Radionuclide Imaging with Coded Apertures and Three-Dimensional Image Reconstruction from Focal-Plane Tomography

Lee-Tzuu Chang

ABSTRACT

Two techniques for radionuclide imaging and reconstruction have been studied; both are used for improvement of depth resolution.

The first technique is called coded aperture imaging, which is a technique of tomographic imaging. The image of an object is first recorded in a coded form by a gamma camera with a special aperture containing a coded hole pattern or opening. A decoding process is needed to reconstruct the image from its coded form. The investigations of this technique are concentrated in theory and a few prospective coding-decoding methods. The results show that a non-redundant pinhole coded aperture can provide us a simple and reliable way to obtain the images of a 3-D object with adequate depth resolutions and low background artifacts.

The second technique is a special 3-D image reconstruction method which is introduced here as an improvement to the so called focal-plane tomography. This method can be applied to a number of tomographic imaging devices, such as the positron camera, the rotating slanted-hole collimator camera (tomocamera), the multiplane tomographic scanner (tomoscanner), and the multiple pinhole camera. Applications of this reconstruction method to the last tomographic imaging device have shown,

through computer simulations and experiments with radioactive phantoms, that within the limit of the depth resolution of the tomographic imaging device, structures from activities above and below the in-focus planes can be removed mostly from the focal-plane tomographic images.

RADIONUCLIDE IMAGING WITH CODED APERTURES AND THREE-DIMENSIONAL IMAGE RECONSTRUCTION FROM FOCAL-PLANE TOMOGRAPHY

CONTENTS

PART II : THREE-DIMENSIONAL IMAGE RECONSTRUCTION FROM FOCAL-PLANE TOMOGRAPHY

0. INTRODUCTION TO RADIONUCLIDE IMAGING

0.1 Review of Radionuclide Imaging in Nuclear Medicine

Nuclear medicine is the application of radioactive materials to the diagnosis and treatment of patients and the study of human disease. (Refs. 1-4) From its early beginings when artificially produced radionuclides first became available[5], nuclear medicine has developed into one of the most exciting and important components of modern biomedical science. Nuclear methods have found application in both the diagnosis and treatment of disease. With the passage of time, however, diagnostic techniques are receiving an increasingly large share of the attention of clinical investigators.

In diagnostic nuclear medicine, examinations of tumors, organs, or other anatomical structures are carried out by detection of the electromagnetic radiation emitted by radioactive substances concentrated in these structures. These electromagnetic radiations may be either emitted by the radionuclide(gamma rays), or they may result from a process secondary to radioactive decay(x rays and annihilation radiation). (For the sake of simplicity, the different electromagnetic radiations will, in this thesis, be called gamma rays). The instruments used for this purpose can be broadly categorized into (1) scanners and (2) gamma cameras.

Scanners utilize radiation detectors collimated to accept photons from a single resolution element of the structure to be

visualized, and the image of the structure is formed by a scanning motion of the detector over the area of interest. Gamma cameras, on the other hand, are stationary devices; their radiation detector continuously observes the whole field to be visualized. Only a discussion of gamma cameras will be given in this section for the purpose of this thesis, descriptions of other instruments for radionuclide imaging should be refered to Ref. 6.

Radionuclide cameras are relatively new instruments for imaging the distribution of radiocompounds inside the body. They have many advantages over scanners.[7] They have the ability to "see" some organs in their entirety and to display visually changes in an organ during a dynamic function test. Most of them are constructed so the detector head can be positioned to do organ imaging either in the recumbent or the sitting position. The construction of a gamma camera consists of three components: (1) a device capable of channeling the gamma-ray photons with the purpose of forming an image, (2) a detector sensitive to the gamma-ray photons, and (3) a system capable of establishing a spatial relationship between the signal supplied by the detector and the position of the gamma-ray photons in the image formed by (1).

Since the index of refraction of optical lenses is very close to one for electromagnetic radiation with wavelengths in gamma-ray region, a large lens aperture provides no means of collimating gamma-ray photons for the image forming. A number of different methods of collimation have been used with the purpose of providing

the detector of a gamma camera with an image formed by gamma-ray photons. Conventional collimating devices are multichannel collimators and pinhole apertures. Other collimating devices such as slanted multichannel collimators,[8,9] converging (or diverging) multichannel collimators,[10] and coded apertures[11, 12] are relatively new designs and their special uses will be explained in Section 0.2.

Multichannel (or parallel-hole) collimators are the most widely used image forming devices in gamma cameras. They consist of a series of holes or channels with parallel axes in a plate made of a dense and heavy material such as lead. The image-forming mechanism of a multichannel collimator is shown in Fig.1-1. The gamma-ray photons emitted from the source pass through the parallel channels of the collimator and form an image of the source distribution with its size and orientation unchanged. The best image resolution is obtained by means of a multichannel collimator in contact with the object observed, and this resolution decreases with distance.

The principle of operation of pinhole apertures for gamma rays is identical to that of the pinhole optical camera. A pinhole aperture consists of a hole that is provided in a plate opaque to gamma ray photons, originating at the object to be visualized and traveling towards the hole as shown in Fig.1-2, form an inverted image of the object at the detector. The size of the image is determined by the object-to-aperture and the aperture-to-detector distances. Thus, a pinhole aperture has the ability to magnify images of small objects for a better image resolution.

There are many position sensitive devices which can be used as detectors in gamma cameras. At least four different types of gamma cameras can be identified according to the detectors used. The Anger-type scintillation cameras[6,7] are the most widely used gamma cameras. Commercially available Anger Cameras use a 1/2-inch thick NaI(Tl) crystal that varies in diameter from 12 to 13.5 inches, depending on the manufacturer, and use 19 or 37 multiplier phototubes arranged in a hexagonal array. Other experimental types of gamma cameras that have been used are image intensifier cameras[3], spark chambers[14] and multi-wire proportional chambers,[15] and semiconductor imaging systems[16]. General descriptions of these gamma camera systems are given in Ref. 6 and 7.

0.2 Tomography and Three-Dimensional Image Reconstruction

An important limitation of the conventional gamma cameras is that gamma cameras with a pinhole aperture or a parallel-hole collimator can not "see" the depth of an object distribution. A conventional gamma camera photo is only a two-dimensional portrayal of the three-dimensional distribution of the object. For the purpose of diagnosis for abnormality in small lesions, it is highly desirable to have imaging devices which can exclude the unwanted activity above and below the plane of interest. During the last 50 years, radiology has developed several methods to exclude unwanted information and improve the radiographic image. These methods are called tomography from the Greek words meaning "to write a cut".

Conventional radiographic tomography enhances the visibility of radiographic structures on or near a chosen plane by blurring all objects off the chosen plane, thereby reducing the confusing effects of overlying and underlying structures. In radionuclide imaging, the same technique is used with more flexibility because of the availability of various radionuclides for radiopharmaceuticals and advanced instruments for imaging. To avoid confusion with the old terms used in radiology, a specific term was suggested by Anger for the radionuclide imaging methods that result in sharp images of radioactive objects located on a chosen plane and blurred images of objects located above or below that plane; it is called focal-plane tomography. The earliest apparatus used for this purpose are rectilinear scanners with wide-angle focused collimators,[17-19] and positron cameras.[20,21] During the last decade many new instruments and techniques have been introduced to perform focal-plane tomography, some of them have been actually in use for quite some time while others are still under study. Examples are the multiplane tomographic scanner (tomoscanner),[22] the rotating slanted-hole collimator camera (tomocamera),[8,9] the magnifying collimator camera,[10] and coded aperture imaging.[11,12] A summary of all these tomographic imaging devices and techniques is given in Ref.23.

Coded aperture imaging was recently introduced in the field of nuclear medicine imaging first by Barrett[11] and later by Rogers[12] as a high resolution tomographic imaging technique. The special feature of this imaging technique is that it is a two-step process.

The image first obtained from a coded aperture imaging system is not in a recognizable form. It must be decoded before the true image can be visualized. The depth information of the object distribution is stored in the coded image and can be brought back through the decoding process. Although in some cases coded aperture imaging provides us net gain in signal collection efficiency, the main advantage of coded aperture imaging is its ability to get tomographic images without any movement of the object or the camera (including the collimating device). Brief descriptions of this imaging technique will be given in Chapter 1, and an intensive analysis is the subject of Part I of this thesis.

Although tomography provides us better image contrast and depth resolution than the conventional methods, it does not eliminate the events from off-plane activities, and such events can still reduce the contrast and the visibility of the in-focus images even through these events are scattered. Since the objects for nuclear medicine imaging are often of low contrast, it is often not possible to identify the true in-focus images from the out-of-focus background images. Removal of the background images would enable detection of smaller lesions and of lesions of lower contrast. It is possible to remove the unwanted background images from the focal-plane tomographic images provided that the direction of each detected gamma-ray event is known. A method of 3-D image reconstruction, which will be given in Part II of this thesis, is for this purpose.

In parallel to the rapid growth of radionuclide imaging, trans-

mission imaging in radiology has also achieved great improvement. A 3-D radiographic imaging technique, called tomosynthesis, has been studied by a number of investigators.[37,38] Instead of a continuous motion (in conventional tomography) a discrete sample procedure is performed in tomosynthesis, i.e. N discrete radiographs are recorded from N different positions of the x-ray source. Tomographic images of an object at various depths are constructed from its sample radiographs. This imaging technique has also been used in radionuclide imaging. A gamma camera with a multiple pinhole array aperture (or a single-pinhole aperture at N different positions) is used in this case. Detailed descriptions of the method of tomosynthesis in radionuclide imaging is given in Chapter 5. Both conventional tomography and tomosynthesis are not capable of imaging fast moving objects because of the duration of the sequential imaging process. To overcome this difficulty, Weiss has suggested a method of coded sources for imaging of objects like the beating heart or fast flowing contrast media injected in the brain, kidneys or other organs. This method uses a single radiograph (coded image) generated by an array of x-ray sources to construct the tomographic images of the object.[39] The principle of operation of a coded source imaging system is exactly the same as that of a coded aperture imaging system.

Aside from tomographic imaging in x-ray transmission study, a number of improved image reconstruction techniques, such as the algebraic reconstruction techniques,[24] the iterative reconstruction techniques,[25] the Fourier reconstruction technique,[26] and the technique

of linear superposition with compensation,[27] have shown great promise for 3-D image reconstruction. All of these techniques have the advantage of being section imaging methods in which a layer of the object is imaged from its sides so that neither the activity nor the radiation of the overlying and underlying structures can influence the image reconstruction of a chosen plane section. Image reconstructions are actually performed in the x-y plane. Budinger has adopted some of these techniques in radionuclide imaging.[28] It is evident that studies of various 3-D imaging and reconstruction techniques as well as improvements of tomographic imaging techniques with better depth resolution and image contrast and minimum background artifacts will be the future trends for both radiographic imaging and radionuclide imaging.

PART I RADIONUCLIDE IMAGING WITH CODED APERTURES

1. INTRODUCTION TO CODED APERTURE IMAGING

1.1 Coded Aperture Imaging in Nuclear Medicine

It was explained in Section 0.1 that a gamma camera requires a device capable of channeling the gamma-ray photons with the purpose of forming an image. A pinhole aperture or a multichannel collimator can provide the detector of the gamma camera a direct mapping of the object distribution. Aside from the size and orientation the image impinging on the detector is an exact 2-D portrayal of the object distribution. Thus, a pinhole aperture or a multichannel collimator gives us a simple and direct way to obtain images of objects in the x-y plane. In order to obtain some information on the object distribution in the z direction one has to take several views of the object at different angles. This means longer exposure time is required.

One way which can enable us to get several views of the object at different angles without increasing the exposure time is to add more pinholes to a single-pinhole aperture plate of a gamma camera. In this case, however, all of the pinhole views impinging on the detector overlap one another and are spread over a large area of the detector. The image of the object is no longer recognizable and a means of image reconstruction is required. This image is said to be coded by the pinhole pattern of the aperture plate. We shall call it a coded image

or a shadowgram.

Coded aperture imaging is an indirect method of imaging in which the conventional pinhole aperture or multichannel collimator is replaced by a plate containing a coded hole pattern or opening, as shown in Fig.1-3. A point gamma-ray source will cast a shadow of this hole pattern on the detector and this shadow is the coded image of the point source. A source distribution is encoded as a superposition of many such patterns. The hole pattern is chosen to facilitate the image reconstruction or decoding process and to ensure a faithful reproduction of the source distribution. The decoding may be accomplished in a number of ways both optically and electronically.

Any aperture plate with opening other than a single pinhole may be called a coded aperture plate, and there exist many different code patterns. The multiple-pinhole-array(MPA) coded apertures and the Fresnel-zone-plate(FZP) coded apertures are the best known ones. A multiple-pinhole-array aperture is simply an aperture consisting of a number of pinholes. The random pinhole array shown in Fig.1-4a was first suggested by Dicke[29] for x-ray astronomy and has been used for solar x-radiation.[30] This type of MPA consists of a large number of pinholes randomly distributed having an average transmission of 50%.

A second type of MPA is called the non-redundant pinhole array and an example of this type is shown in Fig.1-4b. The non-redundant pinhole array was originally proposed for radar antenna.[31-33]

An optical simulation of the use of this type of MPA was shown by Wouters, et al.[34] The positions of the pinholes of a non-redundant pinhole array are carefully chosen so that the vector distance between any two pinholes occurs only once, i.e., the pinholes are spaced non-redundantly. This non-redundancy means that the autocorrelation function of this pinhole array is sharply peaked at the center and uniformly distributed elsewhere.

The reconstruction of a MPA coded aperture system can be made by correlating the coded images with the original coded pinhole aperture array. It is done either with a simple optical system or with a small digital computer. The resulting images from this two-step system have a finite depth of focus, that is, a sharp reconstructed image can be seen at the appropriate image plane and blurred images are seen at image planes some distance away from that plane. Therefore, 3-D information of the object distribution is obtained through this system. There is also gain in geometrical efficiency due to the increased number of pinholes over the single-pinhole aperture system. But as the object to be imaged becomes larger the signal-to-noise ratio of coded aperture imaging decreases and the gain in geometrical efficiency is lost (this will be explained later).

Perhaps the first use of a coded aperture was by Mertz and Young who used a Fresnel zone plate to image x-ray stars.[35] Decoding was done by using coherent light from a laser. Barrett later applied this method to Nuclear Medicine imaging.[36] The FZP coded aperture systems are somewhat more complicated than the MPA coded aperture

systems because of the additional focusing property of Fresnel zone plates. A Fresnel zone plate consists of a number of concentric zones and the radius of the n^{th} zone is given by $r_n = \sqrt{n}\ r_1$, where r_1 is the radius of the first zone. The transmission of each zone is either 0 or 1 as shown in Fig.1-4c. It is known that a Fresnel zone plate behaves somewhat like a lens. When the zone plate is illuminated with a plane wave of monochromatic light, part of the light is diffracted into a converging spherical wave so that it forms a bright focal spot according to the formula $f = r_1^2/\lambda$, where f is the focal length and λ is the wave length of the incoming coherent light. Detailed analysis of the diffraction properties of zone plates will be given in Sections 2.2 and 3.3.

The coded image or zone plate shadowgram is obtained by putting a zone plate mask somewhere between the radioactive object and the detector. For two point sources located at different distances from the zone plate mask, the shadowgram so produced consists of two zone plate patterns of different sizes. Upon passing a beam of coherent light through this zone plate shadowgram, the large zone plate pattern in the shadowgram focuses the light beam to a point further away from the shadowgram than that from the small zone plate pattern because the focal length of a zone plate pattern is proportional to the area of the first zone(r_1^2). Therefore, by coding the object distribution into a zone plate shadowgram, the 3-D information of the object distribution is stored.

1.2 Advantages and Limitations

The coded aperture systems have a number of advantages over the conventional imaging systems. The most important one is that the coded aperture systems can offer us three-dimensional information about the source distribution. Other advantages are :

(1) Since the coding-decoding process has very little correlation with the background and scattered radiation, the resulting image is less sensitive to it. For the same reason, other noise such as electronic noise and film noise (if there is any) can cause only slight problem to the coded aperture systems as compared to the conventional imaging systems.

(2) The effect of nonuniformity of the detector systems is less dominant when the coded aperture imaging technique is used.

(3) For noncompact source distributions such as the skeleton, and for source distributions having a small number of picture elements such as the thyroid glands, the kidneys, and the pancreas, the signal-to-noise ratio is higher when we use the coded aperture systems because of the net gain of geometrical efficiency over the single pinhole camera systems.[53,54,55]

(4) Coded apertures offer us considerable flexibility in terms of field of view and image magnification as compared to the parallel-hole collimators.

There are also limitations and drawbacks to the coded aperture systems. These are :

(1) The coded aperture systems require larger detectors in order to cover its field of view which is larger than that of a conventional imaging system.

(2) The signal-to-noise ratio of a coded aperture system decreases as the size of the object to be imaged increases.

(3) The coded aperture systems require two-step processing which is time consuming when on-line computer processing is not available (in that case, an optical reconstruction method is used).

2. ANALYSIS FOR CODED APERTURE IMAGING SYSTEMS

2.1 Coding and Decoding

As mentioned earlier coded aperture imaging is a two-step process, coding and decoding. Therefore an analysis of some general aspects of the coding-decoding process would enable us to get a better understanding of the prospective and the limiting features of coded aperture imaging in nuclear medicine.

The analysis of a coded aperture imaging system is similar to that of an electrical communication system. Both the imaging system and the communication system are designed to collect or convey information. Although in the latter case the information is of a temporal nature (i.e., a modulated voltage or current waveform) while the former case it is of a spatial nature (i.e., a distribution over space), there is actually no substantial difference between them from an abstract point of view. In the analysis of communication systems the electronic networks which are studied often possess two important properties called linearity and invariance. Any system which is linear and (time- or space-) invariant can be described mathematically with considerable ease. A system is said to be linear if the following superposition property is obeyed for all input functions $f(x,y)$ and $g(x,y)$ and all complex constants a and b :

$$\Psi\left\{a\, f(x,y) + b\, g(x,y)\right\} = a\,\Psi\left\{f(x,y)\right\} + b\,\Psi\left\{g(x,y)\right\},$$

where the mathematical operator $\Psi\{\ \}$ represents the system, which we imagine to operate on input functions to produce output functions. We let the symbol $h(x,y;\eta,\xi)$ denote the response of the system at point (x,y) of the output space to a δ-function input at coordinates (η,ξ) of the input space: that is,

$$h(x,y;\eta,\xi) = \Psi\{\delta(x-\eta,y-\xi)\}$$

The function $h(x,y;\eta,\xi)$ is called the impulse response of the system. A linear imaging system is said to be space-invariant if its impulse response $h(x,y;\eta,\xi)$ depends only on the distances $(x-\eta)$ and $(y-\xi)$. For such a system we can write $h(x,y;\eta,\xi) = h(x-\eta,y-\xi)$ and $h(x,y;0,0) = h(x,y)$. Thus an imaging system is space-invariant if the image of a point-source object changes only in location, not in functional form, as the point source moves from one place (η_1,ξ_1) to another (η_2,ξ_2). A coded aperture system is approximately linear and space-invariant within its field of view if the attenuation of gamma rays between the source and the detector (which is assumed to be linear) does not change much from one straight-line path to another. For the purpose of analysis, we consider all coded aperture systems to be linear and space-invariant and assume the attenuation of gamma rays between the source and the detector is negligible.

In the next two sections studies of the coding-decoding process will be given in both the spatial domain and the frequency domain. Terms such as convolution, correlation, and Fourier transform will be used very often without giving any descriptions of their mathematical properties and physical significance. A few textbooks in

optics and communication theory are listed in Refs. 49,51,58,and 59 for reference.

Although on-line decoding is possible in coded aperture imaging, the decoding of coded images is mostly performed off-line. Therefore a coded aperture system is considered as two subsystems. We define the imaging or coding system as the gamma camera system which makes the shadowgrams and define the reconstruction system as the system which performs the decoding of the shadowgrams. The shadowgram of a point source is the impulse or point response of the imaging system. This impulse response is a binary function (i.e., the function has only two possible values, 0 or 1) because the transmission of the coded aperture plate is binary. However, it is possible to construct an imaging system which gives an impulse response function other than a binary function (this will be explained in the next section). The impulse response of the reconstruction system, on the other hand, may be any type of function(real or complex).

2.2 Spatial Domain Analysis

We consider the imaging of a 2-D source distribution first. We let the distribution of the source object be given by $o(x,y)$ and the projection of the coded aperture pattern on the detector from a point source at (0,0) in the object plane be given by $h(x,y)$. $h(x,y)$ is then the impulse (or point) response of the imaging system. The image of $o(x,y)$ on the detector formed by a hypothetic pinhole at

(0,0) in the aperture plane is given by $\frac{1}{\alpha^2} o(-\frac{1}{\alpha}x,-\frac{1}{\alpha}y)$, where $\alpha = S_o/S_1$ is the magnification factor of the imaging system resulting from the imaging geometry shown in Fig.1-3. Applying the principle of superposition to this linear and space-invariant system, the coded image or shadowgram of the object $o(x,y)$ on the detector plane is given by the convolution integral of the image $\frac{1}{\alpha^2} o(-\frac{1}{\alpha}x,-\frac{1}{\alpha}y)$ and the impulse response $h(x,y)$ as

$$s(x,y) = \frac{1}{\alpha^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} o(-\frac{1}{\alpha}(\xi,\zeta))h(x,y;\xi,\zeta)d\xi d\zeta = \frac{1}{\alpha^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} o(-\frac{1}{\alpha}(\xi,\zeta))h(x-\xi,y-\zeta)d\xi d\zeta$$

Adopting the commonly used short-hand notation for a convolution integral, the shadowgram $s(x,y)$ is expressed as, with $\underline{r} = (x,y)$,

$$s(\underline{r}) = \frac{1}{\alpha^2} o(-\frac{1}{\alpha}\underline{r}) \circledast h(\underline{r}) \qquad (2.1)$$

To get the image of $o(\underline{r})$ we have to perform a decoding operation on the shadowgram $s(\underline{r})$. This decoding operation can also be expressed in a form of convolution integral because the reconstruction system is also linear and space-invariant(this will become apparent later). If $h'(\underline{r})$ is the impulse response of the reconstruction system, the output image $i(\underline{r})$ from this system to an input $\alpha^2 s(-\alpha\underline{r})$ is then given by

$$i(\underline{r}) = \alpha^2 s(-\alpha\underline{r}) \circledast h'(\alpha\underline{r}) = o(\underline{r}) \circledast \left[h(-\alpha\underline{r}) \circledast h'(\alpha\underline{r})\right] \qquad (2.2)$$

Since only $o(\underline{r}) \circledast \delta(\underline{r}) = o(\underline{r})$, a perfect reproduction of the 2-D source distribution $o(\underline{r})$ can be obtained only if $h(-\alpha\underline{r}) \circledast h'(\alpha\underline{r}) = \delta(\alpha r)$.

A commonly used decoding technique for the reconstruction of MFA coded shadowgrams is the optical correlation method. This method takes $h(\underline{r})$, the impulse response of the imaging system, as $h'(\underline{r})$ of Eq.(2.2) so that the resulting image $i(\underline{r})$ becomes

$$i(\underline{r}) = o(\underline{r}) \circledast \left[h(-\alpha\underline{r}) \circledast h(\alpha\underline{r}) \right]$$

Since
$$h(-\underline{r}) \circledast h(\underline{r}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(-\eta, -\xi)\, h(x-\eta, y-\xi)\, d\eta\, d\xi$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\eta', \xi')\, h(x+\eta', y+\xi')\, d\eta'\, d\xi'$$

$$= h(\underline{r}) \circledast h(\underline{r}),$$

the image $i(\underline{r})$ can be expressed as

$$i(\underline{r}) = o(\underline{r}) \circledast \left[h(\alpha\underline{r}) \otimes h(\alpha\underline{r}) \right] = o(\underline{r}) \circledast a(\alpha\underline{r}), \qquad (2.3)$$

where $\otimes$ is used to represent the correlation of two functions and $a(\underline{r})$ is defined as the autocorrelation function of $h(\underline{r})$. We see that in this case the resulting image $i(\underline{r})$ is the convolution of $o(\underline{r})$ and $a(\alpha\underline{r})$ or, in other words, the image of $o(\underline{r})$ is blurred by the coding-decoding process with a point spread function $a(\alpha\underline{r})$. If $a(\underline{r})$ is a delta function, $\delta(\underline{r})$, then there is no spread of $o(\underline{r})$ and the resulting image $i(\underline{r})$ is a perfect reproduction of the original source distribution. However, this has never been the case for any known coded aperture systems. The best autocorrelation function we can have, from various MFA coded aperture functions, only approximately resemble a delta function, as shown in Fig.2-1b, which is from a non-redundant pinhole array function. A study of the autocorrelation

functions of three different types of pinhole arrays as well as the Fresnel zone plate will be given in Section 3.2. Experimental results from MPA coded aperture systems using the optical correlation method will also be given there.

The correlation method may also be applied to the reconstruction of FZP coded shadowgrams. However, the reconstructed images with this method are generally embedded in a background of high intensity due to the distribution of $a(\underline{r})$ of the Fresnel zone plate. If the correlation method will be used to decode the FZP coded shadowgrams a complex coding method can be applied to eliminate the background. This method produces an equivalent impulse response with a δ-function-like autocorrelation (an example is shown in Fig.2-1d). Descriptions of this method will be given later.

The commonly used decoding method for the reconstruction of FZP coded shadowgrams utilizes the diffraction properties of zone plate patterns. The transmission function of an infinitely large Fresnel zone plate can be expanded into a Fourier sine series plus a positive constant (= 1/2). If we use $h_{zp}(\underline{r})$ to denote this transmission function and let $x = r^2$ and $T = 2r_1^2$ = the period, the Fourier coefficients of $h_{zp}(\underline{r})$ are calculated as follows:

$$h_{zp}(x) = h_{zp}(\underline{r}) = \begin{cases} 1, & \text{if } r_{2n} < |\underline{r}| < r_{2n+1} \\ 0, & \text{if } r_{2n+1} < |\underline{r}| < r_{2n+2}, \end{cases}$$

where $r_n = \sqrt{n}\, r_1$ for $n = 0, 1, 2, \ldots$, and r_1 is the radius of the first zone of the zone plate. The Fourier series expansion of $h_{zp}(x)$ is

given by

$$h_{zp}(x) = a_o + a_1\cos(2\pi x/T) + \ldots + a_n\cos(2n\pi x/T) + \ldots + b_1\sin(2\pi x/T) + \ldots + b_n\sin(2n\pi x/T) + \ldots \,, \quad n = 1, 2, \ldots$$

$$a_o = \frac{1}{T}\int_0^T h_{zp}(x)\, dx = \frac{1}{T}\int_0^{T/2} dx = 1/2$$

$$a_n = \frac{2}{T}\int_0^T h_{zp}(x)\cos(2n\pi x/T)\, dx = \frac{1}{n\pi}\int_0^{n\pi} \cos(x')\, dx' = 0$$

$$b_n = \frac{2}{T}\int_0^T h_{zp}(x)\sin(2n\pi x/T)\, dx = \frac{1}{n\pi}\int_0^{n\pi} \sin(x')\, dx' = \begin{cases} \frac{2}{n\pi}, & n=1,3,\ldots \\ 0, & n=2,4,\ldots \end{cases}$$

In terms of r^2 and r_1^2, the zone plate transmission function is given by

$$h_{zp}(\underline{r}) = \frac{1}{2} + \frac{2}{\pi}\sin(\pi r^2/r_1^2) + \ldots + \frac{2}{n\pi}\sin(n\pi r^2/r_1^2) + \ldots \qquad (2.4)$$

$n = 1, 3, 5, \ldots$ = odd positive integers

or (in complex form)

$$h_{zp}(\underline{r}) = \frac{1}{2} + \frac{i}{\pi}e^{-i\pi r^2/r_1^2} - \frac{i}{\pi}e^{i\pi r^2/r_1^2} - \ldots - \frac{i}{n\pi}e^{in\pi r^2/r_1^2} - \ldots \qquad (2.5)$$

$n = \pm 1, \pm 3, \pm 5, \ldots$ = odd integers

Eq.(2.5) implies that, if a plane wave of light of wave length λ is incident on the zone plate, each exponential term of Eq.(2.5) will modulate the phase of the wave by an amount given by $n\pi r^2/r_1^2$ ($n = \pm 1, \pm 3, \pm 5, \ldots$). It is known that, after the incidence of this wave on the zone plate, part of the light is diffracted into both converging and diverging spherical waves which come into a number of focal spots according to the formula $f_n = r_1^2/n\lambda$ (n = odd integers), and the rest of the light is unaffected. The focal lengths f_n can be calculated from a simple equality as follows: In order to have a constructive interfer-

ence at (0,0,z) of wavelets emerging from point sources at (x,y,0) with amplitudes 1 and phase factors $n\pi r^2/r_1^2$ ($r^2 = x^2 + y^2$) the following equality must be satisfied.

$$\sqrt{z^2 + \frac{2r_1^2}{n}} - z = \lambda$$

Expanding the first term of the above equality and making a two-term approximation gives

$$z\,(1 + r_1^2/nz^2) - z \approx \lambda \qquad \text{or} \qquad z \approx r_1^2/n\lambda \ .$$

In the reconstruction of FZP coded shadowgrams only the term $e^{-i\pi r^2/r_1^2}$ in Eq. (2.5) is considered as the "signal carrier"; all the other terms in the complex expression of $h_{zp}(\underline{r})$ are unwanted. The reason for this can be explained easily by use of Eq.(2.2). When the diffraction property of a zone plate is utilized for the reconstruction of FZP coded shadowgrams the impulse response $h'(\underline{r})$ of the reconstruction system is given by the Fresnel wavefuction $-\frac{i}{\lambda z} e^{i\pi r^2/\lambda z}$ (this will be described in detail in Section 3.3). The output image $i(\underline{r})$ on the first focal plane of the zone plate, $z = r_1^2/\lambda$, is given by $o(\underline{r}) \circledast \left[h_{zp}(-\underline{r}) \circledast (-\frac{i}{r_1^2} e^{i\pi r^2/r_1^2}) \right]$. Only the component $e^{-i\pi r^2/r_1^2}$ of $h_{zp}(\underline{r})$ is in autocorrelation with $h'(\underline{r})$ and this autocorrelation produces a delta-function response which, upon convoluting with $o(\underline{r})$, results in a perfect reproduction of $o(\underline{r})$. (The proof of this property will be given in Section 3.3) All the other components of $h_{zp}(\underline{r})$ results in background and artifacts which are superimposed on the true image of $o(\underline{r})$. The largest contribution of the background and artifacts is, of course, due to the undiffracted light from the zero order

component of $h_{zp}(\underline{r})$, i.e. 1/2 in Eq.(2.5). It was proposed that this zero order component of $h_{zp}(\underline{r})$ can be effectively removed by a complex coding technique. The method is described briefly below; a detailed study will be given later. The Fresnel zone plate represented by $h_{zp}(\underline{r})$ is called a positive zone plate (center transparent). If the transmission of $h_{zp}(\underline{r})$ is reversed such that the new transmission function $\tilde{h}_{zp}(\underline{r})$ is 0 wherever $h_{zp}(\underline{r})$ is 1 and is 1 wherever $h_{zp}(\underline{r})$ is 0, then the zone plate represented by $h_{zp}(\underline{r})$ is called a negative zone plate (center opaque). The Fourier series expansion of $\tilde{h}_{zp}(\underline{r})$ is given by

$$\tilde{h}_{zp}(\underline{r}) = \frac{1}{2} - \frac{2}{\pi}\sin(\pi r^2/r_1^2) - \ldots - \frac{2}{n\pi}\sin(n\pi r^2/r_1^2) - \ldots \qquad n = 1, 3, 5, \ldots \tag{2.6}$$

If we form two shadowgrams $s_+(\underline{r})$ and $s_-(\underline{r})$ of the source object $o(\underline{r})$ with a fixed imaging geometry, s_+ from the positive zone plate and s_- from the negative zone plate, the difference of the two shadowgrams gives a new shadowgram which contains no contribution from the zero order component of $h_{zp}(\underline{r})$ because

$$s_c(\underline{r}) = s_+(\underline{r}) - s_-(\underline{r}) = o(\underline{r}) \circledast \left[h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r})\right] = o(\underline{r}) \circledast \left[2h_{zp}(\underline{r}) - 1\right].$$

$s_c(\underline{r})$ is equivalent to the shadowgram formed directly from imaging $o(\underline{r})$ with a coded aperture function $h_c(\underline{r}) = h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r})$. The impulse response of this imaging system is a composite one, $h_c(\underline{r})$, which has three possible values, 0, 1, and -1. The autocorrelation function of $h_c(\underline{r})$ is sharply peaked at $\underline{r} = \underline{0}$ and is very low elsewhere (an example is shown in Fig. 2-1d, which is generated by computer with a bounded $h_c(\underline{r})$, i.e., from zone plates of finite size). Both the auto-

correlation and the diffraction properties of $h_c(\underline{r})$ can be used to decode the shadowgram $s_c(\underline{r})$. It is difficult to construct $s_c(\underline{r})$ from $s_+(\underline{r})$ and $s_-(\underline{r})$ optically, and $s_c(\underline{r})$ can not be recorded directly on film because it contains negative values.

A digital image processing system,on the other hand, has no difficulty of this kind and it can perform all the operations which an optical system can do. We shall assume, therefore, all the image reconstructions from composite shadowgrams will be done with a digital computer in later studies.

A coded aperture function with a sharply peaked autocorrelation function can only guarantee good reproductions of 2-D source distributions. In case of imaging 3-D source distributions, both the autocorrelation and the crosscorrelation properties of the coded aperture function are important in the coding-decoding process. To see this, we consider the following example. We assume that a 3-D object distribution can be represented by a finite number of slices of its sections along the z axis, with the activities in each section concentrating on the midplane of that section. The separations between two adjacent midplanes of the sections are given by the depth resolution of the system. We let this object distribution be represented by $o_i(\underline{r})$, i=1,2, ...,N_p, and let the distance between the object plane o_i and the aperture plate be given by S_i. The aperture to detector distance is S_o. We can formulate an expression for the shadowgram of this object distribution, similar to Eq.(2.1).

$$s(\underline{r}) = \sum_{i=1}^{N_p} \frac{1}{\alpha_i^2}\, o_i(-\frac{1}{\alpha_i}\,\underline{r}) \circledast h_i(\underline{r}) \qquad (2.7)$$

where $h_i(\underline{r})$ is the point response of the detector to a point source at $\underline{r} = \underline{0}$ of object plane i, and $\alpha_i = d_o/S_i$.

If $h(\underline{r})$ is the transmission function of the coded aperture, then $h_i(\underline{r})$ is given by $h_i(\underline{r}) = h(\underline{r}/(1+\alpha_i))$. The decoding operation on $s(\underline{r})$ can be expressed in a form of convolution integral again. The resulting image $i_j(\underline{r})$ of object plane j is given by Eq.(2.2) with α replaced by α_j. In order to express $i_j(\underline{r})$ in a familiar form we let $h_j'(\alpha_j\underline{r})$ be given by $h_j(\alpha_j\underline{r})$, i.e. the correlation method. Thus, the image $i_j(\underline{r})$ becomes

$$i_j(\underline{r}) = \alpha_j^2\, s(-\alpha_j\underline{r}) \circledast h(\alpha_j\underline{r}/1+\alpha_j) \qquad j = 1, 2, \ldots, N_p$$

or

$$i_j(\underline{r}) = \sum_{i=1}^{N_p} \frac{\alpha_j^2}{\alpha_i^2}\, o_i\left(\frac{\alpha_j}{\alpha_i}\underline{r}\right) \circledast \left[h(-\alpha_j\underline{r}/1+\alpha_i) \circledast h(\alpha_j\underline{r}/1+\alpha_j)\right] \tag{2.8}$$

The desired image is given by the autocorrelation term in Eq.(2.8); the term with i = j:

$$o_j(\underline{r}) \circledast \left[h(-\beta_{jj}\underline{r}) \circledast h(\beta_{jj}\underline{r})\right] = o_j(\underline{r}) \circledast a(\beta_{jj}\underline{r}),$$

$$\beta_{ij} \equiv \alpha_j/(1+\alpha_i) \tag{2.9}$$

The undesired background is given by the crosscorrelation terms; the terms with $i \neq j$:

$$\frac{\alpha_j^2}{\alpha_i^2}\, o_i\left(\frac{\alpha_j}{\alpha_i}\underline{r}\right) \circledast \left[h(-\beta_{ij}\underline{r}) \circledast h(\beta_{jj}\underline{r})\right] = \frac{\alpha_j^2}{\alpha_i^2}\, o_i\left(\frac{\alpha_j}{\alpha_i}\underline{r}\right) \circledast c_{ij}(\underline{r}), \tag{2.10}$$

where the crosscorrelation function $c_{ij}(\underline{r})$ is defined as

$$c_{ij}(\underline{r}) \equiv h(-\beta_{ij}\underline{r}) \circledast h(\beta_{jj}\underline{r})$$

Evidently, the autocorrelation function $a(\underline{r})$ should be peaked as sharp as possible to ensure good image contrast and resolution. At the

same time, it is also necessary that the crosscorrelation functions $c_{ij}(\underline{r})$ to be flattened and diminished as much as possible to prevent build-up of undesired background and artifacts.

Examples of the autocorrelation functions of a few coded aperture functions are shown in Fig.2-1. Examples of the crosscorrelation functions are shown in Fig.2-2. The non-redundant pinhole array gives the best autocorrelation function as well as the crosscorrelation function among the coded aperture functions with non-negative transmission. The autocorrelation function of a combined zone plate function ($h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r})$, 10 zones) is quite close to a delta function as compared to the other autocorrelation functions (Fig.2-1d). The crosscorrelation functions of this combined zone plate function, shown in Fig.2-2b, have some small peaks at and near $\underline{r} = \underline{0}$. The cause of these undesirable peaks is mostly due to the interference between the first harmonic and the other harmonics of the zone plate function. It is highly desirable to eliminate the contribution of most of the unnecessary components of $h_{zp}(\underline{r})$ from the FZP coded shadowgrams to suppress the interference. Complex coding with two zone plates, shown in Fig.2-3a, is not enough to give good reproductions of 3-D source distributions because it eliminates only the contribution from the zero order component of $h_{zp}(\underline{r})$. A better complex coding which utilizes three zone plate shadowgrams is given below for future study. (Complex coding with four zone plates is shown in Ref. 66.) If we combine three shadowgrams, which are formed by the imaging system with three different zone plate functions (shown in Fig.2-3b- 1,2,3), in a complex form the

resulting composite shadowgram contains only the terms of Eq.(2.5) with n = 1, 5, 7, 11, 13, 17, ... With $x = r^2$ and $T = 2r_1^2$, the Fourier series expansions of the three zone plate functions of Fig.2-3b are given by

$$h_1(x) = \frac{1}{2} + \frac{2}{\pi}\sum_n \frac{1}{n}\sin(n\pi/3)\cos(2n\pi x/T) + \frac{2}{\pi}\sum_n \frac{1}{n}\cos(n\pi/3)\sin(2n\pi x/T)$$

$$h_2(x) = \frac{1}{2} - \frac{2}{\pi}\sum_n \frac{1}{n}\sin(n\pi/3)\cos(2n\pi x/T) + \frac{2}{\pi}\sum_n \frac{1}{n}\cos(n\pi/3)\sin(2n\pi x/T)$$

$$h_3(x) = \tilde{h}_{zp}(x) = \frac{1}{2} - \frac{2}{\pi}\sum_n \frac{1}{n}\sin(2n\pi x/T), \qquad n = 1, 3, 5, \ldots$$

If we combine the three shadowgrams resulting from h_1, h_2, and h_3 in a complex form given by $s_1(\underline{r})\, e^{-i\pi/3} + s_2(\underline{r})\, e^{i\pi/3} + s_3(\underline{r})\, e^{i\pi}$, then the equivalent impulse response $h_c(\underline{r})$ of this system is given by the following:

$$h_c(\underline{r}) = \frac{1}{2} h_1(\underline{r}) + \frac{1}{2} h_2(\underline{r}) - h_3(\underline{r}) - i\sqrt{\frac{3}{2}}\left[h_1(\underline{r}) - h_2(\underline{r})\right]$$

$$h_c(x) = \frac{3}{\pi}\left[\sin(2\pi x/T) + \frac{1}{5}\sin(10\pi x/T) + \ldots\right] - i\frac{3}{\pi}\left[\cos(2\pi x/T) - \frac{1}{5}\cos(10\pi x/T) + \ldots\right]$$

$$h_c(\underline{r}) = \frac{3}{\pi i}\, e^{i\pi r^2/r_1^2} + \frac{3}{\pi i}\sum_{n=5}^{n=6m+1} \gamma\, \frac{1}{n}\, e^{i\gamma n\pi r^2/r_1^2}, \qquad (2.11)$$

where m is a positive integer and $\gamma = 1$ if $n = 6m+1$ and $\gamma = -1$ if $n = 6m-1$. The first term on the right side of Eq.(2.11) is the desired "signal carrier" and the rest of the terms are unwanted and relatively small. A plot of the real part of $h_c(\underline{r})$ is shown in Fig.2-3b-4.

2.3 Frequency Domain Analysis

Any system which possesses the properties of linearity and invariance can be described mathematically with considerable ease using the techniques of frequency analysis. Thus, just as it is convenient to describe an audio amplifier in terms of its (temporal) frequency response, so likewise it is often convenient to describe an imaging system in terms of its (spatial) frequency response. To apply the frequency analysis to the previous coded aperture imaging system, the object distribution $o(\underline{r})$ is considered to be an input signal to a linear and space-invariant system which has an impulse response $h(-\alpha\underline{r}) \circledast h'(\alpha\underline{r})$.

The two dimensional Fourier transform of a function $g(x,y)$ is defined by

$$\mathcal{F}\{g(x,y)\} = G(\nu_x,\nu_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\, e^{-i2\pi(\nu_x x+\nu_y y)}\, dx\, dy , \tag{2.12}$$

and the Fourier transform of the convolution integral, $f \circledast g$, is given by the product of the two individual Fourier transforms,

$$\mathcal{F}\{f(\underline{r}) \circledast g(\underline{r})\} = \mathcal{F}\{f(\underline{r})\}\mathcal{F}\{g(\underline{r})\} = F(\underline{\nu})\, G(\underline{\nu}) \tag{2.13}$$

Since the transfer function of a linear and invariant system is defined as the Fourier transform of its impulse response, the transfer function of the coding-decoding process is then given by, omitting the scale factors,

$$\mathcal{F}\{h(-\underline{r}) \circledast h'(\underline{r})\} = H^*(\underline{\nu})\, H'(\underline{\nu}) \tag{2.14}$$

Fig.2-4 shows a schematic block diagram of the coded aperture

imaging system. The input-output signals and the system's responses are shown in both the spatial domain and the frequency domain. It is clear that a perfect reproduction of the object distribution can be achieved only if the total system transfer function $H^*(\underline{\nu})H'(\underline{\nu})$ is unity, which is equivalent to $h(-\underline{r}) \circledast h'(\underline{r}) = \delta(\underline{r})$. The absolute magnitude of the transfer function of a system is called the modulation transfer function of that system. For an imaging system, certain spatial frequency components of the object distribution can never be recovered if the modulation transfer function of this system is zero at those frequencies. In coded aperture imaging if $|H(\underline{\nu})|$ is non-zero everywhere within the limit of the object's frequency band, then the best reconstruction can be achieved with $H'(\underline{\nu}) = 1/H^*(\underline{\nu})$. On the other hand, however, if $|H(\underline{\nu})|$ is zero or relatively low at certain spatial frequencies of the object distribution, application of $H'(\underline{\nu}) = 1/H^*(\underline{\nu})$ to the image reconstruction will lead to amplification of the spatial noise present in the coded image (finite photon statistics) because the noise may not be zero or small at those frequencies. Therefore it is desirable to have $|H(\underline{\nu})|$ reasonably constant over its frequency bandwidth. Fig.2-5 gives a few examples of $|H(\underline{\nu})|$ of various coded aperture systems, which were generated with a digital computer. It is seen that the non-redundant pinhole array system gives a better $|H(\underline{\nu})|$ than those of the other systems, because of the uniformity of $|H(\underline{\nu})|$.

It is worthwhile to point out that $|H(\underline{\nu})|$ of a complex FZP coding system is better than that of a simple FZP coding system (see Fig.2-5b). The reason for this is, of course, because of the lack of the zero

order component in the complex FZP coding governed by $h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r})$. In case of the complex coding with three zone plate shadowgrams, which is governed by $h_c(\underline{r})$ of Eq.(2.11), the modulation transfer function $\left|\mathcal{F}\left[h_c(\underline{r})\right](\underline{\nu})\right|$ shall be even better. In fact, if we make an approximation of $h_c(\underline{r})$ by its first term, the magnitude of $\mathcal{F}\left[h_c(\underline{r})\right](\underline{\nu})$ is a constant.

$$\left|\mathcal{F}\left[h_c(\underline{r})\right](\underline{\nu})\right| = \frac{3r_1^2}{\pi}\left|\mathcal{F}\left[\frac{1}{ir_1^2}e^{i\pi r^2/r_1^2}\right]\right| = \frac{3r_1^2}{\pi}\left|e^{-i\pi r_1^2\nu^2}\right| = \frac{3r_1^2}{\pi}$$

In spite of the fact that $|H(\underline{\nu})|$ of the coded aperture systems may have zero responses at a few frequencies $\underline{\nu}$, we may still use $H'(\underline{\nu}) = 1/H^*(\underline{\nu})$ to perform the decoding in the frequency domain. In digital image processing we can determine the spatial frequency components of the object at frequencies where $|H(\underline{\nu})|$ is not zero by letting $H'(\underline{\nu}) = 1/H^*(\underline{\nu}) = H(\underline{\nu})/|H(\underline{\nu})|^2$ because $I(\underline{\nu}) = O(\underline{\nu})\,H^*(\underline{\nu})\,H'(\underline{\nu}) = O(\underline{\nu})$ at these frequencies (see Fig.2-4). Due to the presence of noise in the imaging process one may want to let $H'(\underline{\nu})$ be zero at frequencies where $|H(\underline{\nu})|$ is zero or relatively small to supress the amplification of noise. However, this will result in a poor image resolution and possibly a certain degree of image deformation due to the loss of essential frequency components of the object in the resulting image. On the basis of resolution and signal to noise, an appropriate form of $H'(\underline{\nu})$ may be

$$\begin{aligned} H'(\underline{\nu}) &= H(\underline{\nu})/|H(\underline{\nu})|^2 \quad , \quad \text{if } |H(\underline{\nu})| > \alpha H_{max} \\ H'(\underline{\nu}) &= H(\underline{\nu})/\alpha^2 H_{max}^2 \quad , \quad \text{if } |H(\underline{\nu})| \le \alpha H_{max} \end{aligned} \tag{2.19}$$

where H_{max} is the largest Fourier frequency component of $|H(\underline{\nu})|$, and

α is a variable parameter which is chosen to optimize the resolution and the signal to noise of the system. Digital reconstruction of FZP coded shadowgrams using the above form of $H'(\underline{v})$ has been studied by Dance et al.[56] The study was based on simple zone plate imaging. The reconstruction is very sensitive to the input noise and the choice of α because $|H(\underline{v})|$ of a simple zone plate imaging system has a large number of "zeros". It is expected that a non-redundant pinhole array system is preferable for the application of Eq.(2.19) in image reconstruction. A study of this will be given in Section 4.2.

In two-dimensional image processing the Fourier transform is often encountered. If a digital computer is used to do the image processing the Fourier transform of a 2-D function f(x,y) in an N x N matrix is performed in discrete form, defined as

$$F(u,v) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y)\, e^{-i2\pi(ux+vy)/N} \qquad (2.20)$$

The evaluation of Eq.(2.20) can be broken into two steps. First, the one-dimensional Fourier transform F(u,y), is taken along the x-coordinate of every horizontal line of f(x,y). Then a second one-dimensional Fourier transform is taken in the y-direction of every vertical line of F(u,y) to yield the composite transform, F(u,v).

Until recently, only brute force methods were available for producing Fourier transform digitally. By conventional techniques, the evaluation of Eq.(2.20) would require $2N^3$ complex multiplications and additions. Fortunately, algorithms have been developed which require

only $2N^2\log(N)$ complex additions and complex multiplications using the above two-step technique. The fast Fourier transform (FFT) algorithm[40] was first proposed by Cooley and Tukey[41] in 1965 for efficiently computing the discrete Fourier transform of a series of data samples. The efficiency of this FFT algorithm is such that solutions to many problems can now be obtained substantially more economically than in the past. An excellent introduction of the FFT algorithm has been given by Brigham in his book "The Fast Fourier Transform".[42]

2.4 Efficiency, Resolution, and Signal to Noise

Geometric Efficiency

The geometric efficiency of a coded aperture imaging system is almost always better than that of a conventional single-pinhole camera because of the larger area of opening of the coded aperture. In coded aperture imaging, however, the gain in geometric efficiency is sometimes not enough to give a better signal-to-noise ratio because of the coding -decoding process involved (this will be explained later).

The geometric efficiency of a multiple pinhole array (MPA) coded aperture system is proportional to the total number and the size of the pinholes in the array aperture.

The geometric efficiency of a Fresnel zone plate (FZP) coded aperture system is proportional to the open area of the zone plate, which is given by $\pi r_N^2/2$ (r_N is the radius of the zone plate).

Resolution

The lateral resolution of a MPA coded aperture system is similar

to that of a single-pinhole camera system, and is given by (see Fig. 2-6a)

$$\Delta_x = d\,(1 + S_1/S_o)\ , \qquad (2.21)$$

where d is the diameter of the pinholes and S_1 and S_o are shown in the figure.

The depth resolution of the MPA coded aperture system is determined not only by the imaging geometry but also by the arrangement of the imaging pinhole array. For the purpose of illustration, a circular multiple-pinhole array of diameter D is used here as an example (d is the diameter of each pinhole). Suppose two point sources separated by distance Δ_z are located on the optical axis of an imaging system as shown in Fig.2-6a and the lateral spatial resolution of that imaging system is given by $\Delta_x = d(1+S_1/S_o)$ a condition for the minimum Δ_z exists and is given by the following equations.

$$\frac{\Delta_x}{\Delta_z} = \frac{D/2}{S_1}$$

or

$$\Delta_z = 2\,\Delta_x\, S_1/D = 2\,d\,(1 + S_1/S_o)\, S_1/D \qquad (2.22)$$

To check whether Eq.(2.22) still applies in the reconstruction of the shadowgram of the two point sources, we construct a simple decoding system (shown in Fig.2-6b) basing on the optical correlation method of Section 2.2. The shadowgram is back-projected by a planar light source (incoherent) through a mask which is a replica of the original circular array. By the principle of autocorrelation, the images of the two point sources will be focused on two image planes corresponding to the two object planes.

The light from the in-focus image on plane 1 produces out-of-focus background on plane 2. The light intensity of the in-focus image on plane 2 is proportional to $1/\pi d''^2$ while that of the out-of-focus background on plane 2 is proportional to $1/\pi D''^2$. If c is defined as the intensity ratio of the in-focus image and the out-of-focus image, then c is given by

$$c = \left(\frac{D''}{d''}\right)^2 = \left(\frac{D'\Delta_z/(S_1+S_o)}{d'S_1/S_o}\right)^2 = \left(\frac{\Delta_z}{d(1+S_1/S_o)S_1/D}\right)^2 \tag{2.23}$$

If we let Δ_z be given by Eq.(2.22), then c=4. This means that the first image is attenuated to 1/4 of the intensity of the second image on plane 2. In other words, the depth resolution obtained previously is adequate.

When a non-circular multiple pinhole array is used, Eq.(2.22) is still valid. In this case, however, the appropriate value of D is given by $2r_m$, where r_m is the median radius of the pinhole locations in the array measured from the center.

Calculations of the lateral and depth resolutions of a zone plate system are quite different from that of a MPA system because they are derived from the focusing properties of a zone plate. The lateral resolution of a zone plate system is given by (from Barrett[11])

$$\Delta_x = \beta(N)\,(1 + S_1/S_o)\,\Delta r_N \tag{2.24}$$

in which $\beta(N)$ is a parameter of order of unity which depends on the number of rings in the zone plate.[57] Δr_N is the width of the outermost ring in the zone plate. Eq.(2.24) shows that the zone plate aperture has a lateral resolution equivalent to a single-pinhole aperture having a pinhole diameter of $\beta(N)\Delta r_N \approx \Delta r_N$.

The depth resolution of a zone plate system is given by[11]

$$\Delta_z = (3/8N) S_1 (1 + S_1/S_o) \qquad (2.25)$$

Eq.(2.25) is of the same form as Eq.(2.22), and the factor (3/8N) is roughly the same as (2d/D) in many cases.

Signal to Noise

In radionuclide imaging, the measured source distributions contain spatial noise due to the finite photon emission nature of the sources. When the imaging system is a single-pinhole camera the signal-to-noise ratio of the image is given by a single parameter. If C_o is the average number of photons collected by the pinhole camera from one resolution element of the source object, then the signal-to-noise ratio is given by, from Poisson statistics,

$$\left(\frac{S}{\sigma}\right)_{ph} = \sqrt{C_o} \qquad (2.26)$$

In coded aperture imaging, calculations of the signal-to-noise ratios for various coded aperture systems are somewhat complicated because of the involvement of coding and decoding. A simplified calculation of the signal-to-noise ratio for a non-redundant pinhole array (NRPA) coded aperture system is given in Appendix A. In the calculation we let the total number of pinholes in the array be given by N and let the source object be given by M equally spaced point sources (or resolution elements) of equal intensities C_o. C_o is defined as before and is given by the average number of detected gamma-ray events which pass through one of the N aperture holes.

From Eq.(A.3) the signal-to-noise ratio of NRPA coded aperture imaging, at the center of the image (the worst case), is given by

$$\left(\frac{S}{\sigma}\right)_{np} = \frac{N\sqrt{C_o}}{\sqrt{N+M}} \quad , \quad \text{if } M \leq N(N-1)$$

$$\left(\frac{S}{\sigma}\right)_{np} = \sqrt{C_o} \quad , \quad \text{if } M > N(N-1) \qquad (2.27)$$

For a 27-hole NRPA coded aperture, a relatively large number of holes, the number of object picture elements M is "large" when $M > 27 \times 26 = 702$. The usual number of picture elements in Nuclear Medicine situations is generally larger than this number and so, in these applications, the equations for $M > N(N-1)$ usually apply. The signal-to-noise ratio for a single pinhole camera is $\sqrt{C_o}$ which is just what we obtain for the NRPA aperture system when $M > N(N-1)$. Since the time is the same for both apertures to get C_o count/pinhole, for the same diameter pinholes, we see that for most nuclear medicine applications of NRPA coded aperture imaging there is no net geometrical advantage over the single pinhole when producing pictures with the same resolution and the same signal-to-noise ratio. FZP coded aperture imaging is also no more efficient than the single pinhole when the number of object picture elements exceeds a certain number but this number is a function of the zone plate used.

Formulations of the signal-to-noise ratio of a zone plate system have been made by a number of investigators.[53,54] We let C_z to be the average number of photons collected by the detector of a zone plate camera from one point source or one resolution element in an array of M equal point sources. The expressions of the signal-to-noise ratio of

this zone plate camera are given by, from Barrett and DeMeester,[53]

$$\left(\frac{S}{\sigma}\right)_{zp} = \frac{\sqrt{2}}{\pi}\sqrt{\frac{C_z}{M}} \tag{2.28}$$

Since the geometric efficiency of a zone plate camera relative to a single-pinhole camera of the same lateral resolution is approximately given by

$$\frac{\text{open area of the zone plate aperture}}{\text{area of a pinhole aperture with diameter } \Delta r_N} = \frac{r_N^2/2}{\Delta r_N^2/4}$$

$$= \frac{2N\, r_1^2}{(\sqrt{N}-\sqrt{N-1})^2 r_1^2} \approx 8N^2 ,$$

Eq.(2.28) may be expressed in the following form.

$$\left(\frac{S}{\sigma}\right)_{zp} = \frac{\sqrt{2}}{\eta}\sqrt{\frac{8N^2 C_o}{M}} = \frac{4N}{\pi\sqrt{M}}\left(\frac{S}{\sigma}\right)_{ph} \tag{2.29}$$

Eq.(2.29) indicates that, in case $M > 16N^2/\eta^2$, the zone plate camera is worse than the single-pinhole camera in signal-to-noise ratio.

3. EXPERIMENTS AND RESULTS

3.1 Multiwire Proportional Chamber Camera System

Multi-wire Proportional Chambers (MWPC) were developed at CERN starting in 1968 as position sensitive detectors for particle physics research as successors to the Spark Chamber detectors which had been in use for the previous 10 years. As the name implies, the MWPC is a gas-filled radiation detector operating in the proportional mode of gas multiplication, and containing a number of anode wires. The chamber is usually rectangular in shape, and the anode wires form a parallel grid. Two cathode planes of parallel wires are needed for locating the position of ionizing events. The use of MWPCs in Nuclear Medicine was developed primarily by our group at the Lawrence Berkeley Laboratory. (Refs.44-46) The good spatial resolution, fast count rate, large detection area and relatively low cost make MWPCs very suitable to a few clinical applications in radionuclide imaging.

The MWPC used in coded aperture imaging is shown in Fig.3-1. The effective sensitive area is 48cm x 48cm. There are three wire planes inside the chamber, the central anode plane and the two mutually orthogonal "drift planes" for locating the x and y coordinates of an ionizing event. The drift planes are at 0 volt above ground, while the anode is at about 3000 volts. The chamber is filled with a gas mixture of 90% xenon and 10% CH_4 at a pressure slightly above the atmospheric pressure. Stainless steel wires of diameter 20 microns spaced 2 millimeters apart

are used for the anode, while wires of diameter 50 microns spaced 1 millimeter apart form the drift planes. The distance between the anode and each drift plane is about 4 millimeters.

The principle of operation of this MWPC system is as follows : When an X-ray or γ-ray photon reaches the active volume of the chamber it has a certain probability to interact with a gas atom and expell an electron through the photoelectric effect. The kinetic energy of the emitted electron equals the photon energy minus the binding energy of that electron. This energy will then be deposited within a small region, mostly in the form of ionization. The higher the density of the gas, the larger the probability of interaction, and the smaller the region of energy deposition. This implies that higher sensitivity and better spatial resolution can be achieved by using a heavy and dense gas under high pressure. Due to the electric field between the wire planes, the secondary electrons (from the ionization of gas atoms by the photoelectron) drift to the electric field between the wire planes, these electrons drift towards the anode wires and undergo multiplication in a small region surrounding a wire, producing a voltage pulse on it. Simultaneously, an induced pulse of the opposite polarity is generated on the cathode wires by the motion of the positive ions. The x and y coordinates of the ionizing event are determined by the use of two electromagnetic delay lines. These delay lines consist of a solenoidal winding on a plastic core which forms an inductance. A distributive capacitance is formed by copper strips as shown in Fig.3-2. Since the delay lines are placed over (but not in contact with) the drift planes, signals from these wire planes are coupled capacitively to the delay

lines. The time lag between the arrival of a pulse on a drift plane wire and its arrival at the end of the delay line depends on the position at which the pulse coupled into the delay line. Thus, the spatial positions of ionizing events are translated into time delays. A reference signal that indicates the occurrence of an ionizing event is obtained by reading out the pulse from the central anode plane through an RC net work with a time constant of 500 n sec. This signal is amplified and fed into the MWPC window discriminator and the zero-cross timing discriminator. The window discriminator has the ability to select the upper and lower pulse heights of the incoming signals. The timing discriminator uses the technique of differentiation and zero-crossing to provide a "time-zero" reference pulse that starts the two image forming systems, the analog system and the digital system. The analog system and part of the digital system are shown in Fig.3-3. A detailed description of the digital system will be given in Chapter 7.

The output signals from the discriminators are used to start two Time-to-Amplitude Converters (TAC). Similarly processed signals obtained from the x-coordinate delay line and y-coordinate delay line are used to stop the TACs. The outputs from the TACs are then used to drive the x and y deflection plates of a CRT (Tektronix 602 display unit), while simultaneously a z unblank signal is applied, as shown in Fig.3-3. Event dots are generated sequentially in the CRT screen and are integrated into an image by a Polaroid camera. A pile-up rejector circuit is used to eliminate multiple events occuring within the delay lines' delay interval ($\sim$3 microseconds) and the dead time of the electronics.

The maximum count rate of this system is about 10^5 events/sec due to the dead time of the delay lines, the associated electronics, and the CRT display unit.

3.2 Multiple Pinhole Array Coded Aperture Imaging

Experimental studies of the MPA coded aperture systems were made with the multiwire proportional chamber detector. The optical correlation method (Section 2.2) was used to decode the MPA coded shadowgrams. Because of the importance of autocorrelation functions in the reconstruction, an investigation of the autocorrelation functions of three different types of pinhole arrays as well as the Fresnel zone plate, shown in Fig.3-4, was made. Their autocorrelation functions, obtained optically, are shown in Fig.3-5 and, for similar types of patterns, drawn schematically in Fig.3-6.

Fig.3-4a shows a non-redundant pinhole array (NRPA). This array has the property that its autocorrelation function (Fig.3-5a) is non-redundant. That is, while the central peak, $a(\underline{o})$, has intensity N corresponding to the overlapping of all N holes from two superimposed arrays, non-central locations of $a(\underline{r})$ have at most intensity 1 corresponding to at most one hole from each of the superimposed arrays overlapping at a given displacement $\underline{r}$ (Fig.3-6a).

In contrast to this array, the regularly spaced array of Fig.3-4b having the same number of holes has a highly redundant autocorrelation pattern (Fig.3-5b and 3-6b). The central peak still has intensity N and

the total amount of background is still the same, N(N-1) but for this array a number of holes have overlapped simultaneously giving a localized and relatively high intensity contribution to the image background. The random array (Fig.3-4c). having N holes distributed randomly over the same area as the previous arrays, has an autocorrelation pattern (Fig.3-5c and 3-6c) similar to that of the non-redundant array, its background peaks however, are sometimes larger than 1. For comparison, the Fresnel zone plate coded aperture is given in Fig.3-4d. Its autocorrelation function (Fig.3-5d and 3-6d) is characterized by a high ratio (0.5) of the high total transmission of the zone plate aperture, 50%, compared with $\sim 2\%$ for the other three pinhole arrays.

By its nature the non-redundant pinhole array has the most desirable autocorrelation function and produces the lowest background intensity in the final image. In a very different application these same non-redundant arrays, with radio telescopes replacing the pinholes, are used in astronomy for mapping radio sources. Some of these arrays are described by M.E.Golay.[32] Smaller non-redundant arrays ($N \leq 15$) are fairly compact, that is, with most of the points of $a(\underline{r})$ lying within a circle with few unoccupied locations inside and only a few points lying outside. Large arrays are built up by computer, starting with smaller ones and using various algorithms.[33] The 27-hole non-redundant array shown in Fig.3-4a was provided by Klemperer. A circular (ring) point array with an odd number of points is also non-redundant. However, the autocorrelation function of this type of array is not as compact and uniform as that of Fig.3-4a.

Imaging of radioactive objects using coded apertures of Fig.3-4 and the multiwire proportional chamber detector was made starting with the non-redundant pinhole array. The test objects were gamma-ray sources of 30 and 60 Kev energies. The coded aperture used were made from 1.5 mm thick lead and the multiple pinhole arrays were usually about 8 cm across. Diameters of the pinholes, 3 mm, were chosen so their shadows on the detector were at least 3 times chamber resolution (1.5 mm).

The shadowgram transparency is made by taking a time exposure of the CRT which displays the gamma-ray events. Shadowgrams in coded aperture imaging are much denser in the center than on the edges and some care must be taken not to saturate the film. The developed transparency is placed on a diffusing screen and illumined from behind (Fig. 3-7a), the light source being a 500 watt projector. The viewing screen is placed about a meter downstream. The mask, aluminum foil perforated with a replica of the coded aperture used, is placed between screen and shadowgram. The initial focussing is done by rotating the mask about a longitudinal axis so it lines up with the original aperture, and by moving the mask longitudinally to focus on the screen. When this is done a simple movement of the screen is all that is necessary to bring different planes in the object into focus. If the mask is made so that the ratio of mask size to film size is the same as the ratio of aperture size to detector size then images of different planes in the object have the same sizes.

The first test was the imaging of a standard Picker Thyroid

Phantom, shown in Fig.3-7b, using the non-redundant pinhole array of Fig.3-4a. The shadowgram (shown in Fig.3-7a) consisting of 540,000 events was recorded on Kodak Ektapan film. The reconstructed image as shown in Fig.3-7a was printed on high contrast paper to reduce the background of light and enhance the contrast. The right and left lobes of the thyroid phantom differed in activity, and this difference can be seen in the reconstruction. The smallest of the cold nodules is 5 mm in diameter and is clearly imaged.

The tomographic capability of this coded aperture system is shown in Fig.3-8. Fig.3-8 shows the shadowgram of a radioactive phantom consisting of three geometrical patterns labelled with ^{125}I (a cross, a triangle, and a circle) which were separated by a distance of 2.5 cm in depth. The imaging geometry for the middle object plane was 1:1, i.e. $S_1=S_o$. A total of one million gamma events were recorded on the shadowgram. Fig.3-8 b, c, and d are the reconstructions of each of the patterns obtained by adjusting the screen to mask distance of the reconstruction system so as to bring each pattern into sharp focus. Each of the in-focus images is superimposed on a background arising from the planes at other depths in the phantom. The tomographic capability of this system was further tested by moving the test phantom toward the aperture such that the maximum depth resolving power of the imaging system was reached (a decrease in S_1 produces a smaller Δ_z). The results, from an imaging geometry $S_1 = \frac{1}{2}S_o$, are shown in Fig.3-9. It is seen that background artifacts from off-plane activities in the reconstructions of Fig.3-8 are not seen in the reconstructions of Fig. 3-9.

Experiments were made to evaluate the lateral and depth resolutions of the above coded aperture system. A reconstruction of four point sources separated by 5 mm, 1 cm, and 2 cm is shown in Fig.3-10a. The lateral resolution of the system, calculated from $\Delta_x = d(1+S_1/S_o)$, is about 4 mm. Fig.3-10b shows the response of the system to a source having 5 mm bars separated by 5 mm. For measurement of depth resolutions, sequential image reconstructions were made for the shadowgram of previous 3 geometrical patterns, at intervals of 6.25 mm along the optical axis of the reconstruction system. The depth resolution, calculated from $\Delta_z = 2d(1+S_1/S_o)S_1/D$ with D given as the median diameter of the 27-pinhole NRPA coded aperture, is about [illegible] mm at $S_1 = 125$ mm. This agrees with the experimental result as shown in Fig. 3-11.

As seen from these reconstructions, good resolution and tomographic effects are easily obtainable with the NRPA coded aperture system. Only a strong light source, a diffusing screen, some optical alignment equipment, and photographic materials are needed to reconstruct the shadowgram. In fact with some modifications to the set-up, direct image decoding is also possible without going through the intermediate step of photographing the shadowgram separately. For example the spot of light from the CRT display, which defines the coordinates of the detected radiation, could be made to pass through the reconstruction pinhole mask and be directly observed in the image plane. By a system of lenses, the various planes within the object could be separately focussed and recorded. Since no movement of the

detector is necessary and all the information of the object's shape and size is recorded in a single picture, dynamic studies are possible.

Only two simple experimental tests were made using multiple pinhole arrays other than the non-redundant pinhole array. The results, together with those from the non-redundant pinhole array, are shown in Fig.3-12 and Fig.3-13. Images of point gamma-ray sources taken with the apertures of Fig.3-4 are shown in Fig.3-12. Images of the thyroid phantom obtained with pinhole arrays which differ only in the placement of their 27 holes, non-redundant, redundant, and random arrays, are compared in Fig.3-13. The redundant array image is only barely recognizable because of the poor distribution of its autocorrelation function. The random array image is much better than the redundant array image. However, it still contains somewhat more background artifacts than the non-redundant array image.

3.3 Fresnel Zone Plate Coded Aperture Imaging

The diffraction properties of a zone plate were described in Section 2.2. The diffraction pattern of a zone plate and the reconstruction of a zone plate shadowgram by diffraction can be derived easily with a convolution formulation of Fresnel diffraction. If a plane aperture mask of amplitude transmittance $t(\underline{r})$ is illuminated by a plane wave of amplitude 1 and wavelength λ the diffracted amplitude $g(\underline{r})$ on a plane at a distance z downstream from the aperture mask is given by the Fresnel-Kirchhoff diffraction formula. After adopting the approxi-

mation of parabolic wavefronts and omitting the obliquity factor, we may express this diffracted amplitude as a convolution,[50,51]

$$g(\underline{r}) = \epsilon\, t(\underline{r}) \circledast w(\underline{r}) , \tag{3.1}$$

where ϵ is a constant phase factor ($\epsilon = e^{i2\pi z/\lambda}$) and $w(\underline{r})$ is called the Fresnel wavefunction :

$$w(\underline{r}) = -\frac{i}{\lambda z}\, e^{i\pi r^2/\lambda z}$$

The Fourier transform of $w(\underline{r})$ is given by $W(\underline{\nu}) = e^{-i\pi\lambda z\nu^2}$.

The Fresnel wavefunction $w(\underline{r})$ has two important properties :

(1) $w(\underline{r})$ is self-orthogonal under the convolution operation, i. e. , $w^*(\underline{r}) \circledast w(\underline{r}) = \delta(\underline{r})$. This result follows from $|W(\underline{\nu})|^2 = 1$.

(2) The convolution of $w(\underline{r})$ with unity is equal to unity :

$$w(\underline{r}) \circledast 1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\underline{r})\, dx\, dy = 1 .$$

The diffraction pattern of a positive zone plate of finite radius r_N is derived as follows : The transmission function $h_{zp}(\underline{r})$ of this zone plate may be expressed as, from Eq.(2.5),

$$h_{zp}(\underline{r}) = \left[\frac{1}{2} + \frac{i}{\pi}e^{-i\pi r^2/r_1^2} - \frac{i}{\pi}e^{i\pi r^2/r_1^2} - \frac{i}{\pi}\sum_{\substack{n\neq 1 \\ \text{odd integers}}}\frac{1}{n}e^{in\pi r^2/r_1^2}\right]\mathrm{circ}(\underline{r}/r_N) , \tag{3.2}$$

where the circle function $\mathrm{circ}(\underline{r}/r_N)$ is defined by

$$\mathrm{circ}(\underline{r}/r_N) = \begin{cases} 1, & \text{if } |\underline{r}| \le r_N \\ 0, & \text{otherwise} \end{cases}$$

The diffraction pattern of this zone plate on a plane at $z = r_1^2/\lambda$ (the major real focus of the zone plate) can be obtained by substituting $t(\underline{r})$ in Eq.(3.1) by $h_{zp}(\underline{r})$. The Fresnel wavefunction at $z = r_1^2/\lambda$ is given by

$$w(\underline{r}) = -\frac{i}{r_1^2}\, e^{i\pi r^2/r_1^2} \tag{3.3}$$

The diffraction pattern (amplitude) is then given by

$$g(\underline{r}) = \mathcal{E}\left\{\left[\frac{1}{2} + \frac{r_1^2}{\pi}w^*(\underline{r}) + \frac{r_1^2}{\pi}w(\underline{r}) + \frac{r_1^2}{\pi}\sum_{\substack{n\neq\pm1\\ \text{odd}}}\frac{1}{n}w(\sqrt{n}\underline{r})\right]\text{circ}(\underline{r}/r_N)\right\} \circledast w(\underline{r}) \tag{3.4}$$

The autocorrelation term in Eq.(3.4), $\left[w^*(\underline{r})\text{circ}(\underline{r}/r_N)\right] \circledast w(\underline{r})$, is calculated as follows.

$$\left[\frac{r_1^2}{\pi}w^*(\underline{r})\text{circ}(\underline{r}/r_N)\right] \circledast w(\underline{r}) = \frac{1}{\pi r_1^2}\left\{e^{-i\pi r^2/r_1^2}\left[\mathcal{F}^{-1}\mathcal{F}\text{circ}(\underline{r}/r_N)\right]\right\} \circledast e^{i\pi r^2/r_1^2}$$

$$= \frac{1}{\pi r_1^4}\iint_{-\infty}^{\infty} e^{-i\pi(x_o^2+y_o^2)/r_1^2}\left[\iint_{-\infty}^{\infty}\frac{J_1(2\pi\sqrt{u^2+v^2}\,r_N)}{\sqrt{u^2+v^2}/r_N}e^{i2\pi(ux_o+vy_o)}du\,dv\right]e^{i\pi[(x-x_o)^2+(y-y_o)^2]/r_1^2}dx_o\,dy_o$$

$$= \frac{r_N}{\pi r_1^2}e^{i\pi(x^2+y^2)/r_1^2}\iint_{-\infty}^{\infty}\frac{J_1(2\pi\sqrt{u^2+v^2}\,r_N)}{\sqrt{u^2+v^2}}\left[\iint_{-\infty}^{\infty}e^{i2\pi[(u-\frac{x}{r_1^2})x_o+(v-\frac{y}{r_1^2})y_o]}dx_o\,dy_o\right]du\,dv$$

$$= \frac{N}{\pi r_N}e^{i\pi r^2/r_1^2}\iint_{-\infty}^{\infty}\frac{J_1(2\pi\sqrt{u^2+v^2}\,r_N)}{\sqrt{u^2+v^2}}\,\delta(u-\tfrac{x}{r_1^2})\,\delta(v-\tfrac{y}{r_1^2})\,du\,dv$$

$$= Ne^{i\pi r^2/r_1^2}\frac{J_1(2\pi r r_N/r_1^2)}{\pi r r_N/r_1^2} = N\,e^{iN\pi(r/r_N)^2}\frac{J_1(2N\pi r/r_N)}{N\pi r/r_N} \tag{3.5}$$

The calculations of the other terms of Eq.(3.4) may be done in the same way as that in the above, the results with approximations [60] are given in the following.

$$\tfrac{1}{2}\,\text{circ}(\underline{r}/r_N) \circledast w(\underline{r}) \approx \tfrac{1}{2}\,\text{circ}(\underline{r}/r_N)\left[1 \circledast w(\underline{r})\right] = \tfrac{1}{2}\,\text{circ}(\underline{r}/r_N)$$

$$\left[\frac{r_1^2}{\pi}w(\underline{r})\text{circ}(\underline{r}/r_N)\right] \circledast w(\underline{r}) \approx \frac{1}{2\pi}\,\text{circ}(\underline{r}/2r_N)\,e^{iN\pi\frac{1}{2}(r/r_N)^2}$$

$$\left[\frac{r_1^2}{\pi}\sum_{n\neq\pm1}^{\text{odd}}\frac{1}{n}w(\sqrt{n}\underline{r})\text{circ}(r/r_N)\right] \circledast w(\underline{r}) \approx \frac{1}{\pi}\sum_{n\neq\pm1}^{\text{odd}}\frac{1}{n(n+1)}\text{circ}(\underline{r}/(n+1)r_N)e^{iN\pi\frac{n}{n+1}(\frac{r}{r_N})^2}$$

Thus, the intensity distribution of $g(\underline{r})$ is given by

$$g^*(\underline{r})g(\underline{r}) = \tfrac{1}{4}\mathrm{circ}(\underline{r}/r_N) + N^2\left[\frac{J_1(2N\pi r/r_N)}{N\pi r/r_N}\right]^2 + \frac{1}{4\pi^2}\,\mathrm{circ}(\underline{r}/2r_N) + \frac{1}{\pi^2}\sum_{\substack{n\neq\pm 1}}^{\text{odd}}\left|\frac{1}{n(n+1)}\right|^2 \mathrm{circ}(\underline{r}/|n+1|\,r_N) + \text{Cross-Product Terms} \qquad (3.6)$$

A graphical interpretation of Eq.(3.6) is shown in Fig.3-14. The intensity distribution of $g(\underline{r})$ is viewed as the output of a zone plate system to a point source input. If the input source distribution is given by $o(\underline{r})$ then the output of this system will be given by $\left|o(\underline{r}) \circledcirc g(\underline{r})\right|^2$ because the Fresnel diffraction amplitude of the zone plate shadowgram of $o(\underline{r})$ is given by $\epsilon s(\underline{r}) \circledcirc w(\underline{r}) = \epsilon\, o(\underline{r}) \circledcirc h_{zp}(\underline{r}) \circledcirc w(\underline{r}) = o(\underline{r}) \circledcirc g(\underline{r})$. Thus, the reconstruction of a zone plate shadowgram is completely determined by the distribution of $g(\underline{r})$. A plot of $g(\underline{r})$, generated by computer, is shown in Fig.3-15. The sharp central peak in the figure is, of course, produced by the Bessel function of Eq.(3.5). The background is small and is the result of the rest of the terms in Eq.(3.4).

The first term on the right side of Eq.(3.6) is relatively high in magnitude as compared with the other non-signal terms. The effect of this term is small when the object imaged is small. When the object imaged is large, however, the intensity of the background produced by this term is given by $\left|o(\underline{r}) \circledcirc \tfrac{1}{4}\mathrm{circ}(\underline{r}/r_N)\right|^2$ while the signal is given by $N^2|o(\underline{r})|^2$. This means that, if we assume the object contains M uniformly distributed resolution elements with intensity 1, the

background in the reconstructed image may be as high as $\frac{1}{4}M^2$ in some cases while the signal is only N^2.

Fig.3-16 shows an optical reconstruction system for zone plate shadowgrams. The two mirrors were used to extend the length of the optical path on a small optical bench. The optical stop at the focal plane of the lens may be used for spatial frequency filtering to remove the undiffracted light. The image formed on a plane somewhere behind the focal plane of the lens corresponds to the negative focal length of the zone plate pattern. This real image can be seen through a telescope located at proper distance away from the image plane. Fig.3-17a shows a zone plate shadowgram of four point sources separated at least by twice of the system resolution. The coherent optical reconstruction of this shadowgram is shown in Fig.3-17b.

The reconstruction of zone plate shadowgrams by Fresnel diffraction can also be performed with a digital computer. Aside from the abvious advantages in image processing with the computer, mentioned in Section 2.2, there are some other advantages of using the computer in coded aperture imaging. A computer system is free of the problems associated with film such as small dynamic range,film grain noise. Results with the computer can be obtained quickly without the delay associated with a two-step photographic process. In computer reconstruction of zone plate shadowgrams the convolution integral of Eq.(3.1) is evaluated by use of Fourier transforms. Direct evaluation of the 2-D convolution integral of two N x N matrices requires N^4 multiplications and additions while the indirect method requires only three

Fast Fourier Transforms (Section2.3) and N^2 complex multiplications in the frequency domain.

The Fresnel diffraction amplitude $i(\underline{r})$ at $z = r_1^2/\lambda$ of a zone plate shadowgram $s(\underline{r})$ is given by the convolution of $s(\underline{r})$ with $w(\underline{r})$, where $w(\underline{r})$ is the Fresnel wavefunction at z. The Fourier transform of $i(\underline{r})$ is given by

$$I(\underline{\nu}) = \epsilon\, S(\underline{\nu})\, W(\underline{\nu}) = \epsilon\, S(\underline{\nu})\, e^{-i\pi r_1^2 \nu^2} \qquad (3.7)$$

The reconstruction of $s(\underline{r})$ by computer is done by first multiplying the Fourier transform of $s(\underline{r})$ by $e^{-i\pi r_1^2 \nu^2}$ and then taking the inverse Fourier transform of the product.

$$i(\underline{r}) = \mathcal{F}^{-1}\left\{ S(\underline{\nu})\, e^{-i\pi r_1^2 \nu^2} \right\} \qquad (3.8)$$

A computer generated thyroid phantom was used to test this method The reconstructed image is shown in Fig.3-18, plotted by the computer. It is seen that the reconstruction contains a large amount of background, which is, as explained before, due to the positive bias component (=1/2) of the zone plate. This high level background may be eliminated to certain degree by frequency filtering,[11] at the expense of losing part of the low frequency components of the object in the resulting image. Many investigators have used the off-axis section of a zone plate, for which the undiffracted light pattern does not overlap the desired diffraction pattern and is therefore not a problem. A halftone screen must be used here as a spatial frequency carrier which moves the object frequency spectrum into the passband of the off-axis zone plate.[61] The major drawback to the off-axis zone plate is that

its fine rings require a high-resolution image detector to be used. A much better approach to the background elimination is to use the complex coding method of Section 2.2. Applications of this method to the zone plate system as well as the multiple pinhole array system will be given in Section 4.1.

4. FURTHER STUDIES AND CONCLUSIONS

4.1 Complex Coded Aperture Imaging

One of the problems associated with coded aperture imaging is that the background level in the reconstructed image of an extended object is always high if only simple binary coded apertures are used. As it is shown in Section 2.2, perfect reconstruction can be made only if the coding function $h(\underline{r})$ and the decoding function $h'(\underline{r})$ are related by $h(-\underline{r}) \circledast h'(\underline{r}) = \delta(\underline{r})$. For binary coded aperture functions $h(\underline{r})$, there is no decoding function $h'(\underline{r})$ such that $h(-\underline{r}) \circledast h'(\underline{r}) = \delta(\underline{r})$. The idea of complex coded aperture imaging is that by a combination of two or more shadowgrams we get a new shadowgram which can result in an image without background. The combination of the shadowgrams is made in such a way that the new shadowgram appears to be from a coded aperture with both positive and negative transmissions and the reconstruction of this new shadowgram contains no background.

FZP Coded Aperture Imaging

It was shown in Section 3.3 that the background in the coherent optical reconstruction of a positive zone plate shadowgram is from the positive bias term of the zone plate transmission function, $\frac{1}{2}\,\mathrm{circ}(\underline{r}/r_N)$. This positive bias term can be eliminated, as mentioned in Section 2.2, if we combine two zone plate shadowgrams s_1 and s_2, $s_1(\underline{r})$ from a positive zone plate and $s_2(\underline{r})$ from its corresponding negative zone plate, by $s(\underline{r}) = s_1(\underline{r}) - s_2(\underline{r})$. The reconstruction of $s(\underline{r})$ is the same as that of an ordinary zone plate shadowgram.

The Fresnel diffraction amplitude of $s(\underline{r})$ at $z = r_1^2/\lambda$ is given by

$$i(\underline{r}) = \epsilon\, s(\underline{r}) \circledast w(\underline{r}) = \epsilon\, o(\underline{r}) \circledast \left[h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r}) \right] \circledast w(\underline{r}) \tag{4.1}$$

From Eqs.(2.4), (2.6), and (3.3) we have

$$i(\underline{r}) = \epsilon\, o(\underline{r}) \circledast \left[\frac{2r_1^2}{\pi} \sum_{n=\text{odd}} \frac{1}{n} w(\sqrt{n}\underline{r}) \text{circ}(\underline{r}/r_N) \right] \circledast w(\underline{r}) \tag{4.2}$$

The intensity distribution of $i(\underline{r})$ is given by (see Eq.(3.6))

$$i(\underline{r})i^*(\underline{r}) = 2N^2 \left[\frac{J_1(2N\pi r/r_N)}{N\pi r/r_N} \right]^2 + \frac{1}{2\pi^2} \text{circ}(\underline{r}/2r_N) + \ldots \tag{4.3}$$

Thus, the background term, $\frac{1}{2}\text{circ}(\underline{r}/r_N)$ in Eq.(3.6), is eliminated through the use of this complex coding technique.

The autocorrelation of a combined zone plate function $h_{zp} - \tilde{h}_{zp}$, as described in Section 2.2, can also be used for the reconstruction of the combined shadowgram $s(\underline{r})$. In fact, reconstruction of $s(\underline{r})$ by autocorrelation is almost the same as reconstruction of $s(\underline{r})$ by Fresnel diffraction because their point responses are similar.(compare Fig.2-1d and Fig.3-15) The reconstructed image of $s(\underline{r})$ by the autocorrelation technique is given by

$$i(\underline{r}) = s(\underline{r}) \circledast \left[h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r}) \right] = o(\underline{r}) \circledast a_c(\underline{r}), \tag{4.4}$$

where $a_c(\underline{r})$ is the autocorrelation function of the combined zone plate function $h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r})$. A computer simulation of this complex coded aperture imaging was made with a computer generated thyroid phantom (as the object). The reconstruction of the combined shadowgram $s(\underline{r})$ was done by autocorrelation, and the resulting image is shown in Fig.4-1. It is seen that this image contains no background and is much better than that from the simple coded aperture imaging, as shown in Fig.3-18.

A further test was done to study this complex coding technique with a 3-D source distribution. The source distribution consisted of three sheet patterns, a triangle, a cross and a circle, separated by 2 cm in depth. The diameter of the zone plate (9 zones) was 13 cm. The result from the reconstruction using complex coding with two zone plates (positive and negative) is shown in Fig.4-2. It is seen that the structures from the off-plane activities produce a large amount of background artifacts in the in-focus images. The reason for this unpeasant result is because of the multiple-focusing property of the zone plate. Each $n \neq 0$ component of the zone plate transmission function (see Eq.(2.5)) corresponds to a focus (real or virtual for $n<0$ or $n>0$). In the reconstruction of the zone plate shadowgram, the contribution of the $n \neq 0$ & -1 components is relatively small when the $n=-1$ component is in autocorrelation, and could be large when the $n=-1$ component is in crosscorrelation (in this case one of the $n \neq 0$ & -1 components may be in autocorrelation and therefore result in an unwanted peak). As was explained in Section 2.2, the reconstruction of a 3-D source distribution contains mixtures of autocorrelations and crosscorrelations, therefore, background artifacts are produced in the reconstruction from all the non-signal carriers of the zone plate (i.e. those $n \neq -1$ terms of Eq.(2.5)).

Among all the non-signal carriers of a zone plate, the components with $n=0$, $n=1$, $n=-3$, $n=3$ are most important. Complex coding with two zone plates (positive and negative) can eliminate only the contribution from the component with $n=0$ while complex coding with three zone plates

(shown in Fig.2-3b) can eliminate the contribution from all of them. (see Eq.(2.11)) Therefore, it is expected that by complex coding with three or even more zone plates the background artifacts in the reconstruction of 3-D source distributions can be reduced.

MPA Coded Aperture Imaging

It is also possible to combine two or more MPA coded shadowgrams such that the resulting shadowgram shares the same autocorrelation property as the previous combined zone plate shadowgram. For a large random pinhole array with thousands of holes and 50% transmission, we define its "negative" in the same way as that for a negative zone plate, within the boundary of the array. We let the transmission function of this random pinhole array be given by $h_p(\underline{r})$ and let the transmission function of its negative be given by $h_n(\underline{r})$. Within the boundary of the array $h_n(\underline{r})$ equals 1 if $h_p(\underline{r})$ is 0 and vice versa. The combined function, $h(\underline{r}) = h_p(\underline{r}) - h_n(\underline{r})$, has a sharply peaked autocorrelation function. This can be shown as follows

$$a(\underline{r}) = h(-\underline{r}) \circledast h(\underline{r}) = \left[h_p(-\underline{r}) - h_n(-\underline{r})\right] \circledast \left[h_p(\underline{r}) - h_n(\underline{r})\right] \quad (4.5)$$

$$a(\underline{r}) = a_p(\underline{r}) + a_n(\underline{r}) - c_{pn}(\underline{r}) - c_{np}(\underline{r}) \quad (4.6)$$

where

$a_p(\underline{r}) \equiv h_p(-\underline{r}) \circledast h_p(\underline{r})$ and $a_n(\underline{r}) \equiv h_n(-\underline{r}) \circledast h_n(\underline{r})$ are autocorrelation functions, and $c_{pn}(\underline{r}) \equiv h_p(-\underline{r}) \circledast h_n(\underline{r})$ and $c_{np}(\underline{r}) \equiv h_n(-\underline{r}) \circledast h_p(\underline{r})$ are crosscorrelation functions.

To derive expressions for the autocorrelation and crosscorrelation functions, we first consider the correlation function of two arbitrary

functions f(x,y) and g(x,y).

$$c(x,y) = f(-x,-y) \circledast g(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\eta,\xi)\, g(\eta+x,\xi+y)\, d\eta d\xi \qquad (4.7)$$

If f(x,y) and g(x,y) are aperture transmission functions of binary values, 0 and 1, the correlation function c(x,y) becomes

$$c(x,y) = \text{sum of all 1's of the product } f(\eta,\xi)\, g(\eta+x,\xi+y) \qquad (4.8)$$

Now, we assume the pinholes and the entire array are square in shape, with the width of each pinhole given by w and the width of the entire array given by L. The autocorrelation function $a_p(\underline{r})$ (or $a_n(\underline{r})$) and the crosscorrelation function $c_{pn}(\underline{r})$ (or $c_{np}(\underline{r})$) are given by

$$a_p(x,y) = \text{sum of all 1's of the product } h_p(\eta,\xi)\, h_p(\eta+x,\xi+y)$$

$$c_{pn}(x,y) = \text{sum of all 1's of the product } h_p(\eta,\xi)\, h_n(\eta+x,\xi+y)$$

It is very clear that $a_p(0,0)$=N=the total number of holes in the array and $c_{pn}(0,0)$=0 because $h_n(\eta,\xi)$ is the negative of $h_p(\eta,\xi)$. As $|x|$ and $|y|$ become slightly greater than w, $a_p(x,y)$ drops suddenly to N/2 and $c_{pn}(x,y)$ increases suddenly to N/2 because half of the holes in $h_p(\eta,\xi)$ are expected to overlap to those in either $h_p(\eta+x,\xi+y)$ or $h_n(\eta+x,\xi+y)$. As $|x|$ and $|y|$ become larger and larger, both $a_p(x,y)$ and $c_{pn}(x,y)$ become smaller and smaller, and become zero finally when $|x|$ and $|y|$ are greater than L. Thus, the appropriate expressions for $a_p(x,y)$ (or $a_n(x,y)$) and $c_{pn}(x,y)$ (or $c_{np}(x,y)$) are

$$a_p(x,y) \approx \tfrac{1}{2}N\Lambda(x/L)\Lambda(y/L) + \tfrac{1}{2}N\,\mathrm{rect}(x/w)\,\mathrm{rect}(y/w) + \sigma_p(x,y) \qquad (4.9)$$

$$c_{pn}(x,y) \approx \tfrac{1}{2}N\Lambda(x/L)\Lambda(y/L) - \tfrac{1}{2}N\,\mathrm{rect}(x/w)\,\mathrm{rect}(y/w) + \sigma_{pn}(x,y) \qquad (4.10)$$

where the rectangle function rect(x/w) is defined by

$$\text{rect}(x/w) = \begin{cases} 1, & |x| \le w \\ 0, & \text{otherwise} \end{cases}$$

and the triangle function $\Lambda(x/L)$ is defined by

$$\Lambda(x/L) = \begin{cases} 1-|x|/L, & |x| \le L \\ 0, & \text{otherwise} \end{cases}$$

The noise terms of Eqs.(4.9) and (4.10), $\sigma_p(x,y)$ and $\sigma_{pn}(x,y)$, give measures of the deviations of the correlation functions from their expected values. Since $w \ll L$ for a large array of thousands of holes, the rectangle function rect(x/w) represents a sharply peaked square pulse of small width. The autocorrelation function $a(\underline{r})$ of Eq.(4.6) is given by

$$a(\underline{r}) \approx 2N \text{ rect}(x/w)\text{rect}(y/w) + \sigma(x,y), \qquad (4.11)$$

where $\sigma(x,y) = \left(\sigma_p^2 + \sigma_n^2 + \sigma_{pn}^2 + \sigma_{np}^2 \right)^{1/2}$

We see that the background term, $\Lambda(x/L)\Lambda(y/L)$, in the autocorrelation function of a simple random pinhole array is eliminated through this complex process. Since the distribution of $a(\underline{r})$ is like a delta function, the complex coding technique used for zone plate imaging can be applied here too. However, for random pinhole arrays of small number of holes, complex coding may not work well because of the noise term of Eq.(4.11). This noise term is relatively unimportant when the number of holes in the array is very large, but can be important if this number decreases.

For pinhole arrays other than the random pinhole array, the complex coding technique still applies. For example, one of the

simplest way to do complex coding with a non-redundant pinhole array is to apply Eq.(4.5) using $h_n(\underline{r}) = h_p(-\underline{r})$, where $h_p(\underline{r})$ is the original non-redundant pinhole array and $h_n(\underline{r})$ represents the array obtained by rotating the original array 180°. Computer simulations of this case were made and the results showed that only small improvements were achieved this way. There is a better way to eliminate the background in the reconstruction of NRPA coded shadowgrams. The method will be studied in the next section.

4.2 Decoding by Fourier Transform Deconvolution

The frequency domain analysis of coded aperture imaging in Section 2.3 gives us a method of decoding in the frequency domain. Eq.(2.19) is a form of the decoding transfer function $H'(\underline{\nu})$. Two expressions are used for $H'(\underline{\nu})$ according to the magnitude of the coding transfer function $|H(\underline{\nu})|$.[56]

$$
\begin{aligned}
H'(\underline{\nu}) &= H(\underline{\nu})/|H(\underline{\nu})|^2 , \quad \text{if } |H(\underline{\nu})| \geq \alpha H_{max} = \alpha \,\mathrm{Max}(|H(\underline{\nu})|) \\
H'(\underline{\nu}) &= H(\underline{\nu})/\alpha^2 H_{max}^2 , \quad \text{if } |H(\underline{\nu})| \leq \alpha H_{max}
\end{aligned}
\tag{4.12}
$$

The parameter α should be carefully chosen to supress the amplification of statistical fluctuations in the measurement, without losing too much fidelity. We define the transfer function $H(\underline{\nu})$ at frequency $\underline{\nu}$ to be a "zero" if $|H(\underline{\nu})|$ is smaller than the cut-off value αH_{max}. The total number of zeros in $|H(\underline{\nu})|$ of a coded aperture system gives us a measure of the system's response.

The reconstructed image $i(\underline{r})$ in the frequency domain is given by

$$I(\underline{\nu}) = S^*(\underline{\nu})\, H'(\underline{\nu}) = O(\underline{\nu})\, H^*(\underline{\nu})\, H'(\underline{\nu}) .$$

At those frequencies $\underline{\nu}$ where $|H(\underline{\nu})| \leq \alpha H_{max}$ the frequency components of the image are given by $I(\underline{\nu}) = O(\underline{\nu})\, H^*(\underline{\nu})\, H(\underline{\nu}) \,/\, \alpha^2 H_{max}^2$.

In the spatial domain, the above expression becomes

$$i(\underline{r}) = o(\underline{r}) \circledast \left[h(-\underline{r}) \circledast h(\underline{r}) \right] / \alpha^2 H_{max}^2 = \left[o(\underline{r}) \circledast a(\underline{r}) \right] / \alpha^2 H_{max}^2 \quad (4.13)$$

Thus, if $|H(\underline{\nu})|$ has large percentage of zeros, the reconstruction with Eq.(4.12) is dominated by the autocorrelation.

An investigation of the total number of zeros in $|H(\underline{\nu})|$ as a function of α for a few coded aperture functions was made with a digital computer. The fast Fourier transform algorithm was again used to compute the transform $H(\underline{\nu})$. The test functions were stored in matrices of 128 x 128 elements and the transforms of these functions were stored in complex matrices of 65 x 128 elements. Three zone plate functions and one pinhole array function were used for the investigation. The zone plate functions were $h_{zp}(\underline{r})$, $\tilde{h}_{zp}(\underline{r})$, and $h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r})$, which represented a positive zone plate with 9 zones, a negative zone plate with 10 zones, and a combined zone plate with 10 zones respectively. All of the zone plates have the same r_1 (radius of the first zone). The diameter of the positive zone plate was equivalent to 76 elements (or bins) of the matrix used, and the width of its last annular zone was about 2 bins. The pinhole array function was $h_{np}(\underline{r})$, which represented a non-redundant pinhole array with 9 holes.(shown in Fig.4-5a) The size of the array was about the same as those of the zone plates and the diameter of each

pinhole in the array was slightly less than 2 bins. The results from the computer calculations are listed below.

$h(\underline{r})$ =	$h_{zp}(\underline{r})$	$\tilde{h}_{zp}(\underline{r})$	$h_{zp}(\underline{r})-\tilde{h}_{zp}(\underline{r})$	$h_{np}(\underline{r})$
	Total number of zeros (samples with values $\leq \alpha H_{max}$) in 8320 samples of $\lvert H(\underline{\nu})\rvert$			
α = 0.001	1535	1150	123	--
α = 0.01	6784	6483	970	11
α = 0.05	--	--	--	455
α = 0.1	--	--	--	1789

It is seen that, for the same α, $\lvert H(\underline{\nu})\rvert$ of the non-redundant pinhole array has far fewer zeros than those of the zone plates. Because of this fact and the good autocorrelation property of the non-redundant pinhole array, it is expected that the non-redundant pinhole array can give us better image reconstruction with Eq.(4.12) than the zone plates.

The non-redundant pinhole array was used to study the decoding method given by Eq.(4.12). The initial work was to examine the system's behavior under various conditions with a point gamma-ray source. Computer simulations of the coding-decoding process with different values of α and different levels of input statistical noise were made, and the results are shown in Fig.4-3. It is seen that the width of the signal peak (or resolution) is smaller for smaller values of α while the magnitude of the noisy background is higher for these α.

From Fig.4-3, a compromise between the resolution and the signal to noise is probably given by α=0.05. This value of α was used in a test reconstruction of a 2-D source distribution. The test object was a

computer generated thyroid phantom with 5% statistical noise in it. This 5% noise-to-signal ratio was based on that the system received a random number of events with an average of 400 from each resolution element of the object if the intensity of that element is 1. The result of the reconstruction together with the non-redundant pinhole array used are shown in Fig.4-5. It is seen that good reconstruction without background can be obtained with this decoding method.

Effects of α on the point reconstructions at planes some distances away from the in-focus plane were also studied. The results from computer simulations, as shown in Fig.4-4, indicate that the off-plane reconstructions are not very sensitive to changes of α . The average magnitude of the background fluctuations in each off-plane reconstruction is very small, about 1% of that of the in-focus signal peak. This background to signal ratio is much smaller than that of the zone plate system in the previous section.

A computer simulation of the reconstruction of a 3-D source distribution (consisting a cross, a circle, and a triangle) was made and the results are shown in Fig.4-6. The input statistical noise was 5% for each resolution element of the source object. α=0.05 was used. It is seen that images of Fig.4-6 are much better than those of Fig.4-2.

One important thing has not been explained yet. The results from the computer reconstruction using the method described in this section seem to be not as good as those from the optical reconstruction using the correlation method (see Figs.3-7, -8, and -9). Since the point reconstructions from the method given by Eq.(4.12) are better than those from

the optical correlation method, the reconstructions shown in Figs. 4-5 and -6 should look better than those shown in Figs. 3-7, -8, and -9. The reason for the better appearance of the reconstructions in Figs.3-7, -8, and -9 is because those reconstructions were printed on high contrast paper. The biasing action of the paper suppressed the background artifacts in those reconstructions.

4.3 Conclusions

It was shown in Section 3.2 that a NRPA coded aperture system provides us a simple and economical way to obtain the tomographic images of an extended 3-D source distribution. It was also shown that both the autocorrelation and the crosscorrelation of a non-redundant pinhole array are suitable for coded aperture imaging. The non-redundant nature of the pinholes results in good image resolution and uniform background with few artifacts in the reconstructed images.

Improvement of simple coding by complex coding was discussed in Section 4.1. By use of a positive and a negative zone plates one can effectively remove the background produced by the zero order component of the zone plate from the reconstruction. However, background artifacts are still present in the reconstruction of a 3-D distribution because of the other unwanted components of the zone plate. It is expected that by use of three or even more zone plates one may effectively remove a few of the most unwanted components of the zone plate (see Section 2.2), but this complicates the imaging and reconstruction process.

In theory the best coding with zone plates can be obtained if two sinusoidal (sine and cosine) zone plates are used. In this case a complex coding with the two zone plates can give an equivalent impulse response $e^{i\pi r^2/r_1^2}$ (=cos.zp+i sin.zp), which corresponds a constant modulation transfer function (see Section 2.3).

The Fourier transform deconvolution method in Section 4.2 appears to be very powerful for decoding of NRPA coded shadowgrams. The in-focus point reconstructions (Fig.4-3) are very good on the basis of resolution and signal to noise. The out-of-focus point reconstructions (Fig.4-4) give only a small amount of background which fluctuates above and below zero. This results in very few background artifacts in the reconstructions of 3-D source distributions. Eq.(4.12) is only one of the many possible forms for the decoding transform $H'(\underline{\nu})$. Other forms of $H'(\underline{\nu})$ may be found in the future, which can minimize or eliminate the background artifacts.

It was mentioned in the begining that coded aperture imaging is only a special branch of tomography. In coded aperture imaging the direction of each gamma-ray event can not be determined exactly because of the aperture used. A few other tomographic imaging techniques (or devices, see Section 0.2), on the other hand, operate on events with known directions and form their images by principle of back-projection. The advantage of having the direction of each event know is that better estimate of the 3-D source distribution can be made from enough data samples of the distribution using various algorithms (depending on the sampling). A new algorithm for 3-D image reconstruction from tomography

will be given in Part II. The particular way of sampling for that image reconstruction algorithm will also be given. The direction of each gamma-ray event collected by the detector must be known there.

PART II : THREE-DIMENSIONAL IMAGE RECONSTRUCTION FROM FOCAL-PLANE TOMOGRAPHY

5. FOCAL-PLANE TOMOGRAPHY AND THREE-DIMENSIONAL IMAGE RECONSTRUCTION

5.1 Gamma Camera Systems for Focal-Plane Tomography

As mentioned in the begining, a number of existing radionuclide imaging devices other than the coded aperture imaging system can also give depth information about a 3-D source distribution. These devices have in common that they provide tomographic images of the object, that is, that images of a given object plane have that plane in focus and all other object planes contribute an out-of-focus background superimposed on the in-focus image.

Imaging devices which give depth information about a source point require radiation from the point to be detected from distinctly separate directions. In these imaging devices the direction of each gamma ray event is known and, if separate views of a source distribution have been made, a tomographic image on any plane through the source can be made by back-projecting the gamma rays onto that plane.

Since the direction of each gamma-ray event is known, the depth information content of these imaging devices is much higher than that of a coded aperture imaging system. In fact, the tomographic images obtained from these imaging devices have some unused information about the original 3-D source distribution. In the following chapters, a method of 3-D image reconstruction which utilizes this unused information to improve

the tomographic images will be studied. Three imaging devices which give tomographic images are described below.

Multiple Single-Pinhole Camera

One such tomographic device uses multiple single-pinhole views (Fig. 5-1a). The depth-information properties of these multiple views is illustrated in Fig.5-1b where the source distribution is a single point. An exposure is made using one pinhole selected from the array. Tomographic images on a number of planes are made by back-projecting photons from this exposure through the same pinhole and adding the appropriate intensity to each tomographic plane at its intersection with the line. The process is repeated for the other views and the final tomographic plane image is the sum of contributions from all these views. The plane which actually contained the point source has a sharp image while in other planes the image of the point source is blurred out.

In analyzing image formation it is useful to use the point response function $h_{ij}(\underline{r},\underline{r}')$. This function, characteristic of the imaging device used, describes the response at point $\underline{r}$ in plane j to a point source at point $\underline{r}'$ in plane i. From Fig.5-1b it is seen that this function, or blurring pattern, for the multiple single-pinhole camera has a shape similar to the original array of pinholes but with a size which depends upon the geometry. The non-redundant pinhole arrays which were used previously for coded aperture imaging can be used to obtain multiple projections for this camera system, and a 9-hole non-redundant pinhole array is shown in Fig.5-2a.

Rotating Slanted-Hole Collimator Camera

Another device used to obtain tomographic images in Nuclear Medicine is the rotating slanted-hole collimator(Fig.5-3). The collimator rotates about an axis perpendicular to the detector and the parallel holes are slanted at an angle to this axis, generally about 20 degrees. When the collimator is at a given position the image of a point source is a single point on the detector. When the collimator has rotated 180° the image of this point source has traveled on the arc of a circle to an opposite position. As done previously, tomographic images on a number of transverse planes can be made by back-projecting the detector image obtained at a given position of the collimator along the known direction of the parallel holes and then repeating this process for all positions of the collimator.

In one mode of operation of this camera views are taken at discrete positions of the collimator and the blurring patterns have a shape similar to Fig.5-2b.[62] In another mode the collimator is rotated continuously during data collection and the blurring pattern is a circle (Fig.5-2c).[63]

Positron Camera

Positron cameras are currently under intensive development[64,65] because of their ability to give tomographic images without the use of a collimator and the associated loss of intensity. The two 511 KeV annihilation gamma-rays from a positron source radiate from the source point at 180° to each other (Fig.5-4). Interactions with two detectors determine, as in the previous cases, only a line on which the source lies. Projection of events detected onto a transverse plane gives, again, a

tomographic image of the source distribution with that plane in focus and other planes blurred and superimposed.

The fraction of detected events from a point source in the midplane, say, decreases considerably as the point source moves away from the center of the plane. The three-dimensional image reconstruction method which will be studied later requires that the point response function remains constant in shape, size, and intensity as the point source moves over the camera field of view on a given plane (this is called spatial invariance), although it may be different for different planes. This blurring pattern can be maintained constant over a given area of a plane if the computer which constructs the tomographic image planes accepts data only for those events for which $|x_2-x_1| \leq d$ and $|y_2-y_1| \leq d$ where d is smaller than w, the width of the detectors (square in shape). The region of constant detection efficiency for the midplane which results is a square of width w-d. The blurring pattern is also a square (Fig. 5-2d). If the detectors are of disc shape, the appropriate expression for the acceptable data is $|\underline{r}_1-\underline{r}_2| \leq d$. The blurring pattern is therefore a disc.

5.2 Solution for Three-Dimensional Images From Tomographic Images

Conventional radionuclide focal-plane tomography produces images of radioactive objects within the human body in which structures at a chosen depth appear in focus whereas those at other depths are blurred.

The inherent limitation of these techniques is their inability actually to remove background signal due to objects not in the chosen focal plane. Tomography disperses the events from off-plane activity but such events are still present as a semi-uniform background that can produce artifacts or obscure faint structures even when the structures are located on the focal plane of a given tomographic image or tomogram. In Nuclear Medicine imaging with its poor resolution and frequently low contrast objects it is often not possible to distinguish the in-focus plane from the out-of-focus images. Removal of the background would enable detection of smaller lesions and lesions of lower contrast.

A tomographic image of a plane section through an object has a finite width slab of the object in focus. The thickness of this slab is the depth of field which depends on geometry and detector resolution. Also present in this tomographic image is an out-of-focus background contributed by the rest of the object. If this background is removed from a collection of these tomographic planes which are separated by the system depth of field, then the three-dimensional object is known to an accuracy determined by the longitudinal and lateral resolution of the system.

Now, the remaining problem is how the out-of-focus background can be removed. Let us examine the constituents of a tomographic image first. We assume in the following that the object is located in N_p planes and is represented by the functions $o_i(\underline{r})$, i=1, ..., N_p. N_p tomograms $t_j(\underline{r})$ are formed by a computer from camera data using the backprojection methods discussed previously. Since the point response

function $h_{ij}(\underline{r},\underline{r}')$ represents the system response at $\underline{r}$ on plane j due to a point source of unit intensity at $\underline{r}'$ on plane i, the total contribution from plane i to plane j is just the integral of $o_i(\underline{r}')h_{ij}(\underline{r},\underline{r}')$ over all $\underline{r}'$. If the imaging system is made to have a constant point response function as described previously, then $h_{ij}(\underline{r},\underline{r}')=h_{ij}(\underline{r}-\underline{r}')$ and the above integral becomes a convolution integral of $o_i(\underline{r})$ and $h_{ij}(\underline{r})$. The tomographic image $t_{ij}(\underline{r})$ on plane j due to source $o_i(\underline{r}')$ on plane i is then given by

$$t_{ij}(\underline{r}) = \int o_i(\underline{r}')h_{ij}(\underline{r}-\underline{r}')d\underline{r}' = o_i(\underline{r}) \circledast h_{ij}(\underline{r}) \qquad i,j=1,2,\ldots N_p \quad (5.1)$$

Since the tomographic image has contribution from all object planes we have

$$t_j(\underline{r}) = \sum_{i=1}^{N_p} o_i(\underline{r}) \circledast h_{ij}(\underline{r}) \qquad j=1,2,\ldots..N_p \quad (5.2)$$

Taking Fourier transform on both sides of the above equation, we have

$$T_j(\underline{\nu}) = \sum_{i=1}^{N_p} O_i(\underline{\nu})\, H_{ij}(\underline{\nu}) \qquad j=1,2,\ldots. N_p \quad (5.3)$$

The above equation is linear and can be solved for $O_i(\underline{\nu})$ easily if $\det(H_{ij}(\underline{\nu})) \neq 0$ for all $\underline{\nu}$. The object distribution in spatial domain can be obtained by taking inverse Fourier transform of $O_i(\underline{\nu})$. A detailed description of this 3-D image reconstruction process will be given in the next chapter.

It is worthwhile to point out that the tomographic images obtained from a coded aperture imaging system can not be used for the above image reconstruction because of $\det(H_{ij}(\underline{\nu})) = 0$ for all $\underline{\nu}$ in coded aperture imaging. To prove this, we let $h_i^c(\underline{r})$ represent the response of the

coding system to a point source on object plane i and let $h_j^d(\underline{r})$ represent the response of the decoding system on image plane j to a point on the shadowgram for coded aperture imaging. Assuming the system is linear and space-invariant, the point response function $h_{ij}(\underline{r})$ of the coded aperture system is then given by $h_{ij}(\underline{r}) = h_i^c(\underline{r}) \circledast h_j^d(\underline{r})$. The Fourier transform of $h_{ij}(\underline{r})$ is given by $H_{ij}(\underline{\nu}) = H_i^c(\underline{\nu})\, H_j^d(\underline{\nu})$. The determinant of ($H_{ij}(\underline{\nu})$) of this system is given by

$$\det\,(H_{ij}(\underline{\nu})) = \begin{vmatrix} H_{11} & H_{21} & \cdots & H_{N_p 1} \\ H_{12} & H_{22} & \cdots & H_{N_p 2} \\ \vdots & & & \\ H_{1N_p} & H_{2N_p} & \cdots & H_{N_p N_p} \end{vmatrix} = H_1^c H_2^c \cdots H_{N_p}^c \begin{vmatrix} H_1^d & H_1^d & \cdots & H_1^d \\ H_2^d & H_2^d & \cdots & H_2^d \\ \vdots & & & \\ H_{N_p}^d & H_{N_p}^d & \cdots & H_{N_p}^d \end{vmatrix}$$

Therefore, $\det(H_{ij}(\underline{\nu}))=0$ for all $\underline{\nu}$ in coded aperture imaging.

6. BASIC MATHEMATICS FOR THREE-DIMENSIONAL IMAGE RECONSTRUCTION

6.1 Tomographic Images from Focal-Plane Tomography

As explained in the previous chapter, the formations of tomographic images by the three tomographic imaging camera systems are similar, that is, all the tomographic images are made through back-projection. Therefore, only the multiple single-pinhole camera system is used to illustrate the formation of tomographic images. To produce the tomographic images for a pinhole camera system, a number of separate exposures of the object are obtained with a pinhole aperture and a planar gamma ray detector, and with the pinhole being moved successively to different points of an N_h-point array (Fig.6-1a). The known location of the pinhole and the event positions recorded by the detector determine the direction of emission for each detected gamma-ray. Using these N_h single-pinhole images tomographic images on N_p transverse planes through the object are built up with computer in the following manner (Fig.6-1b). The K^{th} pinhole image is projected back through the K^{th} pinhole onto each of N_p tomographic planes. This is done for all the pinhole images and the final tomographic images t_j, $j=1,\ldots, N_p$, are the sum of the contributions from each of the N_p single-pinhole images. These are true tomographic images, the image t_1, for example, having in-focus all the points of the corresponding object plane o_1 with all other planes contributing out-of-focus background.

If the three-dimensional object can be approximated by intensity distributions $o_i(\underline{r})$ on a finite number of planes, $i=1,\ldots, N_p$, then to

construct these o_i's from the tomographic images one needs to consider the equations which describe the formation of the tomograms.

We assume here and in the following that the pinholes in the array are small (delta functions). The effect of finite width pinholes will be discussed later. Let the function $h_{ij}(\underline{r})$ be the response of the j^{th} tomographic plane to a point source located in the i^{th} object plane. As discussed previously, the response of the tomogram $t_j(\underline{r})$ to $o_i(\underline{r})$ is its convolution with $h_{ij}(\underline{r})$, $o_i(\underline{r}) \circledast h_{ij}(\underline{r})$, and $t_j(\underline{r})$ is the sum of its responses to the N_p object planes.

We can find the form of $h_{ij}(\underline{r})$ and the scale factor in $o_i(\underline{r})$ from Fig.6-1c which shows the response of two tomographic planes to a point source located on an object plane. The figure shows that $h_{ii}(\underline{r})$ is a delta function of intensity N_h and that $h_{ij}(\underline{r})$ ($j \neq i$) is the pattern of the N_h hole array used but with a reduced size depending on geometry. If $h(\underline{r})$ is the function which describes the pinhole array transmission with pinholes located at positions $\underline{r}_k$, k=1,..., N_h, we have

$$h_{ij}(\underline{r}) = h(\underline{r}/m_{ij}) = \sum_{k=1}^{N_h} \delta(\underline{r}-m_{ij}\underline{r}_k) \qquad i,j=1,2,\ldots,N_p \quad (6.1)$$

where $m_{ij} \equiv (s_i-s_j)/s_i$ and s_i is the distance of t_i (and of o_i) from the pinhole array. The object intensity which contributes to the tomogram $t_j(\underline{r})$ is $o_i(\underline{r}/\alpha_{ij})/\alpha_{ij}^2$ where $\alpha_{ij} \equiv s_j/s_i$. The equation for the j^{th} tomogram is then

$$t_j(\underline{r}) = \sum_{i=1}^{N_p} \frac{1}{\alpha_{ij}^2}\, o_i(\underline{r}/\alpha_{ij}) \circledast h_{ij}(\underline{r}) \qquad j=1,2,\ldots,N_p \quad (6.2)$$

Eq.(6.2) is a set of N_p equations in the N_p variables $o_i(\underline{r})$. The t_j's are combinations of single-pinhole image data and the h_{ij}'s depend only

on geometry, that is, on pinhole locations in the array and on placement of the reconstruction planes. All the equations are functions of $\underline{r}$, position relative to the optic axis, and these N_p equations must be solved for each position $\underline{r}$, in digital processing for instance, on a 64 x 64 matrix.

The tomographic images of the other two camera systems can be expressed by the same formula as above, of course, with different point blurring functions $h_{ij}(\underline{r})$ as shown in Fig.5-2.

6.2 Image Reconstruction by Fourier Transform Deconvolution

Taking the Fourier transform of Eq.(6.2) we have

$$T_j(\underline{\nu}) = \sum_{i=1}^{N_p} O_i(\underline{\nu}\,\alpha_{ij})\, H_{ij}(\underline{\nu}), \qquad j=1,2,\ldots,N_p \quad (6.3)$$

where the quantities T_j, O_i, H_{ij} are functions of spatial frequency $\underline{\nu}$ and are the Fourier transform of the corresponding quantities of Eq. (6.2). To exhibit these equations as a set of linear equations in the quantities $O_i(\underline{\nu})$ we need to evaluate each of the equations for $T_j(\underline{\nu})$ at a different point $\underline{\nu}' \equiv s_j\,\underline{\nu}$

$$T_j(\underline{\nu}'/s_j) = \sum_{i=1}^{N_p} O_i(\underline{\nu}'/s_i)\, H_{ij}(\underline{\nu}'/s_j), \qquad j=1,2,\ldots,N_p \quad (6.4)$$

or in matrix form $T = H\,O$ (6.5)

where

$$T = \begin{pmatrix} T_1(\underline{\nu}'/s_1) \\ T_2(\underline{\nu}'/s_2) \\ \vdots \\ T_{N_p}(\underline{\nu}'/s_{N_p}) \end{pmatrix} \qquad O = \begin{pmatrix} O_1(\underline{\nu}'/s_1) \\ O_2(\underline{\nu}'/s_2) \\ \vdots \\ O_{N_p}(\underline{\nu}'/s_{N_p}) \end{pmatrix}$$

$$H = \begin{pmatrix} H_{11}(\underline{\nu}'/s_1) & H_{21}(\underline{\nu}'/s_1) & \cdots & H_{N_p 1}(\underline{\nu}'/s_1) \\ H_{12}(\underline{\nu}'/s_2) & H_{22}(\underline{\nu}'/s_2) & \cdots & H_{N_p 2}(\underline{\nu}'/s_2) \\ \vdots & \vdots & & \vdots \\ H_{1N_p}(\underline{\nu}'/s_{N_p}) & H_{2N_p}(\underline{\nu}'/s_{N_p}) & & H_{N_p N_p}(\underline{\nu}'/s_{N_p}) \end{pmatrix} \tag{6.5}$$

For those (angular) spatial frequencies $\underline{\nu}'$ for which the deteminant of H, $D(\underline{\nu}') \equiv \left| H_{ij}(\underline{\nu}'/s_j) \right|$, is not zero, an inverse matrix, H^{-1}, exists and Eqs.(6.4) can be sloved for $O_i(\underline{\nu}'/s_i)$.

$$H^{-1}\, T = H^{-1}\, H\, O = O \tag{6.6}$$

Inverse Fourier transformations then give the desired background-free images $o_i(\underline{r})$.

$$\mathcal{F}^{-1}\, O = \mathcal{F}^{-1} \begin{pmatrix} O_1(\underline{\nu}) \\ O_2(\underline{\nu}) \\ \vdots \\ O_{N_p}(\underline{\nu}) \end{pmatrix} = \begin{pmatrix} o_1(\underline{r}) \\ o_2(\underline{r}) \\ \vdots \\ o_{N_p}(\underline{r}) \end{pmatrix} \tag{6.7}$$

If the deteminant function $H(\underline{\nu}')$ is zero at some angular spatial frequency $\underline{\nu}'$ then the reconstructed transform images $O_i(\underline{\nu})$ are not determined at the spatial frequency $\underline{\nu}'/s_i$. From Eqs.(6.4) and (6.1) we can get

$$H_{ij}(\underline{\nu}'/s_j) = \sum_{k=1}^{N_h} e^{-i2\pi \underline{\nu}' \cdot \underline{r}_k (s_i - s_j)/s_i s_j} \qquad i,j=1,2,\ldots,N_p \tag{6.8}$$

Since $H_{ij}(\underline{0}) = N_h$=constant for all i's and j's, the deteminant $D(\underline{\nu}')$ is always zero at $\underline{\nu}' = \underline{0}$. This means that the reconstructions $o_i(\underline{r})$ are

undefined by an additive constant. This is not a problem if this is the only zero since this constant can be determined by some subsidiary condition, for instance, that $o_i(\underline{r})$ has no negative value and the background outside the field of view is supposed to be zero.

The general determination of additional zeros of the determinant $D(\underline{\nu}')$ is complicated and has not been done yet. The physical reason for zeros in the binocular vision case (N_h=2), for example, is as follows. Spatial frequency on each of two input planes can be such as to give the same spatial frequency on the detector plane for each of the two single-pinhole exposures. On reconstruction one is uncertain how much of the original frequencies to a ascribe to each plane. A regularly spaced array will increase the number of zero. However numerical calculations of $D(\underline{\nu}')$ for non-redundant pinhole arrays of 9 to 24 pinholes and for 2-, 3-, and 4- plane geometries showed no zeros to exist, other than at $\underline{\nu}'=\underline{0}$. Although no general proof is made it would appear that zeros of the determinant can easily be avoided by the use of non-redundant pinhole arrays.

The determinant function $D(\underline{\nu}')$ is characteristically very low near $\underline{\nu}'=\underline{0}$ but rapidly increases to its maximum value. $D(\underline{\nu}')$ varies with the shape of the pinhole array as well as with the number and the location of the tomographic planes. Fig.6-2 gives the cross-sectional plots of the determinant functions as functions of array types (regular and non-redundant arrays) and plane geometries (2-, 3-, and 4- plane geometries). The non-redundant pinhole arrays are seen to give a determinant $D(\underline{\nu}')$ of more uniform magnitude compared to the regular pinhole array. Fig.6-2-1. shows the pinhole arrays used for computing those determinant functions.

6.3 Systems of Finite Spatial Resolution

To investigate the effect of using an aperture array which has finite width pinholes of diameter d we consider the geometry of Fig.6-1c. A point source in plane i casts a shadow of a given pinhole onto the detector plane which has a diameter $d'=d(s_o+s_i)/s_i$. The diameter of this spot cast on the tomographic plane j through the corresponding zero-width pinhole on the right of the figure (zero-width because this is a mathematical operation not a physical process) is $d_{ij}=d's_j/s_o=d\ s_j(1/s_i+1/s_o)$. The real point response function that occurs in Eq.(6.2) is $h^c_{ij}(\underline{r}) \equiv h_{ij}(\underline{r}) \circledast \mathrm{circ}(\underline{r}/d_{ij})$. $\mathrm{circ}(\underline{r}/d_{ij})$ represents a circle function of diameter d_{ij}, which is defined by

$$\mathrm{circ}(\underline{r}/d_{ij}) = \begin{cases} 1, \text{ if } |\underline{r}| \leq d_{ij}/2 \\ 0, \text{ if } |\underline{r}| > d_{ij}/2 \end{cases} \tag{6.9}$$

If we define

$$d_i \equiv d\ s_i(1/s_i + 1/s_o) = d_{ij}/\alpha_{ij} \tag{6.10}$$

then

$$\mathrm{circ}(\underline{r}/d_{ij}) = \mathrm{circ}(\underline{r}/d_i\alpha_{ij})$$

Upon replacing $h_{ij}(\underline{r})$ in Eq.(6.2) by $h^c_{ij}(\underline{r})$ we have

$$t_j(\underline{r}) = \sum_{i=1}^{N_p} \frac{1}{\alpha_{ij}^2}\, o_i(\underline{r}/\alpha_{ij}) \circledast h_{ij}(\underline{r}) \circledast \mathrm{circ}(\underline{r}/d_i\alpha_{ij}) \qquad j = 1, 2, \ldots, N_p \tag{6.11}$$

If we define a new object distribution function $o^c_i(\underline{r})$ by

$$o^c_i(\underline{r}) \equiv o_i(\underline{r}) \circledast \mathrm{circ}(\underline{r}/d_i), \qquad i = 1, 2, \ldots, N_p \tag{6.12}$$

then the expression of $t_j(\underline{r})$ in Eq.(6.11) becomes

$$t_j(\underline{r}) = \sum_{i=1}^{N_p} \frac{1}{\alpha_{ij}^2}\, o^c_i(\underline{r}/\alpha_{ij}) \circledast h_{ij}(\underline{r}) \qquad j = 1, 2, \ldots, N_p \tag{6.13}$$

Eq.(6.13) has the same form as Eq.(6.2), therefore, the solution of $o_i^c(\underline{r})$ can be obtained by the same deconvolution method as described previously for the infinite resolution system.

The reconstructed images of this finite resolution system are given by $o_i^c(\underline{r})$, which is defined in Eq.(6.12) as a convolution integral of $o_i(\underline{r})$ and $\mathrm{circ}(\underline{r}/d_i)$. The operation of the circle function on $o_i(\underline{r})$ is considered as a blurring operation which occurs whenever an imaging system of finite resolution is used.

Apparently, the image resolution of the entire process, tomographic imaging and 3-D image reconstruction, depends only on the first part of the process. In other words, the resolution of the image is defined when the object is imaged by the camera, and the later image processing by the computer makes no change to this resolution (provided that the computer has enough memory to represent and process the data from the camera). Therefore, the formulas for the lateral and depth resolutions of a MPA coded aperture system can still be applied to this 3-D image reconstruction system. If a pinhole array of mean radius r_m (weighted by the distribution of pinholes in the array) is used for the tomographic imaging and 3-D reconstruction, then the lateral and depth resolutions for object plane i, at s_i from the pinhole array aperture, are given by

$$\Delta_x = d\,(1 + s_i/s_o)\,, \tag{6.14}$$

and

$$\Delta_z = d\,(1 + s_i/s_o)\,s_i/r_m\,, \tag{6.15}$$

where $i = 1, 2, \ldots, N_p$.

7. EXPERIMENTS AND RESULTS

7.1 Construction of Digital Image Read-Out System

The MWPC system described in Section 3.1 is again used here. Since the 3-D image reconstruction method requires the use of fast digital Fourier transforms, the analog system is no longer feasible now. Instead, a digital computer together with an interface controller and some peripheral devices are used to read and store the MWPC output. As shown in Fig.7-1 the output signals from the timing and pulse height discriminators are used to start and stop a Camac digitizer. The function of the pile-up rejector used in the analog system (in Section 3.1) is replaced by a trigger conditioner. A delay gate sets the busy time interval for the trigger conditioner so that any signal pulses coming from the central wire plane within that time interval after the first one, which turns on the trigger conditioner, will be rejected.

The prompt output of the trigger conditioner is used to start the digitizer and to trigger a pulse generator. The pulse generator then generates two output pulses, both are delayed by a short time interval. One pulse is used to set the Look-at-me (LAM) signal for the Camac digitizer. A LAM signal is a signal sent by the Camac module to the Camac crate controller to indicate request for attention. The other pulse is used to stop counting of the digitizer. The time delay of these two pulses from the prompt output of the trigger conditioner is about 5 microseconds while the busy time interval of the trigger conditioner is

about 130 microseconds. This long busy time interval of the trigger conditioner is required because the PDP-11/10 software takes about 120 microseconds to read the x and y data, to locate and increase a particular memory location according to the x and y data, and to check the status of the switch register and the crate controller.

The timing of the Camac digitizer is provided by a 150 MH_z clock, and the digital information of the x and y signals is transmitted through the "read bus" lines of a Camac crate to the crate controller. The PDP-11/Camac crate controller serves as an interface between the MWPC electronics and the PDP-11/10 computer. This type 1533A Interface/crate controller made by Borer Electronics Ag serves to connect a Camac crate directly to a PDP-11 Unibus. Built as a double width Camac module, the crate controller is designed on the principle of offering transparent operation in the read/write mode whereby each crate subaddress appears as a separate peripheral on the Unibus. The 1533A has an interrupt vector generator for 16 vectors with individual priority selection which speeds the handling of LAM's considerably.

The PDP-11/10 central processor controlls the signal reading, the image processing, the data transmission, and the image display of the MWPC system. All the computer system components and peripherals of PDP-11/10 connect to and communicate with each other on a single high-speed bus known as the Unibus, see Fig.7-1. The 28K core memory of the PDP-11/10 provides enough space for temporary data storage. For more processed data and long term storage, the disk unit of a PDP-11/45 system is available. Data transmission between the PDP-11/10 and the PDP-11/45

is through a telephone line with maximum speed ~ 9600 bauds(bits per second). The tape unit of the PDP-11/45 system is also available for the permanent storage of processed data on magnetic tapes. The LA36 DEC-writer is a fast(relative to the teletype) data terminal. It prints from a set of 64 characters at speeds up to 30 characters per second. Data entry is made from the 97-character keyboard. The Tektronix-4012 Computer Display Terminal is used for graphic plotting and input-output characters. Inputs to the computer are made through the Terminal's keyboard. Data from the computer can be displayed on the face or screen of the Terminal's cathode ray storage tube. Fig.7-2 shows a photograph of the entire experimental set-up except the PDP-11/45 system, which is a remote station.

7.2 Software for Digital Image Processing

The intrinsic resolution of the MWPC detector and its associated electronics is about 2 millimeters. Since the active area of this chamber is about 48cm x 48cm, the total number of resolution elements in this chamber is about 240 x 240 = 57.6K which is about twice of the memory size of the PDP-11/10 computer. To run a program with the PDP-11/10 computer, about 8k memory space is required to store the system software, and another 4k memory space should be preserved for the program and subroutines which read and process the MWPC image data. Only 16k out of the 28k memory space is available for storing the digitized image data of the MWPC system. In other words, the output of the MWPC can be

at most represented by a 128 x 128 single-word matrix in the computer programming. Because of this limitation, the spatial resolution of the digital system is at least 2 times worse than the intrinsic resolution of the MWPC detector.

BASIC is used in the digital system because it is convenient. The computer program which reads and bins the MWPC image data contains an assembly language subroutine which runs a loop with a cycle time about 120 microseconds. This subroutine commands the computer to check the LAM signal of the Camac digitizer first. When there is a LAM signal it means that an event is digitized by the Camac digitizer and is ready for computer to read. The computer then start to read the x and y coordinates of that event through the Camac/PDP-11 interface controller. The values of x and y are then used to determine the memory location for that event in the 128 x 128 matrix which is assigned to begin at some particular memory address. The content of the memory location for that event is increased by 1. The subroutine then commands the computer to initialize the Camac digitizer through the Camac/PDP-11 interface controller and wait for the next LAM signal. If the last switch register on the PDP-11/10 operator's console is raised at this time, the read cycle ends, otherwise, the read cycle continues. Since there are only 16 bits in one memory word of the PDP-11/10 computer and the image matrix is represented by single-word elements, the maximum permissible number of events in one element is $2^{16} - 1 = 65535$ which is far above the maximum value obtained experimentally.

The image of the MWPC output can be displayed on the screen of the

Tektronix 4012. The 6" x 8" display area of the Tektronix 4012 contains 780 x 1024 points for the cathode ray beam. Since the width of a light spot on the screen is about 2 points, the total number of resolution elements on the display screen is 390 x 512. To display the intensity distribution of an image which is stored in an $N_1 \times N_2$ matrix, we divide the display screen into $N_1 \times N_2$ squares. Each square contains $\frac{390}{N_1} \times \frac{512}{N_2}$ gray-shade elements which are used to represent the contents of a matrix element. If an image of the MWPC output is stored in a 128 x 128 square matrix, the maximum number of gray shades for each matrix element is $3 \times 3 = 9$. For a better gray-shade display, the image should be rebinned into a 64 x 64 matrix. The gray-shade level is then $6 \times 6 = 36$.

To make tomographic images for the 3-D image reconstruction, a number of image matrices are needed. The PDP-11/10 computer can take only one 128 x 128 matrix and the rest of the matrices must be stored temporarily in a direct accessible medium. The disc unit of the PDP-11/45 computer is used for this purpose. The data processing is done by transmitting data in and out between the PDP-11/10 computer and PDP-11/45 disc, and the final result is transmitted to the PDP-11/45 tape unit and stored on a magnetic tape permanently.

It is very time consuming and inconvenient to have to transfer data in and out from one system to another. To avoid this inconvenience, the last part of the 3-D image reconstruction which involves the use of Fourier transforms and matrix inversion and multiplications is done by a large off-line computer (CDC-7600). The formation of tomographic images for the 3-D image reconstruction is still done by the on-line

PDP-11/10 computer and the resulting tomograms are recorded on a magnetic tape using the PDP-11/45 tape unit. The magnetic tape is then used as an input to the CDC-7600 computer. Tape conversion is necessary because that a PDP-11 computer word contains 16 bits while a CDC-7600 computer word contains 60 bits. The 9-track magnetic tape obtained from the PDP-11 system must be decoded and recoded into CDC-7600 words before it can be used for the second part of the image reconstruction.

The formation of tomographic images from a set of projections is the most time consuming operation in the entire 3-D image reconstruction process. Because it requires N_p coordinate shiftings for each event received by the detector and the total number of events in each projection is usually quite large (about 10^4-10^5). If a computer with storage capacity larger than the PDP-11/10 computer is available, the formation of tomograms can be done while the MWPC data is being taken (provided that the data taking rate is low). In this case, no extra time is required to produce the tomograms. The image reconstruction of the tomograms takes very little time if the inverse matrix H^{-1} is predetermined (H^{-1} may be predetermined because it does not depend on the input object). To illustrate the total computing time required to make a reconstruction we let $N_p = 3$ and let the three tomograms be stored in matrices of size 64 x 64.

The first step of the 3-D image reconstruction process is Fourier transformation of the three tomograms. The fast Fourier transform (FFT) algorithm described in Section 2.3 is used here to perform the digital transforms. It takes less than 20 milli-seconds for a large computer

(CDC-7600) to do one fast Fourier transform using an assembly language subroutine of FFT.

The second step is to perform the complex multiplications and additions of H^{-1} and the transformed tomograms T (both H^{-1} and T are complex matrices).

$$\begin{pmatrix} O_1(\underline{\nu}) \\ O_2(\underline{\nu}) \\ O_3(\underline{\nu}) \end{pmatrix} = O = H^{-1}T = \begin{pmatrix} H_{11}^{-1}(\underline{\nu}) & H_{12}^{-1}(\underline{\nu}) & H_{13}^{-1}(\underline{\nu}) \\ H_{21}^{-1}(\underline{\nu}) & H_{22}^{-1}(\underline{\nu}) & H_{23}^{-1}(\underline{\nu}) \\ H_{31}^{-1}(\underline{\nu}) & H_{32}^{-1}(\underline{\nu}) & H_{33}^{-1}(\underline{\nu}) \end{pmatrix} \begin{pmatrix} T_1(\underline{\nu}) \\ T_2(\underline{\nu}) \\ T_3(\underline{\nu}) \end{pmatrix}$$

$$O_i(\underline{\nu}) = H_{i1}^{-1}(\underline{\nu})T_1(\underline{\nu}) + H_{i2}^{-1}(\underline{\nu})T_2(\underline{\nu}) + H_{i3}^{-1}(\underline{\nu})T_3(\underline{\nu}) , \quad i = 1, 2, 3 \tag{7.1}$$

It takes three complex multiplications and two complex additions to evaluate each $O_i(\underline{\nu})$ at one of the 64 x 64 frequencies $\underline{\nu}$. If we neglect the computing time of complex additions, which is small compared to that of complex multiplications, we have 3 x 64 x 64 complex multiplications in the second step. Since the Fourier transform of a real function has a property that the real part is even and the imaginary part is odd, all the frequency components can be represented by complex matrices of size 33 x 64. The above number of complex multiplications is therefore reduced by a factor of 2. On a large computer (CDC-7600) the time required for doing those complex multiplications is much less than a FFT of 64 x 64 elements (less than 10 milli-second).

The last step is the inverse Fourier transformation of $O_i(\underline{\nu})$. It takes also about 20 milli-seconds for each inverse FFT. Thus, the total computing time required for the reconstruction of three images from three tomograms is roughly given by 20 ms x 3 + 10 ms x 3 + 20 ms x 3

= 150 ms = 0.15 second. It is seen that this computing time is proportional to the total number of tomograms used.

7.3 Computer Simulations and Experimental Results

Computer Simulations

A 9-pinhole array was used in a computer simulation in order to test the 3-D image reconstruction method. The pinhole array is shown in Fig.5-2a. The object (Fig.7-3b) was located in three planes, s_1=8 cm, s_2=10 cm, and s_3=12 cm. The computer generated the tomographic images of Fig.7-3c in the same three planes. The reconstructions produced, using these tomographic images as input, are given in Fig.7-3d. The results show excellent agreement with the original object. The tomographic images were produced, however, with no statistical variation in intensity of the picture elements of the object from one single pinhole exposure to the next.

In the more realistic case when the object picture elements vary statistically it is found that there occurs a small component of background in the reconstructed images in addition to the expected variation in the intensity of the image elements. This is shown in Fig. 7-4 where an average of 400 events total from each object picture element has been collected, distributed statistically over the N_h=9 single pinhole exposures. This gives a 5% statistical fluctuation in object picture element intensities. The corresponding fluctuation measured in the images is 6%. Thus, the reconstruction method introduces a small amount of noise. If the total number of events collected is held fixed

it is found that the amount of introduced noise increases slightly as the number of pinholes is decreased.

In another case, a five-plane object was imaged and reconstructed by the same system. Again, 5% statistical fluctuation was generated in the object picture elements, and the tomographic images of Fig.7-5a were formed. The reconstructions are shown in Fig.7-5b. Only a small amount of noise is seen. A study on statistical noise in the image reconstruction process will be given in Section 8.2 and also in Appendix B.

Experiments with Radioactive Sources

Radioactive sources were used to test the reconstruction method. The detector and the associated digital image read-out system were described in Section 7.1. The software for the image processing was discussed in the previous section. In all the experimental tests, nine projections of the object distribution were taken with a circular pinhole array (like the one shown in Fig.5-2b) for the fomation of focal-plane tomograms. For a given object field of view a circular pinhole array gives the best depth resolution among the multiple pinhole arrays of the same size. The 9-hole circular array used was of diameter 9 cm and each pinhole of the array was of diameter 4 mm.

The first test object, consisting of three geometric patterns, was located on three object planes, s_1=20 cm, s_2=25 cm, and s_3=30 cm. S_o was 39 cm. This test object was labelled with I-125 (average photon energy about 30 Kev). There were about 10^5 events in each object plane and the spatial resolution was about 6.5 mm. The tomograms and their reconstructions are shown in Fig.7-6. Although the choice of objects

was such as to make their nature readily apparent from the tomograms alone, the reconstruction method has clearly removed artifacts and background successfully from the tomograms. No effort was made to improve the image reconstruction by correcting the non-uniformity of the MWPC detector and the variation in transmissions of the pinholes.

The second test object consisting of five numerical patterns was located on five object planes, s_1=20 cm, s_2=25 cm, s_3=30 cm, s_4=35 cm, and s_5=40 cm. All other imaging parameters were unchanged. The point responses of the system were obtained first. They were then used to compute the inverse frequency matrix, H^{-1}. A rough correction of the detector non-uniformity and pinhole transmission was made so that about the same total number of events was found in each pinhole projection.

In this test the background events due to scattering and cosmic radiation were about 10% of the signal events received by the MWPC detector. Angular (spatial) tomograms were formed on line with the PDP-11/10 computer (also with the PDP-11/45 computer). The reconstruction of these tomograms was performed by the CDC-7600 computer at the Lawrence Berkeley Laboratory. The tomograms and their reconstructions are shown in Fig.7-7 (the tomograms are shown in their ordinary way to avoid confusion).

As a demonstration, an object consisting of five point sources was used in place of the previous object. The point sources were regularly spaced at 5 cm intervals in depth. The tomograms and their reconstructions are shown in Fig.7-8. The point blurring patterns of the imaging system (multiple single-pinhole camera) can be seen clearly in the last 3 tomograms. Fig.7-8b shows that the 3-D reconstruction method has

removed all the off-plane point blurring patterns from the tomograms.

From Eqs.(6.14) and (6.15) we can calculate the lateral and depth resolutions of the imaging system (5-plane geometry), using d=4 mm and r_m=4.5 cm. Both the lateral and the depth resolutions are worse for object planes further from the pinhole array aperture. The calculated lateral resolutions are 6.1 mm, 6.6 mm, 7.1 mm, 7.6 mm, and 8.1 mm on object plane 1, 2, 3, 4, and 5 respectively. The calculated depth resolutions are 2.7 cm, 3.6 cm, 4.7 cm, 5.9 cm, and 7.2 cm on object plane 1, 2, 3, 4, and 5 respectively. It is seen that the imaging system can not resolve the last two object planes well because the system's depth resolutions are greater than the plane spacing (5 cm) on the last two object planes. This may be the reason that the reconstructed images on the last three object planes are not as good as those on the first two object planes (see Figs. 7-7b and 7-8b).

8. FURTHER STUDIES AND CONCLUSIONS

8.1 Determinant Function

The discussion given in section 6.2 showed that the determinant function $D(\underline{\nu}')$ is always zero at zero spatial frequency for the multiple pinhole camera systems. In fact, $D(\underline{\nu}')$ is zero at $\underline{\nu}'=0$ for all the camera systems which perform focal-plane tomography. This general property, $D(\underline{0})=0$, arises from the fact that only a projection at 90^{o} will give the total intensity of an object plane. All other projections for an object distribution which has, say, N_p planes of uniform intensity ($o_i(\underline{r})=I_i$) give the same value ($\sum_i I_i$) and assignment of a given plane's intensity is not possible.

$D(\underline{\nu}')$ for various multiple single-pinhole camera systems has been discussed previously. The calculations of $D(\underline{\nu}')$ for the positron camera and the rotating slanted-hole collimator camera (or the tomocamera) systems are as follows: The point response function, $h_{ij}(\underline{r})$, of a tomocamera system is either a delta function ($i=j$) or a ring-shaped function ($i\neq j$, for the continuous type). The Fourier frequency function, $H_{ij}(\underline{\nu})$, of a ring-shaped function is a Bessel function of order zero ($J_o(\nu)$) while that of a delta function is a constant.

As discussed in section 5.1, data from the positron camera should be taken with some maximum allowable difference in the coordinate values, $d \geq \max(|x_1-x_2|)$ and $d \geq \max(|y_1-y_2|)$. If the detectors of the positron camera are of disc shape, the appropriate expression for the above difference in coordinate values is $d \geq \max(|\underline{r}_1-\underline{r}_2|)$, where $\underline{r}_1$ and $\underline{r}_2$ are

positions on the two detectors of the positron camera respectively. When the above is done, in each object plane which is sufficiently near the midplane there is an area where point sources are detected with constant efficiency. The point response function, $h_{ij}(\underline{r})$ with $i \neq j$, is then given either by a 2-D rectangle function (rect(x)rect(y)) or by a circle function (circ($\underline{r}$)), according to the shape of the detectors of the positron camera. The corresponding Fourier frequency functions, $H_{ij}(\underline{\nu})$, are a sinc function $\left(\frac{\sin(\nu_x)}{\nu_x}\frac{\sin(\nu_y)}{\nu_y}\right)$ and a Bessel function of order one $\left(\frac{J_1(2\pi\nu)}{\nu}\right)$. Fig.8-1 shows $D(\underline{\nu}')$ as a function of spatial frequency for the tomocamera system and the positron camera system, based on a regularly spaced 4-plane geometry. The $D(\underline{\nu}')$ of two pinhole camera systems are also shown in Fig.8-1 for comparison. They are based on a regularly spaced 3-plane geometry.

Because the slope of $D(\underline{\nu}')$ is also zero at $\underline{\nu}'=0$, the determinant function has small values near the origin, for instance, at the first harmonic of spatial frequency, $\underline{\nu}'$. Since the reconstruction $O_i(\underline{\nu}')$ has terms in it proportional to $T_j(\underline{\nu}')/D(\underline{\nu}')$, $j=1,\ldots, N_p$. $T_j(\underline{\nu}')$ depends on data from the camera and therefore has statistical fluctuations in it which are magnified by $1/D(\underline{\nu}')$, giving rise to incorrect values for $O_i(\underline{\nu}')$. These low frequency fluctuations have not been a problem so far for up to 5-plane reconstructions but they may turn out to be a limitation of this reconstruction method. One simple way which can give us estimations of the zero and low frequency components of an object distribution from its reconstructed images is described below. The inverse frequency matrix, H^{-1}, is undetermined at $\underline{\nu}'=0$ only, and

therefore, it is set to be zero at $\underline{v}'=0$. A computer reconstruction produces a set of images which are stored in N_p image matrices (for example, N_p 64 by 64 square matrices). These image matrices contain images with wrong zero spatial frequencies and perhaps a few low spatial frequency components. In most of the cases, it is known that the images of an object distribution are located only in the central portions of the matrices (for example, only in the central 48 by 48 area), the rest of the matrices should contain no events (or a few background events which in general are relatively small). The presence of large non-zero values and fluctuations in the marginal areas of the image matrices is clearly due to the undetermined zero frequency components and the noisy low frequency components. Least squares estimations of the zero and low frequency components can now be made in such a way that the estimated values give minimum background activity in the marginal areas of the image matrices. Experiments and computer simulations with $N_p=3$ and 5 have showed that the zero and first few frequency components of the object distribution can be estimated well enough this way.

8.2 Statistical Noise

The calculation of the output signal-to-noise ratio for the 3-D image reconstruction method is rather complicated because many transforms and a matrix inversion are involved. Since more than one projection is needed for the construction of tomographic images, the input statistical noise is not only a function of time and space but also a function of the projections. For the same counting interval, one

of the projections may receive more counts than some of the others from the same resolution element of the source object due to counting statistics. When the 3-D image reconstruction method is applied to a set of tomographic images, the spatial noise in these images is transformed and amplified at the intermediate stages of the reconstruction process. At low spatial frequencies the determinant function $D(\underline{v})$ is small (see Section 8.1), therefore, the amplification of noise is large. To suppress the amplification of noise, modifications of $D(\underline{v})$ may be needed at low spatial frequencies.

An attempt was made to relate the output signal-to-noise ratio of the 3-D image reconstruction process to the input parameters of the system. The calculations and the result are given in Appendix B. Although the result of Appendix B is derived from a multiple pinhole camera system, it can be applied to other similar tomographic imaging systems as well. The pinhole camera system uses N_h separated pinhole projections to form the tomographic images. The input signal-to-noise ratio of this system is given by, from Poisson statistics,

$$\left(\frac{S}{\sigma}\right)_{input} \approx \sqrt{o_i(\underline{r})} \quad , \tag{8.1}$$

where $o_i(\underline{r})$ is the total detected gamma-ray events from one resolution element of the object distribution at $\underline{r}$.

The output signal-to-noise ratio at image plane k is given by, from Appendix B,

$$\left(\frac{S}{\sigma}\right)_{output} \approx \sqrt{\frac{o_i(\underline{r})}{N_p C_k}} = \left(\frac{S}{\sigma}\right)_{input} / \sqrt{N_p C_k} \tag{8.2}$$

where the coefficient C_k is defined as

$$C_k \equiv \frac{1}{\pi} \int_{-\infty}^{\infty} \sum_{j=1}^{N_p} \left| H_{jk}^{-1} \left(\frac{u}{\pi \Delta^2 / L} \right) \right|^2 \frac{\sin u}{u} \, du \qquad (8.3)$$

and $H_{jk}^{-1}(u)$ is an element of the inverse matrix of $H(u)$ (see Appendix B). $\frac{\Delta^2}{L}$ is given as the area of one resolution element of the object divided by the maximum number of detected gamma-ray events from each resolution element. Since there is no simple formula for $H_{jk}^{-1}(u)$, C_k can not be further simplified. Eq.(8.3) indicates that C_k is a function of at least three parameters:the total number of projections N_h, the point blurring pattern $h_{ij}(\underline{r})$, and the imaging geometry (number of object planes and plane separations). The order of magnitude of C_k was estimated to be about 1.

Monte-Carlo calculations of the signal-to-noise ratio of the image reconstruction process have been made with a computer. The input object distribution was a uniformly distributed N_p x 25 x 25 point array with a lateral point separation given by the system resolution (0.5 cm). The camera received on an average of $400/N_h$ counts from one object point for each pinhole in the imaging pinhole array aperture. Since there were N_h pinholes in the array aperture, the input statistical noise was $1/\sqrt{400}=5\%$. The number of object planes was chosen to be 3 and 5 for this calculation. The object plane to aperture distances s_i, i=1, 2, ..., N_p, were given by s_1=20 cm, s_2=25 cm, s_3=30 cm, for N_p=3,

and s_1=20 cm, s_2=22.5 cm, s_3=25 cm, s_4=27.5 cm, s_5=30 cm, for N_p=5.

The output noise-to-signal ratio of the 3-D image reconstruction process was calculated as follows.

$N_p = 3$	object plane 1	output noise	7.1%	
	" 2	"	7.5%	average noise 7.4%
	" 3	"	7.6%	
$N_p = 5$	object plane 1	output noise	9.6%	
	" 2	"	9.2%	
	" 3	"	10.0%	average noise 9.3%
	" 4	"	7.3%	
	" 5	"	10.4%	

It is very difficult to relate the results of N_p=3 to the results of N_p=5 because their C_k are hard to calculate. If we let C_k=0.7 for all k in both cases, we have, from Eq.(8.2),

$$(\frac{\sigma}{S})_{output} \approx 5\% \sqrt{3 \times 0.7} = 7.25\%, \quad \text{for } N_p = 3$$

$$(\frac{\sigma}{S})_{output} \approx 5\% \sqrt{5 \times 0.7} = 9.35\%, \quad \text{for } N_p = 5$$

It is seen that the above values agree with the results from the computer simulations.

8.3 Spatial Invariance and Self-Attenuation

The basic assumption of the 3-D image reconstruction method is that the point response function, $h_{ij}(\underline{r})$, must be a space-invariant function, that is, $h_{ij}(\underline{r},\underline{r}') = h_{ij}(\underline{r}-\underline{r}')$ for any point source located at $\underline{r}'$ on object plane i, and $\underline{r}$ is a location on the tomographic plane j. Due to the inverse-square law and self-attenuation (mostly of tissue in organ imaging), the point response function is actually not a space-

invariant function, the expression $h_{ij}(\underline{r},\underline{r}') = h_{ij}(\underline{r}-\underline{r}')$ is only an approximation which works well when the imaging solid angle is small. When the imaging solid angle is very large, as shown in Fig.8-2, the point blurring pattern for a point source at $\underline{r}'$ = c(point 2) is quite different from that for a point source at $\underline{r}' = \underline{0}$ (point 1) because the photon attenuation of paths L_2 and L_2' is different from that of paths L_1 and L_1'.

The correction of the deviation due to the inverse-square law can be done at the same time when correction of the detector nonuniformity and pinhole transmission is made. Each pinhole image in the detector requires a correction matrix for making corrections. By moving a point source from point to point on the middle object plane in a constant time interval, a pinhole projection can be obtained, and the reciprocal of this projection (normalized to 1) is the desired correction matrix. The correction matrices made this way give exact correction only to those gamma-ray events from the middle object plane. For object planes other than the middle one, the correction is only an approximation.

When tomographic imaging devices such as the tomocamera and the positron camera are used, the inverse-square law does not give spatial variance problems in the imaging system. Since the tomocamera uses a slanted parallel-channel collimator, the response of the camera is the same for point sources at different locations (within the field of view) on a given object plane. In the positron camera case, if data selection is made according to Section 5.1, the response of the camera is also the same for point sources at different positions on an object plane,

although all the point responses fall off at larger values of $|\underline{r}-\underline{r}'|$ because of the inverse-square law.

The correction of self-attenuation due to the source (or the tissue, in organ imaging) itself is very hard to make because the source distribution is supposed to be unknown. Usually the effect of self-attenuation is neglected in focal-plane tomography for imaging of high energy gamma emitters since it is not serious as compared to that in transverse axial tomography. For imaging of low energy photon emitters, however, the effect of self-attenuation may not be neglected in focal-plane tomography. Additional effort is needed for the 3-D image reconstruction from focal-plane tomography.

We may apply an iterative least-squares technique to the reconstructed images of the 3-D image reconstruction method to improve the image quality. The iterative least-squares technique is used mostly in transverse axial tomography for x-ray transmission imaging as mentioned in Section 0.2. A detailed description of this technique is given in Ref.28. Using this iterative technique we improve the estimation of the object distribution $o_i(\underline{r})$ in a least-squares sense, with the reconstructed images (from the focal-plane tomograms) being the first set of estimated $o_i(\underline{r})$, i=1, 2, ..., N_p. The self-attenuation is taken into account by the weighting factors used in the iterative least-squares technique.

8.4 Object Distribution and Image Processing

To study further the 3-D image reconstruction method, a few test reconstructions were made by computer simulation. The purpose of these test reconstructions is to find out what may happen in more realistic situations, for instance, that the number of object planes N_p and the object plane to aperture distances s_i used in the image processing may differ from their true values by certain amounts. These test reconstructions may give us some idea of how sensitive the 3-D image reconstruction is to the imaging parameters.

A 7-plane object distribution was used to make the first test reconstruction. The original object distribution was in 7 regularly spaced planes located at s_1=7, s_2=8, s_3=9, s_4=10, s_5=11, s_6=12, s_7=13. The image processing for the 3-D reconstruction used N'_p=3 and s'_i=8,10,12. The pinhole camera system was used, and the pinhole projections of the 7-plane object were used to form three tomographic images at s'_i=8, 10,12. Thus, only three images were reconstructed from their tomographic images. Fig.8-3 shows pictures of the input object (a), one of the 9 pinhole projections of the object (b), the tomographic images (c), and the reconstructed images (d) of this computer simulation. If one examines the three reconstructed images carefully he will find that each reconstructed image consists of only the images of the object at planes on or near the assigned tomographic plane of that reconstructed image. That is, for instance, the last image of Fig.8-3d (located at s'=12) consists of only the image of the object at s_5=11, s_6=12, and s_7=13. This indicates that, aside from a little noise in the resulting images,

all the off-plane background images in the three tomograms of Fig.8-3c were removed by the 3-D image reconstruction method.

The same test reconstruction was performed again with another 7-plane object. This time the object was made of 7 small square sources located on 7 regularly spaced object planes just like the first test object. Each square source was assigned to a particular location on its plane in such a way that the normal projection of the sources on the detector contained no overlapping images. Fig.8-4a shows the relative positions of these small sources. After going through the same imaging and reconstruction process as in the first test, three tomographic images and three reconstructions were obtained, and they are shown in Fig.8-4b and c. It is seen that, aside from the noise, the 3-D image reconstruction method had clearly removed the off-plane background images from the three tomograms of Fig.8-4b. (Note: In these two test reconstructions, images of the object at $s_i'-1$, s_i', and $s_i'+1$ are considered as the in-focus images of the tomogram at s_i' and all other images are considered as the off-plane background images of that tomogram). The noise in the reconstructed images of Fig.8-4c is somewhat high, which is probably due to the sharp changes in source distribution of the object. A sharply varying distribution has wide frequency bandwidths and therefore requires small sample spacings for digital image processing according to the 3-D sampling theorem (to be described later).

Another test reconstruction was made with a correct N_p and incorrect s_i, $i=1,2,\ldots, N_p$. The test object was three numerical patterns (1, 2, and 3) located at $s_1=8$, $s_2=10$, $s_3=12$. The image

processing used object plane locations $s_i' = s_i + z$ to form the tomographic images and perform the 3-D image reconstruction, where z is a variable parameter used to make s_i' different from s_i. The results of this test are shown in Fig.8-5 with z = -0.5 in (a), z=0.5 in (b), z=1.0 in (c), and z=1.5 in (d). It is seen that false object plane locations for the image processing result in deformation in the reconstructed images. However, the images of the object can still be seen clearly when z is small.

Practically, all the object distributions for radionuclide imaging are continuous. The use of discrete distributions in the 3-D image reconstruction method is not only because of its simplicity, the requirement of digital image processing is also a reason. An object distribution imaged by a finite resolution system is bandlimited by the system's frequency pass-band, and therefore it is considered as a bandlimited function. The 3-D sampling theorem says that exact recovery of a bandlimited function can be achieved from an appropriately spaced 3-D array of its sample values (the maximum x, y, and z spacings of the sample lattice are $\frac{1}{2W_x}$, $\frac{1}{2W_y}$, and $\frac{1}{2W_z}$ respectively, where $2W_x$, $2W_y$, and $2W_z$ represent the bandwidths of its frequency spectrum). Therefore, it is all right to use discrete distributions in the 3-D image reconstruction method, as long as enough samples of the distributions are obtained.

In focal-plane tomography, the samples of an object distribution are obtained by taking projections of the object from one side only. The sampling in the x and y directions may be appropriate this way if many projections of the object are taken at different angles ($<90^o$ from

the z axis). However, it is likely that the sampling in the z direction is not appropriate because the imaging system can not "see" the depth of the object well this way. This relates somehow to the fact that the determinant functions $D(\underline{v})$ of all focal-plane tomographic imaging systems are small and zero at low spatial frequencies (see Section 8.1). It may be possible to make a relatively uniform $D(\underline{v})$ if we modify the point response functions $h_{ij}(\underline{r})$ of the tomographic imaging system with a few projections of the object distribution obtained at 90^{o} angles from the z axis. A study on this is needed in the future.

8.5 Conclusions

The principle of operation of the 3-D image reconstruction method studied in Part II is based on Fourier transform deconvolution, which is in turn based on a set of linear combinations of convolution integrals. Therefore, the basic assumption of this method is that the imaging system is linear and space-invariant. This image reconstruction method can be applied to all tomographic imaging systems which perform focal-plane tomography, as long as they are linear and space-invariant.

On the basis of spatial invariance, the tomocamera system is better than the other two camera systems described in Section 5.1, for the application of this image reconstruction method (because the response of the collimator used by a tomocamera is roughly constant for all depths). The positron camera system with its high detection efficiency and good point blurring characteristic offers a strong potential application of this method. However, in order to apply this method to

a positron camera system, some effort is needed to make the system space -invariant (described in Section 5.1). As discussed in Section 8.3, there is no simple way to make the multiple pinhole camera system exactly space-invariant. Therefore, application of the 3-D image reconstruction method to this camera system is limited.

It was shown in Section 7.2 that the image reconstruction of a 3-plane object takes very little computing time. For an object distribution with 5 x 64 x 64 resolution elements the total computing time required for the image reconstruction with a large computer (CDC-7600) is about one quarter of a second (5 FFT's, complex additions and multiplications, and 5 inverse FFT's). The same image reconstruction may also be performed by a small computer together with an associated disc system. Due to constraints of the core memory size and the speed for performing complex arithmetic, a small computer system usually requires more time to do a reconstruction. However, as explained in Section 7.2, this increase in computing time is still much less than the time required to accumulate the gamma-ray events and construct the tomographic images.

Although there has not yet been an attempt to examine this image reconstruction method under clinical situations, results from experiments and computer simulations indicate that better image contrast and fewer background artifacts over the original focal-plane tomographic images can be achieved without too much effort and computing time. Future studies of this method based on theorems of sampling and projection are necessary for the improvement of image reconstruction.

APPENDIX A.

Background and Statistical Noise in Non-Redundant Pinhole Array Coded Aperture Imaging

The discussions given below are for the coded aperture imaging systems in which a non-redundant pinhole array (NRPA) coded aperture is used and the reconstruction is done using the correlation method of Section 2.2. Because of its simplicity, an incoherent optical system is often used to perform the reconstruction. A detailed description of this type of coded aperture imaging is given in Section 3.2. Fig. A-1 shows a schematic diagram of the imaging and reconstruction of a point source with a 3-hole coded aperture. The pattern on the image plane is given by the autocorrelation function of the pinhole array used because the mask for decoding is a replica of that pinhole array. (see Eq.(2.3))

The image of a single point source using an N-hole coded aperture and incoherent optical reconstruction has a central intensity N corresponding to the matching of the coded shadow of the i^{th} aperture hole with the i^{th} hole in the mask for each of the N holes. Surrounding this image is background which is produced when the coded shadow of the i^{th} aperture hole combines with the j^{th} mask hole ($j \neq i$). The total amount of background for a point source is thus seen to be N(N-1) points, each of intensity 1 as in the non-redundant pinhole array, or over correspondingly fewer points of higher intensity, as for the redundant array (Fig.3-6b).

Background and statistical noise for NRPA coded aperture imaging

can be calculated with the following assumptions :

(1) The object lies in a single transverse plane and consists of M point sources of equal intensities C_o where C_o is the number of detected gamma ray events which pass through one of the N aperture holes. The point sources are located on an equilateral triangular grid with separation given by the resolution length Δ_x (Eq.(2.21)). A triangular grid is used because the non-redundant pinhole arrays are composed on such a grid to make more compact.

(2) The autocorrelation function is approximated by a spike of intensity N surrounded by a dise of radius r_a and intensity 1. r_a is chosen so as to contain the same number of background points, N(N-1) as does the exact autocorrelation function but packed tightly on the triangular grid. Thus, the approximate function is just a compact version of the exact function.

(3) The pinhole diameters are equal to the minimum spacing between holes in the array. This is the least resolution usable with a given non-redundant pinhole array.

The image obtained in the double process of shadowgram exposure and incoherent optical reconstruction is equivalent to the single step $i(\underline{r}) = o(\underline{r}) \circledast a(\underline{r})$. In order to provide a common scale we express the convolution in the detector plane (Fig.A-1) by projecting the object through a point in the center of the aperture plane onto the detector planes. This has been done in Fig.A-2. Assumption 3 assures that the spacing of the object points (Fig.A-2b) is the same as the spacing of the autocorrelation points (Fig.A-2a). The convolution is performed by centering the autocorrelation function on each object point and summing

the results over the M object points. The signal at the center point 0 of the object has a value NC_o. This signal is superimposed on a background B at point 0 due to the overlapping of the autocorrelation functions from neighboring object points. Those neighboring points, such as point 1 in Fig.A-2b which are closer than the distance r_a, each contribute $1C_o$ to the background intensity at 0. Those which are farther away than r_a, such as point 2, contribute nothing to the background at 0 since the autocorrelation function is assumed zero outside r_a. Since a maximum of N(N-1) points occur in the autocorrelation function and therefore lie within the distance r_a of the central point we have the following equation for the intensity of background at the center of the image depending on whether the actual number of object points M is greater than N(N-1) or not.

$$\begin{aligned} B&=MC_o \quad \text{if } M \leq N(N-1) \\ B&=N(N-1)C_o \quad \text{if } M > N(N-1) \end{aligned} \tag{A.1}$$

The background is maximum in the center of the image of the M point sources and falls to some finite value on its edges.

Noise σ in the image is due to statistical fluctuations in the number C_o of gamma-ray events and comes from independent fluctuations in the signal and in the background. Thus we have

$$\begin{aligned} \sigma &= \sqrt{N+M}\,\sqrt{C_o} \quad \text{if } M \leq N(N-1) \\ \sigma &= N\sqrt{C_o} \quad \text{if } M > N(N-1) \end{aligned} \tag{A.2}$$

we obtain the following values for the ratio of signal to background and of signal to noise at the center of the image

$$\frac{S}{B} \approx \frac{N}{M} \quad \text{and} \quad \frac{S}{\sigma} \approx \frac{N\sqrt{C_o}}{\sqrt{N+M}} , \qquad \text{if } M \leq N(N-1)$$

$$\frac{S}{B} \approx \frac{1}{N-1} \quad \text{and} \quad \frac{S}{\sigma} \approx \sqrt{C_o} , \qquad \text{if } M > N(N-1) \tag{A.3}$$

APPENDIX B.

Statistical Noise in Three-Dimensional Image Reconstruction from Focal-Plane Tomography

The multiple pinhole camera system described in Section 5.1 will be used here for the calculations of the statistical noise in the 3-D image reconstruction process described in Section 6.2. The pinhole camera system uses N_h separated pinhole projections to form the tomographic images. We assume the activities of the source object are in N_p separated object planes and are represented by the functions $o_i(\underline{r})$, $i=1, 2, \ldots, N_p$. Function $o_i^{\lambda}(\underline{r})$ is used to represent the events detected by the detector from a resolution element at $\underline{r}$ on the object plane i through the λth imaging pinhole aperture. The point response function, $h_{ij}^{\lambda}(\underline{r}-\underline{r}')$, of the tomographic imaging system is defined as the response of the system at $\underline{r}$ on the tomographic plane j from a point source at $\underline{r}'$ on the object plane i through the λth pinhole aperture. The normalized system point response function, $h_{ij}(\underline{r}-\underline{r}')$, is defined as $\frac{1}{N_h}\sum_{\lambda=1}^{N_h} h_{ij}^{\lambda}(\underline{r}-\underline{r}')$. To find an expression for the input signal and noise, we let the true mean of $o_i^{\lambda}(\underline{r})$ be given as $m_i(\underline{r})/N_h$ and the statistical fluctuation of

$o_i^{\lambda}(\underline{r})$ from its mean be given as $n_i^{\lambda}(\underline{r})$. The detected object distribution *from one pinhole projection*, $o_i^{\lambda}(\underline{r})$, is then given by

$$o_i^{\lambda}(\underline{r}) = m_i(\underline{r})/N_h + n_i^{\lambda}(\underline{r}) \tag{B.1}$$

The RMS value of the input noise, which is given as the standard deviation of the input signal, is calculated by Poisson statistics as

$$\langle n_i(\underline{r})^2 \rangle^{1/2} = \sqrt{E(n_i^{\lambda 2})} = \sqrt{m_i(\underline{r})/N_h} \tag{B.2}$$

The input signal-to-noise ratio to this pinhole camera system is then given by

$$\left(\frac{S}{\sigma}\right)_{\text{input}} = \frac{\sum_{\lambda=1}^{N_h} m_i(\underline{r})/N_h}{\sqrt{\sum_{\lambda=1}^{N_h} \langle n_i^{\lambda}(\underline{r})^2 \rangle}} = \frac{m_i(\underline{r})}{\sqrt{\sum_{\lambda=1}^{N_h} m_i(\underline{r})/N_h}} = \sqrt{m_i(\underline{r})} \approx \sqrt{o_i(\underline{r})} \tag{B.3}$$

where $o_i(\underline{r}) = \sum_{\lambda=1}^{N_h} o_i^{\lambda}(\underline{r})$ is the total detected object distribution (while $m_i(\underline{r})$ is the hypothetic true distribution).

The output signal-to-noise ratio is calculated as follows: The tomographic images $t_j(\underline{r})$, formed by the back-projection method, are given by

$$t_j(\underline{r}) = \sum_{i=1}^{N_p} \left[\sum_{\lambda=1}^{N_h} o_i^{\lambda}(\underline{r}) \circledast h_{ij}^{\lambda}(\underline{r}) \right] \tag{B.4}$$

Substituting $o_i^{\lambda}(\underline{r})$ by Eq.(B.1), we have

$$\begin{aligned} t_j(\underline{r}) &= \sum_{i=1}^{N_p} \left[\sum_{\lambda=1}^{N_h} \frac{m_i(\underline{r})}{N_h} \circledast h_{ij}^{\lambda}(\underline{r}) + \sum_{\lambda=1}^{N_h} n_i^{\lambda}(\underline{r}) \circledast h_{ij}^{\lambda}(\underline{r}) \right] \\ &= \sum_{i=1}^{N_p} m_i(\underline{r}) \circledast h_{ij}(\underline{r}) + \sum_{i=1}^{N_p} \sum_{\lambda=1}^{N_h} n_i^{\lambda}(\underline{r}) \circledast h_{ij}^{\lambda}(\underline{r}) \end{aligned} \tag{B.5}$$

Taking Fourier transforms to both sides of Eq.(B.5), we have

$$T_j(\underline{\nu}) = \sum_{i=1}^{N_p} M_i(\underline{\nu})H_{ij}(\underline{\nu}) + \sum_{i=1}^{N_p}\sum_{\lambda=1}^{N_k} N_i^{\lambda}(\underline{\nu})H_{ij}^{\lambda}(\underline{\nu}) \qquad (B.6)$$

As was given in Section 6.2, an inverse frequency matrix $H^{-1}(\underline{\nu})$ with $\sum_{j=1}^{N_p} H_{ij}(\underline{\nu})H_{jk}^{-1}(\underline{\nu}) = \delta_{ik}$ is required to reconstruct the object distribution. If we ignore the fact that $H^{-1}(\underline{0})$ doesn't exist, we may multiply both sides of Eq.(B.6) by $H_{jk}^{-1}(\underline{\nu})$ and sum them over all j. Thus, we have the following expression for the frequency components of the reconstruction.

$$\sum_{j=1}^{N_p} T_j(\underline{\nu})H_{jk}^{-1}(\underline{\nu}) = \sum_{i=1}^{N_p} M_i(\underline{\nu}) \sum_{j=1}^{N_p} H_{ij}(\underline{\nu})H_{jk}^{-1}(\underline{\nu}) + \sum_{i=1}^{N_p}\sum_{j=1}^{N_p}\sum_{\lambda=1}^{N_k} N_i^{\lambda}(\underline{\nu})H_{ij}^{\lambda}(\underline{\nu})H_{jk}^{-1}(\underline{\nu})$$

$$= M_k(\underline{\nu}) + \sum_{i=1}^{N_r}\sum_{j=1}^{N_p}\sum_{\lambda=1}^{N_k} N_i^{\lambda}(\underline{\nu})H_{ij}^{\lambda}(\underline{\nu})H_{jk}^{-1} \qquad (B.7)$$

Upon taking the inverse Fourier transform of Eq.(B.7), we have

$$\mathcal{F}^{-1}\left\{ \sum_{j=1}^{N_p} T_j(\underline{\nu})H_{jk}^{-1}(\underline{\nu}) \right\} = m_k(\underline{r}) + \sum_{i=1}^{N_p}\sum_{j=1}^{N_p}\sum_{\lambda=1}^{N_k} n_i^{\lambda}(\underline{r}) \circledast h_{ij}^{\lambda}(\underline{r}) \circledast h_{jk}^{-1}(\underline{r}) \qquad (B.8)$$

The first term of the right side of Eq.(B.8) is the output signal, and the last term which contains $n_i^{\lambda}(\underline{r})$ is clearly the output noise. Let the output noise be represented by

$$g_k(\underline{r}) = \sum_{i=1}^{N}\sum_{j=1}^{N_p}\sum_{\lambda=1}^{N_k} n_i^{\lambda}(\underline{r}) \circledast h_{ij}^{\lambda}(\underline{r}) \circledast h_{jk}^{-1}(\underline{r}) \qquad (B.9)$$

The calculation of the RMS value of the output noise, $\langle g_k(\underline{r})^2 \rangle^{1/2}$, requires an understanding of how the noise is transfered through the system. It is necessary to define a few quantities for the random process which gives rise to the noise. First we define a stochastic process $\{ Z_t , t \in T \}$ as a family of random variables indexed by points t in a subset T of the real line. We define the covariance

function of Z_t as $E((Z_t-m_t)(Z_s-m_s)) = R(t,s)$, where $m_t = E(Z_t)$. A process $\{Z_t, -\infty<t<\infty\}$ is called widesense stationary if $E(Z_t)=m$ (independent of t) and $R(t,s) = R(t-s)$. The spectral density function of a widesense stationary process $\{Z_t, -\infty < t <\infty\}$ is defined as

$$S(\nu) = \mathcal{F}\{R(t)\} = \int_{-\infty}^{\infty} e^{-i2\pi\nu t} R(t)\, dt, \tag{B.10}$$

where

$$R(t) = E((Z_{t+t_o} - m)(Z_{t_o} - m)) = \mathcal{F}^{-1}\{S(\nu)\} \tag{B.11}$$

If a widesense stationary process $\{Z_t, -\infty < t < \infty\}$ is the input to a linear and time invariant system with transfer function $H(\nu)$, then the output process $\{Y_t = Z_t \circledast h, -\infty < t < \infty\}$ is also widesense stationary and the spectral density function of the output process is given by

$$S_Y(\nu) = |H(\nu)|^2 S_Z(\nu) \tag{B.12}$$

The proof of Eq.(B.12) can be found in many textbooks of system analysis.

In order to change the 2D spatial coordinates into 1D real numbers, we define a one-to-one mapping which transforms a point, $\underline{r}=(x,y)$ into an area, t, given by

$$t(\underline{r}) \equiv (x-x_{min})(y-y_{min}) + (y-\Delta-y_{min})(x_{max}-x) \tag{B.13}$$

Where x_{min}, x_{max}, y_{min}, and y_{max} give the boundary of the object distribution, and Δ is the width of a resolution element of the object. For the purpose of evaluating the statistical noise, we assume the input object distribution is uniform so that the noise is not affected by the distribution of the object. We let a random process $\{Z_{t(\underline{r})}, 0<t(\underline{r})<x_{max}y_{max}\}$ be the total number of events detected from the resolution elements of the object in area $t(\underline{r})$. Further more, Z_t is a

Poisson process with parameter $\bar{m}t/N_h$. The derivative of Z_t, $\dot{Z}_t$ (events per unit area), can be used to describe the input signal, $o_i^{\lambda}(\underline{r})$. Fig. B-1. shows an example of the distributions of Z_t and $\dot{Z}_t$. Note that $\dot{Z}_t$ is actually a series of delta functions.

We can now relate Z_t to $o_i^{\lambda}(\underline{r})$ by

$$o_i^{\lambda}(\underline{r}) = Z_{t(\underline{r})+\Delta^2} - Z_{t(\underline{r})} = \int_{t(\underline{r})}^{t(\underline{r})+\Delta^2} \dot{Z}_t \, dt = \int_{t(\underline{r})}^{t(\underline{r})+\Delta^2} \sum_k \delta(t-t_k) dt \tag{B.14}$$

To check the above expression, we evaluate the mean function of $o_i^{\lambda}(\underline{r})$ as follows :

$$E(o_i^{\lambda}(\underline{r})) = E(Z_{t+\Delta^2} - Z_t) = \frac{\bar{m}}{N_h}(t+\Delta^2) - \frac{\bar{m}}{N_h}t = \frac{\bar{m}}{N_h}\Delta^2 \tag{B.15}$$

From Eq.(8.1), we have

$$E(o_i^{\lambda}(\underline{r})) = \frac{m_i(\underline{r})}{N_h} \tag{B.16}$$

Since the object distribution is assumed to be uniform, $m_i(\underline{r})$ becomes a constant, m_i, and is given by $m_i = \bar{m}\Delta^2$. It can be shown that the covariance function of Z_t is given by $R_Z(t,s) = \frac{\bar{m}}{2N_h}(t+s-|t-s|)$ and the mean and the covariance function of $\dot{Z}_t$ are given by $E(\dot{Z}_t) = \frac{\bar{m}}{N_h}$ and $R_{\dot{Z}}(t,s) = \frac{\bar{m}}{N_h}\delta(t-s) = R_{\dot{Z}}(t-s)$, for $0 < t,s < x_{max}y_{max}$.

Note that the random process $\dot{Z}_t$ is widesense stationary if the range of t is extended to the entire real line. If we define a process, n_t, by $n_t = \dot{Z}_t - E(\dot{Z}_t) = \dot{Z}_t - \frac{\bar{m}}{N_h}$, then n_t is also widesense stationary because $E(n_t) = 0$ and $R_n(t,s) = R_{\dot{Z}}(t,s) = \frac{\bar{m}}{N_h}\delta(t-s) = R_n(t-s)$.

The relationship between n_t and the noise term, $n_i^{\lambda}(\underline{r})$ in Eq.(B.1), is as the following :

$$n_i^{\lambda}(\underline{r}) = o_i^{\lambda}(\underline{r}) - \frac{m_i}{N_h} = \int_{t(\underline{r})}^{t(\underline{r})+\Delta^2} (\dot{Z}_t - \frac{1}{\Delta^2}\frac{m_i}{N_h})\, dt = \int_{t(\underline{r})}^{t(\underline{r})+\Delta^2} n_t\, dt \quad (B.17)$$

If the original delta function representation of $\dot{Z}_t$ is replaced by a more realistic square-pulse function, p(t) with width $\frac{\Delta^2}{L}$ and magnitude $\frac{L}{\Delta^2}$ (where L is a large integer, large enough to prevent the occurrence of more than one event within $\frac{\Delta^2}{L}$), Eq.(B.17) can then be written in the following form.

$$n_i^{\lambda}(\underline{r}) = \sum_{\ell=1}^{L} (n_{t+\ell\frac{\Delta^2}{L}}) \frac{\Delta^2}{L} \quad (B.18)$$

Making a change of variable and substituting $n_i^{\lambda}(\underline{r})$ by Eq.(B.18), the output noise g_k given by Eq.(B.9) becomes

$$g_k = \frac{\Delta^2}{L} \sum_{\ell=1}^{L} \sum_{i=1}^{N^p} \sum_{j=1}^{N^p} \sum_{\lambda=1}^{N^h} n_{t+\ell\frac{\Delta^2}{L}} \circledast h_{ij}^{\lambda}(t) \circledast h_{jk}^{-1}(t) \quad (B.19)$$

The term $n_t \circledast h_{ij}^{\lambda}(t) \circledast h_{jk}^{-1}(t)$ in Eq.(B.19) may be viewed as the output process of a linear system with impulse response $h_{ij}^{\lambda}(t) \circledast h_{jk}^{-1}(t)$ to an input random process n_t. As an approximation, the random process n_t is still widesense stationary and the spectral density function of n_t is given by

$$S_n(\nu) = \int_{-\infty}^{\infty} e^{-i2\pi\nu t} R_n(t)\, dt = \int_{-\infty}^{\infty} e^{-i2\pi\nu t} \frac{\tilde{m}}{N_h} \delta(t)\, dt = \frac{\tilde{m}}{N_h} \quad (B.20)$$

By the previous assumption, the delta function $\delta(t)$ in Eq.(B.20) should be replaced by the square-pulse function p(t) and the actual spectral density function of n_t is given by

$$S_n(\nu) = \int_{-\infty}^{\infty} e^{-i2\pi\nu t} p(t)\, dt = \int_{-\Delta^2/2L}^{\Delta^2/2L} e^{-i2\pi\nu t} \frac{\tilde{m}}{N_h} (\frac{L}{\Delta^2})\, dt$$

$$S_n(\nu) = \frac{\tilde{m}}{N_h}\left(\frac{\sin(\pi\nu\Delta^2/L)}{\pi\nu\Delta^2/L}\right) \qquad (B.21)$$

Note that Eq.(B.21) reduces to Eq.(B.20) as $L \to \infty$ because

$$\lim_{L\to\infty}\left(\frac{\sin\pi\nu\Delta^2/L}{\pi\nu\Delta^2/L}\right) = 1.$$

From Eq.(B.12) and $\mathcal{F}\left\{h_{ij}^{\lambda}(t) \circledast h_{jk}^{-1}(t)\right\} = H_{ij}^{\lambda}(\nu)\, H_{jk}^{-1}(\nu)$, the spectral density function of $n_t \circledast h_{ij}^{\lambda}(t) \circledast h_{jk}^{-1}(t)$ is given by

$$S_{ijk}^{\lambda}(\nu) = \left| H_{ij}^{\lambda}(\nu)\, H_{jk}^{-1}(\nu) \right|^2 S_n(\nu) \qquad (B.22)$$

Ideally, $h_{ij}^{\lambda}(\underline{r})$ is a delta function centered at a point $\underline{r}$. The one-to-one mapping of $h_{ij}^{\lambda}(\underline{r})$, $h_{ij}^{\lambda}(t)$, is also a delta function and therefore,

$$\left| H_{ij}^{\lambda}(\nu) \right| = \left| \int_{-\infty}^{\infty} e^{-i2\pi\nu t}\, \delta(t-t_o)\, dt \right| = \left| e^{-i2\pi\nu t_o} \right| = 1 \qquad (B.23)$$

Eq.(B.22) can then be written as

$$S_{ijk}^{\lambda}(\nu) = S_{jk}(\nu) = \left| H_{jk}^{-1}(\nu) \right|^2 S_n(\nu) = \left| H_{jk}^{-1}(\nu) \right|^2 \frac{\tilde{m}}{N_h}\left(\frac{\sin\pi\nu\Delta^2/L}{\pi\nu\Delta^2/L}\right) \qquad (B.24)$$

We defind a random process $f_k(t)$ by

$$f_k(t) = \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \sum_{\lambda=1}^{N_h} n_t \circledast h_{ij}^{\lambda}(t) \circledast h_{jk}^{-1}(t) \qquad (B.25)$$

The spectral density of the random process $f_k(t)$ is calculated as follows :

$$S_f(\nu) = \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \sum_{\lambda=1}^{N_h} S_{ijk}^{\lambda}(\nu) = N_p\, N_h \sum_{j=1}^{N_p} S_{jk}(\nu)$$

$$S_f(\nu) = N_p \tilde{m} \sum_{j=1}^{N_p} \left| H_{jk}^{-1}(\nu) \right|^2 \left(\frac{\sin \pi \nu \Delta^2/L}{\pi \nu \Delta^2/L} \right) \tag{B.26}$$

The variance or mean-square value of f_k can be obtained as follows :

$$\left\langle f_k(t)^2 \right\rangle = E\left(f_k(t) - Ef_k(t)\right)^2 = R_f(o) = \int_{-\infty}^{\infty} S_f(\nu)\, d\nu$$

$$\left\langle f_k(t)^2 \right\rangle = N_p \tilde{m} \int_{-\infty}^{\infty} \sum_{j=1}^{N_p} \left| H_{jk}^{-1}(\nu) \right|^2 \frac{\sin \pi \nu \Delta^2/L}{\pi \nu \Delta^2/L}\, d\nu$$

$$\left\langle f_k(t)^2 \right\rangle = N_p \tilde{m} \frac{L}{\Delta^2} C_k \tag{B.27}$$

Where the coefficient C_k is defined as

$$C_k \equiv \frac{1}{\pi} \int_{-\infty}^{\infty} \sum_{j=1}^{N_p} \left| H_{jk}^{-1}\left(\frac{u}{\pi \Delta^2/L}\right) \right|^2 \frac{\sin u}{u}\, du \tag{B.28}$$

Since there is no simple formula for $H_{jk}^{-1}(\nu)$ and the computation of C_k is rather complex, no attem has been made to relate the numerical values of C_k to its parameters such as the total number of projections, the imaging pinhole distribution, and the imaging geometry. The order of magnitude of C_k was estimated to be one, and the value of C_k varies with k, the image plane.

From Eq.(B.19) and Eq.(B.25), we have, for the output noise,

$$g_k = \frac{\Delta^2}{L} \sum_{\ell=1}^{L} f_k(t+\ell \frac{\Delta^2}{L}) \tag{B.29}$$

The variance of g_k is given by

$$\left\langle g_k^2 \right\rangle = E(g_k - Eg_k)^2 = \frac{\Delta^4}{L^2} \sum_{\ell=1}^{L} \sum_{\ell'=1}^{L} E\, f_k(t+\ell \frac{\Delta^2}{L})\, f_k(t+\ell' \frac{\Delta^2}{L}) \tag{B.30}$$

Since the random processes $\left\{ n_{t+\ell\Delta^2/L} \right\}$ and $\left\{ n_{t+\ell'\Delta^2/L} \right\}$ are of zero mean

and independent if $\ell \neq \ell'$, we may at least assume that $f_k(t+\ell\frac{\Delta^2}{L})$ and $f_k(t+\ell'\frac{\Delta^2}{L})$ are uncorrelated if $\ell \neq \ell'$, that is,

$$E\, f_k(t+\ell\frac{\Delta^2}{L})\, f_k(t+\ell'\frac{\Delta^2}{L}) = 0 \qquad \text{if } \ell \neq \ell'$$

The variance of ε_k then becomes

$$\langle \varepsilon_k^2 \rangle = \frac{\Delta^4}{L^2} \sum_{\ell=1}^{L} \left\langle f_k(t+\ell\frac{\Delta^2}{L})^2 \right\rangle = \Delta^2 N_p \tilde{m} C_k = N_p m_i C_k \qquad (B.31)$$

The RMS value of the output noise is given by

$$\langle \varepsilon_k^2 \rangle^{1/2} = \sqrt{N_p m_i C_k} \qquad (B.32)$$

The output signal-to-noise ratio at image plane k is given by

$$\left(\frac{S}{\sigma}\right)_{\text{output}} = \frac{m_i}{\langle \varepsilon_k^2 \rangle^{1/2}} = \sqrt{\frac{m_i}{N_p C_k}} \approx \sqrt{\frac{o_i(\underline{r})}{N_p C_k}} \qquad (B.33)$$

ACKNOWLEDGMENTS

This work was supported by the U.S. Atomic Energy Commission through the Lawrence Berkeley Laboratory. I wish to express my appreciation to Dr. Victor Perez-Mendez, Group Leader of the Physics Division, for providing the research environment and supervising the progress of this work.

A special thanks goes to Prof. Selig Kaplan of the Nuclear Engineering Department for his guidance and assistance in all phases of the degree program.

I am indebted to Dr. Burns Macdonald who was very helpful in the preparation of this work and was invaluable in critically reviewing the entire text.

Finally, I wish to express my sincere gratitude to my wife, Wei-Ching, for her continuing moral support and for her assistance in typing the final draft.

REFERENCES

(1) H.N. Wagner (ed.), "Principle of Nuclear Medicine", (W.B. Saunders Co., 1968).

(2) W.H. Blahd, "Nuclear Medicine", (McGraw-Hill Book Co., 2nd ed., 1971).

(3) E.R. Powsner, D.E. Raeside, "Diagnostic Nuclear Medicine", (Grune & Stratton, Inc., 1971).

(4) C.D. Maynard, "Clinical Nuclear Medicine", (Lea & Febiger, 1969).

(5) F. Joliot and I. Curie, Artifical Production of a New Kind of Radio-Elements, Nature, p.201, 1934.

(6) G.J. Hine and J.J. Erickson, Advance in Scintigraphic Instruments, Chapter 1 in "Instrumentation in Nuclear Medicine", (G.J. Hine & J.A. Sorenson ed., Academic Press, Inc., 1974), v.2, p.2.

(7) H.O. Anger, Radioisotope Cameras, Chapter 19 in "Instrumentation in Nuclear Medicine", v.1, p.486, 1967.

(8) G.S. Freedman, Tomography with a Gamma Camera - Theory, Journal of Nuclear Medicine, v.11, p.602, 1970.

(9) G. Muehllehner, A Tomographic Scintillation Camera, Phys. Med. Biol., v.16, p.87, 1971.

(10) S. Rudin, P.A. Bardfeld, and H.E. Hart, Use of Magnifying Multi-hole Collimators in the Gamma-Ray Camera System, J. Nucl. Med., v.12, p.831, 1971.

(11) H.H. Barrett, Fresnel Zone Plate Imaging in Nuclear Medicine, J. Nucl. Med., v.13, n.6, p.382, 1972.

(12) W.L. Rogers, K.S. Han, and L.W. Jones, et al., Application of a

Fresnel Zone Plate to Gamma-Ray Imaging, J. Nucl. Med., v.13, p. 612, 1972.

(13) Mallard et al., Iis et al., and Thomas et al., on Image Intensifier Gamma Cameras, in "Medical Radioisotope Scintigraphy" (Inter. Atomic Energy Agency, Vienna), v.1, p.3, p.17,p.43, 1969.

(14) N.H. Horwitz, J.E. Lofstrom, and A.L. Forsaith, A Comparison of Clinical Results Obtained with a Spark Imaging Camera and a Conventional Scintillation Scanner, Radiology, v.86, p.830, 1966.

(15) L. Kaufman, V. Perez-Mendez, J.M. Sperinde, and G. Stoker, Multi-wire Proportional Chamber for Low Dose X-Radiography, Amer. J. Roentgenol., Radium Ther. Nucl. Med., v.113, p.378, 1971.

(16) A.D. Beach, Aperture-Indexed Collimators for Gamma Cameras with Semi-Conductor Detector Arrays, J. Phys., E2, p.378, 1969.

(17) M. Brucer, Radioisotope Scanning, ORINS-20, 1958.

(18) G. L. Brownell, Theory of Radioisotope Scanning, Inter. J. Appl. Radiation Isotopes, v.3, p.181, 1958.

(19) B. Cassen, Section Scanning with a Large Solid Angle Collimator, J. Nucl. Med. v.6, p.767, 1965.

(20) H.O. Anger and D.J. Rosenthal, Scintillation Camera and Positron Camera, in "Medical Radioisotope Scanning", IAEA, Vienna, 1959.

(21) H.O. Anger, Gamma-Ray and Positron Scintillation Camera, Nucleonics, v.21, n.10, p.56, 1963.

(22) H.O. Anger, Tomographic Gamma-Ray Scanner with Simultaneous Read-out of Several Planes, Lawrence Radiation Laboratory Report, UCRL-16899, 1966.

(23) G.S. Freedman (ed.), "Tomographic Imaging in Nuclear Medicine", (The Society of Nuclear Medicine, Inc., New York, 1973).

(24) R. Gordon, A Tutorial on Algebraic Reconstruction Techniques, IEEE Trans. on Nucl. Sci., v.NS-21, n.3, p.78, 1974.

(25) B.E. Oppenheim, More Accurate Algorithms for Iterative Three-Dimensional Reconstruction, IEEE Trans. on Nucl. Sci., v.NS-21, n.3, p.72, 1974.

(26) L.A. Shepp and B.F. Logan, The Fourier Reconstruction of a Head Section, IEEE Trans. on Nucl. Sci., v.NS-21, n.3, p.21, 1974.

(27) Z.H. Cho, General Views on 3-D Image Reconstruction and Computerized Transverse Axial Tomography, IEEE Trans. on Nucl. Sci., v.NS-21, n.3, p.44, 1974.

(28) T.F. Budinger and G.T. Gullberg, Three-Dimensional Reconstruction in Nuclear Medicine Emission Imaging, IEEE Trans. on Nucl. Sci., v.NS-21, n.3, p.2, 1974.

(29) R.H. Dicke, Scatter-Hole Cameras for X-Rays and Gamma-Rays, Astrophysical J., v.153, L101, Aug. 1968.

(30) R.L. Blake, A.J. Burek, E. Fenimore, and R. Puetter, Solar X-Ray Photography with Multiplex Pinhole Camera, Rev. Sci. Instrum., v.45, n.4, p.513, 1974.

(31) A.T. Moffet, Minimum-Redundancy Linear Arrays, IEEE Trans. on Antennas and Propagation, v.AP-16, n.2, p.172, Mar. 1968.

(32) M.J.E. Golay, Point Arrays Having Compact, Nonredundant, Autocorrelations, J. Opt. Soci. Amer., v.61, p.272, Feb. 1971.

(33) W.K. Klemperer, Non-Redundant Phased-Array Radar, IEEE Conference Publication # 105, "Radar, Present and Future", London, p.74,

Oct. 1973.

(34) A. Wouters, K.M. Simon, and J.G. Hirschberg, Direct Method of Decoding Multiple Images, Applied Optics, v.12, n.8, p.1871, 1973

(35) L. Mertz and N.O. Young, Fresnel Transformations of Images, Proc. Int. Conf. on Opt. Inst., London, p.305, 1961.

(36) H.H. Barrett and F.A. Horrigan, Fresnel Zone Plate Imaging of Gamma-Rays: Theory, Appl. Opt., v.12, p.2686, 1973.

(37) D.G. Grant, Tomosynthesis: A Three-Dimensional Radiographic Imaging Technique, IEEE Trans. on Biomed. Eng., v.19, p.20, 1972.

(38) M. Kock and U. Tiemens, Tomosynthesis: A Holographic Method for Variable Depth Display, Opt. Communications, v.7,n.3,p.260, 1973.

(39) H. Weiss, Nonredundant Point Distribution for Coded Aperture Imaging with Application to 3-D On-Line X-Ray Information Retrieving, IEEE Trans. on Computers, v.C-24, n.4, p.391, 1975.

(40) AE Group's Subcommittee on Measurement Concepts, What is the Fast Fourier Transform ?, IEEE Trans. Audio and Electroacoustics, v.AU -15, p.45, June 1967, also Proc. IEEE, v.55, n.10, p.1664, 1967.

(41) J.W. Cooley and J.W. Tukey, An Algorithm for the Machine Calculation of Complex Fourier Series, Math. of Comput., v.19, n.90, p.297, Apr. 1965.

(42) E.O. Brigham, "The Fast Fourier Transform", (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974).

(43) G. Charpak, R. Bouclier, T. Bressani, J. Favier, and C. Zupancic, Nucl. Inst. & Methods, Multiwire Proportional Chambers, v.62, p.262, 1968.

(44) L. Kaufman, V. Perez-Mendez, D. Shames, and G. Stoker, A Multiwire Proportional Chamber for Nuclear Medicine Applications, IEEE Trans. on Nucl. Sci., v.NS-19, p.169, June 1972.

(45) S.N. Kaplan, L. Kaufman, V. Perez-Mendez, and K. Valentine, Multiwire Proportional Chambers for Biomedical Applications, Nucl. Inst. and Methods, v.106, p.397, 1973.

(46) L. Kaufman, V. Perez-Mendez, and G. Stoker, Performance of a Pressurized Xenon-Filled Multiwire Proportional Chamber, IEEE Trans. on Nucl. Sci., v.NS-20, p.426, 1973.

(47) G. Groh, G.S. Hayat, and G.W. Stoke, X-Ray and Gamma-Ray Imaging with Multiple-Pinhole Camera Using a Posteriori Image Synthesis, Applied Optics, v.11, n.4, p.931, 1972.

(48) F.A. Jenkins and H.E. White, "Fundamentals of Optics", (McGraw-Hill Book Co., New York, 1957, 3rd ed.). p.360.

(49) A.R. Shulman, "Optical Data Processing", (Wiley & Sons, Inc., New York, 1970), p.387.

(50) J.T. Winthrop and C.R. Worthington, Convolution Formulation of Fresnel Diffraction, J. Opt. Soci. Am., v.56, n.5, p.588, 1964.

(51) J.W. Goodman, "Introduction to Fourier Optics", (McGraw-Hill Book Co., San Francisco, 1968), p.60.

(52) Y.L. Chow, Comparison of Some Correlation Array Configurations for Radio Astronomy, IEEE Trans. on Antennas and Propagation, p.567, July 1970.

(53) H.H. Barrett and G.D. DeMeester, Quantum Noise in Fresnel Zone Plate Imaging, Applied Optics, v.13, n.5, p.1100, 1974.

(54) A.G. Lindgren, O.K. Guha, J.E. Spence, A Noise Analysis of Fresnel Zone Plate Imaging Systems, Digest of Papers, 1974 Int. Optical Computing Conf., Apr. 9-11, Zurich, 1974.

(55) L.T. Chang, B. Macdonald, V. Perez-Mendez, and L. Shiraishi, Coded Aperture Imaging of Gamma-Rays Using Multiple Pinhole Arrays and Multiwire Proportional Chamber Detector, IEEE Nucl. Sci. Symp., Washington, D. C., Dec. 11-13, 1974.

(56) D.R. Dance, B.C. Wilson, and R.P. Parker, Digital Reconstruction of Point Sources Imaged by a Zone Plate Camera, Phys. Med. Biol., v.20, n.5, p.747, 1975.

(57) D.J. Stigliani, R. Mittra, and R.G. Semonin, Resolving Power of a Zone Plate, J. Opt. Soci. Amer., v.57, p.610, 1967.

(58) D.J. Sakrison,"Communication Theory: Transmission of Waveforms and Digital Information",(John Wiley & Sons, Inc., 1968).

(59) G.W. Stroke,"An Introduction to Coherent Optics and Holography", (Academic Press, New York, 2nd ed., 1969).

(60) A. Engel and G. Herziger, Computer-Drawn Modulated Zone Plates, Applied Optics, v.12, n.3, p.471, 1973.

(61) H.H. Barrett, D.T. Wilson, and G.D. DeMeester, Optics Communicat., v.5, p.398, 1972.

(62) G.S. Freedman, Digital Gamma Camera Tomography - Theory, in "Tomographic Imaging in Nuclear Medicine", (Society of Nuclear Medicine, New York, 1973, G.S. Freedman ed.), p.68.

(63) G. Muehllehner, Performance Parameters for a Tomographic Scintillation Camera, in "Tomographic Imaging in Nuclear Medicine", p.76, 1973.

(64) C.B. Lim, D. Chu, V. Perez-Mendez, L. Kaufman, R. Hattner, D.C. Price, S. Swann, and D. Palmer, The U.C.S.F. Multiwire Proportional Chamber Positron Camera I: Design and Technical Specifications, Digest of Papers, Topical Meeting on Image Processing for 2-D and 3-D Reconstruction from Projections, Aug.4-7, 1975, Stanford University, Stanford, California.

(65) L. Eriksson, J.K. Chan, and Z.H. Cho, Emission Tomography with a Circular-Ring Transverse-Axial Positron Camera, Digest of Papers, presented at the Topical Meeting at Stanford University, Calif., Aug. 4-7, 1975.

(66) H.H. Barrett, W.W. Stoner, D.T. Wilson, and G.D. DeMeester, Coded Apertures Derived from the Fresnel Zone Plate, Optical Engineer., v.13, n.6, p.539, 1974.

FIGURE CAPTIONS

1-1 A Multichannel Collimator Gamma Camera

The gamma-ray photons emitted by the source pass through the channels of the collimator and impinge on the detector. A dot image is formed through the electronic system of the camera and can be displayed on the screen of a CRT.

1-2 A Pinhole Aperture Gamma Camera

1-3 A Coded Aperture Imaging System

The single-pinhole aperture of Fig. 1-2 is replaced by an aperture plate containing 27 pinholes.

1-4 Examples of Coded Aperture Patterns

(a) A random pinhole array

(b) A non-redundant pinhole array

(c) A Fresnel zone plate

2-1 Examples of the Autocorrelation Functions of a Few Coded Aperture Patterns

The 2-D gray-shade plots and cross-sectional plots of the autocorrelation functions were generated by computer from the following patterns:

(a) A random pinhole array (50 holes)

(b) A non-redundant pinhole array (9 holes)

(c) A positive zone plate ($h_{zp}(\underline{r})$, 9 zones)

(d) A combined zone plate ($h_{zp}(\underline{r})-\tilde{h}_{zp}(\underline{r})$, 10 zones)

2-2 Examples of the Correlation Functions of Two Patterns

The autocorrelation functions (in-focus point reconstructions by the correlation method) and the crosscorrelation functions (out-of-focus point reconstructions, on planes 2 cm and 4 cm apart from the focal plane) were obtained from the following two patterns:

(a) A non-redundant pinhole array (27 holes)

(b) A combined zone plate ($h_{zp}(\underline{r}) \sim \tilde{h}_{zp}(\underline{r})$, 10 zones)

a-1 and b-1 are autocorrelation functions and the rest are cross-correlation functions. a-1, 2, and 3 were obtained optically and b-1, 2, and 3 were obtained by computer.

2-3 Plots Showing the Constructions of Two Composite Zone Plate Patterns

(a) A composite zone plate made from 1) a positive zone plate and 2) a negative zone plate

(b) A composite zone plate made from 1) a -60°-phase-shifted zone plate, 2) a 60°-phase-shifted zone plate, and 3) a 180°-phase-shifted zone plate (or a negative zone plate)

2-4 Schematic Diagram Showing the Coding and Decoding Process

The input-output relationship of coded aperture imaging is shown in both the spatial domain and the frequency domain.

2-5 Plots of Modulation Transfer Functions $|H(\underline{\nu})|$ of a Few Coded Aperture Systems

The 2-D gray-shade plots and cross-sectional (along x=y line) plots of $|H(\underline{\nu})|$ were generated by computer from the following:

(a), (b), (c) A non-redundant array, a regular (or redundant)

array, and a random array respectively

The transfer functions of the pinhole arrays do not vanish at high spatial frequencies because the pinholes in the arrays are of narrow widths (also because of the property of FFT used).

(d),(e) A positive zone plate ($h_{zp}(\underline{r})$, 9 zones) and a combined zone plate ($h_{zp}(\underline{r})-\tilde{h}_{zp}(\underline{r})$, 10 zones) respectively

(f) A random array of large number of holes

2-6 Imaging of Two Point Sources with a Circular Pinhole Array Coded Aperture

(a) Formation of the shadowgram (showing only two pinholes)

(b) Reconstruction of the shadowgram with a replica of the original pinhole array (with smaller pinhole size)

3-1 Schematic Diagram of a Multiwire Proportional Chamber

(a) A side view. The x- and y- wire planes are mutually orthogonal

(b) A top view showing the wire planes and the delay lines

3-2 Electromagnetic Delay Line

(a) A 3-view engineering drawing of the delay line

(b) Pulse-shape characteristics of a signal coupling into and out of the delay line

3-3 Electronics Block Diagram of a Multiwire Proportional Chamber Camera System

Only part of the digital image reconstruction system is shown in this figure

3-4 Coded Aperture Used with Wire Chamber Camera for the Study of Coded Aperture Imaging

(a) A non-redundant pinhole array (27 holes)

(b) A redundant pinhole array (27 holes)

(c) A random pinhole array (27 holes)

(d) A Fresnel zone plate (9 zones)

3-5 Autocorrelation Functions of the Aperture Patterns Used

The autocorrelation functions of the previous four patterns were obtained optically, and they are shown in the same order as in the previous figure.

3-6 Sketches of the Previous Autocorrelation Functions

The previous autocorrelation functions are drawn schematically to show their important features.

3-7 Simple Decoding System for Non-redundant Array Coded Shadowgrams

(a) A schematic diagram of the decoding system

(b) Construction of Picker Thyroid Phantom. The lobe on the left has twice the activity of that on the right and contains two cold spots of zero activity. The right lobe has an upper cold spot and a lower hot spot of the same activity as that on the left.

3-8 Imaging of 3-D Source Distribution with Non-Redundant Pinhole Array

A radioactive phantom consisting of three geometrical patterns (a triangle, a cross, and a circle) was imaged with a 27-hole non-redundant pinhole array coded aperture.

(a) The coded shadowgram

(b), (c), and (d) The reconstructed images

3-9 Imaging of 3-D Source Distribution with Non-Redundant Pinhole Array

The previous radioactive source was moved toward the aperture plate so that the pinhole projections of the source distribution on the detector plane spread more. The shadowgram and the reconstructed images are shown in the same order as those in the previous figure.

3-10 Reconstructions Showing Lateral Resolution of the System

(a) Image of four point sources. The four point sources having a minimum spacing 5 mm were imaged with a 27-hole non-redundant pinhole array coded aperture.

(b) Image of a bar pattern. 5 mm bars spaced 5 mm apart were imaged with the previous coded aperture.

3-11 Sequential Image Reconstructions Showing Depth Resolution of the System

Coded aperture imaging of 3 objects separated by 25 mm in depth was made with the system of Fig. 3-8. Images were reconstructed sequentially at intervals of 6.25 mm along the optical axis (z axis) of the incoherent optical decoding system. The parameters for imaging and reconstruction were S_o=175 mm, d=2 mm, D=22 mm, and

(4) ∇ : S_1 = 100 mm, Δ_z(calc.) = 14 mm.

(8) X : S_1 = 125 mm, Δ_z(calc.) = 19 mm.

(12) O : S_1 = 150 mm, Δ_z(calc.) = 25 mm.

3-12 Image Produced by Point Gamma-Ray Sources Using Various Apertures

Coded apertures of Fig.3-4 were used to image point gamma-ray sources. The images obtained with wire chamber camera and optical reconstruction are shown in the same order as their apertures in Fig.3-4.

3-13 Images of Picker Thyroid Phantom with 27-Hole Array Coded Apertures

The images were obtained with (a) a non-redundant array, (b) a redundant array, and (c) a random array.

3-14 Graphical Interpretation of Eq.(3.6) Showing Intensity Distribution of Fresnel Diffraction of a Positive Zone Plate

Only the diffraction patterns of n=0, ± 1, ± 3, and -5 are shown. The magnitudes of the diffraction patterns shown in the figure are not exactly proportional to their actual values.

3-15 Fresnel Diffraction Amplitude of a Positive Zone Plate

The diffracted amplitude of a positive zone plate (9 zones) was generated by computer and is shown as a function of r.

3-16 Coherent Optical Reconstruction System for Zone Plate Shadowgrams

(a) A schematic diagram of the coherent optical system

(b) A photograph of the entire reconstruction system

3-17 Optical Reconstruction of Point Sources with a Positive Zone Plate

(a) A shadowgram of 4 point sources separated at least by 2 resolution lengths of the imaging system.

(b) The reconstructed image using the system of Fig.3-16b.

3-18 Image of Picker Thyroid Phantom Obtained with a Positive Zone Plate and Coherent Reconstruction

The reconstruction was performed by computer using the Fresnel diffraction formula.

4-1 Image of Picker Thyroid Phantom from Complex Coding with Two Zone Plates

A positive zone plate (9 zones) and its negative zone plate were used to make this computer simulation of complex coded aperture imaging. The decoding was done by autocorrelation with $h_{zp}(\underline{r}) - \tilde{h}_{zp}(\underline{r})$.

4-2 Image of 3-D Object from Complex Coding with Two Zone Plates

The computer generated 3-D object was a triangle, a cross, and a circle, which were separated by 2 cm in depth. The reconstructed tomographic images of the object are shown in this figure. They were obtained from the same system and decoding method as those described previously.

4-3 In-Focus Point Reconstructions Using Non-Redundant Pinhole Array and Fourier Transform Deconvolution Method

The reconstruction of a point source at the in-focus plane were made using the Fourier transform deconvolution method and a 9-

hole non-redundant pinhole array coded aperture (shown in Fig. 4-5a). The resulting images, under various conditions (given below), are plotted cross-sectionally along the y = 0 line.

(a) No input statistical noise :

1) α = 0.01, 2) α = 0.05, and 3) α = 0.1

(b) Input noise-to-signal ratio $(N/S)_i$ = 10% :

1) α = 0.01, 2) α = 0.05, and 3) α = 0.1

(c) Input noise-to-signal ratio $(N/S)_i$ = 20% :

1) α = 0.01, 2) α = 0.05, and 3) α = 0.1

4-4 Out-of-Focus Point Reconstructions Using Non-Redundant Pinhole Array and Fourier Transform Deconvolution Method

The reconstructions of a point source at two out-of-focus planes were made under various conditions (given below), using the pinhole array and decoding method described previously

(a) α = 0.01 : 1) 2 cm out of focus, 2) 4 cm out of focus

(b) α = 0.05 : 1) 2 cm out of focus, 2) 4 cm out of focus

(c) α = 0.1 : 1) 2 cm out of focus, 2) 4 cm out of focus

4-5 Image of Picker Thyroid Phantom with Fourier Transform Deconvolution Method

The Picker Thyroid Phantom was generated by computer with 5% statistical noise in each picture element. The parameter used for the Fourier transform deconvolution was α = 0.05.

(a) The 9-hole non-redundant pinhole array used

(b) The reconstructed image

4-6 Reconstruction of a 3-D Source Distribution with Fourier Transform Deconvolution Method

The 3-D source distribution consisting a cross, a circle, and a triangle was generated by computer with 5% noise in each picture element. The object planes were separated by 2 cm in depth. The parameter α was 0.05 and the pinhole array was the same as in Fig.4-5. The reconstructed images are shown in (a), (b), and (c).

5-1 Multiple Single-Pinhole Camera

(a) Making the different views with the multiple single-pinhole camera

(b) Response to a point source. The tomographic images of the point source are shown as the sum of back projections of the single-pinhole views.

5-2 Point Blurring Patterns

(a) For multiple pinhole array

(b) For rotating slanted-hole collimator with discrete rotations (or for circular multiple pinhole array)

(c) For rotating slanted-hole collimator with continuous rotation

(d) For positron camera with data selection (square detectors)

5-3 Rotating Slanted-Hole Collimator Camera - Response to a Point Source

The tomographic images of the point source are shown for two positions of the collimator, 0^{o} and 180^{o}.

5-4 Positron Camera - Response to a Point Source

The tomographic images of the point source are shown for two positron events.

6-1 Schematic Diagram Showing the Formation of Tomographic Images from Pinhole Projections

(a) Exposure of the K^{th} image using the K^{th} pinhole of the array

(b) The contribution of the K^{th} single pinhole image to the N_p=3 tomographic planes. The final images on t_1, t_2, and t_3 are sum of the contributions from each of the N_h pinholes.

(c) The response of tomographic planes t_i and t_j to a point source in object plane o_i

6-2 Plots of the Determinant Functions $D(\underline{v}')$ of Three Pinhole Arrays

The determinant functions were computed by computer with different imaging geometries and pinhole arrays. Both 2-D gray-shade plots and cross-sectional (along y=0 line) plots are shown with $v' = 0$ at the left bottom of each plot. The imaging geometries were i) s_1=8cm, and s_2=10cm for N_p=2; ii) s_1=8cm, s_2=10cm, and s_3=12cm for N_p=3, iii) s_1=8cm, s_2=10cm, s_3=12cm, and s_4=14cm for N_p=4.

(a) For a (Y-type) non-redundant pinhole array

(b) For a circular pinhole array of odd number of holes

(a) Tomographic images on the five planes of a five plane object using the point blurring pattern of Fig.5-2a

(b) Reconstructions of these planes showing a small amount of background introduced by photon statistics

7-6 3-D Image Reconstruction - Experiment Using MWPC Digital System and Radioactive Source

A circular pinhole array of 9 holes (similar to the one shown in Fig.5-2b) was used to construct the tomographic images t_1, t_2, and t_3. The reconstructions of the tomograms are o_1, o_2, and o_3.

7-7 3-D Image Reconstruction - Experiment Using MWPC Digital System and Radioactive Source

(a) Tomographic images of five numerical patterns separated by 5 cm in depth, obtained with a 9-hole circular array

(b) The reconstructed images

7-8 3-D Image Reconstruction - Experiment Using MWPC Digital System and Radioactive Source

(a) Tomographic images of five point sources separated by 5 cm in depth, obtained with a 9-hole circular array

(b) The reconstructed images

8-1 Plots of the Determinant Functions $D(\nu')$ of a Few Tomographic Imaging Camera Systems

$R_{smallest}$ is the radius of the smallest point blurring pattern produced by each camera system. δx represents the lateral resolution length of each camera system.

8-2 Schematic Diagram Showing a Cause of Spatial Variance
Because of the inverse-square law, the blurring pattern of a point source at point 2 is different from that of a point source at point 1.

8-3 Image Reconstruction Showing Effect of Fewer Planes for Image Processing
(a) A 3-D object consisting of seven numerical patterns (each of size 23cm x 23cm) located at s_i=7cm, 8cm, 9cm, 10cm, 11cm, 12cm, and 13cm from the pinhole array
(b) One of the 9 single-pinhole projections of the object obtained with a circular array
(c) Three tomograms formed at s_i'=8cm, 10cm, and 12cm
(d) Reconstructions of the above three tomograms

8-4 Image Reconstruction Showing Effect of Fewer Planes for Image Processing
(a) Normal projection of the object showing the relative positions of the square sources. The small square sources were in place of the previous 3-D object, each source was in one of the seven object planes.
(b) Three tomograms formed at s_i' = 8cm, 10cm, and 12cm
(c) Reconstructions of the above three tomograms

8-5 Image Reconstruction Showing Effect of Wrong Plane Locations for Image Processing
A 3-D object consisting of three numerical patterns was imaged with a 9-hole circular array. The correct object plane

locations were s_1=8cm, s_2=10cm, and s_3=12cm. Three tomograms and their reconstructions were made at planes $s_i' = s_i + z$, $i = 1,2,3$.

(a) Reconstructions with z = -0.5 cm

(b) Reconstructions with z = 0.5 cm

(c) Reconstructions with z = 1.0 cm

(d) Reconstructions with z = 1.5 cm

A-1 Point Response for 3-Hole Non-Redundant Pinhole Array Coded Aperture Imaging

The decoding is done by optical correlation. Dotted lines are background.

A-2 Analysis of Image Using the Autocorrelation Function

(a) Sketch of approximate autocorrelation function of a 6-hole non-redundant array as measured in the detector plane

(b) Object points, spaced by the system resolution, projected onto the detector plane

B-1 Example of the Distributions of Random Processes Z_t and $\dot{Z}_t$

(a) Random process Z_t

(b) Random process $\dot{Z}_t$

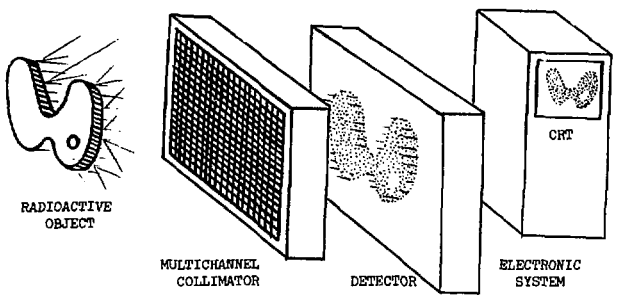


FIGURE 1 - 1

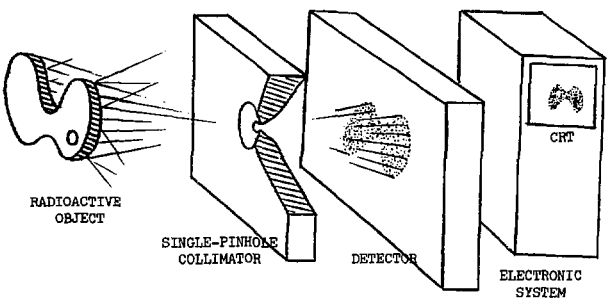


FIGURE 1 - 2

XBL 765-1736

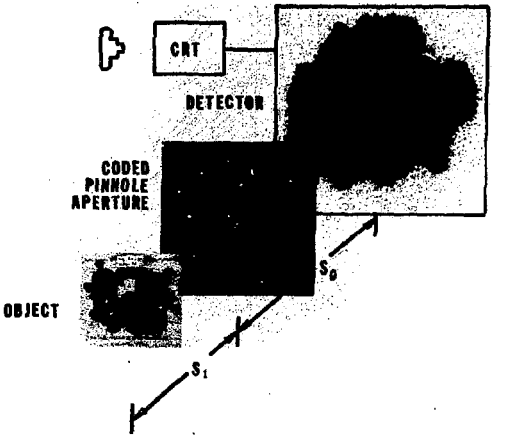


FIGURE 1 - 3

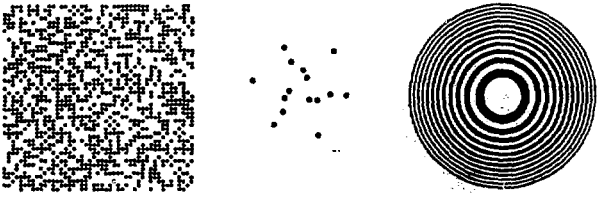
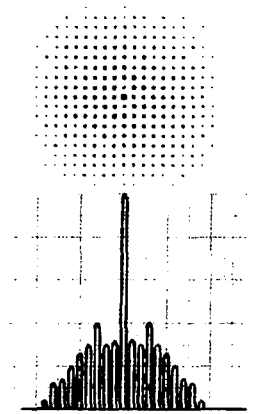
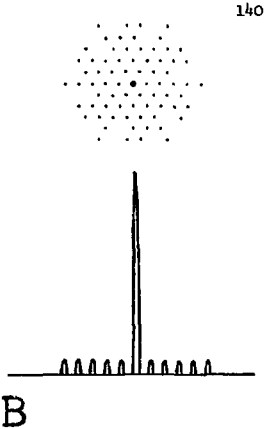
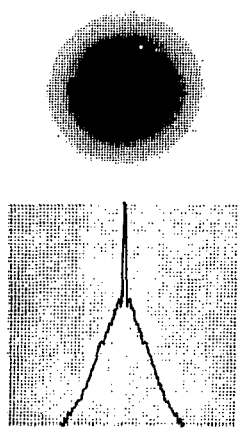


FIGURE 1 - 4

XBB 765-4956







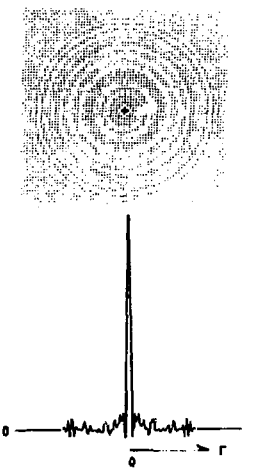


FIGURE 2 - 1

XBB 763-2863

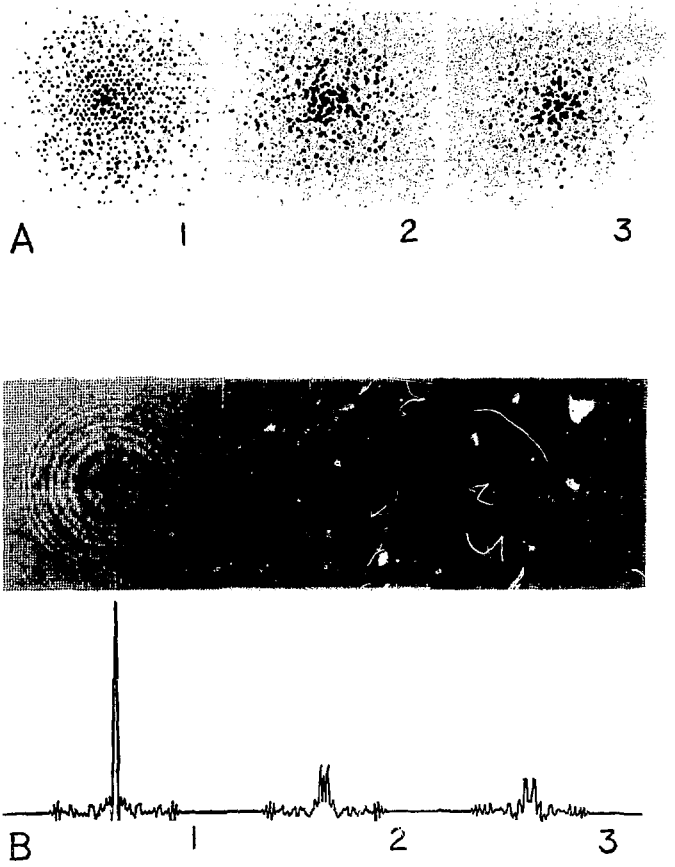


FIGURE 2 - 2

XBB 763-2862

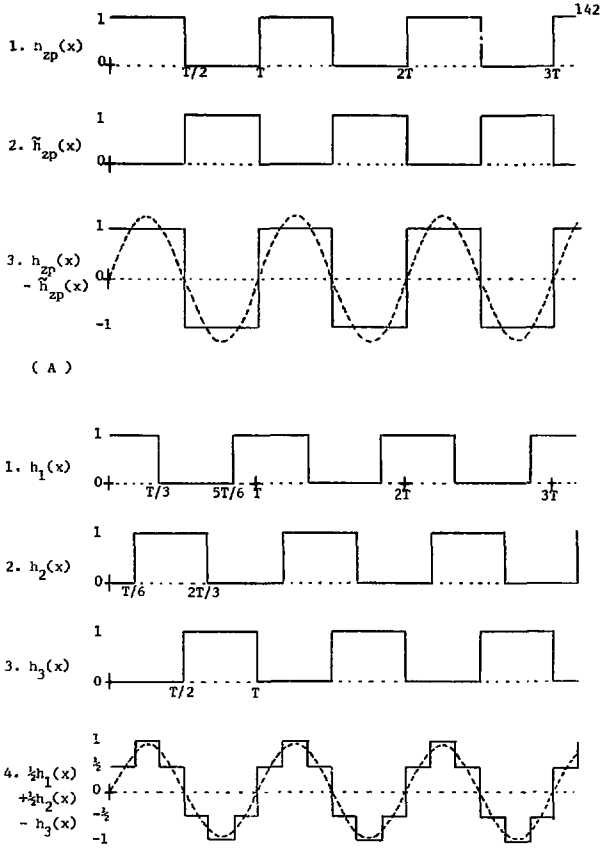
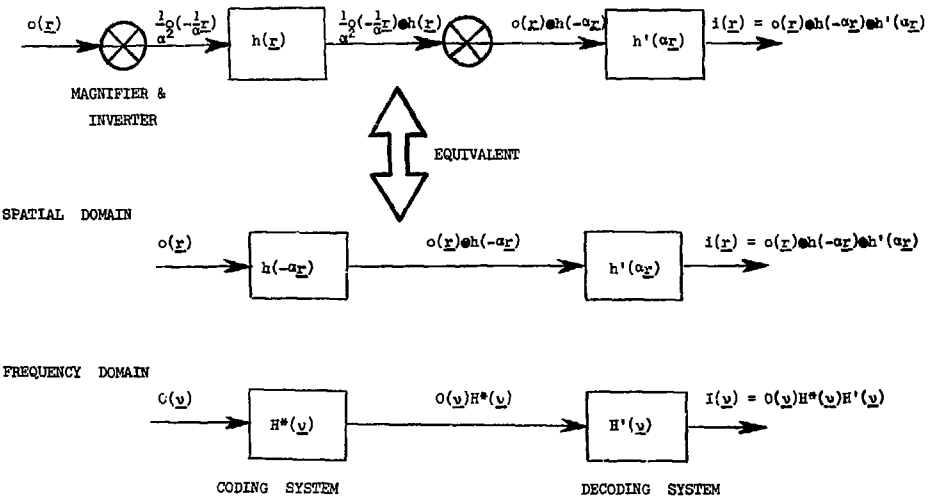


FIGURE 2 - 3

XBL 765-1737

BLOCK DIAGRAM OF THE CODING-DECODING PROCESS



XBL 765-1735

FIGURE 2 - 4

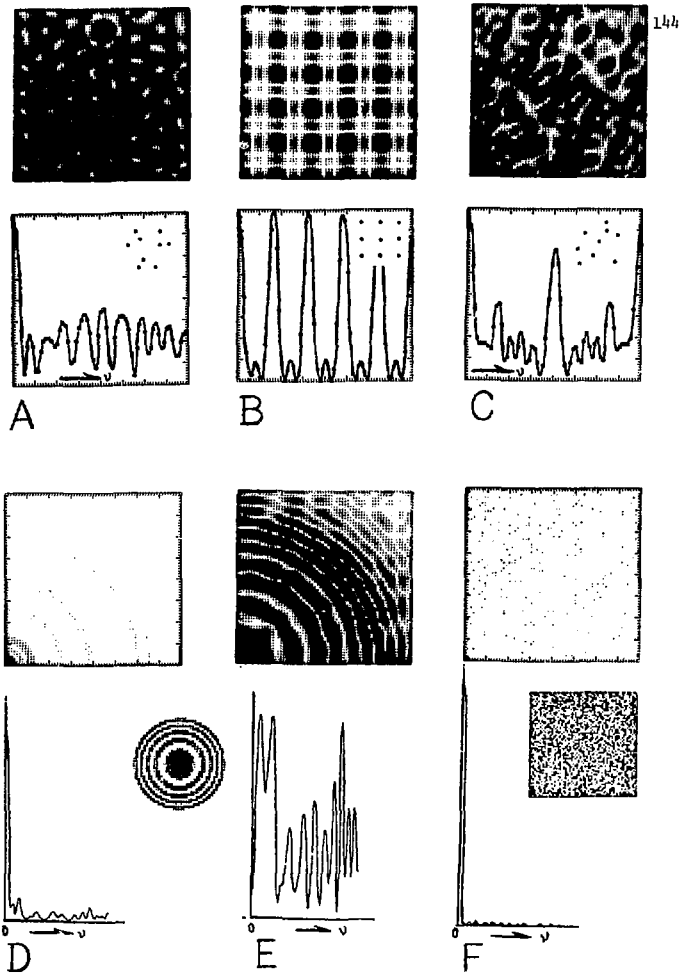
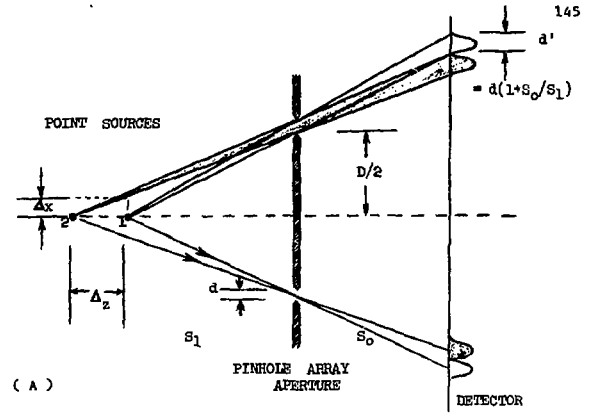
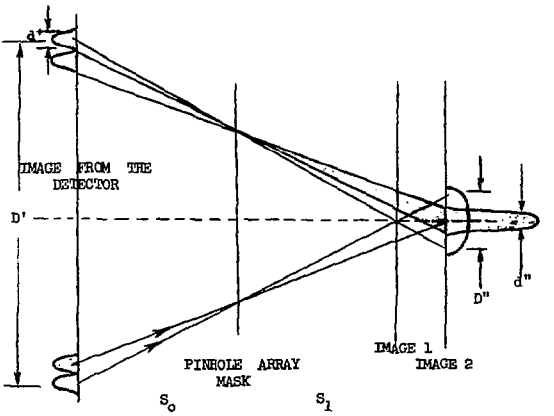


FIGURE 2 - 5

XBB 763-2859



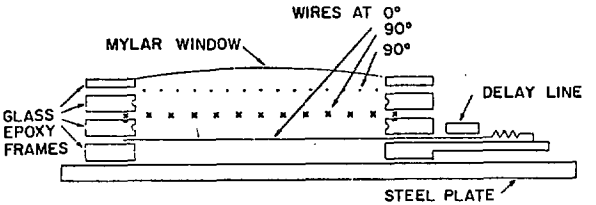
(A)



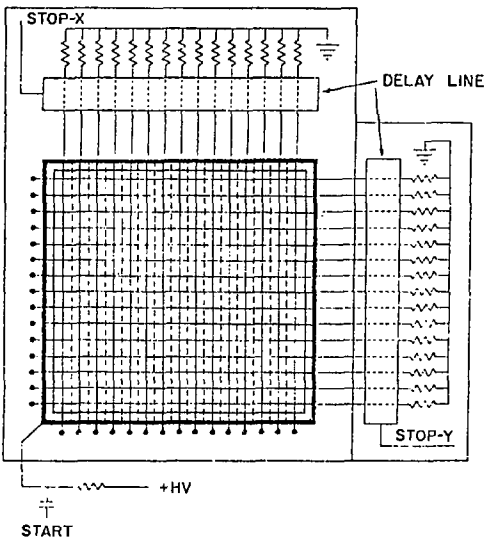
(B)

FIGURE 2 - 6

XBL 765-1738

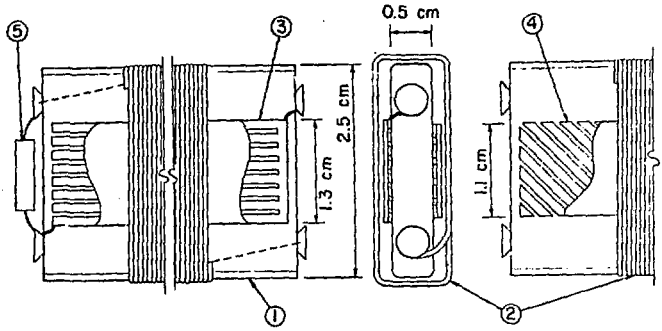

WIRES AT 0°
90°
90°
MYLAR WINDOW
DELAY LINE
GLASS
EPOXY
FRAMES
STEEL PLATE

(A)


STOP-X
DELAY LINE
STOP-Y
+HV
START

(B)

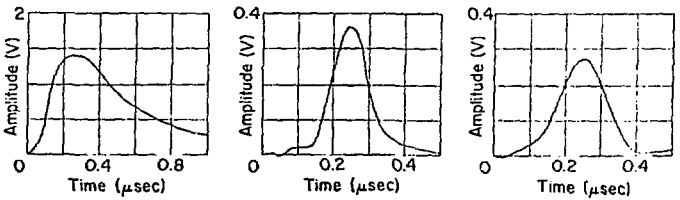
FIGURE 3 - 1



1 Plastic Core
2 Winding (#38 wire, 8 turns/mm)
3 Ground Strips (6.3 strips/cm of 12 μm Al on 25 μm Mylar. Strip width = 1.2 mm, gap width = 0.4 mm)
4 Compensating Strips (Same material as ground strips)
5 Terminating Resistor

(A)

DELAY LINE PULSES



Input Pulse from Chamber

Output Pulses from Delay Line
(10 cm down line) (40 cm down line)

(B)

FIGURE 3 - 2

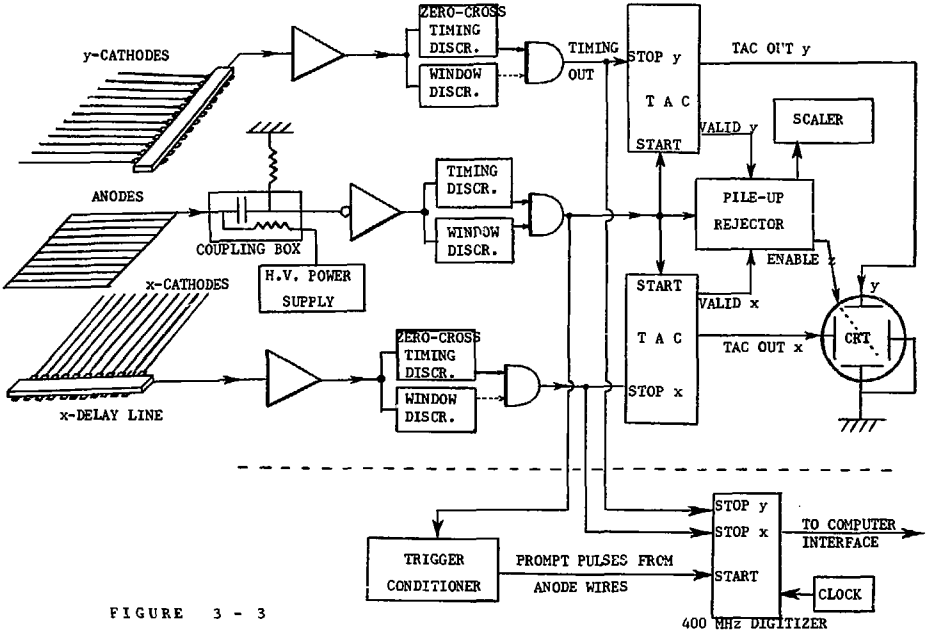
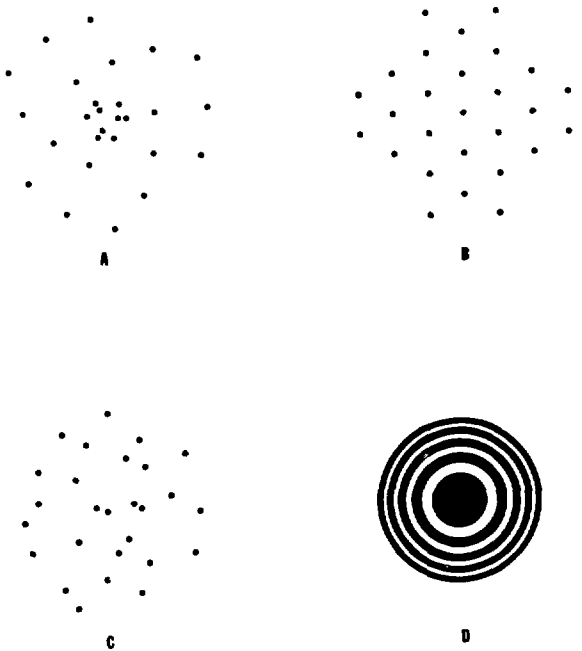


FIGURE 3 - 3

XBL 765-1745



XBL 765-1750

FIGURE 3 - 4

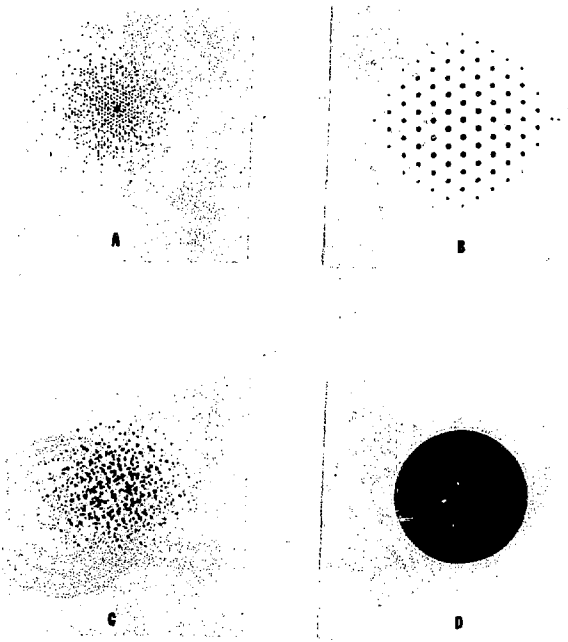
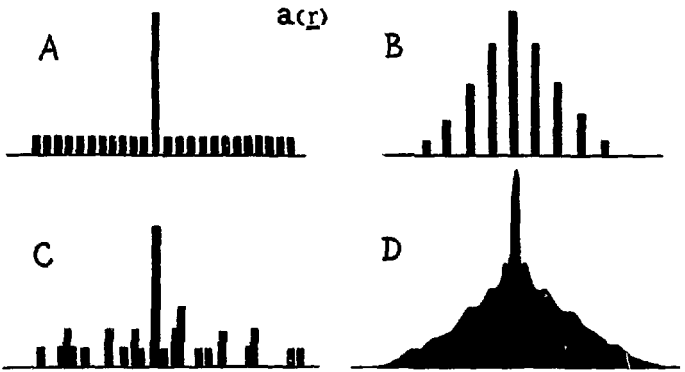


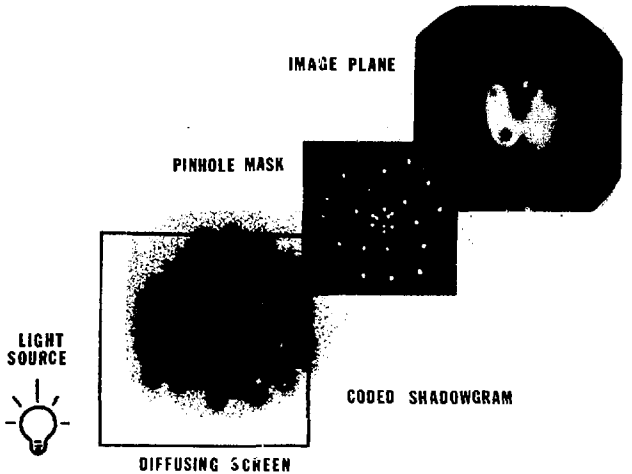
FIGURE 3 - 5

XBB 765-4952

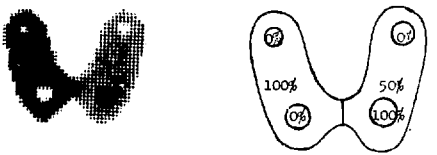


XBL 765-1743

FIGURE 3 - 6



(A)



(B)

FIGURE 3 - 7

XBB 765-4958

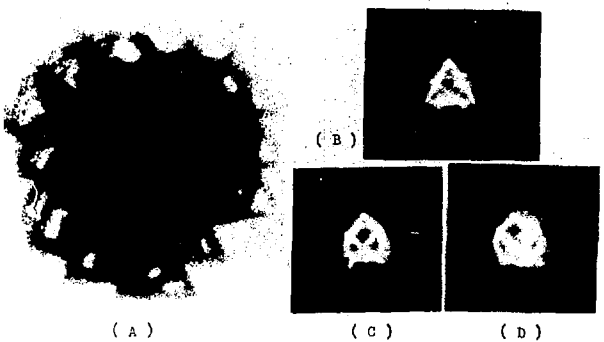


FIGURE 3 - 8

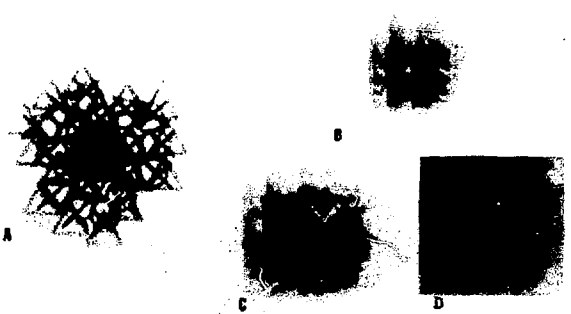
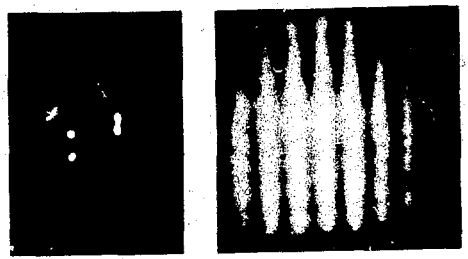


FIGURE 3 - 9

XBB 765-4959



(A) (B)

FIGURE 3 - 10

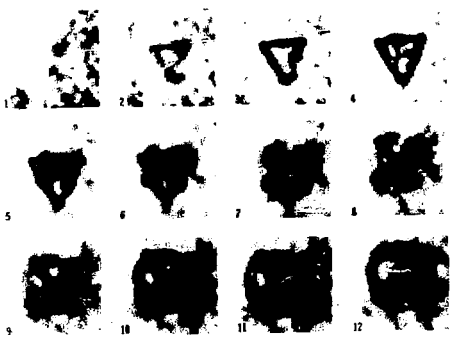


FIGURE 3 - 11

XBB 765-4955

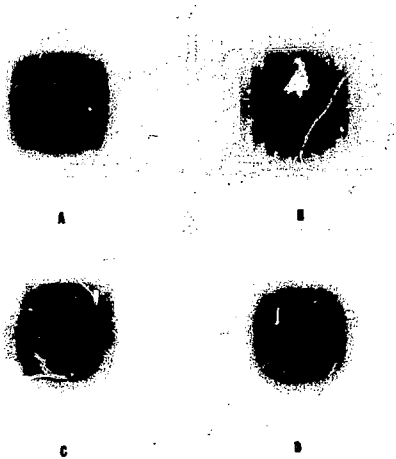
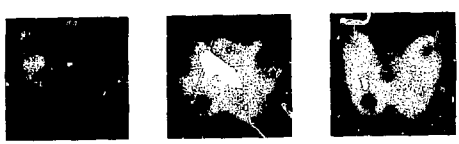


FIGURE 3 - 12



(A) (B) (C)

FIGURE 3 - 13

XBB 765-4954

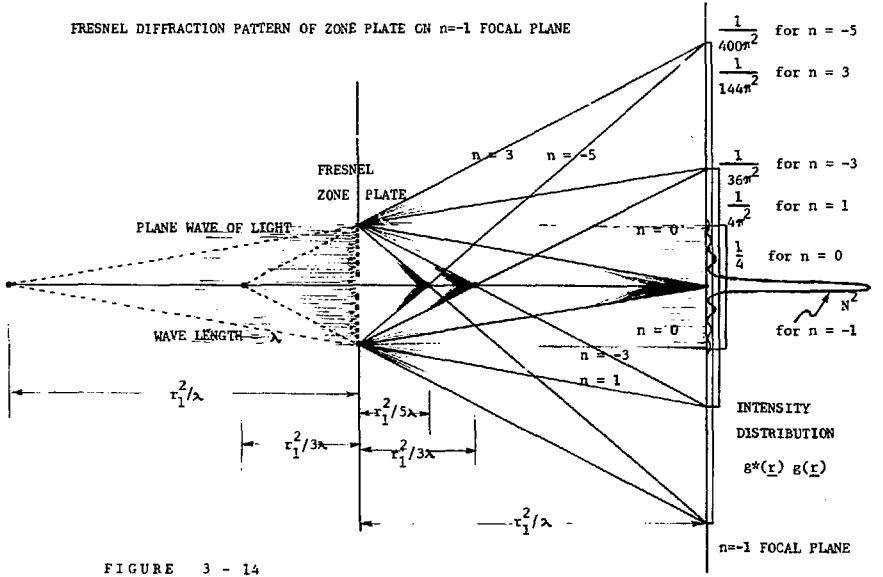


FIGURE 3 - 14

XBL 765-1739

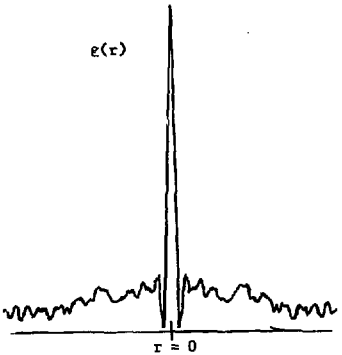
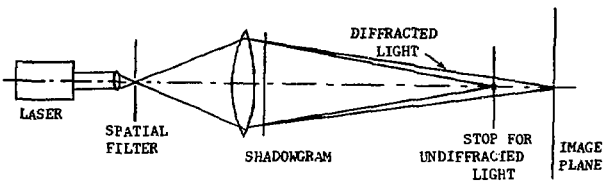


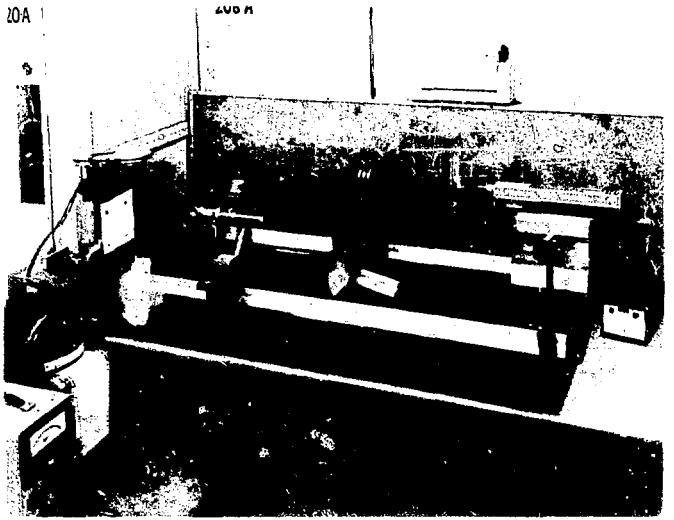
FIGURE 3 - 15



(A)

XBL 765-1740

FIGURE 3 - 16



XBB 735-3242

(B)

FIGURE 3 - 16

(A) (B)

FIGURE 3 - 17

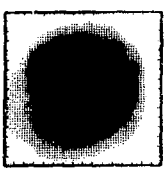


FIGURE 3 - 18

XBB 765-4960

FIGURE 4 - 1

FIGURE 4 - 2

XBB 765-4953

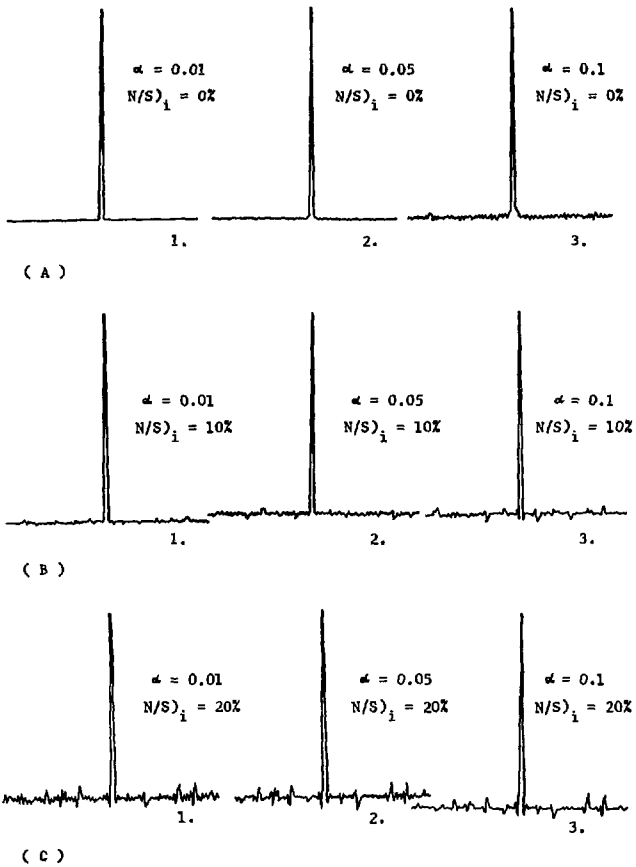
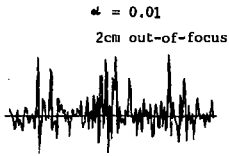
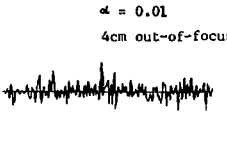


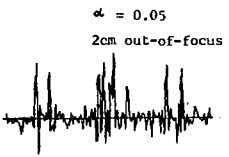
FIGURE 4 - 3

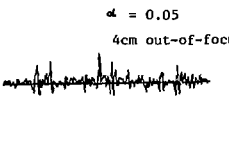
XBL 765-1741



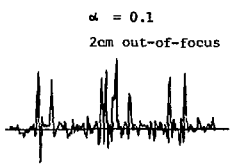


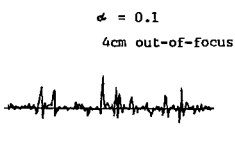
(A)





(B)

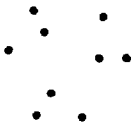




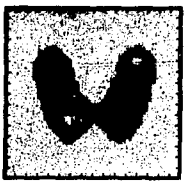
(C)

FIGURE 4 - 4

XBL 765-1742

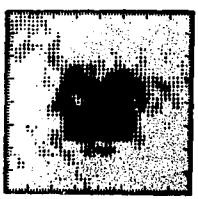


(A)

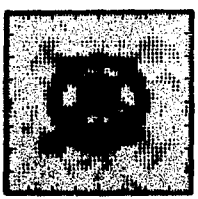


(B)

FIGURE 4 - 5



(A)



(B)

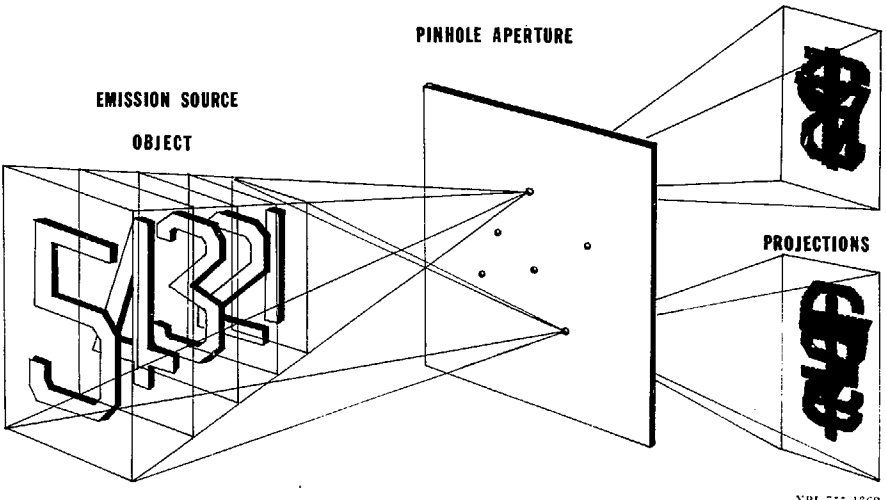
(C)

FIGURE 4 - 6

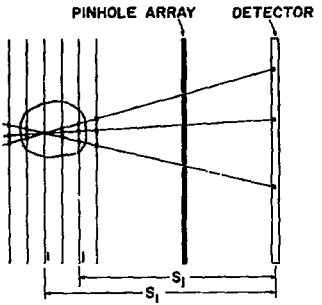
XBB 765-4957

FIGURE 5 - 1

THE IMAGING SYSTEM

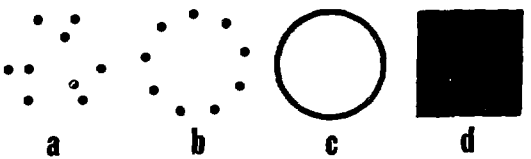


(A)



(B)

FIGURE 5 - 1



XBL 765-1749

FIGURE 5 - 2

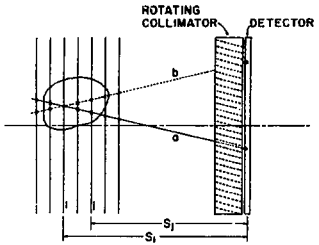
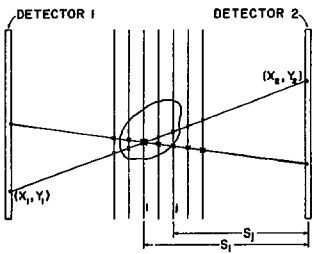
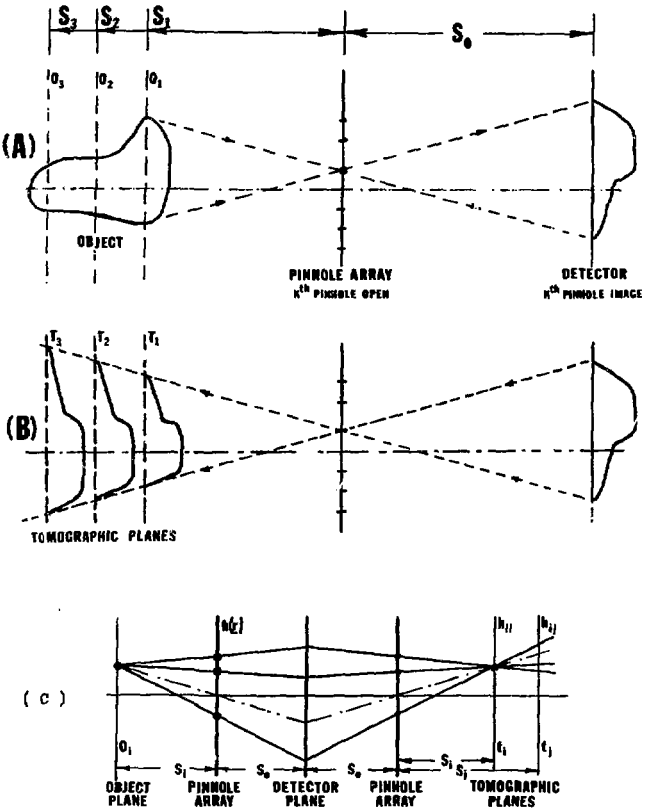


FIGURE 5 - 3



XBL 765-1751

FIGURE 5 - 4



XBL 765-1752

FIGURE 6 - 1

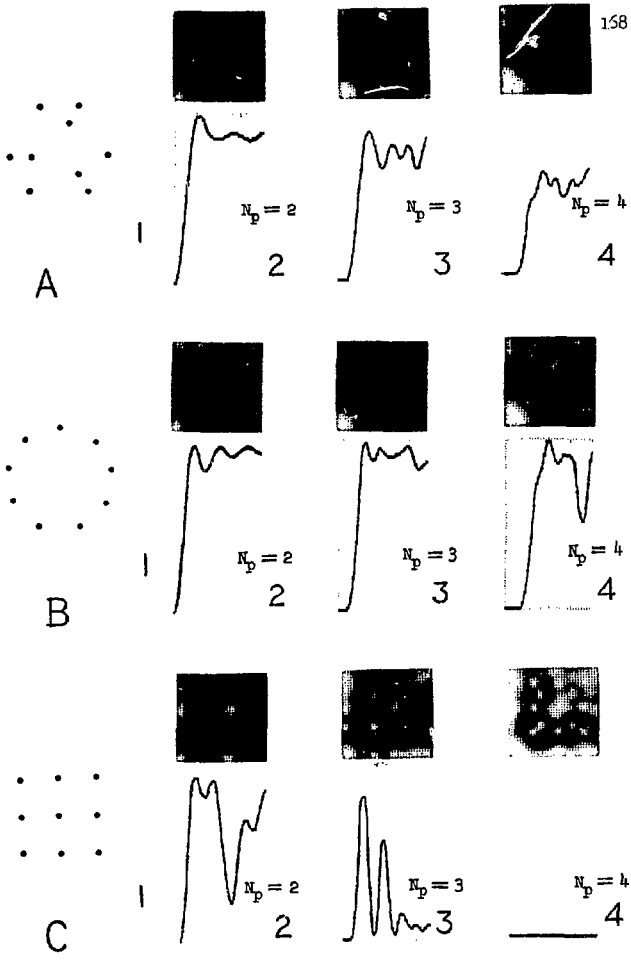
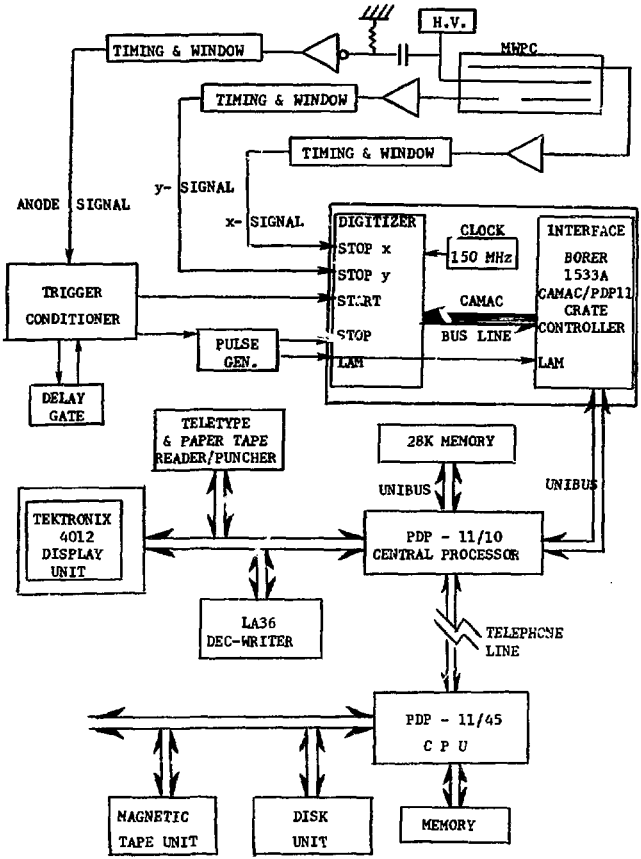


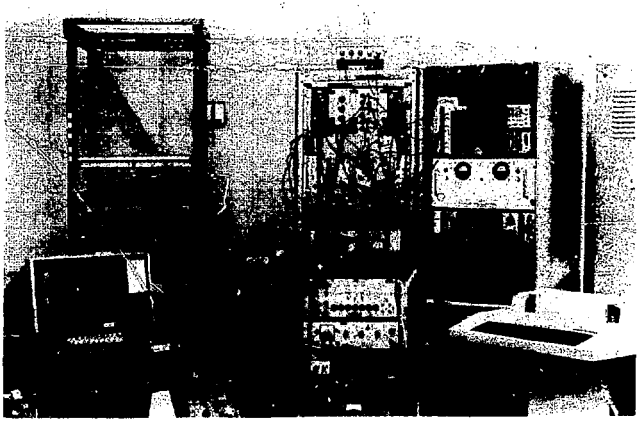
FIGURE 6 - 2

XBB 763-2860



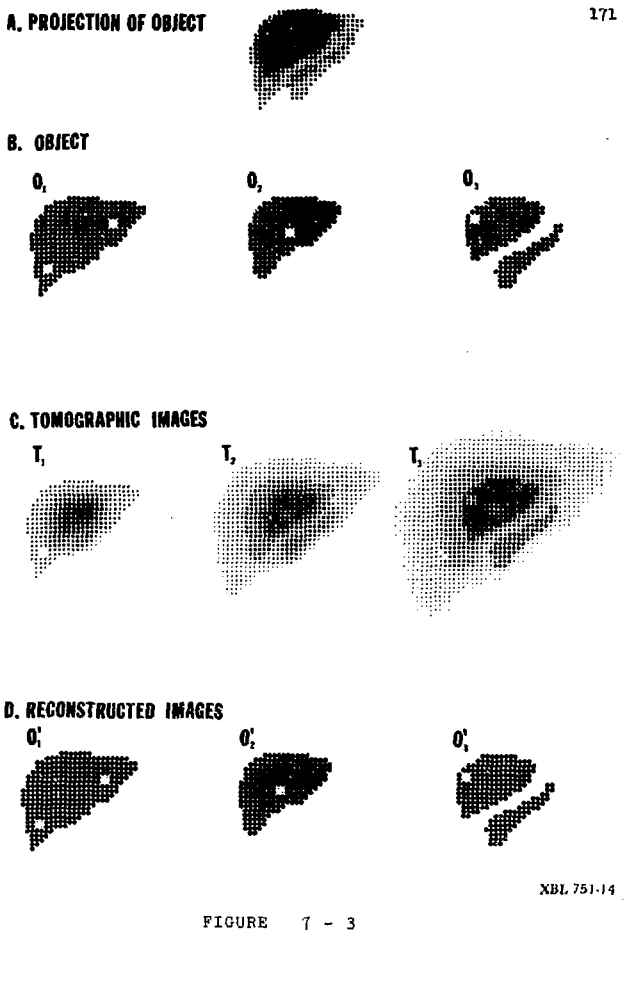
XBL 765-1744

FIGURE 7 - 1



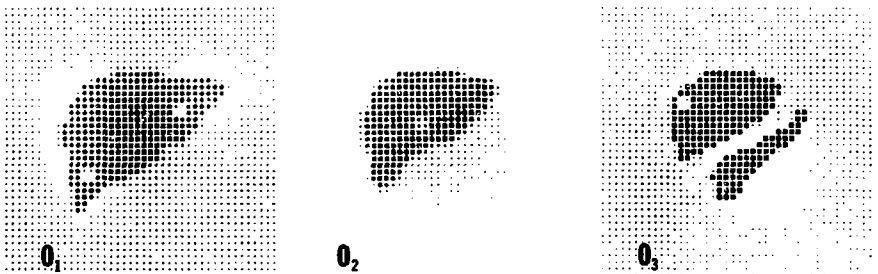
XBB 761-852

FIGURE 7 - 2



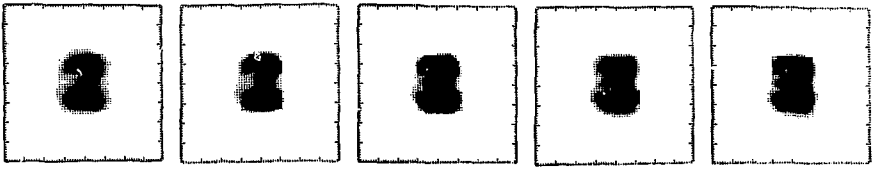
XBL 751-14

FIGURE 7 - 3



XBL 755-1189

FIGURE 7 - 4

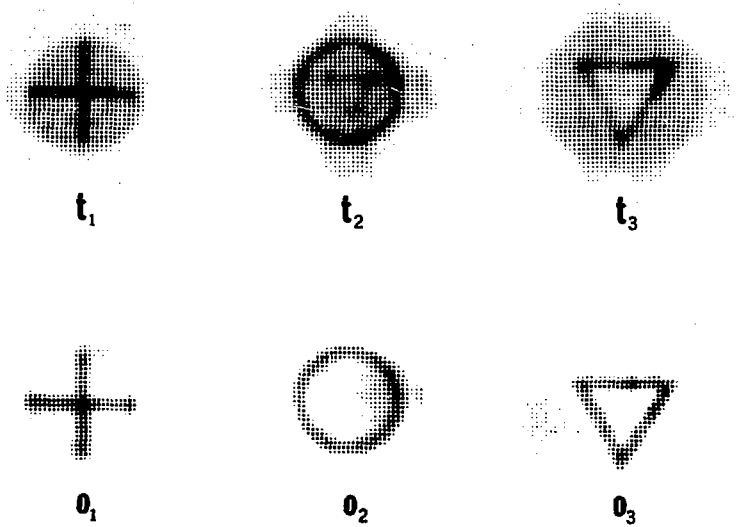


(A) THE TOMOGRAPHIC IMAGES

(B) THE RECONSTRUCTIONS

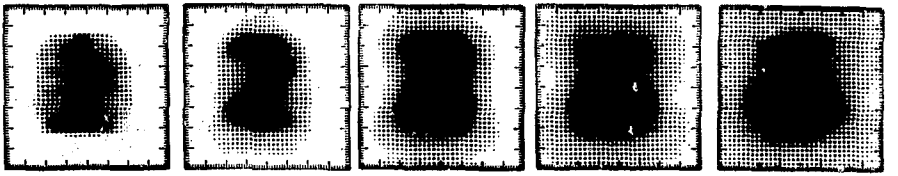
XBL 733-1866

FIGURE 7 - 5

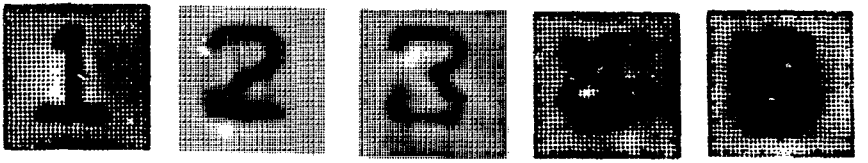


XBL 755-1188

FIGURE 7 - 6



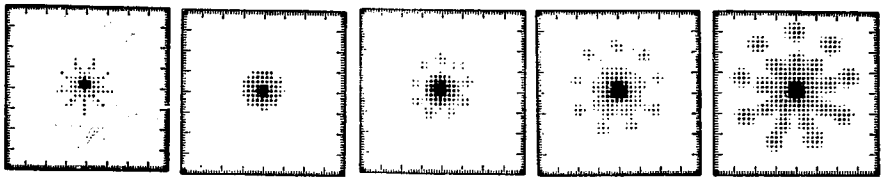
A



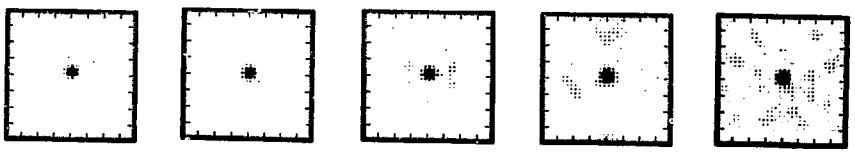
B

FIGURE 7 - 7

XBB 763-2856



A



B

FIGURE 7 - 8

XBB 763-2855

PLOT OF DET(H) AS FUNCTION OF FREQUENCY AND POINT BLURRING PATTERNS (DISC, RING, AND POINT ARRAYS)

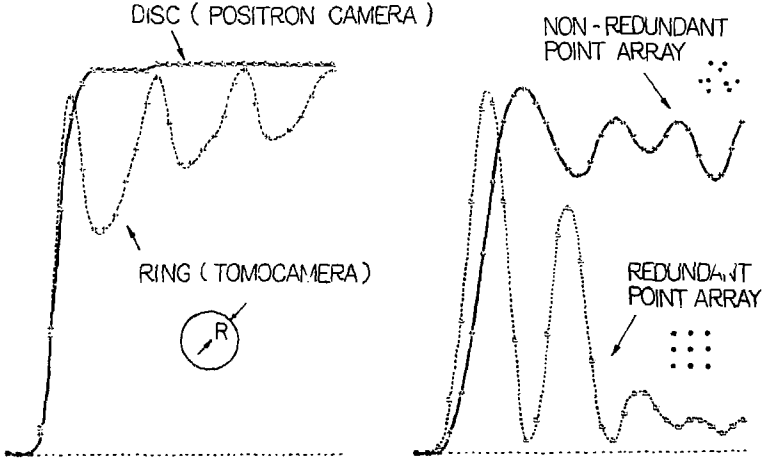


FIGURE 8 - 1

XBL 757-1800

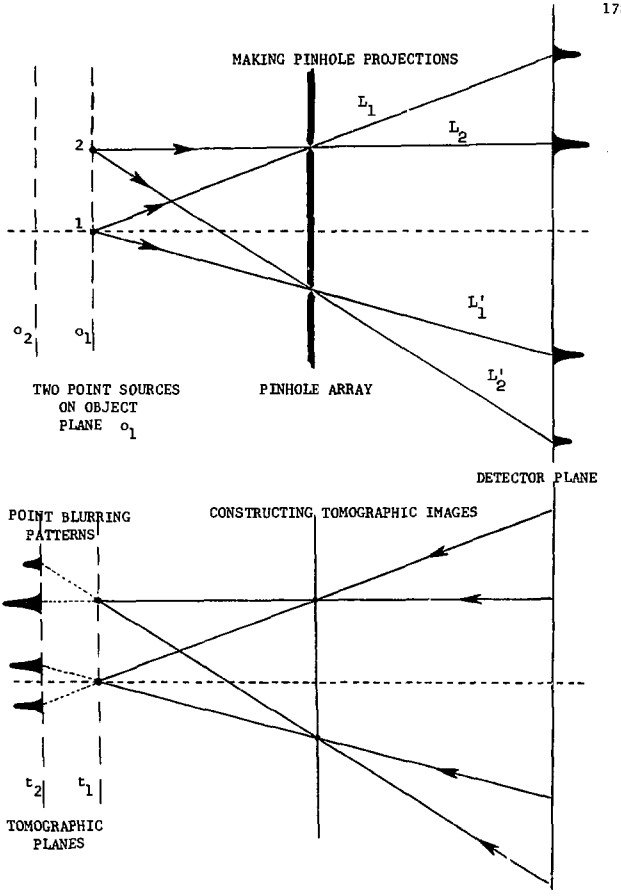
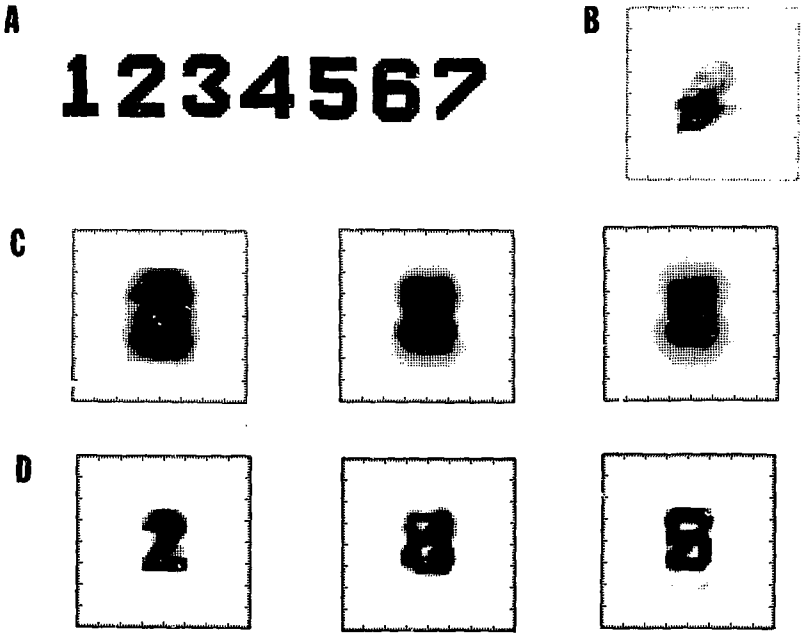


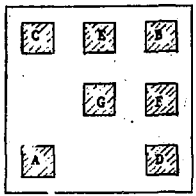
FIGURE 8 - 2

XBL 765-1747



XBL 755-1190

FIGURE 8 - 3



AN OBJECT CONSISTING OF 7 SMALL SOURCES

MINIMUM LATERAL SPACING BETWEEN SOURCES = 4 (X OR Y DIRECTION)

SIZE OF EACH SQUARE SOURCE = 4 x 4

MINIMUM LONGITUDINAL SPACING BETWEEN SOURCES = 1 (Z DIRECTION)

(A) NORMAL PROJECTION OF THE OBJECT ON THE DETECTOR PLANE

(B) TOMOGRAMS

(C) RECONSTRUCTIONS

FIGURE 8 - 4

XBB 765-4951

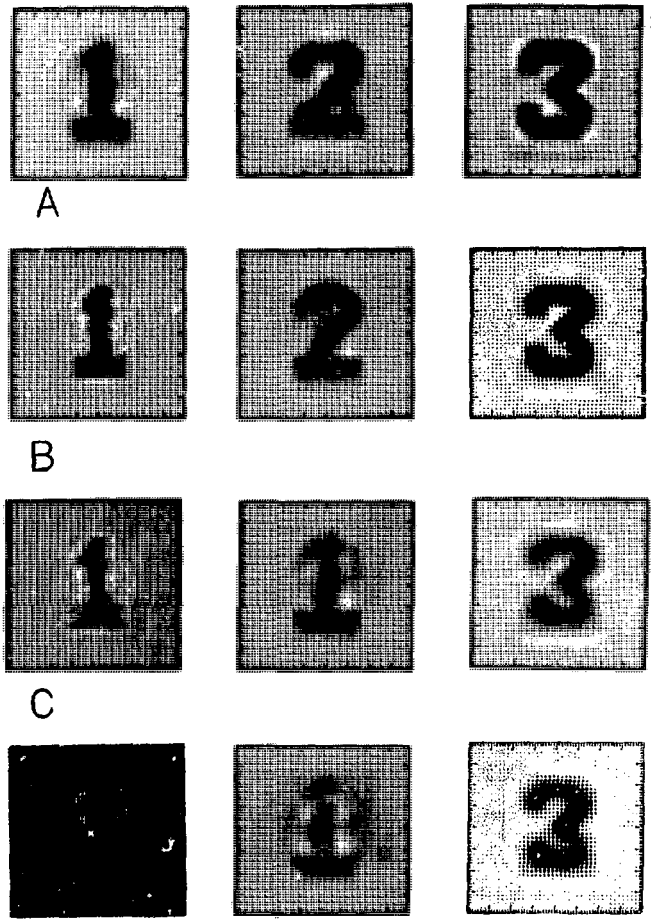


FIGURE 8 - 5

XBB 763-2857

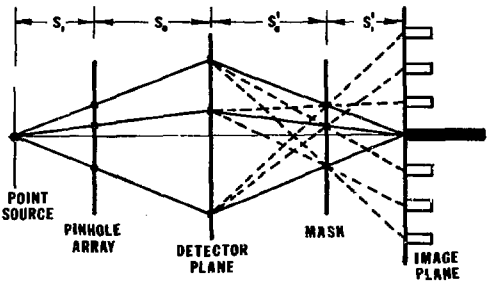
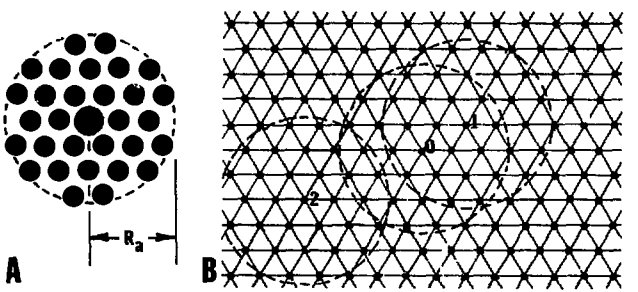
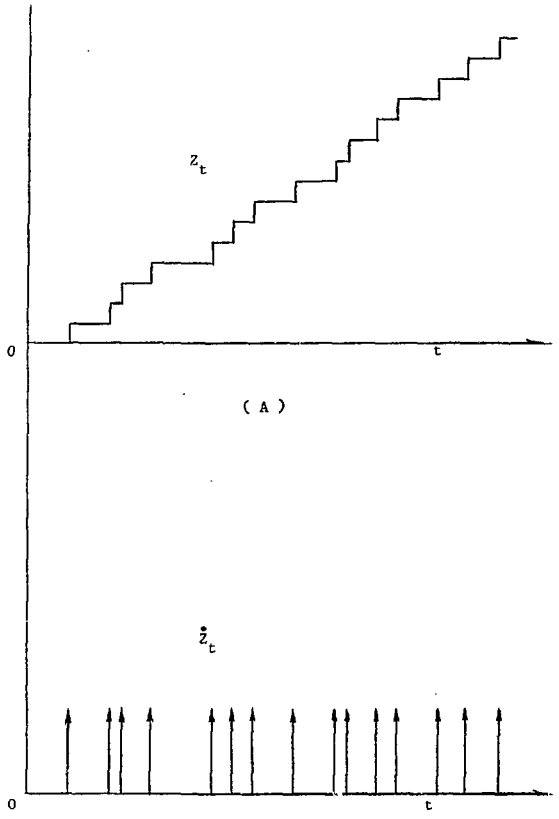


FIGURE A - 1



XBL 765-1748

FIGURE A - 2



XBL 765-1746

FIGURE B - 1