

# Fusion et Stéréoscopie

JEAN-MARC DELVIT<sup>1</sup>, STEPHANIE ARTIGUES<sup>2</sup>

<sup>1</sup> CNES

18 avenue Edouard Belin, 31404 Toulouse Cedex 9, France

<sup>2</sup> Magellium

24 rue Hermès, BP 12113, 31521 Ramonville Saint-Agne, France

<sup>1</sup>jean-marc.delvit@cnes,

<sup>2</sup>stephanie.artigues@magellium.fr

**Résumé** – Nous proposons dans cet article une méthode permettant de coloriser une image panchromatique à l'aide d'une image multispectrale acquise dans des conditions de prise de vue différentes, sans connaissance a priori du relief. La première étape consiste à simuler l'image multispectrale qui aurait été acquise par la rétine panchromatique en recalant l'image multispectrale sur l'image panchromatique. Cette étape nécessite la connaissance des modèles géométriques des produits ainsi que des modèles numériques de terrain que nous proposons de générer directement à partir des produits. La seconde étape consiste à fusionner les images recalées en utilisant une méthode pan-sharpen qui peut intégrer le rapport des Fonctions de Transfert de Modulation (FTM).

**Abstract** – This paper addresses the issue of panchromatic and multispectral image fusion when data are not acquired in the same conditions. The first stage consists in simulating multispectral image seen by the panchromatic retina by registering images. This step requires the knowledge of products geometric models and a digital terrain model, we propose to deduce it from products. The second stage consists in using pansharpen method integrating Modulation Transfer Functions (MTF) for multispectral and panchromatic images fusion.

## 1 Introduction

Classiquement les images panchromatiques et multispectrales utilisées pour les applications de fusion sont acquises simultanément sous des conditions de prise de vue très proches. Dans ce papier, nous développons une approche permettant de fusionner une image panchromatique avec une image multispectrale acquise dans des conditions de prise de vue différentes, dans un contexte très haute résolution.

Chacune des images comporte des distorsions liées à la géométrie et au relief de la scène imagée. Dans un contexte Spot 5 par exemple, en considérant un angle stéréoscopique (B/H) de 0.018 radians entre les acquisitions panchromatique (PAN) et multispectrales (XS), une méconnaissance du terrain de 10 mètres engendre un effet de parallaxe entre les images de  $10 \times 0.018 = 18$  cm soit 0.07 pixel PAN (à la résolution de 2.5 mètres) [1], [7]. Les décalages entre les images ne sont pas perceptibles à l'œil, le seuil de sensibilité étant admis autour de 0.5 pixel. Ce même angle stéréoscopique crée des effets bien plus importants dès lors que l'on traite des produits très haute résolution. Par exemple à 10 cm, le décalage induit par une méconnaissance du relief de 10 mètres pour le même B/H est de 1.8 pixel. Ces décalages entre les images panchromatiques et multispectrales engendrent des glissements de couleur visibles au niveau du produit fusionné, ils sont particulièrement gênant dans le contexte d'images très résolues en milieu urbain.

La méthode développée dans ce papier consiste à corriger les décalages entre les images sans connaissance a priori du relief, mais en exploitant directement les produits images et les métadonnées qui leur sont associées. L'intérêt de cette méthode par rapport à un recalage via un modèle numérique d'élévation (MNE) externe est double. Contrairement à un MNE externe, représentatif d'une scène à un instant  $t$ , le MNE endogène déduit des images de la scène traduit tous les éléments effectivement présents dans la scène (véhicules stationnés), il est cohérent avec les images (pas de sensibilité aux changements entre la date de production du MNE externe et celle des acquisitions). En outre la méthode permet de générer un MNE très résolu et assure une parfaite superposabilité aux images.

## 2 Méthode

La chaîne de traitement comporte deux étapes. La première consiste à recalibrer les images sans connaissance a priori du relief. En effet, les images étant acquises sous des conditions de prise de vue différentes, elles offrent une vision stéréoscopique de la scène qui permet de restituer le relief directement à partir des images et de leurs métadonnées. Pour optimiser le temps de traitement, les images sont projetées en géométrie épipolaire. Une des propriétés de cette géométrie est qu'un point de la scène est imagé sur la même ligne dans les deux images. Les décalages entre les images, liés à la méconnaissance du relief, ont ainsi une seule

composante suivant une direction (colonne). Le calcul de la grille de disparité entre les deux images, permettant de rectifier les effets liés au relief est ainsi accéléré.

Les images panchromatiques et multispectrales ainsi recalées (au facteur de résolution près), sont ensuite fusionnées. Nous utilisons pour cette étape une méthode de pan-sharpening pouvant intégrer le rapport des Fonctions de Transfert de Modulation (FTM) des instruments panchromatique et multispectral.

## 2.1 Recalage

La méthode de calcul des décalages utilisée dans cet article met en place une stratégie de corrélation multi-échelle avec filtrage, particulièrement adaptée à des images très résolues en milieu urbain [2]. Deux filtres sont mis en place, le premier modélise les structures de type bâti en reconstruisant les données jugées peu fiables de la nappe de disparité par des plans. Le second filtre consiste à supprimer les artefacts résiduels de la nappe de disparités en rebouchant les trous en moyennant les valeurs du voisinage.

Les nappes de disparité générées sont de bonne qualité, denses et sans artefacts. Les grandes phases de l'algorithme sont :

- Préparation des données
  - Dézoom de l'image panchromatique à la résolution de l'image multispectrale
  - Génération d'une image panchromatique de synthèse par combinaison linéaire des bandes multispectrales
  - Projection des images en géométrie épipolaire en exploitant les modèles géométriques
  - Génération de la pyramide d'image (dézoom)
- Calcul de la carte de disparité, suivant un processus multi-résolution
  - Corrélation dense des images
  - Filtrage des décalages en fonction de l'écart-type des images et reconstruction par des plans
  - Filtrage des artefacts de la nappe de disparité et reconstruction des trous par moyennage
  - Mise à l'échelle et zoom des disparités. Cette nappe sert de prédicteur à l'échelle suivante
- Conversion de la carte de disparité en grille de rééchantillonnage et application de la grille à l'image XS

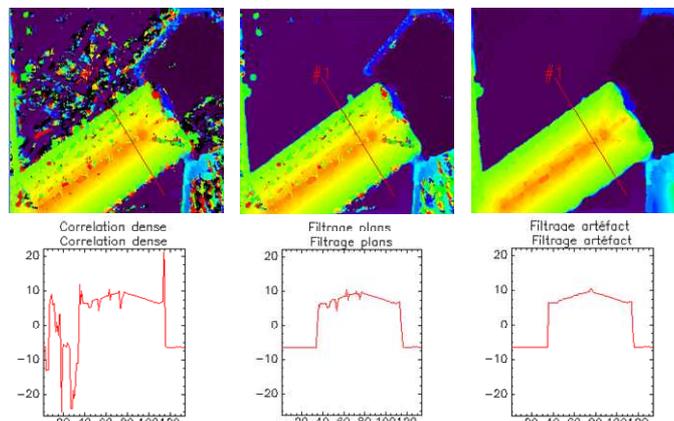


Figure 1 : résultats des opérations de filtrage – cartes et profils

## 2.2 Fusion

L'étape de fusion P+XS consiste à mixer les informations contenues dans l'image panchromatique avec celles contenues dans l'image multispectrale de sorte à obtenir un produit tirant profit des deux sources d'images : la résolution (et les détails) de l'image panchromatique et la couleur de l'image multispectrale.

La méthode de pansharpening utilisée dans ce papier se déroule en trois étapes [3]. L'image multispectrale est tout d'abord zoomée à la résolution de l'image panchromatique tandis que l'image panchromatique est convoluée par un filtre passe-bas, qui peut être défini par le rapport des FTM PAN et XS, afin de rendre son contenu fréquentiel équivalent à celui de l'image multispectrale. Le rapport des deux images permet de moduler l'information contenue dans l'image multispectrale. Il est ensuite multiplié par l'image panchromatique pour obtenir le produit fusionné [8] :

$$Fusion = \frac{(XS)_{recalé\_zoom}}{(PAN)_{mou}} \cdot PAN \quad (1)$$

## 2.3 Méthode d'évaluation des fusions

L'évaluation est réalisée de façon visuelle dans un premier temps puis en calculant un critère de ressemblance entre les produits fusionnés avec le produit de référence [4] [5]. Le critère de ressemblance utilisé est le SSIM (Structural Similarity Index Map) [6]. Il se présente sous forme de carte dont les valeurs au pixel (x,y) indiquent le degré de ressemblance (-1 le moins ressemblant à 1 le plus ressemblant) entre les images I et J sur le voisinage des pixels considérés :

$$SSIM_{I,J}(x,y) = \frac{(2\mu_I(x,y)\mu_J(x,y) + c_1) \cdot (2\sigma_{I,J}(x,y) + c_2)}{(\mu_I(x,y)^2 + \mu_J(x,y)^2 + c_1) \cdot (\sigma_I(x,y)^2 + \sigma_J(x,y)^2 + c_2)} \quad (2)$$

avec  $\mu_I(x,y)$  et  $\sigma_I(x,y)$  respectivement la moyenne et la variance de l'image I au voisinage du point de coordonnées (x,y),  $\sigma_{I,J}(x,y)$  la covariance des images I et J au voisinage du point de coordonnées (x,y).  $c_1$  et  $c_2$  des constantes :  $c_1 = (k_1 \cdot L)^2$  et

$c_2 = (k_2 \cdot L)^2$  avec  $L$  l'étalement dynamique des images (par exemple  $L=255$  pour une image codée sur 8 bits) et  $k_1$  et  $k_2$  des constantes respectivement fixées à 0.01 et 0.03.

La moyenne de cette carte, notée MSSIM, fournit un critère de ressemblance global entre les images.

### 3 Résultats

Nous illustrons les résultats obtenus avec la méthode sur des couples stéréoscopiques aéroportés acquis à la résolution de 10 cm. Le système est composé de deux caméras imageant simultanément la scène sous des angles de prise de vue différents, dans les quatre bandes spectrales rouge, vert, bleu et proche infrarouge. Nous considérons les images panchromatiques acquises par la première caméra et la seconde caméra ( $P_{Ref}$  et  $P_{Sec}$  simulées par combinaisons linéaires des bandes multispectrales), les images multispectrales acquises respectivement par la première et la seconde caméra dézoomées d'un facteur 4 ( $XS_{Ref}$  et  $XS_{Sec}$ ). L'angle stéréoscopique entre les images (Ref et Sec) est de 0.12 radians. Nous comparons les produits suivants :

- $P_{Ref}/XS_{Ref}$  : référence
- $P_{Ref}/XS_{Sec}$  : produit fusionné avec recalage sur MNT externe (cas d'un MNT plat) – Cas 1
- $P_{Ref}/(XS_{Sec})_{recalc}$  : produit fusionné en appliquant la méthode présentée (la carte de disparité est calculée à partir des deux bandes panchromatiques  $P_{Ref}$  et  $P_{Sec}$  à 10 cm) avec un pas de 2 – Cas 2

Les résultats de fusion sont présentés sur la Figure 2. Les cartes de ressemblance entre la référence et chacun des deux produits fusionnés font apparaître une nette amélioration de la qualité liée à l'utilisation de la méthode de recalage. La cas 1 fait apparaître des glissements de couleurs importants, induits par les décalages entre les images panchromatiques et multispectrales, proportionnels à l'erreur altimétrique. Ces glissements ne sont pas perceptibles sur le produit fusionné obtenu en appliquant notre méthode, qui est visuellement très proche de la référence. Le critère de ressemblance global entre les produits est de 0.97 pour le cas 2 contre 0.87 pour le cas 1.

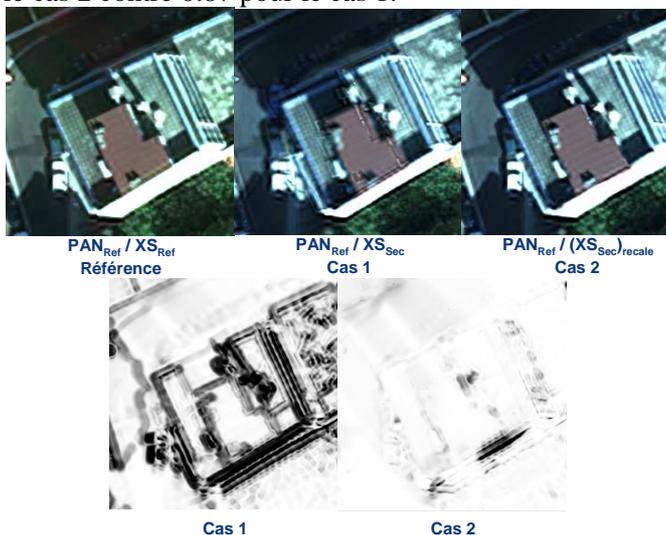


Figure 2 : extraits des produits fusionnés de référence et des cas 1 et 2 et Cartes de ressemblance des produits fusionnés avec la référence

### 3.1 Influence de la précision du MNT

Nous comparons dans cette partie les produits fusionnés  $P_{Ref}/(XS_{Sec})_{recalc}$  générés à partir de trois MNE plus ou moins précis : calculés à partir des deux bandes panchromatiques  $P_{Ref}$  et  $P_{Sec}$  à 10 cm) avec des pas de 2, 10 et 25 pixels correspondant respectivement aux cas 2, 3 et 4. Les différents MNE sont présentés sur la figure 3, moins les MNE sont précis plus le temps de génération est réduit.

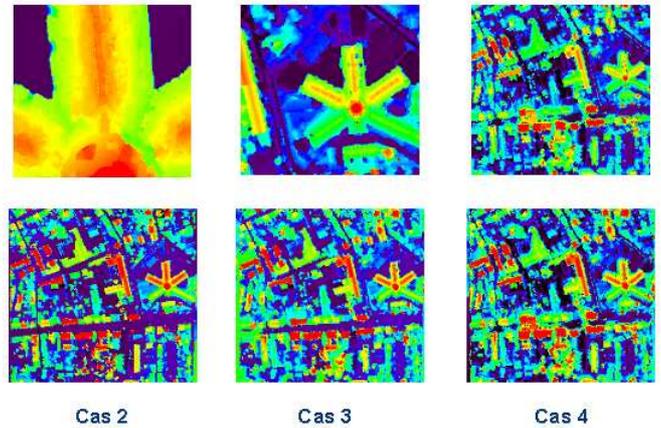


Figure 3 : MNE générés à partir des images en géométrie épipolaire pour les pas de 2, 10 et 25 (cas 2, 3 et 4) – en haut à pleine échelle – en bas vue d'ensemble des MNE.

Les résultats présentés sur la figure 4 illustrent l'impact de la qualité du MNE sur la qualité du produit fusionné. On constate que tous les défauts du MNE présents sur les bords du bâtiment sont visibles sur le produit fusionné et engendrent une déformation des objets. Il est donc important d'avoir un MNE de bonne qualité pour effectuer ce type de colorisation et donc des algorithmes de génération de MNE performants. On remarque en outre sur les cartes de ressemblance entre les différents produits fusionnés et le produit de référence apparaissent au niveau des façades des bâtiments et ceci en raison des conditions de prises de vue différentes des images panchromatiques et multispectrales : les façades vues sur une image sont occultés sur la seconde et vice versa.

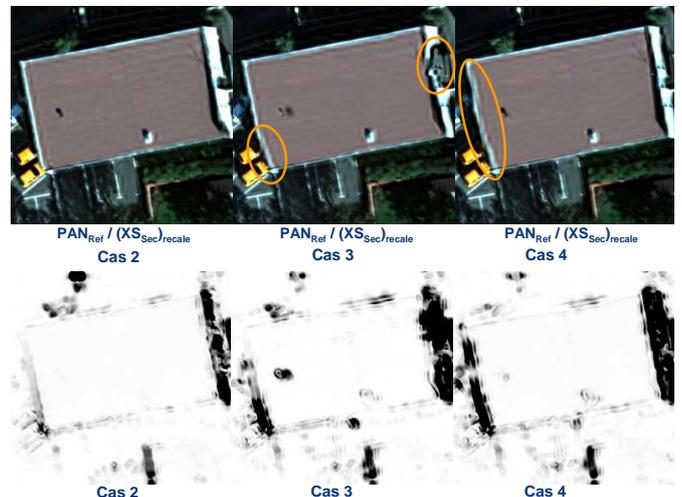


Figure 4 : extraits des produits fusionnés des cas 2, 3 et 4 et Cartes de ressemblance des produits fusionnés avec la référence

Le critère de ressemblance des produits fusionnés avec le produit de référence est calculé sur l'ensemble de la scène pour les différents cas traités. Les scores obtenus sont présentés dans le tableau 1. On constate que plus le MNE est précis plus la ressemblance avec le produit fusionné élevée, mais que pour un pas de 10 (cas 3) qui permet de réduire le temps de calcul du MNE, la dégradation des performances de la fusion est faible ; un compromis peut être trouvé.

**Tab 1 : Valeur moyenne du critère de ressemblance entre les produits fusionnés  $P_{Ref}/(XS_{Sec})_{recalc}$  des différents cas et le produit de référence**

Cas	1	2	3	4
MSSIM	0.8703	0.9672	0.9618	0.9587

## 4 Conclusion

La méthode développée dans ce papier permet de coloriser une image panchromatique avec une image multispectrale acquise dans des conditions de prise de vue différentes, sans connaissance a priori du relief, grâce à un algorithme de recalage performant, adapté à la très haute résolution en milieu urbain.

Les différentes sources d'erreur pouvant être identifiées sur le produit fusionné sont :

- Oclusions : les configurations des prises de vue des produits à fusionner peuvent être telles que des façades de bâtiments visibles sur l'image panchromatique ne sont pas accessibles sur l'image multispectrale. Ces erreurs ne peuvent pas être corrigées, mais ne semblent pas gênantes pour l'interprétation du produit fusionné.

- Objets mobiles : dans le cas où les images multispectrales et panchromatiques ne sont pas acquises simultanément, les objets en mouvement dans la scène ne sont pas imagés aux mêmes positions. Il en résulte des artefacts sur le produit fusionné. Ces erreurs sont inhérentes aux données traitées et ne peuvent pas être corrigées.

- Sursol : les décalages entre les images panchromatiques et multispectrales projetées en géométrie épipolaire sont proportionnels aux erreurs de sursol. Nous avons ainsi montré que la prise en compte du sursol (véhicule, végétation, etc...) dans le processus de fusion permet de réduire la perception des glissements de couleur et d'améliorer ainsi le critère de ressemblance avec le produit de référence. Tout l'intérêt de générer le MNE directement à partir du couple stéréo panchromatique acquis réside dans la capacité de restituer le sursol au niveau des éléments visibles sur la scène (par exemple au niveau des véhicules présents sur l'image pour lesquels un MNE externe ne rendrait pas compte de l'altitude). L'enjeu est alors d'améliorer la précision et la robustesse des algorithmes de restitution automatique d'altitude.



**Figure 5 : exemple sur des véhicules : la référence, le cas 1 sans prise en compte du sursol, le cas 2 en considérant un MNE endogène.**

En outre, la méthode permet la génération simultanée du produit fusionné PAN/XS et du modèle numérique d'élévation de la zone imagée. Elle permet d'imaginer des systèmes d'observation de la terre permettant de répondre simultanément aux deux besoins : restitution 3D et imagerie couleur.

## 5 Références

- [1] C. Latry, « La chaîne de fusion de données SPOT5 ». Bulletin de la SFPT n°169 (2003-1), pp. 46-52
- [2] J.-M. Delvit et S. Artigues « Automatic DEM generation from low B/H stereoscopic acquisition ». Proc. of SPIE Vol. 7831, 78310J, Sep. 2010
- [3] E. Christophe, J. Inglada et A. Giros, « Orfeo ToolBox: a complete solution for mapping from high resolution satellite images ». The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B4. Beijing 2008
- [4] F. Laporterie-Déjean, H. de Boissezon, G. Flouzat et M.-J. Lefèvre-Fonollosa, « Thematic and statistical evaluation of five panchromatic/multispectral fusion methods on simulated PLEIADES-HR images ». Inf. Fusion, vol.6, no. 3, pp. 193-212, Sep. 2005
- [5] L. Alparone, L. Wald, J. Chanussot, Senior Member, IEEE, C. Thomas, P. Gamba, Senior Member, IEEE et L. Mann Bruce, Senior Member, IEEE, « Comparison of Pansharpening Algorithms: Outcome of the 2006 GRS-S Data-Fusion Contest ». IEEE Transactions on Geoscience and Remote Sensing, vol. 45, no. 10, Oct 2007
- [6] Z. Wang, A. C. Bovik, H. R. Sheikh et E. P. Simoncelli « Image quality assessment: From error visibility to structural similarity ». IEEE Transactions on Image Processing, vol. 13 no. 4, pp. 600-612, Avril 2004
- [7] S. May, C. Latry « Digital Elevation Model Computation with SPOT5 Panchromatic and Multispectral images using low stereoscopic angle and geometric model refinement », Proceedings of IGARSS, 2009
- [8] P. Kubik « PLEIADES image quality : from users' needs to products definition », Proceedings of SPIE, 2005