

Extraction conjointe de descripteurs pour l'accès à de bases d'images multispectrales codées dans le domaine transformé en ondelettes

Sarra SAKJI-NSIBI, Amel BENAZZA-BENYAHIA

Unité de Recherche en Imagerie Satellitaire et ses Applications, SUP'COM-Tunis
Université de Carthage, Cité Technologique des Communications, 2083, Tunisie
sakji.sarra@gmail.com, benazza.amel@supcom.rnu.tn

Résumé – Dans cet article, nous étudions l'effet de la compression des images multispectrales dans un système d'accès par le contenu. Nous mettons en évidence l'intérêt d'utiliser des descripteurs issus d'une modélisation statistique multivariée des coefficients d'ondelettes des différentes composantes spectrales ainsi que l'effet de la quantification de ces coefficients sur les performances. Les résultats expérimentaux ont été conduits sur des bases d'images multispectrales SPOT4.

Abstract – In this paper, we are interested in multicomponent image indexing in the wavelet transform domain. In this respect, the joint distribution of the wavelet coefficients through all the spectral channels is modeled by a parametric multivariate model. The contribution of this work consists in emphasizing on the advantage of a multivariate modeling and in studying the effect of the quantization on the retrieval performances. Experimental results are performed on multispectral remote sensing SPOT4 images.

1 Introduction

Ces dernières années ont été marquées par un progrès notable des techniques d'acquisition et de stockage d'images à l'origine de la constitution de bases d'images multispectrales. Ces collections d'images sont régulièrement consultées pour différentes applications comme la télédétection. A cette fin, des systèmes de recherche d'images par le contenu sont conçus afin de permettre à l'utilisateur de retrouver rapidement des images répondant à une requête donnée. Généralement, un système de recherche d'images par le contenu s'articule autour de deux étapes principales [1] : l'extraction des descripteurs du contenu visuel et la procédure de recherche qui vise à trouver les images de la base dont les descripteurs sont les plus proches de ceux de la requête selon une mesure de similarité prédéfinie.

Par ailleurs, les tailles de plus en plus croissantes des bases d'images posent des problèmes au niveau de leur stockage ou de leur transmission sur des canaux de bande passante réduite. Une solution préconisée est le recours à des techniques de compression afin de réduire les redondances de ces données. A cet égard, la compression dans le domaine de la transformée en ondelettes est recommandée car il fournit une représentation parcimonieuse de l'image à différentes résolutions. Celle-ci peut être reconstruite de manière échelonnée en précision ou en résolution [2]. Dans le contexte de recherche d'images, pour éviter de décoder les images de la base avant d'extraire les signatures, une solution est de calculer les signatures *directement* sur les coefficients d'ondelettes. Ceci permet aussi d'exploiter la nature de la représentation échelle-espace en vue d'une meilleure caractérisation visuelle du contenu de l'image [3].

Dans le cas d'images multispectrales, l'approche univariée est

la plus intuitive et consiste à calculer les descripteurs relatifs à chaque canal spectral indépendamment des autres canaux [4]. Une seconde alternative consiste à exploiter la dépendance entre les coefficients des différentes composantes spectrales à travers le choix d'un modèle paramétrique de la distribution conjointe du vecteur des coefficients d'ondelettes de *toutes* les composantes spectrales. Les signatures de l'image sont les paramètres du modèle. Par exemple, une distribution gaussienne généralisée multivariée a été utilisée [5]. Récemment dans [6], nous avons construit des modèles sophistiqués issus de la théorie des copules [7].

A notre connaissance, la plupart des travaux ont mis en œuvre des modèles associés aux coefficients en ondelettes *avant* leur quantification. Or, dans la pratique, accéder à des bases d'images quantifiées est un cas de figure qui se produit couramment. Dans cet article, nous proposons d'étudier l'effet de la quantification des images de la base ou des images requêtes sur les performances du système d'accès par le contenu. L'objectif est de mesurer la perte due à la compression sur les performances du système de recherche d'images multispectrales lorsque les signatures sont extraites à partir de coefficients d'ondelettes quantifiés.

Cet article est organisé comme suit. Dans la section 2, nous présentons le modèle multivarié que nous proposons. Dans la section 3, nous décrivons la procédure de quantification des coefficients d'ondelettes. Nous fournissons les résultats des simulations menées sur des images multispectrales satellitaires SPOT4 dans la section 4 et des éléments de conclusion dans la section 5.

2 Modèles multivariés proposés

Une décomposition en ondelettes sur J niveaux de résolution est appliquée séparément à chaque composante $b = 1, \dots, B$ d'une image multispectrale $(I_0^{(b)})_{b=1}^B$ pour obtenir une image d'approximation et à chaque étage $j = \dots, J$ et 3 sous-bandes de détails $w_{j,o}^{(b)}$ orientées horizontalement ($o = 1$), verticalement ($o = 2$) et diagonalement ($o = 3$). Pour chaque position spatiale (m, n) , le vecteur $\mathbf{w}_{j,o}(m, n)$ de coefficients défini par :

$$\mathbf{w}_{j,o}(m, n) = (w_{j,o}^{(1)}(m, n), \dots, w_{j,o}^{(B)}(m, n))^T \quad (1)$$

est considéré comme la réalisation d'un vecteur aléatoire dont on cherche à modéliser la loi. La méthode la plus simple (dite univariée) consiste à n'utiliser que les lois marginales $f_{w_{j,o}^{(b)}}$. Deux modèles sont couramment employés pour traduire la parcimonie de la décomposition en ondelettes. La première est la distribution gaussienne généralisée (DGG) [2, 8] :

$$\forall x \in \mathbb{R}, \quad f_{w_{j,o}^{(b)}}(x) = \frac{\beta_{j,o}^{(b)}}{2\alpha_{j,o}^{(b)}\Gamma(1/\beta_{j,o}^{(b)})} e^{-(|x|/\alpha_{j,o}^{(b)})^{\beta_{j,o}^{(b)}}} \quad (2)$$

où $\Gamma(z) \triangleq \int_0^{+\infty} t^{z-1} e^{-t} dt$ et $\alpha_{j,o}^{(b)} > 0, \beta_{j,o}^{(b)} > 0$. La seconde loi est la distribution Gamma (DG) et concerne les valeurs absolues des coefficients [9] :

$$\forall x \in \mathbb{R}^+, \quad f_{w_{j,o}^{(b)}}(x) = x^{\gamma_{j,o}^{(b)}-1} \frac{e^{-\delta_{j,o}^{(b)}}}{\delta_{j,o}^{(b)\gamma_{j,o}^{(b)}}\Gamma(\gamma_{j,o}^{(b)})} \quad (3)$$

où $\gamma_{j,o} > 0, \delta_{j,o}^{(b)} > 0$. Notons que les paramètres de ces 2 lois peuvent être facilement estimés notamment par la méthode des moments. Dans l'approche univariée, le vecteur des signatures $\mathbf{f}_{j,o}$ à un étage j et une orientation o donnés est alors formé des paramètres d'échelle et de forme :

$$\mathbf{f}_{j,o} = \begin{cases} (\alpha_{j,o}^{(1)}, \beta_{j,o}^{(1)}, \dots, \alpha_{j,o}^{(B)}, \beta_{j,o}^{(B)})^T & \text{si DGG} \\ (\gamma_{j,o}^{(1)}, \delta_{j,o}^{(1)}, \dots, \gamma_{j,o}^{(B)}, \delta_{j,o}^{(B)})^T & \text{si DG.} \end{cases} \quad (4)$$

De telles signatures présentent l'inconvénient de ne pas tenir compte des dépendances interspectrales. Pour pallier cette lacune, nous proposons d'utiliser un modèle de la loi conjointe $\mathbf{w}_{j,o}$ mettant en œuvre des copules. En effet, le théorème de Sklar [10] stipule que la fonction de répartition conjointe $F_{\mathbf{w}_{j,o}}$ de $\mathbf{w}_{j,o}$ peut s'écrire de la manière suivante :

$$F_{\mathbf{w}_{j,o}}(x_1, \dots, x_B) = C_{j,o}(F_{w_{j,o}^{(1)}}(x_1), \dots, F_{w_{j,o}^{(B)}}(x_B)) \quad (5)$$

pour tout B -uplet (x_1, \dots, x_B) de \mathbb{R}^B et où les $F_{w_{j,o}^{(b)}}$ représentent les fonctions de répartition marginales et où $C_{j,o}$ est la fonction de répartition conjointe de B variables aléatoires ayant des distributions marginales uniformes dans l'intervalle $[0, 1]$. La fonction $C_{j,o}$ est appelée *copule* et traduit les dépendances mutuelles entre les différentes composantes. Dans le cas où $C_{j,o}$ est continue et différentiable, la densité de probabilité conjointe $f_{\mathbf{w}_{j,o}}$ peut s'écrire sous la forme suivante :

$$f_{\mathbf{w}_{j,o}}(x_1, \dots, x_B) = \begin{aligned} & f_{w_{j,o}^{(1)}}(x_1) \dots f_{w_{j,o}^{(B)}}(x_B) \\ & c_j(F_{w_{j,o}^{(1)}}(x_1), \dots, F_{w_{j,o}^{(B)}}(x_B)) \end{aligned} \quad (6)$$

où les $f_{w_{j,o}^{(b)}}$ désignent les densités de probabilité marginales et $c_{j,o}$ la densité de la copule $C_{j,o}$.

A l'instar de l'approche univariée, les distributions marginales $f_{w_{j,o}^{(b)}}$ restent modélisées par les lois DGG ou DG. Concernant la copule $c_{j,o}$, la littérature offre un vaste choix de familles de copules. Pour des raisons de simplicité de l'estimation de ses paramètres, a été retenue la copule gaussienne définie pour tout B -uplet (x_1, \dots, x_B) dans $[0, 1]^B$ comme suit :

$$C_{j,o}(x_1, \dots, x_B) = \varphi_{\mathbf{R}_{j,o}}(\phi^{-1}(x_1), \dots, \phi^{-1}(x_B)) \quad (7)$$

où $\mathbf{R}_{j,o}$ est la matrice de covariance de $\mathbf{w}_{j,o}$, $\varphi_{\mathbf{R}_{j,o}}$ est la fonction de répartition de la gaussienne B -variée $\mathcal{N}(\mathbf{0}, \mathbf{R}_{j,o})$ et ϕ^{-1} est la fonction de répartition inverse de la distribution normale univariée $\mathcal{N}(0, 1)$. Il est important de souligner que dans notre cas, il n'est pas nécessaire d'avoir une expression explicite de la loi $f_{w_{j,o}^{(b)}}$ puisque seuls ses paramètres suffisent à constituer le vecteur des signatures $\mathbf{f}_{j,o}$ défini par :

- les $2B$ paramètres de forme et d'échelle des lois marginales,
- les $B(B-1)/2$ éléments non diagonaux de la partie supérieures de $\mathbf{R}_{j,o}$ reflétant les dépendances interspectrales.

Une fois les descripteurs des images de la base calculés, l'étape suivante est la procédure de recherche qui permet à l'utilisateur de retrouver les images de la base semblables à une image requête. En effet, le système doit comparer le vecteur des signatures de la requête aux vecteurs des signatures stockés dans la base des index. La comparaison se fait selon la mesure de similarité retenue. La distance euclidienne normalisée est souvent utilisée en raison de la rapidité de sa mise en œuvre.

3 Effet de la quantification

En général, les travaux rapportés dans la littérature décrivent l'extraction des caractéristiques pertinentes de l'image à partir des coefficients d'ondelettes *non* quantifiés [8, 6]. Toutefois, notons que certains travaux tiennent compte de l'opération de quantification pour la détermination des descripteurs. Par exemple, dans [11], le nombre de coefficients significatifs dans chaque sous-bande relativement à une séquence de seuils et les histogrammes des approximations successives des coefficients servent de signatures. Dans [12], le corrélogramme de coefficients en ondelettes quantifiés sert de signature, une contribution de ce travail est d'optimiser les valeurs des seuils même si le critère d'optimisation n'a pas été clairement défini. Pour notre part, nous nous intéressons à un autre contexte de travail. En effet, nous supposons qu'est effectuée une compression des images de la collection d'images (ou même des images requêtes) dans le domaine transformé en ondelettes grâce à la quantification des coefficients. Les paramètres de la quantification ne sont pas modifiables par le module qui réalise l'indexation des images comme c'est le cas quand par exemple, la norme de compression de type JPEG2000 est utilisée. Notre objectif est d'étudier l'impact de la quantification sur les performances du systèmes d'accès par le contenu. A ce niveau, il s'avère utile de rappeler que les mesures d'évaluation les plus

courantes sont la précision PR et le rappel RC moyens sont définis par :

$$PR = \frac{1}{Q} \sum_{i_q=1}^Q \frac{R_{i_q}}{R}, \quad RC = \frac{1}{Q} \sum_{i_q=1}^Q \frac{R_{i_q}}{R_{t,i_q}}. \quad (8)$$

où Q est le nombre d'images requêtes, R est le nombre total d'images retournées par le système de recherche, R_{i_q} le nombre d'images retournées considérées comme "pertinentes" par rapport à une image requête i_q selon la vérité-terrain, R_{t,i_q} le nombre total d'images pertinentes par rapport à i_q qui se trouvent dans la base d'apprentissage. Ces critères de performances ne peuvent pas s'exprimer explicitement en fonction des descripteurs retenus. Par conséquent, étudier le problème de l'effet de la quantification s'avère un problème difficile et à notre connaissance, rarement abordé dans la littérature. Pour cela, nous supposons que les vecteurs $\tilde{w}_{j,o}$ issus coefficients d'ondelettes quantifiés $\tilde{w}_{j,o}^{(b)}$ (pour $b = 1, \dots, B$) sont également des réalisations d'un vecteur aléatoire ayant le même type de loi multivariée que le vecteur aléatoire issu des coefficients $w_{j,o}^{(b)}$ non quantifiés. En revanche, les valeurs des paramètres de la loi diffèrent et donc, donnent lieu respectivement à des vecteurs de signatures différents $\tilde{f}_{j,o}$ et $\tilde{f}_{j,o}$. Cette différence est à l'origine d'une dégradation des performances se traduisant par une baisse des valeurs de PR et RC que nous cherchons en premier lieu à évaluer selon le débit retenu.

Pour ce faire, nous avons retenu le quantificateur du codeur *Embedded image coding algorithm using ZeroBlocks of sub-band* (EZBC) [13]. Plusieurs comparaisons entre codeurs échelonnables ont mis en évidence un léger avantage de ce codeur pour la plupart des images naturelles testées. Contrairement au codeur EBCOT de JPEG2000, l'allocation de débit ne se fait au moyen d'une optimisation débit-distorsion. L'algorithme EZBC présente l'avantage d'assurer une complexité opératoire réduite en encodant en un seul symbole les coefficients d'ondelettes non significatifs qui sont mis à zéro. Il a également la particularité d'exploiter les dépendances entre les coefficients d'ondelettes au sein d'une même sous-bande lors du processus de codage.

4 Résultats expérimentaux

Nous utilisons une base d'images qui sont des extraits de taille 256×256 d'images satellitaires SPOT4. Ces images sont relatives à des scènes de différentes régions de la Tunisie. Dans nos simulations, les trois premières composantes ($B = 3$) sont prises en compte et sont séparément décomposées selon un schéma de lifting 5/3 appliqué séparément à chaque canal spectral des images de la base. Les coefficients d'ondelettes ainsi générés sont en premier lieu non quantifiés ensuite quantifiés à 3 débits différents 0.25, 0.5 et 0.75 bpp. Les signatures sont extraites selon les approches univariée et multivariée avec des DG pour les lois marginales. Une recherche exhaustive est appliquée en utilisant un grand nombre d'images requêtes appartenant à la base d'apprentissage. La figure 1 montre les courbes

précision-rappel obtenues dans les cas suivants.

- Cas a : Les coefficients d'ondelettes des images de la base d'apprentissage et ceux des images requêtes ne sont pas quantifiés et l'approche univariée est appliquée.
- Cas b : Les coefficients d'ondelettes des images de la base d'apprentissage et ceux des images requêtes ne sont pas quantifiés et l'approche multivariée est appliquée.
- Cas c : Seuls les coefficients d'ondelettes des images de la base d'apprentissage sont quantifiés à un débit de 0.5 bpp et l'approche univariée est appliquée.
- Cas d : Seuls coefficients d'ondelettes des images de la base d'apprentissage sont quantifiés à un débit de 0.5 bpp et l'approche multivariée est appliquée.
- Cas e : Seuls les coefficients d'ondelettes des images requêtes sont quantifiés à un débit de 0.5 bpp et l'approche univariée est appliquée.
- Cas f : Seuls les coefficients d'ondelettes des images requêtes sont quantifiés à un débit de 0.5 bpp et l'approche multivariée est appliquée.

On remarque une amélioration des performances assurée par l'approche multivariée par rapport à l'approche univariée due à la prise en compte des dépendances interspectrales dans l'étape d'extraction des signatures que les coefficients soient quantifiés ou pas. Comme attendu, la quantification dégrade les performances surtout si elle est appliquée aux coefficients d'ondelettes des images requêtes. L'influence du débit choisi sur les pertes en performances est visible sur les courbes précision-rappel des figures 2 à 5. Il apparaît que plus la perte est significative aux bas débits mais reste acceptable pour un débit de 0.5 bpp pour les images de la collection.

5 Conclusion

Dans ce travail, nous nous sommes intéressés à la recherche par le contenu d'image multicomposantes dans le domaine de codage. Les descripteurs sont les paramètres des lois conjointes employant des copules des coefficients en ondelettes dans toutes les composantes spectrales. D'une part, les expériences menées ont permis de mettre en évidence l'amélioration des performances grâce à la prise en compte des dépendances interspectrales. D'autre part, nous avons pu chiffrer la dégradation des performances entraînée par la quantification des coefficients d'ondelettes des images de la collection ou des images requêtes.

Références

- [1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp. 1349-1380, 2000.
- [2] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Processing*, vol. 1, no. 2, pp. 205-220, 1992.
- [3] J. Beck, A. Sutter, R. Ivry, "Spatial frequency channels and perceptual grouping in texture segregation," *Computer Vision, Graphics, and Image Processing*, vol. 37, no. 2, pp. 299-325, 1987.
- [4] S. D. Newsam, C. Kamath, "Retrieval using texture features in high resolution multi-spectral satellite imagery," *SPIE Defense and Security Symp.*, Orlando, Florida, USA, 2004.

- [5] N. Khelil, A. Benazza-Benyahia, "Wavelet-based multivariate approach for multispectral image indexing," *SPIE Conference on Wavelet Applications in Industrial Processing*, Philadelphia, USA, 2004.
- [6] S. Sakji, A. Benazza-Benyahia, "Copula-based statistical models for multicomponent image retrieval in the wavelet transform domain," *IEEE ICIP*, Cairo, Egypt, pp. 725-728, Nov. 2009.
- [7] R. B. Nelsen, "An introduction to copulas," *Lecture Notes in Statistics*, vol. 139, 1998.
- [8] M. N. Do, M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. on Image Processing*, vol. 11, no. 2, pp. 146-158, 2002.
- [9] R. J. Mathiassen, A. Skavhaug and K. Bo, "Texture similarity measure using Kullback-Leibler divergence between Gamma distributions," *European Conf. on Computer Vision*, Copenhagen, Denmark, vol. 3, pp. 133-147, June 2002.
- [10] A. Sklar, "Fonctions de répartition à n dimensions et leurs marges", *Publ. Inst. Statist. Univ. Paris 8*, pp. 229-231, 1959.
- [11] K. C. Liang, C.-C. J. Kuo, "Progressive image indexing and retrieval based on embedded wavelet coding," *Internat. Conf. on Image Processing*, Santa Barbara, CA, USA, October 1997.
- [12] M. T. Saadatmand, H. A. Moghaddam, "Enhanced wavelet correlogram methods for image indexing and retrieval," *IEEE Internat. Conf. on Image Processing*, Genoa, Italy, September 2005.
- [13] S. T. Hsiang, J. W. Woods, "Embedded image coding using zeroblocks of subband/wavelet coefficients and context modeling," *IEEE Internat. Symp. on Circuits and Systems*, Geneva, Switzerland, May 2000.

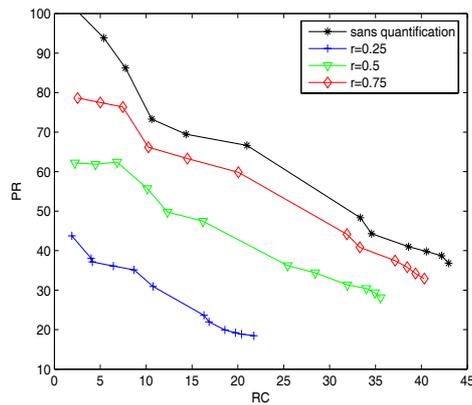


FIG. 3 – Performances de l’approche univariée quand les images de la base et requêtes sont quantifiées.

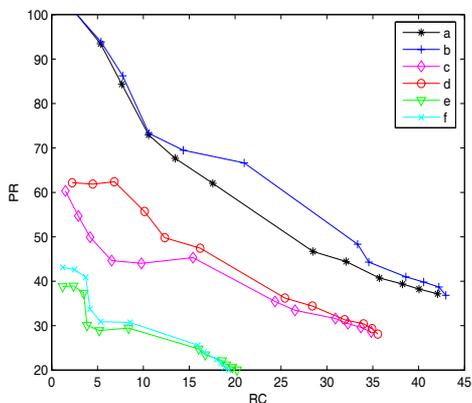


FIG. 1 – Performances des approches univariée et multivariée.

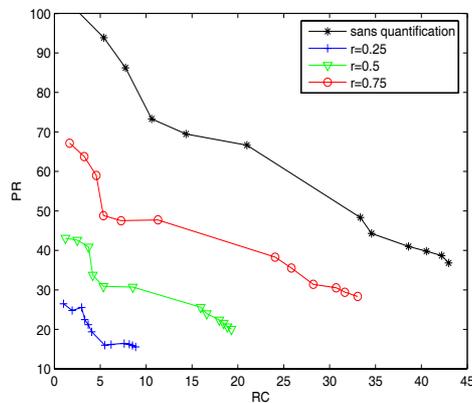


FIG. 4 – Performances de l’approche univariée quand seules les images requêtes sont quantifiées.

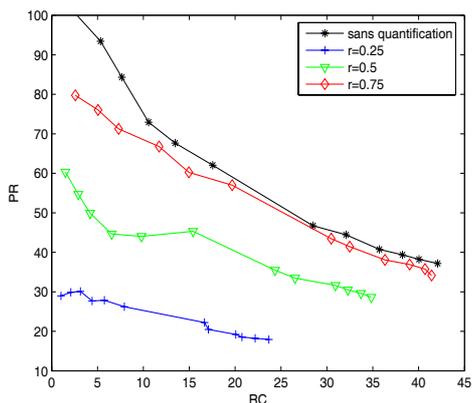


FIG. 2 – Performances de l’approche univariée quand seules les images de la base sont quantifiées et pas les images requêtes.

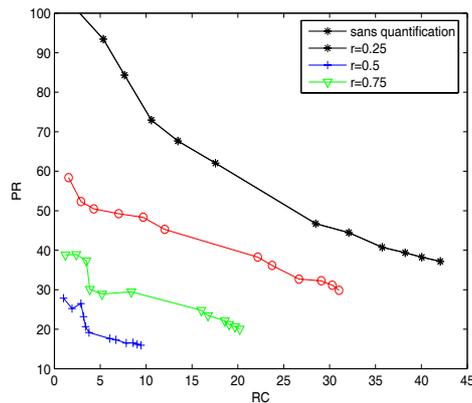


FIG. 5 – Performances de l’approche multivariée pour différents débits quand seules les images requêtes sont quantifiées.