UCRL - 76274
PREPRINT
CONF-741212--5

LAWRENCE LIVERMORE LABORATORY
University of California/Livermore, California

BEST LINEAR DECODING OF RANDOM MASK IMAGES

J. W. Woods, M. P. Ekstrom, T. M. Palmieri, & R. E. Twogood

December 11, 1974

THIS PAPER WAS PREPARED FOR SUBMISSION TO THE
1974 IEEE Nuclear Science Symposium & 14th Scintillation
& Semiconductor Symposium

BEST LINEAR DECODING OF RANDOM MASK IMAGES

J. W. Woods, M. P. Ekstrom, T. M. Palmieri, & R. E. Twogood
Lawrence Livermore Laboratory, University of California
Livermore, CA 94550

Abstract

In 1968 Dicke proposed coded imaging of x and γ rays via random pinholes. Since then, many authors have agreed with him that this technique can offer significant image improvement. We present a best linear decoding of the coded image and show its superiority over the conventional matched filter decoding. Experimental results in the visible light region are presented.

I. Introduction

In 1968 Dicke[1] proposed a random pinhole array be used in x-ray astronomy as a means of increasing the S/N ratio of a pinhole camera. Since then much experimental work has been done on various types of coded apertures. Barrett et al[2] have considered a Fresnel zone plate for the medical imaging application. MacDonald, et al[3] presented at last year's Nuclear Science Symposium an application of the Fresnel zone plate aperture together with a multiwire proportional counter and computer reconstruction. In [1-3] and also [4] & [5], the image reconstruction was performed by a matched filtering of the coded image.

Generally two problems arise when the random pinhole array of the Fresnel zone plate is used for imaging extended objects, i.e., objects defined over extended regions of space. First, a large d.c. background appears in the decoded image. Second, the S/N ratio deteriorates markedly as the object size increases.

In this paper we articulate the coded aperture problem in a system theory setting. Using the concepts of statistical communication theory,[7,8] we formulate the problem of image reconstruction as one of optimally decoding the measured image. We present a best linear decoding algorithm for processing the coded aperture image and show its superiority over the conventional matched filter decoding used by others.

II. Background: Coded Aperture Imaging

To provide adequate background, we first present a brief discussion of Dicke's coded aperture camera. It consists of a perforated entrance plate with each perforation acting as a single pinhole. A point source radiates the face plate and an image of the mask is formed in the image plane. This image is recorded using radiographic film, a multiwire proportional counter, or other types of x-ray detection devices. The basic motivation for using multiple pinholes is simply to improve the signal-to-noise ratio (S/N) of the single pinhole imaging by introducing repeated measurements. This is easy to see for the case above of a single isolated point source. If there are a total of P pinholes in the mask and the noise fields incident to the measurement are such that the noise contributed at each image point is independent, the S/N improvement is achieved by registering and averaging the individual images from each pinhole. This results in an improvement factor of $\sqrt{P}$. Clearly for P on the order of say 10^4, this improvement will be considerable.

For the case of extended objects, (as arises in the medical application), the situation is substantially more complex. Let us assume that the distribution of pinholes in the aperture plane is given by the zero-one function $h(m,n)$.* That is, at those points in the plane which are pinholes, we have $h(m,n) = 1$; otherwise, it is zero. Now, for a point source of magnitude $s(m',n')$ the formed image will be $s(m',n')h(m-m',n-n')$. We can consider an extended object to be an array of point sources $\{s(m',n')\}$ defined over a finite region in the distant object plane. Because a translation of the object results in a translation of the image, we have the image of this extended object as

$$y(m,n) = \sum_{m',n'}\sum s(m',n')\, h(m-m',n-n') \quad . \tag{1}$$

In this representation the coded image field $y(m,n)$ is related to the object field by a discrete two-dimensional (2-D) convolution, where $h(m,n)$ is the so-called point-spread response of the mask. It is important to note that the arrays in (1) represent the magnitude of fields (for example, $g(m,n)$ may be the illuminance of the image field), and are therefore positive functions with non-zero means.

Because error is associated with all physical measurements, the accessible image $a(m,n)$ is made up of the coded image $y(m,n)$ and measurement noise $w(m,n)$. This error or noise is taken to be strictly additive. It arises from insufficient source statistics, limited precision instruments, background effects, sensor noise, etc. This noise may be functionally related to or independent of the coded image.

Now, in order to decode such imagery Dicke suggested a matched filter decoder. His rationale was if $h(m,n)$ was chosen such that its correlation was sharply peaked at the origin, the cross-correlation of $h(m,n)$ with the coded images would give peaks in locations corresponding to point sources. That is Dicke's decoder involved forming

$$d(m,n) = \sum_{m',n'}\sum a(m',n')\, h(m'-m,n'-n). \tag{2}$$

This is, of course, a matched filter or correlation operation and appears to have been attractive to Dicke and others[1-5] to a great extent because of its ease and variety of implementations.

The performance of this decoder, particularly when imaging extended objects, is dependent on $h(m,n)$ having a sharply peaked correlation function. To achieve this Dicke introduced a random pinhole mask (Barrett and others[2,3] have used a Fresnel zone plate). For $h(m,n)$ defined on an N-by-N grid, this mask has a random distribution of pinholes such that the probability of any particular point being a pinhole is 1/2. The average number of pinholes in the total mask is $N^2/2$. While the correlation of this mask is peaked at the origin, its background is nonzero. For the correlation

*For simplicity in notation we assume these functions are defined over a normalized sampling grid; hence, m and n are integers.

peak normalized to 1, the background has a mean of 1/2 and variance of $1/N^2$. The basic Fresnel zone plate correlation has similar artifacts in its background.

The use of Dicke's decoding scheme when imaging with either the random mask or Fresnel zone plate results in objectionable distortion in the decoded image. The degree of distortion is directly related to the spatial extent of the object fields. There are two principal sources of this distortion. The first of these is deterministic and an artifact of the decoding. This is the large background or DC level buildup associated with the correlation function of h(m,n). Consider the imaging of an M-by-M array of point sources. In the above decoding, each point source contributes a background of 1/2 for a total background of $M^2/2$. Thus, the ratio of the decoded image at a point source location to its mean background is

$$\frac{M^2/2 + 1}{M^2/2} \quad .$$

Clearly, for even modest sized M, this background completely dominates the correlation peaks. Unfortunately this distortion cannot be satisfactorily removed from the decoded image by simply subtracting its sample mean, due to the presence of background noise. We will demonstrate below that this can be avoided with proper decoding, however, and is therefore not a fundamental limitation.

The second source of distortion is due to the nature of the discrete, random emissions of x-ray and γ-ray sources. When the counting statistics are such that a limiting number of photons are collected during image formation, the dominant noise source in the measured image is a quantum noise. Assuming a stochastic source model with Poisson statistics (this model will be elaborated below), the noise field is uncorrelated with the source distribution, however, its mean at any point is proportional to the intensity of the image field at that point. The implication here is that as the number of pinholes increases, the noise will increase, thus offsetting to some extent the multiplex advantage of multiple pinholes. This quantum noise is a fundamental limitation of aperture encoding when imaging extended sources.

III. Optimal Linear Mean-Square Error Decoding

In this section we derive a decoding algorithm in which the deterministic degradations involved with Dicke's correlation decoding are eliminated. In addition to providing an improved decoding of this imagery, availability of this algorithm also allows us to quantitatively describe the general limiting performance of multiple aperture cameras.

In formulating our decoder, we consider stochastic models for the images. Both the source s(m,n) and the measured image a(m,n) are taken to be sample arrays from homogeneous random fields with 2-D autocovariances $K_s(m,n)$ and $K_a(m,n)$, respectively. Their crosscovariance is $C_{sa}(m,n)$. As in (2), we assume the form of a convolutional decoder. However, we take a different approach from Dicke in that we design the decoder to obtain (in some useful sense) an optimal estimate of the source s from the decoding. In this articulation, our problem becomes analogous to that of image restoration and we make use of some available restoration formalisms.[6,7]

More formally, the estimate of each member in the source field is constructed according to

$$\hat{s}(m,n) = \sum_{m'} \sum_{n'} a(m',n')\, g_0(m-m', n-n') \qquad (3)$$

where $g_0(m,n)$ is the point spread response of the optimal decoder. This spread response is chosen so as to minimize the mean-square error

$$\varepsilon = E\{[s(m,n) - \hat{s}(m,n)]^2\}, \quad \text{for all } m,n \qquad (4)$$

with $\hat{s}(m,n)$ as given in (3) and the expectation taken over the ensemble of possible x-ray source and noise arrays. Given the statistical models, this error is a natural measure of the optimality of the estimates.

The decoder which minimizes ε in (4) is given by the inverse Fourier transform of

$$G_0(u,v) = \frac{\Phi_{sa}(u,v)}{\Phi_a(u,v)} \quad , \qquad (5)$$

where we have the following 2-D Fourier transform pairs,

$$\begin{aligned} K_a(m,n) &\longleftrightarrow \Phi_a(u,v) \\ C_{sa}(m,n) &\longleftrightarrow \Phi_{sa}(u,v) \\ g_0(m,n) &\longleftrightarrow G_0(u,v) \end{aligned} \qquad (6)$$

Assuming the source and noise processes are uncorrelated, (5) becomes

$$G_0(u,v) = \frac{\overline{H}(u,v)\, \Phi_s(u,v)}{|H(u,v)|^2\, \Phi_s(u,v) + \Phi_w(u,v)} \qquad (7)$$

where, using a notation similar to that above, we have the transform pairs,

$$\begin{aligned} h(m,n) &\longleftrightarrow H(u,v) \\ K_s(m,n) &\longleftrightarrow \Phi_s(u,v) \\ K_w(m,n) &\longleftrightarrow \Phi_w(u,v) \end{aligned} \qquad (8)$$

and $\overline{H}$ = complex conjugate of H.

Thus, our decoder is a classical 2-D Wiener filter. The mean-square error associated with this decoding[8] is given by

$$\varepsilon = \int_{-1/2}^{+1/2} \int \frac{\Phi_s(u,v)\, \Phi_w(u,v)}{|H(u,v)|^2\, \Phi_s(u,v) + \Phi_w(u,v)}\, du\, dv \quad . \qquad (9)$$

The relations in (7) and (9) are the principal results of this section. They specify the optimal linear decoding and quantitatively describe its performance. In addition to having desirable formal properties, these results are also intuitively pleasing. For example, in the absence of noise (i.e., $\Phi_w(u,v) = 0$), we have $G_0(u,v) = 1/H(u,v)$. This decoder is simply an inverse filter and its application to y(m,n) completely removes the effects of the aperture encoding. This is not the case with the correlation decoding scheme. When noise is present, the filter in (7) achieves an optimal balance between the unremoved remnant of the encoding and the distortion due to the noise sources. This is, of course, the case of practical interest and we will now consider the question of just how well this decoding can be performed in the presence of photon-limited noise. This photon limited case would be expected to occur in x-ray astronomy as well as medical imaging.

A. Photon-Limited Imaging

We construct a compound Poisson model for the imaging. As before, we model the coded image y(m,n) as a realization of a homogeneous random field with covariance $K_y(m,n)$ and mean μ'. Here, for each specific realization we take the value at any point, say y(m',n'), to be the mean event rate of a Poisson source. The available or measured image a(m',n') is the count of this source. Accordingly, the error or noise process we consider is simply their difference

$$w(m',n') \triangleq a(m',n') - y(m',n') \quad . \tag{10}$$

With Poisson statistics, the number of events generated by the separate sources are statistically independent. Using this property and (10), we obtain the following relations which we state without proof:

$$K_a(m,n) = K_y(m,n) + K_w(m,n) \tag{11}$$

$$C_{yw}(m,n) = 0 \tag{12}$$

$$K_w(m,n) = \mu' \, \delta(m,n) \tag{13}$$

where $\delta(m,n)$ is the discrete delta function. Note that for h(m,n) a zero-one function, we also have

$$\mu' = \mu ||h||_1 = \mu ||h||_2^2 \tag{14}$$

where μ is the mean of the input source s(m,n) and

$$||h||_1 \triangleq \Sigma |h(k,\ell)| \;\&\; ||h||_2 \triangleq \sqrt{\Sigma\, h^2(k,\ell)} \quad .$$

Using these relations, we can express the mean-square error associated with this decoding as

$$\varepsilon = \int_{-1/2}^{+1/2}\!\!\int \frac{\mu\, \Phi_s(u,v)}{\mu + \dfrac{|H(u,v)|^2}{||h||}\, \Phi_s(u,v)}\, du\,dv \tag{15}$$

Where we have averaged over the ensemble of sources. The specific coding function in (15) is a member of the ensemble of Dicke's random masks. With high probability, $||h||_1 = N^2/2$, and by averaging over this ensemble, we can write

$$E\{\varepsilon\} = \int_{-1/2}^{+1/2}\!\!\int \left\{ E \frac{\mu\, \Phi_s(u,v)}{\mu + \dfrac{2}{N^2}\, |H(u,v)|^2\, \Phi_s(u,v)} \right\} d\Phi\, dv \quad . \tag{16}$$

This expectation can be evaluated by noting that $|H(u,v)|^2$ is approximately Chi-square distributed with two degrees of freedom.

Thus, having a known source density, $\Phi_s(u,v)$, we can quantitatively determine the performance of our decoder. In order to obtain a qualitative estimate of the behavior of this error let us consider the case of $\Phi_s(u,v) = \sigma^2$, i.e., constant. This is the spectrum of a white source field which is a limiting case of equal energy at all frequencies. For this field the decoding error (16) becomes

$$E\{\varepsilon\} = \frac{\mu N^2}{2}\, E\left\{ \frac{1}{\left(\dfrac{\mu N^2}{2\sigma^2} + |H(u,v)|^2\right)} \right\} \quad . \tag{17}$$

Now, most images of practical interest will be scaled such that the mean is on the order of the standard deviation. Thus $\mu << \sigma^2$, and a useful approximation of (17) results in

$$E\{\varepsilon\} = \mu \ell n\left[\frac{\sigma^2}{\mu}\right] \quad . \tag{18}$$

It is of interest to compare this to the mean-square error resulting from no coding-decoding, that is, for a single pinhole:

$$\varepsilon_{UC} = \mu \quad . \tag{19}$$

Evidently, employing aperture encoding in this case involves an average loss of performance. This reasoning applies to the case of infinite extent, i.e., to homogeneous random fields. We next consider the case where the source field is confined to a sub-region to the object plane.

B. Signals of Limited Support

Let the source field be bounded by a smallest rectangle here taken to be a square of side M for simplicity. Now place this signal s in an N x N square, "the object plane", by adding random shift modulo - N x N. Finally, create a homogeneous random field over all space by periodic repetition of this N x N cycle. We are now in a position to apply our previous results. We can speak of averages over all space or what is the same the N x N square. Also we can talk of averages over the M x M support of s. Since photon noise has the same support as the signal we have

$$\varepsilon_{UC}^N = \left(\frac{M}{N}\right)^2 \varepsilon_{UC}^M \tag{20}$$

where $\varepsilon_{UC}^K \equiv$ error in the uncoded case ensemble averaged over the K x K square. We note that ε^N is approximately the homogeneous error ε mentioned above in the infinite support case, when M is large compared to the correlation distances involved.

For the coded case, represented by the subscript "c", we have approximately, for N >> M,

$$E\{\varepsilon_c^M\} = E\{\varepsilon_c^N\} \tag{21}$$

because the random mask with high probability creates a nearly homogeneous noise field in the detector plane. Upon decoding this homogeneity is preserved by the shift-invariant decoder. Combining (18-21) we have finally

$$E\{\varepsilon_c^M\} \approx \left(\frac{M}{N}\right)^2 \ell n\left(\frac{\sigma^2}{\mu}\right) \varepsilon_{uc} \quad . \tag{22}$$

Thus the coding gain is expected to be

$$G = (N/M)^2 / \ell n(\sigma^2/\mu) >> 1. \quad (N >> M) \tag{23}$$

From this average value, we can conclude that at least half of the random masks will achieve half of the coding gain and there exists at least one mask which will achieve it all, for otherwise the average could not be correct. This is analogous to the random coding arguments of Information Theory.

It can also be shown that no linear decoder can do better than

$$\epsilon_0^M = \left(\frac{M}{N}\right)^2 \frac{\mu\sigma^2}{\mu+\sigma^2} = \left(\frac{M}{N}\right)^2 \epsilon_{UC} \quad , \qquad (24)$$

thus the factor $\ln(\sigma^2/\mu)$ describes the non-optimality of choosing a mask at random. For typical images this factor is on the order of 2, thus the random masks are seen to be fairly efficient. However, it turns out there is a very simple way of actually achieving (24). Simply repeat the image, i.e., multiplex it in space $(M/N)^2$ times. Then stack up the results.

However, if one could not anticipate the shape of the object and the value of M ahead of time, some type of adaptive procedure would be necessary to fit the $(M/N)^2$ pictures into the M x M square. Also, for M xM sources that display a marked non-homogeneity, it could be expected that the bright regions of the source would have relatively less error and the dark regions of the source would have relatively more error than the non-overlapping multiplex scheme.

To summarize very succinctly, it happens that the best coded apertures spread out the photon noise so that averages over the N x N object plane are equal with or without coding, and the random mask is off by a factor $\ln(\sigma^2/\mu) \approx 2$ (typically). Future work should center on finding the good codes for both white and non-white source spectra. Minimization of (15) subject to the 0-1 constraint on the range of h, could be approached as a non-linear optimization problem.[9]

IV. Experimental Results

Experimental results were obtained in the optical region using a small (ten grid units high) letter E. The noises here are film fog and grain noise. The grain noise being signal dependent as is the case with photon noise. Figure 1 shows an image of the 100 x 100 random mask. Figure 2 shows the coded image of the letter E. The optical matched filter decoded E was barely visible due to background buildup and hence is not shown. Figure 3 shows a computer matched filtering with means removed. Figure 4 shows an "optimal" estimate where the design SNR was chosen to maintain image resolution. Figure 4 is much sharper than Figure 3 and does not have the horizontal and vertical artifacts of Figure 3. Finally Figure 5 presents the decoded image of a two level E where the bottom half is 10 times intensity of the upper half.

V. Conclusions

We have considered coded aperture imaging with randomly chosen masks. We have shown that the previous decoding methods suffered from deterministic (for a given mask) errors that lead to buildup of background and high frequency errors in the decoded image. System theory was applied via image enhancement to the aperture coding problem. For the photon limited case the optimal linear decoder was derived. We found that the error with the random mask had the correct dependence on (M/N) in the finite support case.

Acknowledgement

This work was performed under the auspices of the U.S. Atomic Energy Commission.

References

[1] H. H. Dicke, "Scatter-Hole Cameras for X-Rays and Gamma-Rays," Astrophysical J., vol. 153, p. 1101, August 1968.

[2] H. H. Barrett, "Fresnel Zone Plate Imaging in Nuclear Medicine," J. Nuc. Med., vol 13, p. 382, July 1972.

[3] B. MacDonald et al, "Gamma-Ray Imaging Using a Fresnal Zone Plate Aperture, Multiwire Proportional Chamber Detector, and Computer Reconstruction," IEEE Trans. Nucl. Sci., vol. 21, p. 678, February 1974.

[4] G. Groh, et al, "X-Ray and Gamma-Ray Imaging with Multiple-Pinhole Cameras Using a Posteriori Image Synthesis," Applied Optics, vol. 11, p. 931, April 1972.

[5] A. Z. Abcasu, et al, "Coded Aperture Gamma-Ray Imaging with Stochastic Apertures," Optical Engr., vol. 13, p. 117, March/April 1974.

[6] C. W. Helstrom, "Image Restoration by the Method of Least-Squares," J. Opt. Soc. Amer., vol. 57, p. 297, March 1967.

[7] M. P. Ekstrom and O. L. Rater, "Computer Processing of Radiographic Images with the Finite Fourier Transform," IEEE Trans. Nucl. Sci., vol. NS-19, p. 40, February 1972.

[8] H. L. Van Trees, Detection, Estimation and Modulation Theory, I, Wiley, New York, 1968.

[9] D. G. Luenberger, Optimization by Vector Space Methods, Wiley, 1969.

Figures

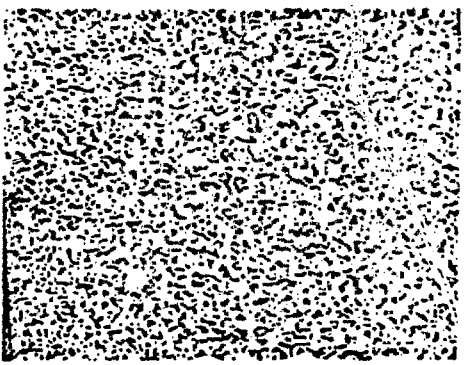

Figure 1 Random Mask

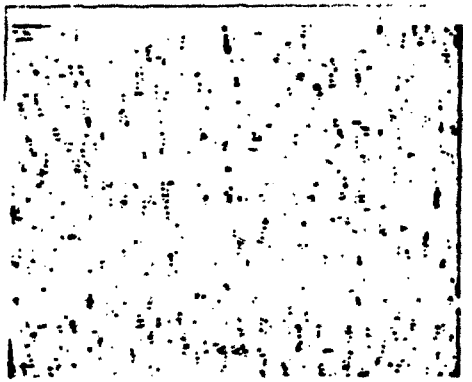

Figure 2 Coded Image of E

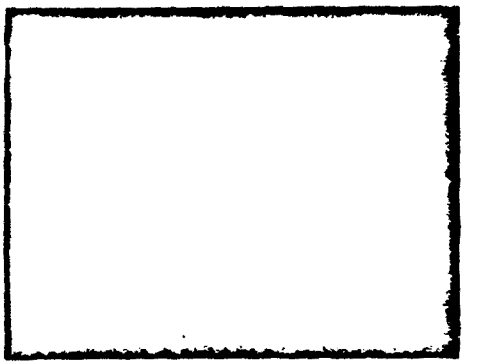

Figure 3 Matched E - Means Removed

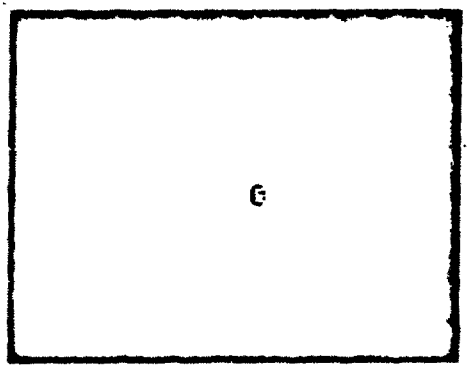

Figure 4 "Optimal" Estimate of E

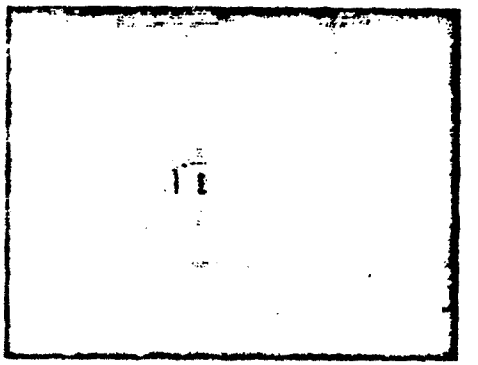

Figure 5 Estimate of Two-Level $\bar{E}$