AN ITERATIVE ALGORITHM FOR 2-D SIGNAL RESTORATION

Kazys KAZLAUSKAS

Institute of Mathematics and Informatics
Akademijos St. 4, 2600 Vilnius, Lithuania

Abstract. Two-dimensional signals of physical phenomena may be inadvertently altered before recording through the system whose bandwidth is smaller than that of the signal. It is often desired to restore later such data by removing the effects of the linear system. This restoration may be accomplished by synthesizing two-dimensional (2-D) inverse filters on computers. Approximations are necessary to insure the stability of the inverse filter.

Key words: two-dimensional, iterative technique, signal restoration, discrete.

1. Introduction. The restoration of a signal that has been distorted is one of the most important problems in signal processing applications (Andrews, 1977). Iterative signal restoration is a common and simple approach in recovering degraded signals. If matrix $F$ presents a linear deterministic distortion, $y$ and $x$ represent, respectively, blurred and original signals, then the distortion may be due to motion between the camera and the scene or due to atmospheric turbulence. The signal restoration problem is then to find a signal as close as possible to the original one, subject to a suitable optimality criterion. Iterative restoration algorithms have a number of advantages over direct restoration techniques, and they have been used extensively (Abramatic, 1982; Dudgeon, 1980; Katsaggelos, 1990; Farruck, 1983; Perry, 1994; Sanz, 1983; Thomas, 1981; Thomas, 1991).

The organization of the paper is as follows. The model of a 2-D system is reviewed in Section 2. In Section 3, we define intervals in which the restored signal exists. Then, using the expansion of the 2-D inverse operator into a power series, we derive formulas for calculating an approximated inverse filter and signal restoration in the time and frequency domains. It is shown that if the number of iterations or the number of the power series terms $N \to \infty$,
then the limit of the restored signal sequence is the input signal. The extension of that method to the case when the 2-D frequency response is normalized is addressed in Section 4. Unfortunately, the iterative algorithm is slowly converging, and the convergence point is usually not the best restoration because of noise amplification. In Section 5 a stopping rule is imposed to maximize the effectiveness of the iterative algorithm. The normalized error is used as a measure of the restoration quality. We analyze convergence conditions and define the optimum number of iterations. We briefly analyze restoration in a noisy environment in Section 6.

An advantage of the iteration algorithm is that some matrix sequence can be computed in advance. The solution sequence is then computed on-line after the distorted signal is available.

2. 2-D digital filters. The transfer function of a 2-D digital filter is

\[ F'(z_1, z_2) = \frac{B'(z_1, z_2)}{A'(z_1, z_2)} = \frac{\sum_{k,l=1}^{M_2 N_2} b'_{kl} z_1^{-k+1} z_2^{-l+1}}{\sum_{i,j=1}^{M_1 N_1} a'_{ij} z_1^{-i+1} z_2^{-j+1}}, \]  

(1)

where \( a'_{ij} \) and \( b'_{kl} \) are constants; \( z_1 \) and \( z_2 \) are 2-D \( z \) transform variables.

If \( x(n_1, n_2) \) is input, then \( z \) transform of the output is

\[ Y(z_1, z_2) = F'(z_1, z_2) X(z_1, z_2), \]  

(2)

where \( Y(z_1, z_2) \) is \( \mathcal{Z}\{y(n_1, n_2)\} \), \( X(z_1, z_2) \) is \( \mathcal{Z}\{x(n_1, n_2)\} \), \( \mathcal{Z} \) is the sign of 2-D \( z \) transform.

Without loss of generality, we can also assume that Eq. 1 is normalized so that \( a_{11} = 1 \). In such a case from Eq. 1, and assuming that \( z_1^{-1} = e^{-j\omega_1} \), \( z_2^{-1} = e^{-j\omega_2} \), we have a 2-D rational frequency response

\[ F(\omega_1, \omega_2) = \frac{B(\omega_1, \omega_2)}{A(\omega_1, \omega_2)} = \frac{\sum_{k,l=1}^{M_2 N_2} b_{kl} \exp[-j\omega_1(k-1) - j\omega_2(l-1)]}{\sum_{i,j=1}^{M_1 N_1} a_{ij} \exp[-j\omega_1(i-1) - j\omega_2(j-1)]}, \]  

(3)

where \( a_{ij} = a'_{ij}/a_{11} \), \( b_{kl} = b'_{kl}/a_{11} \); \( \omega_1 \) and \( \omega_2 \) are spatial frequencies.
3. Inverse filtering. For the ideal measurement system $F(\omega_1, \omega_2)$, it follows that $y(n_1, n_2) = x(n_1, n_2)$ (see Fig. 1).

In all practical cases $y(n_1, n_2) \neq x(n_1, n_2)$, and $y(n_1, n_2)$ is corrupted with noise $\xi(n_1, n_2)$. So the purpose of inverse filtering is to process the signal $v(n_1, n_2)$ so that we get a restored signal $\tilde{x}(n_1, n_2)$ as close to the input signal $x(n_1, n_2)$ as possible.

\[
\tilde{x}(n_1, n_2) = \mathcal{F}^{-1}\left\{\frac{V(\omega_1, \omega_2)}{F(\omega_1, \omega_2)}\right\},
\]

if $F(\omega_1, \omega_2) \neq 0$.

As follows from Fig. 1

\[
V(\omega_1, \omega_2) = F(\omega_1, \omega_2) X(\omega_1, \omega_2) + \xi(\omega_1, \omega_2).
\]  

(4)

Restoration of the signal $x(n_1, n_2)$ by the inverse filtering method is carried out dividing both sides of Eq. 4 by frequency response $F(\omega_1, \omega_2)$ and calculating a 2-D inverse Fourier transform, i.e.,

\[
\tilde{x}(n_1, n_2) = \mathcal{F}^{-1}\left\{\frac{V(\omega_1, \omega_2)}{F(\omega_1, \omega_2)}\right\} = x(n_1, n_2) + \mathcal{F}^{-1}\left\{\frac{\xi(\omega_1, \omega_2)}{F(\omega_1, \omega_2)}\right\},
\]

if $F(\omega_1, \omega_2) \neq 0$.

Eq. 5 cannot be evaluated numerically because of the zeros of $F(\omega_1, \omega_2)$. An approximation must be made. On the basis of the information available, if there is no knowledge of $x(n_1, n_2)$ other than that obtained in $y(n_1, n_2)$, the most one can do is to restore up to their value those frequency components of $x(n_1, n_2)$ which have been reduced by the convolution with $f(n_1, n_2)$. On this basis, the solution is defined in the following section.

4. An iterative implementation of 2-D inverse filters. Real filters don’t pass those frequencies of the input signal $x(n_1, n_2)$ for which $F(\omega_1, \omega_2) = 0$. In the case, when we have no information about the input signal $x(n_1, n_2)$
An iterative algorithm for 2-D signal restoration

and noise $\xi(n_1, n_2) = 0$, it is expedient to restore the input signal only in the intervals where $F(\omega_1, \omega_2) \neq 0$. We define the restored signal as

$$\hat{X}(\omega_1, \omega_2) = \begin{cases} \frac{Y(\omega_1, \omega_2)}{F(\omega_1, \omega_2)}, & \text{if } F(\omega_1, \omega_2) \neq 0, \\ 0, & \text{if } F(\omega_1, \omega_2) = 0. \end{cases}$$

In Eq. 6 we separate a non-ideal part of the filter $F(\omega_1, \omega_2)$

$$\hat{X}(\omega_1, \omega_2) = Y(\omega_1, \omega_2) + [1/F(\omega_1, \omega_2) - 1]Y(\omega_1, \omega_2),$$

where $1/F(\omega_1, \omega_2) - 1$ characterizes the distortion of the filter $F(\omega_1, \omega_2)$. For the ideal filter $1/F(\omega_1, \omega_2) - 1 = 0$.

Expanding a 2-D inverse filter, we have

$$\frac{1}{F(\omega_1, \omega_2)} = 1 + \left[1 - F(\omega_1, \omega_2)\right] + \left[1 - F(\omega_1, \omega_2)\right]^2 + \ldots$$

Then

$$\hat{X}(\omega_1, \omega_2) = Y(\omega_1, \omega_2) \left\{1 + \left[1 - B(\omega_1, \omega_2)\right] + \left[1 - B(\omega_1, \omega_2)\right]^2 + \ldots\right\},$$

where

$$Y_a(\omega_1, \omega_2) = Y(\omega_1, \omega_2) A(\omega_1, \omega_2).$$

In all practical cases we can calculate only a finite number of power series terms, i.e.,

$$\hat{X}_N(\omega_1, \omega_2) = Y_a(\omega_1, \omega_2) \sum_{i=0}^{N} \left[1 - B(\omega_1, \omega_2)\right]^i,$$

or in the time domain

$$\hat{X}_N(n_1, n_2) = F^{-1}\left\{Y_a(\omega_1, \omega_2) \sum_{i=0}^{N} \left[1 - B(\omega_1, \omega_2)\right]^i\right\}.$$
Also
\[ \hat{x}_N(n_1, n_2) = y_a(n_1, n_2) \ast f_N(n_1, n_2), \]  
(12)
where
\[ y_a(n_1, n_2) = y(n_1, n_2) \ast a(n_1, n_2), \]
\[ f_N(n_1, n_2) = \delta(n_1, n_2) + \left[ \delta(n_1, n_2) - b(n_1, n_2) \right] + \ldots \]
\[ + \left[ \delta(n_1, n_2) - b(n_1, n_2) \right]^N, \]
and
\[ \delta(n_1, n_2) = 1, \text{ if } n_1 = n_2 = 1 \text{ and } \delta(n_1, n_2) = 0, \text{ in other cases}, \]
* is the sign of the convolution operation.

Define
\[ l_1(n_1, n_2) = [\delta(n_1, n_2) - b(n_1, n_2)] \ast y_a(n_1, n_2) \]
and
\[ l_N(n_1, n_2) = [\delta(n_1, n_2) - b(n_1, n_2)] \ast l_{N-1}(n_1, n_2), \quad N \neq 1. \]
Then
\[ \hat{x}_N(n_1, n_2) = y_a(n_1, n_2) + \sum_{i=1}^{N} l_i(n_1, n_2) \]  
(13)
or, in the frequency domain,
\[ \hat{X}_N(\omega_1, \omega_2) = Y_a(\omega_1, \omega_2) + \sum_{i=1}^{N} l_i(\omega_1, \omega_2), \]  
(14)
where
\[ l_i(\omega_1, \omega_2) = Y_a(\omega_1, \omega_2) \left[ 1 - B(\omega_1, \omega_2) \right]^i, \quad i = 1, N, \]
\[ |1 - B(\omega_1, \omega_2)| < 1. \]

From Eq. 14 we can calculate estimates of the input signal
\[ \hat{x}_0(n_1, n_2) = y_a(n_1, n_2), \]
\[ \hat{x}_1(n_1, n_2) = \hat{x}_0(n_1, n_2) \]
\[ + [y_a(n_1, n_2) - b(n_1, n_2) \ast \hat{x}_0(n_1, n_2)], \]
\[ \hat{x}_N(n_1, n_2) = \hat{x}_{N-1}(n_1, n_2) \]
\[ + [y_a(n_1, n_2) - b(n_1, n_2) \ast \hat{x}_{N-1}(n_1, n_2)], \]  
(15)
An iterative algorithm for 2-D signal restoration

\[ \hat{x}_0(n_1, n_2) = y_a(n_1, n_2), \]
\[ \hat{x}_1(n_1, n_2) = \hat{x}_0(n_1, n_2) + [\delta(n_1, n_2) - b(n_1, n_2)] * y_a(n_1, n_2), \]
\[ \hat{x}_N(n_1, n_2) = \hat{x}_{N-1}(n_1, n_2) + [\delta(n_1, n_2) - b(n_1, n_2)]^{N-1} * y_a(n_1, n_2). \]  

(16)

If the sequence of restored signals \( \hat{x}_0(n_1, n_2), \ldots, \hat{x}_N(n_1, n_2) \) has a limit as \( N \to \infty \), then the Fourier transform of this limit is \( \hat{X}(\omega_1, \omega_2) = Y(\omega_1, \omega_2)/F(\omega_1, \omega_2) \), if the condition \( \hat{X}(\omega_1, \omega_2) = 0 \) when \( F(\omega_1, \omega_2) = 0 \) is satisfied. We can get the same result as in Eq. 16 solving the Fredholm equation of the first type by a sequential substitution method. As we can see from Eq. 16, the first estimate of the input signal is equal to the output signal. The following estimate is calculated by adding the previous estimate with a correction member. The correction member is equal to the difference between the output signal of the filter \( F(\omega_1, \omega_2) \) and the output signal of the filter \( B(\omega_1, \omega_2) \) (see Fig. 2).

![Fig. 2. Iterative input signal restoration](image)

Note that the input signal restoration by inverse filtering is similar to the input signal restoration using the pseudoinverse method (Kazlauskas, 1977). By substituting \( \hat{x}_0(n_1, n_2) \) into \( \hat{x}_1(n_1, n_2) \), and so on up to \( N \), from equations (15) we obtain the following expression

\[ \hat{x}_N(n_1, n_2) = (N + 1) y_a(n_1, n_2) - U_N(n_1, n_2) * y_a(n_1, n_2), \]  

(17)
where

\[ U_N(n_1, n_2) = \frac{(N + 1)N}{2!} b(n_1, n_2) - \frac{(N + 1)N(N - 1)}{3!} b^{2**}(n_1, n_2) + \ldots (-1)^N b^{N**}(n_1, n_2), \] (18)

in which \( b^{N**}(n_1, n_2) \) is the \( N \)-time 2-D convolution of \( b(n_1, n_2) \). Thus, for the filter \( F(\omega_1, \omega_2) \), we can compute \( U_N(n_1, n_2) \) in advance and use it in the restoration process.

Equation (17) can be written as a 2-D convolution of \( y_a(n_1, n_2) \) with the impulse response of the restoration system \( b_N(n_1, n_2) \)

\[ \hat{x}_N(n_1, n_2) = b_N(n_1, n_2) * y_a(n_1, n_2), \] (19)

where

\[ b_N(n_1, n_2) = (N + 1)\delta(n_1, n_2) - \frac{(N + 1)N}{2!} b(n_1, n_2) + \frac{(N + 1)N(N - 1)}{3!} b^{2**}(n_1, n_2) + \ldots (-1)^N b^{N**}(n_1, n_2). \]

Partial case. Let \( b_{ii} = 1 \). Then the restored signal in a frequency domain

\[ \hat{X}(\omega_1, \omega_2) = Y_a(\omega_1, \omega_2)/(1 - E(\omega_1, \omega_2)), \] (20)

where

\[ E(\omega_1, \omega_2) = \sum_{k=1}^{M} \sum_{l=1}^{N} b_{kl} \exp (-j\omega_1(k - 1) - j\omega_2(l - 1)). \]

By expanding equation (20) into a power series, we get

\[ \hat{X}(\omega_1, \omega_2) = Y_a(\omega_1, \omega_2) \sum_{i=0}^{\infty} E^i(\omega_1, \omega_2), \quad |E(\omega_1, \omega_2)| < 1. \] (21)

The inverse filter frequency response is

\[ \hat{F}(\omega_1, \omega_2) = A(\omega_1, \omega_2) \sum_{i=0}^{\infty} E^i(\omega_1, \omega_2), \quad |E(\omega_1, \omega_2)| < 1. \] (22)
An iterative algorithm for 2-D signal restoration

Using $N$ terms of the power series, from equation (21) we obtain

$$
\begin{align*}
\hat{X}_0(\omega_1, \omega_2) &= Y_a(\omega_1, \omega_2), \\
\hat{X}_1(\omega_1, \omega_2) &= Y_a(\omega_1, \omega_2) + E(\omega_1, \omega_2)\hat{X}_0(\omega_1, \omega_2), \\
\vdots \\
\hat{X}_N(\omega_1, \omega_2) &= Y_a(\omega_1, \omega_2) + E(\omega_1, \omega_2)\hat{X}_{N-1}(\omega_1, \omega_2).
\end{align*}
$$

(23)

It follows from equation (22) that the rational frequency response of the inverse filter can be theoretically implemented by applying iterative computation an infinite number of times, providing the convergence criterion is met. In practice, naturally the iterative computation is applied only a finite number of times, so the frequency response that is actually realized is an approximation to the original inverse rational frequency response.

The iterative implementation seems well suited to digital processors that have the ability to convolve a 2-D signal with a filter kernel of limited extent.

5. Determination of the Iteration number. In any practical implementation, the number of iterations actually computed is finite. we can take the ratio

$$
\left|\frac{X(\omega_1, \omega_2) - \hat{X}_N(\omega_1, \omega_2)}{X(\omega_1, \omega_2)}\right|,
$$

(24)
as a measure of the spectral error introduced by terminating the iterative computation after $N$ iterations. The ratio is complex, in general. However,

$$
\hat{X}_N(\omega_1, \omega_2) = F_N(\omega_1, \omega_2)Y(\omega_1, \omega_2)
= F_N(\omega_1, \omega_2)F(\omega_1, \omega_2)X(\omega_1, \omega_2),
$$

(25)

where

$$
F_N(\omega_1, \omega_2) = \sum_{i=0}^{N} \left[1 - F(\omega_1, \omega_2)\right]^i = \frac{1 - [1 - F(\omega_1, \omega_2)]^{N+1}}{F(\omega_1, \omega_2)}.
$$

(26)

Then, using equations (24), (25), and (26), we have

$$
\left|\frac{X(\omega_1, \omega_2) - \hat{X}_N(\omega_1, \omega_2)}{X(\omega_1, \omega_2)}\right| = |1 - F(\omega_1, \omega_2)|^{N+1}.
$$

(27)

Analogously, for the normalized case, we have

$$
\left|\frac{X(\omega_1, \omega_2) - \hat{X}_N(\omega_1, \omega_2)}{X(\omega_1, \omega_2)}\right| = |E(\omega_1, \omega_2)|^{N+1}.
$$

(28)
If we specify a tolerable degree of spectral error by fixing $\varepsilon$, we can use these relations to tell us how many iterations will be needed for a given value of $|E(\omega_1, \omega_2)|$ or $|1 - F(\omega_1, \omega_2)|$. Conversely, it can be used to determine how $|E(\omega_1, \omega_2)|$ or $|1 - F(\omega_1, \omega_2)|$ must be restricted, if the number of iterations $N$ is preassigned.

Convergence conditions. The power series in (9) converges, if $|1 - B(\omega_1, \omega_2)| < 1$. If the condition $|1 - B(\omega_1, \omega_2)| < 1$ is not satisfied, then we can choose a ratio $B(\omega_1, \omega_2)/k$, where $k = \text{const}$, such that the inequality $|1 - B(\omega_1, \omega_2)/k| < 1$ be satisfied. In such a case we must multiply the restored signal $\hat{x}_N(n_1, n_2)$ by $1/k$.

The power series $1 + [1 - B(\omega_1, \omega_2)] + [1 - B(\omega_1, \omega_2)]^2 + \ldots$ is nonuniformly converging in the domain $|1 - B(\omega_1, \omega_2)| < 1$, therefore a partial amount of the power series

$$B_N(\omega_1, \omega_2) = 1 + \left[1 - B(\omega_1, \omega_2)\right] + \left[1 - B(\omega_1, \omega_2)\right]^2 + \ldots$$

$$+ \left[1 - B(\omega_1, \omega_2)\right]^{N-1}$$

(29)

is unrestrictedly growing as $N \to \infty$ and $B(\omega_1, \omega_2) \to 0$. The same power series uniformly converges in the domain $|1 - B(\omega_1, \omega_2)| < r < 1$. Clearly the inequalities $|1 - [1 - B(\omega_1, \omega_2)]| > 1 - |1 - B(\omega_1, \omega_2)| \geq 1 - r$ are satisfied for all $B(\omega_1, \omega_2)$, if $|1 - B(\omega_1, \omega_2)| < r$. Then $|\hat{F}_N(\omega_1, \omega_2) - \hat{F}(\omega_1, \omega_2)| \leq r^N/(1-r)$ for all $B(\omega_1, \omega_2)$ in the domain $|1 - B(\omega_1, \omega_2)| < r$.

Conclusion. The power series in (9) uniformly converges if $|1 - B(\omega_1, \omega_2)| < r < 1$. The filter $\hat{F}(\omega_1, \omega_2)$ is stable, because $B(\omega_1, \omega_2) \neq 1$.

6. Restoration in the presence of noise. The previously developed results, in general, are applicable to the signal restoration problem when output of the filter is observed with additive noise. We present an approach of modifying the previous filter. We analyze the case when there is no apriori information on the noise. In the presence of noise, an inverse filter can be obtained by applying the Wiener filtering method.

Minimizing the mean-square error between the input signal $x(n_1, n_2)$ and the restored signal $\hat{x}(n_1, n_2)$, we get the Fourier transform of the inverse filter impulse response $f(n_1, n_2)$:

$$\hat{F}(\omega_1, \omega_2) = \frac{1}{\hat{F}(\omega_1, \omega_2)} |\hat{F}(\omega_1, \omega_2)|^2$$

$$+ \hat{F}(\omega_1, \omega_2)/|X(\omega_1, \omega_2)|,$$

(30)
where $\xi(\omega_1, \omega_2)$ is the noise power density spectrum; $F(\omega_1, \omega_2)$ and $X(\omega_1, \omega_2)$ are the Fourier transform of $f(n_1, n_2)$ and $x(n_1, n_2)$, respectively. For the case $\xi(\omega_1, \omega_2) = 0$, (30) reduces to the inverse filter $1/F(\omega_1, \omega_2)$.

If no statistical properties of the process are known, we assume the ratio of the noise and signal power density to be constant. Then equation (30) is of the form:

$$\hat{F}(\omega_1, \omega_2) = \frac{1}{F(\omega_1, \omega_2)} \frac{|F(\omega_1, \omega_2)|^2}{|F(\omega_1, \omega_2)|^2 + c},$$

(31)

where $c$ is constant. The second term on the right side of equation (31) is a smoothing function which additionally smoothes the restored signal, in order to provide an acceptable signal enhancement in the presence of noise.

7. Concluding remarks. Theoretically we obtain the frequency function of IF by accomplishing iterative computation an infinite number of times. However, in fact we calculate only a limited number of iterations, consequently, the frequency function of IF, that is realized, is an approximation to the original frequency function of IF. The iterative implementation of IF is quite convenient in processors that may entail a convolution of 2-D signals with a pulse characteristic of the filter of limited size. Since the same calculations are iterated, such processors are able to operate with longer characteristics of the filter.

The method ensures a sufficiently good quality of restoring signals in the cases when noise is not the main source of distortion. The presence of noise restricts the maximal $N$ to be used. For small $N$ one may essentially increase the informaticity of an output signal even in the presence of noise. The possibility of choosing the optimal number of iterations $N$ depends on a priori information on an input signal of the system. If the noise is great with respect to the input signal, then the output signal of the system is processed by Wiener's filter.

In the iterative realization of 2-D filters a recursive IF is replaced by a feedback filter. Afterwards the feedback filter is represented by the sum of 2-D nonrecursive filters, and instead of the stability condition of the recursive filter the convergence condition is verified.

In addition, this method unlike other methods, gives an experimenterator a possibility of controlling the restoration process to a certain extent by varying the amount of iterations $N$.

The results of the work can be successfully applied in design of 2-D IF's and restoration of 2-D signals.
REFERENCES


Received June 1995
An iterative algorithm for 2-D signal restoration

K. Kazlauskas received Ph.D. degree from the Kaunas Polytechnic Institute (Kaunas, Lithuania) in 1975. He is a senior researcher of the Technological Process Control Department at the Institute of Mathematics and Informatics. His research interests include digital signal processing and speech enhancement.

DVIMAČIŲ SIGNALŲ ATSTATYMO ITERATYVUS ALGORITMAS

Kazys KAZLAUSKAS

Fiziniių procesų dvimačiai signalai keičia savo formą per tiesinę sistemą, kurių pralaidumo juosta siauresnė negu signalo. Dažnai reikia atstatyti šiuos signalus, įvertinant sistemos poveikį. Signalų atstatymą galima atlikti kompiuterio pagalba, sintezuojant dvimačių inversinį filtrą. Inversinis filtras aproksimuojamas tam, kad užtikrinti jo stabilumą.