

# Bayesian MAP Restoration of Scintigraphic Images

Mai. K. Nguyen<sup>1</sup>, Hervé Guillemin<sup>2</sup> and Patrick Duvaut<sup>1</sup>

<sup>1</sup>Equipe Traitement des Images et du Signal (CNRS URA 2235- ENSEA- Université de Cergy-Pontoise), Ecole Nationale Supérieure de l'Electronique et ses Applications, 6 Av. du Ponceau, 95014 Cergy, France  
<sup>2</sup>Laboratoire de Biophysique, Faculté de médecine de Besancon, 2 Place St Jacques, 25000 Besancon, France.

## Abstract

We are interested in the problem of restoring scintigraphic images acquired by the gamma detector in nuclear medicine. The aim is to improve the detectability of possible heterogeneous areas in different organs. We propose to solve the problem in the Bayesian framework with the Maximum *A Posteriori* (MAP) principle. The prior information was modeled by a Markov Random Field (MRF). The optimization is based on two kinds of methods: the stochastic algorithm of simulated annealing with a Gibbs sampler, and the deterministic algorithm of Graduated Non Convexity (GNC). We compared the results to the images restored by the Metz filter, more classical in this field. We applied these methods to the restoration of cold or warm nodules in the thyroid gland. We noticed the superiority of the proposed methods in terms of contrast around the nodules and uniformity in the images.

## 1. Introduction

The scintigraphic images have very poor quality. The small heterogeneous areas are detected only if their size is up to a threshold depending on the used gamma camera. The main degradations are due to [1]:

- The Poisson noise.
- The bad spatial resolution of the detector, which depends on the distance (source-detector).
- The bad energy resolution of the detector, which prevent us from distinguishing primary and scattered photons, near the photoelectric peak. The detection of scattered photons causes a non-stationary spatial blur in the image.

The last two problems above can be approximately taken into account by using a Gaussian spatial impulse response where variance is estimated in function of the used gamma camera.

In the case of the thyroid gland, the spatial impulse response is supposed to be independent of the distance (source-detector). This hypothesis is justified by noting that this organ is relatively flat, so that all the points inside it are located at a distance approximately constant from the camera.

The image formation model is a linear Poisson one. Let us note  $\mathbf{x}$  the vector whose components are obtained by scanning the researched image, and  $\mathbf{C}$  the convolution matrix. The number  $Y_i$  of photons detected at pixel  $i$  in the observed image is a Poisson random variable of parameter  $E[Y_i]=m_i$  where  $m_i$  is the average number of photons detected at pixel  $i$  by the gamma camera:

$$E[Y_i]=m_i = \sum_{j=0}^{N^2-1} C_{ij} \cdot x_j = (\mathbf{C} \cdot \mathbf{x})_i \quad (1)$$

Each component  $C_{ij}$  of the  $\mathbf{C}$  matrix can be considered as the probability that a photon emitted at real position  $j$  in the researched image is detected at position  $i$  by the gamma-camera.

The convolution matrix  $\mathbf{C}$  in (1) is ill conditioned. The restoration of scintigraphic images is typically an ill posed inverse problem [2]. Its resolution goes through the introduction of prior information concerning the researched object. The methods that we use all offer a regularized solution to the problem: the Metz filter and the Bayesian MAP restoration.

## 2. Restoration by the Metz Filter

The expression of the Metz filter [3], [4] is based on another writing of the inverse filter :

$$\frac{1 - (1 - \|H\|^2)^p}{H} \quad (2)$$

$H$  is the Fourier transform of the spatial impulse response. This filter acts as the inverse filter to a certain frequency, then as a smoothing filter. The order  $p$  of the filter defines the cutting frequency and enables us to reach a compromise between a solution very regular but too far from the observed data, and a solution close to the observed data but in which the noise is increased. This regularization method is not entirely satisfactory because the image is globally smooth and the possible local discontinuities in the activity are reduced.

## 3. Bayesian MAP Method

The Bayesian MAP method allows us to regularize the solution by introducing prior information about the researched object [5]. Defining the prior as a MRF permits to model images of different natures. The discontinuities in the image can then be preserved in judiciously choosing the local interaction function associated with the MRF.

The MAP method presumes that the observed image  $\mathbf{y}$  and the researched image  $\mathbf{x}$  are two specific realizations of two random fields  $X$  and  $Y$  with which the density of probability  $p(X=\mathbf{x})$  and  $p(Y=\mathbf{y})$  are associated [6]. The restored image is the one that maximizes the posterior probability from the Bayes' formula following:

$$p(X=\mathbf{x}|Y=\mathbf{y})=p(Y=\mathbf{y}|X=\mathbf{x}) \cdot p(X=\mathbf{x})/p(Y=\mathbf{y}) \quad (3)$$

Noting that the term  $p(Y=y)$  does not depend on  $\mathbf{x}$ , the solution is obtained by maximizing the product of two terms :

$$p(Y=y|X=\mathbf{x}).p(X=\mathbf{x}) \quad (4)$$

The first one,  $p(Y=y|X=\mathbf{x})$  guarantees fidelity of the restored data to the observed ones. It only depends on the image formation model :

$$p(Y = \mathbf{y}|X = \mathbf{x}) = \prod_{i=0}^{N^2-1} \frac{(\mathbf{C}.\mathbf{x})_i^{y_i}}{y_i!} \cdot \exp(-(\mathbf{C}.\mathbf{x})_i) \quad (5)$$

The second term  $p(X=\mathbf{x})$ , reflects our prior knowledge concerning the researched object, and permits to obtain a regularized solution. We modeled  $X$  field by a MRF, the probability density  $p(X=\mathbf{x})$  then takes the following form :

$$p(X=\mathbf{x}) = 1/Z \cdot \exp[-U(\mathbf{x})] \quad \text{with} \quad U(\mathbf{x}) = \sum_{c \in \zeta} \alpha V_c(\mathbf{x}) \quad (6)$$

where  $U(\mathbf{x})$  is the *a priori* energy,  $V_c$  is the local interaction function associated to the Markov field,  $Z$  is a normalizing constant,  $\zeta$  is the set of all cliques associated with a neighborhood.

Since the images usually treated involve homogeneous and transitions areas, the function  $V_c$  is defined as [5].

$$V_c(s) = \frac{s^2}{s^2 + 3.\delta^2} \quad (7)$$

where  $s$  symbolizes numerical approximation of horizontal and vertical gradient in each site pixel  $i$  (in the case of the first-order neighborhood),  $\delta$  is a constant linked with the transition threshold. The maximum of posterior probability is then reached by minimizing the following function:

$$F(\mathbf{x}) = \sum_{i=0}^{N^2-1} (-y_i \cdot \ln((\mathbf{C}.\mathbf{x})_i) + (\mathbf{C}.\mathbf{x})_i) + \alpha \cdot \sum_{c \in \zeta} V_c(\mathbf{x}) \quad (8)$$

The first term insures fidelity of restored data to observed ones, and the second to prior information.  $\alpha$  plays the role of a regularization parameter.

### 3.1 Stochastic Optimization

The function (8) is a non-convex function of  $\mathbf{x}$  since its hessian is a non-definite-positive matrix. We minimized it by a stochastic relaxation method that theoretically guarantees to converge until the global minimum. Relaxation stochastic methods [5], [7], [8] proceed by local changes that do not necessary imply a decreasing of the posterior energy. This feature makes them reach, under some assumptions, the global minimum of energy without being trapped to local minima. Stochastic methods of simulated annealing rely on the introduction of a new parameter  $T$  comparable with a temperature:

$$\frac{1}{T} \cdot \left[ \sum_{i=0}^{N^2-1} (-y_i \cdot \ln((\mathbf{C}.\mathbf{x})_i) + (\mathbf{C}.\mathbf{x})_i) + \alpha \cdot \sum_{c \in \zeta} V_c(\mathbf{x}) \right] \quad (9)$$

Their principle consists of choosing a random sequence of configurations according to a Gibbs distribution of energy (8)

while making parameter  $T$  decrease from a high temperature until another one, close to 0. The posterior Gibbs field converges to the uniform distribution over the set of minimal energy configurations as  $T$  tends to 0, and to the uniform distribution over the set  $\Omega$  of all the configurations allowed when  $T$  tends to infinity [8]. The annealing schedule (i.e. the way the temperature decreases), which insures to converge to the configuration of minimal energy, is chosen like [5] :

$$T(ni) = \frac{\tau}{\ln(1 + ni)} \quad (10)$$

Where  $T(ni)$  is the temperature at iteration  $ni$ ,  $\tau$  is constant.

[5] showed that such a schedule guarantees convergence to the global minimum of the posterior energy, under a certain condition on the parameter  $\tau$ .

For each temperature  $T$ , the random choice of a new configuration  $\hat{\mathbf{x}}$  according to the posterior law can't be driven directly because the number of configurations in the space of all the allowed configurations is too important. The most commonly used approaches to make random choice appeal to Markov chains and some of their properties. The idea of these methods consists in defining an ergodic Markov chain for which the Gibbs distribution represents the unique invariant distribution. We used Gibbs sampler that build a Markov chain with transition probabilities depending on the local characteristics of the posterior Gibbs distribution. The Gibbs sampler builds a Markov chain with (8) as equilibrium distribution. The introduction of the parameter  $T$ , associated with the annealing schedule (10) enables us to get random choice uniformly on the set of minimal posterior energy configurations.

The main drawback of this method is very high calculation cost. A variety of methods based on the expectation maximization (EM) have also been proposed. The slow convergence of EM is well documented in [9].

We introduce a deterministic relaxation algorithm inspired by the GNC proposed first by Blake and Zisserman in [10] for segmentation and extended to the ill-posed inverse problem in [11], [12].

### 3.2 Optimization with GNC algorithm

The principle of this algorithm consists in approximating a non convex function  $F(x)$  with a sequence of continuously derivable functions  $F_r(x)$  converging toward it while  $r$  tends to infinity, by taking care to choose the first  $F_o(x)$  to be convex (Fig. 1).

As far as we are concerned, the succession of functions is obtained by making the parameter  $\delta$  in (7) following a geometric sequence of powers of two from  $\delta_0$  until  $(2^r \cdot \delta_0)$ . The first value  $\delta_0$  is chosen so that the hessian of the posterior energy is positively definite. As the hessian of the first term of (8) is positively definite [1], a sufficient condition for the posterior energy to be convex, is that all the terms in the diagonal hessian matrix of the second term in (8), are strictly positive. This condition is satisfied if all the values  $s$  are within the interval  $]-\delta, \delta[$  on which  $V_c$  is convex. For the images usually processed, one can suppose that all the absolute values of  $s$  are lower than a

bound. Unfortunately, this bound is unknown, and we estimated it by the maximum value  $m$  of the gradient in the observed image.

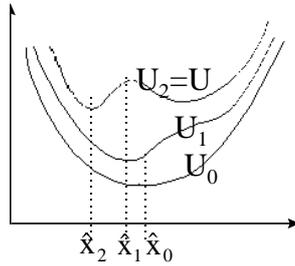


Figure 1 : principle of the GNC algorithm

Contrary to the method of simulated annealing, the GNC algorithm does not insure to converge until the global minimum of the function (8). However, the GNC algorithm always permits to reach the same solution, which remains close to the global minimum. Moreover, the solution is not very sensitive to the choice of the relaxation scheme and to the regularization parameter. In practical experiments, a satisfactory solution is reached for much shorter computation time.

#### 4. Estimation of Hyperparameters

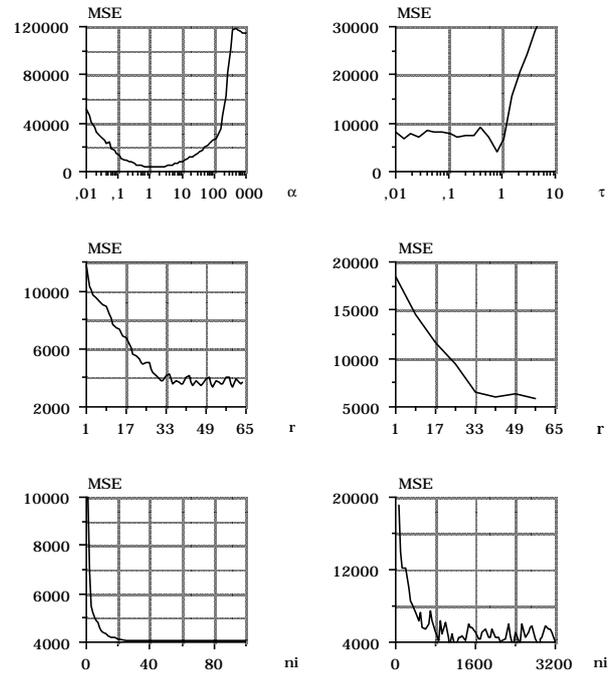
We determined the optimal values of parameters on synthetic data analytically simulated by convolution with a Gaussian impulse response and corrupted by Poissonian noise. The width at half of the maximum (WHM) of the impulse response is 9 pixels (standard deviation = 3.82 pixels). The noise is simulated approximately by noting that a Poisson law with parameter  $n$  converges to a Gaussian law with mean and variance  $n$  as  $n$  increases. The corresponding researched data are mono-dimensional signals of 64 pixels including a central area of 20 pixels of activity  $\omega$  constant in the middle of an inactive area. We made a succession of tests with  $\omega$  varying between 10 and 200 by step of 10. This choice enables to cover most of experiments in clinical routine for the acquisition times and the activity of the radioactive sources usually used. For each value of  $\omega$ , 50 acquisitions of the same object were simulated. For every simulated image, we optimized separately the parameters  $\alpha$  or  $\tau$ ,  $\delta$  and the number of iterations  $ni$ , each separately, while keeping the other constant. The quality of restoration is evaluated by measuring the mean square error between the restored and the original images.

	$\alpha$	$\delta$	$\tau$	$ni$
<b>GNC</b>	1	$m/2^5$		40 maxi
<b>Gibbs</b>	1	$m/2^3$	1	1500

Chart 1: optimized hyper parameters (criterion : mean square error);  $m$  is the maximum of the gradient in the observed image.

The results are presented on figure 2 with  $\omega=100$ . The retained parameters are those for which mean square error becomes minimal (MMSE). Chart 1 sums up the set of the optimal parameters with the MMSE criterion. We only

notice that the optimal parameters barely depend on the activity  $\omega$  in the central area.



line 1	$\delta(r=41)$ ; $\alpha$ varies; $ni=40$ Iterations	$\delta(r=33)$ ; $\alpha=1$ ; $\tau$ varies; $ni=1500$ Iterations
line 2	$\delta(r)$ varies; $\alpha=1$ ; $ni=40$ Iterations	$\delta(r)$ varies; $\alpha=1$ ; $\tau=1$ ; $ni=1500$ Iterations
line 3	$\delta(r=41)$ ; $\alpha=1$ ; Iteration varies	$\delta(r=33)$ ; $\alpha=1$ ; $\tau=1$ ; Iteration varies

Figure 2 : deterministic (left) and stochastic (right) algorithms : estimation of hyper-parameters (criterion : mean square error);  $\delta=m \cdot 2^{-(r-1)/8}$ ;  $m$  is the maximum of the gradient in the observed image.

#### 5. Simulation

We simulated the acquisition of two images by Monte-Carlo method. The first one contains a cylindrical cold nodule 12mm in diameter located at the bottom of a cylinder filled with radio active liquid. The second contains two cold nodules 20mm in diameter located at 6mm from one to another, inside a radio active cylinder. The WHM of the impulse response is about 23 mm. We measured the contrast at the nodules, uniformity in the area of constant activity and Kaufmann distance (only for the first model) which informs us about steep feature of the nodule edges. We present results on chart 2.

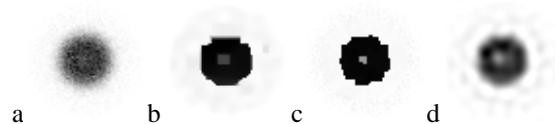


Figure 3 : simulated data with an 12mm diameter nodule. a: observed data-b :GNC -c: stochastic algorithm-d:Metz.

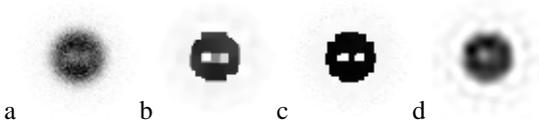


Figure 4 : simulated data with 2 nodules distant to 6mm  
a: observed data-b :GNC -c: stochastic algorithm-d:Metz.

	2 nodules model		1 nodule model		
	Contrast	Uniformity	Contrast	Uniformity	Kaufmann
Metz	0.34	45.98	0.28	66.01	0.61
GNC	0.50	5.38	0.14	5.48	0.25
Gibbs	0.94	0.07	0.41	0.32	0.04

Chart 2 : performance measures for each method for simulated data.

The MAP method reconstructs the nodules with a slightly greater contrast than Metz filtering. The superiority of the MAP method is especially emphasized by the other two measures : the variance in the constant activity areas is much lower in the MAP restored images than those treated by the Metz filter. The uniform areas indeed appear much more regularly according to the prior information. Contrary to the Metz filter, this feature does not exclude the preservation of clean edges as proved by Kaufmann feature (values much closer to 0 for the MAP method than for the Metz filter).

Concerning two optimization algorithms for MAP restoration, we only note that the differences of performance are very slight. On the contrary, the computation time is much greater in the case of stochastic optimization than in the case of the GNC algorithm. It is thus advised to use the GNC algorithm in clinical routine.

Then we restored the acquired image of a conical frustum phantom with 2 cylindrical nodules 10mm in diameter located at 5mm from one to another. The WHM of the impulse response is about 9 mm. The restored images appear on Fig. 5.

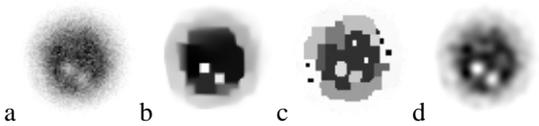


Figure 5 : acquired data of a conical frustum phantom.  
a:observed data-b :GNC -c: stochastic algorithm-d:Metz.

After making a sequence of tests for different sizes of nodules and different distances between nodules, we showed [1] that the images presented in this paper correspond to the limit size and distance of detectability for the impulse response given.

## 6. Conclusion

We have presented a new method of scintigraphic image restoration based on the Bayesian framework and the MRF prior model. The optimization was carried out using two different methods : stochastic and GNC algorithms. The hyper-parameters were calculated by drawing a series of tests which cover most experiences in clinical routine. We applied the different methods of restoration and the different algorithms of optimization in the case of the thyroid gland. The difference of the Metz filter is that the MAP method enables the regularization of solutions while preserving discontinuities. This feature improves the detectability of nodules. The proposed methods can also be used for other kinds of images, in simply making a judicious choice for the local interaction function associated with the MRF.

## 7. References

- [1] H. Guillemin, « Amélioration de la résolution spatiale des images scintigraphiques de médecine nucléaire. Application à la glande thyroïde », Thèse de Sciences, Université de Cergy-Pontoise, 1998.
- [2] G. Demoment, « Reconstruction and restoration: overview of common estimation structures and problems, » IEEE Trans. Acoust., Speech and Signal Processing, vol. 37, pp. 2024-2035, Dec. 1989.
- [3] T. A. Inuma and T. Nagai, « Image restoration in radioisotope imaging systems, » Phys med biol, vol. 12, no. 4, pp. 501-509, 1967.
- [4] P. Duvaut, « Traitement du signal, concepts et applications, » Hermès, 1991.
- [5] S. Geman and D. Geman, « Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, » IEEE Trans Pat Anal Mach Int, vol. PAMI-6, no. 6, March 1984.
- [6] M. K. Nguyen and A. Mohammad-Djafari, « Bayesian approach with the Maximum Entropy principle in image reconstruction from micro wave scattered field data, » IEEE Trans. Med. Im., vol. 3, no. 2, June 1994.
- [7] J. K. Óruanaidh and W. J. Fitzgerald, « Numerical Bayesian methods applied to signal processing, » Springer Verlag, 1996.
- [8] G. Winkler, « Image analysis, random fields and dynamic Monte Carlo methods, » Springer Verlag, 1995.
- [9] C. A. Bouman and K. Sauer, « A unified approach to statistical tomography using coordinate descent optimization, » IEEE Trans. Med. Im., vol. 5, no. 3, March 1996.
- [10] A. Blake and A. Zisserman, « Visual reconstruction, » MIT Press, Cambridge MA, 1987.
- [11] M. Nikolova, A. Mohammad-Djafari and J. Idier, « Inversion of large-support ill-conditioned linear operation using a Markov model with a line process, » IEEE ICASSP 1994, vol. 5, pp. 357-360, 1994.
- [12] J. M. Bruneau, « Résolution de problèmes inverses mal-posés en analyse multirésolution. Application à la restauration d'images et à la stéréovision, » Thèse de Sciences, Université de Nice-Sophia Antipolis, 1993.