HAL
open science

Sur une approche de l'analyse en composantes indépendantes à la compression des images multi composantes

Isidore Paul Akam Bita

▶ **To cite this version:**

Isidore Paul Akam Bita. Sur une approche de l'analyse en composantes indépendantes à la compression des images multi composantes. domain_other. Université Joseph-Fourier - Grenoble I, 2007. Français. NNT : . tel-00151987

HAL Id: tel-00151987

https://theses.hal.science/tel-00151987v1

Submitted on 5 Jun 2007

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

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

Université Joseph Fourier de Grenoble

THESE

pour obtenir le grade de

DOCTEUR DE L'UJF

Spécialité : "mathématiques appliquées"

préparée au laboratoire Jean KUNTZMANN (LJK)
et à Supélec campus de Metz dans l'équipe système de traitement des signaux (STS)

dans le cadre de **l'Ecole Doctorale "mathématiques, sciences et technologie de l'information, informatique"**

présentée et soutenue publiquement par

Isidore Paul AKAM BITA

le 12 Février 2007

Titre :

Sur une approche de l'analyse en composantes indépendantes à la compression des images multi composantes

---

Directeur de thèse : Dinh-Tuan PHAM

Co-directeur de thèse : Michel BARRET

JURY

M.	Christian JUTTEN	Professeur UJF Grenoble	Président
M.	Pierre DUHAMEL	Directeur de recherche CNRS	Rapporteur
M.	Jean-Christophe PESQUET	Professeur Université Marne la vallée	Rapporteur
M.	Dinh-Tuan PHAM	Directeur de recherche CNRS	Directeur de thèse
M.	Michel BARRET	Maître de conférence Supélec	Co-directeur de thèse
Dr.	Cathérine LAMBERT	Chef service au CNES	Examinateur

Université Joseph Fourier de Grenoble

THESE

pour obtenir le grade de

DOCTEUR DE L'UJF

Spécialité : "mathématiques appliquées"

préparée au laboratoire Jean KUNTZMANN (LJK)
et à Supélec campus de Metz dans l'équipe système de traitement des signaux (STS)

dans le cadre de **l'Ecole Doctorale"mathématiques, sciences et technologie de l'information, informatique"**

présentée et soutenue publiquement par

Isidore Paul AKAM BITA

le 12 Février 2007

Titre :

Sur une approche de l'analyse en composantes indépendantes à la compression des images multi composantes

---

Directeur de thèse : Dinh-Tuan PHAM
Co-directeur de thèse : Michel BARRET

JURY

M.	Christian JUTTEN	Professeur UJF Grenoble	Président
M.	Pierre DUHAMEL	Directeur de recherche CNRS	Rapporteur
M.	Jean-Christophe PESQUET	Professeur Université Marne la vallée	Rapporteur
M.	Dinh-Tuan PHAM	Directeur de recherche CNRS	Directeur de thèse
M.	Michel BARRET	Maître de conférence Supélec	Co-directeur de thèse
Dr.	Cathérine LAMBERT	Chef service au CNES	Examinateur

Remerciements

Je remercie le Seigneur DIEU tout puissant, qui m'a donné la force, le courage et l'intelligence pour que je puisse mener à bien mes études depuis ma tendre enfance jusqu'à ce jour.

Je remercie Mr Christian JUTTEN, Professeur à l'unversité Josepf Fourier de Grenoble qui a accepté d'examiner mon rapport de thèse et d'être le président du jury.

Je remercie Mr Pierre DUHAMEL, Directeur de recherche au CNRS et enseignant à Supélec d'avoir accepté d'être rapporteur de mon mémoire de thèse malgré son emploi de temps très chargé.

Je remercie Mr Jean-Christophe PESQUET, Professeur à l'université Marne la vallée d'avoir également accepté, malgré son emploi de temps très chargé, d'être rapporteur de cette thèse.

Je remercie le Dr Cathérine LAMBERT, chef du service Animation Technique et Futur au CNES, qui a accepté d'examiner ma thèse et qui a nous a fourni des images multi spectrales.

Je remercie mon directeur de thèse Mr Dinh Tuan PHAM, Directeur de recherche au CNRS, qui m'a soutenu, m'a aidé durant toute ma thèse et grace à qui, j'ai appris la rigueur mathématique, ainsi que les bases de la modélisation mathématique d'un problème physique.

Je remercie Mr Michel BARRET, co-directeur de ma thèse, Maître de conférence et enseignant-chercheur à Supélec, grace à qui j'ai obtenu cette thèse financée, et qui m'a fait confiance, m'a toujours soutenu et m'a appris la rigueur et la clarté dans la rédaction d'un document scientifique.

Je remercie le LJK et Supélec Campus de Metz, qui ont mis les moyens matériels et humains à ma disposition durant ces années de thèse. Un merci particulier à Claudine Meyrieux, qui a toujours pris un peu de son temps, pour m'aider à résoudre tous mes problèmes administratifs.

Je remercie enfin tous ceux qui de près ou de loin m'ont apporté leur soutien durant ces travaux de thèse et dont la liste est loin d'etre exhaustive.

Dédicaces

Je dédie ce travail de thèse à mon fils Thierry Stéphane Akam Bita, qui a été durant tous ces travaux de thèse, ma principale source d'inspiration et de motivation.

Je dédie ce travail à la famille Bita Akam, mon père , ma mère qui m'ont encouragé et m'ont donné tous les moyens disponibles, pour que je puisse arriver à ce niveau, veuillez trouver en cette thèse, le fruit de votre dur labeur.

Je dédie ce travail à mes frères et soeurs, Jean Paul Edou, Joelle Ekoudi Meye, Jeanne Claire Mbe Bita et Jean Paul Bita Bita qui ont, malgré la distance et l'éloignement, su m'encourager à tenir bon. Trouver en ce travail, non pas un achèvement en soi, mais une oeuvre à atteindre et à dépasser.

Je dédie également ce travail à la famille Ela Evina, mon oncle et ma tante, qui grace à leurs conseils et à leur aide ont toujours été là pour moi quand il le fallait, veuillez trouver en ce travail, le fruit de vos efforts.

Je dédie enfin ce travail à Michelle Atyame et à Monique Nguini, qui ont été là durant tout ce temps, et qui ont parfois, eu à supporter mes caprices.

A mes tous mes amis, qui m'ont assisté de près ou de loin d'une façon quelconque, veuillez trouver ici ma profonde gratitude. Une pensée particulière vont à l'encontre de Nouidui Thierry, Koum Charles, Jongo Dora, Sy Ousmane, Nahal Mehdi, Djeutie Adeline, Tcheungo Martial, Deuheula Eric ...

Résumé

Dans une première partie, nous définissons plusieurs schémas de compression en partant de l'état de l'art qu'est JPEG 2000 aujourd'hui. Les schémas de compressions que nous avons définis proposent d'utiliser une transformation pour la réduction de la redondance spatiale transformée en ondelette discrète (TOD) et une autre transformation pour réduire la redondance spectrale. Les transformations optimales sous les hypothèses faible distortion, permettant de réduire la redondance spectrale, s'obtiennent dans certains cas en minimisant un critère qui peut être interprété comme le critère de l'analyse en composantes indépendantes (ACI) (minimisation de l'information mutuelle) additionné d'un terme toujours positif ou nul qui est une certaine mesure à l'orthogonalité de la transformation obtenue. Les performances obtenues en intégrant ces transformations dans nos schémas de compression montrent une amélioration des performances en comparaison à la transformation de Karhunen Loève (TKL). Dans la deuxième partie, nous proposons un modèle de mélange convolutif pour rechercher une transformation unique réduisant à la fois les redondances spatiales et spectrales. Nous définissons le critère à minimiser sous les hypothèses faibles distortions et nous montrons que ce critère peut s'interprété comme celui de l'ACI pour la séparation et déconvolution lorsque le critère à minimiser est l'information mutuelle auquel s'additionne un terme toujours positif ou nul. Puis nous proposons deux algorithmes permettant d'obtenir d'une part la transformation minimisant le critère dans le cas général, et d'autre part celle qui minimise le critère sous la contrainte que la distorsion dans le domaine transformée est la même que celle du domaine de l'image.

Mots clés : *Compression, TOD, images multi-composantes, TKL, ACI, séparation et déconvolution, débit, distorsion, PRSB (Pic du rapport signal à bruit), allocation optimale de bits, faible distorsion, haut débit, formule d'approximation de la distorsion.*

Abstract

In the first part of this work, we define several compression schemes based on JPEG2000 using two different transforms to reduce the spatial and spectral redundancies. The first one is the DWT (Discrete Wavelet Transform) that is well known with his properties to reduce in a good way the spatial redundancy between the wavelet coefficients. We develop the criterion to be minimized in order to obtain the optimal spectral transform of those compression schemes. This criterion equals the criterion used in independent component analysis (ICA) (when the mutual information is minimized) added with another term that could be seen as a measure of the transformation matrix to the orthogonality. Indeed, this additional term is always positive and becomes equal to zero if and only if the linear transform obtained is orthogonal. The results show that the new transforms perform better than the KLT (Karhunen Loève transform) that is usually used in compression to reduce the spectral redundancy. In the second part of our work, we define one new criterion in compression using a separation deconvolution modeling. This criterion is simply linked to the criteria that is minimized in ICA (when using mutual information like criterion to minimize) with a relation closed to that of the first part. We have evaluated the gradient and the Hessian of that criterion and choose to use a BFGS algorithm for the minimization of the criteria. Two algorithms were settled, the first one that minimizes the criteria in the general case and the second one, that minimizes the criteria with a relative simple constraint.

Keywords : *Compression, DWT, multi component images , KLT, ICA, séparation and déconvolution, rate, distortion, PSNR (Peak of signal noise ratio), optimal bits allocation, high rate, approximation of the distortion formulae.*

Table des matières

Remerciements **ii**

Dédicaces **iii**

Résumé **iv**

Abstract **iv**

Introduction générale **1**

Définitions, notions et rappels en compression d'images et en ACI **3**
0.1 Définitions et notions générales en compression d'images 3
0.2 Codeurs et normes en compression d'images 6
0.2.1 Quelques codeurs usuels en compression 6
0.2.2 La norme JPEG 6
0.2.3 Le standard JPEG2000 7
0.3 Brèves notions d'indépendance et description de quelques méthodes d'ACI 8
0.3.1 Brèves notions de l'indépendance statistique 8
0.3.2 Quelques méthodes d'ACI 9
0.3.2.1 La minimisation de l'information mutuelle 10
0.3.2.2 La maximisation de la non gaussianité 11
0.3.2.3 L'estimation du maximum de vraisemblance 12
0.3.2.4 Les méthodes tensorielles 13
0.4 L'ACI et la compression des images 15

I Étude et évaluation des performances des systèmes de compression utilisant des transformations modifiées à base d'ACI **16**
Introduction 17

1 Description globale du schéma de compression étudié, approximation de la distorsion **18**
1.1 Structures de compression 18
1.1.1 Notations 18
1.1.2 Schéma séparable 19
1.1.3 Schéma en sous-bandes 20

1.1.4 Schéma mixte en sous-bandes . . . 20
1.1.5 Intérêt de ces schémas de compression . . . 22
1.2 Formule d'approximation de la distorsion . . . 23
1.2.1 Cas général . . . 23
1.2.2 Approximation de la distorsion du schéma mixte en sous-bandes . . . 25
1.2.3 Approximation de la distorsion du schéma en sous-bandes . . . 27
1.2.4 Approximation de la distorsion du schéma séparable . . . 28
1.3 Calcul des coefficients de pondération w_m . . . 28

2 Gain de codage et réduction maximale de débit généralisés 33
2.1 Définitions des gains de codage généralisés . . . 33
2.1.1 Réduction maximale de débit généralisée . . . 34
2.1.2 Gain de codage généralisé . . . 34
2.2 Allocation optimale asymptotique de bits du schéma mixte en sous-bandes . . . 35
2.3 Allocation optimale de bits à bas débit . . . 36
2.4 Transformations optimisant le gain de codage à faible distorsion . . . 37
2.4.1 Cas simplifié . . . 37
2.4.2 Cas du schéma en sous-bandes . . . 39
2.4.3 Cas du schéma séparable . . . 40
2.4.4 Complexité algorithmique de *Ica_opt* et *Ica_orth* . . . 42

3 Evaluation des performances sur quelques images multi composantes 44
3.1 Description de la base de données d'images . . . 44
3.2 Validation des formules d'approximation de la distorsion . . . 46
3.3 Évaluation des gains de codage généralisés . . . 51
3.3.1 Résultats obtenus en utilisant un estimateur de l'entropie . . . 52
3.3.1.1 Cas du schéma en sous-bandes . . . 53
3.3.1.2 Schéma mixte en sous-bandes avec $P = 4$ . . . 61
3.3.1.3 Schema séparable avec adaptation sur l'espace ondelette . . . 65
3.3.1.4 Schéma séparable avec adaptation sur l'espace image . . . 69
3.3.2 Résultats obtenus en utilisant le VM9 . . . 73
3.3.2.1 cas du schéma séparable avec adaptation sur l'espace image . . . 73
3.3.3 Discussion et interprétations des résultats . . . 77
3.3.3.1 Schéma séparable . . . 77
3.3.3.2 Schéma en sous-bandes . . . 80
3.3.3.3 Schéma mixte en sous-bandes pour $P = 4$ . . . 82
3.3.3.4 Schéma séparable avec adaptation dans l'espace image + $VM9$ . . . 83

II Étude et évaluation des performances d'un système de compression utilisant des transformations de séparation-déconvolution 86
Introduction . . . 87

4 Définition du critère en compression pour un modèle de mélange convolutif 88
4.1 Rappels sur le modèle convolutif . . . 88
4.2 Description du schéma de compression, formule de la distorsion . . . 89
4.3 Le critère à optimiser en compression . . . 91
4.3.1 Relations entre le filtre de séparation et son inverse . . . 92
4.3.2 Définition du critère à minimiser . . . 93
4.3.3 Lien avec le critère de séparation-déconvolution basé sur l'information mutuelle 94
4.3.4 Invariance du critère . . . 96

5 Minimisation du critère 98
5.1 Calcul du gradient du critère . . . 98
5.1.1 Développement de Taylor au second ordre de C_1 . . . 98
5.1.2 Développement de Taylor au second ordre de C_2 . . . 100
5.1.3 Calcul du gradient du critère . . . 101
5.2 Définition du critère sous contrainte, calcul de son gradient . . . 103
5.2.1 Définition du critère sous contrainte . . . 103
5.2.2 Calcul du gradient du critère $C_{contrainte}$ . . . 104
5.3 Estimation empirique du gradient et du critère . . . 105
5.4 Algorithmes BFGS pour la minimisation de C et $C_{contrainte}$ . . . 107
5.5 Complexité algorithmique . . . 108

6 Evaluation des performances des transformations convolutives 110
6.1 Evaluation des Performances de la transformation *Ica_conv_opt* . . . 110
6.1.1 Validation de la formule d'approximation de la distorsion . . . 110
6.1.2 Evaluation et comparaison des gains de codage généralisés . . . 114
6.1.3 Interprétations et discussions des résultats . . . 123

Conclusion générale 124

Annexe1 132

Annexe2 147

Annexe3 151

Table des figures

1 Schéma général d'un système de compression d'images 5

1.1 TOD d'une image et choix d'une sous-bande 21
1.2 Découpage des sous-bandes en blocs de taille P 21

3.1 Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma en sous-bandes 47
3.2 Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma séparable avec adaptation dans l'espace ondelette 48
3.3 Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma séparable avec adaptation dans l'espace image 49
3.4 Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma mixte en sous-bandes avec $P = 4$ 50
3.5 Courbe débit vs PRSB de l'image Moissac : cas du schéma en sous-bandes 53
3.6 Gain de codage généralisé de l'image Moissac : cas du schéma en sous-bandes 54
3.7 Courbe débit vs PRSB de l'image Vannes : cas du schéma en sous-bandes 55
3.8 Gain de codage généralisé de l'image Vannes : cas du schéma en sous-bandes 56
3.9 Courbe débit vs PRSB de l'image de référence de Montpellier : cas du schéma en sous-bandes 57
3.10 Courbe débit vs PRSB de Montpellier décalée de 0.2 pixel : cas du schéma en sous-bandes 57
3.11 Courbe débit vs PRSB de Montpellier décalée de 0.35 pixel : cas du schéma en sous-bandes 58
3.12 Courbe débit vs PRSB de Montpellier décalée de 0.5 pixel : cas du schéma en sous-bandes 58
3.13 Courbe débit vs PRSB de l'image hyper-spectrale Jasper avec $N_1 = 45$, $N_2 = 44$: cas du schéma en sous-bandes 59
3.14 Gain de codage généralisé de l'image hyper-spectrale Jasper avec $N_1 = 45$, $N_2 = 44$: cas du schéma en sous-bandes 60
3.15 Courbe débit vs PRSB de l'image Moissac : cas du schéma mixte en sous-bandes avec $P = 4$ 61
3.16 Gain de codage généralisé de l'image Moissac : cas du schéma mixte en sous-bandes avec $P = 4$ 62
3.17 Courbe débit vs PRSB de l'image Vannes : cas du schéma mixte en sous-bandes avec $P = 4$ 63

3.18 Gain de codage généralisé de l'image Vannes : cas du schéma mixte en sous-bandes avec $P = 4$ . . . 64
3.19 Courbe débit vs PRSB de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace ondelette . . . 65
3.20 Gain de codage généralisé de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace ondelette . . . 66
3.21 Courbe débit vs PRSB de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace ondelette . . . 67
3.22 Gain de codage généralisé de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace ondelette . . . 68
3.23 Courbe débit vs PRSB de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace image . . . 69
3.24 Gain de codage généralisé de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace image . . . 70
3.25 Courbe débit vs PRSB de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace image . . . 71
3.26 Gain de codage généralisé de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace image . . . 72
3.27 Courbe débit vs PRSB de l'image Moissac : schéma séparable avec adaptation dans sur l'espace image + EBCOT . . . 73
3.28 Gain de codage généralisé de l'image Moissac : schéma séparable avec adaptation sur l'espace image + EBCOT . . . 74
3.29 Courbe débit vs PRSB de l'image Vannes : schéma séparable avec adaptation sur l'espace image + EBCOT . . . 75
3.30 Gain de codage généralisé de l'image Vannes : schéma séparable avec adaptation sur l'espace image + EBCOT . . . 76

4.1 Schéma général d'un système de compression simplifié . . . 89

6.1 Distorsion relative en fonction du PRSB des sous-bandes de l'image Vannes . . . 111
6.2 Distorsion relative en fonction du PRSB de l'image Toulouse décalée de 0.2 pixel . . 112
6.3 Distorsion relative en fonction du PRSB de l'image Toulouse décalée de 0.5 pixel . . 113
6.4 Gain de codage généralisé de la sous-bande HH2 de Montpellier décalée de 0.2 pixel 115
6.5 Gain de codage généralisé de la sous-bande HH2 de Montpellier décalée de 0.2 pixel 116
6.6 Gain de codage généralisé de la sous-bande HL2 de Toulouse décalée de 0.5 pixel . . 117
6.7 Gain de codage généralisé de la sous-bande HL2 de Toulouse décalée de 0.5 pixel . . 118
6.8 Gain de codage généralisé de la sous-bande LH2 de vannes . . . 119
6.9 Gain de codage généralisé de la sous-bande LH2 de Vannes . . . 120
6.10 Gain de codage généralisé de la sous-bande HH1 de Moissac . . . 121
6.11 Gain de codage généralisé de la sous-bande HH1 de Moissac . . . 122
6.12 A gauche Image de Paris, à droite image de Colorado. . . . 147
6.13 A gauche Image de Tokyo, à droite image de Rio. . . . 147
6.14 A gauche Image de Toulouse1, à droite image de Toulouse2. . . . 148
6.15 A gauche Image Jasper, à droite image de Raw. . . . 148
6.16 A gauche Image de Moffett, à droite image de Cuprite. . . . 149
6.17 De gauche à droite, Moissac, Strasbourg, Vannes, Pgx. . . . 149

Liste des tableaux

3.1 Images multi spectrales fournies par le CNES 45
3.2 Images multi spectrales de référence fournies par le CNES 45
3.3 Images hyperspectrales AVIRIS fournies par la NASA 45
3.4 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma séparable avec adaptation dans l'espace ondelette 78
3.5 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma séparable avec adaptation dans l'espace image 78
3.6 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma séparable avec adaptation dans l'espace ondelette 79
3.7 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma séparable avec adaptation dans l'espace image 79
3.8 Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace ondelette pour $N_1 = 20$ 79
3.9 Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace image pour $N_1 = 20$ 80
3.10 Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace ondelette pour $N_1 = 224$ 80
3.11 Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace image pour $N_1 = 224$ 80
3.12 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma en sous-bandes 81
3.13 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma en sous-bandes 81
3.14 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma en sous-bandes avec $N_1 = 224$ 82
3.15 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma en sous-bandes avec $N_1 = 45$ 82
3.16 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma en sous-bandes avec $N_1 = 20$ 82
3.17 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma mixte en sous-bandes 82

3.18 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma mixte en sous-bandes 82
3.19 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma mixte en sous-bandes avec $N_1 = 20$ 83
3.20 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma séparable avec adaptation dans l'espace image + VM9 84
3.21 Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma séparable avec adaptation dans l'espace image + VM9 pour $N_1 = 224$ 84

6.1 PRSB des images de 8 bpp du schéma en sous-bandes 133
6.2 PRSB des images de 12 bpp du schéma en sous-bandes 134
6.3 PRSB des images de 16 bpp du schéma en sous-bandes 135
6.4 PRSB des images de 8 bpp du schéma séparable avec adaptation dans l'espace ondelette 136
6.5 PRSB des images de 12 bpp du schéma séparable avec adaptation dans l'espace ondelette 137
6.6 PRSB des images de 16 bpp du schéma séparable avec adaptation dans l'espace ondelette 138
6.7 PRSB des images de 8 bpp du schéma séparable avec adaptation dans l'espace image 139
6.8 PRSB des images de 12 bpp du schéma séparable avec adaptation dans l'espace image 140
6.9 PRSB des images de 16 bpp du schéma séparable avec adaptation dans l'espace image 141
6.10 PRSB des images de 8 bpp du schéma mixte en sous-bandes pour $P = 4$ 142
6.11 PRSB des images de 12 bpp du schéma mixte en sous-bandes pour $P = 4$ 143
6.12 PRSB des images de 16 bpp du schéma mixte en sous-bandes pour $P = 4$ 144
6.13 PRSB des images de 12 bpp du schéma séparable avec adaptation dans l'espace image + $VM9$ 145
6.14 PRSB des images de 16 bpp du schéma séparable avec adaptation dans l'espace image + VM9 146

Introduction générale

Le développement sans cesse croissant des techniques de communications numériques et de leurs applications nécessite le stockage et la transmission d'une quantité d'informations qui croît exponentiellement. Cette information est stockée sous forme de données numériques. Les méthodes de transmission et de stockage sont de plus en plus performantes, mais elles ont du mal à suivre la croissance de la quantité des données due aux nouvelles applications et à leurs contraintes de qualité. Pour répondre à la gestion de cette quantité très importante de données, il convient d'améliorer les techniques de transmission et de stockage au moyen de méthodes de compression bien adaptées. Les systèmes de communication utilisant les satellites illustrent parfaitement ce problème. En effet, les images satellites sont des données très volumineuses et revêtent aujourd'hui une grande importance de par leur utilisation aussi bien en télédétection, que dans des domaines comme la géographie, l'océanographie, la cartographie des villes et bien d'autres encore. De par la demande sans cesse croissante de ces images avec des spécifications de plus en plus exigeantes, il convient de rechercher des techniques de compression adaptées pour satisfaire cette demande.

Cette thèse s'inscrit dans le cadre du projet[1] ACI^{2}M dont l'objectif principal est de faire une étude des nouvelles transformations réversibles et inversibles, susceptibles d'être utilisées dans les systèmes de compression d'images satellites, aussi bien à bord qu'au sol, dans une chaîne de codage utilisant un codeur progressif pouvant aller jusqu'au codage sans perte ou un codeur moins complexe. La description de ce projet est faite sur le site web[2] où l'on trouve également les différents membres y ayant travaillé. Seul le codage source est étudié dans ce projet, nous n'abordons pas le codage canal intervenant dans les systèmes de transmissions de données. Notre contribution dans ce projet et dont l'objet est cette thèse de doctorat, consiste à faire une étude pour la recherche de nouvelles transformations, en utilisant des techniques à base d'ACI (Analyse en Composantes Indépendantes), pour la compression des images satellites.

L'ACI est un outil qui fait partie du domaine plus vaste et plus général qu'est la séparation de sources. La séparation aveugle de sources, par opposition à la séparation non aveugle de sources, est un nouveau domaine dans la branche des mathématiques, plus précisément en traitement du signal et en statistiques, permettant à partir d'un mélange, de reconstituer des sources sans aucune connaissance a priori sur elles. Lorsque l'indépendance entre les différentes sources est utilisée comme critère, le problème de la séparation aveugle de sources devient similaire et se confond à celui de l'ACI. Pour certains auteurs [18], la séparation aveugle de sources, un peu plus ancienne, serait la même chose que l'ACI dont le concept est très récent et a été mis à jour et résolu de façon très explicite au début des années 90 par Comon [26]. Cependant, la notion d'ACI a d'abord été introduite par les travaux de [45] où une approche comparable à l'ACP (Analyse en Composantes

[1] Action Concertée Incitative Masse de Données

[2] http ://www.metz.supelec.fr/metz/recherche/aci2m/index.htm

Principales) a été utilisée. Pour éviter toute confusion, nous utiliserons indifféremment les notions d'ACI et de séparation aveugle de sources pour parler des techniques de séparation aveugle de sources basées sur l'indépendance statistique.

En compression d'images, l'objectif est de réduire ou d'exploiter autant que faire se peut la redondance qui existe entre les échantillons du signal à comprimer, au moyen de techniques diverses. Certaines méthodes d'ACI supposent implicitement que les observations sont obtenues à partir de sources indépendantes, d'autres méthodes ne font pas cette hypothèse. Elles recherchent une transformation qui minimise l'information mutuelle entre les composantes transformées à partir d'un modèle de mélange. Comme nous pouvons le voir, il y a une similarité entre l'ACI et la compression dans le sens où la première permet de réduire l'information mutuelle ou la dépendance entre les composantes transformées, et la deuxième cherche à minimiser ou à exploiter la redondance entre les composantes transformées ou non. Dans un système de compression sans perte, le codage par transformée consiste à rechercher une transformation réversible permettant de minimiser l'information mutuelle entre composantes transformées. En compression avec pertes, la redondance qu'il s'agit de minimiser pour réduire la taille du flot de bits portant l'information des données n'est plus l'information mutuelle, mais cette dernière augmentée d'un terme additif que nous préciserons par la suite, et qui est dû à l'introduction d'une distorsion acceptable [52]. Fort de cela, nous nous proposons dans cette thèse d'étudier des méthodes basées sur des techniques d'ACI et qui sont susceptibles d'avoir de bonnes performances en compression d'images multi composantes.

Ce rapport se subdivise en deux parties distinctes. Dans la première partie, nous définissons et étudions des schémas de compression dont la particularité est d'utiliser deux transformations pour la réduction de la redondance, l'une pour réduire la redondance dans une composante et l'autre pour réduire la redondance entre les composantes. Puis nous étendons aux images multi composantes les notions de gain de codage généralisé et de réduction maximale de débit généralisée qui ont été introduites par Narozny [52] dans sa thèse. Les travaux de thèse de ce dernier ont permis de mettre à jour, deux nouvelles transformations qui sont optimales, dans le sens qu'elles permettent de maximiser le gain de codage généralisé à faibles distorsions. Enfin, nous comparons les performances obtenues avec l'état de l'art.

La deuxième partie de ce rapport de thèse traite du problème de l'ACI appliquée à la compression en utilisant non plus un modèle de mélange instantané pour la réduction de la redondance, mais un modèle de mélange convolutif. Nous proposons une transformation basée sur de l'ACI convolutif permettant de maximiser le gain de codage généralisé à faibles distorsions. Une comparaison des performances obtenues en intégrant cette transformation dans un système de compression est faite, non seulement par rapport à l'état de l'art, mais aussi par rapport aux schémas utilisés dans la première partie.

Au préalable, nous rappelons quelques définitions et notions générales en compression dont certaines seront utilisées dans la suite du rapport et nous faisons une brève description de quelques méthodes d'ACI.

Définitions, notions et rappels en compression d'images et en ACI

Dans ce chapitre, nous rappelons quelques définitions générales et nous introduisons des notions propres à la compression des images. Par la suite, nous introduisons succinctement quelques codeurs et des normes qui sont bien connus dans la communauté de la compression des images. Après avoir rappelé quelques notions de statistiques et de la théorie de l'information, nous présentons une brève description de quelques méthodes utilisées pour résoudre le problème de l'ACI dans le cadre d'un modèle de mélange linéaire instantané. Ce chapitre s'achève par un bref état de l'art des travaux menés en compression d'images multi-composantes en y incorporant des transformations d'ACI.

0.1 Définitions et notions générales en compression d'images

Une image numérique est un signal 2D constitué de lignes et de colonnes, et dont l'élément de base est appelé pixel. La valeur d'un pixel représente l'intensité lumineuse spatialement localisée sur les coordonnées de ce pixel. Elle est en général un entier naturel borné et, la valeur maximale que peut prendre un pixel est liée à la **profondeur** ou à l'**amplitude** de l'image. Plus la profondeur d'une image est grande, plus grande est la quantité d'information que l'on peut extraire de cette image. Pour une image ayant une profondeur $N_b \in \mathbb{N}$, les valeurs des pixels sont des entiers naturels compris dans l'ensemble $[0, 2^{N_b} - 1]$. Les images naturelles en niveau de gris ont généralement une amplitude de 8 bpp (bits par pixel) alors que les images couleur naturelles ont généralement une amplitude de 24 bpp et sont constituées des composantes rouge, verte et bleue. La **résolution** d'une image est la taille de l'image en terme de nombre de lignes et nombre de colonnes. Pour une image contenant 1000 lignes et 500 colonnes, on dira que sa résolution est de 1000×500. Des définitions plus détaillées de ces notions se trouvent dans [39][50].

Une image satellitaire ou encore image multi-composante est généralement constituée de plusieurs composantes. Chaque composante est soit une image monochromatique prise autour d'une longueur d'onde en général située dans le visible ou l'infrarouge et dont la bande passante est assez étroite, soit une image panchromatique ayant une bande passante couvrant une grande partie du spectre. Des informations générales sur les images multi-composantes et sur le spectre de la lumière se trouvent dans [1]. Il existe plusieurs types d'images satellitaires différant par leurs nombres de composantes et par leurs modes d'acquisition :

- **Les images multi spectrales** : les différentes composantes sont prises les unes après les autres, le balayage sur chaque composante est réalisé en fonction du type de capteur. Des détails sur les modes d'acquisition et les capteurs utilisés se trouvent aux URL [2] [3]. Les images

multi spectrales comprennent généralement, tout au plus, une dizaine de bandes spectrales. La taille d'une composante est souvent assez importante (quelques millions de pixels par composantes) pour ce type d'image et, il existe souvent un problème de déregistration entre les différentes composantes dû non seulement au fait que les différentes composantes ne sont pas acquises en même temps, mais aussi provenant du fait que la terre et les satellites d'acquisition d'images sont perpétuellement en rotation, la première autour d'elle-même et autour du soleil et ces derniers autour de la terre. Le problème de déregistration est partiellement résolu par la bonne connaissance du système d'acquisition d'images et par des traitements au niveau du sol.

- **Les images hyper spectrales** : le capteur effectue une seule acquisition, puis celle-ci est diffractée au travers d'un prisme, ce qui permet d'avoir les différentes bandes spectrales. Le problème de déregistration entre les composantes n'existe pas pour ce type d'image. Une image hyper spectrale comporte généralement plusieurs centaines de bandes spectrales et couvre un spectre plus large avec des bandes contigües allant du visible à l'infrarouge [4] [5]. Ici, la taille d'une composante est moins importante, quelques centaines de milliers de pixels.
- **Les images super spectrales** : elles constituent la prochaine génération d'images satellitaires multi spectrales [6], et comprennent plus de bandes spectrales que les images multi spectrales actuelles, elles sont constituées de 10 à 50 bandes spectrales allant du visible à l'infrarouge. La taille d'une composante est du même ordre de grandeur que celle d'une image multi spectrale et les bandes spectrales ne sont pas contigües. Des détails sur le capteur d'images super spectrales ASTER se trouvent dans [7] [24].

Durant les premières missions d'acquisition des images multi-composantes, celles-ci avaient une profondeur de 8 bpp. De nos jours, avec l'évolution de la technologie, les images sont acquises avec une profondeur de 12 ou 16 bpp. Le point commun aux différentes bandes spectrales d'une image multi composante est que ces dernières représentent une seule et unique scène, mais obtenues pour différentes longueurs d'onde. Il existe naturellement, une ressemblance entre les différentes bandes spectrales d'une image puisque provenant de la prise d'une même scène, et l'on utilisera le terme de **redondance** pour désigner cette ressemblance. Par la suite, nous préciserons cette notion de redondance au moyen de la théorie de l'information. Dans une image multi composante, il existe deux types de redondance :

- La redondance intra-composante : c'est la redondance qui existe entre les différents pixels d'une bande spectrale, elle constitue la **redondance spatiale** dans chacune des composantes.
- La redondance inter-composante : c'est celle qui existe entre les différentes bandes spectrales de l'image multi-composante, elle constitue la **redondance spectrale** entre les différentes composantes de l'image.

En compression d'images multi-composantes, plus un système de compression minimisera ou tiendra compte de ces redondances dans la chaîne de compression, meilleures seront ses performances. Le schéma général d'un système de compression d'images multi-composantes peut se représenter comme sur la Fig. 1.

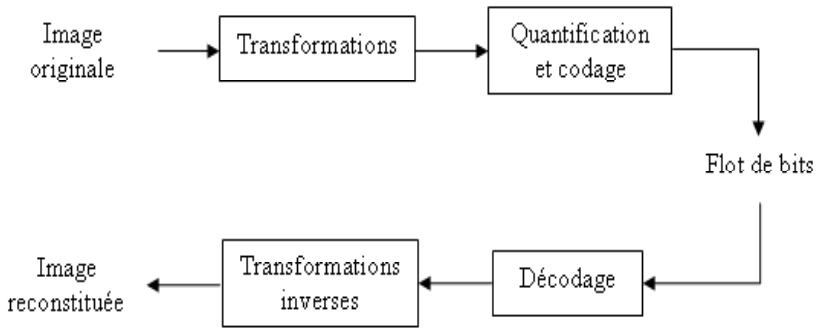


FIG. 1 – Schéma général d'un système de compression d'images

Le but des **transformations** est de représenter l'image dans un nouvel espace de représentation et ce afin de réduire autant que faire se peut les redondances spatiales et spectrales se trouvant dans l'image originale. Les méthodes de compression utilisant des transformations avant la quantification et le codage, sont connues sous le nom de **codage par transformées**. Après la transformation l'image est **quantifiée**, puis celle-ci est codée sans perte pour obtenir un **flot de bits** qui peut alors être transmis ou stocké sur un support. Un panorama exhaustif des méthodes de quantification est fait dans [31]. L'opération de codage et de quantification vise à exploiter la redondance résiduelle entre coefficients transformés tout en ayant une complexité acceptable. Il s'agit de trouver suivant l'application, le meilleur compromis entre complexité et réduction de la taille du flot de bits pour une distorsion donnée ou de façon duale, entre complexité et minimisation de la distorsion pour une taille de flot de bits donnée. Suivant l'utilisation qui en est faite, deux types de transformations sont appliquées : des transformations inversibles en général linéaires, et des transformations réversibles appelées aussi transformations d'entiers en entiers, qui sont généralement non linéaires. Ces dernières permettent de faire une **compression sans perte**, alors que les premières sont utilisées lorsque l'application permet la reconstitution d'une image décodée qui n'est qu'une approximation de l'image originale. Dans ce cas, on parle de **compression avec pertes** et la dégradation introduite est en général quantifiée, il s'agit de la **distorsion**.

Le **débit** est la taille moyenne en bits par pixel (bpp) nécessaire pour coder chaque pixel de l'image. La **distorsion** quant à elle permet d'avoir une mesure de la qualité de l'image reconstituée. Plus elle est faible, meilleure est la qualité de l'image reconstruite. Dans la suite de ce rapport, sauf mention explicite du contraire, nous utiliserons l'erreur quadratique moyenne comme mesure de distorsion. Soit $\mathbf{X}$, l'image originale ayant N composantes spectrales, et $\widehat{\mathbf{X}}$ l'image reconstituée, la distorsion moyenne D entre ces deux images est définie par :

$$D = \frac{1}{N} E\|\mathbf{X} - \widehat{\mathbf{X}}\|^2 \tag{1}$$

où $E(.)$ désigne l'espérance mathématique et $\|.\|$ est la norme euclidienne. Si N_L et N_C désignent les nombres de lignes et de colonnes de chacune des composantes, la distorsion est alors empiriquement évaluée par

$$D = \frac{1}{N N_L N_C} \sum_{i=1}^{M} \sum_{j=1}^{N_L} \sum_{k=1}^{N_C} [X_i(j,k) - \widehat{X}_i(j,k)]^2 \tag{2}$$

où $X_i(j,k)$ (respectivement $\widehat{X}_i(j,k)$) désigne le pixel de la $i^{ème}$ composante de l'image originale (respectivement de l'image reconstituée) spatialement localisé en (j,k). Dans la communauté de la compression d'images, on utilise plus généralement le PSNR (Peak Signal to Noise Ratio) ou PRSB (Pic du Rapport Signal sur Bruit) en Français, comme mesure de la qualité de la compression. Il est fonction de la distorsion D et est défini par :

$$PRSB = 10 \log_{10} \frac{(2^{N_b} - 1)^2}{D} \tag{3}$$

Pour un débit fixé, plus le PRSB est élevé, meilleure est la qualité de l'image reconstruite. Notons cependant qu'un PRSB élevé (pas de façon absolue, mais en fonction du type d'application), ne signifie pas forcément que l'on aura une bonne qualité visuelle de l'image reconstituée, même si cela semble être vrai pour la plupart des images naturelles. Il existe d'autres mesures de distorsion qui mettent en avant d'autres critères comme une préférence fréquentielle [10] ou qui prennent en compte des paramètres de la chaîne globale de compression tels que le débruitage ou la déconvolution, des exemples pour ces derniers cas sont explicités dans [57] [58].

0.2 Codeurs et normes en compression d'images

0.2.1 Quelques codeurs usuels en compression

La partie de quantification et de codage est une partie importante dans un schéma de compression. Comme nous l'avons déja souligné, un codeur efficace essaye d'exploiter autant que faire se peut, la redondance résiduelle dans les données après quantification (si celle-ci a lieu). Pour être efficace, le codeur doit tenir compte du type de quantification (scalaire ou vectorielle) qui précède. Un des codeurs les plus anciens et les plus simples à implanter est le codeur de Huffman, qui est un codeur entropique, et qui contrairement aux codeurs à longueur de code fixe, est un codeur à longueur de code variable. En effet, il attribue à chaque mot de la séquence à coder, un nombre de bits en fonction de sa fréquence d'apparition. Ainsi, plus un mot a une grande occurrence, plus faible est le nombre de bits avec lequel il est codé. Le codeur de Huffman est décrit dans [36]. Une introduction détaillée des notions de codage et de quantification pour la compression d'image est faite dans [12]. Le codeur arithmétique qui a été introduit dans [66] est un codeur entropique à longueur de code variable qui permet d'avoir de meilleures performances que le codeur de Huffman. Cette augmentation des performances du codeur arithmétique est obtenue avec une augmentation de la complexité algorithmique par rapport au codeur de Huffman. Outre ces deux codeurs qui sont largement connus, il y a aussi le codeur LZW qui est à l'origine du Zip qui est très répandu et utilisé de nos jours aussi bien pour la compression des fichiers textes que pour la compression d'autres types de fichiers. Le fonctionnement de ce codeur qui a subi entre temps plusieurs améliorations est décrit sous sa forme primaire dans [78] [79].

0.2.2 La norme JPEG

La norme JPEG [41] est un standard en compression d'images, elle a été finalisée vers le début des années 90. Dans le schéma de compression de base de la norme JPEG, la TCD (Transformée en Cosinus Discrète) est appliquée sur des blocs de taille 8×8 pour réduire la redondance spatiale entre pixels. La quantification est faite suivant une table de quantification psycho-visuelle dont les paramètres sont ajustables en fonction de la qualité visée. Après la quantification, on effectue

un balayage en zigzag de chaque bloc avant le codage. Pour cette dernière partie, le codeur RLE (Run Length Coding) qui est décrit dans [12] est utilisé, suivi du codeur entropique de Huffman. Ce standard continue d'être utilisé de nos jours aussi bien pour les images fixes que pour la vidéo avec les normes MPEG, MPEG2. Pour les images couleur, la transformation YCrCb qui permet d'obtenir les composantes dans le domaine luminance-chrominances, est appliquée sur les différentes composantes en vue de réduire la redondance spectrale, avant l'application de la DCT. Une des limitations de la norme JPEG est l'effet de blocs qui apparaît à bas débits. La norme JPEG, qui a eu beaucoup de succès avec l'avènement de l'Internet pour l'échange et le transfert d'images fixes, tend progressivement à disparaître au profit du nouveau standard JPEG2000.

0.2.3 Le standard JPEG2000

La norme JPEG2000 [42] [71] qui est le nouveau standard de compression d'images est toujours en cours d'élaboration. Dans sa version basique, elle utilise une TOD (Transformée en Ondelette Discrète) sur chaque composante pour la réduction de la redondance spatiale. Elle utilise une structure en arbre de zéros pour le codage, ainsi qu'un codeur arithmétique adaptatif, à modélisation de contexte. Le codeur utilisé est EBCOT (Embedded Bloc Coding Optimized Truncation) [72] dont l'une des particularités est la "scalabilité ", c'est-à-dire que l'on peut tronquer le flot de bits n'importe où, et reconstituer une image avec une qualité optimale par rapport au flot de bits qui a été utilisé par le décodeur. Un autre avantage indéniable de ce nouveau standard est le codage par région d'intérêt qui permet de coder une partie de l'image avec une qualité différente du reste de l'image. Pour les images multi-composantes, plusieurs options sont possibles, l'une d'elles est l'utilisation de la TKL (Transformée de Karhunen-Loeve) correspondant à une ACP (Analyse en Composantes Principales) entre les composantes pour la réduction de la redondance spectrale avant l'application d'une TOD par composante. Des discussions et des tests se poursuivent pour la finalisation complète de ce nouveau standard qui offre des perspectives et des applications à une échelle plus importante que JPEG. L'utilisation de la TOD permet de s'affranchir de l'effet de blocs observé à bas débits avec le standard JPEG.

D'autres codeurs avec des structures en arbre comme SPIHT [67] ont été envisagés et proposés pour la norme JPEG2000, mais ce dernier n'a pas été retenu. Sa structure en arbre, proche de celle de EBCOT, permet également d'envisager la propriété de scalabilité. Des variantes de SPIHT en compression d'images multi-composantes ont été introduites dans [40] avec le codeur SPECK. Des études sont en cours pour une version 3D de SPIHT [25] appliquée au codage des images hyper spectrales.

Notons enfin qu'il existe des normes particulières pour la compression à bord et au sol des images multi-composantes. Ces normes ne sont pas connues du grand public et sont souvent des hybrides de standards de compression, où certaines contraintes sont relâchées ou modifiées en fonction de l'évolution de la technologie. En effet, aux contraintes habituelles d'un standard de compression, il faut ajouter entre autres des contraintes de traitements en temps réel lors de l'acquisition de l'image, mais aussi des contraintes de pré et post-traitements effectués sur l'image lorsque celle-ci arrive au sol.

0.3 Brèves notions d'indépendance et description de quelques méthodes d'ACI

L'ACI est un domaine qui a été introduit récemment et qui présente des champs d'applications potentiels multiples et variés. Avant d'aborder quelques méthodes de résolution directement liées au problème l'ACI, nous rappelons quelques définitions sur la théorie de l'information et des statistiques qui sont utilisées en ACI.

0.3.1 Brèves notions de l'indépendance statistique

Dans ce rapport, $(\Omega, \mathcal{E}, P)$ désigne un espace probabilisé. Soient X et Y deux variables aléatoies définies sur $(\Omega, \mathcal{E}, P)$. Elles sont dites indépendantes si $\forall A \subset \Omega$, $\forall B \subset \Omega$, les événements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants, c'est-à-dire

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \tag{4}$$

où $P(X \in A, Y \in B)$ désigne la probabilité d'avoir l'événement $X \in A$ et $Y \in B$, $P(X \in A)$ est la probabilité d'avoir $X \in A$ et $P(Y \in B)$ est la probabilité d'avoir $Y \in B$.

Si X et Y sont deux variables aléatoires réelles, elles sont dites indépendantes quand $\forall (x, y) \in \mathbb{R}^2$ nous avons :

– $f_{XY}(x, y) = f_X(x)f_Y(y)$

dans le cas de variables à valeurs continues admettant une densité de probabilité jointe $f_{XY}(x, y)$, où $f_X(x)$ et $f_Y(y)$ sont les densités de probabilité marginales de X et de Y.

Si X et Y sont deux variables aléatoires discrètes à réalisation dans un ensemble dénombrable $\mathbb{D}$, elles sont indépendantes quand nous avons $\forall (x, y) \in \mathbb{D}^2$:

– $P_{XY}(x, y) = P_X(x)P_Y(y)$

où $P_{XY}(x, y)$ est la probabilité conjointe des événements $X = x$ et $Y = y$, $P_X(x)$ est la probabilité de $X = x$ et $P_Y(y)$ celle de $Y = y$.

L'entropie d'une variable aléatoire réelle X admettant une densité de probabilité est définie par :

$$H(X) = \int_{-\infty}^{+\infty} -f_X(x) \log_2 f_X(x) dx \tag{5}$$

L'entropie conjointe de deux variables aléatoires continues X et Y admettant une densité de probabilité conjointe est définie par :

$$H(X, Y) = \int_{-\infty}^{+\infty} -f_{XY}(x, y) \log_2 f_{XY}(x, y) dx dy \tag{6}$$

Ces définitions de l'entropie sont formellement les mêmes dans le cas des variables aléatoires discrètes, à la différence près que la densité de probabilité est remplacée par la distribution de probabilité, et l'intégrale par une somme discrète. Dans le cas des variables aléatoires continues, on parlera d'entropie différentielle, et nous parlerons de l'entropie de Shannon dans le cas des variables aléatoires discrètes.

L'entropie relative, encore appelée divergence de Kullback Leibler, de deux fonctions de densité de probabilité f et g est définie par :

$$D_{KL}(f \| g) = \int_{-\infty}^{+\infty} f(x) \log_2 \frac{f(x)}{g(x)} dx \tag{7}$$

Il est démontré dans [27] que la divergence de Kullback Leibler est toujours non négative et, elle est nulle si et seulement si $f(x) = g(x)\ \forall x \in \mathbb{R}$. Elle est parfois appelé distance de Kullback Leibler, mais ce n'est pas à proprement parler une mesure de distance dans le sens mathématique du terme entre les fonctions de densité de probabilité. En effet, par définition, la divergence de Kullback Leibler entre deux fonctions de densité de probabilité n'est pas symétrique de plus, elle ne satisfait pas l'inégalité triangulaire.

L'information mutuelle entre deux variables aléatoires X et Y est définie par :

$$I(X;Y) = D_{KL}(f_{XY} \| f_X f_Y) \tag{8}$$

où f_{XY} est la densité de probabilité conjointe de X et Y, f_X et f_Y sont les densités de probabilités marginales de X et Y respectives. Elle représente la divergence de Kullback Leibler entre la densité de probabilité conjointe de X et Y et le produit des densités marginales de X et Y. En utilisant la définition de l'information mutuelle précédente et en développant cette équation, il est démontré dans [27] que celle-ci peut encore s'écrire comme :

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{9}$$

Plus généralement, l'information mutuelle de N variables aléatoires $X_1, X_2, \ldots, X_N$ en utilisant cette nouvelle définition est donnée par :

$$I(X_1; X_2; \ldots; X_N) = \sum_{i=1}^{N} H(X_i) - H(X_1, X_2, \ldots, X_N) \tag{10}$$

L'information mutuelle est le critère par excellence de mesure de la dépendance entre variables aléatoires. C'est une quantité adimensionnelle qui est toujours positive ou nulle [27] de par la définition de l'équation (8). Elle est nulle si et seulement si les variables aléatoires sont indépendantes. L'information mutuelle peut-être interprétée physiquement dans le cas de variables aléatoires discrètes, comme une augmentation de la taille du flot de bits si l'on décide de faire une compression sans perte incluant un codage séparé des variables aléatoires. La divergence $D_{KL}(p(x)\|q(x))$ peut-être interprétée physiquement dans le cas d'une variable aléatoire discrète X de distribution de probabilités $P_X(x) = p(x)$, comme l'augmentation de la taille du flot de bits (Cf théorème 5.4.3 de [27]) si le codage de la variable aléatoire X est fait en utilisant la distribution de probabilités $q(x)$ au lieu de $p(x)$, d'où l'intérêt lors du codage de rechercher une distribution de probabilités $q(x)$ très proche de la vraie distribution de probabilités $p(x)$, de manière à réduire la taille du flot de bits final.

Deux variables aléatoires X et Y sont dites décorrélées lorsque l'égalité suivante est satisfaite :

$$E(XY) = E(X)E(Y) \tag{11}$$

Deux variables aléatoires indépendantes sont décorrélées, mais l'inverse n'est pas toujours vraie. Par la suite, nous décrivons quelques méthodes utilisées en ACI, auparavant nous rappelons le modèle de mélange qui est utilisée ici.

0.3.2 Quelques méthodes d'ACI

En séparation de sources, un modèle de mélange est choisi en général, en supposant qu'il existe des sources statistiquement indépendantes vérifiant ce modèle. Le modèle d'un mélange linéaire de

base en séparation aveugle de sources est :

$$\mathbf{X}(t) = \mathbf{A}\mathbf{S}(t) \tag{12}$$

où $\mathbf{X}(t) = [X_1(t), X_2(t), \ldots, X_N(t)]^T$ est le vecteur des observations, $\mathbf{S}(t) = [S_1(t), S_2(t), \ldots, S_N(t)]^T$ est le vecteur source, $\mathbf{A}$ est une matrice carrée appelée matrice de mélange, N est le nombre d'observations. Nous nous limitons au cas où il y a autant d'observations que de sources et, la matrice de mélange est inversible. Dans ce cas, les sources $\mathbf{S}$ ne peuvent être retrouvées qu'à un facteur d'échelle près et à une permutation près. En effet, si l'on multiplie la source $S_i(t)$ par un coefficient α_i de façon à obtenir une nouvelle source $\alpha_i S_i(t)$, et que l'on divise tous les éléments de la $i^{ème}$ colonne de $\mathbf{A}$ par α_i (ce qui se traduit mathématiquement par $A_{ji} \leftarrow A_{ji}/\alpha_i$, $1 \leq j \leq N$), l'équation du modèle (12) reste valide, on parle alors d'ambiguïté d'échelle. D'autre part, soit $\mathbf{P}$ une matrice de permutation, l'équation du modèle (12) peut encore s'écrire comme $\mathbf{X}(t) = (\mathbf{A}\mathbf{P}^{-1})(\mathbf{P}\mathbf{S}(t))$, où $\mathbf{P}\mathbf{S}(t)$ représente les sources inconnues, mais dans un ordre différent, on parle alors d'ambiguïté de permutation. Ces deux ambiguïtés peuvent être résolues en se fixant des contraintes assez simples qui ne changent pas fondamentalement le problème de l'ACI.

Dans le cas où il y'a autant d'observations que de sources et que la matrice carrée $\mathbf{A}$ est inversible, alors ne disposant que des observations, le vecteur source est recherché sous la forme :

$$\mathbf{Y}(t) = \mathbf{B}\mathbf{X}(t) \tag{13}$$

sous l'hypothèse que les composantes de $\mathbf{Y}(t) = [Y_1(t), Y_2(t), \ldots, Y_N(t)]^T$ sont statistiquement indépendantes. Par la suite, nous nous affranchirons de la variable temporelle, la raison principale étant qu'en pratique le vecteur d'observation $\mathbf{X}(t)$ peut être considéré comme stationnaire, donc en particulier la loi de $\mathbf{X}(t)$ ne dépend pas de la variable temporelle t, et nous désignerons par $\mathbf{Y}$ et $\mathbf{X}$ les vecteurs $\mathbf{Y}(t)$ et $\mathbf{X}(t)$ respectifs. Plusieurs critères et contrastes ont été proposés [37] pour résoudre ce problème. Dans une certaine mesure et sous certaines conditions, il est possible de démontrer que ces critères et contrastes deviennent équivalents [37]. Dans cette thèse, nous nous baserons sur certaines de ces méthodes pour définir et construire de nouveaux algorithmes à base d'ACI, qui convergent vers des transformations optimales (sous certaines conditions que nous préciserons) en compression d'images. Par la suite, nous allons présenter quelques méthodes généralement utilisées en ACI. Toutes ces méthodes et bien d'autres encore sont décrites et détaillées dans [37].

0.3.2.1 La minimisation de l'information mutuelle

En pratique, à partir d'un modèle de mélange, les sources sont recherchées de telle sorte qu'elles soient les plus indépendantes possible. Le critère par excellence de la mesure de dépendance étant l'information mutuelle, les sources estimées sont recherchées en minimisant leur information mutuelle. En se référant au modèle de l'équation (13), il s'agit de trouver la matrice inversible $\mathbf{B}$ qui minimise l'information mutuelle entre les composantes de $\mathbf{Y}$. En se rappelant d'après la théorie de l'information [27] que la relation entre les entropies jointes de $\mathbf{Y}$ et de $\mathbf{X}$ est donnée par :

$$H(Y_1, Y_2, \ldots, Y_N) = H(X_1, X_2, \ldots, X_N) + \log_2 |\det(\mathbf{B})| \tag{14}$$

il vient qu'il est encore possible d'écrire l'information mutuelle entre les composantes de $\mathbf{Y}$ comme

$$I(Y_1; Y_2; \ldots; Y_N) = \sum_{i=1}^{N} H(Y_i) - H(X_1, X_2, \ldots, X_N) - \log_2 |\det(\mathbf{B})| \tag{15}$$

Ainsi, minimiser l'information mutuelle $I(Y_1; Y_2; \ldots; Y_N)$ est équivalent à minimiser le critère

$$C(\mathbf{B}) = \sum_{i=1}^{N} H(Y_i) - \log_2 |\det(\mathbf{B})| \tag{16}$$

Ce critère a été utilisé par [59] pour mettre au point l'algorithme d'ACI `icainf` basé sur la minimisation de l'information mutuelle. Un panorama des méthodes d'ACI basées sur la minimisation de l'information mutuelle est fait dans [37] (chapitre 10).

0.3.2.2 La maximisation de la non gaussianité

Une des conséquences du théorème de la limite centrale [37] (section 2.5.2) est que densité de probabilité de la somme de plusieurs variables aléatoires indépendantes tend vers une distribution gaussienne. En d'autres termes, la densité de probabilité de la somme de plusieurs variables aléatoires indépendantes est plus gaussienne (dans le sens où elle tend vers une distribution gaussienne) que chacune des variables aléatoires d'origine. Ce résultat est très intéressant dans la mesure où il nous dit que chercher à rendre des variables aléatoires indépendantes est équivalent à maximiser la non gaussianité de ces variables [37] (section 8.1). Une question qu'on peut se poser à ce niveau est de savoir comment faire pour mesurer la non gaussianité d'une variable aléatoire.

Le kurtosis est une mesure classique de la non gaussianité, c'est une quantité adimensionnelle. Soit X une variable aléatoire réelle de moyenne nulle, son kurtosis que nous noterons $kurt(X)$ est défini par :

$$kurt(X) = \frac{E(X^4)}{[E(X^2)]^2} - 3 \tag{17}$$

Le kurtosis d'une variable aléatoire gaussienne est nulle. Si X est une variable non gaussienne, son kurtosis est non nul en général. Les variables aléatoires ayant un kurtosis strictement positif sont dites super gaussiennes et celles ayant un kurtosis strictement négatif sont dites sous gaussiennes. La définition du kurtosis et des cumulants croisés d'une variable aléatoire de moyenne quelconque est rappelée dans [37] (section 2.7). Le kurtosis d'une variable aléatoire X de moyenne nulle est le quotient entre le cumulant d'ordre 4 de X et le carré de la variance de X. Un inconvénient du kurtosis est que son estimation requiert souvent un échantillon de taille importante . Une mesure alternative de la non gaussianité est la néguentropie. La néguentropie d'une variable aléatoire Y que nous noterons $J(Y)$ se définit par

$$J(Y) = H(Y_{gauss}) - H(Y) \tag{18}$$

où Y_{gauss} est une variable aléatoire gaussienne de même variance que Y. Dans la cas d'une variable aléatoire vectorielle, la variance est remplacée par la matrice de variances covariances. La néguentropie est toujours positive et devient nulle si et seulement si Y est gaussienne. En faisant un développement de Gram-Charlier de la néguentropie autour d'une variable gaussienne, il est démontré [37] (pp. 114-115) qu'une approximation de la néguentropie d'une variable aléatoire de moyenne nulle et de variance égal à 1, est obtenue en utilisant les cumulants d'ordre supérieur par

$$J(Y) \approx \frac{1}{12} E(Y^3)^2 + \frac{1}{48} [E(Y^4) - 3]^2 \tag{19}$$

Cette approximation permet d'avoir un calcul simple et peu complexe de la néguentropie.

`FastICA` [38] est un algorithme bien connu utilisant la maximisation de la non gaussianité. Il existe deux versions de `FastICA` pour la résolution du problème de l'ACI, l'une utilisant une méthode séquentielle et l'autre qui utilise une méthode simultanée. La méthode séquentielle consiste à recherche les sources les unes après les autres. Soit $\mathbf{b}_i$ un vecteur ligne et qui représente la $i^{ème}$ ligne de la matrice $\mathbf{B}$ recherchée dans l'équation (13), le vecteur $\mathbf{b}_i$ est recherché de telle sorte qu'il maximise $kurt(Y_i)$ avec $Y_i = \mathbf{b}_i\mathbf{X}$, ce qui nous permet d'avoir une source Y_i. La prochaine source $Y_j,\ i \neq j$ est recherchée de la même façon, c'est-à-dire que l'on cherche le vecteur $\mathbf{b}_j$ qui maximise $kurt(Y_j)$ avec $Y_j = \mathbf{b}_j\mathbf{X}$, sous la contrainte que Y_i et Y_j soient décorrelés. En procédant de cette manière, on arrive ainsi à retrouver la matrice $\mathbf{B}$, la contrainte étant nécessaire pour éviter que l'on retrouve la même source plus d'une fois.

La résolution du problème de l'ACI suivant le modèle de l'équation (13) et en utilisant la méthode simultanée, consiste à rechercher la matrice $\mathbf{B}$ qui maximise en valeur absolue la somme du kurtosis des composantes de $\mathbf{Y}$ sous la contrainte que ses composantes soient décorrelés.

Le kurtosis n'est qu'une mesure de la non gaussianité, il est possible d'utiliser d'autres mesures comme la néguentropie par exemple, ou une fonction quelconque qui peut-être vue comme une mesure de la non gaussianité. En pratique, l'algorithme `FastICA` utilise une phase de blanchiment des données qui consiste à remplacer les observations $\mathbf{X}$ par $\widetilde{\mathbf{X}} = \mathbf{C}\mathbf{X}$ avec $Cov(\widetilde{\mathbf{X}}) = \mathbf{I}$, où $\mathbf{C}$ est une matrice carrée inversible, $\mathbf{I}$ est la matrice identité et $Cov(\widetilde{\mathbf{X}})$ est la matrice de covariance de $\mathbf{X}$. Les sources sont alors recherchées sous la forme $\mathbf{Y} = \mathbf{B}\widetilde{\mathbf{X}}$, $\mathbf{B}$ étant une matrice orthogonale, de telle sorte que leur néguentropie (ou leur kurtosis) soit maximisée d'une façon séquentielle ou d'une façon simultanée.

0.3.2.3 L'estimation du maximum de vraisemblance

L'objectif ici est de résoudre le problème de l'ACI en utilisant la méthode du maximum de vraisemblance. Nous avons posé $\mathbf{S} = \mathbf{B}\mathbf{X}$ en supposant que les composantes de $\mathbf{S}$ sont mutuellement indépendantes, où $\mathbf{B}$ est l'inverse de la matrice de mélange de l'équation (12). La densité de probabilité de la variable aléatoire $\mathbf{X}$ est donnée par :

$$f(\mathbf{x}) = |\det \mathbf{B}| \prod_{i}^{N} f_{S_i}((\mathbf{B}\mathbf{x})_i) \tag{20}$$

où f_{S_i} est densité de probabilité de S_i et $(\mathbf{B}\mathbf{X})_i$ est la $i^{ème}$ composante de $\mathbf{B}\mathbf{X}$. Cette égalité se justifie par l'hypothèse d'indépendance qui est faite entre les composantes de $\mathbf{S}$. Par la suite, pour simplifier les notations, nous posons $f_i = f_{S_i}$ et $\mathbf{B}^T = (\mathbf{b}_1, \mathbf{b}_2, \ldots \mathbf{b}_N)$ l'écriture vectorielle de la matrice $\mathbf{B}^T$. Supposons que nous disposons de M observations indépendantes et identiquement distribuées (i.i.d) de $\mathbf{X}$ que nous notons $\mathbf{X}(m)$ avec $1 \leq m \leq M$, la fonction de vraisemblance est alors donnée par :

$$L(\mathbf{B}) = \prod_{m=1}^{M} \prod_{i=1}^{N} f_i(\mathbf{b}_i^T \mathbf{x}(m)) |\det \mathbf{B}| \tag{21}$$

et la log-vraisemblance obtenue par le logarithme de cette dernière équation vaut

$$\log L(\mathbf{B}) = \sum_{m=1}^{M} \sum_{i=1}^{N} \log f_i(\mathbf{b}_i^T \mathbf{x}(m)) + M \log |\det \mathbf{B}| \tag{22}$$

En normalisant cette dernière expression par rapport au nombre de réalisations nous obtenons :

$$\log L_M(\mathbf{B}) = \frac{1}{M}\sum_{m=1}^{M}\sum_{i=1}^{N}\log f_i(\mathbf{b}_i^T\mathbf{x}(m)) + \log|\det\mathbf{B}| \tag{23}$$

Ainsi, l'estimateur du maximum de vraisemblance de la matrice de séparation $\mathbf{B}$ est $\widehat{\mathbf{B}}_{MV}$ qui maximise $\log L_M(\mathbf{B})$. La différentielle de la log-vraisemblance normalisée d'après [60] vaut :

$$d\log L_M(\mathbf{B}) = \frac{1}{M}\sum_{i,j=1}^{N}\sum_{m=1}^{M}\Psi_i(\mathbf{b}_i^T\mathbf{x}(m))\partial_{ij}\mathbf{b}_j^T\mathbf{x}(m) + \sum_{i=1}^{N}\partial_{ii} \tag{24}$$

où $[\partial_{ij}] = -d\mathbf{B}\mathbf{B}^{-1}$ $(1 \leq i,j \leq N)$ et $\Psi_i(u) = -d\log f_i(u)/du$ est la **fonction score** de la $i^{ème}$ source. Cette différentielle s'annule au maximum de la valeur de la log-vraisemblance, et ce maximum est obtenu en résolvant les $N(N-1)$ équations suivantes :

$$\frac{1}{M}\sum_{m=1}^{M}\Psi_i(\mathbf{b}_i^T\mathbf{x}(m))\mathbf{b}_j^T\mathbf{x}(m) = \delta(i-j),\ (1 \leq i \neq j \leq N) \tag{25}$$

où δ désigne le symbole de Kronecker[3]. Pour M très grand, l'équation (25) d'après la loi des grands nombres devient :

$$E[\Psi_i(\mathbf{b}_i^T\mathbf{x}(m))\mathbf{b}_j^T\mathbf{x}(m)] = \delta(i-j),\ (1 \leq i \neq j \leq N) \tag{26}$$

Il est alors possible de résoudre le problème de l'ACI connaissant la densité de probabilité des sources. Dans la pratique, l'on dispose souvent de très peu d'informations a priori sur cette densité de probabilité. Il est possible de relâcher cette connaissance a priori et de définir des densités de probabilités réelles quelconques pour chaque source comme cela a été fait dans [60] [21].

Le principe de l'Infomax [47] (information maximisation) est très proche du maximum de vraisemblance. Bell et Sejnowski [13] ont appliqué une approche neuronale utilisant l'Infomax pour résoudre le problème de l'ACI en maximisant l'entropie des sources. Certains auteurs [37] considèrent et démontrent que le principe de l'Infomax est quasiment équivalent au maximum de vraisemblance pour la résolution du problème simplifié de l'ACI.

0.3.2.4 Les méthodes tensorielles

Les algorithmes d'ACI utilisant les méthodes tensorielles sont basés sur l'annulation des cumulants croisés d'ordre 4. Faisons d'abord un bref rappel sur la définition de ces cumulants.

Soit X une variable aléatoire continue, de densité de probabilité $f(x)$. La fonction caractéristique $\varphi(w)$ est définie par

$$\varphi(w) = E[\exp(jwx)] = \int_{-\infty}^{+\infty}\exp(jwx)f(x)dx \tag{27}$$

[3]Le symbole de Kronecker est défini par :

$$\begin{aligned}\delta(n) &= 1,\ \ si\ \ n=0\\ \delta(n) &= 0,\ \ si\ \ n\neq 0\end{aligned}$$

avec $j^2 = -1$. La fonction caractéristique peut encore s'exprimer en utilisant le développement en série de Taylor au voisinage de $w = 0$ (quand cela est possible) comme suit :

$$\varphi(w) = \sum_{k=0}^{+\infty} E(X^k) \frac{(jw)^k}{k!} \tag{28}$$

où $E(X^k)$ (quand il existe) est le moment d'ordre k de la variable aléatoire X. La seconde fonction caractéristique se définit à partir de la première comme suit :

$$\phi(w) = \log(\varphi(w)) \tag{29}$$

Le développement en série de Taylor (quand il existe) de cette dernière s'exprime par

$$\phi(w) = \sum_{k=0}^{+\infty} \kappa_k \frac{(jw)^k}{k!} \tag{30}$$

où κ_k est le cumulant d'ordre k de X défini par :

$$\kappa_k = (-j)^k \frac{d^k \phi(w)}{dw^k}|_{w=0} \tag{31}$$

Dans le cas où X est une variable aléatoire de moyenne nulle, les quatre premiers cumulants valent $\kappa_1 = 0$, $\kappa_2 = E(X^2)$, $\kappa_3 = E(X^3)$ et $\kappa_4 = E(X^4) - 3[E(X^2)]^2$. La définition de ces cumulants dans le cas général où X est une variable aléatoire de moyenne non nulle se trouve dans [53] [32]. Dans ces derniers articles, sont également données, les définitions des cumulants croisés d'ordre 2, 3 et 4 d'une variable aléatoire vectorielle $\mathbf{X}$ de moyenne nulle et qui valent respectivement :

$$\begin{aligned} cum(X_i, X_j) &= E(X_i X_j) \\ cum(X_i, X_j, X_k) &= E(X_i X_j X_k) \\ cum(X_i, X_j, X_k, X_l) &= E(X_i X_j X_k X_l) - E(X_i X_j) E(X_k X_l) \\ &- E(X_i X_k) E(X_j X_l) - E(X_i X_l) E(X_j X_k) \end{aligned} \tag{32}$$

Une méthode de l'ACI consiste donc à minimiser les cumulants croisés d'ordre 4 pour trouver des composantes aussi indépendantes que possible. Un des algorithmes bien connu est `JADE` (Joint Approximate Diagonalization of Eigenmatrices) qui utilise une méthode tensorielle de minimisation des cumulants croisés. Cette méthode a été introduite dans [20] et son principe consiste à rechercher la matrice orthogonale $\mathbf{W}$ qui diagonalise conjointement les matrices $\mathbf{Q}_i = \mathbf{W}\mathbf{F}(\mathbf{M}_i)\mathbf{W}^T$, où $\mathbf{F}(\mathbf{M}_i)$ est une matrice propre du tenseur des cumulants d'ordre 4. Notons cependant que dans [20], il y a une première phase de blanchiment des données. L'algorithme simplifié `FOBI` [19] (Fourth Order Blind Identification) utilise également une méthode tensorielle et permet de retrouver les composantes indépendantes sous réserve que celles-ci aient des kurtosis différents, cette méthode est l'une des plus efficientes en ACI, de par sa faible complexité algorithmique. Contrairement à `FOBI`, l'algorithme `SOBI` [14] (Second Order Blind Identification) utilise des statistiques du second ordre pour une méthode de diagonalisation conjointe, et permet d'obtenir des sources indépendantes en présence du bruit.

0.4 L'ACI et la compression des images

L'ACI a été introduite récemment pour résoudre plusieurs problèmes dans le domaine du traitement des images en général, et des images satellitaires en particulier. Des travaux menés par [22] [17] [54] ont montré que la séparation aveugle de sources pouvait être utilisée pour la déconvolution, le débruitage ou l'analyse des images multicomposantes. L'ACI a été utilisée dans [46] pour la résolution du problème de classification en traitement d'images. Liu [48] a utilisé une méthode d'ACI par réseau de neurones pour proposer des techniques d'analyse d'images, pouvant servir à la classification des images couleur RVB. Un lien entre les techniques d'ACI et la compression avec pertes des signaux en général et des images en particulier a été fait dans [52]. Ce lien peut se comprendre de façon assez intuitive : l'un et l'autre recherchant à réduire autant que faire se peut la redondance entre les données. En ACI on veut réduire l'information mutuelle entre les sources, et en compression d'images multi-composantes, on aimerait réduire autant que faire se peut, la redondance spectrale qui existe entre les différentes composantes qui constituent l'image (en supposant que la redondance spatiale a été réduite d'une façon efficace). Ramakrishna et *al.* [65] ont proposé un schéma de compression d'images hyper spectrales où ils ont incorporé l'algorithme `FastICA` [38]. Outre la compression, leur étude permet de faire de la classification des images. Ils ont aussi effectué une comparaison des performances obtenues par `FastICA` avec celles de la TKL et JPEG2000.

Étonnement, les premiers liens entre la compression des images et l'ACI ont été faits à l'issue des travaux de David Hubel[4] [35] sur le fonctionnement du traitement de l'information par le cortex visuel primaire. D'autres travaux ayant un lien avec le domaine de la neurophysiologie [30] [55] ont permis de mieux comprendre la structure des fonctions de bases des images naturelles, ainsi que la perception et le traitement de ces dernières par le cortex visuel. D'autre part, il est apparu à l'issue de ces travaux de recherche de David Hubel que l'information visuelle était traitée par des neurones du cerveau en utilisant une structure proche de celle de l'ACI dans la mesure où chaque détail de l'information visuelle est traité indépendamment par une partie du cerveau.

[4] Prix Nobel de médecine en 1981.

Première partie

Étude et évaluation des performances des systèmes de compression utilisant des transformations modifiées à base d'ACI

Introduction

Cette première partie constituée de trois chapitres introduit une étude de plusieurs systèmes de compression d'images multi-composantes et présente leurs performances. Ces systèmes ont pour particularité l'utilisation de deux transformations pour réduire la redondance, l'une étant la transformée en ondelette discrète qui est utilisée dans la norme JPEG2000. Elle est appliquée à chaque composante pour réduire la redondance spatiale, puis une autre transformation permettant de réduire la redondance spectrale est appliquée. Plusieurs cas seront étudiés suivant l'ordre dans lequel sera appliquée la transformation réduisant la redondance spectrale : avant ou après la TOD.

Le premier chapitre est consacré à la description des différents schémas de compression qui seront étudiés et à l'approximation de la formule de la distorsion globale moyenne sous l'hypothèse haute résolution ou encore à faibles distorsions de ces schémas de compression. Cette formule d'approximation dans les différents cas fait apparaître des facteurs de pondération qui dépendent de la transformation (spatiale et/ou spectrale) appliquée au décodeur. L'identité reliant ces facteurs de pondération aux coefficients des transformations (spatiales et spectrales) est donnée explicitement. Bien que le résultat soit connu [76] dans le cas d'une TOD seule, nous avons choisi de donner une démonstration complète, n'en ayant pas trouvée dans la littérature.

Le deuxième chapitre étend les notions de gain de codage généralisé et de réduction maximale de débit généralisée introduites dans [52] [51] aux images multi-composantes. En se servant de l'approximation de la distorsion faite au premier chapitre, nous donnons une expression des pas de quantification optimaux pour chacune des composantes de l'image obtenue après transformation à haute résolution. Ce chapitre s'achève par un rappel sur les critères à optimiser pour obtenir les deux transformations qui ont été introduites dans [51] [52].

Le troisième et dernier chapitre de cette première partie est consacré à la présentation des résultats obtenus et à leur comparaison avec l'état de l'art. Notre objectif étant de faire de la compression, les résultats sont présentés sous forme de courbes débit vs PRSB et les gains de codage généralisé sont déduits à partir de ces courbes. Des tableaux récapitulatifs de ces résultats pour certains débits sont également donnés. Au préalable dans ce chapitre, nous faisons une vérification expérimentale des formules d'approximations de la distorsion établies au premier chapitre.

Chapitre 1

Description globale du schéma de compression étudié, approximation de la distorsion

Après avoir rappelé les notations qui seront utilisées, nous décrivons les différents schémas de compression qui seront étudiés. Une formule d'approximation de la distorsion moyenne globale est établie pour chaque cas, ainsi que des approximations des différents éléments intervenant dans cette formule. Ce premier chapitre s'achève par un calcul détaillé des coefficients de pondérations provenant de la TOD.

1.1 Structures de compression

1.1.1 Notations

Une image multicomposante $\mathbf{X}$ est constituée de N composantes $\mathbf{X}_i$ avec $1 \leq i \leq N$. Chaque composante est une image 2-D ayant N_L lignes et N_C colonnes (nous supposons que toutes les composantes ont le même nombre de pixels). Pour alléger les notations et simplifier les expressions mathématiques, les pixels d'une composante sont rangés selon un ordre de balayage prédéfini et qui est le même pour chaque composante. Ici, les pixels sont rangés les uns à la suite des autres colonne par colonne et l'on note $\mathbf{X}_i$ le vecteur ligne de dimension $L = N_L N_C$ ainsi obtenu. L'image multi-composante $\mathbf{X}$ est une matrice de dimension $N \times L$, dont la $i^{ème}$ ligne vaut $\mathbf{X}_i$. Par la suite, nous désignerons indifféremment par $i^{ème}$ composante de l'image $\mathbf{X}$, aussi bien la matrice $\mathbf{X}_i$ de dimension $N_L \times N_C$ que le vecteur ligne $\mathbf{X}_i$ de longueur L, la distinction entre ces deux objets se faisant sans ambiguïté suivant le contexte. Le terme **multi-pixel** désigne le vecteur constitué en prenant une des colonnes de $\mathbf{X}$. Dans la suite, nous serons amenés à découper les composantes de l'image en blocs de P pixels adjacents, les blocs ne se recouvrant pas, de sorte que $L = KP$, avec $K \in \mathbb{N}^*$. Nous supposerons sans perte de généralité, que L est multiple de P.

Notons vec l'opérateur qui transforme une matrice de dimension $N \times P$ en un vecteur en rangeant ses colonnes verticalement les unes au dessus des autres (la première se trouvant en haut et la dernière en bas) et ^{T} la transposée d'une matrice. L'opération inverse est réalisée par l'opérateur mat qui transforme un vecteur de taille NP en une matrice de dimension $N \times P$ en rangeant les éléments de la matrice, les uns à la suite des autres, colonne par colonne. Définissons

aussi l'opérateur per qui, appliqué à une matrice $\mathbf{B}$ de dimension $N \times L$ avec $L = KP$, commence par découper $\mathbf{B}$ en K blocs deux à deux disjoints $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_K$ de dimension $N \times P$ puis applique l'opérateur vec à chacun de ces blocs. Ainsi, la matrice per($\mathbf{B}$) est de dimension $NP \times K$.

L'opérateur inverse per^{-1} appliqué à une matrice $\mathbf{B}$ de dimension $NP \times K$ transforme chaque colonne $\mathbf{B}_k$ en une matrice de dimension $N \times P$ en utilisant l'opérateur mat($\mathbf{B}_k$). Ainsi, la matrice per$^{-1}(\mathbf{B})$ obtenue est de dimension $N \times (KP)$.

En particulier, pour un vecteur $\mathbf{x} \in \mathbb{R}^{NP}$ écrit par blocs $\mathbf{x} = (\mathbf{x}_1^T \ldots \mathbf{x}_P^T)^T$, chacun des $\mathbf{x}_u, 1 \leq u \leq P$, étant un vecteur de $\mathbb{R}^N$, nous avons la relation :

$$\mathrm{per}^{-1}(\mathbf{x}\mathbf{e}_k^{'T}) = \sum_{u=1}^{P} \mathbf{x}_u \mathbf{e}_{(k-1)P+u}^T \tag{1.1}$$

où $\mathbf{e}_u$ est le $u^{ème}$ vecteur de la base canonique de $\mathbb{R}^{KP}$ et $\mathbf{e}_k^{'}$ est le $k^{ème}$ vecteur de la base canonique de $\mathbb{R}^K$.

Dans la suite, nous décrivons plusieurs schémas de compression pour les images multi-composantes qui utilisent deux transformations, l'une pour la réduction de la redondance spatiale, et l'autre pour réduire la redondance spectrale. L'une et l'autre pouvant être utilisées dans un ordre différent selon plusieurs cas comme cela a été suggéré partiellement dans [28] et [75].

1.1.2 Schéma séparable

La description du système de compression dans ce premier cas se résume comme suit :

- La même décomposition en ondelette discrète ou TOD est appliquée à chaque composante $\mathbf{X}_i$ de l'image multi-composante. Notons $\mathbf{W}$ la matrice carrée inversible d'ordre L associée à la TOD. Appliquée à l'image entière, la transformation $\mathbf{X} \longmapsto \mathbf{X}\mathbf{W}^T$ correspond à la décomposition en ondelette de cette dernière. Un **multi-coefficient** désigne le vecteur constitué par une colonne de la matrice $\mathbf{X}\mathbf{W}^T$. La TOD a pour but ici de réduire la redondance spatiale dans chacune des différentes composantes de l'image.
- Une transformation linéaire $\mathbf{A}$ dont le but est de réduire la redondance spectrale est ensuite appliquée à l'image transformée $\mathbf{X}\mathbf{W}^T$ et nous obtenons une image $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{W}^T$.

L'expression $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{W}^T$ peut-être réécrite de deux façons différentes. Elle peut encore s'écrire comme $\mathbf{Y} = \mathbf{A}(\mathbf{X}\mathbf{W}^T)$, ce qui suppose que la TOD a été appliquée avant la transformation spectrale, ou encore comme $\mathbf{Y} = (\mathbf{A}\mathbf{X})\mathbf{W}^T$, ce qui suppose que la transformation spectrale a d'abord été appliquée sur l'image $\mathbf{X}$, et que la TOD a été appliquée sur l'image transformée $\mathbf{A}\mathbf{X}$. C'est une des raisons pour laquelle nous parlons de **schéma séparable** car la transformation spectrale et la TOD peuvent s'appliquer l'une avant ou après l'autre sur l'image multi-composante et le résultat obtenu reste le même à savoir $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{W}^T$. Dans les schémas que nous étudions la TOD est fixée, seule la transformation spectrale s'adapte aux données. Pour le schéma séparable, nous distinguons deux variantes suivant l'espace de représentation des données dans lequel est réalisée l'adaptation de la transformation. Cet espace est soit celui de l'image originale, soit celui des coefficients d'ondelette. Dans le premier cas nous parlerons de **schéma séparable avec adaptation dans l'espace image**, ou plus brièvement schéma séparable avec adaptation image, et dans le second cas nous parlerons de **schéma séparable avec adaptation (dans l'espace) ondelette**.

Si l'image transformée $\mathbf{Y}$ subit une quantification scalaire par composante, nous obtenons l'image transformée quantifiée que nous notons $\mathbf{Y}^q$. Cette image quantifiée est ensuite codée séparément par composante pour obtenir un débit moyen de codage. L'opération de codage étant une

opération totalement réversible, on retrouve l'image transformée et quantifiée $\mathbf{Y}^q$ après le décodage. A partir de cette dernière, il est possible de reconstituer une image approchée $\widehat{\mathbf{X}}$ de $\mathbf{X}$ en procédant de la façon suivante :

- La transformation spectrale inverse $\mathbf{A}^{-1}$ est d'abord appliquée à $\mathbf{Y}^q$ et nous obtenons $\mathbf{A}^{-1}\mathbf{Y}^q$
- La TODI (Transformée en Ondelette Discrète Inverse) est ensuite appliquée sur l'image transformée obtenue et nous obtenons l'image approchée $\widehat{\mathbf{X}}$ qui est donnée par $\widehat{\mathbf{X}} = \mathbf{A}^{-1}\mathbf{Y}^q\mathbf{W}^{-T}$ (l'ordre dans lequel nous appliquons les transformations inverses est en fait indifférent, le résultat reste identique à savoir $\widehat{\mathbf{X}} = \mathbf{A}^{-1}(\mathbf{Y}^q\mathbf{W}^{-T}) = (\mathbf{A}^{-1}\mathbf{Y}^q)\mathbf{W}^{-T}$).

En compression, nous nous intéressons au débit moyen de codage obtenu ainsi qu'à la distorsion qui est mesurée entre l'image originale $\mathbf{X}$ et l'image approchée $\widehat{\mathbf{X}}$. La distorsion étant introduite par l'opération de quantification. La mesure de distorsion utilisée ici est rappelons le, l'erreur quadratique moyenne.

1.1.3 Schéma en sous-bandes

Dans ce schéma, la TOD est d'abord appliquée sur chaque composante, puis une transformation spectrale par sous-bande est appliquée sur l'image dans le domaine ondelette. Ce schéma se résume comme suit :

- La même TOD est appliquée à chaque composante $\mathbf{X}_i$ de l'image multi-composante et nous obtenons l'image transformée $\mathbf{X}\mathbf{W}^T$ correspondant à la décomposition en ondelette de $\mathbf{X}$.
- Puis les coefficients d'ondelette de chaque composante sont partitionnés et regroupés par sous-bandes dans un ordre prédéfini des sous-bandes d'ondelette et qui est le même pour toutes les composantes. En supposant que la TOD produit M sous-bandes par composante, l'image obtenue est donnée par :
$$\mathbf{X}\mathbf{W}^T = [(\mathbf{X}\mathbf{W}^T)^{(1)} \dots (\mathbf{X}\mathbf{W}^T)^{(M)}].$$
- Enfin une transformation spectrale est appliquée par sous-bande et l'image finale transformée est obtenue par :
$$\mathbf{Y} = [\mathbf{A}^{(1)}(\mathbf{X}\mathbf{W}^T)^{(1)} \dots \mathbf{A}^{(M)}(\mathbf{X}\mathbf{W}^T)^{(M)}].$$

En quantifiant cette dernière image transformée par sous-bande et par composante, nous obtenons l'image quantifiée $\mathbf{Y}^q = [\mathbf{Y}^{q(1)} \dots \mathbf{Y}^{q(M)}]$. Cette dernière image nous permet de reconstruire une image approchée $\widehat{\mathbf{X}}$ en procédant comme suit :

- Partant de $[\mathbf{Y}^{q(1)} \dots \mathbf{Y}^{q(M)}]$, nous appliquons les transformations spectrales inverses $\mathbf{A}^{(m)-1}$ par sous-bande pour $1 \leq m \leq M$:
$$[\mathbf{Y}^{q(1)} \dots \mathbf{Y}^{q(M)}] \longmapsto [\mathbf{A}^{(1)-1}\mathbf{Y}^{q(1)} \dots \mathbf{A}^{(M)-1}\mathbf{Y}^{q(M)}]$$
- La TODI est enfin appliquée pour obtenir une image approchée $\widehat{\mathbf{X}}$ obtenue par :
$$\widehat{\mathbf{X}} = [\mathbf{A}^{(1)-1}\mathbf{Y}^{q(1)} \dots \mathbf{A}^{(M)-1}\mathbf{Y}^{q(M)}]\mathbf{W}^{-T}$$

La distorsion et le débit moyen de codage peuvent alors être obtenus comme dans la sous-section 1.1.2. Ce schéma est appelé schéma en sous-bandes car une transformation spectrale est appliquée par sous-bande pour réduire la redondance spectrale entre les composantes de l'image multi-composante transformée dans le domaine ondelette.

1.1.4 Schéma mixte en sous-bandes

La description du système de compression dans ce cas se résume comme suit :

- La même TOD est appliquée à chaque composante $\mathbf{X}_i$ de l'image multi composante. Nous obtenons l'image transformée dans le domaine ondelette $\mathbf{X}\mathbf{W}^T$.

- Puis les coefficients d'ondelette de chaque composante sont regroupés par sous-bandes dans un ordre prédéfini des sous-bandes d'ondelette.
$$\mathbf{XW}^T = [(\mathbf{XW}^T)^{(1)} \dots (\mathbf{XW}^T)^{(M)}].$$
Chaque sous-bande est découpée en blocs de P pixels adjacents ($P = 1, 2$ ou 4) sans recouvrement. Notons K_m le nombre de bloc de P coefficients d'ondelette se trouvant dans la $m^{ème}$ sous-bande d'une composante. Sans perdre en généralité, nous supposons que le nombre de coefficients d'ondelette de chaque sous-bande est divisible par P, ceci pour éviter d'avoir des blocs de tailles différentes dans une même sous-bande. Dans chaque bloc, les coefficients d'ondelette sont ordonnés suivant un ordre prédéfini puis, les blocs d'une même sous-bande sont balayés colonne par colonne et enfin, les différentes sous-bandes sont balayées dans un ordre prédéfini (des basses résolutions vers les plus hautes). Notons M le nombre de sous-bandes par composante. Les coefficients ainsi ordonnés associés à la $m^{ème}$ sous-bande et aux N composantes sont notés $(\mathbf{XW}^T)^{(m)}$. Une illustration de ce découpage en blocs pour une sous-bande est faite dans 1.2 où la sous-bande prise pour le découpage a été choisie de façon quelconque après la TOD qui est illustré par le schéma 1.1.

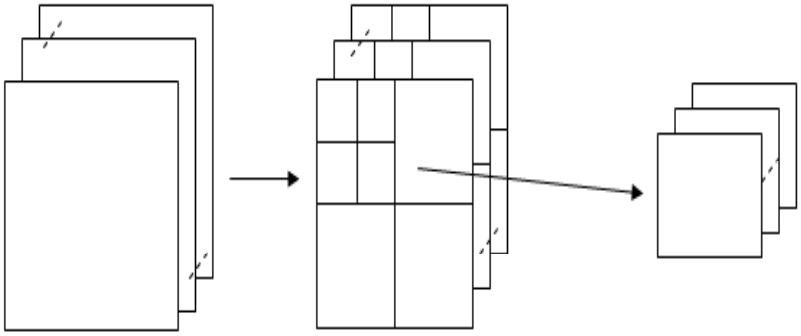


Fig. 1.1 – TOD d'une image et choix d'une sous-bande

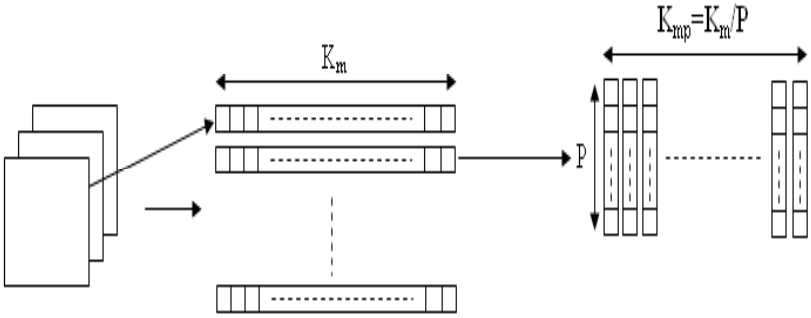


Fig. 1.2 – Découpage des sous-bandes en blocs de taille P

- L'opérateur per est ensuite appliqué à l'image multi-composante de la façon suivante :
$$[(\mathbf{XW}^T)^{(1)} \dots (\mathbf{XW}^T)^{(M)}] \longmapsto [\mathrm{per}((\mathbf{XW}^T)^{(1)}) \dots \mathrm{per}((\mathbf{XW}^T)^{(M)})].$$
Ainsi, une matrice de dimension $NP \times K$ est obtenue, chacune de ses colonnes contenant

pour chaque composante, tous les coefficients d'ondelette d'un bloc de taille P, le bloc étant situé au même endroit pour les différentes composantes.

- Enfin, nous appliquons une transformation linéaire $\mathbf{A}^{(m)}$ dépendant de la sous-bande à chaque colonne de la matrice ainsi obtenue, et dont le but est non seulement de réduire la redondance spectrale entre les différentes composantes, mais aussi de réduire la redondance entre coefficients d'ondelette voisins d'une même sous-bande quand $P > 1$. En effet, après la TOD, il subsiste encore une redondance spatiale entre les coefficients d'ondelette dans une même sous-bande, ce qui justifie ce choix.

$$[\mathrm{per}((\mathbf{X}\mathbf{W}^T)^{(1)}) \dots \mathrm{per}((\mathbf{X}\mathbf{W}^T)^{(M)})] \longmapsto [\mathbf{A}^{(1)}\,\mathrm{per}((\mathbf{X}\mathbf{W}^T)^{(1)}) \dots \mathbf{A}^{(M)}\,\mathrm{per}((\mathbf{X}\mathbf{W}^T)^{(M)})]$$

Nous supposons comme dans les sections précédentes que la TOD utilisée ainsi que les transformations $\mathbf{A}^{(m)}$ sont totalement inversibles. Notons $\mathbf{Y}$ l'image multi composante ainsi obtenue et $\mathcal{A}$ la transformation globale qui transforme l'image $\mathbf{X}$ en $\mathbf{Y}$ avec $\mathrm{vec}(\mathbf{Y}^T) = \mathcal{A}\,\mathrm{vec}(\mathbf{X}^T)$.

Toutes les suppositions et hypothèses faites jusqu'ici dans la description du schéma mixte en sous-bandes sont mineures et non contraignantes.

Si chaque coefficient transformé subit une quantification scalaire[1], nous obtenons la matrice $\mathbf{Y}^q$. A partir de cette dernière, il est possible de reconstituer une image approchée $\widehat{\mathbf{X}}$ en procédant de la façon suivante :

- Partant de $[\mathbf{Y}^{q(1)} \dots \mathbf{Y}^{q(M)}]$, nous appliquons les transformations inverses $\mathbf{A}^{(m)-1}$ pour $1 \leq m \leq M$:
$$[\mathbf{Y}^{q(1)} \dots \mathbf{Y}^{q(M)}] \longmapsto [\mathbf{A}^{(1)-1}\mathbf{Y}^{q(1)} \dots \mathbf{A}^{(M)-1}\mathbf{Y}^{q(M)}]$$
Il est à noter qu'au préalable, l'image $\mathbf{Y}^q$ est partitionnée en regroupant les coefficients d'ondelette qui ont subi une transformation linéaire et une quantification par sous-bande et par bloc de P coefficients.
- Reconstruire les N composantes à partir des NP composantes obtenues en appliquant l'opérateur per^{-1} :
$$[\mathbf{A}^{(1)-1}\mathbf{Y}^{q(1)} \dots \mathbf{A}^{(M)-1}\mathbf{Y}^{q(M)}] \longmapsto [\mathrm{per}^{-1}(\mathbf{A}^{(1)-1}\mathbf{Y}^{q(1)}) \dots \mathrm{per}^{-1}(\mathbf{A}^{(M)-1}\mathbf{Y}^{q(M)})]$$
- Enfin, appliquer la transformation en ondelette discrète inverse (TODI) à chaque composante issue de l'opération précédente pour reconstituer une image $\widehat{\mathbf{X}}$:
$$\widehat{\mathbf{X}} = [\mathrm{per}^{-1}(\mathbf{A}^{(1)-1}\mathbf{Y}^{q(1)}) \dots \mathrm{per}^{-1}(\mathbf{A}^{(M)-1}\mathbf{Y}^{q(M)})]\mathbf{W}^{-T}$$

1.1.5 Intérêt de ces schémas de compression

Le schéma séparable avec adaptation dans l'espace image est celui appliqué dans la norme JPEG pour le codage des images couleurs, et aussi dans la norme JPEG2000, non seulement pour le codage des images couleur, mais aussi pour le codage des images multi-composantes. Son étude ici est nécessaire non seulement pour faire une comparaison avec JPEG2000, mais aussi pour pouvoir comparer les performances des transformations spectrales autre que la TKL qui est préconisée pour cette norme. Le schéma séparable avec adaptation dans l'espace ondelette requiert aussi une étude. En effet, étant donné le premier schéma cité dans cette sous-section, il est nécessaire de comparer et d'étudier, le bénéfice que l'on aurait à appliquer la transformation spectrale adaptative avant ou après la TOD.

Pour le schéma séparable avec adaptation dans l'espace ondelette, la même transformation spectrale est appliquée sur toutes les sous-bandes d'une composante. Ce qui sans doute est loin

[1] A chaque élément d'un bloc de taille P est associé un quantificateur qui dépend de la composante, de la position dans le bloc et de la sous-bande, mais pas du bloc.

d'être optimal dans la mesure où l'on sait que les différentes sous-bandes après la TOD dans une composante présente des caractéristiques différentes aussi bien en termes de détails y apparaissant, qu'en termes de niveau d'énergie correspondant à chaque sous-bande. Ce qui nous a conduit au schéma en sous-bandes, qui préconise d'appliquer une transformation spectrale par sous-bande, ce qui revient à rechercher M transformations spectrales.

Il est bien connu qu'après la TOD (ici nous utilisons la 9/7 de Daubechies), il demeure une redondance spatiale entre coefficients d'ondelette, aussi bien dans une sous-bande, qu'entre les différentes sous-bandes. Cette redondance spatiale remanente a été quantifiée pour certaines images dans [49] et, la méthode utilisée pour évaluer cette redondance y est clairement définie. Le schéma mixte en sous-bandes propose d'appliquer une transformation spectrale par sous-bande (après découpage en blocs de taille P sur chaque sous-bande), non seulement pour réduire la redondance spectrale entre les différentes sous-bandes des composantes, mais aussi pour réduire la redondance spatiale qui demeure entre les coefficients d'ondelette dans une sous-bande. Comme nous pouvons le constater, la démarche qui nous a permis de proposer ces différents schémas de compression est simple et suit une certaine logique. En effet, pour faire une comparaison avec l'état de l'art qui aujourd'hui est JPEG2000, il nous faut un schéma de compression qui lui est similaire, partant de là, apparaît naturellement le schéma séparable avec adaptation dans l'espace ondelette. Le schéma en sous-bande provient des carences et des limites soulignées plus hauts de ce dernier schéma séparable. Le dernier schéma étudié, qui est le schéma mixte en sous-bandes est une généralisation du schéma en sous-bandes et permet d'aller plus loin que ce dernier dans la minimisation des redondances spatiales et spectrales dans une image multi-composante.

1.2 Formule d'approximation de la distorsion

Nous nous intéresserons par la suite à la distorsion entre $\widehat{\mathbf{X}}$ et $\mathbf{X}$ introduite par la quantification pour le schéma de compression mixte en sous-bandes. La mesure de distorsion utilisée ici, rappelons-le, est l'erreur quadratique moyenne. Ce schéma étant un peu plus général, les formules d'approximation des autres schémas seront déduits de celle de ce schéma.

Soit D l'erreur quadratique moyenne entre $\widehat{\mathbf{X}}$ et $\mathbf{X}$. Il est bien connu que $D = (1/NL)\sum_{i=1}^{N}\sum_{\ell=1}^{L}(X_i(\ell) - \widehat{X}_i(\ell))^2$ est une très bonne estimation de la distorsion définie par $D = E[\|\mathbf{X} - \widehat{\mathbf{X}}\|^2]/N$, où $\mathbf{X}$ et $\widehat{\mathbf{X}}$ correspondent à un multi-pixel de l'image originale et de l'image reconstruite (localisé au même endroit) en supposant que les multi-pixels $\mathbf{X}(\ell)$ $(1 \leq \ell \leq L)$ (correspondant aux colonnes de $\mathbf{X}$) sont des vecteurs aléatoires identiquement distribués (idem pour l'image reconstruite). L'opérateur $E(.)$ désignant l'espérance mathématique et $\|\mathbf{x}\|$ est la norme euclidienne du vecteur $\mathbf{x} \in \mathbb{R}^N$. Nous cherchons à évaluer la distorsion D du schéma de compression défini ci-dessus. Avant cela, commençons par évaluer la distorsion dans un schéma général.

1.2.1 Cas général

Soit $\mathbf{X}$ un vecteur aléatoire réel constitué de N composantes auquel on fait subir une transformation linéaire : $\mathbf{X} \longmapsto \mathcal{A}\mathbf{X}$ où $\mathcal{A}$ est une matrice carrée inversible d'ordre N. Le signal obtenu après la transformation est ensuite quantifié. Notons $q(\mathcal{A}\mathbf{X})$ le signal quantifié ainsi obtenu. Nous appliquons ici une **quantification scalaire uniforme** sur chaque composante du signal obtenu. Notons $\widehat{\mathbf{X}} = \mathcal{A}^{-1}q(\mathcal{A}\mathbf{X})$ le signal reconstitué et qui a subi une distorsion par rapport à $\mathbf{X}$, distorsion induite par la quantification. L'erreur de reconstruction entre $\mathbf{X}$ et $\widehat{\mathbf{X}}$ est donnée par :

$$\mathbf{X} - \widehat{\mathbf{X}} = \mathcal{A}^{-1}[\mathcal{A}\mathbf{X} - q(\mathcal{A}\mathbf{X})] = \mathcal{A}^{-1}\mathbf{b}.$$

où $\mathbf{b}$ est le bruit de quantification du signal transformé. La distorsion D entre $\mathbf{X}$ et $\widehat{\mathbf{X}}$ se calcule alors comme suit :

$$D = \tfrac{1}{N} E\|\mathcal{A}^{-1}\mathbf{b}\|^2 = \tfrac{1}{N} E[\mathbf{b}^T \mathcal{A}^{-T} \mathcal{A}^{-1}\mathbf{b}]$$

avec $\mathcal{A}^{-T} = (\mathcal{A}^{-1})^T$. D'après cette dernière expression, on voit que si $\mathcal{A}$ est orthonormale, la distorsion est réduite à $\frac{1}{N}E\|\mathbf{b}\|^2$ c'est-à-dire que la distorsion entre $\mathbf{X}$ et $\widehat{\mathbf{X}}$ est la même que celle entre $\mathbf{Y}$ et $q(\mathbf{Y})$ en ayant posé $\mathbf{Y} = \mathbf{AX}$. Dans le cas général, la distorsion peut encore s'écrire comme suit :

$$D = \tfrac{1}{N} \operatorname{tr}\{E(\mathbf{b}\mathbf{b}^T)\mathcal{A}^{-T}\mathcal{A}^{-1}\}$$

Par la suite on fait et on admet l'hypothèse (qui sera justifiée plus tard) suivante :

$H1$:les composantes du bruit de quantification $\boldsymbol{b}$ sont décorrélées et de moyennes nulles.

Proposition 1.2.1 *Sous l'hypothèse H_1 la distorsion s'écrit :*

$$D = \frac{1}{N}\sum_{i=1}^{N} w_i D_i \tag{1.2}$$

où $D_i = E(b_i^2)$ est la distorsion induite par la quantification portant sur $\boldsymbol{Y}_i$, la $i^{\text{ème}}$ composante de $\boldsymbol{Y}$ et

$$w_i = \textstyle\sum_{j=1}^{N} (\mathcal{A}^{-1})_{ji}^2 = \|\mathcal{A}^{-1}\boldsymbol{e}_i\|^2$$

$(\mathcal{A}^{-1})_{ji}$ désignant l'élément d'indices (j,i) de la matrice $\mathcal{A}^{-1}$ et $\boldsymbol{e}_i$ étant le $i^{\text{ème}}$ vecteur de base de $\mathbb{R}^N$.

Cette formule reste valable sans l'hypothèse $H1$ si la matrice $\mathcal{A}^{-T}\mathcal{A}^{-1}$ est diagonale (en particulier si $\mathcal{A}$ est orthogonale).

Preuve : Sous les hypothèses faites, une des deux matrices $\mathcal{A}^{-T}\mathcal{A}^{-1}$ ou $E(\mathbf{b}\mathbf{b}^T)$ est diagonale. Par suite, la trace de leur produit se réduit à la somme des produits de leurs termes diagonaux. ■

Un cas particulier est lorsque la matrice $\mathcal{A}$ est orthogonale, dans ce cas, la distorsion sur $\mathbf{X}$ est la même que celle sur $\mathbf{Y}$ car $w_i = 1$, $\forall\ 1 \leq i \leq N$.

L'hypothèse $(H1)$ trouve sa justification sous les conditions suivantes :

- C_1 : Le vecteur aléatoire $\mathbf{Y}$ admet une densité de probabilité continue qui ne subit pas de variations trop brusques.
- C_2 : La quantification est faite à haute résolution (faible distorsion) c'est-à-dire que les pas de quantification sont suffisamment petits.

En effet, soit $f_{\mathbf{Y}}$ la densité de probabilité de $\mathbf{Y}$, la loi conditionnelle de $\mathbf{b}$ sachant q($\mathbf{Y}$) admet la densité :

$$f_{\mathbf{b}|q(\mathbf{Y})}(\mathbf{u}) = \begin{cases} f_{\mathbf{Y}}(q(\mathbf{Y}) + \mathbf{u})/C & \text{si } \mathbf{u} \in \prod_{i=1}^{N}[-h_i/2, h_i/2] \\ 0 & \text{sinon} \end{cases}$$

C ici représente une constante de normalisation de telle sorte que $f_{\mathbf{b}|q(\mathbf{Y})}$ soit une densité de probabilité et h_i est le pas de quantification du quantificateur de $\mathbf{Y}_i$. En admettant que les deux conditions précédentes sont vraies, il est alors possible de considérer la densité f_{Y_i} comme constante dans l'hyper-cube $\prod_{i=1}^{N}[q_i(Y_i) - h_i/2, q_i(Y_i) + h_i/2]$ et dans ce cas la densité conditionnelle précédente peut-être approchée par :

$$f_{\mathbf{b}|q(\mathbf{Y})}(\mathbf{u}) = \begin{cases} 1/\prod_{i=1}^{N} h_i & \text{si } \mathbf{u} \in \prod_{i=1}^{N}[-h_i/2, h_i/2] \\ 0 & \text{sinon} \end{cases}$$

En utilisant cette approximation, la densité conditionnelle de $\mathbf{b}$ sachant $q(\mathbf{Y})$ est indépendante de $q(\mathbf{Y})$, de plus, elle n'est rien d'autre que la loi uniforme définie dans l'hyper-cube $\prod_{i=1}^{N}[-h_i/2, h_i/2]$, ce qui justifie le fait que la moyenne des composantes de $\mathbf{b}$ soit nulle car la loi uniforme dans l'intervalle $[-h_i/2, h_i/2]$ est de moyenne nulle. Ainsi le bruit de quantification $\mathbf{b}$ est indépendant de $\mathbf{Y}$ et les composantes du bruit de quantification sont indépendantes car leur densité jointe est le produit des densités marginales. Ce qui justifie l'hypothèse $H1$.

1.2.2 Approximation de la distorsion du schéma mixte en sous-bandes

Pour appliquer la formule (1.2) à la structure de compression qui a été décrite dans la section 1.1.4, il convient de noter que l'espace des matrices $NP \times K$ a pour base la famille des matrices $\mathbf{e}_{i,k} = \mathbf{e}_i\mathbf{e}_k^T$, $1 \leq i \leq NP$, $1 \leq k \leq K$. Ici, $\mathbf{e}_i$ est le $i^{ème}$ vecteur de base de $\mathbb{R}^{NP}$ et $\mathbf{e}_k$ est le $k^{ème}$ vecteur de base de $\mathbb{R}^K$. Si $k(1 \leq k \leq K_m)$ désigne un indice de colonne pour la sous-bande $m(1 \leq m \leq M)$ alors en utilisant l'équation (1.1) nous avons :

$$\begin{aligned} \mathcal{A}^{-1}\operatorname{vec}(\mathbf{e}_{i,k}) &= \operatorname{vec}[\operatorname{per}^{-1}(\mathbf{A}^{(m)-1}\mathbf{e}_i\mathbf{e}_k^T)(\mathbf{W}^{-T})] \\ &= \operatorname{vec}[\sum_{u=1}^{P}(\mathbf{A}^{(m)-1}\mathbf{e}_i)_u(\mathbf{W}^{-1}\mathbf{e}_{(k-1)P+u})^T] \end{aligned}$$

$\mathbf{A}^{(m)-1}\mathbf{e}_i$ étant un vecteur de taille NP, on le décompose en P blocs (chaque bloc ici étant un vecteur) disjoints de taille N et on appelle $(\mathbf{A}^{(m)-1}\mathbf{e}_i)_u$ le $u^{ème}$ bloc ainsi obtenu et qui est un vecteur de taille N. Ainsi, le calcul de la norme de cet élément est

$$\begin{aligned} ||\mathcal{A}^{-1}\operatorname{vec}(\mathbf{e}_{i,k})||^2 &= \operatorname{tr}\{[\sum_{u=1}^{P}(\mathbf{A}^{(m)-1}\mathbf{e}_i)_u(\mathbf{W}^{-1}\mathbf{e}_{(k-1)P+u})^T][\sum_{u=1}^{P}(\mathbf{A}^{(m)-1}\mathbf{e}_i)_u(\mathbf{W}^{-1}\mathbf{e}_{(k-1)P+u}^T)^T]^T\} \\ &= \sum_{u,v=1}^{P}(\mathbf{W}^{-1}\mathbf{e}_{(k-1)P+u})^T(\mathbf{W}^{-1}\mathbf{e}_{(k-1)P+v})\operatorname{tr}\{(\mathbf{A}^{(m)-1}\mathbf{e}_i)_v(\mathbf{A}^{(m)-1}\mathbf{e}_i)_u^T\} \\ &= \sum_{u,v=1}^{P}\mathbf{e}_{(k-1)P+u}^T(\mathbf{W}^{-T}\mathbf{W}^{-1})\mathbf{e}_{(k-1)P+v}(\mathbf{A}^{(m)-1}\mathbf{e}_i)_v^T(\mathbf{A}^{(m)-1}\mathbf{e}_i)_u \end{aligned} \tag{1.3}$$

Cette expression est assez complexe mais se simplifie sous certaines hypothèses particulières comme la suivante :

$H2$: *La matrice* $\boldsymbol{W}^{-T}\boldsymbol{W}^{-1}$ *est diagonale*, ce qui est le cas pour une transformation en ondelette orthogonale par exemple, l'expression précédente se simplifie et devient :

$$||\mathcal{A}^{-1}\operatorname{vec}(\mathbf{e}_{i,k})||^2 = \sum_{u=1}^{P}||\mathbf{W}^{-1}\mathbf{e}_{(k-1)P+u}||^2 \times ||(\mathbf{A}^{(m)-1}\mathbf{e}_i)_u||^2 \tag{1.4}$$

Dans le cas (satisfait en général pour les ondelettes dyadiques) où $||\mathbf{W}^{-1}\mathbf{e}_{(k-1)P+u}||^2$ ne dépend pas de u mais uniquement de la sous-bande m, cette dernière expression devient :

$$||\mathcal{A}^{-1}\,\mathrm{vec}(\mathbf{e}_{i,k})||^2 = ||\mathbf{W}^{-1}\mathbf{e}_{kP}||^2 \times ||\mathbf{A}^{(m)-1}\mathbf{e}_i||^2 \tag{1.5}$$

Compte tenu de ces simplifications, si nous appliquons la formule de la distorsion (1.2) du cas général en posant $i = (\ell - 1)N + j$ avec $(1 \leq \ell \leq P)$ et $(1 \leq j \leq N)$, la distorsion devient sous les hypothèses $H1$ et $H2$

$$D = \frac{1}{NPK}\sum_{j=1}^{N}\sum_{\ell=1}^{P}\sum_{m=1}^{M}\sum_{k}||\mathbf{W}^{-1}\mathbf{e}_{kP}||^2||\mathbf{A}^{(m)-1}\mathbf{e}_{(\ell-1)N+j}||^2 D_{j,\ell,k}$$

Dans la sommation, k parcourt toutes les colonnes de $\mathbf{X} - \widehat{\mathbf{X}}$ associées à la sous-bande m et $D_{j,\ell,k}$ correspond à la distorsion du quantificateur appliqué aux $\ell^{ème}$ éléments des blocs de taille P de la $m^{ème}$ sous-bande de la $j^{ème}$ composante. Si nous appliquons une quantification scalaire uniforme à haute résolution [31] [27] pour chaque élément des blocs de chaque sous-bande et de chaque composante, il devient alors possible de supposer et ce de façon réaliste que dans chacun de ces blocs, la distorsion ne dépend pas de k mais uniquement de la position dans le bloc de la sous-bande. Par conséquent une approximation de la distorsion entre l'image de départ $\mathbf{X}$ et l'image reconstruite $\widehat{\mathbf{X}}$ est :

$$D = \frac{1}{NP}\sum_{j=1}^{N}\sum_{\ell=1}^{P}\sum_{m=1}^{M}\pi_m w_m w_{j,\ell}^{(m)} D_{j,\ell}^{(m)} \tag{1.6}$$

avec $\pi_m = K_m/L$, la proportion de la sous-bande m dans l'image, $w_m = (1/K_m)\sum_k ||\mathbf{W}^{-1}\mathbf{e}_{kP}||^2$, où ici encore k parcourt toutes les colonnes associées à la sous-bande m, $w_{j,\ell}^{(m)} = ||\mathbf{A}^{(m)-1}\mathbf{e}_{(\ell-1)N+j}||^2$ et $D_{j,\ell}^{(m)}$ est la distorsion moyenne du $\ell^{ème}$ bloc $(1 \leq \ell \leq P)$ de la sous-bande m de la $j^{ème}$ composante.

Les coefficients de pondérations w_m dépendent de la TOD utilisée et plus précisément des filtres de synthèse utilisés pour la TODI. Les coefficients de pondérations $w_{j,\ell}^{(m)}$ dépendent des transformations spectrales appliquées pour la réduction non seulement de la redondance spectrale, mais aussi de la redondance spatiale résiduelle entre les coefficients d'ondelette dans une même sous-bande.

Dans le cas particulier où les transformations $\mathbf{A}^{(m)}$ sont toutes orthogonales, l'expression (1.6) se simplifie davantage et devient alors :

$$D = \frac{1}{NP}\sum_{j=1}^{N}\sum_{\ell=1}^{P}\sum_{m=1}^{M}\pi_m w_m D_{j,\ell}^{(m)}$$

En effet, dans ce cas, les coefficients $w_{j,\ell}^{(m)}$ provenant des transformations spectrales sont tous égaux à 1.

L'hypothèse $H2$ est forte, seule l'ondelette de Haar est orthogonale, symétrique et à support compact. En compression d'images avec pertes, l'ondelette 9/7 de Daubechies donne de très bonnes performances [11] et elle est quasi orthogonale, c'est-à-dire que pour cette TOD, l'hypothèse $H2$ est presque vérifiée : $\mathbf{W}^{-T}\mathbf{W}^{-1}$ est quasi diagonale (c'est-à-dire que les éléments hors de la diagonale principale sont négligeables devant les termes diagonaux). Dans nos tests et simulations, nous utiliserons l'ondelette 9/7 de Daubechies.

Dans ce qui précède, nous nous sommes intéressés à l'approximation de la distorsion du schéma mixte en sous-bandes, par la suite, les expressions obtenues pour les schémas en sous-bandes et séparables sont explicitées.

1.2.3 Approximation de la distorsion du schéma en sous-bandes

Le schéma décrit dans la section 1.1.3 correspond au cas particulier du schéma mixte en sous-bandes avec $P = 1$. Dans ce cas particulier, l'hypothèse $H1$ suffit en se servant des calculs assez simples, pour établir une formule d'approximation de distorsion du schéma en sous-bandes semblable à celle de la relation (1.6). C'est l'objet de ce paragraphe. La relation (1.3) pour le schéma en sous-bandes s'écrit

$$\begin{aligned} \|\mathcal{A}^{-1}vec(\mathbf{e}_{ik})\|^2 &= \mathrm{tr}\{\mathbf{A}^{(m)-1}\mathbf{e}_i(\mathbf{W}^{-1}\mathbf{e}_k)^T(\mathbf{W}^{-1}\mathbf{e}_k)(\mathbf{A}^{(m)-1}\mathbf{e}_i)^T\} \\ &= \|\mathbf{A}^{(m)-1}\mathbf{e}_i\|^2\|\mathbf{W}^{-1}\mathbf{e}_k\|^2 \end{aligned}$$

Sous l'hypothèse $H1$, nous pouvons appliquer la relation (1.2) qui donne ici

$$D = \frac{1}{NL}\sum_{i=1}^{N}\sum_{m=1}^{M}\|\mathbf{A}^{(m)-1}\mathbf{e}_i\|^2\sum_{k}\|\mathbf{W}^{-1}\mathbf{e}_k\|^2 D_{ik}$$

où ici encore k parcourt toutes les colonnes de $\mathbf{X} - \widehat{\mathbf{X}}$ associées à la sous-bande m. En général $\|\mathbf{W}^{-1}\mathbf{e}_k\|^2$ ne dépend pas de k, mais seulement de m sauf peut-être pour les indices correspondant à un pixel au bord de l'image. Il est alors possible de poser $w_m = \|\mathbf{W}^{-1}\mathbf{e}_k\|^2$ et la formule de la distorsion précédente se réduit à :

$$D = \frac{1}{N}\sum_{m=1}^{M}\pi_m w_m\left[\sum_{i=1}^{N} w_i^{(m)} D_i^{(m)}\right] \tag{1.7}$$

où $\pi_m = K_m/L$ est la proportion de la sous-bande m dans une composante, $w_i^{(m)} = \|\mathbf{A}^{(m)-1}\mathbf{e}_i\|^2$, $D_i^{(m)} = (1/K_m)\sum_k D_{ik}$. En pratique, une quantification scalaire uniforme à haute résolution est effectuée et les coefficients d'une même sous-bande et d'une même composante sont quantifiés de la même façon. Il devient alors intéressant de souligner le fait que dans ce cas la distorsion $D_i^{(m)}$ dépende uniquement du pas de quantification associé.

Le terme entre crochets de (1.7) peut-être interprété comme la distorsion de la sous-bande m de l'image multi-composante. La distorsion globale est donc une combinaison linéaire des distorsions dans les sous-bandes de chaque composante et comme une quantification scalaire uniforme à haute résolution est faite, la distorsion de chaque sous-bande dans une composante ne dépend que du pas de quantification appliqué à cette dernière.

Discussion : Les formules des équations donnant une estimation de la distorsion et notamment la formule (1.7) reposent uniquement sur l'hypothèse $H1$ qui stipule que les composantes du bruit de quantification $\mathbf{b}$ sont décorrélées et de moyennes nulles. Dans ce qui précède, nul n'est besoin que les composantes de $\mathbf{Y}$ soient indépendantes pour que cette hypothèse soit vérifiée. Toutefois, il ne faudrait pas que ces composantes soient trop dépendantes. En effet, si deux composantes sont fortement dépendantes, leur densité conjointe serait fortement concentrée sur une courbe et donc varierait trop vite. Dans le cas où les composantes de $\mathbf{Y}$ proviennent d'un algorithme à base d'ACI

ou de la TKL, il est réaliste de les supposer très peu dépendantes et, dans ce cas, ces formules restent valides. Dans le cas ou la transformation globale $\mathcal{A}$ est orthogonale, l' hypothèse $H1$ n'est plus nécessaire pour obtenir la formule d'approximation de l'équation (1.7). Enfin, si les composantes de $\mathbf{Y}$ sont indépendantes, alors l'hypothèse $H1$ est vérifiée même quand la quantification ne se fait plus à haute résolution.

1.2.4 Approximation de la distorsion du schéma séparable

Sous l'hypothèse $H1$ la relation (1.3) dans ce cas devient :

$$\begin{aligned} \|\mathcal{A}^{-1}vec(\mathbf{e}_{ik})\|^2 &= \mathrm{tr}\{\mathbf{A}^{-1}\mathbf{e}_i(\mathbf{W}^{-1}\mathbf{e}_k)^T(\mathbf{W}^{-1}\mathbf{e}_k)(\mathbf{A}^{-1}\mathbf{e}_i)^T\} \\ &= \|\mathbf{A}^{-1}\mathbf{e}_i\|^2\|\mathbf{W}^{-1}\mathbf{e}_k\|^2 \end{aligned}$$

En se servant des résultats obtenus dans les sections précédentes, nous déduisons que la distorsion dans ce cas est donnée par l'équation

$$D = \frac{1}{N}\sum_{m=1}^{M}\pi_m w_m[\sum_{i=1}^{N} w_i D_i^{(m)}] \tag{1.8}$$

où $w_i = \|\mathbf{A}^{-1}\mathbf{e}_i\|^2$ représente le coefficient de pondération qui dépend uniquement de la transformation spectrale. Ce coefficient reste le même pour toutes les sous-bandes d'une composante.

1.3 Calcul des coefficients de pondération w_m

Dans la suite nous explicitons le calcul des coefficients de pondération w_m intervenant dans les formules (1.6), (1.7), (1.8), et qui dépendent uniquement du filtre de synthèse de la TODI. Comme nous l'avons mentionné, le résultat de ces calculs est bien connu [76], mais n'ayant pas trouvé de démonstration explicite dans la littérature, nous avons choisi de la présenter dans ce rapport. Dans ce qui suit, les effets de bord pour une image après la TOD sont négligés, leur influence n'étant globalement pas d'un apport important. En se servant de l'équation (1.8), nous déduisons que si la TOD est appliquée et que la transformation spectrale $\mathbf{A} = I$, dans ce cas, la formule d'approximation de la distorsion est réduite à :

$$D = \frac{1}{N}\sum_{i=1}^{N}[\sum_{m=1}^{M}\pi_m w_m D_i^{(m)}] \tag{1.9}$$

L'expression entre crochets représente la distorsion de la $i^{ème}$ composante de l'image qui n'a subit qu'une transformation spatiale, la TOD. Par la suite, nous allons établir cette formule d'approximation de la distorsion pour une composante qui a subi une TOD et nous ferons l'analogie avec cette expression entre crochets pour déduire les valeurs des coefficients de pondération w_m.

Le signal de départ qui est une image est représenté par la suite numérique $\{x_{\ell,k}\}$. La décomposition en ondelette sur un niveau s'effectue via les filtres passe-bas h^0 de réponse impulsionnelle $(h_n^0)_n$, passe-haut h^1 de réponse impulsionnelle $(h_n^1)_n$, suivi d'un sous-échantillonnage $\downarrow 2$ que l'on applique d'abord sur les lignes et qui permet d'obtenir les relations de récurrence :

$$\begin{aligned} C_{\ell,k}^{j,j-1} &= \sum_n h_n^0 C_{\ell,2k-n}^{j-1,j-1} = \sum_n h_{2k-n}^0 C_{\ell,n}^{j-1,j-1}, \\ D_{\ell,k}^{j,j-1} &= \sum_n h_n^1 D_{\ell,2k-n}^{j-1,j-1} = \sum_n h_{2k-n}^1 D_{\ell,n}^{j-1,j-1}. \end{aligned}$$

Les filtres h^0 et h^1 suivis d'un sous-échantillonnage sur les colonnes sont ensuite appliqués aux suites numériques $\{C_{\ell,k}^{j,j-1}\}$ et $\{D_{\ell,k}^{j,j-1}\}$, ce qui permet d'obtenir les valeurs des coefficients d'ondelette à l'issue d'un niveau de décomposition et dont les expressions sont données par :

$$\begin{aligned}
C_{\ell,k}^{j-1,j-1} &= \sum_n h_n^0 C_{\ell,2k-n}^{j,j-1} = \sum_n h_{2k-n}^0 C_{\ell,n}^{j,j-1}, \\
V_{\ell,k}^{j-1,j-1} &= \sum_n h_n^1 C_{\ell,2k-n}^{j,j-1} = \sum_n h_{2k-n}^1 D_{\ell,n}^{j,j-1}, \\
S_{\ell,k}^{j-1,j-1} &= \sum_n h_n^0 D_{\ell,2k-n}^{j,j-1} = \sum_n h_{2k-n}^0 D_{\ell,n}^{j,j-1}, \\
D_{\ell,k}^{j-1,j-1} &= \sum_n h_n^1 D_{\ell,2k-n}^{j,j-1} = \sum_n h_{2k-n}^1 D_{\ell,n}^{j,j-1}.
\end{aligned}$$

En ayant initialisé par $C_{\ell,k}^{0,0} = x_{\ell,k}$, ce schéma de décomposition s'effectue de façon recursive. Si les filtres d'analyse h^0 et h^1 sont choisis de façon à ce que la reconstruction parfaite soit possible, alors le signal $\{x_{\ell,k}\}$ peut être reconstitué à partir des coefficients d'échelle $C_{\ell,k}^{J,J}$ et des coefficients d'ondelette ou détails $V_{\ell,k}^{j,j}$, $S_{\ell,k}^{j,j}$, $D_{\ell,k}^{j,j}$ $(1 \leq j \leq J)$ horizontaux, verticaux et diagonaux respectivement. Pour cela, nous effectuons un sur-échantillonnage $\uparrow 2$ sur les derniers coefficients d'ondelette obtenus, puis les filtres de synthèse passe-bas g^0 de réponse impulsionnelle $(g_n^0)_n$ et passe-haut g^1 de réponse impulsionnelle $(g_n^1)_n$ sont appliqués et ce de façon recursive. Le signal est alors reconstitué à chaque niveau comme suit :

$$\begin{aligned}
C_{\ell,k}^{j,j+1} &= \sum_n g_n^0 C_{l-2n,k}^{j+1,j+1} + \sum_n g_n^1 V_{l-2n,k}^{j+1,j+1} \\
D_{\ell,k}^{j,j+1} &= \sum_n g_n^0 S_{l-2n,k}^{j+1,j+1} + \sum_n g_n^1 D_{l-2n,k}^{j+1,j+1} \\
C_{\ell,k}^{j,j} &= \sum_n g_n^0 C_{\ell,k-2n}^{j,j+1} + \sum_n g_n^1 D_{\ell,k-2n}^{j,j+1}
\end{aligned} \tag{1.10}$$

Étant donné que la reconstruction du signal image $\{x_{\ell,k}\}$ à partir des coefficients d'échelle et de détails se fait d'une façon linéaire, si la décomposition en ondelette a été appliquée sur $J \in \mathbb{N}^*$ niveaux, l'image reconstituée peut s'exprimer de la façon suivante en fonction de ces coefficients :

$$\begin{aligned}
C_{\ell,k}^{0,0} = &\sum_n \sum_m \alpha_{\ell,k,m,n}^{J,J} C_{m,n}^{J,J} + \sum_{j=1}^{J} \sum_n \sum_m \beta_{\ell,k,m,n}^{j,j} V_{m,n}^{j,j} + \\
&\sum_{j=1}^{J} \sum_n \sum_m \gamma_{\ell,k,m,n}^{j,j} S_{m,n}^{j,j} + \sum_{j=1}^{J} \sum_n \sum_m \lambda_{\ell,k,m,n}^{j,j} D_{m,n}^{j,j}
\end{aligned} \tag{1.11}$$

Proposition 1.3.1 *Les termes $\alpha_{\ell,k,m,n}^{J,J}$, $\beta_{\ell,k,m,n}^{J,J}$, $\gamma_{\ell,k,m,n}^{J,J}$, $\lambda_{\ell,k,m,n}^{J,J}$ peuvent se mettre sous la forme :*

$$\begin{aligned}
\alpha_{\ell,k,m,n}^{J,J} &= \phi^J(\ell - 2^J m)\phi^J(k - 2^J n) \\
\beta_{\ell,k,m,n}^{j,j} &= \phi^j(\ell - 2^j m)\psi^j(k - 2^j n) \\
\gamma_{\ell,k,m,n}^{j,j} &= \psi^j(\ell - 2^j m)\phi^j(k - 2^j n) \\
\lambda_{\ell,k,m,n}^{j,j} &= \psi^j(\ell - 2^j m)\psi^j(k - 2^j n)
\end{aligned} \tag{1.12}$$

où les fonctions réelles ϕ^j et ψ^j sont calculées par les relations de récurrence :

$$\phi^{j+1}(k) = \sum_{\ell} \phi^j(k - 2^j\ell) g^0_\ell, \qquad \psi^{j+1}(k) = \sum_{\ell} \phi^j(k - 2^j\ell) g^1_\ell$$

et en initialisant $\phi^1(k) = g^0_k$ et $\psi^1(k) = g^1_k$.

Preuve : Soit $B = \sum_m \sum_n \phi^J(\ell - 2^J m)\phi^J(k - 2^J n) C^{J,J}_{m,n}$, le premier terme du membre de droite de l'équation (1.11) nous allons développer cette expression compte tenu de (1.10) en fonction des sous-bandes du niveau $J+1$. Commençons par remarquer que $C^{J,J}_{m,n}$ peut encore s'écrire :

$$\begin{aligned} C^{J,J}_{m,n} &= \sum_p g^0_{n-2p} C^{J,J+1}_{m,p} + \sum_p g^1_{n-2p} D^{J,J+1}_{m,p} \\ &= \sum_p g^0_{n-2p} (\sum_q g^0_{m-2q} C^{J+1,J+1}_{q,p} + \sum_q g^1_{m-2q} V^{J+1,J+1}_{q,p}) \\ &+ \sum_p g^1_{n-2p} (\sum_q g^0_{m-2q} S^{J+1,J+1}_{q,p} + \sum_q g^1_{m-2q} D^{J+1,J+1}_{q,p}) \end{aligned} \tag{1.13}$$

Utiliser cette dernière expression de $C^{J,J}_{m,n}$ dans celle de B, nous permet d'avoir

$$\begin{aligned} B &= \sum_p \sum_q (\sum_m \phi^J(\ell - 2^J m - 2^{J+1} q) g^0_m)(\sum_n \phi^J(k - 2^J n - 2^{J+1} p) g^0_n) C^{J+1,J+1}_{q,p} \\ &+ \sum_p \sum_q (\sum_m \phi^J(\ell - 2^J m - 2^{J+1} q) g^0_m)(\sum_n \phi^J(k - 2^J n - 2^{J+1} p) g^1_n) V^{J+1,J+1}_{q,p} \\ &+ \sum_p \sum_q (\sum_m \phi^J(\ell - 2^J m - 2^{J+1} q) g^1_m)(\sum_n \phi^J(k - 2^J n - 2^{J+1} p) g^0_n) S^{J+1,J+1}_{q,p} \\ &+ \sum_p \sum_q (\sum_m \phi^J(\ell - 2^J m - 2^{J+1} q) g^1_m)(\sum_n \phi^J(k - 2^J n - 2^{J+1} p) g^1_n) D^{J+1,J+1}_{q,p} \end{aligned}$$

En se servant de cette dernière équation, il devient simple de démontrer par récurrence que (1.11) est vrai $\forall J \in \mathbb{N}^*$ (J représente le niveau de décomposition en ondelette de l'image).

En effet, au premier rang, nous avons $J = 1$, en développant la dernière équation de (1.10) au niveau de l'image reconstituée c'est-à-dire $j = 0$, il est simple de constater que :

$$\begin{aligned} C^{0,0}_{\ell,k} = &\sum_n \sum_m \alpha^{1,1}_{\ell,k,m,n} C^{1,1}_{m,n} + \sum_n \sum_m \beta^{1,1}_{\ell,k,m,n} V^{1,1}_{m,n} + \\ &\sum_n \sum_m \gamma^{1,1}_{\ell,k,m,n} S^{1,1}_{m,n} + \sum_{j=1}^{J} \sum_n \sum_m \lambda^{1,1}_{\ell,k,m,n} D^{1,1}_{m,n} \end{aligned} \tag{1.14}$$

ce qui confirme que l'équation (1.11) est vraie au rang $J = 1$. Supposons que cette dernière équation soit vraie au rang $J \in N^*$ et montrons qu'elle reste vraie au rang $J+1$. En tenant compte des définitions des fonctions $\phi^{j+1}(k)$, $\psi^{j+1}(k)$ et en les injectant dans la dernière expression de B, celle

ci devient :

$$\begin{aligned} B &= \sum_p \sum_q \phi^{J+1}(\ell - 2^{J+1}q)\phi^{J+1}(k - 2^{J+1}p)C_{q,p}^{J+1,J+1} \\ &+ \sum_p \sum_q \phi^{J+1}(\ell - 2^{J+1}q)\psi^{J+1}(k - 2^{J+1}p)V_{q,p}^{J+1,J+1} \\ &+ \sum_p \sum_q \psi^{J+1}(\ell - 2^{J+1}q)\phi^{J+1}(k - 2^{J+1}p)S_{q,p}^{J+1,J+1} \\ &+ \sum_p \sum_q \psi^{J+1}(\ell - 2^{J+1}q)\psi^{J}(k - 2^{J+1}p)D_{q,p}^{J+1,J+1} \end{aligned}$$

Si nous remplaçons le premier terme du membre de droite de l'équation (1.11) par la dernière expression de B ci-dessus, on constate que cette équation est aussi vraie au rang $J+1$, d'où la démonstration par récurrence du calcul des fonction réelles ϕ^j et ψ^j et des termes $\alpha_{\ell,k,m,n}^{J,J}$, $\beta_{\ell,k,m,n}^{J,J}$, $\gamma_{\ell,k,m,n}^{J,J}$, $\lambda_{\ell,k,m,n}^{J,J}$. ■

Nous constatons alors que ϕ^{j+1}(resp ψ^{j+1}) est la convolution de ϕ^j (resp ψ^j) avec la suite g_k^0 (resp g_k^1) sur-échantillonnée d'un facteur de 2, j fois.

Revenons au calcul de la distorsion après une transformation en ondelette d'une composante, une quantification scalaire uniforme est effectuée sur les sous-bandes, un pas de quantification quelconque étant assigné à chaque sous-bande, de façon à respecter l'hypothèse faible distorsion. Après la quantification, la TODI est appliquée aux coefficients d'ondelette quantifiés. Soient $\widehat{\mathbf{X}}$ la variable aléatoire représentant l'image ainsi reconstituée et $\mathbf{X}$, celle qui désigne l'image originale. Appelons $\epsilon_{\ell,k}^J$ $\eta_{\ell,k}^j$ $\tau_{\ell,k}^j$ $\upsilon_{\ell,k}^j$ les bruits de quantification des coefficients d'échelles et des coefficients de détails horizontaux, verticaux, et diagonaux respectifs. D'après les équations (1.11) et (1.12), et compte tenu de ce qui précède, le bruit de quantification entre les variables aléatoires $\mathbf{X}$ et $\widehat{\mathbf{X}}$ peut s'exprimer comme suit :

$$\begin{aligned} &\sum_m \sum_n \phi^J(\ell - 2^J m)\phi^J(k - 2^J n)\epsilon_{\ell,k}^J + \sum_{j=1}^J \sum_m \sum_n \phi^j(\ell - 2^j m)\psi^j(k - 2^j n)\eta_{\ell,k}^j \\ &+ \sum_{j=1}^J \sum_m \sum_n \psi^j(\ell - 2^j m)\phi^j(k - 2^j n)\tau_{\ell,k}^j + \sum_{j=1}^J \sum_m \sum_n \psi^j(\ell - 2^j m)\psi^j(k - 2^j n)\upsilon_{\ell,k}^j \end{aligned}$$

Sous l'hypothèse que les bruits de quantification dans le domaine transformé en ondelette soient décorrélés, de moyennes nulles et de variances identiques dans chaque sous-bande, la variance du bruit de quantification est alors obtenue par :

$$\begin{aligned} &\sum_m \sum_n \phi^J(\ell - 2^J m)^2\phi^J(k - 2^J n)^2 v_\epsilon^J + \sum_{j=1}^J \sum_m \sum_n \phi^j(\ell - 2^j m)^2\psi^j(k - 2^j n)^2 v_\eta^j \\ &+ \sum_{j=1}^J \sum_m \sum_n \psi^j(\ell - 2^j m)^2\phi^j(k - 2^j n)^2 v_\tau^j + \sum_{j=1}^J \sum_m \sum_n \psi^j(\ell - 2^j m)^2\psi^j(k - 2^j n)^2 v_\upsilon^j \end{aligned}$$

où $v_\epsilon^J = \mathrm{var}(\epsilon_{\ell,k}^J)$, $v_\eta^j = \mathrm{var}(\eta_{\ell,k}^j)$, $v_\tau^j = \mathrm{var}(\tau_{\ell,k}^j)$, $v_\upsilon^j = \mathrm{var}(\upsilon_{\ell,k}^j)$, var() ici désignant la variance. La distorsion globale moyenne pour une image ayant L pixels peux s'écrire comme suit :

$$\begin{aligned} D_{glo} &= \frac{L_J}{L}||\phi^J||^4 v_\epsilon^J + \sum_{j=1}^{J} \frac{L_j}{L}||\phi^j||^2||\psi^j||^2 v_\eta^j \\ &+ \sum_{j=1}^{J} \frac{L_j}{L}||\phi^j||^2||\psi^j||^2 v_\tau^j + \sum_{j=1}^{J} \frac{L_j}{L}||\psi^j||^4 v_\upsilon^j \end{aligned} \tag{1.15}$$

Ici $||.||$ désigne la norme euclidienne, L_j représente le nombre de pixels de la sous-bande j et vaut sensiblement $L/2^{2j}$ en supposant que ce rapport soit toujours un entier.

Les coefficients de pondération de la transformée en ondelette dépendent essentiellement des filtres de synthèse qui sont utilisés lors de la reconstruction du signal. Ces coefficients de pondération sont facilement calculables car les suites $\phi^j(k)$ (resp $\psi^j(k)$) ne sont qu'une convolution de g_k^0 (resp g_k^1) avec elle-même sur-échantillonnée d'un facteur 2 un certain nombre de fois.

Pour établir une analogie avec l'expression entre crochets de l'équation (1.8), il convient de remarquer d'une part que π_m n'est rien d'autre que le rapport L_j/L, et d'autre part que la distorsion $D_i^{(m)}$ n'est rien d'autre que ce que nous avons appelé ici v_η^j, v_ϵ^J, v_τ^j ou encore v_υ^j. **Ainsi, les coefficients de pondération w_m provenant de la TOD et qui interviennent dans les formules de la distorsion de la section 1.2 valent respectivement** $w_m = ||\phi^J||^4$ pour la sous-bande basse résolution, $w_m = ||\phi^j||^2||\psi^j||^2$ pour les sous-bandes de détails horizontaux et verticaux et $w_m = ||\psi^j||^4$ pour les sous-bandes de détails diagonaux avec $(1 \leq j \leq J)$.

Notons que dans [64] une telle formule de la distorsion moyenne sur une composante après une transformée en ondelette a été utilisée mais sans expliciter le calcul des coefficients de pondération. Dans [76], le calcul des coefficients de pondération après une transformée en ondelette est complètement explicité pour un signal $1D$, des indications sont données pour la généralisation du calcul des coefficients à un signal $2D$. Les auteurs de ce dernier article font aussi l'hypothèse que les bruits de quantification des différentes sous-bandes sont indépendants. Usevitch [73] a également proposé une méthode de calcul de ces coefficients de pondération pour une ondelette biorthogonale dans le cadre d'une allocations optimale de bits, une étude similaire a été faite dans [15]. Dans les deux derniers cas, le calcul des coefficients de pondération est détaillé pour un signal $1D$, des indications sur l'extension à l'image sont faites en utilisant la propriété de séparabilité de l'ondelette.

Conclusion : Dans ce premier chapitre, nous avons défini et proposé différents schémas de compression qui ont la particularité d'utilise une transformation pour réduire la redondance spatiale et une autre pour réduire la redondance spectrale. Dans certains cas, la transformation permettant de réduire la redondance spectrale permet aussi de réduire une partie de la redondance spatiale, c'est notamment le cas du schéma mixte en sous-bandes. Nous avons ensuite explicité une approximation de la distorsion de ces schémas de compression sous certaines hypothèses qui ont été explicitées et justifiées pour la plupart. Enfin, nous avons fait une démonstration permettant d'avoir une formule explicite des coefficients de pondérations dépendant de la TODI qui interviennent dans ces formules d'approximation de la distorsion, les autres coefficients de pondérations ayant été explicités sous forme d'équations simples. Par la suite, nous étendons les notions de gains de codage et de réduction de debit généralisés

Chapitre 2

Gain de codage et réduction maximale de débit généralisés

Dans ce chapitre, nous étendons les notions de ***gain de codage généralisé et de réduction maximale de débit généralisée*** aux transformations utilisées pour la réduction de la redondance spectrale aux différents schémas de compression définis au premier chapitre. Nous expliciterons les formules uniquement pour le schéma général mixte en sous-bandes. Ensuite, nous donnons les pas de quantifications de chaque sous-bande de chaque composante pour une allocation optimale de bits à haute résolution toujours pour ce schéma. Ce chapitre s'achève par un bref rappel sur les transformations qui permettent d'optimiser ces gains de codage sous les hypothèses haut débit et qui ont été introduites dans [51] [52] et par une étude de leur complexité algorithmique.

2.1 Définitions des gains de codage généralisés

Dans la suite, sauf mention explicite, nous parlerons des gains de codage généralisés pour désigner les deux notions de gain de codage généralisé et de réduction maximale de débit généralisée. Soit $\mathbf{X}$, une image numérique ayant N composantes, chaque composante $\mathbf{X}_i$ $(1 \leq i \leq N)$ étant constituée de L pixels, les pixels sont pris et rangés dans le sens des colonnes par exemple. L'image $\mathbf{X}$ est représentée par une matrice constituée de N lignes et de L colonnes et, chaque ligne représente une composante $\mathbf{X}_i$. Une décomposition en ondelette sur $J \in \mathbb{N}^*$ niveaux est appliquée sur chaque composante de l'image et les sous-bandes obtenues sont notées $\mathbf{X}_i^{(m)}$, $1 \leq i \leq N$, $1 \leq m \leq M = 3 \times J + 1$. Pour réduire la redondance spectrale, une transformation linéaire $\mathbf{A}^{(m)}$ est ensuite appliquée à sous-bande transformé per$(\mathbf{X}_i^{(m)})$ avec $P = 4$ (cas du schéma mixte en sous-bandes), ainsi $\mathbf{A}^{(m)}$ est une transformation de taille $NP \times NP$. Après les transformations spectrales, les sous-bandes $\mathbf{Y}_{i,\ell}^{(m)}$ $(1 \leq \ell \leq P)$ sont obtenues. Notons que les transformations utilisées sont telles que le nombre de composantes de l'image devient NP, et le nombre de pixels (qui deviennent des coefficients après la transformation) par composante et par sous-bande est divisé par P. Les sous-bandes $\mathbf{Y}_{i,\ell}^{(m)}$ sont ensuite quantifiées et codées. Notons $R_{i,\ell}^{(m)}$ le débit moyen de codage de la sous-bande $\mathbf{Y}_{i,\ell}^{(m)}$. Ce débit représente le nombre de bits moyen nécessaire pour coder chaque coefficient transformé dans cette sous-bande. La quantification entraîne une dégradation de l'information, cette dégradation peut être mesurée par la distorsion. Notons $D_{i,\ell}^{(m)}$ la distorsion obtenue dans la sous-bande $\mathbf{Y}_{i,\ell}^{(m)}$ après quantification, la mesure de distorsion utilisée restant l'erreur

quadratique moyenne.

En appliquant les transformations linéaires inverses $(\mathbf{A}^{(m)})^{-1}$ aux sous-bandes quantifiées $\mathbf{Y}_{i,\ell}^{q(m)}$, puis l'opérateur inverse per^{-1} par sous-bande, nous obtenons une approximation des composantes après la TOD. Enfin, en appliquant la TODI sur chaque composante, une image approchée $\widehat{\mathbf{X}}$ est ainsi reconstituée. Notons $D_{glo} = D(\mathbf{X}, \widehat{\mathbf{X}})$, la distorsion globale moyenne obtenue sur l'image, et $R(\mathcal{A}, D_{glo})$ le débit moyen de codage obtenu pour la distorsion moyenne D_{glo} après application de la transformation $\mathcal{A}$ où $\mathcal{A}$ est constituée de l'ensemble des transformations $\mathbf{A}^{(m)}$. Les définitions mathématiques de la distorsion $D(\mathbf{X}, \widehat{\mathbf{X}})$ et du débit moyen de codage R sont données par :

$$\begin{aligned} D(\mathbf{X}, \widehat{\mathbf{X}}) &= \frac{1}{N}\frac{1}{L}\sum_{i=1}^{N}\sum_{j=1}^{L}(X_i(j) - \widehat{X}_i(j))^2 \\ R &= \frac{1}{NP}\sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{\ell=1}^{P}\pi_m R_{i,\ell}^{(m)} \end{aligned}$$

où π_m est le rapport entre le nombre de pixels K_m de la sous-bande (m) et L, $X_i(j)$ (respectivement $\widehat{X}_i(j)$) est le $j^{ème}$ pixel de la composante $\mathbf{X}_i$ (respectivement $\widehat{\mathbf{X}}_i$) en ayant prédéfini un balayage de l'image, colonne par colonne par exemple. Nous faisons fi des problèmes dus aux effets de bord, notamment lorsque le nombre de lignes ou de colonnes n'est pas un nombre pair lors du passage d'un niveau de décomposition en ondelette m, au niveau $m+1$, car il est simple de définir une règle dans ce cas précis.

2.1.1 Réduction maximale de débit généralisée

La **réduction maximale de débit généralisée** exprimée en bits par pixels (bpp) d'une transformation $\mathcal{A}$ utilisée pour la réduction de la redondance spectrale se définit [52] comme suit :

$$R_{max}(\mathcal{A}, D) = \min_{D_{glo} \leq D} R(\mathcal{A}_0, D_{glo}) - \min_{D_{glo} \leq D} R(\mathcal{A}, D_{glo})$$

où $R(\mathcal{A}_0, D_{glo})$ est le débit global moyen de codage obtenu après application d'une transformation étalon sur les sous-bandes, pour une distorsion D_{glo}, $R(\mathcal{A}, D_{glo})$ est le débit global moyen de codage obtenu après application de la transformation $\mathcal{A}$. La transformation étalon devra être au préalable définie, et une transformation étalon possible est la transformation identité (c'est-à-dire que les sous bandes $\mathbf{X}_{i,\ell}^{(m)}$ sont quantifiées après la TOD).

2.1.2 Gain de codage généralisé

Généralement, en compression d'image, la distorsion est mieux appréciée lorsque celle-ci est exprimée en termes de PRSB. Le PRSB, rappelons le, est une quantité adimensionnelle qui est exprimé en dB. Si l'image numérique de départ est codée avec une profondeur de c bits par pixel, pour une distorsion globale moyenne D_{glo}, le PRSB est défini par :

$$PRSB = 10\log_{10}[\frac{(2^c - 1)^2}{D_{glo}}]$$

Le **gain de codage généralisé** d'une transformation $\mathcal{A}$ utilisée pour la réduction de la redondance spectrale et, exprimé en dB, est défini par :

$$G(\mathcal{A}, R) = \max_{R_{glo} \leq R} PRSB(\mathcal{A}, R_{glo}) - \max_{R_{glo} \leq R} PRSB(\mathcal{A}_0, R_{glo})$$

où $PRSB(\mathcal{A}, R_{glo})$ est le PRSB obtenu après application des transformations $\mathbf{A}^{(m)}$ sur les sous-bandes pour un débit moyen global R_{glo} et, $PRSB(\mathcal{A}_0, R_{glo})$ est le PRSB obtenu après application des transformations étalons $\mathbf{A}_0^{(m)}$ sur les sous-bandes $\mathbf{X}_{i,\ell}^{(m)}$ pour un débit moyen global R_{glo}.

La notion de gain de codage généralisé est définie par rapport à une transformation étalon qui doit être choisie. En effet, la notion de gain fait appel à une comparaison entre deux quantités de même nature, d'où les définitions simples et naturelles qui ont été données, et qui ne sont rien d'autre que la différence point à point entre les courbes débit versus PRSB ou vice versa, en supposant qu'une allocation optimale de bits a été faite entre les quantificateurs. Notons qu'il existe une dualité entre la réduction maximale de débit généralisé et le gain de codage généralisé, puisque dépendant tous les deux, des courbes débit versus PRSB dans le cadre d'une allocation optimale de débit ou de distorsion entre les sous-bandes des différentes composantes.

2.2 Allocation optimale asymptotique de bits du schéma mixte en sous-bandes

D'après les définitions des gains de codage, il est indispensable de faire une allocation optimale de bits ou de distorsion (ce problème est dual) pour évaluer l'un ou l'autre des gains de codage. Le problème d'allocation optimale de bits se pose comme suit. *Étant donné un budget de bits global moyen R par pixel et par composante, comment répartir ce budget moyen de bit entre les différents quantificateurs pour que la distorsion globale moyenne D_{glo} entre l'image de départ et l'image reconstruite soit minimale, sous la contrainte que le débit global moyen effectif R_{glo} soit inférieur ou égal à R.*

C'est un problème d'optimisation avec contrainte qui peut être résolu en utilisant la méthode des multiplicateurs de Lagrange. Le problème d'allocation optimale de bits étant dual à celui d'allocation optimale de distorsion, en utilisant la méthode des multiplicateurs de Lagrange, cela revient à minimiser pour le schéma en sous-bandes dont la formule de la distorsion est donnée par l'équation (1.7) l'expression suivante :

$$\mathcal{L} = \frac{1}{NP}\sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{\ell=1}^{P}\pi_m R_{i,\ell}^{(m)} + \mu\left[\frac{1}{NP}\sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{\ell=1}^{P}\pi_m w_m w_{i,\ell}^{(m)} D_{i,\ell}^{(m)} - D\right] \tag{2.1}$$

par rapport aux distorsions $D_{i,\ell}^{(m)}$ et aux débits $R_{i,\ell}^{(m)}$où μ est le multiplicateur de Lagrange à déterminer et D une distorsion cible fixée. En effet, pour une transformation spectrale donnée, les $w_{i,\ell}^{(m)} = \|\mathbf{A}^{(m)-1}\mathbf{e}_{P(i-1)+\ell}\|^2$ deviennent des constantes et sont facilement calculables. Par ailleurs il est bien connu d'après la théorie de l'information, que pour de faibles distorsions (haut débit), le débit de codage [31] dans chaque sous bande $\mathbf{Y}_{i,\ell}^{(m)}$ (nous supposons que la loi de probabilité est la même pour tous les coefficients d'ondelette dans une sous-bande d'une composante et nous appelons $Y_{i,\ell}^{(m)}$ la variable aléatoire qui modélise ces coefficients) est donné par : $R_{i,\ell}^{(m)} \simeq H(Y_{i,\ell}^{(m)}) - \frac{1}{2}\log_2[cD_{i,\ell}^{(m)}]$ où $H(Y_{i,\ell}^{(m)})$ est l'entropie différentielle de $Y_{i,\ell}^{(m)}$ définie avec le logarithme à base deux et c une constante réelle, qui dans le cas d'une quantification scalaire uniforme vaut $c = 12$. Ce débit représente le débit optimal de codage asymptotique sous l'hypothèse que les coefficients d'ondelette dans une sous-bande sont i.i.d (indépendants et identiquement distribués). Le problème

de minimisation consiste dès lors à minimiser :

$$\mathcal{L} = -\frac{1}{NP}\sum_{i=1}^{N}\sum_{m=1}^{M}\sum_{\ell=1}^{P}\pi_m\frac{1}{2}\log_2(D_{i,\ell}^{(m)}) + \mu(\frac{1}{NP}\sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{\ell=1}^{P}\pi_m w_m w_{i,\ell}^{(m)} D_{i,\ell}^{(m)} - D)$$

par rapport aux distorsions $D_{i,\ell}^{(m)}$ et à μ. Pour minimiser cette expression, nous calculons les dérivées partielles par rapport aux $D_{i,\ell}^{(m)}$ et à μ et nous obtenons :

$$\frac{\partial\mathcal{L}}{\partial D_{i,\ell}^{(m)}} = -\frac{1}{NP}\pi_m\frac{1}{2\log(2)D_{i,\ell}^{(m)}} + \mu\frac{1}{NP}\pi_m w_m w_{i,\ell}^{(m)} \quad ; \quad \frac{\partial\mathcal{L}}{\partial\mu} = \frac{1}{NP}\sum_{m=1}^{M}\sum_{i=1}^{N}\sum_{\ell=1}^{P}\pi_m w_m w_{i,\ell}^{(m)} D_{i,\ell}^{(m)} - D$$

et en annulant ces dérivées partielles, l'optimum est atteint lorsque la distorsion de chaque sous-bande de chacune des composantes est calculée selon l'expression :

$$D_{i,\ell,opt}^{(m)} = \frac{D}{w_m w_{i,\ell}^{(m)}} \quad \text{et} \quad \mu_{opt} = \frac{1}{2D\log(2)} \tag{2.2}$$

Cette dernière équation est très intéressante car dans [31], il est démontré que la distorsion s'estime en fonction du pas de quantification $q_{i,\ell}^{(m)}$ du quantificateur de la $m^{ème}$ sous-bande de la $i^{ème}$ composante du $\ell^{ème}$ bloc à faibles distorsions par : $D_{i,\ell}^{(m)} \simeq [q_{i,\ell}^{(m)}]^2/12$, ainsi, pour une distorsion cible donnée, nous pouvons calculer les pas de quantification optimaux à appliquer à chaque quantificateur après les transformations spatiales et spectrales.

Remarque : Les définitions des gains de codage généralisés restent identiques pour les différents schémas de compression définis au premier chapitre. La formule de la distorsion est différente pour chaque type de schéma, ce qui a une influence sur le calcul de l'allocation optimale asymptotique de bits. Nous avons choisi de développer les formules d'allocation asymptotique du schéma mixte en sous-bandes qui est un schéma général, la méthode est la même pour les autres schémas de compression décrits au chapitre précédent. Il est à souligner ici que les définitions du gain de codage ne dépendent pas de P dans le cas du schéma mixte en sous-bandes, elles restent identiques $\forall$ P. Cependant l'estimation de la formule de la distorsion elle dépend de P et des hypothèses qui ont été formulées précédemment pour ce dernier schéma. Notons toutefois que pour le cas P quelconque du schéma mixte en sous-bandes, il a été démontré dans le chapitre précédent qu'il est possible d'exprimer la distorsion globale moyenne en fonction des distorsions des sous-bandes après transformations sous certaines hypothèses, c'est ce que résume l'équation (1.6). Le schéma en sous-bandes peut encore être vu comme un cas particulier du schéma mixte en sous-bandes avec $P = 1$. D'autre part, pour le schéma mixte en sous-bandes, si l'ondelette 9/7 de Daubechies est utilisée (elle est quasi orthogonale), et que les transformations spectrales $\mathbf{A}^{(m)}$ sont toutes orthogonales, l'allocation optimale de bits est atteinte lorsque la distorsion de chaque sous-bande par composante est prise comme $D_{i,l,opt}^{(m)} = \frac{D}{w_m}$ car dans ce cas, les coefficients de pondérations $w_{i,l}^{(m)} = 1$. Enfin, toujours pour le schéma mixte en sous-bandes, si de plus la TOD est orthogonale cas de l'ondelette de Haar par exemple, alors $D_{i,l,opt}^{(m)} = D$, c'est-à-dire que la distorsion devient la même pour toutes les sous-bandes des différentes composantes.

2.3 Allocation optimale de bits à bas débit

Les formules d'allocations optimales de la section 2.2 ne sont valables qu'à faibles distorsions. Pour évaluer les gains de codage généralisés à bas débits, nous utiliserons l'algorithme bien connu

d'allocation optimale de Shoham et Gersho qui est décrit dans [69]. Pour appliquer cet algorithme, nous avons besoin d'une formule d'approximation de la distorsion ainsi qu'une formule d'approximation du débit en fonction de la distorsion et du débit des différentes sous-bandes de chaque composante.

Les formules d'approximation de la distorsion établies au premier chapitre de cette partie comme nous le vérifierons dans le prochain chapitre, sont très bonnes et restent valables aussi bien à hauts débits qu'à bas débits. Celle du débit est somme toute un peu plus triviale car s'exprimant comme une somme pondérée des débits des différentes sous bandes. L'algorithme de Shoham et Gersho prend en entrée un ensemble fini de points débit distorsion de chacune des sous-bandes après transformations, une formule d'approximation de la distorsion et du débit en fonction des valeurs de ces points, ainsi qu'une valeur cible de débit ou de distorsion. Cet algorithme donne en sortie, le point débit distorsion optimal à prendre pour chaque sous-bande, pour atteindre le débit ou la distorsion cible en se servant des multiplicateurs de Lagrange. En effet, le problème d'allocation optimale de bits étant un problème de minimisation sous contrainte, ce problème est ramené à un problème de minimisation sans contrainte en utilisant un multiplicateur de Lagrange. La particularité de cet algorithme est qu'il permet, d'une façon simple et efficace, de trouver aussi bien la valeur optimale du paramètre lagrangien que celles des points débit distorsion permettant de résoudre le problème d'allocation optimale de bits pour un débit ou une distorsion cible donné, et ce dans un ensemble fini d'éléments. Des informations sur la théorie des multiplicateurs de Lagrange généralisés au cas de données discrètes se trouvent dans la publication d'Everett [29].

2.4 Transformations optimisant le gain de codage à faible distorsion

2.4.1 Cas simplifié

Soit $\mathbf{X}$, un vecteur aléatoire constitué de N composantes, et $\mathbf{A}$ une transformation linéaire inversible qui transforme $\mathbf{X}$ en $\mathbf{Y}$ par $\mathbf{Y} = \mathbf{AX}$. Supposons que l'on veuille faire un codage séparé des composantes de $\mathbf{Y}$ après quantification, on aimerait savoir quelle transformation maximise le gain de codage généralisé ci-dessus défini dans un schéma de compression un peu plus simple. La réponse à cette question a été clairement donnée dans [51]. Dans ce rapport, nous ne reprenons pas les développements réalisées dans cet article, nous faisons juste un résumé succinct des étapes importantes qui permettent d'aboutir aux résultats obtenus.

L'hypothèse de quantifications à haute résolution ou faible distorsion est faite à ce niveau. Et dans ce cas, pour une quantification scalaire uniforme d'une variable aléatoire X, le débit s'estime comme énoncé ci-haut en utilisant la formule de Bennett par :

$$H(X) - \frac{1}{2}\log_2(cD)$$

avec $H(X)$ l'entropie différentielle de la variable aléatoire X, D la distorsion introduite du fait de la quantification et qui est faible, et c étant une constante qui vaut 12. Ainsi, la longueur moyenne asymptotique de codage vaut :

$$\frac{1}{N}\sum_{i=1}^{N}[H(Y_i) - \frac{1}{2}\log_2(cD_i)]$$

où $H(Y_i)$ est l'entropie différentielle de Y_i, et D_i est la distorsion associée à la quantification faite sur Y_i. La distorsion qui nous intéresse n'est pas celle sur les coefficients transformés entre $\mathbf{Y}$ et $\mathbf{Y}^q$, mais plutôt celle qui existe entre le vecteur original $\mathbf{X}$ et celui reconstitué $\widehat{\mathbf{X}}$. L'erreur de reconstruction peut être obtenue par :

$$\widehat{\mathbf{X}} - \mathbf{X} = \mathbf{A}^{-1}[\mathbf{Y}^q - \mathbf{Y}] \tag{2.3}$$

En admettant que l'hypothèse $H1$ qui a été énoncée et justifiée au chapitre précédent soient vraie, c'est-à-dire que les composantes du bruit de quantification sur $\mathbf{Y}$ sont décorrélées et de moyennes nulles, la distorsion globale moyenne sur $\mathbf{X}$ s'exprime alors comme suit :

$$D_{glo} = \frac{1}{N}\sum_{j=1}^{N} w_j D_j$$

avec $w_j = \sum_{i=1}^{N}(\mathbf{A}^{-1})_{ij}^2$ et, D_j est la distorsion moyenne sur la $j^{ème}$ composante de $\mathbf{Y}^q - \mathbf{Y}$.

Le problème du gain de codage se pose alors de la façon suivante. Étant donné une distorsion globale moyenne D, comment répartir cette distorsion de telle sorte que la longueur moyenne de codage

$$\sum_{i=1}^{N}[H(Y_i) - \frac{1}{2}\log_2(cD_i)]$$

soit minimale sous la contrainte que $D_{glo} \leq D$, où D_{glo} est la distorsion globale moyenne effective obtenue entre le vecteur original et celui reconstitué. La répartition ici se fait par rapport aux D_i, $(1 \leq i \leq N)$, et par rapport à la transformation $\mathbf{A}$. En d'autres termes, quelle est la transformation $\mathbf{A}$ qui permet de minimiser cette dernière équation sous la contrainte ci-dessus. Dans [51], il est démontré que pour $\mathbf{A}$ fixée cette minimisation conduit à prendre $D_i = D/w_i$ et le problème se ramène donc à la minimisation sans contrainte du critère :

$$C(\mathbf{A}) = \sum_{i=1}^{N}[H(Y_i) + \frac{1}{2}\log_2(w_i)]$$

Il est facile de remarquer que w_i est le $i^{ème}$ terme diagonal de la matrice $\mathbf{A}^{-T}\mathbf{A}^{-1}$, ce qui permet d'écrire le critère précédent comme :

$$C(\mathbf{A}) = \sum_{i=1}^{N} H(Y_i) + \frac{1}{2}\log_2 \det \mathrm{diag}(\mathbf{A}^{-T}\mathbf{A}^{-1}) \tag{2.4}$$

où diag est l'opérateur qui construit une matrice diagonale à partir des termes diagonaux de la matrice qui lui est passée en argument. Ce critère peut encore se mettre sous la forme $C(\mathbf{A}) = C_{ACI}(\mathbf{A}) + Q$ où $C_{ACI}(\mathbf{A}) = \sum_{i=1}^{N} H(Y_i) - \log_2 \det \mathbf{A}$ est le critère utilisé pour résoudre le problème d'ACI en prenant l'information mutuelle comme critère à minimiser, et $Q = \frac{1}{2}\log_2 \frac{\det \mathrm{diag}(\mathbf{A}^{-T}\mathbf{A}^{-1})}{\det(\mathbf{A}^{-T}\mathbf{A}^{-1})}$ est un terme réel toujours positif ou nul [51] [52]. Il est nul si et seulement si la matrice $\mathbf{A}$ est orthogonale (en fait, il faut et il suffit que les vecteurs colonne de la matrice soient 2 à 2 orthogonaux). Dans [51] [52], un algorithme *GCGsup* que nous appellerons *Ica_opt* et qui permet de trouver ladite transformation $\mathbf{A}$ minimisant ce critère est décrit. Il est également détaillé un algorithme *Ica_orth* permettant de trouver la transformation $\mathbf{A}$ qui minimise ce critère sous la contrainte $Q = 0$. Nous

garderons le même nom pour ce dernier. L'article [51] est joint dans l'Annexe3. Notons tout de même que ces algorithmes sont obtenus en utilisant une méthode de type Newton, qui nécessite le calcul du gradient et une approximation du hessien à chaque itération. Les formules explicites du gradient et du hessien du critère dans les deux cas, sont données et expliquées dans [51] [52].

2.4.2 Cas du schéma en sous-bandes

Pour ce qui est du schéma de compression en sous-bandes qui a été explicité dans la sous-section 1.1.3, le problème du gain de codage se pose comme suit : soit D une distorsion globale moyenne fixé, comment répartir cette distorsion entre les sous-bandes $\mathbf{Y}^{q(m)}$ $(1 \leq m \leq M)$ de telle sorte que la longueur moyenne de codage asymptotique

$$R = \frac{1}{N} \sum_{i=1}^{N} \sum_{m=1}^{M} \pi_m [H(Y_i^{(m)}) - \frac{1}{2} \log_2(cD_i^{(m)})] \tag{2.5}$$

soit minimale sous la contrainte que $D_{glo} \leq D$, où D_{glo} est la distorsion globale moyenne effective obtenue entre l'image originale et celle reconstituée. En se servant de la formule d'approximation de la distorsion de l'équation (1.7) et des résultats obtenus dans la section 2.2, nous déduisons que la résolution asymptotique de ce problème de minimisation sous contrainte consiste à prendre $D_{i,opt}^{(m)} = D/w_m w_i^{(m)}$. Et dans ce cas, la longueur moyenne de codage asymptotique devient alors :

$$R = \sum_{m=1}^{M} \pi_m \left[\frac{1}{N} \sum_{i=1}^{N} \left\{ H(Y_i^{(m)}) + \frac{1}{2} \log_2 w_i^{(m)} \right\} + \frac{1}{2} \log_2 w_m \right] - \frac{1}{2} \log_2(cD) \tag{2.6}$$

L'ondelette étant fixée (ici la 9/7 de Daubechies), le problème de la compression pour le schéma en sous-bandes revient donc à rechercher de façon indépendante les M transformations spectrales $\mathbf{A}^{(m)}$ qui minimisent chacune le critère

$$\sum_{i=1}^{N} [H(Y_i^{(m)}) + \frac{1}{2} \log_2(w_i^{(m)})]$$

avec $w_i^{(m)} = \sum_{i=1}^{N} (\mathbf{A}^{-1(m)})_{ij}^2$. Ce qui revient pour une sous-bande $\mathbf{Y}^{(m)}$ à rechercher la transformation qui minimise le même critère que celui de l'équation (2.4). Ainsi, la transformation obtenue par l'algorithme Ica_opt, appliquée sur chaque sous-bande $\mathbf{Y}^{(m)}$, permet de maximiser le gain de codage généralisé à faible distorsion du schéma en sous-bandes. De la même façon et en se servant de ce qui précède, nous pouvons conclure que la transformation obtenue par l'algorithme Ica_orth, appliquée sur chaque sous-bande $\mathbf{Y}^{(m)}$, est la transformation orthogonale qui permet de maximiser le gain de codage généralisé à faible distorsion du schéma en sous-bandes.

Discussion 1 : La formule d'approximation de la distorsion du schéma en sous-bandes comme nous l'avons vue est valide sous l'hypothèse $H1$. Pour le schéma mixte en sous-bandes, la formule d'approximation de la distorsion a été établie sous les hypothèses $H1$ et $H2$. Sous ces hypothèses, il devient simple, en utilisant un raisonnement identique à celui du schéma en sous-bandes, d'arriver à la même conclusion, c'est-à-dire que la transformation obtenue par l'algorithme Ica_opt, appliquée sur chaque sous-bande per($\mathbf{Y}^{(m)}$), permet de maximiser le gain de codage généralisé à faible distorsion du schéma mixte en sous-bandes.

2.4.3 Cas du schéma séparable

Pour le schéma séparable, il est important de distinguer les deux cas, le cas où la transformation spectrale s'adapte aux coefficients d'ondelette et le cas où elle s'adapte à l'image brute. Dans ce dernier cas, posons $\mathbf{Z} = \mathbf{AX}$ où $\mathbf{X}$ est l'image originale et $\mathbf{A}$ la transformation spectrale. Pour ce cas, ne disposant que de l'image brute, nous ne pouvons que minimiser le critère

$$\sum_{i=1}^{N}[H(Z_i) + \frac{1}{2}\log_2(w_i)]$$

où $\mathbf{Z}_i$ est la $i^{ème}$ composante de $\mathbf{Z}$. Cela revient à chercher une transformation optimisant le gain de codage dans un schéma de compression où la TOD n'intervient pas, et à utiliser la transformation trouvée dans le schéma de compression séparable avec adaptation dans l'espace image. Ce qui peut se justifie dans le cas où l'effet de la TOD est très faible. Dans le premier cas cité, si nous minimisons toujours le critère ci-dessus défini, mais cette fois-ci sur les composantes de $\mathbf{Y} = \mathbf{A}(\mathbf{XW}^T)$. Cela revient à ignorer la partition de ces coefficients en sous-bandes et de considérer qu'ils sont homogènes, c'est-à-dire que la distribution dans les différentes sous-bandes est la même. Ce critère correspond donc à un schéma de compression où l'on quantifie et l'on code les coefficients d'ondelette d'une composante de la même façon sans tenir compte de leur appartenance à une sous-bande. Dans les deux cas, en procédant de cette manière, le critère minimisé n'est pas celui qui optimise le gain de codage généralisé à faibles distorsions.

Pour maximiser le gain de codage généralisé à haute résolution, nous devons minimiser pour une distorsion cible donnée D_c la longueur moyenne de codage

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{m=1}^{M}\pi_m[H(Y_i^{(m)}) - \frac{1}{2}\log_2(cD_i^{(m)})]$$

sous la contrainte que la distorsion

$$\frac{1}{N}\sum_{m=1}^{M}\sum_{i=1}^{N}\pi_m w_m w_i D_i^{(m)} \leq D_c$$

Proposition 2.4.1 *La longueur moyenne de codage devient plus courte en codant séparément les sous-bandes des différentes composantes. En effet, nous avons :*

$$H(Y_i) \geq \sum_{m=1}^{M}\pi_m H(Y_i^{(m)})$$

avec égalité si et seulement si la distribution de Y_i est la même dans chaque sous-bande.

Preuve : Introduisons la variable aléatoire J indiquant la sous-bande m à laquelle appartient un coefficient quelconque Y_i de $\mathbf{Y}_i$. Nous savons d'une part que l'entropie de Y_i conditionnellement à $J = m$ est donnée par :

$$H(Y_i^{(m)}) = -\int f(Y_i = y|J = m)\log_2 f(Y_i = y|J = m)dy$$

et d'autre part que $\sum_{m=1}^{M}\pi_m H(Y_i^{(m)}) = H(Y_i|J)$ est l'entropie de Y_i sachant J. Or d'après la théorie de l'information [27], il est démontré que

$$H(Y_i|J) = H(Y_i, J) - H(J)$$

et par conséquent $H(Y_i) - H(Y_i|J) = H(Y_i) + H(J) - H(Y_i, J) = I(Y_i; J)$ constitue l'information mutuelle entre Y_i et J qui est toujours positive ou nulle. Ainsi $H(Y_i) \geq H(Y_i|J)$ avec l'égalité si et seulement si Y_i et J sont indépendants, c'est-à-dire que la loi de Y_i sachant $J = m$ ne dépend pas de m, d'où la preuve. ■

Dans le cas du schéma séparable avec adaptation sur l'espace ondelette, la transformation spectrale $\mathbf{A}$ est appliquée et adaptée à toutes les sous-bandes après la TOD. L'allocation optimale de bits pour $\mathbf{A}$ fixée (en se référant à ce qui a été fait dans la sous-section 2.2) conduit à $\log_2 D_i^{(m)} = -\log_2(\mathbf{A}^{-T}\mathbf{A}^{-1})_{ii}$ plus un terme constant indépendant de $\mathbf{A}$. Ainsi, en utilisant les transformations retournées par les algorithmes *Ica_opt* et *Ica_orth* dans ce schéma, le critère minimisé n'est pas la longueur de codage asymptotiquement atteignable, mais une borne supérieure de celle-ci comme nous le montre la proposition 2.4.1. Le schéma séparable avec adaptation dans le domaine image est intéressant dans la mesure où c'est ce schéma qui a été choisie pour la compression des images multi-composantes dans la norme JPEG2000, il nous permettra donc de faire une comparaison des performances par rapport à ce dernier. Disposant des algorithmes *Ica_opt* et *Ica_orth*, nous nous sommes proposés d'évaluer leurs performances dans le cas du schéma séparable.

Proposition 2.4.2 *Dans le cas du schéma séparable, si la transformation spectrale est la TKL et que la TOD utilisée est orthogonale, le résultat obtenu* $\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X}\boldsymbol{W}^T$ *est identique quelque soit l'espace d'adaptation choisi.*

Preuve

En effet, dans le cas où la base d'ondelette est orthogonale :

$$\langle \mathbf{X}_i, \mathbf{X}_j \rangle = \langle (\mathbf{X}\mathbf{W}^T)_i, (\mathbf{X}\mathbf{W}^T)_j \rangle = \sum_{m=1}^{M} \langle (\mathbf{X}\mathbf{W}^T)_i^{(m)}, (\mathbf{X}\mathbf{W}^T)_j^{(m)} \rangle$$

où $\langle , \rangle$ désigne le produit scalaire. Donc si on ne centre pas les variables, faire la TKL avant ou après la TOD (orthogonale) aboutit à la même transformation spectrale. En pratique, les variables sont centrées et nous devons considérer la matrice de terme général $\langle \mathbf{X}_i - \mu_i \mathbf{1}, \mathbf{X}_j - \mu_j \mathbf{1} \rangle$, où μ_i désigne la moyenne arithmétique de la composante de $\mathbf{X}_i$ et $\mathbf{1}$ est le vecteur ligne ne comportant que des 1. Or $\mathbf{1}\mathbf{W}^T$ est nul partout sauf sur la sous-bande de plus basse fréquence (que nous appellerons sous-bande $m = 1$ pour fixer les idées) où il vaut α. Par suite nous avons

$$\langle \mathbf{X}_i - \mu_i \mathbf{1}, \mathbf{X}_j - \mu_j \mathbf{1} \rangle = \langle (\mathbf{X}\mathbf{W}^T)_i^{(1)} - \alpha\mu_i \mathbf{1}, (\mathbf{X}\mathbf{W}^T)_j^{(1)} - \alpha\mu_j \mathbf{1} \rangle + \sum_{m=2}^{M} \langle (\mathbf{X}\mathbf{W}^T)_i^{(m)}, (\mathbf{X}\mathbf{W}^T)_j^{(m)} \rangle$$

Mais nous savons que

$$0 = \langle \mathbf{X}_i - \mu_i \mathbf{1}, \mathbf{1} \rangle = \langle (\mathbf{X}\mathbf{W}^T)_i^{(1)} - \alpha\mu_i \mathbf{1}, \mathbf{1} \rangle$$

et donc $\alpha\mu_i$ doit être égal à la moyenne arithmétique du vecteur ligne $(\mathbf{X}\mathbf{W}^T)_i^{(1)}$. Ainsi, faire la TKL sur l'image initiale $\mathbf{X}$ conduirait à la même matrice de transformation que faire la TKL sur $\mathbf{X}\mathbf{W}^T$ à condition que la moyenne arithmétique des sous-bandes $(\mathbf{X}\mathbf{W}^T)^{(m)}$, $m > 1$ soit le vecteur nul. En d'autres termes, cette dernière condition revient tout simplement à dire que la moyenne arithmétique du vecteur ligne $(\mathbf{X}\mathbf{W}^T)_i^{(m)}$ est nulle pour $m > 1$. Parisot [56] a utilisé

une modélisation où il suppose effectivement que les sous-bandes (excepté la sous-bande basse fréquence) après une TOD ont une moyenne arithmétique nulle. Ce n'est pas vrai en général, mais c'est approximativement vrai, car l'espérance mathématique de $(\mathbf{X}\mathbf{W}^T)^{(m)}$ est nulle pour $m > 1$. Pour voir cela, il suffit de remarquer que $(\mathbf{X}\mathbf{W}^T)^{(m)} = [(\mathbf{X} - \mu\mathbf{1})\mathbf{W}^T]^{(m)}$ pour tout vecteur colonne constant μ et donc pour le cas particulier où μ est le vecteur des moyennes de $\mathbf{X}$ (qui ne dépend pas de l'indice des pixels), nous obtenons le résultat. Ce qui parachève la démonstration de cette proposition. ■

2.4.4 Complexité algorithmique de *Ica_opt* et *Ica_orth*

Les algorithmes de minimisation permettant d'obtenir *Ica_opt* et *Ica_orth* sont des algorithmes itératifs de descente de type Quasi-Newton. Ils nécessitent à chaque itération le calcul du critère, celui du gradient ainsi que celui de l'hessien. Pour avoir un ordre d'idée de leur complexité algorithmique, nous allons évaluer d'une façon assez sommaire, la complexité algorithmique de ces entités. Les détails sur le calcul du gradient et de l'hessien des critères à minimiser pour obtenir ce algorithmes se trouvent dans [51] [52]. Rappelons que l'article [51] est joint en Annexe3 dans sa version originale.

Soient N le nombre de composantes du signal original $\mathbf{X}$ et L la taille de chacune des composantes de $\mathbf{X}$. Appelons r le nombre de points de discrétisation nécessaire pour le calcul de la fonction score dont l'algorithme est explicité dans [62]. Nous avons en général $r \ll L$ et qui est un entier compris dans l'intervalle [30; 60]. L'estimation de la fonction score des N composantes se fait avec une complexité de l'ordre de $\mathrm{O}(NrL)$. L'estimation de l'entropie différentielle des N composantes se fait avec une complexité algorithmique du même ordre que celle de la fonction score. L'inversion d'une matrice carrée $\mathbf{A}$ de taille N se fait avec une complexité de l'ordre de $\mathrm{O}(N^3)$. Le calcul du déterminant de cette matrice est de l'ordre de $\mathrm{O}(N^3)$, idem pour le produit de deux matrices carrées de taille N. Ainsi l'estimation du critère de l'équation 2.4 se fait avec une complexité en $\mathrm{O}(NrL + 2N^3)$. L'estimation du gradient de ce critère qui nécessite le calcul de la fonction score est faite avec une complexité en $\mathrm{O}(NrL + N^2L)$. Le hessien, est estimé avec une complexité en $\mathrm{O}(NrL + 2N^2L)$. Partant d'un point initial choisi comme solution initiale, nous devons calculer, le gradient, une approximation du hessien et le critère à chaque itération, puis trouver le pas idéal pour faire la mise à jour du point courant qui devient la solution finale recherchée lorsqu'une condition définie au départ est atteinte. En général, pour passer d'une itération à une autre nous avons besoin de calculer ces trois entités une seule fois, il arrive pour certaines itérations que l'algorithme aient besoin de les calculer deux ou trois fois. Cela dépend du pas d'ajustement λ qui est utilisé pour la mise à jour de la solution $x_k = x_{k-1} - \lambda\delta x$, δx dépendant du gradient et de l'hessien en x_{k-1}.

Soient p l'entier représentant le nombre d'itérations au bout duquel l'algorithme converge, cet entier dépend de la condition initiale mais en général, nous avons constaté que l'algorithme converge au bout de $p \in [20;\ 60]$ itérations. Finalement, l'algorithme *Ica_opt* converge vers sa solution avec une complexité en $\mathrm{O}(3NprL + 2pN^3 + 3pN^2L)$. L'algorithme *Ica_orth* a une complexité de calcul similaire à *Ica_opt*.

Conclusion : Dans ce chapitre, nous avons étendu la notion des gains de codage généralisés introduites dans [51] aux schémas de compression défini dans le premier chapitre dans le cadre d'un codage séparé des composantes. Par la suite, en prenant en exemple le schéma mixte en sous-bandes, nous avons évalué les pas de quantification optimaux à appliquer à chaque sous-bande de chaque composante à haute résolution. Enfin, nous avons fait un rappel sur les transformations modifiées à

base d'ACI qui ont été mis à jour par Narozny [52] [51], et nous avons fait une étude simple de leur complexité algorithmique en ayant au préalable démontré que ces transformations permettent de maximiser le gain de codage généralisé non seulement du schéma en sous-bandes, mais aussi celui du schéma mixte en sous-bandes. Par la suite, nous réalisons une étude expérimentale pour évaluer les gains de codage généralisés de quelques images multi-composantes.

Chapitre 3

Evaluation des performances sur quelques images multi composantes

Disposant d'une base de données constituée d'images multi spectrales et hyper spectrales, nous avons évalué les performances des systèmes de compression qui ont été décrits dans le premier chapitre de cette partie. Dans le présent chapitre, nous définissons les images multi composantes qui constituent notre base de données. Après avoir vérifié la validité des formules d'approximation de la distorsion établies au premier chapitre de cette partie, nous donnons les résultats des performances en termes de courbes débit vs PRSB de ces images à partir des systèmes de compression que nous avons décrits et nous déduisons les gains de codage généralisés à partir de ces courbes. Une comparaison des résultats obtenus est faite par rapport à l'état de l'art qui de nos jours est le nouveau standard de compression d'images JPEG2000. Nous abordons aussi le problème de déregistration et de son impact sur le niveau des performances des transformations utilisées.

3.1 Description de la base de données d'images

Nous disposons d'une base de données constituée d'images multi spectrales dont la plupart nous ont été gracieusement fournies par le CNES[1] et qui pour certaines, représentent des portions de villes de France. Nous disposons également de portions d'images de plusieurs villes du monde que nous avons extraites sous Matlab dans le Toolbox Image. Les images hyper spectrales proviennent pour certaines du site web de la NASA[2] et ont été obtenues gratuitement, certaines grâce à l'apport du CNES. Les quatre images extraites sous Matlab sont celles de Rio, Paris, Colorado et Tokyo. Ce sont des images Landsat qui contiennent six composantes mono-chromatiques. Elles ont pour résolution $N_L = N_C = 512$ et ont une profondeur de 8 bpp. Nous résumons les caractéristiques de ces images dans les tableaux 3.1, 3.2 et 3.3.

Outre ces images, nous avons aussi testé les performances des systèmes de compression sur des images naturelles connues et souvent utilisées dans la communauté de la compression d'image. Il s'agit des images couleur Lena, Mandrill et Peppers qui sont toutes de taille 512×512 et constituées des trois composantes rouge, vert et bleu.

Les images multi spectrales des villes de Toulouse et Montpellier ont été décalées de respectivement $0,2$, $0,35$ et $0,5$ pixel par rapport à l'image de référence, nous nous servirons de ces

[1] Centre National d'Etudes Spatiales de Toulouse
[2] National Aeronautics and Space Administration

Nom image	Nombre composantes	Résolution en bpp	Nombre lignes	Nombre colonnes
Moissac	4	12	3152	320
Strasbourg	4	12	3272	352
Vannes	4	12	3736	352
Port_debouc	4	12	1376	320
Toulouse1	3	8	1024	1024
Toulouse2	3	8	1024	1024
Fort_France	6	8	1024	1024

TAB. 3.1 – Images multi spectrales fournies par le CNES

Nom image	Nombre composantes	Résolution en bpp	Nombre lignes	Nombre colonnes
Montpellier_ref	4	12	2448	296
Toulouse_ref	4	12	3672	352

TAB. 3.2 – Images multi spectrales de référence fournies par le CNES

Nom image	Nombre composantes	Résolution en bpp	Nombre lignes	Nombre colonnes
Hyper	224	16	512	614
Hyper1	224	16	512	614
Hawai	224	12	512	680
Hawai2	224	12	512	680
Maine	224	12	512	680
Maine2	224	12	512	680
Jasper	224	16	512	512
Moffett	224	16	512	512
Cuprite	224	16	512	512

TAB. 3.3 – Images hyperspectrales AVIRIS fournies par la NASA

images pour étudier l'impact de la déregistration et du décalage des images sur les performances. Les images décalées ont les mêmes caractéristiques que celles de référence à savoir : même résolution, même profondeur et même nombre de composantes. Ces images sont des simulations d'images PLEIADES[3]. Les images Toulouse1 et Toulouse2 quant à elles sont des images SPOT5[4].

Les images hyper spectrales AVIRIS[5] sont généralement de dimensions $224 \times 614 \times 512$. Certaines ont été tronquées en $224 \times 512 \times 512$ (Jasper, Moffett et Cuprite) pour permettre une comparaison des performances avec les autres partenaires travaillant dans le cadre du projet ACI^{2}M dont fait partie cette thèse. Les autres images hyper spectrales utilisées sont constituées de 224 bandes spectrales, chacune ayant une résolution de 680×512 et une amplitude de 12 bpp. En Annexe2, nous illustrons une composante de quelques-unes de ces images. Soulignons que seuls les résultats des images hyper-spectrales de Moffett, Jasper et Cuprite seront présentés dans ce rapport, les simulations n'ayant pas été réalisées d'une façon complète pour les autres images.

3.2 Validation des formules d'approximation de la distorsion

Pour évaluer la validité des formules d'approximation de la distorsion que nous avons obtenues au premier chapitre, nous avons réalisé la simulation simple suivante : après avoir appliqué la transformation (spatiale et spectrale) sur l'image, nous affectons à chaque sous-bande de chaque composante, un pas de quantification de façon aléatoire, mais de manière à rester sous les hypothèses hautes résolutions (distorsions relativement faibles). Puis nous évaluons d'une part la distorsion, en la calculant grâce aux formules d'approximations par rapport aux distorsions des sous-bandes des différentes composantes, et d'autre part, nous évaluons la distorsion réelle en appliquant les transformations inverses et en reconstruisant une image approchée. Appelons D_{reel} la distorsion réelle ainsi obtenue et D_{app} la distorsion obtenue en utilisant la formule d'approximation. Le PRSB est ensuite obtenu à partir de la distorsion réelle, et nous appelons **distorsion relative**, le rapport $D_r = \frac{D_{app}}{D_{reel}}$. En faisant plusieurs tirages aléatoires pour différents pas de quantification, nous obtenons un ensemble de points pour D_{reel} et D_{app}, ce qui nous permet de tracer le nuage de points D_r en fonction du PRSB.

Nous avons choisi sans aucune raison particulière, de présenter les résultats obtenus avec l'image de Vannes pour les différents cas suivant :

- Le schema séparable .
- Le schéma en sous-bandes.
- Le schéma mixte en sous-bandes avec $P = 4$.

Les transformations étudiées sont l'identité, la TKL, *Ica_orth* et *Ica_opt*. Pour chacun des cas, nous avons mis le nuage de points des différentes transformations sur une seule figure. L'erreur relative moyenne pour chacune des transformations s'obtient en calculant la moyenne arithmétique des erreurs d'approximation. L'erreur d'approximation ϵ étant ici définie par $\epsilon = |1 - D_r|$. En effet, la valeur idéale de la distorsion relative vaut 1 c'est-à-dire lorsque l'approximation de la distorsion coincide avec la distorsion réelle.

[3]http ://smsc.cnes.fr/PLEIADES/Fr/
[4]http ://smsc.cnes.fr/SPOT/Fr/
[5]http ://aviris.jpl.nasa.gov/html/aviris.freedata.html

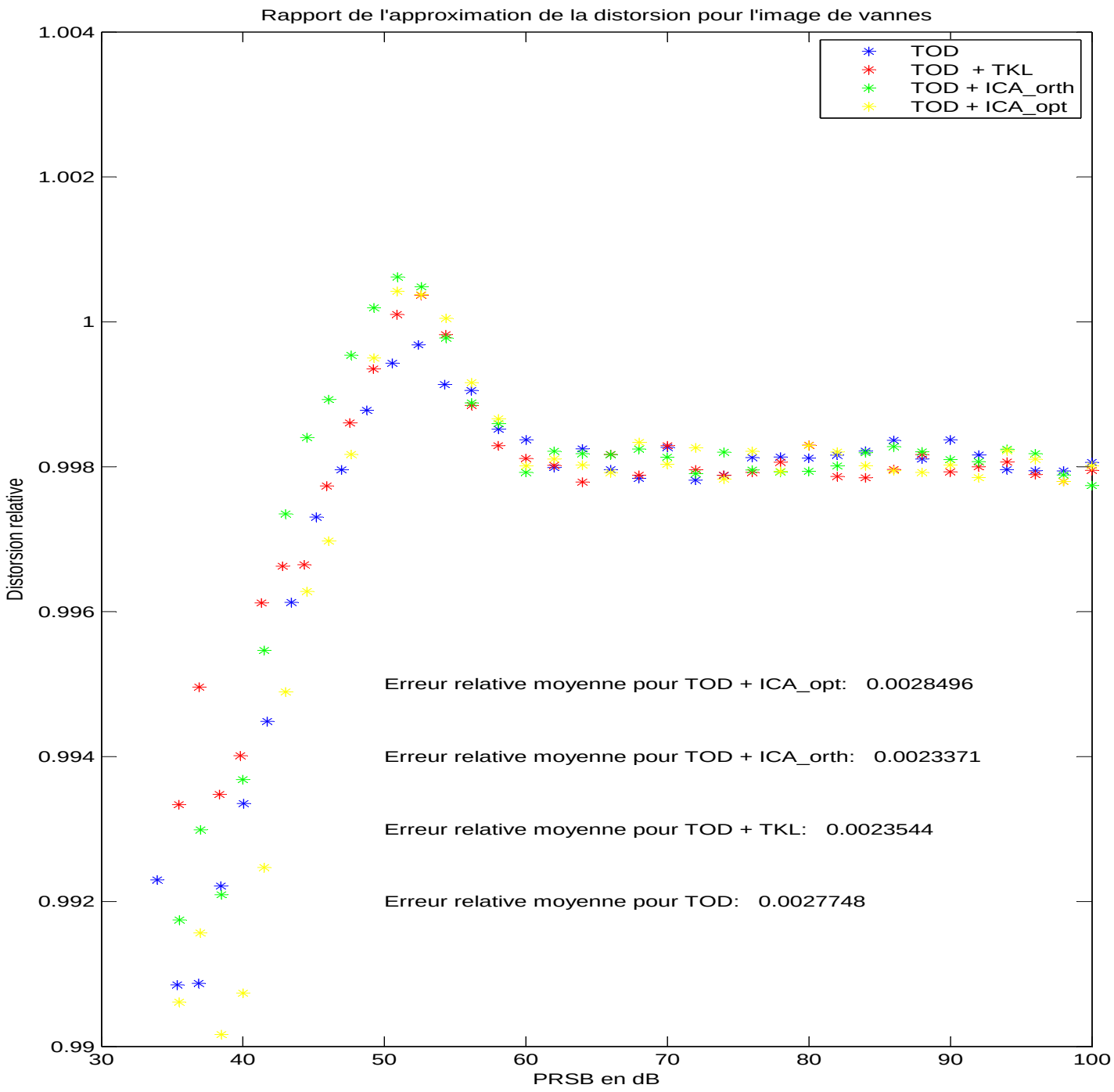


FIG. 3.1 – Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma en sous-bandes

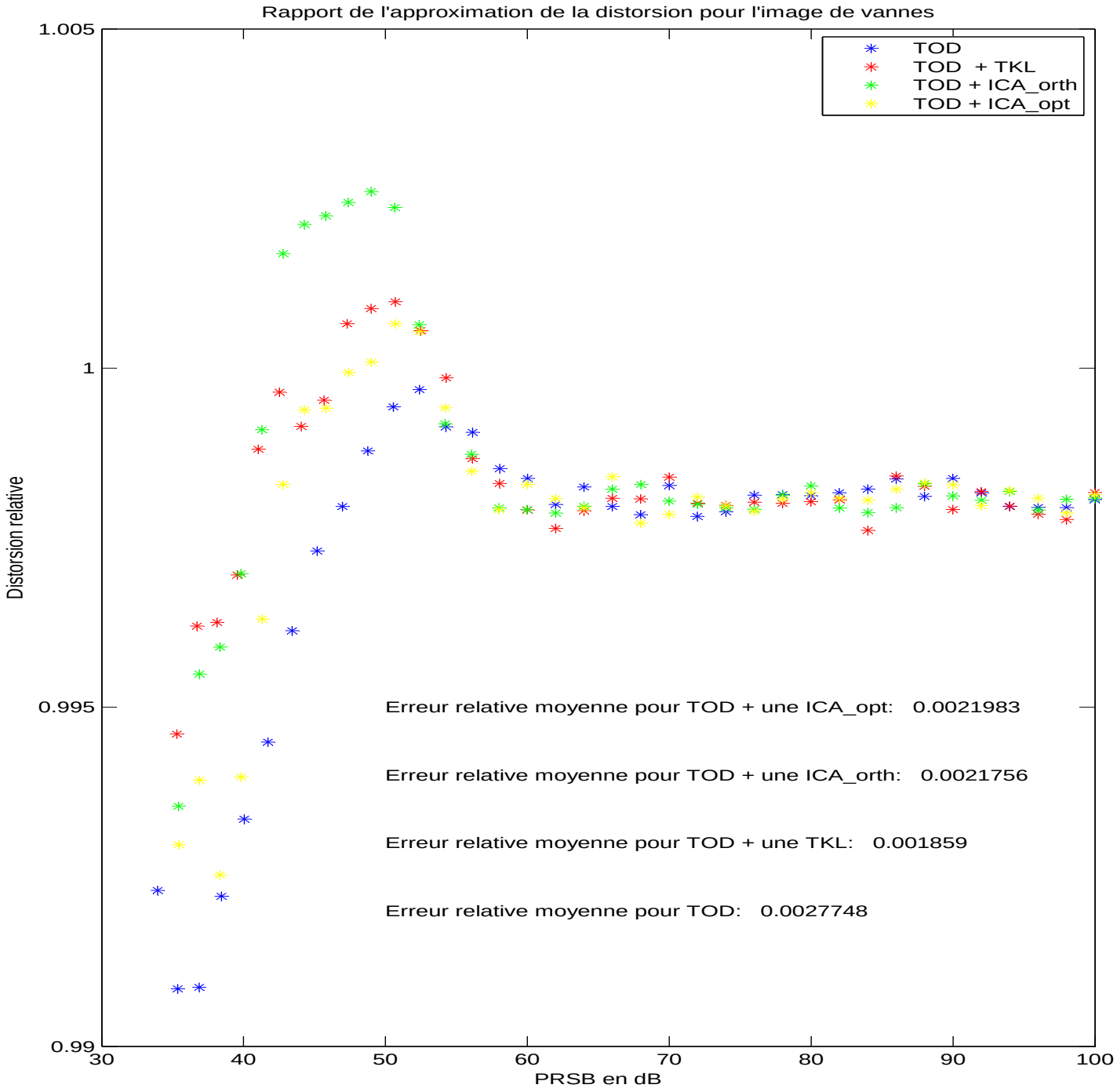


FIG. 3.2 – Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma séparable avec adaptation dans l'espace ondelette

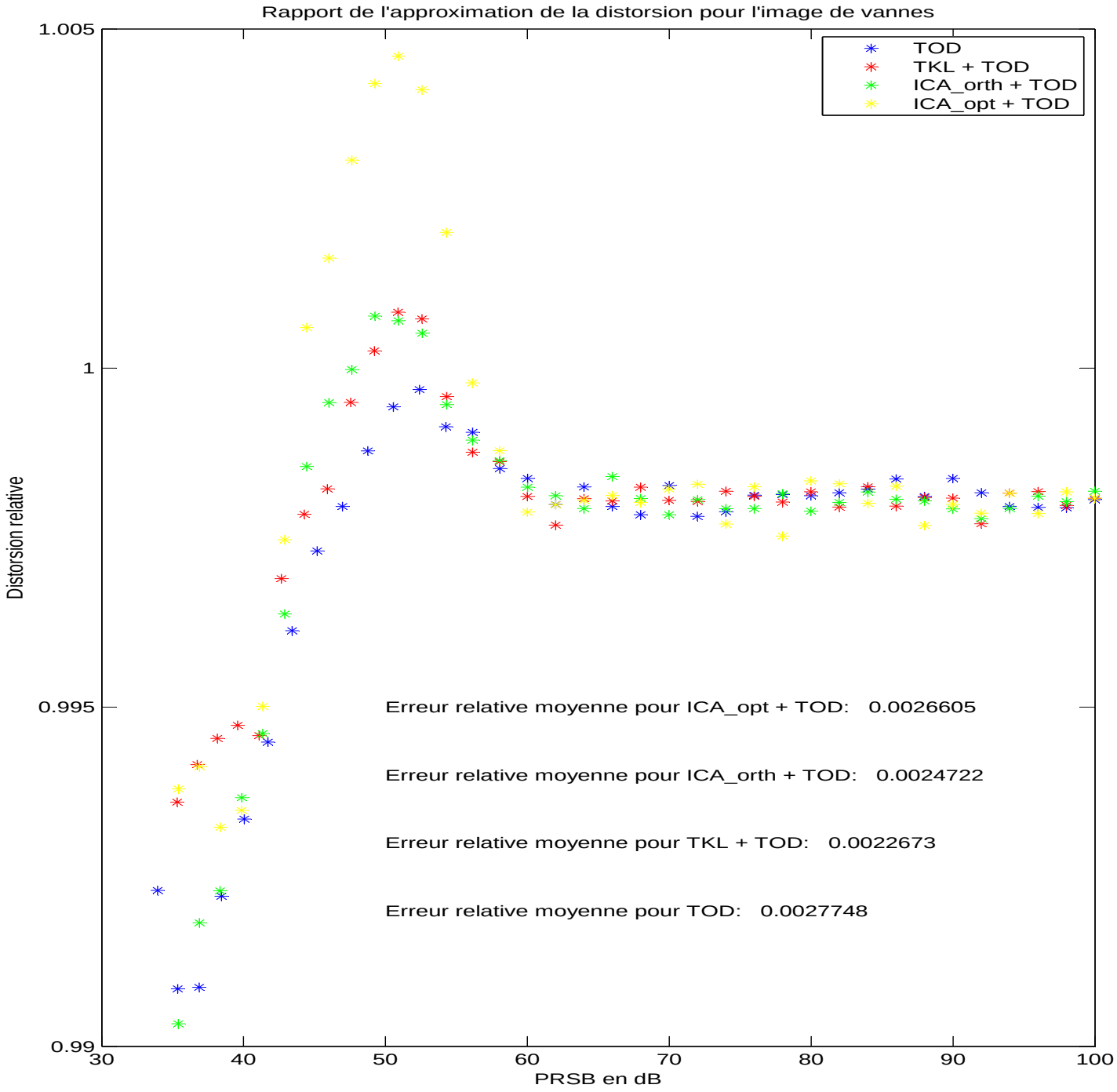


FIG. 3.3 – Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma séparable avec adaptation dans l'espace image

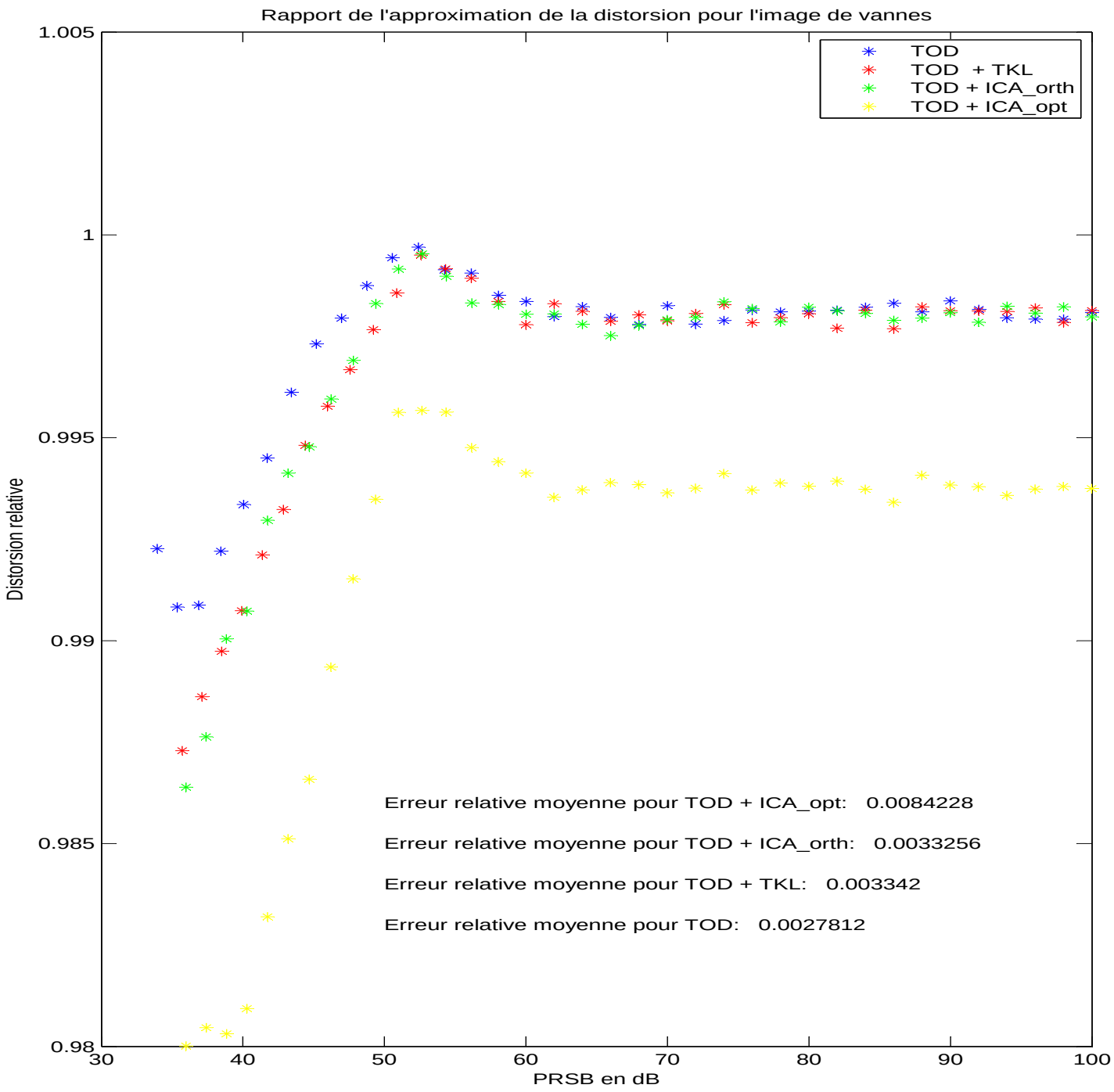


FIG. 3.4 – Distorsion relative en fonction du PRSB de l'image Vannes : cas du schéma mixte en sous-bandes avec $P = 4$

En observant les différentes figures 3.1 à 3.4, on se rend compte que la formule d'approximation permet d'avoir une très bonne estimation de la distorsion : l'erreur moyenne d'approximation dans les différents cas de figure reste sensiblement très faible. Ces formules d'approximation de la distorsion permettent de valider le fait que la transformée en ondelette utilisée, à savoir la 9/7 de Daubechies, est quasi orthogonale dans la mesure où cette hypothèse est faite pour estimer la distorsion en utilisant ces formules dans le cas du schéma mixte en sous-bandes. Par conséquent, il devient légitime d'utiliser ces formules d'approximation pour avoir une bonne estimation de la distorsion lors de la phase d'allocation optimale de bits. Par la suite, en utilisant ces formules d'approximation, nous évaluons les performances des différents schémas de compression définis au premier chapitre et nous en déduisons les gains de codage généralisés associés aux différentes transformations étudiées.

3.3 Évaluation des gains de codage généralisés

Nous avons évalué les gains de codage généralisés des images multi-spectrales pour le schéma de compression séparable, le schéma en sous-bandes et le schéma mixte en sous-bandes. Les transformations spectrales étudiées ici sont *Ica_opt*, *Ica_orth* et la TKL. Comme transformation de référence, nous avons choisi d'une part la transformation Identité, et d'autre part la TKL dont les performances sont comparées avec celles des nouvelles transformations à base d'ACI.

Pour le cas particulier des images hyper-spectrales qui sont constituées de 224 bandes spectrales, pour des problèmes liés à la complexité de calcul, nous avons dans certains cas, subdivisé l'image en blocs de taille $N_1 = 20$ ou $N_1 = 45$ sur le plan spectral[6], puis nous avons évalué les courbes débit distorsion sur chaque bloc, avant de moyenner sur les différents blocs pour avoir la courbe debit distorsion globale de l'image. Pour le cas $N_1 = 20$, le dernier bloc contient $N_2 = 24$ bandes spectrales. Pour le cas $N_1 = 45$, le dernier bloc contient $N_2 = 44$ bandes spectrales. Soit $R_c(k)$, $(1 \leq k \leq K)$ un ensemble fini de débits fixés, et N_b le nombre de blocs de l'image qui a été découpée sur le plan spectral. Notons $(R_c(k), D_i(k))$ les points de la courbe débit distorsion du $i^{ème}$ bloc de l'image hyper spectrale, les images étant découpées en blocs sur le plan spectral. La courbe débit vs distorsion globale de l'image hyper-spectrale peut être obtenue par les points $(R_c(k), D_g(k))$, avec

$$D_g(k) = \sum_i \alpha_i D_i(k) \tag{3.1}$$

où $\alpha_i = N_i/N$, N étant le nombre total de bandes spectrales de l'image, et N_i étant le nombre de bandes spectrales du $i^{ème}$ bloc. Cette méthode utilisant une moyenne pondérée pour avoir la distorsion globale est loin d'etre optimale dans la mesure où la plage des valeurs des débits $R_c(k)$ varie considérablement d'un bloc à un autre, les écarts se situant entre les valeurs les plus basses et les plus hautes. Cela étant du au fait que les différents blocs ont des caractéristiques différentes aussi bien en terme d'énergie qu'en termes de détails contenus dans les composantes d'un bloc. Dans [9] nous avons présenté des résultats utilisant cette méthode d'allocation pour obtenir la courbe débit vs PRSB finale des images hyper-spectrales Hyper et Hyper1. Dans cet article, l'allocation optimale de débit est faite pour tous les débits en utilisant les approximations asymptotiques qui peuvent être mises en défaut lorsque l'on s'intéresse à des faibles débits. Une autre méthode très intéressante consiste à utiliser l'algorithme de Shoham et Gersho [69] qui est une méthode utilisant des multiplicateurs lagrangiens sur un ensemble fini de points. Disposant des points $(R_c(k), D_i(k))$

[6]Pour les images hyper spectrales, un bloc désigne un groupe de N_1 bandes spectrales contiguës et n'a donc rien à voir avec le découpage par blocs de P coefficients dans une composante.

de chaque bloc de l'image hyper-spectrale, nous appliquons l'algorithme de Shoham et Gersho sur cet ensemble fini de points, en utilisant la formule de l'équation (3.1) pour la distorsion globale moyenne. Le debit global moyen lui est estimé par la formule suivante

$$R_g(k) = \sum_i \alpha_i R_i(k) \tag{3.2}$$

où $R_i(k)$ est le débit du $i^{ème}$ bloc de l'image découpée en bloc sur son axe spectral. C'est cette dernière méthode que nous avons retenue pour obtenir la courbe débit vs distorsion globale d'une image hyper-spectrale, les résultats des simulations qui seront présentés ont été obtenues suivant cette méthode.

Pour la partie codage, nous avons dans un premier temps utilisé un estimateur de l'entropie d'ordre 1 pour calculer le débit, puis nous avons utilisé le $VM9$ (Verification Model version 9) décrit dans [43] et qui utilise le codeur EBCOT de JPEG2000. L'utilisation d'un estimateur de l'entropie d'ordre 1 pour avoir le débit se justifie par le fait que des codeurs simples comme le codeur de Huffman ou encore le codeur arithmétique qui offre des performances meilleures que le premier cité, permettent d'atteindre des performances (en terme de débit obtenu après le codage) très proches de cette estimation. Dans l'annexe2, nous explicitons le calcul de cet entropie d'ordre 1 pour une image. Afin de mieux illustrer les résultats, notamment les gains de codage généralisés, nous traçons d'abord pour une image donnée, les courbes débit vs distorsion et débit vs PRSB, et de cette dernière courbe, nous déduisons les gains de codage généralisés qui ont été définis dans le deuxième chapitre. Nous montrons ces courbes pour quelques images multi-composantes, et pour les autres images, les valeurs des points de la courbe débit vs PRSB sont résumées dans des tableaux, pour des débits généralement utilisés dans la communauté de la compression des images. Ces tableaux se trouvent pour la plupart dans l'Annexe1 de ce rapport.

3.3.1 Résultats obtenus en utilisant un estimateur de l'entropie

Nous avons choisi de présenter les résultats en termes de courbe débit vs PRSB pour une partie des images, les résultats obtenus sur les autres images sont résumés dans des tableaux récapitulatifs. Nous avons choisi de présenter sans aucune raison particulière les résultats pour les images multi-spectrales de Vannes et Moissac, ainsi que pour les images de références et les images décalées de Montpellier. Pour ce qui est des images hyper-spectrales, les résultats seront présentés pour l'image de Jasper notamment pour le cas du schéma en sous-bandes. Nous montrons les courbes débit vs PRSB des différentes transformations étudiées, ainsi que la courbe donnant le gain de codage généralisé en fonction du débit. Le gain de codage généralisé est évalué par rapport à la transformation identité prise comme transformation étalon d'une part sur un premier graphe et par rapport à la TKL (prise comme transformation étalon) d'autre part dans un autre graphe. Nous présentons les courbes du PRSB en fonction du débit d'une part parce qu'elles permettent d'évaluer les performances des différents systèmes de compression étudiés, et d'autre part, nous montrons le gain de codage généralisé pour avoir une idée en dB du gain que l'on ferait en utilisant une transformation spectrale par rapport à une autre. Notre objectif ici étant en partie de comparer les performances des transformations spectrales à base d'ACI à celles de la TKL. Toutefois, comme nous l'avons démontré au chapitre 2, les transformations obtenues par la algorithmes Ica_orth et Ica_opt ne sont pas optimales dans le cas du schéma séparable.

3.3.1.1 Cas du schéma en sous-bandes

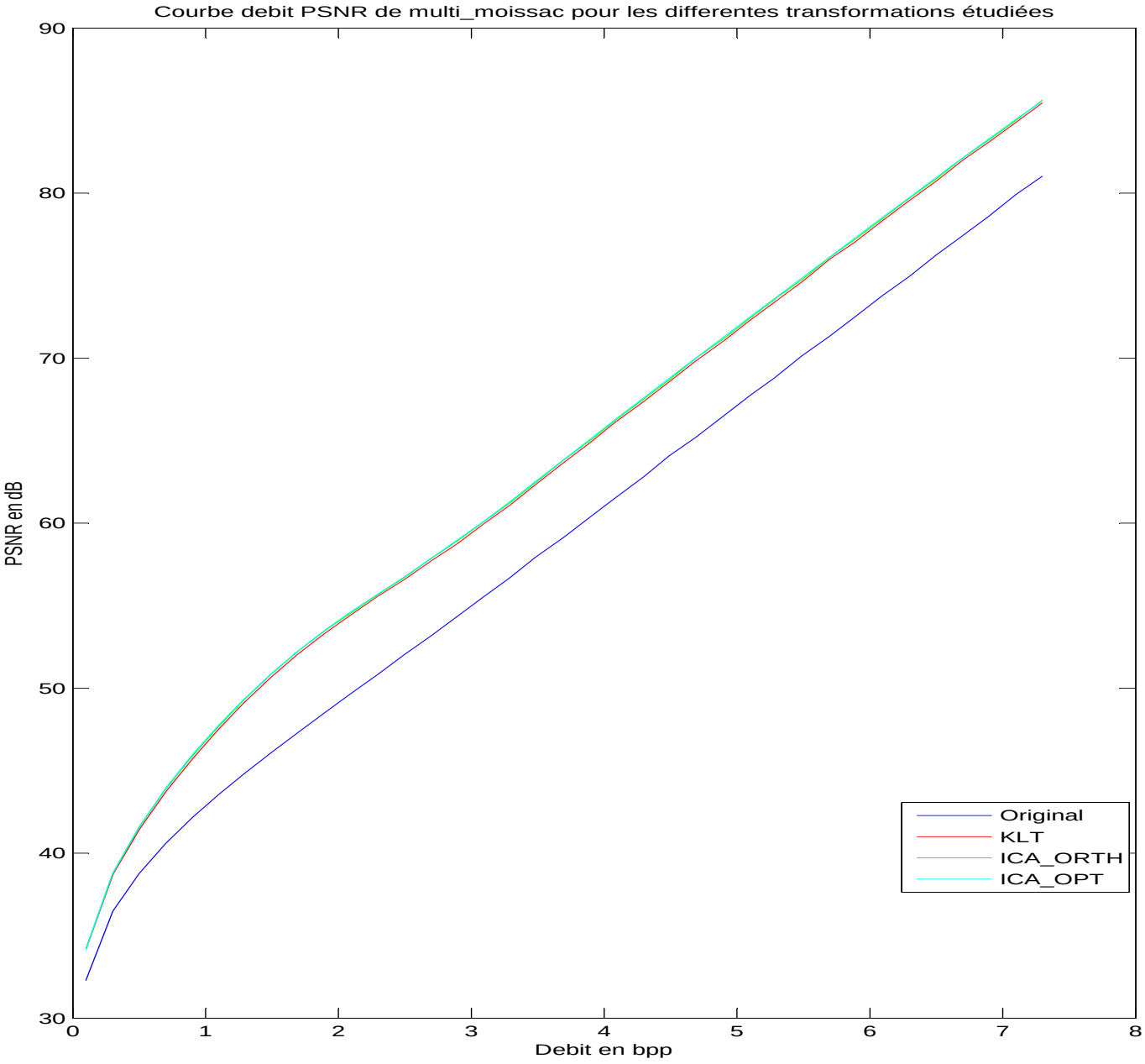
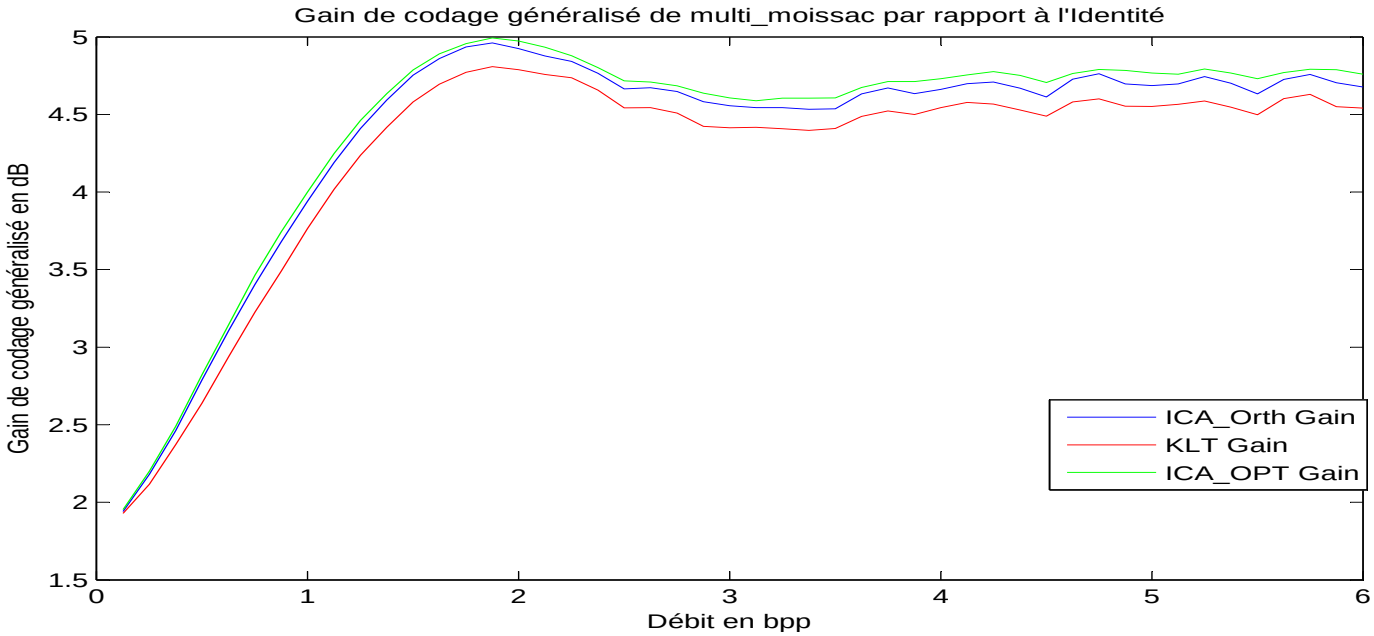


FIG. 3.5 – Courbe débit vs PRSB de l'image Moissac : cas du schéma en sous-bandes



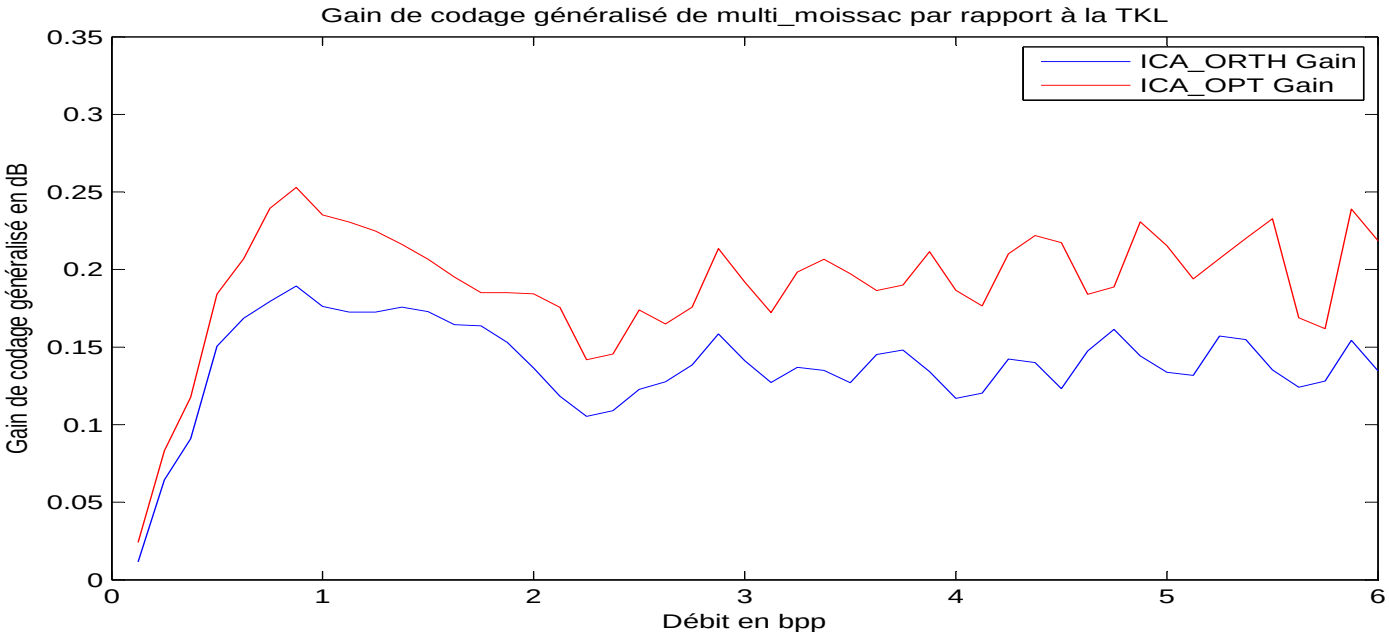


FIG. 3.6 – Gain de codage généralisé de l'image Moissac : cas du schéma en sous-bandes

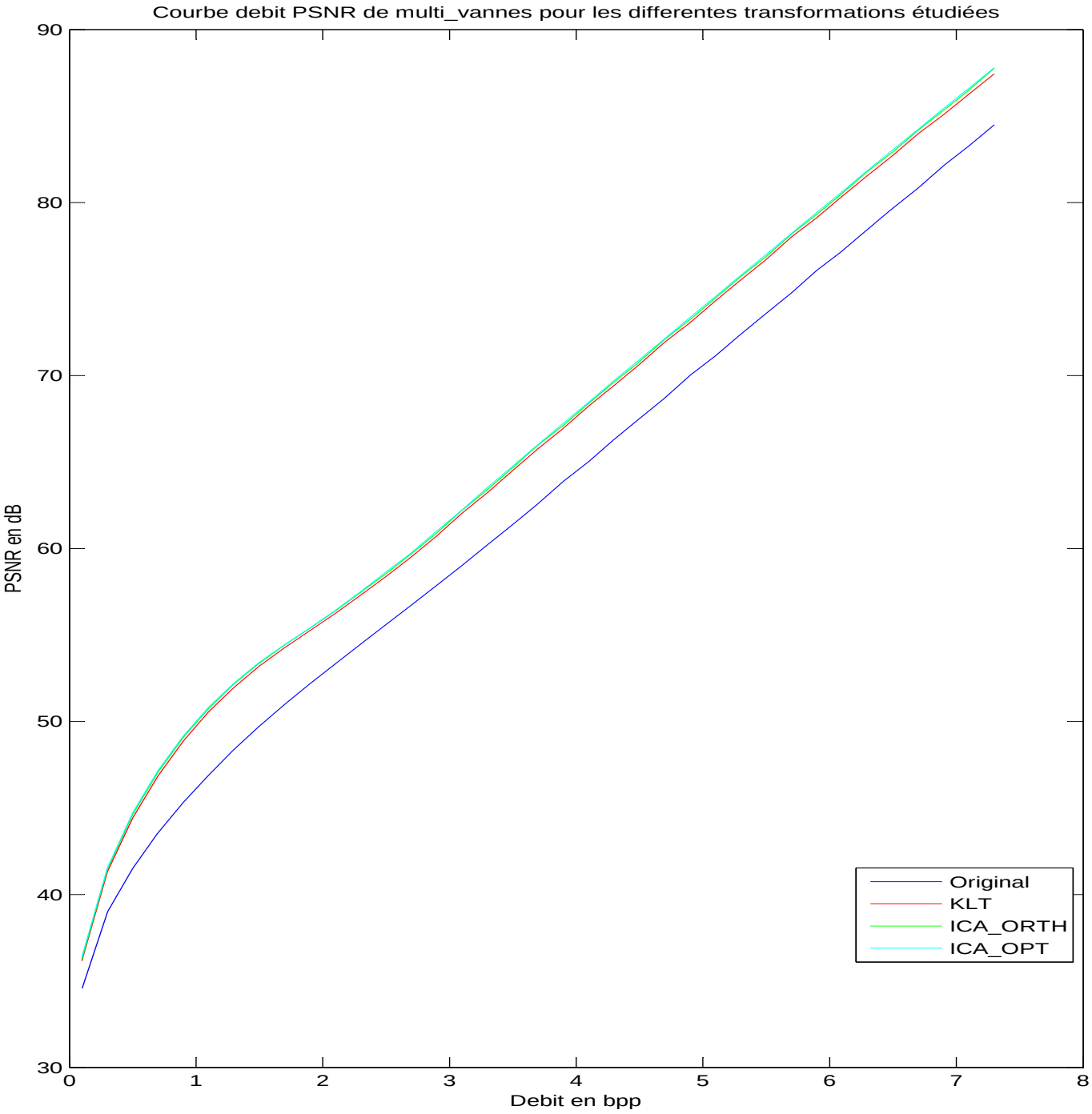
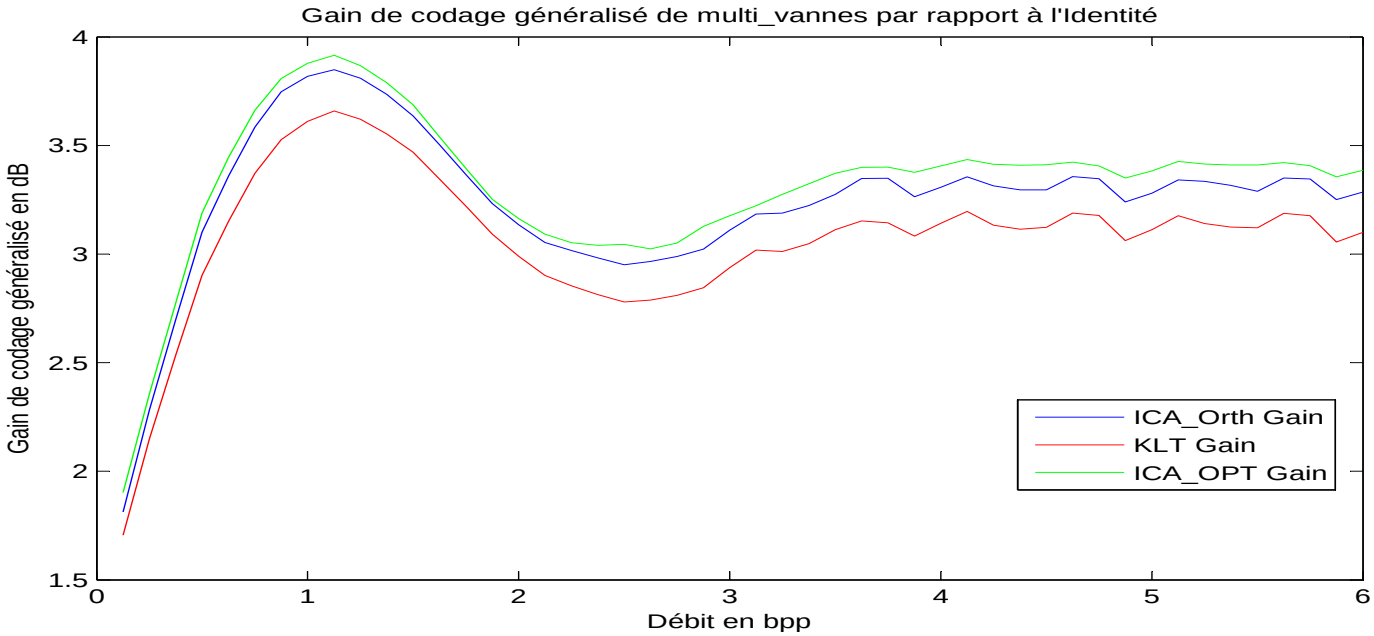


FIG. 3.7 – Courbe débit vs PRSB de l'image Vannes : cas du schéma en sous-bandes



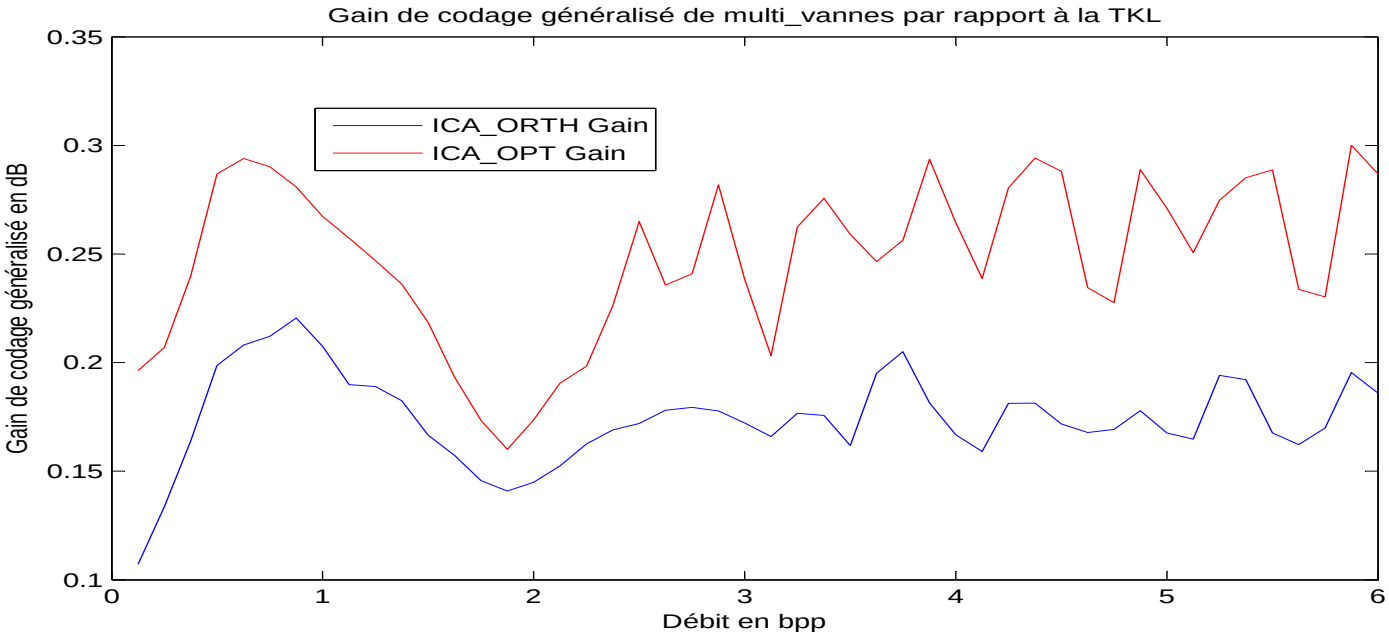


FIG. 3.8 – Gain de codage généralisé de l'image Vannes : cas du schéma en sous-bandes

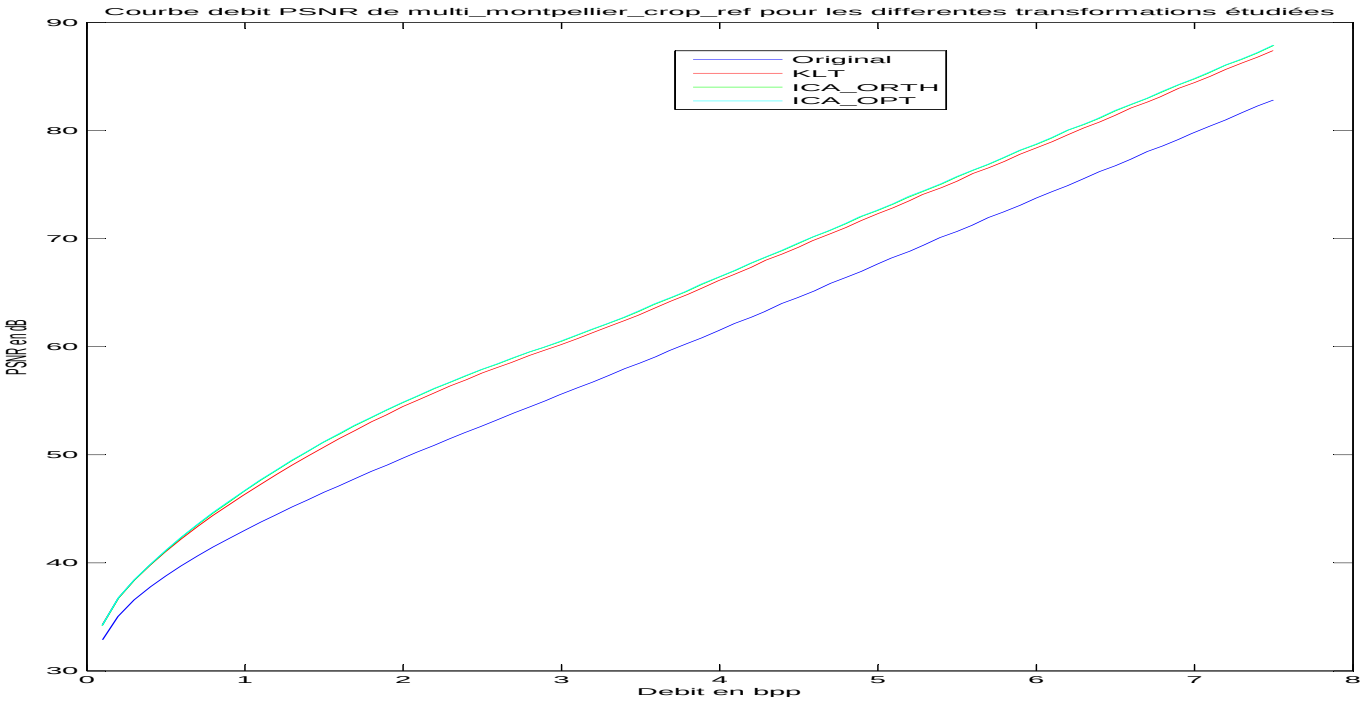


FIG. 3.9 – Courbe débit vs PRSB de l'image de référence de Montpellier : cas du schéma en sous-bandes

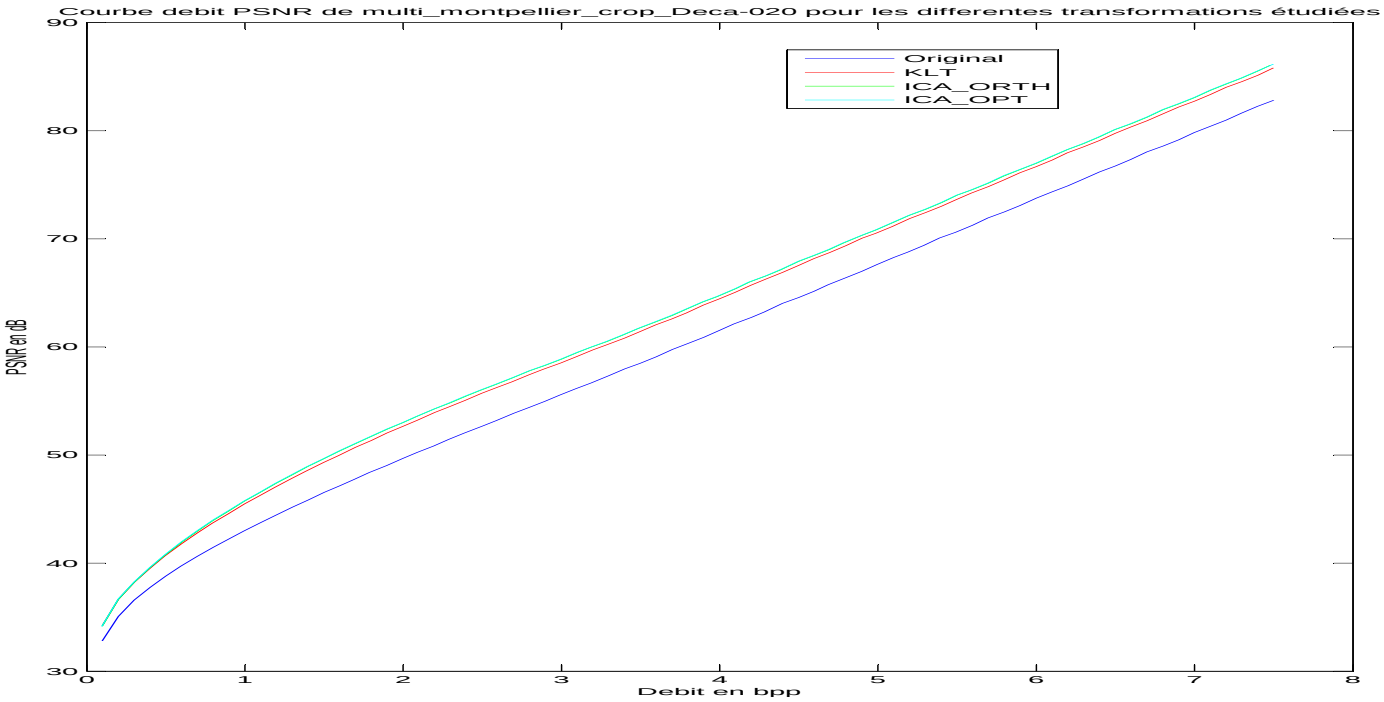


FIG. 3.10 – Courbe débit vs PRSB de Montpellier décalée de 0.2 pixel : cas du schéma en sous-bandes

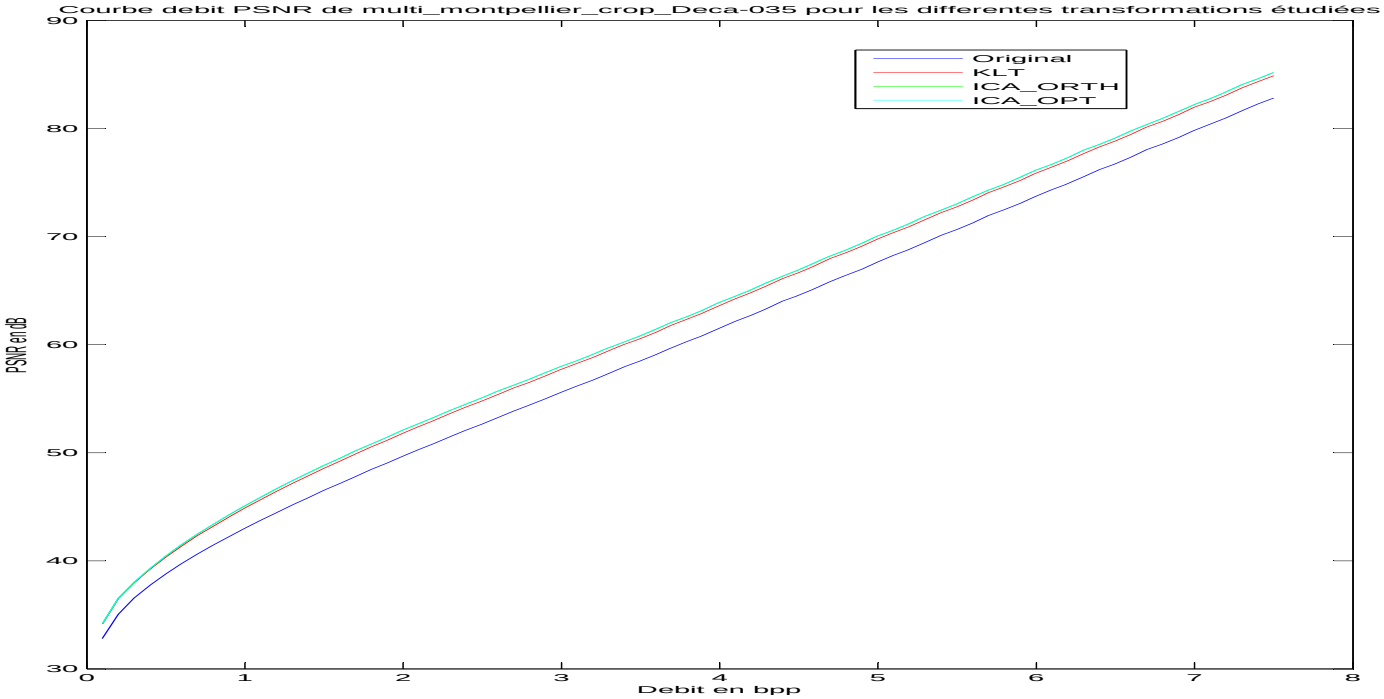


FIG. 3.11 – Courbe débit vs PRSB de Montpellier décalée de 0.35 pixel : cas du schéma en sous-bandes

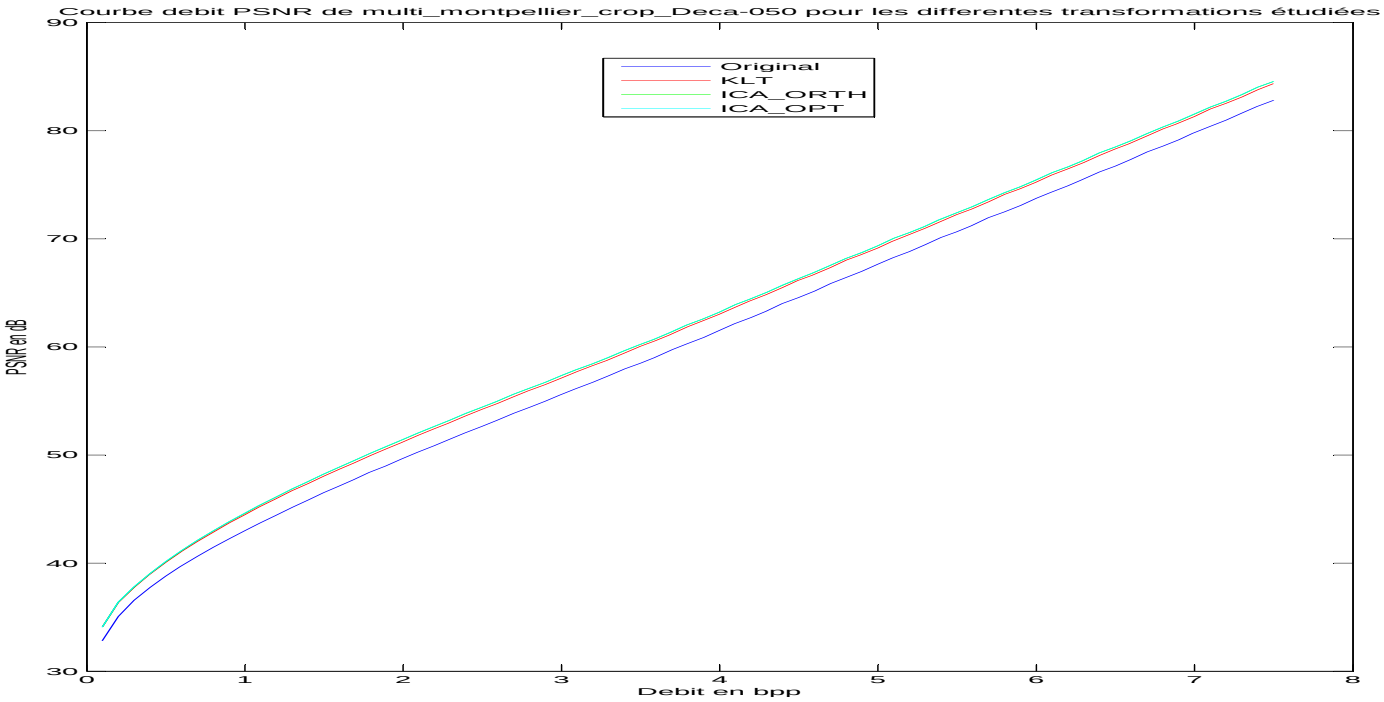


FIG. 3.12 – Courbe débit vs PRSB de Montpellier décalée de 0.5 pixel : cas du schéma en sous-bandes

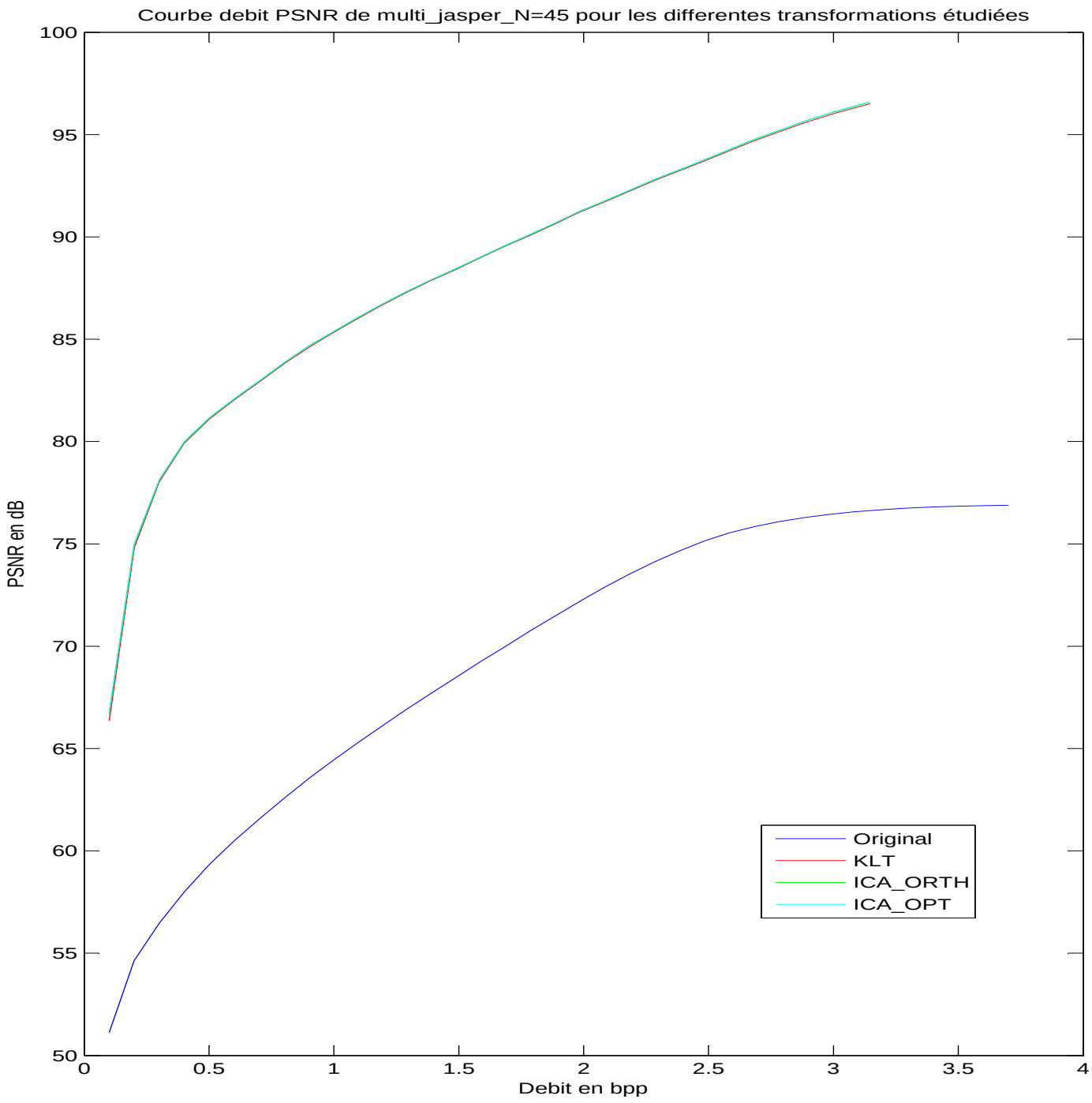
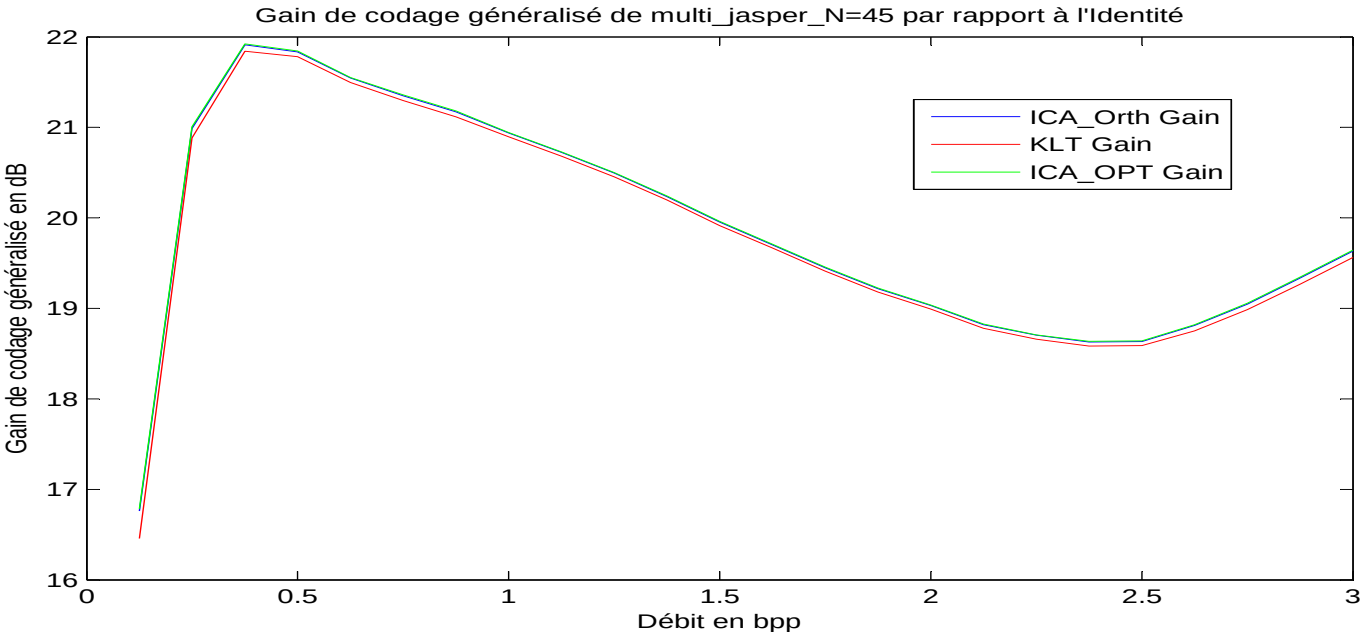


FIG. 3.13 – Courbe débit vs PRSB de l'image hyper-spectrale Jasper avec $N_1 = 45$, $N_2 = 44$: cas du schéma en sous-bandes



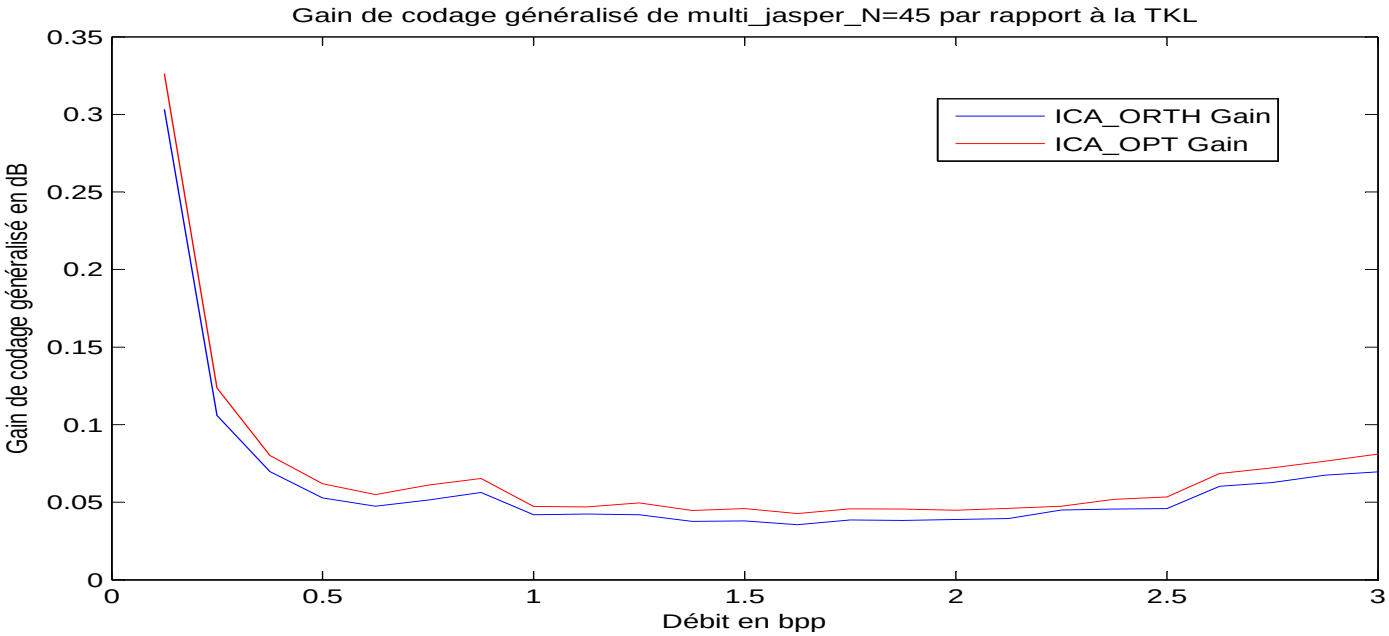


FIG. 3.14 – Gain de codage généralisé de l'image hyper-spectrale Jasper avec $N_1 = 45$, $N_2 = 44$: cas du schéma en sous-bandes

3.3.1.2 Schéma mixte en sous-bandes avec $P = 4$

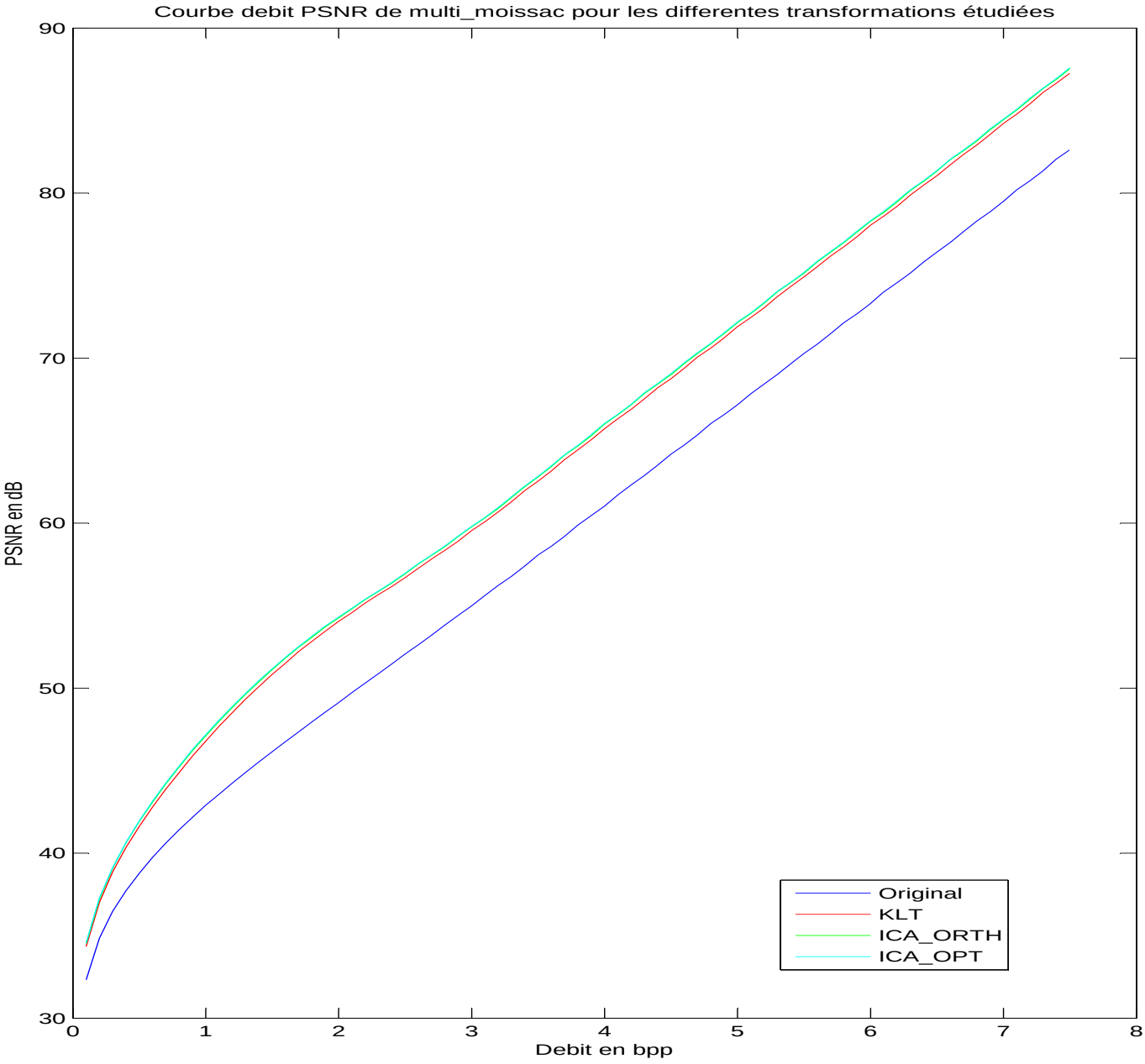
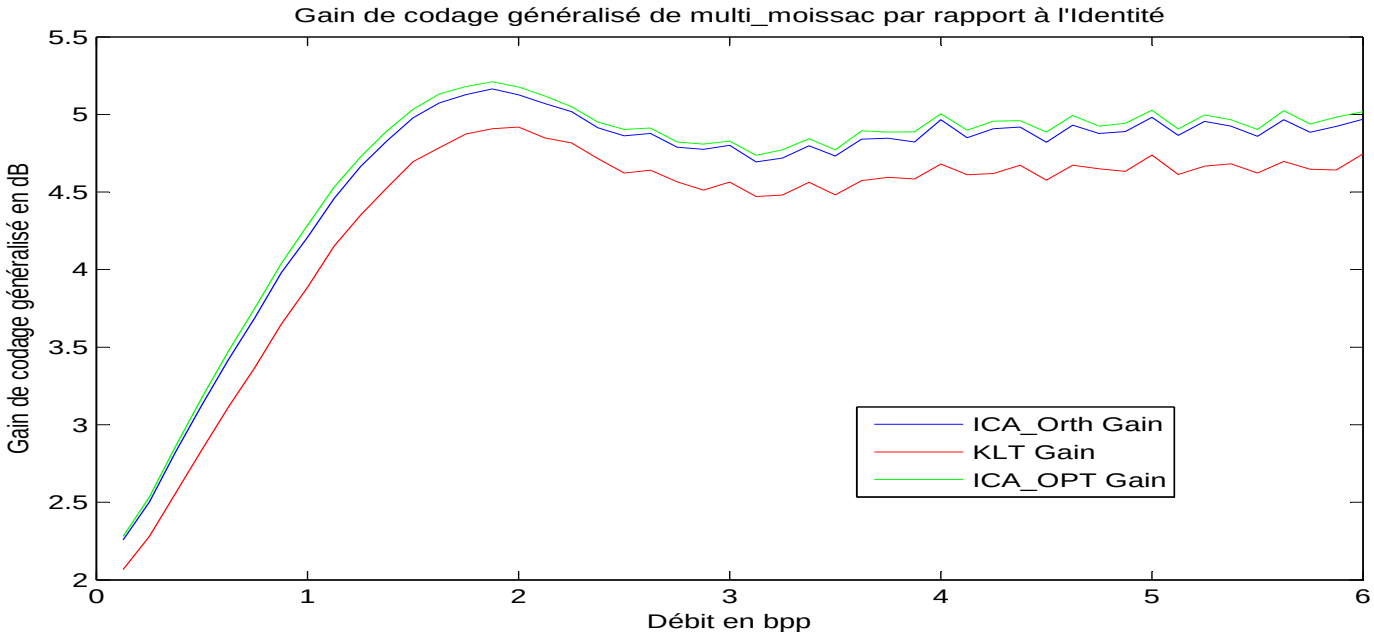


FIG. 3.15 – Courbe débit vs PRSB de l'image Moissac : cas du schéma mixte en sous-bandes avec $P = 4$



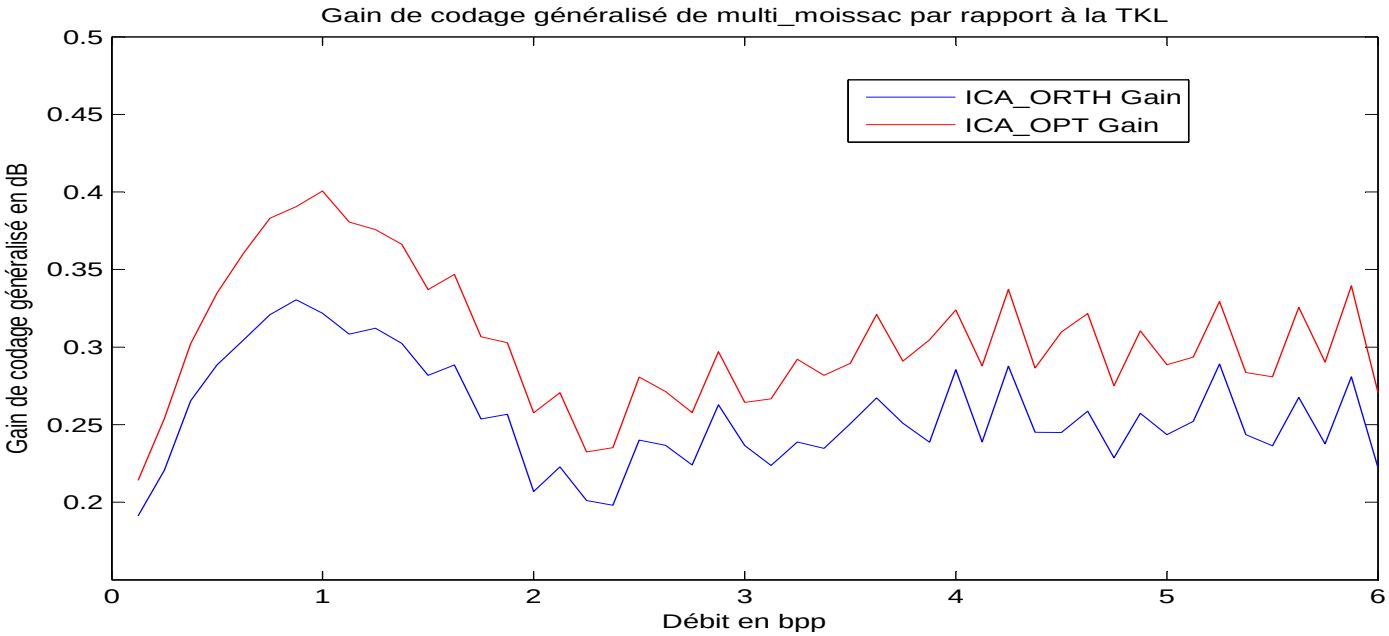


FIG. 3.16 – Gain de codage généralisé de l'image Moissac : cas du schéma mixte en sous-bandes avec $P = 4$

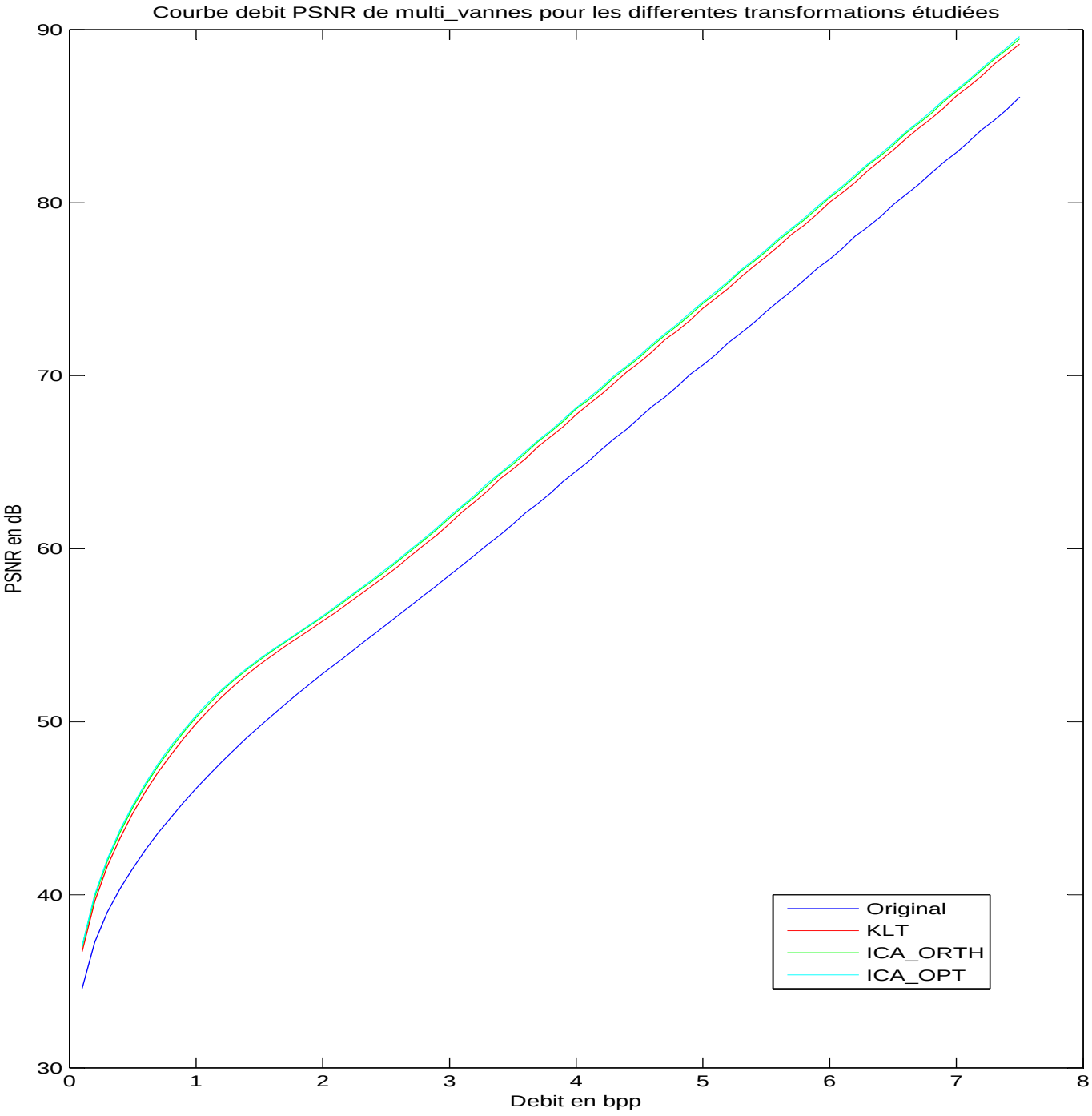
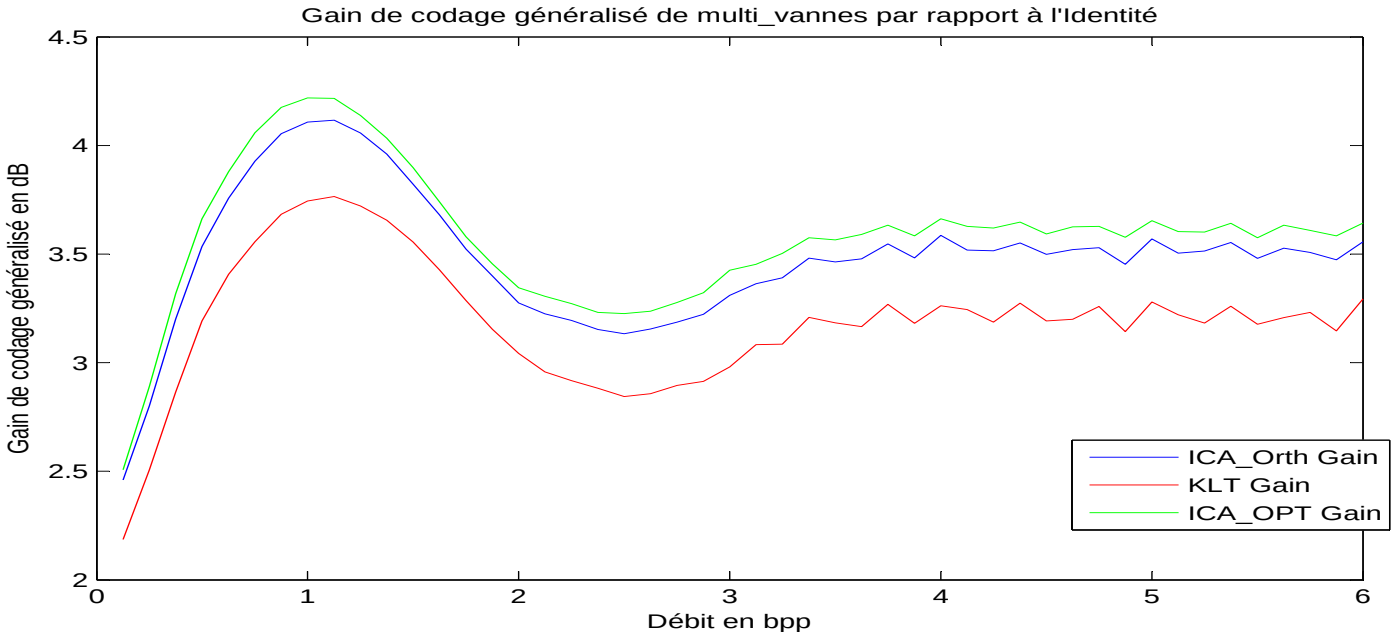


FIG. 3.17 – Courbe débit vs PRSB de l'image Vannes : cas du schéma mixte en sous-bandes avec $P = 4$



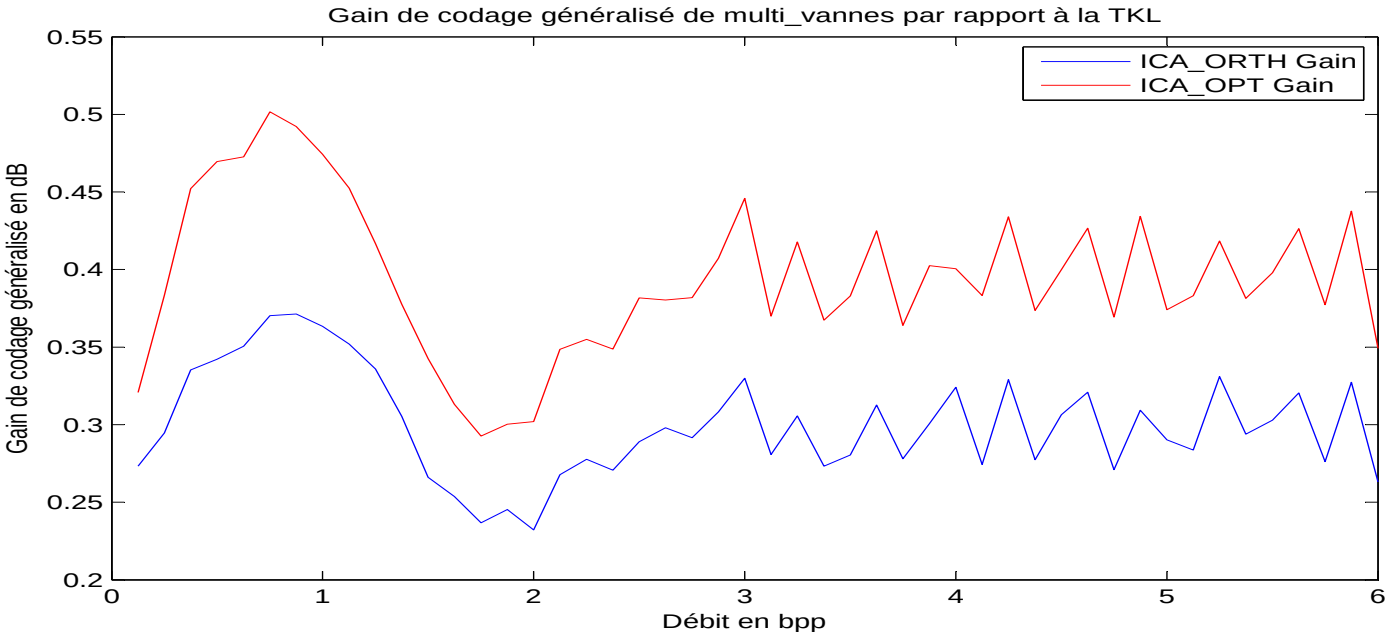


FIG. 3.18 – Gain de codage généralisé de l'image Vannes : cas du schéma mixte en sous-bandes avec $P = 4$

3.3.1.3 Schema séparable avec adaptation sur l'espace ondelette

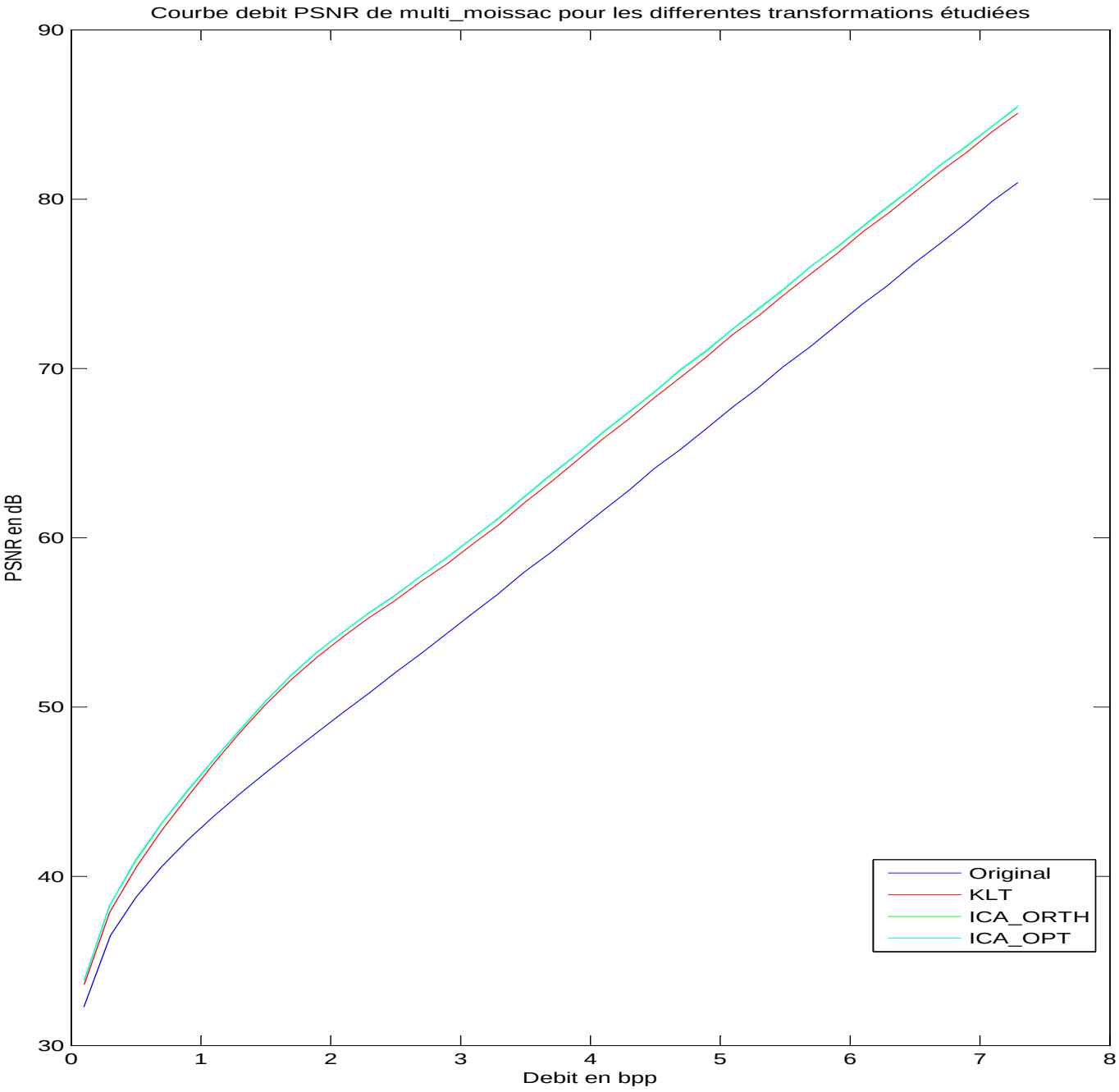
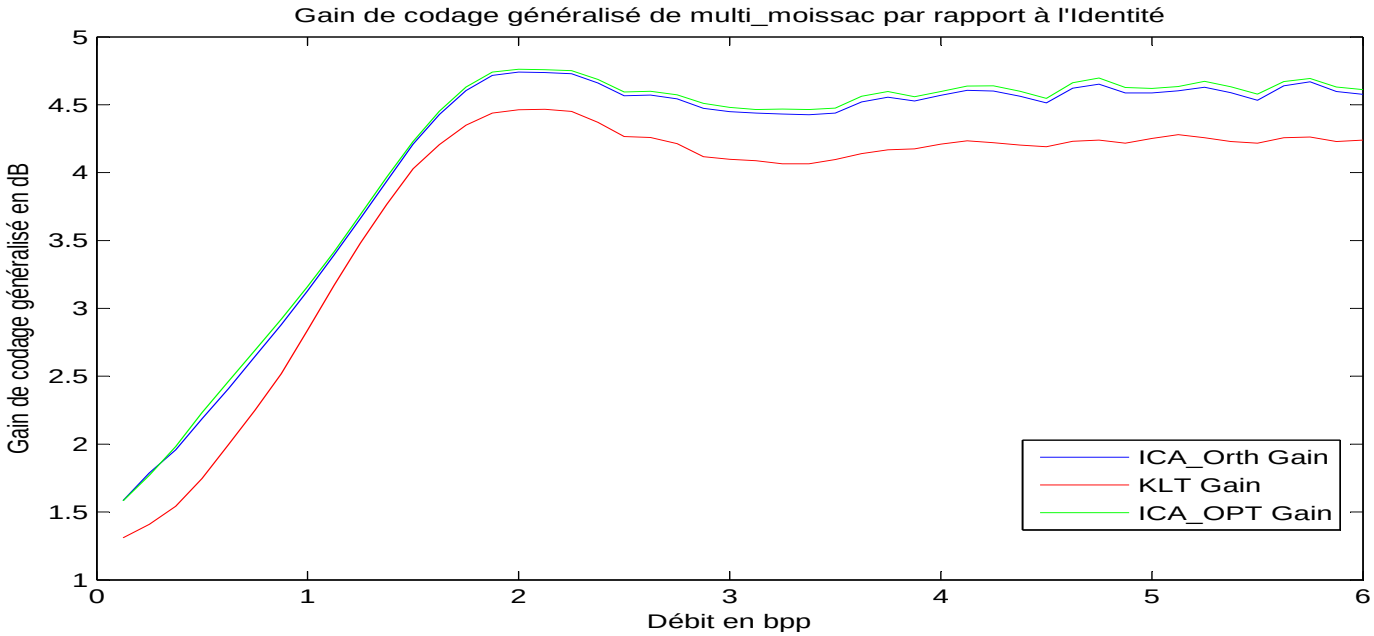


FIG. 3.19 – Courbe débit vs PRSB de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace ondelette



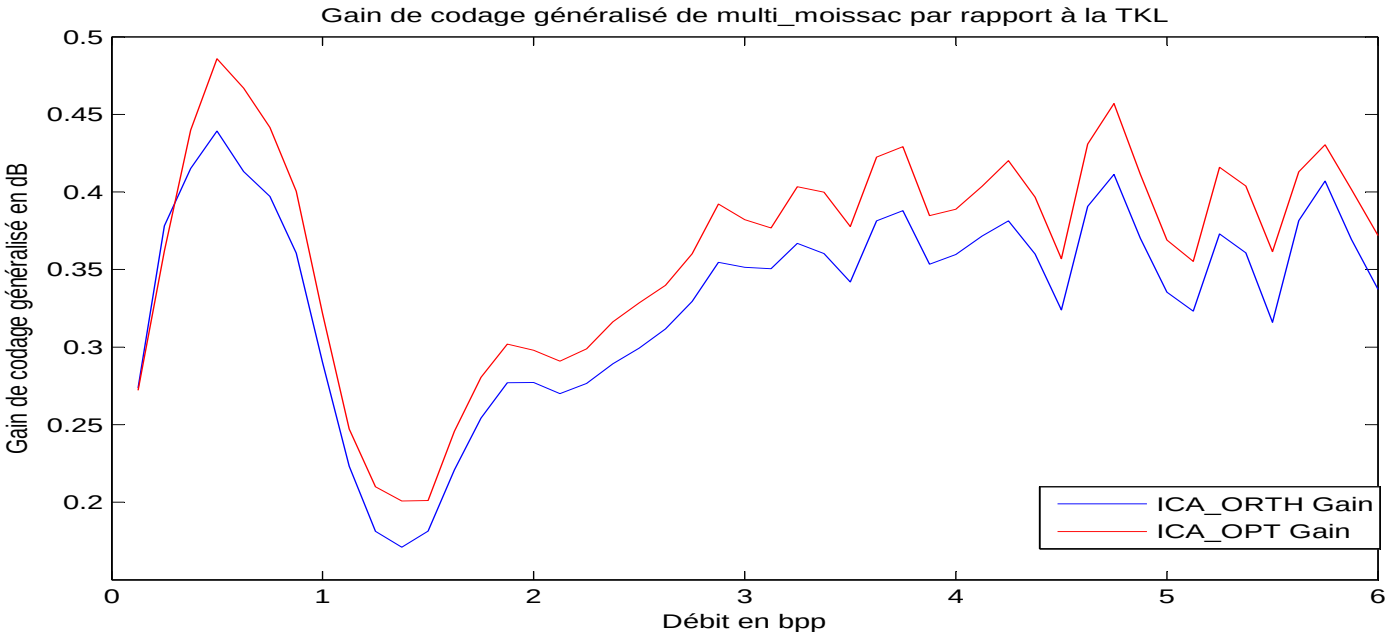


FIG. 3.20 – Gain de codage généralisé de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace ondelette

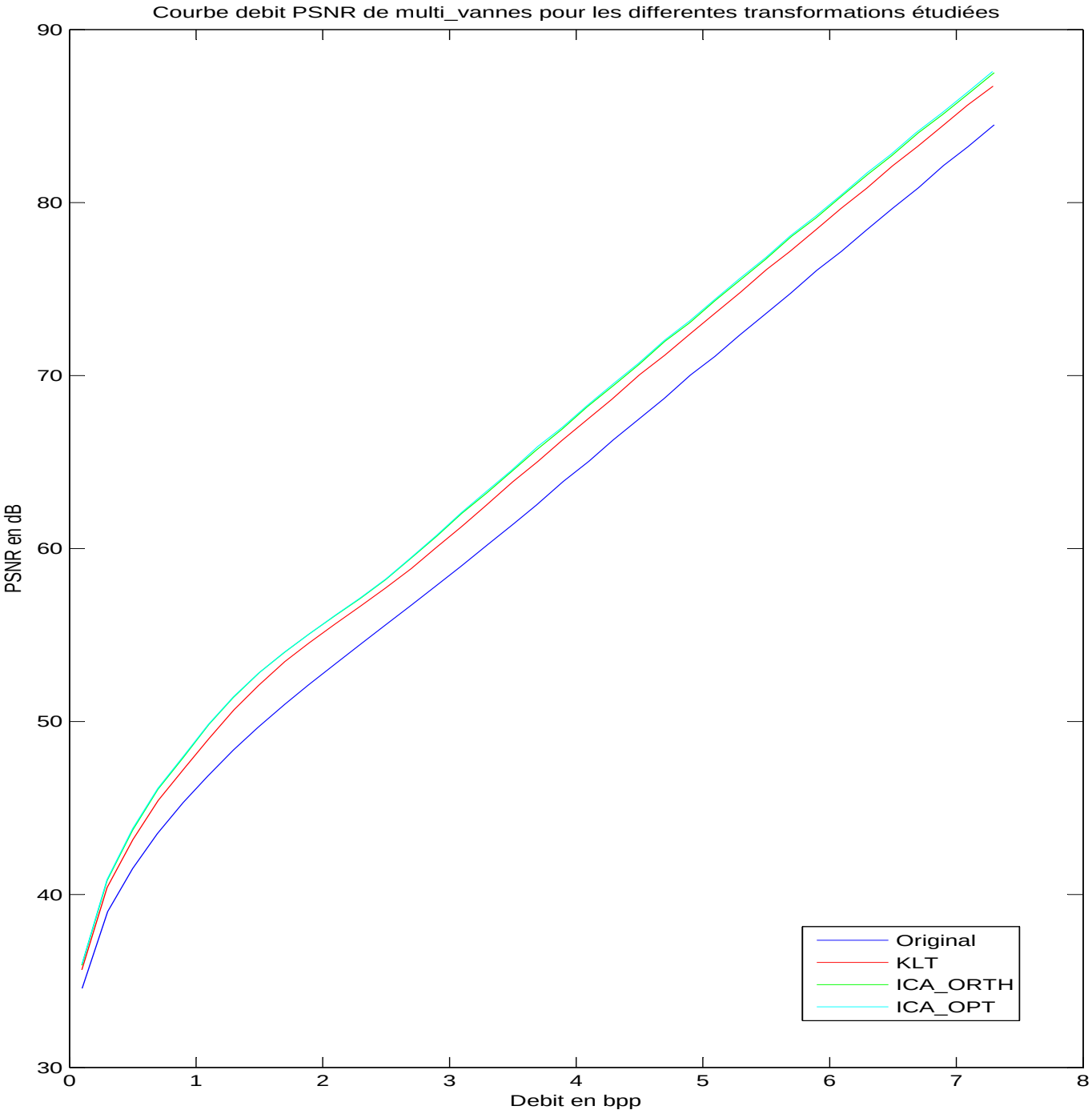
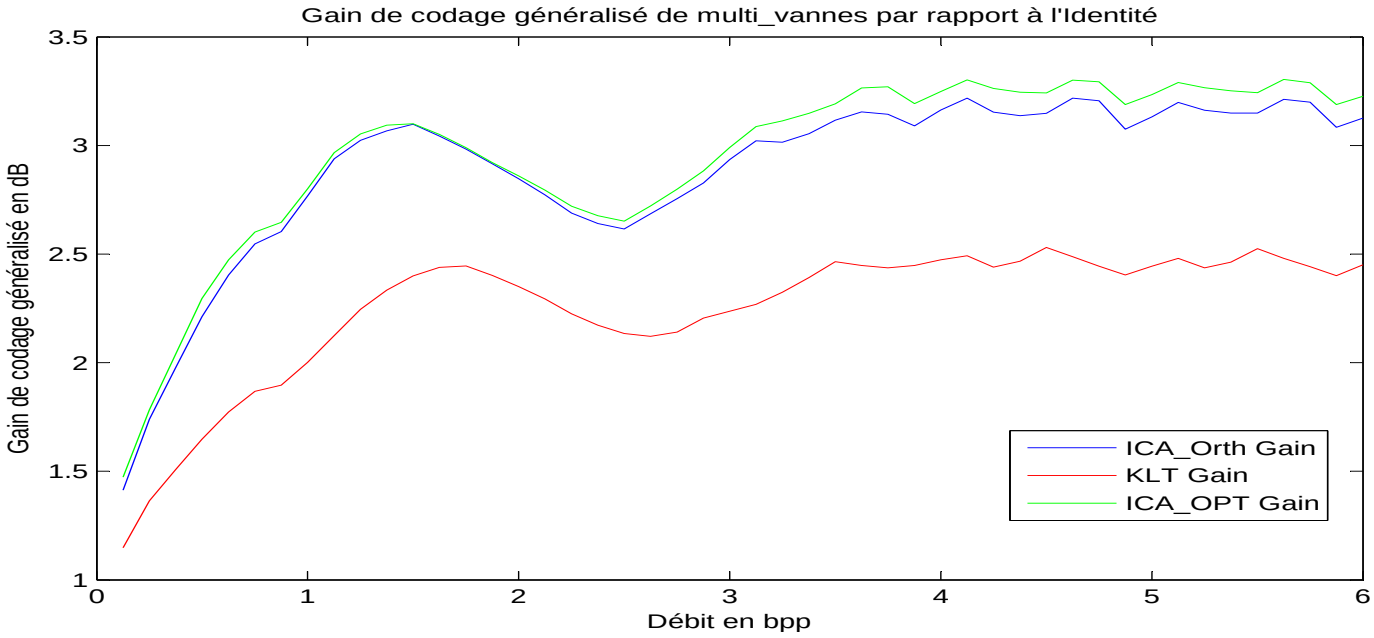


FIG. 3.21 – Courbe débit vs PRSB de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace ondelette



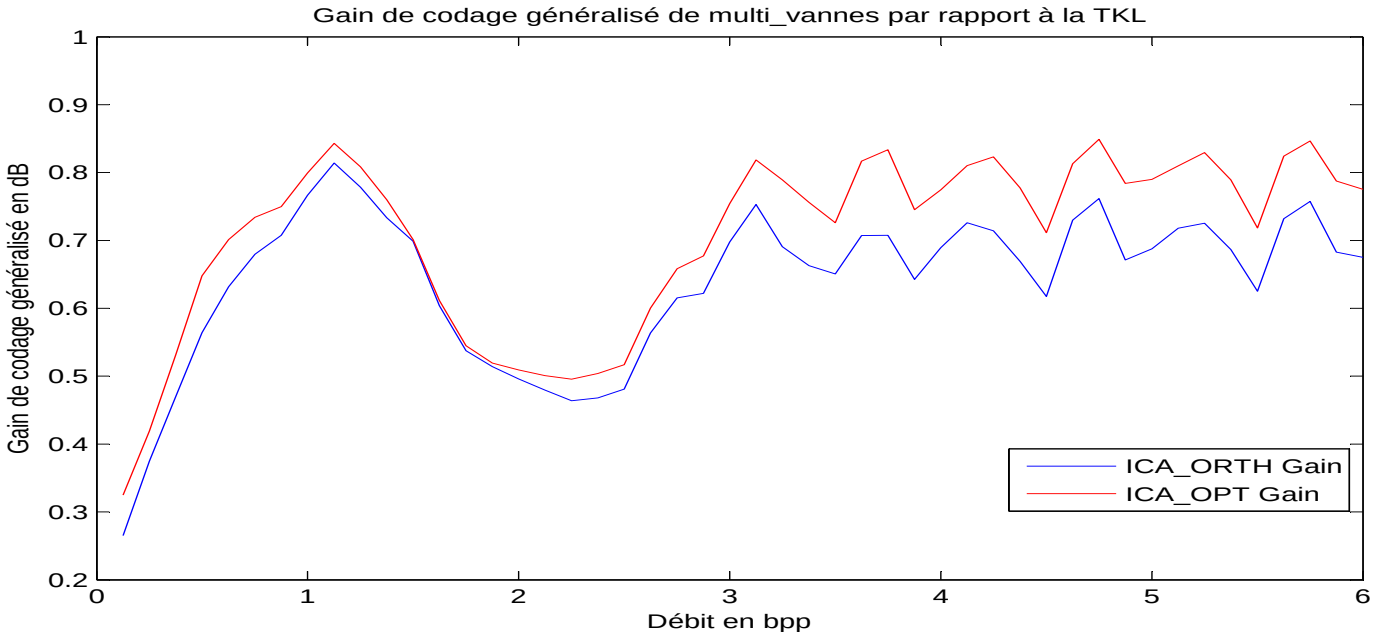


FIG. 3.22 – Gain de codage généralisé de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace ondelette

3.3.1.4 Schéma séparable avec adaptation sur l'espace image

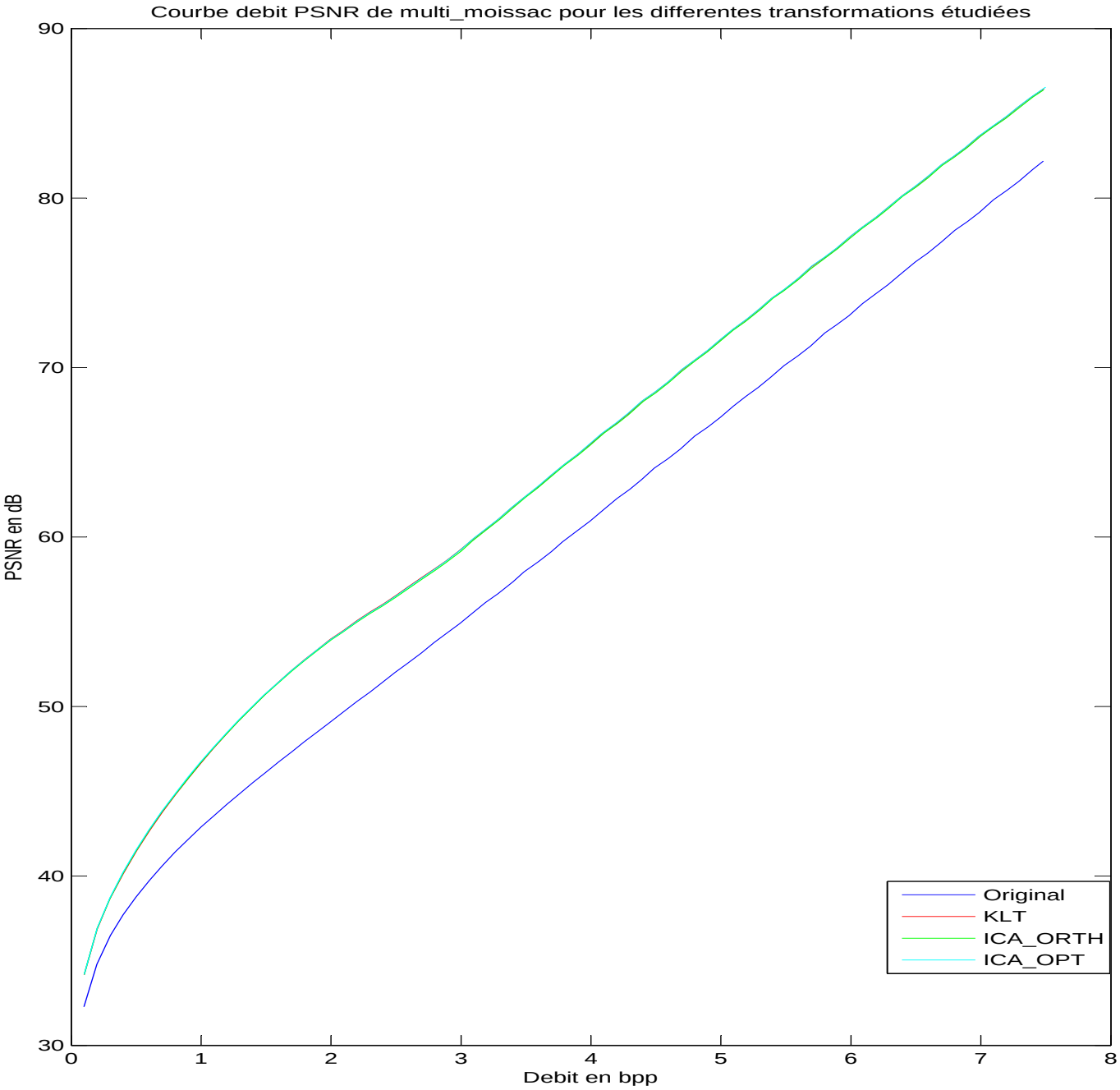
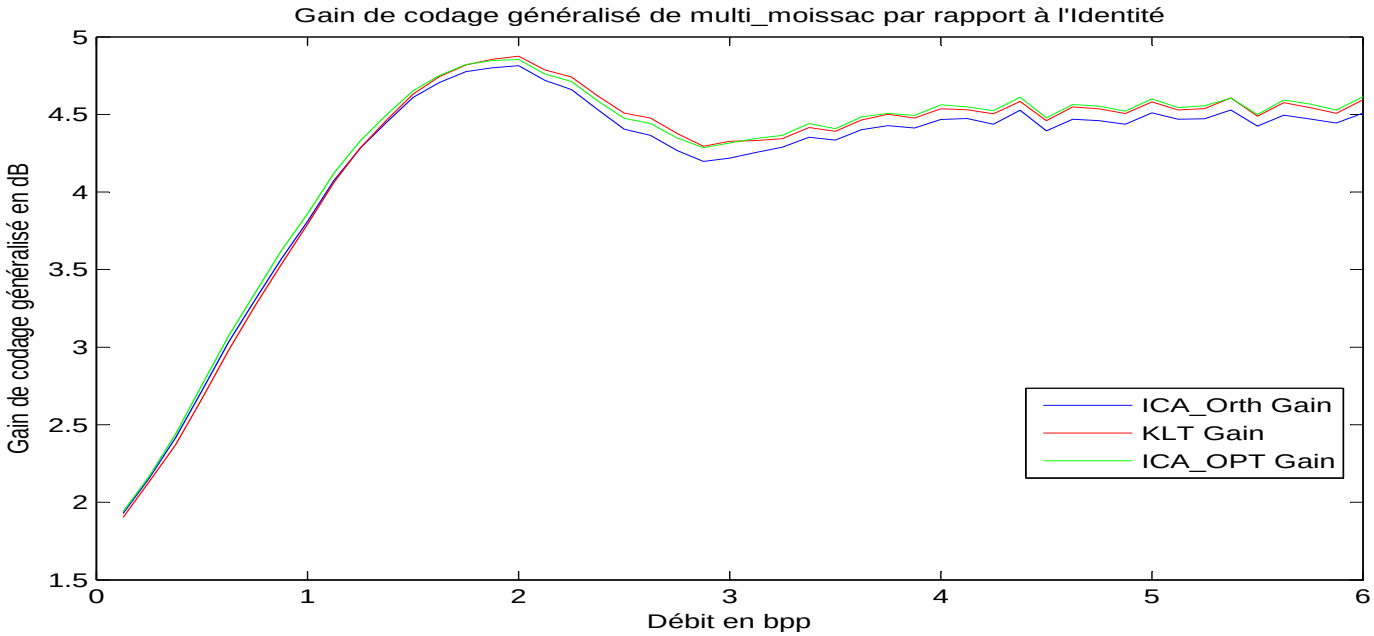


FIG. 3.23 – Courbe débit vs PRSB de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace image



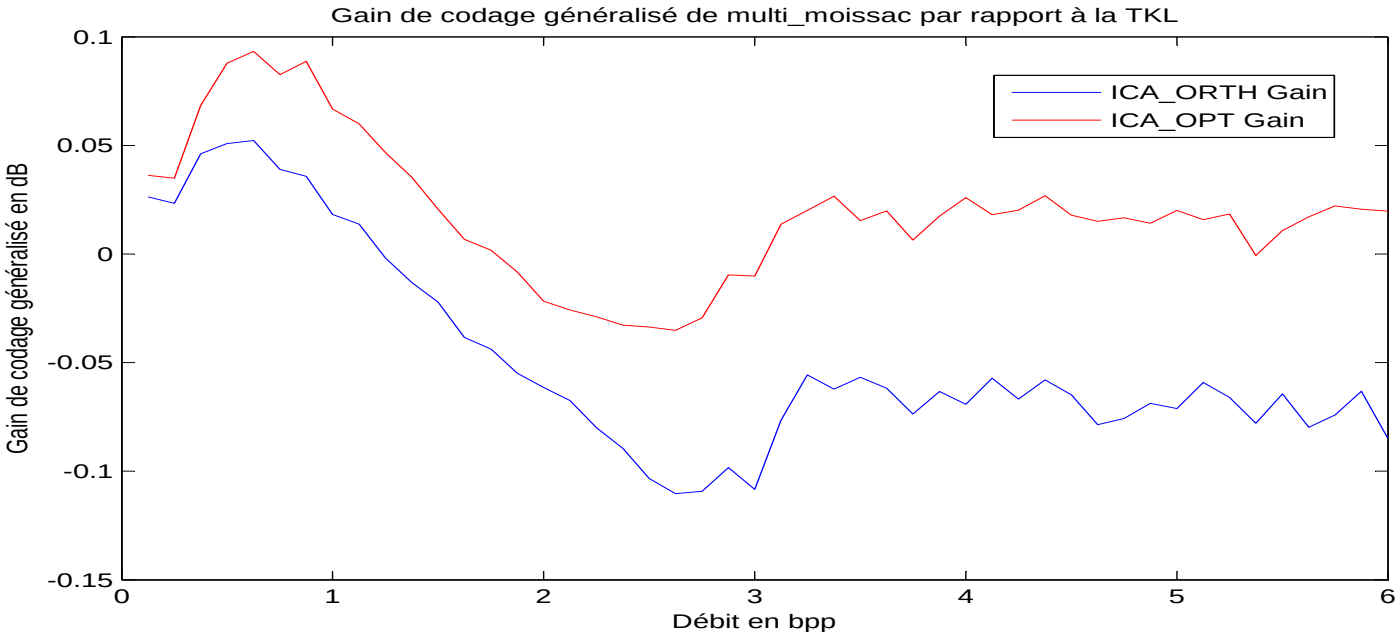


FIG. 3.24 – Gain de codage généralisé de l'image Moissac : cas du schéma séparable avec adaptation sur l'espace image

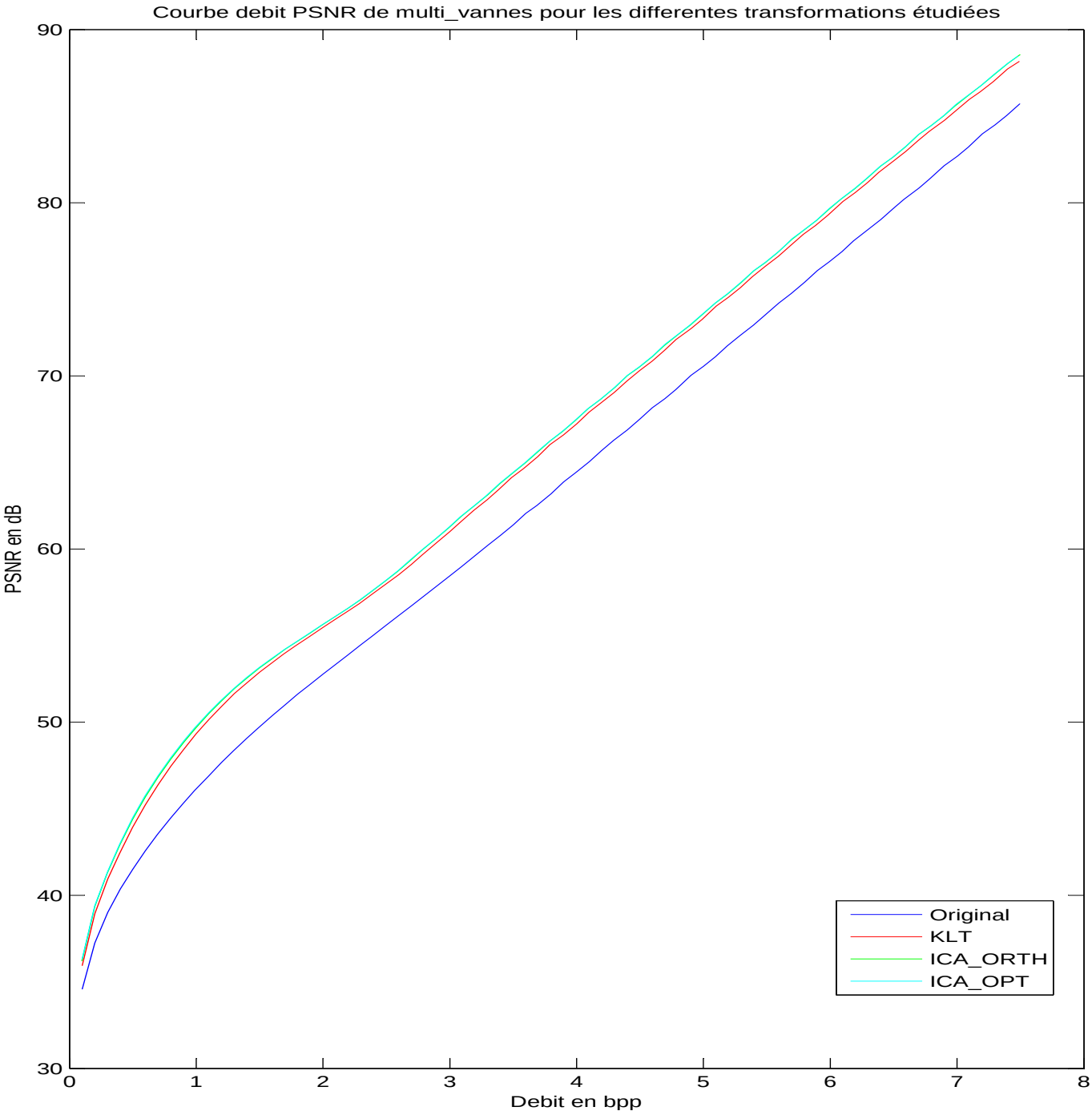
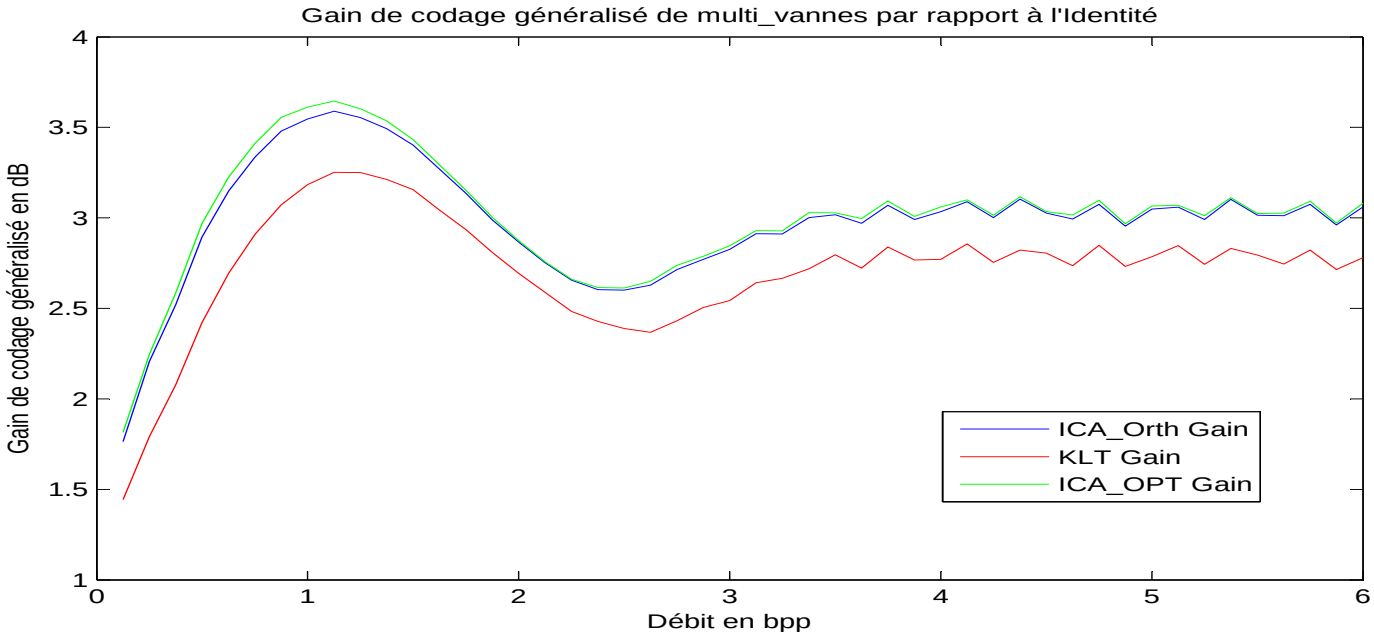


FIG. 3.25 – Courbe débit vs PRSB de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace image



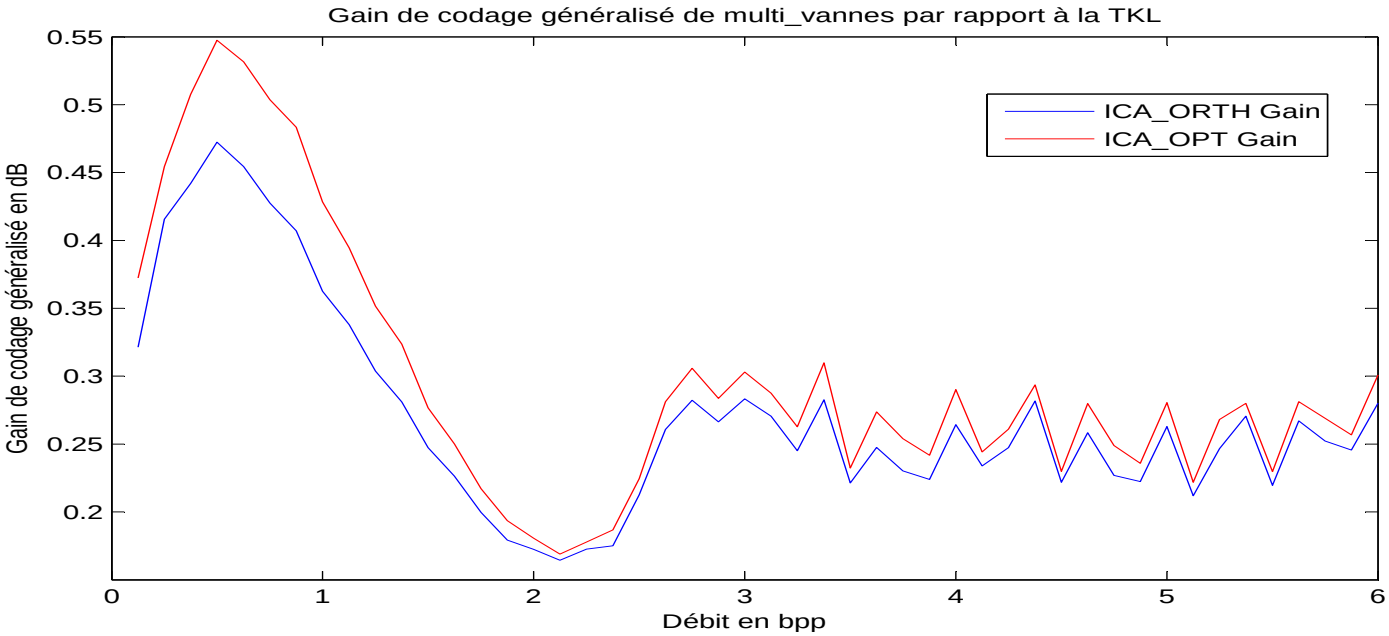


FIG. 3.26 – Gain de codage généralisé de l'image Vannes : cas du schéma séparable avec adaptation sur l'espace image

3.3.2 Résultats obtenus en utilisant le VM9

3.3.2.1 cas du schéma séparable avec adaptation sur l'espace image

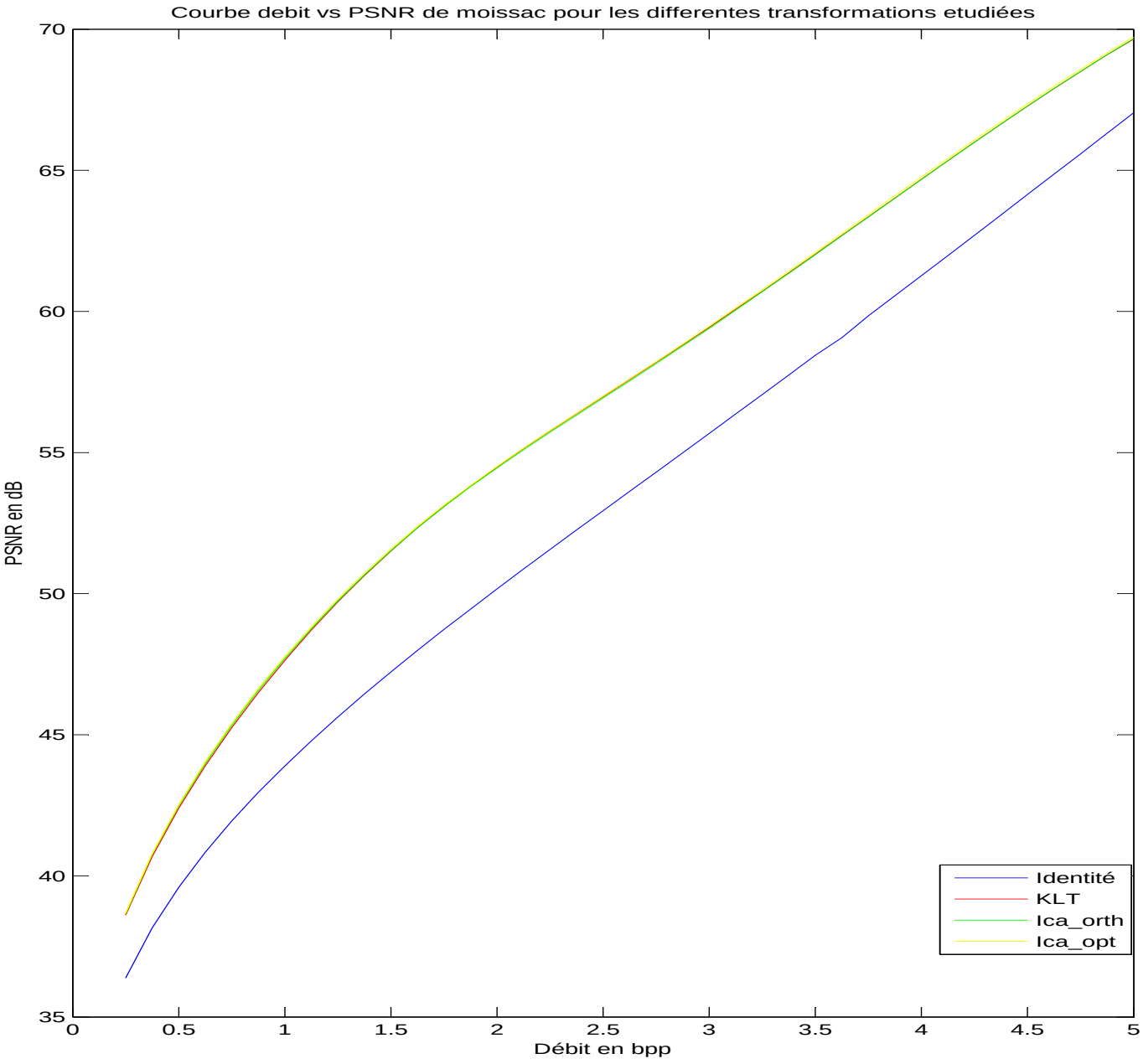
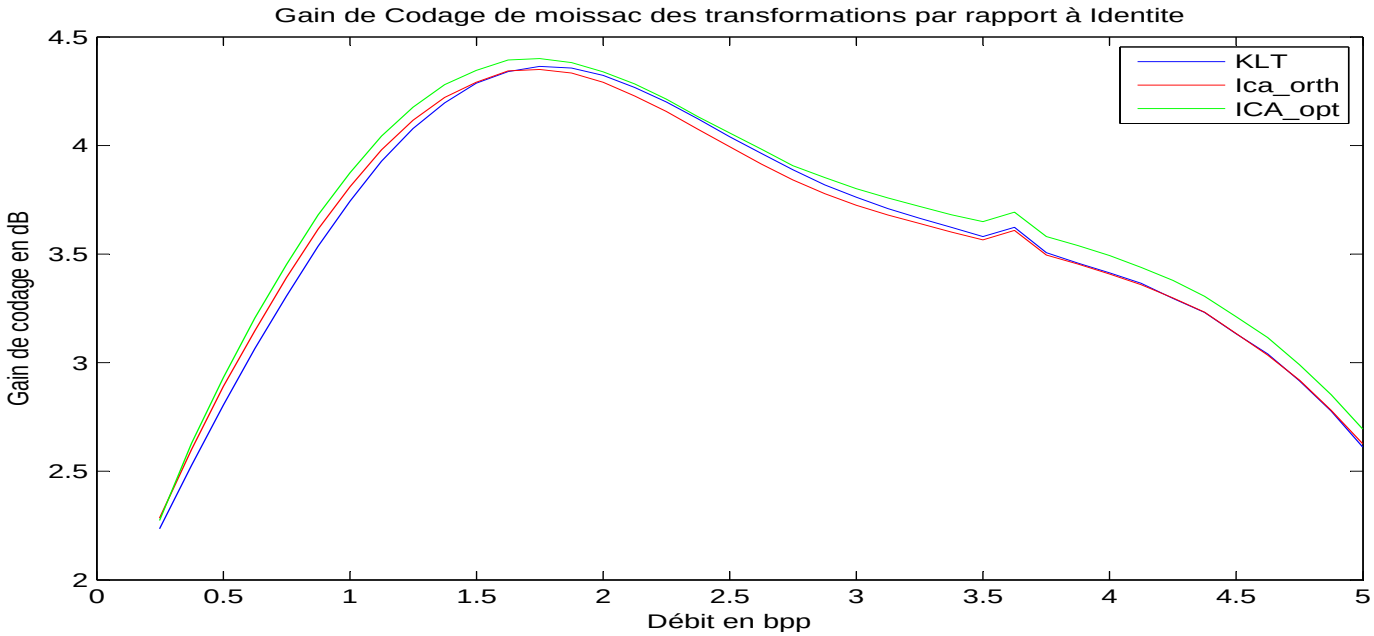


FIG. 3.27 – Courbe débit vs PRSB de l'image Moissac : schéma séparable avec adaptation dans sur l'espace image + EBCOT



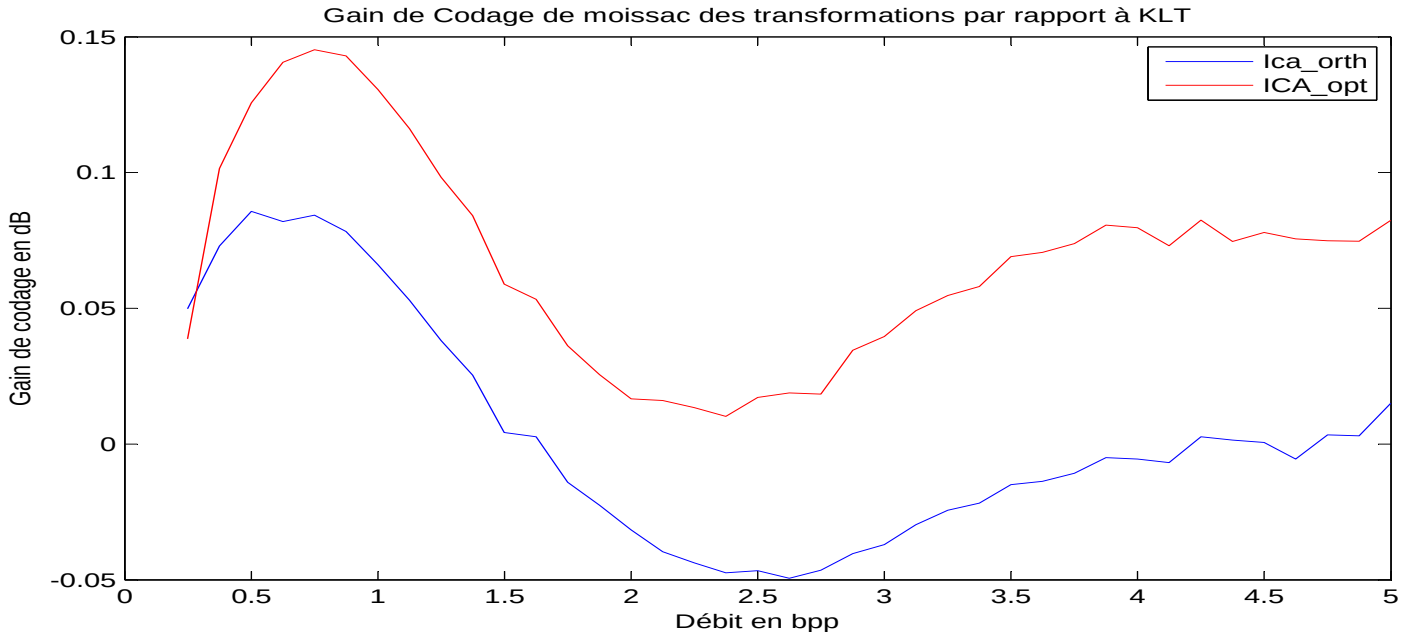


FIG. 3.28 – Gain de codage généralisé de l'image Moissac : schéma séparable avec adaptation sur l'espace image + EBCOT

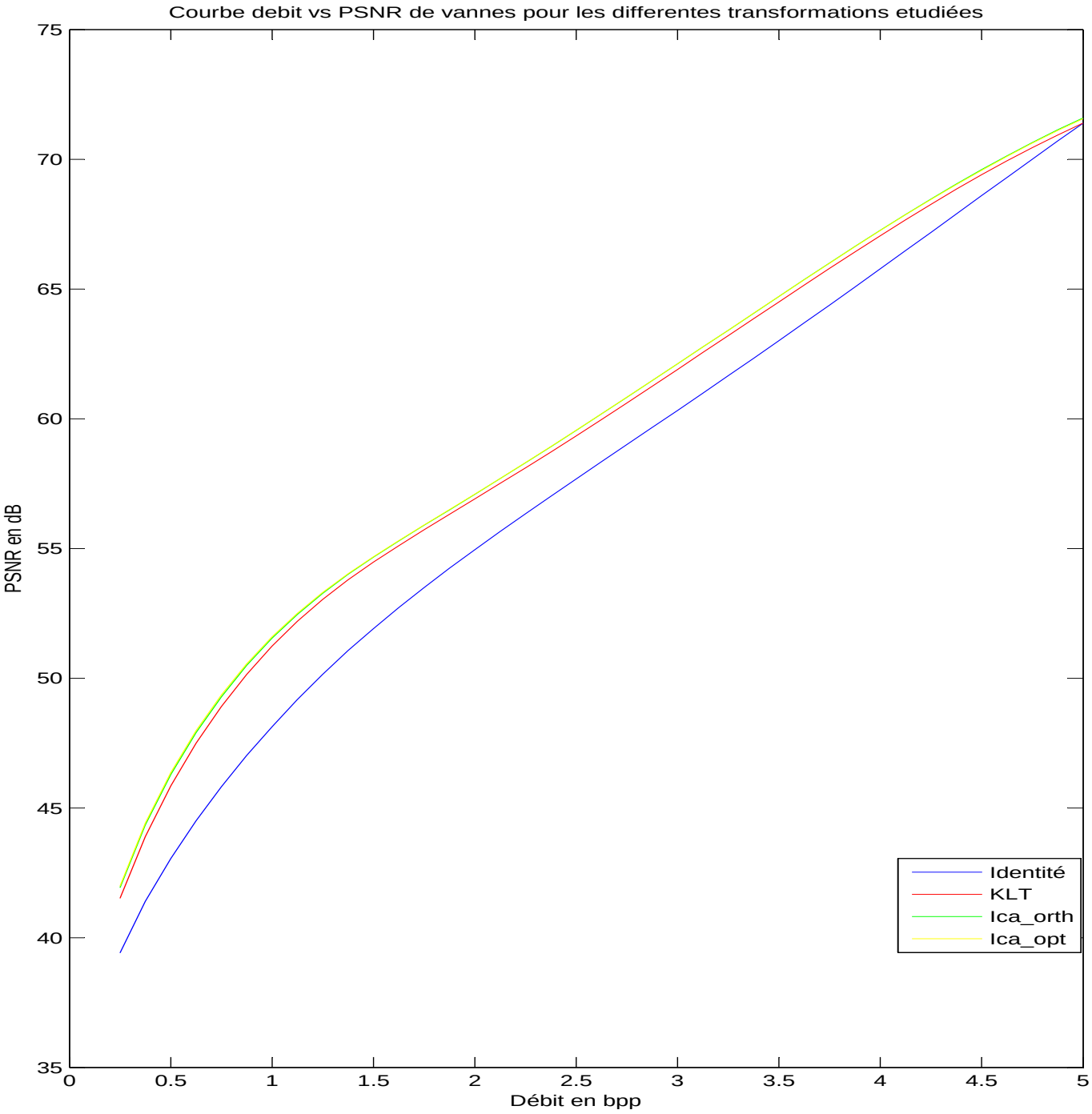
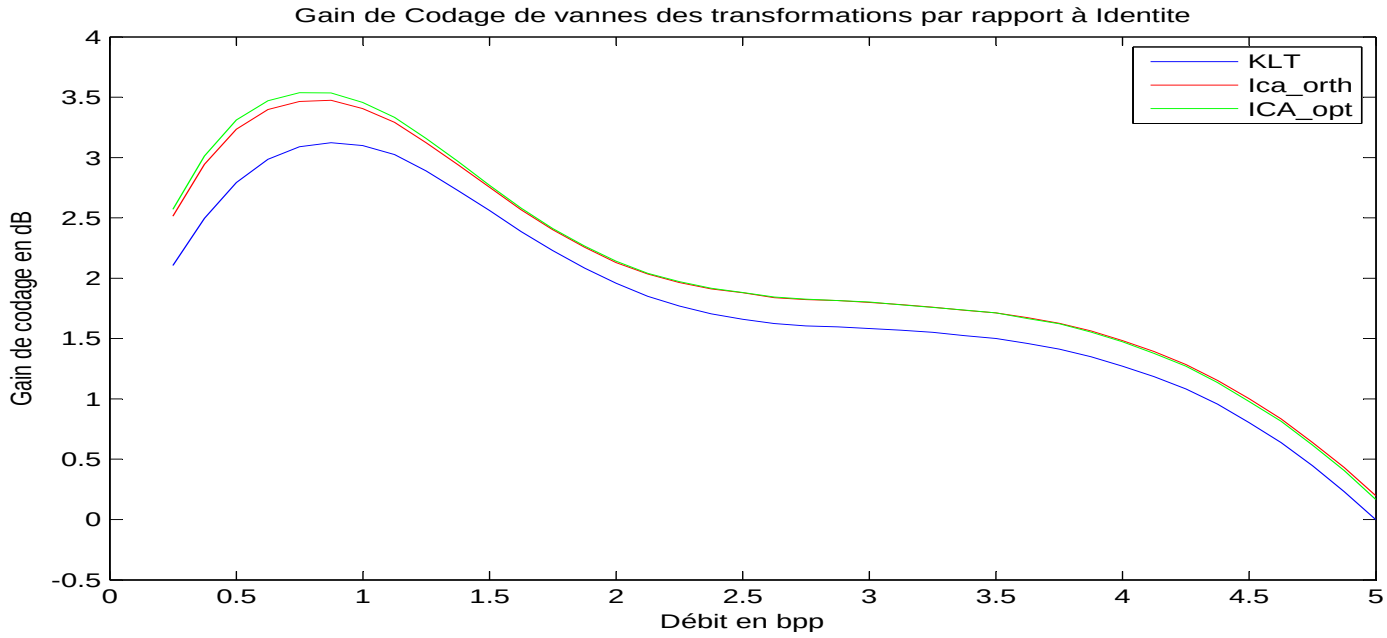


FIG. 3.29 – Courbe débit vs PRSB de l'image Vannes : schéma séparable avec adaptation sur l'espace image + EBCOT



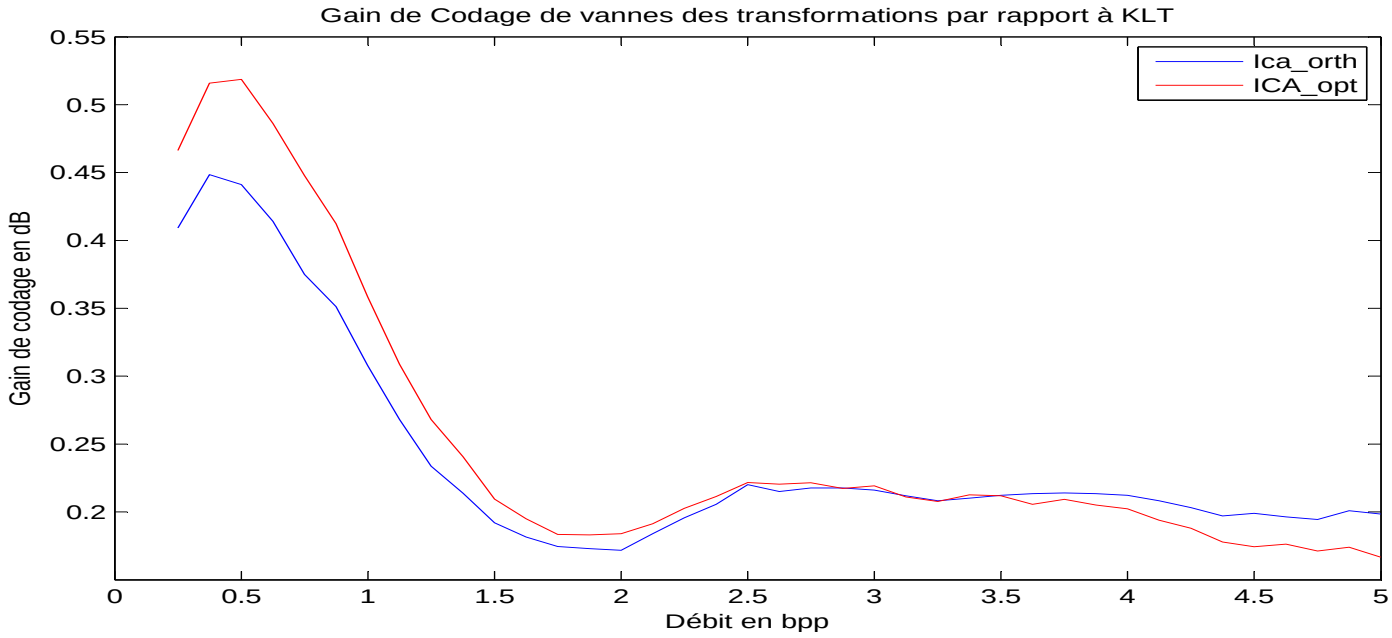


FIG. 3.30 – Gain de codage généralisé de l'image Vannes : schéma séparable avec adaptation sur l'espace image + EBCOT

3.3.3 Discussion et interprétations des résultats

3.3.3.1 Schéma séparable

Dans cette partie, nous nous intéressons uniquement au cas où le débit a été évalué en utilisant un estimateur de l'entropie d'ordre 1. En annexe2, nous avons rappelé la formule permettant d'estimer l'entropie d'ordre 1 d'une image. Les valeurs de débits que nous donnons dans ces tableaux ont une précision de $\Delta d = 0.01$ bpp. En effet, nous avons limité et tronqué le nombre de décimales à 2 en faisant un arrondi au rationnel le plus proche. Cette erreur de précision du débit entraîne une erreur de précision sur les valeurs du PRSB correspondant. Il est bien connu [31] qu'à haute résolution, en utilisant une quantification scalaire uniforme comme dans notre cas, un accroissement du débit de 1 bpp provoque un accroissement du PRSB de 6.02 dB. Nous pouvons donc déduire qu'à haute résolution, le PRSB est donné avec une erreur de précision de $\Delta P = 0.0602$ dB. Ces précisions restent identique dans tous les cas où le débit est estimé en utilisant un estimateur de l'entropie d'ordre 1.

En observant les tableaux correspondants au schéma séparable à l'annexe1 pour les images qui ont une profondeur de 8 bpp, nous constatons dans un premier temps et ce qui est somme toute évident que pour la transformation spectrale identité, les résultats obtenus dans le cas du schéma séparable avec adaptation dans l'espace image sont identiques à ceux obtenus pour le schéma séparable avec adaptation dans l'espace ondelette. En effet, dans ces deux cas, lorsque la transformation spectrale est la transformation identité, le schéma de compression se réduit juste à faire une TOD sur chaque composante. Les images ayant de fortes régions d'homogénéité comme l'image Fort_France (que nous appelons image Raw dans les tableaux) permettent d'avoir un PRSB plus élevé à débit égal que les images comme celles de la ville de Toulouse, qui présentent une grande hétérogénéité de par la diversité de la structure de la ville et de la texture de ces images. Dans le cas du schéma séparable avec adaptation dans l'espace ondelette, les transformations d'ACI modifiées ont en général des performances meilleures que celles de la TKL et qui varient selon les images. Dans les tableaux 3.4 et 3.5 nous avons évalué le gain de codage généralisé moyen par rapport à la transformation identité des images de notre base de données ayant une profondeur de 8 bpp pour différents débits. Comme nous l'avons démontré dans la section 2.4, les transformations d'ACI modifiées ne sont pas optimales dans ce cas dans la mesure où elles n'optimisent pas la maximisation du gain de codage généralisé. C'est ce qui justifie les résultats obtenus pour le schéma séparable avec adaptation dans l'espace image où dans certains cas, la TKL a des performances meilleures que celles des transformations modifiées d'ACI. Les performances des transformations spectrales dans le cas du schéma séparable avec adaptation dans l'espace image sont en général légèrement meilleures que dans le cas du schéma séparable avec adaptation dans le domaine ondelette à bas et moyens débits (débits ≤ 2 bpp). Ce qui n'est plus le cas à hauts débits où les transformations à base d'ACI modifiées ont des performances légèrement meilleures dans le cas du schéma séparable avec adaptation dans l'espace ondelette. Dans les deux cas, nous notons un gain de codage généralisé moyen des transformations modifiées d'ACI par rapport à la TKL de 0.36 dB pour Ica_orth et de 0.41 dB pour Ica_opt pour le schéma séparable avec adaptation dans l'espace ondelette. Lorsque l'adaptation est faite dans l'espace image, ce gain demeure, mais devient inférieur à 0.1 dB, la moyenne arithmétique ici étant faite sur les débits choisis pour la présentation des résultats.

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Gain TKL	1.80	2.17	2.37	2.48	2.62	2.70	2.68	2.62
Gain Ica_orth	2.08	2.51	2.74	2.86	2.98	3.05	3.05	3.01
Gain Ica_opt	2.12	2.56	2.79	2.93	3.04	3.10	3.10	3.07

TAB. 3.4 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma séparable avec adaptation dans l'espace ondelette

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Gain TKL	2.37	2.69	2.86	2.93	2.95	2.93	2.84	2.81
Gain Ica_orth	2.40	2.74	2.92	2.99	3.00	3.00	2.94	2.90
Gain Ica_opt	2.40	2.75	2.95	3.04	3.05	3.03	2.98	2.93

TAB. 3.5 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma séparable avec adaptation dans l'espace image

Pour les images ayant une profondeur de 12 bpp, nous constatons que le schéma séparable avec adaptation dans le domaine image semble en moyenne légèrement meilleur que le schéma séparable avec adaptation dans le domaine ondelette pour la TKL, La différence de performances atteignant parfois 1 dB (Image de Moissac à 0.75 bpp). Les transformations modifiées à base d'ACI, bien que ne minimisant pas vraiment le gain de codage généralisé, permettent néanmoins d'avoir de bonnes performances par rapport à la TKL, surtout pour le schéma séparable avec adaptation dans le domaine ondelette où le gain de codage de Ica_orth et de Ica_opt par rapport à la TKL est à peu près de 0.40 dB en moyenne pour tous les débits étudiés. Nous constatons que pour les images qui ont subi un léger décalage par rapport à l'image de référence, les performances sont quasi-identiques (en terme de debt vs PRSB) lorsqu'aucune transformation spectrale n'est appliquée. Par contre, les performances des transformations spectrales subissent une décroissance importante au fur et à mesure que le décalage devient important. Nous constatons une décroissance moyenne du gain de codage de 1 dB à bas et moyens débits, et à hauts débits, ce gain décroît en moyenne de 2 dB. Pour les deux schémas étudiés ici, le gain de codage généralisé des transformations d'ACI modifiés par rapport à la TKL ne subit pas de variation significative avec le décalage pour les différents débits. Ce qui nous fait croire que le décalage de l'image a une influence sur les performances des transformations spectrales, et par conséquent sur le gain de codage par rapport à la transformation identité. Cette influence devient négligeable lorsque l'on s'intéresse pour ces images (decalées et de reférence) au gain de codage des transformations modifiées d'ACI par rapport à la TKL. Nous pouvons donc conclure dans ce cas, qu'il est important et nécessaire lors des post-traitements au sol, de pouvoir recaler le images pour qu'elles soient aussi proche que possible de l'image de référence idéale. Les tableaux récapitulatifs du gain de codage généralisé moyen de ces images pour certains débits sont illustrés ci-après.

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Gain TKL	1.45	1.71	2.05	2.34	2.76	2.98	2.98	2.93	2.89	2.94
Gain Ica_orth	1.64	1.99	2.39	2.74	3.20	3.41	3.40	3.35	3.33	3.41
Gain Ica_opt	1,.65	2.02	2.43	2.78	3.24	3.45	3.44	3.40	3.40	3.47

TAB. 3.6 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma séparable avec adaptation dans l'espace ondelette

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Gain TKL	1.82	2.08	2.47	2.79	3.14	3.30	3.27	3.18	3.15	3.18
Gain Ica_orth	1.87	2.17	2.58	2.92	3.27	3.41	3.37	3.27	3.26	3.28
Gain Ica_opt	1.90	2.22	2.65	3.00	3.36	3.48	3.43	3.34	3.32	3.35

TAB. 3.7 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma séparable avec adaptation dans l'espace image

Les images hyper-spectrales AVIRIS qui ont une profondeur de 16 bpp ont subi des post-traitements. Ces images ont initialement une profondeur de 12 bpp. C'est l'une des raisons pour lesquelles nous avons des PRSB élevés même à 0.25 bpp. Lorsque la transformation spectrale est appliquée sur les $N_1 = 224$, nous avons à bas débits (débits inférieurs à 1 bpp) un accroissement moyen du PRSB de 2 dB par rapport au découpage en blocs (sur l'axe spectral) avec $N_1 = 20$. Ce qui est sans doute normale dans la mesure où la redondance spectrale dans le premier cas cité est diminuée sur toutes les bandes spectrales, ce qui est sans doute meilleur qu'une diminution par blocs sur l'axe spectral. Ce cas peut par ailleurs s'avérer être complexe à mettre en oeuvre car nécessitant de grosses ressources physiques. Le cas de l'image Moffett pour $N_1 = 224$ nous confirme ce que nous avions déjà démontré, à savoir que les transformations modifiées d'ACI ne maximisent pas dans ce cas le gain de codage généralisé, ce qui justifie le fait que la TKL ait de meilleures performances en général que Ica_orth. Les performances de la TKL pour les deux schémas séparables sont semblables et diffèrent de moins de 0.1 dB en moyenne en faveur du schéma séparable avec adaptation dans l'espace image pour tous les débits, ce qui justifie d'une certaine façon ce qui a été démontré au deuxième chapitre à travers la **proposition** 2.4.2. La différence entre les deux schémas reste du même ordre pour les transformations modifiées d'ACI. Les transformations d'ACI modifiées ne permettent pas d'améliorer d'une façon significative les performances de la TKL, dans certains cas, la TKL a des performances meilleures que ces dernières, cela provenant du fait que ce n'est pas exactement le gain de codage généralisé que ces transformations maximisent. Les tableaux ci-dessous récapitulent les gains de codage généralisés moyens obtenus pour les débits choisis.

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	17.46	18.28	17.88	17.52	16.95	16.29	15.19	14.79	14.48	14.21
Gain Ica_orth	17.30	18.24	17.85	17.52	16.96	16.30	15.20	14.79	14.47	14.21
Gain Ica_opt	17.70	18.46	18.05	17.69	17.11	16.43	15.33	14.92	14.61	14.35

TAB. 3.8 – Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace ondelette pour $N_1 = 20$

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	17.72	18.35	17.95	17.58	17.02	16.37	15.26	14.85	14.52	14.28
Gain Ica_orth	17.68	18.32	17.91	17.55	17.00	16.35	15.23	14.80	14.48	14.23
Gain Ica_opt	17.76	18.40	17.98	17.62	17.06	16.40	15.28	14.86	14.54	14.30

TAB. 3.9 – Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace image pour $N_1 = 20$

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	20.84	20.16	19.57	18.91	18.13	17.38	16.28	15.89	15.57	15.40
Gain Ica_orth	19.58	19.50	18.96	18.36	17.52	16.70	15.57	15.18	14.91	14.72
Gain Ica_opt	20.31	19.89	19.35	18.69	17.85	17.06	15.97	15.59	15.31	15.13

TAB. 3.10 – Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace ondelette pour $N_1 = 224$

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	20.78	20.16	19.58	18.96	18.20	17.47	16.33	15.93	15.60	15.47
Gain Ica_orth	20.25	19.71	19.15	18.56	17.84	17.10	15.95	15.53	15.21	15.05
Gain Ica_opt	20.90	20.22	19.67	19.02	18.23	17.48	16.34	15.94	15.62	15.48

TAB. 3.11 – Gain de codage généralisé moyen par rapport à la transformation Identité des images hyper-spectrales de 16 bpp du schéma séparable avec adaptation dans l'espace image pour $N_1 = 224$

Dans tous les cas de figure, nous constatons qu'il est important, voire même primordial, d'appliquer une transformation spectrale pour avoir de bonnes performances en compression, aussi bien pour les images multi-spectrales que pour les images hyper-spectrales, le gain de codage des transformations spectrales (TKL, Ica_orth et Ica_opt) par rapport à la transformation Identité dépassant parfois les 20 dB.

3.3.3.2 Schéma en sous-bandes

Pour les images ayant une profondeur initiale de 8 bpp, les transformations spectrales à base d'ACI font très légèrement mieux que la TKL, le gain de codage généralisé moyen par rapport à cette dernière étant respectivement de l'ordre de 0.047 dB et 0.068 dB pour Ica_orth et Ica_opt. Les performances de Ica_orth sont très proches de celles de Ica_opt, ce qui tend à montrer que la transformation optimale qui minimise le gain de codage n'est pas très loin d'être orthogonale. Nous constatons dans ce cas, que les performances sont légèrement meilleures que celles du schéma séparable avec adaptation dans le domaine image aussi bien pour la TKL que pour les transformations à base d'ACI, l'amélioration moyenne au niveau du PRSB pour un débit fixé étant de l'ordre de 0.36 dB. Le tableau ci-après donne une estimation du gain de codage généralisé moyen pour différents débits.

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Gain TKL	2.53	2.95	3.18	3.31	3.43	3.42	3.33	3.25
Gain Icaorth	2.54	2.98	3.22	3.36	3.49	3.48	3.40	3.31
Gain Icaopt	2.55	2.99	3.23	3.37	3.50	3.50	3.43	3.36

TAB. 3.12 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma en sous-bandes

Pour les images de 12 bpp, nous confirmons encore ce que nous avons constaté pour le schéma séparable, notamment que les performances des transformations spectrales décroissent avec le décalage, phénomène que nous observons sur les images de référence et les images décalées de Montpellier et de Toulouse. Les transformations à base d'ACI permettent d'améliorer légèrement celles de la TKL, l'amélioration étant plus sensibles à moyens et hauts débits, où l'on obtient parfois des gains de codage généralisés de l'ordre de 0.5 dB. Pour les débits inférieurs à 1 bpp, ce gain de codage demeure en général en deçà de 0.1 dB. Le constat fait pour les images de 8 bpp semble ici se généraliser, c'est-à-dire que ce schéma améliore légèrement les performances du schéma séparable, notamment celui du schéma séparable avec adaptation dans l'espace image, l'amélioration étant plus sensible pour les transformations à base d'ACI et allant jusqu'à 0.2 dB en moyenne. Ce qui est bien normal, dans ce cas les transformations modifiées à base d'ACI permettent effectivement de maximiser le gain de codage généralisé sur chaque sous-bande.

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Gain TKL	1.82	2.19	2.60	2.94	3.30	3.43	3.39	3.30	3.26	3.32
Gain Ica_orth	1.86	2.28	2.74	3.12	3.53	3.67	3.63	3.54	3.50	3.54
Gain Ica_opt	1.88	2.31	2.78	3.16	3.57	3.71	3.68	3.58	3.55	3.59

TAB. 3.13 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma en sous-bandes

Pour les images hyper-spectrales, nous constatons d'une part que les performances des transformations spectrales décroissent lorsque le nombre de bandes spectrales sur lesquelles elles sont estimées décroît (la décroissance moyenne étant de l'ordre de 2 dB à bas débits lorsque l'on passe de $N_1 = 224$ à $N_1 = 20$). L'influence de la taille de N_1 est mise en exergue ici avec 3 valeurs différents et nous confirmons bien que pour un débit fixé, le PRSB décroît lorsque N_1 diminue ce qui est somme toute normale. D'autre part, malgré le fait que les transformations d'ACI modifiées aient des performances très proches de celles de la TKL mais légèrement supérieures (le gain de codage généralisé moyen par rapport à la TKL étant de 0.15 dB), ce schéma en sous-bandes permet d'avoir des performances meilleures que celles du schéma séparable avec adaptation dans le domaine image. Cette amélioration atteignant parfois 1 dB (image moffett avec $N_1 = 224$ pour la transformation Ica_opt) et décroissant lorsque le débit croit. Ci-après, nous donnons une estimation du gain de codage généralisé moyen des images hyper-spectrales à certains débits pour les différents cas étudiés.

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	21.46	20.58	19.99	19.26	18.46	17.74	16.61	16.22	15.90	15.74
Gain Ica_orth	21.64	20.74	20.13	19.37	18.54	17.83	16.70	16.30	16.00	15.84
Gain Ica_opt	21.67	20.75	20.15	19.38	18.56	17.84	16.71	16.31	16.01	15.86

TAB. 3.14 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma en sous-bandes avec $N_1 = 224$

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	19.89	19.59	19.08	18.54	17.83	17.14	16.03	15.69	15.67	15.42
Gain Ica_orth	20.01	19.66	19.16	18.61	17.90	17.20	16.09	15.76	15.74	15.45
Gain Ica_opt	20.03	19.68	19.18	18.62	17.91	17.21	16.11	15.77	15.76	15.48

TAB. 3.15 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma en sous-bandes avec $N_1 = 45$

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	18.06	18.60	18.17	17.78	17.19	16.51	15.42	15.01	14.70	14.42
Gain Ica_orth	18.12	18.65	18.22	17.84	17.24	16.56	15.46	15.05	14.74	14.47
Gain Ica_opt	18.14	18.67	18.23	17.85	17.25	16.57	15.47	15.06	14.75	14.48

TAB. 3.16 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma en sous-bandes avec $N_1 = 20$

3.3.3.3 Schéma mixte en sous-bandes pour $P = 4$

Le schéma mixte en sous-bandes est intéressant dans la mesure où non seulement il permet de réduire la redondance spectrale entre les composantes, mais aussi, il permet de réduire la redondance spatiale remanente entre coefficients voisins après la TOD.

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Gain TKL	2.59	2.90	3.11	3.25	3.36	3.35	3.27	3.18
Gain Ica_orth	2.66	2.98	3.21	3.36	3.47	3.46	3.39	3.31
Gain Ica_opt	2.69	3.00	3.23	3.38	3.50	3.49	3.42	3.36

TAB. 3.17 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 8 bpp du schéma mixte en sous-bandes

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Gain TKL	2.02	2.45	2.82	3.12	3.47	3.58	3.53	3.45	3.40	3.48
Gain Ica_orth	2.11	2.61	3.05	3.39	3.81	3.93	3.89	3.80	3.74	3.82
Gain Ica_opt	2.13	2.66	3.12	3.48	3.88	4.02	3.97	3.88	3.83	3.92

TAB. 3.18 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma mixte en sous-bandes

Dans ce cas, nous constatons que même pour la transformation identité (où l'amélioration est très légère) les performances des transformations spectrales sont nettement meilleures que dans le cas du schéma en sous-bandes aussi bien pour les images de 8 bpp que pour les autres images. Au niveau de la transformation identité, cela s'explique par le fait que le débit est estimé non plus dans une sous-bande complète comme pour le schéma en sous-bandes, mais sur chacun des $P = 4$ blocs de chacune des sous-bandes. L'amélioration des performances des transformations spectrales est surtout sensible à bas et à moyens débits et est de l'ordre de 0.5 dB par rapport au schéma en sous-bandes pour les différentes transformations spectrales étudiés ici. Cependant, le gain de codage des transformations à base d'ACI par rapport à la TKL reste du même ordre pour une image donnée en comparaison au schéma en sous-bandes, ce qui nous permet de dire que le schéma mixte en sous-bandes permet une amélioration globale des performances des transformations spectrales par rapport au schéma en sous-bandes, mais ne permet pas une amélioration du gain de codage généralisé des transformations à base d'ACI par rapport à la TKL.

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	18.01	18.49	17.99	17.59	16.99	16.33	15.22	14.79	14.46	14.20
Gain Ica_orth	18.13	18.56	18.04	17.65	17.05	16.38	15.28	14.85	14.53	14.29
Gain Ica_opt	18.15	18.57	18.07	17.67	17.06	16.39	15.29	14.87	14.55	14.30

TAB. 3.19 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma mixte en sous-bandes avec $N_1 = 20$

3.3.3.4 Schéma séparable avec adaptation dans l'espace image + $VM9$

Dans ce cas, nous n'avons pas pu avoir des résultats intéressants pour les images de 8 bpp, la difficulté étant liée au bon paramétrage du $VM9$ pour ce type d'images.

Pour les images de 12 bpp, nous constatons que l'effet du décalage n'a pas d'influence sur les performances lorsque le codeur EBCOT du $VM9$ est inséré dans la chaîne de compression. Ce qui se justifie en partie par le fait que EBCOT est un codeur complexe qui effectue une modélisation de contexte sur un voisinage proche des coefficients d'ondelette après la TOD, il semble que cette modélisation inhibe l'effet du décalage car le PRSB de l'image de référence varie de moins de 0.04 par rapport à celui des images décalées pour un débit fixé. Notons que les performances du $VM9$ à bas et moyens débits permettent d'avoir une amélioration moyenne du PRSB de 0.70 dB pour un débit fixé par rapport au même schéma où le débit est estimé par l'entropie d'ordre 1. Cette amélioration est moins importante pour les débits supérieurs à 2 bpp. A l'exception de l'image de Vannes qui permet d'avoir un gain de codage moyen des transformations à base d'ACI par rapport à la TKL de l'ordre de 0.4 dB, les performances de la TKL en valeurs absolues restent sensiblement identiques à celles de ces transformations pour les autres images. Le codeur du $VM9$ exploitant la redondance résiduelle entre les coefficients après la TOD dans une sous-bande, il est difficile d'apprécier l'apport réelle de la transformation spectrale. A bas débits, les transformations spectrales permettent d'avoir un gain de codage généralisé de l'ordre de 3 dB par rapport à la transformation identité pour un débit de 1 bpp. Ce gain de codage généralisé devient quelquefois plus important au delà de 1 bpp. Le tableau ci-après résume les valeurs moyennes du gain de codage généralisé de ces images pour différents débits.

Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Gain TKL	1.83	2.23	2.55	2.77	3.01	3.05	2.99	2.85	2.69	2.51
Gain Ica_orth	1.89	2.33	2.69	2.93	3.18	3.23	3.16	3.01	2.84	2.66
Gain Ica_opt	1.89	2.35	2.73	2.99	3.24	3.28	3.20	3.05	2.88	2.7

TAB. 3.20 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 12 bpp du schéma séparable avec adaptation dans l'espace image + VM9

Pour les images hyper-spectrales, le gain apporté par EBCOT par rapport à l'entropie d'ordre 1 diminue, notamment pour l'image de Jasper où l'on constate que pour la transformation identité, le gain de codage de EBCOT par rapport à l'entropie d'ordre 1 est inférieure à 0.5 dB en moyenne. Le PRSB estimé en utilisant un estimateur de l'entropie d'ordre 1 est meilleur que celui du $VM9$ pour les transformations spectrales à base d'ACI ou la TKL pour les autres images. Cela se justifie par le fait que EBCOT pendant la phase de codage pour un débit donné, effectue plusieurs itérations en codant et décodant l'image et en réalisant l'allocation optimale de bits entre les sous-bandes par la méthode des multiplicateurs de Lagrange. Des simplifications existent à ce niveau sur le codeur pour diminuer la complexité et le temps de calcul et pour optimiser l'utilisation des ressources physiques (notamment de la mémoire vive). Ces simplifications s'accompagnent naturellement d'une perte de performances, ce qui justifie ces écarts. Dans nos simulations, lorsque nous estimons le débit par l'entropie d'ordre 1, nous faisons fi de ces préoccupations sur l'utilisation des ressources physiques et sur le temps de calcul.

Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Gain TKL	19.39	18.67	17.88	17.11	16.30	15.53	14.12	13.46	12.89	11.74
Gain Ica_orth	19.23	18.53	17.74	16.98	16.17	15.38	13.96	13.32	12.74	11.62
Gain Ica_opt	19.56	18.80	18.01	17.22	16.38	15.57	14.13	13.49	12.90	11.75

TAB. 3.21 – Gain de codage généralisé moyen par rapport à la transformation Identité des images de 16 bpp du schéma séparable avec adaptation dans l'espace image + VM9 pour $N_1 = 224$

Conclusion : Dans cette première partie, nous avons proposé et défini des schémas de compression des images multi-composantes associant une transformation pour la réduction de la redondance spatiale par composante et une autre pour la réduction de la redondance spectrale. Notons que des schémas de compressions similaires utilisant plusieurs transformations avaient déjà été proposés et étudiés dans [28] [75] avec la TKL comme transformation spectrale. Une étude similaire comparant le schéma schéma séparable avec adaptation dans le domaine image et le schéma en sous-bandes a été présentée dans [9] en incorporant aussi bien la TKL que les transformations à base d'ACI pour la réduction de la redondance spectrale. Ces résultats diffèrent de ceux qui sont présentés ici dans la mesure où l'allocation optimale de bits qui n'est valable qu'à haute résolution a été utilisée pour tous les débits dans cet article. Des formules d'approximation de la distorsion ont été déduites pour les différentes schémas de compression sous certaines hypothèses qui ont été justifiées et précisées. Ces formules d'approximation de la distorsion sont nécessaires pour l'allocation optimale de bits aussi bien à haut débit qu'à bas débit où l'algorithme de Shoham et Gersho est utilisé. Ces formules d'approximation font intervenir des coefficients de pondérations dont nous avons explicité et détaillé les expressions. Nous avons enfin évalué les performances des différents cas de figure proposés d'abord en calculant le débit avec un estimateur de l'entropie d'ordre 1, et dans un autre cas, nous avons utilisé le VM9 dans la chaîne de compression pour avoir une estimation du débit

par le codeur EBCOT de JPEG2000. Les résultats obtenus nous montrent qu'il y'a une nécessité d'appliquer des transformations pour la réduction de la redondance spectrale. Les performances des transformations à base d'ACI modifiées introduites par Narozny [52] [51] permettent en général d'avoir des performances légèrement meilleures que celles de la TKL en terme de gain de codage généralisé par rapport à la transformation Identité. Dans certains cas, ce gain de codage devient assez significatif.

Les schémas de compression proposés dans cette partie nécessitent l'application de deux types de transformations, l'une et l'autre étant indépendantes et permettant de réduire un type de redondance chacune (sauf pour les cas mixtes ou la redondance spectrales permet aussi de réduire une partie de la redondance spatiale). Nous avons réfléchi à la possibilité d'utiliser une transformation unique qui permettrait de réduire au mieux les redondances spatiales et spectrales en compression. L'utilisation d'une transformation de séparation déconvolution aveugle modifiée pour la compression, pourrait être une solution à ce dernier problème. L'étude et l'intégration d'une telle transformation dans un système de compression fait l'objet de la deuxième partie de cette thèse.

Deuxième partie

Étude et évaluation des performances d'un système de compression utilisant des transformations de séparation-déconvolution

Introduction

Notre objectif dans cette thèse est d'étudier et de proposer de nouvelles transformations permettant d'avoir de bonnes performances en compression d'images multi composantes. Dans la première partie, nous avons proposé et étudié un système de compression général utilisant une transformation pour la réduction de la redondance spatiale et une autre pour la réduction de la redondance spectrale. Ces deux transformations indépendantes étant appliquées l'une après l'autre à l'image, nous avons étudié plusieurs scénarios dépendant de l'ordre dans lequel ces transformations sont appliquées. Bien que l'une et l'autre des transformations utilisées pour la réduction des redondances spatiales (TOD pour la réduction de la redondance spatiale dans une composante) et spectrales (Transformations à base d'ACI ou la TKL) permettent chacune de bien réduire un type de redondance, il n'est pas assuré que leur association permette de réduire les redondances dans une image multi-composante de façon optimale.

Dans cette partie, nous nous proposons de rechercher une transformation unique permettant de réduire de façon optimale à haute résolution, les redondances spatiales et spectrales. Un modèle de mélange qui parait bien adapté de par la géométrie d'une image multi composante est un modèle de mélange convolutif. Un tel modèle permettrait de réaliser à la fois la séparation (dans le sens où l'on réduirait la redondance spectrale entre les composantes de l'image) et la déconvolution (dans le sens où l'on réduirait la redondance spatiale dans chaque composante) dans une image multi composante et ce, d'une façon bien adaptée à la compression.

Le premier chapitre de cette partie pose le problème de la séparation-déconvolution appliquée à la compression des images en utilisant un modèle de mélange convolutif et met à jour un nouveau critère à minimiser. Le deuxième chapitre quant à lui est basé sur la mise au point d'un algorithme permettant de minimiser ce critère en se servant de son gradient et d'une approximation de son hessien. Une étude de la complexité de l'algorithme y est faite. Cette deuxième partie s'achève par une étude des performances obtenues en intégrant cette nouvelle transformation dans un schéma de compression. Nous comparons les performances obtenues avec celles des schémas de compression de la première partie.

Chapitre 4

Définition du critère en compression pour un modèle de mélange convolutif

Dans ce chapitre nous posons le problème de la compression d'une image multi-composante en utilisant un modèle de mélange convolutif. Après quelques rappels sur le modèle convolutif en séparation aveugle de sources, nous posons le problème de la compression en utilisant ce modèle dans un schéma de compression assez simple. Une formule de la distorsion est ensuite établie, ce qui nous permet de mettre à jour et de définir un nouveau critère en compression. Enfin, nous établissons une relation entre ce nouveau critère, et le critère à minimiser en ACI, pour un modèle de mélange convolutif.

4.1 Rappels sur le modèle convolutif

Soit $\mathbf{X}$ une image numérique constituée de N composantes, chaque composante $\mathbf{X}_i$ $(1 \leq i \leq N)$ étant une image 2D comprenant N_L lignes et N_C colonnes. Nous notons :

$$\mathbf{X}(k_1, k_2) = [X_1(k_1, k_2), X_2(k_1, k_2), \ \ldots, \ X_N(k_1, k_2)]^T$$

un multi-pixel, $1 \leq k_1 \leq N_L$, $1 \leq k_2 \leq N_C$, constitué des pixels à la position (k_1, k_2) de toutes les composantes prises sur l'axe spectral. Supposons que le signal numérique $\mathbf{X}$ représentant l'image multi-composante soit obtenu à partir d'un modèle de mélange convolutif comme suit :

$$\mathbf{X}(k_1, k_2) = \sum_{\ell_1} \sum_{\ell_2} \mathbf{M}(\ell_1, \ell_2) \mathbf{S}(k_1 - \ell_1, k_2 - \ell_2) \tag{4.1}$$

où $\{\mathbf{M}(\ell_1, \ell_2)\}$ est une suite de matrices carrées de taille $N \times N$ et

$$\mathbf{S}(k_1, k_2) = [S_1(k_1, k_2), S_2(k_1, k_2), \ \ldots, \ S_N(k_1, k_2)]^T$$

est un multi-coefficient du signal source $\mathbf{S}$. Supposons que les sources et les observations soient de même dimension, c'est-à-dire que $\mathbf{S}$ est constitué de N composantes et que chaque composante $\mathbf{S}_i$ $(1 \leq i \leq N)$ est un sinal 2D comprenant N_L lignes et N_C colonnes. Nous faisons les hypothèses suivantes :

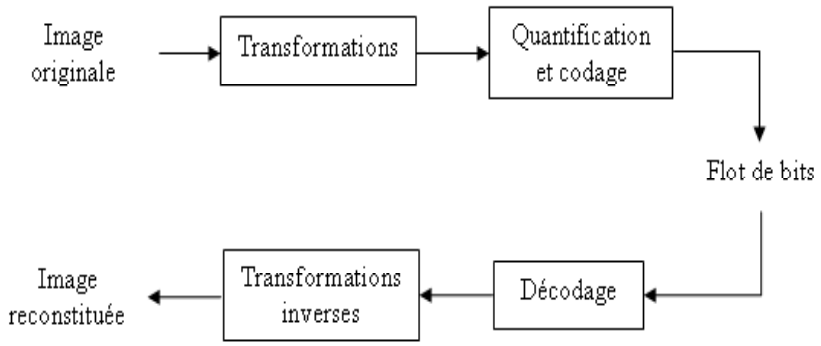


FIG. 4.1 – Schéma général d'un système de compression simplifié

$H3$: Les composantes de $\boldsymbol{S}$ sont mutuellement indépendantes c'est-à-dire que $\boldsymbol{S}_i$ est indépendant de $\boldsymbol{S}_j$ pour $i \neq j$.

$H4$: Dans une composante de l'image multi composante source, les coefficients sont indépendants entre eux, en d'autres termes nous avons $S_i(j,k)$ et $S_i(p,q)$ qui sont indépendants pour $j \neq p$ ou $k \neq q$.

Ces deux hypothèses traduisent le fait que les composantes de l'image source sont déconvoluées dans le sens que dans une composante de l'image source les coefficients sont indépendants les uns des autres, et que celles-ci sont mutuellement indépendantes. Partant des observations dont nous disposons et qui constituent le signal $\mathbf{X}$, nous voulons retrouver des composantes $\mathbf{Y}$ vérifiant autant que possible les hypothèses $H3$ et $H4$ ci-dessus. Ceci est réalisé par application d'une opération de convolution comme suit :

$$\mathbf{Y}(k_1,k_2) = \sum_{\ell_1}\sum_{\ell_2} \mathbf{W}(\ell_1,\ell_2)\mathbf{X}(k_1-\ell_1,k_2-\ell_2) \tag{4.2}$$

où $\mathbf{Y}$ est un signal de même dimension que $\mathbf{X}$ et $\{\mathbf{W}(\ell_1,\ell_2)\}$ est une suite de matrices carrées de taille $N \times N$. Des modèles similaires pour la séparation de sources appliquées aux images ont été utilisés dans [22] [17] pour la déconvolution ou le débruitage des images. Supposons enfin que l'ensemble des matrices $\{\mathbf{W}(\ell_1,\ell_2)\}$ non nulles soit fini, c'est-à-dire que pour $\ell_1 > L_1$ et $\ell_2 > L_2$, $\mathbf{W}(\ell_1,\ell_2) = \mathbf{O}_N$, $\mathbf{O}_N$ étant la matrice carrée nulle d'ordre N. L'objectif de ce que nous appellerons par abus de langage l'ACI convolutif, est donc de trouver les $L_1 \times L_2$ matrices $\mathbf{W}(\ell_1,\ell_2)$ de taille $N \times N$ qui satisfont l'équation (4.2) et qui vérifient autant que faire se peut les hypothèses $H3$ et $H4$.

4.2 Description du schéma de compression, formule de la distorsion

Soit le schéma de compression général de la figure 4.1 où la transformation utilisée est une transformation de séparation et déconvolution. L'image originale $\mathbf{X}$ est transformée en $\mathbf{Y}$ par une

transformation convolutive suivant le modèle de l'équation (4.2), les composantes $\mathbf{Y}_i$ obtenues sont ensuite quantifiées et codées séparément. Après quantification de l'image transformée, nous obtenons une image transformée quantifiée notée $\mathbf{Y}^q$. Nous utilisons une quantification scalaire uniforme par composante et un codeur entropique simple comme le codeur arithmétique ou le codeur de Huffman qui permettent d'avoir des performances proches de celle de l'entropie d'ordre 1. En appliquant le schéma de compression inverse, nous pouvons reconstituer une image approchée $\widehat{\mathbf{X}}$ de $\mathbf{X}$. Nous nous intéressons à la distorsion entre le signal original $\mathbf{X}$ et le signal reconstitué $\widehat{\mathbf{X}}$, la mesure de distorsion étant l'erreur quadratique moyenne. Par la suite nous utilisons les notations vectorielles par souci de simplification et nous notons $\mathbf{k} = (k_1, k_2)$, $\ell = (\ell_1, \ell_2)$ les coordonnées spatiales d'une image ou d'une suite de matrices.

Rappelons que l'opération de codage est une opération réversible et sans perte, la distorsion n'intervenant que par le fait de la quantification. Soient $\{\mathbf{W}(\ell)\}$, le filtre matriciel utilisé pour transformer l'image $\mathbf{X}$ en $\mathbf{Y}$ et $\{\mathbf{V}(\ell)\}$ son inverse (le filtre matriciel inverse), utilisé pour reconstituer $\widehat{\mathbf{X}}$ à partir de $\mathbf{Y}^q$. L'opération de convolution qui n'est rien d'autre qu'une opération de filtrage est traduite dans l'un et l'autre cas par les équations

$$\mathbf{Y}(\mathbf{k}) = \sum_{\ell} \mathbf{W}(\ell)\mathbf{X}(\mathbf{k} - \ell) \tag{4.3}$$

pour la transformation directe et

$$\widehat{\mathbf{X}}(\mathbf{k}) = \sum_{\ell} \mathbf{V}(\ell)\mathbf{Y}^q(\mathbf{k} - \ell) \tag{4.4}$$

pour la transformation inverse. La distorsion entre l'image originale et celle reconstituée est par définition :

$$D(\mathbf{X}, \widehat{\mathbf{X}}) = \frac{1}{N N_L N_C} \sum_{\mathbf{k}} [\mathbf{X}(\mathbf{k}) - \widehat{\mathbf{X}}(\mathbf{k})]^T [\mathbf{X}(\mathbf{k}) - \widehat{\mathbf{X}}(\mathbf{k})] \tag{4.5}$$

Elle peut encore s'écrire plus simplement comme

$$D(\mathbf{X}, \widehat{\mathbf{X}}) = \frac{1}{N} \sum_{i=1}^{N} D(\mathbf{X}_i, \widehat{\mathbf{X}}_i) \tag{4.6}$$

où $D(\mathbf{X}_i, \widehat{\mathbf{X}}_i)$ est la distorsion de la $i^{ème}$ composante de $\mathbf{X}$ et est définie par

$$D(\mathbf{X}_i, \widehat{\mathbf{X}}_i) = \frac{1}{N_L N_C} \sum_{\mathbf{k}} (X_i(\mathbf{k}) - \widehat{X}_i(\mathbf{k}))^2 \tag{4.7}$$

Le bruit de quantification du signal sur l'image $\mathbf{X}$ est donné par :

$$\mathbf{X}(\mathbf{k}) - \widehat{\mathbf{X}}(\mathbf{k}) = \sum_{\ell} \mathbf{V}(\ell)\{\mathbf{Y}(\mathbf{k} - \ell) - \mathbf{Y}^q(\mathbf{k} - \ell)\} \tag{4.8}$$

Partant de cette expression, on déduit que le bruit de quantification de la $i^{ème}$ composante de $\mathbf{X}$ est obtenu par

$$\mathbf{X}_i(\mathbf{k}) - \widehat{\mathbf{X}}_i(\mathbf{k}) = \sum_{\ell} \sum_{j} V_{ij}(\ell)\{Y_j(\mathbf{k} - \ell) - Y_j^q(\mathbf{k} - \ell)\}$$

où $V_{ij}(\ell)$ désigne l'élément (i,j) de la matrice $\mathbf{V}(\ell)$. Si de plus l'hypothèse $H1$ est admise (hypothèse qui a été justifiée dans la section 1.2.1) c'est-à-dire que les composantes du bruit de quantification $\mathbf{b}(\mathbf{k}) = \mathbf{Y}(\mathbf{k}) - \mathbf{Y}^q(\mathbf{k})$ sont décorrélées et de moyennes nulles, alors la distorsion $D(\mathbf{X}_i, \widehat{\mathbf{X}}_i)$ devient :

$$D(\mathbf{X}_i, \widehat{\mathbf{X}}_i) = \sum_{\ell} \sum_{j} V_{ij}^2(\ell) D_j \tag{4.9}$$

où D_j désigne la distorsion de la composante $\mathbf{Y}_j$ de $\mathbf{Y}$. De cette dernière équation, on déduit que la distorsion de l'image $D(\mathbf{X}, \widehat{\mathbf{X}})$ s'obtient suivant l'expression

$$\begin{aligned} D(\mathbf{X}, \widehat{\mathbf{X}}) &= \frac{1}{N} \sum_{i=1}^{N} \sum_{\ell} \sum_{j} V_{ij}^2(\ell) D_j \\ &= \frac{1}{N} \sum_{j=1}^{N} w_j D_j \end{aligned} \tag{4.10}$$

avec $w_j = \sum_i \sum_\ell \mathbf{V}_{ij}^2(\ell)$. Cette dernière expression est très intéressante dans la mesure où elle nous permet, sous l'hypothèse $H1$ (qui est très réaliste et somme toute très peu restrictive), d'écrire la distorsion de l'image $\mathbf{X}$ comme une somme pondérée des distorsions des composantes de l'image transformée $\mathbf{Y}$, les coefficients de pondération dépendant essentiellement du filtre de la transformation inverse.

Nous constatons une adéquation avec la formule de la distorsion qui a été obtenue dans la section 1.2 pour un modèle linéaire qui suppose un mélange instantané. La forme des coefficients de pondération change néanmoins et nécessite le calcul du filtre inverse. Dans le cas d'un mélange instantané, le calcul de ces coefficients de pondération nécessite simplement le calcul de l'inverse de la matrice de transformation. Par la suite, nous verrons qu'il est possible, à partir uniquement du filtre de la transformation directe, de calculer la valeur de ces coefficients de pondération.

4.3 Le critère à optimiser en compression

Notre étude concerne le schéma de compression qui a été décrit dans la section précédente. Nous nous intéressons au codage séparé des composantes de $\mathbf{Y}$. A haute résolution ou faible distorsion, la longueur moyenne de codage, ou débit, d'une variable aléatoire X est donnée d'après la formule d'approximation de Bennett par $H(X) - \frac{1}{2}\log_2(cD)$, où D est la distorsion introduite par la quantification et c est une constante qui dépend du type de quantification. Dans le cas d'une quantification scalaire uniforme, nous avons $c = 12$.

Se pose le problème d'allocation optimale de bits ou de distorsion quand plusieurs quantificateurs sont utilisés. Étant donné un budget de bits total R_c par pixel, comment répartir ce nombre de bits entre les composantes $\mathbf{Y}_j$ pour que la distorsion globale entre $\mathbf{X}$ et $\widehat{\mathbf{X}}$ soit minimale sous la contrainte que le débit final moyen obtenu R soit tel que $R \leq R_c$. Les problèmes d'allocation de bits et de distorsions étant duaux, il est équivalent de savoir, pour une distorsion cible fixée D_c, comment minimiser le débit global moyen par pixel, dont l'expression à haute résolution est

$$\frac{1}{N} \sum_{j=1}^{N} [H(Y_j) - \frac{1}{2}\log_2(cD_j)]$$

sous la contrainte que la distorsion effective $\sum_j w_j D_j/N$ soit telle que $\sum_j w_j D_j \leq N D_c$, par rapport aux distorsions D_j et aux matrices $\mathbf{V}(\ell)$ du filtre inverse. Pour des matrices $\mathbf{V}(\ell)$ fixées, la résolution de ce problème d'allocation revient à prendre (comme cela a été démontré dans la section 2.2) $D_j = D_c/w_j$. Le problème en compression finalement revient pour D_c fixé, à minimiser l'expression

$$\sum_{j=1}^{N}[H(Y_j) + \frac{1}{2}\log_2(w_j)] \tag{4.11}$$

avec, rappelons le, $w_j = \sum_i \sum_\ell \mathbf{V}_{ij}^2(\ell)$. Le terme w_j est le j^{eme} terme diagonal de la matrice $\sum_\ell \mathbf{V}(\ell)^T\mathbf{V}(\ell)$. Compte tenu des propriétés générales de la fonction logarithmique, l'expression de l'équation (4.11) peut aussi s'écrire comme

$$\sum_{j=1}^{N} H(Y_j) + \frac{1}{2}\log_2(\prod_{j=1}^{N} w_j) \tag{4.12}$$

En utilisant le fait que le terme w_j est le $j^{ème}$ terme diagonal de la matrice $\sum_\ell \mathbf{V}(\ell)^T\mathbf{V}(\ell)$, le problème en compression, en utilisant un modèle de mélange convolutif, revient donc à minimiser

$$C = \sum_{j=1}^{N} H(Y_j) + \frac{1}{2}\log_2 \det\{\text{diag}(\sum_\ell \mathbf{V}(\ell)^T\mathbf{V}(\ell))\} \tag{4.13}$$

où diag() est l'opérateur qui construit une matrice diagonale à partir des termes diagonaux de son argument. Le critère de l'équation (4.13) dépend à la fois du filtre de la transformation $\{\mathbf{W}(\ell)\}$ (nécessaire pour le calcul de l'image transformée $\mathbf{Y}$) et de son inverse $\{\mathbf{V}(\ell)\}$.

En pratique, partant des observations qui constituent le signal original $\mathbf{X}$, nous recherchons le signal source $\mathbf{Y}$ sous la forme de l'équation (4.2), ce qui signifie que nous recherchons uniquement la séquence des matrices $\mathbf{W}(\ell)$ qui constituent le **filtre de séparation et de déconvolution** que nous appellerons par la suite le filtre de séparation. La question posée à ce niveau est de savoir comment faire pour déterminer d'une façon simple les matrices $\mathbf{V}(\ell)$ (qui constituent le filtre inverse du filtre de séparation) en fonction des matrices $\mathbf{W}(\ell)$? En d'autres termes, quelles relations lient les matrices $\mathbf{V}(\ell)$ aux matrices $\mathbf{W}(\ell)$? C'est à cette question que nous essayerons de donner des éléments de réponse par la suite.

4.3.1 Relations entre le filtre de séparation et son inverse

Le schéma de compression que nous avons décrit dans la section 4.2 est assez simple et nous avons supposé que le filtre de séparation utilisé pour la transformation réduisant la redondance est totalement inversible, d'où la relation :

$$\mathbf{X}(\mathbf{k}) = \sum_\ell \mathbf{V}(\ell)\mathbf{Y}(\mathbf{k}-\ell) \tag{4.14}$$

En remplaçant dans cette dernière équation l'expression de $\mathbf{Y}(\mathbf{k})$ par celle de l'équation (4.3), nous pouvons encore écrire cette première comme

$$\begin{aligned}\mathbf{X}(\mathbf{k}) &= \sum_\ell \sum_{\mathbf{n}} \mathbf{V}(\ell)\mathbf{W}(\mathbf{n})\mathbf{X}(\mathbf{k}-\ell-\mathbf{n}) \\ &= \sum_{\mathbf{m}} [\sum_\ell \mathbf{V}(\ell)\mathbf{W}(\mathbf{m}-\ell)]\mathbf{X}(\mathbf{k}-\mathbf{m})\end{aligned} \tag{4.15}$$

En observant l'expression de cette dernière équation, il apparaît qu'une relation liant les matrices $\mathbf{V}(\ell)$ aux matrices $\mathbf{W}(\ell)$ est :

$$\sum_{\ell} \mathbf{V}(\ell)\mathbf{W}(\mathbf{m}-\ell) = \delta_{\mathbf{m}} \tag{4.16}$$

où

$$\delta_{\mathbf{m}} = \begin{cases} \mathbf{I}_N \text{ si } \mathbf{m} = (0,0) \\ \mathbf{0}_N \text{ sinon} \end{cases}$$

Une autre relation entre les matrices $\mathbf{V}(\ell)$ et $\mathbf{W}(\ell)$ et qui est une conséquence directe de (4.16) est : $\widetilde{\mathbf{V}}(\mathbf{z})\widetilde{\mathbf{W}}(\mathbf{z}) = \mathbf{I}_N$, où $\widetilde{\mathbf{V}}(\mathbf{z}) = \sum_{\ell} \mathbf{V}(\ell)\mathbf{z}^{-\ell}$ et $\widetilde{\mathbf{W}}(\mathbf{z}) = \sum_{\ell} \mathbf{W}(\ell)\mathbf{z}^{-\ell}$ sont les transformées en $\mathbf{z}$ des séquences de matrices $\{\mathbf{V}(\ell)\}$ et $\{\mathbf{W}(\ell)\}$ respectivement, avec $\mathbf{z} = (z_1, z_2)$. Notons ici que nous ne nous soucions pas des problèmes d'inversibilité de $\mathbf{W}(\mathbf{z})$ et de $\mathbf{V}(\mathbf{z})$, nous supposons que ces deux filtres sont parfaitement inversibles. Dans [22] et dans [61], les conditions d'inversibilité des filtres matriciels sont énoncées et discutées dans une certaine mesure. Des discussions plus élaborées et plus complètes sur les filtres en général, et sur les conditions d'inversibilité en particulier se trouvent dans [74]. Il devient donc légitime de penser que le critère à minimiser peut s'exprimer comme une fonction dépendant uniquement de la séquence de matrices $\{\mathbf{W}(\ell)\}$. Avant d'en arriver à l'expression finale du critère, faisons d'abord quelques rappels.

D'une part, il est connu que la norme au carrée de Frobenius d'une matrice complexe $\mathbf{A}$ est définie comme : $||A||_F^2 = \mathrm{tr}(A^H A)$ où ^{H} désigne la transposée hermitienne et tr() désigne la trace d'une matrice. D'autre part, il est tout aussi bien connu d'après l'égalité de Parseval qu'il y'a conservation d'énergie d'un signal en passant du domaine temporel ou spatial au domaine fréquentiel, en d'autres termes la relation suivante est vérifiée pour le filtre constitué par la séquence de matrices $\{\mathbf{V}(\ell)\}$:

$$\sum_{\ell} \mathbf{V}^T(\ell)\mathbf{V}(\ell) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \widetilde{\mathbf{V}}^H(e^{-i\lambda})\widetilde{\mathbf{V}}(e^{-i\lambda}) d\lambda$$

avec

$$\widetilde{\mathbf{V}}(e^{-i\lambda}) = \sum_{\ell} \mathbf{V}(\ell) e^{-i\ell\lambda}$$

$\widetilde{\mathbf{V}}(e^{-i\lambda})$ est la transformée de Fourier de la séquence de matrices $\{\mathbf{V}(\ell)\}$. Par la suite, pour simplifier les notations, nous utiliserons $\widetilde{\mathbf{V}}(\lambda)$ au lieu de $\widetilde{\mathbf{V}}(e^{-i\lambda})$ pour désigner cette transformée de Fourier. Soulignons ici que nous utilisons les notations vectorielles pour les variables d'indice spatial et fréquentiel pour la clarté et la facilité de notation. Ainsi, on a $\lambda = (\lambda_1, \lambda_2)$ et $d\lambda = d\lambda_1 d\lambda_2$. Dans la suite, sauf cas exceptionnel que nous préciserons, nous adoptons et utilisons ces notations. Une conséquence de la relation liant les matrices $\mathbf{V}(\ell)$ aux $\mathbf{W}(\ell)$ est que nous avons $\widetilde{\mathbf{V}}(\lambda) = \widetilde{\mathbf{W}}^{-1}(\lambda)$, avec $\widetilde{\mathbf{W}}(\lambda) = \sum_{\ell} \mathbf{W}(\ell) e^{-i\ell\lambda}$. Nous en déduisons alors

$$\sum_{\ell} \mathbf{V}^T(\ell)\mathbf{V}(\ell) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda) d\lambda \tag{4.17}$$

4.3.2 Définition du critère à minimiser

Dans la section précédente, nous avons défini un critère traduit par l'équation (4.13) à minimiser en compression d'image en supposant que nous avons un modèle de mélange convolutif. Ce critère

dépend du filtre de transformation inverse. En se servant de l'équation (4.17), ce critère peut encore s'écrire comme

$$C(\{\mathbf{W}(\ell)\}) = \sum_{j=1}^{N} H(Y_j) + \frac{1}{2}\log_2 \det \operatorname{diag}[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\frac{d\lambda}{(2\pi)^2}] \tag{4.18}$$

Ce critère est une fonction qui dépend essentiellement des matrices $\mathbf{W}(\ell)$. En effet, connaissant le filtre de séparation, il devient simple de calculer l'image transformée en se servant de l'équation (4.3). L'entropie différentielle étant une caractéristique propre à une variable aléatoire, le critère dépend donc essentiellement du filtre recherché constitué par la séquence des matrices $\{\mathbf{W}(\ell)\}$. Ce critère à été mis à jour dans [8] où nous avons défini et explicité le calcul de son gradient.

Dans le cas de la séparation-déconvolution aveugle (cas qui ne nous intéresse pas en compression), il est démontré dans [61] que le critère à minimiser en ne se basant que sur la réduction de l'information mutuelle (aussi bien entre coefficients transformés dans une composante de $\mathbf{Y}$ qu'entre les différentes composantes de celle-ci) est donné par :

$$C_{ICA}(\{\mathbf{W}(\ell)\}) = \sum_{j=1}^{N} H(Y_j) - \int_0^{2\pi}\int_0^{2\pi} \log_2 |\det \widetilde{\mathbf{W}}(\lambda)| \frac{d\lambda}{(2\pi)^2} \tag{4.19}$$

Par la suite, nous établissons une relation, entre ce critère et celui défini à l'équation (4.18).

4.3.3 Lien avec le critère de séparation-déconvolution basé sur l'information mutuelle

Le critère défini par l'équation (4.18) peut encore s'écrire de la façon suivante :

$$\begin{aligned} C &= \sum_{j=1}^{N} H(\mathbf{Y}_j) - \int_0^{2\pi}\int_0^{2\pi} \log_2 |\det \widetilde{\mathbf{W}}(\lambda)| \frac{d\lambda}{(2\pi)^2} + \\ & \frac{1}{2}\log_2 \det \operatorname{diag}[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\frac{d\lambda}{(2\pi)^2}] - \\ & \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} \log_2 \det[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]\frac{d\lambda}{(2\pi)^2} \end{aligned} \tag{4.20}$$

Ainsi ce critère peut se mettre sous la forme $C = C_{ICA} + P$ avec :

$$\begin{aligned} P &= \frac{1}{2}\log_2 \det \operatorname{diag}\left[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\frac{d\lambda}{(2\pi)^2}\right] - \\ & \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} \log_2 \det[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]\frac{d\lambda}{(2\pi)^2} \end{aligned} \tag{4.21}$$

Proposition 4.3.1 *Le terme P définie dans (4.21) est un nombre réel positif ou nul et n'est nul que si et seulement si $\widetilde{\boldsymbol{W}}^{-H}(\lambda)\widetilde{\boldsymbol{W}}^{-1}(\lambda)$ est une matrice diagonale et constante, indépendante de λ.*

Démonstration :

Soit X une variable aléatoire réelle, il est connu en se basant sur l'inégalité de Jensen [27] (page 25) que pour toute fonction f convexe on a :

$$E(f(X)) \geq f(E(X)) \tag{4.22}$$

avec égalité si et seulement si $f(X)$ est constante. Considérons le premier terme P_1 dans l'expression de P de (4.21), il peut encore se mettre sous la forme :

$$P_1 = \frac{1}{2}\sum_{j=1}^{N} \log_2 \int_0^{2\pi}\int_0^{2\pi} [\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jj} \frac{d\lambda}{(2\pi)^2} \tag{4.23}$$

avec $[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jj}$ le $j^{ème}$ terme diagonal de la matrice $\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)$. Cette nouvelle expression se déduit facilement en utilisant la propriété des fonctions logarithmiques qui transforme le logarithme d'un produit d'éléments en une somme des logarithmes de ces éléments. Si nous appliquons l'inégalité de Jensen à la fonction $-\log_2$, nous obtenons l'inégalité suivante :

$$\log_2 \int_0^{2\pi}\int_0^{2\pi} [\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jj} \frac{d\lambda}{(2\pi)^2} \geq \int_0^{2\pi}\int_0^{2\pi} \log_2 [\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jj} \frac{d\lambda}{(2\pi)^2} \tag{4.24}$$

avec l'égalité si et seulement si $[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jj}$ est une constante. Afin appliquer l'inégalité de (4.22), nous avons posé $X = [\widetilde{\mathbf{W}}^{-H}(\Lambda)\widetilde{\mathbf{W}}^{-1}(\Lambda)]_{jj}$, où Λ est un vecteur de loi uniforme sur $[0, 2\pi]^2$. Nous pouvons alors déduire que le premier terme P_1 de P vérifie l'inégalité suivante :

$$P_1 \geq \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} \log_2 \det \operatorname{diag}[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)] \frac{d\lambda}{(2\pi)^2} \tag{4.25}$$

Cette inégalité sur P_1 est intéressante dans la mesure où elle nous permet de lui donner une limite inférieure. En se servant de cette inégalité, nous déduisons aussi que :

$$\begin{aligned} P \quad \geq \quad & \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} \log_2 \det\{\operatorname{diag}[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]\} \frac{d\lambda}{(2\pi)^2} \\ - \quad & \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} \log_2 \det\{[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]\} \frac{d\lambda}{(2\pi)^2} \end{aligned} \tag{4.26}$$

Cette dernière inégalité se réduit plus simplement en

$$P \geq \frac{1}{2}\int_0^{2\pi}\int_0^{2\pi} \log_2 \left\{ \frac{\det \operatorname{diag}[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]}{\det[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]} \right\} \frac{d\lambda}{(2\pi)^2} \tag{4.27}$$

Cette nouvelle écriture est plus simple et nous permet d'appliquer l'inégalité d'Hadamard [27] (page 502) qui stipule que pour toute matrice $\mathbf{M}$ hermitienne semi-définie positive, on a $\det(\mathbf{M}) \leq \det \operatorname{diag}(\mathbf{M})$ avec égalité si et seulement si $\mathbf{M}$ est une matrice diagonale. Nous en déduisons donc que $P \geq 0$ car :

$$\frac{\det \operatorname{diag}[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]}{\det[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]} \geq 1,$$

d'où la démonstration de la première partie de la proposition qui stipule que le terme P est un réel toujours positif ou nul.

Le cas particulier $P = 0$ n'intervient que lorsque la matrice $\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)$ est diagonale et constante, c'est-à-dire indépendante de λ. Dans ce cas en effet, l'inégalité d'Hadamard se transforme en une égalité et, $\det \operatorname{diag}[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)] = \det[\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]$. Ce qui parachève la démonstration de la proposition 4.3.1. ■

La solution dans le cas particulier $P = 0$ peut être recherchée parmi les filtres passe-tout. Les filtres passe-tout sont aussi appelés filtres para-unitaires. En effet, d'après [74] (page 23), le filtre constitué par l'ensemble des matrices $\mathbf{W}(\ell)$ est dit para-unitaire si l'une des conditions équivalentes suivantes est satisfaite :

- $\sum \mathbf{W}(\ell)\mathbf{W}(\ell - \mathbf{k})^H = \delta_{\mathbf{k}}\mathbf{Id}$
- $\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}(\lambda)^H = \mathbf{Id}$

Nous pouvons ainsi faire un parallèle avec les travaux de Narozny [51] [52] dans lesquels il a démontré que le critère à minimiser en compression était lié au critère de l'ACI (en prenant l'information mutuelle comme critère à minimiser) en utilisant un modèle de mélange instantané (voir section 2.4). La proposition ci-dessus démontrée est très intéressante et nous montre qu'un algorithme de séparation-deconvolution tel que celui défini dans [61] n'est pas adapté à la compression des images multi-composantes. La relation obtenue ici est similaire à celle obtenue en supposant un mélange instantané, c'est-à-dire que le critère à minimiser en compression peut s'écrire comme le critère utilisé dans le cadre de l'ACI convolutif (en prenant l'information mutuelle comme critère à minimiser) auquel s'ajoute un terme réel qui est toujours positif ou nul. Ce dernier terme dans notre cas correspond à une sorte de distance à la "para-unitarité" du filtre.

4.3.4 Invariance du critère

Théorème 4.3.1 *Le critère défini par l'équation (4.18) est invariant par multiplication à gauche de son argument par une matrice diagonale réelle de déterminant non nul.*

Démonstration : Soit $\mathbf{D}$ une matrice carrée diagonale, notons d_i le $i^{ème}$ élément diagonal de cette matrice et supposons que $\forall i = 1 \dots N, d_i \neq 0$. Soient $\{\mathbf{W}(\ell)\}$ et $\{\mathbf{R}(\ell)\}$ deux séquences matricielles telles que $\mathbf{R}(\ell) = \mathbf{D}\mathbf{W}(\ell)$. Nous allons démontrer que $C(\{\mathbf{W}(\ell)\}) = C(\{\mathbf{R}(\ell)\})$.

Soient $\mathbf{Y}(\mathbf{k}) = \sum_\ell \mathbf{W}(\ell)\mathbf{X}(\mathbf{k}-\ell)$ et $\mathbf{Z}(\mathbf{k}) = \sum_\ell \mathbf{R}(\ell)\mathbf{X}(\mathbf{k}-\ell)$, les images transformées obtenues après filtrage de l'image $\mathbf{X}$ par les séquences $\{\mathbf{W}(\ell)\}$ et $\{\mathbf{R}(\ell)\}$ respectives. $\mathbf{D}$ étant une matrice diagonale, nous pouvons encore écrire tout simplement que $\mathbf{Z}(\mathbf{k}) = \mathbf{D}\mathbf{Y}(\mathbf{k})$, en d'autres termes, nous avons $\mathbf{Z}_i(\mathbf{k}) = d_i\mathbf{Y}_i(\mathbf{k})$. Ce qui nous permet d'avoir une relation entre les entropies marginales des composantes de $\mathbf{Z}$ et de $\mathbf{Y}$ et qui se traduit par :

$$\sum_{i=1}^{N} H(Z_i) = \sum_{i=1}^{N} H(Y_i) + \sum_{i=1}^{N} \log_2 |d_i| \tag{4.28}$$

Le second terme du critère de l'équation (4.18) pour le filtre $\{\mathbf{R}(\ell)\}$ est :

$$A = \frac{1}{2}\log_2 \det \operatorname{diag}\left[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{R}}^{-H}(\lambda)\widetilde{\mathbf{R}}^{-1}(\lambda)\frac{d\lambda}{(2\pi)^2}\right] \tag{4.29}$$

Appelons A_0 son homologue pour le filtre $\{\mathbf{W}(\ell)\}$ et qui est défini par :

$$A_0 = \frac{1}{2}\log_2 \det \operatorname{diag}\left[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\frac{d\lambda}{(2\pi)^2}\right]$$

Compte tenu du fait que $\widetilde{\mathbf{R}}(\lambda) = \mathbf{D}\widetilde{\mathbf{W}}(\lambda)$, nous pouvons écrire l'expression de A comme $A = A_0 - \frac{1}{2}\log_2 \prod_{i=1}^{N} d_i^2$, et qui peut encore se mettre sous la forme $A = A_0 - \sum_{i=1}^{N} \log_2 |d_i|$. En regroupant cette dernière expression qui nous donne le second terme du critère de l'équation (4.18) avec celle de l'équation (4.28), nous obtenons

$$\begin{aligned} C(\{\mathbf{R}(\ell)\}) &= \sum_{i=1}^{N} H(\mathbf{Y}_i) + \sum_{i=1}^{N} \log_2 |d_i| + A_0 - \sum_{i=1}^{N} \log_2 |d_i| && (4.30)\\ &= \sum_{i=1}^{N} H(\mathbf{Y}_i) + A_0 && (4.31)\\ &= C(\{\mathbf{W}(\ell)\}) && (4.32) \end{aligned}$$

ce qui parachève la démonstration du théorème. ■

Le fait que le critère soit invariant par multiplication à gauche de son argument nous montre que la minimisation du critère ne fournit pas à une solution unique, ce qui crée une ambiguïté pour la recherche d'une solution minimisant le critère. Cependant, il est toujours possible d'ajouter des contraintes pour que la solution du problème devienne unique.

Conclusion : Dans ce chapitre, nous avons défini un système de compression assez simple utilisant une transformation de type convlutif, et nous avons mis à jour un nouveau critère adapté à la compression des images. Ce nouveau critère est invariant par multiplication à gauche de son argument par une matrice diagonale réelle non singulière, ce qui exclut l'unicité de la solution pour la minimisation de ce critère. Nous avons trouvé une relation simple entre ce critère et le critère utilisé dans [61] pour la séparation-déconvolution aveugle. Par la suite, nous proposons un algorithme permettant de minimiser ce critère en utilisant une variante de la méthode de descente du gradient de Newton.

Chapitre 5

Minimisation du critère

Afin de minimiser le critère défini par l'équation (4.18), nous nous proposons d'utiliser la méthode BFGS du nom de ces auteurs *Broyden-Fletcher-Goldfarb-Shanno* qui est une variante de la méthode de descente du gradient de Newton. Pour cela, nous avons besoin d'évaluer le gradient de ce critère. Dans ce chapitre, nous commençons par faire un développement de Taylor du critère avec pour objectif, le calcul de son gradient. Par la suite, nous explicitons deux algorithmes, l'un minimisant le critère dans le cas général, l'autre minimisant le critère sous une contrainte que nous définirons. Enfin, nous faisons une étude de la complexité de ces algorithmes.

5.1 Calcul du gradient du critère

Soient C_1 et C_2 le premier et le deuxième terme du membre de droite de l'équation (4.18) qui définit le critère à minimiser. Nous allons effectuer un développement au second ordre de Taylor de ces deux termes pour déterminer le gradient du critère. Nous justifierons par la suite, les raisons du choix d'une méthode BFGS où une estimation utilisant des formules empiriques de l'hessien est faite à chaque itération.

5.1.1 Développement de Taylor au second ordre de C_1

Nous rappelons que : $C_1(\{\mathbf{W}(\ell)\}) = \sum_{j=1}^{N} H(Y_j)$. Soit X une variable aléatoire et Δ un petit incrément de X, dans [62], il est démontré que le développement au second ordre de l'entropie différentielle de $X + \Delta$ est obtenu par l'équation :

$$H(X+\Delta) \approx H(X) + \frac{1}{\log 2} E[\Delta \psi_X(X)] + \frac{1}{2\log 2} E\{[\Delta - E(\Delta|X)]^2 \psi_X'(X) - \frac{1}{\log 2}[E(\Delta|X)]'^2\} \quad (5.1)$$

avec $H(.)$ l'entropie différentielle d'une variable aléatoire, et $\psi_X = -(\log f_X)'$ est la fonction score de X, f_X étant la densité de probabilité de X. Si les matrices $\mathbf{W}(\ell)$ subissent une légère variation $\delta\mathbf{W}(\ell)$, cela se traduit au niveau de $\mathbf{Y}$ par une variation $\delta\mathbf{Y}(\mathbf{k}) = \sum_\ell \delta\mathbf{W}(\ell)\mathbf{X}(\mathbf{k}-\ell)$. Soit $\delta W_{ij}(\ell)$ l'élément (i,j) de la matrice $\delta\mathbf{W}(\ell)$. La variation de la $j^{ème}$ composante de $\mathbf{Y}$ est obtenue par $\delta Y_j(\mathbf{k}) = \sum_\ell \sum_i \delta W_{ji}(\ell) X_i(\mathbf{k}-\ell)$.

Par la suite, nous ne reprendrons pas totalement ces développements, nous utiliserons les résultats des développements explicités en détails dans [61]. Si nous supposons la propriété de stationnarité dans chacune des composantes $\mathbf{Y}_j$, alors la densité de probabilité f_{Y_j} ne dépend pas de la

position spatiale qui ici est $\mathbf{k}$. Ainsi le développement de Taylor au premier ordre de $H(Y_j+\delta Y_j)$ est : $\frac{1}{\log 2}\sum_\ell \sum_{i=1}^N E\{\psi_j[Y_j(0,0)]\delta W_{ji}(\ell)\mathbf{X}_i(-\ell)\}$. Nous en déduisons que le développement au premier ordre du terme C_1 vaut :

$$\frac{1}{\log 2}\,\mathrm{tr}\sum_\ell \delta\mathbf{W}(\ell)E\{\mathbf{X}(-\ell)\psi[\mathbf{Y}(\mathbf{0})]^T\} \tag{5.2}$$

où $\psi[\mathbf{Y}(\mathbf{0})]$ est le vecteur de composantes $\psi_1[Y_1(\mathbf{0})], \psi_2[Y_2(\mathbf{0})], \ldots, \psi_N[\mathbf{Y}_N(0)]^T$ et ψ_i est la fonction score associée à Y_i.

Compte tenu du fait que $E\{\mathbf{X}(\ell)\psi[\mathbf{Y}(0)]^T\} = \int_0^{2\pi}\int_0^{2\pi}\Gamma_{\mathbf{X},\psi(\mathbf{Y})}(\lambda)e^{il\lambda}d\lambda$ où

$$\Gamma_{\mathbf{X},\psi(\mathbf{Y})}(\lambda) = \frac{1}{(2\pi)^2}\sum_\ell E\{\mathbf{X}(\ell)\psi[\mathbf{Y}(0)]^T\}e^{-i\ell\lambda}$$

est l'interspectre entre les processus $\{\mathbf{X}(\ell)\}$ et $\{\psi[\mathbf{Y}(\ell)]\}$, nous pouvons encore écrire l'expression de l'équation (5.2) comme

$$\frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi}\mathrm{tr}[\delta\widetilde{\mathbf{W}}(\lambda)\Gamma_{\mathbf{X},\psi(\mathbf{Y})}(\lambda)]d\lambda \tag{5.3}$$

Sachant que $\Gamma_{\mathbf{X},\psi(\mathbf{Y})}(\lambda) = \widetilde{\mathbf{W}}^{-1}(\lambda)\Gamma_{\mathbf{Y},\psi(\mathbf{Y})}(\lambda)$ de part l'opération de convolution qui lie $\mathbf{X}$ et $\mathbf{Y}$, où $\Gamma_{\mathbf{Y},\psi(\mathbf{Y})}(\lambda)$ est l'interspectre entre les processus $\{\mathbf{Y}(\ell)\}$ et $\{\psi[\mathbf{Y}(\ell)]\}$, l'expression de l'équation (5.3) vaut finalement

$$\frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi}\mathrm{tr}[\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\Gamma_{\mathbf{Y},\psi(\mathbf{Y})}(\lambda)]d\lambda$$

Le terme du second ordre étant assez compliqué, pour simplifier les développements, nous supposons que les hypothèses $H3$ et $H4$ sont vérifiées c'est-à-dire que nous avons à la fois une indépendance spectrale entre les composantes de $\mathbf{Y}$, mais aussi une indépendance spatiale, entre les coefficients d'une composante $\mathbf{Y}_i$, $1 \leq i \leq N$. Ceci est vérifié lorsque la solution obtenue est proche de celle de la séparation-déconvolution aveugle, proche au sens où ces deux hypothèses sont valides. La séquence $\{\mathbf{V}(\ell)\}$ étant l'inverse de $\{\mathbf{W}(\ell)\}$ au sens convolutif du terme, d'après [62], le terme du second ordre de $H(Y_j+\delta Y_j)$ est :

$$\frac{1}{2\log 2}E\psi_j'(Y_j)\sum_{k=1}^N E(Y_k^2)\int_0^{2\pi}\int_0^{2\pi}|[\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jk}|^2\frac{d\lambda}{(2\pi)^2} - \frac{1}{2\log 2}\{E(Y_j^2)E[\psi'(Y_j)]+1\}(\overline{\delta\mathbf{W}\mathbf{W}^{-1}})_{jj}^2$$

avec $(\overline{\delta\mathbf{W}\mathbf{W}^{-1}})_{jj} = [\delta\mathbf{W}(\ell)\star\mathbf{V}(\ell)]_{jj}(0) = \int_0^{2\pi}\int_0^{2\pi}[\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jj}\frac{d\lambda}{(2\pi)^2}$, $\star$ ici désignant l'opération de convolution entre les séquences de matrices. Partant de là, nous pouvons finalement écrire le développement au second ordre du terme C_1 du critère comme :

$$\begin{aligned} C_1(\widetilde{\mathbf{W}}+\delta\widetilde{\mathbf{W}}) \;=\; & C_1(\widetilde{\mathbf{W}}) + \frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi}\mathrm{tr}\{\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\Gamma_{\mathbf{Y},\psi(\mathbf{Y})}(\lambda)\}d\lambda + \\ & \frac{1}{2\log 2}\sum_{j,k=1}^N\int_0^{2\pi}\int_0^{2\pi}E[\psi_j'(Y_j)]E(Y_k^2)|[\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jk}|^2\frac{d\lambda}{(2\pi)^2} - \\ & \frac{1}{2\log 2}\sum_{j=1}^N\{E[\psi_j'(Y_j)]E(Y_j^2)+1\}(\overline{\delta\mathbf{W}\mathbf{W}^{-1}})_{jj}^2 \end{aligned} \tag{5.4}$$

Il est facile de vérifier (une démonstration est faite dans [44] en pp $33-34$) en faisant une intégration par partie, que $E[\psi_j^{'}(Y_j)] = E[\psi_j^2(Y_j)]$, ce qui n'est rien d'autre que l'information de Fischer. Ce développement de Taylor au second ordre de C_1 nous sera utile par la suite, nous nous servirons des termes au premier ordre pour déterminer le gradient de notre critère à minimiser.

5.1.2 Développement de Taylor au second ordre de C_2

Nous rappelons ici que le deuxième terme C_2 du critère est défini par

$$C_2(\{\mathbf{W}(\ell)\}) = \frac{1}{2}\log_2 \det \operatorname{diag}[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\frac{d\lambda}{(2\pi)^2}] \tag{5.5}$$

Cette expression peut se mettre sous la forme suivante :

$$C_2(\{\mathbf{W}(\ell)\}) = \frac{1}{2}\sum_{j=1}^{N}\log_2 \int_0^{2\pi}\int_0^{2\pi} [\widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]_{jj}\frac{d\lambda}{(2\pi)^2}\}$$

Supposons avoir fait subir une petite variation $\delta\widetilde{\mathbf{W}}(\lambda) = \widetilde{\varepsilon}(\lambda)\widetilde{\mathbf{W}}(\lambda)$ à $\widetilde{\mathbf{W}}(\lambda)$ comme suit :

$$\widetilde{\mathbf{W}}(\lambda) \longleftarrow \widetilde{\mathbf{W}}(\lambda) + \widetilde{\varepsilon}(\lambda)\widetilde{\mathbf{W}}(\lambda)$$

Le développement de Taylor ici est fait en supposant une petite variation sur le plan spectral. Il est facile de basculer du plan spatial au plan spectral et vice versa de par les propriétés de la transformation de Fourier. Nous utilisons le plan spectral car il permet de simplifier les développements et les équations. En posant $\widetilde{\mathbf{M}}(\lambda) = \widetilde{\mathbf{W}}^{-H}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)$, le développement de l'expression dans la double intégrale de l'équation (5.5) est donné par :

$$[\widetilde{\mathbf{W}}(\lambda) + \widetilde{\varepsilon}(\lambda)\widetilde{\mathbf{W}}(\lambda)]^{-H}[\widetilde{\mathbf{W}}(\lambda) + \widetilde{\varepsilon}(\lambda)\widetilde{\mathbf{W}}(\lambda)]^{-1} = [\mathbf{I} + \widetilde{\varepsilon}^H(\lambda)]^{-1}\widetilde{\mathbf{M}}(\lambda)[\mathbf{I} + \widetilde{\varepsilon}(\lambda)]^{-1}$$

En faisant un développement du membre de droite de cette dernière expression au second ordre nous obtenons :

$$\widetilde{\mathbf{M}}(\lambda) - \widetilde{\mathbf{M}}(\lambda)\widetilde{\varepsilon}(\lambda) - \widetilde{\varepsilon}^H(\lambda)\widetilde{\mathbf{M}}(\lambda) + \widetilde{\varepsilon}^H(\lambda)\widetilde{\mathbf{M}}(\lambda)\widetilde{\varepsilon}(\lambda) + \widetilde{\mathbf{M}}(\lambda)\widetilde{\varepsilon}^2(\lambda) + \widetilde{\varepsilon}^{2H}(\lambda)\widetilde{\mathbf{M}}(\lambda) \tag{5.6}$$

Le $j^{ème}$ terme diagonal de la matrice obtenu par l'expression de l'équation (5.6) vaut :

$$\begin{aligned}\widetilde{\mathbf{M}}_{jj}(\lambda) \quad - \quad & \sum_{k=1}^{N}\widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda) - \sum_{k=1}^{N}\widetilde{\mathbf{M}}_{kj}(\lambda)\widetilde{\varepsilon}^*_{kj}(\lambda) + \sum_{i,k=1}^{N}\widetilde{\varepsilon}^*_{ij}(\lambda)\widetilde{\mathbf{M}}_{ik}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda) + \\ & \sum_{i,k=1}^{N}\widetilde{\varepsilon}^*_{ij}(\lambda)\widetilde{\varepsilon}^*_{ki}(\lambda)\widetilde{\mathbf{M}}_{kj}(\lambda) + \sum_{i,k=1}^{N}\widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{ki}(\lambda)\widetilde{\varepsilon}_{ij}(\lambda)\end{aligned}$$

L'objectif dans cette section est de pouvoir exprimer le développement de Taylor au second ordre du terme C_2 ci-dessus et défini du critère de l'équation (4.18). Pour arriver à nos fins, nous appelons

A_j, $\delta A_{1,j}$ et $\delta A_{2,j}$ les expressions suivantes :

$$\begin{aligned}
A_j &= \int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{M}}_{jj}(\lambda)\frac{d\lambda}{(2\pi)^2} \\
\delta A_{1,j} &= -\int_0^{2\pi}\int_0^{2\pi} \{\sum_{k=1}^{N} \widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda) + \sum_{k=1}^{N} \widetilde{\mathbf{M}}_{kj}(\lambda)\widetilde{\varepsilon}^*_{kj}(\lambda)\}\frac{d\lambda}{(2\pi)^2} \\
\delta A_{2,j} &= \int_0^{2\pi}\int_0^{2\pi} \{\sum_{i,k=1}^{N} \widetilde{\varepsilon}^*_{ij}(\lambda)\widetilde{\mathbf{M}}_{ik}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda) + \sum_{i,k=1}^{N} \widetilde{\varepsilon}^*_{ij}(\lambda)\widetilde{\varepsilon}^*_{ki}(\lambda)\widetilde{\mathbf{M}}_{kj}(\lambda) \\
&+ \sum_{i,k=1}^{N} \widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{ki}(\lambda)\widetilde{\varepsilon}_{ij}(\lambda)\}\frac{d\lambda}{(2\pi)^2}
\end{aligned}$$

En observant les éléments de $\delta A_{1,j}$, on remarque que les termes dans la double intégrale sont le conjugué l'un de l'autre ; d'autre part, la double intégrale est faite sur le plan spectral entre 0 et 2π, et nous savons que dans ce domaine là, les éléments sont le conjugué l'un de l'autre entre 0 et 2π par symétrie autour de π puisque notre signal de départ $\mathbf{X}$ est à valeurs réelles. Ainsi, les expressions de $\delta A_{1,j}$ et $\delta A_{2,j}$ se simplifient et deviennent :

$$\begin{aligned}
\delta A_{1,j} &= -2\int_0^{2\pi}\int_0^{2\pi} \sum_{k=1}^{N} \widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda)\frac{d\lambda}{(2\pi)^2} \\
\delta A_{2,j} &= \int_0^{2\pi}\int_0^{2\pi} \{\sum_{i,k=1}^{N} \widetilde{\varepsilon}^*_{ij}(\lambda)\widetilde{\mathbf{M}}_{ik}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda) + 2\sum_{i,k=1}^{N} \widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{ki}(\lambda)\widetilde{\varepsilon}_{ij}(\lambda)\}\frac{d\lambda}{(2\pi)^2}
\end{aligned}$$

Pour obtenir le développement de Taylor au second ordre de $C_2(\{\mathbf{W}(\ell)\})$, nous supposons avoir eu une petite variation δA_j de A_j, avec $\delta A_j = \delta A_{1,j} + \delta A_{2,j}$. Les termes $\delta A_{1,j}$ et $\delta A_{2,j}$ représentent respectivement les petites variations au premier et au second ordre autour de A_j. Ainsi, en s'arrêtant au second ordre, nous avons le développement de Taylor suivant :

$$\log_2(A_j + \delta A_j) = \log_2(A_j) + \frac{1}{\log 2}\frac{\delta A_j}{A_j} - \frac{1}{2\log 2}\frac{\delta A_{1,j}^2}{A_j^2}$$

Ce qui nous permet de déduire que le développement de Taylor au second ordre du terme C_2 est :

$$C_2(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) = C_2(\widetilde{\mathbf{W}}) + \frac{1}{2\log 2}\sum_{j=1}^{N}\frac{\delta A_{1,j}}{A_j} + \frac{1}{2\log 2}\sum_{j=1}^{N}\frac{\delta A_{2,j}}{A_j} - \frac{1}{4\log 2}\sum_{j=1}^{N}\frac{\delta A_{1,j}^2}{A_j^2} \tag{5.7}$$

5.1.3 Calcul du gradient du critère

D'après ce qui précède, il est possible théoriquement de calculer le gradient et le hessien de notre critère car nous avons pu avoir son développement de Taylor au second ordre. Si nous observons les développements de Taylor au second ordre des deux termes C_1 et C_2, nous constatons que les expressions des termes au second ordre sont assez complexes et font intervenir les termes croisés des matrices $\widetilde{\varepsilon}(\lambda)$. Au lieu d'utiliser une méthode de minimisation quasi-Newton qui nécessite le

calcul effectif du gradient et du hessien à chaque itération, nous utilisons la méthode BFGS, qui permet d'avoir de très bonnes performances pour la minimisation d'un critère en général. Cette méthode ne nécessite pas le calcul effectif du hessien, ce dernier est estimé de façon empirique. Ainsi, nous aurons juste besoin de calculer le gradient par la suite pour définir l'algorithme BFGS nous permettant de minimiser le critère défini à l'équation (4.18). De ce qui précède, il ressort que le développement de Taylor du critère $C(\widetilde{\mathbf{W}})$ dans le cadre général en ne tenant compte que des éléments au premier ordre est :

$$\begin{aligned} C(\widetilde{\mathbf{W}}+\delta\widetilde{\mathbf{W}}) &= C(\widetilde{\mathbf{W}}) + \frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi} \mathrm{tr}\{\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\Gamma_{\mathbf{Y},\psi(\mathbf{Y})}(\lambda)\}d\lambda - \\ & \frac{1}{\log 2}\sum_{j=1}^{N}\frac{\int_0^{2\pi}\int_0^{2\pi}\sum_{k=1}^{N}\widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda)\frac{d\lambda}{(2\pi)^2}}{\int_0^{2\pi}\int_0^{2\pi}\widetilde{\mathbf{M}}_{jj}(\lambda)\frac{d\lambda}{(2\pi)^2}} \end{aligned} \tag{5.8}$$

En se rappelant des relations entre les interspectres, nous savons que $\Gamma_{\mathbf{X},\psi(\mathbf{Y})}(\lambda) = \widetilde{\mathbf{W}}^{-1}(\lambda)\Gamma_{\mathbf{Y},\psi(\mathbf{Y})}(\lambda)$. Ainsi le deuxième terme du membre de droite de l'équation (5.8) peut encore se mettre sous la forme suivante :

$$\frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi}\mathrm{tr}\{\delta\widetilde{\mathbf{W}}(\lambda)\Gamma_{\mathbf{X},\psi(\mathbf{Y})}(\lambda)\}d\lambda$$

Pour mieux comprendre les simplifications et développements qui vont suivre, rappelons quelques définitions.

Définitions : Soient $X(n)$ et $Y(n)$ deux signaux numériques réels mutuellement stationnaires au sens large. La fonction d'inter-corrélation entre X et Y notée R_{XY} se définit par :

$$R_{XY}(k) = E[X(n)Y(n-k)]$$

L'interspectre entre X et Y, qui n'est rien d'autre que la transformée de Fourier discrète de la fonction d'inter-corrélation est définie par :

$$\Gamma_{XY}(\lambda) = \frac{1}{(2\pi)^2}\sum_{\mathbf{k}} R_{XY}(\mathbf{k})e^{-i\mathbf{k}\lambda}$$

Compte tenu des définitions précédentes, par transformé de Fourier inverse appliquée au deuxième terme du membre de droite de l'equation (5.8), nous obtenons :

$$\frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi}\mathrm{tr}\{\delta\widetilde{\mathbf{W}}(\lambda)\Gamma_{\mathbf{X},\psi(\mathbf{Y})}(\lambda)\}d\lambda = \frac{1}{\log 2}\,\mathrm{tr}\sum_{\ell}\delta\mathbf{W}(\ell)R_{\mathbf{X},\psi(\mathbf{Y})}(-\ell) \tag{5.9}$$

Notons $\mathbf{D} = \mathrm{diag}[\int_0^{2\pi}\int_0^{2\pi}\widetilde{\mathbf{M}}(\lambda)\frac{d\lambda}{(2\pi)^2}]$ la matrice diagonale dont le $j^{ème}$ terme diagonal est $\int_0^{2\pi}\int_0^{2\pi}\widetilde{\mathbf{M}}_{jj}(\lambda)\frac{d\lambda}{(2\pi)^2}$. En l'insérant dans le troisième et dernier terme du membre de droite de l'équation (5.8), celui-ci devient :

$$-\frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi}\mathrm{tr}[\mathbf{D}^{-1}\widetilde{\mathbf{M}}(\lambda)\widetilde{\varepsilon}(\lambda)]\frac{d\lambda}{(2\pi)^2} \tag{5.10}$$

Nous rappelons ici que nous avons préalablement posé $\widetilde{\varepsilon}(\lambda) = \delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)$, en tenant compte de cette égalité dans l'expression de (5.10), cette dernière devient :

$$-\frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi} \mathrm{tr}[\mathbf{D}^{-1}\widetilde{\mathbf{M}}(\lambda)\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)]\frac{d\lambda}{(2\pi)^2}$$

Par la suite, nous appliquons la propriété sur la trace d'un produit de matrices carrées $\mathrm{tr}(\mathbf{ABC}) = \mathrm{tr}(\mathbf{BCA}) = \mathrm{tr}(\mathbf{CAB})$ à l'expression précédente et celle-ci se transforme alors en :

$$-\frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi} \mathrm{tr}[\delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)\mathbf{D}^{-1}\widetilde{\mathbf{M}}(\lambda)]\frac{d\lambda}{(2\pi)^2}$$

Par définition de la transformée de Fourier, il est bien connu que $\delta\widetilde{\mathbf{W}}(\lambda) = \sum_\ell \delta\mathbf{W}(\ell)e^{-i\ell^T\lambda}$; ce qui nous permet finalement d'écrire l'expression de l'équation(5.10) comme suit :

$$-\frac{1}{\log 2}\,\mathrm{tr}\sum_\ell \delta\mathbf{W}(\ell)[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-1}(\lambda)\mathbf{D}^{-1}\widetilde{\mathbf{M}}(\lambda)e^{-i\ell^T\lambda}\frac{d\lambda}{(2\pi)^2}] \tag{5.11}$$

En remplaçant les expressions des deuxième et troisième membres de droite de l'équation (5.8) par celles de (5.9) et de (5.11) dans le développement de Taylor à l'ordre 1 du critère, cela nous permet d'écrire la première équation citée comme :

$$\begin{aligned} C(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) \;&\simeq\; C(\widetilde{\mathbf{W}}) + \frac{1}{\log 2}\,\mathrm{tr}\sum_\ell \delta\mathbf{W}(\ell)\mathbf{R}_{\mathbf{X},\psi(\mathbf{Y})}(-\ell) \\ &-\; \frac{1}{\log 2}\,\mathrm{tr}\sum_\ell \delta\mathbf{W}(\ell)[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-1}(\lambda)\mathbf{D}^{-1}\widetilde{\mathbf{M}}(\lambda)e^{-i\ell^T\lambda}\frac{d\lambda}{(2\pi)^2}] \end{aligned} \tag{5.12}$$

Ce développement du premier ordre peut encore se mettre sous la forme plus intéressante

$$C(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) = C(\widetilde{\mathbf{W}}) + \sum_\ell \mathrm{tr}[\delta\mathbf{W}(\ell)\mathbf{G}^T(\ell)]$$

avec

$$\mathbf{G}(\ell) = \frac{1}{\log 2}\mathbf{R}^T_{\mathbf{X},\psi(\mathbf{Y})}(-\ell) - \frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{M}}^T(\lambda)\mathbf{D}^{-1}\widetilde{\mathbf{W}}^{-T}(\lambda)e^{-i\ell^T\lambda}\frac{d\lambda}{(2\pi)^2} \tag{5.13}$$

$\{\mathbf{G}(\ell)\}$ n'étant rien d'autre que le gradient de notre critère à minimiser. Nous venons ainsi de calculer, d'une façon simple et explicite, le gradient du critère de l'équation (4.18). Par la suite, nous allons définir le critère sous la contrainte $\mathbf{D} = \mathbf{I}$ et calculer le gradient sous cette contrainte.

5.2 Définition du critère sous contrainte, calcul de son gradient

5.2.1 Définition du critère sous contrainte

La matrice diagonale $\mathbf{D}$ est intéressante à juste titre car les termes diagonaux de cette matrice ne sont rien d'autre que les coefficients de pondération w_j permettant d'avoir une approximation de la formule de la distorsion de l'image originale comme le montre les équations (4.10), (4.13) et

(4.18) par rapport à la distorsion de l'image transformée. Résoudre le problème de minimisation du critère en compression sous la contrainte que la matrice $\mathbf{D}$ est la matrice identité est équivalent à imposer que la distorsion calculée sur l'image transformée $\mathbf{Y}$ est la même que celle sur l'image de départ $\mathbf{X}$, du moins lorsque la contrainte est atteinte. Ce qui facilite également l'allocation optimale de bits à haute résolution où la distorsion comme nous l'avons démontré devient la même pour chaque composante pour une distorsion cible D_c. D'autre part, la minimisation du critère sous cette contrainte nous affranchit de l'ambiguïté sur la non unicité de la solution liée au fait que le critère est invariant par multiplication à gauche de son argument par une matrice diagonale non singulière. De façon plus précise, $C(\{\Delta\mathbf{W}(\ell)\}) = C(\{\mathbf{W}(\ell)\})$ pour toute matrice diagonale Δ non singulière. Soit maintenant une fonction $F : \{\mathbf{W}(\ell)\} \mapsto F(\{\mathbf{W}(\ell)\})$ telle que pour toute séquence $\{\mathbf{W}(\ell)\}$ nous ayons

$$\min_{\Delta\ diagonal} F(\{\Delta\mathbf{W}(\ell)\}) = \min_{\{\mathbf{W}'(\ell)\}} F(\{\mathbf{W}'(\ell)\})$$

Dans ce cas, la séquence de matrices $\{\mathbf{W}_*(\ell)\}$ qui minimise $C(\{\mathbf{W}(\ell)\})+F(\{\mathbf{W}(\ell)\})$ minimise également à la fois $C(\{\mathbf{W}(\ell)\})$ et $F(\{\mathbf{W}(\ell)\})$. En effet, si $\{\mathbf{W}_*(\ell)\}$ ne minimise pas $C(\{\mathbf{W}(\ell)\})$, il existe une séquence de matrice $\{\mathbf{W}(\ell)\} \neq \{\mathbf{W}_*(\ell)\}$ telle que $C(\{\mathbf{W}_*(\ell)\}) > C(\{\mathbf{W}(\ell)\})$ et d'après la propriété de F, il existe une matrice diagonale Δ telle que $F(\{\Delta\mathbf{W}(\ell)\}) = \min_{\{\mathbf{W}'(\ell)\}} F(\{\mathbf{W}'(\ell)\}) \leq F(\{\mathbf{W}_*(\ell)\})$. Par suite $C(\{\mathbf{W}_*(\ell)\})+F(\{\mathbf{W}_*(\ell)\}) > C(\{\Delta\mathbf{W}(\ell)\})+F(\{\Delta\mathbf{W}(\ell)\})$, ce qui contredirait le fait que $\{\mathbf{W}_*(\ell)\}$ minimise $C(\{\mathbf{W}(\ell)\})+F(\{\mathbf{W}(\ell)\})$. De même, si $\{\mathbf{W}_*(\ell)\}$ ne minimise pas $F(\{\mathbf{W}(\ell)\}$, il existe une matrice diagonale Δ telle que $F(\{\Delta\mathbf{W}_*(\ell)\}) = \min_{\{\mathbf{W}(\ell)\}} F(\{\mathbf{W}(\ell)\}) < F(\{\mathbf{W}_*(\ell)\})$ et donc $C(\{\Delta\mathbf{W}_*(\ell)\})+F(\{\Delta\mathbf{W}_*(\ell)\}) < C(\{\mathbf{W}_*(\ell)\})+F(\{\mathbf{W}_*(\ell)\})$, ce qui conduirait encore à une contradiction.

Le résultat précédent montre que minimiser $C(\{\mathbf{W}(\ell)\}) + F(\{\mathbf{W}(\ell)\})$ revient à minimiser $C(\{\mathbf{W}(\ell)\})$ sous la contrainte que $\{\mathbf{W}(\ell)\}$ appartient à l'ensemble de points minimisant F. Un choix intéressant de la fonction F serait $F(\{\mathbf{W}(\ell)\}) = [\mathrm{tr}(\mathbf{D}) - \log\det(\mathbf{D})]/(2\log 2)$ où $\mathbf{D} = \mathrm{diag}\int_0^{2\pi}\int_0^{2\pi}\widetilde{\mathbf{M}}(\lambda)d\lambda/(2\pi)^2$. On vérifie que F satisfait la condition requise, car $F(\{\Delta\mathbf{W}(\ell)\}) = [\mathrm{tr}(\Delta^{-2}\mathbf{D}) - \log\det(\Delta^{-2}\mathbf{D})]/(2\log 2)$ et que $\mathrm{tr}(\mathbf{D}) - \log\det\mathbf{D}$ est minimum quand $\mathbf{D}$ est la matrice identité. L'ensemble des points minimisant $F(\{\mathbf{W}(\ell)\})$ n'est autre que l'ensemble des séquences $\{\mathbf{W}(\ell)\}$, tel que la matrice $\mathbf{D}$ associée est la matrice identité. Par la suite, nous définissons le critère sous contrainte par $C_{contrainte}(\{\mathbf{W}(\ell)\}) = C(\{\mathbf{W}(\ell)\}) + F(\{\mathbf{W}(\ell)\})$, ce qui nous donne au final

$$C_{contrainte}(\{\mathbf{W}(\ell)\}) = \sum_{j=1}^{N} H(Y_j) + \frac{1}{2\log 2}\,\mathrm{tr}(\mathbf{D}) \tag{5.14}$$

Ce critère a l'air bien plus simple que le critère initial car il est bien plus simple de manipuler la trace d'une matrice que le logarithme de son déterminant.

5.2.2 Calcul du gradient du critère $C_{contrainte}$

Soient $C_{contrainte_1}$ et $C_{contrainte_2}$ les premier et deuxième termes respectifs du membre de droite de l'équation (5.14), en se servant des développements de la section 5.1 il devient simple de faire le développement de Taylor à l'ordre 1 du critère $C_{contrainte}(\{\mathbf{W}(\ell)\})$. En effet, $C_{contrainte_1} = C_1$ et son développement de Taylor a été fait précédemment, il ne nous reste plus qu'à faire le développement de Taylor à l'ordre 1 de $C_{contrainte_2} = \frac{1}{2\log 2}\,\mathrm{tr}(\mathbf{D})$. Comme dans la sous-section 5.1.2, nous supposons une petite variation de $\widetilde{\mathbf{W}}(\lambda)$ comme suit : $\widetilde{\mathbf{W}}(\lambda) \longleftarrow \widetilde{\mathbf{W}}(\lambda)+\widetilde{\varepsilon}(\lambda)\widetilde{\mathbf{W}}(\lambda)$ avec $\delta\widetilde{\mathbf{W}}(\lambda) = \widetilde{\varepsilon}(\lambda)\widetilde{\mathbf{W}}(\lambda)$.

Il a été démontré dans cette sous section que le $j^{ème}$ terme diagonal du développement de Taylor de l'expression dans la double intégrale de $\mathbf{D}$ au 1er ordre vaut :

$$\widetilde{\mathbf{M}}_{jj}(\lambda) - \sum_{k=1}^{N} \widetilde{\mathbf{M}}_{jk}(\lambda)\widetilde{\varepsilon}_{kj}(\lambda) - \sum_{k=1}^{N} \widetilde{\mathbf{M}}_{kj}(\lambda)\widetilde{\varepsilon}^{*}_{kj}(\lambda)$$

Nous en déduisons donc le développement de Taylor de ce deuxième terme qui est donné par

$$C_{contrainte_2}(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) = C_{contrainte_2}(\widetilde{\mathbf{W}}) - \frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi} \mathrm{tr}[\widetilde{\mathbf{M}}(\lambda)\widetilde{\varepsilon}(\lambda)]\frac{d\lambda}{(2\pi)^2} \tag{5.15}$$

En utilisant d'une part le fait que $\widetilde{\varepsilon}(\lambda) = \delta\widetilde{\mathbf{W}}(\lambda)\widetilde{\mathbf{W}}^{-1}(\lambda)$, et d'autre part la définition de la transformée de Fourrier de la séquence de matrice $\{\mathbf{W}(\ell)\}$ que nous rappelons ici $\delta\widetilde{\mathbf{W}}(\lambda) = \sum_\ell \delta\mathbf{W}(\ell)e^{-i\ell^T\lambda}$, l'équation (5.15) devient finalement

$$C_{contrainte_2}(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) = C_{contrainte_2}(\widetilde{\mathbf{W}}) - \frac{1}{\log 2}\,\mathrm{tr}\sum_\ell \delta\mathbf{W}(\ell)[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-1}(\lambda)\widetilde{\mathbf{M}}(\lambda)e^{-i\ell^T\lambda}\frac{d\lambda}{(2\pi)^2}] \tag{5.16}$$

En regroupant ce résultat avec celui de l'équation (5.4) et en ne tenant compte que des termes au 1er ordre, nous obtenons le développement de Taylor du critère sous contrainte $C_{contrainte}$ au 1er ordre dont l'expression est réduite à :

$$\begin{aligned} C_{contrainte}(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) \;=\; & C_{contrainte}(\widetilde{\mathbf{W}}) + \frac{1}{\log 2}\,\mathrm{tr}\sum_\ell \delta\mathbf{W}(\ell)R_{\mathbf{X},\psi(\mathbf{Y})}(-\ell) - \\ & \frac{1}{\log 2}\,\mathrm{tr}\sum_\ell \delta\mathbf{W}(\ell)[\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{W}}^{-1}(\lambda)\widetilde{\mathbf{M}}(\lambda)e^{-i\ell\lambda}\frac{d\lambda}{(2\pi)^2}] \end{aligned} \tag{5.17}$$

Cette dernière expression peut encore se mettre sous la forme simplifiée

$$C_{contrainte}(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) = C_{contrainte}(\widetilde{\mathbf{W}}) + \sum_\ell \mathrm{tr}[\delta\mathbf{W}(\ell)\mathbf{G}^T_{contrainte}(\ell)]$$

avec

$$\mathbf{G}_{contrainte}(\ell) = \frac{1}{\log 2}R^T_{\mathbf{X},\psi(\mathbf{Y})}(-\ell) - \frac{1}{\log 2}\int_0^{2\pi}\int_0^{2\pi} \widetilde{\mathbf{M}}^T(\lambda)\widetilde{\mathbf{W}}^{-T}(\lambda)e^{-i\ell^T\lambda}\frac{d\lambda}{(2\pi)^2}, \tag{5.18}$$

qui constitue le gradient du critère $C_{contrainte}$.

5.3 Estimation empirique du gradient et du critère

Le calcul numérique effectif du critère défini par l'équation (4.18) ainsi que de son gradient sont des opérations très complexes faisant intervenir des doubles intégrations et nécessite la connaissance exacte des densités de probabilités des signaux. En réalité, la densité de probabilité des signaux est rarement connue dans un cas pratique où nous supposons que les observations ont été obtenues à partir d'un mélange convolutif comme nous l'avons fait ici. En pratique afin de mettre au point un algorithme de minimisation, nous utilisons non pas les expressions mathématiques réelles du

gradient et du critère, mais plutôt une estimation de ces dernières. Pour cela, la double intégrale est remplacée par une double somme de Riemann, la fonction score et l'entropie sont estimées comme cela est suggéré dans [59] et la fonction d'inter-corrélation est estimée en remplaçant l'espérance par une moyenne empirique. Le critère est estimé de façon empirique par :

$$\widehat{C}(\{\mathbf{W}(\ell)\}) = \sum_{j=1}^{N} \widehat{H}(\mathbf{Y}_j) + \frac{1}{2}\log_2\{\det(\widehat{\mathbf{D}})\} \tag{5.19}$$

avec

$$\widehat{\mathbf{D}} = \frac{1}{T_1 T_2}\sum_{m=0}^{T_1-1}\sum_{n=0}^{T_2-1} \mathrm{diag}[\widetilde{\mathbf{M}}(\lambda_{m,n})] \tag{5.20}$$

où pour la clarté et la simplicité des expressions, nous avons choisi d'utiliser les notations vectorielles et la variable fréquentielle discrète notée $\lambda_{m,n}$ est définie par $\lambda_{m,n} = (\frac{2\pi m}{T_1}, \frac{2\pi n}{T_2})$. Nous choisirons T_1 et T_2 assez grands de telle sorte que l'erreur d'approximation des intégrales par la somme de Riemann est faible, mais il ne faut pas non plus prendre des valeurs très grandes car elles n'amélioreraient pas d'une façon significative les estimations des intégrales mais, augmenteraient la complexité de calcul. Le développement de Taylor au premier ordre de ce critère estimé peut se mettre sous la forme suivante :

$$\widehat{C}(\widetilde{\mathbf{W}} + \delta\widetilde{\mathbf{W}}) = \widehat{C}(\widetilde{\mathbf{W}}) + \sum_{\ell} \mathrm{tr}[\delta\mathbf{W}(\ell)\widehat{\mathbf{G}}^T(\ell)] \tag{5.21}$$

avec

$$\widehat{\mathbf{G}}(\ell) = \frac{1}{\log 2}\widehat{\mathbf{R}}^T_{\mathbf{X},\widehat{\psi}(\mathbf{Y})}(-\ell) - \frac{1}{\log 2}\frac{1}{T_1 T_2}\sum_{m=0}^{T_1-1}\sum_{n=0}^{T_2-1} \widetilde{\mathbf{M}}^T(\lambda_{m,n})\widehat{\mathbf{D}}^{-1}\widetilde{\mathbf{W}}^{-T}(\lambda_{m,n})e^{-i2\pi(\frac{ml_1}{T_1}+\frac{nl_2}{T_2})} \tag{5.22}$$

$\widehat{\mathbf{G}}(\ell)$ est le gradient du critère empirique de (5.19). La fonction d'inter-corrélation estimée $\widehat{\mathbf{R}}_{\mathbf{X},\widehat{\psi}(\mathbf{Y})}(\ell)$ entre $\mathbf{X}$ et $\widehat{\psi}(\mathbf{Y})$ est donnée par :

$$\widehat{\mathbf{R}}_{\mathbf{X},\widehat{\psi}(\mathbf{Y})}(\ell) = \frac{1}{N_L N_C}\sum_{\mathbf{k}=(0,0)}^{(N_L-1,N_C-1)} \mathbf{X}(\mathbf{k})\widehat{\psi}^T[\mathbf{Y}(\mathbf{k}-\ell)] \tag{5.23}$$

Le critère sous la contrainte $\mathbf{D} = \mathbf{I}$ est estimé de façon empirique par :

$$\widehat{C}_{contrainte}(\{\mathbf{W}(\ell)\}) = \sum_{j=1}^{N} \widehat{H}(Y_j) + \frac{1}{2\log 2}\mathrm{tr}(\widehat{\mathbf{D}}) \tag{5.24}$$

avec $\widehat{\mathbf{D}}$ définie de la même façon qu'à l'équation (5.20). Son gradient lui est estimé par l'expression

$$\widehat{\mathbf{G}}_{contrainte}(\ell) = \frac{1}{\log 2}\widehat{\mathbf{R}}^T_{\mathbf{X},\widehat{\psi}(\mathbf{Y})}(-\ell) - \frac{1}{\log 2}\frac{1}{T_1 T_2}\sum_{m=0}^{T_1-1}\sum_{n=0}^{T_2-1} \widetilde{\mathbf{M}}^T(\lambda_{m,n})\widetilde{\mathbf{W}}^{-T}(\lambda_{m,n})e^{-i2\pi(\frac{ml_1}{T_1}+\frac{nl_2}{T_2})} \tag{5.25}$$

où $\widehat{\mathbf{R}}^T_{\mathbf{X},\widehat{\psi}(\mathbf{Y})}(\ell)$ est une estimation de la fonction d'inter-corrélation entre $\mathbf{X}$ et $\widehat{\psi}(\mathbf{Y})$ dont la définition est donnée par l'équation (5.23).

Les estimations empiriques du gradient et du critère avec et sans contrainte sont nécessaires dans la mesure où elles permettent de pouvoir mettre en place un algorithme numérique. Dans le cas de la minimisation sous contrainte, celle-ci devient $\widehat{\mathbf{D}} = \mathbf{I}$. En utilisant les définitions des estimations du critère et du gradient, nous explicitons par la suite des algorithmes permettant de minimiser ce critère avec ou sans contrainte. Nous appellerons *Ica_conv_opt* l'algorithme qui minimise le critère général et *Ica_conv_orth* celui qui minimise le critère sous la contrainte $\mathbf{D} = \mathbf{I}$.

5.4 Algorithmes BFGS pour la minimisation de C et $C_{contrainte}$

Nous utilisons une méthode de minimisation BFGS du nom de ces auteurs *Broyden-Fletcher-Goldfarb-Shanno* qui est une méthode itérative, basée sur l'estimation du gradient et une approximation empirique du hessien. Le gradient est évalué à chaque étape en utilisant les formules d'estimations de la section 5.3 et dépend de la solution courante. L'hessien, lui, est estimé à chaque étape, en se basant sur sa valeur précédente et en l'actualisant selon une certaine façon qui est à l'origine de cette méthode. La façon de mettre à jour la valeur du hessien d'une itération à l'autre ressemble fortement à celle de la méthode DFP du nom de ces auteurs *Davidson-Fletcher-Powell*, qui est décrite et détaille dans [80]. Dans [63], on y trouve les détails liés aux différences entre ces deux méthodes d'optimisation, ainsi qu'une implantation de la méthode BFGS dans un cadre général. La différence majeure entre ces deux méthodes résulte en la mise à jour du hessien estimé sur la solution courante. En général, la méthode BFGS converge mieux que la méthode DFP, ce qui justifie le fait que nous l'ayons choisie pour minimiser notre critère. Nous rappelons ici les grandes lignes de cet algorithme en l'adaptant à notre problème.

Nous allons utiliser, dans l'espace des matrices carrées d'ordre N les opérateurs vec et mat qui ont été définis à la section 1.1.1 du premier chapitre et qui permettent de transformer respectivement une matrice en un vecteur et un vecteur en une matrice. L'objectif étant de rechercher un ensemble fini de matrice $\mathbf{W}(\ell)$, nous fixons au préalable le nombre de matrices en faisant l'hypothèse suivante

$$\forall\ \ell = (\ell_1, \ell_2),\ \mathbf{W}(\ell) = \mathbf{O}_N \quad si \ \ |\ell_1| > L_1 \ \ ou \ \ |\ell_2| > L_2$$

ce qui a pour conséquence de limiter la séquence des matrices $\{\mathbf{W}(\ell)\}$ à rechercher au nombre de $N_{\mathbf{w}} = (2L_1 + 1)(2L_2 + 1)$ matrices de taille $N \times N$. Introduisons les vecteurs $\mathbf{g}$ et $\mathbf{x}$ de dimension $N_{\mathbf{w}}N^2$ et qui sont respectivement définis par : $\mathbf{g} = \mathrm{vec}(\{\widehat{\mathbf{G}}(\ell)\}) = [\mathrm{vec}(\widehat{\mathbf{G}}(\ell_1))^T, \ldots, \mathrm{vec}(\widehat{\mathbf{G}}(\ell_{N_{\mathbf{w}}}))^T]^T$ et $\mathbf{x} = \mathrm{vec}(\{\mathbf{W}(\ell)\}) = [\mathrm{vec}(\mathbf{W}(\ell_1))^T, \ldots, \mathrm{vec}(\mathbf{W}(\ell_{N_{\mathbf{w}}}))^T]^T$. Le vecteur $\mathbf{g}$ est constitué par l'ensemble des matrices $\widehat{\mathbf{G}}(\ell)$ transformées en vecteurs par l'opérateur vec et prises dans un ordre prédéfini. Il représente le gradient au point $\mathbf{x}$ qui lui est constitué par l'ensemble des matrices $\mathbf{W}(\ell)$ transformées en vecteurs par l'opérateur vec et prises dans le même ordre que dans le cas de $\mathbf{g}$. Le vecteur $\mathbf{x}$ représente une solution courante. Enfin, appelons $f(\mathbf{x})$ la fonction définie sur $\mathbb{R}^{N_{\mathbf{w}}N^2}$ par : $f(\mathbf{x}) = \widehat{C}(\{\mathbf{W}(\ell)\})$ dans le cas sans contrainte et par $f(\mathbf{x}) = \widehat{C}_{contrainte}(\{\mathbf{W}(\ell)\})$ dans le cas de la minimisation sous contrainte.

Par la suite, pour des raisons de simplicité et d'adaptabilité, nous allons résoudre notre problème en minimisant le critère défini à l'équation (4.18) par rapport au vecteur $\mathbf{x}$. Cela revient au même que de le faire sur la séquence de matrice $\{\mathbf{W}(\ell)\}$ tout en sachant qu'à tout moment, il est très facile à partir du vecteur $\mathbf{x}$ de reconstituer la séquence $\{\mathbf{W}(\ell)\}$ en utilisant l'opérateur mat et en l'appliquant aux $N_{\mathbf{w}}$ vecteurs disjoints de longueur N^2 constitués en prenant à chaque fois N^2 éléments successifs de $\mathbf{x}$. L'algorithme BFGS décrit dans [63] nous permettant de résoudre notre problème de minimisation se résume alors comme suit :

1. Choisir une solution arbitraire initiale que l'on note $\mathbf{x}^0 = \text{vec}(\{\mathbf{W}(\ell)\}^0)$, un choix possible est $\mathbf{W}^0(\ell) = \mathbf{0}_N$ si $\ell_1 \neq 0$ et $\ell_2 \neq 0$, et $\mathbf{W}^0(0,0) = \mathbf{I}_N$. On ne s'intéresse qu'aux $N_\mathbf{w}$ matrices $\mathbf{W}^0(\ell)$ non nulles et la taille du vecteur $\mathbf{x}^0$ est $N_\mathbf{w}N^2$. Notons $\mathbf{g}^0 = \nabla f(\mathbf{x}^0)$ le gradient estimé au point $\mathbf{x}^0$ et $\mathbf{H}_0$ une matrice carrée quelconque de taille $N_\mathbf{w}N^2 \times N_\mathbf{w}N^2$, symétrique définie positive et qui est une estimation empirique de l'inverse de l'hessien du critère, nous choisissons naturellement et simplement d'initialiser comme préconisé pour cette méthode BFGS $\mathbf{H}_0 = \mathbf{I}_{N_\mathbf{w}N^2}$.
2. Pour $v = 0,\ 1,\ 2,\ \ldots$, faire
   (a) $\mathbf{s}^v = \mathbf{H}_v\mathbf{g}^v$,
   (b) $\alpha_v = \min_\alpha[f(\mathbf{x}^v - \alpha\mathbf{s}^v)]$,
   (c) $\mathbf{x}^{v+1} = \mathbf{x}^v - \alpha_v\mathbf{s}^v, \qquad \mathbf{g}^{v+1} = \nabla f(\mathbf{x}^{v+1})$,
   (d) $\mathbf{H}_{v+1} = \mathbf{H}_v + \mathbf{P}_v + \mathbf{Q}_v + \mathbf{E}_v$
3. Continuer jusqu'à ce que $||\mathbf{g}^v|| < \epsilon$.

avec

$$\mathbf{P}_v = -\frac{\alpha_v\mathbf{s}^v(\mathbf{s}^v)^T}{(\Delta\mathbf{g}^v)^T\mathbf{s}^v}, \tag{5.26}$$

$$\mathbf{Q}_v = -\frac{\mathbf{H}_v\Delta\mathbf{g}^v(\Delta\mathbf{g}^v)^T\mathbf{H}_v}{(\Delta\mathbf{g}^v)^T\mathbf{H}_v\Delta\mathbf{g}^v} \tag{5.27}$$

$$\mathbf{E}_v = [(\Delta\mathbf{g}^v)^T\mathbf{H}_v\Delta\mathbf{g}^v]\mathbf{u}^v(\mathbf{u}^v)^T \tag{5.28}$$

où

$$\mathbf{u}^v = \frac{\mathbf{s}^v}{(\mathbf{s}^v)^T\Delta\mathbf{g}^v} - \frac{\mathbf{H}_v\Delta\mathbf{g}^v}{(\Delta\mathbf{g}^v)^T\mathbf{H}_v\Delta\mathbf{g}^v} \tag{5.29}$$

et

$$\Delta\mathbf{g}^v = \mathbf{g}^{v+1} - \mathbf{g}^v \tag{5.30}$$

5.5 Complexité algorithmique

Pour évaluer la complexité de ces deux algorithmes, nous recherchons la complexité de calcul du gradient et du critère à chaque itération. Les autres éléments qui interviennent dans la définition de l'algorithme dépendant fortement de l'estimation de ces deux quantités, notamment du hessien qui est mis à jour d'une façon empirique et qui est propre à la méthode de minimisation BFGS. Soient N le nombre de composantes de l'image de départ, T_1 et T_2 les nombres d'intervalles des subdivisions des deux axes de fréquences spatiales associées aux sommes de Riemann mentionnées ci-dessus, N_L et N_C la taille d'une composante en nombre de lignes et de colonnes avec $L = N_LN_C$, et $N_\mathbf{w}$ le nombre de matrices carrées de taille $N \times N$ recherchées par la transformation *Ica_conv_opt*. Nous allons évaluer la complexité d'une itération et nous déduirons la complexité de l'algorithme en fonction du nombre d'itérations.

Appelons m le nombre de points de discrétisation nécessaires pour le calcul de la fonction score dont l'algorithme est explicité dans [62]. Nous avons en général $m \ll L$ et qui est un entier compris dans l'intervalle [30; 60]. L'estimation de la fonction score des N composantes se fait avec une complexité de l'ordre de $\text{O}(NmL)$. L'estimation de l'entropie différentielle des N composantes se fait avec une complexité algorithmique du même ordre que celle de la fonction score. Le calcul de l'estimation de la fonction d'inter-corrélation de l'équation (5.23) se fait avec une complexité en

$\mathrm{O}(N^2 N_{\mathbf{w}} L)$. Si nous observons les éléments intervenant dans le deuxième terme du membre de droite de l'équation (5.22), $\widetilde{\mathbf{W}}(\lambda_{m,n})$ a une complexité en $\mathrm{O}(N_{\mathbf{w}} N^2)$, sachant que le calcul de l'inverse d'une matrice carré d'ordre N se fait en $\mathrm{O}(N^3)$ et que le produit de deux matrices carrées d'ordre N se fait également en $\mathrm{O}(N^3)$, la matrice $\widetilde{\mathbf{M}}(\lambda_{m,n}) = \widetilde{\mathbf{W}}^{-H}(\lambda_{m,n})\widetilde{\mathbf{W}}^{-1}(\lambda_{m,n})$ a une complexité de calcul en $\mathrm{O}(N_{\mathbf{w}} N^2 + 2N^3)$ alors que $\widehat{\mathbf{D}}$ qui nécessite le calcul de T valeurs de $\widetilde{\mathbf{M}}(\lambda_{m,n})$ a une complexité de calcul en $\mathrm{O}(N_{\mathbf{w}} T N^2 + 2TN^3)$. On déduit alors que le calcul du deuxième terme du membre de droite de l'équation (5.22) se fait en $\mathrm{O}(4N^3 T + N_{\mathbf{w}} N^2 T)$. Ainsi, le calcul du gradient de l'équation (5.22) pour les L matrices recherchées se fait en $\mathrm{O}([Nm + N^2 N_{\mathbf{w}}]L + [4N^3 N_{\mathbf{w}} + N_{\mathbf{w}}^2 N^2]T)$. Le nombre de composantes N des images multi spectrales est tel que $N \in \{3, 4, 6\}$, et le terme Nm dans ce cas n'est pas négligeable devant $N^2 N_{\mathbf{w}}$. Le calcul du critère estimé de façon empirique de l'équation (5.19) se fait en O(NmL), car nous avons déjà inclus la complexité de $\widehat{\mathbf{D}}$ dans le calcul du gradient empirique. La mise à jour de l'inverse de l'hessien après avoir calculé le gradient et le critère au point courant se fait avec une complexité en $\mathrm{O}(2(N^2 N_{\mathbf{w}})^3)$. En effet, comme nous l'avons dit ci-haut, durant la phase pratique de la mise en oeuvre de l'algorithme, le critère est minimisé par rapport au vecteur $\mathbf{x}$ dont la longueur est de $N^2 N_{\mathbf{w}}$ et qui représente les valeurs des $N_{\mathbf{w}}$ matrices recherchées par l'algorithme de minimisation. La mise à jour de l'hessien peut encore se mettre sous la forme d'un produit de trois matrices carrées de taille $N^2 L \times N^2 L$ auquel s'ajoute une matrice carrée de taille identique, ce qui justifie sa complexité donnée précédemment.

Pour une itération, nous évaluons au préalable le critère et le gradient au point courant, puis le critère et le gradient de la solution mise à jour avant de mettre à jour la valeur de l'hessien. Cela revient à calculer le critère et le gradient une fois avant de mettre à jour l'hessien. Certaines itérations nécessitent de faire plusieurs estimations du critère pour la recherche de α_v, surtout au début de l'algorithme. Toutefois, ce nombre reste inférieur à cinq en général à chaque itération. En supposant que l'algorithme converge au bout de p itérations (p étant de l'ordre de la centaine), nous pouvons conclure que l'algorithme `Ica_conv_opt` a une complexité en $\mathrm{O}([Nm + N^2 N_{\mathbf{w}}]pL + [4N^3 N_{\mathbf{w}} + N_{\mathbf{w}}^2 N^2]pT)$. Dans la pratique, nous choisirons $T_1 = N_L$ et $T_2 = N_C$, ce qui réduit la complexité algorithmique à $\mathrm{O}([Nm + N^2 N_{\mathbf{w}} + 4N^3 N_{\mathbf{w}} + N_{\mathbf{w}}^2 N^2]pT)$. La minimisation du critère sous contrainte se fait avec une complexité à peu près du même ordre.

Conclusion : Ce deuxième chapitre propose et explicite un algorithme de minimisation BFGS du critère mis à jour au premier chapitre de cette partie. Pour cela, nous avons calculé le gradient du critère et, nous avons déduit les expressions des estimations empiriques du gradient et du critère d'abord dans le cas général, puis dans le cas particulier où le critère est minimisé sous la contrainte $\widehat{\mathbf{D}} = \mathbf{Id}$. Enfin, nous avons évalué d'une façon assez sommaire la complexité de calcul de ces deux algorithmes. Par la suite, nous présentons les performances en terme de courbes débit vs PRSB obtenues par ces algorithmes.

Chapitre 6

Evaluation des performances des transformations convolutives

Ce dernier chapitre présente et propose une discussion sur les performances obtenues par les transformations convolutives qui ont été définies et explicitées dans les précédents chapitres de cette deuxième partie. Nous nous focaliserons uniquement sur la transformation convolutive sans contrainte, l'autre transformation n'ayant pour intérêt que la préservation de la distorsion après transformation. Le schéma de compression étudié étant celui présenté en 4.1 où la transformation est ici la transformation convolutive obtenue par l'algorithme permettant de minimiser le critère de l'équation (4.18) et que nous appellerons *Ica_conv_opt*. La base de données des images reste la même que celle définie et présenté au chapitre 3 de la première partie et nous utilisons quelques unes de ces images pour réaliser les simulations.

6.1 Evaluation des Performances de la transformation *Ica_conv_opt*

Nous nous intéressons ici au schéma de compression présenté dans 4.1. L'image originale ici étant une sous-bande $\mathbf{X}^{(m)}$ obtenue après application d'une TOD sur chaque composante de l'image $\mathbf{X}$. Par la suite, nous considérons donc une sous bande quelconque $\mathbf{X}^{(m)}$ $(1 \leq m \leq M)$ comme une image originale et nous allons appliquer plusieurs transformations différentes pour réduire la redondance spectrale (dans certains cas la redondance spectrale aussi est réduite dans chaque composante) entre les composantes de $\mathbf{X}^{(m)}$. Avant cela, nous vérifions sur quelques image la validité de la formule d'approximation de la distorsion de l'équation (4.10).

6.1.1 Validation de la formule d'approximation de la distorsion

Pour évaluer la validité de la formule d'approximation de la distorsion obtenue dans cette partie, nous appliquons la transformation convolutive obtenue par l'algorithme *Ica_conv_opt* sur une sous-bande $\mathbf{X}^{(m)}$ de l'image, puis nous affectons un pas de quantification à chaque composante $\mathbf{X}_i^{(m)}$ de façon à toujours être sous les hypothèses faibles distorsion. Puis nous évaluons d'une part la distorsion, en la calculant grâce aux formules d'approximations par rapport aux distorsions des composantes $\mathbf{Y}_i^{(m)}$ obtenues après la transformation, et d'autre part, nous évaluons la distorsion réelle en appliquant la transformation convolutive inverse et en reconstruisant une image approchée

$\widehat{\mathbf{X}}^{(m)}$. Appelons D_{reel} la distorsion réelle ainsi obtenue et D_{app} la distorsion obtenue en utilisant la formule d'approximation. Le PRSB est ensuite obtenu à partir de la distorsion réelle, et nous appelons **distorsion relative**, le rapport $D_r = D_{app}/D_{reel}$. En faisant plusieurs tirages aléatoires pour différents pas de quantification, nous obtenons un ensemble de points pour D_{reel} et D_{app}, ce qui nous permet de tracer le nuage de points D_r en fonction du PRSB. Pour chaque sous-bande, nous calculons l'erreur relative moyenne qui s'obtient en calculant la moyenne arithmétique des erreurs d'approximation. L'erreur d'approximation ϵ étant ici définie par $\epsilon = |1 - D_r|$.

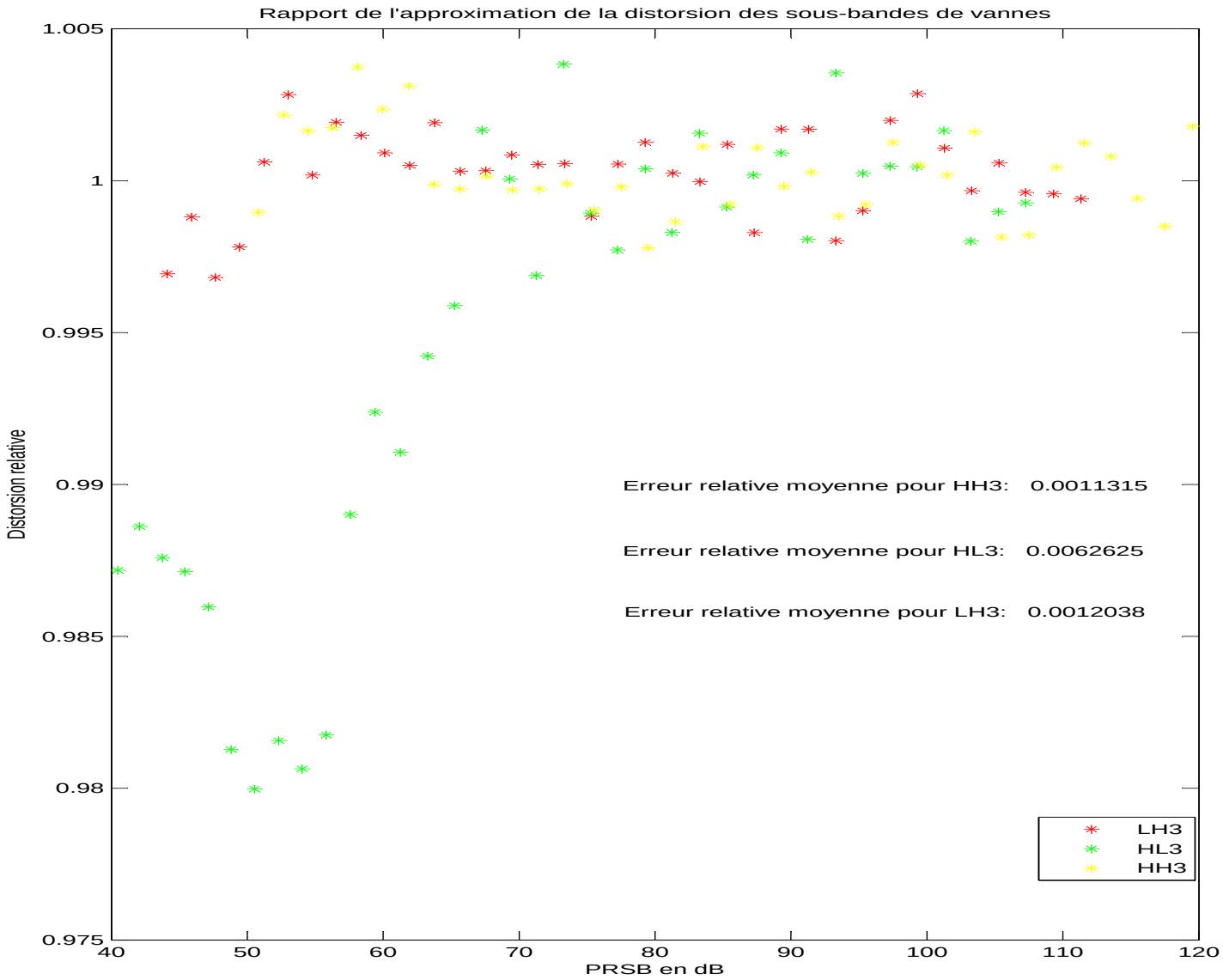


FIG. 6.1 – Distorsion relative en fonction du PRSB des sous-bandes de l'image Vannes

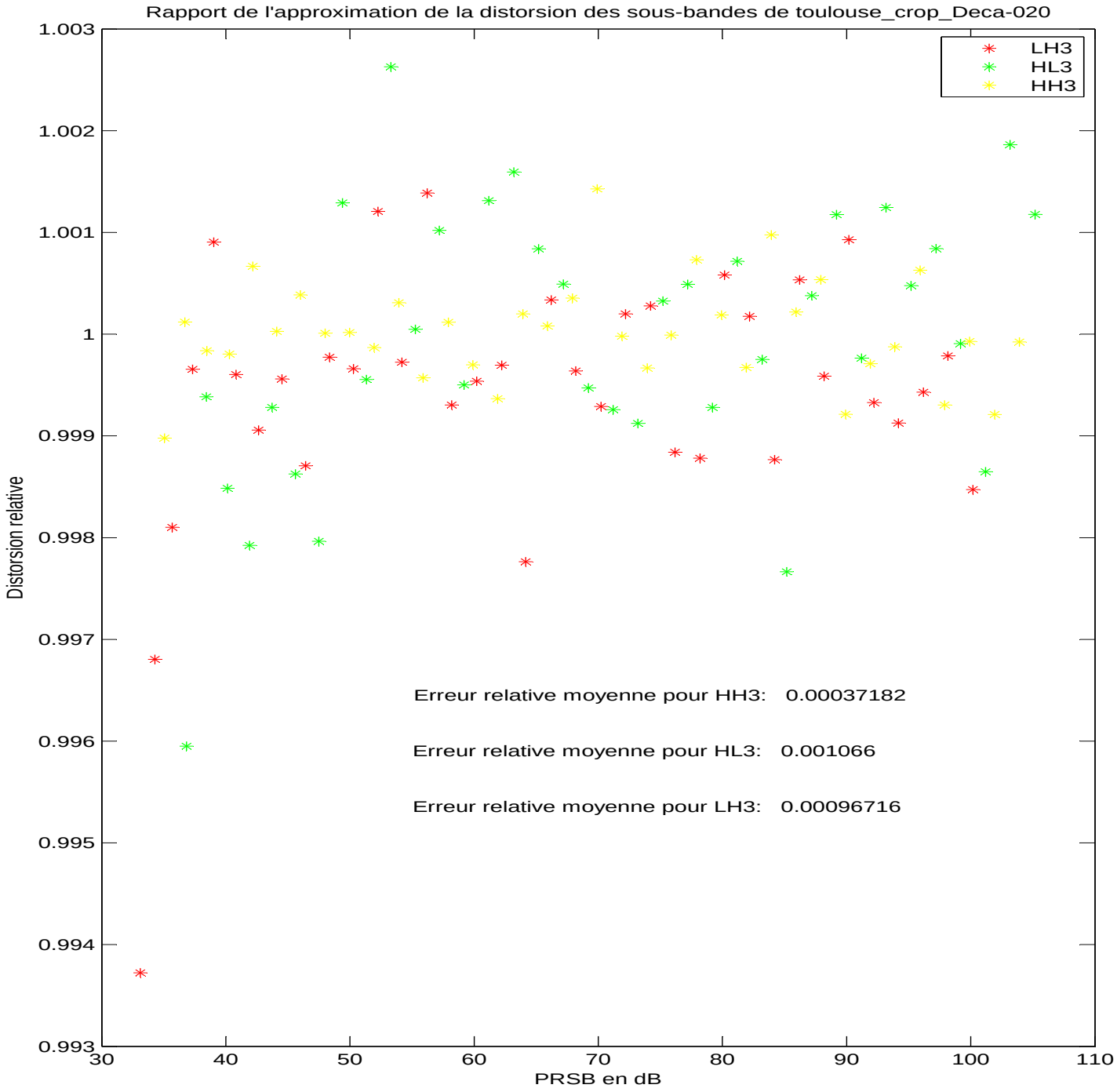


FIG. 6.2 – Distorsion relative en fonction du PRSB de l'image Toulouse décalée de 0.2 pixel

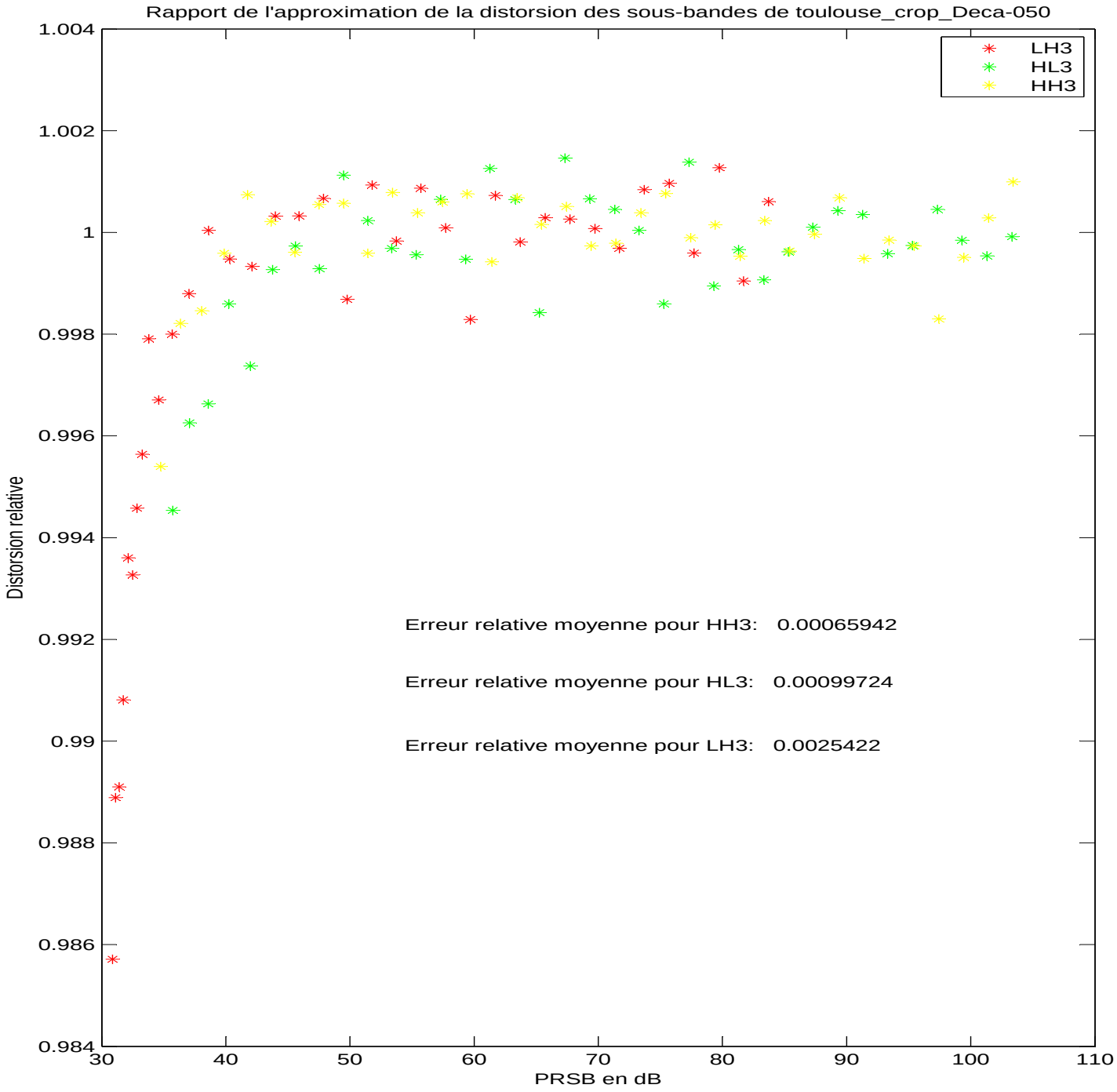


FIG. 6.3 – Distorsion relative en fonction du PRSB de l'image Toulouse décalée de 0.5 pixel

Nous avons présenté les résultats obtenus en utilisant les sous-bandes HH3, HL3 et LH3 des images de Vannes, de référence Toulouse et de Toulouse décalée de 0.5 pixel comme image originale. Ces sous-bandes représentent les coefficients de détails horizontaux, verticaux et diagonaux après trois niveaux de décomposition en ondelette. Pour chacune des figures, nous évaluons l'erreur relative moyenne. L'erreur relative moyenne est obtenue en calculant la moyenne arithmétique du carré des erreurs d'approximation. L'erreur d'approximation ϵ étant ici définie par $\epsilon = 1 - D_r$. En effet, la valeur idéale de la distorsion relative vaut 1 c'est-à-dire lorsque l'approximation de la distorsion coincide avec la distorsion réelle.

Nous constatons en analysant ces figures que la formule d'approximation de la distorsion de l'équation (4.10) est bonne car très proche de la valeur idéale et ce d'autant plus que la mesure de l'erreur moyenne d'approximation est toujours très faible. Nous avons eu des résultats similaires en évaluant la distorsion par cette formule d'approximation pour les différentes sous-bandes non seulement de ces images sur lesquelles nous avons effectué cette simulation, mais aussi sur d'autres images de notre bas de données. Ce qui conforte et justifie le fait que nous utilisions cette formule d'approximation par la suite lors de l'allocation optimale de bits pour l'évaluation des courbes debits vs PRSB en utilisant la transformation retournée par l'algorithme *Ica_conv_opt*.

6.1.2 Evaluation et comparaison des gains de codage généralisés

Pour évaluer les performances de l'algorithme *Ica_conv_opt*, nous avons appliqué la transformation convolutive issue de cet algorithme aux images originales (ici une sous bande $\mathbf{X}^{(m)}$), puis nous avons quantifié et codé les différentes composantes en faisant une allocation optimale de bits entre les différentes composantes au moyen de l'algorithme de Shoham et Gersho [69]. Nous avons dans nos simulations limité la séquence des matrices recherchées de telle sorte que $\mathbf{W}(\ell_1, \ell_1) = \mathbf{O}_N$ si $|\ell_1| > 1$ ou $|\ell_2| > 1$, ce qui limite la taille de la séquence de matrices recherchée à 9 matrices carrées de taille N. Afin d'évaluer la distorsion, il est nécessaire de calculer la transformation inverse de *Ica_conv_opt*. Cette évaluation est fastidieuse et complexe dans la mesure où le filtre matriciel obtenu par *Ica_conv_opt* est à support spatial fini et par conséquent son inverse a un support spatial infini. Pour contourner ce problème, nous avons travaillé sur le plan spectral où, il est possible de reconstituer exactement le signal original si aucune quantification n'est faite sur les composantes après la transformation. Ainsi, après la quantification du signal $\mathbf{Y}^{(m)}$), nous appliquons la FFT (Fast Fourier Transform) pour avoir une représentation du signal sur le plan fréquentiel. Le fait de travailler sur le plan spectral est intéressant dans la mesure où l'opération de convolution sur le plan spatial est transformé par une multiplication sur le plan spectral, ce qui simplifie considérablement les calculs et nous évite d'avoir à calculer de façon explicite le filtre matriciel inverse sur le plan spatial. Notons $\widetilde{\mathbf{Z}}(\lambda) = \widetilde{\mathbf{V}}(\lambda)\widetilde{\mathbf{Y}}^{q(m)}(\lambda)$, l'image obtenue sur le plan fréquentiel et qui est une approximation de l'image originale sur le plan fréquentiel. L'image approchée finale $\widehat{\mathbf{X}}^{(m)}$ de $\mathbf{X}^{(m)}$ est obtenue simplement en appliquant la IFFT (Inverse Fast Fourier Transform) à l'image $\widetilde{\mathbf{Z}}(\lambda)$. Rappelons enfin ici que les relations liant le filtre de séparation-deconvolution et son inverse sont données dans la sous-section 4.3.1.

Le débit est évalué par un estimateur de l'entropie d'ordre 1 et la distorsion est l'erreur quadratique moyenne. Les courbes debit vs PRSB ainsi que les gains de codage généralisés de quelques images sont illustrés non seulement pour la transformation convolutive modifiée pour la compression, mais aussi, pour les transformations spectrales étudiées dans la première partie de ce rapport, notamment pour le schéma en sous-bandes (en utilisant comme transformation spectrale la TKL et

les transformations d'ACI modifiées), et le schéma mixte en sous-bandes avec $P = 4$ (en utilisant la transformation retournée par l'algorithme *Ica_opt* comme transformation spectrale) ce qui nous permet d'avoir une première comparaison visuelle des performances de ces différentes transformations.

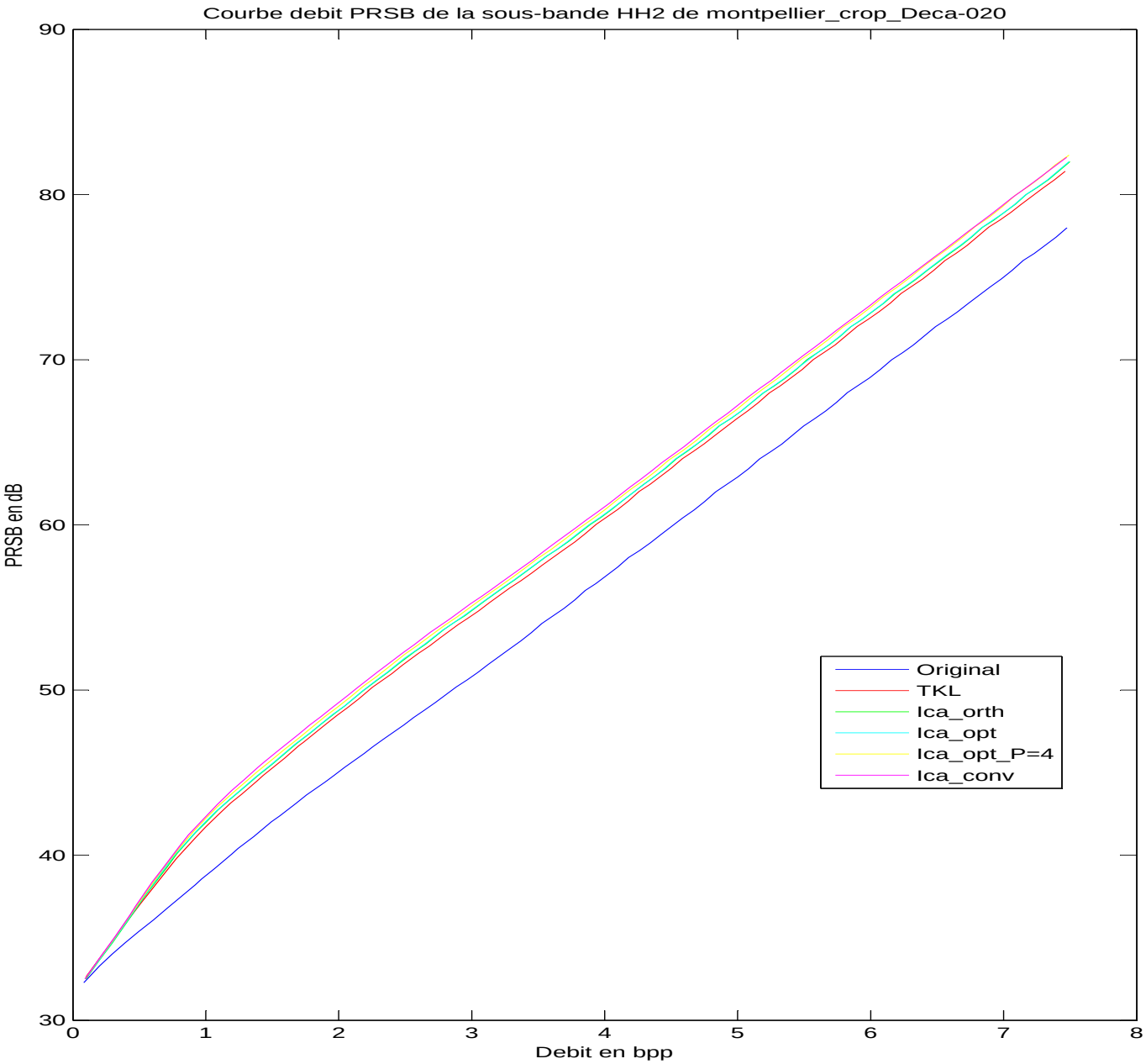


FIG. 6.4 – Gain de codage généralisé de la sous-bande HH2 de Montpellier décalée de 0.2 pixel

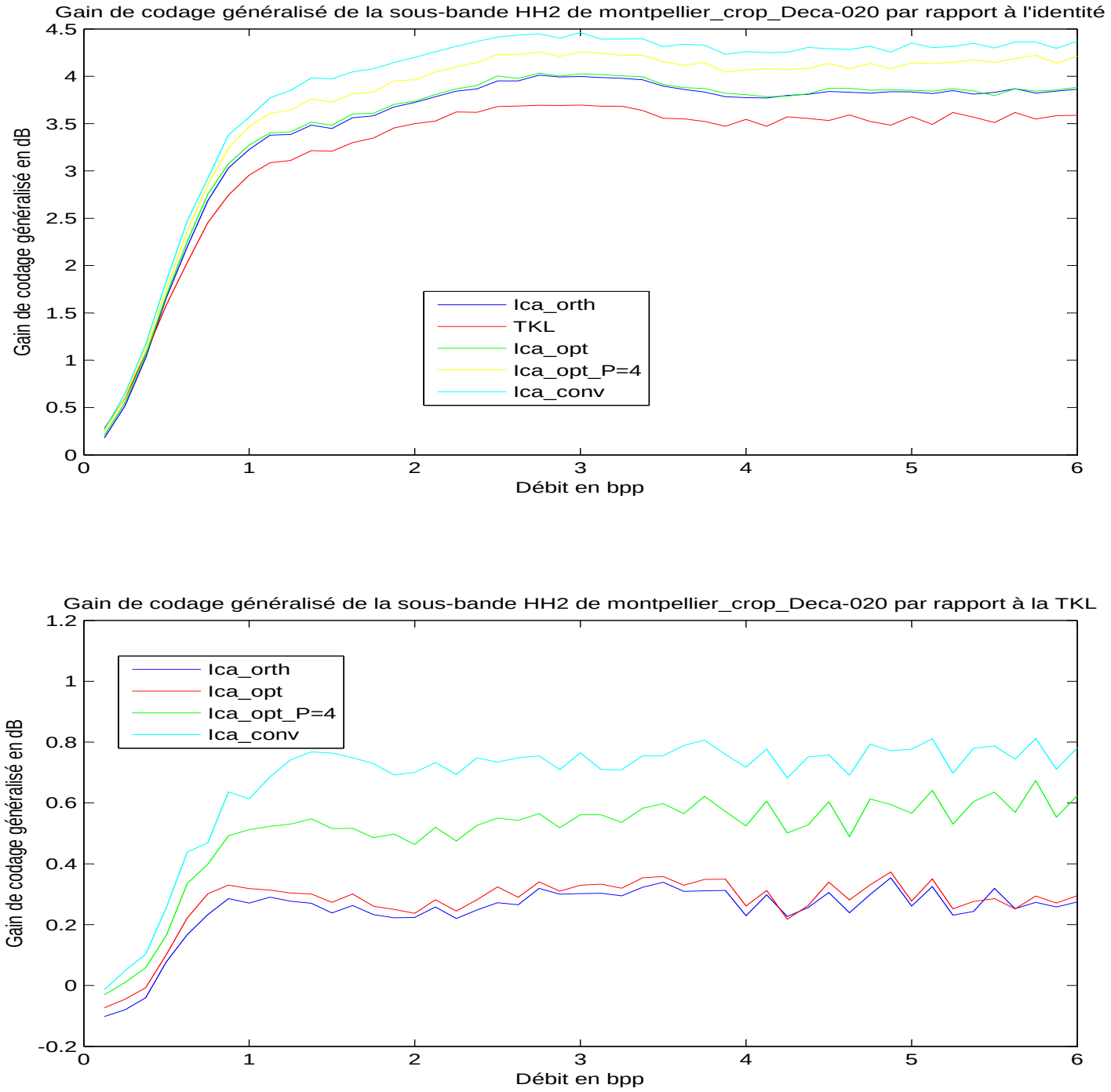


FIG. 6.5 – Gain de codage généralisé de la sous-bande HH2 de Montpellier décalée de 0.2 pixel

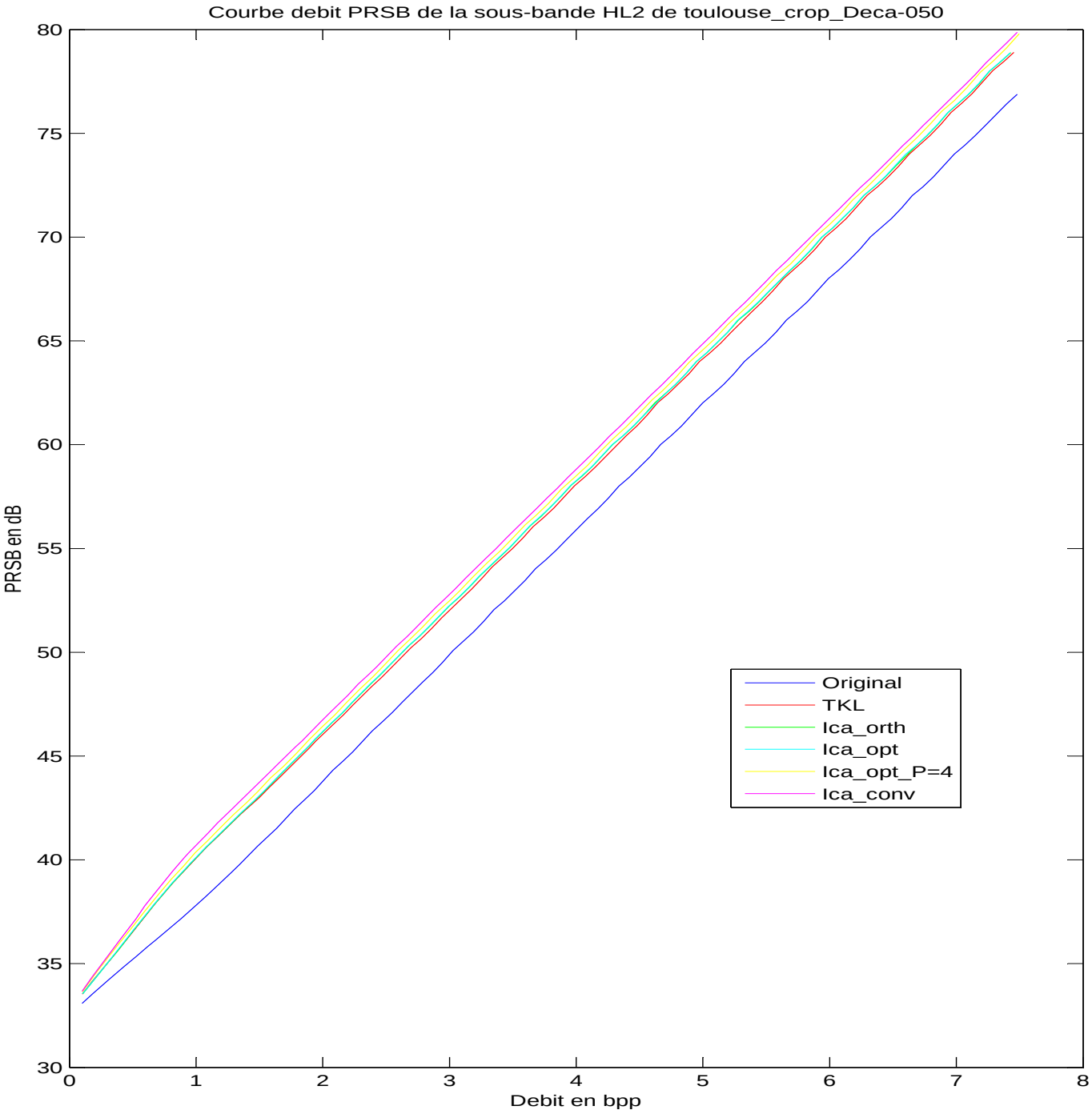


FIG. 6.6 – Gain de codage généralisé de la sous-bande HL2 de Toulouse décalée de 0.5 pixel

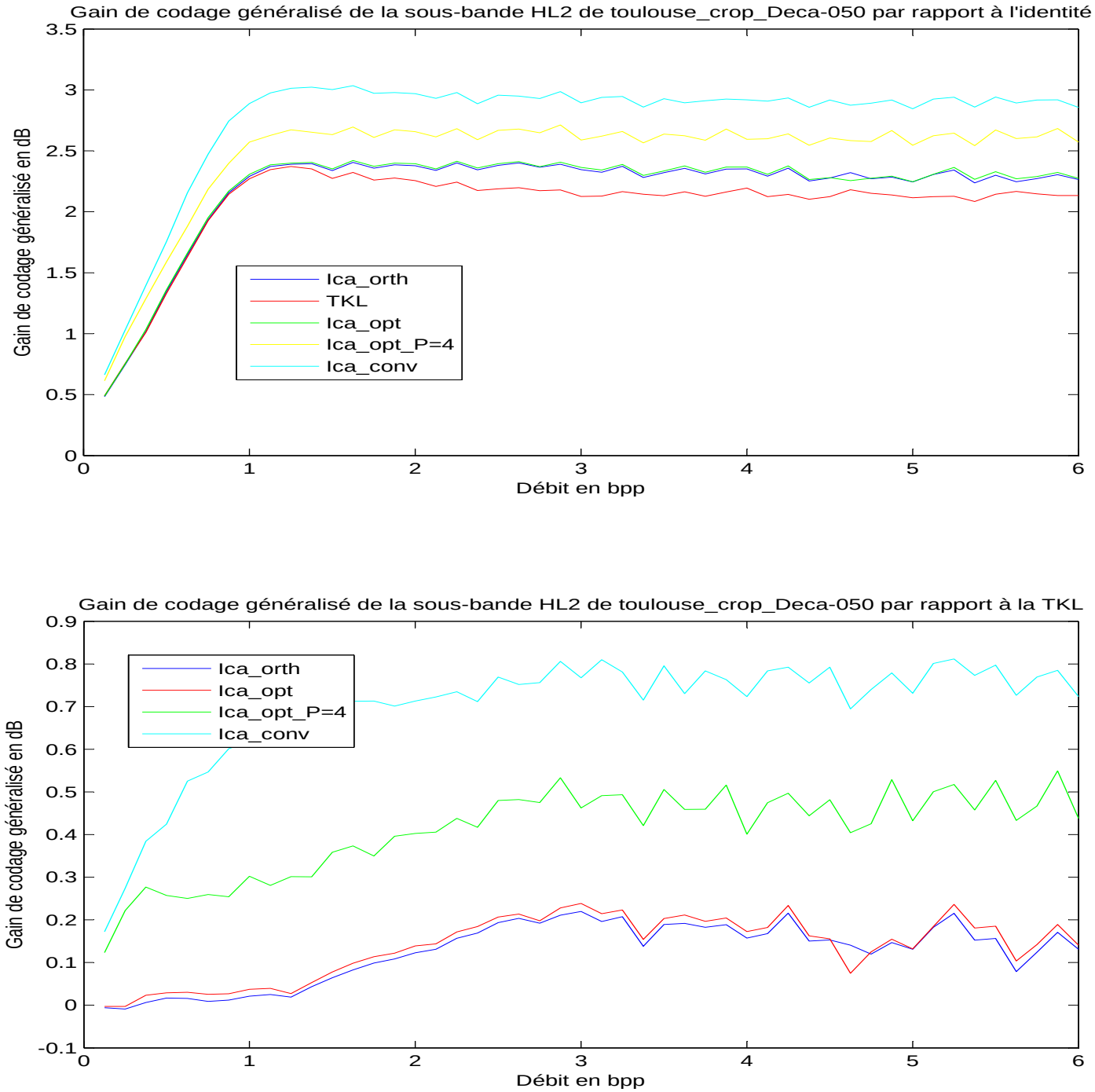


FIG. 6.7 – Gain de codage généralisé de la sous-bande HL2 de Toulouse décalée de 0.5 pixel

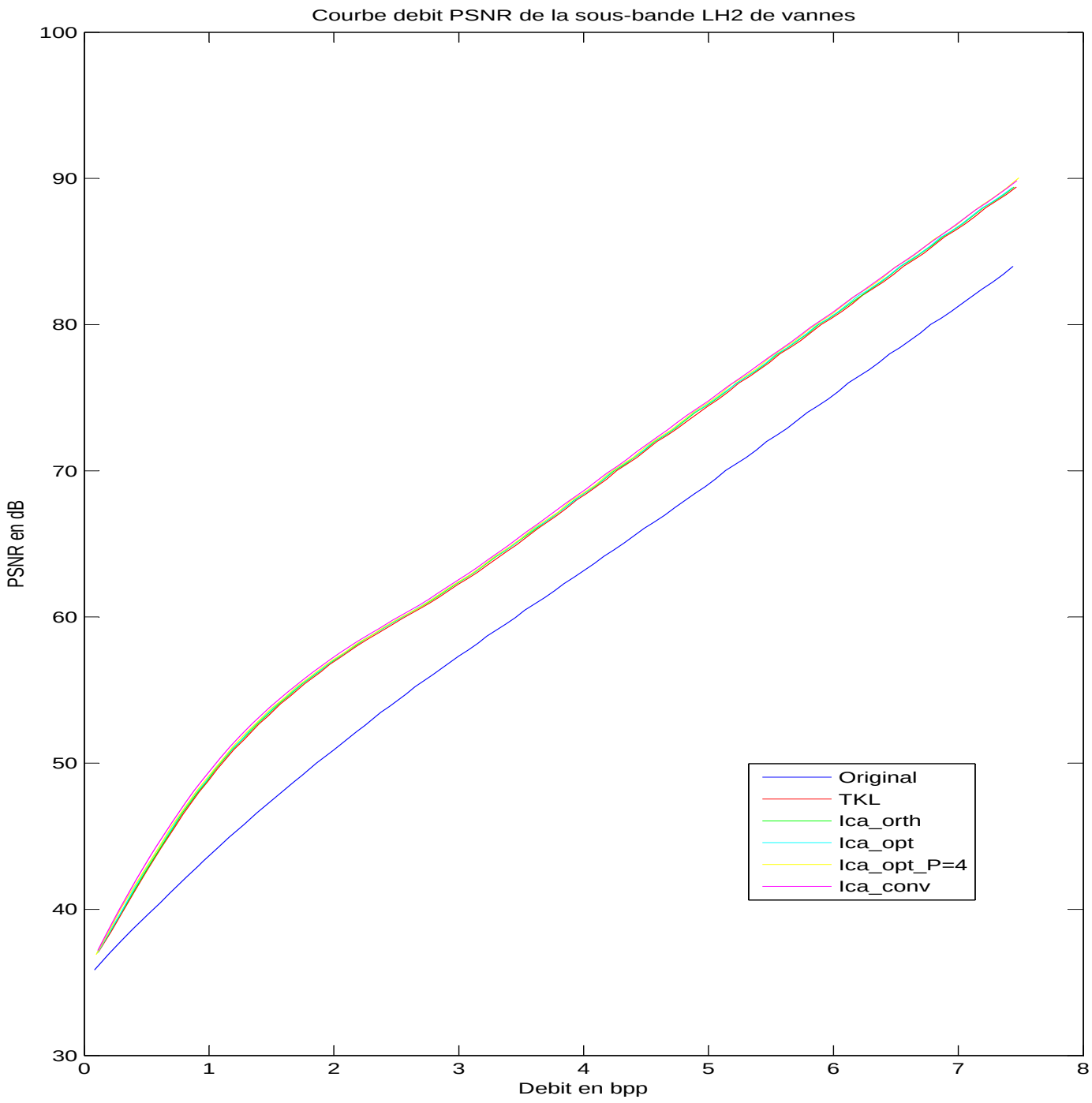


FIG. 6.8 – Gain de codage généralisé de la sous-bande LH2 de vannes

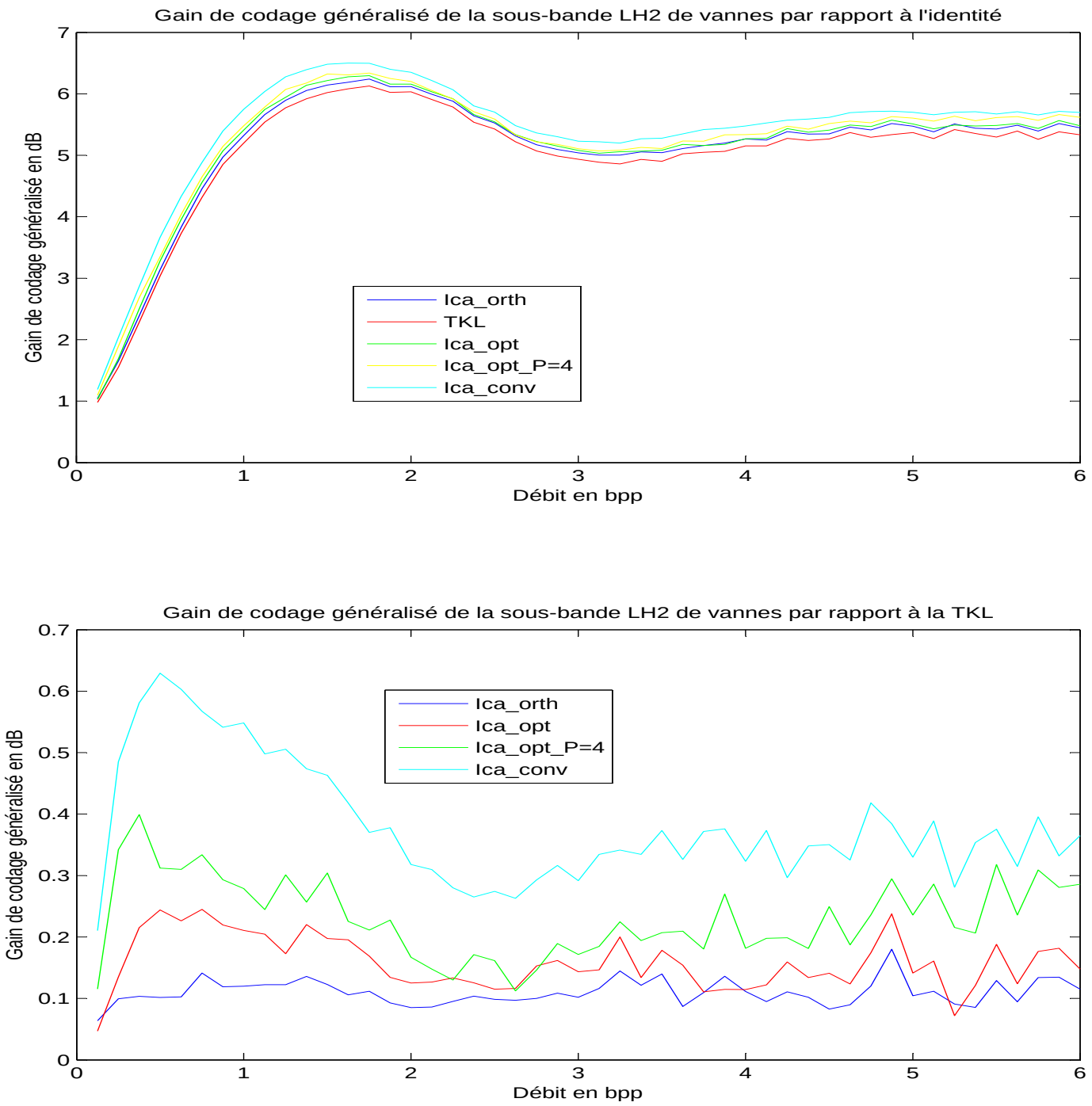


FIG. 6.9 – Gain de codage généralisé de la sous-bande LH2 de Vannes

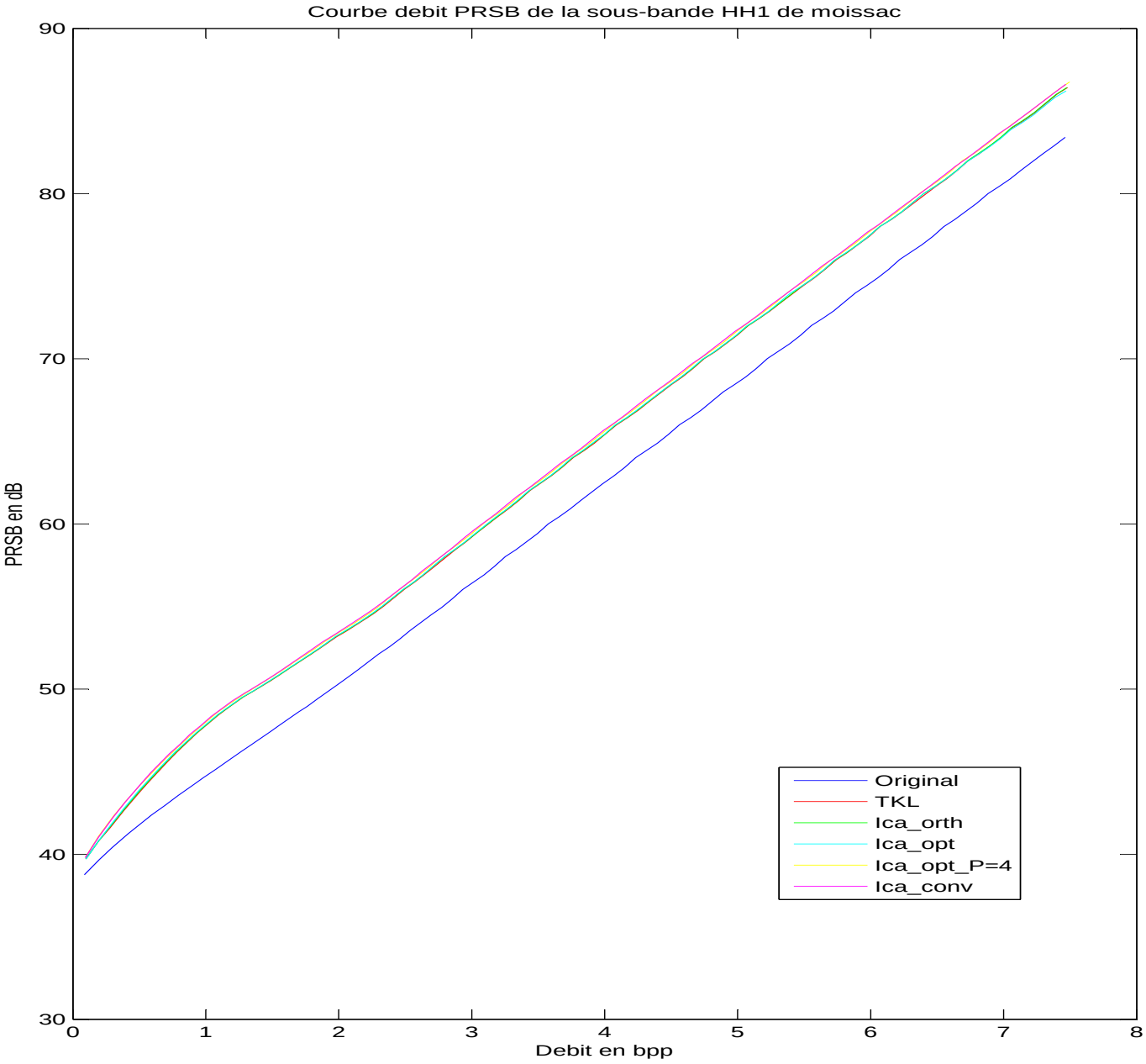
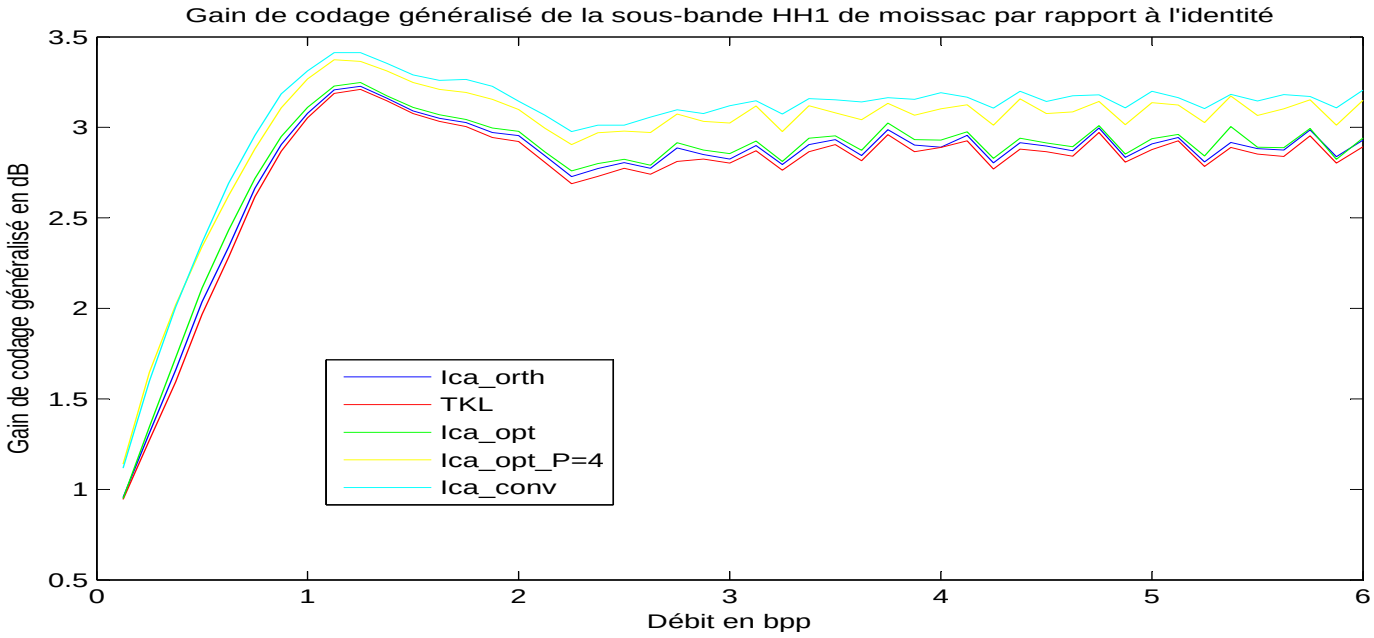


FIG. 6.10 – Gain de codage généralisé de la sous-bande HH1 de Moissac



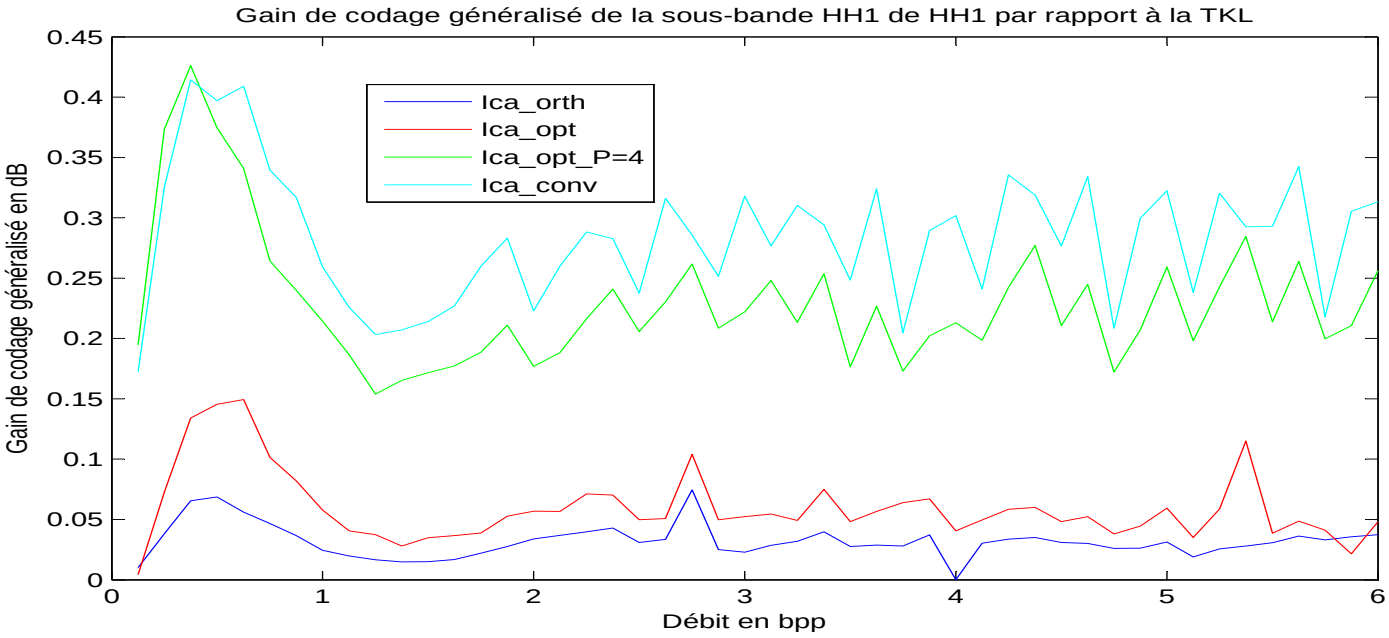


FIG. 6.11 – Gain de codage généralisé de la sous-bande HH1 de Moissac

6.1.3 Interprétations et discussions des résultats

En observant les résultats ci-dessous obtenus qui nous donnent les gains de codage généralisés obtenus sur plusieurs images (une image ici est une sous-bande après la TOD), nous constatons qu'en utilisant la transformation convolutive obtenue par l'algorithme *Ica_conv*, nous obtenons un gain de codage généralisé qui est bien pus élevé que celui obtenu en utilisant les autres transformations spectrales. Ce qui était prévisible dans la mesure où la transformation convolutive permet de réduire à la fois les redondances spatiales et spectrales entre l'image. La valeur du gain de codage obtenue ici dépend de l'image. Nous avons réalisé cette évaluation du gain de codage en utilisant plusieurs images de notre base de données initiales (en nous servant des différentes sous-bandes obtenues après application de la TOD sur les images de notre base de données), le constat est le même, la transformation convolutive permet d'améliorer les performances (en termes de courbe debit vs PRSB) des transformations spectrales utilisées dans la première première partie de ce rapport de thèse, notamment pour le schéma en sous-bandes.

Nous avons constaté que le gain de codage de la transformation convolutive par rapport à la transformation identité (et aussi par rapport aux autres transformations spectrales) obtenu était plus important dans la sous-bande basse fréquence LL3 que dans les autres sous-bandes. Cela s'explique par le fait qu'après la TOD, les coefficients d'ondelette sont bien décorrélés dans toutes les sous-bandes, excepté dans la sous-bande basse-fréquence où il demeure une redondance importante entre coefficients. Ainsi, la transformation convolutive dans ce cas, permet de minimiser non seulement la redondance spatiale remanente, mais aussi la redondance spectrale entre les composantes. Ce qui justifie que ce gain de codage généralisé élevé obtenu dans les sous-bandes basse fréquence.

Notons enfin, que nous avons également essayé d'appliquer la transformation convolutive à l'image brute de départ qui n'a subi aucune transformation. En comparant les courbes débit vs PRSB obtenus dans ce cas, nous constatons qu'elles ont des performances qui sont en deçà de celles des schémas de compressions étudiés dans la première partie, notamment le schéma en sous-bandes. Nous avons choisi de ne pas présenter ces résultats dans cette thèse et de continuer à explorer cette voie avant de tirer des conclusions partielles ou définitives.

Conclusion générale

L'objet de cette thèse est de faire une étude sur l'application des transformations d'ACI en compression d'images multi-composantes. Ce travail s'inscrit dans la suite des travaux de Narozny [52] [51] qui a essayé de comprendre et d'expliquer les performances de transformations d'ACI en compression en général, et pour la compression des images en particulier. A l'issu de ces travaux, deux transformations d'ACI modifiées ont été mises au point et sont optimales en compression à haute résolution pour une compression séparée des composantes. Dans la première partie de ce rapport, nous avons étudié plusieurs scénarios de compression incorporant une TOD par composante pour la minimisation de la redondance spatiale et une transformation linéaire pour la minimisation de la redondance spectrale. Nous avons comparé les performances des différents cas étudiés d'abord en évaluant le débit grâce à un estimateur de l'entropie d'ordre 1, ensuite en incorporant le $VM9$ dans la chaîne de compression. Les résultats obtenus ont permis de voir que les nouvelles transformations d'ACI modifiées, permettent en général d'avoir des performances meilleures que la TKL qui est utilisée et bien connue dans la communauté de la compression d'images. Nous avons aussi mis en évidence le besoin d'appliquer une transformation spectrale pour améliorer les performances en termes de courbes débit vs PRSB aussi bien à bas débits qu'à hauts débits. Enfin, notre étude a permis de voir que la dérégistration a un impact négatif sur les performances des transformations spectrales en termes de courbes débit vs PRSB, cet impact tend à disparaître lorsqu'un codeur comme celui du VM9 de JPEG2000 est utilisée.

Dans la deuxième partie de cette thèse, nous avons proposé une transformation basée sur un modèle convolutif et qui a été modifiée pour la compression. Après avoir montré qu'une telle transformation était liée de par le critère à une transformation de séparation-déconvolution en séparation de sources (en utilisant l'information mutuelle comme critère à minimiser), nous avons proposé deux algorithmes permettant de minimiser le critère obtenu en utilisant et en supposant un tel modèle de mélange. Les premiers résultats montrent que la transformation convolutive permet d'avoir de meilleures performances que les transformations modifiées d'ACI pour une image (ici, une multi sous-bande après la TOD) lorsque celle-ci est utilisée pour réduire la redondance spectrale. D'autre part, la transformation convolutive utilisée toute seule pour réduire les redondances spatio-spectrales dans un schéma de compression, ne permet pas d'avoir des performances meilleures que celles des schémas de compression où deux transformations indépendantes et différentes sont utilisées pour réduire chacune une de ces deux redondances. Afin de tirer des conclusions définitives sur l'utilisation d'une transformation convolutive en compression d'images, il nous parait utile et important d'explorer d'autres voies, l'une étant de ne plus utiliser une méthode BFGS, mas plutôt une méthode quasi-Newton basée sur une estimation réelle du hessien à partir du développement de Taylor à l'ordre 2 du critère obtenu, l'autre étant de rechercher d'autres méthodes de minimisation du critère. Ces différentes voies constituent des pistes à explorer et qui pourraient permettre d'avoir des conclusions un peu plus élaborées. Il nous parait aussi intéressant de voir dans la continuité de

ces travaux de thèse, si les transformations de séparation-déconvolution pourraient avoir une autre utilité que la compression, nous pensons notamment au problème de débruitage des images. Enfin, il nous parait aussi utile, pour les schémas de la première partie, de rechercher les transformations modifiées à base d'ACI qui permettent effectivement de maximiser le gain de codage généralisé pour le schéma séparable adaptatif. Ces pistes pourraient permettre d'avoir une meilleure idée sur l'application des transformations à base d'ACI à la compression des images multi-composantes.

Bibliographie

[1] http ://telsat.belspo.be/beo/fr/guide/deftele.asp ?section=1.

[2] http ://www.spotimage.fr/html/_65_.php.

[3] http ://landsat.gsfc.nasa.gov/.

[4] http ://aviris.jpl.nasa.gov/.

[5] http ://www.ccrs.nrcan.gc.ca/optic/hyper/index_f.php.

[6] http ://www.cnes-tv.com/dossiers/spot5/pdf/pleiades.pdf.

[7] http ://asterweb.jpl.nasa.gov/characteristics.asp.

[8] Isidore Paul Akam Bita, Dinh-Tuan Pham et Michel Barret, A Separation Deconvolution Modelling for Multicomponent Images Compression, Workshop TBICA Juillet 2006, Paris, France, pp. 31–34.

[9] Isidore Paul Akam Bita, Michel Barret et Dinh-Tuan Pham, Compression of Multicomponent Satellite Images Using Independent Component Analysis, Proceedings of the Sixth International Conference on Independent Component Analysis and Blind Signal Separation (ICA 2006), Charleston, Usa, vol. LNCS 3889, pp. 335–342, Mars 2006.

[10] Marc Antonini, Michel Barlaud, Bernard Rougé et Cathérine Lambert-Nebout, Weighted Optimum Bit Allocation for multiresolution Satellite Image Coding, Proceedings of GRETSI (Signal and Image Processing), Juan-les-pins, France, Vol. 1, pp. 445–458, 1993.

[11] Marc Antonini, Michel Barlaud, Pierre Mathieu et Ingrid Daubechies, Image Coding Using Wavelet Transform, IEEE Transation on Image Processing, Vol. 1, No. 2, pp. 205–220, Avril 1992

[12] Michel Barlaud et Claude Labit, *Compression et Codage des Images et des Vidéos*, Lavoisier 2002.

[13] A. J. Bell et T. J. Sejnowski, An Information-Maximisation Approach to Blind Separation and Blind Deconvolution, Neural Computation, Vol. 7, pp. 1129–1159, 1995.

[14] Adel Belouchrani, Karim Abed Meraim, Jean-François Cardoso et Éric Moulines, A Blind Source Separation Technique Based on Second Order Statistics, IEEE Transactions on Signal Processing, Vol. 45, No. 2, pp. 434–443, Février 1997.

[15] Ali Bilgin, Philip J. Sementilli, Fang Sheng, Michael W. Marcellin, Scalable Image Coding Using Reversible Integer Wavelet Transforms, IEEE Transactions On Image Processing, Vol. 9, No. 11, pp. 1972–1977, Novembre 2000.

[16] H. Blume et A. Fand, Reversible and Irreversible Image Data Compression Using thr S-Transform and Lempel-Ziv Coding, Proccedings of SPIE, Vol. 1091, pp. 2–18, 1989.

[17] Michael M. Bronstein, Alexander M. Bronstein, Michael Zibulevsky, Yehoshua Y. Zeevi, Blind Deconvolution of Images using Optimal Sparse Representations, IEEE Transactions on Image Processing, Vol. 14, No. 6, pp. 726–736, Juin 2005.

[18] Jean-François Cardoso, Multidimensional Independent Component Analysis, Proceedings of ICASSP, Seattle, 1998.

[19] Jean-François Cardoso, Source Separation using Higher Order Moments, Proceedings of ICASSP, pp. 2109–2112, 1989.

[20] Jean-François Cardoso et Antoine Souloumiac, Blind Beamforming for non Gaussian Signal, IEE Proceedings-F, Vol. 140, No. 6, pp. 362–370, 1993.

[21] Jean-François Cardoso, Blind Signal Separation : Statistical Principles, IEEE Special Issue on Blind Identification and Estimation, Vol. 9, No. 10, pp. 2009–2025, Octobre 1998.

[22] Marc Castella et Jean-Christophe Pesquet, An Iterative Blind Source Separation Method for Convolutive Mixtures of Images, Proceedings of the 5th International Conference On Independent Component Analysis and Blind Signal Separation, Vol. 3195, pp. 922–929, 2004.

[23] Marc Castella, *Separation de Sources non Linéaires dans le Cas des Mélanges Convolutifs*, Université de Marne la vallée, Décembre 2004.

[24] Mohamed Chikhaoui, Ferdinand Bonn, Amadou Idrissa Bokoye et Abdelaziz Merzouk, A spectral Index for Land Degradation Mapping using ASTER Data : Application to a Semi-Arid Mediterranean Catchment, International Journal of Applied Earth Observation and Geoinformation, Vol. 7, pp. 140–153, Décembre 2005.

[25] Emmanuel Christophe et William A. Pearlman, Three-dimensional SPIHT Coding of Volume Images with Random Access and Resolution Scalability, Submitted to IEEE Transactions on Image Processing, Mai 2006.

[26] Pierre Comon, Independent Component Analysis, A New Concept, Signal Processing, vol. 36, pp. 287–314, 1994.

[27] Thomas M.Cover et Joy A.Thomas, *Elements of Information Theory*, John Wiley and Sons, 1991.

[28] P. L. Dragotti, G. Poggi and A. R P. Ragozini, Compression of Multispectral Images by Three-Dimensional SPIHT Algorith, IEEE Transaction on Geoscience and Remote Sensing, Vol. 38, No. 1, pp. 416–428, 2000.

[29] Hugh Everett, Generalized Lagrange Multiplier Method for Solving Problems of Optimum Allocation of Resources, Operation Research, Vol. 11, pp. 399–417, 1963.

[30] D. J. Field, Relation Between the Statistics of Natural Images and the Response Properties of Cortical Cells, Journal of The Optical Society of America A, Vol. 4, No. 12, pp. 2379–2394, Décembre 1987.

[31] A. Gersho and R. M. Gray, *Vector quantization and signal compression*, Kluwer Academic Publisher, 1992.

[32] M. Girolami, *Self-Organization Neural Networks - Independent Component Analysis and Blind Source Separation*, Springer-Verlag, 1999.

[33] M. J. Gormish, E. L. Schwartz, A. F. Keith, M. P. Boliek et A. Zandi, Lossless and Nearly Lossless Compression for High Quality Images, Proceedings of SPIE, Vol. 3025, pp. 62–70, Mars 1997.

[34] Robert M. Gray et David L. Neuhoff, *Quantization*, IEEE Transactions On Information Theory, Vol. 44, No. 6, pp. 2325–2384, Octobre 1998.

[35] David H. Hubel, *Eye, Brain, and Vision*, Henry Holt and Company, Scientific American Library Series, Janvier 1988.

[36] David A. Huffman, A Method for the Construction of Minimum-Redundancy Codes, Proceedings of the IRE, pp. 1098–1101, Septembre 1952.

[37] Aapo Hyvarinen, Juha Tarhunen et Erkki Oja, *Indepenent Component Analysis*, Wiley Interscience 2001.

[38] Aapo Hyvarinen, Fast and Robust Fixed-point Algorithms for Independent Component Analysis, IEEE Transactions on Neural Network, Vol. 10, No. 3, pp. 626–634, 1999.

[39] Eric Incerti, *Compression d'Images : Algorithmes et Standards*, Vuibert, 2003.

[40] Asad Islam et William A. Pearlman, An Embedded and Efficient Low-Complexity Hierarchical Image Coder, Visual Communications and Image Processing, Proceedigs of SPIE, Vol. 3653, pp. 294–305, Janvier 1999.

[41] ISO et CCITT, *Digital Compression and Coding of Continuous-Tone still Images*, ISO/IEC/ 10918-1,2,3 Information Technology.

[42] ISO et CCITT, JPEG 2000 Image Coding System - Part 1 : Core Coding System, ISO/IEC/ 15444-1 Information Technology.

[43] *JPEG2000 Verification Model 9.1 (Technical description)*, ISO/IEC JTC 1/SC 29/WG 1 WG1 N2165, Juin 2001.

[44] Chritian Jutten et Alain Guérin-Dugué, *De la séparation de source à l'analyse en composantes indépendantes*, Actes de l'école d'été de Printemps, Villard-de-Lans(Isere), 2001.

[45] Christian Jutten et Jeanny Hérault, Independent Component Analysis versus PCA, Proceedings of EUSIPCO, Grenoble, France, pp. 643–646, 1988.

[46] Herve Le Borgne, N. Guyader, Anne Guerin-Dugué et Jeanny Hérault, Classification of Images : ICA Filters vs Human Perception,Proceedings of the 7th International Symposium on Signal Processing and its Applications, ISSPA, Paris, France, Vol. 2, pp. 251–254, Juillet 2003.

[47] R. Linsker, Self-Organization in a Perceptual Network, Computer, Vol. 21, pp. 105–107, 1988.

[48] Xiaoqing LIU, *Analyse d'Images Couleur en Composantes Indépendantes par Réseau de Neurones*, Institut National Polytechnique de Grenoble, Décembre 1991.

[49] J. Liu et P. Moulin, Information-Theoretic Analysis of Interscale and Intrascale Dependencies Between Image Wavelet Coefficients, IEEE Transactions on Image Processing, Vol. 10, No. 11, pp. 1647–1658, 2001.

[50] Henri Maitre, *Le Traitement des Images*, Hermes, 2003.

[51] Michel Narozny, Michel Barret, Dinh-Tuan Pham et Isidore Paul Akam Bita, Modified ICA Algorithm for Finding Optimal Transforms in Transform Coding, IEEE 4th Int. Symp. on Image and Signal Processing Analysis, pp. 111–116, Septembre 2005.

[52] Michel Narozny, *Analyse en composantes Indépendantes et Compression de Données*, Université Paris XI, 2005.

[53] C. Nikias et A. Petropulu, *Higher Order Spectral Analysis - A Nonlinear Signal Processing Framework*, Prentice Hall, 1993.

[54] Danielle Nuzillard et Albert Bijaoui, Blind Source Separation and Analysis of Multispectral Astronomical Images, Astronomy and Astrophysics, Septembre 2000.

[55] B. A. Olshausen et D. J. Field, Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Natural Images, Nature, Vol. 381, No. 13, pp. 607–609, Juin 1996.

[56] Christophe Parisot, *Allocations Basées Modèles et Transformée en Ondelettes au Fil de l'Eau pour le Codage d'Images et de Vidéos*, Université de Nice - Sophia Antipolis, 2003.

[57] C. Parisot, S. Tramini, M. Antonini, M. Barlaud, C. Latry et C. Lambert-Nebout, Optimisation d'une Chaine Image de Télédétection : de la Compression Embarquée aux Post-Traitements Sol, Proceedings of GRETSI (Signal and Image Processing), Toulouse, France, Septembre 2001.

[58] C. Parisot, S. Tramini, M. Antonini, M. Barlaud, C. Latry et C. Lambert-Nebout, Optimisation of the Joint Coding/Decoding Struture, Proceedings of IEEE International Conference on Image Processing, Thessaloniki, Greece, Vol. 3, pp. 470–473, Octobre 2001.

[59] Dinh-Tuan Pham, Fast Algorithms for Mutual Information Based Independent Component Analysis, IEEE Transaction on Signal Processing, Vol. 52, No. 10, pp. 2690–2700, 2004.

[60] Dinh-Tuan Pham, Philippe Garrat et Christian Jutten, Separation of Mixture of Independent Sources through a Maximum Likelihood Approach, Proceedings of European Signal Processing, pp. 771-774, 1992.

[61] Dinh-Tuan Pham, Mutual Information Approach to Blind Separation-Deconvolution, Proceeding of the 13th European Signal Processing Conference (EUSIPCO), Antalya, Turquie, Septembre 2005.

[62] Dinh-Tuan Pham, Entropy of Random Variable Slightly Contaminated with Another, IEEE Letter on Signal Processing, 2005.

[63] William H. Press, William T. Vetterling, Saul A. Teukolsky et Brian P.Flannery, *Numerical Recipes in C*, University of Cambridge, Second Edition, 1992.

[64] Majid Rabbani et Rajan Joshi, An overview of the JPEG2000 still Image compression Standard, ISO WG1 N2233, Juillet 2001.

[65] Bharath Ramakrishna, Jing Wang, Chein-I Chang, Antonio Plaza, Hsuan Ren, Chein-chi Chang, Janet L. Jensen, james O. Jensen, Spectral/Spatial Hyperspectral Image Compression in Conjunction with Virtual Dimensionality, Proceedings of SPIE, Vol. 5806, pp. 772–781, Juin 2005.

[66] Jorma Rissanen, Generalized Kraft Inequality and Arithmetic Coding, IBM Journal of Research and Development, pp. 198–203, Mai 1976.

[67] Amir Said et William A. Pearlman, A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees, IEEE Transactions on Circuits and System for Video Technology, Vol. 6, pp. 243–250, Juin 1996.

[68] Amir Said et William A. Pearlman, An Image Multiresolution Representation for Lossless and Lossy Compression, IEEE Transactions on Image Processing, Vol. 5, pp. 1303–1310, Septembre 1996.

[69] Yair Shoham et Allen Gersho, Efficient Bit Allocation for an Arbitrary Set of Quantizers, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 36, No. 9, pp. 1445–1453, Septembre 1988.

[70] Wim Sweldens, The Lifting Scheme : A Custom-design Construction of Biorthogonal Wavelets, Applied and Computational Harmonic Analysis, Vol. 3, No. 2, pp. 186–200, Avril 1996.

[71] David S. Taubman et Michael W. Marcellin, *JPEG2000 : Image Compression Fundmentals, Standards and Practice*, Kluwer Academic, 2002.

[72] David Taubman, Erik Ordentlich, Marcelo Weinberger et Gadiel Seroussi, Embedded Block Coding in JPEG 2000, Signal Processing : Image Communication, Vol. 17, pp. 49–72, 2002.

[73] Bryan Usevitch, Optimal Bit Allocation for Biorthogonal Wavelet Coding, Proceedings of Data Compression Conference, pp. 387–395, Avril 1996.

[74] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall Signal Processing Series, 1993.

[75] Jacques Vaisey, Michel Barlaud et Marc Antonini, Multispectral Image Coding Using Lattice VQ and the Wavelet Transform, IEEE International Conference on Image Processing, Chicago, USA, pp. 307–311, Octobre 1998.

[76] John W. Woods et T. Naven, A Filter Based Bit Allocation Scheme for Subband Compression of HDTV, IEEE Transactions on Image Processing, Vol. 1, No. 3, pp. 436–440, Juillet 1992.

[77] A. Zandi, J. Allen, E. Schwartz et M. Bolick, CREW : Compresson with Reversible Embedded Wavelets, IEEE Data Compression, pp. 212–221, Mars 1995.

[78] J. Ziv et A. Lempel, A Universal Algorithm for Sequential Data Compression, IEEE Transactions on Information Theory, Vol. 23, No. 3, pp. 337–343, Mai 1977.

[79] J. Ziv et A. Lempel, Compression of Individuals Sequences via Variable Rate Coding, IEEE Transactions on Information Theory, Vol. 24, No. 5, pp. 530–536, Septembre 1978.

[80] G. Zoutendijk, *Mathematical Programming Methods*, North Holland 1976.

Annexe1

Dans cette annexe, nous présentons les tableaux récapitulant le PRSB pour des valeurs particulières de débits pour les différents cas de figure étudiées. Les tableaux sont données par groupe d'images ayant la même profondeur pour un type de schéma de compression. Les images Montp-ref et Toul-ref représentent les images de reférences des villes de Montpellier et de Toulouse. Pour les images décalées de ces images de reférence, nous les avons appelées Toul-020, Toul-035, Toul-050 pour désigner les images décalées de 0.2, 0.35 et 0.5 pixel de l'image de reférence de Toulouse, idem pour celles de Montpellier. Id représente le PRSB pour la transformation Identité, Klt celui de la TKL, Icaorth et Icaopt repésentent le PRSB pour les transformtions ICA_orth et ICA_opt. Nous avons appelé l'image Fort_France Raw et l'image de Port_debouc dans le tableaux est notée Pgx. Nous avons choisi ces noms simplifiés qui représentent l'extension de ces images originales.

Pour ce qui est des images hyper-spectrales,nous désignerons par Jasper_224 l'image qui a été découpé sur l'axe spectral en blocs de $N_1 = 224$ images (c'est-à-dire toutes les bandes spectrales), Jasper_20 représente l'image de Jasper qui a été découpé sur l'axe spectral en blocs de taille $N_1 = 20$ bandes spectrales contiguës à partir de la première bande, la dernière contenant $N_2 = 24$ bandes spectrales. Nous appellerons Jasper_45 l'image découpée en blocs de $N_1 = 45$ bandes spectrales contiguës, la dernier bloc contenant $N_2 = 44$ bandes. Nous utilisons les mêmes notations pour les images Moffett et Cuprite.

Le débit est estimé en faisant une estimation de l'entropie d'ordre 1, sauf dans certains cas qui sont précisés, où nous utilisons le $VM9$ dans la chaîne de codage et de décodage.

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Colorado	Id	30.24	33.28	35.44	37.27	40.46	43.39	46.31	49.23
	Klt	35.80	39.57	41.97	43.86	46.88	49.69	52.44	55.40
	Icaorth	35.78	39.59	42.02	43.93	46.96	49.79	52.54	55.50
	Icaopt	35.80	39.59	42.02	43.94	46.99	49.80	52.59	55.53
Rio	Id	34.82	37.21	39.00	40.63	43.61	46.59	49.51	52.52
	Klt	37.67	40.71	42.87	44.65	47.61	50.38	53.29	56.25
	Icaorth	37.68	40.76	42.93	44.72	47.67	50.44	53.36	56.31
	Icaopt	37.69	40.76	42.94	44.71	47.67	50.49	53.36	56.37
Paris	Id	29.84	31.97	33.58	35.06	37.94	40.81	43.91	46.94
	Klt	32.48	35.33	37.65	39.64	43.02	46.04	49.04	51.99
	Icaorth	32.49	35.39	37.73	39.71	43.09	46.09	49.09	52.03
	Icaopt	32.49	35.40	37.74	39.75	43.13	46.10	49.14	52.06
Tokyo	Id	37.10	38.87	40.38	41.78	44.57	47.53	50.56	53.74
	Klt	39.19	41.71	43.68	45.28	48.10	50.93	53.90	56.94
	Icaorth	39.19	41.72	43.68	45.30	48.13	50.97	53.93	56.96
	Icaopt	39.19	41.72	43.69	45.31	48.16	50.98	53.92	57.02
Lena	Id	31.91	34.51	36.07	37.43	40.11	42.93	45.99	48.98
	Klt	33.65	35.89	37.45	38.90	41.52	44.15	47.22	50.36
	Icaorth	33.65	35.89	37.46	38.91	41.52	44.16	47.24	50.37
	Icaopt	33.65	35.91	37.47	38.91	41.53	44.19	47.28	50.41
Mandrill	Id	21.99	23.95	25.66	27.10	29.85	32.76	35.76	38.78
	Klt	23.67	26.23	27.99	29.46	32.04	34.70	37.77	40.84
	Icaorth	23.67	26.24	28.00	29.46	32.04	34.72	37.79	40.84
	Icaopt	23.71	26.26	28.01	29.48	32.06	34.73	37.81	40.86
Peppers	Id	32.64	36.38	38.43	40.10	43.03	45.63	48.51	51.60
	Klt	34.85	38.56	41.00	42.99	46.83	50.10	52.57	55.03
	Icaorth	34.84	38.64	41.11	43.18	47.10	50.34	52.93	55.29
	Icaopt	34.86	38.65	41.11	43.18	47.11	50.36	52.95	55.30
Raw	Id	39.60	42.80	44.86	46.49	49.21	51.94	54.81	57.86
	Klt	43.57	46.70	48.29	49.43	51.75	54.37	57.43	60.40
	Icaorth	43.57	46.76	48.33	49.47	51.80	54.45	57.52	60.47
	Icaopt	43.65	46.80	48.38	49.51	51.84	54.52	57.61	60.69
Toulouse1	Id	30.55	33.33	35.37	37.14	40.26	43.15	46.00	48.92
	Klt	31.78	35.13	37.41	39.37	42.75	45.64	48.29	51.20
	Icaorth	31.81	35.15	37.43	39.39	42.76	45.64	48.30	51.22
	Icaopt	31.81	35.13	37.43	39.39	42.76	45.66	48.33	51.24
Toulouse2	Id	30.77	33.61	35.66	37.46	40.59	43.47	46.33	49.26
	Klt	32.12	35.54	37.92	39.96	43.42	46.37	49.01	51.89
	Icaorth	32.12	35.55	37.92	39.96	43.43	46.37	49.02	51.89
	Icaopt	32.11	35.55	37.93	39.96	43.43	46.37	49.02	51.89

TAB. 6.1 – PRSB des images de 8 bpp du schéma en sous-bandes

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Pgx	Id	29.79	33.03	35.37	37.33	40.82	43.99	47.09	50.14	53.14	56.17
	Klt	32.64	36.84	39.96	42.51	46.60	49.98	53.02	55.89	58.71	61.74
	Icaorth	32.71	36.93	40.11	42.74	46.96	50.30	53.28	56.13	58.96	62.05
	Icaopt	32.72	36.99	40.22	42.85	47.07	50.44	53.42	56.24	59.15	62.16
Moissac	Id	35.41	38.77	40.98	42.85	46.13	49.12	52.05	54.97	58.03	61.02
	Klt	37.53	41.40	44.20	46.62	50.71	53.91	56.59	59.39	62.44	65.56
	Icaorth	37.59	41.55	44.38	46.79	50.88	54.04	56.71	59.53	62.56	65.68
	Icaopt	37.61	41.59	44.44	46.85	50.92	54.09	56.76	59.58	62.63	65.75
Strasbourg	Id	30.20	33.61	36.02	38.07	41.67	44.96	48.05	50.99	53.92	56.85
	Klt	31.82	35.85	38.83	41.46	45.93	49.57	52.59	55.30	58.11	61.19
	Icaorth	31.81	35.86	38.89	41.55	46.09	49.77	52.77	55.44	58.25	61.36
	Icaopt	31.80	35.88	38.91	41.58	46.14	49.81	52.82	55.48	58.30	61.40
Vannes	Id	37.88	41.51	44.00	46.11	49.73	52.76	55.60	58.45	61.4	64.46
	Klt	40.02	44.41	47.37	49.72	53.20	55.75	58.38	61.39	64.51	67.60
	Icaorth	40.16	44.61	47.59	49.93	53.37	55.90	58.55	61.56	64.67	67.77
	Icaopt	40.23	44.70	47.66	49.99	53.42	55.93	58.64	61.62	64.77	67.86
Montp-ref	Id	35.81	38.78	41.05	43.01	46.54	49.71	52.67	55.61	58.51	61.53
	Klt	37.55	41.05	43.87	46.35	50.71	54.47	57.56	60.20	62.99	66.16
	Icaorth	37.55	41.12	44.07	46.66	51.18	54.84	57.90	60.50	63.33	66.45
	Icaopt	37.58	41.16	44.12	46.71	51.22	54.87	57.93	60.55	63.38	66.49
Montp-020	Id	35.81	38.78	41.04	43.01	46.53	49.70	52.67	55.61	58.51	61.54
	Klt	37.42	40.71	43.26	45.49	49.32	52.67	55.73	58.54	61.44	64.46
	Icaorth	37.44	40.80	43.46	45.75	49.66	53.01	56.05	58.87	61.78	64.76
	Icaopt	37.46	40.83	43.50	45.78	49.69	53.04	56.08	58.90	61.82	64.78
Montp-035	Id	35.80	38.78	41.04	43.01	46.53	49.70	52.66	55.61	58.50	61.54
	Klt	37.24	40.34	42.76	44.89	48.57	51.82	54.80	57.73	60.58	63.62
	Icaorth	37.27	40.43	42.92	45.08	48.83	52.10	55.09	57.99	60.82	63.91
	Icaopt	37.28	40.46	42.94	45.10	48.84	52.12	55.11	58.01	60.84	63.93
Montp-050	Id	35.8	38.78	41.04	43.00	46.53	49.69	52.66	55.61	58.5	61.53
	Klt	37.06	40.07	42.43	44.48	48.05	51.22	54.23	57.12	60.07	63.03
	Icaorth	37.09	40.12	42.53	44.61	48.24	51.44	54.43	57.34	60.24	63.22
	Icaopt	37.10	40.14	42.54	44.62	48.26	51.45	54.44	57.36	60.25	63.23
Toul-ref	Id	34.14	36.97	38.95	40.71	43.95	46.98	50.01	52.98	56.00	58.96
	Klt	35.98	39.36	41.90	44.11	47.75	50.96	54.00	56.88	59.83	62.83
	Icaorth	36.04	39.45	42.02	44.24	47.93	51.24	54.26	57.13	60.08	63.06
	Icaopt	36.05	39.48	42.05	44.29	48.01	51.31	54.34	57.25	60.16	63.16
Toul-020	Id	34.15	36.98	38.95	40.71	43.95	46.98	50.00	52.95	56.00	58.93
	Klt	36.13	39.20	41.63	43.74	47.23	50.38	53.32	56.26	59.17	62.24
	Icaorth	36.17	39.28	41.74	43.87	47.40	50.62	53.65	56.53	59.49	62.47
	Icaopt	36.18	39.29	41.77	43.89	47.43	50.65	53.68	56.56	59.53	62.51
Toul-035	Id	34.15	36.98	38.95	40.71	43.95	46.98	50.00	52.95	56.00	58.92
	Klt	35.90	38.75	40.97	42.94	46.31	49.36	52.35	55.27	58.28	61.25
	Icaorth	35.94	38.80	41.04	43.03	46.42	49.54	52.53	55.51	58.46	61.46
	Icaopt	35.95	38.81	41.06	43.05	46.45	49.56	52.56	55.54	58.48	61.48
Toul-050	Id	34.15	36.98	38.96	40.71	43.95	46.97	50.00	52.95	56.00	58.92
	Klt	35.62	38.29	40.37	42.27	45.51	48.56	51.54	54.49	57.48	60.49
	Icaorth	35.66	38.36	40.49	42.41	45.68	48.75	51.77	54.72	57.74	60.68
	Icaopt	35.67	38.37	40.51	42.43	45.69	48.77	51.80	54.73	57.76	60.71

Tab. 6.2 – PRSB des images de 12 bpp du schéma en sous-bandes

Images	Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Cuprite_224	Id	60.57	64.10	66.76	69.09	71.25	73.29	77.01	78.73	80.40	83.51
	Klt	79.98	82.54	84.77	86.35	87.73	89.10	91.75	93.19	94.64	97.76
	Icaorth	80.05	82.60	84.82	86.39	87.76	89.13	91.78	93.22	94.68	97.81
	Icaopt	80.05	82.61	84.83	86.40	87.76	89.13	91.79	93.22	94.69	97.81
Moffett_224	Id	57.74	62.30	65.65	68.44	70.90	73.17	77.26	79.16	80.95	84.17
	Klt	79.28	82.60	85.12	87.02	88.48	89.85	92.64	94.04	95.47	98.42
	Icaorth	79.66	82.93	85.43	87.24	88.66	90.02	92.83	94.20	95.65	98.61
	Icaopt	79.70	82.96	85.46	87.27	88.69	90.04	92.85	94.23	95.68	98.66
Jasper_224	Id	55.85	59.35	62.08	64.44	66.58	68.57	72.32	74.10	75.83	79.18
	Klt	79.27	82.36	84.56	86.37	87.90	89.32	92.05	93.42	94.78	97.91
	Icaorth	79.39	82.43	84.64	86.43	87.95	89.36	92.09	93.47	94.83	97.96
	Icaopt	79.41	82.45	84.65	86.44	87.96	89.37	92.10	93.48	94.84	97.98
Cuprite_20	Id	60.19	64.01	66.72	69.07	71.24	73.29	77.01	78.72	80.39	83.51
	Klt	77.22	80.76	83.20	85.17	86.69	88.03	90.72	92.09	93.59	96.60
	Icaorth	77.24	80.77	83.21	85.19	86.70	88.03	90.73	92.10	93.59	96.61
	Icaopt	77.25	80.77	83.22	85.19	86.71	88.04	90.74	92.10	93.60	96.61
Moffett_20	Id	57.27	62.22	65.62	68.41	70.89	73.16	77.26	79.15	80.95	84.17
	Klt	75.93	80.77	83.40	85.69	87.37	88.77	91.58	92.98	94.36	97.30
	Icaorth	76.08	80.90	83.51	85.82	87.49	88.89	91.68	93.08	94.46	97.43
	Icaopt	76.11	80.92	83.53	85.84	87.51	88.91	91.69	93.10	94.47	97.45
Jasper_20	Id	55.45	59.26	62.05	64.44	66.57	68.56	72.31	74.09	75.82	79.18
	Klt	73.95	79.76	82.31	84.41	86.20	87.75	90.54	91.92	93.31	96.23
	Icaorth	73.96	79.78	82.33	84.43	86.22	87.77	90.56	91.94	93.32	96.25
	Icaopt	73.99	79.79	82.34	84.44	86.23	87.78	90.57	91.94	93.33	96.26
Cuprite_45	Id	60.26	64.06	66.74	69.08	71.25	73.28	76.93	78.32	79.36	80.40
	Klt	78.83	81.60	84.01	85.75	87.19	88.52	91.21	92.54	93.82	95.93
	Icaorth	78.86	81.62	84.04	85.77	87.21	88.54	91.23	92.56	93.85	95.96
	Icaopt	78.87	81.63	84.05	85.78	87.21	88.54	91.24	92.57	93.85	95.96
Moffett_45	Id	57.41	62.27	65.63	68.41	70.90	73.17	77.28	79.09	80.43	81.69
	Klt	77.63	81.71	84.30	86.45	88.00	89.41	92.09	93.29	94.39	95.96
	Icaorth	77.84	81.87	84.46	86.61	88.15	89.55	92.23	93.42	94.53	96.21
	Icaopt	77.88	81.90	84.49	86.63	88.17	89.57	92.25	93.45	94.56	96.35
Jasper_45	Id	55.53	59.31	62.05	64.44	66.58	68.56	72.30	73.92	75.21	76.46
	Klt	76.42	81.09	83.35	85.34	87.03	88.48	91.29	92.58	93.80	96.02
	Icaorth	76.52	81.14	83.40	85.38	87.07	88.52	91.33	92.63	93.84	96.09
	Icaopt	76.54	81.15	83.41	85.39	87.08	88.52	91.34	92.63	93.85	96.10

Tab. 6.3 – PRSB des images de 16 bpp du schéma en sous-bandes

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Colorado	Id	30.24	33.28	35.44	37.27	40.46	43.39	46.31	49.23
	Klt	34.96	38.72	41.16	43.13	46.36	49.25	51.99	54.87
	Icaorth	35.27	39.08	41.57	43.54	46.69	49.60	52.34	55.27
	Icaopt	35.32	39.11	41.60	43.56	46.71	49.61	52.38	55.29
Rio	Id	34.82	37.21	39.00	40.63	43.61	46.59	49.51	52.52
	Klt	36.89	39.85	41.92	43.71	46.80	49.74	52.63	55.59
	Icaorth	37.39	40.33	42.34	44.04	47.05	50.03	52.93	55.92
	Icaopt	37.48	40.54	42.69	44.46	47.42	50.32	53.24	56.25
Paris	Id	29.84	31.97	33.58	35.06	37.94	40.81	43.91	46.94
	Klt	31.97	34.80	37.03	39.00	42.35	45.42	48.48	51.44
	Icaorth	32.31	35.22	37.56	39.54	42.90	45.92	48.92	51.89
	Icaopt	32.34	35.25	37.59	39.58	42.94	45.94	48.97	51.90
Tokyo	Id	37.10	38.87	40.38	41.78	44.57	47.53	50.56	53.74
	Klt	38.73	41.23	43.18	44.86	47.79	50.62	53.59	56.62
	Icaorth	38.81	41.27	43.19	44.87	47.82	50.66	53.63	56.64
	Icaopt	38.81	41.28	43.20	44.89	47.82	50.70	53.62	56.69
Lena	Id	31.91	34.51	36.07	37.43	40.11	42.93	45.99	48.98
	Klt	32.36	34.63	36.12	37.55	40.30	43.12	46.26	49.39
	Icaorth	32.81	35.07	36.69	38.14	40.87	43.74	46.89	50.09
	Icaopt	32.83	35.07	36.67	38.12	40.84	43.75	46.93	50.10
Mandrill	Id	21.99	23.95	25.66	27.10	29.85	32.76	35.76	38.78
	Klt	23.58	26.10	27.81	29.26	31.86	34.57	37.64	40.70
	Icaorth	23.61	26.13	27.88	29.35	31.95	34.62	37.68	40.74
	Icaopt	23.60	26.14	27.88	29.35	31.95	34.62	37.68	40.73
Peppers	Id	32.63	36.38	38.42	40.08	43.02	45.63	48.51	51.57
	Klt	33.44	37.50	39.91	41.95	45.66	48.87	51.65	54.27
	Icaorth	33.62	37.86	40.33	42.50	46.53	49.69	52.35	54.81
	Icaopt	33.67	37.86	40.37	42.56	46.62	49.75	52.33	54.84
Raw	Id	39.60	42.80	44.86	46.49	49.21	51.94	54.81	57.86
	Klt	42.14	45.73	47.80	49.07	51.32	53.98	57.06	60.08
	Icaorth	42.17	45.82	47.85	49.09	51.36	54.09	57.23	60.24
	Icaopt	42.32	45.95	47.95	49.19	51.41	54.16	57.31	60.32
Toulouse1	Id	30.55	33.33	35.37	37.14	40.26	43.15	46.00	48.92
	Klt	31.14	34.01	36.15	38.02	41.51	44.74	47.61	50.44
	Icaorth	31.65	34.67	36.85	38.72	42.05	45.23	48.12	51.11
	Icaopt	31.66	34.66	36.85	38.73	42.04	45.23	48.14	51.14
Toulouse2	Id	30.77	33.61	35.66	37.46	40.59	43.47	46.33	49.26
	Klt	31.50	34.49	36.75	38.71	42.33	45.60	48.38	51.22
	Icaorth	31.93	35.07	37.31	39.27	42.72	45.97	48.82	51.80
	Icaopt	31.93	35.07	37.31	39.27	42.72	45.97	48.82	51.80

TAB. 6.4 – PRSB des images de 8 bpp du schéma séparable avec adaptation dans l'espace ondelette

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Pgx	Id	29.79	33.03	35.37	37.33	40.82	43.99	47.09	50.14	53.14	56.17
	Klt	32.34	36.44	39.49	42.03	46.25	49.83	52.90	55.75	58.57	61.59
	Icaorth	32.43	36.61	39.69	42.29	46.57	50.12	53.15	55.98	58.80	61.89
	Icaopt	32.43	36.65	39.79	42.38	46.71	50.25	53.24	56.05	58.96	62.00
Moissac	Id	35.41	38.77	40.98	42.85	46.13	49.12	52.05	54.97	58.03	61.02
	Klt	36.82	40.51	43.23	45.69	50.16	53.58	56.31	59.07	62.12	65.23
	Icaorth	37.20	40.95	43.62	45.98	50.34	53.86	56.61	59.42	62.46	65.59
	Icaopt	37.18	41.00	43.67	46.01	50.36	53.88	56.64	59.45	62.50	65.62
Strasbourg	Id	30.20	33.61	36.02	38.07	41.67	44.96	48.05	50.99	53.92	56.85
	Klt	31.40	35.44	38.34	40.90	45.36	49.15	52.28	55.03	57.83	60.90
	Icaorth	31.56	35.58	38.53	41.12	45.65	49.46	52.53	55.27	58.10	61.20
	Icaopt	31.58	35.60	38.55	41.16	45.71	49.52	52.59	55.33	58.16	61.28
Vannes	Id	37.88	41.51	44.00	46.11	49.73	52.76	55.60	58.45	61.40	64.46
	Klt	39.24	43.16	45.87	48.11	52.13	55.11	57.73	60.69	63.86	66.93
	Icaorth	39.61	43.72	46.55	48.88	52.83	55.61	58.21	61.38	64.51	67.62
	Icaopt	39.66	43.81	46.60	48.91	52.83	55.62	58.25	61.44	64.59	67.71
Montp-ref	Id	35.81	38.78	41.05	43.01	46.54	49.71	52.67	55.61	58.51	61.53
	Klt	37.28	40.69	43.45	45.87	50.16	53.88	56.99	59.75	62.52	65.63
	Icaorth	37.42	40.96	43.86	46.43	50.93	54.63	57.71	60.34	63.22	66.40
	Icaopt	37.44	40.97	43.90	46.49	50.96	54.67	57.75	60.41	63.31	66.47
Montp-020	Id	35.81	38.78	41.04	43.01	46.53	49.70	52.67	55.61	58.51	61.54
	Klt	37.09	40.25	42.74	44.88	48.67	52.02	55.06	57.98	60.81	63.90
	Icaorth	37.27	40.60	43.23	45.50	49.36	52.70	55.75	58.59	61.54	64.62
	Icaopt	37.32	40.62	43.25	45.53	49.40	52.74	55.81	58.65	61.61	64.68
Montp-035	Id	35.80	38.78	41.04	43.01	46.53	49.70	52.66	55.61	58.50	61.54
	Klt	36.94	39.92	42.24	44.24	47.82	51.08	54.17	57.06	60.03	62.97
	Icaorth	37.08	40.23	42.69	44.81	48.49	51.74	54.74	57.70	60.61	63.74
	Icaopt	37.11	40.25	42.72	44.84	48.52	51.78	54.78	57.73	60.65	63.79
Montp-050	Id	35.80	38.78	41.04	43.00	46.53	49.69	52.66	55.61	58.50	61.53
	Klt	36.75	39.57	41.79	43.70	47.13	50.35	53.38	56.33	59.25	62.30
	Icaorth	36.90	39.90	42.27	44.28	47.83	50.98	54.03	56.95	59.99	63.02
	Icaopt	36.93	39.92	42.29	44.31	47.85	51.01	54.05	56.97	60.01	63.05
Toul-ref	Id	34.14	36.97	38.95	40.71	43.95	46.98	50.01	52.98	56.00	58.96
	Klt	35.72	39.06	41.55	43.74	47.43	50.70	53.79	56.71	59.63	62.62
	Icaorth	35.91	39.31	41.87	44.08	47.76	51.05	54.10	57.00	59.94	62.93
	Icaopt	35.92	39.34	41.91	44.13	47.83	51.16	54.20	57.13	60.07	63.07
Toul-020	Id	34.15	36.98	38.95	40.71	43.95	46.98	50.00	52.95	56.00	58.93
	Klt	35.84	38.86	41.26	43.35	46.85	50.02	53.00	55.99	58.89	61.96
	Icaorth	36.00	39.13	41.59	43.70	47.22	50.47	53.49	56.43	59.39	62.39
	Icaopt	36.02	39.14	41.60	43.72	47.25	50.49	53.53	56.46	59.43	62.42
Toul-035	Id	34.15	36.98	38.95	40.71	43.95	46.98	50.00	52.95	56.00	58.92
	Klt	35.69	38.53	40.76	42.73	46.09	49.18	52.23	55.19	58.21	61.17
	Icaorth	35.80	38.64	40.87	42.85	46.22	49.33	52.37	55.37	58.36	61.34
	Icaopt	35.81	38.65	40.89	42.87	46.25	49.35	52.38	55.39	58.38	61.36
Toul-050	Id	34.15	36.98	38.96	40.71	43.95	46.97	50.00	52.95	56.00	58.92
	Klt	35.43	38.08	40.19	42.07	45.31	48.37	51.38	54.41	57.41	60.42
	Icaorth	35.52	38.18	40.29	42.18	45.43	48.50	51.52	54.54	57.56	60.54
	Icaopt	35.53	38.20	40.31	42.21	45.45	48.52	51.54	54.56	57.58	60.56

TAB. 6.5 – PRSB des images de 12 bpp du schéma séparable avec adaptation dans l'espace ondelette

Images	Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Cuprite_20	Id	60.19	64.01	66.72	69.07	71.24	73.29	77.01	78.72	80.39	83.51
	Klt	76.67	80.49	82.93	84.94	86.47	87.82	90.51	91.88	93.36	96.39
	Icaorth	76.05	80.24	82.67	84.72	86.26	87.63	90.33	91.71	93.17	96.22
	Icaopt	76.71	80.56	83.01	85.01	86.54	87.87	90.56	91.94	93.42	96.46
Moffett_20	Id	57.27	62.22	65.62	68.41	70.89	73.16	77.26	79.15	80.95	84.17
	Klt	75.33	80.42	83.07	85.42	87.13	88.55	91.35	92.75	94.16	97.09
	Icaorth	75.35	80.52	83.20	85.58	87.28	88.68	91.47	92.87	94.27	97.21
	Icaopt	75.70	80.73	83.34	85.70	87.38	88.79	91.57	92.98	94.37	97.33
Jasper_20	Id	55.45	59.26	62.05	64.44	66.57	68.56	72.31	74.09	75.82	79.18
	Klt	73.30	79.41	82.03	84.13	85.94	87.51	90.29	91.69	93.07	96.00
	Icaorth	73.40	79.44	82.09	84.20	86.01	87.59	90.37	91.75	93.13	96.07
	Icaopt	73.60	79.59	82.18	84.29	86.08	87.65	90.43	91.81	93.20	96.13
Cuprite_224	Id	60.57	64.10	66.76	69.09	71.25	73.29	77.01	78.73	80.40	83.51
	Klt	79.52	82.14	84.40	86.04	87.43	88.74	91.45	92.85	94.33	97.41
	Icaorth	78.19	81.44	83.78	85.43	86.75	88.00	90.65	92.09	93.65	96.69
	Icaopt	78.62	81.63	83.98	85.65	86.99	88.25	90.95	92.37	93.92	96.97
Moffett_224	Id	57.74	62.30	65.65	68.44	70.90	73.17	77.26	79.16	80.95	84.17
	Klt	78.68	82.19	84.71	86.68	88.16	89.53	92.29	93.74	95.14	98.16
	Icaorth	77.35	81.45	84.03	86.19	87.56	88.82	91.60	93.05	94.43	97.45
	Icaopt	78.37	82.10	84.66	86.65	88.02	89.33	92.14	93.62	95.01	98.05
Jasper_224	Id	55.85	59.35	62.08	64.44	66.58	68.57	72.32	74.10	75.83	79.18
	Klt	78.47	81.89	84.10	85.97	87.53	88.92	91.69	93.06	94.42	97.49
	Icaorth	77.34	81.38	83.56	85.42	86.98	88.31	91.04	92.39	93.84	96.87
	Icaopt	78.11	81.68	83.88	85.75	87.28	88.63	91.41	92.78	94.17	97.24

Tab. 6.6 – PRSB des images de 16 bpp du schéma séparable avec adaptation dans l'espace ondelette

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Colorado	Id	30.24	33.28	35.44	37.27	40.46	43.39	46.31	49.23
	Klt	35.69	39.37	41.68	43.52	46.56	49.28	52.02	54.99
	Icaorth	35.68	39.40	41.74	43.59	46.65	49.40	52.11	55.09
	Icaopt	35.69	39.46	41.86	43.75	46.75	49.46	52.20	55.19
Rio	Id	34.82	37.21	39.00	40.63	43.61	46.59	49.51	52.52
	Klt	37.58	40.62	42.77	44.55	47.45	50.26	53.13	56.13
	Icaorth	37.59	40.62	42.77	44.55	47.45	50.26	53.14	56.14
	Icaopt	37.57	40.58	42.71	44.49	47.38	50.20	53.08	56.06
Paris	Id	29.84	31.97	33.58	35.06	37.94	40.81	43.91	46.94
	Klt	32.44	35.31	37.63	39.56	42.84	45.86	48.86	51.82
	Icaorth	32.42	35.29	37.63	39.59	42.87	45.88	48.87	51.83
	Icaopt	32.41	35.28	37.62	39.59	42.86	45.86	48.86	51.81
Tokyo	Id	37.10	38.87	40.38	41.78	44.57	47.53	50.56	53.74
	Klt	39.14	41.56	43.45	45.05	47.89	50.71	53.73	56.74
	Icaorth	39.13	41.64	43.58	45.18	47.94	50.75	53.80	56.81
	Icaopt	39.10	41.63	43.57	45.17	47.94	50.75	53.79	56.81
Lena	Id	31.91	34.51	36.07	37.43	40.11	42.93	45.99	48.98
	Klt	33.39	35.53	36.91	38.16	40.65	43.43	46.53	49.70
	Icaorth	33.33	35.48	36.85	38.09	40.57	43.35	46.46	49.60
	Icaopt	33.34	35.52	36.92	38.19	40.70	43.49	46.58	49.75
Madrill	Id	21.99	23.95	25.66	27.10	29.85	32.76	35.76	38.78
	Klt	23.32	25.73	27.45	28.89	31.54	34.30	37.29	40.40
	Icaorth	23.32	25.68	27.39	28.84	31.44	34.21	37.20	40.32
	Icaopt	23.28	25.77	27.61	29.11	31.75	34.44	37.39	40.50
Peppers	Id	32.64	36.38	38.43	40.10	43.03	45.63	48.51	51.60
	Klt	33.95	37.56	39.78	41.54	44.69	47.61	50.13	52.95
	Icaorth	34.23	37.89	40.12	41.91	45.20	48.20	50.96	53.58
	Icaopt	34.27	37.94	40.18	41.97	45.28	48.24	50.99	53.61
Raw	Id	39.60	42.80	44.86	46.49	49.21	51.94	54.81	57.86
	Klt	43.33	46.38	48.03	49.18	51.31	54.02	57.10	60.12
	Icaorth	43.51	46.56	48.15	49.30	51.38	54.16	57.30	60.37
	Icaopt	43.53	46.54	48.11	49.22	51.30	54.10	57.23	60.31
Toulouse1	Id	30.55	33.33	35.37	37.14	40.26	43.15	46.00	48.92
	Klt	31.96	35.13	37.43	39.38	42.74	45.63	48.27	51.15
	Icaorth	31.96	35.12	37.42	39.37	42.70	45.60	48.24	51.12
	Icaopt	31.96	35.13	37.42	39.36	42.70	45.59	48.24	51.13
Toulouse2	Id	30.77	33.61	35.66	37.46	40.59	43.47	46.33	49.26
	Klt	32.27	35.54	37.91	39.94	43.40	46.32	48.98	51.90
	Icaorth	32.27	35.54	37.91	39.93	43.39	46.31	48.97	51.90
	Icaopt	32.27	35.54	37.91	39.93	43.38	46.31	48.97	51.89

TAB. 6.7 – PRSB des images de 8 bpp du schéma séparable avec adaptation dans l'espace image

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Pgx	Id	29.79	33.03	35.37	37.33	40.82	43.99	47.09	50.14	53.14	56.17
	Klt	32.62	36.76	39.82	42.35	46.44	49.81	52.81	55.64	58.50	61.51
	Icaorth	32.52	36.53	39.49	41.99	46.03	49.42	52.44	55.22	58.12	61.07
	Icaopt	32.51	36.55	39.55	42.08	46.18	49.58	52.55	55.36	58.25	61.26
Moissac	Id	35.41	38.77	40.98	42.85	46.13	49.12	52.05	54.97	58.03	61.02
	Klt	37.77	41.43	44.26	46.68	50.76	53.98	56.55	59.27	62.41	65.52
	Icaorth	37.79	41.49	44.30	46.70	50.74	53.92	56.45	59.17	62.36	65.46
	Icaopt	37.80	41.52	44.34	46.74	50.78	53.96	56.52	59.26	62.43	65.55
Strasbourg	Id	30.20	33.61	36.02	38.07	41.67	44.96	48.05	50.99	53.92	56.85
	Klt	32.00	35.86	38.89	41.53	46.05	49.70	52.72	55.40	58.21	61.26
	Icaorth	31.94	35.78	38.82	41.48	46.01	49.67	52.69	55.37	58.17	61.22
	Icaopt	31.98	35.83	38.87	41.53	46.06	49.71	52.73	55.40	58.21	61.26
Vannes	Id	37.88	41.51	44.00	46.11	49.73	52.76	55.60	58.45	61.40	64.46
	Klt	39.91	43.93	46.92	49.32	52.89	55.47	57.99	61.00	64.19	67.22
	Icaorth	40.32	44.40	47.35	49.68	53.13	55.64	58.20	61.28	64.41	67.49
	Icaopt	40.36	44.48	47.43	49.75	53.16	55.65	58.21	61.30	64.42	67.51
Montp-ref	Id	35.81	38.78	41.05	43.01	46.54	49.71	52.67	55.61	58.51	61.53
	Klt	37.54	41.00	43.80	46.27	50.66	54.49	57.61	60.23	63.01	66.17
	Icaorth	37.45	40.86	43.60	46.06	50.45	54.28	57.46	60.12	62.88	66.04
	Icaopt	37.54	41.14	44.08	46.66	51.13	54.79	57.85	60.47	63.30	66.42
Montp-020	Id	35.81	38.78	41.04	43.01	46.53	49.70	52.67	55.61	58.51	61.54
	Klt	37.39	40.63	43.14	45.34	49.19	52.59	55.67	58.51	61.38	64.38
	Icaorth	37.39	40.76	43.40	45.68	49.54	52.88	55.90	58.72	61.65	64.62
	Icaopt	37.42	40.79	43.43	45.70	49.57	52.90	55.93	58.76	61.68	64.64
Montp-035	Id	35.80	38.78	41.04	43.01	46.53	49.70	52.66	55.61	58.50	61.54
	Klt	37.19	40.17	42.51	44.54	48.13	51.38	54.46	57.39	60.27	63.24
	Icaorth	37.22	40.38	42.83	44.95	48.63	51.85	54.82	57.75	60.60	63.65
	Icaopt	37.24	40.38	42.84	44.95	48.62	51.84	54.82	57.74	60.60	63.65
Montp-050	Id	35.80	38.78	41.04	43.00	46.53	49.69	52.66	55.61	58.50	61.53
	Klt	36.96	39.80	42.05	43.98	47.42	50.65	53.70	56.58	59.53	62.51
	Icaorth	37.04	40.04	42.39	44.39	47.92	51.06	54.07	56.92	59.90	62.85
	Icaopt	37.04	40.02	42.37	44.37	47.89	51.03	54.04	56.90	59.87	62.83
Toul-ref	Id	34.14	36.97	38.95	40.71	43.95	46.98	50.01	52.98	56.00	58.96
	Klt	36.12	39.28	41.83	44.03	47.68	50.98	54.04	56.89	59.84	62.79
	Icaorth	36.21	39.44	42.01	44.23	47.89	51.17	54.21	57.06	60.02	62.97
	Icaopt	36.23	39.45	42.03	44.27	47.94	51.23	54.28	57.16	60.10	63.06
Toul-020	Id	34.15	36.98	38.95	40.71	43.95	46.98	50.00	52.95	56.00	58.93
	Klt	36.06	39.15	41.59	43.71	47.23	50.46	53.45	56.38	59.32	62.33
	Icaorth	36.14	39.26	41.72	43.84	47.37	50.58	53.59	56.50	59.46	62.44
	Icaopt	36.16	39.27	41.74	43.86	47.38	50.60	53.61	56.51	59.47	62.46
Toul-035	Id	34.15	36.98	38.95	40.71	43.95	46.98	50.00	52.95	56.00	58.92
	Klt	35.84	38.69	40.90	42.86	46.23	49.31	52.34	55.3	58.30	61.25
	Icaorth	35.91	38.77	41.00	42.97	46.34	49.43	52.45	55.41	58.39	61.37
	Icaopt	35.92	38.79	41.02	42.99	46.35	49.44	52.46	55.42	58.40	61.38
Toul-050	Id	34.15	36.98	38.96	40.71	43.95	46.97	50.00	52.95	56.00	58.92
	Klt	35.51	38.17	40.23	42.08	45.29	48.35	51.33	54.34	57.29	60.32
	Icaorth	35.62	38.31	40.41	42.29	45.52	48.59	51.60	54.58	57.58	60.56
	Icaopt	35.63	38.32	40.42	42.30	45.53	48.60	51.61	54.58	57.59	60.56

TAB. 6.8 – PRSB des images de 12 bpp du schéma séparable avec adaptation dans l'espace image

Images	Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Cuprite_20	Id	60.19	64.01	66.72	69.07	71.24	73.29	77.01	78.72	80.39	83.51
	Klt	76.95	80.59	82.99	84.97	86.52	87.86	90.55	91.92	93.40	96.43
	Icaorth	76.99	80.62	83.04	85.01	86.55	87.89	90.58	91.94	93.42	96.45
	Icaopt	77.01	80.63	83.04	85.02	86.56	87.9	90.59	91.95	93.43	96.46
Moffett_20	Id	57.27	62.23	65.62	68.41	70.89	73.18	77.29	79.19	81.01	84.17
	Klt	75.51	80.45	83.19	85.51	87.25	88.70	91.50	92.90	94.27	97.23
	Icaorth	75.34	80.33	82.99	85.35	87.14	88.58	91.34	92.71	94.11	97.05
	Icaopt	75.58	80.55	83.19	85.55	87.30	88.73	91.49	92.88	94.26	97.23
Jasper_20	Id	55.45	59.26	62.05	64.44	66.57	68.56	72.31	74.09	75.82	79.18
	Klt	73.62	79.49	82.07	84.18	85.98	87.56	90.34	91.73	93.11	96.03
	Icaorth	73.62	79.51	82.10	84.21	86.01	87.59	90.37	91.75	93.13	96.06
	Icaopt	73.63	79.52	82.11	84.22	86.02	87.60	90.38	91.76	93.14	96.07
Moffett_224	Id	57.74	62.30	65.65	68.44	70.91	73.19	77.29	79.20	81.02	84.18
	Klt	78.39	82.12	84.64	86.73	88.24	89.63	92.37	93.80	95.21	98.23
	Icaorth	76.78	80.68	83.22	85.42	87.11	88.48	91.19	92.56	94.00	96.95
	Icaopt	78.67	82.16	84.75	86.77	88.24	89.57	92.31	93.76	95.18	98.21
Cuprite_224	Id	60.57	64.10	66.76	69.09	71.25	73.28	77.01	78.73	80.40	83.50
	Klt	79.54	82.14	84.43	86.09	87.48	88.81	91.50	92.89	94.37	97.46
	Icaorth	79.71	82.30	84.56	86.18	87.55	88.87	91.55	92.96	94.44	97.54
	Icaopt	79.73	82.31	84.58	86.19	87.56	88.89	91.56	92.98	94.45	97.56
Jasper_224	Id	55.85	59.35	62.08	64.44	66.58	68.57	72.33	74.11	75.83	79.19
	Klt	78.56	81.96	84.16	86.05	87.62	89.02	91.77	93.14	94.49	97.57
	Icaorth	78.41	81.89	84.15	86.04	87.6	88.99	91.74	93.11	94.45	97.52
	Icaopt	78.47	81.92	84.18	86.07	87.63	89.02	91.76	93.13	94.48	97.55

Tab. 6.9 – PRSB des images de 16 bpp du schéma séparable avec adaptation dans l'espace image

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00
Colorado	Id	30.26	33.30	35.46	37.30	40.48	43.42	46.35	49.31
	Klt	36.24	39.88	42.23	44.08	47.15	49.96	52.79	55.84
	Icaorth	36.28	39.95	42.36	44.23	47.29	50.12	52.94	55.97
	Icaopt	36.27	39.97	42.35	44.22	47.29	50.13	52.95	55.95
Rio	Id	34.84	37.23	39.02	40.65	43.64	46.61	49.55	52.57
	Klt	37.99	40.96	43.14	44.89	47.84	50.65	53.56	56.52
	Icaorth	38.06	41.08	43.27	45.04	47.96	50.79	53.68	56.65
	Icaopt	38.06	41.10	43.29	45.06	47.96	50.81	53.68	56.69
Paris	Id	29.85	31.98	33.59	35.07	37.96	40.84	43.94	46.98
	Klt	32.68	35.66	37.99	40.00	43.30	46.34	49.37	52.28
	Icaorth	32.76	35.78	38.12	40.14	43.44	46.44	49.50	52.38
	Icaopt	32.76	35.78	38.15	40.18	43.46	46.47	49.49	52.43
Tokyo	Id	37.11	38.88	40.39	41.80	44.59	47.55	50.59	53.79
	Klt	39.42	42.00	43.92	45.53	48.34	51.21	54.16	57.18
	Icaorth	39.47	42.05	44.00	45.61	48.44	51.34	54.25	57.28
	Icaopt	39.47	42.07	44.01	45.62	48.47	51.35	54.26	57.35
Lena	Id	31.94	34.53	36.09	37.45	40.14	42.97	46.06	49.09
	Klt	34.00	36.15	37.68	39.13	41.69	44.33	47.46	50.59
	Icaorth	34.08	36.20	37.76	39.23	41.76	44.40	47.55	50.66
	Icaopt	34.09	36.22	37.76	39.22	41.76	44.40	47.58	50.67
Mandrill	Id	22.00	23.96	25.67	27.12	29.87	32.78	35.79	38.82
	Klt	24.06	26.57	28.31	29.77	32.36	35.08	38.12	41.15
	Icaorth	24.09	26.60	28.34	29.81	32.38	35.14	38.16	41.22
	Icaopt	24.10	26.60	28.34	29.81	32.38	35.14	38.17	41.21
Peppers	Id	32.76	36.44	38.50	40.16	43.10	45.70	48.61	51.76
	Klt	35.41	39.01	41.57	43.83	47.88	51.01	53.53	55.97
	Icaorth	35.61	39.20	41.84	44.11	48.16	51.29	53.85	56.38
	Icaopt	35.79	39.27	41.88	44.18	48.27	51.40	54.03	56.62
Raw	Id	39.62	42.81	44.88	46.51	49.23	51.98	54.86	57.94
	Klt	44.04	46.97	48.60	49.82	52.16	54.74	57.88	60.86
	Icaorth	44.15	47.08	48.70	49.91	52.27	54.89	57.99	60.98
	Icaopt	44.17	47.12	48.73	49.92	52.29	54.93	58.03	61.11
Toulouse1	Id	30.55	33.34	35.37	37.15	40.26	43.15	46.01	48.92
	Klt	32.04	35.18	37.48	39.42	42.76	45.67	48.31	51.22
	Icaorth	32.12	35.27	37.57	39.52	42.87	45.78	48.40	51.30
	Icaopt	32.13	35.29	37.61	39.57	42.92	45.81	48.45	51.37
Toulouse2	Id	30.77	33.62	35.67	37.46	40.60	43.48	46.35	49.29
	Klt	32.32	35.55	37.91	39.92	43.37	46.32	48.96	51.87
	Icaorth	32.41	35.66	38.04	40.05	43.49	46.43	49.08	51.99
	Icaopt	32.44	35.71	38.09	40.10	43.55	46.47	49.11	52.02

TAB. 6.10 – PRSB des images de 8 bpp du schéma mixte en sous-bandes pour $P = 4$

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Pgx	Id	29.79	33.04	35.38	37.34	40.84	44.00	47.12	50.19	53.19	56.25
	Klt	32.96	37.02	40.09	42.67	46.75	50.15	53.16	56.05	58.87	61.98
	Icaorth	33.00	37.15	40.32	42.95	47.16	50.55	53.50	56.34	59.20	62.29
	Icaopt	32.96	37.23	40.44	43.10	47.31	50.70	53.66	56.49	59.39	62.49
Moissac	Id	35.66	38.78	41.01	42.90	46.14	49.13	52.07	54.98	58.07	61.04
	Klt	37.94	41.62	44.38	46.78	50.84	54.04	56.69	59.54	62.55	65.72
	Icaorth	38.16	41.90	44.70	47.10	51.12	54.25	56.93	59.78	62.80	66.00
	Icaopt	38.19	41.95	44.76	47.18	51.17	54.30	56.97	59.81	62.84	66.04
Strasbourg	Id	30.42	33.62	36.03	38.08	41.67	44.97	48.05	50.99	53.94	56.85
	Klt	32.23	36.06	39.00	41.56	45.94	49.55	52.60	55.31	58.12	61.17
	Icaorth	32.26	36.13	39.13	41.75	46.21	49.85	52.87	55.56	58.35	61.45
	Icaopt	32.26	36.17	39.19	41.86	46.33	49.99	52.98	55.63	58.46	61.60
Vannes	Id	38.12	41.51	44.01	46.14	49.73	52.77	55.60	58.46	61.42	64.48
	Klt	40.63	44.70	47.57	49.88	53.29	55.82	58.44	61.44	64.60	67.74
	Icaorth	40.93	45.05	47.94	50.25	53.55	56.05	58.73	61.77	64.88	68.07
	Icaopt	41.01	45.17	48.07	50.36	53.63	56.12	58.82	61.89	64.98	68.14
Montp-ref	Id	35.80	38.76	41.03	43.00	46.51	49.68	52.66	55.60	58.51	61.55
	Klt	37.71	41.30	44.08	46.53	50.87	54.55	57.63	60.26	63.07	66.26
	Icaorth	37.75	41.41	44.31	46.87	51.36	54.99	58.03	60.66	63.52	66.64
	Icaopt	37.78	41.49	44.41	46.98	51.44	55.08	58.11	60.73	63.62	66.71
Montp-020	Id	35.80	38.76	41.02	43.00	46.51	49.68	52.65	55.60	58.51	61.55
	Klt	37.59	41.00	43.52	45.72	49.51	52.89	55.95	58.76	61.70	64.69
	Icaorth	37.64	41.12	43.74	46.03	49.95	53.34	56.36	59.20	62.09	65.10
	Icaopt	37.68	41.18	43.82	46.11	50.03	53.41	56.43	59.28	62.14	65.18
Montp-035	Id	35.79	38.76	41.02	42.99	46.51	49.68	52.65	55.60	58.51	61.55
	Klt	37.44	40.66	43.04	45.14	48.82	52.11	55.09	58.01	60.86	64.00
	Icaorth	37.51	40.79	43.26	45.44	49.18	52.49	55.52	58.35	61.25	64.31
	Icaopt	37.53	40.84	43.32	45.52	49.26	52.56	55.59	58.42	61.33	64.39
Montp-050	Id	35.79	38.76	41.02	42.98	46.51	49.68	52.64	55.60	58.51	61.55
	Klt	37.27	40.39	42.72	44.78	48.36	51.57	54.55	57.50	60.37	63.40
	Icaorth	37.34	40.52	42.93	45.05	48.71	51.96	54.94	57.85	60.69	63.78
	Icaopt	37.35	40.55	42.97	45.10	48.76	52.01	54.99	57.89	60.75	63.87
Toul-ref	Id	34.15	36.98	38.96	40.72	43.95	46.98	50.02	52.96	56.02	58.95
	Klt	36.27	39.62	42.12	44.26	47.90	51.04	54.07	56.92	59.95	62.88
	Icaorth	36.34	39.73	42.28	44.41	48.17	51.41	54.40	57.30	60.23	63.20
	Icaopt	36.36	39.77	42.35	44.50	48.27	51.53	54.55	57.46	60.36	63.40
Toul-020	Id	34.15	36.98	38.96	40.72	43.96	46.98	50.01	52.96	56.02	58.95
	Klt	36.20	39.42	41.82	43.84	47.30	50.44	53.38	56.31	59.27	62.32
	Icaorth	36.28	39.57	42.02	44.06	47.63	50.81	53.84	56.70	59.71	62.67
	Icaopt	36.29	39.61	42.07	44.12	47.71	50.89	53.89	56.79	59.78	62.75
Toul-035	Id	34.15	36.98	38.96	40.72	43.96	46.98	50.01	52.96	56.02	58.95
	Klt	36.00	38.98	41.17	43.06	46.44	49.49	52.42	55.40	58.38	61.41
	Icaorth	36.07	39.13	41.38	43.31	46.74	49.87	52.82	55.82	58.74	61.82
	Icaopt	36.08	39.15	41.42	43.36	46.79	49.93	52.91	55.89	58.84	61.89
Toul-050	Id	34.15	36.99	38.96	40.72	43.96	46.98	50.01	52.97	56.02	58.95
	Klt	35.76	38.59	40.71	42.56	45.86	48.86	51.85	54.75	57.83	60.80
	Icaorth	35.83	38.74	40.92	42.79	46.14	49.19	52.20	55.12	58.16	61.11
	Icaopt	35.84	38.75	40.96	42.84	46.18	49.24	52.25	55.20	58.22	61.20

TAB. 6.11 – PRSB des images de 12 bpp du schéma mixte en sous-bandes pour $P = 4$

Images	Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Cuprite_20	Id	60.33	64.45	67.20	69.53	71.68	73.67	77.34	79.05	80.70	83.88
	Klt	76.85	80.85	83.22	85.19	86.72	88.07	90.80	92.19	93.67	96.71
	Icaorth	76.94	80.89	83.26	85.23	86.75	88.10	90.84	92.22	93.71	96.76
	Icaopt	76.96	80.89	83.26	85.23	86.75	88.11	90.85	92.24	93.72	96.78
Moffett_20	Id	57.42	62.56	66.05	68.94	71.51	73.85	78.08	80.05	81.88	85.14
	Klt	76.18	80.96	83.59	85.91	87.64	89.09	91.89	93.30	94.67	97.71
	Icaorth	76.38	81.08	83.68	86.03	87.76	89.20	92.00	93.42	94.81	97.86
	Icaopt	76.42	81.10	83.73	86.07	87.77	89.22	92.02	93.44	94.83	97.87
Jasper_20	Id	55.47	59.27	62.06	64.46	66.59	68.59	72.36	74.15	75.89	79.27
	Klt	74.22	79.94	82.48	84.60	86.38	87.93	90.75	92.14	93.53	96.48
	Icaorth	74.28	79.99	82.51	84.63	86.41	87.96	90.78	92.18	93.56	96.53
	Icaopt	74.29	80.00	82.52	84.64	86.42	87.97	90.79	92.19	93.57	96.54

TAB. 6.12 – PRSB des images de 16 bpp du schéma mixte en sous-bandes pour $P = 4$

Images	Debit	0.25	0.50	0.75	1.00	1.50	2.00	2.50	3.00	3.50	4.00
Pgx	Id	30.37	33.69	36.15	38.26	41.94	45.28	48.44	51.53	54.51	57.41
	Klt	33.47	37.74	40.89	43.45	47.53	50.89	53.82	56.53	59.17	61.80
	Icaorth	33.32	37.46	40.56	43.09	47.18	50.57	53.53	56.24	58.84	61.46
	Icaopt	33.34	37.51	40.62	43.15	47.26	50.65	53.59	56.29	58.90	61.52
Moissac	Id	36.38	39.60	41.94	43.89	47.23	50.17	52.94	55.68	58.44	61.27
	Klt	38.61	42.40	45.25	47.64	51.51	54.49	56.98	59.44	62.02	64.68
	Icaorth	38.66	42.49	45.33	47.70	51.52	54.46	56.93	59.40	62.01	64.68
	Icaopt	38.65	42.53	45.39	47.77	51.57	54.51	57.00	59.48	62.09	64.76
Strasbourg	Id	30.96	34.30	36.83	39.00	42.79	46.19	49.30	52.23	55.03	57.80
	Klt	32.60	36.69	39.86	42.57	47.08	50.65	53.58	56.13	58.65	61.28
	Icaorth	32.55	36.63	39.82	42.55	47.08	50.67	53.60	56.15	58.67	61.30
	Icaopt	32.57	36.66	39.85	42.58	47.12	50.70	53.63	56.18	58.70	61.34
Vannes	Id	39.41	43.05	45.82	48.13	51.91	54.95	57.67	60.32	63.00	65.78
	Klt	41.52	45.85	48.91	51.23	54.47	56.91	59.33	61.90	64.50	67.05
	Icaorth	41.93	46.29	49.28	51.54	54.67	57.08	59.55	62.12	64.71	67.26
	Icaopt	41.98	46.37	49.35	51.59	54.68	57.09	59.55	62.12	64.71	67.25
Montp-ref	Id	36.32	39.38	41.75	43.81	47.41	50.61	53.54	56.34	59.08	61.85
	Klt	38.07	41.64	44.48	46.94	51.23	54.86	57.92	60.51	62.94	65.41
	Icaorth	37.96	41.48	44.25	46.70	50.98	54.66	57.74	60.37	62.81	65.28
	Icaopt	37.91	41.56	44.56	47.18	51.57	55.17	58.17	60.73	63.17	65.63
Montp-020	Id	36.32	39.37	41.74	43.81	47.42	50.61	53.54	56.34	59.09	61.85
	Klt	37.88	41.22	43.78	45.98	49.79	53.13	56.15	58.89	61.48	64.04
	Icaorth	37.89	41.40	44.12	46.43	50.33	53.64	56.58	59.27	61.85	64.40
	Icaopt	37.92	41.42	44.14	46.45	50.33	53.64	56.58	59.27	61.86	64.41
Montp-035	Id	36.31	39.38	41.74	43.81	47.42	50.61	53.54	56.34	59.08	61.86
	Klt	37.65	40.75	43.15	45.22	48.83	52.05	55.02	57.79	60.44	63.06
	Icaorth	37.70	40.99	43.56	45.73	49.45	52.67	55.58	58.29	60.92	63.53
	Icaopt	37.72	41.00	43.55	45.72	49.43	52.65	55.56	58.28	60.90	63.52
Montp-050	Id	36.32	39.38	41.74	43.81	47.42	50.61	53.54	56.34	59.09	61.86
	Klt	37.41	40.38	42.69	44.67	48.19	51.35	54.30	57.06	59.75	62.40
	Icaorth	37.51	40.66	43.11	45.19	48.80	51.95	54.84	57.57	60.22	62.87
	Icaopt	37.51	40.64	43.09	45.17	48.77	51.92	54.81	57.53	60.19	62.83
Toul-ref	Id	34.38	37.18	39.28	41.10	44.37	47.40	50.33	53.21	56.05	58.90
	Klt	36.42	39.74	42.31	44.49	48.18	51.42	54.39	57.16	59.84	62.48
	Icaorth	36.52	39.89	42.51	44.71	48.42	51.66	54.62	57.38	60.05	62.68
	Icaopt	36.54	39.88	42.50	44.71	48.44	51.70	54.67	57.44	60.12	62.76
Toul-020	Id	34.38	37.19	39.28	41.10	44.37	47.41	50.34	53.20	56.05	58.9
	Klt	36.36	39.58	42.05	44.14	47.73	50.89	53.83	56.62	59.35	62.03
	Icaorth	36.44	39.69	42.20	44.29	47.87	51.03	53.97	56.78	59.48	62.16
	Icaopt	36.44	39.69	42.19	44.28	47.88	51.03	53.98	56.77	59.49	62.16
Toul-035	Id	34.38	37.19	39.28	41.10	44.37	47.40	50.33	53.20	56.05	58.90
	Klt	36.13	39.09	41.37	43.33	46.72	49.78	52.70	55.54	58.32	61.04
	Icaorth	36.18	39.18	41.49	43.44	46.84	49.9	52.82	55.64	58.41	61.15
	Icaopt	36.18	39.18	41.48	43.43	46.83	49.89	52.81	55.63	58.41	61.14
Toul-050	Id	34.38	37.19	39.28	41.10	44.37	47.40	50.33	53.20	56.05	58.90
	Klt	35.79	38.54	40.68	42.54	45.83	48.82	51.72	54.55	57.36	60.13
	Icaorth	35.89	38.70	40.88	42.75	46.05	49.06	51.96	54.80	57.60	60.36
	Icaopt	35.89	38.70	40.87	42.74	46.05	49.06	51.96	54.79	57.60	60.36

TAB. 6.13 – PRSB des images de 12 bpp du schéma séparable avec adaptation dans l'espace image + $VM9$

Images	Debit	0.25	0.50	0.75	1.00	1.25	1.50	2.00	2.25	2.50	3.00
Cuprite_224	Id	61.25	64.73	67.38	69.68	71.77	73.71	77.27	78.92	80.49	83.45
	Klt	79.06	81.73	83.82	85.43	86.8	88.04	90.36	91.47	92.55	94.56
	Icaorth	79.24	81.90	83.96	85.54	86.88	88.11	90.42	91.53	92.60	94.60
	Icaopt	79.25	81.92	83.99	85.56	86.90	88.13	90.43	91.54	92.62	94.59
Moffett_224	Id	59.57	64.49	68.10	71.07	73.64	75.92	79.95	81.75	83.42	86.40
	Klt	78.35	81.85	84.27	86.20	87.72	89.03	91.39	92.48	93.50	95.35
	Icaorth	77.65	81.19	83.61	85.60	87.15	88.46	90.81	91.91	92.97	94.90
	Icaopt	78.65	82.00	84.43	86.32	87.78	89.04	91.34	92.42	93.46	95.33
Jasper_224	Id	56.33	59.82	62.56	64.91	67.03	69.00	72.67	74.41	76.10	79.34
	Klt	77.92	81.47	83.60	85.36	86.84	88.16	90.51	91.59	92.62	94.51
	Icaorth	77.97	81.55	83.69	85.45	86.91	88.21	90.54	91.61	92.64	94.53
	Icaopt	77.94	81.52	83.67	85.44	86.90	88.20	90.52	91.59	92.62	94.51

Tab. 6.14 – PRSB des images de 16 bpp du schéma séparable avec adaptation dans l'espace image + VM9

Annexe2

Cette deuxième et dernière annexe illustre d'une part une composante de quelques-unes des images que nous avons utilisées pour notre étude durant cette thèse, d'autre part nous explicitons le calcul de l'entropie d'ordre 1 d'une image tel que nous l'avons fait pour obtenir les résultats des simulations.

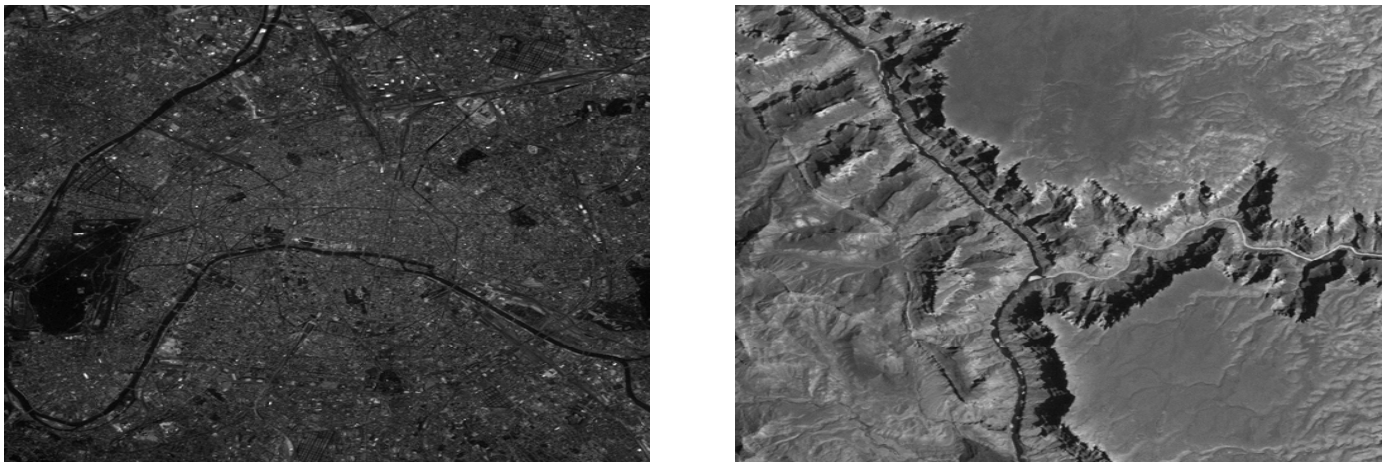


Fig. 6.12 – A gauche Image de Paris, à droite image de Colorado.

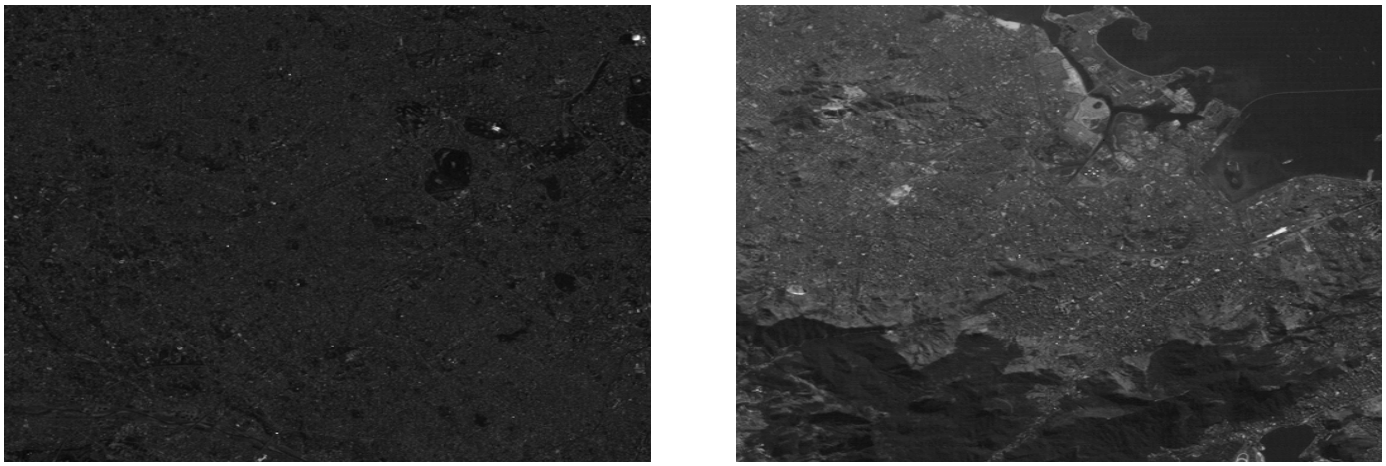


Fig. 6.13 – A gauche Image de Tokyo, à droite image de Rio.

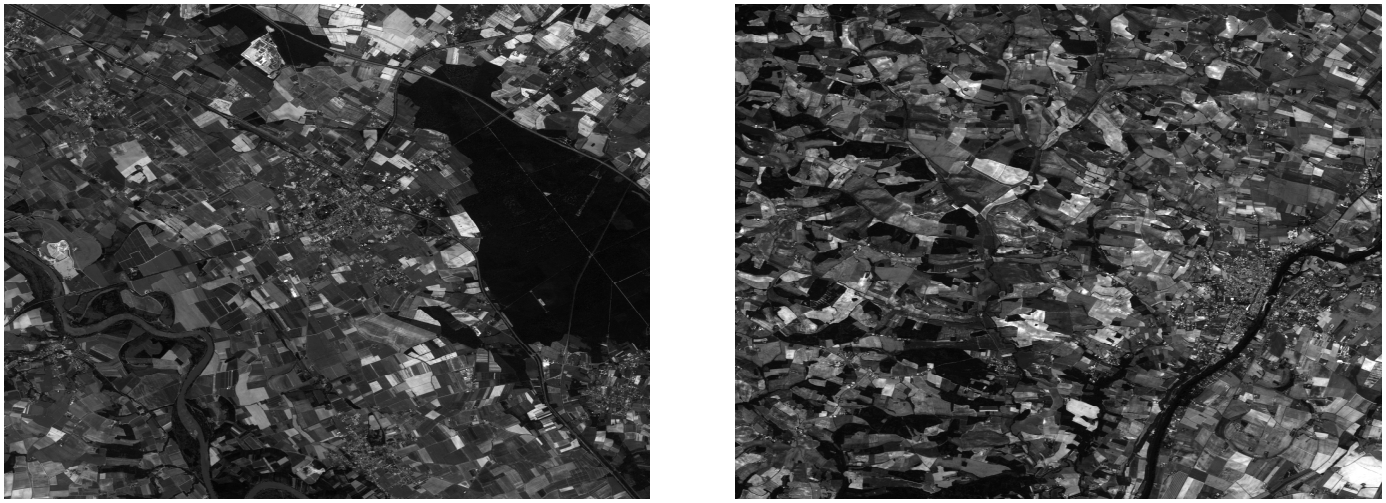


Fig. 6.14 – A gauche Image de Toulouse1, à droite image de Toulouse2.

Fig. 6.15 – A gauche Image Jasper, à droite image de Raw.

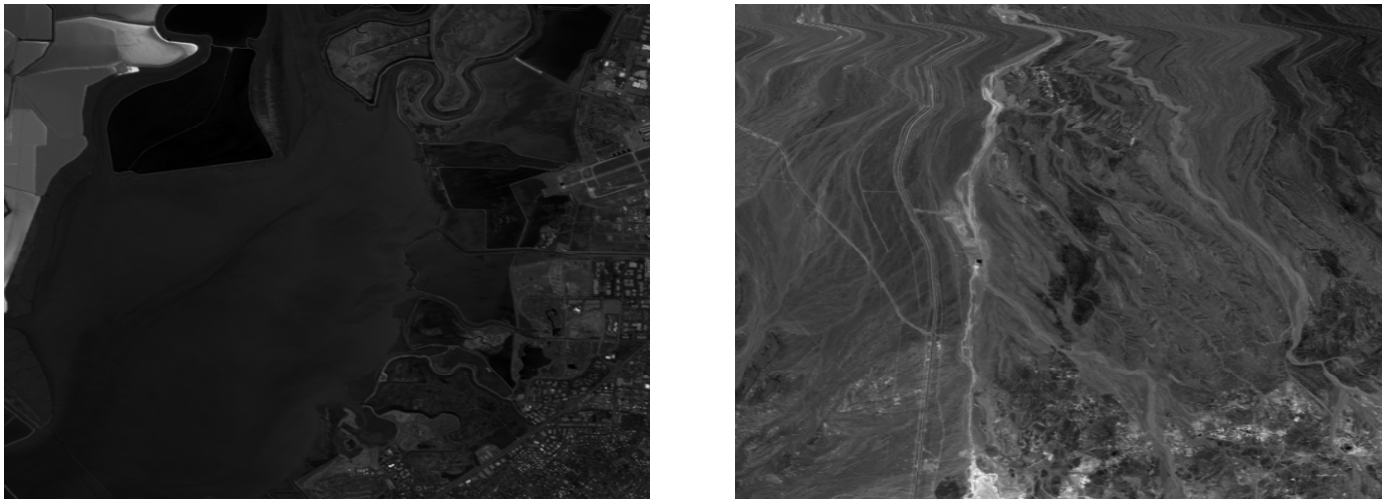


Fig. 6.16 – A gauche Image de Moffett, à droite image de Cuprite.

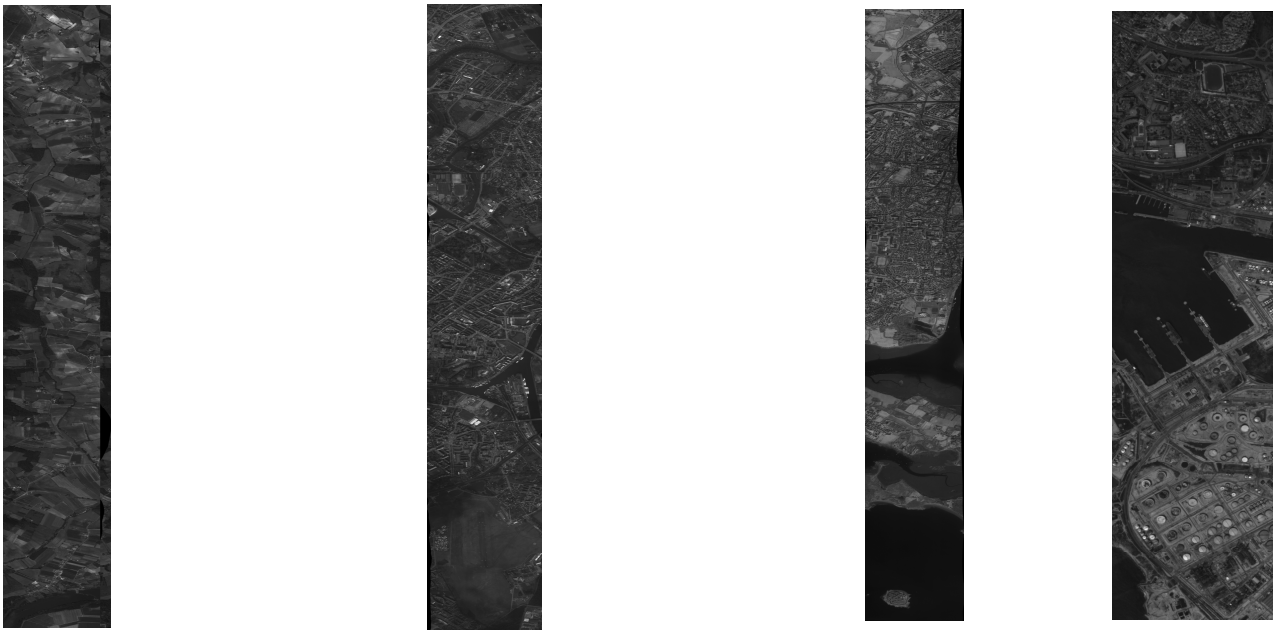


Fig. 6.17 – De gauche à droite, Moissac, Strasbourg, Vannes, Pgx.

Estimation de l'entropie d'ordre 1 d'une image

Soit $\mathbf{X}$ une image numérique constituée d'une seule composante et dont la résolution est $N_L \times N_C$. Soit N_b la profondeur de cette image. Ces pixels sont des valeurs entières appartenant à l'intervalle $[0; 2^{N_b} - 1]$. Notons N_k le nombre de pixels dont la valeur est l'entier naturel k, et $p_k = N_k/L$ (L=$N_L \times N_C$) une estimation de la probabilité discrète d'avoir un pixel dont la valeur est l'entier k. Nous modélisons ainsi l'image comme une variable aléatoire discrète dont les réalisations sont indépendantes et identiquement distribuées. L'estimation de l'entropie d'ordre 1 de cette image est donnée par :

$$r = -\sum_k p_k \log_2 p_k \tag{6.1}$$

où k est un entier naturel tel que $k \in [0; 2^{N_b} - 1]$. Pour nos simulations, nous estimons l'entropie d'ordre 1 de l'image (sous-bandes ou composante entière) après une quantification scalaire uniforme et, il est bien connu dans ce cas que l'image après quantification peut-être représentée entièrement par des entiers relatifs connaissant q le pas de quantification.

Annexe3

Modified ICA algorithms for finding optimal transforms in transform coding

Michel Narozny*, Michel Barret
Équipe Systèmes de Traitement du Signal ,
Supélec, 2, rue É. Belin, 57070, Metz, France
firstname.lastname@supelec.fr
*This work was partially supported by the Lorraine Region.

Dinh-Tuan Pham, Isidore Paul Akam-Bita
Laboratoire de Modélisation et Calcul, IMAG
B.P. 53X, 38041 Grenoble Cedex, France
Dinh-Tuan.Pham@imag.fr
Isidore-Paul.Akambita@imag.fr

Abstract

The Karhunen-Loève transform (KLT) has long been viewed as the best available block transform for a system that orthogonally transforms a Gaussian vector source, scalar quantizes the components of the transformed vector using optimal bit allocation, and then inverse transforms the vector. This paper treats variable-rate orthogonal and non-orthogonal transform codes of non-Gaussian sources. This approach uses a uniform scalar quantizer followed by an optimal entropy code, and each quantized component is encoded separately. In this paper, we provide two new algorithms for finding both the optimal linear transform and the optimal orthogonal transform to be used in a high-rate transform coding system employing entropy-constrained uniform quantization. Both algorithms are based on modified versions of an independent component analysis (ICA) algorithm by Pham. Experiments were carried out on both artificial and real data in order to compare the performances of the transforms returned by our new algorithms with that of the KLT. Results show that, on artificial data, the new optimal transforms outperformed the KLT, while no significant improvements were observed on the real data.

1. Introduction

Common sources such as speech and images have considerable "redundancy" that scalar quantization cannot exploit. Strictly speaking the term "redundancy" refers to the statistical correlation or dependence between the samples of such sources and is usually referred to as *memory* in the information theory literature. It is well known that "removing the redundancy" in the data before scalar quantizing it leads to much improved codes. In transform coding, the redundancy is reduced by using a transform between the data before a scalar quantization. Generally the transform is linear. In this type of transform coding, an input vector X of size N is transformed into another vector S of the same dimension; the components S_i, $i = 1, \ldots, N$, of that vector are then described to the decoder using independent scalar quantizers on the coefficients. High resolution theory shows that the Karhunen-Loève transform (KLT) is optimal for Gaussian sources [1], and the asymptotic low resolution analysis does likewise [2]. Transform coding has been extensively developed for coding images and video (for example, H.261, H.263, JPEG, and MPEG), where the discrete cosine transform (DCT) is most commonly used because of its computational simplicity and its good performance. Special transform codes are subband codes which decompose an image into separate images by using a set of linear filters. The resulting subbands can then be quantized, e.g., by scalar quantizers. The discrete wavelet transform (DWT) is a particular subband code, which is used in the new image compression standard JPEG 2000.

The basic idea that leads to the use of linear preprocessing before coding lies in the de-correlation effect between the pixel values that allows the use of simple source encoders. Yet appealing, this idea is hampered by the fact that linear processing alone may not achieve total independence in the case of non-Gaussian sources. This explains why most of the compression methods (JPEG, JPEG 2000) that perform well use linear pre-processing and some form of context modeling. Through context modeling it is possible to extract the dependencies remaining in the data after linear pre-processing in order to improve the compression performance.

For non Gaussian data, the linear transform that performs the best in high rate transform coding does not decorrelate the data in general, and this result remains valid when the transform is constrained to be orthogonal. Hence, the KLT is not the best linear transform in high rate transform coding for non Gaussian data.In [4], Narozny and Barret show that the optimal linear transform for a high-rate linear transform coding system employing entropy-constrained uniform quantization is the one that minimizes a contrast $\mathcal{C}$ which is equal to the sum of the mutual information between the transform coefficients and another term which may be interpreted as a kind of distance to orthogonality of the transform.

In this paper, we propose an algorithm for the minimization of the contrast $\mathcal{C}$. This algorithm consists of a modified version of the mutual information based ICA algorithm by Pham called `ICAinf` [5]. Furthermore, the algorithm `ICAinf` can also be modified in order to find the orthogo-

nal matrix that minimizes the contrast $\mathcal{C}$.

The remainder of the paper is organized as follows: Section 2 gives a description of the coding structure under consideration and a brief review of the main results presented in [4]. Section 3 gives a presentation of independent component analysis. Section 4 elaborates on the link between transform coding and ICA. The new algorithms are both presented in section 5, and section 6 lists our main experimental results.

2. Generalized coding gain in transform coding

The class of signals to be encoded is represented by a random vector $X = (X_1, \ldots, X_N)^T$ of size N. Let $(\Omega, \mathcal{E}, P)$ be the probability space associated to $X : \Omega \to \mathbb{R}^N$. A transform coder applies an invertible transform $\mathbf{A}$: $\mathbb{R}^N \to \mathbb{R}^N$ to X in order to obtain a random vector S better suited to coding than X. To construct a finite code, each coefficient S_i is approximated by a quantized variable $\widehat{S}_i$. We concentrate on scalar quantizations, which are most often used for transform coding. The decoder then applies a transformation $\mathbf{B}$ to $\widehat{S} = (\widehat{S}_1, \ldots, \widehat{S}_N)^T$ in order to obtain an approximation $\widehat{X}$ of X. In this paper we assume $\mathbf{B} = \mathbf{A}^{-1}$ and the transform $\mathbf{A}$ is linear.

2.1. Entropy-constrained scalar quantization

A scalar quantizer Q approximates a random variable Z by a quantized variable $\widehat{Z}$. It is a mapping from the source alphabet $\mathbb{R}$ to a reproduction codebook $\{\widehat{z}_k\}_{k\in\mathcal{K}} \subset \mathbb{R}$, where $\mathcal{K}$ is an arbitrary countable index set. Quantization can be decomposed into two operations, 1) the lossy encoder $\alpha : \mathbb{R} \to \mathcal{K}$ is specified by a partition of $\mathbb{R}$ into partition cells $S_k = \{z \in \mathbb{R} \,|\, \alpha(z) = k\}$, $k \in \mathcal{K}$, 2) the reproduction decoder $\beta : \mathcal{K} \to \mathbb{R}$ is specified by the codebook $\{\widehat{z}_k\}_{k\in\mathcal{K}}$. We denote $p_k = \Pr\{Z \in S_k\} = \Pr\{\widehat{Z} = \widehat{z}_k\}$. The Shannon theorem [6] proves that the entropy

$$H(\widehat{Z}) = -\sum_k p_k \log_2 p_k \tag{1}$$

is a lower bound of the average number of bits per symbol used to encode the values of $\widehat{Z}$. Arithmetic entropy coding achieves an average bit rate that can be arbitrarily close to the entropy lower bound (see e. g., [12]); therefore, we shall consider that this lower bound is reached. An *entropy constrained scalar quantizer* is designed to minimize $H(\widehat{Z})$ for a fixed mean square distorsion $D = \mathrm{E}[(Z - \widehat{Z})^2]$, where $\mathrm{E}[Z]$ denotes the expectation of Z. Consider $\tilde{Z} = (Z - \mathrm{E}[Z])/\sigma$ the normalized random variable associated to Z, with σ^2 its variance, and let $h(\tilde{Z})$ be the differential entropy of $\tilde{Z}$, $h(\tilde{Z}) = -\int_{-\infty}^{+\infty} p(\tilde{z}) \log_2 p(\tilde{z}) d\tilde{z}$, where $p(\tilde{z})$ denotes the probability density function (pdf) of $\tilde{Z}$. A result from high resolution quantization theory (see e.g., [6]) is that the quantizer performance is described by

$$D \simeq c\,\sigma^2\, 2^{-2R}, \qquad \text{with} \quad c = \frac{2^{2h(\tilde{Z})}}{12}, \tag{2}$$

where $R = H(\widehat{Z})$ is the minimum average bit rate, and the constant c depends only on the pdf shape.

2.2. Generalized coding gain

Coding (quantizing and entropy coding) each transform coefficient S_i separately splits the total number of bits among the transform coefficients in some manner. This bit allocation problem can be stated this way: one is given a set of quantizers described by their high rate-distorsion performances as $D_i \simeq c_i\, \sigma_i^2\, 2^{-2R_i}$, $i = 1, 2, \ldots, N$. The problem is to minimize the end-to-end distorsion

$$D = \frac{1}{N} \sum_{i=1}^{N} \mathrm{E}\left[\left(X_i - \widehat{X}_i\right)^2\right] \tag{3}$$

given a maximum average rate $R = N^{-1} \sum_{i=1}^{N} R_i$.

In [4], the authors show that when the linear transform $\mathbf{A}$ is used, the minimum value of the end-to-end distorsion can be approximated as follows

$$D_{\mathbf{A}}(R) \simeq \left(\prod_{i=1}^{N} w_i c_i\right)^{1/N} \left(\prod_{i=1}^{N} \sigma_i^2\right)^{1/N} 2^{-2R}, \tag{4}$$

where σ_i^2 is the variance of S_i, the constant c_i—defined in the relation (2)—is associated with the normalized variable of S_i, and the weight w_i corresponds to the square euclidean norm of the ith column of $\mathbf{A}^{-1}$.

Let $\mathbf{I}$, $\sigma_i^{\star 2}$ and $c_i^\star$ be respectively the identity transform, the variance of X_i and the constant associated with the normalized random variable of X_i according to the relation (2). The distorsion rate (**??**) was then used to define a figure of merit called the generalized coding gain

$$G^\star = \frac{D_{\mathbf{I}}(R)}{D_{\mathbf{A}}(R)} = \left(\prod_{i=1}^{N} \frac{c_i^\star}{w_i c_i}\right)^{\frac{1}{N}} \left(\prod_{i=1}^{N} \frac{\sigma_i^{\star 2}}{\sigma_i^2}\right)^{\frac{1}{N}}. \tag{5}$$

The generalized coding gain is the factor by which the distorsion is reduced because of the linear transform $\mathbf{A}$, assuming high rate and optimal bit allocation.

2.3. Optimal transform for coding

In [4], the authors show that finding the matrix $\mathbf{A}$ which maximizes $G^\star$ is the same problem as finding the matrix $\mathbf{A}$ which minimizes the contrast

$$\mathcal{C}(\mathbf{A}) = I(S_1; \ldots; S_N) + \frac{1}{2} \log_2 \frac{\det \mathrm{Diag}[\mathbf{A}^{-T}\mathbf{A}^{-1}]}{\det \mathbf{A}^{-T}\mathbf{A}^{-1}}, \tag{6}$$

where the first term of (6) is the mutual information between the random variables $S_1, \ldots, S_N$:

$$I(S_1; \ldots; S_N) = \int_{\mathbb{R}^N} p(s) \log_2 \frac{p(s)}{p(s_1) \cdots p(s_N)}\, ds \tag{7}$$

and for any square matrix $\mathbf{C}$, $\mathrm{Diag}(\mathbf{C})$ denotes the diagonal matrix having the same main diagonal as $\mathbf{C}$. The mutual

information of $S_1, \ldots, S_N$ is a measure of the statistical dependence between the transform coefficients S_i: it is always non-negative, and zero if and only if the variables are statistically independent. As for the second term in (6), it is always non-negative, and zero if and only if $\mathbf{A}^{-1}$ is a transform with orthogonal columns. The columns may not be of unit euclidean norm. In other words, the second term of the contrast $\mathcal{C}(\mathbf{A})$ can be interpreted as a kind of distance to orthogonality for the transform $\mathbf{A}$. Furthermore, if $\mathbf{D}$ is a diagonal matrix, one can verify that $\mathcal{C}(\mathbf{DA}) = \mathcal{C}(\mathbf{A})$, i.e., the contrast is scale invariant. As a consequence, normalizing the rows to have an orthogonal transform and using equal quantizer step sizes is optimal.

3. Independent component analysis

A common problem encountered in a variety of disciplines, including statistical pattern matching, data analysis, signal processing, neural network research and compression, is finding a suitable representation of multivariate data. For computational and conceptual simplicity, such a representation is often sought as a linear transformation of the original data. Well-known linear transformation methods include, for example, principal component analysis (PCA). A recently developed linear transformation method is the independent component analysis, in which the desired representation is the one that minimizes the statistical dependence of the components of the representation. Although non-linear forms of ICA also exist, we shall only consider the linear case here.

3.1. Definition

In the literature, at least three different basic definitions of linear ICA can be found [7, 8]. In [9], Hyvärinen gives the following definition for the noise-free ICA model, which is of primary interest in our study

Definition 1 (Noise-free ICA model) *ICA of a random vector X of size N consists of estimating the following generative model for the data:*

$$X = \mathbf{B}S \tag{8}$$

where $\mathbf{B}$ is a constant $N \times M$ "mixing" matrix, the latent variables (components) S_i in the vector $S = (S_1, \ldots, S_M)^T$ are assumed independent.

The identifiability of the noise-free ICA model has been treated in [7]. By imposing the following fundamental restrictions (in addition to the basic assumption of statistical independence), the identifiability of the model can be assured.

1. All the independent components S_i, with the possible exception of one component, must be non-Gaussian.
2. The number N of observed linear mixtures must be at least as large as the number M of independent components, i.e., $M \leq N$.
3. The matrix $\mathbf{B}$ must be of full column rank.

Under these assumptions, the independent components and the columns of $\mathbf{B}$ can only be estimated up to a multiplicative constant, because any constant multiplying an independent component in (8) could be canceled by dividing the corresponding column of the mixing matrix $\mathbf{B}$ by the same constant. Furthermore, the definition of the noise-free ICA model implies no ordering of the independent components, which is in contrast to, e.g., PCA.

In the following, we make the assumption that the dimension of the observed data equals the number of the independent components, i.e., $N = M$.

3.2. Estimation of the ICA model

The estimation of the data model of ICA is usually performed by formulating an objective function and then minimizing or maximizing it. The mutual information is a natural measure of the dependence between random variables. Finding a transform that minimizes the mutual information between the components S_i is a very natural way of estimating the ICA model [7]. The problem with mutual information is that it is difficult to estimate. One needs a good estimate of the density. This problem has severely restricted the use of mutual information in ICA estimation. Some authors have used approximations of mutual information based on polynomial density expansions [7, 10], which lead to the use of higher-order cumulants. More recently, in [5], Pham has proposed fast algorithms to perform ICA based on the use of mutual information.

4. Link between transform coding and ICA

The criterion (6) may be decomposed into

$$\mathcal{C}(\mathbf{A}) = \mathcal{C}_{\mathrm{ICA}}(\mathbf{A}) + \mathcal{C}_{\mathrm{O}}(\mathbf{A}), \tag{9}$$

where $\mathcal{C}_{\mathrm{ICA}}(\mathbf{A}) = I(S_1; \ldots; S_N)$, and

$$\mathcal{C}_{\mathrm{O}}(\mathbf{A}) = \frac{1}{2}\log_2\left[\frac{\det \mathrm{Diag}(\mathbf{A}^{-T}\mathbf{A}^{-1})}{\det \mathbf{A}^{-T}\mathbf{A}^{-1}}\right]. \tag{10}$$

The first term $\mathcal{C}_{\mathrm{ICA}}(\mathbf{A})$ corresponds to the mutual information criterion in ICA. The second term $\mathcal{C}_{\mathrm{O}}(\mathbf{A})$ measures a pseudo-distance to orthogonality of the transform $\mathbf{A}$. In general, the optimal transform $\mathbf{A}_{\mathrm{opt}}$ in transform coding, i.e., the transform which minimizes the contrast defined in the relation (6), will be different from that $\mathbf{A}_{\mathrm{ICA}}$ which minimizes the first term of (6), i.e., the solution of the ICA problem. Note that the contrast $\mathcal{C}(\mathbf{A})$ is always non-negative, and that it is equal to zero if and only if $\mathbf{A}$ is a transform with orthogonal columns which produces independent components. Therefore, when such a transform exists, it is both the solution of the compression problem and that of the ICA problem. Unfortunately, for most sources, it is very unlikely to find orthogonal transforms that produce independent components.

It is important to notice here that the classical assumption made in blind source separation problems, that is the observations are obtained from a linear mixing of independent sources, is not necessary in order to derive the expression of the contrast in (6), i.e., for finding the transform that maximizes the generalized coding gain.

The expression of the contrast (6) depends on the definition of the distorsion. In this work, we measure the distorsion as mean squared error (MSE). Therefore, it is not surprising that orthogonal transforms are favored over other linear transforms since they are energy-preserving.

5. Modified ICA algorithms for coding

In this section, we propose two algorithms for the minimization of the contrast (6). The first algorithm, called `GCGsup` for *Generalized Coding Gain Supremum*, consists of a modified version of the mutual information based ICA algorithm by Pham [5] called `ICAinf`. The second term of (6) has been incorporated in `ICAinf` in order to find the optimal linear transform $\mathbf{A}_{\text{opt}}$ which minimizes the contrast (6). In the second new algorithm, called `ICAorth` for *Independent Composent Analysis Orthogonal*, the algorithm `ICAinf` has been modified in order to find the optimal orthogonal matrix $\mathbf{A}_{\text{orth}}$ that minimizes the contrast $\mathcal{C}(\mathbf{A})$.

5.1. Algorithm `GCGsup`

The minimization of the criterion (6) can be done through a gradient descent algorithm, but a much faster method is the Newton algorithm (which amounts to using the natural gradient [11]). As in [5], because of the multiplicative structure of our optimization problem, we use multiplicative increment of the parameter $\mathcal{E}$ rather than additive increment. Starting with a current estimator $\widehat{\mathbf{A}}$, it consists of expanding $\mathcal{C}(\widehat{\mathbf{A}}+\mathcal{E}\widehat{\mathbf{A}})$ with respect to the matrix $\mathcal{E}$ up to second order and then minimizing the resulting quadratic form in $\mathcal{E}$ to obtain a new estimate. Note that the parameter $\mathcal{E}$ is a matrix of order N. This method requires the computation of the Hessian[1] of $\mathcal{C}(\widehat{\mathbf{A}}+\mathcal{E}\widehat{\mathbf{A}})$ with respect to $\mathcal{E}$, which is quite involved. For this reason, we will approximate it by the relative Hessian of $\mathcal{C}(\widehat{\mathbf{A}}+\mathcal{E}\widehat{\mathbf{A}})$, computed under the assumption that the transform coefficients $\widehat{S}_i$ are independent. The method is then referred to as quasi-Newton.

Although those simplifications result in a slower convergence speed about the solution, they cause the robustness of the algorithm to be improved by reducing the risk of divergence when the initial estimator $\widehat{\mathbf{A}}_0$ is far from the final solution. Note that the final solution is the same as that obtained without simplification since the algorithm consists of cancelling the first order terms in the expansion of $\mathcal{C}(\mathbf{A}+\mathcal{E}\mathbf{A})$.

Given that (see e. g., [12]) $h(X) = \sum_{i=1}^{N} h(X_i) - I(X_1;\ldots;X_N)$, $h(S) = \sum_{i=1}^{N} h(S_i) - I(S_1;\ldots;S_N)$, $h(S) = h(X) + \log_2|\det\mathbf{A}|$, and the term $h(X)$ does not depend on $\mathbf{A}$, minimizing the contrast (6) is the same as minimizing $\widetilde{\mathcal{C}}_{ICA}(\mathbf{A}) = \mathcal{C}_{\text{O}}(\mathbf{A}) + \widetilde{\mathcal{C}}_{\text{ICA}}(\mathbf{A})$ where

$$\widetilde{\mathcal{C}}_{\text{ICA}}(\mathbf{A}) = \sum_{i=1}^{N} h(S_i) - \log_2|\det\mathbf{A}|. \tag{11}$$

It has been shown in [13] that the Taylor expansion of $\widetilde{\mathcal{C}}_{\text{ICA}}(\mathbf{A}+\mathcal{E}\mathbf{A})$ up to second order may be approximated as follows

$$\widetilde{\mathcal{C}}_{\text{ICA}}(\mathbf{A}+\mathcal{E}\mathbf{A}) = \widetilde{\mathcal{C}}_{\text{ICA}}(\mathbf{A}) + \sum_{1\le i\ne j\le N} \mathrm{E}[\psi_{S_i}(S_i)S_j]\mathcal{E}_{ij} + \\ + \frac{1}{2}\sum_{1\le i\ne j\le N}\{\mathrm{E}[\psi^2_{S_i}(S_i)]\,\mathrm{E}[S_j^2]\mathcal{E}_{ij}^2 + \mathcal{E}_{ij}\mathcal{E}_{ji}\} + \cdots, \tag{12}$$

where the function ψ_{S_i} is equal to the derivative of $-\log_2 p(s_i)$ and is known as the score function, which can be viewed as the gradient of the entropy functional. This approximation lies essentially in the assumption of independent transform coefficients, which may not be valid if the solution of the ICA problem is far from an orthogonal transform, i.e., from the solution that minimizes the contrast (6).

Now let $\mathbf{M} = \mathbf{A}^{-T}\mathbf{A}^{-1}$. One may verify that the Taylor expansion of $\mathcal{C}_{\text{O}}(\mathbf{A}+\mathcal{E}\mathbf{A})$ with respect to $\mathcal{E}$ and around $\mathcal{E} = \mathbf{0}$, up to second order, is given by

$$\mathcal{C}_{\text{O}}(\mathbf{A}+\mathcal{E}\mathbf{A}) = \mathcal{C}_{\text{O}}(\mathbf{A}) - \sum_{1\le i\ne j\le N}\frac{M_{ji}}{M_{ii}}\mathcal{E}_{ji} - \frac{1}{2}\sum_{1\le i\ne j\le N}\mathcal{E}_{ij}\mathcal{E}_{ji} \\ + \sum_{\substack{1\le i,j,k\le N\\ j\ne i \text{ and } k\ne i}}\left[\left(\frac{M_{jk}}{2M_{ii}} - \frac{M_{ij}M_{ik}}{M_{ii}^2}\right)\mathcal{E}_{ji}\mathcal{E}_{ki} + \frac{M_{kj}}{M_{kk}}\mathcal{E}_{ji}\mathcal{E}_{ik}\right] \\ + \sum_{1\le i\ne j\le N}\frac{M_{ji}}{M_{jj}}\mathcal{E}_{ii}\mathcal{E}_{ji} + \cdots \tag{13}$$

The quadratic form associated with the above expansion is quite involved and is not positive. One possible approximation consists in neglecting the non diagonal elements of $\mathbf{M}$, which amounts to assuming that the optimal linear transform is close to an orthogonal transform. Under this hypothesis, one may verify that

$$\mathcal{C}_{\text{O}}(\mathbf{A}+\mathcal{E}\mathbf{A}) \approx \mathcal{C}_{\text{O}}(\mathbf{A}) - \sum_{1\le i\ne j\le N}\frac{M_{ji}}{M_{ii}}\mathcal{E}_{ji} + \\ + \frac{1}{2}\sum_{1\le i\ne j\le N}\left[\frac{M_{jj}}{M_{ii}}\mathcal{E}_{ji}^2 + \mathcal{E}_{ji}\mathcal{E}_{ij}\right] + \cdots \tag{14}$$

Furthermore, the quadratic form associated with the above expansion is positive, but not positive definite. However, this is a sufficient condition for the matrix associated with the quadratic form of the Taylor expansion of $\widetilde{\mathcal{C}}(\mathbf{A})$ to be positive definite, which ensures the stability of the iterative

[1] The Hessian of a function of several variables is the matrix of its second partial derivatives.

algorithm. Finally we have

$$\widetilde{\mathcal{C}}(\mathbf{A}+\mathcal{E}\mathbf{A}) \approx \widetilde{\mathcal{C}}(\mathbf{A}) + \sum_{1\le i\neq j\le N} \mathcal{E}_{ij}\Big[\mathrm{E}[\psi_{S_i}(S_i)S_j] - \frac{M_{ij}}{M_{jj}}\Big] + \\ + \frac{1}{2}\sum_{1\le i\neq j\le N}\Big\{\Big[\mathrm{E}[\psi^2_{S_i}(S_i)]\mathrm{E}[S_j^2] + \frac{M_{ii}}{M_{jj}}\Big]\mathcal{E}_{ij}^2 + 2\mathcal{E}_{ij}\mathcal{E}_{ji}\Big\} \\ + \cdots \tag{15}$$

Explicitly, the iteration consists of solving the linear equations

$$\begin{bmatrix} \mathrm{E}[\psi^2_{S_i}(S_i)]\mathrm{E}[S_j^2] + \frac{M_{ii}}{M_{jj}} & 2 \\ 2 & \mathrm{E}[\psi^2_{S_j}(S_j)]\mathrm{E}[S_i^2] + \frac{M_{jj}}{M_{ii}} \end{bmatrix} \begin{bmatrix} \mathcal{E}_{ij} \\ \mathcal{E}_{ji} \end{bmatrix} \\ = \begin{bmatrix} \frac{M_{ij}}{M_{jj}} - \mathrm{E}[\psi_{S_i}(S_i)S_j] \\ \frac{M_{ji}}{M_{ii}} - \mathrm{E}[\psi_{S_j}(S_j)S_i] \end{bmatrix}. \tag{16}$$

The indeterminate diagonal terms $\mathcal{E}_{ii}$ are arbitrarily fixed to zero. Then the estimator $\widehat{\mathbf{A}}$ is left multiplied by $\mathbf{I}+\mathcal{E}$ in order to update it. In this expression, the probability density functions being unknown, the score function $\psi_{S_i}(s_i)$ is replaced by an estimation (see [5]) and the expectations are estimated by empirical means.

5.2. Algorithm `ICAorth`

In this section, we propose a modified version of the mutual information based ICA algorithm by Pham [5] in order to find the orthogonal transform that minimizes the contrast (6). Finding the orthogonal transform which minimizes the contrast (6) amounts to finding the orthogonal transform which minimizes the first term of (6), or equivalently, finding the orthogonal transform which minimizes $\widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A})$. If the matrix $\mathbf{A}$ is orthogonal, so is $\mathbf{A}+\mathcal{E}\mathbf{A}$, providing that $\mathbf{I}+\mathcal{E}$ be orthogonal. This last condition will be satisfied up to second order if $\mathcal{E}$ is anti-symmetric, since $(\mathbf{I}+\mathcal{E})^T(\mathbf{I}+\mathcal{E}) = \mathbf{I}+\mathcal{E}^T\mathcal{E}$ differs from the identity only by second order terms. Let $\mathcal{E}$ be anti-symmetric. The Taylor expansion of $\widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A}+\mathcal{E}\mathbf{A})$ becomes

$$\widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A}+\mathcal{E}\mathbf{A}) = \widetilde{\mathcal{C}}_{\mathrm{ICA}}(\mathbf{A}) + \\ + \sum_{1\le i<j\le N}\left\{\mathrm{E}[\psi_{S_i}(S_i)S_j] - \mathrm{E}[\psi_{S_j}(S_j)S_i]\right\}\mathcal{E}_{ij} + \\ + \frac{1}{2}\sum_{1\le i<j\le N}\mathcal{E}_{ij}^2\Big[\mathrm{E}[\psi^2_{S_i}(S_i)]\mathrm{E}[S_j^2] + \mathrm{E}[\psi^2_{S_j}(S_j)]\mathrm{E}[S_i^2] - 2\Big] \\ + \cdots, \tag{17}$$

and the minimization of the second term in the above expansion yields

$$\mathcal{E}_{ij} = \frac{\mathrm{E}[\psi_{S_j}(S_j)S_i] - \mathrm{E}[\psi_{S_i}(S_i)S_j]}{\mathrm{E}[\psi^2_{S_i}(S_i)]\,\mathrm{E}[S_j^2] + \mathrm{E}[\psi^2_{S_j}(S_j)]\,\mathrm{E}[S_i^2] - 2}. \tag{18}$$

6. Simulation results

In this section, we are interested in comparing the performances of $\mathbf{A}_{opt}$, $\mathbf{A}_{orth}$, $\mathbf{A}_{ICA}$, and the KLT for the independent coding of N information sources which are both statistically dependent and non-Gaussian. Our objective metric to measure the performance of a transform is the generalized coding gain (5). Given both N sources and a transform, the estimation of (5) requires good estimates of the pdfs of each source as well as each transformed component, which may be very difficult to obtain. In the next section, we shall elaborate on a more "practical" way of evaluating (5) which consists in actually coding the transformed components and measuring both the actual bit-rate (i.e., the first order entropy of quantized data) and the actual end-to-end distortion.

6.1. Evaluation methodology

We carried out our experiments on N information sources. These sources are either 1D signals or multidimensional-like images. In the latter, each image is first converted into a 1D-signal by scanning vertically column by column. The multidimensional signal $X = (X_1,\ldots,X_N)^T$ resulting from this pre-processing is then linearly transformed using an invertible linear transform $\mathbf{A}$ to produce the vector $S = (S_1,\ldots,S_N)^T$ whose components are coded separately. The i-th component S_i is first high-rate quantized with a uniform scalar quantizer of quantization step q_i. This gives $\widehat{S}_i$. The bit-rate R_i is then estimated by computing the empirical first order entropy of $\widehat{S}_i$ and the inverse transform $\mathbf{A}^{-1}$ is applied to $\widehat{S} = (\widehat{S}_1,\ldots,\widehat{S}_N)^T$ in order to reconstruct an approximation $\widehat{X} = (\widehat{X}_1,\ldots,\widehat{X}_N)^T$ of X. The distorsion is the end-to-end one, given in the relation (3) and the total average rate is the empirical mean of the N rates R_i $(1 \le i \le N)$. It is well known that under the high-resolution hypotheses the optimal allocation of rates between the transform coefficients results in equal weighted distortions w_iD_i (see e.g., [4]). Moreover, using uniform scalar quantizers, bit allocation amounts to choosing a quantization step q_i for each of the N components, and for small q_i (i.e., satisfying the hypothesis of high resolution quantization for S_i, [6]) the distortion D_i may be well approximated by $q_i^2/12$. Therefore, the equal-weighted distortion property of the analytical bit allocation solution gives a simple rule (for high bit-rate): let c be a constant, then make all the quantization steps $q_1,\ldots,q_N$ such that $w_iD_i = c$. This gives $q_i = \sqrt{\frac{12c}{w_i}}$, for $i = 1, \ldots, N$. When the constant c varies (under the assumption of high resolution quantization for each component) we obtain the classical asymptotical curve (distorsion versus bit-rate, or equivalently bit-rate versus distorsion). In our tests, we consider that the hypothesis of high resolution quantization is valid when for each component S_i, the relative deviation between the actual distorsion $\mathrm{E}[(S_i-\widehat{S}_i)^2]$ (where the expectation is estimated by empirical mean) and $q_i^2/12$ is not greater than 5%. For a given high bit-rate, the ratio between the end-to-end distortion read on the asymptotical curve obtained using the identity transform and that read on the asymptotical curve obtained using $\mathbf{A}$ yields the generalized coding gain of $\mathbf{A}$.

We have tested two kinds of data, one consists in synthetical memoryless sources and the sources of the second are

wavelet coefficients in the same sub-band (HH sub-band) of a multispectral satellite image Landsat.

In the first test, we have $N = 6$ and the data size is 2^{16}. First we produce a white vector $(\widetilde{S}'_1, \ldots, \widetilde{S}'_N)^T = \widetilde{S}'$ whose the i-th component is the normalized random variable associated to $S'_i = \mathrm{Sign}(Y_i).|Y_i|^\alpha$, where $(Y_1, \ldots, Y_N)^T$ is a normalized white Gaussian random vector. The exponent α is an arbitrary positive real number. When $\alpha > 1$ (resp. $\alpha < 1$), $\widetilde{S}'_i$ is super- (resp. sub-) Gaussian. Then, the vector X is obtained by the operation $X = \mathbf{M}\widetilde{S}'$, where $\mathbf{M}$ is an arbitrary orthogonal matrix. The random vector X being white, the KLT does nothing on it and the generalized coding gain $G^\star$ of the KLT is equal to $0\,\mathrm{dB}$. However, the components of X are not independent, and any algorithm among `GCGsup`, `ICAorth` and `ICAinf` gives the same result: the components S_i are independent, and the generalized coding gain $G^\star$ is the same for $\mathbf{A}_{ICA}$, $\mathbf{A}_{ort}$ and $\mathbf{A}_{opt}$. In Tab. 1 we present $G^\star$ for different values of α. We remark that when α increases the hypothesis of high-resolution quantization is satisfied for an increasing rate (e. g., when $\alpha = 2.5$, the rate must be greater than about 7 bits per coefficient).

α	0.5	1	1.5	2	2.5
$G^\star$ (dB)	1.4	0.0	1.65	3.2	4.8

Table 1. Generalized coding gain of $\mathbf{A}_{ort}$.

The second test deals with real data, obtained from a satellite multispectral image Landsat of dimension $512 \times 512 \times 6$ and coded with 8 bits per pixel. Each component is decomposed in wavelet coefficients using the Daubechies 9-7 DWT. In our test, $N = 6$ and the six information sources we use are the six HH-subbands obtained by this way.

	KLT	$\mathbf{A}_{orth}$	$\mathbf{A}_{opt}$	$\mathbf{A}_{ICA}$
$G^\star$ (dB)	3.05	3.05	3.1	1.55

Table 2. Generalized coding gain when the sources are sub-band signals.

7. Conclusion

In this paper we have presented two new algorithms that compute the linear optimal transform for high-rate transform coding, in the general case of non Gaussian data when the distorsion is measured with the mean square error and the rate measured by the empirical first order entropy of quantized transform coeficients. One algorithm computes the optimal orthogonal transform, and the other the optimal linear transform. These algorithms are both derived from an algorithm by D. T. Pham that minimizes the mutual information of the transformed components. Comparison of the performances in high-rate transform coding between the classical Karhunen-Loève Transform (KLT) and the transforms returned by our algorithms are given. On synthetical data, the transforms given by the new algorithms perform significantly better that the KLT, however on real data all the transforms, included KLT, give roughly the same coding gain.

References

[1] J.-Y. Huang and P. M. Schultheiss, "Block quantization of correlated Gaussian random variables," *IEEE Trans. Commun.*, vol. COM-11, pp. 289-296, Sept. 1963.

[2] V. K. Goyal, J. Zhuang, and M. Vetterli, "Transform coding with backward adaptive updates," *IEEE Trans. Inform. Theory*, vol. 46, pp. 1623-1633, July 2000.

[3] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using wavelet transform", *IEEE Trans. Image Processing*, vol 1, pp. 205–220, 1992.

[4] M. Narozny and M. Barret, "Generalized coding gain in transform coding", submitted to EUSIPCO 2005, Sept. 2005.

[5] D. T. Pham, "Fast algorithms for mutual information based independent component analysis", IEEE Transaction on Signal Processing, Jan. 2004.

[6] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Boston, Kluwer, 1992.

[7] P. Comon, "Independent component analysis – a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.

[8] C. Jutten and J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, pp. 1–10, 1991.

[9] A. Hyvärinen, "Survey on independent component analysis," *Neural Computing Surveys*, vol. 2, pp. 94–128, 1999.

[10] S. Amari, A. Cichocki, and H. H. Yang, "A new learning algorithm for blind source separation," in *Advances in Neural Information Processing 8*, pp. 757–763, MIT Press, Cambridge, MA, 1996.

[11] S. Amari, "Natural gradient works eficiently in learning", *Neural Computation*, vol. 10, no. 2, pp. 251–276, 1998.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York, Wiley & Sons, 1991.

[13] D. T. Pham, "Blind separation of instantaneous mixture of sources via an independent component analysis", *IEEE Trans. Signal Processing*, vol. 44, no. 11, pp. 2768–2779, 1996.