

Post-transformée dans le domaine ondelettes appliquée à la compression d'images satellite

Xavier DELAUNAY¹, Marie CHABERT², Géraldine MORIN², Carole THIEBAUT³

¹TéSA/CNES/NOVELTIS, 14-16 Port Saint Etienne, 31000 Toulouse

²IRIT/ENSEEIH, 2 rue C. Camichel, 31071 Toulouse Cedex 7

³CNES, 18 avenue Edouard Belin, 31401 Toulouse Cedex 9

xavier.delaunay@tesa.prd.fr

Résumé – Cet article s'intéresse à la compression d'images satellite. La technique proposée met en œuvre la transformée en ondelettes couplée à une post-transformée par blocs sélectionnée dans un dictionnaire de bases. L'originalité du travail réside dans le dictionnaire utilisé. Ce dictionnaire est constitué de bases construites à partir de l'Analyse en Composantes Principales (ACP) sur une base d'apprentissage de blocs de coefficients d'ondelettes. La post-transformée proposée permet d'éliminer les corrélations entre coefficients d'ondelettes voisins identifiés dans [1]. Elle bénéficie aussi de la propriété de concentration d'énergie de l'ACP qui est exploitée au moment du codage.

Abstract – This article is about satellite image compression. A post-transform by blocks following the wavelet transform is proposed. On each block, the best basis of a dictionary is chosen to express the wavelet coefficients. The dictionary of bases is built by Principle Component Analysis (PCA) on training sets of wavelet coefficients blocks. Remaining correlations between wavelet coefficients identified in [1] are removed thanks to the post-transform. This post-transform also benefits from the energy compaction property due to the PCA. A proposed coding scheme exploit this property.

1 Introduction

En compression d'image, des transformées linéaires sont couramment employées afin d'éliminer les corrélations entre pixels. Un quantificateur scalaire suivi d'un codeur entropique permet ensuite d'obtenir de bonnes performances en compression avec une complexité plus faible que celle d'un quantificateur vectoriel. Les systèmes de compression les plus récents, tels que JPEG2000 ou la recommandation du CCSDS (The Consultative Committee for Space Data Systems) pour le spatial [2], utilisent une transformée en ondelettes séparables. Les ondelettes possèdent les propriétés de multi-résolution et de localisation spatiale et fréquentielle conduisant à une structuration adéquate de l'information et à une concentration de l'énergie sur quelques coefficients. Cependant, bien qu'elles soient adaptées à la description de zones homogènes et des contours de l'image dans des directions particulières, elles ne représentent pas tous les contours de façon parcimonieuse. Ceci se traduit par des agglomérats de coefficients d'ondelettes relativement forts au niveau des contours de l'image. Nous avons également observé des corrélations entre coefficients d'ondelettes voisins dans [1].

Pour exploiter ces corrélations, plusieurs solutions ont déjà été proposées soit au niveau du codage, soit au niveau de la transformée elle-même. Par exemple, EBCOT [3], le codeur de JPEG2000, prend en compte ces corrélations en introduisant la notion de contextes. De nouvelles transformées géométriques [4, 5, 6, 7] ont également été proposées. Les éléments de base de ces transformées sont directionnels et anisotropes et donc mieux adaptés à la

représentation des contours. Dans cet article, une autre approche est explorée : une post-transformée linéaire par blocs est utilisée à la suite de la transformée en ondelettes. Cette méthode s'inspire de la transformée en bandelettes par groupements proposée dans [7], mais contrairement à cette dernière, les éléments de base ne sont pas nécessairement géométriques. Ils sont construits par ACP sur des blocs de coefficients d'ondelettes comme expliqué dans la section 2. Le choix de la meilleure base utilisée sur chaque bloc se fait par minimisation d'un critère débit-distorsion détaillé dans la section 3. Enfin, dans la section 4, plusieurs résultats en compression sont présentés.

2 Construction du dictionnaire de bases

Dans cet article, les post-transformées utilisées ne sont pas directionnelles comme dans les bandelettes, mais adaptées aux caractéristiques des ondelettes. Ce sont les transformées dans les bases obtenues par ACP sur les blocs de coefficients d'ondelettes. Ainsi, la propriété de concentration d'énergie de l'ACP pourra être exploitée au moment du codage.

Le dictionnaire de bases contient N_B bases d'apprentissage constituées de blocs de M coefficients d'ondelettes. Une ACP est effectuée sur chacune des bases d'apprentissage. Chaque ACP permet de constituer une base orthonormée $\mathcal{B}_b = \{\phi_{b,m}\}_{m=0}^{M-1}$ de l'espace à M dimensions où les $\phi_{b,m}$ sont les vecteurs de base. Les post-transformées

linéaires consistent en la projection des blocs f de coefficients d'ondelettes sur les vecteurs $\phi_{b,m}$ des bases \mathcal{B}_b .

$$f_b = \sum_{m=0}^{M-1} \langle f, \phi_{b,m} \rangle \phi_{b,m}$$

L'intérêt de construire un dictionnaire de bases $\mathcal{D} = \{\mathcal{B}_b\}_{b=1}^{N_B}$ est de mettre en compétition plusieurs représentations f_b du même bloc de coefficients d'ondelettes f . La minimisation d'un critère débit-distorsion permet ensuite la sélection de la représentation la plus appropriée pour la compression comme expliqué dans la section 3. De plus, les post-transformées proposées bénéficient des propriétés de concentration d'énergie de l'ACP qui pourront être exploitées par un codeur approprié.

Plusieurs tests ont été effectués sur la constitution des bases d'apprentissage. Par exemple, pour l'un d'eux, les bases d'apprentissage ont été constituées de blocs de coefficients d'ondelettes sélectionnés aléatoirement dans n'importe quelle sous-bande des transformées en ondelettes de plusieurs images d'observation de la Terre. Pour les résultats reportés dans la section 4, chaque base d'apprentissage était constituée de blocs de coefficients issus d'une unique sous-bande (HL₁ ou LH₁ ou HH₁ ...) de la transformée en ondelettes de ces images. Les résultats obtenus par ces deux tests sont très proches.

3 Sélection de la post-transformée

3.1 Principe

Le principe est de tester les post-transformées d'un dictionnaire pour décorrélérer les blocs de coefficients d'ondelettes. Le dictionnaire est connu par le codeur et par le décodeur. La sélection de la meilleure transformée s'effectue par minimisation d'un critère débit-distorsion indépendamment pour chaque bloc de coefficients d'ondelettes.

Le critère utilisé est emprunté à la théorie des bandelettes [5, 6]. Pour chaque bloc, la transformée retenue est celle pour laquelle la distorsion après quantification est minimale en fixant une contrainte sur le débit. Ce problème se ramène donc à la minimisation du Lagrangien :

$$\mathcal{L} = D + \lambda R \quad (1)$$

où D est une mesure de distorsion après quantification des coefficients du bloc et R est une estimation du débit nécessaire au codage du bloc de coefficients.

3.2 Calcul du Lagrangien

La mesure de distorsion utilisée est l'erreur quadratique entre les coefficients f_b et les coefficients f_{bR} reconstruits après quantification.

$$D = \|f_b - f_{bR}\|^2$$

Les coefficients f_b sont obtenus en appliquant la transformée numérotée b au bloc de coefficients d'ondelettes f .

Le débit R se décompose en deux parties :

$$R = R_C + R_b$$

Le nombre R_C représente le nombre de bits nécessaire au codage des coefficients transformés et quantifiés. La quantification est effectuée par un quantificateur Q_Δ scalaire, uniforme à pas Δ et à zone morte. Le débit R_C est estimé par l'entropie des coefficients quantifiés $Q_\Delta f_b$.

$$R_C = \sum_{m=0}^{M-1} \log_2 \frac{1}{p(Q_\Delta f_b(m))}$$

où M est le nombre de coefficients par bloc. Les probabilités $p(Q_\Delta f_b(m))$ sont calculées à partir de l'histogramme des coefficients d'ondelettes dans chaque sous-bande.

Le nombre R_b est le nombre de bits nécessaire au codage de l'identifiant b de la transformée utilisée sur ce bloc. L'identifiant $b = 0$ correspond à la transformée identité. Cette transformée, qui correspond en réalité à une absence de transformée, est privilégiée lors du calcul du Lagrangien en diminuant le coût R_b qui lui est associé. Ainsi, on fixe la probabilité *a priori* p_b que la transformée b soit utilisée sur ce bloc :

$$p_b = \begin{cases} 0,6 & \text{si } b = 0 \\ 0,4/N_B & \text{si } b \in \{1, \dots, N_B\} \end{cases} \quad (2)$$

où N_B est le nombre de transformées du dictionnaire autre que la transformée identité. Le coût de codage de l'identifiant b est alors estimé par :

$$R_b = -\log p_b \quad (3)$$

Comme dans [5], le paramètre λ est fixé à :

$$\lambda = \frac{3\Delta^2}{4\gamma_0} \quad \text{avec } \gamma_0 = 7.$$

4 Résultats de compression

Les filtres 9/7 CDF (Cohen-Daubechies-Feauveau) sont utilisés pour la transformée en ondelettes. Dans l'optique d'une application aux images satellite, trois niveaux de décomposition sont utilisés selon la recommandation du CCSDS. Les études sont effectuées sur des images d'observation de la Terre simulées PLEIADES (résolution 70 cm) et PELICAN (capteur aéroporté de résolution 20 cm) de taille 1024×1024 codées sur 12 bits.

Le dictionnaire a été créé à partir d'ACP sur des blocs issus des sous-bandes de la transformée en ondelettes de sept de ces images. Il comprend une base par sous-bande de la décomposition en ondelettes. Il y a donc $N_B = 9$ bases dans ce dictionnaire. De plus, la post-transformée est appliquée à des blocs de coefficients d'ondelettes de taille 4×4 . L'espace de cette transformée est donc de dimension $M = 16$.

Après transformation et quantification, les coefficients $Q_\Delta f_b$ et les identifiants b des post-transformées utilisées sur chaque bloc sont codés par un codeur arithmétique. Le train binaire final est donc formé de deux parties : un en-tête correspondant aux identifiants b et le corps correspondant aux coefficients transformés $Q_\Delta f_b$.

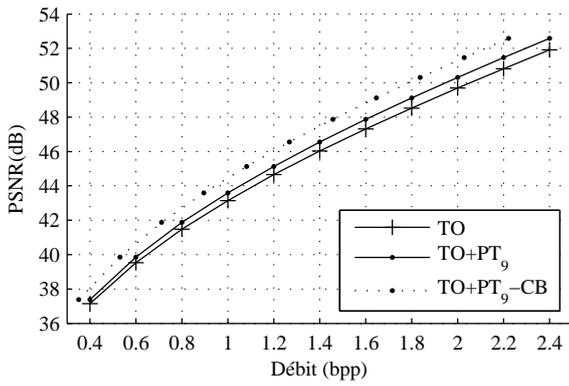


FIG. 1 – Performances en compression de la transformée en ondelettes (TO) suivie ou non de la post-transformée (PT₉) avec et sans le codage du choix des bases (CB).

4.1 Compression en utilisant le dictionnaire complet

Dans cette section et contrairement à la section 4.2, les N_B bases du dictionnaire créé peuvent être utilisées pour la post-transformation des blocs de coefficients d'ondelettes issus de n'importe quelle sous-bande. Pour chaque bloc, la minimisation du Lagrangien (1) permet la sélection de la meilleure transformée parmi les N_B proposées.

La figure 1 illustre les performances moyennes de cette post-transformée appliquée à six autres images d'observation de la Terre en comparaison avec les performances de la transformée en ondelettes seule. Le même codeur arithmétique est utilisé dans les deux cas. Un débit inférieur à 2 bpp est visé en compression bord. A ce débit, l'utilisation de la post-transformée permet d'obtenir un gain de 0,6 dB en PSNR. La courbe tracée en pointillés présente les performances obtenues sans compter l'en-tête dans le débit final. A 2 bpp, l'en-tête représente environ 0,16 bpp soit en moyenne 2,56 bits par bloc 4×4 . Cette valeur est légèrement supérieure à l'entropie R_b moyenne *a priori* calculée à partir des équations (2) et (3) qui est de 2,24 bits pour un dictionnaire de $N_B = 9$ bases. En d'autres termes, à 2 bpp, le codage de l'en-tête réduit les performances de 1,0 dB (cf. figure 2). Sans codage de l'en-tête, le gain en PSNR est de 1,6 dB par rapport à la transformée en ondelettes. Ceci traduit l'efficacité de la post-transformée à éliminer les corrélations dans la transformée en ondelettes.

Sur la figure 2, une comparaison des résultats est effectuée avec la transformée en bandelettes par groupements proposée dans [6]. Les courbes sont tracées relativement à la courbe débit-distorsion de la transformée en ondelettes de la figure 1. La transformée en bandelettes utilise un dictionnaire constitué de 15 bases. Bien que le nombre de bases utilisée par la post-transformée proposée soit plus faible, les résultats approchent ceux des bandelettes.

4.2 Compression dans la base adaptée

4.2.1 Concentration d'énergie

Les vecteurs dans les bases d'ACP sont ordonnés par contribution décroissante d'énergie. Les coefficients transformés sont donc, en moyenne, rangés selon leur ampli-

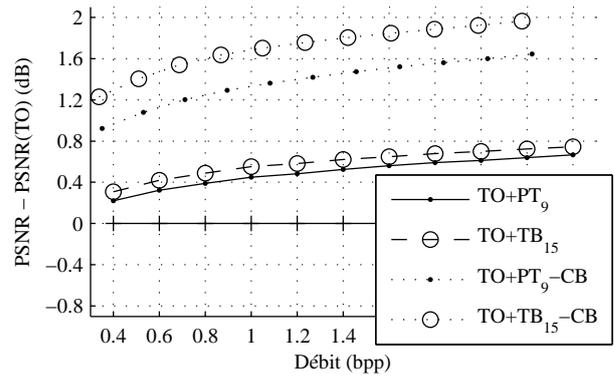


FIG. 2 – Performances de la post-transformée (PT₉) et de la transformée en bandelettes (TB₁₅) avec et sans le codage du choix des bases (CB) relativement à la transformée en ondelettes.

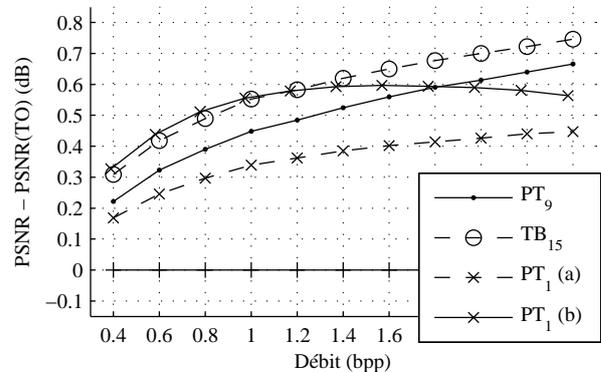


FIG. 3 – Performances de la post-transformée avec le dictionnaire de 9 bases (PT₉), de la transformée en bandelettes (TB₁₅), de la post-transformée avec la base adaptée (PT₁ (a)) et de la post-transformée avec la base adaptée suivie d'un codage exploitant la compaction d'énergie (PT₁ (b)) relativement à la transformée en ondelettes.

tude. Lorsque le dictionnaire complet de 9 bases est utilisé, cette propriété n'est pas garantie de manière systématique puisque la base sélectionnée par le critère débit-distorsion (1) n'est pas toujours la base ACP adaptée à la sous-bande traitée. Ceci diminue donc le potentiel de compaction d'énergie.

Dans cette section, le dictionnaire est réduit à une seule base : la base ACP adaptée à la sous-bande traitée. Comme ce dictionnaire ne permet de mettre en compétition que deux représentations de chaque bloc de coefficients d'ondelettes (identifiants $b = 0$ ou $b = 1$), les résultats en compression obtenus sont moins bons que précédemment (courbe PT₁ (a) de la figure 3). Cependant, comme seule la base adaptée est utilisée dans chaque sous-bande, il est possible d'exploiter efficacement la concentration d'énergie obtenue par cette post-transformée en utilisant simplement la propriété que les vecteurs sont ordonnés par contribution d'énergie décroissante dans la base ACP.

Pour cela, les coefficients des blocs non transformés sont codés par le codeur arithmétique indépendamment des coefficients des blocs transformés. De plus, dans les blocs

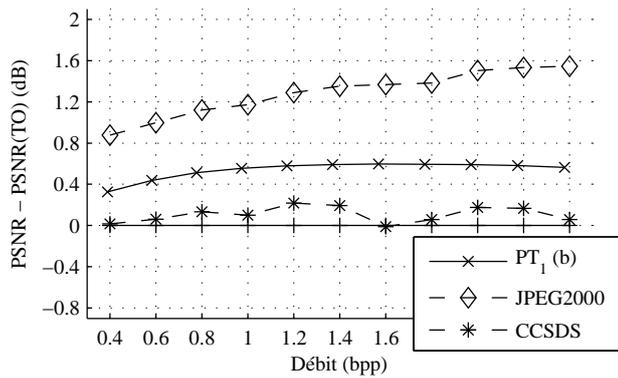


FIG. 4 – Performances de la post-transformée, de JPEG2000 et d'un codeur suivant la recommandation du CCSDS relativement à la transformée en ondelettes.

transformés dans la base ACP, un codage arithmétique est réalisé indépendamment pour chacune des positions des M coefficients du blocs. Ces positions correspondent à la classification des différents niveaux d'énergie effectuée par l'ACP. Ainsi, lors du codage, une séparation des niveaux d'énergie est effectuée. Le but recherché est le même lorsque les sous-bandes d'une transformée en ondelettes sont codées indépendamment les unes des autres.

Cette stratégie de codage permet de diminuer le train binaire de 1% à 3%. Les résultats obtenus en compression sont tracés sur la figure 3 (courbe PT_1 (b)). A 2 bpp, le débit est diminué de 1,42% ce qui correspond à une augmentation du PSNR de 0,16 dB. A 1,8 bpp les performances sont les mêmes que pour l'utilisation du dictionnaire complet de 9 bases. A bas débit, les performances sont comparables à celles des bandelettes qui utilisent un dictionnaire de 15 bases. L'exploitation de la compaction d'énergie obtenue grâce à la transformée dans des bases ACP permet de réduire le nombre de bases du dictionnaire, et donc la complexité calculatoire, tout en gardant de bonnes performances en compression.

4.2.2 Comparaison aux standards

Sur la figure 4, les résultats du schéma proposé sont comparés aux résultats donnés par JPEG2000 [8] et par un codeur répondant aux recommandations du CCSDS [9]. A 2 bpp, par rapport à la transformée en ondelettes, le gain introduit par la post-transformée est de 0,6 dB, celui de JPEG2000 est de 1,5 dB et celui du codeur du CCSDS est de 0,1 dB. Comme dans JPEG2000, un codeur par plans de bits est utilisé dans la recommandation du CCSDS. Cependant, dans JPEG2000, c'est un codeur arithmétique couplé à une optimisation débit-distorsion qui est utilisé. Le coût calculatoire est bien plus élevé que celui du codage entropique par tables de la recommandation du CCSDS [2]. JPEG2000 ne peut donc pas être implanté en matériel pour la compression à bord avec les technologies actuelles. La post-transformée dans la base adaptée ne nécessite que $\mathcal{O}(M)$ opérations par pixel alors que les autres transformées par blocs telles que la transformée dans le dictionnaire complet et la transformée en bandelettes nécessitent $\mathcal{O}(N_B M)$ opérations par pixel. Aussi,

un codeur par plans de bits permettant d'exploiter les propriétés de la post-transformée proposée tout en prenant en compte les contraintes d'un traitement embarqué peut être envisagé.

5 Conclusion

Les bandelettes ont déjà montré que l'utilisation d'une seconde transformée adaptée permet d'exploiter les corrélations entre coefficients d'ondelettes voisins. La construction de post-transformées par ACP est prometteuse car les résultats obtenus sont très proches des bandelettes. De plus, il est possible d'exploiter la propriété de concentration de l'énergie qu'offrent les ACP au moment du codage. Cependant, il reste beaucoup de paramètres à régler dans cet algorithme de codage, en particulier au niveau de la sélection de la meilleure base. Dans l'optique d'un codeur par plans de bits, ce processus de sélection pourrait être différent. Ce sont des pistes actuellement explorées. Néanmoins, pour une faible complexité supplémentaire, la post-transformée présentée permet d'améliorer sensiblement les résultats en compression par rapport à la transformée en ondelettes.

Remerciements :

Ce travail a été effectué avec le support de NOVELTIS (www.noveltis.fr) et du CNES (www.cnes.fr), ainsi que grâce aux composantes logicielles fournies par Gabriel Peyré.

Références

- [1] X. Delaunay, C. Thiebaut, V. Charvillat. Compression embarquée d'images satellites : Vers l'exploitation de la géométrie. *CORESAs*, Nov. 2006, Cean.
- [2] *Image data compression. Recommended standard.* CCSDS 122.0-B-1, Blue Book, Nov. 2005.
- [3] D.S. Taubman and M. Marcellin. *JPEG2000 - Image compression fundamentals, standards and practice.* Springer, 2002.
- [4] E.J. Candès, D.L. Donoho. Curvelets - A surprisingly effective nonadaptive representation for objects with edges. *Curves and Surfaces*, L.L. Schumaker et al. (eds), Vanderbilt University Press, Nashville, TN, 1999.
- [5] E. Le Pennec, S. Mallat. Sparse geometric image representations with bandelets, *IEEE Trans. on Image Processing*, vol. 14, no. 4, pp. 423-438, Apr. 2005.
- [6] G. Peyré, S. Mallat. Surface compression with geometric bandelets. *ACM Trans. on Graphics (Proc. of SIGGRAPH'05)*, vol. 24, no.3, pp. 601-608, Jul. 2005.
- [7] G. Peyré. *Géométrie multi-échelles pour les images et les textures.* Thèse de Doctorat, Ecole Polytechnique, Déc. 2005.
- [8] *Kakadu software*, Dec. 2006 (V5.2.4), D.S. Taubman, <http://www.kakadusoftware.com>.
- [9] *CCSDS Image Data Compression Simulation Software*, Apr. 2006 (V2.0.3), NASA's Goddard Space Flight Center.