UNCLASSIFIED

NRL Report 8447

# A Method for Specifying the Noise Suppression-Resolution Tradeoff in Digital Image Filtering with Local Statistics

WARREN W. WILLMAN

*Systems Research Branch*
*Space Systems Divisions*

December 11, 1980

**NAVAL RESEARCH LABORATORY**
**Washington, D.C.**

Approved for public release; distribution unlimited.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>NRL Report 8447 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>A METHOD FOR SPECIFYING THE NOISE SUPPRESSION-RESOLUTION TRADEOFF IN DIGITAL IMAGE FILTERING WITH LOCAL STATISTICS | | 5. TYPE OF REPORT & PERIOD COVERED<br>Final report on one phase of a continuing NRL problem.<br>6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>Warren W. Willman | | 8. CONTRACT OR GRANT NUMBER(s) |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Naval Research Laboratory<br>Washington, DC 20375 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>Program Element 35-111N<br>Project X0922CC<br>NRL Problem 79-0927-A-0 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Naval Electronic Systems Command<br>Arlington, VA 20360 | | 12. REPORT DATE<br>December 11, 1980<br>13. NUMBER OF PAGES<br>18 |
| 14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED<br>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

| | |
|---|---|
| Digital image processing | Adaptive noise filtering |
| Image processing | Noise suppression |
| Image enhancement | Local statistics method |
| Image noise filtering | |

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A modification of Lee's local-statistics method is developed for digital image noise filtering without the need for a user-specified noise level. Instead, the user specifies an approximate upper size limit for image features he is willing to regard as noise. A noise level is then estimated automatically for each local region of the image by partitioning the region into subregions and exploiting the differing degrees of spatial correlation for signal and noise. Extensions for use with multiplicative noise, third-moment statistics, and edge-detection schemes are discussed briefly. Examples are given for an image of 256 X 256 pixels.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601



Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

## CONTENTS

INTRODUCTION ..... 1

ALGORITHM DEVELOPMENT ..... 2

Simplified Analysis ..... 3

More-Realistic Analysis ..... 5

Overall Filtering Procedure ..... 8

PERFORMANCE CONSIDERATIONS ..... 9

EXTENSIONS ..... 14

Use of Third-Moment Estimates ..... 14

Use with Multiplicative Noise ..... 15

Use with Edge-Detection Refinement ..... 16

REFERENCES ..... 16

UNCLASSIFIED

# A METHOD FOR SPECIFYING THE NOISE SUPRESSION–RESOLUTION TRADEOFF IN DIGITAL IMAGE FILTERING WITH LOCAL STATISTICS

## INTRODUCTION

Recently Lee [1] developed a noise-suppression algorithm for digital images composed of a square array of picture elements, or pixels. In its basic form, this algorithm is based on the theory that the noise is additive, random (with zero mean), and independent for different pixels, and that it has a known variance which is the same for all pixels in the image. The algorithm processes each pixel separately. Except near the image boundaries, where no processing is performed, the output for each pixel is an estimated gray level which is a nonlinear function of all the pixel gray levels within the square neighborhood of a user-specified size of which that pixel is the center. The estimated gray level of such a pixel is computed according to the equation

$$\hat{x} = \bar{x} + \frac{m}{m + r}(z - \bar{x}), \tag{1}$$

where

$\hat{x}$ = estimated gray level,
$z$ = observed gray level (of this pixel),
$\bar{x}$ = sample mean of (observed) gray levels in surrounding neighborhood,
$r$ = user-specified noise variance,
$m$ = $\max\{0, \mathrm{v} - r\}$, and
$\mathrm{v}$ = sample variance of gray levels in surrounding neighborhood.

One justification for this estimate is the fact that $\hat{x}$ of Eq. (1) would be the conditional mean gray level of the pixel in question in the underlying noise-free image, given the observed noise-corrupted gray level $z$, if this noise-free gray level had a *known* prior Gaussian distribution with mean $\bar{x}$ and variance $m$. Actually, of course, such an $\bar{x}$ and $m$ are not known, but if it is further assumed that all the pixel gray levels in the local neighborhood are independent and have the same distribution, and that the above *local statistics* $\bar{x}$ and $m$ achieve their expected values, then they can be substituted for these parameters of the prior distribution, as suggested by the notation. These local statistics would typically fluctuate somewhat about their expected values under these conditions, so a key point here is that these fluctuations would be relatively minor for this choice of statistics, since these statistics are basically averages over a large ensemble of pixels.

One requirement of this whole procedure is that the user must specify the variance $r$ of the noise. In most practical applications, this noise variance is unknown and is also spatially varying. To deal with this problem, Lee [2] has developed an adaptive refinement of this scheme, in which $r$ is not user-specified, but rather is estimated for each pixel's local neighborhood from the sample variance of a relatively smooth subneighborhood prior to the aforementioned processing. Beyond this difficulty, however, there is the somewhat philosophical problem that a given feature in a given image can often play the role of either signal or noise, depending on what kind of information happens to be of interest. In an air-search radar, for example, ground clutter is regarded as noise, even though it consists of actual returns which are neither random nor spurious. Thus there is merit in allowing the user to have some control over the noise suppression, but in a way which allows efficient discrimination between interesting and superfluous information.

Manuscript submitted August 12, 1980.

A different modification of the same basic algorithm is developed here as a means of achieving such discrimination. This modification also uses Eq. (1), with the local mean statistic $\bar{x}$ computed as before. The variance estimates $m$ and $r$, however, are generated by partitioning the local neighborhood, or window, into a number of subregions in a certain way. Except for some correction factors, the "noise" variance $r$ is then taken as the average sample variance *within* these subregions and the "signal" variance $m$ as the sample variance *among* the subregion sample means. The only quantity controlled by the user is the size of the window to be used. To a first approximation, the width of this window sets an upper limit on the size of features which the algorithm might treat as noise rather than signal. The algorithm is meant then to suppress noise to the maximum extent that is consistent with this limitation, essentially by exploiting the difference in spatial correlation between signal and noise. Hence, specifying a larger window size results in more noise suppression, but with a greater loss of resolution for subtle details. Further explanation and elaborations of this tradeoff are given in a later section.

## ALGORITHM DEVELOPMENT

The algorithm examined here estimates the gray level of each pixel from the observed gray levels of the surrounding $N \times N$ square neighborhood of pixels. The number $N$ is always chosen as odd, so the pixel whose gray level is being estimated is at the center of this neighborhood. This local neighborhood of pixels, excluding the one at the center, is then divided into $M$ subregions, which are all approximately square and equal in size. An example of such a subdivision is shown in Fig. 1 for a $5 \times 5$ region with four subregions. This pattern of four subregions can clearly be extended to any $N \times N$ region with $N$ odd. For a reason explained later, however, it is better to keep $M$ approximately equal to $N$. Thus, if the size of the local neighborhood were increased to $9 \times 9$, it would be better to subdivide it into nine $3 \times 3$ subregions (except that the center subregion would have the center pixel removed).

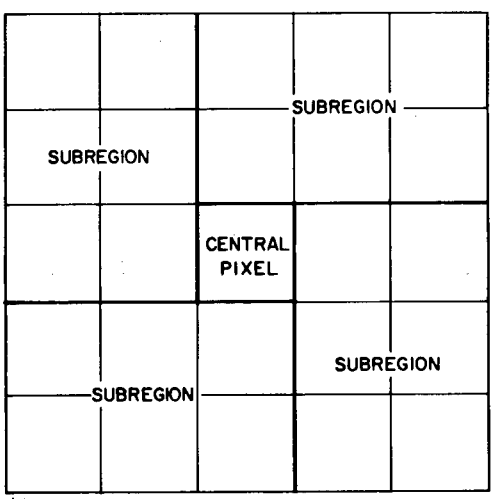


Fig. 1 — Subdivision of $5 \times 5$ window into four subregions

The $\bar{x}_j$ and $s_j$ statistics (i.e., functions of the observed pixel gray levels) are computed for each subregion $j$:

$$\bar{x}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} z_{ji} \tag{2}$$

and

$$s_j = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (z_{ji} - \bar{x}_j)^2, \tag{3}$$

UNCLASSIFIED

where

$n_j$ = number of pixels in subregion $j$

and

$z_{ji}$ = observed gray level of the $i$ th pixel in subregion $j$.

In addition, the $\bar{x}$ and $s$ statistics are computed for the region as a whole (excluding the center pixel):

$$\bar{x} = \frac{1}{M} \sum_{j=1}^{M} \bar{x}_j \tag{4}$$

and

$$s = \frac{1}{M-1} \sum_{j=1}^{M} (\bar{x}_j - \bar{x})^2. \tag{5}$$

These statistics are then used to form estimates of the local mean and variance of the image signal and the local variance of the added noise. These estimates in turn are used in Eq. (1) to estimate the gray level of the central pixel. This entire process is repeated with each pixel in the image serving as the central pixel.

Since the observation noise is presumed to have a zero mean, the local mean of the (noise-free) image is simply estimated as the sample mean of the observed gray levels in the neighborhood surrounding the central pixel, i.e., $\bar{x}$ as given by Eq. (4). The local variance estimates for the image and noise are made on the basis that the noise values tend to fluctuate more rapidly with position than do the gray levels of the underlying image. In such a case, the noise tends to contribute to the $s_j$ statistics of Eq. (3) relatively more, and the signal relatively more to the $s$ statistic of Eq. (5). The particular way in which these statistics are used to distinguish between the local signal and noise variance can also be regarded as an operational definition of which aspects of the observed image constitute noise and which constitute signal. Some implications of this view will be discussed later.

To determine a way of estimating these variances, we now consider two probabilistic analyses, one simple and the other more realistic. The simplified analysis leads to serious inaccuracies, but it is nevertheless instructive because it introduces some basic effects in a more direct way and because it suggests a way of extending the algorithm to include local third-moment statistics of the image.

## Simplified Analysis

Since the noise is considered to fluctuate more rapidly with position than does the signal, we adopt the crude approximation that the noise values for different pixels in the $N \times N$ region are statistically independent, identically distributed random variables and that the image signal is constant *within* each of the $M$ subregions, but that these values are independent, identically distributed random variables for the different subregions. Of course, it's impossible for all the gray-level changes in an image to occur exactly at the subregion boundaries for all possible locations of the central pixel, so this approximation is guaranteed to contain some errors. If these are accepted, however, it follows from Eqs. (2) and (3) and standard sampling theory that

$$E(s_j) = r, \; j = 1, \cdots, M, \tag{6}$$

where $E$ denotes expected value and $r$ is the noise variance. Also, if the observed pixel gray level $z_{ji}$ is written as

$$z_{ji} = u_j + w_{ji}, \tag{7}$$

where

$u_j$ = the (constant) signal gray level with subregion $j$

and

$w_{ji}$ = noise value for pixel $i$ in subregion $j$,

then it follows from Eq. (2) that

$$\bar{x}_j = u_j + \frac{1}{n_j} \sum_{i=1}^{n_j} w_{ji}. \tag{8}$$

Hence, it follows from decomposing expectations that

$$\begin{aligned} E(\bar{x}_j) &= E[E(x_j/u_j)] \\ &= E(u_j) = \mu \end{aligned} \tag{9}$$

and

$$\begin{aligned} E(x_j^2) &= E[E(x_j^2/u_j)] \\ &= E\left(u_j^2 + \frac{r}{n_j}\right) = \mu^2 + m + \frac{r}{n_j}, \end{aligned} \tag{10}$$

where $\mu$ and $m$ are the mean and variance of $u_j$. Since both signal and noise values are statistically independent in distinct subregions, the $\bar{x}_j$ are independent random variables. Assuming that $n_j$ is the same number $n$ for all subregions, these random variables are also identically distributed, with mean $\mu$ and variance $m + r/n$. From standard sampling theory, therefore,

$$E(s) = m + \frac{r}{n}, \tag{11}$$

for $s$ as defined by Eqs. (4) and (5).

It follows from Eqs. (6) and (11) that

$$r = \frac{1}{M} \sum_{j=1}^{M} E(s_j)$$

and

$$m = E(s) - \frac{r}{n}$$

under the approximations used here, where

$r$ = local noise variance

and

$m$ = local image signal variance.

This suggests that a reasonable procedure for constructing estimates of these variances would be to ignore fluctuations of the statistics $s_j$ and $s$ from their expected values to obtain

$$\hat{r} = \frac{1}{M} \sum_{j=1}^{M} s_j \tag{12}$$

and

$$\hat{m} = s - \frac{\hat{r}}{n}, \tag{13}$$

where $s_j$ and $s$ are given by Eqs. (2) to (5), and where

$\hat{r}$ = estimate of $r$

and

$\hat{m}$ = estimate of $m$.

UNCLASSIFIED

Because of possible fluctuations of the $s$ and $s_j$ statistics from their expected values, the estimate $\hat{m}$ might assume a negative value on any given implementation of this procedure. However, $m$ is a variance, and therefore necessarily nonnegative. Also, Eq. (1) is not reasonable for negative values of $m$. Hence, it is important in practice to modify this estimate of $m$ slightly to guarantee nonnegativity, for instance by using the maximum of zero and $\hat{m}$ of Eq. (13).

## More-Realistic Analysis

As mentioned earlier, the approximations on which the variance estimates of Eqs. (12) and (13) are based cannot be accurate. A more realistic approach is to base these estimates on the expected values of the $s_j$ and $s$ statistics, as given by Eqs. (2) to (5), when the gray levels of the noise-free image have a joint probability distribution with a bona fide correlation function. As a simple but nontrivial way of doing this, we adopt the approximation that, at least within a region larger than the local neighborhood being considered, the marginal distribution of each pixel's gray level (without noise) is identical to that of every other pixel, with mean $\mu$ and variance $m$, and that for two pixels $i$ and $j$ the correlation coefficient $\rho_{ij}$ for the gray levels is

$$\rho_{ij} = e^{-\frac{|x_i - x_j| + |y_i - y_j|}{L}}, \tag{14}$$

where, in an $x$-$y$ coordinate system aligned with the pixel grid,

$(x_i,y_i)$ = location of center of pixel $i$,
$(x_j,y_j)$ = location of center of pixel $j$, and
$L$ = a scale distance (corresponds heuristically to an average object size in the image).

As before, the noise values are assumed to be zero-mean, independent, identically distributed random variables for each pixel in the local neighborhood, with variance $r$.

The form of the correlation function corresponding to Eq. (14) is shown in Fig. 2, where $\rho(x,y)$ denotes $\rho_{ij}$ for $x_i - x_j = x$ and $y_i - y_j = y$. This form is not completely ideal since it is not isotropic. It does at least have 90-degree symmetry, however, and the anisotropy is relatively mild. This form of correlation function is used here because it is mathematically convenient and can be specified by the single distance parameter $L$, which corresponds roughly to an average object size, or resolution, in the noise-free image.

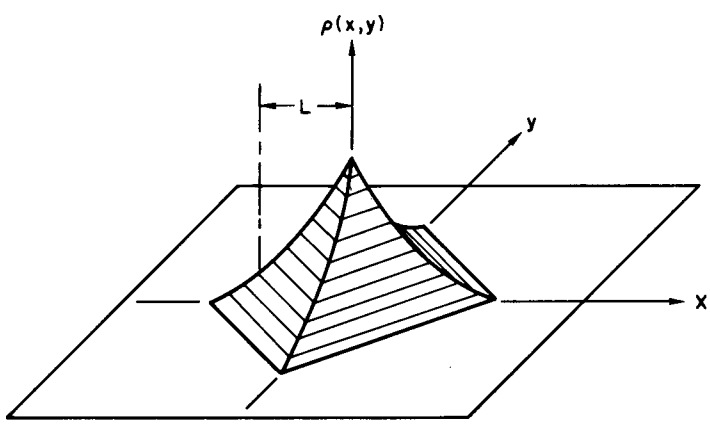


Fig. 2 — Exponential correlation function

Now consider $s_j$ for a particular subregion $j$, and delete the $j$ subscript in the notation. From Eqs. (2) and (3),

$$s = \frac{1}{n-1} \sum_{i=1}^{n} [z_i - \frac{1}{n}(z_1 + \cdots + z_n)]^2.$$

After some rearrangement of terms, this can be rewritten as

$$s = \frac{1}{n-1} \left( \sum_{i=1}^{n} z_i^2 \right) - \frac{1}{n(n-1)} (z_1 + \cdots + z_n)^2. \tag{15}$$

Assuming that the signal and noise are uncorrelated, it follows in this case that

$$E(z_i^2) = \mu^2 + m + r$$

by a standard result of probability theory. Since expectation is a linear operation, therefore,

$$E\left( \frac{1}{n-1} \sum_{i=1}^{n} z_i^2 \right) = \frac{n}{n-1} (\mu^2 + m + r). \tag{16}$$

Also, $(z_1 + \cdots + z_n)$ can be expressed as $b\ y\ y^T\ b^T$, where

$$y = \begin{bmatrix} z_1 \\ \hline z_2 \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline z_n \end{bmatrix}$$

and

$$b = [1\ 1 \cdots 1].$$

From standard results of multivariate probability theory, therefore,

$$\begin{aligned} E(z_1 + \cdots + z_n) &= b\ E(y\ y^T)\ b^T \\ &= b\ (a\ a^T + A + r\ I_n)\ b^T, \end{aligned} \tag{17}$$

under the assumptions of this analysis, where $I_n$ is the $n$-dimensional identity matrix, $a$ is the $n$-vector

$$\begin{bmatrix} \mu \\ \hline \mu \\ \hline \cdot \\ \cdot \\ \cdot \\ \hline \mu \end{bmatrix}$$

and $A$ is the $n \times n$ matrix whose $(i,j)$th component is $m\rho_{ij}$. Evaluating Eq. (17) at the component level then gives

$$E(z_1 + \cdots + z_n) = n^2\mu^2 + nr + m \sum_{\alpha,\beta=1}^{n} \rho_{\alpha\beta}. \tag{18}$$

Combining Eqs. (16) and (18) with Eq. (15) for $s$ gives

$$\begin{aligned} E(s) &= \mu \left[ \frac{n}{n-1} - \frac{n^2}{n(n-1)} \right] + r \left[ \frac{n}{n-1} - \frac{n}{n(n-1)} \right] + m \left[ \frac{n}{n-1} - \frac{1}{n(n-1)} \sum_{\alpha,\beta=1}^{n} \rho_{\alpha\beta} \right] \\ &= r + m \left( \frac{n}{n-1} \right) \left( 1 - \frac{1}{n^2} \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} \rho_{\alpha\beta} \right). \end{aligned} \tag{19}$$

UNCLASSIFIED

This expression is difficult to evaluate exactly, but a relatively simple approximation can be obtained for the case in which the subregion is square, contains a large number $n$ of pixels, and is oriented with its edges at 45° to the pixel grid (which is presumed to be a square array). In this case, the summations in Eq. (19) can be approximated by integrals, and the equation reduces to

$$E(s_j) = r + m - \frac{m}{4R^4} \iint_{|x| + |y| \leqslant R} e^{-\frac{|x| + |y|}{L}} dx\, dy, \tag{20}$$

where the $j$ subscript has been reinstated in the notation, and where $R$ is one-half the diagonal width of the subregion $j$. Evaluating the integral of Eq. (20) gives

$$E(s_j) = r + m - mG(R/L), \tag{21}$$

where $G$ is used for later convenience to denote the function

$$G(x) = \frac{1}{2x^4} [4x^2 - 6x + 3 + e^{-2x} (2x^2 - 3)]. \tag{22}$$

This approximation is used from now on for the expected value of $s_j$ under the conditions of this analysis, but it should be borne in mind that it is somewhat in error for subregions which are not quite square, have a different orientation, or have only a small number of pixels. The anisotropy of the correlation function $\rho(x,y)$ is rather mild, however, so the accuracy of the approximation should depend only weakly on orientation.

To find the expected value of the statistic $s$ of Eq. (5), we note from Eq. (2) for $\bar{x}_j$ and a rearrangement of terms that

$$s = \frac{1}{M-1} \sum_{j=1}^{M} \frac{1}{n_j} \sum_{i=1}^{n_j} [(z_{ji} - \bar{x})^2 - (z_{ji} - \bar{x}_j)^2].$$

Therefore, if each of the $M$ subregions has an equal number $n$ of pixels,

$$s = I_1 - I_2, \tag{23}$$

where

$$I_1 = \frac{1}{n(M-1)} \sum_{j=1}^{M} \sum_{i=1}^{n} (z_{ij} - \bar{x})^2 \tag{24}$$

and

$$I_2 = \frac{1}{M-1} \sum_{j=1}^{M} \left(\frac{n-1}{n}\right) s_j. \tag{25}$$

Taking expected values and using Eq. (21) gives

$$E(I_2) = \left(\frac{M}{M-1}\right) \left(\frac{n-1}{n}\right) [r + m - mG(R/L)], \tag{26}$$

where $2R$ is the diagonal width of the subregions. The statistic corresponding to $s_j$ for the entire local neighborhood is

$$\frac{1}{Mn-1} \sum_{j=1}^{M} \sum_{i=1}^{n} (z_{ji} - \bar{x})^2 = \frac{Mn-n}{Mn-1} I_1.$$

Since this neighborhood is also square except for the exclusion of the central pixel, which is presumed to have a negligible effect for a large number of pixels, the analysis leading to Eq. (21) can be applied to this expression to give

$$E(I_1) = \frac{Mn-1}{Mn-n}\,[r + m - mG(R\sqrt{M}/L)]. \tag{27}$$

The diagonal width of this neighborhood is $\sqrt{M}$ times that of the subregions because it is a square consisting of $M$ equally sized square subregions.

Using Eqs. (26) and (27) to evaluate the expected value of Eq. (23), we obtain

$$E(s) = \frac{M(n-1)}{n(M-1)}\,m[G(R/L) - G(R\sqrt{M}/L)] + \frac{1}{n}[r + m - mG(R\sqrt{M}/L)]. \tag{28}$$

Under the assumptions of this analysis, if the estimates $\hat{r}$ and $\hat{m}$ are constructed from the local statistics by Eq. (12) and

$$\hat{m} = \frac{n(M-1)\,(s - \hat{r}/n)}{(Mn-1)\,[G(R/L) - G(R\sqrt{M}/L)]}, \tag{29}$$

where $s_j$ and $s$ are defined by Eqs. (2) to (5), then it follows from Eqs. (21) and (28) that

$$E(\hat{r}) = r + m[1 - G(R/L)] \tag{30}$$

and

$$E(\hat{m}) = m. \tag{31}$$

As before, it is important in practice to modify the estimate of Eq. (29) slightly to insure that it is nonnegative. The estimate $\hat{r}$ here is biased, since $E(\hat{r}) \neq r$. This defect could be eliminated by the use of a more complicated estimate, but such procedures seem to make the estimates more sensitive to fluctuations of the statistics from their average behavior; hence they are not adopted here.

**Overall Filtering Procedure**

Comparing the more realistic estimate of Eq. (29) with that based on the simplified analysis—i.e., Eq. (13)—shows that the only change is multiplication of the more simplistic estimate of $m$ by a correction factor. The estimate of $r$ is the same in both cases. Unfortunately, this correction factor depends on the scale distance $L$ in the correlation function for the underlying noise-free image, a quantity which is not really specified. Hence the approach taken here is to multiply the estimate of Eq. (13) by a correction factor which depends only on $M$ and $n$, and to select a value which gives an empirically reasonable contribution to expectations of $\hat{m}$ and $\hat{r}$ for values of $L$ in the transition region where $G(R/L)$ and $G(R\sqrt{M}/L)$ change from one to zero. Figure 3 includes a graphical display of $G(R/L)$.

For the case of no noise, it follows from Eqs. (28) and (30) that

$$E[f(s - \hat{r}/n)] = f\left[\left(\frac{n-1}{n}\right)\left(\frac{M}{M-1}\right) + \frac{1}{n}\right][G(R/L) - G(R\sqrt{M}/L)]m, \tag{32}$$

where $f$ is an arbitrary correction factor. Figure 3 shows the ratio of this expectation and that of $\hat{r}$ ( as given by Eq. (30)) to $m$ as functions of $R/L$ for the limiting case of large $n$. This figure shows the results both for four subregions ($M = 4$) with $f = 5$ and for nine subregions with $f = 4$. These values of $f$ are judged to give reasonable average behavior of the resulting estimates in these two cases. No further cases were examined because the corresponding local neighborhoods would be too large for computational interest. Finally, the variation in Eq. (32) with $n$ is used to compensate for finite $n$ to yield the following estimation procedure:

$$\hat{r} = \frac{1}{M}\sum_{j=1}^{M} s_j, \tag{33}$$

UNCLASSIFIED

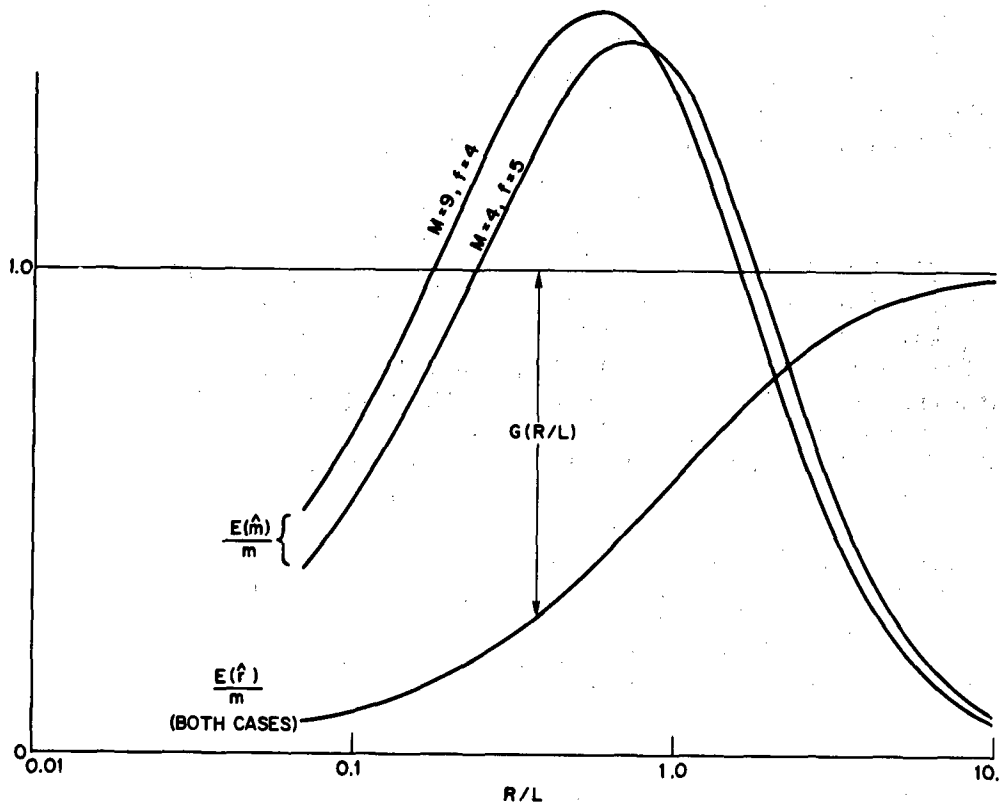


Fig. 3 — $G$, $E(\hat{m})$, and $E(\hat{r})$ as functions of $R/L$ in the transition region

$$\tilde{m} = \max\left\{0, f_M\left(\frac{Mn}{Mn-1}\right)(s - \hat{r}/n)\right\}, \tag{34}$$

and

$$\hat{m} = \tilde{m} + \frac{1}{Mn}\left[(z - \bar{x})^2 - (\tilde{m} + r)\right], \tag{35}$$

where the statistics $s_j$ and $s$ are defined by Eqs. (2) to (5), and where

$$f_4 = 5$$

and

$$f_9 = 4.$$

The modification of Eq. (35) is an ad hoc procedure used to give a better response to an isolated pixel in a locally uniform background with a different gray level.

## PERFORMANCE CONSIDERATIONS

Assuming that $M$ is always chosen to be approximately equal to $n$, the only parameter left to the discretion of the user in the final noise-suppression algorithm of Eqs. (1) to (5) and (33) to (35), with $m$ and $r$ in Eq. (1) replaced by $\hat{m}$ and $\hat{r}$, is the size of the square neighborhood from which the local statistics are formed. Since Eq. (1) would be a Bayesian estimate of the central pixel's gray level if $m$ and $r$ were known signal and noise variances, the use of the variance estimates $\hat{m}$ and $\hat{r}$ determined by Eq. (33) to (35) can be viewed as specifying which features of the observed image will be treated as signal, i.e., as part of an underlying noise-free image, and which will be treated as noise. This specification does not follow directly from the derivation of Eqs. (33) to (35), because this derivation is

based on average responses to a statistical property—the effective diameter of a correlation function—of an ensemble of images, not to particular types of features in any single image. From Eq. (1), however, it can be seen that features giving rise to large ratios of $\hat{m}/\hat{r}$ for pixels in its vicinity will be preserved as observed, whereas those producing small ratios will be smoothed out by local spatial averaging. This is because Eq. (1) can be rewritten, with $\hat{m}$ and $\hat{r}$ replacing $m$ and $r$, as

$$\hat{x} = \frac{\hat{r}}{\hat{m} + \hat{r}}\bar{x} + \frac{\hat{m}}{\hat{m} + \hat{r}}z.$$

In turn, Eqs. (33) to (35) show that the $\hat{m}/\hat{r}$ ratio is large for a pixel when the statistic $s$ is large compared to the average of the $s_j$ statistics for its local neighborhood. From Eqs. (2) to (5), this means that the gray-level variations *among* subregions are generally large compared to variations *within* a subregion. Roughly speaking, therefore, those features which get smoothed out by local spatial averaging (i.e., that are suppressed as noise) will be those whose gray-levels fluctuate rapidly with position compared to the user-chosen local neighborhood width and do so consistently over a region whose size is comparable to or greater than such a neighborhood. Edges of larger scale objects, even though they constitute a rapid fluctuation, will contribute more to $\hat{m}$ than to $\hat{r}$ and will not be smoothed out as noise, because this fluctuation is highly localized and is also associated with a large-scale contrast. The localization limits the contribution to $\hat{r}$ by confining the effects of the edge to a minority of the subregions, and the large-scale contrast causes a larger contribution to $\hat{m}$ by creating differences between many subregion averages. Isolated objects which are small compared to the local neighborhood size give intermediate values of the $\hat{m}/\hat{r}$ ratio in the absence of noise. The lack of any large-scale contrast results in a small contribution to $\hat{m}$ as well as to $\hat{r}$.

A feature such as a gridwork pattern of streets with a small spacing tends to get smoothed out as noise, however, even though it contains large linear components. Evidently, the distinction between such a pattern and a random gray-level pattern such as television "snow" is too subtle for this algorithm. Since periodicity is a key aspect of such a feature, however, it might be possible to use frequency-domain techniques to spare it from suppression. For example, one might proceed as follows for the local neighborhood of each pixel in turn:

Step 1. Fourier transform the (local) image.

Step 2. "Threshold" step 1 image to retain only high peaks.

Step 3. Inverse transform the image of step 2.

Step 4. Subtract step 3 image from original image.

Step 5. Process step 4 image as described in the preceding sections.

Step 6. Add step 3 image to step 5 image.

Such advanced procedures were not examined here, however.

Figure 4 shows the results obtained if we apply this algorithm to a digitized synthetic aperture radar image using a variety of local neighborhood or window sizes. This image measures 256 × 256 pixels. The local neighborhoods used were square in each case shown, and they were subdivided into four subregions in the manner indicated in Fig. 1. This same image is also shown in Fig. 5 for 9 × 9 local neighborhoods divided into both four and nine subregions. Figure 4 shows a general trend toward progressively greater suppression of noise (such as in the ocean area) and small-scale features (especially ones which are also complex) as the user-selected window size is increased. Another penalty for increasing the window size is the greater computational effort required. The case of a 3 × 3 window is included in Fig. 4 largely for completeness. Each of the four subregions in this case contains only two pixels, which is a serious departure from the approximations of squareness and large number of pixels on which the algorithm is based.

UNCLASSIFIED

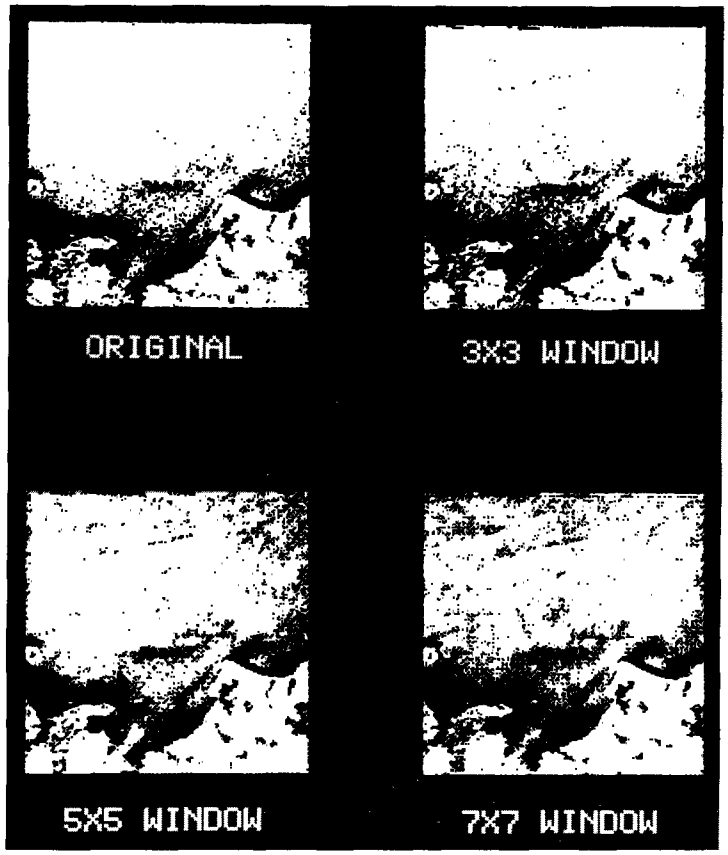


Fig. 4 — Performance vs window size

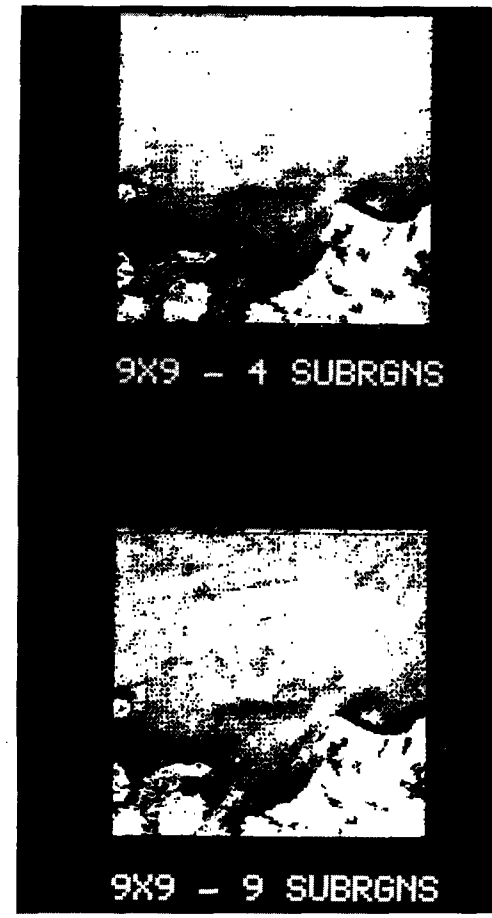


Fig. 5 — Effect of using more subregions

The operation of this algorithm is thus controlled by the user simply by the choice of a window width that is basically the approximate size of features he is willing to specify as noise rather than signal (i.e., a resolution limit). To a first approximation, the algorithm then operates to suppress the noise to the extent that is consistent with this specification. Another obvious way of performing this same kind of resolution/noise-suppression tradeoff is simply to use local spatial averaging, i.e., to set $x = \bar{x}$ in the notation here. The question naturally arises as to whether the added complexity of the current algorithm offers any significant improvement in performance. The algorithm of Lee [1], of which this is an extension, showed a rather dramatic improvement over local spatial averaging, but the user must also supply additional information, namely, the measurement noise variance. It is clear from a comparison of the examples of Fig. 6, however, that there is still a rather dramatic improvement with the current algorithm, even though the measurement noise variance is now estimated from the data rather than specified by the user. This improvement is in both the degree of noise suppression and the retention of detail without blurring. A partial explanation for this improvement is evident from the response to a semi-infinite edge in the absence of any measurement noise, shown in Fig. 7. The response curve for local spatial averaging is from Ref. 3. The curve for the current algorithm is for the case of an edge aligned with a window side, a large number of pixels in the window, and four subregions. In this case, the summations of Eqs. (2) to (5) can be approximated by integrals, which leads to the equation

$$\frac{(x - \hat{x}) \operatorname{sgn}(x - \hat{x})}{a} = \frac{\left|\frac{u}{L}\right| \left|\frac{L - u}{L}\right|}{f + (2 - f)\left|\frac{u}{L}\right|}$$

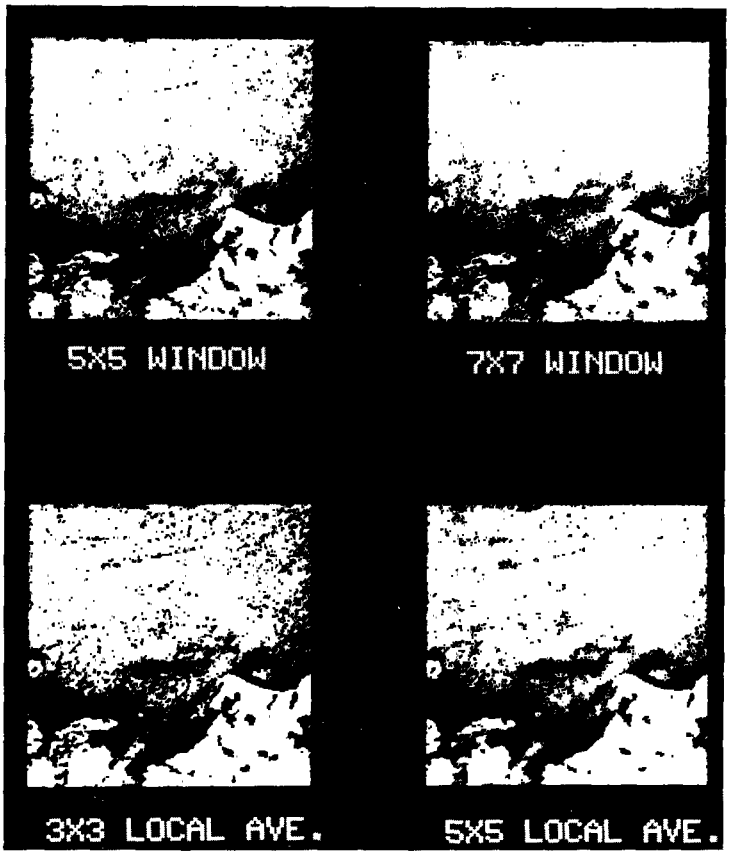


Fig. 6 — Comparison with local spatial averaging

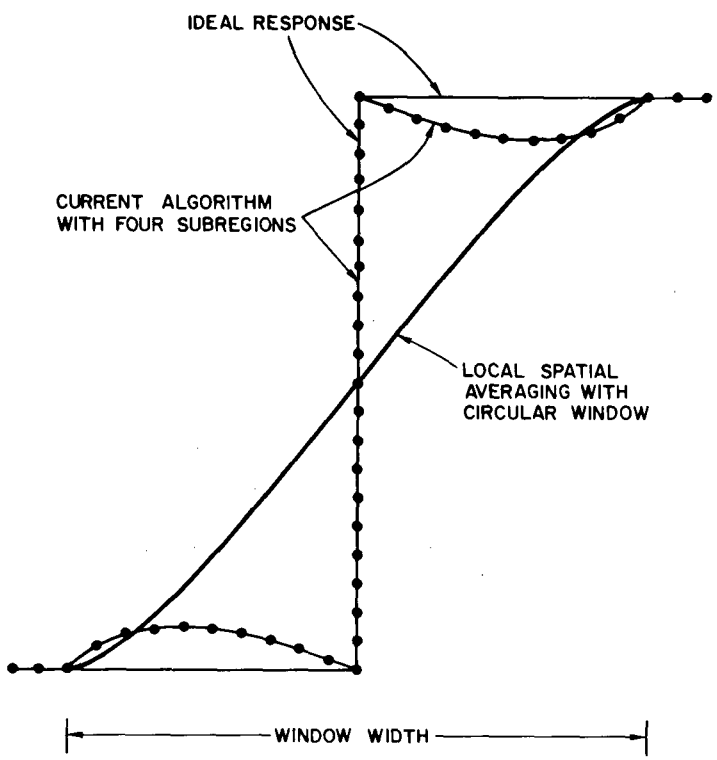


Fig. 7 — Response to semi-infinite edge (large number of pixels—four subregions)

UNCLASSIFIED

where

$a$ = gray-level change across edge,
$u$ = distance from window center,
$L$ = $R/\sqrt{2}$ = one-half window edge width,
$x$ = actual gray level at coordinate $u$, and
$\hat{x}$ = estimated gray level at $u$.

This equation is not quite accurate, because the number of pixels per subregion will not be very large until the number of subregions is increased beyond four (the two numbers are kept approximately equal). The case of a larger number of subregions is more difficult to analyze, however.

The rule of thumb that the number of subregions be approximately equal to the number of pixels per subregion has been adopted because of the occurrence of an undesirable phenomenon when the local neighborhood is divided into only four subregions, as in Fig. 1, and the number of pixels grows large. Consider such a case with no measurement noise and a small object within the window, say for simplicity a square of uniform gray level $x$ on a zero background, such that the sides of the square are aligned with those of the window. Also assume that the center of the object is displaced, parallel to an edge for simplicity, a distance $u$ from the center of the window. For values of $u$ such that the object is entirely within the window, the gray level estimated by the algorithm can be obtained if we approximate the sums by integrals; this gives

$$\bar{x} = \lambda x,$$
$$\hat{r} = [\lambda + \lambda^2(1 + \theta^2)]x^2,$$
$$\hat{m} = F\lambda^2\theta^2x^2,$$

and finally

$$\frac{\hat{x}}{x} = \left[\frac{F\theta^2 + 1 + \lambda(1 + \theta^2)}{F\lambda\theta^2 + 1 + \lambda(1 + \theta^2)}\right]\lambda$$

for $\theta < 1$, where

$\theta$ = $2u/D$,
$\lambda$ = $D^2/L^2$ (assumed $< 1/4$),
$D$ = length of object sides,
$L$ = length of window sides, and
$F$ = $\frac{4}{3} f_4 = 6\frac{2}{3}$.

As the number of pixels in the window increases, $\lambda$ decreases, for a fixed object, because $L$ grows and $D$ stays the same. The limiting behavior as the window grows is therefore

$$\frac{\hat{x}}{x} = \left(1 + 6\,\frac{2}{3}\,\theta^2\right)\lambda.$$

From its definition, $\theta$ varies from zero, when the object is centered in the window, to unity when the object edge is at the window center. Hence, as the window size increases, such a small object will not only be de-emphasized overall by the algorithm (because $\lambda$ decreases), but it will also be "hollowed out" in the center, ultimately by a factor of about eight. This hollowing out effect begins to be noticeable for the ships of Fig. 4 in the cases of the $5 \times 5$ and $7 \times 7$ windows.

As a remedy for this effect, we somewhat arbitrarily adopt the procedure of increasing the number of subregions as the number of pixels in the entire window increases. To suppress random fluctuations in the estimates of the variances and ultimately the gray level, however, it is also advantageous to have a large number of pixels in each subregion. Hence, both numbers are increased as roughly the square

root of the total number of pixels in the window. Arranging the subregions as a square array of squares is done for simplicity and even distribution. Figure 5 shows the image of Fig. 4 after processing with a $9 \times 9$ window for the cases of both four subregions and nine subregions. The hollowing out of the ship continues in the case of four subregions but is eliminated by using nine subregions. As an added benefit the other small object nearby is brightened considerably. A small price is paid in the form of less background noise suppression with nine subregions than with four, but the degree of noise suppression is still greater than in the case of a $7 \times 7$ window with four subregions (compare with Fig. 4).

A finer subdivision than nine subregions is probably too burdensome computationally to be of interest at present. The next step would be a $4 \times 4$ array of $4 \times 4$ pixel subregions, meaning a $16 \times 16$ window. At that point there would probably be so little advantage in resolution over processing the image in $2 \times 2$ blocks that this refinement would be computationally unattractive.

## EXTENSIONS

### Use of Third-Moment Estimates

It happens that a perturbation theory exists for the use of the third central moment of a prior gray-level distribution to refine the computation of a pixel's posterior mean gray level, given a noise-corrupted observed value. If the prior distribution of a pixel's noise-free gray level is almost normal, but with a small third central moment $\lambda$, then the refinement of Eq. (1) for the conditional mean is

$$\hat{x} = \bar{x} + \frac{m}{m+r}(z - \bar{x}) + \frac{r\lambda}{2(m+r)^3}[(z - \bar{x})^2 - (m + r)]. \tag{36}$$

This refinement is developed in Ref. 3 and is based on first-order Edgeworth expansions of the probability densities involved. The noise is still treated as a zero-mean normal random variable.

We use this refined formula for each pixel by replacing the parameters $\bar{x}$, $m$, $\lambda$, and $r$ in Eq. (36) by estimates formed from local statistics of a surrounding square neighborhood. As before, $\bar{x}$, $m$, and $r$ are estimated according to Eqs. (2) to (5) and (33) to (35). Under the assumptions of the simplified analysis, the statistics $\bar{x}_j$ are independent, identically distributed random variables if the number of pixels in each subregion is the same. Since the noise is assumed to have no third central moment here, from standard sampling theory $\lambda$ is the expected value of the local statistic $\bar{\lambda}$, where

$$\bar{\lambda} = \frac{M}{(M-1)(M-2)} \sum_{j=1}^{M} (\bar{x}_j - \bar{x})^3. \tag{37}$$

This statistic is not used directly as an estimate of $\lambda$ because the approximations inherent in the simplified analysis lead to errors of a factor of about five, when compared to the results of a more realistic analysis, in estimating the second moment $m$. In the case of the second-moment estimates, however, the main change resulting from this greater accuracy is equivalent to multiplication of the quantity $(\bar{x}_j - \bar{x})$ by the factor

$$\sqrt{f_M \left[\frac{Mn}{Mn-1}\right]}.$$

Since it is difficult to define a natural counterpart to a reasonable correlation function for third central moments of image pixels, compensation for the probable error resulting from the simplifications leading to the estimate of Eq. (37) is made here simply by application of the same correction factor to $(\bar{x}_j - \bar{x})$ in Eq. (37). In this extension, therefore, $\lambda$ in Eq. (36) is replaced by the estimate $\hat{\lambda}$, where

$$\hat{\lambda} = f_M^{3/2} \left[\frac{Mn}{Mn-1}\right]^{3/2} \bar{\lambda}, \tag{38}$$

with $f_4 = 5$ and $f_9 = 4$ as before.

UNCLASSIFIED

Figure 8 shows a comparison of this refinement with the basic algorithm. The incorporation of third-moment estimates produces only a slight change in this example, but it also requires only about 15% more computing time. The advantages seem to be that edges are less blurred and that isolated pixels show up better. On the other hand, there seems to be less noise suppresion in relatively uniform regions, such as the ocean in this example.

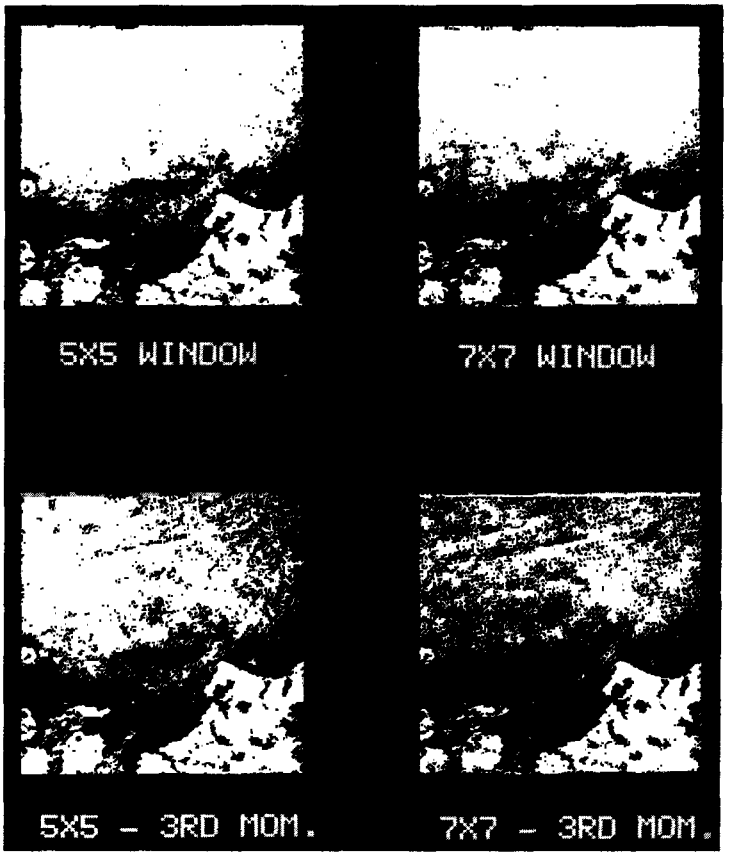


Fig. 8 — Effect of using third-moment estimates

## Use with Multiplicative Noise

Figures 4 to 6 are in fact synthetic aperture radar images. Investigations by Lee [4] have shown that the noise in such images is more accurately described as multiplicative noise than additive noise, i.e., that instead of an equation like Eq. (7) for a given pixel, one should use one of the form

$$z = u + uw. \tag{39}$$

where $w$ is approximately normally distributed with zero mean and known variance. Also, the noise variance for these synthetic aperture radar images has been determined in Ref. 4 to be 0.08. This reformulation has two implications. One is that Eq. (1) can no longer be interpreted as the Bayes rule for the conditional mean of $u$, given $z$ and prior normal densities for $u$ and the noise $w$. The other is that subregion statistics are no longer needed to estimate a local variance for $u$, since the noise variance is known. Lee [4] has exploited this extra knowledge to implement a noise-suppression filter for synthetic aperture radar images which is based on local statistics (without subregions) and a linear approximation to the conditional mean filter for multiplicative noise.

Eq. (39) can be rewritten as

$$z = u + \left[1 + \frac{u - \bar{x}}{\bar{x}}\right] v, \tag{40}$$

where

$v$ = $\bar{x}w$, a zero-mean normal random variable with variance $r$ ($r = 0.08\bar{x}^2$ for synthetic aperture radar)

and

$\bar{x}$ = local statistic of Eq. (4).

Now, given $u$, the expected value of $z$ is just $u$, from Eq. (39). Since $\bar{x}$ is a local spatial average of $z$, the quantity

$$\frac{u - \bar{x}}{\bar{x}} \tag{41}$$

should usually be small compared to unity, by the law of large numbers. Hence, the use of the noise-suppression algorithm of this report in the context of multiplicative noise can be viewed as the following procedure. Make the approximation of deleting the small quantity (41) from Eq. (40), estimate the variance of $v$ with subregion statistics rather than relying on its relation to $\bar{x}$, and then follow the same rationale as in the case of additive noise. This seems to be a reasonable procedure, the main drawback being that some known information, the relation of $r$ to $\bar{x}$, is not used. Of course if the variance of the *multiplicative* noise $w$ is not known, then this is an advantage.

**Use with Edge-Detection Refinement**

In a refinement of the algorithm of Ref. 1, Lee [2] has improved its operation by using a scheme to detect the presence and location of any significant edge in the local neighborhood of each pixel being processed. If such an edge is detected, he then alters the processing for that pixel by using as the operative local neighborhood only those pixels in the original local neighborhood which lie on the same side of the estimated edge. The basic procedure of estimating both image signal and noise variances by use of subregion statistics could be incorporated into this refinement simply by making the subregions a partition of such a modified local neighborhood instead of the entire original one. Of course, a more flexible partitioning scheme would have to be used, because the size and shape of this restricted local neighborhood is variable.

## REFERENCES

1. J.-S. Lee, "Digital Image Enhancement and Noise Filtering by Use of Local Statistics," IEEE Trans. Pat. Anal. and Mach. Int. **PAMI-2**, No. 2, 165-168, Mar. 1980.

2. J.-S. Lee, "Refined Filtering of Image Noise Using Local Statistics," NRL Report 8374; Jan. 16, 1980; also Computer Graphics and Image Processing, to be published.

3. W. W. Willman, "A Nonlinear Filtering Technique for Digitized Images Degraded by Film-Grain Noise," NRL Report 8225, Aug. 30, 1978.

4. J.-S. Lee, "Speckle Analysis and Smoothing of Synthetic Aperture Radar Images," Computer Graphics and Image Processing; to be published.