Restoration of Severely Blurred High Range Images Using Stochastic and Deterministic Relaxation Algorithms in Compound Gauss Markov Random Fields *

Rafael Molina[1], Aggelos K. Katsaggelos[2], Javier Mateos[1] and Aurora Hermoso[3]

[1] Departamento de Ciencias de la Computación e I.A. Universidad de Granada. 18071 Granada, España.
[2] Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois 60208-3118.
[3] Departamento de Estadística e I.O. Universidad de Granada. 18071 Granada, España.

Abstract. Over the last few years, a growing number of researchers from varied disciplines have been utilizing Markov random fields (MRF) models for developing optimal, robust algorithms for various problems, such as texture analysis, image synthesis, classification and segmentation, surface reconstruction, integration of several low level vision modules, sensor fusion and image restoration. However, not much work has been reported on the use of this model in image restoration.
In this paper we examine the use of compound Gauss Markov random fields (CGMRF) to restore severely blurred high range images. For this deblurring problem, the convergence of the Simulated Annealing (SA) and Iterative Conditional Mode (ICM) algorithms has not been established. We propose two new iterative restoration algorithms which extend the classical SA and ICM approaches. Their convergence is established and they are tested on real and synthetic images.

1 Introduction

Image restoration refers to the problem of recovering an image, f, from its blurred and noisy observation, g, for the purpose of improving its quality or obtaining some type of information that is not readily available from the degraded image.

It is well known that translation linear shift invariant (LSI) image models do not, in many circumstances, lead to appropriate restoration methods. Their main problem is their inability to preserve discontinuities. To move away from simple LSI models several methods have been proposed.

The CGMRF theory provides us with a mean to control changes in the image model using a hidden random field. A compound random field has two levels; an

* This work has been supported by the "Comisión Nacional de Ciencia y Tecnología" under contract PB93-1110.

upper level which is the real image that has certain translation invariant linear sub-models to represent image characteristics like border regions, smoothness, texture, etc. The lower or hidden level is a finite range random field to govern the transitions between the sub-models. The use of the underlying random field, called the line process, was introduced by Geman and Geman in [4] in the discrete case and extended to the continuous case by Jeng [5], Jeng and Woods [6, 7] and Chellapa, Simchony and Lichtenstein [3].

Given the image and noise models, the process of finding the maximum *a posteriori* (MAP) estimate for the CGMRF is much more complex, since we no longer have a convex function to be minimized and methods like simulated annealing (SA) (see [4]) have to be used. Although this method leads to the MAP estimate, it is a very computationally demanding method. A faster alternative is deterministic relaxation which results in local MAP estimation, also called *iterative conditional mode* (ICM) [1].

In this paper we extend the use of SA to restore high dynamic range images in the presence of blurring, a case where convergence of this method has not been shown (see [5, 6, 7] for the continuous case without blurring).

In Sect. 2 we introduce the notation we use and the proposed model for the image and line processes as well as the noise model. Both, stochastic and deterministic relaxation approaches to obtain the MAP estimate without blurring are presented in Sect. 3. Reasons why these algorithms may be unstable in the presence of blurring are studied in Sect. 4. In Sect. 5 we modify the SA algorithm and its corresponding relaxation approach in order to propose our modified algorithms. Convergence proofs are given in Sect. 7. In Sect. 6 we test both algorithms on real images and Sect. 7 concludes the paper.

2 Notation and Model

We will distinguish between f, the 'true' image which would be observed under ideal conditions and g, the observed image. The aim is to reconstruct f from g. Bayesian methods start with a prior distribution, a probability distribution over images f by which they incorporate information on the expected structure within an image. It is also necessary to specify $p(g \mid f)$, the probability distribution of observed images g if f were the 'true' image. The Bayesian paradigm dictates that the inference about the true f should be based on $p(f \mid g)$ given by

$$\mathrm{p}(f \mid g) = \mathrm{p}(g \mid f)p(f)/\mathrm{p}(g) \propto \mathrm{p}(g \mid f)\mathrm{p}(f). \tag{1}$$

Maximization of (1) with respect to f yields

$$\hat{f} = \arg\max_f \mathrm{p}(f \mid g), \tag{2}$$

the maximum *a posteriori* estimator. For the sake of simplicity, we will denote by $f(i)$ the intensity of the true image at the location of the pixel i on the lattice. We regard f as a $p \times 1$ column vector of values $f(i)$. The convention applies equally to the observed image g. Let us now examine the image and noise models.

The use of CGMRF was first presented in [4] using an Ising model to represent the upper level and a line process to model the abrupt transitions. Extensions to continuous range models using GMRF were presented in [5]. The CGMRF model used in this paper was proposed by Chellapa, Simchony and Lichtenstein in [3] and it is an extension of the Blake and Zisserman's weak membrane model [2] used for surface interpolation and edge detection. The convergence proof that will be given here can also be extended to the CGMRF defined in [5, 6, 7].

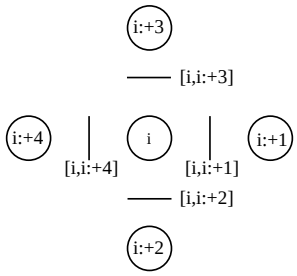

Fig. 1. Image and line sites.

The probability density function (pdf) of the CGMRF is given by

$$
\begin{aligned}
&-\log \mathrm{p}(f,l) = \text{const} \\
&+\frac{1}{2\sigma_{\mathrm{w}}^2}\sum_i \left[\phi(f(i)-f(i:+1))^2(1-l([i,i:+1])+\beta l([i,i:+1])\right. \\
&\left.+\ \phi(f(i)-f(i:+2))^2(1-l([i,i:+2]))+\beta l([i,i:+2])+(1-4\phi)f^2(i)\right], \qquad (3)
\end{aligned}
$$

where $l([i,j])$ takes the value zero if pixels i and j are not separated by an active line and one otherwise and $i:+1$, $i:+2$, $i:+3$, $i:+4$ denote the four pixels around pixel i as described in figure 1. That is, we penalize the introduction of the line element between pixels i and j (see figure 1) by the term $\beta l([i,j])$. The intuitive interpretation of this line process is simple; it acts as an activator or inhibitor of the relation between two neighboring pixels depending on whether or not the pixels are separated by an edge. In this paper we shall use this simple image model. The convergence proofs that follow can be easily extended to more complex image model including more neighboring pixels or interactions in the line process.

Let us now describe the noise model. A simplified but very realistic noise model for many applications is the Gaussian model with mean zero and variance σ_{n}^2. This means that the observed image corresponds to the model $g(i) = (Df)(i)+n(i) = \sum_j d(i-j)f(j)+n(i)$, where D is the $p\times p$ matrix defining the systematic blur, assumed to be known and approximated by a block circulant

matrix, $n(i)$ is the additive Gaussian noise with zero mean and variance σ_{n}^2 and $d(j)$ are the coefficients defining the blurring function.

Then, the probability of the observed image g if f were the 'true' image is

$$\mathrm{p}(g \mid f) \propto \exp\left[-\frac{1}{2\sigma_{\mathrm{n}}^2}\left\|g - Df\right\|^2\right]. \tag{4}$$

3 MAP Estimation Using Stochastic and Deterministic Relaxation

The MAP estimates of f and l, $\hat{f}, \hat{l}$ are given by,

$$\hat{f}, \hat{l} = \arg\max_{f,l} p(f, l \mid g). \tag{5}$$

This is an obvious extension of (2) where now we have to estimate both the image and line processes. The modified simulated annealing (MSA), algorithm we are going proposing in this work, ensures convergence to a local MAP estimate regardless of the initial solution. We start by examining the SA procedure as defined in [5].

Since $\mathrm{p}(f, l \mid g)$ is nonlinear it is extremely difficult to find $\hat{f}$ and $\hat{l}$ by any conventional method. Simulated annealing is a relaxation technique to search for MAP estimates from degraded observations. It uses the distribution

$$\begin{aligned}\mathrm{p}_T(f, l \mid g) = \frac{1}{Z_T}\exp\Bigg\{&-\frac{1}{T}\frac{1}{2\sigma_{\mathrm{n}}^2}\left\|g - Df\right\|^2 \\ &-\frac{1}{T}\frac{1}{2\sigma_{\mathrm{w}}^2}\sum_i\Big[\phi(f(i) - f(i:+1))^2(1 - l([i, i:+1])) + \beta l([i, i:+1]) \\ &+\ \phi(f(i) - f(i:+2))^2(1 - l([i, i:+2])) + \beta l([i, i:+2]) + (1 - 4\phi)f^2(i)\Big]\Bigg\}\end{aligned} \tag{6}$$

where T is the temperature and Z_T a normalization constant.

We shall need to simulate the conditional *a posteriori* density function for $l([i, j])$, given the rest of l, f and g and the conditional *a posteriori* density function for $f(i)$ given the rest of f, l and g. To simulate the line process conditional *a posteriori* density function, $\mathrm{p}_T(l([i, j]) \mid l([k, l]) : \forall [k, l] \neq [i, j], f, g)$, we have

$$\mathrm{p}_T(l([i, j]) = 0 \mid l([k, l]) : \forall [k, l] \neq [i, j], f, g) \propto \exp\left[-\frac{1}{T}\frac{\phi}{2\sigma_{\mathrm{w}}^2}(f(i) - f(j))^2\right], \tag{7}$$

$$\mathrm{p}_T(l([i, j]) = 1 \mid l([k, l]) : \forall [k, l] \neq [i, j], f, g) \propto \exp\left[-\frac{1}{T}\frac{\beta}{2\sigma_{\mathrm{w}}^2}\right]. \tag{8}$$

Furthermore, for our Gaussian noise model,

$$\mathrm{p}_T(f(i) \mid f(j) : \forall j \neq i, l, g) \sim \mathcal{N}\left(\mu^{\underline{l}[i]}(i), T\sigma^{2\underline{l}[i]}(i)\right), \tag{9}$$

where $\mu^{\underline{l}[i]}(i)$ and $\sigma^{2\underline{l}[i]}(i)$ are given by

$$\mu^{\underline{l}[i]}(i) = \lambda^{\underline{l}[i]}(i)\phi \sum_{j \text{ nhbr } i} \frac{f(j)(1-l([i,j]))}{nn^{\underline{l}[i]}(i)} + (1-\lambda^{\underline{l}[i]}(i)) \left(\frac{(D^T g)(i) - (D^T Df)(i)}{c} + f(i) \right), \tag{10}$$

$$\sigma^{2\underline{l}[i]}(i) = \frac{\sigma_{\text{w}}^2 \sigma_{\text{n}}^2}{nn^{\underline{l}[i]}(i)\,\sigma_{\text{n}}^2 + c\,\sigma_{\text{w}}^2}, \tag{11}$$

where c is the sum of the square of the coefficients defining the blur function, that is, $c = \sum_j d(j)^2$, $nn^{\underline{l}[i]}(i) = \phi \sum_{j \text{ nhbr } i}(1 - l([i,j])) + (1 - 4\phi)$ and

$$\lambda^{\underline{l}[i]}(i) = \frac{nn^{\underline{l}[i]}(i)\,\sigma_{\text{n}}^2}{nn^{\underline{l}[i]}(i)\,\sigma_{\text{n}}^2 + c\,\sigma_{\text{w}}^2},$$

and $\underline{l}[i]$ is the four dimensional vector representing the line process configuration around image pixel (i).

Then the sequential SA to find the MAP, with no blurring ($D = I$), proceeds as follows (see [5]):

Algorithm 1 Sequential SA procedure. *Let i_t, $t = 1, 2, \ldots$, be the sequence in which the sites are visited for updating.*

1. *Set $t = 0$ and assign an initial configuration denoted as f_{-1}, l_{-1} and initial temperature $T(0) = 1$.*
2. *The evolution $l_{t-1} \rightarrow l_t$ of the line process can be obtained by sampling the next point of the line process from the raster-scanning scheme based on the conditional probability mass function defined in (7) and (8) and keeping the rest of l_{t-1} unchanged.*
3. *Set $t = t+1$. Go back to step 2 until a complete sweep of the field l is finished.*
4. *The evolution $f_{t-1} \rightarrow f_t$ of the observed system can be obtained by sampling the next value of the observation of the line process from the raster-scanning scheme based on the conditional probability mass function given in (9) and keeping the rest of l_{t-1} unchanged.*
5. *Set $t = t + 1$. Go back to step 4 until a complete sweep of the field f is finished.*
6. *Go to step 2 until $t > t_f$, where t_f is a specified integer.*

The following theorem from [5] guarantees that the SA algorithm converges to the MAP estimate in the case of no blurring.

Theorem 1. *If the following conditions are satisfied:*

1. $|\phi| < 0.25$
2. $T(t) \rightarrow 0$ *as* $t \rightarrow \infty$, *such that*
3. $T(t) \geq C/\log(1 + k(t))$,

then for any starting configuration f_{-1}, l_{-1}*, we have*

$$\mathrm{p}(f_t, l_t \mid f_{-1}, l_{-1}, g) \rightarrow \mathrm{p}_0(f, l) \text{ as } t \rightarrow \infty,$$

where $\mathrm{p}_0(.,.)$ *is the uniform probability distribution over the MAP solutions and* $k(t)$ *is the sweep iteration number at time* t*.*

Instead of using a stochastic approach, we can use a deterministic method to search for a local maximum. An advantage of the deterministic method is that its convergence is much faster than the stochastic approach, since instead of simulating the distributions, the mode from the corresponding conditional distribution is chosen. The disadvantage is the local nature of the solution obtained. This method can be seen as a particular case of SA where the temperature is always set to zero.

4 Instability of the SA and ICM Solutions

Unfortunately, due to the presence of blurring the convergence of SA has not been established for this problem. The main problem of the methods is that, if c is small, as is the case for severely blurred images, the term $[(D^T g)(i) - (D^T Df)(i)]/c$ in (10) is highly unstable. For the ICM method the problem gets worse because sudden changes in the first stages, due to the line process, become permanent (see [9]).

Let us examine intuitively and formally why we may have convergence problems with algorithm 1 and its deterministic relaxation approximation when severe blurring is present. Let us assume for simplicity no line process and examine the iterative procedure where we update the whole image at the same time; it is important to note that this is not the parallel version of SA but an iterative procedure. We have,

$$\begin{aligned} f_t &= \lambda\phi N f_{t-1} - (1-\lambda)\left[\frac{D^T D}{c} f_{t-1} - f_{t-1}\right] + (1-\lambda)\frac{D^T g}{c} \\ &= A f_{t-1} + \text{const}, \end{aligned} \tag{12}$$

where t is the iteration number, understood as sweep of the whole image, and

$$A = \left[I - \lambda(I - \phi N) - (1-\lambda)\frac{D^T D}{c}\right]. \tag{13}$$

For the method to converge A must be a contraction mapping. However this may not be the case. For instance, if the image suffers from severe blurring then c is close to zero and the matrix $[D^T D/c]$ has eigenvalues greater than one. Furthermore, if the image has a high dynamic range, like astronomical images where ranges $[0, 7000]$ are common, it is natural to assume that σ_{w}^2 is big and thus, $(1-\lambda)[D^T D/c]$ has eigenvalues greater than one. Therefore, this iterative method may not converge. It is important to note that, when there is no blurring, $c = 1$ and A is a contraction mapping.

Let us modify A in order to have a contraction. Adding $[(1-\lambda)(1-c)/c]f$ to both sides of (12) we have, in the iterative procedure,

$$\begin{aligned}(1+[(1-\lambda)(1-c)/c])\, f_t = {} & [(1-\lambda)(1-c)/c]f_{t-1} \\ & + Af_{t-1} + \text{const}\end{aligned}$$

or

$$f_t = \omega f_{t-1} + (1-\omega)[Af_{t-1} + \text{const}],$$

with $\omega = (1-c)\sigma_{\text{w}}^2/(\sigma_{\text{n}}^2 + \sigma_{\text{w}}^2)$. We then have for this new iterative procedure

$$f_t = \tilde{A} f_{t-1} + (1-\omega)\text{const},$$

where

$$\tilde{A} = \left[I - \rho(I - \phi N) - (1-\rho)D^T D\right],$$

with $\rho = \sigma_{\text{n}}^2/(\sigma_{\text{n}}^2 + \sigma_{\text{w}}^2)$, is now a contraction mapping.

5 The Modified Simulated Annealing Algorithm

Let us now examine how to obtain a contraction for our iterative procedure. Let us rewrite (10) as an iterative procedure and add $(\alpha(1 - nn^{\underline{l}[i]}(i)) + \beta(1-c))f(i)$ to each side of the equation, we have

$$\begin{aligned}(\alpha+\beta)f_t(i) = {} & (\alpha(1 - nn^{\underline{l}[i]}(i)) + \beta(1-c))f_{t-1}(i) \\ & + \alpha\phi \sum_{j \text{ nhbr } i} f_{t-1}(j)(1 - l([i,j])) \\ & + \beta((D^T g)(i) - (D^T Df)_{t-1}(i) + cf_{t-1}(i)),\end{aligned} \tag{14}$$

where $\alpha = 1/\sigma_{\text{w}}^2$ and $\beta = 1/\sigma_{\text{n}}^2$, or,

$$\begin{aligned}& f_t(i) = \\ & \omega^{\underline{l}_{t-1}[i]}(i) f_{t-1}(i) + (1 - \omega^{\underline{l}_{t-1}[i]}(i)) \times \left(\lambda^{\underline{l}[i]}(i)\phi \sum_{j \text{ nhbr } i} \frac{f_{t-1}(j)(1 - l([i,j]))}{nn^{\underline{l}[i]}(i)} \right. \\ & \left. + (1 - \lambda^{\underline{l}[i]}(i)) \left(\frac{(D^T g)(i) - (D^T Df)_{t-1}(i)}{c} + f_{t-1}(i) \right) \right),\end{aligned}$$

where $\omega^{\underline{l}_{t-1}[i]}(i) = (\sigma_{\text{n}}^2(1 - nn^{\underline{l}_{t-1}[i]}(i)) + (1-c)\sigma_{\text{w}}^2/(\sigma_{\text{n}}^2 + \sigma_{\text{w}}^2)$.

So, in order to have a contraction, we update the whole image at the same time using the value of $f(i)$ obtained in the previous iteration, $f_{t_{k-1}}(i)$, and, instead of simulating from the normal distribution defined in (9) to obtain the new value of $f(i)$, we simulate from the distribution

$$\mathcal{N}\left(\mu_m^{\underline{l}[i]}(i), T\sigma_m^2 {}^{\underline{l}_{t-1}[i]}(i)\right) \tag{15}$$

with mean

$$\mu_m^{\underline{l}[i]}(i) = \omega^{\underline{l}_{t-1}[i]}(i) f_{t_{k-1}}(i) + (1 - \omega^{\underline{l}_{t-1}[i]}(i)) \mu^{\underline{l}_{t-1}[i]}(i) \tag{16}$$

and

$$\sigma_m^{2\ \underline{l}_{t-1}[i]}(i) = (1 - (\omega^{\underline{l}_{t-1}[i]})^2) \sigma^{2\underline{l}_{t-1}[i]}(i). \tag{17}$$

The reason to use this modified variance is clear if we take into account that, if

$$X \sim \mathcal{N}(m, \sigma^2)$$

and

$$Y|X \sim \mathcal{N}\left(\lambda X + (1-\lambda)m, (1-\lambda^2)\sigma^2\right),$$

where $0 < \lambda < 1$, then

$$Y \sim \mathcal{N}(m, \sigma^2).$$

We then have for this iterative method that the transition probabilities are

$$\pi_{T(t_k)}(f_{t_k} \mid f_{t_{k-1}}, l_{t_k}, g) \propto$$
$$\exp[-\frac{1}{2T(t_k)}[f_{t_k} - \underline{M}^{l_{t_k}} f_{t_{k-1}} - \underline{Q}^{l_{t_k}} g]^t [\underline{Q}_1^{l_{t_k}}]^{-1} [f_{t_k} - \underline{M}^{l_{t_k}} f_{t_{k-1}} - \underline{Q}^{l_{t_k}} g]], \tag{18}$$

where

$$\underline{M}^{l_{t_k}} = \Omega^{l_{t_k}} + (I - \Omega^{l_{t_k}})(C^{l_{t_k}} - (D^T D)_*^{l_{t_k}}), \tag{19}$$
$$\underline{Q}^{l_{t_k}} = (I - \Omega^{l_{t_k}}) B^{l_{t_k}}, \tag{20}$$

where

$$C^{\underline{l}_{t_k}[i]} * f_{t_k}(i) = \phi \lambda^{\underline{l}_{t_k}[i]}(i) \sum_{j \text{ nhbr } i} \frac{(1 - l([i,j]))}{nn^{\underline{l}_{t_k}[i]}(i)} f_{t_k}(j),$$

and

$$(D^T D)_*^{\underline{l}_{t_k}[i]} * f_{t_k}(i) = (1 - \lambda^{\underline{l}_{t_k}[i]}(i)) \left(\frac{(D^T D f)(i)}{c} - f(i) \right),$$

$\Omega^{l_{t_k}}$ is a diagonal matrix with entries $\omega^{\underline{l}_{t_k}[i]}(i)$ and $\underline{Q}_1^{l_{t_k}}$ is a diagonal matrix with entries $\sigma_m^{2\ \underline{l}[i]}(i)$.

In the coming section we apply the modified SA and ICM algorithms, whose convergence is established in the appendix to restore astronomical images.

The algorithms are the following:

Algorithm 2 Sequential MSA procedure. *Let i_t, $t = 1, 2, \ldots$, be the sequence in which the sites are visited for updating.*

1. *Set $t = 0$ and assign an initial configuration denoted as f_{-1}, l_{-1} and initial temperature $T(0) = 1$.*

2. *The evolution $l_{t-1} \rightarrow l_t$ of the line process can be obtained by sampling the next point of the line process from the raster-scanning scheme based on the conditional probability mass function defined in (7) and (8) and keeping the rest of l_{t-1} unchanged.*
3. *Set $t = t+1$. Go back to step 2 until a complete sweep of the field l is finished.*
4. *The evolution $f_{t-1} \rightarrow f_t$ of the observed system can be obtained by sampling the next value of the whole image based on the conditional probability mass function given in (15)*
5. *Go to step 2 until $t > t_f$, where t_f is a specified integer.*

The following theorem guarantees that the MSA algorithm converges to a local MAP estimate, even in the presence of blurring.

Theorem 2. *If the following conditions are satisfied:*

1. $|\phi| < 0.25$
2. $T(t) \rightarrow 0$ *as* $t \rightarrow \infty$, *such that*
3. $T(t) \geq C/\log(1 + k(t))$,

then for any starting configuration f_{-1}, l_{-1}, we have

$$\mathrm{p}(f_t, l_t \mid f_{-1}, l_{-1}, g) \rightarrow \mathrm{p}_0(f, l) \text{ as } t \rightarrow \infty,$$

where $\mathrm{p}_0(.,.)$ is a probability distribution over local MAP solutions and $k(t)$ is the sweep iteration number at time t.

The modified ICM procedure is obtained by selecting in steps 2 and 4 of Algorithm 2 the mode of the corresponding transition probabilities.

6 Test Examples

Let us examine how the modified ICM algorithm works on a synthetic star image, blurred with an atmospherical point spread function (PSF), D, given by

$$d(i) \propto \left(1 + (u^2 + v^2)/R^2\right)^{-\delta}, \tag{21}$$

with $\delta = 3$, $R = 3.5$, $i = (u, v)$, and Gaussian noise with $\sigma_{\mathrm{n}}^2 = 64$. If we use $\sigma_{\mathrm{w}}^2 = 24415$, which is realistic for this image, and take into account that, for the PSF defined in (21), $c = 0.02$, A defined in (13) is not a contraction. Figures 2a and 2b depict the original and corrupted image, respectively. Restorations from the original and modified ICM methods with $\beta = 2$ for 2500 iterations are depicted on Fig. 2c and Fig. 2d, respectively. Similar results are obtained with 500 iterations.

The methods were also tested on images of Saturn which were obtained at the Cassegrain f/8 focus of the 1.52-m telescope at Calar Alto Observatory (Spain) on July, 1991. Results are presented on a image taken through a narrow-band interference filter centered at the wavelength 9500 Å.

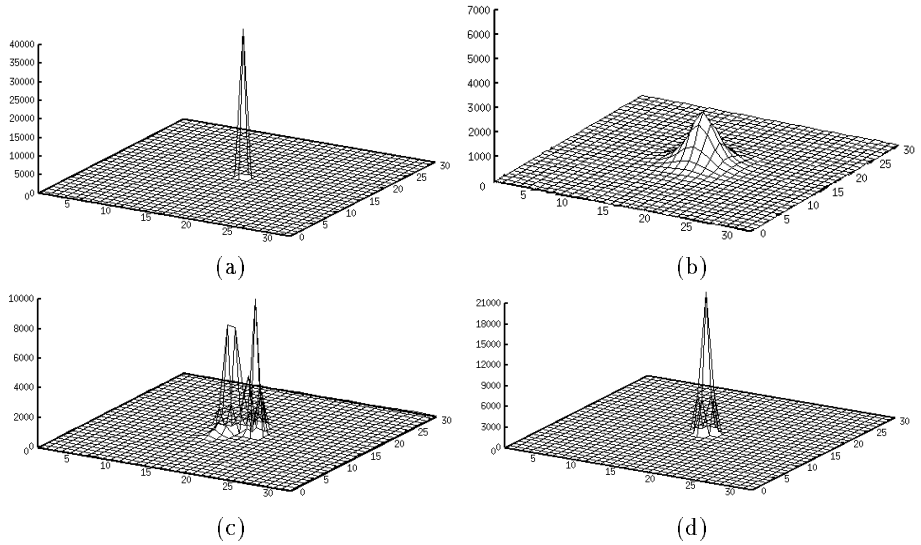

Fig. 2. (a) Original image. (b) Observed image. (c) ICM restoration. (d) Restoration with the proposed ICM method.

The blurring function defined in (21) was used. The parameters δ and R were estimated from the intensity profiles of satellites of Saturn that were recorded simultaneously with the planet and of stars that were recorded very close in time and airmass to the planetary images. We found $\delta \sim 3$ and $R \sim 3.4$ pixels.

Figure 3 depicts the original image and the restorations after running the original ICM and our proposed ICM methods for 500 iterations and the original SA and our proposed SA methods for 5000 iterations. In all the images the improvement in spatial resolution is evident. To examine the quality of the MAP estimate of the line process we compared it with the position of the ring and disk of Saturn, obtained from the Astronomical Almanac, corresponding to our observed image. Although all the methods detect a great part of the ring and the disk, the ICM method, Fig. 4a, shows thick lines. The SA method, on the other hand, gives us thinner lines and the details are more resolved, Fig. 4b. Obviously there are some gaps in the line process but better results would be obtained by using 8 neigbors instead of 4 or, in general, adding more l-terms to the energy function.

Table 1 shows the computing time per iteration of the studied methods. They are referred to the computing time of the ICM method. The little difference between the ICM and SA methods is due to the fact that most of the time is spent in convolving images.

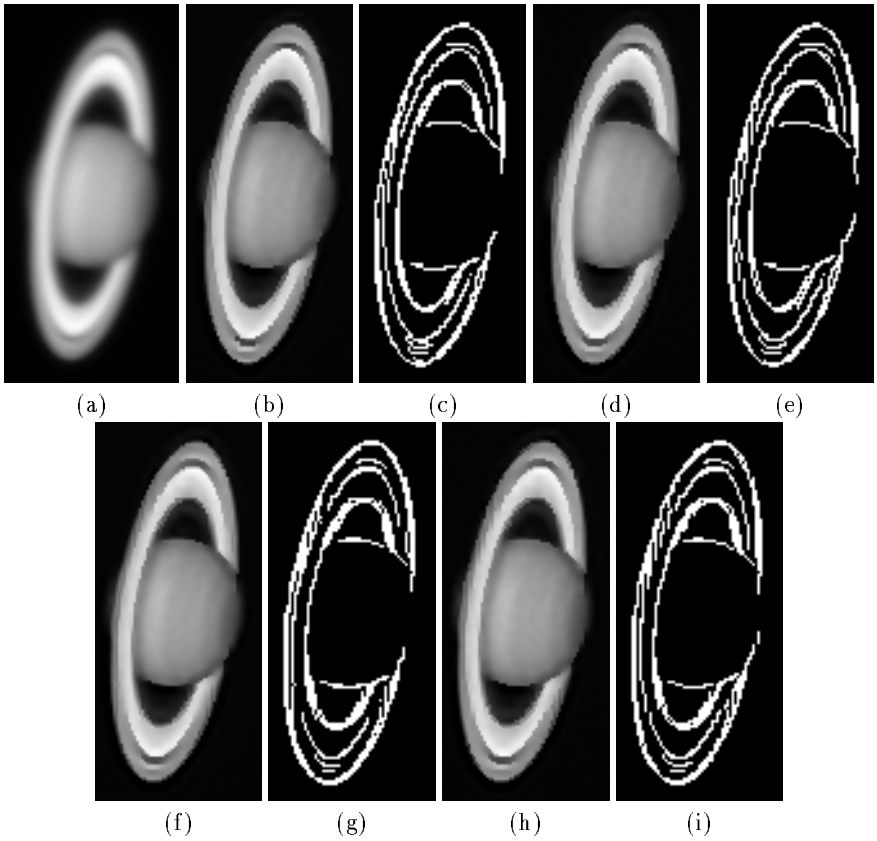

Fig. 3. (a) Original image, (b) restoration with the original ICM method and (c) its line process, (d) restoration with the original SA method and (e) its line process. (f) restoration with the proposed ICM method and (g) its line process, (h) restoration with the proposed SA method and (i) its line process.

7 Conclusions

In this paper we have presented two new methods that can be used to restore high dynamic range images in the presence of severe blurring. These methods extend the classical ICM and SA procedures, so that convergence of the algorithms is now guaranteed. The experimental results verify the derived theoretical results. Further extensions of the algorithms are under consideration.

Appendix: Convergence of the MSA Procedure

In this section we shall examine the convergence of the MSA algorithm. It is important to make clear that in this new iterative procedure we simulate $f(i)$

(a) (b)

Fig. 4. Comparison between the real edges (light) and the obtained line process (dark). (a) Proposed ICM method, (b) Proposed SA method

Table 1. Computing times per iteration of the methods referred to the ICM computing time.

Method	Original ICM	Original SA	Proposed ICM	Proposed SA
Relative Time	1.00	1.13	0.12	0.20

using (15) and to simulate $l([i,j])$ we keep using (7) and (8). We shall denote by π_T the corresponding transition probabilities. That is, $\pi_{T(t_k)}(f_{t_k}|f_{t_{k-1}}, l_{t_k}, g)$ is obtained from (18) and $\pi_{T(t_k)}(l_{t_k}|f_{t_{k-1}}, l_{t_{k-1}})$ is obtained from (7) and (8).

Since updating the whole image at the same time prevents us from having $P_t \mathrm{p}_{T(t)} = \mathrm{p}_{T(t)}$ we will not be able to show the convergence to the global MAP estimates using the same proofs as in [4, 6].

To prove the convergence of the chain we need some lemmas and definitions as in [5, 6].

We assume a measure space (Ω, Σ, μ) and a conditional density function $\pi_n(s_n|s_{n-1})$ which defines a Markov chain $s_1, s_2, \ldots, s_n, \ldots$. In our application, the s_i are vectors valued with a number of elements equal to the number of pixels in the image. For simplicity, we assume Ω is R^d and μ is a Lebesgue measure on R^d. Define a Markov operator $P_n : L^1 \rightarrow L^1$ as follows

$$P_n f(s_n) = \int_\Omega \pi_n(s_n|s_{n-1}) f(s_{n-1}) ds_{n-1}. \tag{22}$$

By P_n^m we mean the composite operation $P_{n+m} P_{n+m-1} \ldots P_{n+2} P_{n+1}$. The convergence problem we are dealing with is the same as the convergence of P_0^m as $m \rightarrow \infty$.

Definition 3. Let x be a vector with components $x(i)$ and Q be a matrix with components $q(i,j)$. We define $\| x \|_2$ and $\| Q \|_2$ as follows

$$\| x \|_2 = (\sum_i |x(i)|^2)^{1/2},$$

$$\| Q \|_2 = \sup_x \frac{\| Qx \|_2}{\| x \|_2} = \max_i (\rho(i))^{1/2},$$

where $\rho(i)$ are the eigenvalues of matrix $Q^t Q$.

Definition 4. A continuous nonnegative function $V : \Omega \to R$ is a Liapunov function if

$$\lim_{\| s \| \to \infty} V(s) = \infty, \tag{23}$$

where $\| s \|$ is a norm of s.

Denote by D the set of all pdfs with respect to Lebesgue measure and the L_1 norm defined as follows:

$$\| f \|_1 = \int_\Omega |f(s)| ds \quad \forall f \in L^1.$$

Definition 5. Let $P_n : L^1 \to L^1$ be a Markov operator. Then $\{P_n\}$ is said to be asymptotically stable if, for any $f_1, f_2 \in D$,

$$\lim_{m \to \infty} \| P_0^m (f_1 - f_2) \|_1 = 0. \tag{24}$$

Let us prove the following lemma.

Lemma 6. *If $|\phi| < 0.25$ then, $\forall l$,*

$$\| \underline{M}^l \|_2 < 1,$$

where $\underline{M}^l$ has been defined in (19).

Proof. First we note that from (14)

$$f_t(i) = f_{t-1}(i) - \rho(\phi \sum_{j \,\mathrm{nhbr}\, i} (f_{t-1}(i) - f_{t-1}(j))(1 - l([i,j])) + (1 - 4\phi) f_{t-1}(i))$$
$$+ (1 - \rho)((D^T g)(i) - (D^T D f)_{t-1}(i)),$$

where $\rho = \alpha/(\alpha + \beta)$.

So, $\underline{M}^l$ is symmetric and for any vector x

$$x^T \underline{M}^l x = \sum x(i)^2 - \rho(\sum_i \phi(x(i) - x(i:+1))^2 (1 - l([i, i:+1]))$$
$$- \rho(\sum_i \phi(x(i) - x(i:+2))^2 (1 - l([i, i:+2])) + (1 - 4\phi) \sum_i x^2(i))$$
$$- (1 - \rho) x^T D^T D x$$

Obviously if $|\phi| < 0.25$, $\forall x \neq 0$, $x^T \underline{M}^l x < \sum x(i)^2$. Furthermore,

$$\begin{aligned} x^T \underline{M}^l x &\geq \sum x(i)^2 - \rho(\sum_i \phi(x(i) - x(i:+1))^2 + \sum_i \phi(x(i) - x(i:+2))^2 \\ &\quad +(1-4\phi)\sum_i x^2(i)) - (1-\rho)x^T D^T D x \\ &= x^T(I - \rho(I - \phi N) - (1-\rho)D^T D)x \end{aligned}$$

and if $|\phi| < 0.25$, $-I < (I - \rho(I - \phi N) - (1-\rho)D^T D)$. So, if $|\phi| < 0.25$,

$$-I < \underline{M}^l < I$$

and

$$x^T \underline{M}^{l^T} \underline{M}^l x < x^T x,$$

which proves that $\underline{M}^l$ is a contraction matrix for $|\phi| < 0.25$. □

We shall also use the following lemma from [5, 6].

Lemma 7. *Assume B is a d-dimensional positive definite matrix with eigenvalues $\rho(1) \geq \rho(2) \geq \ldots \geq \rho(d) > 0$ and $B = J^t D J$, where D is a diagonal matrix which consists of the eigenvalues. Let $b > 0$, then*

$$\frac{1}{\sqrt{(2\pi)^d |B|}} \int_{\|x\|_2 > b} \exp[-x^t B^{-1} x] dx \geq q \left(\frac{b}{\sqrt{\rho(d)}} \right)^{d-2} \exp\left[-\frac{b^2}{2\rho(d)} \right].$$

The following theorem from [6] gives the sufficient conditions on the convergence of P_0^m in terms of transition density functions.

Theorem 8. *Let (Ω, Σ, μ) be a measure space and μ be Lebesgue measure on R^d. If there exists a Liapunov function, $V : \Omega \to R$, such that*

$$\int_\Omega V(s_n)\pi_n(s_n|s_{n-1})ds_n \leq \alpha V(s_{n-1}) + \beta \quad \text{for } 0 \leq \alpha < 1 \quad \text{and} \quad \beta \geq 0 \tag{25}$$

and

$$\sum_{i=1}^{\infty} \| h_{m_i} \|_1 = \infty, \quad m_i = i\tilde{m} \quad \text{for any integer } \tilde{m} > 0, \tag{26}$$

where

$$h_{m_i}(s_{m_i}) = \inf_{\|s_{m_i-1}\| \leq r} \pi_{m_i}(s_{m_i} \mid s_{m_i-1}) \tag{27}$$

and r is a positive number satisfying the following inequality:

$$V(s) > 1 + \frac{\beta}{1-\alpha} \qquad \forall \| s \| > r$$

then, for the Markov operator, $P_n : L^1 \to L^1$, defined by (22) we have that P_0^m is asymptotically stable.

We are going to show that the sufficient conditions of Theorem 8 are satisfied by the Markov chain defined by our MSA procedure. The proof follows the same steps as the one given in [5, 6].

Let $V(f,l)$ be the Liapunov function

$$V(f,l) = \| f \|_2 + \| l \|_2 \ . \tag{28}$$

Step 1: Show that

$$\sum_{l_{t_k}} \int_\Omega V(f_{t_k}, l_{t_k}) \pi_{T(t_k)}(f_{t_k}, l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k} \leq \beta + \alpha V(f_{t_{k-1}}, l_{t_{k-1}}).$$

First we show that

$$\int_\Omega \| f_{t_k} \|_2 \ \pi_{T(t_k)}(f_{t_k} | f_{t_{k-1}}, l_{t_k}, g) df_{t_k} \leq \beta_1 + \alpha \| f_{t_{k-1}} \|_2 \quad \forall l_{t_k}. \tag{29}$$

We have

$$\begin{aligned}
&\int_\Omega \| f_{t_k} \|_2 \ \pi_{T(t_k)}(f_{t_k} | f_{t_{k-1}}, l_{t_k}, g) df_{t_k} = \qquad \text{(by change of variable)} \\
&= \int_\Omega \| \bar{f}_{t_k} + \underline{M}^{l_{t_k}} f_{t_{k-1}} + \underline{Q}^{l_{t_k}} g \|_2 \ \text{const} \ \exp\left[-\frac{1}{2T(t_k)} \bar{f}_{t_k}^t [\underline{Q}_1^{l_{t_k}}]^{-1} \bar{f}_{t_k} \right] d\bar{f}_{t_k} \\
&\leq \int_\Omega (\| \bar{f}_{t_k} \|_2 + \| \underline{M}^{l_{t_k}} f_{t_{k-1}} \|_2 + \| \underline{Q}^{l_{t_k}} g \|_2) \text{const} \\
&\quad \exp\left[-\frac{1}{2T(t_k)} \bar{f}_{t_k}^t [\underline{Q}_1^{l_{t_k}}]^{-1} \bar{f}_{t_k} \right] d\bar{f}_{t_k} \\
&\leq \beta_1 + \alpha \| f_{t_{k-1}} \|_2,
\end{aligned} \tag{30}$$

where

$$\begin{aligned}
\alpha &= \max_l \| \underline{M}^l \|_2, \\
\beta_1 &= \max_l \left[T(t_k)]^{1/2} [trace(\underline{Q}_1^l)]^{1/2} + \| \underline{Q}^{l_{t_k}} g \|_2 \right],
\end{aligned}$$

with $\alpha < 1$, since for Lemma 6, $\underline{M}^l$ is a contraction, $\forall l$.

Furthermore, it can be easily shown that

$$\sum_{l_{t_k}} \pi_{T(t_k)}(l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}) \| l_{t_k} \|_2 \leq \beta_1 + \alpha \| l_{t_{k-1}} \|_2, \tag{31}$$

since l_{t_k} has only a finite number of levels, choosing β_1 big enough, the above inequality obviously holds.

Let us now show that (29) holds. We have, using (30) and (31),

$$\sum_{l_{t_k}} \int_{\Omega} V(f_{t_k}, l_{t_k}) \pi_{T(t_k)}(f_{t_k}, l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k} =$$

$$= \sum_{l_{t_k}} \int_{\Omega} \| f_{t_k} \|_2 \, \pi_{T(t_k)}(f_{t_k}, l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k}$$

$$+ \sum_{l_{t_k}} \int_{\Omega} \| l_{t_k} \|_2 \, \pi_{T(t_k)}(f_{t_k}, l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k}$$

$$\leq \beta_1 + \alpha \| f_{t_{k-1}} \|_2 + \beta_2 \| l_{t_{k-1}} \|_2 = \beta + \alpha V(f_{t_{k-1}}, l_{t_{k-1}}).$$

Step 2: Step 2 is the same that in [5, 6] using

$$\delta_{\max} = \max \left\{ \sup_{\|f\|_2 \leq a} [\frac{\phi}{2\sigma_{\mathrm{w}}^2}(f(i) - f(j))^2, \frac{\beta}{2\sigma_{\mathrm{w}}^2}] \right\}$$

and

$$\delta_{\min} = \min \left\{ \inf_{\|f\|_2 \leq a} ([\frac{\phi}{2\sigma_{\mathrm{w}}^2}(f(i) - f(j))^2, \frac{\beta}{2\sigma_{\mathrm{w}}^2}] \right\}.$$

References

1. Besag, J. (1986), "On the Statistical Analysis of Dirty Pictures", *J. Royal Statistics Soc.* B l 48, 259–302.
2. Blake, A. and Zisserman, A. (1987), "Visual Reconstruction", Cambridge, MIT Press.
3. Chellapa, R., Simchony, T. and Lichtenstein, Z. (1991), "Image Estimation Using 2D Noncausal Gauss-Markov Random Field Models", in *Digital Image Restoration*, Katsaggelos, A.K. (ed.), Springer Series in Information Science, vol. 23, Springer–Verlag.
4. Geman, S., Geman D. (1984), "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images", *IEEE Trans. on PAMI*, vol. PAMI-9, (l 6), 721–742.
5. Jeng, F.C., (1988), "Compound Gauss-Markov Random Fields for Image Estimation and Restoration", *Ph.D Thesis*, Rensselaer Polythenic Institute.
6. Jeng, F.C., and Woods, J.W. (1988), "Simulated Annealing in Compound Gaussian Random Fields", *IEEE Trans. Inform. Theory*, l 36, 94–107.
7. Jeng, F.C. and Woods, J.W. (1991), "Compound Gauss-Markov Models for Image Processing", in *Digital Image Restoration*, Katsaggelos, A.K. (ed.), Springer Series in Information Science, vol. 23, Springer–Verlag.
8. Molina, R., Ripley, B.D., Molina, A., Moreno, F. and Ortiz, J.L. (1992), "Bayesian Deconvolution with Prior Knowledge of Object Location: Applications to Ground-Based Planetary Images", *AJ*, l 104, 1662–1668.
9. Molina, R., Katsaggelos, A.K., Mateos, J. and Abad, J. (1996), "Restoration of Severely Blurred High Range Images Using Compound Models", *Proceeding of ICIP-96.*
10. Ripley, B.D. (1981) "Spatial Statistics", Wiley, New York.

This article was processed using the LaTeX macro package with LLNCS style