A Pan-sharpening method appropriate to vegetation applications

Ying Zhang (张 瑛)*, Binbin He (何彬彬), and Xiaowen Li (李小文)*****

Institute of Geo-Spatial Information Science and Technology,
University of Electronic Science Technology of China, Chengdu 610054, China
**E-mail: zhangyingxf@hotmail.com; **e-mail: binbinhe@uestc.edu.cn; ***e-mail: lix@bnu.edu.cn*

Received December 8, 2008

An improved Pan-sharpening algorithm appropriate to vegetation applications is proposed to fuse a set of IKONOS panchromatic (PAN) and multispectral image (MSI) data. The normalized difference vegetation index (NDVI) is introduced to evaluate the quality of fusion products. Compared with other methods such as principal component analysis (PCA), wavelet transform (WT), and curvelet transform (CT), this algorithm has a better trade-off between keeping the spatial and spectral information. The NDVI performances indicate that the fusion product of this method is more suitable for vegetation applications than the other methods.

OCIS codes: 100.0100, 110.0110, 350.0350, 350.2660.
doi: 10.3788/COL20090709.0781.

Light remains less energy when passing a narrow path than passing a broad one. In optical remote sensing, it is necessary to extend the instantaneous field of view (IFOV) so that multispectral sensors can gather enough energy to form images. This induces that the spatial resolution of multispectral images (MSIs) is lower than that of panchromatic (PAN) images. However, spatial information and spectral information in remote sensing images are all necessary for many applications such as researches on urban, vegetation, and geology. Therefore, it is very useful for remote sensing applications to obtain images with higher quality by merging PAN and MSI images. This processing is called Pan-sharpening or image fusion.

There are three categories of fusion methods[1], projection-substitution, relative contribution, and methods that belong to the *amelioration de la resolution spatiale par injection de structures* (ARSIS) concept[2−11]. Related experiments showed that the former two types of methods often caused color distortion (or spectral distortion), and the filtering operation in ARSIS fusion may produce ringing artifacts[9]. Recent research[1] has shown that the weakness of fusion methods are mainly caused by the assumptions when developing different fusion algorithms, and the physical principle should be taken into account. On the other hand, the commonly used statistical indices are statistically strong but physically plausible for remote sensing applications. Our experiments show that fusion results are not appropriate to some applications (especially vegetation applications), though the statistical indices values indicate a good quality. Hence, how to choose quantitative indices to evaluate the quality of fusion product is also an important issue. In this letter, a fusion method taking into account the physical principle of image formation is proposed. We aim to produce the fusion products which can preserve the initial spectral information and be suitable for vegetation applications. Moreover, the normalized difference vegetation index (NDVI)[12−14] is introduced as a quantitative index of "spectral information" to evaluate the quality of fusion results because of its importance in vegetation studies.

In optical remote sensing, if the atmospheric effects are neglected, the physical principle of optical image formation can be approximately written as[7]

$$\mathrm{DN_b}(x,y) = k_\mathrm{b}\rho_\mathrm{b}(x,y)\cos[\gamma(x,y)] + O_\mathrm{b}, \tag{1}$$

where $\mathrm{DN_b}(x,y)$ denotes the digital number of pixel (x,y) in band b, k_b is a constant related to the gain factor of the sensor calibration coefficients, ρ_b is the pixel reflectance observed at the top of atmosphere, $\gamma(x,y)$ is the solar incidence angle to the terrain surface at pixel (x,y), and O_b is the offset factor of the sensor calibration coefficients.

Based on this physical principle, Wang *et al.* proposed the general image fusion (GIF) framework[7]:

$$\frac{\mathrm{DN_p^h}(\Omega)}{\mathrm{DN_p^l}(\Omega)} = \frac{\mathrm{DN_b^h}(\Omega)}{\mathrm{DN_b^l}(\Omega)}, \tag{2}$$

where $\mathrm{DN_p^h}(\Omega)$, $\mathrm{DN_p^l}(\Omega)$, $\mathrm{DN_b^h}(\Omega)$, and $\mathrm{DN_b^l}(\Omega)$ are DN values of the object Ω which is represented by pixel (x,y); the superscripts h and l represent high-resolution and low-resolution, respectively; and the subscripts p and b represent p-band (panchromatic band) and b-band (a band in the original MSI, such as blue band), respectively.

In fusion problems, $\mathrm{DN_p^h}(\Omega)$ and $\mathrm{DN_b^l}(\Omega)$ are known while $\mathrm{DN_p^l}(\Omega)$ and $\mathrm{DN_b^h}(\Omega)$ are unknown. Then $\mathrm{DN_b^h}(\Omega)$ can be solved as long as $\mathrm{DN_p^l}(\Omega)$ is obtained. So how to estimate $\mathrm{DN_p^l}(\Omega)$ is a key problem in the GIF fusion method. From this point of view, principal component analysis (PCA) and wavelet transform (WT)[7] fusion can be regarded as examples of GIF fusion. More detailed examples can be found in Ref. [7]. Spectral characteristics of objects are not embodied in the commonly used methods. Here we introduce another method taking into account the spectral characteristics.

Suppose there is a linear relationship among the reflectances of special earth surface on different bands:

 © 2009 Chinese Optics Letters

$$\rho_{\mathrm{p}}(\Omega) = [\alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_n] \cdot [\rho_{\mathrm{b1}}(\Omega) \quad \rho_{\mathrm{b2}}(\Omega) \quad \cdots \quad \rho_{\mathrm{bn}}(\Omega)]^{\mathrm{T}} = \sum_{i=1}^{n} \alpha_i \cdot \rho_{\mathrm{b}i}(\Omega), \tag{3}$$

where ρ_{p} represents the reflectance of Ω on p-band light, and α_i $(i = 1, 2, \cdots, n)$ represents the reflectance relationship. At a certain time, ρ_{p} and $\rho_{\mathrm{b}i}(\Omega)$ (for $i = 1, 2, \cdots, n$) are constants for a certain object Ω. So the coefficients can be found.

To be simple, let $O_{\mathrm{b}} = 0$ and we obtain

$$\mathrm{DN}_{\mathrm{b}}(\Omega) = \sum_{i=1}^{n} \beta_i \cdot \mathrm{DN}_{\mathrm{b}i}(\Omega), \tag{4}$$

where $\beta_i = \alpha_i \cdot k_{\mathrm{p}}/k_{\mathrm{b}i}$ for $i = 1, 2, \cdots, n$. Equation (4) means that if there is a linear relationship among reflectances, the relationship among the corresponding DN values is also linear.

For a set of PAN and n-bands MSI at the same spatial resolution l, we have

$$\mathrm{DN}_{\mathrm{p}}^{\mathrm{l}}(\Omega) = [\beta_1 \ \beta_2 \ \cdots \ \beta_n] \cdot \left[\mathrm{DN}_{\mathrm{b1}}^{\mathrm{l}}(\Omega) \ \mathrm{DN}_{\mathrm{b2}}^{\mathrm{l}}(\Omega) \ \cdots \ \mathrm{DN}_{\mathrm{bn}}^{\mathrm{l}}(\Omega)\right]^{\mathrm{T}}. \tag{5}$$

Denoting $A = [\beta_1 \ \beta_2 \ \cdots \ \beta_n]$, $\mathrm{DN}_{\mathrm{p}}^{\mathrm{l}}(\Omega)$ can be obtained from Eq. (5) if A is estimated.

The new MSI with higher spatial resolution can be obtained through the following method. Firstly, the PAN image is downsampled to the lower resolution and A is estimated by total least square (TLS) algorithm with both the downsampled PAN data and the original MSI data. Secondly, MSI is upsampled to the higher resolution and we obtain $\mathrm{DN}_{\mathrm{p}}^{\mathrm{l}}(\Omega)$ via Eq. (5). Thirdly, we can get fused images by Eq. (2). Moreover, since we have $\beta_i = \alpha_i \cdot k_{\mathrm{p}} k_{\mathrm{bi}}$, A will depend on the object Ω. The pixels of original images should be identified to different objects via classification algorithm such as K-mean cluster. We call this method a TLS-GIF with classification (TLS-GIF-WC). The details can be described as the following four steps.

Step 1: downsample the original PAN data and get $\mathrm{DN}_{\mathrm{p}}^{\mathrm{downsample}}$.

Step 2: classify pixels of MSI to $m$objects (Ω_i, for $i = 1, 2, \cdots, m$) by K-mean classification. Then estimate A by Eq. (5) via the TLS algorithm for every object.

Step 3: upsample MSI and get $\mathrm{DN}_{\mathrm{bi}}^{\mathrm{upsample}}$ (for $i = 1, 2, \cdots, m$). Then classify pixels of the new MSI to m objects and estimate $\mathrm{DN}_{\mathrm{p}}^{\mathrm{l}}$ via Eq. (5) using $\mathrm{DN}_{\mathrm{bi}}^{\mathrm{upsample}}$ and A.

Step 4: get $\mathrm{DN}_{\mathrm{b}i}^{\mathrm{h}}$ (for $i = 1, 2, \cdots, m$) via Eq. (2).

It is worth mentioning that the above method is based on the condition $O_{\mathrm{b}} = 0$. When this condition does not hold for some sensors, we can subtract O_{b} from the image data before fusion. With respect to the resampling techniques, we adopt the cubic convolution algorithm to upsample and the pixel aggregate algorithm to downsample the MSI.

In the following, a set of IKONOS images are fused via TLS-GIF-WC. For comparison, the projection-substitution method, PCA fusion, and two common ARSIS methods, WT and curvelet transform (CT)[15], are adopted to produce fusion products.

Figure 1 shows the fusion results produced by the four methods. Visually, the fused images produced by WT (Fig. 1(d)), CT (Fig. 1(e)), and TLS-GIF-WC (Fig. 1(f)) look clearer than that by PCA (Fig. 1(c)). In other words, there is more spatial information in the three results than PCA. However, some artifacts can be found in Figs. 1(d) and (e). This means that WT and CT methods get more spatial information at the cost of more artifacts and TLS-GIF-WC gets a better trade-off.

Regarding the quantitative evaluation of quality, more and more attention has been paid to the performance of keeping spectral information. As defined by Wald[16], "the 'greater quality' depends upon the application". In vegetation applications such as plant extraction and vegetation changing detection, NDVI is a very influential parameter. So we introduce NDVI as a quantitative index of spectral information. Researches show that the NDVI can be calculated from near-infrared (NIR) band and red band images[17]:

$$\mathrm{NDVI} = \frac{\mathrm{DN}_{\mathrm{NIR}} - \mathrm{DN}_{\mathrm{red}}}{\mathrm{DN}_{\mathrm{NIR}} + \mathrm{DN}_{\mathrm{red}}}. \tag{6}$$

Table 1 lists the range of NDVI values calculated from the original MSI, upsampled MSI, and fused images. Table 1 shows that the range of NDVI values in TLS-GIF-WC result is consistent with that of upsampled MSI, but the ranges in PCA, WT, and CT results are changed. The boundary situations in WT and CT results, –1 and 1, indicate that the DN value of the corresponding pixels

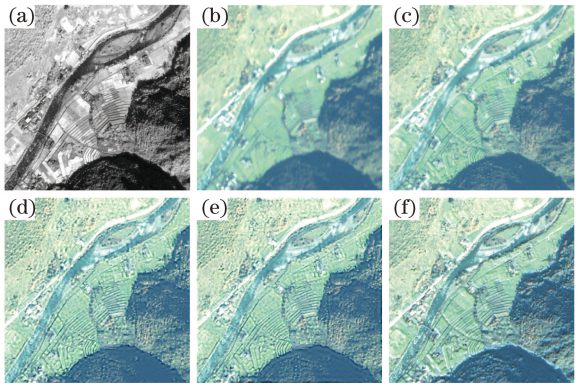

Fig. 1. IKONOs images of a mountain area in China. (a) PAN, 1 m; (b) MSI, 4 m upsampled to 1 m; (c) PCA fusion result; (d) WT fusion result; (e) CT fusion result; (f) TLS-GIF-WC fusion result.

Table 1. Ranges of NDVI Calculated from MSI

NDVI	Original MSI	Upsampled MSI	PCA
Minimum	–0.5079	–0.4961	–1.0000
Maximum	0.6856	0.6846	0.7359
NDVI	WT	CT	TLS-GIF-WC
Minimum	–1.0000	–1.0000	–0.4961
Maximum	1.0000	1.0000	0.6846

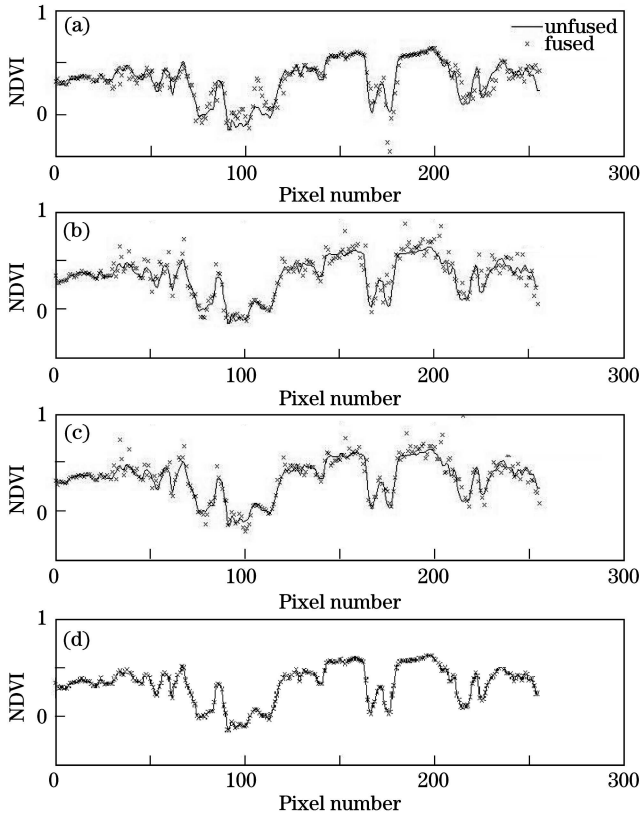

Fig. 2. NDVIs in fused and unfused MSI. (a) PCA fused MSI compared with upsampled MSI; (b) WT fused MSI compared with upsampled MSI; (c) CT fused MSI compared with upsampled MSI; (d) TLS-GIF-WC fused MSI compared with upsampled MSI.

in NIR band image or in red band image is zero. This is not consistent with the original MSI or the upsampled MSI. Furthermore, 256 pixels are randomly selected from PCA, WT, CT, and TLS-GIF-WC fusion results separately, and the corresponding NDVI values in fused and unfused MSI are ploted in Fig. 2. It is shown that NDVI values in TLS-GIF-WC result are the same as those of unfused MSI but the NDVI values in other three results do not. This consistency in TLS-GIF-WC result can be verified through mathematical deduction.

From Eq. (2) we have

$$\begin{aligned} \mathrm{DN}^{\mathrm{h}}_{\mathrm{NIR}} &= \frac{\mathrm{DN}^{\mathrm{h}}_{\mathrm{p}}}{\mathrm{DN}^{\mathrm{l}}_{\mathrm{p}}} \cdot \mathrm{DN}^{\mathrm{l}}_{\mathrm{NIR}}, \\ \mathrm{DN}^{\mathrm{h}}_{\mathrm{red}} &= \frac{\mathrm{DN}^{\mathrm{h}}_{\mathrm{p}}}{\mathrm{DN}^{\mathrm{l}}_{\mathrm{p}}} \cdot \mathrm{DN}^{\mathrm{l}}_{\mathrm{red}}. \end{aligned} \tag{7}$$

Combining Eqs. (6) and (7), we obtain

$$\mathrm{NDVI}^{\mathrm{h}} = \mathrm{NDVI}^{\mathrm{l}}, \tag{8}$$

which means the GIF fusion will not change the NDVI while enhancing the spatial resolution.

Therefore, the mathematical analysis and experiments show that the TLS-GIF-WC method preserves more spectral information than PCA, WT, and CT methods.

Table 2. ERGAS Values of Fused Images

PCA	WT	CT	TLS-GIF-WC
1.2635	1.3267	1.4079	1.3141

A statistical index, the *erreur relative globale adimensionnelle de synthese* (ERGAS)[11], is adopted to evaluate the performance of keeping spectral information. The ERGAS values of the fusion results are listed in Table 2. In Table 2, the ERGAS value of PCA result, is the closest to the ideal value 0, and the TLS-GIF-WC result is better than WT and CT results. Combining the performances in Tables 1 and 2, we can find that PCA has the best statistical performance, but its NDVI is changed during fusion process. Since NDVI is one of the most popular indices reflecting the spectral information of MSI in vegetation studies, we had better choose the fusion method which does not change the NDVI value or some other vegetation indices[17]. From this perspective, the proposed method works best among the four methods.

In conclusion, the TLS-GIF-WC fusion result preserves more spectral information and TLS-GIF-WC is more suitable for vegetation application than the other three methods. As far as fusion method is mentioned, it is necessary to choose a suitable performance index to indicate whether the fusion products are appropriate to applications or not. In future work, fusion methods will be designed by optimizing these special indices.

This work was supported by the National Natural Science Foundation of China (No. 40701146), the National "863" Program of China (No. 2007AA12Z227), and the National "973" Program of China (No. 2007CB714406).

References

1. C. Thomas, T. Ranchin, L. Wald, and J. Chanussot, IEEE Trans. Geosci. Remote Sens. **46,** 1301 (2008).
2. T.-M. Tu, S.-C. Su, H.-C. Shyu, and P. S. Huang, Information Fusion **2,** 177 (2001).
3. T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, and L. Wald, ISPRS J. Photogramm. Remote Sens. **58,** 4 (2003).
4. L. Deng, Y. Chen, and J. Li, J. Infrared Millim. Waves (in Chinese) **24,** 34 (2005).
5. X. Sun, J. Geomatics (in Chinese) **33,** (3) 40 (2008).
6. L. Alparone, B. Aiazzi, S. Baronti, and A. Garzelli, in *Proc. Int. Geosci. Remote Sens. Symp.* 458 (2003).
7. Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, IEEE Trans. Geosci. Remote Sens. **43,** 1391 (2005).
8. W. Dou, Y. Chen, X. Li, and D. Z. Sui, Comput. Geosci. **33,** 219 (2007).
9. A. Garzelli, F. Nencini, and L. Capobianco, IEEE Trans. Geosci. Remote Sens. **46,** 228 (2008).
10. K. Liu, L. Guo, and W. Chang, Acta Opt. Sin. (in Chinese) **28,** 681 (2008).
11. Q. Zhang and B. Guo, Act. Opt. Sin. (in Chinese) **28,** 74 (2008).
12. Q. Tian and X. Min, Advance in Earth Sciences (in Chinese) **13,** 327 (1998).
13. E. Swinnen and F. Veroustraete, IEEE Trans. Geosci. Remote Sens. **46,** 558 (2008).
14. Z. Wu, W. Li, Y. Zeng, J. Li, and Y. Xue, Science of Surveying and Mapping (in Chinese) **33,** 98 (2008).
15. N. Jiang, Y. Wang, and J. Mao, Chin. J. Sci. Instrum. (in Chinese) **29,** 61 (2008).
16. L. Wald, IEEE Trans. Geosci. Remote Sens. **37,** 1190 (1999).
17. X. Xu, *Remote Sensing Physics* (in Chinese) (Peking University Press, Beijing, 2005).