TWO DETERMINISTIC HALF-QUADRATIC REGULARIZATION ALGORITHMS FOR COMPUTED IMAGING

Pierre Charbonnier*, Laure Blanc-Féraud*, Gilles Aubert †, Michel Barlaud*, member IEEE,

* Laboratoire I3S, URA 1376 du CNRS,
GDR 134: Traitement du Signal et Images
Université de NICE-SOPHIA ANTIPOLIS
Bât. 4 UNSA-CNRS , 250 Av. A. Einstein, Les lucioles 1
06560 Valbonne (France)

† Laboratoire J.A. DIEUDONNÉ
URA 168 DU CNRS
Université de NICE-SOPHIA ANTIPOLIS
Parc Valrose
06108 NICE cédex 2 (France)

ABSTRACT

Many image processing problems are *ill-posed* and must be regularized. Usually, a roughness penalty is imposed on the solution. The difficulty is to avoid the smoothing of edges, which are very important attributes of the image.

In this paper, we first give sufficient conditions for the design of such an *edge-preserving regularization*. Under these conditions, it is possible to introduce an auxiliary variable which role is twofold. Firstly, it marks the discontinuities and ensures their preservation from smoothing. Secondly, it makes the criterion *half-quadratic*. The optimization is then easier. We propose a deterministic strategy, based on alternate minimizations on the image and the auxiliary variable. This yields two algorithms, ARTUR and LEGEND. In this paper, we apply these algorithms to the problem of SPECT reconstruction.

1. INTRODUCTION

In many computed imaging applications, the observed data p can be related to the unknown image f through a linear model of the form:

$$p = Rf + \eta \qquad (1)$$

where η is assumed to be a white Gaussian noise, and R is a linear operator. R can be a space-invariant operator, as in image restoration. In this paper, we address the problem of tomographic reconstruction. In this case, R is spatially variant and represents the Radon transform.

In image Processing, reconstructing the image f from the data p is an ill-posed problem: the knowledge of the direct model (1) is not sufficient to determine a satisfying solution. It is then necessary to regularize the problem, i.e. to introduce an *a priori* constraint on the solution.

This can be done by enforcing a roughness penalty. The difficulty is then to avoid the smoothing of discontinuities, which are important attributes of images. Markov Random Fields are a way to design such an *edge-preserving regularization* [1]. The Gibbs energy involves a first or second order difference operator (discrete derivative) to regularize the solution. The potential function applied on these differences is selected such as edges on the solution are preserved.

In this paper, we first give sufficient conditions on the potential functions that ensure an edge-preserving regularization. The potential function can be convex or non convex, leading to a convex or non convex global objective function (or criterion). For medical tomographic reconstruction, the computation time is limited so a deterministic strategy is preferable. Additional difficulty comes from the non linearity of the gradient of an edge preserving criterion, whenever it is convex or not.

Based on results of [2,3], we propose in a second part to transform this non quadratic and possibly non convex criterion into an *half-quadratic* criterion by introducing an auxiliary variable. This permits to *linearize* the problem and to derive a *deterministic* algorithm, based on alternate minimizations on the image and the auxiliary variable. This yields a relaxation algorithm based on the general principle of minimization of a sequence of criteria. Well-known deterministic algorithms such as GNC (Graduated Non Convexity [4]) and MFA (Mean Field Annealing [5]) use this principle to minimize non convex functional. However, the proposed method can be applied to any criterion involving edge preserving regularization potential function.

Two different ways to introduce the auxiliary variables are proposed. Then, the associated algorithms are concurrently presented. Auxiliary variables in both algorithms represent the image edges. Therefore the algorithms simultaneously reconstruct the image and its associated edge map. Some results are presented in the case of tomographic reconstruction.

2. CONDITIONS FOR EDGE PRESERVATION

In a Bayesian framework, the posterior distribution $P(f/p)$ is defined by :

$$P(f/p) \propto P(p/f).P(f) \qquad (2)$$

The likelihood distribution is defined in (1) as a white gaussian noise, and the a priori distribution $P(f)$ is supposed to be in this paper a Gibbs distribution with an energy $E(f)$. The MAP (Maximum a posteriori) estimate is given by the minimum of a criterion $J(f)$ of the form

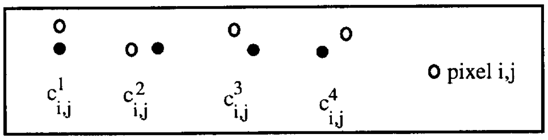
$$J(f) = \|p - Rf\|^2 + E(f) \qquad (3)$$

The energy $E(f)$ corresponds to the energy of a Markov random field such as :

$$\mathrm{E}(f) = \lambda^2 \sum_{c \in C} \varphi(d_c f) \qquad (4)$$

0-8186-6950-0/94 $4.00 © 1994 IEEE

$\mathcal{C}$ is the set of cliques with two pixels of a second order Markov model. In pixel *(i,j)*, the four cliques $c^l \,/\, l = 1,..,4$ are :



d_c is a difference operator on the clique c, corresponding to a first order derivative on *f*. In pixel *(i,j)* for l=1, $d_{c^l_{i,j}}$ is equal to :

$$d_{c^1_{i,j}} f = \mu_1 \left(f_{i+1,j} - f_{i,j} \right) \tag{5}$$

and similar definitions hold for l=2,3,4. μ_l is defined by :

$$\mu_l = \begin{cases} 1/\delta & \text{if } l = 1,2 \\ 1/(\delta\sqrt{2}) & \text{if } l = 3,4 \end{cases} \tag{6}$$

δ in (6) and λ^2 in (4) are the parameters of the model (see section 5).

The potential function φ is chosen such that edges are preserved. Several functions have been proposed (see for instance [2,4,6]). In many papers, however, the potential functions are chosen according to the conditions of validity of the algorithm used. We state the problem of the design of an edge preserving regularization in a general way, using the criterion defined by (3). We propose sufficient conditions on φ to ensure edge preservation independantly of the algorithm. Based on the analysis of the normal equation associated with (3), it can be shown that if the potential function (supposed to be even) verifies :

- $\lim\limits_{u \to 0^+} \dfrac{\varphi'(u)}{2u} = M , \quad 0 < M < +\infty$
- $\lim\limits_{u \to +\infty} \dfrac{\varphi'(u)}{2u} = 0 \qquad (7)$
- $\dfrac{\varphi'(u)}{2u}$ *is continious and strictly decreasing on* $[0, +\infty[$

then φ defines an edge preserving regularization [7]. φ' stands for the first derivative of φ.

Note that the class of function defined by (7) includes some convex functions such as $\varphi_{GR}(u) = logcosh(u)$ proposed in [6], or the function $\varphi_{HS}(u) = \sqrt{1+u^2}$ which have same behavior as φ_{GR}. Experimental results are shown in section 5 with φ_{GR} and a non convex function used in [2] $\varphi_{GM}(u) = \dfrac{u^2}{1+u^2}$.

3. HALF-QUADRATIC REGULARIZATION

The MAP estimate is given by the minimization of *J(f)* defined by (3)-(4).

In the non convex case, problems arise due to the existence of multiple local extrema. Stochastic relaxation algorithms give a global optimal solution but are very slow, mainly in the case of a large support operator R as in tomographic reconstruction. For a medical application, the computation time is limited so a deterministic scheme is preferable.

Three deterministic relaxation algorithms are well-known and widely used for the non convex minimization in Markovian estimation : the ICM (Iterated Conditional Mode) algorithm [8], the GNC (Graduated non Convexity) [4], and the MFA (Mean Field Annealing) [5]. Like stochastic algorithms, the ICM requires a great deal of computations in the case of a large support operator R. The GNC is designed for noisy image restoration $(R = I)$ and for $\varphi(u) = \varphi_{BL}(u) = min\left(u^2, 1\right)$. Adaptations can be done for other applications or potential function but this requires a new study for each new case. It is quite the same for the MFA : edges are modeled by a Boolean variable, that corresponds to the function φ_{BL}. Note however that constraints on edges can be introduced in the MFA. The model proposed hereafter does not take into account such constraints.

An additional difficulty in the minimization of *J(f)* is the non linearity of the gradient, whenever *J(f)* is convex or not. So in the convex case, due to the non linearity and the large size of the problem, standard optimization algorithms based on direct minimization are not practicable.

Based on results of [2,3], we propose to transform this non quadratic and possibly non convex criterion into an *half-quadratic* criterion by introducing an auxiliary variable. This permits to *linearize* the problem and to derive a *deterministic* algorithm, based on alternate minimizations on the image and the auxiliary variable.

The general principle of half quadratic regularization is to introduce an auxiliary variable b such as :

$$\varphi(u) = \underset{b}{Inf} \left\{ \varphi^*(u,b) \right\} \tag{8}$$

and such as $\varphi^*(u,b)$ is *convex* with respect to u when b is fixed.

Applying (8) on (4), we define a new (auxiliary) variable b_c on each clique $c^l_{i,j} / l = 1,..4$ and an augmented criterion $J^*(f,b)$ by :

$$J^*(f,b) = \|p - Rf\|^2 + \sum_{c \in \mathcal{C}} \varphi^*\left((d_c f), b_c \right) \tag{9}$$

$J^*(f,b)$ must be minimized in *f* and *b*.

The conditions for which (8) holds, and the expressions of φ^* are given in Tables 1 and 2, for two different cases. Note that these conditions are verified for all potential functions satisfying (7). The first case is an extension of results of [2] to a more general class of potential functions. The second case is based on the Legendre transform [9], introduced for half quadratic regularization in [3]. The algorithms derived in [2,3] are stochastic while we propose a deterministic optimization. The criterion $J^*(f,b)$ has two important properties.

On one hand, if b is fixed, $\varphi^*(u,b)$ is quadratic with respect to u, and then the energy $J^*(f,b)$ is quadratic with respect to f. So the minimization over f, when b is fixed, reduces to the resolution of linear normal equations given in Tables 1 and 2 (line 3).

Condition on φ	$\varphi(\sqrt{u})$ is strictly concave
Expression of $\varphi^*(u,b)$	$\exists\psi$ *function of* $b \in [0,\mathrm{M}]$, $\varphi^*(u,b) = bu^2 + \psi(b)$
b fixed Normal equation in $\hat{f}$ $\hat{f} = \arg\min_f \{J^*(f,b)\}$	$\left(R^t R - \lambda^2 \Delta(b)\right)\hat{f} = R^t p$ see (10)
f fixed Expression of $\hat{b}$ $\hat{b} = \arg\min_b \{J^*(f,b)\}$	$\hat{b}_c = \dfrac{\varphi'(d_c f)}{2 d_c f} \quad \forall c \in \mathcal{C}$

Table 1: First transformation into an half quadratic criterion, leading to the algorithm ARTUR.

Condition on φ	$u^2 - \varphi(u)$ is strictly convex
Expression of $\varphi^*(u,b)$	$\exists\zeta$ *function of* $b \in \Re$, $\varphi^*(u,b) = (b-u)^2 + \zeta(b)$
b fixed Normal equation in $\hat{f}$ $\hat{f} = \arg\min_f \{J^*(f,b)\}$	$\left(R^t R - \lambda^2 \Delta\right)\hat{f} = R^t p + \sum_{l=1}^{4} \lambda^2 D_l^t b_l$ see (10)
f fixed Expression of $\hat{b}$ $\hat{b} = \arg\min_b \{J^*(f,b)\}$	$\hat{b}_c = \left[1 - \dfrac{\varphi'(d_c f)}{2 d_c f}\right](d_c f) \quad \forall c \in \mathcal{C}$

Table 2: Second transformation into an half quadratic criterion, leading to the algorithm LEGEND.

In the normal equations, Δ is a discrete approximation of the Laplacian operator, $\Delta(b)$ is a weighted discrete Laplacian, D_l a matrix operator of differences, and Λ_l a diagonal weighting matrix, defined by :

$$\Delta = \sum_{l=1}^{4} D_l^t D_l \qquad \Delta(b) = \sum_{l=1}^{4} D_l^t \Lambda_l D_l \tag{10}$$

where $(D_l f)_{i,j} = d_{c_{i,j}^l} f$ and $\Lambda_l = diag\left\{ b_{c_{i,j}^l} \,/\, i,j = 1,..N \right\}$

On the other hand, it can be shown that if f is fixed, then the criterion is convex in b [7,10]. Moreover, the minimum value is given by analytical expressions (see Tables 1,2 line 4).

These considerations lead us to propose a simple deterministic algorithm based on alternative minimization on the two variables f and b.

4. TWO DETERMINISTIC ALGORITHMS

The MAP estimate is obtained by minimizing $J(f)$ over f- or, equivalently, $J^*(f,b)$ over f and b. We propose to use the principle of alternate minimizations :

$$\begin{aligned} & f^0 \equiv 0 \\ & \textit{Repeat} \\ & \quad \left| \begin{array}{l} 1)\ b^{n+1} = \arg\min_b \left[J^*\left(f^n, b\right) \right] \\ 2)\ f^{n+1} = \arg\min_f \left[J^*\left(f, b^{n+1}\right) \right] \end{array} \right. \\ & \textit{Until convergence} \end{aligned} \tag{11}$$

b^{n+1} and f^{n+1} are computed using the results given in Table 1 and 2, leading to the algorithms named ARTUR [11] and LEGEND. Both algorithms solve a sequence of normal equations that evolve through the b variable at each step n.

In the case of ARTUR, the system evolves with the matrix to invert. We use the iterative successive over relaxation algorithm to compute f^{n+1}.

In the LEGEND algorithm, the matrix to invert is constant at each step and the second member of the system evolves. The matrix can be inverted at the beginning of the algorithm. It results in faster iterations for the algorithm LEGEND than for ARTUR, but more iterations are needed due to a slower evolution of the solution.

In both cases, the auxiliary variable b models the edges of the image, in the same way as the line-process introduced in [1]. However, b is here is a continuous while it is a Boolean variable in [1]. Both algorithms give an estimation of the image and of its edges in the same time.

Regardless the auxiliary variable b, the algorithm (11) gives in f a relaxation algorithm based on the general principle of minimization of a sequence of criteria :

$$J^n(f) = J^*\left(f, b^n\right) \tag{12}$$

Contrarily to the GNC where the criteria evolution is arbitrary, each criterion evolves by itself according to the data. Note also that each criterion is convex.

When φ (and so $J(f)$) is convex, it can be shown theoretically that the sequence f^n converges [7]. The global minimum is then reached.

In the non convex case, the convergence is only established for the sequence of criteria $J\left(f^n, b^n\right)$, which forms a strictly decreasing and bounded below sequence. Experimentally, the convergence is always observed using both algorithms. The estimated solution is different from the global minimum and depends on the initial guess, f^0. However for any initial condition, ARTUR and LEGEND give satisfying solutions [10].

5. EXPERIMENTAL RESULTS

Some results are shown for Single Photon Emission Computed Tomography (SPECT) reconstruction on a synthetic phantom named Nice presented on Figure 1. This phantom is a 64x64 pixels image derived from Shepp & Logan's phantom, and models a cross-section of the brain. Synthetic projections are corrupted by Poisson noise, leading to a total of 6 million counts in the projection vector. The Signal to Noise Ratio on projection vector is 26.7dB (variance ratio).

The model (3)-(4) involves two parameters λ^2 and δ. λ^2 is the regularization parameter or smoothing parameter : it tunes the importance of the a priori in the solution. So in homogeneous area, it tunes the importance of the smoothing action in the solution. The second parameter, δ, is linked to the value of the differences $d_{c_{i,j}^l} f$ above which a discontinuity is introduced and there is no smoothing. These parameters are experimentally fixed to $\lambda^2 = 525$ and $\delta = 7$ in the simulations.

Figure 1 shows the estimate f^n for several steps n. The potential function used is φ_{GM}. By choosing the first initial guess $f^0 \equiv 0$ then $b^1 \equiv M$. It means that there is no discontinuity in the image. Then the first normal equation to be solved (Table 1 or 2, line 3) corresponds to the standard Tikhonov regularization (without edge preservation) :

$$\left(R^t R - \lambda^2 \Delta\right)\hat{f} = R^t p \tag{13}$$

This result is shown in (1), Figure 1. Then edges are gradually introduced (see steps 2,5,10,27). Step (27) is the final estimate, given by the algorithm ARTUR after convergence. We can check that the local minimum given in (27) is a good approximation of the original image Nice. The SNR (variance ratio) between the two images is 24.9dB.

Figure 2-(b) shows a reconstruction using the convex potential function φ_{GR}. Compared to the two reconstructions obtained by using the non convex potential function φ_{GM} with the algorithms ARTUR (c) and LEGEND (d), it can be noted that edges in (b) are less sharp than in (c) and (d). Nevertheless, (b) is really an edge preserving reconstruction. In the non convex case, the algorithms ARTUR and LEGEND give two different estimates (see images 2-(c) and 2-(d)). However the results are similar in terms of visual quality.

6. CONCLUSION

A deterministic scheme is proposed to minimize the non convex and non linear criteria involved in edge-preserving regularization. This scheme is based on the introduction of an auxiliary variable, which role is to mark discontinuities and to facilitate the estimation of the solution. The augmented criterion is quadratic with respect to the unknown image and convex with respect to the auxiliary variable.

The introduction of the auxiliary variable can be performed two ways, yielding two reconstruction algorithms: ARTUR and LEGEND. These algorithms are based on alternate minimizations on the image and the auxiliary variable. They simultaneously estimate the image and an edge map, given by the auxiliary variable.

Some results are given in the case of tomographic reconstruction. These results illustrate the behavior of the algorithm: starting from a smooth solution, they gradually introduce discontinuities into the solution. Both algorithms yield very good results, even when the regularization function is non convex.

REFERENCES

[1] S. Geman and D. Geman "Stochastic Relaxation, Gibbs Distribution, and The Bayesian Restoration of Images" *IEEE Trans. Pattern Anal. Machine Intell.*, Vol PAMI-6, pp. 721-741, Nov. 1984.

[2] S. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities", *IEEE Trans. Pattern Anal. Machine Intell.*, Vol PAMI-14, N°3, pp. 367-383, March 1992.

[3] D. Geman and C. Yang, "Nonlinear Image Recovery with Half-Quadratic Regularization and FFTs", *preprint*, March 1993.

[4] A. Blake, A. Zisserman, "Visual Reconstruction", *The MIT Press Series in artificial intelligence*, 1987.

[5] D. Geiger, F. Girosi, "Parallel and Deterministic Algorithms from MRF's: Surface Reconstruction", *IEEE Trans. on Pattern Anal. and Mach. Intell.,* Vol. PAMI-13, n°5, pp. 401-412, May 1991.

[6] P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm", *IEEE Trans. Med. Imaging*, Vol. MI-9, N°1, pp. 84-93, March 1990.

[7] G. Aubert, M. Barlaud, L. Blanc-Féraud and P. Charbonnier, "Deterministic Edge-Preserving Regularization in Computed Imaging", *Submitted to IEEE Transactions on Image Processing*, tech. rep., Research Report N° 94-01, I3S, University of Nice-Sophia Antipolis, 1994.

[8] J. Besag, "On the Statistcal Analysis of Dirty Pictures", *J. Roy. Statist. Soc.*, N°3, pp. 259-302, 1986.

[9] R. Rockafellar, "Convex Analysis", *Princeton University Press*, 1970.

[10] P. Charbonnier, "Reconstruction d'image : régularisation avec prise en compte des discontinuités" PhD. Thesis, University of Nice-Sophia Antipolis, september 1994.

[11] P. Charbonnier, L. Blanc-Féraud and M. Barlaud, "ARTUR: An Adaptive Deterministic Relaxation Algorithm for Edge-Preserving Tomographic Reconstruction", *Submitted to IEEE Transactions on Image Processing*, tech. rep., Research Report N° 93-76, I3S, University of Nice-Sophia Antipolis, 1993.

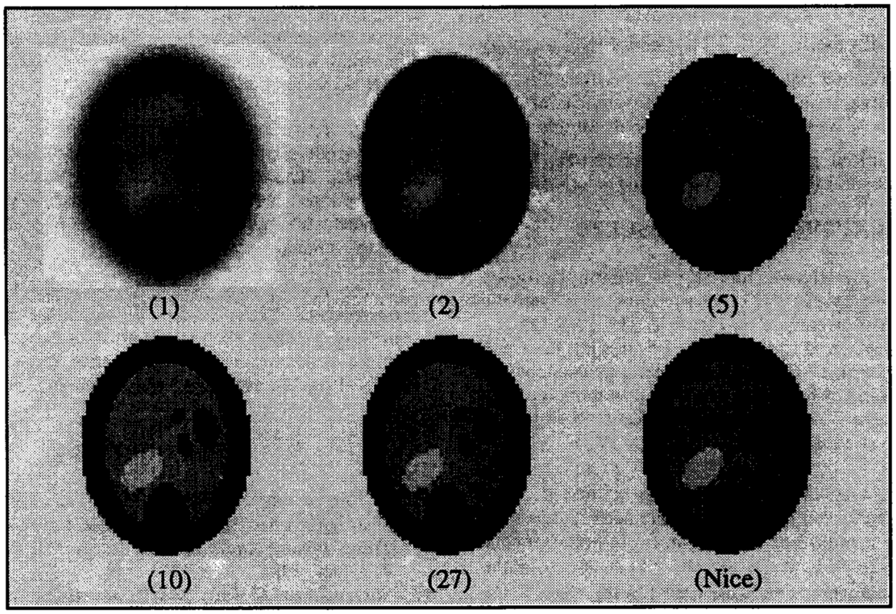


Figure 1: Estimated image after 1, 2, 5, 10 and 27 steps, and synthetic phantom (Nice).

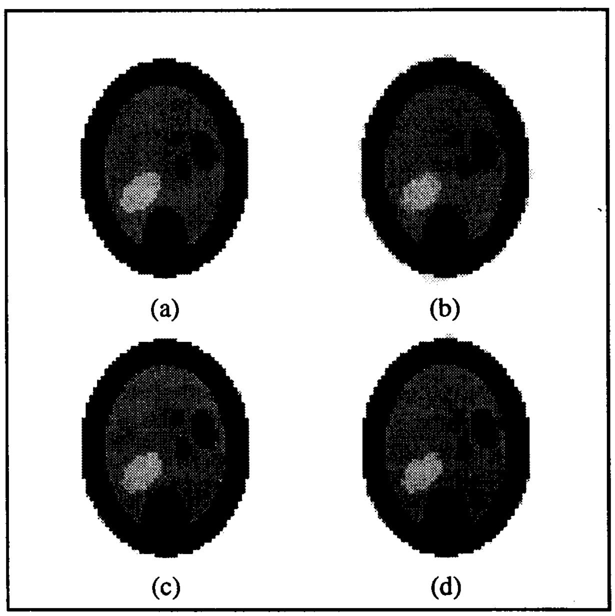


Figure 2: Synthetic phantom Nice (a). Reconstruction with ARTUR, using φ_{GR} (b) and φ_{GM} (c). Reconstruction with LEGEND using φ_{GM} (d).