

Extraction d'Information, Amélioration du SNR Et Compression des Données Dans Les Images SAR Multifréquences Multipolarisées

S. Chitroub et B. Sansal

Institut d'Electronique

Université des Sciences et de la Technologie Houari Boumedienne

BP. 32, El-Alia, Bab-Ezzouar, Alger, ALGERIE.

Fax: (213) (02) 51 - 50 - 39.

Résumé

Une méthode statistique permettant d'obtenir une nouvelle représentation optimale des images SAR multifréquences multipolarisées est proposée. La méthode est inspirée de l'analyse en composantes principales à bruit additif ajusté (NAPC) et de la théorie de diagonalisation simultanée de deux matrices symétriques par une seule matrice. Un nombre restreint des nouvelles images transformées décrivant fidèlement les données image SAR originales avec un rapport signal sur bruit amélioré sont alors retenues. Le développement théorique ainsi que la mise en oeuvre de la méthode en question, avec les résultats obtenus sur des images SAR réelles des capteurs radar du système SIR-C, font l'objet de ce présent article.

Abstract

A statistical method for establishing an optimum representation of multifrequency multipolarization SAR images is proposed. This consists in recovering a set of n uncorrelated images from m original images, where $n < m$. The relevant images should contain a maximum of original information with an enhancement in signal-to-noise ratio. The method is inspired from a Noise-Adjusted Principal Components transform (NAPC) and the theory of simultaneous diagonalization of two symmetric matrices by a single matrix. The theoretical analysis, the procedure of implementation, and the results obtained using a real SAR images, provided by the SIR-C system radar sensors, are the subject of the present paper.

1. Introduction

L'imagerie radar à ouverture synthétique (SAR) est aujourd'hui très utilisée dans le domaine de traitement d'images de télédétection. L'acquisition des images SAR, sur une même scène, en plusieurs fréquences et en différentes polarisations apporte des connaissances sur la nature des cibles et leurs orientations et ouvre la porte à des nouvelles applications géophysiques et autres [1][2]. Cependant, les images SAR multifréquences multipolarisées présentent des redondances dans l'information indiquées par les fortes valeurs des coefficients de corrélation entre ces images. Par ailleurs, ces images sont affectées par le speckle qui possède les caractéristiques d'un bruit multiplicatif [2][3]. Cette représentation des images SAR multifréquences multipolarisées rend l'analyse et la classification de ces images difficiles à réaliser.

Récemment, Lee et Hoppel [3] ont présenté la transformation en composantes principales à bruit speckle ajusté des images SAR multifréquences multipolarisées. La transformation est inspirée de l'analyse en composantes principales à bruit additif ajusté (NAPC) [4] adaptée à la nature des images SAR. Cependant, aucune analyse théorique justificative de l'algorithme de calcul de la matrice de la transformation n'est proposée, par conséquent, des difficultés ont été rencontrées dans l'implémentation et l'application de la transformation. Roger [5] a montré que la transformation NAPC est basée sur la théorie de diagonalisation simultanée de deux matrices symétriques par une seule matrice [6], mais aucun résultat pratique n'a été exposé.

Nous proposons ici une méthode statistique, inspirée des références [5] et [6], susceptible de fournir des nouvelles images transformées décrivant fidèlement les images SAR

multifréquences multipolarisées originales. La conception de la matrice de la transformation linéaire de la méthode est faite à partir de la matrice de covariance des images SAR originales et celle du bruit speckle dépendant du signal image. La méthode diagonalise en même temps la matrice de covariance des images originales et celle du bruit speckle. Cette dernière se transforme en une matrice identité. Par conséquent, les nouvelles images obtenues sont affectées par un même bruit décorrélié du signal image et de variance égale à 1. Le rapport signal sur bruit (SNR) se caractérise dans notre méthode par les solutions de l'équation aux valeurs propres généralisées engendrée. Ces valeurs propres sont la diagonale de la matrice de covariance des nouvelles images classées dans l'ordre décroissant. Ces nouvelles images sont alors statistiquement non corrélées et le rapport signal sur bruit est amélioré dans les premières images. Un algorithme pratique permettant l'implémentation de la méthode est proposé. Une version rapide de l'algorithme est aussi avancée en utilisant la décomposition de *Cholesky* de la matrice de covariance du bruit, en supposant que l'estimation de cette matrice est bonne.

Le reste de cet article est organisé comme suit. Le modèle et les statistiques utilisées sont présentés dans la section 2. La section 3 est consacrée à l'analyse théorique de la méthode proposée. L'algorithme d'obtention de la matrice de la transformation linéaire est donné dans la section suivante. Dans la section 5, les résultats commentés sont présentés. Une conclusion du travail présenté ici est donnée à la fin.

2. Modèle et statistiques

Le speckle dans les images SAR (amplitudes ou intensités) a les caractéristiques d'un bruit multiplicatif dans le sens où sa présence est proportionnelle à la valeur du niveau de gris du

pixel et dépendante de la nature de la cible visée [2][3]. On définit $x_i(j, k)$ comme étant le contenu du pixel (j, k) dans la $i^{\text{ème}}$ image SAR originale, $s_i(j, k)$ comme étant le signal image non bruité et $b_i(j, k)$ comme étant le bruit multiplicatif speckle. On supprime l'indice relatif à la position du pixel, on a alors la relation suivante :

$$x_i = s_i \cdot b_i \quad (1)$$

Sous l'hypothèse que le speckle a une moyenne statistique égale à 1, et que x_i et b_i sont statistiquement indépendants [2][3], la relation (1) peut alors écrire sous la forme :

$$x_i = s_i + s_i \cdot (b_i - 1) \quad (2)$$

Le terme $s_i(b_i - 1)$ caractérise la variation du bruit speckle dépendant du signal image. On note alors par X le processus aléatoire stationnaire constitué par les images SAR multifréquences multipolarisées originales, S le processus stationnaire signal image constitué par les images non bruitées, et B le processus de bruit speckle dont la variation est dépendante de celle du signal image. Le modèle de départ est alors le suivant :

$$X = S + B \quad (3)$$

La matrice de covariance du processus X , Σ_X , est alors égale à la somme de la matrice de covariance du processus signal image, Σ_S , et celle du processus du bruit speckle dépendant du signal image, Σ_b . Les statistiques utilisées dans le concept de la matrice de la transformation linéaire sont les matrices de covariance Σ_X et Σ_b .

3. Extraction des composantes principales des images SAR

L'extraction des composantes principales des images SAR multifréquences multipolarisées consiste, via une matrice de transformation linéaire A , à résumer l'information portée par les images originales en un nombre plus restreint de nouvelles images statistiquement non corrélées et de SNR amélioré. Dans la méthode que nous présentons ici, le critère « C » de détermination de la matrice A est le suivant : « *trouvant A pour laquelle la matrice Σ_b se transforme en une matrice identité (bruit de variance égale à 1 dans les nouvelles images obtenues) et la matrice Σ_X se transforme en même temps à une matrice diagonale (nouvelles images statistiquement non corrélées) dont les éléments de la diagonale sont classés dans l'ordre décroissant (variance maximale dans les premières nouvelles images)* ». Le problème posé suivant le critère « C » peut être formaliser mathématiquement comme suit. Les vecteurs ligne de la matrice A sont les vecteurs, a_i , qui maximisent le rapport :

$$\lambda_i = (a_i^T \cdot \Sigma_X \cdot a_i) / (a_i^T \cdot \Sigma_b \cdot a_i) \quad (4)$$

en tenant compte de ces deux contraintes suivantes :

$$\begin{cases} a_i^T \cdot \Sigma_b \cdot a_i = 1 \\ \text{et } a_i^T \cdot a_j = 0 \quad \text{pour } i \neq j \end{cases} \quad (5)$$

L'équation (4) est équivalente à trouver le vecteur a_i tel que :

$$\partial \lambda_i / \partial a_i = (\Sigma_b^{-1} \cdot \Sigma_X - \lambda_i I) \cdot a_i = 0 \quad (6)$$

ou I est la matrice identité. En généralisant les équations (4), (5) et (6) pour le reste des vecteurs ligne de la matrice A , on obtient le système d'équations suivant :

$$\begin{cases} (\Sigma_X - \Sigma_b \cdot \Lambda) \cdot A = 0 \\ A^T \cdot \Sigma_b \cdot A = I \end{cases} \quad (7)$$

ou Λ est la matrice diagonale des valeurs propres, λ_i . Il s'agit donc d'un problème des valeurs propres généralisée mais avec une contrainte de blanchiment de la matrice de covariance de bruit Σ_b . Trouver la matrice A consiste alors à résoudre le système d'équation (7). Nous décrivons dans le paragraphe suivant l'algorithme pratique qui nous permet d'obtenir la matrice A . Deux versions de l'algorithme sont proposées. La première nécessite deux procédures de calcul des valeurs propres et des vecteurs propres. La deuxième est une version rapide de la première, en remplaçant la procédure de calcul des valeurs propres et des vecteurs propres de la matrice Σ_b par sa décomposition de *Cholesky*.

4. Algorithme d'obtention de la matrice A

L'Algorithme d'obtention de matrice A que nous présentons ici est déduit d'un théorème sur les matrices carrées réelles symétriques [5][6]. Le théorème assure l'existence de la matrice A recherchée selon le critère « C » ci-dessus.

Théorème :

Soient C et D deux matrices carrées réelles symétriques. Si C est définie positive, alors il existe une matrice non singulière H telle que : $H^t \cdot C \cdot H = I$ et $H^t \cdot D \cdot H = F$, ou F est une matrice diagonale dont les éléments de la diagonale sont les racines λ de l'équation polynomiale $|D - \lambda C| = 0$.

Comme les deux matrices de covariance, Σ_X et Σ_b qui entrent dans la résolution du système d'équation (7) sont deux matrices carrées réelles symétriques et définies positive, alors et par identification $\Sigma_b \equiv C$ et $\Sigma_X \equiv D$ la matrice A recherchée existe et elle est non singulière. Cependant, du au fait que la matrice de covariance du bruit dépendant du signal image est estimée à partir des images filtrées dans le domaine spatial et les images brutes [3], il arrive, pour des raisons pratiques, que la matrice Σ_b soit une matrice singulière. Dans ces conditions, l'équation polynomiale à considérer dans le théorème ci-dessus est $|I - \lambda \cdot \Sigma_X^{-1} \cdot \Sigma_b| = 0$. L'algorithme qui nous permet d'obtenir la matrice A consiste à réaliser en même temps l'opération de blanchiment de la matrice Σ_b et la diagonalisation de la matrice Σ_X . L'implémentation d'un tel algorithme s'effectue selon les étapes pratique suivantes :

Algorithme :

- (1) Estimation de la matrice Σ_X à partir des images brutes.
- (2) Estimation de la matrice Σ_b à partir des images différences entre les images brutes et les images filtrées.
- (3) A partir de Σ_b , on calcule la matrice des vecteurs propres orthonormés, U , et la matrice diagonale des valeurs propres correspondante, V .
- (4) Construisant la matrice de blanchiment, Φ , telle que :

$$\Phi = U.V^{-1/2} \quad \text{et} \quad \Phi^T . \Sigma_b . \Phi = I \quad (8)$$

(5) Calculant la matrice de covariance des images à bruit speckle ajusté, Ψ , pour laquelle on a :

$$\Psi = \Phi^T . \Sigma_X . \Phi \quad (9)$$

(6) A partir de Ψ on calcule la matrice des vecteurs propres orthonormés, E , et la matrice diagonale des valeurs propres correspondante, P .

(7) La matrice de la transformation linéaire recherchée est alors égale à :

$$A = (\Phi . E)^T = (U.V^{-1/2} . E)^T \quad (10)$$

(8) Les images composantes principales, Y , sont obtenues en projetant les images brutes originales sur la nouvelle base orthogonale A telle que :

$$Y = A.(X - M) \quad (11)$$

ou M est le vecteur des moyennes des images brutes.

Les éléments de la diagonale de la matrice, P , caractérisent le SNR dans les nouvelles images. Ces valeurs propres sont les solutions de l'équation aux valeurs propres généralisées de l'expression (7). Le SNR est alors amélioré dans les premières nouvelles images de variances importantes. La rapidité de l'algorithme peut être amélioré en utilisant la décomposition de *Cholesky* de Σ_b . Ceci consiste à remplacer les étapes (3) à (7) de l'algorithme par les étapes suivantes:

(1) Déterminant la décomposition de *Cholesky* de la matrice Σ_b telle que:

$$\Sigma_b = T^T . T \quad (12)$$

T est la matrice triangulaire supérieure.

(2) Calculant la matrice intermédiaire suivante:

$$R = (T^{-1})^T . \Sigma_X . T^{-1} \quad (13)$$

(3) Calculant la matrice des vecteurs propres orthonormés W , de la matrice R .

(4) La matrice de la transformation linéaire est alors égale:

$$A = (T^{-1} . W)^T \quad (14)$$

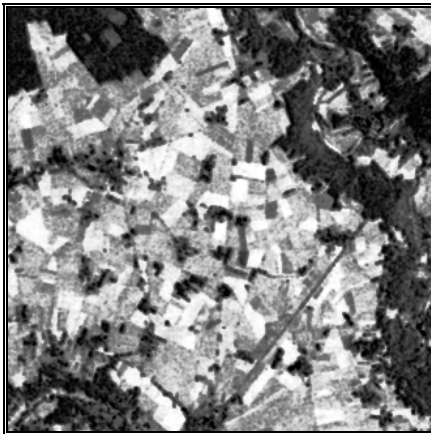
Le nombre d'opérations des étapes (3) à (7) de l'algorithme original est ($\cong 11.N^3$) opérations, ou N est le nombre des images SAR originales, alors que la version rapide demande ($\cong 7.N^3$) opérations. Cependant, la deuxième version donne de bon résultat si l'estimation de Σ_b est bonne. Si Σ_b est mal conditionnée, pour des raisons pratiques, l'algorithme original est préféré.

5. Résultats expérimentaux

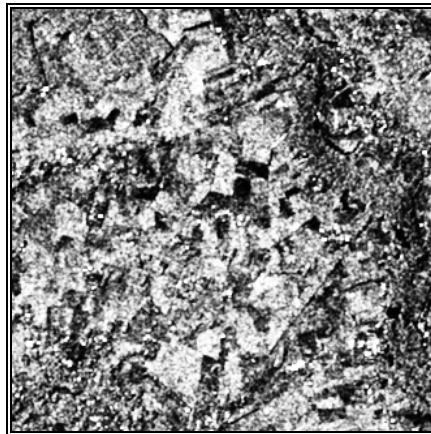
La méthode a été appliquée sur des images puissances SAR prises par les capteurs radar du système SIR-C. Les images couvrent le site de l'Orgeval (Paris) de 500X500 pixels. Ces images sont prises en deux bandes fréquentielles (bande C et bande L), en deux polarisations (HH et HV) pour chacune et d'angle d'incidence 57°. Nous remarquons que ces images ne fournissent pas des informations précises sur le terrain observé et le bruit speckle est très fort. La première image composante principale obtenue (CP1) est très contrastée. Le bruit speckle de départ est largement réduit (SNR $\cong 101.437$, [Tab: 2]) en comparaison avec le meilleur SNR dans les images originales (bande L :HV de SNR $\cong 70.087$, [Tab: 1]). Cette composante présente une amélioration dans les détails locaux de la scène observée ce qui rendra l'exactitude de la classification plus précise en utilisant cette image. La deuxième composante principale CP2 est moins contrastée et un peu bruitée car la variance de CP2 est faible par rapport à celle de CP1. La troisième et la quatrième composante principale sont très dominées par le bruit et l'information portée, différente de celle existant dans CP2, est négligeable. L'information caractérisant l'état de surface du terrain observé est alors extraite et concentrée dans les images CP1 et CP2 avec un SNR très élevé dans l'image CP1.

6. Conclusion

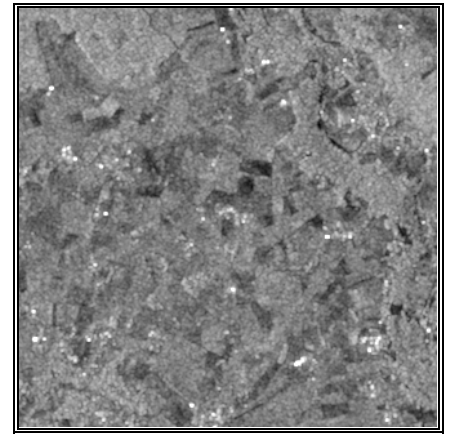
Une méthode statistique pour l'extraction des composantes principales des images SAR multifréquences multipolarisées est proposée. La méthode est basée sur le théorème de diagonalisation des deux matrices réelles symétriques par une seule matrice. Le théorème permet de résoudre l'équation aux valeurs propres généralisée engendrée en tenant compte de la contrainte de blanchiment de la matrice de covariance du bruit speckle dépendant du signal image. Un algorithme pratique d'implémentation de la méthode est donné. Une version rapide de l'algorithme est avancée. Les résultats expérimentaux montrent l'efficacité de la méthode dans l'extraction, compression et amélioration du SNR dans les images SAR multifréquences multipolarisées brutes.



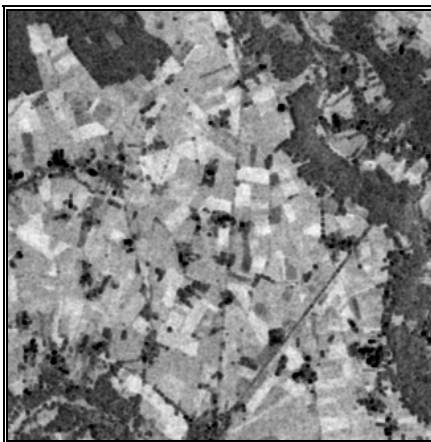
1^{ère} Image Composante Principale



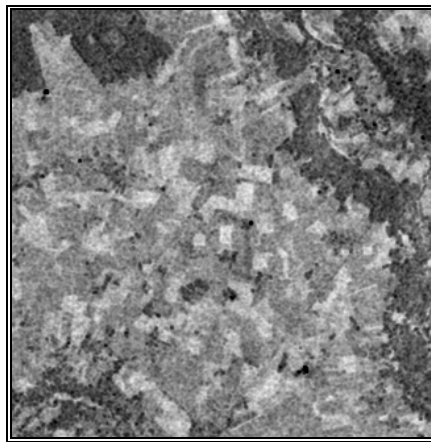
2^{ème} Image Composante Principale



Bande C: Polarisation HH



Bande L: Polarisation HH



Bande C: Polarisation HV



Bande L: Polarisation HV

Image originale	Bande C:HH	Bande L: HH	Bande C:HV	Bande L:HV
SNR	12.925	49.593	21.870	70.087

[Tab : 1] : SNR dans les images brutes originales

Composante Principale	Image CP1	Image CP2	Image CP3	Image CP4
SNR	101.437	14.573	11.130	6.355

[Tab : 2] : SNR dans les images composantes principales

Remerciements

Le premier auteur tient à remercier vivement Monsieur Youcef Chibani membre du Laboratoire de Traitement du signal, Institut d'Electronique, U. S. T. H. B et Docteur Sylvie Le Hégarat-Mascle pour le filtrage et la fourniture des images SAR traitées dans cet article ainsi que pour leur discussion fructueuses.

Références

[1] Rolando L. Jordan, Bryan L., Huneycutt, and Marian Werner, « The SIR-C/X SAR synthetic aperture radar system », *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 33, NO. 4, July 1995, PP. 829-839.

[2] Sylvie Le Hégarat-Mascle, *Classification non supervisée d'images SAR polarimétriques*, Thèse de Doctorat, Spécialité: Signal et Images, ENST 96 E 024, Paris, Septembre 1996.

[3] Jong-Sen Lee and Karl Hoppel, « Principal components transformation of multifrequency polarimetric SAR imagery », *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, NO. 4, July 1992, PP. 686-696.

[4] J. B. Lee, A. S. Woodyatt and M. Berman, « Enhancement of high spectral resolution remote sensing data by a Noise-Adjusted Principal Components Transform », *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 28, NO. 3, May 1990, PP. 295-304.

[5] R. E. Roger, « A faster way to compute the Noise-Adjusted Principal Components Transform Matrix », *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, NO. 6, November 1994, PP. 1194-1196.

[6] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic press, New-York INC, 1972, Chapter 2, PP. 33-36.

[7] W. H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes, The Art of Science Computing*, New-York : Cambridge University Press, 1986.