans cette optique, le premier objectif est d'analyser et de mieux comprendre dans quelle mesure l'environnement marin peut impacter les résultats d'une chaîne ATR. Le second objectif est ensuite de mettre en place une méthode de description du fond marin visant à mieux décrire l'environnement marin dans lequel se passe la mission de chasse aux mines. Enfin, le troisième objectif est d'utiliser cette description de l'environnement comme une source d'information pour améliorer les performances de la chaîne ATR. Il est alors nécessaire de trouver une manière d'intégrer ces descripteurs environnementaux dans le processus de reconnaissance automatique de cibles. Pour cela, nous pensons qu'il faut remettre en cause le séquençement de cette chaîne.

Plusieurs groupes de recherche travaillent actuellement sur cette idée d'introduire dans la chaîne ATR de l'information concernant l'environnement proche d'une cible. Nous pouvons citer par exemple l'université d'Heriot-Watt (Edinburgh - Écosse), le CMRE (*Center for Maritime Research and Experimentation* - La Spezia - Italie), ainsi que le FFI norvégien (Ministère de la Défense Norvégien). Cependant, jusqu'à présent, les travaux présentés consistent à quantifier les performances des algorithmes ATR existants en fonction de l'environnement. En d'autres termes, les algorithmes de détection ne s'appuient pas sur des caractérisations environnementales mais leurs résultats sont évalués lors d'un post-traitement qui les prends alors en compte. Dans ces travaux de thèse, nous considérons une autre stratégie qui propose d'intégrer cette connaissance environnementale dès le début de la chaîne ATR. De cette manière, cette étape initiale peut permettre de conditionner et d'adapter directement les phases de détection et de classification à l'environnement marin de la mission.

## 1.4 Organisation

La suite de ce manuscrit s'organise de la façon suivante. Dans le second chapitre nous introduirons les principes de base de l'imagerie acoustique. Ce chapitre permettra de mieux comprendre les images sonar que nous traiterons tout au long de ces travaux, et d'aborder leurs caractères complexes en terme de détection.

Le troisième chapitre est consacré à la problématique de détection et de classification de cibles en imagerie sonar. Nous y introduirons les différentes techniques utilisées jusqu'à

présent, ainsi que leurs limites. Nous y aborderons également les premières études proposant l'introduction d'informations environnementales dans le processus ATR. Ces dernières montrent que le type de fond marin impacte les performances d'une chaîne ATR, et proposent de définir et de qualifier les fonds en trois types, à savoir les fonds homogènes, anisotropes ou complexes. La position de nos travaux vis-à-vis de ces études sera alors explicitée.

Dans les chapitres 4 et 5, nous introduirons les outils mathématiques utilisés pour répondre à nos objectifs. Le chapitre 4 introduit le signal monogène qui permet une représentation différente de l'image en séparant de manière orthogonale les informations énergétiques, géométriques et structurelles qui y sont contenues, le tout dans un cadre multi-échelle. Nous verrons que cette dernière est efficace pour décrire les images sonar à travers notamment une application à la classification des fonds marins. Cependant, ce nouvel outil n'est qu'une représentation de l'information, il est donc nécessaire de trouver un moyen de l'exploiter et d'en extraire l'information recherchée pour décrire l'environnement marin. Pour cela, le chapitre 5 propose d'étudier le concept de dimension intrinsèque. Nous verrons que ce dernier permet de décrire de manière locale la texture d'une image et qu'il est relié à la notion d'homogénéité, d'anisotropie et de complexité de cette texture. De plus, la version dite "continue" de la dimension intrinsèque permet de fournir des indices de confiance sur ces différentes notions. Nous verrons que ces indices peuvent être estimés à partir des informations issues du signal monogène et par conséquent définis dans un cadre multi-échelle.

Le chapitre 6 est consacré à l'application du signal monogène et de la dimension intrinsèque pour décrire une image sonar sous la forme de cartes d'homogénéité, d'anisotropie et de complexité des fonds marins. Ces cartes sont définies comme des degrés de confiance vis à vis de ces trois descripteurs. Nous verrons que ces trois cartes sont alors représentatives des difficultés de détection automatique des mines sous-marines. De plus, nous montrerons que ces trois cartes permettent de réaliser une classification spécifique des fonds marins de manière non-supervisée, sans l'utilisation de classifieur. Cette classification en zones homogènes, anisotropes et complexes d'une image sonar permet de localiser précisément les zones de l'image où les performances de la détection pourraient être impactées. Enfin, nous proposerons de caractériser les zones homogènes et anisotropes de manière spécifique afin d'extraire des informations potentiellement utiles lors d'une phase de détection ou pour replanifier une mission.

Le chapitre 7 propose l'architecture d'une nouvelle chaîne ATR. Cette dernière place la description environnementale précédente en tout début de chaîne, afin qu'elle puisse s'y adapter automatiquement. Ensuite, nous aborderons le cas particulier de l'application de nos outils à une détection spécifique dans les zones de rides de sable (ou anisotrope).

Finalement, le chapitre 8 conclura ce manuscrit et proposera certaines perspectives à ces travaux de thèse.

# Chapitre 2

## Principes de l'imagerie sonar

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>8</b>
<b>2.2</b>	<b>Principes généraux des sonars de cartographie</b>	<b>9</b>
2.2.1	Émission et acquisition des signaux	10
2.2.1.1	Module d'émission	10
2.2.1.2	Module antenne	11
2.2.1.3	Module de réception	14
2.2.2	Étude et mise en forme des signaux en réception	15
2.2.2.1	Étude d'un ping sonar	15
2.2.2.2	Formation des images	16
<b>2.3</b>	<b>Imagerie acoustique des fonds marins</b>	<b>19</b>
2.3.1	Origine du speckle et modélisations statistiques associées	20
2.3.1.1	Généralités	20
2.3.1.2	Modélisation statistique de la réverbération : Loi de Rayleigh	21
2.3.1.3	Autres modélisations statistiques	23
2.3.2	Images texturées et complexes	24
2.3.3	Apparition d'objets	26
<b>2.4</b>	<b>Présentation des bases de données utilisées dans la thèse</b>	<b>28</b>
<b>2.5</b>	<b>Conclusion</b>	<b>30</b>

---

## 2.1 Introduction

Tout au long de ce manuscrit, nous traiterons des images issues d'un système sonar particulier appelé sonar latéral. Ce chapitre a pour vocation d'en présenter un aperçu, et de comprendre comment sont formées les images.

Tout d'abord, dans le milieu sous-marin, les fortes absorptions des ondes électromagnétiques limitent l'utilisation des technologies optiques traditionnelles de formation d'images. En effet, les portées de détection se révèlent trop faibles pour la majorité des applications souhaitées. Alors qu'un capteur optique permettra d'observer un objet jusqu'à quelques mètres, un capteur magnétique pourra quant-à-lui porter jusqu'à une dizaine de mètres [10, 11]. Ainsi, les ondes électromagnétiques, dont la lumière, se propagent difficilement dans ce milieu, interdisant les communications radios et la vision optique trop lointaine. En revanche, les ondes mécaniques, telles que les ondes acoustiques, subissent une faible atténuation dans une large bande de fréquences, entraînant leur utilisation généralisée dans le milieu marin.

L'étude de ce phénomène ondulatoire a contribué de manière significative à l'exploration de l'environnement sous-marin. Les premières expériences dans le domaine de l'acoustique sous-marine remontent au début du 19<sup>ème</sup> siècle par la détermination de la vitesse de propagation du son dans l'eau [12, 13]. C'est à l'aube de la première guerre mondiale que les premiers systèmes de détection d'objets à distance furent mis au point, et avec eux, l'apparition des premiers hydrophones piézoélectriques [14]. Ces systèmes acoustiques sous-marins sont spécifiquement désignés par le terme "sonar", acronyme de l'anglais *SOund Navigation And Ranging*. Depuis cette période, les progrès en matière d'électronique et de traitement d'antenne ont permis d'élargir les champs d'application et de démocratiser ces systèmes. Les traitements de données acoustiques permettent aujourd'hui d'accéder à des informations variées sur le milieu sous-marin, comme l'imagerie du fond, la structure des sols, la bathymétrie ou la bathycélérité.

Dans cette thèse, nous nous intéresserons particulièrement au système sonar latéral. Nous verrons que les ondes acoustiques émises par ce dernier, puis rétro-diffusées par les fonds, peuvent être représentées sous forme 2D, et donc d'images. De cette manière, l'interprétation des données acoustiques est facilitée, et permet de visualiser de manière précise les fonds marins, mais également les objets qui peuvent potentiellement y être posés. Les progrès technologiques en matière de sonar font que ces images ont aujourd'hui des résolutions de plus en plus fines et permettent d'avoir accès à toujours plus d'informations sur les fonds.

Dans ce chapitre, la section 2.2 commencera par présenter succinctement ce procédé d'imagerie, en partant de l'émission de l'onde sonore jusqu'à la formation de l'image sonar. La section 2.3 sera consacrée à la description des images sonar afin de voir ce qui y apparaît, comme les objets ou encore l'environnement marin. Nous aborderons notamment la complexité de l'étude de ces images. Puis la section 2.4 sera consacrée à la présentation des bases de

données que nous utiliserons tout au long de cette thèse, et sur lesquelles nous illustrerons le développement de nos outils.

## 2.2 Principes généraux des sonars de cartographie

Le sonar latéral est un système spécifique de cartographie destiné à obtenir une image détaillée de l'interface du fond marin. En particulier, il apporte une solution efficace à la détection d'objets de faibles dimensions posés sur le fond. De plus, le sonar latéral est un dispositif généralement léger et facile à mettre en œuvre.

L'architecture d'un sonar latéral se compose de deux antennes, chacune placée de part et d'autre d'un porteur (*i.e.* d'un poisson tracté par un bateau, d'un AUV etc.). Ces dernières sont généralement d'ouverture large dans le plan vertical pour couvrir un maximum de terrain, et étroite dans le plan horizontal afin d'obtenir une résolution fine. L'exemple d'un sonar latéral Klein 5000 est proposé sur la Figure 2.1 où l'on peut constater la forme spécifique de l'une de ses antennes (la bande noire horizontale couvrant la quasi totalité du flanc bâbord ici). La Figure 2.2 propose quant à elle d'illustrer la géométrie d'acquisition d'un sonar latéral lorsque ce dernier est placé sur un poisson. On y observe alors la forme des ouvertures verticales et horizontales des antennes bâbord et tribord.

Ces antennes sont utilisées en phase d'émission mais également en phase de réception. Elles sont connectées à plusieurs modules qui permettent, tout d'abord, de générer un signal électrique correspondant au signal que l'on désire propager dans le milieu sous-marin ; c'est le **module d'émission**. La transformation de ce signal électrique en une onde sonore qui se propagera ensuite dans la masse d'eau est assurée par le **module antenne** avec des caractéristiques de directivité spécifiques. Pendant que l'onde sonore se propage dans le milieu marin, le **module de réception** s'active et commence à enregistrer les ondes rétro-diffusées par les fonds marins. En effet, lorsque l'onde sonore émise se propage et rencontre l'interface (généralement



FIGURE 2.1: Exemple d'un sonar latéral Klein 5000 monté sur un poisson (documentation Klein).

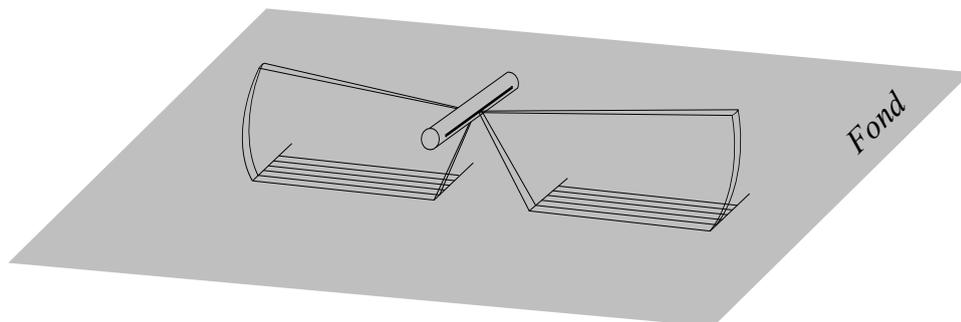


FIGURE 2.2: Géométrie d'acquisition d'un sonar latéral positionné sur un poisson tracté par un bateau.

rugueuse) du fond marin, cette dernière est diffusée dans toutes les directions, et en particulier, vers le sonar (*i.e.* rétro-diffusée). L'antenne reçoit donc cette onde, la transforme en signal électrique, puis le module de réception réalise les premiers traitements avant la numérisation des données.

Cette section est organisée de la manière suivante. Dans un premier temps, au paragraphe 2.2.1, nous allons nous attacher à décrire comment sont émis puis enregistrés les signaux acoustiques. Nous y verrons des explications plus approfondies sur les différents modules que nous venons de mentionner afin de mieux comprendre leurs influences sur le signal émis puis reçu. Ensuite, au paragraphe 2.2.2, nous analyserons plus précisément le contenu d'un signal reçu par l'antenne sonar, à savoir les phénomènes qui influencent son niveau énergétique comme la propagation à travers le canal ou l'interaction avec les obstacles (fonds, objets, ...). Nous verrons également qu'une mise sous forme 2D d'une succession de ces signaux en réception permet alors de créer une image du fond marin.

### 2.2.1 Émission et acquisition des signaux

Tout d'abord, nous allons ici nous intéresser aux différents modules qui permettent de générer, propager puis enregistrer un signal acoustique.

#### 2.2.1.1 Module d'émission

Ce module permet de générer une impulsion de durée très courte (de l'ordre de la microseconde) qui possède une fréquence comprise généralement entre 300 et 900 kHz pour les systèmes de cartographie [15]. La fréquence de répétition de ces impulsions est de l'ordre de quelques millisecondes. La modification de ces paramètres influe sur les performances du sonar. En effet, d'après l'équation du sonar, l'index de réflexion d'une cible dépend non-seulement de sa forme, de sa rugosité, de l'angle d'incidence de l'onde par rapport à elle, mais également de la fréquence utilisée [14]. Plus une fréquence sera élevée et plus l'on pourra visualiser des détails fins sur la zone insonifiée. En contrepartie, la portée de celle-ci sera diminuée puisque le

coefficient d'absorption augmente approximativement avec le carré de la fréquence [16]. Pour fixer un ordre de grandeur, à 100 kHz, cette absorption est d'environ 30 dB/km, soit une portée espérée supérieure à 100 m. Alors qu'à 1 MHz l'absorption sera de l'ordre de 500 dB/km soit une portée n'atteignant pas les 50 m.

Le niveau de puissance émis par ce module est également un élément important puisque l'intensité sonore est limitée par l'apparition de la cavitation. La pression d'apparition de la cavitation croît rapidement avec la fréquence selon R. Urick [14]. Ainsi, en haute fréquence, ces niveaux sont surtout limités par les contraintes mécaniques et thermiques à l'intérieur du transducteur.

Les oscillateurs, modulateurs et amplificateurs qui forment ce module sont ensuite connectés à l'antenne sonar qui se charge de transmettre l'onde générée au milieu sous-marin.

### 2.2.1.2 Module antenne

L'antenne d'un sonar latéral est formée de transducteurs piézo-électriques qui permettent de convertir un signal électrique en un signal acoustique qui se propagera dans le canal de transmission sous-marin, c'est le **mode émission**. A l'inverse, ces transducteurs convertiront également le signal sonore rétro-diffusé par les fonds marins en signal électrique lors du **mode réception**.

Les antennes les plus récentes sont généralement constituées par un regroupement de transducteurs élémentaires dont les signaux individuels sont combinés. Cependant, certaines antennes peuvent également être constituées d'un seul et même transducteur directif (de dimension caractéristique grande devant la longueur d'onde du signal).

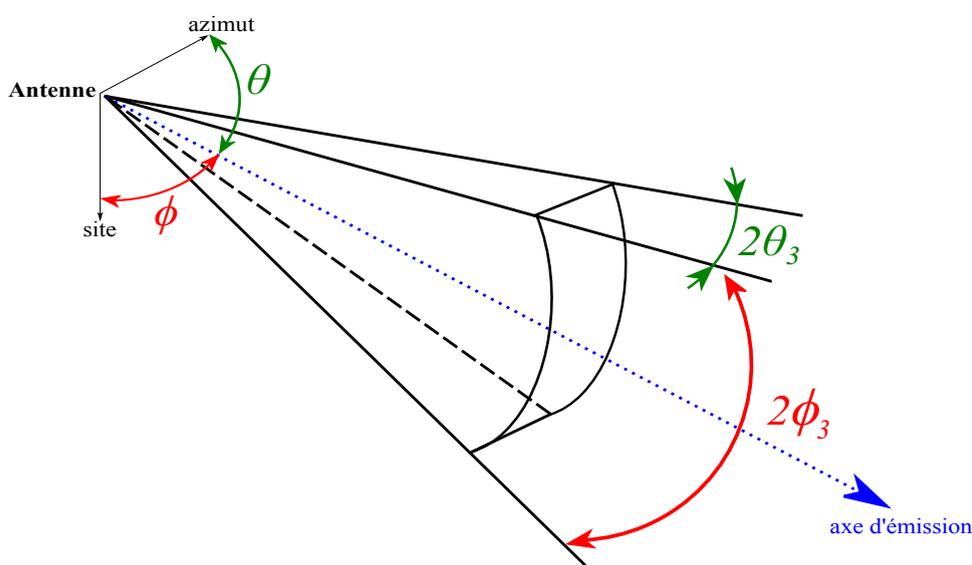


FIGURE 2.3: Géométrie d'une antenne considérée comme ponctuelle.

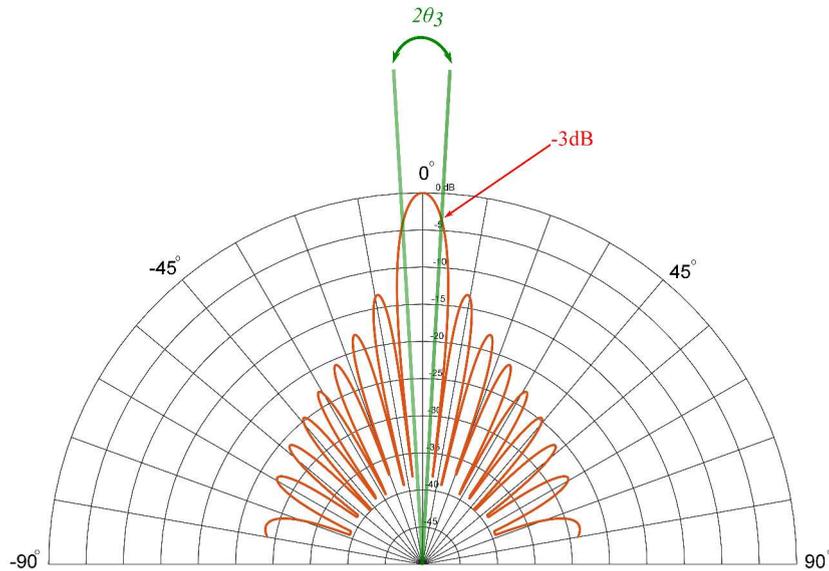


FIGURE 2.4: Exemple de diagramme de directivité en azimut en coordonnées polaires.

Une antenne se caractérise à l'aide de deux angles en azimut  $\theta$  et en site  $\phi$  que la Figure 2.3 propose d'illustrer pour une antenne considérée comme ponctuelle (à l'origine du repère). Alors que l'azimut est l'angle formé entre l'axe de propagation du signal et un axe horizontal de référence, le site est l'angle formé entre ce même axe de propagation et un axe de référence vertical. Les performances de l'antenne peuvent être représentées à l'aide d'un diagramme de directivité pour chacun de ces angles. L'exemple d'un tel diagramme est proposé sur la Figure 2.4 dans une représentation polaire (exprimée ici en décibels). On peut y observer un lobe principal important qui est centré sur la direction d'émission de l'antenne (ici 0 degré).

On remarque que si l'on dévie légèrement de la direction principale d'émission, la fonction de directivité ne s'atténue pas brutalement, mais progressivement. Ainsi, le faisceau principal d'énergie a une largeur angulaire finie. Celle-ci est traditionnellement définie par la largeur à  $-3$  dB, c'est-à-dire le secteur angulaire contigu dans lequel l'intensité est supérieure à la moitié de l'intensité maximale. Cette grandeur, notée ici  $2\theta_3$ , est appelée ouverture en azimut de l'antenne. Il existe également l'équivalent pour l'ouverture en site de l'antenne notée  $2\phi_3$ .

Par ailleurs, on peut noter que si l'on s'écarte toujours plus de la direction principale d'émission, et que l'on quitte le lobe principal, on obtient des maxima locaux équirépartis en angles, et de plus en plus faibles lorsque l'on s'écarte de la direction principale. Ces lobes sont appelés lobes secondaires. Sur cet exemple, les deux premiers lobes secondaires (ceux qui entourent le lobe principal) ont un niveau d'environ  $-13$  dB. Cette valeur est particulièrement importante ici mais reflète bien le cas général, ce qui fait que les lobes secondaires sont des éléments non-négligeables, qui peuvent très facilement polluer des mesures. Au-delà de ces

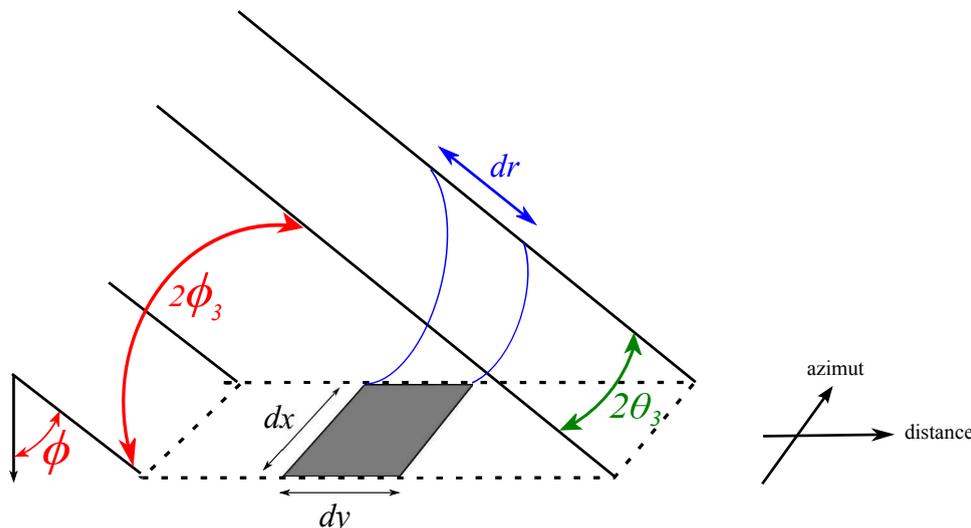


FIGURE 2.5: Cellule de résolution d'un sonar.

deux premiers lobes secondaires, les lobes suivants s'atténuent selon une loi en  $\frac{1}{\theta}$  dans le cas d'antennes sonar.

Le diagramme de directivité dépend de nombreux paramètres tels que la forme de l'antenne, du nombre et de la disposition des capteurs, et de la fréquence utilisée. Cette dernière modifie l'ouverture angulaire de l'antenne. Plus elle augmente, et plus l'ouverture sera étroite. Mais le niveau des lobes secondaires sera accentué par la même occasion. Ainsi, pour chaque application, il est nécessaire de réaliser un compromis entre l'ouverture d'antenne et le niveau des lobes secondaires.

L'une des caractéristiques importantes de l'antenne réside en son pouvoir de résolution azimutale. Ce dernier correspond à la distance minimale qui sépare deux objets situés sur l'axe des azimuts (cf. Figure 2.5) pour pouvoir être discernés sur le signal de retour. Cette résolution  $dx$  s'exprime à une distance  $r$  du sonar par la relation :

$$dx = 2r\theta_3. \quad (2.1)$$

Par la même occasion, nous pouvons également définir le pouvoir de résolution en distance du sonar. Cette fois c'est la distance minimale qui doit séparer deux objets situés sur l'axe de propagation pour pouvoir les discernés sur le signal de retour. Cette résolution  $dr$  s'exprime quant à elle dans le plan oblique (cf. Figure 2.5) par la relation :

$$dr = \frac{c}{2B}, \quad (2.2)$$

où  $B$  désigne la bande passante du signal émis et  $c$  la célérité du son dans l'eau (environ 1500 m/s). La cellule de résolution du sonar est donc définie comme la surface d'aire  $dx \times dr$ . Lorsque l'on traite une image issue d'un tel dispositif, c'est cette cellule de résolution qui est prise en

compte lors du dimensionnement d'un pixel dans la future image sonar. Cependant, si l'on veut connaître la surface insonifiée du fond correspondante, on doit projeter cette quantité  $dr$  sur le fond, on définit alors la résolution dite *transversale* du sonar :

$$dy = \frac{c}{2B \sin(\phi)}. \quad (2.3)$$

Puisque les zones du fond proche de l'antenne sont à la verticale sous le sonar, alors qu'en portée maximale l'incidence tend vers  $\pi/2$ , nous avons une résolution transversale plus fine à grande distance qu'à proche distance.

**Remarque importante :** dans le cas où l'onde émise est monochromatique, il existe une relation entre la bande passante  $B$  et la durée d'impulsion  $T$  par la relation  $BT = 1$ , ainsi les résolutions  $dr$  et  $dy$  dépendent de la durée d'impulsion. Cependant, la majorité des systèmes sonars actuels utilisent une onde modulée linéairement en fréquence (*i.e.* un chirp) et nous avons alors  $BT > 1$ . Dans ce cas, un traitement cohérent est réalisé dans le module de réception (dont nous parlons au prochain paragraphe) et permet de rendre les quantités  $dr$  et  $dy$  indépendantes du temps d'impulsion  $T$ .

Une fois que le module antenne a émis l'onde acoustique, cette dernière se propage dans le canal de transmission sous-marin et le mode réception de l'antenne s'active. Cette fois, les transducteurs piézo-électriques de l'antenne permettent de transformer l'onde sonore rétro-diffusée par les fonds marins en un signal électrique. Il est à noter que les propriétés de directivité de l'antenne en réception restent inchangées par rapport à l'émission.

### 2.2.1.3 Module de réception

Une fois que les transducteurs de l'antenne ont capté le signal rétro-diffusé, une étape d'amplification est alors nécessaire. Elle permet de conserver un niveau de signal reçu relativement constant en fonction de la portée. En effet, du fait des pertes par propagation (absorption et divergence géométrique), l'onde acoustique s'atténue au cours des trajets aller et retour dans le milieu marin. De plus, un ensemble de phénomènes de bruitage vient perturber le signal tel que le bruit ambiant de la mer, les bruits de réverbération, mais également les bruits propres à la chaîne de traitement électronique.

Avant de convertir le signal sous une forme numérique, beaucoup de sonar modernes possèdent un gain variable en réception (soit une TVG, de l'anglais *Time Varying Gain*, qui suit une loi déterministe, soit un AGC, de l'anglais *Automatic Gain Control*, qui est un asservissement du niveau du signal). Le but de cet étage est de conditionner au mieux les signaux pour la numérisation malgré les fortes dynamiques couramment rencontrées.

Le signal est ensuite numérisé afin d'être traité numériquement pour une mise en forme (image, distance de détection, etc...). Il est à noter que c'est en fonction de cet échantillonnage

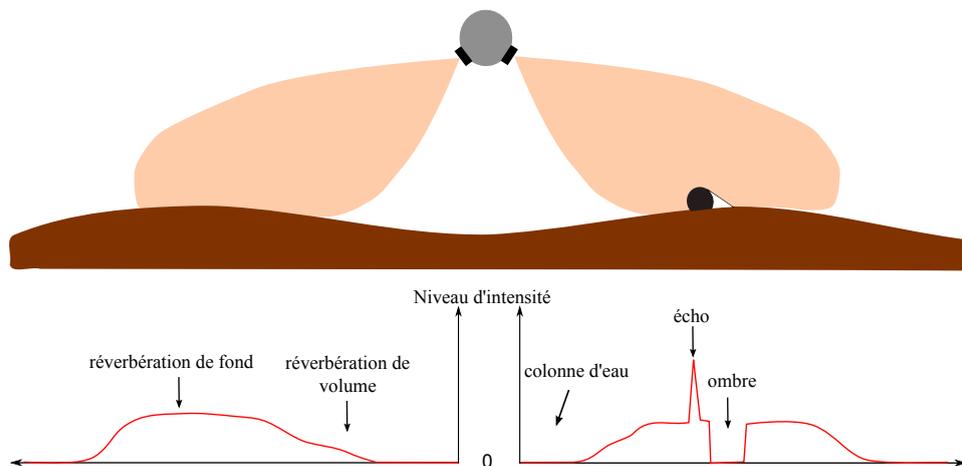


FIGURE 2.6: Géométrie d'acquisition d'une voie de sonar latéral et le signal reçu correspondant.

(qui peut être temporel et spatial) que la dimension des pixels d'une image sonar est définie. Cette dimension (en mètres) est fixée pour être inférieure à celle de la cellule de résolution du sonar afin d'éviter une perte d'information.

Ces différentes étapes de la réception d'un signal entraînent un nombre important de modifications du signal brute, qui peuvent mener à une dégradation de l'information portée initialement. La modélisation de ces phénomènes à l'intérieur des systèmes sonar reste un problème très complexes.

## 2.2.2 Étude et mise en forme des signaux en réception

Dans cette section, nous allons tout d'abord nous intéresser à l'étude d'une seule impulsion afin de mieux comprendre le signal reçu par une antenne sonar, c'est-à-dire un **ping sonar**. Ensuite, nous verrons que le mouvement du porteur, couplé à une succession d'impulsions, permettra de former une image.

### 2.2.2.1 Étude d'un ping sonar

Pour une antenne et une impulsion donnée, le signal sonore va se propager perpendiculairement à cette dernière et intersecter le fond marin, et ce, depuis la quasi-verticale sous le sonar jusqu'à de grandes distances de part et d'autre du poisson comme le montre la Figure 2.6. L'onde sonore va alors interagir avec le fond marin puis être diffusée dans toutes les directions et, en particulier, rétro-diffusée en direction de l'antenne sonar. Cette dernière la capte par ses éléments piézo-électriques. Ces derniers mesurent les variations de pression incidente en les transformant en signaux électriques (mode réception). C'est ensuite l'amplitude complexe de ces signaux qui est étudiée.

Sur ce signal, le niveau énergétique dépend de la nature et de la géométrie de l'obstacle sur lequel s'est rétro-diffusée l'onde sonore, c'est **l'écho**. La position d'un écho sur ce signal de réception dépend de l'éloignement de l'obstacle en question par rapport au sonar. En effet, plus la distance est importante et plus le temps de propagation sera long.

La Figure 2.6 propose l'exemple schématique d'un signal de réception pour une configuration donnée du fond marin. On y observe que le signal reçu est d'abord quasi-nul, ce qui correspond au fait que l'onde sonore n'a pas encore eu le temps de toucher un obstacle. En fait, cette durée correspond au temps de propagation dans la colonne d'eau, c'est-à-dire au temps que l'onde met pour toucher le fond marin. Une fois que l'onde a touché le fond, l'amplitude des signaux en réception augmente. Dans le cas où l'onde touche un objet posé sur le fond, l'écho est alors plus intense, et apparaît comme un pic très fort dans le signal retour (dû à une réflexion spéculaire de l'onde sonore). Du fait de la géométrie d'acquisition, et de la propagation sphérique des ondes acoustiques (en champ lointain), une zone ne pourra pas être insonifiée derrière l'objet, c'est une **zone d'ombre**. Globalement, on observe que l'amplitude du signal reçu diminue avec la portée. Ce phénomène est principalement dû aux pertes par propagation et absorption de l'onde acoustique, mais également à la décroissance de l'énergie rétro-diffusée en fonction de l'angle d'arrivée de l'onde sur le fond. Une modélisation de ce dernier a été faite par Jackson comme le décrit par exemple X. Lurton dans [16].

Après un traitement de ce signal par le module de réception, on obtient une première information concernant un segment étroit du fond sous la forme d'un signal 1D (un ping). Notons que pour chaque émission, nous avons en réception ce signal 1D pour chacune des deux antennes, *i.e.* pour le côté bâbord et tribord.

### 2.2.2.2 Formation des images

Nous venons de voir qu'une émission du sonar permet de scanner un segment étroit du fond marin dans le sens d'avancée du porteur, et long dans sa direction perpendiculaire (selon la portée du système).

Maintenant, si l'on considère que le porteur avance, on peut supposer que les impulsions suivantes permettront d'insonifier des segments voisins, et donc de réaliser un balayage du fond de la mer. En concaténant toutes ces segments successifs, on peut les représenter sous une forme 2D, correspondant à une image. Les variations d'intensité donnent à cette image l'illusion d'une vue de dessus du fond marin, sous un éclairage rasant.

Traditionnellement, les images bâbord et tribord sont représentées côte-à-côte, où la trajectoire du sonar doit être imaginée au centre de l'image, comme le montre la Figure 2.7. Cette dernière reprend l'exemple de la géométrie d'acquisition du sonar latéral de la Figure 2.2 en la complétant par l'image du fond marin ainsi formée. On y constate l'apparition d'une

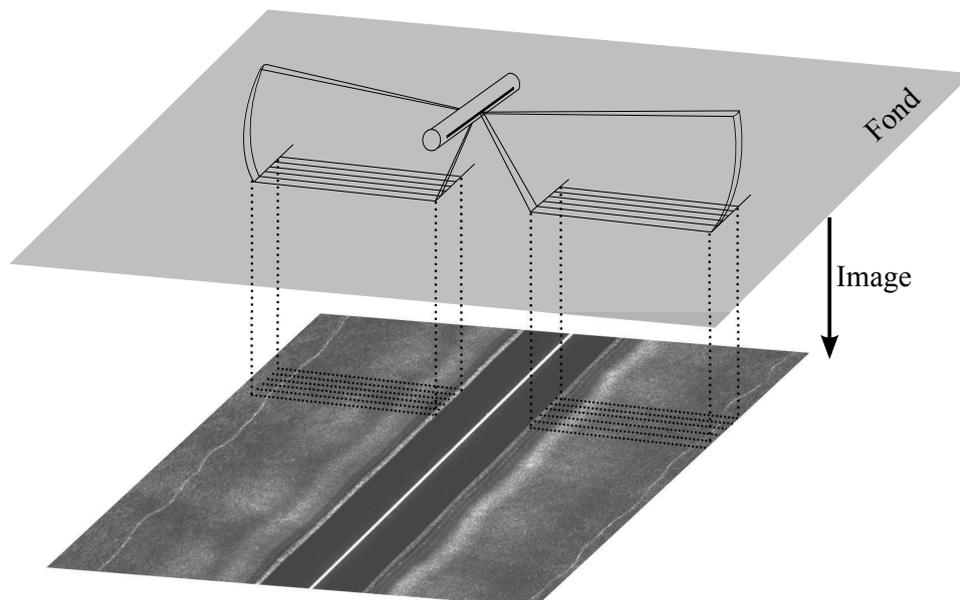


FIGURE 2.7: Géométrie d'acquisition d'un sonar latéral positionné sur un poisson tracté par un bateau. L'image sonar résultante est représentée juste en dessous.

bande noire au centre de l'image qui correspond au temps de propagation de l'onde acoustique à travers la colonne d'eau.

Une image typique du sonar latéral est illustrée plus précisément sur la Figure 2.8. Ses caractéristiques les plus notables y sont présentées. L'axe horizontal correspond à la ligne insonifiée par le sonar, ou à la distance perpendiculaire à la trajectoire du sonar (*across-track* en anglais). L'axe vertical correspond à la direction de mouvement du porteur du sonar (*along-track* en anglais). On trouve sur cette image la bande centrale noire correspondant à la colonne d'eau, dont la taille dépendant directement de la hauteur d'eau sous le sonar et de l'angle d'inclinaison de l'antenne. Les lobes secondaires de cette antenne apparaissent également sur l'image par la création de bandes verticales d'intensités variables.

On peut également observer parfois ce qu'on appelle un retour de la surface. En effet, suivant la conception du sonar, il se peut que certains artéfacts amènent la création d'une onde sonore à la verticale du sonar en direction de la surface. La réflexion de cette dernière sur la surface et son retour sur le sonar crée un liseré très intense comme on peut le voir sur la Figure 2.8. Ce phénomène est surtout visible sur les très petits fonds quand la hauteur d'eau au-dessus du sonar est faible. On remarque également sur cette même figure que la navigation du porteur impacte la qualité des données puisqu'ici ce dernier réalise un virage pendant l'acquisition. L'accès aux données de navigation est donc crucial si l'on veut éviter toute confusion lors d'une analyse automatique de telles images.

L'évolution des technologies en traitements d'antennes permet aujourd'hui de réaliser des images du fond marin avec de moins en moins d'artéfacts et des résolutions toujours plus fines. Par exemple, nous avons pu voir dans le paragraphe 2.2.1.2 que les antennes sonar peuvent être

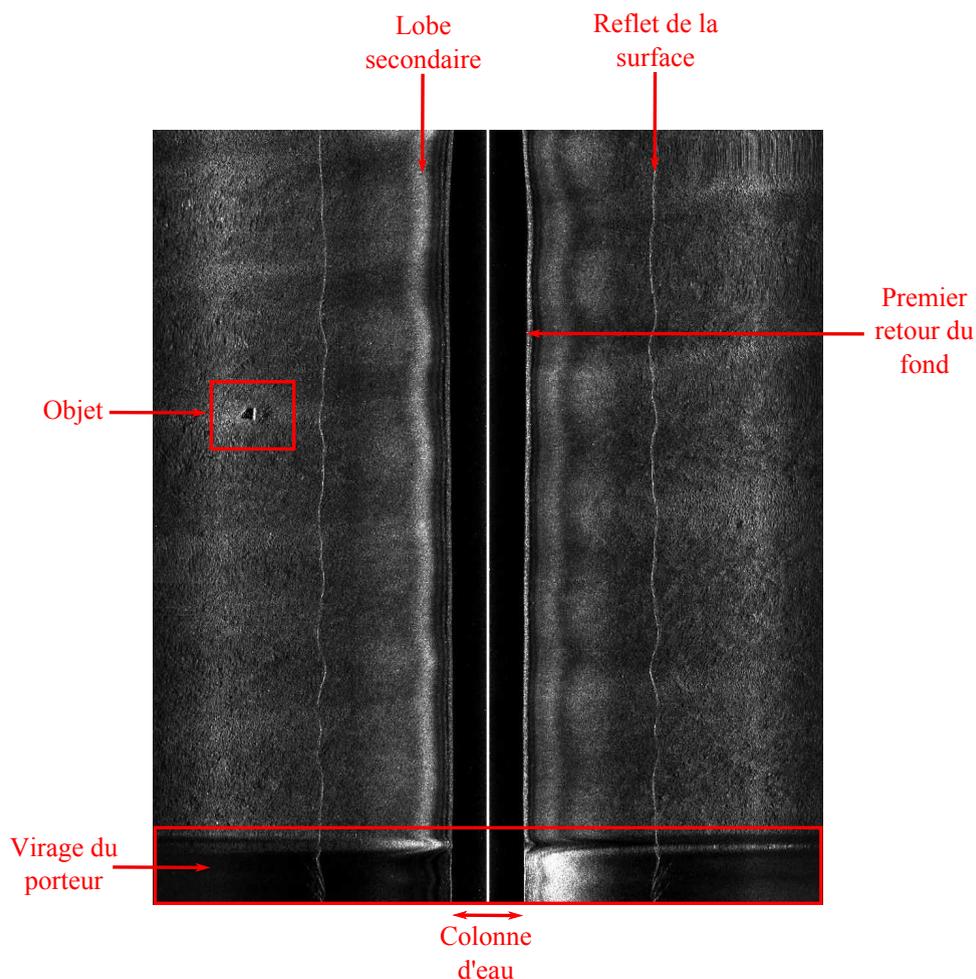


FIGURE 2.8: Exemple d'une image acoustique de sonar latéral avec certaines de ses caractéristiques.

réalisées en réseaux de transducteurs plutôt que par l'utilisation d'un transducteur unique. Ce procédé permet par exemple d'améliorer la directivité de l'antenne par le biais d'une **formation de voies**, et donc de diminuer les phénomènes de bruit venant des directions non choisies. Par ailleurs, ces réseaux de transducteurs permettent également une intégration cohérente des signaux des voies d'une antenne réelle sur un nombre  $N$  fixé de récurrences successives constituant alors la possibilité de créer une antenne synthétique. C'est le principe de base des images sonar dites SAS (de l'anglais *Synthetic Aperture Sonar*). L'avantage est de simuler une antenne d'une grande longueur, et par conséquent, d'augmenter la résolution des images sonar.

L'objectif de ces travaux de thèse n'ayant pas pour but l'amélioration des technologies d'imagerie sonar, nous invitons le lecteur à se référer aux ouvrages de P. Blondel et de X. Lurton [16, 17] pour toutes précisions supplémentaires sur ces différents systèmes sonars et traitements de l'onde acoustique.

## 2.3 Imagerie acoustique des fonds marins

Dans la section précédente, nous nous sommes attachés à décrire la manière dont l'image sonar est formée. Dans cette section, nous allons maintenant chercher à comprendre ce qu'on peut retrouver dans ces images, et quels types d'informations y sont contenues.

La Figure 2.9 propose tout d'abord plusieurs exemples d'images réelles issues de différents systèmes sonars latéraux. On remarque que cette représentation en niveaux de gris permet une interprétation rapide de l'environnement sous-marin. Par l'intermédiaire de différents contrastes, on peut y reconnaître certains types de fonds marins particuliers tels que du sable fin ou de la vase. De plus, l'arrangement des pixels dans ces images forme différentes textures qui permettent d'apercevoir d'autres types de fonds, comme par exemple des textures

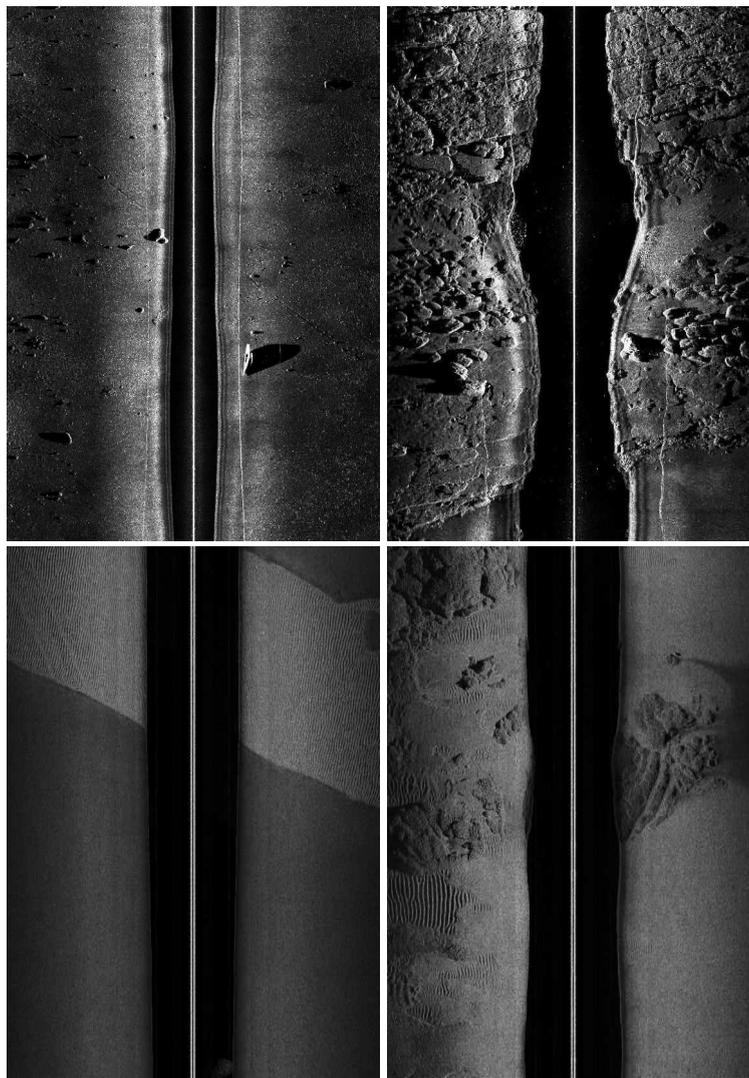


FIGURE 2.9: Mosaïque d'images acoustiques provenant de différents systèmes sonars. En haut, un sonar latéral Marine Sonic Technology Ltd de fréquence 900 kHz. En bas, un sonar latéral Klein 5000 de fréquence 455 kHz.

ondulées typiques de la présence de rides de sable. Il est également possible d'observer des structures à différentes échelles, comme par exemple des roches ou des objets posés sur les fonds, aux tailles et formes diverses.

Ainsi, il est possible de considérer plusieurs manières pour caractériser l'information contenue à l'intérieur d'une image sonar. Une première méthode serait de caractériser l'image au niveau de l'intensité des pixels par une approche statistique. En effet, si l'on regarde de plus près ces images, les détails fins qu'elles contiennent, montrent un aspect granulaire. Cela provient de la constitution physique des fonds marins à l'intérieur des cellules de résolution du sonar. Il serait alors intéressant de connaître la distribution statistique de ces ondes réfléchies si l'on souhaite une interprétation fiable de ces données. C'est ce que nous nous attacherons à décrire dans le paragraphe 2.3.1, en introduisant la notion de speckle. Une autre manière de caractériser ces images serait de ne plus s'intéresser uniquement à l'intensité du pixel, mais à l'arrangement spatial de ces intensités, c'est la notion de texture que nous verrons dans le paragraphe 2.3.3. Ensuite, puisque ces travaux concernent la problématique de la détection d'objets, le paragraphe 2.3.3 aura pour objectif de décrire comment les objets apparaissent sur ces images sonar, et comment on peut alors les caractériser.

### 2.3.1 Origine du speckle et modélisations statistiques associées

L'observation des détails ou des contours présents dans les images sonar montrent un aspect granulaire, comme nous pouvons le constater sur la Figure 2.10. On observe dans les deux blocs proposés une variabilité du niveau d'intensité des pixels. Ce phénomène connu sous le nom de *speckle*, ou chatolement, a fait l'objet de nombreux travaux au sein des applications nécessitant des systèmes d'émission-réception cohérents. Nous pouvons citer les travaux [18, 19, 20, 21] en imagerie laser, [22, 23, 24, 25] en imagerie radar SAR (Synthetic Aperture Radar), [26, 27, 28, 29, 30] en imagerie médicale, encore [31, 32, 33] en imagerie sonar.

#### 2.3.1.1 Généralités

Le bruit de speckle apparaît lorsqu'une onde électromagnétique ou acoustique interagit avec une surface rugueuse constituée par un grand nombre de diffuseurs dont les dimensions sont petites devant la longueur d'onde incidente. Chaque diffuseur se comporte alors comme une source émettant une onde, de même longueur d'onde, mais dont l'amplitude, la phase et la direction d'émission sont quant à elles aléatoires.

Le module de réception enregistre alors un signal qui est le résultat de l'interférence de toutes ces ondes réfléchies dans une même cellule de résolution. Ainsi, un tel phénomène ne peut être modélisé d'une autre manière que de façon statistique.

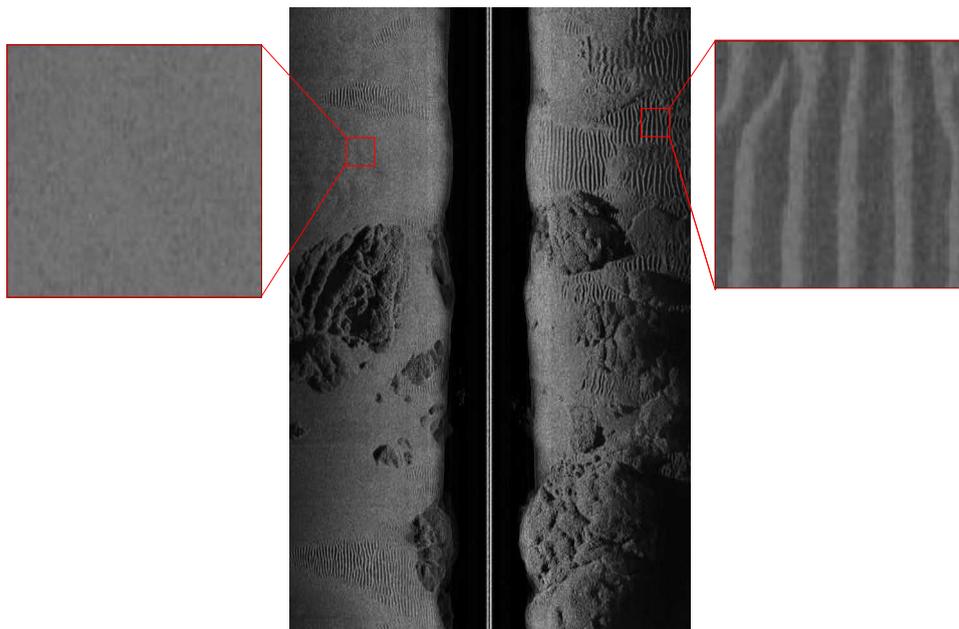


FIGURE 2.10: Illustration du speckle pour deux blocs rectangulaires d'une image sonar Klein 5000.

### 2.3.1.2 Modélisation statistique de la réverbération : Loi de Rayleigh

Avant de définir les propriétés de ce bruit, il est nécessaire de connaître la distribution statistique des ondes réfléchies sur le fond marin à l'intérieur d'une même cellule de résolution.

#### Rappels sur la distribution d'une onde rétro-diffusée

Considérons qu'une source  $k$  émette une onde complexe  $A_k e^{j\phi_k}$ , d'amplitude  $A_k$  et de phase  $\phi_k$ . Alors si l'on suppose que  $E$  est l'amplitude de l'onde reçue par un capteur, elle s'écrit comme la somme des contributions de chacune des  $n$  sources présentes dans une même cellule de résolution :

$$E = E_r + jE_i \quad \text{avec} \quad E_r = \sum_{k=1}^n A_k \cos(\phi_k) \quad \text{et} \quad E_i = \sum_{k=1}^n A_k \sin(\phi_k), \quad (2.4)$$

où  $j^2 = -1$  désigne l'unité complexe.

Certaines hypothèses sont alors couramment formulées sur les diffuseurs à l'intérieur d'une cellule de résolution. Elles sont généralement au nombre de trois, et portent sur le nombre, la répartition et les dimensions de ces diffuseurs :

**Hypothèse 1 :** Dans une même cellule de résolution, le fond marin est constitué d'un grand nombre  $n$  de diffuseurs indépendants et répartis de manière aléatoire.

**Hypothèse 2 :** Dans une même cellule de résolution, si le fond marin est constitué de zones homogènes alors les conditions de réflectivité des ondes sont considérées comme uniformes.

**Hypothèse 3 :** La distance qui sépare les diffuseurs d'une cellule de résolution et l'antenne sonar doit être très supérieure à la longueur d'onde rétrodiffusée.

Sauf cas particuliers, d'après les fréquences généralement utilisées par les systèmes sonar, les dimensions des cellules de résolutions sont de l'ordre de quelques centimètres carrés, ou tout au plus, de quelques dizaines de centimètres carrés. Ce qui permet de vérifier la première et la troisième hypothèse. De plus, les études réalisées sur différentes rugosités des fonds marins permettent d'affirmer que l'hypothèse 2 se vérifie également [34, 35, 36].

### La distribution de Rayleigh

Les hypothèses précédentes permettent de considérer que les amplitudes  $A_k$  des différentes sources peuvent être modélisées par des variables indépendantes, de même moyenne et de même variance grâce à l'uniformité locale à l'intérieur de la cellule de résolution. De plus, la phase  $\phi_k$  est proportionnelle au rapport entre la distance sonar-diffuseur  $r_k$  et la longueur d'onde rétrodiffusée :

$$\phi_k = 2\pi \frac{r_k}{\lambda}. \quad (2.5)$$

Ainsi, puisque les diffuseurs sont indépendants, et qu'il existe une grande variation des distances diffuseurs-sonar, alors il est possible de modéliser les phases  $\phi_k$  par des variables aléatoires indépendantes et uniformes sur l'intervalle  $[-\pi, \pi]$ .

Puisque  $A_k$  et  $\phi_k$  sont deux variables aléatoires et indépendantes, qui respectent une loi de distribution identique, il est possible d'appliquer le théorème central limite au couple  $(E_r, E_i)$ . La densité conjointe de  $E_r$  et de  $E_i$  est alors une loi gaussienne circulaire (de moyenne nulle et de variance  $\sigma^2$ ) :

$$f_{E_r, E_j}(u, v) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-(u^2 + v^2)}{2\sigma^2}\right). \quad (2.6)$$

L'amplitude du signal reçu par le capteur  $A = \sqrt{E_r^2 + E_j^2}$  suit alors une loi de Rayleigh  $\mathcal{R}_A$  définie par :

$$\mathcal{R}_A(x) = \frac{x}{\alpha^2} \exp\left(\frac{-x^2}{2\alpha^2}\right). \quad (2.7)$$

Le paramètre  $\alpha > 0$  caractérise la loi de Rayleigh (voir Figure 2.11) en terme d'étalement. Les moments d'ordre 1 et 2 de cette loi s'expriment en fonction de ce paramètre par les expressions suivantes :

$$m_A = \mathbb{E}[x] = \alpha \sqrt{\frac{\pi}{2}} \quad \text{et} \quad \sigma_A^2 = \alpha^2 \frac{(4 - \pi)}{2}. \quad (2.8)$$

où  $\mathbb{E}[\cdot]$  est l'espérance mathématique.

Cette amplitude admet donc un caractère aléatoire qui donne un aspect granulaire aux zones de réverbération du fond. On peut noter que la moyenne  $m_A$  et l'écart-type  $\sigma_A$  de la loi de Rayleigh sont proportionnels. M. Tur montre dans [37] que c'est la caractéristique d'un

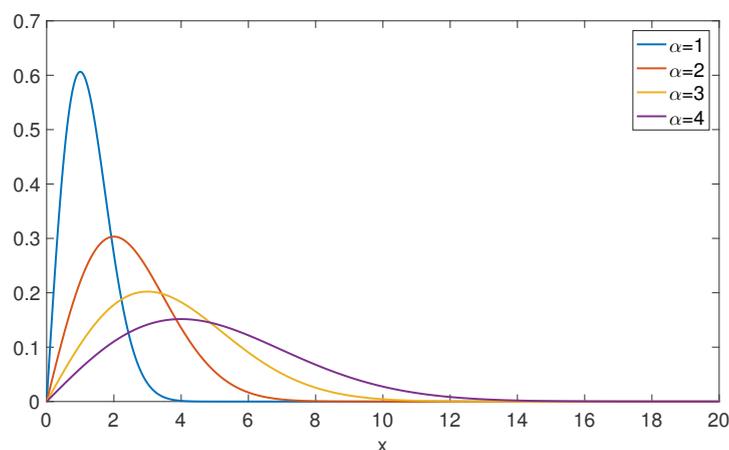


FIGURE 2.11: Représentation de la loi de Rayleigh pour différentes valeurs de paramètres.

bruit multiplicatif. Ces deux valeurs varient en fonction de la nature des fonds marins au niveau macroscopique (rugosité ou taille du sédiment qui recouvre le fond) mais également en fonction des caractéristiques du signal émis [38].

Ainsi, dans le cas du sonar, et lorsque les hypothèses précédentes sont vérifiées, on peut supposer qu'une loi de Rayleigh permet de modéliser le lien entre l'interface du fond marin, les caractéristiques techniques du sonar et la distribution des niveaux de gris d'une image acoustique.

### 2.3.1.3 Autres modélisations statistiques

Cependant, cette modélisation de la rétro-diffusion du fond par une loi de Rayleigh n'est pas toujours réaliste. Bien qu'elle corresponde à de nombreuses situations en imagerie acoustique, sur certains types de fonds comme les roches ou encore les rides de sables de grandes dimensions, d'autres modèles doivent être envisagés. En effet dans ce cas, les dimensions des rugosités deviennent alors du même ordre de grandeur que les cellules de résolution. Les hypothèses exprimées précédemment ne sont alors plus vérifiées. D'autres modélisations doivent être envisagées.

En effet, nous avons supposé précédemment que le nombre de diffuseurs par cellule de résolution est supposé très grand et implicitement constant d'une cellule à une autre. C'est l'hypothèse d'homogénéité locale des zones de fond marin. Cependant, lorsque la taille des grains, qui constituent le fond, devient du même ordre de grandeur que la longueur d'onde incidente, alors le nombre de diffuseurs peut évoluer d'une cellule de résolution à une autre. Différentes modélisations ont alors été proposées pour ce phénomène [39]. La loi K en est une, et propose une modélisation qui se base sur une justification physique du phénomène. Elle

exprime la distribution de l'amplitude du champ rétro-diffusé par la relation suivante :

$$f_K(x) = \frac{2b}{\Gamma(\alpha)} \left[ \frac{bx}{2} \right]^\alpha K_{\alpha-1}(bx), \quad (2.9)$$

où

- $K_{\alpha-1}(\cdot)$  est la fonction de Bessel modifiée de seconde espèce d'ordre  $\alpha - 1$  ;
- $\Gamma(\alpha)$  est la loi gamma de paramètre  $\alpha$  ;
- $b = 2\sqrt{\frac{1}{\beta E[A_k^2]}}$  avec  $\beta$  le paramètre d'échelle de la loi K.

D'autres lois proposent quant à elle une modélisation purement mathématique du phénomène, telles que la loi de Weibull, ou encore la loi log-normale. Elles permettent de mieux modéliser les distributions à queues lourdes, mais également d'obtenir une plus grande souplesse d'analyse. En effet, la loi de Weibull permet dans certains cas de pouvoir considérer les lois exponentielles ou encore la loi de Rayleigh comme de simples cas particuliers.

Ainsi, nous disposons à présent de modélisations possibles de la distribution des niveaux de gris qui forment les images sonar. L'utilisation de ces distributions statistiques ont permis par exemple des applications telles que la classification des fonds marins grâce à leur simplicité de mise en œuvre. Cependant, puisqu'elles se limitent uniquement à une analyse de l'image au niveau de l'intensité du pixel, les méthodes statistiques se heurtent à des difficultés lors de l'apparition de textures dans les images [39]. En effet, la valeur d'un pixel n'est alors pas assez informative pour pouvoir extraire toute l'information contenue dans l'image. Il est donc nécessaire de considérer l'information contenue dans les textures.

### 2.3.2 Images texturées et complexes

Comme nous l'avons déjà remarqué précédemment, les images sonar font apparaître des textures qui peuvent être informatives sur le type de fond marin présent dans l'image. Lorsque l'on observe une image sonar, notre oeil permet de les distinguer rapidement. Cependant, il est extrêmement difficile d'en donner une définition puisque qu'aucun consensus n'a été trouvé au sein de la communauté du traitement des images. Beaucoup de définitions existent donc comme le montre H. Chen dans [40]. Nous considérerons ici qu'une texture est une région de l'image d'intensité variable formant des motifs répétés avec un certain degré d'aléatoire. Ces motifs puisent leur origine des propriétés physiques de la surface du fond marin (*i.e.* des rugosités, de la bathymétrie, etc.) et des propriétés hydrodynamiques (*i.e.* les courants) locales du milieu sous-marin.

Ainsi, les textures que l'on observe sur ces images sont reliées à la physique des fonds marins insonifié par le sonar. Ces fonds marins constituent ce qu'on appellera dans cette thèse **l'environnement marin**. Ces derniers sont très variables en fonction des conditions météorologiques, des saisons, ou encore des lieux géographiques. Les sédiments qui les forment

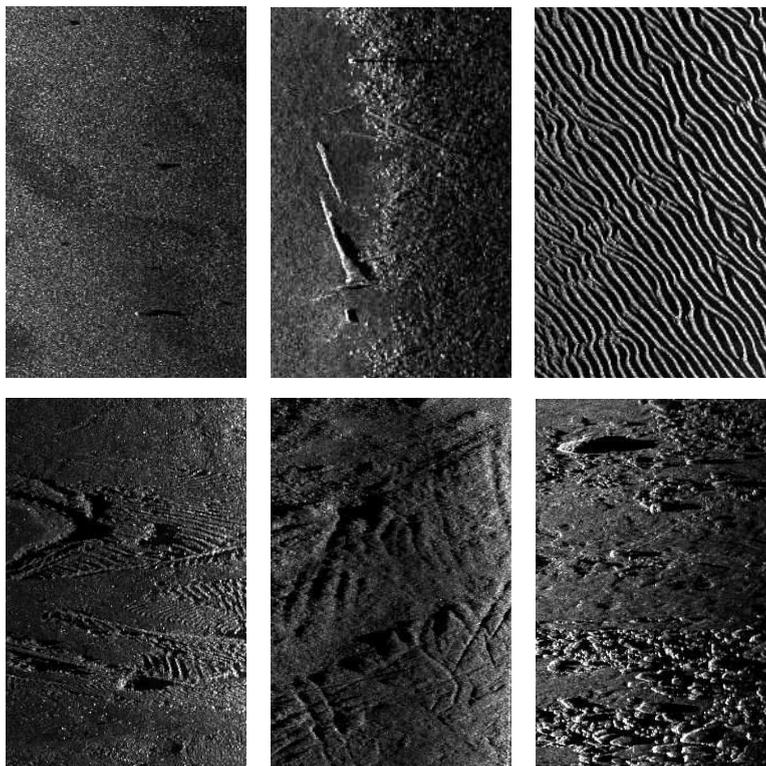


FIGURE 2.12: Mosaïque représentant différents types de fonds marins. Sur ces différentes images, nous pouvons apercevoir différentes structures sous-marines comme des zones de sable fin, des rides de sables, des roches, des traces de dragage, ou encore des objets posés sur le fonds.

n'interagissent pas tous de la même manière ni avec les phénomènes océaniques, ni même avec les ondes acoustiques. Ainsi, leur composition matérielle, leur rugosité ou leur bathymétrie influent l'onde acoustique rétro-diffusée. Leur apparence sur les images sonar est alors variée et bien souvent difficiles à exploiter pour prédire leur type.

La Figure 2.12 propose d'illustrer ces propos à l'aide d'une mosaïque composées de différents fonds marins extraits d'images sonar latéral. On peut y observer que les fonds marins apparaissent comme des textures pouvant être très variées. Ils peuvent alors aller de textures purement stochastiques (zones de l'image qui apparaissent comme composées uniquement de bruit, comme l'exemple en haut à gauche qui n'est d'autre que du sable fin), jusqu'à des textures très répétitives spatialement en termes de niveaux de gris (comme dans le cas des rides de sables, en haut à droite). La présence de roches, de cailloutis, de traces de dragage ou encore d'objets manufacturés, fait également parti de cet environnement marin apparaissant sur une image, et qui contribue à sa complexité.

On constate alors que l'information portée par une image sonar latéral peut être de plusieurs types. Premièrement, c'est une **information énergétique** de par le niveau de rétro-diffusion de l'onde sonore enregistrée par le sonar puis projetée sur l'image. Deuxièmement, c'est une **information structurelle** puisqu'on observe sur ces images des arrangements entre

pixels qui forment des motifs pouvant être reliés à des structures connues (roches, objets, rides de sable etc.). Troisièmement, l'image sonar contient des **informations géométriques** puisque nous pouvons avoir accès à la direction / l'orientation de ces différentes structures (orientation des rides de sable, du bord d'une roche, etc.). De plus, l'image sonar contient des **informations multi-échelles**, puisque toutes les structures observées ne sont pas de la même taille / dimension. Par exemple, on peut y observer des cailloux comme des rochers, des rides de sable comme des dunes. Nous verrons de manière plus approfondie au chapitre suivant, que de nombreux travaux proposent de caractériser et d'analyser l'image sonar par le biais d'analyses de la texture.

Ainsi, l'environnement marin est une notion complexe en termes énergétiques, structurales et géométriques, dans lequel détecter une mine sous-marine sera une tâche potentiellement très difficile. Dans le paragraphe suivant, nous allons nous attacher à décrire comment un objet apparaît sur ces images sonar, et comment il peut être alors caractérisé.

### 2.3.3 Apparition d'objets

La Figure 2.13 montre un objet insonifié dans le champ lointain d'une antenne sonar, pour un angle de rasance  $\phi$ . Sous cet objet, le niveau d'intensité de l'onde acoustique rétro-diffusée est affiché par l'intermédiaire d'un niveau de gris (l'ordre de grandeur de ces niveaux est affiché sur une *colorbar* en haut à droite de la figure). On y remarque que le niveau de rétro-diffusion n'est pas le même en fonction de la surface de l'objet insonifié. En effet, ce niveau de rétro-diffusion dépend de plusieurs facteurs. Premièrement, il est dépendant de la constitution physique de l'objet puisque l'onde acoustique n'interagit pas de la même manière avec tous les matériaux. Deuxièmement, le niveau de l'onde sonore rétro-diffusée dépend également de l'orientation de la surface insonifiée. Si celle-ci a une normale orientée vers le sonar, le niveau rétro-diffusé sera fort (le **spéculaire**). Au contraire, plus la différence entre l'angle d'inclinaison de l'antenne et la normale de cette surface sera importante, moins l'énergie rétro-diffusée sera forte. De plus, l'état de la surface joue également un rôle sur le niveau de rétro-diffusion. En

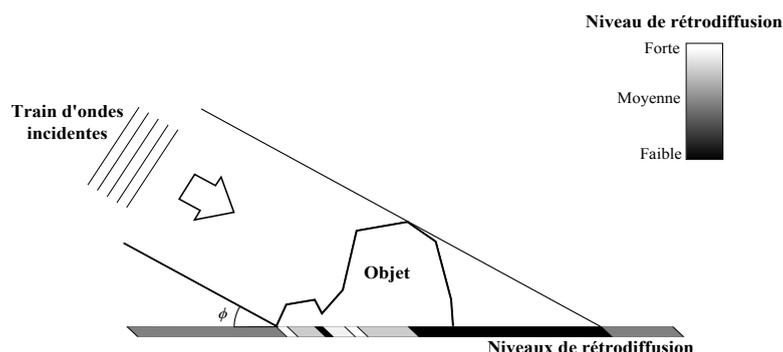


FIGURE 2.13: Représentation schématique de l'insonification d'un objet en champ lointain (inspirée de [1]).

effet, plus une surface sera lisse, plus l'onde réfléchiée sera concentrée dans une direction précise (réflexion spéculaire). Au contraire d'une surface rugueuse qui diffusera l'onde réfléchiée dans toutes les directions.

Ces différentes observations permettent de constater que les surfaces d'un objet orientées vers le sonar fournissent des niveaux de rétro-diffusion forts qui constitueront **l'écho de cet objet**. Les surfaces orientées à l'opposée de la direction d'incidence des ondes sonores auront quant à elles des niveaux très faibles, contribuant à la notion **d'ombre de l'objet**. Cette ombre sera également fonction de la hauteur de l'objet. En effet, plus l'objet aura une hauteur importante, et plus la zone de fond marin derrière lui ne pourra être insonifiée comme le montre la Figure 2.13. Les niveaux de rétro-diffusion dans ces zones seront majoritairement dus aux réverbérations de volume.

Ainsi, sur les images issues d'un sonar latéral, les objets apparaissent par une paire écho-ombre caractéristique. La Figure 2.14 permet d'observer ce phénomène pour un ensemble d'exemples extraits d'image sonar réelles. On y constate la variabilité des formes de ces échos faisant apparaître la silhouette des objets, mais également la variabilité des formes des ombres associées qui renseignent, entres autres, sur la hauteur et forme de l'objet.

La résolution avec laquelle nous pouvons observer ces objets dépend directement du système sonar utilisé, comme nous l'avons vu dans la section 2.2. Dans le passé, les résolutions sonar étaient telles que l'écho de l'objet était très rarement de bonne qualité, et ne permettait pas d'être utilisé pour caractériser un objet posé sur le fond. Uniquement l'ombre de ce dernier (généralement de bien plus grande dimension) était alors traitée pour définir la nature de cet objet. De nos jours, l'évolution des systèmes sonar, et en particulier l'arrivée du sonar à antenne

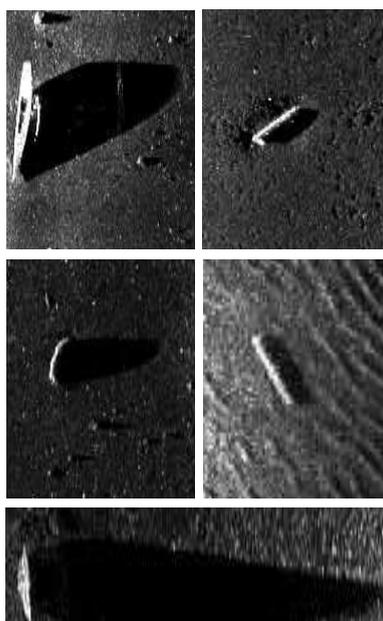


FIGURE 2.14: Mosaïque de différents objets apparaissant dans des image sonars.

synthétique (SAS - *Synthetic Aperture Sonar*), permet une résolution bien plus importante de l'image. Par conséquent, le contenu informatif concernant l'écho augmente et permet d'être pris en compte dans la caractérisation de l'objet.

Sur la mosaïque d'images précédente, nous avons présenté des images sonar où les objets posés sur les fonds sont clairement identifiables en termes d'échos et d'ombres. Ce phénomène n'est qu'un cas particulier relativement simple. En effet, ces objets sont posés sur des fonds constitués de sable fin (*i.e.* relativement homogènes), qui permettent à l'écho et à l'ombre de clairement se distinguer du fond marin. Cependant, l'objet peut être potentiellement posé sur d'autres types de fonds bien différents, qui rendront l'observation du motif écho-ombre bien plus difficile. Par exemple, on observe sur la Figure 2.15 deux exemples d'imagettes sonar illustrant des rides de sable. Sur chacune, un rectangle vert indique la position où une mine sous-marine est posée. On constate alors que le motif écho-ombre de la cible peut se confondre facilement avec le motif écho-ombre provoqué par le caractère ondulé du fond marin. Ainsi, si dans l'imagette (a) on peut tout de même constater la présence de la mine sous-marine (son écho et son ombre), dans l'imagette (b) on ne la distingue pas (l'objet est ici caché derrière une ride de sable).

De ce fait, l'observation d'un objet sur l'image sonar est une opération qui peut devenir complexe, puisque ce dernier peut se confondre avec son environnement proche. Nous verrons dans le chapitre suivant les problématiques liées à la détection de ces objets de manière plus approfondie.

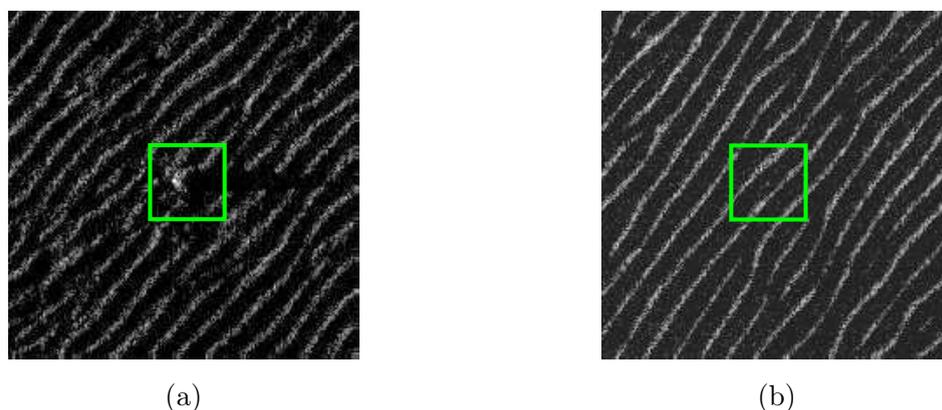


FIGURE 2.15: Exemple de la difficulté de détection d'une cible posée sur un fond non-homogène (rides de sable ici). Le rectangle vert indique la position d'une mine sous-marine. En (a), paire écho-ombre de l'objet visible. En (b), paire écho-ombre non-visible.

## 2.4 Présentation des bases de données utilisées dans la thèse

Dans ces travaux, nous utiliserons des données issues de systèmes sonar différents afin d'étudier la robustesse de nos algorithmes. Ces données sont de deux types : des données

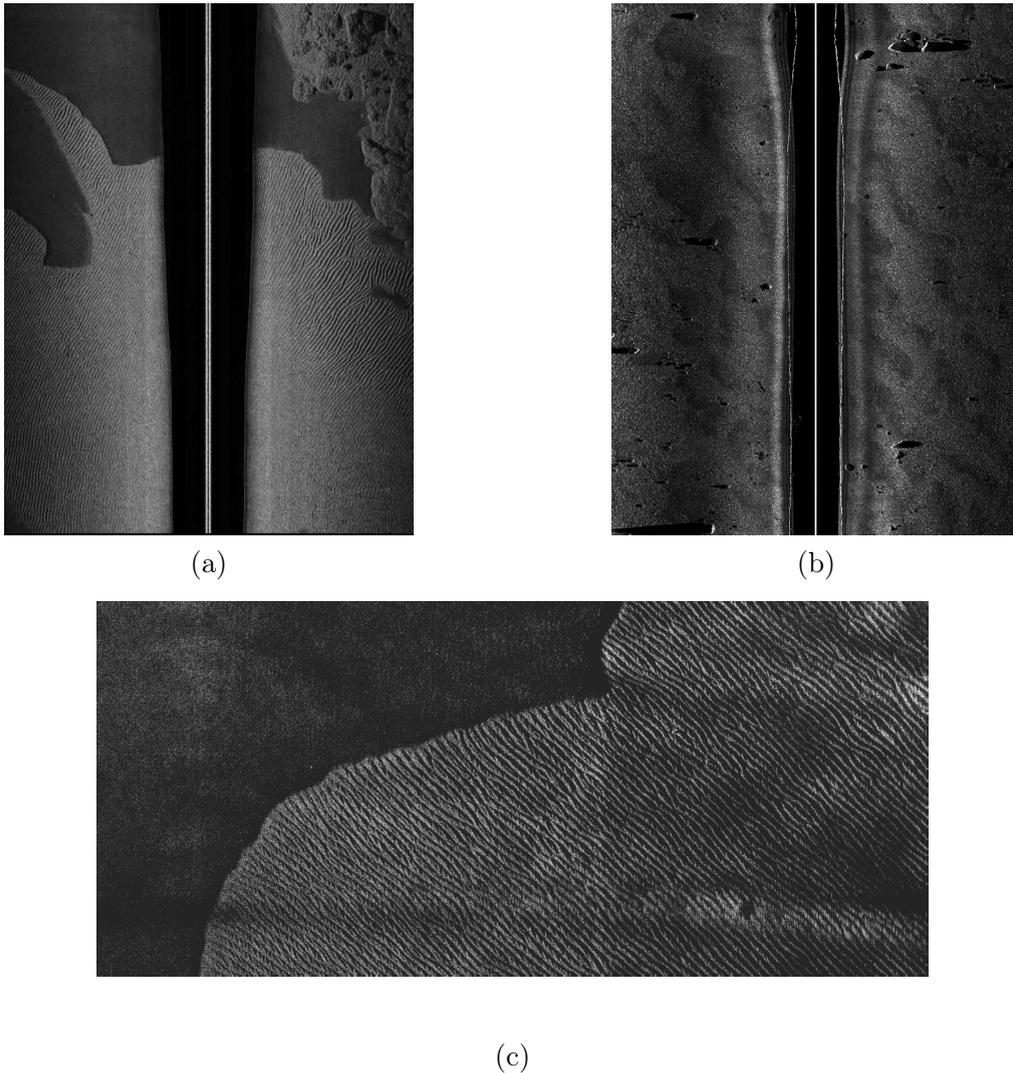


FIGURE 2.16: Exemples d'images sonar pour chacune des trois bases de données, (a) base de données DSTO, (b) base de données DGA, et (c) base de données CMRE.

de sonars latéraux classiques et des données de sonar à antenne synthétique SAS. Elles sont réparties en trois bases de données distinctes.

La première a été fournie par la Direction Générale de l'Armement (DGA). Les images sonar contenues dans cette base ont été collectées lors d'une mission réalisée en 2002 à l'aide d'un sonar latéral Klein 5000 monté sur un poisson et tracté par un navire. Sa fréquence d'émission est de 455 kHz, et sa portée maximale de 75m. La dimension des pixels est fixée à 10 cm en azimut (*i.e.* sur l'axe *along-track*) et en distance (*i.e.* sur l'axe *across-track*).

La seconde a été fournie par le *Defense Science and Technology Organisation* australien (DSTO) et provient d'un challenge ATR que le DSTO a organisé en 2013. Ces données sont divisées en une base d'entraînement (environ 500 images où la position des cibles est spécifiée), et en base de test (5000 images où la position des cibles est inconnue). Elles ont été collectées à l'aide de l'AUV REMUS 100 sur lequel un sonar latéral Marine Sonic Technology Ltd a été

monté. Sa fréquence d'émission est de 900 kHz, pour une portée maximale qui oscille entre 30 et 50 m en fonction des images. Ces images sont toutes de taille 1024 x 1000 pixels dont les dimensions oscillent entre 12 et 15 cm en azimuth et entre 6 et 10 cm en distance selon la portée. Il est également à noter que pour ces images, les données de navigation de l'AUV sont fournies.

La troisième et dernière base de données d'images sonar a été fournie par le CMRE (*Centre for Maritime Research and Experimentation*) de l'OTAN situé à La Spezia en Italie. Cette base de données fournit des images sonar collectées par un système SAS. De ce fait, la dimension des pixels est nettement améliorée puisqu'elle atteint 2.5 cm en azimuth et 1.5 cm en distance à l'aide d'un sonar interférométrique SAS de 300 kHz de fréquence centrale, et de 60 kHz de bande passante. Ce sonar fut monté sur l'AUV MUSCLE durant la campagne ARISE de 2011 dont l'on pourra étudier environ 300 images. Sur ces images, la position des mines d'exercice posées sur le fond sont accessibles.

Un exemple de ces trois types de données est présenté sur la Figure 2.16. Notons ici que l'exemple de la base de données du CMRE (cf. Figure 2.16) est une visualisation de l'image tribord où la colonne d'eau (bande noire centrale) a été supprimée.

## 2.5 Conclusion

Dans ce chapitre, les principes de base de l'imagerie acoustique ont été abordés avec une attention particulière pour les données issues de sonars latéraux. Ces systèmes sont, pour la plupart du temps, des moyens simples et à faible coût pour imager le fond de la mer de manière détaillée.

Cependant, la résolution toujours plus fine de ces images que permettent les progrès technologiques, font de l'image sonar un élément très complexe à analyser de manière automatique. La variabilité de l'environnement marin qui y apparaît, tant en termes énergétiques, structurels, ou géométriques, fait que cette analyse doit être réalisée minutieusement, afin de permettre une étude correcte et efficace des fonds marins.

Dans le chapitre suivant, nous allons nous intéresser à une application spécifique de l'analyse de ces images sonar à savoir le domaine de la chasse aux mines, et particulièrement les chaînes de reconnaissance automatique de cibles.

# Chapitre 3

## Détection et classification de cibles en imagerie sonar

### Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>32</b>
<b>3.2</b>	<b>Les processus ATR</b>	<b>32</b>
3.2.1	Phase de détection	33
3.2.1.1	Détection par filtrage adapté	34
3.2.1.2	Segmentation orientée détection	34
3.2.1.3	Détection par saillance	35
3.2.2	Phase d'extraction d'attributs	36
3.2.2.1	Attributs statistiques	36
3.2.2.2	Attributs géométriques	37
3.2.2.3	Attributs de Fourier	37
3.2.3	Phase de Classification	37
3.2.3.1	Classification supervisée	38
3.2.3.2	Classification basée modèles	38
3.2.4	Limites et perspectives	39
<b>3.3</b>	<b>Nécessité d'une caractérisation environnementale</b>	<b>40</b>
3.3.1	Trois grandes catégories de fonds marins pour l'ATR	40
3.3.2	Descripteurs d'anisotropie et de complexité des fonds marins	41
3.3.2.1	Définition des descripteurs	42
3.3.2.2	Estimation des descripteurs	44
3.3.3	Exploitation des descripteurs environnementaux dans l'ATR	52
3.3.3.1	Intégration en classification de mines sous-marines	53
3.3.3.2	Intégration pour la réduction des taux de fausses alarmes	55
<b>3.4</b>	<b>Approche proposée dans cette thèse</b>	<b>57</b>

---

3.4.1	Position par rapport à la littérature . . . . .	57
3.4.1.1	Nouvelle représentation de l'information . . . . .	58
3.4.1.2	Descripteurs environnementaux plus adaptés . . . . .	59
3.4.1.3	Chaîne de traitement différente . . . . .	59

---

## 3.1 Introduction

Dans le chapitre 1, nous avons introduit les opérations de contre-mesures sous-marines (MCM) et expliqué en quoi les systèmes autonomes pouvaient, à terme, être utilisés pour les rendre plus rapides, plus sûres, et surtout plus efficaces. S'en est suivie, dans le chapitre 2, une introduction aux systèmes imageurs du fonds marins, avec une attention particulière portée au système sonar latéral. Dans la suite de cette thèse, nous nous intéresserons aux problématiques liées à l'analyse automatique des images issues de tels capteurs, pour des applications de reconnaissance automatique de cibles. Dans ce chapitre, nous allons introduire ce processus ATR et les challenges qui lui sont associés, par le biais de la littérature, dans le contexte particulier de l'imagerie sonar.

Ce chapitre est organisé de la manière suivante. Le paragraphe 3.2 propose une introduction aux chaînes ATR où nous verrons que ce processus particulier est constitué de trois phases distinctes et successives, à savoir une phase de détection, une phase d'extraction d'attributs puis une phase de classification. Nous y verrons également que ce traitement doit être complété par une prise en compte de l'environnement marin. Les premières études sur l'extraction et l'intégration de données environnementales dans le processus ATR seront alors présentées à la section 3.3. Enfin, dans la section 3.4, nous présenterons l'approche que nous allons développer à travers ces travaux de thèse.

## 3.2 Les processus ATR

De manière générale, l'acronyme ATR désigne les algorithmes informatiques qui visent à détecter, classer et identifier la signature d'une cible dans des données issues d'un système d'observation. Dans ces travaux de thèse, les processus ATR que nous étudierons concernent la détection de cibles sous-marines par le biais d'images issues de systèmes sonars à balayage latéral. Nous avons vu dans le chapitre précédent que de tels systèmes ont pour avantage de pouvoir scanner rapidement une vaste zone de fond marin.

La chasse aux mines sous-marines est divisée de manière historique en trois tâches successives : la détection, la classification, puis l'identification. Les frontières entre ces trois opérations successives sont relativement floues et dépendent souvent des termes techniques utilisés dans les différentes marines. Généralement, la phase de détection revient à discriminer la signature

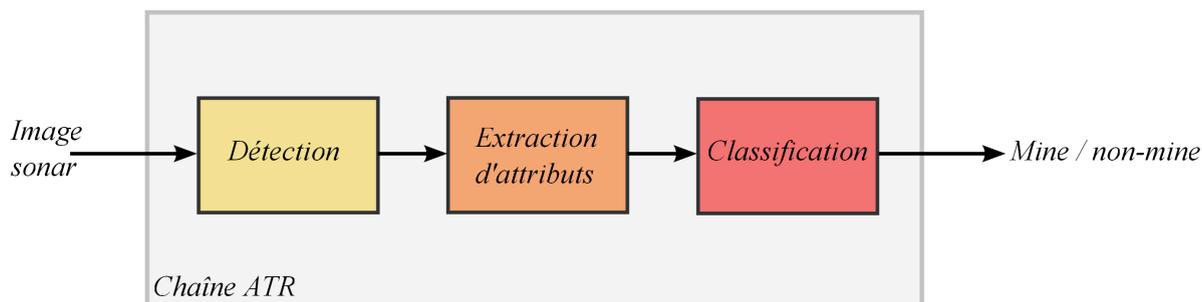


FIGURE 3.1: Séquence de traitements d'un processus ATR.

d'une cible par rapport aux réverbérations du fond, alors que la phase de classification consiste à séparer ces détections en deux classes : mine / non-mine. La phase d'identification introduit quant à elle la possibilité de discerner les différents types de mines (par exemple les mines Manta, Rockan, sphériques, etc.).

Dans cette thèse, nous limiterons notre définition des chaînes ATR aux étapes de détection et de classification des cibles sous-marines. La Figure 3.1 propose le schéma d'un tel processus. A partir d'une image sonar, le traitement commence par une **phase de détection** des cibles potentielles sur l'intégralité de l'image sonar. Ces détections (ou "contacts") sont ensuite décrites à l'aide d'une **phase d'extraction d'attributs** afin de pouvoir finalement les distinguer suivant deux classes : mine / non-mine, lors d'une **phase de classification**.

On note ici le caractère séquentiel d'une telle chaîne de traitement. Cette succession d'étapes permet de réduire la taille des données au fur et à mesure du traitement, en partant d'une image très importante, pour finir une information sur la présence ou non de mines sous-marines. L'avantage de cette architecture est de pouvoir augmenter la complexité des algorithmes de traitement tout en gardant un coût de calcul constant.

Nous allons maintenant décrire ces différentes étapes qui forment la chaîne ATR.

### 3.2.1 Phase de détection

La phase de détection est la première opération d'une chaîne ATR. Elle consiste à traiter directement l'image sonar de manière à identifier les zones qui contiennent potentiellement des cibles. De telles zones sont couramment appelées **régions d'intérêts** ou ROI (acronyme anglais de *region of interest*). Puisque cette phase doit analyser des images de grandes dimensions, il est nécessaire que ces algorithmes soient peu gourmands en temps de calculs.

La littérature propose différentes approches pour mener à bien cette étape. Une première catégorie regroupe les méthodes de segmentation orientées détection, qui visent à segmenter les images sonar en trois classes : les échos, les ombres et le fond. Une seconde catégorie regroupe les algorithmes basés sur la saillance. Ce sont des approches plus variées, qui considèrent généralement les mines sous-marines comme des anomalies au sein d'une région texturée.

En d'autres termes, les zones de l'image contenant des cibles sont considérées comme des "erreurs de comportement" du fond. Les approches par l'utilisation de filtres adaptés (plus connues sous le nom de *matched filtered*) constituent une troisième catégorie et se basent sur une modélisation particulière du motif formé par une mine dans les images sonar. Nous introduirons ces différentes méthodes dans les paragraphes suivants.

### 3.2.1.1 Détection par filtrage adapté

La détection de mines sous-marines à l'aide de filtres adaptés est la première à avoir fait l'objet d'un brevet. Il fut déposé par Stevens et. al. en 1977 [41]. Il est basé sur l'observation qu'une mine sous-marine possède un motif particulier dans le signal (ou "ping") enregistré par un sonar. En effet, comme nous l'avons vu dans le chapitre précédent (cf. section 2.3.3), ce signal 1D est composé d'un pic très intense (l'écho), suivi d'une zone de très basse intensité, de dimension variable (l'ombre). L'idée est alors de créer un filtre 1D qui modélise une telle réponse, puis de le convoluer avec le signal initial. Pour déterminer la présence d'une cible, il faut alors étudier la présence de maxima locaux dans des pings successifs.

T. Aridgides [42], J. Hyland [43] et G. Dobeck [44] ont popularisé et généralisé cette méthode directement sur des images sonar en 2D.

### 3.2.1.2 Segmentation orientée détection

En 1989, M. Doherty propose d'utiliser une approche segmentation de la détection d'objets sur les images sonar [45]. Il observe que les pixels associés à une cible n'ont pas la même distribution statistique (en termes de niveaux de gris) que les pixels associés au fond marin. Il propose alors de réaliser un seuillage adapté des images pour mettre en évidence les échos. Afin de valider la détection, il met en place une recherche des ombres associées à ces échos, par un ensemble d'opérations de lissages et de moyennages.

Ces prémices montrent ce que sera l'une des approches les plus couramment utilisées en détection d'objets au cours des décennies suivantes, à savoir une segmentation de l'image en trois classes : l'écho d'une cible, l'ombre et le fond.

En 1995, M. Bello introduit les champs de Markov aléatoires dans ce contexte, et démontre qu'ils sont adaptés à une telle segmentation [46]. Ces résultats poussent M. Mignotte et C. Collet à approfondir cette méthode [47, 48, 49, 50]. Cependant, du fait des temps de calculs importants et des capacités informatiques du moment, il faut attendre 2003 et l'étude proposée par S. Reed et collègues, pour obtenir un algorithme efficace [51, 52, 53]. En utilisant des informations spatiales *a priori* (sur les tailles et signatures géométriques de cibles), un modèle de champ de Markov orienté-détection est alors développé pour segmenter l'image en

ces trois classes. En 2014, O. Daniell propose dans [4] de revoir cette méthode en lui appliquant la méthode des *graph-cuts* [54] pour accélérer toujours plus la convergence des résultats.

Une autre approche initiée par B. Calder dans [55, 56] propose quant à elle d'utiliser un modèle stochastique bayésien afin de classer chaque pixel d'une image. Pour cela, il réalise une modélisation bayésienne des données et fait appel à un champ de Gibbs pour modéliser les cibles. L'approche est considérée comme robuste mais très gourmande en temps de calculs.

F. Maussang et al. ont également investi l'approche statistique pour segmenter les images sonar et en particulier les images SAS [57, 58]. Ces méthodes sont basées sur la relation entre la moyenne et l'écart-type de la distribution de Rayleigh des niveaux de gris dans les images sonar. En modélisant les réponses du fond par une loi de Weibull (nous avons vu dans le paragraphe 2.3.1 que la loi de Rayleigh en est un cas particulier), ils observent que la moyenne et l'écart-type sont reliés par une constante multiplicative. Cependant, pour des zones d'écho et d'ombre, les auteurs notent que cette relation n'est plus aussi stricte. Ils proposent alors d'en tirer partie en appliquant deux seuillages dans un plan défini par la moyenne et l'écart-type. Ces derniers permettent alors d'extraire les zones d'échos et d'ombres.

### 3.2.1.3 Détection par saillance

Les méthodes de détection par saillance sont aujourd'hui couramment utilisées et proposent de voir un objet comme une anomalie dans une région texturée. Elles cherchent à modéliser les différences entre une zone du fond donnée et son voisinage.

Par exemple dans [59], L. Linett se base sur le fait que les réverbérations du fond peuvent être modélisées par un processus Brownien fractionnaire [60] dont la distribution est décrite par sa dimension fractale. Il considère alors qu'une région dont la dimension fractale est différente de ces voisines, peut contenir un objet.

Dans [61], L. Attalah utilise l'entropie de Shannon pour détecter les régions saillantes. En effet, il constate que les régions contenant des ombres ou des échos ont une plus forte entropie que de simples zones de fond marin. Ainsi, l'auteur propose de détecter les objets via une détection de pics dans la fonction d'entropie.

Dans [62], l'image sonar est décomposée en blocs (de la taille typique d'une mine sous-marines, *i.e.* de l'ordre du mètre) et l'auteur propose de calculer la corrélation, selon les colonnes, entre les différents blocs d'une même région. Selon l'auteur, les mines sous-marines auront alors une forte valeur de corrélation, et un simple seuil permettrait de les détecter.

Dans [8], D.P. Williams propose un algorithme de détection par saillance dont les calculs se font par l'intermédiaire du concept d'image intégrale. Ce dernier permet d'atteindre des vitesses de détection proches du temps réel sur des images SAS malgré leurs très grandes dimensions. A partir d'un ensemble de fenêtres glissantes aux tailles variables, l'auteur propose

d'estimer des zones d'échos et d'ombres par une architecture en cascade. La simplicité et la rapidité de l'algorithme sont ses principaux atouts qui lui permettent d'être l'un des rares à pouvoir être implémenté à bord des AUVs.

### 3.2.2 Phase d'extraction d'attributs

Une fois la phase de détection terminée, l'image sonar initiale est réduite en un nombre fini de contacts. Le voisinage autour de chacune de ces détections définit une ROI. Pour vérifier que ces contacts sont bien des mines sous-marines, chacun d'entre-eux est envoyé vers une phase de classification. Cependant, pour pouvoir réaliser une telle opération, il est nécessaire d'en extraire un ensemble d'attributs permettant de les décrire. Ces attributs sont généralement calculés sur l'objet détecté, mais parfois également dans son voisinage. De nombreux travaux ont permis de définir différentes classes de descripteurs pouvant être invariants par rapport à la position, à l'orientation et à la dilatation des objets.

#### 3.2.2.1 Attributs statistiques

Les attributs statistiques permettent de représenter la distribution de l'intensité des pixels de la région d'intérêt. Les plus simples sont ceux issus des statistiques d'ordre 1, comme la moyenne, la variance, l'énergie ou encore l'entropie. Cependant, ces quantités ne dépendent que de l'histogramme, et prennent donc des valeurs arbitraires sur un même objet, comme par exemple en fonction du contraste, de sa position dans l'image ou encore d'une transformation telle qu'une rotation ou dilatation.

La théorie des moments a alors été introduite pour pallier ces défauts. Les moments sont calculés à partir des pixels contenus à l'intérieur des frontières d'un objet. Ils ont été introduits par M.K. Hu en 1962 [63]. Différents combinaisons de ces moments permettent alors d'obtenir des descripteurs invariants aux transformations de base pouvant intervenir sur les objets (par exemple la translation, le changement d'échelle et la rotation). M. Teague [64] propose quant à lui de favoriser l'utilisation des moments de Zernike. Ces moments sont basés sur la théorie des polynômes orthogonaux, et permettent d'offrir une représentation des propriétés de l'image en évitant toute redondance. Il faut cependant noter que la description d'une forme par ces méthodes est bien souvent limitée car les moments d'ordre supérieur à deux sont très sensibles aux dégradations mineurs que peuvent subir les contours, comme le montre A. Sluzek [65]. Malgré cela, A. Isaacs propose tout de même de combiner ces différents moments afin de décrire, puis de classer les détections sous-marines [66].

Nous allons voir maintenant que d'autres approches proposent quant à elles d'extraire des descripteurs directement sur la forme des objets. Il a d'ailleurs été remarqué que la classification des motifs basée sur cette technique est intimement liée aux modes de fonctionnement du système visuel humain.

### 3.2.2.2 Attributs géométriques

Les attributs géométriques sont extraits à partir d'images segmentées (en termes d'échos, d'ombres et de fond). Ils sont essentiellement utilisés pour la description des zones d'ombre via des paramètres de longueur, de largeur ou d'orientation. Souvent, ces méthodes sont basées sur la modélisation de la zone d'ombre par une ellipse dont les dimensions fournissent ces différents paramètres [67, 68]. L'amélioration des technologies sonar, et en particulier l'apparition des systèmes SAS, permettent aujourd'hui d'avoir une résolution suffisante sur les échos des cibles et donc d'en extraire également ces attributs.

Notons ici que I. Quidu [69] propose quant à elle d'utiliser une combinaison entre les attributs statistiques et géométriques sous la forme d'un "ensemble hybride". Cette méthode permet d'utiliser les points forts de chacun d'entre eux. Par exemple, elle réduit l'influence du bruit ainsi que les distorsions géométriques pouvant intervenir sur les images sonar et qui réduisent souvent l'efficacité des descripteurs statistiques et géométriques respectivement. Dans ces travaux, une normalisation par rapport aux angles de rasance est également réalisée afin que chaque objet détecté apparaisse sous une rasance de  $45^\circ$  rendant alors plus robuste l'extraction des descripteurs.

Cependant, les performances de ces méthodes pour distinguer les différents objets sont limitées. En effet, elles dépendent fortement de la précision avec laquelle on peut calculer toutes ces caractéristiques, en fonction de la portée ou du type de fond.

### 3.2.2.3 Attributs de Fourier

Les descripteurs de Fourier réalisent la transformation du contour d'un objet dans le domaine fréquentiel. Ils constituent une représentation multi-résolution de la forme de l'objet et un nombre limité d'entre-eux est suffisant pour reconstruire l'objet. De plus, les opérations de translation, rotation, dilatation et décalage du contour s'expriment de façon très simple dans le domaine fréquentiel. Ils permettent alors de fournir des descripteurs invariants. Plusieurs études mentionnent l'utilisation de tels descripteurs dans la classification de mines sous-marines comme les études de D. Boulinguez [70], R. Fandos [71] ou encore I. Quidu [72].

## 3.2.3 Phase de Classification

Une fois que les descripteurs sont extraits à partir des contacts, il est possible de réaliser une phase de classification. A l'image des phases de détection et d'extraction d'attributs, la phase de classification a fait l'objet de nombreuses études et peut être réalisée de différentes manières. Généralement, les algorithmes de classification de mines sous-marines se scindent en deux groupes à savoir les approches supervisées et les approches basées sur des modèles.

### 3.2.3.1 Classification supervisée

Les approches de classification de mines sous-marines supervisées se basent sur l'utilisation de données d'entraînement dont on connaît la vérité terrain (*i.e.* les labels mine / non-mine). Lors de la phase d'entraînement (ou d'apprentissage), les attributs précédents sont calculés pour chaque exemple de la base de données. Le principe étant alors de faire apprendre au classifieur la distinction entre une mine et une non-mine dans l'espace défini par ces descripteurs.

Sur des types de fonds similaires à la base d'entraînement, une telle approche permet d'atteindre de très bons taux de classification comme le montre T. Cobb [73]. Cependant, si les fonds marins observés sont différents pendant la phase de test, aucune garantie n'est alors donnée aux résultats du classifieur [4].

Ce désavantage peut être corrigé en parti par l'utilisation des simulateurs qui intègrent des cibles dans des images réelles [52, 74, 75]. On peut alors entraîner des algorithmes sur un ensemble d'environnements très divers en augmentant de manière synthétique la base de données. Cependant, il est possible que le classifieur apprenne également des artefacts qui peuvent être dus au processus de simulation.

Afin d'éviter de recourir à une base d'entraînement, il est possible de faire appel à des méthodes de classification qui sont alors basées sur l'utilisation de modèles.

### 3.2.3.2 Classification basée modèles

La classification basée sur des modèles repose entièrement sur des connaissances *a priori* que l'on possède sur les objets à détecter. Ces informations permettent alors de créer un modèle pour chaque cible que l'on désire mettre en évidence. De telles approches sont utiles pour classer des cibles sur lesquelles on ne possède pas beaucoup d'exemples (*i.e.* pas d'apprentissage possible).

Dans la littérature sur l'ATR, de nombreuses méthodes ont été proposées, et reposent toutes exclusivement sur l'appariement de modèles, nous pouvons citer par exemple [51, 52, 53, 76, 77, 78, 79]. Seules les implémentations sont alors différentes, avec une forte variété d'algorithmes.

Certaines de ces approches proposent d'utiliser des algorithmes d'appariement basés sur les contours comme celle de S. Reed [76]. Son algorithme permet une coopération entre contours actifs pour extraire l'écho et l'ombre des cibles de manière conjointe, dans un environnement complexe. La théorie de décision de Dempster Shafer est ensuite utilisée pour classer les détections observées vis-à-vis des modèles connus.

Autre exemple, dans [78], E. Hasanbelliu part du principe que les échos et les ombres des objets sont segmentés au préalable et proposent alors d'utiliser une corrélation non-linéaire pour associer des modèles aux détections réalisées.

### 3.2.4 Limites et perspectives

Jusqu'à présent dans ce chapitre, nous avons introduit de manière générale les différentes approches liées à la problématique de l'ATR pour des données issues de systèmes sonar à balayage latéral. Pour cela, nous avons suivi le cheminement séquentiel de la chaîne ATR en commençant par la phase de détection, pour aboutir à celle de classification des objets. Ces deux étapes fondamentales sont alors reliées par une phase d'extraction d'attributs.

Nous avons pu constater qu'un ensemble important d'études se font efforcées de mettre en place des algorithmes de traitement autonomes des images sonar afin d'éviter le plus possible l'intervention humaine dans le processus. Cependant, les études sur les chaînes ATR sont toujours d'actualité du fait de la complexité d'une analyse automatique de telles images. En effet, au chapitre 2, nous avons vu que les images du sonar latéral font apparaître les objets posés sur les fonds marins mais également leur environnement proche, par le biais de textures diverses et variées. Cette variabilité et complexité de l'information environnementale font de l'image sonar une donnée difficile à traiter, et en particulier pour y détecter les mines sous-marines.

Des études récentes, telles que celles de D.P. Williams [80] du CMRE, de M. Geilhufe [81] du FFI, ou de O. Daniell [4] de l'Université d'Heriot-Watt (Seebyte), montrent que de manière générale, les conditions environnementales dans lesquelles se passent une mission de chasse aux mines sous-marines conditionnent généralement ses résultats et performances. D'ailleurs, O. Daniell mentionne dans son étude le rôle des textures ressemblant à des cibles ou MLTs (de l'anglais *Mine Like Textures*) que l'on peut définir comme des zones de l'image sonar qui sont propices aux fausses alarmes. Or, selon l'auteur, les taux de fausses alarmes doivent nécessairement être bas pour les missions de chasse aux mines, et en particulier, pour les missions autonomes. En effet, si l'AUV détecte trop de contacts, il ne pourra pas se dérouter pour analyser plus précisément chacun d'entre eux individuellement, sans perdre énormément d'énergie et de temps. De ce fait, aujourd'hui, la littérature préconise de coupler les chaînes ATR avec une caractérisation de l'environnement. L'objectif est alors de fournir à l'algorithme de reconnaissance de cibles des informations sur l'environnement qu'il doit prendre en compte afin de réaliser sa mission de manière performante.

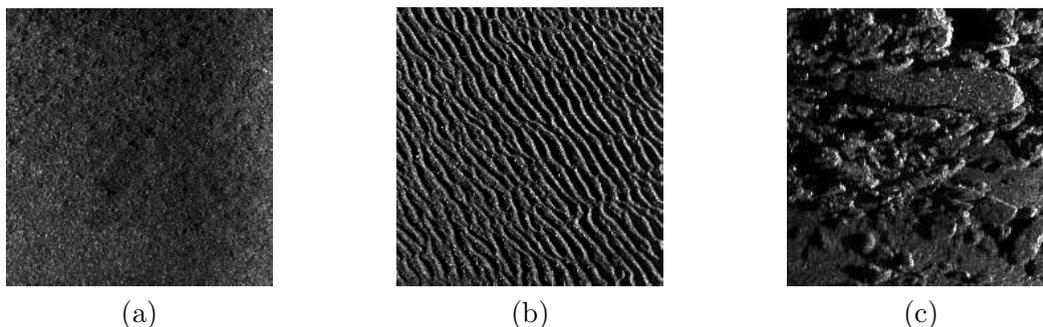


FIGURE 3.2: Exemple d'un fond homogène (a), anisotrope (b), et complexe (c).

### 3.3 Nécessité d'une caractérisation environnementale

Dans [82], O. Daniell est l'un des premiers à introduire la notion de caractérisation des fonds marins dans le processus de reconnaissance de cibles sous-marines. Il démontre que l'augmentation des taux de fausses alarmes sont majoritairement dus à l'influence de deux types de fonds marins, qu'il appelle fonds anisotropes et fonds complexes. Alors que le premier est relié à la présence de rides de sable, le second est quant à lui relié à la notion de fonds marins hautement texturés (*i.e.* avec de fortes variations de niveaux d'intensité des pixels) ou encombrés (*i.e.* une forte densité de structures sous-marines telles que des amas rocheux).

Depuis, la littérature sur l'ATR s'accorde alors à dire qu'il est nécessaire de regrouper les différents types de fonds marins en trois grandes catégories, qui représenteront non plus les fonds en termes sédimentaires, mais en termes de complexités de leurs textures (portées sur l'image). Cette complexité est alors liée à la difficulté d'y détecter des mines sous-marines.

#### 3.3.1 Trois grandes catégories de fonds marins pour l'ATR

Dans ce paragraphe, nous allons introduire les trois grandes catégories de fonds marins qui impactent de manières distinctes les performances d'une chaîne ATR en termes de probabilité de détection et de fausses alarmes selon [2, 3, 4, 80, 81, 82, 83]. Ces trois types de fonds sont les fonds **homogènes**, **anisotropes** et **complexes**.

##### Les fonds homogènes

La catégorie des fonds homogènes regroupe les fonds marins composés de structures sous-marines dont les tailles sont inférieures à celles de la cellule de résolution du sonar. Elle est généralement constituée par les fonds de sable fin ou de vase. Par conséquent, une cible posée sur un fond homogène sera clairement visible (en terme d'écho et d'ombre comme nous l'avons vu au paragraphe 2.3.3) et détectable par un opérateur, ou par une chaîne de traitement automatique. La Figure 3.2.(a) propose l'exemple d'une zone homogène où l'on remarque bien l'absence évidente de structure sous-marine digne d'intérêt d'un point de vue ATR.

### Les fonds anisotropes

Les fonds anisotropes sont des fonds constitués par des structures sous-marines très directives, comme les rides de sables. Ces types de fonds marins sont redoutés par les opérateurs et les algorithmes autonomes du fait des successions "écho-ombre" présentes dans ces zones. Alors qu'un opérateur mettra un temps très important à analyser une telle zone, une chaîne ATR aura quant-à-elle un taux de fausses alarmes très élevé. De tels fonds marins peuvent également rendre des objets moins détectables (confusion entre l'écho de l'objet et celui des rides, objet caché dans l'ombre d'une ride etc.), et ainsi mettre en défaut la phase de détection. La Figure 3.2.(b) propose l'exemple d'un fond anisotrope sur lequel on peut observer les successions d'échos et d'ombres dans une direction particulière (selon une diagonale de l'imagette ici).

L'intérêt de mettre en évidence une telle zone est qu'il existe des possibilité pour la traiter différemment via des post-traitement ou une ré-acquisition d'une image sonar sur la zone pour y détecter les objets potentiellement cachés. Nous y reviendrons dans la suite de ce manuscrit.

### Les fonds complexes

Les fonds complexes sont les fonds marins constitués de motifs plus difficiles à interpréter. Ils sont généralement constitués de cailloutis, de roches, de posidonie etc. Ce sont des fonds encombrés qui nécessitent un effort d'analyse très important. Une mine sous-marine peut facilement s'y confondre avec son environnement proche, et devenir très difficile à détecter. Dans ces zones, les probabilités de détections risquent de décroître significativement. La Figure 3.2.(c) illustre un fond considéré comme complexe, où l'écho d'un objet peut se confondre avec l'écho des roches ou se cacher dans une grande zone d'ombre. De plus, la densité des structures dans cette zone peut amener un algorithme de détection à y fournir un taux très important de fausses alarmes (beaucoup de succession écho-ombre dans la zone).

Ces trois catégories de fonds marins représentent alors différents niveaux de difficulté auxquels doit faire face un opérateur humain ou un système autonome. Différentes méthodes ont été proposées pour les extraire des images sonar.

### 3.3.2 Descripteurs d'anisotropie et de complexité des fonds marins

La question de l'extraction de ces fonds marins peut directement faire penser aux problématiques de classification des images sonar. Cependant, les méthodes générales de classification des fonds marins se basent sur le principe qu'il existe des frontières nettes entre les différents types de fonds marins. Or, ne serait-ce que la difficulté de segmentation manuelle des images sonar prouve que ces frontières sont souvent subjectives et laborieuses à définir. A l'inverse, nous pouvons supposer que les conditions environnementales résident plutôt dans un continuum

que dans des classes discrètes. Un tel continuum irait donc de pair avec une caractérisation des fonds marins par l'intermédiaire de descripteurs continus. Ces descripteurs doivent ainsi témoigner d'une interprétabilité aisée de la texture du fond, afin de pouvoir les intégrer aussi bien à une chaîne de traitement automatique qu'à un outil d'aide à l'opérateur à bord des navires militaires.

Afin de mettre en évidence les différentes catégories de fonds marins présentées, la littérature propose de définir deux descripteurs aux noms évocateurs : l'*anisotropie* et la *complexité* de la texture du fond marin.

### 3.3.2.1 Définition des descripteurs

Afin de distinguer ces deux types de fond, O. Daniell propose de tirer parti d'un aspect particulier des zones anisotropes, à savoir leur caractère orienté. Il propose alors l'utilisation de filtres orientés pour permettre de les discerner. Pour cela, il analyse les réponses énergétiques d'un banc de filtres  $\psi_{s,k}$ , orientés selon différentes directions  $k \in \{0^\circ, 0^\circ + \Delta\theta, \dots, 180^\circ\}$ , où  $\Delta\theta$  définit le pas de discrétisation en angle des filtres, et où le paramètre  $s \in \{1, \dots, S\}$  joue le rôle de coefficient de dilatation pour tenir compte des différentes tailles des structures sous-marines possibles (*i.e.* un rôle d'échelle). On note alors la réponse énergétique d'un filtre, à une orientation donnée  $k$  et une échelle donnée  $s$ , par :

$$E_{s,k} = |\psi_{s,k}(x, y)|^2, \quad (3.1)$$

et la moyenne  $N_s$  de ces réponses énergétiques sur les différentes orientations pour une échelle  $s$  donnée :

$$N_s = \frac{1}{K} \sum_{k=1}^K E_{s,k}, \quad (3.2)$$

où  $K$  désigne le nombre d'orientations.

O. Daniell part du constat que les rides de sable ont une direction privilégiée, comme le montre les Figures 3.3.(b) dans le domaine spatial (successions écho-ombre dans une direction donnée) et 3.3.(e) dans le domaine fréquentiel (concentration du spectre dans une direction donnée). Il suppose alors qu'il est possible de définir de tels filtres orientés qui répondront de manière forte dans une direction proche de celle des rides de sable et de manière très faible dans une direction perpendiculaire à cette dernière. Il définit l'expression d'un descripteur d'anisotropie  $A$  par :

$$A = \sum_{s=1}^S \frac{\text{std}_k(E_{s,k})}{N_s}, \quad (3.3)$$

où  $\text{std}_k(E_{s,k})$  correspond à l'écart-type des énergies  $E_{s,k}$  suivant les orientations  $k$  pour une échelle  $s$  fixe. Le descripteur d'anisotropie  $A$  réalise donc la somme des écarts-types suivant les orientations pour toutes les échelles d'analyse. Si une texture possède une direction privilégiée,

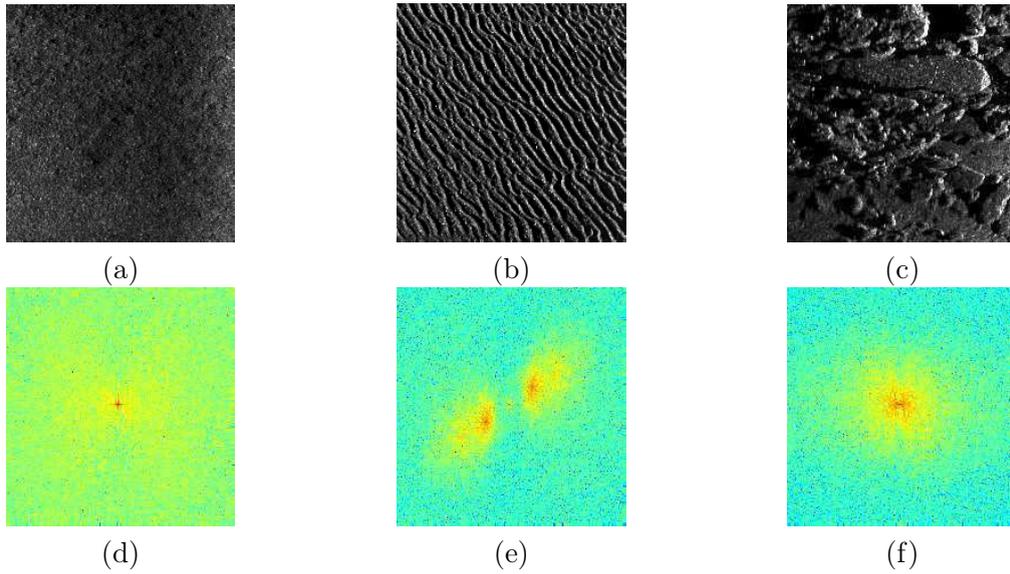


FIGURE 3.3: Trois exemples de fonds marins : homogène (a), anisotrope (b), complexe (c), et leurs transformées de Fourier (TF) respectives (d), (e), et (f). L'origine du plan fréquentiel est considéré au centre de l'image.

alors l'écart-type des réponses énergétiques du banc de filtres orientés sera important, et par conséquent l'anisotropie  $A$  le sera également. A l'inverse, l'écart-type sera faible si la texture ne possède pas de direction privilégiée, et donc le descripteur  $A$  sera également faible. On remarque que l'auteur propose une division par  $N_s$  de chaque écart-type à une échelle donnée. Ceci permet de rendre le descripteur indépendant de l'énergie et donc du contraste dans l'image.

Dans le cas d'un fond homogène, on constate sur la Figure 3.3.(a) et (d) qu'aucune orientation n'est discernable ni sur la texture du fond ni dans le domaine spectral (spectre concentré à l'origine, *i.e.* au centre de l'image). Ainsi, les filtres orientés répondront que de manière faible dans toutes les directions et échelles. De ce fait, l'anisotropie  $A$  sera également faible.

Contrairement aux rides de sables et aux fonds homogènes, les fonds marins complexes auront tendance à fournir des réponses énergétiques bien plus importantes pour différentes orientations et échelles des filtres. En effet, on peut constater à partir des Figures 3.3.(c) et (f) que de nombreuses roches, aux tailles variées et dont les bords ont des orientations diverses, peuvent être aperçues. Le spectre de cette imagerie n'est alors ni concentré autour d'une orientation spécifique (comme dans le cas de rides de sable), ni autour de l'origine (comme dans le cas d'un fond homogène). Ainsi, les écarts-types en orientations seront plus faibles et le descripteur  $A$  également.

Puisque l'anisotropie  $A$  sera faible dans le cas de fonds homogènes et complexes, l'auteur propose de définir un coefficient supplémentaire pour pouvoir les discerner : la complexité. Il le

définit comme la moyenne des réponses énergétiques sur les différentes orientations et échelles :

$$C = \frac{1}{S} \sum_{s=1}^S N_s. \quad (3.4)$$

Comme nous l'avons vu précédemment, sur une texture complexe, les filtres orientés répondent fortement pour différentes orientations et échelles. Ainsi, la moyenne des réponses énergétiques du banc de filtres sur l'ensemble des échelles sera importante. Au contraire, pour une zone de fond anisotrope ou homogène, la valeur de  $C$  sera faible. Ainsi selon l'auteur, ces deux descripteurs sont adaptés à la recherche des fonds anisotropes et complexes qui rendent délicates les opérations de chasses aux mines.

Suite à ces définitions, plusieurs études ont suivi afin d'analyser l'anisotropie et la complexité de la texture des fonds marins, comme celles de E. Fakiris (Université de Patras - Grèce) [2], de D.P. Williams (CMRE - Italie) [3] ou encore M. Geilhufe (FFI - Norvège) [81]. Toutes proposent les mêmes définitions pour ces deux descripteurs, mais avec des filtres et des transformées différentes.

### 3.3.2.2 Estimation des descripteurs

A travers la littérature, plusieurs méthodes ont été proposées pour estimer ces coefficients. La première d'entre-elles est celle utilisée par O. Daniell dans l'étude initiale [82], qui repose sur l'utilisation des filtres de Gabor.

#### Les filtres de Gabor

Dans le domaine spatial, un filtre de Gabor est une sinusoïde modulée par une gaussienne. Un tel filtre est généralement exprimé de la façon suivante :

$$g(x, y) = \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right) \cos(-2\pi(u_0x + v_0y)). \quad (3.5)$$

Le filtre  $g$  est donc une sinusoïde centrée autour de la fréquence  $\mathbf{u} = (u_0, v_0)$  dont l'étendue spatiale est limitée par une enveloppe gaussienne fixée par les paramètres  $\sigma_x$  et  $\sigma_y$  qui représentent les variances suivant les deux axes orthogonaux du plan. La Figure 3.4.(a) propose l'exemple d'un tel filtre en 3D, qui illustre le phénomène d'atténuation de la sinusoïde par la gaussienne.

Le filtre de Gabor est particulièrement bien adapté pour repérer une fréquence spécifique en un lieu donné de l'image. Le lecteur pourra se référer à l'étude de J. Daugman [84] pour de plus amples informations sur les propriétés d'incertitudes en localisation dans les domaines spatiaux et fréquentiels de ces filtres.

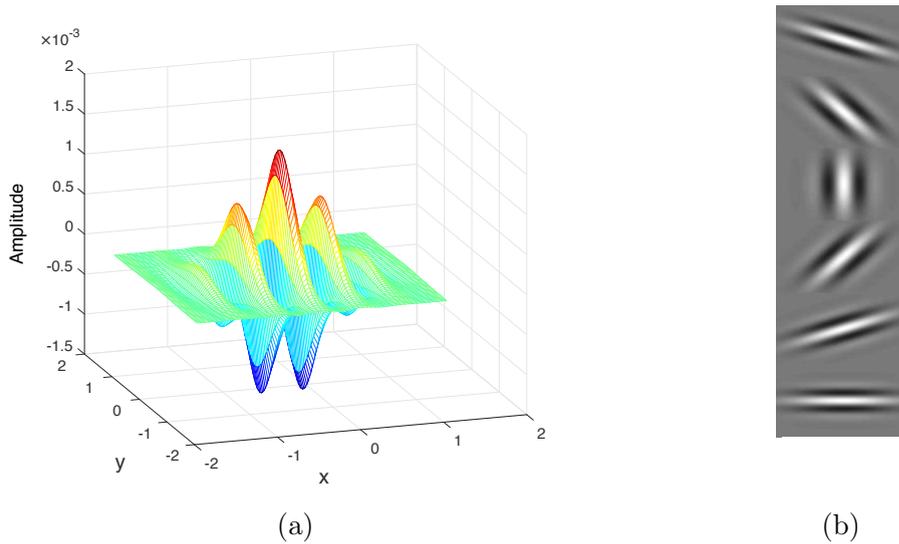


FIGURE 3.4: Illustration d'un filtre de Gabor en 3D (a) et de ses versions orientées (b).

Cependant, l'autre avantage du filtre de Gabor est qu'il est directionnel. Pour une fréquence donnée  $u_0$ , on peut faire intervenir un angle  $\theta$  dans la sinusoïde de l'équation 3.5 :

$$g(x, y) = \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right) \cos(-2\pi u_0 (\cos(\theta)x + \sin(\theta)y)). \quad (3.6)$$

Cet angle définit alors l'orientation du filtre. La Figure 3.4.(b) illustre différents filtres de Gabor pour 6 valeurs de  $\theta$  distinctes.

Dans le domaine fréquentiel, l'expression d'un tel filtre est également une gaussienne centrée autour de la fréquence fournie par la sinusoïde :

$$G(u, v) = \exp(-2\pi^2(\sigma_x^2(u - u_0)^2 + \sigma_y^2(v - v_0)^2)). \quad (3.7)$$

Ainsi les filtres de Gabor sont des filtres passe-bande, dont la fréquence centrale est fixée par la fréquence de la sinusoïde et dont les caractéristiques de la bande passante (*i.e.* sa largeur) sont définies par la gaussienne modulante. Un pavage du plan fréquentiel peut alors être réalisé en fixant un nombre d'orientations et de fréquences centrales distinctes. La Figure 3.5 illustre l'empreinte d'un ensemble de filtres de Gabor pour 6 orientations (chacune correspondant à une couleur), et pour trois fréquences centrales distinctes. Le paramétrage du banc de filtres consistera alors à définir le taux de recouvrement acceptable entre les filtres.

Les filtres de Gabor ont été fréquemment utilisés pour l'analyse de textures et en particulier pour la description des fonds marins comme dans l'étude de I. Leblond [32]. O. Daniell remarque quant à lui que ces filtres sont relativement bien adaptés pour pouvoir estimer les coefficients d'anisotropie et de complexité. En effet, les rides de sable peuvent être grossièrement modélisées comme des sinusoïdes. De ce fait, si la fréquence et l'orientation du filtre de Gabor épousent parfaitement la structure des rides de sable, sa réponse énergétique sera

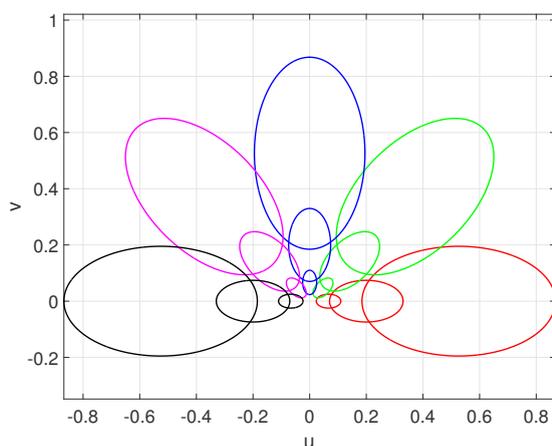


FIGURE 3.5: Emprunte des filtres de Gabor dans le domaine fréquentiel pour 6 orientations (une par couleur) et 3 échelles.

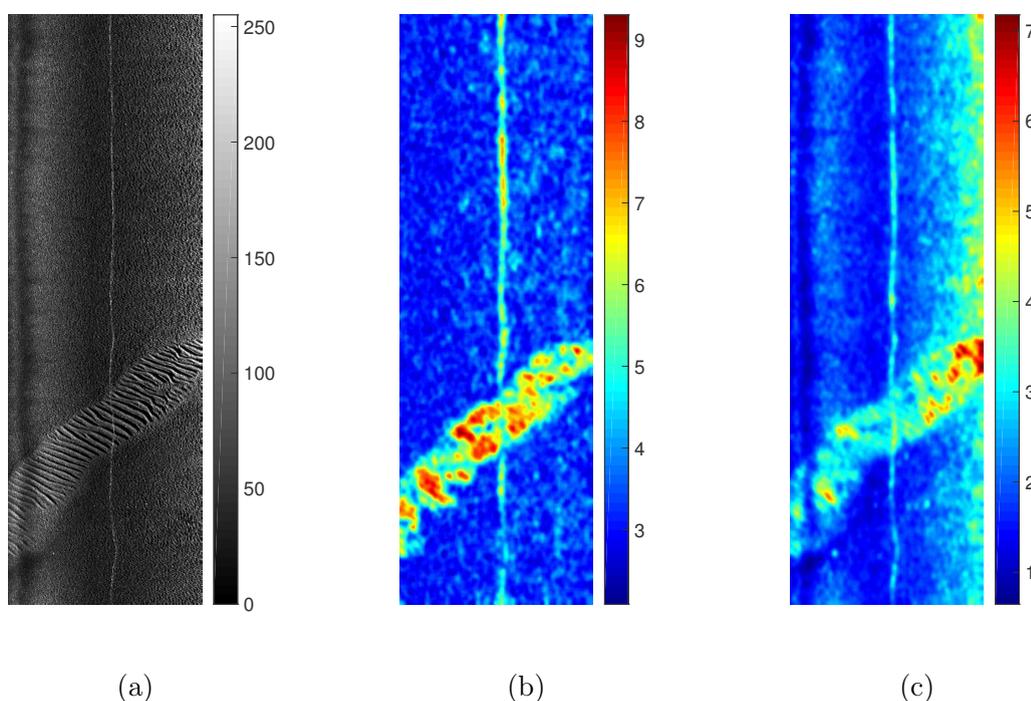


FIGURE 3.6: (a) Exemple d'une image sonar réelle issue d'un sonar latéral Marine Sonic sur laquelle sont extraits deux descripteurs à l'aide des filtres de Gabor : (b) l'anisotropie et (c) la complexité.

très forte, contrairement à celle obtenue dans une direction perpendiculaire. Il relie alors les paramètres de la gaussienne, ainsi que la fréquence de la sinusoïde, aux dimensions des structures qu'il veut mettre en évidence dans les images et discrétise l'ensemble de ces orientations  $\theta \in \{0^\circ, 0^\circ + \Delta\theta, \dots, 180^\circ\}$  avec un pas  $\Delta\theta = 45^\circ$ . Une fréquence centrale importante permettra d'analyser les détails "haute fréquence" (*i.e.* de faibles dimensions), alors qu'une fréquence centrale basse permettra d'analyser les détails "basse fréquence" (*i.e.* de grandes dimensions).

En utilisant les réponses énergétiques de ce banc de filtres, on peut estimer les descripteurs

d'anisotropie et de complexité précédemment définis. Par rapport au paragraphe 3.3.2.1, les indices  $k$  et  $s$  indexent alors les différents angles d'orientation ainsi que les fréquences centrales des filtres de Gabor. La Figure 3.6 propose l'estimation de ces deux descripteurs sur une image sonar réelle issue d'un sonar latéral Marine Sonic (base de données DSTO). Cette image représente une zone de rides de sable (dans la moitié inférieure de l'image) entourée de sable fin homogène. On constate alors que le coefficient d'anisotropie réagit fortement dans cette zone de rides, alors qu'il fournit des valeurs faibles sur les fonds homogènes. Le coefficient de complexité met en évidence la zone de rides de sable de manière moins marquée, et réagit surtout en fin de portée. Dans cette zone, les structures du fond sont légèrement déformées à cause de l'angle de rasance et de l'augmentation du niveau de bruit, ce qui induit une complexification de la texture. On note également que le reflet de la surface (ligne verticale très intense traversant toute l'image) est également considéré comme une structure de l'image fortement anisotrope, et peut polluer l'analyse automatique de l'image.

L'utilisation des filtres de Gabor est reconnue comme étant très gourmande en temps de calcul dans le domaine spatial. Pour pallier ce défaut, il est alors nécessaire de réaliser le filtrage dans le domaine fréquentiel ou de considérer l'approximation de ces filtres de Gabor par les filtres de Haar comme le mentionne O. Daniell [82].

### Les filtres de Haar

En 2D, les filtres de Haar sont des filtres rectangulaires modulés par une fonction créneau de telle sorte que :

$$\psi_{s,k}(x, y) = \begin{cases} 1 & \text{si} & 0 \leq x' < s_1 & \text{et} & -s_2 \leq y' < s_2 \\ -1 & \text{si} & -s_1 \leq x' < 0 & \text{et} & -s_2 \leq y' < s_2 \\ 0 & \text{sinon} & & & \end{cases} \quad (3.8)$$

où  $x' = x \cos(\theta) + y \sin(\theta)$  et  $y' = -x \sin(\theta) + y \cos(\theta)$ . Alors que les paramètres  $s_1$  et  $s_2$  définissent la dimension du filtre (horizontale et verticale), le paramètre  $\theta$  définit son orientation. La Figure 3.7 illustre 4 filtres de Haar selon différentes directions. On peut y remarquer la simplicité de ces filtres, qui ne sont composés chacun que deux 2 rectangles de valeurs 1 (en blanc) et  $-1$  (en noir). Ces filtres agissent comme une différence entre la somme des pixels dans un des rectangles (le blanc par exemple) et la somme des pixels dans le rectangle adjacent (le noir). De ce fait, ce processus peut être très fortement accéléré par l'utilisation du concept d'image intégrale [85]. Les filtres de Haar furent popularisés dans le contexte du traitement d'image par P. Viola et M. Jones [86] dans un but de détection des visages. Dans le domaine de la détection d'objets en imagerie acoustique, les travaux de B. Lehmann [87] et de D.P. Williams [88] proposent également l'intégration de ces filtres.

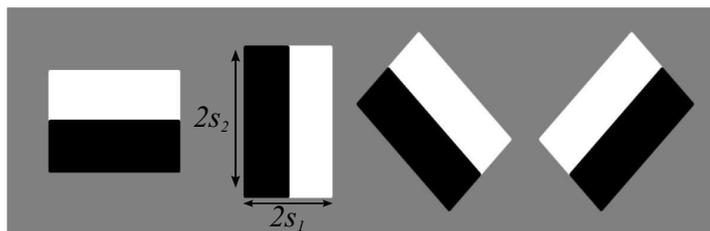


FIGURE 3.7: Différents filtres de Haar orientés selon 4 directions distinctes  $0^\circ$ ,  $90^\circ$ ,  $45^\circ$  et  $135^\circ$ .

Dans le domaine de la classification des fonds, O. Daniell propose dans [82] de comparer l'utilisation de ces filtres dans le calcul des coefficients d'anisotropie et de complexité vis-à-vis des filtres de Gabor. Cependant, l'auteur constate que malgré la vitesse largement supérieure de calcul de ces filtres, ils conduisent à une sélectivité angulaire bien trop faible par rapport aux filtres de Gabor. En effet, les filtres de Haar ne sont implémentés généralement que pour les 4 directions  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  et  $135^\circ$  par raison de simplicité. Pour remédier à ce problème, E. Fakiris propose dans [2] de tirer parti de l'implémentation des caractéristiques pseudo-Haar orientées de C. Messom [89]. Cette dernière permet de calculer les filtres de Haar dans un nombre de directions plus important (12 au total), à l'aide du concept d'image intégrale orientée [90]. Ainsi, E. Fakiris estime les descripteurs d'anisotropie et de complexité de la texture du fond marin de manière rapide et efficace.

La Figure 3.8 est extraite de la publication de E. Fakiris, et propose d'illustrer les différences entre l'utilisation de 4 et de 12 directions distinctes pour les filtres de Haar dans l'estimation des descripteurs d'anisotropie et de complexité. L'image sonar traitée est une image SAS où l'on peut noter la présence d'une large zone de rides de sable au centre de l'image, une zone très hétérogène à gauche de celle-ci, et une zone homogène en haut à droite. On remarque que pour 12 orientations, le descripteur d'anisotropie permet de mettre en évidence de façon claire la zone de rides de sable. Le descripteur de complexité permet quant à lui de pointer une zone complexe à gauche du champ de rides. Ce même descripteur, dans le cas de 4 orientations, sous-estime fortement la complexité de cette zone en ne réagissant uniquement que sur le bord des rides de sable.

Ce résultat est logique puisque les coefficients d'anisotropie et de complexité reposent sur l'estimation de moyennes et de variances selon les angles d'analyse. Ainsi, plus le nombre d'échantillons sera grand et plus ces informations statistiques seront alors représentatives. De même, le voisinage de chaque pixel sera alors plus finement analysé.

Bien que les filtres de Gabor ou de Haar permettent de mettre en évidence les zones anisotropes et complexes d'une image sonar, il demeure que leurs paramétrages en orientations et en échelles n'appartiennent à aucune stratégie formelle de décomposition d'images. Pour y remédier, certaines études abordent de manière plus structurée le problème à l'aide de l'analyse par ondelettes.

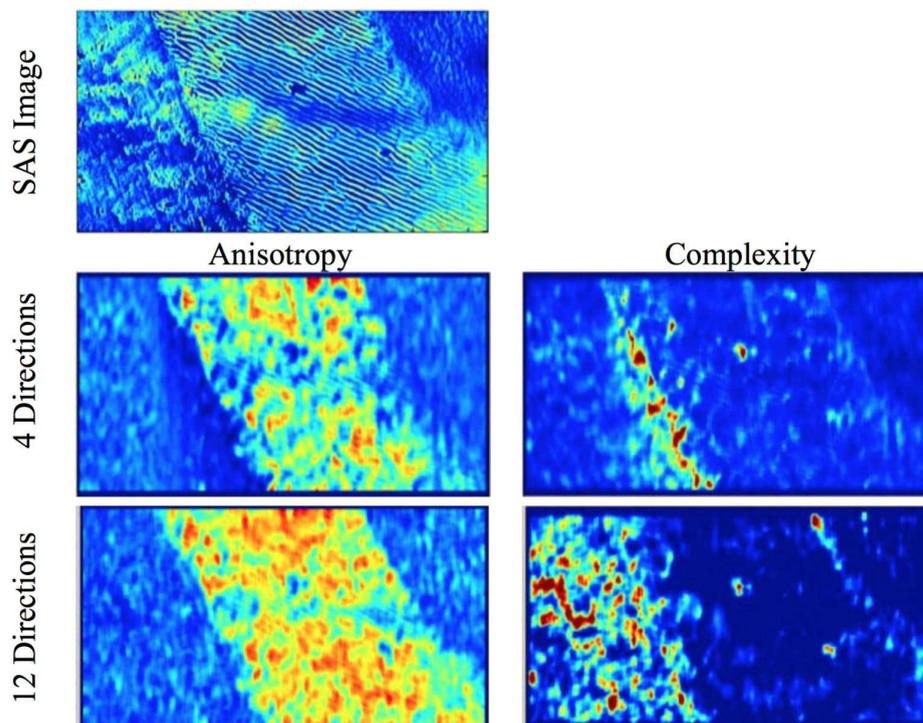


FIGURE 3.8: Résultats des descripteurs d'anisotropie et de complexité extraits de l'étude de E. Fakiris [2], basée sur l'utilisation des filtres de Haar. Ces résultats sont proposés pour 4 et 12 orientations distinctes de ces filtres.

### Analyse par ondelettes complexes

L'analyse par ondelettes fournit un cadre formel de décomposition de l'image pour différents niveaux de résolutions. L'idée proposée par M. Geilhufe dans [81] est donc d'utiliser ce cadre théorique et les transformées en ondelettes discrètes (DWT - de l'anglais *Discret Wavelet Transform*) pour estimer les descripteurs d'anisotropie et de complexité. Par la suite, O. Daniell a également proposé dans [4] de les estimer à l'aide des coefficients en ondelettes mais son approche utilise la transformée en ondelettes complexes par la méthode de l'arbre dual (DTCWT - de l'anglais *Dual Tree Complex Wavelet Transform*). Cette méthode permet d'éviter les problèmes récurrents liés à la transformée en ondelettes discrète 2D classique. En effet, cette dernière souffre de 4 limitations fondamentales :

- **Oscillation** : Les ondelettes étant des filtres passe-bande, leurs applications sur des points singuliers (comme des bords par exemple) mène à une réponse oscillante qui complique les traitements d'extraction de ces points ou la modélisation des signaux comme le montre H. Choi dans [91]. Alors qu'il est largement reconnu que des points singuliers fournissent de forts coefficients d'ondelettes, I. Selesnick montre que ce n'est pas vrai dans le cas général en fournissant un contre-exemple [92].

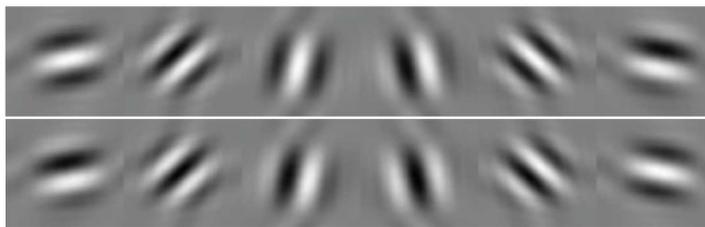


FIGURE 3.9: Orientation des ondelettes réelles (en haut), et imaginaires (en bas).

- **Repliement spectral** : Les algorithmes de décomposition en ondelettes sont pour la grande majorité basés sur l’algorithme proposé par S. Mallat dans [93]. Le processus de décomposition en ondelettes discrètes fait alors intervenir des opérations de sous-échantillonnage et sur-échantillonnage séparées par des filtrages passe-bas et passe-haut. Cependant, le théorème de Shannon stipule que le sous-échantillonnage des signaux va de pair avec un repliement spectral. Ce repliement apparaît si la fréquence d’échantillonnage n’est pas deux fois supérieure à la fréquence maximale contenue dans le signal. Bien qu’il soit reconnu que le procédé de transformée inverse en ondelettes permet de supprimer cet effet de repliement, ce dernier n’est vrai que si les coefficients en ondelettes n’ont pas été modifiés ou filtrés. Dans le cas contraire, la reconstruction du signal fera intervenir des artéfacts.
- **Manque d’invariance par translation** : Ce phénomène résulte du repliement spectral mentionné précédemment. En effet, une transformée est dite invariante par translation si l’énergie totale dans une sous-bande est indépendante d’une légère translation du signal initial. Or, le repliement spectral introduit par le sous-échantillonnage induit naturellement cette dépendance à la translation de la transformée en ondelettes discrètes.  
De plus, autour des points singuliers, J. Romberg remarque qu’un léger décalage du signal perturbe les coefficients en ondelettes [94].
- **Manque de directivité** : La transformée en ondelettes discrètes se réalise à l’aide de filtres séparables selon les lignes et colonnes de l’image. Ces derniers sont des filtres réels qui ne permettent pas de distinguer les fréquences positives des fréquences négatives contenues dans l’image. Ainsi, à chaque niveau de résolution, les sous-bandes contiennent des fréquences positives et négatives qui font de la décomposition en ondelettes discrètes un outil très pauvre en directivité comme le mentionne M. Victor dans [95].

La solution pour remédier à ces limites a été proposée N. Kingsbury dans [96]. Ce dernier a observé que la transformée de Fourier ne souffrait d’aucun des problèmes cités ci-dessus. Il remarque alors que la majeure différence entre la transformée en ondelettes et la transformée de Fourier réside dans l’expression complexe de cette dernière.

En effet, la transformée de Fourier n’est rien d’autre qu’une projection du signal sur une base de fonctions sinusoïdales complexes. Il propose alors de définir les ondelettes complexes

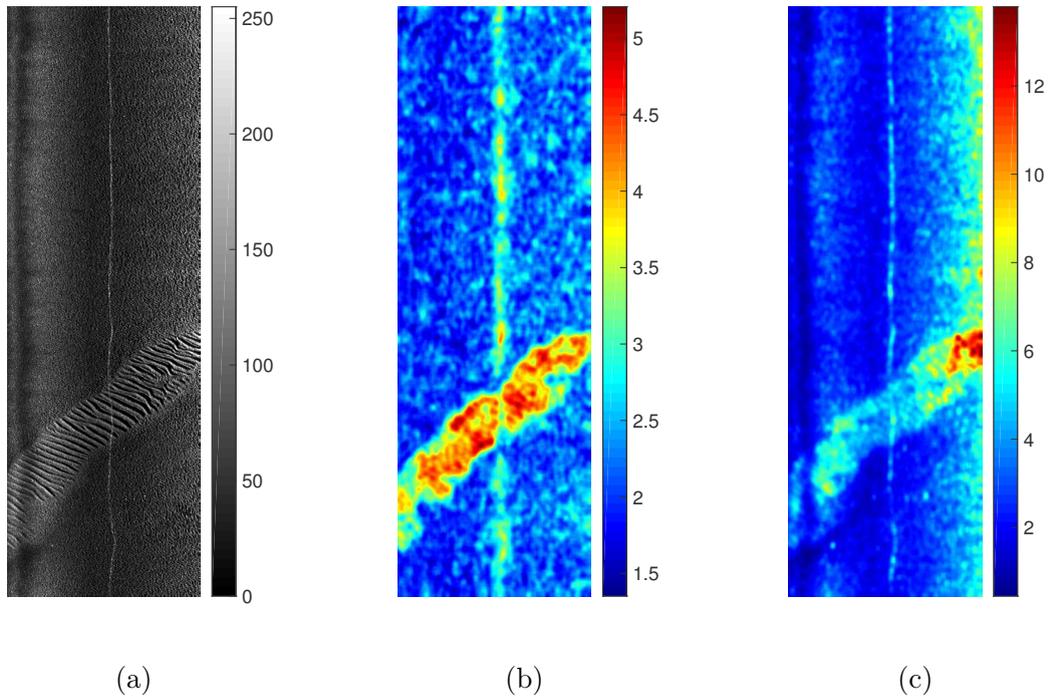


FIGURE 3.10: Exemple d'une image sonar réelle issue d'un sonar latéral Marine Sonic (a) sur laquelle sont extraits deux descripteurs à l'aide de la DTCWT : l'anisotropie (b) et la complexité (c).

$\psi^{\mathbb{C}}(t)$  qui seront composées d'une partie réelle et imaginaire telles que :

$$\psi^{\mathbb{C}}(t) = \psi^{\mathbb{R}}(t) + j\psi^{\mathbb{I}}(t), \quad (3.9)$$

avec l'unique condition que la partie imaginaire soit la transformée de Hilbert de la partie réelle (transformée que nous verrons plus précisément au chapitre 4). Cette condition permet alors de remédier aux problèmes de directivité (suppression des composantes à fréquences négatives par la transformée de Hilbert), et d'oscillation des ondelettes [92] par l'utilisation du module des coefficients en ondelettes complexes.

Les deux limites fondamentales restantes sont quant à elles résolues par l'utilisation d'un arbre dual. C'est-à-dire que la partie réelle de l'ondelette complexe réalise une décomposition en ondelettes discrètes classique, alors que la partie imaginaire de l'ondelette réalise également une décomposition en ondelettes discrètes de manière parallèle. Ce processus permet de rendre globalement le traitement invariant par translation selon M. Victor [95].

Ici, la propriété intéressante de cet arbre de décomposition est qu'il fournit à chaque niveau de résolution 12 sous-bandes (6 sous-bandes dans la branche réelle et 6 sous-bandes dans la branche imaginaire). Ces filtres sont représentés sur la Figure 3.9. On remarque que la diversité en orientations de ces ondelettes est plus importante que dans le cas des ondelettes

réelles, qui ne sont généralement que de trois (horizontale, verticale, et diagonale). Cette remarque renforce l'idée qu'une telle décomposition est adaptée à l'estimation des coefficients d'anisotropie et de complexité.

L'implémentation de la transformée en ondelettes complexes d'une image  $I(x, y) \in \mathbb{R}$ , avec  $x \in \{1, \dots, N\}$  et  $y \in \{1, \dots, N\}$ , donne différentes sous-bandes  $\psi_{s,k}^{\mathbb{C}} \in \mathbb{C}$ . Chacune de ces sous-bande est caractérisée par un indice  $s$  correspondant au niveau de résolution ( $s = 1, \dots, S$ ) ainsi que par un indice  $k$  correspondant à son orientation (parmi  $15^\circ, 45^\circ, 75^\circ, 105^\circ, 135^\circ, 165^\circ$ ). Chaque sous-bande est ensuite post-traitée afin qu'elles soient toutes de taille  $N/2 \times N/2$ . L'énergie des coefficients d'ondelettes peut être définie par la relation :

$$E_{s,k}(x, y) = |\psi_{s,k}^{\mathbb{C}}(x, y)|^2 = \psi_{s,k}^{\mathbb{R}}(x, y)^2 + \psi_{s,k}^{\mathbb{I}}(x, y)^2. \quad (3.10)$$

A partir de ces coefficients, on peut alors évaluer les descripteurs d'anisotropie et de complexité selon les expressions explicitées dans le paragraphe 3.3.2.1.

Une illustration de ces descripteurs est proposée à la Figure 3.10 sur la même image sonar que celle utilisée à la Figure 3.6. On remarque que l'utilisation de la DTCWT permet au coefficient d'anisotropie d'avoir une réponse plus homogène sur l'ensemble de la zone de rides de sable, par rapport à ce qu'il laissait entrevoir lorsqu'il était estimé par biais des filtres de Gabor. Cependant, cette méthode d'estimation semble fournir plus d'artéfacts sur les fonds homogènes et être plus sensible au motif réalisé par reflet de la surface. Le coefficient de complexité donne quant à lui sensiblement les mêmes résultats.

Parmi les trois méthodes proposées pour estimer ces descripteurs d'anisotropie et de complexité de la texture du fond, c'est cette dernière méthode, basée sur la transformée en ondelettes complexes, qui semble la plus aboutie. Elle permet d'intégrer l'analyse de la complexité de l'image dans le cadre multi-résolution, tout en fournissant un ensemble d'orientations assez grand pour rendre les définitions de l'anisotropie et de la complexité statistiquement plus significatives.

### 3.3.3 Exploitation des descripteurs environnementaux dans l'ATR

Dans ce qui précède, nous avons présenté les différentes approches proposées par la littérature pour estimer des descripteurs environnementaux. Dans cette partie, nous relatons les travaux qui ont été développés pour intégrer et exploiter ces informations dans les chaînes ATR.

Du fait de la nature récente de ces recherches, cette problématique n'a été que très peu abordée dans la littérature. On recense uniquement deux approches, l'une qui vise une intégration de ces descripteurs pour améliorer la phase de classification des mines sous-marines,

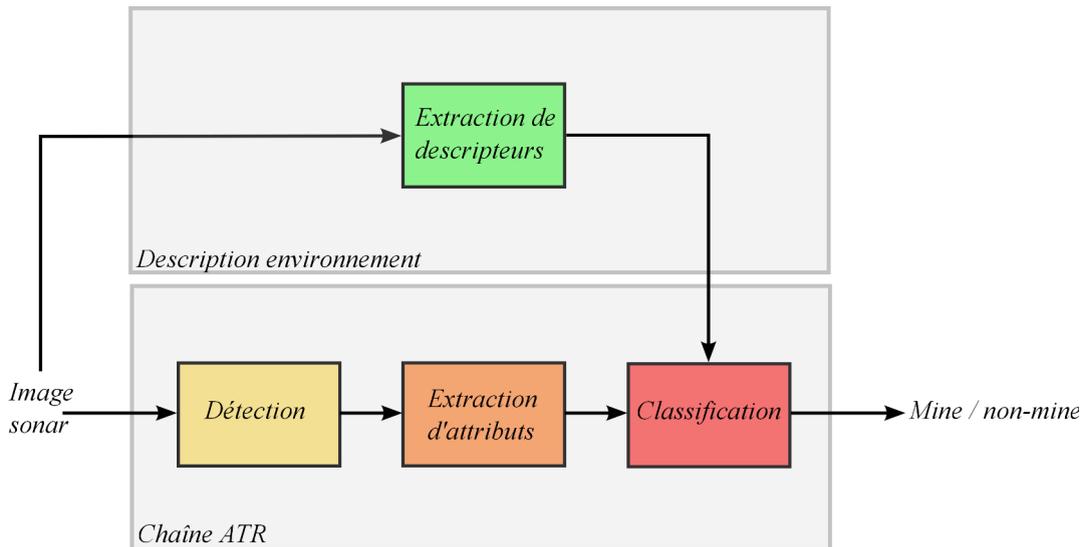


FIGURE 3.11: Schématisation de la stratégie d'intégration mise en place par D.P. Williams pour intégrer des descripteurs environnementaux dans le processus ATR.

et la seconde qui propose de réduire le taux de fausses alarmes d'une chaîne ATR à l'aide de ces derniers.

### 3.3.3.1 Intégration en classification de mines sous-marines

D.P Williams fut le premier à intégrer des descripteurs environnementaux au sein des chaînes ATR [3]. Le schéma présenté sur la Figure 3.11 illustre la stratégie alors mise en place par l'auteur. On constate que la description de l'environnement marin (*i.e.* le calcul de l'anisotropie et de la complexité) se fait en parallèle des étapes de détection et d'extraction d'attributs pour n'intervenir qu'à la fin du traitement lors de la phase de classification (vue au paragraphe 3.2.3). Dans cette étude, D.P. Williams s'intéresse donc uniquement à l'amélioration de la phase de classification des mines sous-marines par l'intégration de données environnementales. Il suppose que les étapes de détection et d'extraction d'attributs sont préalablement définies et indépendantes de cette description.

Concernant le problème spécifique de classification de cibles sous-marines, il part du principe que la plupart des algorithmes d'apprentissage statistiques se basent sur l'hypothèse implicite que les données d'entraînement et de test sont générées par une même distribution statistique. Ainsi, les données labellisées qui permettent l'entraînement d'un classifieur sont supposées représentatives des données non-labellisées qui devront être traitées.

Or, nous avons vu jusqu'à présent qu'en imagerie sonar, la variabilité de l'environnement marin ne peut justifier une telle hypothèse d'homogénéité des statistiques suivant les zones géographiques ou les conditions d'acquisition. Ainsi, D.P. Williams considère que la phase de classification doit tenir compte de l'environnement marin, et par conséquent, que le classifieur utilisé ne doit pas être le même suivant le type de fond marin. Il tente alors de démontrer que les

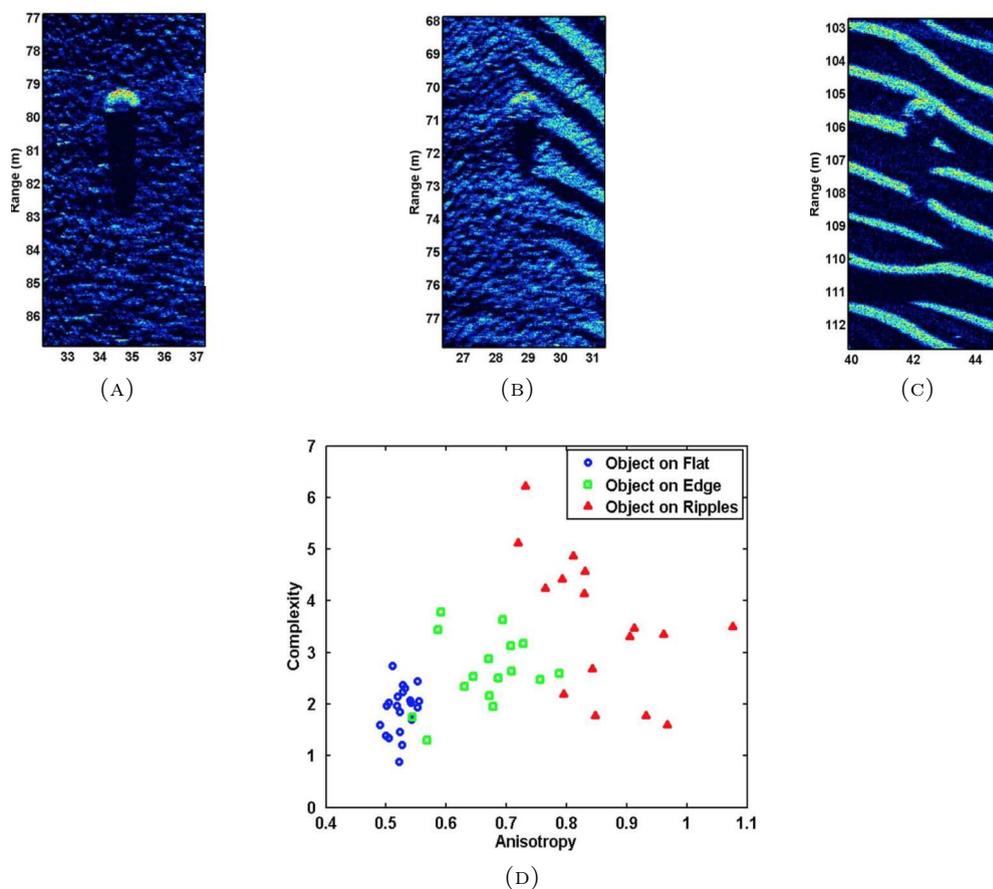


FIGURE 3.12: Exemples de trois environnements marins distincts dans lesquels une mine sous-marine est positionnée : (a) sur un fond homogène, (b) sur une frontière entre un fond homogène et un fond de rides de sable, et (c) sur un fond de rides de sable. La distribution des coefficients d'anisotropie et de complexité autour de ces trois cibles est représentée en (d) (résultats extraits de [3]).

coefficients d'anisotropie et de complexité du fond marin peuvent être utiles pour distinguer les différentes conditions environnementales et améliorer cette phase de classification. Par exemple, il propose de représenter la distribution de ces deux descripteurs autour d'une même cible mais posée dans trois environnements distincts à savoir : 1) sur un fond homogène, 2) sur un fond composé de rides de sable et 3) sur une frontière entre un fond homogène et un fond de rides de sable. Nous retrouvons sur la Figure 3.12 l'illustration de ces trois scénarios (a),(b),(c), ainsi que la distribution des descripteurs d'anisotropie et de complexité associée (d). On y constate qu'une séparation peut être trouvée à partir de ces trois distributions permettant de discerner un objet posé sur un fond homogène d'un autre posé sur des rides de sable ou sur une frontière entre les deux.

Pour intégrer ces deux descripteurs dans la classification, D.P. Williams propose d'utiliser, non pas un seul classifieur, mais un ensemble de  $N_C > 2$  classifieurs, chacun spécialisé pour une condition environnementale particulière. Il définit **une condition environnementale** comme étant **un couple** de valeurs  $(A, C)$  (*i.e.* un couple (Anisotropie, Complexité))

donné. Cependant, puisque les deux coefficients  $A$  et  $C$  sont continus, il est nécessaire de découper leur dynamique en  $n$  ensembles. Par exemple, dans le cas particulier où ces derniers sont discrétisés en un même nombre  $n$  de sous-ensembles, on obtient alors  $n^2$  conditions environnementales et par conséquent un nombre classifieurs  $N_C = n^2$ . L'apprentissage de ces  $N_C$  classifieurs est alors réalisé de manière supervisée à l'aide d'une base d'entraînement. Ainsi, le nombre  $n$  de sous-ensemble est directement relié au nombre de classifieurs à utiliser, et aura une influence sur l'apprentissage de ces derniers. Par exemple, plus  $n$  sera grand (et donc  $N_C$ ), plus les conditions environnementales étudiées deviendront précises. Par conséquent pour chacune d'elles, la quantité d'exemples dans la base d'entraînement sera restreinte et l'apprentissage des classifieurs ne pourra potentiellement pas se faire correctement. De même, si ce nombre est trop petit, il y aura un manque de diversité en termes de conditions environnementales, et l'intérêt de l'approche sera directement impactée. Le nombre de classifieurs à utiliser doit donc être fixé avec soin. L'auteur propose d'apprendre ce nombre de manière automatique sur la base d'apprentissage, via la maximisation d'une fonction d'entropie.

Outre les descripteurs environnementaux, l'ensemble de ces classifieurs sont entraînés par une extraction de 27 attributs sur les détections (mélange d'attributs statistiques, géométriques et de Fourier) dont la liste complète peut être retrouvée dans [3].

Lors de la phase de test, un poids est donné à chacun de ces classifieurs en fonction de la condition environnementale de l'objet détecté. Un poids important est alors donné au classifieur correspondant à cette dernière. Au contraire, un poids faible est accordé aux classifieurs très éloigné de la condition environnementale observée pour l'objet détectée. La moyenne pondérée des prédictions de l'ensemble des classifieurs est ensuite réalisée pour estimer si la détection est une mine ou non.

L'auteur présente les performances de l'algorithme sur un ensemble de données réelles issues de différentes campagnes d'acquisition. Pour de plus amples précisions sur l'approche proposée et ses résultats, nous invitons le lecteur à se référer à [3].

### 3.3.3.2 Intégration pour la réduction des taux de fausses alarmes

La seconde méthode fut proposée par O. Daniell dans [4]. Elle se distingue de la méthode précédente par le fait qu'elle considère qu'une chaîne ATR complète (*i.e.* détection - extraction d'attributs - classification) existe déjà. L'objectif est alors de rajouter une phase supplémentaire consistant à filtrer les fausses alarmes de cette dernière. Pour pouvoir adapter ce module de manière simple à n'importe quelle chaîne ATR, le cahier des charges impose qu'aucune connaissance sur les algorithmes de la chaîne ATR n'est accessible. La seule information disponible est alors l'indice de confiance sur la classification des contacts qui en sortent. Comme dans le cas précédent, nous proposons de résumer schématiquement cette méthode dans la Figure 3.13. On y remarque que cette fois, la chaîne ATR n'intègre pas directement d'informations

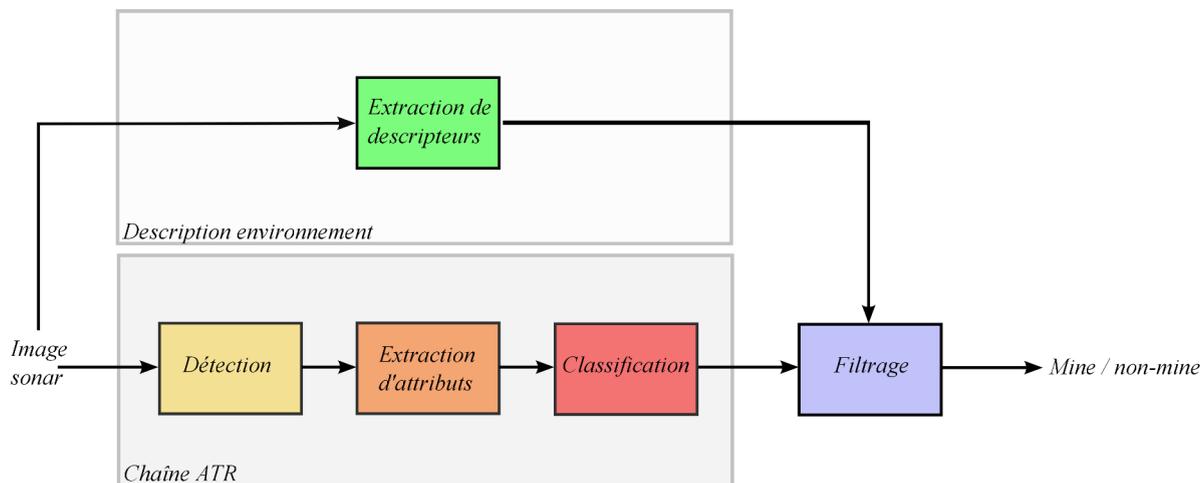


FIGURE 3.13: Schématisation de la stratégie mise en place par O. Daniell pour intégrer des descripteurs environnementaux dans le processus ATR.

environnementales. C'est un processus entièrement décorrélé qui vient s'ajouter à la suite de celle-ci par un filtrage de ses résultats.

La méthode proposée par O. Daniell repose alors sur l'observation que le taux de fausses alarmes d'une chaîne ATR donnée dépend du type de fond marin. L'auteur considère les coefficients d'anisotropie et de complexité comme les descripteurs du fonds. En plus de ces deux coefficients, il ajoute un coefficient d'échelle et de hauteur des structures sous-marines. Alors que ce premier est obtenu à l'aide de la transformée en ondelettes complexes (la même qu'utilisée pour l'estimation de l'anisotropie et de la complexité), le descripteur de hauteur est calculé quant à lui à partir de l'ombre des structures présentes dans l'image. Cette dernière est détectée grâce à un algorithme de segmentation en 3 classes de l'image sonar : écho, ombre, et fond, publié par S. Reed [76]. Il remarque que certaines combinaisons de ces 4 descripteurs sont associées à de forts taux de fausses alarmes. Cependant, il met en évidence que la relation qui existe entre la probabilité de fausses alarmes ( $P_{fa}$ ) et ces descripteurs de la texture du fond, est complexe et fortement non-linéaire. Ainsi, il en conclut que seul un algorithme d'apprentissage supervisé peut permettre de distinguer les fausses alarmes des véritables cibles par l'intermédiaire de tels descripteurs. Il propose alors d'entraîner un algorithme perceptron multi-couches (*i.e.* un réseau de neurones). Afin de générer une base d'entraînement, l'algorithme ATR est appliqué aux données d'une mission où les vérités terrains sont connues. Les descripteurs du fond y sont également calculés. Ainsi, pour chaque détection de la chaîne ATR, il est possible de relier l'indice de confiance qu'elle fournit sur ces sorties, les descripteurs du fonds, et la vérité terrain. La phase d'apprentissage, qui est réalisée par le réseau de neurones, permet alors d'estimer cette relation.

O. Daniell montre que la phase de test de l'algorithme sur des données inconnues est très encourageante. En effet, il mentionne que le taux de fausses alarmes peut être radicalement réduit, jusqu'à 90% dans certains cas. La Figure 3.14 montre un exemple de résultats

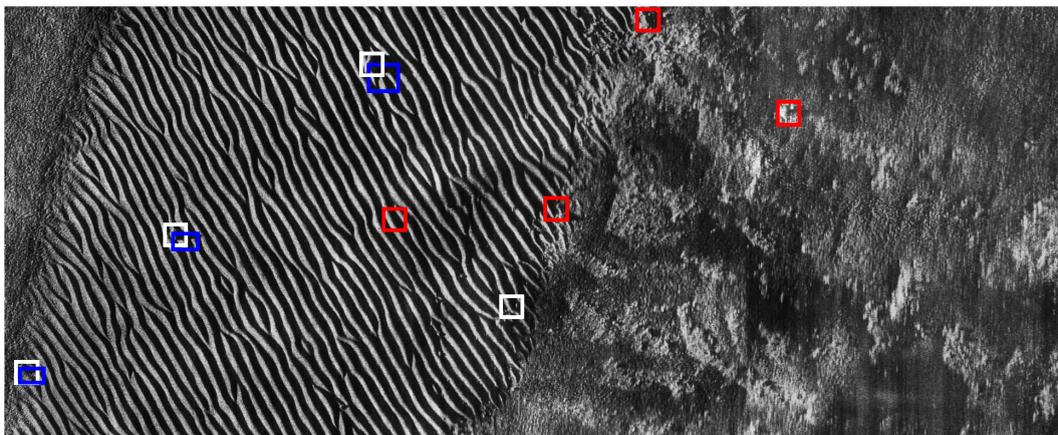


FIGURE 3.14: Exemple d'une image sonar où les résultats d'une chaîne ATR sont illustrés. L'ensemble des rectangles correspond aux détections fournies par cette dernière. Les rectangles bleus définissent les vérités terrains, les rectangles rouges correspondent aux détections filtrées par l'algorithme de O. Daniell, et les blancs sont les détections gardées par ce même algorithme (Résultats extraits de [4]).

de l'algorithme proposé, extrait de la publication de O. Daniell [4]. Sur une image sonar l'algorithme ATR fournit un ensemble de détections représentées par différents rectangles. Alors que les rectangles bleus représentent les vérités terrains, les rectangles rouges représentent les fausses alarmes filtrées par l'algorithme. Les rectangles blancs sont quant à eux les détections que l'algorithme juge comme de véritables détections. On remarque donc qu'une seule fausse alarme n'est pas supprimée par l'étape de filtrage dans cet exemple.

### 3.4 Approche proposée dans cette thèse

A travers ce chapitre, nous avons pu constater que les chaînes ATR sont associées à une littérature conséquente, qui évolue en parallèle des améliorations technologiques.

A l'origine, ces processus cherchaient uniquement à détecter des paires écho-ombre dans les images sonar pour déclarer la présence de mines, ou non, dans ces données. Cependant, nous avons vu que l'environnement marin est un frein majeur aux performances de telles chaînes de traitement. Son intégration dans le processus est alors devenue une priorité, afin que la détection et la reconnaissance de cibles prennent en compte le contexte dans lequel la cible se situe.

L'analyse des solutions proposées par la littérature, pour décrire et caractériser l'environnement marin, nous permet de proposer de nouvelles pistes de recherche.

#### 3.4.1 Position par rapport à la littérature

Tout d'abord, nous proposons d'analyser le contenu d'une image sonar par le biais d'une nouvelle représentation, afin d'en extraire davantage d'informations.

### 3.4.1.1 Nouvelle représentation de l'information

Jusqu'à présent, nous avons pu constater que la littérature propose différentes transformations de l'image afin d'en extraire divers descripteurs, comme l'anisotropie et la complexité.

Cependant, ces transformées nécessitent généralement un ensemble de paramètres qu'il faut gérer et fixer. Par exemple, les bancs de filtres orientés (comme ceux de Gabor ou de Haar) exigent de fixer un nombre donné d'orientations et de dilatations. A notre connaissance, aucune étude paramétrique n'a alors été réalisée lors de la définition des coefficients d'anisotropie et complexité. Il est donc difficile de connaître ceux qui permettront d'analyser au mieux une image sonar.

L'étude par ondelettes possède quant à elle l'avantage d'avoir une notion d'orientation, et d'approche multi-résolution, directement intégrées dans sa propre théorie. Il y a donc moins de paramètres à gérer. Comme nous avons pu le voir, l'approche proposée par O. Daniell [4] utilise une définition complexe des ondelettes sensée corriger de nombreuses limites de l'analyse par ondelettes réelles 2D. Cependant, certains défauts de l'approche peuvent être soulignés. En effet, cette méthode se base sur la transformée de Hilbert 2D, dont la définition a fait l'objet de nombreuses études mais de peu de consensus au cours des dernières décennies (nous y reviendrons dans le chapitre 4). De plus, le nombre limité d'orientations pour les ondelettes est toujours un frein à une étude générale de l'image.

Dans cette étude, nous voulons proposer une autre représentation de l'image, qui soit avec un nombre de paramètres très limité afin qu'elle s'adapte facilement aux différentes données issues de systèmes sonar latéral variés. De ce fait, il serait intéressant d'avoir une analyse isotrope de l'image (*i.e.* sans définir un nombre d'orientations pour analyser l'image). De plus, comme nous avons pu le constater, une image sonar n'est pas constituée que d'informations énergétiques, on y observe également des structures aux formes diverses dont les géométries sont des informations significatives et pertinentes. Il serait alors intéressant que cette nouvelle représentation en tienne compte.

Dans la littérature, un outil est connu pour extraire d'un signal des informations énergétiques mais également structurelles sur la forme d'un signal, c'est le concept de signal analytique. Pour analyser des données sous forme d'images, nous proposons alors d'étudier son extension isotrope multi-dimensionnelle connue sous le nom de **signal monogène**. Nous verrons dans le chapitre 4 qu'il permet de représenter un signal 2D en trois informations locales (*i.e.* en chaque pixel) et orthogonales à savoir respectivement une information **énergétique**, **structurelle** et **géométrique**. De plus, il permet de tenir compte de la variabilité des structures dans l'image par une approche **multi-échelle**.

L'avantage d'un tel outil serait donc de fournir une représentation plus complète de l'image, sans l'intervention d'un paramétrage excessif.

Cependant, une telle représentation ne résout pas le problème de description des fonds marins, il permet uniquement une meilleure représentation de l'information 2D. Ainsi, il est nécessaire de trouver un moyen de l'utiliser à bon escient afin d'en extraire ce que l'on désire.

### 3.4.1.2 Descripteurs environnementaux plus adaptés

Les descripteurs d'anisotropie et de complexité des fonds marins, proposés par la littérature, se basent tous sur une même définition (cf. paragraphe 3.3.2.1). Cependant, cette définition utilise uniquement les réponses énergétiques des filtres orientés, ou des coefficients en ondelettes. Aucune autre considération qu'énergétique n'est alors intégrée.

De ce fait, ces coefficients ont une dynamique qui dépend directement de celle de l'image. C'est à dire qu'ils ne peuvent être interprétés que pour un seul type de données à la fois. De plus, cette interprétation est rendue difficile par le fait que leurs dynamiques ne sont pas connues *a priori*. En effet, ils peuvent prendre n'importe quelle valeur dans un intervalle  $[0, +\infty[$ . Si l'on possède une base de données suffisamment représentative de la variété des fonds marins, un algorithme peut s'y adapter. Sinon, il sera très difficile de rendre significatifs de tels descripteurs, y compris pour des opérateurs sensés utiliser l'algorithme.

Au contraire, nous proposons dans ces travaux de redéfinir ces descripteurs d'anisotropie et de complexité en les dotant de dynamiques significatives aussi bien pour un opérateur humain que pour un traitement automatique. De plus, nous les compléterons par l'ajout d'un descripteur sur l'homogénéité du fond marin. En effet, puisque les fonds homogènes sont reliés à une absence de structures sous-marines, et en particulier de mine, ils peuvent être interprétés comme des fonds où, non-seulement la navigation des bâtiments de surface est sécurisée, mais également où les opérations de chasse aux mines sont inutiles. Pour définir ces trois nouveaux descripteurs, nous utiliserons un concept qui n'a jamais été introduit en imagerie sonar : **la dimension intrinsèque**. Entre autres, nous verrons que l'utilisation de ce concept permet de fournir des indices de confiance bornés entre 0 et 1 concernant l'homogénéité, l'anisotropie et la complexité d'une image, et ce, en chaque pixel (*i.e.* localement). De plus, ces indices peuvent être estimés à l'aide des informations issues du signal monogène. Par conséquent, il reposera sur l'architecture multi-échelle proposée. De plus, contrairement aux études précédentes, nous aurons à disposition un indice sur l'homogénéité du fond marin. Ainsi, nous pourrions quantifier l'absence de structure sous-marine, et par conséquent, l'absence de mine dans une région donnée du fond. Des zones de navigation sécurisée (pour les différents bâtiments de surface ou véhicules autonomes) pourront ainsi être définies.

### 3.4.1.3 Chaîne de traitement différente

Outre cette représentation et description originale de l'image sonar, nous proposons dans ces travaux de repenser la structure de la chaîne de traitement ATR. En effet, nous avons pu

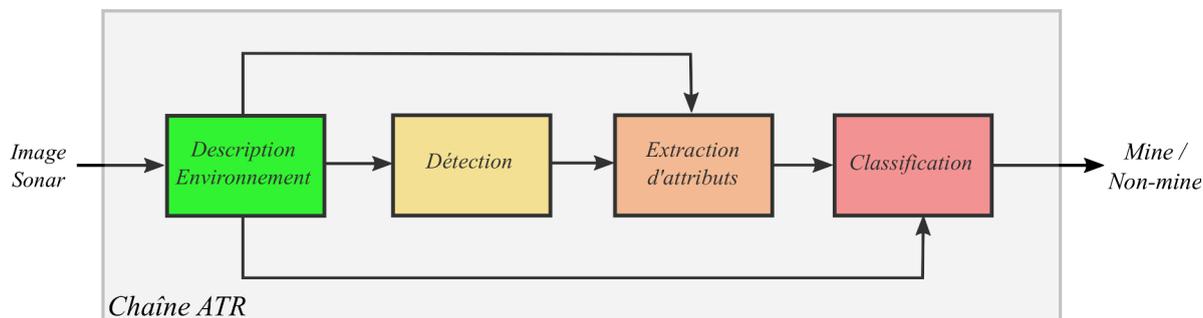


FIGURE 3.15: Schématisation de la stratégie mise en place par O. Daniell pour intégrer des descripteurs environnementaux dans le processus ATR.

constater que cette dernière est constituée de trois blocs successifs correspondant aux phases de détection, d'extraction d'attributs sur ces détections, puis d'une phase de classification.

Actuellement, avec la prise en compte d'informations environnementales, cette architecture évolue légèrement. En effet, on voit apparaître une phase de description de l'environnement que vient se greffer en parallèle de cette chaîne. La littérature propose de la prendre en considération en toute fin de traitement, lors d'un post-traitement. Au contraire, dans ces travaux, nous proposons de considérer cette intégration de manière plus précoce, dès le début de la chaîne ATR. Ainsi, les informations environnementales seront disponibles pour chacune des étapes du processus de reconnaissance de cibles, afin que ces dernières les prennent en compte et s'adaptent au contexte local. L'architecture d'une telle chaîne est proposée à la Figure 3.15.

Ces travaux de thèse étant fortement orientés sur la description de l'information contenue dans l'image sonar, la Figure 3.16 représente schématiquement la phase de description de l'environnement que nous proposons ici, ainsi que des outils qui y interviennent. Les trois blocs successifs qui y apparaissent correspondent respectivement aux trois prochains chapitres de ce manuscrit.

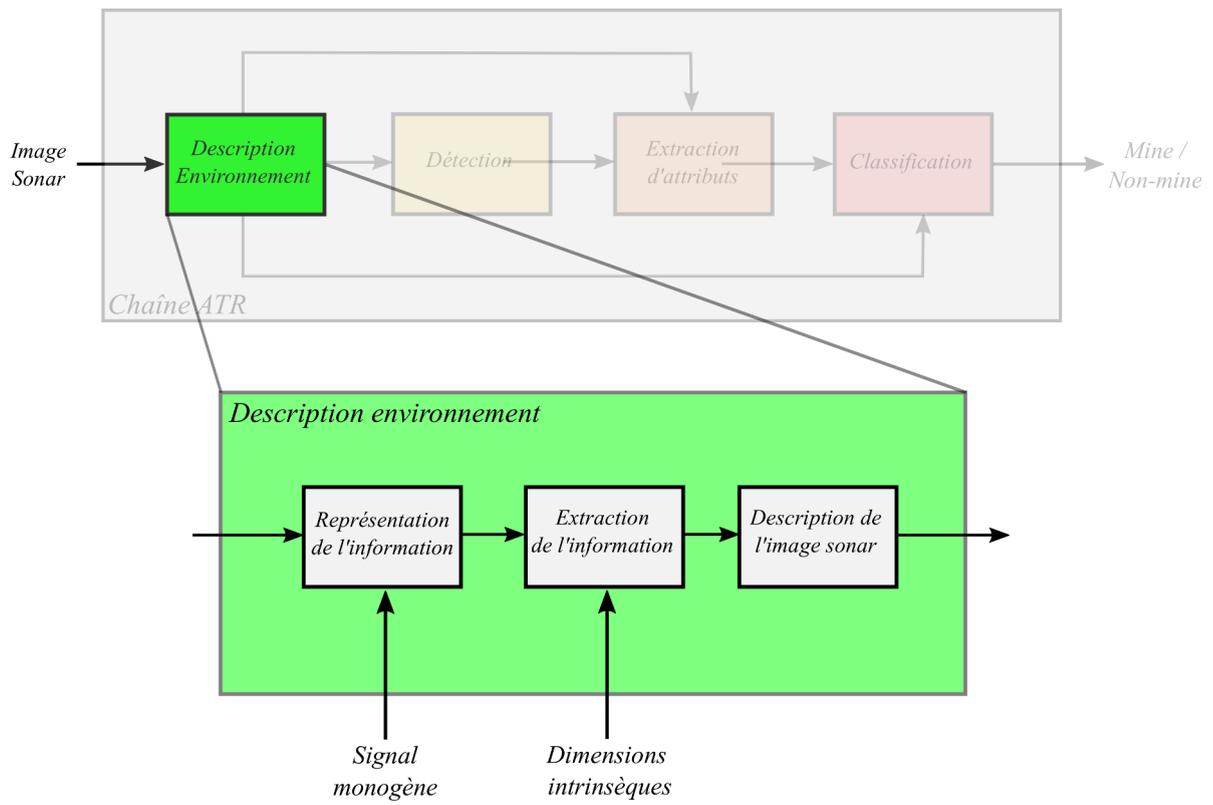


FIGURE 3.16: Zoom sur l'architecture de la description de l'environnement proposée.



# Chapitre 4

## Représentation monogène d'une image

### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>65</b>
<b>4.2</b>	<b>Analyse complexe des signaux 1D</b>	<b>66</b>
4.2.1	La transformée de Hilbert	66
4.2.2	Le signal analytique	67
4.2.3	L'analyse complexe	68
4.2.4	Discussion	70
<b>4.3</b>	<b>Les extensions 2D du signal analytique</b>	<b>71</b>
4.3.1	La transformée de Hilbert 2D	71
4.3.2	Au delà des nombres complexes	72
4.3.3	L'algèbre géométrique	72
<b>4.4</b>	<b>Définition du signal monogène</b>	<b>73</b>
4.4.1	La transformée de Riesz	73
4.4.1.1	Définition	73
4.4.1.2	Propriétés	75
4.4.1.3	Application à une onde sinusoïdale 2D	76
4.4.2	Le signal monogène	78
4.4.2.1	Première définition	78
4.4.2.2	Application à une image synthétique	79
4.4.2.3	Discussion	80
4.4.3	Définition générale du signal monogène	81
<b>4.5</b>	<b>Signal monogène et Analyse multi-échelle</b>	<b>83</b>
4.5.1	Théorie de l'espace d'échelle	83
4.5.2	Espace d'échelle de Poisson	86
4.5.2.1	Les noyaux de Poisson 2D	86
4.5.2.2	Noyaux de Poisson et théorie de l'espace d'échelle	88

4.5.2.3	L'espace d'échelle de Poisson . . . . .	90
4.5.3	Analyse multi-échelle . . . . .	91
4.5.3.1	Introduction aux différences de Gaussiennes . . . . .	92
4.5.3.2	Application aux différences de Poisson . . . . .	97
4.5.4	Représentation monogène multi-échelle . . . . .	100
4.5.4.1	Définition . . . . .	100
4.5.4.2	Application à une image synthétique . . . . .	101
<b>4.6</b>	<b>Application à l'image sonar . . . . .</b>	<b>103</b>
4.6.1	Paramétrage de l'analyse multi-échelle . . . . .	103
4.6.2	Représentation monogène d'une image sonar latéral . . . . .	107
4.6.3	Application à la classification des fonds marins . . . . .	111
4.6.3.1	Présentation de l'approche . . . . .	111
4.6.3.2	Matrices de Confusion . . . . .	112
4.6.3.3	Classification d'une image sonar . . . . .	113
<b>4.7</b>	<b>Résumé du Chapitre . . . . .</b>	<b>115</b>

---

## 4.1 Introduction

Dans le chapitre 2, nous avons vu que l'environnement marin apparaissait comme une information complexe et variable sur une image sonar. De fait, le traitement automatique d'une image sonar est une tâche difficile, qui nécessite l'utilisation d'outils d'analyse divers comme nous avons pu le constater dans le chapitre 3. A travers ce chapitre, nous proposons d'introduire une nouvelle représentation de l'information contenue dans une image, appelée signal monogène. Ce concept est apparu au début des années 2000 pour proposer une extension multi-dimensionnelle au signal analytique. En particulier, le signal monogène permet alors de séparer de manière orthogonale les caractéristiques énergétiques, structurelles et géométriques d'une image, le tout dans un cadre multi-échelle.

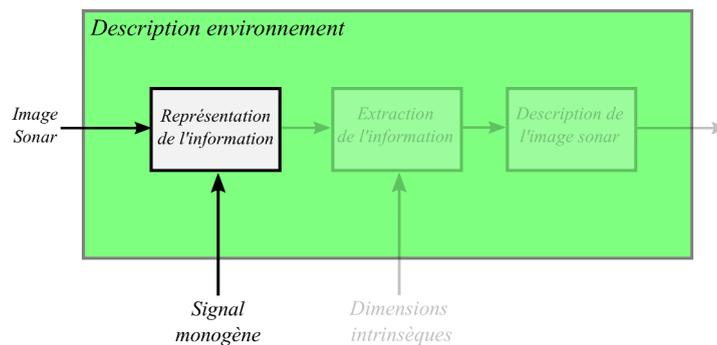


FIGURE 4.1: Position de la représentation monogène dans le processus.

Cette représentation de l'image est la première phase de notre processus de description automatique d'une image sonar, comme schématisé à la Figure 4.1. Cette étape permettra ensuite d'extraire et de décrire l'environnement marin contenu dans cette image.

Ce chapitre est organisé de la manière suivante. Afin de mieux comprendre le concept de signal monogène, nous proposons d'introduire au préalable la notion de signal analytique 1D dans la section 4.2. Après avoir mentionné les différentes propositions d'extension de ce concept pour des signaux 2D ainsi que leurs limites dans la section 4.3, nous verrons dans la section 4.4 que c'est aujourd'hui le signal monogène qui est considéré comme l'extension multi-dimensionnelle du signal analytique. A la section 4.5, nous verrons qu'il est possible d'intégrer le signal monogène dans le cadre de l'analyse multi-échelle afin de tenir compte des différents niveaux de détails dans l'image. Finalement, dans la section 4.6, nous appliquerons ce concept aux images sonar afin de discuter de son exploitation dans nos travaux.

## 4.2 Analyse complexe des signaux 1D

Dans les années 1940, la volonté de trouver un outil permettant d'évaluer précisément les grandeurs instantanées d'un signal a conduit D. Gabor puis J. Ville [97, 98] à la définition du *signal analytique*. Cette représentation complexe d'un signal réel obtenue à l'aide de la transformée de Hilbert (présentée au paragraphe suivant) fournit des informations d'amplitude instantanée et de fréquence instantanée sur le signal. Ces propriétés ont popularisé son utilisation dans de nombreux contextes comme la modulation des signaux (en amplitude (AM) ou en fréquence (FM)) ou l'analyse des signaux (comme par exemple la caractérisation d'électro-cardiogrammes dans le domaine médical ou encore la détection de cibles dans les signaux radar).

### 4.2.1 La transformée de Hilbert

La transformée de Hilbert tient son nom du mathématicien David Hilbert. Elle est apparue dans la communauté des mathématiques comme la relation entre la partie réelle et imaginaire des fonctions dites *holomorphes* (ou *analytiques*). Elle se définit de la manière suivante :

**Definition 4.1.** La transformée de Hilbert d'une fonction  $f : \mathbb{R} \rightarrow \mathbb{R}$  est définie par :

$$f_{\mathcal{H}}(t) = \frac{1}{\pi} \text{v.p} \int_{-\infty}^{+\infty} \frac{f(\tau)}{t - \tau} d\tau, \quad (4.1)$$

où l'opérateur v.p est la valeur principale de Cauchy [99].

En d'autres termes, la transformée de Hilbert d'une fonction  $f$  est obtenue en convoluant cette dernière par la distribution tempérée  $h_1(t) = \frac{1}{\pi} \text{v.p} \frac{1}{t}$ . Cette opération de convolution fournit à la transformée une nature linéaire. Son expression dans l'espace fréquentiel, notée  $F_{\mathcal{H}}$ , est obtenue en multipliant la transformée de Fourier de la fonction  $f$ , notée  $F$ , par la transformée de Fourier de  $h_1$ , notée  $H_1$ . On a alors :

$$F_{\mathcal{H}}(u) = F(u) H_1(u) = -F(u) i \text{sign}(u), \quad (4.2)$$

où  $u$  désigne la variable fréquentielle,  $i$  le nombre complexe imaginaire et  $\text{sign}(u) = \frac{u}{|u|}$  est la fonction signe (avec  $|\cdot|$  qui représente la valeur absolue). L'expression  $H_1$  dans laquelle intervient cette fonction signe permet d'interpréter la transformée de Hilbert comme un processus réalisant une rotation de l'argument des composantes fréquentielles d'un angle de  $\frac{\pi}{2}$  dans le sens trigonométrique pour les fréquences positives, et dans le sens inverse (anti-trigonométrique) pour les fréquences négatives. Le signal et sa transformée de Hilbert sont alors dits en quadrature (terme qui renvoie à la notion d'orthogonalité  $\langle f, f_{\mathcal{H}} \rangle = 0$ ). La Table 4.1 propose quelques

Fonctions	Transformées de Hilbert
$\cos(t)$	$\sin(t)$
$\sin(t)$	$-\cos(t)$
$\delta(t)$	$\frac{1}{\pi t}$
$\text{rect}(t)$	$\frac{1}{\pi} \ln \left  \frac{2t+1}{2t-1} \right $

TABLE 4.1: Quelques exemples classiques de fonctions et leurs transformées de Hilbert associées .

exemples classiques de fonctions et de leurs transformées de Hilbert respectives. Notons ici que sur les exemples de fonctions sinusoïdales proposés, il est possible d'apercevoir la notion de quadrature (*i.e.* un déphasage de  $\frac{\pi}{2}$ ) puisqu'un cosinus est transformé en sinus et inversement. Il faut également noter que l'expression de  $H_1$  montre une singularité en  $u = 0$ . Par conséquent, la transformée de Hilbert n'est pas adaptée à l'étude des signaux de moyenne non-nulle.

Cette introduction de la transformée de Hilbert va maintenant nous permettre de définir la notion de *signal analytique*.

### 4.2.2 Le signal analytique

Dans la communauté du traitement du signal, les signaux mono-dimensionnels (1D) étudiés sont généralement des fonctions d'une variable réelle, typiquement le temps. Pour les analyser dans le domaine temporel, le concept de *signal analytique* a été défini de la manière suivante :

**Definition 4.2.** Soit  $s : \mathbb{R} \rightarrow \mathbb{R}$  un signal réel d'une variable temporelle  $t$ . Son signal analytique  $s_a$  est alors défini par :

$$s_a(t) = s(t) + is_{\mathcal{H}}(t). \quad (4.3)$$

Le signal analytique est donc une fonction complexe, dont la partie réelle est formée par le signal et la partie imaginaire par sa transformée de Hilbert. Ce nombre complexe peut également être exprimé sous forme polaire :

$$s_a(t) = A_s(t) \exp(-i\phi_s(t)), \quad (4.4)$$

où  $A_s(t)$  et  $\phi_s(t)$  sont respectivement appelées l'amplitude et la phase instantanées. Ces deux quantités sont données par les expressions suivantes :

$$A_s(t) = \sqrt{s(t)^2 + s_{\mathcal{H}}(t)^2} \quad (4.5)$$

$$\phi_s(t) = \text{atan} \left( \frac{s_{\mathcal{H}}(t)}{s(t)} \right). \quad (4.6)$$

L'expression du signal analytique peut également être définie dans le domaine fréquentiel :

$$S_a(u) = S(u) + i(-i \text{sign}(u) S(u)) \quad (4.7)$$

$$= 2 S(u) U(u), \quad (4.8)$$

où  $U$  désigne la fonction échelon et  $S$  la transformée de Fourier du signal  $s$ . Cette expression permet de constater que le calcul du signal analytique revient à supprimer les composantes de fréquences négatives du signal  $s$  et à multiplier par 2 l'amplitude des composantes de fréquences positives. Par conséquent, un signal analytique ne peut jamais être une fonction à valeurs réelles. C'est pourquoi l'utilisation du signal analytique porte le nom *d'analyse complexe* en théorie du signal.

### 4.2.3 L'analyse complexe

Dans cette partie, nous proposons d'expliquer les raisons pour lesquelles le signal analytique a été introduit pour l'analyse des signaux. Pour cela, supposons que le signal réel  $s$  est une sinusoïde pure, de fréquence  $f_0$  et d'amplitude  $A > 0$  constantes :

$$s(t) = A \cos(2\pi f_0 t). \quad (4.9)$$

On remarque que si l'on s'intéresse à une mesure prise à un instant  $t = t_0$  de ce signal  $s$ , cette dernière donnera uniquement accès à une valeur comprise dans l'intervalle  $[-A, A]$ , dépendant de  $t_0$ . Ainsi, toute information sur le signal (telle que son amplitude  $A$  ou encore sa fréquence  $f_0$ ) ne pourra être estimée de manière instantanée. Cependant, le concept de signal analytique va quant à lui le permettre.

D'après la définition de la transformée de Hilbert, nous avons donc :

$$s_{\mathcal{H}}(t) = A \sin(2\pi f_0 t). \quad (4.10)$$

La Figure 4.2.(a) propose d'illustrer sur un même graphe le signal  $s(t)$  et  $s_{\mathcal{H}}(t)$ . On y constate le phénomène de déphasage de  $\pi/2$  engendré par la transformée de Hilbert. D'après l'équation (4.3), le signal analytique associé à  $s(t)$  est donné par :

$$s_a(t) = A[\cos(2\pi f_0 t) + i \sin(2\pi f_0 t)] = A \exp(i2\pi f_0 t). \quad (4.11)$$

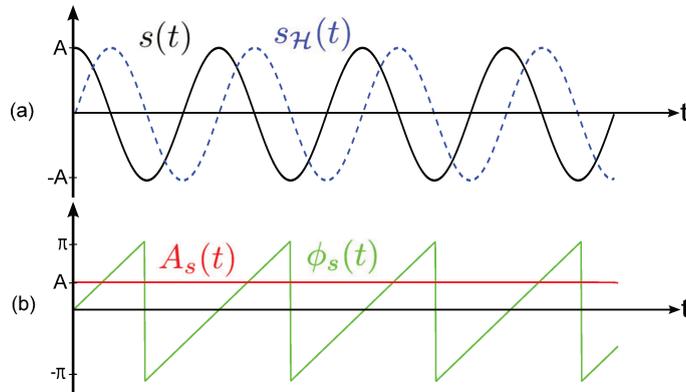


FIGURE 4.2: (a) Visualisation du signal  $s(t) = A \cos(2\pi f_0 t)$  (en trait plein) et de sa transformée de Hilbert  $s_H(t)$  (en trait discontinu). (b) Illustration de l'évolution de l'amplitude instantanée  $A_s(t)$  et de la phase instantanée  $\phi_s(t)$  du signal  $s(t)$ .

Phase $\varphi$	$0$	$\frac{\pi}{2}$	$\pm\pi$	$-\frac{\pi}{2}$
Forme locale				

FIGURE 4.3: Valeurs caractéristiques de la phase instantanée et formes associées du signal.

On note ici que l'amplitude  $A$  de la sinusoïde correspond directement à l'amplitude instantanée du signal analytique :

$$A_s(t) = |s_a(t)| = A. \quad (4.12)$$

De ce fait, grâce à l'utilisation de la transformée de Hilbert, on peut obtenir en chaque instant  $t$  une grandeur caractéristique du signal sinusoïdal étudié telle que son amplitude. Puisque l'énergie d'un signal sinusoïdal peut être reliée à son amplitude, il est courant de retrouver dans la littérature que l'amplitude instantanée du signal analytique est considérée comme une information **énergétique**.

De manière similaire, la valeur d'un échantillon de  $s$  à un instant  $t_0$  ne nous renseigne pas sur la forme de ce signal. Pourtant, il peut être utile de savoir si ce dernier se trouve sur un maximum ou un minimum local par exemple. Nous allons maintenant constater que la notion de phase instantanée permet d'avoir accès à une telle caractéristique du signal. D'après l'équation (4.6), la phase instantanée du signal  $s$  est donnée par :

$$\phi_s(t) = \arg(s_a(t)) = 2\pi f_0 t. \quad (4.13)$$

On constate sur la Figure 4.2.(b) que cette phase instantanée prend des valeurs qui sont caractéristiques des formes locales du signal (*i.e.* sur un maxima ou minima local, sur une transition montante ou descendante). On peut alors interpréter la phase comme un descripteur continu de la structure locale d'un signal, c'est pourquoi la littérature qualifie cette information de **structurelle** [100, 101]. La Figure 4.3 recense les 4 valeurs caractéristiques qu'une phase instantanée peut prendre en fonction de la forme locale du signal. Par ailleurs, on remarque

dans l'équation (4.13) que la phase est une fonction linéaire de la variable  $t$ . Il est alors possible d'en extraire une information de fréquence instantanée  $f_i(t)$  du signal par différentiation de la phase instantanée par rapport à la variable  $t$  :

$$f_i(t) = \frac{1}{2\pi} \frac{d}{dt} \arg[s_a(t)]. \quad (4.14)$$

Le signal  $s$  utilisé dans cet exemple étant une sinusoïde pure, on obtient une fréquence instantanée constante  $f_i(t) = f_0$ .

Nous venons de voir que l'analyse complexe d'un signal permet d'extraire plusieurs caractéristiques instantanées que nous avons illustrées sur un signal sinusoïdal pur. Mais qu'en est-il pour un signal quelconque ? Selon G. Grandlund et H. Knutsson [101], la notion de phase, ou de fréquence instantanée, ne peut décrire instantanément un signal que si il est purement sinusoïdal (ou monochromatique). Néanmoins, il n'y a pas de définition précise pour juger si un signal est mono-composante ou pas. A défaut, il est généralement admis que toutes les informations issues de l'analyse complexe restent valables et interprétables pour des signaux à bande-étroite (*i.e.* pour des signaux donc les composantes fréquentielles sont proches les unes des autres). Pour de plus amples informations sur ce concept, nous invitons le lecteur à se reporter à [101].

#### 4.2.4 Discussion

Nous venons de voir que le concept de signal analytique permet une analyse complexe d'un signal 1D sinusoïdal (ou du moins à bande-étroite) et d'en extraire des informations instantanées. Ces informations sont alors reliées à des caractéristiques énergétiques (l'amplitude instantanée), structurelles (la phase instantanée) mais également fréquentielles (la fréquence instantanée).

Outre sa capacité à décrire instantanément un signal, le signal analytique permet de séparer les informations extraites de manière orthogonale. En effet, d'après G. Grandlund et H. Knutsson [101], l'amplitude et la phase du signal analytique vérifient les propriétés dites *d'invariance-équivariance*, c'est-à-dire que la phase et l'amplitude instantanées sont indépendantes l'une de l'autre. Pour le comprendre, on peut par exemple considérer la Figure 4.4 qui propose d'illustrer le signal analytique dans le plan complexe. A un instant  $t$  donné, le signal analytique  $s_a$  peut être représenté sous la forme d'un vecteur dont les coordonnées polaires sont l'amplitude instantanée  $A_s(t)$  et la phase instantanée  $\phi_s(t)$ . Une telle visualisation permet alors de constater que l'amplitude instantanée peut varier indépendamment de la phase instantanée, et inversement. Ainsi, l'information énergétique portée par  $A_s(t)$  est invariante par rapport l'information structurelle portée par  $\phi_s(t)$ .

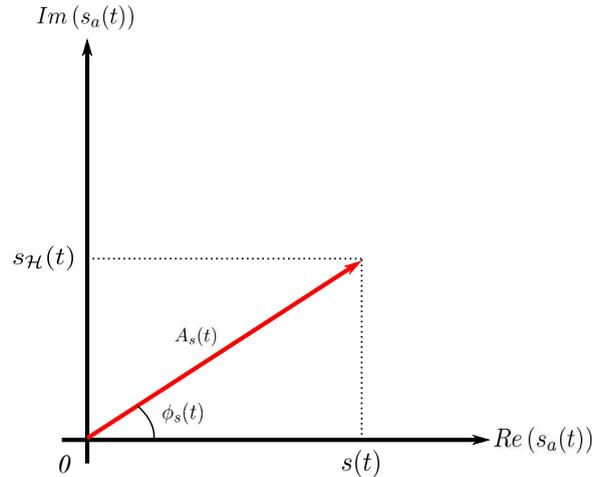


FIGURE 4.4: Représentation du signal analytique dans le plan complexe.

Malgré sa restriction à l'analyse de signaux à bande-étroite, l'exploitation de ces caractéristiques instantanées se fait aujourd'hui dans divers domaines, allant du codage de l'information jusqu'à la détection d'objets sur des données radar, en passant par l'étude de données sismiques. Cependant, par définition, l'utilisation de ce signal analytique est limitée aux signaux unidimensionnel (1D). Or, de nombreuses applications pourraient être intéressées par une adaptation de ce concept à des données de dimension supérieure, comme dans le cas du traitement d'image. En particulier dans ce domaine, il est reconnu que des informations locales en termes de formes ou de fréquences sont des propriétés importantes. Il serait alors utile d'avoir accès à un outil similaire pour l'analyse d'images.

### 4.3 Les extensions 2D du signal analytique

Comme nous l'avons vu dans le paragraphe précédent, le signal analytique est particulièrement bien adapté à l'étude des signaux 1D. La communauté du traitement d'image s'est donc intéressée à définir un tel concept pour des signaux 2D. Pour cela, il est alors nécessaire de redéfinir la transformée de Hilbert afin qu'elle puisse s'adapter à de tels signaux.

#### 4.3.1 La transformée de Hilbert 2D

A travers la littérature, il est possible de trouver un certain nombre d'approches proposant une extension bi-dimensionnelle de la transformée de Hilbert. Partant du principe que la transformée de Hilbert est bien définie dans le cas 1D, H. Stark propose de calculer deux transformées de Hilbert distinctes, l'une suivant les lignes, et l'autre suivant les colonnes de l'image. Il aboutit alors à une formulation de la transformée de Hilbert "totale" [102]. S. Hahn [103] propose en 1992 une extension appelée transformée de Hilbert "single orthant". Cette approche part du constat que la transformée de Hilbert ne conserve que les fréquences positives

du spectre du signal d'origine. L'auteur propose alors de construire une transformée de Hilbert 2D en ne conservant uniquement que le spectre des orthants<sup>1</sup> correspondant à des fréquences positives, le tout, en évitant la redondance des informations. Cependant, cette approche ne permet pas de conserver la relation de quadrature entre le signal et sa transformée de Hilbert. En 1996, le même auteur propose alors de définir une transformée de Hilbert partielle [104], qui n'est autre qu'une transformée de Hilbert 1D dans une direction spécifiée. Cependant, une telle méthode se base sur l'hypothèse que le signal ne varie que dans une direction bien définie.

### 4.3.2 Au delà des nombres complexes

T. Bulöw mentionne dans [105] que les méthodes précédentes ne permettent pas de définir une généralisation viable de la transformée de Hilbert. L'auteur suppose que ces différents échecs reposent en grande partie sur l'algèbre utilisée. En effet, puisque la popularité du signal analytique est basée sur l'analyse complexe qui lui est associée, toutes ces propositions d'extension 2D se sont restreintes à l'algèbre complexe (*i.e.* une formulation du signal analytique sous forme de nombres complexes). Or selon l'auteur, dans des espaces de dimensions supérieures à 2, l'analyse complexe des signaux n'est plus adaptée. Pour y remédier, l'auteur propose de combiner les transformées de Hilbert partielles et totales précédentes dans le formalisme des quaternions pour aboutir à un signal analytique hypercomplexe.

Cependant, M. Felsberg remarque dans [106], que ce signal analytique hypercomplexe ne respecte pas toutes les propriétés du signal analytique. Il propose alors de repenser entièrement la théorie du signal analytique, mais cette fois dans l'algèbre géométrique.

### 4.3.3 L'algèbre géométrique

L'algèbre géométrique est une structure mathématique dont l'utilisation devient de plus en plus courante dans de nombreux domaines, que ce soit en physique ou en sciences industrielles. Cette popularité repose sur leurs capacités à retranscrire de manière simple et concise des opérations géométriques complexes et sophistiquées dans des espaces de grandes dimensions. Naturellement, le domaine du traitement de l'image commence à s'y intéresser comme le démontre l'ouvrage de L. Dorst [107].

L'avantage qui intéresse M. Felsberg est la capacité de cette algèbre à étendre des notions simples comme la rotation, la translation ou encore la dilatation à des espaces de dimensions supérieures à 1. L'auteur propose alors de redéfinir certains outils classiques du traitement du signal dans cette algèbre afin de bénéficier de ses propriétés intéressantes. En particulier,

---

1. Un orthant est un sous-espace d'un espace  $nD$ . Par définition, un tel espace sera alors formé de  $2^n$  orthants. Par exemple dans le cas 1D, il y a deux orthants qui représentent un demi-axe chacun. Dans un espace 2D, il y a quatre orthants qui représentent chacun un quadrant du plan (par exemple les quadrants du plan fréquentiel)

ceci le conduit à redéfinir la transformée de Hilbert et de lui trouver une extension multidimensionnelle dans un cadre mathématique formel. Cette extension, **le signal monogène**, est détaillée dans le paragraphe suivant et sera exploitée tout au long de ce manuscrit.

Le lecteur pourra se reporter à l'annexe A pour des descriptions plus détaillées des différentes extensions de la transformée de Hilbert 2D. De plus, à l'annexe B est proposée une introduction à algèbre géométrique et à certaines de leurs propriétés.

## 4.4 Définition du signal monogène

La théorie du signal exprimée par M. Felsberg dans le cadre de l'algèbre géométrique permet de mettre en évidence que la transformée de Hilbert n'est qu'un cas particulier d'une transformée plus générale, appelée transformée de Riesz. Les détails mathématiques liés à cette extension ne faisant pas l'objet de modification et n'étant pas d'intérêt particulier pour la suite du manuscrit, nous invitons les lecteurs qui seraient intéressés de se référer à l'annexe C ou aux travaux de M. Felsberg [106, 108, 109].

### 4.4.1 La transformée de Riesz

La transformée de Riesz doit son nom au mathématicien M. Riesz qui propose, en 1929, une généralisation de la transformée de Hilbert à des espaces Euclidiens de dimension supérieure à 1 [110]. Bien que popularisée par la théorie de Fourier et en particulier dans la théorie des intégrales singulières (cf. théorie des intégrales singulières de Calderon-Zygmund [111]), cette transformée n'est apparue en traitement d'images qu'avec les travaux de M. Felsberg au début des années 2000. Dans la suite de cette section, nous allons définir et décrire certaines propriétés de la transformée de Riesz dans le cas particulier d'un espace bidimensionnel (2D), sachant que cette définition peut être étendue à tout espace de dimension  $N \in \mathbb{N}$ .

#### 4.4.1.1 Définition

Si l'on considère une fonction  $f$  de deux variables réelles  $\mathbf{x} = (x, y)$ , sa transformée de Riesz, notée  $\mathcal{R}\{f\}(\mathbf{x})$ , est définie par l'intermédiaire de deux transformées indépendantes  $\mathcal{R}_1\{f\}$  et  $\mathcal{R}_2\{f\}$  que l'on regroupe généralement sous la forme d'un nombre complexe :

$$\mathcal{R}\{f\}(\mathbf{x}) = \mathcal{R}_1\{f\}(\mathbf{x}) + i\mathcal{R}_2\{f\}(\mathbf{x}). \quad (4.15)$$

Les deux transformées utilisées ont pour expression :

$$\mathcal{R}_1\{f\}(\mathbf{x}) = \text{v.p.} \iint_{\mathbb{R}^2} \frac{x'}{2\pi\|\mathbf{x}'\|^3} f(\mathbf{x} - \mathbf{x}') d\mathbf{x}', \quad (4.16)$$

$$\mathcal{R}_2\{f\}(\mathbf{x}) = \text{v.p.} \iint_{\mathbb{R}^2} \frac{y'}{2\pi\|\mathbf{x}'\|^3} f(\mathbf{x} - \mathbf{x}') d\mathbf{x}'. \quad (4.17)$$

où  $\|\mathbf{x}'\|$  représente la norme du vecteur  $\mathbf{x}'$ . En d'autres termes, la transformée de Riesz d'une fonction bidimensionnelle  $f$  est la convolution de cette dernière avec les deux noyaux  $r_1$  et  $r_2$  (dits noyaux de Riesz) suivants :

$$r_1(\mathbf{x}) = \frac{x}{2\pi(x^2 + y^2)^{3/2}}, \quad (4.18)$$

$$r_2(\mathbf{x}) = \frac{y}{2\pi(x^2 + y^2)^{3/2}}. \quad (4.19)$$

Une représentation 3D (cf. Figure 4.5) montre que ces deux noyaux ont une même enveloppe mais une orientation différente (un angle de  $\pi/2$  les distingue). Ces noyaux peuvent également être exprimés dans l'espace des fréquences :

$$R_1(\mathbf{u}) = -i \frac{u_1}{\|\mathbf{u}\|}, \quad (4.20)$$

$$R_2(\mathbf{u}) = -i \frac{u_2}{\|\mathbf{u}\|}, \quad (4.21)$$

où  $u_1, u_2$  représentent les variables fréquentielles et  $\mathbf{u} = (u_1, u_2)$ .

On peut y constater leurs ressemblances avec le noyau de Hilbert de l'équation (4.2). En particulier, on remarque la singularité à l'origine (*i.e.* pour  $\mathbf{u} = 0$ ) de ces deux noyaux, qui

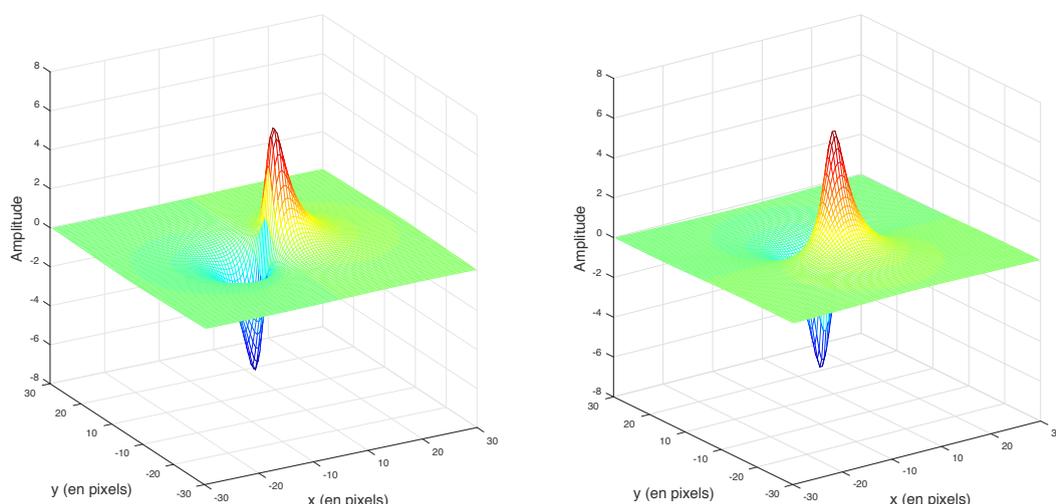


FIGURE 4.5: Représentation 3D des deux noyaux de Riesz  $r_1$  (à gauche) et  $r_2$  (à droite).

implique que la transformée de Riesz n'est pas adaptée à l'étude des signaux de moyenne non-nulle, tout comme la transformée de Hilbert.

Si l'on note  $F$  la transformée de Fourier ( $TF$ ) de la fonction  $f$ , la transformée de Riesz s'exprime dans le domaine fréquentiel par :

$$\mathcal{R}\{f\}(\mathbf{x}) \xleftrightarrow{TF} -i \frac{u_1 + iu_2}{\|\mathbf{u}\|} F(\mathbf{u}). \quad (4.22)$$

Nous allons maintenant décrire plus précisément la transformée de Riesz à travers certaines de ses propriétés les plus notables.

#### 4.4.1.2 Propriétés

Nous avons choisi l'écriture de l'équation 4.15 de manière à pouvoir interpréter  $\mathcal{R}$  comme un opérateur que nous appellerons dans la suite **opérateur de Riesz**.

La première propriété de l'opérateur de Riesz est qu'il est **unitaire**. En effet, on peut montrer [112] que son adjoint est égal à son inverse et vaut alors :

$$\mathcal{R}^{-1}\{f\}(\mathbf{x}) \xleftrightarrow{TF} i \frac{u_1 - iu_2}{\|\mathbf{u}\|} F(\mathbf{u}). \quad (4.23)$$

L'avantage d'un opérateur unitaire est qu'il conserve la structure algébrique et métrique entre l'ensemble de départ et l'ensemble d'arrivée. Par exemple, des vecteurs qui sont orthogonaux le resteront après transformée de Riesz.

L'opérateur de Riesz possède également la propriété d'invariance par translation et changement d'échelle. En effet, il vérifie la relation :

$$\mathcal{R}\{f(a\mathbf{y} + \mathbf{b})\}(\mathbf{x}) = \mathcal{R}\{f(\mathbf{y})\}(a\mathbf{x} + \mathbf{b}). \quad (4.24)$$

L'invariance par translation est fondamentale, puisqu'elle exprime le fait qu'une information sera traitée de la même manière quelle que soit sa localisation dans l'espace.

Ces propriétés d'invariance sont aussi complétées par l'invariance en rotation de l'opérateur (à un facteur constant près) qui s'exprime par :

$$\mathcal{R}\{f(\mathbf{R}_\theta \mathbf{y})\}(\mathbf{x}) = e^{i\theta} \mathcal{R}\{f(\mathbf{y})\}(\mathbf{R}_\theta \mathbf{x}) \quad \text{avec} \quad \mathbf{R}_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (4.25)$$

Toutefois, si l'on considère le module de l'opérateur de Riesz, on constate que l'invariance en rotation est alors pleinement vérifiée (la norme de l'exponentielle complexe est unitaire). Cette

propriété permet à la transformée de Riesz de ne pas altérer l'orientation des structures lors de son application à une image par exemple.

De plus, l'expression complexe de la transformée de Riesz de l'équation (4.15) nous permet d'extraire deux informations : une norme  $\mathcal{N}_{\mathcal{R}}$  et un angle  $\theta_{\mathcal{R}}$  :

$$\mathcal{N}_{\mathcal{R}}(\mathbf{x}) = \|\mathcal{R}\{f\}(\mathbf{x})\| = \sqrt{\mathcal{R}_1\{f\}(\mathbf{x})^2 + \mathcal{R}_2\{f\}(\mathbf{x})^2} \quad (4.26)$$

$$\theta_{\mathcal{R}}(\mathbf{x}) = \arg(\mathcal{R}\{f\}(\mathbf{x})) = \arctan\left(\frac{\mathcal{R}_2\{f\}(\mathbf{x})}{\mathcal{R}_1\{f\}(\mathbf{x})}\right), \quad (4.27)$$

Nous verrons au paragraphe suivant que c'est  $\theta_{\mathcal{R}}$  qui porte une information particulièrement intéressante puisqu'elle est reliée à la notion d'orientation d'un signal 2D.

Des propriétés complémentaires sur la transformée de Riesz pourront être retrouvées par le lecteur dans la thèse de R. Soulard [113].

#### 4.4.1.3 Application à une onde sinusoïdale 2D

Pour vérifier que la transformée de Riesz réalise bien une généralisation de la transformée de Hilbert, considérons une onde plane sinusoïdale  $s$ , d'amplitude  $A$  et de fréquence pure  $u_0$ , orientée dans la direction  $\alpha$  :

$$s(\mathbf{x}) = A \cos(2\pi \mathbf{u}_0^T \mathbf{x}) \quad \text{où} \quad \mathbf{u}_0 = (u_0 \cos(\alpha), u_0 \sin(\alpha)). \quad (4.28)$$

La transformée de Fourier de ce signal 2D est :

$$S(\mathbf{u}) = \frac{A}{2} (\delta(\mathbf{u} - \mathbf{u}_0) + \delta(\mathbf{u} + \mathbf{u}_0)). \quad (4.29)$$

où  $\delta$  est l'impulsion de Dirac qui signifie que  $\delta(m - n) = 1$  si  $m = n$  et que  $\delta(m - n) = 0$  sinon. D'après les équations (4.20) et (4.21), on obtient l'expression des deux transformées  $\mathcal{R}_j$   $j \in \{1, 2\}$  (dans le domaine fréquentiel) qui interviennent dans l'expression de la transformée de Riesz :

$$\mathcal{R}_1\{s\}(\mathbf{u}) = \frac{A \cos(\alpha)}{2} (\delta(\mathbf{u} - \mathbf{u}_0) - \delta(\mathbf{u} + \mathbf{u}_0)), \quad (4.30)$$

$$\mathcal{R}_2\{s\}(\mathbf{u}) = \frac{A \sin(\alpha)}{2} (\delta(\mathbf{u} - \mathbf{u}_0) - \delta(\mathbf{u} + \mathbf{u}_0)), \quad (4.31)$$

Dans le membre de droite de ces deux équations, nous retrouvons l'expression de la transformée de Fourier d'un sinus 2D. D'après les propriétés de la transformée de Fourier inverse, nous obtenons alors la transformée de Riesz du signal  $s(\mathbf{x})$  :

$$\mathcal{R}\{s\}(\mathbf{x}) = A \sin(2\pi \mathbf{u}_0^T \mathbf{x}) e^{i\alpha}. \quad (4.32)$$

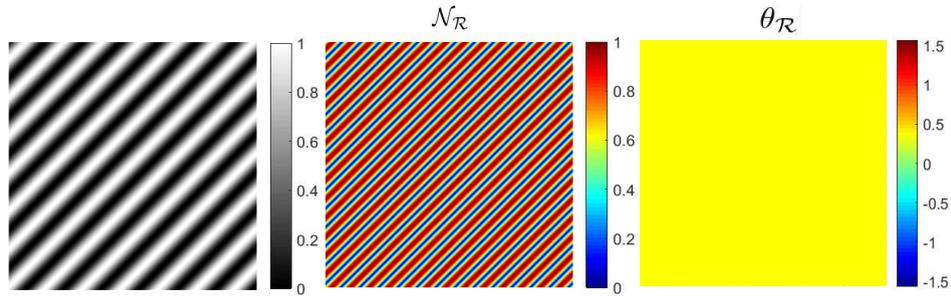


FIGURE 4.6: Exemple d'une sinusoïde 2D de fréquence et d'orientation fixe (à gauche) pour laquelle la norme  $\mathcal{N}_R$  et l'argument  $\theta_R$  de la transformée de Riesz sont illustrées (au milieu et à droite respectivement).

On constate que la fonction cosinus initiale est donc devenue un sinus, ce qui correspond à un déphasage de  $\pi/2$ , comme dans le cas de la transformée de Hilbert. Par ailleurs, on remarque la présence du facteur  $e^{i\alpha}$  correspondant à l'invariance par rotation de la transformée à un facteur constant près (précédemment mentionné à l'équation (4.25)). Le calcul de la norme de la transformée de Riesz mène à l'expression :

$$\mathcal{N}_R(\mathbf{x}) = A |\sin(2\pi \mathbf{u}_0^T \mathbf{x})| \quad (4.33)$$

qui montre que cette transformée ne crée aucune amplification par rapport au signal de départ mais aussi que son contenu fréquentiel est inchangé (*i.e.* que la transformée de Riesz peut être assimilée à un filtrage "passe-tout"). On peut également en conclure que la transformée de Riesz est un opérateur isotrope puisque le déphasage qu'elle provoque est indépendant de l'orientation du signal. Ce caractère isotrope permet une application de cette transformée à tout type de signaux orientés de manière quelconque, et il est intéressant de noter que l'information  $\theta_R$  qui en est extraite (cf. équation (4.27)) donne en plus accès cette orientation :

$$\theta_R(\mathbf{x}) = \arg(\mathcal{R}\{f\}(\mathbf{x})) = \alpha \quad (4.34)$$

Ainsi, la transformée de Riesz est non seulement un outil isotrope mais elle permet d'accéder également à une information d'orientation concernant le signal 2D. Il est également à noter que ces informations de norme et d'orientation sont disponibles en chaque position  $\mathbf{x}$  (*i.e.* en chaque pixel de l'image). Ces informations seront donc considérées comme **locales**.

La Figure 4.6 propose d'illustrer l'exemple que nous venons de détailler ci-dessus. La fonction sinusoïdale est paramétrée de façon à avoir une amplitude  $A = 1$ , une orientation  $\alpha = 45^\circ$  (suivant une diagonale de l'image) et une période fixée à  $\lambda = 10$  pixels (soit une fréquence  $u_0 = 0.1 \text{ pixel}^{-1}$ ). Les informations de norme et d'orientation issues de la transformée de Riesz sont également représentées (au milieu et à droite respectivement). Entre ces deux informations, c'est l'information d'orientation qui nous intéressera particulièrement dans la suite de ce manuscrit. On constate que cette dernière donne bien accès à l'orientation du signal

2D en chaque pixel de l'image (cette information étant donnée en radians, l'information est présentée comme une image dont la valeur de tous les pixels est de 0.78 rad).

Maintenant que nous venons de définir la transformée de Riesz, nous allons nous intéresser à l'extension du concept de signal analytique qu'elle permet : le **signal monogène**.

## 4.4.2 Le signal monogène

### 4.4.2.1 Première définition

Nous avons vu au paragraphe 4.2 que le signal analytique peut être considéré comme un couple formé du signal 1D étudié et de sa transformée de Hilbert. D'après M. Felsberg, l'extension multi-dimensionnelle de ce signal analytique, le signal monogène, est composé par le signal  $nD$  étudié et de sa transformée de Riesz. En particulier, dans le cas 2D, le signal monogène  $f_M$  est donc un triplet :

$$f_M(\mathbf{x}) = [f(\mathbf{x}), \mathcal{R}_1\{f\}(\mathbf{x}), \mathcal{R}_2\{f\}(\mathbf{x})]. \quad (4.35)$$

Cette formalisation sous forme d'un vecteur 3D permet d'en extraire 3 informations, à savoir une norme et deux angles (*i.e.* ses coordonnées sphériques) :

$$A(\mathbf{x}) = |\mathbf{f}_M(\mathbf{x})| = \sqrt{\mathcal{R}_1\{f\}(\mathbf{x})^2 + \mathcal{R}_2\{f\}(\mathbf{x})^2 + f(\mathbf{x})^2} \quad (4.36)$$

$$\phi(\mathbf{x}) = \arg(f(\mathbf{x}) + j|\mathcal{R}\{f\}(\mathbf{x})|) = \text{atan} \left( \frac{\sqrt{\mathcal{R}_1\{f\}(\mathbf{x})^2 + \mathcal{R}_2\{f\}(\mathbf{x})^2}}{f(\mathbf{x})} \right) \quad (4.37)$$

$$\theta(\mathbf{x}) = \arg(\mathcal{R}\{f\}(\mathbf{x})) = \text{atan} \left( \frac{\mathcal{R}_2\{f\}(\mathbf{x})}{\mathcal{R}_1\{f\}(\mathbf{x})} \right) \quad (4.38)$$

On remarque que par rapport à ce que l'on a pu voir dans le cas du signal analytique et des signaux 1D, une troisième composante apparaît ici en plus de l'amplitude  $A$  et de la phase  $\phi$  : c'est l'angle  $\theta$ . Cependant, cet angle est défini comme l'argument de la transformée de Riesz, et nous venons de voir dans le paragraphe précédent qu'il est relié à la notion d'orientation. En conséquence, le signal monogène fournit trois informations qui représentent une amplitude  $A$ , une orientation  $\theta$  et une phase  $\phi$ . Alors que dans le cas 1D, ces informations étaient dites **instantanées**, dans le cas 2D du traitement d'images nous les qualifierons de **locales**. En effet, les trois informations sont accessibles à chaque position  $\mathbf{x}$ , c'est-à-dire en chaque pixel de l'image. L'amplitude, l'orientation et la phase locales seront appelées **informations monogènes** dans le reste de ce manuscrit.

Comme dans le cas du signal analytique, les informations extraites du signal monogène portent sur des caractéristiques distinctes du signal étudié. L'information d'amplitude renseigne sur l'énergie locale de ce signal. La phase représente une information de forme locale

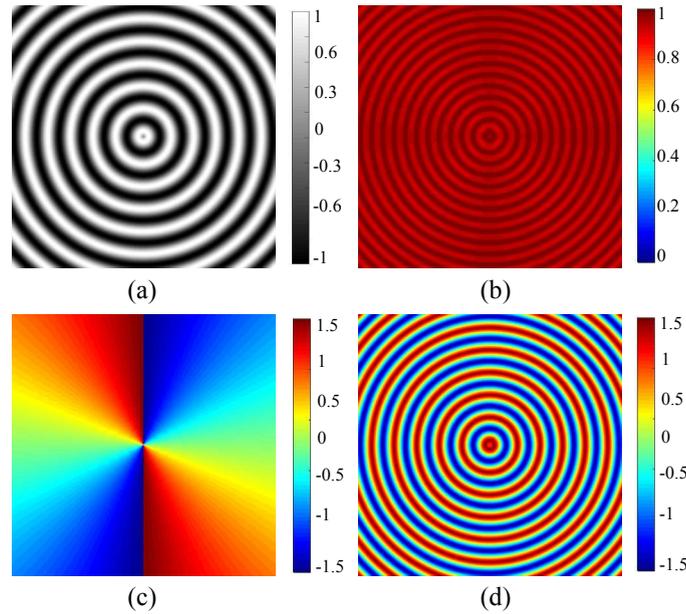


FIGURE 4.7: Illustration d'une sinusoïde 2D isotrope (a) sur laquelle les informations d'amplitude (b), de phase (c) et d'orientation locale (d) sont extraites.

et conserve alors son qualificatif d'information **structurelle**. L'orientation apporte une information **géométrique** supplémentaire qui n'avait pas lieu d'être en 1D. Elle permet de décrire l'orientation du signal 2D localement.

Puisque les informations monogènes jouent le rôle de coordonnées sphériques, elles sont par définition mutuellement orthogonales. Elles peuvent donc varier indépendamment l'une de l'autre. En particulier, l'orientation et la phase sont orthogonales toutes deux à l'amplitude. Elles ont donc la spécificité d'être indépendantes de l'énergie et par conséquent du contraste dans l'image.

#### 4.4.2.2 Application à une image synthétique

Afin d'illustrer le signal monogène, considérons un signal 2D sinusoïdal (d'amplitude et de fréquence constantes) qui se propage de manière isotrope depuis le centre de l'image (cf. Figure 4.7.(a)). Les trois informations monogènes extraites sont représentées sur les Figures 4.7.(b) (c) et (d). Comme attendu, on constate que l'information d'amplitude représente le niveau énergétique du signal initial qui est ici quasi-constant sur l'image entière puisque la sinusoïde 2D a une amplitude constante. Concernant l'information de phase locale, on remarque qu'elle suit les variations du signal. Cette information étant angulaire, toutes ses valeurs sont comprises dans l'intervalle  $[-\pi/2, \pi/2]$ , et on constate que la périodicité de la sinusoïde induit également une périodicité dans la phase locale. Comme nous l'avons vu lors de l'étude de la phase du signal analytique, la phase monogène prend également des valeurs singulières en fonction de la forme locale du signal. Par exemple, on constate que la phase vaut  $\pi/2$  lorsque la sinusoïde est sur un maximum local et vaut  $-\pi/2$  sur un minimum local. Concernant l'orientation locale,

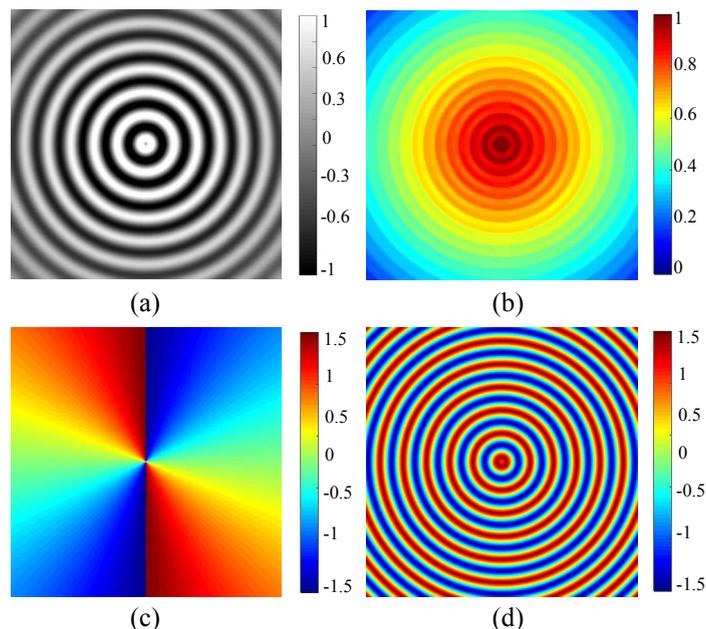


FIGURE 4.8: Illustration d'une sinusoïde 2D isotrope dont l'amplitude décroît de manière radiale à partir du centre de l'image (a) sur laquelle les informations d'amplitude (b), de phase (c) et d'orientation locale (d) sont extraites.

on observe qu'elle nous donne accès à la direction de propagation du signal. La sinusoïde étant de propagation isotrope, l'information d'orientation est constante selon chaque direction.

Pour illustrer le phénomène d'indépendance au contraste, la Figure 4.8 propose exactement le même signal, mais cette fois dont l'amplitude décroît de manière radiale (à partir du centre de l'image). L'amplitude du signal monogène justifie son qualificatif d'information énergétique en suivant cette décroissance. La phase et l'orientation ne sont quant à elles pas modifiées compte tenu de l'orthogonalité entre les trois informations.

#### 4.4.2.3 Discussion

Jusqu'à présent, nous avons pu constater que le signal monogène justifie l'extension multidimensionnelle du signal analytique en fournissant trois informations locales et orthogonales reliées aux caractéristiques énergétiques, structurelles et géométriques d'un signal 2D. Cependant, ces deux concepts partagent également la même limite : ils se restreignent tous les deux à l'étude de signaux monochromatiques. Or de manière générale, qu'ils soient 1D ou nD, les signaux réels sont composés d'une multitude de composantes fréquentielles aux diverses amplitudes et phases. Que ce soit la transformée de Hilbert ou la transformée de Riesz, le déphasage qu'elles impliquent sur ces signaux n'est alors pas remis en question. C'est l'interprétation que l'on peut faire des informations extraites (amplitude, phase ou orientation) qui sera quant à elle erronée. En effet, M. Felsberg mentionne que du fait de la superposition de toutes les fréquences, l'information contenue par chaque composante est réduite à une simple moyenne dominée par les fréquences des composantes de fortes amplitudes. Toutefois, au paragraphe

4.2.2, nous avons mentionné que la littérature admet que les informations issues du signal analytique restent significatives pour des signaux à bande-étroite. Cependant, très peu de signaux réels sont naturellement à bande-étroite, et en particulier les signaux 2D que nous traiterons dans ces travaux de thèse. Pour analyser les signaux quelconques, il est alors nécessaire de séparer les composantes fréquentielles.

Le moyen le plus populaire pour séparer des composantes fréquentielles d'un signal est le filtrage passe-bande. Ce dernier permet de sélectionner une partie restreinte du spectre d'un signal pour l'étudier séparément. De manière générale, un tel filtrage est réalisé par l'application de deux filtres passe-bas, dont l'un définit la fréquence de coupure basse et l'autre la fréquence de coupure haute. De nombreux filtres passe-bande existent dans la littérature mais nous allons voir dans le paragraphe suivant qu'il est possible d'en construire un directement à partir du signal monogène.

### 4.4.3 Définition générale du signal monogène

Tout d'abord, il est nécessaire de revenir rapidement au concept de signal analytique. En mathématique, une fonction est qualifiée d'*analytique* si elle vérifie les équations de Cauchy-Riemann 2D. Ainsi, bien que le signal analytique soit un concept 1D, il se base donc sur une théorie développée dans un espace de dimension 2. Si l'on note  $x$  et  $y$  ces deux dimensions, M. Felsberg montre dans [109] que le signal analytique (que nous avons décrit à la section 4.2.2) n'est en fait qu'une restriction de la solution des équations de Cauchy-Riemann 2D à l'axe  $y = 0$  (*i.e.* l'établissement d'une condition aux limites). Il démontre alors que pour  $y > 0$  des noyaux de Poisson apparaissent dans l'expression du signal analytique.

Pour trouver l'expression du signal analytique 2D, l'auteur considère qu'il faut donc étudier les équations de Cauchy-Riemann 3D (appelons ces trois dimensions  $x$ ,  $y$  et  $z$ ). Les fonctions qui vérifient ces équations sont appelées fonctions *monogènes*, d'où le nom de signal monogène. L'ajout de cette troisième dimension ne s'est pas fait ressentir dans la définition du signal monogène du paragraphe précédent puisque nous étions dans la restriction à  $z = 0$  (dans la condition aux limites). Pour une valeur  $z > 0$ , la définition du signal monogène va donc être nécessairement différente. Si l'on considère une fonction bidimensionnelle  $f(x, y)$ , son signal monogène est alors défini par :

$$f_M(x, y, z) = [P(x, y, z) * f(x, y), P(x, y, z) * \mathcal{R}_1\{f\}(x, y), P(x, y, z) * \mathcal{R}_2\{f\}(x, y)] \quad (4.39)$$

où l'opérateur "\*" désigne la convolution. Ainsi, on observe l'apparition d'un filtre  $P$  sur chacune des trois composantes du vecteur définissant le signal monogène. M. Felsberg montre

que l'expression de ce filtre est la suivante :

$$P(x, y, z) = \frac{z}{(x^2 + y^2 + z^2)^{3/2}}. \quad (4.40)$$

On constate que cette expression est celle d'un noyau de Poisson 2D. En effet, la dimension  $z$  n'est pas une dimension spatiale pour ce noyau, mais un paramètre qui influence sa taille (comme l'écart-type pour une gaussienne par exemple). Afin d'éviter toute confusion entre les variables spatiales et ce paramètre, nous proposons de renommer  $z$  par  $s$  et de ré-exprimer le noyau de Poisson 2D de la manière suivante :

$$P(x, y, s) = \frac{s}{(x^2 + y^2 + s^2)^{3/2}}. \quad (4.41)$$

La représentation fréquentielle de ce noyau est :

$$P(x, y, s) \xleftrightarrow{TF} \exp(-2\pi s (u_1^2 + u_2^2)). \quad (4.42)$$

On constate que le noyau de Poisson est un filtre passe-bas dont la sélectivité fréquentielle est définie par  $s$  (nous reviendrons plus loin sur l'étude plus précise de ces noyaux).

A présent, si l'on note par  $L$  le résultat de la convolution du signal  $f$  par le noyau de Poisson  $P$  :

$$L(x, y, s) = f(x, y) * P(x, y, s), \quad (4.43)$$

on obtient, par linéarité de la convolution, une nouvelle expression du signal monogène :

$$f_M(x, y, s) = [L(x, y, s), \mathcal{R}_1\{L\}(x, y, s), \mathcal{R}_2\{L\}(x, y, s)]. \quad (4.44)$$

On retrouve alors la définition de l'équation 4.35 dans laquelle le signal  $f$  est remplacé par sa version filtrée passe-bas  $L$ .

Néanmoins, nous avons noté au paragraphe précédent que les informations issues du monogènes ne sont significatives que pour des signaux à bande-étroite. A première vue, les filtres passe-bas de Poisson ne semblent donc pas être la solution au problème de sélectivité fréquentielle du signal monogène. Cependant, nous allons voir dans la section suivante que l'apparition de ces noyaux permet de faire un lien avec la **théorie de l'espace d'échelle** qui lui même permet une **analyse multi-échelle**. C'est cette dernière qui est reconnue pour réaliser la décomposition d'un signal (1D ou 2D) en différentes bandes-passantes successives et de les relier à la notion d'*échelle*. Nous verrons que ces noyaux de Poisson permettront une telle décomposition, et par conséquent, l'extraction d'informations monogènes significatives.

## 4.5 Signal monogène et Analyse multi-échelle

Dans le domaine du traitement d'images, le découpage en différentes bandes-passantes d'un signal 2D est plus connue sous le nom d'analyse multi-échelle. Ce cadre théorique propose de tirer parti de la réciprocité entre le domaine fréquentiel et spatial en liant chaque bande de fréquence à un niveau de détails (*i.e.* une taille de structure) dans l'image, c'est ce qu'on appelle une échelle. Cependant, avant de pouvoir se focaliser sur une échelle particulière, il est tout d'abord nécessaire d'introduire le concept d'espace d'échelle.

### 4.5.1 Théorie de l'espace d'échelle

Toute image est bornée par deux limites physiques que l'on appelle *échelle interne* et *échelle externe*. La première correspond à la plus petite des tailles que l'on va retrouver dans l'image : la taille du pixel (*i.e.* sa résolution). La seconde correspond à la dimension de l'image. Ces deux limites sont donc fixées par le matériel et le scénario d'acquisition de l'image. Entre les échelles interne et externe, des détails vont apparaître sur l'image avec des tailles diverses : c'est la notion d'échelle. Alors que les échelles fines caractérisent les détails de petites dimensions (*i.e.* hautes fréquences), les échelles larges caractérisent les détails grossiers (*i.e.* basses fréquences). Lorsqu'on ne possède aucun *a priori* sur ce qui peut apparaître dans l'image, la difficulté est donc d'analyser l'image à toutes les échelles simultanément.

L'idée a donc été de considérer l'échelle comme un paramètre *continu* afin de générer un *espace d'échelle*. Si d'après J. Weickert [114] cette idée prend racine en 1959 par T. Iijima dans [115], c'est A. Witkin [116] qui, en 1984, est le premier à avoir développé cette idée pour aboutir à la théorie formelle et unifiée que nous connaissons aujourd'hui.

Cette théorie consiste à analyser une image  $I$  telle que :

$$\begin{aligned} \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\rightarrow I(x, y) \end{aligned} \quad (4.45)$$

et de considérer que sa représentation dans l'espace d'échelle est une autre image  $L$  :

$$\begin{aligned} \mathbb{R}^2 \times \mathbb{R}^+ &\rightarrow \mathbb{R} \\ (x, y, \sigma) &\rightarrow L(x, y, \sigma) \end{aligned} \quad (4.46)$$

de telle sorte que  $L(x, y, 0) = I(x, y)$  et  $L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$ ,  $\forall \sigma > 0$ . L'image résultante est alors une fonction des deux coordonnées de positions  $x$  et  $y$ , mais également d'un paramètre additionnel  $\sigma \in \mathbb{R}^+$  qui représente une dimension dite d'échelle. Ce paramètre  $\sigma$  modélise une opération de lissage de l'image d'origine exprimée par le filtre  $G$ , qui supprime progressivement les détails (cela correspond par exemple à un éloignement progressif de la scène

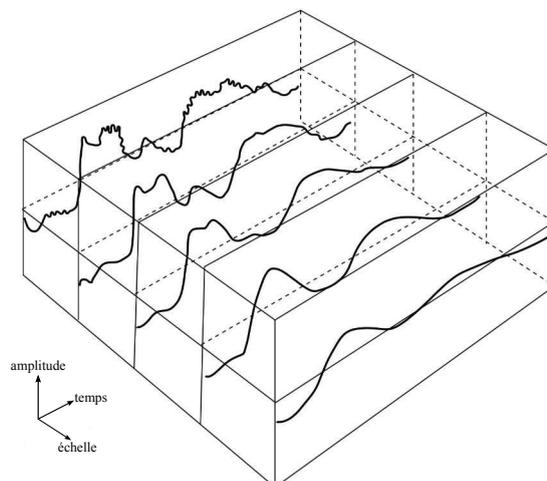


FIGURE 4.9: Illustration schématique de la représentation en espace d'échelle d'un signal 1D (extraite de [5]).

en vision par ordinateur). Si l'on se fie aux travaux de T. Linderberg [117, 118, 119], le filtrage d'une image par un noyau gaussien d'écart-type  $\sigma$  (en pixels) aura pour effet de supprimer les structures de l'image (*i.e.* les détails) dont les dimensions sont inférieures à  $2\sigma$  (en pixels).

Ainsi, par variation de ce paramètre d'échelle, l'image initiale peut être représentée par une famille d'images. Selon J. Koenderink [120], cette famille est gouvernée par les lois issues de l'équation de diffusion de la chaleur [121] :

$$\frac{\partial^2 L}{\partial x^2} + \frac{\partial^2 L}{\partial y^2} = \frac{\partial L}{\partial \sigma} \quad (4.47)$$

avec la condition aux limites  $L(x, y, 0) = I(x, y)$ . Le paramètre d'échelle remplace alors la variable temporelle dans l'expression classique. Cette analogie peut fournir une interprétation physique au lissage, où la représentation en espace d'échelle  $L$  est vue comme l'évolution au cours du temps d'une distribution de température initiale  $I$  dans un milieu homogène. La Figure 4.9 propose d'illustrer un tel phénomène de lissage dans le cas d'un signal temporel 1D. On y observe que les filtrages successifs du signal d'origine (le plus lointain sur l'illustration) font disparaître les détails fins (*i.e.* les hautes fréquences) et ne préservent que les détails grossiers (*i.e.* les basses fréquences).

Cependant, l'opération de lissage n'est pas arbitraire. En effet, le choix du noyau de lissage est d'une importance cruciale puisque l'intérêt de l'espace d'échelle réside dans son principe de *non-crédation de nouvelles structures*. En d'autres termes, les opérations de lissage successives doivent uniquement simplifier le signal et ne surtout pas créer des artefacts. J. Koenderink formule ce critère à l'aide de deux principes fondamentaux que doit respecter tout espace d'échelle :

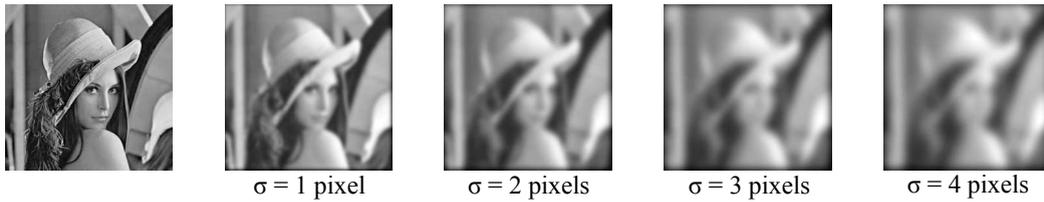


FIGURE 4.10: Résultats de l'application d'un filtrage gaussien sur l'image Lena (image la plus à gauche) pour des valeurs de  $\sigma$  croissantes.

1) *Principe de causalité* : toute information à une échelle grossière a nécessairement une cause à une échelle fine, bien que l'inverse ne soit pas forcément vrai. Le principe de causalité stipule donc l'irréversibilité d'une diminution de la finesse d'analyse : on ne peut pas créer des structures qui n'existaient pas à une échelle plus fine.

2) *Principes d'homogénéité et d'isotropie* : Ces deux principes décrivent les invariances par translation, rotation et changement d'échelle. Le manque d'information *a priori* sur le contenu de l'image (taille, orientation et localisation des différentes structures) nous impose de devoir traiter chaque pixel de l'image de la même manière. Ainsi, ces principes d'invariance nous assurent une généralité du traitement.

Ces principes sont complétés par certains axiomes assurant, entre autres, la linéarité et la séparabilité des noyaux  $G$  générant l'espace d'échelle (le lecteur pourra retrouver une définition précise de ces axiomes dans [115]).

Puisque le noyau gaussien est la fonction de Green de l'équation de diffusion de la chaleur sur un domaine infini [121], il est considéré comme l'unique noyau permettant de générer un espace d'échelle [5, 116, 120, 122]. Son expression bidimensionnelle est alors la suivante :

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (4.48)$$

où  $\sigma$  est l'écart-type de la gaussienne.

La figure 4.10 illustre l'application de ces filtres gaussiens pour des valeurs d'écart-types croissants (valeurs en pixels), appliqués à l'image Lena. On y remarque que les détails fins, comme les cheveux ou encore les différents bords (le bord des épaules par exemple) sont très bien définis pour le plus petit  $\sigma$ . Par contre, pour des valeurs plus grandes de l'écart-type, ces détails disparaissent ou deviennent flous. L'image correspondant à la plus grande variance nous montre uniquement de manière grossière la personne, avec un fort lissage du visage, des épaules, des cheveux et du chapeau. On constate donc que le principe de l'espace d'échelle est de simplifier l'information au fur et à mesure en supprimant les détails.

L'espace d'échelle est donc un cadre théorique développé pour tenir compte de la nature multi-échelle des images. Il y a deux avantages importants de l'approche. Premièrement, la représentation en espace d'échelle permet de multiples interprétations de la donnée 2D partant

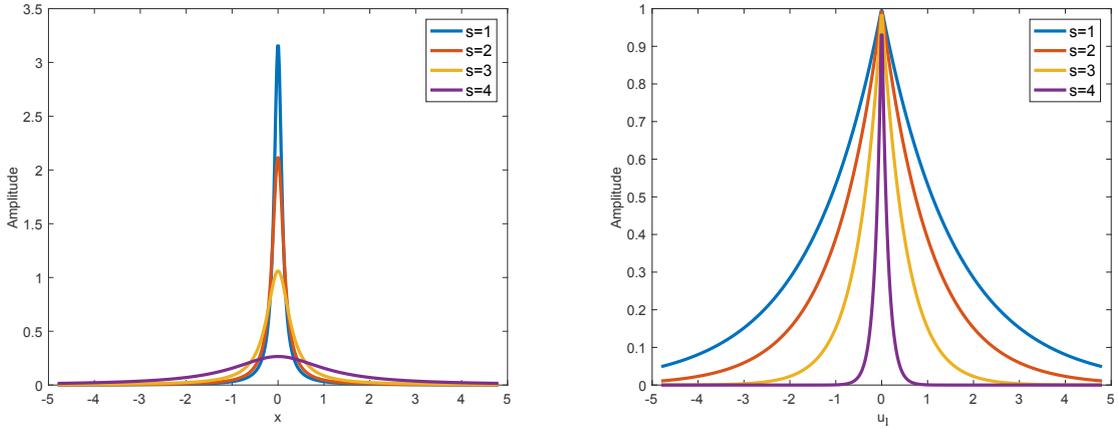


FIGURE 4.11: Représentation spatiale et fréquentielle des filtres de Poisson 2D (coupe horizontale selon la direction  $x$  et  $u_1$  respectivement) pour un ensemble de paramètres distincts.

d'une échelle fine jusqu'à une description de haut niveau des structures grossières contenues dans l'image. Deuxièmement, l'espace d'échelle fournit une flexibilité pour sélectionner une échelle, ou un ensemble d'échelles, qui caractérisent le mieux l'objet ou la structure présente dans l'image. De nombreux algorithmes se sont alors basés sur ce principe pour proposer des solutions dans les domaines de la reconnaissance d'objets ou de scènes, de pistage ou encore de reconstruction 3D.

Jusqu'à très récemment, les noyaux gaussiens étaient reconnus comme les seuls à respecter les principes de création d'un espace d'échelle. Cependant, M. Felsberg démontre dans [106] que les noyaux de Poisson sont également des candidats crédibles.

## 4.5.2 Espace d'échelle de Poisson

Dans cette partie, nous allons voir que les noyaux de Poisson peuvent également vérifier les principes d'un espace d'échelle. Le seul intérêt de ces noyaux par rapport aux noyaux gaussiens précédemment étudiés, réside dans le lien étroit qu'ils entretiennent avec la théorie du signal monogène.

### 4.5.2.1 Les noyaux de Poisson 2D

Rappelons que les noyaux de Poisson 2D sont des filtres dont l'expression est la suivante :

$$P(x, y, s) = \frac{s}{(x^2 + y^2 + s^2)^{3/2}}, \quad (4.49)$$

dont la représentation fréquentielle est donnée par la relation :

$$P(x, y, s) \xleftrightarrow{TF} \exp(-2\pi s (u_1^2 + u_2^2)). \quad (4.50)$$

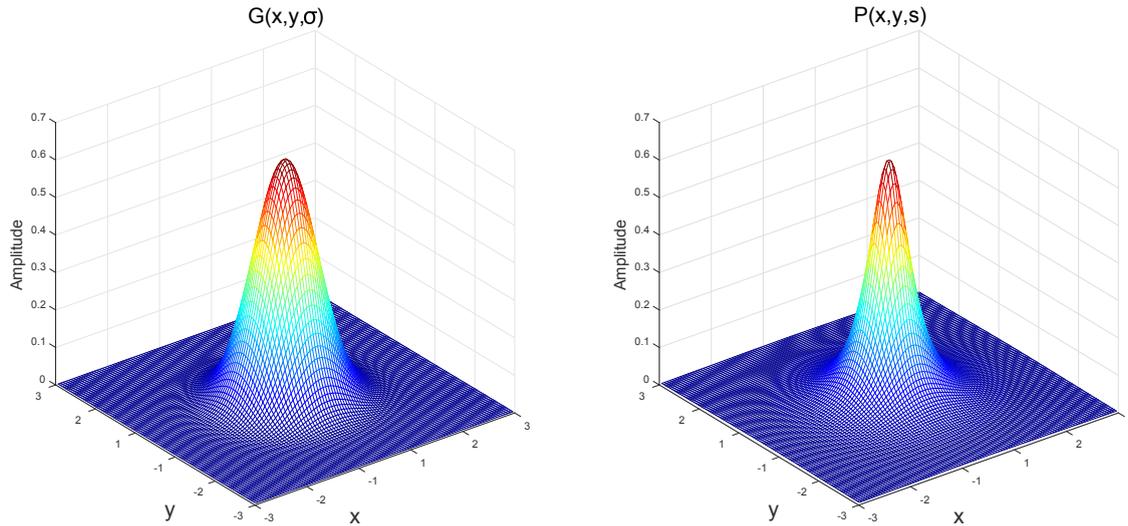


FIGURE 4.12: Illustration d'un noyau gaussien (à gauche) et d'un noyau de Poisson (à droite).

L'influence du paramètre  $s$  de ces noyaux dans le domaine spatial et fréquentiel est illustrée sur la Figure 4.11 (coupes 1D des noyaux selon l'axe défini par la variable spatiale  $x$  et fréquentielle  $u_1$ ). On y constate que les noyaux de Poisson sont donc des filtres dont la largeur dans le domaine spatial croît avec  $s$ . Dans le domaine fréquentiel, lorsque le paramètre  $s$  augmente la sélectivité du filtre (*i.e.* la finesse du noyau) devient plus forte. Ce phénomène est connu sous le nom de principe d'incertitude d'Heisenberg [123] qui stipule qu'on ne peut être localisé précisément dans les domaines spatial et fréquentiel simultanément.

Ainsi, on constate une similarité entre le paramètre  $s$  d'un noyau de Poisson et l'écart-type  $\sigma$  d'un noyau gaussien. Par conséquent, le paramètre  $s$  laisse à penser qu'il pourrait agir comme une dimension d'échelle. Dans un premier temps, nous pouvons tout d'abord comparer qualitativement ces deux types de noyaux afin d'observer leurs différences.

Afin de faciliter cette comparaison, il est nécessaire que leurs maximums respectifs soient égaux. Cette condition est réalisée si l'équation suivante est respectée :

$$P(0, 0, s) = G(0, 0, \sigma) \iff s = \sqrt{2\pi} \sigma. \quad (4.51)$$

La Figure 4.12 propose une représentation 3D de ces noyaux dans le domaine spatial. On peut voir que ces deux noyaux possèdent un lobe important et un caractère isotrope commun. Pour poursuivre cette comparaison, la Figure 4.13 propose une coupe 1D de ces noyaux dans les domaines spatial et fréquentiel. Dans le domaine spatial, le noyau de Poisson (en bleu) est moins large à mi-hauteur que le gaussien (en rouge). Cependant, ce dernier présente un étalement spatial plus court que le noyau de Poisson. Dans le domaine spectral, les mêmes observations sur la forme des noyaux peuvent être réalisées.

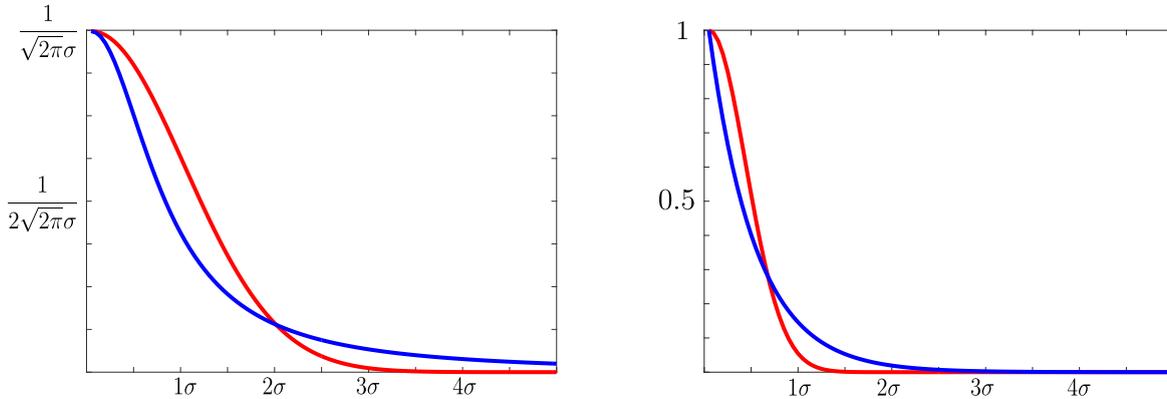


FIGURE 4.13: Comparaison entre le noyau gaussien (en rouge) et de Poisson (en bleu) dans le domaine spatial et fréquentiel, de gauche à droite respectivement.

#### 4.5.2.2 Noyaux de Poisson et théorie de l'espace d'échelle

Nous venons de voir que les noyaux gaussiens et de Poisson sont extrêmement proches en termes de formes et de caractéristiques. Cependant, ces seules observations ne suffisent pas à conclure que les noyaux de Poisson permettent de réaliser un espace d'échelle. En effet, nous avons mentionné dans le paragraphe 4.5.1 que des principes théoriques doivent être nécessairement respectés pour être en mesure de l'affirmer. Cependant dans [124], J. Babaud démontre l'unicité des noyaux gaussiens à vérifier l'ensemble des axiomatiques d'un espace d'échelle. Cette unicité est basée sur le fait que le noyau gaussien est la seule et unique solution de l'équation de diffusion de la chaleur. Par conséquent, c'est l'unique noyau à respecter le principe fondamental de causalité stipulé par la théorie de l'espace d'échelle. Ainsi, bien que les noyaux de Poisson soient des filtres linéaires, isotropes, invariants par rotation et changement d'échelle [106], ils ne sont pas solutions de l'équation de diffusion de la chaleur. De ce fait, ils ne respectent tout simplement pas le principe de causalité.

Néanmoins, M. Felsberg remet en cause ce principe de causalité en cherchant à le nuancer. Selon lui, il y a deux types de causalité : la *causalité forte* et la *causalité faible*. Et selon l'auteur, toutes deux permettent de définir un espace d'échelle.

Pour traduire la notion de causalité, J. Koenderink [120] a introduit la notion d'*isophote* (ou *courbe de niveau*).

**Definition 4.3.** Une isophote est une courbe de niveau définie dans l'espace d'échelle de telle sorte que :  $L(x, y, s) = \lambda \in \mathbb{R}$ .

En d'autres termes, l'isophote est une courbe qui relie les points de l'image ayant le même niveau d'intensité à travers l'espace d'échelle. Selon J. Koenderink, ces courbes de niveau doivent présenter leurs côtés convexes uniquement vers les échelles larges aux extrémums locaux, comme le montre la Figure 4.14.a. Pour respecter le principe de causalité, les isophotes ne peuvent donc pas présenter une concavité dirigée vers les échelles grossières, qui témoignerait

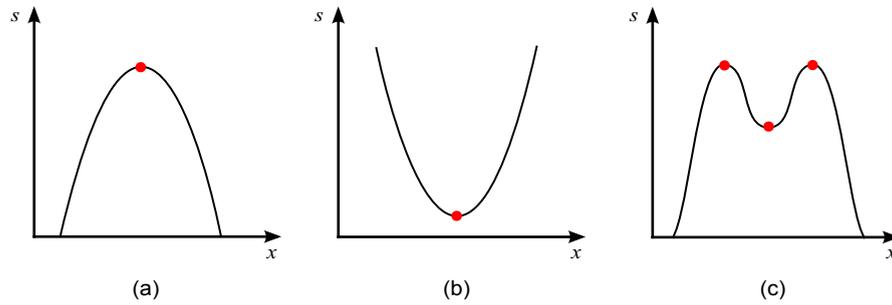


FIGURE 4.14: Illustration de plusieurs isophotes dans le cas 1D : en abscisse la variable de position  $x$  et en ordonnée la variable d'échelle  $s$ . (a) Isophote dont l'extremum local possède une convexité vers les échelles grossières. (b) Isophote dont l'extremum local possède une concavité vers les échelles grossières. (c) Isophote dont les deux extrémités sont reliées à l'image d'origine ( $s = 0$ ).

d'une apparition de structure lors de l'augmentation de l'échelle d'analyse (c'est le cas présenté sur la Figure 4.14.b). Ce principe est appelé principe de *causalité forte* par M. Felsberg.

Cependant, ce dernier définit également une notion de *causalité faible* qui consiste à dire que si l'isophote est connectée au plan  $s = 0$  (*i.e.* à l'image d'origine) à ses deux extrémités, alors on respecte quand même le principe de causalité. C'est le cas illustré sur la Figure 4.14.c, où on observe un minimum local concave dans le sens des échelles croissantes sur l'isophote, mais qui est connectée au plan  $s = 0$  aux extrémités. Il démontre dans [106] que les noyaux de Poisson respectent cette causalité faible, et donc qu'ils peuvent engendrer un espace d'échelle. Cet espace d'échelle n'est alors plus dirigé par les lois de l'équation de diffusion de la chaleur, mais par l'équation que vérifient les noyaux de Poisson, à savoir l'équation de Laplace.

R. Duits va même plus loin dans [125] en démontrant que les espaces d'échelle gaussiens et de Poisson sont reliés par l'opérateur Laplacien fractionnaire. Cette opérateur crée un continuum d'espaces d'échelle, appelés les *espaces d'échelle  $\alpha$*  (de l'anglais  $\alpha$  *scale-space*), où  $0 < \alpha \leq 1$ , et dont les espaces d'échelle gaussiens et de Poisson ne sont que des cas particuliers (pour  $\alpha = 1$  et  $\alpha = 1/2$  respectivement).

La Figure 4.15 illustre l'application de tels filtres de Poisson pour des valeurs d'échelles croissantes. Comme dans le cas des filtres gaussiens (voir Figure 4.10), on y remarque que les détails fins, comme les cheveux de la personne ou encore les différents bords (par exemple le bord des épaules) sont très bien définis pour de petites valeurs d'échelles. Au contraire, pour de grandes valeurs d'échelles, les détails sont supprimés et les contours sont rendus flous (*i.e.* une localisation moins précise).

Ainsi, les noyaux de Poisson respectent les axiomes de l'espace d'échelle et peuvent donc permettre une décomposition d'image sous une forme de famille d'images lissées à un paramètre (le paramètre du noyau de Poisson, noté  $s$ ). Cet espace d'échelle est simplement nommé espace d'échelle de Poisson.

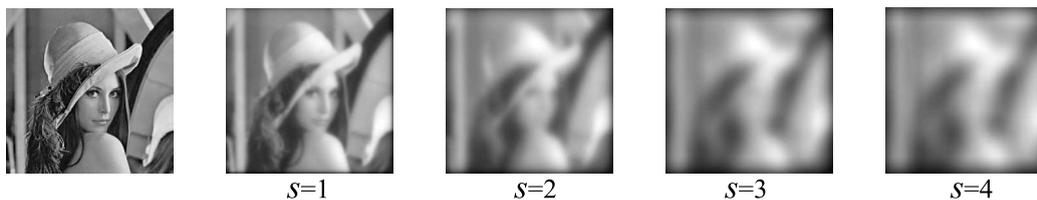


FIGURE 4.15: Résultats de l'application d'un noyau de Poisson sur l'image Lena (image la plus à gauche) pour des valeurs de  $s$  croissantes.

### 4.5.2.3 L'espace d'échelle de Poisson

Tout comme l'espace d'échelle gaussien présenté dans la section 4.5.1, l'espace d'échelle de Poisson permet donc de décomposer une image en une succession d'images lissées respectant les principes et axiomes définis par Iijima [115], Witkin [116] ou encore Koenderink [120].

Cette décomposition consiste à analyser une image  $I$  telle que :

$$\begin{aligned} \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\rightarrow I(x, y) \end{aligned} \quad (4.52)$$

et de considérer que sa représentation dans l'espace d'échelle est une autre image  $L$  :

$$\begin{aligned} \mathbb{R}^2 \times \mathbb{R} &\rightarrow \mathbb{R} \\ (x, y, s) &\rightarrow L(x, y, s) \end{aligned} \quad (4.53)$$

de telle sorte que  $L(x, y, 0) = I(x, y)$  et  $L(x, y, s) = P(x, y, s) * I(x, y)$ . L'image résultante est alors une fonction des deux variables de position  $x$  et  $y$ , mais également d'un paramètre d'échelle  $s$ . Dans ce qui précède, nous avons vu que ce paramètre  $s$  est l'équivalent du paramètre  $\sigma$  pour une gaussienne. Ainsi, on peut faire une analogie entre le filtrage par des noyaux gaussiens (cf. paragraphe 4.5.1) et ceux de Poisson en supposant que le filtrage d'une image par un filtre de Poisson de paramètre  $s$  a pour effet de supprimer les détails dont les dimensions ont une dimension inférieure à  $2s$ .

La Figure 4.16 propose l'illustration d'un tel phénomène sur une image synthétique composée de 3 disques dont les diamètres respectifs sont de 10, 20 et 40 pixels. On propose de réaliser 5 filtrages à l'aide de filtres de Poisson pour des paramètres  $s = 2$ ,  $s = 5$ ,  $s = 10$ ,  $s = 20$  et  $s = 30$ . On remarque que pour  $s = 2$  (inférieure au plus petit rayon des trois cercles), les trois structures sont toujours clairement visibles et n'ont pas été altérées par le lissage. Par contre, lorsque  $s$  devient de l'ordre de grandeur du rayon d'un disque, celui-ci commence à être lissé (valeur maximale de la convolution au centre du disque, puis décroissance radiale). Lorsque  $s$  devient plus important que le rayon, le disque commence à disparaître. Il est à noter que la suppression d'un niveau de détail n'est pas directe d'une échelle à une autre (le disque de

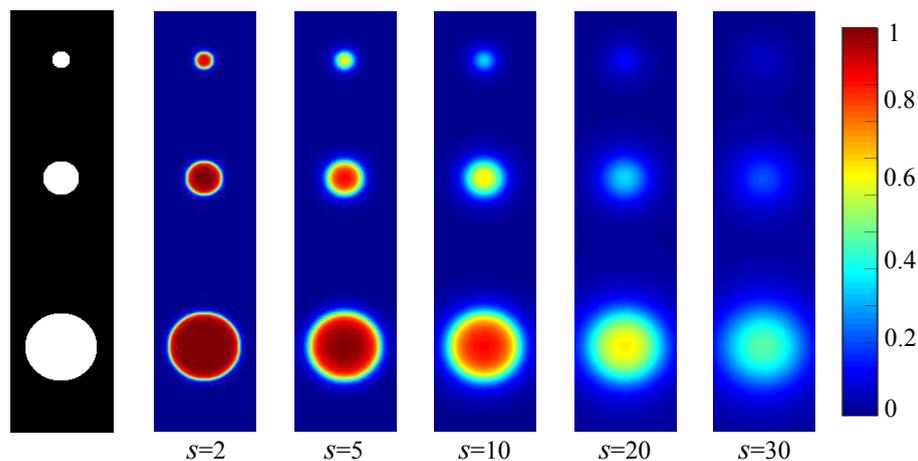


FIGURE 4.16: A gauche : Exemple d'une image contenant 3 disques de diamètres respectifs 10, 20 et 40 pixels. Cette image est lissée pour 4 valeurs croissantes du paramètre d'échelle  $s$  (de gauche à droite).

dimension 10 pixels est toujours légèrement observable pour une valeur  $s = 10$  par exemple), ce qui est dû à la sélectivité fréquentielle non-idéale des filtres de Poisson.

L'espace d'échelle est un outil dont l'utilité a été largement démontrée par des applications allant de la détection de points d'intérêt, jusqu'aux mosaïques d'images [126], en passant par les méthodes SIFT (Shift Invariant Feature Transform) [122]. La plupart de ces méthodes proposent d'utiliser l'espace d'échelle pour faire une analyse multi-échelle.

### 4.5.3 Analyse multi-échelle

Jusqu'à présent, nous avons vu que l'espace d'échelle correspondait à un cadre théorique pour supprimer au fur et à mesure les détails d'une image. La littérature a alors remarqué qu'il serait possible d'en tirer parti pour permettre de mettre en évidence les détails d'une dimension donnée. C'est ce que l'on appelle une analyse multi-échelle. Nous allons voir que cette analyse multi-échelle consiste dans le domaine fréquentiel à décomposer le spectre d'une image en différentes bandes-passantes.

Pour réaliser une analyse multi-échelle à partir de l'espace d'échelle, la méthode la plus efficace a été popularisée par les travaux de L.G. Lowe [122] à travers les méthodes SIFT. Dans cette étude, réalisée dans le cadre de l'espace d'échelle gaussien, l'auteur propose d'utiliser les différences de gaussiennes pour analyser une image à des niveaux de détails particuliers. Initialement, ces différences de gaussiennes furent introduite par les travaux de T. Lindeberg [5] et de J. Crowley [127].

Nous allons introduire dans un premier temps ces filtres particuliers, afin d'appliquer le même concept aux noyaux de Poisson. On réalisera ainsi une analyse multi-échelle via des différences de Poisson.

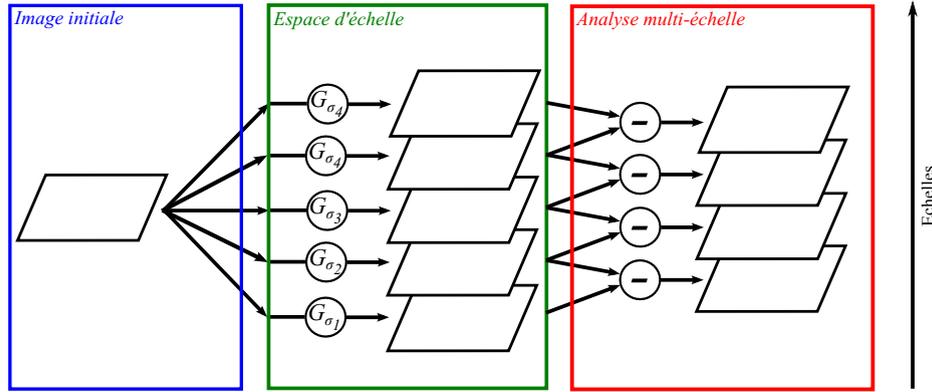


FIGURE 4.17: Processus de construction de l'analyse multi-échelle.

#### 4.5.3.1 Introduction aux différences de Gaussiennes

Rappelons que l'espace d'échelle gaussien est engendré par une variation du paramètre d'échelle  $\sigma$ , correspondant à l'écart-type d'une gaussienne  $G$  par laquelle on convolue une image de départ  $I$  :

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y). \quad (4.54)$$

Le noyau gaussien étant isotrope (comme nous l'avons vu sur la Figure 4.12), cette convolution a pour effet de filtrer les détails dont les dimensions ont un rayon inférieur à  $2\sigma$ . Maintenant, si l'on suit le raisonnement de L.G. Lowe et que l'on considère un second filtrage de l'image  $I$  par une gaussienne d'écart-type  $k\sigma$ , où  $k$  est un nombre réel tel que  $k > 1$ , on supprime alors les détails de dimensions inférieures à  $2k\sigma$ . Selon l'auteur, il est alors possible de conserver uniquement les détails dont les dimensions sont approximativement dans l'intervalle  $[2\sigma, 2k\sigma]$  en réalisant la différence de gaussiennes (de l'anglais *Difference of Gaussians* - DoG) :

$$D(x, y, \sigma, k) = L(x, y, \sigma) - L(x, y, k\sigma) = (G(x, y, \sigma) - G(x, y, k\sigma)) * I(x, y) \quad (4.55)$$

où  $k$  est donc un paramètre qui va influencer la finesse de la discrétisation en échelles que l'on désire. Plus  $k$  sera petit, plus l'intervalle  $[2\sigma, 2k\sigma]$  sera restreint. De ce fait, on se focalisera alors sur un niveau de détail particulier dans l'image.

**Remarque importante :** On note ici la différence entre une analyse en espace d'échelle et une analyse multi-échelle. Dans l'espace d'échelle, on supprime au fur et à mesure des détails. Dans l'analyse multi-échelle, on ne conserve que les détails d'une dimension souhaitée en réalisant des soustractions adéquates depuis l'espace d'échelle. La Figure 4.17 illustre ce processus de construction.

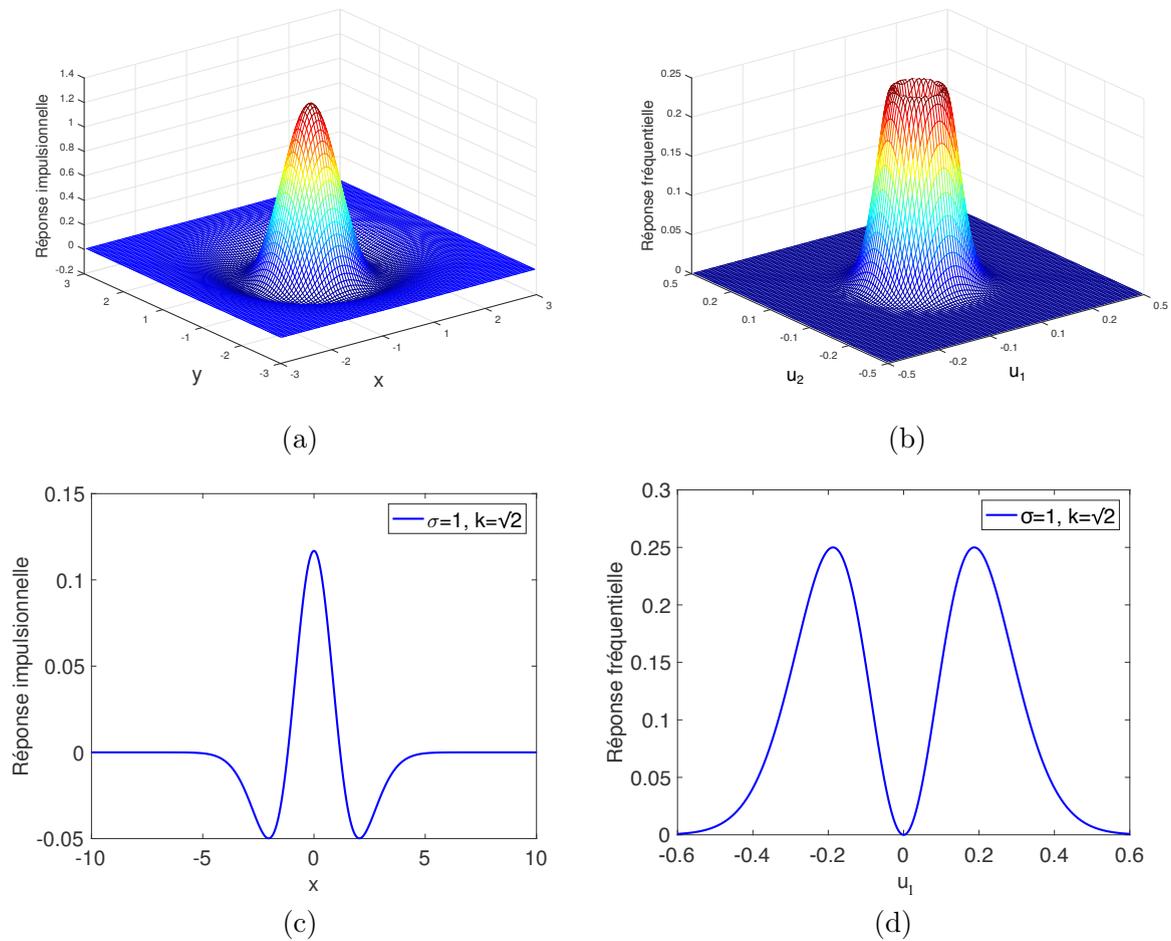


FIGURE 4.18: Représentation 3D spatiale (a) et fréquentielle (b) d'une différence de gaussienne pour  $\sigma = 1$  et  $k = \sqrt{2}$ . Une coupe 1D de ces deux noyaux est proposée en (c) et (d) respectivement.

### Le filtre DoG

Un filtre DoG est donc le résultat de la soustraction de deux noyaux gaussiens d'écart-types distincts. La Figure 4.18 représente un de ces filtres DoG pour des valeurs de  $\sigma$  et  $k$  fixées (par exemple ici  $\sigma = 1$  et  $k = \sqrt{2}$ ). On remarque qu'un DoG est une sorte de chapeau mexicain dans le domaine spatial. Dans le domaine des fréquences, ce DoG présente deux lobes symétriques par rapport à la fréquence nulle dans cette coupe 1D selon  $u$  (en 2D on verrait donc un anneau), ce qui est caractéristique d'un filtre passe-bande. Ainsi, lorsqu'on se focalise sur un niveau de détail dans l'image en réalisant une différence de gaussiennes, on étudie une bande spectrale particulière du signal 2D. Le paramétrage de  $\sigma$  et de  $k$  influence ce niveau de détail et par conséquent les caractéristiques de la bande-passante étudiée. En faisant varier ces paramètres, on peut alors espérer analyser une image à différents niveaux de détails revenant à décomposer le spectre en différentes bandes-passantes, c'est **l'analyse multi-échelle**. Il est cependant nécessaire de fixer ce paramétrage.

## Les paramètres de l'analyse multi-échelle

Dans l'espace d'échelle, nous avons déjà discuté du paramétrage de  $\sigma$ , qui est relié au filtrage des détails de dimension inférieure  $2\sigma$ . La nouveauté ici est le filtrage des détails de dimension supérieure à  $2k\sigma$ . Plus  $k$  sera proche de 1 et plus les détails conservés dans l'image seront strictement proches d'une dimension de  $2\sigma$ . Dans le domaine spectral, cela signifie que la bande de fréquence étudiée est alors très sélective et donc très étroite. Dans ce cas, si l'on veut étudier tout le spectre du signal 2D, il est nécessaire de considérer de nombreuses autres bandes-passantes. Au contraire si l'on fixe une valeur de  $k$  plus grande, on aura une précision moins importante sur le niveau de détail associé à la différence de gaussienne (*i.e.* une bande-passante plus large) et donc la nécessité d'utiliser moins de filtres pour couvrir tout le spectre du signal.

Ainsi, le choix de  $k$  revient à un compromis entre la précision désirée et le nombre de filtres. Il a fait l'objet de nombreuses études afin de trouver la valeur qui discrétise au mieux les détails d'une image, parmi lesquelles celles de L.G. Lowe [122] ou D. Marr [128]. Alors que ce dernier propose une valeur de  $k = 1.6$  en se basant sur des études de sensibilités des filtres, L.G. Lowe décrit une manière efficace de construire l'analyse multi-échelle basée sur une division en octave du spectre, ce que nous allons maintenant expliquer.

## Le découpage en octaves

L'analyse d'un spectre en octaves est surtout connue dans le domaine l'acoustique [129] pour permettre une découpe du spectre et une analyse simplifiée d'un son, et en particulier d'un bruit. Le découpage en octaves propose de discrétiser le spectre en une succession de bandes  $[f_1, f_2]$  dont la largeur relative  $(f_2 - f_1)/f_1$  est constante, ce qui est reconnu pour être le processus de l'audition humaine. Pour respecter ce critère, il faut dans chaque bande que la plus haute fréquence soit le double de la plus basse fréquence (*i.e.*  $f_2 = 2f_1$ ). Chacune de ces bandes est alors appelée une octave. Ainsi, si l'on considère une fréquence initiale  $f_0$ , le découpage en octaves sera réalisée par un intervalle  $[f_0, 2f_0]$  suivi des intervalles  $[2f_0, 4f_0]$ ,  $[4f_0, 8f_0]$  et ainsi de suite. Le schéma de la Figure 4.19 propose de représenter un tel découpage pour un axe fréquentiel. On constate que le principal avantage d'une telle décomposition du

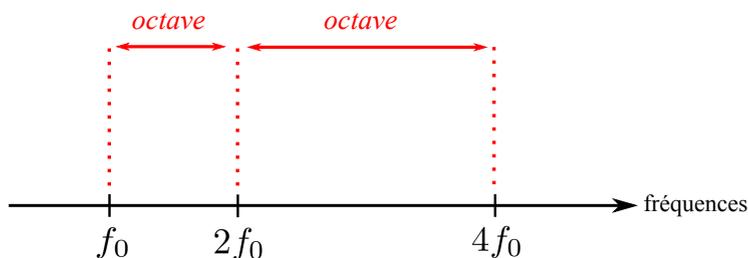


FIGURE 4.19: Schématisation d'une découpe en octaves d'un axe fréquentiel.

spectre est qu'elle nécessite uniquement la donnée d'une fréquence  $f_0$  initiale. Le découpage se réalise ensuite automatiquement à partir de ses multiples.

Pour réaliser sa décomposition multi-échelle, L.G. Lowe propose de partir de ce principe mais en discrétisant également chaque octave en un nombre  $n \in \mathbb{N}$  de sous-bandes. Il relie alors le facteur  $k$  de la différence de gaussiennes à  $n$  par  $k = 2^{1/n}$ . Ainsi, en faisant varier le nombre de sous-bandes  $n$  par octave, on discrétise le spectre avec plus ou moins de finesse.

### La décomposition multi-échelle

En pratique, l'avantage d'une décomposition en octaves, c'est que seuls trois paramètres doivent être fixés : le nombre  $m$  d'octaves sur lesquels on désire analyser l'image, le nombre  $n$  de décompositions par octave et enfin une première valeur de l'écart-type, que l'on notera  $\sigma^{(1)}$ , pour définir le premier niveau de décomposition que l'on désire (c'est la donnée d'entrée à partir de laquelle la décomposition se fera ensuite automatiquement en octaves, cf. rôle de  $f_0$  précédemment). Cette dernière est reliée à la taille minimale des détails  $2\sigma^{(1)}$  que nous souhaitons mettre en évidence dans l'image. Ainsi, l'analyse multi-échelle contient  $N = m \times n$  sous-bandes (ou niveaux de décomposition). Chacune de ces sous-bandes correspond alors à un filtrage par une différence de gaussiennes. Si l'on note  $i \in \{1, \dots, N\}$  l'indice d'une sous-bande (sachant que  $i = 1$  correspond à la première sous-bande *i.e.* le niveau de détail le plus fin), celle-ci est alors associée à la différence de gaussienne dont les deux écarts-types sont  $\sigma^{(i)}$  et  $k\sigma^{(i)}$  avec la relation de récurrence  $\sigma^{(i)} = k^{(i-1)}\sigma^{(1)}$  (cf. Figure 4.20). Par conséquent, les niveaux de détails associés à cette dernière sont compris dans l'intervalle  $[2\sigma^{(i)}, 2k\sigma^{(i)}]$ .

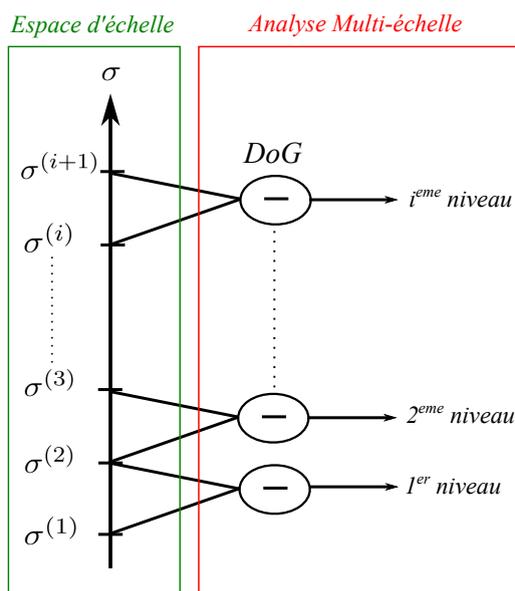


FIGURE 4.20: Construction successive des filtres DoG.

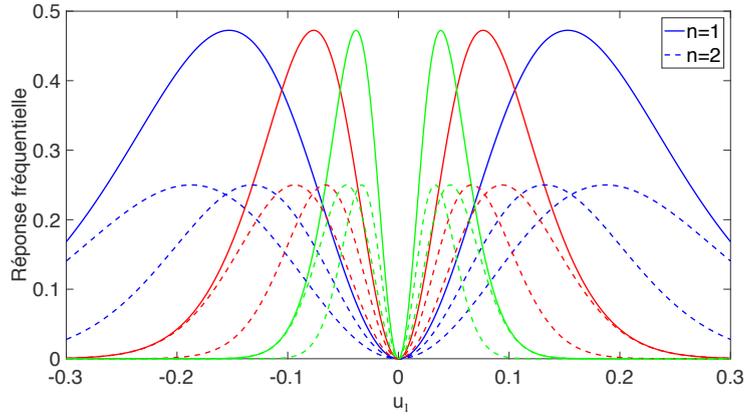


FIGURE 4.21: Discrétisation en fréquence des différences de gaussiennes en fonction du paramètre  $n$ . En trait plein, une seule décomposition par octave est réalisée (un octave par couleur). En pointillé, deux décompositions par octave est cette fois réalisée (même couleur pour les décompositions appartenant au même octave).

Par exemple, si l'on se fixe  $n = 1$  et  $m = 3$ , on utilise une seule décomposition pour chacune des trois octaves. On a donc par définition  $k = 2^{1/1} = 2$ . Dans ce cas, la première différence de gaussiennes se fait à l'aide des deux écarts-types  $\sigma^{(1)}$  et  $k\sigma^{(1)}$ . Cette première décomposition correspond alors à des niveaux de détails compris dans l'intervalle  $[2\sigma^{(1)}, 4\sigma^{(1)}]$ . La seconde décomposition est réalisée par une différence de gaussiennes dont les écarts-types sont  $\sigma^{(2)} = k\sigma^{(1)}$  et  $k\sigma^{(2)}$ . On obtient alors ici des détails compris dans l'intervalle  $[4\sigma^{(1)}, 8\sigma^{(1)}]$ . La troisième décomposition utilise ensuite les écarts-types  $\sigma^{(3)} = k\sigma^{(2)}$  et  $k\sigma^{(3)}$  pour obtenir des détails compris dans l'intervalle  $[8\sigma^{(1)}, 16\sigma^{(1)}]$ . La Figure 4.21 propose de représenter ces trois décompositions en traits pleins (avec une couleur différente pour chacun). La première octave est tracée en vert, la seconde en rouge et enfin la troisième en bleu. Du fait de la décomposition en octaves, on observe que la bande de fréquence grandit avec l'augmentation du niveau de détail.

Maintenant, si l'on désire deux décompositions par octave, on fixe alors  $n = 2$  et par conséquent  $k = 2^{1/2} = \sqrt{2}$ . On observe alors sur cette même Figure que les 3 octaves précédentes sont alors subdivisées chacune en 2 sous-bandes, chacune représentée de la même couleur que l'octave en question, mais en pointillés. Il y a alors maintenant 6 niveaux de décomposition ici au lieu de 3 précédemment. Les bandes de fréquence sont alors plus sélectives (étroites) mais avec plus de recouvrement.

Nous venons de démontrer ici qu'il est possible de réaliser une décomposition d'images sur différentes bandes-passantes par l'intermédiaire de soustraction dans l'espace d'échelle. Cependant, nous nous sommes contentés d'étudier cette analyse multi-échelle dans le cas d'un espace d'échelle gaussien. Or, nous avons vu que l'espace d'échelle de Poisson est également possible. De plus, les noyaux de Poisson apparaissent naturellement dans l'expression du signal monogène que nous souhaitons utiliser dans ces travaux. De ce fait, nous proposons de réaliser

cette même analyse multi-échelle, mais en remplaçant les noyaux gaussiens par des noyaux de Poisson. Par conséquent, nous pourrions décomposer notre image en différentes bandes-passantes et ensuite rendre les informations issues du signal monogène enfin significatives.

### 4.5.3.2 Application aux différences de Poisson

Puisque les noyaux de Poisson permettent une représentation sous forme d'espace d'échelle, nous pouvons envisager de réaliser une analyse multi-échelle des images à l'aide, cette fois, de différences de noyaux de Poisson (appelée DoP - *difference of Poisson*). Un DoP peut être exprimé de la manière suivante :

$$B(x, y, s, k) = P(x, y, s) - P(x, y, ks), \quad (4.56)$$

où  $ks$  et  $s$  sont les paramètres des deux noyaux de Poisson, avec  $k$  un nombre réel tel que  $k > 1$ .

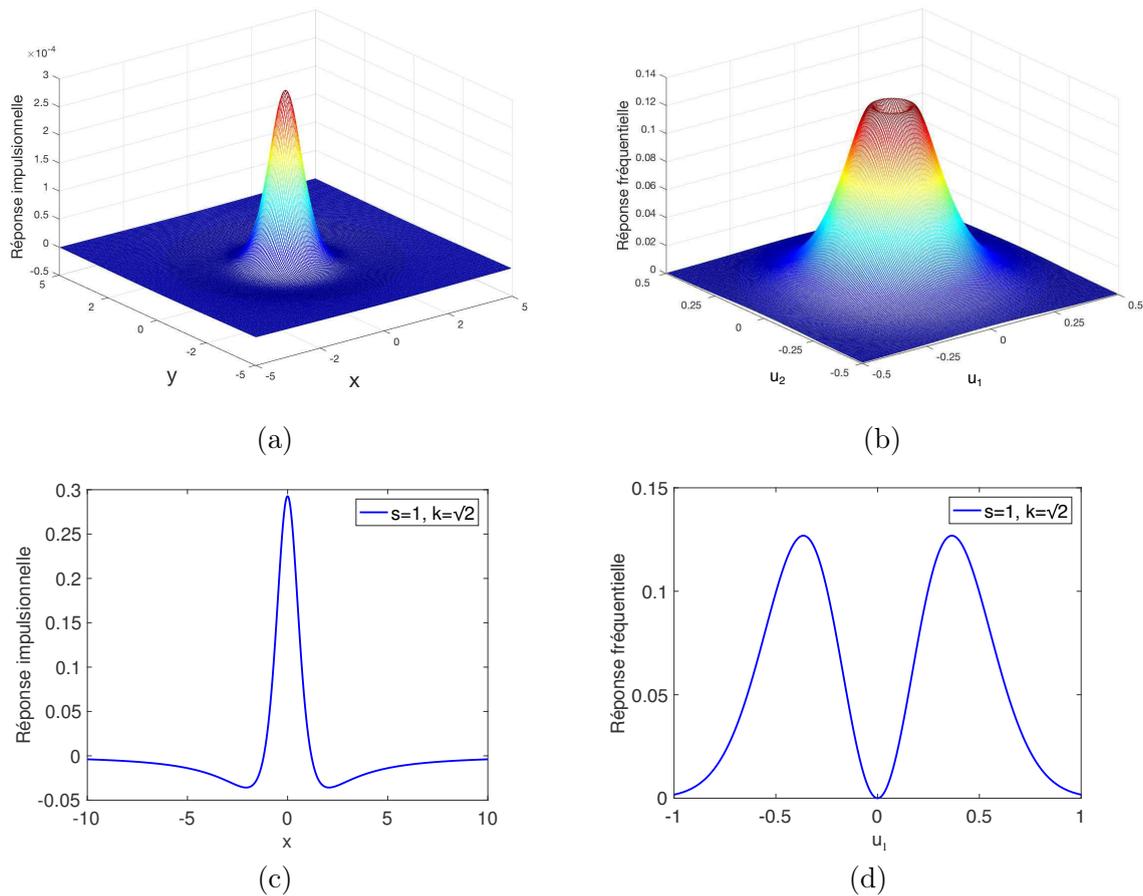


FIGURE 4.22: Représentation 3D spatiale (a) et fréquentielle (b) d'une différence de Poisson pour  $s = 1$  et  $k = \sqrt{2}$ . Une coupe 1D de ces deux noyaux est proposée en (c) et (d) respectivement.

La Figure 4.22 propose une représentation spatiale (a) et fréquentielle (b) d'un filtre DoP en 3D, ainsi qu'une coupe 1D de ces derniers en (c) et (d) respectivement. Tout comme le filtre DoG, on remarque que la forme spatiale du filtre fait penser au chapeau mexicain. Cependant, la partie négative du filtre est moins marquée que dans le cas du DoG illustré à la Figure 4.18(c). La représentation fréquentielle est également semblable au DoG puisqu'on observe que ce filtre est un filtre passe-bande. Cependant, on peut constater que l'étalement d'un filtre DoP est légèrement plus importante que celle d'un DoG dans le domaine fréquentiel. Ce qui induit une sélectivité fréquentielle moins forte.

Nous avons vu dans la section 4.5.2.3 qu'un noyau de Poisson, de paramètre  $s$ , permet de supprimer les détails de l'image dont les dimensions sont inférieures à  $2s$ . Nous pouvons alors réaliser le même raisonnement que L.G. Lowe concernant les différences de noyaux gaussiens mais avec cette fois la différence de noyaux de Poisson. Ainsi, le filtre  $B$  précédent permet de mettre en évidence sur l'image les détails ayant des dimensions appartenant approximativement à l'intervalle  $[2s, 2ks]$ .

De ce fait, nous proposons d'appliquer toute la théorie multi-échelle que nous venons d'expliquer dans le paragraphe précédent en remplaçant uniquement les filtres gaussiens  $G$  par des filtres de Poisson  $P$ . La seule différence est que le paramètre de ces filtres n'est plus  $\sigma$  mais  $s$ . Le découpage en octave est également conservé pour paramétrer le facteur  $k$ , à savoir  $k = 2^{1/n}$  où  $n \in \mathbb{N}$  est le nombre de subdivisions dans une octave. La Figure 4.23 propose d'illustrer le cas d'une décomposition en 3 octaves du spectre pour  $n = 1$  (trait plein) et  $n = 2$  (pointillé). Elle permet de constater que la dimension des bandes-passantes est pour importante dans le cas de filtres DoP que dans le cas de filtres DoG que nous avons vu à la Figure 4.21. On constate également qu'il y a recouvrement assez important entre les différents filtres DoP d'une même octave (pour  $n = 2$ ).

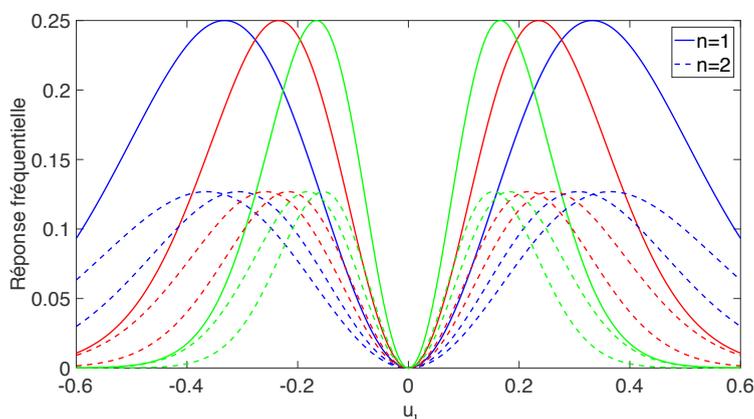


FIGURE 4.23: Discretisation en fréquence des différences de Poisson en fonction du paramètre  $n$ . En trait plein, une seule décomposition par octave est réalisée (un octave par couleur). En pointillé, deux décompositions par octave est cette fois réalisée (même couleur pour les décompositions appartenant au même octave).

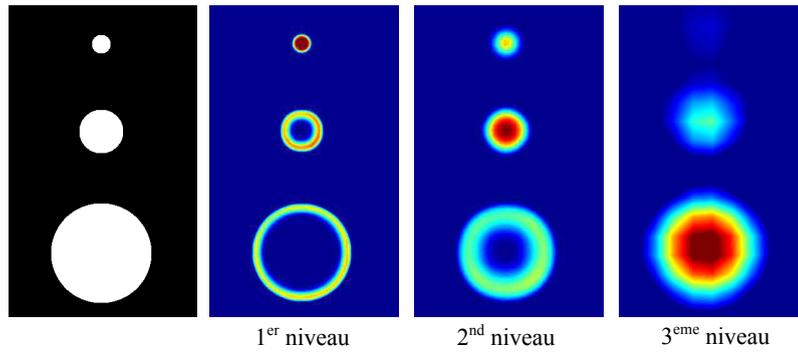


FIGURE 4.24: A gauche : Exemple d'une image contenant 3 disques de diamètres respectifs 5, 10 et 20 pixels. A droite : Résultats du filtrage de cette image par un filtre DoP pour 3 niveaux de décomposition.

Il est à noter que M. Felsberg mentionne ces filtres DoP dans [109]. Cependant, ses travaux se situent majoritairement dans le domaine de la recherche mathématique plutôt que dans l'application spécifique au traitement d'images. De ce fait, il ne propose pas une analyse multi-échelle par l'intermédiaire de ces filtres, il s'arrête à la notion d'espace d'échelle de Poisson dans [106]. Ainsi, ce que nous proposons ici est d'aller un cran plus loin dans l'utilisation de ces filtres en analysant une image pour des niveaux de détails.

Afin d'illustrer notre approche multi-échelle, nous proposons de considérer l'image synthétique de la Figure 4.24 (à gauche). Cette dernière est composée de 3 disques de diamètres 5, 10 et 20 pixels respectivement. On propose ici de réaliser une analyse multi-échelle sur  $m = 3$  octaves chacune décomposée en  $n = 1$  sous-bande. On fixe la valeur initiale  $s^{(1)}$  à 2 pixels. Ainsi, le premier niveau de décomposition est réalisé par un filtre DoP dont les deux paramètres sont  $s^{(1)} = 2$  pixels et  $ks^{(1)} = 4$  pixels. En théorie, cette première décomposition doit alors correspondre à des niveaux de détails compris dans l'intervalle  $[4, 8]$  pixels. On constate alors que l'image filtrée met en évidence le disque le plus petit. En effet, sa dimension (diamètre de 5 pixels) appartient à l'intervalle correspondant au niveau de détail étudié ici. On note également que le bord des deux autres disques réagissent aussi puisque un bord est composé de nombreuses fréquences et en particulier de très hautes fréquences, le filtre passe-bande DoP va donc réagir à ces détails fins ici. Le second niveau de décomposition est réalisé par un filtre DoP dont les paramètres sont cette fois  $s^{(2)} = ks^{(1)} = 4$  pixels et  $ks^{(2)} = k^2s^{(1)} = 8$  pixels, soit des niveaux de détails compris dans l'intervalle  $[8, 16]$  pixels. On remarque que cette fois c'est le disque du milieu qui réagit le plus à ce filtrage. En effet, son diamètre de 10 pixels appartient à l'intervalle précédent. De plus, on constate que le petit disque commence à disparaître de l'image (il se "floute"). Le troisième et dernier niveau de décomposition est réalisé par un filtre DoP dont les paramètres sont  $s^{(3)} = ks^{(2)} = k^2s^{(1)} = 8$  pixels et  $ks^{(3)} = k^2s^{(2)} = k^3s^{(1)} = 16$  pixels, soit des niveaux de détails compris dans l'intervalle  $[16, 32]$  pixels. La encore, on constate que c'est le disque dont la dimension appartient à cet intervalle qui réagit de manière forte au filtre, c'est-à-dire le troisième ici.

A travers ce paragraphe, nous avons constaté que notre analyse multi-échelle basée sur l'utilisation de filtres DoP permet de mettre en évidence des détails d'une dimension donnée. Maintenant que nous avons un tel outil, nous allons voir qu'il est possible d'étendre la notion de signal monogène à l'analyse multi-échelle.

#### 4.5.4 Représentation monogène multi-échelle

Nous avons mentionné au paragraphe 4.4.2.1 que le signal monogène permettait d'extraire d'un signal des informations significatives d'amplitude, d'orientation et de phase, uniquement si ce dernier est d'une fréquence bien définie, ou du moins à bande-étroite. Nous venons de voir que les noyaux de Poisson qui apparaissent naturellement dans la définition du signal monogène, permettent de définir un espace d'échelle puis une analyse multi-échelle via la différence des noyaux de Poisson. Nous allons voir que ces filtres passe-bande vont fournir au signal monogène le cadre multi-échelle qui permet de rendre significatives les informations d'amplitude, d'orientation et de phase.

##### 4.5.4.1 Définition

Rappelons que la définition générale du signal monogène d'un signal 2D  $f$ , peut être donnée par l'expression suivante :

$$f_M(\mathbf{x}, s) = [P(\mathbf{x}, s) * f(\mathbf{x}), P(\mathbf{x}, s) * \mathcal{R}_1\{f\}(\mathbf{x}), P(\mathbf{x}, s) * \mathcal{R}_2\{f\}(\mathbf{x})], \quad (4.57)$$

où  $\mathbf{x} = (x, y)$ ,  $s$  est un paramètre de lissage du filtre de Poisson  $P$  et où les opérateurs  $\mathcal{R}_j$  définissent les transformées de Riesz selon les dimensions  $j \in \{1, 2\}$  de l'image.

Pour un paramètre de lissage  $ks$ , l'expression du signal monogène de  $f$  devient le suivant :

$$f_M(\mathbf{x}, ks) = [P(\mathbf{x}, ks) * f(\mathbf{x}), P(\mathbf{x}, ks) * \mathcal{R}_1\{f\}(\mathbf{x}), P(\mathbf{x}, ks) * \mathcal{R}_2\{f\}(\mathbf{x})]. \quad (4.58)$$

On peut alors faire intervenir un filtre DoP, noté  $B$  dans le paragraphe précédent, en réalisant la différence des deux équations précédentes :

$$f_M(\mathbf{x}, s) - f_M(\mathbf{x}, ks) = [B(\mathbf{x}, s, k) * f(\mathbf{x}), B(\mathbf{x}, s, k) * \mathcal{R}_1\{f\}(\mathbf{x}), B(\mathbf{x}, s, k) * \mathcal{R}_2\{f\}(\mathbf{x})], \quad (4.59)$$

où  $B(\mathbf{x}, s, k) = P(\mathbf{x}, s) - P(\mathbf{x}, ks)$ . Si l'on note  $L(\mathbf{x}, s, k) = B(\mathbf{x}, s, k) * f(\mathbf{x})$  le filtrage du signal  $f$  par le filtre DoP, on obtient l'expression suivante par linéarité de la convolution :

$$f_M(\mathbf{x}, s) - f_M(\mathbf{x}, ks) = [L(\mathbf{x}, s, k), \mathcal{R}_1\{L\}(\mathbf{x}, s, k), \mathcal{R}_2\{L\}(\mathbf{x}, s, k)]. \quad (4.60)$$

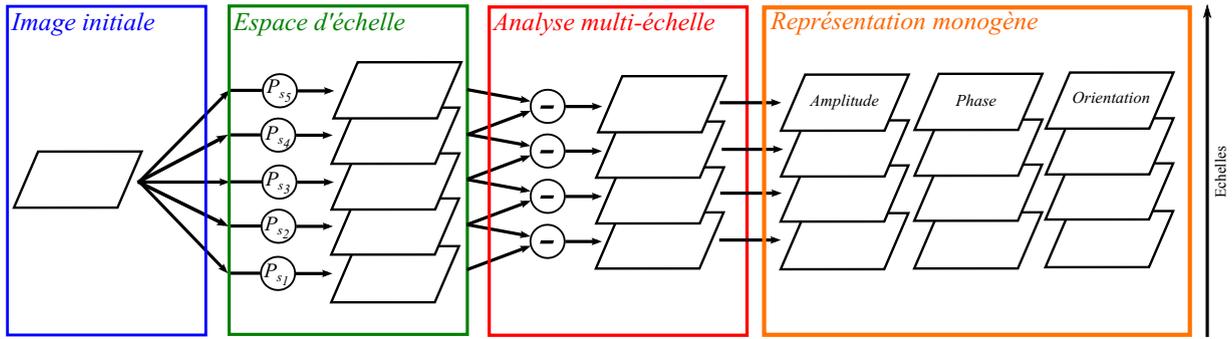


FIGURE 4.25: Processus de construction de la représentation monogène multi-échelle.

Si l'on regarde le membre de droite de cette équation, on peut interpréter cette quantité comme le signal monogène associé au signal  $L$ . Afin de distinguer cette expression du signal monogène de celle présentée aux équations (4.57) et (4.58), nous la noterons dans la suite de ce manuscrit :

$$f_M^{(s)}(\mathbf{x}) = [L(\mathbf{x}, s), \mathcal{R}_1\{L\}(\mathbf{x}, s), \mathcal{R}_2\{L\}(\mathbf{x}, s)]. \quad (4.61)$$

$f_M^{(s)}(\mathbf{x})$  est donc le signal monogène de  $f$  pour un niveau de détail associé au paramètre  $s$ . Nous oublions volontairement d'y mentionner le rôle de  $k$  puisque ce dernier peut être vu comme un hyper-paramètre dans le traitement. Il ne dépend uniquement que du nombre de niveaux de décomposition que l'on désire et est constant durant tout l'algorithme.

Puisque le signal  $L$  est à bande-étroite, nous pouvons extraire de son signal monogène  $f_M^{(s)}$  les informations d'amplitude, d'orientation et de phase qui seront cette fois considérées comme significatives. Nous les noterons alors  $A^{(s)}$ ,  $\theta^{(s)}$  et  $\phi^{(s)}$  respectivement :

$$A^{(s)}(\mathbf{x}) = \left| \mathbf{f}_M^{(s)}(\mathbf{x}) \right|, \quad (4.62)$$

$$\phi^{(s)}(\mathbf{x}) = \arg(L(\mathbf{x}, s) + j|\mathcal{R}\{L\}(\mathbf{x}, s)|), \quad (4.63)$$

$$\theta^{(s)}(\mathbf{x}) = \arg(\mathcal{R}\{L\}(\mathbf{x}, s)). \quad (4.64)$$

On peut alors imaginer que pour chaque niveau de détail de l'analyse multi-échelle, le signal monogène de l'image filtrée nous fournisse ces trois informations. L'ensemble de ces informations formeront la **représentation monogène multi-échelle** (cf. Figure 4.25).

#### 4.5.4.2 Application à une image synthétique

Nous proposons d'appliquer cette représentation à une image synthétique (cf. Figure 4.26) formée d'une sinusoïde 2D de propagation isotrope (à partir du centre de l'image). La sinusoïde est modulée de telle sorte que sa période au centre de l'image soit de 60 pixels et à ses extrémités qu'elle soit de 10 pixels.

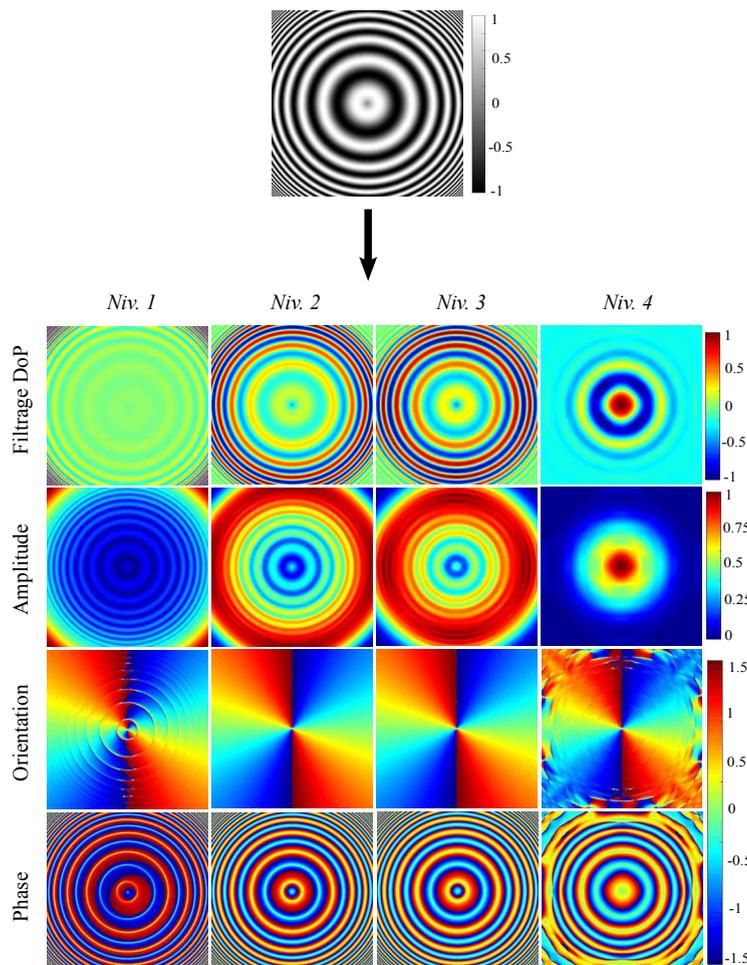


FIGURE 4.26: Illustration d'une représentation monogène multi-échelle à quatre niveaux de détail pour une image synthétique.

On configure alors notre analyse multi-échelle pour réaliser quatre niveaux de détails sur  $m = 4$  octaves (*i.e.*  $n = 1$  division par octave, soit  $k = 2$ ). Le niveau de détail initial est fixé par  $s^{(1)} = 2$  pixels de telle sorte que ces différents niveaux mettent successivement en évidence les détails de dimension 4, 8, 16 puis 32 pixels approximativement.

On constate alors que les réponses des filtres DoP sur cette sinusoïde sont différentes sur les 4 octaves. En effet, sur le premier niveau de décomposition, le filtre DoP réagit fortement sur les hautes fréquences de la sinusoïde (là où la longueur d'onde est proche du niveau de détail fixé à 4 pixels). En augmentant le niveau de détail, les filtres DoP réagissent ensuite sur les longueurs d'onde plus importantes. Pour chacun de ces niveaux de détails, on observe que l'information d'amplitude monogène représente l'énergie de l'image filtrée. Elle renseigne donc sur la présence d'un signal dans la bande passante considérée. Plus l'amplitude locale est forte, et plus cela signifie que le signal initial possède une fréquence locale proche de la fréquence centrale du filtre passe-bande.

Les informations d'orientation et de phase semblent quant à elles moins dépendantes des

décompositions fréquentielles. Cela provient de leurs indépendances à l'énergie. En effet, si la sinusoïde n'est pas totalement filtrée (c'est le cas aux niveaux de détail 2 et 3, comme nous pouvons le voir par exemple au centre de l'image DoP), des ondulations sont toujours visibles, ainsi l'orientation et la phase fourniront des informations correctes. Cependant, on observe que dans les zones où le signal initial est totalement filtré, ces informations ne sont plus correctes. Par exemple, aux extrémités de l'image pour le dernier niveau de détail, la phase et l'orientation sont totalement déformées.

## 4.6 Application à l'image sonar

Dans cette section, nous allons nous intéresser à la représentation monogène des images sonar et analyser les informations que nous pouvons en extraire. Cependant, avant toute chose, il est nécessaire de paramétrer l'analyse multi-échelle pour qu'elle s'adapte à ce type de données particulier.

### 4.6.1 Paramétrage de l'analyse multi-échelle

Nous avons vu précédemment que les filtres DoP permettent de décomposer le spectre d'un signal en bandes-étroites, et ainsi d'extraire des informations d'amplitude, d'orientation et de phase à différents niveaux de détails. Puisque l'objectif du projet dans lequel se place ce travail est la reconnaissance automatique de cibles sous-marines, nous avons pris le parti de contraindre notre analyse multi-échelle en fonction des dimensions de ces dernières. Au cours des dernières décennies, une large gamme de mines sous-marines a été produite par l'industrie de l'armement. Parmi celles-ci, les plus classiques sont les mines sphériques, Mantas, Rockans ou encore cylindriques. Quelques exemples sont proposés sur la Figure 4.27. Nous allons donc maintenant chercher à paramétrer les bornes de notre analyse multi-échelle en fonction de leurs tailles minimales et maximales. Dans [130], on peut retrouver leurs dimensions typiques, à savoir de l'ordre de 0.30 m pour la plus petite des mines (la sphérique) jusqu'à 2 m pour la

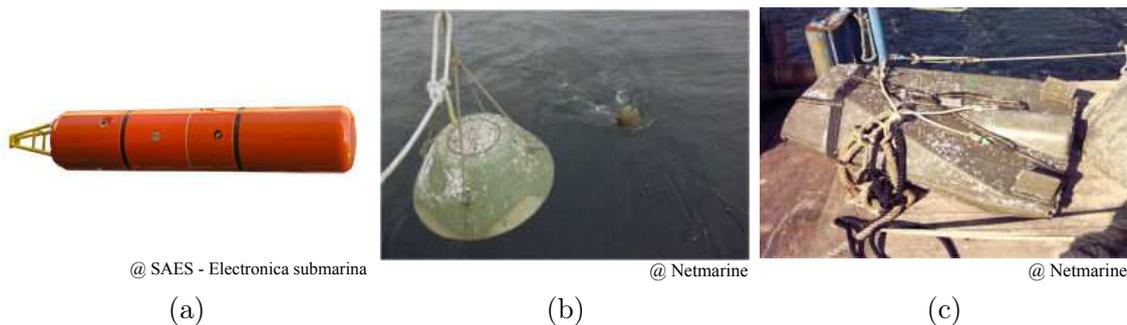


FIGURE 4.27: Exemple de trois mines sous-marines : (a) une cylindrique, (b) une Manta et (c) une Rockan.

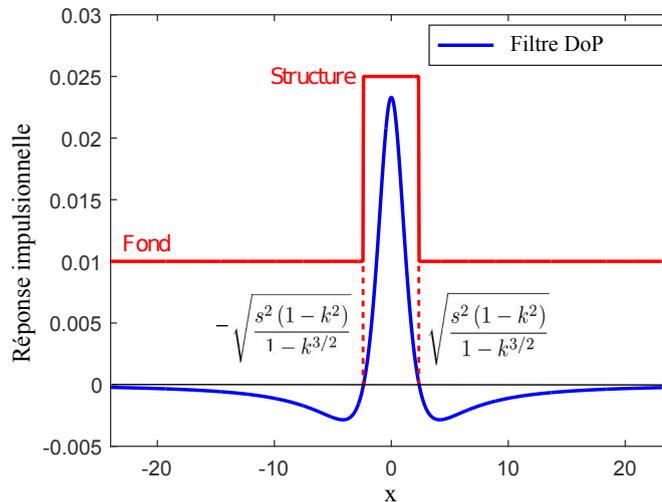


FIGURE 4.28: Dimensions d'une structure de l'image à la plus forte réponse du DoP en fonction des paramètres  $s$  et  $k$ .

plus grande (la mine cylindrique). Nous proposons alors de réaliser une analyse multi-échelle dans l'intervalle allant de 0.25 m à 2.5 m, afin d'être certain de pouvoir analyser toute cette gamme. Grâce à la dimension des pixels fournie par les données sonar, nous pourrions convertir en mètres les différents paramètres qui étaient jusqu'ici exprimés en pixels, et en particulier le paramètre  $s$  des filtres de Poisson.

Pour cela, considérons un modèle simple de structure sous-marine en une dimension, à savoir un écho plus fort que la réverbération du fond sur lequel elle est posée (cf. schéma rouge sur la Figure 4.28). Si l'on considère un filtre DoP (cf. courbe bleue sur la Figure 4.28), on peut alors définir ses paramètres  $s$  et  $k$  pour lesquels il aura une réponse maximale sur une structure de dimension  $l$  (en pixels) :

$$l = 2\sqrt{\frac{s^2(1-k^2)}{1-k^{3/2}}}. \quad (4.65)$$

Cette équation fait intervenir  $s$  et  $k$  telles deux inconnues et il n'est donc pas possible de la résoudre directement. Or, comme nous l'avons mentionné,  $k$  est un hyper-paramètre qui dépend de la discrétisation du spectre que l'on désire. Ainsi, il est possible de le fixer manuellement. Dans cette étude, nous proposons de fixer  $k = 2^{1/4}$ , ce qui revient à discrétiser chaque octave d'analyse en 4 sous-bandes. Il est très difficile de quantifier l'apport d'un tel paramétrage. Cependant, nous proposons ce nombre plutôt élevé afin d'obtenir une description fine de l'information contenue dans l'image, quitte à avoir un certain recouvrement entre les bandes-passantes pour des niveaux de détails proches.

Une fois ce paramètre fixé, il suffit de fixer la dimension  $l$  du plus petit détail que l'on souhaite analyser dans l'image ( $l = 0.25$  m ici). L'équation précédente ne devient alors qu'une

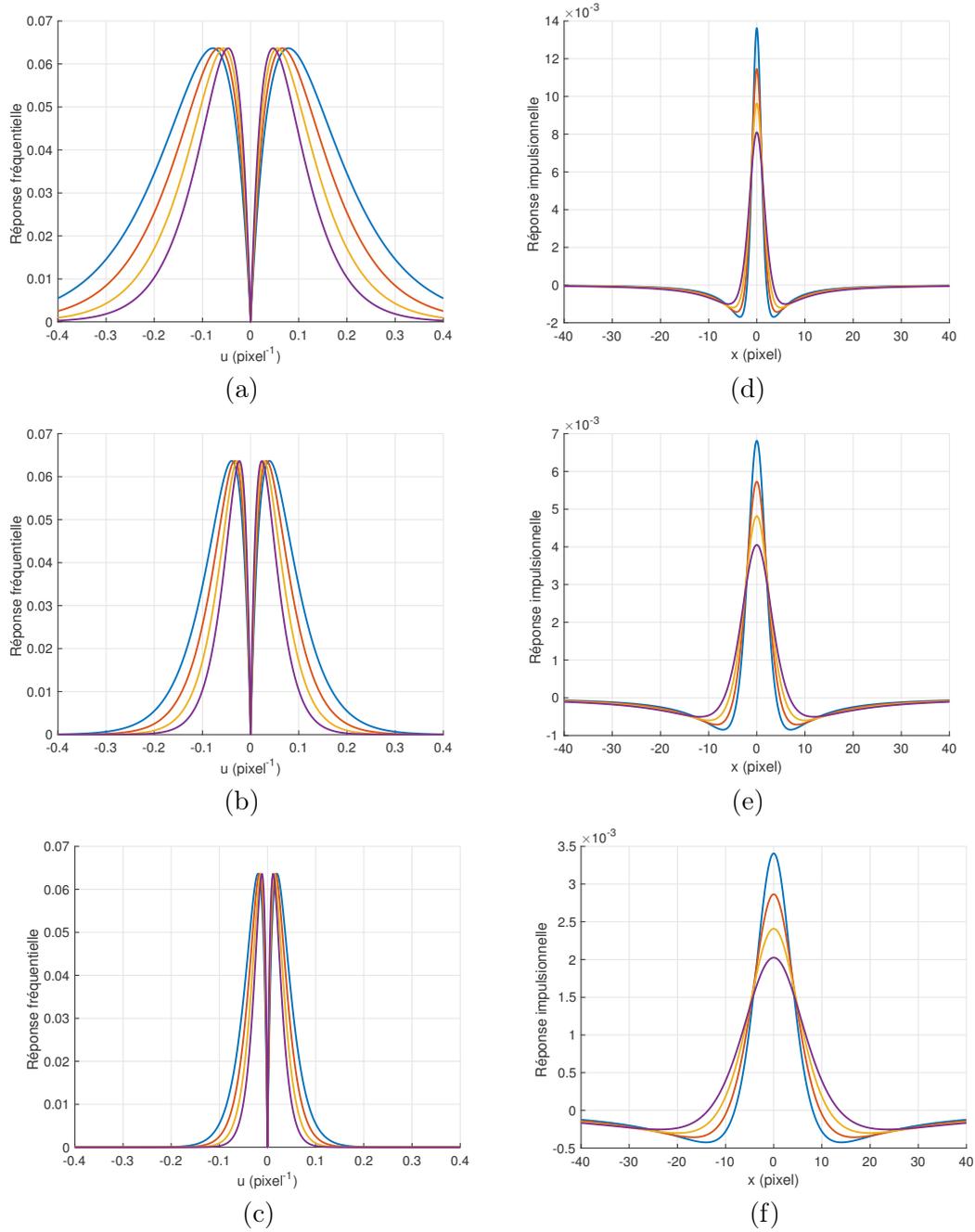


FIGURE 4.29: Représentation des 4 filtres passe-bande DoP pour chacun des 3 octaves dans le domaine fréquentiel en (a), (b) et (c), puis dans le domaine spatial en (d), (e) et (f).

équation à une seule inconnue, dont on extrait le paramètre  $s^{(1)}$  qui configure le premier noyau DoP de la décomposition. Le second noyau de la décomposition sera alors automatiquement défini comme étant le noyau DoP de paramètre  $s^{(2)} = ks^{(1)}$ , puis successivement  $s^{(3)} = k^2s^{(1)}$  pour le troisième etc... jusqu'à obtenir un niveau de détail correspondant à  $l = 2.5$  m. En prenant  $k = 2^{1/4}$ , l'intervalle  $[0.25\text{m } 2.5\text{m}]$  est alors discrétisé en 12 sous-bandes fréquentielles, qui représentent 3 octaves. Au final, nous obtenons alors 12 niveaux de détails dans notre analyse multi-échelle qui sont : 0.25, 0.32, 0.41, 0.5, 0.63, 0.8, 1, 1.25, 1.5, 1.8, 2.15 et 2.5 mètres.

La Figure 4.29 illustre, pour chacun des 3 octaves, les 4 décompositions dans les domaines spatial et fréquentiel. On y remarque la plus grande sélectivité de ces filtres en basses fréquences. De plus, on observe que ce sont bien les hautes fréquences qui correspondent aux détails les plus fins (faibles ouvertures spatiales des filtres DoP), alors que les basses fréquences correspondent aux détails grossiers (grandes ouvertures spatiales des filtres DoP).

Afin de mettre en évidence cette sélectivité des structures en fonction de leurs tailles, la Figure 4.30 illustre 4 de ces 12 niveaux de détail pour deux imagerie extraites de données sonar. Notons que chaque niveau de détail est donné en fonction de valeur de  $l$  correspondante (cf. équation (4.65)). Sur la première ligne, on observe qu'un objet de faible dimension réagit aux filtres DoP pour des détails fins, puis disparaît lorsque ces niveaux de détails deviennent trop grossiers. Pour observer les effets de ces filtres sur un objet plus large, la seconde ligne propose une imagerie représentant un petit voilier (un dériveur) posé sur le fond. On remarque qu'aux niveaux de fin détail, ce sont les bords de ce dernier qui réagissent, alors qu'en augmentant le niveau de détail, on met en évidence l'objet entier. Sur ces deux exemples, on observe également que le phénomène de bruit est surtout localisé aux bas niveaux de détail (caractère granulaire des images).

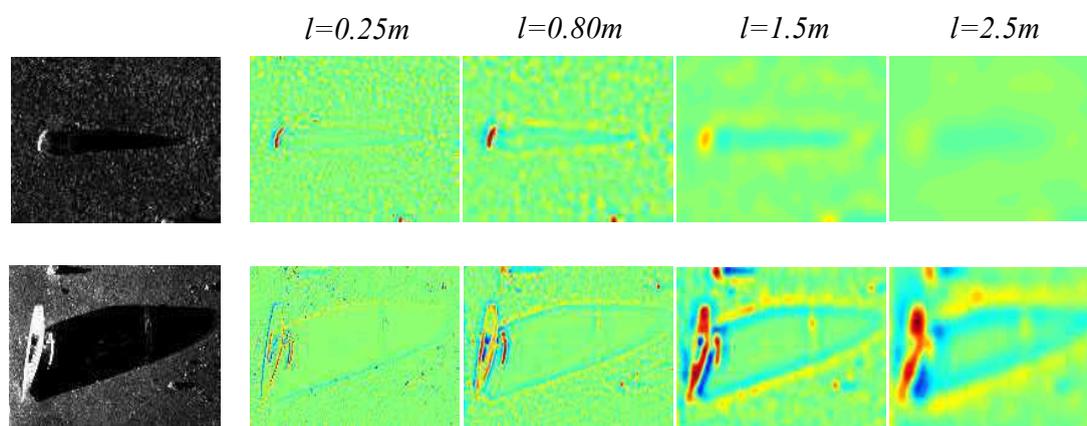


FIGURE 4.30: Sur la première ligne, exemple d'image sonar contenant un objet de petite dimension et ses réponses aux filtres DoP pour différents niveaux de détails. Sur la seconde ligne, les mêmes informations sont représentées mais pour une structure plus importante (on peut ici reconnaître un voilier).

### 4.6.2 Représentation monogène d'une image sonar latéral

Maintenant que le paramétrage de l'aspect multi-échelle a été réalisé, nous pouvons illustrer les informations issues du signal monogène pour des données issues d'un sonar latéral. Dans cette section, nous proposons de présenter trois cas distincts d'image sonar, à savoir le cas d'un objet posé sur un fond homogène, le cas d'un champ de rides de sable et enfin le cas d'une zone de fond plus complexe, encombrée par un amas de roches.

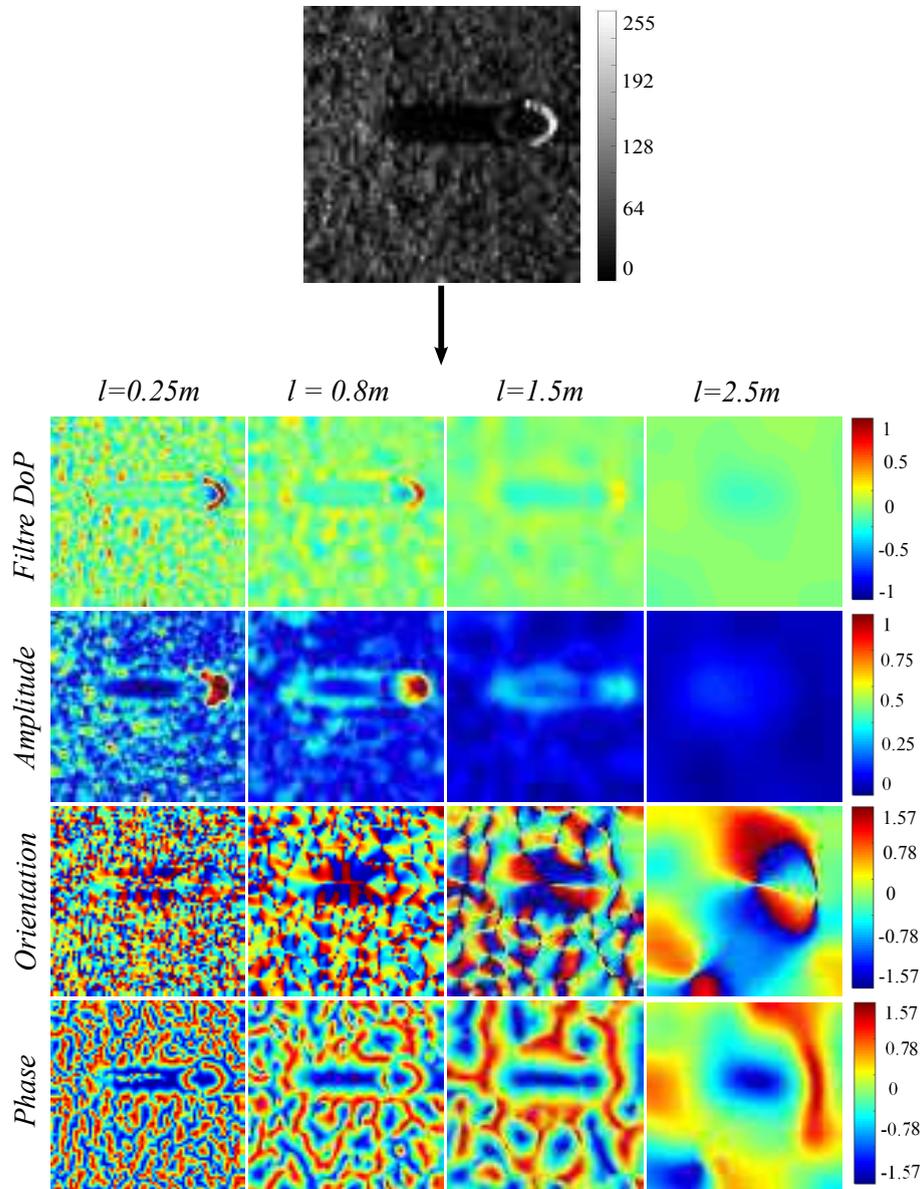


FIGURE 4.31: Représentation monogène multi-échelle d'un objet posé sur un fond homogène (Données DSTO).

### Cas d'un objet posé sur un fond homogène

La première illustration, proposée sur la Figure 4.31, est le cas d'une imagerie issue d'une image sonar latéral (base de données DSTO) qui représente un objet posé sur un fond sableux homogène. On y distingue alors aisément l'objet, caractérisé par son écho de forte amplitude suivi d'une ombre. Les informations monogènes sont alors proposées pour 4 niveaux de détails parmi les 12 possibles. On remarque que les informations énergétiques, à savoir les réponses des filtres DoP et les amplitudes monogènes associées, mettent en évidence l'écho de l'objet pour les deux premiers niveaux présentés ici. L'écho de cet objet étant de dimension faible, les niveaux de détails plus grossiers filtrent fortement l'objet jusqu'à le faire disparaître. Sur ces informations énergétiques, on peut également noter la présence de bruit. Ce dernier provoque dans l'image une variation d'intensité très forte et très localisée. Ainsi, on remarque que son influence dans la décomposition multi-échelle est limitée aux bas niveaux de détail. Dans le cas présenté ici, on peut noter que c'est la première échelle qui est principalement affectée par le phénomène. Si l'on observe maintenant l'information d'orientation locale, on remarque que ses valeurs sont relativement aléatoires dans l'image. En effet, le fond marin qui entoure l'objet est une zone constituée de sable où aucune structure ne faisant apparaître d'orientation particulière n'est présente. On peut enfin noter que la présence de l'objet dans cette zone n'est pas vraiment observable via cette information géométrique.

L'information de phase est également une quantité bruitée sur la zone de fond homogène entourant l'objet. Là aussi, l'absence de structure remarquable sur ce fond marin résulte en des valeurs de la phase aléatoires. Cependant, on remarque que dans la zone de l'objet et pour les niveaux de fin détail, la forme de ce dernier ressort par l'intermédiaire de valeurs particulières, autour de  $\pm\pi/2$  pour cet exemple. La forme circulaire de l'objet apparaît comme deux arcs de cercle que l'on observe également dans l'image initiale. De plus, on note que l'ombre de l'objet possède aussi une valeur particulière de la phase autour de  $-\pi/2$  ici.

### Cas d'un champ de rides de sable

Nous proposons ici d'illustrer ces informations monogènes sur un autre type de fond à savoir les rides de sable. La Figure 4.32 propose la même représentation monogène multi-échelle que précédemment pour une imagerie où un champ de rides de sable traverse une zone de sable homogène. Les filtres DoP réagissent particulièrement sur les rides de sables pour des niveaux de détails qui correspondent à la taille de ces dernières dans l'image sonar (*i.e.* ici entre 0.25 et 0.75m). Puisque l'amplitude monogène représente l'information énergétique des images filtrées, on peut noter qu'elle met en évidence les zones du fond constituées de structures dont les tailles correspondent au niveau de détail fixé. On remarque ici que le bruit intervient très fortement au premier niveau de décomposition, mais son influence décroît rapidement avec l'augmentation du niveau de détail. Les rides de sable étant des motifs orientés, l'information d'orientation

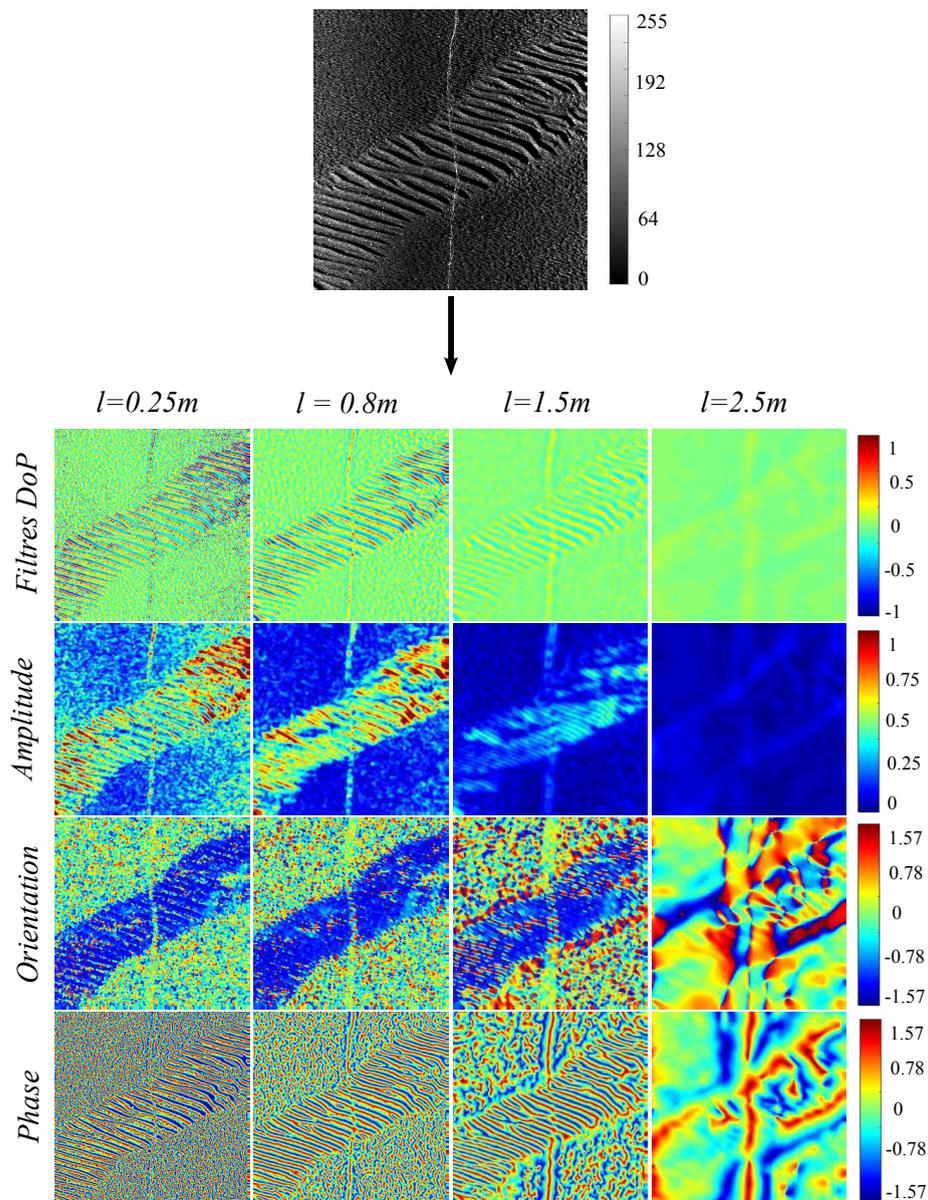


FIGURE 4.32: Représentation monogène multi-échelle d'un champ de rides de sable traversant l'image (Données DSTO).

monogène est cette fois significative. Si l'on observe cette information sur les niveaux de détail qui correspondent à la taille des rides de sable (*i.e.* entre le premier et le second niveau ici), l'orientation locale fournit des valeurs notables. En effet, le champ de rides de sable y apparaît comme une zone orientée de manière assez homogène. Au contraire, les zones de sable fin qui l'entourent possèdent quant à elles une orientation monogène aléatoire. Aux niveaux de détail grossiers, puisque les rides de sables sont filtrées, l'information d'orientation perd de sa signification. Concernant la phase locale, on observe que cette information est elle aussi d'intérêt sur les rides de sable. En effet, toujours aux niveaux de détail correspondants à la taille de ces rides, la phase locale suit les variations des rides dans un intervalle  $[-\pi/2 \pi/2]$ . L'analyse de la structure du champ de ride de sables semble donc être possible par l'intermédiaire de

cette information qui renseigne sur la position des crêtes et des vallées de la structure. Comme nous l'avons déjà remarqué sur la Figure 4.31, la phase locale ne semble pas être significative dans les zones homogènes, ce qui conduit à des valeurs aléatoires dans les régions entourant le champ de rides de sable ici.

### Cas d'une zone de fond complexe

La troisième et dernière zone que nous proposons dans ce paragraphe est le cas d'un fond marin plus complexe comme celui illustré sur la Figure 4.33. On y constate une forte densité de roches aux formes et tailles variées. Il y a alors sur l'image sonar une forte densité de paires écho-ombre, ce qui rend difficile une opération de chasse aux mines sous-marines.

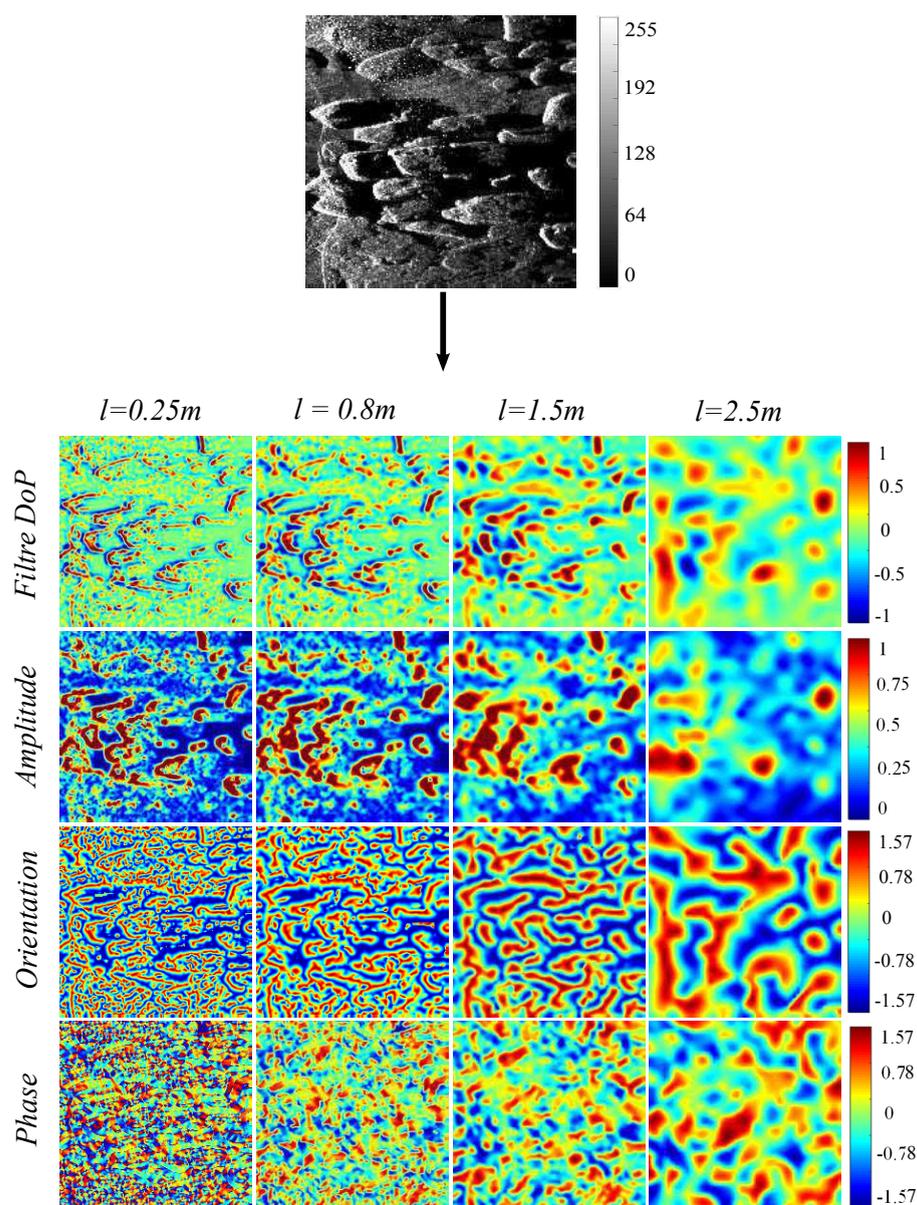


FIGURE 4.33: Représentation monogène multi-échelle d'un amas de roches (Données DSTO).

On remarque que les réponses des filtres DoP sont différentes en fonction du niveau de détail étudié. Alors qu'aux niveaux de fin détail ce sont les bords et contours des roches qui sont mis en évidence, aux niveaux de détail grossier ce sont les motifs de la roche dans sa globalité qui est mise en évidence. Ces observations sont les mêmes sur l'information d'amplitude. Concernant l'information de phase, on remarque son caractère très bruité aux niveaux de fin détail. Pour des niveaux plus grossiers, on voit alors apparaître dans cette information le contour des roches et des formes de structures que l'on retrouve dans l'image sonar initiale. On voit la aussi que les zones ombres semblent fournir des informations de phase au alentours de  $-\pi/2$  (cf. cas de l'objet posé sur un fond homogène). Concernant l'information d'orientation, on observe que celle-ci est également très bruitée sur l'ensemble des niveaux ici.

### 4.6.3 Application à la classification des fonds marins

Dans les sections précédentes, nous avons pu observer que le signal monogène permet de fournir des informations d'amplitude, d'orientation et de phase en chaque pixel de l'image sonar, dans un cadre multi-échelle. Cependant, l'ensemble de ces informations forment seulement une nouvelle représentation de l'image via une décomposition orthogonale multi-échelle, mais ne permettent pas encore d'extraire des informations qualitatives sur les fonds marins. Afin de démontrer que cette représentation est porteuse d'informations significatives pour l'analyse d'images sonar, nous proposons ici de réaliser une classification supervisée des fonds marins à l'aide de la représentation monogène.

Bien qu'une telle approche ne soit pas l'objectif visé par ces travaux, nous souhaitons tirer parti d'une base d'apprentissage disponible au laboratoire pour valider les potentialités du signal monogène, et ainsi démontrer que les informations qui en sont issues sont pertinentes et discriminantes.

#### 4.6.3.1 Présentation de l'approche

Au sein de l'ENSTA Bretagne, une base de données d'images sonar latéral KLEIN (base de données DGA présentée à la section 2.4 du chapitre 2) a été étudiée de manière à en extraire des imagerie représentatives des fonds marins. Ceux-ci sont regroupés en cinq classes correspondant à de la posidonie, de la roche, du sable, de la vase et des rides de sable. Chacune de ces classes est constituée de 60 imagerie. Un exemple d'imagerie pour ces 5 types de fonds est proposé sur la Figure 4.34.

Pour chaque imagerie, nous proposons alors d'extraire les informations monogènes multi-échelles. Chacune de ces informations est ensuite décrite à l'aide de sept descripteurs statistiques : la moyenne, l'écart-type, le skewness, le kurtosis et trois quantiles (3%, 75% et 97%). Ces différents descripteurs, ainsi que le label du fond marin associé, sont utilisés en entrée

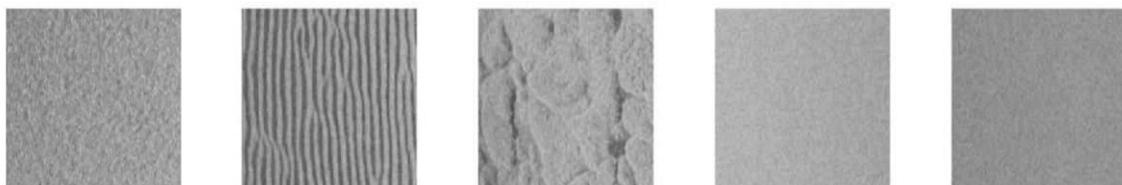


FIGURE 4.34: Illustration de la base d'apprentissage : (de gauche à droite) posidonie, rides de sable, roche, sable, et vase.

d'un classifieur Bayésien Naïf (par le biais du logiciel WEKA) qui crée alors un modèle de classification. On propose d'estimer la fiabilité du modèle ainsi trouvé par une technique de validation croisée.

#### 4.6.3.2 Matrices de Confusion

Afin d'illustrer les performances du classifieur, nous pouvons dresser la matrice de confusion sur les 5 types de fonds marins (cf. Table 4.2). On y remarque alors que les taux de bonne classification sont élevés pour l'ensemble des classes. Ils atteignent même 100% dans le cas des rides de sable. Ces résultats sont alors comparables à ceux obtenues par d'autres méthodes et en particulier celle de N. Lasmar [131] réalisée sur la même base de données. On peut alors en conclure que les informations monogènes semblent être dignes d'intérêt pour analyser les structures du fond. En particulier, de simples descripteurs statistiques issus de ces informations permettent de distinguer ces classes.

Nous avons vu dans la section 3.3 que d'un point de vue ATR, trois grandes catégories de fonds marins existent, à savoir les fonds homogènes, anisotropes, et complexes. Ici, nous proposons de réaliser une classification supervisée en ces 3 classes en regroupant certaines des classes précédentes. La classe homogène sera alors constituée par le sable et la vase, la classe anisotrope par les rides de sable, et les fonds complexes seront le regroupement de la posidonie et des roches.

Classes	Posidonie	Rides	Roche	Sable	Vase
Posidonie	97	0	3	0	0
Rides	0	100	0	0	0
Roche	2	0	98	0	0
Sable	0	0	0	94	6
Vase	0	0	0	3	97

TABLE 4.2: Matrice de confusion du classifieur Bayésien pour 5 classes de fonds marins.

Classes	Homogène	Anisotrope	Complexe
Homogène	98	2	0
Anisotrope	0	100	0
Complexe	0	0	100

TABLE 4.3: Matrice de confusion du classifieur Bayésien pour 3 classes de fonds marins.

Pour chacune de ces classes, on crée une base d'apprentissage composée de 60 imagettes qui représentent de manière équi-répartie les classes regroupées. A nouveau, les descripteurs statistiques des informations monogènes multi-échelle sont introduits dans un classifieur Bayésien Naïf qui fournit les taux de confusion présentés à la Table 4.3. On y observe à nouveau de très bons résultats de classification, qui mènent à un taux global de 99% de bonne classification.

Puisque l'étude de cette base de données semble fournir de bons résultats, il est envisageable d'appliquer le modèle appris par le classifieur à des images sonar entières.

#### 4.6.3.3 Classification d'une image sonar

Dans cette section, nous proposons d'utiliser les résultats précédents pour classifier une image sonar. La méthode utilisée est relativement simple car notre objectif est juste d'illustrer le fait que les informations portées par le signal monogène sont pertinentes pour notre étude. Pour cela, nous considérons que le modèle de classification trouvé précédemment peut s'appliquer à des blocs d'une image sonar. Ce découpage est réalisé de manière à ce que chaque bloc possède une taille similaire aux imagettes de la base d'apprentissage (ici 64x64 pixels). On définit également un taux de recouvrement entre les blocs successifs, à savoir 1/2 ici (les différents blocs se chevauchent à moitié en horizontal et en vertical). Pour chacun d'entre eux, on réalise alors le même protocole d'extraction de descripteurs que pour la base de données, mais cette fois les descripteurs sont utilisés pour tester le modèle appris et estimer son label.

Le résultat d'une telle classification est proposé sur la Figure 4.35 dans le cas de 5 classes. On remarque sur ces deux exemples que les rides de sable (de couleur bleue foncée) ainsi que les zones rocheuses (en rouge) sont les zones les mieux définies. Alors que sur la Figure 4.35.(a) les zones de sable semblent correctement classées, sur la Figure 4.35.(b) il semblerait y avoir une confusion entre le sable (en vert) et la vase (en jaune) en fonction de la portée. Celle-ci est certainement due aux pertes en propagation de l'onde acoustique qui résulte en une perte d'intensité du niveau des pixels en bout de portée (à droite de l'image). La vase étant une zone homogène de plus faible intensité que le sable, ces deux classes peuvent se confondre si le modèle de classification ne tient pas compte de la portée. On note également sur ces deux images, que les frontières des zones de rides de sable et des roches sont injustement classées

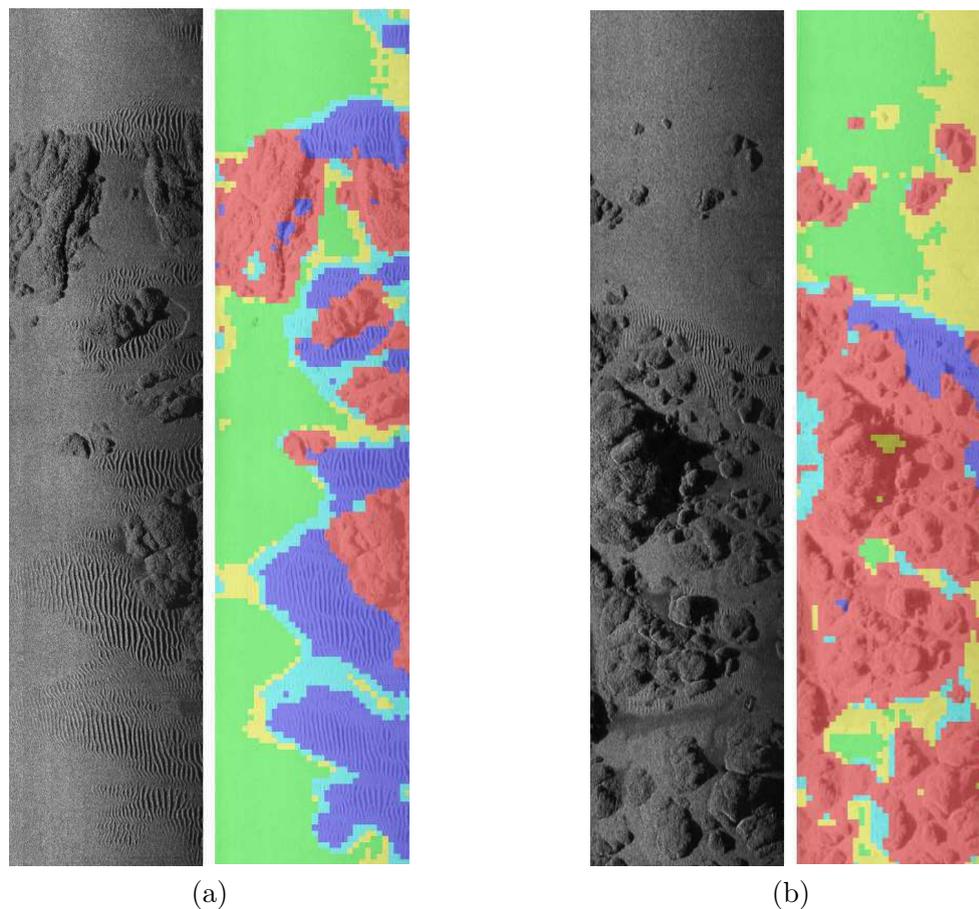


FIGURE 4.35: Deux exemples d'image sonar et leur classification respective en 5 classes (a) et (b). Les couleurs correspondant aux différentes classes sont les suivantes, jaune = vase, vert = sable, bleu clair = posidonie, bleu foncé = rides de sable et rouge = roche.

comme de la posidonie (en bleu ciel). Cela provient du découpage par blocs qui introduit une faible précision sur la définition des frontières (dans un même bloc plusieurs classes peuvent apparaître au niveau d'une frontière et biaiser l'estimation de la classe).

On propose également sur la Figure 4.36 le résultat de la classification en 3 classes sur ces mêmes images. On remarque alors que les différentes zones de fond marin sont classées correctement, à l'exception de certains bords de roche qui sont classés en zones anisotropes.

Cette classification supervisée de l'image permet de se rendre compte que les informations monogènes multi-échelle sont dignes d'intérêt pour réaliser une description des fonds marins. Nous avons utilisé ici des descripteurs statistiques du premier ordre et fait appel à un simple classifieur Bayésien. Cette méthode de classification supervisée n'étant pas le but de la thèse, aucune étude des performances en fonction de la taille des imagerie, du recouvrement entre les fenêtres ou encore du classifieur utilisé n'a été faite. Cependant les différents résultats de classification qui en ressortent, peuvent être considérés comme très corrects en comparaison avec d'autres approches utilisées sur ces mêmes données [132].

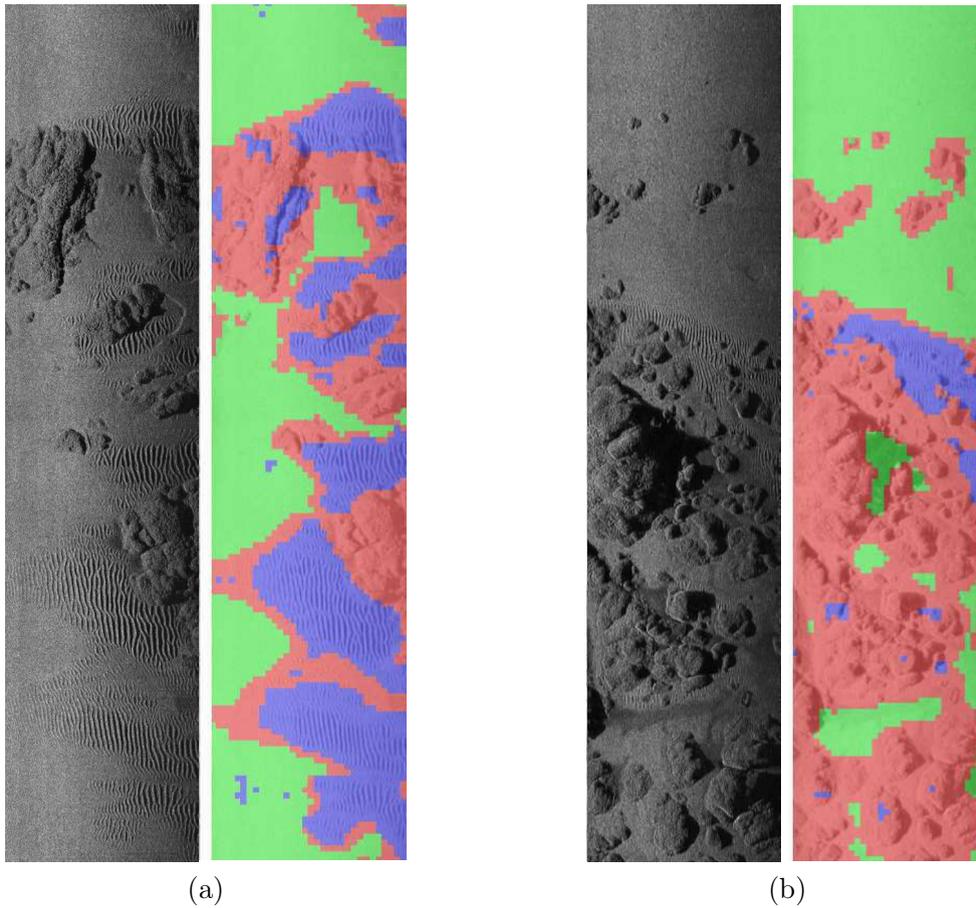


FIGURE 4.36: Deux exemples d'image sonar et leur classification respective en 3 classes (a) et (b). Les couleurs correspondant aux différentes classes sont les suivantes, vert = homogène, bleu = anisotrope, et rouge = complexe.

## 4.7 Résumé du Chapitre

Dans les chapitres 2 et 3 précédents, nous avons pu constater que l'information portée par une image sonar peut être de différentes natures, à savoir énergétique, structurelle, géométrique et multi-échelle. Nous nous sommes alors attachés dans ce chapitre à proposer un outil de traitement d'image qui permette d'extraire de telles informations : le signal monogène. L'avantage de la représentation particulière d'une image que le signal monogène fournit est qu'elle limite le nombre de paramètres à fixer grâce à son caractère isotrope. En effet, jusqu'à présent, nous avons vu que les seuls et uniques paramètres qui interviennent dans l'étude correspondent aux paramètres liés à l'analyse multi-échelle. Contrairement à d'autres approches que nous pouvons retrouver dans la littérature telles que les matrices de co-occurrences [133], les bancs de filtres de Gabor ou encore les ondelettes [134, 135, 136, 137, 138, 139, 140] nous n'avons pas besoin de fixer par exemple un certain nombre d'orientations pour paver de manière plus ou moins complète le plan spatial ou fréquentiel d'une image. Ainsi cela permet de rendre l'approche plus générale, comme par exemple une adaptation aisée à différentes bases données.

Cependant, il est maintenant nécessaire d'utiliser cette nouvelle représentation de manière à en extraire l'information qui nous intéresse. Les chapitres qui suivent proposent donc de se baser sur cette représentation monogène multi-échelle pour décrire et caractériser une image sonar dans un contexte ATR.

# Chapitre 5

## La dimension intrinsèque des images

### Sommaire

---

<b>5.1</b>	<b>Introduction</b>	<b>119</b>
<b>5.2</b>	<b>La dimension intrinsèque discrète</b>	<b>120</b>
5.2.1	Introduction	120
5.2.2	La dimension intrinsèque en traitement d'image	120
5.2.3	Estimation de la dimension intrinsèque discrète	122
5.2.4	Potentialités pour l'étude de la complexité des images sonar	122
<b>5.3</b>	<b>La dimension intrinsèque continue</b>	<b>124</b>
5.3.1	Topologie triangulaire	124
5.3.1.1	Considérations spectrales	124
5.3.1.2	Construction d'un triangle	125
5.3.1.3	Estimation en espace triangulaire	126
5.3.1.4	Topologie triangulaire et signal monogène	127
5.3.2	Topologie conique	129
5.3.2.1	Position du problème	129
5.3.2.2	Construction d'un cône	130
5.3.2.3	Estimation en espace conique	131
5.3.2.4	Liens avec le signal monogène	131
5.3.2.5	Exemples sur des données synthétiques	133
5.3.2.6	Exemple sur une image réelle	135
5.3.3	Dimensions intrinsèques continues multi-échelles	136
5.3.3.1	Définition	136
5.3.3.2	Illustration des dimensions intrinsèques multi-échelles	136
<b>5.4</b>	<b>Application à l'imagerie sonar</b>	<b>138</b>
5.4.1	Paramétrage spécifique aux images sonar	138
5.4.2	Analyse des fonds marins homogènes	139

5.4.3	Analyse des fonds marins anisotropes . . . . .	141
5.4.4	Analyse des fonds marins complexes . . . . .	143
<b>5.5</b>	<b>Résumé du chapitre . . . . .</b>	<b>144</b>

---

## 5.1 Introduction

Au chapitre précédent, nous avons vu que la représentation monogène multi-échelle des images sonar nous fournit un ensemble d'informations énergétiques, structurales et géométriques sur les images à différents niveaux de détail. La richesse de ces informations nous laisse à penser qu'il est envisageable d'en extraire des descripteurs environnementaux significatifs sur les fonds marins. Pour ce faire, nous proposons de suivre les travaux initiés au sein du CMRE, de l'Université d'Heriot-Watt ou encore du FFI norvégien qui définissent des coefficients d'anisotropie et de complexité estimés par le biais de différentes approches (bancs de filtres de Gabor, de Haar, analyse en ondelettes, etc.) comme nous l'avons vu au paragraphe 3.3. A travers ce chapitre, nous proposons de redéfinir ces descripteurs en exploitant le concept de dimension intrinsèque. Celui-ci est directement relié aux degrés de liberté d'un signal 2D. Nous verrons que plus la dimension intrinsèque sera élevée, plus le signal possèdera des degrés de liberté, et par conséquent, plus il sera complexe. Le principal avantage de cet outil d'analyse est que son estimation peut être directement réalisée via les informations monogènes. Ainsi, puisque le signal monogène possède une extension multi-échelle, nous pourrions également étendre la notion de dimension intrinsèque à notre cadre multi-échelle. Cette nouvelle phase dans le traitement de l'image sonar s'inscrit donc après le calcul du signal monogène dans notre chaîne de description de l'environnement marin comme l'illustre la Figure 5.1.

Ce chapitre est organisé de la manière suivante. Après une présentation des origines du concept et de sa définition discrète dans la section 5.2, nous aborderons le cas plus général de la dimension intrinsèque continue dans la section 5.3. En particulier, nous y verrons l'extension multi-échelle de ce concept que nous proposons à travers ces travaux. Finalement, la section 5.4 sera quant à elle consacrée à l'illustration de ce concept sur des images sonar réelles.

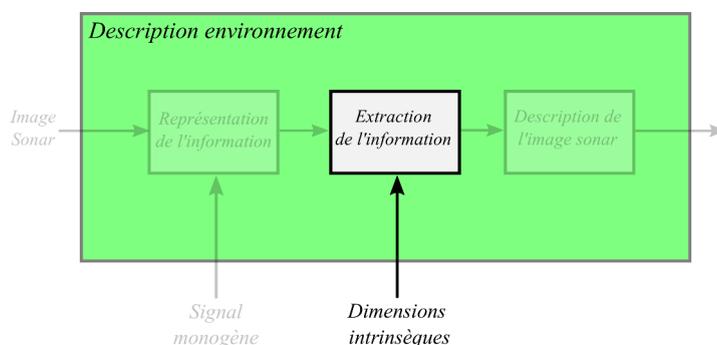


FIGURE 5.1: Position de l'extraction d'informations dans le processus.

## 5.2 La dimension intrinsèque discrète

### 5.2.1 Introduction

Pour trouver les origines du concept de dimension intrinsèque, il est nécessaire de quitter le domaine du traitement d'image et de s'intéresser tout d'abord au domaine des statistiques. Dans ce dernier, il est courant de traiter des ensembles de données de très grandes dimensions. De ce fait, pour réduire les espaces de stockage, ou encore le temps de calcul des algorithmes, il est pertinent de se poser la question de la dimension réelle qu'occupe ces données.

En effet, si les données sont représentées sous la forme de vecteurs de dimension  $D$ , il est généralement admis que ces données appartiennent à un espace  $\mathbb{R}^D$ . Cependant, cela ne signifie pas pour autant que la dimension réelle des données est de  $D$ . La dimension de l'ensemble des données est en fait le nombre minimal de variables qui permettent de représenter les données sans perte d'information. C'est dans cette optique que K. Fukunaga propose la première définition de notion de dimension intrinsèque dans [141] :

**Definition 5.1.** Un ensemble de données  $\Omega \in \mathbb{R}^D$  possède une dimension intrinsèque (*iD - intrinsic dimensionality*) de  $M$ , si ses éléments sont contenus intégralement dans un  $M$ -sous-espace de  $\mathbb{R}^D$  (*i.e.* un sous-espace de dimension  $M < D$ ).

De nombreuses études se sont ensuite penchées sur la manière d'estimer une telle dimension, comme par exemple [142, 143, 144, 145]. Les études les plus récentes, mentionnant cette dimension intrinsèque, proviennent des méthodes "d'apprentissage de variétés" (de l'anglais *manifold learning*), qui visent à analyser les données de très haute dimension et à les transformer en un espace plus petit facilitant le traitement [142]. En d'autres termes, l'étude de la dimension intrinsèque peut être reliée aux méthodes connues sous le nom d'ACP (Analyse en Composantes Principales).

La communauté du traitement d'image s'est alors intéressée à ce concept en considérant que les structures contenues dans une image sont des variétés mathématiques d'un espace 2D.

### 5.2.2 La dimension intrinsèque en traitement d'image

Partant du constat que les images sont principalement constituées de zones homogènes séparées par des bords ou des coins, C. Zetsche [146] propose d'étudier le concept de dimension intrinsèque pour séparer ces différentes formes. Puisque l'image est un espace de dimension 2, il considère qu'il n'y a que 3 sous-espaces contenus dans celui-ci, chacun représentés par trois valeurs distinctes de dimensions intrinsèques. L'auteur considère donc que les structures contenues dans une image peuvent être caractérisées selon trois valeurs de dimension intrinsèque :

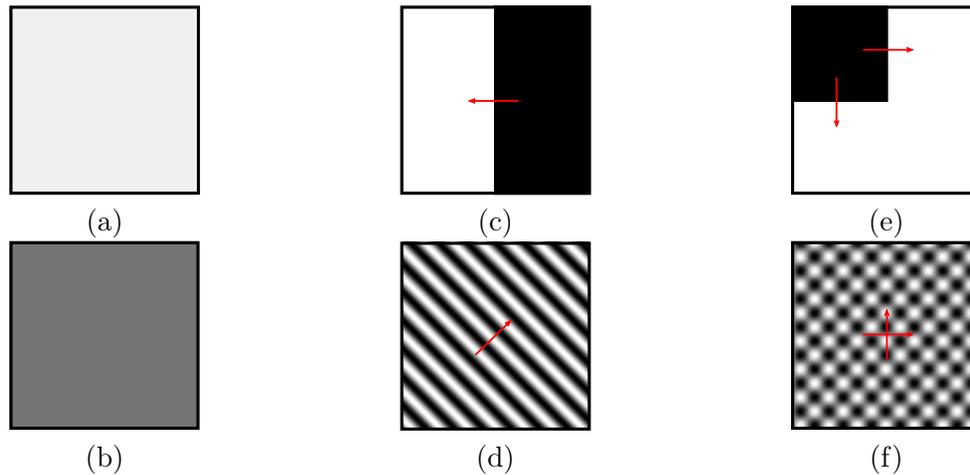


FIGURE 5.2: Exemple de 6 images aux dimensions intrinsèques différentes : (a)-(b) images de signaux homogènes i0D, (c)-(d) images de signaux unidirectionnel i1D, et (e)-(f) images de signaux ni constant ni unidirectionnel i2D.

- Une dimension intrinsèque de 0 (i0D) représente un signal 2D constant  $u(x, y) = cst$ . Il ne dépend alors d'aucune coordonnée spatiale, et n'a donc aucun degré de liberté.
- Une dimension intrinsèque de 1 (i1D) représente un signal 2D qui ne varie que dans une direction donnée. Un tel signal 2D peut s'exprimer uniquement par l'intermédiaire d'une seule variable  $z$  telle que :  $u(x, y) = u(z)$ , où  $z = a \cos(\theta)x + b \sin(\theta)y$ . Il n'a alors qu'un seul degré de liberté.
- Une dimension intrinsèque de 2 (i2D) représente un signal 2D qui sera quant à lui ni i0D, ni i1D. Un tel signal possède alors deux degrés de liberté, soit le nombre maximum en 2D.

La Figure 5.2 propose d'illustrer des signaux 2D typiques qui correspondent aux 3 dimensions intrinsèques précédentes. Les signaux constants et homogènes comme représentés en (a) et (b) correspondent au cas i0D. Le cas i1D correspond aux signaux qui ne varient que dans une direction donnée, comme un bord (c) ou une sinusoïde 2D (d). Sur ces exemples, la direction de variation est représentée par une flèche rouge. Les exemples (e) et (f) représentent quant à eux des signaux plus complexes, qui varient dans plusieurs directions, comme un coin ou une superposition de deux sinusoïdes.

Naturellement, la dimension intrinsèque de l'image entière n'ayant qu'un intérêt limité, C. Zetsche propose de l'analyser par bloc de manière à choisir pour chacun d'entre eux la dimension intrinsèque du signal 2D contenu dans ce dernier. Une telle analyse des images revient donc à réaliser une sorte classification de l'image en trois valeurs distinctes 0, 1 et 2 qui représentent la dimension intrinsèque des différents blocs de celle-ci.

### 5.2.3 Estimation de la dimension intrinsèque discrète

En traitement d'image, l'estimation de la dimension intrinsèque est généralement proposée par une analyse spectrale des blocs de l'image [146, 147]. En effet, les trois cas possibles de la dimension intrinsèque peuvent se distinguer via le contenu spectral des blocs :

- Cas i0D : le spectre d'un bloc homogène sera concentré autour de l'origine (*i.e.* un dirac).
- Cas i1D : le spectre d'un bloc contenant uniquement un signal ne variant que dans une direction donnée, sera concentré autour d'une ligne.
- Cas i2D : le spectre d'un bloc qui n'est ni homogène, ni orienté de façon homogène, sera donc ni concentré autour de l'origine, ni autour d'une ligne. Sa représentation fréquentielle sera répartie dans l'espace fréquentiel entier.

Nous pourrions constater ces allures particulières du contenu fréquentiel dans le paragraphe suivant dans le cas d'imagettes sonar illustrant ces trois cas de dimensions intrinsèques. Dans la littérature, l'évaluation de la dimension intrinsèque à partir de ces observations fréquentielles peuvent être retrouvées dans les travaux de G. Krieger [147, 148, 149] par l'utilisation d'opérateurs non-linéaires complexes, tels que les opérateurs de Volterra.

### 5.2.4 Potentialités pour l'étude de la complexité des images sonar

D'après les paragraphes précédents, la dimension intrinsèque serait un moyen de distinguer les zones homogènes, les zones orientées ainsi que les zones plus complexes d'une image. Ainsi, on peut supposer que ce concept permette de distinguer les textures du fond marin présentes sur les images sonar, et particulier celles qui nous intéressent dans ces travaux. En effet, nous avons vu qu'en ATR, deux types de fonds marins sont généralement redoutés, à savoir les fonds anisotropes et complexes. Sur la Figure 5.3, on constate que le premier cité possède une forme particulière du fait de son orientation privilégiée (b). L'observation de son spectre (e) montre que cette orientation privilégiée se retrouve également dans sa transformée de Fourier. Cette allure particulière du spectre correspond à la caractérisation fréquentielle que nous avons précédemment donnée d'un bloc i1D. Dans le cas d'une zone de fond que l'on peut considérer comme complexe (présentée à la Figure 5.3.(c)), on remarque que le contenu spectral illustré sur la Figure 5.3.(f) n'est ni concentré autour d'une ligne ni autour de l'origine, ce qui correspond à la définition d'un bloc i2D. Concernant le cas d'un fond homogène (cf. Figure 5.3.(a) et (d)), on remarque que son spectre est quant à lui concentré autour de l'origine du plan fréquentiel ce qui est caractéristique d'un bloc i0D.

A travers ces quelques exemples, on peut supposer que l'étude de la dimension intrinsèque semble pertinente pour pouvoir décrire la texture des fonds marins dans le cadre de l'ATR. En effet, elle permettrait de quantifier la complexité du contenu d'un bloc de l'image à l'aide de considérations géométriques. La dimension intrinsèque proposerait donc une alternative à la

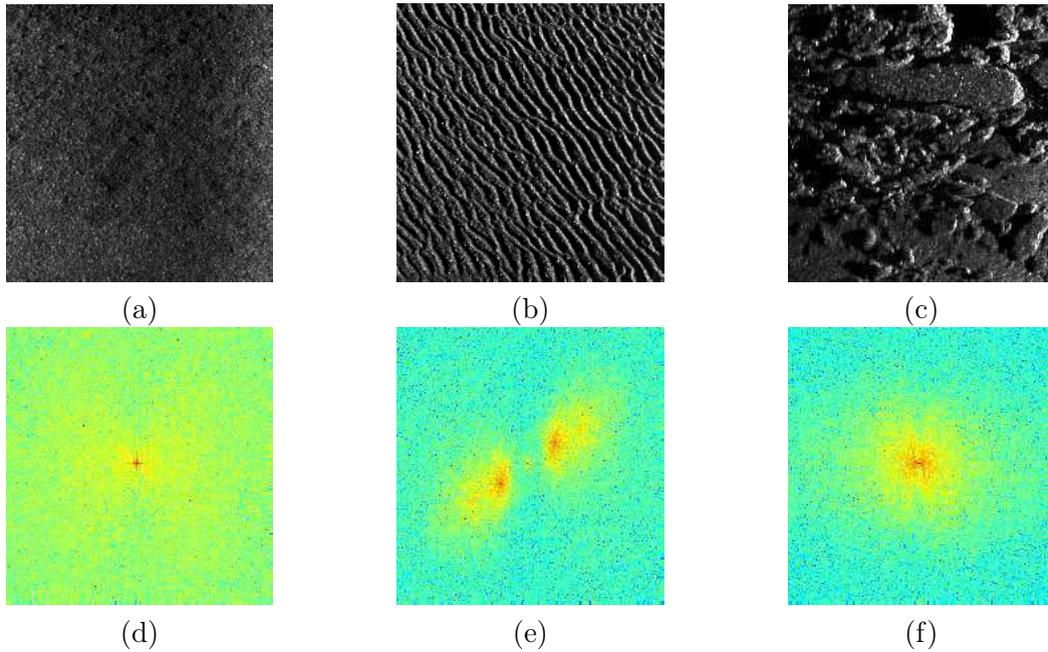


FIGURE 5.3: Trois exemples de fonds marins : homogène (a), anisotrope (b), complexe (c), et leurs transformées de Fourier (TF) respectives (d), (e), et (f). L'origine du plan fréquentiel est considéré au centre de l'image.

définition des coefficients d'anisotropie et de complexité que l'on retrouve dans la littérature, avec l'avantage d'appartenir à un cadre mathématique formel.

Cependant, l'étude des images réelles montre qu'il n'y a que très peu de signaux purement  $i0D$ ,  $i1D$  ou  $i2D$ . Les exemples précédents suffisent à s'en convaincre puisque la notion de bruit peut rendre confuse l'analyse d'un bloc et de son spectre. En effet, d'après les définitions précédentes, le bruit n'appartient ni à la classe  $i0D$  (un bruit blanc affecte par exemple toutes les fréquences et ne possède donc pas un spectre concentré parfaitement autour de l'origine) ni à la classe  $i1D$  (aucune orientation n'est privilégiée par un tel bruit). Le bruit est donc un phénomène  $i2D$  qui affecte l'ensemble des images. Cela s'observe sur la Figure 5.3 qui montre des spectres jamais parfaitement concentrés ni autour de l'origine ni autour d'une ligne. La question est donc de séparer le bruit des structures véritablement  $i2D$  d'une image. Bien qu'il existe des moyens d'estimation du niveau de bruit d'une image [150], la frontière entre le bruit et les structures  $i2D$  reste floue.

Selon N. Krüger [151], pour permettre une telle discrimination, la définition de la dimension intrinsèque doit évoluer et ne plus se restreinte à un choix discret, mais doit plutôt être vue comme un continuum de valeurs.

### 5.3 La dimension intrinsèque continue

Partant du constat que les images réelles sont davantage une combinaison de structures de différentes dimensions intrinsèques que des cas idéaux, N. Krüger propose de redéfinir cette dimension intrinsèque dans un cadre continu.

Pour ce faire, N. Krüger introduit une nouvelle représentation de la dimension intrinsèque par l'intermédiaire d'une topologie triangulaire et d'associer à chaque bloc de l'image trois indices de confiance. Ces indices exprimeront le degré d'appartenance du signal contenu dans le bloc vis-à-vis des cas i0D, i1D et i2D idéaux.

#### 5.3.1 Topologie triangulaire

Nous avons vu précédemment que les différentes dimensions intrinsèques se distinguent par les formes spécifiques du spectre d'un bloc de l'image. Ainsi, l'auteur propose de caractériser la dispersion de ce spectre selon l'origine et selon une droite dans le plan fréquentiel.

##### 5.3.1.1 Considérations spectrales

Toujours par le biais d'une analyse par blocs de l'image, N. Krüger utilise la notion de variance afin de caractériser la dispersion du spectre du signal.

Dans cette section, nous considérons le spectre  $F$  d'un signal 2D  $f$  dans le plan fréquentiel  $\Omega$ . Les coordonnées fréquentielles seront exprimées par un vecteur  $\mathbf{u} = (u_1, u_2)$ .

Dans un système de coordonnées polaires, ces coordonnées fréquentielles sont exprimées par un vecteur  $\mathbf{u} \rightarrow (q, \theta)$  qui représente respectivement une distance et un angle. La variance du spectre  $F$  par rapport à l'origine (*i.e.* une variance radiale) peut alors être estimée par la relation suivante :

$$\sigma_O^2 = \frac{1}{K} \int_0^Q q^2 \int_0^{2\pi} |F(q \cos \theta, q \sin \theta)|^2 d\theta dq, \quad (5.1)$$

où  $K$  est un facteur de normalisation tel que :

$$K = \int_0^Q \int_0^{2\pi} |F(q \cos(\theta), q \sin(\theta))|^2 d\theta dq,$$

et  $Q$  représente la dimension radiale de  $\Omega$ .

De même, la dispersion selon une droite du spectre est donnée par la variance suivante (*i.e.* une variance angulaire) :

$$\sigma_L^2 = \min_{\theta_0} \left( \frac{1}{N} \int_{\theta_0 - \pi}^{\theta_0 + \pi} (\theta - \theta_0)^2 \int_0^Q |F(q \cos(\theta_0 - \pi), q \sin(\theta_0 - \pi))|^2 d\theta dq \right) \quad (5.2)$$

Dimension intrinsèque	i0D	i1D	i2D
$\sigma_O^2$	0	Grande	Grande
$\sigma_L^2$	inconnue	0	Grande

TABLE 5.1: Ordre de grandeur des variances  $\sigma_O^2$  et  $\sigma_L^2$  en fonction des dimensions intrinsèques idéales.

qui représente la valeur minimale de la variance du spectre selon une droite d'orientation  $\theta_0$ . Par la même occasion, cette variance minimale définit l'orientation du signal 2D.

### 5.3.1.2 Construction d'un triangle

Pour permettre de comprendre les variations de ces quantités selon les différentes dimensions intrinsèques, la Table 5.1 propose un ordre de grandeur de ces variances pour les trois cas idéaux de dimension intrinsèque. Une analyse rapide de ce tableau permet de noter que la variance selon l'origine permet de séparer le cas i0D des autres cas i1D et i2D. Ainsi, si cette variance est élevée, elle indique que le signal n'est pas homogène et qu'il y a une variation des niveaux d'intensité des pixels dans le bloc analysé. On remarque également que la variance  $\sigma_L^2$  est indéfinie dans le cas d'un bloc totalement homogène. Cela provient du fait que le spectre est concentré autour d'un dirac au centre du plan fréquentiel. L'estimation d'une variance selon une droite est alors imprécise. Cependant, cette dernière permet de bien séparer les cas i1D et i2D.

Il est intéressant de remarquer que ces différentes variances possèdent des dynamiques très importantes et non-bornées. Il est donc difficile de les interpréter telles quelles pour caractériser les blocs d'une image. C'est pourquoi N. Krüger propose d'introduire dans les équations (5.1) et (5.2) une fonction de normalisation. Cette dernière permet de confiner la dynamique de ces variances entre 0 et 1. Pour définir une telle fonction, M. Felsberg propose d'utiliser un seuillage doux tel que pour toute valeur  $A \in [0, +\infty[$  on ait :

$$\tilde{A} = N(A) \in [0, 1]. \quad (5.3)$$

Le lecteur trouvera de plus amples détails sur cette fonction de normalisation  $N$  dans [152].

Une telle opération permet d'obtenir deux variances  $\tilde{\sigma}_L^2$  et  $\tilde{\sigma}_O^2$ , que l'on peut représenter sur deux axes orthogonaux comme le montre la Figure 5.4.a. La topologie de la dimension intrinsèque devient alors un carré de dimensions un sur un. Chaque coin du carré définit alors une dimension intrinsèque idéale, à l'exception de la dimension intrinsèque de 0 dont la position est toujours indéfinie sur l'axe  $\tilde{\sigma}_L^2$ .

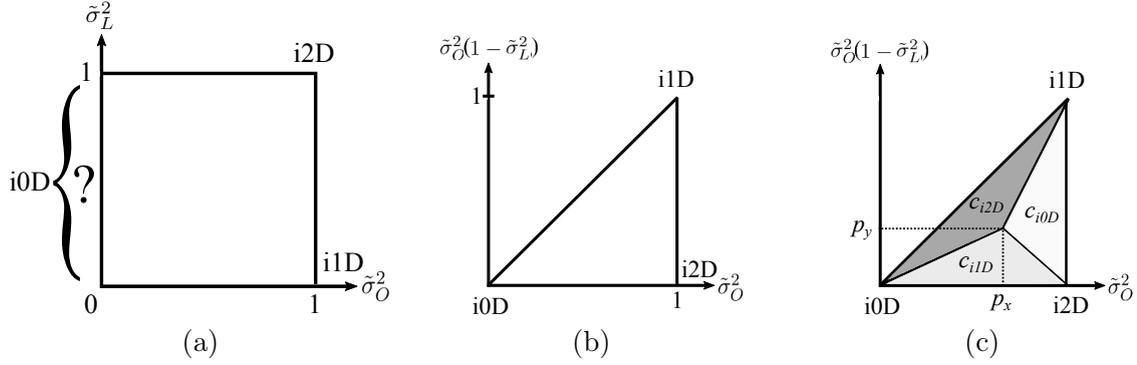


FIGURE 5.4: Topologie de la dimension intrinsèque. (a) Topologie carrée initiale, (b) topologie triangulaire et (c) paramétrisation de la topologie triangulaire.

Pour permettre de fixer cette position, une modification de cet axe est alors proposée. Elle consiste à y représenter la quantité  $\tilde{\sigma}_O^2(1-\tilde{\sigma}_L^2)$ , ce qui a pour effet de passer d'une topologie carrée à une topologie triangulaire (cf. Figure 5.4.b). Il est à noter que les coins  $i1D$  et  $i2D$  sont alors inversés, et qu'une moitié du carré initial devient une zone "interdite". Cette dernière correspond au fait qu'un bloc de l'image ne peut pas avoir simultanément son spectre concentré autour de l'origine et autour d'une ligne.

Maintenant que l'on a vu que la dimension intrinsèque évolue dans un espace triangulaire, il est nécessaire de passer à la manière d'estimer cette dernière.

### 5.3.1.3 Estimation en espace triangulaire

Pour estimer la dimension intrinsèque d'un bloc donné de l'image, il suffit de voir le couple  $(\tilde{\sigma}_O^2, \tilde{\sigma}_L^2)$  extrait de son spectre comme un point de coordonnées  $(p_x, p_y)$  dans le triangle. La position de ce point vis-à-vis des coins du triangle, représentent alors les similarités du contenu du bloc par rapport aux cas idéaux  $i0D$ ,  $i1D$  et  $i2D$ .

Ces similarités peuvent alors être représentées sous la forme de degrés de confiance en utilisant les coordonnées barycentriques de ce point dans le triangle, ou de manière équivalente, aux aires relatives des sous-triangles opposés à chaque coin comme le montre la Figure 5.4. Si on note  $A_j$  l'aire du triangle opposé au coin de dimension intrinsèque  $j$  idéale, et  $T$  l'aire du triangle entier ( $T = 1/2$ ), on obtient 3 degrés de confiance :

$$c_{i0D} = \frac{A_0}{T} = 1 - p_x \quad (5.4)$$

$$c_{i1D} = \frac{A_1}{T} = p_y \quad (5.5)$$

$$c_{i2D} = \frac{A_2}{T} = 1 - p_y. \quad (5.6)$$

Par exemple, on remarque que si le point représentatif  $(p_x, p_y)$  se rapproche du coin  $i0D$ , (*i.e.*  $p_x$  se rapproche de zéro), l'aire  $A_0$  va augmenter, et par conséquent  $c_{i0D}$  sera proche de un.

Ainsi, plus le point représentatif s'éloigne horizontalement du coin inférieur gauche, moins son degré de confiance  $c_{i0D}$  sera élevé. De même, plus ce point s'éloigne de l'axe horizontal, plus le signal possèdera un indice  $c_{i1D}$  important signifiant son caractère directif, jusqu'à devenir purement i1D dans le coin supérieur droit. Dans le cas où le point se situe proche du coin inférieur droit, le bloc analysé possède alors un contenu complexe et par conséquent un fort indice de confiance  $c_{i2D}$ .

Une telle définition des indices de similarité permet de vérifier l'équation :

$$c_{i0D} + c_{i1D} + c_{i2D} = 1. \quad (5.7)$$

Ainsi, la représentation continue de la dimension intrinsèque revient à estimer trois coefficients complémentaires qui peuvent chacun être vu comme un degré de confiance sur les trois caractéristiques intrinsèques possibles du signal 2D contenu dans le bloc analysé.

#### 5.3.1.4 Topologie triangulaire et signal monogène

Nous venons de voir que la dimension intrinsèque continue peut être estimée à partir de considérations spectrales. Cependant, N. Krüger [151] mentionne que de nombreuses études se sont également penchées sur la manière d'estimer efficacement les variances de ces spectres (selon l'origine et selon une ligne) par le biais d'opérateurs spatiaux tels que les tenseurs de structure [153, 154], les extensions polynomiales [155] ou encore les harmoniques sphériques [108]. L'auteur remarque que le signal monogène semble également fournir les informations adaptées à l'estimation de ces dimensions intrinsèques continues. En effet, on rappelle que cet outil permet d'obtenir une décomposition orthogonale du signal en trois informations locales à savoir une amplitude, une orientation et une phase. N. Krüger montre que l'amplitude locale  $A$  est un moyen d'estimer la variance du spectre selon l'origine. On obtient alors la relation suivante :

$$\tilde{\sigma}_O^2 = N(A(\mathbf{x})), \quad (5.8)$$

qui n'est donc que l'amplitude locale issue du signal monogène, normalisée par une fonction  $N$  qui est la même que celle mentionnée dans la section 5.3.1.2 (*i.e.* un seuillage doux). On remarque que le caractère local de l'amplitude monogène permet de définir cette variance en chaque pixel  $\mathbf{x}$  de l'image.

Pour définir une telle variance du spectre, l'utilisation de cette information monogène peut trouver une explication dans les observations réalisées dans le chapitre précédent, au paragraphe 4.5.4.2. En effet, nous avons vu que l'amplitude monogène est une information énergétique qui permet de renseigner sur la présence d'un signal ou non, dans une bande de fréquence spécifique. Ainsi, si l'amplitude monogène est forte dans une bande de fréquence

donnée, une partie du spectre du signal original est contenu dans celle-ci, ce qui montre si sa concentration est autour de l'origine ou non.

Le signal monogène fournit également une information d'orientation locale  $\theta$  qui semble être adéquate pour définir la notion de directivité du signal dans un bloc donné. Ainsi, N. Krüger propose de l'utiliser pour estimer la notion de variance selon une droite du spectre. Pour cela, l'auteur considère une région  $R$  (par exemple une fenêtre rectangulaire) autour d'un pixel  $\mathbf{x}$  donné de l'image, et y calcule une orientation moyenne. Il définit la variance  $\tilde{\sigma}_L^2$  comme la somme pondérée des distances entre les orientations locales et leur moyenne dans cette région. La pondération est constituée par l'amplitude locale normalisée selon l'expression suivante :

$$\tilde{\sigma}_L^2 = \sum_{\mathbf{x} \in R} N(A(\mathbf{x})) d(\theta(\mathbf{x}), E_\theta^R[\theta]), \quad (5.9)$$

où  $E_\theta^R[\theta]$  est l'espérance de  $\theta$  calculée sur le voisinage  $R$  et la fonction  $d$  correspond au calcul de la distance entre l'orientation locale et sa moyenne dans le voisinage  $R$ . Dans la suite de ce manuscrit, cette fonction ne sera pas utilisée et nous proposons donc au lecteur intéressé par de plus amples détails de se référer à l'étude de N. Krüger [156].

Le signal monogène permet ainsi d'estimer les variances du spectre par l'intermédiaire des informations amplitude et d'orientation locales qu'il fournit. L'avantage de son utilisation réside justement dans ce caractère local, qui procure à la dimension intrinsèque une estimation possible en chaque pixel de l'image, et non plus sur un découpage par blocs de l'image.

Ainsi, alors que le découpage par blocs nécessitait un paramétrage en termes de taux de recouvrement entre les différents blocs et affecte à chaque pixel d'un bloc une même valeur de dimension intrinsèque, l'estimation de la dimension intrinsèque continue revient ici à l'utilisation d'une fenêtre glissante. La taille de cette fenêtre définit le voisinage et influe uniquement sur l'estimation de  $\tilde{\sigma}_L^2$ , c'est-à-dire sur l'estimation d'une orientation privilégiée dans le voisinage du pixel courant. D'après l'équation (5.8), la variance  $\tilde{\sigma}_O^2$  est quant à elle toujours calculée à partir de l'information d'amplitude locale fournie par le signal monogène (*i.e.* pixel-à-pixel) et dépend donc seulement du voisinage induit par l'étendue des noyaux de Riesz.

Dans cette section, seule l'estimation des variances a été modifiée. Ainsi les expressions des coefficients de confiance sur les trois cas idéaux de dimensions intrinsèques sont inchangées par rapport à la section précédente. On obtient toujours pour chaque pixel de l'image, trois coefficients complémentaires qui représentent la confiance qu'un pixel appartienne à une zone i0D, i1D ou i2D. La complémentarité de ces informations n'est pas sans rappeler le cadre des probabilités, ce que M. Felsberg mentionne dans [152]. Cependant, cette topologie triangulaire ne satisfait pas l'auteur, qui propose quant à lui une représentation de la dimension intrinsèque continue sous la forme de cône.

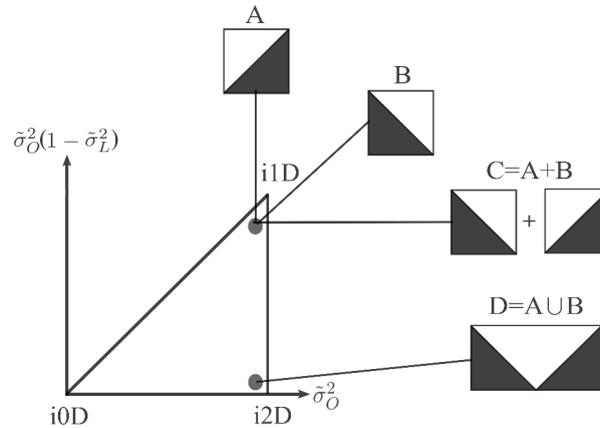


FIGURE 5.5: Positions dans le triangle de différents blocs  $i1D$ , A et B, orientés selon des directions opposées. La position de leur dimension intrinsèque moyenne C ainsi que celle de l'union des deux blocs D sont également présentées.

### 5.3.2 Topologie conique

M. Felsberg remarque que la topologie triangulaire ne permet pas d'analyser toute la variété des structures possibles dans une image, et en particulier les structures directives  $i1D$ .

Notons dès à présent que par souci de clarté, nous revenons temporairement au cas général de l'analyse par blocs d'une image. Cela permet de localiser plus facilement les problèmes liés à la topologie triangulaire.

#### 5.3.2.1 Position du problème

Dans son étude, M. Felsberg remarque que le problème principal de la topologie triangulaire provient du fait que son expression n'est pas consistante. Si l'on considère deux blocs voisins d'une image (*i.e.* adjacents), la consistance signifierait que la moyenne de leurs dimensions intrinsèques soit la même que la dimension intrinsèque des deux blocs joints. Or, si l'on imagine deux blocs adjacents  $i1D$ , mais chacun d'orientation distincte, la moyenne des deux dimensions intrinsèques serait  $i1D$ , ce qui est faux. En effet, la jonction des blocs ferait alors apparaître un signal  $2D$  ayant deux orientations distinctes, et serait théoriquement  $i2D$ .

Ce phénomène est illustré à la Figure 5.5, où deux blocs  $i1D$  d'orientations opposées sont présentés en A et en B. Ces deux blocs étant identiques, à l'exception de leur orientation propre, ils possèdent le même point représentatif dans le triangle. La position moyenne de ces deux points est donc toujours le même point noté C. Cependant, on remarque que l'union de ces deux blocs fait apparaître un coin D qui est une structure  $i2D$  idéale.

Selon l'auteur, ce problème provient du fait que les points représentatifs de tous les blocs  $i1D$  sont projetés dans la même région du triangle (dans le coin supérieur droit), et qu'aucune distinction n'est alors faite entre leurs orientations respectives. Ainsi, il propose de faire intervenir l'orientation des structures contenues dans l'image dans la représentation.

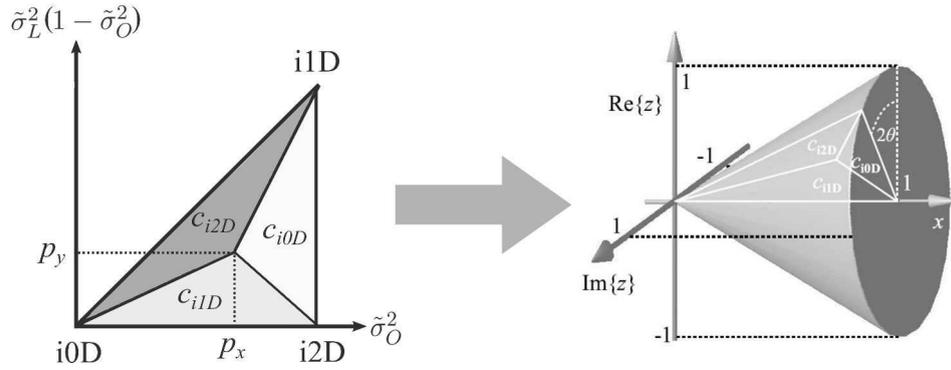


FIGURE 5.6: Topologie conique de la dimension intrinsèque.

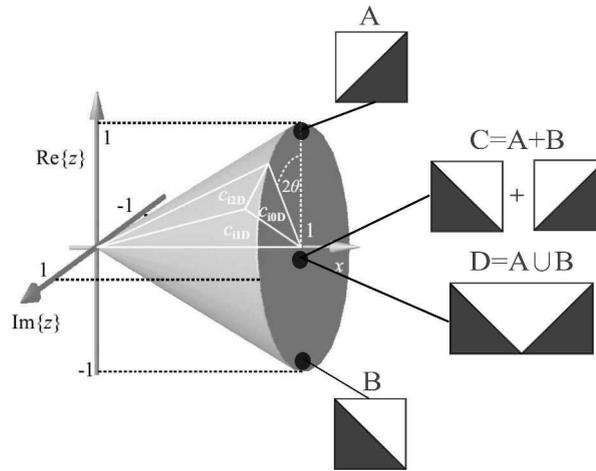


FIGURE 5.7: Positions dans le cône de différents blocs i1D, A et B, orientés selon des directions opposées. La position de leur dimension intrinsèque moyenne C ainsi que celle de l'union des deux blocs D sont également présentées.

### 5.3.2.2 Construction d'un cône

A partir de la topologie triangulaire, M. Felsberg introduit une topologie conique en réalisant la rotation de ce triangle autour de l'axe horizontal par l'intermédiaire d'une représentation en angle double. Pour ce faire, il suffit de remplacer la coordonnée  $p_y$  précédente par la valeur complexe suivante :

$$z = p_y \exp(j2\theta), \quad (5.10)$$

où  $\theta$  représente l'orientation de la structure contenue dans le bloc analysé. Une telle représentation permet aux structures i1D d'être réparties dans le cône en fonction de leurs orientations. La Figure 5.6 propose d'illustrer cette topologie.

L'avantage de cette représentation 3D réside dans sa capacité à répartir dans un espace 3D les structures i1D en fonction de leurs orientations locales. Ainsi, cette représentation devient consistante. Les blocs i1D de même orientation auront des points représentatifs proches, alors qu'ils seront projetés chacun à l'opposé du cône dans le cas de deux blocs d'orientations opposées, comme le montre la Figure 5.7 avec les blocs A et B, leur moyenne et leur union.

Dans le cas de ce modèle topologique, on remarque que le volume occupé par les signaux constants (*i.e.* proches du coin  $i0D$ ) est relativement faible puisqu'on se situe dans la pointe du cône. Cela provient du fait que la variété des signaux constants est très faible par rapport à celle des signaux non-constants, qui occupent eux un volume beaucoup plus important dans le cône.

### 5.3.2.3 Estimation en espace conique

Dans ce nouvel espace 3D, le point représentatif du bloc analysé sera alors placé suivant ses coordonnées  $(p_x, Re\{z\}, Im\{z\})$ . Les indices de confiance  $c_{i0D}$ ,  $c_{i1D}$  et  $c_{i2D}$  peuvent à nouveau être estimés. Pour cela, il suffit de tracer le plan d'orientation  $2\theta$  passant par ce point représentatif et de remarquer que son intersection avec le cône forme à nouveau un triangle, illustré en blanc sur la Figure 5.6. Dans ce cas, les formules exprimées dans le paragraphe précédent restent valables dans le cas de la topologie conique.

Dans la suite de ce manuscrit, nous utiliserons la dimension intrinsèque via la topologie conique introduite dans cette section. Dans le paragraphe suivant, nous allons expliciter la manière dont nous implémentons l'estimation de la dimension intrinsèque continue dans cette topologie, et en particulier l'utilisation des informations monogènes pour y parvenir.

### 5.3.2.4 Liens avec le signal monogène

D'après M. Felsberg, la topologie conique possède certains avantages liés à son modèle convexe. En particulier, il mentionne le fait qu'il n'est plus nécessaire d'étudier l'image par blocs mais directement au niveau du pixel. Pour ce faire, il est nécessaire d'avoir un outil fournissant, en chaque pixel de l'image, une information d'amplitude et d'orientation. Ces informations peuvent être fournies par différents outils comme par exemple le gradient. Cependant, dans le chapitre 4, nous avons vu que le signal monogène permet également l'extraction de telles informations locales. C'est donc ce dernier que nous utiliserons dans la suite pour permettre l'estimation de la dimension intrinsèque en espace conique. Nous allons maintenant voir que chaque pixel de l'image possède un point représentatif dans le cône et comment on peut définir la dimension intrinsèque continue à partir de son voisinage.

Si l'on suit le raisonnement de M. Felsberg, dans le cas de la topologie conique, l'amplitude locale (normalisée entre 0 et 1) d'un pixel  $\mathbf{x}$  fournie par le signal monogène joue toujours le rôle de  $\sigma_{\mathcal{O}}^2 = p_x$  comme nous l'avons vu lors de l'étude de la topologie triangulaire à l'équation (5.8). C'est-à-dire que cette amplitude locale définit la position du point représentatif du pixel  $\mathbf{x}$  dans le cône par rapport au coin  $i0D$ . Si on se réfère à la Figure 5.6, la position de ce point représentatif est donc quelque part sur un disque (intersection du cône avec un plan  $p_x = cst$ ). Théoriquement, puisqu'un pixel est associé à une seule orientation locale  $\theta(\mathbf{x})$ , ce

point représentatif se situe donc forcément sur la surface du cône (chaque pixel est considéré comme un élément i1D pris un par un, *i.e.* on suppose implicitement que  $p_y = p_x$  pour un pixel). Ainsi, la position finale de ce point représentatif est fixée par son orientation en angle double  $2\theta$  que nous avons vu à l'équation (5.10). Maintenant, si l'on souhaite savoir si ce pixel appartient à une zone de l'image plutôt *i0D*, *i1D* ou *i2D*, il est nécessaire de tenir compte de son voisinage (comme nous l'avons vu dans l'équation (5.9) dans le cas de la topologie triangulaire). Pour cela, M. Felsberg propose de projeter également dans le cône les points représentatifs de tous les pixels de son voisinage. La convexité du modèle conique stipule alors que si l'on réalise un moyennage de tous ces points représentatifs, la position de ce point résultant sera forcément sur ou à l'intérieur du cône. Ainsi, dans le cas où tous les pixels voisins ont tous une même orientation, tous les points représentatifs du voisinage seront proches les uns des autres sur la surface du cône, et leur moyenne sera également sur la surface du cône (*i.e.* un voisinage fortement i1D). Au contraire, si les pixels n'ont pas la même orientation, les points représentatif seront projetés aléatoirement sur la surface du cône et leur position moyenne sera proche de son axe (*i.e.* un voisinage fortement i2D).

Afin de mieux comprendre le fonctionnement de cette méthode d'estimation, nous proposons de résumer ces différentes étapes de manière successive :

1. **Représentation complexe de l'image** : On extrait les informations d'amplitude locale et d'orientation locale à partir du signal monogène de l'image initiale  $I$ . On peut alors les représenter sous la forme d'un nombre complexe  $G$  :

$$G(\mathbf{x}) = A(\mathbf{x}) \exp(j\theta(\mathbf{x})),$$

où  $A$  représente l'amplitude locale et  $\theta$  l'orientation locale.

2. **Normalisation et représentation en angle double** : On normalise l'amplitude locale pour qu'elle se situe dans l'intervalle  $[0, 1]$  par la fonction de "seuillage doux"  $N$  précédemment mentionnée. Puisque nous avons vu que  $p_x = p_y = N(A(\mathbf{x}))$  pour un pixel  $\mathbf{x}$  donné, la représentation en angle double est donc la suivante :

$$z(\mathbf{x}) = N(A(\mathbf{x})) \exp(j2\theta(\mathbf{x})).$$

3. **Représentation conique** : On peut alors positionner le point représentatif du pixel  $\mathbf{x}$  dans le cône en utilisant les coordonnées  $\mathbf{c}(\mathbf{x}) = (c_1, c_2, c_3)(\mathbf{x}) = (|z|, \text{Re}\{z\}, \text{Im}\{z\})(\mathbf{x})$ . A ce stade, chaque pixel de l'image possède un point représentatif sur la surface du cône.
4. **Moyennage dans le cône** : Pour savoir si un pixel appartient à une région *i0D*, *i1D* ou *i2D*, il faut maintenant tenir compte de la notion de voisinage. Pour cela, M. Felsberg propose d'utiliser une fenêtre de pondération gaussienne sur les coordonnées coniques des pixels voisins de  $\mathbf{x}$  (*i.e.* un moyennage à l'intérieur du cône). Les dimensions de cette gaussienne définiront la taille du voisinage. Ainsi pour un pixel donné  $\mathbf{x}$ , on réalise la

moyenne pondérée des coordonnées coniques des pixels contenus dans ce voisinage par la relation :

$$\mathbf{c}'(\mathbf{x}) = \sum_{\mathbf{i}} \omega_{\mathbf{i}} \mathbf{c}(\mathbf{i}),$$

où  $\mathbf{i}$  est la coordonnée d'un pixel voisin de  $\mathbf{x}$ , et  $\omega_{\mathbf{i}}$  son poids associé à la gaussienne 2D centrée sur le pixel  $\mathbf{x}$ . Le paramètre d'écart-type  $\sigma$  de la gaussienne est fixé à 1 pixel par l'auteur. Par conséquent, ce sont les pixels à l'intérieur du lobe principal de la gaussienne qui auront le plus d'importance sur le résultat. Plus le pixel sera à l'extérieur du lobe principal, et moins son orientation monogène sera influente sur la notion d'orientation privilégiée dans le voisinage du pixel courant  $\mathbf{x}$ . Nous reviendrons sur cette notion de voisinage dans la suite de ce manuscrit, et en particulier lors de l'adaptation de cette estimation des dimensions intrinsèques continues pour le cas particulier des images sonar à la section 5.4.1.

5. **Représentation triangulaire** : Une fois la coordonnée moyenne du voisinage estimée, on projette le point  $\mathbf{c}'(\mathbf{x})$  dans le triangle d'orientation  $2\theta$  :

$$(\hat{x}(\mathbf{x}), \hat{y}(\mathbf{x})) = \left( c'_1(\mathbf{x}), \sqrt{(c'_2(\mathbf{x}))^2 + (c'_3(\mathbf{x}))^2} \right).$$

6. **Coordonnées barycentriques** : D'après les équations (5.4), (5.5) et (5.6) nous obtenons les coordonnées barycentriques du point, et donc par conséquent, les 3 indices de confiance concernant sa dimension intrinsèque continue :

$$\begin{aligned} c_{i0D}(\mathbf{x}) &= 1 - \hat{x}(\mathbf{x}) \\ c_{i1D}(\mathbf{x}) &= \hat{y}(\mathbf{x}) \\ c_{i2D}(\mathbf{x}) &= \hat{x}(\mathbf{x}) - \hat{y}(\mathbf{x}). \end{aligned}$$

Au final, l'utilisation des informations locales fournies par le signal monogène nous permet d'estimer la notion de dimension intrinsèque continue en chaque pixel de l'image. Ainsi, à partir d'une image initiale  $I$ , nous obtenons donc trois images résultantes  $c_{i0D}$ ,  $c_{i1D}$  et  $c_{i2D}$  qui nous renseignent sur le degré de confiance avec lequel chaque pixel appartient à une zone  $i0D$ ,  $i1D$  et  $i2D$  de l'image respectivement. Dans le paragraphe suivant, nous proposons d'illustrer ces informations sur des données synthétiques et réels

### 5.3.2.5 Exemples sur des données synthétiques

A la section 5.2, nous avons pu voir que dans le cadre du traitement d'images, le concept de dimension intrinsèque vise à discriminer les signaux 2D homogènes ( $i0D$ ), les signaux orientés selon une direction spécifique ( $i1D$ ) des autres signaux plus complexes ( $i2D$ ), n'étant ni  $i0D$  ni  $i1D$ .

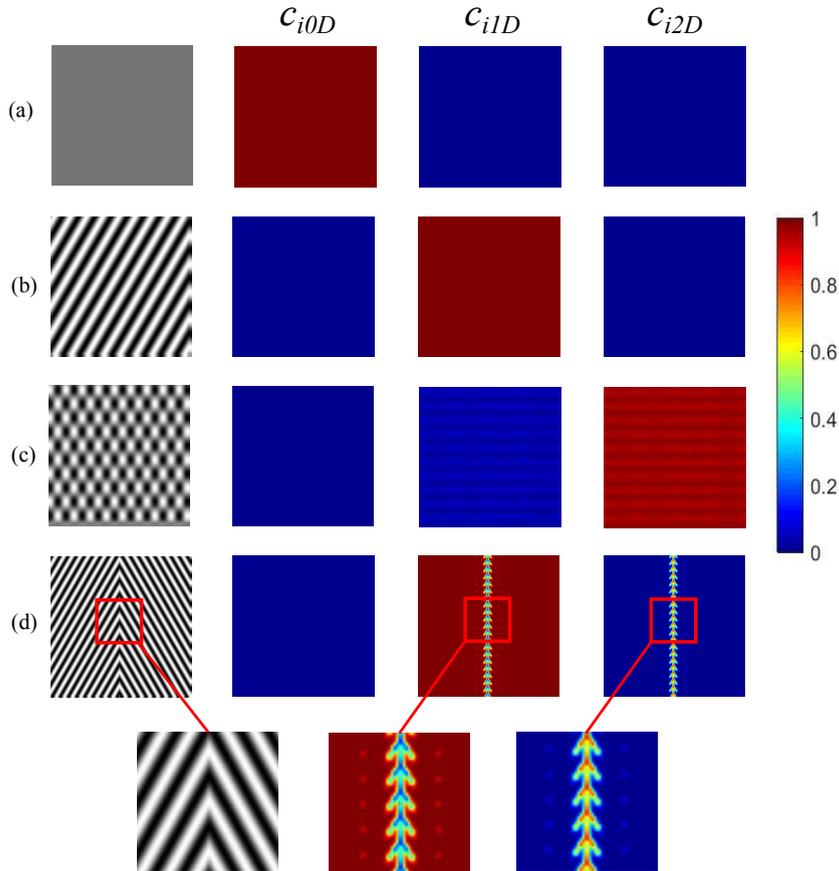


FIGURE 5.8: Illustration des indices de confiance  $c_{i0D}$ ,  $c_{i1D}$  et  $c_{i2D}$  pour quatre images synthétiques.

Pour illustrer le concept de dimension intrinsèque continue, nous proposons alors d'estimer les trois indices de confiances  $c_{i0D}$ ,  $c_{i1D}$  et  $c_{i2D}$  précédemment introduit sur quatre exemples synthétiques représentés sur la Figure 5.8. Ces derniers sont constitués d'un signal homogène, d'une sinusoïde 2D, d'une superposition de deux sinusoïdes 2D d'orientations orthogonales, et finalement d'une image formée d'une sinusoïde 2D orientée dans le plan de gauche et de son symétrique par rapport à l'axe central vertical dans le plan de droite. Cela fait apparaître un motif en chevron au niveau de cet axe central.

Pour le signal homogène (ligne (a) de la Figure 5.8), on remarque que l'ensemble des pixels ont une forte confiance au cas  $i0D$ . Comme attendu, l'ensemble des valeurs  $c_{i0D}$  sont proches de 1, et par complémentarité, les valeurs  $c_{i1D}$  et  $c_{i2D}$  sont alors proches de 0.

Dans le cas de la sinusoïde 2D illustrée sur la ligne (b) de la Figure 5.8, on note que cette fois, l'ensemble des pixels de l'image ont une forte confiance au cas  $i1D$ . En effet, leurs valeurs  $c_{i1D}$  sont toutes proches de 1. Ce qui apparaît naturel, puisqu'une sinusoïde 2D est un signal d'orientation constante, qui vérifie la définition d'une texture  $i1D$ .

Par contre, lorsqu'on superpose deux sinusoïdes 2D d'orientations distinctes, la texture de l'image devient plus complexe. Le signal résultant n'est alors ni constant, ni orienté selon

une seule et même direction. En d'autres termes, il vérifie la définition d'une texture i2D. Cela se vérifie sur la Figure 5.8.(c) dans laquelle l'ensemble des pixels de l'image possède une forte confiance  $c_{i2D}$  proche de 1.

Si l'on considère l'image proposée à la Figure 5.8.(d), on remarque que pour chaque sinusoïde prise séparément, on peut réaliser les mêmes observations faites précédemment dans le cas d'un signal i1D. La majorité de leurs pixels ont un fort indice de confiance au cas i1D. Cependant, au niveau de la jonction des deux sinusoïdes, le motif en chevron rompt le caractère i1D du signal. La texture locale de l'image n'est alors ni i0D ni i1D, elle possède deux orientations majeures. De ce fait, l'indice de confiance  $c_{i2D}$  augmente dans cette zone. Le zoom proposé sur cette jonction montre clairement le motif de chevron qui apparaît sur les indices  $c_{i2D}$ .

### 5.3.2.6 Exemple sur une image réelle

Nous proposons ici d'illustrer le résultat de l'estimation de ces trois indices de confiance sur une image réelle à savoir une image représentant une ruelle ici (cf. Figure 5.9). L'intérêt de cette image est qu'elle laisse entrevoir un certain nombre de motifs classiques, tels que des zones homogènes (les murs des différents bâtiments), des bords très droits (le cadre des fenêtres par exemple) et des angles (par exemple le coin de ces mêmes fenêtres). On constate que l'estimation des dimensions intrinsèques continues permet de mettre en évidence ces différents motifs de manière distincte sur les trois cartes proposées. En effet, alors que les zones homogènes de

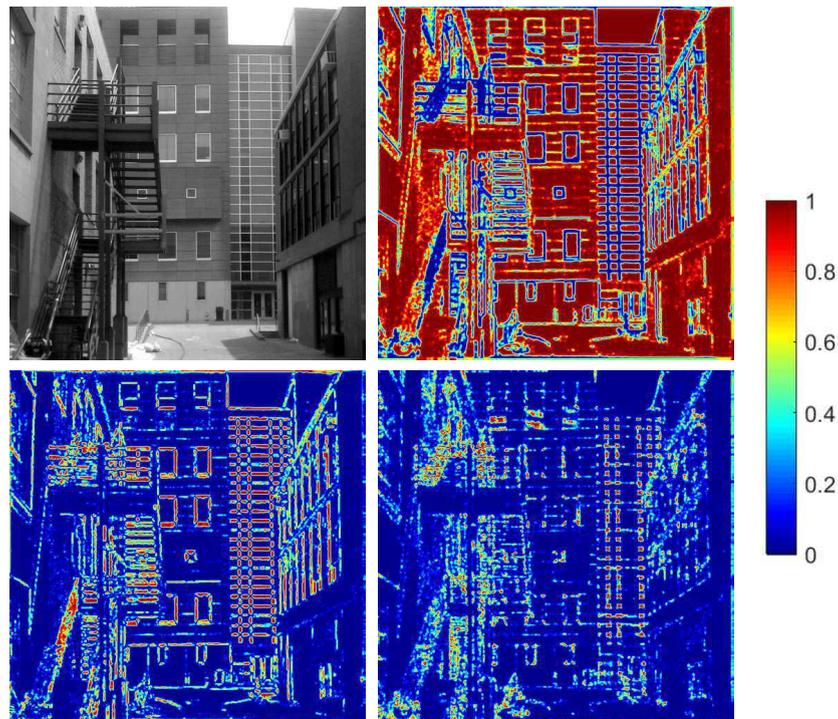


FIGURE 5.9: Illustration des indices de confiance  $c_{i0D}$ ,  $c_{i1D}$  et  $c_{i2D}$  pour une image réelle.

l'image ont des indices de confiance  $c_{i0D}$  très importants (image en haut à droite), les bords des différentes fenêtres ou ceux des marches de l'escalier (à gauche de l'image initiale) ont de forts indices de confiance  $c_{i1D}$  (image en bas à gauche). Concernant les dimensions intrinsèques continues  $c_{i2D}$ , on constate que ce sont les coins des fenêtres qui réagissent fortement puisque ces motifs sont par définition  $i2D$ .

Sur ces différents exemples nous nous sommes attachés à illustrer l'estimation des dimensions intrinsèques continues par le biais des informations fournies par signal monogène. Cependant dans la section 4.3, nous avons vu que le signal monogène permet également une analyse multi-échelle d'une image. Dans le paragraphe suivant, nous proposons d'en tirer parti afin de définir des dimensions intrinsèques continues dans le cadre multi-échelle.

### 5.3.3 Dimensions intrinsèques continues multi-échelles

#### 5.3.3.1 Définition

En utilisant les informations issues du signal monogène pour estimer les dimensions intrinsèques continues, nous avons unifié la représentation monogène d'une image (que nous avons vu au chapitre 4) et l'analyse des dimensions intrinsèques (que nous venons de voir dans ce chapitre). En particulier, c'est l'amplitude et l'orientation locales fournies par le signal monogène qui permettent une telle estimation. Or, à la section 4.5.4, nous avons constaté que l'extension multi-échelle du signal monogène permet à ces deux informations d'être également disponibles pour un ensemble d'échelles d'analyse. Pour rappel, nous avons noté ces deux informations  $A^{(s)}$  et  $\theta^{(s)}$ , où  $s$  désigne le paramètre d'un filtre passe-bande DoP qui spécifie le niveau de détail mis en évidence dans l'espace des échelles. Ce que nous proposons dans ce paragraphe est donc de reprendre exactement la démarche précédente pour l'estimation des dimensions intrinsèques continues de M. Felsberg (que nous avons explicité dans le paragraphe 5.3.2.4), mais cette fois, en remplaçant l'amplitude et l'orientation locales monogènes  $A$  et  $\theta$  par leurs versions multi-échelles  $A^{(s)}$  et  $\theta^{(s)}$ . Ainsi, pour chaque niveau de détail que l'on souhaite analyser, nous obtenons des indices de confiance  $c_{i0D}^{(s)}$ ,  $c_{i1D}^{(s)}$  et  $c_{i2D}^{(s)}$  en chaque pixel de l'image : c'est ce que nous proposons d'appeler les **dimensions intrinsèques continues multi-échelles**.

Afin d'illustrer ces propos, nous proposons de représenter ces nouvelles informations multi-échelle sur une image classique en traitement d'images à savoir Lena.

#### 5.3.3.2 Illustration des dimensions intrinsèques continues multi-échelles

Comme nous pouvons le voir sur la Figure 5.10, cette image est constituée de régions aux propriétés diverses, aux motifs de différentes tailles et orientations. Certaines d'entre elles sont d'ailleurs notoires, comme par exemple l'épaule de la personne, le mur derrière elle ou encore la surface des poutres qui correspondent à des régions relativement homogènes. Les

contours de ces régions, mais également ceux du cadre du miroir (en noir) à droite de Lena, sont caractéristiques de zones de fort contraste que nous appellerons des bords. La chevelure et les plumes du chapeau sont quant à elles des exemples de régions très texturées.

Nous proposons d'étudier cette image à travers quatre niveaux de détail. Si l'on se réfère à la décomposition multi-échelle réalisée dans le paragraphe 4.5.3.2, cette décomposition multi-échelle est effectuée sur 4 octaves chacun subdivisé en un seul intervalle (*i.e.*  $k = 2$ ). En proposant un premier niveau de détail pour un paramètre  $s = 1$ , nous avons donc les niveaux de détail suivant pour des valeurs de  $s$  égales à 2, 4 puis 8 respectivement. Ces quatre niveaux de détail sont illustrés du plus fin ( $s = 1$ ) au plus grossier ( $s = 8$ ) selon les lignes de la Figure 5.10. Pour chacun de ces niveaux de détail, nous représentons quatre informations, à savoir l'image filtrée par les filtres DoP de l'analyse multi-échelle et l'estimation des 3 indices de confiance  $c_{i0D}^{(s)}$ ,  $c_{i1D}^{(s)}$  et  $c_{i2D}^{(s)}$  pour chacune de ces images filtrées.

Au niveau de détail le plus fin, ce sont les détails hautes fréquences qui sont conservés sur l'image. Les bords des structures de l'image sont alors localisés de manière précise. Les pixels correspondant à de forte valeurs  $c_{i0D}$  appartiennent aux zones homogènes comme l'épaule, le chapeau, ou encore le mur derrière la personne. Au contraire, les pixels appartenant aux bords de ces structures possèdent de forte valeurs  $c_{i1D}$  qui reflètent leur caractère orienté. La texture des régions relatives à la chevelure de la personne étant plus complexe, elle apparaît de manière plus nette dans l'image  $c_{i2D}^{(s=1)}$ . Les yeux de Lena ont également une forte confiance au cas i2D.

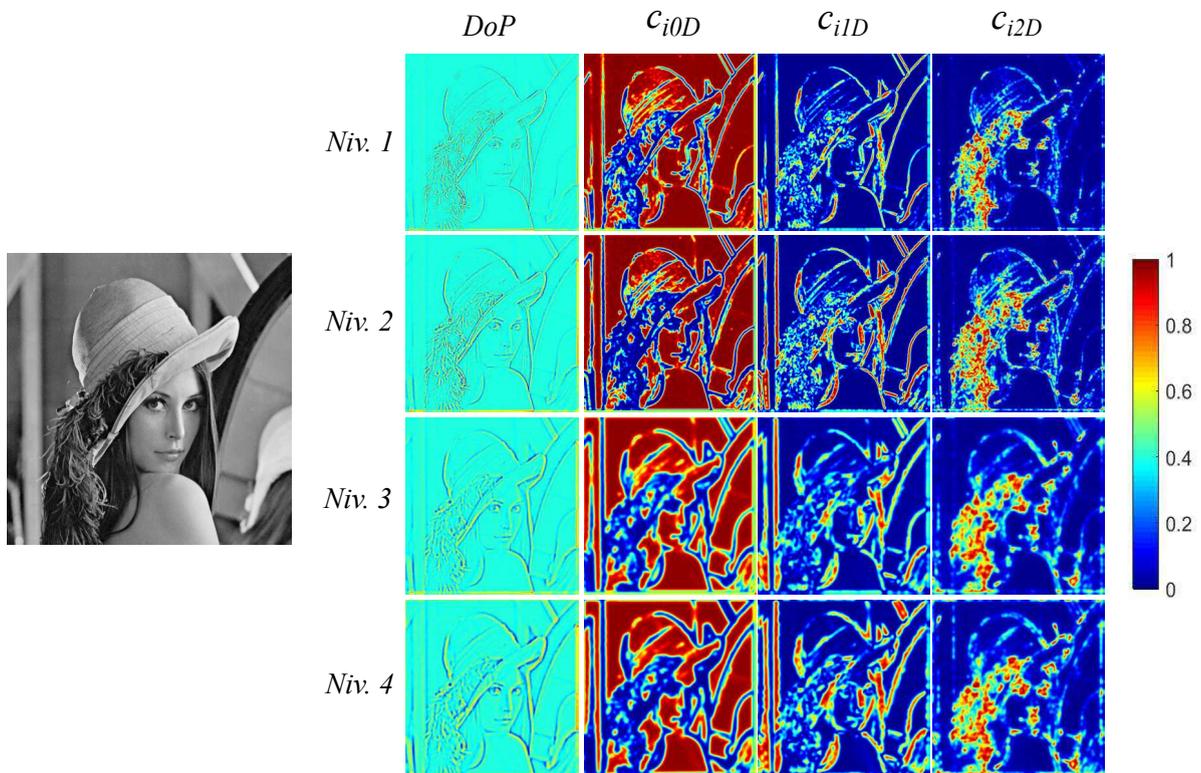


FIGURE 5.10: Illustration de la dimension intrinsèque multi-échelle appliquée à l'image Lena.

Plus l'on augmente le niveau de détail plus les hautes fréquences sont filtrés de l'image et les détails basses fréquences (*i.e.* plus larges) sont mis en évidence. Ceci se répercute clairement sur les indices de confiance à travers les niveaux de détail. En particulier, nous pouvons prendre l'exemple des cheveux de Lena. Dans cette zone, les fortes valeurs de  $c_{i2D}$  se regroupent en une seule et même région à mesure qu'on augmente le niveau de détail. De même sur les bords du miroir, les fortes valeurs de  $c_{i1D}$  forment des motifs de plus en plus larges. Ainsi, on constate que l'on dégrade au fur et à mesure la résolution de l'image et par conséquent la précision sur la localisation des motifs qu'elle contient.

Nous proposons maintenant d'appliquer le concept de dimension intrinsèque continue aux images relatives à ces travaux de thèse, à savoir les images sonar.

## 5.4 Application à l'imagerie sonar

Au cours de ce manuscrit, nous avons pu voir que dans le cadre d'une application ATR il est d'intérêt de décrire les fonds marins en trois catégories, à savoir les fonds homogènes, les fonds anisotropes et les fonds complexes. Or, dans ce chapitre, nous avons constaté que le concept de dimension intrinsèque est un outil qui permet de distinguer les textures d'une image en terme de complexité de son contenu, et semble donc pertinent pour notre étude. Ainsi, nous proposons d'étudier dans ce qui suit l'application des dimensions intrinsèques continues multi-échelles sur des images sonar afin d'analyser les informations que nous pourrions en extraire dans le but de définir des descripteurs suivant les trois types de fonds précédemment cités. Cependant, avant toute chose, il est tout d'abord nécessaire de paramétrer notre algorithme d'estimation des dimensions intrinsèques continues multi-échelles pour qu'il s'adapte aux données particulières que représentent les images issues d'un sonar latéral.

### 5.4.1 Paramétrage spécifique aux images sonar

Nous avons vu à la section 4.6.2 que le signal monogène peut être utilisé pour représenter l'image sonar d'une manière différente, par le biais d'une séparation orthogonale des informations locales d'amplitude, d'orientation et de phase, le tout dans le cadre d'une analyse multi-échelle. L'accès aux résolutions spatiales de l'image sonar permet à cette analyse multi-échelle de se focaliser sur des détails de l'image ayant des dimensions spécifiques (en mètres), et que nous supposons être dignes d'intérêts pour la suite de l'étude. Dans la section précédente, nous avons vu que les informations issues du signal monogène peuvent permettre d'estimer les dimensions intrinsèques continues de manière multi-échelle afin de fournir des indices de confiance quant à la complexité locale de la texture de l'image.

Dans la section 5.3.2.3, nous avons vu que la notion de voisinage est un aspect important lors de l'estimation des dimensions intrinsèques continues. Jusqu'à présent, nous avons suivi

le paramétrage proposé par M. Felsberg, à savoir l'utilisation d'une moyenne pondérée sur un voisinage donné par une gaussienne 2D d'écart-type 1 pixel, et ce, pour tous les niveaux de détail de l'analyse multi-échelle (cf. étape 4 du processus explicité à la section 5.3.2.4). Or, dans ce cadre multi-échelle et si l'on veut conserver un voisinage donné par une gaussienne 2D, il serait sûrement plus judicieux d'adapter la taille de ce voisinage et l'écart-type de la gaussienne en fonction du niveau de détail traité. En effet pour un niveau de détail donné, les structures de l'image ayant la dimension correspondante sont mises en évidence. Ainsi, si l'on désire savoir si cette structure appartient au cas i0D, i1D ou i2D, il serait nécessaire de considérer un voisinage qui englobe entièrement cette structure. De plus, afin de donner plus d'importance aux pixels de cette structure, il serait intéressant que cette dernière soit dans le lobe principale de la gaussienne 2D, ce qui signifie adapter son écart-type  $\sigma$ . Ce que nous proposons est donc de modifier légèrement l'étape 4 du processus d'estimation des dimensions intrinsèques continues (cf. section 5.3.2.4) en réalisant une moyenne pondérée par une gaussienne dont l'étendue du voisinage et l'écart-type dépendent, cette fois, du niveau de détail traité. Ainsi, si l'on considère un niveau de détail  $l$  (en mètres) et une taille de pixel de  $\Delta x = \Delta y$  dans l'image en mètres (pixels carrés), l'écart-type de la gaussienne à l'échelle  $l$  sera définie en nombre de pixels par :

$$\sigma^{(l)} = \frac{1}{2} \times \frac{l}{\Delta x}, \quad (5.11)$$

de telle sorte que le lobe principal de cette gaussienne soit alors approximativement d'un rayon de dimension  $\sigma^{(l)}$ , ce qui lui permet d'intégrer entièrement la structure ayant la dimension  $l$ . Nous choisissons ensuite de définir la taille du voisinage induit par une telle pondération de manière à tronquer la gaussienne à  $2\sigma^{(l)}$ . Pour chaque échelle  $l$ , la dimension du voisinage est donc définie par une gaussienne d'écart-type  $\sigma^{(l)}$  sur un support de dimension  $2\sigma^{(l)} \times 2\sigma^{(l)}$ .

La Table 5.2 propose un exemple de calcul des tailles de voisinages pour les images sonar. Pour les 12 niveaux de détail utilisés dans l'analyse multi-échelle, on y retrouve la dimension du voisinage (en mètres et en pixels) pour une résolution d'image sonar fixée à  $\Delta x = \Delta y = 0.1$  m. Il est à noter que la conversion entre les mètres et les pixels est réalisée de manière à avoir une dimension impaire en termes de pixels.

Maintenant que nous avons vu le paramétrage de l'estimation des dimensions intrinsèques continues multi-échelles que nous proposons dans le cas particulier de l'imagerie sonar, nous allons analyser séparément ses résultats sur des fonds marins typiques (*i.e.* homogènes, anisotropes et complexes).

#### 5.4.2 Analyse des fonds marins homogènes

Les fonds marins homogènes ont été définis à la section 3.3.1 comme les fonds constitués de diffuseurs dont les dimensions sont inférieures à la taille de la cellule de résolution du sonar. C'est donc le cas pour le sable et la vase. Dans de telles régions, aucune structure sous-marine

Niveaux de détail $l$ (en mètres)	Voisinages (en mètres / pixels)
0.25	$0.5 \times 0.5 / 5 \times 5$
0.32	$0.64 \times 0.64 / 7 \times 7$
0.41	$0.82 \times 0.82 / 9 \times 9$
0.5	$1 \times 1 / 11 \times 11$
0.63	$1.26 \times 1.26 / 13 \times 13$
0.8	$1.6 \times 1.6 / 17 \times 17$
1	$2 \times 2 / 19 \times 19$
1.25	$2.5 \times 2.5 / 25 \times 25$
1.5	$3 \times 3 / 31 \times 31$
1.8	$3.6 \times 3.6 / 37 \times 37$
2.15	$4.3 \times 4.3 / 43 \times 43$
2.5	$5 \times 5 / 51 \times 51$

TABLE 5.2: Liste des échelles utilisées dans cette étude et les dimensions du voisinage associé pour une résolution de l'image sonar supposée de  $\Delta x = \Delta y = 0.1$  m.

n'est donc identifiable. Un exemple d'une telle région est illustrée sur la Figure 5.11 et nous proposons d'estimer et d'analyser les dimensions intrinsèques continues multi-échelle de cette image. Nous avons vu à la section 4.6 que notre analyse multi-échelle se présente comme une succession de 12 filtrages passe-bande DoP se focalisant chacun sur des détails de l'image compris dans l'intervalle  $[0.25\text{m}, 2.5\text{m}]$ . Nous proposons ici de représenter quatre niveaux de détail en affichant la sortie des filtres DoP pour chacun d'entre eux ainsi que l'estimation des trois indices de confiance  $c_{i0D}$ ,  $c_{i1D}$  et  $c_{i2D}$  correspondants.

On peut alors remarquer que pour une zone de fond homogène, l'absence de structure sous-marine mène à de forts indices de confiance sur le caractère i0D de la zone, et ce sur l'ensemble des pixels ainsi que sur la quasi-totalité des niveaux de détail. On remarque tout de même que sur le niveau de détail le plus fin, le bruit vient légèrement polluer l'information, en fournissant de très faibles confiances au cas i2D (le bruit est par nature un élément i2D). Cependant, son influence décroît très rapidement avec l'augmentation du niveau de détail analysé.

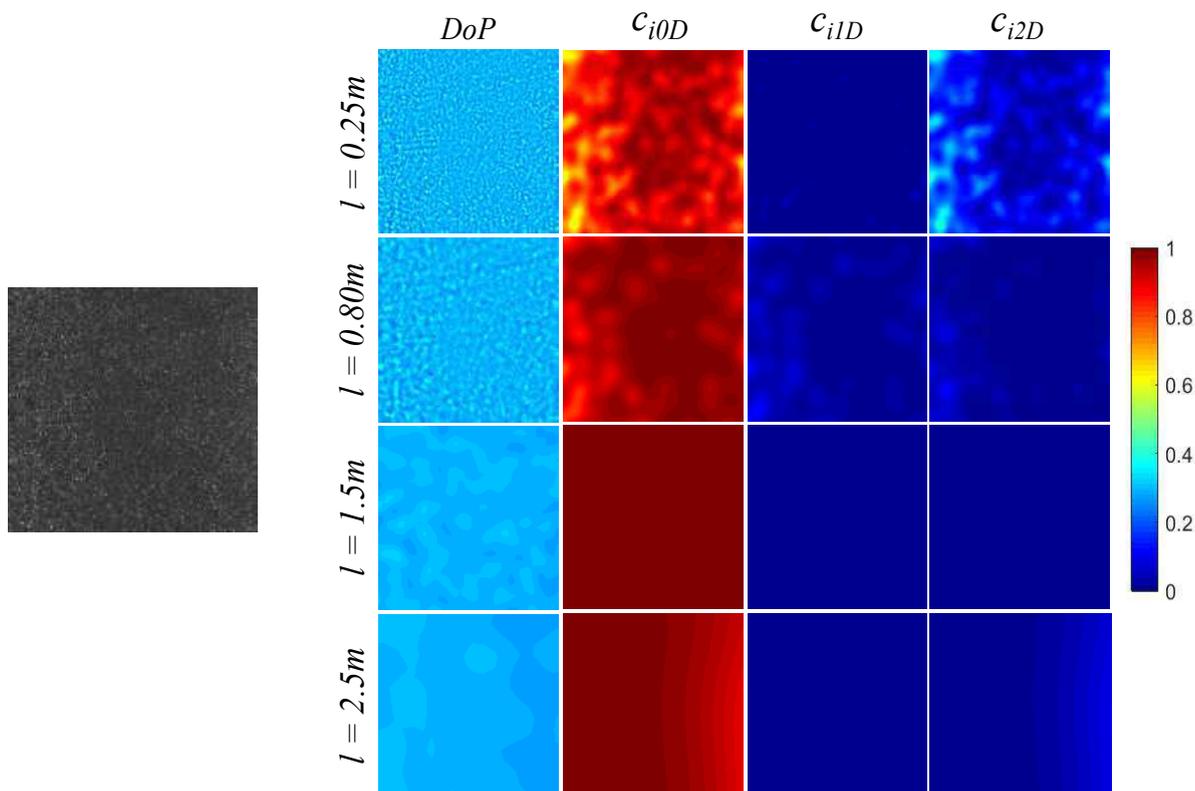


FIGURE 5.11: Illustration de la dimension intrinsèque continue multi-échelle appliquée à un fond marin homogène.

On peut alors en conclure qu'une zone homogène peut facilement être décrite par le biais des indices de confiances  $c_{i0D}$  et donc par sa nature i0D.

### 5.4.3 Analyse des fonds marins anisotropes

Comme dans le cas des fonds homogènes précédents, les fonds anisotropes ont été définis à la section 3.3.1 comme les fonds marins constitués de rides de sable. Ce sont des fonds qui possèdent donc une forte directionnalité, avec une succession d'échos et d'ombres importante.

Un exemple d'une telle région est proposée sur la Figure 5.12. On y observe un champ de ride de sable sur la diagonale ascendante de l'image entouré par du sable fin homogène qui permettra d'aider l'analyse des résultats dans la suite. Notons également qu'un trait vertical très intense est visible sur cette image. Nous avons vu au chapitre 2 que ce motif est dû à un reflet de la surface. Nous présentons sur cette même figure le résultat du filtrage passe-bande pour 4 niveaux de détails distincts, ainsi que les dimensions intrinsèques continues correspondantes.

On remarque que la sélection fréquentielle réalisée par l'analyse multi-échelle ne permet pas d'observer de la même manière les rides de sable sur l'ensemble des niveaux de détail. En effet sur cet exemple, les réponses des filtres DoP sur les rides de sables semblent être plus fortes pour  $l = 0.8$  m. Les rides de sables y sont bien plus contrastées qu'au niveau de détail

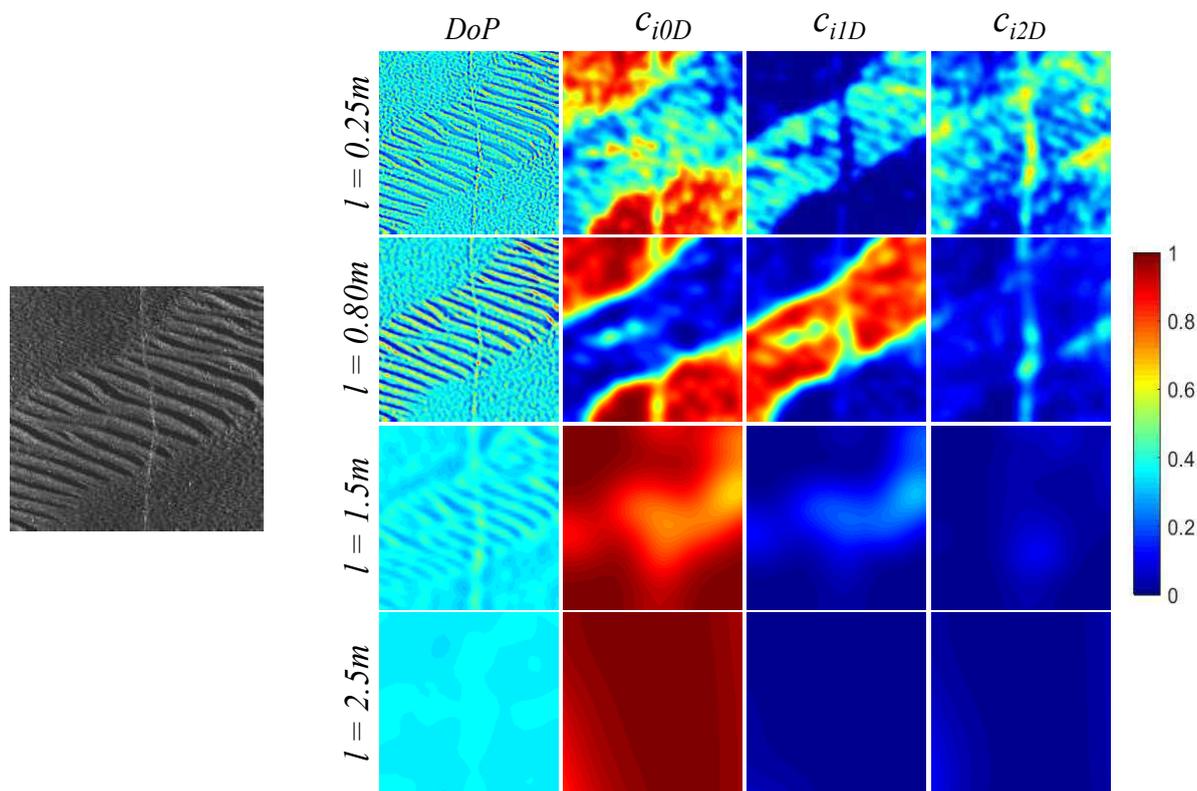


FIGURE 5.12: Illustration de la dimension intrinsèque continue multi-échelle appliquée à un fond marin anisotrope.

$l = 0.25m$ , et beaucoup moins filtrées qu'aux niveaux supérieurs. Finalement, on peut alors en conclure que ce niveau de détail correspond à la taille des rides de sable dans l'image sonar.

Par ailleurs pour  $l = 0.8$ , on observe que l'indice de confiance  $c_{i1D}$  est très important sur la majorité des pixels du champ de rides de sable. On note tout de même quelques pixels de faible valeur sur la partie inférieure du champ au niveau de l'artefact lié au reflet de la surface, ainsi qu'à gauche de celui-ci, puisqu'une petite zone homogène y est observable.

Aux niveaux de détail supérieurs, les rides de sable sont filtrées au fur et à mesure. Cette zone devient alors au fur et à mesure homogène. Cet effet est clairement visible sur l'information  $c_{i0D}$ . En effet, les pixels du champ de rides de sable voient leurs confiances en la dimension intrinsèque de 0 augmenter jusqu'à approcher une valeur de 1 à le niveau de détail le plus grossier. Ainsi, l'indice de confiance  $c_{i0D}$  peut être vu comme une information sur l'effet d'un filtrage. Plus cette information sera élevée, et plus le signal 2D aura été filtré.

La zone homogène autour du champ de rides de sable respecte toujours les propriétés vues dans la section précédente, à savoir une constance des forts indices  $c_{i0D}$  sur l'ensemble des niveaux de détail.

On note également que pour  $l = 0.25$  (niveau de détail le plus fin), certains pixels de l'image ont un fort indice de confiance  $c_{i2D}$ . Ce phénomène est dû en partie au bruit, mais

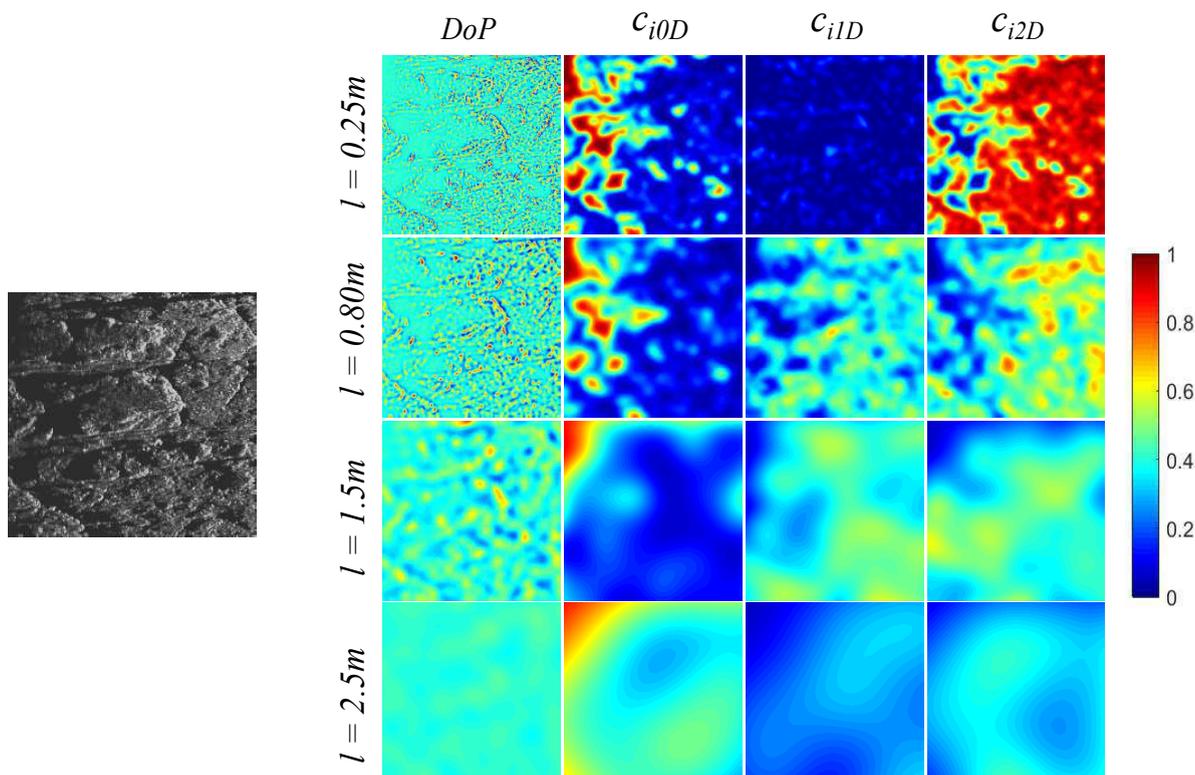


FIGURE 5.13: Illustration de la dimension intrinsèque continue multi-échelle appliquée à un fond marin complexe.

également à l'extrémité des rides de sable qui sont vues comme des coins et donc des motifs par définition  $i2D$ .

A travers cette exemple, on retiendra que les pixels d'une image appartenant à un champ de rides de sable possèdent la particularité d'avoir un fort indice de confiance  $c_{i1D}$  pour des niveaux de détail correspondant à la taille des rides de sable. Pour des détails plus grossiers, ces rides sont filtrées et ces mêmes pixels ont alors des valeurs de  $c_{i0D}$  très fortes.

#### 5.4.4 Analyse des fonds marins complexes

Nous avons défini à la section 3.3.1 les fonds marins complexes comme les régions d'une image sonar dans lesquelles il sera difficile de détecter une mine. Ces fonds particuliers sont généralement constitués de cailloutis, de roches, de végétations (ex : la posidonie).

L'exemple proposé sur la Figure 5.13 illustre le cas d'un fond marin encombré par un amas de roches aux tailles variables. La complexité de l'image provient non seulement des motifs complexes mais également des niveaux d'intensité très variables dans la zone. Dans une telle région, la présence d'une mine sous-marine sera très compliquée à détecter.

Comme précédemment, sur cette même Figure, nous proposons l'illustration de 4 niveaux de détails et l'estimation des dimensions intrinsèques continues correspondantes.

On observe tout d'abord que la grande diversité dans la taille des structures présentes dans l'image influe sur les réponses énergétiques des filtres DoP. En effet, ces dernières sont importantes sur les 3 premiers niveaux proposés. Ainsi dans l'image, il y a un mélange de structures allant de 0.25m jusqu'à 1.5m au minimum.

Concernant les dimensions intrinsèques continues, on remarque qu'au niveau de détail le plus fin, la zone apparaît quasiment exclusivement comme une texture i2D. En effet, la majorité des pixels possèdent alors un indice de confiance très fort  $c_{i2D}$ . Seules les zones d'ombres sont considérées comme fortement homogènes par le biais de forts indices de confiance  $c_{i0D}$ .

Plus l'on monte dans les niveaux de détail, et plus le phénomène de lissage prend de l'ampleur. Ainsi, les bords et les coins des roches s'arrondissent et laissent place à des indices de confiance de plus en plus répartis entre les cas i1D et i2D. Au dernier niveau de détail, les roches ayant commencé à être fortement filtrées, l'image se rapproche du cas i0D.

## 5.5 Résumé du chapitre

Dans ce chapitre, nous nous sommes intéressés à l'étude de la dimension intrinsèque, outil très peu utilisé dans le domaine du traitement d'image mais reconnu depuis de nombreuses années dans le domaine des statistiques. Pourtant, dans le contexte du traitement d'image, nous avons vu que ce concept mathématique peut être intéressant pour mettre en évidence la complexité locale d'une texture. Cette notion de complexité est basée sur l'analyse des degrés de liberté d'un signal 2D. Elle propose de distinguer trois cas de signaux 2D, à savoir les signaux constants i0D, les signaux qui ne varient que dans une seule et même direction i1D, et les signaux plus complexes i2D, qui ne sont ni i0D ni i1D. Cependant, puisque ces trois classes discrètes ne permettent pas de décrire des signaux réels, une définition de dimension intrinsèque continue a été utilisée.

Cette extension fournit trois indices de confiance complémentaires sur le comportement du signal 2D. Ces trois indices représentent la confiance avec laquelle on peut considérer un signal comme étant i0D, i1D et i2D. L'avantage de cette formulation continue est qu'elle peut être facilement calculée à partir des informations monogènes que nous avons introduites dans le chapitre 4. L'estimation de ces indices de confiance peut alors être réalisée en chaque pixel de l'image. Nous avons alors proposé d'exploiter ce lien entre le signal monogène et les dimensions intrinsèques continues pour définir une dimension intrinsèque continue multi-échelle, ce qui n'a jamais été traitée de manière formelle dans la littérature.

Nous avons proposé d'étudier d'analyser leur application sur des cas réels d'images sonar. Ainsi, nous avons alors pu constater que certains types de fonds marins peuvent se distinguer via l'exploitation de ces dimensions intrinsèques continues multi-échelles à différents niveaux de détail. En effet, alors que les fonds homogènes possèdent des textures fortement i0D sur

l'ensemble des niveaux de détails utilisés, les fonds anisotropes sont caractérisés par de fortes confiances au cas i1D sur certains niveaux. Au contraire, les fonds complexes mènent à des indices de confiance plus forts sur le cas i2D.

Ainsi, il semblerait que ces indices de confiance puissent être utilisés pour définir des descripteurs représentatifs de la difficulté de détection de cibles dans un contexte de chasse aux mines sous-marines. L'objectif du chapitre suivant est donc de mettre en place un tri automatique de l'ensemble des informations multi-échelles disponibles afin de fournir des descripteurs pertinents à une chaîne ATR.



# Chapitre 6

## Approche proposée pour la description des fonds marins

### Sommaire

---

<b>6.1</b>	<b>Introduction</b>	<b>149</b>
<b>6.2</b>	<b>Description continue des fonds marins</b>	<b>150</b>
6.2.1	Formulation des objectifs	150
6.2.2	Homogénéité des fonds	150
6.2.2.1	Carte d'homogénéité	151
6.2.2.2	Carte d'inhomogénéité	154
6.2.3	Anisotropie des fonds	156
6.2.3.1	Première approche	156
6.2.3.2	Etude sédimentologique des rides de sable	157
6.2.3.3	Intégration de données sédimentologiques dans la recherche d'anisotropie	159
6.2.3.4	Carte d'anisotropie	161
6.2.4	Complexité des fonds	162
6.2.5	Résultats	163
6.2.6	Comparaison avec une méthode existante	166
6.2.6.1	Comparaison algorithmique	166
6.2.6.2	Comparaison qualitative	167
6.2.6.3	Comparaison quantitative	171
6.2.7	Liens avec l'ATR	173
<b>6.3</b>	<b>Classification non-supervisée des fonds marins</b>	<b>175</b>
6.3.1	Introduction	175
6.3.2	Classification des fonds homogènes	175
6.3.2.1	Seuillage de la carte d'homogénéité	176

6.3.2.2	Paramétrage du seuillage . . . . .	178
6.3.3	Classification des fonds anisotropes et complexes . . . . .	179
6.3.4	Résultats . . . . .	180
<b>6.4</b>	<b>Conclusion . . . . .</b>	<b>183</b>

---

## 6.1 Introduction

Dans les chapitres précédents, nous nous sommes intéressés à deux outils mathématiques originaux. L'un propose une nouvelle représentation de l'image par une décomposition orthogonale des informations énergétiques, géométriques et structurelles, dans un cadre multi-échelle : le signal monogène. L'autre est un concept permettant d'étudier le contenu géométrique d'une image, c'est la dimension intrinsèque. Cette dernière possède une extension continue, appelée dimension intrinsèque continue, qui peut être estimée de manière locale dans le cadre multi-échelle par le biais du signal monogène.

Dans le chapitre précédent, nous avons pu voir que cette dimension intrinsèque peut potentiellement permettre de distinguer les fonds marins homogènes, anisotropes et complexes. Cependant, toutes les observations que nous avons faites dans l'espace multi-échelle ne permettent pas encore de fournir des descripteurs précis à propos de l'homogénéité, de l'anisotropie ou de la complexité d'une image sonar. En d'autres termes, jusqu'à présent, nous nous sommes limités à une extraction de l'information, et une étape supplémentaire dans le traitement est alors nécessaire, c'est la phase de description. Cette nouvelle étape du traitement est illustrée sur la Figure 6.1.

Dans ce chapitre, nous proposons une nouvelle approche de description de l'image sonar basée sur les différents outils que nous avons précédemment définis. Comme nous le verrons dans la section 6.2, l'un des objectifs principaux de ce chapitre est de pouvoir mettre en évidence les fonds marins où les opérations de chasse aux mines seront difficiles. Pour cela, nous verrons la définition de trois nouveaux descripteurs concernant la texture d'une image sonar. Ceux-ci fournissent chacun un degré de confiance sur le caractère homogène, anisotrope et complexe des fonds marins. Au delà de leurs dynamiques bornées, ces descripteurs sont également complémentaires ce qui nous permet d'envisager une classification de l'image en ces trois types de fonds. Nous n'aurons plus un continuum de valeurs continues pour représenter les fonds marins mais des valeurs discrètes liées aux labels de classification. La section 6.3 proposera de réaliser une telle classification non-supervisée des images sonar. Nous verrons

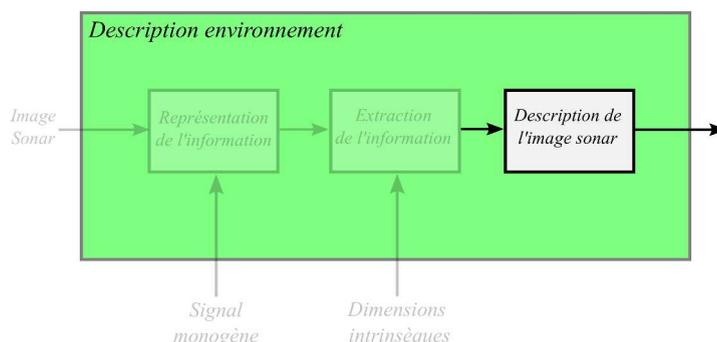


FIGURE 6.1: Position de la phase de description dans le processus.

que cette prise de décision sur les fonds homogènes, anisotropes et complexes, permet de les localiser précisément sur l'image sonar.

## 6.2 Description continue des fonds marins

### 6.2.1 Formulation des objectifs

Dans le chapitre précédent, nous avons pu constater que le concept de dimension intrinsèque continue semble être adéquat pour une description des images sonar en termes d'homogénéité, d'anisotropie et de complexité. Cependant, il faut maintenant convertir ces analyses et interprétations sous la forme de descripteurs. Par conséquent, nous chercherons à travers cette section à définir trois descripteurs qui seront reliés aux textures contenues dans une image sonar en termes d'homogénéité, d'anisotropie et de complexité. Néanmoins, afin de rendre ces descripteurs significatifs pour des algorithmes automatiques, voire utilisables pour des opérateurs à bord de navires, nous nous imposons certaines contraintes supplémentaires. En effet, puisque nous avons constaté que le concept de dimension intrinsèque continue possède des caractéristiques d'indices de confiance et de complémentarités, il serait intéressant que nos descripteurs en bénéficient également. Pour répondre à ces attentes, cette phase de description doit donc nous fournir trois descripteurs d'homogénéité, d'anisotropie et de complexité, bornés chacun dans un intervalle  $[0, 1]$ , et dont la somme en chaque pixel de l'image vaudra 1. La Figure 6.2 propose la schématisation des objectifs de cette approche de description du fond marin.

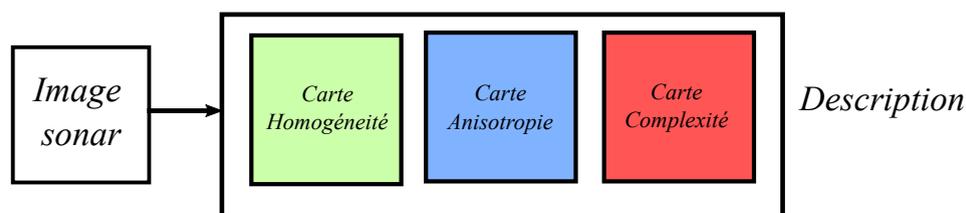


FIGURE 6.2: Synoptique de la première étape.

Dans le paragraphe suivant, nous allons nous intéresser à la définition d'un premier descripteur qui va nous fournir un indice de confiance sur l'homogénéité d'un fond marin.

### 6.2.2 Homogénéité des fonds

Nous avons vu à la section 3.3.1 que la littérature définit les fonds homogènes comme étant les types de fonds marins sur lesquels la détection et la reconnaissance d'une cible seront les plus performantes. Ils sont généralement constitués par le sable fin et la vase, où la paire écho-ombre de la mine sera clairement visible. Lorsqu'aucune structure sous-marine ou objet n'est posé dessus, ces types de fonds résultent en une image sonar relativement constante en

termes de niveaux d'intensité (si l'on ne considère ni le bruit de speckle ni les atténuations en distance dues aux pertes par propagation). D'ailleurs, nous avons pu constater au chapitre précédent (cf. paragraphe 5.4.2) que ces fonds marins peuvent être reliés à la notion de dimension intrinsèque de 0 (i0D). Il semble donc possible de quantifier cette homogénéité du fond marin à l'aide de la dimension intrinsèque de 0 et par conséquent, de quantifier l'absence de structure sous-marine (et en particulier de mine) dans une zone de l'image sonar. Ainsi, nous proposons de définir notre propre vision d'un fond marin homogène comme étant un fond marin où aucune structure sous-marine n'est présente. De ce fait, un fond homogène ne représente aucun danger pour la navigation et une perte de temps pour la détection de mines aussi bien pour un opérateur que pour un algorithme ATR. A travers les paragraphes qui suivent, nous proposons de définir un descripteur qui nous renseignera sur l'homogénéité d'un fond marin et qui respecte la contrainte d'une dynamique comprise entre 0 et 1 pour qu'il puisse être identifié à un indice de confiance.

### 6.2.2.1 Carte d'homogénéité

Pour définir un tel descripteur d'homogénéité, nous avons proposé une première méthode en étudiant le comportement des dimensions intrinsèques continues multi-échelles. Celle-ci est basée sur le moyennage des indices de confiances  $c_{i0D}$  sur l'ensemble des niveaux de détail que nous proposons dans notre analyse multi-échelle. Ainsi, si l'on considère  $c_{i0D}^{(k)}(\mathbf{x})$  comme étant l'indice de confiance  $c_{i0D}$  d'un pixel  $\mathbf{x}$  au niveau de détail d'indice  $k$ , le coefficient homogénéité est donné par :

$$H_1(\mathbf{x}) = \frac{1}{N} \sum_{k=1}^N c_{i0D}^{(k)}(\mathbf{x}), \quad (6.1)$$

où  $N$  correspond au nombre de niveaux de détail dans notre analyse, soit 12 ici. Les résultats obtenus ont été présentés lors d'une conférence [157].

Cependant, le moyennage est une opération reconnue comme étant très dépendante des valeurs extrêmes. De plus, nous avons pu voir que le bruit est un facteur qui dégrade plus ou moins fortement l'estimation des dimensions intrinsèques aux niveaux des détails fins de l'image. Ainsi, dans ce manuscrit, nous proposons une autre définition de l'homogénéité basée sur la valeur médiane de ces indices de confiance multi-échelles :

$$H(\mathbf{x}) = \text{Mediane}_k \left[ c_{i0D}^{(k)}(\mathbf{x}) \right]. \quad (6.2)$$

La Figure 6.3 propose le résultat de ces deux descripteurs d'homogénéité sous la forme de cartes (descripteurs disponibles en chaque pixels de l'image) pour une même image sonar issues de la base de données DGA (*i.e.* collectée par l'intermédiaire d'un système sonar latéral KLEIN 5000). Sur cette dernière, on y observe une large zone de sable fin sur sa moitié supérieure. Au milieu, on y remarque une zone rocheuse. Quant à sa moitié inférieure, on y note la présence de

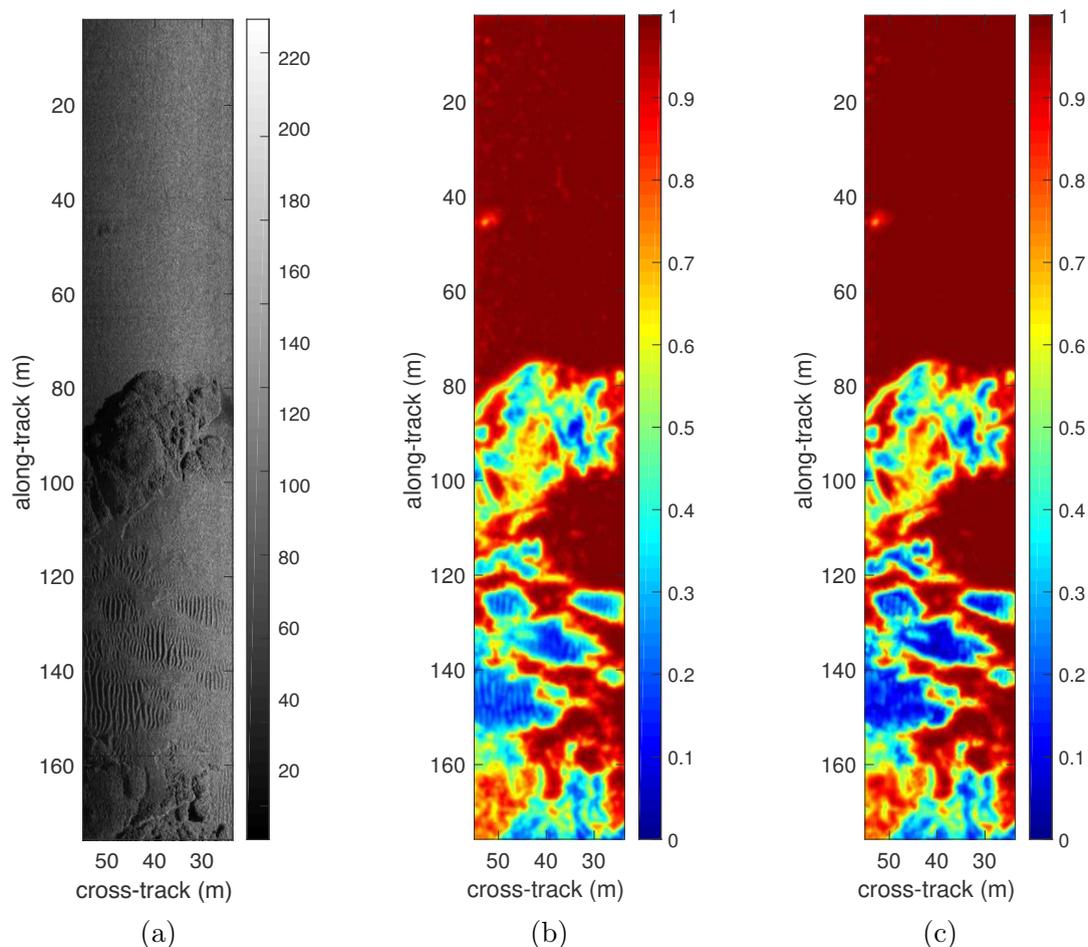


FIGURE 6.3: Exemple de deux cartes d'homogénéité pour une image sonar latéral Klein 5000 (a). En (b), une carte basée sur la moyenne, et en (c) une carte basée sur la médiane.

plusieurs champs de rides de sable. On peut constater sur cette figure que les deux descripteurs d'homogénéité fournissent pratiquement la même information. En effet, que ce soit en utilisant la moyenne ou la médiane, les descripteurs ont une dynamique comprise dans l'intervalle  $[0, 1]$ . Cette dernière permet de les voir comme deux indices de confiance sur l'homogénéité du fond marin. Ainsi, plus un pixel aura une valeur proche de 1 plus la confiance sera grande sur le fait qu'il appartienne à une zone de fond homogène. On observe ici que cette confiance sur l'homogénéité est maximale dans les zones sableuses, contrairement aux zones rocheuses et de rides de sable qui fournissent un indice de confiance faible.

Cependant, comme nous l'avions présumé, on remarque que la méthode basée sur la moyenne est légèrement plus sensible au bruit (ce qu'on observe en particulier sur les zones de sable). De plus, elle fournit également un contraste moins important entre les zones du fond sans structure sous-marine et celles avec. En effet, par exemple, la valeur de l'homogénéité sur les rides de sables sur la Figure 6.3(b) est plus importante que sur la Figure 6.3(c). Ceci conforte l'utilisation de la médiane comme méthode d'estimation de l'homogénéité.

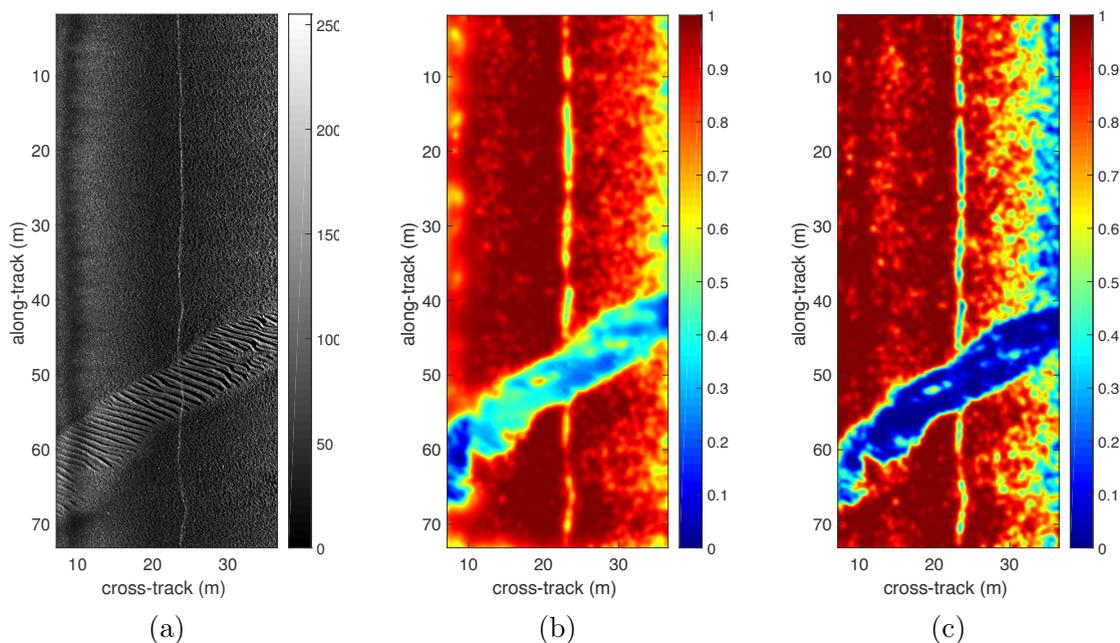


FIGURE 6.4: Exemple de deux cartes d'homogénéité pour une image sonar latéral Marine Sonic (a). En (b), une carte basée sur la moyenne, et en (c) une carte basée sur la médiane.

Enfin, nous pouvons noter que ces deux méthodes proposent une description locale (*i.e.* en chaque pixel) de l'image, qui résulte en des frontières relativement précises entre les différentes zones homogènes de l'image et les autres.

La Figure 6.4 propose d'illustrer les mêmes informations que sur la Figure 6.3 précédente, mais pour un autre système sonar. L'image est cette fois issue d'un système sonar latéral Marine Sonic (base de données DSTO). Elle présente un fond homogène traversé par un champ de rides de sable. Sur les deux cartes d'homogénéité présentées, on peut alors faire exactement les mêmes observations que précédemment. Le point important ici est que pour obtenir ces cartes, nous avons conservé les mêmes paramètres que pour l'exemple précédent. La seule chose qui change en entrée de notre algorithme est la dimension des pixels spécifique à chacune de ces images. Ceci illustre donc la robustesse de notre approche par rapport aux différents systèmes d'observation sonar.

Cependant, sur ces données, on note tout de même que les indices de confiance fournissent des valeurs légèrement moins importantes sur les fonds homogènes. Ce phénomène provient en grande partie du niveau de bruit du système, répercuté sur l'image sonar. Au niveau des détails fins, il impacte alors fortement l'estimation des dimensions intrinsèques continues, et il en résulte une réduction de la confiance en l'homogénéité d'une texture. De plus, on observe également une décroissance de l'indice d'homogénéité en fonction de la portée. Ce phénomène provient non seulement des effets de rasance en fin de portée, qui viennent déformer les structures de l'image, mais également de l'impact plus important du bruit à ces distances. Ainsi,

l'algorithme met en évidence une certaine inhomogénéité de la texture dans cette zone. Une telle décroissance n'est pas forcément un problème. Au contraire, nous avons défini précédemment les fonds homogènes comme étant des zones de l'image dans lesquelles aucune mine ni structure sous-marine ne sera présente. Or, en bout de portée, il est plus difficile de discerner la présence ou non d'un objet puisque son écho  $y$  possède un niveau d'intensité plus faible en raison des pertes par propagation et peut se cacher plus facilement dans le niveau de bruit (ambiant ou électronique). De ce fait, la décroissance de la confiance en l'homogénéité dans ces zones illustre cette incertitude sur la présence de mines.

Dans ce paragraphe, nous avons tiré parti de l'analyse des dimensions intrinsèques continues multi-échelles pour permettre de construire une carte d'homogénéité des fonds marins. Cette carte possède l'avantage d'avoir une dynamique comprise dans l'intervalle  $[0, 1]$ , assimilable à un indice de confiance comme le sont les dimensions intrinsèques continues. Son interprétation devient alors aisée pour un opérateur ou pour un système informatique. Par ailleurs, on peut extraire d'une telle carte son complémentaire, à savoir une carte d'inhomogénéité sur les fonds.

### 6.2.2.2 Carte d'inhomogénéité

L'inhomogénéité des fonds marins sera définie comme le complémentaire de la carte d'homogénéité. On l'exprime donc de la manière suivante :

$$\bar{H}(\mathbf{x}) = 1 - H(\mathbf{x}). \quad (6.3)$$

Ainsi, alors que la carte d'homogénéité exprime la simplicité de la texture du fond marin, la carte d'inhomogénéité exprime quant à elle une complexification de la texture. La Figure 6.5 propose d'illustrer de telles cartes pour les deux mêmes exemples d'images sonar étudiés précédemment. En particulier, cette figure permet d'observer la complémentarité entre  $H$  et  $\bar{H}$ . De plus, on y constate que les zones rocheuses et de rides de sable ont bien des indices d'inhomogénéité très importants (proche de 1).

A l'intérieur de ces zones fortement inhomogènes, il y a une grande population de fonds marins distincts qui peut intervenir (roche, rides de sable, posidonie etc.). Afin de les distinguer, nous proposons de définir un nouveau descripteur permettant de mettre en évidence les zones de rides de sable.

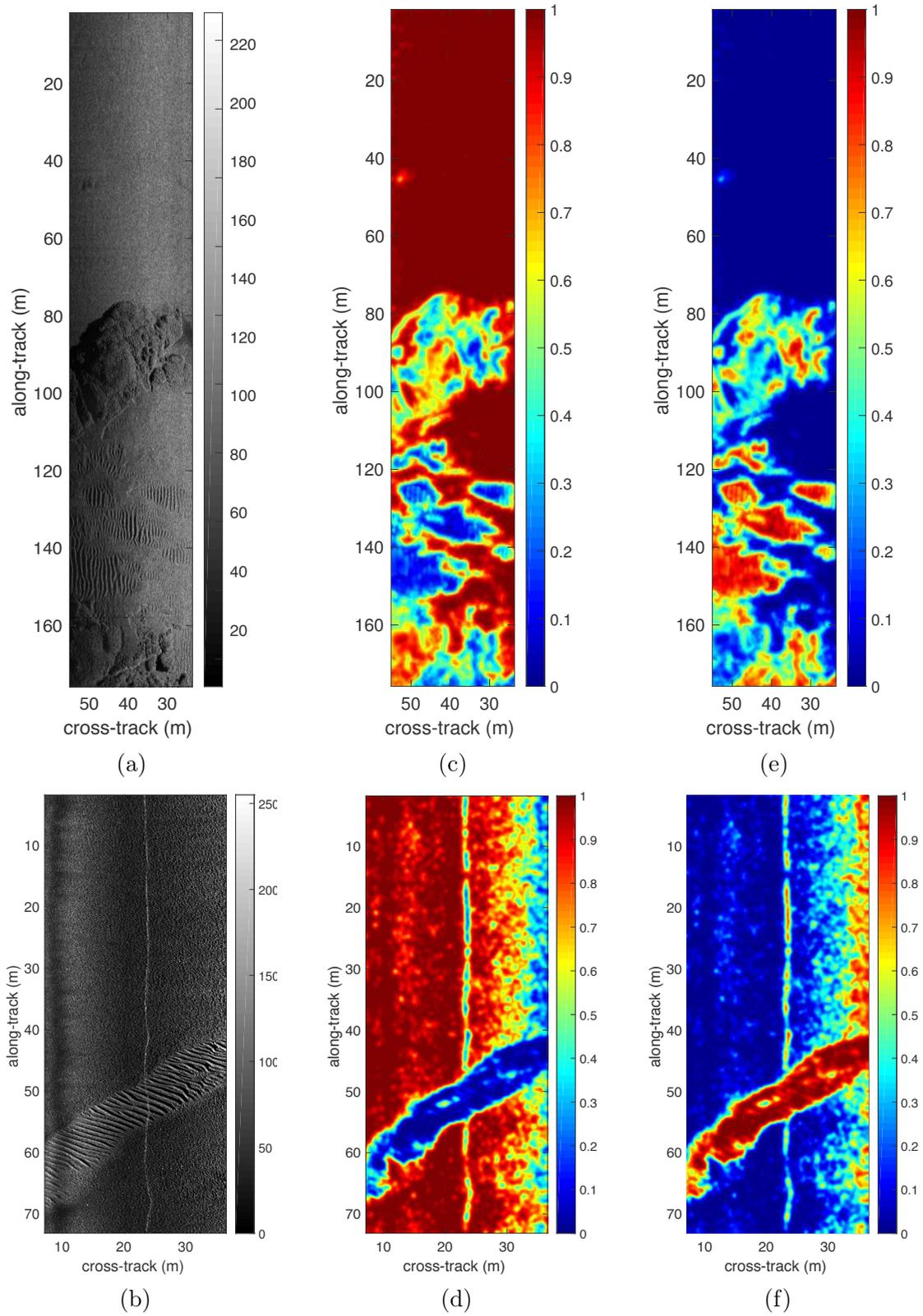


FIGURE 6.5: Exemple de deux images sonar (a) (b), ainsi que leur carte d'homogénéité (c) (d) et d'inhomogénéité (e)(f) respectives.

## 6.2.3 Anisotropie des fonds

### 6.2.3.1 Première approche

Les rides de sable apparaissent sur l'image sonar comme des textures dites directives ou anisotropes. La forte succession d'échos et d'ombres qui les caractérise est la raison pour laquelle les détecteurs de mines sous-marines y sont peu performants. En effet, les taux de fausses alarmes y sont généralement très importants, tout comme la probabilité de non détection d'une cible. Par conséquent, il est utile de définir un descripteur concernant la présence (ou non) de telles textures dans l'image sonar.

Comme nous l'avons observé au paragraphe 5.4.3, les pixels d'une image qui appartiennent à un champ de rides de sable possèdent de fortes confiances en la dimension intrinsèque de 1 pour le niveau de détail qui correspond à la longueur d'onde de ces dernières.

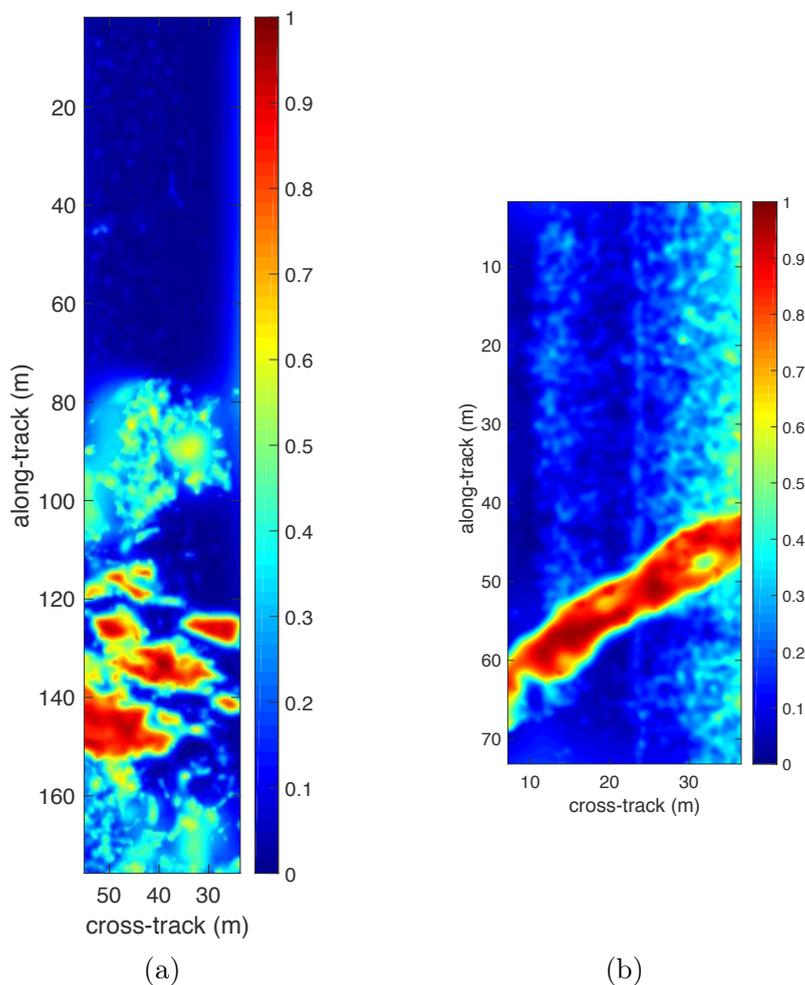


FIGURE 6.6: Exemple de deux cartes d'anisotropie (a) et (b) correspondantes aux images sonar présentées à la Figure 6.5.

Puisque que nous n'avons aucune information *a priori* sur la longueur d'onde des rides de sable, nous proposons dans un premier temps d'extraire une première information des images, à savoir le maximum des indices de confiance  $c_{i1D}$  à travers les échelles pour chaque pixel de l'image :

$$A_1(\mathbf{x}) = \max_{k \in [1;N]} \left( c_{i1D}^{(k)}(\mathbf{x}) \right), \quad (6.4)$$

où  $c_{i1D}^{(k)}(\mathbf{x})$  correspond à l'indice de confiance  $c_{i1D}$  du pixel  $\mathbf{x}$  de l'échelle  $k$ , et  $N$  est le nombre total d'échelles à savoir 12 ici.

La Figure 6.6 propose l'illustration d'une telle information pour les deux exemples d'images sonar présentées dans la section précédente.

A l'instar du descripteur d'homogénéité précédemment défini, on note tout d'abord que le descripteur  $A_1$  conserve la dynamique des dimensions intrinsèques continues, à savoir une information de confiance dans un intervalle  $[0, 1]$ . Sur les deux images présentées, on constate alors que les différents champs de rides de sable sont mis en évidence par de fortes valeurs de ce descripteur, témoignant d'une forte confiance en l'anisotropie de ces zones. De plus, on remarque que son caractère local procure lui aussi une localisation précise de ces régions et en particulier de leurs frontières.

Cependant sur l'image (a) de cette Figure, nous pouvons noter que les roches au centre et en bas de l'image répondent également à ce descripteur. Ce phénomène a déjà été abordé au paragraphe 5.4.4, où l'exemple d'une image sonar illustrant une zone rocheuse a été analysée en termes de dimensions intrinsèques continues multi-échelles. Nous y avons vu que les filtres utilisés dans l'analyse multi-échelle provoquent des lissages successifs. Les extrémités des roches s'arrondissent au fur et à mesure que l'on monte dans les échelles, ce qui résulte en une augmentation des indices de confiance  $c_{i1D}$ .

Ainsi, même si ces indices sont plus faibles que pour des rides de sable très directives, elles diminuent la distinction entre les zones anisotropes et rocheuses. Il serait donc nécessaire d'affiner la définition du descripteur d'anisotropie. Pour cela, nous devons mieux définir ce que nous sous-entendons par le terme "ride de sable", en abordant succinctement le domaine de la sédimentologie.

### 6.2.3.2 Etude sédimentologique des rides de sable

Comme le décrit G. Nichols dans [6], les rides de sable sont des structures formées en grande partie sous l'influence des courants sous-marins. Sous l'effet de ces courants fluides, du matériel (généralement de petite dimension comme des grains de sable) est alors transporté et participe à la construction d'amas locaux sur le fond. Sommairement, ces amas déforment localement le flux du courant. Ces déformations entraînent alors la création de motifs ondulés que l'on peut observer sur les images sonar.

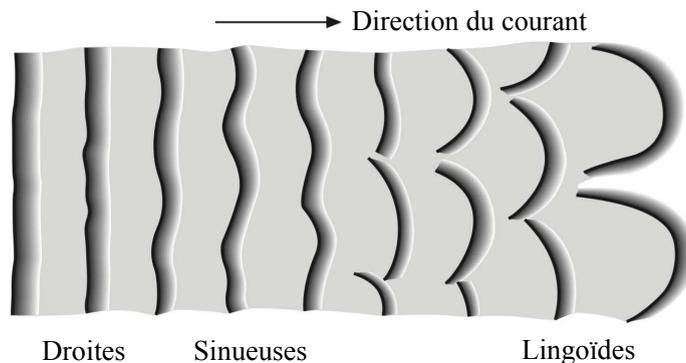


FIGURE 6.7: Schéma des différentes formes de rides possibles dans une image sonar (extrait et adapté de [6]).

Ainsi, le caractère directif (ou anisotrope) de ces rides de sable est directement reliée aux propriétés du courant. Ces dernières entraînent une large variété de formes possibles pour les rides de sable sous-marines. Ces rides peuvent alors posséder des crêtes relativement droites, sinueuses, voire sous forme d'arcades isolées (appelées forme lingoïde) comme le montre la Figure 6.7. Plus le courant sera intense et long, plus les rides de sable deviendront lingoïdes.

Cependant selon G. Nichols, les structures directives que l'on peut retrouver sur les fonds marins ne sont pas uniquement composées par les rides de sable. En effet, il mentionne également le cas des dunes sous-marines. Néanmoins, il souligne que leurs formations hydrodynamiques restent distinctes et qu'il ne faut pas les confondre.

Pour le démontrer, J.D. Collinson [7] propose d'étudier la répartition des longueurs d'onde et des hauteurs pour l'ensemble des structures directives des fonds marins. Cela résulte alors en deux graphes présentés sur la Figure 6.8. On y observe alors clairement deux modes qui ne se recouvrent pas. Ainsi l'auteur en conclut que ces structures ne sont pas issues d'un même processus de formation et qu'il n'existe donc pas de continuum entre les rides et les dunes.

Un autre fait remarquable peut ressortir de ces deux graphiques. En effet, on observe que les rides de sable sous-marines ont des longueurs d'onde allant jusqu'à un ordre de grandeur du

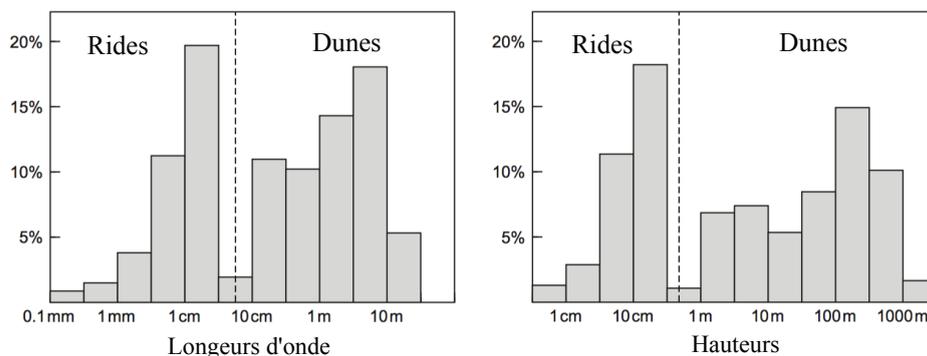


FIGURE 6.8: Répartitions des longueurs d'onde et des hauteurs pour les structures directives du fond marin (extraite et adaptée de [7]).

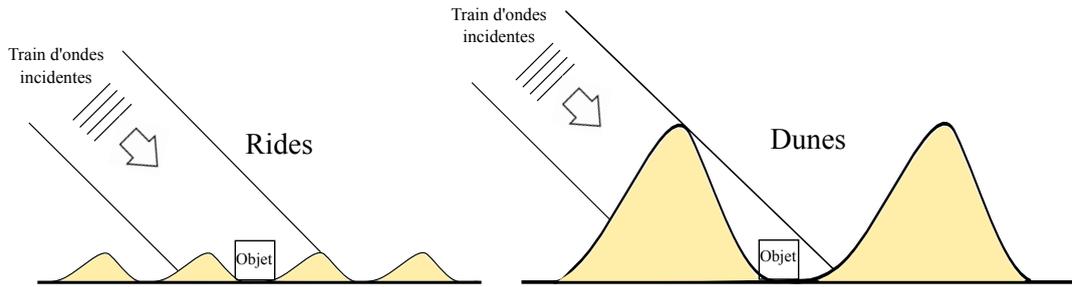


FIGURE 6.9: Comparaison entre un objet présent dans un champ de rides de sable et dans une zone de dunes sous-marines.

mètre, pour des hauteurs de l'ordre de la dizaine de centimètres. Cependant, les dunes sous-marines peuvent quant à elles culminer à des hauteurs proches de la dizaine de mètres pour des longueurs d'onde d'une centaine de mètres. Ainsi une mine sous-marine, dont la hauteur typique est de l'ordre de quelques dizaines de centimètres, sera facilement cachée par une dune sous-marine si cette dernière se situe entre l'objet et le sondeur. Mais dans un champ de rides de sable, sa hauteur sera potentiellement suffisante pour permettre de l'apercevoir sur l'image comme le montre le schéma proposé à la Figure 6.9.

Puisque les dunes de sable sous-marines peuvent facilement cacher une mine et rendre la phase de détection impossible, nous considérons qu'il est plus raisonnable de caractériser les dunes comme des structures complexes du fond marin. Nous proposons ainsi de les exclure des zones anisotropes que nous cherchons à mettre en évidence. De ce fait, cette hypothèse va permettre d'utiliser les graphiques présentés à la Figure 6.8 pour introduire une connaissance sédimentaire sur les rides de sable, et ainsi de fixer une limite en termes de longueurs d'onde et de dimensions pour les structures sous-marines que nous voulons mettre en évidence via le coefficient d'anisotropie.

### 6.2.3.3 Intégration de données sédimentologiques dans la recherche d'anisotropie

Nous venons de voir que les rides de sable sont des structures sédimentologiques dont les longueurs d'onde n'excèdent pas l'ordre du mètre. Puisque notre analyse multi-échelle de l'image sonar permet d'analyser des détails qui vont de 0.25 à 2.5 m, nous considérerons alors qu'au dernier niveau de détail les rides de sable doivent être totalement filtrées par le filtre passe-bande DoP. Nous avons pu constater au paragraphe 5.4.3 (cf. Figure 5.12) que cette suppression total des rides de sable a pour effet de rendre cette région du fond constante et homogène sur l'image filtrée. Par conséquent au niveau de détail le plus grossier, chaque pixel d'une telle région voit son indice de confiance en la dimension intrinsèque de 0 être très important (*i.e.*  $c_{i0D}^{(N)}$  proche de 1). Par conséquent, nous proposons de modifier l'expression du descripteur  $A_1$  afin de confirmer que les pixels qu'il considère comme appartenant à des zones anisotropes soient bien dus à des rides de sable. Pour cela, il suffit de multiplier  $A_1$  par l'indice

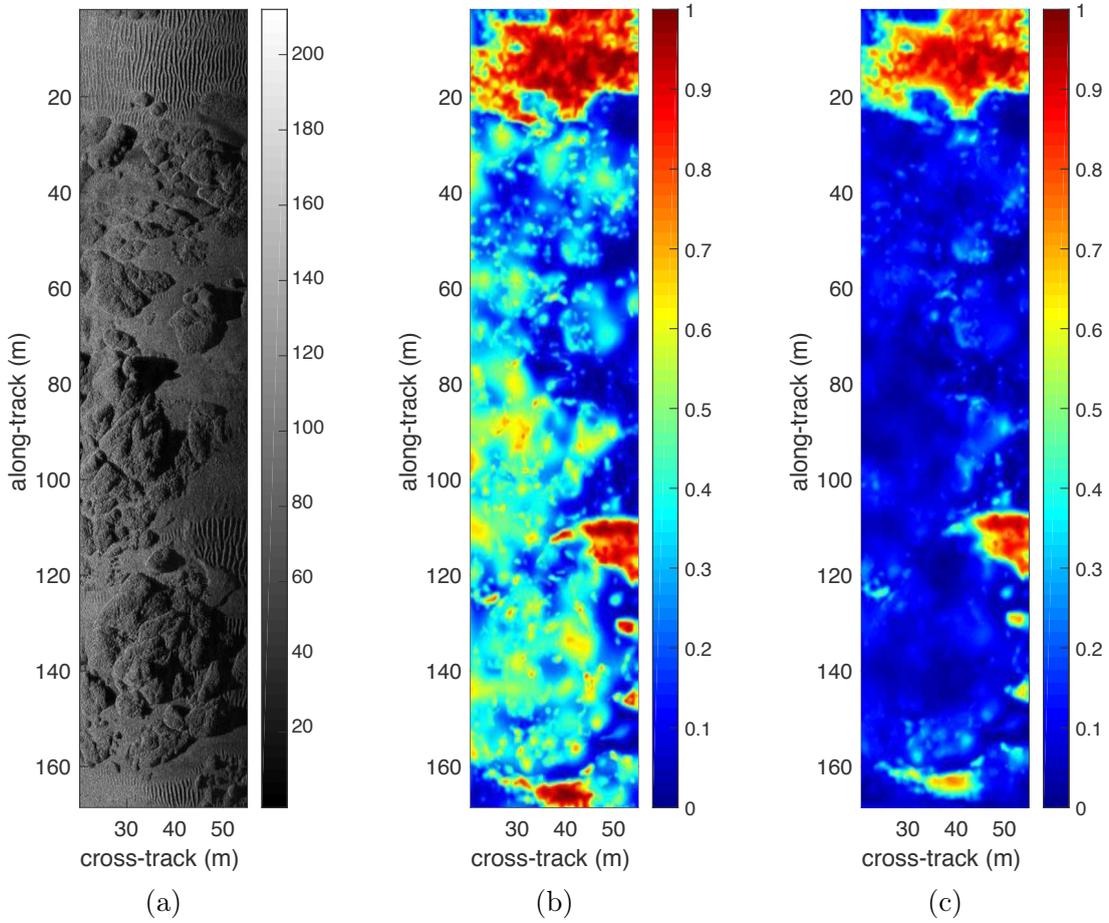


FIGURE 6.10: Exemple d'une image sonar (a) et des deux cartes d'anisotropie  $A_1$  (b) et  $A_2$  (c) correspondantes.

$c_{i0D}$  au niveau de détail le plus grossier  $N$  en chaque pixel de l'image :

$$A_2(\mathbf{x}) = A_1(\mathbf{x}) \times c_{i0D}^{(N)}(\mathbf{x}) = \max_{k \in \llbracket 1; N \rrbracket} \left( c_{i1D}^{(k)}(\mathbf{x}) \right) \times c_{i0D}^{(N)}(\mathbf{x}). \quad (6.5)$$

Sur cette expression, on constate alors que si un pixel appartient à une zone totalement filtrée alors le descripteur  $A_1$  sera simplement multiplié par 1. Au contraire, sa valeur sera réduite si  $c_{i0D}^{(N)}$  est faible.

La Figure 6.10 présente pour une même image sonar (a) les deux cartes d'anisotropie proposées  $A_1$  (b) et  $A_2$  (c). L'image sonar présente trois champs de rides de sable (en haut, dans la moitié inférieure droite et en bas), ainsi que des ensembles rocheux présents sur la majorité de l'image. On remarque rapidement que ces dernières biaisent fortement le résultat d'anisotropie proposé par le descripteur  $A_1$ . En effet, bien que les champs de rides de sable fournissent de plus forts indices de confiance, l'anisotropie des roches est tout de même proche de 0.5. Le contraste entre les rides de sable et les roches n'est alors pas satisfaisant. Le descripteur  $A_2$  propose quant à lui une information beaucoup plus proche de ce qu'on attend d'un estimateur d'anisotropie

puisque l'influence de la roche a été fortement réduite. Les roches étant des structures sous-marines dont le contenu fréquentiel est plus varié et moins localisé que celui des rides de sable, elles n'ont pas été totalement filtrées au niveau de détail le plus grossier ici, ce qui permet de supprimer leurs influences.

On note tout de même que l'utilisation d'une information provenant d'un niveau de détail grossier dans la définition de  $A_2$  (*i.e.*  $c_{i0D}^{(N)}$  ici) a pour conséquence de dégrader la localisation des frontières des zones anisotropes. Ces dernières sont alors plus étalées dans la carte  $A_2$  que dans  $A_1$ . De ce fait les zones anisotropes de faible étendue, comme par exemple le champ de rides de sable en bas de l'image, sont bien moins mis en évidence.

Au final, une analyse multi-échelle des dimensions intrinsèques continues couplée à une étude sédimentologique des rides de sable nous a permis de dresser une carte d'anisotropie. Cette information prend la forme d'un indice de confiance sur la présence de rides de sable dans une image sonar. Cependant, au début de cette section, nous avons fixé comme objectif d'extraire des descripteurs du fonds complémentaires entre eux. Or, ici aucun lien n'a encore été réalisé avec le descripteur d'homogénéité défini précédemment.

#### 6.2.3.4 Carte d'anisotropie

Puisque notre objectif initial est de définir trois descripteurs d'homogénéité, d'anisotropie et de complexité complémentaires, il est nécessaire de trouver un moyen de relier la carte précédente  $A_2$  à la carte d'homogénéité  $H$ . Pour ce faire, nous proposons de passer par la carte d'inhomogénéité  $\overline{H}$ , complémentaire de la carte d'homogénéité, et de la pondérer par la carte  $A_2$ . De cette manière, on définit une nouvelle carte d'anisotropie  $A$  qui représente le degré de confiance qu'un pixel appartienne à la fois à une zone inhomogène et à une zone anisotrope :

$$A(\mathbf{x}) = \overline{H}(\mathbf{x})A_2(\mathbf{x}). \quad (6.6)$$

La Figure 6.11 propose d'illustrer la différence entre l'ancien descripteur  $A_2$  et le nouveau  $A$  pour une image sonar issues de la base de donnée DGA. Cette dernière est essentiellement constituée de nombreux champs de rides de sable et d'une zone rocheuse dans la moitié inférieure droite de celle-ci. Les résultats obtenus montrent que sous l'effet de la multiplication, les indices de confiance de la carte d'anisotropie  $A$  sont légèrement diminués par rapport à la carte  $A_2$ . D'un autre côté, les frontières des régions semblent être mieux définies.

Finalement, nous avons donc réussi à définir un descripteur d'anisotropie qui répond aux exigences que nous nous étions fixés au début de ce chapitre, à savoir un descripteur sous la forme d'un indice de confiance en chaque pixel de l'image (*i.e.* une carte) et qui est en plus en lien avec le descripteur d'homogénéité. Nous allons voir maintenant que c'est ce lien qui

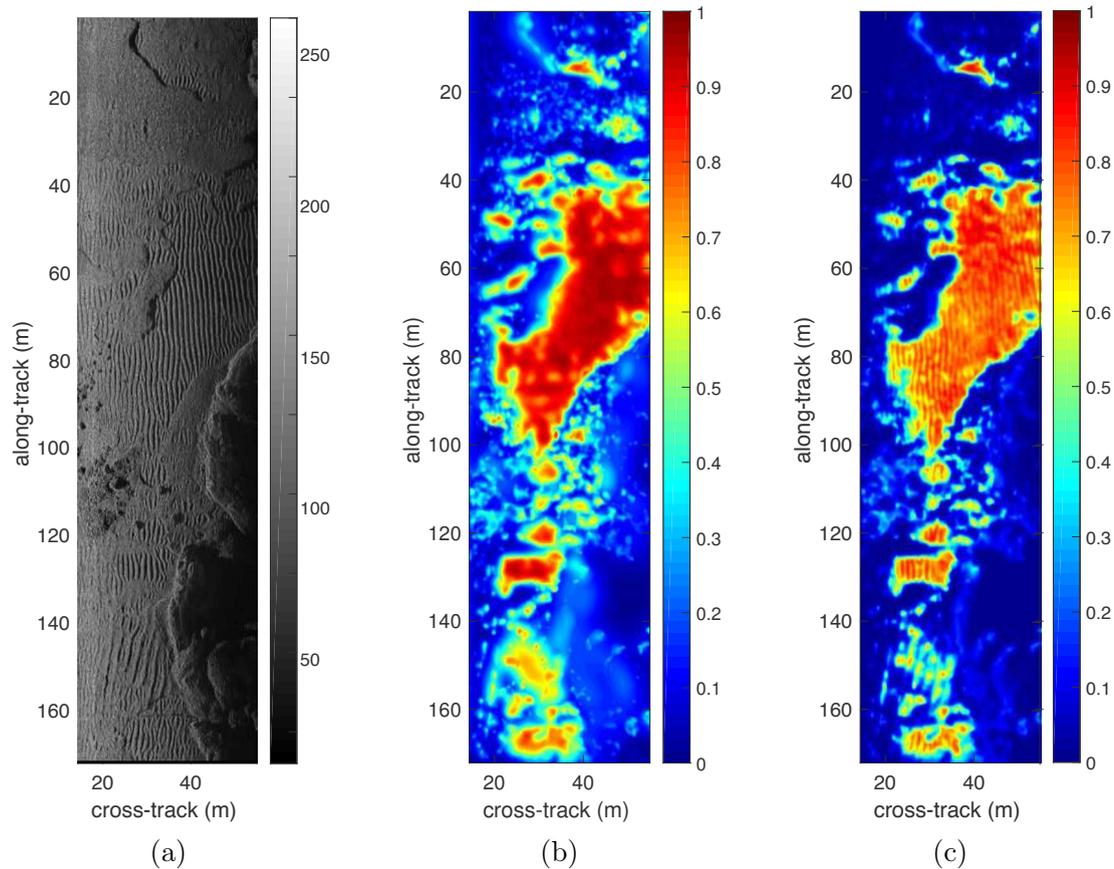


FIGURE 6.11: Exemple d'une image sonar (a) et des deux cartes d'anisotropie  $A_2$  (b) et  $A$  (c) correspondantes.

va nous permettre de définir le troisième et dernier descripteur représentant la complexité de l'image sonar.

#### 6.2.4 Complexité des fonds

Les textures que l'on peut considérer comme complexes sont d'une telle variabilité qu'il est impossible de définir un modèle général, ou d'étudier tous les cas possibles qui peuvent intervenir. Or, jusqu'à présent nous avons réussi à définir deux cartes qui représentent les degrés de confiance quant à l'homogénéité et à l'anisotropie du fond marin. Puisque ces deux cartes ont été spécialement définies de manière à ce qu'elles soient liées l'une à l'autre, nous allons pouvoir déterminer un indice de confiance sur la complexité de l'image sans avoir à définir mathématiquement la complexité d'une image. En effet, il suffit de voir le descripteur de complexité comme la confiance sur le fait qu'un pixel appartienne à une zone inhomogène sans pour autant appartenir à une zone anisotrope. Si l'on reprend la formulation que nous avons vu à l'équation 6.6, nous pouvons définir la carte de complexité  $C$  comme étant :

$$C(\mathbf{x}) = \overline{H}(\mathbf{x}) (1 - A_2(\mathbf{x})). \quad (6.7)$$

Une telle opération revient alors à diviser la carte d'inhomogénéité en deux cartes complémentaires puisqu'on a alors  $C(\mathbf{x}) + A(\mathbf{x}) = \overline{H}(\mathbf{x})$  pour tout pixel  $\mathbf{x}$  de l'image.

Par conséquent, les 3 cartes d'homogénéité  $H$ , d'anisotropie  $A$  et de complexité  $C$  sont également complémentaires entre elles puisque on vérifie aisément la relation suivante :

$$H(\mathbf{x}) + A(\mathbf{x}) + C(\mathbf{x}) = 1. \quad (6.8)$$

De cette manière, nous avons donc réussi à extraire d'une image, trois descripteurs qui représentent des indices de confiance complémentaires quant à l'homogénéité, l'anisotropie et la complexité du fond marin. Nous allons voir dans le paragraphe suivant plusieurs illustrations de ces trois cartes.

### 6.2.5 Résultats

Dans ce paragraphe, nous allons illustrer les trois descripteurs que nous avons défini à travers cette section. Puisque nous avons à disposition trois bases de données distinctes issues de différents systèmes sonar, nous proposons un exemple pour chacune d'elles afin d'étudier en parallèle la robustesse de l'approche vis-à-vis des différents systèmes.

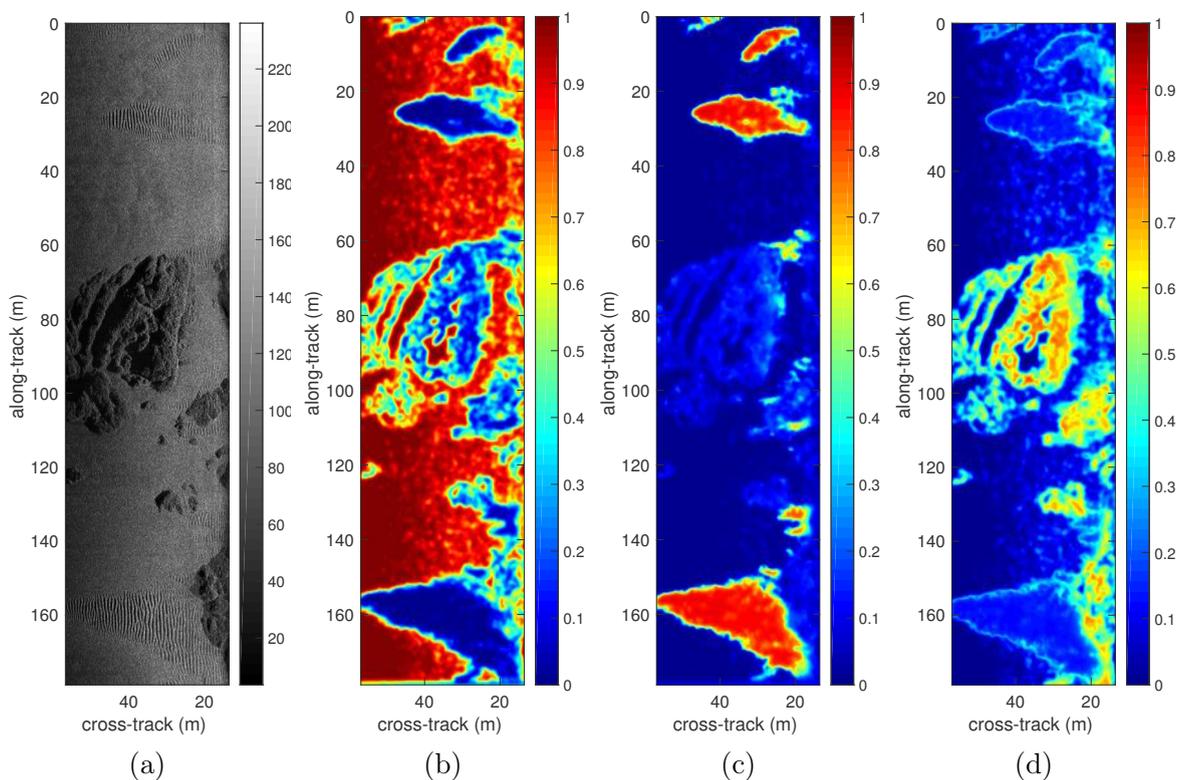


FIGURE 6.12: Exemple d'une image sonar issue de la base de données DGA (a) pour laquelle les cartes d'homogénéité (b), d'anisotropie (c) et de complexité sont extraites.

La Figure 6.12 illustre une première image sonar d'un système Klein 5000 (base de données DGA) et de ses trois cartes d'homogénéité, d'anisotropie et de complexité correspondantes. Sur l'image sonar initiale (a) on peut constater la présence de trois champs de rides de sable (deux en haut de l'image et un en bas) et de plusieurs amas rocheux (au centre de l'image et en début de portée à droite de l'image). Ces deux types de fonds particuliers sont séparés par des zones de sable fin et homogène. On constate que la carte d'homogénéité fournit des indices de confiance très forts sur les zones de sable fin où aucune structure sous-marine n'est présente. Par complémentarité des trois cartes, sur ces mêmes zones, les indices d'anisotropie et de complexité sont faibles. Les champs de rides sont quant à eux clairement mis en évidence dans la carte d'anisotropie avec des valeurs très importantes de l'indice de confiance (proche de 1). Au contraire, les zones de roches qui apparaissent sur l'image sonar par des textures plus complexes possèdent des indices de confiance plus importants dans la carte de complexité.

La Figure 6.13 propose également ces mêmes informations mais cette fois pour un autre système sonar latéral, à savoir un Marine Sonic (base de données DSTO). On y remarque que l'indice d'homogénéité est très élevé sur les zones non-texturées, alors que le champ de rides de sable est quant à lui fortement anisotrope. La complexité de l'image s'avère être plus forte en bout de portée (*i.e.* à droite de l'image) pour ce type de capteur, ce qui est sûrement dû, encore une fois, à un niveau de bruit plus intense et aux effets de rasance des ondes acoustiques. En début de portée, on remarque une bande noire verticale symptomatique des effets d'un lobe secondaire de l'antenne sonar. On observe qu'il y a dégradation de l'estimation de l'anisotropie sur l'extrémité du champ de rides de sable.

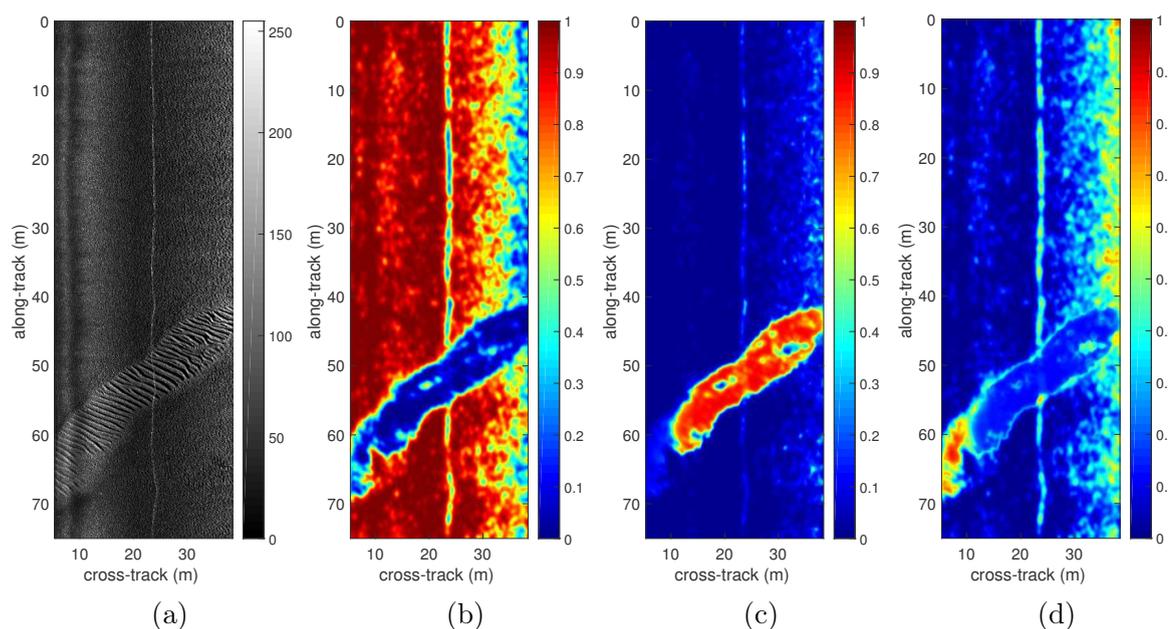


FIGURE 6.13: Exemple d'une image sonar issue de la base de données DSTO (a) pour laquelle les cartes d'homogénéité (b), d'anisotropie (c) et de complexité sont extraites.

Finalement, la Figure 6.14 propose des résultats pour une image sonar SAS de la base de données du CMRE. On y observe que le fond homogène est clairement identifiable via de forts indices de confiance dans la carte d'homogénéité et que la zone de rides de sable est identifiable dans la carte d'anisotropie. On y observe également une autre caractéristique majeure de notre approche concernant les frontières des régions, et particulièrement celles des rides de sable. En effet, on remarque que les frontières des rides de sable sont ici considérées comme des zones complexes. Cela provient du fait que l'extrémité d'une ride de sable n'est pas un motif 1D, mais bien 2D puisqu'elles se rapprochent d'un motif en forme de coin. Par ailleurs de manière générale, la définition des frontières entre deux régions texturées différentes d'une image est un problème récurrent dans le domaine de la classification. Nous avons pu le constater dans le paragraphe 4.6.3 lorsque nous avons réalisé la classification d'images sonar par le biais des informations monogènes. Nous avons alors constaté que la classification d'un bloc de l'image contenant deux types de fond (*i.e.* contenant une frontière) fournissait une mauvaise estimation du label de la zone. Ici, le fait que les frontières entre les différents types de fond soient définies comme des zones complexes évite toute confusion ou erreur sur la nature du fond. De plus, l'adjectif de "complexe" qualifie bien ce que représente une frontière puisque la détection qu'une cible y sera très difficile comme le mentionne D.P. Williams dans [3].

A travers ces différents exemples, nous pouvons affirmer que les trois descripteurs que nous avons définis ici permettent de mettre en évidence, de manière précise, les trois types de fonds marins que nous cherchions à décrire au début de cette étude. Leur originalité est basée sur la complémentarité qui les relie, mais également sur la notion d'indice de confiance qu'ils fournissent. Le mélange de ces deux caractéristiques permet d'obtenir trois informations significatives et comparables entres-elles.

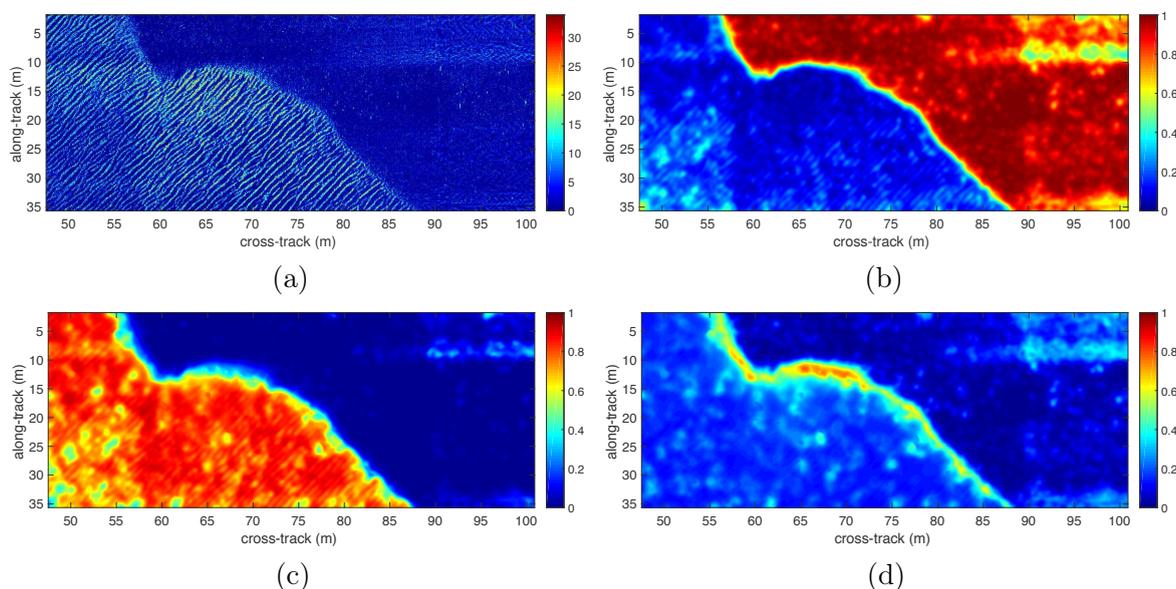


FIGURE 6.14: Exemple d'une image sonar issue de la base de données CMRE (a) pour laquelle les cartes d'homogénéité (b), d'anisotropie (c) et de complexité sont extraites.

De plus, puisqu'aucune modification des paramètres de l'algorithme n'a été réalisée entre ces trois exemples, on peut en conclure que l'approche proposée à travers ce chapitre s'adapte parfaitement aux différentes bases de données. La seule modification nécessaire est liée à la dimension des pixels de l'image (*i.e.* à la résolution). Afin de démontrer l'apport de ces descripteurs, nous proposons maintenant de les comparer avec une méthode existante.

## 6.2.6 Comparaison avec une méthode existante

Au paragraphe 3.3.2.1, nous avons vu que la littérature propose également trois algorithmes visant à extraire des informations d'anisotropie et de complexité dans les images sonar. Nous proposons ici de comparer les résultats de notre approche par rapport au dernier algorithme publié. Celui-ci a été proposé par O. Daniell dans [4]. Dans un premier temps, avant toute comparaison qualitative ou quantitative, nous proposons de comparer ces deux approches au niveau algorithmique. C'est-à-dire au niveau des méthodes et des paramètres utilisés.

### 6.2.6.1 Comparaison algorithmique

Comme nous l'avons présenté au paragraphe 3.3.2.1, la méthode de O. Daniell se base sur une analyse en ondelettes complexes par un arbre dual. Ce cadre permet d'implémenter une décomposition en ondelettes pour 6 orientations (nombre fixé par la théorie) à chaque niveau de résolution souhaité. Si l'on note  $\psi_{s,k}^{\mathbb{C}}(x, y)$  le coefficient de l'ondelette complexe au pixel de coordonnées  $(x, y)$  pour une orientation d'indice  $k \in \{1, \dots, 6\}$  et au niveau de résolution d'indice  $s \in \{1, \dots, m\}$ , O. Daniell propose d'étudier l'énergie de ces coefficients en ondelettes  $E_{s,k}(x, y) = |\psi_{s,k}^{\mathbb{C}}(x, y)|^2$  mais en la normalisant à l'aide de deux paramètres  $\alpha$  et  $\beta$  :

$$N_{s,k}(x, y) = \alpha^s E_{s,k}(x, y)^\beta. \quad (6.9)$$

C'est en étudiant les moyennes et les écarts-types de ces réponses énergétiques normalisées  $N_{s,k}$  suivant les orientations et les niveaux de résolution que l'auteur propose une définition des coefficients d'anisotropie et de complexité que nous avons vu aux équations (3.3) et (3.4). Ainsi, cette méthode de caractérisation des fonds marins nécessite que l'utilisation de trois paramètres, à savoir le nombre  $m$  de niveaux de résolution et les paramètres de normalisation  $\alpha$  et  $\beta$  que l'auteur fixe expérimentalement à 0.25 et 0.75 respectivement. Il est à noter qu'aucune justification n'a été fournie par l'auteur pour ces valeurs et en particulier concernant leurs validités pour différentes bases de données ou systèmes sonars.

Notre approche se base quant à elle sur une représentation monogène de l'image sonar qui est isotrope (contrairement aux 6 orientations d'analyse de la décomposition en ondelettes complexes de O. Daniell). Cette représentation étant multi-échelle, nous avons du définir cette décomposition. Nous avons alors proposé une certaine manière de faire au paragraphe 4.5.3.2

qui nécessite de fixer le premier et le dernier niveau de détail (0.25 m et 2.5 m respectivement) et le nombre de niveaux décompositions entre les deux (au nombre de 12 ici, cf. paramétrage de  $k$ ). Dans cette analyse multi-échelle et pour accéder à des niveaux de détail particuliers, il est également nécessaire d'avoir une connaissance sur la dimension des pixels dans l'image. Ensuite, nous avons proposé d'utiliser les informations issues de ce signal monogène multi-échelle pour estimer les dimensions intrinsèques continues, puis les descripteurs d'homogénéité, d'anisotropie et de complexité, le tout sans l'utilisation d'autres paramètres. Ainsi, tout comme l'approche de O. Daniell, notre proposition n'utilise que trois paramètres internes au traitement (définissant l'analyse multi-échelle), mais également une connaissance extérieure, propre à l'image sonar étudiée, à savoir sa résolution. Cependant, les trois paramètres internes sont fixés une fois pour toute et indépendamment de l'image traitée. Ainsi, puisque les paramètres à fixer dans notre approche ne sont jamais spécifiques à une image sonar ou à une base de données, nous pouvons considérer que c'est cet avantage qui permet la robustesse de nos algorithmes vis-à-vis des différentes bases de données (ce que nous avons pu constater jusqu'à présent).

Maintenant que nous avons comparé ces deux algorithmes au point de vue de l'implémentation, nous allons nous intéresser à leurs différences qualitatives, c'est-à-dire une analyse visuelle des résultats qu'ils fournissent.

### 6.2.6.2 Comparaison qualitative

La Figure 6.15 propose une première illustration d'une image sonar déjà rencontrée dans le paragraphe précédent (issue de la base de données du DSTO) qui représente un champ de rides de sable traversant l'image. Alors que la première ligne de cette figure présente les cartes d'anisotropie et de complexité issues de l'algorithme de O. Daniell, la seconde présente ceux de notre approche. Notons ici que nous ne présentons pas notre carte d'homogénéité puisque la littérature n'en propose pas actuellement.

Dans les deux ensembles de résultats, on peut remarquer visuellement que le champ de rides de sable est bel et bien mis en évidence dans les deux cartes d'anisotropies, en (b) et en (d). Cependant, on y constate également certaines différences notamment dans le contraste qu'elles fournissent avec les régions frontalières, dans la précision de ces frontières, mais également dans l'homogénéité du descripteur d'anisotropie sur la zone de rides de sable. Ces distinctions semblent toutes être à l'avantage de notre approche (d).

La dynamique du coefficient d'anisotropie est également différente, puisque nous proposons un indice de confiance là où O. Daniell propose un descripteur sans dynamique connue a priori. L'avantage d'un indice de confiance est sa facilité d'interprétation et en particulier lorsque l'on étudie une image donnée. On comprend alors que plus le descripteur sera proche de 1 plus il sera représentatif du type de fond présent dans l'image. En revanche, puisque les

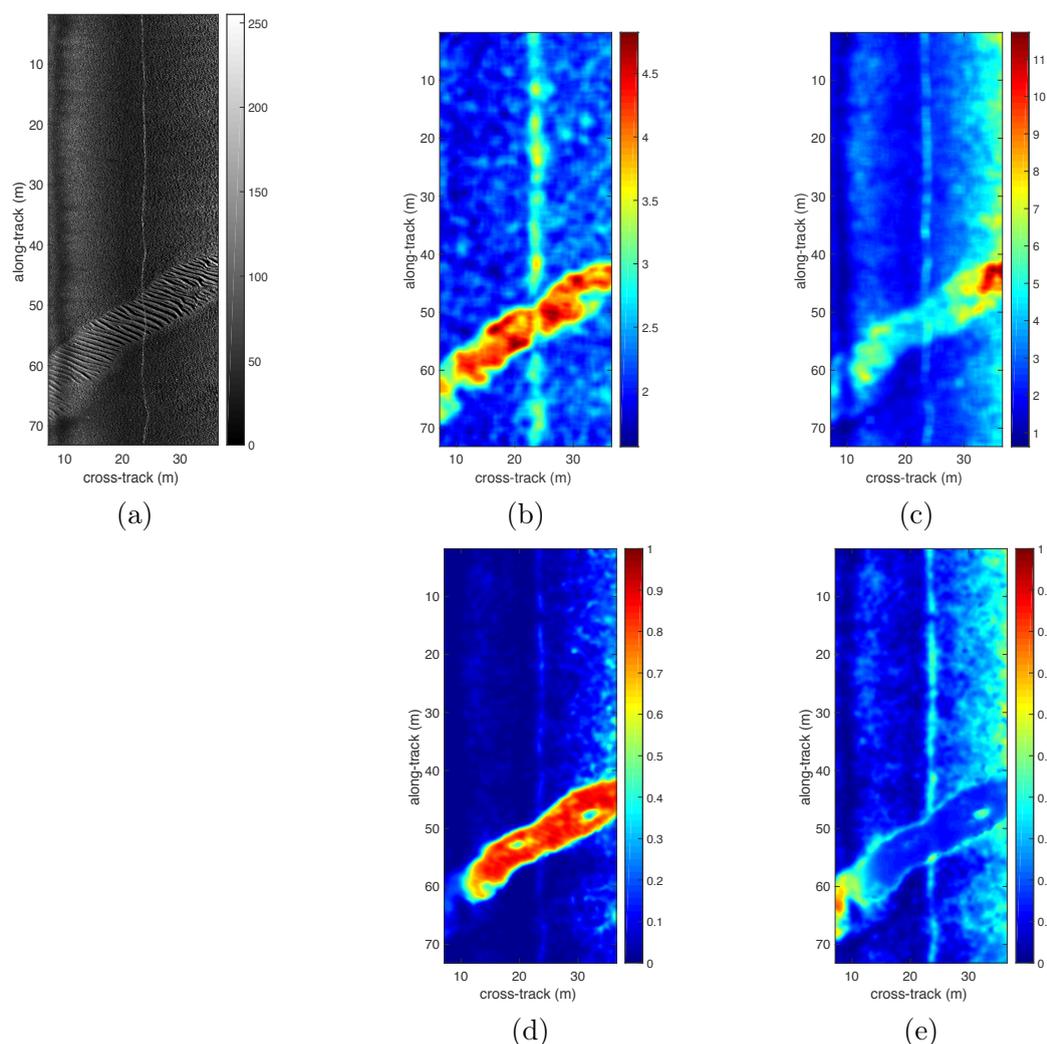


FIGURE 6.15: Exemple d'une image sonar (a) issues de la base DSTO. Les cartes d'anisotropie et de complexité sont illustrées pour l'algorithme de O. Daniell sur la première ligne, en (b) et (c) respectivement, et pour notre approche en (d) et (e).

coefficients de O. Daniell ont des valeurs non-bornées, lorsque l'on observe une image, il est difficile de se faire une idée de la signification d'une valeur d'anisotropie ou de complexité donnée. Pour permettre leurs interprétations, il est alors nécessaire d'étudier un ensemble d'exemples pour connaître leurs valeurs typiques (par exemple du coefficient d'anisotropie sur des rides de sable ou de complexité sur les roches).

Une autre distinction concerne le motif provoqué par un reflet de la surface (que l'on retrouve au milieu de l'image par un liseré vertical blanc). Alors que ce dernier ne peut ni être associé à une zone homogène de l'image ni à une zone de rides de sable, il serait opportun qu'il soit décrit majoritairement par le coefficient de complexité. En particulier, si une cible se trouve à la même position que ce reflet dans l'image, il y aura une forte probabilité qu'elle ne soit pas détectée. Dans ce cas, on constate que notre approche permet bel et bien de mettre en évidence ce reflet comme une structure de l'image possédant une forte complexité et

une anisotropie faible. Ce résultat peut apparaître contradictoire puisque un tel motif semble directif et donc s'interpréter comme une structure 1D. Or, nous avons intégré au descripteur d'anisotropie des connaissances sédimentaires afin qu'il mette uniquement en évidence les rides de sable. Pour ce faire, rappelons que nous avons forcé ce descripteur à analyser les structures de l'image qui sont totalement filtrées au niveau de détail le plus grossier de notre analyse multi-échelle. Ce reflet de la surface étant un motif qui s'apparente localement à un dirac (ou à une fonction créneau de faible dimension), il possède des composantes fréquentielles depuis les basses fréquences jusqu'aux très hautes fréquences (la transformée de Fourier d'un dirac est une fonction constante de valeur 1). Il ne sera donc pas totalement filtré au dernier niveau de détail de notre analyse multi-échelle, et ne sera donc pas présent dans la carte d'anisotropie. Au contraire, dans l'approche de O. Daniell, ce motif apparaît de manière plus claire sur la carte d'anisotropie plutôt que dans la carte de complexité ce qui peut engendrer de la confusion.

Néanmoins, l'approche de O. Daniell semble moins sensible aux effets des lobes secondaires en début de portée. En effet, en bas à gauche de l'image, on remarque que son descripteur d'anisotropie réagit de manière plus importante que le nôtre. Dans cette zone, c'est notre descripteur de complexité qui y prédomine. A contrario, en fin de portée, c'est l'inverse qui se produit. Notre descripteur d'anisotropie y réagit de manière plus forte là où son descripteur de complexité réagit de manière maximale.

Ces différences entre les deux algorithmes sont amplifiées sur le second exemple proposé à la Figure 6.16. Cette fois, les résultats d'anisotropie et de complexité sont proposés pour une image sonar issue d'une autre base de données (DGA). Par rapport à l'exemple précédent, aucune modification des deux algorithmes n'a été faite à l'exception de la dimension des pixels pour notre approche. Comme précédemment, on y remarque que les zones anisotropes sont définies de manière bien plus précises dans le cas de notre algorithme (d). De plus, on observe que le descripteur de complexité proposé par O. Daniell fournit ici des valeurs peu discriminantes entre les rides de sable et sur les zones rocheuses. L'interprétation de ces deux descripteurs est alors beaucoup plus difficile que les indices de confiance que nous proposons. Cette difficulté est encore une fois renforcée par la dynamique des descripteurs proposés par O. Daniell. En effet, on observe par exemple que les valeurs du coefficient de complexité sont bien différentes pour cette image que pour celle présentée à la Figure 6.15 précédente. Ainsi, suivant le système sonar utilisé, les résultats de la description du fond proposé par O. Daniell seront différents et devront faire l'objet d'une étude spécifique à la base de données traitée. Au contraire, notre approche fournit quant à elle des résultats robustes au changement de la base de données.

Afin de mieux comparer ces deux algorithmes, nous proposons d'étudier de manière plus précise leurs résultats dans les zones homogènes, anisotropes et complexes via une comparaison quantitative.

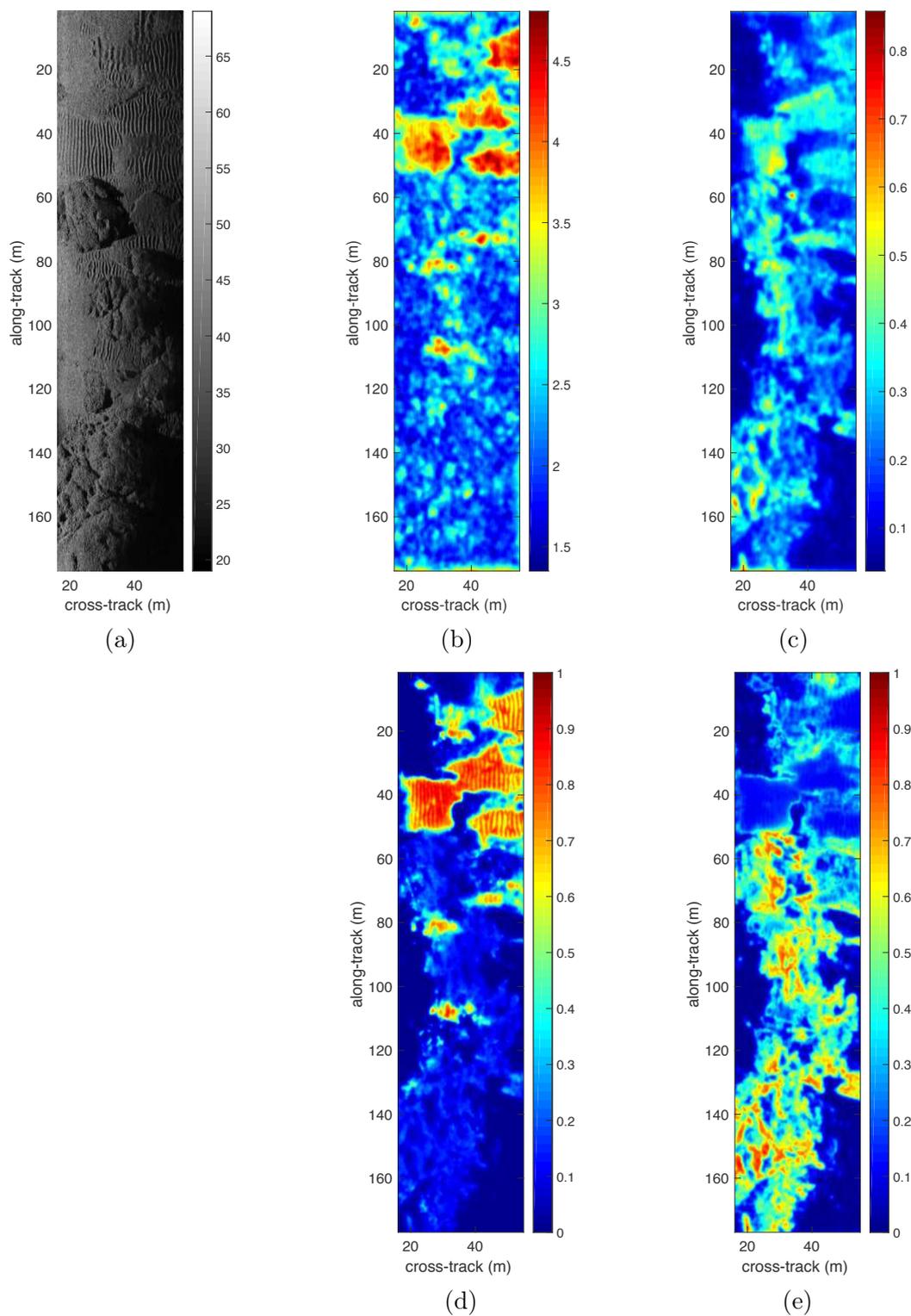


FIGURE 6.16: Exemple d'une image sonar (a) issues de la base DGA. Les cartes d'anisotropie et de complexité issues l'algorithme de O. Daniell sont illustrées sur la première ligne, en (b) et (c) respectivement, et celles issues de notre approche en (d) et (e).

### 6.2.6.3 Comparaison quantitative

Afin de fournir une comparaison quantitative, nous proposons de segmenter à la main une image sonar en zones homogènes, anisotropes et complexes. La Figure 6.17 illustre une telle opération où les zones colorées en vert représentent les fonds que l'on considère comme homogènes, les zones colorées en bleu représentent les fonds que l'on considère comme anisotropes, et finalement les zones colorées en rouge représentent les fonds que l'on considère comme complexes, à savoir les roches ici. Puisque nous n'avons pas encore étudié la problématique liée aux zones d'ombre dans l'image sonar (que nous traiterons au paragraphe 7.2.2), nous proposons de les exclure de notre analyse. Elles sont alors représentées par des zones en noir.

Pour chacune de ces trois zones, on trace sur la Figure 6.18 un histogramme 2D qui représente la répartition de leurs pixels en fonction de leurs valeurs d'anisotropie et de complexité estimées par O. Daniell. On remarque que les zones homogènes sont décrites par une anisotropie non-nulle, puisque la répartition moyenne de leurs pixels est concentrée autour d'une valeur d'anisotropie de 1 et de complexité proche de 0 sur la Figure 6.18.(a). On note

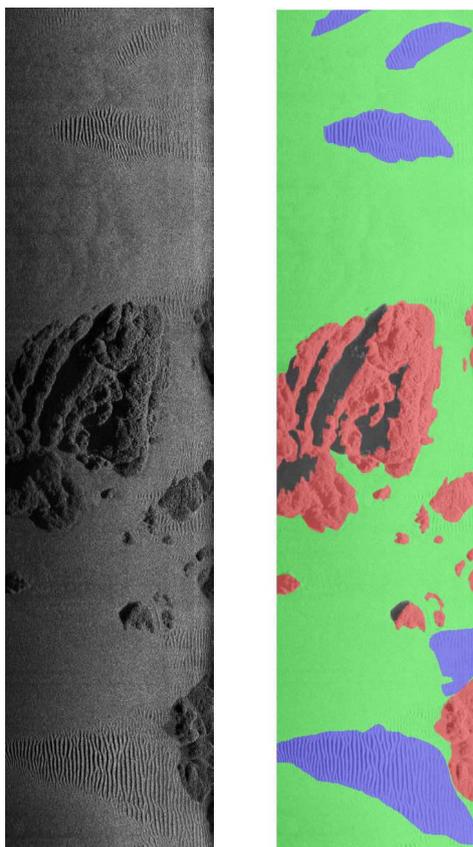


FIGURE 6.17: A gauche, une image sonar issues de la base de donnée DGA. A droite, sa segmentation manuelle en termes de zones homogènes (en vert), anisotropes (en bleu) et complexes (en rouge). Les grandes zones d'ombre ont également été segmentées et apparaissent sans couleur.

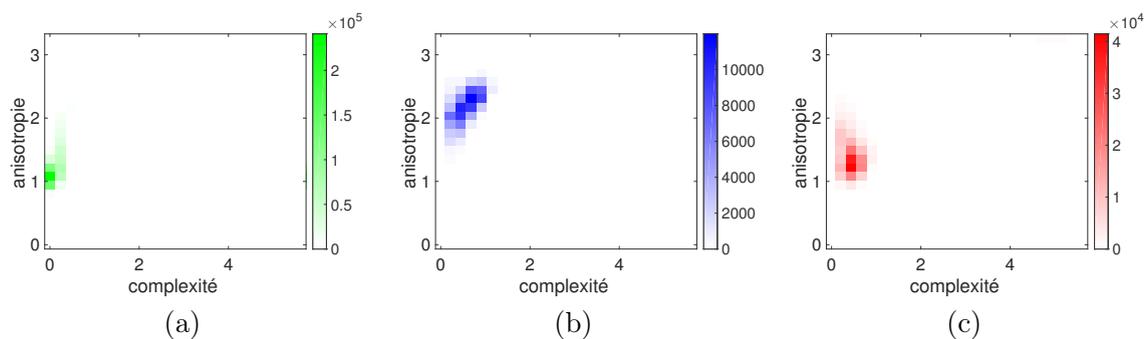


FIGURE 6.18: Histogrammes 2D représentant la répartition des pixels en fonction des descripteurs d'anisotropie et de complexité calculés par O. Daniell dans les zones homogènes (a), anisotropes (b) et complexes (c) de l'image sonar présentée à la Figure 6.17.

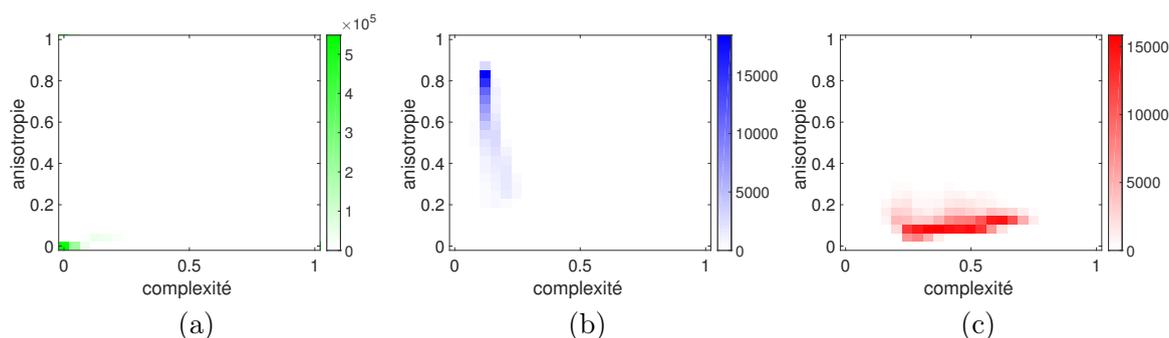


FIGURE 6.19: Histogrammes 2D représentant la répartition des pixels en fonction des descripteurs d'anisotropie et de complexité que nous proposons dans ces travaux pour des zones considérées comme homogènes (a), anisotropes (b) et complexes (c) de l'image sonar présentée à la Figure 6.17.

également que cette répartition est très proche de celle illustrée sur l'histogramme (c) qui représente la répartition des pixels des zones complexes. De ce fait, ces deux descripteurs ne semblent pas avoir un fort pouvoir de discrimination entre les zones homogènes et les zones complexes. Cependant, sur la Figure 6.18.(b), on observe tout de même que la répartition des pixels des zones anisotropes est plus caractéristique, avec de fortes valeurs du descripteur d'anisotropie. Mais, cette répartition fait intervenir des valeurs plus élevées de la complexité que les zones complexes elles-mêmes. Ce qui démontre que ces deux coefficients ne permettent pas une description efficace des trois types de fonds que nous souhaitons distinguer.

La Figure 6.19 propose exactement les mêmes histogrammes mais cette fois pour les descripteurs estimés via notre approche. On remarque cette fois que les répartitions des pixels sur ces trois histogrammes sont bien différentes. Alors que dans le cas des zones homogènes, ces pixels sont concentrés autour de valeurs d'anisotropie et de complexité proches de 0, les pixels des zones anisotropes possèdent de fortes valeurs d'anisotropies et des complexités plutôt faibles. Les pixels appartenant à des zones complexes ont quant à eux une répartition centrée autour de valeurs de complexité fortes et d'anisotropie faibles. Ainsi, ces différents histogrammes 2D montrent que notre définition des descripteurs d'anisotropie et de complexité

permettent de mieux discriminer les zones homogènes, anisotropes et complexes d'une image sonar, en comparaison de la méthode de O. Daniell. Par ailleurs, leurs analogies à des degrés de confiance favorisent leur signification et leur compréhension.

### 6.2.7 Liens avec l'ATR

Maintenant que nous avons à disposition trois indices de confiance qui décrivent localement la texture d'un fond marin sur l'image sonar, nous pouvons commencer à aborder la problématique de détection de mines sous-marines. Nous allons illustrer ici le fait que la détection d'une cible est fortement conditionnée à son environnement proche.

Pour ce faire, nous avons à disposition un des algorithmes de détection automatique les plus récents, à savoir celui proposé par D.P Williams dans [80]<sup>1</sup>. Comme nous l'avons mentionné à la section 3.2.1, ce détecteur nous semble être actuellement le plus abouti, et possède l'avantage d'être déjà implémenté à bord des AUV du CMRE. Nous proposons donc d'étudier ses performances en fonction des descripteurs que nous venons de définir. Pour ce faire, nous utiliserons les données SAS du CMRE sur lesquelles les positions des mines sont connues et pour lesquelles l'algorithme de détection a été spécialement étudié.

Autour de chacune des détections fournies par l'algorithme, nous proposons d'estimer la moyenne des descripteurs d'anisotropie et de complexité dans un voisinage de taille 5m x 5m. Afin que l'objet détecté n'affecte pas l'estimation, une fenêtre de garde est utilisée autour du motif écho-ombre comme cela est illustré sur la Figure 6.20. Sur cette dernière, le rectangle rouge représente le voisinage de la détection, tandis que le vert correspond à la fenêtre de garde. La dimension verticale de la fenêtre de garde est fixée à 1 mètre, alors que sa dimension horizontale dépend de la distance de détection dans l'image à cause du phénomène d'élongation de l'ombre. Ainsi, si l'on considère une cible dont la distance au sonar est de  $R$  (en mètres), de hauteur  $h$  (en mètres), et pour une altitude du sonar de  $H$  (en mètres), on peut estimer par l'intermédiaire du théorème de Thalès que la dimension horizontale  $s$  de son ombre sera de :

$$s = \frac{hR}{H - h}. \quad (6.10)$$

Lors du chapitre suivant nous verrons de plus amples détails sur les relations entre la position des cibles et la taille des ombres associées.

La Figure 6.21 montre la répartition des fausses alarmes (a) et des non détections (b) en fonction de la moyenne des descripteurs d'anisotropie et de complexité dans leurs voisinages. On peut alors constater qu'il existe deux foyers nets de fausses alarmes et de non détection, à savoir les zones de l'image ayant de fortes valeurs d'anisotropie (au-delà de 0.3), ainsi que celles

1. Nous profitons de ce passage pour remercier D.P Williams de nous avoir fourni cet algorithme de détection. Ce dernier nous a permis d'évaluer nos propres algorithmes, que ce soit dans ce chapitre pour la description de l'environnement marin, que dans le chapitre suivant lors d'une phase de détection

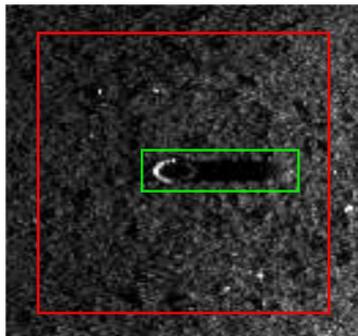


FIGURE 6.20: Illustration du voisinage autour d'une détection donnée (en rouge) et de sa fenêtre garde (en vert).

dont la complexité est supérieure à 0.25 ici. Par ailleurs, on peut également noter que ces fausses alarmes et non détections sont quasiment inexistantes sur les fonds fortement homogènes, qui se situent au niveau de l'origine de l'histogramme.

En conclusion, ces différentes illustrations montrent que nos descripteurs d'anisotropie et de complexité du fond marin sont bien reliés aux difficultés de détection de cibles sous-marines. De plus, nous avons pu constater que certains fonds sont prédisposés à engendrer des fausses alarmes voire à mettre en échec la phase de détection. Ainsi, pour réaliser une détection performante et d'obtenir de meilleurs taux de détections et de fausses alarmes, il nous semble nécessaire qu'une chaîne ATR tienne compte de l'environnement local de la mission.

Jusqu'à présent, nous avons proposé à travers ce chapitre une description de l'environnement marin par le biais de trois nouveaux descripteurs. Nous avons souligné l'intérêt de ces derniers que ce soit dans la description des fonds marins mais également pour les problématiques liées à l'ATR. Dans ce contexte, il peut être également intéressant, aussi bien pour un opérateur que pour un système autonome, d'avoir une classification des fonds marins en

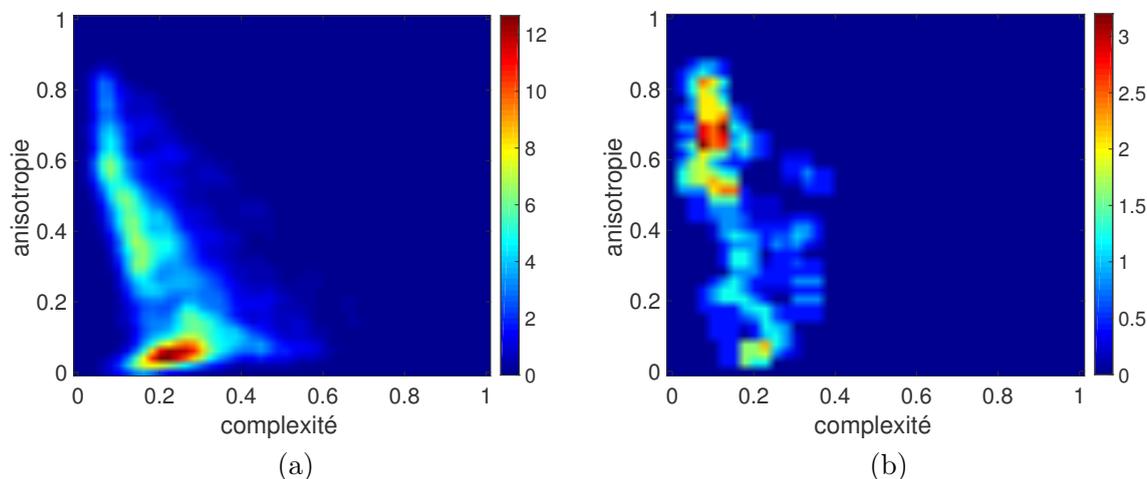


FIGURE 6.21: Répartition des fausses alarmes (a) et des manques (b) du détecteur de D.P. Williams en fonction des coefficients d'anisotropie et de complexité.

termes de zones homogènes, anisotropes ou complexes. Par classification, nous entendons ici une image sonar segmentée et labellisée avec une prise de décision dure sur le type de fond marin présent (à l'inverse des cartes "continues" que nous avons vu jusqu'à présent). Le paragraphe suivant propose d'aborder cette problématique de classification des fonds marins dans un cadre non-supervisé.

## 6.3 Classification non-supervisée des fonds marins

### 6.3.1 Introduction

Les descripteurs d'homogénéité, d'anisotropie et de complexité que nous proposons à travers ce chapitre sont reliés aux caractéristiques des trois types de fonds qui impactent les performances d'une chaîne ATR. Nous avons pu constater à la section précédente que ces descripteurs permettent une distinction entre les différents types de fonds marins qui nous intéressent. Nous proposons ici de tirer parti de la notion d'indice de confiance et de complémentarité qui les caractérisent pour réaliser une classification des images sonar en fonds homogènes, anisotropes et complexes. C'est-à-dire que nous allons prendre une décision binaire sur le fait qu'une zone de fond marin appartienne ou pas à une classe donnée (contrairement à ce que nous avons fait jusqu'à présent). Nous allons voir que cette opération de classification est rendue possible par le biais de considérations simples, qui évitent toute utilisation de base d'apprentissage ou de classifieur. En d'autres termes, nous pouvons dire que cette opération est non-supervisée. La Figure 6.22 propose le synoptique de la suite de notre chaîne de description des image sonar.

Puisque chacun des trois descripteurs que nous avons défini est relié à un type de fond particulier, leur complémentarité pourrait nous laisser penser qu'une simple sélection de la classe correspondante au descripteur le plus fort suffirait. Une telle stratégie permettrait de rendre la classification entièrement non-paramétrique. Or, selon nous, la classification d'une région de l'image sous le qualificatif "homogène" doit être réalisée avec précaution. En effet, le qualificatif d'homogène étant synonyme d'absence de structure dans l'image, il renvoie donc naturellement à l'absence d'objet potentiellement malveillant comme une mine sous-marine. Ainsi, comme nous allons le voir dans les paragraphes suivants, nous préférons étudier la classification des zones homogènes de manière spécifique d'une part, et de discriminer ensuite les zones restantes en fonds marins anisotropes et complexes.

### 6.3.2 Classification des fonds homogènes

Puisque les différentes cartes que nous avons définies dans ce chapitre sont basées sur des indices de confiance, il semble possible de réaliser une prise de décision via de simples seuillages.

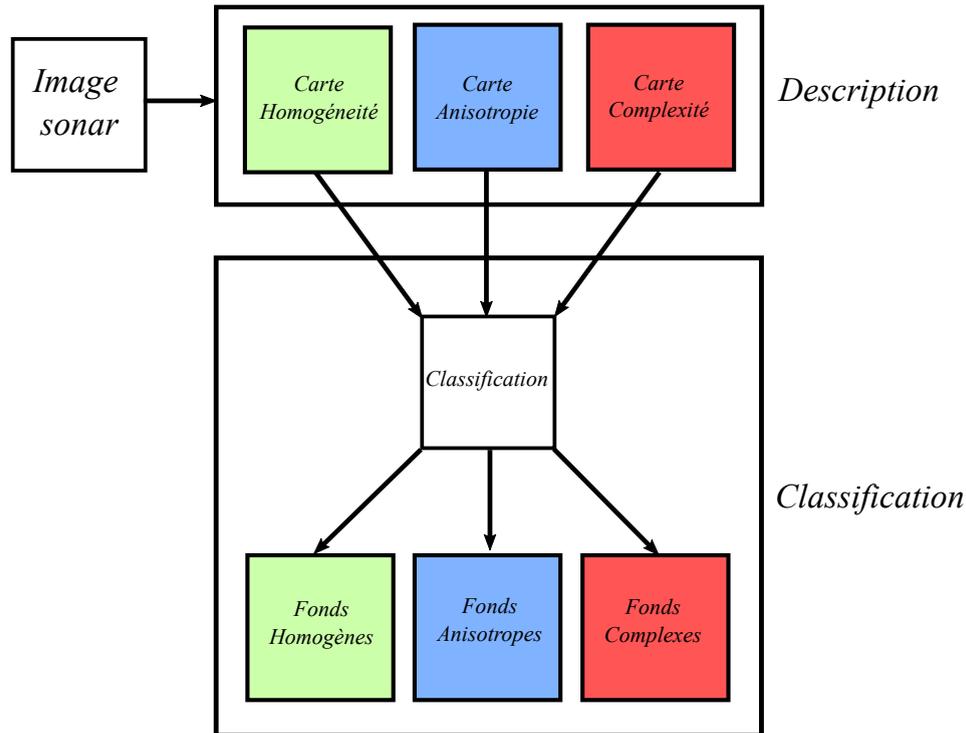


FIGURE 6.22: Synoptique de l'étape de classification.

### 6.3.2.1 Seuillage de la carte d'homogénéité

Au paragraphe 7.2.2, nous avons défini une carte d'homogénéité  $H$  qui se base sur la médiane des indices  $c_{i0D}$  à travers les niveaux de détail. Cette carte nous fournit un indice de confiance sur l'absence de structure dans l'image sonar. Ainsi, plus cette confiance est forte (*i.e.* proche de 1) et plus on peut considérer que le fond marin est homogène. Par conséquent, nous proposons une prise de décision sur l'homogénéité par le biais d'un simple seuillage :

$$\text{Homog}(\mathbf{x}) = H(\mathbf{x}) > T, \quad (6.11)$$

où le seuil  $T \in [0, 1]$  doit être fixé en fonction de la confiance vis-à-vis de l'homogénéité du fond marin que l'on désire.

La Figure 6.23 propose l'illustration d'une telle extraction pour quatre valeurs distinctes du seuil  $T$  pour une image sonar issue de la base de données du DSTO. Cette image sonar (a) est formée d'un fond de sable fin où un certain nombre d'échos y sont visibles, et en particulier celui d'une mine sous-marine dans la moitié supérieure (au milieu de la portée). On y observe que plus on augmente le seuil  $T$ , plus la sélectivité devient forte, et plus on réduit le nombre de zones homogènes de l'image.

On observe également que pour un seuil faible, un objet posé sur un fond homogène peut lui-même être considéré comme appartenant à la classe homogène, alors que pour un seuil élevé, il en est directement exclu. Cette observation est fondamentale si l'on désire intégrer une

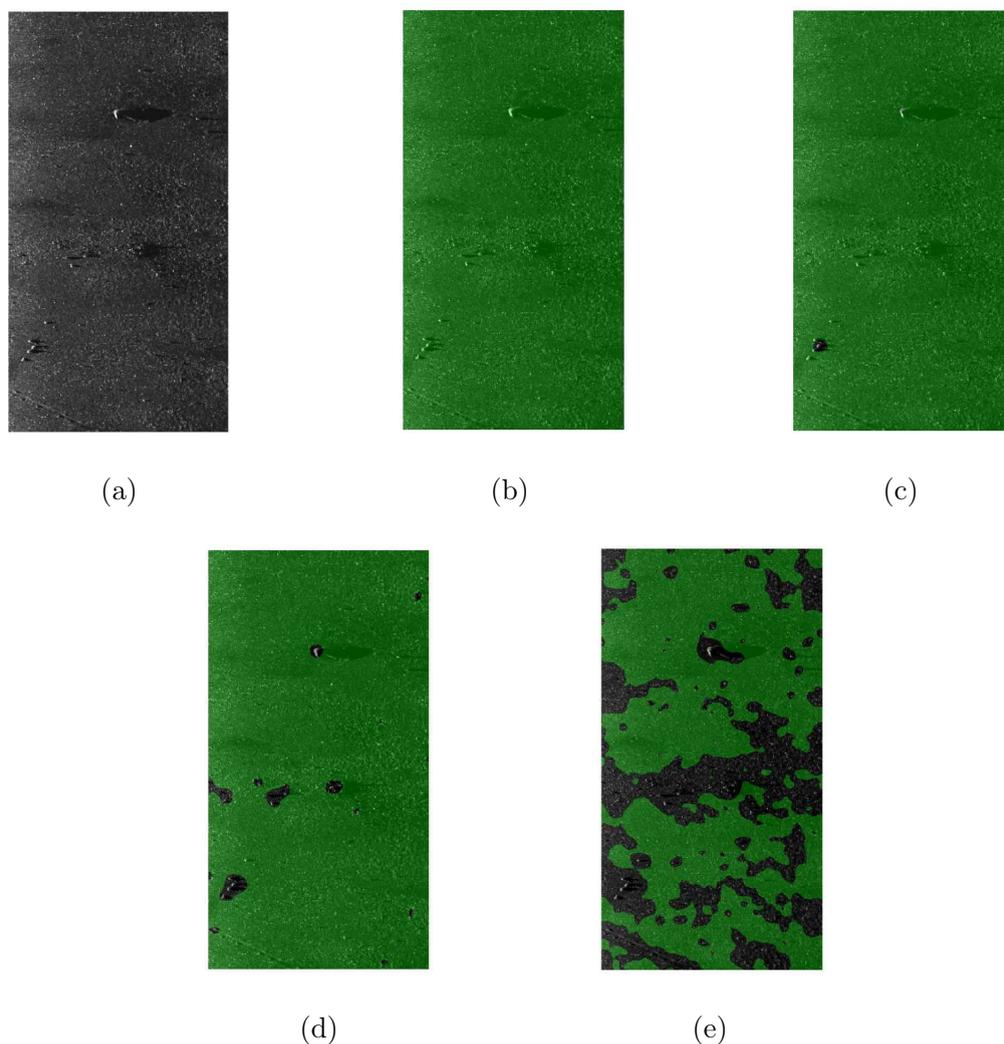


FIGURE 6.23: Effets de l'augmentation du seuil  $T$  sur l'extraction des zones homogènes (en transparence verte) d'une image sonar (a) issue de la base de données DSTO. De (b) à (e),  $T = 0, 0.25, 0.5, 0.75$  respectivement.

telle extraction de zones homogènes dans une opération de détection de mines sous-marines. En effet, cette opération ne doit surtout pas introduire de confusion entre la présence et l'absence de mines dans la région. Ainsi, on peut choisir d'extraire les zones grossièrement homogènes de l'image pour y appliquer un algorithme de détection où l'on sait qu'il sera performant. C'est-à-dire que dans tous les cas, on utilisera un détecteur sur la zone, et donc un seuil  $T$  bas, peu sélectif, peut alors suffire (*i.e.* un objet classé comme homogène ne pose pas de problème).

Une autre option consiste à extraire les zones homogènes de manière à ce qu'elles représentent les zones où l'on est sûr qu'il n'y a pas de structure sous-marine, d'objet, et encore moins de mine. Dans ce cas, les zones homogènes seront supposées comme des régions bénignes, que l'on ne traitera plus afin d'accélérer le processus de détection, et de concentrer les ressources informatiques (ou humaines) sur les zones plus complexes qui peuvent potentiellement contenir

des objets. De ce fait, le seuil  $T$  doit être très sélectif afin de ne surtout pas classer une mine comme homogène.

Dans ces travaux, nous nous plaçons dans ce second cas. Ce choix est bien sûr discutable, mais dans un contexte opérationnel, il permet de réduire fortement le pourcentage de zones de l'image à étudier pour y détecter la présence éventuelle de mines, comme nous allons le voir dans la suite. Par conséquent, ceci se traduira par diminution des temps de calcul nécessaires à la phase de détection. On propose donc de définir les zones homogènes comme les zones où il y aura une absence totale de structure sous-marine, et par conséquent de mine. Ces zones seront alors considérées comme sécurisées, et on n'y réalisera pas de phase de détection. De ce fait, il est nécessaire d'étudier le paramétrage du seuil  $T$  afin de réaliser une extraction performante.

### 6.3.2.2 Paramétrage du seuillage

Afin d'étudier le paramétrage de ce seuillage, nous avons à disposition la base de données CMRE qui nous fournit 360 images SAS accompagnées de leurs vérités terrains respectives (*i.e.* la position et le type de mines sous-marines présentes dans l'image). Pour chacune de ces images sonar, nous faisons varier la valeur du seuil  $T$  entre 0 et 1, et nous étudions le pourcentage des cibles qui ne seront pas classées comme des zones homogènes de l'image. De ce fait, elles seront considérées comme appartenant à des zones anisotropes ou complexes.

Ces résultats sont proposés sur la Figure 6.24. On y observe alors clairement que pour des seuils bas, énormément de cibles sont classées comme des fonds homogènes, il en reste alors que très peu dans les zones restantes de l'image. Plus on devient sélectif en augmentant ce seuil, et plus le pourcentage augmente. On remarque pour des valeurs du seuil au-dessus

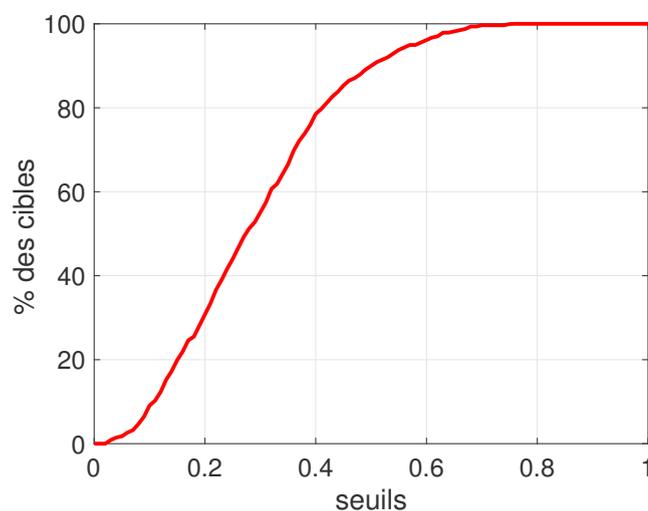


FIGURE 6.24: Pourcentage des cibles considérées comme n'appartenant pas à des zones homogènes de l'image en fonction du seuil  $T$ .

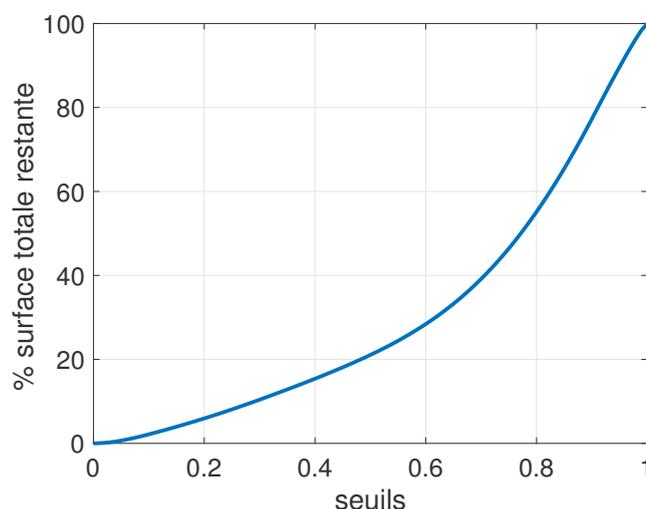


FIGURE 6.25: Pourcentage de la surface totale des images sonar de la base de données dus CMRE considérée comme non-homogènes (*i.e.* restantes à étudier) en fonction du seuil  $T$ .

de 0.75, l'intégralité des cibles de la base de données sont dans des zones qu'il sera nécessaire d'étudier plus profondément.

Par ailleurs, nous pouvons également comprendre que ce seuil joue sur la quantité de données qu'il sera nécessaire d'étudier par la suite. En effet, la Figure 6.25 propose d'illustrer le pourcentage de la surface totale des images sonar qui ne sera pas considérée comme homogène, et qu'il restera donc à étudier pour y détecter les mines sous-marines. On remarque alors que plus le seuil  $T$  sera élevé, plus il restera des zones à analyser. Ainsi, plus le volume de données sera important. Par conséquent, il serait intéressant de réaliser un compromis entre le volume de données à conserver et la présence des mines dans ces dernières. Pour cela, nous proposons d'utiliser un seuil  $T = 0.8$ , qui correspond à un pourcentage de cibles restantes de 100%, et une suppression de près de 45% de la surface des images sonar de la base de données.

De ce fait, après un tel seuillage, nous avons classifié dans l'image sonar les fonds homogènes où l'on considère l'absence de mines et de structures sous-marines. Il ne reste donc dans les données que les fonds anisotropes et complexes qu'il est nécessaire de classifier afin d'obtenir une classification complète des images.

### 6.3.3 Classification des fonds anisotropes et complexes

Afin de classer les autres zones de l'image suivant les classes anisotropes et complexes, il nous semble opportun de conserver la même démarche et d'exploiter la caractéristique de complémentarité de nos descripteurs. Nous proposons donc de classer chaque pixel des fonds marins restants par l'intermédiaire de sa valeur maximale dans les cartes d'anisotropie et de complexité correspondantes. Ainsi, un pixel  $\mathbf{x}$  est considéré comme appartenant à une zone

anisotrope si son indice d'anisotropie est plus important que son indice de complexité :  $A(\mathbf{x}) \geq C(\mathbf{x})$ . A l'inverse, il est considéré comme appartenant à une zone complexe si  $C(\mathbf{x}) > A(\mathbf{x})$ .

L'algorithme 1 illustre l'intégralité du processus de classification de l'image sonar que nous proposons. On y remarque alors que les règles utilisées sont simples et ne nécessitent aucun apprentissage, ni classifieurs évolués, à l'exception du paramétrage du seuil sur l'homogénéité  $T$ . De cette manière, nous obtenons une classification complète de l'image en trois classes de fonds marins.

---

**Algorithme 1** : Algorithme de classification

---

```

1 fonction Classification ( $H, A, C$ );
   Entrées : Cartes d'homogénéité  $H$ , d'anisotropie  $A$ , de complexité  $C$ 
   Sorties : Image sonar classifiée  $Clas$ 
2 pour tous pixels  $(x, y)$  de l'image faire
3   si  $H(x, y) \geq T$  alors
4     |  $Clas(x, y) \leftarrow Homogene$ ;
5   sinon
6     | si  $A(x, y) \geq C(x, y)$  alors
7       |  $Clas(x, y) \leftarrow Anisotrope$ ;
8     | sinon
9       |  $Clas(x, y) \leftarrow Complexe$ ;
10    | fin
11  fin
12 fin

```

---

### 6.3.4 Résultats

Des résultats de cette approche de classification des fonds marins sont proposés sur les Figures 6.26, 6.27 et 6.28. Chacune de ces trois figures illustre deux exemples pour chaque base de données utilisée au cours de ces travaux, à savoir respectivement la base de données DGA, CMRE et DSTO. Sur l'ensemble de ces résultats, la convention d'attribution des couleurs est la suivante : les zones de transparence verte, bleue et rouge représentent les fonds estimés comme homogènes, anisotropes et complexes respectivement.

Globalement, sur ces différents exemples, on remarque que la classification des zones homogènes, anisotropes et complexes est très satisfaisante. On y observe que les frontières entre les zones homogènes et anisotropes sont constamment classées comme complexes. Ce phénomène qui peut sembler problématique au premier abord est en fait intéressant, puisque le processus de détection d'objet est très délicat aux frontières entre des régions aux textures différentes.

On note aussi que la méthode de classification que nous proposons est relativement précise en terme de définition des frontières entre les différentes régions de l'image, qui provient de la représentation monogène de l'image qui nous fournit des informations en chaque pixel.

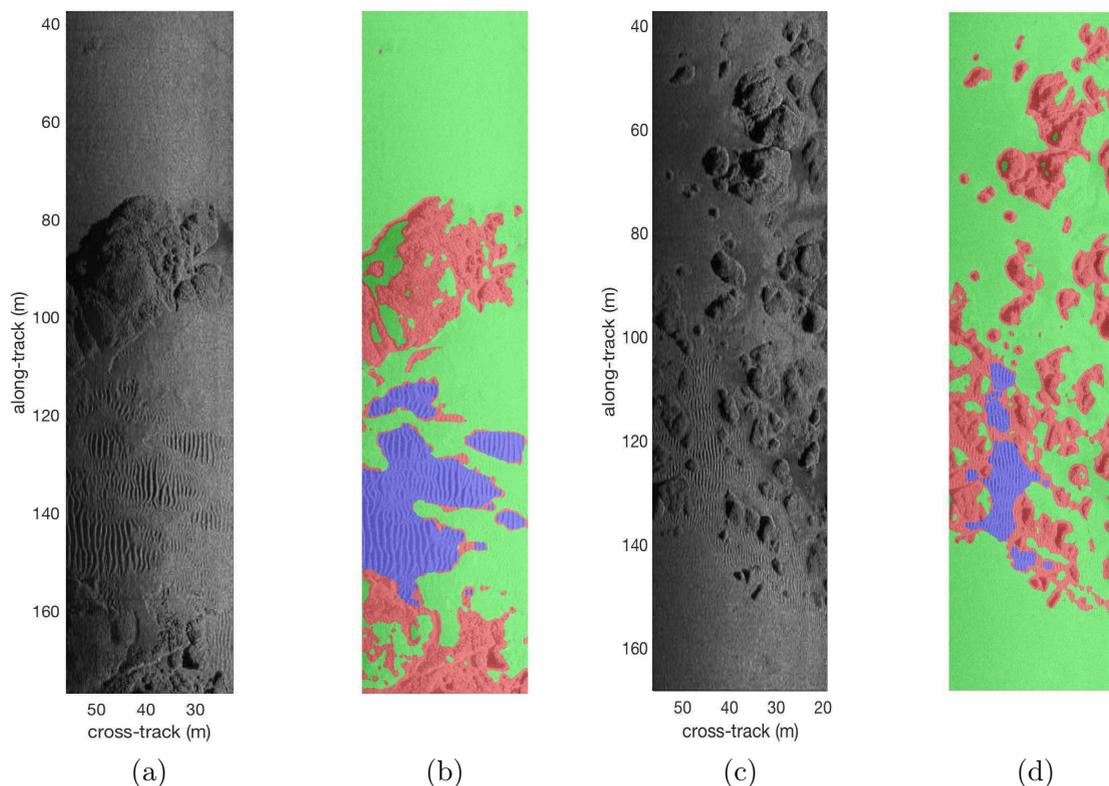


FIGURE 6.26: Illustration de résultats de classification pour deux images sonar issues de la base de données DGA.

Ceci est à mettre en parallèle avec les méthodes classiques de classification des images basées sur l'utilisation de fenêtres glissantes et qui conduisent à des effets de bloc dans l'image classifiée. Cet effet n'apparaît pas ici bien qu'une notion de voisinage est tout de même prise en considération comme nous l'avons vu à la section 5.4.1.

On remarque tout de même que les résultats sont perfectibles, en particulier sur la Figure 6.28.(d). En effet sur les données du DSTO, on observe que cette classification est assez dépendante de la portée et ne tient pas compte des effets qui lui sont dus dans l'image comme l'augmentation du bruit ou les phénomènes de déformations géométriques.

Par ailleurs, il est important de noter que les structures isolées posées sur les fonds marins, dont les dimensions sont proches de celles de cibles sous-marines, apparaissent comme des zones classées complexe. Ainsi, puisque nous sommes dans une logique de chaîne de traitement ATR, il sera nécessaire plus tard d'en tenir compte lors de la phase de détection de cibles, et de trouver un moyen de ne pas les traiter de la même manière qu'un fond marin relié à une véritable texture complexe.

On note également que la notion d'ombre n'est pas prise en compte par cette approche de classification. On remarque que les zones d'ombre de faibles étendues sont généralement classées de la même manière que les structures qui l'ont créée, à savoir anisotrope ou complexe. Cependant, lorsque l'ombre est très importante, on observe qu'elle est classée comme une zone

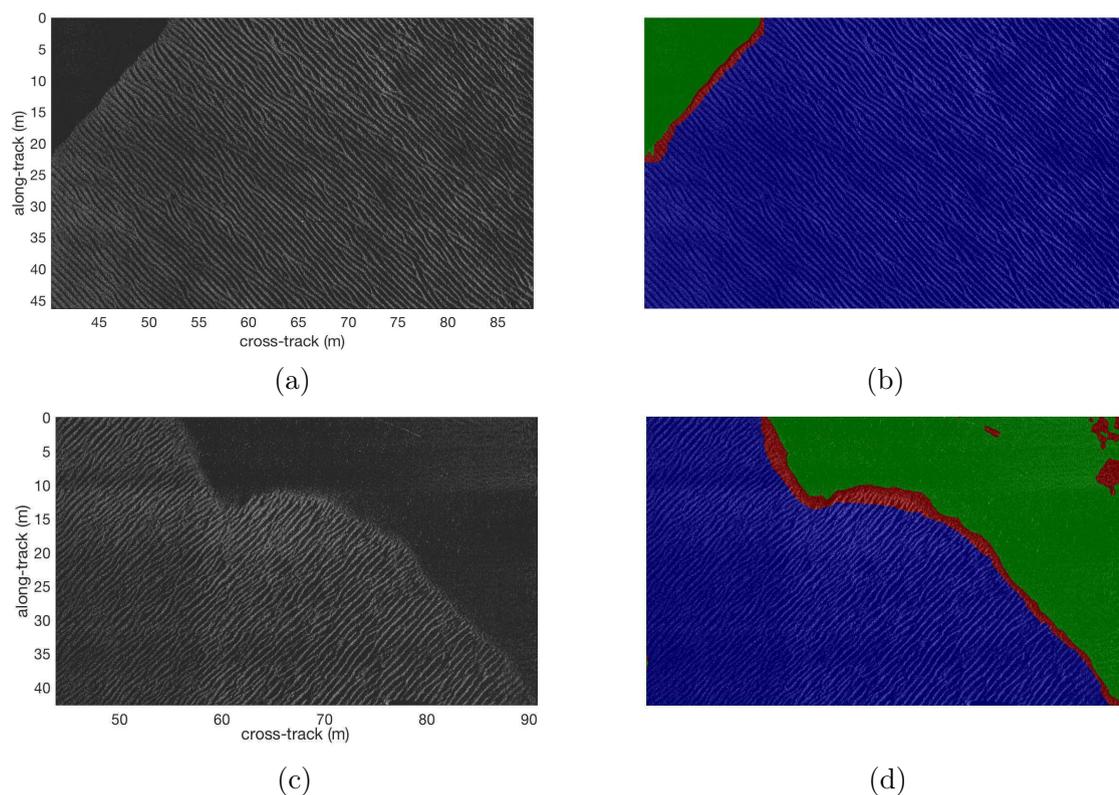


FIGURE 6.27: Illustration de résultats de classification pour deux images sonar issues de la base de données CMRE.

homogène. C'est le cas sur la Figure 6.26.(b) pour l'ombre d'une grande roche présente sur la gauche de la moitié supérieure de l'image. Bien que les zones d'ombre sont des régions de l'image qui ne possèdent aucune structure apparente, elles n'en demeurent pas moins des zones où un objet peut facilement s'y cacher lorsqu'elles sont de tailles importantes. Il serait donc nécessaire de traiter ce problème de manière à mettre en évidence de telles zones notamment en vue d'une possible replanification d'une mission ou d'un AUV.

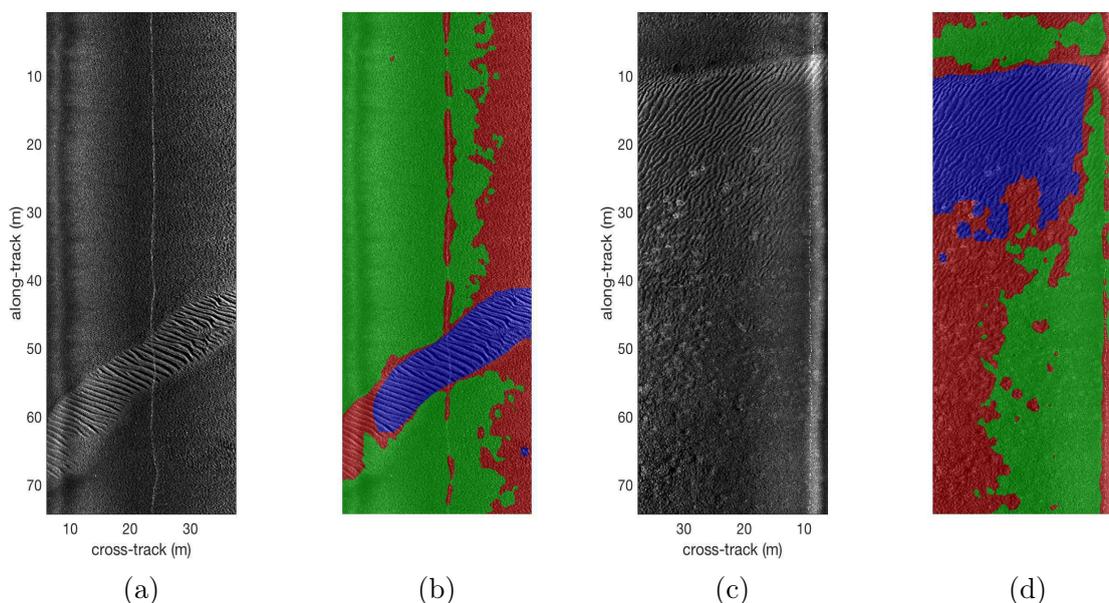


FIGURE 6.28: Illustration de résultats de classification pour deux images sonar issues de la base de données DSTO.

## 6.4 Conclusion

A travers ce chapitre, nous nous sommes intéressés à la description des images sonar par le biais de la complexité de leurs textures d'un point de vue détection de mines sous-marines. Pour ce faire, nous avons défini trois descripteurs reliés à la notion d'homogénéité, d'anisotropie et de complexité des fonds marins. Ces derniers, estimés par l'intermédiaire des dimensions intrinsèques continues multi-échelles, ont la particularité d'être complémentaires et de représenter des indices de confiance sur le type de fond marin présent dans l'image. Après comparaison avec les descripteurs déjà proposés dans la littérature, nous avons pu constater que notre approche permet une meilleure discrimination (quantitative ou qualitative) des fonds homogènes, anisotropes et complexes, tout en fournissant des valeurs de descripteurs plus significatives (dynamiques connues *a priori*).

En étudiant un détecteur de mines sous-marines de la littérature et déjà implémenté à bord des AUVs du CMRE, nous avons remarqué que les performances de détection sont directement reliées aux différents types de fonds (homogènes, anisotropes, complexes) obtenus par l'intermédiaire de nos descripteurs. En particulier, nous avons constaté que les fonds marins décrits par de forts indices d'anisotropie et de complexité sont à l'origine d'importants taux de fausses alarmes et non détections. Ces observations démontrent que les différents types de fonds marins ne peuvent pas être analysés de la même manière en ATR. Ainsi, selon nous, un seul et même détecteur ne pourra pas atteindre des performances élevées pour tous les types de fonds marins. Nous supposons qu'il est donc nécessaire de tenir compte des caractéristiques propres à chaque type de fond lors de la phase de détection. Cependant pour pouvoir imaginer

un détecteur spécifique aux zones anisotropes ou complexes par exemple, il est nécessaire de pouvoir localiser ces types de fonds dans l'image sonar au préalable. Nous avons alors proposé d'aborder la problématique de classification des images sonar.

La définition et la localisation précises des trois types de fonds homogènes, anisotropes et complexes a été réalisée par l'intermédiaire de considérations simples sur nos trois descripteurs continus. Cette approche permet une classification non-supervisée de l'image sonar nécessitant peu de paramètres (un seul seuillage y est réalisé pour la classification des zones homogènes). Nous avons pu remarquer que cette méthode s'adapte facilement aux trois bases de données que nous avons à disposition, sans modification du moindre paramètre dans nos algorithmes.

Une fois ces trois types de fonds localisés de manière précise, nous verrons dans le chapitre suivant qu'il est possible de caractériser chacun d'entre eux séparément. Cette caractérisation permettra de fournir des informations utiles pour la phase de détection de cibles, et de repenser l'architecture d'une chaîne ATR afin qu'elle s'adapte le mieux possible au type de fond marin.

# Chapitre 7

## Exploitation de la description environnementale en ATR

### Sommaire

---

<b>7.1</b>	<b>Introduction</b>	<b>187</b>
<b>7.2</b>	<b>Application à la caractérisation des fonds marins</b>	<b>187</b>
7.2.1	Motivations	187
7.2.2	Caractérisation des fonds homogènes	188
7.2.2.1	Observations	188
7.2.2.2	Analyse de l'histogramme des images sonar	189
7.2.2.3	Détection des zones d'ombre par segmentation d'histogramme	192
7.2.2.4	Détection des zones d'ombre par les quantiles	194
7.2.2.5	Résumé de la chaîne de détection d'ombres et résultats	195
7.2.3	Caractérisation des fonds anisotropes	198
7.2.3.1	Longueur d'onde des rides de sable	198
7.2.3.2	Orientation des rides de sable	200
7.2.3.3	Mise en évidence des crêtes et vallées des rides de sable	203
<b>7.3</b>	<b>Architecture ATR adaptée au type de fond</b>	<b>205</b>
7.3.1	Prise en compte de la caractérisation environnementale	205
7.3.2	Stratégie proposée	206
<b>7.4</b>	<b>Détection de cibles dans les zones anisotropes</b>	<b>207</b>
7.4.1	Analyse des zones anisotropes avec et sans objets	207
7.4.2	Détection d'échos	209
7.4.2.1	Carte des échos	209
7.4.2.2	Le détecteur CA-CFAR	211
7.4.2.3	Résultats de la détection des échos	213
7.4.3	Détection d'ombres	215

7.4.4 Résultats finaux de la détection . . . . .	216
<b>7.5 Conclusion et perspectives . . . . .</b>	<b>220</b>

---

## 7.1 Introduction

Jusqu'à présent, nous avons réalisé une description des fonds marins dans l'image sonar en termes d'homogénéité, d'anisotropie et de complexité. Nous avons tout d'abord proposé une description continue avec la définition de trois nouveaux descripteurs, suivie d'une description discrète par l'intermédiaire d'une classification non-supervisée de l'image sonar (en trois classes). Cette dernière opération permet de localiser précisément les différents types de fonds marins que nous avons identifiés comme impactant les performances d'une chaîne ATR. Dans ce chapitre, nous proposons d'exploiter cette description de l'environnement marin pour extraire des caractéristiques significatives sur ces différents types de fonds puis de les intégrer à une chaîne ATR afin qu'elle adapte le processus de détection de cibles à son environnement.

Dans un premier temps, à la section 7.2, nous allons nous attacher à caractériser deux types de fonds particuliers, à savoir les fonds homogènes et anisotropes. Alors que pour le premier cité nous verrons que c'est la détection des zones d'ombre qui sera pertinente en termes de caractérisation, pour les fonds anisotropes nous pourrions extraire des informations de longueur d'onde, d'orientation ou encore de localisation des crêtes et des vallées de ces rides de sable. Une fois que ces informations seront disponibles, nous verrons qu'il est alors possible de repenser l'architecture d'un processus ATR dans la section 7.3. Cette architecture proposera d'utiliser un détecteur spécifique pour chaque type de fond marin et sera basée sur l'utilisation des informations environnementales extraites précédemment. Dans la section 7.4, nous proposerons de réaliser un premier détecteur de cibles sous-marines spécialisé aux champs de rides de sable.

## 7.2 Application à la caractérisation des fonds marins

### 7.2.1 Motivations

Dans le chapitre précédent, nous avons proposé de réaliser une classification de l'image en trois classes : les fonds homogènes, anisotropes et complexes. À travers cette section, nous proposons d'analyser et de caractériser en particulier deux d'entre eux, à savoir les fonds homogènes et anisotropes. Cette nouvelle étape dans le traitement permettra d'extraire des informations qui seront potentiellement utiles pour une chaîne ATR. Pour le permettre, nous avons à disposition toutes les informations issues du signal monogène et des dimensions intrinsèques continues dans le cadre multi-échelle.

## 7.2.2 Caractérisation des fonds homogènes

Dans ce paragraphe, nous allons nous focaliser sur l'analyse des fonds marins classés comme homogènes dans l'image sonar. Après quelques observations préliminaires, nous verrons qu'il est possible de caractériser ces types de fonds particuliers à partir de l'histogramme des niveaux de gris afin d'y détecter la présence de grandes zones d'ombre. Ces dernières sont dignes d'intérêt en chasse aux mines puisqu'elles cachent de grandes portions de l'image sonar et qu'il y est impossible de réaliser une détection de cible. Nous compléterons cette étude en abordant également la problématique de la détection de zones d'ombre de petites dimensions.

### 7.2.2.1 Observations

Comme nous l'avons vu au chapitre précédent, les fonds marins classés comme homogènes représentent les régions de l'image sonar où l'on considère l'absence de structure sous-marine ou de mine. Cependant, si l'on observe les résultats de classification proposés sur la Figure 7.1, on peut constater que les régions classées comme homogènes ne correspondent pas précisément à la définition des fonds homogènes que nous avons donné à la section 6.2 (*i.e.* des fonds correspondants aux zones de sable fin ou de vase). Comme nous avons déjà pu le noter lors de l'analyse des résultats de classification (cf. paragraphe 6.3.4), on constate que les zones de fonds classées comme homogènes (en transparence verte sur les figures) font également apparaître des zones d'ombre.

Si l'on observe l'image sonar de la Figure 7.1.(a), on peut constater la présence d'une très grande zone d'ombre dans le coin inférieur droit. Les grandes zones d'ombre de ce type sont généralement dues au relief ou à des structures sous-marines très hautes qui empêchent la propagation des ondes acoustiques et donc l'insonification de la zone immédiatement derrière. Une telle structure sous-marine occulte donc une partie du fond marin où la détection d'objet est alors rendue impossible. Cependant, si l'on observe ces zones de l'image sonar, on constate qu'elles sont par définition des zones homogènes en termes de dimensions intrinsèques (signal d'amplitude très faible mais constant) et qu'il est donc naturel de les retrouver classées comme homogènes ici. Or dans ces travaux, nous avons défini un fond homogène comme étant une zone de l'image qui ne contient pas de mine sous-marine. Une telle classification des zones d'ombre est donc contradictoire puisqu'on ne peut affirmer l'absence d'une cible dans cette zone. Il serait donc adéquat de distinguer les véritables zones de fonds homogènes (sable ou vase) vis-à-vis de ces grandes zones d'ombre. En termes opérationnels, l'avantage d'une détection de ces grandes zones d'ombre serait de permettre une re planification d'une trajectoire pour le porteur du sonar. De la sorte, il serait possible d'analyser la zone avec un point de vue différent, permettant d'observer le fond marin précédemment occulté.

Puisque nous avons été amené à aborder la problématique des zones d'ombre, une seconde observation peut être réalisée sur les images présentées à la Figure 7.1.(b). Cette fois,

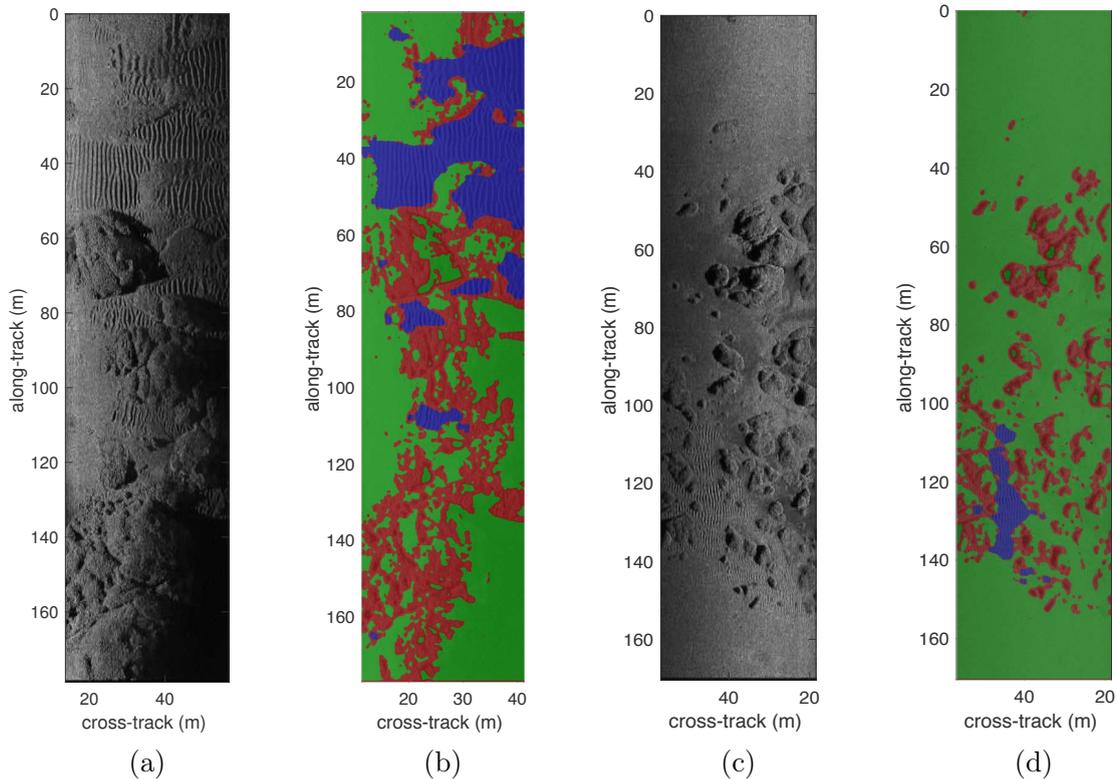


FIGURE 7.1: Illustration de résultats de classification pour deux images sonar issues de la base de données DGA.

elle concerne les zones d'ombre de petites dimensions. Celle-ci sont généralement dues à des structures sous-marines également de petites dimensions, et l'on constate qu'elles sont bien souvent classées de la même manière que les structures qui les ont provoquées (*i.e.* anisotropes ou complexes), telles que des cailloux, des rides de sables ou encore des objets. Ainsi, bien que ces zones particulières de l'image n'appartiennent pas directement aux zones classées comme homogènes, il serait tout de même intéressant de trouver un moyen de les mettre en évidence pour disposer d'une méthode complète de détection des zones d'ombre de l'image sonar.

Afin d'aborder ces problématiques liées à la présence des zones d'ombre dans l'image sonar, nous proposons dans un premier temps d'analyser l'histogramme des images sonar.

### 7.2.2.2 Analyse de l'histogramme des images sonar

Comme nous venons de le constater, les zones d'ombre se distinguent de la réverbération du fond marin par le niveau d'intensité bien plus faible. On peut donc émettre l'hypothèse que la présence de ces zones d'ombre puisse impacter la forme d'un histogramme, c'est pourquoi nous proposons d'analyser ce point spécifique ici. Par ailleurs, puisque nous cherchons à décrire les zones sombres de l'image, nous proposons de pré-traiter les images sonar en réalisant un rehaussement logarithmique du contraste (*i.e.* images sonars converties en décibels). Cela a pour effet d'éclaircir les images sonar en augmentant l'intervalle des intensités sombres. La Figure

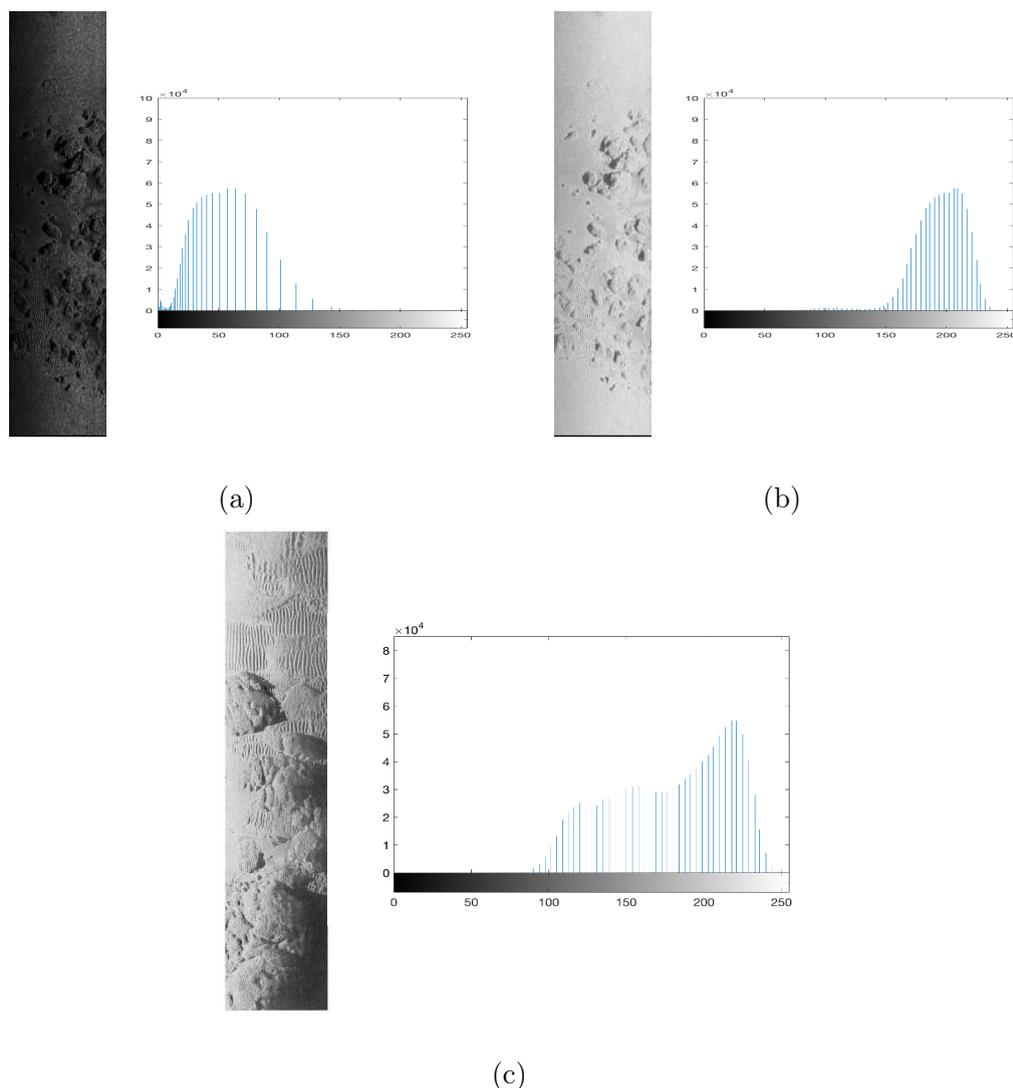


FIGURE 7.2: En (a) et (b) : illustration de l'histogramme en niveaux de gris pour une même image sonar avant et après rehaussement du contraste. En (c) : autre exemple d'image sonar et de son histogramme rehaussé.

7.2 propose d'illustrer cette opération sur une image sonar issue de la base de données DGA et d'observer son effet sur l'histogramme. Avant le rehaussement du contraste (a) et conformément à ce que nous avons vu au paragraphe 2.3.1, on constate que l'histogramme en niveaux de gris fait apparaître une distribution sous la forme d'une loi de Rayleigh. Après rehaussement du contraste (b), on observe que l'image sonar est alors bien éclaircie. Son histogramme est également modifié puisque la concentration des niveaux de gris se fait maintenant autour des fortes intensités. Par conséquent, la distribution de Rayleigh ne permet plus de caractériser cet histogramme.

Au contraire de l'image 7.2.(b) qui ne présente que des zones d'ombre peu étendues, la Figure 7.2.(c) illustre l'histogramme rehaussé d'une image sonar qui présente quant à elle une grande zone d'ombre en bas à droite. On aurait pu penser que le grand nombre de pixels

appartenant aux zones sombres aurait une influence sur l'aspect de son histogramme. Pourtant, on constate qu'il n'y a qu'un seul et même mode sur ces deux histogrammes, ce qui ne permet pas d'envisager une détection des ombres simplement à partir de ce type d'histogramme.

Cependant, nous avons vu que les grandes zones d'ombre sont des régions de l'image sonar initiale classées comme homogènes. Nous proposons donc de réaliser à nouveau un histogramme, mais cette fois, uniquement à l'intérieur des régions classées comme homogènes. La Figure 7.3 illustre de tels histogrammes dans le cas des mêmes images sonar que nous avons vu précédemment. Alors que dans le cas de l'image (a), ce nouvel histogramme ne diffère que très peu par rapport au précédent, dans le cas de l'image (b), deux modes apparaissent clairement. Alors que l'un de ces modes est relié aux zones claires de l'image (*i.e.* aux réverbérations du fond), le second mode correspond aux zones sombres de l'image (*i.e.* aux zones d'ombre). Ce phénomène provient du fait que de grandes zones d'ombre sont présentes à l'intérieur des zones classées comme homogènes. Il y a donc assez de pixels appartenant aux zones d'ombre pour créer un second mode.

Finalement, l'histogramme relatif aux zones homogènes d'une image sonar permet de constater la présence ou non de grandes zones d'ombre dans l'image. Si cet histogramme est bimodal alors il y a une forte probabilité que de grandes zones d'ombre soient présentes. Si au contraire il est unimodal, il y a une forte probabilité que l'image sonar ne contienne pas de grandes zones d'ombre, mais seulement des petites.

Dans les deux paragraphes suivants, nous allons nous intéresser à la manière de détecter les zones d'ombre dans ces deux cas de figure.

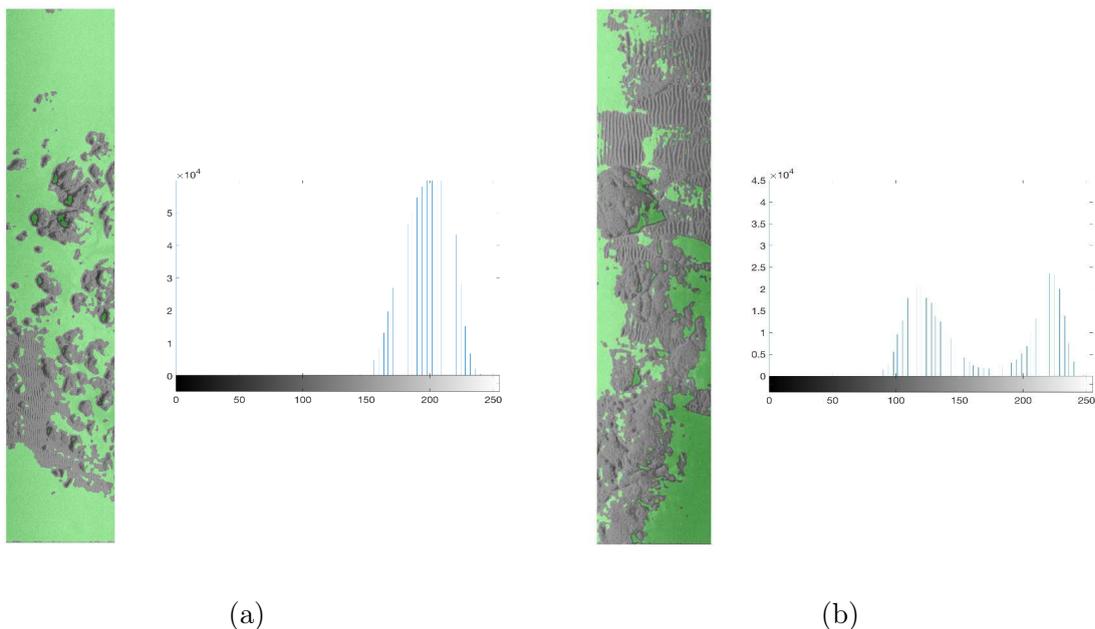


FIGURE 7.3: Histogrammes des zones homogènes (en transparence verte) de deux images sonar de la base de données DGA.

### 7.2.2.3 Détection des zones d'ombre par segmentation d'histogramme

Nous venons de constater que l'histogramme des zones homogènes affiche deux modes lorsque de grandes zones d'ombre sont présentes dans l'image sonar. Il serait donc intéressant d'analyser un tel histogramme de manière automatique afin de caractériser le nombre de mode et de pouvoir potentiellement détecter les ombres présentes dans l'image.

Les études liées à l'analyse d'histogramme sont aujourd'hui nombreuses puisque ce dernier permet de représenter de manière compacte une grande quantité de données et d'en extraire un ensemble de propriétés statistiques sur celles-ci. L'une des descriptions les plus efficaces est la liste des modes présents dans l'histogramme, soient les intervalles de valeurs autour desquels les niveaux de gris sont concentrés. Cependant, il n'est pas toujours évident de quantifier la concentration des données dans un intervalle, et par conséquent, de dénombrer le nombre de modes.

On recense deux types d'approches pour segmenter un histogramme, les approches paramétriques et les non-paramétriques. Alors que les premières se basent sur des hypothèses concernant la distribution des données, comme par exemples des mélanges de Gaussiennes (voir [158]), les approches non-paramétriques reposent quant à elles sur aucune hypothèse de distribution et réalisent différents seuillages afin de diviser l'histogramme en intervalles, le tout en cherchant à minimiser un critère énergétique, comme la variance, l'entropie etc. (voir [159, 160, 161, 162, 163]).

Jusqu'à présent, nous avons pu constater que le traitement de l'image sonar que nous réalisons dans ces travaux évite au maximum l'utilisation de paramètres. De ce fait, nos algorithmes s'adaptent facilement aux différentes bases de données que nous avons à disposition. De plus, nous ne connaissons pas *a priori* la composition de l'histogramme et le nombre de modes. Ainsi, nous privilégions une approche non-paramétrique.

D'après J. Delon [164], les données réelles créent des histogrammes complexes où le phénomène de bruit intervient par de nombreux pics auxquels les méthodes non-paramétriques sont très sensibles. Ainsi, ces méthodes fournissent généralement une surévaluation importante du nombre de modes. L'auteur a donc proposé une nouvelle méthode non-paramétrique, robuste aux faibles variations dans l'histogramme et assez locale pour y détecter les petits modes. Cette méthode, appelée algorithme de segmentation FTC (*Fine to Coarse*), est basée sur une définition rigoureuse d'une segmentation dite *admissible*, qui évite les problèmes de sur- et sous-segmentation des histogrammes. Puisqu'aucune modification de l'algorithme publié par l'auteur n'a été apportée ici, nous invitons donc le lecteur à se référer à l'étude originale [164] pour plus de détails sur son implémentation.

La Figure 7.4 propose le résultat de l'algorithme FTC sur les histogrammes de la Figure 7.3. Alors que dans le cas de l'image sans grande zone d'ombre l'algorithme détecte bien qu'un

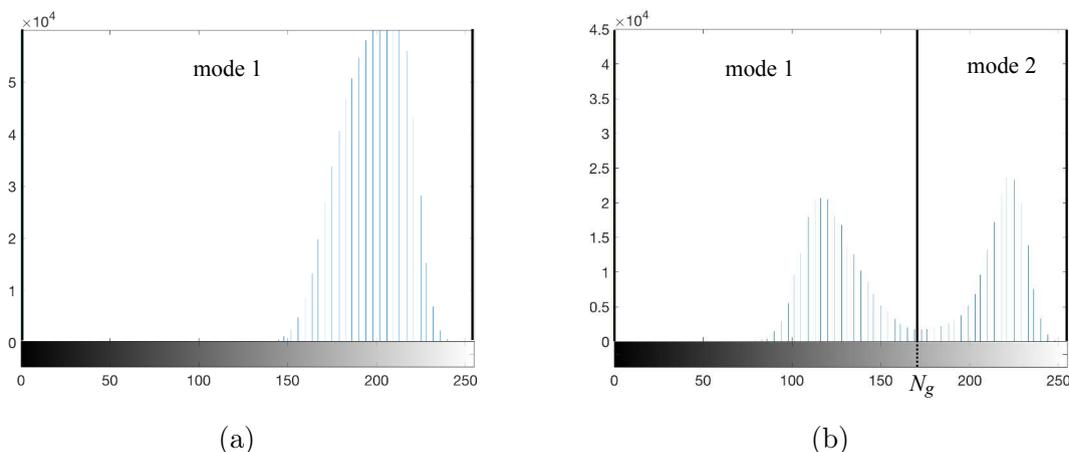


FIGURE 7.4: Segmentation de deux histogrammes par la méthode FTC.

seul mode, dans le cas de l'image sonar avec une grande zone d'ombre, l'algorithme détecte deux modes, l'un étant donc relié aux zones d'ombre.

D'après ce qui précède, si on s'intéresse au mode qui est centré sur l'intensité la plus faible, ce dernier devrait correspondre aux pixels des zones d'ombre. Pour le vérifier, nous proposons de seuiller l'image sonar de la Figure 7.3.(b), notée  $I$ , de telle sorte que :

$$I(\mathbf{x}) < N_g, \quad (7.1)$$

où  $N_g$  est le niveau de gris correspondant à la première séparation entre les modes détectés par l'algorithme FTC (voir Figure 7.4.(b)).

La Figure 7.5.(b) illustre cette détection des zones d'ombre de l'image sonar 7.5.(a). On remarque que la grande zone d'ombre en bas de l'image est clairement mise en évidence. De plus, on peut observer que les petites zones d'ombre profitent également du fait que la grande zone d'ombre fournisse un mode significatif dans l'histogramme pour être également détectées (c'est particulièrement le cas en bas à gauche de l'image).

On constate également sur la Figure 7.5.(b) que les ombres associées aux rides de sable (dans la partie supérieure de l'image) ne sont pas correctement estimées. Nous verrons dans la suite de ce manuscrit une méthode de détection des ombres spécifiques aux champs de rides de sable. Il est donc normal de ne pas observer de zones d'ombre dans les zones classées anisotropes de la Figure 7.5.(c).

Dans le paragraphe suivant, nous allons nous intéresser à la détection des ombres dans le cas où l'histogramme est cette fois unimodal.

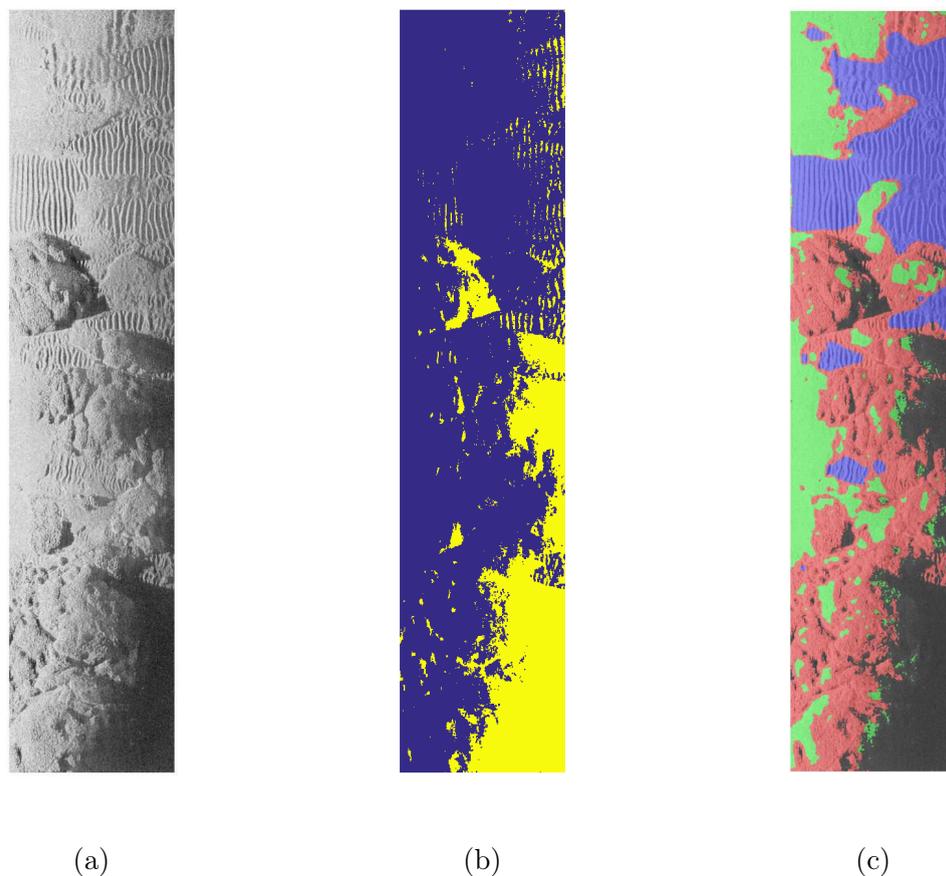


FIGURE 7.5: Résultats d'extraction des zones d'ombre par segmentation d'histogramme : (a) image sonar initiale, (b) zones d'ombre, (c) prise en compte dans l'image classée, où les zones sombres correspondent aux ombres détectées.

#### 7.2.2.4 Détection des zones d'ombre par les quantiles

Nous avons pu voir aux Figures 7.2 et 7.3 que l'unique présence d'ombres de petites dimensions est imperceptible que ce soit sur l'histogramme de l'image entière ou sur celui des zones homogènes. De ce fait, lorsque l'algorithme FTC nous informe qu'il n'y a qu'un seul mode dans l'histogramme, il est nécessaire de procéder autrement pour détecter les éventuelles ombres dans l'image sonar.

Puisque la quantité de ces ombres dans l'image n'est pas suffisante pour créer un mode dans l'histogramme, nous considérons donc qu'elles sont rares dans l'image. Cette rareté peut être quantifiée par l'utilisation des quantiles.

Par exemple, considérons un ensemble de valeurs aléatoires que l'on trie de la plus petite à la plus grande. Les quantiles seront alors un moyen de diviser ce jeu de données en différents intervalles qui dépendent de l'ordre du quantile souhaité. Par exemple, si on considère que notre ensemble d'observations contient 100 échantillons, un quantile d'ordre 50% (ou 1/2) aura pour effet de séparer cet ensemble en deux intervalles de 50 échantillons. Ce quantile particulier est ce que l'on appelle la médiane.

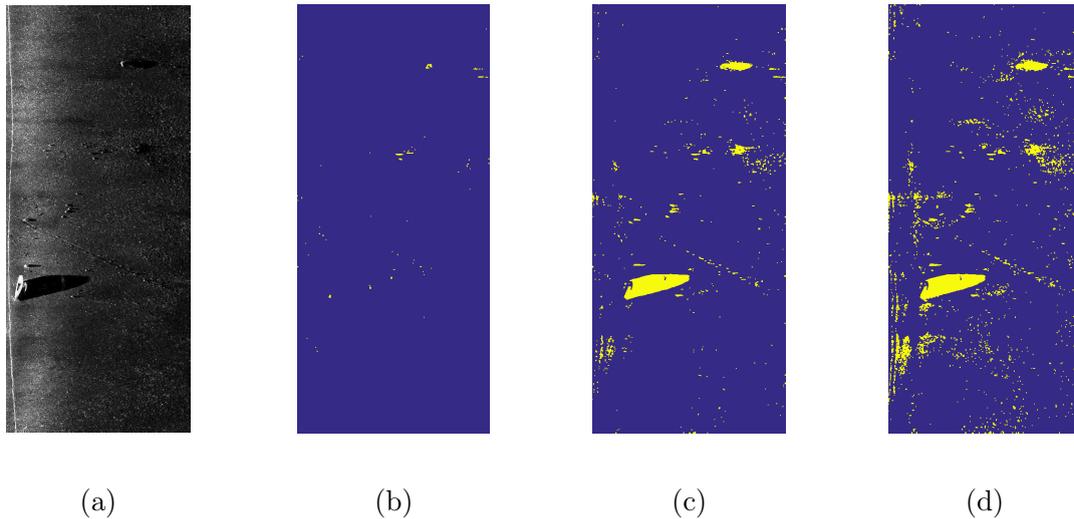


FIGURE 7.6: Résultats d'extraction des zones d'ombre par utilisation des quantiles pour une image sonar de la base de données du DSTO (a) pour trois valeurs distinctes de percentiles, à savoir pour le 1<sup>er</sup> (b), 5<sup>ème</sup> (c) et le 10<sup>ème</sup> (d).

Certains jeux de quantiles possèdent des noms bien particuliers, comme les percentiles, les déciles ou encore les quartiles. Ici, nous nous intéresserons particulièrement aux percentiles qui sont des multiples du centième. Par exemple, le 3<sup>ème</sup> percentile est une valeur qui sépare les échantillons de manière à ce qu'il y en est 3% en dessous et 97% au-dessus. C'est ce que nous allons utiliser pour détecter les pixels d'ombre considérés comme rares.

Afin que cette étude ne soit pas dépendante de la portée, nous proposons de définir, pour chaque colonne de l'image, le  $n^{\text{ème}}$  percentile de ces niveaux de gris. Ainsi, pour chaque colonne de l'image, nous obtenons donc un seuil par l'intermédiaire de ce percentile. Tous les pixels de colonne qui sont inférieurs à ce seuil seront alors considérés comme appartenant à des ombres.

La Figure 7.6 propose l'illustration d'une telle estimation dans le cas de trois valeurs distinctes de percentiles, à savoir le 1<sup>er</sup> (b), 5<sup>ème</sup> (c) et le 10<sup>ème</sup> (d). L'image (a) est une image sonar issue de la base de données du DSTO qui illustre le cas d'un fond homogène où sont posés un certain nombre d'objets (un petit voilier dans la partie inférieure gauche de l'image, et une mine dans la partie supérieure droite) ainsi que des cailloux de petites dimensions. On note alors que plus on augmente la valeur du percentile, plus on aura une sur-estimation des zones ombres. Cependant lorsque le percentile est trop bas, l'estimation est trop stricte et conduit à une sous-estimation des zones d'ombre. Il est donc nécessaire de réaliser un compromis, et nous proposons alors d'utiliser le 5<sup>ème</sup> percentile.

#### 7.2.2.5 Résumé de la chaîne de détection d'ombres et résultats

Au final, nous avons étudié dans cette section la caractérisation des zones homogènes. Nous avons pu constater que les zones de l'image classées comme homogènes sont en fait un

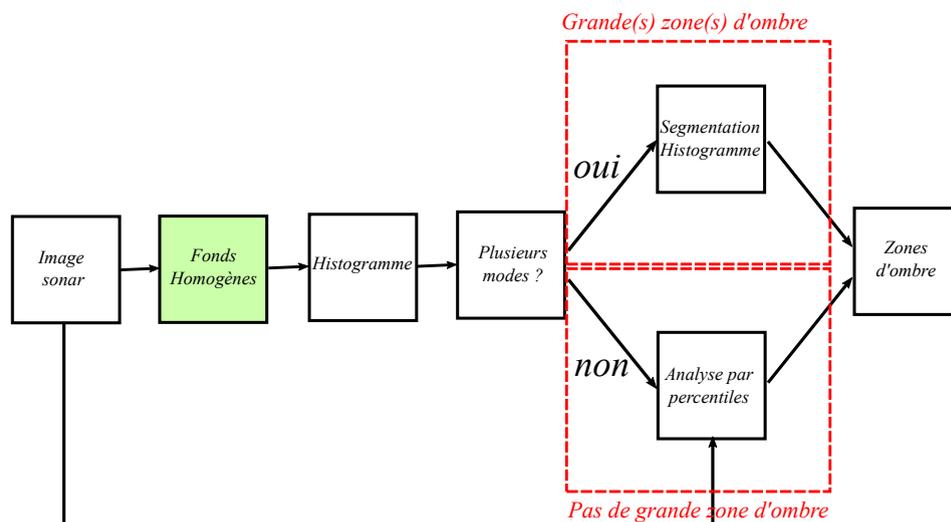


FIGURE 7.7: Synoptique de la chaîne de détection des ombres.

mélange des fonds sableux et vaseux que nous voulions mettre en évidence, mais également des zones d'ombre. Or ces ombres ne peuvent pas être considérées comme des fonds homogènes puisqu'elles représentent l'absence d'information sur le fond, et par conséquent, sur la présence éventuelle de mines sous-marines. Nous avons alors proposé une méthode pour détecter les zones d'ombre dans l'image sonar. Celle-ci utilise en fait deux chaînes de détection (cf. schématisation de la chaîne de détection des ombres de la Figure 7.7), à savoir une détection spécifique en présence de grandes zones d'ombre et une détection spécifique lorsque l'image ne contient que de petites zones d'ombre. Le choix de la chaîne à utiliser est réalisé par l'intermédiaire d'une analyse de l'histogramme des fonds homogènes. Ce dernier est segmenté de manière non-paramétrique par l'algorithme FTC. Lorsque ce dernier détecte plus d'un mode, on considère qu'il existe au moins une grande zone d'ombre dans l'image. Un simple seuillage sur l'histogramme permet alors de détecter simplement les zones d'ombre. Par contre, si un seul mode est détecté, on considère qu'il n'y a que très peu d'ombre dans l'image sonar. On propose alors une détection basée sur une analyse des percentiles de chaque colonne de l'image (*i.e.* pour une distance donnée).

Dans ces deux cas, on obtient finalement une carte des ombres. On peut alors en tenir compte dans la classification de l'image sonar, comme cela est illustré sur les Figures 7.8 et 7.9 pour des images sonar de la base de données DGA et DSTO. Les zones d'ombre détectées y sont représentées en noir et on remarque que leur estimation semble alors très correcte.

Nous avons pu constater que ce processus d'extraction des zones d'ombre dans les images sonar fonctionne pour les bases de données DGA et DSTO. Pour les images SAS issues de la base de données du CMRE, les résultats ne sont pas satisfaisants et la détection de ces zones d'ombre reste toujours à ce jour un problème ouvert. La difficulté majeure provient très probablement d'un pré-traitement de ces images interne au CMRE, qui ne permet pas d'étudier de cette manière les histogrammes qui nécessiteraient alors d'un traitement particulier.

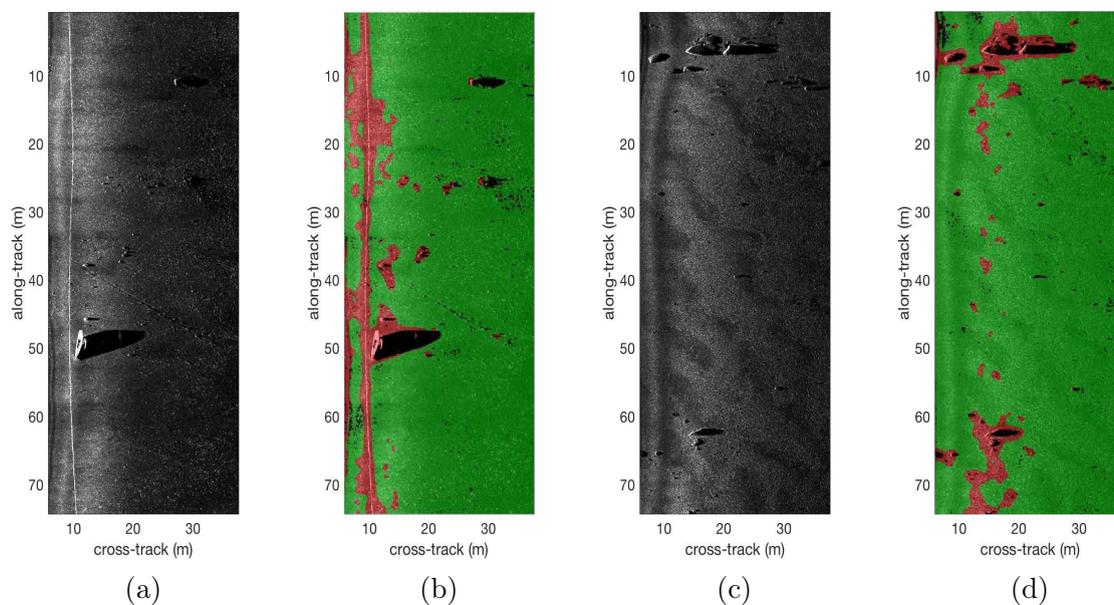


FIGURE 7.8: Résultats finaux de classification pour deux images sonar (a) et (c) de la base de données DSTO : la transparence verte indique les zones classées comme homogènes, la bleue les zones classées comme anisotropes, la rouge les zones classées comme complexes et la noire les zones classées comme de l'ombre.

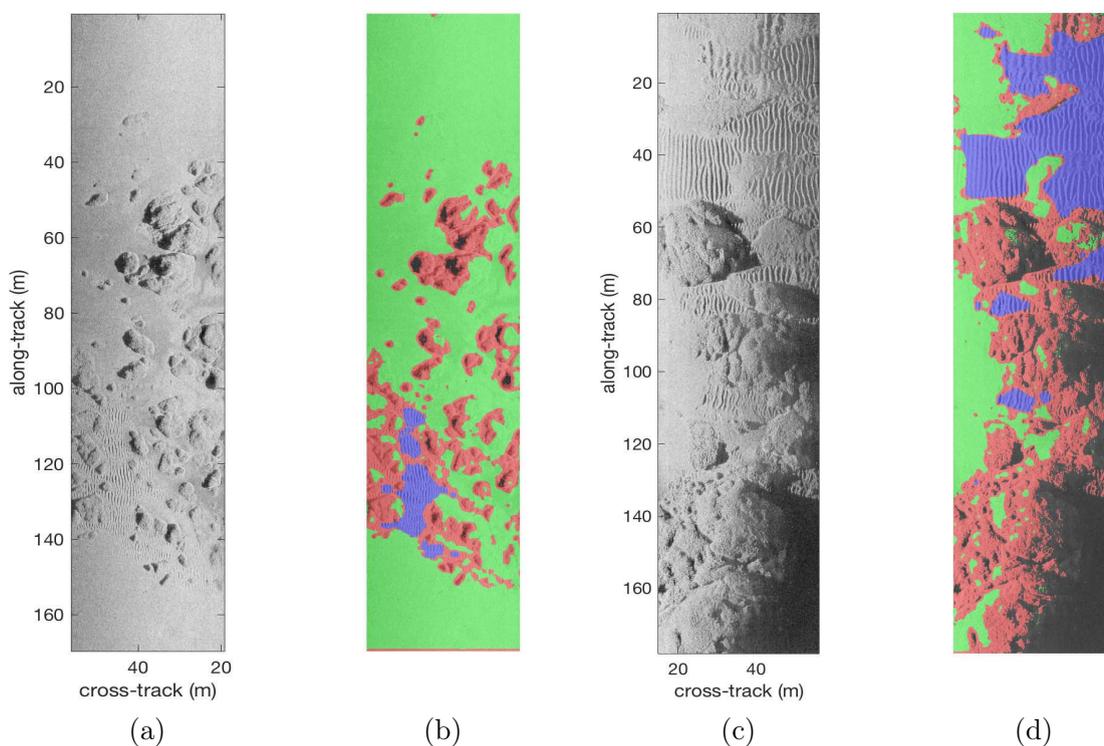


FIGURE 7.9: Résultats finaux de classification pour deux images sonar (a) et (c) de la base de données DGA.

### 7.2.3 Caractérisation des fonds anisotropes

Dans les paragraphes précédents, nous nous sommes attachés à caractériser les zones de l'image classées comme homogènes. Ici, nous allons nous intéresser aux zones de l'image classées comme anisotropes qui correspondent aux champs de rides de sable.

Les rides de sable sont des structures sous-marines qui apparaissent dans les images sonar comme des motifs périodiques avec des successions d'échos et d'ombres. Dans un contexte de chasse aux mines sous-marines, ces rides de sable posent alors certaines difficultés pour la détection d'objets. En effet, ces objets, qui sont généralement détectés et identifiés par l'intermédiaire de leur paire "écho-ombre", se retrouvent sur un fond marin qui les déforme ou les dissimule. Alors que l'ombre d'un objet peut s'y fragmenter ou se combiner avec celles des rides de sable, les caractéristiques de taille ou de forme des échos des cibles peuvent également y être confondues voire masquées par ceux des rides.

De ce fait, en présence de rides de sable, la détection de cibles peut devenir problématique même pour des opérateurs entraînés, et mène très souvent à des non détections ou à des fausses alarmes.

Pour une opération de chasse aux mines sous-marines, il serait donc utile d'obtenir des informations sur ces champs de rides de sable afin d'en tenir compte, soit directement dans la détection des cibles, soit dans une opération de replanification de la mission afin de permettre une exploration différente de la zone.

Pour ce faire, nous proposons ici de caractériser ces rides de sable par l'intermédiaire de trois informations, à savoir leur longueur d'onde, leur orientation et une information reliée à l'organisation structurelle du champ de rides via une détection des crêtes et des vallées.

Puisque nous avons réussi à localiser de manière précise ces champs de rides, nous proposons ici de nous focaliser sur ces zones, et d'estimer ces caractéristiques par l'intermédiaire des informations multi-échelles que nous avons à disposition.

#### 7.2.3.1 Longueur d'onde des rides de sable

Les rides de sable apparaissent grossièrement comme des textures périodiques de type sinusoïdal. On peut donc les caractériser par leur longueur d'onde. L'approche multi-échelle que nous proposons semble donc être pertinente pour caractériser le contenu fréquentiel d'un tel type de fond marin. En effet, elle permet de discrétiser le plan fréquentiel en différentes bandes passantes, où chacune définit un niveau de détail spécifique dans l'image sonar.

Au paragraphe 5.4.3, nous avons noté que plus un filtre passe-bande aura sa fréquence centrale proche de la fréquence des rides de sable, plus sa réponse énergétique sera maximale. C'est ce que nous pouvons constater à la Figure 7.10, où les réponses énergétiques des filtres

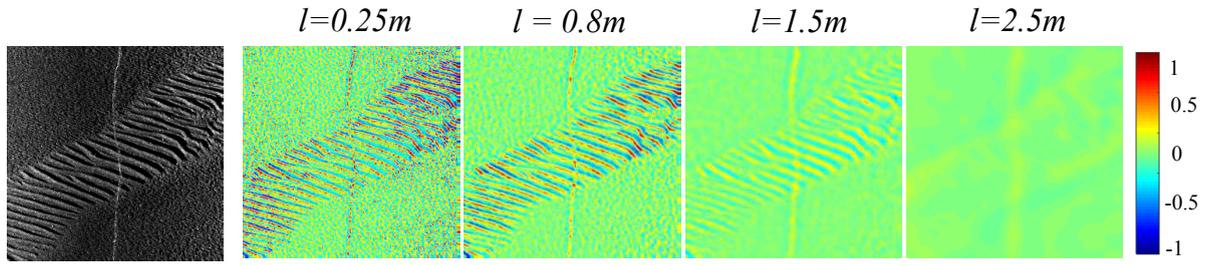


FIGURE 7.10: illustration des réponses énergétiques des filtres passe-bande DoP pour quatre niveaux de détail.

DoP sont présentées pour quatre niveaux de détail parmi les douze que nous utilisons dans cette étude. On y remarque que c'est le second niveau ici  $l = 0.8$  m) qui fournit la réponse maximale sur le champ de rides de sable. De ce fait, c'est la bande fréquentielle associée à ce niveau de détail qui fournit la longueur d'onde des rides de sable. Dans l'espace multi-échelle, nous appellerons ce niveau de détail spécifique "l'échelle des rides de sable".

D'autre part, à la Figure 7.11, nous pouvons constater qu'il n'y a pas seulement que la réponse énergétique des filtres passe-bande qui est maximale à l'échelle des rides de sable, mais également l'indice de confiance en la dimension intrinsèque de 1. Puisque cette information semble encore plus discriminante que la réponse maximale des filtres DoP, il est tout à fait possible d'estimer l'échelle des rides de sable par la relation suivante :

$$S(\mathbf{x}) = \operatorname{argmax}_{k \in [1, N]} \left( c_{i1D}^{(k)}(\mathbf{x}) \right), \quad (7.2)$$

où  $N = 12$  est le nombre d'échelles utilisées, et  $c_{i1D}^{(k)}(\mathbf{x})$  correspond à l'indice  $c_{i1D}$  de l'échelle  $k$  pour un pixel  $\mathbf{x}$  donné.

Une telle définition nous fournit alors l'indice du niveau de détail qui correspond le plus à la longueur d'onde des rides de sable. De plus, puisque cette information est locale (*i.e.* disponible en chaque pixel), elle permet de respecter la variabilité de la longueur d'onde des rides de sable à l'intérieur même d'un champ de rides.

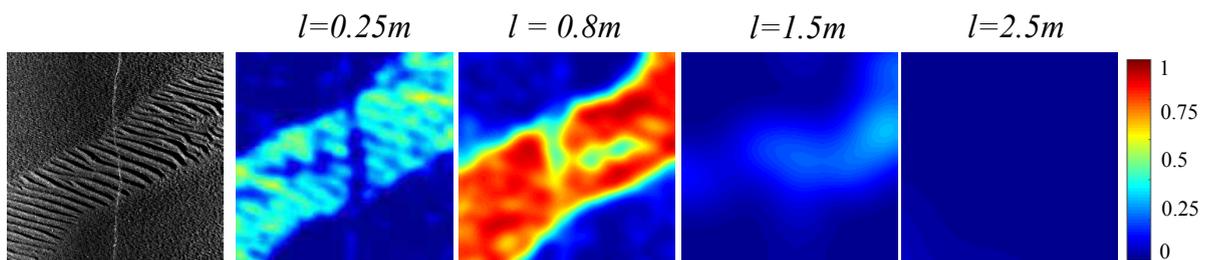


FIGURE 7.11: illustration des indices  $c_{i1D}$  pour quatre niveaux de détail.

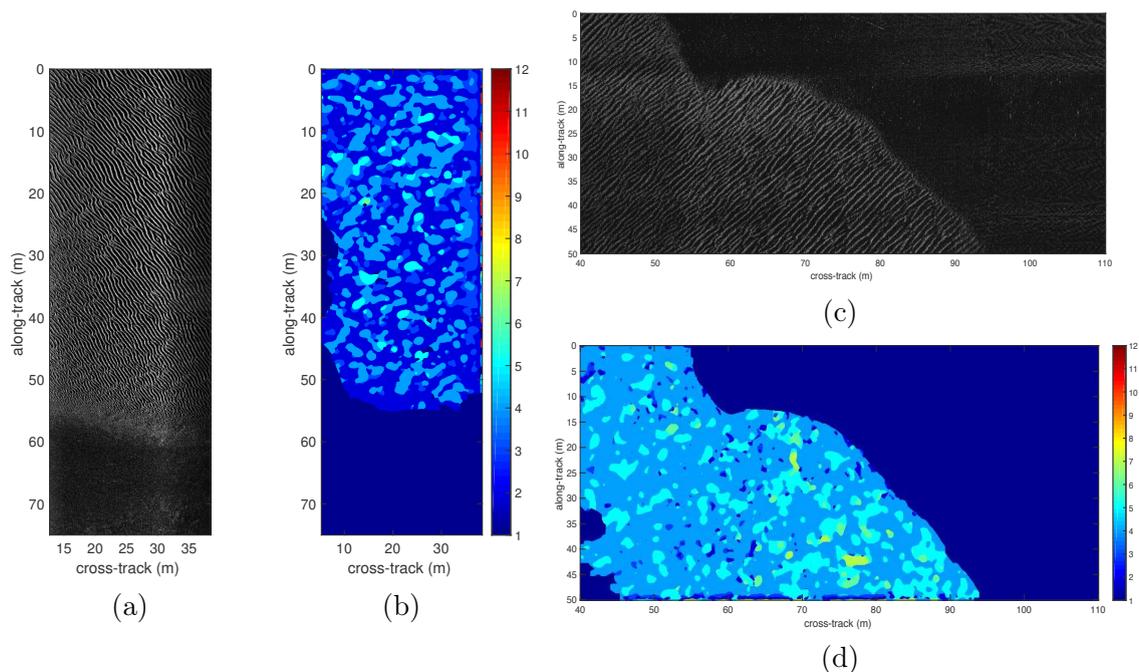


FIGURE 7.12: Image sonar issues de la base DSTO (a) et du CMRE (c) ainsi leurs cartes des échelles  $S$  respectives (b) et (d).

La Figure 7.12 illustre cette information pour des images sonars issues de la base de données DSTO et CMRE. On y remarque que l'échelle à l'intérieur des différents champs de rides de sable tournent autour d'une valeur de 3 et 4 sur les cartes (b) et (d) respectivement. Ceci correspond à des longueurs d'onde de l'ordre 0.5 m et 0.8 m. On observe également que les déformations à l'intérieur même des champs de rides provoquent des modifications dans l'estimation de ces informations. Ce phénomène est lié au caractère local de l'information.

L'avantage d'une telle carte est qu'elle fournit, pour chaque pixel du champ de rides de sable, l'indice du niveau de détail pour lequel les rides de sable peuvent être analysées de manière significative. En effet, ces indices correspondent aux filtres passe-bande qui ont le moins déformé le contenu informatif du signal 2D localement. Ainsi, si l'on désire obtenir davantage d'informations sur les rides de sable, il faudra tenir compte de cette notion de niveau de détail. C'est ce que nous allons maintenant illustrer par l'étude de l'orientation de ces structures sous-marines particulières.

### 7.2.3.2 Orientation des rides de sable

Dans l'étude sédimentologique que nous avons réalisé au paragraphe 6.2.3.2, nous avons pu constater que la formation des rides de sable est due aux effets des courants sous-marins. Ces courants possèdent une direction privilégiée qui se retrouve sur le motif des rides de sable. En effet, la succession d'échos et d'ombres visibles sur les images sonar se fait de manière parallèle à la direction du courant et fournit un motif global qui possède une orientation privilégiée dans

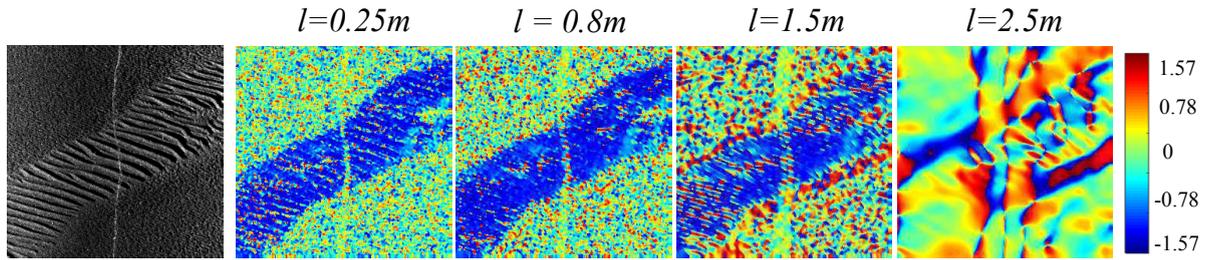


FIGURE 7.13: Illustration des orientations monogènes pour quatre niveaux de détail.

l'image. Une telle orientation contribue à la caractérisation du champ de rides et pourrait donc être une donnée intéressante.

Dans ce cas, l'information d'orientation locale fournie par le signal monogène peut être pertinente. La Figure 7.13 propose une illustration de cette information pour quatre des douze niveaux de détail de notre analyse multi-échelle. On y remarque alors que cette orientation semble bel et bien fournir l'information désirée. En effet, on observe aux trois premiers niveaux une forte homogénéité de l'orientation dans le champ de rides de sable. Cependant, c'est bien au second niveau ici (*i.e.* celui qu'on avait considéré comme étant l'échelle des rides de sable dans le paragraphe précédent) que l'on observe la meilleure définition des frontières et une homogénéité des valeurs d'orientation à travers cette zone. Au premier niveau, l'orientation locale est plus bruitée, alors qu'au troisième les frontières sont mal définies.

De ce fait, notre analyse multi-échelle nous procure cette même information d'orientation en chaque pixel de l'image et pour douze niveaux de détail distincts. Il est alors nécessaire de choisir la plus juste et de manière automatique. Nous allons donc exploiter l'information sur l'échelle des rides de sable que nous avons proposé précédemment.

En effet, nous pouvons supposer que c'est à l'échelle des rides de sable que l'orientation monogène sera la plus juste et la plus significative. Comme nous l'avons mentionné, le filtre passe-bande correspondant y est sensé moins filtrer l'information et donc permettre une caractérisation plus précise. Nous proposons donc une carte des orientations des rides de sable par l'expression suivante :

$$O(\mathbf{x}) = \theta^{(S(\mathbf{x}))}(\mathbf{x}), \quad (7.3)$$

où  $\theta^{(k)}$  représente l'orientation locale estimée au niveau de détail d'indice  $k$ . C'est-à-dire qu'en chaque pixel  $\mathbf{x}$ , nous choisirons la valeur d'orientation monogène donc l'indice a été déterminé dans la carte des échelles  $S$ .

Une telle information est illustrée sur la Figure 7.14 pour les mêmes images sonar que nous avons présentées au paragraphe précédent. On constate que ces valeurs sont comprises dans l'intervalle  $[-\pi/2, \pi/2]$ , et qu'elles sont globalement homogènes sur l'ensemble des zones anisotropes. Cependant, à l'intérieur de ces zones, les rides de sable peuvent parfois se croiser, se fusionner voire disparaître, ce qui fournit certaines valeurs "anormales" d'orientation.

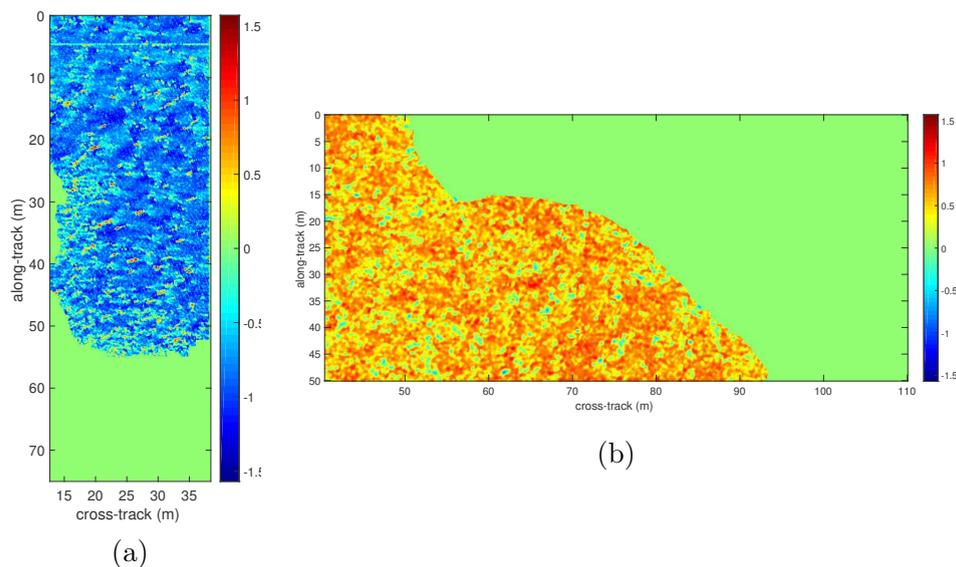


FIGURE 7.14: Carte d'orientation des rides de sables pour les deux images sonar de la Figure 7.12.

On peut alors concevoir que la valeur moyenne de ces orientations, pour un champ de rides de sable donné, peut également être une information utile, et en particulier pour la robotique autonome sous-marine. En effet, si un AUV possède cette information, il sera en mesure de replanifier une trajectoire, plus adaptée à l'analyse du champ de rides de sable. Cette nouvelle trajectoire permettra de réduire l'influence des rides de sable dans l'image, en passant par exemple dans une direction parallèle à leur orientation moyenne. De ce fait, la forte densité de motifs écho-ombre des rides de sable sera réduite et les objets potentiellement cachés auparavant seront plus facilement discernables comme le mentionne D.P. Williams dans [165].

La Figure 7.15 illustre cette possibilité de replanification. L'image sonar a été acquise par l'intermédiaire d'un premier passage de l'AUV, noté 1. La flèche bleue représente alors l'orientation moyenne à l'intérieur du champ de rides de sable, qui correspond à la direction

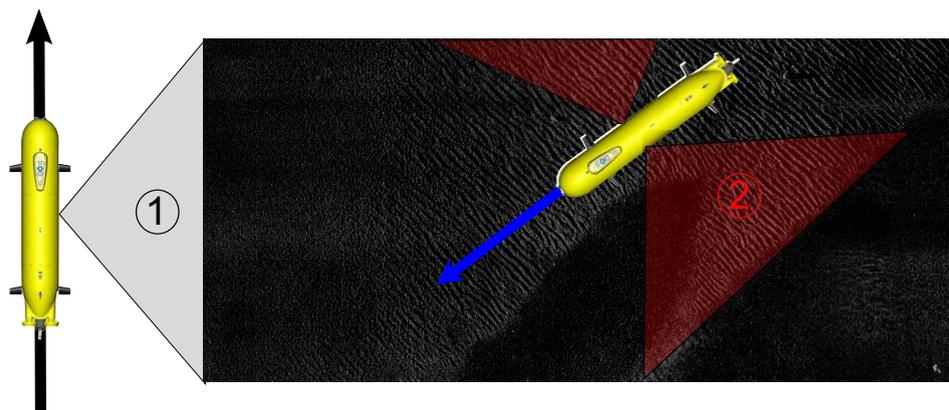


FIGURE 7.15: Schéma d'une possibilité de replanification dans les zones anisotropes.

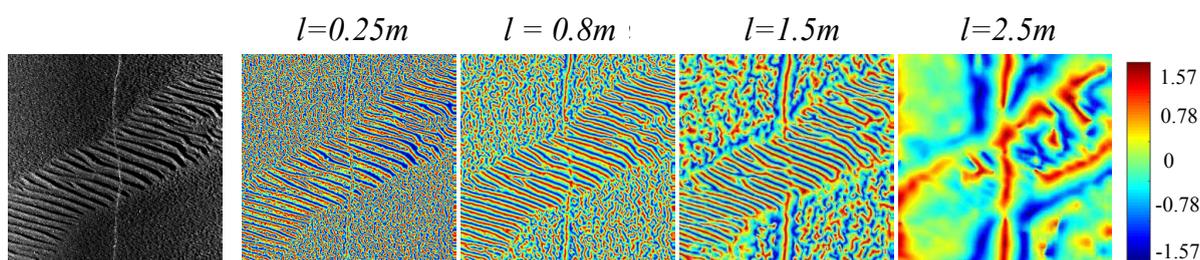


FIGURE 7.16: Illustration des phases monogènes pour 4 échelles parmi les 12 totales.

que l'AUV devra prendre lors d'un second passage si l'on désire réduire l'effet des rides de sable dans l'image, et ainsi inspecter leurs zones d'ombre.

Au delà des informations fréquentielles et géométriques que nous venons de présenter, nous allons voir que le signal monogène permet d'extraire de ces rides de sable une autre caractéristique reliée cette fois à la structure du champ de rides de sable.

### 7.2.3.3 Mise en évidence des crêtes et vallées des rides de sable

Si l'on se réfère à la Figure 7.16 qui illustre l'information de phase locale, on remarque que les ondulations caractéristiques d'un champ de rides de sable y ressortent. En effet, rappelons que M. Felsberg a défini cette phase locale comme une information structurelle. On remarque ici que c'est bien la structure générale du champ de rides de sable qui est mise en évidence, par l'intermédiaire de valeurs particulières de la phase locale en fonction des crêtes et des vallées de ces rides. L'avantage de cette information structurelle réside dans sa dynamique bornée dans l'intervalle  $[-\pi/2, \pi/2]$ , ainsi que dans son indépendance au contraste que nous avons mentionnée au chapitre 4.

Tout comme l'orientation locale, la phase locale est disponible en chaque pixel de l'image et pour chaque niveau de détail de notre analyse multi-échelle. Il est donc nécessaire de choisir celle qui sera la plus significative. Pour cela, nous allons à nouveau choisir pour chaque pixel de l'image la valeur de la phase qui correspond au niveau de détail qui lui est affectée dans la

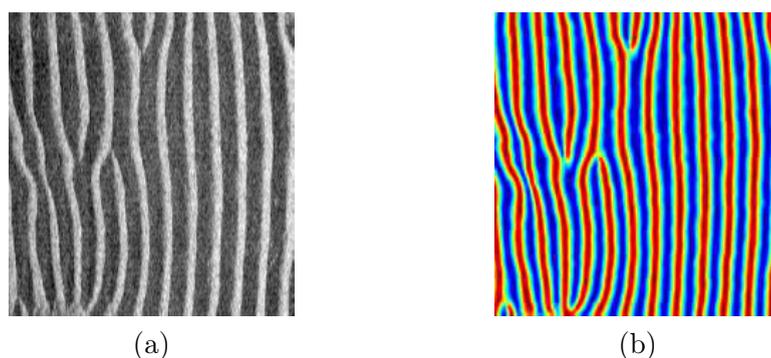


FIGURE 7.17: Représentation d'une image issue de la base de données DGA (a) et de sa carte  $P$  des phases monogènes.

carte  $S$  :

$$P(\mathbf{x}) = \phi^{(S(\mathbf{x}))}(\mathbf{x}). \quad (7.4)$$

La Figure 7.17 illustre une telle information pour une imagerie qui représente une zone de rides de sable. On observe alors clairement que les valeurs de cette carte  $P$  suivent les motifs induits par les rides de sable. La Figure 7.18 propose une coupe horizontale de ces deux imageries (en bleu celle de l'image initiale, et en rouge celle de la phase  $P$ ). Dans cette représentation, on remarque directement que la phase  $P$  possède une valeur maximale (*i.e.*  $\pi/2$ ) sur les crêtes et minimale dans le creux des vallées (*i.e.*  $-\pi/2$ ). On observe également sur cette illustration que le bruit n'impacte pas les valeurs de  $P$ .

On propose alors de classer la zone de rides de sable en termes de crêtes et de vallées par un simple seuillage de l'information de phase  $P$  :

$$C_r(\mathbf{x}) = P(\mathbf{x}) \geq 0 \quad (7.5)$$

$$V_a(\mathbf{x}) = P(\mathbf{x}) < 0 \quad (7.6)$$

Dans le cas de l'imagerie sonar utilisée à la Figure 7.17, nous obtenons la classification présentée à la Figure 7.19, où la transparence bleue désigne les zones considérées comme des crêtes, et la rouge comme les zones considérées comme des vallées. On remarque que cette classification est très précise et qu'elle permet de bien mettre en évidence la structure du champ de rides de sable.

Nous pourrions ici penser que le terme "vallée" se rapproche de la notion d'ombre. Or, puisque que le signal monogène réalise une séparation orthogonale entre les informations d'amplitude, d'orientation et de phase, aucune notion énergétique n'est portée par la phase monogène. Ainsi, il serait très risqué de parler ici de classification écho-ombre dans la zone de ride de sable.

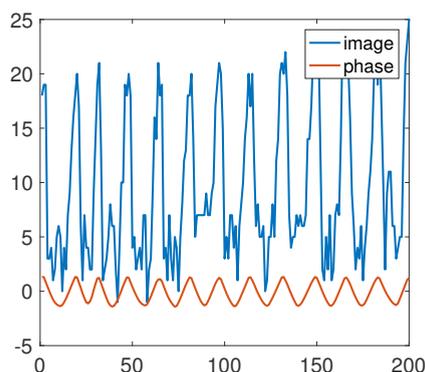


FIGURE 7.18: Coupes 1D horizontales de l'imagerie (a) (en bleu) et de sa carte  $P$  (b) (en rouge) de la Figure 7.17.

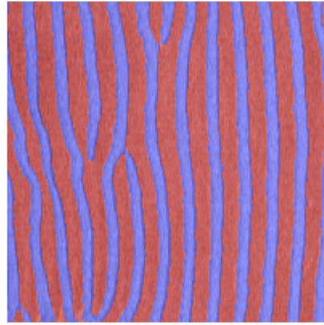


FIGURE 7.19: Classification d'une image sonar en termes de crêtes (en bleu) et de vallées (en rouge).

A travers cette première section, nous nous sommes attachés à caractériser deux classes de types de fonds, à savoir les fonds homogènes et anisotropes <sup>1</sup>. Nous avons alors constaté qu'il était possible d'en extraire des informations intéressantes pour le domaine de l'ATR comme la position des zones d'ombre ou des caractéristiques sur les champs de rides de sable. Dans la section suivante, nous allons maintenant proposer une première manière d'intégrer une telle caractérisation de l'environnement marin dans le processus ATR. En particulier, nous allons maintenant proposer une nouvelle architecture de ce dernier.

### 7.3 Architecture ATR adaptée au type de fond

Comme nous l'avons constaté au chapitre 3, le terme de chaîne ATR est consacré aux trois processus automatiques et successifs de détection, d'extraction de descripteurs et de classifications de cibles. Cependant, la variabilité de l'environnement marin impacte considérablement leurs performances. De ce fait, depuis quelques années, les recherches se sont focalisées sur la manière d'intégrer le contexte environnemental dans le processus ATR. Des post-traitements ont alors été mis en place en aval de la chaîne ATR afin de filtrer les fausses alarmes liées à l'environnement, ou de réaliser une classification des cibles plus performante (cf. paragraphe 3.3.3). Néanmoins, selon nous, cette intégration de l'environnement doit se faire en amont de la chaîne ATR, afin que celle-ci s'y adapte automatiquement.

#### 7.3.1 Prise en compte de la caractérisation environnementale

Si l'on souhaite intégrer des informations liées à l'environnement dans un processus ATR, il serait opportun que sa caractérisation soit utile pour chacune des phases qui la compose. De ce

---

1. La caractérisation des fonds complexes n'a pas été traitée au cours de ces travaux. Dans de futures études, elle nécessitera une attention particulière du fait de la forte variabilité des structures sous-marines qui pourront y apparaître. En effet, notre approche vise à définir comme complexe tout fond marin n'étant pas homogène ni anisotrope. De ce fait, des méthodes différentes seront nécessaires pour les caractériser et pour tenir compte de l'ensemble des scénarios possibles.

fait, nous proposons d'insérer la phase de description et de caractérisation de l'environnement, présentée au chapitre précédent, au tout début de la chaîne ATR, au plus près de l'image sonar initiale. Ainsi, les phases de détection et de classification des mines sous-marines pourront exploiter ces informations et adapter leurs paramétrages, leurs façons de détecter, voire de classifier selon le contexte de la mission. De plus, ces informations pourront servir aux véhicules autonomes qui pourront alors mieux comprendre l'environnement dans lequel ils se situent, et adapter leurs trajectoires afin d'exploiter le fond marin de la manière la plus adéquate possible.

### 7.3.2 Stratégie proposée

Une fois l'environnement marin décrit, il est nécessaire de trouver une stratégie pour que le reste de la chaîne ATR en tienne compte. Pour ce faire, nous proposons ici de tirer parti de la classification en trois classes que nous avons présentée au chapitre précédent. En effet, puisque les différents types de fond n'ont pas les mêmes impacts sur les processus ATR, nous proposons que ces derniers utilisent un algorithme spécialisé pour chacun de ces types de fond. Par conséquent, l'algorithme ATR s'appuiera sur les caractéristiques de chaque type de fond, afin de les prendre en considération et de les utiliser à bon escient. La Figure 7.20 propose l'illustration schématique d'une telle architecture. On y remarque qu'après la phase de description étudiée dans le chapitre précédent, le traitement se sépare en deux, à savoir une chaîne de traitement spécifique aux zones anisotropes (en bleue) et une autre spécialisée dans les zones complexes. Chaque étape de ces traitements pourra également utiliser la description de l'image sonar réalisée au préalable ou des éléments de cette description.

Puisque nous considérons les fonds marins homogènes comme des zones où il ne peut y avoir d'objet, aucun traitement y est réalisé. Ceci se traduit par une augmentation de la vitesse du traitement ATR. Cependant, les objets posés sur des fonds homogènes sont quant à

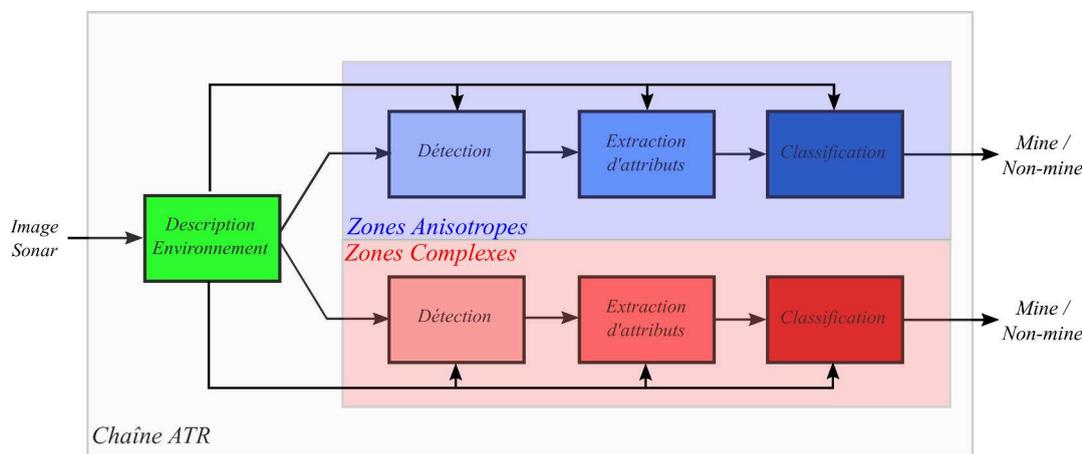


FIGURE 7.20: Architecture de la stratégie proposée.

eux classés comme complexes (cf. Figure 6.28), il sera donc nécessaire de tenir compte de cette observation dans la chaîne de traitement des fonds complexes.

Dans ces travaux, nous axons nos efforts de manière privilégiée dans les zones de rides de sable (l'étude des zones complexes pourra faire l'objet de futurs travaux). Comme nous l'avons vu, ces types de fonds marins sont des zones particulières qui nécessitent une grande attention pour des opérations de chasses aux mines. En effet, ils sont sources de taux de fausse alarme très élevés, mais également de fortes probabilités de non détection. De ce fait, nous proposons dans la section suivante, la réalisation d'une phase détection de mines sous-marines spécialisée dans les zones anisotropes.

## 7.4 Détection de cibles dans les zones anisotropes

Afin de mettre en place une détection spécifique pour les zones de rides de sable, nous proposons dans un premier temps d'observer les différences entre une zone de rides de sable qui ne présente pas de cible sous-marine et une autre où une mine y est posée. Nous verrons qu'à partir de telles observations il est possible de concevoir de nouvelles cartes qui mettront en évidence les échos et les ombres des cibles. Il sera alors plus aisé de réaliser leur détection. Ces différentes cartes se baseront sur la caractérisation des rides de sable que nous avons mis en place dans le paragraphe 7.2.3. Par conséquent, une telle méthode de détection s'adaptera automatiquement aux rides de sable observées.

### 7.4.1 Analyse des zones anisotropes avec et sans objets

A travers le chapitre 4 de ce manuscrit, nous avons mis en place un outil de représentation de l'information par le biais du signal monogène, qui permet de fournir diverses informations dans le cadre multi-échelle. Afin de constater les modifications potentielles de ces informations dans l'image sonar entre la présence et l'absence d'une mine sous-marine dans un champ de rides de sable, nous proposons d'illustrer ces informations dans les deux cas.

La Figure 7.21 présente les résultats du filtrage passe-bande DoP, de l'amplitude, de l'orientation et de la phase locale monogènes pour quatre des douze niveaux de détail dans le cas d'un fond anisotrope typique. Ici, on constate que ces rides de sable possèdent des réponses énergétiques plus importantes pour le niveau de détail  $l = 0.8$  m, ce qui correspond à ce que nous avons défini comme l'échelle des rides de sable (*i.e.* à la bande de fréquence qui correspond à leur longueur d'onde). A ce niveau de détail, on remarque que l'orientation locale est fortement homogène dans l'intégralité de l'imagette, ce que nous avons relié à l'orientation des rides de sable. Toujours à ce même niveau, on note que la phase locale met clairement en évidence la structure du champ de rides, ce que nous avons relié à la notion de crêtes et de vallées. Aux

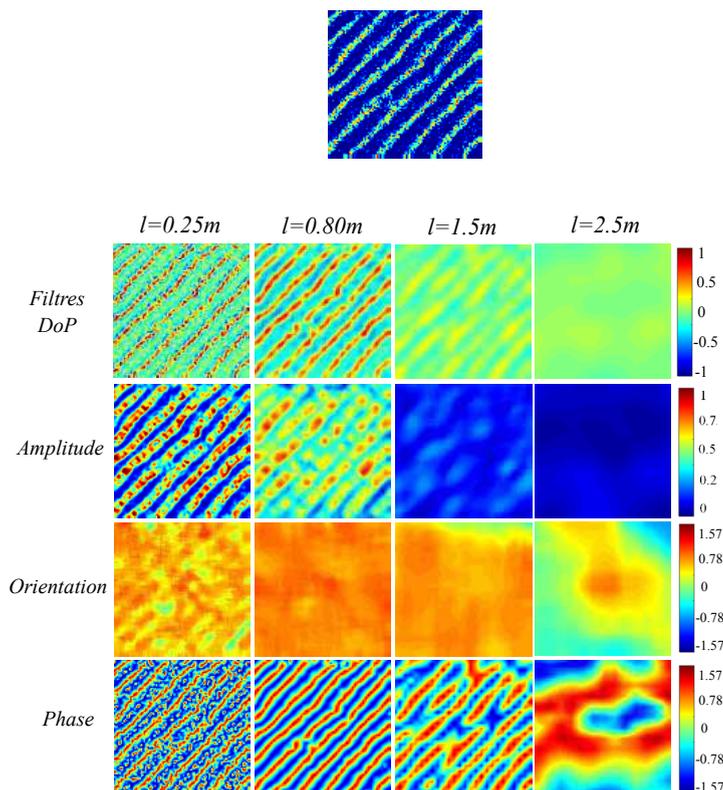


FIGURE 7.21: Représentation monogène de l'imagette d'un fond marin anisotrope.

niveaux de détail immédiatement supérieurs, les rides de sable sont progressivement filtrées et toutes ces informations perdent de leurs significations.

Maintenant, observons exactement les mêmes informations, mais cette fois pour une imagette représentant une mine située dans un champ de rides de sable. A partir de la Figure 7.22, nous pouvons remarquer que l'objet présent dans les rides modifie au niveau énergétique les observations précédentes mais également la structure du champ de rides de sable dans l'information de phase locale. En effet, au niveau des réponses énergétiques (*i.e.* des filtres DoP et de l'amplitude monogène), on constate que l'écho de la mine persiste sur un ensemble plus important de niveaux de détail que les rides de sable. Effectivement, jusqu'à l'échelle des rides de sable ( $l = 0.8$  m), l'écho de la cible est comme noyée dans les réponses énergétiques des rides de sable. Cependant, au delà, les rides de sable sont atténuées alors que l'écho de la cible persiste et la rend plus identifiable. Cette caractéristique peut s'expliquer par le biais de deux considérations. Premièrement, l'écho d'un objet manufacturé comme une mine sous-marine est plus intense que celui provoqué par une structure sableuse telle qu'une ride de sable, provoquant ainsi une réponse plus intense des filtres DoP, et qui s'atténue plus lentement. Deuxièmement, dans les images sonar, les mines sous-marines possèdent des échos dont les dimensions sont généralement supérieures aux échos des rides de sable. De ce fait, ils ne répondent pas sur les mêmes échelles. Si l'on observe maintenant l'information de phase locale, on remarque que l'ombre de l'objet rompt la structure du champ de rides de sable. Ceci est particulièrement

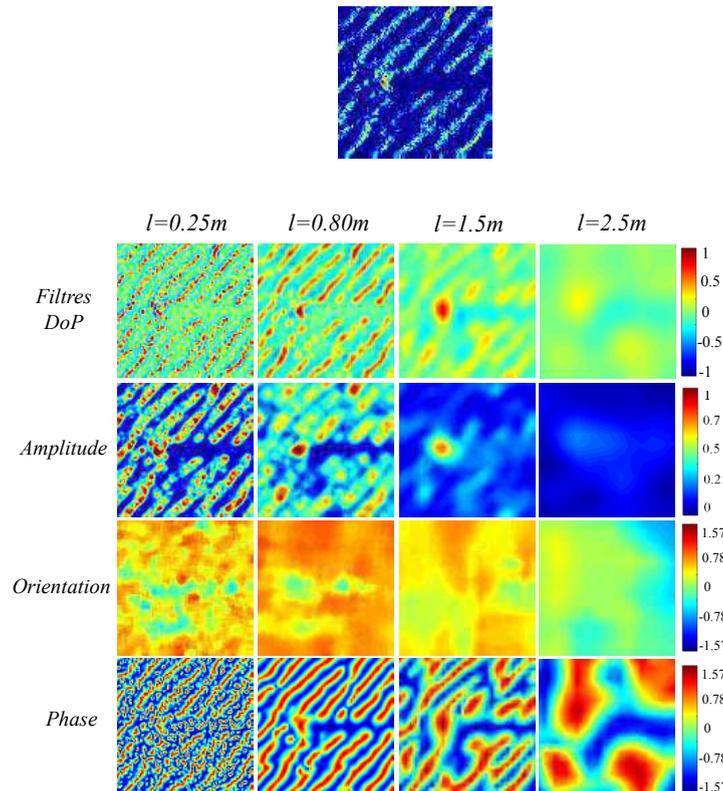


FIGURE 7.22: Représentation monogène de l'imagette d'un fond marin anisotrope où une mine sous-marine se situe.

visible au niveau de détail  $l = 0.8$  m. En effet, les rides de sable forment de longues crêtes quasi-parallèles qui sont interrompues par l'ombre de la cible. Cette dernière apparaît alors avec une phase locale proche de  $-\pi/2$ .

A partir de ces diverses remarques et observations, il semble donc possible de détecter l'écho et l'ombre d'un objet à partir des informations monogènes.

## 7.4.2 Détection d'échos

Nous avons pu constater que l'écho d'une cible sous-marine possède l'avantage de persister sur les échelles supérieures à celles des rides de sable. Afin de tenir compte de cette spécificité, nous pouvons utiliser la caractérisation des rides de sable que nous avons réalisé au paragraphe 7.2.3.1. Nous allons voir qu'elle permet de dresser une carte des échos dans laquelle, l'influence des rides de sable est diminuée, et où l'écho d'un objet est mis en évidence.

### 7.4.2.1 Carte des échos

Nous avons donc caractérisé les champs de rides de sable en termes de longueurs d'onde et donc de niveaux de détail. Ceci a permis de connaître le niveau de détail jusqu'auquel les rides de sable ont un impact énergétique dans l'analyse multi-échelle. Cette information étant

disponible localement, en chaque pixel, elle prend la forme d'une carte que nous avons notée  $S$ .

Puisque l'écho d'une cible persiste pour des niveaux de détail supérieurs à celles des rides de sable (dans l'information d'amplitude monogène), nous proposons d'en tenir compte par la définition d'une carte des échos  $E$  qui est la moyenne de l'amplitude monogène sur les niveaux de détail supérieurs à l'échelle des rides de sable :

$$E(\mathbf{x}) = \frac{1}{N - \hat{S}} \sum_{k=\hat{S}+1}^N A^{(k)}(\mathbf{x}), \quad (7.7)$$

où  $A^{(k)}(\mathbf{x})$  est l'amplitude monogène au pixel  $\mathbf{x}$  de l'échelle  $k$ , et où  $\hat{S}$  désigne la valeur moyenne des échelles des rides de sable dans zone anisotrope détectée.

La Figure 7.23 illustre le résultat d'une telle opération (b) pour une image sonar issue de la base de donnée CMRE (a). Cette image présente un champ de rides de sable dans lequel une mine sous-marine est posée (dans la partie inférieure de l'image, en milieu de portée). On remarque que la carte des échos  $E$  permet de mettre en évidence de manière plus évidente l'écho de cette cible, et réduit l'influence de la rétrodiffusion des rides de sable. Cependant, telle quelle, cette carte ne permet pas encore une détection simple et efficace de cet écho.

En effet, dans le but de procéder à cette détection par l'intermédiaire d'un algorithme classique comme le CFAR (de l'anglais *Constant False Alarm Rate* ou Taux de Fausses Alarmes Constant en français), il est nécessaire que les données vérifient certaines propriétés. En effet, la théorie concernant le détecteur CA-CFAR (CA - de l'anglais *Cell Averaging*) est très largement documentée dans le cas où la distribution des données est exponentielle comme le mentionne H. Rohling dans [166]. Ainsi, il est nécessaire de configurer nos données si l'on désire utiliser ce type détecteur. Afin de rendre la distribution de notre carte des échos  $E$  grossièrement exponentielle (aucun test d'ajustement n'a été opéré ici), nous proposons de calculer  $E^\alpha$ , où  $\alpha$  est l'exposant. La Figure 7.24 illustre alors ces cartes  $E^\alpha$  pour des valeurs de  $\alpha = 2$  et 3, ainsi

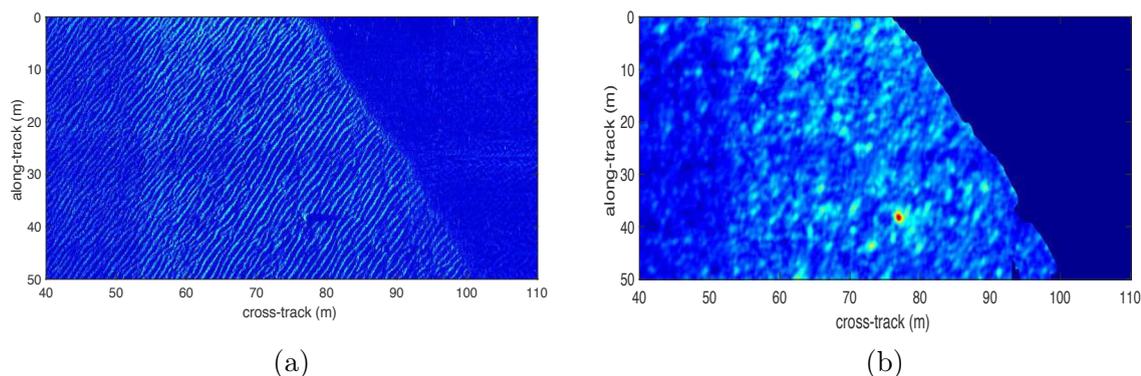


FIGURE 7.23: Image sonar SAS issues de la base de données CMRE (a) et sa carte des échos  $E$  associées (b).

que leurs histogrammes associés. On remarque dans ces cartes que la cible sous-marine apparaît de manière plus significative à mesure qu'on augmente la valeur de  $\alpha$ . De plus, visuellement, on constate que les distributions associées se rapprochent de plus en plus d'une loi exponentielle.

Les cartes  $E^\alpha$  proposent donc une nouvelle représentation d'un champ de rides de sable par une mise en évidence des anomalies qu'il contient. Elles pourraient donc être fournies directement à un opérateur pour l'aider à analyser une telle zone de fond marin rapidement. Cependant, puisque cette thèse se positionne nettement dans un contexte d'automatisation de chaîne de traitement, nous proposons maintenant d'utiliser un détecteur de type CA-CFAR sur la carte  $E^3$  afin de détecter automatiquement l'écho d'une mine sous-marine.

#### 7.4.2.2 Le détecteur CA-CFAR

Les algorithmes CFAR ont été introduits pour pallier la difficultés de détermination des seuils de détection. En effet, alors qu'en l'absence de bruit de fond dans les données un seuil global peut facilement permettre de réaliser une tâche de détection, cette dernière devient critique lors de la présence d'un bruit important. Afin de maintenir des performances constantes dans un environnement non-stationnaire, la phase de détection doit estimer localement les paramètres de ce bruit pour y fixer un seuil local. Pour cela, on utilise généralement une fenêtre glissante qui détermine le seuil de détection en chaque pixel.

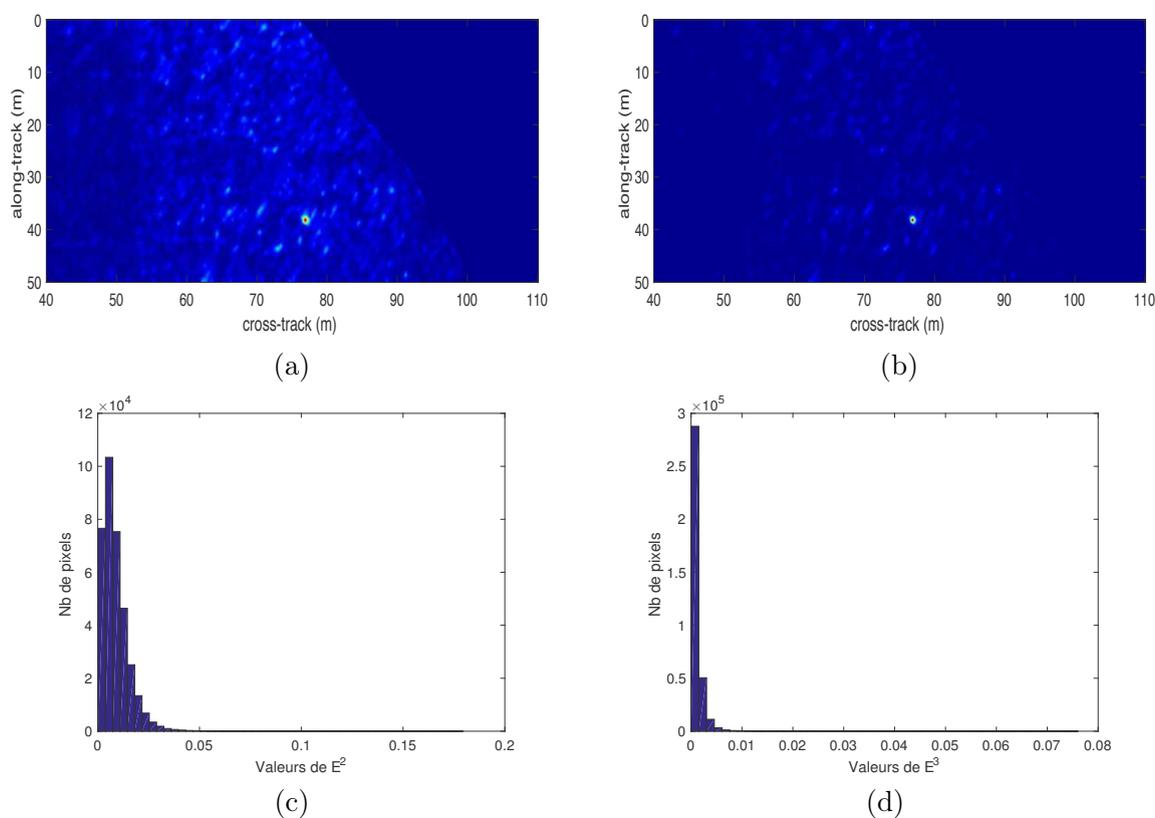


FIGURE 7.24: Cartes  $E^2$  (a) et  $E^3$  (b) ainsi que leurs distributions associées en (c) et (d).

En 2D, une telle fenêtre d'analyse est constituée de deux ensembles comme le montre la Figure 7.25. Autour du pixel courant, le premier ensemble est appelé fenêtre de garde, l'autre fenêtre de test. La fenêtre de garde permet de réduire l'influence d'une cible étendue dans l'estimation des paramètres statistiques qui se fait quant à elle dans la fenêtre de test.

La détection CFAR, partant du principe que la puissance du bruit dans les images est inconnue, propose de l'estimer dans la fenêtre de test et de fixer le seuil de détection  $T$  par la relation :

$$T = \beta P_n, \quad (7.8)$$

où  $P_n$  correspond à la puissance du bruit estimée, et  $\beta$  est le facteur de seuil qui permet d'obtenir une probabilité de fausses alarmes constante.

De nombreuses versions du détecteurs CFAR existent comme le mentionne M. Richards dans [167]. Cependant, parmi eux, c'est le détecteur CA-CFAR qui est certainement le plus employé du fait de sa simplicité. De plus, il est généralement utilisé comme une base de comparaison pour les autres détecteurs. Pour un CA-CFAR, l'estimation de la puissance du bruit autour du pixel courant est fourni par la relation suivante :

$$P_n = \frac{1}{N} \sum_{j=1}^N x_j^2, \quad (7.9)$$

où  $N$  est le nombre de pixels dans la fenêtre de test et  $x_j$  la valeur en amplitude du pixel d'indice  $j$  dans cette fenêtre. M. Richards montre que pour obtenir un détecteur dont la probabilité de fausse alarme est constante, le facteur  $\beta$  doit nécessairement vérifier l'équation suivante :

$$\beta = N \left( Pfa^{-1/N} - 1 \right), \quad (7.10)$$

où  $Pfa$  est la probabilité de fausse alarme désirées.

Cependant, cette dernière expression n'est valide que sous certaines hypothèses, et en particulier que les données ont une distribution exponentielle.

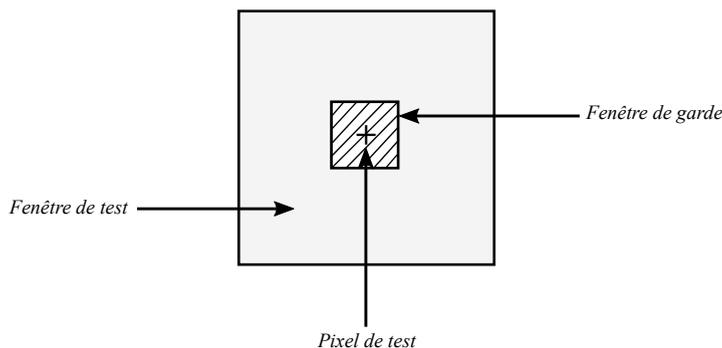


FIGURE 7.25: Configuration de la fenêtre TFAC.

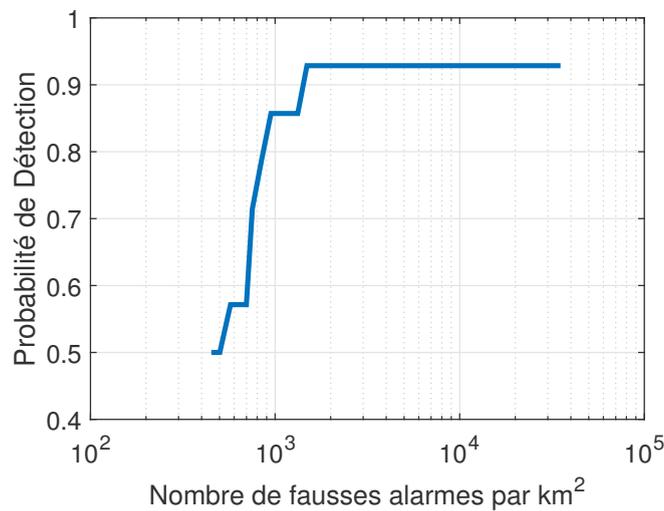


FIGURE 7.26: Courbe ROC du détecteur d'écho proposé. L'axe des abscisses représente le nombre de fausses alarmes par km<sup>2</sup> de rides de sable, alors que l'axe des ordonnées représente la probabilité de détection.

#### 7.4.2.3 Résultats de la détection des échos

L'hypothèse de distribution exponentielle des données est la raison pour laquelle il n'est pas possible d'appliquer le détecteur directement sur la carte des échos  $E$ , mais plutôt à son cube  $E^3$ . Pour réaliser cette détection, il est toutefois nécessaire de fixer les dimensions de la fenêtre de test et de garde. Pour cela, nous proposons de tenir compte du fait que les mines sous-marines "classiques" (à savoir les mines cylindriques, Manta, sphériques, Rockan etc.) possèdent des dimensions qui n'excèdent pas 2.5m de long ou de large. Ainsi, la fenêtre de garde aura une taille de 2.5m x 2.5m centrée autour du pixel de test (afin d'être certain que tous les types de mines sous-marines n'en sortiront pas), alors que la fenêtre de test aura quant à elle une taille de 5m x 5m centrée, elle aussi, autour du pixel de test.

Un autre paramètre doit également être fixé, il s'agit de la probabilité de fausse alarme  $Pfa$  qui apparaît dans l'équation (7.10). Pour cela, nous proposons de construire une courbe semblable à une courbe ROC (de l'anglais *Receiver Operating Characteristic*), où la probabilité de fausses alarmes du détecteur est remplacée par le nombre de fausses alarmes par km<sup>2</sup> de rides de sable (sur l'axe des abscisses) (cf. Figure 7.26). Ce choix de représentation permet d'obtenir des informations plus concrètes sur le détecteur, et favorisera sa comparaison avec les méthodes existantes. Afin de réaliser une telle courbe, les 360 images de la base de données CMRE sont utilisées. A travers cet ensemble, de nombreuses images sonar font intervenir des champs de rides de sable aux orientations et longueurs d'onde variées. L'algorithme de détection d'échos, proposé dans cette section, est alors exécuté automatiquement sur ces images, pour différentes valeurs de  $Pfa$  entre  $10^{-12}$  à  $10^{-1}$ . On constate alors que la probabilité de détection augmente avec le nombre de fausses alarmes par km<sup>2</sup>.

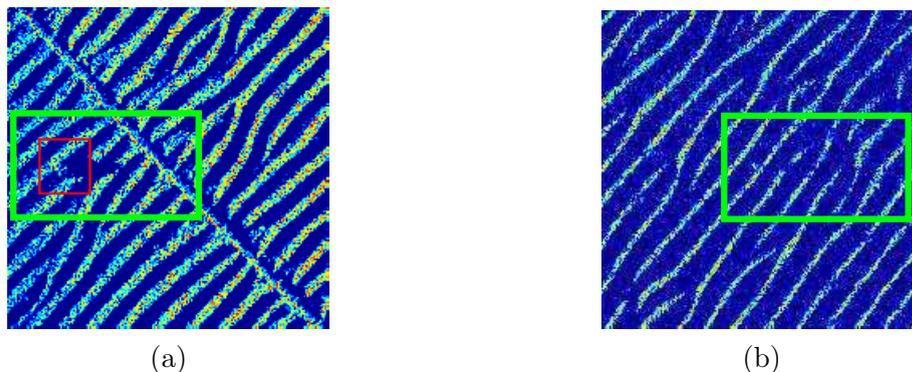


FIGURE 7.27: Exemples de la difficulté de détection d'un écho. En vert la vérité terrain, en rouge le résultat de détection de l'algorithme proposé.

Nous avons pu remarquer au cours de nos différents tests que pour une valeur de  $Pfa = 10^{-5}$ , on obtient de manière globale une probabilité de détections de 94%, pour un nombre moyen de fausses alarmes de 2500 par  $\text{km}^2$  de rides de sable. Si l'on ramène cette valeur au nombre d'images dans la base de données, on obtient un nombre moyen de 5 fausses alarmes par image sonar. Parmi les 6% de cibles qui sont manquées, on recense un certain nombre de cas où la détection d'un écho sera très compliquée, y compris pour un œil expérimenté comme nous pouvons le voir sur la Figure 7.27. Alors que sur l'imagette présentée en (a), l'écho de l'objet est clairement visible et détecté par l'algorithme (entouré en rouge), sur l'imagette (b) l'objet est rendu pratiquement invisible par le champ de rides de sable. La détection de l'écho est alors mise en échec.

La Figure 7.28 illustre le résultat de la phase de détection d'échos pour deux images sonar entières issues de la base de données du CMRE. Sur ces images, les vérités terrains sont entourées en vert, et les résultats de la détection d'écho sont pointés en rouge. Cette Figure permet alors de constater que l'on a bien une détection efficace des mines sous-marines dans le champ de rides de sable, mais que ces résultats sont encore légèrement ternis par l'apparition d'un certain nombre de fausses alarmes.

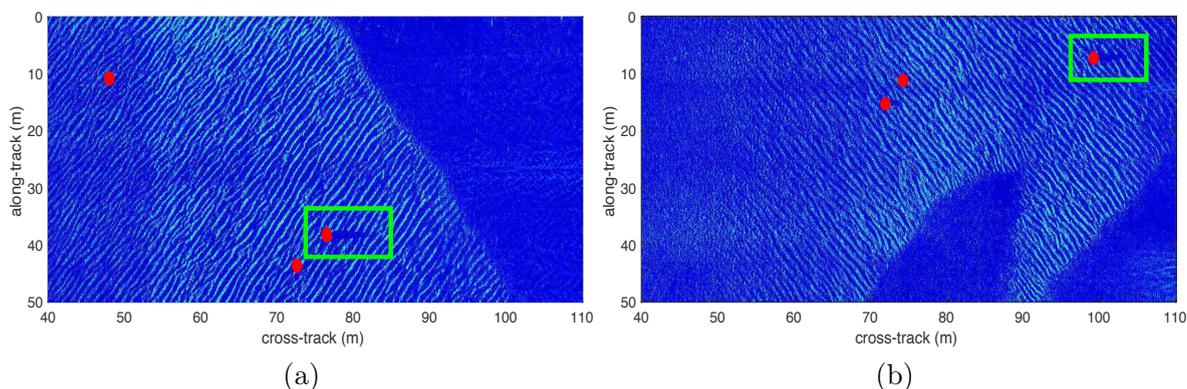


FIGURE 7.28: Résultats de la détection d'échos pour deux images sonar de la base de données CMRE. La vérité terrain est entourée en vert alors que les résultats de la détection d'échos sont pointés en rouge.

Cependant, cette phase de détection n'a pour l'instant tenu compte que d'une seule des caractéristiques d'une cible en imagerie sonar, à savoir son écho. Or, son ombre est également une information pertinente à prendre en considération. C'est ce que nous allons maintenant intégrer dans la phase de détection.

### 7.4.3 Détection d'ombres

Dans le paragraphe 7.4.1, nous avons pu constater que l'écho d'une cible dans un champ de rides de sable possède un caractère particulier, qui permet de la mettre en évidence parmi les échos des rides de sable. Concernant l'ombre de ces cibles, nous avons remarqué qu'elle possède également un aspect particulier dans la représentation monogène multi-échelle. En effet, l'ombre d'une cible peut être caractérisée par une valeur de la phase locale monogène très faible, autour de  $-\pi/2$ , à l'échelle des rides de sable. De ce fait, dans une zone anisotrope, l'ombre d'une cible possède la même valeur de phase locale que ce que nous avons défini comme les "vallées" des rides de sable dans le paragraphe 7.2.3.3. Par conséquent, si l'on considère la classification en termes de crêtes et de vallées que nous avons mis en place dans ce même paragraphe, l'ombre d'une mine sous-marine sera considérée comme une vallée. Ce phénomène est illustré sur la Figure 7.29 qui présente, pour deux imageries sonar (a) et (d) (issues de la base de données CMRE), leurs phases locales monogènes respectives (à l'échelle des rides de sable) en (b) et (e), et leurs classifications en termes de crêtes et de vallées en (c) et (f). On y constate alors que l'ombre des objets est considérée comme une vallée (en rouge), et que

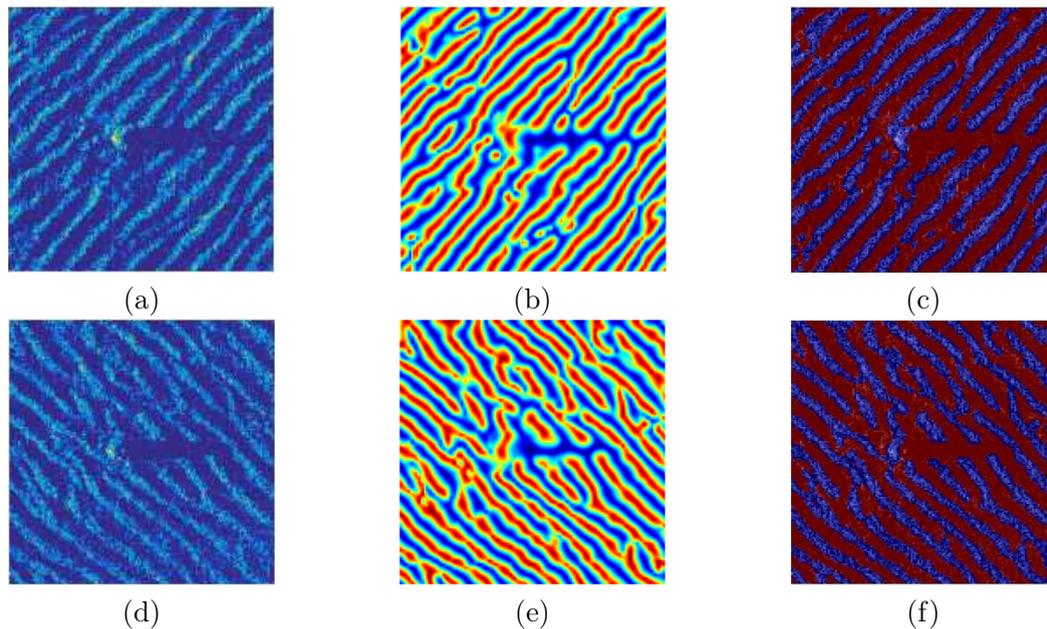


FIGURE 7.29: Classification en termes de crêtes (en bleu) et de vallées (en rouge) pour deux imageries sonar (a) et (d), où une mine sous-marine est présente à l'intérieur du champ de rides de sable. En (b) et (e), l'information de phase monogène est également représentée.

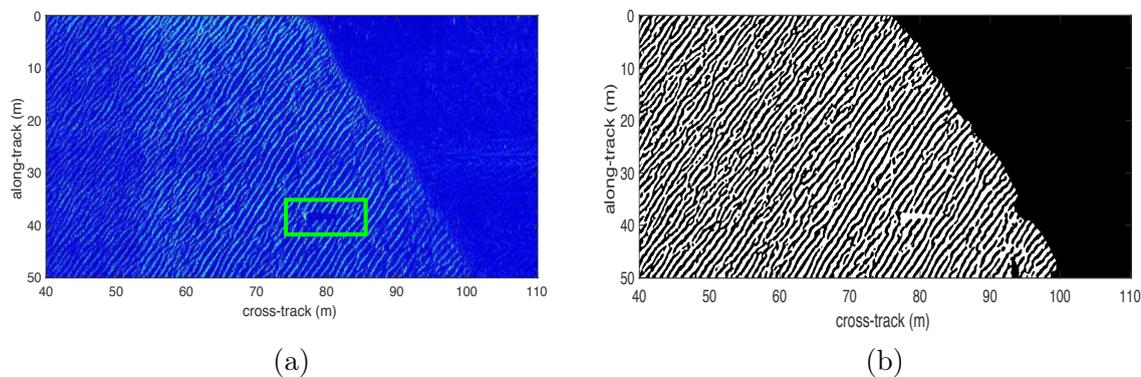


FIGURE 7.30: Image sonar issues de la base de données CMRE et sa carte des vallées associée à la zone de rides de sable.

par la même occasion, l'écho de la cible est quant à lui classé en tant que crête (en bleu). Par conséquent, cette classification spécifique aux rides de sable permet de détecter les vallées des rides, mais également l'ombre des objets qui y sont déposés. Certes, une telle approche ne permet pas une discrimination entre l'ombre des objets et les vallées des rides, mais puisque les vallées des rides de sable sont généralement liées à des zones d'ombre (comme nous pouvons l'observer sur les différentes imageries proposées au cours de cette section), nous confondrons de manière volontaire la notion de vallée et d'ombre, uniquement dans le contexte très particulier d'une détection de cibles.

De ce fait, la carte des vallées d'une zone anisotrope sera considérée comme une détection des ombres, ce que nous illustrons sur la Figure 7.30. De telles cartes permettent de constater que l'ombre associée à un objet se distingue par une élongation plus importante que celles de rides de sable, et nous allons voir maintenant que c'est cette information de longueur qui sera essentielle pour éliminer une grande partie des fausses alarmes du détecteur d'échos précédent.

#### 7.4.4 Résultats finaux de la détection

A travers cette section, nous avons mis en place une détection des échos et des ombres, chacune spécifique aux zones anisotropes, et qui s'appuient toutes deux sur une caractéristique locale de la zone, à savoir l'échelle des rides de sable. Nous avons pu remarquer que la détection d'échos est performante, mais fournit encore un nombre de fausses alarmes qu'il serait possible de diminuer en prenant en compte l'élongation des ombres. Pour ce faire, nous proposons que chaque contact issu du détecteur d'écho soit associé à son ombre dans la carte des vallées. De cette manière, il est possible d'estimer la hauteur de ce dernier.

En effet, si l'on considère un contact à une distance  $R$  du sonar (*i.e.* sur l'axe *cross-track*), pour une altitude du porteur  $H$  au dessus du fond, et une ombre dont la distance extrême est

$S$  (cf. Figure 7.31), la hauteur  $h$  de l'objet détecté est alors donné par la formule suivante :

$$h = H \left( \frac{S - R}{S} \right). \quad (7.11)$$

Ainsi, pour chaque contact, il est possible d'estimer sa hauteur et donc d'y appliquer une décision quant à sa nature de fausse alarme ou non. Dans ces travaux, nous proposons d'appliquer un seuil sur cette information de hauteur. En effet, si l'on se réfère à [130], les mines sous-marines ont pour une grande majorité des hauteurs minimales de 0.25m. De ce fait, nous fixons une telle valeur de seuil pour nos détections.

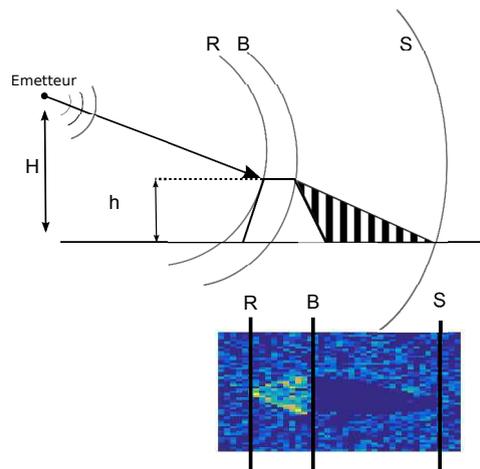


FIGURE 7.31: Configuration pour l'estimation de la hauteur d'un objet.

La Figure 7.32 présente deux résultats de détection qui mêlent la détection d'échos dans les zones anisotropes et la prise en compte de l'ombre de ces dernières. On remarque alors pour ces différents exemples que le nombre de fausses alarmes est nul et que la probabilité de détection est de 100%.

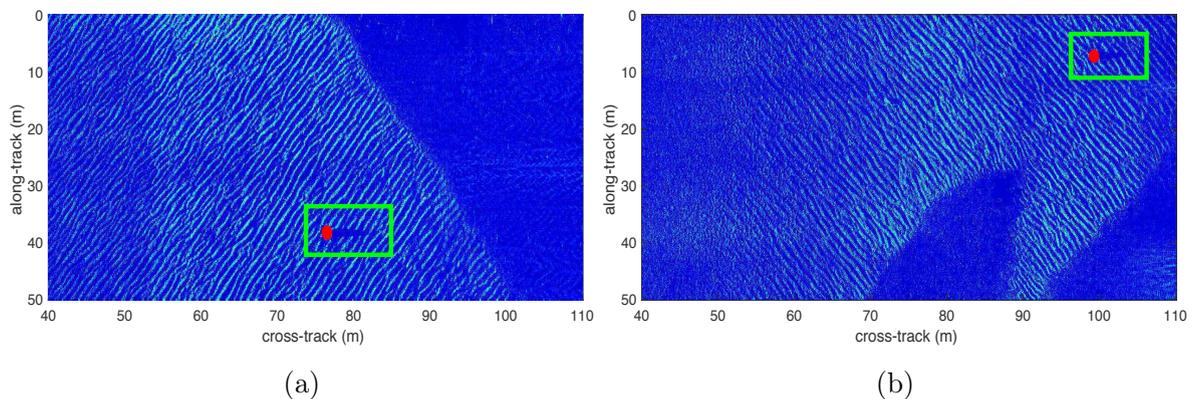


FIGURE 7.32: Exemple de détections finales sur des images sonar issues de la base de données CMRE : En rouge les détections de l'algorithme, en vert, les vérités terrain.

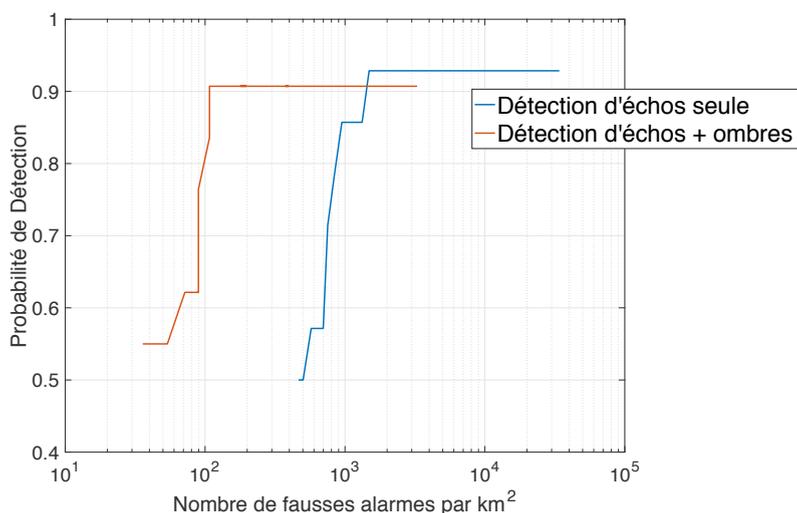


FIGURE 7.33: Courbe ROC du détecteur proposé (en rouge) et celle du détecteur d'écho seul (en bleu). L'axe des abscisses représente le nombre de fausses alarmes par km<sup>2</sup> de rides de sable alors que l'axe des ordonnées représente la probabilité de détection.

Afin de qualifier ce détecteur, nous proposons de réaliser sa fonction ROC sur l'ensemble de la base de données CMRE. La Figure 7.33 représente la probabilité de détection de notre algorithme en fonction du nombre de fausses alarmes par km<sup>2</sup> de rides de sable. A des fins de comparaison, nous y présentons également la courbe du détecteur d'échos seul (vu à la Figure 7.26). On constate que le nombre de fausses alarmes par km<sup>2</sup> est fortement réduit lorsque l'on prend en compte la notion d'ombre dans la détection.

Précédemment nous avons vu que le détecteur d'échos seul propose un nombre moyen de fausses alarmes de 2500 par km<sup>2</sup> pour un paramètre  $Pfa = 10^{-5}$ . Ici, la prise en compte de l'ombre dans le processus réduit ce nombre à environ 250 par km<sup>2</sup>, ce qui représente une division par 10 du nombre de fausses alarmes. Cependant, on remarque que l'ajout d'une phase d'estimation des ombres ne permet pas d'atteindre les mêmes performances en termes de

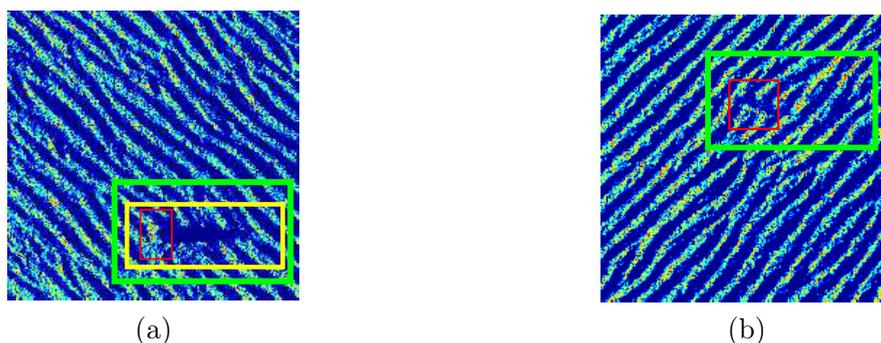


FIGURE 7.34: Exemples de la difficulté de détection d'une ombre. En vert la vérité terrain, en rouge le résultat de la détection d'échos, et en jaune le résultat de la détection d'échos et d'ombres.

probabilité de détection que pour une détection d'échos uniquement. En effet, on passe d'une asymptote atteignant 94% à une asymptote aux alentours de 91%, ce qui reste tout de même un bon résultat. Cette diminution provient de la nature ondulée des rides de sable qui peut potentiellement impacter la forme de l'ombre derrière un objet. De ce fait, l'estimation de la hauteur de l'objet peut être biaisée, et mener à un manque lors de la détection finale. La Figure 7.34 illustre cette difficulté par l'intermédiaire de deux exemples. Alors que pour l'imagette (a), l'objet laisse apparaître une ombre très importante qui lui permet d'être détecté, pour l'imagette (b) ce n'est pas le cas. Ainsi la détection de l'écho, pourtant correctement réalisée (entourée en rouge), n'est pas validée par la présence d'une ombre suffisante.

La Figure 7.35 propose enfin une comparaison des performances de notre détecteur par rapport à celles du détecteur de D.P. Williams [8], créé spécifiquement pour ces données. Nous y présentons également les performances du détecteur d'échos seul. On y constate que les différentes probabilités de détection sont extrêmement proches entre ces algorithmes (a), mais qu'une différence importante se situe dans le nombre de fausses alarmes au  $\text{km}^2$  qu'ils fournissent (b). En effet, nous obtenons une diminution d'un rapport 7.

Ces différentes observations permettent de conclure que la prise en compte du type de fond marin est d'une aide significative dans une phase de détection, et particulièrement d'un point de vue de la réduction du nombre de fausses alarmes.

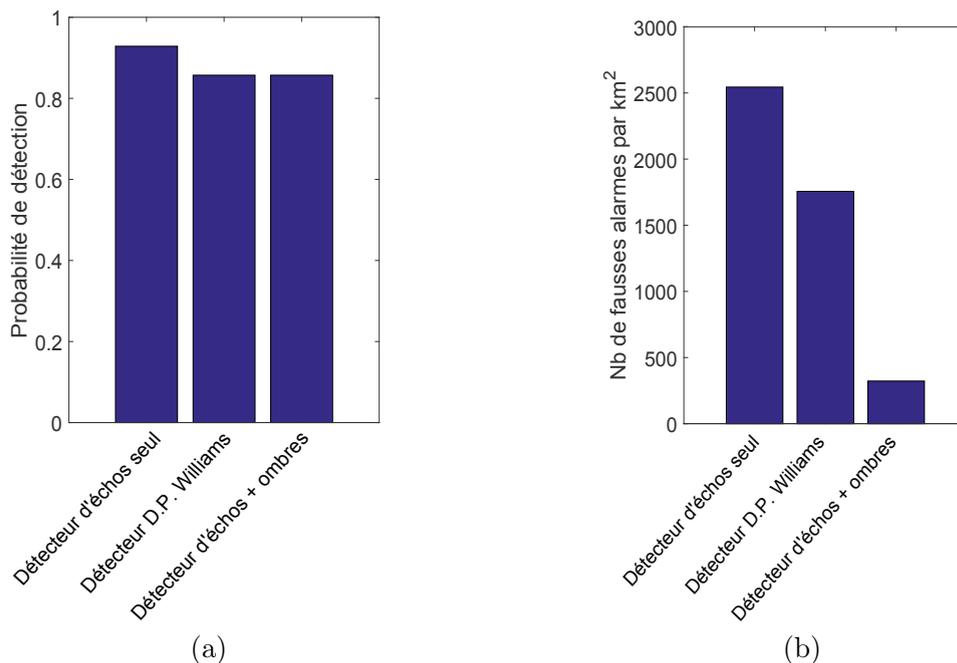


FIGURE 7.35: Comparaison de la probabilité de détection et du nombre de fausses alarmes par  $\text{km}^2$  entre le détecteur d'écho seul, le détecteur d'écho et d'ombre que nous proposons dans cette étude et le détecteur de D. Williams [8].

## 7.5 Conclusion et perspectives

Suite au chapitre 6 que nous avons axé sur la description de l'environnement dans l'image sonar, nous nous sommes intéressés à travers ce chapitre à sa caractérisation puis à son intégration dans le cadre de l'ATR.

Pour ce faire, nous nous sommes tout d'abord attachés à caractériser les fonds marins homogènes et anisotropes afin d'en ressortir des informations significatives en vue d'un processus ATR. Ensuite, nous avons proposé de définir une nouvelle stratégie de construction d'une chaîne ATR. Celle-ci part du constat que chacun des trois types de fonds marins nécessite une attention particulière et n'affecte pas les performances d'une chaîne ATR de la même manière. Nous avons alors proposé une architecture qui utilise une détection spécifique pour chacun d'entre eux.

Dans cette optique, nous avons défini un détecteur spécifique aux zones anisotropes qui permet une extraction de l'écho des cibles et de leur ombre en s'appuyant sur la caractérisation des rides de sable précédente. De ce fait, notre détection est non seulement adaptée au type de fond, mais s'adapte également automatiquement aux caractéristiques du champ de rides de sable traité. Nous avons alors pu constater que ce détecteur permet d'obtenir des performances bien supérieures, dans les zones de rides de sable, aux détecteurs pouvant intervenir dans la littérature, et en particulier celui proposé par D.P. Williams du CMRE.

Les travaux futurs devront maintenant se concentrer sur la réalisation d'un détecteur spécifique aux zones de fond marin complexes, qui nécessitera des procédés plus coûteux et plus complexes du fait de la grande variété des structures sous-marines susceptibles d'y intervenir.

## Chapitre 8

# Conclusion générale

Ces travaux de thèse se situent dans le contexte particulier de la détection et de la reconnaissance automatique de cibles sous-marines. Depuis maintenant plusieurs décennies, les chaînes ATR (de l'anglais *Automatic Target Recognition*) sont l'objet d'attentions particulières par les opérationnels de la chasse aux mines sous-marines. En effet, la dangerosité et le caractère répétitif de leurs missions poussent les différentes marines à aider ou remplacer l'opérateur humain lors de l'analyse des images du fond marin et de la détection des mines sous-marines.

Cependant, les performances des chaînes ATR n'atteignent pas encore des niveaux suffisants pour espérer les rendre totalement opérationnels. Le principal frein à ce développement est identifié par la communauté comme étant la variabilité de l'environnement marin. Ce dernier apparaît sur les images sonar du fait du relief ou de la nature du fond marin, en impactant directement l'onde acoustique rétro-diffusée. De ce fait, l'environnement devient une source potentielle de fausses alarmes sur l'image sonar. De plus, la signature même d'une cible sous-marine peut également se retrouver modifiée par cet environnement et potentiellement s'y confondre. Dans ce cas, la phase de détection mène généralement à un échec qui peut être très critique pour la sécurité et la sûreté des voies de navigation.

Pour remédier à ce problème, l'approche classique consiste le plus souvent à tenter de gommer les effets de cet environnement par différents outils de traitement d'image tels que des filtrages et de la segmentation. L'idée est alors de conserver uniquement l'information dite "utile" pour la reconnaissance de cible, *i.e.* les zones d'échos et d'ombres. Cependant, après des années d'efforts consacrés à la recherche d'une chaîne capable de s'affranchir de tous les artefacts dus à l'environnement d'une cible, force est de constater qu'une chaîne universelle n'existe pas.

L'objectif de ces travaux thèse était de repenser le problème en considérant l'environnement marin comme une source d'informations, et de trouver la manière de l'intégrer dans processus de reconnaissance automatique de cibles marines.

Dans un premier temps, afin de pouvoir extraire le maximum d'informations sur l'environnement marin à partir d'images sonar, nous avons proposé une nouvelle représentation de ces dernières en se basant sur l'étude du signal monogène. En tant qu'extension multidimensionnelle du signal analytique, cet outil nous permet de séparer orthogonalement les informations énergétiques, géométriques et structurelles d'une image dans un cadre multi-échelle. L'intégration de ce nouvel outil au contexte particulier de l'analyse d'images sonar est la première contribution de ces travaux. Nous avons alors démontré l'intérêt et le potentiel des informations locales fournies par le signal monogène à travers une opération de classification supervisée des images sonar. De plus, puisque les informations monogènes sont extraites à partir d'opérateurs isotropes (cf. transformée de Riesz lors du calcul du signal monogène ou d'utilisation des filtres DoP lors de l'analyse multi-échelle), le nombre de paramètres à fixer lors de l'analyse d'une image est réduit par rapport aux méthodes plus classiques telles que les ondelettes ou les bancs de filtres orientés. Nous avons alors constaté que l'utilisation de cet outil sur des données issues de systèmes sonars différents est facilitée.

Afin de relier toutes ces nouvelles informations à des notions concrètes sur les fonds marins, nous avons ensuite proposé d'utiliser le concept de dimension intrinsèque dans sa version continue. La dimension intrinsèque continue a deux avantages majeurs qui nous ont poussé à s'y intéresser. Premièrement, elle permet de décrire la complexité de la texture d'une image sous la forme de descripteurs continus et appartenant à une théorie mathématique formelle. Deuxièmement, l'estimation de ces descripteurs peut être réalisée directement à partir des informations issues du signal monogène. Nous avons alors proposé d'utiliser le cadre multi-échelle fourni par ce dernier pour définir des dimensions intrinsèques continues multi-échelles. A partir d'une analyse de ces informations, nous avons proposé trois nouveaux descripteurs de l'image sonar qui renseignent quant à l'homogénéité, l'anisotropie et la complexité des fonds marins. Contrairement à ce type de descripteurs existant déjà dans la littérature, la caractéristique principale ici est qu'ils représentent chacun un indice de confiance quant à l'appartenance d'un pixel de l'image à un fond homogène, anisotrope ou complexe. De ce fait, leurs dynamiques sont connues *a priori* et bornées dans l'intervalle  $[0, 1]$ . Nous avons également pris le parti de rendre ces descripteurs complémentaires afin que leur somme soit toujours égale à un en chaque pixel de l'image. Selon nous, cette propriété permettrait soit à un opérateur humain d'interpréter rapidement le contenu de l'image sonar, soit à un ordinateur de prendre une décision quant aux types de fond marin qui y sont présents. Pour démontrer ce dernier point, nous avons proposé de mettre en place un protocole simple de classification non-supervisée du fond (*i.e.* basé uniquement sur l'utilisation de seuillages) suivant les classes homogène, anisotrope et complexe. Nous avons alors remarqué que la propriété intéressante de cette classification réside dans sa précision en termes de localisation de ces différentes classes. De plus, il est à noter qu'une frontière entre deux types de fond distinct est classée automatiquement comme une zone complexe. D'un point de vue ATR, une telle décision est intéressante puisque les

frontières sont des zones où la détection de cibles est généralement considérée comme difficile dans la littérature.

Que ce soient les algorithmes d'extraction des descripteurs ou ceux de classification, ils ont tous été évalués sur différentes bases de données issues de capteurs sonar divers, tels que des sonars latéraux KLEIN 5000 de la DGA, Marine Sonic Technology Ltd du DSTO australien, ou encore sur des systèmes SAS du CMRE. Les résultats montrent alors des performances constantes entre les différentes bases de données bien qu'aucune modification des algorithmes, et surtout de leurs paramètres, m'a été faite. Ceci montre que le choix d'utiliser des opérateurs isotropes, sans paramétrage excessif, permet de s'adapter facilement aux différentes images sonar.

Puisqu'une opération de détection de mines sous-marines ne fait pas face aux mêmes difficultés sur les fonds homogènes, anisotropes ou complexes, nous avons proposé d'utiliser leur localisation précise afin que le processus ATR s'y adapte automatiquement. Dans cette optique, nous nous sommes intéressé à analyser et à caractériser de manière spécifique les zones classées comme homogènes et anisotropes pour pouvoir en extraire des informations pertinentes à des fins de détection. L'idée étant ensuite de proposer une nouvelle architecture ATR qui utilise ces informations pour améliorer les performances de la détection de cibles.

En premier lieu, l'analyse des zones classées comme homogènes nous a permis de constater que nos résultats de classification montraient une confusion entre les fonds véritablement homogènes, constitués de sable ou de vase et que nous considérons comme bénins dans ces travaux, et les zones d'ombre dans lesquelles une mine sous-marine peut potentiellement se cacher. Par conséquent, il a été nécessaire de mettre en place un protocole de détection de ces zones d'ombre. Nous l'avons basé sur une segmentation non-paramétrique d'histogrammes dans le cas où de grandes zones d'ombre sont présentes dans l'image sonar, et sur l'utilisation des quantiles dans le cas contraire. La détection de ces zones d'ombre peut alors être très utile pour l'ATR mais également pour les opérationnels de la chasse aux mines puisqu'elle permet de pointer une zone du fond qui n'a pas été insonifiée et qu'il faudra nécessairement ré-inspecter si l'on juge qu'une mine peut s'y cacher.

Nous avons ensuite proposé une caractérisation des zones classées comme anisotropes. Pour ce faire, nous avons utilisé les informations multi-échelles disponibles (*i.e.* issues du signal monogène et des dimensions intrinsèques continues) pour les caractériser en termes de longueur d'onde, d'orientation mais aussi de crête et de vallée. Ces caractéristiques, propres au champ de rides de sable observé, nous ont permis de voir qu'une cible posée à l'intérieur de ce champ peut être considérée comme une anomalie causée aussi bien par son écho que par son ombre. Nous avons alors réalisé un détecteur spécifique aux zones anisotropes qui s'adapte à leurs caractéristiques environnementales locales. Grâce aux données fournies par le CMRE et à leurs vérités terrains, nous avons pu remarquer que les performances de cet algorithme surpassent celles de l'algorithme le plus récemment publié en termes de taux de fausses alarmes, bien que

la probabilité de détection soit équivalente. Ce résultat permet de constater que la prise en compte de l'environnement marin dans le processus de reconnaissance automatique de cibles est une piste intéressante pour les futurs systèmes de chasse aux mines.

### **Perspectives :**

Ce travail de thèse ouvre sur différentes perspectives et travaux futurs comme par exemple :

- Une étude spécifique de la détection de mines sous-marines dans les zones complexes. En effet, ces zones n'ont pas été traitées au cours de ces travaux, et il est nécessaire de mettre en place un protocole de détection spécialisé dans ces régions. Cependant, la variabilité des structures sous-marines qui peuvent y apparaître laisse à penser que cette tâche sera très difficile et que des méthodes évoluées seront alors nécessaires. De plus, il serait adéquate d'avoir à disposition une base de données d'images sonar représentant des conditions environnementales complexes dans lesquelles des cibles auront été posées. Une possibilité serait également de labelliser ces zones complexes en tant que fonds "non-chassables" après une étude plus approfondie de ces caractéristiques.
- Une analyse plus approfondie des informations portées par le signal monogène et les dimensions intrinsèques dans l'espace d'échelle. En effet, ces deux outils nous fournissent un ensemble d'informations très important. Dans ces travaux, nous avons proposé une première approche afin de les analyser et de les comprendre le mieux possible, mais cette étude peut certainement être améliorée. Par exemple, il est probable que l'analyse de toutes ces informations puisse être menée par directement par des méthodes d'apprentissage qui rendraient alors la caractérisation environnementale des images plus automatique que ce qu'elle est actuellement, et peut être également la détection des cibles s'y cachant.
- La paramétrisation de l'analyse multi-échelle utilisée. En effet, nous nous sommes intéressés dans ces travaux à analyser des dimensions de structures sous-marines dont les tailles vont de 0.25m à 2.5m. Sur les données que nous avons à disposition, ce paramétrage semblait adéquate en fournissant des descripteurs d'homogénéité, d'anisotropie, de complexité ou des résultats de classification très précis. Cependant, sur d'autres images sonar, collectées sur d'autres zones géographiques avec des fonds marins que nous n'avons eu l'opportunité d'analyser ici, il serait intéressant d'observer l'influence de ce paramétrage. Par exemple, on peut imaginer aller à des niveaux de détail plus important afin obtenir des informations supplémentaires sur des structures sous-marines importantes.





## Annexe A

# Les transformées de Hilbert 2D classiques

Cette annexe propose d'introduire les différentes méthodes d'extension de la transformée de Hilbert, et de leur signal analytique associé, pour des espaces à deux dimensions. On se focalise ici sur les méthodes proposées avant l'introduction par M. Felsberg du concept de signal monogène [106].

Dans ce qui suit,  $f(\mathbf{x}) : \mathbb{R}^2 \rightarrow \mathbb{R}$  désignera un signal réel 2D de variable  $\mathbf{x} = (x_1, x_2)$ , et  $F(\mathbf{u})$  sera sa transformée de Fourier, où  $\mathbf{u} = (u_1, u_2)$ .

### A.1 La transformée de Hilbert partielle

En 1996, l'approche de S.L. Hahn [104] est de considérer l'extension 2D de la transformée de Hilbert comme l'application d'une transformée de Hilbert 1D dans une direction spécifiée par un vecteur unitaire  $\mathbf{n}_k = [\cos(\theta_k) \sin(\theta_k)]^T$ . Cette transformée de Hilbert, qu'il appelle *transformée de Hilbert partielle*, notée  $F_{\mathcal{H}}^{\mathbf{n}_k}$ , est donnée par l'expression suivante dans le domaine fréquentiel :

$$F_{\mathcal{H}}^{\mathbf{n}_k}(\mathbf{u}) = -F(\mathbf{u}) \, i \, \text{sign}(\mathbf{u}^T \cdot \mathbf{n}_k), \quad (\text{A.1})$$

où  $i$  est le nombre complexe imaginaire,  $\text{sign}(\cdot) = \frac{\cdot}{|\cdot|}$  est la fonction signe (avec  $|\cdot|$  qui représente la valeur absolue),  $T$  désigne la transposée et l'opérateur " $\cdot$ " correspond au produit scalaire. La transformée de Hilbert partielle est alors une transformée de Hilbert 1D projetée dans une direction  $\theta_k$  fixée par le vecteur unitaire  $\mathbf{n}_k$ .

Le signal analytique 2D  $s_f^{\mathbf{n}_k}$  issu de cette transformée est appelé *signal analytique partiel*. Il est exprimé sous la forme complexe suivante :

$$s_f^{\mathbf{n}_k}(\mathbf{x}) = f(\mathbf{x}) + i f_{\mathcal{H}}^{\mathbf{n}_k}(\mathbf{x}), \quad (\text{A.2})$$

où  $f_{\mathcal{H}}^{\mathbf{n}_k}$  est la transformée de Fourier inverse de  $F_{\mathcal{H}}^{\mathbf{n}_k}$ .

Ce signal analytique partiel conserve la particularité de séparer le spectre en deux parties (analogues aux fréquences positives et négatives du cas 1D) comme le montre l'expression de son spectre :

$$S_f^{\mathbf{n}_k}(\mathbf{u}) = \begin{cases} 2F(\mathbf{u}) & \text{si } \mathbf{u}^T \cdot \mathbf{n}_k > 0 \\ F(\mathbf{u}) & \text{si } \mathbf{u}^T \cdot \mathbf{n}_k = 0 \\ 0 & \text{sinon} \end{cases}$$

Cependant, une telle approche nécessite l'hypothèse que le signal contenu dans l'image possède un sens de variation privilégié et fixe que l'on est capable d'estimer. Cette approche ne peut donc pas être retenue pour généraliser le signal analytique.

Pour pallier à ces défauts, la transformée de Hilbert totale est introduite en proposant de réaliser une transformée de Hilbert partielle suivant les deux coordonnées  $x_1$  et  $x_2$ .

## A.2 La transformée de Hilbert totale

H. Stark [102] propose en 1971 l'expression de la transformée de Hilbert totale  $f_{\mathcal{H}}^{\text{tot}}$ , qui est donnée par la relation suivante dans le domaine spatial :

$$f_{\mathcal{H}}^{\text{tot}}(\mathbf{x}) = \text{v.p} \int_{\mathbb{R}^2} \frac{f(x'_1, x'_2)}{\pi^2(x_1 - x'_1)(x_2 - x'_2)} dx'_1 dx'_2, \quad (\text{A.3})$$

où v.p signifie la valeur principale de Cauchy.

Elle est ainsi définie comme la convolution du signal par la distribution tempérée  $h_2(x_1, x_2) = \frac{1}{\pi^2 x_1 x_2}$ . On peut également l'exprimer dans le domaine fréquentiel par l'intermédiaire de l'expression :

$$F_{\mathcal{H}}^{\text{tot}}(\mathbf{u}) = - \prod_{k=1}^2 \text{sign}(u_k) F(\mathbf{u}) \quad (\text{A.4})$$

Ce qui est l'expression de la transformée de Hilbert suivant les deux axes orthogonaux de l'image.

Ainsi, le signal analytique total  $s_f^{\text{tot}}(\mathbf{x})$  de  $f(\mathbf{x})$  est défini par :

$$s_f^{\text{tot}}(\mathbf{x}) = f(\mathbf{x}) + i f_{\mathcal{H}}^{\text{tot}}(\mathbf{x}), \quad (\text{A.5})$$

et par la relation suivante dans le domaine fréquentiel :

$$S_f^{\text{tot}}(\mathbf{u}) = F(\mathbf{u}) \left( 1 - i \prod_{k=1}^2 \text{sign}(u_k) \right) \quad (\text{A.6})$$

Plusieurs désavantages de cette approche peuvent être notés. Tout d'abord dans la définition de la transformée de Hilbert totale (équation (A.3)), on remarque que cette approche n'est pas isotrope. En effet, le signal est convolué par le noyau  $\frac{1}{\pi^2 x_1 x_2}$ . De ce fait, les informations qui seront extraites du signal analytique total seront affectées, de manière systématique, par des erreurs qui dépendront de l'angle entre les axes orthogonaux de l'image et l'orientation du signal 2D. En outre, comme nous pouvons le voir dans l'équation (A.4), cette approche perd la notion d'orthogonalité entre le signal et sa transformée de Hilbert. En effet, nous n'avons plus de déphasage entre  $F$  et  $F_{\mathcal{H}}^{tot}$  qui était symbolisé par le nombre complexe imaginaire  $i$  dans l'équation A.1. Ainsi, les informations de phases ou de fréquences locales ne pourront pas être correctement estimées. Par conséquent, cette extension bidimensionnelle du signal analytique n'est pas suffisante.

### A.3 La transformée de Hilbert 2D par approche "Single Orthant"

Dans cette approche, l'idée proposée par S.L. Hahn [103] consiste à ne conserver que le spectre des orthants correspondant aux fréquences positives, tout en évitant les redondances. De manière générale, un orthant est un sous-espace d'un espace nD. Un espace nD sera composé de  $2^n$  orthants. En particulier, dans le cas 1D, un orthant correspondra à un demi-axe, alors qu'en 2D se sera un quadrant. La Figure A.1 illustre les quatre orthants de l'espace fréquentiel dans le cas 2D.

La méthode consiste alors à conserver uniquement l'information d'un seul des orthants (par exemple, on considère ici le numéro 1 de la Figure A.1 correspondant aux fréquences positives). Pour le signal  $f$ , cette opération est définie par l'équation suivante :

$$F_{SO_1}(u_1, u_2) = F(u_1, u_2) (1 + \text{sign}(u_1)) (1 + \text{sign}(u_2)). \quad (\text{A.7})$$

Notons que  $F_{SO_1}(u_1, u_2)$  est égale à  $4F(u_1, u_2)$  dans le premier orthant et vaut 0 ailleurs. Dans [103], S.L. Hahn démontre que la transformée de Fourier inverse de l'équation (A.7) mène à un

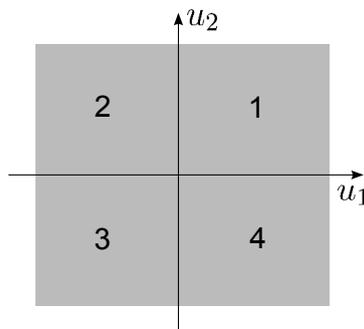


FIGURE A.1: Les quatre orthants du plan fréquentiel.

signal complexe  $s_f^{SO}$  qui fait intervenir les transformées de Hilbert totale et partielle définies précédemment :

$$s_f^{SO}(\mathbf{x}) = f(\mathbf{x}) - f_{\mathcal{H}}^{tot}(\mathbf{x}) + j [f_{\mathcal{H}}^{\mathbf{n}1}(\mathbf{x}) + f_{\mathcal{H}}^{\mathbf{n}2}(\mathbf{x})], \quad (\text{A.8})$$

où  $f_{\mathcal{H}}^{\mathbf{n}1}(\mathbf{x})$  et  $f_{\mathcal{H}}^{\mathbf{n}2}(\mathbf{x})$  définissent les transformées de Hilbert partielles de  $f$  selon l'axe horizontal et vertical de l'image.

Bien que ce signal analytique "single orthant" permette également de définir des informations d'amplitude et de phase locales, il ne constitue pas une extension 2D convenable du signal analytique. En effet, une reconstruction complète du signal d'origine ne pourra pas être effectuée en étudiant un seul et même orthant du plan fréquentiel (contrairement aux approches proposées par le signal analytique partiel et total, où il suffit de prendre la partie réelle dans les équations (A.2) et (A.6) pour reconstruire le signal initial).

## A.4 Le signal analytique hypercomplexe

Les 3 approches précédentes restreignent le signal analytique 2D à un nombre complexe afin d'en extraire facilement les informations d'amplitude et de phase locale de manière analogue au cas 1D. Or en 2D, il semblerait qu'une telle représentation manque de degré de liberté. Afin de pallier à ces limites et d'offrir une plus grande liberté au signal analytique 2D, T. Bülow [168] propose une approche hypercomplexe. Cette approche consiste en une combinaison entre l'approche "single orthant" précédente et la transformée de Fourier hypercomplexe [105]. Cette transformée permet à un signal nD d'être reconstruit de manière complète à partir d'un seul orthant. Cette dernière est définie par l'expression suivante :

$$F^h(\mathbf{u}) = \int_{\mathbb{R}^n} f(\mathbf{x}) \prod_{k=1}^n \exp(-\mathbf{e}_k 2\pi u_k x_k) d^n \mathbf{x}, \quad (\text{A.9})$$

où les symboles  $\mathbf{e}_k$ ,  $k \in \llbracket 1, n \rrbracket$ , sont les unités imaginaires qui définissent une  $\mathbb{R}$ -algèbre de dimension  $2^n$  [169]. Cette transformée est inversible et son expression est donnée par :

$$f(\mathbf{x}) = \int_{\mathbb{R}_n} F^h(\mathbf{u}) \prod_{k=1}^n \exp(\mathbf{e}_{n+1-k} 2\pi u_{n+1-k} x_{n+1-k}) d^n \mathbf{u}. \quad (\text{A.10})$$

Ainsi, le signal hypercomplexe de  $f$ , restreint au premier orthant, est définie dans le domaine hypercomplexe de Fourier par :

$$F_{SO}^h(\mathbf{u}) = F^h(\mathbf{u}) \prod_{k=1}^n (1 + \text{sign}(u_k)). \quad (\text{A.11})$$

Dans le cas 2D, le signal hypercomplexe peut donc être obtenu par la transformée de Fourier hypercomplexe inverse de l'équation A.11 et est exprimée à l'aide des transformées de Hilbert

partielles (selon l'axe horizontal et vertical de l'image) et totale de la manière suivante :

$$f_{\text{SO}}^h(\mathbf{x}) = f(\mathbf{x}) + i f_{\mathcal{H}}^{\mathbf{n}1}(\mathbf{x}) + j f_{\mathcal{H}}^{\mathbf{n}2}(\mathbf{x}) + k f_{\mathcal{H}}^{\text{tot}}(\mathbf{x}), \quad (\text{A.12})$$

qui est l'expression d'un quaternion de Hamilton, où  $i^2 = j^2 = k^2 = ijk = -1$ . Le signal initial est cette fois reconstituable via la partie réelle du quaternion.

Cependant, bien que cette représentation fournisse également des informations d'amplitude et de phase instantanées, elle ne permet pas une extension stricte du concept de signal analytique en 2D.

En effet, tout comme les 3 approches précédentes, cette dernière n'est pas isotrope comme le souligne M. Felsberg dans [9]. C'est-à-dire que la propriété fondamentale d'invariance-équivariance du signal analytique 1D n'est pas respectée.

L'auteur propose une solution à ce problème en étudiant les fonctions dites *monogènes* qui sont l'extension des fonctions analytiques dans des espaces de dimension supérieure. Il en résulte une généralisation isotrope du signal analytique basée sur la transformée de Riesz. Ce dernier, appelé *signal monogène*, tire partie de l'algèbre géométrique pour étendre l'analyse complexe à des signaux 2D.



## Annexe B

# Algèbre géométrique

Cette annexe propose une introduction à algèbre géométrique, en se focalisant sur les outils nécessaires à la compréhension des travaux de M. Felsberg pour l'extension du signal analytique à des données multi-dimensionnelles.

### B.1 Origines de l'algèbre géométrique

L'algèbre géométrique est une structure mathématique qui repose sur l'algèbre de Clifford. Ce dernier vise à créer un langage géométrique de haut niveau unifiant le produit scalaire, le produit vectoriel, les rotations et autres dilations sur tout objet de dimension quelconque comme des "demis ou quarts plans orientés", des "segments de droite orientés", ou tout simplement des vecteurs. Les capacités de l'algèbre géométrique à retranscrire de manière simple et concise des opérations géométriques complexes et sophistiquées dans un espace de grande dimension, en font un outil à la popularité grandissante. La physique ou les sciences de l'ingénieur l'utilisent de plus en plus systématiquement.

Nous allons tenter de décrire une telle algèbre, en nous focalisant sur les points clefs qui ont permis à M. Felsberg de construire l'extension multidimensionnelle du signal analytique. Plus de précisions sont apportées dans l'ouvrage de C. Perwass [170].

#### B.1.1 Premières définitions

Afin de définir cette algèbre, on part d'un espace vectoriel  $E$  sur  $\mathbb{R}$  de dimension  $n$ , muni de la base orthonormée canonique  $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ . La construction d'une algèbre dépend de sa loi de composition interne, c'est-à-dire du produit entre deux éléments  $u$  et  $v$  de  $E$ . Dans le cadre de l'algèbre géométrique, le produit  $uv$  est appelé *produit géométrique*.

**Definition B.1.** Soient  $u$  et  $v$  deux éléments de  $E$ , le produit géométrique entre  $u$  et  $v$  sera défini par la relation :

$$uv = u.v + u \wedge v, \quad (\text{B.1})$$

où les opérateurs "." et " $\wedge$ " sont respectivement le produit scalaire (ou produit intérieur) et produit vectoriel (ou produit extérieur).

Puisque le produit vectoriel est une forme anti-symétrique, nous obtenons dans le cas particulier où les éléments  $u$  et  $v$  sont orthogonaux :

$$u.v = 0 \Rightarrow uv = -vu \quad (\text{B.2})$$

$$\Rightarrow (uv)^2 = -(uv)(vu) = -u(vv)u = -Q(u)Q(v), \quad (\text{B.3})$$

où  $Q$  est une forme quadratique, c'est-à-dire une application  $Q : E \rightarrow \mathbb{R}$ .

Dans ce cas particulier, on remarque que l'objet construit par le produit géométrique  $uv$  possède un carré scalaire de signe négatif. Ce dernier ne peut donc ni être un scalaire, ni un vecteur. On appellera cet objet un *multivecteur*. Ce concept peut être vu comme une entité globale qui considère les scalaires et les vecteurs comme des cas particuliers.

**Definition B.2.** Soient  $m$  vecteurs  $\mathbf{u}_1, \dots, \mathbf{u}_m \in E$ , on appelle multivecteur (ou  $m$ -vecteur) l'élément :

$$M = \mathbf{u}_1 \wedge \mathbf{u}_2 \dots \wedge \mathbf{u}_m. \quad (\text{B.4})$$

On peut alors constater par exemple qu'un 0-vecteur sera en fait un scalaire, qu'un 1-vecteur symbolisera un vecteur (donc un axe et une direction dans l'espace), et qu'un 2-vecteur (ou plus couramment *bivecteur*) représentera quant-à-lui un plan (*orienté*) et ainsi de suite.

Pour définir une base de  $E$  dans l'algèbre géométrique, il est nécessaire de tenir compte de ces nouveaux éléments spécifiques à cette algèbre. Pour cela, nous partons de la base orthonormée canonique  $(\mathbf{e}_1, \dots, \mathbf{e}_n)$  de  $E$ . On peut alors définir les bivecteurs  $\mathbf{e}_{ij}$  de base de l'algèbre de la manière suivante :

$$\mathbf{e}_{ij} = \mathbf{e}_i \wedge \mathbf{e}_j, \quad i \neq j, \quad (\text{B.5})$$

ainsi que les *trivecteurs* de base par la relation :

$$\mathbf{e}_{ijk} = \mathbf{e}_i \wedge \mathbf{e}_j \wedge \mathbf{e}_k, \quad i \neq j \neq k. \quad (\text{B.6})$$

où  $i, j$  et  $k$  appartiennent à l'intervalle  $[[1, n]]$ .

L'espace vectoriel  $E$  étant de dimension  $n$ , on peut alors aller jusqu'à former l'unique  $n$ -vecteur :  $I_n = \mathbf{e}_1 \wedge \dots \wedge \mathbf{e}_n$ , que l'on nomme *pseudoscalaire*. L'ensemble de ces multivecteurs particuliers forme la base de l'algèbre géométrique, notée  $\mathbb{G}_n$ .

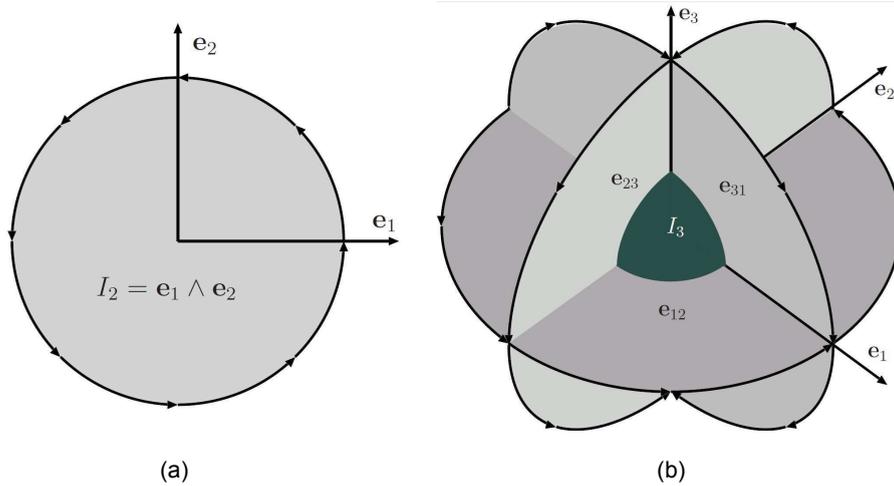


FIGURE B.1: (a) Espace euclidien 2D composé des vecteurs de base  $\mathbf{e}_1$  et  $\mathbf{e}_2$  ainsi que du pseudoscalaire  $\mathbf{I}_2$ . (b) Espace Euclidien 3D composé des vecteurs de base  $\mathbf{e}_1$ ,  $\mathbf{e}_2$ ,  $\mathbf{e}_3$ , de la base des bivecteurs  $\mathbf{e}_{12}$ ,  $\mathbf{e}_{13}$ ,  $\mathbf{e}_{23}$ , et du pseudoscalaire  $\mathbf{I}_3 = \mathbf{e}_{123}$  (extrait de [9]).

### B.1.2 Cas de l'espace Euclidien 2D

Dans le cas 2D (*i.e.*  $n = 2$ ), l'espace Euclidien  $E$  est défini par une base orthonormée canonique formée des vecteurs  $\mathbf{e}_1$  et  $\mathbf{e}_2$ , comme le montre la Figure B.1.(a). A partir de ces vecteurs, un seul bivecteur  $\mathbf{e}_{12}$  peut être formé, et sera par la même occasion le pseudoscalaire  $\mathbf{I}_2 = \mathbf{e}_{12}$ . Un multivecteur  $M$  quelconque de  $E$  sera alors exprimé par la relation :

$$M = a + b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_{12}, \quad (\text{B.7})$$

où  $a, b, c, d \in \mathbb{R}$ . Cet exemple 2D nous permet de se rendre compte que le concept de bivecteur représente en fait la notion de rotation entre deux vecteurs, comme ici entre  $\mathbf{e}_1$  et  $\mathbf{e}_2$ . En effet, l'action du bivecteur sur l'élément de base  $\mathbf{e}_1$  est donnée par la relation  $\mathbf{e}_1\mathbf{e}_{12} = \mathbf{e}_2$ . Cette action définit donc un angle de  $\pi/2$ , qui n'est pas sans rappeler l'unité imaginaire  $j = e^{j\pi/2}$  de l'algèbre complexe.

### B.1.3 Cas de l'espace Euclidien 3D

Dans le cas 3D (*i.e.*  $n = 3$ ), l'espace Euclidien  $E$  est défini par une base orthonormée canonique formée des vecteurs  $\mathbf{e}_1$ ,  $\mathbf{e}_2$  et  $\mathbf{e}_3$ , comme le montre la Figure B.1.b. A partir de ces derniers, 3 bivecteurs  $\mathbf{e}_{12}$ ,  $\mathbf{e}_{31}$  et  $\mathbf{e}_{23}$  sont formés, ainsi qu'un seul trivecteur  $\mathbf{e}_{123} = \mathbf{I}_3$  qui est le pseudoscalaire correspondant à cet espace. Un multivecteur  $M$  quelconque de  $E$  sera alors exprimé par la relation :

$$M = a + b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_3 + e\mathbf{e}_{12} + f\mathbf{e}_{31} + g\mathbf{e}_{23} + g\mathbf{I}_3, \quad (\text{B.8})$$

	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{12}$	$\mathbf{e}_{23}$	$\mathbf{e}_{13}$	$\mathbf{I}_3$
1	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{12}$	$\mathbf{e}_{23}$	$\mathbf{e}_{13}$	$\mathbf{I}_3$
$\mathbf{e}_1$	$\mathbf{e}_1$	1	$\mathbf{e}_{12}$	$-\mathbf{e}_{31}$	$\mathbf{I}_3$	$-\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_{23}$
$\mathbf{e}_2$	$\mathbf{e}_2$	$-\mathbf{e}_{12}$	1	$\mathbf{e}_{23}$	$\mathbf{e}_3$	$\mathbf{I}_3$	$-\mathbf{e}_1$	$\mathbf{e}_{31}$
$\mathbf{e}_3$	$\mathbf{e}_3$	$\mathbf{e}_{31}$	$-\mathbf{e}_{23}$	1	$-\mathbf{e}_2$	$-\mathbf{e}_1$	$\mathbf{I}_3$	$\mathbf{e}_{12}$
$\mathbf{e}_{23}$	$\mathbf{e}_{23}$	$\mathbf{I}_3$	$-\mathbf{e}_3$	$\mathbf{e}_2$	-1	$-\mathbf{e}_{12}$	$\mathbf{e}_{31}$	$-\mathbf{e}_1$
$\mathbf{e}_{31}$	$\mathbf{e}_{31}$	$\mathbf{e}_3$	$\mathbf{I}_3$	$-\mathbf{e}_1$	$\mathbf{e}_{12}$	-1	$-\mathbf{e}_{23}$	$-\mathbf{e}_2$
$\mathbf{e}_{12}$	$\mathbf{e}_{12}$	$-\mathbf{e}_2$	$\mathbf{e}_1$	$\mathbf{I}_3$	$-\mathbf{e}_{31}$	$\mathbf{e}_{23}$	-1	$-\mathbf{e}_3$
$\mathbf{I}_3$	$\mathbf{I}_3$	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	$-\mathbf{e}_1$	$-\mathbf{e}_2$	$-\mathbf{e}_3$	-1

TABLE B.1: Table des produits géométriques entre les éléments de base  $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_{12}, \mathbf{e}_{23}, \mathbf{e}_{13}, \mathbf{e}_{123}\}$ . La première colonne et la première colonne indiquent les membres de gauche et de droite du produit géométrique, respectivement.

où  $a, b, c, d, e, f, g, h \in \mathbb{R}$ .

La Figure précédente permet également de se rendre compte de la notion de rotation amenée par le concept de bivecteur. En 3D, ces bivecteurs définissent 3 plans orientés qui représentent les différentes rotations possibles entre chaque vecteurs de base de l'algèbre géométrique. Le pseudoscalaire représente quant-à-lui le volume de l'espace dans lequel il est défini (dans le cas 2D c'est un plan, en 3D c'est un volume 3D).

Dans l'annexe C (au paragraphe C.2), nous pourrons voir que l'extension monogène du signal analytique nécessite de se placer dans l'espace Euclidien 3D. C'est pour cette raison que dans la suite de ce paragraphe, nous allons poursuivre l'étude de l'algèbre géométrique dans cet espace. La Table B.1 permet de récapituler les produits géométriques entre les différents multivecteurs de base de cet espace Euclidien 3D.

#### B.1.4 Les spineurs

Nous avons vu précédemment que les bivecteurs permettaient la rotation des vecteurs de base par création de plans orientés. Pour étudier la notion de rotation en 3D, il est nécessaire d'introduire également un type très particulier de multivecteurs, nommés *spineurs*. Ils ont la particularité de n'avoir qu'une partie scalaire et bivectorielle, c'est-à-dire qu'ils ont la forme suivante :

$$s = a + b\mathbf{e}_{12} + c\mathbf{e}_{23} + d\mathbf{e}_{13}. \quad (\text{B.9})$$

L'avantage de tels multivecteurs réside dans la sous-algèbre qu'ils représentent. En effet, cette dernière est isomorphe aux quaternions comme le montre P. Lounesto dans [171]. Pour rappel, un quaternion  $q$  est un élément qui est défini par  $q = a + ib + jc + kd$ , où  $i$ ,  $j$  et  $k$  sont les unités quaternioniques telles que  $i^2 = j^2 = k^2 = ijk = -1$ . En algèbre géométrique 2D, nous avons vu qu'un lien existait entre l'unité imaginaire et la notion de bivecteur. En 3D, les bivecteurs de base ont les mêmes propriétés que ces unités quaternioniques. C'est-à-dire que les spineurs possèdent les mêmes propriétés de groupe que les quaternions. En particulier, ils ont la possibilité de représenter de manière simple et compacte des rotations 3D.

Si l'on considère un vecteur  $u = x_1\mathbf{e}_1 + y_1\mathbf{e}_2 + z_1\mathbf{e}_3$ ,  $(x, y, z) \in \mathbb{R}^3$ , de  $E$  et que l'on réalise l'opération :

$$r = (a - b\mathbf{e}_{12} - c\mathbf{e}_{23} - d\mathbf{e}_{13})u(a + b\mathbf{e}_{12} + c\mathbf{e}_{23} + d\mathbf{e}_{13}) \quad (\text{B.10})$$

$$= x_2\mathbf{e}_1 + y_2\mathbf{e}_2 + z_2\mathbf{e}_3 \quad (\text{B.11})$$

le vecteur  $r$  est alors la rotation du vecteur  $u$  autour de l'axe dirigé par le vecteur  $b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_3$ . L'angle  $\alpha$  de la rotation est donné par les relations suivantes :

$$a = \cos \frac{\alpha}{2}, \quad b = c = d = \frac{1}{\sqrt{b^2 + c^2 + d^2}} \sin(\alpha/2). \quad (\text{B.12})$$

## B.2 Algèbre géométrique et traitement du signal

Nous avons défini un certain nombre de concepts nécessaires à la compréhension de l'extension multidimensionnelle du signal analytique proposée par M. Felsberg [9]. La première contribution de l'auteur a été de proposer une formalisation de certains outils classiques du traitement du signal dans le cadre des algèbres géométriques.

Tout d'abord dans de telles algèbres, il est possible de distinguer algébriquement les signaux et les filtres. En effet, alors que les signaux sont à valeurs vectorielles :

$$s(\mathbf{x}) = s_1(\mathbf{x})\mathbf{e}_1 + s_2(\mathbf{x})\mathbf{e}_2 + s_3(\mathbf{x})\mathbf{e}_3, \quad (\text{B.13})$$

où les signaux réels classiques sont portés par l'axe  $\mathbf{e}_3$  (i.e.  $s_1 = s_2 = 0$ ), et les signaux 'plus sophistiqués' peuvent être construits en utilisant les axes  $\mathbf{e}_1$  et  $\mathbf{e}_2$  pour former des parties 'imaginaires'.

Les filtres sont quant-à-eux formulés sous forme de spineurs. L'expression générale des filtres dans l'algèbre géométrique sera donc de la forme :

$$h(\mathbf{x}) = h_1(\mathbf{x}) + h_2(\mathbf{x})\mathbf{e}_{12} + h_3(\mathbf{x})\mathbf{e}_{23} + h_4(\mathbf{x})\mathbf{e}_{31}. \quad (\text{B.14})$$

En particulier, un filtre réel sera alors un simple salaire qui modifiera le "vecteur signal" sans rotation. Si l'expression du filtre nécessite des composantes imaginaire, la partie bivectorielle sera alors non-nulle, et impactera en terme de rotation le vecteur signal.

Le filtrage d'un signal  $s$  par un filtre  $h$  est alors défini par la convolution suivante :

$$(h * s)(\mathbf{x}) = \int_{\mathbb{R}^2} h(\mathbf{x}')s(\mathbf{x} - \mathbf{x}')d\mathbf{x}', \quad (\text{B.15})$$

où le produit dans l'intégrale est le produit géométrique.

L'analyse fréquentielle étant devenue incontournable dans le domaine du traitement du signal, il est nécessaire de formuler la transformée de Fourier dans cette nouvelle algèbre. Cette dernière est ainsi reformulée de la manière suivante :

$$\hat{s}(\mathbf{u}) = \int_{\mathbb{R}^2} s(\mathbf{x}) \exp(-\mathbf{I}_3 2\pi \mathbf{u}^T \mathbf{x}) d\mathbf{x}, \quad (\text{B.16})$$

où  $\mathbf{u} = [u_1 \ u_2]^T$ . Là encore le produit dans l'intégrale correspond au produit géométrique, et l'on peut remarquer que l'imaginaire complexe de la définition classique de la transformée de Fourier est remplacé par le pseudoscalaire  $\mathbf{I}_3$ . Cette transformée de Fourier peut s'interpréter comme la projection du vecteur signal sur une base d'ondes planes orientées.

Cette transformée de Fourier peut être divisée en une partie paire et impaire :

$$\hat{s}^e(\mathbf{u}) = \frac{\hat{s}(\mathbf{u}) + \hat{s}(-\mathbf{u})}{2} = \int_{\mathbb{R}^2} s(\mathbf{x}) \cos(-\mathbf{I}_3 2\pi \mathbf{u}^T \mathbf{x}) d\mathbf{x}, \quad (\text{B.17})$$

$$\hat{s}^o(\mathbf{u}) = \frac{(\hat{s}(\mathbf{u}) - \hat{s}(-\mathbf{u}))^*}{2} = - \int_{\mathbb{R}^2} s(\mathbf{x}) \sin(-\mathbf{I}_3 2\pi \mathbf{u}^T \mathbf{x}) d\mathbf{x}, \quad (\text{B.18})$$

qui permet de voir cette transformée de Fourier sous la forme du multivecteur  $\hat{s}(\mathbf{u}) = \hat{s}^e(\mathbf{u}) + \mathbf{I}_3 \hat{s}^o(\mathbf{u})$ . On y remarque une fois de plus que l'imaginaire complexe, de la version classique, est ici remplacé par le pseudoscalaire.

Évidemment, les propriétés de la transformée de Fourier (TF) classique sont conservées, comme la linéarité ou encore le théorème du décalage :

$$s(\mathbf{x} - \mathbf{x}') \xleftrightarrow{TF} \hat{s}(\mathbf{u}) \exp(-\mathbf{I}_3 2\pi \mathbf{u}^T \mathbf{x}') \quad (\text{B.19})$$

Ainsi, la transformée de Fourier du signal  $s$  (à valeurs vectorielles) sera décomposée de la manière suivante dans les algèbres géométriques :

$$\hat{s}(\mathbf{u}) = F\{s(\mathbf{x})\}(\mathbf{u}) \quad (\text{B.20})$$

$$= F\{s_1(\mathbf{x})\}(\mathbf{u})\mathbf{e}_1 + F\{s_2(\mathbf{x})\}(\mathbf{u})\mathbf{e}_2 + F\{s_3(\mathbf{x})\}(\mathbf{u})\mathbf{e}_3 \quad (\text{B.21})$$

$$= \hat{s}_1^e(\mathbf{u})\mathbf{e}_1 + \hat{s}_1^o(\mathbf{u})\mathbf{e}_{23} + \hat{s}_2^e(\mathbf{u})\mathbf{e}_2 + \hat{s}_2^o(\mathbf{u})\mathbf{e}_{31} + \hat{s}_3^e(\mathbf{u})\mathbf{e}_3 + \hat{s}_3^o(\mathbf{u})\mathbf{e}_{12} \quad (\text{B.22})$$

Le spectre d'un filtre (i.e. d'un spineur) aura une expression répartie différemment sur les axes, avec en particulier les composantes scalaires et pseudoscalaires non-nulles :

$$\hat{h}(\mathbf{u}) = F\{h(\mathbf{x})\}(\mathbf{u}) \quad (\text{B.23})$$

$$\begin{aligned} &= F\{h_1(\mathbf{x})\}(\mathbf{u}) + F\{h_2(\mathbf{x})\}(\mathbf{u})\mathbf{e}_{12} + F\{h_3(\mathbf{x})\}(\mathbf{u})\mathbf{e}_{23} + F\{h_4(\mathbf{x})\}(\mathbf{u})\mathbf{e}_{31} \\ &= \hat{h}_1^e(\mathbf{u}) + \hat{h}_1^o(\mathbf{u})\mathbf{I}_3 + \hat{h}_2^e(\mathbf{u})\mathbf{e}_{12} + \hat{h}_2^o(\mathbf{u})\mathbf{e}_3 + \hat{h}_3^e(\mathbf{u})\mathbf{e}_{23} + \hat{h}_3^o(\mathbf{u})\mathbf{e}_1 \\ &\quad + \hat{h}_4^e(\mathbf{u})\mathbf{e}_{31} + \hat{h}_4^o(\mathbf{u})\mathbf{e}_2 \end{aligned} \quad (\text{B.24})$$

La notion de filtrage dans le domaine fréquentiel peut alors être définie. Elle correspond au produit géométrique entre le spectre du filtre et celui du signal :  $\hat{h}\hat{s}$ . Le produit géométrique n'étant pas commutatif, il est important de respecter l'ordre du produit. Dans le cas particulier (mais le plus fréquent) d'un signal réel (i.e. restreint à  $\mathbf{e}_3$ ) et d'un filtre réel (i.e. restreint à la partie scalaire du spineur), l'opération de filtrage s'exprime alors de la manière suivante :

$$(h_1) * (s_3\mathbf{e}_3) \quad \xleftrightarrow{TF} \quad (\hat{h}\hat{s})\mathbf{e}_3 \quad (\text{B.25})$$

A partir de ces premières définitions des outils de traitement du signal dans l'algèbre géométrique, il est possible définir un cadre mathématique rigoureux à l'extension bidimensionnelle du signal analytique.



## Annexe C

# Généralisation isotrope du signal analytique

Cette section reprend les étapes de formalisation du signal monogène [9], généralisation isotrope du signal analytique, dans l'algèbre géométrique précédemment introduite.

### C.1 Généralisation de la transformée de Hilbert

Le signal analytique portant son nom des fonctions dites analytiques, M. Felsberg en a profité pour baser sa généralisation à des dimensions supérieures dans le contexte des fonctions analytiques. Cette étude a abouti à la généralisation de la transformée de Hilbert 1D par la transformée de Riesz en tant qu'opérateur de déphasage. Le signal analytique 2D, dit signal monogène, résulte finalement d'un triplet constitué du signal et de ces deux transformées de Riesz. Dans la suite, nous allons reprendre les grandes lignes de cette généralisation dans l'algèbre géométrique.

#### C.1.1 Équations de Cauchy-Riemann

Le signal analytique 1D est défini mathématiquement comme la restriction à l'axe réel d'une *fonction analytique*. Une fonction analytique  $f$  est une fonction à valeur complexe d'une variable complexe  $f(x + jy) = f_R(x + jy) + if_I(x + jy)$  qui est dérivable. Cette dérivabilité est définie par les équations de Cauchy-Riemann suivantes :

$$\left\{ \frac{\partial f_R}{\partial x} = \frac{\partial f_I}{\partial y} \quad \text{et} \quad \frac{\partial f_R}{\partial y} = -\frac{\partial f_I}{\partial x} \right\} \iff \left\{ \nabla_{\mathbb{C}} f = 0 \quad \text{où} \quad \nabla_{\mathbb{C}} = \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} \right\}, \quad (\text{C.1})$$

dont le membre de droite est appelé "équation de Dirac" et fait intervenir un gradient complexe.

Un lien peut être établi entre ces fonctions analytiques et les fonctions harmoniques. En effet, soit  $p : \mathbb{R}^2 \rightarrow \mathbb{R}$  une fonction harmonique, elle vérifie l'équation dite de Laplace (2D) :

$$\Delta_2 p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} = 0. \quad (\text{C.2})$$

On peut alors vérifier aisément que le gradient complexe conjugué de  $p$  est une fonction analytique :

$$\left\{ \Delta_2 p = 0 \quad \text{et} \quad f = \frac{\partial p}{\partial x} - j \frac{\partial p}{\partial y} \right\} \Rightarrow \nabla_{\mathbb{C}} f = 0. \quad (\text{C.3})$$

On dit alors que  $f$  "dérive d'un potentiel harmonique".

Ainsi, étant donné un signal 1D réel  $s(x)$ , il existe une unique fonction analytique  $f$  qui contient dans la partie réelle le signal, *i.e.*  $\mathcal{R}e\{f(x + jy)\} = s(x)$ . Après développement, il apparaît alors que la partie imaginaire de cette fonction analytique n'est d'autre que la transformée de Hilbert de sa partie réelle pour un  $y$  fixé :

$$\mathcal{I}m\{f(x + jy)\} = \mathcal{H}\{\mathcal{R}e\{f(x + jy)\}\}(x). \quad (\text{C.4})$$

Ainsi, la transformée de Hilbert est partie intégrante de la théorie des fonctions analytiques puisqu'elle représente le lien entre la partie réelle et la partie imaginaire de ces dernières.

L'idée de M. Felsberg est donc d'étendre la notion de fonction analytique à des dimensions supérieures, et observer le lien qu'il existe entre les parties réelles et imaginaires pour observer s'il est possible de trouver une extension à la transformée de Hilbert.

Puisque nous venons de voir que les fonctions analytiques sont des fonctions complexes à variable complexe, elles sont incluses dans un espace 2D (une variable réelle et une variable imaginaire). Par conséquent, si l'on désire étendre le concept de signal analytique à des signaux 2D, on doit considérer l'ajout d'une dimension supplémentaire, soit l'analyse des fonctions analytiques dans un espace 3D. Il est donc nécessaire avant tout de redéfinir les équations de Cauchy-Riemann dans un tel espace. Cependant, ces équations sont figées dans l'algèbre des nombres complexes, et il est donc difficile de mener une telle extension. L'algèbre géométrique apparaît alors comme une solution crédible et efficace pour étendre ces équations et les résoudre en 3D.

### C.1.2 Équations de Cauchy-Riemann généralisées

Dans la section précédente, nous avons remarqué que les fonctions analytiques 2D dérivent d'un potentiel harmonique.

En 3D, un potentiel  $p$  est dit harmonique s'il vérifie l'équation de Laplace (3D) suivante :

$$\Delta_3 p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} = 0. \quad (\text{C.5})$$

La dérivation de ce potentiel harmonique est alors donnée, dans l'algèbre géométrique, par la fonction  $\mathbf{f}$  suivante :

$$\mathbf{f} = \nabla_3 p = \mathbf{e}_1 \frac{\partial p}{\partial x} + \mathbf{e}_2 \frac{\partial p}{\partial y} + \mathbf{e}_3 \frac{\partial p}{\partial z} \quad (\text{C.6})$$

$$= \mathbf{e}_1 f_1 + \mathbf{e}_2 f_2 + \mathbf{e}_3 f_3 \quad (\text{C.7})$$

qui doit respecter l'équation de Dirac 3D pour conserver sa nature *analytique* :

$$\nabla_3 \mathbf{f} = 0 \quad (\text{C.8})$$

Dans l'algèbre géométrique, cette équation est lue comme le produit géométrique entre l'opérateur  $\nabla_3$  et  $\mathbf{f}$ . Cette dernière équation peut alors être divisée en deux parties, l'une vectorielle et l'autre bivectorielle :

$$\nabla_3 \cdot \mathbf{f} = 0 \quad (\text{C.9})$$

$$\nabla_3 \wedge \mathbf{f} = 0 \quad (\text{C.10})$$

qui correspondent respectivement aux faits que  $\mathbf{f}$  est à divergence nulle, et irrotationnelle. Sachant qu'il peut être démontré que chacune des composantes d'une fonction harmonique est également une fonction harmonique, M. Felsberg propose les équations suivantes comme une généralisation des équations de Cauchy-Riemann en 3D :

$$\left\{ \begin{array}{l} \frac{\partial}{\partial x} f_1 + \frac{\partial}{\partial y} f_2 + \frac{\partial}{\partial z} f_3 = 0 \\ \frac{\partial}{\partial y} f_3 - \frac{\partial}{\partial z} f_2 = 0 \\ \frac{\partial}{\partial z} f_1 - \frac{\partial}{\partial x} f_3 = 0 \\ \frac{\partial}{\partial x} f_2 - \frac{\partial}{\partial y} f_1 = 0 \end{array} \right. \quad (\text{C.11})$$

Les fonctions qui vérifient un tel système sont appelées fonctions monogènes et représentent l'extension des fonctions analytiques pour des dimensions supérieures à 1.

Cependant, M. Felsberg montre que la résolution de ces équations peut se faire selon la valeur de  $z$ , et il y a alors deux cas : soit  $z = 0$  et soit  $z > 0$ .

## C.2 Résolution des équations de Cauchy-Riemann pour $z = 0$

M. Felsberg démontre que pour une valeur fixée de la dimension  $z = 0$ , les trois composantes  $f_1$ ,  $f_2$  et  $f_3$  de ces fonctions monogènes sont reliées par la transformée de Riesz. Cette transformée a été définie plus précisément au paragraphe 4.4.1.1, mais on rappelle ici

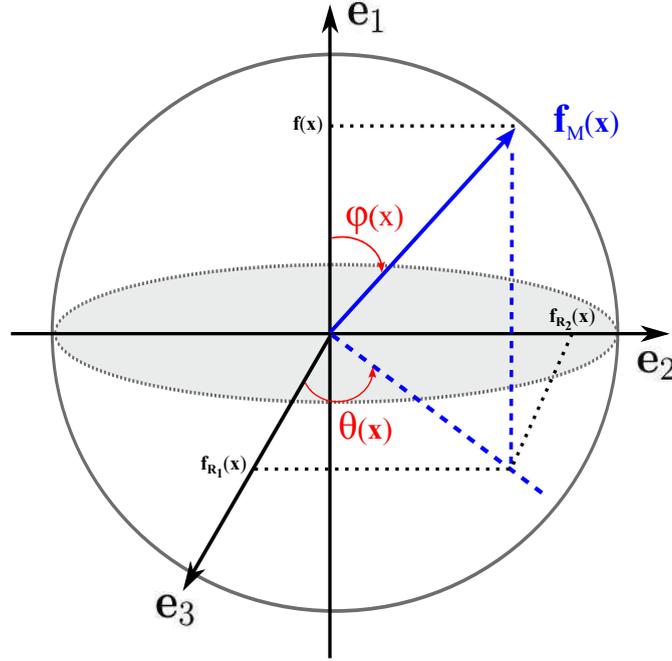


FIGURE C.1: Représentation 3D du vecteur signal monogène dans une base de l'algèbre géométrique.

qu'elle est connue du domaine mathématique comme l'extension multi-dimensionnelle de la transformée de Hilbert. Il propose de définir l'extension du signal analytique par le biais de cette transformée, et l'appelle signal monogène.

Puisque que le signal analytique correspond au nombre complexe formé par le signal (en partie réelle) et sa transformée de Hilbert (en partie imaginaire), M. Felsberg propose de définir le signal monogène comme le vecteur composé du signal et des deux parties de sa transformée de Riesz. Afin de généraliser l'écriture complexe, il utilise l'algèbre géométrique pour définir le signal monogène associé au signal  $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2) = f(x, y)\mathbf{e}_3$  par l'écriture suivante :

$$\mathbf{f}_M(\mathbf{x}) = \mathcal{R}_1\{f\}(\mathbf{x})\mathbf{e}_1 + \mathcal{R}_2\{f\}(\mathbf{x})\mathbf{e}_2 + f(\mathbf{x})\mathbf{e}_3, \quad (\text{C.12})$$

où  $\mathcal{R}_1\{f\}$  et  $\mathcal{R}_2\{f\}$  sont les deux parties de la transformée de Riesz du signal  $f$ . Pour  $k \in [1, 2]$ , leurs expressions sont données par :

$$\mathcal{R}_k\{f\}(\mathbf{x}) = (r_k * f)(\mathbf{x}) = \text{v.p} \iint_{\mathbb{R}^2} \frac{x'_k}{2\pi\|\mathbf{x}'\|^3} f(\mathbf{x} - \mathbf{x}') d\mathbf{x}' \quad (\text{C.13})$$

où  $r_k$  sont les noyaux de Riesz.

On constate que l'expression du signal monogène  $\mathbf{f}_M$  est analogue à celle du signal analytique qui se veut être une juxtaposition du signal et de sa version déphasée de  $\pi/2$  dans une sorte de "partie imaginaire". Ici la transformée de Riesz joue le rôle de la transformée de Hilbert.

Ce vecteur peut être représenté dans l'espace 3D comme le montre la Figure C.1. On y observe alors que 3 informations peuvent en être extraites, à savoir les coordonnées sphériques de ce dernier : une norme et deux angles. Ces informations sont données par les relations suivantes :

$$A(\mathbf{x}) = |\mathbf{f}_M(\mathbf{x})| = \sqrt{f_{R_1}(\mathbf{x})^2 + f_{R_2}(\mathbf{x})^2 + f(\mathbf{x})^2} \quad (\text{C.14})$$

$$\phi(\mathbf{x}) = \arg(f(\mathbf{x}) + j|\mathcal{R}\{f\}(\mathbf{x})|) = \text{atan} \left( \frac{\sqrt{f_{R_1}(\mathbf{x})^2 + f_{R_2}(\mathbf{x})^2}}{f(\mathbf{x})} \right) \quad (\text{C.15})$$

$$\theta(\mathbf{x}) = \arg(\mathcal{R}\{f\}(\mathbf{x})) = \text{atan} \left( \frac{f_{R_2}(\mathbf{x})}{f_{R_1}(\mathbf{x})} \right) \quad (\text{C.16})$$

Ces trois informations  $A$ ,  $\phi$  et  $\theta$  sont appelées respectivement *amplitude*, *phase* et *orientation locale* du signal 2D.

### C.3 Résolution des équations de Cauchy-Riemann pour $z > 0$

Nous avons abordé la résolution de des équations de Cauchy-Riemann 3D uniquement dans le cas  $z = 0$ . Or, pour être tout-à-fait général, il est nécessaire de les résoudre également dans le demi-espace supérieur ( $z > 0$ ).

Dans cet espace, selon M. Felsberg [9], la solution de ces équations est toujours une fonction monogène mais cette fois les noyaux de Riesz sont substitués par des noyaux de Poisson conjugués. Le signal monogène d'un signal  $I(x, y)$  dans ce demi-espace est alors formulé de la manière suivante :

$$\mathbf{f}_M(x, y, z) = (H_1 * f)(x, y, z)\mathbf{e}_1 + (H_2 * f)(x, y, z)\mathbf{e}_2 + (P * f)(x, y, z)\mathbf{e}_3, \quad (\text{C.17})$$

où  $H_k(x, y, z) = (r_k * P)(x, y, z)$ ,  $k \in \{1, 2\}$ , sont appelés noyaux de Poisson conjugués, et  $P$  est un noyau de Poisson 2D dont l'expression dans le domaine spatial est la suivante :

$$P(x, y, z) = \frac{z}{(x^2 + y^2 + z^2)^{3/2}}, \quad (\text{C.18})$$

Les noyaux de Poisson conjugués ne sont donc qu'une convolution entre un noyau de Poisson 2D et un noyau de Riesz.

Grâce aux lois de permutation autorisées par la convolution, l'équation (C.17) peut être réécrite de la manière suivante :

$$\mathbf{f}_M(x, y, z) = \mathcal{R}_1\{L\}(x, y, z)\mathbf{e}_1 + \mathcal{R}_1\{L\}(x, y, z)\mathbf{e}_2 + L(x, y, z)\mathbf{e}_3, \quad (\text{C.19})$$

où  $L(x, y, z) = P(x, y, z) * f(x, y)$ .

On peut alors considérer que dans le demi-espace supérieur  $z > 0$ , le signal monogène d'un signal  $f$  correspond au vecteur qui est formé par la version filtrée  $L$  (dans la composante  $\mathbf{e}_3$ ), et de ses deux transformées de Riesz (dans les composantes  $\mathbf{e}_1$  et  $\mathbf{e}_2$ ).

Il est possible de donner une interprétation physique du demi-espace engendré par des valeurs positives de la dimension  $z$ . En effet, puisque l'expression donnée à l'équation (C.18) définit des noyaux 2D, la dimension  $z$  agit seulement comme un paramètre. Ce paramètre permet alors de configurer le noyau de Poisson et une possible décomposition de l'image à différents niveaux de détails.

# Bibliographie

- [1] O. J. Daniell, *Sonar image interpretation for sub-sea operations*. PhD thesis, Heriot-Watt University, 2015. [x](#), [26](#)
- [2] E. Fakiris and D. Williams, “Seafloor acoustic anisotropy and complexity assessment towards prediction of ATR performance,” in *1st International Conference and Exhibition on Underwater Acoustics*, (Greece), 2013. [x](#), [40](#), [44](#), [48](#), [49](#)
- [3] D. Williams and E. Fakiris, “Exploiting environmental information for improved underwater target classification in sonar imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6284–6297, 2014. [xi](#), [40](#), [44](#), [53](#), [54](#), [55](#), [165](#)
- [4] O. Daniell, Y. Petillot, S. Reed, J. Vazquez, and A. Frau, “Reducing false alarms in automated target recognition using local sea-floor characteristics,” in *Sensor Signal Processing for Defence (SSPD)*, 2014. [xi](#), [35](#), [38](#), [39](#), [40](#), [49](#), [55](#), [57](#), [58](#), [166](#)
- [5] T. Lindeberg, *Scale-space theory in computer vision*, vol. 256. Springer Science & Business Media, 2013. [xi](#), [84](#), [85](#), [91](#)
- [6] G. Nichols, *Sedimentology and stratigraphy*. John Wiley & Sons, 2009. [xv](#), [157](#), [158](#)
- [7] J. Collinson, N. Mountney, and D. Thompson, *Sedimentary Structures*. Terra Publishing, London, 2006. [xv](#), [158](#)
- [8] D. P. Williams, “Fast target detection in synthetic aperture sonar imagery : A new algorithm and large-scale performance analysis,” *IEEE Journal of Oceanic Engineering*, vol. 40, no. 1, pp. 71–92, 2015. [xviii](#), [35](#), [219](#)
- [9] M. Felsberg and G. Sommer, “The monogenic signal,” *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, 2001. [xviii](#), [231](#), [235](#), [237](#), [241](#), [245](#)
- [10] M. P. Strand, “Underwater electro-optical system for mine identification,” in *SPIE’s 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics*, pp. 487–497, International Society for Optics and Photonics, 1995. [8](#)
- [11] J. S. Taylor and M. C. Hulgán, “Electro-optic identification research program,” in *OCEANS’02 MTS/IEEE*, vol. 2, pp. 994–1002, IEEE, 2002. [8](#)

- [12] H. L. Van Trees, *Detection, estimation, and modulation theory, optimum array processing*. John Wiley & Sons, 2004. 8
- [13] W. S. Burdic, *Underwater acoustic system analysis*. Prentice Hall, 1991. 8
- [14] R. J. Urick, *Principles of underwater sound for engineers*. McGraw-Hill Education, 1967. 8, 10, 11
- [15] M. Bouvet, *Traitements des signaux pour les systèmes sonar*. Collection technique et scientifique des télécommunications, Masson, 1992. 10
- [16] X. Lurton, *An introduction to underwater acoustics : principles and applications*. Springer Science & Business Media, 2002. 11, 16, 18
- [17] P. Blondel, *The handbook of sidescan sonar*. Springer Science & Business Media, 2010. 18
- [18] P. A. Kelly, H. Derin, and K. D. Hartt, “Adaptive segmentation of speckled images using a hierarchical random field model,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 10, pp. 1628–1641, 1988. 20
- [19] R. Martinsen, K. Kennedy, and A. Radl, “Speckle in laser imagery : efficient methods of quantification and minimization,” in *LEOS’99. IEEE Lasers and Electro-Optics Society 1999 12th Annual Meeting*, vol. 1, pp. 354–355, IEEE, 1999. 20
- [20] J. W. Goodman, “Some fundamental properties of speckle,” *JOSA*, vol. 66, no. 11, pp. 1145–1150, 1976. 20
- [21] A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser, “Optical coherence tomography-principles and applications,” *Reports on progress in physics*, vol. 66, no. 2, p. 239, 2003. 20
- [22] M. Simard, G. DeGrandi, K. P. Thomson, and G. B. Benie, “Analysis of speckle noise contribution on wavelet decomposition of SAR images,” *IEEE transactions on geoscience and remote sensing*, vol. 36, no. 6, pp. 1953–1962, 1998. 20
- [23] J.-S. Lee, M. R. Grunes, D. L. Schuler, E. Pottier, and L. Ferro-Famil, “Scattering-model-based speckle filtering of polarimetric SAR data,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 1, pp. 176–187, 2006. 20
- [24] J.-S. Lee, J.-H. Wen, T. L. Ainsworth, K.-S. Chen, and A. J. Chen, “Improved sigma filter for speckle filtering of SAR imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 202–213, 2009. 20

- [25] L. Gagnon and A. Jouan, "Speckle filtering of SAR images : a comparative study between complex-wavelet-based and standard filters," in *Optical Science, Engineering and Instrumentation'97*, pp. 80–91, International Society for Optics and Photonics, 1997. [20](#)
- [26] A. Achim, A. Bezerianos, and P. Tsakalides, "Novel bayesian multiscale method for speckle removal in medical ultrasound images," *IEEE transactions on medical imaging*, vol. 20, no. 8, pp. 772–783, 2001. [20](#)
- [27] O. V. Michailovich and A. Tannenbaum, "Despeckling of medical ultrasound images," *IEEE Transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 53, no. 1, pp. 64–78, 2006. [20](#)
- [28] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation : a survey," *IEEE Transactions on medical imaging*, vol. 25, no. 8, pp. 987–1010, 2006. [20](#)
- [29] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A versatile wavelet domain noise filtration technique for medical imaging," *IEEE transactions on medical imaging*, vol. 22, no. 3, pp. 323–331, 2003. [20](#)
- [30] T. Loupas, W. McDicken, and P. Allan, "An adaptive weighted median filter for speckle suppression in medical ultrasonic images," *IEEE transactions on Circuits and Systems*, vol. 36, no. 1, pp. 129–135, 1989. [20](#)
- [31] J. Chanussot, F. Maussang, and A. Hétet, "Scalar image processing filters for speckle reduction on synthetic aperture sonar images," in *OCEANS'02 MTS/IEEE*, vol. 4, pp. 2294–2301, IEEE, 2002. [20](#)
- [32] I. Leblond, M. Legris, and B. Solaiman, "Use of classification and segmentation of sidescan sonar images for long term registration," in *OCEANS'05 MTS/IEEE*, vol. 1, pp. 322–327, IEEE, 2005. [20](#), [45](#)
- [33] G. Padmavathi, P. Subashini, M. M. Kumar, and S. K. Thakur, "Comparison of filters used for underwater image pre-processing," *IJCSNS*, vol. 10, no. 1, pp. 58–65, 2010. [20](#)
- [34] A. P. Lyons and D. A. Abraham, "Statistical characterization of high-frequency shallow-water seafloor backscatter," *The Journal of the Acoustical Society of America*, vol. 106, no. 3, pp. 1307–1315, 1999. [22](#)
- [35] B. Kasatkin, "Anomalous phenomena in sound propagation near the sea floor : A review," *Acoustical Physics*, vol. 48, no. 4, pp. 379–387, 2002. [22](#)
- [36] D. R. Jackson, K. B. Briggs, K. L. Williams, and M. D. Richardson, "Tests of models for high-frequency seafloor backscatter," *IEEE journal of oceanic engineering*, vol. 21, no. 4, pp. 458–470, 1996. [22](#)

- [37] M. Tur, K.-C. Chin, and J. W. Goodman, "When is speckle noise multiplicative?," *Applied optics*, vol. 21, no. 7, pp. 1157–1159, 1982. [22](#)
- [38] R. L. Martin and R. W. Farwell, "High-frequency acoustic modeling," in *Proceedings of the IEEE OCEANS'94 Conference on Oceans Engineering for Today's Technology and Tomorrow's Preservation*, vol. 2, pp. II-272, IEEE, 1994. [23](#)
- [39] G. Le Chenadec, *Analyse de descripteurs énergétiques et statistiques de signaux sonar pour la caractérisation des fonds marins*. PhD thesis, Université de Bretagne Occidentale, 2004. [23](#), [24](#)
- [40] C.-h. Chen, *Handbook of pattern recognition and computer vision*. World Scientific, 2015. [24](#)
- [41] W. E. Stevens and F. Brimberg, "Automatic target detector," June 14 1977. US Patent 4,030,096. [34](#)
- [42] T. Aridgides, D. Antoni, M. F. Fernandez, and G. J. Dobeck, "Adaptive filter for mine detection and classification in side-scan sonar imagery," in *SPIE's 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics*, pp. 475–486, International Society for Optics and Photonics, 1995. [34](#)
- [43] J. C. Hyland and G. J. Dobeck, "Sea mine detection and classification using side-looking sonar," in *SPIE's 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics*, pp. 442–453, International Society for Optics and Photonics, 1995. [34](#)
- [44] G. J. Dobeck, J. C. Hyland, *et al.*, "Automated detection and classification of sea mines in sonar imagery," in *AeroSense'97*, pp. 90–110, International Society for Optics and Photonics, 1997. [34](#)
- [45] M. Doherty, J. Landowski, P. Maynard, G. Uber, D. Fries, and F. Maltz, "Side scan sonar object classification algorithms," in *Unmanned Untethered Submersible Technology, 1989. Proceedings of the 6th International Symposium on*, pp. 417–424, IEEE, 1989. [34](#)
- [46] M. G. Bello, "Markov random-field-based anomaly screening algorithm," in *SPIE's 1995 Symposium on OE/Aerospace Sensing and Dual Use Photonics*, pp. 466–474, International Society for Optics and Photonics, 1995. [34](#)
- [47] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Markov random field and fuzzy logic modeling in sonar imagery : application to the classification of underwater floor," *Computer Vision and Image Understanding*, vol. 79, no. 1, pp. 4–24, 2000. [34](#)
- [48] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Three-class markovian segmentation of high-resolution sonar images," *Computer Vision and Image Understanding*, vol. 76, no. 3, pp. 191–204, 1999. [34](#)

- [49] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Statistical model and genetic optimization : application to pattern detection in sonar images," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 5, pp. 2741–2744, IEEE, 1998. [34](#)
- [50] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy, "Unsupervised markovian segmentation of sonar images," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 2781–2784, IEEE, 1997. [34](#)
- [51] S. Reed, Y. Petillot, and J. Bell, "An automatic approach to the detection and extraction of mine features in sidescan sonar," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 1, pp. 90–105, 2003. [34](#), [38](#)
- [52] J. Bell, Y. Petillot, K. Lebart, S. Reed, E. Coiras, P. Mignotte, and H. Rohou, "Target recognition in synthetic aperture and high resolution sidescan sonar," in *High Resolution Imaging and Target Classification, 2006. The Institution of Engineering and Technology Seminar on*, pp. 99–106, IET, 2006. [34](#), [38](#)
- [53] S. Reed, Y. Petillot, and J. Bell, "Automated approach to classification of mine-like objects in sidescan sonar using highlight and shadow information," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 151, no. 1, pp. 48–56, 2004. [34](#), [38](#)
- [54] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001. [35](#)
- [55] B. Calder, L. Linnett, and D. Carmichael, "Spatial stochastic models for seabed object detection," in *AeroSense'97*, pp. 172–182, International Society for Optics and Photonics, 1997. [35](#)
- [56] B. Calder, L. Linnett, and D. Carmichael, "Bayesian approach to object detection in sidescan sonar," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 145, no. 3, pp. 221–228, 1998. [35](#)
- [57] F. Maussang, J. Chanussota, and A. Histetb, "Automated segmentation of sas images using the mean–standard deviation plane for the detection of underwater mines," in *in Proc. of MTS/IEEE Oceans 03 conference*, 2003. [35](#)
- [58] F. Maussang, J. Chanussot, A. Hétet, and M. Amate, "Mean-standard deviation representation of sonar images for echo detection : Application to sas images," *IEEE Journal of Oceanic Engineering*, vol. 32, no. 4, pp. 956–970, 2007. [35](#)
- [59] L. Linnett, D. Carmichael, S. Clarke, and A. Tress, "Texture analysis of sidescan sonar data," in *Texture analysis in radar and sonar, IEE Seminar on*, pp. 2–1, IET, 1993. [35](#)

- [60] J. M. Bell and L. Linnett, "Simulation and analysis of synthetic sidescan sonar images," *IEE Proceedings-radar, sonar and navigation*, vol. 144, no. 4, pp. 219–226, 1997. [35](#)
- [61] L. Atallah, C. Shang, and R. Bates, "Object detection at different resolution in archaeological side-scan sonar images," in *Europe Oceans 2005*, vol. 1, pp. 287–292, IEEE, 2005. [35](#)
- [62] T. G-Michael and J. D. Tucker, "Canonical correlation analysis for coherent change detection in synthetic aperture sonar imagery," *Institute of Acoustics Proceedings*, vol. 32, no. 4, pp. 117–122, 2010. [35](#)
- [63] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE transactions on information theory*, vol. 8, no. 2, pp. 179–187, 1962. [36](#)
- [64] M. R. Teague, "Image analysis via the general theory of moments," *JOSA*, vol. 70, no. 8, pp. 920–930, 1980. [36](#)
- [65] A. Sluzek, "Identification and inspection of 2-d objects using new moment-based shape descriptors," *Pattern Recognition Letters*, vol. 16, no. 7, pp. 687–697, 1995. [36](#)
- [66] J. C. Isaacs, "Sonar automatic target recognition for underwater uxo remediation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 134–140, 2015. [36](#)
- [67] E. Dura, J. Bell, and D. Lane, "Superellipse fitting for the recovery and classification of mine-like shapes in sidescan sonar images," *IEEE Journal of Oceanic Engineering*, vol. 33, no. 4, pp. 434–444, 2008. [37](#)
- [68] D. Köhntopp, B. Lehmann, and D. Kraus, "Efficient superellipse fitting based contour extraction for mine-like shape recognition," in *Proceedings of the 2st international conference and exhibition on Underwater Acoustics (UA2014)*, 2014. [37](#)
- [69] I. Quidu, J.-P. Malkasse, G. Burel, and P. Vilbé, "Mine classification using a hybrid set of descriptors," in *OCEANS 2000 MTS/IEEE Conference and Exhibition*, vol. 1, pp. 291–297, IEEE, 2000. [37](#)
- [70] D. Boulinguez and A. Quinquis, "Classification of underwater objects using fourier descriptors," in *Image Processing And Its Applications, 1999. Seventh International Conference on (Conf. Publ. No. 465)*, vol. 1, pp. 240–244, IET, 1999. [37](#)
- [71] R. Fandos and A. M. Zoubir, "Optimal feature set for automatic detection and classification of underwater objects in sas images," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 454–468, 2011. [37](#)

- [72] I. Quidu, J.-P. Malkasse, G. Burel, and P. Vilbé, "Mine classification based on raw sonar data : an approach combining fourier descriptors, statistical models and genetic algorithms," in *OCEANS 2000 MTS/IEEE Conference and Exhibition*, vol. 1, pp. 285–290, IEEE, 2000. [37](#)
- [73] J. T. Cobb and J. R. Stack, "In situ adaptive feature extraction for underwater target classification," in *36th Applied Imagery Pattern Recognition Workshop (aipr 2007)*, pp. 42–47, IEEE, 2007. [38](#)
- [74] E. Coiras, P.-Y. Mignotte, Y. Petillot, J. Bell, and K. Lebart, "Supervised target detection and classification by training on augmented reality data," *IET Radar, Sonar & Navigation*, vol. 1, no. 1, pp. 83–90, 2007. [38](#)
- [75] P.-Y. Mignotte, E. Coiras, H. Rohou, Y. Petillot, J. Bell, and K. Lebart, "Adaptive fusion framework based on augmented reality training," *IET Radar, Sonar & Navigation*, vol. 2, no. 2, pp. 146–154, 2008. [38](#)
- [76] S. Reed, I. T. Ruiz, C. Capus, and Y. Petillot, "The fusion of large scale classified side-scan sonar image mosaics," *IEEE Transactions on Image Processing*, vol. 15, no. 7, pp. 2049–2060, 2006. [38](#), [56](#)
- [77] C. Rao, K. Mukherjee, S. Gupta, A. Ray, and S. Phoha, "Underwater mine detection using symbolic pattern analysis of sidescan sonar images," in *2009 American Control Conference*, pp. 5416–5421, IEEE, 2009. [38](#)
- [78] E. Hasanbelliu, J. Principe, and C. Slatton, "Correntropy based matched filtering for classification in sidescan sonar imagery," in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, pp. 2757–2762, IEEE, 2009. [38](#), [39](#)
- [79] J. Groen, E. Coiras, J. D. R. Vera, and B. Evans, "Model-based sea mine classification with synthetic aperture sonar," *IET radar, sonar & navigation*, vol. 4, no. 1, pp. 62–73, 2010. [38](#)
- [80] D. P. Williams, "Fast unsupervised seafloor characterization in sonar imagery using lacunarity," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, pp. 6022–6034, Nov. 2015. [39](#), [40](#), [173](#)
- [81] M. Geilhufe and Ø. Midtgaard, "Quantifying the complexity in sonar image for mcm performance estimation," in *Proc. 2nd International Conference and Exhibition on Underwater Acoustics*, pp. 1041–1048, 2014. [39](#), [40](#), [44](#), [49](#)
- [82] O. Daniell, Y. Petillot, and S. Reed, "Unsupervised seafloor classification for automatic target recognition," in *International Conference on Detection and Classification of Underwater Targets*, (Brest), Cambridge Scholars, 2012. [40](#), [44](#), [47](#), [48](#)

- [83] D. P. Williams, "Unsupervised seabed segmentation of synthetic aperture sonar imagery via wavelet features and spectral clustering," in *16th IEEE International Conference on Image Processing*, pp. 557–560, 2009. [40](#)
- [84] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *JOSA A*, vol. 2, no. 7, pp. 1160–1169, 1985. [44](#)
- [85] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf : Speeded up robust features," in *European conference on computer vision*, pp. 404–417, Springer, 2006. [47](#)
- [86] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004. [47](#)
- [87] B. Lehmann, K. Siantidis, I. Aleksis, and D. Kraus, "Efficient pre-segmentation algorithm for sidescan-sonar images," in *Image and Signal Processing and Analysis (ISPA), 2011 7th International Symposium on*, pp. 471–475, IEEE, 2011. [47](#)
- [88] D. P. Williams, "On adaptive underwater object detection," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4741–4748, IEEE, 2011. [47](#)
- [89] C. Messom and A. Barczak, "Fast and efficient rotated haar-like features using rotated integral images," in *Australian Conference on Robotics and Automation*, pp. 1–6, 2006. [48](#)
- [90] A. L. Barczak, "Toward an efficient implementation of a rotation invariant detector using haar-like features," *Proceedings of the IVCNZ'05, Dunedin, New Zealand*, pp. 31–36, 2005. [48](#)
- [91] H. Choi, J. Romberg, R. Baraniuk, and N. Kingsbury, "Hidden markov tree modeling of complex wavelet transforms," in *Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on*, vol. 1, pp. 133–136, IEEE, 2000. [49](#)
- [92] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, "The dual-tree complex wavelet transform," *IEEE signal processing magazine*, vol. 22, no. 6, pp. 123–151, 2005. [49](#), [51](#)
- [93] S. G. Mallat, "A theory for multiresolution signal decomposition : the wavelet representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 11, no. 7, pp. 674–693, 1989. [50](#)
- [94] J. Romberg, H. Choi, R. Baraniuk, and N. Kingbury, "Multiscale classification using complex wavelets and hidden markov tree models," in *Image Processing, 2000. Proceedings. 2000 International Conference on*, vol. 2, pp. 371–374, IEEE, 2000. [50](#)

- [95] M. Victor and P. Ale, "Complex wavelet transform in signal and image analysis," *Institute of Chemical Technology, Department of Computing and Control Engineering, Technicka*, 2004. [50](#), [51](#)
- [96] N. Kingsbury, "Image processing with complex wavelets," *Philosophical Transactions of the Royal Society of London A : Mathematical, Physical and Engineering Sciences*, vol. 357, no. 1760, pp. 2543–2560, 1999. [50](#)
- [97] D. Gabor, "Theory of communication," *Journal of the Institution of Electrical Engineers - Part III : Radio and Communication Engineering*, vol. 93, no. 26, pp. 429–441, 1946. [66](#)
- [98] J. d. Ville, "Théorie et applications de la notion de signal analytique," *Cables et transmission*, vol. 2, no. 1, pp. 61–74, 1948. [66](#)
- [99] G. H. Hardy, "The elementary theory of cauchy's principal values," *Proceedings of the London Mathematical Society*, vol. 1, no. 1, pp. 16–40, 1901. [66](#)
- [100] H. Knutsson and G. H. Granlund, "Fourier domain design of line and edge detectors," in *The 5th International Conference on Pattern Recognition, Miami, Florida, USA*, 1980. [69](#)
- [101] G. H. Granlund and H. Knutsson, *Signal Processing for Computer Vision*. Boston, MA : Springer US, 1995. [69](#), [70](#)
- [102] H. Stark, "An extension of the hilbert transform product theorem," *Proceedings of the IEEE*, vol. 59, no. 9, pp. 1359–1360, 1971. [71](#), [228](#)
- [103] S. L. Hahn, "Multidimensional complex signals with single-orthant spectra," *Proceedings of the IEEE*, vol. 80, no. 8, pp. 1287–1300, 1992. [71](#), [229](#)
- [104] S. L. Hahn, *Hilbert Transforms in Signal Processing*. Artech House signal processing library, Artech House, 1996. [72](#), [227](#)
- [105] T. Bülow, *Hypercomplex spectral signal representations for the processing and analysis of images*. Universität Kiel. Institut für Informatik und Praktische Mathematik, 1999. [72](#), [230](#)
- [106] M. Felsberg and G. Sommer, "The monogenic scale-space : A unifying approach to phase-based image processing in scale-space," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 1-2, pp. 5–26, 2004. [72](#), [73](#), [86](#), [88](#), [89](#), [99](#), [227](#)
- [107] L. Dorst, D. Fontijne, and S. Mann, *Geometric algebra for computer science : an object-oriented approach to geometry*. Morgan Kaufmann Publishers Inc., 2009. [72](#)

- [108] M. Felsberg and G. Sommer, "Image features based on a new approach to 2d rotation invariant quadrature filters," in *European Conference on Computer Vision*, pp. 369–383, Springer, 2002. [73](#), [127](#)
- [109] M. Felsberg, *Low-level Image Processing with the Structure Multivector*. Inst. für Informatik und Praktische Mathematik, 2002. [73](#), [81](#), [99](#)
- [110] M. Riesz, "Sur les fonctions conjuguées," *Mathematische Zeitschrift*, vol. 27, no. 1, pp. 218–244, 1928. [73](#)
- [111] E. M. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*. Princeton university press, 1971. [73](#)
- [112] M. Unser, D. Sage, and D. Van De Ville, "Multiresolution monogenic signal analysis using the riesz-laplace wavelet transform," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2402–2418, 2009. [75](#)
- [113] R. Soulard, *Ondelettes analytiques et monogènes pour la représentation des images couleur*. PhD thesis, Université de Poitiers, 2012. [76](#)
- [114] J. Weickert, S. Ishikawa, and A. Imiya, "Linear scale-space has first been proposed in japan," *Journal of Mathematical Imaging and Vision*, vol. 10, no. 3, pp. 237–252, 1999. [83](#)
- [115] T. Iijima, "Basic theory of pattern observation," *Papers of Technical Group on Automata and Automatic Control, IECE*, december 1959. [83](#), [85](#), [90](#)
- [116] A. Witkin, "Scale-space filtering : A new approach to multi-scale description," in *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP'84*, vol. 9, pp. 150–153, IEEE, 1984. [83](#), [85](#), [90](#)
- [117] T. Lindeberg, *Scale-space theory in computer vision*, vol. 256. Springer Science & Business Media, 2013. [84](#)
- [118] T. Lindeberg, "Scale-space theory : A basic tool for analyzing structures at different scales," *Journal of applied statistics*, vol. 21, no. 1-2, pp. 225–270, 1994. [84](#)
- [119] 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*, vol. 11, no. 3, pp. 283–318, 1993. [84](#)
- [120] J. J. Koenderink, "The structure of images," *Biological cybernetics*, vol. 50, no. 5, pp. 363–370, 1984. [84](#), [85](#), [88](#), [90](#)
- [121] D. V. Widder, *The heat equation*, vol. 67. Academic Press, 1976. [84](#), [85](#)

- [122] D. Lowe, "Object recognition from local scale-invariant features," in *The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999*, vol. 2, pp. 1150–1157 vol.2, 1999. [85](#), [91](#), [94](#)
- [123] E. Schrödinger, "About Heisenberg uncertainty relation," *Proceedings of The Prussian Academy of Sciences Physics-Mathematical Section*, pp. 296–303, 1930. [87](#)
- [124] 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*, no. 1, pp. 26–33, 1986. [88](#)
- [125] R. Duits, L. Florack, J. De Graaf, and B. ter Haar Romeny, "On the axioms of scale space theory," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 3, pp. 267–298, 2004. [89](#)
- [126] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Transactions on Graphics (TOG)*, vol. 2, no. 4, pp. 217–236, 1983. [91](#)
- [127] J. L. Crowley and A. C. Parker, "A representation for shape based on peaks and ridges in the difference of low-pass transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 2, pp. 156–170, 1984. [91](#)
- [128] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London B : Biological Sciences*, vol. 207, no. 1167, pp. 187–217, 1980. [94](#)
- [129] T. Houtgast and H. J. Steeneken, "A review of the mtf concept in room acoustics and its use for estimating speech intelligibility in auditoria," *The Journal of the Acoustical Society of America*, vol. 77, no. 3, pp. 1069–1077, 1985. [94](#)
- [130] N. R. Council, *Naval Mine Warfare : Operational and Technical Challenges for Naval Forces*. The National Academies Press, 2001. [103](#), [217](#)
- [131] N.-E. Lasmar, A. Baussard, and G. Le Chenadec, "Asymmetric power distribution model of wavelet subbands for texture classification," *Pattern Recognition Letters*, vol. 52, pp. 1–8, 2015. [112](#)
- [132] A. Baussard, "Bayesian texture classification using steerable riesz wavelets : Application to sonar images," in *OCEANS 2015-MTS/IEEE Washington*, pp. 1–6, IEEE, 2015. [114](#)
- [133] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979. [115](#)
- [134] T. Randen and J. H. Husoy, "Filtering for texture classification : A comparative study," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 21, no. 4, pp. 291–310, 1999. [115](#)

- [135] P. P. Ohanian and R. C. Dubes, "Performance evaluation for four classes of textural features," *Pattern recognition*, vol. 25, no. 8, pp. 819–833, 1992. [115](#)
- [136] M. F. Augusteijn, L. E. Clemens, and K. A. Shaw, "Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, no. 3, pp. 616–626, 1995. [115](#)
- [137] S. Li and J. Shawe-Taylor, "Comparison and fusion of multiresolution features for texture classification," *Pattern Recognition Letters*, vol. 26, no. 5, pp. 633–638, 2005. [115](#)
- [138] M. Unser, "Texture classification and segmentation using wavelet frames," *IEEE Transactions on image processing*, vol. 4, no. 11, pp. 1549–1560, 1995. [115](#)
- [139] I. Karoui, R. Fablet, J.-M. Boucher, and J.-M. Augustin, "Seabed segmentation using optimized statistics of sonar textures," *IEEE Transactions on geoscience and remote sensing*, vol. 47, no. 6, pp. 1621–1631, 2009. [115](#)
- [140] A. Nait-Chabane, B. Zerr, and G. Le Chenadec, "Sidescan sonar imagery segmentation with a combination of texture and spectral analysis," in *OCEANS MTS/IEEE conference*, pp. 1–6, 2013. [115](#)
- [141] K. Fukunaga, "Intrinsic dimensionality extraction, in : Classification, pattern recognition and reduction of dimensionality," *Vol. 2 of Handbook of Statistics*,, pp. 347–362, 1982. [120](#)
- [142] E. Levina and P. J. Bickel, "Maximum likelihood estimation of intrinsic dimension," in *Advances in neural information processing systems*, pp. 777–784, 2004. [120](#)
- [143] P. J. Verwee and R. P. W. Duin, "An evaluation of intrinsic dimensionality estimators," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 17, no. 1, pp. 81–86, 1995. [120](#)
- [144] K. Fukunaga, "15 intrinsic dimensionality extraction," *Handbook of Statistics*, vol. 2, pp. 347–360, 1982. [120](#)
- [145] M. Hein and J.-Y. Audibert, "Intrinsic dimensionality estimation of submanifolds in  $\mathbb{R}^d$ ," in *Proceedings of the 22nd international conference on Machine learning*, pp. 289–296, ACM, 2005. [120](#)
- [146] C. Zetsche and E. Barth, "Fundamental limits of linear filters in the visual processing of two-dimensional signals," *Vision Research*, vol. 30, no. 7, pp. 1111–1117, 1990. [120](#), [122](#)
- [147] G. Krieger and C. Zetsche, "Nonlinear image operators for the evaluation of local intrinsic dimensionality," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp. 1026–1042, 1996. [122](#)

- [148] G. Krieger, I. Rentschler, G. Hauske, K. Schill, and C. Zetsche, "Object and scene analysis by saccadic eye-movements : an investigation with higher-order statistics," *Spatial vision*, vol. 13, no. 2, pp. 201–214, 2000. [122](#)
- [149] G. Krieger, C. Zetsche, and E. Barth, "Higher-order statistics of natural images and their exploitation by operators selective to intrinsic dimensionality," in *Higher-Order Statistics, 1997., Proceedings of the IEEE Signal Processing Workshop on*, pp. 147–151, IEEE, 1997. [122](#)
- [150] W. Förstner, "Image preprocessing for feature extraction in digital intensity, color and range images," in *Geomatic Method for the Analysis of Data in the Earth Sciences*, pp. 165–189, Springer, 2000. [123](#)
- [151] N. Krüger and M. Felsberg, "A continuous formulation of intrinsic dimension," in *British Machine Vision Conference* (R. Harvey and A. Bangham, eds.), 2003. [123](#), [127](#)
- [152] M. Felsberg and N. Krüger, "A probabilistic definition of intrinsic dimensionality for images," in *Joint Pattern Recognition Symposium*, pp. 140–147, Springer, 2003. [125](#), [128](#)
- [153] J. Bigun, "Optimal orientation detection of linear symmetry," 1987. [127](#)
- [154] W. Förstner and E. Gülch, "A fast operator for detection and precise location of distinct points, corners and centres of circular features," in *Proc. ISPRS intercommission conference on fast processing of photogrammetric data*, pp. 281–305, 1987. [127](#)
- [155] G. Farneback, "Fast and accurate motion estimation using orientation tensors and parametric motion models," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 1, pp. 135–139, IEEE, 2000. [127](#)
- [156] N. Krüger, "Learning object representations using a priori constraints within orassyll," *Neural Computation*, vol. 13, no. 2, pp. 389–410, 2001. [128](#)
- [157] L. Picard, A. Baussard, G. Le Chenadec, and I. Quidu, "Potential of the intrinsic dimensionality for characterizing the seabed in the atr context," in *OCEANS 2015-Genova*, pp. 1–7, IEEE, 2015. [151](#)
- [158] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012. [192](#)
- [159] W. D. Fisher, "On grouping for maximum homogeneity," *Journal of the American statistical Association*, vol. 53, no. 284, pp. 789–798, 1958. [192](#)
- [160] J. Esclarín and L. Alvarez, "Image quantization using reaction-diffusion equations," *SIAM Journal on Applied Mathematics*, vol. 57, no. 1, pp. 153–175, 1997. [192](#)

- [161] C.-I. Chang, K. Chen, J. Wang, and M. L. Althouse, "A relative entropy-based approach to image thresholding," *Pattern recognition*, vol. 27, no. 9, pp. 1275–1289, 1994. [192](#)
- [162] H.-D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2071–2082, 2000. [192](#)
- [163] J. N. Kapur, P. K. Sahoo, and A. K. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer vision, graphics, and image processing*, vol. 29, no. 3, pp. 273–285, 1985. [192](#)
- [164] J. Delon, A. Desolneux, J.-L. Lisani, and A. B. Petro, "A nonparametric approach for histogram segmentation," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 253–261, 2007. [192](#)
- [165] D. P. Williams, "Auv-enabled adaptive underwater surveying for optimal data collection," *Intelligent Service Robotics*, vol. 5, no. 1, pp. 33–54, 2012. [202](#)
- [166] H. Rohling, "Radar cfar thresholding in clutter and multiple target situations," *IEEE transactions on aerospace and electronic systems*, no. 4, pp. 608–621, 1983. [210](#)
- [167] M. Richards, *Fundamentals of Radar Signal Processing*. Mcgraw Hill, 2005. [212](#)
- [168] T. Bülow and G. Sommer, "Hypercomplex signals - a novel extension of the analytic signal to the multidimensional case," *Signal Processing, IEEE Transactions on*, vol. 49, no. 11, pp. 2844–2852, 2001. [230](#)
- [169] I. L. Kantor and A. S. Solodovnikov, *Hypercomplex numbers : an elementary introduction to algebras*. Springer, 1989. [230](#)
- [170] C. Perwass, *Geometric Algebra with Applications in Engineering*, vol. 4 of *Geometry and Computing*. Berlin, Heidelberg : Springer Berlin Heidelberg, 2009. [233](#)
- [171] P. Lounesto, *Clifford algebras and spinors*. London Mathematical Society lecture note series, Cambridge, New York : Cambridge University Press, 2001. [237](#)