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], Aurora Hermoso[3] and C. Andrew Segall[2]

[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, no much work has been reported on the use of Simulated Annealing (SA) and Iterative Conditional Mode (ICM) algorithms for Maximum a Posteriori estimation in image restoration problems when severe blurring is present.
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 and Iterative Conditional Mode (ICM) algorithms has not been established. We propose two new iterative restoration algorithms which can be considered as extensions of the classical SA and ICM approaches and whose convergence is established. Finally, they are tested on real and synthetic images and the results compared with the restorations obtained by other iterative schemes.

Keywords: Image restoration, Compound Gauss-Markov Random Fields, Simulated Annealing, Iterative Conditional Mode.

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

* This work has been supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIC97-0989.

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 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[1] in the discrete case and extended to the continuous case by Jeng[2], Jeng and Woods[3,4] and Chellapa, Simchony and Lichtenstein[5].

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 Geman and Geman[1]) have to be used. Although this method, when blurring is not present, 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)[6].

When blur is present, different approaches have been proposed to find the MAP. Blake and Zisserman[7] propose the use of Gradually Non Convexity, which can be extended to the blurring problem, Molina and Ripley[8] propose the use of a log-scale for the image model, Green[9], Bouman and Sauer[10], Schultz and Stevenson[11] and Lange[12] use convex potentials in order to ensure uniqueness of the solution.

Additional solutions are motivated by the application of nonlinear partial differential equations in image processing, particularly anisotropic diffusion[13]. These operators are designed to smooth an image while preserving discontinuities (e.g. You et al.[14], Catte[15], St. Marc[16]). While traditionally utilized for image enhancement, embedding the diffusion inhibitor within a variational framework facilitates image restoration. Restoration using different variants of the diffusion mechanism has been presented by Geman and McClure(17), Hebert and Leahy[18], Green[9] and Bouman and Sauer[10]. Design of the edge preserving function is critical for solution convergence and quality. Selection of suitable expressions has been explored by Charbonnier et al.[19].

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 Jeng[2] and Jeng and Woods[3,4] 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 de-

terministic 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 compare the results with other existing methods. Sect. 7 concludes the paper.

2 Notation and Model

2.1 Bayesian 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.

2.2 Incorporation of the Edge Locations in the Prior Model

The use of CGMRF was first presented by Geman and Geman[1] 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 by Jeng[2]. The CGMRF model used in this paper was proposed by Chellapa, Simchony and Lichtenstein in[5] and it is an extension of the Blake and Zisserman's weak membrane model[7] used for surface interpolation and edge detection. The convergence proof that will be given here can also be extended to the CGMRF defined by Jeng[2] and Jeng and Woods[3,4].

Let us first describe the prior model without any edges. Our prior knowledge about the smoothness of the object luminosity distribution makes it possible to model the distribution of f by a CAR (Conditional Auto-Regressive) model (see Ripley[20]). Thus,

$$\mathrm{p}(f) \propto \exp\left\{-\frac{1}{2\sigma_{\mathrm{w}}^2} f^T(I - \phi N)f\right\}, \tag{3}$$

where $N_{ij} = 1$ if cells i and j are spatial neighbors (pixels at distance one), zero otherwise and ϕ just less than 0.25. The parameters can be interpreted by the following expressions describing the conditional distribution

$$\mathrm{E}(f(i) \mid f(j),\ j \neq i) = \phi \sum_{j\ \mathrm{nhbr}\ i} f(j), \qquad \mathrm{var}(f(i) \mid f(j),\ j \neq i) = \sigma_{\mathrm{w}}^2, \quad (4)$$

where the suffix 'j nhbr i' denotes the four neighbor pixels at distance one from pixel i (see Fig. 1). The parameter σ_{w}^2 measures the smoothness of the 'true' image.

INSERT FIGURE 1 HERE.

From (3) we have

$$-\log \mathrm{p}(f) = \mathrm{const} + \frac{1}{2\sigma_{\mathrm{w}}^2} \sum_i f(i)(f(i) - \phi(Nf)(i))$$

Then, if $i:+1$, $i:+2$, $i:+3$, $i:+4$ denote the four pixels around pixel i as described in Fig. 1 and we assume a 'toroidal edge correction', we have

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

This expression can be rewritten as

$$\begin{aligned} &-\log \mathrm{p}(f, l) = \mathrm{const} \\ &+\frac{1}{2\sigma_{\mathrm{w}}^2} \sum_i \left[\phi(f(i) - f(i:+1))^2 (1 - l([i, i:+1]) + \beta l([i, i:+1]) \right. \\ &\left. + \phi(f(i) - f(i:+2))^2 (1 - l([i, i:+2])) + \beta l([i, i:+2]) + (1 - 4\phi) f^2(i)\right], \quad (5) \end{aligned}$$

where $l([i, j]) \equiv 0$ for all i and j.

We now introduce a line process by simply redefining the function $l([i, j])$ as taking the value zero if pixels i and j are not separated by an active line and one otherwise. We then penalize the introduction of an active line element in the position $[i, j]$ (see Fig. 1) by the term $\beta l([i, j])$ since otherwise the expression in (5) would obtain its minimum value by setting all line elements equal to one. The intuitive interpretation of this line process is simple; it acts as an activator or inhibitor of the relation between two neighbor pixels depending on whether or not the pixels are separated by an edge.

2.3 Noise Models

A simplified but very realistic noise model for many applications is to assume that it is Gaussian with mean zero and variance σ_{n}^2. That 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

$$\mathrm{p}(g \mid f) \propto \exp\left[-\frac{1}{2\sigma_{\mathrm{n}}^2}\left\|g - Df\right\|^2\right]. \tag{6}$$

3 MAP Estimation Using Stochastic and Deterministic Relaxation

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

$$\hat{f}, \hat{l} = \arg\max_{f,l} p(f, l \mid g). \tag{7}$$

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 by Jeng[2].

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

$$\begin{aligned}\mathrm{p}_T(f, l \mid g) = \frac{1}{Z_T}\exp\Bigg\{ & -\frac{1}{T}\frac{1}{2\sigma_{\mathrm{n}}^2}\left\|g - Df\right\|^2 \\ & -\frac{1}{T}\frac{1}{2\sigma_{\mathrm{w}}^2}\sum_i \big[\phi(f(i) - f(i:+1))^2(1 - l([i, i:+1])) + \beta l([i, i:+1]) \\ & + \phi(f(i) - f(i:+2))^2(1 - l([i, i:+2])) + \beta l([i, i:+2]) + (1 - 4\phi)f^2(i)\big]\Bigg\}\end{aligned} \tag{8}$$

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

$$\mathrm{p}_T(l([i,j]) = 0 \mid l([k,l]) : \forall [k,l] \neq [i,j], f, g) \propto \exp\left[-\frac{1}{T}\frac{\phi}{2\sigma_{\mathrm{w}}^2}(f(i) - f(j))^2\right], \tag{9}$$

$$\mathrm{p}_T(l([i,j]) = 1 \mid l([k,l]) : \forall [k,l] \neq [i,j], f, g) \propto \exp\left[-\frac{1}{T}\frac{\beta}{2\sigma_{\mathrm{w}}^2}\right]. \tag{10}$$

Furthermore, for our Gaussian noise model,

$$\mathrm{p}_T(f(i) \mid f(j) : \forall j \neq i, l, g) \sim \mathcal{N}\left(\mu^{\underline{l}[i]}(i), T\sigma^{2\underline{l}[i]}(i)\right), \tag{11}$$

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

$$\begin{aligned}\mu^{\underline{l}[i]}(i) = \lambda^{\underline{l}[i]}(i)\phi \sum_{j \text{ nhbr } i} \frac{f(j)(1 - l([i,j]))}{nn^{\underline{l}[i]}(i)} \\ +(1 - \lambda^{\underline{l}[i]}(i)) \left(\frac{(D^T g)(i) - (D^T D f)(i)}{c} + f(i)\right),\end{aligned} \tag{12}$$

$$\sigma^{2\underline{l}[i]}(i) = \frac{\sigma_{\mathrm{w}}^2 \sigma_{\mathrm{n}}^2}{nn^{\underline{l}[i]}(i)\,\sigma_{\mathrm{n}}^2 + c\,\sigma_{\mathrm{w}}^2}, \tag{13}$$

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

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

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 Jeng[2]):

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 $l_{t-1} \to l_t$ 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 (9) and (10) 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 $f_{t-1} \to f_t$ of the image f can be obtained by sampling the next value of the image process from the raster-scanning scheme based on the conditional probability mass function given in (11) and keeping the rest of f_{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.*

Note that steps 2–3 consist in an exact and direct sampling from the independent conditional law $\mathrm{p}_T(l([i,j]) \mid f_{t-1}, g) = \prod_{<i,j>} \mathrm{p}_T(l([i,j]) \mid f_{t-1}(i), f_{t-1}(j)$, while steps 4–5 consist in one sweep of Gibbs sampling from conditional distribution $\mathrm{p}_T(f \mid l_t, g)$ with initial configuration f_{t-1}.

The following theorem from Jeng[2] 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. $T(t) \to 0$ *as* $t \to \infty$, *such that*
3. $T(t) \geq C_T / \log(1 + k(t))$,

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

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

where $\mathrm{p}_0(.,.)$ *is the uniform probability distribution over the MAP solutions,* C_T *is a positive constant 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 $[(D^T g)(i) - (D^T D f)(i)]/c$ in (12) 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 Molina et al.[21]).

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,

$$\begin{aligned} f_t &= \lambda\phi N f_{t-1} - (1-\lambda)\left[\frac{D^T D}{c} f_{t-1} - f_{t-1}\right] + (1-\lambda)\frac{D^T g}{c} \\ &= A f_{t-1} + \text{const}, \end{aligned} \tag{14}$$

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

$$A = \left[I - \lambda(I - \phi N) - (1-\lambda)\frac{D^T D}{c}\right]. \tag{15}$$

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 $\sigma_{\rm w}^2$ is big and thus, $(1-\lambda)[D^TD/c]$ 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 (14) we have, in the iterative procedure,

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

or

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

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

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

where

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

with $\rho = \sigma_{\rm n}^2/(\sigma_{\rm n}^2 + \sigma_{\rm w}^2)$, 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 (12) 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

$$\begin{aligned}(\alpha+\beta) f_t(i) = {} & (\alpha(1 - nn^{\underline{l}[i]}(i)) + \beta(1-c)) f_{t-1}(i) \\ & + \alpha\phi \sum_{j \text{ nhbr } i} f_{t-1}(j)(1 - l([i,j])) \\ & + \beta((D^Tg)(i) - (D^TDf)_{t-1}(i) + c f_{t-1}(i)),\end{aligned} \tag{16}$$

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

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

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

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 (11) to obtain the new value of $f(i)$, we simulate from the distribution

$$\mathcal{N}\left(\mu_m^{\underline{l}_{t-1}[i]}(i), T\sigma_m^{2\,\underline{l}_{t-1}[i]}(i)\right) \tag{17}$$

with mean

$$\mu_m^{\underline{l}_{t-1}[i]}(i) = \omega^{\underline{l}_{t-1}[i]}(i) f_{t_{k-1}}(i) + (1 - \omega^{\underline{l}_{t-1}[i]}(i))\mu^{\underline{l}_{t-1}[i]}(i) \tag{18}$$

and

$$\sigma_m^{2\,\underline{l}_{t-1}[i]}(i) = (1 - (\omega^{\underline{l}_{t-1}[i]}(i))^2)\sigma^{2\underline{l}_{t-1}[i]}(i). \tag{19}$$

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

$$\pi_{T(t_k)}(f_{t_k} \mid f_{t_{k-1}}, l_{t_k}, g) \propto$$
$$\exp[-\frac{1}{2T(t_k)}[f_{t_k} - \underline{M}^{l_{t_k}} f_{t_{k-1}} - \underline{Q}^{l_{t_k}} g]^t [\underline{Q}_1^{l_{t_k}}]^{-1} [f_{t_k} - \underline{M}^{l_{t_k}} f_{t_{k-1}} - \underline{Q}^{l_{t_k}} g]], \tag{20}$$

where

$$\underline{M}^{l_{t_k}} = \Omega^{l_{t_k}} + (I - \Omega^{l_{t_k}})(C^{l_{t_k}} - (D^T D)_*^{l_{t_k}}), \tag{21}$$
$$\underline{Q}^{l_{t_k}} = (I - \Omega^{l_{t_k}})B^{l_{t_k}}, \tag{22}$$

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 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 (9) and (10) 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 image system can be obtained by sampling the next value of the whole image based on the conditional probability mass function given in (17)*
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. $T(t) \geq C_T / \log(1 + k(t))$,

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, C_T is a constant and $k(t)$ is the sweep iteration number at time t.

We notice that if the method converges to a configuration $(\overline{f}, \overline{l})$, then

$$\overline{f} = \arg\max_f p(f|\overline{l}, g)$$

Furthermore,

$$\overline{l} = \arg\max_l p(l|\overline{f}, g)$$

We conjecture that the method we are proposing converges to an distribution over global maxima. However, the difficulty of using synchronous models prevent us from proving that result (See Younes[22]).

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

$$d(i) \propto \left(1 + (u^2 + v^2)/R^2\right)^{-\delta}, \tag{23}$$

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 (23), $c = 0.02$, A defined in (15) 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.

INSERT FIGURE 2 HERE.

The proposed methods were also tested on real images and compared with ARTUR, the method proposed by Charbonnier et al.[19]. ARTUR minimizes energy functions of the form

$$J(f) = \lambda^2 \left\{ \sum_i \varphi[f(i) - f(i:+1)] + \sum_i \varphi[f(i) - f(i:+2)] \right\} + \| g - Df \|^2 \tag{24}$$

where λ is a positive constant and φ is a potential function satisfying some edge preserving conditions. The potential functions we used in our experiments, φ_{GM}, φ_{HL}, φ_{HS} and φ_{GR}, are shown in table 1.

Charbonnier et al.[19] show that, for those φ functions, it is always possible to find a function J^* such that

$$J(f) = \inf_l J^*(f, l)$$

where J^* is a dual energy which is quadratic in f when l is fixed. l can be understood as a line process which, for those potential functions, takes values in the interval $[0, 1]$.

To minimize (24), Charbonnier et al. propose the following iterative scheme:

1. $n = 0$, $f^0 \equiv 0$
2. Repeat
   3. $l^{n+1} = \arg\min_l [J^*(f^n, l)]$
   4. $f^{n+1} = \arg\min_f [J^*(f, l^{n+1})]$
   5. $n = n + 1$
6. Until convergence.

In our experiments the convergence criterion used in step 6 above was

$$\| f^{n+1} - f^n \|^2 / \| f^n \|^2 < 10^{-6}.$$

The solution of step 4 was found by a Gauss-Seidel algorithm. The stopping criterion was

$$\| f^{n+1,m+1} - f^{n+1,m} \|^2 / \| f^{n+1,m} \|^2 < 10^{-6}$$

where m is the iteration number.

We use images of Saturn which were obtained at the Cassegrain f/8 focus of the 1.52m. 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.

The blurring function defined in (23) 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.

INSERT TABLE 1 HERE.

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. In particular, ring light contribution has been successfully removed from equatorial regions close to the actual location of the rings and amongst the rings of Saturn, the Cassini division is enhanced in contrast, and the Encke division appears on the ansae of the rings in all deconvolved images. 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 neighbors instead of 4 or, in general, adding more l-terms to the energy function.

Figure 5 depicts the results after running ARTUR using potential functions φ_{GM}, φ_{HL}, φ_{HS} and φ_{GR} on the Saturn image together with the results obtained by the proposed ICM method. Note that line processes obtained by the potential functions used with ARTUR are presented in inverse gray levels. The results suggest that φ_{GM} and φ_{HL} capture better the lines of the image than φ_{HS} and φ_{GR}. Lines captured by all these functions are thicker than the obtained by the proposed ICM method, notice that the line process produced by these potential

functions is continuous on the interval $[0,1]$. Furthermore, φ_{GM} also captures some low intensity lines, due to the noise, that creates some artifacts on the restoration, specially on the Saturn rings, see figure 5b. Finally, the potential functions used with ARTUR have captured the totality of the planet contour although the line process intensity on the contour is quite low.

INSERT FIGURE 3 HERE.
INSERT FIGURE 4 HERE.
INSERT FIGURE 5 HERE.

The methods were also tested on images of Jupiter which were also obtained at the Cassegrain f/8 focus of the 1.52-m telescope at Calar Alto Observatory (Spain) on August, 1992.

The blurring function was the same as in the previous experiment. Figure 6 depicts the original image and the restorations after running the original ICM and our proposed ICM method for 500 iterations and our proposed SA method for 5000 iterations. In all the images the improvement in spatial resolution is evident. Features like the Equatorial plumes and great red spot are very well detected. ARTUR was also tested on these images obtaining similar results to the obtained with Saturn.

INSERT FIGURE 6 HERE.

In order to obtain a numerical comparison, ARTUR and our methods were tested and compared using the cameraman image. The image was blurred using the PSF defined in (23) with the parameters $\delta = 3$ and $R = 4$. Gaussian noise with variance 62.5 was added obtaining a image with $SNR = 20dB$. The original and observed image are shown on figure 7.

INSERT FIGURE 7 HERE.

Figures 8 and 9 depicts the restorations after running our proposed SA method for 5000 iterations, our proposed ICM method for 500 iterations and ARTUR with different potential functions.

INSERT FIGURE 8 HERE.

INSERT FIGURE 9 HERE.

In order to compare the quality of the restorations we used the peak signal-to-noise ratio ($PSNR$) that, for two images f and g of size $M \times N$, is defined as

$$PSNR = 10 \log_{10} \left[\frac{M \times N \times 255^2}{\| g - f \|^2} \right].$$

Results, shown in table 2, are quite similar for all the methods but they suggests that better results are obtained when our proposed SA method is used. For the two best methods in terms of the $PSNR$, our proposed SA and ARTUR with φ_{GM}, we have included cross-sections of the original and restored images in Fig. 10. It can be observed that, although both profiles are quite similar, our proposed SA method obtain sharper edges than the ones obtained with φ_{GM}.

INSERT TABLE 2 HERE.
INSERT FIGURE 10 HERE.

Table 3 shows the total computing time of the studied methods after running them on one processor of a Silicon Graphics Power Challenge XL. It also shows the relative execution time 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.

INSERT TABLE 3 HERE.

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, and the convergence of the algorithms is 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 $f(i)$

using (17) and to simulate $l([i,j])$ we keep using (9) and (10). 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 (20) and $\pi_{T(t_k)}(l_{t_k}|f_{t_{k-1}}, l_{t_{k-1}})$ is obtained from (9) and (10).

Since updating the whole image at the same time prevents us from having a stationary distribution we will not be able to show the convergence to the global MAP estimates using the same proofs as in Geman and Geman[1] and Jeng and Woods[3].

To prove the convergence of the chain we need some lemmas and definitions as in Jeng[2] and Jeng and Woods[3].

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 $P_n : L^1 \to L^1$ as follows

$$P_n\pi(s_n) = \int_\Omega \pi_n(s_n|s_{n-1})\pi(s_{n-1})ds_{n-1}. \tag{25}$$

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 $m \to \infty$.

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^tQ.

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

$$\lim_{\|s\|\to\infty} V(s) = \infty, \tag{26}$$

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:

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

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 $\pi_1, \pi_2 \in D$,

$$\lim_{m\to\infty} \| P_0^m(\pi_1 - \pi_2) \|_1 = 0. \tag{27}$$

The following theorem from Jeng and Woods[3] gives the sufficient conditions on the asymptotic stability of P_0^m in terms of transition density functions.

Theorem 6. *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*

$$\int_\Omega V(s_n)\pi_n(s_n|s_{n-1})ds_n \leq \alpha V(s_{n-1}) + \beta \quad \text{for } 0 \leq \alpha < 1 \quad \text{and} \quad \beta \geq 0 \tag{28}$$

and

$$\sum_{i=1}^{\infty} \| h_{m_i} \|_1 = \infty, \quad m_i = i\tilde{m} \quad \text{for any integer} \quad \tilde{m} > 0, \tag{29}$$

where

$$h_{m_i}(s_{m_i}) = \inf_{\|s_{m_i-1}\| \leq r} \pi_{m_i}(s_{m_i} \mid s_{m_i-1}) \tag{30}$$

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 (25) we have that P_0^m is asymptotically stable.

We are going to show that the sufficient conditions of Theorem 6 are satisfied by the Markov chain defined by our MSA procedure when the parameters describing the algorithm 2 satisfy the conditions of theorem 2. Since asymptotic stability of an inhomogeneous Markov sequence implies convergence, this will imply for our restoration problem that for any starting configuration f_{-1}, l_{-1}, we have

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

where $\mathrm{p}_0(.,.)$ is the probability distribution over the MAP solutions.

Let us prove the following lemma.

Lemma 7. *If $|\phi| < 0.25$ then, $\forall l$,*

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

where $\underline{M}^l$ has been defined in (21).

Proof. First we note that from (16)

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

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

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

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

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 Jeng[2] and Jeng and Woods[3].

Lemma 8. *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*

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

Using these two lemmas, let us now show that the sufficient conditions of Theorem 6 are satisfied.

The proof follows the same steps as the one given in Jeng[2] and Jeng and Woods[3].

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

$$V(f,l) = \| f \|_2 + \| l \|_2 . \tag{31}$$

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

$$\int_{\Omega} \| f_{t_k} \|_2 \, \pi_{T(t_k)}(f_{t_k}|f_{t_{k-1}}, l_{t_k}, g) df_{t_k} \leq \beta_1 + \alpha \| f_{t_{k-1}} \|_2 \quad \forall l_{t_k}. \tag{32}$$

We have

$$\begin{aligned}
&\int_{\Omega} \| f_{t_k} \|_2 \, \pi_{T(t_k)}(f_{t_k}|f_{t_{k-1}}, l_{t_k}, g) df_{t_k} = \qquad \text{(by change of variable)} \\
&= \int_{\Omega} \| \bar{f}_{t_k} + \underline{M}^{l_{t_k}} f_{t_{k-1}} + \underline{Q}^{l_{t_k}} g \|_2 \, \text{const} \, \exp\left[-\frac{1}{2T(t_k)} \bar{f}^t_{t_k} [\underline{Q}_1^{l_{t_k}}]^{-1} \bar{f}_{t_k} \right] d\bar{f}_{t_k} \\
&\leq \int_{\Omega} (\| \bar{f}_{t_k} \|_2 + \| \underline{M}^{l_{t_k}} f_{t_{k-1}} \|_2 + \| \underline{Q}^{l_{t_k}} g \|_2) \text{const} \\
&\quad \exp\left[-\frac{1}{2T(t_k)} \bar{f}^t_{t_k} [\underline{Q}_1^{l_{t_k}}]^{-1} \bar{f}_{t_k} \right] d\bar{f}_{t_k} \\
&\leq \beta_1 + \alpha \| f_{t_{k-1}} \|_2,
\end{aligned} \tag{33}$$

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 7, $\underline{M}^l$ is a contraction, $\forall l$.

Furthermore, it can be easily shown that

$$\sum_{l_{t_k}} \pi_{T(t_k)}(l_{t_k}|f_{t_{k-1}}, l_{t_{k-1}}) \| l_{t_k} \|_2 \leq \beta_1 + \alpha \| l_{t_{k-1}} \|_2, \tag{34}$$

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 (32) holds. We have, using (33) and (34),

$$\begin{aligned}
&\sum_{l_{t_k}} \int_{\Omega} V(f_{t_k}, l_{t_k}) \pi_{T(t_k)}(f_{t_k}, l_{t_k}|f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k} = \\
&= \sum_{l_{t_k}} \int_{\Omega} \| f_{t_k} \|_2 \, \pi_{T(t_k)}(f_{t_k}, l_{t_k}|f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k} \\
&+ \sum_{l_{t_k}} \int_{\Omega} \| l_{t_k} \|_2 \, \pi_{T(t_k)}(f_{t_k}, l_{t_k}|f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k} \\
&\leq \beta_1 + \alpha \| f_{t_{k-1}} \|_2 + \beta_2 \| l_{t_{k-1}} \|_2 = \beta + \alpha V(f_{t_{k-1}}, l_{t_{k-1}}).
\end{aligned}$$

Step 2: Show that if temperature $T(t)$ decreases as $\frac{C_T}{\log(1+k(t))}$, then for any $n_0 > 0$, we have

$$\sum_{m=1}^{\infty} ||h_{mt_{n_0}}||_1 = \infty,$$

where C_T is a positive constant and $k(t)$ is the iteration number of sweeps up to time t.

This result can be obtained by establishing a lower bound for $||h_{t_k}||_1$. First, recall that

$$h_{t_k} = \inf_{(f_{t_{k-1}}, l_{t_{k-1}}) \in R_a} \pi_{T(t_k)}(f_{t_k}, l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}, g),$$

where R_a is defined as

$$R_a = \{(f, l) | V(f, l) \leq a\}.$$

By definition of L_1 norm, we have

$$||h_{t_k}||_1 = \sum_{l_{t_k}} \int \inf_{(f_{t_{k-1}}, l_{t_{k-1}}) \in R_a} \pi_{T(t_k)}(f_{t_k}, l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k}$$

and, from the definition of the iterative procedure,

$$\begin{aligned}
&||h_{t_k}||_1 = \sum_{l_{t_k}} \int \inf_{(f_{t_{k-1}}, l_{t_{k-1}}) \in R_a} \pi_{T(t_k)}(f_{t_k}, l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}, g) df_{t_k} \\
&= \sum_{l_{t_k}} \int \inf_{(f_{t_{k-1}}, l_{t_{k-1}}) \in R_a} \pi_{T(t_k)}(f_{t_k} | f_{t_{k-1}}, l_{t_k}, g) \pi_{T(t_k)}(l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}) df_{t_k} \\
&\geq \inf_{l_{t_k}} \left[\inf_{(f_{t_{k-1}}, l_{t_{k-1}}) \in R_a} \pi_{T(t_k)}(l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}) \right. \\
&\quad \left. \int \inf_{(f_{t_{k-1}}, l_{t_{k-1}}) \in R_a} \pi_{T(t_k)}(f_{t_k} | f_{t_{k-1}}, l_{t_k}, g) df_{t_k} \right] \\
&\geq \inf_{(l_{t_k}, l_{t_{k-1}})} \inf_{||f_{t_{k-1}}||_2 \leq a} \pi_{T(t_k)}(l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}) \\
&\quad \inf_{l_{t_k}} \left[\int \inf_{||f_{t_{k-1}}||_2 \leq a} \pi_{T(t_k)}(f_{t_k} | f_{t_{k-1}}, l_{t_k}, g) df_{t_k} \right].
\end{aligned}$$

Let

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

and

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

then, from (9) and (10),

$$\begin{aligned}
\inf_{(l_{t_k}, l_{t_{k-1}})} \inf_{||f_{t_{k-1}}||_2 \leq a} \pi_{T(t_k)}(l_{t_k} | f_{t_{k-1}}, l_{t_{k-1}}) &\geq \left(\frac{1}{2} \exp \left[-\frac{\delta_{\max} - \delta_{\min}}{T(t_k)} \right] \right)^{\tilde{p}} \\
&= C_2 \exp \left[-\frac{C_1}{T(t_k)} \right],
\end{aligned}$$

where $C_1 = \tilde{p}(\delta_{\max} - \delta_{\min})$ and $C_2 = (1/2)^{\tilde{p}}$ with $\tilde{p}$ the number of line sites.

Hence,

$$||h_{t_k}||_1 \geq C_2 \exp\left[-\frac{C_1}{T(t_k)}\right] \inf_{l_{t_k}} \int \inf_{||f_{t_{k-1}}||_2 \leq a} \pi_{T(t_k)}(f_{t_k}|f_{t_{k-1}}, l_{t_k}, g) df_{t_k}.$$

Now, from (20),

$$\pi_{T(t_k)}(f_{t_k}|f_{t_{k-1}}, l_{t_k}, g) = \text{const}$$
$$\times \exp\left[-\frac{1}{2T(t_k)}[f_{t_k} - \underline{M}^{l_{t_k}} f_{t_{k-1}} - \underline{Q}^{l_{t_k}} g]^t [\underline{Q}_1^{l_{t_k}}]^{-1} [f_{t_k} - \underline{M}^{l_{t_k}} f_{t_{k-1}} - \underline{Q}^{l_{t_k}} g]\right].$$

It can be shown that there exists a finite number b with $b >> a$ such that

$$\int \inf_{||f_{t_{k-1}}||_2 \leq a} \pi_{T(t_k)}(f_{t_k}|f_{t_{k-1}}, l_{t_k}, g) df_{t_k} \geq$$
$$\frac{1}{\left[(2\pi)^d T(t_k)|\underline{Q}_1^{l_{t_k}}|\right]^{\frac{1}{2}}} \int_{||f_{t_k}||_2 > b} \exp\left[-\frac{1}{2T(t_k)} f_{t_k}^t [\underline{Q}_1^{l_{t_k}}]^{-1} f_{t_k}\right] df_{t_k}.$$

Then, from Lemma 8, we can establish the following inequality:

$$\frac{1}{\left[(2\pi)^d T(t_k)|\underline{Q}_1^{l_{t_k}}|\right]^{\frac{1}{2}}} \int_{||f_{t_k}||_2 > b} \exp\left[-\frac{1}{2T(t_k)} f_{t_k}^t [\underline{Q}_1^{l_{t_k}}]^{-1} f_{t_k}\right] df_{t_k} \geq$$
$$q\left(\frac{b}{\sqrt{T(t_k)\lambda_d}}\right)^{d-2} \exp\left[-\frac{b^2}{2T(t_k)\lambda_d}\right],$$

where λ_d is the smallest eigenvalue of $\underline{Q}_1^{l_{t_k}}$ and b is a positive constant.

Obviously λ_d and b depend on l_{t_k}. Denoting λ_d^* and b^* such that

$$\left(\frac{b^*}{\sqrt{T(t_k)\lambda_{d^*}}}\right)^{d-2} \exp\left[-\frac{b*^2}{2T(t_k)\lambda_{d^*}}\right] =$$
$$\inf_{l_{t_k}} \left(\frac{b}{\sqrt{T(t_k)\lambda_d}}\right)^{d-2} \exp\left[-\frac{b^2}{2T(t_k)\lambda_d}\right],$$

we have

$$||h_{t_k}||_1 \geq C_2 \exp\left[-\frac{C_1}{T(t_k)}\right] q\left(\frac{b^*}{\sqrt{T(t_k)\lambda_{d^*}}}\right)^{d-2} \exp\left[-\frac{b^{*2}}{2T(t_k)\lambda_{d^*}}\right].$$

Then, if we take $T(t_k) = \frac{C_T}{\log(1+k)}$, we have

$$||h_{t_k}||_1 \geq \frac{C^*(\log(1+k))^{(d-2)/2}}{1+k},$$

where

$$C^* = C_2 q \left(\frac{b^*}{\sqrt{T(t_k)\lambda_{d^*}}} \right)^{d-2}$$

and

$$C_T = C_1 + \frac{b^{*2}}{2\lambda_{d^*}},$$

and so

$$\sum_{m=1}^{\infty} ||h_{mt_{n_0}}||_1 = \infty.$$

References

1. Geman, S. and Geman D., Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, *IEEE Trans. on PAMI*, **9**, 721–742 (1984).
2. Jeng, F.C., Compound Gauss-Markov Random Fields for Image Estimation and Restoration, *Ph.D Thesis*, Rensselaer Polytechnic Institute (1988).
3. Jeng, F.C., and Woods, J.W. , Simulated Annealing in Compound Gaussian Random Fields, *IEEE Trans. Inform. Theory*, **36**, 94–107 (1988).
4. Jeng, F.C. and Woods, J.W., Compound Gauss-Markov Models for Image Processing, *Digital Image Restoration*, Katsaggelos, A.K. (ed.), Springer Series in Information Science, **23**, Springer–Verlag (1991).
5. Chellapa, R., Simchony, T. and Lichtenstein, Z., Image Estimation Using 2D Noncausal Gauss-Markov Random Field Models, *Digital Image Restoration*, Katsaggelos, A.K. (ed.), Springer Series in Information Science, **23**, Springer–Verlag (1991).
6. Besag, J., "On the Statistical Analysis of Dirty Pictures", *J. Royal Statistics Soc.* B **48**, 259–302 (1986).
7. Blake, A. and Zisserman, A., "Visual Reconstruction", Cambridge, MIT Press (1987).
8. Molina R. and Ripley B. D., Using spatial models as priors in astronomical images analysis, *J. Appl. Statist.* **16**, 193–206 (1989).
9. Green, P.J., Bayesian reconstruction from emission tomography data using a modified EM algorithm, *IEEE Trans. on Medical Imaging*, **9**, 84–92 (1990).
10. Bouman, C. and Sauer, K., A Generalized Gaussian Image Model for Edge-Preserving MAP Estimation, *IEEE Trans. on Image Processing* **2**, 296–230 (1993).
11. Schultz, R. R. and Stevenson, R. L., Stochastic modeling and estimation of multiespectral image data, *IEEE Trans. on Image Processing*, **4**, 1109–1119 (1995).
12. Lange, K. Convergence of EM image reconstruction algorithms with Gibbs smoothing, *IEEE Trans. on Image Processing*, **4**, 439–446 (1990).
13. Perona, P., and Malik, J., Scale-space and edge detection using anisotropic diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, 629–639 (1990).
14. You, Y.L., Xu, W, Tannenbaum, A. and Kaveh, M., Behavioral analysis of anisotropic diffusion in image processing, *IEEE Transactions on Image Processing*, **5**, 1539–1553 (1996).

15. Catte, F., Lions, P. L., Morel, J.M., and Coll, T., Image selective smoothing and edge detection by nonlinear diffusion, *SIAM Journal of Numerical Analysis*, **29**, 182–198 (1992).
16. Saint-Marc, P., Chen, J. S. and Medioni, G., Adaptive smoothing: a general tool for early vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**, 514–529 (1991).
17. Geman, S., and McClure, D. E., Bayesian image analysis: An application to single photon emission tomography, *the Proc. Stat. Comput. Sect. Washington, D.C.: Amer. Stat. Assoc.*, 12–18, (1985).
18. Hebert, T. and Leahy, R., A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors, *IEEE Transactions on Medical Imaging*, **8**, 194–202 (1989).
19. Charbonnier, P., Blanc-Feraud, L., Aubert G. and Barlaud M. Deterministic edge-preserving regularization in computed imaging, *Transactions on Image Processing* **5**, 298–311 (1997).
20. Ripley, B.D., Spatial Statistics, Wiley, New York (1981).
21. Molina, R., Katsaggelos, A.K., Mateos, J. and Abad, J., Restoration of Severely Blurred High Range Images Using Compound Models, *Proceeding of ICIP-96* **2**, 469–472 (1996).
22. Younes, L., Synchronous Random Fields and Image Restoration, *CMLA, Ecole Normales Supérieure de Cachan.* Technical Report. (1998)

Rafael Molina was born in 1957 and received his degree in Mathematics (Statistics) in 1979. He completed his Ph. D. Thesis in 1984 in Optimal Design in Linear Models and became Associate Professor in Computer Science and Artificial Intelligence at the University of Granada in 1989. His areas of research interest are image restoration (applications to astronomy and medicine), parameter estimation, image compression, and blind deconvolution. He is a member of the IEEE, SPIE, Royal Statistical Society and AERFAI (Asociación Española de Reconocimento de Formas y Análisis de Imágenes) and currently the Dean of the Computer Engineering Faculty at the University of Granada.

Aggelos K. Katsaggelos received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1979 and the M.S. and Ph.D. degrees both in electrical engineering from the Georgia Institute of Technology, Atlanta, Georgia, in 1981 and 1985, respectively.
In 1985 he joined the Department of Electrical Engineering and Computer Science at Northwestern University, Evanston, IL, where he is currently professor, holding the Ameritech Chair of Information Technology. During the 1986-1987 academic year he was an assistant professor at Polytechnic University, Department of Electrical Engineering and Computer Science, Brooklyn, NY. His current research interests include image and video recovery, video compression, motion estimation, boundary encoding, computational vision, and multimedia signal processing. Dr. Katsaggelos is a Fellow of the IEEE, an Ameritech Fellow, a member of the Associate Staff, Department of Medicine, at Evanston Hospital, and a member of SPIE. He is a member of the Steering Committee of the *IEEE Transactions on Medical Imaging*, the IEEE Technical Committees on Visual Signal Processing and Communications, Image and Multi-Dimensional Signal Processing, and Multimedia Signal Processing, and the editor-in-chief of the *IEEE Signal Processing Magazine.* He has served as an Associate editor for the *IEEE Transactions on Signal Processing* (1990-1992), an area editor for the journal *Graphical Models and Image Processing* (1992-1995), a member of the Steering Committee of the *IEEE Transactions on Image Processing* (1992-1997). He is the editor of *Digital Image Restoration* (Springer-Verlag, Heidelberg, 1991), co-author of *Rate-Distortion Based Video Compression* (Kluwer Academic Publishers, 1997), and co-editor of *Recovery Techniques for Image and Video Compression,* (Kluwer Academic Publishers, 1998). He has served as the General Chairman of the 1994 Visual Communications and Image Processing Conference (Chicago, IL), and as technical program co-chair of the 1998 IEEE International Conference on Image Processing (Chicago, IL).

Javier Mateos was born in Granada (Spain) in 1968. He received his Diploma and M. S. degrees in Computer Science from the University of Granada in 1990 and 1991, respectively. He completed his Ph. D. in Computer Science at the University of Granada in July 1998.
Since September 1992, he has been Assistant Professor at the Department of Computer Science and Artificial Intelligence of the University of Granada. His research interest includes image restoration and image and video recovery and

compression. He is a member of the AERFAI (Asociación Española de Reconocimento de Formas y Análisis de Imágenes).

Aurora Hermoso was born in 1957 and received her degree in Mathematics (Statistics) in 1979. She completed her Ph. D. Thesis in 1984 in Likelihood test In Log Normal Processes and became Associate Professor in Statistics at the University of Granada in 1986. Her area of research interest is Estimation in Stochastic Systems. She is a member of SEIO (Sociedad Española de Estadística e Investigación Operativa).

C. Andrew Segall is a Ph.D. student at Northwestern University, Evanston, Illinois, and a member of the Image and Video Processing Laboratory. His research interests are in image processing and include scale space theory, nonlinear filtering and target tracking.

Captions of Figures and Tables

Fig. 1. Image and line sites.

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

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.

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

Fig. 5. (a) Restoration with the proposed ICM method (f) and (k) its corresponding horizontal and vertical line process. Figures (b), (c), (d) and (e) show the restoration when ARTUR is run for the potentials φ_{GM}, φ_{HL}, φ_{HS} and φ_{GR}, respectively. Their corresponding horizontal line processed are shown in figures (g)–(j) and their vertical processed are shown in figures (l)–(o).

Fig. 6. (a) Original image, (b) restoration with the original ICM method and (c) its line process, (d) restoration with the proposed ICM method and (e) its line process, (f) restoration with the proposed SA method and (g) its line process.

Fig. 7. (a) original cameraman image. (b) observed image.

Fig. 8. (a) restoration with the proposed SA method and (b), (c) its horizontal and vertical line process, (d) restoration with the proposed ICM method and (e), (f) its horizontal and vertical line process, (g) restoration with ARTUR with φ_{GM} function and (h), (i) its horizontal and vertical line process.

Fig. 9. (a) restoration with ARTUR with φ_{HL} function and (c), (c) its horizontal and vertical line process, (d) restoration with ARTUR with φ_{HS} function and (e), (f) its horizontal and vertical line process, (g) restoration with ARTUR with φ_{GR} function and (h), (i) its horizontal and vertical line process.

Fig. 10. Cross section of line 153 of the original cameraman image (solid line) and reconstructed images (dotted line) with (a) proposed SA and (b) ARTUR with φ_{GM}.

Table 1. Edge preserving potential functions used with ARTUR.

Table 2. Comparison of the different restoration methods in terms of $PSNR$.

Table 3. Total computing time of the methods and relative time per iteration referred to the ICM.

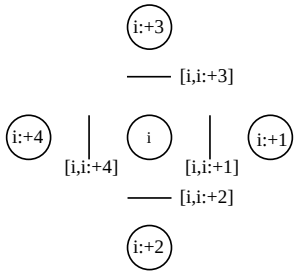
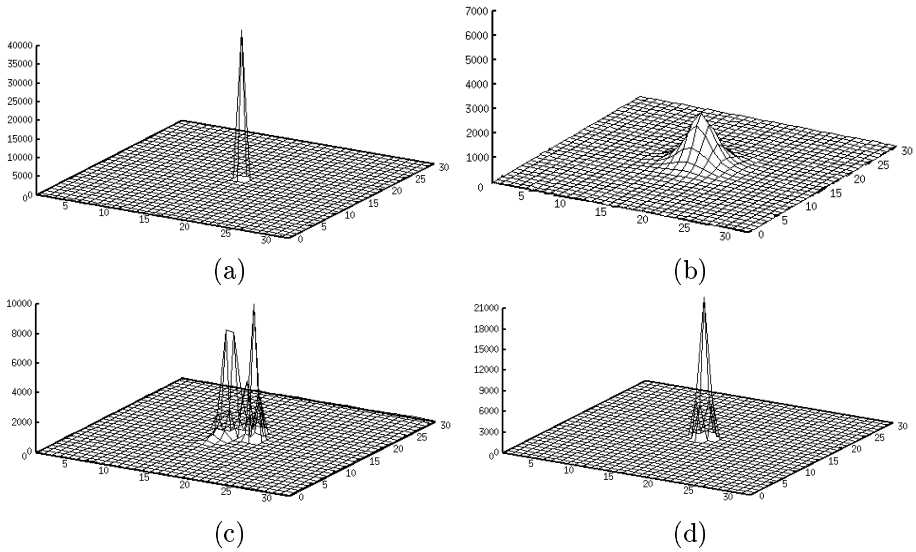


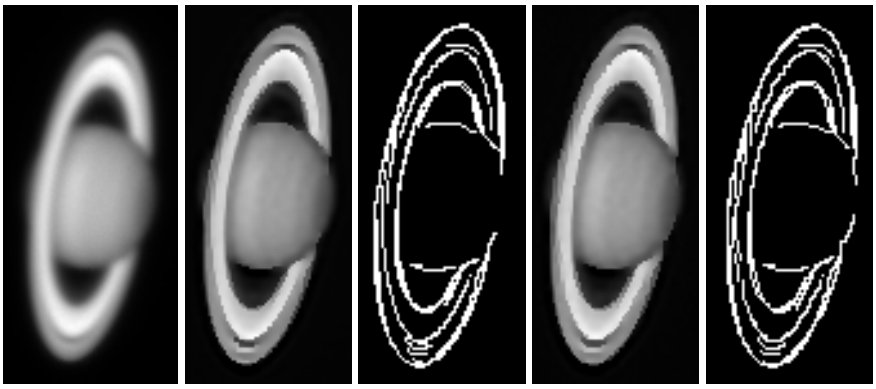
Fig. 1


40000
35000
30000
25000
20000
15000
10000
5000
0
7000
6000
5000
4000
3000
2000
1000
0
10000
8000
6000
4000
2000
0
21000
18000
15000
12000
9000
6000
3000
0
5
10
15
20
25
30

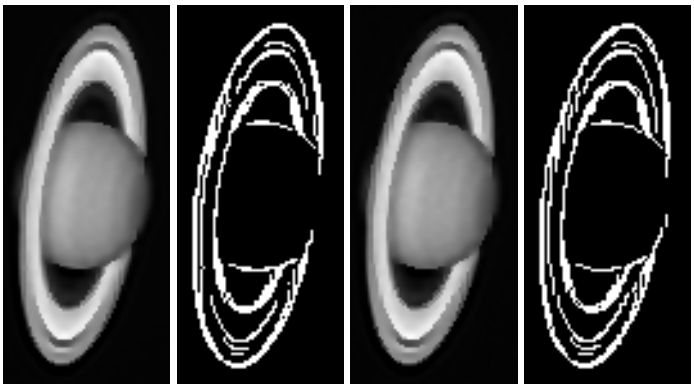
(a) (b)

(c) (d)

Fig. 2



(a) (b) (c) (d) (e)



(f) (g) (h) (i)

Fig. 3

(a) (b)

Fig. 4

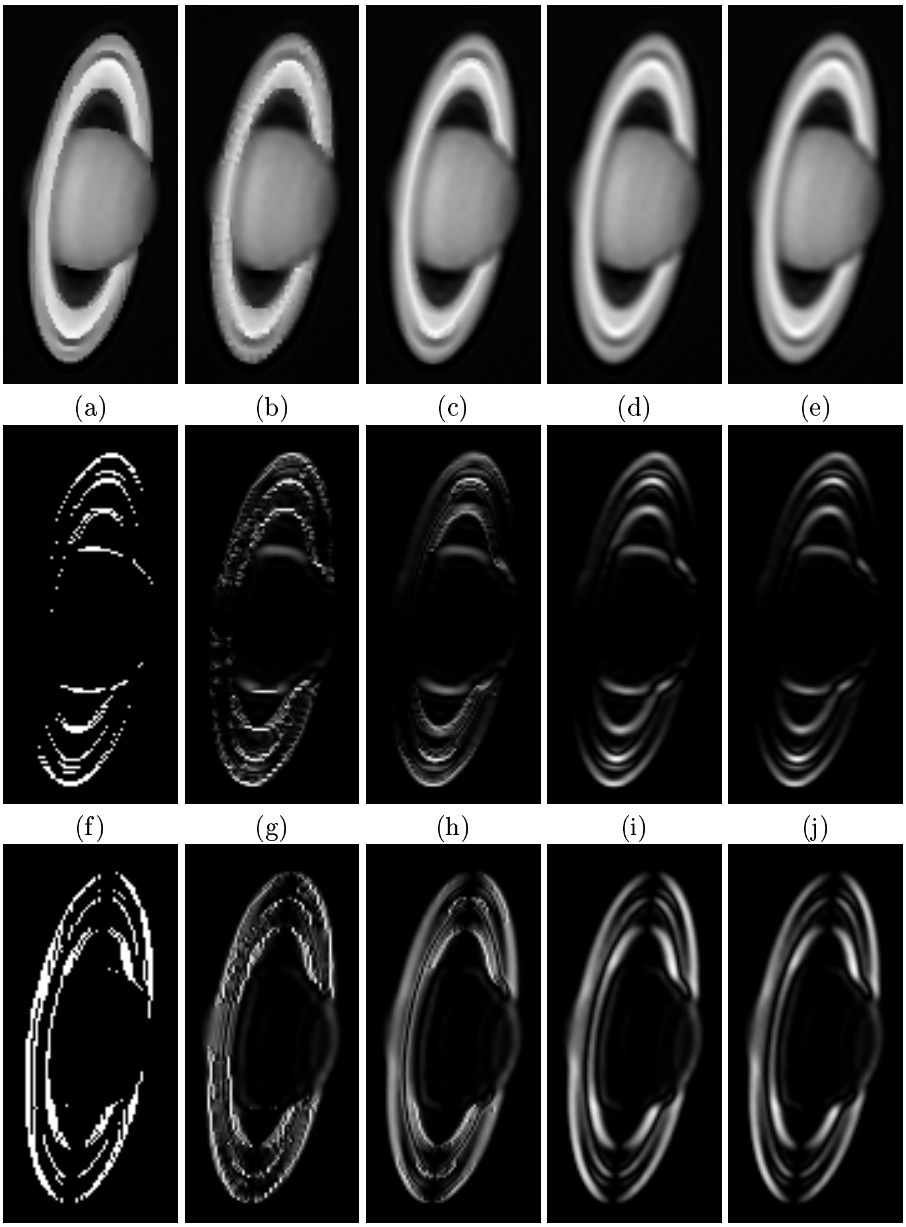
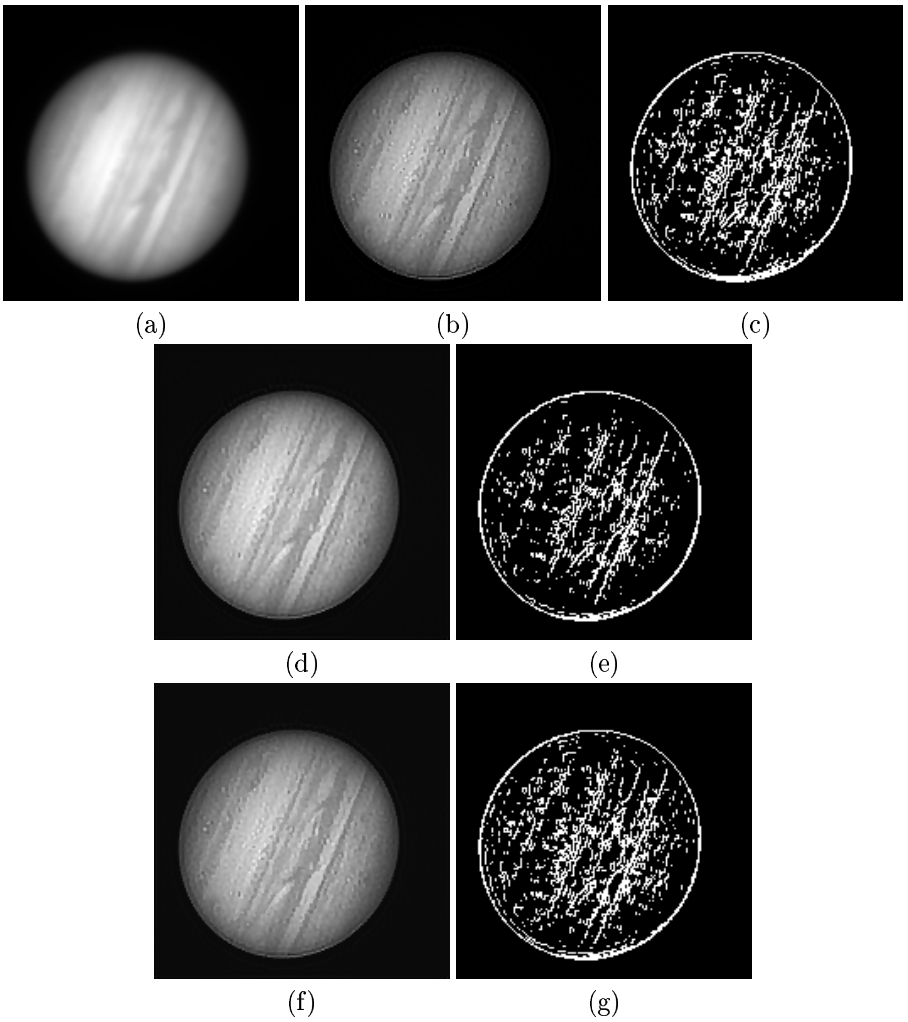


Fig. 5



(a) (b) (c)

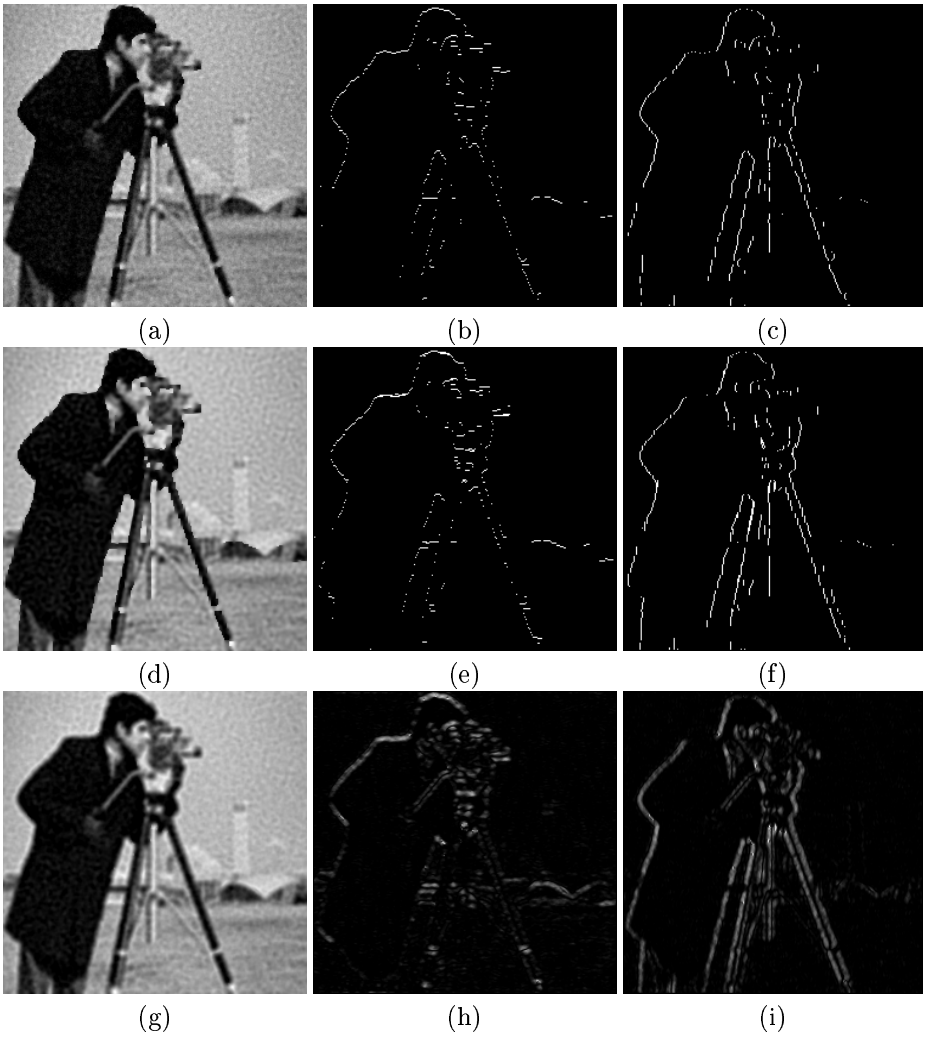
(d) (e)

(f) (g)

Fig. 6

(a) (b)

Fig. 7

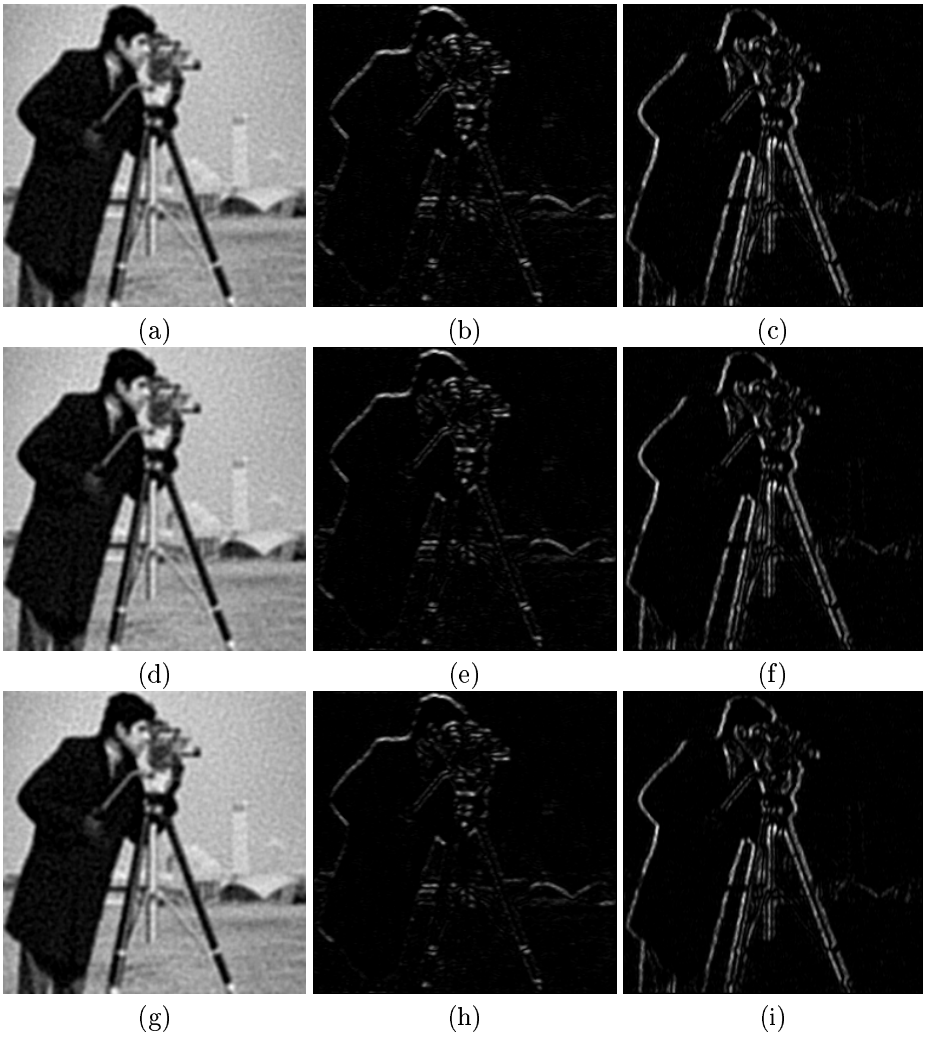


(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Fig. 8

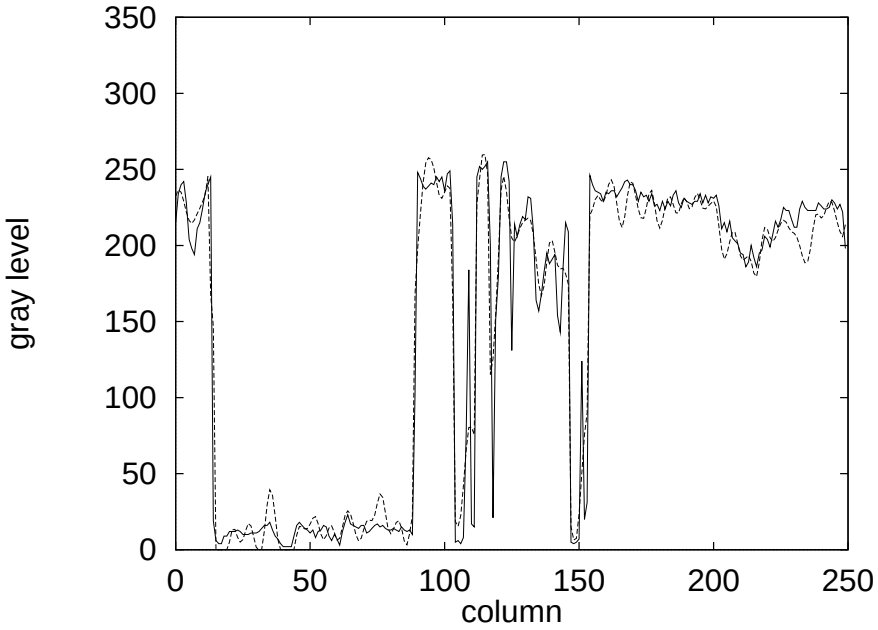


(a) (b) (c)

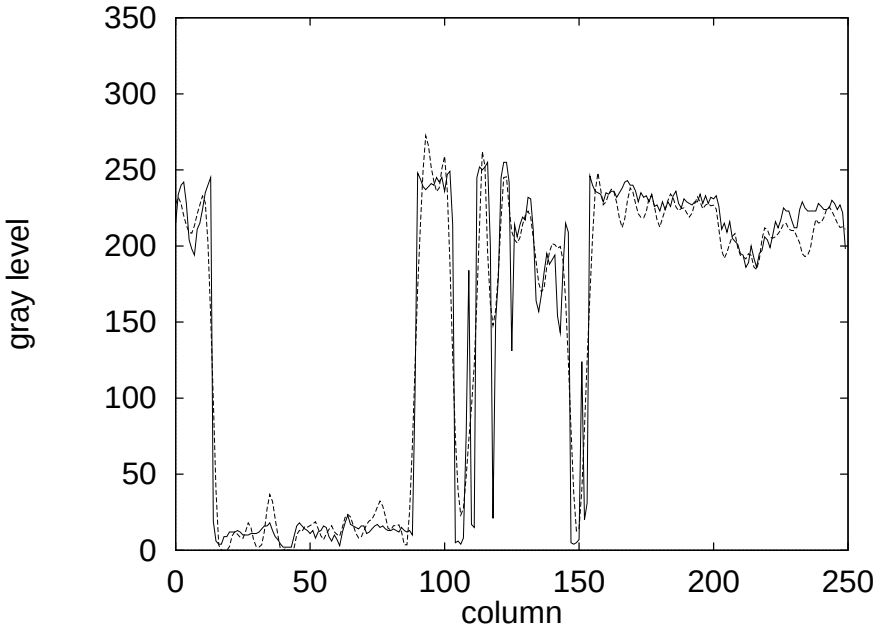
(d) (e) (f)

(g) (h) (i)

Fig. 9


gray level
column
0
50
100
150
200
250
300
350

(a)


gray level
column
0
50
100
150
200
250
300
350

(b)

Fig. 10

Potential function	φ_{GM}	φ_{HL}	φ_{HS}	φ_{GR}
Expression of $\varphi(t)$	$\frac{t^2}{1+t^2}$	$\log(1+t^2)$	$2\sqrt{1+t^2}-2$	$2\log[\cosh(t)]$

Table 1

				ARTUR with			
	Observed	Proposed ICM	Proposed SA	φ_{GM}	φ_{HL}	φ_{HS}	φ_{GR}
PRNR (dB)	18.89	20.72	21.11	20.75	20.64	20.72	20.51

Table 2

	Original		Proposed		ARTUR with			
Method	ICM	SA	ICM	SA	φ_{GM}	φ_{HL}	φ_{HS}	φ_{GR}
Total Time (sec.)	1149	12852	140	2250	198	38	29	29
Relative Time	1.00	1.13	0.12	0.20	0.17	0.17	0.17	0.17

Table 3