BAYESIAN DECONVOLUTION METHODS IN ASTRONOMY

R. Molina
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada.
18071 Granada. España.

A. K. Katsaggelos
Department of Electrical and Computing Engineering
Northwestern University.
Evanston, IL 60208.

J. Mateos
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad de Granada.
18071 Granada. España.

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 image restoration. While linear-shift invariant (LSI) models have been generally used for image restoration in Astronomy, no much work has been reported on the use of more complex models in this area. In a previous paper [1] we examined several methods within the Bayesian paradigm to perform image restoration in Astronomy. Here, we describe the use of Compound Gaussian Markov Random Fields (CGMRF), a non LSI model that preserves image discontinuities, to restore astronomical images. Problems on the application of the model arising from the high dynamic range and severe blurring of astronomical images are addressed and new methods to estimate the real underlying image based on stochastic relaxation and mean field approximation are proposed.

1 Introduction

During recent years several techniques for the deconvolution of observational *psf*'s have been developed based on different algorithms, linear or non-linear. So far, the most used method for restoration of HST images seems to be the Richardson-Lucy (R-L) method, although maximum entropy methods, in particular the MemSys package, are also frequently used in the astronomical community (see the issue edited by Núñez [2] for an excellent overview of image reconstruction and restoration in astronomy).

In this paper we concentrate on the use of CGMRF, a non LSI model that preserves image discontinuities, to restore astronomical images. The CGMRF theory provides us a way to control the change on the image model using a hidden random field. The use of the underlying random field, called line

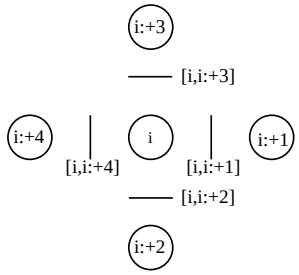

Figure 1: Image and line sites

process, was introduced by Geman and Geman[3]. Given the image and the 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. In section 2 we introduce the used notation and the proposed model for the image and line processes as well as the noise models. Both, stochastic and deterministic relaxation approaches to obtain the MAP estimate are presented in section 3. Algorithms based on the mean field approximation are described in section 4. Finally, in section 5 we test those algorithms on real astronomical images.

2 Notation and Model

Bayesian methods start with a prior distribution, a probability distribution over images $\mathbf{f}$, $p(\mathbf{f})$ (it is here that we incorporate information on the expected structure within an image). It is also necessary to specify the probability distribution $p(\mathbf{g}|\mathbf{f})$, of observed images $\mathbf{g}$ if $\mathbf{f}$ were the true image. The Bayesian paradigm dictates that inference about the true $\mathbf{f}$ should be based on $p(\mathbf{f}|\mathbf{g})$ given by

$$p(\mathbf{f}|\mathbf{g}) \propto p(\mathbf{f})p(\mathbf{g}|\mathbf{f}).$$

To show just one restoration, it is common to choose the mode of this posterior distribution, that is, to display the image $\hat{\mathbf{f}}$ which satisfies

$$\hat{\mathbf{f}} \text{ maximizes } p(\mathbf{f})p(\mathbf{g}|\mathbf{f}).$$

This is known as the maximum *a posteriori* estimate of $\mathbf{f}$.

Let us now introduce the CGMRF image model. The idea is to build a prior model consisting of two processes, one accounting for the intensity values and the other for the location of edges in the image. Its probability distribution

is given by

$$-\log \mathrm{p}(\mathbf{f},\mathbf{l}) = \mathrm{const} + \sum_i \left[\phi(\mathbf{f}_i - \mathbf{f}_{i:+1})^2(1-\mathbf{l}_{<i,i:+1>}) + \beta\mathbf{l}_{<i,i:+1>}\right.$$
$$\left. + \phi(\mathbf{f}_i - \mathbf{f}_{i:+2})^2(1-\mathbf{l}_{<i,i:+2>}) + \beta\mathbf{l}_{<i,i:+2>} + (1-4\phi)\mathbf{f}_i^2\right]/2\sigma_{\mathrm{w}}^2 \quad (1)$$

where $i{:}+1$, $i{:}+2$, $i{:}+3$, $i{:}+4$ denote the four pixels around pixel i as described in Fig. 1, ϕ just less than 0.25 and the line process is introduced by simply defining the function $\mathbf{l}_{<i,j>}$ as taking the value zero if pixels i and j are not separated by an active line and one otherwise (see Fig. 1). We penalize the introduction of the line element by the term $\beta\mathbf{l}_{<i,j>}$ since otherwise the expression in Eq. 1 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.

Let us now describe the noise model. The main sources of noise are the discrete nature of photons and electronic noise in the CCD detector. Physical considerations suggest that this noise is independent from pixel to pixel and depending on the average number of photons arriving at the CCD cell. A simplified noise model 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 $\mathbf{g}_i = (\mathbf{D}\mathbf{f})_i + \mathbf{n}_i = \sum_j d_{i-j}\mathbf{f}_j + n_i$, where $\mathbf{D}$ is the $p \times p$ matrix defining the systematic blur, assumed to be known and approximated by a block circulant matrix, $\mathbf{n}_i$ 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 $\mathbf{g}$ if $\mathbf{f}$ were the 'true' image is

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

Although in this paper we use this Gaussian noise model, the theory we develop here can be easily extended to the Poisson noise model.

3 Stochastic and Deterministic Relaxation for Estimating the MAP

Let us now proceed to find $\hat{\mathbf{f}}$, $\hat{\mathbf{l}}$, the MAP estimate of $\mathbf{f}$ and $\mathbf{l}$, that is

$$\hat{\mathbf{f}},\hat{\mathbf{l}} = arg\max_{\mathbf{f},\mathbf{l}} \mathrm{p}(\mathbf{f},\mathbf{l} \mid \mathbf{g}). \quad (3)$$

Since $\mathrm{p}(\mathbf{f},\mathbf{l} \mid \mathbf{g})$ is nonlinear it is extremely difficult to find $\hat{\mathbf{f}}$ and $\hat{\mathbf{l}}$ by any conventional method. The relaxation technique to search for MAP estimates

we are going to propose, which is called simulated annealing (SA), uses the distribution

$$\begin{aligned} \mathrm{p}_T(\mathbf{f},\mathbf{l}\mid\mathbf{g}) = \frac{1}{Z_T}\exp\left\{-\frac{1}{T}U(\mathbf{f},\mathbf{l}|\mathbf{g})\right\} = \frac{1}{Z_T}\exp\left\{-\frac{1}{T}\frac{1}{2\sigma_{\mathrm{n}}^2}\left\|\mathbf{g}-\mathbf{D}\mathbf{f}\right\|^2\right. \\ -\frac{1}{T}\sum_i\left[\phi(\mathbf{f}_i-\mathbf{f}_{i:+1})^2(1-\mathbf{l}_{<i,i:+1>})+\beta\mathbf{l}_{<i,i:+1>}\right. \\ \left.\left.+\ \phi(\mathbf{f}_i-\mathbf{f}_{i:+2})^2(1-\mathbf{l}_{<i,i:+2>})+\beta\mathbf{l}_{<i,i:+2>}+(1-4\phi)\mathbf{f}_i^2\right]/2\sigma_{\mathrm{w}}^2\right\}, \quad (4) \end{aligned}$$

where T is the temperature and Z_T is a normalization constant.

Then the classical simulated annealing algorithm proceeds as follows.

Algorithm 1 Sequential SA procedure.

1. *Set* $T=1$. *Set* $k=0$.
2. *Let* u *denote the sites in the line process. For each site,* u, *simulate its posterior probability distribution,* $\mathrm{p}_T(\mathbf{l}_u \mid$ *rest of* $\mathbf{l},\mathbf{f},\mathbf{g})$.
3. *For each pixel* i *update* $\mathbf{f}_i$ *by simulating* $\mathrm{p}_T(\mathbf{f}_i \mid$ *rest of* $\mathbf{f},\mathbf{l},\mathbf{g})$.
4. *Set* $k=k+1$. *Decrease the temperature* T *and go back to the step 2 until* k *is greater than a specified number.*

Note that the temperature is decreased only after a full sweep of the line process and the picture (after going through all lines and pixels and update them according to the algorithm). Conditions for the algorithm to converge to the MAP are established in Geman and Geman[3] and Chellapa *et al.* [4].

Hence, we have to simulate the conditional *a posteriori* density function for $\mathbf{l}_{<i,j>}$, given the rest of $\mathbf{l}$, $\mathbf{f}$ and $\mathbf{g}$ and the conditional *a posteriori* density function for $\mathbf{f}_i$ given the rest of $\mathbf{f}$, $\mathbf{l}$ and $\mathbf{g}$. Lets us now examine these distributions.

In order to simulate the line process conditional *a posteriori* density function, $\mathrm{p}_T(\mathbf{l}_{<i,j>} \mid$ *rest of* $\mathbf{l},\mathbf{f},\mathbf{g})$, we have

$$\mathrm{p}_T(\mathbf{l}_{<i,j>}=0 \mid \textit{rest of } \mathbf{l},\mathbf{f},\mathbf{g}) \quad\propto\quad \exp\left[-\frac{1}{T}\frac{\phi}{2\sigma_{\mathrm{w}}^2}(\mathbf{f}_i-\mathbf{f}_j)^2\right], \quad (5)$$

$$\mathrm{p}_T(\mathbf{l}_{<i,j>}=1 \mid \textit{rest of } \mathbf{l},\mathbf{f},\mathbf{g}) \quad\propto\quad \exp\left[-\frac{1}{T}\frac{\beta}{2\sigma_{\mathrm{w}}^2}\right]. \quad (6)$$

Furthermore, for Gaussian noise, $\mathrm{p}_T(\mathbf{f}_i \mid$ *rest of* $\mathbf{f},\mathbf{l},\mathbf{g}) \sim \mathcal{N}(\mu_i,\sigma_i^2)$, where μ_i and σ_i are given by

$$\mu_i = \lambda_i \sum_{j\ \mathrm{nhbr}\ i}\phi\frac{\mathbf{f}_j(1-\mathbf{l}_{<i,j>})}{nn_i}+(1-\lambda_i)\left(\frac{(\mathbf{D}^T\mathbf{g})_i-(\mathbf{D}^T\mathbf{D}\mathbf{f})_i}{c}+\mathbf{f}_i\right), \quad (7)$$

$$\sigma_i^2 = \frac{T\sigma_{\rm w}^2\sigma_{\rm n}^2}{nn_i\,\sigma_{\rm n}^2 + c\,\sigma_{\rm w}^2}, \tag{8}$$

where c is the sum of the square of the coefficients defining the blur function, that is, $c = \sum_j d_j^2$, $nn_i = \sum_{j\,\mathrm{nhbr}\,i} \phi(1 - \mathbf{l}_{<i,j>}) + (1 - 4\phi)$ and $\lambda_i = nn_i\,\sigma_{\rm n}^2/(nn_i\,\sigma_{\rm n}^2 + c\,\sigma_{\rm w}^2)$.

Unfortunately, classical simulated annealing for the Gibbs sampler may not converge for high dynamic range and severe blurring [5]. In particular, no convergence proof of algorithm 1 in presence of blurring has been given.

To solve this problem we use the value of $\mathbf{f}_i$ obtained in the previous iteration, $\mathbf{f}_i^{old}$, and, in algorithm 1, instead of simulating from the normal distribution defined in Eq. 7 and Eq. 8 to obtain the new value of $\mathbf{f}_i$, we simulate from the normal distribution with mean μ_i' and variance $\sigma_i^{2\prime}$ defined as [5]

$$\mu_i' = \omega_i \mathbf{f}_i^{old} + (1 - \omega_i)\mu_i, \tag{9}$$

$$\sigma_i^{2\prime} = (1 - \omega_i^2)\sigma_i^2, \tag{10}$$

where $\omega_i = ((1 - nn_i)\sigma_{\rm n}^2 + (1 - c)\sigma_{\rm w}^2)/(\sigma_{\rm n}^2 + \sigma_{\rm w}^2)$. We also update the whole image at the same time.

Instead of using a stochastic approach, we can use a deterministic method, called iterative conditional mode (ICM), 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 instead. The disadvantage is the local nature of the solution obtained. This method can be seen as a particular case of simulated annealing where the temperature is always set to zero.

In the test example section we apply the following modified SA and ICM algorithms, whose convergence have been established to restore astronomical images [6].

Algorithm 2 Sequential Modified SA procedure

1. *Set $T = 1$. Set $k = 0$.*
2. *Let u denote the sites in the line process. For each site, u, simulate its posterior probability distribution,* $\mathrm{p}_T(\mathbf{l}_u \mid \textit{rest of } \mathbf{l}, \mathbf{f}, \mathbf{g})$.
3. *The evolution $\mathbf{f}_{k-1} \rightarrow \mathbf{f}_k$ of the observed system can be obtained by sampling in parallel the next value of the whole image based on the conditional probability mass function defined in Eq. 9 and Eq. 10.*
4. *Set $k = k + 1$. Decrease the temperature T and go back to the step 2 until k is greater than a specified number.*

The modified ICM procedure is obtained by selecting, in steps 2 and 3 of algorithm 2, the mode of the corresponding transition probabilities.

4 Mean Field Annealing

To find the MAP we can also use the mean field approximation (MFA) (see Chandler [7] for a clear treatment of the mean field theory). The mean field theory concerns the estimation of $E[\mathbf{x}_i|\mathbf{g}]$. By definition

$$E[\mathbf{x}_i|\mathbf{g}] = \sum_{\mathbf{x}} \mathbf{x}_i p(\mathbf{x}|\mathbf{g}) = \frac{1}{Z_T} \sum_{\mathbf{x}} \mathbf{x}_i \exp[-\frac{1}{T} U(\mathbf{x}|\mathbf{g})], \tag{11}$$

where $\mathbf{x} = (\mathbf{f}, \mathbf{l})$ and $U(\mathbf{x}|\mathbf{g})$ is defined in Eq. 4.

However, it is well known that the calculation of Z and the sum above involve all possible realizations of the MRF which is complex and computationally not feasible.

The mean field theory suggests an approximation to Eq. 11 based on the following assumption: the influence of $\mathbf{x}_j, j \neq i$, in the calculation of $E[\mathbf{x}_i|\mathbf{g}]$ can be approximated by the influence of $E[\mathbf{x}_j|\mathbf{g}]$.

Using the mean field approximation we have for our problem that $E[\mathbf{f}_i|\mathbf{g}]$ is given by Eq. 7 and

$$E[\mathbf{l}_{<i,j>}|\mathbf{g}] = \frac{1}{1 + \exp(-\frac{1}{T}(\frac{\phi}{2\sigma_{\mathbf{w}}^2}(\mathbf{f}_i - \mathbf{f}_j)^2 - \frac{\beta}{2\sigma_{\mathbf{w}}^2}))}.$$

The mean field is usually computed iteratively, to find the mean field at i, the mean field at the neighbors of i is needed. However, in our iterative method, we do not use Eq. 7 because the matrix defining the iterative procedure is not a contraction matrix and use instead the mean defined in Eq. 9 as the iterative procedure. Finally, $\frac{1}{T}$ can be fixed or increased gradually (annealing) in the iterations of the mean field calculations.

5 Results

The proposed methods were 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.

The blurring function, $\mathbf{D}$ can be approximated by $d_i \propto (1 + i^2/R^2)^{-\delta}$. 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 2 depicts the original image and the restorations after running our proposed ICM, SA and MFA method. 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 proposed ICM and MFA method show thick lines. The proposed SA method, on the other hand, gives us thinner lines and the details are more resolved. However, it needs more iterations and is slower than the other two methods.

Acknowledgments

This work has been supported by the "Comisión Nacional de Ciencia y Tecnología" under contract PB93-1110 and the ESF network Converging Computing Methodologies in Astronomy.

References

1. R. Molina, J. Mateos, J. Abad, N. Pérez de la Blanca, A. Molina and F. Moreno, *Int. J. of Imaging Systems and Technology*, **4**, 370 (1995).
2. J. Núñez (Ed.), *Int. J. of Imaging Systems and Technology*, **4** (1995).
3. S. Geman and D. Geman, *IEEE Trans. on PAMI*, **6**, 721 (1984).
4. R. Chellapa, T. Simchony and Z. Lichtenstein in *Digital Image Restoration*, ed. Katsaggelos, A.K. (Springer Series in Information Science, Springer–Verlag, 1991).
5. R. Molina, A.K. Katsaggelos, J. Mateos and J. Abad, *Proceeding of ICIP-96*, 469, (1996).
6. R. Molina, A.K. Katsaggelos, J. Mateos and A. Hermoso, submitted to *EMMCVPR'97*, (1997).
7. D. Chandler, in *Introduction to Modern Statistical Mechanics* (Oxford University Press, 1987).

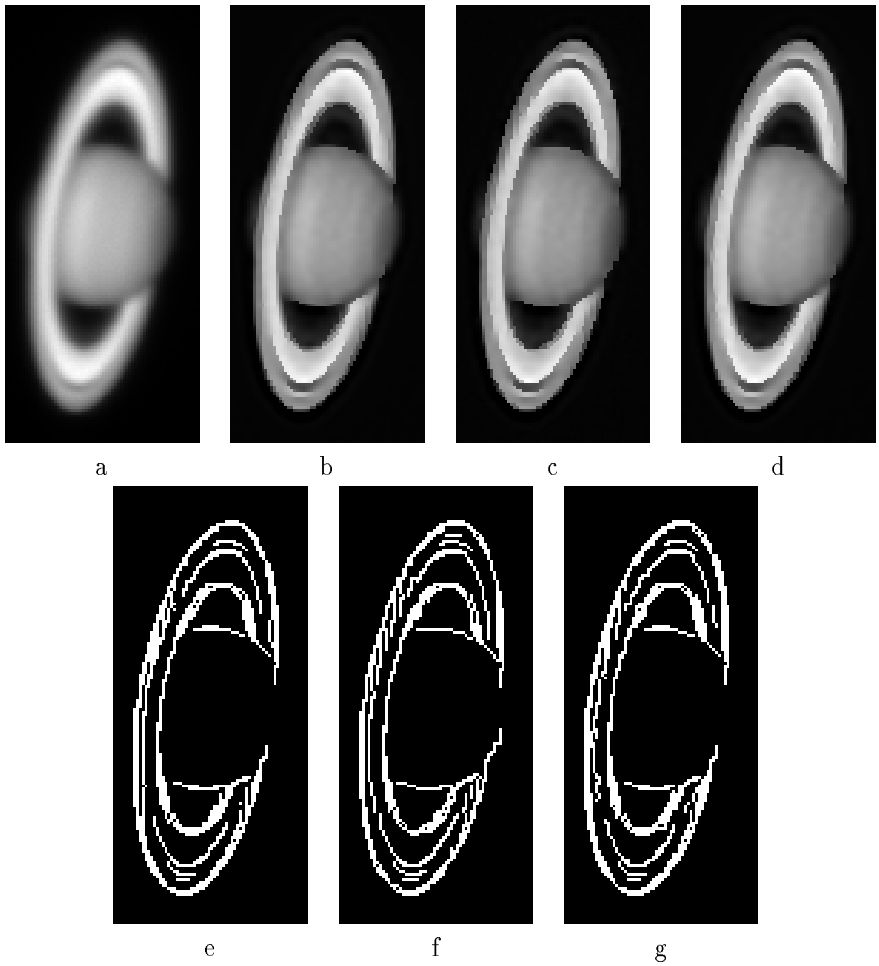

Figure 2: (a) Original image. Restoration: (b) proposed ICM method, (c) proposed SA method and (d) proposed MFA method. Line Process: (e) proposed ICM method, (f) proposed SA method and (g) proposed MFA method.