

BAYESIAN SUPER-RESOLUTION PANSHARPENING USING CONTOURLETS

I. Amro, J. Mateos *

M. Vega

Dept. de Ciencias de la Computación e I. A.
Universidad de Granada, España

Dept. de Lenguajes y Sistemas Informáticos
Universidad de Granada, España

ABSTRACT

Pansharpening is a technique that fuses a low resolution multispectral image and a high resolution panchromatic image, to obtain a multispectral image with the spatial resolution and quality of the panchromatic image while preserving spectral information of the multispectral image. In this paper, we present a new pansharpening method based on super-resolution and contourlet transform. The experimental results show that the proposed method not only enhances the spatial resolution of the pansharpened image, but also preserves the spectral information of the original multispectral image.

Index Terms— pansharpening, super-resolution, contourlets, multispectral image, remote sensing

1. INTRODUCTION

Pansharpening is a technique for fusing the information of a low resolution multispectral (MS) image and a high resolution panchromatic (PAN) image, usually remote sensing images, to provide a high resolution MS image with the level of detail of the PAN image. In the literature, this task has been addressed from different points of view (see [1] and [2] for a description and comparison of pansarpening methods). In recent years, Contourlet-based algorithms, specifically non-subsampled contourlet transform (NSCT)[3] based ones, are becoming popular[4]. The contourlet transform provides a complete shift-invariant and multiscale representation, with a fast implementation. The main problems of the classical methods based on contourlet are the control of the noise in the images and their dependency on an initial interpolation but they efficiently preserve texture and contour information.

Recently, a new super-resolution (SR) approach has been proposed [5]. Within the Bayesian formulation, this method incorporates prior knowledge on the expected characteristics of MS images using a Total Variation prior, and considers the sensor characteristics to model the observation process of both PAN and MS images. However, the MS bands which are not covered by PAN image cannot be improved properly using this method and some color bleeding may appear.

In this paper we propose a new pansharpening method that combines the super-resolution technique presented in [5] with non-subsampled contourlet transform in order to obtain a method that efficiently preserves the texture and contour information of the PAN image while improving all the bands of the image, even those that are not covered by the PAN image. This paper is organized as follows. In section 2 the Bayesian SR using contourlet approach is described and the used notation introduced. Section 4 describes the inference of the high resolution MS image. Experimental results

and comparison are presented in section 5 for synthetic and SPOT images and finally, section 6 concludes the paper.

2. PROBLEM FORMULATION

Let us assume that y , the unknown high resolution MS image we would have observed under ideal conditions, has B bands, y_b , $b = 1, \dots, B$, each one centered on a narrow spectral band, of size $p = m \times n$, that is, $y = [y_1^t, y_2^t, \dots, y_B^t]^t$, where each band of this image is expressed as a column vector by lexicographically ordering the pixels in the band, and t denotes the transpose of a vector or matrix. The observed low resolution MS image Y has B bands Y_b , $b = 1, \dots, B$, each of size $P = M \times N$ pixels, with $M < m$ and $N < n$. These images are also stacked into the vector $Y = [Y_1^t, Y_2^t, \dots, Y_B^t]^t$, where each band of this image is also expressed as a column vector by lexicographically ordering the pixels in the band. The sensor also provides us with a panchromatic image x of size $p = m \times n$, that contains reflectance data in a single band that covers a wide area of the spectrum.

Using the contourlet transform, these images can also be expressed in NSCT domain, as

$$x = \sum_j C_j x + C_r x, \quad y_b = \sum_j C_j y_b + C_r y_b, \quad (1)$$

where $C_j x$ and $C_j y_b$ are the NSCT coefficients for the PAN and MS bands, respectively, at a level of decomposition j and $C_r x$ and $C_r y_b$ are the residual (low pass filtered version of original) image of PAN and MS bands, respectively.

Following [5], we define the relationship between high resolution MS image and its low resolution counterpart as

$$Y_b = H y_b + n_b, \quad (2)$$

where the $P \times p$ degradation matrix H combines the subsampling, integration and blur present in the image and n_b is the capture noise assumed to be independent white Gaussian of known variance β_b^{-1} .

The panchromatic image contains the details of the high resolution MS image but lacks of its spectral information. Following the wavelet and contourlet based pansharpening methods, see [2, 4], the PAN image is considered here as a combination of the high frequency details of the high resolution MS bands, plus a residual, low pass filtered image, which have an unknown relationship with the MS bands. The relationship between the PAN and the high resolution MS images is defined in this paper as

$$x_d = \sum_j C_j x = \frac{1}{B} \sum_{b=1}^B \sum_j C_j y_b + v, \quad (3)$$

where x_d contains the details of the PAN image, obtained using the contourlet transform, which is able to effectively extract the details

*This work has been supported by the Consejería de Innovación, Ciencia y Empresa of the Junta de Andalucía under contract P07-TIC-02698.

of an image and v is the noise that is assumed to be Gaussian with zero mean and known variance γ^{-1} . Note that this model doesn't take into account the residuals of the NSCT, in a similar way as the additive wavelet or contourlet methods do. So, we are assuming that the objects structure is present in all the bands but, since the PAN image doesn't cover all the spectral range of the multispectral image covers, its intensities will not necessarily coincide with the MS bands intensities and, hence, we are not forcing this similarity.

3. BAYESIAN FORMULATION

The Bayesian formulation starts with the definition of the joint distribution $p(y, Y, x_d) = p(y)p(Y, x_d|y)$ and then performs the inference based on $p(y|Y, x_d) = p(y, Y, x_d)/p(Y, x_d)$. Also assuming that Y and x_d are independent, for a given y , $p(Y, x_d|y) = p(Y|y)p(x_d|y)$.

Let us now describe these probability distributions. For the MS image, we choose the Total Variation (TV) prior given by

$$p(y) = \prod_{b=1}^B p(y_b) \propto \prod_{b=1}^B \alpha_b^{p/2} \exp[-\alpha_b TV(y_b)], \quad (4)$$

with $TV(y_b) = \sum_{i=1}^p \sqrt{(\Delta_i^h(y_b))^2 + (\Delta_i^v(y_b))^2}$, where $\Delta_i^h(y_b)$ and $\Delta_i^v(y_b)$ represent the horizontal and vertical first order differences at pixel i respectively, and α_b is the model parameter of the band b .

From the degradation model in Eq. (2), we obtain the probability distribution of the observed image Y , given y ,

$$p(Y|y) = \prod_{b=1}^B p(Y_b|y_b) \propto \prod_{b=1}^B \beta_b^{P/2} \exp\left\{-\frac{1}{2}\beta_b \|Y_b - Hy_b\|^2\right\}. \quad (5)$$

The probability distribution of the details of the PAN image, x_d , given y , from Eq. (3), is written as

$$p(x_d|y) \propto \gamma^{p/2} \exp\left\{-\frac{1}{2}\gamma \left\|x_d - \frac{1}{B} \sum_{b=1}^B \sum_j C_j y_b\right\|^2\right\}. \quad (6)$$

4. BAYESIAN INFERENCE

In order to extract conclusions from the posterior distribution $p(y|Y, x_d)$ we need to calculate $p(Y, x_d)$. However, $p(Y, x_d)$ cannot be calculated analytically and we will apply the variational methodology to approximate the posterior distribution by another one, $q(y)$, that minimizes the Kullback-Leibler (KL) divergence [6], defined as

$$\begin{aligned} C_{KL}(q(y)||p(y|Y, x_d)) &= \int q(y) \log\left(\frac{q(y)}{p(y|Y, x_d)}\right) dy \\ &= \int q(y) \log\left(\frac{q(y)}{p(y, Y, x_d)}\right) dy + const = \mathcal{M}(q(y)) + const, \end{aligned} \quad (7)$$

which is always non negative and equal to zero only when $q(y) = p(y|Y, x_d)$.

Unfortunately, the integral in Eq. (7) cannot be directly evaluated due to the TV prior but we can approximate it by using the Majorization-Minimization approach[7] that converts a non-quadratic problem to a quadratic problem by the introduction of a new parameter that also needs to be estimated. Let us consider the inequality, also used in [5], which states that, for any $w \geq 0$ and $z > 0$

$$\sqrt{w} \leq \frac{w+z}{2\sqrt{z}}, \quad (8)$$

and use it with $w = (\Delta_i^h(y_b))^2 + (\Delta_i^v(y_b))^2$ and $z = u_b(i)$ to approximate the probability distribution $p(y_b)$ in Eq. (4) as

$$M(y_b, u_b) = \alpha_b^{p/2} \exp\left[-\alpha_b \sum_i \frac{(\Delta_i^h(y_b))^2 + (\Delta_i^v(y_b))^2 + u_b(i)}{2\sqrt{u_b(i)}}\right], \quad (9)$$

where u_b is a p -dimensional vector, $u_b \in (R^+)^p$ with components $u_b(i)$, $i = 1, \dots, p$ and that, as we will show later, has a tight relationship with the image. Substituting Eq. (9) into Eq. (4), we obtain $p(y) \geq c \cdot \prod_{b=1}^B M(y_b, u_b)$. This leads to the following lower bound for the joint probability distribution

$$p(y, Y, x_d) \geq c \cdot \left[\prod_{b=1}^B M(y_b, u_b)\right] p(Y|y)p(x_d|y) = F(y, Y, x_d, u), \quad (10)$$

where $u = [u_1^t, u_2^t, \dots, u_B^t]^t$.

By defining, $\tilde{\mathcal{M}}(q(y), u) = \int q(y) \log\left(\frac{q(y)}{F(y, Y, x_d, u)}\right) dy$, and using Eq. (10), we obtain

$$\mathcal{M}(q(y)) \leq \min_u \tilde{\mathcal{M}}(q(y), u). \quad (11)$$

The Majorization-Minimization algorithm repeatedly constructs an upper concave bound for $\mathcal{M}(q(y), u)$ that is iteratively tightening until convergence. In particular, it computes a sequence of parameters vector u^k that minimizes $\tilde{\mathcal{M}}(q(y), u)$ and then, for a fixed u^k , it creates the new estimation of $q(y)$, $q^k(y)$, that minimizes $C_{KL}(q(y)||p(y|Y, x_d))$. Hence the following algorithm can be used to find the estimations of $q(y)$ and u .

Algorithm 1 Majorization-Minimization algorithm for the estimation of the posterior distribution.

Given $u^1 \in (R^+)^{Bp}$, for $k = 1, 2, \dots$

1. Find

$$q^k(y) = \arg \min_{q(y)} \int q(y) \times \log\left(\frac{q(y)}{F(y, Y, x_d, u^k)}\right) dy. \quad (12)$$

2. Find

$$u^{k+1} = \arg \min_u \int q^k(y) \times \log\left(\frac{q^k(y)}{F(y, Y, x_d, u)}\right) dy. \quad (13)$$

until a stopping criterion is met.

The solution of Eq. (13) is given by

$$u_b^{k+1}(i) = E_{q^k(y)} \left[(\Delta_i^h(y_b))^2 + (\Delta_i^v(y_b))^2 \right], \quad (14)$$

for $i = 1, \dots, p$. Note that u_b^{k+1} represents the local spatial activity of y_b and they will be high for pixels in a neighborhood with high level of detail, thus preserving the image structures, and low for zones with low spatial activity where the quadratic term in Eq. (9) will keep it smooth.

By differentiating the integral on the right hand side of Eq. (12) with respect to $q(y)$ and setting it equal to zero, we obtain

$$q^k(y) = \mathcal{N}(y|E_{q^k(y)}[y], cov_{q^k(y)}[y]), \quad (15)$$

with

$$cov_{q^k(y)}[y] = \mathcal{A}^{-1}(u^k) \quad (16)$$

and

$$E_{q^k(y)}[y] = cov_{q^k(y)}[y] \phi^k, \quad (17)$$

Measure	Band	NSCT in [4]	SR in [5]	Proposed
COR	R	0.91	0.84	0.97
	G	0.91	0.98	0.97
	B	0.90	0.62	0.95
SSIM	R	0.79	0.90	0.96
	G	0.81	0.94	0.96
	B	0.81	0.85	0.95
PSNR	R	26.75	32.68	37.25
	G	27.17	35.50	37.51
	B	27.65	30.15	36.13
ERGAS	-	5.76	3.12	1.85

Table 1. Synthetic Image Quantative Results

where ϕ^k is the $(B \times p) \times 1$ vector,

$$\phi^k = (\text{diag}(\beta) \otimes H^t)Y + \frac{\gamma}{B}(1_B \otimes \sum_j C_j^t x_d), \quad (18)$$

where 1_B is the column vector of size $1 \times B$ with all its elements equal to one, and

$$\mathcal{A} = \begin{pmatrix} \alpha_1 \zeta(u_1^k) & 0_p & \cdots & 0_p \\ 0_p & \alpha_2 \zeta(u_2^k) & \cdots & 0_p \\ \vdots & \vdots & \ddots & \vdots \\ 0_p & 0_p & \cdots & \alpha_B \zeta(u_B^k) \end{pmatrix} + \text{diag}(\beta) \otimes H^t H + \frac{\gamma}{B^2}(1_{B \times B} \otimes \sum_{j,k} C_j^t C_k), \quad (19)$$

where $1_{B \times B}$ is a $B \times B$ matrix with all its elements equal to one, I_B is a $B \times B$ identity matrix, \otimes is the Kronecker product, $\beta = (\beta_1, \beta_2, \dots, \beta_B)^t$ and

$$\zeta(u_b^k) = (\Delta^h)^t W(u_b^k) (\Delta^h) + (\Delta^v)^t W(u_b^k) (\Delta^v), \quad (20)$$

for $b = 1, \dots, B$, where Δ^h and Δ^v represent $p \times p$ convolution matrices associated with the first order horizontal and vertical differences, respectively, and

$$W(u_b^k) = \text{diag} \left(u_b^k(i)^{-\frac{1}{2}} \right), \quad (21)$$

is a $p \times p$ diagonal matrix, for $i = 1, \dots, p$. This is a spatial adaptivity matrix since it controls the amount of smoothing at each pixel location depending on the strength of the intensity variation at that pixel, as expressed by the horizontal and vertical intensity gradient [5].

5. EXPERIMENTAL RESULTS

In order to test the proposed method, we used both a synthetic color image and a real SPOT5 image. We compared the proposed SR using contourlets method with the SR method in [5] and the additive NSCT method [4]. To assess the spatial improvement of the pansharpened images we use the correlation of the high frequency components (COR) [1] which takes values between zero and one (the higher the value the better the quality of the pansharpened image). Spectral fidelity was assessed by means of the peak signal-to-noise ratio (PSNR), the Structural Similarity Index Measure (SSIM) [8], an index ranging from -1 to $+1$ with $+1$ corresponding to exactly equal images, and the *erreur relative globale adimensionnelle de*

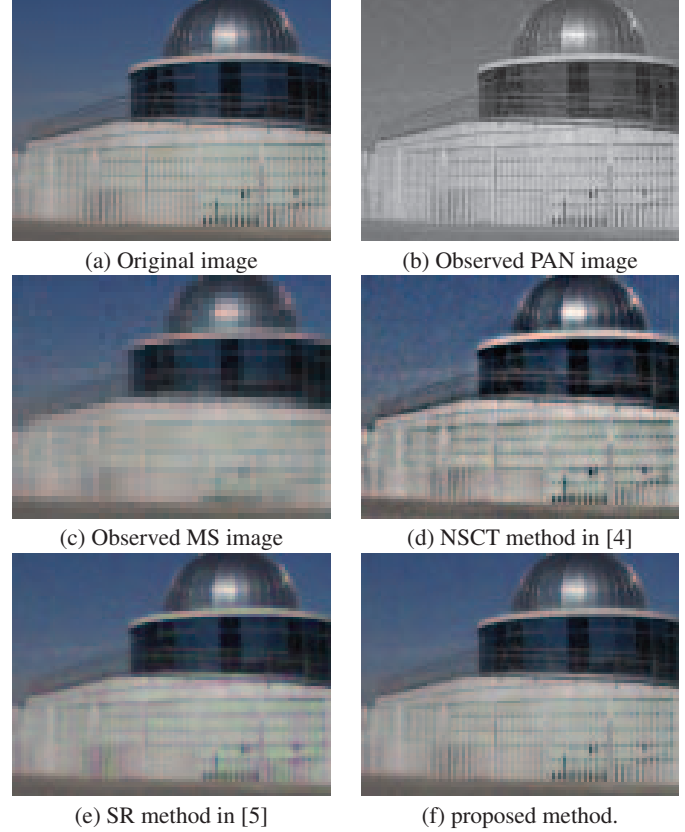


Fig. 1. Results for the synthetic image

synthèse (ERGAS) [9] index, a global criterion for what the lower the value, specially a value lower than the number of bands in the image, the higher the quality of the pansharpened image.

In order to conduct experiments where the ground truth is known, we used synthetic multispectral observations, obtained from the color image, displayed in Fig. 1(a), by convolving it with mask $0.25 \times 1_{2 \times 2}$ to simulate sensor integration, and then downsampling it by a factor of two by discarding every other pixel in each direction and adding zero mean Gaussian noise with variance 16. For the PAN image we used the luminance of the original color image and zero mean Gaussian noise of variance 9 was added. The observed PAN image and MS image, scaled to the size of the PAN image for displaying purposes, are shown in Fig. 1(b) and (c), respectively.

The proposed algorithm was run until the criterion $\|y^k - y^{k-1}\|^2 / \|y^{k-1}\|^2 < 10^{-4}$ was satisfied, where y^k denotes the mean of $q^k(y)$, which typically is reached within 5 iterations. The values of parameters were experimentally chosen to be $\alpha_b = 0.045$, $\beta_b = 1/16$, $b = 1, 2, 3$ and $\gamma = 1/9$. We are working on the method that it can estimate the parameters automatically. The resulted images corresponding to the reconstruction of the synthetic image using the NSCT method in [4], the SR method in [5], and the proposed method are displayed in Fig. 1(d)-(f), respectively, and Table 1 shows the corresponding quantitative results. The highlighted value in the table presents the highest value for each measure. The proposed method provides better results for each measure except for the COR of band 2 (green) where a very high value, similar to the one obtained by the method in [5] is obtained. The COR values

reflect that all methods are able to incorporate the details of the PAN image into the pansharpened one, although the SR method in [5], see Fig. 1(e), introduced less details in the band 3 (blue) since the blue band contributes only a 10% to the PAN image and more into the band 2 since it has the highest contribution, a 60%, which is reflected as a greenish color near the edges of the image. The NSCT method in [4] incorporates details in all the bands but produces a noisy image, see Fig. 1(d). The proposed method (Fig. 1(f)) is able to incorporate detail in all the bands while controlling the noise. The spectral fidelity measures show that the proposed method performs better than the competing method, which is also clear from the image in Fig. 1(f), producing an image that is not as noisy as the NSCT method in [4] (Fig. 1(d)) and preserves better the colors than the SR method in [5] (Fig. 1(e)), while better controlling the noise. It is remarkable the high SSIM and low ERGAS values which reflect the high quality of the resulting images. Note also that the PSNR for the proposed method is about 10dB higher than NSCT method in [4] and from 2 to almost 6 dB higher than for the SR method in [5].

In a second experiment, the method was tested on real SPOT5 dataset, where the MS image covers a region of interest of 256 by 256 pixels with pixel resolution of 10 m, while the PAN image is 512 by 512 pixels with a pixel resolution of 5 m. The MS image consists of four bands from the visible and infrared region corresponding to green (b1: 0.50-0.59 μm), red (b2: 0.61-0.68 μm), Near IR (b3: 0.78-0.89 μm), Mid IR (b4: 1.58-1.75 μm), while the PAN image consists of a single band covering the visible and NIR (0.48-0.71 μm). Figure 2(a) shows a region of the RGB color image representing bands 1 to 3 of the MS image. Its corresponding PAN image is depicted in Fig. 2(b).

Visual inspections of the resulting images, displayed in Figs. 2(c)-(f), reveals similar conclusions to the obtained for the synthetic image. The proposed method provides the best result, preserving the spectral properties of MS image while incorporating the high frequencies from the panchromatic image and controlling the noise in the image. Figure 2(c) depicts the bicubic interpolation of the MS image and although the colors are perfect, no detail is appreciated. The NSCT method in [4] (Fig.2(d)) provides a detailed image but quite noisy, the SR method in [5] provides good details for bands 1 and 2, see Fig.2(e), but not for bands 3 and 4 since the PAN image does not cover those bands. This is why the blue color in Fig.2(e), seems to be vanished. The proposed method successfully preserves the colors, incorporates the details from the PAN image into the pansharpened image and controls the noise in the images.

6. CONCLUSIONS

In this paper, a new pansharpening method based on super-resolution reconstruction and non subsampled contourlet transform has been presented. The proposed method preserves the spectral properties of MS image while incorporating the high frequencies from the panchromatic image and controlling the noise in the image.

The efficiency of pansharpening methods has been evaluated by means of visual and quantitative analysis, for synthetic and real data. Based on the presented experiments, the proposed method does significantly outperform NSCT-based and TV-based super-resolution methods.

7. REFERENCES

[1] V. Vijayaraj, "A quantitative analysis of pansharpened images," M.S. thesis, Mississippi State University, August 2004.

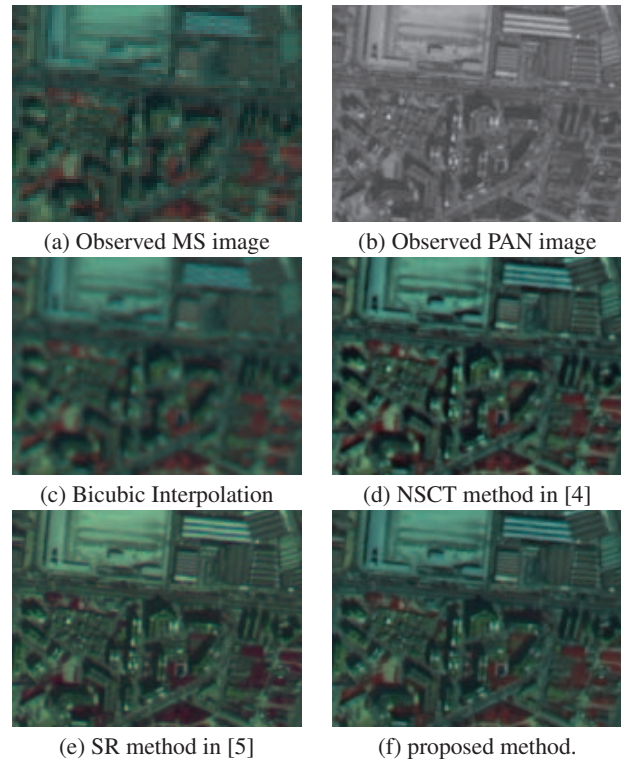


Fig. 2. Results for the SPOT5 image

- [2] L. Alparone, L. Wald, J. Chanussot, P. Gamba, and L. M. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 10, pp. 3012–3021, 2007.
- [3] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampling contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, 2006.
- [4] M. Lillo-Saavedra and C. Gonzalo, "Multispectral images fusion by a joint multidirectional and multiresolution representation," *Int. J. Remote Sens.*, vol. 28, no. 18, pp. 4065 – 4079, 2007.
- [5] M. Vega, J. Mateos, R. Molina, and A.K. Katsaggelos, "Super resolution of multispectral images using TV image models," in *21th Intern. Conf. on Knowledge-Based and Intelligent Information & Engineering Systems*, I. Lovrek, R.J. Howlett, and L.C. Jain, Eds., 2008, vol. LNAI-5179, Part III, pp. 408–415.
- [6] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, pp. 79–86, 1951.
- [7] N. Galatsanos, "A majorization-minimization approach to total variation reconstruction of super-resolved images," in *16th European Signal Processing Conf. (EUSIPCO)*, August 2008.
- [8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600 – 612, 2004.
- [9] L. Wald, T. Ranchin, and M. Mangolini, "Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images," *Photogramm. Eng. Remote Sens.*, vol. 63, no. 6, pp. 691 – 699, June 1997.