Pansharpening of High and Medium Resolution Satellite Images Using Bilateral Filtering

Diego Renza, Estibaliz Martinez, Agueda Arquero, and Javier Sanchez

Polytechnic University of Madrid
d.renza@alumnos.upm.es,
{emartinez,aarquero}@fi.upm.es,
jarsahe@hotmail.com

Abstract. We provide and evaluate a fusion algorithm of remotely sensed images, i.e. the fusion of a panchromatic (PAN) image with a multi-spectral (MS) image using bilateral filtering, applied to images of three different sensors: SPOT 5, Landsat ETM+ and Quickbird. To assess the fusion process, we use six quality indexes, that confirm, along with visual analysis, good overall results for the three sensors.

Keywords: Bilateral filter, Image fusion, Pansharpening.

1 Introduction

The term "image fusion" usually implies the integration of images acquired by multiple sensors with the intention of providing a better perspective of a scene that contains more content. In remote sensing there are many sensors that have a set of multispectral bands and a co-registered higher spatial resolution panchromatic band. Examples of this sensors are SPOT, Landsat ETM+, QuickBird, or IKONOS. With appropriate algorithms it is possible to combine these data and produce multispectral imagery with higher spatial resolution. This concept is known as multispectral or multisensor merging, fusion or sharpening (of the lower-resolution image) [1].

In remote sensing, the fusion schemes can be grouped into three classes: (1) color related techniques, (2) statistical/numerical methods and (3) combined approaches [2]. The first covers the color composition of three image channels in the RGB color space as well as more sophisticated color transformations, e.g., HSV (hue-saturation-value), IHS (intensity-hue-saturation) [3]. The second group includes methods that use different mathematical tools, like channel statistics including correlation and filters. High pass filtering (HPF) [1], principal component analysis (PCA) [3] and multiresolution analysis (MRA) [4] belong to this category. The combined approaches include methods that are not limited to follow one approach.

In particular, fusion methods based on injection of high-frequency components into resampled versions of the MS bands have shown better spectral results, attracting the interest of researchers in recent years [5]. Within these methods

I. Bloch and R.M. Cesar, Jr. (Eds.): CIARP 2010, LNCS 6419, pp. 311–318, 2010.

can highlight the different variants of MRA (e.g. wavelets), or a less complex scheme, the high pass filtering, so we propose a new fusion scheme that makes use of bilateral filter as principal element for the extraction of features.

The next section describes a non-linear technique appropriate for image processing, known as bilateral filter (BF) which has been used for some image processing applications like denoising, texture editing or optical flow estimation [6]; also BF has been used for merging video (Visible (RGB) + IR) [7] and detail enhancement in multi-light image collections [8]. BF is an effective way to smooth an image while preserving its discontinuities and also to separate image structures of different scales, hence we propose its use in pansharpening.

2 Bilateral Filter

A bilateral filter is an edge-preserving smoothing filter. BF operates both in the domain and range of the image (i.e. pixel values). In the image domain, the core component of many filters is the kernel convolution. At each pixel position the filter estimates the local average of intensities, which corresponds to low-pass filtering. For instance, the Gaussian filtering (GF) is a weighted average of the intensity of the adjacent pixels where the weights decrease with the spatial distance to the center position; GF is a simple approach to smooth images, but with blurred edges, because pixels across discontinuities are averaged together.

Bilateral filter was proposed based on the definition of Gaussian convolution, taking into account both the image domain as the range image. BF is also defined as a weighted average of nearby pixels, the difference with GF is that BF takes into account the difference in value with the neighbors to preserve edges while smoothing. The key idea of the bilateral filter is that for a pixel influences another pixel, it should not only occupy a nearby location but also have a similar value [6]. The bilateral filter is simple: each pixel is replaced by a weighted average of its neighbors; BF depends only on two parameters that indicate the size and contrast of the features to preserve; furthermore, BF can be used in a non-iterative manner. Mathematically BF is given by [6]:

$$BF\left[I\right]_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}\left(\|p-q\|\right) G_{\sigma_r}\left(|I_p - I_q|\right) I_q \tag{1}$$

$$W_p = \sum_{q \in S} G_{\sigma_s}\left(\|p-q\|\right) G_{\sigma_r}\left(|I_p - I_q|\right) \tag{2}$$

Where normalization factor W_p ensures pixel weights sum to 1.0.

$\sum\limits_{q \in S}$ denotes a sum over all image pixels indexed. $|.|$ is used for the absolute value and $\|.\|$ for the L2 norm. G_σ denotes a 2D Gaussian Kernel.

G_{σ_r} is a range Gaussian function that decreases the influence of pixels q when their intensity values differ from I_p, therefore as the range parameter σ_r increases, the bilateral filter gradually approximates Gaussian convolution. G_{σ_s} is a spatial Gaussian weighting function that decreases the influence of distant pixels, therefore increasing the spatial parameter σ_s smooths larger features.

3 Image Fusion with Bilateral Filter

The potential benefit of a fused image is that the single resulting image both has a high spatial resolution and contains the spectral information, hence, the result of image fusion is a new image which is more suitable for human and machine perception or further image-processing tasks such a classification, segmentation, feature extraction or object recognition. For image fusion with BF, we first resample the MS image so that its bands have the same pixel size as the PAN image. These bands, along with PAN image are decomposed by means of BF.

The BF can split an image into two parts: (1) a filtered image, equivalent to low frequency component and in particular to spectral information, which holds only the large-scale features, as the bilateral filter smooths away local variations without affecting strong edges and (2) a "residual" image, made by subtracting the filtered image from the original, which holds only the image portions that the filter removed, i.e. detail or small-scale components that can represent texture or structures within image.

Therefore, the idea of applying the bilateral filter is to extract the details of the PAN image and the approximation (large-scale component) of the MS bands. A key factor when extracting these components is filter-parameters determination. Experimentally we obtained the best results with a $\sigma_s = R/2$, for both the PAN and MS images filter, where R is the scale ratio between PAN and MS images. In the case of σ_r, best results were obtained with $\sigma_r = 0.1(2^{nbits} - 1)$ for MS image and $\sigma_r = 0.4(2^{nbits} - 1)$ for the PAN image, where $nbits$ is the radiometric resolution of each image.

To combine these two components it is necessary to define an injection model to establish how high frequency information will be merged with the MS bands. Such a model can be global over the whole image or can depend on the spectral or spatial context [5]. In our case it was only necessary to apply a weight factor to the high frequency component, adding it to the low frequency component, thus minimizing spectral distortions. For the injection of these details the gain factor is defined by the ratio of standard deviation of each MS band approximation to the standard deviation of the PAN. The fusion process is showed in the Fig. 1.

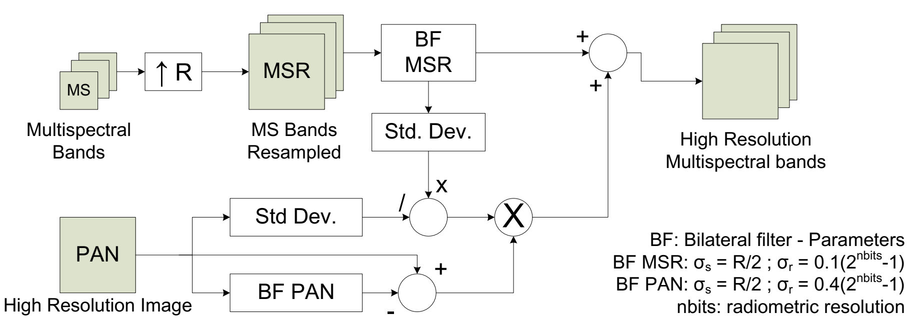


Fig. 1. Block diagram of bilateral filter-based image fusion approach (scale ratio 1:R)

4 Data and Methodology

4.1 Study Area

The proposed approach was applied to images from sensors of medium and high resolution: SPOT 5, Landsat ETM+ and Quickbird. The cases presented in this paper are:

1. Spot5: PAN 2048 x 2048 px (2,5 m), MS 512x512x4 (10 m). Madrid (Spain), 461990 E, 4480340 N (WGS84 UTM Zone 30)
2. Landsat ETM+: PAN 768 x 768 px (15 m), MS 384x384x6 (30 m) (Bands 1-5 & 7). Madrid (Spain), 409012,5 E, 4419247,5 N (WGS84 UTM Zone 30)
3. Quickbird: PAN 2048 x 2048 px (0,7 m), MS 512x512x4 (2,8 m). Madrid (Spain), 434753,2 E, 4479221,6 N (WGS84 UTM Zone 30)

4.2 Quality Determination - Comparison with Other Techniques

For comparison we employed four different image-fusion algorithms: two standard fusion techniques implemented in commercial software (ENVI), these were Gram-Schmidt spectral sharpening [9] and principal component (PC) fusion [10]. The third algorithm was a multiresolution analysis-based method, proposed recently [11,12]. The fourth method, bilateral filter-based fusion, is proposed in this work. All images were visually and statistically evaluated.

For quality evaluation, the first step was a visual analysis of the fused images. Therefor it was necessary take into account the color preservation of the fused image with respect to the original MS image. In the same way the quality evaluation took into account the spatial improvement of fused image compared to the original PAN image. Although the visual analysis is subjective and depends on the interpreter, it usually gives a first idea of spatial and spectral distortions.

To evaluate the quality of the merged image, it is usually compared against a reference image. In practice a true reference image does not exist, therefore, it is necessary to create it from the original PAN (with resolution h) and MS (with resolution l) images. Therefor, the PAN image is degraded to the low resolution l and the MS image is degraded to the resolution l^2/h. The fusion process is applied to new degraded images and the quality can be assessed using the original MS image as a reference image [13].

To assess the fusion process we use the following metrics [4,13]: (a) Relative Dimensionless Global Error (ERGAS), used to estimate the overall spectral quality of fused images. (b) Spectral Angle Mapper (SAM), which determines the degree of spectral similarity of an image against a known or reference image, expressed in terms of the average angle between the two spectra. (c) Mean bias (MB), which measures the difference in central tendency of two images. (d) Variance Difference (VD) for estimating the change in variance during the enhancement of the spatial resolution. (e) Standard deviation difference (SDD), which provides a global indication of the level of error at any pixel. (f) Correlation Coefficient (CC), which shows the similarity in small size structures between the original and synthetic images.

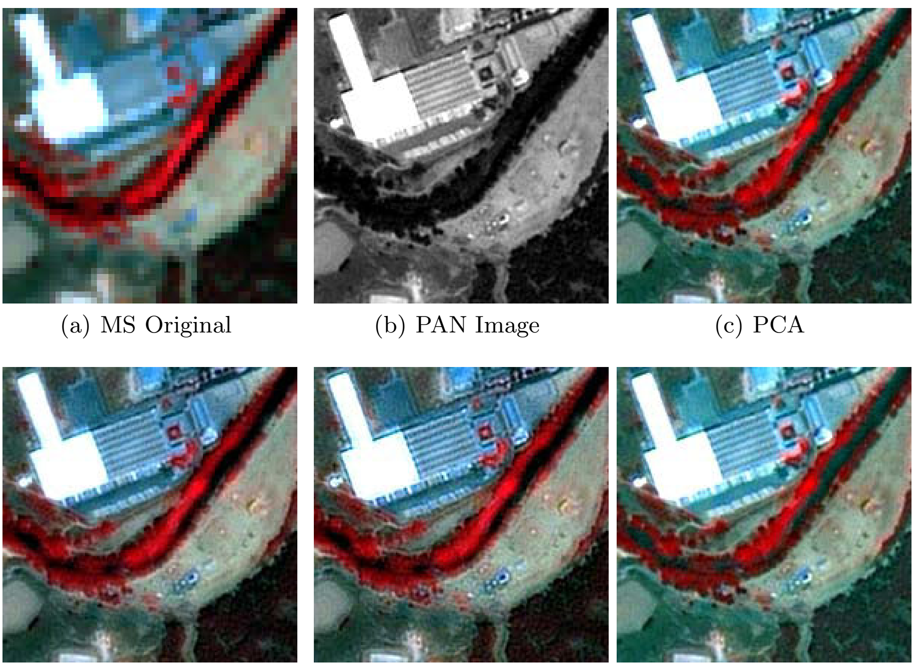
5 Results

Tables 1,2,3 show the quality measures for the whole image of three cases studied. Figures 2,3,4 show only a sub-scene of the corresponding merged images. The visual analysis shows a perceptible improvement in spatial quality for fused images compared with the original MS image. Although slight variations occur, all methods clearly sharpen the original image and improving the spatial quality.

By comparing the resulting images, one can also see that the fused images retain a high degree of spectral information. For color preservation, 432 and 321

Table 1. Test Case No. 1, SPOT5. Quality metrics. [a]Average for all four bands.

Type	ERGAS	SAM (*rad*)	MB[a]	VD[a]	SDD[a]	CC[a]
Ideal	0	0	0	0	0	1
Bilateral filter	**1,3316**	**0,0260**	**-6,38E-04**	**0,0375**	**0,0523**	**0,9736**
Gram-Schmidt	1,9021	0,0329	1,64E-04	0,1164	0,0756	0,9476
PCA	1,7581	0,0317	-2,96E-04	0,1405	0,0698	0,9549
DT-CWT	1,6392	0,0273	2,59E-05	0,0050	0,0642	0,9607



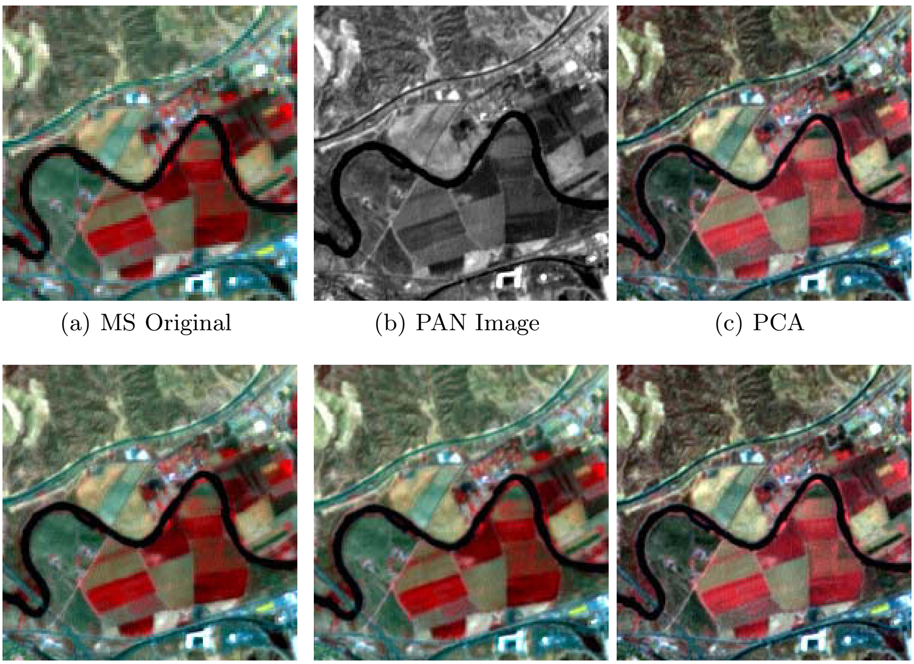
(a) MS Original (b) PAN Image (c) PCA

(d) BF (e) DT-CWT (f) Gram-Schmidt

Fig. 2. Test Case No. 1, SPOT5. A fragment of the original MS (10 m), PAN (2,5 m) and fused images (2,5 m). (200x200 pixels false color (NIR-Red-Green) composition).

Table 2. Test Case No. 2, Landsat ETM+. Quality metrics. [a]Average for all six bands.

Type	ERGAS	SAM (*rad*)	MB[a]	VD[a]	SDD[a]	CC[a]
Ideal	0	0	0	0	0	1
Bilateral filter	**2,1755**	**0,0243**	**-1,02E-05**	**0,0691**	**0,0427**	**0,9800**
Gram-Schmidt	4,0632	0,0352	4,34E-04	0,0722	0,0794	0,9329
PCA	4,5065	0,0368	-2,76E-04	0,0577	0,0878	0,9185
DT-CWT	2,3891	0,0294	-9,01E-06	0,0532	0,0473	**0,9749**



(a) MS Original (b) PAN Image (c) PCA

(d) BF (e) DT-CWT (f) Gram-Schmidt

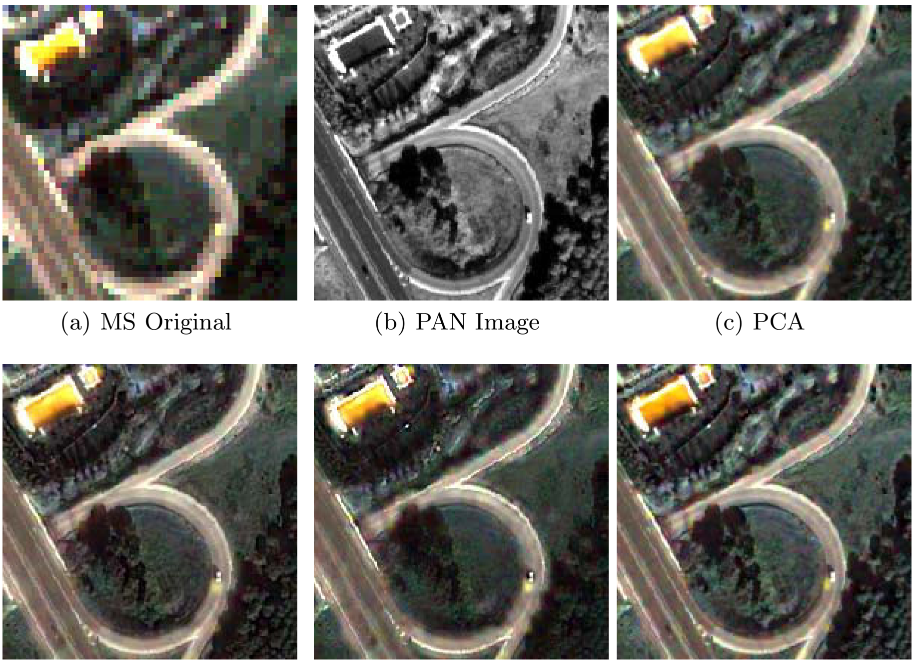
Fig. 3. Test Case No. 2, Landsat ETM+. A fragment of the original MS (30 m), PAN (15 m) and fused images (15 m). (200x200 pixels false color 432 composition).

band RGB combinations are showed for spectral analysis. The spectral distortion of the PCA and Gram-Schmidt methods is most visible in the red color of vegetation (false color). The BF and DT-CWT methods produce an image whose colors better match the original image.

Regarding the quality indexes, (Tables 1,2,3) the values obtained for ERGAS, combined with the angles obtained by the SAM metric, were acceptable, so the spectral information is largely preserved, particularly according to BF and DT-CWT for SPOT and Landsat images. As for the mean bias, results close to ideal were obtained since the injection of spatial information resulted in a near-zero mean. The difference in variance suggests an advantage for the BF and

Table 3. Test Case No. 3, Quickbird. Quality metrics. [a]Average for all four bands.

Type	ERGAS	SAM (*rad*)	MB[a]	VD[a]	SDD[a]	CC[a]
Ideal	0	0	0	0	0	1
Bilateral filter	**3,8319**	**0,0908**	**-6,30E-04**	**0.1907**	**0.1476**	**0.9288**
Gram-Schmidt	4,4896	0.1067	4,54E-04	0,3162	0.1726	0.9069
PCA	4,5392	0.1015	8,36E-04	0.2458	0.1745	0.9032
DT-CWT	3,8623	0.0868	-7,40E-05	0.1378	0.1486	0.9290



(a) MS Original (b) PAN Image (c) PCA
(d) BF (e) DT-CWT (f) Gram-Schmidt

Fig. 4. Test Case No. 3, Quickbird. A fragment of the original MS (2,8 m), PAN (0,7 m) and fused images (0,7 m). (200x200 pixels true color (RGB) composition).

DT-CWT, which preserve slightly more information. The overall estimate of error in a pixel (SDD) shows again better results for the BF and DT-CWT. Although visually the results are very similar in spatial quality, correlation coefficient shows slightly higher results for the BF in two of three cases.

6 Conclusion

The proposed fusion scheme showed good results applied to three different types of sensors. It was compared against traditional methods and showed that BF is an appropriate alternative for image fusion, presenting the best results for SPOT and Landsat images and results similar to a wavelet approach. In the Quickbird case it is necessary to study alternatives to the merger scheme adapted to the pixel size to preserve the spectral information more accurately. However, it is important to consider the influence of image degradation when implementing quality indices. A future work will consider alternatives such as multiscale approach and fast approximations of the bilateral filter.

References

1. Schowengerdt, R.A.: Remote Sensing: Models and Methods for Image Processing, 3rd edn. Academic Press, San Diego (2007)
2. Pohl, C., Van Genderen, J.L.: Multisensor image fusion in Remote sensing: concepts, methods and applications. Int. Journal of Remote Sensing 19, 823–854 (1998)
3. Gonzalez-A, M., Saleta, J.L., Garcia-C, R., Garcia, R.: Fusion of Multispectral and Panchromatic Images Using Improved IHS and PCA Mergers Based on Wavelet Decomposition. IEEE Trans. on Geos. and Rem. Sens. 42(6), 1291–1299 (2004)
4. Alparone, L., Wald, L., Chanussot, J., Thomas, C., Gamba, P., Bruce, L.M.: Comparison of Pansharpening Algorithms Outcome of the 2006 GRS-S Data-Fusion Contest. IEEE Trans. on Geoscience and Rem. Sensing 45(10) (2007)
5. Aiazzi, B., Baronti, S., Selva, M.: Image fusion through multiresolution oversampled decompositions. In: Stathaki, T. (ed.) Image Fusion: Algorithms and Applications. Academic Press, London (2008)
6. Paris, S., Kornprobst, P., Tumblin, J., Durand, F.: Bilateral Filtering: Theory and Applications. Foundations and Trends in Computer Graphics and Vision 4(1), 1–73 (2008)
7. Bennett, E.P., Mason, J.L., McMillan, L.: Multispectral Bilateral Video Fusion. IEEE Transactions on Image Processing 16(5), 1185–1194 (2007)
8. Fattal, R., Agrawala, M., Rusinkiewicz, S.: Multiscale Shape and Detail Enhancement from Multi-light Image Collections. In: ACM SIGGRAPH (2007)
9. Laben, C.A., Brower, B.V.: Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. US Patent No. 6,011,875. Eastman Kodak Company (2000)
10. Chavez, W.J., Sides, S.C., Anderson, J.A.: Comparison of three different methods to merge multiresolution and multispectral data. Phot. Eng. Rem. Sens. 57, 295–303 (1991)
11. Ioannidou, S., Karathanassi, V.: Investigation of the Dual-Tree Complex and Shift-Invariant Discrete Wavelet Transforms on Quickbird Image Fusion. IEEE Geoscience and Remote Sensing Letters 4(1), 166–170 (2007)
12. Renza, D., Martinez, E., Arquero, A.: Optimizing classification accuracy of remotely sensed imagery with DT-CWT fused images. In: Bayro-Corrochano, E., Eklundh, J.O. (eds.) CIARP 2009. LNCS, vol. 5856, pp. 1031–1038. Springer, Heidelberg (2009)
13. Wald, L.: Data fusion: Definitions and Architectures - Fusion of Images of Different Spatial Resolutions, Paris, ch. 8 (2002)