FAST TOTAL VARIATION IMAGE RESTORATION WITH PARAMETER ESTIMATION USING BAYESIAN INFERENCE

Bruno Amizic[1], S. Derin Babacan[1], Michael K. Ng[2], Rafael Molina[3], Aggelos K. Katsaggelos[1]

[1]Department of Electrical Engineering
and Computer Science
Northwestern University, USA
{amizic,sdb}@northwestern.edu
aggk@eecs.northwestern.edu

[2]Dept. of Mathematics
Hong Kong Baptist University
mng@math.hkbu.edu.hk

[3]Departamento de Ciencias
de la Computación e I.A.
Universidad de Granada, Spain
rms@decsai.ugr.es

ABSTRACT

In this paper we propose two fast Total Variation (TV) based algorithms for image restoration by utilizing variational posterior distribution approximation. The unknown image and the hyperparameters for the image and observation models are formulated and estimated simultaneously within a hierachical Bayesian framework, rendering the algorithms fully-automated without any free parameters. Experimental results demonstrate that the proposed algorithms provide restoration results competitive to existing methods in terms of image quality while achieving superior computational efficiency.

Index Terms— Bayesian methods, image restoration, total variation, variational methods, parameter estimation.

1. INTRODUCTION

The image degradation process is often represented by a linear model as

$$\mathbf{y} = \mathbf{Hx} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{n}$ represent the original image, the observed image, and the noise, respectively, all ordered lexicographically. The matrix $\mathbf{H}$ represents the blurring matrix, which is assumed to be known. It is also assumed that $\mathbf{n}$ is sampled from a zero-mean independent Gaussian random process with variance β^{-1}.

The image restoration problem is to find an estimate of $\mathbf{x}$ from $\mathbf{y}$ and $\mathbf{H}$ using prior knowledge about $\mathbf{n}$ and $\mathbf{x}$. The literature on image restoration is rich (a review and classification of the major approaches can be found for example in [1]).

In this paper we adopt the Bayesian formulation of the restoration problem (see, for example, [2, 3, 4]) to jointly estimate the image and the parameters using a TV image prior. In order to do so we introduce a hidden unknown image $\mathbf{u}$ which is assumed to be obtained from a TV prior and whose contribution is to produce a final restoration which is close, in the Gaussian sense, to $\mathbf{x}$. This idea was first presented in [5] as a way to perform a fast TV based image restoration from a regularization point of view, We explore here the same model but from the Bayesian point of view, which allows the estimation of all the parameters involved in the model and achieve the same computational efficiency as the method proposed in [5].

This paper is organized as follows. In Section 2 we present the hierarchical Bayesian model. Section 3 describes the variational inference methods and the derivation of the proposed methods. We present the experimental results in Section 4 and conclude in Section 5.

2. BAYESIAN MODELING

In this work, we utilize an hierarchical Bayesian model which consists of two stages: First, the image and the observation noise are modeled using a hidden unknown image and unknown hyperparameters; and second, a prior model on the unknown image and hyperprior distributions on the hyperparameters are introduced.

2.1. Prior and observation models

Given the observation model in Eq. (1), the corresponding probability distribution can be stated as

$$p(\mathbf{y}|\mathbf{x}, \beta) \propto \beta^{N/2} \exp\left[-\frac{\beta}{2} \parallel \mathbf{y} - \mathbf{Hx} \parallel^2\right]. \tag{2}$$

Instead of using a TV prior directly as the image model, we assume that the original image is a noisy Gaussian realization of another image $\mathbf{u}$, that is,

$$p(\mathbf{x}|\mathbf{u}, \alpha) \propto \alpha^{N/2} \exp\left[-\frac{\alpha}{2} \parallel \mathbf{x} - \mathbf{u} \parallel^2\right], \tag{3}$$

The properties of the image $\mathbf{u}$ will be described later. Then, the image $\mathbf{u}$ is modeled as a realization of a TV prior distribution, that is,

$$p(\mathbf{u}|\gamma) \propto \frac{1}{Z_{\mathrm{TV}}(\gamma)} \exp\left[-\gamma \mathrm{TV}(\mathbf{u})\right], \tag{4}$$

where $Z_{\mathrm{TV}}(\gamma)$ is the partition function and

$$\mathrm{TV}(\mathbf{u}) = \sum_i \sqrt{(\Delta_i^h(\mathbf{u}))^2 + (\Delta_i^v(\mathbf{u}))^2}, \tag{5}$$

where the operators $\Delta_i^h(\mathbf{u})$ and $\Delta_i^v(\mathbf{u})$ correspond to, respectively, horizontal and vertical first order differences, at pixel i, that is, $\Delta_i^h(\mathbf{u}) = u_i - u_{l(i)}$ and $\Delta_i^v(\mathbf{u}) = u_i - u_{a(i)}$, where $l(i)$ and $a(i)$ denote the nearest neighbors of i, to the left and above, respectively. We can approximate the partition function $Z_{\mathrm{TV}}(\gamma)$ using

$$\int_u \int_v \exp\left[-\gamma\sqrt{u^2 + v^2}\right] du\, dv = 2\pi/\gamma^2, \tag{6}$$

as proposed in [6], to obtain

$$p(\mathbf{u}|\gamma) \propto \gamma^{N/2} \exp\left[-\gamma \mathrm{TV}(\mathbf{u})\right], \tag{7}$$

where N is the size of the original image $\mathbf{x}$ and the hidden image $\mathbf{u}$.

2.2. Hyperprior on the hyperparameters

In this work we use flat improper hyperpriors on γ, α and β, that is, we utilize

$$\mathrm{p}(\gamma) \propto \mathrm{const}, \qquad \mathrm{p}(\alpha) \propto \mathrm{const}, \qquad \mathrm{p}(\beta) \propto \mathrm{const}. \quad (8)$$

Note that with this choice of the hyperpriors the observation $\mathbf{y}$ is fully responsible for the estimation of the image and hyperparameters.

The joint distribution of all unknowns can finally be found by combining Eqs. (2), (3), (4) and (8) as

$$
\begin{aligned}
\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{y}) &= \mathrm{p}(\gamma)\mathrm{p}(\alpha)\mathrm{p}(\beta)\mathrm{p}(\mathbf{u}|\gamma)\mathrm{p}(\mathbf{x}|\mathbf{u}, \alpha)\mathrm{p}(\mathbf{y}|\mathbf{x}, \beta) \\
&\propto \gamma^{N/2}\alpha^{N/2}\beta^{N/2} \exp\left[-\gamma \mathrm{TV}(\mathbf{u})\right] \\
&\times \exp\left[-\frac{\alpha}{2}\parallel \mathbf{x}-\mathbf{u}\parallel^2\right]\exp\left[-\frac{\beta}{2}\parallel \mathbf{y}-\mathbf{Hx}\parallel^2\right]. \quad (9)
\end{aligned}
$$

3. BAYESIAN INFERENCE AND VARIATIONAL APPROXIMATION

The Bayesian paradigm dictates that inference on $(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x})$ should be based on the posterior distribution

$$\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x} \mid \mathbf{y}) = \frac{\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{y})}{\mathrm{p}(\mathbf{y})}, \quad (10)$$

which is intractable. Therefore, we resort to variational methods to approximate it by a tractable distribution of the form $\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}) = \mathrm{q}(\gamma, \alpha, \beta)\mathrm{q}(\mathbf{u})\mathrm{q}(\mathbf{x})$. The variational criterion used to find $\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x})$ is the minimization of the Kullback-Leibler divergence [7], given by

$$
\begin{aligned}
&C_{KL}(\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}) \parallel \mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}|\mathbf{y})) \\
&= \int \mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}) \log\left(\frac{\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x})}{\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}|\mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \\
&= \int \mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}) \log\left(\frac{\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x})}{\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \\
&+ \mathrm{const},
\end{aligned}
$$
$$(11)$$

which is always non negative and equal to zero only when the distributions $\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x})$ and $\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}|\mathbf{y})$ coincide.

We now examine two approximations of the posterior distribution.

3.1. Approximation by degenerate distributions

We first assume that the distributions $\mathrm{q}(\gamma, \alpha, \beta)$, $\mathrm{q}(\mathbf{u})$ and $\mathrm{q}(\mathbf{x})$ are degenerate distributions, that is, they take one value with probability one and the rest of the values with probability zero. Using this assumption, the following procedure is obtained to estimate γ, α, β, $\mathbf{u}$, $\mathbf{x}$.

Algorithm 1 *Given initial estimates $\gamma^1, \alpha^1, \beta^1$ of the hyperparameters and the initial estimate $\mathbf{u}^1$ of the hidden image,*
For $k = 1, 2, \ldots$ until a stopping criterion is met:

1. *Find*

$$\mathbf{x}^k = \arg\min_{\mathbf{x}} \frac{\alpha^k}{2}\parallel \mathbf{x}-\mathbf{u}^k\parallel^2 + \frac{\beta^k}{2}\parallel \mathbf{y}-\mathbf{Hx}\parallel^2 \quad (12)$$

2. *Find*

$$\mathbf{u}^{k+1} = \arg\min_{\mathbf{u}} \gamma^k\mathrm{TV}(\mathbf{u}) + \frac{\alpha^k}{2}\parallel \mathbf{x}^k-\mathbf{u}\parallel^2 \quad (13)$$

3. *Find*

$$\gamma^{k+1} = \frac{N/2}{\mathrm{TV}(\mathbf{u}^{k+1})} \quad (14)$$

$$\alpha^{k+1} = \frac{N}{\parallel \mathbf{u}^{k+1}-\mathbf{x}^k\parallel^2} \quad (15)$$

$$\beta^{k+1} = \frac{N}{\parallel \mathbf{y}-\mathbf{Hx}^k\parallel^2} \quad (16)$$

It is very interesting to note that, without the estimation of the hyperparameters, the procedure to estimate the original as well as the sparse hidden variable coincides with the model proposed in [5]. Consequently, the proposed Bayesian algorithm represents a formulation of the method proposed in [5] which also allows the estimation of the involved parameters. Notice also that procedure based on the method in [8] to find the restoration proposed in [5] can also be used here to find the solution of Eqs. (12) and (13). As demonstrated in Section 4, this procedure significantly improves the computational speed of the algorithm in combination with the priors in Eqs. (3) and (4).

3.2. Approximation by non-degenerate distributions

Let us now examine the case where the distributions $\mathrm{q}(\gamma, \alpha, \beta)$, $\mathrm{q}(\mathbf{u})$ and $\mathrm{q}(\mathbf{x})$ are not constrained to be degenerate. Due to the form of the TV prior, it is difficult to evaluate the integral in Eq. (11). We therefore utilize a minorization of the TV prior. Let us define, for γ, $\mathbf{u}$ and a $N-$dimensional vector $\mathbf{v} \in (R^+)^N$, with components v_i, $i = 1, \ldots, N$, the following functional

$$\mathrm{M}(\gamma, \mathbf{u}, \mathbf{v}) = \gamma^{N/2} \exp\left[-\frac{\gamma}{2}\sum_i \frac{(\Delta_i^h(\mathbf{u}))^2 + (\Delta_i^v(\mathbf{u}))^2 + v_i}{\sqrt{v_i}}\right] \quad (17)$$

Next consider the following inequality [6], which states that for two real numbers $z \geq 0$ and $v > 0$

$$\sqrt{z} \leq \sqrt{v} + \frac{1}{2\sqrt{v}}(z-v). \quad (18)$$

Utilizing this inequality in Eq. (17), it is clear that $\mathrm{M}(\gamma, \mathbf{u}, \mathbf{v})$ is a lower bound of the prior $\mathrm{p}(\mathbf{u}|\gamma)$, that is,

$$\mathrm{p}(\mathbf{u}|\gamma) \geq \mathrm{const} \cdot \mathrm{M}(\gamma, \mathbf{u}, \mathbf{v}), \quad (19)$$

which leads to the following lower bound of the joint probability distribution

$$
\begin{aligned}
\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{y}) &\geq \mathrm{p}(\gamma)\mathrm{p}(\alpha)\mathrm{p}(\beta)\mathrm{M}(\mathbf{u}, \mathbf{v})\mathrm{p}(\mathbf{x}|\mathbf{u}, \alpha)\mathrm{p}(\mathbf{y}|\mathbf{x}, \beta) \\
&= \mathrm{F}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{v}, \mathbf{y}). \quad (20)
\end{aligned}
$$

Utilizing this lower bound, we obtain

$$
\begin{aligned}
&\int \mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}) \log\left(\frac{\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x})}{\mathrm{p}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}|\mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \\
&\leq \int \mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}) \log\left(\frac{\mathrm{q}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x})}{\mathrm{F}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{v}, \mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \\
&\hspace{8cm} + \mathrm{const} \quad (21)
\end{aligned}
$$

which presents an upper bound of the KL distance in Eq. (11). This upper bound is easier to evaluate than the actual KL distance. Therefore, $\mathrm{F}(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{v}, \mathbf{y})$ can be used as the majorizing functional of the joint distribution in Eq. (9). Proceeding in this fashion, we obtain the the following algorithm for evaluating the approximating posteriors $\mathrm{q}(\mathbf{u})$, $\mathrm{q}(\mathbf{x})$ and $\mathrm{q}(\gamma, \alpha, \beta)$.

Algorithm 2 *Given* $q^1(\gamma, \alpha, \beta)$*, an initial estimate of the distribution* $q(\gamma, \alpha, \beta)$*, and* $q^1(\mathbf{u})$*, an initial estimate of* $q(\mathbf{u})$ *and* $\mathbf{v}^1 \in (R^+)^N$*,*

For $k = 1, 2, \ldots$ *until a stopping criterion is met:*

1. Find

$$q^k(\mathbf{x}) = \arg\min_{q(\mathbf{x})} \int q^k(\gamma, \alpha, \beta) q^k(\mathbf{u}) q(\mathbf{x})$$

$$\times \log\left(\frac{q^k(\gamma, \alpha, \beta) q^k(\mathbf{u}) q(\mathbf{x})}{F(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{v}^k, \mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \quad (22)$$

2. Find

$$q^{k+1}(\mathbf{u}) = \arg\min_{q(\mathbf{u})} \int q^k(\gamma, \alpha, \beta) q(\mathbf{u}) q^k(\mathbf{x})$$

$$\times \log\left(\frac{q^k(\gamma, \alpha, \beta) q(\mathbf{u}) q^k(\mathbf{x})}{F(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{v}^k, \mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \quad (23)$$

3. Find

$$\mathbf{v}^{k+1} = \arg\min_{\mathbf{v}} \int q^k(\gamma, \alpha, \beta) q^{k+1}(\mathbf{u}) q^k(\mathbf{x})$$

$$\times \log\left(\frac{q^k(\gamma, \alpha, \beta) q^{k+1}(\mathbf{u}) q^k(\mathbf{x})}{F(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{v}, \mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \quad (24)$$

4. Find

$$q^{k+1}(\gamma, \alpha, \beta) = \arg\min_{q(\alpha, \beta)} \int_q q(\gamma, \alpha, \beta) q^{k+1}(\mathbf{u}) q^k(\mathbf{x})$$

$$\times \log\left(\frac{q(\gamma, \alpha, \beta) q^{k+1}(\mathbf{u}) q^k(\mathbf{x})}{F(\gamma, \alpha, \beta, \mathbf{u}, \mathbf{x}, \mathbf{v}^{k+1}, \mathbf{y})}\right) d\gamma d\alpha d\beta d\mathbf{u} d\mathbf{x} \quad (25)$$

Set

$$q(\gamma, \alpha, \beta) = \lim_{k \to \infty} q^k(\gamma, \alpha, \beta), q(\mathbf{u}) = \lim_{k \to \infty} q^k(\mathbf{u}), q(\mathbf{x}) = \lim_{k \to \infty} q^k(\mathbf{x}).$$
$$(26)$$

Let us now explicitly state the forms of the distributions found at each step of the algorithm. We denote by $\mathrm{E}_{q(\cdot)}$ the expectation with respect to the distribution $q(\cdot)$. In order to find $q(\mathbf{x})$, we differentiate the integral on the right-hand side of Eq. (22) with respect to $q(\mathbf{x})$ and set it equal to zero to obtain

$$q^k(\mathbf{x}) \propto \exp \mathrm{E}_{q^k(\gamma, \alpha, \beta) q^k(\mathbf{u})}[\ln F(\gamma, \alpha, \beta, \mathbf{x}, \mathbf{u}, \mathbf{v}^k)]. \quad (27)$$

Therefore, $q^k(\mathbf{x})$ is an N-dimensional Gaussian distribution with parameters

$$\mathrm{cov}_{q^k(\mathbf{x})}[\mathbf{x}] = \left(\mathrm{E}_{q^k(\beta)}[\beta]\mathbf{H}^t\mathbf{H} + \mathrm{E}_{q^k(\alpha)}[\alpha]\mathbf{I}\right)^{-1}, \quad (28)$$

$$\mathrm{E}_{q^k(\mathbf{x})}[\mathbf{x}] = \mathrm{cov}_{q^k(\mathbf{x})}[\mathbf{x}] \left(\mathrm{E}_{q^k(\alpha)}[\alpha]\mathrm{E}_{q^k(\mathbf{u})}[\mathbf{u}] + \mathrm{E}_{q^k(\beta)}[\beta]\mathbf{H}^t\mathbf{y}\right). \quad (29)$$

Similarly, $q(\mathbf{u})$ is found by differentiating the integral on the right-hand side of Eq. (23) with respect to $q(\mathbf{u})$ and by setting it equal to zero, which results in

$$q^{k+1}(\mathbf{u}) \propto \exp \mathrm{E}_{q^k(\gamma, \alpha, \beta) q^k(\mathbf{x})}[\ln F(\gamma, \alpha, \beta, \mathbf{x}, \mathbf{u}, \mathbf{v}^k)]. \quad (30)$$

Therefore $q^{k+1}(\mathbf{u})$ is an N-dimensional Gaussian distribution with parameters

$$\mathrm{cov}_{q^{k+1}(\mathbf{u})}[\mathbf{u}] = \left(\mathrm{E}_{q^k(\alpha)}[\alpha]\mathbf{I} + \mathrm{E}_{q^k(\gamma)}[\gamma](\Delta^h)^t W(\mathbf{v}^k)(\Delta^h)\right.$$
$$\left. + \ \mathrm{E}_{q^k(\gamma)}[\gamma](\Delta^v)^t W(\mathbf{v}^k)(\Delta^v)\right)^{-1}, \quad (31)$$

$$\mathrm{E}_{q^{k+1}(\mathbf{u})}[\mathbf{u}] = \mathrm{cov}_{q^{k+1}(\mathbf{u})}[\mathbf{u}]\mathrm{E}_{q^k(\alpha)}[\alpha]\mathrm{E}_{q^k(\mathbf{x})}[\mathbf{x}], \quad (32)$$

where $W(\mathbf{v})$ is the $N \times N$ diagonal matrix of the form

$$W(\mathbf{v}) = \mathrm{diag}\left(\frac{1}{\sqrt{v_i^k}}\right), \ i = 1, \ldots, N \quad (33)$$

Similarly, we have from Eq. (24)

$$\mathbf{v}_i^{k+1} = \mathrm{E}_{q^{k+1}(\mathbf{u})}[(\Delta_i^h(\mathbf{u}))^2 + (\Delta_i^v(\mathbf{u}))^2], \ i = 1, \ldots, N. \quad (34)$$

where

$$\mathrm{E}_{q^{k+1}(\mathbf{u})}[(\Delta_i^h(\mathbf{u}))^2 + (\Delta_i^v(\mathbf{u}))^2]$$
$$= \ (\Delta_i^h(\mathrm{E}_{q^{k+1}(\mathbf{u})}[\mathbf{u}]))^2 + ((\Delta_i^v(\mathrm{E}_{q^{k+1}(\mathbf{u})}[\mathbf{u}]))^2$$
$$+ \ \mathrm{E}_{q^{k+1}(\mathbf{u})}[(\Delta_i^h(\mathbf{u} - \mathrm{E}_{q^{k+1}(\mathbf{u})}[\mathbf{u}]))^2]$$
$$+ \ \mathrm{E}_{q^{k+1}(\mathbf{u})}[(\Delta_i^v(\mathbf{u} - \mathrm{E}_{q^{k+1}(\mathbf{u})}[\mathbf{u}]))^2] \quad (35)$$

Finally we find $q^{k+1}(\gamma, \alpha, \beta)$ by differentiating the integral on the right hand side of Eq. (25) with respect to $q(\gamma, \alpha, \beta)$ and setting it equal to zero, which results in

$$q^{k+1}(\gamma, \alpha, \beta) = q^{k+1}(\gamma) q^{k+1}(\alpha) q^{k+1}(\beta) \quad (36)$$

where $q^{k+1}(\gamma)$, $q^{k+1}(\alpha)$, and $q^{k+1}(\beta)$ are Gamma distributions, given respectively by

$$q^{k+1}(\gamma) \propto \gamma^{N/2} \exp\left[-\gamma \sum_i \sqrt{v_i^{k+1}}\right], \quad (37)$$

$$q^{k+1}(\alpha) \propto \alpha^{N/2} \exp\left[-\alpha \frac{\mathrm{E}_{q^{k+1}(\mathbf{u}) q^k(\mathbf{x})} \| \mathbf{u} - \mathbf{x} \|^2}{2}\right], \quad (38)$$

$$q^{k+1}(\beta) \propto \beta^{N/2} \exp\left[-\beta \frac{\mathrm{E}_{q^k(\mathbf{x})} \| \mathbf{y} - \mathbf{Hx} \|^2}{2}\right]. \quad (39)$$

The means of these distributions are used as the estimates of the hyperparameters.

Finally, we note here that this proposed restoration algorithm represents an extension of the previous one (and so of the method proposed in [5]) which takes into account the uncertainty of the estimated image and parameters. This incorporation of uncertainty is performed through the utilization of the posterior covariances of the unknowns. It should be pointed out that the method in [8] can again be utilized to find the estimates given in Eqs. (29) and (32), and therefore Algorithm 2 is also computationally very efficient as Algorithm 1.

4. EXPERIMENTAL RESULTS

In this section we present experimental results obtained by the use of the proposed algorithms. Performance of our algorithms will be evaluated on two images (i.e. Lena and Cameraman) with two different linear space invariant degradations, Gaussian blur with variance 9 and the uniform blur of size 9x9, respectively. In all cases, white Gaussian noise was added to the blurred images to obtain degraded images with blurred signal to noise (BSNR) ratios of 30 and 40dB.

In order to compare different restorations we have used improvement in signal to noise Ratio (ISNR) as our comparison metric. The ISNR is defined as $10 \log_{10} \left(\|\mathbf{x} - \mathbf{y}\|^2 / \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right)$, where $\mathbf{x}$, $\mathbf{y}$ and $\hat{\mathbf{x}}$ are the original, observed and estimated images, respectively. Additionally, we report algorithm run times to evaluate computational complexity of the proposed methods. Algorithms 1 and 2 are denoted ALG1 and ALG2, respectively. BMK1 and BMK2 represent first and second methods from [4], respectively. Note that both BMK1 and BMK2 are image restoration methods also based on the TV priors, whereas the computations are performed using the conjugate gradient method. For all experiments, $\|\mathbf{x}^k - \mathbf{x}^{k-1}\| / \|\mathbf{x}^{k-1}\| < 2 \times 10^{-4}$; (or $E_{\mathbf{q}^k(\mathbf{x})}[\mathbf{x}^k]$ instead of $\mathbf{x}^k$) is used as a termination criteria. The comparison between the algorithms are presented in Table 1. It should be pointed out that the parameters of the proposed methods are estimated automatically as described in Section 3. However, based on [5] and empirical data, we found that local minima can be avoided by setting $E_{\mathbf{q}^k(\alpha)}[\alpha]$ equal to the $E_{\mathbf{q}^k(\gamma)}[\gamma]$, which is the method utilized in the experiments reported here.

It is clear from Table 1 that the proposed methods ALG1 and ALG2 provide restoration results competitive to BMK1 and BMK2 methods in terms of image quality, as indicated by the ISNR metric. In the Gaussian blur case, the proposed methods provided results similar to BMK1 and BMK2, whereas in the uniform blur case, they resulted in slightly decreased ISNR values. On the other hand, it is clear from Table 1 that the proposed methods are computationally much more efficient than BMK1 and BMK2, converging 6-20 times faster. The restoration results of the Cameraman image in the case of the uniform blur of size 9x9 with 40dB BSNR, as well as the degraded version, are shown in Figure 1. As can be seen from the Figure 1, ALG1 and ALG2 provide restored images with very high visual quality, and the difference as indicated by the ISNR values are not significant.

In summary, the proposed methods ALG1 and ALG2 provide ISNR results comparable with existing TV-based image restoration methods such as BMK1 and BMK2, whereas they are computationally much more efficient. Moreover, both methods show little or no difference in the restored images in terms of visual quality.

Table 1. Performance of the proposed algorithms

BSNR	Method	Cameraman (9x9 uniform)		Method	Lena (Gaussian variance 9)	
		ISNR[dB]	time[min]		ISNR[dB]	time[min]
40dB	BMK1	7.93	13.31	BMK1	4.55	4.59
	BMK2	8.08	14.64	BMK2	4.51	4.19
	ALG1	6.81	0.55	ALG1	4.50	0.47
	ALG2	6.58	0.79	ALG2	4.49	0.64
30dB	BMK1	3.55	4.66	BMK1	3.55	3.44
	BMK2	3.77	5.72	BMK2	3.54	3.78
	ALG1	3.23	0.58	ALG1	3.49	0.52
	ALG2	3.20	0.79	ALG2	3.48	0.70

5. CONCLUSIONS

In this paper we presented two Bayesian image restoration methods based on total variation image priors. The unknown image and all algorithm parameters are modeled using a hierarchical Bayesian framework and estimated simultaneously which results in fully automated methods. We utilized a two-stage image model where the TV image prior is imposed on a hidden image. The utilization of this image priors makes the use of computationally efficient methods possible. Experimental results demonstrate that the proposed

(a)   (b)

(c)   (d)
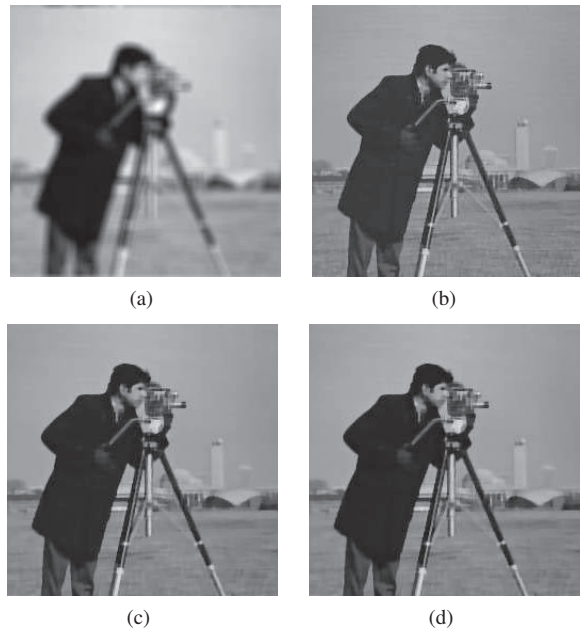
Fig. 1. (a) Image degraded by the uniform size 9x9 blur (BSNR=40dB), (b) Restored image using BMK2 (ISNR=8.08dB), (c) Restored image using ALG1 (ISNR=6.81dB), (d) Restored image using ALG2 (ISNR=6.58dB).

algorithms provide restoration results competitive in image quality and superior in computational efficiency.

6. REFERENCES

[1] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Processing Mag.*, vol. 14, no. 2, pp. 24–41, Mar 1997.

[2] N. P. Galatsanos, V. Z. Mesarovic, R. Molina, A. K. Katsaggelos, and J. Mateos, "Hyperparameter estimation in image restoration problems with partially-known blurs," *Optical Eng.*, vol. 41, no. 8, pp. 1845–1854, 2002.

[3] R. Molina, J. Mateos, and A.K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. on Img Proc.*, vol. 15, no. 12, pp. 3715–3727, Dec 2006.

[4] 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.

[5] Y. Huang, M. Ng, and Y. Wen, "A fast total variation minimization method for image restoration," *SIAM Journal on Multiscale Modeling and Simulation*, vol. 7, pp. 774–795, 2008.

[6] J. Bioucas-Dias, M. Figueiredo, and J. Oliveira, "Total-variation image deconvolution: A majorization-minimization approach," in *ICASSP'2006*, 2006.

[7] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.

[8] A. Chambolle and P. Lions, "Image recovery via total variation minimization and related problems," *Numer. Math.*, vol. 76, pp. 167–188, 1997.