Bayesian fusion of hyperspectral and multispectral images

Qi Wei, Nicolas Dobigeon, Jean-Yves Tourneret

BAYESIAN FUSION OF HYPERSPECTRAL AND MULTISPECTRAL IMAGES

Qi Wei, Nicolas Dobigeon, and Jean-Yves Tourneret

University of Toulouse, IRIT/INP-ENSEEIHT, 31071 Toulouse cedex 7, France

ABSTRACT

This paper presents a Bayesian fusion technique for multi-band images. The observed images are related to the high spectral and high spatial resolution image to be recovered through physical degradations, e.g., spatial and spectral blurring and/or subsampling defined by the sensor characteristics. The fusion problem is formulated within a Bayesian estimation framework. An appropriate prior distribution related to the linear mixing model for hyperspectral images is introduced. To compute Bayesian estimators of the scene of interest from its posterior distribution, a Gibbs sampling algorithm is proposed to generate samples asymptotically distributed according to the target distribution. To efficiently sample from this high-dimensional distribution, a Hamiltonian Monte Carlo step is introduced in this Gibbs sampler. The efficiency of the proposed fusion method is evaluated with respect to several state-of-the-art fusion techniques.

Index Terms— Fusion, multispectral and hyperspectral images, Bayesian estimation, Gibbs sampler, Hamiltonian Monte Carlo.

1. INTRODUCTION

Fusing a high spatial and low spectral resolution image with an auxiliary image of higher spectral but lower spatial resolution, also known as multi-resolution image fusion, has been explored for many years [1, 2]. When considering remotely sensed images, an archetypal fusion task is the pansharpening, which consists of fusing a high spatial resolution panchromatic (PAN) image and low spatial resolution multispectral (MS) image [1, 3]. More recently, hyperspectral (HS) imaging, which consists of acquiring a same scene in several hundreds of contiguous spectral bands, has opened a new range of relevant applications, such as target detection [4] and spectral unmixing [5]. Naturally, to take advantage of the newest benefits offered by HS images, the problem of fusing HS and PAN images has been explored [6]. Capitalizing on decades of experience in MS pansharpening, most of the HS pansharpening approaches merely adapt existing algorithms for PAN and MS fusion [7]. Other methods are specifically designed to the HS pansharpening problem such as [8] or [9, 10] invoking super-resolution techniques. Conversely, the fusion of MS and HS images has been considered in fewer research works and is still a challenging problem because of the high dimensionality of the data to be processed. The fusion of MS and HS differs from traditional MS or HS pansharpening since both spatial and spectral information is contained in multi-band images. Therefore, a lot of pansharpening methods, such as component substitution [11] and relative spectral contribution [12] are inapplicable or inefficient for the HS/MS fusion problem. Since the fusion problem is ill-posed, Bayesian inference offers a convenient way to regularize the problem by defining an appropriate prior distribution for the scene of interest. Following this strategy, Hardie *et al.* proposed a Bayesian estimator for fusing the co-registered high spatial-resolution MS and high spectral-resolution HS images [13]. The estimator of [13] was implemented in the wavelet domain to improve denoising performance [14].

This work proposes to take into account constraints related to the fusion problem via appropriate prior distributions used to build a new Bayesian fusion model. Many works have defined a Bayesian model for the unknown scene by exploiting *a priori* information provided by one of the sensors [13][15]. In this paper, all the sensor specifications (including spectral and spatial responses) are exploited to properly design the spatial and spectral degradations affecting the image to be recovered (see [16] for more details about these degradations). To define the prior distribution assigned to the unknown image, we resort to geometrical considerations well admitted in the HS imaging literature devoted to the linear unmixing problem [5, 17]. In particular, the high spatial resolution HS image to be estimated is assumed to live in a lower dimensional subspace, which is a suitable hypothesis when the observed scene is composed of a finite number of macroscopic materials. Two Bayesian estimators are classically considered in the literature: the minimum mean square error (MMSE) and maximum a posteriori (MAP) estimators defined as the mean and mode of the posterior distribution. The determination of these estimators requires to solve multi-dimensional integration or optimization problems, which can be difficult to handle. For instance, algorithms designed to maximize the posterior distribution may suffer from the presence of local extrema preventing convergence towards the actual maximum of the posterior. In this paper, we propose to compute the MMSE and MAP estimators of the unknown scene by using samples generated according to the posterior distribution of interest. To bypass the difficulty of sampling in a high-dimensional subspace, a suitable Hamiltonian Monte Carlo (HMC) algorithm is investigated [18, 19]. It differs from the standard Metropolis-within-Gibbs algorithm by exploiting Hamiltonian evolution dynamics to propose states with higher acceptance ratio, reducing the correlation between successive samples.

The paper is organized as follows. Section 2 formulates the fusion problem in a Bayesian framework. In Section 3, we propose a new hierarchical Bayesian model defined by the joint posterior distribution of the unknown image, its parameters and hyperparameters. Section 4 studies a hybrid Gibbs sampler based on an HMC method to sample the posterior distribution of this Bayesian model. Simulations results are presented in Section 5 and conclusions are reported in Section 6.

2. PROBLEM FORMULATION

Let $\mathbf{Z}_1$ and $\mathbf{Z}_2$ denote the HS and MS images acquired by two different sensors for a same scene $\mathbf{X}$. The observed data are supposed to be degraded versions of the high-spectral and high-spatial resolution

Part of this work has been supported by the Chinese Scholarship Council (CSC, 201206020007) and by the Hypanema ANR Project n°ANR-12-BS03-003.

scene $\mathbf{X}$, according to the following observation model

$$\mathbf{Z}_p = \mathcal{F}_p(\mathbf{X}) + \mathbf{E}_p \tag{1}$$

with $p \in \{1,2\}$. In (1), $\mathcal{F}_p(\cdot)$ is a linear transformation that models the degradations affecting $\mathbf{X}$ for the pth sensor. These degradations may include spatial blurring, spectral blurring and decimation operation. In what follows, the images $\mathbf{Z}_p$ ($p=1,2$) and the unobserved scene $\mathbf{X}$ are assumed to be pixelated images of sizes $N_p = n_{\mathrm{x},p} \times n_{\mathrm{y},p} \times n_{\lambda,p}$ and $M = m_{\mathrm{x}} \times m_{\mathrm{y}} \times m_{\lambda}$, where $\cdot_{\mathrm{x}}$ and $\cdot_{\mathrm{y}}$ both refer to spatial dimensions and $\cdot_{\lambda}$ is for the spectral dimension and N_p is the total number of measurements in $\mathbf{Z}_p$. Classically, the observed image $\mathbf{Z}_p$ can be lexicographically ordered to build the $N_p \times 1$ vector $\mathbf{z}_p$. Without loss of generality, the band interleaved by pixel (BIP)-like vectorization scheme [20, pp. 103–104] is adopted in this work (see paragraph 3.2). As a consequence, the observation equation (1) can be rewritten as

$$\mathbf{z}_p = \mathbf{F}_p\mathbf{x} + \mathbf{e}_p \tag{2}$$

where the $M \times 1$ and $N_p \times 1$ vectors $\mathbf{x}$ and $\mathbf{e}_p$ are ordered versions of the scene $\mathbf{X}$ and noise $\mathbf{E}_p$. The noise vector $\mathbf{e}_p$ is assumed to be a white Gaussian sequence, i.e., $\mathbf{e}_p \sim \mathcal{N}\left(\mathbf{0}_{N_p}, s_p^2\mathbf{I}_{N_p}\right)$ where $\mathbf{0}_{N_p}$ is the $N_p \times 1$ vector made of zeros and $\mathbf{I}_{N_p}$ is the $N_p \times N_p$ identity matrix. Moreover, in (2), $\mathbf{F}_1$ is an $N_1 \times M$ matrix that reflects the spatial degradation $\mathcal{F}_1(\cdot)$ and $\mathbf{F}_2$ is an $N_2 \times M$ matrix that reflects the spectral degradation $\mathcal{F}_2(\cdot)$. Note that when using a single-band image $\mathbf{z}_1$, $\mathbf{F}_1$ is an $n_x n_y \times n_x n_y$ (generally sparse) Toeplitz matrix, that is symmetric for a symmetric convolution kernel $\boldsymbol{\kappa}_1$.

The problem addressed in this paper consists of estimating the high-spectral and high-spatial resolution image $\mathbf{x}$ by fusing the spatial and/or spectral information provided by the set of observed images $\mathbf{z} = \{\mathbf{z}_1, \mathbf{z}_2\}$. For this, we introduce a new hierarchical Bayesian model detailed in the next section.

3. HIERARCHICAL BAYESIAN MODEL

3.1. Likelihood function

Using the statistical properties of the noise vectors $\mathbf{e}_p$ ($p=1,2$), the distribution of $\mathbf{z}_p$ is clearly Gaussian with mean vector $\mathbf{F}_p\mathbf{x}$ and covariance matrix $s_p^2\mathbf{I}_{N_p}$, i.e.,

$$f\left(\mathbf{z}_p|\mathbf{x}, s_p^2\right) = \left(\frac{1}{2\pi s_p^2}\right)^{\frac{N_p}{2}} \exp\left(-\frac{1}{2s_p^2}\left\|\mathbf{z}_p - \mathbf{F}_p\mathbf{x}\right\|^2\right) \tag{3}$$

where $\|\mathbf{x}\| = \left(\mathbf{x}^T\mathbf{x}\right)^{\frac{1}{2}}$ is the ℓ_2-norm of $\mathbf{x}$. As mentioned in the previous section, the measurements are generally acquired by different sensors. Therefore, the observed vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ are assumed to be independent, conditionally upon the unobserved scene $\mathbf{x}$ and the noise variances s_1^2 and s_2^2, i.e., $f\left(\mathbf{z}|\mathbf{x}, \boldsymbol{s}^2\right) = f\left(\mathbf{z}_1|\mathbf{x}, s_1^2\right) f\left(\mathbf{z}_2|\mathbf{x}, s_2^2\right)$ with $\boldsymbol{s}^2 = (s_1^2, s_2^2)^T$.

3.2. Prior distributions

The likelihood of $\mathbf{z}$ is parameterized by the unknown scene $\mathbf{x}$ to be recovered and the vector of noise variances $\boldsymbol{s}^2$. In this section, prior distributions are introduced for these parameters.

Scene prior: Following a BIP strategy, the vectorized image $\mathbf{x}$ can be decomposed as $\mathbf{x} = \left(\boldsymbol{x}_1^T \cdots \boldsymbol{x}_{m_{\mathrm{x}}m_{\mathrm{y}}}^T\right)^T$, where $\boldsymbol{x}_i = (x_{i,1} \cdots x_{i,m_\lambda})^T$ is the $m_\lambda \times 1$ vector corresponding to the ith spatial location (with $i = 1, \cdots, m_{\mathrm{x}}m_{\mathrm{y}}$). Due to the linear mixing model, $\boldsymbol{x}_i$ lives in a subspace $\mathbb{R}^{\widetilde{m}_\lambda \times 1}$ where $\widetilde{m}_\lambda$ is much smaller than the number of bands m_λ [5, 21]. Therefore, we introduce a linear transformation from $\mathbb{R}^{m_\lambda \times 1}$ to $\mathbb{R}^{\widetilde{m}_\lambda \times 1}$ such that $\boldsymbol{u}_i = \mathbf{V}\boldsymbol{x}_i$, where $\boldsymbol{u}_i$ is the projection of the vector $\mathbf{x}_i$ onto the subspace of interest and the transformation matrix $\mathbf{V}$ is of size $\widetilde{m}_\lambda \times m_\lambda$. Using the notation $\mathbf{u} = \left(\boldsymbol{u}_1^T \cdots \boldsymbol{u}_{m_{\mathrm{x}}m_{\mathrm{y}}}^T\right)^T$, we have $\mathbf{u} = \mathfrak{V}\mathbf{x}$, where $\mathfrak{V}$ is an $\widetilde{M} \times M$ block-diagonal matrix whose blocks are equal to $\mathbf{V}$ and $\widetilde{M} = m_{\mathrm{x}}m_{\mathrm{y}}\widetilde{m}_\lambda$. A Gaussian prior is then assigned to the projected vectors $\boldsymbol{u}_i$ ($i = 1, \cdots, m_{\mathrm{x}}m_{\mathrm{y}}$)

$$\boldsymbol{u}_i|\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}} \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}}\right). \tag{4}$$

This Gaussian prior has been used successfully in many image processing applications including image denoising [22] and image restoration [23]. It has also the advantage of being a conjugate distribution for the likelihood function. Consequently, as it will be shown in Section 4, coupling this Gaussian prior distribution with the Gaussian likelihood leads to simpler estimators constructed from the posterior distribution of interest.

Noise variance priors: A non-informative Jeffreys' prior is assigned to the noise variances s_p^2 for $p=1,2$, i.e., $f\left(s_p^2\right) \propto \frac{1}{s_p^2}\mathbf{1}_{\mathbb{R}^+}\left(s_p^2\right)$, where $\mathbf{1}_{\mathbb{R}^+}(\cdot)$ is the indicator function defined on $\mathbb{R}^+$ (see [24] for motivations).

3.3. Hyperparameter priors

The hyperparameter vector associated with the parameter priors defined above is $\boldsymbol{\Phi} = \{\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}}\}$. The quality of the fusion algorithm investigated in this paper clearly depends on the value of the hyperparameters that need to be adjusted carefully. Instead of fixing these hyperparameters *a priori* , we propose to estimate them from the data by using a hierarchical Bayesian algorithm. This approach requires to define priors for the different hyperparameters (usually referred to as hyperpriors) which are summarized in this section.

Hyperparameter $\boldsymbol{\mu}_{\boldsymbol{u}}$: The hyperparameter $\boldsymbol{\mu}_{\boldsymbol{u}}$ is assigned a conjugate Gaussian distribution $\mathcal{N}\left(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0\right)$, where $\boldsymbol{\mu}_0$ and $\boldsymbol{\Sigma}_0$ are fixed to ensure a non-informative prior for $\boldsymbol{\mu}_{\boldsymbol{u}}$.

Hyperparameter $\boldsymbol{\Sigma}_{\boldsymbol{u}}$: Assigning a conjugate inverse-Wishart (IW) distribution to the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{u}}$ has provided interesting results in the signal and image processing literature [25,26]. Following these works, we have chosen $\boldsymbol{\Sigma}_{\boldsymbol{u}} \sim \mathcal{W}^{-1}(\boldsymbol{\Psi}, \eta)$, where $\mathcal{W}^{-1}$ is the IW distribution and its parameters $(\boldsymbol{\Psi}, \eta)^T$ are fixed to provide a non-informative prior for $\boldsymbol{\Sigma}_{\boldsymbol{u}}$.

3.4. Posterior distribution

The unknown parameter vector associated with the proposed hierarchical Bayesian fusion model is composed of the projected scene $\mathbf{u}$, the noise vector $\boldsymbol{s}^2$, i.e., $\boldsymbol{\theta} = \left\{\mathbf{u}, \boldsymbol{s}^2\right\}$, and the hyperparameters in $\boldsymbol{\Phi} = \{\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}}\}$. The joint posterior distribution of the unknown parameters and hyperparameters can be computed using the following hierarchical structure

$$f\left(\boldsymbol{\theta}, \boldsymbol{\Phi}|\mathbf{z}\right) \propto f\left(\mathbf{z}|\boldsymbol{\theta}\right) f\left(\boldsymbol{\theta}|\boldsymbol{\Phi}\right) f\left(\boldsymbol{\Phi}\right). \tag{5}$$

By assuming prior independence between the hyperparameters $\boldsymbol{\mu}_{\boldsymbol{u}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{u}}$ and the parameters $\mathbf{u}$ and $\boldsymbol{s}^2$ conditionally upon $\boldsymbol{\mu}_{\boldsymbol{u}}$ and

$\boldsymbol{\Sigma}_{\boldsymbol{u}}$, the following results can be obtained

$$\begin{aligned} f\left(\boldsymbol{\theta}|\boldsymbol{\Phi}\right) &= f\left(\mathbf{u}|\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}}\right) f\left(\boldsymbol{s}^2\right) \\ f\left(\boldsymbol{\Phi}\right) &= f\left(\boldsymbol{\mu}_{\boldsymbol{u}}\right) f\left(\boldsymbol{\Sigma}_{\boldsymbol{u}}\right). \end{aligned} \tag{6}$$

The posterior distribution $f\left(\mathbf{u}|\mathbf{z}\right)$, which is obtained by marginalizing out the hyperparameter vector $\boldsymbol{\Phi}$ and the noise variances $\boldsymbol{s}^2$ from the joint posterior $f\left(\boldsymbol{\theta}, \boldsymbol{\Phi}|\mathbf{z}\right)$ is difficult to handle to compute the MAP or MMSE estimators of the image $\mathbf{u}$. Instead, this paper proposes to use a Markov chain Monte Carlo (MCMC) method to generate a collection of N_{MC} samples $\mathcal{U} = \left\{\mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(N_{\text{MC}})}\right\}$ that are asymptotically distributed according to the posterior $f\left(\mathbf{u}|\mathbf{z}\right)$. The Bayesian estimators of $\mathbf{u}$ can be computed using these generated samples. For instance, the MMSE estimator of $\mathbf{u}$ can be approximated by an average of the generated samples $\hat{\mathbf{u}}_{\text{MMSE}} \approx \frac{1}{N_{\text{MC}} - N_{\text{bi}}} \sum_{t=N_{\text{bi}}+1}^{N_{\text{MC}}} \mathbf{u}^{(t)}$, where N_{bi} is the number of burn-in iterations required to reach the sampler convergence. The highly-resolved HS image can finally be computed as $\hat{\mathbf{x}}_{\text{MMSE}} = \left(\mathfrak{Y}^T\mathfrak{Y}\right)^{-1}\mathfrak{Y}^T\hat{\mathbf{u}}_{\text{MMSE}} = \mathfrak{Y}^T\hat{\mathbf{u}}_{\text{MMSE}}$. Since it is not easy to directly sample from $f\left(\mathbf{u}|\mathbf{z}\right)$, we propose to sample according to the joint posterior $f\left(\mathbf{u}, \boldsymbol{s}^2, \boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}}|\mathbf{z}\right)$ by using a Metropolis-within-Gibbs sampler, which can be easily implemented since all the conditional distributions associated with $f\left(\mathbf{u}, \boldsymbol{s}^2, \boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}}|\mathbf{z}\right)$ are simple. These conditional distributions are provided in Section 4.

4. HYBRID GIBBS SAMPLER

The Gibbs sampler has received much attention in the statistical community to solve Bayesian estimation problems [27] [28]. However, it cannot be used for generating samples distributed according to (5) because it is not possible to sample the conditional distribution of the projected image $\mathbf{u}$ (see Section 4.3 for more details). As a consequence, we propose to study a hybrid Gibbs sampler defined by the 4-step procedure detailed below.

4.1. Sampling the mean of the image $\boldsymbol{\mu}_{\boldsymbol{u}}$

Combining the prior of $\boldsymbol{\mu}_{\boldsymbol{u}}$ with (4), we obtain

$$\boldsymbol{\mu}_{\boldsymbol{u}}|\boldsymbol{\Sigma}_{\boldsymbol{u}}, \mathbf{u}, \boldsymbol{s}^2, \mathbf{z} \sim \mathcal{N}\left(\boldsymbol{\mu}_{\boldsymbol{\mu}_{\boldsymbol{u}}|\mathbf{u}}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}_{\boldsymbol{u}}|\mathbf{u}}\right) \tag{7}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\mu}_{\boldsymbol{u}}|\mathbf{u}} &= \boldsymbol{\Sigma}_{\boldsymbol{\mu}_{\boldsymbol{u}}|\mathbf{u}}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\mu}_0 + \boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}\sum_{i=1}^{m_{\text{x}}m_{\text{y}}}\boldsymbol{u}_i\right) \\ \boldsymbol{\Sigma}_{\boldsymbol{\mu}_{\boldsymbol{u}}|\mathbf{u}} &= \left(\boldsymbol{\Sigma}_0^{-1} + m_{\text{x}}m_{\text{y}}\boldsymbol{\Sigma}_{\boldsymbol{u}}^{-1}\right)^{-1}. \end{aligned} \tag{8}$$

The Gaussian distribution (7) is easy to sample since the size of $\boldsymbol{\Sigma}_{\boldsymbol{\mu}_{\boldsymbol{u}}|\mathbf{u}}$ is $\widetilde{m}_\lambda \times \widetilde{m}_\lambda$, where $\widetilde{m}_\lambda$ is the actual dimension of the space where the data live, which is generally lower than 10.

4.2. Sampling the covariance matrix of the image $\boldsymbol{\Sigma}_{\boldsymbol{u}}$

Standard computations yield the following IW distribution as conditional distribution for the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{u}}$

$$\boldsymbol{\Sigma}_{\boldsymbol{u}}|\boldsymbol{\mu}_{\boldsymbol{u}}, \mathbf{u}, \boldsymbol{s}^2, \mathbf{z} \sim \mathcal{W}^{-1}\left(\boldsymbol{\Psi} + \sum_{i=1}^{m_{\text{x}}m_{\text{y}}}(\boldsymbol{u}_i - \boldsymbol{\mu}_{\boldsymbol{u}})^T(\boldsymbol{u}_i - \boldsymbol{\mu}_{\boldsymbol{u}}), m_{\text{x}}m_{\text{y}} + \eta\right).$$

This IW distribution is easy to sample using standard generators.

4.3. Sampling the projected image $\mathbf{u}$

Choosing the conjugate distribution (4) as prior for the projected image $\mathbf{u}$ leads to the conditional posterior distribution

$$\mathbf{u}|\boldsymbol{\mu}_{\boldsymbol{u}}, \boldsymbol{\Sigma}_{\boldsymbol{u}}, \boldsymbol{s}^2, \mathbf{z} \sim \mathcal{N}\left(\boldsymbol{\mu}_{\mathbf{u}|\mathbf{z}}, \boldsymbol{\Sigma}_{\mathbf{u}|\mathbf{z}}\right) \tag{9}$$

with

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathbf{u}|\mathbf{z}} &= \left(\boldsymbol{\Sigma}_{\mathbf{u}}^{\star\,-1} + \sum_{p=1}^{P}\tfrac{1}{s_p^2}\mathfrak{Y}\mathbf{F}_p^T\mathbf{F}_p\mathfrak{Y}^T\right)^{-1} \\ \boldsymbol{\mu}_{\mathbf{u}|\mathbf{z}} &= \boldsymbol{\Sigma}_{\mathbf{u}|\mathbf{z}}\left(\sum_{p=1}^{P}\tfrac{1}{s_p^2}\mathfrak{Y}\mathbf{F}_p^T\mathbf{z}_p + \boldsymbol{\Sigma}_{\mathbf{u}}^{\star\,-1}\boldsymbol{\mu}_{\mathbf{u}}^{\star}\right) \end{aligned}$$

and $\boldsymbol{\mu}_{\mathbf{u}}^{\star} = (\underbrace{\boldsymbol{\mu}_{\boldsymbol{u}}^T \cdots \boldsymbol{\mu}_{\boldsymbol{u}}^T}_{m_{\text{x}}m_{\text{y}}})^T, \boldsymbol{\Sigma}_{\mathbf{u}}^{\star} = \text{diag}(\underbrace{\boldsymbol{\Sigma}_{\boldsymbol{u}} \cdots \boldsymbol{\Sigma}_{\boldsymbol{u}}}_{m_{\text{x}}m_{\text{y}}})$.

Note that $\boldsymbol{\Sigma}_{\mathbf{u}|\mathbf{z}}$ is difficult to obtain since $\mathbf{F}_p$ is a high-dimensional matrix and thus inverting the matrix in $\boldsymbol{\Sigma}_{\mathbf{u}|\mathbf{z}}$ is very complicated. As a consequence, sampling directly from the above distribution is not possible and makes the standard Gibbs sampler inapplicable here. In this paper, we propose to use an HMC method to generate vectors distributed according to the Gaussian distribution (9). More details about the proposed HMC method are available in [29] and are omitted here for space limitations.

4.4. Sampling the vector of noise variances $\boldsymbol{s}^2$

The conditional distributions of the noise variances s_p^2 for $p = 1, 2$ are the following inverse-gamma distributions

$$s_p^2|\mathbf{u}, \mathbf{z} \sim \mathcal{IG}\left(\frac{N_p}{2}, \frac{\left\|\mathbf{z}_p - \mathbf{F}_p\mathfrak{Y}^T\mathbf{u}\right\|^2}{2}\right) \tag{10}$$

which are straightforward to sample.

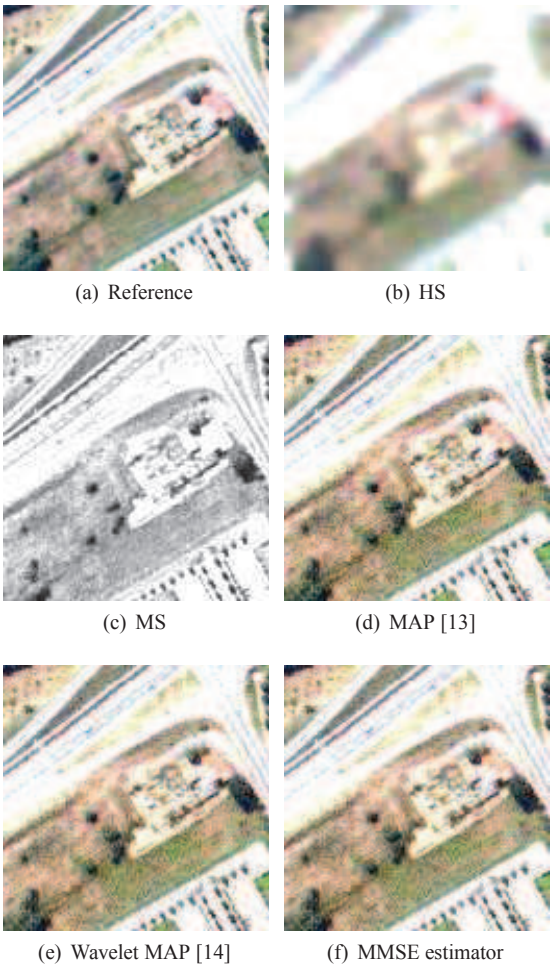
5. SIMULATION RESULTS

This section studies the performance of the proposed Bayesian fusion algorithm. The reference image, considered here as the high spatial and high spectral image, is a $128 \times 128 \times 103$ HS image acquired over Pavia, Italy, by the Reflective Optics System Imaging Spectrometer (ROSIS). This image was initially composed of 115 bands that have been reduced to 103 bands after removing the water vapor absorption bands.

5.1. Simulation scenario

We propose to reconstruct the reference HS image $\mathbf{u}$ from two HS and MS images $\mathbf{z}_1$ and $\mathbf{z}_2$. First, a high-spectral and low-spatial resolution image $\mathbf{z}_1$ (HS image) has been generated by applying a 17×17 blurring filter and down-sampling every 4 pixels in both vertical and horizontal direction on each band of the reference image. Second, a 4-band MS image $\mathbf{z}_2$ is obtained by filtering $\mathbf{u}$ with the IKONOS reflectance spectral responses. The HS and MS images have been both contaminated by zero-mean additive Gaussian noises with signal to noise ratios $\text{SNR}_1 = 30$dB and $\text{SNR}_2 = 20$dB, where $\text{SNR}_p = 10\log_{10}\left(\frac{\|\mathbf{z}_p\|_2^2}{N_p s_p^2}\right)$ for $p = 1, 2$. A composite color image, formed by selecting the red, green and blue bands of the reference image is shown in Fig. 1(a). The noise-contaminated HS and MS images are depicted in Fig. 1(b) and Fig. 1(c) (the HS image has been interpolated for better visualization).

To learn the matrix $\mathbf{V}$, we propose to use the principal component analysis (PCA). Note that other dimensionality reduction techniques such as [30, 31] could also be used. However, the PCA has



(a) Reference (b) HS (c) MS (d) MAP [13] (e) Wavelet MAP [14] (f) MMSE estimator

Fig. 1. (a) Reference. (b) HS image. (c) MS image. (d) MAP estimator. (e) Wavelet MAP estimator. (f) Proposed estimator.

been used here for simplicity. Precisely, the dimension of the projection subspace $\widetilde{m}_\lambda$ is defined as the minimum integer $\widetilde{m}_\lambda$ satisfying the condition $(\sum_{i=1}^{\widetilde{m}_\lambda} d_i)/(\sum_{i=1}^{m_\lambda} d_i) \geq 0.985$ where $d_1 \geq d_2 \geq ... \geq d_{m_\lambda}$ are the eigenvalues of the sample covariance matrix. For this example, the first $\widetilde{m}_\lambda = 5$ eigenvectors contain 98.5% of the information.

5.2. Fusion performance

To evaluate the quality of the proposed fusion strategy, different image quality measures can be investigated. Referring to [14], we propose to use the reconstruction error (RE), the averaged spectral angle mapper (SAM) and the universal image quality index (UIQI) as quantitative measures. The reconstruction error (RE) is related to the difference between the actual and fused images $\text{RE}(\mathbf{x}, \hat{\mathbf{x}}) = 10 \log_{10} \left(\frac{\|\mathbf{x}-\hat{\mathbf{x}}\|_2^2}{\|\mathbf{x}\|^2} \right)$. The smaller RE, the better the fusion and vice versa. The definition of SAM and UIQI can be found in [14].

Our experiments compare the proposed hierarchical Bayesian method with two state-of-the-art fusion algorithms for MS and HS images [13,14]. Fusion results obtained with the different algorithms are depicted in Fig. 1(d), 1(e) and 1(f). Graphically, the proposed algorithm performs competitively with the other methods. Quantitative results reported in Table 1 show that the proposed method slightly outperforms the methods of [13] and [14] (at the price of a higher computational complexity). This result can be explained by the fact that the proposed method explicitly exploits the sensor characteristics. Of course this improvement also depends on the nature of the images to be fused and on the degradation operator. Moreover, an interesting property of the proposed Bayesian method is that it allows noise variances to be estimated from the samples generated by the Gibbs sampler. The performance of the MMSE estimator of s_1^2 (resp. s_2^2) for a fixed value of s_2^2 (resp. s_1^2) is illustrated in Fig. 2. These results show that the noise variances can be estimated with good performance, especially at high values of the SNRs.

Table 1. Performance of different fusion methods.

Methods	RE(dB)	UIQI	SAM(deg)	Time(s)
Hardie	-16.534	0.9452	9.0501	0.8
Zhang	-17.022	0.9519	8.5039	16.9
Proposed	-17.340	0.9558	8.1980	4186.5

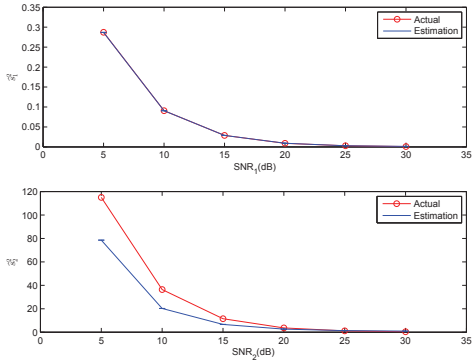


Fig. 2. Noise variances and their MMSE estimates. Top: HS Image ($\text{SNR}_2 = 30$dB). Bottom: MS Image ($\text{SNR}_1 = 30$dB).

6. CONCLUSIONS

This paper proposed a hierarchical Bayesian model for the fusion of multispectral and hyperspectral images. The image to be recovered was assumed to be degraded by physical transformations included within a forward model. We defined an appropriate prior distribution exploiting geometrical concepts used for the spectral unmixing of hyperspectral images. This prior distribution was used to define a new hierarchical Bayesian fusion model. The posterior distribution of this model was sampled using a Hamiltonian Monte Carlo algorithm. Simulations conducted on realistic multispectral and hyperspectral images showed that the proposed method gives slightly better results than two state-of-the-art fusion techniques. These improvements can be attributed to the proposed observation model, which explicitly exploits the sensor characteristics. Moreover, the proposed method has several advantages: 1) it allows the noise variances to be estimated jointly with the image to be recovered, 2) it can be generalized to more complicated fusion models such as those based on non-Gaussian image priors. Future work includes the development of similar fusion algorithms accounting for band-dependent noise variances or/and imperfect knowledge about the linear operators.

7. REFERENCES

[1] I. Amro, J. Mateos, M. Vega, R. Molina, and A. K. Katsaggelos, "A survey of classical methods and new trends in pansharpening of multispectral images," *EURASIP J. Adv. Signal Process.*, vol. 2011, no. 1, pp. 1–22, 2011.

[2] L. Wald, "Some terms of reference in data fusion," *IEEE Trans. Geosci. and Remote Sens.*, vol. 37, no. 3, pp. 1190 –1193, May 1999.

[3] D. Liu and P. T. Boufounos, "Dictionary learning based pansharpening," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Kyoto, Japan, March 2012, pp. 2397–2400.

[4] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, jan 2002.

[5] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observations and Remote Sens.*, vol. 5, no. 2, pp. 354–379, 2012.

[6] M. Cetin and N. Musaoglu, "Merging hyperspectral and panchromatic image data: qualitative and quantitative analysis," *Int. J. Remote Sens.*, vol. 30, no. 7, pp. 1779–1804, 2009.

[7] C. Chisense, J. Engels, M. Hahn, and E. Gülch, "Pansharpening of hyperspectral images in urban areas," in *Proc. XXII Congr. of the Int. Society for Photogrammetry, Remote Sens.*, Melbourne, Australia, 2012.

[8] G. Chen, S.-E. Qian, J.-P. Ardouin, and W. Xie, "Super-resolution of hyperspectral imagery using complex ridgelet transform," *Int. J. of Wavelets, Multiresolution and Inform. Processing*, vol. 10, no. 03, 2012.

[9] Y. Zhao, J. Yang, Q. Zhang, L. Song, Y. Cheng, and Q. Pan, "Hyperspectral imagery super-resolution by sparse representation and spectral regularization," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, pp. 1–10, 2011.

[10] H. Zhang, L. Zhang, and H. Shen, "A super-resolution reconstruction algorithm for hyperspectral images," *Signal Processing*, vol. 92, no. 9, pp. 2082–2096, 2012.

[11] V. Shettigara, "A generalized component substitution technique for spatial enhancement of multispectral images using a higher resolution data set," *Photogramm. Eng. Remote Sens.*, vol. 58, no. 5, pp. 561–567, 1992.

[12] J. Zhou, D. Civco, and J. Silander, "A wavelet transform method to merge landsat tm and spot panchromatic data," *Int. J. Remote Sens.*, vol. 19, no. 4, pp. 743–757, 1998.

[13] R. C. Hardie, M. T. Eismann, and G. L. Wilson, "MAP estimation for hyperspectral image resolution enhancement using an auxiliary sensor," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1174–1184, Sept. 2004.

[14] Y. Zhang, S. De Backer, and P. Scheunders, "Noise-resistant wavelet-based Bayesian fusion of multispectral and hyperspectral images," *IEEE Trans. Geosci. and Remote Sens.*, vol. 47, no. 11, pp. 3834 –3843, Nov. 2009.

[15] M. Joshi and A. Jalobeanu, "MAP estimation for multiresolution fusion in remotely sensed images using an IGMRF prior model," *IEEE Trans. Geosci. and Remote Sens.*, vol. 48, no. 3, pp. 1245–1255, March 2010.

[16] X. Otazu, M. Gonzalez-Audicana, O. Fors, and J. Nunez, "Introduction of sensor spectral response into image fusion methods. Application to wavelet-based methods," *IEEE Trans. Geosci. and Remote Sens.*, vol. 43, no. 10, pp. 2376–2385, 2005.

[17] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, 2009.

[18] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid Monte Carlo," *Physics Lett. B*, vol. 195, no. 2, pp. 216–222, Sept. 1987.

[19] R. M. Neal, "MCMC using Hamiltonian dynamics," *Handbook of Markov Chain Monte Carlo*, vol. 54, pp. 113–162, 2010.

[20] J. B. Campbell, *Introduction to remote sensing*, 3rd ed. New-York, NY: Taylor & Francis, 2002.

[21] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[22] P. J. Liu, "Using Gaussian process regression to denoise images and remove artefacts from microarray data," Ph.D. dissertation, University of Toronto, 2007.

[23] N. A. Woods, N. P. Galatsanos, and A. K. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 201–213, Jan. 2006.

[24] C. P. Robert and G. Casella, *Monte Carlo statistical methods*. New York, NY, USA: Springer-Verlag, 2004.

[25] S. Bidon, O. Besson, and J.-Y. Tourneret, "The adaptive coherence estimator is the generalized likelihood ratio test for a class of heterogeneous environments," *IEEE Signal Process. Lett.*, vol. 15, pp. 281–284, 2008.

[26] M. Bouriga and O. Féron, "Estimation of covariance matrices based on hierarchical inverse-wishart priors," *J. of Stat. Planning and Inference*, 2012.

[27] G. Casella and E. I. George, "Explaining the Gibbs sampler," *The American Statistician*, vol. 46, no. 3, pp. 167–174, 1992.

[28] C. P. Robert, *The Bayesian Choice: from Decision-Theoretic Motivations to Computational Implementation*, 2nd ed., ser. Springer Texts in Statistics. New York, NY, USA: Springer-Verlag, 2007.

[29] Q. Wei, N. Dobigeon, and J.-Y. Tourneret, "Bayesian fusion of multi-band images," *arXiv preprint arXiv:1307.5996*, 2013.

[30] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. and Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, 2008.

[31] N. Acito, M. Diani, and G. Corsini, "Hyperspectral signal subspace identification in the presence of rare signal components," *IEEE Trans. Geosci. and Remote Sens.*, vol. 48, no. 4, pp. 1940–1954, 2010.