



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Institut National Polytechnique de Toulouse*
Discipline ou spécialité : *Signal, Image, Acoustique et Optimisation*

Présentée et soutenue par *Xavier DELAUNAY*
Le 12 novembre 2008

Titre :
*Compression d'images satellite
par post-transformées dans le domaine ondelettes*

JURY

<i>Sylvie ROQUES, Directeur de recherche CNRS, OMP Toulouse</i>	<i>Président</i>
<i>Christine GUILLEMOT, Directeur de recherche INRIA, INRIA Rennes</i>	<i>Rapporteur</i>
<i>Marc ANTONINI, Directeur de recherche CNRS, Université de Nice Sophia Antipolis</i>	<i>Rapporteur</i>
<i>Francis CASTANIE, Professeur des universités, INP Toulouse</i>	<i>Membre</i>
<i>Gabriel PEYRE, Chargé de recherche CNRS, Université de Paris Dauphine</i>	<i>Membre</i>
<i>Marie CHABERT, Maître de conférence, INP Toulouse,</i>	<i>Membre</i>
<i>Carole THIEBAUT, Ingénieur, CNES Toulouse,</i>	<i>Membre</i>
<i>Vincent GERMAIN, Ingénieur, société NOVELTIS, Ramonville St Agne</i>	<i>Membre</i>

Ecole doctorale : *Mathématiques, Informatique et Télécommunications de Toulouse*
Unité de recherche : *Laboratoire TéSA*
Directeur(s) de Thèse : *Francis CASTANIE*

COMPRESSION D'IMAGES SATELLITE PAR POST-TRANSFORMÉES DANS LE DOMAINE ONDELETTES

Résumé

Le CNES s'intéresse aux nouvelles transformées pour accroître les performances de la compression d'images à bord des satellites d'observation de la Terre. Dans cette thèse nous étudions les post-transformées. Elles sont appliquées après la transformée en ondelettes. Chaque bloc de coefficients d'ondelettes est de nouveau transformé dans une base sélectionnée dans un dictionnaire par minimisation d'un critère débit-distorsion. Nous commençons par mettre en évidence des dépendances entre coefficients d'ondelettes qui limitent les performances en compression. Nous étudions ensuite la transformée en bandelettes par blocs, qui est à l'origine des post-transformées, et nous en optimisons les paramètres pour la compression d'images satellite. En particulier nous adaptons la méthode d'optimisation de Shoham et Gersho à la sélection de la meilleure base de bandelettes. Nous en déduisons une formule du multiplicateur de Lagrange optimal qui intervient dans le critère de sélection. Dans un deuxième temps, nous analysons les dépendances entre coefficients d'ondelettes qui ne sont pas prises en compte par les bandelettes et nous définissons de nouvelles bases de post-transformées. Les bases construites par ACP minimisent les corrélations entre coefficients post-transformés et compactent l'énergie de chaque bloc sur un petit nombre de coefficients. Cette propriété est exploitée lors du codage entropique. Enfin, nous modifions le critère de sélection des bases pour adapter la post-transformée à une compression progressive. Nous employons alors la post-transformée de Hadamard dans le codeur du CCSDS le tout ayant une faible complexité calculatoire.

Mots-clés : Compression d'images, applications satellites, post-transformées, transformée en bandelettes, transformée en ondelettes, optimisation débit-distorsion, codage par plans de bits.

SATELLITE IMAGE COMPRESSION BY POST-TRANSFORMS IN THE WAVELET DOMAIN

Abstract

The French Space Agency, CNES, is interested in the transforms derived from the wavelets in order to increase the image compression efficiency on-board of Earth observation satellites. In this thesis, the post-transforms are studied. They are employed after the wavelet transform. Each block of wavelet coefficients is further transformed in a basis selected among a dictionary by minimization of a rate-distortion criterion. First, we emphasize dependencies between wavelet coefficients limiting the compression efficiency. Then, we study the bandelet transform by blocks, from which the post-transforms derive, and we optimize its parameter for the compression of satellite images. Particularly, we adapt Shoham and Gersho optimization method to the problem of the selection of the best bandelet basis. We deduce from these results an expression of the optimal Lagrangian multiplier used in the rate-distortion criterion. Next, we analyze dependencies between wavelet coefficient which are not exploited by the bandelet transform and we define new post-transform bases. Bases build by PCA minimize the correlations between post-transformed coefficients and compact the energy of each block on a small number of coefficients. This feature is exploited during the entropy coding process. Last, we modify the bases selection criterion to adapt the post-transform to progressive compression schemes. We then employ the Hadamard post-transform with the CCSDS image encoder to obtain a low computational complexity yet efficient compression scheme.

Keywords: Image coding, satellite applications, post-transforms, bandelet transform, wavelet transform, rate-distortion optimization, bit-plane coding.

Remerciements

C'est peut-être la première page que vous lisez mais pour moi ces dernières pages clôturent trois années de dur labeur. J'ai mené ces travaux de thèse accompagné par une véritable équipe d'encadrement constituée de six personnes. Cela n'a pas toujours été facile de trouver des dates convenant à tous pour les réunions d'avancement. J'ai souvent dû ré-expliquer plusieurs fois mes travaux. Cela m'a poussé à écrire de nombreux rapports qui se sont révélés être très utiles pour la rédaction de publications et surtout pour la rédaction de ce manuscrit.

Je tiens d'abord à remercier tout particulièrement Carole Thiebaut qui est l'instigatrice de ces travaux. Je la remercie pour son accueil au CNES et pour m'avoir toujours soutenu dans les pistes de recherche que j'ai choisi d'explorer. Un grand merci à Marie Chabert qui a toujours relu mes écrits avec énormément d'attention et y a apporté d'innombrables corrections. En particulier, je la remercie pour avoir pris le temps de corriger ce manuscrit pendant ses vacances, pour ses nombreux encouragements et ses notes d'humour. J'adresse aussi mes remerciements à Vincent Charvillat pour son enthousiasme souvent vif et manifeste au cours des réunions et pour les nombreuses idées de voies à explorer qu'il a suggérées, bien que je ne les aie pas toutes suivies. Je le remercie aussi pour ses conseils qui m'ont aidé à choisir mon poste après la thèse. Merci à Géraldine Morin pour m'avoir donnée une méthode de travail très efficace en cas de double emploi. Cela m'a été très utile pour préparer ma soutenance et pour finir d'écrire ce manuscrit. Je la remercie aussi pour ses corrections, en particulier sur mes publications en anglais. Je remercie aussi Rosa Ruiloba pour son accueil à Noveltis et pour son soutien tout au long de ma thèse. Enfin, je remercie Emmanuel Christophe qui est venu renforcer cette équipe d'encadrement après avoir soutenu sa thèse. Je le remercie grandement pour son aide technique précieuse. Merci à tous les six pour votre formidable travail d'encadrement et pour les nombreuses relectures et corrections apportées à toutes mes publications.

Je remercie aussi Francis Castanié pour avoir accepté le rôle de directeur de thèse et pour m'avoir accueilli au laboratoire TéSA. J'adresse aussi mes remerciements à Christine Guillemot et Marc Antonini pour leur lecture minutieuse de ce manuscrit et pour avoir accepté de rapporter cette thèse. Je remercie également Sylvie Roques qui m'a fait l'honneur de présider le jury. Merci à Gabriel Peyré pour les nombreuses discussions à propos des bandelettes et des post-transformées, pour les composantes logicielles qu'il m'a fournies ainsi que pour avoir accepté de participer au jury. Je remercie aussi Vincent Germain qui a participé au jury.

J'ai eu la chance de voir trois environnements de travail différents : le CNES, le la-

boratoire TéSA et Noveltis. Cela a permis énormément de rencontres enrichissantes. Je suis d'abord passé au CNES. Je remercie Gilles pour m'avoir accueilli dans son service, l'indispensable Mylène, Mathieu, Fabien, Bruno, Vivient, Roberto, Florent, sans oublier Michel qui a partagé mon bureau, Pierre, Lorraine, Laurent, Benoît, François, Sébastien et aussi Christophe.

Merci aux collègues du TéSA, à commencer par Marie-Jo et Sarah qui ont été d'une efficacité redoutable pour gérer mes problèmes de missions, Manu qui a partagé mon bureau et m'a souvent distrait avec des calembours plus que douteux (non, ne cherche pas il n'y a pas de contrepèterie), Mariana et Anchalee qui ont voyagé avec moi et avec qui j'ai pas mal randonné dans les Pyrénées. Merci aussi à Patrice, Lucile, Ferdinand, Wil, David, Corine et Julien pour leur sympathie. Merci aussi aux personnes de l'autre côté du canal avec qui j'ai pu voyager à Troyes mais aussi à Hawaï et à Las Vegas : Florent, Nicolas, Jean-Yves et Martial.

Je tiens à remercier Richard Bru pour m'avoir accueilli à Noveltis, pour la confiance qu'il m'a accordée au début de cette thèse et qu'il a depuis renouvelée. Merci aussi aux collègues de Noveltis pour leur sympathie : Jérôme, François et Julien qui ont usé les prises de la salle d'escalade avec moi, Jean-Yves qui a pris le temps de relire mon manuscrit, Annie, David, Gwenaëlle, Tristan, Emmanuel, Arnaud, Christian, Elodie et Emilie.

Merci aussi aux stagiaires qui ont contribué à cette thèse : Yan, David, Florin, Fabien, Reza et Thomas.

Enfin, je souhaite remercier mes parents qui m'ont aidé à me hisser jusque là et qui m'ont toujours conforté dans mes choix professionnels.

Maintenant, je vous invite à lire la suite. Aller, un peu de courage. Au moins le premier chapitre. Il n'y a rien de bien compliqué. Bonne lecture.

Gendibal watched intently. There was little redundancy, if any, in mentalist-vision. In true vision, or even in physical hypervision across the parsecs, one saw enormously more in the way of information bits than was absolutely necessary for comprehension and one could miss a great deal without losing anything significant.

Isaac Asimov

Table des matières

Introduction	1
1 Images d’observation de la Terre et compression à bord des satellites	5
1.1 Un peu d’histoire	6
1.2 Les missions d’observation de la Terre au CNES	6
1.2.1 SPOT 1 à SPOT 4	6
1.2.2 SPOT 5	7
1.2.3 PLEIADES	7
1.3 Les images satellite	8
1.3.1 Caractéristiques des images satellite	8
1.3.2 La chaîne de traitement d’image bord-sol	11
1.4 La compression d’images satellite	12
1.4.1 Les étapes de compression	12
1.4.2 Mesurer les performances en compression	14
1.4.3 Compression à bord : Besoins et contraintes	15
1.5 Préparer l’avenir	17
2 Dépendances et compression	19
2.1 La transformée en ondelettes	20
2.1.1 Pour la compression d’images satellite	20
2.1.2 Dépendances entre coefficients d’ondelettes	21
2.2 Exploitation des dépendances dans les codeurs par plans de bits	26
2.2.1 BPE : le codeur du standard CCSDS	26
2.2.2 EBCOT : le codeur de la norme JPEG2000	29
2.2.3 Les codeurs morphologiques	32
2.3 Transformées dérivées des ondelettes	34
2.3.1 Transformées redondantes	35
2.3.2 Les bandelettes	36
2.4 Conclusion	38
3 Les bandelettes appliquées à la compression d’images satellite	41
3.1 La transformée en bandelettes par blocs	42
3.1.1 Construction du dictionnaire de bases	42
3.1.2 Transformée en bandelettes d’un bloc de coefficients d’ondelettes	45

3.1.3	Sélection de la meilleure base de bandelettes	47
3.1.4	Algorithme de compression en bandelettes	50
3.2	Choix de la taille des blocs	52
3.2.1	Corrélations entre coefficients d'ondelettes	52
3.2.2	Extension des bases de bandelettes à des blocs de tailles diverses . .	53
3.3	Sélection des meilleures bases de bandelettes par optimisation débit-distorsion	54
3.3.1	Le problème d'allocation débit-distorsion	54
3.3.2	Calcul théorique du multiplicateur de Lagrange optimal	56
3.3.3	Recherche empirique du multiplicateur de Lagrange optimal	59
3.3.4	Conclusion	68
3.4	Études expérimentales des performances de la transformée en bandelettes .	68
3.4.1	Analyse de la sélection des bases directionnelles	68
3.4.2	Estimation du débit des coefficients de bandelettes	70
3.4.3	Estimation du débit de signalisation	71
3.4.4	Performances du critère de sélection des bases	75
3.4.5	Performances des bases de bandelettes	77
3.5	Conclusion	81
4	Nouvelles bases pour la post-transformée	83
4.1	Analyse des redondances exploitées par les bases de bandelettes	84
4.1.1	Construction des classes de blocs	85
4.1.2	Informations mutuelles entre coefficients d'ondelettes dans un bloc .	85
4.1.3	Corrélations entre coefficients d'ondelettes dans un bloc	86
4.2	Bases construites par extensions des groupements	88
4.2.1	Construction des nouvelles bases	89
4.2.2	Analyse de l'inter-corrélation entre coefficients transformés	89
4.2.3	Performances en compression	92
4.3	Bases distantes entre elles	93
4.3.1	Cohérence d'un dictionnaire	93
4.3.2	Dissimilarité entre bases	95
4.4	Bases construites par ACP	96
4.4.1	ACP par sous-bandes	96
4.4.2	ACP sur différentes classes de blocs	98
4.4.3	Exploitation de la compaction d'énergie	101
4.5	Conclusion	105
5	Post-transformées et codage par plans de bits	107
5.1	Sélection de la post-transformée pour un codage par plans de bits	108
5.1.1	Position du problème	108
5.1.2	Sélection de la post-transformée par minimisation l_1	109
5.1.3	Performances en compression	110
5.1.4	Conclusion	112
5.2	La post-transformée dans le codeur du CCSDS	112

5.2.1	Post-transformées et arbres de zéros	112
5.2.2	Choix du dictionnaire de post-transformées	113
5.2.3	Post-transformée des petits-enfants	113
5.2.4	Post-transformée des enfants	117
5.2.5	Perspectives de modification de l'algorithme du CCSDS	118
5.3	Conclusion	119
Conclusions et perspectives		121
A Images		125
A.1	Ensemble d'images de test	125
A.2	Ensemble d'images d'apprentissage	127
B Exemple de codage avec le BPE du standard CCSDS		129
B.1	Exemple utilisé	129
B.2	Codage avec l'algorithme du CCSDS original	130
B.3	Codage avec l'algorithme modifié	137
C Exemple de codage avec EBCOT		141
C.1	Exemple utilisé	141
C.2	Codage du premier plan de bits	141
C.3	Codage du second plan de bits	145
C.4	Suite du codage	148
D Preuves diverses		151
D.1	Théorème d'optimisation de l'allocation débit-distorsion	151
D.2	Calcul de la distorsion sous hypothèse de quantification haute résolution .	152
D.3	Calcul du débit sous hypothèse de quantification haute résolution	153
D.4	Cohérence d'un dictionnaire de vecteurs	153
D.5	Dissimilarité entre bases	154
Bibliographie		157

Introduction

LE volume des images d'observation de la Terre acquises par satellites est actuellement en nette augmentation. Cela est principalement dû à l'évolution des besoins utilisateurs. Aujourd'hui ces derniers souhaitent des images disponibles dans un délai très court après leur acquisition et actualisées fréquemment. Ils ont besoin d'images multi-spectrales et d'une résolution sub-métrique. La compression des images à bord des satellites permet de réduire le volume des données transmises au sol et donc de transmettre un plus grand nombre d'images dans un laps de temps identique. Les techniques de compression actuellement utilisées à bord des satellites dérivent des techniques employées pour la compression d'images au sol. Ayant le souci de préparer les techniques futures pour répondre aux besoins croissants de compression à bord, le CNES s'intéresse aux nouvelles transformées et techniques de codage qui permettraient d'obtenir un gain en taux de compression pour une qualité-image identique. Les post-transformées présentées dans cette thèse s'inspirent d'une version de la transformée en bandelettes adaptée à la compression d'images satellite permettant de conserver une grande qualité d'image. Le but de ce travail n'est pas de proposer un algorithme de compression qui soit implantable dès aujourd'hui sur des circuits imprimés « qualifiés spatial » mais d'explorer les possibilités offertes par cette nouvelle technique de compression.

Le premier chapitre est la véritable introduction à ce travail doctoral et présente la compression bord des satellites d'observation de la Terre. L'augmentation du besoin en compression est mis en évidence par un historique des missions civiles d'observation de la Terre du programme SPOT conduit par le CNES. Les caractéristiques des différents types d'images satellite sont présentées ainsi que les traitements conduisant à un produit commercialisable. Les méthodes de compression et les contraintes spécifiques à bord des satellites sont également détaillées.

Le second chapitre porte sur la transformée en ondelettes qui est utilisée dans les systèmes de compression les plus récents, à bord des satellites comme au sol. Nous mettons en évidence les corrélations et l'information résiduelle entre coefficients d'ondelettes pour montrer les limites de cette transformée pour la compression. Plusieurs familles d'algorithmes de codage par plans de bits emploient différentes techniques afin d'exploiter ces dépendances. En particulier, nous analysons les performances des contextes définis dans le codeur de JPEG2000. Enfin, de nouvelles transformées dérivées des ondelettes sont décrites. Parmi elles, la transformée en bandelettes possède des propriétés intéressantes

pour une application en compression. La transformée en bandelettes par blocs est plus efficace que la transformée en ondelettes pour la compression d'images satellite.

Dans le troisième chapitre, nous étudions cette transformée en bandelettes et nous optimisons les paramètres pour la compression des images satellite. L'algorithme de la transformée en bandelettes par blocs est d'abord détaillé. Ensuite, nous définissons la meilleure taille des blocs à transformer à partir d'une étude des corrélations entre coefficients d'ondelettes séparés par plus d'un coefficient et à partir de résultats en compression obtenus par extension de la transformée en bandelettes à des blocs de grande taille. La meilleure transformée en bandelettes sur chaque blocs est sélectionnée parmi un dictionnaire de transformées possibles par minimisation d'un critère débit-distorsion. Sous hypothèse haut débit, nous proposons une formule théorique pour le calcul du multiplicateur de Lagrange optimal qui intervient dans ce critère. Nous effectuons ensuite une recherche empirique du multiplicateur de Lagrange optimal à partir de la compression d'images satellite en adaptant la méthode d'optimisation de Shoham et Gersho à la problématique des bandelettes. Les résultats obtenus par cette méthode confirment les résultats théoriques. Nous analysons aussi la précision des estimations intervenant dans le critère débit-distorsion. En particulier, nous améliorons l'estimation du débit de signalisation par un processus adaptatif. Nous observons ensuite que les transformées en bandelettes choisies ne sont pas forcément représentatives des structures géométriques présentes dans les images. Enfin, nous étudions l'efficacité des différentes bases de la transformée en bandelettes et nous donnons des résultats de la compression d'images satellite avec différentes combinaisons de bases de bandelettes.

La transformée en bandelettes choisie sur chaque bloc n'étant finalement pas liée à la géométrie du bloc, nous proposons dans le quatrième chapitre des bases non-directionnelles pour la post-transformée des blocs de coefficients d'ondelettes. Nous commençons par analyser quelles sont les dépendances qui ne sont pas exploitées par les bases de bandelettes directionnelles. A partir des résultats de cette analyse et en vue d'exploiter plus de dépendances, nous proposons de nouvelles bases inspirées par la construction des bases de bandelettes. Cependant, certaines bases de cet ensemble sont similaires et paraissent donc superflues. Nous proposons alors une mesure de dissimilarité entre bases et nous construisons un ensemble de bases maximisant cette mesure. Les résultats en compression obtenus avec cet ensemble ne sont pas suffisamment satisfaisants. Nous construisons alors des bases adaptées aux blocs de coefficients d'ondelettes par Analyse en Composantes Principales. Ces bases permettent de minimiser les corrélations entre coefficients après la post-transformée. De plus, elles compactent l'énergie des blocs de coefficients d'ondelettes sur un petit nombre de coefficients post-transformés dont on connaît la position. Nous définissons alors un nouvel algorithme de compression permettant d'exploiter cette connaissance au cours du codage entropique. Enfin, nous montrons que les post-transformées en cosinus discrets et de Hadamard compactent aussi l'énergie des blocs et ont l'avantage de pouvoir être calculées avec une faible complexité.

Finalement, dans le cinquième chapitre, nous retenons un algorithme de compression par post-transformées avec une faible complexité calculatoire. Pour cela, nous utilisons la

post-transformée de Hadamard et un codeur par plans de bits spécialement conçu pour la compression à bord des satellites. Chaque bloc de coefficients d'ondelettes peut alors être post-transformé ou non. Cependant, le processus de sélection de la post-transformée par minimisation du critère débit-distorsion n'est pas compatible avec un schéma de compression progressive. Nous remplaçons alors ce critère par celui de minimisation de la norme l_1 et nous étudions les conséquences sur les performances en compression. Nous modifions ensuite le codeur par plans de bits pour inclure l'information de signalisation de la post-transformée dans le train binaire emboîté et nous réordonnons les vecteurs dans la base de Hadamard afin que la compaction d'énergie soit exploitée par le codeur. Enfin, nous étendons ce schéma de compression en proposant de post-transformer simultanément des blocs de coefficients sur deux niveaux de la décomposition en ondelettes et nous modifions une fois encore le codeur par plans de bits pour que les dépendances inter-échelles, renforcées par la compaction d'énergie due à la post-transformée, soient correctement exploitées.

Chapitre 1

Images d'observation de la Terre et compression à bord des satellites

1.1	Un peu d'histoire	6
1.2	Les missions d'observation de la Terre au CNES	6
1.2.1	SPOT 1 à SPOT 4	6
1.2.2	SPOT 5	7
1.2.3	PLEIADES	7
1.3	Les images satellite	8
1.3.1	Caractéristiques des images satellite	8
1.3.2	La chaîne de traitement d'image bord-sol	11
1.4	La compression d'images satellite	12
1.4.1	Les étapes de compression	12
1.4.2	Mesurer les performances en compression	14
1.4.3	Compression à bord : Besoins et contraintes	15
1.5	Préparer l'avenir	17

Ce chapitre introductif dresse un panorama des missions d'observation de la Terre par satellite et des méthodes de compression des images à bord. Dans un premier temps, nous replaçons le contexte historique de l'observation de la Terre par satellite. Nous donnons ensuite quelques caractéristiques des grandes missions françaises d'observation de la Terre conduites par le CNES, des techniques de compression employées et des images produites. Nous expliquons les traitements qui sont appliqués à ces images pour arriver à un produit final à forte valeur ajoutée. Enfin, nous introduisons des notions de compression d'images et nous expliquons pour quelles raisons la compression à bord des satellites est nécessaire et doit être améliorée pour anticiper les besoins futurs.

1.1 Un peu d'histoire

L'observation de la Terre par satellite a commencé dès les années soixante dans un contexte de guerre froide et était limitée à des applications de renseignements militaires. Comme les photographies étaient prises sur des films argentiques, il était nécessaire de les redescendre physiquement sur Terre. Ainsi, les films de la série de satellites de surveillance américains Corona étaient renvoyés dans des capsules équipées de parachutes et interceptées par des avions de la CIA munis de grands filets. Du côté soviétique, les satellites de la série Zenit étaient principalement constitués d'une capsule de rentrée atmosphérique de 2,3 m de diamètre et d'environ 2400 kg contenant l'appareil photographique, les films, une balise de localisation, les parachutes et une charge de destruction. La durée des missions était de 8 à 15 jours et la production de ces satellites industrielle : 144 satellites Corona ont été lancés entre 1960 et 1972 et plus de 500 satellites Zenit entre 1962 et 1994 [1, 2]. Le satellite américain KH-11, lancé en 1976, fut le premier satellite d'observation à emporter un capteur CCD [2]. Les données pouvaient alors être transmises par lien de télémesure radio. Le coût de production des images devint alors plus faible. De plus, les images d'observation de la Terre purent être traitées numériquement permettant la compression à bord et la restauration au sol. Le premier satellite civil français d'observation de la Terre SPOT 1 fut lancé dix ans plus tard en 1986.

1.2 Les missions d'observation de la Terre au CNES

Le programme SPOT (Système Probatoire d'Observation de la Terre [3] ou Satellite Pour l'Observation de la Terre) a été décidé par la France en 1978. Il est réalisé par le CNES.

1.2.1 SPOT 1 à SPOT 4

La première génération de satellites du programme SPOT comprend trois satellites : SPOT 1 lancé en 1986, SPOT 2 en 1990 et SPOT 3 en 1993. Quatre détecteurs optiques sont utilisés pour l'acquisition des images : un détecteur dans le vert (bande B1 : 500 nm – 590 nm), dans le rouge (bande B2 : 610 nm – 680 nm), dans le proche infrarouge (bande

B3 : 790 nm – 890 nm) à la résolution de 20 m par pixel et un détecteur panchromatique¹ (510 nm – 730 nm) à la résolution de 10 m par pixel. Les images panchromatiques ont une largeur de 6000 pixels soit une fauchée de 60 km de large. Le débit binaire en sortie du module d'acquisition d'images est de 32 Mb/s.

Afin de permettre la transmission d'un maximum d'images dans un canal de transmission limité, les images sont compressées à bord. La décorrélation des images est effectuée par un codage prédictif DPCM (Differential Pulse Code Modulation) et des codes de longueur fixe. La quantification du résidu est non-uniforme et adaptée à une hypothèse de distribution Laplacienne. Un pixel sur trois, la référence, est codé sur 8 bits. Les deux autres constituent le résidu et sont codés sur 5 bits. Ceci assure la compression des images avec un taux de compression de 1,33. Bien que les performances en compression soient faibles, cet algorithme est très simple à mettre en oeuvre puisque seulement 3 opérations par pixel sont nécessaires. Il était adapté aux performances limitées de l'électronique embarquée au moment de la conception de SPOT 1 [4, 5].

Le satellite SPOT 4 lancé en 1998 est la deuxième génération du programme SPOT. Il intègre un détecteur de plus dans le moyen infrarouge (bande B4 : 1580 nm – 1750 nm) mais ne dispose pas de canal panchromatique. Par contre la bande rouge B2 a une résolution de 10 m par pixel. La technique de compression utilisée est la même que sur la première génération.

1.2.2 SPOT 5

Le satellite d'observation de la Terre SPOT 5, lancé en mai 2002, permet l'acquisition d'images multispectrales avec quatre bandes spectrales. Les bandes B1 à B3 dans le vert, le rouge et le proche infrarouge ont une résolution de 10 m par pixel. La bande B4 dans le moyen infrarouge a une résolution de 20 m par pixel. Les images panchromatiques (480 nm – 710 nm) ont une résolution de 5 m. La fauchée est de 60 km. La résolution des images peut-être améliorée à 2,5 m après traitement grâce à un mode super-résolution et à deux barrettes (lignes de détecteurs) de 12000 pixels décalées d'un demi pixel : acquisition en mode quinconce [6, 7]. Le débit binaire en sortie du module d'acquisition d'images est de 128 Mb/s.

Le compresseur utilise la DCT (Discrete Cosine Transform) et un code de longueur variable. Grâce à un algorithme de régulation de débit asservissant la quantification, le taux de compression est maintenu constant et égal à 2,8 [4, 8].

1.2.3 PLEIADES

Le lancement du premier satellite PLEIADES est prévu en 2010. La fauchée sera réduite à 20 km mais la résolution du mode panchromatique (480 nm – 820 nm) sera de 70 cm. Dans ce mode, les images auront une largeur de 30000 pixels. Une bande B0 bleu (450 nm – 530 nm) dans le spectre de la lumière visible a été ajoutée dans le mode multispectral ce qui permet la production d'images en couleurs naturelles à partir des trois

¹Ce terme est défini dans la section 1.3.1

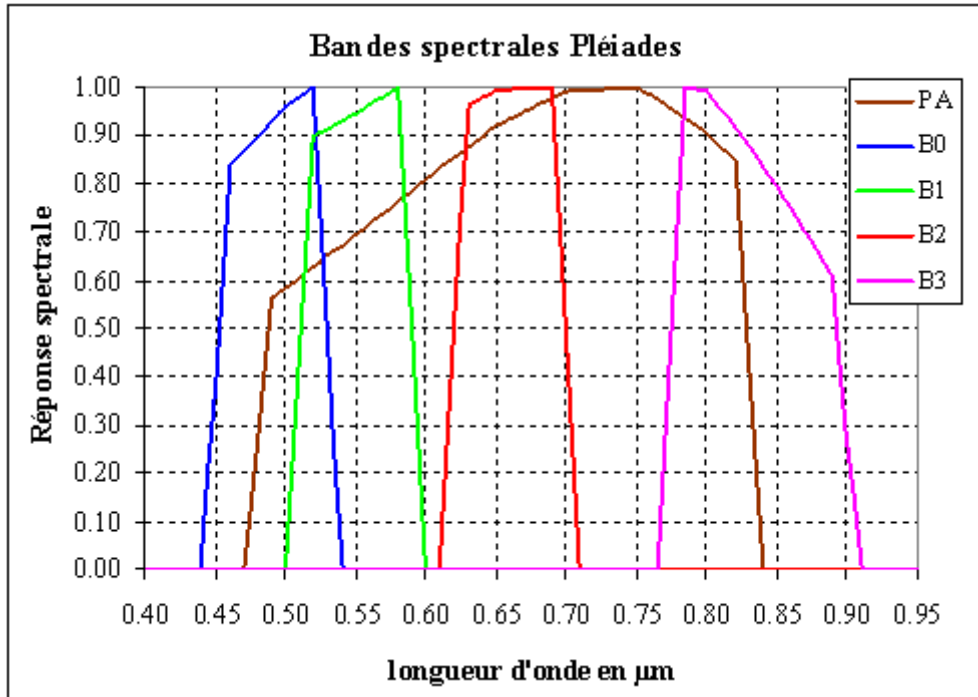


FIG. 1.1 – Bandes B0 à B3 et bande panchromatique de PLEIADES (figure extraite de [9]).

bandes B0, B1 et B2. Il n'y aura plus de bande B4 dans le moyen infrarouge. Les quatre bandes spectrales notées B0 à B3 dans le bleu, le vert, le rouge et le proche infrarouge sont représentées sur la figure 1.1. Elles auront une résolution de 2,8 m. Le débit binaire en sortie du module d'acquisition d'images est de 4,5 Gb/s.

Le compresseur MRCPB (Multi-Résolution par Codage de Plans Binaires) utilisera une transformée en ondelettes par « lifting scheme » (schéma de lissage) [4, 10, 11] qui présente une complexité calculatoire plus faible que le schéma de convolution et qui ne nécessite pas de mémoire auxiliaire. Ce schéma *lifting* permet donc d'économiser deux ressources critiques à bord. De plus, il permet une compression de l'image au fil de l'eau. Le codage entropique sera effectué par codage RLR (Run-Length Rice) des plans de bits [12, 13]. La régulation de débit sera donc implicitement effectuée par troncature du train binaire emboîté généré par le codeur. Le compresseur permettra d'augmenter le taux de compression à 5 pour une qualité d'image identique à celle des images SPOT 5.

1.3 Les images satellite

1.3.1 Caractéristiques des images satellite

Des images panchromatiques Une image numérique est une matrice de pixels. A chaque pixel de cette matrice est associée une valeur correspondant à une intensité lumineuse. Une palette de couleurs est utilisée pour représenter cette intensité lumineuse. Dans le cas d'images monochromatiques, ces intensités sont en général représentées par



FIG. 1.2 – Image de Toulouse de taille 945×450 pixels acquise par le mode panchromatique de SPOT 5 à la résolution de 5 m par pixel.

une gamme de gris allant du noir pour la plus faible intensité au blanc pour la plus forte. En imagerie satellite, on parle d'image panchromatique lorsque l'acquisition est effectuée sur une bande spectrale large. Par exemple, dans le mode panchromatique de SPOT 5, l'acquisition est effectuée sur une bande spectrale allant d'une longueur d'onde électromagnétique de 480 nm à 710 nm. On rappelle que le spectre électromagnétique de la lumière visible s'étend de 380 nm pour le violet à 780 nm pour le rouge. Le mode panchromatique de SPOT 5 acquiert donc une grande partie de la lumière visible. La figure 1.2 est une image panchromatique de SPOT 5 à la résolution de 5 m par pixel. Des images panchromatiques issues d'autres capteurs sont données en annexe A.

Des images couleurs, multispectrales et hyperspectrales Pour une image en couleurs, ce n'est pas une valeur qui est associée à chaque pixel mais un vecteur de trois valeurs. Ce vecteur est couramment représenté dans la base RVB (Rouge, Vert, Bleu) mais il existe aussi d'autres bases très utilisées telles que les bases YCbCr (Luminance, Chrominance bleue, Chrominance rouge) ou HSV (Hue/Teinte, Saturation, Valeur). On parle d'images multispectrales lorsque le vecteur de valeurs associé à un pixel est au moins de dimension deux. Par exemple, les satellites d'observation de la Terre PLEIADES permettront l'acquisition d'images multispectrales avec les 4 bandes spectrales B0 à B3 dans le bleu, le vert, le rouge et le proche infrarouge. L'image en couleurs représentée sur la figure 1.3 a été créée à partir des bandes B0, B1 et B2. La multiplication du nombre de bandes spectrales permet de différencier la nature des objets observés. Par exemple, la bande verte et la bande proche infrarouge sont sensibles à la chlorophylle alors que la bande moyen infrarouge est sensible à l'amidon, à la lignine, un composé du bois, et permet de séparer les nuages de la neige et de la glace [14]. On parle généralement d'images hyperspectrales lorsque le nombre de bandes spectrales est supérieur à la centaine ou lorsque



FIG. 1.3 – Image couleurs simulée PLEIADES à la résolution de 70 cm par pixel de taille 945×450 (image extraite de [9] – © CNES 2006).

l'ensemble des bandes spectrales acquises est continu. Hyperion est le premier capteur hyperspectral lancé à bord du satellite EO-1 (Earth Observing-1) de la NASA en 2000. Il est capable d'acquérir 220 bandes spectrales entre 400 nm et 2500 nm [15, 16]. Une image hyperspectrale est représentée sous la forme d'un cube hyperspectral sur la figure 1.4. La troisième dimension correspond à l'axe du spectre électromagnétique.

La résolution numérique La résolution numérique est le nombre de bits nécessaire au codage binaire d'un vecteur de valeurs associé à un pixel. En général, la résolution des images monochromatiques est de 8 bits par pixel. Les valeurs d'intensité des pixels sont comprises entre 0 pour le noir et 255 pour le blanc. Les images couleurs ont en général une résolution de 24 bits par pixel : 8 bits sur chacune des trois composantes de couleurs. Les images panchromatiques PLEIADES, elles, sont codées sur 12 bits et ont donc des valeurs comprises entre 0 et 4095. *C'est sur des images panchromatiques 12 bits à haute résolution spatiale que nous travaillons dans cette thèse.*

Les différents modes d'acquisition On distingue aujourd'hui trois modes d'acquisition d'images par satellite : l'acquisition par matrices CCD, l'acquisition par balayage et l'acquisition de type scanner. Les matrices de CCD sont plus généralement utilisées pour des missions d'observation astronomiques telles que celle du satellite américain Clementine, lancé en 1994, qui a cartographié l'ensemble de la surface de la Lune et dont le module de compression d'images a été fourni par le CNES ou le télescope spatial Corot, lancé fin 2006, dont la maîtrise d'oeuvre a été conduite par le CNES. Les barrettes CCD sont plus utilisées pour l'acquisition des images d'observation de la Terre par balayage. L'image est construite ligne par ligne au cours du défilement du satellite sur son orbite. *Ce*

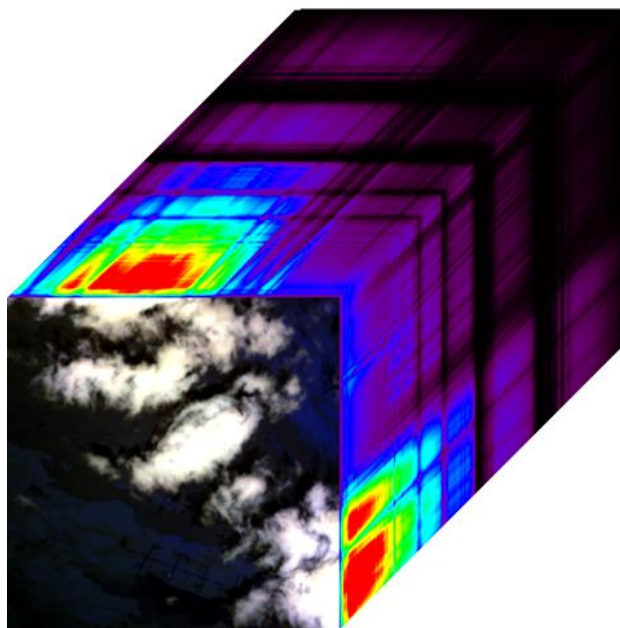


FIG. 1.4 – Représentation d’une image hyperspectrale sous forme d’un cube.

principe, appelé push-broom, est utilisé sur les satellites SPOT et PLEIADES. L’image formée par balayage est de largeur fixe et de longueur virtuellement infinie. Sur les satellites Landsat de la NASA, l’acquisition des images est effectuée par un scanner [17]. Pour ce type d’acquisition, un seul détecteur est utilisé pour la création de l’image. Les lignes de l’image sont produites par rotation d’un miroir qui reflète l’image du sol dans la direction perpendiculaire au sens de déplacement du satellite. Les colonnes de l’image sont produites au cours du déplacement du satellite comme pour l’acquisition par balayage.

1.3.2 La chaîne de traitement d’image bord-sol

La chaîne de traitement des images satellite est représentée sur la figure 1.5. Un satellite acquiert les images par balayage en défilant autour de la Terre. Les informations géométriques sur le satellite telles que son orbite, son altitude et sa direction de visée sont enregistrées au cours de l’acquisition. Elles permettront de localiser précisément la zone imagée au sol. De même les informations radiométriques seront transmises au sol afin d’effectuer une première série de corrections sur l’image brute. Les images sont compressées à bord puis, soit directement transmises au sol lorsque le satellite est en visibilité directe avec une station de réception, comme sur SPOT 5, soit stockées à bord. Le débit vidéo en sortie de l’instrument optique est très supérieur au débit de la liaison télémesure. Par exemple, sur PLEIADES, le débit vidéo est de 4,5 Gb/s alors que le débit de liaison télémesure nominal n’est que de 465 Mb/s lorsque le satellite est en visibilité avec une station de réception au sol ce qui n’est pas toujours le cas. Il y a un facteur 10 entre ces débits. On comprend alors la nécessité de la compression à bord et de la mémoire de masse. Sur PLEIADES, le taux de compression standard est de 5 et peut atteindre 7 et la capacité

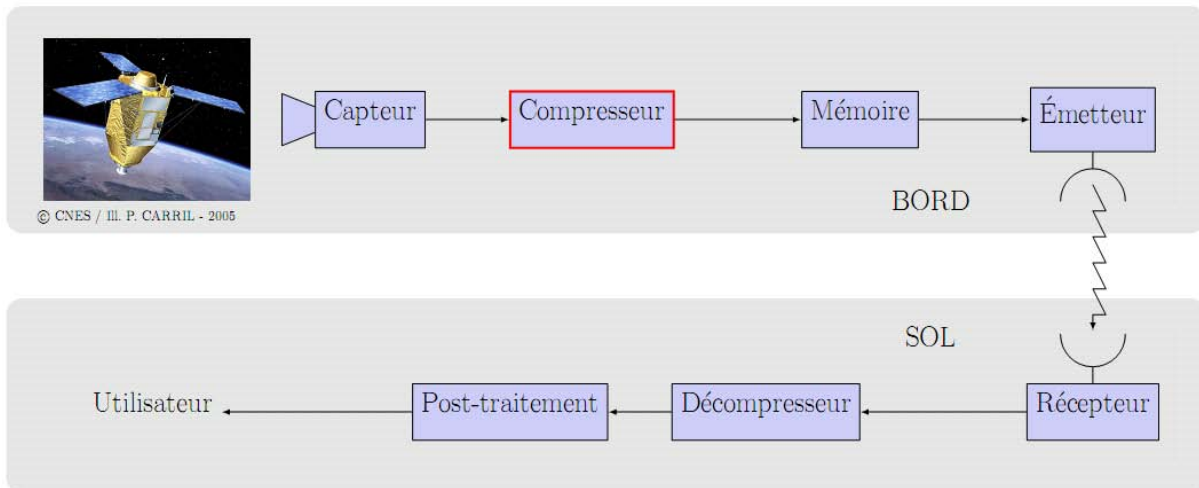


FIG. 1.5 – Chaîne de traitement des images bord/sol.

de stockage à bord est de 600 Gb. *Dans cette thèse, nous travaillons à l'amélioration de la compression à bord par l'utilisation de nouvelles techniques de compression.*

Au sol, plusieurs séries de traitements sont effectuées. La première consiste à décoder et décompresser l'image. On obtient alors une image brute dite de *niveau 0*. Ensuite des corrections radiométriques sont appliquées à l'image pour obtenir une image de *niveau 1* utilisable pour des applications de défense. L'image au niveau dit *capteur parfait* est obtenue après traitements géométriques. Cette image peut déjà être commercialisée. Enfin, au niveau *ortho-image*, l'image est projetée au sol avec des corrections sur le relief et est localisée dans un système de coordonnées géographiques ou cartographiques. Ce produit final est destiné aux systèmes d'informations géographiques [18, 19].

1.4 La compression d'images satellite

1.4.1 Les étapes de compression

Le schéma d'un système de compression classique est représenté sur la figure 1.6. La première étape est la décorrélation. Elle permet d'exploiter les redondances statistiques de l'image. Ensuite vient l'étape de quantification. C'est dans cette étape que la perte d'information a lieu. Enfin, le codage entropique des données quantifiées constitue la dernière étape. Elle permet d'organiser le train binaire en compressant les données sans altérer l'information.

La décorrélation La décorrélation a pour objectif de réduire les redondances statistiques de l'image et ainsi réduire les entropies d'ordre faible. Par la suite, cela permet d'utiliser une quantification et un codage scalaire, c'est-à-dire le traitement des coefficients un par un pour la compression [20]. Cette étape est réversible sans pertes ou quasiment sans pertes. Elle fait intervenir une transformée. Par exemple la DCT est utilisée dans le très populaire standard de compression d'images JPEG mais aussi sur SPOT 5 ; la

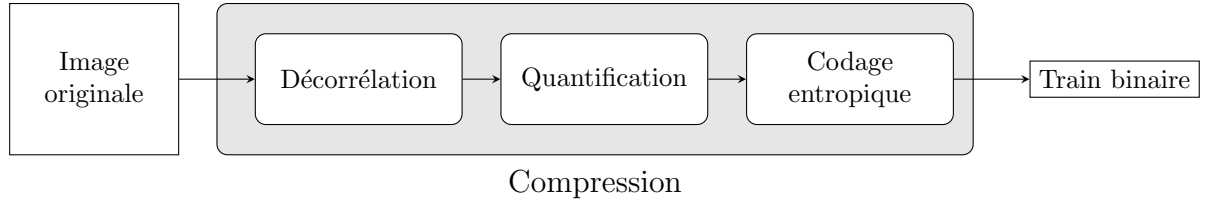
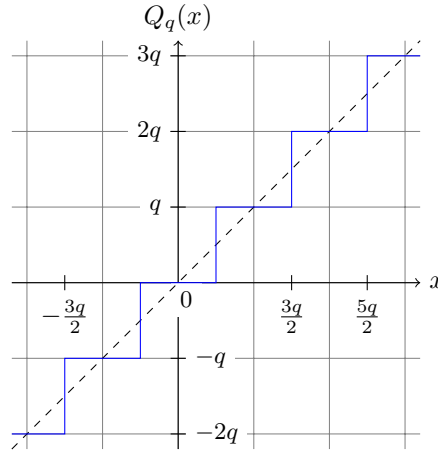


FIG. 1.6 – Schéma général d'un système de compression.

FIG. 1.7 – Quantification scalaire uniforme de pas q .

transformée en ondelettes est à la base du standard JPEG2000 [21] et sera implantée à bord de PLEIADES [5] ; la transformée optimale de Karhunen-Loeve (KLT) peut aussi être utilisée pour la décorrélacion d'images multi- ou hyperspectrales suivant l'axe spectral [22, 23]. En principe, après la transformée, une grande partie de l'énergie de l'image est compactée sur un faible nombre de coefficients. *Dans cette thèse, nous étudions les post-transformées dans le domaine ondelettes afin d'améliorer la décorrélacion spatiale.*

La quantification C'est la seule étape non réversible du schéma de compression. La quantification peut être vectorielle ou scalaire. La quantification vectorielle permet la quantification simultanée de plusieurs coefficients. Elle est coûteuse en termes de complexité calculatoire et n'est utile que si les données ne sont pas décorrélées. La quantification scalaire est moins complexe. On distingue la quantification scalaire uniforme et la quantification scalaire non uniforme. Dans le cas de la quantification scalaire uniforme représentée sur la figure 1.7, les intervalles de quantification sont tous de la même taille. Cette quantification est optimale au sens de la distorsion pour une distribution uniforme des coefficients. Il est possible de construire une quantification scalaire non-uniforme optimale au sens de la distorsion pour une statistique marginale donnée [24]. Par exemple, pour la quantification des coefficients de la transformée en ondelettes d'une image, il est montré dans [25] que le quantificateur scalaire à zone morte double est un très bon choix au sens de la distorsion engendrée.

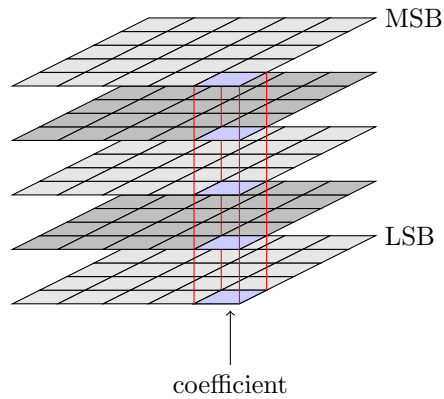


FIG. 1.8 – Représentation de plans de bits pour des coefficients codés sur 5 bits.

Le codage entropique Le codage entropique permet la compression des coefficients quantifiés en générant un train binaire de longueur proche de la limite basse théorique de Shannon. Plusieurs types de codage entropique existent. Le codage par Run-Length est utilisé dans le fax mais aussi dans le standard JPEG [26]. Le codage par dictionnaire comme LZW (Lempel-Ziv-Welch) est utilisé dans le format d'image GIF (Graphics Interchange Format). Le codage par statistique comme le codage de Huffman est aussi utilisé dans le format JPEG. Enfin, un codage arithmétique est effectué dans le standard JPEG2000 par le codeur MQ [21, 27].

Plutôt que de coder des coefficients quantifiés à valeurs dans $\mathbb{Z}$, il est possible de coder les coefficients par plans de bits en représentant l'amplitude des coefficients quantifiés sous forme binaire comme illustré sur la figure 1.8. Ainsi les coefficients sont codés progressivement du plan de bits le plus significatif (MSB) au plan de bits le moins significatif (LSB). Cette technique de codage est souvent utilisée avec la transformée en ondelettes. En particulier, elle est utilisée dans EBCOT (Embedded Block Coding with Optimal Truncation points) [27, 28] le module de codage de JPEG2000, dans le module de compression MR-CNES (Multi-Résolution CNES) [5], dans MRCPB le futur module de compression de PLEIADES, et dans le standard de compression d'images du CCSDS (Consultative Committee for Space Data Systems) [29, 30]. L'avantage des codeurs par plans binaires est qu'ils permettent une compression et une décompression progressives. Le train binaire produit peut être tronqué à n'importe quel endroit et l'image peut toujours être intégralement reconstruite avec une qualité plus ou moins élevée selon la position du point de troncature.

1.4.2 Mesurer les performances en compression

Taux de compression et débit binaire Le taux de compression est le rapport entre la taille de l'image originale et la taille de l'image compressée en bits. Par exemple, une image à la résolution numérique de 12 bits et de taille 1024×1024 a un volume de 1,5 Mo. Avec un taux de compression de 5, le volume de cette image est réduit à 307,4 ko. Ce taux de compression peut être relié au débit binaire exprimé en bits par pixel (bpp) : c'est le volume des données compressées divisé par le nombre de pixels dans l'image ou encore la

résolution numérique de l'image divisée par le taux de compression. Dans le cas d'images 12 bits, avec un taux de compression égal à 5, le débit binaire est de 2,4 bpp. *Ce sont des débits proches de 2,4 bpp que nous visons dans cette thèse.*

Compromis débit-distorsion et critères de qualité Le taux de compression n'est pas le seul critère de performance d'un système de compression. Dans le cas d'une compression avec pertes, la qualité de l'image reconstruite doit aussi être prise en compte. Il y a donc un compromis à trouver entre le taux de compression et la qualité. On parle de compromis débit-distorsion. La distorsion est une mesure de l'erreur commise entre l'image originale et l'image reconstruite. Pour les mesures de distorsion, on utilisera l'erreur quadratique moyenne EQM entre l'image originale I_0 et l'image compressée I_C de taille $M \times N$:

$$EQM(I_0, I_C) = \frac{1}{M N} \sum_{i=1}^M \sum_{j=1}^N (I_0(i, j) - I_C(i, j))^2.$$

Pour les mesures de qualité, on utilisera le $PSNR$ (Peak Signal to Noise Ratio) en décibels défini à partir de l' EQM par :

$$PSNR(I_0, I_C) = 20 \log_{10} \left(\frac{2^r - 1}{\sqrt{EQM(I_0, I_C)}} \right)$$

où r est la résolution numérique de l'image. Les images sur lesquelles nous travaillons (cf. annexe A) sont des images 12 bits. Nous avons donc $r = 12$. Bien que ce critère ne soit pas forcément représentatif des évaluations subjectives visuelles [31], nous ne nous attarderons pas sur les critères de qualité subjectifs parce que les images d'observation de la Terre sont soumises à des traitements post-compression qui permettent d'en améliorer la qualité visuelle. Dans ce travail, nous préférons donc un critère objectif avant ces traitements. D'autres critères de qualité objectifs existent comme l'erreur maximale ou l'erreur moyenne absolue mais nous utiliserons le PSNR car c'est une référence pour la mesure des performances en compression.

Complexité Un troisième critère de performance d'un système de compression est la complexité. La complexité calculatoire peut être mesurée par le temps d'exécution du processus de compression ou en nombre d'opérations par pixel : c'est le nombre moyen d'opérations qui sont nécessaires à la compression de l'image. La complexité d'implantation est aussi importante dans le cadre de la compression à bord des satellites. En effet, pour pouvoir être implanté sur des circuits électroniques sur lesquels le nombre de portes logiques est limité, il faut que cette complexité soit la plus faible possible. Enfin, les besoins en mémoire nécessaires au cours de la compression sont aussi à prendre en compte puisque la mémoire est une ressource critique à bord des satellites.

1.4.3 Compression à bord : Besoins et contraintes

Les satellites d'observation de la Terre sont utilisés pour diverses applications : agriculture, urbanisme, cartographie, défense. Ces applications ont besoin d'une excellente

qualité d'image. Le bruit de compression doit être du même ordre que le bruit du capteur et le PSNR doit donc être élevé. L'augmentation de la résolution des images entraîne une augmentation des débits en entrée des compresseurs. Les besoins en compression sont donc de plus en plus importants avant la transmission des données au sol.

Il existe trois contraintes majeures en compression embarquée. La première est liée au mode d'acquisition des images d'observation de la Terre. Les capteurs balayant la surface de la planète produisent une image de largeur fixe mais de longueur virtuellement infinie. Cette image doit donc être compressée et transmise au fur et à mesure de son acquisition. La technique employée consiste à traiter l'image au fil de l'eau, c'est-à-dire, bloc de lignes par bloc de lignes. Sur PLEIADES la compression au fil de l'eau sera effectuée par blocs de 16 lignes. Il faut donc que l'algorithme de compression soit compatible avec le schéma de *lifting* nécessaire pour effectuer la transformée en ondelettes des images acquises dans le mode *push-broom*.

La seconde contrainte lors de la compression à bord est liée au débit de télémesure pour la transmission des données vers le sol et au débit maximum en entrée de la mémoire de masse. Ces débits sont fixes et limités. En conséquence, le débit en sortie du compresseur doit lui aussi être limité. Cependant, pour une qualité d'image donnée, le taux de compression peut dépendre de la statistique des images. Il peut être faible sur des images « complexes » et élevé sur des zones uniformes. Pour cette raison un algorithme de régulation de débit est utilisé sur SPOT 5 [8]. En outre, un codage progressif sera utilisé sur PLEIADES.

La troisième contrainte est liée aux circuits électroniques. En effet, pour assurer un traitement en temps réel, les algorithmes sont implantés de façon matérielle. Les capacités de calcul dépendent donc du nombre de portes disponibles sur les circuits. C'est pourquoi, l'algorithme de compression doit être peu complexe. L'écart technologique, en nombre de portes par centimètre carré, entre les ASIC (Application Specific Integrated Circuits) commerciaux et les ASIC qualifiés spatial est de 5 ans. Cet écart est dû à la tolérance aux radiations requise par les ASIC qualifiés spatial. Néanmoins, grâce à l'évolution rapide des technologies, la contrainte de complexité est peu à peu relâchée. En 2004, le nombre de portes par centimètre carré de circuit était de 2 millions, ce qui permettait une complexité calculatoire de plus de 200 opérations par pixels.

Un algorithme améliorant les performances en compression ne pourra donc être accepté à bord des satellites qu'à condition qu'il respecte les contraintes à bord et que sa complexité et les besoins en mémoire requis restent raisonnables. Dans cette thèse, nous travaillons sur des images d'observation de la Terre 12 bits avec une résolution inférieure à 70 cm. Le taux de compression visé est d'environ 6. Nous donnons des résultats en compression pour une gamme de débits entre 0,5 bpp et 3,0 bpp. De plus, nous veillons à ce que nos algorithmes permettent une compression compatible avec le mode d'acquisition *push-broom* et que leur complexité reste faible.

1.5 Préparer l'avenir

Les diverses applications liées à l'observation de la Terre, telles que la défense, la sécurité civile, la prévention des catastrophes naturelles et d'autres, ont besoin de données toujours plus précises. De nouvelles applications nécessitent des images multi- ou hyperspectrales [22] ou encore deux prises de vues (stéréo-vision) [6, 32]. Dans tous les cas, la résolution des images augmente. Dans les futurs satellites d'observation de la Terre, la résolution des images sera supérieure à celle de PLEIADES. C'est pourquoi les ensembles d'images utilisés dans cette thèse sont formés des images à très haute résolution représentées en annexe A. Les images simulées PLEIADES ont une résolution de 70 cm par pixel et les images issues du capteur aéroporté PELICAN ont une résolution de 25 cm par pixel. Les débits d'acquisition seront aussi plus importants. De plus, pour un taux de compression de 6, le PSNR devra être supérieur à 50 dB. L'utilisation de la transformée en ondelettes seule ne suffira plus pour compresser les données au débit imposé par le canal de transmission. Il est nécessaire d'anticiper ces nouveaux besoins en compression en étudiant des méthodes novatrices permettant d'augmenter les performances [33]. C'est dans ce cadre que s'inscrit cette thèse. Le but de ce travail n'est pas de développer un algorithme qui puisse être implanté à bord aujourd'hui. Il est plus prospectif. Il consiste à explorer une nouvelle méthode de compression d'images à haute résolution par l'utilisation de nouvelles transformées. Notre choix s'est porté sur les post-transformées effectuées après la transformée en ondelettes.

Chapitre 2

Dépendances et compression

2.1	La transformée en ondelettes	20
2.1.1	Pour la compression d'images satellite	20
2.1.2	Dépendances entre coefficients d'ondelettes	21
2.2	Exploitation des dépendances dans les codeurs par plans de bits	26
2.2.1	BPE : le codeur du standard CCSDS	26
2.2.2	EBCOT : le codeur de la norme JPEG2000	29
2.2.3	Les codeurs morphologiques	32
2.3	Transformées dérivées des ondelettes	34
2.3.1	Transformées redondantes	35
2.3.2	Les bandelettes	36
2.4	Conclusion	38

La transformée en ondelettes est utilisée par les systèmes de compression d'images satellites les plus récents. Afin de pouvoir augmenter les performances en compression, une analyse des faiblesses de la transformée en ondelettes doit être effectuée. Dans ce chapitre, nous effectuons une analyse des dépendances entre coefficients d'ondelettes dans la transformée d'images satellite. Cette analyse a été publiée dans [34]. Ensuite, nous décrivons différentes techniques employées par les codeurs pour exploiter ces dépendances parmi lesquelles on trouve le codage contextuel utilisé dans JPEG2000. Dans [35], nous avons analysé l'efficacité des contextes définis dans le codeur de JPEG2000. Enfin, un autre moyen pour réduire les dépendances entre coefficients transformés est d'utiliser une transformée plus efficace que la transformée en ondelettes. Nous décrivons brièvement quelques transformées dérivées des ondelettes, en particulier les transformées en bandelettes à partir desquelles les post-transformées ont été construites.

2.1 La transformée en ondelettes

2.1.1 Pour la compression d'images satellite

La transformée en ondelettes retenue pour la compression d'images satellite dans le standard du CCSDS [29, 36] est la transformée en ondelettes discrète 9/7 biorthogonale. C'est celle qui permet d'atteindre les meilleures performances en compression [37]. C'est aussi la transformée en ondelettes utilisée dans la norme JPEG2000 [21]. Le nombre de niveaux de décomposition en ondelettes a été fixé à 3. Ce nombre de niveaux de décomposition est un compromis entre performances en compression et complexité calculatoire [36]. De plus, un faible niveau de décomposition réduit le nombre de pixels affectés par la perte de données compressées. La décomposition en ondelettes sur trois niveaux d'un morceau d'une image PELICAN est représentée sur la figure 2.1.

L'acquisition par balayage des images satellite pose certains problèmes pour effectuer la transformée en ondelettes. En effet, comme la longueur des images est très grande et virtuellement infinie, il est impératif de commencer la compression avant que la totalité d'une image ait été acquise pour ne pas saturer la mémoire disponible à bord. Pour cela, la méthode proposée dans la partie 1 du standard JPEG2000 [21] est de découper l'image en « tuiles » et de compresser chacune des tuiles indépendamment. Cette méthode permet de ne traiter que les morceaux de l'image qui ont été acquis mais peut faire apparaître des artefacts de compression à la jonction des tuiles quand l'image est reconstruite à bas débit. Une autre méthode existe. Elle consiste à traiter l'image en continu au cours de son acquisition et à effectuer la transformée en ondelettes au fil de l'eau. Les problèmes liés à cette méthode pour la transformée en ondelettes, ses avantages et les besoins en mémoire requis ont été traités dans [38]. La transformée au fil de l'eau est effectuée à partir du schéma de lissage ou « lifting scheme » [10]. Ce schéma est celui utilisé dans le standard de compression d'image du CCSDS [29]. Il est aussi présent dans la partie 2 de JPEG2000 [39].

Dans le reste de cette thèse, la transformée en ondelettes utilisée sera celle du standard

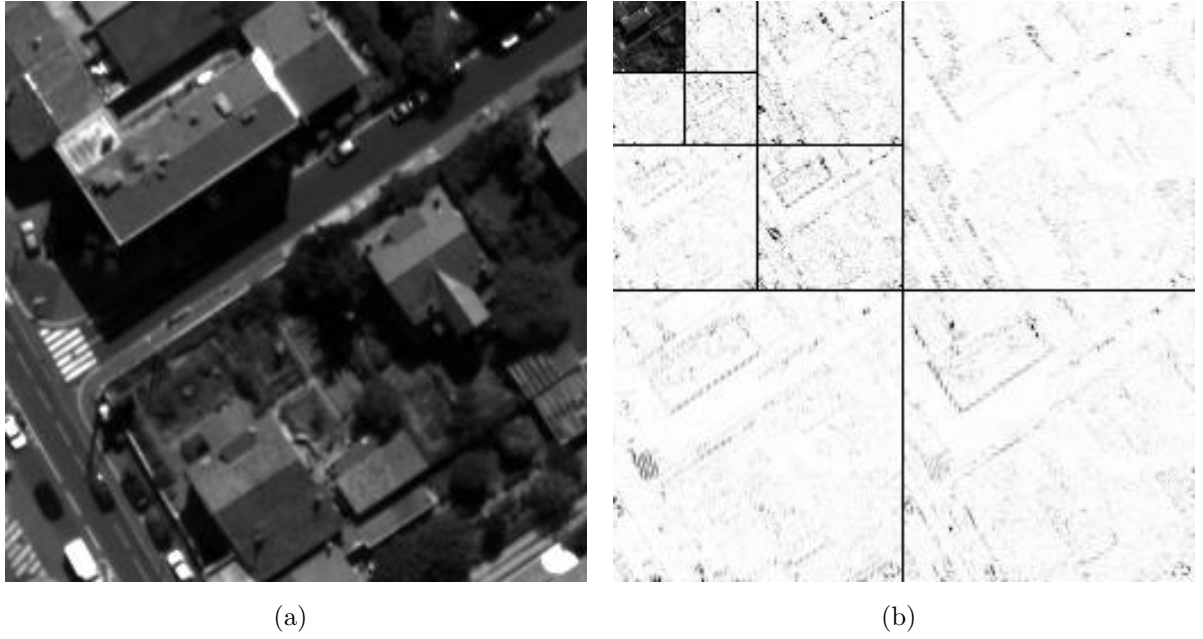


FIG. 2.1 – Image PELICAN (a) et la transformée en ondelettes de cette image (b).

du CCSDS [29] pour la compression d’images avec pertes c’est-à-dire la transformée en ondelettes discrète 9/7 en flottants sur trois niveaux de décomposition.

2.1.2 Dépendances entre coefficients d’ondelettes

Bien que la transformée en ondelettes soit utilisée pour décorréler les images, il est encore possible d’observer des structures dans les images transformées. La figure 2.1(b), représente la transformée en ondelettes de l’image de la figure 2.1(a). Sur cette figure, les coefficients d’ondelettes sont représentés en valeur absolue et la palette de couleurs a été ajustée de sorte que toute la gamme de gris soit utilisée dans chaque sous-bande. Ceci permet de faire apparaître les structures de l’image.

Des dépendances statistiques ont aussi été observées entre coefficients d’ondelettes dans un voisinage très local. Liu et Moulin [40] ont étudié les dépendances inter- et intra-échelles en mesurant l’information mutuelle entre deux ou plusieurs coefficients d’ondelettes. Simoncelli a modélisé les probabilités d’apparition conjointes entre coefficients d’ondelettes voisins [41] et développé un codeur basé sur cette modélisation avec Bucci-grossi [42].

Identification quantitative des dépendances entre coefficients d’ondelettes voisins

Dans [34], nous nous intéressons aussi aux dépendances statistiques entre coefficients d’ondelettes dans un voisinage très local, mais plutôt que d’étudier l’information mutuelle ou les probabilités d’apparition conjointe, nous étudions les coefficients de corrélation entre coefficients d’ondelettes non quantifiés. Une mesure d’information mutuelle est aussi ef-

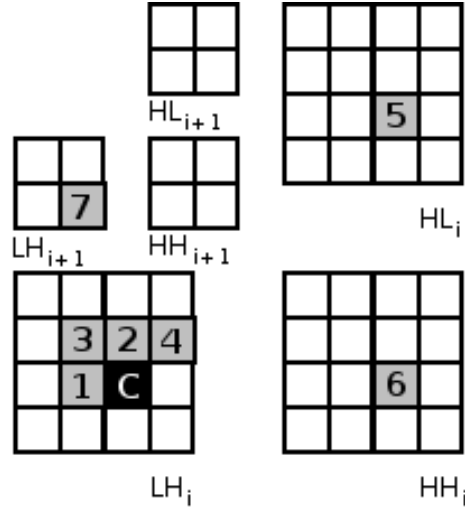


FIG. 2.2 – Définition d'un voisinage dans le domaine ondelettes.

fectuée afin de valider les résultats obtenus. Pour un coefficient d'ondelettes C considéré, on définit un ensemble de sept voisins dans la transformée en ondelettes. Ce voisinage est représenté sur la figure 2.2 et diffère légèrement de celui utilisé dans [42]. Les voisins numérotés de 1 à 4 sont situés dans la même sous-bande que le coefficient C . Les voisins numérotés 5 et 6 sont les cousins du coefficient C dans les sous-bandes à la même résolution, et le voisin numéroté 7 est son parent dans la sous-bande à la résolution inférieure. Ce dernier n'est pas défini pour un coefficient dans les sous-bandes à la résolution 3.

La figure 2.3 présente les coefficients de corrélation entre le coefficient C et chacun de ses sept voisins sur les transformées en ondelettes des images de l'ensemble d'apprentissage représenté en annexe A. Ces coefficients de corrélation ont été obtenus en considérant successivement l'ensemble des coefficients de chaque sous-bande comme les réalisations d'une variable aléatoire. Nous avons donc fait abusivement l'hypothèse d'ergodicité et de stationnarité des coefficients d'ondelettes comme dans [41]. On constate qu'il existe des coefficients de corrélation de plus de 0,4 en valeur absolue entre coefficients directement voisins (numérotés 1 et 2). Les coefficients de corrélation inter- sous-bandes et inter-échelles sont eux toujours inférieurs à 0,05 en valeur absolue. De plus, la corrélation entre les coefficients d'ondelettes diminue lorsque l'échelle de la transformée augmente. Enfin, le signe des coefficients de corrélation est à mettre en relation avec les filtres d'ondelettes appliqués à l'image comme l'a observé Azimifar dans sa thèse [43]. Les corrélations les plus fortes sont celles entre coefficients voisins dans la direction du filtrage en ondelettes passe-haut, c'est-à-dire dans la direction horizontale dans les sous-bandes HL et dans la direction verticale dans les sous-bandes LH.

Sur les transformées en ondelettes d'images de bruit gaussien et d'images de bruit impulsionnel, les coefficients de corrélation observés sont inférieurs à 0,01 en valeur absolue, ce qui suggère que ces corrélations entre coefficients d'ondelettes voisins sont liées à la structure des images. Sur des images synthétiques contenant des contours mais sans textures, on observe le même type de corrélations que sur les images naturelles mais dans

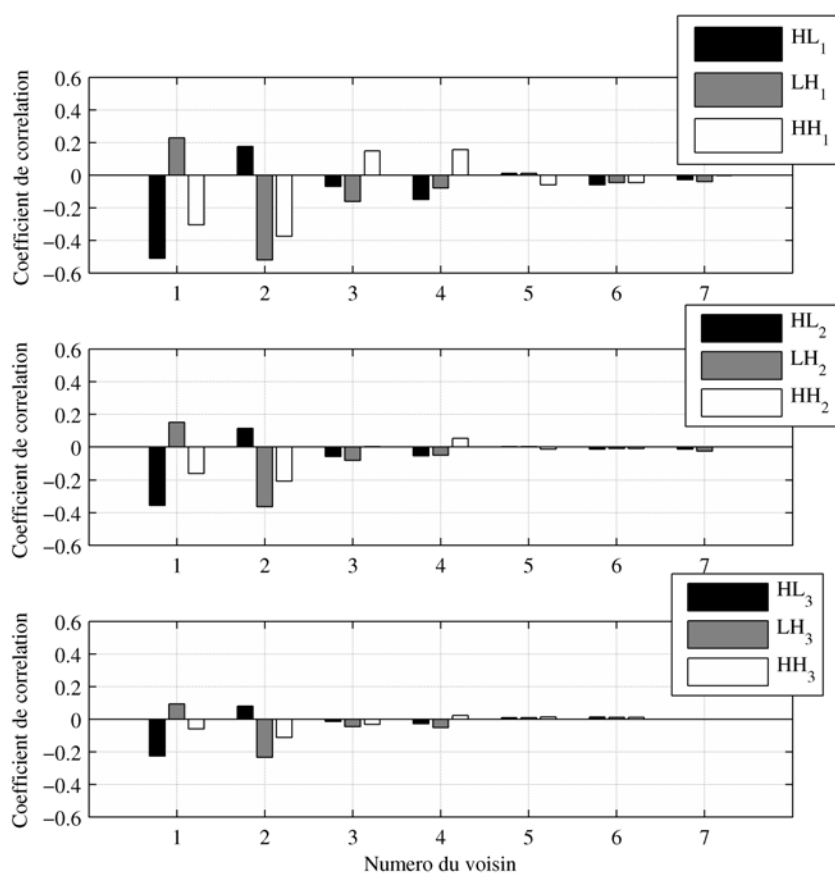


FIG. 2.3 – Coefficients de corrélation entre coefficients d'ondelettes voisins dans les diverses sous-bandes.

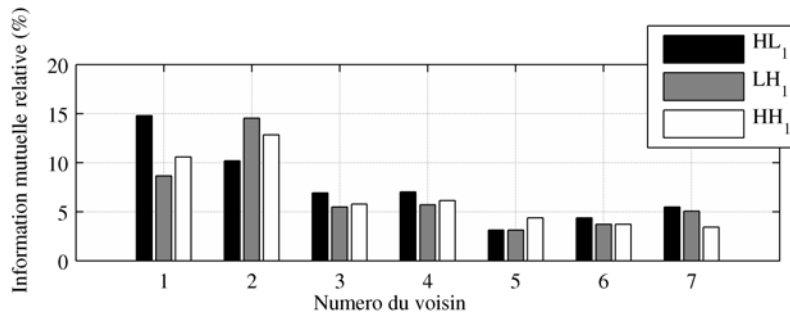


FIG. 2.4 – Information mutuelle relative entre coefficients d’ondelettes voisins dans les sous-bandes à la résolution la plus fine.

ce cas, les corrélations augmentent avec l’échelle de la transformée. Sur des images de texture sans aucun contour telles que des images satellite de mer ou de champ, on a pu observer, dans les sous-bandes à la résolution 1, des coefficients de corrélation atteignant 0,6 pour les voisins numérotés 1 et 2. Cependant, dans ce cas, les corrélations diminuent très rapidement avec l’échelle de la transformée. Elles sont toutes inférieures à 0,2 à la résolution 2. On peut donc faire l’hypothèse que ces corrélations sont issues des très petites structures qui constituent la texture et qui n’existent que dans les sous-bandes haute fréquences. Seules les corrélations issues des grandes structures, comme les contours, peuvent être observées à faible résolution.

La figure 2.4 présente l’information mutuelle relative [44] entre coefficients d’ondelettes voisins après quantification de pas $q = 32$. Une mesure relative est préférée à une mesure absolue pour comparer des sous-bandes d’entropies différentes. L’information mutuelle relative $I_r(X, Y)$ peut être interprétée comme la *proportion* d’information qu’il est possible d’économiser pour le codage de la variable aléatoire X si on connaît parfaitement la variable aléatoire Y . Elle s’exprime en fonction des entropies d’ordre 0 de la manière suivante :

$$I_r(X, Y) = \frac{2(H_0(X) - H_0(X|Y))}{H_0(X) + H_0(Y)} \quad (2.1)$$

La figure 2.4 confirme qu’il est possible d’exploiter les relations entre coefficients d’ondelettes dans un voisinage très local. En effet, dans les sous-bandes à la résolution la plus fine, l’information mutuelle relative approche les 15% pour les voisins 1 et 2. Elle est autour de 5% pour le voisinage inter- sous-bandes (voisins 5, 6 et 7).

Cette étude a montré que la plupart des dépendances statistiques très locales liées à la structure des images naturelles sont observables à l’intérieur d’une même sous-bande.

Dépendances liées à la géométrie

Dans la section précédente, nous avons mis en évidence l’existence de dépendances entre coefficients d’ondelettes voisins. Ces dépendances sont observables dans la transformée en ondelettes de n’importe quelle image naturelle et ont un caractère structurel lié aux filtres d’ondelettes. Dans cette section, nous revenons sur les structures visibles dans la transformée en ondelettes des images (cf. figure 2.1(b)). Elles sont dues à l’inefficacité



FIG. 2.5 – Détection d'un contour (en blanc).

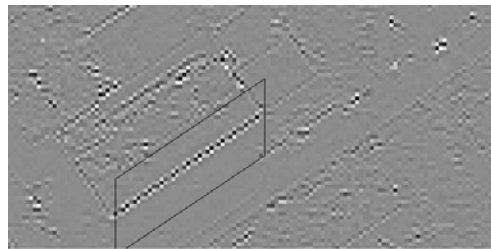


FIG. 2.6 – Sous-bande HL_1 et zone du contour étudié.

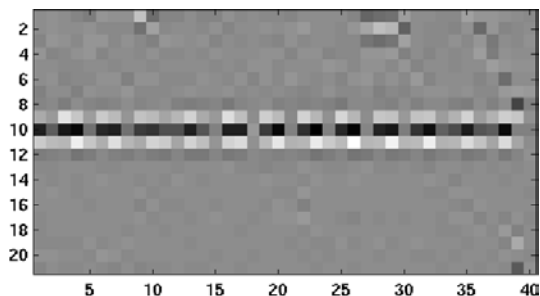


FIG. 2.7 – Lecture colonne par colonne de la zone du contour étudié.

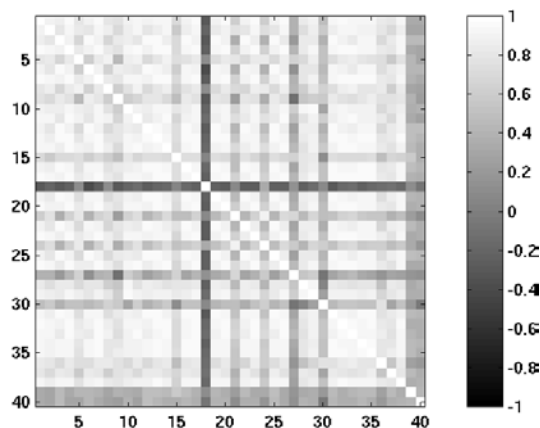


FIG. 2.8 – Matrice des coefficients de corrélation entre les colonnes de coefficients d'ondelettes de la zone étudiée.

de la transformée en ondelettes 2D séparable à représenter correctement les contours. En effet, cette transformée n'est capable d'exploiter les régularités géométriques que dans les directions des deux axes puisque les filtres sont appliqués horizontalement et verticalement, alors que des contours interviennent avec toutes les orientations. Nous mettons ici en évidence les corrélations entre coefficients d'ondelettes qui peuvent exister le long d'un contour dans une image. Ces corrélations ont donc un caractère exceptionnel puisqu'elles n'existent que dans certaines zones bien précises de la transformée en ondelettes.

Sur la figure 2.5, un contour a été détecté en utilisant un filtrage de Sobel vertical. Les coefficients d'ondelettes de la sous-bande HL_1 de la transformée en ondelettes de cette image sont représentés sur la figure 2.6. Sur cette figure, la zone des coefficients d'ondelettes étudiés autour du contour est représentée par le parallélogramme. Ces coefficients réordonnés le long de ce contour, représentés sur la figure 2.7 permettent de mettre en évidence les dépendances géométriques. En considérant chaque colonne de cette figure comme une variable aléatoire et chaque ligne comme une réalisation, il est possible de tracer la matrice des coefficients de corrélation de toutes ces variables aléatoires 2 à 2 (figure 2.8). Nous constatons qu'il existe des corrélations supérieures à 0,7 même à longue distance, par exemple entre la colonne 1 et la colonne 38.

Comme la résolution des capteurs embarqués à bord des satellites augmente, les structures observées sont de plus en plus longues. En particulier, les structures liées à l'activité

humaine ont des formes régulières. L'exploitation de ces régularités suscite un intérêt évident. L'objectif pour des codeurs en ondelettes orientés vers l'exploitation de la géométrie serait donc de pouvoir tenir compte des structures de l'image afin d'exploiter les redondances d'information. L'objectif d'une transformée géométrique serait d'adapter l'orientation des éléments de base à la direction des contours. Cependant, il ne faut pas oublier le caractère exceptionnel de ces corrélations à longue distance qui ne peuvent être observées que le long des contours. Les dépendances statistiques observées dans la section précédente ont plus de poids. En effet, les contours ne représentent qu'une très petite partie des images naturelles. Les textures occupent une place beaucoup plus importante.

2.2 Exploitation des dépendances dans les codeurs par plans de bits

Afin d'exploiter les dépendances entre coefficients d'ondelettes, plusieurs méthodes de codage ont été mises au point dans la littérature. Ici, nous décrivons trois techniques de codage par plans de bits : le codeur par plan de bits BPE (Bit Plane Encoder) du standard du CCSDS qui est dérivé des techniques par arbres de zéros, le codeur EBCOT de la norme JPEG2000 et les codeurs morphologiques. Pour décrire les techniques de codage par plans de bits, il faut introduire la notion de *coefficient significatif*.

Définition. On dit que le coefficient d'ondelettes x est significatif au plan de bit b si :

$$|x| \geq 2^b.$$

Soit un coefficient d'ondelettes x tel que $2^b \leq |x| < 2^{b+1}$. Dans la représentation binaire de x , le premier bit valant 1 en partant du plan de bits le plus significatif se trouve au plan de bits b . Le processus de codage est effectué du plan de bits le plus significatif vers les plans de bits les moins significatifs. Le coefficient x n'est pas significatif à tous les plans de bits supérieurs à b : il n'y a que des bits 0 à coder. Ce coefficient x devient significatif au plan de bits b : le bit à coder est 1. Dès lors on sait que $2^b \leq |x| < 2^{b+1}$. C'est la fin de la partie *codage de la signifiante* pour ce coefficient. Le bit de signe du coefficient x est codé dès que ce coefficient devient significatif, c'est-à-dire dès que le premier bit 1 a été codé pour ce coefficient. Commence ensuite la partie *affinage de l'amplitude*. Le coefficient x est significatif aux plans de bits inférieurs à b . Les bits à coder sont 0 ou 1 pour affiner la connaissance de l'amplitude de $|x|$ dans l'intervalle $[2^b, 2^{b+1}]$.

2.2.1 BPE : le codeur du standard CCSDS

Le codeur par plans de bits du CCSDS [29, 30] a été spécialement conçu pour la compression à bord des satellites. Les contraintes d'implantation spécifiques, décrites dans la section 1.4.3, ont donc été prises en compte lors de sa conception. C'est pourquoi nous décrivons ici cet algorithme et la stratégie employée pour exploiter les dépendances entre coefficients d'ondelettes.

Le BPE est dérivé de la famille des codeurs par arbres de zéros sans en faire vraiment partie. Les codeurs de cette famille tels que le codeur EZW (Embedded Zerotree Wavelet algorithm) [45] ou SPIHT (Set Partitioning In Hierarchical Trees) [46] exploitent les redondances d'information inter- et intra-bandes dans les plans de bits des coefficients d'ondelettes. La plupart des coefficients d'ondelettes de grande amplitude proviennent des contours de l'image et sont donc généralement regroupés. Ces groupes de forts coefficients sont séparés par de larges zones dans lesquelles les coefficients d'ondelettes sont faibles. Les bits correspondant à ces zones sont donc tous à zéros sur plusieurs des plans de bits les plus significatifs. Les arbres de zéros exploitent la redondance dans ces plans de bits de façon comparable à un *quadtrees*. De plus, les grandes structures de l'image se répètent à toutes les échelles. Les zones de faibles coefficients d'ondelettes aussi mais dans le sens des échelles de résolution grossière vers les échelles de résolution fine. A un plan de bits b , si un coefficient d'ondelettes n'est pas significatif dans une sous-bande de résolution grossière alors, statistiquement, les coefficients d'ondelettes correspondant à la même zone de l'image dans les sous-bandes à une résolution plus fine ont une forte probabilité de ne pas être significatifs. Les arbres de zéros exploitent aussi cette redondance inter-échelles dans les plans de bits.

Après une transformée en ondelettes sur trois niveaux, l'algorithme du CCSDS traite des *blocs* de 64 coefficients d'ondelettes correspondant à une région de taille 8×8 pixels dans l'image originale, aux dépassements des filtres en ondelettes près. Chaque bloc comporte :

- 1 coefficient de la sous-bande LL_3 ;
- 3×1 coefficients des sous-bandes HL_3 , LH_3 et HH_3 : les parents ;
- 3×4 coefficients des sous-bandes HL_2 , LH_2 et HH_2 : les enfants ;
- 3×16 coefficients des sous-bandes HL_1 , LH_1 et HH_1 : les petit-enfants.

La figure 2.9 illustre cette répartition dans un bloc. Le codage des coefficients d'ondelettes est effectué par plans de bits sur ces blocs de coefficients. Les coefficients de la sous-bande LL_3 sont codés séparément et nous ne décrivons pas ici comment ils sont codés. Le lecteur pourra se référer à la description du standard dans [29]. Ici nous nous intéressons au codage des coefficients des autres sous-bandes.

Le BPE du CCSDS exploite des plages de zéros par l'utilisation des *mots de transition*. Aussi, afin de localiser ces plages de zéros, plusieurs ensembles de coefficients sont définis sur la figure 2.9 :

- les coefficients parents sont notés p_i avec $i \in [0, 2]$;
- les ensembles de 4 coefficients enfants de la même sous-bande c_{ij} avec $j \in [0, 3]$ sont notés C_i ;
- les ensembles de 16 coefficients petits-enfants de la même sous-bande sont notés G_i ;
- chaque ensemble G_i est partitionné en 4 sous-ensembles de 4 coefficients notés H_{ij} ;
- la concaténation des ensembles C_i et G_i forme l'ensemble des descendants noté D_i ;
- la concaténation de tous les ensembles de descendants D_0 , D_1 et D_2 forme l'ensemble B .

Le plan de bits le plus significatif est déterminé par le coefficient d'ondelettes de plus

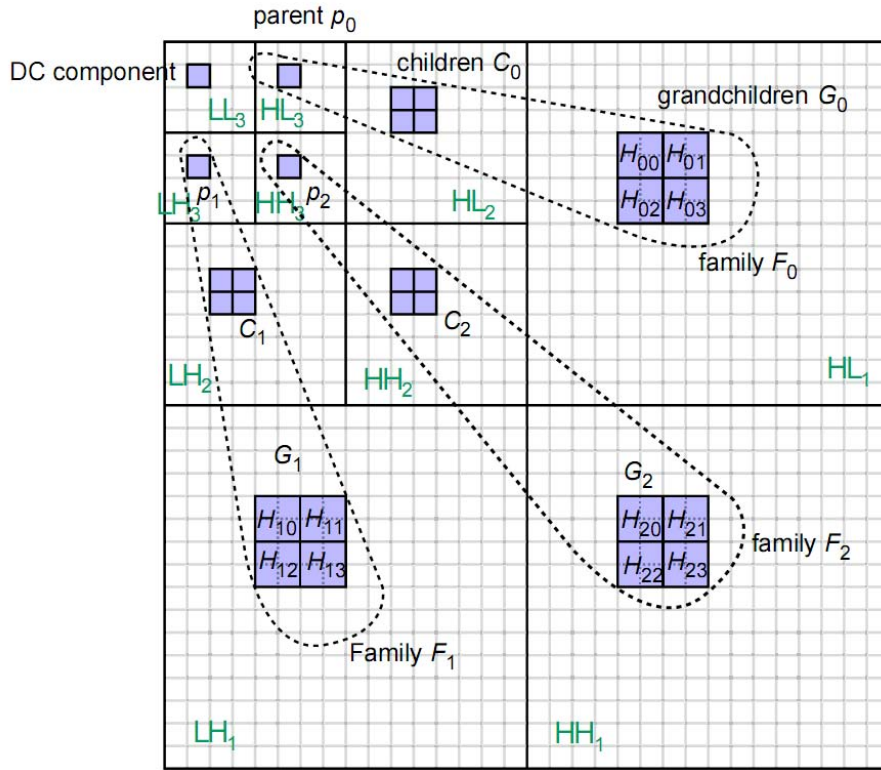


FIG. 2.9 – Bloc de 64 coefficients d’ondelettes tel que défini dans le standard du CCSDS (figure extraite de [29]).

grande amplitude dans le bloc (mis à part le coefficient de la sous-bande LL_3). Le codage de la signifiante est effectué comme décrit ci-dessous :

A un plan de bits b , après le codage des coefficients parents, des mots de transition sont utilisés pour signaler les plages de zéros, c’est-à-dire les ensembles dans lesquels tous les coefficients sont non signifiants au plan de bits b . Ainsi, le premier mot de transition $tran_B$ est un mot de 1 bit indiquant si tous les coefficients de l’ensemble B sont encore non signifiants au plan de bits b . Sinon, le mot de transition $tran_D$ permet d’indiquer les ensembles D_i dans lesquels tous les coefficients sont encore non signifiants. Dans les ensembles D_i contenant des coefficients signifiants, les coefficients enfants des ensembles C_i sont codés et le mot de transition $tran_G$ indique les ensembles G_i dans lesquels tous les coefficients petits-enfants sont encore non signifiants. Enfin, dans les ensembles G_i contenant des coefficients signifiants, le mot de transition $tran_{Hi}$ indique les ensembles H_{ij} dans lesquels tous les coefficients sont encore non signifiants. Les petits-enfants des groupes H_{ij} signifiants sont ensuite codés. Ce processus de codage est détaillé sur un exemple dans l’annexe B.

Le BPE du CCSDS n’est pas à proprement parler un codeur par arbres de zéros dans la mesure où il n’exploite pas les relations entre la signifiante des coefficients d’un niveau de résolution à l’autre. Par exemple, il n’exploite pas la relation entre un coefficient p_i et l’ensemble C_i comme dans un arbre de zéro. De même, la relation entre le coefficient c_{ij}

et l'ensemble H_{ij} n'est pas exploitée. En revanche, les redondances sont exploitées dans des plages de zéros multi-échelles ($tran_B$ et $tran_D$) ou intra-bande ($tran_{Hi}$).

Dans le chapitre 5 et dans [47], nous modifions ce codeur pour l'adapter à une post-transformée dans un schéma de compression ayant une faible complexité calculatoire. Nous verrons que dans ce schéma, la post-transformée permet d'exploiter des dépendances de type « textures » et ainsi de créer plus de plages de zéros dans les sous-bandes haute fréquence.

2.2.2 EBCOT : le codeur de la norme JPEG2000

Avec SPIHT, le codeur EBCOT est la référence pour la compression d'images à base de la transformée en ondelettes. Les performances en compression de cet algorithme figurent parmi les meilleures de l'état de l'art [48]. C'est pourquoi nous décrivons la technique employée dans cet algorithme pour exploiter les dépendances entre coefficients d'ondelettes.

Des contextes pour exploiter les dépendances

EBCOT utilise une stratégie différente des arbres de zéros pour exploiter les dépendances statistiques entre coefficients. Les relations entre coefficients d'ondelettes voisins dans la même sous-bande sont exploitées au moyen d'un codage contextuel.

EBCOT code les plans de bits en trois passes : la passe de propagation de la significance, la passe d'affinage de l'amplitude, et la passe de nettoyage. Cette dernière permet de coder les bits qui n'ont pas été codés par les deux autres passes. Trois primitives de codage sont utilisées : le codage de la significance, le codage du signe, et l'affinage de l'amplitude [49]. 18 contextes permettent de décrire l'environnement des bits codés par chacune des primitives et ainsi de tenir compte des dépendances entre coefficients d'ondelettes voisins lors du codage entropique. La définition des contextes est donnée dans [27, 28] et est reproduite dans l'exemple de codage en annexe C.

La primitive de codage de la significance est utilisée pour le codage des coefficients qui ne sont pas significants au plan de bits b ou qui deviennent significants à ce plan de bits. Cette primitive utilise 9 contextes de codage. La détermination du contexte pour le codage d'un coefficient d'ondelettes x dans les sous-bandes HL ou LH dépend plus des voisins dans la direction du filtre passe-haut que des voisins dans la direction du filtre passe-bas ou des voisins diagonaux. Ces choix sont cohérents avec les corrélations et les informations mutuelles observées entre coefficients d'ondelettes voisins sur les figures 2.3 et 2.4. Dans les sous-bandes HH, la significance des voisins diagonaux a plus d'importance que la significance des autres voisins pour la détermination du contexte. Dans cette primitive de codage de la significance, un contexte additionnel dit « run-length » est utilisé pour le codage de séquences de zéros.

La primitive de codage du signe utilise 5 contextes et est basée sur une prédiction effectuée à partir d'arguments spectraux [28]. Les images ont généralement un spectre passe-bas. Les filtrages passe-bas suivis des décimations préservent cette allure du spectre de l'image. Les coefficients voisins dans la direction du filtre passe-bas auront donc ten-

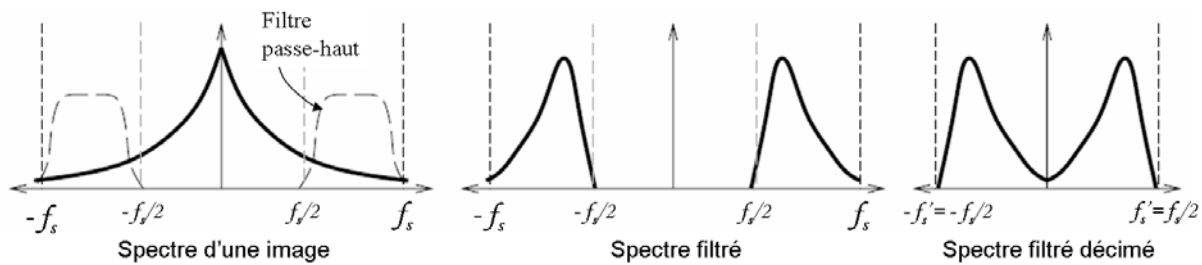


FIG. 2.10 – Spectre d’une image filtrée par un passe-haut et décimée. Figure extraite de [28].

dance à avoir le même signe. Par contre, les filtrages passe-haut suivis des décimations impliquent des repliements de spectre comme illustré sur la figure 2.10. Les coefficients voisins dans la direction du filtrage passe-haut auront donc tendance à avoir des signes opposés. Ceci est aussi en accord avec le signe des corrélations obtenues sur la figure 2.3.

Dans la primitive d’affinage de l’amplitude les 3 contextes sont déterminés selon la signifiante des coefficients voisins et selon le raffinement déjà effectué sur le coefficient codé.

Analyse de l’efficacité des contextes

C’est en utilisant ces contextes de codage qu’EBCOT exploite les dépendances entre coefficients d’ondelettes. Pourtant, les contextes utilisés dans EBCOT ont été définis de façon heuristique dans [28]. Plusieurs études ont été menées pour évaluer l’efficacité de ces contextes. Dans [50], une quantification contextuelle optimale au sens de l’information mutuelle est recherchée à partir de tous les états possibles du voisinage.

La quantification contextuelle est l’opération qui permet de regrouper des états $\mathbf{c}$ du voisinage en un nombre réduit de contextes. Par exemple, dans la primitive de codage de la signifiante, il existe 2^8 états $\mathbf{c}$ correspondant à la signifiante des huit voisins du coefficient codé. Après quantification contextuelle chaque état de l’ensemble $A_m = \{\mathbf{c} \mid Q(\mathbf{c}) = m\}$, $m = 1, \dots, M$ est associé au contexte m . La quantification contextuelle optimale est celle qui permet de maximiser l’information mutuelle entre les bits à coder et leur contexte. La meilleure quantification en termes d’information mutuelle est celle qui associe à chaque état possible du voisinage un contexte distinct. En effet, si on fusionne plusieurs contextes en un seul, l’information donnée par le contexte est réduite. L’information mutuelle entre les bits à coder et les contextes est donc plus faible [50]. Cependant, il est préférable d’utiliser un nombre réduit de contextes. En effet, on est très vite confronté au problème de *dilution* de l’information dans les contextes [40] : plus il y a de contextes, plus le nombre d’échantillons nécessaires à l’estimation des probabilités d’apparition conjointes des bits dans leur contexte est important. En conséquence, si on ne possède qu’un nombre limité d’échantillons, l’augmentation du nombre de contextes peut diminuer l’efficacité du codage parce que les probabilités d’apparition conjointes des bits dans leur contexte sont mal évaluées.

Pour le codage de la signifiante, une segmentation en quatre contextes optimaux permet d'obtenir la même quantité d'information mutuelle que la segmentation en neuf contextes définie dans EBCOT. De plus, une segmentation en neuf contextes optimaux permet d'obtenir une quantité d'information mutuelle quasiment égale à la segmentation en 256 contextes [50]. Enfin, dans la sous-bande HH, les voisins verticaux et horizontaux contribuent plus à l'information mutuelle que les voisins diagonaux [50], ce qui ne semble pas avoir été prévu lors de la définition des contextes d'EBCOT dans [27, 28]. Pourtant, cette constatation est cohérente avec les résultats que nous avons obtenus sur les figures 2.3 et 2.4 si l'on compare les corrélations ou les informations mutuelles des voisins numérotés de 1 à 4 dans la sous-bande HH.

Dans [35], nous avons proposé une mesure de l'efficacité des contextes d'EBCOT : un contexte m est efficace s'il permet de coder un grand nombre de symboles X avec une entropie conditionnelle $H_m(X|\mathbf{c} \in A_m)$ faible. L'indice m indique que cette entropie n'est calculée que sur les symboles pour lesquels le contexte est m c'est-à-dire tel que l'état $\mathbf{c} \in A_m$:

$$H_m(X|\mathbf{c} \in A_m) = - \sum_{k=\{0,1\}} P(X = k|\mathbf{c} \in A_m) \log_2 P(X = k|\mathbf{c} \in A_m).$$

L'entropie conditionnelle $H_m(X|m \in A_m)$ n'est jamais supérieure à l'entropie $H_m(X)$ des symboles pour lesquels le contexte est m . Nous avons alors proposé la mesure de l'efficacité $\eta(X, m)$ du contexte m définie par :

$$\eta(X, m) = P(\mathbf{c} \in A_m) (H_m(X) - H_m(X|\mathbf{c} \in A_m)).$$

Si le contexte m apparaît peu souvent, les probabilités $P(X|\mathbf{c} \in A_m)$ sont mal estimées. Dans ce cas, la probabilité $P(\mathbf{c} \in A_m)$ est faible et réduit l'efficacité $\eta(X, m)$ du contexte. Dans le cas contraire, il faut que l'entropie conditionnelle $H_m(X|\mathbf{c} \in A_m)$ soit faible pour que l'efficacité du contexte soit élevée. Cette mesure d'efficacité nous a permis de montrer que certains contextes utilisés par la primitive de codage de la signifiante ne sont pas efficaces. En utilisant une procédure d'optimisation basée sur la minimisation de la réduction d'information mutuelle au cours de la fusion de contextes [50], nous avons pu réduire le nombre des contextes utilisés par cette primitive de codage de 9 à 4 avec une perte inférieure à 0,01 dB quel que soit le débit. Ceci confirme l'inefficacité de certains contextes dans la primitive de codage de la signifiante.

Dans [51], des contextes « texture+énergie » sont créés. Ils prennent en compte non seulement la signifiante des voisins, mais aussi l'amplitude des coefficients voisins. L'utilisation de l'information énergétique n'augmente que peu la compression. Elle permet de diminuer le débit de 0,34% par rapport à EBCOT. Dans [52], une autre approche est proposée : effectuer la quantification contextuelle en ligne. Les contextes sont redéfinis à chaque plan de bits codé. Les résultats obtenus sont 0,6 dB en PSNR supérieurs à EBCOT au prix d'une augmentation de la complexité. Cependant, les contextes utilisés pour le codage de la signifiante prennent en compte la signifiante du coefficient parent, le coefficient à la même position spatiale mais à une résolution plus faible. On ne sait donc pas

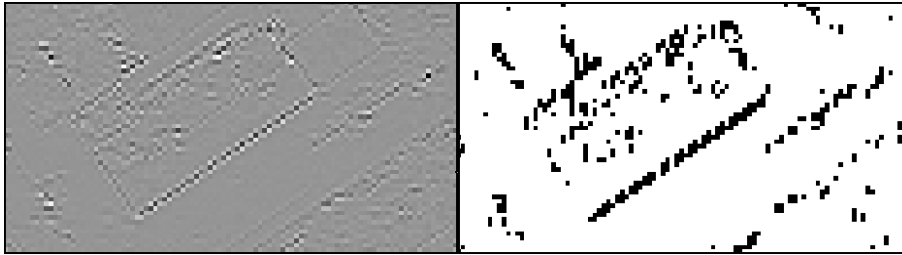


FIG. 2.11 – Sous-bande LH_1 de la transformée en ondelettes et coefficients significants au 3^e plan de bits.

si l’augmentation des performances provient de la quantification contextuelle en ligne ou de la prise en compte du coefficient parent.

Ces études ont montré que les contextes utilisés dans EBCOT permettent d’exploiter une grande partie de l’information contextuelle donnée par la signifiante des huit voisins du coefficient codé. Les contextes définis dans EBCOT ne sont cependant pas optimaux et il est possible de réduire le nombre de ces contextes pour obtenir des résultats équivalents.

Une approche contextuelle est envisagée pour exploiter les dépendances entre coefficients dans le schéma de compression par post-transformées. Dans le chapitre 4 et dans [53], nous utilisons un schéma de codage contextuel basique pour exploiter des dépendances énergétiques créées par les post-transformées.

2.2.3 Les codeurs morphologiques

Dans un schéma de compression progressif, les bits susceptibles de diminuer le plus la distorsion doivent être inclus en premier dans le train binaire [12]. C’est pourquoi les coefficients d’ondelettes sont codés plan de bits par plan de bits en commençant par les plans de bits les plus significatifs. A l’intérieur d’un plan de bits, ce sont les bits des coefficients significants qui permettent de réduire la distorsion. C’est sur le codage de la localisation des coefficients d’ondelettes significants que les codeurs efficaces se distinguent des codeurs inefficaces [48]. Dans les plans de bits les plus significatifs, la plupart des bits sont nuls. Le codage de ces bits nuls n’apporte aucune information sur les coefficients d’ondelettes mais sert à localiser implicitement les coefficients significants. Cependant, si on connaît la position des coefficients significants, alors il est inutile de coder ces bits nuls. Dans les codeurs morphologiques, au lieu de coder des plages de zéros, la position de nouveaux coefficients significants est recherchée à partir de la connaissance de la position des coefficients significants précédemment codés. Ils exploitent ainsi le fait que les coefficients significants sont en général agglomérés le long des contours ou dans les régions très structurées de l’image (cf. figure 2.11). Cette approche permet donc d’exploiter directement les faiblesses de la transformée en ondelettes pour la représentation de contours obliques. C’est pourquoi nous décrivons brièvement quelques codeurs de cette famille.

Les codeurs morphologiques codent les plans de bits suivant un parcours adaptatif : dès qu’un coefficient signifiant a été codé, le codage est propagé aux voisins de ce coefficient. Les codeurs MRWD (Morphological Representation of Wavelet Data) [54], SLCCA

(Significance-Linked Connected Component Analysis) [55], et EMDC (Embedded Morphological Dilatation Coding) [56] utilisent une dilatation pour propager le codage de la signifiante par croissance de régions.

Morphological Representation of Wavelet Data

Le codage d'ondelettes par approche morphologique a été introduit par Servetto *et al.* dans MRWD [54]. Dans une première version de ce codeur les coefficients d'ondelettes ne sont pas codés par plans de bits mais sont envoyés en pleine amplitude dans un codeur arithmétique contextuel adaptatif. Le parcours des coefficients d'ondelettes est effectué suivant un processus adaptatif récursif : dès qu'un coefficient signifiant (supérieur à un seuil prédéfini en valeur absolue) a été codé, le codage est propagé aux voisins de ce coefficient par dilatation. Ce processus s'apparente à une croissance de régions, la forme des régions étant transmise implicitement par les coefficients codés. Les relations inter-échelles sont aussi exploitées pour prédire la position des coefficients signifiants.

Significance-Linked Connected Component Analysis

Cependant, dans MRWD, la position d'un coefficient signifiant par agglomérat, le *germe*, doit être transmise explicitement. Pour diminuer le coût de codage lié aux germes, les agglomérats les plus petits sont éliminés dans SLCCA [55]. Des liens inter-échelles sont aussi définis pour coder la position des agglomérats d'une sous-bande enfant sachant la position des agglomérats de la sous-bande parent. Ainsi, deux agglomérats ont un *lien de signifiante* s'il existe dans l'un un parent signifiant qui a un enfant signifiant dans l'autre. Cela permet d'éviter le codage explicite des positions des germes pour la plupart des agglomérats de coefficients signifiants.

Les agglomérats dits multi-résolutions introduits dans cet algorithme comprennent des agglomérats à différentes échelles liés par des liens de signifiante. Les bonnes performances en débit-distorsion sont en partie dues à l'élimination des petits agglomérats qui ont un coût de codage élevé. Il peut donc exister de forts coefficients d'ondelettes isolés qui ne sont pas codés et cela peut entraîner une erreur maximale élevée. Ceci pose un problème pour la compression des images satellites puisque une forte erreur peut entraîner une mauvaise interprétation de la part des analystes, en particulier dans le cadre d'une exploitation militaire des images.

Embedded Morphological Dilatation Coding

EMDC [56] est un codeur morphologique plus abouti que MRWD [54] et SLCCA [55] : le codeur est basé sur un schéma de compression progressif. Dans EMDC [56], Lazzaroni *et al.* font l'hypothèse que les agglomérats de coefficients signifiants intra-bandes ne sont pas forcément connectés. C'est pourquoi ils utilisent une étape de dilatation supplémentaire pour analyser les régions autour des frontières des agglomérats. De plus, comme EBCOT, EMDC est un codeur par plans de bits fractionnaires. Cinq passes de codage sont ordonnées selon leurs performances respectives en débit-distorsion. La passe

de dilatation intra-bande est appliquée immédiatement après chaque coefficient significatif codé car c'est cette passe qui présente le meilleur compromis débit-distorsion. Les trois passes suivantes permettent d'augmenter la taille des agglomérats codés à chaque plan de bits. La cinquième passe permet le codage explicite de la position des nouveaux germes d'agglomérats. Les performances de ce codeur sont de l'ordre de 0,35 dB supérieures à JPEG2000.

Retour sur EBCOT

L'approche par dilatation est présente dans la passe de propagation de la signifiante d'EBCOT. En effet, dans cette passe, les coefficients codés sont les coefficients voisins de coefficients déjà significatifs. Cependant, la croissance des régions n'est pas effectuée totalement car le parcours des coefficients ne se fait que dans un seul sens (cf. annexe C). En conséquence, si un coefficient est trouvé significatif au plan de bits b , les coefficients au dessus et à gauche de ce coefficient ne seront pas forcément codés par la passe de propagation de la signifiante au plan de bits b . Pour y remédier, une passe supplémentaire de *reverse significance propagation* avait été prévue dans les premières versions d'EBCOT. Cette passe parcourait les coefficients en sens inverse. Elle a été abandonnée parce qu'elle ne contribuait que très peu à la performance en compression et complexifiait le parcours des coefficients [28].

Conclusions

Les codeurs morphologiques permettent d'adapter le codage au contenu des images. En effet, ils propagent le codage des coefficients d'ondelettes sur les structures. Le problème le plus évident pour les utiliser à bord des satellites est que ces codeurs demandent de grandes ressources en mémoire au moment du codage puisque le parcours adaptatif des coefficients requiert l'utilisation de listes. De plus, ce parcours adaptatif rend l'implantation difficile voire impossible lorsque la transformée en ondelettes est effectuée au fil de l'eau.

Dans les compresseurs actuels, les faiblesses de la transformée en ondelettes sont donc compensées par l'utilisation de codeurs plus ou moins complexes permettant d'exploiter les dépendances entre coefficients d'ondelettes avec ingéniosité. Cependant, des améliorations au niveau de la transformée peuvent être envisagées.

2.3 Transformées dérivées des ondelettes

Plutôt que d'exploiter les dépendances dans la transformée en ondelettes avec des codeurs complexes, une autre approche consiste à en chercher les sources afin de les éliminer. La transformée en ondelettes est optimale pour la représentation de singularités ponctuelles. Malheureusement, dans les images, la plupart des singularités rencontrées sont des contours [57] et le schéma de transformée en ondelettes séparable n'est pas adapté à la représentation de ces contours. Les seules directions bien appréhendées sont les directions horizontales et verticales. Beaucoup de coefficients d'ondelettes de forte

amplitude sont nécessaires à la représentation des contours obliques. Afin de représenter ces contours avec plus de parcimonie, de nouvelles transformées dérivées des ondelettes ont été proposées. Do et Vetterli [58] ont listé les propriétés recherchées :

1. *la multirésolution* ;
2. *la localisation* spatiale et fréquentielle des éléments de base ;
3. *l'échantillonnage critique* ;
4. *la multi-directionnalité* : les éléments de base doivent être orientés dans de multiples directions ;
5. *l'anisotropie* : les éléments de base doivent être plus ou moins allongés.

La transformée en ondelettes possède déjà les trois premières propriétés. Cependant, sa déficience à représenter les contours avec parcimonie provient de la faible directionnalité et anisotropie de ses éléments de base. La plupart des nouvelles transformées proposées dans la littérature possèdent ces deux propriétés mais souvent au détriment de la propriété d'échantillonnage critique. Cette dernière est la plus importante du point de vue de la compression. En effet, il est actuellement impossible d'avoir une représentation très parcimonieuse de toutes les images naturelles. Aussi, l'introduction d'une faible redondance lors de la transformée est un handicap difficile à surmonter pour la compression. Cependant, pour une compression à très bas débit, comme beaucoup de coefficients transformés sont annulés par la quantification, la représentation des images peut être parcimonieuse dans une transformée dérivée des ondelettes. Dans [59] de bons résultats visuels ont été obtenus à très bas débit avec un dictionnaire redondant de fonctions de bases orientées et un algorithme de *Matching Pursuit*. Les bases orientées permettent de mieux représenter les contours avec moins d'effets de rebonds que la transformée en ondelettes. Cependant, les résultats en PSNR ne sont supérieurs à ceux de JPEG2000 que pour des taux de compression supérieurs à 60. Pour la compression d'images satellite, nous visons une grande qualité et donc des taux de compression beaucoup plus faible entre 5 et 6. La propriété d'échantillonnage critique est donc nécessaire.

2.3.1 Transformées redondantes

Les contourlettes [58] et les curvelettes [57, 60] sont deux exemples de transformées géométriques qui sont proches dans leur concept. Elles constituent des familles de « frames » d'ondelettes géométriques conçues pour représenter les contours de façon parcimonieuse. Le problème majeur pour utiliser ces transformées pour des applications de compression est qu'elles sont redondantes.

Les curvelettes

La première génération des curvelettes [57] est construite à partir d'une décomposition en sous-bandes de l'image, d'un partitionnement par blocs de ces sous-bandes avec une taille de bloc dépendant de la sous-bande, et de la transformée en ridgelettes [61] de ces blocs. La transformée en ridgelettes consiste en une transformée de Radon suivie d'une

transformée en ondelettes 1D. La décomposition en sous-bandes et le partitionnement par blocs permettent la localisation fréquentielle et spatiale ainsi que la multirésolution et la transformée de Radon permet la multi-directionnalité. Les contours mis en évidence dans les sous-bandes hautes fréquences sont détectés dans toutes les directions par la transformée de Radon et représentés de façon parcimonieuse grâce à la transformée en ondelettes 1D.

La seconde génération des curvelettes [60] est définie directement à partir d'une fonction d'échelle φ et d'une forme d'onde γ . Les éléments de base

$$\gamma_\mu(x) = 2^{3j/2} \gamma(D_j R_{\theta_j} x - k)$$

sont obtenus par dilatation anisotrope $D = \begin{pmatrix} 2^{2j} & 0 \\ 0 & 2^j \end{pmatrix}$, rotation d'angle $\theta_j = \pi l 2^{-j}$ et translation de paramètre k de la forme d'onde γ . Cette nouvelle transformée en curvelettes est transposée dans le domaine discret et des implantations numériques sont proposées dans [62].

Les contourlettes

A l'inverse des curvelettes, les contourlettes sont construites directement dans le domaine discret [58]. L'image est décomposée dans une pyramide Laplacienne qui permet l'analyse multirésolution. Ensuite, un banc de filtres directionnels est utilisé sur chacun des niveaux de la pyramide ce qui permet l'analyse multi-directionnelle. La pyramide Laplacienne introduit un facteur de redondance de 4/3. Aussi, dans [63], une version des contourlettes ayant la propriété d'échantillonnage critique a été développée. Elle est basée sur des bancs de filtres permettant une analyse multirésolution et multi-directionnelle conjointe. Bien que cette dernière version paraisse adaptée à des problématiques de compression, nous avons préféré nous intéresser à la transformée en bandelettes.

2.3.2 Les bandelettes

Contrairement aux transformées en curvelettes et en contourlettes qui sont des transformées à bases fixes et qui ne dépendent donc pas de l'image, les transformées en bandelettes sont des transformées à bases adaptées [64, 65]. Elles nécessitent donc l'envoi d'une information supplémentaire pour signaler le choix qui a été fait.

Première génération de bandelettes

La première génération de la transformée en bandelettes a été proposée par Le Penec dans sa thèse [66]. L'image est d'abord divisée en régions contenant chacune un *flot géométrique* parallèle. Ces régions sont régulières le long des lignes de flot. Autour d'un contour, le flot est parallèle aux tangentes de la courbe du contour. Les paramètres géométriques doivent être transmis au décodeur. En pratique, les régions sont construites par *quadtree* pour minimiser le coût de la segmentation et une optimisation est effectuée permettant une affectation optimale du débit lié à la géométrie par rapport au débit lié au

codage des coefficients de bandelettes [67]. Les bandelettes sont des ondelettes déformées le long des lignes de flot. L'opération revient à effectuer une transformée en ondelettes 2D séparable sur une région déformée afin que les lignes de flots soient horizontales ou verticales. Cependant, si le flot est horizontal, les fonctions d'échelles dans le sens horizontal $\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$, où 2^j est le facteur d'échelle et k le paramètre de translation, sont remplacées par les ondelettes $\psi_{l,k}(x) = 2^{l/2}\psi(2^l x - k)$ avec $l < j$ qui possèdent l'avantage d'avoir des moments évanescents ce qui permet d'exploiter les régularités. Ceci est possible puisque la base orthonormale $\{\varphi_{j,k}(x)\}_k$ engendre le même espace $V_j = \bigcup_{l < j} W_l$ que la base $\{\psi_{l,k}(x)\}_{l < j, k}$. Dans le contexte d'un flot géométrique horizontal, l'étape de bandelettisation consiste donc à changer la famille d'ondelettes séparables $\{\varphi_{j,k_1}(x_1)\psi_{j,k_2}(x_2)\}_{j,k_1,k_2}$ par la famille $\{\psi_{l,k_1}(x_1)\psi_{j,k_2}(x_2)\}_{j,l < j, k_1,k_2}$ [67]. Les fonctions $\psi_{l,k_1}(x_1)\psi_{j,k_2}(x_2)$ sont appelées bandelettes car leur support est parallèle aux lignes de flot et elles sont allongées dans la direction du flot géométrique ($2^l < 2^j$). Cette transformée en bandelettes peut être implémentée dans le domaine discret à partir de bancs de filtres d'ondelettes.

Seconde génération de bandelettes

La seconde génération de bandelettes développée par Peyré [64] ne suit pas le même raisonnement. Cette nouvelle transformée en bandelettes est construite au-dessus d'une transformée en ondelettes 2D. L'objectif de ces bandelettes est d'exploiter les régularités géométriques le long des contours dans les sous-bandes de la transformée en ondelettes. Pour cela, une approximation polynomiale est effectuée sur des bandes qui suivent la géométrie. Peyré étend la théorie des multi-ondelettes de Alpert [68] au cas 2D : il utilise des polynômes déformés 2D à la place des polynômes 1D.

Dans la théorie des multi-ondelettes, plusieurs ondelettes mères $\psi^1, \dots, \psi^n$ polynomiales par morceaux et nulles en dehors de l'intervalle $[0, 1]$ sont utilisées. Elles sont orthogonales entre elles et forment une base de l'espace W_0^n . Les bases d'ondelettes $\{\psi_{j,k}^n(x)\}_k$ engendrent les espaces W_j^n orthogonaux aux espaces V_j^n qui sont définis comme l'ensemble des fonctions polynomiales de degré inférieur à n sur l'intervalle $[2^{-j}k, 2^{-j}(k+1)]$, pour $k = 0, \dots, 2^j - 1$ et nulles en dehors de cet intervalle. Les ondelettes construites sont donc orthogonales à des fonctions polynomiales par morceaux de degré inférieur à n sur des intervalles multirésolution de taille dyadique.

Là où Alpert effectue une subdivision dyadique des intervalles, Peyré effectue une subdivision dyadique de bandes déformées. Il construit ainsi les bandelettes de seconde génération qui permettent l'approximation multirésolution de fonctions dans des bandes à partir de fonctions polynomiales par morceaux.

Une première version de cette transformée en bandelettes pour la compression est la transformée en bandelettes d'ordre zéro [64, 69]. Dans cette version, les géométries considérées sont linéaires à l'intérieur des blocs issus d'une segmentation en *quadtrees* des sous-bandes de la transformée en ondelettes de l'image. La transformée de Haar monodimensionnelle est effectuée sur les coefficients d'ondelettes de chaque bloc réordonnés selon la direction de la géométrie. Cette transformée est équivalente à la transformée de Alpert d'ordre 0 [64].

La transformée en bandelettes par groupements est une seconde version de la transformée en bandelettes utilisée pour la compression [64]. Dans cette version, les sous-bandes de la transformée en ondelettes sont segmentées en blocs de taille 4×4 ce qui permet d'économiser le calcul et la transmission du *quadtrees*. Dans ces blocs, des groupes de coefficients d'ondelettes sont transformés indépendamment les uns des autres dans des bases orthogonales de polynômes discrets. Cette transformée en bandelettes ne correspond plus aux multi-ondelettes de Alpert car il n'y a pas de propriété de multirésolution sur les bandes. C'est à partir de cette transformée en bandelettes par groupements que les post-transformées étudiées dans cette thèse sont construites. La transformée en bandelettes par groupements est expliquée en détails dans le chapitre suivant.

2.4 Conclusion

La transformée en ondelettes possède des propriétés intéressantes du point de vue de la compression. En particulier, elle est utilisée dans les récents codeurs pour la compression d'images à bord des satellites. Cependant, nous avons vu qu'il existe des dépendances statistiques entre coefficients d'ondelettes. Ces dépendances ont été décrites dans la littérature. Dans ce chapitre et dans [34], nous avons mis en évidence des corrélations et de l'information mutuelle entre coefficients d'ondelettes voisins dans la transformée en ondelettes d'images satellite ainsi que des corrélations entre coefficients d'ondelettes dans le voisinage d'un contour. Dans les codeurs par plans de bits, différentes stratégies sont développées pour exploiter ces dépendances entre coefficients d'ondelettes : exploiter les plages de zéros, utiliser des contextes et modifier le parcours des coefficients pour suivre les contours. Dans [35], nous avons montré qu'il est possible de réduire le nombre de contextes dans EBCOT. Une approche différente consiste à utiliser une transformée dérivée des ondelettes afin d'éliminer ces dépendances dès l'étape de transformée. La figure 2.12 résume les étapes dans lesquelles les dépendances peuvent être exploitées dans un schéma de compression. Dans la suite, nous avons choisi de nous intéresser à la transformée en bandelettes par groupements qui constitue un moyen peu complexe permettant une amélioration des résultats en compression par rapport à ceux obtenus avec la transformée en ondelettes.

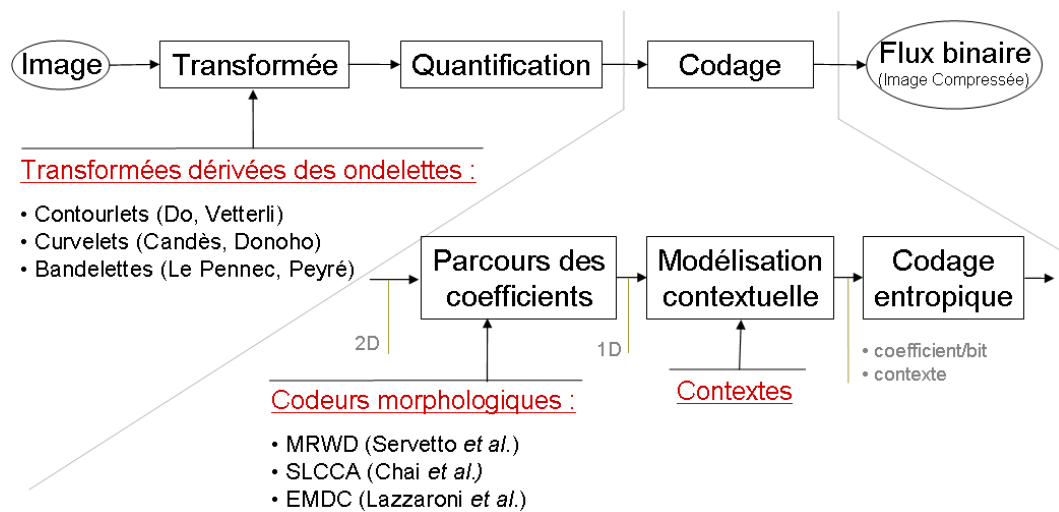


FIG. 2.12 – Exploitation des dépendances dans un schéma de compression.

Chapitre 3

Les bandelettes appliquées à la compression d'images satellite

3.1	La transformée en bandelettes par blocs	42
3.1.1	Construction du dictionnaire de bases	42
3.1.2	Transformée en bandelettes d'un bloc de coefficients d'ondelettes	45
3.1.3	Sélection de la meilleure base de bandelettes	47
3.1.4	Algorithme de compression en bandelettes	50
3.2	Choix de la taille des blocs	52
3.2.1	Corrélations entre coefficients d'ondelettes	52
3.2.2	Extension des bases de bandelettes à des blocs de tailles diverses	53
3.3	Sélection des meilleures bases de bandelettes par optimisation débit-distorsion	54
3.3.1	Le problème d'allocation débit-distorsion	54
3.3.2	Calcul théorique du multiplicateur de Lagrange optimal	56
3.3.3	Recherche empirique du multiplicateur de Lagrange optimal . .	59
3.3.4	Conclusion	68
3.4	Études expérimentales des performances de la transformée en bandelettes	68
3.4.1	Analyse de la sélection des bases directionnelles	68
3.4.2	Estimation du débit des coefficients de bandelettes	70
3.4.3	Estimation du débit de signalisation	71
3.4.4	Performances du critère de sélection des bases	75
3.4.5	Performances des bases de bandelettes	77
3.5	Conclusion	81

Dans sa thèse [64], Peyré propose une adaptation de la théorie des bandelettes pour une application à la compression d'images. C'est de ce travail qu'est issue la transformée en bandelettes par blocs et par groupements. Dans ce chapitre, nous reprenons toutes les étapes de cette transformée afin d'en optimiser les paramètres pour la compression d'images d'observation de la Terre à haute résolution. Dans la section 3.1, nous décrivons comment est effectuée la transformée en bandelettes par blocs. En particulier nous expliquons comment les bases de bandelettes sont construites et quelle est la base de bandelettes retenue pour transformer un bloc de coefficients d'ondelettes. Ensuite, dans la section 3.2, nous étudions la taille des blocs qu'il convient de transformer à partir d'une analyse des corrélations entre coefficients d'ondelettes distants dans les images satellite. De plus nous étendons la construction des bases de bandelettes directionnelles à des blocs de tailles diverses. Dans la section 3.3 nous nous intéressons au problème d'optimisation du critère débit-distorsion pour le choix de la meilleure base de bandelettes. Nous avons publié une partie de cette étude dans [70]. Enfin, dans la section 3.4 nous étudions les performances en compression des différentes étapes de la transformée en bandelettes par blocs et nous apportons des améliorations au processus d'estimation du débit de signalisation.

3.1 La transformée en bandelettes par blocs

Après la transformée en ondelettes, les sous-bandes sont segmentées en blocs de 4×4 coefficients. Un dictionnaire de bases orthonormales est défini. Chaque bloc de coefficients d'ondelettes est alors transformé dans chacune des bases de ce dictionnaire. Pour un bloc de coefficients d'ondelettes, la meilleure représentation de ce bloc en termes de débit-distorsion est retenue pour la compression.

3.1.1 Construction du dictionnaire de bases

Bases de bandelettes par groupements directionnels

Les coefficients de bandelettes étant les coefficients obtenus par projection des coefficients d'ondelettes de chaque bloc dans une base polynomiale, la reconstruction des coefficients d'ondelettes à partir des coefficients de bandelettes quantifiés est une approximation polynomiale des coefficients d'ondelettes originaux. De plus, cette approximation polynomiale est directionnelle.

Douze directions discrètes sont définies sur une grille de 4×4 pixels dans le tableau 3.1. Pour chacune de ces directions, les pixels voisins dans la direction considérée sont liés et définissent *un groupe*. Les 12 configurations de groupement suivant ces directions et telles que définies par Peyré dans [64] sont représentées sur la figure 3.1. Si l'on suit strictement les directions définies dans le tableau 3.1, certains pixels de la grille 4×4 se retrouvent isolés. Dans [64] et sur la figure 3.1, il a été choisi de regrouper les pixels isolés suivant la direction horizontale ou verticale la plus proche de la direction des autres groupes.

Configuration	#1	#2	#3	#4	#5	#6
Angle (rad)	$\text{atan}(1/3)$	$\text{atan}(1/2)$	$\pi/4$	$\text{atan}(2)$	$\text{atan}(3)$	$\pi/2$
Angle (°)	18,4	26,6	45	63,4	71,6	90
Configuration	#7	#8	#9	#10	#11	#12
Angle (rad)	$-\text{atan}(3)$	$-\text{atan}(2)$	$-\pi/4$	$-\text{atan}(1/2)$	$-\text{atan}(1/3)$	0
Angle (°)	-71,6	-63,4	-45	-26,6	-18,4	0

TAB. 3.1 – Angles correspondant aux différentes directions des groupes.

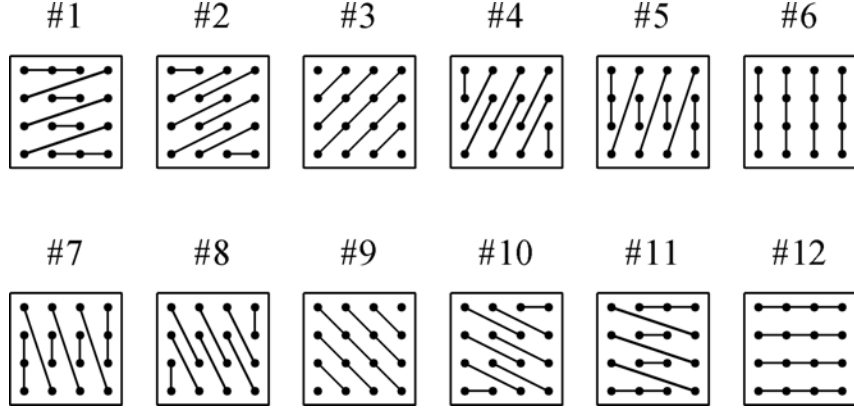


FIG. 3.1 – 12 configurations de groupes directionnels.

A une configuration correspond un ensemble de groupes, noté G , et constitué de N_G groupes : $G = \{G_i\}_{i=1}^{N_G}$. Soit G_i un groupe de cet ensemble contenant g_i pixels. Les bases directionnelles sont construites à partir des bases de polynômes de Legendre discrets de degré 0 à $g_i - 1$. Comme les groupes peuvent contenir de un à quatre pixels, quatre bases orthonormales de polynômes discrets sont définies. Elles sont représentées sur la figure 3.2.

Dans le cas où $g_i = 1$, le groupe n'est constitué que d'un pixel. Il n'existe évidemment qu'une seule base dans $\mathbb{R}^1$. Dans le cas où $g_i = 2$, on obtient la base de Haar. Elle correspond à la moyenne et à la différence des pixels (les deux vecteurs normalisés de cette base sont $[1/\sqrt{2}, 1/\sqrt{2}]$ et $[1/\sqrt{2}, -1/\sqrt{2}]$). Les cas où $g_i > 2$ sont plus difficiles à interpréter bien qu'il y ait toujours une composante continue correspondant à la moyenne des pixels.

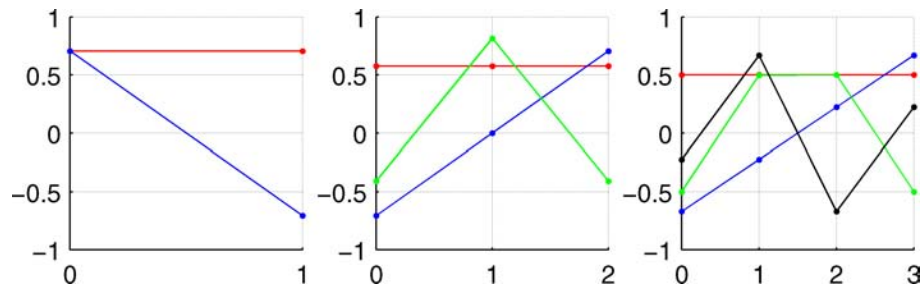
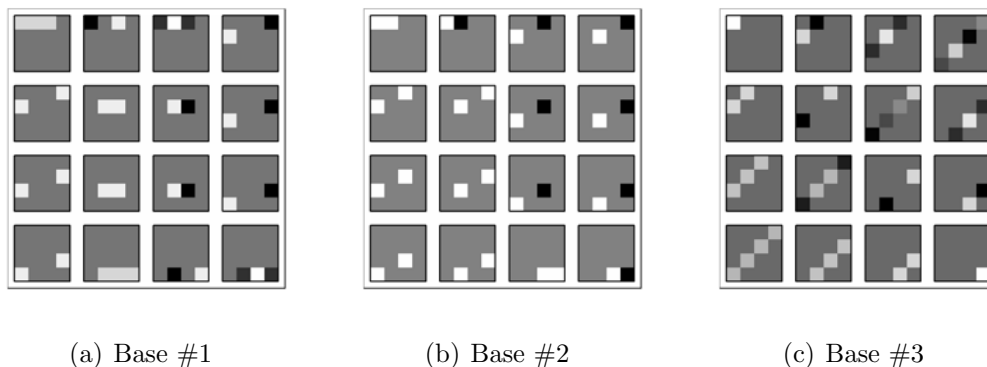


FIG. 3.2 – Bases de polynômes discrets de Legendre d'ordre 2 à 4.


 FIG. 3.3 – Base de bandelettes directionnelles sur $\mathbb{R}^{16}$.

Par concaténation des différentes bases de $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$ ou $\mathbb{R}^4$ dans l'espace $\mathbb{R}^{16}$ en suivant les sous-espaces définis par les groupements de la figure 3.1, on obtient des bases directionnelles de $\mathbb{R}^{16}$. Les bases correspondantes aux directions #1, #2 et #3 sont représentées sur la figure 3.3.

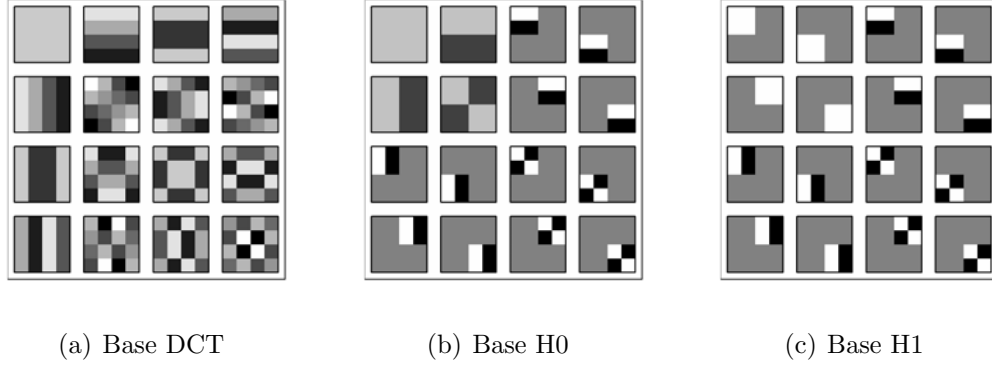
Chaque base est composée de 16 vecteurs représentés sous forme de bloc de 4×4 coefficients. Notons que ces bases sont toutes orthonormales. En effet, les vecteurs sont construits sur des sous-espaces orthogonaux entre eux et les bases de Legendre dans ces sous-espaces sont orthonormales. La somme des g_i sur la grille 4×4 est égale à 16 et les groupes ne se chevauchent pas. En conséquence, à une direction correspond une base de 16 vecteurs orthogonaux constitués par les bases polynomiales dans les différents groupes G_i . Cette base est une base de bandelettes. Comme 12 directions discrètes sont définies pour des blocs de taille 4×4 , on obtient 12 bases directionnelles.

Bases complémentaires

Pour compléter ces 12 bases directionnelles, la base de DCT et deux bases de Haar notées H0 et H1 sont utilisées. Ces 3 bases complémentaires permettent de prendre en compte des configurations qui ne sont pas décrites par les bases directionnelles et doivent permettre de mieux décrire les textures informes. Elles sont représentées sur la figure 3.4. Le dictionnaire de bandelettes construit par Peyré est donc composé de 15 bases orthonormales : les 12 bases directionnelles et les 3 bases complémentaires.

A ce dictionnaire s'ajoute la base canonique de $\mathbb{R}^{16}$. La transformée dans cette base est l'identité. Elle correspond en fait à une absence de transformée. Comme elle ne requiert aucun calcul mis à part le calcul du critère débit-distorsion (c.f. section 3.1.3), cette transformée sera toujours en concurrence avec les autres même si cela n'est pas systématiquement mentionné dans la suite.

Le dictionnaire de bases de bandelettes $\mathcal{D}$ contient donc $N_{\mathcal{B}} = 15$ bases notées $\mathcal{B}_b$ avec $b \in [1, N_{\mathcal{B}}]$. La base canonique sera identifiée par l'indice $b = 0$ dans la suite.


 FIG. 3.4 – Bases complémentaires sur $\mathbb{R}^{16}$.

3.1.2 Transformée en bandelettes d'un bloc de coefficients d'ondelettes

La transformée en bandelettes d'un bloc de coefficients d'ondelettes se fait par simple projection du bloc de coefficients d'ondelettes sur les vecteurs de la base de bandelettes.

Soit M la dimension de l'espace. Les M vecteurs de la base $\mathcal{B}_b$ sont notés ϕ_m^b avec $m \in [1, M]$. Un bloc de $M = 4 \times 4$ coefficients d'ondelettes constitue un vecteur de l'espace $\mathbb{R}^M$. Il est noté f . L'objectif de la transformée en bandelettes est de trouver la base $\mathcal{B}_b$ du dictionnaire de bandelettes $\mathcal{D}$ dans laquelle ce vecteur f sera le mieux représenté pour la compression. On note f^b les différentes représentations du vecteur f dans les $N_B + 1$ bases $\mathcal{B}_b$ du dictionnaire. La transformée en bandelettes du bloc de coefficients d'ondelettes f est effectuée par le changement de base :

$$f^b = \sum_{m=1}^M \langle f, \phi_m^b \rangle \phi_m^b. \quad (3.1)$$

Les coefficients de bandelettes notés $a_b[m]$ correspondent aux produits scalaires entre le vecteur f et les vecteurs ϕ_m^b de la base $\mathcal{B}_b$:

$$a_b[m] = \langle f, \phi_m^b \rangle.$$

Un exemple de la transformée en bandelettes d'un bloc de 4×4 coefficients d'ondelettes dans la base directionnelle #3 est représenté sur la figure 3.5. Un coefficient de forte amplitude positive (en blanc ligne 3, colonne 3) apparaît dans la transformée en bandelettes. Ce coefficient résulte du produit scalaire entre le bloc de coefficients d'ondelettes et le vecteur de la base #3 à la même position que ce coefficients. Les coefficients non nuls de ce vecteur coïncident avec ceux du bloc de coefficients d'ondelettes d'où le fort coefficient de bandelettes obtenu par produit scalaire. Sur la figure 3.6(b), les blocs marqués d'un trait rouge sont transformés dans une base de bandelettes directionnelle. Ceux marqués C sont transformés dans la base DCT et ceux marqués H0 ou H1 en bleu sont transformés dans les bases de Haar représentées sur la figure 3.4. Les blocs non marqués sont les blocs de coefficients d'ondelettes qui ne sont pas transformés en bandelettes. Le bloc entouré en vert est le bloc de l'exemple de la figure 3.5. Ce choix de la transformée en bandelettes sur

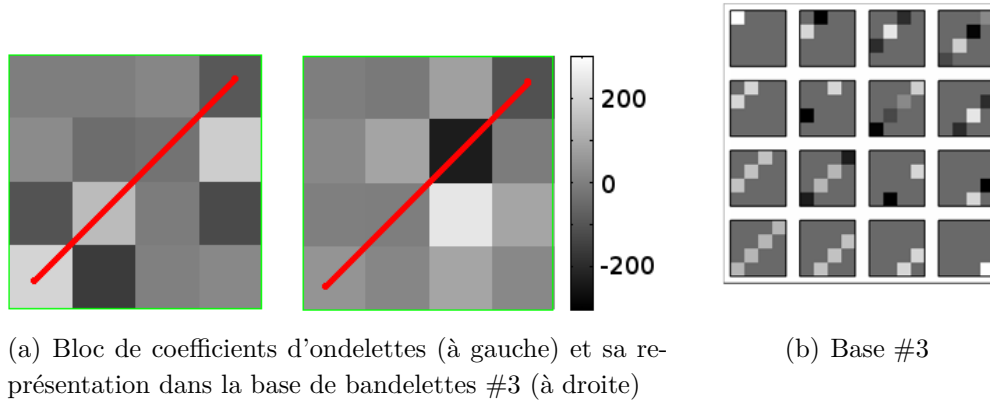


FIG. 3.5 – Exemple de transformée en bandelettes dans la base #3.

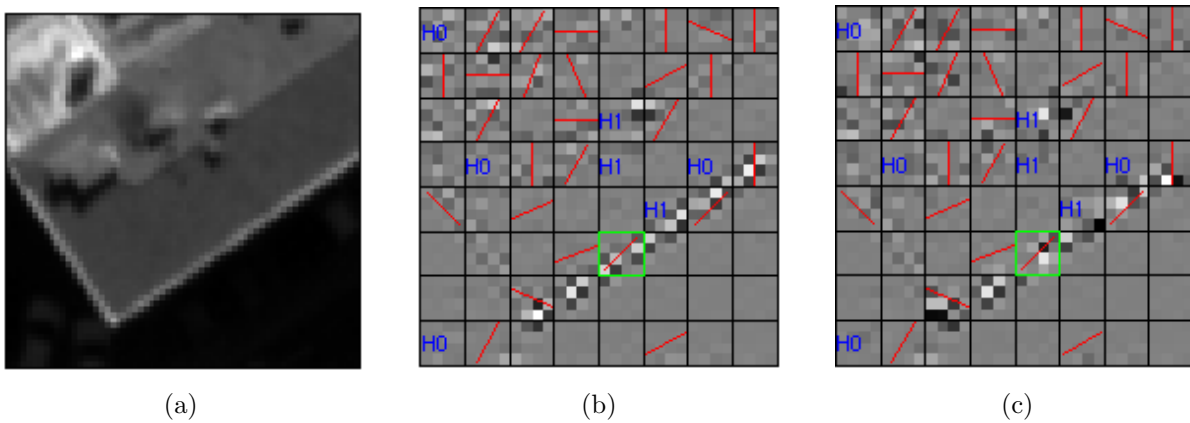
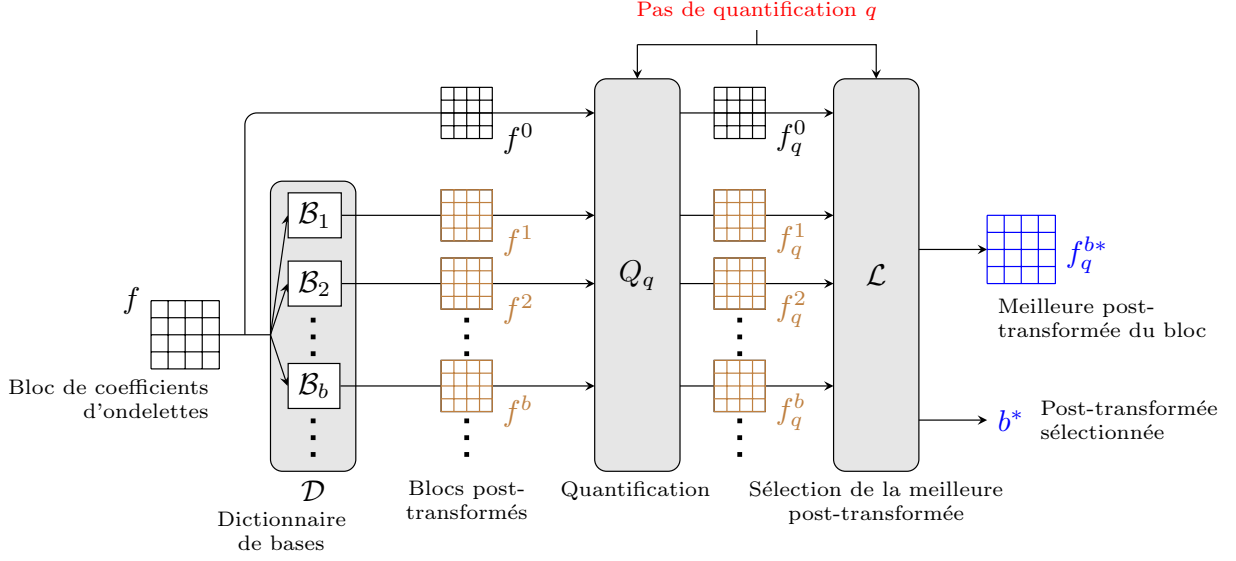


FIG. 3.6 – (a) Coin d'une toiture vue par le capteur PELICAN. (b) Sous-bande LH_1 de la transformée en ondelettes. (c) Sous-bande LH_1 de la transformée en bandelettes.


 FIG. 3.7 – Processus de transformée en bandelettes d'un bloc f de coefficients d'ondelettes.

un bloc f est le résultat d'une mise en concurrence de toutes les représentations f^b de ce bloc dans toutes les bases $\mathcal{B}_b$ du dictionnaire. Comme expliqué dans la section suivante, c'est la représentation f^b qui minimise un critère débit-distorsion qui sera conservée.

3.1.3 Sélection de la meilleure base de bandelettes

La figure 3.7 illustre le processus de transformée en bandelettes d'un bloc f de coefficients d'ondelettes. Ce bloc est d'abord représenté dans toutes les bases $\mathcal{B}_b$ du dictionnaire $\mathcal{D}$. Les représentations f^b du bloc f sont ensuite quantifiées par un quantificateur scalaire uniforme à zone morte et de pas q . Enfin, la meilleure représentation en termes de débit et de distorsion f_q^{b*} est sélectionnée parmi toutes les représentations quantifiées f_q^b pour être transmise au décodeur. L'indice b^* de la base correspondante est aussi transmise afin que le décodeur puisse correctement inverser la transformée en bandelettes. Cette information supplémentaire est appelée information de *signalisation* dans la suite.

Quantification scalaire

L'étape de transformation en bandelettes est parfaitement inversible puisque les bases utilisées sont orthogonales. Cependant des pertes sont introduites par quantification des coefficients de bandelettes $a_b[m]$.

La quantification effectuée est une quantification scalaire uniforme de pas q avec une zone morte de largeur $2q$:

$$Q_q(x) = \begin{cases} 0 & \text{si } |x| < q, \\ \text{signe}(x)(k + 1/2)q & \text{si } kq \leq |x| \leq (k + 1)q. \end{cases} \quad (3.2)$$

Cette quantification est représentée sur la figure 3.8. C'est la quantification couramment employée pour la compression en ondelettes. Falzon et Mallat ont montré théoriquement

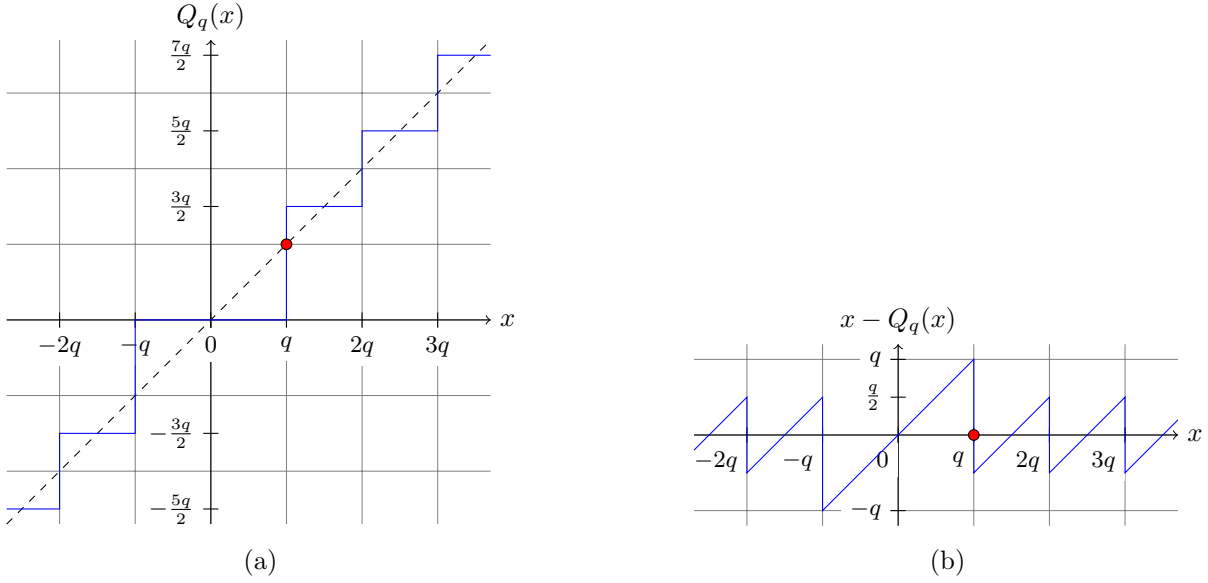


FIG. 3.8 – (a) Quantification scalaire quasi-uniforme à zone morte. (b) Erreur de quantification.

et numériquement dans [25] que ce choix de zone morte est proche du choix optimal pour minimiser la distorsion sur les coefficients d'ondelettes. Comme la distribution des coefficients de bandelettes est proche de la distribution des coefficients d'ondelettes (cf. figure 3.9 et section 3.4.2), cette quantification est aussi presque optimale pour les coefficients de bandelettes. On note $f_q^b = Q_q(f^b)$ la représentation quantifiée du bloc f dans la base b .

Compromis débit-distorsion

En vue de l'optimisation de la compression, le choix de la meilleure base $\mathcal{B}_b$ pour la représentation du bloc de coefficients d'ondelettes f est effectué par minimisation d'un critère de débit-distorsion. Une diminution du débit R ayant très souvent comme corollaire une augmentation de la distorsion D , un lagrangien est utilisé :

$$\mathcal{L}(f_q^b) = D(f_q^b) + \lambda R(f_q^b) \quad (3.3)$$

où $D(f_q^b)$ est l'erreur quadratique due à la quantification de la transformée en bandelettes f^b , $R(f_q^b)$ est une estimation du débit nécessaire au codage des coefficients quantifiés de f_q^b et λ est le multiplicateur de Lagrange optimisé pour la compression. L'indice b^* de la meilleure base $\mathcal{B}_{b^*}$ pour la représentation du bloc de coefficients d'ondelettes f parmi toutes les bases du dictionnaire $\mathcal{D} = \{\mathcal{B}_b\}_{b=0}^{N_{\mathcal{B}}}$ est donné par :

$$b^* = \arg \min_{b \in [0, N_{\mathcal{B}}]} \mathcal{L}(f_q^b) \quad (3.4)$$

Dans [67], Le Pennec et Mallat donnent une expression théorique du multiplicateur de Lagrange λ qui ne dépend que du pas de quantification utilisé :

$$\lambda = \frac{3}{4\gamma_0} q^2 \quad \text{avec} \quad \gamma_0 = 6, 5. \quad (3.5)$$

Nous reviendrons en détail sur le calcul de ce multiplicateur de Lagrange dans la section 3.3.

Calcul de la distorsion

Puisque toutes les transformées utilisées sont orthonormales, l'erreur quadratique due à la quantification peut être calculée dans le domaine transformé. Elle est donnée par :

$$D(f_q^b) = \sum_{m=1}^M |a_b[m] - Q_q(a_b[m])|^2.$$

Estimation du débit

Le débit $R(f_q^b)$ se décompose en deux parties : le débit $R_c(f_q^b)$ nécessaire au codage des coefficients quantifiés de la représentation f_q^b et le débit R_b nécessaire à la signalisation de la base de bandelettes sélectionnée parmi les $N_B + 1$ bases du dictionnaire. Cette information de signalisation est nécessaire au décodeur afin d'inverser correctement la transformée en bandelettes. Le débit $R(f_q^b)$ s'exprime donc par :

$$R(f_q^b) = R_c(f_q^b) + R_b.$$

Le terme $R_c(f_q^b)$ est estimé par le débit des coefficients de bandelettes quantifiés :

$$R_c(f_q^b) = - \sum_{m=1}^M \log_2 \Pr(Q_q(a_b[m])). \quad (3.6)$$

Les probabilités $\Pr(Q_q(a_b[m]))$ sont estimées à partir de l'histogramme des coefficients d'ondelettes dans chaque sous-bande de la transformée en ondelettes. En effet, les probabilités des coefficients de bandelettes ne peuvent pas être directement évaluées à cette étape du processus puisque les coefficients de bandelettes vont dépendre de la sélection de la meilleure représentation de chaque bloc. Or la sélection de la meilleure représentation sur chaque bloc, ne peut se faire qu'après l'estimation du débit $R(f_q^b)$ (cf. équations 3.3 et 3.4). D'autre part, nous verrons que près de la moitié des blocs n'est pas transformée en bandelettes. En conséquence, l'histogramme des coefficients d'ondelettes est une bonne estimation de l'histogramme des coefficients de bandelettes (cf. figure 3.9). Des résultats expérimentaux viendront appuyer cette affirmation dans la section 3.4.2.

Le second terme R_b est estimé par :

$$R_b = -\log_2 \Pr(b) \quad \text{avec} \quad \Pr(b) = \begin{cases} 0,5 & \text{si } b = 0, \\ 0,5/N_B & \text{si } b \in [1, N_B]. \end{cases} \quad (3.7)$$

La base identifiée par l'indice $b = 0$ est la base canonique. Cet indice permet d'identifier les blocs de coefficients d'ondelettes qui ne sont pas transformés en bandelettes. Les autres indices $b \in [1, N_B]$ identifient les bases de bandelettes utilisées pour transformer les autres blocs de coefficients d'ondelettes. La probabilité $\Pr(b)$ *a priori* affectée à l'indice $b = 0$ est plus forte que celles affectées aux autres indices correspondant aux transformées en

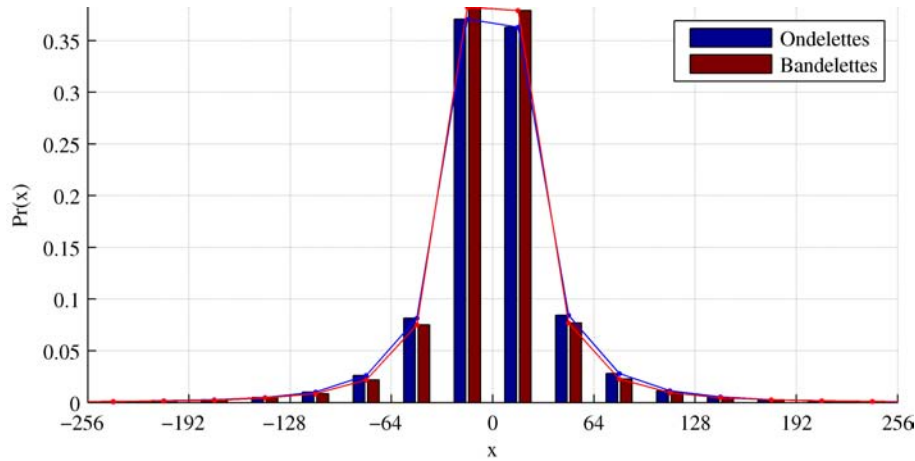


FIG. 3.9 – Histogrammes de la transformée en ondelettes et en bandelettes de la sous-bande HL_1 de l'image « PELICAN Toulouse 3 » représentée en annexe A.

bandelettes. En conséquence, le coût de signalisation *a priori* d'un bloc de coefficients d'ondelettes non transformé est plus faible que le coût de signalisation *a priori* d'un bloc transformé en bandelettes. Ce coût est de 1 bit dans le premier cas ($b = 0$) et de $1 + \log_2(N_B)$ dans les autres cas ($b \in [1, N_B]$). Cela influence le critère débit-distorsion $\mathcal{L}(f_q^b)$ en faveur des blocs de coefficients d'ondelettes non transformés. Les blocs non transformés sont alors plus probables que les blocs transformés dans une base de bandelettes donnée. Cela peut se constater sur l'exemple de la figure 3.6.

Il y a deux raisons pour favoriser les blocs non transformés. La première est que, dans la plupart des blocs, les coefficients d'ondelettes peuvent déjà être compressés efficacement. C'est le cas par exemple sur les régions homogènes. La seconde raison est que cela permet aussi de diminuer le coût de signalisation R_b par rapport à une distribution *a priori* uniforme sur le choix des bases de bandelettes. En effet, lorsqu'un indice est beaucoup plus probable que les autres, l'entropie de la source est plus faible. Par exemple, pour un nombre de $N_B = 15$ bases dans le dictionnaire de bandelettes, le coût moyen de R_b donné dans l'équation 3.7 est de

$$E[R_b] = \sum_{b=0}^{N_B} \Pr(b) (-\log_2(\Pr(b))) = 2,95 \text{ bits par bloc.}$$

Avec une distribution *a priori* uniforme, ce coût moyen est de 4 bits par bloc.

3.1.4 Algorithme de compression en bandelettes

Le schéma de compression en bandelettes est illustré sur la figure 3.10. L'image est d'abord transformée en ondelettes. Les blocs de 4×4 coefficients d'ondelettes sont ensuite transformés en bandelettes et quantifiés pour obtenir la transformée en bandelettes de l'image et la signalisation de la transformée utilisée sur chaque bloc. Enfin, les coefficients de bandelettes quantifiés et la signalisation sont codés entropiquement pour former le train binaire compressé.

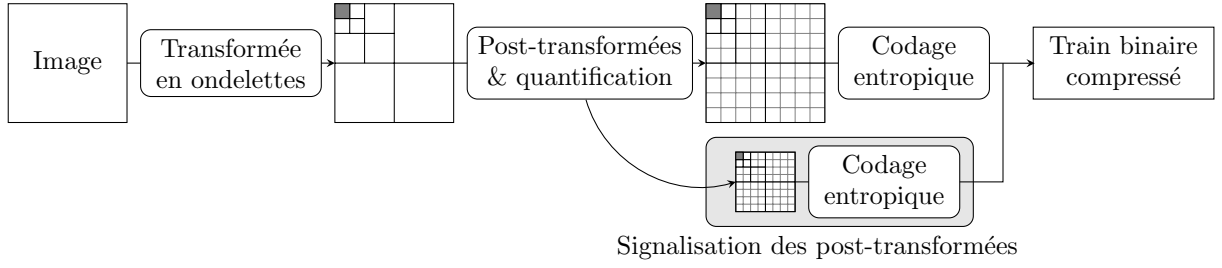


FIG. 3.10 – Schéma de compression par transformée en bandelettes par blocs. L'étape « Post-transformées & quantification » est représentée par le schéma de la figure 3.7.

Algorithme 1 : Compression d'image par transformée en bandelettes par blocs.

Entrées : Une image et un pas de quantification q

Sorties : Le train binaire compressé

Transformée en ondelettes de l'image ;

Quantification et codage entropique de la sous-bande LL ;

pour chaque sous-bande de la transformée en ondelettes exceptée la sous-bande LL **faire**

Calcul de l'histogramme de la sous-bande pour évaluer

les probabilités $\Pr(Q_q(a_b[m]))$ utilisées dans l'équation (3.6) ;

Partition de la sous-bande en blocs de 4×4 coefficients ;

pour chaque bloc f **faire**

pour chaque base $\mathcal{B}_b$ du dictionnaire $\mathcal{D}$ **faire**

Transformée du bloc f en utilisant la base de bandelettes $\mathcal{B}_b$ comme décrit dans l'équation (3.1) ;

Quantification des coefficients de bandelettes $\{a_b[m]\}_{m=1}^M$ de la représentation f^b avec le quantificateur scalaire de l'équation (3.2) ;

Évaluation du coût débit-distorsion $\mathcal{L}(f_q^b)$ défini dans l'équation (3.3) ;

fin

Mémorisation de la représentation f_q^{b*} qui minimise le critère débit-distorsion $\mathcal{L}(f_q^b)$;

Mémorisation de la signalisation b^* de la base utilisée pour obtenir cette représentation ;

fin

Codage entropique de la signalisation des bases utilisées sur chaque bloc de la transformée en bandelettes ;

Codage entropique des coefficients de bandelettes ;

fin

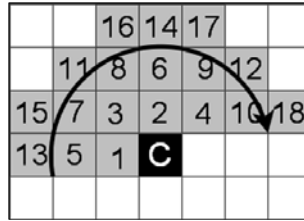


FIG. 3.11 – Définition d'un voisinage intra-bande dans le domaine ondelettes.

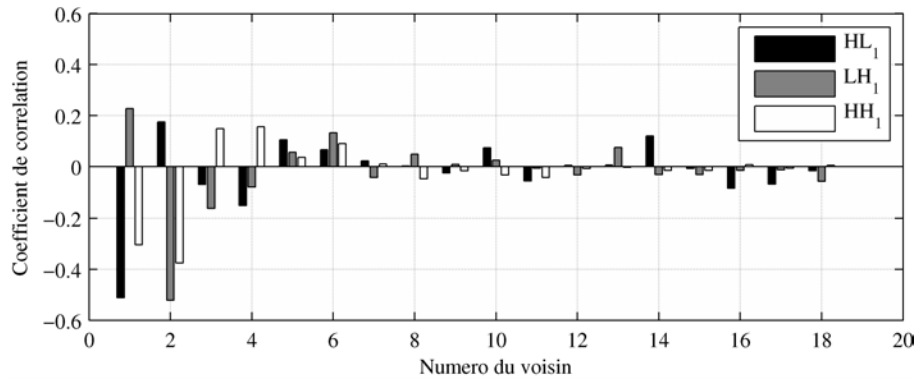


FIG. 3.12 – Coefficients de corrélation entre un coefficient d'ondelettes C et chacun de ses dix-huit plus proches voisins intra-bande.

L'algorithme 1 résume les étapes de la compression. Le codeur entropique utilisé est un codeur arithmétique adaptatif extrait de la QccPack [71].

3.2 Choix de la taille des blocs

Dans cette section, nous analysons quelle taille de blocs permet d'obtenir les meilleurs résultats en compression pour la transformée en bandelettes. Pour choisir une taille de blocs, nous commençons par étudier les dépendances entre coefficients d'ondelettes voisins. En effet, le but de la transformée est de réduire ces dépendances selon le principe du codage par transformée [20].

3.2.1 Corrélations entre coefficients d'ondelettes

Dans la section 2.1.2, nous avons identifié des corrélations intra- et inter- sous-bandes entre coefficients d'ondelettes voisins. Ici, nous nous intéressons aux corrélations intra-sous-bande. Ce sont celles qui peuvent être éliminées par des post-transformées par blocs. Dans la section 2.1.2, le voisinage intra- sous-bande considéré était limité aux quatre plus proches coefficients au dessus et à gauche du coefficient d'ondelettes C considéré (cf. figure 2.2). Nous élargissons ici ce voisinage intra-bande. Le nouveau voisinage considéré est schématisé sur la figure 3.11. Les voisins du coefficient C sont numérotés en fonction de la distance qui les sépare du coefficient C et de gauche à droite.

La figure 3.12 présente les coefficients de corrélation entre le coefficient d'ondelette

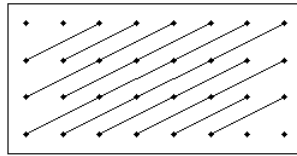


FIG. 3.13 – Configuration de groupements directionnels #2 pour un bloc de taille 4×8 .

C et chacun de ses dix-huit voisins. Plus la distance est grande entre deux coefficients d'ondelettes, plus les corrélations sont faibles. Ceci est vrai dans n'importe quelle sous-bande de la transformée en ondelettes. Au delà d'une distance supérieure à 3 pixels entre deux coefficients d'ondelettes (voisins du coefficient C ayant un numéro supérieur ou égal à 13 sur la figure 3.11), les corrélations sont presque nulles. On s'attend donc à ce que l'exploitation de corrélations sur des blocs de taille supérieure à 4×4 coefficients n'apporte que très peu de gain.

De plus, on constate sur la figure 3.12 que les corrélations sont différentes selon les sous-bandes comme cela avait été mentionné dans la section 2.1.2. Les corrélations entre voisins horizontaux numérotés 1 et 5 sont plus fortes dans les sous-bandes HL et les corrélations entre voisins verticaux numérotés 2 et 6 sont plus fortes dans les sous-bandes LH. Cela correspond aux effets du filtre d'ondelettes passe-haut. Cette différence entre sous-bandes peut être exploitée en modifiant la taille des blocs transformés par les bandelettes.

3.2.2 Extension des bases de bandelettes à des blocs de tailles diverses

Bien que peu de redondances entre coefficients d'ondelettes distants de plus de trois pixels aient été identifiées en dehors des zones correspondant aux contours des images, augmenter la taille des blocs peut permettre de diminuer le coût de la signalisation. En effet, plus les blocs sont grands, moins il y en a dans une sous-bande de taille fixée et donc moins il y a de signalisation sur les bases de bandelettes choisies à transmettre.

Plusieurs tailles de blocs ont été considérées : des blocs de taille 4×8 , 4×16 , 8×8 , 8×16 et 16×16 . Dans chaque cas, 12 configurations de groupements directionnels ont été définis. La configuration de groupements directionnels #2 pour un bloc de taille 4×8 est représentée sur la figure 3.13. Les nouvelles bases de bandelettes directionnelles sont construites à partir de ces groupements en utilisant les polynômes discrets de Legendre de façon similaire à la section 3.1.1. La base correspondant à la configuration de groupements de la figure 3.13 est représentée sur la figure 3.14.

Comme les corrélations entre coefficients d'ondelettes diffèrent selon les sous-bandes, pour tenir compte de l'orientation des sous-bandes, le dictionnaire de bases directionnelles est différent dans chaque sous-bande. Les blocs transformés ont une taille rectangulaire orientée verticalement, 8×4 par exemple, dans les sous-bandes HL. La taille des blocs est orientée horizontalement, 4×8 par exemple, dans les sous-bandes LH et les blocs sont carrés dans les sous-bandes HH.

Ces nouvelles bases de bandelettes exploitent alors avantageusement les corrélations le

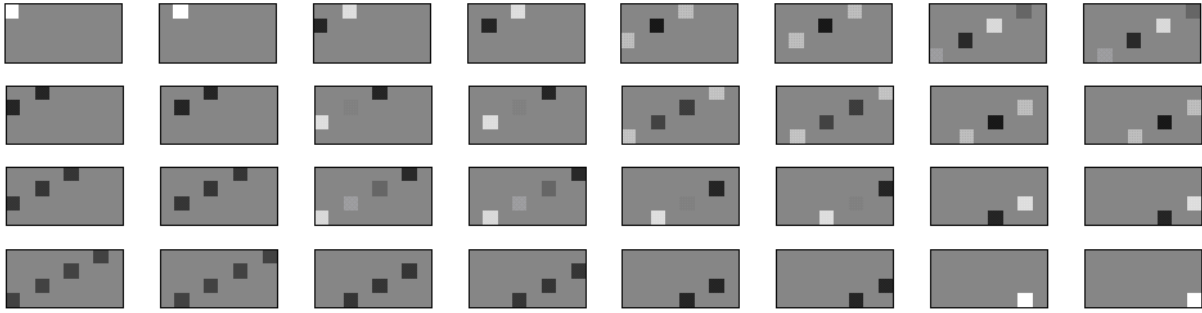


FIG. 3.14 – Base de bandelettes directionnelles #2 pour un bloc de taille 4×8 .

long des grand contours telles que celles identifiées à la fin de la section 2.1.2. Cependant, les corrélations entre coefficients d'ondelettes des proches voisins sont moins bien exploitées et les résultats en compression sont moins bons qu'avec des blocs de taille 4×4 . Nous avons constaté une perte moyenne de 0,3 dB en PSNR avec des blocs de taille 16×16 , de 0,12 dB avec des blocs de taille 8×8 et de moins de 0,05 dB avec des blocs de taille 4×8 .

De plus, plus la taille des blocs est grande, plus la dimension de la base est grande. Par exemple, pour des blocs de taille 4×8 , l'espace de travail est $\mathbb{R}^{32}$. Il y a donc plus de vecteurs à stocker dans la mémoire du codeur et plus de mémoire est consommée au moment de la transformée. Les transformées sont aussi plus longues à effectuer. Enfin, comme il y a moins de blocs, la transformée en bandelettes est moins parallélisable.

Au vu des résultats en compression, du coût de signalisation des indices des bases, des critères matériels et des corrélations résiduelles dans le domaine ondelettes, la meilleure taille des blocs transformés est 4×4 .

3.3 Sélection des meilleures bases de bandelettes par optimisation débit-distorsion

Pour effectuer une transformée en bandelettes efficace du point de vue de la compression, chaque bloc de coefficients d'ondelettes doit être transformé dans la base du dictionnaire de bandelettes qui permet d'optimiser la compression. Ainsi, le choix de la meilleure base de bandelettes sur chaque bloc de coefficients d'ondelettes est un problème d'optimisation débit-distorsion décrit par l'équation (3.3). Nous allons ici expliquer et discuter le choix du multiplicateur de Lagrange (3.5) effectué dans [67].

3.3.1 Le problème d'allocation débit-distorsion

Le problème d'allocation débit-distorsion est un problème d'optimisation non linéaire récurrent dans la littérature. On note S le vecteur d'allocation. Dans le cas de la recherche des bases de bandelettes optimales sur chacun des blocs numérotés i , le vecteur d'allocation est $S = (b_1, b_2, \dots, b_i, \dots, b_{N_b})$ où b_i est la base utilisée pour transformer le bloc f_i et N_b le nombre de blocs. En effet, avec un pas de quantification q fixé, le débit de chaque

bloc varie en fonction de la base de bandelettes sélectionnée. La fonction objectif est la distorsion $D(S)$. Elle est définie pour toute allocation S de l'ensemble $\mathcal{S}$ de tous les vecteurs d'allocation possibles. Comme le dictionnaire contient N_B bases, le cardinal de cet ensemble est $N_B^{N_b}$. La fonction de contrainte est le débit binaire $R(S)$ défini pour tout S de $\mathcal{S}$. Le problème est alors de trouver le vecteur d'allocation S qui minimise la distorsion $D(S)$ (l'objectif) sous la contrainte d'un débit cible R_T :

$$\begin{cases} \min_{S \in \mathcal{S}} D(S) \\ R(S) \leq R_T. \end{cases} \quad (3.8)$$

Théorème principal

Dans [72], Shoham et Gersho donnent un équivalent non-contraint au problème contraint décrit par les équations (3.8) en utilisant le théorème suivant :

Théorème. *Pour tout $\lambda \geq 0$, la solution $S^*(\lambda)$ au problème non-contraint*

$$\min_{S \in \mathcal{S}} D(S) + \lambda R(S) \quad (3.9)$$

est aussi la solution au problème contraint (3.8) avec la contrainte $R(S) \leq R(S^(\lambda))$.*

La preuve de ce théorème peut être trouvée dans [72]. Elle est retranscrite en annexe D. Ce théorème nous dit que pour tout $\lambda \geq 0$, il existe un problème contraint dont la solution est identique à la solution au problème non-contraint. Il ne garantit pas l'existence d'une solution au problème contraint (3.8). Cependant, si $R(S^*(\lambda)) = R_T$, alors $S^*(\lambda)$ est la solution au problème contraint.

Pour simplifier l'écriture, on note $R^*(\lambda) = R(S^*(\lambda))$. En faisant varier λ de 0 à l'infini, on crée des ensembles de solutions $S^*(\lambda)$ et de contraintes $R^*(\lambda)$. La question est de savoir si l'une des valeurs $R^*(\lambda)$ coïncide avec le débit cible R_T et de trouver la valeur λ correspondante. Ainsi, on ne recherche plus directement le vecteur d'allocation optimal $S^*(\lambda)$ mais on l'obtient par l'intermédiaire du multiplicateur de Lagrange λ . L'allocation optimale $S^*(\lambda)$ est définie lorsque le multiplicateur de Lagrange λ est fixé. L'objectif maintenant est de trouver le multiplicateur de Lagrange λ tel que le débit donné par l'allocation optimale $R^*(\lambda)$ soit le plus proche possible du débit R_T que l'on s'est fixé. On passe donc de la recherche d'un vecteur optimal $S^*(\lambda)$ à la recherche d'un scalaire optimal λ^* . Le problème est « simplifié ».

Le problème d'allocation dans la littérature

Dans [72], Shoham et Gersho proposent un algorithme d'allocation de débit à des blocs de coefficients étant donné un ensemble arbitraire de quantificateurs. L'allocation est optimisée pour la performance mesurée à partir des données. Ils répondent à la question : *Quel quantificateur appliquer à chacun des blocs afin de minimiser la distorsion totale sous la contrainte du débit fixé ?*

Dans [73], Ramchandran et Vetterli étendent les algorithmes de Shoham et Gersho pour la recherche de la décomposition optimale d'une image dans des bases de paquets

d'ondelettes étant donné un ensemble de quantificateurs différents applicables à chaque paquet. L'allocation est aussi optimisée pour la performance mesurée à partir des données. Ils répondent à la question : *Quelle est la meilleure décomposition en paquets d'ondelettes étant donné l'ensemble des quantificateurs possibles* afin de minimiser la distorsion totale sous la contrainte du débit fixé ?

Dans sa thèse [38], Parisot cherche à optimiser la zone morte des quantificateurs utilisés sur chacune des sous-bandes de la transformée en ondelettes. Pour cela, l'allocation est calculée à partir de modélisations de la distribution des coefficients d'ondelettes des sous-bandes en gaussiennes généralisées. Il utilise ensuite des formules du débit et de la distorsion sous hypothèses asymptotique (haute résolution) ou non asymptotique (bas et moyen débits). Il répond aux questions : *Quel quantificateur appliquer à chacune des sous-bandes* afin de minimiser la distorsion ou le débit sous la contrainte du débit ou respectivement de la distorsion ?

Dans l'étude expérimentale menée dans la section 3.3.3, nous cherchons l'allocation optimale, c'est-à-dire, les bases de bandelettes optimales sur chacun des blocs de la transformée en ondelettes d'une image avec un pas de quantification q fixé. Cet ensemble est optimisé à partir de la performance mesurée sur les données. Nous répondons à la question : *Quelle est la meilleure base de bandelettes sur chaque bloc* afin de minimiser la distorsion sous la contrainte du débit fixé ? Comme ce débit fixé peut être atteint soit en faisant varier le paramètre λ et donc l'allocation $S^*(\lambda)$, soit en faisant varier le pas de quantification q , nous cherchons les couples (λ, q) optimaux. Commençons par des analyses théoriques.

3.3.2 Calcul théorique du multiplicateur de Lagrange optimal

Dans [67], Le Pennec et Mallat ne fixent pas de débit cible R_T . Cependant, ils cherchent l'allocation optimale $S^*(\lambda, q)$ pour le pas de quantification q utilisé i.e. celle qui minimise le coût débit-distorsion défini par le lagrangien :

$$J(\lambda, q) = D(S, q) + \lambda R(S, q). \quad (3.10)$$

Pour résoudre ce problème, Le Pennec et Mallat cherchent à relier les paramètres λ et q donnant l'allocation optimale $S^*(\lambda, q)$ par une formule. L'allocation optimale $S^*(\lambda, q)$ ne dépend alors plus que d'un seul paramètre, le paramètre q par exemple. Une fois le paramètre optimal q trouvé, on obtient le paramètre λ optimal par l'intermédiaire de cette formule et ainsi l'allocation optimale $S^*(\lambda, q)$. En faisant l'hypothèse qu'une formule reliant λ et q existe lorsque l'allocation est optimale et en faisant l'hypothèse que l'allocation a été fixée et est optimale, le lagrangien (3.10) s'écrit :

$$J(q) = D(q) + \lambda R(q). \quad (3.11)$$

Sous hypothèse de l'existence d'un minimum de la fonction (3.11), lorsque ce lagrangien est minimum, alors la dérivée de cette fonction s'annule et on a :

$$\frac{\partial D}{\partial q} = -\lambda \frac{\partial R}{\partial q}. \quad (3.12)$$

Hypothèse bas et moyen débits

La résolution de l'équation (3.12) nécessite une expression du débit R et de la distorsion D en fonction du pas de quantification q .

Dans [67], Le Pennec et Mallat utilisent une estimation du débit tirée de [25]. Dans [25], Falzon et Mallat ont montré théoriquement et numériquement que pour la plupart des images transformées dans une base de cosinus discrets ou d'ondelettes, le débit binaire moyen R nécessaire au codage des coefficients quantifiés est quasiment proportionnel au nombre moyen de coefficients quantifiés non nuls M .

$$R \approx \gamma_0 M. \quad (3.13)$$

Cette relation est valable sous hypothèse de compression à bas et moyen débits, c'est-à-dire lorsque la proportion de coefficients non nuls est comprise entre 2^{-7} et 2^{-1} . Dans [25], il a été montré numériquement que $\gamma_0 = 6,5$ sur la base des quatre images « Lenna », « Barbara », « Mandrill », et « Aerial ». Le Pennec et Mallat font l'hypothèse que cette relation reste valable pour les coefficients de bandelettes.

Nous avons vérifié cette formule expérimentalement pour la transformée en bandelettes de plusieurs *images d'observation de la Terre avec une résolution de 12 bits par pixel*. Nous avons obtenu des valeurs stables de γ_0 lorsque la proportion de coefficients non nuls est inférieure à 2^{-3} , c'est-à-dire à bas débit. Selon les images γ_0 vaut entre 6,5 et 8. Pour des proportions de coefficients non nuls supérieures, les valeurs de γ_0 décroissent lentement. Lorsque la proportion de coefficients non nuls est de 2^{-1} , elles sont comprises entre 5,5 et 7 selon les images. On considère donc que l'approximation $\gamma_0 \approx 6,5$ reste valable dans le cas de la compression en bandelettes d'images 12 bits à bas et moyen débits.

Pour estimer la variation de la distorsion à bas débit, Le Pennec et Mallat [67] prennent en compte le quantificateur scalaire utilisé. C'est le quantificateur Q_q de l'équation (3.2) représenté sur la figure 3.8. On considère un coefficient d'ondelettes d'amplitude x proche de q . La distorsion D induite par une variation infinitésimale de q est de $q^2/4$ si le coefficient x est quantifié à $3q/2$. Elle est de q^2 si le coefficient x est quantifié à 0 (cf. figure 3.8(b)). En conséquence, mettre un coefficient à 0 par variation infinitésimale du pas de quantification q revient à augmenter la distorsion de $3q^2/4$ et à diminuer le nombre M de coefficients non nuls. La variation de la distorsion en fonction de la variation du pas de quantification est donc approchée par :

$$\frac{\partial D}{\partial q} \approx -\frac{3q^2}{4} \frac{\partial M}{\partial q}. \quad (3.14)$$

Finalement en combinant les équations (3.12), (3.13) et (3.14) on obtient :

$$-\frac{3q^2}{4} \frac{\partial M}{\partial q} \approx -\lambda \gamma_0 \frac{\partial M}{\partial q}.$$

D'où l'approximation (3.5) du multiplicateur de Lagrange λ en fonction du pas de quantification q .

Hypothèse haut débit

Dans sa thèse [38], Parisot utilise l'hypothèse haute résolution pour calculer une allocation optimale. Son but est de trouver les pas de quantification à affecter à chacune des sous-bandes de la transformée en ondelettes d'une image afin d'optimiser le débit et la distorsion totale. Pour un λ donné, il calcule un ensemble de pas de quantification optimaux q_i pour chaque sous-bande i de la transformée en ondelettes. Nous avons adapté son raisonnement au cas de la recherche des meilleures bases de bandelettes pour un pas de quantification unique q et nous obtenons une autre expression du multiplicateur de Lagrange λ en fonction de q .

Pour résoudre l'équation (3.12), il faut une expression de la distorsion D et du débit R en fonction du pas de quantification q . Sous hypothèse d'une quantification haute résolution, on a une expression simple de l'erreur quadratique moyenne en fonction du pas de quantification [74] :

$$D \approx \frac{q^2}{12} \quad (3.15)$$

De même sous hypothèse haute résolution, le débit R s'exprime en fonction du pas de quantification [75] :

$$R \approx h(X) - \log_2 q \quad (3.16)$$

où $h(X)$ est l'entropie différentielle de la source X , l'image dans notre cas. Les preuves de (3.15) et (3.16) sont retranscrites en annexe D.

En insérant les expressions de la distorsion (3.15) et du débit (3.16) sous hypothèse haute résolution dans l'équation (3.12), nous obtenons :

$$\frac{2q}{12} \approx -\lambda \left(-\frac{1}{q \ln(2)} \right),$$

d'où l'expression du multiplicateur de Lagrange λ en fonction du pas de quantification q sous l'hypothèse d'une quantification scalaire uniforme haute résolution :

$$\lambda \approx \frac{\ln(2)}{6} q^2. \quad (3.17)$$

Discussion

Le Pennec et Mallat ont fourni une approximation du multiplicateur de Lagrange λ en fonction du pas de quantification q permettant de trouver une solution optimale $S^*(\lambda, q)$ à l'équation (3.10) pour tout pas de quantification q . Cette expression du multiplicateur de Lagrange λ (3.5) est valable à bas et moyen débits c'est-à-dire lorsque la proportion de coefficients non nuls est comprise entre 2^{-7} et 2^{-1} . Elle est aussi issue de deux approximations successives. La première est l'estimation du débit (3.13). La seconde est la variation de la distorsion en fonction de la variation du pas de quantification (3.14).

Sous hypothèse haute résolution, nous avons obtenu une expression très similaire en suivant le raisonnement décrit par Parisot. En effet, sous hypothèse bas et moyen débits comme sous hypothèse haut débit on a :

$$\lambda \propto q^2,$$

avec comme coefficient multiplicatif à bas et moyen débits :

$$\frac{3}{4\gamma_0} = 0,11538 \quad \text{pour } \gamma_0 = 6, 5,$$

et comme coefficient multiplicatif à haut débit :

$$\frac{\ln(2)}{6} = 0,11552.$$

Les deux expressions sont donc identiques bien qu'elles proviennent d'approximations et d'hypothèses différentes.

Dans la section suivante, nous montrons la validité de ces deux expressions, obtenues avec des hypothèses contradictoires, par une recherche empirique des couples (λ, q) optimaux pour la compression d'un ensemble d'images d'observation de la Terre. Cette étude a été publiée dans [70].

3.3.3 Recherche empirique du multiplicateur de Lagrange optimal

La méthode d'optimisation est issue de l'article de Shoham et Gersho [72]. Nous l'avons adaptée au cas des bandelettes et elle permet de donner des valeurs optimales pour le multiplicateur de Lagrange qui intervient dans le choix de la meilleure base de bandelettes. Deux formules théoriques identiques du multiplicateur de Lagrange ont été données dans la section précédente. Cette méthode permet d'en donner une expression plus précise. De plus, nous montrons que les résultats en compression obtenus avec cette nouvelle expression sont très proches des résultats optimaux.

Algorithmes d'allocation débit-distorsion

Deux algorithmes sont détaillés par Ramchandran et Vetterli [73] pour obtenir l'allocation optimale. On rappelle que dans les travaux de Ramchandran et Vetterli [73] comme dans les travaux de Shoham et Gersho [72], l'allocation est optimisée pour la performance mesurée à partir des données. Le but ici est de trouver le multiplicateur de Lagrange optimal par une recherche quasi-exhaustive. Pour des raisons de complexité, il est préférable de calculer un multiplicateur de Lagrange quasi-optimal à l'aide d'une formule simple telles que les formules théoriques trouvées précédemment. C'est pour cette raison que nous déduisons une formule empirique à partir des résultats donnés par la recherche exhaustive des multiplicateurs de Lagrange optimaux sur plusieurs images et pour différents pas de quantification. Nous effectuons cette recherche sur des images d'observation de la Terre 12 bits.

La première étape consiste à calculer tous les couples débit-distorsion associés à chaque constituant élémentaire. Dans le cas des bandelettes, il faut commencer par calculer le débit et la distorsion associés à toutes les représentations f_{iq}^b possibles de chaque bloc f_i , c'est-à-dire dans chaque base de bandelettes $\mathcal{B}_b$ du dictionnaire D et quantifiées par tous les quantificateurs Q_q possibles.

Algorithme 2 : Point-Optimal($\lambda, \{(r_i^b, d_i^b)\}_{b,i}$).

Entrées : Le multiplicateur de Lagrange λ et les points débit-distorsion $\{(r_i^b, d_i^b)\}_{b,i}$ de chaque bloc f_i transformé dans chaque base b

Sorties : L'allocation optimale $S^*(\lambda)$, le débit $R^*(\lambda)$ et la distorsion $D^*(\lambda)$

pour chaque bloc f_i faire

// Sélection de la base donnant le point de fonctionnement optimal
pour le bloc f_i

$b_i^*(\lambda) = \arg \min_{b \in [1, N_B]} (j_i^b(\lambda))$ avec $j_i^b(\lambda) = d_i^b + \lambda r_i^b$;

fin

$S^*(\lambda) \leftarrow \{b_i^*(\lambda)\}_{i=1}^{N_i}$;

$R^*(\lambda) \leftarrow \sum_{i=1}^{N_i} r_{b_i^*(\lambda), i}$;

$D^*(\lambda) \leftarrow \sum_{i=1}^{N_i} d_{b_i^*(\lambda), i}$

Fixons le pas de quantification q . Tous les blocs sont quantifiés par la même quantification. Le choix d'un pas de quantification différent pour chaque bloc serait trop coûteux en terme de débit de signalisation. Chaque bloc de coefficients d'ondelettes f_i est transformé dans chacune des N_B bases du dictionnaire de bandelettes $\mathcal{D}$. Pour chaque représentation f_{iq}^b , $b \in [0, N_B]$ du bloc numéro i , on calcule la distorsion d_i^b et une estimation du débit r_i^b nécessaire au codage de f_{iq}^b . Le calcul de la distorsion et l'estimation du débit sont détaillés dans la section 3.1.3. La méthode d'estimation du débit utilisée n'est pas critique ici puisque ce sont les positions relatives des points débit-distorsion (r_i^b, d_i^b) , l'un par rapport à l'autre, qui sont importantes et non les performances absolues. L'algorithme 2 permet de calculer l'allocation optimale $S^*(\lambda)$ étant donnée n'importe quelle valeur positive du multiplicateur de Lagrange λ . Il permet de résoudre l'équation suivante :

$$S^*(\lambda) = \arg \min_{S \in \mathcal{S}} (D(S) + \lambda R(S))$$

avec λ fixé. Le débit optimal $R^*(\lambda) = R(S^*(\lambda))$ et la distorsion optimale $D^*(\lambda) = D(S^*(\lambda))$ pour ce multiplicateur de Lagrange λ découlent de l'allocation optimale $S^*(\lambda)$ trouvée.

Cet algorithme utilise la propriété d'additivité des coûts $j_i^b(\lambda) = d_i^b + \lambda r_i^b$ de chaque bloc f_i . Ces coûts sont minimisés un par un sur chaque bloc afin de minimiser le coût total. La figure 3.15 illustre la minimisation de ce coût $j_i^b(\lambda)$ sur le bloc f_i . Les croix matérialisent les couples débit-distorsion associés à chaque représentation de ce bloc dans les différentes bases de bandelettes. Les droites sont de pente $-\lambda$. Le coût $j_i^b(\lambda)$ se lit à l'intersection de la droite de pente $-\lambda$ passant par le point (r_i^b, d_i^b) avec l'axe des ordonnées d_i . La base optimale pour la représentation de ce bloc et pour le paramètre λ fixé est celle qui permet de minimiser ce coût. L'allocation optimale pour l'ensemble des blocs et pour le paramètre λ fixé est l'ensemble des bases qui permettent de minimiser les coûts $j_i^b(\lambda)$ sur chaque bloc f_i .

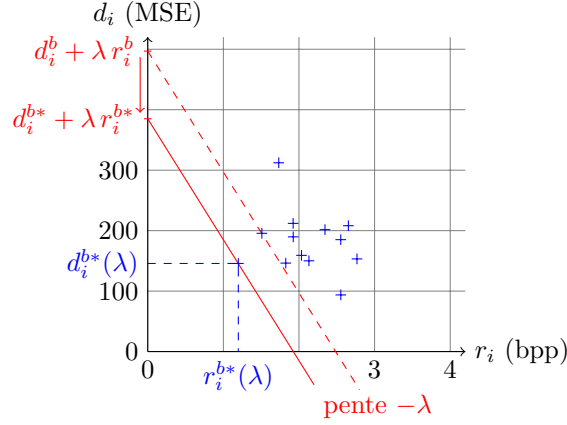
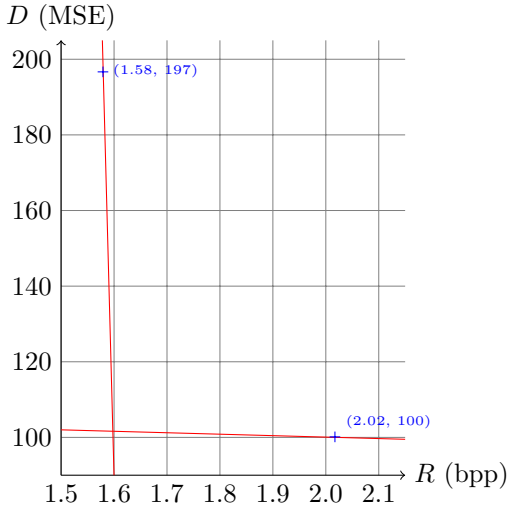


FIG. 3.15 – Interprétation graphique du processus de sélection de la meilleure base sur un bloc. Chaque croix correspond à un point débit-distorsion (r_i^b, d_i^b) de la représentation quantifiée f_{iq}^b du bloc f_i . La meilleure base pour une pente $-\lambda$ donnée est trouvée par minimisation de la quantité $d_i^b + \lambda r_i^b$. Cette quantité est lue à l'intersection de la droite de pente $-\lambda$ avec l'axe des ordonnées.

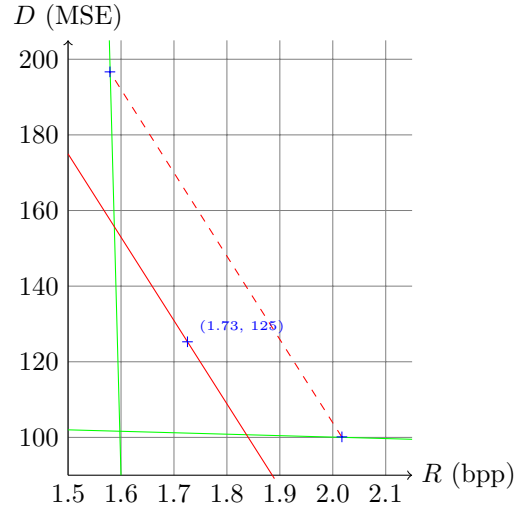
L'algorithme 3 est un algorithme itératif qui permet finalement de déterminer le paramètre λ à utiliser sur une image afin d'optimiser la distorsion D avec un débit inférieur au débit cible R_T donné en entrée. Le paramètre λ est déterminé par dichotomie. A chaque itération, l'algorithme 2 est utilisé afin de déterminer le débit et la distorsion pour le paramètre temporaire λ_{next} . L'algorithme est initialisé avec deux valeurs extrêmes du paramètre λ_{next} . Un paramètre λ_{next} très grand correspond à une minimisation du débit sans tenir compte de la distorsion. Un paramètre λ_{next} très petit correspond à une minimisation de la distorsion sans tenir compte du débit. L'algorithme converge vers le plus petit paramètre λ qui permet de respecter le débit maximum imposé R_T .

Le terme $+\epsilon$ qui intervient dans le calcul de λ_{next} est un nombre positif proche de 0 qui permet d'augmenter la pente $-\lambda_{\text{next}}$ afin de s'assurer que le prochain point de fonctionnement débit-distorsion sélectionné est toujours celui ayant le plus petit débit parmi l'ensemble des points de fonctionnement débit-distorsion optimaux pour ce paramètre λ . En effet, il peut exister plus d'un point de fonctionnement débit-distorsion optimal pour un paramètre λ_{next} donné. Par exemple, à la fin des itérations, lorsque l'algorithme a convergé, les points de fonctionnement $(R^*(\lambda_l), D^*(\lambda_l))$ et $(R^*(\lambda_u), D^*(\lambda_u))$ sont optimaux pour le paramètre $\lambda_{\text{next}} - \epsilon$. Le terme $+\epsilon$ qui intervient dans le calcul de λ_{next} permet de sélectionner le plus petit débit optimal $R^*(\lambda_u)$. En effet on a toujours $R^*(\lambda_u) \leq R_T \leq R^*(\lambda_l)$.

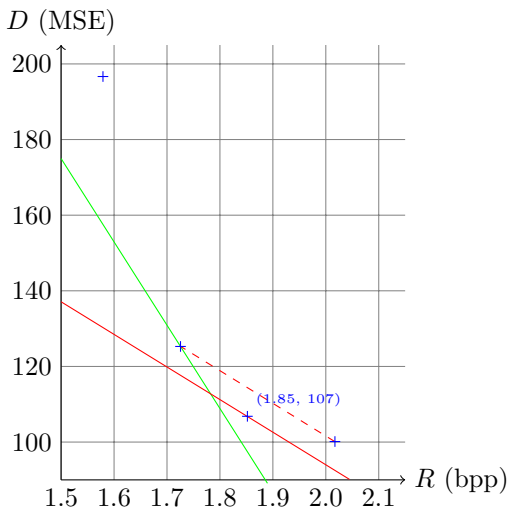
La figure 3.16 illustre l'initialisation et trois itérations de l'algorithme dichotomique 3. Cet algorithme converge très rapidement vers la consigne R_T .



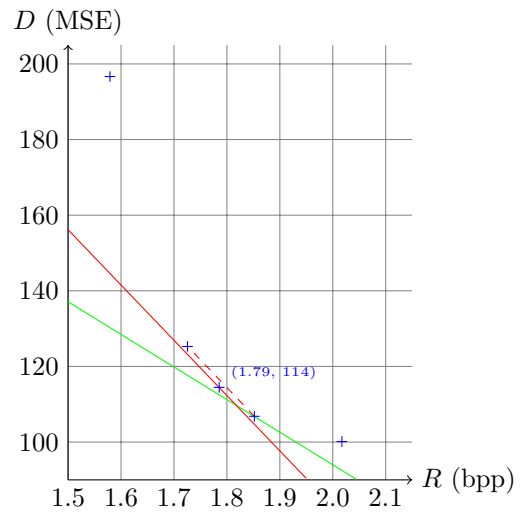
(a) Initialisation : Deux points de fonctionnement extrêmes sont trouvés pour les paramètres $\lambda_l = \epsilon$ et $\lambda_u = 1/\epsilon$. Ils correspondent respectivement à la distortion minimale et au débit minimum.



(b) Première itération: Un nouveau point de fonctionnement est trouvé pour le nouveau paramètre $\lambda_{\text{next}} = 220$



(c) Seconde itération: $\lambda_{\text{next}} = 86$



(d) Troisième itération: $\lambda_{\text{next}} = 146$

FIG. 3.16 – Exemple de trois itérations de l'algorithme dichotomique 3 vers le débit cible fixé à $R_T = 1.8$ bpp ici. Les croix correspondent aux points de fonctionnement optimaux trouvés pour les différentes valeurs de λ . Ces points de fonctionnement sont calculés dans l'algorithme 2 par $R^*(\lambda) = \sum_{i=1}^{N_i} r_i^*(\lambda)$ et $D^*(\lambda) = \sum_{i=1}^{N_i} d_i^*(\lambda)$ à partir des données de la transformée en bandelettes d'une image satellite.

Algorithme 3 : Algorithme dichotomique.

Entrées : Le débit cible R_T et les points débit-distorsion $\{(r_i^b, d_i^b)\}_{b,i}$ de chaque bloc f_i transformé dans chaque base b

Sorties : Le multiplicateur de Lagrange optimal λ^* pour le débit cible R_T

```

// Initialisation: calcul des points extrêmes
 $\lambda_l \leftarrow \epsilon$ ;  $\lambda_u \leftarrow 1/\epsilon$ ;
 $(R^*(\lambda_l), D^*(\lambda_l)) \leftarrow \text{Point-Optimal}(\lambda_l, \{(r_i^b, d_i^b)\}_{b,i})$ ;
 $(R^*(\lambda_u), D^*(\lambda_u)) \leftarrow \text{Point-Optimal}(\lambda_u, \{(r_i^b, d_i^b)\}_{b,i})$ ;

// Calcul de la nouvelle pente entre les deux points de
// fonctionnement entourant le débit cible
 $\lambda_{\text{next}} \leftarrow \left| \frac{D^*(\lambda_l) - D^*(\lambda_u)}{R^*(\lambda_l) - R^*(\lambda_u)} \right| + \epsilon$ ;

// Exécution de l'algorithme 2
 $(R^*(\lambda_{\text{next}}), D^*(\lambda_{\text{next}})) \leftarrow \text{Point-Optimal}(\lambda_{\text{next}}, \{(r_i^b, d_i^b)\}_{b,i})$ ;

// Itérations jusqu'au débit cible
tant que  $R^*(\lambda_{\text{next}}) \neq R^*(\lambda_u)$  faire
    si  $R^*(\lambda_{\text{next}}) > R_T$  alors
        |  $\lambda_l \leftarrow \lambda_{\text{next}}$ 
    sinon
        |  $\lambda_u \leftarrow \lambda_{\text{next}}$ 
    fin
     $\lambda_{\text{next}} \leftarrow \left| \frac{D^*(\lambda_l) - D^*(\lambda_u)}{R^*(\lambda_l) - R^*(\lambda_u)} \right| + \epsilon$ ;
     $(R^*(\lambda_{\text{next}}), D^*(\lambda_{\text{next}})) \leftarrow \text{Point-Optimal}(\lambda_{\text{next}}, \{(r_i^b, d_i^b)\}_{b,i})$ ;
fin

 $\lambda^* \leftarrow \lambda_{\text{next}}$ ;

```

Application aux bandelettes

Dans le cas des bandelettes, le débit cible R_T n'est pas fixé. Nous n'utilisons donc pas l'algorithme 3. Cependant, nous cherchons l'allocation optimale $S^*(\lambda, q)$ pour le pas de quantification q utilisé.

L'algorithme 2 sert à sélectionner les bases de bandelettes optimales étant donné un pas de quantification q et une valeur du multiplicateur de Lagrange λ . Les points débit-distorsion $\{(r_i^b, d_i^b)\}_{b,i}$ de chaque bloc f_i transformé dans chaque base b sont déterminés pour cette valeur de q . L'algorithme 2 permet donc d'obtenir l'allocation optimale $S^*(\lambda, q)$.

Pour simplifier l'écriture, on utilise la notation $R^*(\lambda, q) = R^*(S^*(\lambda, q))$ et $D^*(\lambda, q) = D^*(S^*(\lambda, q))$. La courbe débit-distorsion optimale pour un pas de quantification q fixé est calculée en faisant varier λ de 0 à $+\infty$. Cette courbe est un ensemble de points débit-distorsion $(R^*(\lambda, q), D^*(\lambda, q))$ optimaux pour le paramètre q fixé. Des exemples de courbes

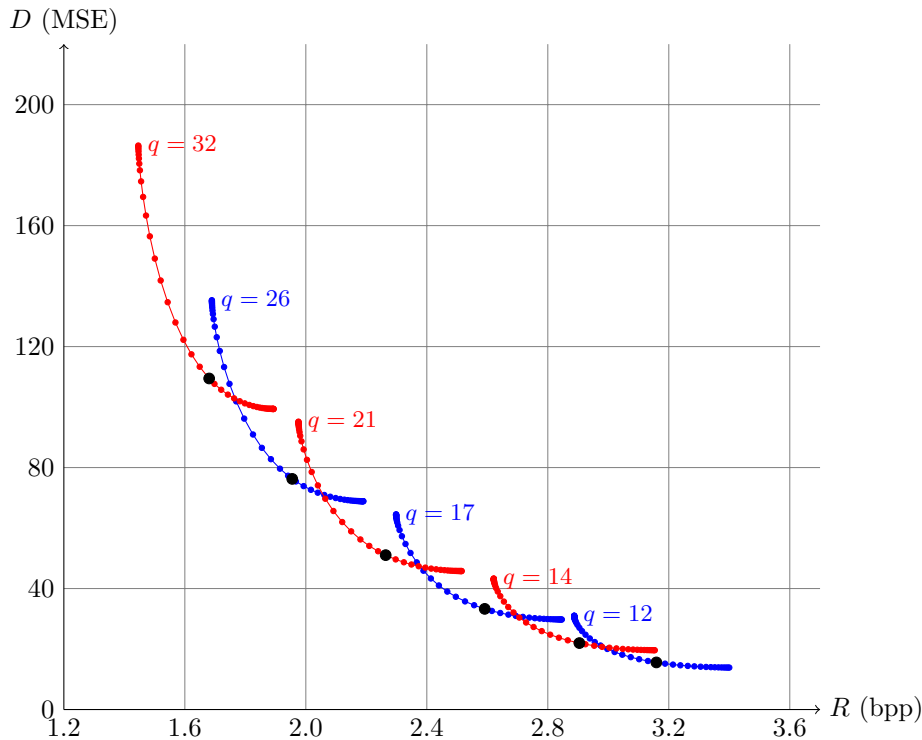


FIG. 3.17 – Courbes débit-distorsion calculées sur une image avec différents pas de quantification q . Ces courbes ont été obtenues avec des valeurs de λ variant de 1 à 5000. Chaque courbe est un ensemble de points $\{(R^*(\lambda, q), D^*(\lambda, q))\}_\lambda$ optimaux pour un pas de quantification q . Les gros points noirs sont obtenus avec la valeur de λ calculée à partir de la formule de l'équation (3.5).

optimales sont donnés sur la figure 3.17 pour différents pas de quantification variant de 12 à 32. Ce sont toutes des courbes obtenues par transformées en bandelettes de la même image.

Afin d'obtenir la courbe débit-distorsion optimale pour n'importe quel pas de quantification, le même procédé est répété avec plusieurs pas de quantification différents. La figure 3.18 est un exemple des résultats obtenus. Elle a été obtenue avec un plus grand nombre de pas de quantification. L'enveloppe inférieure de ces courbes est la courbe débit-distorsion optimale pour n'importe quel pas de quantification q . Elle est représentée en rouge sur la figure 3.18. Sur la figure 3.17, on peut aussi constater que les points obtenus pour la valeur λ calculée à partir de la formule théorique de l'équation (3.5) sont toujours proches de l'optimal, c'est-à-dire proche de l'enveloppe convexe contenant les différentes courbes.

Afin de vérifier la validité des formules théoriques pour le calcul du multiplicateur de Lagrange, on cherche à minimiser les coûts débit-distorsion globaux $J(\lambda, q)$ de l'équation (3.10) en fonction de différents pas de quantification q et de différents multiplicateurs de Lagrange λ . On ne considère plus que des allocations optimales $S^*(\lambda, q)$ trouvées grâce à l'algorithme 2 et on note :

$$J^*(\lambda, q) = D^*(\lambda, q) + \lambda R^*(\lambda, q).$$

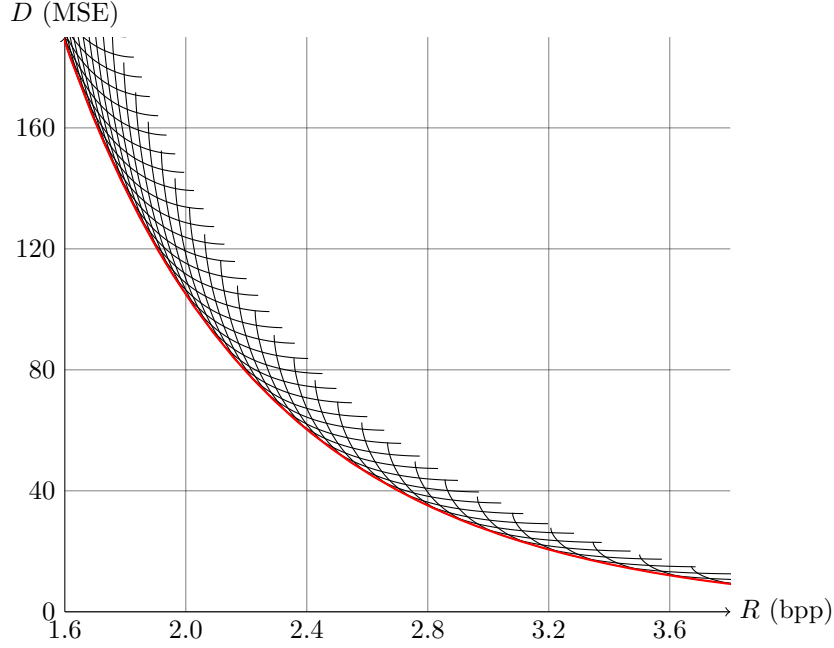


FIG. 3.18 – Courbes débit-distorsion calculées sur une image avec différents pas de quantification q . Les points optimaux sont représentés en rouge.

Étant donné un multiplicateur de Lagrange λ , le coût $J^*(\lambda, q)$ pour le pas de quantification q optimal est inférieur à tout autre coût $J^*(\lambda, q')$ calculé pour n'importe quel autre pas de quantification q' :

$$\forall q', \quad D^*(\lambda, q) + \lambda R^*(\lambda, q) \leq D^*(\lambda, q') + \lambda R^*(\lambda, q'). \quad (3.18)$$

Ainsi, à partir des points débit-distorsion calculés pour obtenir les courbes tracées sur les figures 3.17 et 3.18, il est possible de trouver quelles sont les valeurs de q optimales étant donnés différents paramètres λ . Ce sont les valeurs $q^*(\lambda)$ pour lesquelles le critère débit-distorsion est minimum :

$$q^*(\lambda) = \arg \min_q (D^*(\lambda, q) + \lambda R^*(\lambda, q)).$$

Le même procédé a été utilisé sur les six images de l'ensemble de test représenté en annexe A. Sur la figure 3.19, les pas de quantification optimaux obtenus $q^*(\lambda)$ ont été tracés en fonction de λ . On constate que les valeurs optimales de $q^*(\lambda)$ ne dépendent quasiment pas de l'image. De plus, on constate que λ est proportionnel à q^2 mais avec un coefficient multiplicatif différent de celui donné dans l'équation (3.5) ou (3.17). Les courbes de $q^*(\lambda)$ tracées sur la figure 3.20 sont mieux approchées par la droite d'équation :

$$\log_2(\lambda) = 2 \log_2(q) - 2,7$$

d'où la nouvelle formule empirique du multiplicateur de Lagrange :

$$\lambda \approx 0,15 q^2. \quad (3.19)$$

La figure 3.21 montre que le rapport λ/q^2 est relativement stable pour une gamme de

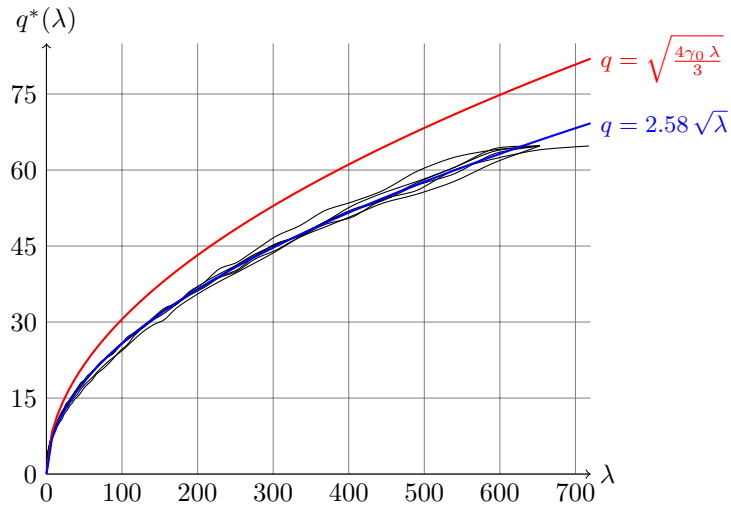


FIG. 3.19 – Tracés des pas de quantification optimaux en fonction de différentes valeurs du multiplicateur de Lagrange λ et sur différentes images.

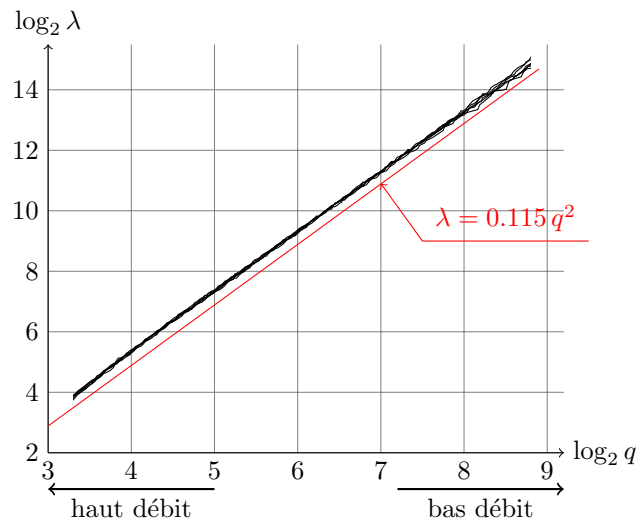


FIG. 3.20 – Multiplicateurs de Lagrange optimaux obtenus sur six images d'observation de la Terre comparés à la courbe théorique $\lambda = 0,115 q^2$.

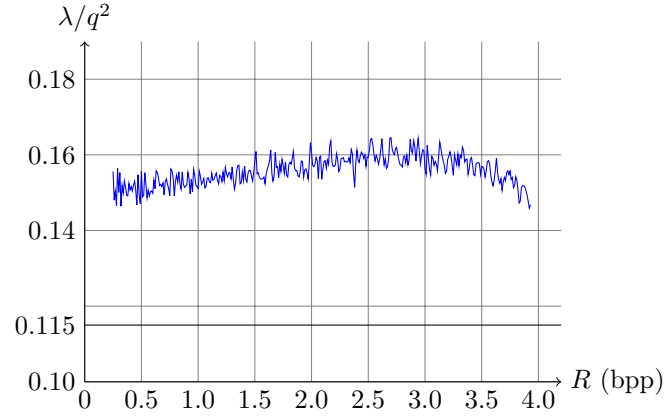


FIG. 3.21 – Rapports entre les multiplicateurs de Lagrange optimaux λ et le carré des pas de quantification q^2 moyennés sur six images d’observation de la Terre. Ce rapport est comparé à la valeur théorique 0,115 sur une large gamme de débits.

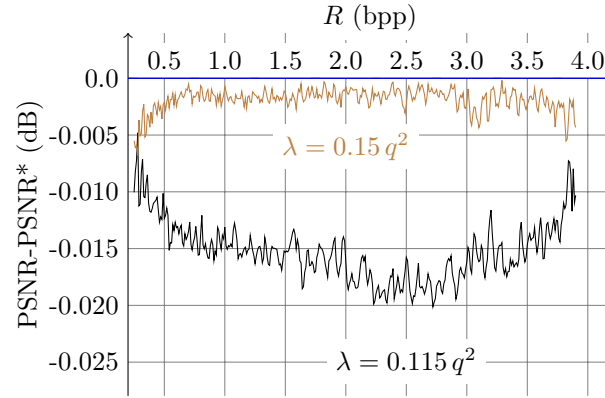


FIG. 3.22 – Comparaison des pertes en PSNR par rapport aux résultats obtenus par recherche quasi-exhaustive du meilleur multiplicateur de Lagrange λ . Les résultats obtenus avec la formule empirique et la formule théorique sont comparés. Ces résultats sont moyennés sur six images d’observation de la Terre.

débit allant de 0,2 bpp à 3,5 bpp.

Sur la figure 3.22 les résultats en compression obtenus par recherche quasi-exhaustive du meilleur multiplicateur de Lagrange λ sont comparés aux résultats obtenus lorsque ce paramètre est calculé avec la formule théorique $\lambda = 0,115 q^2$ ou avec la formule empirique $\lambda = 0,15 q^2$. Les résultats obtenus avec la formule empirique ne sont que très peu inférieurs aux résultats donnés par le recherche exhaustive : -0,005 dB au maximum. De plus, la formule empirique donne toujours de meilleurs résultats que les formules théoriques (3.5) ou (3.17). Le gain de la formule empirique par rapport à la formule théorique est d’environ 0,015 dB pour des débits entre 1 bpp et 3,5 bpp.

Cette amélioration des performances est peu significative mais permet de confirmer les formules empirique et théorique pour le calcul d’un multiplicateur de Lagrange quasi-optimal. On peut noter que, même si le multiplicateur de Lagrange λ utilisé n’est pas tout à fait optimal, les points débit-distorsion calculés, eux, sont optimaux pour le multiplicateur

de Lagrange utilisé. C'est le cas de tous les points calculés pour obtenir la figure 3.17. L'impact de l'erreur commise sur le multiplicateur de Lagrange λ utilisé reste donc faible puisque la sélection des bases de bandelettes sur chaque bloc est toujours effectuée par le processus d'optimisation décrit dans l'algorithme 2.

3.3.4 Conclusion

Nous avons vu que le problème de la sélection des meilleures bases de bandelettes est un problème similaire au problème d'allocation débit-distorsion récurrent dans le domaine de la compression. Ce problème peut être résolu par la minimisation d'un lagrangien et se résume à la recherche d'un multiplicateur de Lagrange optimal. Nous avons vu que sous hypothèse bas et moyen débits, Le Pennec et Mallat ont donné une formule théorique de ce multiplicateur de Lagrange optimal. Étrangement, en partant d'une hypothèse différente, l'hypothèse haute-résolution, nous avons obtenu la même formule avec la même constante multiplicative après application numérique. Nous avons vérifié ces formules expérimentalement en menant une recherche quasi-exhaustive du multiplicateur de Lagrange optimal. Les résultats obtenus montrent que les formules théoriques sont très proches de l'optimal. Comme il n'est pas souhaitable d'effectuer une recherche quasi-exhaustive du multiplicateur de Lagrange en ligne, nous avons déterminé une formule empirique à partir des résultats obtenus. Cette formule est la même que la formule théorique à la constante multiplicative près. Le calcul du multiplicateur de Lagrange par la formule empirique permet d'améliorer très légèrement les résultats en compression par rapport à la formule théorique et ces résultats sont très proches des résultats optimaux. L'utilisation de la formule empirique (3.19) pour le calcul du multiplicateur de Lagrange λ en fonction du pas de quantification q utilisé est donc recommandée pour la compression des images satellites.

3.4 Études expérimentales des performances de la transformée en bandelettes

3.4.1 Analyse de la sélection des bases directionnelles

La figure 3.23 présente les résultats de l'analyse en bandelettes de la sous-bande LH_1 d'une image. Cette portion d'image a été spécialement choisie parce qu'elle contient des contours bien distincts qui se retrouvent visuellement dans la transformée en ondelettes. C'est un cas assez rare qui n'est pas représentatif des images d'observation de la Terre. En général, les contours ne sont pas aussi bien marqués et les segments droits sont plus courts. Le pas de quantification a été fixé à $q = 32$. Les bases de bandelettes ont été sélectionnées par minimisation du critère débit-distorsion de l'équation (3.3) et sont représentées sur chaque bloc de 4×4 coefficients d'ondelettes. Les directions correspondant aux bases de bandelettes directionnelles ne sont pas en accord avec les directions des contours dans l'image. Plusieurs raisons peuvent expliquer ce résultat. Sur des blocs de taille 4×4 pris indépendamment les uns des autres, il est difficile de distinguer une direction même pour



(a) Image PELICAN

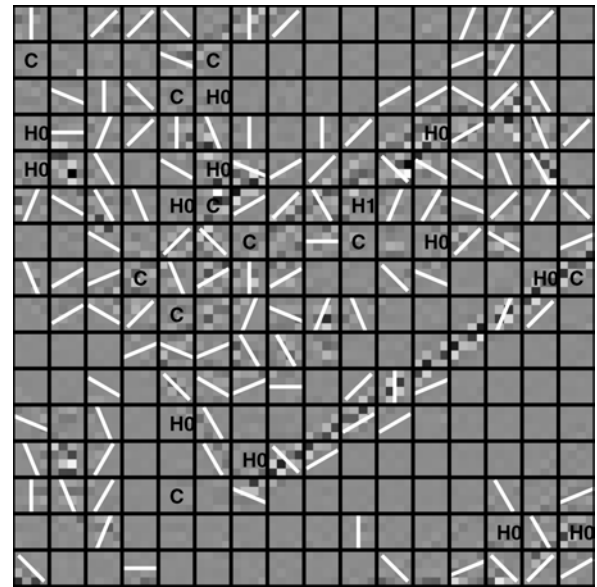

(b) Sous-bande LH_1 de la transformée en ondelettes et bases de bandelettes sélectionnées sur chaque bloc

FIG. 3.23 – Image PELICAN de résolution 25 cm et les bases de bandelettes sélectionnées sur la sous-bande LH_1 de la transformée en ondelettes.

l'oeil humain. La taille des blocs est trop petite. De plus, l'information géométrique n'est pas prise en compte dans le processus de sélection de la meilleure base de bandelettes. La base choisie dépend de la distorsion engendrée par la quantification. Elle n'est donc pas toujours liée à la direction naturelle des coefficients d'ondelettes.

Par ailleurs, une base de bandelettes est susceptible d'exploiter des corrélations dans la direction pour laquelle elle a été construite ou dans la direction perpendiculaire. En effet, il y a deux sources de régularités dans la transformée en ondelettes d'une image : celles introduites par les contours et celles introduites par le filtrage en ondelettes. Les bases directionnelles ont été créées afin d'exploiter les régularités le long des contours. En fait, elles peuvent exploiter les deux types de régularités : les régularités basses fréquences le long des contours et les oscillations perpendiculaires aux contours qui sont dues au filtrage d'ondelettes passe-haut sur les contours. Par exemple, les vecteurs de la partie rouge de la base d'ondelettes représentées sur la figure 3.24 sont efficaces pour exploiter les régularités basses fréquences le long d'un contour dans la direction #2 définie sur la figure 3.1. Les vecteurs de la partie bleue de cette base, eux sont efficaces pour exploiter les oscillations qui seraient perpendiculaires à un contour dans la direction #8 définie sur la figure 3.1.

Dans le cas des images satellites, l'information géométrique ne peut donc pas être extraite des bases de bandelettes utilisées. Même dans le cas du contour net de la figure 3.23, la direction des bases de bandelettes est rarement en accord avec la direction du contour. La combinaison d'une analyse géométrique et d'une compression efficace semble être difficile à effectuer dans le domaine ondelettes. Ces observations montrent aussi qu'il

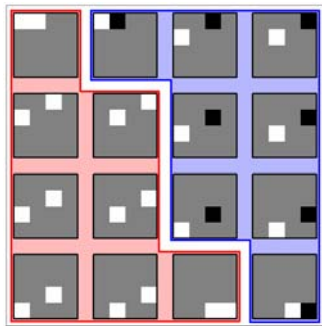


FIG. 3.24 – Base directionnelle #2. Les vecteurs de la partie rouge permettent d'exploiter les régularités le long d'un contour dans la direction #2. Les vecteurs de la partie bleue permettent d'exploiter les oscillations perpendiculaires à un contour dans la direction #8.

Il y a peu de relations entre les bases de bandelettes sélectionnées sur des blocs adjacents. Les indices des bases sélectionnées peuvent donc être codés indépendamment.

3.4.2 Estimation du débit des coefficients de bandelettes

Nous avons vu que, pour la sélection de la meilleure base de bandelettes en terme de débit-distorsion, le débit $R_c(f_q^b)$ est estimé à partir des histogrammes des sous-bandes de la transformée en ondelettes. Pourtant, pour un pas de quantification q donné, on observe expérimentalement que le kurtosis d'une sous-bande transformée en bandelettes est supérieur au kurtosis des coefficients d'ondelettes de cette sous-bande. Le kurtosis, ou coefficient d'aplatissement de Pearson, est une mesure de la forme de la densité de probabilité d'une variable aléatoire. Plus cette densité de probabilités est « plate », plus le kurtosis est faible. Le kurtosis β_2 est donné par :

$$\beta_2 = \frac{\mu_4}{\sigma^4}$$

où μ_4 est le moment centré d'ordre 4 et σ l'écart type de la distribution. Par exemple pour la sous-bande HL_1 de la transformée en bandelettes de l'image PLEIADES « Port de Bouc » représentée en annexe A, le kurtosis vaut 44.9. Il vaut 30.2 pour la sous-bande HL_1 de la transformée en ondelettes de cette même image. La transformée en bandelettes permet donc une compaction de l'énergie. L'entropie des coefficients de bandelettes est plus faible que l'entropie des coefficients d'ondelettes. C'est ce qu'on observe en comparant les résultats reportés dans les tableaux 3.2 et 3.3. Ces résultats ont été obtenus pour la compression en ondelettes et en bandelettes de l'image « Port de Bouc » avec un pas de coefficient $q = 32$. L'entropie de la transformée en bandelettes est plus faible que l'entropie de la transformée en ondelettes : 1,584 bpp contre 1,675 bpp. En revanche, le débit total de la transformée en bandelettes est augmenté par le coût de la signalisation des bases. Le débit total de la transformée en bandelettes de 1.794 bpp est alors supérieur au débit de la transformée en ondelettes qui est de 1.714 bpp. Cette augmentation du débit est compensée par une augmentation de la qualité (50.04 dB avec la transformée en ondelettes contre 51.17 dB avec la transformée en bandelettes).

H_0	R	PSNR
1.675 bpp	1.714 bpp	50.04 dB

TAB. 3.2 – Résultats de la compression en ondelettes de l'image PLEIADES « Port de Bouc » avec le pas de quantification $q = 32$: entropie H_0 , débit R après codage arithmétique adaptatif et PSNR.

$\hat{R}_c$	H_0	R_c	$R = R_c + R_b$	PSNR
1.589 bpp	1.584 bpp	1.627 bpp	1.794 bpp	51.17 dB

TAB. 3.3 – Résultats de la compression en bandelettes de l'image PLEIADES « Port de Bouc » avec le pas de quantification $q = 32$: entropie estimée $\hat{R}_c$ à partir des histogrammes des coefficients d'ondelettes, entropie vraie H_0 , débit vrai R_c après codage arithmétique adaptatif des coefficients de bandelettes, débit total vrai R incluant le coût de la signalisation R_b et PSNR.

Les résultats en compression de plusieurs images satellite avec le codeur arithmétique adaptatif nous ont montré que $\hat{R}_c$ est une bonne estimation du débit vrai R_c . Dans le tableau 3.3, l'entropie estimée $\hat{R}_c$ est comparée à l'entropie vraie H_0 et à débit vrai R_c . L'entropie estimée à partir des histogrammes des coefficients d'ondelettes $\hat{R}_c$ est très proche de l'entropie vraie H_0 et le débit moyen R_c est relativement bien estimé. L'erreur n'est que de 2,3 %.

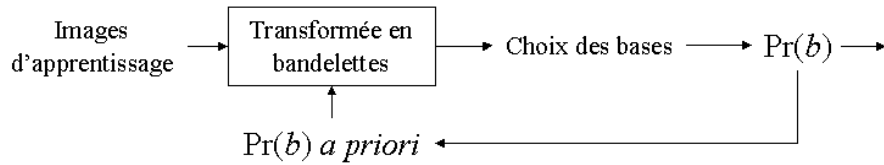
3.4.3 Estimation du débit de signalisation

Dans l'équation (3.7) pour l'estimation du débit R_b de la signalisation, les probabilités d'utilisation des différentes bases du dictionnaire de bandelettes sont fixées *a priori*. Ces probabilités *a priori* peuvent être ajustées hors ligne par un processus d'apprentissage dépendant du dictionnaire de bases $\mathcal{D}$ utilisé, où en ligne, adaptativement pendant la transformée en bandelettes.

Processus d'apprentissage hors ligne

Dans l'équation (3.7), les probabilités *a priori* $\Pr(b)$ sur le choix des bases de bandelettes ont été fixées arbitrairement pour l'estimation du coût de codage R_b . Cependant, en pratique, à la fin de la transformée en bandelettes d'une image, on observe que les probabilités des bases effectivement sélectionnées sont différentes des probabilités *a priori* que l'on a fixées. L'estimation du coût R_b est donc biaisée ce qui peut entraîner un mauvais choix de bases de bandelettes pour certains blocs.

Afin d'éviter ces erreurs, nous proposons un apprentissage hors-ligne de ces probabilités $\Pr(b)$. Pour cela, le schéma d'apprentissage illustré sur la figure 3.25 est utilisé. On commence par effectuer la transformée en bandelettes des images d'un ensemble d'apprentissage avec les probabilités *a priori* $\Pr(b)$ fixées comme dans l'équation (3.7). Après la transformée en bandelettes, on estime les probabilités de chaque base b , en comptant


 FIG. 3.25 – Processus d'apprentissage hors ligne des probabilités *a priori* $\Pr(b)$.

le nombre de blocs transformés dans chacune des bases. Une nouvelle transformée en bandelettes est alors effectuée en utilisant ces nouvelles probabilités *a priori* $\Pr(b)$. De plus, comme chaque sous-bande est codée indépendamment, on distingue les probabilités $\Pr(b)$ dans les différentes sous-bandes. Après 7 itérations, les probabilités obtenues n'évoluent plus. Par exemple, pour un pas de quantification $q = 32$ correspondant à un débit d'environ 2 bpp pour des images 12 bits et sur l'ensemble d'apprentissage d'images d'observation de la Terre, la probabilité $\Pr(b = 0)$ vaut 0,36 sur la sous-bande HL_1 , 0,34 sur la sous-bande LH_1 et 0,70 sur la sous-bande HH_1 . Ces probabilités sont différentes de celles données *a priori*. Cependant, l'utilisation de ces nouvelles probabilités *a priori* pour la compression en bandelettes de l'ensemble de test d'images d'observation de la Terre ne permet d'augmenter les résultats que de 0,05 dB en moyenne.

Apprentissage adaptatif en-ligne

Pour avoir une estimation encore plus précise du coût de codage R_b , nous proposons de l'évaluer adaptativement au cours du codage. Comme la signalisation, c'est-à-dire les symboles b identifiant la base utilisée sur chaque bloc, est encodée grâce à un codeur arithmétique adaptatif, l'idée est d'utiliser l'adaptabilité de ce codeur. Ainsi, la sélection de la base sur le bloc f_i dépendra des bases qui ont été sélectionnées au moment du codage des blocs précédant le bloc f_i . Cela implique que le codage du symbole b de signalisation soit effectué dès que la base a été choisie, c'est-à-dire dès que la minimisation du critère débit-distorsion (3.3) a été effectuée sur un bloc.

Le codeur arithmétique n'émet pas un mot binaire à chaque fois qu'un symbole b est codé. Il n'est donc pas possible d'accéder directement au coût de codage R_b . Ce coût de codage est alors estimé par l'entropie calculée à partir des probabilités issues de l'état courant du codeur arithmétique adaptatif. En effet, le codeur arithmétique adaptatif fait évoluer des tables de probabilités concernant chaque symbole pouvant être codé. A la fin du codage entier d'images de taille 1024×1024 , nous avons obtenu des écarts de moins de 8 bits entre la longueur du train binaire en sortie du codeur arithmétique et le coût de codage R_b estimé de cette façon. L'estimation de R_b est donc excellente.

Cette solution donne des résultats en compression corrects mais on ne maîtrise pas la base de bandelettes qui est la plus souvent sélectionnée. Cela peut être la base $b = 2$ par exemple ou n'importe quelle autre base et non la base canonique $b = 0$. De plus, ce n'est

Algorithme 4 : Apprentissage adaptatif du coût de signalisation.

Entrées : Les probabilités initiales $\text{Pr}(b)$ pour $b \in [0, N_{\mathcal{B}}]$

Sorties : Le train binaire compressé

```

pour chaque sous-bande faire
    // Initialisation du codeur arithmétique
    pour chaque symbole  $b \in [0, N_{\mathcal{B}}]$  faire
        |  $\text{freq}(b) \leftarrow \text{Pr}(b) \times \sqrt{n}$  ;
    fin
     $\text{cum\_freq} \leftarrow \text{Pr}(b) \times \sqrt{n}$  ;
    pour chaque bloc  $f$  faire
        pour chaque base  $b \in [0, N_{\mathcal{B}}]$  faire
            // Évaluation adaptative du coût de signalisation  $R_b$ 
             $P_b \leftarrow \text{freq}(b) \times \text{cum\_freq}$  ;
             $R_b \leftarrow -\log_2(P_b)$  ;
            Transformée du bloc  $f$  dans la base  $\mathcal{B}_b$  ;
            Évaluation du coût débit-distorsion  $\mathcal{L}(f_q^b)$  ;
        fin
        Sélection de la meilleure représentation  $f_q^{b^*}$  ;
        Codage arithmétique du symbole  $b^*$  ;
        // Adaptation du codeur arithmétique
         $\text{freq}(b^*) \leftarrow \text{freq}(b^*) + 1$  ;
         $\text{cum\_freq} \leftarrow \text{cum\_freq} + 1$  ;
    fin
    Codage des coefficients de bandelettes ;
fin

```

pas toujours la même base qui est la plus souvent sélectionnée dans toutes les sous-bandes. En effet, le processus mis en œuvre ici est « auto-satisfait ». Prenons un exemple.

Au début du processus, toutes les bases $b \in [0, N_{\mathcal{B}}]$ sont équiprobables. La base $b = 2$ est choisie pour la transformée en bandelettes du premier bloc de coefficients d'ondelettes. Les probabilités $\text{Pr}(b)$ sont mises à jour et la probabilité $\text{Pr}(b = 2)$ est supérieure aux autres. Le coût R_b calculé pour $b = 2$ est donc inférieur aux autres. Ainsi, grâce au critère débit-distorsion (3.3), le deuxième bloc a plus de chances d'être transformé dans la base $b = 2$. Si tel est le cas, cela renforce encore la probabilité $\text{Pr}(b = 2)$. Finalement, la base de bandelettes qui est la plus souvent sélectionnée dépend des premiers blocs transformés.

Cependant, nous préférons que la base $b = 0$ reste la plus probable. Pour cela, les probabilités initiales dans le codeur arithmétique sont modifiées et on leur affecte les valeurs utilisées pour les probabilités *a priori* $\text{Pr}(b)$ de l'équation (3.7). Il faut cependant veiller à ce que le poids affecté à cette initialisation soit raisonnable. Si le poids de l'initialisation est trop grand, alors le processus n'est plus adaptatif. S'il est trop faible, alors l'initialisation n'a plus d'importance et on ne maîtrise pas la base de bandelettes qui est la

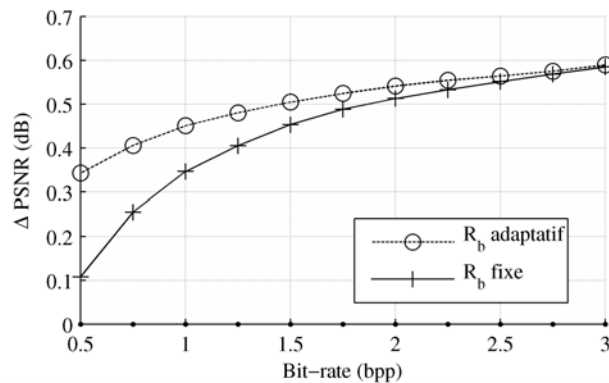


FIG. 3.26 – Performances de la compression avec un dictionnaire ne contenant que la base de Hadamard, avec et sans estimation adaptative du coût de la signalisation R_b relativement aux ondelettes.

plus souvent sélectionnée. Nous avons donc choisi un poids d'initialisation égal à la racine carré du nombre de blocs de coefficients d'ondelettes dans la sous-bande transformée $\sqrt{n}$. Nous avons modifié l'algorithme de compression en bandelettes pour prendre en compte ces modifications. L'algorithme 4 a été simplifié pour plus de clarté. Les détails du codage arithmétique peuvent être trouvés dans [76]. Avec cette méthode d'initialisation, les résultats en compression obtenus sont similaires à ceux obtenus sans cette initialisation, mais la base $b = 0$ reste la plus probable.

Dans le cas du dictionnaire de 15 bases de bandelettes, les performances en compression ne sont augmentées que de 0,05 dB en moyenne. Néanmoins, cette méthode d'estimation des probabilités présente l'avantage d'être adaptative. Nous n'aurons donc pas besoin de ré-apprendre les probabilités *a priori* à affecter à chaque base, à chaque fois que l'on change les bases du dictionnaire ou que l'on en retire. Une petite erreur au moment de l'initialisation des probabilités *a priori* aura peu d'impact puisqu'elle sera corrigée par l'adaptation. Au contraire, dans le cas où les probabilités *a priori* $\Pr(b)$ sont fixes, une petite erreur peut conduire à de mauvais résultats en termes de compression. Par exemple, les performances en compression obtenues avec un dictionnaire ne contenant que la base de Hadamard (cf. chapitre 4) sont tracées sur la figure 3.26. Dans un cas, les probabilités *a priori* sont fixes et valent $\Pr(b = 0) = 0,5$ et $\Pr(b = 1) = 0,5$. Dans l'autre cas, le coût de la signalisation R_b est estimé adaptativement. A bas débit, les résultats sont améliorés jusqu'à 0,2 dB sans avoir à effectuer un processus d'apprentissage.

Codage adaptatif contextuel de la signalisation

Pour réduire le coût de la signalisation R_b , il est aussi possible d'exploiter les dépendances entre les bases sélectionnées sur des blocs voisins. En effet, on constate que sur les régions homogènes des images, quelle que soit la sous-bande, la base sélectionnée est généralement la base canonique $b = 0$: il n'y a pas de forts coefficients d'ondelettes dans le bloc, il n'est donc pas transformé en bandelettes. En plus de ces dépendances intra-sous-bandes, il existe des dépendances inter-sous-bandes entre les bases sélectionnées sur

les blocs. Par exemple, la base canonique $b = 0$ peut être souvent sélectionnée pour les blocs de coefficients d'ondelettes correspondant à la même région homogène de l'image mais dans des sous-bandes différentes.

Ces dépendances peuvent être quantifiées en termes d'information mutuelle. Nous utilisons la mesure d'information mutuelle relative $I_r(X, Y)$. Elle est définie dans l'équation (2.1). L'information mutuelle relative représente la proportion d'information qui peut-être économisée sur le codage de Y si X est déjà connu.

Pour la compression des images de l'ensemble de test, le pas de quantification a été fixé à $q = 32$. Dans les sous-bandes HL_1 et LH_1 , l'entropie des bases sélectionnées est d'environ 3,5 bits par bloc. L'information mutuelle relative intra- sous-bande est de 2,24% et l'information mutuelle relative inter- sous-bandes est de 2,22%. Sur la sous-bande HH_1 , l'entropie des bases sélectionnées est d'environ 1,8 bits par bloc. L'information mutuelle relative inter- sous-bandes avec les sous-bandes HL_1 et LH_1 est d'environ 1,8%.

Il est possible de prendre en compte ces dépendances inter- sous-bandes par un codage arithmétique contextuel avec une modélisation contextuelle adaptée. Cependant, comme le débit nécessaire au codage de la signalisation est faible devant le débit nécessaire au codage des coefficients et comme les dépendances inter- sous-bandes entre les bases sélectionnées sont faibles, les performances en compression ne sont que très peu améliorées.

3.4.4 Performances du critère de sélection des bases

Afin d'évaluer les performances du critère de sélection des bases seul, c'est-à-dire sans les performances dues aux bases de bandelettes, nous construisons un dictionnaire de bases orthonormales aléatoires. Ainsi, les performances observées ne seront pas dues aux propriétés des bases de bandelettes, mais à la capacité du critère de sélection à choisir la meilleure base dans le dictionnaire afin de maximiser les performances en compression. Notons qu'un dictionnaire de bases aléatoires n'a clairement pas d'intérêt dans le cadre de la compression. Il n'est utilisé ici que dans un but d'évaluation des performances du processus de sélection de bases qui est inclus dans le processus de compression global.

Une base aléatoire est construite par orthogonalisation QR de $M = 16$ vecteurs générés aléatoirement dans l'espace $\mathbb{R}^{16}$. Ces vecteurs aléatoires sont composés de $M = 16$ coefficients tirés aléatoirement dans une distribution uniforme sur l'intervalle $[-1, 1]$. Un exemple de base obtenue est représenté sur la figure 3.27.

Les performances en compression obtenues avec un dictionnaire de $N_B = 1$ base aléatoire et un dictionnaire de $N_B = 15$ bases aléatoires sont comparées aux performances de la transformée en ondelettes sur la figure 3.28. Comme expliqué dans la section 3.1.3, certains blocs de coefficients d'ondelettes ne sont pas transformés.

Sur la figure 3.28, on peut voir que les performances obtenues avec un dictionnaire de 15 bases aléatoires sont meilleures que les performances de la transformée en ondelettes. Au débit de 0,5 bpp, le gain en PSNR est de 0,22 dB. Il approche les 0,7 dB à 3 bpp. De même avec un dictionnaire ne contenant qu'une seule base aléatoire, les performances sont meilleures que celles de la transformée en ondelettes. Cependant le gain en PSNR est toujours inférieur à 0,22 dB pour des débits inférieurs à 3 bpp. Cela montre qu'il est

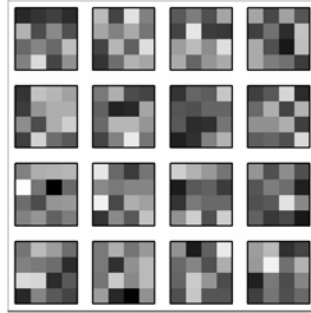


FIG. 3.27 – Base orthonormale aléatoire dans $\mathbb{R}^{16}$.

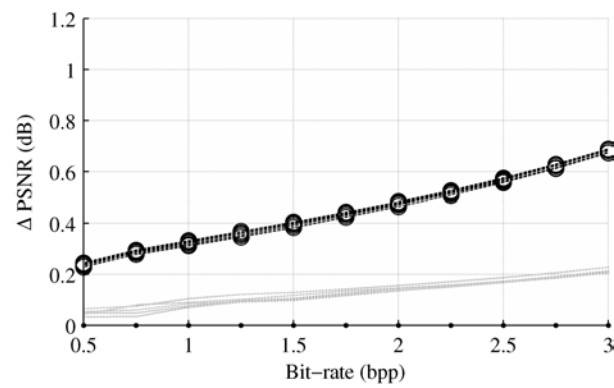


FIG. 3.28 – Différences entre les performances en compression obtenues avec la transformée en ondelettes et les performances en compression obtenues avec 5 dictionnaires constitués de 15 bases aléatoires (traits pointillés marqués d'un rond) et 5 dictionnaires constitués d'une base aléatoire (traits fins).

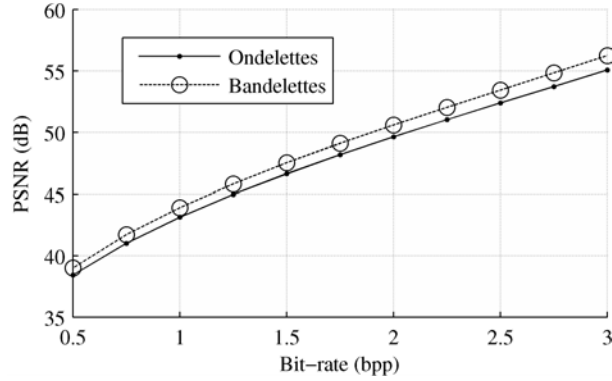


FIG. 3.29 – Résultats de la compression de la transformée en ondelettes et de la transformée en bandelettes moyennés sur l'ensemble des images de test.

possible de réduire la distorsion même en utilisant des bases aléatoires. Nous allons voir que les bases de bandelettes permettent d'améliorer encore ces résultats.

3.4.5 Performances des bases de bandelettes

Comparons maintenant les résultats obtenus avec les transformées en bandelettes et en ondelettes suivies du codage arithmétique adaptatif des coefficients. Dans chaque cas, la transformée en ondelettes est effectuée sur trois niveaux de décomposition comme recommandé dans la norme du CCSDS [29]. Ce nombre de niveaux de décomposition a été choisi parce qu'il représente un bon compromis entre les performances en compression (débit et distorsion) et la complexité calculatoire [36]. La transformée en bandelettes est effectuée comme décrit dans l'algorithme 1. Le multiplicateur de Lagrange vaut $\lambda = 0,15q^2$ comme cela a été expliqué dans la section 3.3. Le débit affecté au codage de la signalisation est évalué adaptativement comme cela a été discuté dans la section 3.4.3. Les images de l'ensemble de test sont compressées successivement en utilisant divers pas de quantification q . Les résultats obtenus sont reportés sur la figure 3.29. Les résultats en PSNR sont moyennés sur l'ensemble des images.

Pour plus de lisibilité, les performances obtenues avec la transformée en bandelettes sont tracées relativement aux performances obtenues avec la transformée en ondelettes sur la figure 3.30 et sur toutes les figures suivantes. Les résultats tracés sur la figure 3.29 sont les résultats de référence permettant d'avoir accès aux valeurs de PSNR absolues. On constate sur la figure 3.30(a) que la transformée en bandelettes permet un gain de 0,58 dB en PSNR à 0,5 bpp jusqu'à 1,18 dB à 3 bpp. Ces gains de qualité ont leur corollaire au niveau du débit. Ainsi, la figure 3.30(b) montre que la transformée en bandelettes permet une réduction de débit de 0,07 bpp pour un débit en ondelettes de 0,75 bpp soit 9,3% et une réduction de 0,21 bpp pour un débit en ondelettes de 3 bpp soit 7,0% à qualité constante.

La figure 3.31 permet de relier les différences de qualité et de débit obtenues entre la compression en bandelettes et la compression en ondelettes. Ainsi, un gain de 0,8 dB en PSNR correspond à une réduction du débit de 0,1 bpp.

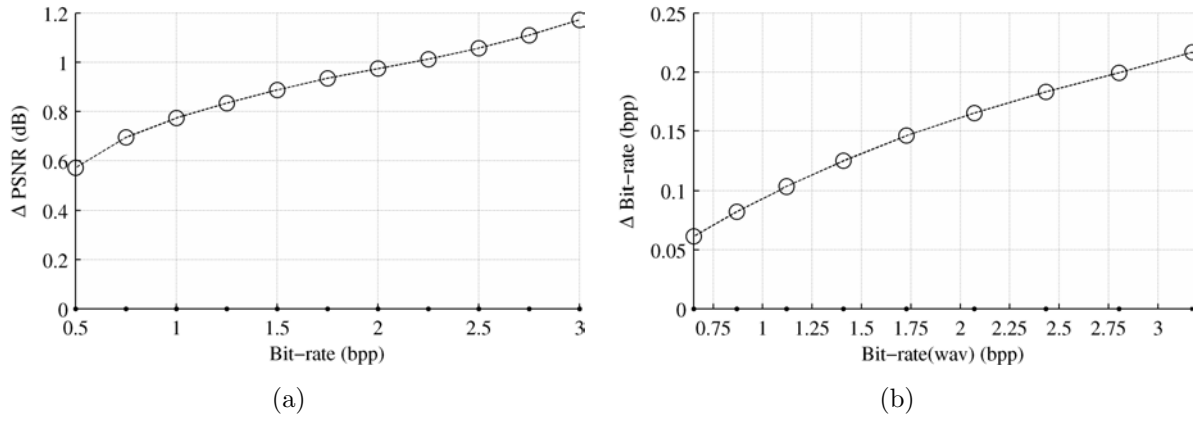


FIG. 3.30 – (a) Différences de qualité obtenues entre la compression en bandelettes et la compression en ondelettes à débits constants. (b) Différences de débit obtenues entre la compression en ondelettes et la compression en bandelettes à qualité constante.

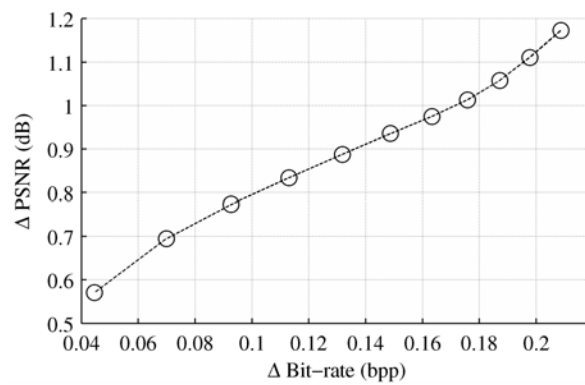


FIG. 3.31 – Différences de qualité obtenues entre la compression en bandelettes et la compression en ondelettes en fonction des différences de débit.

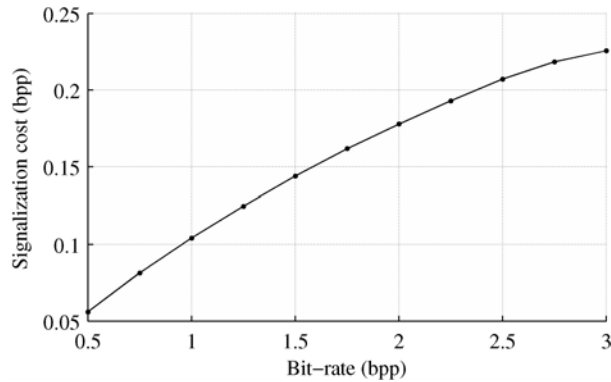


FIG. 3.32 – Coût de la signalisation des bases de bandelettes en fonction du débit total.

Sur la figure 3.32, on peut constater que le débit de signalisation croît régulièrement avec le débit total de la compression en bandelettes. Pour un débit total de 0,5 bpp, le coût moyen de la signalisation est de 0,056 bpp. A 3,0 bpp, il est de 0,23 bpp. Ces résultats étaient attendus et montrent que le processus de sélection des bases de bandelettes permet d'adapter le débit alloué à la signalisation en fonction du débit total. A haut débit, un plus grand nombre de blocs est transformé en bandelettes qu'à bas débit. Cependant, la proportion du débit total alloué à la signalisation diminue lorsque le débit total augmente : à un débit de 0,5 bpp, 11,2% du débit total est alloué à la signalisation ; à 3,0 bpp, seulement 7,5% du débit total est alloué à la signalisation.

L'estimation adaptative du débit de signalisation permet de ne pas avoir à se préoccuper de l'exactitude des probabilités *a priori* $\Pr(b)$ quel que soit le nombre de bases dans le dictionnaire. On peut donc modifier aisément le nombre de bases du dictionnaire afin de tester les performances en compression obtenues avec les différentes bases de bandelettes.

Sur la figure 3.33, les performances obtenues en utilisant diverses constitutions du dictionnaire de bandelettes sont tracées relativement aux performances obtenues avec la transformée en ondelettes. La figure 3.33(a) montre que l'utilisation des 3 bases complémentaires dans le dictionnaire de bandelettes permet une augmentation des performances d'environ 0,15 dB en PSNR. La figure 3.33(b), elle, montre que l'utilisation des 12 bases directionnelles dans le dictionnaire de bandelettes permet une augmentation des performances de seulement 0,16 dB à 0,5 bpp et de 0,35 dB à 3 bpp par rapport au dictionnaire composé des 3 bases complémentaires uniquement. Les performances en compression liées à l'une des bases complémentaires apparaissent donc supérieures aux performances moyennes des bases directionnelles. Les figures 3.33(c) et 3.33(d) montrent que les performances obtenues avec un dictionnaire composé des deux bases de Haar sont identiques aux performances obtenues avec un dictionnaire composé de la base DCT uniquement. La base DCT est donc la base la plus performante pour la compression dans le dictionnaire des bases de bandelettes. De plus, on constate sur la figure 3.33(d) que le dictionnaire composé uniquement de la base DCT permet d'atteindre plus de la moitié des gains en PSNR obtenus avec le dictionnaire des 15 bases de bandelettes par rapport à la transformée en ondelettes.

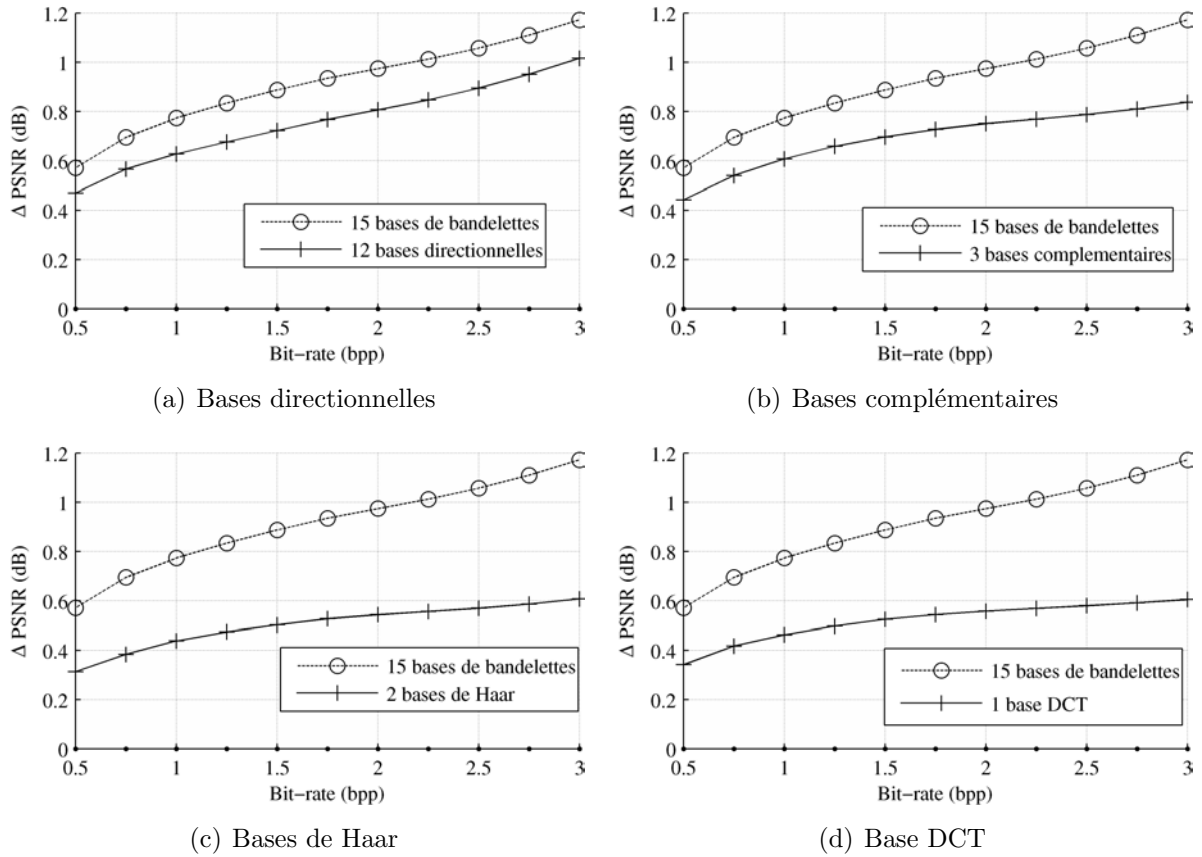


FIG. 3.33 – Différences de qualité obtenues entre la compression avec un dictionnaire de bases de bandelettes et la compression en ondelettes à débits constants. (a) Le dictionnaire est constitué des 12 bases directionnelles uniquement. (b) Le dictionnaire est constitué des 3 bases complémentaires représentées sur la figure 3.4 uniquement. (c) Le dictionnaire est constitué des 2 bases de Haar H0 et H1 représentées sur la figure 3.4 uniquement. (d) Le dictionnaire n'est constitué que de la base DCT.

3.5 Conclusion

La transformée en bandelettes par blocs, tout en étant relativement peu complexe, permet d'améliorer les résultats en compression par rapport à la transformée en ondelettes. Dans la construction de la transformée en bandelettes, nous avons vu que deux étapes sont fondamentales afin d'améliorer la compression : la constitution du dictionnaire de bases de bandelettes et l'optimisation du critère de choix de la base de bandelettes utilisée pour transformer chaque bloc.

Dans [64], des bases de bandelettes directionnelles ont été construites afin d'exploiter les régularités dans la direction des contours dans les images compressées. Ces bases constituent une application pratique de la théorie des bandelettes [64, 67, 77] dans le cadre de transformées par blocs. Malheureusement, comme elles sont utilisées conjointement à un processus d'optimisation débit-distorsion nécessaire à la compression, l'extraction d'information géométrique n'est pas possible. Des bases complémentaires ont été adjointes au dictionnaire de bases directionnelles. Ces bases complémentaires permettent d'améliorer les résultats en compression. En particulier, la base de DCT est la base la plus performante pour la compression en bandelettes. Pour limiter la complexité calculatoire, il est possible de limiter le nombre de bases de bandelettes dans le dictionnaire. L'utilisation de la base de DCT seule est alors envisagée.

Le critère de choix de la base de bandelettes utilisée pour transformer chaque bloc est issu d'une formule solution du problème d'allocation débit-distorsion adapté à la transformée en bandelettes. Cette allocation est obtenue par minimisation d'un lagrangien. Les études que nous avons menées sur ce problème ont abouti à une formule empirique quasi-optimale du multiplicateur de Lagrange à utiliser. Cette formule est très proche de la formule théorique obtenue sous hypothèse bas et moyen débits dans [67] et de la formule que nous avons obtenue sous hypothèse haut débit dans ce chapitre. Cette étude vient donc confirmer la validité des hypothèses théoriques. Nous avons montré à partir des résultats en compression d'images d'observation de la Terre et en utilisant un dictionnaire de bases aléatoires que ce critère joue pour moitié dans le gain en PSNR apporté par l'algorithme de compression en bandelettes.

Chapitre 4

Nouvelles bases pour la post-transformée

4.1	Analyse des redondances exploitées par les bases de bandes- lettes	84
4.1.1	Construction des classes de blocs	85
4.1.2	Informations mutuelles entre coefficients d'ondelettes dans un bloc	85
4.1.3	Corrélations entre coefficients d'ondelettes dans un bloc	86
4.2	Bases construites par extensions des groupements	88
4.2.1	Construction des nouvelles bases	89
4.2.2	Analyse de l'inter-corrélation entre coefficients transformés . . .	89
4.2.3	Performances en compression	92
4.3	Bases distantes entre elles	93
4.3.1	Cohérence d'un dictionnaire	93
4.3.2	Dissimilarité entre bases	95
4.4	Bases construites par ACP	96
4.4.1	ACP par sous-bandes	96
4.4.2	ACP sur différentes classes de blocs	98
4.4.3	Exploitation de la compaction d'énergie	101
4.5	Conclusion	105

LE critère de sélection des bases parmi un dictionnaire ayant été optimisé dans le chapitre précédent, nous nous intéressons maintenant à l'amélioration des résultats en compression par la modification des bases du dictionnaire elles-mêmes. Comme les bases que nous allons construire ne sont pas forcément directionnelles, il n'y a plus de lien avec la théorie des bandelettes si ce n'est pour le procédé de sélection de la meilleure base du dictionnaire. Nous appelons donc *post-transformées* les transformées par blocs effectuées dans le domaine ondelettes. Néanmoins, l'algorithme de compression reste identique à l'algorithme 1 décrit dans la section 3.1.

Dans ce chapitre, de nouvelles bases sont proposées pour la compression par post-transformées de blocs de coefficients d'ondelettes. Nous commençons par analyser quelles sont les redondances effectivement exploitées par les bases de bandelettes directionnelles dans la section 4.1. Nous avons publié dans [78] l'analyse des corrélations entre coefficients d'ondelettes exploitées par les bases de bandelettes. Nous la complétons par une analyse en termes d'information mutuelle. Suite à cette analyse, nous proposons trois approches différentes pour construire des dictionnaires de bases permettant de mieux exploiter les redondances identifiées dans les blocs de coefficients d'ondelettes. La première approche, développée dans la section 4.2, consiste à étendre les groupements des bases de bandelettes directionnelles dans le sens des corrélations identifiées. Nous l'avons proposée dans [78]. La seconde consiste à construire un dictionnaire avec des bases aussi « différentes » les unes des autres que possible. Ainsi, n'importe quel bloc de coefficients pourrait être efficacement représenté dans l'une de ces bases. Nous développons cette approche dans la section 4.3. La troisième approche, développée dans la section 4.4, consiste à construire des bases par Analyse en Composantes Principales (ACP) afin de décorréliser les coefficients d'ondelettes dans les blocs. Nous l'avons proposée dans [78] et [53]. Enfin, nous montrons qu'il est possible d'exploiter la propriété de compaction d'énergie des bases construites par ACP. Pour cela, nous utilisons le schéma de codage que nous avons proposé dans [53]. Nous montrons aussi que la base de DCT et la base de Hadamard possèdent cette propriété de compaction d'énergie.

4.1 Analyse des redondances exploitées par les bases de bandelettes

Avant de procéder à la construction de nouveaux dictionnaires de bases pour la post-transformée, nous allons quantifier les redondances exploitées par les bases de bandelettes directionnelles et surtout identifier les redondances qui ne sont pas exploitées. Pour cela, nous évaluons la capacité des bases de bandelettes directionnelles à exploiter les redondances statistiques en effectuant une analyse de l'information mutuelle et de l'inter-corrélation entre toutes les paires de coefficients d'ondelettes à l'intérieur de blocs de taille 4×4 .

4.1.1 Construction des classes de blocs

Pour évaluer les performances d'une base de bandelettes directionnelles, nous considérons uniquement les blocs de coefficients d'ondelettes qui doivent être transformés dans cette base selon le critère débit-distorsion (3.3). Une analyse en bandelettes a été effectuée sur les sept images d'observation de la Terre de taille 1024×1024 de l'ensemble d'apprentissage représenté en annexe A. Pour le calcul du critère débit-distorsion (3.3), le pas de quantification a été fixé à $q = 32$ ce qui correspond à un débit d'environ 2 bpp pour la compression d'images 12 bits. Ce débit est légèrement inférieur au débit cible visé sur les satellites PLEIADES. L'analyse a été effectuée séparément sur les trois sous-bandes du premier niveau de la transformée en ondelettes : HL_1 , LH_1 et HH_1 . Chacune de ces sous-bandes contient 16384 blocs de coefficients d'ondelettes de taille 4×4 . L'analyse en bandelettes permet de classer ces blocs de coefficients d'ondelettes : les blocs qui sont mieux représentés dans la base de bandelettes $\#b$ d'après le critère débit-distorsion (3.3) forment la classe b . Nous effectuons ensuite une analyse statistique sur chacune de ces classes de blocs de coefficients d'ondelettes.

4.1.2 Informations mutuelles entre coefficients d'ondelettes dans un bloc

L'information mutuelle relative a été calculée entre toutes les paires de coefficients d'ondelettes de chacune des classes de blocs après quantification. L'information mutuelle relative $I_r(X, Y)$ représente la proportion d'information qui peut-être économisée sur le codage de Y si X est déjà connu. Elle est définie dans l'équation 2.1.

La figure 4.1 représente l'information mutuelle entre paires de coefficients d'ondelettes dans les blocs qui sont mieux représentés dans les bases de bandelettes directionnelles $\#1$ et $\#2$ dont les groupements sont représentés sur la figure 3.1. La classe 2 par exemple est composée de 4335 blocs dans la sous-bande HL_1 , 4340 blocs dans la sous-bande LH_1 et 1909 blocs dans la sous-bande HH_1 . Sur la figure 4.1, les traits pleins représentent une information mutuelle relative $I_r(X, Y)$ supérieure à 20%. Les traits discontinus représentent $16\% \leq I_r(X, Y) < 20\%$ et les traits fins $10\% \leq I_r(X, Y) < 16\%$. L'information mutuelle est élevée entre paires de coefficients d'ondelettes qui sont groupés dans les bases de bandelettes directionnelles $\#1$ et $\#2$. Ces dépendances entre coefficients d'ondelettes sont donc bien exploitées par ces bases de bandelettes dans toutes les sous-bandes. Cependant, il existe aussi de l'information mutuelle entre coefficients adjacents dans la direction horizontale dans la sous-bande HL_1 et dans la direction verticale dans la sous-bande LH_1 , c'est-à-dire dans la direction du filtrage d'ondelettes passe-haut. Ces dépendances entre coefficients d'ondelettes sont observables dans toutes les classes de blocs avec le même ordre de grandeur. Ici, elles sont peu exploitées par les bases de bandelettes directionnelles $\#1$ et $\#2$.

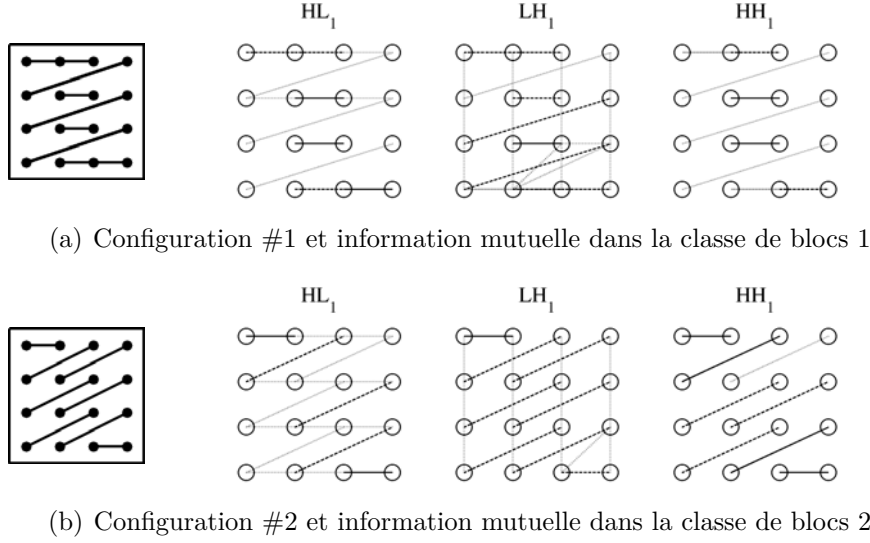


FIG. 4.1 – Information mutuelle relative entre chaque paire de coefficients d'ondelettes des blocs de la classe 1 et 2.

4.1.3 Corrélations entre coefficients d'ondelettes dans un bloc

Une matrice de coefficients de corrélation est calculée pour chaque classe de blocs. Elle représente les corrélations intra-bloc entre toutes les paires de coefficients d'ondelettes. Le coefficient de corrélation $r_{X,Y}$ entre deux coefficients d'ondelettes X et Y d'une classe de n blocs est défini par :

$$r_{X,Y} = \frac{1}{(n-1)} \frac{1}{\sigma_x \sigma_y} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (4.1)$$

où x_i représente la valeur du coefficient d'ondelettes X dans le bloc numéro $i = 1, 2, \dots, n$, $\bar{x}$ est la moyenne des x_i et σ_x^2 est la variance des x_i .

Les matrices des valeurs absolues des coefficients de corrélation des classes de blocs 1 et 2 pour les sous-bandes HL_1 , LH_1 et HH_1 sont représentées sur la figure 4.2. La numérotation des coefficients dans un bloc de 4×4 coefficients d'ondelettes s'effectue colonne par colonne comme défini sur la figure 4.3.

Sur la figure 4.2, les corrélations les plus fortes apparaissent entre coefficients d'ondelettes adjacents horizontalement dans la sous-bande HL_1 et adjacents verticalement dans la sous-bande LH_1 . Ces corrélations sont négatives et sont dues au filtrage passe-haut de la transformée en ondelettes. Sur la figure 4.2, les coefficients de corrélation marqués d'une croix correspondent aux groupements effectués dans les bases de bandelettes directionnelles. Par exemple, dans la base de bandelettes #2, les coefficients 2 et 9 sont groupés. Les coefficients aux positions (2,9) et (9,2) dans les matrices de corrélations sont donc marqués d'une croix. Comme ces coefficients d'ondelettes sont regroupés dans les bases de bandelettes directionnelles, de fortes corrélations étaient attendues. Pourtant, les coefficients de corrélation entre ces coefficients d'ondelettes sont faibles.

Les corrélations entre paires de coefficients d'ondelettes des blocs de la classe 1 et 2 sont aussi représentées sur la figure 4.4. Les traits pleins représentent des coefficients de

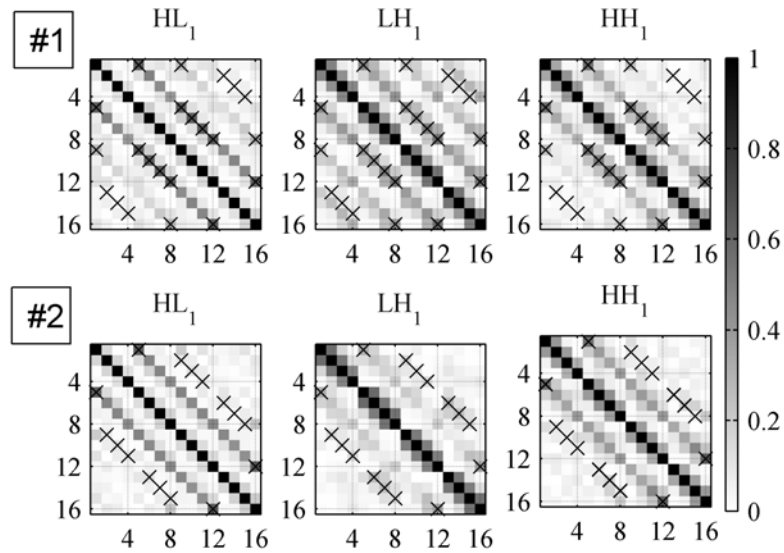


FIG. 4.2 – Matrices des coefficients de corrélation intra-blocs entre coefficients d'ondelettes des blocs des classes 1 et 2 dans les trois sous-bandes HL_1 , LH_1 et HH_1 .

1	5	9	13
2	6	10	14
3	7	11	15
4	8	12	16

FIG. 4.3 – Numérotation des coefficients d'ondelettes dans un bloc 4×4 .

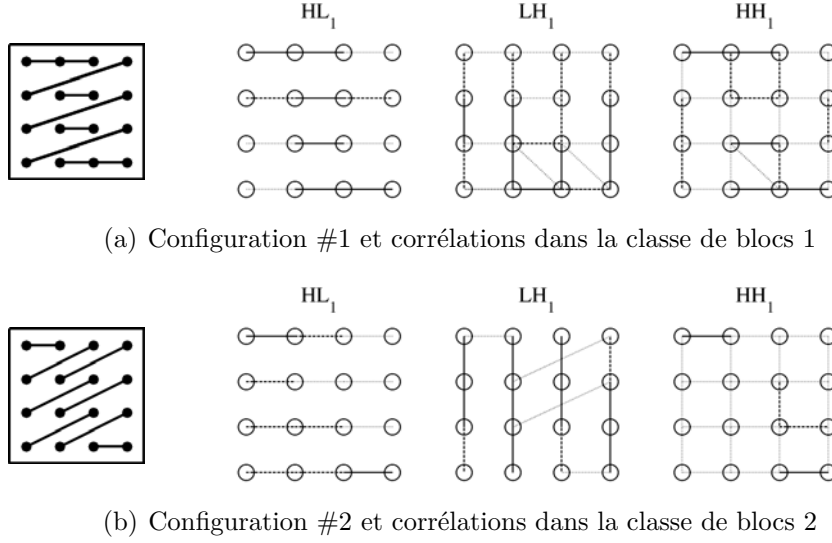


FIG. 4.4 – Corrélations entre paires de coefficients d'ondelettes des classes de blocs #1 et #2.

corrélation $r_{X,Y}$ supérieurs à 0,5 en valeur absolue. Les traits discontinus représentent $0,4 \leq |r_{X,Y}| < 0,5$ et les traits fins $0,25 \leq |r_{X,Y}| < 0,4$. Peu de corrélations apparaissent dans la direction des groupements des bases directionnelles #1 et #2. Les deux coefficients de corrélation qui apparaissent entre paires de coefficients d'ondelettes dans la direction des groupements #2 dans la sous-bande LH₁ ne valent que 0,33 et 0,27. Comme il y a peu de corrélations entre coefficients d'ondelettes dans la direction des groupements, on en déduit que les bases directionnelles #1 et #2 exploitent peu de corrélations. De plus, elles exploitent peu les corrélations horizontales et verticales qui ont été mises en évidence dans ces blocs. Ces corrélations négatives correspondent à la direction du filtrage passe-haut de la transformée en ondelettes. Les mêmes observations ont été faites pour toutes les classes de blocs correspondant à des bases de bandelettes directionnelles.

Les bases de bandelettes directionnelles ne sont donc pas efficaces pour exploiter des corrélations directionnelles. Cela peut s'expliquer par le fait que les corrélations entre coefficients non adjacents dans les directions autres que l'horizontale et la verticale sont faibles. A partir de ces observations, de nouvelles bases ont été construites afin de mieux décorrélérer les coefficients d'ondelettes en prenant en compte les différences statistiques entre les sous-bandes de la transformée en ondelette.

4.2 Bases construites par extensions des groupements

Dans la section précédente, les corrélations prédominantes ont été identifiées entre coefficients d'ondelettes adjacents horizontalement dans les sous-bandes HL et adjacents verticalement dans les sous-bandes LH, c'est-à-dire dans les directions du filtrage passe-

haut de la transformée en ondelettes. Partant de ce constat, nous proposons de nouveaux groupements prenant en compte ces directions.

4.2.1 Construction des nouvelles bases

Pour accroître les performances en compression, de nouveaux groupements sont proposés afin d'exploiter les corrélations dans le sens du filtrage passe-haut. Bien qu'elles n'exploitent que peu de corrélations directionnelles, les bases de bandelettes donnent de bonnes performances en compression. C'est pourquoi les nouveaux groupements sont construits à partir des groupements des bases de bandelettes directionnelles représentés sur la figure 3.1 en leur associant des groupements horizontaux et verticaux. Comme les coefficients d'ondelettes ne sont pas corrélés dans la même direction selon la sous-bande, les nouveaux groupements proposés sont différents dans les sous-bandes HL, LH et HH. Dans les sous-bandes HL, les groupements horizontaux sont privilégiés. Dans les sous-bandes LH, ce sont les groupements verticaux qui sont privilégiés. Enfin, dans les sous-bandes HH, nous avons choisi de privilégier les groupements dans les deux directions. Les différentes configurations de groupements pour les sous-bandes HL, LH et HH sont représentées sur la figure 4.5.

Comme les groupements sont maintenant bi-dimensionnels, des bases de polynômes de Legendre 2D ont été construites par produits tensoriels des bases de polynômes 1D. Un dictionnaire de 12 bases a été construit pour chacune des orientations HL, LH, et HH. Les bases #2 des dictionnaires des sous-bandes HL et LH sont représentées sur la figure 4.6. Il est à noter que l'utilisation d'un dictionnaire différent par sous-bande n'augmente pas la complexité calculatoire ni le coût de signalisation.

4.2.2 Analyse de l'inter-corrélation entre coefficients transformés

Il y a $C_{16}^2 = 120$ combinaisons de paires de coefficients dans un bloc de taille 4×4 et donc autant de coefficients de corrélation entre paires de coefficients. Sur la figure 4.7, seuls les 64 coefficients de corrélation avec la plus forte amplitude $|r_{X,Y}|$ sont représentés. Ils sont triés par amplitude décroissante. Les courbes noires correspondent aux coefficients d'inter-corrélation entre coefficients d'ondelettes, les courbes bleues à l'inter-corrélation entre coefficients de bandelettes, et les courbes marrons à l'inter-corrélation entre coefficients dans les bases de groupements étendus. Cette figure montre que les corrélations entre coefficients des blocs transformés dans les bases construites par groupements étendus sont plus faibles que les corrélations entre coefficients des blocs transformés dans les bases de bandelettes directionnelles. Comme les groupements de la base #1 n'ont pas été modifiés pour la sous-bande HL₁ (cf. figure 4.5), les résultats sont identiques dans ce cas là.

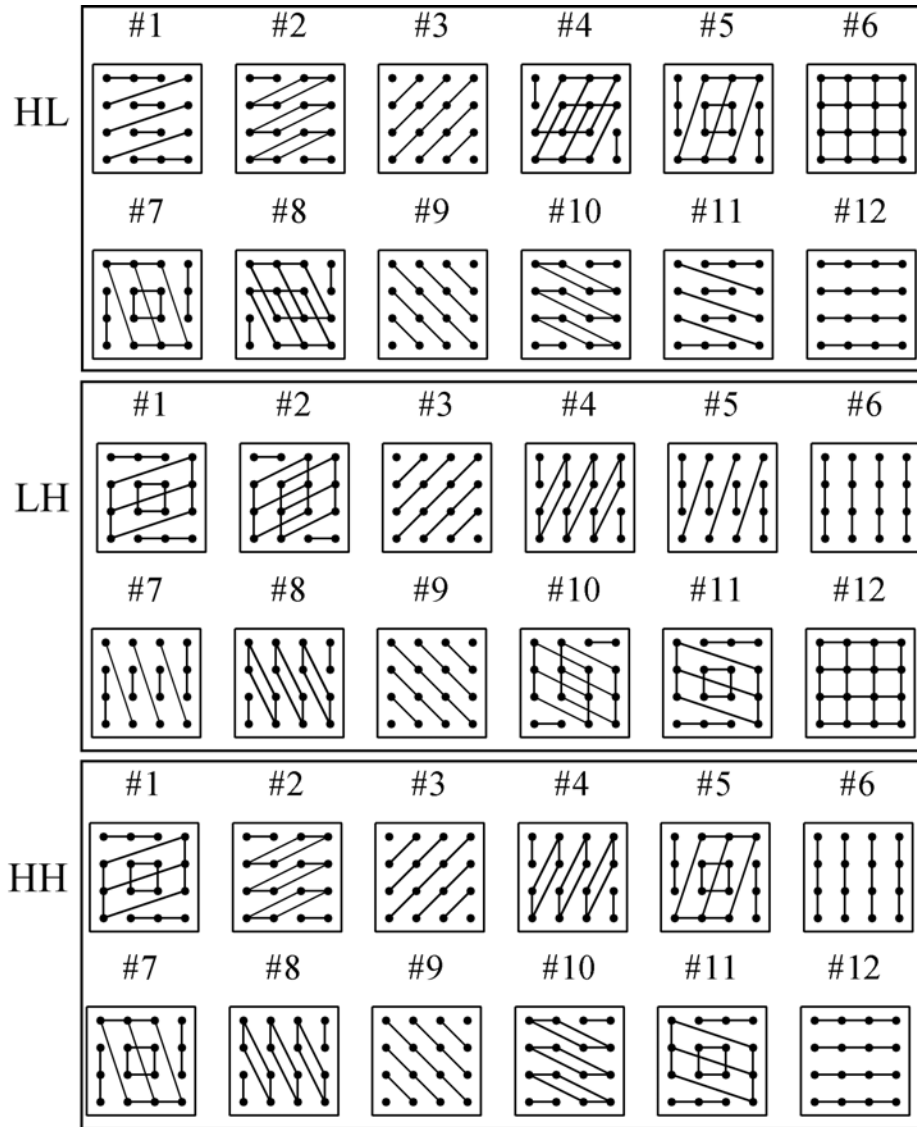


FIG. 4.5 – Nouveaux groupements tenant compte des dépendances dans la direction du filtre d'ondelettes passe-haut.

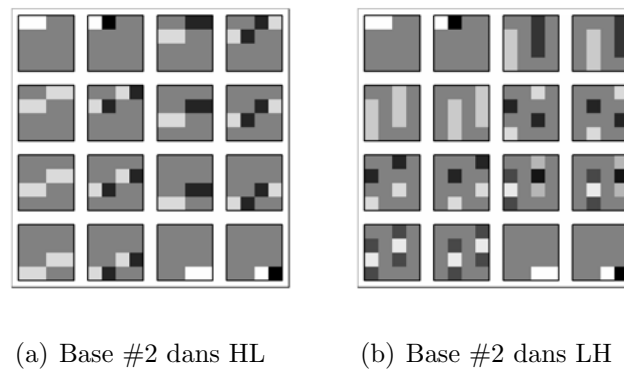


FIG. 4.6 – Bases #2 des nouveaux dictionnaires pour les sous-bandes HL et LH.

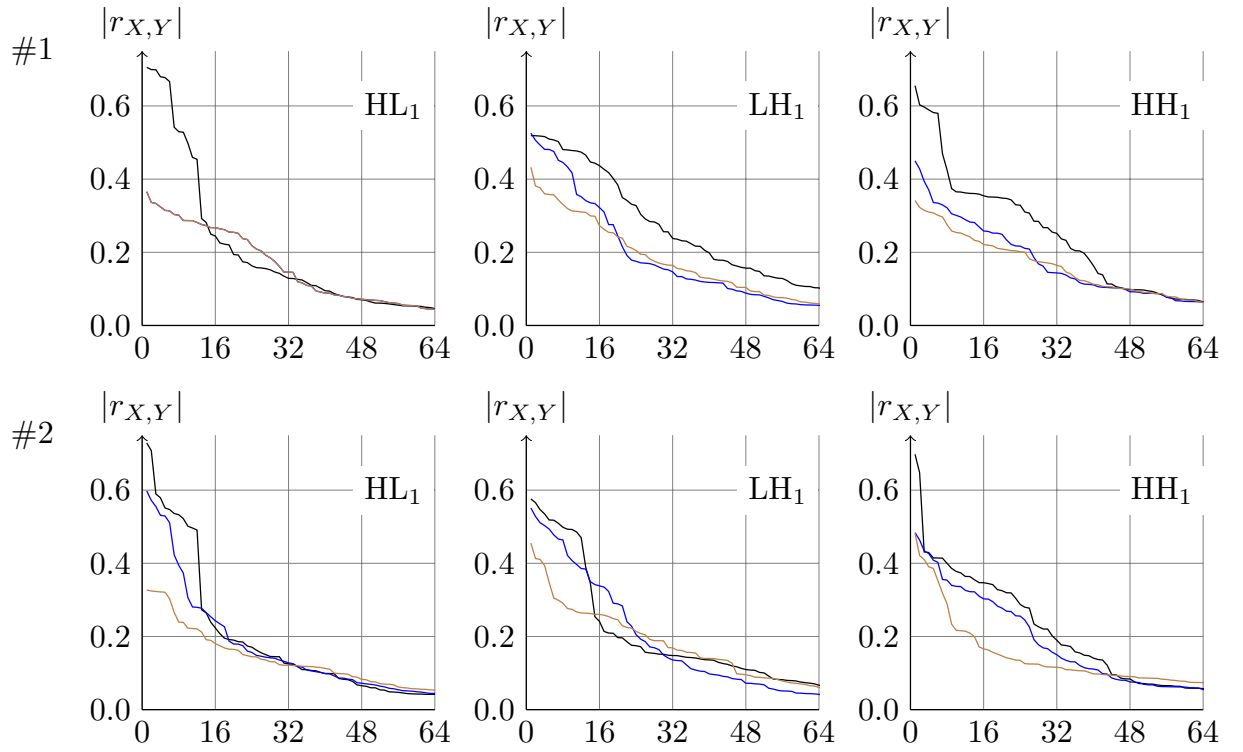


FIG. 4.7 – Les 64 plus forts coefficients de corrélation triés par ordre d’amplitude décroissante pour les classes de blocs 1 et 2 dans les sous-bandes HL₁, LH₁ et HH₁. Les courbes noires correspondent aux corrélations entre coefficients d’ondelettes, les courbes bleues aux corrélations entre coefficients de bandelettes et les courbes marrons aux corrélations entre coefficients dans les bases de groupements étendus.

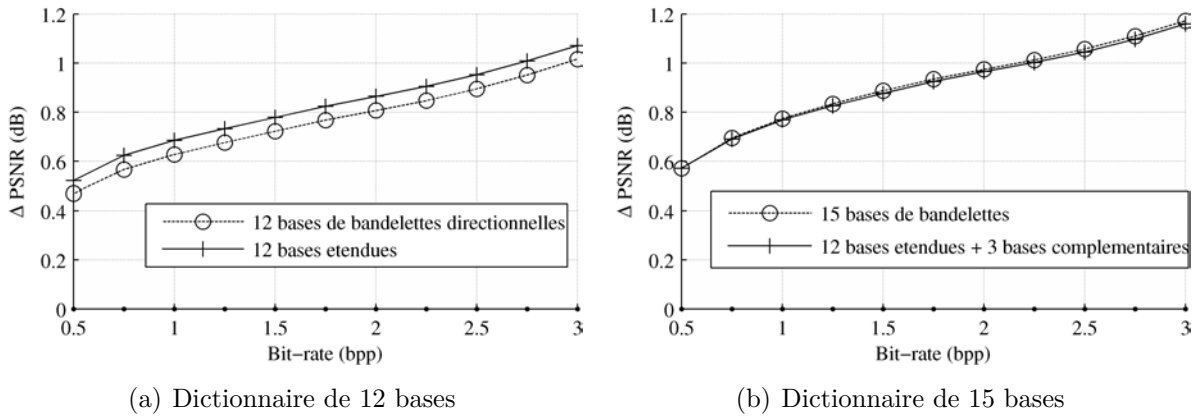


FIG. 4.8 – Performances obtenues par compression avec un dictionnaire de bases de bandelettes et avec un dictionnaire de bases étendues relativement aux performances obtenues avec les ondelettes.

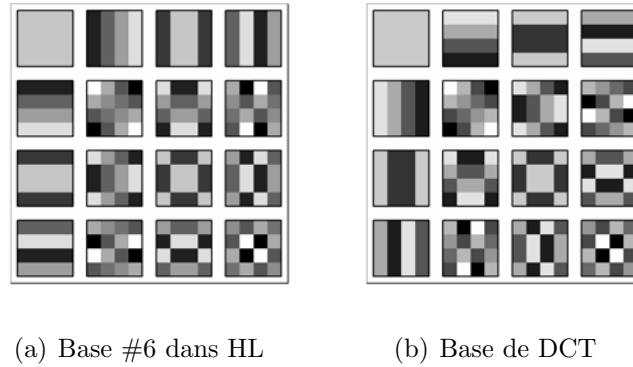


FIG. 4.9 – Base #6 du nouveau dictionnaire pour la sous-bande HL et base de DCT.

4.2.3 Performances en compression

Sur la figure 4.8(a), les résultats en compression obtenus avec les 12 bases construites à partir de ces nouveaux groupements sont comparés aux résultats obtenus avec les 12 bases directionnelles de la transformée en bandelettes. Les nouveaux groupements permettent d'augmenter légèrement les performances en compression par rapport aux bases de bandelettes directionnelles. Cependant, les résultats de la figure 4.8(b) montrent que lorsque les bases complémentaires, c'est-à-dire la base de DCT et les deux bases de Haar, sont ajoutées à ces deux dictionnaires, les résultats en compression sont similaires. Ces bases additionnelles sont plus performantes lorsqu'elles sont utilisées conjointement aux bases de bandelettes directionnelles originales. Cela peut être dû à la similitude de certaines des nouvelles bases avec les bases complémentaires. En effet, comme le montre la figure 4.9, la nouvelle base #6 de la sous-bande HL est similaire à la base de DCT. Il y a donc une redondance dans le dictionnaire de bases. Ces deux bases peuvent être utilisées presque indifféremment. L'une n'apporte rien par rapport à l'autre. Mais parce qu'elles sont toutes les deux présentes dans le dictionnaire les performances sont pénalisées par le coût de la signalisation.

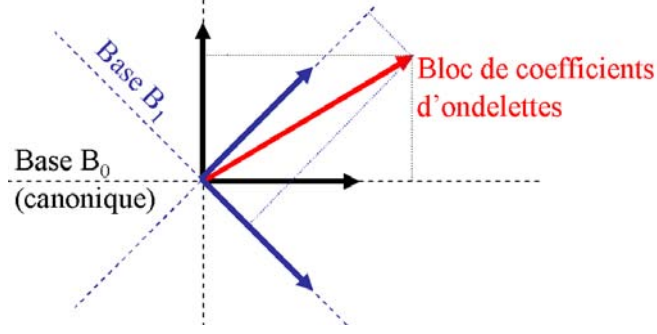


FIG. 4.10 – Exemple en 2D.

On peut alors penser à construire un dictionnaire de bases « non similaires » dans $\mathbb{R}^{16}$.

4.3 Bases distantes entre elles

Les blocs de 4×4 coefficients d'ondelettes peuvent être considérés comme des vecteurs de l'espace $\mathbb{R}^{16}$. Si ces vecteurs étaient uniformément répartis dans cet espace alors, afin d'obtenir la meilleure représentation possible au sens débit-distorsion (3.3), il faudrait que les bases contenues dans le dictionnaire soient le plus espacées possible dans $\mathbb{R}^{16}$. Cette idée est schématisée sur la figure 4.10 dans un espace à deux dimensions. La base la plus « éloignée » de la base canonique $\mathcal{B}_0$ est la base $\mathcal{B}_1$ obtenue par rotation de $\pi/4$ de la base $\mathcal{B}_0$. Un bloc de deux coefficients d'ondelettes est représenté par un vecteur en rouge sur cette figure. Les coefficients de la post-transformée sont les coordonnées de ce vecteur dans la base $\mathcal{B}_1$. Dans cette section, nous discutons de la définition de bases « éloignées » entre elles ce qui nous conduit à la construction d'un nouveau dictionnaire.

4.3.1 Cohérence d'un dictionnaire

Dans [79], Gribonval et Nielsen considèrent le problème de représentation parcimonieuse d'un signal dans un dictionnaire redondant de vecteurs formé de l'union de plusieurs bases. Le problème est différent de celui des post-transformées pour lequel on considère un dictionnaire de bases et non un *dictionnaire de vecteurs*. Dans le cas des post-transformées, le signal est représenté dans des bases indépendantes les unes des autres et non dans l'ensemble des vecteurs du dictionnaire. Néanmoins, comme le concept est assez proche, nous nous inspirons de leurs travaux.

Gribonval et Nielsen utilisent la mesure de *cohérence* d'un dictionnaire de vecteurs :

Définition. La cohérence $M(\mathbf{D})$ du dictionnaire de vecteurs $\mathbf{D} = [\mathbf{g}_1, \dots, \mathbf{g}_K]$ où $K \geq N$ et la famille $\{\mathbf{g}_k\}_{k=1}^K$ est génératrice de l'espace $\mathbb{R}^N$ est définie par :

$$M(\mathbf{D}) = \max_{k \neq k'} |\langle \mathbf{g}_k, \mathbf{g}_{k'} \rangle|.$$

Ils montrent la propriété suivante :

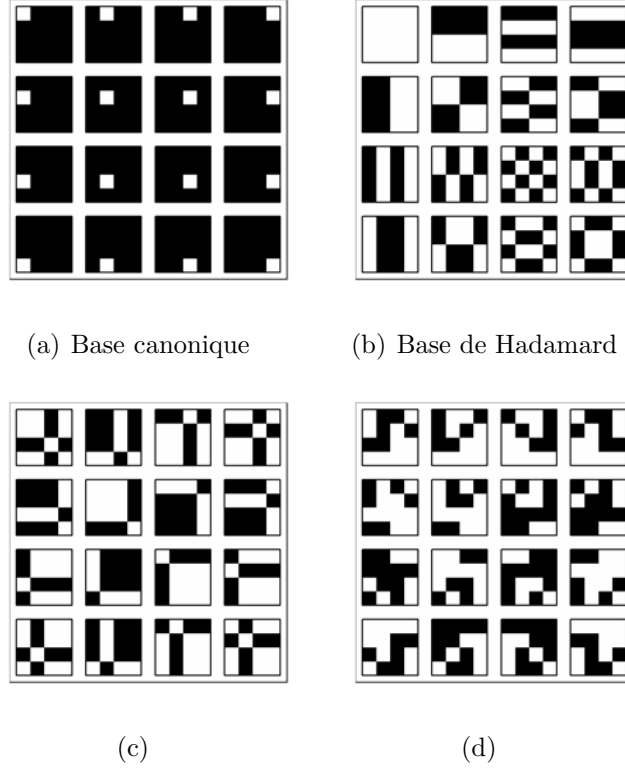


FIG. 4.11 – Dictionnaire de 4 bases dans $\mathbb{R}^{16}$ avec une cohérence minimale égale à $1/4$. Dans la base canonique, un coefficient noir vaut 0 et un coefficient blanc 1. Dans les bases (b), (c) et (d), un coefficient noir vaut $-1/4$ et un coefficient blanc $1/4$.

Propriété. *Dès lors que le dictionnaire $\mathbf{D}$ contient une base orthonormale et un vecteur normé additionnel, alors la cohérence $M(\mathbf{D})$ respecte l'inégalité suivante :*

$$M(\mathbf{D}) \geq \frac{1}{\sqrt{N}}.$$

La preuve est retranscrite en annexe D. Le théorème suivant donne l'existence d'un dictionnaire contenant plusieurs bases et minimisant la cohérence pour certaines valeurs de N [79, 80].

Théorème. *Soient $N = 2^{j+1}$, $j \geq 0$ et $\mathcal{H} = \mathbb{R}^N$. Il existe un dictionnaire $\mathbf{D}$ dans $\mathcal{H}$ formé de l'union de $L = 2^j = N/2$ bases orthonormées de $\mathcal{H}$ tel que pour chaque paire $\mathbf{u}, \mathbf{v} \in \mathbf{D}$, $\mathbf{u} \neq \mathbf{v}$: $|\langle \mathbf{u}, \mathbf{v} \rangle| \in \{0, 1/\sqrt{N}\}$.*

Pour $N = 16$, il existe donc un dictionnaire formé de 8 bases orthonormées tel que la cohérence de ce dictionnaire est égale à $1/4$. Un dictionnaire de 4 bases de $\mathbb{R}^{16}$ avec une cohérence minimale de $1/4$ est représenté sur la figure 4.11.

Minimiser la cohérence d'un dictionnaire est équivalent à maximiser les angles entre les différents vecteurs du dictionnaire et à construire un dictionnaire de vecteurs équiangulaires [80]. Ainsi, dans un dictionnaire avec une cohérence de $M(\mathbf{D}) = 1/\sqrt{N}$, le plus petit angle θ entre deux vecteurs est $\cos \theta = 1/\sqrt{N}$. Dès lors, en dimension $N = 16$, dans

un dictionnaire contenant plus d'une base, le plus petit angle entre deux vecteurs sera forcément inférieur à $\theta = \arccos(1/4) \approx 75,5^\circ$.

Cette mesure montre que les dictionnaires de bases aléatoires tels que ceux construits dans la section 3.4.4 ne sont pas incohérents. Avec 8 bases aléatoires, on obtient une cohérence d'environ 0,8. L'angle minimal entre deux vecteurs du dictionnaire est d'environ 37° .

Cependant, cette mesure n'est pas adaptée aux bases de bandelettes. En effet, dès que deux bases ont un vecteur en commun, la cohérence du dictionnaire est maximale et vaut 1. Certaines bases de bandelettes ont des vecteurs communs. Par exemple les bases directionnelles #1 et #11 ont 4 vecteurs en commun ; elles sont pourtant différentes. Une mesure sur un dictionnaire de bases et non un dictionnaire de vecteurs est donc nécessaire.

4.3.2 Dissimilarité entre bases

Nous définissons une mesure de dissimilarité entre bases à partir de la valeur absolue du produit scalaire entre les vecteurs des différentes bases comme dans la mesure de cohérence. Au lieu de prendre le maximum des produits scalaires, nous prenons la somme.

Définition. La mesure de dissimilarité $d(A, B)$ entre deux bases orthonormées $A = \{\mathbf{a}_i\}_{i=1}^N$ et $B = \{\mathbf{b}_j\}_{j=1}^N$ est définie comme la somme de la valeur absolue des produits scalaires entre chaque paire de vecteurs des bases A et B moins N .

$$d(A, B) = \left(\sum_{i=1}^N \sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle| \right) - N. \quad (4.2)$$

Le terme $-N$ permet d'avoir une mesure de dissimilarité nulle entre deux bases identiques : $d(A, A) = 0$. La réciproque est vraie : si $d(A, B) = 0$, alors la base A est identique à la base B au sens des vecteurs près. Enfin, la dissimilarité entre deux bases A et B ne peut excéder $N(\sqrt{N} - 1)$:

$$d(A, B) \leq N(\sqrt{N} - 1)$$

Les preuves sont fournies en annexe D. En dimension $N = 16$, la mesure de dissimilarité entre bases ne peut donc pas excéder 48. Cette dissimilarité maximale est obtenue entre les bases représentées sur la figure 4.11.

La dissimilarité moyenne entre les bases aléatoires des dictionnaires construits dans la section 3.4.4 est de 36. Sur la figure 4.12, on constate aussi que les dissimilarités entre les bases de bandelettes directionnelles sont faibles. Par contre, la dissimilarité entre la base de DCT et la base canonique (#0) est élevée. Elle vaut 43,2. Cela peut expliquer en partie les bonnes performances en compression obtenues avec un dictionnaire constitué seulement de la base de DCT et de la base canonique (figure 3.33(d)). Pourtant, ces résultats en compression sont entre 0,05 dB et 0,1 dB supérieurs aux résultats en compression obtenus avec un dictionnaire formé des quatre bases représentées sur la figure 4.11 et qui ont une dissimilarité maximale. On peut donc en conclure que les performances en compression ne sont pas forcément dues à une bonne répartition des bases dans l'espace $\mathbb{R}^{16}$. L'hypothèse que les blocs de coefficients d'ondelettes sont uniformément répartis dans l'espace $\mathbb{R}^{16}$

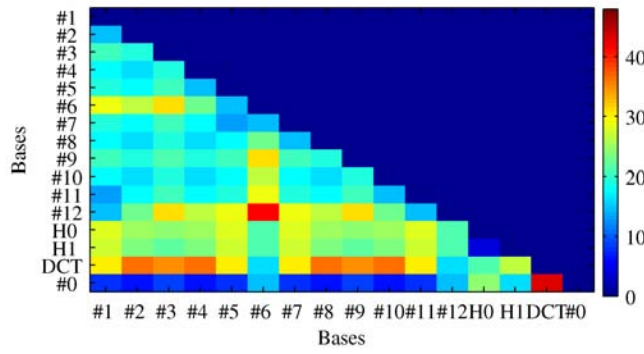


FIG. 4.12 – Mesures de dissimilarités entre bases de bandelettes.

peut être fausse. Il est donc préférable de construire des bases adaptées aux spécificités des blocs de coefficients d'ondelettes.

4.4 Bases construites par ACP

Nous avons vu dans les sections précédentes que les bases construites par extensions des groupements de bandelettes dans le sens du filtrage passe-haut ou les bases construites dans le but d'optimiser la répartition des vecteurs dans l'espace $\mathbb{R}^{16}$ ne permettent pas d'améliorer les résultats en compression. Dans cette section, nous proposons de nouvelles bases adaptées aux données.

Afin de mieux exploiter les corrélations identifiées dans la partie 4.1.3, il est possible d'utiliser des bases construites à partir d'Analyses en Composantes Principales (ACP). Il ne s'agit pas à proprement parlé d'une ACP puisque nous ne cherchons pas à réduire la dimension des données. Toutes les composantes seront utilisées pour reconstruire les blocs de coefficients d'ondelettes. Ce n'est pas non plus une transformée de Karhunen-Loeve (KLT) puisque les bases sont construites à partir d'un ensemble d'images différent de celui sur lequel elles sont appliquées pour la compression. Pour une compression à bord de satellites d'observation de la Terre, les bases pourraient être apprises au sol sur un ensemble d'images pertinentes et téléchargées à bord pour la compression.

4.4.1 ACP par sous-bandes

Comme les statistiques des sous-bandes sont différentes, une idée clé pour augmenter les performances en compression est d'utiliser un dictionnaire de bases différent par sous-bande comme dans la section 4.2. En effet, l'utilisation d'un dictionnaire différent par sous-bande n'augmente pas la complexité calculatoire ni le coût de signalisation.

Dans un premier temps, une base est construite au moyen d'une ACP sur l'ensemble des blocs d'une même sous-bande de la transformée en ondelettes de toutes les images de l'ensemble d'apprentissage représenté dans l'annexe A. La figure 4.13 illustre le schéma de construction des bases par ACP. Les bases obtenues pour les sous-bandes HL, LH et HH sont représentées sur la figure 4.14. Les directions des filtres passe-bas et passe-haut

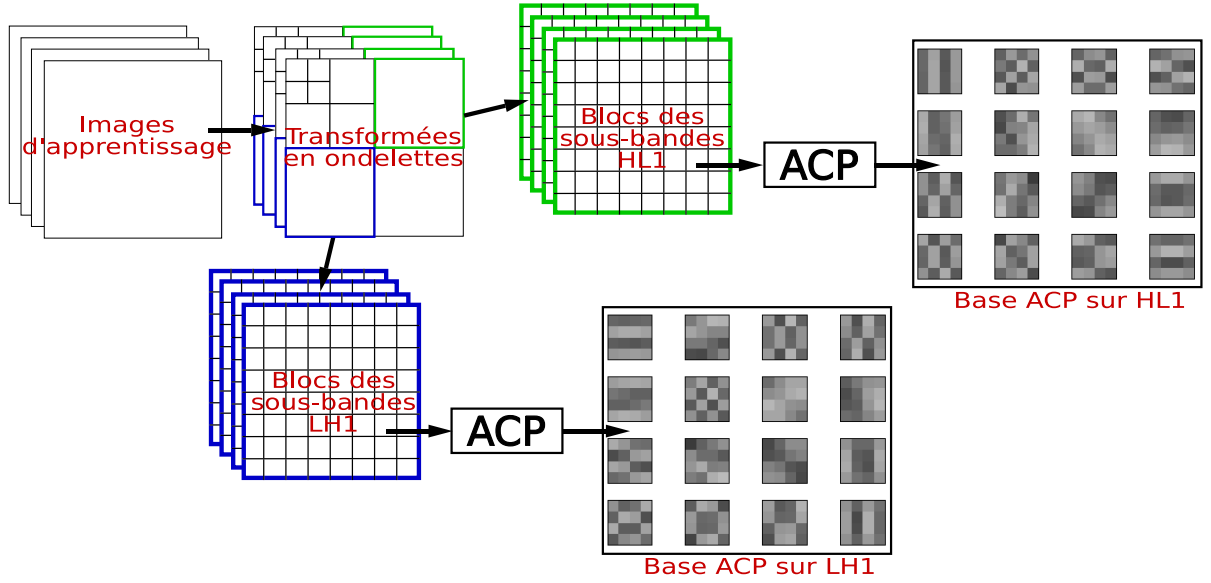
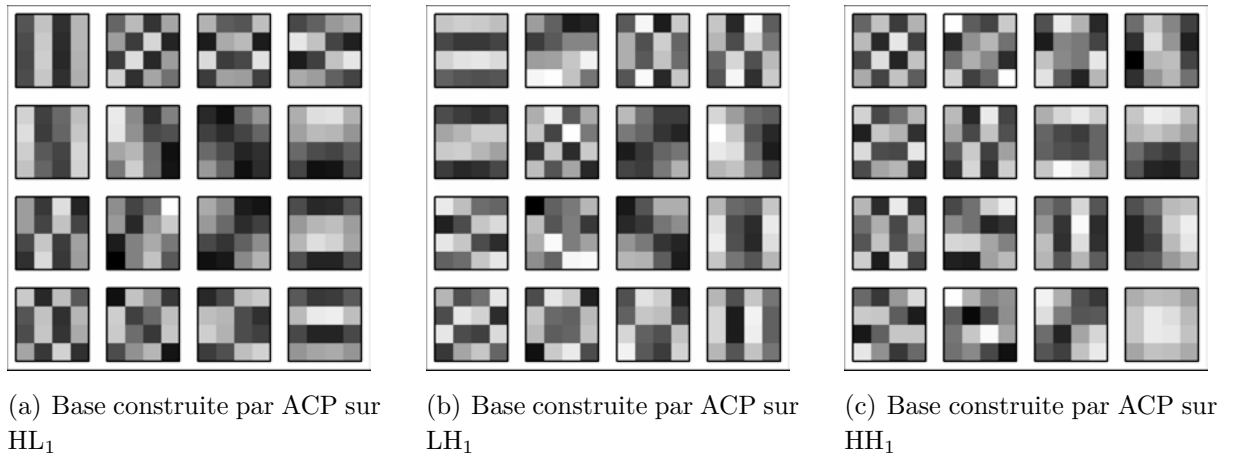


FIG. 4.13 – Schéma de construction des bases par ACP.

FIG. 4.14 – Bases construites par ACP sur les blocs de coefficients d'ondelettes extraits des sous-bandes HL₁, LH₁ and HH₁. Les vecteurs sont rangés colonne par colonne par ordre de valeur propre décroissante.

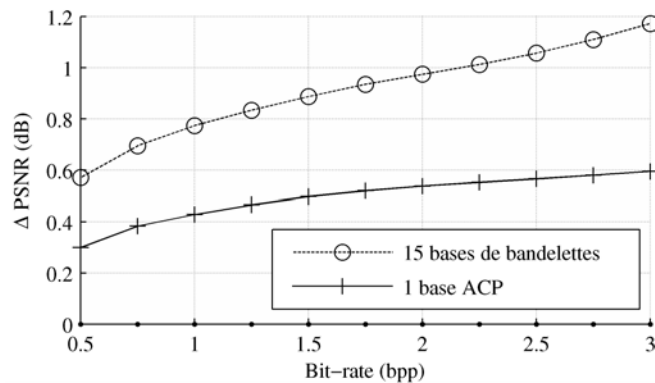


FIG. 4.15 – Différences de qualités obtenues entre la compression par post-transformée avec une base ACP par sous-bande et la compression en ondelettes. Les performances de la compression en bandelettes sont aussi tracées.

sont clairement visibles sur les vecteurs principaux des différentes sous-bandes.

Pour la compression, les dictionnaires de chaque sous-bande ne sont composés que d'une seule base : la base construite par ACP adaptée à la sous-bande traitée. Les blocs de coefficients d'ondelettes sont alors soit post-transformés dans cette base construite par ACP, soit non post-transformés.

Une transformée dans une base construite par ACP est plus complexe qu'une transformée dans une base de bandelettes. En effet, tous les coefficients des vecteurs des bases construites par ACP sont des réels flottants. Une transformée dans cette base requiert donc 256 opérations en flottants. Au contraire, la plupart des coefficients des vecteurs d'une base de bandelette directionnelle sont nuls ce qui limite fortement le nombre d'opérations à réaliser. Cependant, ici, le coût calculatoire est limité puisque une seule post-transformée est réalisée.

Les résultats de la compression par post-transformée avec une base ACP adaptée à chaque sous-bande sont tracés sur la figure 4.15. Ces résultats sont similaires à ceux obtenus avec la base de DCT seule sur la figure 3.33(d). Cependant, la compaction d'énergie due à l'ACP peut facilement être exploitée. En effet, les vecteurs dans les bases d'ACP sont ordonnés par contribution décroissante d'énergie. Les coefficients transformés correspondant aux vecteurs principaux ont en général une amplitude plus élevée que les coefficients transformés correspondant aux autres vecteurs. Après quantification, un codeur entropique peut donc exploiter efficacement cet *a priori* en codant séparément les coefficients inhérents aux différents vecteurs de la base ACP.

4.4.2 ACP sur différentes classes de blocs

Plutôt que d'effectuer une ACP sur tous les blocs de coefficients d'ondelettes d'une même sous-bande, les blocs de coefficients d'ondelettes peuvent aussi être organisés sur un nombre de classes plus important et ayant des propriétés différentes. Une base peut alors être construite par ACP sur chacune de ces classes. Dans la section 4.1.1, les blocs de coefficients d'ondelettes ont été classés par une analyse en bandelettes. Nous avons vu que

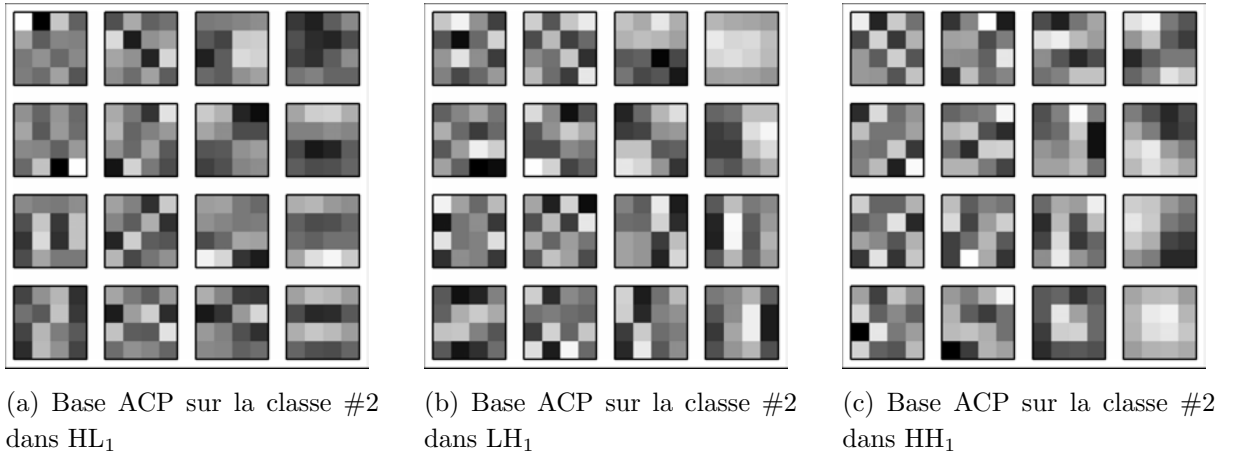


FIG. 4.16 – Bases construites par ACP sur la classe de blocs de coefficients d'ondelettes #2 construite dans la section 4.1.1.

les transformées en bandelettes directionnelles ne permettent pas de bien décorréler les coefficients d'ondelettes dans ces classes de blocs mais que les performances en compression après la transformée en bandelettes sont intéressantes. C'est pourquoi, une ACP a été effectuée sur chacune de ces classes. Les bases obtenues pour les blocs de la classe #2 dans les sous-bandes HL_1 , LH_1 et HH_1 sont représentés sur la figure 4.16. La figure 4.17 montre que ces nouvelles bases construites par ACP sont efficaces pour réduire les corrélations entre coefficients d'ondelettes dans les blocs des différentes classes qui ont été identifiées dans la section 4.1.3. En effet, les coefficients de corrélation les plus forts sont inférieurs à 0,2 dans ces bases et supérieurs à 0,5 dans les bases de bandelettes.

Sur la figure 4.18, les performances en compression obtenues avec le dictionnaire formé des 15 bases de bandelettes et le dictionnaire formé des 12 bases de bandelettes directionnelles (c'est-à-dire sans la base de DCT et les bases de Haar) sont comparées aux performances obtenues avec le dictionnaire formé des 12 bases construites par ACP sur les 12 classes de blocs directionnels. Les résultats obtenus avec les bases construites par ACP sont environ 0,1 dB supérieurs aux résultats obtenus avec les 12 bases de bandelettes directionnelles. De plus, les résultats obtenus avec les 12 bases construites par ACP sont proches des résultats obtenus avec les 15 bases de bandelettes comprenant les 12 bases directionnelles, la base de DCT et les deux bases de Haar. Cependant, l'ajout des bases DCT et Haar au dictionnaire de bases construites par ACP n'augmente pas significativement les résultats en compression. Cela peut être dû au fait que les bases construites par ACP sont plus similaires à la DCT que les bases de bandelettes directionnelles. En conséquence, la base de DCT est plus efficace et plus souvent utilisée avec les bases de bandelettes qu'avec les bases construites par ACP. L'ajout d'autres bases construites par ACP donne les mêmes résultats parce que les bases ACP sont aussi similaires entre elles.

Comme il y a 964 coefficients non nuls dans les 15 bases de bandelettes et 3072 ($12 \times 16 \times 16$) coefficients non nuls dans les 12 bases ACP, la complexité calculatoire de la transformée en bandelettes pourrait être inférieure à la complexité calculatoire de la post-transformée dans les 12 bases ACP à condition d'avoir une implémentation efficace de la

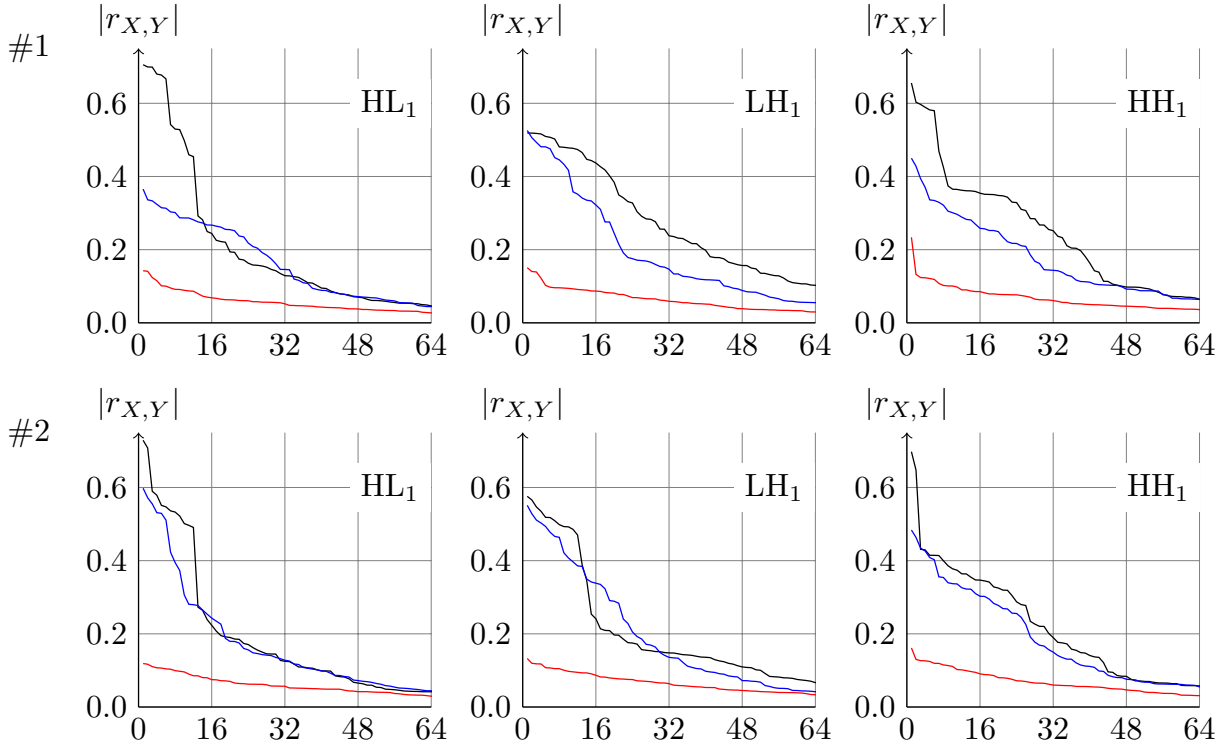


FIG. 4.17 – Les 64 plus forts coefficients de corrélation triés par ordre d’amplitude décroissante pour les classes de blocs 1 et 2 dans les sous-bandes HL_1 , LH_1 et HH_1 . Les courbes noires correspondent aux corrélations entre coefficients d’ondelettes, les courbes bleues aux corrélations entre coefficients de bandelettes et les courbes rouges aux corrélations entre coefficients dans les bases construites par ACP.

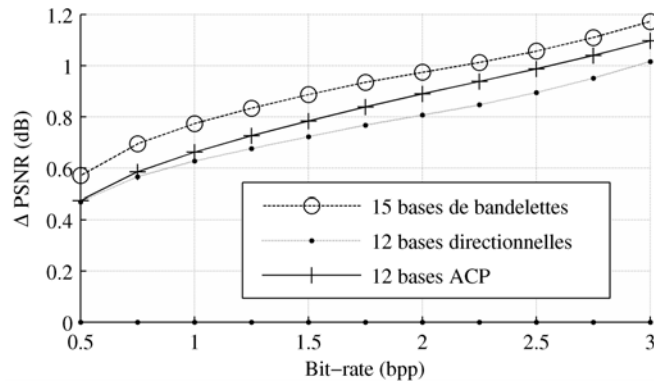


FIG. 4.18 – Qualités obtenues avec la compression par post-transformée avec 12 bases ACP par sous-bande, 12 bases directionnelles, et 15 bases de bandelettes, relativement aux ondelettes.

transformée en bandelettes. Néanmoins, il est possible d'augmenter les performances en compression obtenus avec les 12 bases ACP en exploitant leur propriété de compaction d'énergie dans le codeur entropique. Les bases de bandelettes n'ont pas cette propriété de compaction d'énergie.

De plus, en construisant plus de classes de blocs et donc plus de bases par ACP, il est aussi possible d'améliorer les résultats en compression. Cependant, la complexité augmente presque linéairement avec le nombre de bases dans le dictionnaire alors que le gain en PSNR progresse de plus en plus lentement.

D'autres dictionnaires pourraient aussi être créés à partir d'autres critères pour former les classes de blocs. Par exemple, des bases spécialisées pourraient être construites pour les zones forestières, les mers, les champs ou les zones urbaines.

4.4.3 Exploitation de la compaction d'énergie

Les bases construites par ACP dans les sections précédentes possèdent la propriété d'avoir leur vecteurs triés par contribution d'énergie décroissante. Nous allons montrer qu'il est possible d'exploiter cette propriété au moment du codage.

Principe de codage

La propriété de compaction d'énergie des bases ACP se traduit par des coefficients post-transformés forts, en valeur absolue et en moyenne, au niveau des vecteurs principaux et plus faibles, en valeur absolue et en moyenne, au niveau des autres vecteurs. La variance des coefficients post-transformés est plus forte au niveau des vecteurs principaux qu'au niveau des autres vecteurs. Cette différence statistique peut-être exploitée en codant séparément les coefficients associés aux différents vecteurs des bases ACP.

Un tel codeur a été mis en oeuvre. Il s'agit d'un codeur arithmétique adaptatif contextuel pour lequel un contexte correspond à une position particulière dans le bloc de 4×4 coefficients post-transformés. Un codage arithmétique est réalisé indépendamment pour chacune de ces 16 positions associées à différents niveaux d'énergie par l'ACP. Ainsi, lors du codage, une séparation des niveaux d'énergie est effectuée. Le but recherché est le même lorsque les sous-bandes d'une transformée en ondelettes sont codées indépendamment les unes des autres. De plus les coefficients des blocs non transformés sont codés par le codeur arithmétique indépendamment des coefficients des blocs transformés. Il y a donc $16 + 1$ contextes par sous-bande. Le principe de ce codeur est schématisé sur la figure 4.19.

Cette stratégie de codage permet de diminuer la taille du train binaire sans altérer la qualité. A 1,5 bpp, le débit est diminué de 0,5 %. A débit constant, ceci se traduit par une augmentation de la qualité de 0,1 dB en PSNR. La figure 4.20 présente le gain en qualité obtenu à débit constant grâce à l'exploitation de la propriété de compaction d'énergie des 12 bases construites par ACP précédemment. Pour des débits inférieurs à 2 bpp, les performances obtenues sont similaires à celles obtenues avec les 15 bases de bandelettes. La baisse progressive des performances obtenues avec le codeur schématisé sur la figure

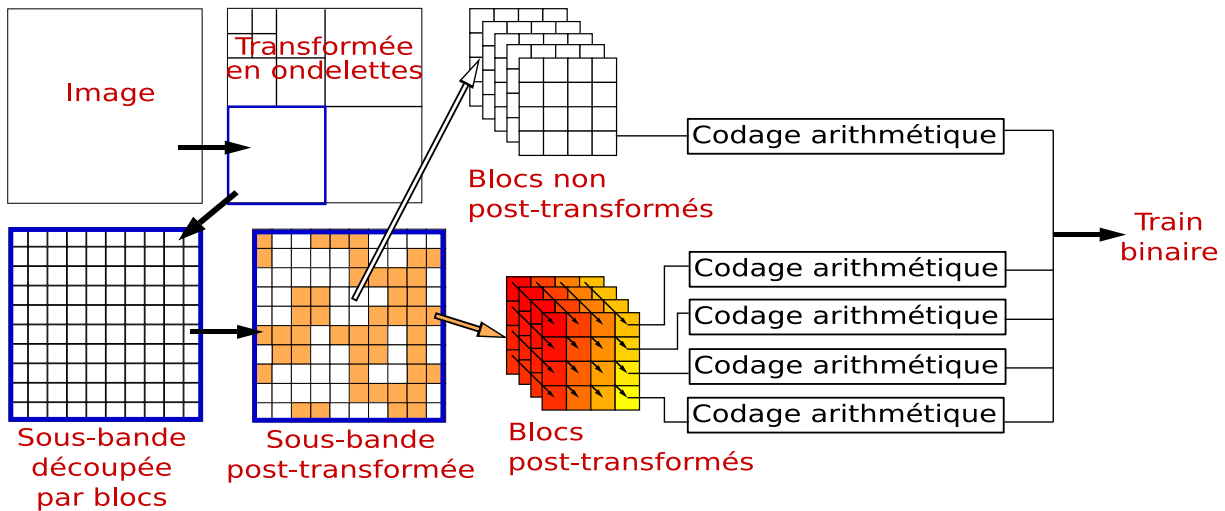


FIG. 4.19 – Schéma de principe d'un codeur arithmétique contextuel exploitant la propriété de compaction d'énergie dans les blocs transformés dans les bases construites par ACP.

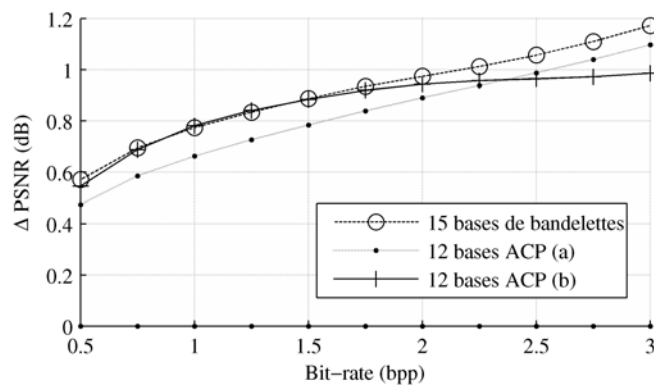


FIG. 4.20 – Performances de la compression par post-transformée avec 12 bases ACP par sous-bande (a) sans et (b) avec exploitation de la compaction d'énergie, et de la compression en bandelettes relativement aux ondelettes.

4.19 à haut débit s'explique par le fait que les pas de quantification sont de plus en plus faibles et qu'il y a donc un nombre croissant de valeurs quantifiées différentes. Il faut donc plus d'échantillons pour que les probabilités correspondant à l'état de chaque contexte du codeur arithmétique adaptatif soient estimées correctement.

Deux solutions pourraient être envisagées pour remédier à ce problème. La première serait de diminuer le nombre de contextes par exemple en les regroupant par 4 comme c'est le cas sur le schéma de la 4.19. La deuxième serait d'initialiser les probabilités associées aux contextes grâce à des valeurs apprises. Ces solutions n'ont pas été mises en oeuvre car nous préférons utiliser un codeur progressif. Néanmoins, ce codeur a permis de montrer qu'il est possible d'exploiter la compaction d'énergie obtenue grâce à la post-transformée dans des bases ACP. Cela permet aussi de réduire le nombre de bases du dictionnaire, et donc la complexité calculatoire, tout en gardant de bonnes performances en compression.

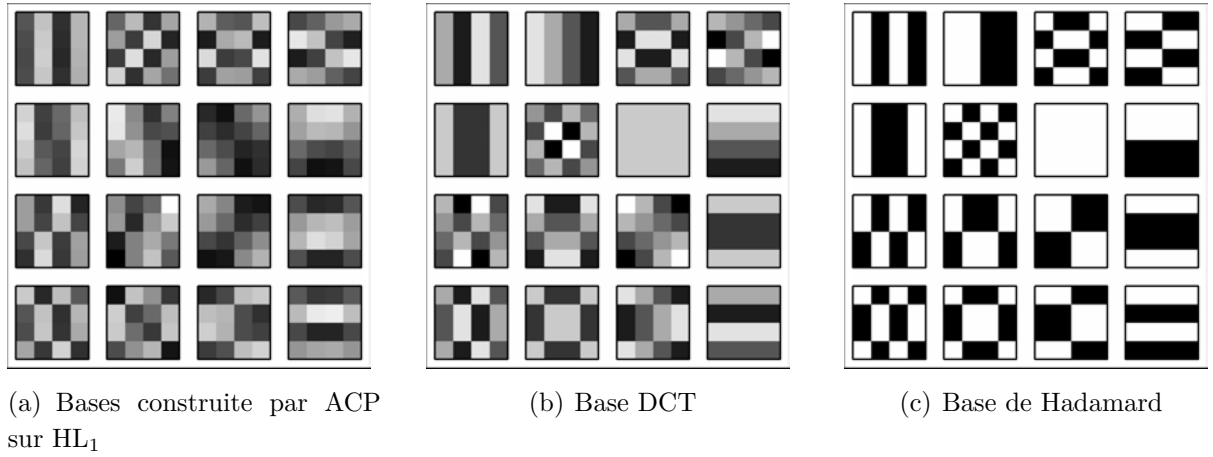


FIG. 4.21 – Base construite par ACP sur des blocs de coefficients d’ondelettes extraits de sous-bandes HL_1 et base de DCT et Hadamard dont les vecteurs ont été triés par contribution d’énergie décroissante pour la post-transformée de blocs de sous-bandes HL_1 .

Dans la section 5.2.3 nous exploiterons cette propriété de compaction d’énergie dans le cadre d’un codeur progressif par plans de bits.

Compaction d’énergie dans les bases de DCT et de Hadamard

Dans la section 4.4.1, nous avons vu que les performances obtenues par post-transformées avec une seule base ACP par sous-bande sont très proches des performances obtenues par post-transformées avec la base de DCT seule. La mesure de dissimilarité définie dans l’équation 4.2 entre ces deux bases est faible. Par exemple, elle vaut 11,8 entre une base construite par ACP sur les sous-bandes HL_1 et la base de DCT. Une autre base ressemble aussi à ces deux bases : la base de Hadamard. La mesure de dissimilarité vaut 5,3 entre la base de Hadamard et la base de DCT et 12,6 entre la base de Hadamard et la base construite par ACP sur les sous-bandes HL_1 .

La figure 4.21(a) présente la base construite par ACP sur des blocs issus de sous-bandes HL_1 de la transformée en ondelettes. Cette base est comparée à la base de DCT pour laquelle les vecteurs ont été réorganisés par contribution d’énergie décroissante dans les sous-bandes HL_1 (figure 4.21(b)). La base DCT est bien similaire à la base construite par ACP. L’avantage de la DCT par rapport à la transformée dans une base construite par ACP est qu’elle peut être calculée en utilisant des algorithmes rapides bien connus [81]. Lorsque les vecteurs de la base de Hadamard sont réorganisés (figure 4.21(c)), cette base ressemble aussi à la base construite par ACP. La transformée de Hadamard est encore moins complexe que la DCT puisqu’elle peut être effectuée par additions et soustractions des coefficients d’ondelettes sans aucune multiplication. Seule une normalisation par 4 est nécessaire et peut être réalisée par un décalage de 2 bits.

Sur la figure 4.21, les vecteurs de la base de DCT et de la base de Hadamard ont été triés par contribution d’énergie décroissante dans les sous-bandes HL_1 de l’ensemble d’images d’apprentissage. On peut observer que le vecteur de plus grande énergie (coin

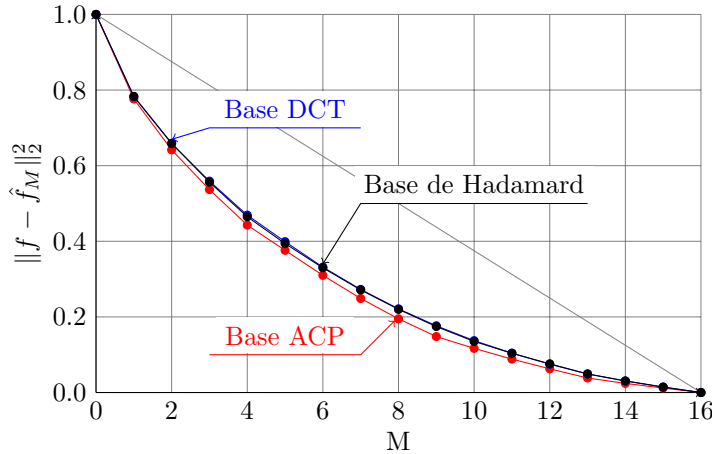


FIG. 4.22 – Vitesse d'approximation des coefficients d'ondelettes de la sous-bande HL_1 dans les bases ACP, DCT et de Hadamard. $\hat{f}_M$ représente le bloc de coefficients d'ondelettes f reconstruit à partir des M coefficients post-transformés les plus énergétiques.

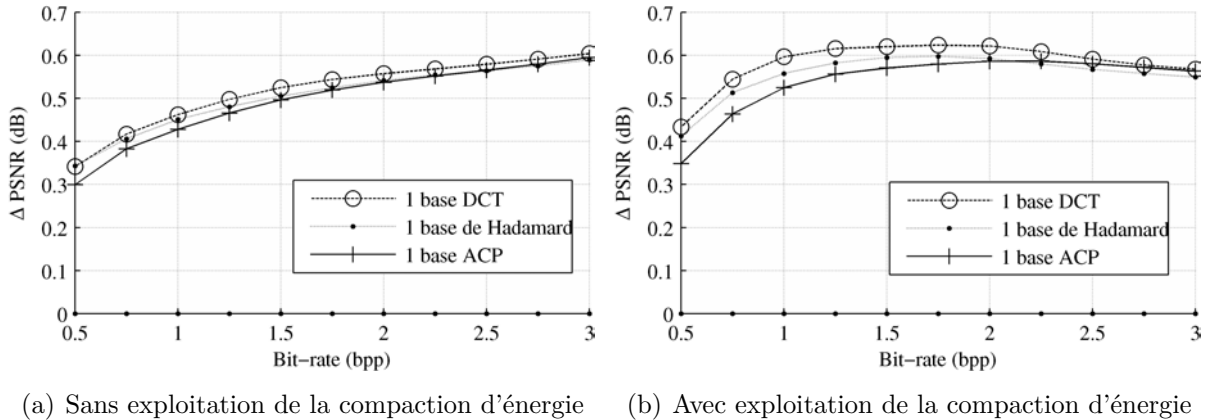


FIG. 4.23 – Performances en compression par post-transformée avec les bases DCT, de Hadamard et une base ACP par sous-bande (a) sans et (b) avec exploitation de la compaction d'énergie.

supérieur gauche de la base) présente un motif basse fréquence verticale et haute fréquence horizontale. Au contraire, le vecteur de plus faible énergie (coin inférieur droit de la base), présente un motif haute fréquence verticale et basse fréquence horizontale. Ces motifs correspondent bien avec le sens des filtres d'ondelettes passe-haut et passe-bas. La propriété de compaction d'énergie peut donc aussi être exploitée avec la base de DCT et la base de Hadamard.

La figure 4.22 présente la vitesse d'approximation des blocs de coefficients d'ondelettes des sous-bandes HL_1 dans la base ACP, la base DCT et la base de Hadamard. La vitesse d'approximation est similaire dans ces trois bases. L'avantage d'une base construite par ACP est donc limitée : les deux autres bases requièrent une complexité calculatoire plus faible et peuvent être aussi efficaces puisqu'elles possèdent aussi la propriété de compaction d'énergie.

La figure 4.23 présente les résultats en compression obtenus sur l'ensemble d'images

de test présenté en annexe A. Les résultats en compression en utilisant les dictionnaires contenant une seule post-transformée, i.e. la post-transformée dans la base ACP, dans la base DCT ou dans la base de Hadamard sont comparés relativement aux résultats de la transformée en ondelettes. Ces résultats ont été obtenus par simple codage arithmétique des coefficients transformés, sans exploiter la compaction d'énergie 4.23(a). Les résultats en compression obtenus avec les trois bases sont similaires. Les performances moyennes des bases construites par ACP sont légèrement inférieures aux performances moyennes de la DCT et de la base de Hadamard. Cela est dû au fait que les bases ACP ont été construites sur un ensemble d'images différent de celui qui a été utilisé pour la compression. En fait sur la plupart des images de l'ensemble utilisé pour la compression, les performances de la post-transformée dans les bases ACP sont légèrement meilleures que les performances de la post-transformée dans la base de DCT ou la base de Hadamard. Il n'y a que peu d'images de l'ensemble de compression sur lesquelles les performances de la post-transformée dans les bases ACP sont significativement en deçà des autres. Ceci conduit à des performances moins bonnes en moyenne.

La figure 4.23(b) montre que les performances obtenues en exploitant la propriété de compaction d'énergie de la base DCT, de la base de Hadamard et des bases construites par ACP sont similaires dans les trois cas. Le gain apporté par l'exploitation de la compaction d'énergie est légèrement plus important pour la DCT que pour les deux autres bases.

Si les contraintes de complexité calculatoire ne permettent l'utilisation que d'une seule base pour la post-transformée, alors il n'est pas recommandé d'utiliser des bases construites par ACP sur chaque sous-bande à moins que l'ensemble d'images d'apprentissage soit suffisamment représentatif de l'ensemble d'images compressé. Avec une seule post-transformée, la base de Hadamard est conseillée puisqu'elle est la moins complexe et que les performances en compression sont similaires aux autres bases. De plus, nous avons montré que la propriété de compaction de l'énergie peut aussi être exploitée dans cette base.

4.5 Conclusion

D'après cette étude, la prise en compte de la géométrie par post-transformée directionnelle ne profite pas particulièrement aux performances en compression dans une application satellite. La constitution du dictionnaire de départ semble devoir d'avantage être guidée par des considérations statistiques telles que la décorrélation et éventuellement l'indépendance. Dans cette optique, nous avons envisagé d'utiliser l'Analyse en Composantes Indépendantes (ACI) pour construire de nouvelles bases. Le premier problème qui se pose alors est que l'ACI ne fournit pas *a priori* des vecteurs orthogonaux. Certains algorithmes rapides d'ACI sont basés sur la minimisation de l'information mutuelle [82]. Pour une application en compression, on cherchera plutôt à minimiser un critère prenant en compte l'information mutuelle et l'écart par rapport à l'orthogonalité [83]. Une autre approche consiste à construire un dictionnaire permettant une représentation parcimonieuse. Nous avons utilisé la méthode décrite dans [84] qui permet de construire un

dictionnaire de bases orthonormales avec une contrainte de parcimonie. Malheureusement, les résultats en compression obtenus avec une base apprise par sous-bande sont similaires à ceux obtenus avec une base ACP.

La construction d'un dictionnaire de bases de post-transformées permettant d'optimiser la compression est donc un problème difficile. Il semble que ce problème soit particulier à la distribution statistique des blocs de coefficients d'ondelettes dans l'espace $\mathbb{R}^{16}$. Cette distribution n'est pas uniforme et nous avons montré dans ce chapitre qu'un dictionnaire formé de bases « distantes » entre elles ne permet pas d'améliorer les résultats en compression. Bien qu'elles soient performantes du point de vue de la compression, nous avons aussi montré que les bases de bandelettes directionnelles proposées par Peyré dans sa thèse [64] ne permettent pas d'exploiter toutes les redondances entre coefficients d'ondelettes dans les blocs 4×4 . En particulier, nous avons identifié des corrélations entre coefficients d'ondelettes qui ne sont pas exploitées par les bases directionnelles.

Pour exploiter ces corrélations nous avons étendu les groupements des bases de bandelettes directionnelles et construit des bases par ACP. Les corrélations entre coefficients post-transformés dans ces bases sont inférieures aux corrélations entre coefficients de bandelettes. De plus, les résultats en compression sont légèrement améliorés par rapport aux bases de bandelettes directionnelles. Nous avons aussi défini un schéma de compression basique permettant de montrer qu'il est possible d'exploiter la propriété de compaction d'énergie des bases construites par ACP. Les résultats en compression obtenus avec la post-transformée dans 12 bases construites par ACP sont alors similaires aux résultats obtenus avec les 15 bases de bandelettes. Cette approche permet de diminuer le nombre de bases du dictionnaire.

Enfin, nous avons montré que la post-transformée dans la base de DCT et la base de Hadamard, moins complexes que la post-transformée dans des bases construites par ACP, sont aussi performantes pour la compression. Elles permettent une amélioration non négligeable de plus de 0,5 dB en PSNR par rapport à la transformée en ondelettes. De plus, elles possèdent aussi la propriété de compaction d'énergie qui peut être exploitée au moment du codage. C'est pourquoi nous choisissons la post-transformée de Hadamard pour constituer un schéma de compression de faible complexité dans le chapitre suivant.

Chapitre 5

Post-transformées et codage par plans de bits

5.1	Sélection de la post-transformée pour un codage par plans de bits	108
5.1.1	Position du problème	108
5.1.2	Sélection de la post-transformée par minimisation l_1	109
5.1.3	Performances en compression	110
5.1.4	Conclusion	112
5.2	La post-transformée dans le codeur du CCSDS	112
5.2.1	Post-transformées et arbres de zéros	112
5.2.2	Choix du dictionnaire de post-transformées	113
5.2.3	Post-transformée des petits-enfants	113
5.2.4	Post-transformée des enfants	117
5.2.5	Perspectives de modification de l'algorithme du CCSDS	118
5.3	Conclusion	119

Dans les chapitres précédents, nous avons étudié la compression par post-transformées avec différents dictionnaires de bases. Nous avons vu que la post-transformée dans la base de Hadamard permet d'obtenir de bons résultats en compression avec une faible complexité. De plus, elle possède des propriétés de compaction de l'énergie. Dans les chapitres précédents, la compression était effectuée par quantification puis codage arithmétique des coefficients quantifiés. Le train binaire généré n'était donc pas emboîté. Dans ce chapitre, nous étudions la possibilité d'adapter un codeur progressif par plans de bits pour le codage des post-transformées. Nous avons publié une partie de cette étude dans [47]. Dans la section 5.1, nous commençons par introduire les problèmes posés par une compression progressive pour le choix de la post-transformée ce qui nous amène à modifier le critère de sélection de la post-transformée sur chaque bloc. Dans la section 5.2, nous adaptons le codeur du CCSDS pour l'utiliser avec la post-transformée de Hadamard dans un algorithme de compression ayant une faible complexité calculatoire.

5.1 Sélection de la post-transformée pour un codage par plans de bits

5.1.1 Position du problème

Dans un algorithme de compression produisant un train binaire emboîté, comme le codage est effectué par plans de bits, le pas de quantification q n'est pas défini. Ceci pose problème dans le processus de sélection de la post-transformée. En effet, la post-transformée sélectionnée dépend fortement du pas de quantification q : le critère débit-distorsion (3.3), le calcul de la distorsion $D(f_q^b)$, l'estimation du débit $R(f_q^b)$ et le multiplicateur de Lagrange $\lambda = \alpha q^2$ dépendent de ce pas de quantification.

Une première solution pour contourner ce problème est de fixer le paramètre q à une valeur correspondant au débit cible. L'unique post-transformée de chaque image est ensuite compressée à différents débits en faisant varier le pas de quantification. Dans ce paragraphe, nous notons q' le pas de quantification utilisé pour la quantification des coefficients post-transformés. Il est différent du paramètre q qui intervient dans le critère débit-distorsion pour la sélection des post-transformées sur chaque bloc. Les résultats en compression obtenus en fixant le paramètre q à 32 lors du choix de la post-transformée sont présentés sur la figure 5.1. Les différents débits ont été obtenus par codage arithmétique des coefficients post-transformés, quantifiés avec différents pas de quantification q' . La post-transformée de chaque image est fixe et ne dépend pas du pas de quantification q' . La figure 5.1 présente les résultats moyens obtenus sur les six images de l'ensemble de test représenté en annexe A.

Les meilleures performances ont été obtenues entre 2 bpp et 2,25 bpp ce qui correspond aux débits obtenus avec le pas de quantification $q' = 32$. Avec ce pas de quantification, les performances en compression obtenues sur chaque image sont identiques aux performances obtenues avec l'algorithme original. En effet, dans ce cas là, la post-transformée est adaptée à la quantification : $q' = q$. Cependant, comme le débit obtenu avec le pas de quantification

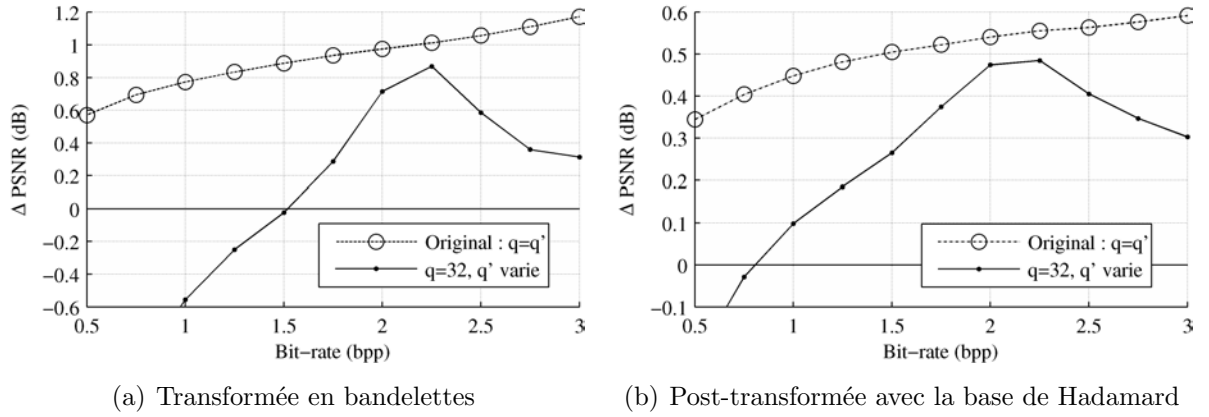


FIG. 5.1 – Comparaison des résultats en compression obtenus avec le paramètre q fixé à 32 lors du choix de la post-transformée avec les résultats de l’algorithme original dans lequel le paramètre q pour le choix de la post-transformée est toujours égal au pas de quantification q' .

$q' = 32$ n’est pas le même sur chaque image, les performances obtenues à débit fixe, moyennées sur toutes les images ne sont pas aussi bonnes que les performances originales moyennes à ce débit. Cette solution ne donne donc pas entière satisfaction puisque le débit obtenu pour une même valeur de q' est variable d’une image à l’autre et que les performances ne sont pas bonnes à tous les débits.

Une autre idée est de fixer ce pas de quantification q en fonction du débit cible R_T . A haut débit, il est possible de faire l’hypothèse haute résolution. Comme nous l’avons vu dans la section 3.3.2, le pas de quantification et le débit sont alors reliés par l’équation (3.16) : $R_T \approx h(X) - \log_2 q$, où $h(X)$ est l’entropie différentielle de la source X . Cependant, l’entropie $h(X)$ est difficile à obtenir et varie d’une image à l’autre. De plus, cette solution n’est pas optimale si on souhaite décoder l’image à un débit différent de R_T comme cela est permis par un train binaire emboîté.

5.1.2 Sélection de la post-transformée par minimisation l_1

La solution retenue est la modification du critère de sélection de la post-transformée afin que ce critère ne dépende plus du pas de quantification. Le but est de sélectionner la base b dans laquelle la représentation post-transformée f^b du bloc f minimise à la fois le débit et la distorsion. Pour minimiser le débit nécessaire au codage de la représentation f^b , il est souhaitable que cette représentation ait beaucoup de coefficients de faible amplitude. En effet, ces coefficients ne deviendront significatifs que dans les plans de bits les moins significatifs. Ils ne coûtent donc pas cher à coder avant que ces plans de bits soient atteints par le processus de codage, c’est-à-dire, à la fin de la formation du train binaire si le processus de codage est effectué jusqu’aux plans de bits les moins significatifs. Bien sûr, le débit cible R_T peut avoir été atteint avant et le processus de codage stoppé avant les plans de bits dans lequel ces coefficients deviennent significatifs. Ceci permet de diminuer fortement le coût de codage de ces coefficients. Au contraire, pour minimiser la distorsion,

il faut qu'un maximum de l'énergie du bloc ait été transmise au décodeur. La compaction d'énergie doit donc être favorisée : un maximum d'énergie du bloc doit être concentrée sur un minimum de coefficients. Ces coefficients seront significatifs et encodés dès les plans de bits les plus significatifs. Une grande partie de l'énergie du bloc sera donc disponible dès le début du décodage.

Ceci s'assimile à la recherche de la représentation la plus parcimonieuse et nous choisissons d'utiliser le critère de parcimonie donné par la norme l_1 . La norme l_1 de la représentation post-transformée f^b est simplement la somme des amplitudes des coefficients de cette représentation f^b :

$$\|f^b\|_1 = \sum_{m=1}^M |a_b[m]|$$

avec $a_b[m] = \langle f, \phi_m^b \rangle$. La post-transformée est sélectionnée par minimisation de ce critère selon l'équation (5.1).

$$b^* = \arg \min_{b \in [0, N_B]} \|f^b\|_1 \quad (5.1)$$

Notons que ce critère ne dépend pas du pas de quantification q . En minimisant la somme de l'amplitude des coefficients, ce critère favorise les représentations post-transformées qui ont beaucoup de coefficients d'amplitude faible. Le débit nécessaire au codage de la représentation sélectionnée f^{b^*} est donc minimisé. Comme les post-transformées sont orthonormales, l'énergie des représentations post-transformées f^b est la même que l'énergie du bloc de 4×4 coefficients d'ondelettes f . La recherche de la représentation f^b ayant la plus faible norme l_1 est donc équivalente à la recherche de la représentation f^b dont la variance de la valeur absolue des coefficients est maximale. Les représentations ayant des coefficients de grande amplitude sont donc sélectionnées et la distorsion est minimisée puisqu'une grande partie de l'énergie est restaurée dans les premiers plans de bits décodés. Enfin, ce critère est moins complexe que le critère débit-distorsion original qui nécessitait l'estimation du débit des représentations post-transformées. Cette estimation était coûteuse du point de vue calculatoire et n'était pas précise parce qu'elle était basée sur l'estimation des probabilités $\Pr(Q_q(a_b[m]))$ à partir de l'histogramme des coefficients d'ondelettes dans chaque sous-bande de la transformée en ondelettes (cf. chapitre 3).

5.1.3 Performances en compression

Sur la figure 5.2, les résultats obtenus avec ce nouveau critère de sélection de la post-transformée sont comparés aux résultats obtenus par minimisation du critère débit-distorsion. Les résultats présentés sur ces figures sont des performances moyennes obtenues par compression des images de l'ensemble de test. Le processus d'évaluation adaptative du coût de la signalisation R_b décrit dans la section 3.4.3 a été utilisé avec le critère débit-distorsion. La signalisation de la post-transformée sélectionnée sur chaque bloc est insérée dans l'en-tête du train binaire.

Les résultats obtenus sur la figure 5.2(a) avec la minimisation de la norme l_1 pour la sélection des bases de bandelettes sont catastrophiques à bas débit et de l'ordre de 0,6 dB en PSNR inférieurs aux résultats obtenus par minimisation du critère débit-distorsion

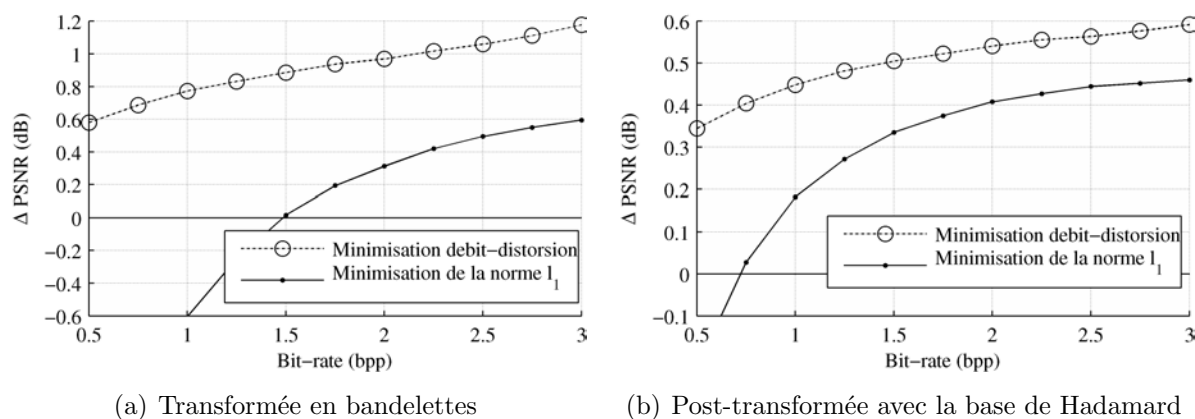


FIG. 5.2 – Comparaison des résultats en compression obtenus avec le critère débit-distorsion ou le critère de la norme l_1 .

pour des débits supérieurs à 2 bpp. *A priori*, la minimisation de la norme l_1 ne paraît donc pas un bon choix pour la sélection des bases de bandelettes.

Pourtant les résultats obtenus sur la figure 5.2(b) avec la minimisation de la norme l_1 pour la sélection de la post-transformée dans la base de Hadamard ne sont pas si catastrophiques. Les performances obtenues ne sont inférieures aux performances des ondelettes que pour des débits inférieurs à 0,7 bpp et ils sont de l'ordre de 0,1 dB en PSNR inférieurs aux résultats obtenus par minimisation du critère débit-distorsion pour des débits supérieurs à 2 bpp. La baisse des performances est surtout due au coût de codage de la signalisation de la post-transformée comme le montrent les mauvais résultats à bas débit. Le coût de la signalisation est beaucoup plus faible dans le cas d'un dictionnaire constitué d'une seule base que dans le cas d'un dictionnaire de 15 bases. C'est pour cela que les performances sont moins impactées avec la post-transformée dans la base de Hadamard.

Avec le critère de minimisation de la norme l_1 , la post-transformée sélectionnée ne dépend plus du débit final puisqu'elle ne dépend pas du pas de quantification. Les bases sélectionnées sont donc les mêmes quel que soit le débit final et le coût de la signalisation est constant. Avec ce critère, nous avons constaté expérimentalement qu'environ un bloc sur deux est post-transformé dans la base de Hadamard et un bloc sur deux n'est pas post-transformé. Le coût de signalisation est donc de 1 bit par bloc soit 0,0615 bpp lorsque la post-transformée est effectuée sur les trois niveaux de la décomposition en ondelettes. Après codage arithmétique adaptatif le débit de signalisation reste le même : les blocs post-transformés sont répartis aléatoirement dans les sous-bandes. Ce coût est significatif à bas débit. A 0,5 bpp, cela représente 12,3 % du débit total. A 3 bpp, cela ne représente plus que 2,1 % du débit total. Au contraire, avec le critère débit-distorsion, la sélection de la post-transformée dépend du pas de quantification utilisé et donc du débit. La minimisation de ce critère débit-distorsion a pour effet d'ajuster le débit alloué au codage de la signalisation avec le débit alloué au codage des coefficients post-transformés. A 3 bpp, le coût moyen de codage de la signalisation est toujours proche de 1 bit par bloc. A 0,5 bpp, ce coût est réduit à 0,3 bit par bloc soit 0,019 bpp ce qui représente seulement 3,8 % du débit total.

5.1.4 Conclusion

Le critère de minimisation de la norme l_1 permet de sélectionner les post-transformées indépendamment du pas de quantification et est donc adapté à une compression progressive par plans de bits. Ce critère donne des résultats satisfaisants à haut débit mais ne permet pas d'obtenir de bons résultats à bas débit à cause du coût de codage de la signalisation qui est incluse dans l'en-tête du train binaire. Cependant, lorsque le coût de la signalisation n'est pas pris en compte dans le calcul des débits, nous avons constaté que les qualités obtenues avec la minimisation de la norme l_1 sont proches de celles obtenues avec la minimisation du critère débit-distorsion, en particulier à bas débit. Le problème pour utiliser ce critère de minimisation de la norme l_1 réside donc dans le coût trop important de la signalisation à bas débit.

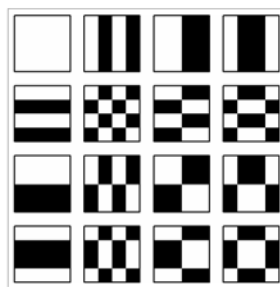
Ce problème peut être résolu en envoyant l'information de signalisation non pas en en-tête du train binaire, mais progressivement. Cela est possible dans le cadre d'une compression par plan de bits. En effet, l'information sur la base sélectionnée pour la post-transformée n'est pas nécessaire tant qu'aucun coefficient du bloc n'est significatif. A bas débit, peu de blocs ont des coefficients significatifs. Le coût de la signalisation peut donc être fortement réduite. Dans la suite, nous insérons la signalisation dans le train binaire emboîté produit par un codeur par plans de bits dès qu'un coefficient d'un bloc est trouvé significatif.

5.2 La post-transformée dans le codeur du CCSDS

Les post-transformées permettent d'exploiter les relations de dépendance entre coefficients d'ondelettes voisins. Ce faisant, elles détruisent les structures de l'image, comme les contours, qui étaient encore présentes dans la transformée en ondelettes. Nous avons vu dans le chapitre 2 que les codeurs morphologiques exploitent ces structures. L'utilisation des post-transformées est donc incompatible avec ces codeurs. Le codeur EBCOT exploite aussi les dépendances intra-bande entre coefficients d'ondelettes voisins. Comme cette tâche a été confiée à la post-transformée, nous pensons que le codeur EBCOT ne donnera pas d'aussi bons résultats pour le codage des coefficients post-transformés que pour le codage des coefficients d'ondelettes. Pour le codage progressif de la post-transformée, nous choisissons donc d'utiliser le codeur par plan de bits du CCSDS [29, 30] qui est dérivé des codeurs par arbres de zéros et qui possède l'avantage d'avoir été conçu spécialement pour le codage à bord des satellites.

5.2.1 Post-transformées et arbres de zéros

Bien que les post-transformées soient susceptibles de détruire les arbres de zéros puisqu'elles modifient les coefficients indépendamment dans les sous-bandes, nous allons tester si le gain en compression qu'elles apportent est supérieur à la perte de performance due à la destruction des arbres de zéros. Dans [85], le gain de codage dû aux dépendances entre parents et enfants dans les codeurs par arbres de zéros a été quantifié expérimentalement.

FIG. 5.3 – La base de Hadamard dans $\mathbb{R}^{16}$.

Une baisse de performance de 0,25 dB à 0,4 dB a été obtenue en détruisant les relations parents-enfants et en utilisant le codeur SPIHT [85]. Comme les gains en compression dûs aux post-transformées sont supérieurs à ces valeurs, il devrait être possible d’obtenir un gain en compression en insérant des post-transformées dans l’algorithme du CCSDS. De plus, nous avons vu dans la section 2.2.1 que le codeur de l’algorithme standard du CCSDS n’exploite pas les dépendances entre enfants et petits-enfants que nous risquons de détruire avec les post-transformées.

5.2.2 Choix du dictionnaire de post-transformées

Nous cherchons à limiter la complexité calculatoire du processus de post-transformée en vue d’obtenir un algorithme de compression suffisamment simple pour une implémentation matérielle et « qualifiée spatial » à bord des satellites. Pour cela, le dictionnaire de post-transformées est réduit à une seule base. De plus, nous choisissons une post-transformée peu complexe : la transformée de Hadamard. La base de Hadamard, représentée sur la figure 4.11(b), est reproduite sur la figure 5.3.

On rappelle que la transformée dans cette base peut être effectuée par additions et soustractions des coefficients d’ondelettes et la normalisation par 4 ne nécessite qu’un décalage de 2 bits. De plus, le gain en PSNR relativement aux ondelettes obtenu par la post-transformée de Hadamard seule (+0,56 dB à 2,5 bpp) est supérieur à 50 % du gain obtenu avec les 15 bases du dictionnaire de bandelettes (+1,06 dB à 2,5 bpp). Avec une complexité calculatoire très réduite, les performances en compression restent correctes. La base de DCT est aussi une bonne candidate car la DCT peut être calculée à l’aide d’algorithmes rapides [81]. Cependant, avec cette post-transformée les résultats sont similaires à ceux de la post-transformée avec la base de Hadamard.

5.2.3 Post-transformée des petits-enfants

Le codeur du CCSDS a été décrit dans la section 2.2.1. Dans cet algorithme, les petits-enfants forment des ensembles de 4×4 coefficients d’ondelettes. Aussi, nous pouvons appliquer les post-transformées sur ces groupes de coefficients dans les sous-bandes HL_1 , LH_1 et HH_1 . Les post-transformées ne sont donc appliquées que sur le premier niveau de résolution de la transformée en ondelettes des images.

Algorithme 5 : Signalisation de la post-transformée dans l'algorithme du CCSDS.

Étape 3 (petits-enfants) :

1. $tran_G$;
 2. PT_i , pour tout i tel que $t_{max}(G_i) = 1$ (signalisation de la post-transformée) ;
 3. $tran_{H_i}$, pour tout i tel que $t_{max}(G_i) \neq 0, -1$;
 4. $types_b[H_{ij}]$ et $signs_b[H_{ij}]$ pour tout i tel que $t_{max}(G_i) \neq 0, -1$ et tout j tel que $t_{max}(H_{ij}) \neq 0, -1$.
-

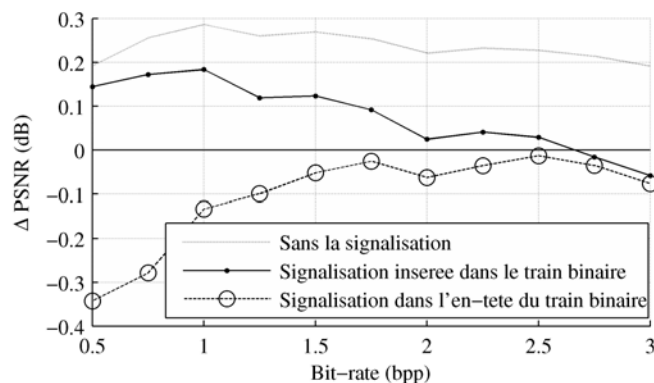


FIG. 5.4 – Performances en compression obtenues par utilisation de la post-transformée et codage avec le BPE.

Nous avons mentionné à la fin de la section 5.1 que le codage de la signalisation des post-transformées en en-tête du train binaire est une mauvaise stratégie. En effet, dans un codage par plans de bits, les coefficients de certains blocs sont susceptibles de ne devenir significatifs que dans les plans de bits les moins significatifs. Si le codage n'est pas effectué jusqu'à ces plans de bits, il n'est pas nécessaire de signaler les post-transformées utilisées pour ces blocs. Ainsi, la post-transformée ne doit être signalée que pour les ensembles de petits-enfants G_i qui ont au moins un coefficient significatif. Dans le cas d'une seule post-transformée dans le dictionnaire, cette information coûte 1 bit par bloc et n'est pas codée entropiquement. Nous avons vu dans la section 5.1.3 que lorsque la sélection de la post-transformée est effectuée par minimisation de la norme l_1 et que la seule post-transformée utilisée est la transformée de Hadamard, alors le coût moyen de codage de la signalisation est d'environ 1 bit par bloc même après codage arithmétique adaptatif. Ce bit indiquant la post-transformée utilisée sur chaque ensemble G_i significatif est inséré dans le train binaire directement après les mots de transition $tran_G$ qui ont au moins un bit valant 1 indiquant un nouvel ensemble de petits-enfants G_i significatif. Nous avons modifié l'algorithme du CCSDS (cf. [29] section 4.5.3.1.8) comme indiqué dans l'algorithme 5.

La figure 5.4 illustre les performances en compression obtenues avec la post-transformée de Hadamard suivie du codage avec le BPE. Ces performances sont tracées relativement aux performances de l'algorithme standard du CCSDS original (ligne 0 dB), c'est-à-dire avec la transformée en ondelettes seule. La courbe tracée en trait fin présente

les résultats obtenus lorsque la signalisation sur les post-transformées utilisées n'est pas transmise. Bien entendu, dans ce cas là, l'image n'est pas décodable. La courbe tracée en trait plein correspond aux performances obtenues par insertion de la signalisation dans le train binaire comme décrit précédemment. La courbe tracée en trait discontinu correspond aux performances obtenues avec toute la signalisation transmise comme en-tête du train binaire. A bas-débit, les résultats obtenus avec la stratégie d'insertion de la signalisation dans le train binaire sont très proches de ceux obtenus sans transmission de la signalisation. En effet, peu d'ensembles de petits-enfants G_i sont significants. Très peu d'information sur la signalisation est donc transmise. De plus, les ensembles de petits-enfants G_i qui sont significants à bas débit sont significants dans les plans de bits les plus significatifs et contribuent donc fortement à la réduction de la distorsion. A haut débit, les résultats obtenus sont proches de ceux obtenus lorsque la signalisation est transmise en en-tête : la plupart des ensembles de petits-enfants G_i sont significants et la quasi-totalité de la signalisation a été transmise.

Cependant, même si l'information de signalisation a été correctement insérée dans le train binaire, on peut observer sur la figure 5.4 que les résultats obtenus avec les post-transformées sont environ 0,1 dB meilleurs que ceux obtenus avec la transformée en ondelettes et même moins bons à haut débit. En effet, les post-transformées modifient la distribution d'énergie dans les ensembles G_i . Les petits-enfants post-transformés sont susceptibles de devenir significants n'importe où dans l'ensemble G_i . Cela influe sur les plages de zéros qui pouvaient être formées dans les groupes H_{ij} . Le mot de transition $tran_{H_i}$ perd alors son efficacité.

Il faut aussi noter que cette modification de la distribution de l'énergie a un impact sur les dépendances entre la signifiante des coefficients enfants et la signifiante des groupes H_{ij} . Cependant cela n'a pas d'impact sur le codage puisque les dépendances entre la signifiante des coefficients enfants et la signifiante des groupes H_{ij} ne sont pas exploitées dans l'algorithme du CCSDS. Comme cela a été mentionné dans le chapitre 2, le codeur du CCSDS n'est pas réellement un codeur par arbres de zéros comme EZW [45] et SPIHT [46]. Dans ces codeurs, la destruction de ces dépendances aurait eu un impact négatif sur le codage, bien que léger selon [85].

Il est possible de limiter la perte d'efficacité due à la modification de la distribution d'énergie en exploitant justement les propriétés énergétiques de la base de Hadamard. Nous avons vu dans la section 4.4.3 que la distribution moyenne de l'énergie sur les vecteurs de la base de Hadamard n'est pas uniforme et que cette distribution dépend de la sous-bande dans laquelle se trouvent les blocs de coefficients post-transformés. Ainsi, il est possible de trier les vecteurs de la base de Hadamard par contribution d'énergie décroissante. Ce tri est différent d'une sous-bande à l'autre. Il est alors possible d'augmenter les performances en compression en regroupant les vecteurs par contribution d'énergie afin que les petits-enfants d'un groupe H_{ij} aient à peu près la même énergie. *Cela permet de recréer artificiellement des plages de zéros dans les groupes H_{ij} de faible énergie.*

Les bases de Hadamard dans lesquelles l'ordre des vecteurs a été modifié selon la sous-bande post-transformée sont représentées sur la figure 5.5. Le groupement des quatre

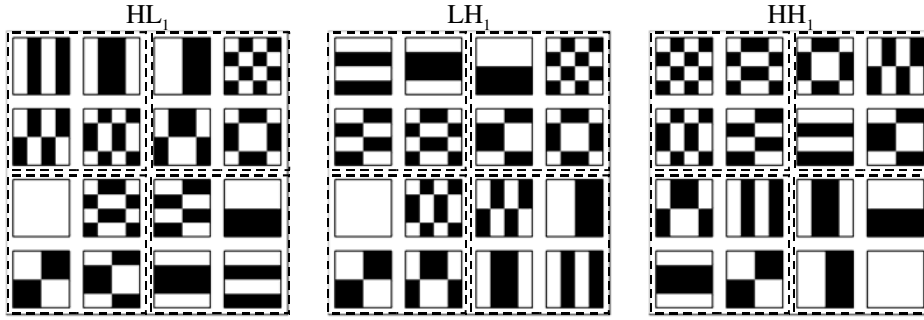


FIG. 5.5 – Bases de Hadamard dont les vecteurs ont été groupés par contribution d'énergie dans les sous-bandes HL_1 , LH_1 et HH_1 . Les vecteurs formant les groupes de petits enfants H_{ij} sont délimités par des traits discontinus.

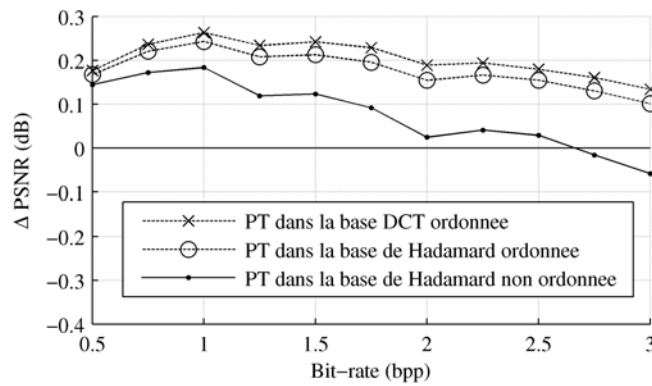


FIG. 5.6 – Comparaison des performances obtenues avec la post-transformée de Hadamard avec et sans ordonnancement des vecteurs. Les performances sont aussi comparées à celles obtenues avec la DCT.

vecteurs en haut à gauche de chaque base correspond aux vecteurs contribuant le plus à l'énergie totale. Bien que l'ordre des vecteurs soit différent dans chaque sous-bande, il faut remarquer que le coût de signalisation de la post-transformée n'a pas augmenté. Il est toujours de 1 bit par bloc puisque chaque bloc ne peut être post-transformé que dans une base : la base adaptée à la sous-bande dans laquelle il se trouve.

Les performances en compression obtenues en utilisant cette stratégie d'ordonnancement de l'énergie sont présentées sur la figure 5.6. Cette stratégie est surtout payante à haut débit, c'est-à-dire lorsqu'un maximum d'ensembles de petits-enfants G_i sont significants et post-transformés. A 3 bpp cela a permis un gain de 0,15 dB en PSNR et les performances à ce débit sont maintenant supérieures à celle de l'algorithme standard du CCSDS. Sur la figure 5.6, les performances obtenues avec la DCT sont également tracées. Cette modification ne permet qu'un gain de moins de 0,05 dB en PSNR par rapport à la post-transformée de Hadamard. Si une très faible complexité est requise, la post-transformée de Hadamard est un meilleur choix.

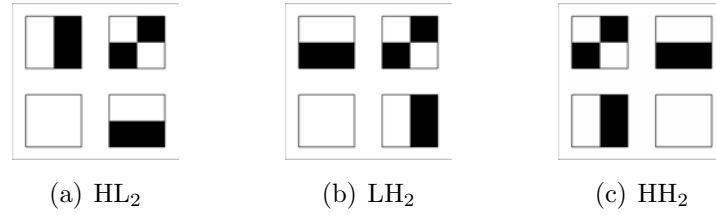


FIG. 5.7 – Bases de Hadamard dans $\mathbb{R}^4$ dont les vecteurs ont été ordonnés par contribution d'énergie dans les sous-bandes HL₂, LH₂ et HH₂.

5.2.4 Post-transformée des enfants

La post-transformée s'applique bien aux ensembles de 4×4 petits-enfants G_i . Pourtant, on peut envisager d'appliquer une post-transformée aux ensembles de coefficients enfants C_i . Pour cela, la base de Hadamard dans $\mathbb{R}^4$ peut être utilisée. De même que précédemment, les vecteurs de cette base peuvent être ordonnés par contribution d'énergie dans les différentes sous-bandes comme cela est représenté sur la figure 5.7.

Pour chaque famille F_i , $i \in [0, 2]$, on note c_{ij} avec $j \in [0, 3]$ les quatre coefficients enfants de l'ensemble C_i . L'idée est alors de renforcer les dépendances entre les coefficients post-transformés enfants c_{ij} et les coefficients post-transformés dans les groupes de petits-enfants H_{ij} . Par exemple, dans la sous-bande HL₂, l'amplitude moyenne des coefficients post-transformés enfants c_{00} sera élevée puisque ces coefficients sont obtenus par projections des coefficients d'ondelettes des ensembles C_0 sur le vecteur de plus forte énergie de la base de Hadamard dans la sous-bande HL₂ (figure 5.7). De même, les amplitudes moyennes des coefficients post-transformés petits-enfants dans l'ensemble H_{00} seront élevées puisque ces coefficients sont obtenus par projections des coefficients d'ondelettes des ensembles G_0 sur les quatre vecteurs de plus grande énergie de la base de Hadamard dans la sous-bande HL₁ (figure 5.5).

Pour renforcer ces dépendances, il faut post-transformer simultanément les coefficients enfants C_i et les coefficients petits-enfants G_i . Les enfants C_i et les petits-enfants G_i sont post-transformés dans la base de Hadamard si la somme des valeurs absolues des coefficients post-transformés dans C_i et G_i est inférieure à la somme des valeurs absolues des coefficients d'ondelettes dans C_i et G_i . Cela revient à choisir la post-transformée f^{b*} qui minimise la norme $l_1 \|f^b\|_1$ où f est la concaténation des ensembles de coefficients d'ondelettes de C_i et de G_i .

Nous définissons les ensembles de 5 coefficients K_{ij} par la concaténation d'un coefficient enfant noté c_{ij} et les quatre coefficients petits-enfants du groupe H_{ij} :

$$K_{ij} = \{c_{ij}, H_{ij}\}$$

Comme les coefficients des ensembles C_i et G_i ont été post-transformés simultanément et triés par ordre d'énergie, la dépendance statistique entre un coefficient c_{ij} et les coefficients du groupe H_{ij} est renforcée.

Cependant, nous avons vu que le codeur de l'algorithme standard du CCSDS n'exploite pas les dépendances entre les coefficients enfants et les coefficients petits-enfants. Pour

Algorithme 6 : Modification de l'algorithme du CCSDS.

- Étape 2 (enfants)
 1. $tran_B$;
 2. $tran_D$, si $tran_B \neq 0$ et $t_{max}(B) \neq -1$;
 3. PT_i , pour tout i tel que $t_{max}(D_i) = 1$ (signalisation de la post-transformée) ;
 4. $tran_{Ki}$ pour tout i tel que $t_{max}(D_i) \neq 0, -1$;
 5. $types_b[\{c_{ij}\}]$ et $signs_b[\{c_{ij}\}]$ pour tout i tel que $t_{max}(D_i) \neq 0, -1$ et tout j tel que $t_{max}(K_{ij}) \neq 0, -1$.
 - Étape 3 (petits-enfants)
 1. $tran_{Hi}$, pour tout i tel que $t_{max}(D_i) \neq 0, -1$;
 2. $types_b[H_{ij}]$ et $signs_b[H_{ij}]$ pour tout i tel que $t_{max}(D_i) \neq 0, -1$ et tout j tel que $t_{max}(H_{ij}) \neq 0, -1$.
-

tirer profit de ces nouvelles dépendances entre les coefficients enfants c_{ij} et les coefficients petits-enfants H_{ij} , il est nécessaire de modifier le codeur de l'algorithme standard du CCSDS.

5.2.5 Perspectives de modification de l'algorithme du CCSDS

Nous définissons les nouveaux mots de transition $tran_{Ki}$ avec $i \in [0, 2]$ permettant de signaler les ensembles K_{ij} avec $j \in [0, 3]$ dans lesquels aucun coefficient n'est significatif :

$$tran_{Ki} = tword[\{t_{max}(K_{i0}), t_{max}(K_{i1}), t_{max}(K_{i2}), t_{max}(K_{i3})\}] \text{ pour } i = 0, 1, 2.$$

De plus, nous modifions la définition des mots de transition $tran_{Hi}$:

$$tran_{Hi} = tword[\{t_{max}(H_{ij}) : j = 0, 1, 2, 3, \text{ tel que } t_{max}(K_{ij}) > 0 \text{ et } t_b(c_{ij}) > 0 \\ \text{dans le plan de bits courant ou tout autre plan de bits plus significatif}\}] \text{ pour } i = 0, 1, 2.$$

Désormais, les mots de transition $tran_{Hi}$ servent à indiquer quels sont les groupes de petits-enfants H_{ij} de la famille i qui deviennent significatifs au plan de bits codé. Il suffit d'indiquer les groupes H_{ij} pour lesquels $t_{max}(K_{ij}) > 0$. En effet, si $t_{max}(K_{ij}) \leq 0$, le mot $tran_{Ki}$ aura déjà indiqué qu'aucun coefficient du groupe $K_{ij} = \{c_{ij}, H_{ij}\}$ n'est significatif. De plus, si $t_{max}(K_{ij}) > 0$ et c_{ij} n'est pas significatif, il existe forcément un petit-enfant du groupe H_{ij} qui est significatif. Dans ce cas là, il n'est pas nécessaire de signaler que le groupe H_{ij} est significatif. Notons qu'avec cette modification, les mots de transition $tran_{Hi}$ peuvent désormais prendre la valeur $\{0, 0, 0, 0\}$ dans le cas où tous les enfants sont significatifs et aucun petit-enfant n'est significatif.

Nous avons choisi de ne pas conserver le mot de transition $tran_G$ car, après cette modification, il ne permet d'économiser des bits que dans les cas où beaucoup d'enfants sont significatifs et très peu de petits-enfants sont significatifs. Ces cas là sont peu probables. Les

étapes de codage 2 et 3 correspondant au codage des enfants et des petits-enfants (cf. [29] section 4.5.3.1.8) sont modifiées comme indiqué dans l'algorithme 6. Cette modification permet de prendre en compte les dépendances entre enfants et petits-enfants dans les ensembles K_{ij} avec les mots de transition $tran_{K_i}$. Un exemple de codage est donné en annexe B. La modification du codeur a pour effet de générer plus de mots binaires de longueur 1. Il risque donc d'y avoir une perte d'efficacité car ces mots de longueur 1 ne sont pas codés entropiquement. De plus, la modification de l'algorithme nécessite une étude statistique des mots binaires générés afin de déterminer les nouvelles tables pour le codage entropique de ces mots. Sur l'exemple donné en annexe B et sans le codage entropique, cette modification permet de diminuer le nombre de bits émis sur le codage du premier plan de bits mais pas sur les suivants. Sans la post-transformée, cette modification pourrait améliorer les résultats en compression à bas débit.

5.3 Conclusion

Pour adapter le schéma de compression par post-transformées à un codage progressif par plan de bits nous avons dû modifier le critère de sélection de la post-transformée sur chaque bloc. Nous choisissons la post-transformée minimisant la norme l_1 de la représentation du bloc. Ce critère est généralement utilisé pour la recherche de représentations parcimonieuses. Si on omet le coût de la signalisation, ce critère est quasiment aussi performant que le critère de minimisation débit-distorsion à bas débit et légèrement moins performant à haut débit. Néanmoins, il présente l'avantage d'être peu complexe du point de vue calculatoire et de ne pas dépendre du pas de quantification. On peut aussi envisager la recherche d'un autre critère permettant d'améliorer les résultats à haut débit et ne dépendant pas du pas de quantification.

La post-transformée avec la base de Hadamard sélectionnée par minimisation de la norme l_1 et codée par l'algorithme du CCSDS constitue un schéma de compression de faible complexité. Nous avons modifié l'algorithme du CCSDS afin d'introduire la signalisation des post-transformées utilisées dans le train binaire emboîté et non en en-tête. Cela permet d'obtenir de bons résultats en compression à bas débit. Nous avons aussi utilisé la propriété de compaction de l'énergie dans la base de Hadamard afin de créer artificiellement des plages de zéros qui sont bien exploitées par l'algorithme du CCSDS. Cela a permis de rehausser les performances à haut débit. La compression de la post-transformée par le codeur du CCSDS donne alors des résultats améliorés de l'ordre de 0,2dB en PSNR par rapport à la compression de la transformée en ondelettes. Le codage progressif des post-transformées est donc possible et améliore les résultats en compression.

Enfin, nous avons envisagé de renforcer artificiellement les dépendances inter-échelles avec les post-transformées et d'exploiter ces dépendances dans un algorithme du CCSDS modifié. Le principe est présenté dans la section 5.2.5 mais ces modifications n'ont pas encore été implantées.

Conclusions et perspectives

Au cours de cette thèse, nous avons étudié les possibilités offertes par les post-transformées pour la compression d'images. Ce travail couvre tous les aspects d'une chaîne de compression à bord, des dépendances entre composantes du signal jusqu'au codage progressif.

Nous avons d'abord mis en évidence les dépendances qui subsistent entre coefficients après transformée en ondelettes des images. Les post-transformées ont été créées dans le but de les exploiter par un deuxième niveau de transformation.

Nous avons ensuite optimisé plusieurs paramètres de la transformée en bandelettes par blocs à partir de la compression d'images satellites. Nous avons montré que la post-transformée de blocs de taille supérieure à 4×4 n'a pas d'intérêt puisqu'en général les corrélations ne sont significatives qu'entre coefficients d'ondelettes séparés par moins de deux coefficients. De plus, la compression avec des blocs de plus grande taille est moins performante. Nous avons aussi proposé une nouvelle formule théorique optimale sous hypothèse haut-débit du multiplicateur de Lagrange qui intervient dans le critère de sélection de la post-transformée. Cette formule s'est révélée être identique à celle qui était déjà utilisée à bas-débit. Nous avons alors vérifié ces formules empiriquement pour la compression d'images satellites en adaptant le processus d'optimisation de Shoham et Gersho à la problématique de sélection de la meilleure post-transformée. Nous en avons déduit une formule empirique du multiplicateur de Lagrange qui est presque identique aux formules théoriques et qui permet d'atteindre des performances en compression extrêmement proches des performances optimales. Nous avons aussi modifié l'estimation du débit de signalisation. Ce débit est maintenant estimé de façon adaptative et les résultats de l'estimation sont excellents. Cela permet de ne pas avoir à effectuer un apprentissage des probabilités de la sélection de chaque post-transformée. Les résultats en compression sont améliorés, en particulier à bas débit lorsque le dictionnaire ne contient qu'une seule post-transformée.

Une analyse des dépendances entre coefficients d'ondelettes exploitées par les transformées en bandelettes a permis de distinguer quelles sont les dépendances qui ne sont pas exploitées. Nous avons alors créé de nouvelles bases afin de mieux exploiter ces dépendances. Les bases construites par ACP ont l'avantage de minimiser les corrélations entre coefficients transformés. De plus, nous avons montré à l'aide d'un algorithme de compression que nous avons défini, qu'il est possible d'exploiter la compaction d'énergie obtenue grâce aux bases d'ACP. Avec douze bases construites par ACP, les performances

obtenues sont similaires à celles obtenues avec les 15 bases de bandelettes. Cependant, la complexité calculatoire est plus élevée pour ces bases ACP. Nous avons alors étudié deux post-transformées ayant une faible complexité calculatoire : la DCT et la transformée de Hadamard. Ces deux post-transformées sont similaires. Il est donc inutile de les utiliser dans le même dictionnaire. De plus, elles permettent toutes les deux une compaction de l'énergie.

Dans le but d'utiliser les post-transformées dans un schéma de compression progressive indispensable pour réguler facilement le débit à bord des satellites, nous avons modifié le critère de sélection de la post-transformée. Nous sélectionnons la post-transformée pour laquelle la norme l_1 est minimale. Ce critère présente l'avantage d'être moins complexe que le critère débit-distorsion. De plus, il ne dépend pas du taux de compression. Cependant, afin de conserver de bonnes performances en compression, la signalisation doit-être insérée dans le train binaire emboîté. Nous avons modifié l'algorithme du CCSDS pour y inclure la signalisation de la post-transformée et réordonné les vecteurs de la base de Hadamard afin que la compaction d'énergie soit correctement exploitée par le codeur du CCSDS. La post-transformée de Hadamard utilisée conjointement avec l'algorithme du CCSDS améliore la qualité de l'image d'environ 0,2 dB en PSNR par rapport à la transformée en ondelettes. Nous pensons pouvoir améliorer ces résultats en effectuant la post-transformée simultanément sur deux niveaux de la transformée en ondelettes et en modifiant le codeur du CCSDS afin d'exploiter les dépendances inter-échelles créées par la compaction d'énergie.

Les résultats obtenus avec les post-transformées sont donc supérieurs aux résultats de la transformée en ondelettes dans le cas d'un simple codage arithmétique des coefficients quantifiés comme dans le cas d'un codage progressif par plans de bits. Cependant, les bases de post-transformées n'ont pas l'aspect géométrique des bases de bandelettes. Nous pensons que la détection et l'exploitation de la géométrie des structures dans un schéma de compression efficace n'est possible qu'à très bas débit. Dans ce cas là, la qualité visuelle est meilleure en préservant les contours. Pour obtenir les niveaux de qualité requis par les images satellite, les gains en compression sont obtenus en exploitant les dépendances liées aux textures car ce sont elles qui constituent la plus grande partie de l'image. Bien que le calcul de la meilleure décomposition en paquets d'ondelettes ne soit pas possible dans un schéma de compression au fil de l'eau, il serait aussi intéressant de comparer les performances des post-transformées aux performances de la meilleure décomposition en paquets d'ondelettes dans un schéma de compression par plans de bits. D'autres bases de post-transformées pourraient aussi être créées à partir des méthodes par ACI en veillant à prendre en compte les contraintes d'orthogonalité ou bien à partir des méthodes d'apprentissage de représentations parcimonieuses dans des bases. Bien que positif, le gain en compression obtenu avec les post-transformées est insuffisant pour répondre aux besoins élevés de la compression à bord. Le CNES envisage une division par deux du débit moyen. Pour cela il étudie parallèlement des méthodes de compression sélective qui font intervenir des techniques de détection de nuages et de codage par régions d'intérêt. Ces derniers éléments sont les perspectives principales de ce travail. Une étude sur l'intégration de

régions d'intérêt dans le codeur du CCSDS sera d'ailleurs lancée chez Noveltis pour le CNES dans le prolongement de cette thèse.

Annexe A

Images

Les images utilisées pour les tests en compression au cours de cette thèse sont des images d'observation de la Terre issues de deux capteurs différents : PLEIADES et PELICAN. Comme les lancements des deux satellites PLEIADES ne sont prévus que pour 2010 et 2011, les images PLEIADES sont simulées à partir d'images issues du capteur aéroporté RAMI. La résolution au sol des images RAMI est de 60 cm par pixel. Pour simuler la résolution de PLEIADES, la résolution de ces images a été réduite à 70 cm par pixel. Les images acquises avec le capteur aéroporté PELICAN ont une résolution au sol de 25 cm par pixel. Dans les deux cas, les images sont en gamme de gris (acquisition panchromatique) avec 4096 niveaux soit 12 bits par pixel. Les segments d'images de taille 1024×1024 représentés ci-dessous ont été utilisés pour les tests en compression et les apprentissages statistiques.

A.1 Ensemble d'images de test



PLEIADES Perpignan 1



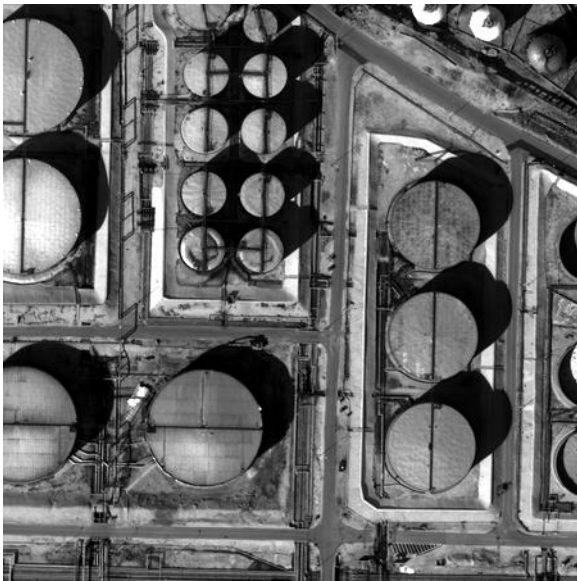
PLEIADES Port de Bouc 1



PLEIADES Toulouse 1



PELICAN Port de Bouc 1

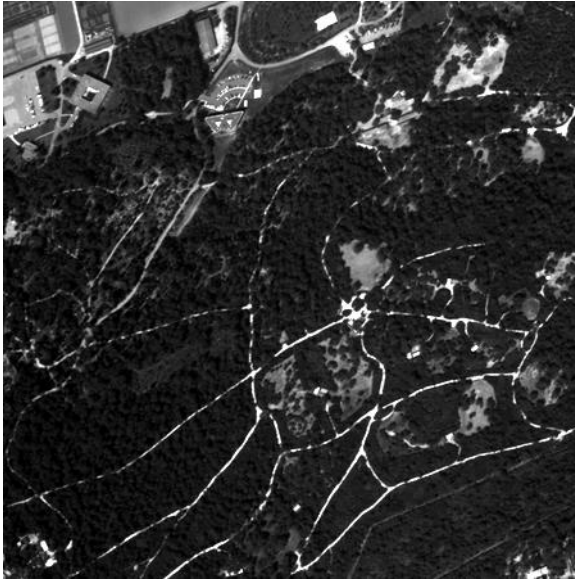


PELICAN Port de Bouc 4



PELICAN Toulouse 3

A.2 Ensemble d'images d'apprentissage



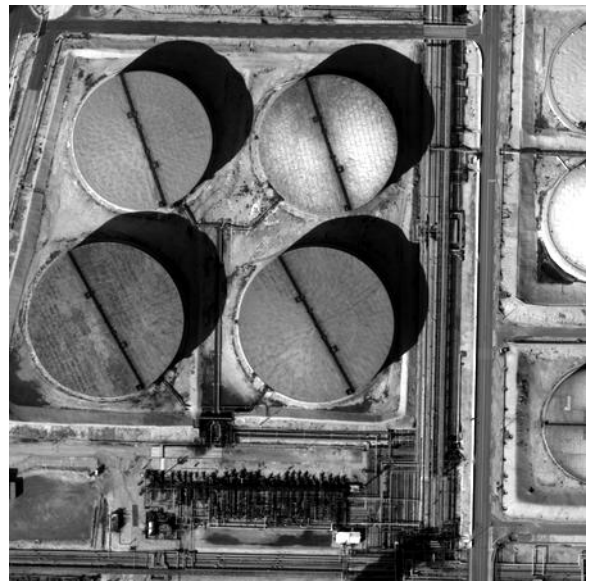
PLEIADES Montpellier



PLEIADES Perpignan 2



PLEIADES Port de Bouc 2



PELICAN Port de Bouc 2



PELICAN Port de Bouc 3



PELICAN Toulouse 1



PELICAN Toulouse 2

Annexe B

Exemple de codage avec le BPE du standard CCSDS

Le codage par plans de bits du BPE (Bit Plane Encoder) est décrit dans la partie 4.5 de la recommandation standard du CCSDS pour la compression d'images [29]. On s'intéresse ici au codage des coefficients d'ondelettes dits « AC », c'est-à-dire les coefficients en dehors de la sous-bande LL. Le codage de l'imagette basse fréquence « DC » n'est pas pris en compte. Les particularités du codage liées au « scaling » non plus.

Le BPE utilise une structure d'arbre pour coder les plans de bits. Trois familles sont définies par *bloc* et numérotées par l'indice $i \in \{0, 1, 2\}$. Il y a trois générations; chacune correspondant à l'un des trois niveaux de la décomposition en ondelettes :

1. Les *parents* correspondent aux coefficients du 3^e niveau de décomposition en ondelettes. Il y a 3 coefficients parents par bloc et on note le groupe $P = \{p_0, p_1, p_2\}$. Dans l'exemple $P = \{-34, -31, 23\}$.
2. Les enfants (*children*) correspondent aux coefficients du 2^e niveau de décomposition en ondelettes. Il y a 3 groupes de 4 coefficients enfants par bloc notés C_i avec $i \in \{0, 1, 2\}$.
3. Les petits-enfants (*grandchildren*) correspondent aux coefficients du 1^{er} niveau de décomposition en ondelettes. Il y a 3 ensembles de 16 coefficients petits-enfants par bloc notés G_i avec $i \in \{0, 1, 2\}$. Chaque ensemble G_i est divisé en 4 groupes de 4 coefficients d'ondelettes notés H_{ij} avec $j \in \{0, 1, 2, 4\}$.

On définit aussi la liste des descendants d'une famille i notée $D_i = \{C_i, G_i\}$ et la liste de tous les descendants dans un bloc notée $B = \{D_0, D_1, D_2\}$.

B.1 Exemple utilisé

L'exemple des coefficients d'ondelettes est tiré de l'article de Shapiro [45] dans lequel il décrit l'exécution de l'algorithme EZW. Ce même exemple a été utilisé par Said et Pearlman pour illustrer le codage SPIHT [86] et SPECK (Set Partitioned Embedded bloCK coder) [87]. Ces quatre différents codeurs peuvent donc être aisément comparés.

63	-34	49	10	7	13	-12	7
-31	23	14	-13	3	4	6	-1
15	14	3	-12	5	7	3	9
-9	-7	-14	8	4	-2	3	2
-5	9	-1	47	4	6	-2	2
3	0	-3	2	3	-2	0	4
2	-3	6	-4	3	6	3	6
5	11	5	6	0	3	-4	4

FIG. B.1 – Exemple de coefficients d'ondelettes utilisé

L'exemple de coefficients d'ondelettes est représenté sur la figure B.1. Ceci constitue un *bloc* de 64 coefficients tel que défini dans le standard du CCSDS. Le coefficient DC est représenté dans la case grisée. Les autres sont des coefficients AC. Le plus fort coefficient AC en valeur absolue vaut 49. Il y a donc 5 plans de bits à coder.

B.2 Codage avec l'algorithme du CCSDS original

Codage du premier plan de bits

Le plan de bits $b = 5$ est représenté sur la figure B.2.

Étape 1 : Codage des parents

Deux mots de code sont générés à cette étape : $types_b[P]$ et $sign_b[P]$.

- $types_b[P]$ est le mot correspondant à l'amplitude des coefficients $x \in P$ qui n'étaient pas significants aux plans de bits précédents : $t_b(x) = 0$ si $|x| < 2^b$ ou $t_b(x) = 1$ si $2^b \leq |x| < 2^{b+1}$. Comme $b = 5$ est le premier plan de bits codé, aucun coefficient n'est encore significatif. On a $P = \{-34, -31, 23\}$ et donc $types_5[P] = 1, 0, 0$.
- $signs_b[P]$ correspond aux bits de signe des coefficients $x \in P$ qui sont devenus significatifs à ce plan de bits c'est-à-dire tels que $2^b \leq |x| < 2^{b+1}$ ($t_b(x) = 1$). Ici, seul le coefficient $p_0 = -34$ est devenu significatif. Il est négatif. On a donc $signs_5[P] = 1$.

Étape 2 : Codage des enfants

Plusieurs mots de code peuvent être générés à cette étape : $tran_B$, $tran_D$, et les $types_b[C_i]$ et $signs_b[C_i]$.

- $tran_B$ est un mot de *transition* servant à indiquer qu'aucun descendant du bloc n'est significatif à ce plan de bits. Ce mot de 1 bit peut prendre la valeur 0 ou 1 ou bien ne pas être généré. S'il a déjà pris la valeur 1 dans un plan de bits précédents, il n'est pas généré. Cela signifie que des descendants ont déjà été trouvés significatifs

et qu'il faut continuer le codage des descendants dans les plans de bits suivants. $tran_B$ prend la valeur 0 si aucun descendant n'est significatif dans le bloc au plan de bits b codé. Il prend la valeur 1 si au moins un descendant du bloc est significatif : $\exists x \in B, 2^b \leq |x| < 2^{b+1} (t_b(x) = 1)$. Dans l'exemple, deux descendants sont significatifs au plan de bits $b = 5$. On a donc $tran_B = 1$.

- $tran_D$ est aussi un mot de *transition* servant à indiquer quels sont les ensembles de descendants D_i qui deviennent significatifs à ce plan de bits. Ce mot n'est généré que si $tran_B \neq 0$ (il existe des descendants significatifs). C'est un mot de longueur variable de 0 à 3 bits. Si un ensemble de descendants D_i a déjà été trouvé significatif à un plan de bits précédent, alors cet ensemble n'est pas pris en compte dans le mot binaire $tran_D$. Dans l'exemple, comme $b = 5$ est le premier plan de bits codé, aucun ensemble de descendants D_i a déjà été trouvé significatif. $tran_D$ sera donc un mot de 3 bits. De plus, comme $tran_B = 1$, le mot $tran_D$ est généré. Au plan de bits $b = 5$, deux descendants sont significatifs. Ils font partie des ensembles de descendants D_0 et D_1 . On a donc $tran_D = 1, 1, 0$.
- $types_b[C_i]$ et $signs_b[C_i]$ sont les mots correspondant respectivement à l'amplitude des enfants $x \in C_i$ qui n'étaient pas significatifs aux plans de bits précédents et aux signes des coefficients $x \in C_i$ qui sont devenus significatifs à ce plan de bits. Le mot $tran_D$ a déjà indiqué qu'il existe des descendants significatifs uniquement dans les familles $i = 0$ et $i = 1$. Seuls les groupes d'enfants C_0 et C_1 sont donc codés. On a $types_5[C_0] = 1, 0, 0, 0$, $signs_5[C_0] = 0$ ($C_0[0] = 49$) et $types_5[C_1] = 0, 0, 0, 0$. Puisqu'aucun coefficient du groupe C_1 n'est significatif, $signs_5[C_1]$ est vide.

Étape 3 : Codage des petits-enfants

Si le mot $tran_B = 0$, cette étape est vide puisque cela signifie qu'aucun descendant, petits-enfants compris, n'est significatif à ce plan de bit. Plusieurs mots de code peuvent être générés à cette étape : $tran_G$, $tran_{H_i}$, et les $types_b[H_{ij}]$ et $signs_b[H_{ij}]$.

- $tran_G$ est un mot de *transition* servant à indiquer quels sont les ensembles de petits-enfants G_i qui deviennent significatifs à ce plan de bits. C'est un mot de longueur variable de 0 à 3 bits. Si un ensemble de descendants D_i a déjà été trouvé significatif, alors les petits-enfants de la famille i sont codés. $tran_G$ indique dans quelle famille on peut trouver des petits-enfants significatifs. Au plan de bits $b = 5$, le mot $tran_D$ a indiqué que des descendants des familles $i = 0$ et $i = 1$ sont significatifs. Seuls les petits-enfants de ces deux familles sont considérés. Seul le petit-enfant $H_{11}[1] = 47$ est significatif. Il fait partie de la famille $i = 1$. On a donc $tran_G = 0, 1$.
- $tran_{H_i}$ est aussi un mot de *transition* servant à indiquer quels sont les groupes de petits-enfants H_{ij} de la famille i qui deviennent significatifs à ce plan de bits. $tran_G$ indique déjà qu'il existe des petits-enfants significatifs uniquement dans la famille $i = 1$. Seul le mot $tran_{H1}$ est donc généré. Seul le petit-enfant $H_{11}[1] = 47$ est significatif. On a donc $tran_{H1} = 0, 1, 0, 0$.
- $types_b[H_{ij}]$ et $signs_b[H_{ij}]$ sont les mots correspondant respectivement à l'amplitude des petits-enfants dans H_{ij} qui n'étaient pas significatifs aux plans de bit précédents

et aux signes des coefficients $x \in H_{ij}$ qui sont devenus significants à ce plan de bits. Le mot $tran_{H1}$ a déjà indiqué qu'il n'existe des petits-enfants significants que dans le groupe H_{11} . Seul le groupe de petits-enfants H_{11} est donc codé. On a $types_5[H_{11}] = 0, 1, 0, 0$ et $signs_5[H_{11}] = 0$ ($H_{11}[1] = 47$).

Codage du deuxième plan de bits

Le plan de bits $b = 4$ est représenté sur la figure B.3.

Étape 1 : Codage des parents

Deux mots de code sont générés à cette étape : $types_b[P]$ et $sign_b[P]$.

- $types_4[P] = 1, 1$. Comme p_0 était déjà signifiant au plan de bits précédent, il n'est pas codé ici.
- $signs_4[P] = 1, 0$ car p_1 est négatif ($p_1 = -31$) et p_2 est positif ($p_2 = 23$).

Étape 2 : Codage des enfants

- $tran_B$ n'est pas généré parce qu'il vaut déjà 1 au plan de bits précédent. On sait donc déjà qu'il y a des coefficients significants parmi les descendants.
- $tran_D$ ne prend plus qu'un bit. On sait déjà qu'il y a eu des descendants significants au plan de bits précédent dans les ensembles D_0 et D_1 (cf. $tran_D$ au plan de bits $b = 5$). Au plan de bits $b = 4$, il n'y a pas de descendant signifiant dans l'ensemble D_2 . On a donc $tran_D = 0$.
- Les mots $types_4[C_0]$ et $types_4[C_1]$ sont générés. En effet, les mots $tran_D$ aux plans de bits $b = 5$ et $b = 4$ ont indiqué qu'il n'existe des descendants significants que dans les familles $i = 0$ et $i = 1$. Comme le coefficient $C_0[0] = 49$ était déjà signifiant au plan de bits précédent $t_4(C_0[0]) = 2$, le mot $types_4[C_0]$ est un mot de trois bits. On a donc $types_4[C_0] = 0, 0, 0$ et $types_4[C_1] = 0, 0, 0, 0$. Aucun coefficient enfant est nouvellement signifiant au plan de bits $b = 4$. Les mots $signs_4[C_i]$ ne sont donc pas générés.

Étape 3 : Codage des petits-enfants

- $tran_G$ est maintenant un mot de 1 bit puisque, d'après le mot $tran_D$, il ne peut pas y avoir de petits-enfants significants dans la famille $i = 2$ et on sait déjà qu'il y a des petits-enfants significants dans la famille $i = 1$. Ici, il n'y a pas de petit-enfant signifiant dans la famille $i = 0$ au plan de bits $b = 4$. On a donc $tran_G = 0$.
- Comme le mot $tran_G$ au plan de bits $b = 5$ a indiqué au moins un coefficient signifiant parmi les petits-enfants de la famille $i = 1$, le mot $tran_{H1}$ précise leur localisation. De plus, au plan de bits $b = 5$, le mot $tran_{H1}$ a déjà indiqué des coefficients significants dans le groupe 1. Il n'est donc pas nécessaire de signaler ce groupe. Le mot $tran_{H1}$ est donc de trois bits au plan de bits $b = 3$: $tran_{H1} = 0, 0, 0$.

	0	1	.	.	.	.	.
-1	+1	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	0	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.

Plan de bits $b = 4$.

Étapes	Mots binaires	Commentaires
Parents	$types_4[P] = 1, 1$ $signs_4[P] = 1, 0$	$p_1 = -31$ et $p_2 = 23$ deviennent significants. p_0 déjà significatif au plan de bits précédent. p_1 négatif et p_2 positif.
Enfants	$tran_D = 0$ $types_4[C_0] = 0, 0, 0$ $types_4[C_1] = 0, 0, 0, 0$	Pas de descendant significatif dans la liste D_2 . $C_0[0]$ déjà significatif au plan de bits précédent. Pas de coefficient significatif dans C_1 .
Petits-enfants	$tran_G = 0$ $tran_{H_1} = 0, 0, 0$ $types_4[H_{11}] = 0, 0, 0$	Pas de petit-enfant significatif dans G_0 . Pas de petit-enfant significatif dans les groupes autre que H_{11} . $H_{11}[1]$ déjà significatif au plan de bits précédent.
Affinage	$0, 1, 0$	Bits des coefficients p_0 , $C_0[0]$ et $H_{11}[1]$ au plan de bits $b = 4$.

FIG. B.3 – Codage du plan de bits $b = 4$. Les coefficients devenant significatifs au plan de bits $b = 4$ sont marqués par un « 1 » précédé du signe du coefficient. Les coefficients non significatifs sont marqués par un « . ». Les coefficients significatifs aux plans de bits précédents sont marqués par un bit « 0 » ou « 1 » non signé.

- Seul le groupe H_{11} est codé puisque selon le mot $tran_{H1}$ c'est le seul groupe qui peut posséder des coefficients significants. Comme le coefficient $H_{11}[1] = 47$ était déjà significatif au plan de bits précédent, le mot $types_4[H_{11}]$ est un mot de trois bits codant la signifiante des autres coefficients du groupe. On a donc $types_4[H_{11}] = 0, 0, 0$.

Étape 4 : Affinage de l'amplitude des coefficients déjà significants

Dans cette partie, l'amplitude des coefficients trouvés significants aux plans de bits précédents est affinée. Les bits de ces coefficients au plan de bits $b = 4$ sont émis en commençant par les parents, puis les enfants et enfin les petits-enfants et en suivant la numérotation des familles et des groupes. Trois coefficients étaient significants avant le plan de bits $b = 4$: p_0 , $C_0[0]$ et $H_{11}[1]$. Le train binaire généré à l'étape 4 est donc 0, 1, 0.

Codage du troisième plan de bits

Le plan de bits $b = 3$ est représenté sur la figure B.4.

Étape 1 : Codage des parents

Tous les parents ont déjà été trouvés significants aux plans de bits précédents. Leur amplitude sera affinée à l'étape 4.

Étape 2 : Codage des enfants

- Il y a des descendants significants dans l'ensemble D_2 . On a donc $tran_D = 1$.
- Tous les $types_3[C_i]$, $i \in \{0, 1, 2\}$ sont générés. En effet, les mots binaires $tran_D$ aux plans de bits $b = 5$, $b = 4$ et $b = 3$ ont indiqué qu'il existe des descendants significants dans trois familles $i \in \{0, 1, 2\}$. On a donc $types_3[C_0] = 1, 1, 1$ et $signs_3[C_0] = 0, 0, 1$. $types_3[C_1] = 1, 1, 1, 0$ et $signs_3[C_1] = 0, 0, 1$. $types_3[C_2] = 0, 1, 1, 1$ et $signs_3[C_2] = 1, 1, 0$.

Étape 3 : Codage des petits-enfants

- $tran_G$ est maintenant un mot de deux bits puisqu'on sait déjà qu'il y a des petits-enfants significants dans la famille $i = 1$. Ici, il y a des petits-enfants significants dans la famille $i = 0$ et aucun dans la famille $i = 2$ au plan de bits $b = 3$. On a donc $tran_G = 1, 0$.
- $tran_{H0} = 1, 1, 0, 1$ précise la localisation des petits-enfants significants dans l'ensemble G_0 . $tran_{H1} = 1, 1, 0$ est un mot de trois bits puisqu'on savait déjà qu'il y a au moins un coefficient significatif dans le groupe H_{11} .
- Les mots $types_3[H_{ij}]$ et $signs_3[H_{ij}]$ générés sont donnés dans le tableau de la figure B.4.

	0	0 +1	· +1 -1 ·
1	0	+1 -1	· · · ·
+1 +1	· -1	· · · +1	
-1 ·	-1 +1	· · · ·	
· +1 ·	· 1	· · · ·	
· · ·	· ·	· · · ·	
· · ·	· ·	· · · ·	
· +1 ·	· ·	· · · ·	

 Plan de bits $b = 3$.

Étapes	Mots binaires	Commentaires
Parents		
Enfants	$tran_D = 1$ $types_3[C_0] = 1, 1, 1$ $signs_3[C_0] = 0, 0, 1$ $types_3[C_1] = 1, 1, 1, 0$ $signs_3[C_1] = 0, 0, 1$ $types_3[C_2] = 1, 1, 1, 0$ $signs_3[C_2] = 1, 1, 0$	Descendant significants dans la liste D_2 . $C_0[0]$ déjà significatif au plan de bits précédent. Trois nouveaux coefficients significants dans C_1 . Trois nouveaux coefficients significants dans C_2 .
Petits-enfants	$tran_G = 1, 0$ $tran_{H_0} = 1, 1, 0, 1$ $tran_{H_1} = 1, 1, 0$ $types_3[H_{00}] = 0, 1, 0, 0$ $signs_3[H_{00}] = 0$ $types_3[H_{01}] = 1, 0, 0, 0$ $signs_3[H_{01}] = 1$ $types_3[H_{03}] = 0, 1, 0, 0$ $signs_3[H_{03}] = 0$ $types_3[H_{10}] = 0, 1, 0, 0$ $signs_3[H_{10}] = 0$ $types_3[H_{11}] = 0, 0, 0$ $types_3[H_{12}] = 0, 0, 0, 1$ $signs_3[H_{12}] = 0$	Petit-enfants significants dans G_0 Pas de petit-enfant significatif dans G_2 . Petit-enfants significants dans les groupes H_{00} , H_{01} et H_{03} . Petit-enfants significants dans les groupes H_{10} , H_{12} .
Affinage	0, 1, 0, 0, 1	Coefficients p_0 , p_1 , p_2 , $C_0[0]$ et $H_{11}[1]$.

FIG. B.4 – Codage du plan de bits $b = 3$. Les coefficients devenant significants au plan de bits $b = 3$ sont marqués par un « 1 » précédé du signe du coefficient. Les coefficients non significants sont marqués par un « · ». Les coefficients significants aux plans de bits précédents sont marqués par un bit « 0 » ou « 1 » non signé.

Étape 4 : Affinage de l'amplitude des coefficients déjà significants

Cinq coefficients étaient significants avant le plan de bits $b = 3$: p_0 , p_1 , p_2 , $C_0[0]$ et $H_{11}[1]$. Le train binaire généré à l'étape 4 est donc 0, 1, 0, 0, 1.

Codage entropique

Les mots binaires générés précédemment sont codés entropiquement à l'aide de tables de codage. Dans un premier temps, ils sont convertis en symboles en utilisant des tables différentes en fonction de la longueur du mot binaire et du type de mot. D'autres tables de codage sont ensuite utilisées pour convertir ces symboles en mots binaires de longueur variable qui sont inclus dans le train binaire. Les mots binaire de 1 bit ne sont pas codés mais inclus directement dans le train binaire.

Redondance identifiée

Une redondance peut être identifiée dans cet algorithme. Lors du codage du plan de bits $b = 5$, le mot $tran_D$ indique qu'il existe des descendants significants dans la famille $i = 1$. Lors du codage de la signifiante des enfants, on s'aperçoit qu'aucun enfant de l'ensemble C_1 n'est significatif. On en déduit qu'il existe forcément au moins un petit-enfant significatif dans la famille $i = 1$. Pourtant, cette information est retransmise dans le mot $tran_G$. Heureusement, la probabilité de cette configuration des coefficients d'ondelettes est faible. Peu de bits sont donc redondants. De plus, l'exploitation de cette redondance complique l'implantation de l'algorithme.

B.3 Codage avec l'algorithme modifié

Le codage décrit dans les tableaux des figures B.5, B.6 et B.7 est effectué avec l'algorithme modifié de la section 5.2.5. La signalisation de la post-transformée (mot PT_i) n'est pas prise en compte dans cet exemple.

Sans le codage entropique des mots de code, l'algorithme du CCSDS produit 28 bits pour le codage du plan de bits $b = 5$, 22 bits pour le codage du plan de bits $b = 4$, et 65 bits pour le codage du plan de bits $b = 3$. La version modifiée de l'algorithme produit respectivement 25, 23 et 73 bits. Elle semble plus adaptée pour la compression à bas débit. Pour la comparaison avec les codeurs EZW [45], SPIHT [86] et [87], il faut prendre en compte le codage de la sous-bande LL.

	-1	+1	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.

Plan de bits $b = 5$.

Étapes	Mots binaires	Commentaires
Parents	$types_5[P] = 1, 0, 0$ $signs_5[P] = 1$	$p_0 = -34$ devient significatif au plan de bits $b = 5$. p_0 négatif.
Enfants	$tran_B = 1$ $tran_D = 1, 1, 0$ $tran_{K0} = 1, 0, 0, 0$ $tran_{K1} = 0, 1, 0, 0$ $types_5[c_{00}] = 1$ $signs_5[c_{00}] = 0$ $types_5[c_{11}] = 0$	Première fois qu'un descendant est significatif. Descendants significatifs dans D_0 et D_1 . Pas de codage des coefficients de D_2 . Enfants significatifs dans l'ensemble K_{00} Petits-enfants significatifs dans l'ensemble K_{11} $c_{00} = 49$ devient significatif au plan de bits $b = 5$. c_{00} est positif. c_{11} n'est pas significatif.
Petits-enfants	$tran_{H0} = 0$ $tran_{H1} = \emptyset$ $types_5[H_{11}] = 0, 1, 0, 0$ $signs_5[H_{11}] = 0$	Pas de petit-enfant significatif dans le groupe H_{00} . c_{11} n'est pas significatif. $H_{11}[1] = 47$ devient significatif au plan de bits $b = 5$. $H_{11}[1]$ est positif.

FIG. B.5 – Codage du plan de bits $b = 5$.

	0	1	.	.	.	.
-1	+1	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	0	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.

 Plan de bits $b = 4$.

Étapes	Mots binaires	Commentaires
Parents	$types_4[P] = 1, 1$ $signs_4[P] = 1, 0$	$p_1 = -31$ et $p_2 = 23$ deviennent significants. p_0 déjà significatif au plan de bits précédent. p_1 négatif et p_2 positif.
Enfants	$tran_D = 0$ $tran_{K_0} = 0, 0, 0$ $tran_{K_1} = 0, 0, 0$ $types_4[c_{00}] = \emptyset$ $types_4[c_{11}] = 0$	Pas de descendant significatif dans la liste D_2 . K_{00} déjà significatif au plan de bits précédent. K_{11} déjà significatif au plan de bits précédent. c_{00} sera affiné à l'étape 4. c_{11} n'est pas significatif.
Petits-enfants	$tran_{H_0} = 0$ $tran_{H_1} = \emptyset$ $types_4[H_{11}] = 0, 0, 0$	Pas de petit-enfant significatif dans H_{00} . c_{11} n'est pas significatif. $H_{11}[1]$ déjà significatif au plan de bits précédent.
Affinage	$0, 1, 0$	Bits des coefficients p_0 , $C_0[0]$ et $H_{11}[1]$ au plan de bits $b = 4$.

 FIG. B.6 – Codage du plan de bits $b = 4$.

	0	0 +1	· +1 -1 ·
1	0	+1 -1	· · · ·
+1 +1	· -1	· · · +1	
-1 ·	-1 +1	· · · ·	
· +1 ·	· 1	· · · ·	
· · · ·	· · · ·	· · · ·	
· · · ·	· · · ·	· · · ·	
· +1 ·	· ·	· · · ·	

 Plan de bits $b = 3$.

Étapes	Mots binaires	Commentaires
Parents		
Enfants	$tran_D = 1$ $tran_{K0} = 1, 1, 1$ $tran_{K1} = 1, 1, 0$ $tran_{K2} = 0, 1, 1, 1$ $types_3[\{c_{01}, c_{02}, c_{03}\}] = 1, 1, 1$ $signs_3[\{c_{01}, c_{02}, c_{03}\}] = 0, 0, 1$ $types_3[\{c_{10}, c_{11}, c_{12}\}] = 1, 1, 1$ $signs_3[\{c_{10}, c_{11}, c_{12}\}] = 0, 0, 1$ $types_3[\{c_{21}, c_{22}, c_{23}\}] = 1, 1, 1$ $signs_3[\{c_{21}, c_{22}, c_{23}\}] = 1, 1, 0$	Descendant significants dans la liste D_2 . K_{00} déjà significatif au plan de bits précédent. K_{11} déjà significatif au plan de bits précédent. c_{00} déjà significatif au plan de bits précédent. K_{13} non significatif K_{20} non significatif
Petits-enfants	$tran_{H0} = 1, 1, 0, 1$ $tran_{H1} = 1, 1$ $tran_{H2} = 0, 0, 0$ $types_3[H_{00}] = 0, 1, 0, 0$ $signs_3[H_{00}] = 0$ $types_3[H_{01}] = 1, 0, 0, 0$ $signs_3[H_{01}] = 1$ $types_3[H_{03}] = 0, 1, 0, 0$ $signs_3[H_{03}] = 0$ $types_3[H_{10}] = 0, 1, 0, 0$ $signs_3[H_{00}] = 0$ $types_3[H_{11}] = 0, 0, 0$ $types_3[H_{12}] = 0, 0, 0, 1$ $signs_3[H_{12}] = 0$	Petit-enfants significants dans les groupes H_{00} , H_{01} et H_{03} . H_{11} déjà significatif au plan de bits précédent et K_{13} non significatif K_{20} non significatif
Affinage	0, 1, 0, 0, 1	Coefficients p_0 , p_1 , p_2 , $C_0[0]$ et $H_{11}[1]$.

 FIG. B.7 – Codage du plan de bits $b = 3$.

Annexe C

Exemple de codage avec EBCOT

Le codeur par plans de bits de JPEG2000 est EBCOT (Embedded Block Coding with Optimal Truncation Points). C'est un codeur contextuel. Sur chaque plan de bits, trois passes de codage sont utilisées : une passe de propagation de la signifiante (*Significance Propagation*), une passe d'affinage de l'amplitude (*Magnitude Refinement*), et une passe de nettoyage (*Cleanup*). Il existe quatre primitives de codage utilisées par ces passes : la primitive RL (Run-Length), la primitive ZC (Zero Coding), la primitive MR (Magnitude Refinement), et la primitive SC (Sign Coding) [21, 27, 28, 49]. Dans ce document, nous allons montrer sur un exemple comment ces passes de codage sont utilisées conjointement aux primitives de codage.

C.1 Exemple utilisé

Le codeur EBCOT est un codeur par blocs, c'est-à-dire qu'il code les coefficients d'ondelettes bloc par bloc. En général, les blocs sont de taille 64×64 . La portion de code-bloc que nous allons utiliser dans notre exemple est présentée sur la figure C.1. On fait l'hypothèse que ce bloc est dans une sous-bande LH. Les nombres représentent les valeurs des coefficients d'ondelettes après la quantification. Le parcours des coefficients est effectué par colonnes de 4 coefficients, et bande par bande, comme illustré sur la figure C.1. Sur cette figure, la première ligne de coefficients d'ondelettes appartient à la bande précédente, la dernière à la bande suivante.

Le coefficient le plus fort en valeur absolue dans ce code-bloc vaut 61 (en position (4,7)). Il y a donc 6 plans de bits dans ce code-bloc ($2^6 = 64$). On les numérote de 0 à 5, du plan de bits le moins significatif au plan de bits le plus significatif. Ce nombre de plans de bits est inclus dans le train binaire émis afin de signaler au décodeur à partir de quel plan de bits le codage est effectué.

C.2 Codage du premier plan de bits

Dans le premier plan de bits ($b = 5$) représenté sur la figure C.2, seule la passe de nettoyage est utilisée. En effet, la passe de propagation de la signifiante ne peut être

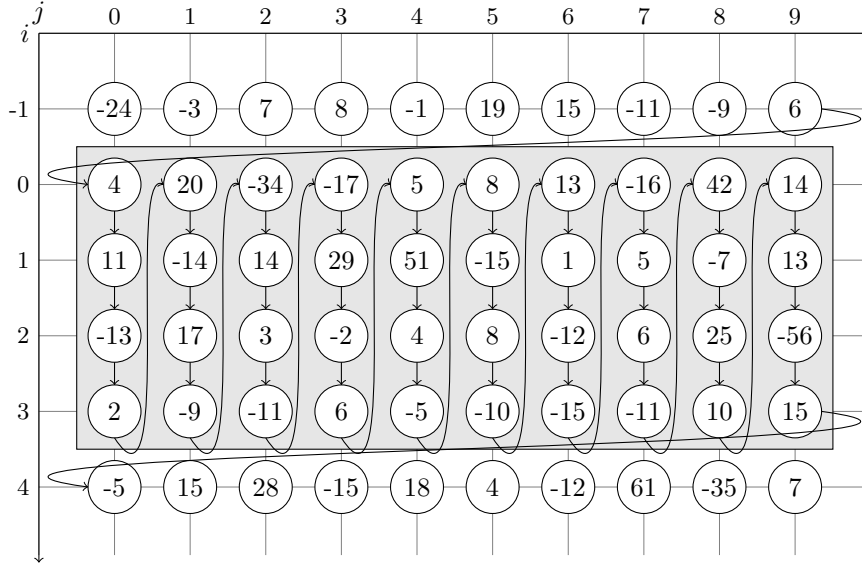


FIG. C.1 – Portion de code-bloc utilisée dans notre exemple.

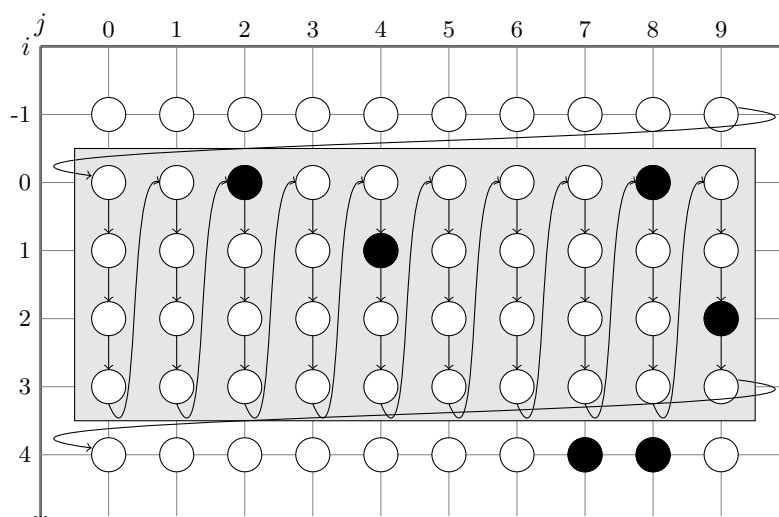
utilisée puisqu'aucun coefficient d'ondelette n'est significatif avant ce plan de bits. Comme il n'y a pas de coefficient significatif connu, on ne peut pas propager le codage de la significiance autour des coefficients significatifs. Pour la même raison, la passe d'affinage de l'amplitude ne peut-être utilisée. En effet, il faut d'abord trouver les coefficients significatifs pour pouvoir ensuite affiner leur amplitude dans les plans de bits moins significatifs. La seule passe utilisée pour coder le premier plan de bits est donc la passe de nettoyage. Cette passe est nommée ainsi parce qu'elle sert à coder tous les bits qui n'ont pas été codés par la passe de propagation de la significiance ou la passe d'affinage de l'amplitude.

Passe de nettoyage

On commence par le codage de la colonne $j = 0$. Avant le codage de cette colonne, on ne connaît pas de coefficient significatif dans cette colonne. De plus, on ne connaît pas de coefficient voisin à cette colonne qui soit significatif. Ces deux conditions réunies, le codeur passe en mode *Run-Length*. Le codeur parcourt alors la colonne $j = 0$. Au plan de bits $b = 5$, aucun coefficient n'est significatif dans cette colonne. Le codeur émet donc un symbole 0 dans le contexte (RL) pour signaler une colonne de zéros.

De même, sur la colonne $j = 1$, avant le codage, il n'y a pas de coefficient significatif connu, ni de voisin significatif *connu*. On reste donc dans le mode *Run-Length*. Comme aucun coefficient de cette colonne n'est significatif au plan de bits $b = 5$, le codeur émet un symbole 0 dans le contexte (RL).

Pour le codage de la colonne $j = 2$, le codeur est toujours dans le mode *Run-Length* car les deux conditions sont réunies. Cependant, le coefficient (0,2) est significatif au plan de bits $b = 5$. Le codeur émet alors le symbole 1 dans le contexte (RL) pour signaler la sortie du mode Run-Length. De plus, le codeur code la position du premier coefficient significatif de la colonne. Il émet les symboles 00 dans le contexte (UNI). Cela permet de signaler au décodeur la position exacte de sortie du mode Run-Length. Dans ce contexte UNIFORM

Plan de bits $b = 5$.

	Symboles codés et contextes	Commentaires
$j = 0$	0(RL)	Run-Length.
$j = 1$	0(RL)	Run-Length.
$j = 2$	1(RL) 00(UNI) 1(SC0) 0(ZC3) 0(ZC0) 0(ZC0)	Run-Length jusqu'au coefficient (0,2). Position du premier coefficient significatif. Le signe prédit par le contexte (SC0) est incorrect. Reste de la colonne codé par la primitive ZC.
$j = 3$	0(ZC5) 0(ZC1) 0(ZC0) 0(ZC0)	Pas de Run-Length parce que le coefficient (0,2) est significatif.
$j = 4$	1(RL) 01(UNI) 0(SC0) 0(ZC3) 0(ZC0)	Run-Length jusqu'au coefficient (1,4). Position du premier coefficient significatif. Le signe prédit par le contexte (SC0) est correct. Reste de la colonne codé par la primitive ZC.
$j = 5$	0(ZC1) 0(ZC5) 0(ZC1) 0(ZC0)	Pas de Run-Length parce que le coefficient (1,4) est significatif.
$j = 6$	0(RL)	Run-Length.
$j = 7$	0(RL)	Run-Length.
$j = 8$	1(RL) 00(UNI) 0(SC0) 0(ZC3) 0(ZC0) 0(ZC0)	Run-Length jusqu'au coefficient (0,8). Position du premier coefficient significatif. Le signe prédit par le contexte (SC0) est correct. Reste de la colonne codé par la primitive ZC.
$j = 9$	0(ZC5) 0(ZC1) 1(ZC0) 1(SC0) 0(ZC3)	Pas de Run-Length parce que le coefficient (0,8) est significatif. Un coefficient significatif dans le contexte (ZC0). Le signe prédit par le contexte (SC0) est incorrect.

FIG. C.2 – Codage du plan de bits $b = 5$ par la passe de nettoyage.

$\bar{\chi}^h$	$\bar{\chi}^v$	κ^{SC}	$\hat{\chi}$
1	1	SC4	1
1	0	SC3	1
1	-1	SC2	1
0	1	SC1	1
0	0	SC0	1
0	-1	SC1	-1
-1	1	SC2	-1
-1	0	SC3	-1
-1	-1	SC4	-1

TAB. C.1 – Prédiction du signe $\hat{\chi}$ et contextes κ^{SC} de la primitive SC.

les symboles sont considérés équiprobables. Le signe de chaque coefficient d'ondelette est codé dès que le coefficient a été trouvé significatif. Il faut donc maintenant coder le signe du coefficient (0,2).

Pour le codage du signe, le codeur EBCOT utilise une prédiction basée sur les signes des quatre coefficients voisins. Les prédictions sont données dans le tableau C.1 pour les sous-bandes LL, LH ou HH. Les valeurs $\bar{\chi}^h$ et $\bar{\chi}^v$ représentent le signe des deux voisins respectivement horizontaux et verticaux. Ces quantités sont inversées dans le tableau de prédiction pour le codage des sous-bandes HL. Une valeur $\bar{\chi}^x$ vaut 1 si les deux voisins sont positifs ou si un voisin est positif et l'autre n'est pas encore signé, 0 si les deux voisins ne sont pas encore signés ou ont des signes opposés, et -1 si les deux voisins sont négatifs ou si un voisin est négatif et l'autre n'est pas encore signé. La valeur $\hat{\chi}$ représente la prédiction du signe du coefficient codé. Si cette prédiction est exacte, le symbole 0 est émis dans le contexte SC respectif, sinon le symbole 1 est émis dans ce même contexte. Dans le cas du coefficient (0,2), comme il n'y a aucun voisin signé, on est dans le contexte (SC0). Le signe prédit est donc positif. Cette prédiction étant incorrecte, le symbole 1 est émis dans le contexte (SC0).

Dans la colonne $j = 2$, il reste à coder les symboles (1,2) à (3,2). Comme on est sorti du mode Run-Length, c'est la primitive ZC qui est utilisée. Neuf contextes, notés ZC0 à ZC8, sont utilisés dans cette primitive. Ces contextes prennent en compte la signifiante connue des huit coefficients voisins du bit à coder. Ils sont définis dans le tableau C.2. Les valeurs κ^h , κ^v et κ^d représentent respectivement le nombre de coefficients voisins horizontaux, verticaux, et diagonaux qui sont déjà significatifs au moment du codage. Le coefficient (1,2) n'a qu'un coefficient voisin significatif : le coefficient (0,2). De plus, le coefficient (1,2) n'est pas significatif au plan de bits $b = 5$. Le symbole 0 est donc émis dans le contexte (ZC3). Les coefficients (2,2) et (3,2) n'ont pas de voisins significatifs et ne sont pas significatifs au plan de bits $b = 5$. Deux symboles 0 sont donc codés dans le contexte (ZC0).

Pour le codage de la colonne $j = 3$, comme le coefficient (0,2), voisin à cette colonne, est connu significatif au plan de bits $b = 5$, le mode Run-Length n'est pas activé. C'est

Sous-bandes LL et LH			Sous-bandes HL			Sous-bandes HH		κ^{ZC}
κ^h	κ^v	κ^d	κ^h	κ^v	κ^d	κ^d	$\kappa^h + \kappa^v$	
0	0	0	0	0	0	0	0	ZC0
0	0	1	0	0	1	0	1	ZC1
0	0	≥ 2	0	0	≥ 2	0	≥ 2	ZC2
0	1	x	1	0	x	1	0	ZC3
0	2	x	2	0	x	1	1	ZC4
1	0	0	0	1	0	1	≥ 2	ZC5
1	0	≥ 1	0	1	≥ 1	2	0	ZC6
1	≥ 1	x	≥ 1	1	x	2	≥ 1	ZC7
2	x	x	x	2	x	≥ 3	x	ZC8

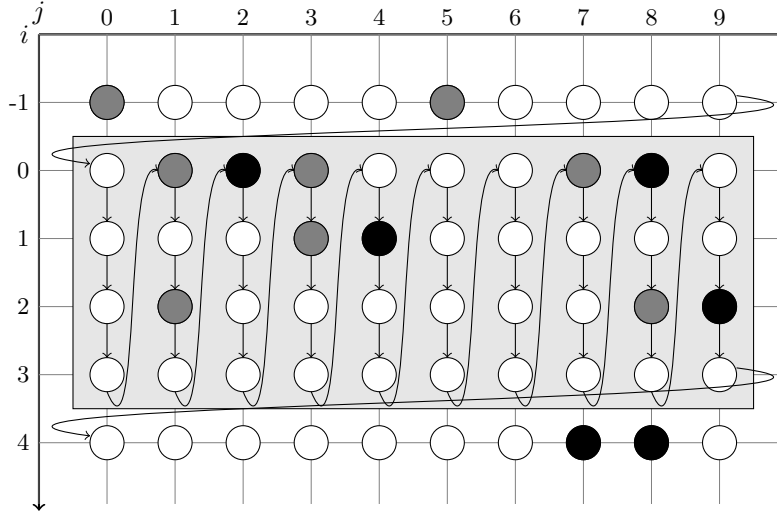
TAB. C.2 – Contextes κ^{ZC} de la primitive ZC. « x » signifie n'importe quelle valeur.

la primitive ZC qui est utilisée pour coder les quatre bits de cette colonne. Les symboles codés sont 0(ZC5), 0(ZC1), 0(ZC0), et 0(ZC0).

Il n'y a pas de coefficient voisin de la colonne $j = 4$ connu signifiant. De plus, aucun des coefficients de cette colonne n'est connu signifiant. Le codeur entre donc dans le mode *Run-Length*. Comme le coefficient (1,4) est signifiant au plan de bits $b = 5$, le codeur émet le symbole 1 dans le contexte (RL) et signale la position du premier coefficient signifiant de la colonne en codant sa position par les symboles 01 dans le contexte (UNI). Le signe de ce coefficient est ensuite codé. Le coefficient (1,4) est positif. Le contexte est (SC0). La prédiction d'un signe positif est correcte. Le symbole 0 est donc codé dans le contexte (SC0). Les deux bits suivants sont codés par la primitive ZC. Les symboles codés sont 0(ZC3) et 0(ZC0). Le reste du codage du plan de bits $b = 5$ par la passe de nettoyage est décrit dans la figure C.2.

C.3 Codage du second plan de bits

Le second plan de bits ($b = 4$) est représenté sur la figure C.3. Les coefficients signifiants au plan de bits précédent sont représentés en noir. Les coefficients trouvés signifiants au plan de bits $b = 4$ sont représentés en gris. Sur ce second plan de bits, les trois passes de codage sont utilisées : la passe de propagation de la signifiante est d'abord utilisée pour coder les bits voisins des coefficients déjà identifiés signifiants au plan de bits $b = 4$ ou aux plans de bits plus significatifs. En effet, les forts coefficients d'ondelettes sont souvent agglomérés. C'est pourquoi on commence par chercher des coefficients signifiants autour des coefficients les plus forts déjà connus. Ensuite, la passe d'affinage de l'amplitude est utilisée pour affiner l'amplitude des coefficients qui étaient déjà signifiants aux plans de bits précédemment codés. Enfin, la passe de nettoyage est utilisée pour coder les bits restants.

FIG. C.3 – Plan de bits $b = 4$.

Passe de propagation de la signifiante

Lors de la lecture du plan de bits $b = 4$, on trouve que le coefficient $(0,0)$ a un voisin déjà signifiant : le coefficient $(-1,0)$ qui est devenu signifiant au plan de bits $b = 4$. Le bit à la position $(0,0)$ est donc codé dans la passe de propagation de la signifiante. Dans la passe de propagation de la signifiante, seule la primitive ZC est utilisée. Le symbole émis pour coder le bit $(0,0)$ est donc 0 dans le contexte (ZC3).

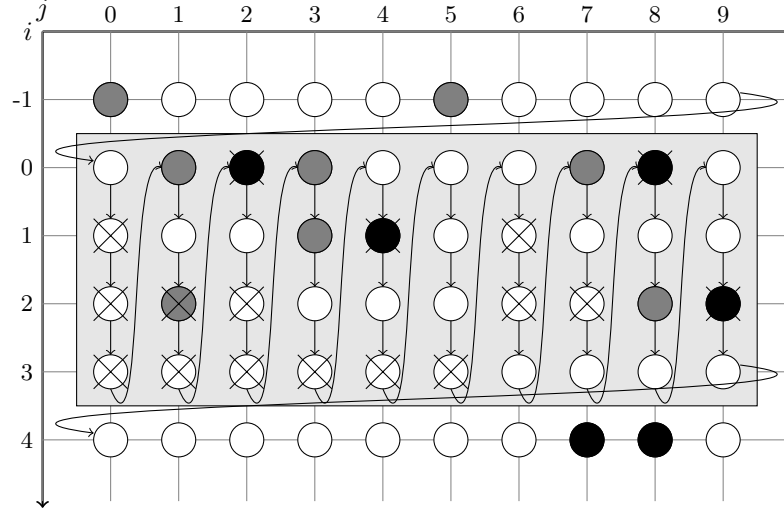
Les coefficients $(1,0)$ à $(3,0)$ n'ont pas de voisins signifiants *connus*. Ils ne sont donc pas codés par la passe de propagation de la signifiante. Ces coefficients sont marqués d'une croix sur la figure C.4.

Comme le coefficient $(0,2)$ est signifiant, les bits aux positions $(0,1)$ et $(1,1)$ sont codés. Le bit $(0,1)$ est signifiant au plan de bit $b = 4$. Le symbole 1(ZC6) est émis. Le signe du coefficient d'ondelette est ensuite codé par le symbole 1(SC3) car le signe prédit est négatif alors que le coefficient codé est positif. Le symbole 0(ZC3) est ensuite émis pour coder le coefficient $(1,1)$.

Les coefficients $(2,1)$ et $(3,1)$ ne sont pas codés. Le coefficient $(0,2)$ n'est pas codé non plus car il était déjà signifiant au plan de bits précédent. Le codeur émet le symbole 0(ZC3) pour coder le bit $(1,2)$. Les coefficients $(2,2)$ et $(3,2)$ ne sont pas codés. Le reste du codage du plan de bits $b = 4$ par la passe de propagation de la signifiante est décrit dans la figure C.4.

Passe d'affinage de l'amplitude

Dans la passe d'affinage de l'amplitude, seule la primitive de codage MR est utilisée. Cette primitive code les bits des coefficients déjà signifiants aux plans de bits précédents en utilisant les trois contextes du tableau C.3 notés MR0 à MR2. La variable $\tilde{\sigma}$ vaut 0 si c'est la première fois que la passe d'affinage de l'amplitude est appliquée au coefficient



Passé de propagation de la signifiante sur le plan de bits $b = 4$.

	Symboles codés et contextes	Commentaires
$j = 0$	0(ZC3)	Codage du coefficient (0,0).
$j = 1$	1(ZC6) 1(SC3) 0(ZC3)	Codage du coefficient (0,1). Le signe prédit dans le contexte (SC3) est incorrect. Codage du coefficient (1,1).
$j = 2$	0(ZC3)	Codage du coefficient (1,2).
$j = 3$	1(ZC6) 0(SC3) 1(ZC7) 0(SC2) 0(ZC3)	Codage du coefficient (0,3). Codage du coefficient (1,3). Codage du coefficient (2,3).
$j = 4$	0(ZC7) 0(ZC3)	Codage des coefficients (0,4) et (2,4).
$j = 5$	0(ZC3) 0(ZC5) 0(ZC1)	Codage des coefficients (0,5) à (2,5).
$j = 6$	0(ZC1) 0(ZC1)	Codage des coefficients (0,6) et (3,6).
$j = 7$	1(ZC5) 1(SC3) 0(ZC3) 0(ZC3)	Codage du coefficient (0,7). Codage des coefficients (1,7) et (3,7).
$j = 8$	0(ZC3) 1(ZC5) 1(SC3) 0(ZC4)	Codage du coefficient (1,8). Codage du coefficient (2,8). Codage du coefficient (3,8).
$j = 9$	0(ZC5) 0(ZC3) 0(ZC3)	Codage des coefficients (0,9), (2,9), et (3,9).

FIG. C.4 – Codage du plan de bits $b = 4$ par la passe de propagation de la signifiante.

$\tilde{\sigma}$	$\kappa^h + \kappa^v$	κ^{MR}
0	0	MR0
0	$\neq 0$	MR1
1	x	MR2

TAB. C.3 – Contextes κ^{MR} de la primitive MR.

codé, sinon la variable $\tilde{\sigma}$ vaut 1. Les variables κ^h et κ^v sont les mêmes que pour la primitive de codage ZC.

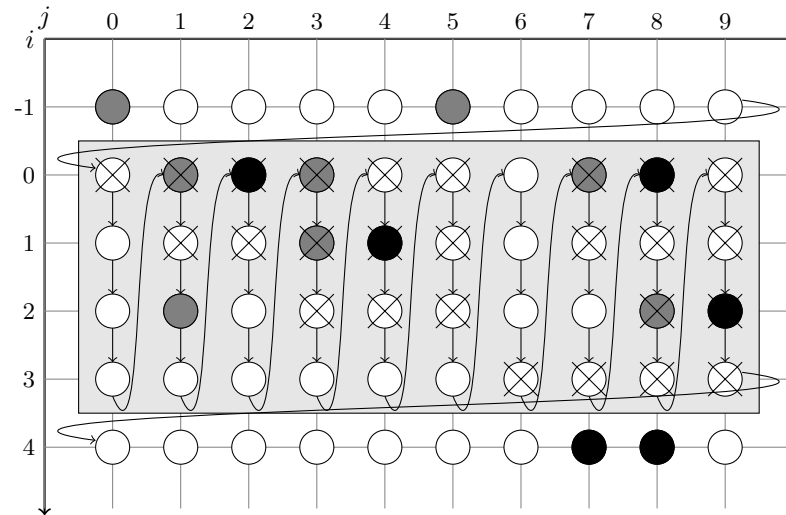
Les coefficients qui étaient significants dans les plans de bits précédemment codés ont leur amplitude affinée dans cette passe. Le symbole 0(MR1) est émis pour affiner l'amplitude du coefficient (0,2). De même, les symboles 1(MR1), 0(MR1), et 1(MR1) sont émis pour affiner l'amplitude des coefficients respectivement en positions (1,4), (0,8), et (2,9).

Passe de nettoyage

La passe de nettoyage est utilisée pour coder les bits qui n'ont pas été codés par les deux passes précédentes. Ce sont les bits qui ne sont pas marqués d'une croix sur la figure C.5. Le codage est décrit dans la figure C.5.

C.4 Suite du codage

Dans la suite, le déroulement du codage est similaire jusqu'au plan de bits $b = 0$. Le codeur arithmétique MQ code chaque bit émis en utilisant des probabilités d'apparitions différentes pour chaque contexte. Ces probabilités conditionnées par les contextes sont ajustées à chaque bit codé. Ce processus de codage des codes-blocs constitue la première partie du codeur EBCOT. Elle est nommée Tier1. Dans Tier2, la seconde partie du codeur, les morceaux de trains binaires, issus du codage des plans de bits des codes-blocs dans Tier1, sont organisés de façon à optimiser le train binaire final en termes de débit-distorsion. Ce processus est nommé PCRD-opt (*Post-Compression Rate-Distorsion OPTimization*).



Passé de nettoyage sur le plan de bits $b = 4$.

	Symboles codés et contextes	Commentaires
$j = 0$	0(ZC1) 0(ZC0) 0(ZC0)	
$j = 1$	1(ZC0) 0(SC0) 0(ZC3)	Codage du coefficient (2,1).
$j = 2$	0(ZC6) 0(ZC1)	
$j = 3$	0(ZC0)	
$j = 4$	0(ZC0)	
$j = 5$	0(ZC0)	
$j = 6$	0(ZC6) 0(ZC1) 0(ZC0)	
$j = 7$	0(ZC5)	

FIG. C.5 – Codage du plan de bits $b = 4$ par la passe de nettoyage.

Annexe D

Preuves diverses

D.1 Théorème d'optimisation de l'allocation débit-distorsion

La preuve de ce théorème peut être trouvée dans [72]. Elle est retranscrite ici.

Théorème. *Soit le problème d'allocation contraint suivant :*

$$\min_{S \in \mathcal{S}} D(S) \tag{D.1}$$

$$R(S) \leq R_T. \tag{D.2}$$

Pour tout $\lambda \geq 0$, la solution $S^(\lambda)$ au problème non contraint*

$$\min_{S \in \mathcal{S}} D(S) + \lambda R(S) \tag{D.3}$$

est aussi la solution au problème contraint (D.1-D.2) avec la contrainte $R(S) \leq R(S^(\lambda))$.*

Démonstration. Pour la solution S^* au problème non contraint (D.3) on a

$$\forall S \in \mathcal{S}, \quad D(S^*) + \lambda R(S^*) \leq D(S) + \lambda R(S)$$

ce qui est équivalent à

$$\forall S \in \mathcal{S}, \quad D(S^*) - D(S) \leq \lambda (R(S) - R(S^*)).$$

Comme cette inégalité est vérifiée pour tout $S \in \mathcal{S}$, elle est aussi vérifiée pour tout S de n'importe quel sous-ensemble de $\mathcal{S}$. En particulier, elle est vérifiée pour tout S du sous-ensemble $\mathcal{S}^*$ qui permet de respecter la contrainte $R(S) \leq R(S^*(\lambda))$. Ce sous-ensemble est défini par

$$\mathcal{S}^* = \{S \mid R(S) \leq R(S^*(\lambda))\}.$$

Comme $\lambda \geq 0$, on a

$$\forall S \in \mathcal{S}^*, \quad D(S^*) - D(S) \leq 0.$$

$D(S^*)$ est donc minimum sur le sous-ensemble $\mathcal{S}^*$ qui permet de respecter la contrainte $R(S) \leq R(S^*(\lambda))$. L'objectif est donc atteint. $\square$

D.2 Calcul de la distorsion sous hypothèse de quantification haute résolution

Sous hypothèse d'une quantification scalaire uniforme à haute résolution, l'erreur quadratique moyenne D s'exprime en fonction du pas de quantification q [74] :

$$D \approx \frac{q^2}{12}. \quad (\text{D.4})$$

Démonstration. L'erreur quadratique moyenne d_k sur l'intervalle de quantification $I_k = [kq, (k+1)q)$ de largeur q vaut :

$$d_k = \int_{I_k} (x - Q_q(x))^2 p(x) dx$$

où $p(x)$ est la densité de probabilité des coefficients x . L'erreur quadratique moyenne due à la quantification scalaire uniforme infinie vaut donc :

$$D = \sum_{k=-\infty}^{+\infty} d_k = \sum_{k=-\infty}^{+\infty} \int_{I_k} (x - Q_q(x))^2 p(x) dx.$$

Sous hypothèse haute résolution, la densité de probabilité $p(x)$ est presque constante sur chaque intervalle de quantification. On a donc $p(x) \approx p_k$ pour tout $x \in I_k$ avec p_k la densité de probabilité moyenne sur l'intervalle I_k . L'erreur quadratique moyenne est donc approchée par :

$$D \approx \sum_{k=-\infty}^{+\infty} p_k \int_{I_k} (x - Q_q(x))^2 dx.$$

Or un coefficient quantifié $x \in I_k$ est reconstruit au milieu de l'intervalle de quantification $I_k = [kq, (k+1)q)$: $Q_q(x) = (k+1/2)q$. La densité de probabilité étant constante sur chaque intervalle de quantification, cette valeur de reconstruction permet de minimiser la distorsion. On a donc :

$$D \approx \sum_{k=-\infty}^{+\infty} p_k \int_{kq}^{(k+1)q} (x - (k+1/2)q)^2 dx.$$

Par changement de variable $t = x - (k+1/2)q$ cette intégrale s'écrit :

$$\begin{aligned} D &\approx \sum_{k=-\infty}^{+\infty} p_k \int_{-q/2}^{q/2} t^2 dt \\ &\approx \sum_{k=-\infty}^{+\infty} p_k \frac{q^3}{12}. \end{aligned}$$

Or $q p_k$ est la probabilité d'apparition d'un coefficient x dans l'intervalle I_k . Pour un quantification scalaire uniforme infinie on a donc $\sum_{k=-\infty}^{+\infty} q p_k = 1$ d'où l'expression de la distorsion D en fonction du pas de quantification q sous hypothèse haute résolution et pour un quantification scalaire uniforme infinie :

$$D \approx \frac{q^2}{12}.$$

□

D.3 Calcul du débit sous hypothèse de quantification haute résolution

Sous hypothèse d'une quantification scalaire uniforme à haute résolution, le débit moyen R s'exprime en fonction du pas de quantification q et de l'entropie différentielle du signal à quantifier $h(X)$ [75] :

$$R \approx h(X) - \log_2 q. \quad (\text{D.5})$$

Démonstration. Le débit entropique d'une source sans mémoire et produisant des tirages indépendants et identiquement distribués est :

$$R = - \sum_{k=-\infty}^{+\infty} \text{Pr}(k) \log_2 \text{Pr}(k)$$

où $\text{Pr}(k)$ est la probabilité d'apparition d'un coefficient x dans l'intervalle I_k . Sous hypothèse haute résolution, cette probabilité peut être approchée par $\text{Pr}(k) \approx q p_k$ où p_k est la densité de probabilité moyenne sur l'intervalle I_k . Pour une quantification scalaire uniforme infinie, le débit entropique est donc approché par :

$$\begin{aligned} R &\approx - \sum_{k=-\infty}^{+\infty} q p_k \log_2(q p_k) \\ &\approx - \sum_{k=-\infty}^{+\infty} q p_k \log_2(q) - \sum_{k=-\infty}^{+\infty} q p_k \log_2(p_k). \end{aligned}$$

Or on a

$$\sum_{k=-\infty}^{+\infty} q p_k = 1$$

et

$$\sum_{k=-\infty}^{+\infty} q p_k \log_2(p_k) \approx \int_{-\infty}^{+\infty} p(x) \log_2(p(x)) dx.$$

L'entropie différentielle du signal à quantifier est définie par :

$$h(X) \triangleq \int_{-\infty}^{+\infty} p(x) \log_2(p(x)) dx$$

d'où l'expression du débit entropique R en fonction du pas de quantification q et de l'entropie différentielle $h(X)$:

$$R \approx h(X) - \log_2(q).$$

□

D.4 Cohérence d'un dictionnaire de vecteurs

Propriété. Dès lors que le dictionnaire $\mathbf{D}$ contient une base orthonormale et un vecteur normé additionnel, alors la cohérence $M(\mathbf{D})$ respecte l'inégalité suivante :

$$M(\mathbf{D}) \geq \frac{1}{\sqrt{N}}.$$

Démonstration. On considère que les N premiers vecteurs du dictionnaire $\mathbf{D} = [\mathbf{g}_1, \dots, \mathbf{g}_N, \mathbf{g}_{N+1}]$ forment une base orthonormale de $\mathbb{R}^N$. Le vecteur $\mathbf{g}_{N+1}$ est normé et donc $\|\mathbf{g}_{N+1}\|^2 = 1$. Comme la famille $\{\mathbf{g}_k\}_{k=1}^N$ est une base orthonormée $\mathbb{R}^N$, le vecteur $\mathbf{g}_{N+1}$ se décompose dans cette base et sa norme s'exprime par :

$$\|\mathbf{g}_{N+1}\|^2 = \sum_{k=1}^N |\langle \mathbf{g}_{N+1}, \mathbf{g}_k \rangle|^2 = 1.$$

On a donc forcément :

$$\max_{k \in [1, N]} |\langle \mathbf{g}_{N+1}, \mathbf{g}_k \rangle|^2 \geq 1/N,$$

d'où :

$$M(\mathbf{D}) \geq \frac{1}{\sqrt{N}}.$$

□

D.5 Dissimilarité entre bases

Propriété (Séparation). Soient deux bases orthonormées $A = \{\mathbf{a}_i\}_{i=1}^N$ et $B = \{\mathbf{b}_j\}_{j=1}^N$. La mesure de dissimilarité entre ces deux bases $d(A, B)$ définie dans l'équation (4.2) est nulle si et seulement si ces deux bases sont identiques, c'est-à-dire si et seulement si elles ont les mêmes vecteurs au sens près.

Démonstration. Par définition de la mesure de dissimilarité entre bases (4.2),

$$d(A, B) = 0 \Leftrightarrow \sum_{i=1}^N \sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle| = N.$$

Or pour tout vecteur normé $\mathbf{a}_i$ et toute base orthonormée $B = \{\mathbf{b}_j\}_{j=1}^N$ on a :

$$\sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle| \geq 1.$$

L'égalité est valable si et seulement si

$$\exists j \in [1, N], \mathbf{a}_i = \pm \mathbf{b}_j.$$

Dans ce cas là, un vecteur de la base B est colinéaire au vecteur $\mathbf{a}_i$. Les produits scalaires avec les autres vecteurs de cette base sont donc nuls. On a donc :

$$\sum_{i=1}^N \sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle| = N \Leftrightarrow \forall j \in [1, N] \quad \exists i \in [1, N], \mathbf{a}_i = \pm \mathbf{b}_j$$

Les deux bases A et B sont donc identiques.

□

Propriété. Soient deux bases orthonormées $A = \{\mathbf{a}_i\}_{i=1}^N$ et $B = \{\mathbf{b}_j\}_{j=1}^N$ de $\mathbb{R}^N$.

$$d(A, B) \leq N(\sqrt{N} - 1)$$

Démonstration. On considère un vecteur normé $\mathbf{a}_i$ exprimé dans la base $B = \{\mathbf{b}_j\}_{j=1}^N$. On cherche à maximiser $\sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle|$ avec la contrainte $\sum_{j=1}^N \langle \mathbf{a}_i, \mathbf{b}_j \rangle^2 = 1$. La somme $\sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle|$ est maximisée lorsque tous les $|\langle \mathbf{a}_i, \mathbf{b}_j \rangle|$ sont égaux entre eux et comme $\sum_{j=1}^N \langle \mathbf{a}_i, \mathbf{b}_j \rangle^2 = 1$, il faut que pour tout $j \in [1, N]$, $|\langle \mathbf{a}_i, \mathbf{b}_j \rangle| = 1/\sqrt{N}$. Il faut donc que la cohérence du dictionnaire formé de l'union des vecteurs des deux bases A et B soit minimale. Dans ce cas là,

$$\sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle| = \frac{N}{\sqrt{N}} = \sqrt{N}$$

et

$$\sum_{i=1}^N \sum_{j=1}^N |\langle \mathbf{a}_i, \mathbf{b}_j \rangle| = N \sqrt{N}.$$

On a donc :

$$d(A, B) \leq N(\sqrt{N} - 1)$$

□

Bibliographie

- [1] P. Legai. *Impact de l'imagerie spatiale commerciale haute résolution sur la sécurité internationale dans sa dimension de défense - Perspectives pour l'Europe de la défense*. Thèse de doctorat, Université de Marne La Vallée, janvier 2003.
- [2] L. Nardon. L'arsenalisation de l'espace : projets américains, réactions européennes. Note de l'Ifri, octobre 2007.
- [3] L'aventure du programme SPOT. CNES Presse, novembre 2003.
- [4] C. Lambert-Nebout, C. Latry, G. Moury, C. Parisot, M. Antonini et M. Barlaud. On-board optical image compression for future high resolution remote sensing systems. Dans *Proceedings of SPIE*, tome 4115, p. 332–346. 2000.
- [5] C. Lambert Nebout, C. Latry et G. Moury. La compression embarquée d'images pour les systèmes optiques d'observation spatiale. Dans *GRETSI '01*. 2001.
- [6] C. Latry et B. Rouge. SPOT5 THR mode. Dans *Earth observing systems III*, p. 480–491. SPIE, San Diego, CA, 1998.
- [7] C. Latry et B. Rouge. Super resolution: quincunx sampling and fusion processing. Dans *International Geoscience and Remote Sensing Symposium*, tome 1, p. 315–317. IEEE, juillet 2003.
- [8] P. Lier, G. Moury, C. Latry et F. Cabot. Selection of the SPOT-5 image compression algorithm. Dans W. L. Barnes, rédacteur, *Earth Observing Systems III*, tome 3439-70, p. 541–552. SPIE, San Diego, octobre 1998.
- [9] CNES. Satellite PLEIADES-HR. URL <http://smc.cnes.fr/PLEIADES/>.
- [10] W. Sweldens. The lifting scheme: a construction of second generation wavelets. *SIAM Journal on Mathematical Analysis*, 29(2), p. 511–546, mars 1998.
- [11] C. Parisot, M. Antonini, M. Barlaud, C. Lambert-Nebout, C. Latry et G. Moury. On board strip-based wavelet image coding for future space remote sensing missions. Dans *International Geoscience and Remote Sensing Symposium*, tome 6, p. 2651–2653. IEEE, Honolulu, HI, jul 2000.

- [12] E. Ordentlich, M. Weinberger et G. Seroussi. A low-complexity modeling approach for embedded coding of wavelet coefficients. Dans *Data Compression Conference*, p. 408–417. Snowbird, UT, USA, mars 1998.
- [13] H. S. Malvar. Fast progressive wavelet coding. Dans *Data Compression Conference*, p. 336–343. Snowbird, UT, USA, mars 1999.
- [14] F. Gascon et M. Berger. GMES Sentinel-2 mission requirements document. Rapport technique, ESA, 2007.
- [15] J. Pearlman, S. Caramen, C. Segal, P. Jarecke, P. Clancy et W. Browne. Overview of Hyperion image spectrometer for the NASA EO-1 mission. Dans *International Geoscience and Remote Sensing Symposium*, tome 7, p. 3036–3038. IEEE, juillet 2001.
- [16] E. Christophe. *Compression des images hyperspectrales et son impact sur la qualité des données*. Thèse de doctorat, École Nationale Supérieure de l’aéronautique et de l’espace, octobre 2006.
- [17] B. L. Markham, J. C. Storey, D. L. Williams et J. R. Irons. Landsat sensor performance: history and current status. *IEEE Transactions on Geoscience and Remote Sensing*, 42(12), p. 2691–2694, 2004.
- [18] F. de Lussy, P. Kubik, D. Greslou, V. Pascal, P. Gigord et J. P. Cantou. PLEIADES-HR image system products, quality and geometric accuracy. Dans *ISPRS*. Hannover, 2005.
- [19] A. Baudoin. Beyond SPOT 5: PLEIADES, part of the french-italian program ORFEO. Dans *XXth ISPRS Congress*, p. 260–267. Istanbul, Turkey, juillet 2004.
- [20] V. K. Goyal. Theoretical foundations of transform coding. *IEEE Signal Processing Magazine*, p. 9–21, septembre 2001.
- [21] ISO/IEC JTC 1/SC 29/WG 1. *JPEG 2000 Part I Final Committee Draft Version 1.0*. JPEG 2000 Editor Martin Boliek, mars 2000.
- [22] C. Thiebaut, D. Lebedeff, C. Latry et Y. Bobichon. On-board compression algorithm for satellite multispectral images. Dans *Data Compression Conference*, p. 467. Snowbird, mars 2006.
- [23] I. P. Akam Bitá, M. Barret, F. Dalla Vedova et J. L. Guttwiler. Onboard compression of hyperspectral images using an exogenous orthogonal quasi-optimal transform. Dans *On-Board Payload Data Compression Workshop*. ESA/ESTEC, Noordwijk, The Netherlands, juin 2008.
- [24] S. Lloyd. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 28(2), p. 129–137, mars 1982.

-
- [25] F. Falzon et S. Mallat. Analysis of low bit image transform coding. *IEEE Transactions on Signal Processing*, 46(4), p. 1027–1042, avril 1998.
 - [26] G. K. Wallace. The JPEG still picture compression standard. *IEEE Transactions on Consumer Electronics*, 38(1), p. 18–34, février 1992.
 - [27] D. S. Taubman et M. W. Marcellin. *JPEG2000: Image compression fundamentals, standards and practice*. Kluwer Academic Publishers, 2001.
 - [28] D. Taubman. High performance scalable image compression with EBCOT. *IEEE Transactions on Image Processing*, 9(7), p. 1158–1170, juillet 2000.
 - [29] CCSDS. *Image Data Compression Recommended Standard - CCSDS 122.0-B-1 Blue Book*, novembre 2005.
 - [30] P.-S. Yeh, P. Armbruster, A. Kiely, B. Masschelein, G. Moury, C. Schaefer et C. Thiebaut. The new CCSDS image compression recommendation. *Aerospace Conference*, 5(12), p. 4138–4145, 2005.
 - [31] P. S. Eskicioglu, Ahmet M.; Fisher. A survey of quality measures for gray-scale image compression. Dans *Space and Earth Science Data Compression Workshop*, p. 49–61. NASA, Goddard Space Flight Center, 1993.
 - [32] A. Baudoin, M. Schroeder, C. Valorge, M. Bernard et V. Rudowski. The HRS-SAP initiative: A scientific assessment of the high resolution stereoscopic instrument on board of SPOT 5 by ISPRS investigators. Dans *XXth ISPRS Congress*. ISPRS, 2004.
 - [33] C. Thiebaut, X. Delaunay, C. Latry et G. Moury. CNES studies for on-board compression of high-resolution satellite images. Dans *Satellite Data Compression, Communication, and Processing IV*. SPIE Optics + Photonics, San Diego, CA, août 2008.
 - [34] X. Delaunay, C. Thiebaut et V. Charvillat. Compression embarquée d’images satellites : vers l’exploitation de la géométrie. Dans *COmpression et REpresentation des Signaux Audiovisuels*, p. 12–17. France Telecom R&D, Caen, France, novembre 2006.
 - [35] X. Delaunay, M. Chabert, G. Morin et V. Charvillat. Bit-plane analysis and contexts combining of JPEG2000 contexts for on-board satellite image compression. Dans *International Conference on Acoustics, Speech and Signal Processing*, p. 1057–1060. IEEE, Honolulu, HI, avril 2007.
 - [36] CCSDS. *Image Data Compression Informational Report CCSDS 120.1-G-1 Green Book*, juin 2007.
 - [37] J. Villasenor, B. Belzer et J. Liao. Wavelet filter evaluation for image compression. *IEEE Transactions on Image Processing*, 4(8), p. 1053–1060, août 1995.
 - [38] C. Parisot. *Allocations basées modèles et transformée en ondelettes au fil de l’eau pour le codage d’images et de vidéos*. Thèse de doctorat, Université de Nice-Sophia Antipolis, janvier 2003.

- [39] ISO/IEC JTC 1/SC 29/WG 1. *JPEG 2000 Part II Final Committee Draft*. JPEG 2000 Editor Martin Boliek, décembre 2000.
- [40] J. Liu et P. Moulin. Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients. *IEEE Transactions on Image Processing*, 10(11), p. 1647–1658, novembre 2001.
- [41] E. P. Simoncelli. Modeling the joint statistics of images in the wavelet domain. Dans *Proc SPIE, 44th Annual Meeting*, tome 3813, p. 188–195. Denver, CO, juillet 1999.
- [42] R. W. Buccigrossi et E. P. Simoncelli. Image compression via joint statistical characterization in the wavelet domain. *IEEE Transactions on Image Processing*, 8(12), p. 1688–1701, décembre 1999.
- [43] Z. Azimifar. *Image Models for Wavelet Domain Statistics*. Thèse de doctorat, University of Waterloo, Ontario, Canada, 2005.
- [44] J. Malo, I. Epifanio, R. Navarro et E. P. Simoncelli. Non-linear image representation for efficient perceptual coding. *IEEE Transactions on Image Processing*, 15(1), p. 68–80, janvier 2006.
- [45] J. M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12), p. 3445–3462, 1993.
- [46] A. Said et W. A. Pearlman. A new fast and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 6, p. 243–250, 1996.
- [47] X. Delaunay, C. Thiebaut, E. Christophe, R. Ruiloba, M. Chabert, V. Charvillat et G. Moin. Lossy compression by post-transforms in the wavelet domain. Dans *On-Board Payload Data Compression Workshop*. ESA/ESTEC, Noordwijk, The Netherlands, juin 2008.
- [48] R. Sudhakar, R. Karthiga et S. Jayaraman. Image compression using coding of wavelet coefficients - A survey. *ICGST International Journal on Graphics, Vision and Image Processing*, 5(6), p. 25–38, 2005.
- [49] D. Taubman, E. Ordentlich, M. Weinberger et G. Seroussi. Embedded block coding in JPEG2000. *Signal Processing: Image Communication*, 17(1), Elsevier, janvier 2002.
- [50] Z. Liu et L. Karam. Mutual information-based analysis of JPEG2000 contexts. *IEEE Transactions on Image Processing*, 14(4), p. 411–422, avril 2005.
- [51] A. Krivoulets, X. Wu et S. Forchhammer. On optimality of context modeling for bit-plane entropy coding in the JPEG2000 standard. Dans *Visual Content Processing and Representation*, tome 2849 de *Lecture Notes in Computer Science*, p. 208–216. Springer Verlag, septembre 2003.

-
- [52] H. Cai, X. San et J. Li. Embedded image coding with context partitioning and quantization. Dans *International Conference on Image Processing*, tome 2, p. 241–244. IEEE, Genova, Italy, septembre 2005.
- [53] X. Delaunay, M. Chabert, G. Morin et C. Thiebaud. Post-transformée dans le domaine ondelettes appliquée à la compression d’images satellite. Dans *GRETSI '07*, p. 1185–1188. CNRS, Troyes, France, septembre 2007.
- [54] S. D. Servetto, K. Ramchandran et M. T. Orchard. Image coding based on a morphological representation of wavelet data. *IEEE Transactions on Image Processing*, 8(9), p. 1161–1174, septembre 1999.
- [55] B.-B. Chai, J. Vass et X. Zhuang. Significance-linked connected component analysis for wavelet image coding. *IEEE Transactions on Image Processing*, 8(6), p. 774–784, juin 1999.
- [56] F. Lazzaroni, R. Leonardi et A. Signoroni. High-performance embedded morphological wavelet coding. *IEEE Signal Processing Letters*, 10(10), p. 293–295, octobre 2003.
- [57] E. J. Candès et D. L. Donoho. Curvelets - a surprisingly effective nonadaptive representation for objects with edges. Dans L. S. et al., rédacteur, *Curves and Surfaces*, p. 105–120. Vanderbilt University Press, Nashville, TN, 2000.
- [58] M. N. Do et M. Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing*, 14(12), p. 2091–2106, décembre 2005.
- [59] R. M. Figueras i Ventura. *Sparse image approximation with application to flexible image coding*. Thèse de doctorat, École Polytechnique Fédérale de Lausanne, 2005.
- [60] E. Candès et D. L. Donoho. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. *Communications on pure and applied mathematics*, 57(2), p. 219–266, 2004.
- [61] E. J. Candès. *Ridgelets: theory and applications*. Thèse de doctorat, Stanford University, 1998.
- [62] E. Candès, L. Demanet, D. Donoho et L. Ying. Fast discrete curvelet transforms. *Multiscale Modeling and Simulation*, 5(3), p. 861–899, 2006.
- [63] Y. Lu et M. N. Do. CRISP-contourlets: a critically sampled directional multiresolution image representation. Dans *Wavelets: Applications in Signal and Image Processing X*, tome 5207, p. 655–665. SPIE, San Diego, USA, août 2003.
- [64] G. Peyré. *Géométrie multi-échelle pour les images et les textures*. Thèse de doctorat, École Polytechnique, 2005.

- [65] G. Jeannic, V. Ricordel et D. Barba. Extraction de contours multirésolution pour un codage d'images par bandelettes. Dans *COmpression et REprésentation des Signaux Audiovisuels*. 2006.
- [66] E. Le Pennec. *Bandelettes et représentation géométrique des images*. Thèse de doctorat, École Polytechnique, décembre 2002.
- [67] E. Le Pennec et S. Mallat. Sparse geometric image representations with bandelets. *IEEE Transactions on Image Processing*, 14(4), p. 423–438, avril 2005.
- [68] B. K. Alpert. A class of bases in L^2 for the sparse representation of integral operators. *SIAM Journal on Mathematical Analysis*, 24, p. 246–262, 1993.
- [69] G. Peyré et S. Mallat. Surface compression with geometric bandelets. *ACM Transactions on Graphics, Proc. of SIGGRAPH'05*, 24(3), p. 601–608, 2005.
- [70] X. Delaunay, E. Christophe, C. Thiebaud et V. Charvillat. Best post-transforms selection in a rate distortion sense. Dans *International Conference on Image Processing*. IEEE, San Diego, CA, octobre 2008.
- [71] J. E. Fowler. QccPack: An open-source software library for quantization, compression, and coding. Dans A. G. Tescher, rédacteur, *Applications of Digital Image Processing XXIII*, p. 294–301. Proc. SPIE 4115, San Diego, CA, août 2000.
- [72] Y. Shoham et A. Gersho. Efficient bit allocation for an arbitrary set of quantizers. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(9), p. 1445–1453, septembre 1988.
- [73] K. Ramchandran et M. Vetterli. Best wavelet packet bases in a rate-distortion sense. *IEEE Transactions on Image Processing*, 2(2), p. 160–175, avril 1993.
- [74] B. M. Oliver, J. R. Pierce et C. E. Shannon. The philosophy of PCM. *Proceedings of the IRE*, 36(11), p. 1324–1331, novembre 1948.
- [75] H. Gish et J. Pierce. Asymptotically efficient quantizing. *IEEE Transactions on Information Theory*, 14(5), p. 676–683, 1968.
- [76] I. H. Witten, R. M. Neal et J. G. Cleary. Arithmetic coding for data compression. *Communications of the ACM, Computing Practice*, 30(6), p. 520–540, juin 1987.
- [77] G. Peyré et S. Mallat. Discrete bandelets with geometric orthogonal filters. Dans *International Conference on Image Processing*, tome 1, p. 65–68. IEEE, septembre 2005.
- [78] X. Delaunay, M. Chabert, V. Charvillat, G. Morin et R. Ruiloba. Satellite image compression by directional decorrelation of wavelet coefficients. Dans *International Conference on Acoustics, Speech, and Signal Processing*, p. 1193–1196. IEEE, Las Vegas, NV, avril 2008.

- [79] R. Gribonval et M. Nielsen. Sparse representations in unions of bases. Rapport de recherche de l'INRIA RR-4642, INRIA, Equipe METISS, Rennes, novembre 2002.
- [80] T. Strohmer et J. R. W. Heath. Grassmannian frames with applications to coding and communications. *Applied and Computational Harmonic Analysis*, 14(3), p. 257–275, mai 2003.
- [81] F. Kamangar et K. Rao. Fast algorithms for the 2-D discrete cosine transform. *IEEE Transactions on computers*, 31(9), p. 899–906, septembre 1982.
- [82] D.-T. Pham. Fast algorithms for mutual information based independent component analysis. *IEEE Transactions on Signal Processing*, 52(10), p. 2690–2700, octobre 2004.
- [83] I. P. Akam Bitá. *Sur l'application de l'analyse en composantes indépendantes à la compression des images multi composantes*. Thèse de doctorat, Université J. Fourier, février 2007.
- [84] S. Lesage, R. Gribonval, F. Bimbot et L. Benaroya. Learning unions of orthonormal bases with thresholded singular value decomposition. Dans *International Conference on Acoustics, Speech and Signal Processing*, tome V, p. 293–296. IEEE, mars 2005.
- [85] M. W. Marcellin et A. Bilgin. Quantifying the parent-child coding gain in zero-tree-based coders. *IEEE Signal Processing Letters*, 8(3), p. 67–69, mars 2001.
- [86] A. Said. Example of application for image compression, 1999. URL http://www.cipr.rpi.edu/~pearlman/papers/ex_spiht-ezw.pdf.
- [87] W. A. Pearlman. Example with speck algorithm, 2001. URL http://www.cipr.rpi.edu/~pearlman/papers/speck_example.pdf.