ON INCREASING THE CONVERGENCE RATE OF REGULARIZED ITERATIVE IMAGE RESTORATION ALGORITHMS

Reginald L. Lagendijk, Russell M. Mersereau*[+] and Jan Biemond

Delft University of Technology
Dept. of Electrical Engineering
Information Theory Group
P.O. Box 5031, 2600 GA Delft
The Netherlands

* Georgia Institute of Technology
School of Electrical Engineering
Atlanta, Georgia 30332 USA

ABSTRACT

In [1] a regularized iterative algorithm was described which has been shown to be very suitable for solving the ill-posed image restoration problem. By incorporating deterministic constraints and adaptivity this very general algorithm is capable of achieving both noise suppression and ringing reduction in the restoration process. It consumes, however, considerable computation to obtain a (visually) stable solution due to the low convergence speed of the algorithm. The purpose of this paper is to investigate the possibilities for speeding up the convergence of this restoration method. To this end we compare the classical steepest descent algorithm (with linear convergence) with a conjugate gradients based method (superlinear convergence) and a new Q-th order converging algorithm. The latter solution method has the highest convergence rate, but is restricted in its application to space-invariant image restoration with a linear constraint. Although the actual convergence speed of the algorithms involved generally depends on the image data to be restored, it will be shown that for real-life images the constrained conjugate gradients algorithm yields a considerable convergence speed improvement.

1. INTRODUCTION

In many practical situations image degradations may be modeled by a linear blur and an additive noise term which is uncorrelated with the signal. The noisy blurred image can then be described by the following algebraic model:

$$g = Df + n, \tag{1}$$

where the linear distortion operator D is known or can be satisfactory identified. The original and noisy blurred images are denoted by the lexicographically ordered vectors f and g, respectively. The characteristics of the noise term n are only partially known in practice, hence the exact original image cannot be computed from the distorted version. Image restoration concentrates on removing the degradations caused by the blur and the noise to obtain an improved image $\hat{f}$ which is an acceptable approximation to the original image. Since the inverse problem formulated in eq. (1) is ill-posed, the solution method has to be regularized in order to obtain physically meaningful solutions. To this end we require the restored image $\hat{f}$ to satisfy the following 3 conditions in which adaptivity and a deterministic constraint are introduced to achieve both noise suppression and ringing reduction [1], [2]:

$$\text{(i)} \quad ||g-D\hat{f}||_R^2=(g-D\hat{f})^t R(g-D\hat{f}) \leq \varepsilon^2, \tag{2}$$

where R is a diagonal weight matrix which locally regulates the restoration process.

$$\text{(ii)} \quad ||L\hat{f}||_S^2 = (L\hat{f})^t S(L\hat{f}) \leq E^2, \tag{3}$$

where L represents a high-pass filter. Equation (3) imposes a smoothness condition on the restored image $\hat{f}$ which is locally regulated by the weight matrix S. (iii) The restored image $\hat{f}$ satisfies the (possibly nonlinear) constraint C, representing certain deterministic a priori information about the original image f. The nonexpansive projection onto the closed convex set described by C is denoted by P.

To compute a solution $\hat{f}$ satisfying the conditions (i)-(iii), first eqs. (2) and (3) are combined into a single quadrature formula

$$\Phi(\hat{f}) = ||g-D\hat{f}||_R^2 + \alpha||L\hat{f}||_S^2 \tag{4}$$

where the regularization parameter has the fixed value $\alpha=(\varepsilon/E)^2$.

The solution to the image restoration problem is now given by the vector $\hat{f}$ which minimizes the functional $\Phi(\hat{f})$ subject to the deterministic constraint C. Because this minimization problem is nonlinear and space-variant, an iterative solution method is used. The iterative image restoration procedures described so far, however, are based on the steepest descent method [1]-[4], and turn out to be very time consuming due to their low convergence speed. For this reason more efficient iterative image restoration algorithms, which preserve the advantages of the iterative approach and have a higher convergence speed, are desirable. In 1-D iterative signal processing [5]-[9] some earlier work concerning the convergence speed of iterative procedures has been reported by Marucci et al. [7], Prost and Goutte [8], and Singh et al. [9] among others.

In this paper we investigate possibilities for speeding up the convergence of the described image restoration method. In Section 2 we briefly review

---

[+] R.M. Mersereau was partially supported by the Joint Services Electronics Program under Contract DAAG29-84-0024.

CH2396-0/87/0000-1183 $1.00 © 1987 IEEE

the classical steepest descent algorithm as was presented in [1],[2]. Next, in Section 3 the method of conjugate gradients is considered, and extended with a projection operator. Basically, the methods of steepest descent and conjugate gradients have a linear convergence rate. In Section 4 we describe a new iterative algorithm with a Q-th order convergence rate. The convergence performance of the described algorithms will be illustrated by some numerical examples in Section 5.

2. STEEPEST DESCENT METHOD

In [1],[2] we have shown that $\Phi(\hat{f})$ is minimized subject to the constraint C by the following procedure: compute the gradient of $\Phi(\hat{f})$ with respect to $\hat{f}$, and define the mapping $G(\hat{f})$

$$G(\hat{f}) = \hat{f}-\tfrac{1}{2}\beta\ \nabla\Phi(\hat{f})=\hat{f}+\beta[D^tR(g-D\hat{f})-\alpha L^tSL\hat{f}], \qquad (5)$$

Then the fixed point iteration

$$\hat{f}_{k+1} = P.G(\hat{f}_k) \qquad (6)$$

converges monotonically to the required solution, provided that G is a contraction mapping, or

$$0 < \beta < 2\ ||D^tRD+\alpha L^tSL||^{-1}, \qquad (7)$$

where $||.||$ is the regular Euclidean norm. The corrections made by G to an iteration $\hat{f}_k$ are in the direction of the negative gradient, hence eq. (6) represents a steepest descent type of algorithm, extended with a projection operator. Since the convergence speed of steepest descent algorithms is controlled by the parameter β, it can be optimized by selecting an appropriate β in every iteration step. Substituting eq. (5) into (6), and introducing a variable β yields

$$r_k = -\tfrac{1}{2}\ \nabla\Phi(\hat{f}_k) = D^tR(g-D\hat{f}_k)-\alpha L^tSL\hat{f}_k \qquad (8a)$$

$$\hat{f}_{k+1} = P[\hat{f}_k + \beta_k\ r_k], \qquad (8b)$$

where r_k is called the direction vector for the k-th iteration step, and where β_k is chosen to minimize $\Phi(\hat{f}_{k+1})$. For a linear projection operator an explicit expression for the optimal β_k can be derived:

$$\beta_k = \frac{(r_k, Pr_k)}{||DPr_k||_R^2 + \alpha||LPr_k||_S^2} \quad . \qquad (9)$$

For a nonlinear projection operator such an explicit relation cannot be obtained. However, as a consequence of the convexity of the set described by the constraint C and the convexity of the set $K = \{h|\Phi(h) \le \Phi(\hat{f}_k)\}$, it can be shown that $\Phi(\hat{f}_{k+1})$ is a convex function of β_k. A (linear) search method, such as repeated quadratic interpolation or the golden-section rule, can therefore successfully be used to obtain the optimal β_k which minimizes $\Phi(\hat{f}_{k+1})$ [11].

Although the projection operator incorporated in the algorithm may increase the convergence rate (for example when C is a very tight constraint) [8], the steepest descent method has basically a linear convergence rate, and is known to converge slowly in practice. The attractiveness of the algorithm is, however, its straightforwardness, simple implementation which is independent of the projection operator to be used and its mild convergence conditions.

3. CONJUGATE GRADIENTS METHOD

Motivated by the desire to accelerate the method of steepest descent, the concept of conjugate directions has been introduced successfully in optimization theory [11]. Conjugate direction methods, which were invented for purely quadratic problems, can be viewed as a special orthogonal expansion of the solution of the minimization problem. We will focus on the most important conjugate direction method, namely the method of the conjugate gradients. One of the advantages of this method is the convergence in a finite number of iterations when exact arithmetic is assumed (superlinear convergence).For non-exact arithmetic the conjugate gradients method does not converge in a finite number of iterations because the conjugacy will no longer hold, but the method has still a considerably increased convergence speed compared with the steepest descent method.

Since the use of nonlinear constraints does not directly fit into conjugate direction algorithms, several ideas have been proposed for extending these methods with nonlinear constraints (for example the gradient projection method [11]). The major disadvantages of these extensions comprise the computational complexity, the strong dependence of the implementation on the constraints to be used and the fact that solutions of the truncated iterative process (as is always done in practice) do not always satisfy the constraint(s). We therefore apply the more direct extension as suggested by Marucci et al. in [7] to project the iterates themselves after every iteration step. Hence, we obtain a flexible algorithm whose complexity is comparable to that of the method of steepest descent. The extended conjugate gradients algorithm is defined as

$$r_k=-\tfrac{1}{2}\nabla\Phi(\hat{f}_k)=D^tR(g-D\hat{f}_k)-\alpha L^tS\ L\hat{f}_k \qquad (10a)$$

$$p_k = r_k + (||r_k||^2/||r_{k-1}||^2).p_{k-1} \qquad (10b)$$

$$\hat{f}_{k+1}=P[\hat{f}_k + \beta_k\ p_k]. \qquad (10c)$$

Here p_k is called the direction vector, which is based on the current steepest descent vector r_k and the preceding direction vector p_{k-1}. Eq. (10) represents a true conjugate gradients algorithm only when the projection operator is omitted. Clearly, by incorporating the projection operator the concept of the orthogonal solution decomposition will no longer hold. Fortunately, in the practice of image restoration the modifications made to an iterate by the projection operator are relatively small, which holds particularly for the restoration of real-life images. Hence, it can be expected that as long as the constraint C is not too tight, the direction vector of the previous iteration step will still be useful in determining the current direction vector.

The optimal β_k value for the extended conjugate gradients algorithm is obtained by minimizing $\Phi(\hat{f}_{k+1})$ with respect to β_k, yielding

$$\beta_k = \frac{(Pp_k, r_k)}{||DPp_k||_R^2 + \alpha||LPp_k||_S^2} \ . \qquad (11)$$

Linear search methods can be employed for nonlinear projection operators since $\Phi(\hat{f}_{k+1})$ is again a convex function of β_k. We cannot use, however, a fixed β_k value for every iteration step, since a general upperbound to β_k, which ensures the convergence of the iterations, does not exist.

4. A Q-th ORDER CONVERGING ALGORITHM

Although the method of the (extended) conjugate gradients has a considerably higher convergence speed than the steepest descent method, both techniques have basically a linear convergence rate. In [9] Singh et al. proposed an iterative restoration technique with a quadratic convergence rate, which we generalize and extend to a regularized iterative image restoration algorithm with a Q-th order convergence rate (Q=2,3,...) in this section. We remark that recently it was brought to our attention that an iterative deconvolution algorithm based on the same generalization has been independently developed by Morris et al. [10].

Consider the minimization of the function $\Phi(\hat{f})$ in eq. (4). The solution $\hat{f}$ of this unconstrained problem can be formulated as

$$(D^tRD + \alpha L^tSL)\, \hat{f} = D^tRg, \tag{12}$$

which we rewrite as

$$\hat{f} = (I - B_0)^{-1} \beta D^tRg \quad , \tag{13}$$

where $\beta \neq 0$ and the operator B_0 is defined by

$$B_0 = I - \beta(D^tRD + \alpha L^tSL). \tag{14}$$

Since the inversion of the matrix $(I-B_0)$ may be (nearly) impossible or time consuming, we approximate $\hat{f}$ in eq. (13) by a Taylor expansion with Q_0 terms ($Q_0 \geq 2$):

$$\hat{f}_1=(I+B_0+\ldots+B_0^{Q_0-1})\beta D^tRg = \sum_{j=0}^{Q_0-1} B_0^j \beta D^tRg. \tag{15}$$

By combining eqs. (13) and (15) to eliminate the term βD^tRg, and by defining $B_1 = B_0^{Q_0}$ we arrive at

$$\hat{f} = (I - B_1)^{-1} \hat{f}_1, \tag{16}$$

which is of the same form as eq. (13). Hence we can repeat the above process infinitely to obtain the following Q-th order converging algorithm

$$\hat{f}_{k+1} = \sum_{j=0}^{Q_k-1} B_k^j \hat{f}_k \; ; \qquad \hat{f}_0 = \beta D^tRg \tag{17a}$$

$$B_{k+1} = B_k^{Q_k} \qquad (Q_k \geq 2). \tag{17b}$$

Here Q_k determines the degree of convergence at iteration step k. To ensure the convergence of the iterations to the solution of eq. (13), B_k has to be a contraction mapping, yielding the same conditions for β as were obtained in the steepest descent algorithm (eq. (7)). The optimization of β in every iteration step to minimize $\Phi(\hat{f}_{k+1})$ results in solving an unattractive high-order polynomial equation, hence we prefer to use a fixed value of β. To interpret the iterative scheme in eq. (17), the following expression for the (k+1)-st iterate of eq. (17) is derived (Q_k=Q, for all k)

$$\hat{f}_{k+1} = \sum_{i=0}^{(Q^{k+1}-1)} (I-\beta(D^tRD + \alpha L^tSL))^i \beta D^tRg. \tag{18}$$

By comparing this result with the corresponding expression obtained from the unconstrained steepest descent algorithm with a fixed β value

$$\hat{f}_{k+1} = \sum_{i=0}^{k+1} (I-\beta(D^tRD + \alpha L^tSL))^i \beta D^tRg, \tag{19}$$

we observe that the proposed iterative scheme eq. (17) and the unconstrained steepest descent method compute exactly the same solution. However, the steepest descent algorithm requires $Q^{k+1}-1$ iterations to obtain the same solution which the Q-th order converging algorithm reaches after only k+1 iterations. The extra expense for the enormous reduction in the required number of iterations are more complicated computations in a single iteration step and extra memory requirements to store B_k. The efficiency of the proposed algorithm therefore depends on the choice of the convergence rate parameter Q and the way by which the algorithm has been implemented.

The most simple and direct method for incorporating the projection operator P in the Q-th order converging algorithm is to project the iterates as was proposed in [9]

$$\hat{f}_{k+1} = P\, [\sum_{j=0}^{Q_k-1} B_k^j \hat{f}_k]. \tag{20}$$

This extension will, however, inevitably lead to diverging iterations and erroneous results since the unaltered iterations on the matrix B_k in eq. (17b) would progress independently of the projection operator. Consequently, the incorporation of the projection operator in eq. (20) must be followed by a modification of eq. (17b) as well. By observing that in the k-th iteration step the following relation holds (cf. eq. (16))

$$\hat{f} = (I - B_k)^{-1} \hat{f}_k, \tag{21}$$

and by using eq. (20), some mathematical manipulations yield the following relation between two successive B_k matrices

$$P(I - B_k^{Q_k})\, \hat{f} = (I - B_{k+1})\, \hat{f}. \tag{22}$$

For a linear projection operator eq. (22) reduces to

$$B_{k+1} = P\, B_k^{Q_k} + (I-P), \tag{23}$$

which includes eq. (17b) as a special case. Eq.(22) cannot be solved for a nonlinear projection operator, therefore the extension in eq. (20) holds merely for linear projections.

The proposed Q-th order converging algorithm can generally be used only for space-invariant image restoration, since space-variant operators (as introduced by the R and S weight matrices) would require the storage of the (sparse) N^2xM^2 matrix B_k for NxM images. In 1-D signal processing the matrix B_k is only of the size N^2, which disposes of the

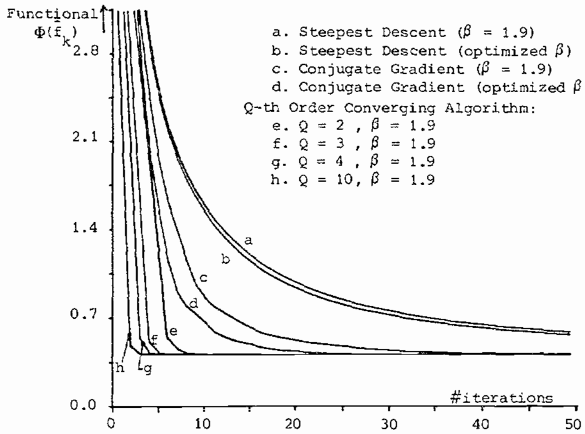

Fig. 1. Convergence behavior of $\Phi(\hat{f}_k)$ (unconstrained).

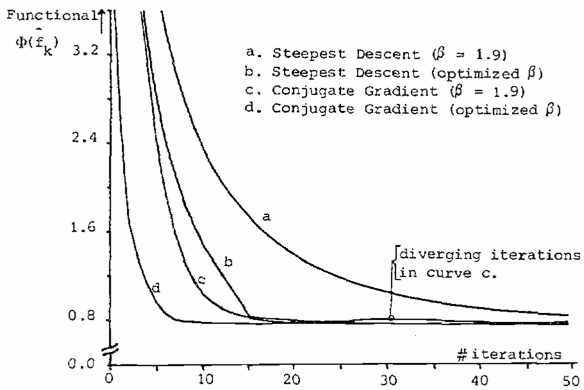

Fig. 2. Convergence behavior of $\Phi(\hat{f}_k)$ (constrained).

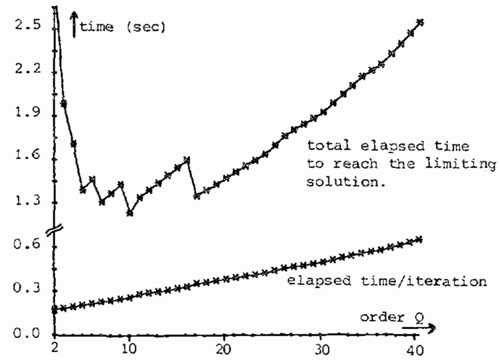

Fig. 3. Order Q of the algorithm in eq.(17) versus the required convergence time.

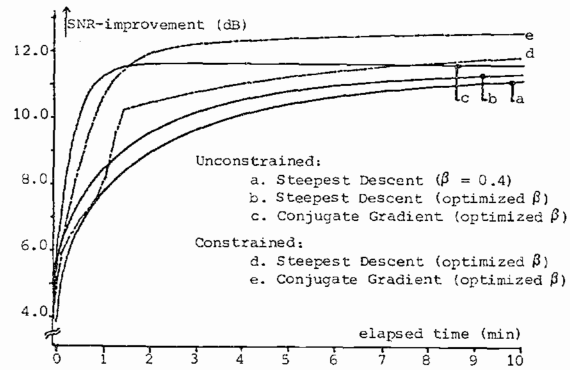

Fig. 4. SNR improvements versus elapsed iteration time for a real-life scene.

above restriction.

5. EXPERIMENTAL RESULTS

A synthetic "G"-image was blurred with a pillbox defocusing model (squared radius = 5), and uncorrelated random noise was added with SNR = 30 dB. In figure 1 the convergence behavior of $\Phi(\hat{f}_k)$ is shown for the unconstrained steepest descent, conjugate gradients and the Q-th order converging algorithm (α=0.01, L=Laplace operator, R=S=Identity). Clearly, the Q-th order converging algorithm has the best convergence performance; the optimal value(s) for Q considering the total elapsed time can be obtained from figure 3. In figure 2 the performance of the steepest descent and conjugate gradients method are compared when a very tight constraint is used, namely bounding the upper and lower intensity values. Note that in this case the Q-th order converging algorithm cannot be applied (see section 4).

In the next example an image of a real-life scene ("cameraman") was blurred with motion blur over 5 pixels and noise was added with SNR=30 dB. We used the adaptive regularized algorithm (using the R and S matrices) to achieve both noise suppression and ringing reduction in this real-life scene. In figure 4 the SNR improvements versus the elapsed time, which is the most important in practice, are shown for both the steepest descent and conjugate gradients algorithm with and without an intensity constraint. For the image under consideration this constraint is moderately tight. Undoubtly the convergence performance of the (extended) conjugate gradients algorithm is superior to the steepest descent algorithm in both the constrained and unconstrained case.

6. ACKNOWLEDGEMENTS

The assistance of prof. dr. D.E. Boekee in various aspects of this work is greatly acknowledged.

REFERENCES

[1] Biemond, J. and R.L. Lagendijk, "Regularized Iterative Image Restoration in a Weighted Hilbert Space", ICASSP'86.

[2] Lagendijk, R.L., J. Biemond and D.E. Boekee, "Regularized Iterative Image Restoration with Ringing Reduction", subm. IEEE trans. on Acoust., Speech and Signal Processing, 1987.

[3] Ichioka, Y. and N. Nakajima, "Iterative Image Restoration Considering Visibility", J. Opt. Soc. Am., vol. 71, 1981.

[4] Katsaggelos, A.K., J.Biemond, R.M.Mersereau and R.W. Schafer, "Nonstationary Iterative Image Restoration", ICASSP'85.

[5] Schafer, R.W., R.M. Mersereau and M.A. Richards, "Constrained Iterative Restoration Algorithms", Proc. IEEE, vol. 69, 1981.

[6] Trussell, H.J. and M.R. Civanlar, "The Feasible Solution in Signal Restoration", IEEE trans. ASSP., vol. 32, 1984.

[7] Marucci, R., R.M. Mersereau and R.W. Schafer, "Constrained Iterative Deconvolution using a Conjugate Gradient Algorithm", ICASSP'82.

[8] Prost, R. and R. Goutte, "Discrete Constrained Iterative Deconvolution Algorithms with Optimized Rate of Convergence", Signal Processing, vol. 7, 1984.

[9] Singh, S., S.N. Tandon and H.M. Gupta, "An Iterative Restoration Technique", Signal Processing, vol. 11, 1986.

[10] Morris, C.E., M.A. Richards and M.H. Hayes, "An Iterative Deconvolution Algorithm with Pth-order Convergence", Proc. of the 1986 DSP Workshop, Chatham, Mass., 1986.

[11] Luenberger, D.G., Introduction to Linear and Nonlinear Programming, Addison Wesley, Reading Mass., 1973.