

# Transformations optimales à haut débit pour la compression d'images multi-composantes selon la norme JPEG2000

Isidore Paul AKAM BITA<sup>1</sup>, Michel BARRET<sup>2</sup> et Dinh-Tuan PHAM<sup>1</sup>

<sup>1</sup>Laboratoire Jean KUNTZMANN

51 rue des mathématiques, BP 53, 38041 Grenoble Cedex 9, France

<sup>2</sup>SUPELEC, Campus de Metz, 2 rue Edouard Belin, 57070 Metz, France

Isidore-Paul.Akambita@imag.fr, Michel.Barret@supelec.fr, Dinh-Tuan.Pham@imag.fr

Cette étude a été financée par l'ANR, via l'ACI masse de données ACI2M.

**Résumé** – Dans cet article nous proposons deux transformations linéaires optimales à faible distorsion pour la réduction de la redondance spectrale en utilisant le codeur EBCOT de JPEG2000, l'une étant orthogonale et l'autre quelconque. Le schéma de compression étudié propose d'appliquer une transformation en ondelette discrète par composante pour réduire la redondance spatiale, après avoir réduit la redondance spectrale en utilisant une transformation linéaire. En appliquant ces transformations et en utilisant le codeur EBCOT de JPEG2000, nous constatons une amélioration des performances en compression par rapport à la transformée de Karhunen Loeve.

**Abstract** – In this paper we propose two optimal linear transforms to reduce the spectral redundancy using the coder EBCOT of JPEG2000 with high rate assumptions. One of these transforms is orthogonal and the other is not. We are interested with a multicomponent compression scheme that proposes to apply a linear transform to reduce the spectral redundancy, then a discrete wavelet transform is applied in each component to reduce the spatial redundancy. Using these transforms with EBCOT for image compression, we show that they perform better than the well known Karhunen Loeve Transform.

## 1 Introduction

Pour la compression des images multi-composantes (multi ou hyper-spectrale), il est nécessaire de réduire ou d'exploiter au mieux les redondances spatiales et spectrales qui existent à l'intérieur des différentes composantes et entre elles. Dans la norme JPEG2000, il est recommandé d'appliquer deux transformations différentes pour réduire ces redondances avant le codage, l'une pour réduire la redondance spatiale et une transformation linéaire pour réduire la redondance spectrale.

La TOD (Transformée en Ondelette Discrète) permet de réduire d'une manière efficace la redondance spatiale dans une composante. Des études préliminaires [2] [7] ont montré que la 9/7 (ondelette recommandée pour la norme JPEG2000) de Daubechies permettait en général d'avoir de bonnes performances en compression aussi bien aux bas débits qu'aux moyens et hauts débits. D'autres travaux [4] [8] proposent d'utiliser comme pour JPEG2000, une transformation linéaire pour réduire la redondance spectrale et la TOD pour la redondance spatiale. Des travaux [3] ont montré que la TOD peut également être utilisée pour réduire la redondance spectrale des images hyper-spectrales. La Transformation de Karhunen-Loève (TKL) est souvent recommandée pour réduire la redondance spectrale entre composantes, même s'il est bien connu que pour des données non gaussiennes (comme les images), elle n'est pas optimale en codage par transformée. Dans [5] [1] des critères caractérisant les transformations optimales en

compression ont été explicitées pour différents schémas de compression, sous l'hypothèse d'une quantification à haute résolution, avec une distorsion mesurée par l'erreur quadratique moyenne et un codage entropique, et des algorithmes donnant les transformations spectrales optimales y sont présentés.

L'objet de cet article est de rechercher la transformation linéaire optimale à faible distorsion en utilisant le schéma de codage de JPEG2000 qui consiste à appliquer une transformation linéaire pour la réduction de la redondance spectrale avant la TOD dans chaque composante. Dans la deuxième section, nous définissons le schéma de compression qui nous intéresse et nous calculons une formule d'approximation de la distorsion. Puis nous définissons le critère à minimiser pour obtenir la transformation linéaire optimale. La troisième section détaille deux algorithmes quasi-Newton pour la minimisation du critère (l'un sans contrainte sur la transformation, l'autre en imposant l'orthogonalité de la transformation). La dernière partie illustre et discute de quelques résultats obtenus en appliquant ces transformations optimales pour la compression des images multi et hyper-spectrales et en les comparant à ceux de la TKL.

## 2 Définition du critère

Dans la suite, une image multi-spectrale constituée de  $N$  composantes ayant chacune  $N_L$  lignes et  $N_C$  colonnes est représentée par une matrice  $\mathbf{X}$  de dimension  $N \times L$ ,

où  $L = N_L \times N_C$ . Ainsi, les pixels d'une composante sont rangés dans une ligne de  $\mathbf{X}$ , selon un balayage pré-défini, ligne après ligne par exemple. Notons  $\mathbf{W}$  la matrice de la TOD appliquée à chaque composante,  $\mathbf{X}\mathbf{W}^T$  (où  $\mathbf{W}^T$  désigne la transposée de  $\mathbf{W}$ ) est l'image transformée dans le domaine ondelette. Soit  $\mathbf{A}$  une matrice inversible de taille  $N \times N$  représentant la transformation spectrale, l'image transformée à l'aide de la TOD et de cette transformation s'écrit  $\mathbf{Y} = \mathbf{A}\mathbf{X}\mathbf{W}^T$ .

Dans le schéma de codage que nous étudions, les coefficients transformés  $\mathbf{Y}$  sont quantifiés avec un quantificateur scalaire uniforme par composante et par sous-bande, puis codés avec un codeur entropique par composante et par sous-bande également. Les coefficients de la TOD sont fixes (ce sont ceux de la 9/7 de Daubechies) et la transformation spectrale  $\mathbf{A}$  s'adapte aux données. Soit  $\mathbf{Y}^q$  la matrice des coefficients transformés et quantifiés. Une approximation  $\hat{\mathbf{X}}$  de l'image originale est obtenue en appliquant les transformations inverses :  $\hat{\mathbf{X}} = \mathbf{A}^{-1}\mathbf{Y}^q\mathbf{W}^{-T}$ . La distorsion  $D$  est mesurée par l'erreur quadratique moyenne et le débit  $R$  (exprimé en bits par pixel—bpp) est mesuré par :  $R = \frac{1}{N} \sum_{i=1}^N \sum_{m=1}^M \pi_m H(Y_i^{(m)q})$ , où  $M$  est le nombre de sous-bandes,  $\pi_m$  désigne la proportion des coefficients de la  $m^{\text{ème}}$  sous-bande par rapport à une composante et  $H(Y_i^{(m)q})$  est l'entropie d'ordre 1 des coefficients quantifiés de la  $m^{\text{ème}}$  sous-bande et  $i^{\text{ème}}$  composante. Sous l'hypothèse

$H_1$  : les composantes du bruit de quantification  $\mathbf{Y} - \mathbf{Y}^q$  sont décorréliées et centrées,

on montre [1] que la distorsion distorsion peut s'exprimer en fonction des distorsions  $D_i^{(m)}$  de la sous-bande  $m$  de la  $i^{\text{ème}}$  composante comme

$$D = \frac{1}{N} \sum_{m=1}^M \sum_{i=1}^N \pi_m w_m w_i D_i^{(m)}, \quad (1)$$

où  $w_i$  est le carré de la norme 2 de la  $i^{\text{ème}}$  colonne de  $\mathbf{A}^{-1}$  et  $w_m$  représentent les coefficients de pondération de chaque sous-bande ( $m$ ) (qui dépendent uniquement du filtre de synthèse de la TOD). Dans [10] [?], il est proposé une méthode simple permettant sous l'hypothèse  $H_1$  de calculer les coefficients de pondérations  $w_m$  pour un signal 1D. Étant donné la propriété de séparabilité de la TOD (la 9/7 de Daubechies que nous utiliserons par la suite), il n'est pas très compliqué d'adapter le calcul de ces coefficients pour une image.

Pour une quantification scalaire, il est bien connu qu'à faible distorsion, le débit peut être approximé, d'après la formule de Bennett, par la relation :

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

où  $c$  est une constante qui dépend uniquement du type de quantification (dans le cas d'une quantification scalaire uniforme  $c = 12$ ) et  $H(Y_i^{(m)})$  l'entropie différentielle de la sous-bande transformée  $Y_i^{(m)}$ . Se pose alors le problème de l'allocation optimale de bits qui consiste à déterminer les valeurs des distorsions  $D_i^{(m)}$  qui minimisent le débit pour une distorsion maximale fixée ou vice versa qui

minimisent la distorsion pour un débit maximal fixé. Ce problème d'allocation optimale de bits pour une distorsion cible  $D_c$  consiste à minimiser le débit moyen (2) sous la contrainte que la distorsion moyenne satisfait  $D \leq D_c$ . Pour une transformation  $\mathbf{A}$  donnée, la résolution de ce problème de minimisation sous contrainte est équivalente à la minimisation de

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \sum_{m=1}^M \pi_m \frac{1}{2} \log_2 D_i^{(m)} + \mu \left( \frac{1}{N} \sum_{m=1}^M \sum_{i=1}^N \pi_m w_m w_i D_i^{(m)} - D_c \right) \quad (3)$$

par rapport aux distorsions  $D_i^{(m)}$  et à  $\mu$ . En annulant les dérivées partielles

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial D_i^{(m)}} &= \frac{-\pi_m}{2N \log(2) D_i^{(m)}} + \mu \frac{1}{N} \pi_m w_m w_i \\ \frac{\partial \mathcal{L}}{\partial \mu} &= \frac{1}{N} \sum_{m=1}^M \sum_{i=1}^N \pi_m w_m w_i D_i^{(m)} - D_c, \end{aligned}$$

nous obtenons les valeurs optimales suivantes :

$$\mu_{opt} = \frac{1}{2D_c \log(2)} \quad \text{et} \quad D_{i,opt}^{(m)} = \frac{D_c}{w_m w_i}. \quad (4)$$

En injectant les valeurs optimales de  $D_i^{(m)}$  dans (2) et en remarquant que  $w_i$  est le  $i^{\text{ème}}$  élément diagonal de la matrice  $\mathbf{A}^{-T}\mathbf{A}^{-1}$ , nous démontrons que la transformation  $\mathbf{A}$  est optimale à haut débit (i.e., sous les conditions d'application de la formule de Bennett et l'hypothèse  $H_1$ ) si et seulement si elle minimise le critère

$$C(\mathbf{A}) = \sum_{i=1}^N \sum_{m=1}^M \pi_m H(Y_i^{(m)}) + \frac{1}{2} \log_2 \det \text{diag}(\mathbf{A}^{-T}\mathbf{A}^{-1}) \quad (5)$$

où  $\det$  et  $\text{diag}$  désignent respectivement le déterminant et la matrice diagonale obtenue en ne prenant que les termes diagonaux de son argument. Remarquons que ce critère est invariant par multiplication à gauche de son argument par une matrice diagonale non singulière, ce qui revient à dire que la solution trouvée en minimisant ce critère est obtenue à une matrice diagonale non singulière multipliative près.

### 3 Minimisation du critère

Pour minimiser le critère de l'équation (5), nous avons utilisé une variante de la méthode de Newton qui consiste à effectuer un développement de Taylor au second ordre du critère autour du point courant et à rechercher le minimum de la forme quadratique ainsi obtenue. Plus précisément, nous effectuons un développement de Taylor au second ordre de  $C(\mathbf{A} + \mathcal{E}\mathbf{A})$  par rapport à la matrice  $\mathcal{E}$  autour de  $\mathcal{E} = \mathbf{0}_N$ , où  $\mathbf{0}_N$  est la matrice nulle d'ordre  $N$ . Puis comme dans [5], pour simplifier l'expression du hessien nous supposons que la matrice  $\mathbf{M} = \mathbf{A}^{-T}\mathbf{A}^{-1}$  est quasi-diagonale obtenant ainsi une forme quadratique des

éléments  $\mathcal{E}_{ij}$  de  $\mathcal{E}$  qui est toujours définie positive :

$$C(\mathbf{A} + \mathcal{E}\mathbf{A}) = C(\mathbf{A}) + \sum_{1 \leq i \neq j \leq N} \left[ \frac{\mathbf{M}_{ii}}{2\mathbf{M}_{jj}} \mathcal{E}_{ij}^2 + \mathcal{E}_{ij} \mathcal{E}_{ji} \right] + \sum_{1 \leq i \neq j \leq N} \left[ \sum_{m=1}^M \pi_m E \left[ Y_j^{(m)} \psi_{Y_i^{(m)}}(Y_i^{(m)}) \right] - \frac{\mathbf{M}_{ji}}{\mathbf{M}_{jj}} \right] \mathcal{E}_{ij} + \frac{1}{2} \sum_{1 \leq i \neq j \leq N} \sum_{m=1}^M \pi_m \mathcal{E}_{ij}^2 E \left[ Y_j^{(m)2} \right] E \left[ \psi_{Y_i^{(m)}}^2(Y_i^{(m)}) \right], \quad (6)$$

où  $\psi_{Y_i^{(m)}}$  est la fonction score [6] de  $Y_i^{(m)}$  et  $E$  désigne l'espérance mathématique. La minimisation du critère (5) se fait par une méthode itérative quasi-Newton qui consiste à calculer  $\mathcal{E}$  en résolvant les systèmes linéaires d'ordre 2 suivants

$$\begin{bmatrix} \Psi_{ij} & 2 \\ 2 & \Psi_{ji} \end{bmatrix} \begin{pmatrix} \mathcal{E}_{ij} \\ \mathcal{E}_{ji} \end{pmatrix} = \begin{pmatrix} \Phi_{ij} \\ \Phi_{ji} \end{pmatrix}, \quad (7)$$

où  $\Psi_{ij} = M_{ii}/M_{jj} + \sum_{m=1}^M \pi_m E[\psi_{Y_i^{(m)}}^2(Y_i^{(m)})] E[Y_i^{(m)2}]$  et  $\Phi_{ij} = M_{ij}/M_{jj} - \sum_{m=1}^M \pi_m E[\psi_{Y_i^{(m)}}(Y_i^{(m)}) Y_j^{(m)}]$  (les termes diagonaux  $\mathcal{E}_{ii}$  étant indéterminés, ils sont mis à zéro), et à remplacer la solution courante  $\mathbf{A}$  par  $\mathbf{A} + \mathcal{E}\mathbf{A}$ . En pratique, les fonctions score et l'entropie différentielle sont calculées avec les estimateurs décrits dans [6]. Nous appelons TLOJP2 (Transformation Linéaire Optimale de JPEG2000) cet algorithme.

Nous avons également recherché la transformation orthogonale qui minimise le critère (5) en procédant comme dans [5]. Dans ce cas, il convient de remarquer que si  $\mathbf{A}$  est orthogonale, alors  $\mathbf{A} + \mathcal{E}\mathbf{A}$  l'est aussi quand  $\mathbf{I} + \mathcal{E}$  est orthogonale. Cette dernière condition est satisfaite au second ordre si  $\mathcal{E}$  est anti-symétrique et la méthode consiste alors à calculer les éléments de  $\mathcal{E}$  avec la relation :

$$\mathcal{E}_{ij} = \frac{C_{ij} - C_{ji}}{D_{ij} + D_{ji} - 2}, \quad (8)$$

où  $D_{ij} = \sum_{m=1}^M E[\psi_{Y_i^{(m)}}^2(Y_i^{(m)})] E[Y_i^{(m)2}]$ , et

$C_{ij} = \sum_{m=1}^M E[\psi_{Y_j^{(m)}}(Y_j^{(m)}) Y_i^{(m)}]$ . La matrice  $\mathbf{A} + \mathcal{E}\mathbf{A}$  obtenue ainsi n'étant pas tout à fait orthogonale, nous y remédions en la remplaçant par  $e^{\mathcal{E}}\mathbf{A}$  qui coïncide avec elle au second ordre. Nous appellerons T00JP2 (Transformation Orthogonale Optimale de JPEG2000) cet algorithme.

## 4 Évaluation des performances

Disposant d'une base de données contenant à la fois des images multi et hyper-spectrales, nous avons évalué les performances de ces deux transformations en utilisant le VM9 (Verification Model version 9) [9] de JPEG2000 pour le codage effectif après la transformation. La mesure de distorsion est l'erreur quadratique moyenne entre l'image originale et l'image reconstituée par le VM9 après le codage pour un débit fixé.

Nous avons utilisé le VM9 (Verification Model version 9) [9]<sup>1</sup> de JPEG2000 pour le codage de six images multi-spectrales (simulations d'images PLEIADES codées sur

12 bpp et ayant 4 bandes monochromatiques). Nous utilisons la transformation identité, la TKL et les transformations retournées par TLOJP2 et T00JP2 pour la transformation spectrale. Ci-dessous nous illustrons une image multi-spectrale de la ville de Toulouse<sup>2</sup> contenant quatre bandes spectrales.

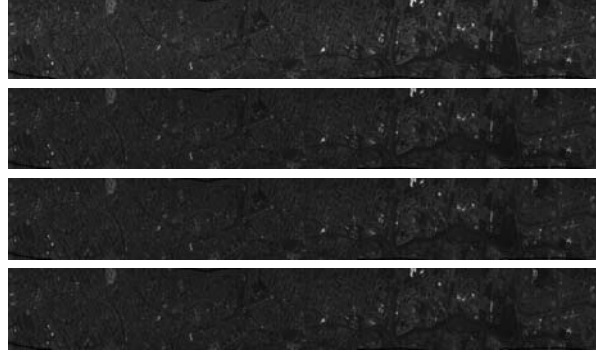


FIG. 1 – Image PLEIADES de Toulouse, de haut en bas, bande 1 à bande 4, B1=[450–530]nm, B2=[510–590], B3=[620–700], B4=[775–915].

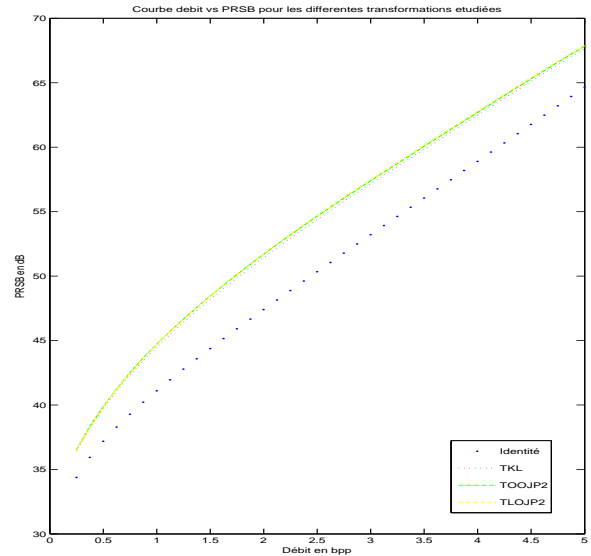


FIG. 2 – Courbes débit vs PRSB de l'image de Toulouse

En se servant du VM9 pour le codage des images, nous relevons pour différents débits le PRSB (Pic du rapport signal à bruit) obtenu pour chacune de ces transformations spectrales. La transformation spectrale Identité est obtenue lorsqu'aucune transformation spectrale n'est appliquée. Nous traçons les courbes débits vs PRSB pour chacune des transformations.

Dans la Fig. 2 nous illustrons les courbes débits vs PRSB des différentes transformations spectrales étudiées. La Fig. 3 quant à elle représente le gain de codage généralisé des transformations spectrales par rapport à une transformation étalon. Le gain de codage généralisé n'est rien d'autre que la différence pour un débit fixé, du PRSB d'une trans-

<sup>1</sup>Les auteurs remercient Jean-Louis Gutwiller de Supélec Metz qui a compilé le VM9 et a facilité son utilisation

<sup>2</sup>Les simulations d'images PLEIADES ont été fournies par le CNES

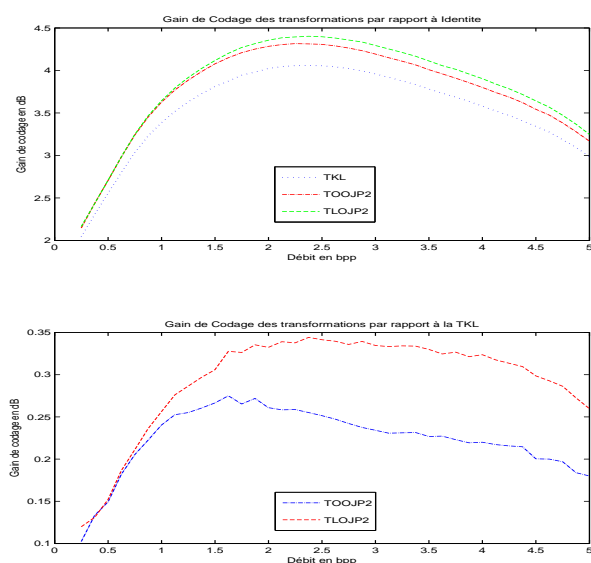


FIG. 3 – Gain de codage généralisé de l'image de Toulouse

TAB. 1 – Valeurs moyennes du PRSB en utilisant le VM9

Debit	0.25	0.50	0.75	1.00	1.50	2.00
Identité	34.64	37.87	40.29	42.37	45.94	49.10
TKL	36.78	40.68	43.62	46.05	50.00	53.20
TOO_JP2	36.84	40.84	43.85	46.33	50.30	53.49
TLO_JP2	36.86	40.87	43.89	46.37	50.35	53.55

formation spectrale par rapport à une transformation étalon fixée (la transformation identité ou la TKL).

Le tableau 1 donne les valeurs moyennes du PRSB à un débit fixé pour les différentes images multi-spectrales dont nous disposons. En observant ce tableau, nous constatons que les nouvelles transformations spectrales que nous proposons permettent d'avoir des performances meilleures que la TKL pour les bas et moyens débits étudiés. Le gain de codage généralisé moyen par rapport à la TKL valant respectivement 0.22 dB et 0.27 dB pour les transformations TOOJP2 et TLOJP2. Ce gain de codage devient légèrement plus important lorsque l'on regarde des débits situés entre 2 et 4 bpp. Nous avons également appliquées ces transformations aux images couleur RVB. Les performances obtenues par le VM9 sont biaisées dans la mesure où le VM9 effectue une quantification des données après la transformation spectrale et avant la TOD. Si cette quantification a un effet négligeable sur les images multi-spectrales qui sont initialement codées sur 12 bpp, l'effet sur les images couleur codées sur 8 bpp devient important et ne permet pas donc de faire une étude objective des performances des transformations spectrales dans la mesure où la distorsion due à cette quantification supplémentaire devient importante. Pour les images hyperspectrales, cette approche permet également d'avoir une amélioration des performances des nouvelles transformations spectrales par rapport à la TKL. L'inconvénient dans ce cas étant la complexité de calcul pour obtenir les transformations spectrales. En effet, les images hyper-spectrales AVIRIS

dont nous disposons contiennent 224 bandes spectrales, chacune étant de taille  $512 \times 512$ .

## 5 Conclusion

Nous avons étudié un schéma de compression d'images multi-composantes compatible avec la norme JPEG2000 et présenté un critère caractérisant la transformation spectrale optimale sous les hypothèses classiques de faible distorsion. Nous avons présenté deux algorithmes permettant de calculer la transformation optimale, l'un sans contrainte et l'autre en imposant l'orthogonalité. Ces transformations offrent des performances légèrement meilleures que la TKL, mais au prix d'un accroissement de la complexité.

## Références

- [1] I.P. Akam Bita, *Sur l'application de l'analyse en composantes indépendantes à la compression des images multi composante*, thèse de l'Université J. Fourier, Grenoble 2007.
- [2] M. Antonini, M. Barlaud, P. Mathieu & I. Daubechies, "Image Coding Using Wavelet Transform", *IEEE Trans. Im. Proc.*, **1** (2), 205–220, 1992.
- [3] E. Christophe, *Compression des images hyperspectrales et son impact sur la qualité des données*, thèse de l'École Nat. Sup. Aéro. & Espace, Toulouse 2006.
- [4] P.L. Dragotti, G. Poggi & A.R.P. Ragozini, "Compression of Multispectral Images by Three-Dimensional SPIHT Algorithm", *IEEE Tran. Geoscience & Remote Sensing*, **38** (1), 416–428, 2000.
- [5] M. Narozny, M. Barret, D.T. Pham & I.P. Akam Bita, "Modified ICA algorithms for finding optimal transforms in transform coding", *4th Int. Symp. Im. & Sig. Proc. & Anal.*, Zagreb (Croatie), 15–17 sept. 2005.
- [6] D.T. Pham, "Fast algorithms for mutual information based independent component analysis", *IEEE Trans. Sig. Proc.*, **52**, (10), 2690–2700, 2004.
- [7] D.S. Taubman & M.W. Marcellin, *JPEG2000 : Image Compression Fundamentals, Standards and Practice*, Kluwer Academic, 2002.
- [8] J. Vaisey, M. Barlaud & M. Antonini, "Multispectral Image Coding Using Lattice VQ and the Wavelet Transform", *IEEE Int. Conf. Im. Proc.*, 307–311, Chicago (USA), 1998.
- [9] *JPEG2000 Verification Model 9.1 (Technical description)*, ISO/IEC JTC 1/SC 29/WG 1 WG1 N2165, 2001.
- [10] 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.