PARAMETER ESTIMATION IN BAYESIAN SUPER-RESOLUTION PANSHARPENING USING CONTOURLETS

Israa Amro, Javier Mateos* †

Depto. de Ciencias de la Computación e I. A.
Universidad de Granada, Spain

Miguel Vega

Depto. de Lenguajes y Sistemas Informáticos
Universidad de Granada, Spain

ABSTRACT

In this paper, we consider the problem of parameter estimation on the super resolution and Bayesian methodology for pansharpening using contourlet transform. The used methodology is able to incorporate prior knowledge on the expected characteristics of the multispectral images, include information on the unknown parameters in the form of hyperprior distributions and estimate the unknown parameters together with the high resolution multispectral image. 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, parameter estimation

1. INTRODUCTION

Pansharpening is a remote sensing technique for fusing a low resolution (LR) multispectral (MS) image and a high resolution (HR) panchromatic (PAN) image to provide a HR MS image with the level of detail of the PAN image. While many approximations to the problem has been developed (see [1] for a complete review), multiscale, and particulary non-subsampled contourlet transform (NSCT)[2] based techniques like the one proposed in [3], are becoming popular. Also, super-resolution (SR) techniques are being used to pansharpen MS images. Recently, we proposed a new pansharpening method [4], 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 MS image, even those that are not covered by the PAN image. However, the method depends on a series of parameters that have to be adjusted making its use cumbersome.

In this paper, we extend the work in [4] to automatically estimate the unknown parameters, together to the HR MS image, within the hierarchical Bayesian framework, thus overcoming the main disadvantage of the technique in [4]. The proposed method also allows to include information on the expected value of the unknown parameters into the hyperprior distributions. This paper is organized as follows. In section 2 the Bayesian SR using contourlet approach is described. Section 3 describes the variational approach to distribution approximation for Bayesian SR pansharpening using contourlets and how inference is performed. Experimental results and comparison are presented in section 4 for synthetic and SPOT images and finally, section 5 concludes the paper.

*Also Al-Quds Open University, Palestine

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

2. BAYESIAN PROBLEM FORMULATION

The unknown HR MS image we would have observed under ideal conditions, y, has B bands, y_b, $b = 1, ..., B$, each one centered on a narrow spectral band, of size $p = m \times n$. Using matrix vector notation, y is expressed as the column vector $y = [y_1^t, y_2^t, ..., y_B^t]^t$, 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 MS image Y has B bands Y_b, $b = 1, ..., B$, each of size $P = M \times N$ pixels, with $M < m$ and $N < n$ that are also stacked into the vector Y, by lexicographically ordering the pixels in each 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. The objective is to reconstruct the HR MS image y from the LR MS image Y and the PAN image x.

In the hierarchical Bayesian approach to our HR MS image reconstruction problem we have two stages; in the first stage, knowledge about the structural form of LR MS and PAN image observation noise and the structural behavior of the HR MS image is used in forming $p(Y, x|y, \Omega)$ and $p(y|\Omega)$, respectively. These noise and image models depend on the unknown hyper-parameters Ω. In the second stage, a hyper-prior on the hyper-parameters is defined, allowing the incorporation of information about the hyper-parameters into the process. So, following the Bayesian paradigm we can define the joint distribution on the observation, hyper-parameters and HRMS image, $p(\Omega, y, Y, x) = p(\Omega)p(y|\Omega)p(Y, x|y, \Omega)$, and the inference will based on $p(\Omega, y|Y, x)$.

Following [4, 5], for the HR MS image y we chose the Total Variation (TV) prior given by

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

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.

The probability distribution of the observed images Y and x given the HR MS image y can be written, taking into account that Y and x are independent for a given y, as $p(Y, x|y, \Omega) = p(Y|y, \Omega)p(x|y, \Omega)$. The conditional distribution of the LR MS image Y is defined as [4, 5]

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

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

Finally, let us consider that the PAN image x and each band of the HR MS image y_b are decomposed using the non-subsampled contourlet transform as $x = \sum_j C_j x + C_r x$, and $y_b = \sum_j C_j y_b + C_r y_b$ where $C_j x$ and $C_j y_b$ are the NSCT coefficients for the PAN image and HRMS band b, respectively, at a level of decomposition j and $C_r x$ and $C_r y_b$ are the residual image of PAN and MS bands, respectively. Following [3] we draw our attention to the relation between the details of the PAN image and of the HR MS bands. We consider here the details of the PAN image, as a combination of the high frequency details of the HR MS bands. Hence, the probability distribution of the details of the PAN image, x_d, given y, is written as

$$p(x|y,\Omega) \propto \gamma^{p/2} \exp\left\{-\frac{1}{2}\gamma \left\| x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j y_b \right\|^2\right\}, \quad (3)$$

where $x_d = \sum_j C_j x$ contains the details of the PAN image, obtained using the contourlet transform, and the noise is assumed to be Gaussian with zero mean and known variance γ^{-1}.

Having defined the prior and noise distributions, we now define the second stage of the hierachichal Bayesian framework, that is, the distribution on the parameters, henceforth named hyperparameters. To model the hyperparameters we are going to employ a gamma distribution,

$$p(w|a_w, c_w) = \Gamma(w|a_w, c_w), \quad (4)$$

where $w > 0, w \in \Omega = (\alpha_1, \dots, \alpha_B, \beta_1, \dots, \beta_B, \gamma)$ denotes a hyperparameter, and $a_w > 0$ and $c_w > 0$ are, respectively, the shape and the inverse scale parameters of the distribution.

Finally, combining the first and second stage of the problem modeling we have the global distribution,

$$\begin{aligned} p(\Omega, y, Y, x) &= p(\Omega)p(y|\Omega)p(Y|y,\Omega)p(x|y,\Omega) \\ &= p(\gamma)p(x|y,\gamma)\prod_{b=1}^{B} p(\alpha_b)p(\beta_b)p(y_b|\alpha_b)p(Y_b|y_b,\beta_b), \quad (5) \end{aligned}$$

where $p(y_b|\alpha_b)$, $p(Y_b|y_b,\beta_b)$ and $p(x|y,\gamma)$ are given in Eqs. (1), (2), and (3), respectively.

3. BAYESIAN INFERENCE AND VARIATIONAL APPROXIMATION OF THE POSTERIOR DISTRIBUTION FOR SR RECONSTRUCTION OF MS IMAGE

For our selection of hyper-parameters in the previous section, the set of all unknowns is $(\Omega, y) = (\alpha_1, \dots, \alpha_B, \beta_1, \dots, \beta_B, \gamma, y)$. The Bayesian paradigm dictates that inference on (Ω, y) should be based on, $p(\Omega, y|Y, x) = p(\Omega, y, Y, x)/p(Y, x)$. However, $p(\Omega, y|Y, x)$ can not be found in closed form and thus we will apply variational methods to approximate this distribution by the distribution $q(\Omega, y)$.

The variational criterion used to find $q(\Omega, y)$ is the minimization of the Kullback-Leibler divergence

$$C_{KL}(q(\Omega,y)||p(\Omega,y|Y,x)) = \int\int q(\Omega,y)\log\left(\frac{q(\Omega,y)}{p(\Omega,y|Y,x)}\right)d\Omega dy.$$

We choose to approximate the posterior distribution $p(\Omega, y|Y, x)$ by the distribution $q(\Omega, y) = q(\Omega)q(y)$, where $q(y)$ and $q(\Omega)$ denote distributions on y and Ω, respectively.

Due to the form of the TV prior, the above integral cannot be directly evaluated so, following [5], we approximate it by using the Majorization-Minimization approach [6]. Let us define the functional

$$M(\alpha_b, y_b, u_b) = \alpha_b^{\frac{p}{2}} \exp\left[-\frac{\alpha_b}{2}\sum_{i=1}^{p}\frac{(\Delta_i^h(y_b))^2 + \Delta_i^v(y_b))^2 + u_b(i)}{\sqrt{u_b(i)}}\right]. \quad (6)$$

where $u_b \in (R^+)^p$ is a p-dimensional vector with components $u_b(i), i = 1, \dots, p$, that need to be calculated and have, as we will show later, an intuitive interpretation related to the unkown images y_b. Comparing Eq. (6) with Eq. (1), we obtain $p(y_b|\alpha_b) \geq c.M(\alpha_b, y_b, u_b)$, which leads to the following lower bound for the joint probability distribution,

$$\begin{aligned} p(\Omega, y, Y, x) &\geq c.p(\Omega)\prod_{b=1}^{B}\left(M(\alpha_b, y_b, u_b)p(Y|y,\beta)\right)p(x|y,\gamma) \\ &= F(\Omega, y, Y, x_d, u), \quad (7) \end{aligned}$$

where $u = [u_1^t, u_2^t, \dots, u_B^t]^t$. Hence, we can find an upper bound for the KL divergence as

$$C_{KL}(q(\Omega,y)||p(\Omega,y|Y,x)) \leq C_{KL}(q(\Omega)q(y)||F(\Omega,y,Y,x_d,u)). \quad (8)$$

The following algorithm that extends the one presented in [4] to deal with parameter estimation, can be used for calculating the approximating posteriors $q(\Omega)$ and $q(y)$.

Algorithm 1 *Estimation of the posterior distributions of the HRMS image and of the parameters*

Given $u^1 \in (R^+)^p$ and $q^1(\Omega)$, an initial estimate of the distribution $q(\Omega)$.
For $k = 1, 2, \dots$ until a stopping criterion is met.

1. *Find* $q^k(y) = \arg\min_{q(y)} C_{KL}(q^k(\Omega)q(y)||F(\Omega,y,Y,x_d,u^k))$ (9)
2. *Find* $u^{k+1} = \arg\min_{u} C_{KL}(q^k(\Omega)q^k(y)||F(\Omega,y,Y,x_d,u))$ (10)
3. *Find* $q^{k+1}(\Omega) = \arg\min_{q(\Omega)} C_{KL}(q(\Omega)q^k(y)||F(\Omega,y,Y,x_d,u^{k+1}))$ (11)

Set $q(\Omega) = \lim_{k\to\infty} q^k(\Omega), q(y) = \lim_{k\to\infty} q^k(y)$.

In Eq. (9), we obtain for $q^k(y)$ the p-dimensional Gaussian distribution $q^k(y) = \mathcal{N}(y|E^k(y)[y], cov^k[y])$, with

$$cov^k[y] = \mathcal{A}^{-1}(u^k), \quad E^k[y] = cov^k[y]\phi^k, \quad (12)$$

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

$$\phi^k = (diag(\beta^k) \otimes H^t)Y + \frac{\gamma^k}{B}(1_B \otimes \sum_j C_j^t x_d), \quad (13)$$

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

$$\mathcal{A}(u^k) = \varsigma(u^k) + diag(\beta^k) \otimes H^t H + \frac{\gamma^k}{B^2}(1_{B\times B} \otimes \sum_{j,k} C_j^t C_k),$$

with $\varsigma(u^k)$ a $B \times p \times B \times p$ block diagonal matrix whose diagonal blocks are the $p \times p$ matrices $\alpha_b^k[(\Delta^h)^t W(u_b^k)(\Delta^h) +$

$(\Delta^v)^t W(u_b^k)(\Delta^v)]$, for $b = 1, ..., B$, where Δ^h and Δ^v representing $p \times p$ convolution matrices associated with the first order horizontal and vertical differences, respectively, $1_{B\times B}$ is a $B \times B$ matrix with all its elements equal to one, $\otimes$ is the Kronecker product, $\beta = (\beta_1, \beta_2, ..., \beta_B)^t$ and $W(u_b^k) = diag(u_b^k(i)^{-1/2})$, for $i = 1, ..., p$, is a $p \times p$ diagonal matrix, that represents 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.

To calculate u^{k+1} from Eq. (10), we obtain

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

Once we know $q^k(y)$ and u^{k+1}, the next step is to calculate the distributions of the hyperparameters from Eq. (11). First we note that

$$q^{k+1}(\Omega) = q^{k+1}(\gamma)\prod_{b=1}^{B} q^{k+1}(\alpha_b)q^{k+1}(\beta_b)$$

where $q(\omega) = \Gamma(\omega|\bar{a}_\omega, \bar{c}_\omega)$ which produces

$$\frac{1}{E^{k+1}[\alpha_b]} = \frac{\lambda_{\alpha_b}}{\overline{\alpha_b}} + (1-\lambda_{\alpha_b})\frac{2}{p}\sum_{i=1}^{p}\left[\left(\Delta_i^h(E^k[y_b])\right)^2 + \left(\Delta_i^v(E^k[y_b])\right)^2 + \frac{1}{p}tr\left[(cov^k[y_b])^{-1}((\Delta^h)^t(\Delta^h) + (\Delta^v)^t(\Delta^v))\right]\right]^{\frac{1}{2}} \quad (14)$$

$$\frac{1}{E^{k+1}[\beta_b]} = \frac{\lambda_{\beta_b}}{\overline{\beta_b}} + (1-\lambda_{\beta_b}) \times \frac{\left\|Y_b - HE^k[y_b]\right\|^2 + tr\left(H^t H(cov^k[y_b])^{-1}\right)}{P} \quad (15)$$

$$\frac{1}{E^{k+1}[\gamma]} = \frac{\lambda_\gamma}{\overline{\gamma}} + (1-\lambda_\gamma)\left[\frac{\left\|x_d - \frac{1}{B}\sum_{b=1}^{B}\sum_j C_j E^k[y_b]\right\|^2}{p} + \frac{tr\left(\frac{1}{B^2}(1_{B\times B}\otimes\sum_{j,k}C_j^t C_k)(cov^k[y_b])^{-1}\right)}{p}\right]. \quad (16)$$

where we took into account that the mean of the prior distribution on the parameter ω, $\overline{\omega} = \bar{a}_\omega/\bar{c}_\omega$ and that $\lambda_{\alpha_b} = \bar{a}_{\alpha_b}/(\bar{a}_{\alpha_b} + p/2)$, $\lambda_{\beta_b} = \bar{a}_{\beta_b}/(\bar{a}_{\beta_b} + P/2)$ and $\lambda_\gamma = \bar{a}_\gamma/(\bar{a}_\gamma + p/2)$.

Those equations indicate that the inverse of the means of the parameters can be expressed as convex combinations of their prior values and their maximum likelihood (ML) estimates and that λ_ω, taking values in the interval $[0, 1)$, can be understood as the confidence on the inverse of the mean of the prior distribution on the parameters. So when λ_ω are equal to zero, no confidence is placed on the prior values of the hyper-parameters and ML estimates are used, making the observations fully responsible of the parameters estimation, while when they are asymptotically equal to one, the prior knowledge of the mean is fully enforced (i.e., no estimation of the hyper-parameters is performed).

From Eqs. (14)-(16), we can see that $cov^k[y]$ is explicitly needed to implement the proposed algorithm. However, since the calculation of $cov^k[y]$ is very intense, following [7], we propose to approximate the covariance matrix $cov^k[y]$ as

$$(cov^k[y_b])^{-1} \approx E^k[\beta_b]H^t H + E^k[\gamma]\frac{1}{B^2}\sum_{j,k}C_j^t C_k + E^k[\alpha_b]z(u_b^k)\left((\Delta^h)^t(\Delta^h) + (\Delta^v)^t(\Delta^v)\right), \quad (17)$$

where $z(u_b^k) = 1/p\sum_{i=1}^{p} 1/\sqrt{u_b^k(i)}$.

Measure	NSCT in [3]	SR in [5]	Proposed 1 (no prior)	Proposed 2 (with prior)
COR	0.91	0.80	**0.99**	**0.99**
SSIM	0.80	0.90	0.96	**0.97**
PSNR	27.19	32.78	36.96	**37.01**
ERGAS	5.76	3.12	2.73	**2.25**

Table 1. Synthetic Image Quantative Results

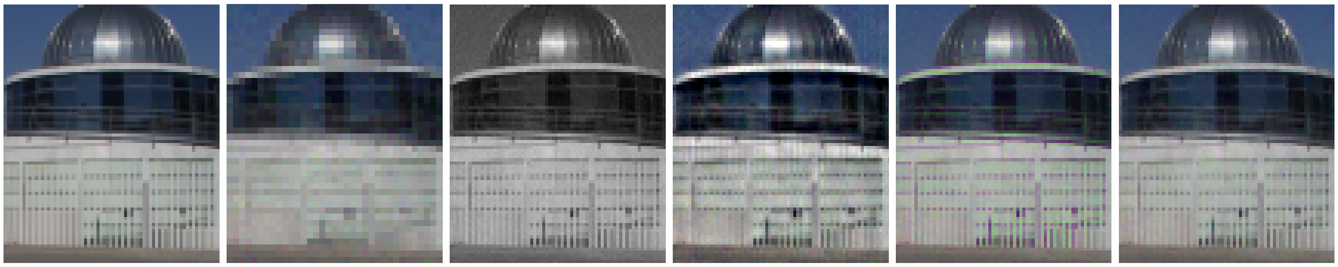
4. EXPERIMENTAL RESULTS

We tested the proposed method on 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 [3]. To assess the spatial improvement of the pansharpened images we use the correlation of the high frequency components (COR) 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 mean peak signal-to-noise ratio (PSNR), the mean Structural Similarity Index Measure (SSIM), an index ranging from -1 to $+1$, with $+1$ corresponding to exactly equal images, and the *erreur relative globale adimensionnelle de synthése* (ERGAS) 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. See, for instance [1] for a detailed description of those measures.

In order to conduct experiments where the ground truth is known, we used synthetic MS observations, obtained from the color image displayed in Figure 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. A ROI of the observed PAN image and MS image, scaled to the size of the PAN image for displaying purposes, are shown in Figure 1(b) and (c), respectively.

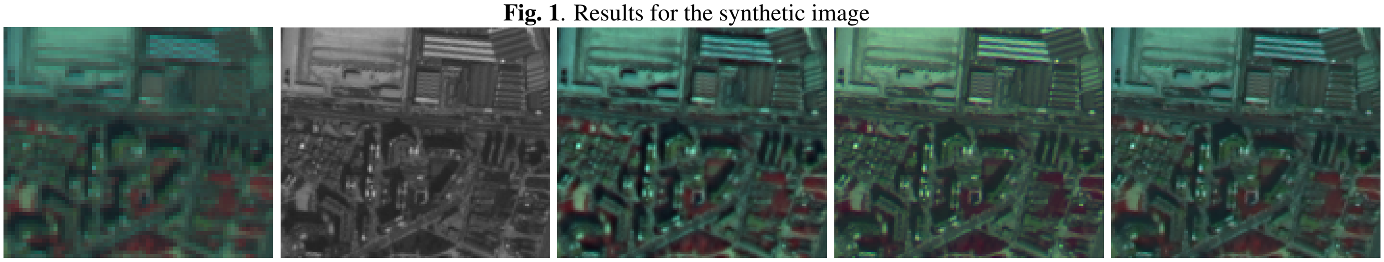
The criterion $\left\|E^k[y] - E^{k-1}[y]\right\|^2 / \left\|E^{k-1}[y]\right\|^2 < 10^{-4}$ was used to stop the algorithm, which typically is reached within 5 iterations. The values of all parameters were automatically estimated using the proposed algorithm. We examined the effect of the introduction of additional information about the unknown hyperparameters through the use of the confidence parameters on the performance of the algorithm. We provided the observed MS and PAN images to the algorithm and run the algorithm with different values of λ_ω. We found that very good results are obtained even in the absence of prior knowledgement on the value of the parameters (see Proposed 1 in Table 1) although introducing prior knowledgement on them improves the performance, increasing COR and SSIM and decreasing ERGAS, and making the method to converge in less iterations. The best values are obtained when $\lambda_{\alpha_b} = 0.9$, $\lambda_{\beta_b} = 0.9$ and $\lambda_\gamma = 0.0$ (see Proposed 2 in Table 1).

The resulting images corresponding to the reconstruction of the synthetic image using the NSCT method in [3], the SR method in [5], and the best result using proposed method are displayed in Figure 1(d)-(f), respectively, and Table 1 shows the corresponding quantitative results. The highlighted value in the table presents the best value for each measure. The proposed method provides better results for each measure. 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 Figure 1(e), introduced



(a) Original (b) Observed MS (c) Observed PAN (d) NSCT in [3] (e) SR in [5] (f) Proposed.

Fig. 1. Results for the synthetic image



(a) Observed MS image (b) Observed PAN image (c)NSCT method in [3] (d) SR method in [5] (e) proposed method.

Fig. 2. Results for the SPOT5 image

less details. The NSCT method in [3] incorporates details in all the bands but produces a noisy image, see Figure 1(d). The proposed method (Figure 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 Figure 1(f), producing an image that is not as noisy as the NSCT method in [3] (Figure 1(d)) and preserves better the colors than the SR method in [5] (Figure 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 [3] and from 2 to almost 6 dB higher than for the SR method in [5].

In a second experiment, the method was tested on a real SPOT5 dataset. 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 Figure 2(b). Visual inspection of the resulting images, displayed in Figures 2(c)-(e), reveals similar conclusions to the obtained for the synthetic image. The proposed method calculates the parameters automatically 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. The NSCT method in [3] (Fig.2(c)) provides a detailed image but quite noisy, the SR method in [5] provides good details for bands 1 and 2, see Figure 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 Figure 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.

5. CONCLUSIONS

In this paper, a new pansharpening method based on SR reconstruction and NSCT transform has been presented. The proposed method estimates automatically the HRMS image and all the unknown hyperparameters. The proposed method preserves the spectral properties of MS image while incorporating the high frequencies from the PAN 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.

6. REFERENCES

[1] I. Amro, *Multispectral Image fusion using Multiscale and Super-resolution methods*, Ph.D. thesis, Computer Science and Artificial Intelligence Dep., Universidad de Granada, 2011.

[2] A. L. da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, 2006.

[3] M. Lillo-Saavedra and C. Gonzalo, "Multispectral images fusion by a joint multidirectional and multiresolution representation," *Int. J. Remote Sens.*, vol. 28, pp. 4065 – 4079, 2007.

[4] I. Amro, J. Mateos, and M. Vega, "Bayesian super-resolution pansharpening using contourlets," in *International Conference on Image Processing 2010 - Finalist of the Huawei Best Student Paper Award*. Honk Kong (China), 2010, pp. 809–812.

[5] M. Vega, J. Mateos, R. Molina, and A.K.. Katsaggelos, "Super resolution of multispectral images using TV image models," in *2th Int. Conf. on Knowledge-Based and Intelligent Information & Eng. Sys.*, 2008, pp. 408–415.

[6] N. Galatsanos, "A majorization-minimization approach to total variation reconstruction of super-resolved images," in *16th European Signal Processing Conf.*, 2008.

[7] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Parameter estimation in tv image restoration using variational distribution approximation," *IEEE Transactions on Image Processing*, vol. 17, no. 3, pp. 326–339, 2008.