Bayesian-Based Iterative Method of Image Restoration*

William Hadley Richardson
Visibility Laboratory, University of California, San Diego, San Diego, California 92152
(Received 15 September 1970)

An iterative method of restoring degraded images was developed by treating images, point spread functions, and degraded images as probability-frequency functions and by applying Bayes's theorem. The method functions effectively in the presence of noise and is adaptable to computer operation.

Index Headings: Spread function; Image restoration; Deconvolution.

This paper reports the results of applying probability methods to restoration of noisy degraded images. Fourier-transform methods of image restoration[1,2] have been successful when noise content in the degraded image is moderate or small. At increased noise levels, however, Fourier methods have failed to produce recognizable images.

ASSUMPTIONS

It was assumed that the degraded image H was of the form $H=W*S$, where W is the original image, S is the point spread function, and $*$ denotes the operation of convolution. It was also assumed that W, S, and H are discrete probability-frequency functions, not necessarily normalized. That is, the numerical value of a point of W, S, or H is considered as a measure of the frequency of the occurrence of an event at that point. S is usually in normalized form. Units of energy (which may be considered unique events) originating at a point in W are distributed at points in H according to the frequencies indicated by S. H then represents the resulting sums of the effects of the units of energy originating at all points of W. In what follows, each of the three letters has two uses when subscripted. For example, W_i indicates either the ith location in the array W or the value associated with the ith location. The unsubscripted letter refers to the entire array or the value associated with the array as in $W=\sum_i W_i$. The double-subscripted $W_{i,j}$ in two dimensions is interpreted similarly to W_i in one dimension. In the approximation formulas, a subscript r appears, which is the number of the iteration.

DISCUSSION

Given the degraded image H, the point spread function S, and the requirement to find the original image W, Bayes's theorem comes readily to mind. In the notation of this problem the usual form of the Bayes's theorem[3] may be stated as the conditional probability of an event at W_i given an event at H_k,

$$P(W_i|H_k)=\frac{P(H_k|W_i)P(W_i)}{\sum_j P(H_k|W_j)P(W_j)};\quad i=\{1,I\},\quad (1)$$

$$j=\{1,J\},\quad k=\{1,K\},$$

where H_k is for the moment an arbitrary cell of H.

Considering all the H_k and their dependence on all W_i in accordance with S, we can say

$$P(W_i)=\sum_k P(W_iH_k)=\sum_k P(W_i|H_k)P(H_k),\quad (2)$$

since $P(W_i|H_k)=P(W_iH_k)/P(H_k)$. Substituting Eq. (1) in Eq. (2) gives

$$P(W_i)=\sum_k \frac{P(H_k|W_i)P(W_i)P(H_k)}{\sum_j P(H_k|W_j)P(W_j)}.\quad (3)$$

In the right side of this equation, the term $P(W_i)$ is also the desired solution. But in many applications of Bayes's theorem, when this $P(W_i)$ term is not known, an accepted practice[4,5] is to make the best of a bad situation and use an estimate of the $P(W_i)$ function to obtain, from Eq. (1), an approximation of $P(W_i|H_k)$. When this practice is applied here, Eq. (3) becomes

$$P_{r+1}(W_i)=P_r(W_i)\sum_k \frac{P(H_k|W_i)P(H_k)}{\sum_j P(H_k|W_j)P_r(W_j)};\quad (4)$$

$$r=\{0,1,\cdots\}.$$

This results in an iterative procedure where the initial $P_0(W_i)$ is estimated. An estimation often used is Bayes's postulate (also known as the equidistribution of ignorance), which assumes a uniform distribution so that $P_0(W_i)=1/I$ or $W_{i,0}=W/I$.

Equation (4) can be reduced to a more easily workable form by $P(W_i)=W_i/W$ and $P(H_k)=H_k/H=H_k/W$, since the restoration is a conservative process and $W=H$, and also $P(H_k|W_i)=P(S_{i,k})=S_{i,k}/S$,

$$S=\sum_j S_j,\quad j=\{1,J\}.$$

Then Eq. (4) becomes

$$W_{i,r+1}/W=(W_{i,r}/W)\sum_k \frac{(S_{i,k}/S)\cdot(H_k/W)}{\sum_j (S_{j,k}/S)\cdot(W_{j,r}/W)}$$

or

$$W_{i,r+1}=W_{i,r}\sum_k \frac{S_{i,k}H_k}{\sum_j S_{j,k}W_{j,r}}.\quad (5)$$

As this stands, investigation shows that in programming Eq. (5) for a computer, the finite size of the arrays

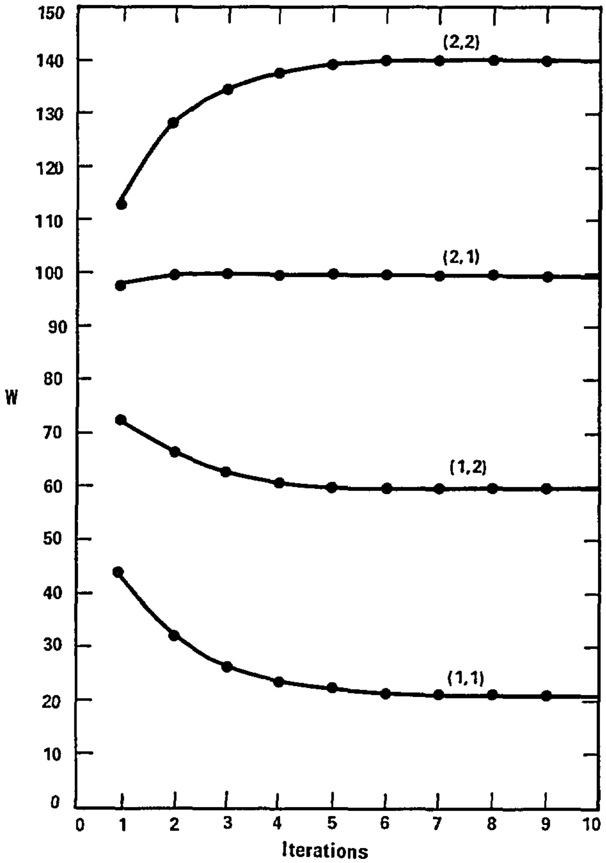

FIG. 1. Convergence of restoration toward true values of undegraded image.

allows Eq. (5) to be rewritten

$$W_{i,r+1}=W_{i,r}\sum_{k=i}^{c}\frac{S_{k-i+1}H_k}{\sum_{j=a}^{b}S_{k-j+1}W_{j,r}},\tag{6}$$

where $a=(1,k-J+1)_{\max}$, $b=(k,I)_{\min}$, and $c=i+J-1$. When Bayes's postulate is used for the initial estimate, Eq. (6) becomes

$$W_{i,1}=\sum_{k=i}^{c}\frac{S_{k-i+1}H_k}{\sum_{j=a}^{b}S_{k-j+1}}\tag{7}$$

for the first iteration. In Eq. (6), it appears that the summation over k is a corrective factor on $W_{i,r}$. If $W_{m,r}$ were incorrect, the $W_{m,r}$ in the denominator of the summation tends to correct the $W_{m,r}$ in the first term of the right side of the equation. For example, if in Eq. (6) $W_{m,r}$ were too large, the recurrences of $W_{m,r}$ in the denominator where $a\leq m\leq b$ would tend to reduce the value of $W_{m,r}$ in calculating $W_{m,r+1}$. This does not always occur, for the value of $W_{m,r}$ may change in the wrong sense initially to compensate for a relatively large change in the value of another $W_{i,r}$ in the neighborhood.

Equations (6) and (7) expanded to two dimensions are

$$W_{i,j,r+1}=W_{i,j,r}\sum_{m=i}^{e}\sum_{n=j}^{f}\frac{H_{m,n}S_{m-i+1,n-j+1}}{\sum_{p=a}^{b}\sum_{q=c}^{d}W_{p,q,r}S_{m-p+1,n-q+1}},\tag{8}$$

and

$$W_{i,j,1}=\sum_{m=i}^{e}\sum_{n=j}^{f}\frac{H_{m,n}S_{m-i+1,n-j+1}}{\sum_{p=a}^{b}\sum_{q=c}^{d}S_{m-p+1,n-q+1}},\tag{9}$$

where

$a=(1,m-K+1)_{\max}$; $b=(m,I)_{\min}$;
$c=(1,n-L+1)_{\max}$; $d=(n,J)_{\min}$;
$e=i+K-1$; $f=j+L-1$;
$i=\{1,I\}$; $j=\{1,J\}$.
K, L are the dimensions of $S_{k,l}$.
I, J are the dimensions of $W_{i,j}$.

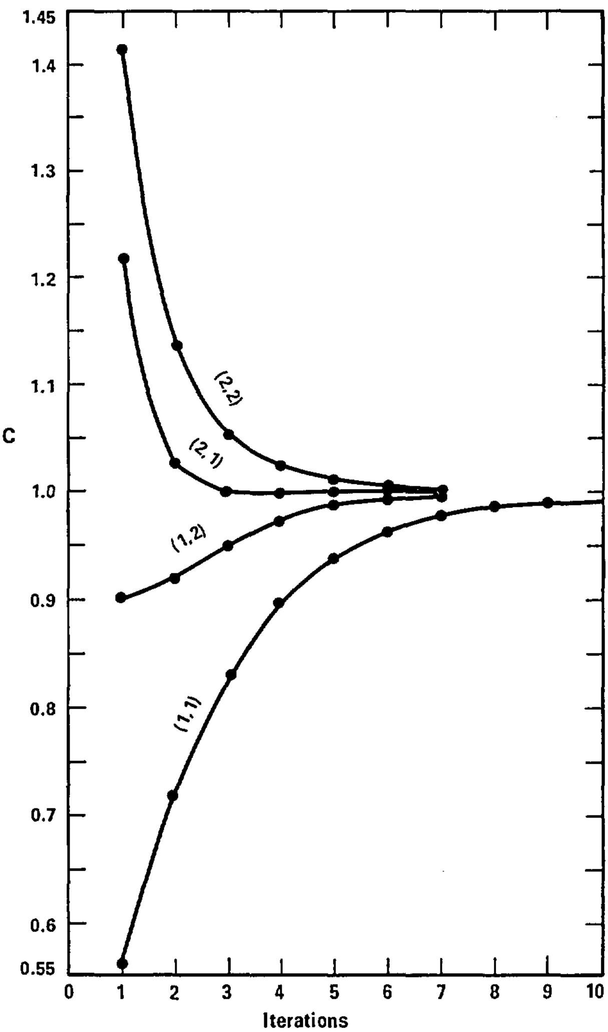

FIG. 2. Convergence of correction-factor values (C) toward unity.

An example of the convergence of the process is shown in a simple two-dimensional, noiseless system in Fig. 1, where

$$H=\begin{bmatrix} 2 & 10 & 12 \\ 16 & 60 & 52 \\ 30 & 82 & 56 \end{bmatrix}; \quad S=\begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}.$$

The undegraded image was

$$W=\begin{bmatrix} 20 & 60 \\ 100 & 140 \end{bmatrix}.$$

Figure 2 shows the trend toward unity of the correction factors that gave the results for each iteration.

Figure 3 shows a restoration of two targets without noise and Fig. 4 shows two targets with 0.1 random, multiplicative noise added after degradation. The degradation was by a uniform point spread function five units square. In these figures, the pairs are in the order (A) undegraded image, (B) degraded image, (C) 10-iteration restoration, (D) 20 iterations, and (E) 30 iterations. The specification of added noise is defined by r in

$$H'=H(1+rd), \tag{10}$$

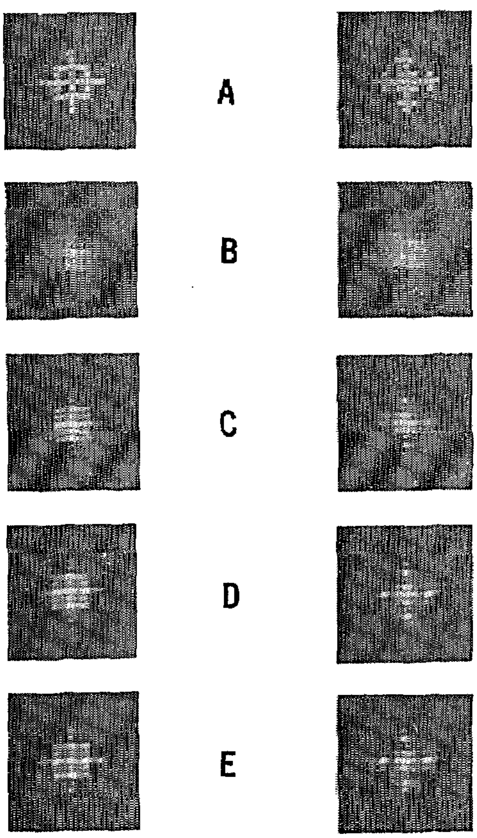

FIG. 3. Restoration with no noise. (A) Original image, (B) degraded image, (C) 10-iteration restoration, (D) 20 iterations, and (E) 30 iterations.

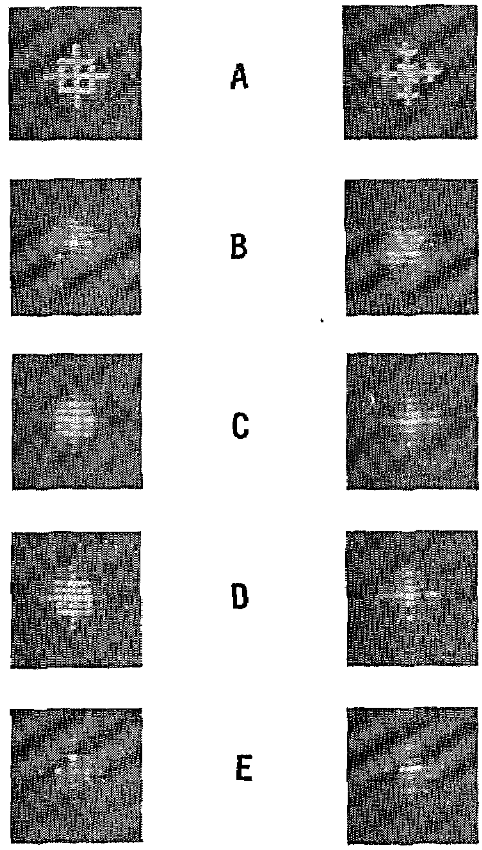

FIG. 4. Restoration with 0.1 noise. (A) Original image, (B) degraded image, (C) 10-iteration restoration, (D) 20 iterations, and (E) 30 iterations.

where H' is the noisy value, H is the noise-free value, r is the fraction of added noise, and d is the random deviate, normal (0,1). In this example, r is 0.1. The two targets were designed with the idea of simplicity, but including a certain amount of detail. They were designed also to make it difficult to relate either degraded image with either undegraded image. The approximating process appears successful because the restorations are readily identifiable after only 10 iterations. In the case of no noise, the procedure tends toward a perfect restoration, whereas with noise the restoration continues to improve.

Figure 5 shows restoration of the same noisy degraded images by the Fourier-transform method using a least-squares filter with estimated parameters to reduce the effects of noise. The presence of background noise in the restoration should be noted in contrast to the lack of it in the iterative restorations. This could be a detriment in more-complex images.

The effect of noise introduced after degradation was investigated. One element of a noiseless degraded image was doubled and the resulting image was restored. The undegraded image was a square uniform field five points on a side. The point spread function was a square uniform field three points on a side, resulting in a

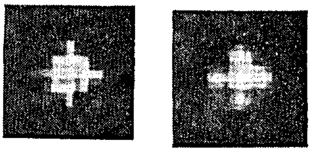

FIG. 5. Restoration with 0.1 noise by the Fourier-transform method with estimated least-squares filter.

seven-by-seven degraded image. First, successive points for noise addition were chosen on a diagonal of the seven-by-seven degraded image, progressing point by point from a corner to the center of the image. Next, points were chosen on an axis, progressing from the center of a side to the center of the image. Figures 6 and 7 show the results after 10 iterations of the restoration process.

This process has been programmed and run extensively on the CDC 3600 computer at the University of California, San Diego, Computer Center. In programming, the massive calculations indicated in Eq. (8) were reduced by a very large factor by removing redundant calculations.

The restorations shown in Figs. 3 and 4 averaged 7.4 s per iteration.

CONCLUSIONS

Although no proof of convergence on a solution has been devised, the process did converge in all cases for which it was used. This seems reasonable because the process is essentially self-correcting in the course of a sequence of interdependent adjustments. In effect, the

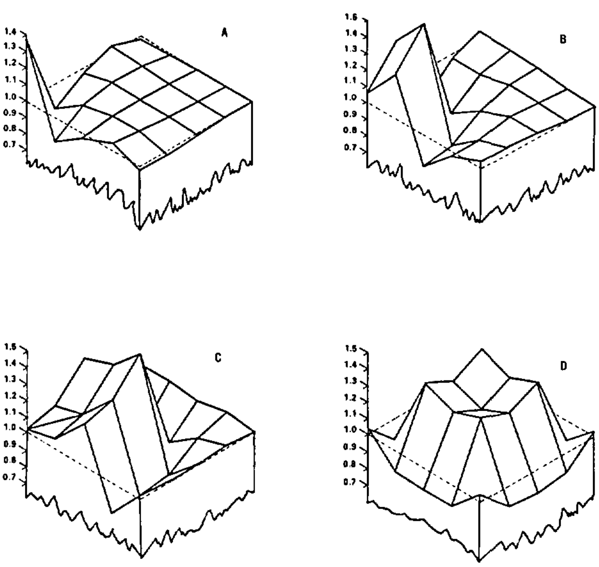

FIG. 6. Restoration of originally uniform image (all values fo $W_{i,j}=1.0$). One value on diagonal of degraded image doubled. (A) $H'(0,0)=2H(0,0)$, corner: maximum $W=1.380$, minimum $W=0.850$; (B) $H'(1,1)=2H(1,1)$: maximum $W=1.474$, minimum $W=0.807$; (C) $H'(2,2)=2H(2,2)$: maximum $W=1.494$, minimum $W=0.819$; (D) $H'(3,3)=2H(3,3)$, center: maximum $W=1.320$, minimum $W=0.863$.

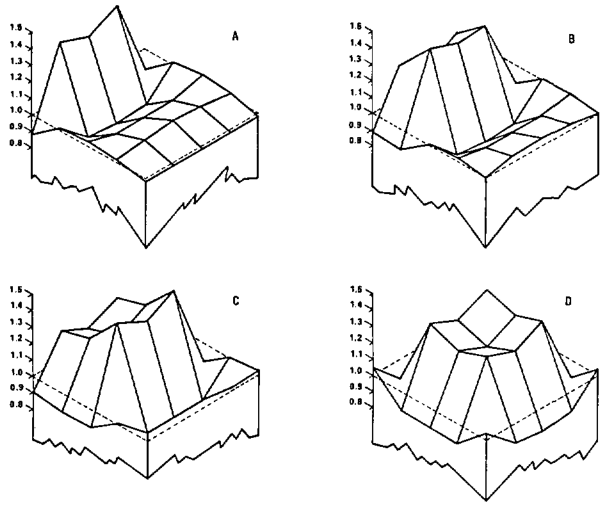

FIG. 7. Restoration of originally uniform image (all values of $W_{i,j}=1.0$). One value on an axis of degraded image doubled. (A) $H'(3,0)=2H(3,0)$, middle of side: maximum $W=1.348$, minimum $W=0.837$; (B) $H'(3,1)=2H(3,1)$: maximum $W=1.307$, minimum $W=0.876$; (C) $H'(3,2)=2H(3,2)$: maximum $W=1.315$, minimum $W=0.882$; (D) $H'(3,3)=2H(3,3)$, center: maximum $W=1.320$, minimum $W=0.863$.

process may be considered an iterative deconvolution approximation.

In the examples of noise added at a single point in the degraded image, a certain amount of disturbance results in the restoration. However, this does not appear to be out of control, as would often be the case with the Fourier type of restoration. In the case of generally distributed random noise, the disturbances tend to compensate each other. On the other hand, the cases in which noise was added at or near the edge of the degraded image give warning of what may happen when a part of a degraded image is processed. Here, the edges of the degraded image may be contaminated by energy coming from an area in the original image outside the area of restoration. In practice, this results in a very bright edge around the restoration, almost entirely confined to the outer row or two. Effects of the extraneous noise in the central part of the restoration are very slight. This result can be countered by eliminating the outer elements of the restored image after restoration or by judicious tapering of values in the edges of the degraded image.

The value of this process is that it can give intelligible results in some cases where the Fourier process cannot, although at higher cost than the Fourier process.

Equations (1)–(4), adapted with the notation of Eqs. (6) and (7), indicate how this type of process may be applied to some statistical uses of Bayes's theorem when the *a priori* probabilities $P(W_i)$ are not known.

ACKNOWLEDGMENTS

The author is grateful to J. L. Harris, Sr., for suggesting this problem and affording the opportunity to

work on it. Dr. R. W. Preisendorfer has made many valuable suggestions during the development of this process. A. Shaules and D. P. MacAdam have acted as helpful sounding boards during the investigation.

REFERENCES

* This research was supported by the Advanced Research Projects Agency under Air Force Contract F08606-68-C-0017.

[1] J. L. Harris, Sr., J. Opt. Soc. Am. **56**, 569 (1966).

[2] B. L. McGlamery, J. Opt. Soc. Am. **57**, 293 (1967).

[3] E. Parzen, *Modern Probability Theory and Its Applications* (Wiley, New York, 1960), Ch. 3, Sec. 4, p. 119, Eq. 4.16.

[4] *The International Dictionary of Mathematics* (Van Nostrand, Princeton, N. J., 1960), p. 70, "Bayes's Theorem."

[5] R. E. Machol and P. Gray, *Recent Developments in Information and Decision Processes* (Macmillan, New York, 1962), p. 175, "Stable Estimation."