Pansharpening of multispectral images using a TV-based super-resolution algorithm

Javier Mateos[1], **Miguel Vega**[1], **Rafael Molina**[1] **and Aggelos K. Katsaggelos**[2]

[1] Universidad de Granada, Granada 18071, SPAIN.
[2] Northwestern University, Evanston, IL 60208, USA.

E-mail: `jmd@decsai.ugr.es`

Abstract. In this paper we propose a novel algorithm for the pansharpening of multispectral images based on the use of a Total Variation (TV) image prior. Within the Bayesian formulation, the proposed methodology incorporates prior knowledge on the expected characteristics of multispectral images, and uses the sensor characteristics to model the observation process of both panchromatic and multispectral images. The pansharpened multispectral images are compared with the images obtained by other parsharpening methods and their quality is assessed both qualitatively and quantitatively.

1. Introduction

Remote sensing systems include sensors able to capture, simultaneously, several low resolution (LR) images of the same area on different wavelengths, forming a multispectral image, along with a high resolution (HR) panchromatic image. The technique called pansharpening is a multispectral image reconstruction technique that jointly process the multispectral and panchromatic images in order to obtain a new multispectral image that, ideally, exhibits the spectral characteristics of the observed multispectral image and the resolution of the panchromatic image. In the literature, a number of pansharpening methods has been proposed (see, for instance, [1, 2] and the comparison of algorithms in [3]) some of them based on super-resolution techniques [4]. Recently a new Bayesian variational framework for total variation (TV) based image restoration problems has been presented [5]. Here we explore and adapt this TV framework and propose a new super-resolution based reconstruction method to the pansharpening of multispectral images.

The paper is organized as follows. In section 2 the Bayesian modeling and inference for super-resolution reconstruction of multispectral images is presented. Section 3 describes the variational approximation of the posterior distribution of the HR multispectral image and unknown hyperparameters and how inference is performed. Section 4 presents experimental results and section 5 concludes the paper.

2. Bayesian Modeling and Inference

Let us assume that $\mathbf{y}$, the unknown HR multispectral image we would have observed under ideal conditions, has B bands $\mathbf{y}_b$, $b = 1, \ldots, B$, each of size $p = m \times n$, that is, $\mathbf{y} = [\mathbf{y}_1^t, \mathbf{y}_2^t, \ldots, \mathbf{y}_B^t]^t$, 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 LR multispectral image $\mathbf{Y}$ has B bands $\mathbf{Y}_b$, $b = 1, \ldots, B$, each of size $P = M \times N$ pixels, with $M < m$ and $N < n$. These images are also stacked into the vector $\mathbf{Y} = [\mathbf{Y}_1^t, \mathbf{Y}_2^t, \ldots, \mathbf{Y}_B^t]^t$, 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 $\mathbf{x}$ of size $p = m \times n$, obtained by spectrally averaging the HR images $\mathbf{y}_b$.

The objective of the HR multispectral image reconstruction is to obtain an estimate of the unknown HR multispectral image $\mathbf{y}$ given the panchromatic HR observation $\mathbf{x}$ and the LR multispectral observation $\mathbf{Y}$. The Bayesian formulation of this problem requires the definition of the joint distribution $\mathrm{p}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x})$. We define this joint distribution as $\mathrm{p}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}) = \mathrm{p}(\Omega)\mathrm{p}(\mathbf{y}|\Omega)\mathrm{p}(\mathbf{Y}, \mathbf{x}|\mathbf{y}, \Omega)$, where Ω denotes the set of hyperparameters needed to describe the required probability density functions, and inference is based on $\mathrm{p}(\Omega, \mathbf{y}|\mathbf{Y}, \mathbf{x})$. Let us now describe those probability distributions.

In this paper we use a TV image prior [6] for each band (the correlation among HR bands is not taken into account), thus defining the multispectral image prior

$$\mathrm{p}(\mathbf{y}|\Omega) = \prod_{b=1}^{B} \mathrm{p}(\mathbf{y}_b|\alpha_b) \propto \prod_{b=1}^{B} \alpha_b^{p/2} \exp\left[-\alpha_b \mathrm{TV}(\mathbf{y}_b)\right] , \quad (1)$$

with $\mathrm{TV}(\mathbf{y}_b) = \sum_{i=1}^{p} \sqrt{(\Delta_i^h(\mathbf{y}_b))^2 + (\Delta_i^v(\mathbf{y}_b))^2}$, where $\Delta_i^h(\mathbf{y}_b)$ and $\Delta_i^v(\mathbf{y}_b)$ represent the horizontal and vertical first order differences at pixel i respectively, α_b is the model parameter of the band b, and the partition function has been approximated using the approach in [5].

We assume that $\mathbf{Y}$ and $\mathbf{x}$, for a given $\mathbf{y}$, are independent and write $\mathrm{p}(\mathbf{Y}, \mathbf{x}|\mathbf{y}, \Omega) = \mathrm{p}(\mathbf{Y}|\mathbf{y}, \Omega)\mathrm{p}(\mathbf{x}|\mathbf{y}, \Omega)$.

For each multispectral image band, we consider the model $\mathbf{Y}_b = \mathbf{H}\mathbf{y}_b + \mathbf{n}_b$, $b = 1, \ldots, B$, where the degradation matrix $\mathbf{H}$ can be written as $\mathbf{H} = \mathbf{D}\mathbf{B}$, with $\mathbf{B}$ a $p \times p$ blurring matrix and $\mathbf{D}$ a $P \times p$ decimation operator, and $\mathbf{n}_b$ is the noise term assumed to be independent white Gaussian of known variance β_b^{-1}. The distribution of the observed image $\mathbf{Y}$ given $\mathbf{y}$ is

$$\mathrm{p}(\mathbf{Y}|\mathbf{y}, \Omega) = \prod_{b=1}^{B} \mathrm{p}(\mathbf{Y}_b|\mathbf{y}_b, \beta_b) \propto \prod_{b=1}^{B} {\beta_b}^{P/2} \exp\left\{-\frac{1}{2}\beta_b \parallel \mathbf{Y}_b - \mathbf{H}\mathbf{y}_b \parallel^2\right\} . \quad (2)$$

The panchromatic image $\mathbf{x}$ is modeled as $\mathbf{x} = \sum_{b=1}^{B} \lambda_b \mathbf{y}_b + \mathbf{v}$, where $\lambda_b \geq 0$, $b = 1, 2, \ldots, B$, are known quantities that can be obtained from the sensor spectral characteristics, and $\mathbf{v}$ is the capture noise that is assumed to be Gaussian with zero mean and known variance γ^{-1}. Based on this model, the distribution of the panchromatic image $\mathbf{x}$ given $\mathbf{y}$, then becomes

$$\mathrm{p}(\mathbf{x}|\mathbf{y}, \gamma) \propto \gamma^{p/2} \exp\left\{-\frac{1}{2}\gamma \parallel \mathbf{x} - \sum_{b=1}^{B} \lambda_b \mathbf{y}_b \parallel^2\right\} . \quad (3)$$

Although the estimation of the parameter vector $(\beta_1, \ldots, \beta_B, \gamma)$ can be incorporated in the estimation process to be described next, we will assume here that these parameters are known or have been estimated in advance and concentrate on gaining insight into the estimation of the prior parameters. The set of hyperparameters then becomes $\Omega = (\alpha_1, \ldots, \alpha_B)$. We will assume the following distribution on the hyperparameters

$$\mathrm{p}(\Omega) = \prod_{b=1}^{B} \Gamma(\alpha_b \mid a_{\alpha_b}^o, c_{\alpha_b}^o) , \quad (4)$$

where each of the hyperparameters ω in Ω has as hyperprior the gamma distribution, that is, $\Gamma(\omega \mid a_\omega^o, c_\omega^o) \propto [\omega]^{a_\omega^o - 1} \exp[-c_\omega^o \omega]$, where $c_\omega^o > 0$ and $a_\omega^o > 0$. This gamma distribution has the mean $\mathbf{E}[\omega] = a_\omega^o / c_\omega^o$ and variance $\mathbf{var}[\omega] = a_\omega^o / (c_\omega^o)^2$.

3. Bayesian Inference and Variational Approximation of the Posterior Distribution

As already known, the Bayesian paradigm dictates that inference on $\mathbf{y}$ should be based on

$$\mathrm{p}(\Omega, \mathbf{y}|\mathbf{Y}, \mathbf{x}) = \mathrm{p}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x})/\mathrm{p}(\mathbf{Y}, \mathbf{x}). \tag{5}$$

with $\mathrm{p}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}) = \mathrm{p}(\Omega)\mathrm{p}(\mathbf{y}|\Omega)\mathrm{p}(\mathbf{Y}|\mathbf{y})\mathrm{p}(\mathbf{x}|\mathbf{y})$, where $\mathrm{p}(\Omega)$, $\mathrm{p}(\mathbf{y}|\Omega)$, $\mathrm{p}(\mathbf{Y}|\mathbf{y})$ and $\mathrm{p}(\mathbf{x}|\mathbf{y})$ have been defined in Eqs. (4), (1), (2) and (3), respectively.

Since $\mathrm{p}(\Omega, \mathbf{y}|\mathbf{Y}, \mathbf{x})$ can not be found in closed form, we apply variational methods to approximate this distribution by a distribution $\mathrm{q}(\Omega, \mathbf{y})$. The variational criterion used to find $\mathrm{q}(\Omega, \mathbf{y})$ is the minimization of the Kullback-Leibler (KL) divergence, given by [7]

$$\begin{aligned} C_{KL}(\mathrm{q}(\Omega, \mathbf{y})||\mathrm{p}(\Omega, \mathbf{y}|\mathbf{Y}, \mathbf{x})) &= \int \mathrm{q}(\Omega, \mathbf{y}) \log\left(\frac{\mathrm{q}(\Omega, \mathbf{y})}{\mathrm{p}(\Omega, \mathbf{y}|\mathbf{Y}, \mathbf{x})}\right) d\Omega d\mathbf{y} \\ &= \int \mathrm{q}(\Omega, \mathbf{y}) \log\left(\frac{\mathrm{q}(\Omega, \mathbf{y})}{\mathrm{p}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x})}\right) d\Omega d\mathbf{y} + \mathrm{const} \\ &= \mathcal{M}(\mathrm{q}(\Omega, \mathbf{y})) + \mathrm{const}\,, \end{aligned} \tag{6}$$

which is always non negative and equal to zero only when $\mathrm{q}(\Omega, \mathbf{y}) = \mathrm{p}(\Omega, \mathbf{y}|\mathbf{Y}, \mathbf{x})$.

Due to the form of the TV prior, the above integral can not be evaluated. We can however majorize the TV prior by a function which renders the integral easier to calculate. Let us consider the following inequality, also used in [8], which states that, for any $w \geq 0$ and $z > 0$

$$\sqrt{w} \leq \frac{w + z}{2\sqrt{z}}\,. \tag{7}$$

Let us define, for $\mathbf{y}_b$ and $\mathbf{u}_b$, where $\mathbf{u}_b$ is any p-dimensional vector $\mathbf{u}_b \in (R^+)^p$, with components $\mathbf{u}_b(i)$, $i = 1, \dots, p$, the following functional

$$\mathrm{M}(\alpha_b, \mathbf{y}_b, \mathbf{u}_b) = \alpha_b^{p/2} \exp\left[-\frac{\alpha_b}{2} \sum_i \frac{(\Delta_i^h(\mathbf{y}_b))^2 + (\Delta_i^v(\mathbf{y}_b))^2 + \mathbf{u}_b(i)}{\sqrt{\mathbf{u}_b(i)}}\right]. \tag{8}$$

Now, using the inequality in Eq. (7) with $w = (\Delta_i^h(\mathbf{y}_b))^2 + (\Delta_i^v(\mathbf{y}_b))^2$ and $z = \mathbf{u}_b(i)$, and comparing Eq. (8) with Eq. (1), we obtain $\mathrm{p}(\mathbf{y}|\Omega) \geq c \cdot \prod_{b=1}^{B} \mathrm{M}(\alpha_b, \mathbf{y}_b, \mathbf{u}_b)$. This leads to the following lower bound for the joint probability distribution

$$\mathrm{p}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}) \geq c \cdot \mathrm{p}(\Omega) \left[\prod_{b=1}^{B} \mathrm{M}(\alpha_b, \mathbf{y}_b, \mathbf{u}_b)\right] \mathrm{p}(\mathbf{Y}|\mathbf{y})\mathrm{p}(\mathbf{x}|\mathbf{y}) = \mathrm{F}(\mathrm{p}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}, \mathbf{u})\,, \tag{9}$$

where $\mathbf{u} = [\mathbf{u}_1^t, \mathbf{u}_2^t, \dots, \mathbf{u}_B^t]^t$.

Hence, by defining

$$\tilde{\mathcal{M}}(\mathrm{q}(\Omega, \mathbf{y}), \mathbf{u}) = \int \mathrm{q}(\Omega, \mathbf{y}) \log\left(\frac{\mathrm{q}(\Omega, \mathbf{y})}{\mathrm{F}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}, \mathbf{u})}\right) d\Omega d\mathbf{y}, \tag{10}$$

and using Eq. (9), we obtain

$$\mathcal{M}(\mathrm{q}(\Omega, \mathbf{y})) \leq \min_{\mathbf{u}} \tilde{\mathcal{M}}(\mathrm{q}(\Omega, \mathbf{y}), \mathbf{u})\,. \tag{11}$$

Therefore, by finding a sequence of distributions $\{\mathrm{q}^k(\Omega, \mathbf{y})\}$ that monotonically decreases $\tilde{\mathcal{M}}(\mathrm{q}(\Omega, \mathbf{y}), \mathbf{u})$ for a fixed $\mathbf{u}$, a sequence of an ever decreasing upper bound of $C_{KL}(\mathrm{q}(\Omega, \mathbf{y})||\mathrm{p}(\Omega, \mathbf{y}|\mathbf{Y}, \mathbf{x}))$ is also obtained due to Eq. (6). However, also minimizing $\mathcal{M}(\mathrm{q}(\mathbf{y}))$

with respect to $\mathbf{u}$, generates a sequence of vectors $\{\mathbf{u}^k\}$ that tightens the upper-bound for each distribution $\mathrm{q}^k(\Omega, \mathbf{y})$. Therefore, the two sequences $\{\mathrm{q}^k(\Omega, \mathbf{y})\}$ and $\{\mathbf{u}^k\}$ are coupled. We develop the following iterative algorithm to find such sequences for calculating the approximating posteriors $\mathrm{q}(\Omega, \mathbf{y}) = \mathrm{q}(\Omega)\mathrm{q}(\mathbf{y})$. We note that the process to find the best posterior distribution approximation of the image in combination with $\mathbf{u}$ is a very natural extension of the Majorization-Minimization approach to function optimization [9].

Algorithm 1 *Posterior image distribution estimation.*

Given $\mathbf{u}^1 \in (R^+)^{Bp}$, *for* $k = 1, 2, \ldots$ *until a stopping criterion is met:*
Find

$$\begin{aligned}
\mathrm{q}^k(\mathbf{y}) &= \arg\min_{\mathrm{q}(\mathbf{y})} \int \mathrm{q}^k(\Omega)\, \mathrm{q}(\mathbf{y}) \times \log\left(\frac{\mathrm{q}^k(\Omega)\, \mathrm{q}(\mathbf{y})}{\mathrm{F}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}, \mathbf{u}^k)}\right) d\Omega d\mathbf{y}\,, && (12)\\
\mathbf{u}^{k+1} &= \arg\min_{\mathbf{u}} \int \mathrm{q}^k(\Omega)\, \mathrm{q}^k(\mathbf{y}) \times \log\left(\frac{\mathrm{q}^k(\Omega)\, \mathrm{q}^k(\mathbf{y})}{\mathrm{F}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}, \mathbf{u})}\right) d\Omega d\mathbf{y}\,, && (13)\\
\mathrm{q}^{k+1}(\Omega) &= \arg\min_{\mathrm{q}(\Omega)} \int \mathrm{q}(\Omega)\, \mathrm{q}^k(\mathbf{y}) \times \log\left(\frac{\mathrm{q}(\Omega)\, \mathrm{q}^k(\mathbf{y})}{\mathrm{F}(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}, \mathbf{u}^{k+1})}\right) d\Omega d\mathbf{y}\,. && (14)
\end{aligned}$$

Set $\mathrm{q}(\Omega) = \lim_{k\leftarrow\infty} \mathrm{q}^k(\Omega)$ *and* $\mathrm{q}(\mathbf{y}) = \lim_{k\leftarrow\infty} \mathrm{q}^k(\mathbf{y})$.

To calculate $\mathbf{u}_b^{k+1}$, for $b = 1, \ldots, B$, we have from Eq. (14) that

$$\mathbf{u}_b^{k+1} = \arg\min_{\mathbf{u}_b} \sum_i \frac{\mathbf{E}_{\mathrm{q}^k(\mathbf{y})}\left[\Delta_i^h(\mathbf{y}_b))^2 + (\Delta_i^v(\mathbf{y}_b))^2\right] + \mathbf{u}_b(i)}{\sqrt{\mathbf{u}_b(i)}}\,, \quad (15)$$

and consequently $\mathbf{u}_b^{k+1}(i) = \mathbf{E}_{\mathrm{q}^k(\mathbf{y})}\left[\Delta_i^h(\mathbf{y}_b))^2 + (\Delta_i^v(\mathbf{y}_b))^2\right]$, for $i = 1, \ldots, p$. It is clear from this equation that the vector $\mathbf{u}_b^{k+1}$ is a function of the spatial first order differences of the unknown image $\mathbf{y}$ under the distribution $\mathrm{q}^k(\mathbf{y})$ and represents the local spatial activity of $\mathbf{y}_b$.

To calculate $\mathrm{q}^k(\mathbf{y})$, we observe that differentiating the integral on the right-hand side of Eq. (13) with respect to $\mathrm{q}(\mathbf{y})$ and setting it equal to zero, we obtain that

$$\mathrm{q}^k(\mathbf{y}) = \mathcal{N}\left(\mathbf{y} \mid \mathbf{E}_{q^k(\mathbf{y})}[\mathbf{y}], \mathbf{cov}_{q^k(\mathbf{y})}[\mathbf{y}]\right), \quad (16)$$

with

$$\mathbf{cov}_{q^k(\mathbf{y})}[\mathbf{y}] = \mathcal{A}^{-1}(\Omega, \mathbf{u}^k)\,, \quad \text{and} \quad \mathbf{E}_{q^k(\mathbf{y})}[\mathbf{y}] = \mathbf{cov}_{q^k(\mathbf{y})}[\mathbf{y}]\phi^k\,, \quad (17)$$

where ϕ^k is the $(B \times p) \times 1$ vector $\phi^k = \left(\mathrm{diag}(\beta) \otimes \mathbf{H}^t\right) \mathbf{Y} + \gamma\left(\mathrm{diag}(\lambda) \otimes \mathbf{I}_p\right)\left(\mathbf{x}^t, \mathbf{x}^t, \ldots, \mathbf{x}^t\right)^t$, and

$$\mathcal{A}(\Omega, \mathbf{u}^k) = \begin{pmatrix} \alpha_1\mathcal{G}(\mathbf{u}_1^k) & \mathbf{0}_p & \cdots & \mathbf{0}_p \\ \mathbf{0}_p & \alpha_2\mathcal{G}(\mathbf{u}_2^k) & \cdots & \mathbf{0}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_p & \mathbf{0}_p & \cdots & \alpha_B\mathcal{G}(\mathbf{u}_B^k) \end{pmatrix} + \mathrm{diag}(\beta) \otimes \mathbf{H}^t\mathbf{H} + \gamma(\lambda\lambda^t) \otimes \mathbf{I}_p\,,$$

where $\otimes$ is the Kronecker product, $\beta = (\beta_1, \beta_2, \ldots, \beta_B)^t$, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_B)^t$, and

$$\mathcal{G}(\mathbf{u}_b^k) = (\Delta^h)^t W(\mathbf{u}_b^k)(\Delta^h) + (\Delta^v)^t W(\mathbf{u}_b^k)(\Delta^v)\,, \quad \text{for} \quad b = 1, \ldots, B\,, \quad (18)$$

where Δ^h and Δ^v represent $p \times p$ convolution matrices associated with the first order horizontal and vertical differences, respectively, and $W(\mathbf{u}_b^k)$ is a $p \times p$ diagonal matrix of the form $W(\mathbf{u}_b^k) = \mathrm{diag}\left(\mathbf{u}_b^k(i)^{-\frac{1}{2}}\right)$, for $i = 1, \ldots, p$. This matrix $W(\mathbf{u}_b^k)$ can be interpreted as 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.

By differentiating the integral on the right hand side of Eq. (14) with respect to $\mathrm{q}(\Omega)$ and setting it equal to zero we obtain

$$\mathrm{q}^{k+1}(\Omega) = \prod_{b=1}^{B} \mathrm{q}^{k+1}(\alpha_b) = \prod_{b=1}^{B} \Gamma\left(\alpha_b | a^o_{\alpha_b} + \frac{p}{2}, c^o_{\alpha_b} + \sum_i^p \sqrt{\mathbf{u}_b^{k+1}(i)}\right). \tag{19}$$

For the means of these distributions the following expressions can be found

$$\frac{1}{\mathbf{E}_{q^{k+1}(\Omega)}[\alpha_b]} \;=\; \mu_{\alpha_b}\frac{1}{\bar{\alpha}^o_b} + (1-\mu_{\alpha_b})\frac{2\sum_i^p \sqrt{\mathbf{u}_b^{k+1}(i)}}{p},\ b = 1, \ldots, B, \tag{20}$$

where $\bar{\alpha}^o_b = \frac{a^o_{\alpha_b}}{c^o_{\alpha_b}}$ and $\mu_{\alpha_b} = \frac{a^o_{\alpha_b}}{p/2 + a^o_{\alpha_b}}$ for $b = 1 \ldots, B$. The above equations indicate that μ_{α_b}, $b = 1, \ldots, B$, can be understood as normalized confidence parameters taking values in the interval $[0, 1)$. That is, when they are zero no confidence is placed on the given hyperparameters, while when the corresponding normalized confidence parameter is asymptotically equal to one it fully enforces the prior knowledge of the mean (no parameter estimation is performed).

4. Experimental Results

Results are presented on a multispectral LandSat ETM+ image. We simulate a multispectral image and a panchromatic image by convolving the original multispectral image (depicted in Fig. 1(a)) and its corresponing panchromatic image with the mask $0.25 \times \mathbf{1}_{2\times 2}$ and downsampling them by a factor of two in each direction by discarding every other pixel, obtaining the observed multispectral and panchromatic images shown in Fig. 1(b) and 1(c), respectively.

The proposed algorithm was ran until the criterion $\|\mathbf{y}^k - \mathbf{y}^{k-1}\|^2/\|\mathbf{y}^{k-1}\|^2 < 10^{-4}$ was satisfied, where $\mathbf{y}^k$ denotes the mean of $\mathrm{q}^k(\mathbf{y})$.The proposed algorithm required 5 iterations to converge. According to the ETM+ sensor spectral response, the panchromatic image covers only the spectrum of a part of the first four bands of the multispectral image. Hence, we apply the proposed method with $B = 4$. The values of λ_b, $b = 1, 2, 3, 4$, calculated from the spectral response of the ETM+ sensor, are equal to 0.0078, 0.2420, 0.2239, and 0.5263, for bands one to four, respectively [4]. The value of the parameters β_b, $b = 1, 2, 3, 4$ and γ where estimated by the method in [4] since it uses the same degradation models than the proposed method. The method in [4], however, cannot provide with accurate information for the parameters α_b, $b = 1, \ldots, 4$ and, hence, their values are estimated using the proposed algorithm with $\mu_{\alpha_b} = 0$, $b = 1, \ldots, 4$.

The quality of the reconstruction is assessed by means of the peak signal-to-noise ratio (PSNR) between each band of the reconstructed and original multispectral images, and the standard ERGAS index [10]. The lower the value of this index the higher the quality of the multispectral reconstructed image. Table 1 shows the resulting PSNR and ERGAS values for the reconstructions of the image using the proposed method, the pansharpening method in [2] and the pansharpening method in [4]. The reconstructed images corresponding to those methods are displayed in Fig. 1(d)–1(f). The proposed model produces better results recovering the HR structures of the original image while reducing at the same time the amount of noise.

Table 1. Values of PSNR,and ERGAS for the image in Fig. 1.

	PSNR				ERGAS
Method / Band	1	2	3	4	
Method in [2]	31.39	28.97	25.64	30.41	6.91
Method in [4]	34.06	31.59	21.28	32.34	8.80
Proposed model	33.79	31.91	28.81	27.31	5.99

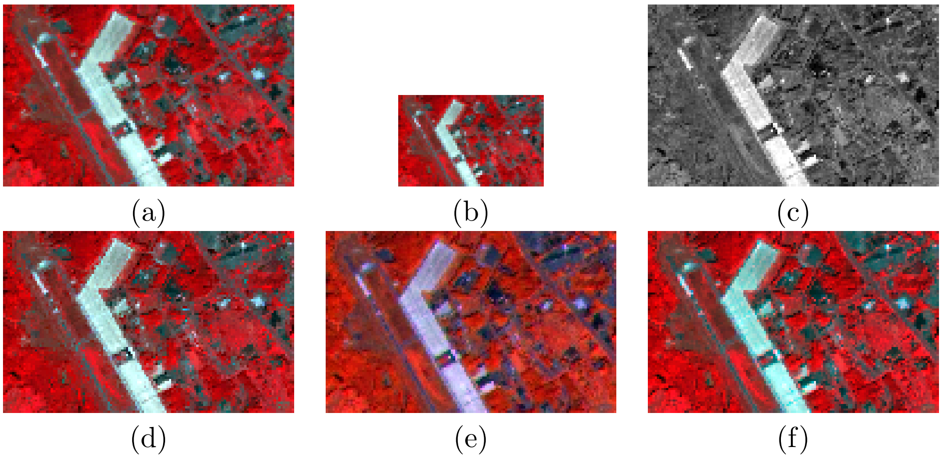

Figure 1. (a) False RGB color image composed of the LR bands 3, 4, and 2 of the original multispectral image; (b) Observed (simulated) multispectral image; (c) Observed (simulated) panchromatic image; (d) Reconstruction using the method in [2]; (e) Reconstruction using the method in [4]; (f) Reconstruction using the proposed model.

5. Conclusions

We have presented a new method for TV-based pansharpening of multispectral images using a super-resolution approach. The proposed method takes into account the sensor characteristics in the image formation model. We have used the variational approach to approximate the posterior distribution of the pansharpened multispectral image. Based on the presented experimental results, the proposed method performs better the methods in [2] and [4].

Acknowledgments

This work has been supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIN2007-65533 and the Consejería de Innovación, Ciencia y Empresa of the Junta de Andalucía under contract P07-TIC-02698.

References

[1] Nuñez J, Otazu X, Fors O, Prades A, Pala V and Arbiol R 1999 *IEEE Trans on Geosc. & Rem. Sens.* **37** 1204–1211
[2] Price J 1999 *IEEE Trans. on Geosc. & Rem. Sens.* **37** 1199–1203
[3] Alparone L, Wald L, Chanussot J, Thomas C, Gamba P and Bruce L 2007 *IEEE Trans. on Geosc. & Rem. Sens.* **45** 3012–3020
[4] Molina R, Vega M, Mateos J and Katsaggelos A 2008 *Applied and Comp. Harmonic Analysis* **24** 251–267
[5] Babacan S, Molina R and Katsaggelos A 2008 *IEEE Trans. on Image Proc.* **17** 326–339
[6] Rudin L I, Osher S, and Fatemii E 1992 *Physica D* **63** 259–268
[7] Kullback S and Leibler R A 1951 *Annals of Math. Stat.* **22** 79–86
[8] Bioucas-Dias J, Figueiredo M and Oliveira J 2006 *ICASSP'2006*
[9] Lange K 2004 *Springer texts in Statistic* (New York, Springer Verlag)
[10] Wald L, Ranchin T and Mangolini M 1997 *Phot. Eng. Rem. Sens.* **63** 691–699