

FROM GLOBAL TO LOCAL BAYESIAN PARAMETER ESTIMATION IN IMAGE RESTORATION USING VARIATIONAL DISTRIBUTION APPROXIMATIONS

Rafael Molina^a, Miguel Vega^b, and Aggelos K. Katsaggelos^c

- a) Dept. Ciencias de la Computación e I. A., Univ. de Granada, 18071 Granada, Spain.
b) Dept. de Lenguajes y Sistemas Informáticos, Univ. de Granada, 18071 Granada, Spain.
c) Dept. of Electrical Engineering and Computer Science,
Northwestern University, Evanston, Illinois 60208-3118.
rms@decsai.ugr.es, mvega@ugr.es, aggk@eecs.northwestern.edu

ABSTRACT

In this paper we present a new Bayesian methodology for the restoration of blurred and noisy images. Bayesian methods rely on image priors that encapsulate prior image knowledge and avoid the ill-posedness of the image restoration problems. Some of these priors depend on global variance parameters, unable to account for local characteristics. Here we first use variational methods to approximate probability posterior distributions for the global model to later use those distributions to define local and more realistic image models which lead to better restored images as it is shown in the experimental section.

Index Terms— Image restoration, regularization, Bayesian models, parameter estimation, variational methods.

1. INTRODUCTION

A standard formulation of the image degradation model is given in lexicographic form by [1]

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

where the $P \times 1$ vectors \mathbf{x} , \mathbf{y} , and \mathbf{n} represent, respectively, the original image, the available noisy and blurred image, and the white Gaussian noise with independent elements of variance $\sigma_n^2 = \beta^{-1}$, and \mathbf{H} represents the known blurring matrix. The images are assumed to be of size $m \times n$, with $P = m \times n$. The restoration problem considered here is to find an estimate of \mathbf{x} from \mathbf{y} and \mathbf{H} and some knowledge about \mathbf{n} and possibly \mathbf{x} .

Bayesian image restoration techniques are based on the introduction of a prior image model on \mathbf{x} whose aim is to encapsulate our prior image knowledge and consequently to avoid the ill-posedness of the image restoration problem. Image prior models like the ones based on Simultaneous Auto-Regression (SAR) [2], Conditional Auto-Regression (CAR) [3] or Total Variation (TV) [4] depend on a global parameter which is related to the global variance of the image model. However, for most images it is not realistic to assume that the variance of the model is the same for the whole image and consequently it should be adapted to its local characteristics. In this paper we propose a two-step approach towards the construction of local image models. In the first step a global image model is used along with variational methods to approximate the probability posterior distribution of the image and global parameters. In the second

step these distributions are used to define local image models which lead to improved quality restored images.

The paper is organized as follows. In section 2 we discuss the global model and inference we use as a first step in the construction of the local modelling and inference. Then the construction and analysis of the local model is performed in section 3. Experimental results are described in section 4. Finally, section 5 concludes the paper.

2. GLOBAL BAYESIAN MODELING AND INFERENCE

The Bayesian formulation of our global image restoration problem requires the definition of the joint distribution $p(\Omega_g, \mathbf{x}, \mathbf{y})$ of the observation \mathbf{y} , the unknown image \mathbf{x} , and the hyperparameters Ω_g . Then, the posterior distribution of the unknowns given the observed image $p(\Omega_g, \mathbf{x}|\mathbf{y})$ has to be calculated and used to estimate the unknown image and hyperparameters.

Given the degradation model of Eq. (1), the distribution of the observed image \mathbf{y} given \mathbf{x} and a set of parameters Ω_g is

$$p(\mathbf{y}|\mathbf{x}, \beta) \propto \beta^{\frac{P}{2}} \exp \left[-\frac{1}{2}\beta \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \right]. \quad (2)$$

In our global image model we assume that the difference between the gray level at location i and the mean of the gray levels at the four closest neighbors of pixel i is a Gaussian random variable with zero mean and variance α^{-1} . We also assume that all the differences are independent. Then we have that

$$p(\mathbf{x}|\alpha) \propto \alpha^{\frac{P}{2}} \exp \left[-\frac{1}{2}\alpha \|\mathbf{C}\mathbf{x}\|^2 \right], \quad (3)$$

where \mathbf{C} denotes the Laplacian operator.

Finally, a flat prior $p(\alpha, \beta)$ is assumed on the unknown parameters α and β .

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

$$p(\alpha, \beta, \mathbf{x}|\mathbf{y}) = \frac{p(\alpha, \beta, \mathbf{y}, \mathbf{x})}{p(\mathbf{y})}. \quad (4)$$

It is difficult to evaluate this distribution and therefore we apply variational methods to approximate it by the distribution $q(\alpha, \beta, \mathbf{x})$.

This work has been partially supported by the Greece-Spain Integrated Action HG2004-0014, and by the "Instituto de Salud Carlos III" project PI 040857.

The variational criterion used to find $q(\alpha, \beta, \mathbf{x})$ is the minimization of the Kullback-Leibler divergence, given by [5, 6]

$$\begin{aligned} C_{KL}(q(\alpha, \beta, \mathbf{x}) \parallel p(\alpha, \beta, \mathbf{x}|\mathbf{y})) \\ &= \int_{\alpha, \beta, \mathbf{x}} q(\alpha, \beta, \mathbf{x}) \log \left(\frac{q(\alpha, \beta, \mathbf{x})}{p(\alpha, \beta, \mathbf{x}|\mathbf{y})} \right) d\alpha d\beta d\mathbf{x} \\ &= \int_{\alpha, \beta, \mathbf{x}} q(\alpha, \beta, \mathbf{x}) \log \left(\frac{q(\alpha, \beta, \mathbf{x})}{p(\alpha, \beta, \mathbf{x}, \mathbf{y})} \right) d\alpha d\beta d\mathbf{x} + \text{const}, \end{aligned} \quad (5)$$

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

We choose to approximate $p(\alpha, \beta, \mathbf{x}|\mathbf{y})$ by the distribution

$$q(\alpha, \beta, \mathbf{x}) = q(\alpha, \beta)q(\mathbf{x}), \quad (6)$$

where $q(\mathbf{x})$ and $q(\alpha, \beta)$ denote distributions on \mathbf{x} and α, β respectively. We propose the following iterative procedure to find $q(\alpha, \beta, \mathbf{x})$.

Algorithm 1

Given $q^1(\alpha, \beta)$, an initial estimate of the distribution $q(\alpha, \beta)$,
for $k = 1, 2, \dots$ until a stopping criterion is met:

1. Find

$$q^k(\mathbf{x}) = \arg \min_{q(\mathbf{x})} C_{KL}(q^k(\alpha, \beta)q(\mathbf{x}) \parallel p(\alpha, \beta, \mathbf{x}|\mathbf{y})) \quad (7)$$

2. Find

$$q^{k+1}(\alpha, \beta) = \arg \min_{q(\alpha, \beta)} C_{KL}(q(\alpha, \beta)q^k(\mathbf{x}) \parallel p(\alpha, \beta, \mathbf{x}|\mathbf{y})). \quad (8)$$

As stopping criterion of the above iterations, the convergence of the parameters defining the distributions $q^k(\mathbf{x})$ and $q^{k+1}(\alpha, \beta)$ can be used.

Differentiating the function to be minimized on the right hand side of Eq. (7) with respect to $q(\mathbf{x})$ and setting it equal to zero we obtain

$$q^k(\mathbf{x}) \propto \exp \mathbf{E}[\log p(\alpha, \beta, \mathbf{x}, \mathbf{y})]_{q^k(\alpha, \beta)}, \quad (9)$$

that is, $q^k(\mathbf{x})$ is a normal distribution

$$q^k(\mathbf{x}) = N(\mathbf{x} \mid \mathbf{E}^k[\mathbf{x}], \mathbf{cov}^k[\mathbf{x}]). \quad (10)$$

with

$$\mathbf{E}^k[\mathbf{x}] = \mathbf{cov}^k[\mathbf{x}]\beta^k \mathbf{H}^t \mathbf{y}, \quad (11)$$

and

$$(\mathbf{cov}^k[\mathbf{x}])^{-1} = \alpha^k \mathbf{C}^t \mathbf{C} + \beta^k \mathbf{H}^t \mathbf{H}. \quad (12)$$

where

$$\mathbf{E}[\alpha]_{q^k(\alpha, \beta)} = \alpha^k, \quad (13)$$

$$\mathbf{E}[\beta]_{q^k(\alpha, \beta)} = \beta^k. \quad (14)$$

Differentiating the function to be minimized on the right hand side of Eq. (8) with respect to $q(\alpha, \beta)$ and setting it equal to zero we obtain

$$q^{k+1}(\alpha, \beta) = q^{k+1}(\alpha)q^{k+1}(\beta) \quad (15)$$

with

$$q^{k+1}(\alpha) = \Gamma \left(\alpha \mid 1 + \frac{P}{2}, \frac{1}{2} \mathbf{E}[\|\mathbf{C}\mathbf{x}\|^2]_{q^k(\mathbf{x})} \right), \quad (16)$$

$$q^{k+1}(\beta) = \Gamma \left(\beta \mid 1 + \frac{P}{2}, \frac{1}{2} \mathbf{E}[\|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2]_{q^k(\mathbf{x})} \right), \quad (17)$$

where for a positive quantity the Gamma distribution is defined by

$$\Gamma(\omega \mid a_\omega^o, b_\omega^o) \propto \omega^{a_\omega^o - 1} \exp[-b_\omega^o \omega], \quad (18)$$

where the $\omega > 0$ denotes a hyperparameter, and $b_\omega^o > 0$ and $a_\omega^o > 0$ are parameters of the distribution. This Gamma distribution has the following mean, mode (which only exists when $a_\omega^o > 1$), and variance

$$\mathbf{E}[\omega] = \frac{a_\omega^o}{b_\omega^o}, \quad \text{Mode}[\omega] = \frac{a_\omega^o - 1}{b_\omega^o}, \quad \text{var}[\omega] = \frac{a_\omega^o}{(b_\omega^o)^2}. \quad (19)$$

If we now want to obtain point estimates of the image and the parameters we can use

$$\hat{\alpha} = \lim_{k \rightarrow \infty} \mathbf{E}^k[\alpha], \quad \hat{\beta} = \lim_{k \rightarrow \infty} \mathbf{E}^k[\beta], \quad \hat{\mathbf{x}} = \lim_{k \rightarrow \infty} \mathbf{E}^k[\mathbf{x}]. \quad (20)$$

Note that in this case the derived algorithm is equivalent to selecting $\hat{\alpha}$ and $\hat{\beta}$ as the mean of $p(\alpha, \beta \mid \mathbf{y})$ and then selecting *maximum a posteriori* image estimate for those parameters. Note also that the mode of $p(\alpha, \beta \mid \mathbf{y})$ can instead be used. However, the mean and mode of $p(\alpha, \beta \mid \mathbf{y})$ almost coincide for large P .

3. LOCAL BAYESIAN MODELING AND INFERENCE

Once we have obtained a global model for restoring the image and estimating the parameters we refine it with local image information. To model the variation in the original image locally we assume that ϵ_i , the difference between the gray level at location i and the mean of the gray levels at its four closest neighbors, which can be expressed using the Laplacian operator \mathbf{C} as

$$\epsilon_i = (\mathbf{C}\mathbf{x})_i, \text{ for } i = 1, 2, \dots, P, \quad (21)$$

is a Gaussian random variable with zero mean and variance α_i^{-1} [7]. We also assume that the ϵ_i 's, $i = 1, \dots, P$ are independent. Then if

$$\bar{\alpha} = (\alpha_1, \dots, \alpha_P) \quad (22)$$

we have

$$p_I(\mathbf{x}|\bar{\alpha}) \propto \left(\prod_{i=1}^P \alpha_i^{\frac{1}{2}} \right) \exp \left[-\frac{1}{2} \sum_{i=1}^P \alpha_i \epsilon_i^2 \right]. \quad (23)$$

Let us now consider the observation model. We already have an estimate $\hat{\beta}$ of β which has been provided by the global model and thus we use the observation model

$$p_I(\mathbf{y}|\mathbf{x}) \propto \hat{\beta}^{\frac{P}{2}} \exp \left[-\frac{1}{2} \hat{\beta} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \right], \quad (24)$$

where $\hat{\beta}$ is treated from now on as a known parameter (we note here that we could also provide local estimates for β by following similar steps to the ones followed next for α).

We now want to use

$$q(\alpha) = \lim_{k \rightarrow \infty} q^k(\alpha) \quad (25)$$

from Eq. (16) to define a prior model for the α_i 's, $i = 1, \dots, P$. To do so we first define

$$M = \lim_{k \rightarrow \infty} \mathbf{E}[\|\mathbf{C}\mathbf{x}\|^2]_{q^k(\mathbf{x})} \quad (26)$$

and observe that

$$\mathbf{E}[\alpha]_{q(\alpha)} = \frac{P/2 + 1}{M/2} \approx \frac{P/2}{M/2} \quad (27)$$

$$\mathbf{var}[\alpha]_{q(\alpha)} = \frac{P/2 + 1}{[M/2]^2} \approx \frac{P/2}{[M/2]^2}. \quad (28)$$

Let us now assume that

$$\alpha = \frac{1}{P} \sum_i \alpha_i \quad (29)$$

and that the α_i 's, $i = 1, \dots, P$ are independent and identically distributed. Then we have that

$$\mathbf{E}[\alpha_i] \approx \frac{P/2}{M/2} \text{ and } \mathbf{var}[\alpha_i] \approx \frac{P^2/2}{[M/2]^2}. \quad (30)$$

Now fitting a Gamma distribution to each α_i with the mean and variance above we can use as prior model for the vector $\bar{\alpha}$

$$p_l(\bar{\alpha}) = p_l(\alpha_1, \dots, \alpha_P) = \prod_i p_l(\alpha_i) \quad (31)$$

where

$$p_l(\alpha_i) \propto \alpha_i^{1/2-1} \exp\left[-\alpha_i \frac{M}{2P}\right]. \quad (32)$$

Finally we have the global distribution given by

$$p_l(\bar{\alpha}, \mathbf{x}, \mathbf{y}) = p_l(\bar{\alpha}) p_l(\mathbf{x}|\bar{\alpha}) p_l(\mathbf{y}|\mathbf{x}) \quad (33)$$

We now approximate $p_l(\bar{\alpha}, \mathbf{x}|\mathbf{y})$ by the distribution

$$q_l(\bar{\alpha}, \mathbf{x}) = q_l(\bar{\alpha}) q_l(\mathbf{x}), \quad (34)$$

where $q_l(\mathbf{x})$ and $q_l(\bar{\alpha})$ denote respectively distributions on \mathbf{x} and $\bar{\alpha}$, and propose the following iterative procedure to find $q_l(\bar{\alpha}, \mathbf{x})$.

Algorithm 2

Given $q_l^1(\bar{\alpha})$, an initial estimate of the distribution $q_l(\bar{\alpha})$,
for $k = 1, 2, \dots$ until a stopping criterion is met:

1. Find

$$q_l^k(\mathbf{x}) = \arg \min_{q_l(\mathbf{x})} C_{KL}(q_l^k(\bar{\alpha}) q_l(\mathbf{x}) \parallel p_l(\bar{\alpha}, \mathbf{x}|\mathbf{y})), \quad (35)$$

2. Find

$$q_l^{k+1}(\bar{\alpha}) = \arg \min_{q_l(\bar{\alpha})} C_{KL}(q_l(\bar{\alpha}), q_l^k(\mathbf{x}) \parallel p_l(\bar{\alpha}, \mathbf{x}|\mathbf{y})). \quad (36)$$

As stopping criterion of the above iterations, the convergence of the parameters defining the distributions $q_l^k(\mathbf{x})$ and $q_l^{k+1}(\bar{\alpha})$ can be used.

To explicitly calculate the distributions $q_l^k(\mathbf{x})$ and $q_l^{k+1}(\bar{\alpha})$ we use the same procedure as the one described for algorithm 1 to find $q^k(\mathbf{x})$ and $q^{k+1}(\alpha, \beta)$. Let

$$\mathbf{E}[\alpha_i]_{q_l^k(\bar{\alpha})} = \alpha_i^k, \text{ for } i = 1, 2, \dots, P \quad (37)$$

and

$$\bar{\alpha}^k = (\alpha_1^k, \alpha_2^k, \dots, \alpha_P^k). \quad (38)$$

Then

$$q_l^k(\mathbf{x}) = N(\mathbf{x} \mid \mathbf{E}_l^k[\mathbf{x}], \mathbf{cov}_l^k[\mathbf{x}]). \quad (39)$$

where

$$\mathbf{E}_l^k[\mathbf{x}] = \mathbf{cov}_l^k[\mathbf{x}] \hat{\beta} \mathbf{H}^t \mathbf{y}, \quad (40)$$

and

$$(\mathbf{cov}_l^k[\mathbf{x}])^{-1} = \mathbf{C}^t \text{diag}(\bar{\alpha}^k) \mathbf{C} + \hat{\beta} \mathbf{H}^t \mathbf{H}. \quad (41)$$

Furthermore

$$q_l^{k+1}(\bar{\alpha}) = \left[\prod_{i=1}^P q_l^{k+1}(\alpha_i) \right], \quad (42)$$

with

$$q_l^{k+1}(\alpha_i) = \Gamma\left(\alpha_i \mid 1, \frac{1}{2} \left(\frac{M}{P} + \mathbf{E}[\epsilon_i^2]_{q_l^k(\mathbf{x})} \right)\right) \quad (43)$$

where

$$\mathbf{E}[\epsilon_i^2]_{q_l^k(\mathbf{x})} = \mathbf{E}_l^k[\mathbf{x}]^t \mathbf{C}^t \mathbf{u}(i) \mathbf{C} \mathbf{E}_l^k[\mathbf{x}] + \text{tr}[\mathbf{C}^t \mathbf{u}(i) \mathbf{C} \mathbf{cov}_l^k[\mathbf{x}]], \quad (44)$$

and $\mathbf{u}(i)$ is a matrix with only one element distinct from zero, $\mathbf{u}(i)_{ii} = 1$.

Note that utilizing the mean of the Gamma distribution we have

$$\frac{1}{\alpha_i^k} = \frac{1}{2} \frac{M}{P} + \frac{1}{2} \mathbf{E}[\epsilon_i^2]_{q_l^k(\mathbf{x})}, \quad (45)$$

that is, the local image variance at pixel i is the average of the global variance provided by the global restoration model and the local variance provided by the restoration.

If we now want to obtain point estimates of the image and the parameters we can use

$$\hat{\alpha} = \lim_{k \rightarrow \infty} \mathbf{E}_l^k[\bar{\alpha}], \quad \hat{\mathbf{x}} = \lim_{k \rightarrow \infty} \mathbf{E}_l^k[\mathbf{x}]. \quad (46)$$

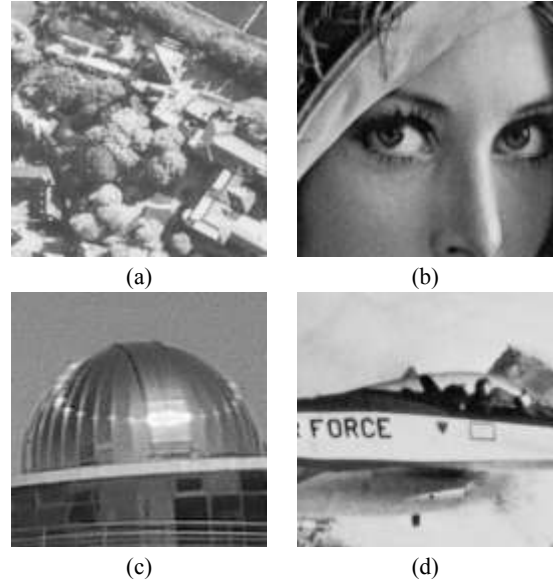


Fig. 1. Image set used in the experiments.

4. EXPERIMENTAL RESULTS

In this section we will compare the global model (Algorithm I), the local model (Algorithm II) and the method consisting on replacing $\mathbf{E}[\epsilon_i^2]_{q_l^k(\mathbf{x})}$ in Eq. (43) by the mean of local variance of pixel i and the local variances of its four closest neighbors (Algorithm III).

The Matlab implementation of the Preconditioned Conjugate Gradients method with preconditioning by an incomplete Cholesky factorization has been used to calculate the mean value of $q^k(\mathbf{x})$. To

blur	img	Method I			Method II			Method III		
		10 dB	20dB	30dB	10 dB	20dB	30dB	10 dB	20dB	30dB
G	a	8.64	8.42	4.56	8.73	8.49	4.19	8.73	8.48	4.63
	b	9.34	10.62	6.90	9.35	10.67	6.71	9.35	10.67	7.03
	c	10.54	10.44	5.29	10.52	10.46	5.28	10.52	10.46	5.49
	d	6.86	8.53	6.84	6.96	8.62	6.63	6.96	8.62	7.05
U	a	8.54	6.85	2.98	8.60	6.86	2.99	8.60	6.85	2.99
	b	9.55	10.02	5.01	9.57	10.02	5.02	9.57	10.02	5.02
	c	10.64	9.45	3.86	10.60	9.45	3.87	10.60	9.45	3.87
	d	6.86	8.33	4.79	6.92	8.37	4.81	6.92	8.37	4.81
H	a	8.75	8.30	4.13	8.82	8.35	3.59	8.82	8.35	4.14
	b	9.69	10.33	5.89	9.72	10.36	5.70	9.72	10.36	5.93
	c	10.67	10.46	5.28	10.64	10.48	3.99	10.64	10.48	5.44
	d	6.90	8.51	6.44	6.98	8.58	5.46	6.98	8.58	6.51

Table 1. Δ_{SNR} (dB) obtained by the three methods when restoring the images in Fig. 1(a-d), degraded with blurs: Gaussian 5x5 (G), Uniform 7x7 (U), and Horizontal displacement of 5 pixels (H).

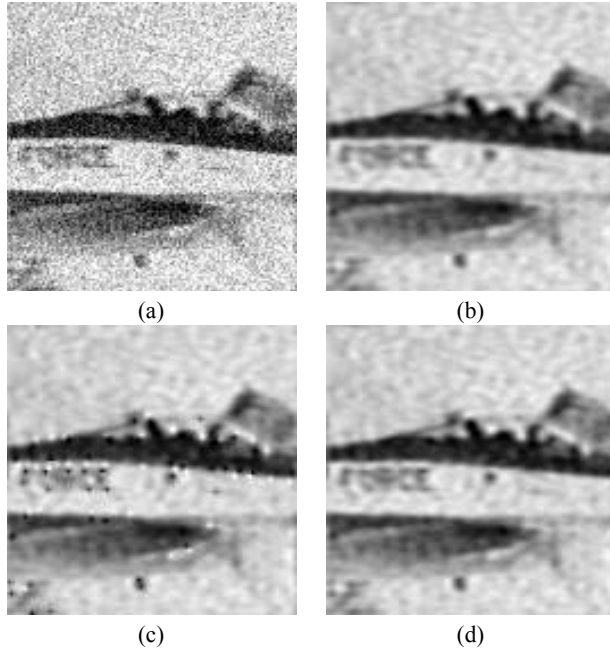


Fig. 2. (a) Image in Fig. 1(b) degraded by motion blur and additive noise of 30 dB. Its restorations using Algorithm I (b), Algorithm II (c), and Algorithm III (d).

calculate the trace in Eq. (44) $\text{diag}(\bar{\alpha}^k)$ is replaced by $\frac{\sum_i \alpha_i^k}{N} \mathbf{I}$, thus allowing for the calculation of the local variances using the Fourier Transform. The actual implementations of Algorithms II and III take 9 secs per iteration step to execute on a 3,4 Mhz intel Xeon desktop PC, for images of size 128×128 , and the blurs used in the experiments. Algorithm I can be implemented in the frequency domain in 0.016 sec per iteration.

The proposed Algorithms have been tested on the set of 128×128 images shown in Fig. 1. Three kinds of blurs have been utilized: Gaussian 3×3 , uniform 7×7 and horizontal motion with a displacement of 5 pixels. Gaussian white noise has been added to obtain degraded images with signal to noise ratios of 10 dB, 20 dB and 30 dB, respectively. The stopping criterion for the three algorithms was $\| \mathbf{E}[\bar{\alpha}]_{q^k(x)} - \mathbf{E}[\bar{\alpha}]_{q^{k-1}(x)} \|^2 / \| \mathbf{E}[\bar{\alpha}]_{q^{k-1}(x)} \|^2 < 10^{-6}$.

Table 1 shows the improvement SNR (Δ_{SNR}) values of the different reconstructions applying Algorithms I, II and III. In most

cases the local variability of the prior variances improves the restoration quality.

Figure 2(a) shows the image in Fig. 1(b) degraded by motion blur and noise (30 dB). The restoration resulting from the application of Algorithms I, II, and III, are shown in Figs. 2(b-d), respectively. In Fig. 2(c) annoying black or white spots in regions of fluctuating luminosity are observed. This problem also appears in other restorations by Algorithm II. In Fig. 2(d) this problem disappears.

5. CONCLUSIONS

In this paper a new Bayesian methodology for the restoration of blurred and noisy images has been developed. This methodology starts by using a global model whose probability posterior distribution has been approximated by applying variational methods. This approximating probability distribution is used to define a more realistic local image model, with spatially dependent prior variances. The new local model better captures the non-homogeneity of the image field, that leading to a better restoration.

6. REFERENCES

- [1] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice Hall, New York, 1977.
- [2] R. Molina, J. Mateos, and A.K. Katsaggelos, "Blind deconvolution using a variational approach to parameter, image, and blur estimation," *IEEE Trans. on Image Processing*, vol. 15, no. 12, pp. 3715–3727, December 2006.
- [3] C.A. Segall, A.K. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Transactions on Image Processing*, vol. 13, no. 7, pp. 898–911, 2004.
- [4] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, pp. 259–268, 1992.
- [5] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Math. Stat.*, vol. 22, pp. 79–86, 1951.
- [6] S. Kullback, *Information Theory and Statistics*, New York, Dover Publications, 1959.
- [7] G. Chantas, N.P. Galatsanos, and A. Likas, "Non stationary bayesian image restoration," in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on Volume 4*, pp. 689 – 692. 2004.