

La chaîne image SPOT5 THR : un exemple d'optimisation globale

Christophe Latry, Bernard Rougé, Simon Baillarin
Centre National d'Etudes Spatiales
18, avenue Edouard Belin. 31055. Toulouse. France

Email : Christophe.Latry@cnes.fr ; Bernard.Rouge@cnes.fr ; Simon.Baillarin@cnes.fr

Résumé – Issu de la filière SPOT, le satellite de télédétection optique SPOT5 s'en démarque par un mode d'échantillonnage quinquonce adapté à la Fonction de Transfert de Modulation (FTM) de l'instrument et généré par un détecteur spécifique comportant deux barrettes CCD décalées dans le plan focal. En raison des faibles valeurs de la FTM pour les hautes fréquences spatiales, les images produites par un tel système sont floues et nécessitent un traitement complexe d'interpolation quinquonce, de déconvolution linéaire et de débruitage par seuillage de coefficients de paquets d'ondelettes.

Abstract – Following SPOT optical remote sensing satellites, SPOT5 has a specific quincunx sampling scheme well suited to the Modulation Transfer Function (MTF) and based upon a double shifted CCD linear array located in the focal plane. Because of low MTF values for high spatial frequencies, such an acquisition system produces blurred images which need a complex ground processing consisting in quincunx resampling, linear deconvolution and denoising, using wavelet packet thresholding.

1. Introduction

Issu de la famille SPOT, le satellite d'observation optique SPOT5, dont le lancement est prévu en avril 2002, permettra d'atteindre une résolution de 3 mètres dans la bande spectrale panchromatique. Si le principe d'acquisition demeure basé sur l'utilisation de barrettes CCD au plan focal d'un télescope générant les images ligne par ligne par déplacement du satellite sur son orbite, SPOT5 se démarque cependant de ses prédécesseurs par une adaptation de l'échantillonnage permettant de satisfaire au premier ordre la condition de Shannon pour un échantillonnage bidimensionnel correct. La grille d'échantillonnage n'est plus la grille usuelle carrée mais une grille quinquonce, obtenue par entrelacement de deux images carrées de pas d'échantillonnage 5 mètres mais décalées de 2,5 mètres en ligne et en colonne. Ce type d'échantillonnage, noté Supermode et obtenu via l'utilisation d'une double barrette dans le plan focal, permet d'accéder pratiquement à la totalité de l'information fréquentielle non éliminée par le système d'acquisition et génère donc un minimum de repliement de spectre.

Les images sont ensuite comprimées à bord puis transmises au sol et décompressées.

Le choix du Supermode entraîne une sophistication notable de la chaîne image au sol, avec trois phases de traitement : l'entrelacement/interpolation des deux images décalées fournissant une image selon une grille carrée de pas 2,5 mètres, la déconvolution compensant l'atténuation des hautes fréquences spatiales par la Fonction de Transfert de Modulation (FTM) instrumentale, et enfin le débruitage réduisant le niveau de bruit amplifié par la déconvolution dans la partie haute fréquence du spectre.

L'objet de ce papier est de décrire chacun des éléments de la chaîne SPOT5.

2. La chaîne image bord SPOT5

2.1 Rappel des principes d'acquisition

SPOT5 est basé sur le principe d'acquisition PUSHBROOM selon lequel une barrette linéaire de détecteurs photosensibles intègre la lumière provenant d'une ligne au sol, la succession des lignes étant réalisée par le déplacement du point situé à la verticale du satellite perpendiculairement à la barrette de détecteurs. On ajuste le temps d'échantillonnage de telle sorte que le pas d'échantillonnage dans le sens de la vitesse soit égal au pas d'échantillonnage dans le sens de la barrette, c'est-à-dire à la distance entre deux détecteurs élémentaires projetés au sol.

Dans le cas de SPOT5, on utilise ainsi une barrette de 12000 détecteurs élémentaires de taille $a = 6.5\mu\text{m}$. Compte tenu de la focale (1.082m) et de l'altitude du satellite (832 km), chacun des détecteurs élémentaires correspond à un carré de côté 5m au sol en visée verticale. Comme la vitesse du point situé à la verticale du satellite vaut, pour l'orbite SPOT5, $V = 6.6\text{km/s}$, on ajuste le temps d'échantillonnage à $t_e = 0.75\text{ms}$ de telle sorte que $Vt_e = 5\text{m}$, ce qui permet d'obtenir une grille d'échantillonnage à maille carrée de pas 5 mètres.

La particularité de SPOT5 est de disposer d'une double barrette dans le plan focal, la deuxième barrette étant décalée de $0.5a$ dans le sens barrette et de $(n+0.5)a$ dans le sens vitesse. Afin de limiter l'impact des perturbations d'attitude de la plateforme, n est la plus petite valeur technologiquement accessible, soit 3 dans le cas de SPOT5.

2.2 Analyse détaillée de la chaîne bord

2.2.1 Module FTM/bruit instrumental

On peut modéliser l'image en sortie de l'instrument et avant échantillonnage spatial par :

$$s(x, y) = h \hat{A} P + b(x, y) \text{ dans le domaine spatial}$$

$$\hat{s}(n_x, n_y) = FTM \hat{P} + \hat{B} \text{ dans le domaine de Fourier}$$

$$h(x, y) = \text{réponse impulsionnelle instrumentale}$$

$$\hat{A} = \text{opérateur de convolution}$$

$$P = \text{paysage} \quad b = \text{bruit}$$

$$\hat{s}(n_x, n_y) = \iint_{\infty} e^{-j2\pi(xn_x + yn_y)} s(x, y) dx dy$$

L'instrument de prise de vue se comporte avec une très bonne approximation comme un système linéaire et spatialement invariant auquel on associe une réponse impulsionnelle $h(x,y)$, dont la transformée de Fourier est nommée Fonction de Transfert de Modulation (FTM). Image d'un objet ponctuel (dirac spatial), $h(x,y)$ est une tâche dont l'extension caractérise le pouvoir de séparation de l'instrument. L'interprétation est plus naturelle dans le domaine de Fourier : les composantes fréquentielles sont multipliées par la FTM, fonction normalisée à 1 pour les fréquences nulles et globalement décroissante. On trouvera dans [1] une analyse détaillée de la FTM dont les courbes de niveaux sont représentées figure 1.

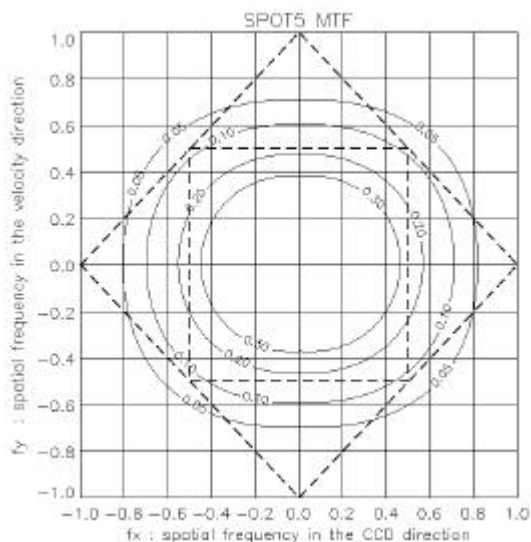


Fig 1 courbes de niveau de la FTM

Le bruit provient de différentes sources : bruit de signal lié à la statistique d'arrivée des photons sur le détecteur, bruit de chaîne, bruit de quantification dépendant du nombre de bits de codage et enfin bruit de compression.

2.2.2 Module échantillonnage

Le principe Supermode de SPOT5 [2][3] consiste à acquérir deux images selon une grille carrée de pas 5 mètres, avec un décalage entre les grilles de 2,5 mètres en ligne et

en colonne. Après entrelacement des images, cela revient à disposer d'un échantillonnage quinconce, c'est à dire d'un échantillonnage selon une grille carrée tournée de 45 degrés par rapport aux axes (vitesse, barrette), avec un pas réduit de $\sqrt{2}$, soit 3,53m.

On constate sur la figure 1 que la FTM SPOT5 est principalement concentrée dans le carré fréquentiel en pointillés tourné de 45 degrés par rapport aux axes f_x et f_y et défini par

$$|f_x| + |f_y| < \frac{1}{5} m^{-1}$$

Ce carré correspond précisément à la zone fréquentiellement accessible par l'échantillonnage quinconce SPOT5, qui satisfait ainsi au premier ordre la condition de Shannon.

2.2.3 Module compression

La compression SPOT5 [4] s'effectue indépendamment sur les deux images 5 mètres décalées pour des raisons de simplicité d'implémentation. Elle comprend les trois opérations classiques de décorrélation, de quantification et de codage. La décorrélation est une transformée DCT par blocs 8x8, la quantification est scalaire et uniforme pour tous les blocs 8x8 appartenant à tous les blocs d'une même ensemble de 8 lignes. Le pas de quantification est choisi par un algorithme de régulation de débit de telle sorte que le débit généré par cet ensemble de 8 lignes oscille faiblement autour d'un débit de consigne correspondant à un taux de compression de 2.815. Il évolue bien sûr pour chaque paquets de 8 lignes en fonction de la complexité de la fauchée.

Cependant, ce pas de quantification de consigne peut s'avérer localement inadapté : par exemple, une zone faiblement texturée noyée dans une fauchée très complexe va être quantifiée avec un pas adapté à la partie complexe. Cela fera apparaître des blocs totalement uniformes ou des damiers (fonctions de base de la DCT) selon qu'un seul ou quelques coefficients DCT auront survécu à la quantification. Pour remédier à ce problème, un traitement d'exception permet, après identification de ces zones « sensibles », de leur appliquer un facteur de quantification plus faible que celui déterminé par la régulation.

3. La chaîne sol SPOT5

3.1 Analyse des modules

3.1.1 Décompression

La décompression consiste, à partir des codes à longueur variable produits par le compresseur, à régénérer les coefficients DCT quantifiés et à effectuer une DCT inverse pour chacune des deux images 5 mètres.

3.1.2 Insertion dans une grille à 2,5m

Le produit final THR sera présenté selon une grille de pas 2,5 mètres. Chacune des deux images est donc insérée dans

une image deux fois plus fine par ajout de zéros aux points manquants, à raison de trois zéros pour un point renseigné.

3.1.3 Interpolation

L'objectif de cette étape est d'obtenir une information radiométrique sur tous les points de la grille de pas 2,5m, alors que seulement la moitié de cette information est disponible dans les deux images 5 mètres.

La phase d'interpolation consiste à convoluer chacune des deux images par deux filtres et à sommer le résultat. La définition de ces filtres dépend du décalage effectif entre les deux images. L'analyse théorique [2] montre en effet qu'un décalage différent d'un demi pixel en ligne et en colonne conduit à la même information que dans le cas nominal : on ne parle plus d'échantillonnage quinconce mais d'échantillonnage entrelacé. Les filtres d'interpolation ont alors une expression dépendant du décalage [2]. Il faut toutefois que le décalage soit constant, au moins localement, c'est à dire sur le support des filtres d'interpolation. En toute rigueur, ces filtres ont un support infini et se comportent comme la fonction sinc définie par :

$$\text{sinc}(x) = \frac{\sin(x)}{x}$$

Dans la pratique on construit des filtres à décroissance beaucoup plus rapide par régularisation de leur transformée de Fourier.

Par ailleurs, le décalage évolue lentement dans le champ et n'est différent de la valeur nominale que pour des visées hors trace. Son estimation est possible par modélisation géométrique mais sa détermination exacte par corrélation automatique est envisageable dans le cas contraire.

A l'issue de cette étape, on dispose d'une image dont le spectre est très atténué sur la frontière de la zone fréquentiellement accessible par l'échantillonnage quinconce (carré fréquentiel tourné à 45 degrés), du fait de la faible valeur de la FTM. On remarque sur les courbes de niveau de la FTM (cf Fig 1) que le repliement de spectres est nul selon les directions f_x et f_y et maximal (mais limité) pour les directions fréquentielles orientées à 45 degrés. La quasi absence de repliement de spectres justifie d'ailleurs le processus d'interpolation d'un point de vue mathématique.

3.1.4 Déconvolution

On cherche ici à compenser le flou associé à l'atténuation des fréquences spatiales par la FTM en bande passante, c'est-à-dire à l'intérieur du carré fréquentiel tourné à 45 degrés.

La déconvolution consiste à appliquer un filtre dont la transformée de Fourier $D(f_x, f_y)$ soit telle que $D \times \text{FTM} = I$, I représentant la FTM d'un capteur idéal vérifiant les propriétés suivantes :

- (a) I est voisin de 1 sur le carré fréquentiel tourné à 45 degrés, 0 à l'extérieur : l'atténuation en bande passante est minimale et il n'y a pas de repliement de spectres
- (b) $\text{TF}^{-1}(I)$ est positive : la réponse impulsionnelle du capteur « idéal » est positive

- (c) $\text{TF}^{-1}(I)$ est la plus compacte possible : la réponse impulsionnelle du capteur « idéal » doit posséder un bon pouvoir de résolution

Les propriétés (a) et (c) sont contradictoires : une fonction ne peut-être à support compact simultanément dans les domaines spatiaux et fréquentiels. Des compromis sont donc nécessaires entre forte déconvolution et limitation des artefacts, du type rebonds au voisinage des transitions associées à des réponses impulsionnelles oscillantes et à décroissance lente. Une solution envisageable pour I est la fonction « prolate sphéroïdale » [5] qui permet une optimisation de la concentration conjointe dans les deux espaces en termes de norme L_2 .

3.1.5 Débruitage

L'opération de déconvolution ne change pas le rapport signal sur bruit : elle remonte indifféremment les composantes fréquentielles du bruit et du signal utile. Or ce rapport est mauvais dans les hautes fréquences car le signal utile est fortement atténué par la FTM alors que la densité spectrale du bruit est constante (on parle de « bruit blanc »).

Visuellement, la déconvolution rehausse les contrastes mais sur les zones de paysage uniforme, elle amplifie le bruit. Sur une zone uniforme ne contenant que du bruit, l'écart-type de celui-ci peut être quadruplé pour des fonctions de déconvolution « forte » et devenir très perceptible.

La technique de débruitage adoptée pour la chaîne de traitement sol SPOT5 [6][7] a pour objectif d'obtenir sur les zones uniformes de l'image finale un bruit blanc dont l'écart-type est égal à K fois l'écart-type du bruit de l'image de départ, non déconvoluée. Dans le cas de SPOT5, on adoptera probablement la valeur 1 pour K : après déconvolution, le bruit sur les zones uniformes sera blanc et de même amplitude qu'avant déconvolution.

L'algorithme de débruitage s'appuie sur une transformation en paquets d'ondelettes, c'est à dire un développement de l'image selon une base de fonctions localisées à la fois spatialement et fréquentiellement :

- la localisation spatiale permet à la fois de discriminer les zones quasi uniformes des zones fortement texturées et de calculer un niveau de bruit instrumental **qui dépend du signal local**. Les coefficients d'ondelettes supérieurs à un seuil variant avec le bruit local (par exemple 3,5 fois l'écart type du bruit local) sont considérés comme représentatifs d'une véritable information et ne sont pas modifiés. Seuls les « petits » coefficients subissent une atténuation en amplitude.
- la localisation fréquentielle permet d'appliquer un seuillage différent selon les fréquences : comme on veut un bruit blanc en sortie, les coefficients d'ondelettes correspondant aux fréquences pour lesquelles la fonction de déconvolution D prend des valeurs élevées doivent être plus fortement atténués. On voit en particulier que la localisation fréquentielle doit être performante dans les zones où la fonction de déconvolution varie rapidement, c'est à dire dans les

hautes fréquences. Pour disposer d'une bonne finesse fréquentielle notamment dans les hautes fréquences, on choisit une décomposition en paquets d'ondelettes en qui assure un découpage régulier du domaine fréquentiel plutôt qu'une décomposition dyadique qui ne découpe finement que les basses fréquences. Le choix des ondelettes (splines) participe aussi à la bonne sélectivité fréquentielle.

Dans le cas de SPOT5, on décompose l'image déconvoluée sur 5 niveaux de décomposition et on répartit ensuite les opérations de seuillage sur les niveaux de recombinaison 5, 4, 3 2 et 1 afin de bénéficier du compromis de localisation spatio-fréquentielle.

Le seuillage multiplicatif adopté assure (contrairement au seuillage « dur » ou « doux ») que les faibles textures ne disparaissent jamais totalement. Les coefficients de seuillage à chaque niveau ne sont pas déduits directement de la fonction de déconvolution mais issus d'une phase dite d'apprentissage, consistant à débruiter une image de bruit blanc déconvoluée : les seuils sont alors déduits des rapports d'écart-type entre image de bruit déconvoluée et image de bruit non déconvoluée.

La décomposition en paquets d'ondelettes est un traitement lourd en termes de nombre d'opérations par point, impliquant des phases de filtrages et de décimation lors de la décomposition et des phases de suréchantillonnage et de filtrage lors de la recombinaison.

Cette méthode de débruitage sera implantée dans la chaîne de traitement des images Supermode SPOT5.

4. Conclusions

Le satellite SPOT5 est le premier de la famille SPOT à posséder un mode d'échantillonnage quinconce lui permettant d'approcher de façon satisfaisante la condition de Shannon. L'adaptation de l'échantillonnage à l'instrument permet d'accéder à la quasi totalité des fréquences spatiales non filtrées par l'instrument et d'améliorer significativement sa résolution par rapport à un échantillonnage classique à maille carré. La chaîne image est alors notablement plus complexe que celle des précédents SPOT puisqu'après avoir subi une compression/décompression de type JPEG, les images quinconces doivent être interpolées, déconvoluées et débruitées : on approche ainsi les 1000 opérations/pixel, complexité néanmoins compatible des performances des processeurs actuels puisque le traitement d'une image SPOT5 de 24000x24000 pixels prendra une demi-heure. Le choix de cette chaîne et sa mise au point s'appuient sur un grand nombre d'analyses de sensibilité, elles-mêmes basées sur des simulations globales et des évaluations quantitatives. L'évaluation qualitative par les futurs utilisateurs est cependant incontournable en l'absence de critères qualité image totalement fiables.

SPOT5 préfigure les futures chaînes bord/sol des satellites d'observation à haute résolution, caractérisées par un échantillonnage approchant de façon satisfaisante la condition de Shannon, des compresseurs fonctionnant à

des taux de compression plus élevés basés sur une transformée en ondelettes et une chaîne de restauration sol pouvant inclure des traitements sophistiqués tels que la fusion d'image panchromatique à haute résolution et couleurs à plus basse résolution, le mosaïquage ... La simulation de la globalité de la chaîne et le recours aux expérimentateurs seront tout aussi indispensables.

References

- [1] C. Latry, B. Rougé. "SPOT5 THR mode". *Proceedings of SPIE 1998 (San Diego)*. Vol 3439, p 480-492
- [2] C. Latry, B. Rougé. "Optimized sampling for CCD instruments : the Supermode scheme". *Proceedings of IGARSS 2000 (Honolulu)*.
- [3] C. Latry, B. Rougé. Brevet national. « Procédé d'acquisition d'une image par balayage pousse-balai ». 28 Juillet 1995. Brevet n° 95 09242.
- [4] P. Lier, G. Moury, C. Latry and F. Cabot, "Selection of the SPOT-5 Image Compression algorithm", *Proceedings of SPIE 98(San Diego)*, vol.3439-70
- [5] D Slepian, H.O.Pollack. « Prolate spheroidal wave functions, Fourier analysis and Uncertainty ». *Bell Syst. Techn J*, Vol 40, pages 43-64, 1961.
- [6] B Rougé. "Théorie de la Chaîne image optique et restauration." *PhD thesis, Université Paris-Dauphine, 1997. Thèse d'habilitation.*
- [7] B. Rougé. « Space-scale frequency representation and fixed chosen noise restoration ». *Proceedings of the IEEE-SP International Symposium on time-frequency and time-scale analysis. October 25-28 1994. Philadelphia, Pennsylvania, USA.*