

GENERALIZED GAUSSIAN MARKOV RANDOM FIELD IMAGE RESTORATION USING VARIATIONAL DISTRIBUTION APPROXIMATION

S. Derin Babacan¹, Rafael Molina², Aggelos K. Katsaggelos¹

¹Department of Electrical Engineering
and Computer Science
Northwestern University, Evanston, IL 60208, USA
sdb@northwestern.edu, aggk@eecs.northwestern.edu

²Departamento de Ciencias
de la Computación e I.A.
Universidad de Granada, 18071 Granada, Spain
rms@decsai.ugr.es

ABSTRACT

In this paper we propose novel algorithms for image restoration and parameter estimation with a Generalized Gaussian Markov Random Field (GGMRF) prior utilizing variational distribution approximation. The restored image and the unknown hyperparameters for both the image prior and the image degradation noise are simultaneously estimated within a hierarchical Bayesian framework. We develop two algorithms resulting from this formulation which provide approximations to the posterior distributions of the latent variables. Experimental results are provided to demonstrate the performance of the algorithms.

Index Terms— Image restoration, Generalized Gaussian Markov Random Fields, variational methods, parameter estimation, Bayesian methods.

1. INTRODUCTION

A standard formulation of the image degradation model is given in matrix-vector form by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (1)$$

where the $N \times 1$ vectors \mathbf{x} , \mathbf{y} , and \mathbf{n} represent respectively the original image, the available noisy and blurred image, and the noise, all ordered lexicographically. The noise is assumed Gaussian with independent elements of variance $\sigma_n^2 = \beta^{-1}$, and \mathbf{H} represents the known blurring matrix.

The image restoration problem is to find an estimate of \mathbf{x} given \mathbf{y} , \mathbf{H} , and knowledge about \mathbf{n} and possibly \mathbf{x} [1]. In a Bayesian formulation, knowledge about the unknown parameters are introduced in the estimation process by incorporating prior image and observation models. Examples of such prior models include Simultaneous Autoregression (SAR), Conditional Autoregression (CAR), or Total Variation (TV). In this paper we propose to use a GGMRF [2][3] as the image prior. In addition to the unknown image and noise, their prior models introduce parameters that are related to their variances. These parameters, referred to as *hyperparameters*, determine the performance of the restoration algorithm and therefore play an important role in Bayesian image restoration. In [2] and [3] point estimates for the unknown image are found and the hyperparameters are not estimated explicitly, but they are instead marginalized (evidence approach). In addition, in [3] a Poisson noise model is utilized.

Recently, there has been a growing interest in variational methods, where the posterior distribution is approximated with the use of the Kullback-Leibler cross-entropy [4]. Several methods tackle

the deconvolution problem using the variational approach (see, for example, [5][6][7][8]).

Using the variational framework we utilize a hierarchical Bayesian paradigm (see, for example, [7][9]) to jointly provide estimates of the posterior distributions of the restored image and the hyperparameters when a GGMRF prior is used. We develop two algorithms using our framework.

This paper is organized as follows. The hierarchical Bayesian model is presented in Sec. 2. Section 3 describes the variational approach to distribution approximation and the derivation of our algorithms. We present the experimental results in Sec. 4 and conclude in Sec. 5.

2. BAYESIAN MODELING

The Bayesian modeling of the GGMRF restoration problem requires first the definition of a joint distribution $p(\alpha, \beta, \mathbf{x}, \mathbf{y})$ of the observation, \mathbf{y} , the unknown image, \mathbf{x} , and the hyperparameters α (to be defined below) and β . We utilize the hierarchical Bayesian paradigm where in the first stage we form the prior distributions $p(\mathbf{y}|\mathbf{x}, \beta)$ and $p(\mathbf{x}|\alpha)$ for the unknowns, and in the second stage we define hyperpriors on the hyperparameters. The joint probability model is shown in graphical form in Fig. 1(a) using a directed acyclic graph.

2.1. First stage: prior models on image and observation

The probability distribution corresponding to the observation model in Eq. (1) is given by

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

As the image model we use the GGMRF prior, given by

$$p(\mathbf{x}|\alpha) \propto \frac{1}{Z_{\mathbf{GG}}(\alpha)} \exp [-\alpha \mathbf{G}\mathbf{G}(\mathbf{x})], \quad (3)$$

where $Z_{\mathbf{GG}}(\alpha)$ is the partition function and

$$\mathbf{G}\mathbf{G}(\mathbf{x}) = \sum_i \sum_{d=1}^4 \left[|\Delta_i^d(\mathbf{x})|^p \right],$$

where the first summation is over all pixels i , $p \in [1, 2]$, and $\Delta_i^d(\mathbf{x})$ denotes the first order difference in the d direction, such that

$$\Delta_i^d(\mathbf{x}) = x_i - x_{i+d}, \quad d = 1, \dots, 4$$

Figure 1(b) shows the directions $d = 1, \dots, 4$ along which the first order differences are taken.

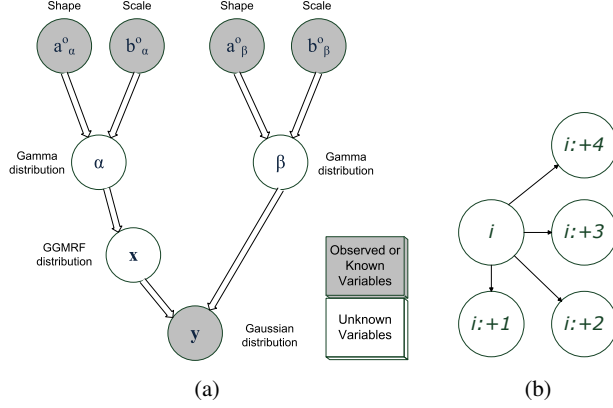


Fig. 1. (a) Graphical model showing relationships between variables, (b) the directions for the first order differences around the pixel i .

Using $u^p = v$ and taking into account that

$$\int_0^\infty \exp[-\alpha u^p] du = \frac{1}{p} \int_0^\infty \exp[-\alpha v] v^{\frac{1-p}{p}} dv \propto \alpha^{-\frac{1}{p}},$$

we use the approximation $\alpha^{-N/p}$ to the partition function to obtain

$$p(\mathbf{x}|\alpha) \propto \alpha^{N/p} \exp[-\alpha \mathbf{G}\mathbf{G}(\mathbf{x})]. \quad (4)$$

2.2. Second stage: hyperprior on the hyperparameters

We use Gamma distributions as our model for the hyperparameters $\omega \in \{\alpha, \beta\}$, given by

$$p(\omega) = \Gamma(\omega|a_\omega^0, b_\omega^0) = \frac{(b_\omega^0)^{a_\omega^0}}{\Gamma(a_\omega^0)} \omega^{a_\omega^0-1} \exp[-\omega b_\omega^0]. \quad (5)$$

Combining the first and second stage, the joint distribution can be written as

$$p(\alpha, \beta, \mathbf{x}, \mathbf{y}) = p(\alpha)p(\beta)p(\mathbf{x}|\alpha)p(\mathbf{y}|\mathbf{x}, \beta). \quad (6)$$

3. INFERENCE AND VARIATIONAL APPROXIMATION

The Bayesian inference on $(\alpha, \beta, \mathbf{x})$ should be based on

$$p(\alpha, \beta, \mathbf{x} | \mathbf{y}) = \frac{p(\alpha, \beta, \mathbf{x}, \mathbf{y})}{p(\mathbf{y})}. \quad (7)$$

However, since the posterior $p(\alpha, \beta, \mathbf{x} | \mathbf{y})$ cannot be found in closed form, we approximate it by a simpler parametric form $q(\alpha, \beta, \mathbf{x}) = q(\alpha, \beta)q(\mathbf{x})$. This distribution can be found in a variational framework by minimizing the Kullback-Leibner (KL) distance, that is,

$$\begin{aligned} C_{KL}(q(\alpha, \beta)q(\mathbf{x}) \| p(\alpha, \beta, \mathbf{x}|\mathbf{y})) \\ &= \int_{\alpha} \int_{\beta} \int_{\mathbf{x}} q(\alpha, \beta)q(\mathbf{x}) \log \left(\frac{q(\alpha, \beta)q(\mathbf{x})}{p(\alpha, \beta, \mathbf{x}|\mathbf{y})} \right) d\alpha d\beta d\mathbf{x} \\ &= \int_{\alpha} \int_{\beta} \int_{\mathbf{x}} q(\alpha, \beta)q(\mathbf{x}) \log \left(\frac{q(\alpha, \beta)q(\mathbf{x})}{p(\alpha, \beta, \mathbf{x}, \mathbf{y})} \right) d\alpha d\beta d\mathbf{x} + \text{const}, \end{aligned} \quad (8)$$

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

Due to the form of our image prior, the KL distance cannot be minimized directly. We define the functional $M(\alpha, \mathbf{x}, \mathbf{v})$ for α, x and $\mathbf{v} \in (R^4+)^N$, with components $(v_{i,1}, \dots, v_{i,4}), i = 1, \dots, N$

$$M(\alpha, \mathbf{x}, \mathbf{v}) = \alpha^{N/p} \exp \left[-\frac{\alpha p}{2} \sum_i \sum_{d=1}^4 \left[\frac{(\Delta_i^d(\mathbf{x}))^2 + \frac{2-p}{p} v_{i,d}}{v_{i,d}^{1-p/2}} \right] \right].$$

Next, using the following inequality for $w \geq 0, z > 0$, and $p \in [1, 2]$

$$w^{p/2} \leq z^{p/2} + \frac{p}{2z^{1-p/2}}(w - z) = \frac{p}{2} \frac{(w + \frac{2-p}{p} z)}{z^{1-p/2}}, \quad (9)$$

we find a lower bound for the image prior, given by

$$p(\mathbf{x}|\alpha) \geq c \cdot M(\alpha, \mathbf{x}, \mathbf{v}),$$

where c is a constant. This inequality can be used to find a lower bound for the joint probability distribution

$$\begin{aligned} p(\alpha, \beta, \mathbf{x}, \mathbf{y}) &\geq c \cdot p(\alpha)p(\beta)M(\alpha, \mathbf{x}, \mathbf{v})p(\mathbf{y}|\mathbf{x}, \beta) \\ &= F(\alpha, \beta, \mathbf{x}, \mathbf{v}, \mathbf{y}). \end{aligned} \quad (10)$$

Using these lower bounds in Eq. (8), we can find an upper bound for the KL distance as follows:

$$\begin{aligned} C_{KL}(q(\alpha, \beta)q(\mathbf{x}) \| p(\alpha, \beta, \mathbf{x}|\mathbf{y})) \\ \leq \min_{\mathbf{v}} \int_{\alpha} \int_{\beta} \int_{\mathbf{x}} q(\alpha, \beta)q(\mathbf{x}) \log \left(\frac{q(\alpha, \beta)q(\mathbf{x})}{F(\alpha, \beta, \mathbf{x}, \mathbf{v}, \mathbf{y})} \right) d\alpha d\beta d\mathbf{x}. \end{aligned} \quad (11)$$

Finally, we employ a minimization of the right-hand side of Eq. (11) and obtain the following iterative procedure to estimate the unknowns:

Algorithm 1 Posterior parameter and image distributions estimation by approximating $p(\alpha, \beta, \mathbf{x} | \mathbf{y})$ by $q(\alpha, \beta)q(\mathbf{x})$.

Given $\mathbf{v}^1 \in (R^4+)^N$ and $q^1(\alpha, \beta)$,

For $k = 1, 2, \dots$ until convergence:

1. Find

$$\begin{aligned} q^k(\mathbf{x}) &= \arg \min_{q(\mathbf{x})} \int_{\alpha} \int_{\beta} \int_{\mathbf{x}} q^k(\alpha, \beta)q(\mathbf{x}) \\ &\quad \times \log \left(\frac{q^k(\alpha, \beta)q(\mathbf{x})}{F(\alpha, \beta, \mathbf{x}, \mathbf{v}^k, \mathbf{y})} \right) d\alpha d\beta d\mathbf{x} \end{aligned} \quad (12)$$

2. Find

$$\begin{aligned} \mathbf{v}^{k+1} &= \arg \min_{\mathbf{v}} \int_{\alpha} \int_{\beta} \int_{\mathbf{x}} q^k(\alpha, \beta)q^k(\mathbf{x}) \\ &\quad \log \left(\frac{q^k(\alpha, \beta)q^k(\mathbf{x})}{F(\alpha, \beta, \mathbf{x}, \mathbf{v}, \mathbf{y})} \right) d\alpha d\beta d\mathbf{x} \end{aligned} \quad (13)$$

3. Find

$$\begin{aligned} q^{k+1}(\alpha, \beta) &= \arg \min_{q(\alpha, \beta)} \int_{\alpha} \int_{\beta} \int_{\mathbf{x}} q(\alpha, \beta)q^k(\mathbf{x}) \\ &\quad \log \left(\frac{q(\alpha, \beta)q^k(\mathbf{x})}{F(\alpha, \beta, \mathbf{x}, \mathbf{v}^{k+1}, \mathbf{y})} \right) d\alpha d\beta d\mathbf{x} \end{aligned} \quad (14)$$

Now we proceed to give the explicit solutions at each step of the algorithm. Note that in the first step we have

$$q^k(\mathbf{x}) \propto \exp \left\{ E_{q^k(\alpha, \beta)} [\ln F(\alpha, \beta, \mathbf{x}, \mathbf{v}^k)] \right\}, \quad (15)$$

which corresponds to a multivariate Gaussian distribution with the mean and the covariance given by

$$E_{q^k(\mathbf{x})}[\mathbf{x}] = \text{cov}_{q^k(\mathbf{x})}[\mathbf{x}] E_{q^k(\beta)}[\beta] \mathbf{H}^t \mathbf{y}, \quad (16)$$

$$\begin{aligned} \text{cov}_{q^k(\mathbf{x})}[\mathbf{x}] &= [E_{q^k(\beta)}[\beta] \mathbf{H}^t \mathbf{H} \\ &+ p E_{q^k(\alpha)}[\alpha] \sum_{d=1}^4 (\Delta^d)^t W_d(\mathbf{v}^k) (\Delta^d)]^{-1} = [\mathbf{C}^k(\mathbf{v}^k)]^{-1}, \end{aligned} \quad (17)$$

where

$$W_d(\mathbf{v}^k) = \text{diag} \left(\frac{1}{v_{i,d}^{1-p/2}} \right), \quad d = 1, \dots, 4, \quad i = 1, \dots, N.$$

In the second step, we have

$$\mathbf{v}_d^{k+1} = \arg \min_{\mathbf{v}_d} \sum_i \frac{E_{q^k(\mathbf{x})}[(\Delta_i^d(\mathbf{x}))^2] + \frac{2-p}{p} v_{i,d}}{v_{i,d}^{1-p/2}} \quad d = 1, \dots, 4$$

and therefore

$$\mathbf{v}_{i,d}^{k+1} = E_{q^k(\mathbf{x})}[(\Delta_i^d(\mathbf{x}))^2], \quad i = 1, \dots, N \quad d = 1, \dots, 4 \quad (18)$$

where

$$\begin{aligned} E_{q^k(\mathbf{x})}[(\Delta_i^d(\mathbf{x}))^2] &= (\Delta_i^d(E_{q^k(\mathbf{x})}[\mathbf{x}]))^2 \\ &+ \frac{1}{N} \text{trace} [\text{cov}_{q^k(\mathbf{x})}[\mathbf{x}] \times ((\Delta^d)^t (\Delta^d))]. \end{aligned}$$

Finally to find $q^{k+1}(\alpha, \beta)$ we differentiate the integral on the right hand side of Eq. (14) with respect to $q(\alpha, \beta)$ and set it equal to zero to obtain

$$q^{k+1}(\alpha, \beta) \propto \exp \left\{ E_{q^k(\mathbf{x})} [\ln F(\alpha, \beta, \mathbf{x}, \mathbf{v}^{k+1})] \right\}.$$

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

$$\begin{aligned} q^{k+1}(\alpha) &\propto \alpha^{N/p + a_\alpha^o - 1} \exp \left[-\alpha \left(\sum_i \sum_{d=1}^4 ([v_{i,d}^{k+1}]^{p/2}) + b_\alpha^o \right) \right], \\ q^{k+1}(\beta) &\propto \beta^{N/2 + a_\beta^o - 1} \exp \left[-\beta \left(\frac{E_{q^k(\mathbf{x})} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2}{2} + b_\beta^o \right) \right]. \end{aligned}$$

As the estimates to these hyperparameters, we use the means of these distributions, which can be given as

$$(E_{q^{k+1}(\alpha)}[\alpha])^{-1} = \gamma_\alpha \frac{1}{\bar{\alpha}^o} + (1 - \gamma_\alpha) \frac{p \sum_{d=1}^4 \sum_i [v_{i,d}^{k+1}]^{p/2}}{N}, \quad (19)$$

$$(E_{q^{k+1}(\beta)}[\beta])^{-1} = \gamma_\beta \frac{1}{\bar{\beta}^o} + (1 - \gamma_\beta) \frac{E_{q^k(\mathbf{x})} [\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2]}{N}, \quad (20)$$

where $\bar{\alpha}^o = a_\alpha^o / b_\alpha^o$, $\bar{\beta}^o = a_\beta^o / b_\beta^o$, $\gamma_\alpha = \frac{a_\alpha^o}{a_\alpha^o + \frac{N}{p}}$, and $\gamma_\beta = \frac{a_\beta^o}{a_\beta^o + \frac{N}{2}}$. The parameters γ_α and γ_β , both taking values in the interval $[0, 1]$, can be understood as normalized confidence parameters. According

to Eqs. (19) and (20), when they are equal to zero, no confidence is placed on the inverse of the mean of the corresponding hyperprior, while when they are asymptotically equal to one, the prior knowledge of the mean is fully enforced, i.e., no estimation of the hyperparameters is performed.

The only remaining task is the calculation of $E_{q^k(\mathbf{x})} [\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2]$ which can be given as

$$\begin{aligned} E_{q^k(\mathbf{x})} [\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2] &= \|\mathbf{y} - \mathbf{H} E_{q^k(\mathbf{x})}[\mathbf{x}]\|^2 \\ &+ \text{trace} (\text{cov}_{q^k(\mathbf{x})}[\mathbf{x}] \mathbf{H}^t \mathbf{H}). \end{aligned}$$

The estimate $q^k(\mathbf{x})$ in Algorithm 1 is the best approximation to the posterior in the KL divergence sense. However, we can also consider a suboptimal case where we assume a degenerate distribution for $q(\mathbf{x})$, that is, $q(\mathbf{x})$ takes one value, \mathbf{x}^k , with probability one and the rest of the values with probability zero. This approach leads to an alternative algorithm, referred to as Algorithm 2, where the expectations involving the parameter $q^k(\mathbf{x})$ are removed. Thus, the covariances in Eqs. (18), (19), and (20) are set equal to zero.

As the estimate to the unknown image \mathbf{x} , we use the mean of $q^k(\mathbf{x})$ shown in Eq. (16) in both algorithms, which requires the inversion of a very large matrix $\mathbf{C}^k(\mathbf{v}^k)$. This, however, introduces a big computational challenge since the last terms in Eq. (17) cannot be represented as block-circulant matrices with circulant blocks (BCCB), and therefore the inverse cannot be computed in the Fourier domain. We therefore employ a gradient descent approach to compute the image estimates without explicitly calculating the image covariance.

Note, however, that the explicit form of $\text{cov}_{q^k(\mathbf{x})}[\mathbf{x}]$ is needed in Eqs. (19)-(20) in Algorithm 1. To overcome this computational difficulty, we use the following approximation

$$\begin{aligned} \text{cov}_{q^k(\mathbf{x})}[\mathbf{x}] &\approx [E_{q^k(\beta)}[\beta] \mathbf{H}^t \mathbf{H} \\ &+ p E_{q^k(\alpha)}[\alpha] \sum_{d=1}^4 z_d(\mathbf{v}^k) (\Delta^d)^t (\Delta^d)]^{-1} = \mathbf{B}^{-1}, \end{aligned}$$

where $W_d(\mathbf{v}^k) \approx z_d(\mathbf{v}^k) \mathbf{I}$ and $z_d(\mathbf{v}^k) = \frac{1}{N} \sum_i \frac{1}{[v_{i,d}^k]^{1-p/2}}$. Note that in this approximation, matrix \mathbf{B} is BCCB, and therefore its inversion can be carried out very efficiently in the Fourier domain.

4. EXPERIMENTAL RESULTS

We performed a number of experiments with the proposed algorithms using several images and several types of blurring functions. The results of some of them are presented here. Since we developed two different algorithms resulting from our framework, we will present results for both of them.

For the experiments presented here, the ‘‘Lena’’ image (shown in Fig. 2(a)) is blurred with a Gaussian shaped blur with variance 9 and a 9x9 uniform blur. Gaussian noise is added to the blurred images to obtain degraded images with blurred-signal-to-noise (BSNR) ratios of 20 and 40dB. An example degraded image is shown in Fig. 2(b) where the blur is Gaussian-shaped with variance 9 and BSNR = 40dB.

The parameters of both algorithms are initialized as follows: The observed image is used as initial estimate for the unknown image \mathbf{x} . The initial values of the hyperparameters and \mathbf{v} are determined using this initial \mathbf{x} in Eqs. (18)-(20). Note that all parameters of the algorithms are initialized using the observation \mathbf{y} so that no manual input is needed, i.e., both algorithms are initialized and run automatically. For all experiments, the criterion



Fig. 2. (a) Original Lena Image, (b) Image degraded by a Gaussian shaped PSF with variance 9 and Gaussian noise of variance 0.16 (BSNR=40dB), (c) Restored image using Algorithm 1 with $p = 1.8$ (ISNR = 4.15dB), (d) Restored image using Algorithm 2 with $p = 1.6$ (ISNR = 3.78dB).

$\| \mathbf{x}^k - \mathbf{x}^{k-1} \|^2 / \| \mathbf{x}^{k-1} \|^2 < 10^{-4}$ is used to terminate the iterative procedure, where \mathbf{x}^k is the mean of $q^k(\mathbf{x})$ in Algorithm 1 and the point estimate in Algorithm 2.

The restoration results of the Lena image in the case of Gaussian blur with 40dB BSNR are shown in Fig. 2(c) for Algorithm 1 and 2(d) for Algorithm 2. Note that Algorithm 1 is more successful at removing the blur whereas the restored image has less pronounced ringing artifacts in Algorithm 2. In both cases the restoration quality is good considering that the parameters of both algorithms are estimated using only the degraded observation without any prior knowledge about the noise. Also, in all cases the estimated value of β^{-1} was very close to the original noise variance.

Figure (3) shows the ISNR evolution in the case of Gaussian and uniform blurs with Algorithm 1 and BSNR = 40dB and 20dB with varying p -values, where ISNR is defined as $10 \log_{10}(\| \mathbf{x} - \mathbf{y} \|^2 / \| \mathbf{x} - \hat{\mathbf{x}} \|^2)$, where $\hat{\mathbf{x}}$ is the estimated image. We experimented with two cases, where in the first case we initialize and estimate the hyperparameters from the observation, and in the second case we compute them using the original unknown image and noise. As can be seen from Fig. (3), the highest ISNR values are achieved with different p -values for different noise levels and blur functions, and the ISNR values are comparable for both cases. It can also be seen that with fixed parameters the performance of the algorithms as a function of p is in agreement with the results reported in [2] and [3].

5. CONCLUSIONS

A novel GGMRF based image restoration methodology has been proposed to simultaneously estimate the reconstructed image and the hyperparameters of the Bayesian formulation. We have adopted a variational approach to approximate the posterior distributions of the unknown parameters to estimate the posterior distributions of unknowns so that the uncertainty of the estimates can be evaluated and different values from these distributions can be used in the restoration process. Two algorithms are provided resulting from this ap-

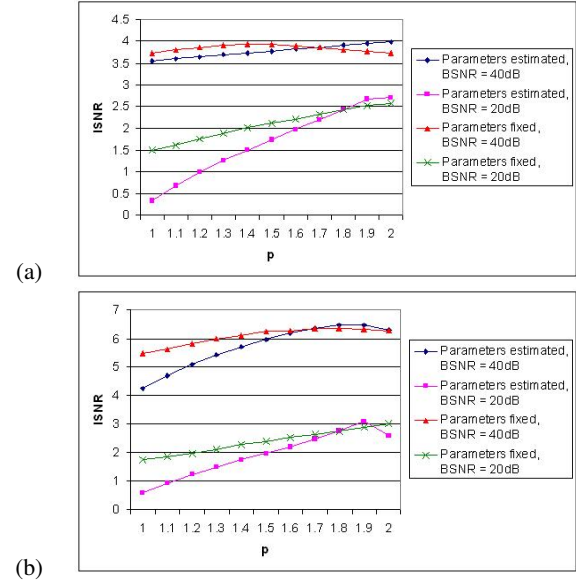


Fig. 3. ISNR values obtained by different p values with Lena image degraded by (a) a Gaussian blur with variance 9 and (b) a 9×9 uniform blur with Gaussian noise (BSNR = 40dB and 20dB).

proach. We have shown that the unknown parameters of the Bayesian formulation can be calculated automatically using only the observation or initial knowledge can be incorporated with different confidence values. Experimental results demonstrated the performance of the proposed algorithms.

6. REFERENCES

- [1] A.K. Katsaggelos, Ed., *Digital Image Restoration*, Springer Series in Information Sciences, vol. 23, Springer-Verlag, 1991.
- [2] C. A. Bouman and K. Sauer, "Generalized Gaussian image model for edge-preserving MAP estimation," *IEEE Trans. Image Processing*, vol. 2, no. 3, pp. 296–310, 1993.
- [3] A. López, R. Molina, A. K. Katsaggelos, A. Rodriguez, J. M. López, and J. M. Llamas, "Parameter estimation in bayesian reconstruction of spect images: An aide in nuclear medicine diagnosis," *International Journal of Imaging Systems and Technology*, vol. 14, no. 1, pp. 21–27, June 2004.
- [4] S. Kullback, *Information Theory and Statistics*, New York, Dover Publications, 1959.
- [5] K. Z. Adami, "Variational methods in Bayesian deconvolution," *PHYSTAT2003 ECONF*, vol. C030908, pp. TUGT002, 2003.
- [6] A. C. Likas and N. P. Galatsanos, "A variational approach for Bayesian blind image deconvolution," *IEEE Trans. Signal Processing*, vol. 52, no. 8, pp. 2222–2233, 2004.
- [7] R. Molina, J. Mateos, and A. K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. Image Processing*, vol. 15, no. 12, pp. 3715–3727, December 2006.
- [8] S. D. Babacan, R. Molina, and A.K. Katsaggelos, "Total variation image restoration and parameter estimation using variational posterior distribution approximation," in *ICIP 2007*, San Antonio, USA, Sept. 2007.
- [9] 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, Aug. 2002.