New Developments and Challenges in Remote Sensing, Z. Bochenek (ed.)
©2007 Millpress, Rotterdam, ISBN 978-90-5966-053-3

MS + Pan image fusion by an enhanced Gram-Schmidt spectral sharpening

B. Aiazzi, S. Baronti & M. Selva
Institute of Applied Physics "Nello Carrara" of the National Research Council IFAC-CNR, Florence, Italy

L. Alparone
Department of Electronics & Telecommunications, University of Florence, Florence, Italy

Keywords: component substitution, Gram-Schmidt spectral sharpening, IKONOS satellite data, image fusion, multi-spectral imagery, multivariate regression

ABSTRACT: In this work, a simple pre-processing patch is introduced before the Gram-Schmidt (GS) spectral sharpening method (as implemented in ENVI) such that the resulting fused multi-spectral (MS) data exhibit higher sharpness and spectral quality. This is achieved by defining a generalized intensity (GI) component as a weighted average of the MS bands, with weights taken either as percentages of overlap between the spectral responses of individual bands and the spectral response of panchromatic (Pan), or better as regression coefficients between the MS bands and the decimated Pan image. In the former case, the weights are pre-calculated for each sensor. In the latter case, the weights are calculated by applying a multivariate regression to the data that are being fused. The above GI component is used as low-resolution approximation of the Pan image. Experimental results carried out on very-high resolution IKONOS data demonstrate that the proposed enhanced GS adaptive (GSA) method visually outperforms both modes of the ENVI implementation of GS, especially in true colour displays. Quantitative scores performed on spatially degraded data by means of such parameters as Wald's ERGAS and the novel Q4 score index based on quaternion theory, confirm the superiority of the enhanced GS method over its baseline.

1 INTRODUCTION

Remote-sensing image fusion techniques aim at integrating the information conveyed by data acquired with different spatial and spectral resolutions from satellite or aerial platforms. The most straightforward goal is photo-analysis, but also such automated tasks as features extraction and segmentation/classification have been found to benefit from fusion. A variety of image fusion techniques is devoted to merge multi-spectral (MS) and panchromatic (Pan) images, which exhibit complementary characteristics of spatial and spectral resolutions (Aiazzi *et al*. 2002; Wang *et al*. 2005; Garzelli and Nencini 2005; Garzelli and Nencini 2006).

When exactly three MS bands are concerned, the simplest fusion method is to resort to the Intensity-Hue-Saturation (IHS) transformation. This procedure is equivalent to inject, i.e. add, the difference between the sharp Pan and the smooth intensity into the re-sampled MS bands (Tu *et al*. 2001). Since the histogram-matched Pan and the intensity component I do not generally have the same radiometry, i.e. local mean, when the fusion product is displayed in colour composition, large spectral distortion, i.e. colour changes, may be noticed. This occurs because the spectral response of *I*, as synthesised by means of the MS bands, may be far different from that of Pan. Thus, also radiance offsets, slowly space-varying, and not only spatial details, are locally injected. When more than three spectral bands are available, IHS fusion may be applied to three consecutive spectral components at a time, or better the IHS transformation may be replaced with principal component analysis (PCA). The latter does not avoid spectral distortion, even if it may be less noticeable. Generally speaking, if the spectral responses of the MS bands are not perfectly overlapped with the bandwidth of Pan, as it happens with the most advanced very-high resolution imagers, IKONOS and QuickBird, IHS- and PCA-based methods may yield poor results in terms of spectral fidelity (Zhang 2004).

One of the most widespread and performing MS + Pan fusion method is the Gram-Schmidt (GS) spectral sharpening, invented by Laben and Brover in 1998 and patented by Eastman Kodak (Laben and Brower 2000). It relies on the component substitution strategy and is widely used, being implemented in the ENVI package. The GS method has two operational modes, depending on how the low-resolution version of the Pan image used in the forward GS transformation is defined. In one case, let us call it "mode 1", it is obtained as the pixel average of the MS bands, which are given as input to the procedure. In the other case, say "mode 2", the approximation is preliminarily obtained by low-pass filtering and decimating the Pan image and is directly given as input to the procedure. The major difference in results, mostly noticeable in true colour display, is that mode 1 exhibits outstanding spatial quality, but spectral distortions may occur, because the average of the MS spectral bands is likely not to have the same radiometry as the Pan image. Instead, mode 2 is unaffected by spectral distortion but generally suffers from a lower sharpness and spatial enhancement in general. This is due to the injection mechanism of high-pass details taken from Pan, which is embedded in the inverse GS transformation, carried out by using the full resolution Pan, while the forward transformation uses the low resolution approximation of Pan obtained by resampling the decimated Pan image provided by the user.

GS spectral sharpening will be briefly reported in the next section. Afterwards, the proposed enhancement is motivated and described. Eventually, results will be shown and some conclusions will be drawn.

2 GRAM-SCHMIDT SPECTRAL SHARPENING

In the GS method, as described by its inventors Laben and Brower (2000), the spatial resolution of the MS image is enhanced by merging the high resolution Pan image with the low spatial resolution MS bands. According to the authors' description the main steps of the methods are the following:

1. A lower spatial resolution Pan image is simulated.
2. The Gram-Schmidt transformation is performed on the simulated lower spatial resolution Pan image and the plurality of lower spatial resolution spectral band images. The simulated lower spatial resolution Pan image is employed as the first band in the Gram-Schmidt transformation.

3. The statistics of the higher spatial resolution Pan image is adjusted to match the statistics of the first transform band resulting from the Gram-Schmidt transformation to produce a modified higher spatial resolution Pan image.
4. The modified higher spatial resolution Pan image is substituted for the first transform band resulting from the Gram-Schmidt transformation to produce a new set of transformed bands.
5. The inverse Gram-Schmidt transformation is performed on the new set of transform bands to produce the enhanced spatial resolution MS image.

Some freedom exists in step 1) and determines the two options of the ENVI package described in the introduction as "mode 1" and "mode 2".

In "mode 1" the procedure is responsible to compute the simulated low-resolution Pan image that is obtained as the pixel average of the MS bands. The image that is produced in the fusion process is spatially sharp but some changes in colours are noticeable with respect the re-sampled MS image.

In "mode 2", it is responsibility of the user to produce the synthetic Pan image that is usually obtained by preliminarily low-pass filtering and then decimating the Pan image. Concerning colours, the resulting fused image appears very similar to the expanded MS image. However, spatial details that are injected are not so sharp as in the case of "mode 1", due to *improper* digital filtering of the Pan image (Aiazzi *et al*. 2006). This effect is noticeable for small objects appearing on a quasi-constant background.

3 ENHANCED GRAM-SCHMIDT IMAGE FUSION

To understand the influence of the spectral response on the fusion of IKONOS images, the relative spectral responses plotted in Fig. 1 are to be considered. Ideally, the MS bands (B, G, R and NIR) should be disjoint and should entirely fall within the bandwidth

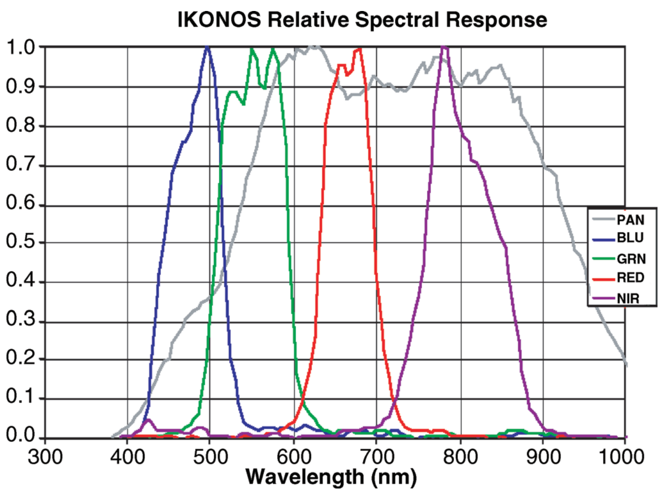

Figure 1. IKONOS spectral responses.

of Pan. From Fig. 1, however, it appears that the G and B bands are largely overlapped and that the B band mostly falls outside the 3 dB cutoff of the Pan band. Furthermore, the spectral response of Pan is extended beyond the NIR band. The colour distortion problem in fusion stems from such mismatches and in particular from the fact that the synthetic Panchromatic in "mode 1" does not reflect the spectral response of Fig 1, since it is obtained as plain average of the B, G, R and NIR bands. As an example, vegetation appears of relatively high reflectance in the NIR and Pan bands, while its reflectance is low in the visible (RGB) bands. Since all bands are equally weighted, the radiometric values of vegetation areas is likely to be much smaller in the synthetic Pan than in the true Pan. This effect causes the injection in the fused MS bands of a radiance offset, which may give rise to colour distortion, mostly noticeable in vegetated regions of the fused image.

In order to avoid this drawback, the idea is to generate the synthetic Pan in such a way that the spectral response of the sensor (IKONOS, in the present case) is considered, by differently weighting the contributions coming from the different spectral channels (Alparone *et al*. 2004b). The preprocessing steps to be performed before the standard GS spectral sharpening are the following:

1. Reduce the original full resolution Pan image to the size of the MS images by low-pass filtering and decimation. Let denote this image as $P_{\downarrow}$.
2. Assume a model according to which, at every pixel position it holds that:

$$P_{\downarrow} = \alpha \cdot B + \beta \cdot G + \gamma \cdot R + \delta \cdot NIR + \varepsilon \quad (1)$$

$= \sum \alpha_i HS + \dot{\varepsilon}$

   where $\alpha, \beta, \gamma, \delta$ and ε are scalar constants.

3. Find $\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}$ and $\hat{\varepsilon}$ that determine an estimate $\hat{P}_{\downarrow}$ of $P_{\downarrow}$ such that the mean square value of the residue $\|P_{\downarrow} - \hat{P}_{\downarrow}\|$ is minimized.
4. Take $\hat{P}_{\downarrow}$ as the simulated Pan image to be given as input to the GS procedure in mode 2.

The implicit assumption is that the regression coefficients computed at the scale of the MS image are equal to those that would be computed at the scale of the original Pan image, if R, G, B and NIR observations were available at the same resolution.

In the next section, the proposed method will be assessed through a comparison with the two operational modes of the ENVI implementation of GS spectral sharpening. Also the fixed weighting coefficients, proposed by Tu *et al.* (2004) to improve the IHS method, i.e. $\alpha = 1/12$, $\beta = 1/4$, $\delta = 1/3$, $\gamma = 1/3$ and $\varepsilon = 0$, have been implemented and the results compared to those found with the proposed regression approach.

4 EXPERIMENTAL RESULTS

Fig. 2 shows a 512×512 detail of the whole 2048×2048 IKONOS image that has been processed. The full-resolution Pan image is reported in Fig. 2(a) while the original true-colour MS image, expanded to the scale of Pan is reported in Fig. 2(b). Fig. 2(c) shows the result of "mode 1", i.e. the synthetic low-resolution Pan is obtained as pixel average of the four MS bands. The spatial enhancement is impressive; however, some colour

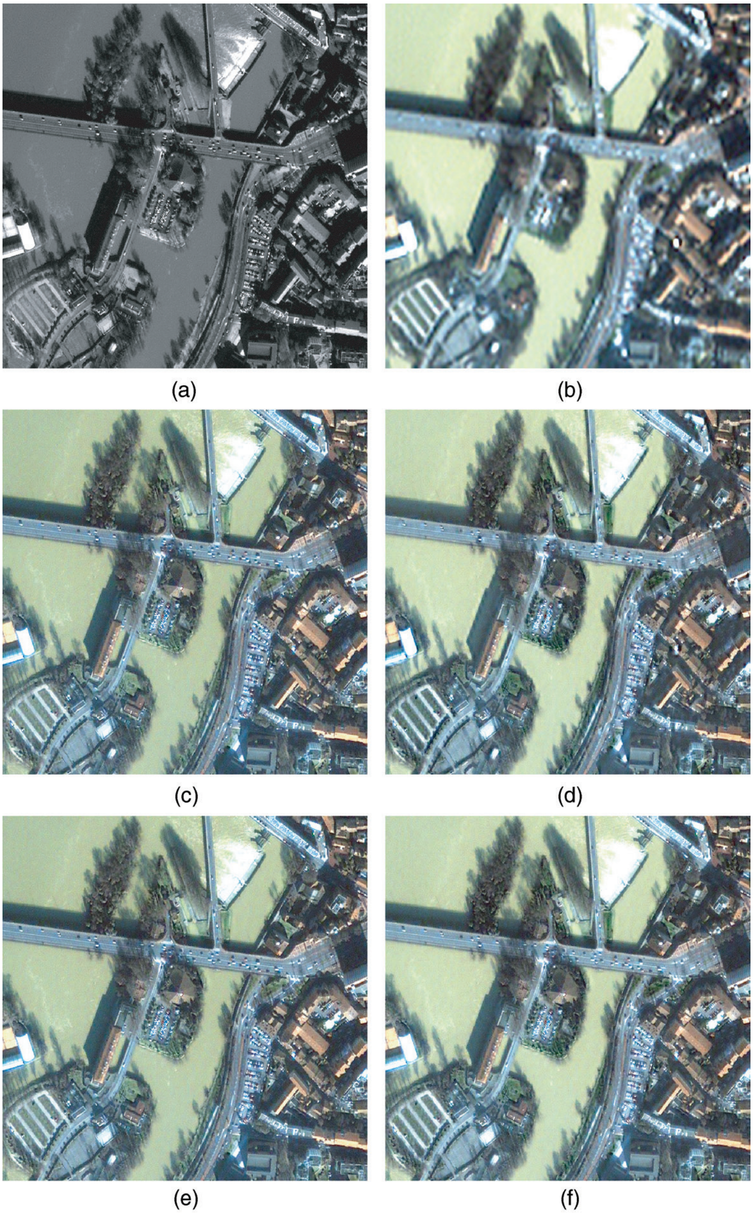

Figure 2. Examples of full-scale spatial enhancement of fusion algorithms displayed as 512 × 512 true colour compositions at 1 m pixels spacing for the IKONOS image. (a): original Pan image; (b): original MS bands (4 m) resampled to the scale of Pan image (1 m); (c): GS ''mode 1'' fusion (GS1); (d): GS ''mode 2'' (GS2); (e): GS with fixed weights (GSF); (f): GS with adaptive weights (GSA).

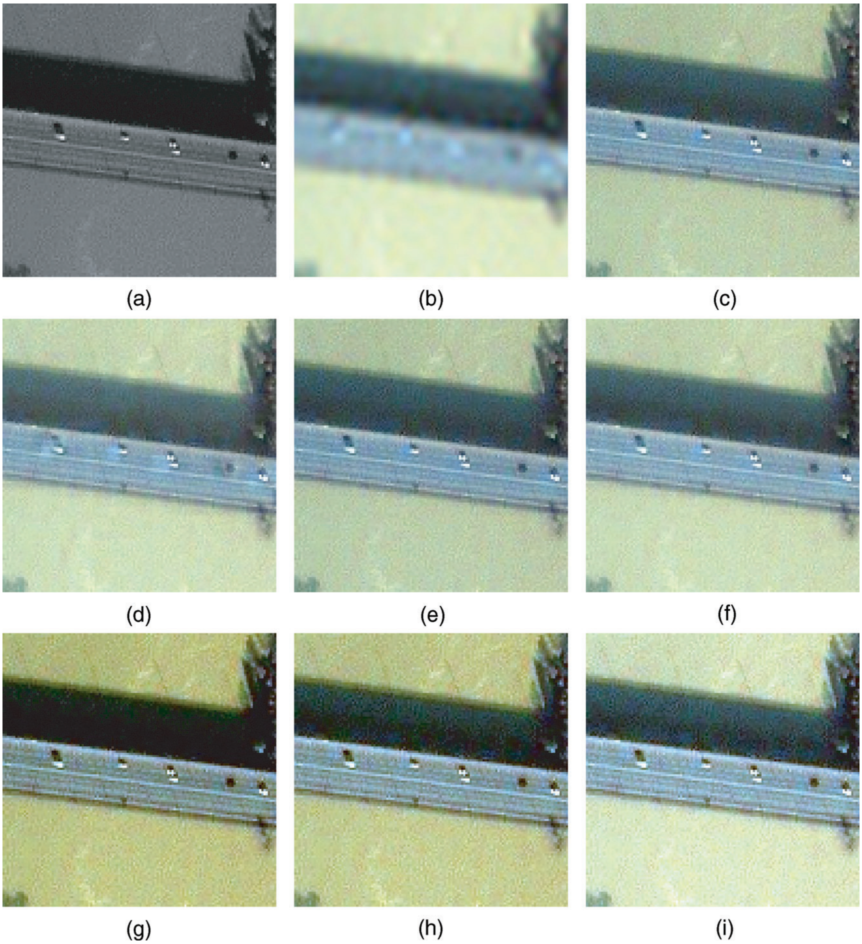

Figure 3. Results of fusion algorithms displayed as 128 $\times$ 128 true colour compositions at 1 m pixel spacing. (a): original Pan image; (b): original MS bands (4 m) re-sampled to the scale of Pan image (1 m); (c): GS "mode 1" fusion; (d): GS "mode 2" fusion; (e): GS with fixed weights fusion; (f): GS with adaptive weights; (g): generalised IHS with equal weights; (h): generalised IHS with unequal fixed weights (IHF); (i): generalised IHS with adaptive weights (IHA).

distortion appears on the green areas on the left of the river and in the river itself. The green areas are too bright while the colour of the river is too dark. These spectral distortion are missing in Fig. 2(d) showing the outcome of GS "mode 2", i.e. the synthetic Pan is obtained as a low-pass filtered version of the full resolution Pan. As a counterpart, Fig. 2(d) is less sharp than Fig. 2(c). Both Figs. 2(e) and (f) result from the proposed modified versions of GS "mode 2". Fig. 2(e) has been obtained by synthesizing the low-resolution Pan as a weighted sum of all spectral bands, according to the fixed weights proposed by Tu *et al*. (2004). Fig. 2(e) appears as sharp as Fig. 2(c), but the spectral distortion is mitigated. The result obtained with the adaptive preprocessing

Table 1. Average cumulative distortion indices between original 4 m MS bands and fused images obtained from 16 m MS and 4 m Pan. Q4 is a quality index and should be as high as possible; the average angular error SAM and ERGAS are distortion indices and should be as low as possible.

4:1	EXP	GS1	GS2	GSF	GSA	IHS	IHF	IHA
Q4	0.630	0.857	0.849	0.860	0.864	0.653	0.834	0.850
SAM	4.85°	4.19°	4.03°	4.06°	3.82°	4.52°	3.92°	3.94°
ERGAS	5.94	3.83	3.81	3.71	3.55	5.26	3.38	3.23

is reported in Fig. 2(f). The weights calculated from the whole image are $\alpha = 0.067$, $\beta = 0.189$, $\gamma = 0.228$, $\delta = 0.319$ and $\varepsilon = 25.028$. The spectral fidelity with respect to the original MS data in Fig. 2(b) is impressive, while spatial details appear as sharp as those of Fig. 2(c). Small 128×128 details of all methods, both GS-based and IHS-based, are displayed in Fig. 3. Spectral distortions are noticeable as changes in colour hues of the river with respect to the re-sampled low-resolution originals.

Visual judgement is corroborated by quantitative assessment obtained by degrading the spatial resolution of all original images by a factor of 4 and performing fusion on such images (Wald *et al*. 1997). The comparison of the fused images with the originals allows such scores as Q4 (Alparone *et al*. 2004a), SAM and ERGAS (Ranchin and Wald 2000; Ranchin *et al*. 2003) to be computed. Numerical values are reported in Table 1 and show that the GS approach is superior to the IHS approach in general; more specifically, that the proposed regression-based enhanced GS turns out to be the most performing method among those reviewed.

5 CONCLUSIONS

The multivariate regression with constant offset brings advantages with respect to all other preprocessing methods adopted to create the synthetic low-resolution Pan image, in both GS and IHS fusion methods. The proposed enhanced GS is effective in improving the quality of the fused images and requires negligible extra computation with respect to that of the standard GS procedure. Furthermore, the pre-processing block can be easily written in IDL and integrated within the GS ENVI module.

REFERENCES

Aiazzi, B., L. Alparone, S. Baronti, and A. Garzelli (2002). Context-driven fusion of high spatial and spectral resolution data based on oversampled multiresolution analysis. *IEEE Trans. Geosci. Remote Sensing 40*(10), 2300–2312.

Aiazzi, B., L. Alparone, S. Baronti, A. Garzelli, and M. Selva (2006). MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogramm. Eng. Remote Sens. 72*(5), 591–596.

Alparone, L., S. Baronti, A. Garzelli, and F. Nencini (2004a). A global quality measurement of Pan-sharpened multispectral imagery. *IEEE Geosci. Remote Sensing Lett. 1*(4), 313–317.

Alparone, L., S. Baronti, A. Garzelli, and F. Nencini (2004b). Landsat ETM+ and SAR image fusion based on generalized intensity modulation. *IEEE Trans Geosci. Remote Sensing 42*(12), 2832–2839.

Garzelli, A. and F. Nencini (2005). Interband structure modeling for Pan-sharpening of very high resolution multispectral images. *Inform. Fusion 6*(3), 213–224.

Garzelli, A. and F. Nencini (2006). PAN-sharpeningof very high resolution multispectral images using genetic algorithms. *Int. J. Remote Sensing 27*(15), 3273–3292.

Laben, C.A. and B.V Brower (2000). Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. US Patent # 6,011,875, Eastman Kodak Company.

Ranchin, T., B. Aiazzi, L. Alparone, S. Baronti, and L. Wald (2003). Image fusion – the ARSIS concept and some successful implementation schemes. *ISPRSJ. Photogramm. Remote Sensing 58*(1-2), 4–18.

Ranchin, T. and L. Wald (2000). Fusion of high spatial and spectral resolution images: the ARSIS concept and its implementation. *Photogramm. Eng. Remote Sens. 66*(1), 49–61.

Tu, T.-M., P. S. Huang, C.-L. Hung, and C.-P Chang (2004). A fast Intensity-Hue-Saturationfusion technique with spectral adjustment for IKONOS imagery. *IEEE Geosci. Remote Sensing Lett. 1*(4), 309–312.

Tu, T.-M., S.-C. Su, H.-C. Shyu, and P. S. Huang (2001). A new look at IHS-like image fusion methods. *Information Fusion 2*(3), 177–186.

Wald, L., T. Ranchin, and M. Mangolini (1997). Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogramm. Eng. Remote Sens. 63*(6), 691–699.

Wang, Z., D. Ziou, C. Armenakis, D. Li, and Q. Li (2005). A comparative analysis of image fusion methods. *IEEE Trans. Geosci. Remote Sensing 43*(6), 1391–1402.

Zhang, Y. (2004). Understanding image fusion. *Photogramm. Eng. Remote Sens. 70*(6), 657–661.