TREIZIÈME COLLOQUE GRETSI - JUAN-LES-PINS DU 16 AU 20 SEPTEMBRE 1991

Sequential and Parallel Iterative Image Restoration

M. Teles de Figueiredo *José M. N. Leitão*

Centro de Análise e Processamento de Sinais - Instituto Nacional de Investigação Científica
Departamento de Engenharia Electrotécnica e de Computadores - Instituto Superior Técnico
Lisboa - PORTUGAL

RÉSUMÉ

Cet article aborde la restauration d'images floues et contaminées par bruit blanc gaussien. Les approches bayesienes et de regularization mennent à des problèmes d'optimisation de trés grande dimension. Si l'image originale est modélisée par un champ de Markov gaussien et le bruit est blanc et gaussien, le problème d'optimisation est quadratique, donc equivalent à un système linear d'équations. Une classe d'algorithmes itératifs, dont le Gauss-Seidel et le Jacobi sont des cas particuliers, est considérée. La convergence de l'algorithme de Gauss-Seidel est démontrée. On propose une version modifiée de la methode de Jacobi, et on établie sa convergence. Cet algorithme est bien adapté à mise en oeuvre parallèle.

ABSTRACT

In this paper, the restoration of images degraded by a linear blur followed by additive white Gaussian noise, is considered. The Bayesian and regularization approaches to restoration both lead to high dimension optimization problems. If the undegraded image model is a Gauss-Markov random field and the noise is white and Gaussian the resulting optimization problem is quadratic, thus equivalent to a linear system of equations. A class of iterative techniques, of which the Gauss-Seidel and the Jacobi algorithms are special cases, is considered. The Gauss-Seidel scheme is proved to converge. A modified version of the Jacobi algorithm, which is shown to converge, is developed. This algorithm is suitable for parallel implementation.

1 Introduction

The goal of image restoration is to recover an image that was degraded, e.g. blurred and corrupted by noise. Recovering the original image from the observed one is a severely ill-conditioned inverse problem [1], [2], [9]. To overcome this difficulty *a priori* constraints are imposed, leading to solutions that are compromises between closeness to the data and obedience to the constraints [1]. Techniques like regularization [2], [4], [7], [9] and Bayesian estimation [1], [3], [5] are two approaches supported on different ways of expressing prior knowledge.

Image restoration formulated in the Bayesian framework requires statistical models of the original image and of the degradation mechanism [1], [3], [5]. The uncorrupted image is modeled as a 2D noncausal Gauss Markov random field (GMRF), and the degradation mechanism is assumed to be a linear blur (LB) followed by additive white Gaussian noise (AWGN). The adoption of the maximum *a posteriori* probability (MAP) estimation criterion leads to a large dimension quadratic optimization problem, equivalent to a linear system of equations. The Tikhnov-Miller regularization approach, with a quadratic stabilizing functional, leads to a similar quadratic optimization problem.

A class of iterative methods, of which Gauss-Seidel and Jacobi algorithms are special cases, is considered. These methods are strictly local, i.e. the updating process depends only on a neighborhood of each pixel, thus being easily implementable in parallel hardware. Other faster algorithms require special purpose architectures (e.g. [8]). It is proved that, for the problem under study, the Gauss-Seidel algorithm converges and is equivalent to the sequential iterative relaxation method – *iterated conditional modes* (ICM) – proposed by Besag [3]. Aiming at parallel implementation (all sites updating their values simultaneously), and since the original Jacobi method can not be guaranteed to converge, a modified version is proposed. The structure of this algorithm allows parallel implementation namely on convolution oriented hardware.

2 Models

2.1 Markov Random Fields

Consider the images defined on a finite lattice $\mathcal{Z}_{MN}$, with $M \times N$ sites (pixels) [5]

$$\mathcal{Z}_{MN} = \{(ij) : 1 \leq i \leq M, 1 \leq j \leq N\}.$$

A neighborhood system $\mathcal{N} = \{N_{ij} \; : (ij) \in \mathcal{Z}_{MN}\}$, defined on $\mathcal{Z}_{MN}$, is any collection of subsets such that

$$(ij) \notin N_{ij} \quad \text{and} \quad (ij) \in N_{kl} \Leftrightarrow (kl) \in N_{ij}.$$

A subset with only one site or a set of mutually neighbor sites is called a *clique*. Let $\mathcal{C}$ denote the set of all cliques.

Since the lattice is finite, special rules have to be provided for the boundary sites. The *free boundary* condition [5] is here adopted.

Let $\mathbf{X} = \{X_{(ij)} : (ij) \in \mathcal{Z}_{MN}\}$ be a family of random variables defined on a space Λ. Let Ω be the set of all possible configurations. If all configurations have nonzero probability and

$$P\left(x_{(ij)} | \{x_{(kl)}, (kl) \neq (ij)\}\right) = P\left(x_{(ij)} | \{x_{(kl)}, (kl) \in N_{ij}\}\right),$$

for all (ij) in $\mathcal{Z}_{MN}$, the family $\mathbf{X}$ is said to be a Markov random field (MRF) with respect to the neighborhood system $\mathcal{N}$. If Λ is continuous, $P(\cdot)$ stands for probability density functions; if Λ is discrete $P(\cdot)$ denotes probability masses. According to the Hammersley-Clifford theorem [3], [5], the joint probabilty of a MRF has the Gibbs form

$$P(\mathbf{x}) = \frac{1}{Z} \exp\left\{-\sum_{C \in \mathcal{C}} V_C(\mathbf{x})\right\}, \qquad (1)$$

where Z, the *partition function* (PF), is a normalizing constant. Each *clique potential* $V_C(\mathbf{x})$ has the property of only depending on the sites in clique C, that is, $V_C(\mathbf{x}) = V_C(\{x_{(ij)} : (ij) \in C\})$.

The local conditional probabilities are obtained from the clique potentials as

$$P\left(x_{(ij)} | \{x_{(kl)}, (kl) \in N_{ij}\}\right) = \frac{\exp\left\{-\sum_{C:(ij)\in C} V_C(\mathbf{x})\right\}}{Z_{(ij)}} \qquad (2)$$

where $Z_{(ij)}$, the local PF, is a normalizing constant.

If clique potentials of the form

$$V_C(\mathbf{x}) = \left(\sum_{(ij)\in C} \alpha_{ij}^C \, x_{(ij)}\right)^2, \qquad (3)$$

are considered, the joint probability of the field can be written in vector notation as

$$P(\mathbf{x}) \propto \exp\{-\vec{\mathbf{x}}^T \mathbf{A} \vec{\mathbf{x}}\}, \qquad (4)$$

where $\vec{\mathbf{x}}$ is the $(MN \times 1)$ lexicographically ordered vector

$$\vec{\mathbf{x}} = \left[x_{(11)}, x_{(12)}, \ldots, x_{(1N)}, x_{(21)}, \ldots, x_{(2N)}, \ldots, x_{(MN)}\right]^T$$

and $\mathbf{A}$ is a $(MN \times MN)$ symmetric positive definite matrix (PDM). Equation (4) shows that, if Λ is the real line, $\vec{\mathbf{X}}$ is a zero mean Gaussian vector with covariance matrix $2\mathbf{A}^{-1}$. The elements of matrix $\mathbf{A}$ are given by

$$A_{(ij)(kl)} = \sum_{C \in \mathcal{C}} \beta_{ij}^C \beta_{kl}^C \qquad (5)$$

where

$$\beta_{ij}^C = \begin{cases} \alpha_{ij}^C & \Leftarrow \quad (ij) \in C \\ 0 & \Leftarrow \quad (ij) \notin C. \end{cases} \qquad (6)$$

Notation:

1. The lexicographically ordered vector of any 2D field $\mathbf{Z}$ will be denoted by $\vec{\mathbf{Z}}$.
2. $Z_{(ij)}$ represents the element $(i + (j-1)M)$ of vector $\vec{\mathbf{Z}}$.
3. $A_{(ij)(kl)}$ stands for the element $(i + (j-1)M, k + (l-1)M)$ of matrix $\mathbf{A}$.

2.2 Observation Model

The observed image is modeled as sample of a GMRF, linearly blurred and contaminated by AWGN. In vector notation

$$\vec{\mathbf{Y}} = \mathbf{B}\vec{\mathbf{X}} + \vec{\mathbf{W}}, \qquad (7)$$

where $\mathbf{B}$ is the $(MN \times MN)$ blur matrix, and $\vec{\mathbf{W}}$ is a $(MN \times 1)$ zero mean Gaussian vector with covariance matrix $\sigma^2\mathbf{I}$. The conditional probability of the observed image given the original one is then

$$P(\mathbf{y}|\mathbf{x}) \propto \exp\left\{-\frac{1}{2\sigma^2} \| \mathbf{B}\vec{\mathbf{x}} - \vec{\mathbf{y}} \|^2\right\}. \qquad (8)$$

3 MAP Restoration

The MAP estimate is defined as [3], [5]

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg\max_{\mathbf{x}} \{P(\mathbf{x}|\mathbf{y})\} = \arg\max_{\mathbf{x}} \{P(\mathbf{y}|\mathbf{x}) \cdot P(\mathbf{x})\}. \qquad (9)$$

Introducing (4) and (8) in (9) leads to

$$\hat{\mathbf{x}}_{\text{MAP}} = \arg\min_{\mathbf{x}} \left\{\vec{\mathbf{x}}^T \mathbf{C} \vec{\mathbf{x}} - 2\vec{\mathbf{b}}^T \vec{\mathbf{x}}\right\} \qquad (10)$$

where $\mathbf{C}$ is the $(MN \times MN)$ symmetric PDM

$$\mathbf{C} = \mathbf{A} + \frac{1}{2\sigma^2}\mathbf{B}^T\mathbf{B}, \qquad (11)$$

and $\vec{\mathbf{b}}$ is given by

$$\vec{\mathbf{b}} = \frac{1}{2\sigma^2}\mathbf{B}^T\vec{\mathbf{y}}.$$

Since (10) is convex it can be minimized by searching the zero of the gradient,

$$\mathbf{C}\hat{\vec{\mathbf{x}}}_{\text{MAP}} = \vec{\mathbf{b}}. \qquad (12)$$

4 Relation with Regularization

The Tikhonov-Miller regularization approach to the inversion of (7) leads to the minimization problem

$$\hat{\vec{\mathbf{x}}}_{\text{REG}} = \arg\min_{\vec{\mathbf{x}}} \{\| \mathbf{B}\vec{\mathbf{x}} - \vec{\mathbf{y}} \|^2 + \gamma \| \mathbf{D}\vec{\mathbf{x}} \|^2\}, \qquad (13)$$

where $\mathbf{D}$, the constraint operator, incorporates *a priori* information about the solution, and γ is the regularization parameter [7], [9]. Equation (13) can be rewritten as

$$\hat{\vec{\mathbf{x}}}_{\text{REG}} = \arg\min_{\vec{\mathbf{x}}} \left\{\vec{\mathbf{x}}^T \left(\mathbf{B}^T\mathbf{B} + \gamma\mathbf{D}^T\mathbf{D}\right)\vec{\mathbf{x}} - 2\left(\mathbf{B}^T\vec{\mathbf{y}}\right)^T \vec{\mathbf{x}}\right\}$$

which has the same form as (10),(11).

This means that all the techniques developed to perform MAP restoration of images under the GMRF-LB-AWGN assumption can also be used in the regularization approach.

5 Iterative Solutions

According to the preceding considerations, the image estimate is, in general, the solution of a linear system $\mathbf{C}\vec{\mathbf{x}} = \vec{\mathbf{b}}$. The huge dimension of matrix $\mathbf{C}$ ($(MN \times MN)$ for a M by N pixel image) strongly suggests the use of iterative schemes.

Splitting matrix $\mathbf{C}$ as $\mathbf{C} = \mathbf{G} - \mathbf{H}$ leads to the equivalent system $\mathbf{G}\vec{\mathbf{x}} = \mathbf{H}\vec{\mathbf{x}} + \vec{\mathbf{b}}$ and to the iteration

$$\mathbf{G}\vec{\mathbf{x}}(k+1) = \left(\mathbf{H}\vec{\mathbf{x}}(k) + \vec{\mathbf{b}}\right) \quad (14)$$

with initial condition $\vec{\mathbf{x}}(0)$ [6]. Matrix $\mathbf{G}$ has to be such that system (14) can be easily solved. Defining the error vector $\vec{\mathbf{e}}(k) = \vec{\mathbf{x}}(k) - \vec{\mathbf{x}}$, where $\vec{\mathbf{x}}$ is the solution of the system, it follows that $\vec{\mathbf{e}}(k) = (\mathbf{G}^{-1}\mathbf{H})^k\, \vec{\mathbf{e}}(0)$; therefore, iteration (14) converges if and only if matrix $\mathbf{M} \equiv \mathbf{G}^{-1}\mathbf{H}$ is *convergent* [6], i.e.

$$\lim_{k\to\infty} (\mathbf{G}^{-1}\mathbf{H})^k = 0. \quad (15)$$

Matrix $\mathbf{M}$ is convergent if and only if $\rho(\mathbf{M}) < 1$, where $\rho(\mathbf{M})$ stands for the *spectral radius* of $\mathbf{M}$ (the maximum of the magnitudes of all the eigenvalues).

5.1 Sequential Algorithm

Choosing matrix $\mathbf{G}$ to be the lower triangular part of $\mathbf{C}$ yields the Gauss-Seidel algorithm, also known as *successive iterations*. This scheme can be written in a simple way if the iteration is redefined as follows. Let a cyclic sequential visit schedule to the field sites be given, $\{(11), (12), \ldots, (1N), (21), \ldots, (MN), (11), \ldots\}$ (at each iteration, only the visited site is allowed to change its value); if, at time t, site (mn) is being visited, its value is changed according to

$$x_{(mn)}(t+1) = \frac{b_{(mn)} - \sum\limits_{(kl)\neq(mn)} C_{(mn)(kl)}\, x_{(kl)}(t)}{C_{(mn)(mn)}}. \quad (16)$$

The neighborhood defined by matrix $\mathbf{C}$, i.e. $N'_{ij} = \{(kl) : C_{(ij)(kl)} \neq 0\}$, depends on the blur matrix $\mathbf{B}$ and on the original neighborhood system $\mathcal{C}$.

Convergence of this method is guaranteed by the following theorem:

Theorem 1 *Consider the system $\mathbf{C}\vec{\mathbf{x}} = \vec{\mathbf{b}}$; if matrix $\mathbf{C}$ is symmetric, has positive diagonal elements, and is a PDM, then the Gauss-Seidel method converges (see [6], page 71).*

Since matrix $\mathbf{C}$ has positive diagonal elements,

$$C_{(mn)(mn)} = \sum_{C\in\mathcal{C}} (\beta^C_{mn})^2 + \frac{1}{2\sigma^2} \sum_{(ij)} (B_{(ij)(mn)})^2 > 0$$

and it is also a symmetric PDM, the conditions of the theorem are verified and convergence is demonstrated.

The deterministic relaxation algorithm ICM [3], that iteratively maximizes the a *posteriori* probability with respect to each field site, works as follows:

A visiting schedule to the field sites is given (e.g. sequential). If at time t, site (kl) is being visited, $x_{(kl)}$ is replaced by the value that maximizes the local conditional a posteriori probability,

$$x_{(kl)}(t+1) = \arg\max_{x_{(kl)}} \left\{ P(x_{(kl)} | \{x_{(ij)}, (ij) \neq (kl)\}, \mathrm{y}) \right\}.$$

Invoking the Markovian nature of $\mathbf{X}$, and using expressions (2) and (3), it is easily found that ICM leads to updating rule (16). This proves that ICM, under the GMRF-AWGN assumption, is nothing more than the Gauss-Seidel algorithm. A relation between ICM and the iterative solution of a system of equations was already suggested by Besag [3].

5.2 Parallel Algorithm

Choosing matrix $\mathbf{G}$ to be the diagonal of $\mathbf{C}$, the resulting scheme is the Jacobi algorithm or *simultaneous iterations* [6]. As convergence can not be guaranteed for the system under study, a modification has to be introduced. Instead of the diagonal of $\mathbf{C}$, let $\mathbf{G} = \mathrm{diag}\{\varepsilon_{(11)}, \varepsilon_{(12)}, \ldots, \varepsilon_{(MN)}\}$. Since matrix $\mathbf{G}$ is diagonal, this iterative process can be written explicitly as

$$x_{(ij)}(t+1) = \frac{1}{\varepsilon_{(ij)}} \left(b_{(ij)} - \sum_{(kl)} C_{(ij)(kl)}\, x_{(kl)}(t) \right) + x_{(ij)}(t). \quad (17)$$

The conditions on the parameters $\varepsilon_{(ij)}$ that assure convergence are given by the following theorem:

Theorem 2 *Let $\mathbf{C}\vec{\mathbf{x}} = \mathbf{b}$ be the system to be solved. The iterative algorithm defined by equation (17) converges if*

$$\varepsilon_{(ij)} > \frac{1}{2} \left(\sum_{(kl)} |C_{(ij)(kl)}| \right), \quad \forall_{(ij)}. \quad (18)$$

See proof in Appendix.

Since all the image pixels update their values simultaneously, the algorithm defined by (17) can be parallelly implemented. Although Jacobi-type algorithms converge slower than the Gauss-Seidel scheme, they are specially adequate to parallel implementation.

An important special case is obtained when the field is homogeneous and the blur is space-invariant (a convolution). In this case, matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are block-Toeplitz and equation (17) can be rewritten as a 2D convolution,

$$x_{(ij)}(t+1) = b_{(ij)} + \sum_k \sum_l f_{(kl)}\, x_{(i-k\; j-l)}(t), \quad (19)$$

and implemented on convolution oriented hardware. The kernel $f_{(kl)}$ of equation (19) is given by

$$f_{(kl)} = \begin{cases} -C_{(ij)(i-k\; j-l)}/\varepsilon_{(ij)} & \Leftarrow \quad (kl) \neq (00) \\ -C_{(ij)(i-k\; j-l)}/\varepsilon_{(ij)} + 1 & \Leftarrow \quad (kl) = (00). \end{cases}$$

6 Example

In the examples presented in figures 1 and 2 a first order neighborhood is adopted, i.e. $N_{(ij)} = \{(i+1\; j), (i-$

$1\ j), (i\ j+1), (i\ j-1)\}$, leading to cliques of the form $C_A = \{(ij), (i+1\ j)\}$ and $C_B = \{(ij), (i\ j+1)\}$. The clique potentials are

$$V_{C_A}(\mathbf{x}) = \frac{1}{\lambda^2}(x_{(ij)} - x_{(i+1\,j)})^2$$
$$V_{C_B}(\mathbf{x}) = \frac{1}{\lambda^2}(x_{(ij)} - x_{(i\,j+1)})^2.$$

Both figures show restored windows of a degraded image.

Figure 1: Blur: 5×5 uniform low pass. Additive white Gaussian noise variance: $\sigma^2 = 20^2$.

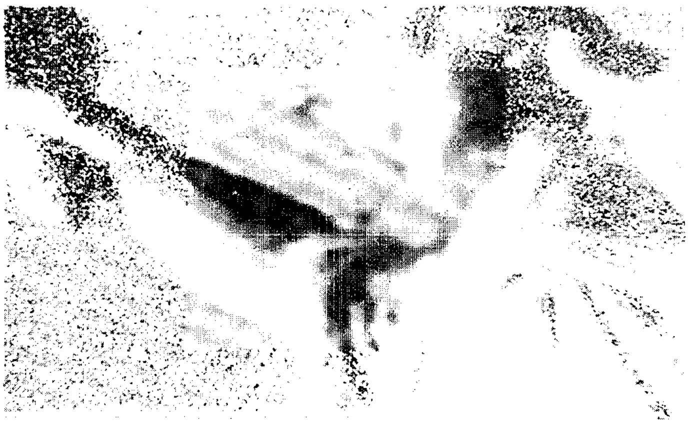

Figure 2: Blur: 9×9 uniform low pass. Additive white Gaussian noise variance: $\sigma^2 = 50^2$.

Appendix

The proof of theorem 2 is based on the following theorems:

Theorem 3 *Let* $\mathbf{G}$ *be a non-singular and symmetric matrix, and* $\mathbf{C} = \mathbf{G} - \mathbf{H}$ *be positive definite. Then,* $\mathbf{M} = \mathbf{G}^{-1}\mathbf{H}$ *is convergent if and only if* $\mathbf{Q} \equiv \mathbf{G} + \mathbf{H} = 2\mathbf{G} - \mathbf{C}$ *is a PDM (see [6], page 72).*

Theorem 4 (Gerschgorin) *Let* $\mathbf{M}$ *be a* $L \times L$ *matrix with eigenvalues* λ_k *and define the absolute row sum as*

$$r_i \equiv \sum_{j=1;j\neq i}^{L} |M_{ij}|.$$

Then, all eigenvalues lie in the union of the row circles,

$$\lambda_k \in \bigcup_{i=1}^{L} \{z\ :\ |z - M_{ii}| \leq r_i\}.$$

(see [6], page 135-137).

By theorem 3, $\mathbf{M}$ is convergent if and only if $\mathbf{Q}$ is a PDM, i.e. all its eigenvalues are positive. Since $\mathbf{C}$ and $\mathbf{Q}$ are symmetric, their eigenvalues are real. The elements of $\mathbf{Q}$ are

$$Q_{(ij)(kl)} = \begin{cases} 2\varepsilon_{(ij)} - C_{(ij)(ij)} & \Leftarrow \quad (ij) = (kl) \\ -C_{(ij)(kl)} & \Leftarrow \quad (ij) \neq (kl). \end{cases} \tag{20}$$

Using Gerschgorin's theorem and (20) it can be stated that, for any eigenvalue λ, of $\mathbf{Q}$,

$$\lambda \geq \min_{(ij)} \left\{ \left(2\varepsilon_{(ij)} - C_{(ij)(ij)}\right) - \sum_{(kl)\neq(ij)} |C_{(ij)(kl)}| \right\}$$
$$\geq \min_{(ij)} \left\{ 2\varepsilon_{(ij)} - \sum_{(kl)} |C_{(ij)(kl)}| \right\},$$

because $C_{(ij)(ij)} > 0$. A sufficient condition for all the eigenvalues to be positive is then (18). This concludes the proof of theorem 2.

References

[1] H. C. Andrews and B. R. Hunt. *Digital Image Restoration.* Prentice Hall, New Jersey, 1977.

[2] M. Bertero, T. Poggio, and V. Torre. "Ill-posed problems in early vision". *Proc. of the IEEE*, vol. 76, pp. 869–889, 1988.

[3] J. Besag. "On the statistical analysis of dirty pictures". *J. of the Royal Stat. Soc. B*, vol. 48, pp. 259–302, 1986.

[4] J. Biemond, R. Lagendijk, and R. Mersereau. "Iterative methods for image deblurring". *Proc. of the IEEE*, vol. 78, pp. 856–883, 1990.

[5] S. Geman and D. Geman. "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images". *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. PAMI-6, pp. 721–741, 1984.

[6] E. Isaacson and H. Keller. *Analysis of Numerical Methods.* John Wiley & Sons, New York, 1966.

[7] A. Katsaggelos, J. Biemond, R. Schafer, and R. Mersereau. "A regularized iterative image restoration algorithm". *IEEE Trans. on Signal Proc.*, vol. ASSP-39, pp. 914–929, 1991.

[8] T. Simchony, R. Chellappa, and Z. Lichtenstein. "Pyramid implementation of optimal-step conjugate-search algorithms for some low-level vision problems". *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-20, pp. 1408–1425, 1989.

[9] A. Tikhonov, A. Goncharsky, and V. Stepanov. "Inverse problems in image processing". In A. Tikhonov and A. Goncharsky, editors, *Ill-Posed Problems in the Natural Sciences*, pp. 220–232. Mir Publishers, Moscow, 1987.