

Super Resolution of Multispectral Images using TV Image Models

Miguel Vega^{1*}, Javier Mateos², Rafael Molina², and Aggelos K. Katsaggelos³

¹ Dpto. de Lenguajes y Sistemas Informáticos,
Universidad de Granada, Granada 18071, SPAIN,

² Dpto. de Ciencias de la Computación e Inteligencia Artificial,
Universidad de Granada, Granada 18071, SPAIN,

³ Department of Electrical and Computer Engineering,
Northwestern University, Evanston, IL 60208, USA

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. Using real and synthetic data, the pansharpened multispectral images are compared with the images obtained by other pansharpening methods and their quality is assessed both qualitatively and quantitatively.

1 Introduction

Remote sensing systems include sensors able to capture, simultaneously, several low resolution images of the same area on different wavelengths, forming a multispectral image, along with a high resolution panchromatic image. Throughout this paper the term *multispectral image reconstruction* will refer to the joint processing of 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.

A few approaches to the super resolution of multispectral images problem, also called pansharpening, have been proposed in the literature. See, for instance, [1, 2] and the comparison of algorithms in [3]. Recently a new Bayesian variational framework for TV-based image restoration problems has been presented [4]. Here we explore and adapt this framework 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

* 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.

the high resolution multispectral image 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 high resolution multispectral image we would have observed under ideal conditions, has B bands \mathbf{y}_b , $b = 1, \dots, B$, each of size $p = m \times n$, that is, $\mathbf{y} = [\mathbf{y}_1^t, \mathbf{y}_2^t, \dots, \mathbf{y}_B^t]^t$, 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 low resolution multispectral image \mathbf{Y} has B bands \mathbf{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 $\mathbf{Y} = [\mathbf{Y}_1^t, \mathbf{Y}_2^t, \dots, \mathbf{Y}_B^t]^t$, 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 high resolution images \mathbf{y}_b .

The objective of the high resolution multispectral image reconstruction problem is to obtain an estimate of the unknown high resolution multispectral image \mathbf{y} given the panchromatic high resolution observation \mathbf{x} and the low resolution multispectral observation \mathbf{Y} . The Bayesian formulation of this problem requires the definition of the joint distribution $p(\mathbf{y}, \mathbf{Y}, \mathbf{x})$. We define this joint distribution as $p(\mathbf{y}, \mathbf{Y}, \mathbf{x}) = p(\mathbf{y})p(\mathbf{Y}, \mathbf{x}|\mathbf{y})$ and inference is based on $p(\mathbf{y}|\mathbf{Y}, \mathbf{x})$. Let us now describe those probability distributions.

In this paper we use TV priors [5] for each band (the correlation among high resolution bands is not taken into account), thus defining the multispectral image prior

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

with $\text{TV}(\mathbf{y}_b) = \sum_{i=1}^p \sqrt{(\Delta_i^h(\mathbf{y}_b))^2 + (\Delta_i^v(\mathbf{y}_b))^2}$, 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 [4].

We assume that \mathbf{Y} and \mathbf{x} , for a given \mathbf{y} , are independent and write

$$p(\mathbf{Y}, \mathbf{x}|\mathbf{y}) = p(\mathbf{Y}|\mathbf{y})p(\mathbf{x}|\mathbf{y}). \quad (2)$$

For each multispectral image band, we consider the model $\mathbf{Y}_b = \mathbf{H}\mathbf{y}_b + \mathbf{n}_b$, $b = 1, \dots, B$, 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

$$p(\mathbf{Y}|\mathbf{y}) = \prod_{b=1}^B p(\mathbf{Y}_b|\mathbf{y}_b) \propto \prod_{b=1}^B \beta_b^{P/2} \exp\left\{-\frac{1}{2}\beta_b \|\mathbf{Y}_b - \mathbf{H}\mathbf{y}_b\|^2\right\}. \quad (3)$$

The panchromatic image \mathbf{x} is modeled as

$$\mathbf{x} = \sum_{b=1}^B \lambda_b \mathbf{y}_b + \mathbf{v},$$

where $\lambda_b \geq 0$, $b = 1, 2, \dots, B$, 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} , is given by

$$p(\mathbf{x}|\mathbf{y}) \propto \gamma^{p/2} \exp \left\{ -\frac{1}{2} \gamma \left\| \mathbf{x} - \sum_{b=1}^B \lambda_b \mathbf{y}_b \right\|^2 \right\}. \quad (4)$$

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

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

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

Since $p(\mathbf{y}|\mathbf{Y}, \mathbf{x})$ can not be found in closed form, we apply variational methods to approximate this distribution by a distribution $q(\mathbf{y})$. The variational criterion used to find $q(\mathbf{y})$ is the minimization of the Kullback-Leibler (KL) divergence, given by [6]

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

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

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 [7], which states that, for any $w \geq 0$ and $z > 0$

$$\sqrt{w} \leq \frac{w + z}{2\sqrt{z}}. \quad (7)$$

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

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

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 $p(\mathbf{y}) \geq c \cdot \prod_{b=1}^B M(\mathbf{y}_b, \mathbf{u}_b)$. As will be shown later, vectors \mathbf{u}_b are quantities that need to be computed and have an intuitive interpretation related to the unknown images \mathbf{y}_b . This leads to the following lower bound for the joint probability distribution

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

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

Hence, by defining

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

and using Eq. (9), we obtain

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

Therefore, by finding a sequence of distributions $\{q^k(\mathbf{y})\}$ that monotonically decreases $\tilde{\mathcal{M}}(q(\mathbf{y}), \mathbf{u})$ for a fixed \mathbf{u} , a sequence of an ever decreasing upper bound of $C_{KL}(q(\mathbf{y})||p(\mathbf{y}|\mathbf{Y}, \mathbf{x}))$ is also obtained due to Eq. (6). However, also minimizing $\mathcal{M}(q(\mathbf{y}))$ with respect to \mathbf{u} , generates a sequence of vectors $\{\mathbf{u}^k\}$ that tightens the upper-bound for each distribution $q^k(\mathbf{y})$. Therefore, the two sequences $\{q^k(\mathbf{y})\}$ and $\{\mathbf{u}^k\}$ are coupled. We develop the following iterative algorithm to find such sequences. 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 [8].

Algorithm 1 *Posterior image distribution estimation.*

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

1. Find

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

2. Find

$$\mathbf{u}^{k+1} = \arg \min_{\mathbf{u}} \int q^k(\mathbf{y}) \times \log\left(\frac{q^k(\mathbf{y})}{F(\mathbf{y}, \mathbf{Y}, \mathbf{x}, \mathbf{u})}\right) d\mathbf{y}. \quad (13)$$

Set $q(\mathbf{y}) = \lim_{k \rightarrow \infty} q^k(\mathbf{y})$.

To calculate \mathbf{u}_b^{k+1} , for $b = 1, \dots, B$, we have from Eq. (13) that

$$\mathbf{u}_b^{k+1} = \arg \min_{\mathbf{u}_b} \sum_i \frac{\mathbf{E}_{q^k(\mathbf{y})} [\Delta_i^h(\mathbf{y}_b))^2 + (\Delta_i^v(\mathbf{y}_b))^2] + \mathbf{u}_b(i)}{\sqrt{\mathbf{u}_b(i)}}, \quad (14)$$

and consequently

$$\mathbf{u}_b^{k+1}(i) = \mathbf{E}_{q^k(\mathbf{y})} [\Delta_i^h(\mathbf{y}_b)^2 + (\Delta_i^v(\mathbf{y}_b))^2], \quad \text{for } i = 1, \dots, p. \quad (15)$$

It is clear from Eq. (15) 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 $q^k(\mathbf{y})$ and represents the local spatial activity of \mathbf{y}_b .

To calculate $q^k(\mathbf{y})$, we observe that differentiating the integral on the right-hand side of Eq. (12) with respect to $q(\mathbf{y})$ and setting it equal to zero, we obtain that

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

with

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

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

$$\phi^k = (\text{diag}(\beta) \otimes \mathbf{H}^t) \mathbf{Y} + \gamma (\text{diag}(\lambda) \otimes \mathbf{I}_p) (\mathbf{x}^t, \mathbf{x}^t, \dots, \mathbf{x}^t)^t,$$

and

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

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

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

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 $W(\mathbf{u}_b^k) = \text{diag}(\mathbf{u}_b^k(i)^{-\frac{1}{2}})$, for $i = 1, \dots, p$. 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. That is, for pixels with high spatial activity the corresponding entries of $W(\mathbf{u}_b^k)$ are very small, which means that no smoothness is enforced, while for pixels in a flat region the corresponding entries of $W(\mathbf{u}_b^k)$ are very large, which means that smoothness is enforced.

4 Experimental Results

Although we performed a wide set of experiments to assess the quality of the proposed approach, here we report results only on a synthetic color image and

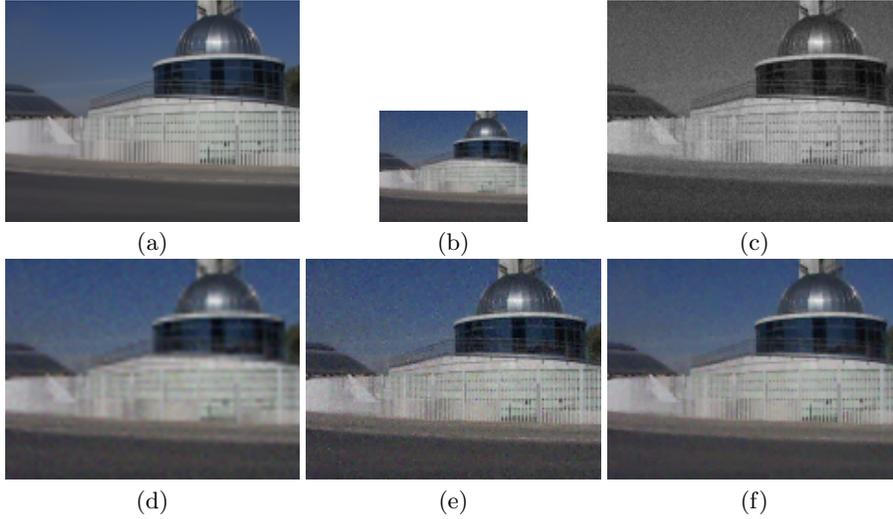


Fig. 1. (a) Original image; (b) Observed multispectral image; (c) Observed panchromatic image; (d) Bicubic interpolation of (a); (e) Reconstruction using the method in [2]; (f) Reconstruction using the proposed model.

a real Landsat multispectral image. In our first experiment, the synthetic multispectral observations are obtained from the color image displayed in Fig. 1(a) by first convolving it with mask $0.25 \times \mathbf{1}_{2 \times 2}$ 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. The panchromatic image was obtained from the original high resolution color image using the model in Eq. (4), with $\lambda_b = 1/3$, for $b = 1, 2, 3$, and additive zero mean Gaussian noise of variance 25. The observed multispectral and panchromatic images are depicted in Figs. 1(b) and 1(c), respectively.

Spatial improvement was assessed by means of the correlation of the high frequency components (COR) [9] which takes values between 0 and 1 (the higher the value the better the quality of the reconstruction), while spectral fidelity was 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, especially a value under 3.0, the higher the quality of the multispectral reconstructed image.

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 $q^k(\mathbf{y})$. It typically required 4 iterations. The value of the parameters were experimentally chosen to be $\alpha_b = 0.001$, $\beta_b = 1/16$, $b = 1, 2, 3$ and $\gamma = 1/25$. Table 1 shows the resulting PSNR, COR, and ERGAS values for the reconstructions of the image using the proposed method, bicubic interpolation and the pansharpening method in [2]. The reconstructed images corresponding to those methods are displayed in Fig. 1(d)–1(f). For all the studied cases, the bicubic interpolation produces over-smoothed reconstructions. The method in [2] produces noisy reconstructions and

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

Method / Band	PSNR			COR			ERGAS
	1	2	3	1	2	3	
Bicubic interpolation	29.09	29.01	29.03	0.49	0.50	0.49	4.62
Method in [2]	32.36	32.55	32.24	0.89	0.90	0.89	3.14
Proposed model	35.31	35.31	35.05	0.92	0.93	0.92	2.26

some color bleeding at the edges, as can be observed at the center of the fence in Fig. 1(e). However the proposed model produces both numerically (see table 1) and visually better results recovering the high resolution structures of the original image while reducing at the same time the amount of noise.

In a second experiment, the method was tested on real Landsat ETM+ image. Figure 2(a) and 2(b) depict a 128×128 pixels false RGB color region of interest composed of bands 3, 4, and 2 of a Landsat ETM+ multispectral image, together with their corresponding 256×256 panchromatic image. 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 [11]. Visual inspection of the results displayed in Figs. 2(c)–2(e) reveals similar conclusions to the ones for the synthetic image. The proposed method provides the best result, preserving the spectral properties of the multispectral image while incorporating the high frequencies from the panchromatic image and controlling the noise in the image. Figure 2(f) depicts the value of $W(\mathbf{u}_4)$ (in log scale), that is, the spatial adaptivity matrix for band 4 in Eq. (18), in which high spatial activity areas are clearly depicted corresponding to lower image values.

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 outperforms bicubic interpolation and the method in [2].

References

1. Nuñez, J., Otazu, X., Fors, O., Prades, A., Pala, V., Arbiol, R.: Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Trans on Geosc. & Rem. Sens.* **37**(3) (1999) 1204–1211
2. Price, J.: Combining multispectral data of different spatial resolution. *IEEE Trans. on Geosc. & Rem. Sens.* **37**(3) (1999) 1199–1203

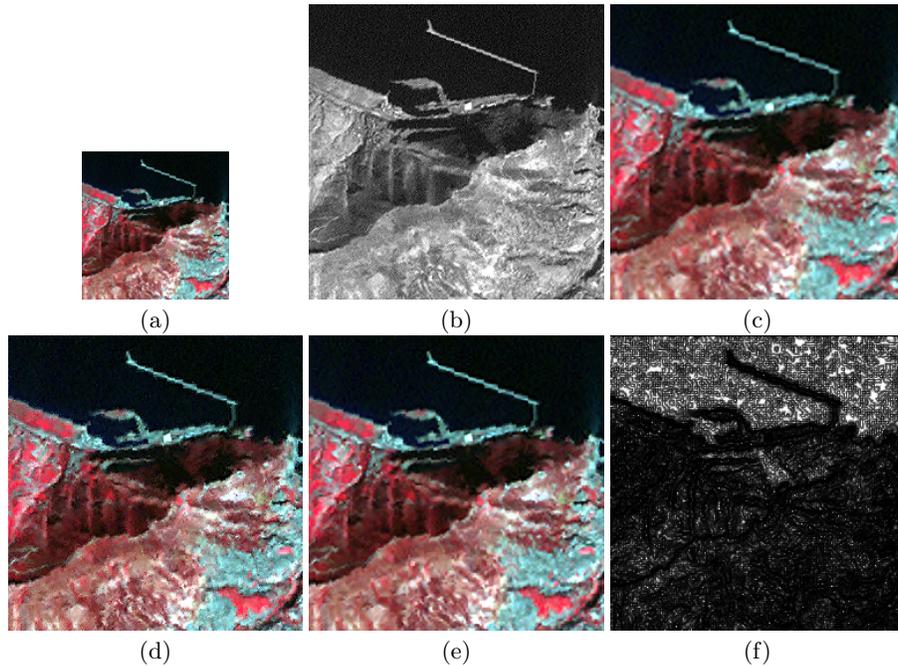


Fig. 2. (a) Observed multispectral image; (b) Observed panchromatic image; (c) Bicubic interpolation of (a); (d) Reconstruction using the method in [2]; (e) Reconstruction using the proposed model; (f) Spatial adaptivity matrix $W(\mathbf{u}_4)$ in log scale.

3. Alparone, L., Wald, L., Chanussot, J., Thomas, C., Gamba, P., Bruce, L.: Comparison of Pansharpening Algorithms: Outcome of the 2006 GRS-S Data-Fusion Contest. *IEEE Trans. on Geosc. & Rem. Sens.* **45**(10) (2007) 3012–3020
4. Babacan, S., Molina, R., Katsaggelos, A.: Parameter estimation in TV image restoration using variational distribution approximation. *IEEE Trans. on Image Proc.* **17**(3) (2008) 326–339
5. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D* **63**(6) (1992) 259–268
6. Kullback, S., Leibler, R.A.: On information and sufficiency. *Annals of Math. Stat.* **22** (1951) 79–86
7. Bioucas-Dias, J., Figueiredo, M., Oliveira, J.: Total-variation image deconvolution: A majorization-minimization approach. In: *ICASSP'2006*. (2006)
8. Lange, K.: Optimization. In: *Springer texts in Statistic*. New York, Springer Verlag (2004)
9. Vijayaraj, V.: A quantitative analysis of pansharpened images. Master's thesis, Mississippi St. University (2004)
10. Wald, L., Ranchin, T., Mangolini, M.: Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Phot. Eng. Rem. Sens.* **63**(6) (1997) 691–699
11. Molina, R., Vega, M., Mateos, J., Katsaggelos, A.: Variational posterior distribution approximation in Bayesian super resolution reconstruction of multispectral images. *Applied and Computational Harmonic Analysis* **24**(2) (2008) 251–267