# Fourier spectrum techniques for characterization of spatial noise in imaging arrays

Glenn D. Boreman, MEMBER SPIE
University of Central Florida
Department of Electrical Engineering
    and
Center for Research in Electro-Optics
    and Lasers
Orlando, Florida 32816

**Abstract.** Noise performance in imaging arrays is often specified simply by the variance of the pixel levels. In this paper a more complete characterization technique is presented, based on the spatial-frequency power spectrum of the noise data on the detectors. This is seen to provide additional information for cases in which the noise spectrum is nonwhite. Experimental data demonstrate the nonwhite nature of the spectrum under certain conditions, especially resulting from spatial correlation of detector nonuniformity and from sampling-lattice-related artifacts in the data.

## CONTENTS

## 1. INTRODUCTION

Noise in imaging sensors usually is described simply as the variance of the pixel levels across the array or as the variance of the time record of a representative single pixel. Although they provide a useful estimate of the noise performance of such a device, these methods of characterization ignore the fact that the noise has a particular spatial distribution. There exist artifacts of array performance, either in the realm of

fixed-pattern noise or occurring in the interaction between the array and the signal acquisition electronics, that are most conveniently displayed and recognized in the Fourier domain.

For systems that operate on the Fourier transform of data acquired by an array sensor, a most important question in describing the sensor's effect on the overall system performance is, in what way does the noise of the sensor manifest itself in the transform data? The quantification of noise artifacts arising in the detector array itself is conveniently accomplished by considering the spatial-frequency power spectrum of the noise.

Characterization of the response of a sensor array in the Fourier domain can also point out spurious harmonic components in the spectrum of the data, which likely would remain unseen in the direct spatial domain. We present data to show that such harmonic components (and their aliases) often have appreciable magnitude compared with the value of the image transform itself. Thus, these components must be accounted for in any attempt to estimate the spectrum of the image data from the Fourier transform of the actual pixel values from the sensor.

The applications that motivated this characterization of sensor noise artifacts by Fourier domain techniques are related to the use of laser speckle patterns as objects for testing optical and electro-optical systems.[1] In these applications, the quantity to be measured is the power spectrum of the speckle pattern after it has passed through the system under test. The sensor noise has been found to compete significantly with the spectrum of the speckle pattern, especially in cases where the speckle pattern is of low contrast. An example of this behavior is presented in Sec. 5.7. Since both the speckle pattern and the sensor noise are random waveforms, they are difficult to

subtract directly in the spatial domain. A characterization of sensor noise artifacts in the transform domain was deemed necessary so that a subtraction might be performed on the power spectra. A power spectrum of the speckle field alone could then be calculated, under the assumption that the sensor noise is additive and signal-independent. The following sections concentrate on the characterization of spatial noise in imaging arrays by means of the spatial-frequency power spectrum.

## 2. EXPERIMENTAL SETUP

The data reported in this paper were collected with a video digitization system, seen in block diagram form in Fig. 1. The system is based on an Imaging Technology IP-512 frame buffer and an 8 bit analog-to-digital converter. The sensor under test produces a standard RS-170 video signal. This analog waveform is sampled and placed into the frame buffer. The buffer size is 512×512 pixels, into which 512 sampled values are placed horizontally and 480 sampled values are placed vertically. The remaining 32 lines of the buffer were not used in this configuration.

A General Electric TN2505 charge-injection device (CID) camera was tested in this study. The manufacturer's specification of the array is 376 pixels horizontally and 576 pixels vertically. This specification is the result of vertical and horizontal interlacing. Each video frame contains two fields of 288 lines. The pixel sites on successive fields are staggered, so that an effective resolution of 376 is obtained by combining two lines of 188 pixels each.

The focal plane was illuminated with a diffuse, uniform white-light source of variable irradiance. An option to completely block light reaching the array was also available. The host computer allowed the display of any particular row or column of the data set, and in addition performed a one-dimensional fast Fourier transform (FFT) on any selected row or column. The squared magnitude of this transform then was displayed on the video monitor.

## 3. DESCRIPTORS OF SPATIAL NOISE

To show how the Fourier domain point of view of sensor noise relates to the more usual specification of pixel variance, this section presents the analytical framework for the interpretation of sensor noise by means of the spatial-frequency power spectrum. For present purposes, we need assume only a rectangular data array of L by M digitized values. Let i be the counting parameter in the L coordinate and j be the counting parameter in the M coordinate. The specifics of how this array of digital data values corresponds to the actual photosite data will be taken up in Sec. 5.4. A typical observation of the array response at the (i,j) position may be denoted

$$d(i,j) = s(i,j) + n(i,j) \, , \tag{1}$$

where $s(i,j)$ is a digitized version of the spatial input scene and $n(i,j)$ is one observation of the two-dimensional noise process.

### 3.1. Variance

If the noise of the sensor is the quantity of primary interest, we assume that the signal input $s(i,j)$ can be made sufficiently uniform that it can be replaced by a constant, which for simplicity we take to be zero. In this case, the variance of the
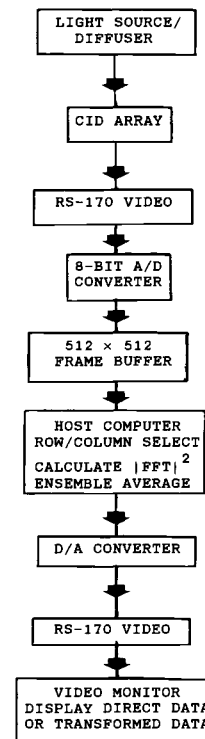


Fig. 1. Diagram of the video digitization system used in the sensor noise experiments.

pixel levels can be expressed as

$$\sigma^2 = \frac{\sum_{i=1}^{L} \sum_{j=1}^{M} \left[ n(i,j) - \sum_{i=1}^{L} \sum_{j=1}^{M} \frac{n(i,j)}{LM} \right]^2}{LM} \, . \tag{2}$$

If we define the mean of the array data as $\mu$, we can write Eq. (2) more compactly as

$$\sigma^2 = \frac{\sum_{i=1}^{L} \sum_{j=1}^{M} [n(i,j) - \mu]^2}{LM} \, . \tag{3}$$

### 3.2. Spatial-frequency power spectrum

Starting with the same digitized array of observed noise values, the discrete Fourier transform (DFT) of the sensor noise data can be expressed as[2]

$$N(p,q) = \sum_{i=1}^{L} \sum_{j=1}^{M} n(i,j) \exp\left[ - i2\pi \left( \frac{p}{L} i + \frac{q}{M} j \right) \right] \, , \tag{4}$$

where p and q are the discrete spatial-frequency components corresponding to the L and M dimensions of the original data array. The transform operation can be carried out conveniently by means of an FFT algorithm, with the region of validity restricted to within the Nyquist limits,

$$p_{Ny} = \frac{L}{2} \, , \quad q_{Ny} = \frac{M}{2} \, . \tag{5}$$

Performing an absolute value and a squaring operation on one observation of the DFT of the noise data produces an estimate of W(p,q), the spatial-frequency power spectrum. We obtain a convergent (stable) estimate of the power spectrum only by averaging several individual $|N(p,q)|^2$ spectral estimates over an ensemble of observations of the measured noise data[3]:

$$W(p,q) = <|N(p,q)|^2> . \tag{6}$$

Ideally, we would like to have the true power spectrum of the noise process available, representing the variability of the array data. In practice, however, we can obtain a good approximation by averaging a number of independently observed spectra since the estimation error decreases as the square root of the number of independent observations.[3] In Sec. 5.1, we note the effect of such ensemble averaging on typical noise spectra.

The spatial-frequency power spectrum W(p,q) has the interpretation of characterizing the fractional contribution to the total noise "power" of each spatial frequency component of the noise process. It is of particular importance when the noise of the sensor array is spatial-frequency dependent (nonwhite).

## 4. RELATIONSHIP OF POWER SPECTRUM TO VARIANCE

This section shows how W(p,q) relates to the more common specification of sensor noise, the variance of the data values. To show the analytical relationship between the spatial-frequency power spectrum and the variance, it is necessary first to consider the spatial autocorrelation of the array noise:

$$R(i',j') = \frac{\sum_{i=1}^{L} \sum_{j=1}^{M} n(i,j)n(i-i',j-j')}{LM} . \tag{7}$$

With reference to Eq. (3), it is seen that variance and autocorrelation are related as follows:

$$\sigma^2 = R(i' = 0, j' = 0) - \mu^2 . \tag{8}$$

By the Wiener-Khintchine theorem, autocorrelation and power spectrum are a Fourier transform pair:

$$R(i',j') \overset{\mathscr{F}}{\longrightarrow} W(p,q) . \tag{9}$$

Applying the central ordinate theorem to the DFT,

$$R(0,0) = \sum_{p=-L/2}^{L/2} \sum_{q=-M/2}^{M/2} W(p,q) . \tag{10}$$

Finally, combining results from Eqs. (7), (8), and (10), we obtain the desired relationship between the spatial-frequency power spectrum and the variance:

$$\sigma^2 = 2 \sum_{p=0}^{L/2} \sum_{q=0}^{M/2} W(p,q) - \mu^2 . \tag{11}$$

The interpretation of Eq. (11) is chiefly to relate the variance of the array data, the usual specification of noise performance, to a more general means of description, the spatial-frequency power spectrum. While the variance has the advantage of simplicity in such a specification, the power spectrum allows an interpretation of the spatial-frequency content of the array noise. That is, if the noise is spatially correlated, certain spatial frequencies in the image have less competition from noise than do others. The variance, as a system descriptor, ignores this since the spatial-frequency dependence of the noise power spectrum has been integrated out if the variance is used. A specification of variance alone for the noise performance of the array essentially forces the interpretation of "white noise," that is, noise power that is independent of spatial frequency. As can be seen from Eq. (9), white noise would result only if there were no spatial correlation (a purely random distribution) of array noise data values.

Data are presented in Sec. 5 to show that this often is not the case and that the specification of noise from the point of view of its spatial-frequency power spectrum is indeed a useful generalization to the concept of variance. Also, certain artifacts in the data are clearly pointed out from the Fourier domain point of view, which otherwise might be difficult to detect in the direct array data.

The spatial-frequency power spectrum is the means by which we can characterize the effect of the sensor noise on the various spatial frequencies of the image. The variance of the pixel levels has been seen to be easily related to the spatial-frequency power spectrum. The variance is proportional to the area under the power spectrum curve (in the discrete case, the sum of the extended frequency components).

## 5. EXPERIMENTAL RESULTS

In this section we consider data that demonstrate the utility of the Fourier domain description of sensor noise. The power spectra that are shown demonstrate that the spatial noise in an imaging array sensor can be nonwhite, that is, spatial-frequency dependent. It also is seen that the Fourier domain point of view facilitates detection and identification of harmonic sampling artifacts that are difficult to discern in the direct array data.

### 5.1. Effect of ensemble averaging on the noise spectrum

To obtain some intuitive feel for Eq. (11), it is instructive at this point to consider briefly a typical observation of the spatial noise seen on a sensor. Figure 2 shows n(i), the digitized noise values (one video line of 512 points), obtained with a zero value of incident irradiance. The ordinate of the plot was expanded so that the pixel-to-pixel variation may be seen more clearly. The waveform shown can be considered as a single observation of a random process. The value of each pixel varies with time, producing other data records for the same line of pixels that are statistically similar but different in detail. This is seen in Fig. 3, which shows another observation of the noise data along the same video line.

The proper means of displaying such statistical similarity is provided by the power spectrum. By taking the absolute value squared of the DFT of an observation record, we obtain an estimate of the true power spectrum of the random process. In this context, the power spectrum W(p,q) provides information regarding the spatial-frequency content of a statistical ensemble of such observations.
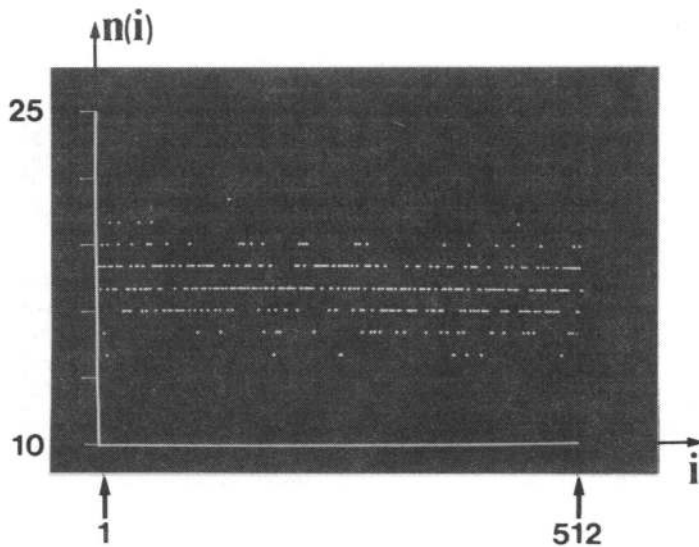
Fig. 2. One typical observation of n(i), the digitized values of sensor noise, for a single video line (512 points). These data were obtained with a zero value of incident irradiance on the sensor. The range of pixel values displayed is between 10 and 25 units, with the total possible range of digitized data values between 0 and 255.
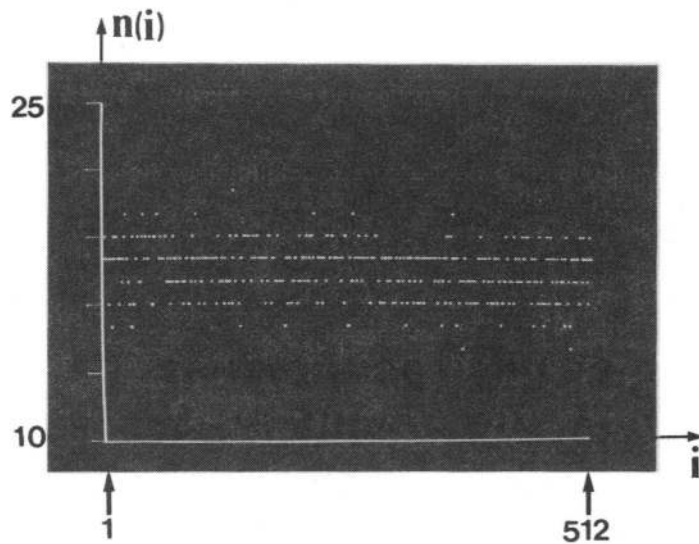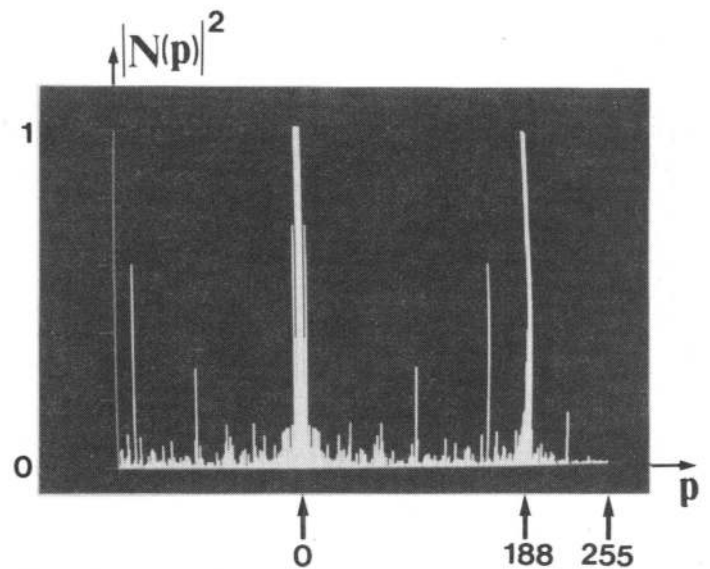


Fig. 4. Estimate of spatial-frequency power spectrum of sensor noise resulting from a single observation. $|N(p)|^2$ is the squared modulus of the DFT of a single noise observation n(i), such as was seen in Figs. 2 and 3. The spatial Nyquist frequency is at $p = 256$. The spectrum exhibits pronounced harmonic content at $p = 188$.
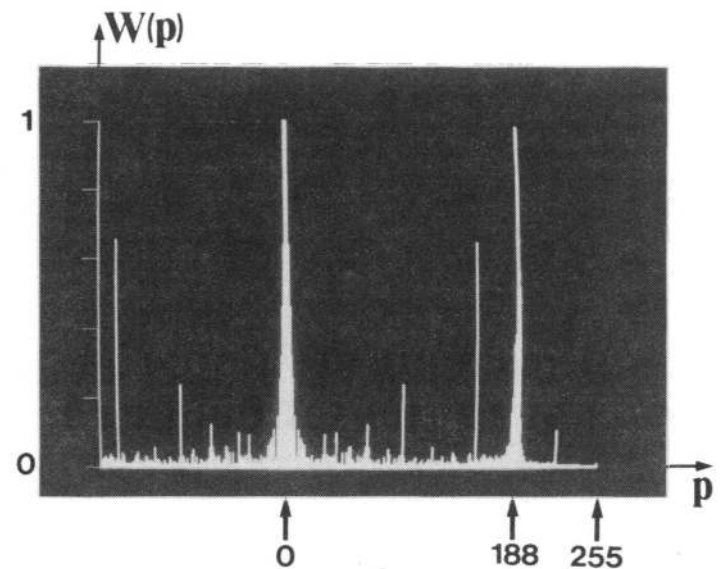


Fig. 3. Another observation of digitized noise values n(i) for the same video line as in Fig. 2. By comparison of the data from the two observations, evidence is seen of both a random variation and a statistically similar structure in the spatial noise.



Fig. 5. Estimate of spatial-frequency power spectrum resulting from an ensemble of 100 observations. W(p) is an average of $100|N(p)|^2$ spectral estimates obtained from a single video line. One such estimate $|N(p)|^2$ was seen in Fig. 4. One can observe some smoothing of the random variations in the spectrum, while the structure characteristic of the true spatial-frequency power spectrum has remained.

The advantage of ensemble averaging is seen by comparing Figs. 4 and 5. Figure 4 shows $|N(p)|^2$, the squared modulus of the DFT of a single individual noise record such as was seen in Figs. 2 and 3. It has an excess variability introduced by the fact that it is the spectrum of a single observation of a random process. The observations seen in Figs. 2 and 3 will have different DFTs due to the differences in the direct data. If one takes many individual spectra and averages them, the random variation in the spectrum decreases, converging as the square root of the number of independent observations to the true value of the power spectrum for that process as a whole. Figure 5 shows W(p), the power spectrum estimate resulting from averaging the squared modulus of the DFTs obtained

from 100 independent observations of the noise on a single line of video data.

The emergence of the true shape of the spatial-frequency power spectrum can be seen by comparing Figs. 4 and 5, although the difference is quite subtle. The ensemble averaging operation has introduced a smoothing effect, most visible in the distribution of low amplitude components in the spectrum. The structure that remains after ensemble averaging is a characteristic of the power spectrum of the sensor noise and cannot be averaged out. Its characterization is necessary for an accurate estimate to be made of an image spectrum.

## 5.2. Interpretation of the overall nonwhite nature of the spatial-frequency power spectrum

From the ensemble-averaged power spectrum in Fig. 5, it is readily seen that the sensor noise waveform has a power spectrum that is spatial-frequency dependent. The noise data exhibit a spatial correlation across the array, which determines the overall shape of the spectrum, in agreement with Eq. (9). The Fourier domain point of view is useful in this case since a specification of variance alone to characterize the array would ignore the fact that the sensor has relatively more noise at certain spatial frequencies than at others. This fact could be important in the choice of the spatial-frequency range of data to be presented to the sensor if an optimum signal-to-noise ratio is to be achieved.

## 5.3. Presence of a spike in the spectrum

Another example of the nonwhite nature of the sensor noise can be seen as a spike in the power spectrum data of Fig. 5. The power spectrum shows a major spike at $p = 188$ (and higher harmonics aliased into lower frequencies), corresponding to a pronounced harmonic content in the array data. The Fourier domain point of view is of particular utility in this case since this aspect of the nonwhite nature of the spectrum is quite difficult to notice in the direct data waveforms of Figs. 2 and 3. The existence of this harmonic would be completely masked by the use of variance to describe the noise performance of the sensor array.

The Fourier domain viewpoint also is useful in leading to the identification of the source of this spatial harmonic. In Eq. (4), we assumed implicitly that the number of photosites is the same as the number of digitized values in the data array. In the actual system used, 512 samples were taken along each horizontal line of analog video data from the sensor array. There were, however, only 188 actual photosites along a horizontal row of the sensor array. This mismatch between the size of the digital data array and the size of the array of photogenerated data values will be shown to be the source of the spatial harmonic seen in Fig. 5.

## 5.4. Origin of the spike in the horizontal spectrum

To investigate the origin of the spike seen in the noise power spectrum, let us first consider Fig. 6, which describes the actual processes the image data undergo in their conversion to an array of digitized values. Even though the analog video waveform is strictly a function of time, for the purposes of this discussion we consider it to be properly formatted into a rectangular raster format.

The first stage, describable by the operator $\mathscr{O}_1$, performs the conversion of continuous scene data $s(x,y)$ into photosite pixel values $f(l,m)$. This can be thought of as a local averaging of $s(x,y)$ over the photosite areas and an assignment of a single data value associated with the average scene energy falling upon each individual photosite:

$$s(x,y) \xrightarrow{\mathscr{O}_1} f(l,m) . \qquad (12)$$

The process of local averaging is describable[4] as a convolution of the scene $s(x,y)$ with a function describing the spatial responsivity of the photosite. We model this spatial responsivity as $\text{rect}(x_p, y_p)$, a two-dimensional rectangle of widths $x_p$
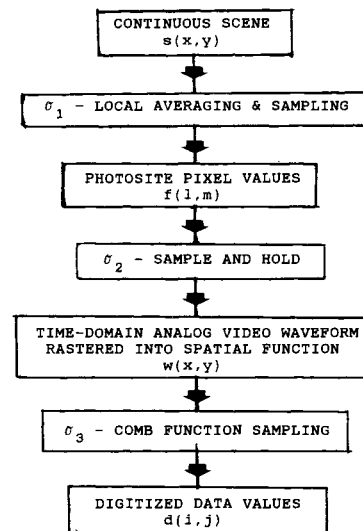


Fig. 6. Processes undergone by the data upon conversion from a continuous spatial scene s(x, y) to an array of digitized values d(i, j). In general, the number of actual photosite data values differs from the number of digitized values in the final data array. This is seen to lead to the spurious harmonic content in the data.

and $y_p$. The assignment of a value to each pixel is modeled as a comb-function sampling of the result of the convolution process. If the photosites are spaced at distances of $x_\Delta$ and $y_\Delta$, we can write the following expression for the action of operator $\mathscr{O}_1$:

$$f(l,m) = \left[ s(x,y) * \text{rect}\left( \frac{x}{x_p} , \frac{y}{y_p} \right) \right] \text{comb}\left( \frac{x}{x_\Delta} , \frac{y}{y_\Delta} \right) , \qquad (13)$$

where the $*$ denotes convolution. The data values produced by each photosite are converted in the sensor electronics to an analog video signal by means of a sample-and-hold operation. Figure 6 denotes this operation by $\mathscr{O}_2$. Considering the time waveform describing the video signal to be rastered into an equivalent two-dimensional spatial function, we can write operationally

$$f(l,m) \xrightarrow{\mathscr{O}_2} w(x,y) . \qquad (14)$$

This can be written explicitly as

$$w(x,y) = f(l,m) * \text{rect}\left( \frac{x}{x_\Delta} , \frac{y}{y_\Delta} \right) . \qquad (15)$$

The sampled-and-held analog video signal was described in the coordinates of the sensor plane. The array of pixel values, located at photosite locations, was convolved with a rectangle function equal to the photosite spacing in order to model the sampled-and-held rastered video waveform in the two-dimensional coordinates of the sensor array surface.

At this point, we are ready to consider the conversion of this analog waveform into an array of digitized values. We can write operationally

$$w(x,y) \xrightarrow{\mathscr{O}_3} d(i,j) . \qquad (16)$$

A comb-function sampling is used to model the A/D conversion of the sampled-and-held analog video waveform into the array of digitized values d(i,j), seen previously in Eq. (1):

$$d(i,j) = w(x,y)\mathrm{comb}\left(\frac{x}{X}, \frac{y}{Y}\right) .$$

(17)

The spacings X and Y are the effective spatial sampling interval, in sensor array coordinates, of the frame buffer data array.

The noise power spectrum of Fig. 5 indicates that the spike occurs at spatial frequency component p = 188. At this point, we note the existence of two different sampling intervals involved in the acquisition of digitized data from an array sensor. This aspect of the nonwhite nature of the spectrum can be explained as a beat frequency between the sampling lattice on the original detector array and the sampling lattice of the frame buffer.

The nature of this beat frequency can be clarified by combining Eqs. (13), (15), and (17). For compactness, we consider the one-dimensional version of these equations:

$$d(i) = \left(\left\{\left[s(x) * \mathrm{rect}\left(\frac{x}{x_p}\right)\right]\mathrm{comb}\left(\frac{x}{x_\Delta}\right)\right\} * \mathrm{rect}\left(\frac{x}{x_\Delta}\right)\right)\mathrm{comb}\left(\frac{x}{X}\right) .$$

(18)

The array of digitized values has an effective spacing on the sensor array surface of X. The sampled-and-held video waveform exhibits level changes only at a spacing of $x_\Delta$, the spatial sampling interval of the actual photosite locations. The final sampled array of values exhibits harmonic content corresponding to the fundamental spatial frequency of the detector spacing.

In the length of one line of sensors, there are 188 photosites. Thus, the fundamental spatial frequency of the photosite locations is exhibited at frequency component p = 188 of the power spectrum, which is what was observed in the data seen in Fig. 5.

### 5.5. Origin of the spike in the vertical spectrum

Power spectra of typical vertical data sets show behavior similar to that seen in Fig. 5 except that the spike occurs at a different frequency, q = 192. The origin of this harmonic content again is a beat phenomenon between sampling frequencies. The sensor produces 288 rows of data per video frame time. The frame buffer stores the entire video image on 480 lines, leaving the last 32 unused, so the effective sampling lattice of the frame buffer in this case is only 480 elements long. From Eq. (5), the spatial Nyquist limit in the vertical direction is

$$q_{Ny} = \frac{M}{2} = 240 .$$

(19)

The fundamental frequency of the vertical photosite spacing is beyond the Nyquist at q = 288. This frequency has an alias at

$$q_{alias} = q_{photosite} - 2(q_{photosite} - q_{Ny}) = 192 ,$$

(20)

which is the frequency at which a spike is seen in vertical spectra analogous to Fig. 5.

### 5.6. Discussion of spike artifacts

This type of artifact also can be considered a manifestation of a nonwhite power spectrum of the sensor response. While not strictly a noise effect per se, artifacts of this type often are termed "fixed-pattern noise." This is another case in which a specification of array performance by means of the variance alone would ignore the harmonic content inherent in the response of the sensor system. This also must be taken into account before an accurate estimate is made of the power spectrum of an image being acquired by such a sensor. This artifact could be alleviated with the proper sampling electronics, which would sample the output of each photosite only once, at the center of the individual response waveform for each pixel. However, for electronics that sample the analog video waveform from the sensor at some other periodicity than that of the sensor itself, this type of artifact is likely to corrupt the image spectrum.

### 5.7. Example from speckle experiments

One example of the use of Fourier domain techniques to characterize sensor noise performance is in the estimation of the spatial-frequency power spectrum of a laser speckle field. Figure 7 shows a line of direct array data for a speckle pattern having low contrast and low overall flux level. By comparison with Figs. 2 and 3, it is seen that the signal due to the speckle is masked almost completely by the spatial noise of the sensor. Averaging an unchanging speckle pattern over several observations would decrease the contribution from purely random noise in the sensors but would not alleviate the spatially structured nature of the noise spectrum. Figure 8 shows $|D(p)|^2$, the squared modulus of the DFT of a single observation of speckle data. By comparison with Fig. 5, it can be seen that there is a noticeable difference in the power spectra, corresponding to the spectrum of the received speckle pattern. The spatially structured noise of the sensor is seen to be characterized more easily in the Fourier domain than in the spatial domain. After suitable ensemble averaging is performed on both noise and speckle spectra,[5] an estimate of the spectrum of the speckle alone can be made by subtraction.

### 6. CONCLUSIONS

The spatial-frequency power spectrum of the noise seen on an imaging sensor is an important system descriptor. It conveniently characterizes the variance per unit spatial frequency interval, the total variance being proportional to the integral of the power spectrum. A measurement of variance alone for an imaging sensor does not take into account how the noise artifacts are distributed spatially and essentially assumes a random distribution. The more realistic model of spatial noise provided by the spatial-frequency power spectrum is of interest particularly for applications involving Fourier transformed data sets from imaging sensor arrays. Quantitative characterization of spatial noise artifacts in the Fourier domain is essential if accurate estimates of scene spectra are to be made.
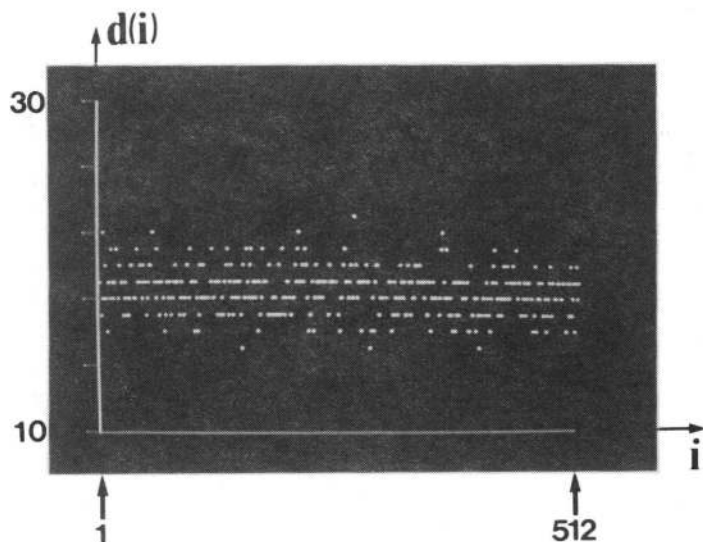
### 7. ACKNOWLEDGMENTS

Fig. 7. Data from one video line, showing digitized values for a low contrast speckle pattern. Plotted is d(i), the sum of the signal and the sensor noise. By comparison to Figs. 2 and 3, it is seen that the sensor noise almost completely masks the speckle signal in this case.
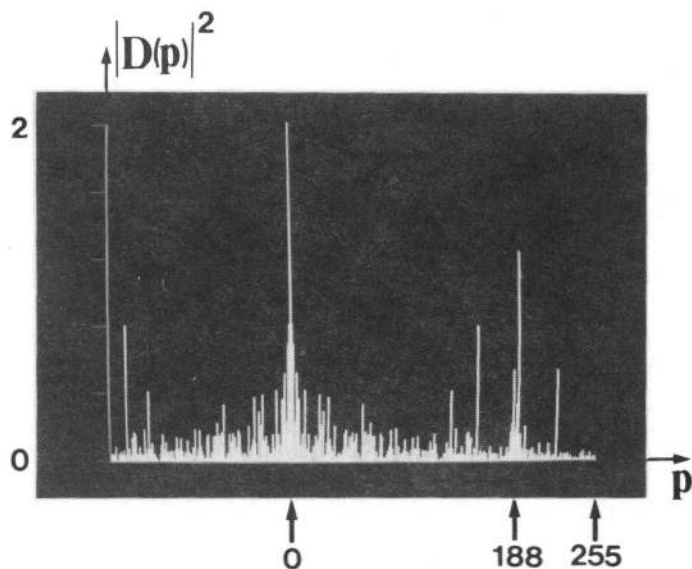
who were of great assistance in the programming of the host computer used in this study.

## 8. REFERENCES

1. G. Boreman and E. L. Dereniak, "Method for measuring modulation transfer function of charge-coupled devices using laser speckle," Opt. Eng. 25(1), 148–150 (1986).
2. W. K. Pratt, *Digital Image Processing*, pp. 235–236, Wiley-Interscience, New York (1978).
3. J. S. Bendat and A. G. Piersol, *Random Data, Analysis and Measurement Procedures*, pp. 189–193, Wiley-Interscience, New York (1971).
4. J. M. Lloyd, *Thermal Imaging Systems*, pp. 369–376, Plenum, New York (1975).
5. J. W. Goodman, "Statistical properties of laser speckle patterns," in *Laser Speckle and Related Phenomena*, J. C. Dainty, ed., pp. 35–40, Springer-Verlag, Berlin (1975).

**Glenn D. Boreman:** Biography and photograph appear with the paper "Compensation for gain nonuniformity and nonlinearity in HgCdTe charge-coupled-device focal planes" in this issue.

Fig. 8. Plotted is $|D(p)|^2$, the squared modulus of the DFT of the line of data d(i) seen in Fig. 7. Note that the scale in this spectrum plot is 0 to 2, which upon comparison with the spectrum of Fig. 5 (scale 0 to 1) shows clear evidence of the presence of the speckle signal. The presence of the speckle signal is seen much more clearly in the Fourier domain than in the direct spatial domain data of Fig. 7.