SAND78-1625
(Unlimited Release)
Printed October 1978

Coded Aperture Imaging - A Comparison Between Fresnel Zone Apertures and Uniformly Redundant Array Apertures

Kwong W. Chu
Systems Research Division III (1253)
Sandia Laboratories
Albuquerque, New Mexico 87185

Abstract

Using computer simulations, comparisons are made between a one-dimensional Fresnel zone aperture and a one dimensional uniformly redundant array aperture. In the former case, artifacts may produce severe distortions in the reconstructed image, and an iterative technique to remove these artifacts is described. In comparison to the Fresnel zone aperture, images reconstructed from uniformly redundant array apertures have few artifacts. A quantum noise analysis for the uniformly redundant array is presented to indicate the conditions under which a uniformly redundant array will produce a signal-to-noise advantage compared to an equivalent pinhole aperture.

Table of Contents

Figures

Figures (Continued) Page

I. Introduction

Sandia Laboratories is currently developing a coded aperture imaging system for investigating fuel motion within a nuclear reactor core.[1] The position and amount of fuel is determined by recording the shadowgram produced by fission gamma rays passing through a coded aperture and then decoding the shadowgram to recover the original source distribution. This paper reports the results from computer models which simulate the coded aperture imaging process. The two problems of major concern are (1) investigation of the severity of artifacts (or distortions) in the reconstructed images and (2) evaluation of conditions under which a coded aperture will improve the signal-to-noise ratio compared to a pinhole aperture with equivalent resolution.

Two one-dimensional apertures are investigated, the Fresnel zone aperture and a uniformly redundant array (URA) aperture. The Fresnel zone aperture has the unique property of propagating a coherent light beam through the shadowgram to reconstruct the source image; however, this technique suffers from the presence of artifacts. As will be shown, the distortions in the reconstructed image are so severe that some source functions are unrecognizable based on the reconstructed image produced by a one-dimensional on-axis Fresnel zone aperture. Because these artifacts are uniquely determined given a certain source function, it is in theory possible to remove them in the absence of noise (i.e. any stochastic process that degrades the signal). In practice, the elimination of artifacts may be difficult even in

situations where the noise is small. Furthermore, the added computational requirements detract from the original motivation for using Fresnel zone apertures, i.e. image reconstruction by simple light propagation.

Image reconstructions from shadowgrams produced by uniformly redundant array apertures are nearly free from artifacts. This is possible because the autocorrelation function of a uniformly redundant array closely approximates a delta function. In the absence of quantum noise, the reconstructed images are good representations of the original source function. Thus from the viewpoint of artifacts, the uniformly redundant array is superior to a Fresnel zone code and light propagation reconstruction.

Quantum noise studies were carried out using the uniformly redundant array. The radiation from a rectangular step function was considered to be the signal. Along with this signal, background radiation was assumed to be present but originating in two different ways: (1) modulated by the aperture, to be referred to as "source background" (this background radiation originates on the same side of the aperture as the fuel pin being studied); or (2) not modulated by the aperture, to be referred to as "stray background". It will be shown that in the former case, as background intensity becomes comparable to the signal intensity, the URA aperture will have a net signal-to-noise disadvantage compared to an equivalent pinhole. On the other hand, if background radiation unmodulated by the aperture becomes dominant, there will usually be a signal-to-noise advantage by using a URA aperture.

Thus, the final advantage gained by using a URA aperture, as opposed to a pinhole, can only be ascertained after a careful assessment of the nature of the background radiation.

II. Computer Models

Computer models to simulate coded aperture imaging were developed to evaluate the relative merits of various aperture codes and different reconstruction techniques. Due to the suitability of one dimensional apertures for the imaging of nuclear fuel pins (which are long, thin objects), primary effort was devoted to one-dimensional codes. This is fortuitous because the computer resources required are substantially less than for two-dimensional coded apertures, but at the same time, many of the significant characteristics of coded aperture imaging are illustrated by one-dimensional models.

The two codes that were investigated are the one-dimensional Fresnel one plate and the uniformly redundant array (URA). The URA's periodic autocorrelation function has flat "sidelobes."[2] That is:

let $U_0\ U_1\ U_2\ \cdots\ U_{N-1}$ represent a binary URA with N numbers (U_K is either 1 or 0); define its periodic autocorrelation function by

$$\rho(K) = \frac{1}{N}\sum_{j=0}^{N-1}(-1)^{U_j + U_{j+K}} \qquad (1)$$

where all subscripts are reduced modulo N if they are less than 0 or greater than N-1: The autocorrelation function has the property

$$\rho(K=0) = 1$$

$$\rho(K) = \frac{-1}{N} \text{ for } 1 \leq K \leq N-1 \qquad (2)$$

The flat sidelobes of the autocorrelation function makes the reconstruction of the source function from URA shadowgrams straight-forward; however, the periodicity property restricts the permissible spatial extent of the sources. As viewed in the detector plane, the entire shadowgram produced by all sources within the field of view must extend no more than twice the size of the aperture shadow produced by a single point source.

Because of the aforementioned restriction, the arrays representing the source functions and the aperture codes were made equal in length. The shadowgrams were generated by simply convolving the source function with the aperture code, i.e. if $\Psi(X)$ is the shadowgram measured in the detector plane, $S(X)$ is the source function, and $A(X)$ is the coded aperture, then

$$\Psi(X) = \int_{-\infty}^{\infty} S(X')\ A(X-X')\ dX' \qquad (3)$$

As actually implemented in the computer models, the aperture codes and the source functions were defined by arrays 512 elements (pixels) long, and the above integral must be approximated by

a discrete summation. Thus, after convolution, the resultant shadowgrams are contained in 1024 element arrays. It should be carefully noted that each array element represents one pixel or sample of the shadowgram; the pixels would correspond, for example, to the spatial sampling rate at which a microdensitometer scans photographic film on which the shadowgram is recorded. The pixel must not be confused with the resolution of the imaging system. One resolution element will span several pixels and is equivalent to the size of the smallest structure in the aperture.

Figure 1 depicts a one-dimensional Fresnel zone plate containing 16 zones and a 63 element URA code, both of which were used in the models. Although Fig. 1 shows a Fresnel code with an opaque center, the complementary code with a transparent center was also employed. The two codes were selected such that the narrowest zones in each code are approximately equal thereby making the resolution of both apertures the same. The Fresnel zone boundaries are given by ($n \leq 16$ for a 16 zone Fresnel code)

$$X_n = \sqrt{n}\, X_1 \tag{4}$$

where X_1 equals 64 pixels. The Fresnel code with the opaque center has a 41% transparency as compared to a clear aperture. The complementary code with a transparent center will consequently have a 59% transparency. The narrowest zone is 8.1 pixels wide ($(\sqrt{16} - \sqrt{15})\ 64 = 8.1$).

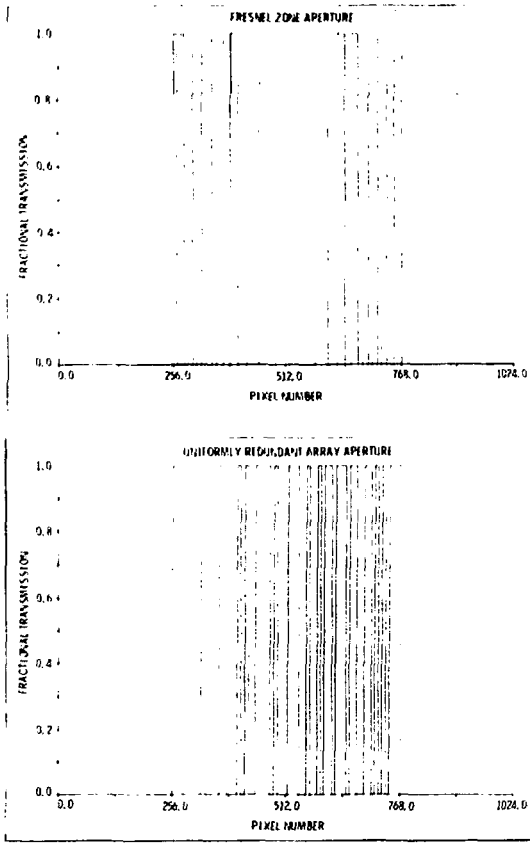

FIGURE 1. CODED APERTURES. This figure shows the transmission function of the two coded apertures employed in the computer simulations. Both apertures are 512 pixels wide, and the smallest structure in each aperture, corresponding to one resolution element, is approximately 8 pixels wide. The top graph shows a Fresnel code with an opaque center zone. The bottom graph shows a 63-element uniformly redundant array.

The URA code is represented by the sequence (11100000100000 1100010100111101000111001001011011101100110101011l). This sequence could be cyclically shifted to produce any one of 63 equivalent URA's; however, this particular sequence was selected to keep the outermost transparent regions as wide as possible. A real coded aperture must have a finite thickness; assuming the source is near the center of the field of view, it is desirable to make the outer zones as wide as possible to minimize vignetting affects. The URA is 512 pixels wide and the narrowest zones are 512/63 = 8.1 pixels wide. The transparency for the URA is 51% compared to a completely clear aperture.

The source function which was selected for much of the work is shown in Fig. 2. If the source function is convolved with the Fresnel zone aperture shown in Fig. 1, the resulting shadowgram is as shown in Fig. 2. If no noise is added, only relative amplitudes are important. Otherwise, the value of each pixel in the shadowgram is interpreted as the mean number of photons expected in that pixel. Noise is added by replacing the value of each shadowgram pixel by a new random number having a Poisson distribution whose mean value is equivalent to the no-noise value. The following sections will discuss how shadowgrams such as that shown in Fig. 2 may be decoded to recover the original source function.

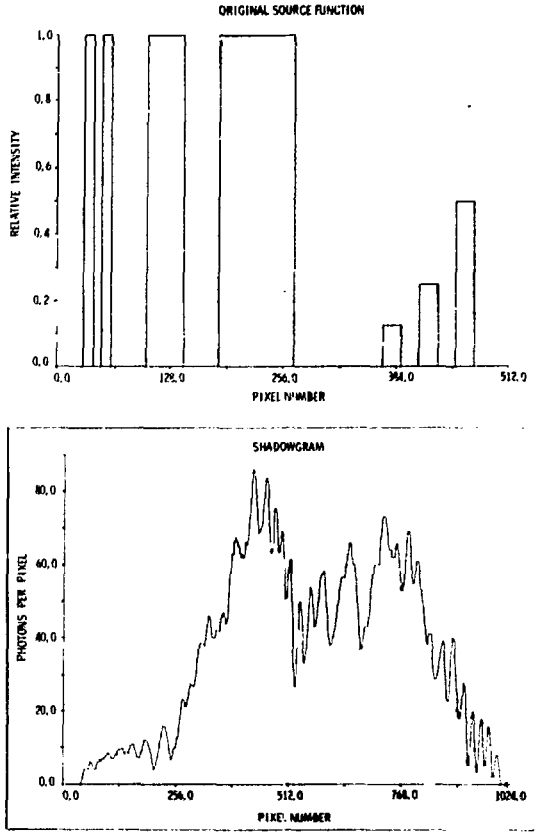

FIGURE 2. FRESNEL ZONE APERTURE SHADOWGRAM. The bottom graph illustrates a typical shadowgram produced by convolving the source function, shown in the top graph, with a Fresnel zone aperture shown in Figure 1. The absolute value of the shadowgram is not important unless quantum noise properties are being investigated. In such a case, the absolute value of each pixel of the shadowgram is interpreted as the expected number of photons received by the pixel.

III. Fresnel Zone Plate Propagation Reconstruction

The properties of Fresnel zone plates and their application in coded aperture imaging have been extensively discussed and will not be considered in detail here.[3] In the case of one-dimensional codes, the aperture acts as a pinhole camera along one axis and as a Fresnel code along the orthogonal direction. An analog propagation reconstruction may be carried out as sketched in Fig. 3. Very briefly, the mathematical description is as follows:

If the Kirchoff formulation for the diffraction of scalar waves is used to approximate the diffraction of light, one obtains for the one-dimensional case (after making several simplifying assumptions)

$$\Psi(X) = \frac{\exp\left[i\pi\left(\frac{2z_o}{\lambda} + \frac{X^2}{\lambda z_o} - \frac{1}{2}\right)\right]}{\lambda z_o} \int \Psi(X') \exp\left[i \frac{\pi}{\lambda z_o} X'^2\right] \exp\left[-i \frac{2\pi X X'}{\lambda z_o}\right] dX' \qquad (5)$$

where Ψ is a complex function representing the amplitude and phase of the light

X' lies in the shadowgram plane such that $\Psi(X')$ describes the one-dimensional shadowgram

X lies in the reconstruction plane such that $\Psi(X)$ describes the reconstructed source function

λ is the wavelength of light

Z_o is the distance between the shadowgram plane and the reconstruction plane

The shadowgram, $\Psi(X')$, is merely the coded aperture, $A(X)$, convolved with the source function, $S(X)$, i.e.

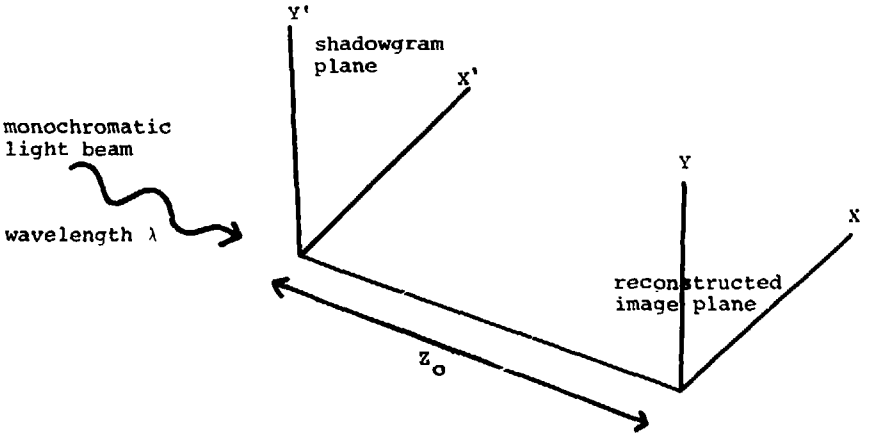

FIGURE 3. FRESNEL ZONE APERTURE - LIGHT PROPAGATION RECONSTRUCTION. In order to reconstruct the source function, the shadowgram is placed in the X'-Y' plane. A monochromatic light beam will be diffracted as described by equation 5. This diffraction pattern will reconstruct the original source function in the X-Y plane where $\frac{\pi}{z_o \lambda} = \frac{-P}{X_1^2}$.

The reconstruction produced by the term of order P will then be in-focus at the X-Y plane.

$$\Psi(X') = \int_{-\infty}^{\infty} S(X)\ A(X'-X)\ dX\ .$$

The Fresnel code aperture can be represented as an exponential series

$$A(X) = \frac{1}{2} + \sum_{\substack{-\infty \\ P\ \text{odd}}}^{\infty} \frac{\exp\left(i\pi P \frac{X^2}{X_1^2}\right)}{i\pi P} \quad \text{central zone transparent}$$

$$= \frac{1}{2} - \sum_{\substack{-\infty \\ P\ \text{odd}}}^{\infty} \frac{\exp\left(i\pi P \frac{X^2}{X_1^2}\right)}{i\pi P} \quad \text{central zone opaque} \qquad (6)$$

where

X_1 = radius of first zone.

For the moment, A(X) is assumed to be unbounded as a function of X in order to demonstrate image reconstruction by the use of a light wave. Later, the effect of a finite aperture size will be discussed.

The reconstructive property of the Fresnel code can be demonstrated by setting $\frac{\pi}{Z_o \lambda}$ in equation (5) equal to $-\frac{\pi}{X_1^2}$ and by considering only the P = 1 term in equation (6), i.e., let

$$A(X) = \frac{\exp\left(i\pi \frac{X^2}{X_1^2}\right)}{i\pi} \tag{7}$$

If these substitutions are made in equation (5), one finds

$$\Psi(X) = \frac{\exp\left(\frac{i2\pi Z_o}{\lambda}\right)}{\pi} S(X) \tag{8}$$

In this case, $\frac{\pi}{\lambda Z_o} = \frac{-\pi}{X_1^2}$ has been arbitrarily selected so that the P = 1 term of the Fresnel code reconstructs the image. Any term of order P could have been selected by setting $\frac{\pi}{Z_o \lambda} = -\frac{P}{X_1^2}$.

The unfortunate aspect of propagation reconstruction is that only one term of the expansion series for the Fresnel code contributes to the reconstruction of the source. Any physically realizable Fresnel code aperture will contain other non-reconstructing terms, and these terms may produce serious distortions in the reconstructed image. As will be shown, the removal of artifacts produced by these non-reconstructing terms is the primary problem with using propagation reconstruction. To study this problem, several shadowgrams of the source function shown in Fig. 2 were made to reconstruct the original source function from these shadowgrams.

Figure 4 shows the reconstructed source function that occurs if a physically realizable Fresnel code aperture with a transparent center (the complement of the Fresnel code in Fig. 1) is used. The results are rather poor. In comparison, if the complex reconstructing aperture function

$$A(X) = \frac{\exp\left(i\pi \frac{X^2}{X_1^2}\right)}{i\pi} \qquad 0 \leq |X| \leq 4\,X_1$$

$$A(X) = 0.0 \qquad |X| > 4\,X_1 \tag{9}$$

is used to create the shadowgram, the reconstructed source function shown in Fig. 5 is obtained. It depicts the original source with good fidelity. Truncation of the aperture code after 16 zones is mainly responsible for inaccuracies present (i.e. the finite aperture reduces the high spatial frequency response). Unfortunately, it is not physically possible to fabricate a complex aperture when imaging gamma ray sources, and unless a means is found to remove the affects of non-reconstructing terms, propagation reconstruction is not an attractive reconstruction technique.

Figure 6 shows the contributions of the various terms using a point source. The DC term and the P = -1 term are the two most important contributors toward the generation of undesired artifacts. The phases change slowly compared to the resolution of the aperture, and thus destructive and constructive interference

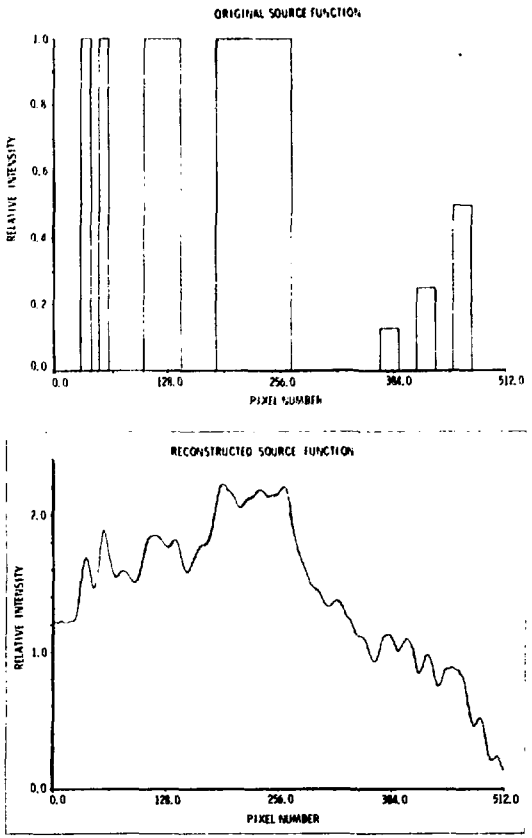

FIGURE 4. PROPAGATION RECONSTRUCTION - ALL TERMS OF FRESNEL CODE PRESENT. The bottom graph shows the reconstructed image obtained if the aperture is a simple Fresnel zone function. The actual Fresnel zone code used had a transparent central zone and is the complement of that shown in Figure 1.

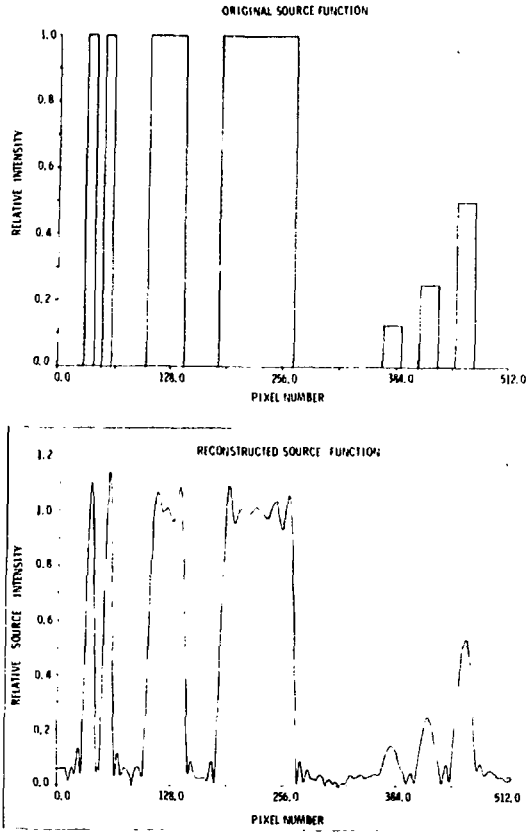

FIGURE 5. PROPAGATION RECONSTRUCTION - COMPLEX FRESNEL CODE APERTURE. The bottom graph shows the reconstructed image obtained if the aperture includes only the reconstructing term i.e., the aperture function is given by equation (9). This aperture is a complex function and is not physically realizable.

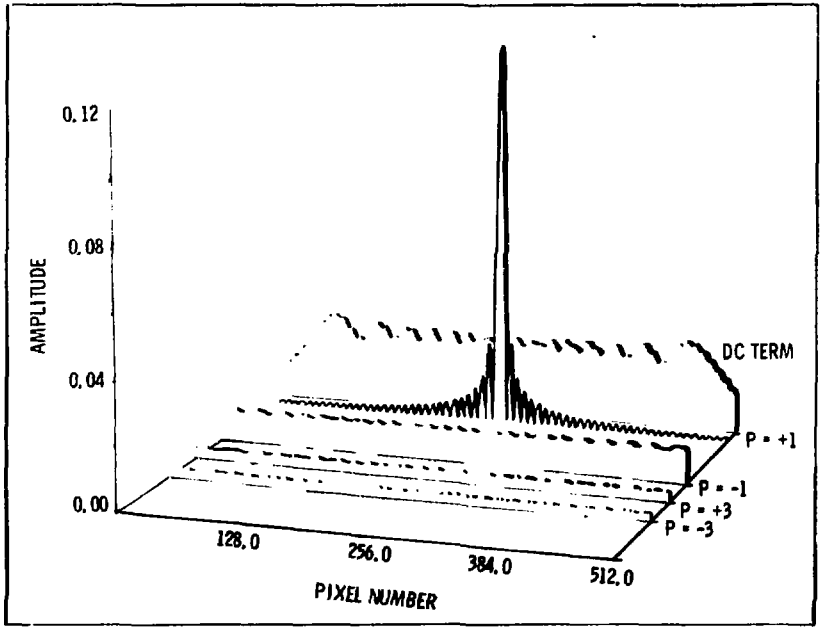

FIGURE 6a. PROPAGATION RECONSTRUCTION - POINT SPREAD FUNCTIONS - AMPLITUDE. This figure shows the relative amplitude of the point spread function for individual terms of the Fresnel zone expansion series up through the third order. The P = +1 term produces reconstruction, but unfortunately other terms are also present in a physically realizable aperture and produce distortions in the reconstructed image.

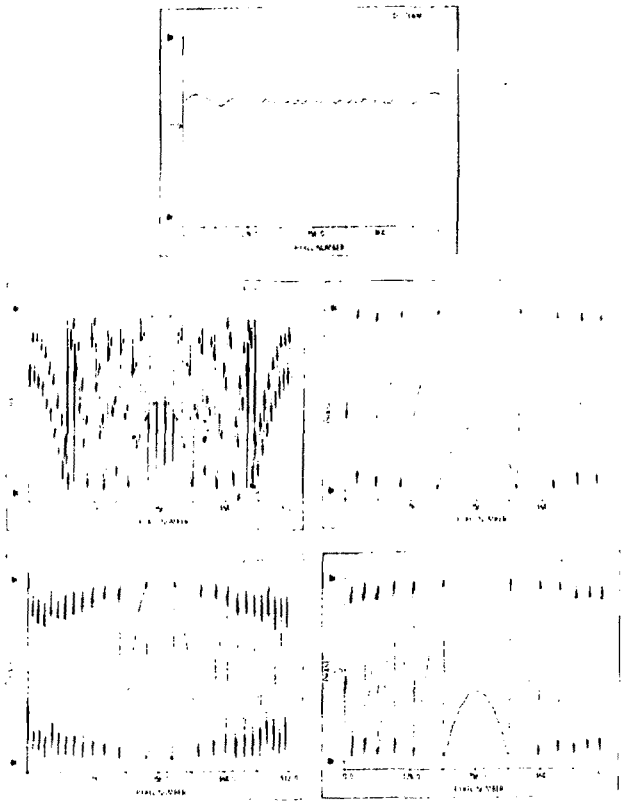

FIGURE 6b. PROPAGATION RECONSTRUCTION - POINT SPREAD FUNCTIONS - PHASE. This figure shows the phase of the point spread function for individual terms of the Fresnel zone expansion series up through the third order. Note that the analomous spikes in the phase function of the reconstructing term (P = +1) are a consequence of the amplitude of this term passing through zero at which point the phase is undefined. For values of the amplitude sufficiently close to zero, the calculated phase is dominated by computational roundoff errors.

between the various terms can produce substantial sidelobe structure. The reconstructing term, P = 1, can be calculated theoretically for the case of a finite Fresnel zone aperture and a single point source; let the source function be an ideal point source ($S(X) = S_o\delta(X)$), and use equation (9) to represent a complex Fresnel code aperture. If these conditions are applied to equation (5), the resulting point response function is

$$\Psi(X) = -\exp\left(-i\pi \frac{X^2}{X_1^2}\right) \frac{\sin\left(\frac{8\pi X}{X_1}\right)}{\pi X} S_o$$

where a constant phase term has been dropped. To compare with the P = 1 term in Figure 6, X_1 (the radius of first Fresnel zone) must be set to 64 pixels in the above equation. In terms of the relative amplitude and position of the zeroes, the computer results agree with the above equation.

Figure 7a shows the point response for a physically realizable Fresnel code, and as can be seen, it is far from an ideal point function. The importance of terms of order 3 and higher may be assessed by using an aperture containing only the DC term, the P = 1 term and the P = -1 term, i.e.,

$$A(X) = 0.5 + 0.5 \sin\left(\pi \frac{X^2}{X_1^2}\right). \qquad (10)$$

The propagation reconstruction using this aperture is shown in Figs. 7b and 8. This reconstruction is not significantly better than that in Fig. 4. It may thus be concluded that the DC term and the diverging P - -1 term are the dominant sources of undesirable artifacts and that the use of a physically realizable sinusoidal Fresnel aperture will not significantly improve the image reconstruction by removing the third and higher order terms.

The DC term may be removed by using the physically unrealizable aperture

$$A(X) = \sin\left(\pi \frac{X^2}{X_1^2}\right) \tag{11}$$

and the resulting reconstruction is shown in Figs. 7c and 9. Substantial improvements are obtained; however, major distortions still exist. The central broad source is highly distorted, and none of the smaller amplitude sources can be distinguished. The conclusion is that unless a means can be found to reliably remove the effects of the DC term and the diverging term, a straightforward propagation reconstruction will not provide a reliable representation of the source function.

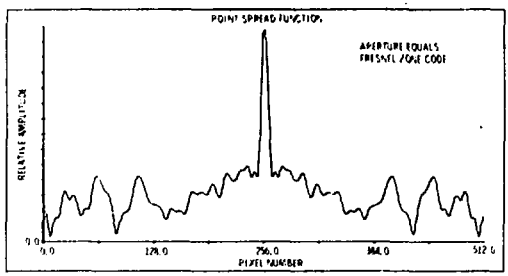

Figure 7a

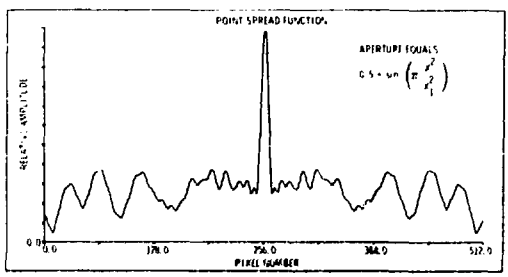

Figure 7b

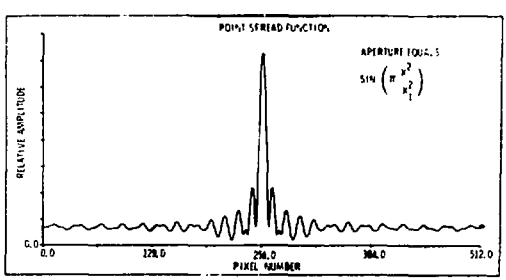

Figure 7c

FIGURE 7. PROBAGATION RECONSTRUCTION OF A POINT SOURCE. The propagation reconstruction of a point source for three different apertures is shown. The aperture is physically realizable in the top two graphs, but not in the bottom graph. In the top graph, all orders of the exponential expansion series for the Fresnel code are present. In the middle graph, only the DC term, and the positive and negative first order terms are present. In the bottom graph, the DC term has been removed and only the first order terms are present.

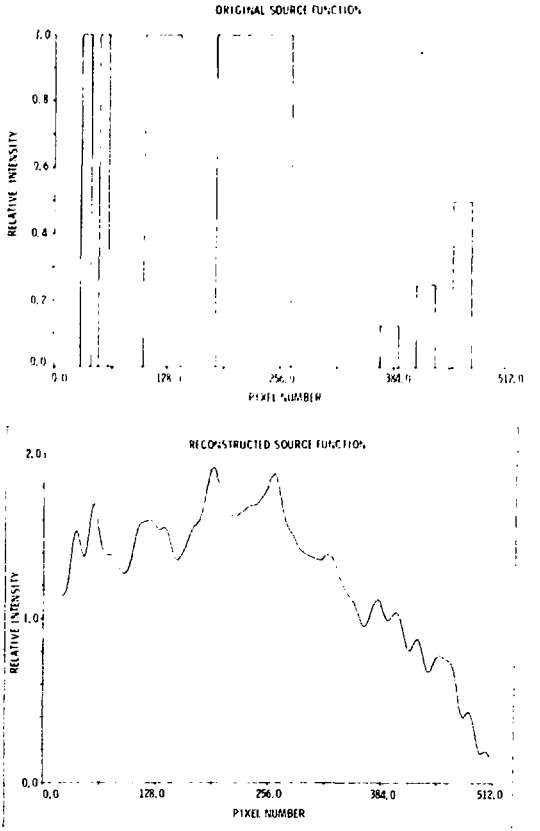

FIGURE 8. PROPAGATION RECONSTRUCTION - ONLY DC AND FIRST ORDER TERMS. Lower graph shows reconstructed image where the aperture function is given by $A(X) = 0.5 + 0.5 \sin (\pi X^2/X_1^2)$. This aperture eliminates terms of order three and higher; however, the reconstructed image is not significantly improved over the case where all higher order terms are present (see Figure 4).

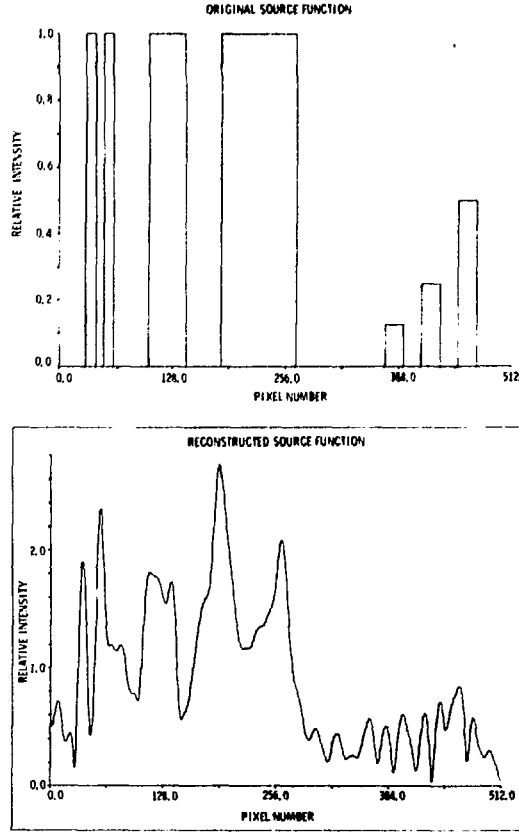

FIGURE 9. PROPAGATION RECONSTRUCTION - ONLY FIRST ORDER TERMS. Lower graph shows reconstructed image where the aperture function is given by $A(X) = \sin(\pi X^2/X_1^2)$. This aperture function eliminates the DC term; however, note that it is not physically realizable. The quality of this reconstructed image should be compared to cases where other higher order terms and DC terms are present (see Figure 4 and Figure 8).

IV. Fresnel Zone Plate Correlation Reconstruction

An alternative to the propagation reconstruction which may be applied to any coded aperture is the correlation reconstruction. The basic idea is to determine a decoding function such that when it is convolved with the coded aperture, the result approximates a delta function. In the case of the Fresnel zone code, various decoding functions have been tried (e.g. $\sin(\pi x^2/x_1^2)$), but all suffer the common problem of producing significant sidelobes. A simple correlation reconstruction applied to Fresnel zone plate shadowgrams will not yield good reconstructions. The results are roughly comparable to propagation reconstruction where the interfering effects of the non-reconstructing terms are not removed (compare Figures 4 and 11).

Figure 10b shows the point source responses using the correlation reconstruction where the decoding function is the Fresnel code. The average value of the sidelobes can be made flatter by choosing a "balanced" decoding function (i.e., a decoding function that is both positive and negative); the differences in the point source responses are illustrated in Fig. 10. In both cases, however, substantial sidelobe structure remains, and the reconstructed images are poor. A technique for improving the reconstructions is still to be illustrated. Because this technique works better using the Fresnel code itself as the decoding function, balanced decoding functions will not be considered further.

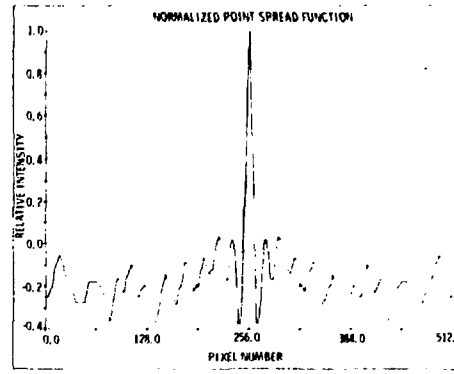

Figure 10a

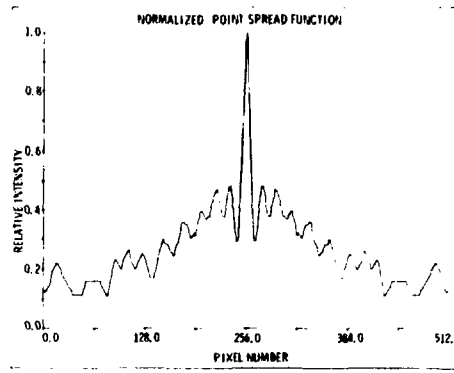

Figure 10b

FIGURE 10. CORRELATION RECONSTRUCTION - POINT SPREAD FUNCTION. This figure shows the point spread function for a Fresnel zone aperture in which two decoding functions are used. In the top figure, the decoding function is the Fresnel zone aperture modified through a DC level shift such that the decoding function is either +1 (where the aperture is transparent) or -1 (where the aperture is opaque). In the bottom graph, the decoding function is simply the Fresnel aperture, and the decoding function is either +1 (where the aperture is transparent) or 0 (where the aperture is opaque).

A correlation reconstruction is illustrated in Fig. 11, and it does not provide a satisfactory reconstruction. Some technique is desirable to compensate for the sidelobes. Various iterative techniques have been used.[4] The one to be demonstrated here will be called a "point source decomposition" technique. Basically an attempt is made to find a set of delta functions which, when convolved with the point response function (Fig. 10b), will approximate the correlation reconstruction in Fig. 11. This set of delta functions may be considered as a point source approximation to the original source function, or in other words, the original source function has been decomposed into a series of point sources. This is achieved by an iterative technique, as follows:

1) The correlated reconstruction (Fig. 11) is scanned to determine the position with the largest absolute value
2) Centered at this position, the point response function (Fig. 10b) is subtracted from the correlation reconstruction. The amplitude of the point response function to be subtracted is scaled to some fixed percentage of the amplitude of the correlated reconstruction at that location. The precise percentage is not critical. Usually some value from about 20% to about 5% is used. The major tradeoff to be made is if the percentage is high, the iterative procedure may not converge properly, but if the percentage is too low, the slow convergence wastes computer time without any commensurate improvement of the final image.

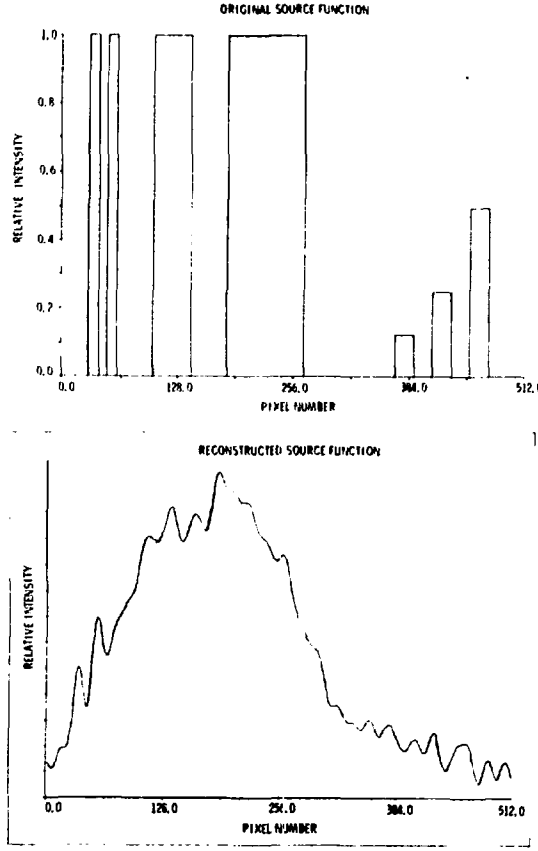

FIGURE 11. CORRELATION RECONSTRUCTION. The bottom graph shows the results of a simple correlation reconstruction where the aperture and the decoding function are a Fresnel zone function. The reconstructed image should be compared with the propagation reconstruction shown in Figure 4.

3) The computer stores the position and amplitude of the subtracted point response function. These amplitudes and positions represent the point sources which define the final reconstructed image.

4) After the above steps are completed, the original correlated reconstruction has been modified through subtracting the point response function. This modified correlated reconstruction is now used in place of the original, and all four of these steps are applied iteratively.

The process is terminated either after a fixed number of iterations or until the residual correlated reconstruction is in the noise level. As the final step, a running spatial average (usually a box window) is made over the point sources determined by the above technique and this constitutes the final reconstructed source. Figure 12 shows the resulting reconstruction when this technique is applied.

For many situations, this technique provides an acceptable reconstruction. There is a tendency for this technique to produce oscillations in regions where the original source function is flat. Note, however, these oscillations are at a wavelength comparable to or less than the resolution of the aperture. Small artifacts may also be generated, as can be seen from the feature just to the right of the central source (see arrow Fig. 12).

In some circumstances the tomographic capability of coded aperture imaging is important. Tomographic capability is possible

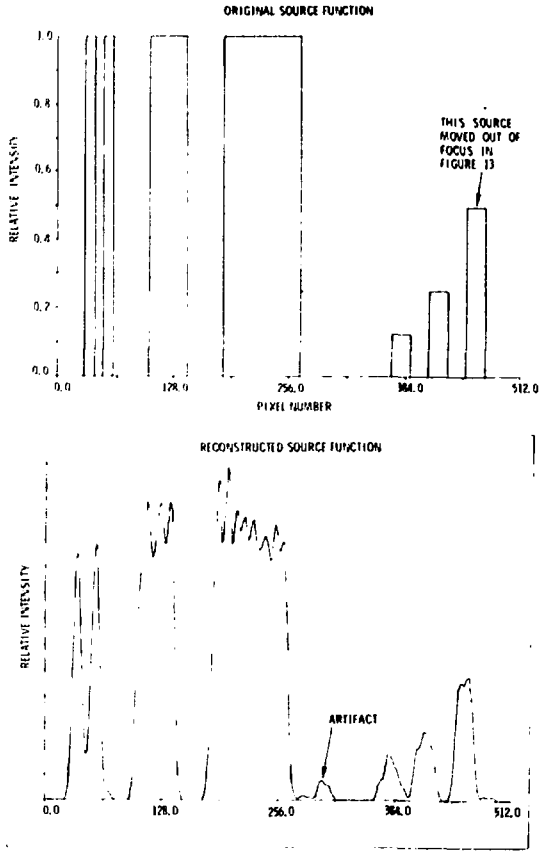

FIGURE 12. CORRELATION RECONSTRUCTION WITH POINT-SOURCE-DECOMPOSITION. Bottom graph shows the reconstructed image obtained after a point-source-decomposition algorithm as described in the text is applied. Three hundred iterations were used to produce the reconstructed image. Note the presence of an artifact (see arrow). Top graph represents original source function and also indicates which source is to be moved out of focus for investigating the depth of field.

because the size of the cast shadow will depend on the distance between the source and aperture. If this tomographic capability is to be useful in situations where several sources lie in separate planes, sources should go out of focus smoothly. To investigate this, the rightmost source in Fig. 12 was moved back by reducing the size of the shadows it casts by 2%, 5%, 10%, and 20%. These shadowgrams were then reconstructed in a manner identical to that shown in Fig. 12. The results are shown in Fig. 13.

For this particular reconstruction technique, the tomographic capability is poor. Note the source which was moved represents only 6% of the total number of photons collected at the detector plane; yet, when it is moved out of focus, it has significant impact on the reconstruction of the sources that remained in focus. If one were imaging three sources in three different planes, there might be confusion regarding the shape of each source due to the other out-of-focus sources. The conclusion is that for this Fresnel code and the point-source-decomposition reconstruction technique, tomographic capability is poor.

The particular reconstruction technique described here works best in situations where there are several compact sources and little background. In situations where there are broad continuous sources covering most of the field of view, the point source decomposition fails. Figure 14 shows the same source function as before, superimposed on a background. The reconstructed image is also shown in Fig. 14. This reconstructive technique fails because too many point sources are required to represent the background

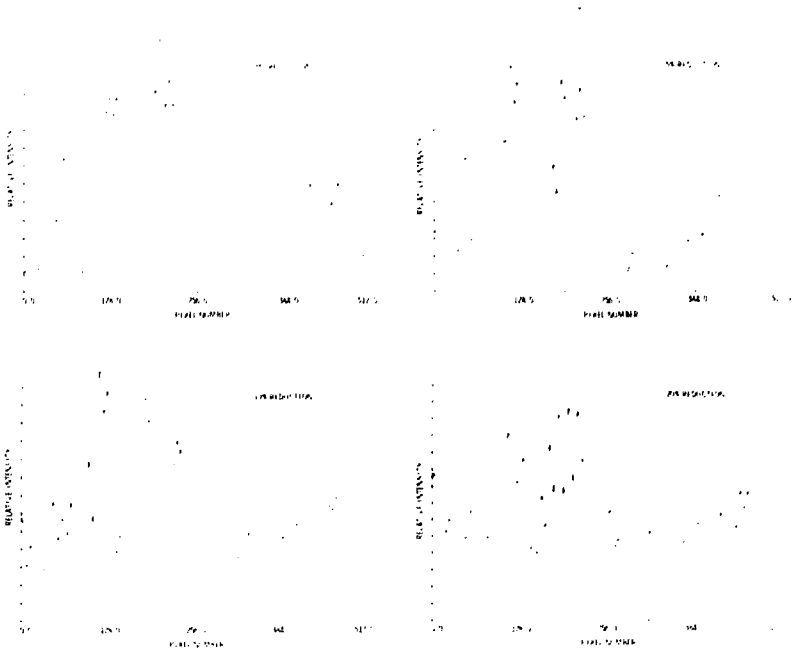

FIGURE 13. DEPTH OF FIELD - FRESNEL CODE. This figure demonstrates the affect of an out-of-focus source on the reconstructed image. The rightmost source (see top graph, Fig. 12) was placed out of focus by reducing the spatial scale of the aperture shadows it produced by the percentages indicated. The reconstruction procedure was the same as for Fig. 12.

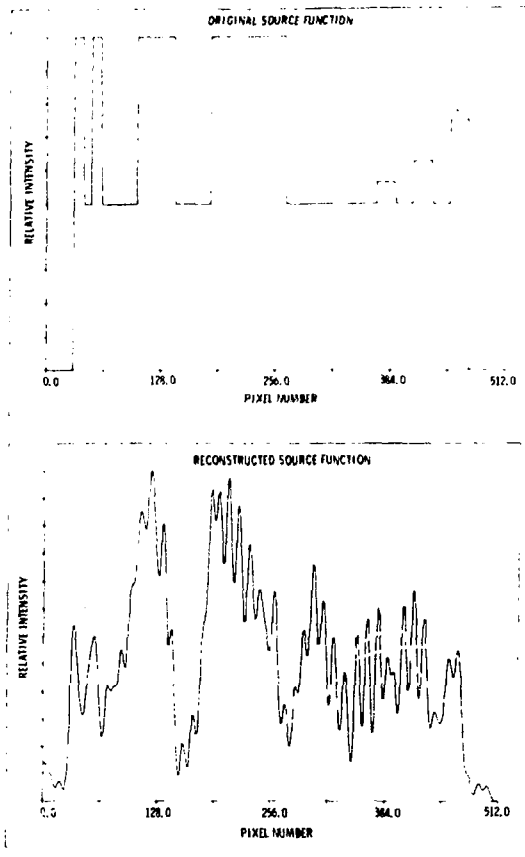

FIGURE 14. CORRELATION RECONSTRUCTION WITH SOURCE BACKGROUND. Lower graph shows the failure of the point-source-decomposition technique in situations where the source function has a broad background component. The image reconstruction procedure was identical to that used for Fig. 12; only the original source function was modified.

and the iterations fail to converge. Unless the background is removed prior to reconstruction, this reconstructive technique will not work under high background conditions.

V. Uniformly Redundant Array

The preceding two sections discussed some of the problems associated with image reconstruction using a Fresnel zone aperture code. As an alternate code, the uniformly redundant array is particularly well suited to correlation reconstruction techniques due to the flat sidelobes of its autocorrelation function. The URA code to be considered has already been discussed and is shown in Fig. 1.

The decoding function is basically the same as the URA aperture itself except that it is made nearly symmetric about zero so that the sidelobes average out to zero; i.e., if U(X) is the URA code shown in Fig. 1, the decoding function U'(X), is

$$U'(X) = 2.0\ U(X) - C \qquad (12)$$

where C is a constant nearly equal to one, adjusted so that all the sidelobes average to zero. (As used in the computer models, C = 0.988). In this manner, no zero baseline shifts are produced as the strength or number of sources are increased or decreased.

If the decoding function is convolved with the coded aperture, the point spread function obtained is as shown in Fig. 15. The side lobes are not perfectly flat because the 63 element URA

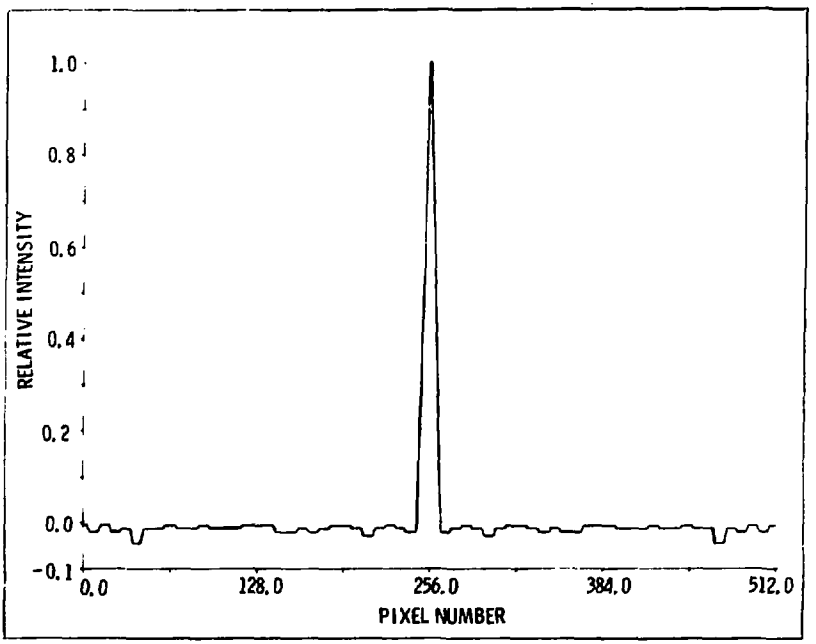

FIGURE 15. UNIFORMLY REDUNDANT ARRAY - POINT SPREAD FUNCTION. This figure shows the reconstructed image of a point source using a uniformly redundant array as the aperture. Note the flatness of the sidelobes as compared to a Fresnel zone aperture (see Figures 7 and 10).

code is actually represented by 512 pixels. Thus when the convolution is calculated, one does not always shift by an integral number of elements. As will be shown, the non-flat sidelobes may generate some artifacts, but these artifacts are insignificant as compared to a Fresnel code (compare point spread function in Figs. 7 and 10 to Fig. 15).

The source functions with and without backgrounds were reconstructed using the URA code. Figures 16 and 17 show the reconstructed image. Even in the presence of background, the reconstructions are for the most part good. Note however, the smallest source (see arrow) is stronger in comparison to the other sources than it should be. This is a chance artifact of this particular source function created mainly by the dominant source situated near the center.

The tomographic qualities of this imaging technique were investigated in the same manner as the Fresnel code. The rightmost source was "moved" or "placed out of focus" by reducing the size of its shadowgrams by 2%, 5%, 10% and 20%. The results are shown in Fig. 18 and should be compared with Fig. 13. Again, an out-of-focus source will distort the in-focus sources being imaged; however, the distortions are significantly less than for Fresnel code using a correlation and point-source-decomposition reconstruction technique.

Considered in terms of ease of reconstruction, fidelity of reconstruction, and depth of field response, the URA code appears to be superior to the Fresnel code. The greatest attraction of

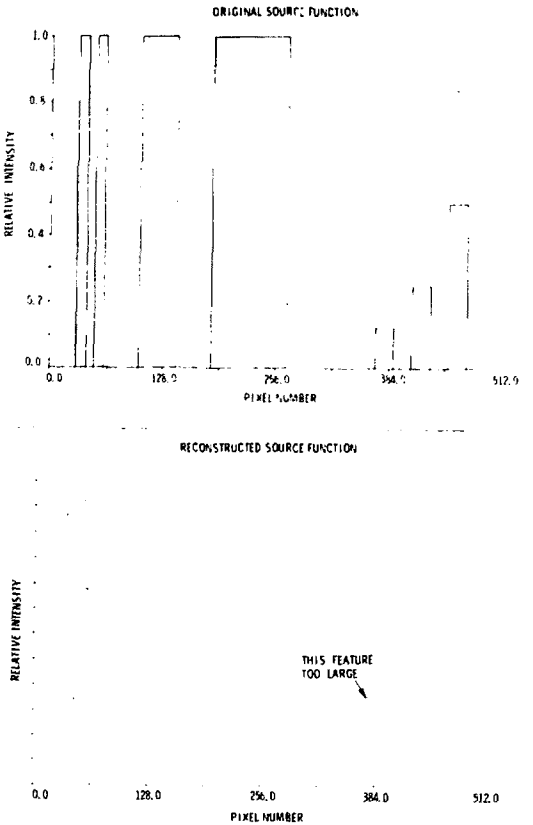

FIGURE 16. URA RECONSTRUCTION- NO SOURCE BACKGROUND. The bottom graph shows the reconstructed image obtained by using the 63 element URA code. Note the presence of some distortion, as indicated by one source which is too strong relative to the other source (see arrow). Top graph shows original source function.

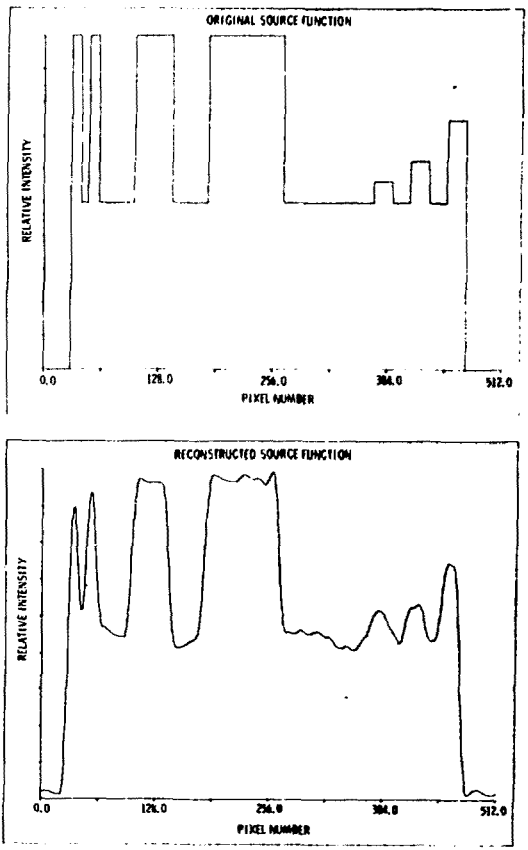

FIGURE 17. URA RECONSTRUCTION WITH SOURCE BACKGROUND. Reconstruction procedure is the same as for Figure 16, except that source function has been modified to include a background. Contrast the good fidelity of this reconstruction with that of the Fresnel code using the point-source-decomposition procedure (see Fig. 14).

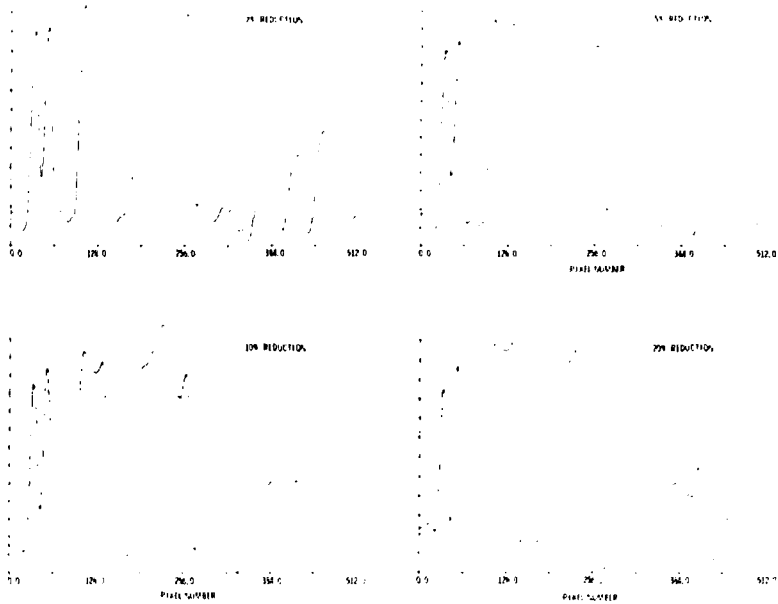

FIGURE 18. DEPTH OF FIELD - URA CODE. This figure demonstrates the affect of an out-of-focus source on the reconstructed image. Following the same procedure as in Fig. 13, the rightmost source was placed out of focus by reducing the spatial scale of the aperture shadows it produced by the percentages indicated above. A comparison with Fig. 13 should be made.

the Fresnel code is the possibility of having a simple and direct analog reconstruction technique (i.e., light propagation). To realize an acceptable reconstruction, the DC term and the diverging term artifacts must be eliminated. Unless this can be realized on a practical basis, the URA code represents a better reconstruction method. In principal, both the Fresnel zone and the URA aperture can be utilized to provide tomographic source reconstruction. In practice, however, a three-dimensional source distribution actually increases the difficulty of obtaining images free from artifacts, because in addition to any problems associated with the two-dimensional in-focus source function, additional artifacts will be present due to other out-of-focus sources. For the rather limited number of sources employed in this study, the URA reconstructions are distorted less by out-of-focus sources than the Fresnel code using a point-source-decomposition reconstruction.

VI. Quantum Noise

The primary goal of any imaging system is to provide the best possible estimate of the source distribution given certain physical limitations (such as integration time, aperture size, detector resolution, etc.) In the case of coded aperture imaging of gamma ray sources, two primary factors influence the fidelity of the reconstructed image: (1) systematic distortions inherent in the imaging system and (2) statistical uncertainties in the signal. In order to facilitate the analysis, these two factors are being treated separately; however, the ability to correct for systematic

distortions is intimately related to the presence of statistical uncertainties in the signal, and a complete analysis should properly treat both problems simultaneously.

In the previous sections, computer simulations demonstrated that the reconstructed images contain distortions due to artifacts produced by non-zero sidelobes. Because these distortions are systematic, they may theoretically be completely eliminated if the aperture transmission and the signal (shadowgram) were continuous functions measured with infinite precision. As a practical problem, this theoretical idealization cannot be achieved, and some artifacts will always persist in the reconstructed images. Thus, depending on the extent to which artifacts can be eliminated, an upper limit will be placed on the dynamic range of the reconstructed image (the ratio of the strongest to the weakest source which may be distinguished in an image). In comparison to the Fresnel zone code, the URA point response function has nearly flat sidelobes, and the artifacts in the reconstructed image can often be ignored. In this sense, the URA aperture is superior to the Fresnel zone aperture, and for this reason, the URA code will be selected for further evaluation with respect to its quantum noise performance (see Barrett and DeMeester: 1974 for Fresnel zone aperture noise analysis).

Quantum noise fluctuations are important in any physical measurement that counts photons, and the statistical fluctuations will be characterized by the Poisson distribution. A shadowgram represents the integrated gamma ray photon flux (number of photons

per unit area) at the detector plane as a function of position. Due to quantum noise, the integrated photon flux measured over any finite time interval and any finite area will randomly fluctuate about the "true" mean (the value that would be measured assuming an infinite integration time and a time invariant physical situation). Consequently, the reconstructed image will also be subject to random variations. Thus, as used in this section, noise strictly refers to statistical fluctuations of the reconstructed image due to quantum noise. In this sense, a reconstructed image may have an excellent signal-to-noise ratio, yet be a poor representation of the actual source function if there are significant artifacts present. The standard deviation of the reconstructed image will be used as the quantitative measure of noise.

In order to calculate the noise, the reconstruction process should first be considered. The one-dimensional shadowgram represents the number of photons detected per unit distance (or per pixel). Let the array $\{A_i\}$ represent the shadowgram where A_i is the number of photons in the i^{th} pixel. Let the array $\{D_i\}$ be the decoding function where $D_i \simeq \pm 1$ for all i $(0 \leq i \leq M-1)$ where M is the total number of pixels in the shadowgram. The k^{th} pixel of the reconstructed image is given by

$$R_k = \sum_{i=0}^{M-1} D_{i+k} A_i \qquad (13)$$

where the subscript of the decoding array is interpreted as modulo M. In the presence of quantum noise, each A_i is a random variable

with a Poisson distribution having some mean value, $\bar{A}_i$, and a standard deviation $\sigma_{A_i} = \sqrt{\bar{A}_i}$. Since the statistical variations from pixel to pixel are independent and since $|D_i| \simeq 1$ for all i, the standard deviation of R_k is

$$\sigma^2_{R_k} = \sum_{i=0}^{M-1} \sigma^2_{A_i} = \sum_{i=0}^{M-1} \bar{A}_i \tag{14}$$

or

$$\sigma_{R_k} = \sqrt{\begin{array}{c}\text{expected total number of gamma} \\ \text{rays in the entire shadowgram}\end{array}} = \sqrt{\gamma_T}$$

Thus, the result is very simple: The standard deviation of the reconstructed image is the square root of an expected total number of gamma photons in the entire shadowgram. This noise is position invariant, and the analysis applies to all gamma photons recorded in the shadowgram irrespective of whether or not they reach the shadowgram by passing through the URA aperture.

Next the reconstructed image amplitude should be considered. A single point source will cast a shadow which is a scaled replica of the URA aperture. This point source is reconstructed when k in equation 13 is such that the decoding function matches the aperture shadow (let $k = k_o$ for this fiducial point source). In this case, R_{k_o} will equal the total number of photons from the fiducial point source which pass through the aperture and reach the shadowgram. Let this number of photons be P. Now consider the addition of a

nearby point source having the same amplitude as the fiducial point source. The number of photons reaching the shadowgram from this nearby point source is also P; however, because the second point source is in a slightly different position, the aperture shadow it casts is displaced. As a consequence, when R_{k_o} is calculated as in equation 13, this second point source will contribute some number less than P due to the mismatch between its aperture shadow and the decoding function. In fact, as point sources are positioned further and further from the fiducial point source, they will contribute less and less to R_{k_o}. Once the point sources are displaced greater than one resolution element from the fiducial source, they will, on the average, contribute nothing more to R_{k_o}. One resolution element corresponds to the displacement between two point sources such that their respective aperture shadows are shifted by one element (i.e. the smallest structure in the aperture).

The preceding analysis for point sources may be extended to calculate the reconstructed image amplitude from continuum sources with constant intensity. Any continuum source may be approximated by a series of point sources spaced much closer than the resolution of the coded aperture. The key point is that photons emitted from a region of the order of a resolution element will add coherently to one another in equation 13. A detailed analysis provides a simple result which is described as follows: Partition the source function into segments having the size of one resolution element. Each resolution element will emit a certain number of photons per unit time, out of which some fraction will actually pass through

the aperture and reach the shadowgram. It is this fractional number of photons that equation 13 calculates.

To quantify this result, let γ_{URA} represent the number of photons collected in the shadowgram which were emitted by one resolution element of a uniform source with unit intensity I_o. For any other uniform source of intensity I, the reconstructed image amplitude is

$$R_{URA} = \gamma_{URA} I \tag{15}$$

where I is normalized to I_o. Note that R_{URA} is a number representing a count of photons. γ_{URA} is a constant number which is characteristic of the coded aperture imaging system under consideration. If the geometry of the imaging system and the transmission function of the aperture are known, the absolute value of γ_{URA} may be calculated; however, for this analysis, its absolute value is irrelevant.

The signal-to-noise ratio of the reconstruction will now be defined as the ratio between the reconstructed image amplitude divided by the RMS fluctuation, or, using equations 14 and 15:

$$\left(\frac{S}{N}\right)_{URA} = \frac{\gamma_{URA} I}{\sqrt{\gamma_T}} \tag{16}$$

Note that this result is independent of the spatial sample rate at which the shadowgram is measured.

The primary motivation for using a coded aperture as opposed to an equivalent pinhole aperture is to improve the quantum-limited signal-to-noise ratio. Therefore, an important problem is the assessment of conditions under which a URA aperture will have a signal-to-noise advantage over an equivalent pinhole aperture. A pinhole aperture and a URA aperture are considered equivalent if they provide the same resolution. Basically, this implies that the pinhole aperture is the same size as the smallest structures in the URA aperture.

In contrast to a coded aperture, a pinhole aperture forms an image directly on the detector plane, and the individual resolution elements of this projected image may be considered as being independent. In other words, the image amplitude corresponding to any one resolution element of the source function contains no noise originating from other portions of the source function separated by more than a resolution element. Thus a pinhole aperture might be perferred under a situation where a relatively weak source is to be imaged in the presence of a very strong source located at least a few resolution elements away.

Mathematically, the results may be expressed as follows. Similar to the analysis for the URA aperture, let γ_{PH} represent the number of photons reaching the detector plane from a resolution element of the gamma ray source. γ_{PH} is to be referenced to the same source intensity unit, I_o, as for the URA aperture. For a continuous source of intensity I, the pinhole image will have an amplitude

$$R_{PH} = \gamma_{PH} I \tag{17}$$

where I is normalized to I_o. R_{PH} represents the photon count per resolution element of the projected image from a source with intensity I. Thus, defining the signal-to-noise ratio of the pinhole aperture as the image amplitude divided by the RMS fluctuation, we have

$$\left(\frac{S}{N}\right)_{PH} = \frac{\gamma_{PH} I}{\sqrt{\gamma_R}} = \frac{\text{total number signal photons per resolution element}}{\sqrt{\text{total number of photons per resolution element}}} \tag{18}$$

In order to compare the pinhole aperture with the URA aperture, the quantities γ_{URA}, γ_{PH}, γ_R and γ_T must be related to one and another. Because this work is in support of imaging nuclear fuel pins, a specific source function as shown in Fig. 19 will be considered. The rectangular object of intensity A (representing the fuel pin) is superimposed on a uniform background of intensity B_1 spanning the entire field of view. The rectangular object spans W resolution elements (assume $W > 1$), and the entire field of view spans F resolution elements. A and B_1 are normalized to I_o so that equations 15 and 17 are applicable. Besides the source background radiation B_1 which must pass through the aperture, let the detector be subject to a uniform flux B_2, where B_2 is the number of photons per resolution element at the detector plane and represents stray background radiation which is not modulated by the

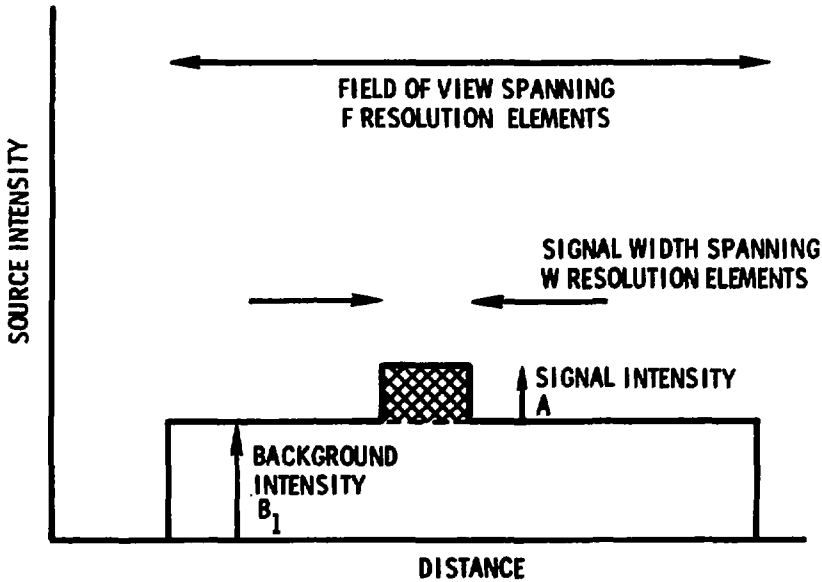

FIGURE 19. SOURCE FUNCTION FOR SIGNAL-TO-NOISE RATIO ANALYSIS. This figure shows the type of source function assumed for the signal-to-noise analysis. The signal consists of the cross-hatched step function with intensity A. It corresponds to the nuclear fuel pin being studied. It was assumed that a uniform source background with intensity B_1 and spanning the entire field of view would also be present.

aperture. Having made these definitions, the signal-to-noise ratio of the URA and the pinhole aperture may be compared.

The signal will be defined as the amplitude of just the rectangular object whose intensity is A (see Fig. 19). For the pinhole aperture

$$\left(\frac{S}{N}\right)_{PH} = \frac{\text{number of photons reaching the detector from one resolution element of rectangular object}}{\sqrt{\text{total number of photons per resolution element at detector plane}}}$$

$$= \frac{\gamma_{PH}A}{\sqrt{\gamma_{PH}A + \gamma_{PH}B_1 + B_2}} \tag{19}$$

For the URA aperture

$$\left(\frac{S}{N}\right)_{URA} = \frac{\text{number of photons reaching the detector from one resolution element of rectangular object}}{\sqrt{\text{total number of photons in entire shadowgram}}}$$

$$= \frac{\gamma_{URA}A}{\sqrt{W\gamma_{URA}A + F\gamma_{URA}B_1 + 2FB_2}} \tag{20}$$

Further define

$$\alpha = \frac{\gamma_{URA}}{\gamma_{PH}} = \frac{\text{transmitting area of URA aperture}}{\text{transmitting area of pinhole aperture}} \tag{21}$$

Equations 19, 20 and 21 may be combined to yield

$$\frac{(S/N)_{URA}}{(S/N)_{PH}} = \sqrt{\frac{\alpha}{W}} \frac{\left(1 + \frac{B_1}{A} + \alpha \frac{B_2}{A\,\gamma_{URA}}\right)^{\frac{1}{2}}}{\left(1 + \frac{F}{W}\frac{B_1}{A} + 2F \frac{B_2}{A\gamma_{URA}}\right)^{\frac{1}{2}}} \tag{22}$$

In order to better understand the implications of the above equation, consider first the case where all background radiation is modulated by the aperture, i.e., let $B_2 = 0$. For a 63 element URA code,

$$F = 63$$

$$\alpha = 32$$

For a 127 element URA code,

$$F = 127$$

$$\alpha = 64$$

Fig. 20 plots $\frac{(S/N)_{URA}}{(S/N)_{PH}}$ versus $\frac{B_1}{A}$

for a case where the signal spans one resolution element (W=1) and the signal spans 10 resolution elements W=10). It can be seen that the signal-to-noise advantage of the URA aperture decreases rapidly as the background builds up, and once the background intensity is greater than the signal intensity, there is

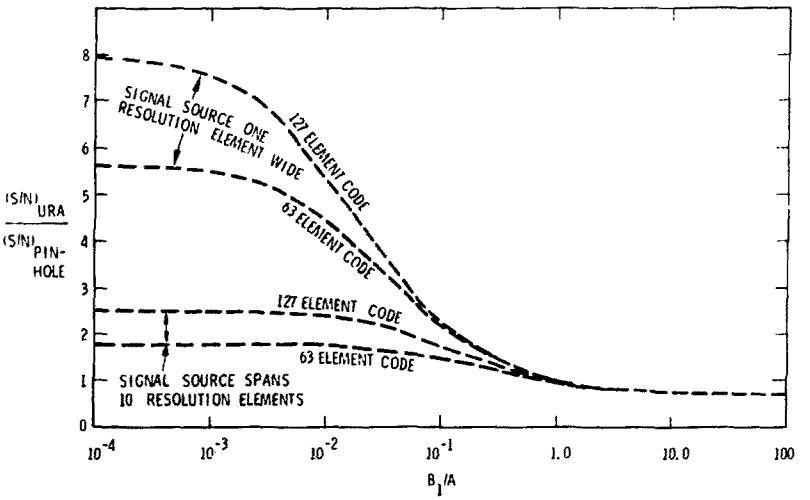

FIGURE 20. SIGNAL-TO-NOISE RATIO VERSUS SOURCE BACKGROUND INTENSITY. This figure plots the signal-to-noise ratio advantage of a URA as compared to an equivalent pinhole as a function of the source-background-intensity to signal-intensity ratio. Note that once the source background intensity is comarable to or greater than the signal intensity, *there is a net disadvantage* in using the URA aperture. The stray background intensity is assumed to be zero.

a slight net disadvantage to using a URA aperture versus an equivalent pinhole.

Next consider the case where $B_1 = 0$. Plots of $\frac{(S/N)_{URA}}{(S/N)_{PH}}$ versus $\frac{B_2}{A\gamma_{URA}}$ are shown in Fig. 21. Note that

$$\frac{B_2}{A\gamma_{URA}} = \frac{\text{number background photons per resolution element}}{\text{number of signal photons per resolution element using URA aperture.}}$$

In this case, even in the limit of $B_2 \to \infty$, the URA will have a signal-to-noise advantage as long as the signal spans no more than 8 resolution elements for a 63 element URA, or 16 resolution elements for a 127 element URA. Thus, the signal-to-noise advantage (or disadvantage) of a URA versus a pinhole aperture depends on the relative portion of background radiation passing through the aperture versus stray background radiation.

VII. Summary and Conclusions

Using computer simulations, comparisons between a one-Fresnel zone aperture and a one-dimensional uniformly redundant array aperture have been made. The primary purpose was to assess the relative quality of the reconstructed images obtained from these two coded apertures. A theoretical analysis of the quantum noise limit for the URA aperture was also made. In this case, the major emphasis was to evaluate the conditions under which a URA aperture would

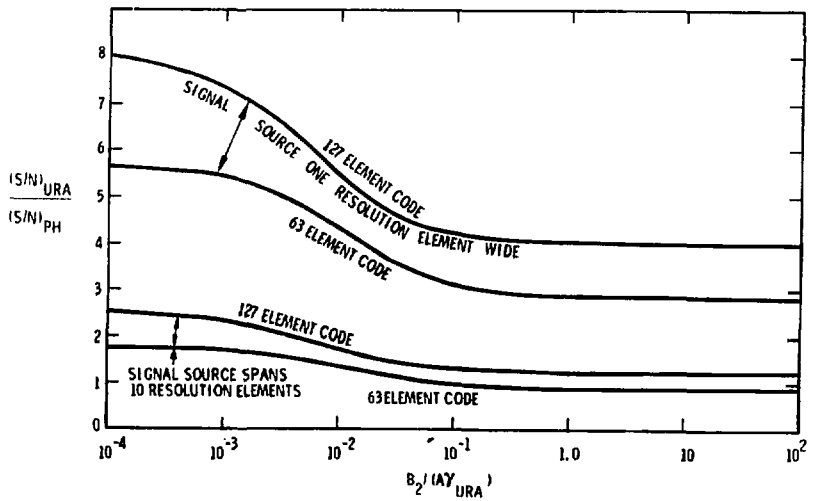

FIGURE 21. SIGNAL-TO-NOISE RATIO VERSUS STRAY BACKGROUND INTENSITY. This figure plots the signal-to-noise ratio advantage of a URA as compared to an equivalent pinhole as a function of the stray-background-intensity to signal-intensity ratio. Note that in contrast with the source background situation, there will usually be a signal-to-noise advantage using the URA as long as the signal itself is not too wide.

improve the signal-to-noise ratio as compared to an equivalent pinhole aperture.

The foremost difficulty that should be initially resolved with any Fresnel zone imaging system is the removal of unwanted artifacts generated by the non-image-reconstructing terms in the exponential expansion series representing the Fresnel code. The DC term and the first order diverging term are the predominant causes of distortions in the reconstructed images. For a relatively complex source function of the type used in this study, the distortions are sufficiently severe that the original source function is nearly unrecognizable based on the reconstructed image. Fortunately these distortions are systematic, and corrections may be made to eliminate the artifacts. One possible iterative technique which is effective in removing artifacts for some source functions (referred to as a point-source-decomposition technique) was described. Using this method, fairly accurate reconstructed images are obtained as long as the original source function contains no substantial broad background component.

In comparison with the Fresnel code, the uniformly redundant array aperture produces images nearly free from artifacts. This is a natural consequence of the flat sidelobe structure in the URA autocorrelation function. Evaluated in terms of the seriousness of undesired artifacts, the URA aperture is superior to the Fresnel code.

Both the Fresnel code and the URA were used to assess the affects of a 3-dimensional source distribution. All shadow-cast

imaging techniques have depth of field capability because the spatial scale of the shadow will depend on the source-to-aperture distance. On the one hand, this may be regarded as a desirable attribute giving the coded aperture imaging technique tomographic capability. However, on a more practical basis, out-of-focus sources can only complicate the problem of removing artifacts from the image of the in-focus sources. In this respect, a correlation reconstruction using the point-source-decomposition procedure is significantly more prone than the URA aperture to produce distorted in-focus images due to the presence of an out-of-focus source.

Quantum noise investigations were performed for the URA aperture. The comparative signal-to-noise advantage of the URA aperture versus an equivalent pinhole aperture was evaluated as a function of background radiation intensity. In this respect, it is important to distinguish source background radiation, which is modulated by the aperture, from stray background radiation, which is not affected by the aperture. Assuming the source background uniformly spans the field of view, the URA quickly loses its signal-to-noise advantage once the source background intensity exceeds or is comparable to the signal intensity. In contrast, a URA aperture will usually retain a signal-to-noise advantage over a pinhole regardless of how intense the stray background radiation becomes. Thus, in order to evaluate the signal-to-noise advantage of the URA aperture, not only the amount but the origin of all background radiation must be determined.

Acknowledgements

The author wishes to thank K. T. Stalker and J. G. Kelly for many useful discussions.

References

1. Kelly, J. G. and K. T. Stalker, "Coded Aperture Imaging of Reactor-Illuminated LMFBR Fuel Pins", American Nuclear Society Winter Meeting: San Francisco, CA, Nov. 27-Dec. 2, 1977.

2. Mac Williams, F. J. and N. J. A. Sloane, "Pseudo-Random Sequences and Arrays", Proceedings IEEE, 64, 1715, Dec. 1976.

3. Barrett, H. H. and F. A. Harrigan, "Fresnel Zone Plate Imaging of Gamma Rays; Theory", Applied Optics 12 2686, Nov. 1973.

4. Simpson, R. G., H. H. Barrett, J. G. Kelly, and K. T. Stalker, "Some Applications of One-Dimensional Coded Apertures", SPIE Symposium, Reston, VA, Apr. 18-21, 1977.

5. Barrett, H. H. and G. D. DeMeester, "Quantum Noise in Fresnel Zone Plate Imaging", Applied Optics 13 1100, 1974.

DISTRIBUTION:

H. H. Barrett
Optical Sciences Center
University of Arizona
Tucson, Arizona 85721

1200 L. D. Smith
1250 R. G. Clem
1253 G. W. Gobeli
1253 K. W. Chu (15)
1253 K. T. Stalker
1255 P. A. Stokes
2163 K. R. Hessel
4420 J. V. Walker
4423 J. E. Powell
4423 L. M. Choate
4423 J. G. Kelly
4423 D. A. McArthur
8266 A. E. Aas (2)
3141 C. A. Pepmueller (Actg) (5)
3151 W. L. Garner (3)
For: DOE/TIC (Unlimited Release)
DOE/TIC (25)
(R. P. Campbell, 3171-1)