

3.9

Iterative Image Restoration

Aggelos K. Katsaggelos

*Northwestern University,
Evanston*

Chun-Jen Tsai

*National Chiao
Tung University,
Taiwan*

1	Introduction.....	235
2	Iterative Recovery Algorithms	236
3	Spatially Invariant Degradation.....	236
3.1	Degradation Model • 3.2 Basic Iterative Restoration Algorithm •	
3.3	3.3 Convergence • 3.4 Reblurring • 3.5 Experimental Results	
4	Matrix-Vector Formulation	241
4.1	4.1 Basic Iteration • 4.2 Least-Squares Iteration • 4.3 Constrained Least-Squares Iteration • 4.4 Spatially Adaptive Iteration	
5	Use of Constraints	247
5.1	5.1 Experimental Results	
6	Additional Considerations.....	248
6.1	6.1 Other Forms of the Iterative Algorithm • 6.2 Hierarchical Bayesian Image Restoration • 6.3 Additional Applications	
7	Discussion	252
	References.....	252

1 Introduction

In this chapter we consider a class of iterative image restoration algorithms. Let \mathbf{g} be the observed noisy and blurred image, \mathbf{D} the operator describing the degradation system, \mathbf{f} the input to the system, and \mathbf{v} the noise added to the output image. The input-output relation of the degradation system is then described by [2]

$$\mathbf{g} = \mathbf{D}\mathbf{f} + \mathbf{v}. \quad (1)$$

The image restoration problem therefore to be solved is the inverse problem of recovering \mathbf{f} from knowledge of \mathbf{g} , \mathbf{D} and \mathbf{v} .

There are numerous imaging applications which are described by (1) [2, 3, 7, 17]. \mathbf{D} , for example, might represent a model of the turbulent atmosphere in astronomical observations with ground-based telescopes, or a model of the degradation introduced by an out-of-focus imaging device. \mathbf{D} might also represent the quantization performed on a signal or a transformation of it, for reducing the number of bits required to represent the signal.

The success in solving any recovery problem depends on the amount of the available prior information. This information refers to properties of the original image, the degradation system (which is in general only partially known), and the noise process. Such prior information can, for example, be represented by the fact that the original image is a sample of a stochastic field, or that the image is “smooth,” or that it takes only non-negative values. Besides defining the amount of prior information, equally critical is the ease of incorporating it into the recovery algorithm.

After the degradation model is established, the next step is the formulation of a solution approach. This might involve the stochastic modeling of the input image (and the noise), the determination of the model parameters, and the formulation of a criterion to be optimized. Alternatively it might involve the formulation of a functional to be optimized subject to constraints imposed by the prior information. In the simplest possible case, the degradation equation defines directly the solution approach. For example, if \mathbf{D} is a square invertible matrix, and the noise is ignored in (1), $\mathbf{f} = \mathbf{D}^{-1}\mathbf{g}$, is the desired unique solution. In most cases, however, the solution of (1) represents an ill-posed problem [25].

Application of regularization theory transforms it to a well-posed problem which provides meaningful solutions to the original problem.

There are a large number of approaches providing solutions to the image restoration problem. For recent reviews of such approaches refer, for example, to [3, 17]. This chapter concentrates on a specific type of iterative algorithms, the *successive approximations* algorithm, and its application to the image restoration problem.

2 Iterative Recovery Algorithms

Iterative algorithms form an important part of optimization theory and numeric analysis. They date back to the Gauss's time, but they also represent a topic of active research. A large part of any textbook on optimization theory or numeric analysis deals with iterative optimization techniques or algorithms [22].

Out of all possible iterative recovery algorithms we concentrate on the successive approximations algorithms, which have been successfully applied to the solution of a number of inverse problems ([23] represents a very comprehensive paper on the topic). The basic idea behind such an algorithm is that the solution to the problem of recovering a signal which satisfies certain constraints from its degraded observation, can be found by the alternate implementation of the degradation and the constraint operator. Problems reported in [23] which can be solved with such an iterative algorithm are the phase-only recovery problem, the magnitude only recovery problem, the bandlimited extrapolation problem, the image restoration problem, and the filter design problem [8]. Reviews of iterative restoration algorithms are also presented in [4, 14]. There are a number of advantages associated with iterative restoration algorithms, among which [14, 23]: (i) there is no need to determine or implement the inverse of an operator; (ii) knowledge about the solution can be incorporated into the restoration process in a relatively straightforward manner; (iii) the solution process can be monitored as it progresses; and (iv) the partially restored signal can be utilized in determining unknown parameters pertaining to the solution.

In the following we first present the development and analysis of two simple iterative restoration algorithms. Such algorithms are based on a linear and spatially invariant degradation, when the noise is ignored. Their description is intended to provide a good understanding of the various issues involved in dealing with iterative algorithms. We adopt a "how-to" approach; it is expected that no difficulties will be encountered by anybody wishing to implement the algorithms. We then proceed with the matrix-vector representation of the degradation model and the iterative algorithms. The degradation systems described now are linear but not necessarily spatially invariant. The relation between the

matrix-vector and scalar representation of the degradation equation and the iterative solution is also presented. Experimental results demonstrate the capabilities of the algorithms.

3 Spatially Invariant Degradation

3.1 Degradation Model

Let us consider the following degradation model

$$g(n_1, n_2) = d(n_1, n_2) * f(n_1, n_2), \quad (2)$$

where $g(n_1, n_2)$ and $f(n_1, n_2)$ represent, respectively, the observed degraded and original image, $d(n_1, n_2)$ the impulse response of the degradation system, and $*$ denotes two-dimensional (2D) convolution. It is mentioned here that the arrays $d(n_1, n_2)$ and $f(n_1, n_2)$ are appropriately padded with zeros, so that the result of 2D circular convolution equals the result of 2D linear convolution in (2) (see Chapter 2.3). Henceforth, in the following all the convolutions involved are circular convolutions and all the shifts are circular shifts.

We rewrite (2) as follows

$$\Phi(f(n_1, n_2)) = g(n_1, n_2) - d(n_1, n_2) * f(n_1, n_2) = 0. \quad (3)$$

The restoration problem therefore, of finding an estimate of $f(n_1, n_2)$ given $g(n_1, n_2)$ and $d(n_1, n_2)$, becomes the problem of finding a root of $\Phi(f(n_1, n_2)) = 0$.

3.2 Basic Iterative Restoration Algorithm

The following identity holds for any value of the parameter β

$$f(n_1, n_2) = f(n_1, n_2) + \beta\Phi(f(n_1, n_2)). \quad (4)$$

Equation (4) forms the basis of the successive approximations iteration, by interpreting $f(n_1, n_2)$ on the left-hand side as the solution at the current iteration step, and $f(n_1, n_2)$ on the right-hand side as the solution at the previous iteration step. That is, with $f_0(n_1, n_2) = 0$

$$\begin{aligned} f_{k+1}(n_1, n_2) &= f_k(n_1, n_2) + \beta\Phi(f_k(n_1, n_2)) \\ &= \beta g(n_1, n_2) + (\delta(n_1, n_2) \\ &\quad - \beta d(n_1, n_2)) * f_k(n_1, n_2), \end{aligned} \quad (5)$$

where $f_k(n_1, n_2)$ denotes the restored image at the k -th iteration step, $\delta(n_1, n_2)$ the discrete delta function and β the relaxation parameter which controls the convergence, as well as, the rate of convergence of the iteration. Iteration (5) is the basis of a large number of iterative recovery algorithms,

and is therefore analyzed in detail. Perhaps the earliest reference to iteration (5) with $\beta=1$ was by Van Cittert [26] in the 1930s.

3.3 Convergence

Clearly if a root of $\Phi(f(n_1, n_2))$ exists, this root is a *fixed point* of iteration (5), that is, a point for which $f_{k+1}(n_1, n_2) = f_k(n_1, n_2)$. It is not guaranteed however that iteration (5) will converge, even if (3) has one or more solutions. Let us, therefore, examine under what condition (sufficient condition) iteration (5) converges. Let us first rewrite it in the discrete frequency domain, by taking the 2D discrete Fourier transform (DFT) of both sides. It then becomes

$$F_{k+1}(u, v) = \beta G(u, v) + (1 - \beta D(u, v))F_k(u, v), \quad (6)$$

where $F_k(u, v)$, $G(u, v)$, and $D(u, v)$ represent respectively the 2D DFT of $f_k(n_1, n_2)$, $g(n_1, n_2)$ and $d(n_1, n_2)$. We express next $F_k(u, v)$ in terms of $F_0(u, v)$. Clearly

$$\begin{aligned} F_1(u, v) &= \beta G(u, v) \\ F_2(u, v) &= \beta G(u, v) + (1 - \beta D(u, v))\beta G(u, v) \\ &= \sum_{\ell=0}^1 (1 - \beta D(u, v))^{\ell} \beta G(u, v) \\ &\vdots \\ F_k(u, v) &= \sum_{\ell=0}^{k-1} (1 - \beta D(u, v))^{\ell} \beta G(u, v) \\ &= H_k(u, v)G(u, v) \end{aligned} \quad (7)$$

We therefore see that the restoration filter at the k -th iteration step is given by

$$H_k(u, v) = \beta \sum_{\ell=0}^{k-1} (1 - \beta D(u, v))^{\ell}. \quad (8)$$

The obvious next question is under what conditions the series in (8) converges and what is this convergence filter equal to. Clearly if

$$|1 - \beta D(u, v)| < 1, \quad (9)$$

then

$$\lim_{k \rightarrow \infty} H_k(u, v) = \lim_{k \rightarrow \infty} \beta \frac{1 - (1 - \beta D(u, v))^k}{1 - (1 - \beta D(u, v))} = \frac{1}{D(u, v)}. \quad (10)$$

Notice that (9) is not satisfied at the frequencies for which $D(u, v) = 0$. At these frequencies

$$H_k(u, v) = k \cdot \beta, \quad (11)$$

and therefore, in the limit $H_k(u, v)$ is not defined. However, since the number of iterations run is always finite, $H_k(u, v)$ is a large but finite number.

Having a closer look at the sufficient condition for convergence, we see that (9) can be rewritten as

$$\begin{aligned} |1 - \beta \operatorname{Re}\{D(u, v)\} - \beta \operatorname{Im}\{D(u, v)\}|^2 &< 1 \\ \Rightarrow (1 - \beta \operatorname{Re}\{D(u, v)\})^2 + (\beta \operatorname{Im}\{D(u, v)\})^2 &< 1. \end{aligned} \quad (12)$$

Inequality (12) defines the region inside a circle of radius $1/\beta$ centered at $c=(1/\beta, 0)$ in the $(\operatorname{Re}\{D(u, v)\}, \operatorname{Im}\{D(u, v)\})$ domain, as shown in Fig. 1. From this figure it is clear that the left half-plane is not included in the region of convergence. That is, even though by decreasing β the size of the region of convergence increases, if the real part of $D(u, v)$ is negative, the sufficient condition for convergence cannot be satisfied. Therefore, for the class of degradations that this is the case, such as the degradation due to motion, iteration (5) is not guaranteed to converge.

The following form of (12) results when $\operatorname{Im}\{D(u, v)\} = 0$, which means that $d(n_1, n_2)$ is symmetric

$$0 < \beta < \frac{2}{D_{\max}(u, v)}, \quad (13)$$

where $D_{\max}(u, v)$ denotes the maximum value of $D(u, v)$ over all frequencies (u, v) . If we now also take into account that $d(n_1, n_2)$ is typically normalized, i.e., $\sum_{n_1, n_2} d(n_1, n_2) = 1$, and represents a low pass degradation, then $D(0, 0) = D_{\max}(u, v) = 1$. In this case (13) becomes

$$0 < \beta < 2. \quad (14)$$

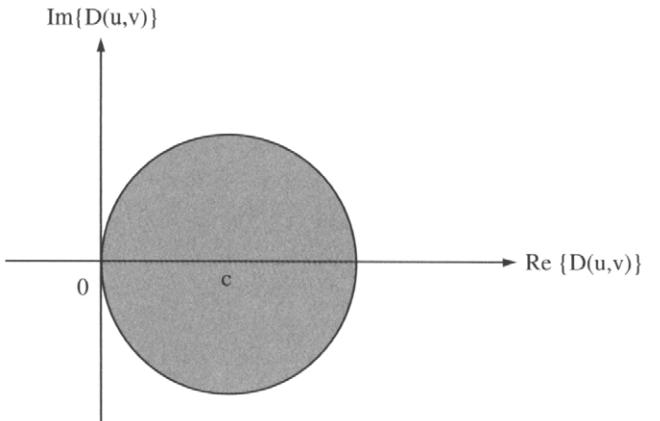


FIGURE 1 Geometric interpretation of the sufficient condition for convergence of the basic iteration, where $c=(1/\beta, 0)$.

From the above analysis, when the sufficient condition for convergence is satisfied, the iteration converges to the original signal. This is also the inverse solution obtained directly from the degradation equation. That is, by rewriting (2) in the discrete frequency domain

$$G(u, v) = D(u, v) \cdot F(u, v), \quad (15)$$

we obtain,

$$F(u, v) = \begin{cases} \frac{G(u, v)}{D(u, v)}, & \text{for } D(u, v) \neq 0 \\ 0, & \text{otherwise,} \end{cases} \quad (16)$$

which represents the *pseudo-inverse* or *generalized inverse* solution.

An important point to be made here is that, unlike the iterative solution, the inverse solution (16) can be obtained without imposing any requirements on $D(u, v)$. That is, even if (2) or (15) has a unique solution, that is, $D(u, v) \neq 0$ for all (u, v) , iteration (5) may not converge, if the sufficient condition for convergence is not satisfied. It is therefore not the appropriate iteration to solve the problem. Actually iteration (5) may not offer any advantages over the direct implementation of the inverse filter of (16), if no other features of the iterative algorithms are used, as will be explained later. One possible advantage of (5) over (16) is that the noise amplification in the restored image can be controlled by terminating the iteration before convergence, which represents another form of regularization, as will also be demonstrated experimentally. An iteration which will converge to the inverse solution of (2) for any $d(n_1, n_2)$ is described in the next section.

3.4 Reblurring

The degradation equation (2) can be modified so that the successive approximations iteration converges for a larger class of degradations. That is, the observed data $g(n_1, n_2)$ are first filtered (reblurred) by a system with impulse response $d^*(-n_1, -n_2)$, where $*$ denotes complex conjugation. Since circular convolutions have been adopted, the impulse response of the degradation system is equal to $\tilde{d}((N_1 - n_1)_{N_1}, (N_2 - n_2)_{N_2})$, where $(\cdot)_{N_i}$ denotes modulo N_i operation, assuming the images are of size $N_1 \times N_2$ pixels. The degradation equation (2) therefore becomes

$$\begin{aligned} \tilde{g}(n_1, n_2) &= g(n_1, n_2) * d^*(-n_1, -n_2) \\ &= d^*(-n_1, -n_2) * d(n_1, n_2) * f(n_1, n_2) \\ &= \tilde{d}(n_1, n_2) * f(n_1, n_2). \end{aligned} \quad (17)$$

If we follow the same steps as in the previous section substituting $g(n_1, n_2)$ by $\tilde{g}(n_1, n_2)$ and $d(n_1, n_2)$ by $\tilde{d}(n_1, n_2)$ the iteration providing a solution to (17) becomes

$$\begin{aligned} f_{k+1}(n_1, n_2) &= \beta d^*(-n_1, -n_2) * g(n_1, n_2) + (\delta(n_1, n_2) \\ &\quad - \beta \tilde{d}(n_1, n_2)) * f_k(n_1, n_2), \end{aligned} \quad (18)$$

with $f_0(n_1, n_2) = 0$. Following similar steps to the ones shown in the previous section we find that the restoration filter at the k -th iteration step is now given by

$$\begin{aligned} H_k(u, v) &= \beta \sum_{\ell=0}^{k-1} (1 - \beta |D(u, v)|^2)^\ell D^*(u, v) \\ &= \beta \frac{1 - (1 - \beta |D(u, v)|^2)^k}{1 - (1 - \beta |D(u, v)|^2)} D^*(u, v). \end{aligned} \quad (19)$$

Therefore, the sufficient condition for convergence, corresponding to condition (9), becomes

$$|1 - \beta |D(u, v)|^2| < 1, \quad \text{or} \quad 0 < \beta < \frac{2}{\max_{u, v} |D(u, v)|^2}. \quad (20)$$

In this case

$$\lim_{k \rightarrow \infty} H_k(u, v) = \begin{cases} \frac{1}{D(u, v)}, & D(u, v) \neq 0 \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

3.5 Experimental Results

In this section the performance of the iterative image restoration algorithms presented so far is demonstrated experimentally. We use a relatively simple degradation model in order to clearly analyze the behavior of the restoration filters. The degradation is due to one-dimensional (1D) horizontal motion between the camera and the scene, due, for example, to camera panning or fast object motion. The impulse response of the degradation system is given by

$$d(n_1, n_2) = \begin{cases} \frac{1}{L}, & -\frac{L-1}{2} \leq n_1 \leq \frac{L-1}{2}, L \text{ odd, } n_2 = 0 \\ \frac{1}{L}, & -\frac{L}{2} + 1 \leq n_1 \leq \frac{L}{2}, L \text{ even, } n_2 = 0 \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

The blurred signal-to-noise ratio (BSNR) is typically used in the restoration community to measure the degree of the degradation (blur plus additive noise). This figure is given by

$$BSNR = 10 \log_{10} \frac{\sigma_{Df}^2}{\sigma_v^2}, \quad (23)$$

where σ_{Df}^2 and σ_v^2 are respectively the variance of the blurred image and the additive noise.

For the purpose of objectively testing the performance of image restoration algorithms, the improvement SNR (ISNR) is often used. This metric using the restored image at the k -th iteration step is given by

$$ISNR = 10 \log_{10} \frac{\|f - g\|^2}{\|f - f_k\|^2}. \quad (24)$$

Obviously, this metric can only be used for simulation cases when the original image is available. While mean squared error (MSE) metrics such as ISNR do not always reflect the visual quality of the restored image, they serve to provide an objective standard by which to compare different techniques. However, in all cases presented here, it is important to consider the behavior of the various algorithms from the viewpoint of ringing and noise amplification, which can be a key indicator of improvement in quality for subjective comparisons of restoration algorithms.

In Fig. 2a the image blurred by the 1D motion blur of extent 8 pixels ($L=8$ in Eq. (22)) is shown, along with $|D(u, 0)|$, a slice of the magnitude of the 256×256 -point DFT of $d(n_1, n_2)$ (notice that all slices of the DFT are the same, independently of the value of v). No noise has been added. The extent of the blur and the size of the DFT were chosen in such a way that exact zeros exist in $D(u, v)$. The next three images represent the restored images using (18) with $\beta=1.0$, along with $|H_k(u, 0)|$ in (19), after 20, 50, and 465 iterations (at convergence). The criterion

$$\frac{\sum_{n_1, n_2} (f_{k+1}(n_1, n_2) - f_k(n_1, n_2))^2}{\sum_{n_1, n_2} (f_k(n_1, n_2))^2} \leq 10^{-8} \quad (25)$$

is used for terminating the iteration. Notice that (5) is not guaranteed to converge for this particular degradation since $D(u, v)$ takes negative values. The restored image of Fig. 2e is the result of the direct implementation of the pseudo-inverse filter, which can be thought of as the result of the iterative restoration algorithm after infinitely many iterations assuming infinite precision arithmetic. The corresponding ISNRs are: 4.03 dB (Fig. 2b), 6.22 dB (Fig. 2c), 11.58 dB (Fig. 2d), and 15.50 dB (2e). Finally, the normalized residual error shown in (25), versus the number of iterations is shown in Fig. 3. The iteration steps at which the restored images are shown in the previous figure are indicated by circles.

We repeat the same experiment when noise is added to the blurred image, resulting in a BSNR of 20 dB, as shown in Fig. 4a. The restored images after 20 iterations (ISNR = 1.83 dB), 50 iterations (ISNR = -0.40 dB), and at convergence after 1376 iterations (ISNR = -9.06 dB) are shown respectively in Figs. 4 b, c, and d. Finally, the restoration based on

the direct implementation of the pseudo-inverse filter is shown in Fig. 4e. The iterative algorithm converges slower in this case.

What becomes evident from these experiments is that:

- As expected, for the noise-free case the visual quality, as well as, the objective quality in terms of ISNR, of the restored images increase as the number of iterations increases.
- For the noise-free case the inverse filter outperforms the iterative restoration filter. Based on this experiment there is no reason to implement this particular filter iteratively, except possibly for computational reasons.
- For the noisy-blurred image the noise is amplified and the ISNR decreases as the number of iterations increases. Noise completely dominates the image restored by the pseudo-inverse filter. In this case, the iterative implementation of the restoration filter offers the advantage that the number of iterations can be used to control the amplification of the noise, which represents a form of regularization. The restored image, for example, after 50 iterations (Fig. 4c) represents a reasonable restoration.
- The iteratively restored image exhibits noticeable *ringing artifacts*, which will be further analyzed below. Such artifacts can be masked by noise, as demonstrated, for example, with the image in Fig. 4d.

3.5.1 Ringing Artifacts

Let us compare the magnitudes of the frequency response of the restoration filter after 465 iterations (Fig. 2d) and the inverse filter (Fig. 2e). First of all, it is clear that the existence of spectral zeros in $D(u, v)$ does not cause any difficulty in the determination of the restoration filter in both cases, since the restoration filter is also zero at these frequencies. The main difference is that the values of $|H(u, v)|$, the magnitude of the frequency response of the inverse filter, at frequencies close to the zeroes of $D(u, v)$ are considerably larger than the corresponding values of $|H_k(u, v)|$. This is because the values of $H_k(u, v)$ are approximated by a series according to (19). The important term in this series is $(1 - \beta|D(u, v)|^2)$, since it determines whether the iteration converges or not (sufficient condition). Clearly, this term for values of $D(u, v)$ close to zero is close to one, and therefore, it approaches zero much slower when raised to the power of k , the number of iterations, than the terms for which $D(u, v)$ assumes larger values and therefore the term $(1 - \beta|D(u, v)|^2)$ is close to zero. This means that each frequency component is restored independently and with different convergence rates. Clearly the larger the values of β the faster the convergence.

Let us denote by $h(n_1, n_2)$ the impulse response of the restoration filter and define

$$h_{all}(n_1, n_2) = d(n_1, n_2) * h(n_1, n_2). \quad (26)$$

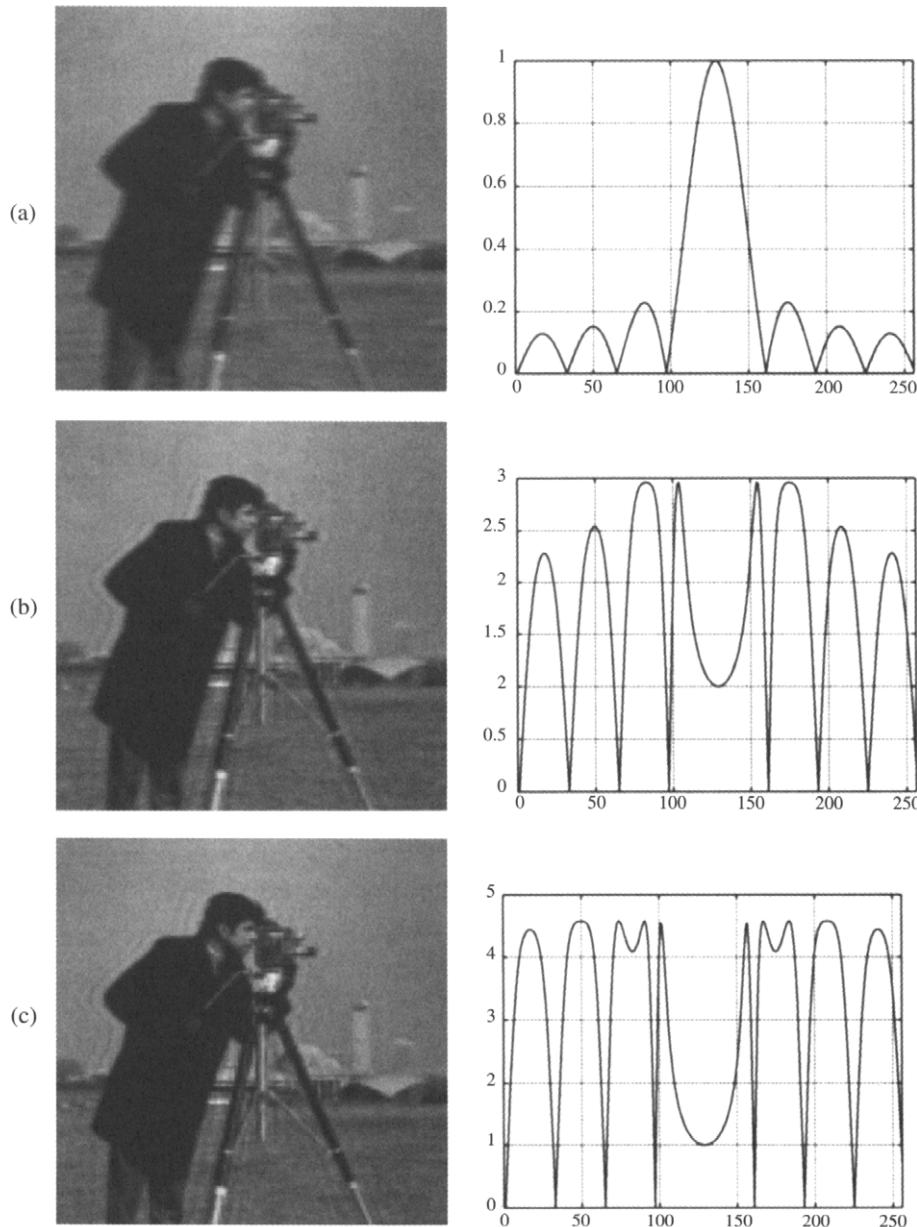


FIGURE 2 (a) Blurred image by a 1D motion blur over 8 pixels and the corresponding magnitude of the frequency response of the degradation system; (b)–(d): images restored by iteration (18), after 20 iterations (ISNR = 4.03dB), 50 iterations (ISNR = 6.22dB) and at convergence after 465 iterations (ISNR = 11.58dB), and the corresponding magnitude of $H_k(u, 0)$ in (19); (e) image restored by the direct implementation of the generalized inverse filter in (16) (ISNR = 15.50dB), and the corresponding magnitude of the frequency response of the restoration filter.

Ideally, $h_{all}(n_1, n_2)$ should be equal to an impulse, or its DFT $H_{all}(u, v)$ should be a constant, that is, the restoration filter is precisely undoing what the degradation system did. Due to the spectral zeros, however, in $D(u, v)$, $H_{all}(u, v)$ deviates from a constant. For the particular example under consideration $|H_{all}(u, 0)|$ is shown in Figs. 5a and 5c, for the inverse filter and the iteratively implemented inverse filter by (18), respectively. In Figs. 5b and 5d the corresponding impulse responses are shown. Due to the periodic zeros of $D(u, v)$ in this particular case, $h_{all}(n_1, n_2)$ consists of the sum of

an impulse and an impulse train (of period 8 samples). The deviation from a constant or an impulse is greater with the iterative restoration filter than with the direct inverse filter.

Now, in the absense of noise the restored image $\hat{f}(n_1, n_2)$ is given by

$$\hat{f}(n_1, n_2) = h_{all}(n_1, n_2) * f(n_1, n_2). \quad (27)$$

Clearly, due to the shape of $h_{all}(n_1, n_2)$ shown in Figs. 5b and 5d (only $h_{all}(n_1, 0)$ is shown, since it is zero for the rest of

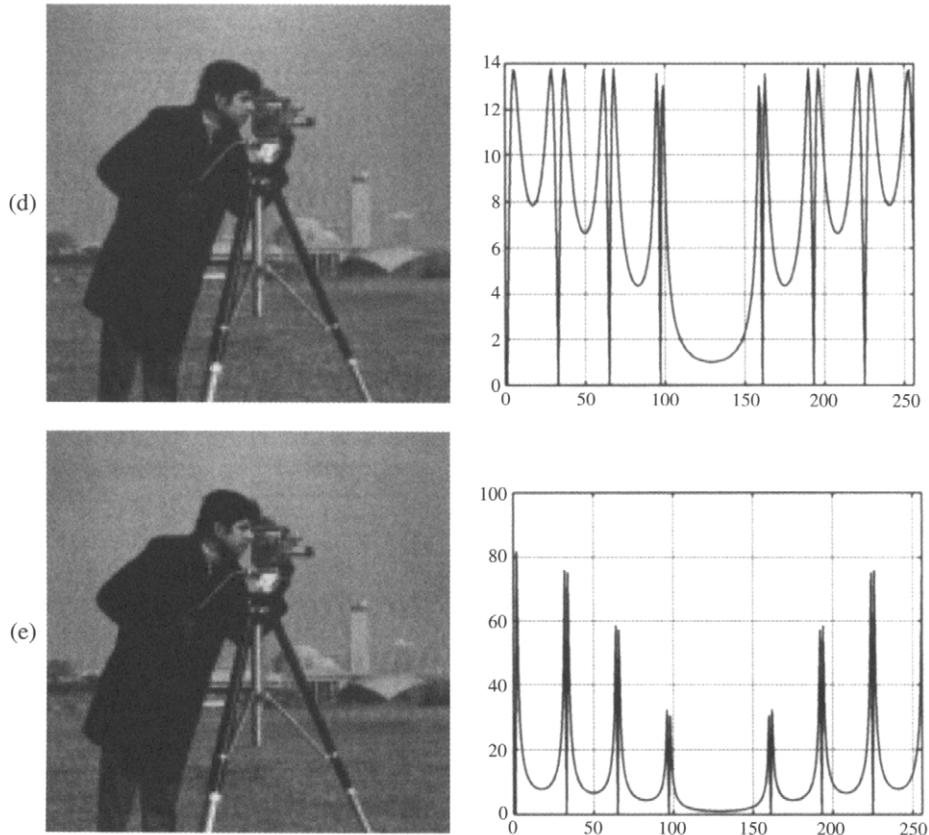


FIGURE 2 (Continued).

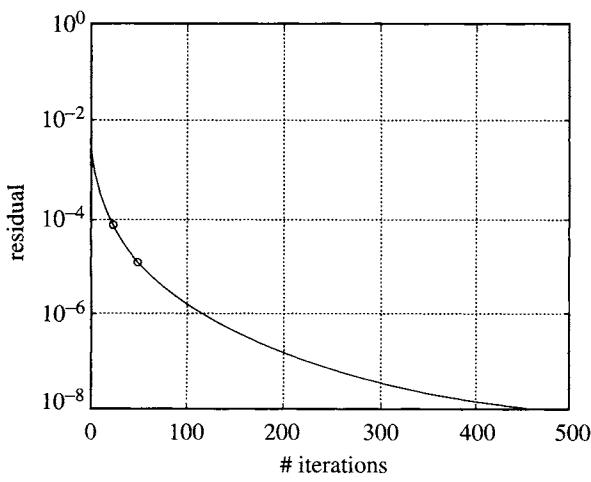


FIGURE 3 Normalized residual error as a function of the number of iterations.

the values of n_2), the existence of the periodic train of impulses gives rise to ringing. In the case of the inverse filter (Fig. 5b) the impulses of the train are small in magnitude and therefore ringing is not visible. In the case of the iterative filter, however, the few impulses close to zero have larger amplitude and therefore ringing is noticeable in this case.

4 Matrix-Vector Formulation

The presentation so far has followed a rather simple and intuitive path. It hopefully demonstrated some of the issues involved in developing and implementing an iterative algorithm. In this section we present the matrix-vector formulation of the degradation process and the restoration iteration. More general results are therefore obtained, since now the degradation can be spatially varying, while the restoration filter may be spatially varying as well, but even nonlinear. The degradation actually can be nonlinear as well (of course it is not represented by a matrix in this case), but we do not focus on this case, although most of the iterative algorithms discussed below would be applicable.

What became clear from the previous sections is that in applying the successive approximations iteration, the restoration problem to be solved is brought first into the form of finding the root of a function [see (3)]. In other words, a solution to the restoration problem is sought which satisfies

$$\Phi(\mathbf{f}) = 0, \quad (28)$$

where $\mathbf{f} \in \mathcal{R}^N$ is the vector representation of the signal resulting from the stacking or ordering of the original signal,

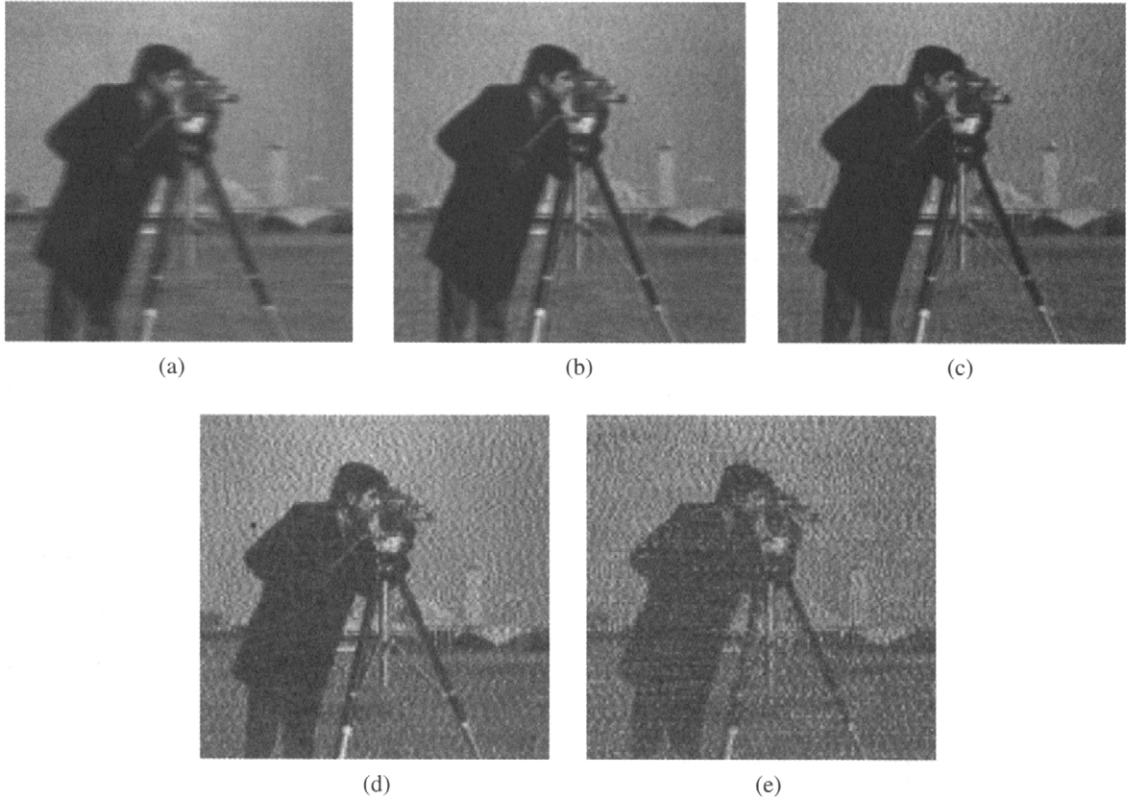


FIGURE 4 (a) Noisy-blurred image; 1D motion blur over 8 pixels, BSNR = 20dB; (b)–(d): images restored by iteration (18), after 20 iterations (ISNR = 1.83 dB), 50 iterations (ISNR = -0.30 dB), and at convergence after 1376 iterations (ISNR = -9.06 dB); (e) image restored by the direct implementation of the generalized inverse filter in (16) (ISNR = -12.09 dB).

and $\Phi(\mathbf{f})$ represents a nonlinear in general function. The row-by-row from left-to-right stacking of an image, is typically referred to as *lexicographic ordering*. For a 256×256 image, for example, vector \mathbf{f} is of dimension $64K \times 1$.

Then the successive approximations iteration which might provide us with a solution to (28) is given by

$$\mathbf{f}_{k+1} = \mathbf{f}_k + \beta\Phi(\mathbf{f}_k) = \Psi(\mathbf{f}_k), \quad (29)$$

with $\mathbf{f}_0 = 0$. Clearly if \mathbf{f}^* is a solution to $\Phi(\mathbf{f}) = 0$, i.e., $\Phi(\mathbf{f}^*) = 0$, then \mathbf{f}^* is also a fixed point of the above iteration, that is, $\mathbf{f}_{k+1} = \mathbf{f}_k = \mathbf{f}^*$. However, as was discussed in the previous section, even if \mathbf{f}^* is the unique solution to (28), this does not imply that iteration (29) will converge. This again underlines the importance of convergence when dealing with iterative algorithms. The form iteration (29) takes for various forms of the function $\Phi(\mathbf{f})$ is examined next.

4.1 Basic Iteration

From (1) when the noise is ignored, the simplest possible form $\Phi(\mathbf{f})$ can take is,

$$\Phi(\mathbf{f}) = \mathbf{g} - \mathbf{D}\mathbf{f} \quad (30)$$

Then (29) becomes

$$\mathbf{f}_{k+1} = \beta\mathbf{g} + (\mathbf{I} - \beta\mathbf{D})\mathbf{f}_k = \beta\mathbf{g} + \mathbf{G}_1\mathbf{f}_k, \quad (31)$$

where \mathbf{I} is the identity operator.

4.2 Least-Squares Iteration

According to the least-squares approach, a solution to (1) is sought by minimizing

$$M(\mathbf{x}) = \|\mathbf{g} - \mathbf{D}\mathbf{f}\|^2. \quad (32)$$

A necessary condition for $M(\mathbf{f})$ to have a minimum is that its gradient with respect to \mathbf{f} is equal to zero. That is, in this case

$$\Phi(\mathbf{f}) = \nabla_{\mathbf{f}} M(\mathbf{f}) = \mathbf{D}^T(\mathbf{g} - \mathbf{D}\mathbf{f}) = 0, \quad (33)$$

where T denotes the transpose of a matrix or vector. Application of iteration (29) then results in

$$\mathbf{f}_{k+1} = \beta\mathbf{D}^T\mathbf{g} + (\mathbf{I} - \beta\mathbf{D}^T\mathbf{D})\mathbf{f}_k = \beta\mathbf{D}^T\mathbf{g} + \mathbf{G}_2\mathbf{f}_k = \mathbf{T}_2\mathbf{f}_k. \quad (34)$$

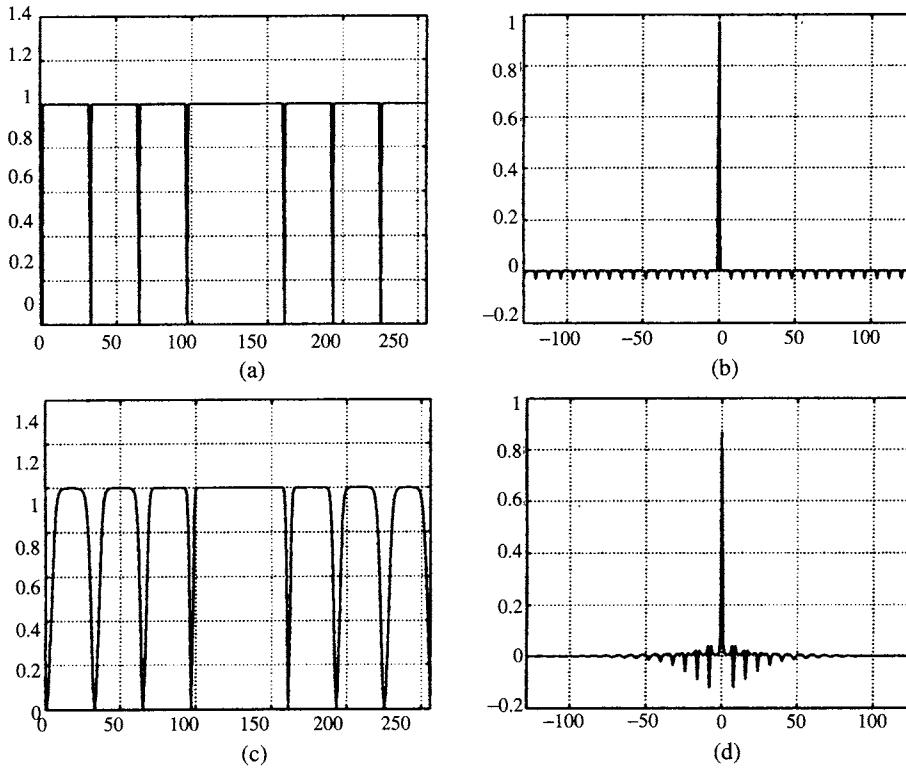


FIGURE 5 (a) $|H_{all}(u, 0)|$ for direct implementation of the inverse filter; (c) $|H_{all}(u, 0)|$ for the iterative implementation of the inverse filter; (b), (d): $h_{all}(n, 0)$ corresponding to Figs. 5a and 5c.

It is mentioned here that the matrix-vector representation of an iteration does not necessarily determine the way the iteration is implemented. In other words, the pointwise version of the iteration may be more efficient from the implementation point of view, than the matrix-vector form of the iteration. Now when (2) is used to form the matrix-vector equation $\mathbf{g} = \mathbf{D}\mathbf{f}$, matrix \mathbf{D} is a block-circulant matrix [2]. A square matrix is circulant when a circular shift of one row produces the next row, and the circular shift of the last row produces the first row. A square matrix is block-circulant when it consists of circular submatrices, which when circularly shifted produce the next row of circulant matrices. This implies that the singular values of \mathbf{D} are the DFT values of $d(n_1, n_2)$, and the eigenvectors are the complex exponential basis functions of the DFT. Iterations (31) and (34) can therefore be written in the discrete frequency domain, and they become identical to iterations (6) and the frequency domain version of iteration (18), respectively [14].

4.3 Constrained Least-Squares Iteration

The image restoration problem is an ill-posed problem, which means that matrix \mathbf{D} is ill-conditioned. A regularization method replaces an ill-posed problem by a well-posed problem, whose solution is an acceptable approximation to the solution of the ill-posed problem [25]. Most regularization approaches transform the original inverse problem into

a constrained optimization problem. That is, a functional needs to be optimized with respect to the original image, and possibly other parameters. By using the necessary condition for optimality, the gradient of the functional with respect to the original image is set equal to zero, therefore determining the mathematic form of $\Phi(\mathbf{f})$. The successive approximations iteration becomes in this case a gradient method with a fixed step (determined by β).

As an example, a restored image is sought as the result of the minimization of [10]

$$\|\mathbf{C}\mathbf{f}\|^2 \quad (35)$$

subject to the constraint that

$$\|\mathbf{g} - \mathbf{D}\mathbf{f}\|^2 \leq \epsilon^2. \quad (36)$$

Operator \mathbf{C} is a high-pass operator. The meaning then of the minimization of $\|\mathbf{C}\mathbf{f}\|^2$ is to constrain the high frequency energy of the restored image, therefore requiring that the restored image is smooth. On the other hand, by enforcing inequality (36) the fidelity to the data is preserved.

Following the Lagrangian approach which transforms the constrained optimization problem into an unconstrained one, the following functional is minimized

$$M(\alpha, \mathbf{f}) = \|\mathbf{D}\mathbf{f} - \mathbf{g}\|^2 + \alpha\|\mathbf{C}\mathbf{f}\|^2. \quad (37)$$

The necessary condition for a minimum is that the gradient of $M(\alpha, \mathbf{f})$ is equal to zero. That is, in this case

$$\Phi(\mathbf{f}) = \nabla_{\mathbf{f}} M(\alpha, \mathbf{f}) = (\mathbf{D}^T \mathbf{D} + \alpha \mathbf{C}^T \mathbf{C})\mathbf{f} - \mathbf{D}^T \mathbf{g}, \quad (38)$$

is used in iteration (29). The determination of the value of the regularization parameter α is a critical issue in regularized restoration, since it controls the trade-off between fidelity to the data and smoothness of the solution, and therefore the quality of the restored image. A number of approaches for determining its value are presented and compared in [9].

Since the restoration filter resulting from (38) is widely used it is worth looking further into its properties. When the degradation matrices \mathbf{D} and \mathbf{C} are block-circulant (38) the resulting successive approximations iteration can be written in the discrete frequency domain. The iteration takes the form

$$\begin{aligned} F_{k+1}(u, v) &= \beta D^*(u, v)G(u, v) + (1 - \beta(|D(u, v)|^2 \\ &\quad + \alpha|C(u, v)|^2))F_k(u, v), \end{aligned} \quad (39)$$

where $C(u, v)$ represents the 2D DFT of the impulse response of a high-pass filter, such as the 2D Laplacian. Following steps similar to the ones presented in Section 3.3, it is straightforward to verify that in this case the restoration filter at the k -th iteration step is given by

$$H_k(u, v) = \beta \sum_{\ell=0}^{k-1} (1 - \beta(|D(u, v)|^2 + \alpha|C(u, v)|^2))^{\ell} D^*(u, v). \quad (40)$$

Clearly if

$$|1 - \beta(|D(u, v)|^2 + \alpha|C(u, v)|^2)| < 1 \quad (41)$$

then

$$\begin{aligned} \lim_{k \rightarrow \infty} H_k(u, v) &= \lim_{k \rightarrow \infty} \beta \frac{1 - (1 - \beta(|D(u, v)|^2 + \alpha|C(u, v)|^2))^k}{1 - (1 - \beta(|D(u, v)|^2 + \alpha|C(u, v)|^2))} D^*(u, v) \\ &= \frac{D^*(u, v)}{|D(u, v)|^2 + \alpha|C(u, v)|^2}. \end{aligned} \quad (42)$$

Notice that condition (41) is not satisfied at the frequencies for which $H_d(u, v) = |D(u, v)| + \alpha|C(u, v)|^2 = 0$. It is therefore now not the zeros of the degradation matrix which need to be considered, but the zeros of the regularized matrix, with DFT values $H_d(u, v)$. Clearly if $H_d(u, v)$ is zero at certain frequencies, this means that both $D(u, v)$ and $C(u, v)$ are zero

at these frequencies. This demonstrates the purpose of regularization, which is to remove the zeros of $D(u, v)$ without altering the rest of its values, or in general to make the matrix $\mathbf{D}^T \mathbf{D} + \alpha \mathbf{C}^T \mathbf{C}$ better conditioned than the matrix $\mathbf{D}^T \mathbf{D}$.

For the frequencies at which $H_d(u, v) = 0$

$$\lim_{k \rightarrow \infty} H_k(u, v) = \lim_{k \rightarrow \infty} k \cdot \beta \cdot D^*(u, v) = 0, \quad (43)$$

since $D^*(u, v) = 0$.

4.3.1 Experimental Results

The noisy and blurred image of Fig. 4a (1D motion blur over 8 pixels, BSNR = 20 dB) is now restored using iteration (39), with $\alpha = 0.01$, $\beta = 1.0$, and C the 2D Laplacian operator. It is mentioned here that the regularization parameter is chosen to be equal to $\frac{\sigma_e^2}{\sigma_{df}^2}$, as determined by a set theoretic restoration approach presented in [16]. The restored images after 20 iterations (ISNR = 2.12 dB), 50 iterations (ISNR = 0.98 dB) and at convergence after 330 iterations (ISNR = -1.01 dB) with the corresponding $|H_k(u, v)|$ in (40), are shown respectively in Figs. 6 a,b,c. In Fig. 6d the restored image (ISNR = -1.64 dB) by the direct implementation of the constrained least-squares filter in (42) is shown, along with the magnitude of the frequency response of the restoration filter. It is clear now by comparing the restoration filters of Figs. 2d and 6c and 2e and 6d, that the high frequencies have been suppressed, due to regularization, that is the addition in the denominator of the filter of the term $\alpha|C(u, v)|^2$. Due to the iterative approximation of the constrained least-squares filter, however, the two filters shown in Figs. 6c and 6d differ primarily in the vicinity of the low frequency zeros of $D(u, v)$. Ringing is still present, as it can be primarily seen in Figs. 6a and 6b, although is not as visible in Figs. 6c and 6d. Due to regularization the results in Figs. 6c and 6d are preferred over the corresponding results with no regularization ($\alpha = 0.0$), shown in Figs. 4d and 4e.

The value of the regularization parameter is very critical for the quality of the restored image. The restored images with three different values of the regularization parameter are shown in Figs. 7 a-c, corresponding to $\alpha = 1.0$ (ISNR = 2.4 dB), $\alpha = 0.1$ (ISNR = 2.96 dB), and $\alpha = 0.01$ (ISNR = -1.80 dB). The corresponding magnitudes of the error images, i.e., $|\text{original} - \text{restored}|$, scaled linearly to the 32–255 range are shown in Figs. 7 d-f. What is observed is that for large values of α the restored image is “smooth” while the error image contains the high frequency information of the original image (large bias of the estimate), while as α decreases the restored image becomes more noisy and the error image takes the appearance of noise (large variance of the estimate). It has been shown in [9] that the bias of the constrained least-squares estimate is a monotonically increasing function of the regularization parameter, while the variance of the estimate is a monotonically decreasing function of the estimate. This implies that the mean-squared error (MSE) of the estimate,

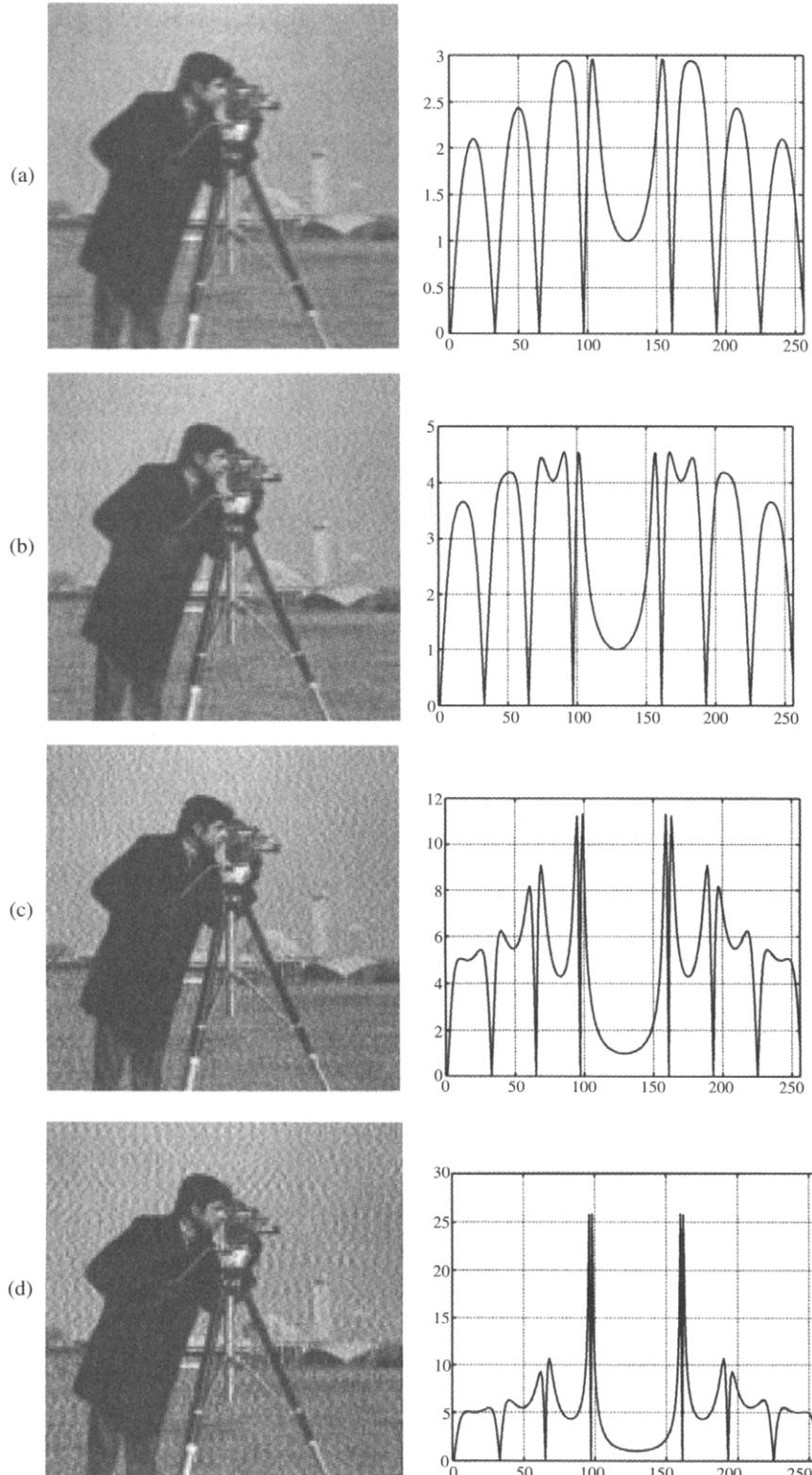


FIGURE 6 Restoration of the noisy-blurred image in Fig. 5a (motion over 8 pixels, BSNR = 20dB); (a)–(c): images restored by iteration (39), after 20 iterations (ISNR = 2.12dB), 50 iterations (ISNR = 0.98dB) and at convergence after 330 iterations (ISNR = −1.01dB), and the corresponding $|H_k(u, 0)|$ in (40); (d): images restored by the direct implementation of the constrained least-squares filter (ISNR = −1.64dB), and the corresponding magnitude of the frequency response of the restoration filter (Eq. (42)).

the sum of the bias and the variance, has a unique minimum for a specific value of α .

4.4 Spatially Adaptive Iteration

Spatially adaptive image restoration is the next natural step in improving the quality of the restored images. There are various ways to argue the introduction of spatial adaptivity, the most commonly used ones being the nonhomogeneity or nonstationarity of the image field and the properties of the human visual system. In either case, the functional to be minimized takes the form [4, 14]

$$M(\alpha, \mathbf{f}) = \|\mathbf{D}\mathbf{f} - \mathbf{g}\|_{\mathbf{W}_1}^2 + \alpha \|\mathbf{C}\mathbf{f}\|_{\mathbf{W}_2}^2, \quad (44)$$

in which case

$$\begin{aligned} \Phi(\mathbf{f}) &= \nabla_{\mathbf{f}} M(\alpha, \mathbf{f}) \\ &= (\mathbf{D}^T \mathbf{W}_1^T \mathbf{W}_1 \mathbf{D} + \alpha \mathbf{C}^T \mathbf{W}_2^T \mathbf{W}_2 \mathbf{C}) \mathbf{f} - \mathbf{D}^T \mathbf{W}_1 \mathbf{g}. \end{aligned} \quad (45)$$

The choice of the diagonal weighting matrices \mathbf{W}_1 and \mathbf{W}_2 can be justified in various ways. In [14] both matrices are

determined by the diagonal noise visibility matrix \mathbf{V} [1]. That is, $\mathbf{W}_1 = \mathbf{V}^T \mathbf{V}$ and $\mathbf{W}_2 = \mathbf{I} - \mathbf{V}^T \mathbf{V}$. The entries of \mathbf{V} take values between 0 and 1. They are equal to 0 at the edges (noise is not visible), equal to 1 at the flat regions (noise is visible) and take values in between at the regions with moderate spatial activity.

4.4.1 Experimental Results

The resulting successive approximations iteration from the use of $\Phi(\mathbf{f})$ in (45) has been tested with the noisy and blurred image we have been using so far in our experiments, which is shown in Fig. 4a. It should be emphasized here that although matrices \mathbf{D} and \mathbf{C} are block circulant, the iteration cannot be implemented in the discrete frequency domain, since the weight matrices \mathbf{W}_1 and \mathbf{W}_2 , are diagonal, but not circulant. Therefore, the iterative algorithm is implemented exclusively in the spatial domain, or by switching between the frequency domain (where the convolutions are implemented) and the spatial domain (where the weighting takes place). Clearly, from an implementation point of view the use of iterative algorithms offers a distinct advantage in this particular case.

The iteratively restored image with $\mathbf{W}_1 = 1 - \mathbf{W}_2$, $\alpha = 0.01$ and $\beta = 0.1$ is shown in Fig. 8a, at convergence after 381

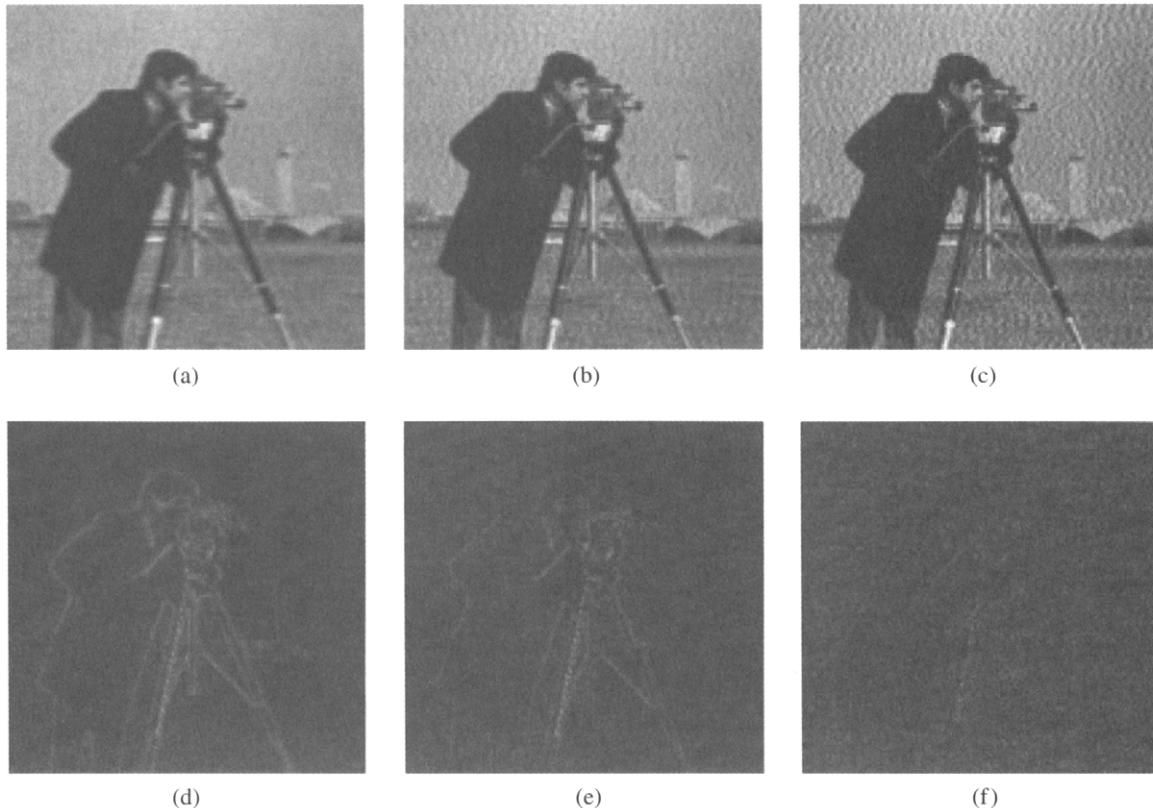


FIGURE 7 Direct constrained least-squares restorations of the noisy-blurred image in Fig. 5a (motion over 8 pixels, BSNR = 20dB) with α equal to: (a) 1, (b) 0.1, (c) 0.01; (d)–(f): corresponding $|\text{original}-\text{restored}|$ linearly mapped to the range [32, 255].

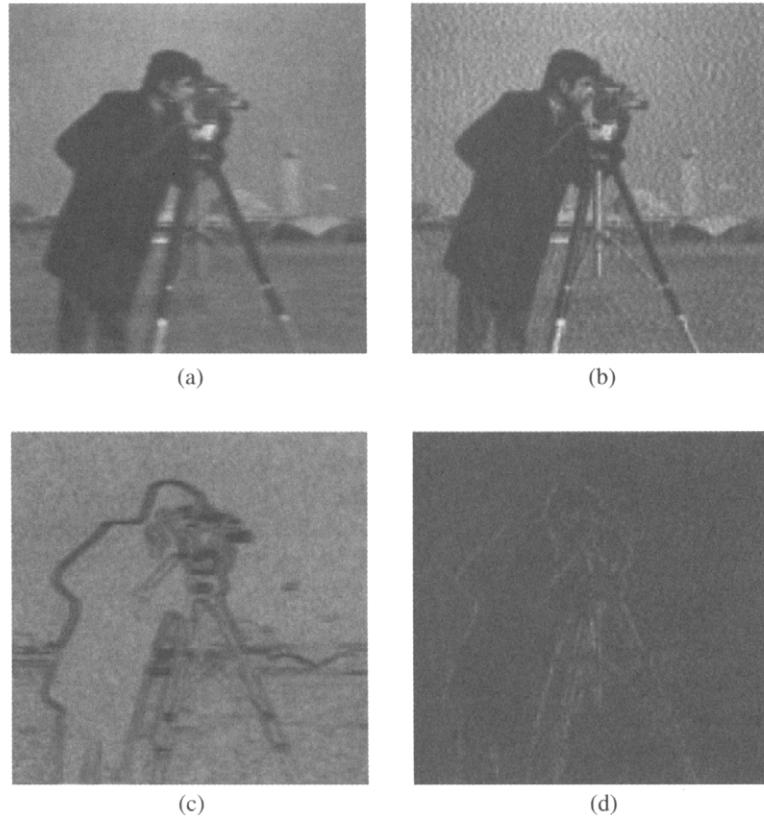


FIGURE 8 Restoration of the noisy-blurred image in Fig. 5a (motion over 8 pixels, BSNR = 20dB), using (a) the adaptive algorithm of (45); (b) the nonadaptive algorithm of iteration (39); (c) values of the weight matrix in Eq. (46); (d) amplitude of the difference between Figs. (a) and (b) linearly mapped to the range [32, 255].

iterations and ISNR = 0.61 dB. The entries of the diagonal matrix \mathbf{W}_2 , denoted by $w_2(i)$ are computed according to

$$w_2(i) = \frac{1}{\theta\sigma^2(i) + 1}, \quad (46)$$

where $\sigma^2(i)$ is the local variance at the ordered i -th pixel location, and θ a tuning parameter. The resulting values of $w_2(i)$ are linearly mapped into the [0,1] range. These weights computed from the degraded image are shown in Fig. 8c, linearly mapped to the [32, 255] range, using a 3×3 window to find the local variance and $\theta = 0.001$. The image restored by the non-adaptive algorithm, that is, $\mathbf{W}_1 = \mathbf{W}_2 = \mathbf{I}$ and the rest of the parameters the same, is shown in Fig. 8b (ISNR = -0.20 dB). The absolute value of the difference between the images in Figs. 8a and 8b, linearly mapped in the [32, 255] range is shown in Fig. 8d. It is clear that the two algorithms differ primarily at the vicinity of edges, where the smoothing is downweighted or disabled with the adaptive algorithm. Spatially adaptive algorithms in general can greatly improve the restoration results, since they can adopt to the local characteristics of each image.

5 Use of Constraints

Iterative signal restoration algorithms regained popularity in the 1970s due to the realization that improved solutions can be obtained by incorporating prior knowledge about the solution into the restoration process. For example, we may know in advance that \mathbf{f} is bandlimited or space-limited, or we may know on physical grounds that \mathbf{f} can only have nonnegative values. A convenient way of expressing such prior knowledge is to define a constraint operator \mathcal{C} , such that

$$\mathbf{f} = \mathcal{C}\mathbf{f}, \quad (47)$$

if and only if \mathbf{f} satisfies the constraint. In general, \mathcal{C} represents the concatenation of constraint operators. With the use of constraints, iteration (29) becomes [23]

$$\begin{aligned} \mathbf{f}_0 &= 0, \\ \tilde{\mathbf{f}}_k &= \mathcal{C}\mathbf{f}_k, \\ \mathbf{f}_{k+1} &= \Psi(\tilde{\mathbf{f}}_k). \end{aligned} \quad (48)$$

As already mentioned, a number of recovery problems, such as the bandlimited extrapolation problem, and the reconstruction from phase or magnitude problem, can be solved with the use of algorithms of the form (48), by appropriately describing the distortion and constraint operators [23].

The *contraction mapping theorem* [22] usually serves as a basis for establishing convergence of iterative algorithms. Sufficient conditions for the convergence of the algorithms presented in Section 4 are presented in [14]. Such conditions become identical to the ones derived in Section 3, when all matrices involved are block-circulant. When constraints are used, the sufficient condition for convergence of the iteration is that at least one of the operators C and Ψ is contractive while the other is nonexpansive. Usually, it is harder to prove convergence and determine the convergence rate of the constrained iterative algorithm, taking also into account that some of the constraint operators are nonlinear, such as, the positivity constraint operator.

5.1 Experimental Results

We demonstrate the effectiveness of the positivity constraint with the use of a simple example. An 1D impulsive signal is shown in Fig. 9a. Its degraded version by a motion blur over 8 samples is shown in Fig. 9b. The blurred signal is restored by iteration (18) ($\beta=1.0$) with the use of the positivity constraint (Fig. 9c, 370 iterations, ISNR = 41.35), and without the use of the positivity constraint (Fig. 9d, 543 iterations,

ISNR = 11.05). The application of the positivity constraint, which represents a non-expansive mapping, simply sets to zero all negative values of the signal. Clearly a considerably better restoration is represented by Fig. 9c.

6 Additional Considerations

In the previous sections we dealt exclusively with the image restoration problem, as described by Eq. (1). As was mentioned in the introduction there is a plethora of inverse problems, i.e., problems described by Eq. (1), for which the iterative algorithms presented so far can be applied. Inverse problems are representative examples of more general recovery problems, i.e., problems for which information which is lost (due, for example, to the imperfections of the imaging system or the transmission medium, or the specific processing the signal is undergoing, such as compression), is attempted to be recovered. A critical step in solving any such problem is the modeling of the signals and systems involved, or in other words, the derivation of the degradation model. After this is accomplished the solution approach needs to be decided (of course these two steps do not need to be independent). In this chapter we dealt primarily with the image restoration problem under a deterministic formulation and a successive approximations based iterative solution approach. In the following three subsections we describe respectively some additional forms the successive

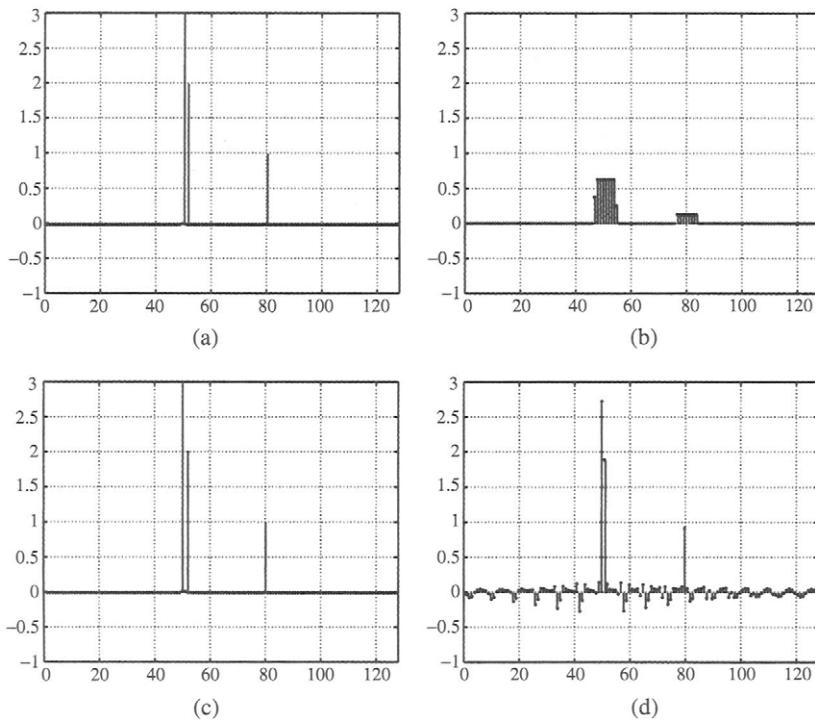


FIGURE 9 (a) original signal; (b) blurred signal by motion blur over 8 samples; signals restored by iteration (18); (c) with positivity constraint; (d) without positivity constraint.

approximations iteration can take, a stochastic modeling of the restoration problem which results in successive approximations type of iterations, and finally additional recent image recovery applications.

6.1 Other Forms of the Iterative Algorithm

The basic iteration presented in the previous sections can be extended in a number of ways. One such way is to utilize the partially restored image at each iteration step in evaluating unknown problem parameters or refining our prior knowledge about the original image. A critical such parameter which directly controls the quality of the restoration results, as was also experimentally demonstrated in Fig. 7, is the regularization parameter α in Eq. (37). As was already mentioned in Section 4.3, a number of approaches have appeared in the literature for the evaluation of α [9]. It depends on the value of $\|\mathbf{g} - \mathbf{D}\mathbf{f}\|^2$ or its upper bound ϵ in Eq. (36), but also on the value of $\|\mathbf{C}\mathbf{f}\|^2$ or an upper bound of it, or in other words on the value of \mathbf{f} . This dependency of α on the unknown original image \mathbf{f} is expressed explicitly in [12], by rewriting the functional to be minimized in Eq. (37) as

$$M(\alpha(\mathbf{f}), \mathbf{f}) = \|\mathbf{g} - \mathbf{D}\mathbf{f}\|^2 + \alpha(\mathbf{f})\|\mathbf{C}\mathbf{f}\|^2. \quad (49)$$

The desirable properties of $\alpha(\mathbf{f})$ and various functional forms it can take are investigated in detail in [12]. One such choice is given by

$$\alpha(\mathbf{f}) = \frac{\|\mathbf{g} - \mathbf{D}\mathbf{f}\|^2}{(1/\gamma) - \|\mathbf{C}\mathbf{f}\|^2}, \quad (50)$$

with γ constrained so that the denominator in Eq. (50) is positive. The successive approximations iteration in this case then becomes

$$\mathbf{f}_{k+1} = \mathbf{f}_k + \beta[\mathbf{D}^T \mathbf{g} - (\mathbf{D}^T \mathbf{D} + \alpha(\mathbf{f}_k) \mathbf{C}^T \mathbf{C}) \mathbf{f}_k]. \quad (51)$$

Sufficient conditions for the convergence of iteration (51) are derived in [12] in terms of the parameter γ , and also conditions which guarantee $M(\alpha(\mathbf{f}), \mathbf{f})$ to be convex (the relaxation parameter β can be set equal to one since it can be combined with the parameter γ). Iteration (51) represents a major improvement towards the solution of the restoration problem because (i) no prior knowledge, such as knowledge of the noise variance, is required for the determination of the regularization parameter, but instead such information is extracted from the partially restored image; and (ii) the determination of the regularization parameter does not constitute a separate, typically iterative step, but it is performed simultaneously with the restoration of the image. The performance of iteration (51) is studied in detail in [12] for

various forms of the functional $\alpha(\mathbf{f})$ and various initial conditions.

This framework of extracting information required by the restoration process at each iteration step from the partially restored image, has also been applied to the evaluation of the weights \mathbf{W}_1 and \mathbf{W}_2 in iteration (45) and in deriving algorithms which use a different iteration-dependent regularization parameter for each discrete frequency component [11].

Additional extinctions of the basic form of the successive approximations algorithm are represented by algorithms with higher rates of convergence [15, 21], algorithms with a relaxation parameter β which depends on the iteration step (steepest descent and conjugate gradient algorithms are examples of this), algorithm which depend on more than one previous restoration steps (multi-step algorithms), and algorithms which utilize the number of iterations as a means of regularizing the solution.

6.2 Hierarchical Bayesian Image Restoration

In the presentation so far we have assumed that the degradation and the images in Eq. (1) are deterministic and the noise only represents a stochastic signal. A different approach towards the derivation of the degradation model and a restoration solution is represented by the Bayesian paradigm. According to it, knowledge about the structural form of the noise and the structural behavior of the reconstructed image is used in forming respectively $p(\mathbf{g}|\mathbf{f}, \tau)$ and $p(\mathbf{f}|\delta)$, where $p(\cdot|\cdot)$ denotes a conditional probability density function (pdf). For example, the following conditional pdf is typically used to describe the structural form of the noise

$$p(\mathbf{g}|\mathbf{f}, \tau) = \frac{1}{Z_{noise}(\tau)} \exp\left[-\frac{1}{2}\tau\|\mathbf{g} - \mathbf{D}\mathbf{f}\|^2\right], \quad (52)$$

where $Z_{noise}(\tau) = (2\pi/\tau)^{N/2}$, with N , as mentioned earlier, the dimension of the vectors \mathbf{f} and \mathbf{g} . Smoothness constraints on the original image can be incorporated under the form of

$$p(\mathbf{f}|\delta) \propto \delta^{q/2} \exp\left\{-\frac{1}{2}\delta S(\mathbf{f})\right\}, \quad (53)$$

where $S(\mathbf{f})$ is a non negative quadratic form which usually corresponds to a conditional or simultaneous autoregressive model in the statistical community or to placing constraints on the first or second differences in the engineering community and q is the number of positive eigenvalues of S [19]. A form of $S(\mathbf{f})$ widely in the engineering community and also in this chapter is

$$S(\mathbf{f}) = \|\mathbf{C}\mathbf{f}\|^2,$$

with \mathbf{C} the Laplacian operator.

The parameters δ and τ are typically referred to as hyperparameters. If they are known, according to the Bayesian paradigm, the image $\mathbf{f}_{(\delta,\tau)}$ is selected as the restored image, defined by

$$\begin{aligned}\mathbf{f}_{(\delta,\tau)} &= \arg \max_{\mathbf{f}} p(\mathbf{f}|\delta)p(\mathbf{g}|\mathbf{f},\tau) \\ &= \arg \min_{\mathbf{f}} [\alpha S(\mathbf{f}) + \tau ||\mathbf{g} - \mathbf{D}\mathbf{f}||^2].\end{aligned}\quad (54)$$

If the hyperparameters are not known then they can be treated as random variables and the hierarchical Bayesian approach can be followed. It consists of two stages. In the first stage the conditional probabilities shown in Eqs. (52) and (53) are formed. In the second stage the hyperprior $p(\delta, \tau)$ is also formulated, resulting in the distribution $p(\delta, \tau, \mathbf{f}, \mathbf{g})$. With the so-called evidence analysis, $p(\delta, \tau, \mathbf{f}, \mathbf{g})$ is integrated over \mathbf{f} to give the likelihood $p(\delta, \tau|\mathbf{g})$, which is then maximized over the hyperparameters. Alternatively, with the maximum *a posteriori* (MAP) analysis, $p(\delta, \tau, \mathbf{f}, \mathbf{g})$ is integrated over δ and τ to obtain the true likelihood, which is then maximized over \mathbf{f} to obtain the restored image.

Although in some cases it would be possible to establish relationships between the hyperpriors, the following model of the global probability is typically used

$$p(\delta, \tau, \mathbf{f}, \mathbf{g}) = p(\delta)p(\tau)p(\mathbf{f}|\delta)p(\mathbf{g}|\mathbf{f}, \tau). \quad (55)$$

Flat or non-informative hyperpriors are used for $p(\delta)$ and $p(\tau)$ if no prior knowledge about the hyperpriors exists. If such knowledge exists, as an example, a gamma distribution can be used [19]. As expected, the form of these probability density functions impacts the subsequent calculations.

Clearly the hierarchical Bayesian analysis offers a methodical procedure to evaluate unknown parameters in the context of solving a recovery problem. A critical step in its application is the determination of $p(\delta)$ and $p(\tau)$ and the above mentioned integration of $p(\delta, \tau, \mathbf{f}, \mathbf{g})$ either over \mathbf{f} , or δ and τ . Both flat and gamma hyperpriors $p(\delta)$ and $p(\tau)$ have been considered in [19], utilizing both the evidence and MAP analyses. They resulted in iterative algorithms for the evaluation of δ , τ , and \mathbf{f} . The important connection between the hierarchical Bayesian approach and the iterative approach presented in Sec. 6.1 is that iteration (51) with $\alpha(\mathbf{f})$ given by Eq. (50) or any of the forms proposed in [11, 12] can now be derived by the hierarchical Bayesian analysis with the appropriate choice of the required hyperpriors and the integration method. It should be made clear that the regularization parameter α is equal to the ratio (τ/δ) . A related result has been obtained in [5] by deriving through a Bayesian analysis the same expressions for the iterative evaluation of the weight matrices \mathbf{W}_1 and \mathbf{W}_2 as in iteration (45) and Eq. (46). It is therefore significant that there is a precise interpretation of the framework briefly described in

the previous section, based on the stochastic modeling of the signals and the unknown parameters.

6.3 Additional Applications

In this chapter we have concentrated on the application of the successive approximations iteration to the image restoration problem. However, as already mentioned multiple times already a number of recovery problems can find solutions with the use of a successive approximations iteration. Two recovery problems which have been actively pursued in the last 10–15 years due to their theoretical challenge but also their commercial significance, are the removal of compression artifacts and resolution enhancement.

The problem of removing compression artifacts addresses the recovery of information lost due to the quantization of parameters during compression. More specifically, in the majority of existing image and video compression algorithms the image (or frame in an image sequence) is divided into square blocks which are processed independently from each other. The discrete cosine transform (DCT) of such blocks (representing either the image intensity when dealing with still images or intra-coding of video blocks or frames, or the displaced frame difference when dealing with inter-coding of video blocks or frames) is taken and the resulting DCT coefficients are quantized. As a result of this processing annoying blocking artifacts result, primarily at high compression ratios. A number of techniques have been developed for removing such blocking artifacts for both still images and video. For example, in [18, 27] the problem of removing the blocking artifacts is formulated as a recovery problem, according to which an estimate of the blocking artifact-free original image is estimated by utilizing the available quantized data, knowledge about the quantizer step size, and prior knowledge about the smoothness of the original image.

A deterministic formulation of the problem is followed in [27]. Two solutions are developed for the removal of blocking artifacts in still images. The first one is based on the CLS formulation and a successive approximations iteration is utilized for obtaining the solution. The second approach is based on the theory of projections onto convex sets (POCS), which has found applications in a number of recovery problems. The evidence analysis within the hierarchical Bayesian paradigm, mentioned above, is applied to the same problem in [18]. Expressions for the iterative evaluation of the unknown parameters and the reconstructed image are derived. The relationship between the CLS-iteration adaptive successive approximations solution and the hierarchical Bayesian solution discussed in the previous section is also applicable here.

Resolution enhancement (also referred to as super-resolution) is a problem which has also seen considerable activity recently (for a recent review see [6, 13] and references therein). It addresses the problem of increasing the resolution

of a single image utilizing multiple aliased low-resolution images of the same scene with sub-pixel shifts among them. It also addresses the problem of increasing the resolution of a video frame (and consequently the whole sequence) of a dynamic video sequence by utilizing a number of neighboring frames. In this case the shifts between any two frames are expressed by the motion field. The low resolution images and frames can be noisy and blurred (due to the image acquisition system), or compressed, which further complicates the problem. There are a number of potential applications of this technology. It can be utilized to increase the resolution of any instrument by creating a number of images of the same scene, but also to replace an expensive high resolution instrument by one or more low resolution ones, or it can serve as a compression mechanism. Some of the techniques developed in the literature address, in addition to the resolution enhancement problem, the simultaneous removal of blurring and compression artifacts, i.e., they combine the objectives of multiple application mentioned in this chapter. For illustration purposes consider the example shown in Fig. 10 [20]. In Fig. 10a the original high resolution image is shown. This image is blurred with a 4×4 non-causal uniform blur function and downsampled by a factor of 4 in each direction to generate 64 low resolution images with global sub-pixel shifts which are integer multiples of $1/4$ in each direction. Noise of the same variance was added to all low resolution images (resulting in SNR of 30 dB for this example). One of

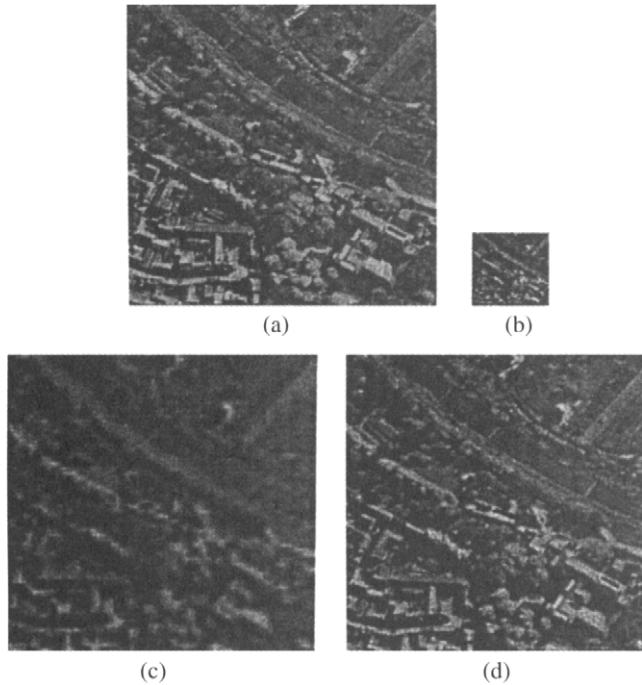


FIGURE 10 Resolution enhancement example: (a) original image; (b) one of the 16 low resolution images with $\text{SNR} = 30 \text{ dB}$ (all 16 images have the same SNR); (c) best bilinearly interpolated image; (d) estimated high resolution image by the hierarchical Bayesian approach [20].

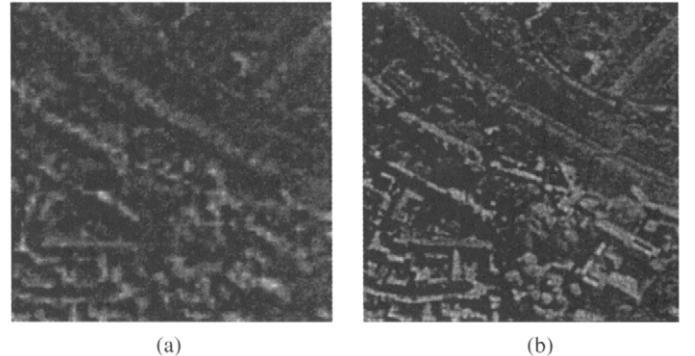


FIGURE 11 Resolution enhancement example: (a) best bilinearly interpolated low resolution image (the SNR for the low resolution images is at random either 10 dB or 20 dB or 30 dB); (b) estimated high resolution image by the hierarchical Bayesian approach (sixteen noise parameters) [20].

the low resolution images (the one with zero shifts in each direction) is shown in Fig. 10b. The best bilinearly interpolated image is shown in Fig. 10c. A hierarchical Bayesian approach is utilized in [20] in deriving iterative algorithms for estimating the unknown parameters (image model parameter similar to δ in Eq. (53) and the additive noise variance) and the high resolution image by utilizing the 16 low resolution images, assuming that the shifts and the blur are known. The resulting high resolution image is shown in Fig. 10d. The same experiment was performed but with a different amount of noise added to each low resolution image from a set of 10 dB, 20 dB, or 30 dB SNR. The best bilinearly interpolated image is shown in this case in Fig. 11a while the one resulting from the hierarchical Bayesian approach in Fig. 11b. Finally, the same last experiment was repeated with the resolution chart image. One of the 16 low resolution images is shown in Fig. 12a, while the one generated by the hierarchical Bayesian approach is shown in Fig. 12b. The hierarchical Bayesian approach was also used for the recovery of a high resolution sequence from low resolution and compressed observations [24].

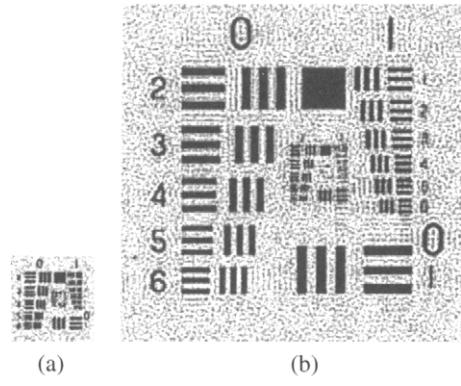


FIGURE 12 Resolution enhancement example: (a) one of the 16 low resolution images (the SNR for the low resolution images is at random either 10 dB or 20 dB or 30 dB); (b) estimated high resolution image by the hierarchical Bayesian approach [20].

7 Discussion

In this chapter we briefly described the application of the successive approximations-based class of iterative algorithms to the problem of restoring a noisy and blurred image. We presented and analyzed in some detail the simpler forms of the algorithm, and briefly described an iteration-adaptive form of the algorithm following a deterministic approach but also a hierarchical Bayesian approach. In addition we briefly described two other inverse problems, the removal of blocking artifacts and the enhancement of resolution, which have been solved using the techniques described in this chapter. With this presentation we have simply touched the “tip of the iceberg.” We only covered a small amount of the material on the topic. More sophisticated forms of iterative image restoration algorithms were left out, since they were deemed to be beyond the scope and the level of this chapter.

It is the hope and the expectation of the authors that the presented material will form a good introduction to the topic for the engineer or the student who would like to work in this area.

References

- [1] G. L. Anderson and A. N. Netravali, “Image Restoration Based on a Subjective Criterion,” *IEEE Trans. Sys. Man. and Cybern., SMC-6*, 845–853, Dec. 1976.
- [2] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice Hall, 1977.
- [3] M. R. Banham and A. K. Katsaggelos, “Digital Image Restoration,” *IEEE Signal Processing Magazine*, 14, 2, 24–41, March 1997.
- [4] J. Biemond, R. L. Lagendijk, and R. M. Mersereau, “Iterative Methods for Image Deblurring,” *Proc. IEEE*, 78, 5, 856–883, May 1990.
- [5] G. Chantas, N. P. Galatsanos, and A. Likas, “Non Stationary Bayesian Image Restoration,” *Proc. Int. Conf. on Pattern Recognition*, 4, 689–692, Aug. 23–26, 2004.
- [6] S. Chaudhuri, ed., *Super-Resolution Imaging*, Kluwer Academic Publishers, 2001.
- [7] G. Demoment, “Image Reconstruction and Restoration: Overview of Common Estimation Structures and Problems,” *IEEE Trans. Acoust., Speech, Signal Processing*, 37, 12, 2024–2036, Dec. 1989.
- [8] D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice Hall, 1984.
- [9] N. P. Galatsanos and A. K. Katsaggelos, “Methods for Choosing the Regularization Parameter and Estimating the Noise Variance in Image Restoration and their Relation,” *IEEE Trans. Image Processing*, 1, 322–336, July 1992.
- [10] B. R. Hunt, “The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computers,” *IEEE Trans. Computers*, C-22, 805–812, Sept. 1973.
- [11] M. G. Kang and A. K. Katsaggelos, “Frequency Domain Adaptive Iterative Image Restoration and Evaluation of the Regularization Parameter,” *Optical Engineering*, 33, 10, 3222–3232, Oct. 1994.
- [12] M. G. Kang and A. K. Katsaggelos, “General Choice of the Regularization Functional in Regularized Image Restoration,” *IEEE Trans. Image Processing*, 4, 5, 594–602, May 1995.
- [13] M. G. Kang and S. Chaudhuri, eds., “Super Resolution Image Reconstruction,” *IEEE Signal Processing Magazine*, 20, 3, May 2003.
- [14] A. K. Katsaggelos, “Iterative Image Restoration Algorithm,” *Optical Engineering*, 28, 7, 735–748, July 1989.
- [15] A. K. Katsaggelos and S. N. Efstratiadis, “A Class of Iterative Signal Restoration Algorithms,” *IEEE Trans. Acoust., Speech, Signal Processing*, 38, 778–786, May 1990 (reprinted in *Digital Image Processing*, R. Chellappa, ed., IEEE Computer Society Press).
- [16] A. K. Katsaggelos, J. Biemond, R. M. Mersereau, and R. W. Schafer, “A Regularized Iterative Image Restoration Algorithm,” *IEEE Trans. Signal Processing*, 39, 4, 914–929, April 1991.
- [17] A. K. Katsaggelos, editor, *Digital Image Restoration*, Springer Series in Information Sciences, 23, Springer-Verlag, Heidelberg, 1991.
- [18] J. Mateos, A. K. Katsaggelos, and R. Molina, “A Bayesian Approach for the Estimation and Transmission of Regularization Parameters for Reducing Blocking Artifacts,” *IEEE Trans. Image Processing*, 9, 7, 1200–1215, July 2000.
- [19] R. Molina, A. K. Katsaggelos and J. Mateos, “Bayesian and Regularization Methods for Hyperparameter Estimation in Image Restoration,” *IEEE Trans. Image Processing*, 8, 2, 231–246, Feb. 1999.
- [20] R. Molina, J. Abad, M. Vega, and A. K. Katsaggelos, “Parameter Estimation in Bayesian High-Resolution Image Reconstruction with Multisensors,” *IEEE Trans. Image Processing*, 12, 12, 1642–1654, Dec. 2003.
- [21] C. E. Morris, M. A. Richards and M. H. Hayes, “Fast Reconstruction of Linearly Distorted Signals,” *IEEE Trans. Acoust., Speech, Signal Processing*, 36, 1017–1025, July 1988.
- [22] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, New York, Academic Press, 1970.
- [23] R. W. Schafer, R. M. Mersereau, and M. A. Richards, “Constrained Iterative Restoration Algorithms,” *Proc. IEEE*, 69, 432–450, April 1981.
- [24] C. A. Segall, A. K. Katsaggelos, R. Molina, and J. Mateos, “Bayesian Resolution Enhancement of Compressed Video,” *IEEE Trans. on Image Processing*, 13, 7, 898–911, July 2004.
- [25] A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*, Winston, Wiley, 1977.
- [26] P. H. Van Cittert, “Zum Einfluss der Spaltbreite auf die Intensitatswerteilung in Spektrallinien II,” *Z. Physik*, 69, 298–308, 1931.
- [27] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, “Regularized Image Reconstruction from Incomplete Block Discrete Cosine Transform Data,” *IEEE Trans. Circuits and Systems for Video Technology*, 3, 6, 421–432, Dec. 1993.