HAL
open science

Destriping d'images satellitaires push-broom pour un modèle affine de réponse des détecteurs

Mehdi Chahine Amrouche, Hervé Carfantan, Jérôme Idier, Aurelie Michel, Vincent Martin

To cite this version:

Mehdi Chahine Amrouche, Hervé Carfantan, Jérôme Idier, Aurelie Michel, Vincent Martin. Destriping d'images satellitaires push-broom pour un modèle affine de réponse des détecteurs. Colloque GRETSI, Aug 2019, Lille, France. hal-02337756

HAL Id: hal-02337756

https://hal.science/hal-02337756v1

Submitted on 29 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Destriping d'images satellitaires *push-broom* pour un modèle affine de réponse des détecteurs

Mehdi Chahine AMROUCHE[1], Hervé CARFANTAN[1], Jérôme IDIER[2], Aurélie MICHEL[1], Vincent MARTIN[3]

[1]Institut de Recherche en Astrophysique et Planétologie, CNRS/UPS
14 avenue Édouard Belin, 31400 Toulouse, France

[2]Laboratoire des Sciences du Numérique de Nantes, CNRS/ECN
École Centrale de Nantes, 1 rue de la Noë, 44321 Nantes Cedex 3, France

[3]Centre National d'Études Spatiales
18 avenue Édouard Belin, 31 401 Toulouse Cedex 9, France

mamrouche@irap.omp.eu, Herve.Carfantan@irap.omp.eu, Jerome.Idier@ls2n.fr,
Aurelie.Michel@onera.fr, Vincent.Martin@cnes.fr

Résumé – Cet article propose une nouvelle méthode de suppression de rayures (*destriping*) d'images acquises par un satellite à balayage (*push-broom*) pour un modèle affine de réponse des détecteurs, dans un cadre d'estimation bayésienne. Un algorithme *Majorize-Minimize* a été développé permettant la minimisation du critère du maximum *a posteriori* pour un coût de calcul très réduit.

Abstract – This paper proposes a new *destriping* method for satellite push-broom-based images, in a Bayesian framework, considering an affine model of the sensor response. We rely on a dedicated *Majorize-Minimize* algorithm to obtain the *maximum a posteriori* estimator at a reduced computation cost.

1 Introduction

Les satellites d'observation de la terre de type *pushbroom* font l'acquisition d'images par l'intermédiaire de barrettes de détecteurs CCD, formant une ligne de l'image. L'avancement du satellite dans une direction orthogonale à ces barrettes permet à chaque détecteur de former une colonne de l'image. En pratique, un manque d'uniformité de la réponse des détecteurs produit sur l'image un effet de rayures (effet dit de *striping*). Lors de la recette en vol du satellite, les défauts des détecteurs sont identifiés et partiellement corrigés (phase de calibration des détecteurs). Néanmoins, l'évolution des détecteurs au cours du temps peut nécessiter une adaptation de ces corrections.

Dans la littérature, les techniques de *destriping* peuvent être classées en trois catégories. Les premières reposent sur les propriétés statistiques de l'image, et supposent que les statistiques de la portion de scène acquise par chaque capteur sont identiques pour tous les capteurs (la moyenne et variance pour la méthode de *Moment Matching* [6] ; la densité de probabilité pour l'approche *Histogram Matching* [1]). Malheureusement, ces approches sont sensibles à la structure de la scène observée. Une deuxième classe de méthodes est basée sur le filtrage[8, 10, 7] dans le but de tronquer les composantes dues aux rayures dans le domaine de filtrage choisi. Cependant, les composantes filtrées contiennent également de l'information utile sur la scène, qui se retrouve supprimée, induisant des artefacts (phénomène de Gibbs, effets de lissage). Récemment, une troisième approche est apparue, modélisant le phénomène à l'origine des rayures, puis corrigeant ce phénomène après estimation des paramètres du modèle via l'optimisation d'un critère [2, 3]. Dans [2], un défaut de la réponse linéaire des détecteurs est considéré pour le *destriping* d'images de type SPOT5. Dans [3], les rayures sont supposées additives et de faible rang. Hormis l'adéquation nécessaire du modèle pris en compte, l'inconvénient de ces méthodes est le coût calculatoire des algorithmes utilisés, notamment dans le cas d'images de grandes tailles.

Les travaux présentés ici se placent dans la lignée de [2]. Nos contributions principales sont d'une part la prise en compte d'un modèle affine des détecteurs, qui s'avère utile pour de nombreux imageurs *pushbroom* ; d'autre part, la proposition d'un algorithme d'optimisation de type *Majorize-Minimize* (MM) de coût de calcul limité, permettant la correction d'images à haute résolution des satellites de type PLEIADES.

La suite de cet article est organisée comme suit : nous présentons d'abord le modèle affine des détecteurs considéré. La modélisation statistique de l'image et des paramètres, ainsi que l'estimateur au sens du *maximum a posteriori* (MAP) sont présentés en Section 3. Les algorithmes d'optimisation proposés font l'objet de la Section 4. Finalement, la section 5

Ce travail a été partiellement financé par le CNES. Aurélie Michel travaille maintenant à l'ONERA, 2 avenue Édouard Belin, 31055 Toulouse Cedex 4.

est une synthèse des résultats obtenus.

2 Correction pour une réponse affine des détecteurs

Les rayures étant dues à un défaut de calibration de la réponse affine des détecteurs, il est naturel de vouloir les corriger via un modèle affine. Pour une image $\boldsymbol{w}$ de R lignes et C colonnes acquise par l'instrument, l'image idéale de la scène observée est $\boldsymbol{z}$ telle que :

$$z_{r,c} = g_c\, w_{r,c} - o_c, \quad r = 1, \ldots, R,\ c = 1, \ldots, C, \tag{1}$$

avec, r et c les indices de ligne et de colonne. Pour chaque colonne d'indice c, g_c est le « gain » et o_c « l'offset » de correction du détecteur.

L'objectif est donc d'estimer les coefficients de gains $\boldsymbol{g} = \{g_c\}_{c=1\ldots C}$ et d'offsets $\boldsymbol{o} = \{o_c\}_{c=1\ldots C}$ à partir de la seule image acquise $\boldsymbol{w}$. Pour cela, nous faisons appel à une modélisation statistique de l'image idéale $\boldsymbol{z}$, des gains $\boldsymbol{g}$ et offsets $\boldsymbol{o}$ et, dans un cadre bayésien, définissons les estimateurs de $\boldsymbol{g}$ et $\boldsymbol{o}$ au sens du maximum a posteriori. Le calcul de ces estimateurs nécessite ainsi l'optimisation d'un critère. La correction de l'image une fois les paramètres $\boldsymbol{g}$ et $\boldsymbol{o}$ estimés est alors immédiate via l'équation (1).

3 Estimation

3.1 Modèle pour l'image

Comme proposé dans [2], l'image est modélisée par un champ de Markov non gaussien du premier ordre. Pour le modèle considéré ici, la densité de probabilité de l'image s'écrit :

$$f_z(\boldsymbol{z}) \propto \exp\left(-\frac{1}{T}\sum_{(r,c)\sim(r',c')} \phi_s(z_{r,c} - z_{r',c'})\right) \tag{2}$$

où $(r,c) \sim (r',c')$ signifie que la somme est effectuée sur l'ensemble des pixels voisins verticaux et horizontaux.

Concernant la fonction dite de « potentiel » ϕ_s, nous nous sommes concentrés sur les deux fonctions classiques suivantes :

$$\phi^{2-1}(x) = \sqrt{x^2 + s^2} - s, \qquad \phi^{2-0}(x) = \frac{x^2}{x^2 + s^2}.$$

D'un point de vue de la modélisation, la fonction ϕ^{2-0} est préférable car elle permet de mieux préserver les fortes variations radiométriques au niveau des contours de l'image. Cependant, cette fonction n'est pas convexe, ce qui pourra causer des problèmes pour le calcul de la solution proposée. De son côté, la fonction ϕ^{2-1} a l'avantage d'être convexe, tout en pénalisant moins les contours qu'une simple fonction quadratique.

T et s sont les hyper-paramètres du modèle markovien liés à image $\boldsymbol{z}$ qu'il faudra donc régler pour utiliser notre méthode. Le paramètre s joue le rôle du seuil entre le comportement quadratique et asymptotique et T est un paramètre d'échelle.

3.2 Modèle pour les paramètres

Dans le cas idéal (acquisition sans effet de rayure), l'image acquise est égale à l'image idéale. En d'autres termes, les gains de correction sont unitaires et les offset de correction sont nuls :

$$(g_c, o_c) = (1, 0), \quad \forall\, c = 1, \ldots, C. \tag{3}$$

Si l'on considère que la radiométrie globale de l'image est préservée, malgré l'effet des rayures, on doit avoir $\mathbb{E}[z_{r,c}] = \mathbb{E}[g_c]\,\mathbb{E}[w_{r,c}] + \mathbb{E}[o_c]$. Il est donc naturel de considérer des gains de moyenne unité et des offsets de moyenne nulle. Pour forcer cette préservation de la radiométrie, on considérera en pratique :

$$\sum_c g_c = C \ \text{ et } \ \sum_c o_c = 0. \tag{4}$$

En dehors d'information particulière sur ces paramètres, hormis leur ordre de grandeur, nous les avons considérés indépendants et identiquement distribués suivant une loi gaussienne :

$$\forall c,\ g_c \sim \mathcal{N}(1, \sigma_g^2) \ \text{ et } \ o_c \sim \mathcal{N}(0, \sigma_o^2),$$

l'ordre de grandeur des variances σ_g^2 et σ_o^2 pouvant être supposé connu pour un instrument donné.

3.3 Estimation bayésienne

A partir des modèles probabilistes présentés ci-dessus, et par le biais de la règle de Bayes, on peut écrire la loi *a posteriori* des paramètres $\boldsymbol{g}$ et $\boldsymbol{o}$ sachant les données $\boldsymbol{w}$:

$$f(\boldsymbol{g}, \boldsymbol{o}|\boldsymbol{w}) = f(\boldsymbol{w}|\boldsymbol{g}, \boldsymbol{o}) f(\boldsymbol{g}) f(\boldsymbol{o}) / f(\boldsymbol{w}) \tag{5}$$

avec la vraisemblance :

$$f(\boldsymbol{w}|\boldsymbol{g}, \boldsymbol{o}) \propto \prod_{r,c} |g_c| \cdot f_z(\{g_c w_{r,c} - o_c\}_{r,c}), \tag{6}$$

où le terme $|g_c|$ est lié au Jacobien du changement de variable de la loi $f(\boldsymbol{z})$ à $f(\boldsymbol{w}|\boldsymbol{g}, \boldsymbol{o})$.

L'estimateur du maximum *a posteriori* (MAP) correspond donc aux paramètres $\boldsymbol{g}$ et $\boldsymbol{o}$ qui maximisent cette loi *a posteriori*, ou de façon équivalente minimisent l'opposé de son logarithme, soit le critère à minimiser :

$$\begin{aligned} J(\boldsymbol{g}, \boldsymbol{o}) &= \frac{1}{T}\sum_{r,c} \phi_s\left(g_c\left(w_{r,c} - w_{r+1,c}\right)\right) \\ &+ \frac{1}{T}\sum_{r,c} \phi_s\left((g_c w_{r,c} - o_c) - (g_{c+1} w_{r,c+1} - o_{c+1})\right) \\ &- R\sum_c \log|g_c| + \lambda_g \sum_c (g_c - 1)^2 + \lambda_o \sum_c o_c^2 \end{aligned} \tag{7}$$

où $\lambda_g = 1/2\sigma_g^2$ et $\lambda_o = 1/2\sigma_o^2$. Ce critère doit être minimisé sous les contraintes (4). Cependant, il est aisé de montrer que la solution satisfait toujours la condition $\sum_c o_c = 0$; aussi seule la contrainte $\sum_c g_c = C$ doit être prise en compte en pratique.

Notons que ce critère est strictement convexe si la fonction ϕ utilisée est convexe (cas de ϕ^{2-1}), ce qui nous assure l'unicité de la solution et l'absence de minima locaux. Dans le cas

contraire (cas de ϕ^{2-0}), les algorithmes de descente locale tel que celui que nous proposons dans la suite ne garantissent pas d'obtenir un minimiseur global s'il existe des minima locaux. En pratique, nous les initialiserons toujours avec la solution nominale (3).

4 Optimisation

4.1 Gradient conjugué et préconditionnement

Une approche classique pour la minimisation de (7) sous la contrainte linéaire $\sum_c g_c = C$ consiste à utiliser un algorithme de descente tel que l'algorithme du gradient conjugué, avec projection de la direction de descente sur le domaine admissible. Les paramètres $\boldsymbol{x}^{(k)}$ (dans notre cas $\boldsymbol{g}$ et $\boldsymbol{o}$), à l'itération k sont mis à jour de la façon suivante :

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \alpha^{(k)} \boldsymbol{d}^{(k)}$$

où les directions de descente $\boldsymbol{d}^{(k)}$ sont choisies de façon classique pour des gradients conjugués (de type Polak-Ribière par exemple), puis projetées, et la recherche du pas $\alpha^{(k)}$ se fait par une minimisation 1D. Afin de s'assurer de la convergence de l'algorithme, nous avons proposé de prendre en compte de façon rigoureuse la barrière logarithmique lors de la recherche du pas $\alpha^{(k)}$ en exploitant une approximation majorante adaptée, comme proposé dans [5].

L'utilisation d'un préconditionneur $\mathbf{P}$ consiste à adopter une direction de descente de type $\mathbf{P}\boldsymbol{d}^{(k)}$, dans l'idée de se rapprocher de la vitesse de convergence de la méthode de Newton, en considérant une matrice de préconditionnement $\mathbf{P}$ définie positive et d'un coût de calcul réduit. On propose ici de choisir pour préconditionneur $\mathbf{P}$ l'inverse du Hessien calculée à la valeur nominale des paramètres (3). Les temps de calcul donnés Tab. 1 illustrent le fait que cette approche de préconditionnement (PGC) permet effectivement de réduire le coût de calcul par rapport à un algorithme du gradient conjugué (GC) standard.

TABLE 1 – Comparaison du temps de calcul (en secondes) des algorithmes de gradient conjugué (GC), gradient conjugué préconditionné (PGC) et MM proposés sur deux images de type PLEIADES.

Image	ϕ	GC	PGC	MM
Image 1	ϕ^{2-1}	72.32	34.30	**13.37**
(4000 × 2000)	ϕ^{2-0}	43.31	20.98	**7.39**
Image 2	ϕ^{2-1}	288.81	62.12	**19.20**
(4000 × 2000)	ϕ^{2-0}	1224.24	52.84	**7.87**

4.2 Simplification du critère

La minimisation du critère (7) est alourdie par la présence de la barrière logarithmique. Aussi, nous nous sommes intéressés à l'influence de chacun des termes qui composent ce critère. D'une part, comme les paramètres restent constants suivant les colonnes, le calcul des différences suivant la direction verticale n'apporte que peu d'information. De même, les paramètres $\boldsymbol{g}$ variant peu autour de 1, les variations du terme logarithmique sont limitées. Des résultats de simulations (comme ceux correspondants au Tab. 2) confirment que ces termes jouent un rôle secondaire et peuvent être éliminés du critère tout en gardant la même qualité d'égalisation. On peut donc en pratique estimer les paramètres $\boldsymbol{g}$ et $\boldsymbol{o}$ minimisant le critère simplifié :

$$K(\boldsymbol{g}, \boldsymbol{o}) = \lambda_g \sum_c (g_c - 1)^2 + \lambda_o \sum_c o_c^2 + \frac{1}{T} \sum_{r,c} \phi_s \left((g_c w_{r,c} - o_c) - (g_{c+1} w_{r,c+1} - o_{c+1}) \right) \quad (8)$$

sous la contrainte $\sum_c g_c = C$. Ce critère ne possédant plus de barrière logarithmique, on peut espérer utiliser un algorithme plus rapide que ceux basés sur le gradient conjugué pour le minimiser.

TABLE 2 – Indices de qualité (PSNR en dB et SSIM [9]) sur des images reconstruites grâce aux paramètres estimés par minimisation du critère original (7) et du critère simplifié (8) sur deux images de type SPOT 5.

ϕ	Critère	PSNR	SSIM
ϕ^{2-1}	Original $J(\boldsymbol{g}, \boldsymbol{o})$, Eq. (7)	58.40	0.99566
	Simplifié $K(\boldsymbol{g}, \boldsymbol{o})$, Eq. (8)	**58.66**	**0.99657**
ϕ^{2-0}	Original $J(\boldsymbol{g}, \boldsymbol{o})$, Eq. (7)	61.18	0.99877
	Simplifié $K(\boldsymbol{g}, \boldsymbol{o})$, Eq. (8)	**61.44**	**0.99888**

4.3 Algorithme MM sous contrainte

L'utilisation d'un algorithme de type *Majorize-Minimize* (MM) est parfaitement adaptée à la minimisation itérative du critère simplifié (8). L'algorithme que nous proposons s'appuie sur le formalisme semi-quadratique et suit le même schéma que l'algorithme ARTUR [4]. Le principe de cet algorithme est de trouver une approximation quadratique majorante $\mathcal{K}(\cdot, \boldsymbol{t})$ simple à minimiser en introduisant des variables auxiliaires $\boldsymbol{t}$.

Pour développer cet algorithme, on peut écrire le critère simplifié (8) sous forme matricielle en notant le vecteur des paramètres $\boldsymbol{x} = [\boldsymbol{g}^T, \boldsymbol{o}^T]^T$:

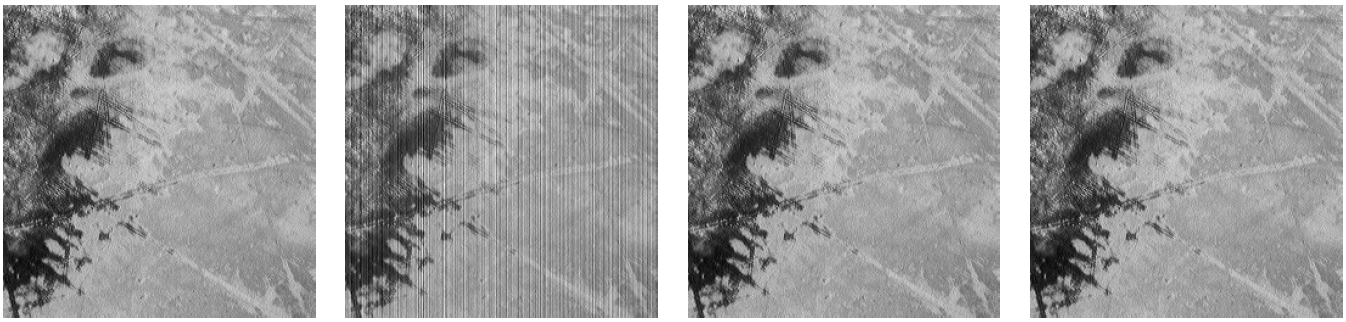
$$K(\boldsymbol{x}) = \frac{1}{T} \sum_{r,c} \phi([\mathbf{V}_r^T \boldsymbol{x}]_c) + \boldsymbol{x}^T \mathbf{Q} \boldsymbol{x} - 2\lambda_g \boldsymbol{e}^T \boldsymbol{x} \quad (9)$$

avec

$$\mathbf{V}_r = \begin{bmatrix} \mathbf{V}_r^1 \\ \mathbf{V}^2 \end{bmatrix}, \quad \boldsymbol{e} = \begin{bmatrix} \mathbb{1}_{[C\times 1]} \\ \mathbb{0}_{[C\times 1]} \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} \lambda_g I_C & 0 \\ 0 & \lambda_o I_C \end{bmatrix},$$
$$\mathbf{V}_r^1 = \mathrm{diag}_0\{w_{r,1}, \ldots, w_{r_C-1}\} - \mathrm{diag}_1\{w_{r,2}, \ldots, w_{r_C}\},$$
$$\mathbf{V}^2 = \mathrm{diag}_0\{1, \ldots, 1\} - \mathrm{diag}_1\{1, \ldots, 1\},$$

$\mathrm{diag}_k\{\boldsymbol{u}\}$ désignant la matrice dont la k-ième diagonale contient les éléments du vecteur $\mathbf{u}$, le reste de ses éléments étant nuls.

L'algorithme MM proposé minimise le critère simplifié (8) sous la contrainte $\sum_c g_c = C$ en procédant à chaque itération



(a) Image Idéale (b) Image Observée (c) Image corrigée pour ϕ^{2-1} (d) Image corrigée pour ϕ^{2-0}

FIGURE 1 – Résultats de correction de l'effet de rayure sur une image de type PLEIADES de taille 4000×2000. Extrait de taille 200×200 de l'image originale, de l'image observée avec rayures, et des images corrigées par la méthode proposée avec l'algorithme MM pour les deux fonctions ϕ.

en deux étapes :

$$\textit{Majorize :}\ t_{r,c}^{(k+1)} = \frac{\phi'([\mathbf{V}_r^T \boldsymbol{x}^{(k)}]_c)}{2[\mathbf{V}_r^T \boldsymbol{x}^{(k)}]_c},$$
$$\text{avec} \quad \mathbf{L}_r^{(k+1)} = \operatorname{diag}\{t_{r,c}^{(k+1)}\}_{c=1,\ldots,C}$$

$$\textit{Minimize :}\ \boldsymbol{x}^{(k+1)} = \frac{C(\mathbf{B}^{(k+1)})^{-1}\boldsymbol{e}}{\boldsymbol{e}^T(\mathbf{B}^{(k+1)})^{-1}\boldsymbol{e}}$$
$$\text{avec} \quad \mathbf{B}^{(k+1)} = \mathbf{Q} + \frac{1}{T}\sum_r \mathbf{V}_r \mathbf{L}_r^{(k+1)} \mathbf{V}_r^T.$$

L'approximation majorante étant quadratique, l'étape de minimisation se fait directement en résolvant un système linéaire. De plus, cette étape se fait ici à faible coût grâce à la structure de la matrice $\mathbf{B}^{(k+1)}$ qui est tridiagonale par bloc. On peut en pratique constater, tel qu'illustré Tab. 1, que l'algorithme MM proposé est beaucoup plus rapide que les algorithmes de type gradient conjugué présentés précédemment.

5 Résultats et Conclusion

De nombreux tests sur des données simulées nous ont permis d'analyser le comportement de l'estimateur et des algorithmes proposés et d'en tirer quelques conclusions. D'une part, la simplification du critère (7) en (8) ne dégrade que rarement les résultats de correction et les améliore même parfois, comme sur l'exemple du Tab. 2. Malgré la non convexité du critère et donc la possibilité de minima locaux, les résultats obtenus pour la fonction non convexe ϕ^{2-0} sont toujours meilleurs que ceux de la fonction convexe ϕ^{2-1} (comme sur l'exemple de Tab. 2). Le temps de calcul est toujours plus faible avec l'algorithme MM qu'avec les algorithmes GC et PGC. Il est généralement légèrement supérieur pour la fonction ϕ^{2-0} que pour la fonction ϕ^{2-1}. Lorsque l'ordre de grandeur des gains et offsets (σ_g^2 et σ_o^2) sont connus, il est relativement aisé de régler les hyper paramètres T et s pour une image donnée en prenant en compte sa dynamique.

A titre illustratif, la Fig. 1 illustre un résultat de correction d'une image via l'estimation des paramètres par l'algorithme MM, pour les deux fonctions de potentiel (la différence de correction est ici peu visible).

Références

[1] B. Cao, Y. Du, D. Xu, H. Li, and Q. Liu. An improved histogram matching algorithm for the removal of striping noise in optical remote sensing imagery. *Optik*, 126(23) :4723–4730, Dec. 2015.

[2] H. Carfantan and J. Idier. Statistical linear destriping of satellite-based pushbroom-type images. *IEEE Trans. Geosci. Remote Sensing*, 48(4) :1860–1871, Apr. 2010.

[3] Y. Chang, L. Yan, T. Wu, and S. Zhong. Remote sensing image stripe noise removal : From image decomposition perspective. *IEEE Trans. Geosci. Remote Sensing*, 54(12) :7018–7031, Dec. 2016.

[4] P. Charbonnier, L. Blanc-Feraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.*, 6(2) :298–311, Feb. 1997.

[5] E. Chouzenoux, S. Moussaoui, and J. Idier. A majorize-minimize line search algorithm for barrier function optimization. In *Proc. EUSIPCO*, Glasgow, UK, Aug. 2009.

[6] F. L. Gadallah, F. Csillag, and E. J. M. Smith. Destriping multisensor imagery with moment matching. *Int. J. Remote Sens.*, 21(12) :2505–2511, Jan. 2000.

[7] B. Munch, P. Trtik, F. Marone, and M. Stampanoni. Stripe and ring artifact removal with combined wavelet-Fourier filtering. *Optics Express*, 17(10) :8567, May 2009.

[8] J.-J. Pan. Destriping of landsat MSS images by filtering techniques. *Photogram. Eng. Remote Sens.*, 58(10) :1417–123, Oct. 1992.

[9] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment : from error visibility to structural similarity. *IEEE Trans. Image Process.*, 13(4) :600–612, Apr. 2004.

[10] B. Zhao, B. He, and Y. Cong. Destriping method using lifting wavelet transform of remote sensing image. In *Proc. Int. Conf. Comput. Mechatron. Control Electron. Eng.*, Changchun, China, Aug. 2010.