HIERARCHICAL BAYESIAN SUPER RESOLUTION RECONSTRUCTION OF MULTISPECTRAL IMAGES

Rafael Molina[a], Miguel Vega[b], Javier Mateos[a], and Aggelos K. Katsaggelos[c]

a) Departmento de Ciencias de la Computación e Inteligencia Artificial,
Universidad de Granada, 18071 Granada, Spain.
b) Departamento de Lenguajes y Sistemas Informáticos,
Universidad de Granada, 18071 Granada, Spain.
c) Department of Electrical Engineering and Computer Science,
Northwestern University, Evanston, Illinois 60208-3118.
mvega@ugr.es, jmd@decsai.ugr.es, rms@decsai.ugr.es, aggk@ece.northwestern.edu

ABSTRACT

In this paper we present a super resolution Bayesian methodology for pansharpening of multispectral images which: a) incorporates prior knowledge on the expected characteristics of the multispectral images, b) uses the sensor characteristics to model the observation process of both panchromatic and multispectral images, c) includes information on the unknown parameters in the model, and d) allows for the estimation of both the parameters and the high resolution multispectral image. Using real data, the pansharpened multispectral images are compared with the images obtained by other parsharpening methods and their quality assessed both qualitatively and quantitatively.

1. INTRODUCTION

Nowadays most remote sensing systems include sensors able to capture, simultaneously, several low resolution images of the same area on different wavelengths, thus forming a multispectral image, along with a high resolution panchromatic image. For instance, the Landsat 7 satellite, equipped with the ETM+ sensor, allows for the capture of a multispectral image with six bands (three bands in the visible spectrum plus three bands in the infrared) with a resolution of 30 meters per pixel, a thermal band with a resolution of 60 meters per pixel and a panchromatic band (covering a large zone on the visible spectrum and the near infrared), with a resolution of 15 meters per pixel.

The term multispectral image reconstruction using super resolution as used in this paper, refers to the joint processing of the multispectral and panchromatic images in order to obtain a new multispectral image that, ideally, will present the spectral characteristics of the observed multispectral image and the resolution and quality of the panchromatic image.

A few approximations to this problem have been proposed in the literature (see [11] for a comparative study). Among them Principal Component Analysis (PCA) (see [5]) and wavelets based approaches [8] have been used. Price [9] proposed a method relying on the statistical relationships between the radiances in the low and high spatial resolution bands. Recently a few super-resolution based methods have also been proposed [4, 1].
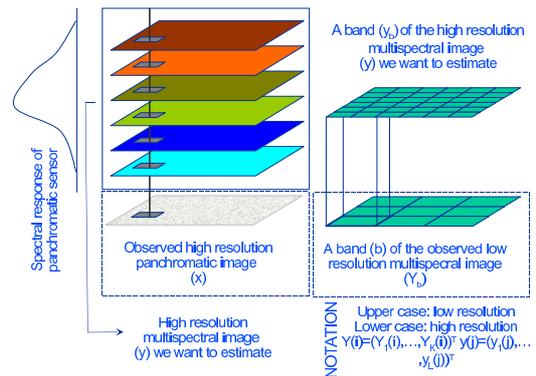
Figure 1: Problem formulation, acquisition model and used notation.

In this paper we formulate the reconstruction of a multispectral image using super resolution techniques from a hierarchical Bayesian perspective and derive a new method to simultaneously estimate the parameters of the model and the high resolution multispectral image from the observed images. The method extends the results already presented in [7], where an iterative algorithm for the reconstruction problem, assuming the values of the parameters were known, was proposed.

The paper is organized as follows. The acquisition model is presented in section 2. In section 3 the hierachical Bayesian paradigm for super resolution applied to multispectral image reconstruction is presented and the required probability distributions are formulated. The Bayesian analysis is performed in section 4 to obtain the reconstruction and parameter estimation algorithm. Experimental results on a real Landsat 7 ETM+ image are described in section 5. Finally section 6 concludes the paper.

2. ACQUISITION MODEL

Let us assume that $\mathbf{y}$, the multispectral image we would observe under ideal conditions with a high resolution sensor, has B bands

$$\mathbf{y} = [\mathbf{y}_1^t, \mathbf{y}_2^t, \ldots, \mathbf{y}_B^t]^t, \qquad (1)$$

each one of size $p = m \times n$ pixels, with t denoting the transpose of a vector or matrix. Each band of this image can be expressed as a column vector by lexicographically ordering the pixels in the band.
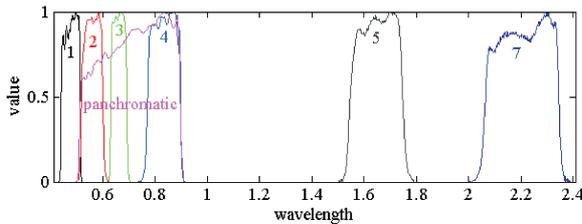
Figure 2: Landsat 7 ETM+ band spectral response normalized to one.

In real applications, this image is not available and, instead, we observe a low resolution multispectral image $\mathbf{Y}$ with B bands

$$\mathbf{Y} = [\mathbf{Y}_1^t, \mathbf{Y}_2^t, \ldots, \mathbf{Y}_B^t]^t,$$

each one of size $P = M \times N$ pixels, $M < m$ and $N < n$. Each band of this image can be expressed as a column vector by lexicographically ordering the pixels in the band. Figure 1 illustrates the acquisition model and the notation used.

Each band, $\mathbf{Y}_b$, is related to its corresponding high resolution image by

$$\mathbf{Y}_b = \mathbf{H}\mathbf{y}_b + \mathbf{n}_b, \quad \forall b = 1, \cdots, B, \qquad (2)$$

where $\mathbf{H}$ is a $P \times p$ matrix representing the blurring, the sensor integration function and the spatial subsampling (we assume that this process is the same over the whole set of spectral images), and $\mathbf{n}_b$ is the capture noise, assumed to be Gaussian with zero mean and variance $1/\beta_b$.

A simple but widely used model for the matrix $\mathbf{H}$ is to consider that each pixel (i,j) of the low resolution image is obtained according to (for $m = 2M$ and $n = 2N$)

$$Y_b(i,j) = \frac{1}{4} \sum_{(u,v) \in E_{i,j}} y_b(u,v) + n_b(i,j), \qquad (3)$$

where $E_{i,j}$ consists of the indices of the four high resolution pixels $E_{i,j} = \{(2i,2j),(2i+1,2j),(2i,2j+1),(2i+1,2j+1)\}$ (see right hand side of figure 1).

We note here that $\mathbf{H}$ can be written as

$$\mathbf{H} = \mathbf{D}\mathbf{B}, \qquad (4)$$

where $\mathbf{B}$ is a $p \times p$ blurring matrix and $\mathbf{D}$ is a decimation operator.

The sensor also provides us with a panchromatic image $\mathbf{x}$ of size $p = m \times n$, obtained by spectral averaging the unknown high resolution images $\mathbf{y}_b$. This relation can be modelled as

$$\mathbf{x} = \sum_{b=1}^{B} \lambda_b \mathbf{y}_b + \nu, \qquad (5)$$

where $\lambda_b \geq 0$, $b = 1, 2, \cdots, B$, are known quantities that can be obtained, as we will see later, from the sensor spectral characteristics, and weight the contribution of each band $\mathbf{y}_b$ to the panchromatic image $\mathbf{x}$ (see left hand side of figure 1), and ν is the capture noise that is assumed to be Gaussian with zero mean and variance γ^{-1}.

3. BAYESIAN MODELING

Our first goal is to define a joint distribution $p(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x})$ of the high resolution panchromatic observed image $\mathbf{x}$, the low

resolution multispectral observations $\mathbf{Y}$, the unknown high resolution multispectral image $\mathbf{y}$, and the hyperparameters describing their distributions Ω.

To model the joint distribution, we utilize the hierarchical Bayesian paradigm (see, for example, [6]). In the hierarchical approach to super resolution of multispectral images we have at least two stages. In the first stage, knowledge about the structural form of the observation process and the structural behavior of the high resolution multispectral image is used in forming $p(\mathbf{Y}, \mathbf{x}|\mathbf{y}, \Omega)$, and $p(\mathbf{y}|\Omega)$, respectively. These models depend on the unknown hyperparameters Ω. In the second stage, a hyperprior on the hyperparameters, $p(\Omega)$, is defined, thus allowing the incorporation of information about these hyperparameters into the process.

For $\Omega, \mathbf{y}, \mathbf{Y}$, and $\mathbf{x}$ the following joint distribution is then defined

$$p(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}) = p(\Omega)p(\mathbf{y}|\Omega)p(\mathbf{Y}, \mathbf{x}|\mathbf{y}, \Omega). \qquad (6)$$

We note here that each of the two above mentioned conditional distributions will depend only on a subset of Ω, but we use this more general notation until we precisely describe the hyperparameters that define Ω.

3.1 Degradation model

We assume that $\mathbf{Y}$ and $\mathbf{x}$, for a given $\mathbf{y}$, and a set of parameters Ω are independent and consequently

$$p(\mathbf{Y}, \mathbf{x}|\mathbf{y}, \Omega) = p(\mathbf{Y}|\mathbf{y}, \Omega)p(\mathbf{x}|\mathbf{y}, \Omega). \qquad (7)$$

Given the degradation model for multispectral image super resolution described by Eq. (2) the distribution of the observed $\mathbf{Y}$ given $\mathbf{y}$ and a set of parameters Ω is equal to

$$\begin{aligned} p(\mathbf{Y}|\mathbf{y}, \Omega) &= \prod_{b=1}^{B} p(\mathbf{Y}_b|\mathbf{y}_b, \beta_b) \\ &= \frac{1}{Z_\beta} \exp\left\{-\frac{1}{2}\sum_{b=1}^{B} \beta_b \parallel \mathbf{Y}_b - \mathbf{H}\mathbf{y}_b \parallel^2\right\}, \end{aligned} \qquad (8)$$

where $Z_\beta = \prod_{b=1}^{B}(\frac{2\pi}{\beta_b})^{P/2}$.

Using the degradation model in Eq. (5), the distribution of the panchromatic image $\mathbf{x}$ given $\mathbf{y}$ and a set of parameters Ω is given by

$$p(\mathbf{x}|\mathbf{y}, \gamma) = \frac{1}{Z_\gamma} \exp\left\{-\frac{1}{2}\gamma \parallel \mathbf{x} - \sum_{j=1}^{B} \lambda_j \mathbf{y}_j \parallel^2\right\}, \qquad (9)$$

where $Z_\gamma = (\frac{2\pi}{\gamma})^{p/2}$.

3.2 Image model

It is also necessary to specify the prior distribution of the high resolution image and a set of parameters Ω, $p(\mathbf{y}|\Omega)$. Although other models are possible, in this paper we assume no correlation between the different high resolution bands. Then, our prior knowledge about the smoothness of the object luminosity distribution within each band makes it possible to model the distribution of $\mathbf{y}$ given a set of parameters Ω by

$$p(\mathbf{y}|\Omega) = \prod_{b=1}^{B} p(\mathbf{y}_b|\alpha_b) = \frac{1}{Z_\alpha} \prod_{b=1}^{B} \exp\left\{-\frac{1}{2}\alpha_b \parallel \mathbf{C}\mathbf{y}_b \parallel^2\right\}, \qquad (10)$$

where $\mathbf{C}$ denotes the Laplacian operator, $1/\alpha_b$ is the variance of the Gaussian distribution and $Z_\alpha \propto \prod_{b=1}^{B}(\frac{2\pi}{\alpha_b})^{p/2}$.

3.3 Hyperprior on the hyperparameters

As mentioned earlier Ω denotes the set of hyperparameters describing the distributions introduced above, that is,

$$\Omega = (\gamma, \beta_1, \ldots, \beta_B, \alpha_1, \ldots, \alpha_B). \tag{11}$$

A large part of the Bayesian literature is devoted to finding hyperprior distributions $p(\Omega)$ for which $p(\Omega, \mathbf{y}|\mathbf{x}, \mathbf{Y})$ can be calculated in a straightforward way or be approximated. These are the so called conjugate priors [2]. Conjugate priors have, as we will see later, the intuitive feature of allowing one to begin with a certain functional form for the prior and end up with a posterior of the same functional form, but with the parameters updated by the sample information.

Taking the above considerations about conjugate priors into account, we will assume that the hyperparameters, $\omega \in \Omega$, are independent and follow a gamma distribution defined by

$$p(\omega|a_\omega^o, c_\omega^o) = \frac{((a_\omega^o - 1)c_\omega^o)^{a_\omega^o}}{\Gamma(a_\omega^o)} \omega^{a_\omega^o - 1} \exp[-(a_\omega^o - 1)c_\omega^o \, \omega], \tag{12}$$

where $\omega > 0$ denotes a hyperparameter, and the two parameters $c_\omega^o > 0$ and $a_\omega^o > 1$ are assumed known. This gamma distribution has the following mean, mode, and variance

$$\mathrm{E}[\omega] = \frac{a_\omega^o}{(a_\omega^o - 1)c_\omega^o} \quad , \quad \mathrm{Mode}[\omega] = \frac{1}{c_\omega^o} \tag{13}$$

$$\mathrm{Var}[\omega] = \frac{a_\omega^o}{((a_\omega^o - 1)c_\omega^o)^2}. \tag{14}$$

Note that the mean and mode do not coincide.

4. BAYESIAN INFERENCE

Having defined the degradation, image, and hyperprior models, Bayesian inference is applied.

The inference approach we will follow consists of estimating the hyperparameters in Ω by using

$$\hat{\Omega} = \arg\max_{\Omega} p(\Omega|\mathbf{Y}, \mathbf{x}) = \arg\max_{\Omega} \int_{\mathbf{y}} p(\Omega, \mathbf{y}, \mathbf{Y}, \mathbf{x}) d\mathbf{y}, \tag{15}$$

and then estimating the high resolution multispectral image by solving

$$\mathbf{y}(\hat{\Omega}) = \arg\max_{\mathbf{y}} p(\mathbf{y}|\hat{\Omega}, \mathbf{Y}, \mathbf{x}). \tag{16}$$

Let us redefine $\mathbf{y}(\Omega)$ by maximizing the log of a function instead of the function itself, that is,

$$\mathbf{y}(\Omega) = \arg\max_{\mathbf{y}} \log p(\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega). \tag{17}$$

To estimate the hyperparameters and image, we note that

$$\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega \sim \mathcal{N}\left(\mathbf{E}[\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega], \mathbf{cov}[\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega]\right). \tag{18}$$

The covariance of the posterior satisfies

$$\mathbf{cov}[\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega] = \mathbf{Q}(\Omega)^{-1}, \tag{19}$$

with

$$\mathbf{Q}(\Omega) = \begin{pmatrix} \alpha_1 \mathbf{C}^t\mathbf{C} & \mathbf{0}_p & \ldots & \mathbf{0}_p \\ \mathbf{0}_p & \alpha_2 \mathbf{C}^t\mathbf{C} & \ldots & \mathbf{0}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_p & \mathbf{0}_p & \ldots & \alpha_B \mathbf{C}^t\mathbf{C} \end{pmatrix}$$

$$+ \begin{pmatrix} \beta_1 \mathbf{H}^t\mathbf{H} & \mathbf{0}_p & \ldots & \mathbf{0}_p \\ \mathbf{0}_p & \beta_2 \mathbf{H}^t\mathbf{H} & \ldots & \mathbf{0}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_p & \mathbf{0}_p & \ldots & \beta_B \mathbf{H}^t\mathbf{H} \end{pmatrix}$$

$$+ \quad \gamma \mathbf{\Lambda} \otimes \mathbf{I}_p, \tag{20}$$

the operator $\otimes$ is the Kronecker product, and

$$\mathbf{\Lambda} = \begin{pmatrix} (\lambda_1)^2 & \lambda_1\lambda_2 & \ldots & \lambda_1\lambda_B \\ \lambda_2\lambda_1 & (\lambda_2)^2 & \ldots & \lambda_2\lambda_B \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_B\lambda_1 & \lambda_B\lambda_2 & \ldots & (\lambda_B)^2 \end{pmatrix}. \tag{21}$$

The mean of the posterior distribution $\mathbf{E}[\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega]$, which coincides with the mode of the posterior distribution, satisfies

$$\mathbf{E}[\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega] = \mathbf{Q}(\Omega)^{-1}\phi(\Omega) \tag{22}$$

where $\mathbf{Q}(\Omega)$ has been defined in Eq. (20) and $\phi(\Omega)$ is the $(B \times p) \times 1$ vector

$$\phi(\Omega) = \begin{pmatrix} \beta_1 \mathbf{H}^t & \mathbf{0}_p & \ldots & \mathbf{0}_p \\ \mathbf{0}_p & \beta_2 \mathbf{H}^t & \ldots & \mathbf{0}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_p & \mathbf{0}_p & \ldots & \beta_B \mathbf{H}^t \end{pmatrix} \mathbf{Y}$$

$$+ \quad \gamma \begin{pmatrix} \lambda_1 \mathbf{I}_p & \mathbf{0}_p & \ldots & \mathbf{0}_p \\ \mathbf{0}_p & \lambda_2 \mathbf{I}_p & \ldots & \mathbf{0}_p \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_p & \mathbf{0}_p & \ldots & \lambda_B \mathbf{I}_p \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{x} \\ \vdots \\ \mathbf{x} \end{pmatrix}. \tag{23}$$

For simplicity, the dependency of $\phi(\Omega)$ on both $\mathbf{x}$ and $\mathbf{Y}$ is not denoted explicitly.

To estimate Ω we now apply the E-M algorithm [3]. Let us assume that we already have an estimate Ω' of the unknown parameters, then

$$\mathbf{U}(\Omega|\Omega') = \mathbf{E}_{\mathbf{y}|\mathbf{Y}, \mathbf{x}, \Omega'}[\log p(\mathbf{y}, \mathbf{x}, \mathbf{Y}, \Omega)]$$

$$= \quad \mathrm{const} + (a_\gamma^o - 1)\log\gamma - (a_\gamma^o - 1)b_\gamma^o\gamma$$

$$+ \quad \sum_b ((a_{\alpha_b}^o - 1)\log\alpha_b - (a_{\alpha_b}^o - 1)c_{\alpha_b}^o\alpha_b)$$

$$+ \quad \sum_b ((a_{\beta_b}^o - 1)\log\beta_b - (a_{\beta_b}^o - 1)c_{\beta_b}^o\beta_b)$$

$$+ \quad p\sum_{b=1}^{B}\log\alpha_b + P\sum_{b=1}^{B}\log\beta_b + p\log\gamma$$

$$- \quad \| \mathbf{C}\mathbf{y}_b(\Omega') \|^2 - \mathbf{tr}[\mathbf{C}^t\mathbf{C}\mathbf{cov}(\mathbf{y}_b|\mathbf{Y}, \mathbf{x}, \Omega')]$$

$$- \quad \| \mathbf{Y}_b - \mathbf{H}_b\mathbf{y}_b(\Omega') \|^2 - \mathbf{tr}[\mathbf{H}^t\mathbf{H}\mathbf{cov}(\mathbf{y}_b|\mathbf{Y}, \mathbf{x}, \Omega')]$$

$$- \quad \| \mathbf{x} - \sum_{b=1}^{B}\lambda_j\mathbf{y}_b(\Omega') \|^2$$

$$- \quad \sum_{i=1}^{B}\sum_{j=1}^{B}\lambda_i\lambda_j\mathbf{tr}[\mathbf{cov}(\mathbf{y}_i, \mathbf{y}_j|\mathbf{Y}, \mathbf{x}, \Omega')]$$

Image	Date	Path	Row
A	2000-07-30	200	031
B	2000-08-08	199	031

Table 1: Landsat 7 ETM+, L1G Orthorectified image sets.

Differentiating $\mathbf{U}(\Omega|\Omega')$ with respect to Ω we obtain at its maximum

$$\frac{1}{\alpha_b} = \frac{a_{\alpha_b}^o - 1}{p + a_{\alpha_b}^o - 1} c_{\alpha_b}^o + \frac{p}{p + a_{\alpha_b}^o - 1} \frac{\| \mathbf{C}\mathbf{y}_b(\Omega') \|^2}{p}$$

$$+ \frac{p}{p + a_{\alpha_b}^o - 1} \frac{\mathbf{tr}[\mathbf{C}^t\mathbf{C}\mathbf{cov}(\mathbf{y}_b|\mathbf{Y},\mathbf{x},\Omega')]}{p} \qquad \forall b \quad (24)$$

$$\frac{1}{\beta_b} = \frac{a_{\beta_b}^o - 1}{P + (a_{\beta_b}^o - 1)} c_{\beta_b}^o + \frac{P}{P + a_{\alpha_b}^o - 1} \frac{\| \mathbf{Y}_b - \mathbf{H}\mathbf{y}_b(\Omega') \|^2}{P}$$

$$+ \frac{P}{P + a_{\alpha_b}^o - 1} \frac{\mathbf{tr}[\mathbf{H}^t\mathbf{H}\mathbf{cov}(\mathbf{y}_b|\mathbf{Y},\mathbf{x},\Omega')]}{P} \quad \forall b \quad (25)$$

$$\frac{1}{\gamma} = \frac{a_\gamma^o - 1}{p + a_\gamma^o - 1} c_\gamma^o + \frac{p}{p + a_\gamma^o - 1} \frac{\| \mathbf{x} - \sum_{b=1}^B \lambda_b \mathbf{y}_b(\Omega') \|^2}{p}$$

$$+ \frac{p}{p + a_\gamma^o - 1} \sum_{i=1}^B \frac{\sum_{j=1}^B \lambda_i\lambda_j \mathbf{tr}[\mathbf{cov}(\mathbf{y}_i,\mathbf{y}_j|\mathbf{Y},\mathbf{x},\Omega')]}{p} \quad (26)$$

These equations provide a very interesting interpretation of the estimated hyperparameters at convergence. The hyperparameter estimates are a weighted sum of their maximum likelihood estimates ($a_\omega^o = 1, \forall \omega \in \Omega$) and the inverse of the modes of their hyperprior distributions (see Eq. (13)).

Based on the above development we propose the following algorithm for the estimation of the hyperparameters and the high-resolution image

Algorithm 1 Iterative estimation of $\hat\Omega$ and $\mathbf{y}(\hat\Omega)$.

1. *Choose Ω^0. Set $k = 0$.*
2. *Compute $\mathbf{y}(\Omega^0)$ using Eq. (17) for $\Omega = \Omega^0$.*
3. *Repeat*
   i) *Set $k=k+1$,*
   ii) *Use $\Omega' = \Omega^{k-1}$ in the right hand side of Eqs. (24), (25) and (26) to obtain Ω^k in the left hand side of these equations.*
   iii) *Compute $\mathbf{y}(\Omega^k)$ using Eq. (17) for $\Omega = \Omega^k$.*
until

$$\frac{\| \mathbf{y}(\Omega^k) - \mathbf{y}(\Omega^{k-1}) \|^2}{\| \mathbf{y}(\Omega^k) \|^2} < \varepsilon \qquad (27)$$

where ε is a prescribed bound.

5. EXPERIMENTAL RESULTS

The proposed algorithm has been tested on the set of Landsat ETM+ images [10] shown in Table 1. Multispectral image regions of interest of size 128×128 pixels, and their corresponding 256×256 pixel regions of the panchromatic image, were used in the experiments. Figure 3 displays the region of interest for image A in Table 1. The image on the left is a false (RGB) color image composed of bands 4, 3 and 2 of the Landsat ETM+ multispectral image.

(a)   (b)
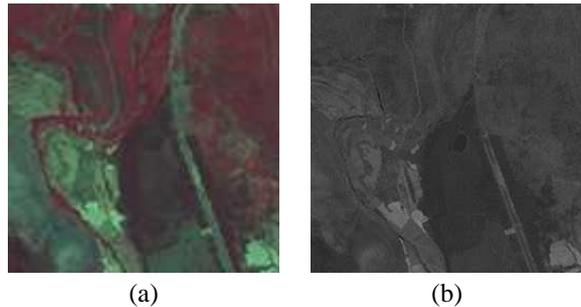
Figure 3: Details of the (a) false color multispectral image *A* and (b) panchromatic image *A* in Table 1.

Image	λ_1	λ_2	λ_3	λ_4
A	0.0467	0.1759	0.1883	0.5891
B	0.0118	0.1918	0.1495	0.4990

Table 2: Estimated λ_b values for the used images.

According to the ETM+ sensor spectral response in Fig. 2, the panchromatic image only covers the spectrum of a part of the first four bands of the multispectral image. Hence, we apply the proposed method with $B = 4$.

In order to apply Algorithm 1, we need to know the contribution of each band to the panchromatic image, that is, the values of λ_b, $b = 1, 2, \ldots, B$ in Eq. (9). These values have been obtained by solving for $\lambda = (\lambda_1, \ldots, \lambda_4)$

$$\hat\lambda = \arg\min_\lambda \| \mathbf{x} - \sum_{b=1}^4 \lambda_b \mathbf{Y}_b^{up} \|^2, \qquad (28)$$

where $\mathbf{Y}_b^{up}$ denotes the upsampled (by bicubic interpolation) version of $\mathbf{Y}_b$ to the size of the panchromatic image $\mathbf{x}$. The values obtained for λ for each image in Table 1 are shown in Table 2.

The initial set of parameters Ω^0 in Algorithm 1 were chosen from the bicubic interpolation of the multispectral image, $\mathbf{y}^0$, using the following equations

$$\alpha_b^0 = p / \| \mathbf{C}\mathbf{y}_b^0 \|^2, \quad b = 1, \ldots, B, \qquad (29)$$

$$\beta_b^0 = P / \| \mathbf{Y}_b - \mathbf{H}\mathbf{y}_b^0 \|^2, \quad b = 1, \ldots, B, \qquad (30)$$

$$\gamma^0 = p / \| \mathbf{x} - \sum_{b=1}^B \lambda_b \mathbf{y}_b^0 \|^2 . \qquad (31)$$

The value of $\varepsilon = 10^{-6}$ was used in Eq. (27) as the prescribed bound.

One important issue with Algorithm 1 is the selection of the hyperprior parameters. As we have already commented, if we select $a_\omega^o = 1, \forall \omega \in \Omega$, Algorithm 1 obtains the maximum likelihood estimate of the hyperparameters, thus making the observed data fully responsible for the estimation of the parameters. However, we usually have some information about the possible values of the parameters and, with the proposed method, we can incorporate this information into the estimation process. In this paper we chose such parameters by applying Algorithm 1 to each band b independently using $\lambda_b = 1$, $b = 1, \ldots, 4$, thus justifying the panchromatic image only from the spectral band b. This will provide us with estimates for α_b and β_b for each band $b = 1, \ldots, 4$, that we will use as values for $c_{\alpha_b}^o$ and $c_{\beta_b}^o$, $b = 1, \ldots, 4$, respectively. The

(a)

(b)

(c)

(d)
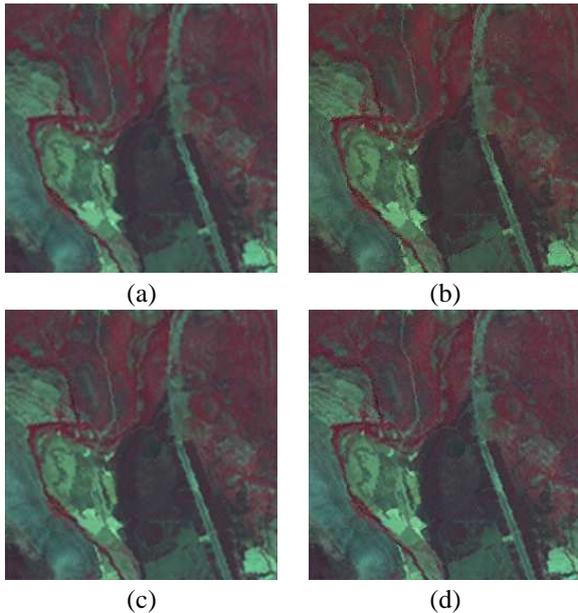
Figure 4: Reconstructions of image A using (a) bicubic interpolation, (b) the Price method, (c) the proposed method with $a_\omega^o = 1, \forall \omega \in \Omega$, and (d) the best reconstruction with the proposed method.

Image	band	MI	MII	Alg. 1, $a_\omega^o = 1$	Alg. 1
A	1	21.0	19.1	21.7	21.7
	2	19.8	18.1	20.4	20.6
	3	17.2	15.5	17.5	17.8
	4	20.1	17.8	20.1	20.6
B	1	18.1	17.0	18.8	18.9
	2	16.8	15.8	17.2	17.4
	3	14.8	13.7	15.1	15.3
	4	17.4	16.4	17.8	18.1

Table 3: PSNR (dB) obtained by bicubic interpolation (MI), the Price method [9] (MII), the proposed method with $a_\omega^o = 1, \forall \omega \in \Omega$ (Alg. 1, $a_\omega^o = 1$), and the best reconstruction by the proposed method (Alg. 1) for each band.

performs both bicubic interpolation and the method in [9]. In our future work non stationary image priors and the use of total variation methods will be considered.

algorithm also provides us with four estimates of γ_b, one for each band, that we combine as $c_\gamma^o = \frac{1}{4} \sum_{b=1}^4 \gamma_b$.

In our experiments we chose a set of five values for a_ω^o such that $a_\omega^o - 1/(N + a_\omega^o - 1)$, with $N = p$ or $N = P$ depending on the parameter ω, $\omega \in \Omega$, ranges from 0 to 1. Note that $(a_\omega^o - 1)/(N + a_\omega^o - 1)$ is equal to zero for $a_\omega^o = 1$ and it is equal to one for $a_\omega^o = \infty$. In this case, we practically do not take into account the maximum likelihood estimates of the hyperparameters and all the information on the hyperparameters is provided by c_ω^o.

Table 3 shows a numerical comparison in terms of PSNR$_b = 10 \log_{10} \left\{ 255^2 \times P / \| \mathbf{Y}_b - \mathbf{D}\mathbf{y}_b(\hat{\mathbf{\Omega}}) \|^2 \right\}$ of the reconstructions using bicubic interpolation, the Price method [9], the proposed method with $a_\omega^o = 1, \forall \omega \in \Omega$, and the proposed method when a_ω^o is chosen to obtain the best reconstruction. The best reconstruction is obtained, for image A with $a_{\alpha_b}^o = p - 1$, $a_{\beta_b}^o = P - 1, b = 1, \ldots, 4$, and $a_\gamma^o = \infty$ and, for image B, with $a_{\alpha_b}^o = 1$, $a_{\beta_b}^o = 1, b = 1, \ldots, 4$, and $a_\gamma^o = \infty$. Figure 4 shows the reconstructions corresponding to image A.

From the reported data it is clear that the proposed method outperforms bicubic interpolation and the Price method, both in quantitative and qualitative terms when no information for the hyperparameters is provided although better results are obtained when some prior knowledge is included into the estimation process.

6. CONCLUSIONS

In this paper the reconstruction of multispectral images has been formulated from a super resolution point of view. A hierachical Bayesian method for estimating both the reconstructed pansharpened images and the model parameters, within the Bayesian framework, has been proposed. Based on the presented experimental results, the new method out-

REFERENCES

[1] T. Akgun, Y. Altunbasak, and R.M. Mersereau. Super-resolution reconstruction of hyperspectral images. *IEEE Trans. on Image Proc.*, 14(11):1860–1875, 2005.

[2] J. O. Berger. *Statistical Decision Theory and Bayesian Analysis*, chapter 3 and 4. New York, Springer Verlag, 1985.

[3] A. D. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the E-M algorithm. *J. Roy. Stat. Soc. B*, 39:1–37, 1977.

[4] M.T. Eismann and R.C. Hardie. Hyperspectral resolution enhancement using high-resolution multispectral imaginary with arbitray response functions. *IEEE Trans. on Geosc. & Rem. Sens.*, 43(3):455–465, 2005.

[5] P. Chavez Jr., S. Sides, and J. Anderson. Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic. *Phot. Eng. & Rem. Sens.*, 57(3):295–303, 1991.

[6] R. Molina, A. K. Katsaggelos, and J. Mateos. Bayesian and regularization methods for hyperparameter estimation in image restoration. *IEEE Trans. on Image Proc.*, 8(2):231–246, 1999.

[7] R. Molina, J. Mateos, A.K. Katsaggelos, and R. Zurita-Milla. A new super resolution Bayesian method for pansharpening landsat ETM+ imagery. In S. Liang, J. Liu, X. Li, R. Liu, and M. Schaepman, editors, *9th Int. Sym. on Phys. Meas. and Signatures in Rem. Sens. (ISPMSRS)*, pages 280–283. Beijing (China), Oct 2005.

[8] J. Nuñez, X. Otazu, O. Fors, A. Prades, V. Pala, and R. Arbiol. Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Trans on Geosc. & Rem. Sens.*, 37(3):1204–1211, 1999.

[9] J.C. Price. Combining multispectral data of different spatial resolution. *IEEE Trans. on Geosc. & Rem. Sens.*, 37(3):1199–1203, 1999.

[10] NASA Landsat Program. Landsat ETM+ Scenes. In *Global Land Cover Facility*. U.S. Geological Survey, Sioux Falls, South Dakota, http://glcf.umiacs.umd.edu.

[11] V. Vijayaraj. A quantitative analysis of pansharpened images. Master's thesis, Mississippi St. Univ., 2004.