This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, D Clin Psychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
FOURIER TRANSFORM METHODS OF DECONVOLVING SCINTIGRAMS USING A GENERAL PURPOSE DIGITAL COMPUTER

A. KEITH BOARDMAN B.Sc.

I hereby declare that this thesis contains the results of my own work, and that it has been composed by myself.
# CONTENTS

## ABSTRACT

### CHAPTER 1  INTRODUCTION

1.1 INTRODUCTION TO IMAGE RESTORATION  
1.2 NUCLEAR MEDICINE IMAGING EQUIPMENT  
1.3 DATA PROCESSING EQUIPMENT  
1.4 RESEARCH PROGRAMME  

### CHAPTER 2  DATA ACQUISITION EQUIPMENT

2.1 PDP 12 INPUT/OUTPUT STRUCTURE  
2.2 J and P ENGINEERING CASSETTE RECORDER SYSTEM  
2.3 DIGIDECK SOFTWARE  
2.4 GAMMA CAMERA INTERFACE  
2.5 GAMMA CAMERA SOFTWARE  

### CHAPTER 3  DISPLAY EQUIPMENT

3.1 INTRODUCTION  
3.2 REFRESH OSCILLOSCOPE  
3.3 TELEVISION DISPLAY  
3.3.1 SCANNING SYSTEM  
3.3.2 MEMORY REFRESH SYSTEM  
3.3.3 TIMING CONTROL  
3.3.4 VIDEO CIRCUITRY  
3.3.5 COMPUTER INTERFACE  
3.3.6 PDP 12/TELEVISION SIMULATOR
CHAPTER 4  IMAGE FORMATION  
AND THE SPATIAL  
FREQUENCY DOMAIN  
4.1  BASIC CHARACTERISTICS OF THE ELEMENTAL  
SCINTIGRAM  66  
4.2  MATHEMATICAL MODEL  67  
4.3  DISCRETE MATHEMATICAL MODEL  71  
4.4  TWO DIMENSIONAL FFT COMPUTER PROGRAM  76  
4.5  POWER SPECTRUM  78  
4.6  PIXEL SIZE  81  
4.7  POINT SPREAD FUNCTION  84  
4.8  PHANTOM MODEL  88  
4.9  SUMMARY  91  

CHAPTER 5  CONSTRAINED  
OPTIMISATION  
5.1  INTRODUCTION  92  
5.2  DECONVOLUTION  93
CHAPTER 9
DIGITAL FILTERING
9.1 INTRODUCTION 145
9.2 TYPES OF DIGITAL FILTER 146
9.3 DESIGN OF NON-RECURSIVE FILTERS 149
9.4 REALISATION OF TWO DIMENSIONAL
NON-RECURSIVE FILTERS 151
9.5 A METHOD FOR OPTIMISING THE LOWPASS
FILTER BANDWIDTH 153
9.6 BANDPASS FILTERING 156
9.7 SUMMARY 158

CHAPTER 10
CLINICAL EXAMPLES
10.1 INTRODUCTION 160
10.2 PROCESSING DATA FROM MORPHIC
STRUCTURES 160
10.3 PROCESSING DATA FROM AMORPHOUS ORGANS 164
10.4 PROBLEMS OF DECONVOLVING CERTAIN TYPES
OF SCINTIGRAM 167
10.4.1 PHYSIOLOGICAL ARTEFACTS 167
10.4.2 INSTRUMENTAL ARTEFACTS 170

CHAPTER 11
SUMMARIES
11.1 HARDWARE 173
11.2 SOFTWARE 174
ABSTRACT

The adaptation of a general purpose laboratory minicomputer for nuclear medicine imaging is described. Electronic interfaces have been designed and constructed to link nucleonic equipment to a PDP 12 computer. A computer television display system has been developed to facilitate interactive processing of scintigraphic data. The main features of the television system are that it is relatively inexpensive and reliable. A domestic quality receiver has been adapted for use as a colour monitor.

Any instrument that records data will produce a distorted or degraded version of the input signal. Generally, imaging equipment will produce a blurred image of the object, and in the case of scintigraphic imaging the blurs may be comparable to the size of the physiological structures being investigated. The process of refocussing the recorded data is called, in mathematical terms, deconvolution. In this study Fourier transform techniques have been developed as methods of implementing deconvolution. It is shown that the restoration of images in the presence of noise is likely to be a mathematically unstable process. Four methods of accommodating the problems associated with noise are described. Each method has built in optimisation of one
form or another so that mathematically stable algorithms are used to implement deconvolution. This means that all the parameters used by the computer programs are determined automatically so that the computer operator is not required to select any parameters manually.

A brief description of two dimensional digital filtering is given to enable comparison between filtering and deconvolution of scintigrams. A two dimensional lowpass filter is developed which automatically defines the passband frequency response appropriate to a particular scintigram.

Finally, all the signal processing methods are tested on both simulated and clinical data. Results show that deconvolution offers advantages over digital filtering particularly for scintigrams obtained from morphic structures. Some of the problems of deconvolving certain types of scintigram are discussed.
CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION TO IMAGE RESTORATION

Image degradation is not unique to biological systems. Scientists have often demonstrated that optical systems can perform spatial filtering and are therefore capable of image restoration (Vander-Lugt 1968). Photographers are familiar with techniques which deliberately distort grey levels in a picture in order to bring out detail. However such techniques cannot increase the sharpness of a photograph beyond that of the original negative. Processing a blurred image spatially in order to improve its focus is a type of image restoration, while deliberately distorting grey levels of an image is a form of image enhancement. From about 1950 onwards a copious amount of literature has been presented on restoration techniques for different types of degradation in imaging systems. However it was the space programme of the 1960's that provided the necessary motivation for the development of the field. The success of the effort must be apparent to any lay observer who has witnessed the improvement in image quality of pictures received from inter-planetary space vehicles over the last ten years. It is only
recently that the restoration, enhancement and analysis of multi-dimensional signals has moved into such fields as medicine, geology, radar, seismology and astronomy.

In the medical scintigraphy field any imaging system will inevitably introduce distortion which may arise from a variety of causes. Probably the most obvious is that due to noise which originates from the measuring system itself, although statistical fluctuations in radioactive count rate usually constitute the most important source of noise. There is always a compromise between detector resolution and sensitivity; the finite width of collimator channels and septal penetration are associated with the blurring of an image. If the data are to be refocussed then whatever mathematical method is used it must be capable of accommodating the presence of noise.

In this study only digital image processing of radionuclide images has been implemented. Although optical image processing is instantaneous, experiments in optical image processing are difficult and often time consuming to set up. Furthermore radionuclide images are only visualised optically; the signal source itself normally consists of gamma quanta which occur as discrete events. Other analogue methods can be devised but digital methods offer the advantage of the relative ease in which
parameters in the restoration algorithm can be adjusted. This is a particularly valuable feature when new methods are being evaluated.

It was not until about 1965 that digital image processing became practical. This was largely due to the advent of the Fast Fourier Transform (Cooley 1965). Using the discrete Fourier transform the time required to transform N data values is proportional to $N^2$; with the Fast Fourier Transform (FFT) the time is proportional to $N \log_2 N$. Ideally the FFT requires N to be equal to $2^n$ where n is a positive integer power. This form is consistent with the 64*64 data arrays used in this study. The work reported here uses the FFT as a means of implementing deconvolution and filtering on a general purpose digital mini-computer.

In this study each deconvolution method has been tested rigorously from a mathematical viewpoint and then demonstrated using model data generated on the computer. The advantages and shortcomings of all the techniques are elaborated upon. Time and lack of sufficient clinical data prevented an objective assessment of the methods using, for example, Receiver Operating Characteristic plots. In any event a definitive assessment can only be obtained by detailed follow up after surgery or autopsy.
There has been much debate about the usefulness of using phantom data to evaluate image processing methods which have been highlighted in the IAEA intercomparison of image processing techniques (IAEA (1972), IAEA [1] (1976), IAEA [2] (1976)). The main argument against the use of phantoms is that organs in the human body do not consist of discrete well defined volumes containing uniform concentrations of radioactivity, and so the phantom data may not be a realistic model of a particular class of scintigram. However it may be that the deficiencies of the model data lie not so much in the model itself but in the fact that in a clinical situation the data may be subject to several errors which may not be accounted for in the model. Such errors or artefacts may arise from physiological or instrumental reasons. Examples of the former type may be simply due to lesions, although refocussed, being masked by biological noise due to irregular background structure within the organ, or, organ motion may continually "smear" the data being collected. Other errors may be introduced by the data collection system, for example, the instrumentation may be non-linear. Under these conditions it may be understandable that restoration methods do not perform as well as they may be expected to from tests with simulated data. Model data can be justified when the mathematical
viability of a deconvolution method is being tested. Some objective comparisons of restoration algorithms can be reached because the true radioactive distributions are known a priori. It should also be noted that the IAEA study concentrates its attention on enhancement rather than restoration techniques. It remains an open question whether deconvolution can offer any improvement in the number of false interpretations of clinical scintigrams.

In the past scintiographic image restoration methods have mainly been concentrated on developing "Wiener type" filters. Some examples can be found in the work of Pistor (1972,1972), Lorenz (1972), Bone (1973), Kirch (1973) and Pullan (1976). The motivation for using Wiener techniques is one of mathematical tractability; it is rather unfortunate that a true Wiener filter for scintiographic deconvolution is not easy to implement. Because approximations have to be made most workers produce Wiener type algorithms which are intuitive rather than based on sound mathematical foundations. A quotation from Gustafsson (1975) is characteristic of the way Wiener type filters are implemented:

"Our application of this Wiener type filter has, in fact, not been to use it as a deconvolution filter, but rather as a smoothing filter with good preservation of picture resolution. In other words, we have deliberately
mismatched the filter relative to the collimator response function and used it as an operator to increase signal to noise ratio almost independently of the characteristics of the imaging device."

A detailed account of the problems of Wiener filtering is presented in chapter 7. Brownell and Chesler (1971) also give an introduction to the basic concepts and problems of resolution enhancement.

Other restoration methods have been used in scintigraphy of which perhaps the most notable is that of Inuma (1967, 1971). He developed an iterative approximation method which has been shown to give encouraging results particularly for one dimensional data. The method uses an estimate of the noise in the data to control the number of iterations used to converge to an estimate of the true radioactive distribution. The method has been implemented in the spatial domain and the author admits that the iterative convolution integral is computationally time consuming; however if the optimum filter functions are determined empirically the convolution integral need be evaluated only once.

1.2 NUCLEAR MEDICINE IMAGING EQUIPMENT
Image restoration techniques have been applied to data from a J and P Engineering Multipoise scanner, and Nuclear Enterprise's upgraded MK IV and MK V HR gamma cameras which are all located within Edinburgh Royal Infirmary. Most mathematical development work has centred on data which has been collected from the J and P Engineering scanner. This is a dual detector instrument capable of several modes of operation. For the purposes of this study it has been used as a conventional rectilinear scanner. The maximum scanning field of each detector is 40*40 cm. The detectors are two 4*2 in. NaI crystals which can be used with one of three sets of collimators. Image processing has been tried on images obtained using Tc99m (140 keV). For this low energy isotope a pair of 109 hole collimators are available which offer a full width at half maximum (FWHM) resolution of 1.1 cm. at a focal distance of 12 cm. in air. The detector heads and the displays on the console are driven by stepping motors which in turn are controlled from a master clock. Division of the clock pulses within the scanner facilitates the linear formation of square pixels with 0.2, 0.4 or 0.8 cm. widths. In effect the nucleonic counts are integrated within each pixel before the appropriate signals are output to the displays. This digital operation of the scanner makes interfacing to a computer a relatively straightforward procedure.
The scanner was originally supplied with a cheap audio cassette tape drive which serves as a pulse recorder that can be used for recording any two outputs from four pulse height analysers in the scanner console. Because of bandwidth limitations on the tape recorder the data are divided by a factor of four and derandomised prior to input to the recorder. The data from the cassette may be replayed into the display system of the scanner so that either a monochrome tapper picture or photoscan can be reproduced. Initially it was decided to use the same cassette as a method of transferring data to the Department of Medical Physics and Medical Engineering PDP 12 computer. A second cassette drive was purchased and the author interfaced this drive to the computer. Although in theory this method of data transfer should be quite adequate, in practice it was fraught with problems. All these problems hinged around the inadequacy of the tape mechanism and electrical interference within the scanner corrupting data on the tape. The data error rate was so high and the system so unreliable, that funds were sought to enable the purchase of a more reliable digital cassette system. Funds were made available and a Racal Thermionic P72 Digideck was purchased. This digital tape system receives 8 bit bytes from the scanner and encodes this data serially on the tape. The actual tape format
was based on the specification of the Manchester Royal Infirmary which a year before had purchased a similar system. The main advantages offered by the new recorder were improved reliability and the provision of writing not only two channels of nucleonic data but also coordinate information on tape. This tape system is suitable for recording all data relevant to both rectilinear scanning and tomographic scanning. A comprehensive description of the Digideck data transfer system is described in chapter 2.

The two gamma cameras have comparable resolution. Because the MK V HR camera was purchased primarily as a research instrument most data have been taken from this camera. Using the low energy parallel hole collimator (NE 8922) and Tc99m a FWHM resolution of 1.2 cm. is obtained at a distance of 12 cm. in air. The cameras are equipped with video tape recorders and from the outset it was decided to use video tape as a data transfer vehicle between either gamma camera and the PDP 12 computer. Because the two gamma cameras use different tape recording formats an interface was constructed for the PDP 12 with a front end multiplexer so either type of tape can be replayed into the computer. If necessary the data could also be fed on-line to the computer, although for practical reasons this method has not been adopted. The
main objections to on-line connection lie in the facts that the computer and the gamma cameras are geographically separated, and that the PDP 12 is a general purpose single task computer system supporting a variety of projects, so that the practicalities of assuring computer availability during any patient scanning session are formidable. A full description of the gamma camera interface is presented in chapter 2.

1.3 DATA PROCESSING EQUIPMENT

An enhanced Digital Equipment Company PDP 12 computer system has been used to implement all the image processing methods discussed in this dissertation. The basic machine has 32K of 12 bit word 1.6 \( \mu \)s cycle time memory, a floating point processor (DEC:PPP12), a removable cartridge disc giving 1.6M words of on-line backing storage (DEC:RK8E) and, two Linc tape magnetic tape drives which are used primarily for archiving (DEC:TU56). Some of the auxiliary peripheral equipment which has been used in this study include the console teleprinter (DEC:LA36), an electrostatic plotter (Versatec:D1100A), a refresh type visual display unit (DEC:VR12), and a real time clock (DEC:KW12A). The operating system used is OS/8 V3C and most of the hundred or so programs developed during the course of this study are Fortran IV based. Assembly
language was used for several programs that have particular timing constraints; the real time data input of gamma camera data is an example where a Fortran program could not accommodate potentially high input/output data rates. Other Fortran callable assembly language routines were written to enable Fortran programs to have access to non standard peripherals such as the electrostatic plotter.

As described the basic PDP 12 computer system lacks the facilities to allow easy data entry from the J and P Engineering scanner and the gamma cameras, and there is no provision for an interactive high quality display to allow a user to examine the effects of manipulating image data within the computer. The refresh display is a bistable device so that grey tones have to be mimicked by repeated refresh or dot density of points. Section 3.2 will discuss in more detail the software written for this display.

The lack of an interactive high quality display led to the design and construction of a television system (Boardman 1976) which would be suitable for both grey tone and colour display. Section 3.3 is devoted entirely to this television system.
1.4 RESEARCH PROGRAMME

The project naturally falls into two parts. The first is primarily concerned with the design and development of both hardware and software to enable the computer system to be used easily and interactively in a nuclear medicine environment. This part of the project is therefore a study in computer engineering, computer programming and electronic engineering. Having assembled such a system the second phase of the programme consists of an examination of the problems of refocussing images and the development of solutions using Fourier transform techniques. This second phase is mainly mathematical but wherever possible a practical rather than a purely theoretical approach has been adopted.
2.1 PDP 12 INPUT/OUTPUT STRUCTURE

Before any of the specialised interfaces associated with this study are described it is appropriate if a slight digression is made to cover some of the basic principles of the PDP 12 Input/Output system so that the reader can more readily appreciate some of the intricacies of the various interfaces to be described later. A more detailed exposition of this section can be found in the computer reference manual (DEC 1970). The Input/Output system is also similar to that found on the DEC PDP 8 and 15, and Data General Nova computer systems.

All communications between external equipment and computer are accomplished through a computer instruction which characterises Input/Output operations and an associated address system which identifies external devices by a six bit binary number. The twelve bit Input/Output instruction format is shown in figure 2.1. Each peripheral piece of equipment has a device selector whose job it is to respond when the computer addresses that device. Once a device acknowledges that the computer
PDP 12 BIT ASSIGNMENT FOR I/O TRANSFERS

INSTRUCTION REGISTER:

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Equals 110
- Device Code
- Function (IOP pulses)

for IOT

Typical allocation of bits 9, 10, 11:

<table>
<thead>
<tr>
<th>Instruction bit</th>
<th>IOP Pulse</th>
<th>IOT Pulse</th>
<th>Used primarily for, but not restricted to</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>IOP1</td>
<td>IOT1</td>
<td>Flag checking, skipping</td>
</tr>
<tr>
<td>10</td>
<td>IOP2</td>
<td>IOT2</td>
<td>Clearing flags, clearing AC</td>
</tr>
<tr>
<td>9</td>
<td>IOP4</td>
<td>IOT4</td>
<td>Reading, loading registers</td>
</tr>
</tbody>
</table>

Any combination of bits 9 to 11 may be microprogrammed.

Figure 2.1
is addressing it, the least significant three bits of the instruction may be used to implement a particular function within the device.

As in most modern computers Input/Output transfers may take place either under program control or via data break transfers. The majority of transfers take place under program control. The main advantages of this method are that the interfaces are relatively easy to design and they are usually easily programmed. On the other hand the data break method is ideal when large amounts of data have to be transferred or if a very rapid response is required. The penalty paid is that the electronics in the interface are usually much more complex. Of the three interfaces to be described, the J and P Engineering cassette recorder and the television display use programmed transfers, and the gamma camera operates via data break.

All twelve bit programmed transfers take place via the accumulator of the computer. A device may strobe data into the accumulator or receive data from it. In order to ensure that data are transferred at the correct time the Input/Output instruction will also initiate the production of three timing pulses known as IOP 1, IOP 2, IOP 4, each of duration 1 μs. These IOP pulses are gated with the three least significant bits of the instruction in the
computer memory buffer so that their appearance in the device can be made conditional. There are also several auxiliary signal lines of which perhaps the most important is the skip line. This can be used to test the status of a device flag and to continue or skip the next instruction in the program based on the result of this test. This facility allows the computer program to determine whether a particular device is available to perform a specified operation.

The data break facility allows a device to transfer or receive data directly into or from the computer memory on a cycle stealing basis. The actual transfer takes place autonomously and so does not involve the computer central processing unit. Because the actual device controls information transfers the interface is necessarily more complex than that of a programmed transfer interface. There are several types of data break transfer of which only one, called memory increment, need be described here because it is relevant to the operation of the gamma camera interface. In this type of data break the contents of core memory at a device specified address are read into the memory buffer, incremented by one, and rewritten back at the same address within a 1.6 μs standard memory cycle. This feature is ideal for producing two dimensional histograms for data originating,
for example, from a gamma camera. In practice the total time to complete the transfer can exceed one memory cycle. This arises because the direct memory access on the PDP 12 is not truly cycle stealing, but must wait for the completion of an instruction before a cycle is made available. Minimum event processing time occurs when the break request is made 700 ns before the end of a standard cycle instruction. The worst case will occur when the break request is made 900 ns after the start of an Input/Output instruction, which uses an extended cycle of 4.5 µs, resulting in a transfer time of about 7 µs. These response times are much shorter than the times associated with the gamma camera video tape and analogue to digital converter system.

All interfaces have been fabricated using DEC equipment practice specifications. The printed circuit boards were made so that they could be housed in DEC expansion racks (DEC type H911K). Each rack is supplied with 64 blank printed circuit board edge connectors which are prewired with bus strips for power distribution. Wire wrap techniques have been used for interconnections between connectors and some special purpose logic circuit boards. Paddle boards with multi-cored ribbon cables have been used to connect the interfaces to the central processing unit. The logic used is TTL which is the same
as that used in the PDP 12.

Figure 2.2 contains pictures of one of the racking systems. The Input/Output bus from the computer enters the rack and before any of these signals are used they are fully buffered. This is done to prevent excessive loading on the signals from and into the computer. The buffered output signals are derived from open collector transistors so that in general several sets of interface signals can be connected in parallel; this is usually achieved by daisy-chaining devices together. The author has been involved in interfacing other pieces of specialised equipment to the computer, apart from those described here, and the buffering has prevented the occurrence of unacceptable pulse distortion.

2.2 J AND P ENGINEERING CASSETTE RECORDER SYSTEM

Although the J and P Engineering cassette tape recorder was interfaced last to the computer it is by far the simplest computer interface of those discussed here so it will be described first.

The Racal P72 Digideck records data as 8 bit bytes incrementally at speeds up to 350 bytes per second. The capacity of one cassette is approximately 50000 bytes.
I/O EXPANSION RACK

Figure 2.2

(1) GAMMA CAMERA       (2) DIGIDECK
The format of the data in each byte is shown in figure 2.3. Each byte is composed of bits PO0 to PO8 where PO0 represents the least significant and PO8 the most significant bits. Each complete datum word from the scanner is 16 bits long and comprises two consecutive bytes; bit PO0 indicates whether a byte is the first or second of a word. Bit PO1 of the first byte determines whether the word is numeric or control information.

In the case of numeric data PO2 defines one of two scanner data channels. The remaining bits of numeric data contain 12 bits so that each channel can accommodate values in the range 0 to 4095. The 12 bit data are obtained from 12 bit counters within the scanner. There is no protection against overflow from the counters and it is up to the user to ensure that the nucleonic count rate will not cause overflow. This problem is circumvented by judicious settings of the scanner pixel size and speed. When PO1 of a first byte is set the remaining bits are encoded as section angle, line coordinates or end of scan code as shown in the figure. The line coordinate word also allows three bits to be used for defining pixel sizes of 0.2, 0.4, or 0.8 cm. The control information is written on to the tape at the beginning of a new scan line. The end of scan code, of course, only appears at the termination of a scanning procedure.
SECTION COORDINATE:

<table>
<thead>
<tr>
<th>Byte 1</th>
<th>T0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byte 2</td>
<td>T7</td>
<td>T6</td>
<td>T5</td>
<td>T4</td>
<td>T3</td>
<td>T2</td>
<td>T1</td>
<td>1</td>
</tr>
</tbody>
</table>

T0->T7 = ANGULAR COORDINATE

LINE COORDINATES:

<table>
<thead>
<tr>
<th>Byte 1</th>
<th>X0</th>
<th>XI2</th>
<th>XI1</th>
<th>XI0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byte 2</td>
<td>X7</td>
<td>X6</td>
<td>X5</td>
<td>X4</td>
<td>X3</td>
<td>X2</td>
<td>X1</td>
<td>1</td>
</tr>
<tr>
<td>Byte 1</td>
<td>Y0</td>
<td>YI2</td>
<td>YI1</td>
<td>YI0</td>
<td>D</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Byte 2</td>
<td>Y7</td>
<td>Y6</td>
<td>Y5</td>
<td>Y4</td>
<td>Y3</td>
<td>Y2</td>
<td>Y1</td>
<td>1</td>
</tr>
</tbody>
</table>

X0->X7 = X COORDINATE
XI0->XI2 = X INCREMENT
Y0->Y7 = Y COORDINATE
YI0->YI2 = Y INCREMENT
D = DIRECTION

Figure 2.3 Part 1
J AND P ENGINEERING MAGNETIC TAPE FORMAT

DATA:

<table>
<thead>
<tr>
<th></th>
<th>P08</th>
<th>P07</th>
<th>P06</th>
<th>P05</th>
<th>P03</th>
<th>P02</th>
<th>P01</th>
<th>P00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byte 1</td>
<td>A4</td>
<td>A3</td>
<td>A2</td>
<td>A1</td>
<td>A0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Byte 2</td>
<td>A11</td>
<td>A10</td>
<td>A9</td>
<td>A8</td>
<td>A7</td>
<td>A6</td>
<td>A5</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>P08</th>
<th>P07</th>
<th>P06</th>
<th>P05</th>
<th>P03</th>
<th>P02</th>
<th>P01</th>
<th>P00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byte 1</td>
<td>B4</td>
<td>B3</td>
<td>B2</td>
<td>B1</td>
<td>B0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Byte 2</td>
<td>B11</td>
<td>B10</td>
<td>B9</td>
<td>B8</td>
<td>B7</td>
<td>B6</td>
<td>B5</td>
<td>1</td>
</tr>
</tbody>
</table>

A0→A11 = CHANNEL A
B0→B11 = CHANNEL B

END:

<table>
<thead>
<tr>
<th></th>
<th>P08</th>
<th>P07</th>
<th>P06</th>
<th>P05</th>
<th>P03</th>
<th>P02</th>
<th>P01</th>
<th>P00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byte 1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

FORMAT AFTER EACH PROCEDURE

Figure 2.3 Part 2
Figure 2.4 shows the schematic of the PDP 12 interface. Signals prefixed by INB or IBB are computer signals and those prefixed with JP are associated with the Digideck. Figure 2.5 contains a summary of the chosen instruction set together with a simple program sequence which may be used to read a byte from the cassette. The Digideck has been given device code 37 on the computer Input/Output bus. It should be noted that the program shows only the basic steps; as presented this program could "hang" if for example the cassette contained no data. A practical program should contain several safe-guards to give protection in the event of an error condition. The logical sequence of steps used to read a byte from the tape will be briefly described with reference to figure 2.5. Instruction DIGF (6372) will initiate the Digideck to fetch a byte from the tape. This instruction clears the DONE flip-flop and issues JP PI signal, which is the pulse that initiates the fetch cycle within the Digideck. During the fetching process instruction DIGS (6371) may be issued but it is not honoured because INB SKIP BUS L is always false while DONE is cleared. When the Digideck has returned a byte it will issue signal JP IDAV so that DONE is clocked to the true state, whereupon the issuing of instruction DIGS will now be honoured and the program will skip the following
FIGURE 2.4 Cassette recorder computer interface
DIGIDECK INSTRUCTION SET

<table>
<thead>
<tr>
<th>MNEMONIC</th>
<th>CODE</th>
<th>FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIGS</td>
<td>6371</td>
<td>Skip if done (ie. if ready)</td>
</tr>
<tr>
<td>DIGF</td>
<td>6372</td>
<td>Fetch byte</td>
</tr>
<tr>
<td>DIGR</td>
<td>6373</td>
<td>Rewind tape</td>
</tr>
<tr>
<td>DIGB</td>
<td>6376</td>
<td>Cassette buffer to accumulator</td>
</tr>
</tbody>
</table>

PROGRAM EXAMPLES:

(a) Initialisation:

| DIGB     | /Buffer to accumulator |
| SMA      | /Skip if minus         |
| JMP NOTRDY | /Not minus, so Digideck not ready |

(b) Fetching and reading a byte:

| DIGF     | /Fetch byte            |
| DIGS     |                          |
| JMP .-1  | /Wait for flag         |
| DIGB     | /Buffer to accumulator |
| SMA      | /Skip if minus         |
| JMP ERROR| /Not minus - an error condition |
| AND (377)| /Byte in bits 4 to 11 of accumulator |

Figure 2.5
instruction, indicating that data are available. Instruction DIGB (6376) can then be issued so that at IOP 2 the accumulator is cleared because INB AC CLEAR BUS L is true, and then at IOP 4 bits PO0 to PO8 are strobed into bits 4 to 11 of the computer accumulator using the INB I/O BUS lines. Bit 00 of the accumulator has also been used to define the Digideck status. When the recorder is in an operable condition this bit will be set. This enables a computer program to determine whether for example the recorder is plugged into the computer, or if there is a cassette on the transport. Assuming data transfer has been successful the accumulator may be deposited in memory and instruction DIGF may be issued again and so the cycle repeats. One further instruction, DIGR (6373), on IOP 2 will rewind the tape. The instruction is usually given after the program has detected an end of scan code.

2.3 DIGIDECK SOFTWARE

Several diagnostic assembler programs have been written for checking the performance and identifying possible faults with the recorder. Program DIGC is an example of such a program and an example of the output from the program is shown in figure 2.6. This program presents the user with a console listing of the binary bit
The above is a listing of a short portion of data from a cassette which contained a formatting error. Near the top of the listing the identification BAD PATN shows that the first byte from CHANNEL A is missing.
patterns of each pair of bytes, determines what type of data is present, and also presents both decimal and octal representations of numeric data. The program will check the bit patterns of consecutive bytes and determine whether there are any inconsistencies between the observed and expected data. This type of checking can be accomplished because often the format of a byte can be predicted from the history of preceding bytes. A careful study of figure 2.3 will show this to be true. DIGC along with several other diagnostic programs were proved invaluable during the commissioning of the Digideck system; the analyses of many incorrect bit patterns on the magnetic tape were able to localise teething faults within the scanner. To date, the interface itself has not failed since installation.

Most of the application assembler software for the cassette system has been developed as Fortran callable modules. In this way several users have been able to make immediate use of the Digideck without any knowledge of how the tape system functions or assembly language. An example of a Fortran subroutine is DIGDEK which is summarised in figure 2.7. When a user wishes to acquire data from the tape he simply calls the subroutine and the data and diagnostic information are returned to his program. The routine automatically reads two consecutive
DIGIDECK SUBROUTINE

This subroutine is written in RALF assembly language but is callable from Fortran IV as

CALL DIGDEK(ARG1,ARG2,ARG3)

The arguments are undefined on entry. On exit the arguments contain the following information:

ARG1 is an integer specifying the type of datum, which can take values from 0 to 8:
0 - Incorrect byte format, or the recorder indicates an error condition (STATUS = 0).
1 - End of scan.
2 - Section coordinate.
3 - X coordinate.
4 - Data channel A.
5 - Data channel B.
6 - Y coordinate.
7 - No data, indicating that a four second search for data has produced a negative result.
8 - Undefined bit pattern in second byte.

ARG2 is an integer which can take values between 0 and 4095 of type ARG1.

ARG3 is an integer specifying the X or Y increment. It is therefore only applicable if ARG1 equals 3 or 6.

Figure 2.7
bytes from the cassette and assembles them as one 16 bit word which is then separated out into its component values and the appropriate numbers are returned via the arguments in the subroutine call. The routines are all written in such a manner that they cannot fail; an error condition is always returned in the event of a recorder failure or if an incorrect bit pattern is detected. In the event of there being no data on the tape the subroutines will time-out when the search does not return data within 4 seconds. The DIGDEK subroutine also has an entry point to subroutine DIGRWD, which requires no arguments and will cause the tape cassette to rewind to the beginning of the tape.

2.4 GAMMA CAMERA INTERFACE

The NE MK IV and MK V HR gamma cameras have been interfaced to the PDP 12 using standard Nuclear Enterprises analogue to digital converter (P.C.B. 415) and video playback (P.C.B. 469) circuit boards that are normally used in the MK V HR camera. Basically these two boards enable the pulse width modulated signal containing X and Y coordinate information on the video tape to be replayed from the recorder, so that sets of two six bit bytes are obtained which define the X and Y positions of each detected scintillation within the gamma camera.
crystal. The main deficiency of this system is the relatively high dead-time associated with the video tape recording system. This dead-time is primarily due to the process of pulse width modulation, which requires a maximum of 22 µs per scintillation, and the 500 µs 50 Hz field sync pulses required for synchronising the inhibition of information during frame flyback. The dead-time of the gamma camera itself is approximately 7 µs. This time is comparable with the maximum data acquisition time of the direct memory transfer into the PDP 12 (section 2.1). As would be expected, during a video tape playback the observed count rates at the gamma camera console and count rates measured by the computer system for the same recording are for practical purposes identical. As an example of the losses imposed by the video tape system a true count rate of 10K counts per second produces an observed count rate of approximately 8K counts per second. Except for high activity dynamic studies the data loss is usually of no great hardship.

The gamma camera interface has been built with the following characteristics:
(a) An ability to collect two dimensional histograms in either 32*32 or 64*64 format. Unfortunately a resolution better than 64*64, such as for example 128*128, is not practical with a computer that has a 12 bit word format.
The coarse matrix of 32*32 format is sometimes useful for dynamic studies when only quantitative data are required from a procedure.

(b) If a location in the histogram overflows, that is if a number of events greater than 4096 is detected in any memory location, the interface will produce a program detectable overflow condition. A maximum count of 4096 is more than adequate for most clinical studies.

(c) A maintenance mode has been included. This enables a diagnostic program to check the functioning of the direct memory access interface without real gamma camera data. In the author's opinion this is an important and necessary facility because it is difficult to produce repeatable well defined gamma camera data for test purposes.

(d) The interface has the ability to identify from which of the two energy channels on the MK V HR camera a particular scintillation has arisen.

(e) The interface has a program switchable multiplexer which allows input from either the MK IV or MK V HR cameras.

Figure 2.8 shows the basic circuit diagram of the interface while figure 2.9 is concerned with the multiplexer and the LED displays used for indicating signal states. A photograph (figure 2.10) shows the front panel of the interface; because the functioning of the
PART OF GAMMA CAMERA INTERFACE

Figure 2.10
interface can be completely operated by software there are no manually operated controls. The interface itself is identified in figure 2.2

The interface has been given PDP 12 device code 31. A summary of the instruction set appears in figure 2.11. A brief description of the operation of the interface will be given by outlining the logic operations associated with each instruction. Named logic elements and signals will be found in figure 2.8.

GSKP (6311): In the event of any location in the computer memory overflowing, B WC OVERFLOW (0) H will be produced so that GC FLAG will be set. The output of GC FLAG directly controls the SKIP BUS. If the program interrupt facility was enabled (GEPI) then ENAB INT will be true so that the program will automatically vector to an interrupt address.

GLCR (6312): INTERFACE CONTROL is a 12 bit register within the interface that can be loaded from the accumulator at the occurrence of IOP 2. Bits 0 to 2 define one of eight 4K memory fields within the computer into which data are to be loaded. In the case of 32*32 histogram collection, bits 3 and 4 are used to define which 1K segment within a 4K memory field is to be used. Finally, bits 8 to 11 select the collection mode of the interface. Bits 8 and 9 define the type of camera and, in the case of the MK V HR
### G. MA CAMERA INSTRUCTION SET

<table>
<thead>
<tr>
<th>MNEMONIC</th>
<th>CODE</th>
<th>FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSKP</td>
<td>6311</td>
<td>Skip on overflow flag</td>
</tr>
<tr>
<td>*GLCR</td>
<td>6312</td>
<td>Load control register</td>
</tr>
<tr>
<td>GEPI</td>
<td>6313</td>
<td>Enable program interrupt</td>
</tr>
<tr>
<td>GEDB</td>
<td>6314</td>
<td>Enable data break</td>
</tr>
<tr>
<td>GMBR</td>
<td>6315</td>
<td>Maintenance data break</td>
</tr>
<tr>
<td>GLMR</td>
<td>6316</td>
<td>Load maintenance register</td>
</tr>
<tr>
<td>GCLR</td>
<td>6317</td>
<td>Clear all flags</td>
</tr>
</tbody>
</table>

* BIT ALLOCATION FOR GLCR:

<table>
<thead>
<tr>
<th>BIT</th>
<th>FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Extended memory address bit 00</td>
</tr>
<tr>
<td>01</td>
<td>Extended memory address bit 01</td>
</tr>
<tr>
<td>02</td>
<td>Extended memory address bit 02</td>
</tr>
<tr>
<td>03</td>
<td>Memory address bit 02</td>
</tr>
<tr>
<td>04</td>
<td>Memory address bit 01</td>
</tr>
<tr>
<td>05</td>
<td>Undefined</td>
</tr>
<tr>
<td>06</td>
<td>Undefined</td>
</tr>
<tr>
<td>07</td>
<td>Undefined</td>
</tr>
<tr>
<td>08</td>
<td>0 for MK IV, 1 for MK V</td>
</tr>
<tr>
<td>09</td>
<td>0 for MK V energy 1, 1 for MK V energy 2</td>
</tr>
<tr>
<td>10</td>
<td>0 for normal mode, 1 for maintenance mode</td>
</tr>
<tr>
<td>11</td>
<td>0 for 64<em>64 matrix, 1 for 32</em>32 matrix</td>
</tr>
</tbody>
</table>

Figure 2.11
camera, which energy channel is to be used. The signals ISOTOPE 1 H and SELECT MKV H are generated which are routed via connector A28 to the multiplexer circuit board (figure 2.9). Bit 10 defines collection from a gamma camera or simulated data collection using the maintenance mode. Bit 11 controls the matrix size. The memory addressing system used for the two matrix sizes is shown in figure 2.12. Modules A26, B25 and B26 act as three line to one line multiplexers. Each line of each pole of the multiplexer can receive data from either pair of analogue to digital converters or from a maintenance register. For 64*64 data collection the 6 bit X coordinate data select the least six significant memory address lines (EXT DATA ADD 06L-11L) and similarly the Y data select the most significant memory address lines (EXT DATA ADD 00L-05L). A similar ten bit addressing system operates for 32*32 data collection except that EXT DATA ADD 00L-01L of the memory address are now determined by bits 3 and 4 of the GLCR instruction.

GEPI (6313): This instruction will set the ENAB INT flip flop so that an overflow will produce a program interrupt. GEDEB (6314): This instruction will set the INENBL flip flop, so enabling the BRK REQ flip flop to be set when direct memory access is required by an analogue to digital converter or a maintenance instruction. After a break request has been honoured by the PDP 12 it will return
GAMMA CAMERA MEMORY ADDRESSING FORMAT

64*64 DATA COLLECTION:

<table>
<thead>
<tr>
<th>Y ADDRESS</th>
<th>X ADDRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>00 01</td>
<td>64 65</td>
</tr>
<tr>
<td>00 01</td>
<td>127 127</td>
</tr>
<tr>
<td>00 01</td>
<td>63 63</td>
</tr>
</tbody>
</table>

X ADDRESS
(6 bits)

MEMORY ADDRESS:

| 00 01 02 03 04 05 06 07 08 09 10 11 |

Y ADDRESS X ADDRESS

Addresses are relative to the start of a 4K memory field.

Figure 2.12 Part 1
GAMMA CAMERA MEMORY ADDRESSING FORMAT

32x32 DATA COLLECTION:

<table>
<thead>
<tr>
<th>Y ADDRESS (5 bits)</th>
<th>X ADDRESS (5 bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>32 33</td>
<td>63</td>
</tr>
<tr>
<td>00 01</td>
<td>31</td>
</tr>
</tbody>
</table>

MEMORY ADDRESS:

<table>
<thead>
<tr>
<th>00 01 02 03 04 05 06 07 08 09 10 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>* Y ADDRESS</td>
</tr>
<tr>
<td>X ADDRESS</td>
</tr>
</tbody>
</table>

*Bits from control register

Addresses are relative to the start of a 1K memory segment.

Figure 2.12 Part 2
signal IOO ADD ACCEPTED (0) H which will reset BRK REQ so that further requests can be made.

GMBR (6315): Providing maintenance mode has been selected using the GLCR instruction then the issuing of GMBR will produce a break request at IOP 4. In this mode the PDP 12 central processor can be made to produce its own break request; the processor effectively inhibits itself from executing further instructions while it steals its own cycle.

GLMR (6316): At the occurrence of IOP 4 the accumulator is loaded into the 12 bit maintenance register. When instruction GMBR is issued the maintenance register specifies an address within a 4K memory field whose contents will be incremented by one.

GCLR (6317): The instruction resets or clears the interface flags. The skip, interrupt and data break enable flip flops are cleared at IOP 1.

For most routine use once the interface has been enabled there is virtually no software overhead (see figure 2.13). Indeed the central processing unit can be performing other tasks throughout the data collection period. The gamma camera programs make use of this time for updating the television display (chapter 3) with the current scintigraphic data, and buffering data onto disc in the case of dynamic studies.
INSTRUCTION SEQUENCE TO ENABLE GAMMA CAMERA DATA INPUT

In this example only four instructions are required to set up data collection into memory field 0 from the MK IV camera using a 64*64 matrix size:

GLCR /Clear interface
CLA /Clear accumulator
GLCR /Load accumulator into control register
GEDB /Enable data break
/
/From this point on data will be accumulated automatically without /program intervention.

Figure 2.13
In order to complete a description of the interface it is instructive to consider the sequence of events that occur after a scintillation has been detected and the analogue to digital converters have assembled the appropriate X and Y coordinates. When the analogue to digital converters have data available in their output registers the interface multiplexer produces ADC AVAILABLE H (figure 2.9) which, providing maintenance mode is not selected, will set the BRK REQ flip flop. At the end of the current PDP 12 cycle the break request will be honoured and IOO ADD ACCEPTED (0) H will reset BRK REQ. EXT INCREMENT MB L will be true so that the memory buffer control will generate a carry insert signal with the result that the memory buffer, whose contents are from an address determined by the value of the analogue to digital output registers, will be incremented. After the data break cycle, signal IOO B BREAK (0) H will be asserted so that the analogue to digital converters are released for another event.

2.5 GAMMA CAMERA SOFTWARE

An assembly language gamma camera diagnostic program has been written which tests the basic functions of the interface system. All the tests are performed with the
interface logic switched into maintenance mode. The program checks the data break request facility in different memory fields. A specific data pattern is written in each memory field using the interface and afterwards the data are checked, by software routines, against the expected results. Any discrepancies are logged on the computer console. If these tests are completed satisfactorily each memory location is incremented past 4095 to check the overflow facility. In practice this software is rarely run because after three and half years use the interface has not yet presented a fault condition.

Interface maintenance mode does not provide a facility for checking the Nuclear Enterprises circuit boards. To enable fault finding on these boards one dual-in-line two pole change over switch has been included on the multiplexer circuit board (figure 2.9) which, when placed in the test position, allows the analogue to digital converters to free run. This is achieved by routing the NE ADC AVAIL IV H and NE ADC AVAIL V H signals back to the respective NE MKIV RESET H and NE MKV RESET H lines. In this configuration the computer is completely passive. Hence an oscilloscope can be used to monitor signals in the analogue to digital converters without consideration of the asynchronous timing constraints
imposed by the computer.

Two basic data collection programs have been written: one for static images called GCS and one for dynamic studies called GCD. The timing constraints imposed by gamma camera data capture and problems of operating direct memory access devices as a background task within a Fortran program necessitated the development of assembly language programs for all gamma camera data collection. GCS will collect one image at a time. During data collection the user is kept informed about the nature of the image using the grey tone/colour television display. After data collection is complete each twelve bit integer in the two dimensional histogram is floated into the three word format used for OS/8 Fortran IV variables and saved in a Fortran compatible random access file. Hence, apart from the actual data collection, all processing on the data can be Fortran based. A general purpose Fortran scintigram display program allows stored data to be retrieved and manipulated before output to the television or electrostatic printer/plotter displays.

Program GCD is similar in nature to GCS except that it allows the collection of multiple images or frames of data. The computer real time clock is used to control frame collection times from 200 ms up to several minutes.
The lower limit of 200 ms is determined by the maximum time required to write the contents of one 4K memory field onto disc. Two memory buffer areas are used to collect data: when one memory field is being used to collect data the other is written onto the disc, so that when a frame period is terminated the memory fields are switched and the last completed frame is always written onto the disc. The data are written onto the disc as 12 bit integers because this format requires the minimum of time. During non critical periods the television is updated with data from the current memory field being used for data collection. There is a relatively small data collection dead-time when input has to be switched from one memory field to another. However this software dead-time amounts to less than 100 µs, when using the program interrupt facility of the real time clock, and is insignificant compared with the minimum allowable frame period. After data collection is complete the integer data are read back from the disc into memory one frame at a time, floated, and written back as a Fortran compatible random access file. Existing software allows 67 64*64 frames to be written on one half the usable area of the disc. By choice, the remaining disc area is dedicated to operating system and user programs.
Whenever possible data collection from the J and P Engineering scanner is kept to a 64*64 format and the data are stored as Fortran random access files. In this way, most of the image processing programs can be used either on scanner or gamma camera data.
CHAPTER 3

DISPLAY EQUIPMENT

3.1 INTRODUCTION

One of the problems of digital signal processing is that it is usually necessary to visualise the results of processing in some way. Several different types of device have been tried to achieve this end. The choice of an optimum display for the presentation of radionuclide scintigrams has been a subject of much debate (Budinger 1972, Clifton 1971, Todd-Pokropek 1972). The merits of different systems have recently been summarised by Todd-Pokropek (1976). The main requirement for the PDP 12 system was to provide an interactive computerised cathode ray tube display capable of producing grey tone images. In general most workers favour television type displays for interactive work. With the advent of fast solid state memories for raster refresh purposes, relatively inexpensive television systems are now practical. Hard copy images can be obtained by photographing the television screen, or other devices such as electrostatic plotters or incremental plotters can be used. The IAEA intercomparison of image processing techniques (IAEA 1972, IAEA [1] 1976, IAEA [2] 1976) has shown that while
incremental plotter results do not look similar to the original type of image obtained from a gamma camera display, they are of relatively high diagnostic value. Both the electrostatic and incremental plotter methods overcome the problems of photography, which may be tedious, or in the case of Polaroid film, expensive.

3.2 REFRESH OSCILLOSCOPE

A refresh oscilloscope, such as the PDP 12 VR12, can be used as a scintigram display. This display has to be refreshed by the computer program if a permanent image is to be maintained. The oscilloscope is a bistable device that requires approximately 23 μs for the completion of a single point. The bistable nature of the display means that brightness exists as a two valued intensity, either on or off, so that grey tone reproduction has to be mimicked by repeated refresh of points. Software has been written so that a single point can be displayed a maximum of seven times per frame so that up to eight grey levels can be produced. The resolution of the two digital to analogue converters on the PDP 12 is nine bits, so that there are 512 addressable dot positions along both X and Y directions. For presenting scintigrams on a 64*64 grid each pixel in the image can consist of a square array of 64 dots. Simple calculations show, that for a picture
containing 4000 pixels with each pixel composed of an 8*8 dot matrix, many seconds are required to produce a single complete frame. While this is quite adequate for time exposure photography, it is impossible to use as an interactive computer display. A further handicap is the fact that the VR12 is a low quality device not designed for the reproduction of high quality images. Although the display is driven by nine bit digital to analogue converters which might suggest that resolution is more than adequate for scintigraphic work, the deflection amplifiers and power supplies in the oscilloscope are inherently noisy. The noise manifests itself as jittering of each point on the cathode ray tube face. Figure 3.1 shows an image of a 64*64 data set. The data are linearly scaled to eight levels, and then each pixel is refreshed a number of times directly proportional to the scaled grey level, and the result recorded using time exposure photography. Using Polaroid 107C film (3000 ASA) a total exposure time of about 10 seconds is required to produce a complete picture. The result is not particularly good by modern display standards and the effects of electronic noise produce a mosaic texture to the picture.

An alternative method of mimicking brightness is to use single frame exposure and an arrangement whereby
Figure 3.1
density of dots is proportional to the intensity of the picture. Dots can be distributed geometrically or randomly in each pixel. The latter approach has the effect of removing the cellular appearance of the picture and the result looks more like a conventional gamma camera picture. However, this is probably a retrograde step as quantification of intensity between regions in a picture becomes more difficult for the observer to judge. These methods along with several others have been used on the VR12, but none match the quality demonstrated in figure 3.1. The data are from a scan of the D.H.S.S. William's phantom.

3.3 TELEVISION DISPLAY

The inadequacies of the PDP 12 cathode ray tube display led to the design and construction of a television display. A television picture provides a flicker free grey tone image which can be used interactively from the computer console. Using the system to be described, data can be manipulated within the computer easily and output to the television as required. As a bonus, a small amount of extra electronic circuitry enables pseudo-colour images to be produced. Pseudo-colour in the image processing literature merely implies a colour scale in which colours are associated with a predefined data scale. In
scintigraphy colour displays have become popular among some clinicians because colour helps to quantify relative radioactive count rates.

The television system described here uses a standard British Radio Corporation Series 3500 chassis (such as Ultra, H.M.V., Ferguson) or a Sony KV1350 receiver to display a 64x64 image. To save unnecessary confusion only the video circuits used to drive the Sony receiver will be described. The modifications required in either type of receiver are trivial. The cost of the electronic components was about £300 in 1974 and a further £200 was required for the purchase of each television receiver.

3.3.1 SCANNING SYSTEM

The British broadcasting television systems use a conventional 2:1 interlaced raster scan system based on the use of 625 scanning lines. Each field is scanned from top to bottom of the picture in 20 ms so with interlacing a complete television picture or frame is produced at a 25 Hz rate. In this system 64 rows of data are displayed as 512 raster lines. The remaining raster lines are kept at black level and are used for frame flyback. Because interlacing is employed the complete image is produced on 256 lines per field so that during the scanning process
each row of data is repeated onto four raster lines.

Each raster line contains information pertaining to 64 data values. The television line period is 64 µs, but 12 µs of this is usually allowed for line flyback. In this system information is displayed in a 32 µs portion of a line period. This has the effect of producing an approximately square aspect ratio on the cathode ray tube and, with slight adjustment of the receiver height control, a perfectly square aspect ratio can be obtained.

3.3.2 MEMORY REFRESH SYSTEM.

Figure 3.2 shows a block diagram of the display system. It is a bare bones representation of the basic units in the system. The diagram indicates the data flow pattern and in order to prevent confusion no timing or control signals have been indicated.

The bulk of the electronics comprises three memories and a memory multiplexer. The Main Memory and the Line Memory are shown in figure 3.3. The former memory consists of 4K words. Each word has a length of 4 bits. The memory consists of serially connected static shift registers (Fairchild 3355). The Line and Interface Buffer Memories are two auxiliary memories each of 64 words.
FIGURE 3.2 Block diagram of television system
FIGURE 3.3 Main and Line Memories
(Signetics 2518). The Memory Multiplexer has three input ports, so that the Video Control section can receive data from any one of the three memories.

To understand the operation of the individual units shown in figure 3.2 it is instructive to consider the sequence of events assuming, for the moment, that the first field of a new frame is about to be displayed, and that the Main Memory has already been loaded with data. After 32 lines, which are held at black level, data from the Main Memory are clocked out through the Memory Multiplexer to produce signals VIDEO 00 to 03 H which are routed to the Video Control section. After suitable processing this information is used to drive the cathodes of the television cathode ray tube. At the same time as data leave the Memory Multiplexer the information is also clocked into the input of the Main and Line Memories. During the next three raster lines information is taken from the Line memory by appropriate switching of the MEM MULX 00 and 01 gating signals. In this period the Main Memory clock (MEM SHIFT A and B H) is inhibited and data circulate only around the Line Memory. On the fifth line the Main Memory port of the multiplexer is opened again and the second row of information is clocked out for display. This line is followed by a three line refresh from the Line Memory. This sequence of events repeats
until 64 rows of data, or 256 raster lines of information have been displayed.

The remaining lines in the field are held at black level. The second field of the frame is displayed in exactly the same way; only the line and frame synchronising signals are adjusted so that an interlaced picture is produced.

The above discussion assumed that data had already been loaded into the Main Memory. To input new data into the display memory from the computer the Interface Buffer Memory is used. This memory forms part of the interface which is described in more detail in section 3.3.5. For the present discussion it is sufficient to say that it can be loaded under program control from the computer accumulator. Data from the Interface Buffer Memory arrives at the Memory Multiplexer as signals INT ROW 00 to 03 H. When the appropriate row is about to be displayed the Interface Buffer Memory contents can be clocked out through the Memory Multiplexer into the Main and Line Memories. The previous contents of the Main Memory will be clocked out and lost. During this process the normal television refresh cycle is not interrupted and data appear to change instantaneously on the television screen.
3.3.3 TIMING CONTROL

To save the construction of a television frame and line sync generator a commercial unit made by Manor Engineering was used. This uses a 2 MHz crystal controlled master clock which is divided down to produce the appropriate frame and line timing signals. As a bonus the equipment can also be used as a cross-hatch generator which is necessary for accurate alignment of the dynamic convergence circuits of a television receiver. The display system master timing signals are shown in figure 3.4. Three signals, FRAME PULSE, LINE PULSE and 2 MHz CLOCK are taken from the cross-hatch generator via three coaxial cables. Noise is removed from the signals by three Schmitt trigger circuits. FRAME PULSE H is of 800 μs duration but FRAME H and FRAME L are reduced to 250 ns by monostable B9/E1. LINE PULSE H is 12 μs wide and this is also reduced to 250 ns and is available as LINE H and LINE L. Figure 3.4 also contains the memory clearing circuitry. When the power is turned on to the system A9/E3 will always fire after a delay produced by the charging of the 220 μf capacitor, so that the memory is automatically cleared when power is applied. A computer instruction can also be used to clear the memory by the generation of CLEAR PULSE L which will produce one
FIGURE 3.4 Master timing and memory clear
pulse from A9/E1. The durations of the quasi-stable states of A9/E1 and A9/E3 are greater than 20 ms (ie. one television frame period) and as MEM CLR (1) L deselects the Main Memory inputs (figure 3.3) the effect is to toggle zeros into the memory registers.

The master timing signals are used as the input to the horizontal and vertical zone circuits (figure 3.5). The horizontal and vertical timing zones define an active area on the television screen in which data can appear. The diagram in figure 3.6 shows the active zone on the television screen and its relationship to frame and line timing. The active zone occurs when signal VIDEO UP H is true. VIDEO UP H is obtained when both the V ZONE and H ZONE flip flops are set. At the beginning of a new frame the eight bit counter (A10) is cleared and then after 32 line pulses V ZONE will be set by A10/C1 going high. After a further 256 line pulses, overflow from the counter will reset V ZONE to its false state. A second eight bit counter (B10) is used to control the H ZONE flip flop. At the start of a new line the counter is cleared and after 48 clock pulses (24 µs) H ZONE is set and after a further 64 pulses (32 µs), during which data are displayed, it is reset. One further signal VEND (1) H is a 250 ns pulse produced at the end of the V ZONE period which is required by the computer interface. Figure 3.5
FIGURE 3.5 Horizontal and Vertical Zone circuits
TELEVISION RASTER ACTIVE ZONE PERIOD

1 Field = 312.5 raster lines

The duration of one field is 20 ms, so that each complete line period is 64 µs.

Figure 3.6
also shows that the memory clock pulses, MEM CLK L, MEM SHIFT A H and MEM SHIFT B H are obtained by gating the 2 MHz clock with the VIDEO UP H signal. Separate memory clocks are required because the Main Memory is only clocked every fourth raster line. The CHANGE DATA multiplexer generates the signal which inhibits the clocking of the Main Memory. If new data are to be introduced into the Line and Main Memories then ROW SELECT (1) H from the computer interface will be enabled for the duration of the appropriate line scan.

3.3.4 VIDEO CIRCUITRY

The Video circuits receive the four bit binary data from the Memory Multiplexer. The data are processed in one of two ways. For displaying grey tone pictures the digital information is converted into analogue form so that the same signal is applied to the red, green and blue guns of the cathode ray tube. For displaying colour pictures the three most significant bits of the data are fed to a colour decoder which produces the appropriate weightings of the red, green and blue components of the image. Seven colours plus black are available for display.
A brief description of the video circuits will be given with reference to figure 3.7. Data from the Memory Multiplexer are taken directly to a 4 bit latch (El). Data are strobed in about 20 ns after the beginning of each MEM CLK H pulse. This arrangement helps to overcome skew problems between different bits of information arriving from the memories at slightly different times. In practical terms it cleans up the vertical edges of the pixels on the television screen. At the end of each raster line it is important that the latch is reset to zero so that black level is maintained on the following line before the first pixel position. The reset is obtained by producing a short pulse from the falling edge of the VIDEO UP H signal.

For displaying colour pictures, COLOUR (1) H from the interface will be true so that E8, a three line to eight line decoder, is enabled. Only one output of the decoder is active at a time, and the resistors connected to each output provide the correct current contributions for the red, green and blue summing amplifiers. The ordering of colours with respect to numerical value is arbitrary as adjustment of the decoder weighting resistors will permit any sequence. Figure 3.8 shows the sequence adopted which is based on the order in which colours occur in the electromagnetic spectrum. This is an unambiguous colour
FIGURE 3.7 part 1  Colour and B/W decoders for Sony receiver
Relative primary colour weights used to produce a constant saturation colour scale

<table>
<thead>
<tr>
<th>Datum Value</th>
<th>Colour</th>
<th>Red (%)</th>
<th>Green (%)</th>
<th>Blue (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 or 1</td>
<td>Black</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2 or 3</td>
<td>Magenta</td>
<td>50</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>4 or 5</td>
<td>Blue</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>6 or 7</td>
<td>Cyan</td>
<td>0</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>8 or 9</td>
<td>Green</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>10 or 11</td>
<td>Yellow</td>
<td>60</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>12 or 13</td>
<td>Orange</td>
<td>70</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>14 or 15</td>
<td>Red</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3.8
scale which clinicians have found acceptable.

For displaying grey tone pictures COLOUR (1) H will be false so that the decoder is disabled. However COLOUR (0) H will be true so that during the active zone of the picture E2 and E3 will transfer data to transistors T8 to T11 which form a four bit digital to analogue converter. Collector clamping is used on these transistors so that clean TTL levels are generated. T1 is used to sum the weighted currents derived from each bit of the four bit word. T1 provides equal drive for each colour summing amplifier so that a grey scale can be produced on the television screen.

The colour circuits comprise transistors T2 and T3, T4 and T5 and, T6 and T7. Each pair of transistors is used to drive one of the video output transistors in the Sony receiver. The circuits used for each primary colour are identical. Each circuit comprises a grounded base summing transistor (T2, T4 and T6) followed by a common collector transistor (T3, T5 and T7) which provides a 47 ohm output impedance for driving the receiver via a coaxial cable. Each output stage has a preset gain control (VRr, VRg and VRb) on the circuit board so that the receiver can be set up for correct grey scale tracking. If a good grey tone picture is produced, a
colour display will automatically be set up for correct hues and saturation. The dc. potential of each output is about +6V; this is required for correct biasing of the video output transistors in the Sony receiver. This bias level need only be approximate because the receiver brightness control can provide control from cut-off to saturation of the cathode ray tube.

3.3.5 COMPUTER INTERFACE

A set of instructions enables the television to be controlled completely from the computer. The interface therefore has the provision for loading and clearing the television memories, and switching its mode from monochrome to colour and vice versa. All these operations have to take place so that normal television raster scanning is always maintained. This implies that the computer, via the interface, must be able to synchronise itself to the line and frame rates so that data are loaded correctly into the Main Memory. For this to occur the Interface Buffer Memory has to be arranged so that it may be clocked at either the television rate or at the slower computer rate.
The circuit diagrams of the interface are shown in parts 1 to 3 of figure 3.9 and a summary of the instruction set is listed in figure 3.10. An understanding of the operation of the interface can be obtained by considering data transfer from computer to interface, followed by transfer from interface to television system. Data are transferred from the computer accumulator using a bit allocation defined in figure 3.10. To transfer a complete row of 64 data values instruction TLAI (6327) has to be issued 64 times. The TLAI instruction makes use of all three computer IOP pulses.

At IOP 1 INT LOAD flip flop is set so that the 8 bit ROW ADDR counter is enabled. At IOP 2 a clock pulse is routed via SHIFT SELECT to the counter so that data bits 00 to 05 are loaded into the counter. At the same time CLOCK SELECT will provide a clock pulse for the Interface Buffer Memory so that data in bits 08 to 11 of the accumulator are toggled in parallel into four static shift registers. Finally at IOP 4 INT LOAD is cleared. This completes the loading of one row of data into the interface. At this point the interface contains 64 four bit words in its shift registers and the row address contains a number between 0 and 63. It should be noted that although the television field period starts at the top of the screen the matrix rows are numbered from 0 to 63 from the bottom upwards. For this reason the row
FIGURE 3.9 part 1  Television computer interface
(Row addressing and timing logic)
FIGURE 3.9 part 2  Television computer interface
(Memory and colour selection logic)
FIGURE 3.9 Part 3 Television computer interface (Device selector)
TELEVISION INSTRUCTION SET

<table>
<thead>
<tr>
<th>MNEMONIC</th>
<th>CODE</th>
<th>FUNCTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSKP</td>
<td>6321</td>
<td>Skip if busy</td>
</tr>
<tr>
<td>TCMS</td>
<td>6322</td>
<td>Television colour mode switch</td>
</tr>
<tr>
<td>TBUF</td>
<td>6323</td>
<td>Load television buffer</td>
</tr>
<tr>
<td>TCLI</td>
<td>6324</td>
<td>Clear interface</td>
</tr>
<tr>
<td>TCLM</td>
<td>6325</td>
<td>Clear memory</td>
</tr>
<tr>
<td>*TLAI</td>
<td>6327</td>
<td>Load accumulator into interface</td>
</tr>
</tbody>
</table>

*BIT ALLOCATION FOR TLAI:

<table>
<thead>
<tr>
<th>00</th>
<th>01</th>
<th>02</th>
<th>03</th>
<th>04</th>
<th>05</th>
<th>06</th>
<th>07</th>
<th>08</th>
<th>09</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FRAME ROW ADDRESS  DATUM VALUE

Figure 3.10
address counter is loaded with the row number and not its
two's complement. The motivation for this convention
stems from the fact that the gamma camera addressing
system and other Department of Medical Physics and Medical
Engineering equipment use the bottom left corner as the
origin for an array of data.

Instruction TBUF (6323) is used to transfer data from
the interface into the display itself. At IOP 1 TRANS REQ
is set and this will enable REQ to be set at the beginning
of a new television field. No further signals are
generated until the active zone of the display is reached,
whereupon SHIFT SELECT will route the MOVE DATA H signal
to increment the row address (ROW ADDR) counter. This
process repeats until at some point the counter will reach
63 with the result that the ROW SELECT flip flop will be
set, and further MOVE DATA H pulses are prevented from
incrementing the counter. ROW SELECT (1) H will open the
CLOCK SELECT multiplexer so that 64 MEM CLK L pulses from
the display can toggle the Interface Buffer Memory
contents into the Line Memory and the selected row of the
Main Memory. During this process the previous contents of
the same row in the Main Memory are lost and data appear
to change instantaneously on the screen. The REQ flip
clop will be cancelled by INT RESET L signal, and at the
commencement of the following line LINE H will cancel
ROW SELECT (1). To prevent further transfers taking place during the current field the TRANS and BUSY flip flops are only reset at the end of the vertical zone period using V END (1) H. This ensures that instruction TSKP (6321) will only indicate that the display is ready to receive more data at the end of the current field. This requirement is necessary because the row address counter must always count from the first row in a field, so that only one row of data may be transferred per field period.

In order to switch the mode of the display from monochrome to colour an instruction, TCMS (6322), has been incorporated. Each time the instruction is issued the COLOUR flip flop will toggle from one state to the other. Instruction TCLI (6324) provides a reset pulse for the whole display. At IOP 4 it will generate INT RESET L and also force colour mode. Finally instruction TCLM (6325) will erase the contents of the Line and Main Memories as discussed in section 3.3.3.

3.3.6 PDP 12 - TELEVISION SIMULATOR

The development of any relatively complex piece of electronic equipment needs a certain amount of time for testing and elimination of teething problems. In order to commission and fully analyse the operation of the
television system it was necessary to design and construct a PDP 12 - television simulator, which enables the display system to be operated without encroaching on valuable PDP 12 computer time. The simulator emulates certain signals that would normally be generated by the PDP 12 and will load a predefined data pattern into the Interface Buffer Memory. The whole test unit is portable and mounted in a diecast box. The normal Input/Output bus lines of the television system are plugged into the simulator.

The simulator box contains switches to preset bits 00 to 05 of the accumulator, bits 09 to 11 of the memory buffer, and one switch labelled EXECUTE which will execute the instruction set in the memory buffer switches. Bits 00 to 08 of the memory buffer are prewired to 632(8); this represents the three most significant octal digits in every Input/Output instruction associated with the television. The schematic of the simulator is shown in figure 3.11. Providing instruction seven (simulated 6327) is not being issued the action of pressing the EXECUTE switch is to produce three pulses from the chain of monostables comprising E1, E2 and E3, which are gated with memory buffer bits 09 to 11 to produce the respective IOP 1 H, IOP 2 H and IOP 4 H pulses. The timing relationships of these pulses are identical to that
FIGURE 3.11 part 2  PDP 12 - Television simulator
normally produced by a PDP 12 Input/Output instruction.

The simulator gives a slightly different interpretation to instruction code 6327; it issues this instruction 64 times each time the EXECUTE switch is pressed. Hence the whole Interface Buffer Memory can be loaded with a single closure of the switch. The way this is achieved is as follows. If memory buffer bits 09 to 11 are all set to ones the CLOCK ENABLE flip flop is set so that E4, a free running 1K Hz clock, is gated into the clock input of an eight bit counter (E6 and E7). Each clock pulse produces EXEC H so that a normal set of IOP pulses are produced. After 64 clock pulses CLOCK ENABLE will be reset so inhibiting the production of further Input/Output instructions. Three bits of the counter are connected to bits 8,9 and 10 of the simulated accumulator data lines, so that numbers between 1 and 15 are loaded in groups of eight into the Interface Buffer Memory. On the display the result will be a set of colour bars or grey tones, each eight pixels wide, ranging from black on the left to red (or white on a grey scale) on the right of the television screen. All the other simulated instructions behave normally. Figure 3.12 shows the test bar patterns produced by the simulator.
Figure 3.12
3.3.7 SOFTWARE

Little diagnostic software has been necessary because the simulator can be used to produce test data. However similar test patterns can be loaded into the display using assembly language programs. One program in particular, which produces a repeating square grid pattern on the screen, is useful for setting up the aspect ratio of the display and checking the convergence of the primary colours.

Application software consists of a Fortran callable routine which enables any row in the display memory to be loaded from a Fortran array. This subroutine also has other entry points so that the memory and the mode of the display can be controlled. Using Fortran IV a 64*64 array can be transferred to the display in about five seconds. This compares with about one second for assembly language programs. Of course once the data have been transferred there is no CPU overhead. Figure 3.13 shows two photographs taken from the television screen of scintigraphic data. The data is the same as used in figure 3.1.
Figure 3.13
3.3.8 SUMMARY

The display system is a prototype design which has now given reliable service for three years. The ease with which the display can be used from Fortran has proved most popular amongst inexperienced programmers. The display uses a programmed transfer interface which, because of the large amounts of data transferred, would be better implemented by direct memory access. Unfortunately the lack of a spare memory port on the current PDP 12 configuration prevented this approach. The display circuits are modular in design and could relatively easily be upgraded. For example, the incorporation of a 16K memory running at twice the clock frequency would enable the presentation of 128*128 matrices. The number of colours could be doubled if four bit decoders were employed. A second display system is currently being designed.

3.4 ELECTROSTATIC PLOTTER DISPLAY

Although the television display is used for interactive image processing sessions with the computer, an electrostatic plotter has been connected to the PDP 12 to provide a hard copy facility. It should be noted that
image interpretation is usually done using the television display; the hard copy output is used primarily as a convenient easy to use method of producing grey tone images for inclusion, for example, in a patient's case notes. The software is written such that processed data, as displayed on the television, may be output to the plotter.

The plotter is capable of producing black dots on dielectrically coated paper. The output is therefore two valued (on or off) so that grey tones have to be mimicked by dot density. Electrostatic displays have been criticised in the past for not being capable of producing large uniform black areas (similar to the differentiating property of the Xerox process) with the result that images tended to lack sufficient contrast. However modern instruments have dual array writing heads which produce overlapping dot patterns which give greater density and contrast. Although the initial capital expenditure is relatively high the running costs are substantially cheaper than, for example, Polaroid film. Multiple copies of a picture can be produced quickly. Picture annotation can be produced by either software or hardware character generators.
The plotter used is a model D1100/A manufactured by Versatec. This instrument uses 11 inch wide paper and has a dot density of 100 nibs per inch, with a total of 1024 across the page. The paper drive operates in increments of 0.01 inch so that square aspect images are automatically produced. The plotter can be operated in a printing mode where a read-only-memory (ROM) character generator is used to produce a 7*9 dot matrix for each character. Up to 132 characters can be printed across the page at speeds up to 500 lines per minute. The Versatec plotter therefore fulfils several needs on the computer system. It can be used as a fast alpha-numeric listing device offering all the facilities of a typical line printer, it may be used to produce grey tone images, and finally it can be used as a graphics hard copy device (see for example figures 4.5 to 4.8). In this context the electrostatic plotter is a cost effective computer peripheral.

The computer interface for the plotter was also purchased from Versatec. The software for producing grey tone images has been developed by the author and this will be described in some detail. Several different grey tone shading arrangements have been tried. Not all of them aim at producing a good visual grey tone scale; for example, a numerical output has been used when more quantitative
representations of data are required. The software to be described obviously involves the manipulation of bit patterns to produce grey scale tones which is very time consuming, and sometimes impossible to do, using a high level language such as Fortran. For this reason all the grey scale software has been coded as Fortran callable assembly language routines.

3.4.1 REGULAR GREY SCALE FOR REPRODUCING 64*64 IMAGES

A regular grey scale is used to describe a grey scale that uses a set of predefined shading patterns for each level of density. The image displayed consists of a set of black dots displayed on a 512*512 grid. So for a 64*64 format image, each pixel is composed of dots distributed within an 8*8 sub-array. This means that in theory 65 different dot densities are available. In practice, for a typical scintigram which will contain noise, eight levels have been found to be adequate. Generally speaking it has been found difficult to see more than about eight identifiable levels for any toning arrangement. The exception is for noise free data, as for example in a picture containing a set of grey scale bars, where contrast is accentuated due to the effects of the Mach phenomenon. In this situation between 12 and 16 levels can be discerned by most observers.
The obvious method of distributing dots within a pixel is to use a number of dots proportional to the datum value. So if data are scaled from 0 to 7, then dot densities per pixel could be 0, 9, 18, ......63. In fact such a scale was found to give most unsatisfactory results (figure 3.14). Generally pictures produced this way lack contrast. If some consideration is taken into account of the logarithmic response of the human eye, then a quadratic dot density scale appears to give a reasonably linear response. It is not relevant here to digress into psychophysics, however the interested reader is referred to Stockham's (1972) paper for a short account of some of the problems of evaluating image processing in relation to subjective distortion introduced by human vision.

The chosen dot pattern distributions are shown diagrammatically in figure 3.15. On an actual image a dot is approximately 0.0075 inch in diameter. Each pattern contains the dot pattern of the preceding level. Each pattern also has a slanting structure; this is quite pleasing to the eye and prevents a checker-board appearance of the image as a whole. For normal display purposes the data are scaled to eight levels before output to the plotter subroutine. The subroutine accepts data from a Fortran array and converts each value into its
PROGRAM: TIMJ (REV N)  DATE: 1-4-78
IDENTIFICATION: JA4

MAX COUNT  367  (ROW 19  COL 34)
TOTAL BGD SUB  40
TOTAL COUNT  310906
MULTI-CYCLES  1
THRESHOLD (X)  90

FIGURE 3.14
REGULAR QUADRATIC DOT DENSITY
GREY SCALE USING 8*8 PIXELS

- NO DOT + BLACK DOT

```
Level 0
--------
 Level 4
--------
```

```
Level 1
--------
 Level 5
--------
```

```
Level 2
--------
 Level 6
--------
```

```
Level 3
--------
 Level 7
--------
```

Figure 3.15
respective dot pattern. The paper flow is unidirectional so that dot patterns have to be arranged as a raster scan in units of 512 bits, before output to the plotter. A 64*64 image can be plotted from a Fortran program in 20 seconds, although in practice the software is deliberately slowed down to reduce paper speed and so increase the darkness of dots produced by the electrostatic printing process. A plotting time of 60 seconds gives reasonable results. An example of the output obtained using this method is shown in figure 3.16.

3.4.2 REGULAR GREY SCALE FOR REPRODUCING 128*128 IMAGES

Some of the coarse cellular nature of the 64*64 display can be reduced if the data are displayed as a 128*128 image. This can readily be implemented by linearly interpolating the original data so that the number of values in the X direction is doubled by inserting extra numbers whose values are equal to the mean of adjacent pixels in the original sequence. A similar process is then performed on the augmented X data in the Y direction, resulting in a new array containing four times the number of data values. If the paper display area is the same as that used for the 64*64 images then the new pixels have one quarter the area and the data are displayed with finer texture. Figure 3.17 demonstrates
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX COUNT</td>
<td>367</td>
<td>(ROW 19 COL 34)</td>
</tr>
<tr>
<td>TOTAL BGD SUB</td>
<td>40</td>
<td>TOTAL COUNT</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
<td>THRESHOLD (%)</td>
</tr>
</tbody>
</table>

**FIGURE 3.16**
PROGRAM: TIMJ (REV N)

IDENTIFICATION: JA4

DATE: 9-3-78

MAX COUNT 367 (ROW 19 COL 34)
TOTAL BGD SUB 40
TOTAL COUNT 310906
MULTI-CYCLES 1
THRESHOLD (X) 90

FIGURE 3.17
this effect. The image is the result of interpolating the data in figure 3.16. Each pixel is composed of a 4*4 dot structure as shown in figure 3.18. A quadratic dot density scale is used on the lines of that described in the preceding section.

Of all the electrostatic plotter grey scale techniques available the 128*128 interpolated display is the method usually chosen by most PDP 12 users. Many examples using this method will be found in later chapters in this dissertation.

3.4.3 REGULAR GREY SCALE FOR REPRODUCING 256*256 IMAGES

The interpolation just described can be taken further. The next logical choice of matrix size is 256*256. However, using the toning methods described for 64*64 and 128*128 arrays, a pixel would have to be composed of only four dots if a similar area display is to be achieved. Hence only five grey levels are available. To overcome this problem the 256*256 display makes use of all 1024 nibs across the plotter page width, with the result that pixels are based on the format used for the 128*128 display. An example of this method is shown in figure 3.19.
REGULAR QUADRATIC DOT DENSITY
GREY SCALE USING 4x4 PIXELS

- NO DOT  + BLACK DOT

<table>
<thead>
<tr>
<th>Level 0</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>----</td>
<td>----+</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>----</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>----</td>
<td>+----</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>----</td>
<td>+----</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level 2</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>++++</td>
<td>++++</td>
</tr>
<tr>
<td>++++</td>
<td>++++</td>
</tr>
<tr>
<td>++++</td>
<td>++++</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level 3</th>
<th>Level 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>++++</td>
<td>++++</td>
</tr>
<tr>
<td>++++</td>
<td>++++</td>
</tr>
<tr>
<td>++++</td>
<td>++++</td>
</tr>
</tbody>
</table>

Figure 3.18
Although the results are quite pleasing to the eye, this method suffers from the disadvantage that a complete picture takes over two minutes to produce. This extended time is mainly due to interpolation and pattern generation routines which assemble the image, which itself contains over one million bits of information.

3.4.4 NUMBER DISPLAY

As a variation on the grey tone patterns already described a toning method has been developed in which the numbers of dots used to print a character, in the range 0 to 7, is quadratically related to the digit value. An example of this using a 64*64 matrix is shown in figure 3.20. This method is not a gimmick: clinicians have been known to request this type of output for some scintigraphic procedures. For example in pulmonary function studies some quantitative information on regional ventilation or perfusion may be readily seen by comparison of number distributions within left and right lungs and between zones in the same lung.

3.4.5 RANDOM DOT DISPLAY
PROGRAM: TIMJ (REV N)  DATE: 1-4-78
IDENTIFICATION: JA4

MAX COUNT 367  (ROW 19  COL 34)
TOTAL BGD SUB 40
TOTAL COUNT 310906
MULTI-CYCLES 1
THRESHOLD (%) 90

FIGURE 3.20
The random dot density display has the ability to obliterate completely the cellular nature of the image. An example of a picture using this method is shown in figure 3.21. The data are displayed as a 64*64 matrix so that each pixel area is defined by an 8*8 dot region. The number of dots in a pixel region is quadratically related to one of eight display levels. The positioning of dots within each region is distributed pseudo-randomly using the following method.

The random dot density subroutine contains a lookup table which consists of 64 unique randomly arranged numbers in the range 0 to 63. As each scaled datum from the image is about to be plotted the subroutine converts the number into a dot count so that the quadratic relationship between display level and dot density is produced. This dot count determines the number of times random numbers are read from the table. Consecutive values in the table are used for each lookup, so that after 64 entries the table is re-entered from the first entry, and the cycle repeats. Each number from the table consists of a six bit integer which is broken down into two three bit parts. The most significant three bits are used to define a row address and the remaining three bits define a column within the pixel. A bit is then set in the appropriate word of the raster scan buffer. When a
PROGRAM: TIMJ (REV N)
IDENTIFICATION: JA4

---

MAX COUNT    367 (ROW 19 COL 34)
TOTAL BGD SUB 40
TOTAL COUNT   310906
MULTI-CYCLES  1
THRESHOLD (%) 90

FIGURE 3.21
complete row of image data has been assembled in the buffer the eight lines of dots are plotted. The cycle repeats until all 64 rows of the matrix have been output. The lookup table is only initialised once at program load time so that for a typical (noisy) scintigram the eye is seldom able to detect a pattern within the picture.

The merit of this type of display is that the image closely resembles the type of picture obtained from the analogue display of a gamma camera. Some clinicians favour this method because they are familiar with interpreting images from gamma camera displays. It is doubtful whether such pictures of unprocessed data are of great diagnostic value. Intuitively it seems that the transfer of analogue gamma camera data into digital form and back into analogue form must result in a deterioration of image signal to noise ratio. This conclusion is confirmed in the IAEA intercomparison of scintigraphic data processing methods (IAEA (1972), IAEA [1] 1976, IAEA [2] 1976).

3.4.6 ORDERED DITHER DISPLAY

The ordered dither display provides a method for reproducing 512*512 grey tone pictures on the plotter which differs from the straightforward methods of linear
interpolation already described. To produce a picture within an A4 page width would necessitate a pixel composed of a maximum of one dot. In other words a truly bistable display would be produced. Even if the full page width of the paper were used this would only provide a dot matrix of 1024*1024 points, so that a maximum of four dots per pixel would be available, resulting in a maximum of five grey tones. These problems led the author to consider the work done by Limb (1969), which was later surveyed by Jarvis (1976), who showed that a technique known as "ordered dither" could produce some of the best hard copy computer generated images using a bilevel display.

Basically the ordered dither consists in comparing the intensity level of each scaled datum with a threshold value. If the intensity is greater than the threshold then the corresponding dot on the plotter is darkened, otherwise that part of the paper is left blank. The threshold value is spatially dependent, such that a set of threshold values is contained within a dither matrix, $D_{pq}$. A matrix element $D_{pq}$ is selected such that

$$ p = j \text{ modulo } 4 + 1 \quad 1 \leq j, k \leq 512 $$
$$ q = k \text{ modulo } 4 + 1 $$

If the original 64*64 picture data are linearly interpolated up to 512*512 pixels and the resultant data quadratically scaled within a range of 1 to 16 levels,
then each pixel intensity can be represented as

\[ I_{jk} \]

where \( 1 \leq I_{jk} < 16 \) and \( 1 \leq j, k \leq 512 \)

A dot location \( j,k \) on the paper is blackened if \( I_{jk} > D^4_{\text{th}} \).

The dither matrix used was

\[
D^4 = \begin{pmatrix}
1 & 9 & 3 & 11 \\
13 & 5 & 15 & 7 \\
4 & 12 & 2 & 10 \\
16 & 8 & 14 & 6
\end{pmatrix}
\]

Integers in the range 1 to 16 are distributed randomly within the matrix. It is now apparent that this arrangement will offer, in theory, 16 grey tones. The size of the dither matrix is to some extent arbitrary; a 4*4 matrix was chosen to be consistent with the 16 tones available on the television display. It should be noted that the plotted picture does not lose spatial resolution as the size of the dither matrix is increased, even though the threshold intensity resolution increases. Of course, in practice, the intensity resolution is governed by the ability of the eye to differentiate between small changes in dot density.

In the author's opinion the grey tone images produced by the ordered dither method are superior to the linearly interpolated ones using a defined toning pattern. Figure 3.22 is an example of the ordered dither method. The main
FIGURE 3.22
draw back to this method is the fact that the Fortran version on the PDP 12 takes over five minutes to produce a single picture; most of this time the program is CPU bound and actual plotting time is relatively small. This excessive time arises because the 64*64 array has first to be interpolated to 512*512 and then some 262144 calculations are performed to produce results for all possible dot positions within the picture. The large amount of processing time required tends to be intolerable for clinical use and under normal circumstances the regular grey scale 128*128 format has been adopted.
CHAPTER 4

IMAGE FORMATION AND THE SPATIAL FREQUENCY DOMAIN

4.1 BASIC CHARACTERISTICS OF THE ELEMENTAL SCINTIGRAM.

Before presenting a mathematical model of the digital image it is instructive to consider some of the properties of the data. Since gamma quanta beams cannot be focussed by an optical lens, an image is formed by using the geometrical properties of lead collimators. The images so produced are characterised by rather poor spatial resolution. The measured radioactive distribution has therefore been blurred and the aim of image processing is to refocus the data in an attempt to obtain a more faithful representation of the object distribution. In the literature this process is referred to as restoration, de-blurring, and in mathematical terminology, deconvolution.

In a clinical scintigram the situation is further aggravated by the poor signal to noise ratio of data. Since the amount of radioisotope which can be administered to a patient must usually be small to minimise biological radiation effects, a compromise between image resolution
and noise must be made. Noise is added to the data by the nature of radioactive decay which is a random process described by Poisson statistics. For a Poisson distribution the fractional standard deviation is inversely proportional to the square root of the number of detected events. Hence for a given data collection time and photon count rate the fractional standard deviation increases as the pixel size decreases. The signal to noise ratio will therefore improve with increasing pixel size. Whether noise is introduced by the measuring equipment, or by nature of the data itself, it will be shown in this dissertation that noise is the main obstacle to perfect restoration. By contrast, even if the data channel has a limited frequency response, this does not in itself rule out a reasonable restoration of the object data.

4.2 MATHEMATICAL MODEL

A scintigram may be considered as an image \( g(x,y) \), which is a two dimensional reproduction of an object distribution \( f(x,y) \). The mathematical relationship between these two functions for a linear system may be described as
The function $h(x, y, x_1, y_1)$ is called the point spread function which describes the degradation or blurring introduced by the measuring system. In equation 4.1 it should be noted that $h$ is written as a function of $(x, y)$ and $(x_1, y_1)$; that is to say the weighting of $f(x_1, y_1)$ depends not only on $(x_1, y_1)$ but also on points $(x, y)$ in the vicinity of $(x_1, y_1)$. Practically, the point spread function defines the output from the system when the input is an impulse, $d(x, y)$. In scintigraphy $h(x, y, x_1, y_1)$ therefore describes the blurring produced at the output of the imaging system for a point source of radioactivity. The terms point spread function and impulse response are used synonymously. For a causal system and an isotropic point spread function

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x'-x_1, y'-y_1) f(x_1, y_1) \, dx_1 \, dy_1 \quad 4.1$$

Then under these circumstances

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x'-x_1, y'-y_1) f(x_1, y_1) \, dx_1 \, dy_1 \quad 4.3$$

or

$$g(x, y) = h(x, y) \ast f(x, y) \quad 4.4$$

The above equation is in the form of a standard convolution integral suggesting that the image is really a convolution of the object with the point spread function.
The problem of restoration consists in solving equation 4.3 for $f(x,y)$. Mathematically the equation is a two dimensional form of the Fredholm equation of the first kind, where $h(x,y)$ is the kernel of the integral equation. Even in the ideal case where noise is assumed to be negligible there is a problem in finding a solution. It can be shown by the Rieman-Lebesque theorem that even if a solution does exist it does not depend continuously on $g(x,y)$ (Miller 1974, Phillips 1962). The consequence of this result is that small changes in the measured data, $g(x,y)$, may produce large perturbations in the solution, $f(x,y)$. Hence a small error or noise in $g(x,y)$ may lead to a grossly distorted solution to the equation.

For a function $f(x,y)$, the Fourier transform, $F(u,v)$, of $f(x,y)$ is defined as

$$F(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \exp\{-i(ux + vy)\} \, dx \, dy$$  \hspace{1cm} 4.5

$F(u,v)$ represents the frequency components of $f(x,y)$. The units of $u$ and $v$ are therefore cycles per unit distance. In general $F(u,v)$ is a complex function. The inverse Fourier transform of equation 4.5 is

$$f(x,y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u,v) \exp\{i(ux + vy)\} \, du \, dv$$  \hspace{1cm} 4.6
Similarly $H(u,v)$ and $G(u,v)$ are the Fourier transforms of $h(x,y)$ and $g(x,y)$ respectively. It can be shown (Bracewell 1965) that if and only if equation 4.4 is true then

$$G(u,v) = H(u,v) F(u,v) \tag{4.7}$$

Hence the Fourier transform maps convolution in the space domain into multiplication in the frequency domain. It is this mapping property of the Fourier transform that makes it useful for implementing deconvolution. In theory, providing $H(u,v)$ never equals zero and noise is negligible, it should be possible to recover the object from the image by

$$F(u,v) = G(u,v)/H(u,v) \tag{4.8}$$

and inverse Fourier transforming to obtain $f(x,y)$.

In a practical situation $g(x,y)$ is not measured because the data available is contaminated with noise, $e(x,y)$, so that what is actually measured is

$$g'(x,y) = g(x,y) + e(x,y) \tag{4.9}$$

$$= h(x,y) * f(x,y) + e(x,y)$$

The noise function is a random process which represents the difference between the observed and expected values of the data. If $E(u,v)$ is the Fourier transform of the noise
then

\[ G(u,v) - E(u,v) = H(u,v) F(u,v) \]

so

\[ F(u,v) = \frac{G'(u,v)}{H(u,v)} - \frac{E(u,v)}{H(u,v)} \quad 4.10 \]

The transform of the quotient \( G'(u,v)/H(u,v) \) is well behaved. For the scintigraphic system \( H(u,v) \) becomes smaller for increasing frequencies and \( G'(u,v) \) behaves in a similar fashion. However \( E(u,v) \) is not related to \( H(u,v) \) in any way so that zeros may not coincide. Generally \( E(u,v) \) is roughly constant so that the effect of division by \( H(u,v) \) is to amplify any noise present in the measured data to the extent that \( f(x,y) \) may be obliterated.

4.3 DISCRETE MATHEMATICAL MODEL

In order to implement digital processing of a radionuclide image it is necessary to represent the image as an array of intensity values or samples over the image area. For the present discussion a set of data will be considered to be a square array of \( N^2 \) samples which describe the function \( f(x,y) \) over the image coordinates \((x,y)\) in the space domain. Then if the sample intervals in the \( x \) and \( y \) directions are \( dx \) and \( dy \) respectively

\[ f_{jk} = f(jdx, kdy) \quad 0 \leq j, k \leq N-1 \]
Further if \( f_{jk} \) is constrained to be periodic with period \( N \) then
\[
f_{j+pN, k+pN} = f_{jk} \quad -\infty < p < \infty
\]
Similar representations of the discrete versions of \( g(x,y) \) and \( h(h,y) \) may be formulated. The discrete linear convolution becomes a summation rather than an integral and
\[
g_{jk} = \sum_{j'=-\infty}^{\infty} \sum_{k'=-\infty}^{\infty} h_{j-j', k-k'} f_{j'k'}
\]
This equation describes the aperiodic convolution of \( f \) and \( h \). However for a discrete system it is more useful to consider a periodic or circular convolution which is related to the discrete Fourier transform. The circular convolution for a unit sample interval may be written as
\[
g_{jk} = \sum_{j'=-N}^{N-1} \sum_{k'=-N}^{N-1} h_{j-j', k-k'} f_{j'k'}
\]
The two dimensional discrete Fourier transform is (Oppenheim 1975)
\[
F_{\alpha\beta} = \frac{1}{N} \sum_{j=-N}^{N-1} \sum_{k=-N}^{N-1} f_{jk} \exp \left\{ -\frac{2\pi i}{N} (j\alpha + k\beta) \right\}
\]
One period of \( f_{jk} \) and \( F_{\alpha\beta} \) occurs in the image range 0 to \( N-1 \). The theory assumes that both \( f_{jk} \) and \( F_{\alpha\beta} \) are periodic functions which repeat themselves indefinitely in both directions. Since each period of a periodic function is identical to any other, only one period will be shown in an illustration. Similarly the inverse Fourier transform is
The scaling factor $1/N$ appears in both transforms. In the literature the scaling factor may appear as $1/N^2$ in one transform. Since equations 4.11 and 4.12 constitute a transform pair, these variants are both valid providing consistency is maintained throughout their manipulation.

At this point it is worth considering some of the implications of discrete convolution. It is straightforward to verify that if a discrete linear two dimensional convolution is carried out on two finite area sequences $f_{jk}$ and $h_{jk}$, where $j$ and $k$ take values up to $N$, then a sequence with a maximum of $(2N-1)^2$ nonzero points must be computed. In this situation $f$ and $h$ need not be periodic. However when the discrete Fourier transform is used to convolve two functions, then by definition the convolution is periodic. In reality any signal that comes from an imaging system will have been convolved aperiodically. So if the discrete Fourier transform is to be used for practical convolution it is necessary to implement an aperiodic convolution with a periodic or circular convolution. The linear two dimensional convolution of two finite two dimensional sequences $f_{jk}$ and $h_{jk}$, each of $N^2$ elements, is obtained by adding zeros to both of them until they are both of area $P^2$, where

$$f_{jk} = \frac{1}{N} \sum_{l \in \mathbb{Z}} \sum_{m \in \mathbb{Z}} f_{l,m} \exp \left\{ \frac{2\pi i}{N} (jl + km) \right\}$$

4.12
P \gg 2N-1, and convolving the augmented functions periodically. Then the convolution process can be represented as

\[ G_{mn} = N H_{mn} F_{mn} \quad 4.13 \]

The two dimensional convolution can be accomplished by repeated evaluation of one dimensional transforms using the following identities

\[ \tilde{f}_{\lambda K} = \sqrt{\frac{i}{N}} \sum_{j=0}^{N-1} f_{jk} \exp\left\{-\frac{2\pi i j \lambda}{N}\right\} \quad 4.14 \]

\[ f_{\lambda m} = \sqrt{\frac{i}{N}} \sum_{k=0}^{N-1} \tilde{f}_{\lambda K} \exp\left\{-\frac{2\pi i k \lambda}{N}\right\} \quad 4.15 \]

In the following chapters various optimisation procedures for finding solutions to equation 4.4 will be discussed. By nature of the Fourier transform techniques used to implement deconvolution, this optimisation is performed in the frequency domain. It is therefore appropriate to consider some other properties of the Fourier data when \( f_{jk} \) is known to be a real positive function.

When equation 4.11 is used to transform a function \( f_{jk} \) the resultant frequency components \( l \) and \( m \) have their origin at one corner of the two dimensional spectrum. For display purposes it is more convenient if the origin of
the data is shifted so that zero frequency is at the centre of the display. The origin has therefore been moved by $N/2$ on both axes. This is simply achieved by multiplying the spatial data by $(-1)^{j+k}$ prior to transforming.

By letting

$$F'_{\ell m} = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} (-1)^{j+k} f_{jk} \exp \left\{ -\frac{2\pi\ell}{N} (j\ell + km) \right\}$$

and using the identity that

$$(-1)^{j+k} = \exp \left\{ i\pi (j + k) \right\}$$

by substitution

$$F'_{\ell m} = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f_{jk} \exp \left\{ -\frac{2\pi\ell}{N} \left[ (j\ell - N\frac{k}{2}) + (m - N\frac{\ell}{2})k \right] \right\} \quad 4.16$$

Equation 4.16 defines the origin as $N/2, N/2$.

Even though $f_{jk}$ is a real positive function, $F_{\ell m}$ is in general complex. However $F_{\ell m}$ will demonstrate conjugate symmetry for a real $f_{jk}$; $F_{\ell m}$ may be expressed in terms of real and imaginary components

$$F_{\ell m} = \Re \{ F_{\ell m} \} + \Im \{ F_{\ell m} \}$$

in which case it can be shown (Oppenheim 1975)

$$\Re \{ F_{\ell m} \} = \Re \{ F_{N-\ell, N-k} \} \quad 1 \leq \ell, m \leq N-1$$

and

$$\Im \{ F_{\ell m} \} = -\Im \{ F_{N-\ell, N-k} \} \quad 1 \leq \ell, m \leq N-1$$
So although the scintigram contains $N^2$ samples, and its Fourier transform $2N^2$ members which define the real and imaginary components of each spatial frequency, all the information is contained in one half of the transformed data or $N^2$ data components. This property is illustrated in figure 4.2: the picture shows rotational symmetry about the origin.

4.4 TWO DIMENSIONAL FFT COMPUTER PROGRAM

Equation 4.14 is the transform of the rows of the image matrix, and equation 4.15 that of the columns. All two dimensional Fourier transforms on the PDP 12 have been programmed as repeated one dimensional operations on the rows and columns of complex data matrices. The programming of these processes will now be described.

The PDP 12 floating point processor provides the normal add, subtract, multiply, divide and normalise functions for operating on real variables. In order to represent variables with mantissa and exponent parts a total of 36 bits or three words are used. Two words define the mantissa and the remaining word the exponent. This structure provides a number range of $\pm 10^{\pm 620}$ with
seven figure significance. A 64*64 array of data therefore requires 12k of core. It is apparent that once two such matrices are stored in the computer memory there is little core available for the program and operating system. From the outset it was therefore decided to make use of the decomposability of the two dimensional Fourier transform.

All image files on the PDP 12 are stored with sequential row access. By using this method one unit of data structure is one row of a matrix and is also one physical record of the PDP 12 magnetic storage system (this applies to both disc and magnetic tape backing storage). Hence a 64*64 array of data is held as a 64 record file. Performing the row Fourier transform is easily programmed; a row is read from the file, the FFT of the row performed, and the transformed row is then written back into the same file. This process continues on all the rows of the matrix. The columnwise transformation is incompatible with the sequential storage of the transformed rows. However if the matrix generated by the row transforms is transposed, and then repeated on the new file, the rows in the new file are the columns that are required to perform equation 4.15. A normal sequence of row transforms is then performed. The resultant stored matrix consists of a two dimensional Fourier transform
whose coordinate system has been reflected about the diagonal. This method of performing the FFT on large matrices has one main drawback in that a substantial amount of input/output time is necessary to fetch a row of the matrix from disc and store the row back on the disc after FFT processing. Further input/output time is required for taking the matrix transpose. Typically the complete transform takes 120s on the PDP 12. The FFT subroutine is based on the Fortran IV version available from the Numerical Algorithms Group (1973).

4.5 POWER SPECTRUM

The Fourier transform of the impulse response is sometimes called the modulation transfer function because it describes the frequency modulation of the object function by the imaging system. The modulation transfer function may be considered analogous to the frequency response of an electronic system. The modulation transfer function is a complex matrix which has magnitude and phase associated with each element. Throughout this dissertation it will be convenient to consider a scalar function which will describe the density of frequency components, called the power spectrum (in the literature it may be referred to as the power density spectrum or Wiener spectrum). The power spectrum gives a measure of
"the energy" in each frequency component. It is obtained by calculating the product of \( F_{\ell m} \) with its complex conjugate, or alternatively the square of the modulus of \( F_{\ell m} \). It can be readily shown (Lynn 1973) that this approach is equivalent to the more conventional method of using the Wiener-Khinchin relationship in which the power spectrum is expressed as the Fourier transform of the autocorrelation function.

As a result of phase cancelling out in the power spectrum a useful result to be used in subsequent chapters is the fact that the power spectrum densities are independent of the positioning of the object distribution within the \((x,y)\) coordinate system. A simple proof is shown by considering a coordinate transformation.

From equation 4.11 a coordinate transformation
\[
j' = j + a, \quad a \text{ is an integer constant}
\]
\[
k' = k + b, \quad b \text{ is an integer constant}
\]
will yield a new Fourier transform
\[
F'_{\ell m} = \frac{1}{N} \sum_{j=a}^{j+a} \sum_{k+b}^{k+b} \hat{f}_{j'k'} \exp \left\{ -\frac{2\pi i}{N} \left[ (j' - a)\ell + (k' - b)m \right] \right\}
\]
or
\[
F'_{\ell m} = \exp \left\{ \frac{2\pi i}{N} \left( a \ell + b m \right) \right\} F_{\ell m} \quad 4.17
\]
and
\[
\tilde{F}'_{\ell m} = \exp \left\{ -\frac{2\pi i}{N} \left( a \ell + b m \right) \right\} \tilde{F}_{\ell m} \quad 4.18
\]
The power spectrum is given by \((F_{k'n'})(\overline{F}_{k'n'})\) which is readily seen to equal \((F_{k'n'})(\overline{F}_{k'n'})\) if equations 4.17 and 4.18 are multiplied together.

The dynamic range of values in the Fourier domain is typically very large. This is appreciated by considering an upper bound on \(f_{jk}\) of \(f_{jk}\) \(\text{max}\). From equation 4.11 \(F_{k'n'}\) will be bounded above by \(Nf_{jk}\) \(\text{max}\). The PDP 12 interfaces to the radionuclide imaging equipment allow a maximum value of \(2^{12}\), so for \(N\) of 64 the maximum power spectral component is bounded by \(10^{10}\). However for a realistic scintigram with say \(5.10^5\) counts, a maximum component of \(6.10^7\) would be produced. This is just compatible with the seven significant figures carried by the PDP 12 floating point processor. It is apparent that the frequency domain contains numbers of much greater amplitude than the spatial domain. However, while the dynamic range of numbers is large it should be noted that few spectral components can actually take large values. From Parsavel’s relationship (Oppenheim 1975), the total energies in the space and frequency domain must be equal.

\[
\sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left\{f_{jk}\right\}^2 = \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} \left\{F_{k'm'}\right\}^2
\]

Thus a few large valued spectral components will account for most of the energy of the spatial function. Typical pictorial representations of power spectra are shown in figures 4.1 to 4.4. These plots were produced from data.
PROGRAM: PSD (VER C)  LIN PLOT  DATE: 31-12-77

IDENTIFICATION: JA4 POWER SPECTRUM

SUM  8.51011E+08
MAX CELL VALUE  3.93026E+08  (ROW 33 COL 33)
MIN CELL VALUE  7.77539E-01  (ROW 21 COL 44)

FIGURE 4.1
ID: JA4 POWER SPECTRUM

SUM: 0.85101E+08
MAX CELL VALUE: 0.383026E+08 (ROW 33 COL 33)
MIN CELL VALUE: 0.777539E-01 (ROW 21 COL 44)

FIGURE 4.2
LINEAR SCALE

Maximum frequency component at (0,0) of 0.383026E + 08

FIGURE 4.3 Power Spectrum of Williams phantom
(The frequency origin is at the centre of the display)
CUBE ROOT SCALE

Maximum frequency component at (0,0) of $0.383026 \times 10^8$

FIGURE 4.4 Power spectrum of Williams phantom
(The frequency origin is at the centre of the display)
collected by scanning the William's phantom (filled with \textsuperscript{Tc99m}) on the J and P scanner using the \textsuperscript{109} hole collimator. The phantom face was placed perpendicular to the collimator axis at the focal distance. A \(0.4\) cm. sample width was used. The total count of the recorded image was approximately \(4.1 \times 10^5\). The linear display scale of figures 4.1 and 4.3 emphasises the large low frequency components but the cube root display scale of figures 4.2 and 4.4 enables more high frequency detail to be seen. It is customary to display data covering a large dynamic range using a logarithmic display scale. For displaying data greater than unity the cube root scale produces similar results to the logarithmic scale. The drawbacks of the logarithmic scale are that precautions have to be taken to cover the cases where data tend to zero (one large negative logarithm can completely dominate the scale) or are negative. As well as overcoming these problems the cube root scale is the natural choice for data that may be equally disposed about zero. In subsequent two dimensional power spectral plots only the cube root of the power spectrum will be displayed.

4.6 PIXEL SIZE
In radionuclide imaging sample size is usually a compromise between acceptable signal to noise ratio and having a small enough pixel area to provide a faithful reproduction of spatial frequencies in the object and point spread functions. Sampling theory states that if the smallest occurring distance in the spatial domain is $dx$ then there exists a frequency $u_c = 1/2dx$ which is called the cut-off or Nyquist frequency. Any frequencies higher than $u_c$ in the data will be folded back into frequencies lower than $u_c$, the outcome of digital signal processing will then be unpredictable. Unfortunately the sampled count density is inversely related to the pixel area for a given photon count rate. It is fruitless having a small pixel if in clinical procedures the counts collected per pixel have small statistical significance.

W.A. Hunt (1970) has suggested that an optimum sample size may be determined by eliminating spatial frequencies where signal energy is less than noise energy. By modelling the point spread function as a gaussian function he concludes that resolution for a single point of activity is proportional to the logarithm of the number of counts. A more common approach is to ensure that frequency components in the point spread function are small beyond the Nyquist frequency. If resolution enhancement is attempted then the pixel size should also
be small enough to minimise aliasing in significant portions of the object spectrum. This latter condition may be difficult to determine for in vivo imaging of radionuclide distributions.

In this study pixel size has largely been determined by hardware constraints. In the case of the gamma camera, the only realistic way of interfacing to a 12 bit word computer is to use 6 bits for both the x and y coordinates. By definition this defines a 64*64 image structure. If one considers one dimension of the cartesian grid then the frequency represented by the j th element is

\[
\text{frequency } = \frac{j}{(N \cdot dx)} \quad 0 \leq j \leq N/2
\]

N represents the total number of samples across the camera face, and dx refers to the sample width in centimetres. The maximum frequency represented is

\[
\frac{1}{2 \cdot dx} \quad \text{cycles per cm.}
\]

The sample width may be computed as the quotient of the camera viewing width and the number of samples across the field. In practice it is readily determined by measuring the number of pixels that separate two fine (0.05 cm. diameter capillary tubes were used) line sources placed perpendicularly to the x axis and of known separation.
For the NE MK V HR camera this corresponds to a sample width of 0.43 cm. The maximum frequency represented is therefore 1.16 cycles per cm. Measurements in the y direction produce a similar result.

For the J and P scanner the sample width is defined by the hardware of the equipment, it being designed to have widths of 0.2, 0.4, 0.8 cm. The maximum frequencies represented are

for \( dx = 0.2 \) cm., maximum frequency = 2.5 cycles per cm. and \( dx = 0.4 \) cm., maximum frequency = 1.25 cycles per cm.

The 0.8 cm. pixel was never used as it would be incapable of faithfully reproducing spatial frequencies propagated by the 109 hole collimator. Normally a 0.4 cm. sample width has been used for most clinical investigations; this size allows most organs to be represented within the computer as a 64*64 array of data.

4.7 POINT SPREAD FUNCTION

The solution of equation 4.4 for \( f(x,y) \) requires a priori knowledge of the point spread function \( h(x,y) \). This condition is true even if noise is assumed negligible. MacIntyre (1968) gives a review of manual methods suitable for the determination of modulation
transfer function. He generally favours derivation from measurements on a fine line source of radioactivity. In the computerised system used here it has been found directly from measurements with a point source.

The theory used in the following chapters will assume that the point spread function is noise free and is invariant under translation through the object plane. In the case of the scanning detector system the former requirement is met by scanning a high specific activity "point" source sufficiently slowly. This must be done such that the count rate is not so high as to produce non-linearities in the detector system and so that overflow does not occur in the Digideck interface. Providing that the point source is not disturbed, repeated scans can be made and the results added to produce a high accuracy image. To some extent checks on the positional accuracy of the point source in consecutive scans can be made by ensuring that the centre of gravity of each scan lies at the same pixel location within the computer matrices.

The other requirement that the point spread function is isotropic is not so easily justified. For a given isotope energy; scattering medium, and collimator, the modulation transfer function will vary with distance
between the source and collimator. Furthermore in a clinical study the object plane and scattering medium depth will vary. If changes in the point spread function are small over the thickness of structures being visualised, then the isotropic behaviour of the point spread function will be approximately true. In the studies carried out here, if the point spread function is thought to vary appreciably for a particular patient investigation then the wider power spectrum was used. This ensures that the mathematical deconvolution techniques, to be discussed later, do not over compensate for the apparent loss of high spatial frequencies resulting in mathematical instabilities in the deconvolution process. In the author's opinion this phenomenon is one of the main factors which results in only approximate refocussing being possible. Figures 4.5 and 4.6 show the point spread functions for the 109 hole collimator on the J and P scanner with and without 5 cm. of hardboard between the source and collimator. A slightly broader response is shown for the case where an absorbing medium is present. The power spectrum profile along the x axis for figure 4.5 is shown in figure 4.7. As the point spread function approximates to a gaussian, and the Fourier transform of a gaussian is another gaussian, the power spectrum of figure 4.7 also closely resembles a gaussian form. Under normal circumstances,
POINT SPREAD FUNCTION

J&P SCANNER 109 HOLE COLLIMATOR.
RESPONSE AT FOCAL POINT USING A 0.2CM PIXEL.

FIGURE 4.5
POINT SPREAD FUNCTION

J&P SCANNER 109 HOLE COLLIMATOR (0.2CM PIXEL).
RESPONSE AT FOCAL POINT WITH 5CM HARDBOARD.

FIGURE 4.6
POWER SPECTRUM OF P.S.F.

J&P SCANNER 109 HOLE COLLIMATOR RESPONSE AT FOCAL POINT.

FIGURE 4.7
for brain and liver scans a modulation transfer function obtained using 5 cm. of hardboard with the object in the focal plane of the collimator was used for the deconvolution of the patient image.

In the gamma camera system similar problems arise, although the situation is further aggravated by the fact that the stationary detector system may exhibit variations in the point spread function over the image plane above those already described. Normally these variations are small for a well adjusted camera compared with those introduced by the variable depth scattering medium. An exception to this rule occurs at the extreme edge of the field of view of the detector system. Generally little useful information is presented in this area so that distortion introduced by data processing is not troublesome.

For two reasons deconvolution has been developed and tested using data from the J and P scanner. The first is that the problem of variation of point spread function across the field of view of the gamma camera is eliminated. Secondly, the scanner provides a choice of pixel size for the accumulation of data. The small sample size of 0.2 cm. is particularly valuable for making accurate measurements of the modulation transfer function.
under different scanning conditions.

4.8 PHANTOM MODEL

When developing any new mathematical processing technique it is always helpful to have simulated or model data whose characteristics are well defined. The model data may also be deliberately distorted enabling the properties of the processing techniques to be assessed. In this respect distortion of data due to noise is particularly important when developing deconvolution techniques. This section describes the ways in which simulated scanner data were set up on the computer.

In order to solve equation 4.13 two functions are required: $H_{lm}$ and $F_{lm}$. The point spread function, $h_{jk}$, was modelled by defining a modified gaussian function

$$h_{jk} = h_{oo} \exp \left\{ -\frac{2.77 (j^2 + k^2)}{FWHM^2} \right\} \quad h_{jk} \geq 0 \quad h_{oo} \quad 4.19$$

$$h_{jk} = 0 \quad h_{jk} < 0.1 \quad h_{oo}$$

where $h_{oo}$ is an arbitrary scaling constant, and FWHM is the full width at half maximum of the impulse response. Figures 4.5 and 4.8 show profiles through genuine data collected using a 0.2 cm. pixel width and the modelled point spread function. The latter figure portrays a one
MODEL P.S.F.

MODEL FOR J&P SCANNER 109 HOLE COLLIMATOR.
(THE X AXIS SCALE IS EQUIVALENT TO A 0.2CM PIXEL)
dimensional cross section calculated from equation 4.19 with the peak of the gaussian moved to the centre of the display. The data are normalised to a maximum of 100 to facilitate comparison between profiles. The noise free model is a realistic approximation to the real data. The object data, \( f_{jk} \), was modelled by manually transferring to the computer the relative cavity thickness of the William's phantom using a cartesian grid. Although this is a tedious task for a matrix comprising over 4000 elements it is a relatively simple procedure. As the phantom has only three different cavity thicknesses the computer program used for the digitisation process could be optimised such that for any row in the matrix only the first and last location of a particular cavity depth need be specified. Unspecified locations were filled with zeros by default. Each square on the cartesian grid that contained a boundary between different cavity thicknesses was given an interpolated value approximately equal to the relative cavity thickness enclosed by the square. The geometrical and the computer grey scale representations of the phantom are shown in figures 4.9 and 4.10.

By using equation 4.13 and taking the inverse FFT produces a result shown in figure 4.11. This picture represents a noise free convolution of the phantom model with the modified gaussian function.
FIGURE 4.9 Geometrical representation of Williams Phantom.
(0.67 actual size)
PROGRAM: TIMJ (REV N)                  DATE: 1-4-78
IDENTIFICATION: MOD

MAX COUNT  200  (ROW 49  COL 37)
TOTAL BGD SUB  0
TOTAL COUNT  177350
MULTI-CYCLES  1
THRESHOLD (X)  100

FIGURE 4.10  Williams model
FIGURE 4.11  Convolved Williams model
Simple deconvolution of figure 4.11 using $H_{l,m}$ results in the image shown in figure 4.12. The result is very similar to that of the original phantom (figure 4.10). The slight patterning effect merely shows that rounding errors within the computer produce slight fluctuations in the data around the 50 percent count density. With an eight level display scale the boundary between the fourth and fifth levels occurs at 50 percent of the maximum image value. The refocussed image shows that deconvolution of the simulated noise free data is quite satisfactory.

The addition of noise to the convolved model and subsequent deconvolution presents a different situation. Statistically normal noise was added to the convolved model using a process based on the Central Limit Theorem. Briefly, this can be summarised as stating that in the limit as $P$ goes to infinity, the sum of $P$ uniformly distributed variables tends to a normal distribution. Hence by using a random number generator, normal noise can be generated. It can be shown that a value of $P$ equal to 12 gives a good approximation to a normal distribution. The random number generator used was based on a special bit manipulation algorithm developed by Palmer (1973) suitable for the 36 bit variables used in the PDP 12 computer. A Chi-Squared test and an autocorrelation
PROGRAM: TIMJ (REV N)
IDENTIFICATION: MODCD

MAX COUNT  200  (ROW 19  COL 38)
TOTAL BGD SUB  0
TOTAL COUNT  177410
MULTI-CYCLES  1
THRESHOLD (X)  100

FIGURE 4.12  Deconvolution of data shown in figure 4.11
program were developed to check the randomness of the generator. The results of these tests were satisfactory. The process of adding noise to the convolved image therefore consisted in using the Central Limit Theorem to modify each element value in the convolved image. The mean of each generated normal variable was made equal to the original pixel value and a standard deviation equal to its square root. The result of this process is depicted in figure 4.13. The data can be compared with figure 3.17 which depicts the results of displaying data from a scan of the real William's phantom. The complex quotient of the Fourier transform of the noisy convolved model with $H_{lm}$ followed by an inverse Fourier transform produced the result shown in figure 4.14. The unpredictable results which were described in section 4.2 are demonstrated.

4.9 SUMMARY

In this chapter the mathematical foundations of image formation have been described. The processes of two dimensional convolution and deconvolution have been developed. The problem of finding a reasonable solution to the Fredholm integral equation in the presence of noise has been demonstrated. Subsequent chapters will describe methods that have been implemented to overcome the problems of mathematical instability.
**PROGRAM:** TIMJ (REV N)  
**DATE:** 12- 4-78

**IDENTIFICATION:** MODCN

<table>
<thead>
<tr>
<th>MAX COUNT</th>
<th>233</th>
<th>(ROW 16 COL 38)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>177526</td>
<td></td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>THRESHOLD (X)</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 4.13** The effect of adding noise to the data in figure 4.11
**IDENTIFICATION:** MODCND

<table>
<thead>
<tr>
<th>SUM</th>
<th>0.177469E+06</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX CELL VALUE</td>
<td>0.103632E+06</td>
</tr>
<tr>
<td>MIN CELL VALUE</td>
<td>-0.942415E+05</td>
</tr>
</tbody>
</table>

**FIGURE 4.14** "Deconvolution" of data in figure 4.13
5.1 INTRODUCTION

In chapter 4 it was shown that the inversion of an equation in the form

\[ g_{jk} = h_{jk} * f_{jk} \]  

5.1

presents problems if large oscillations in a solution are to be avoided. The problem is essentially ill-conditioned in that several solutions satisfy the integral equation for a slightly perturbed \( g_{jk} \). If noise, \( e_{jk} \), is present then there is no unique solution because both \( f_{jk} \) and \( e_{jk} \) are unknown. Phillips (1962) proposed that it may be better to seek a smooth solution to the equation, the assumption being that a smooth solution will tend to be a close approximation to \( f_{jk} \). Such a proposal can be justified if the distribution of \( f_{jk} \) is known a priori to be a smooth function. In the case of in vivo measurements of radioactive distributions this situation will normally hold. A constraint could be introduced such that a smoothness measure is minimised. Phillips suggested that a solution \( \hat{f}_{jk} \), of equation 5.1 may be such that
\[
\sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \{ f''_{jk}\}
\]

is minimised, where \( f''_{jk} \) represents the second derivative of \( f_{jk} \). The solutions developed by Phillips and later refined by Twomey (1965) were presented for a one dimensional case and the implementation of their techniques required matrix inversions. The two dimensional case can be reduced to a one dimensional problem but the implementation on a digital computer is formidable owing to the large size of matrices involved. Hunt (1970, 1971) recognised the problems of implementation and successfully circumvented them by making use of the properties of circulant matrices and expressing the solution in terms of the discrete Fourier transform.

In this chapter two dimensional constrained deconvolution will be developed in terms of the discrete Fourier transform. The statistical properties of scintigraphic data will be incorporated into the solution.

5.2 DECONVOLUTION

For the general case, in which noise is not negligible, equation 5.1 may be expressed as
\[ g_{jk} = h_{jk} * f_{jk} + e_{jk} \] 5.3

so that in terms of the discrete Fourier transform equation 5.3 becomes

\[ G_{\lambda m} = NH_{\lambda m} \hat{F}_{\lambda m} + E_{\lambda m} \] 5.4

It is proposed that the solution to equation 5.4, should be smooth and have errors of similar magnitude to the errors in \( G_{\lambda m} \). Although the noise sequence is unknown it is possible to make an estimate of the total expected error in the image. An estimate of the error in \( g_{jk} \) will be given by the variance of the data, and by Parsavel's relationship this value will also represent the total expected error in the Fourier domain. It will be assumed that

\[
\sum_{j=0}^{N-1} \sum_{k=0}^{N-1} e_{jk}^2 = \epsilon \] 5.5

where \( \epsilon \) is a known constant. If \( \hat{F}_{\lambda m} \) is an estimate of a solution to equation 5.4, and if the errors in \( \hat{F}_{\lambda m} \) and \( G_{\lambda m} \) are to be of similar magnitudes then we may expect

\[
\sum_{j=0}^{N-1} \sum_{m=0}^{N-1} (G_{\lambda m} - NH_{\lambda m} \hat{F}_{\lambda m})(G_{\lambda m} - NH_{\lambda m} \hat{F}_{\lambda m}) = \epsilon \] 5.6

The problem then consists in finding a sequence \( \hat{F}_{\lambda m} \) which satisfies equation 5.6 and which obeys the smoothness criterion.
A numerical approximation to the second derivative of a sequence \( \hat{f}_j \) at a point \( j \) is

\[
\hat{f}_{j-1} - 2\hat{f}_j + \hat{f}_{j+1}
\]

which in two dimensions becomes

\[
\hat{f}_{j-1,k} + \hat{f}_{j+1,k} + \hat{f}_{j,k-1} + \hat{f}_{j,k+1} - 4\hat{f}_{j,k}
\]

Hence the smoothness criterion can be expressed as

\[
\min \left\{ \sum_{\lambda=0}^{N-1} \sum_{\nu=0}^{N-1} \left( \hat{f}_{j-1,k} + \hat{f}_{j+1,k} + \hat{f}_{j,k-1} + \hat{f}_{j,k+1} - 4\hat{f}_{j,k} \right) \right\}
\]

Expression 5.9 may be evaluated by convolving \( \hat{f}_{jk} \) with a matrix \( c \) such that \( c \) takes the form

\[
\begin{pmatrix}
-4 & 1 \\
1 & 0 \\
1 & \phi
\end{pmatrix}
\]

The convolution of \( \hat{f}_{jk} \) and \( c \) may be mapped into multiplication by taking the Fourier transform of the two sequences. It can then be shown that we wish to

\[
\min \left\{ \sum_{\lambda=0}^{N-1} \sum_{\nu=0}^{N-1} \left( \mathcal{F}_k \hat{f}_{\lambda m} \mathcal{F}_m \hat{f}_{\lambda m} \right) \right\}
\]

where \( \mathcal{F}_m \) represents the Fourier transform of a second difference operator. The minimisation must be done so that inequality 5.6 is maintained. Hence the restoration problem is to find a solution \( \hat{F}_{\lambda m} \) that minimises 5.11 subject to expression 5.6 being obeyed. A solution may be found by the method of undetermined multipliers. If \( \lambda \) is
a Lagrangian multiplier then the Lagrangian, \( L \), is defined as
\[
L = N^2 \sum_{f=0}^{N-1} \sum_{m=0}^{N-1} C_{fm} c_{zm} \hat{f}_{zm} \hat{p}_{zm} + \lambda \left\{ \sum_{f=0}^{N-1} \sum_{m=0}^{N-1} \left( G_{zm} - NH_{zm} \hat{p}_{zm} \right) \left( G_{fm} - NH_{fm} \hat{p}_{fm} \right) - \varepsilon \right\}
\]
5.12
By differentiating \( L \) with respect to the real and imaginary parts of \( \hat{p}_{zm} \) and setting the result to zero gives
\[
2 C_{zm} c_{zm} \hat{p}_{zm} N^2 + \lambda \left\{ 2NH_{zm} \left( G_{zm} - NH_{zm} \hat{p}_{zm} \right) \right\} = 0
\]
5.13
Note that when \( \lambda \) is zero the constraint is inactive which implies that the data are not statistically significantly different from zero. This result is therefore a trivial solution. So if \( \lambda \) is greater than zero it follows that
\[
\hat{p}_{zm} = \frac{-H_{zm} G_{zm}}{N(H_{zm} H_{zm} + \gamma C_{zm} c_{zm})}
\]
5.14
where \( \gamma = 1/\lambda \).

In this study two different two dimensional operators have been successfully implemented. The first is based on Phillip's operator generalised for two dimensions, which in fact is the Laplacian operator
\[
\nabla^2 = \Delta_x^2 + \Delta_y^2
\]
5.15
The digital Laplacian operator is represented by the matrix \( c \) in 5.10. It should be noted that the choice of \( c \) is such that it is a real even sequence. The complete
matrix has five non zero elements; the remaining \( N^2 - 5 \) elements are set to zero. Because the matrix is real and even the FFT will contain no imaginary components. The complex conjugate of \( C_{jm} \) (see equation 5.14) is then itself. Figure 5.1 shows an isometric plot of \( C \bar{C} \). The origin of the frequency scale is at the centre of the image. It should be noted that the amplitude of \( C \bar{C} \) increases from zero at the origin to a maximum at a displacement of \( N/2 \) elements from the origin. Inspection of equation 5.14 shows that at low frequencies \( C \bar{C} \) will have little influence on the value of \( \hat{F}_{jm} \), and an increasing influence at higher frequencies. The amount of influence or smoothing depends on the value of the constant \( \gamma \). In the limit as \( \lambda \) approaches infinity, \( \gamma \) will approach zero, and the restoration becomes identical to inverse filtering. This solution is equivalent to \( \epsilon \) approaching zero, in other words there are no data errors and the solution will become the exact one.

Figure 5.1 shows that the discrete Laplacian operator is not radially symmetric in the frequency domain. A second difference radially symmetric operator may be represented by developing a sequence whose magnitude is proportional to the square of the radius from the origin in the frequency domain. The motivation for this stems from the fact that in the continuous case
FIGURE 5.1 Two dimensional Laplacian operator.
The function rises from zero at the centre to unity at the corners.
\[ \mathcal{I} \left\{ \nabla^2 \mathcal{I}(x,y) \right\} = (-1)^{\frac{\partial}{2}} q \mathcal{F}(u,v) \]  
5.16

where \( \frac{\partial}{2} \) is a positive integer and

\[ q = (u^2 + v^2)^{1/2} \]  
5.17

A plot of \( q^2 \) for the case when \( u,v \) take discrete values \( 1,m \) in the range \( 0 \) to \( N-1 \) is shown in figure 5.2.

5.3 SELECTION OF GAMMA

The value of \( \gamma \) in equation 5.14 has to be found such that the constraint equation is satisfied. Assuming that the constraint is active then if the expression for \( \hat{F}_{x_m} \) is substituted into equation 5.6

\[
\sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \left( \frac{G_{x_m} - \frac{H_{x_m} H_{x_m} G_{x_m}}{(H_{x_m} H_{x_m} + \gamma C_{x_m} C_{x_m})}}{G_{x_m} - \frac{H_{x_m} H_{x_m} G_{x_m}}{(H_{x_m} H_{x_m} + \gamma C_{x_m} C_{x_m})}} \right) \right\} = \varepsilon \]  
5.18

After rearranging terms it reduces to

\[ \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \left( \frac{\gamma C_{x_m} C_{x_m}}{H_{x_m} H_{x_m} + \gamma C_{x_m} C_{x_m}} \right)^2 G_{x_m} \overline{G_{x_m}} \right\} = \varepsilon \]  
5.19

The actual value of \( \gamma \) has to be found iteratively. An initial value of \( \gamma \) is chosen and equation 5.19 is used to evaluate the residual

\[ \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \left( \frac{\gamma C_{x_m} C_{x_m}}{H_{x_m} H_{x_m} + \gamma C_{x_m} C_{x_m}} \right)^2 G_{x_m} \overline{G_{x_m}} \right\} \]  
5.20
FIGURE 5.2 Radially symmetric operator.

The function rises from zero at the centre to unity at the corners.
which can then be compared with the statistically estimated value of the error in the object distribution. If the residual and $E$ are within a predefined tolerance of each other then the value of $Y$ is acceptable. If not, and the residual is less than $E$, $Y$ is incremented and the residual recalculated. Similarly if the residual is greater than $E$, $Y$ is decremented and the residual recalculated. This procedure is based on the fact that for $Y$ greater than zero the substitution of equation 5.14 into 5.6 results in expression 5.20 being an increasing function of $Y$. The iteration should proceed until a suitable value of $Y$ is found. At this point $\hat{F}_{km}$ may be calculated using equation 5.14. The inverse FFT will yield a solution $\hat{f}_{jk}$.

In order to try and optimise the time required to find $Y$ two iteration methods have been tried. The first was the Newton-Raphson technique and the second the Method of Bisection. The Newton-Raphson technique is a numerical method for finding the roots of an equation. It is particularly suitable when approximate values of the roots are known a priori. To implement the technique the iteration for $Y$ can be formulated in the following way

$$Y_{d+1} = Y_d - \frac{f(Y_d)}{f'(Y_d)}$$

5.21
where
\[
f'(Y_d) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \left( \frac{Y_d C_{cm} \bar{C}_{em}}{H_{cm} H_{em} + Y_d C_{cm} \bar{C}_{em}} \right)^2 G_{am} \bar{G}_{em} \right\} - \varepsilon \tag{5.22}
\]
and \( f'(Y_d) \) is the first derivative of \( f(Y_d) \) with respect to \( Y_d \). Equation 5.21 states that if \( Y_d \) is an approximate root then a second estimate of the root can be obtained in \( Y_{d+1} \). Usually the iteration procedure will converge on a value of \( Y \); the iteration is allowed to proceed until little change occurs between consecutive estimates of \( Y_d \) or, as in this case, until \( f(Y_d) \) represents a sufficiently close approximation to \( \varepsilon \). Although differentiation of \( f(Y_d) \) is straightforward the last term in equation 5.21 becomes rather unwieldy. After differentiation and manipulation of terms the result is
\[
Y_{d+1} = Y_d - \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \left\{ \left( \frac{Y_d C_{cm} \bar{C}_{em}}{H_{cm} H_{em} + Y_d C_{cm} \bar{C}_{em}} \right)^2 G_{am} \bar{G}_{em} \right\} - \varepsilon \tag{5.23}
\]
Typically, using scintigrams as data, the iteration would terminate after between 2 and 20 cycles. However although this iteration method apparently works successfully and converges quite rapidly there have been occasions when consecutive estimates of \( Y \) would diverge away from the correct value. This was subsequently found to be a problem of rounding errors in the PDP 12 implementation. Evaluation of the cubic factor in the denominator of the
second term of equation 5.23 over the frequency spectrum shows that it can take particularly large values which are outside the seven figure significance of the PDP 12 floating point processor. The problem could be circumvented by the purchase of double precision hardware for the PDP 12 or alternatively by using a different iteration method such as the Method of Bisection.

Although the Method of Bisection may take more iterations to converge to a good estimate of \( \gamma \), it has always been found to give credible results on the PDP 12. In this method an approximate value of \( \gamma \) is chosen and the residual computed in the normal way. The value of \( \gamma \) is then doubled if the residual is less than \( \mathcal{E} \), and a new residual evaluated. The process repeats until a value of \( \gamma \) is found which over-estimates the probable error. At this point the required value of \( \gamma \) has been bracketed between zero and some top value. The process then continues by bisecting the interval between the current evaluation and the previous one until a solution to equation 5.19 is found which is within a predefined tolerance of \( \mathcal{E} \). During this procedure each iteration will bracket \( \gamma \) within an interval half as large as the previous one.
5.4 COMPUTER IMPLEMENTATION

The only extra information required to implement constrained deconvolution over that required for inverse filtering is the matrix $\mathbf{C} \mathbf{C}^*$, and some knowledge of the expected error in the object distribution, $g_{jk}$. The program used for the two dimensional deconvolution will be described as a set of separate computational steps. It should be remembered that the program has been written for a mini-computer system; it could be made much more efficient if run on a large machine with the program and data permanently resident in memory.

STEP 1: Computation of the FFT of the image data, $g_{jk}$, and the impulse response, $h_{jk}$. The FFT implementation is by the method described in section 4.4. This process produces two complex arrays, $G_{lm}$ and $H_{lm}$, which are stored as four disc files.

STEP 2: Computation of the error, $E$. For a Poisson distribution the total expected error in the object distribution is given by the total number of detected quanta within the image. Inspection of equation 4.11 shows that the required sum can be derived from the FFT of $g_{jk}$. If $l$ and $m$ are zero, then

$$G_{00} = \frac{1}{N} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} g_{jk}$$

5.24
In other words the variance of the data is given by $N$ times the components at the origin of the discrete Fourier transform. By definition the complex part of $G_{\lambda m}$ must be zero. In the program $G_{\infty}$ is obtained by reading the first record of the file containing the real part of $G_{\lambda m}$.

STEP 3: Scale $H_{\lambda m}$ so that the restoration sequence $\hat{f}_{jk}$ will be of similar magnitude to $g_{jk}$. This can most easily be achieved by scaling the integral of the impulse response to unity. The scaling factor can be derived, by reasoning similar to that described in step 2, from the components at the origin of the discrete Fourier transform of $h_{jk}$. Each component in the sequence $H_{\lambda m}$ is divided by $NH_{\infty}$.

STEP 4: Compute the complex products of $\overline{H}_{\lambda m} \cdot G_{\lambda m}$, $\overline{H}_{\lambda m} \cdot \overline{H}_{\lambda m}$ and $G_{\lambda m} \cdot \overline{G}_{\lambda m}$ which will be required for implementation of equation 5.19 and 5.14. The results of these calculations are stored as temporary disc files, which physically occupy the same space as $G_{\lambda m}$ and $H_{\lambda m}$. The array $\overline{C C}$ is held permanently on the disc. This file is image independent, so it is more efficient if this matrix is created once by an auxiliary program.

STEP 5: Perform the iteration for $\gamma$ using equation 5.19. An acceptable value of $\gamma$ is taken as one that results in a residual which is within $\pm 10\%$ of $\varepsilon$.

STEP 6: Use equation 5.14 to evaluate $\hat{F}_{\lambda m}$. The result is held as two disc files which physically occupy the same
STEP 7: Use the inverse FFT to produce the final image, \( \hat{f}_{jk} \).

5.5 RESULTS WITH THE PHANTOM MODEL

In line with the data presented in chapter 4 each deconvolution method will be tested with the data from the model William's phantom (figure 4.13) and the modified gaussian point spread function of equation 4.19. Figures 5.3 and 5.4 show results of constrained deconvolution for the two smoothing functions in figures 5.1 and 5.2 respectively. The results are very similar suggesting that the different shapes of the two smoothing functions do not significantly affect the deconvolution. This is not unexpected because the smoothing functions diverge from each other only at high frequencies where the object data have only minimal spectral contributions.

From equation 5.14 the restoration filter response may be written as

\[
R_{\lambda \mu} = \frac{\overline{H}_{\lambda \mu}}{N \left( H_{\lambda \mu} \overline{H}_{\lambda \mu} + \gamma C_{\lambda \mu} \overline{C}_{\lambda \mu} \right)}
\]

5.25

A plot of this filter function used to derive the result in figure 5.3 is shown in figure 5.5. The Method of Bisection, described earlier, gave a value of \( \gamma \) of
FIGURE 5.3 Example of deconvolution using constrained optimisation with the Laplacian operator
FIGURE 5.4 Example of deconvolution using constrained optimisation with the radially symmetric operator
FIGURE 5.5  Restoration filter response used to obtain the result in figure 5.3.

The response rises from unity at the centre to a maximum of 1.32 at approximately 0.3 cycles per cm.
0.003906. The filter response moves from unity at the origin, follows the approximate path of the inverse filter, reaches a maximum of 1.32 and then falls smoothly to zero. The point at which the inverse filter "folds over" is determined by $\gamma$ which in turn is determined by the estimated noise in the data. It is instructive to consider restorations of the same image with manually selected values of $\gamma$. Two examples are shown for $\gamma$ values of one tenth and ten times the value computed for figure 5.5. When $\gamma$ is artificially reduced the restoration filter resembles the inverse filter to higher frequencies (see figure 5.7) the result of which is to allow higher frequencies to have some influence on the restoration process. This result is demonstrated in figure 5.6. Although the refocussing of the data is theoretically sharper it is at the expense of image noise. Conversely if $\gamma$ is increased the restoration filter will behave more like a low pass filter; the response, shown in figure 5.9, shows little amplification of any of the spatial frequencies. Figure 5.8 contains the associated deconvolved image; the data are heavily smoothed. It is interesting to note that $\hat{F}_{jk}$ changes quite slowly for relatively large changes in $\gamma$. This characteristic is useful in that it implies that the estimation of $\gamma$ is not too critical. For example an estimate of $\gamma$ to within a few percent of the optimum value will be quite adequate
PROGRAM: TIMJ (REV N)  DATE: 17- 4-78
IDENTIFICATION: MODP3  (GAMMA=8.3906E-03)

MAX COUNT  222  (ROW 17  COL 35)
TOTAL BGD SUB  8
TOTAL COUNT  179704
MULTI-CYCLES  1
THRESHOLD (%)  100

FIGURE 5.6
FIGURE 5.7 Restoration filter response used to obtain the result in figure 5.6.

The response rises from unity at the centre to a maximum of 2.25 at approximately 0.5 cycles per cm.
<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Count</td>
<td>220</td>
</tr>
<tr>
<td>Total BGD Sub</td>
<td>0</td>
</tr>
<tr>
<td>Total Count</td>
<td>179883</td>
</tr>
<tr>
<td>Multi-Cycles</td>
<td>1</td>
</tr>
<tr>
<td>Threshold (X)</td>
<td>100</td>
</tr>
</tbody>
</table>

FIGURE 5.8
FIGURE 5.9 Restoration filter response used to obtain the result in figure 5.8.

The response rises from unity at the centre to a maximum of 1.05 at approximately 0.1 cycles per cm.
for producing a good estimate of \( \hat{f}_{jk} \).

5.6 SUMMARY

The two dimensional implementation of deconvolution using constrained optimisation has been developed for the case where noise is Poisson distributed. The technique has been successfully implemented on a mini-computer and results have been given for the model William's phantom. In chapter 10 the constrained optimisation along with other restoration methods will be used on clinical data.
CHAPTER 6

DECONVOLUTION BY SERIES EXPANSION

6.1 INTRODUCTION

As long ago as 1931 Van Cittert suggested that if an estimate of the deconvolution equation, \( \hat{f}_{jk} \), is convolved with the impulse response then the resulting \( \hat{g}_{jk} \), should be a good approximation of the original observed data. This simple postulation leads to a method of deconvolution which incorporates within it successive convolutions. It will also be shown in this chapter that the method leads to a restoring algorithm which is very similar to that obtained by developing the inversion equation as a binomial series expansion. It is this latter method which has been implemented on a digital computer.

6.2 VAN CITTERT'S METHOD

Van Cittert's (1931) method is an iterative technique which starts by using the observed data as a first approximation to \( \hat{f}_{jk} \). This approximation is then convolved with the impulse response to yield a second
solution. The difference between the two solutions is used to apply a correction to the second approximation to give a third estimate of the object data. This procedure can be repeated a number of times until a reasonable solution is obtained. Usually, after a given number of iterations the solution will not change significantly from one iteration to another.

If superscript, $\beta$, denotes iteration number then the mathematical operations may be summarised as

$$\hat{G}_{\beta m} = N H_{\beta m} \hat{F}_{\beta m}$$  \hspace{1cm} 6.1

and

$$\hat{F}_{\beta+1}^{\beta+1} = \hat{F}_{\beta m} + (G_{\beta m} - \hat{G}_{\beta m})$$  \hspace{1cm} 6.2

then by substitution

$$\hat{F}_{\beta+1}^{\beta+1} = \hat{F}_{\beta m} (1 - NH_{\beta m}) + G_{\beta m}$$  \hspace{1cm} 6.3

By starting with $F_{\beta m} = G_{\beta m}$ on the right side of equation 6.3, and by successively substituting the left side into the right, a geometric series is developed such that

$$\hat{F}_{\beta m} = G_{\beta m} \sum_{p=0}^{\beta} (1 - NH_{\beta m})^p$$  \hspace{1cm} 6.4

which may be analytically summed to
\[ \hat{f}_{\lambda m}^\lambda = \frac{G_{\lambda m}}{N H_{\lambda m}} \left\{ 1 - (1 - N H_{\lambda m})^{\beta+1} \right\} \text{ for } H_{\lambda m} \neq 0 \quad 6.5 \]

As would be expected, in the limit as \( \beta \) goes to infinity equation 6.5 reverts to an inverse filter restoration, providing \(|1-NH_{\lambda m}|\) is always less than unity. It is most important that the latter condition holds otherwise the series will not converge. Smaller values of \( \beta \) will produce an element of smoothing which controls the instability problems associated with the inverse filter. At first sight it might appear that equation 6.5 may be potentially ill-behaved because as \( H_{\lambda m} \) approaches zero the denominator will also approach zero. Generally problems do not arise because the solution becomes proportional to the noise; this can be verified by application of l'Hopital's rule. If equation 6.5 is split into numerator and denominator components such that

\[ U_{\lambda m} = G_{\lambda m} (1 - (1 - N H_{\lambda m})^{\beta+1}) \quad 6.6 \]

and

\[ V_{\lambda m} = N H_{\lambda m} \quad 6.7 \]

then

\[ \frac{d U_{\lambda m}}{d H_{\lambda m}} = NG_{\lambda m} (1 + \beta)(1 - N H_{\lambda m})^\beta \quad 6.8 \]

and
\[ \frac{dV_{em}}{d H_{em}} = N \] 6.9

So
\[ \frac{dW_{em}}{d H_{em}} = G_{em} (\beta + 1) (1 - NH_{em})^\beta \] 6.10
\[ \frac{d V_{em}}{d H_{em}} / \frac{d V_{em}}{d H_{em}} \]

In the limit as \( H_{em} \) approaches zero
\[ \lim_{H_{em} \to 0} \left\{ \frac{dW_{em}}{d H_{em}} \right\} = G_{em} (\beta + 1) \] 6.11
and by l'Hopital's rule it follows that
\[ \lim_{H_{em} \to 0} \left\{ \hat{f}_{em}^\beta \right\} = G_{em} (\beta + 1) \] 6.12

Using equation 5.4 to substitute for \( G_{em} \) yields
\[ \lim_{H_{em} \to 0} \left\{ \hat{f}_{em}^\beta \right\} = E_{em} (\beta + 1) \] 6.13

Equation 6.13 confirms that as \( H_{em} \) approaches zero the corresponding \( \hat{f}_{em}^\beta \) will be proportional to \( E_{em} \).

Jansson (1970) tried Van Cittert's technique to deconvolve infra-red spectra. He noted that Van Cittert's method could produce large erroneous oscillations in the deconvolved spectra, particularly if the recorded data were strongly contaminated with noise. Jansson subsequently modified the technique slightly by incorporating a variable relaxation parameter which transforms the method into a non-linear constrained restoration. The constraint prevents \( \hat{f}_{em}^\beta \) from taking
unphysical values; for example, it prevents $\hat{f}_{jk}$ from becoming negative.

6.3 DECONVOLUTION BY SERIES EXPANSION

The simple summation that evolves from considering the successive convolution method of Van Cittert led this author to consider a similar summation method in which the inverse filter is expressed as a binomial series expansion. Remembering that in an ideal situation the restoring filter would be

$$ R_{\alpha} = 1/NH_{\alpha} $$

then the expansion technique consists in rearranging the denominator so that it is real and can be expressed as a polynomial. It will be assumed that $H_{\alpha}$ is always non zero.

A derivation of the restoration function follows. Let

$$ R_{\alpha} = \frac{H_{\alpha}}{N|H_{\alpha}|^2} $$

Expression 6.16 may be expanded as a polynomial in $|H_{\alpha}|^2$ so that
The summation is convergent for \(-1 < (1 - |H_{\xi m}|^2) < 1\) or \(|H_{\xi m}| < \sqrt{2}\). Expression 6.17 may be rearranged so that it consists of two separate summations.

\[
R_{\xi m} = \frac{\bar{H}_{\xi m}}{N} \sum_{\beta=0}^{n} (1 - |H_{\xi m}|^2)^{\beta} + \frac{\bar{H}_{\xi m}}{N} \sum_{\beta=n+1}^{\infty} (1 - |H_{\xi m}|^2)^{\beta} \tag{6.18}
\]

The first term of expression 6.18 is a geometric series that can be simplified so that

\[
R_{\xi m} = \frac{\bar{H}_{\xi m}}{N} \frac{(1 - (1 - |H_{\xi m}|^2)^{n+1})}{(1 - (1 - |H_{\xi m}|^2))} + \frac{\bar{H}_{\xi m}}{N} \sum_{\beta=n+1}^{\infty} (1 - |H_{\xi m}|^2)^{\beta} \tag{6.19}
\]

\[
= \frac{\bar{H}_{\xi m}}{N} \frac{(1 - (1 - |H_{\xi m}|^2)^{n+1})}{|H_{\xi m}|^2} + \frac{\bar{H}_{\xi m}}{N} \sum_{\beta=n+1}^{\infty} (1 - |H_{\xi m}|^2)^{\beta} \tag{6.20}
\]

If \((1 - |H_{\xi m}|^2)\) is small then the second term in expression 6.20 contributes little. If \(H_{\xi m}\) approaches zero then the first term of expression 6.20 is well defined but the second term approaches infinity, so that it is the latter term which contributes the pole in expression 6.14 when \(H_{\xi m}\) is zero. The second term may be dropped and \(R_{\xi m}\) redefined so that

\[
R_{\xi m}^n = \frac{\bar{H}_{\xi m}}{N} \frac{(1 - (1 - |H_{\xi m}|^2)^{n+1})}{|H_{\xi m}|^2} \tag{6.21}
\]

thus removing the original pole in \(R_{\xi m}\) but retaining a value close to the original at other values of \(H_{\xi m}\). Hence the solution to the restoration problem is given by
The similarity between the Van Cittert result (equation 6.5) and the series expansion is now apparent. Application of l'Hopital's rule shows that for finite $n$, $R^n_{\hat{f}_{\lambda m}}$ tends to zero as $H_{\lambda m}$ approaches zero, so that equation 6.22 is usually well-behaved. This is probably an advantage of the method. Equation 6.13 shows that the Van Cittert method lets $\hat{F}_{\lambda m}$ become proportional to the noise. At low frequencies where $H_{\lambda m}$ is relatively large, $R^n_{\hat{x}_{\lambda m}}$ will produce amplification of $G_{\lambda m}$. The basic shape of $R^n_{\hat{x}_{\lambda m}}$ will therefore produce enhancement at only those frequencies where the signal to noise ratio is high. The filter responses for $n$ equal to 0, 2, 5 and 10 are shown in figures 6.1 to 6.4. For these examples $H_{\lambda m}$ has been obtained from the Fourier transform of the modified Gaussian function (equation 4.19). Figure 6.1, in which $n$ equals zero, just corresponds to a plot of $H_{\lambda m}$; there is no enhancement and the processed image will be a low pass filtered version of the original. As $n$ increases, $R^n_{\hat{x}_{\lambda m}}$ peaks at both higher frequencies and greater magnitudes. In the limit as $n$ approaches infinity the pure inverse filter response will evolve and $R^n_{\hat{x}_{\lambda m}}$ will also tend to infinity. This of course is what would be expected because the derivation employed a series expansion for the representation of the filter.
FIGURE 6.1 Binomial restoration filter for $n = 0$

The zero frequency component peaks at a value of unity.
FIGURE 6.2 Binomial restoration filter for $n = 2$

The response is unity at the centre and rises to a maximum of 1.24 at approximately 0.3 cycles per cm.
FIGURE 6.3 Binomial restoration filter for $n = 5$.
The response is unity at the centre and rises to a maximum of 1.65 at approximately 0.4 cycles per cm.
FIGURE 6.4 Binomial restoration filter for $n = 10$.
The response is unity at the centre and rises to a maximum of 2.18 at approximately 0.5 cycles per cm.
6.4 CHOICE OF n

The way in which $R_{\lambda m}^n$ behaves for different values of $n$ has been described but so far no criterion has been established to fix $n$ for a particular image restoration. The method adopted here is to iterate on $n$ until a value is found which produces an error in $\hat{F}_{\lambda m}^n$ of similar magnitude to that of $G_{\lambda m}$. If $\epsilon$ is the total expected error then we require that

$$\epsilon \approx \sum_{\lambda=0}^{N-1} \sum_{m=0}^{N-1} \left( G_{\lambda m} - NH_{\lambda m} \hat{F}_{\lambda m}^n \right) \left( G_{\lambda m} - NH_{\lambda m} \hat{F}_{\lambda m}^n \right)$$

6.23

The derivation of equation 6.23 is covered in section 5.2. The result of multiplying out the right side of equation 6.23 and using equation 6.22 is

$$\epsilon \approx \sum_{\lambda=0}^{N-1} \sum_{m=0}^{N-1} \left\{ G_{\lambda m} \bar{G}_{\lambda m} \left(1 - |H_{\lambda m}|^2\right)^{2(n+1)} \right\}$$

6.24

If $\epsilon$ is estimated from the variance of the data, then all the variables in equation 6.24 are known except for $n$. The value of $n$ can be found by repeated evaluation of the right side of the equation until both sides of the equation are in approximate agreement. Choosing $n$ in this way has worked quite successfully at controlling noise amplification during scintigraphic restorations.
Subsequently it has been pointed out to the author that the binomial series expansion is similar to the approach adopted by Metz (1969, 1974). His implementation is intended to be an iterative convolution process. He discusses the effects of varying $n$ in relation to the processed resolution and noise characteristics. However he does not define a criterion for the optimum value of $n$.

6.5 COMPUTER IMPLEMENTATION

Of all the deconvolution methods developed the binomial series expansion requires the least a priori information; only knowledge of the impulse response and the variance of the observed data are required. As for other deconvolution techniques the PDP 12 computer implementation will be described as a set of computational steps.

**STEP 1:** Computation of the FFT of the image data, $g_{jk}$, and the impulse response, $h_{jk}$. The FFT implementation is by the method described in section 4.4. This process produces two complex arrays, $G_{lm}$ and $H_{lm}$, which are stored as four disc files.

**STEP 2:** Computation of the error, $\mathcal{E}$. The method used is the same as that described for the constrained optimisation. Equation 5.24 is used to evaluate $\mathcal{E}$. 
STEP 3: Scale $H_{lm}$ for all $l,m$ such that $H_{oo}$ takes a value of unity. This ensures that $\hat{f}_{jk}$ will be of similar magnitude to $g_{jk}$.

STEP 4: Compute the complex products of $H_{lm}.G_{jm}$, $H_{km}.H_{jm}$ and $G_{km}.\tilde{G}_{jm}$ which will be required for implementation of equations 6.22 and 6.24. The results of these calculations are stored as temporary disc files, which physically occupy the same space as $G_{km}$ and $H_{km}$.

STEP 5: Perform the iteration for $n$ using equation 6.24. The value of $n$ is incremented and the right side of equation 6.24 is evaluated. If this value is within $\pm10\%$ of $\mathcal{E}$ the current value of $n$ is accepted. Otherwise $n$ is incremented and a new estimate made. It has been found that usually $n$ can be located within 10 iterations.

STEP 6: Use equation 6.22 to evaluate $\hat{F}_{lm}^{n}$. The result is held as two disc files which physically occupy the same space as $G_{km}$.

STEP 7: Use the inverse FFT to produce the final image, $\hat{f}_{jk}^{n}$.

6.6 RESULTS WITH THE PHANTOM MODEL
The restoration of the Williams phantom is depicted in Figure 6.5. The computer program produced a value of $n$ equal to 2 (as in Figure 6.2); this value satisfies the requirement that the left and right sides of equation 6.24 agree to within 10% of each other. In this instance the restoring filter is similar to that produced by the constrained optimisation method. The similarities of the two types of restoration are apparent if Figure 6.5 is compared with Figure 5.3.

6.7 SUMMARY

The binomial series expansion method of deconvolution has been compared to the work of Van Cittert, and adapted to scintigraphic image restorations where the limit of summation of the expansion has been equated to the total expected error of the observed data. As an initial example deconvolution of a model phantom has been demonstrated. Further examples of the technique using clinical data will follow in Chapter 10.
PROGRAM: TIMJ (REV N)          DATE: 17-4-78
IDENTIFICATION: MODP8 (BINOMIAL SERIES EXPANSION)

MAX COUNT  220 (ROW 16 COL 39)
TOTAL BGD SUB  0
TOTAL COUNT  179211
MULTI-CYCLES  1
THRESHOLD (x)  100

FIGURE 6.5
CHAPTER 7

WIENER FILTERING

7.1 INTRODUCTION

The classical signal processing technique known as Wiener filtering was first described by Wiener in 1942 and later Levinson (1946, 1947) published two papers on the root mean square criterion for filter design. Since then the technique has been used in many fields of study, and in medical scintigraphy, Lorenz (1972), Pistor (1972, 1972), Bone (1973), Kirch (1973) and Pullan (1976) have all tried their own methods of implementation. Most workers in scintigraphy have built on the formulations derived by Helstrom (1966) and Slepian (1967) who were primarily concerned with optical data processing. In particular they both assumed some knowledge of the noise characteristics, the ratio of the signal to noise power, and the impulse response. The main criticism, from the point of view of implementation in scintigraphy, which will be enlarged upon later, is that the noise was modelled as additive and uncorrelated with the object data. These constraints hardly hold for data originating from radioactive decay, in which the noise or variance of the data is equal to the number of detected quanta. Budinger (1972) recognises the problem by pointing out
that the conventional Wiener filter assumes that cross-correlation between object and noise is negligible. He offers no solution to the problem other than saying that the approach becomes more appropriate when the pixel counts are high and the relative standard deviation low.

To this author's knowledge no derivation for the generalised two dimensional Wiener filter has been presented in terms of the discrete Fourier transform. For this reason a fairly detailed account of the filter design will be given, paying particular attention to the properties of noise.

7.2 DERIVATION OF THE M.M.S.E. FILTER

One way to optimise the image restoration process is to minimise the mean squared error between the restored data and the original image. The filter has been used frequently by many workers because of its mathematical tractability.

If \( \hat{f}_{jk} \) represents an estimate of the object distribution, \( f_{jk} \), then the requirement for the minimum mean squared error is to minimise
\[ \varepsilon_w^2 = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} (\hat{f}_{jk} - f_{jk})^2 \]  

7.1

By virtue of Parseval's relationship the minimisation can be done in the frequency domain

\[ \varepsilon_w^2 = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} |F_{jm} - F_{jm}|^2 \]  

7.2

If \( R_{W_{km}} \) is defined as the Wiener restoring filter, then using the same notation as in equation 5.2

\[ \hat{F}_{km} = R_{W_{km}} \left( N H_{km} F_{km} + E_{km} \right) \]  

7.3

Hence \( \varepsilon_w^2 \) may be written as

\[ \varepsilon_w^2 = \sum_{j=0}^{N-1} \sum_{m=0}^{N-1} \left| R_{W_{km}} \left( N H_{km} F_{km} + E_{km} \right) - F_{km} \right|^2 \]  

7.4

Equation 7.4 may be expressed in terms of real and imaginary parts

\[ \varepsilon_w^2 = \sum_{j=0}^{N-1} \sum_{m=0}^{N-1} \left[ \left( R_{W_{km}} \left( N H_{km} F_{km} + E_{km} \right) - F_{km} \right) \left( R_{W_{km}} \left( N H_{km} F_{km} + E_{km} \right) - F_{km} \right) \right] \]  

7.5

To find a solution, \( R_{W_{km}} \), which minimises \( \varepsilon_w^2 \) it is necessary to differentiate equation 7.5 with respect to \( R_{W_{km}} \):

\[ \frac{\partial \varepsilon_w^2}{\partial R_{W_{km}}} = 2 \left( N H_{km} F_{km} + E_{km} \right) \left( R_{W_{km}} \left( N H_{km} F_{km} + E_{km} \right) - F_{km} \right) \]  

7.6

The reader is referred to the appendix at the end of this chapter for a treatment of the differentiation of real functions of a complex variable. At the minimum \( \frac{\partial \varepsilon_w^2}{\partial R_{W_{km}}} \) is zero so that
After taking the complex conjugate of equation 7.7 and rearranging terms

$$R_{W_{\lambda m}} = \frac{F_{\lambda m}}{NH_{\lambda m}F_{\lambda m} + E_{\lambda m}}$$  \text{(7.8)}

This might seem a surprisingly obvious result because the denominator of equation 7.8 is just $G_{\lambda m}$, the observed data. In section 4.2 it has already been shown that division of Fourier transforms in the presence of noise can (and usually will) lead to unpredictable results.

The important conclusion from equation 7.8 is that the minimum mean square error filter is just the inverse filter. The reason that the Wiener filter can apparently be made to work is by introducing constraints on the characteristics of the noise. It is necessary to define the cross-correlation of $F_{\lambda m}$ and $E_{\lambda m}$ as zero for all $\lambda,m$. By manipulating equation 7.8 further:

$$R_{W_{\lambda m}} = \frac{F_{\lambda m}(NH_{\lambda m}F_{\lambda m} + E_{\lambda m})}{(NH_{\lambda m}F_{\lambda m} + E_{\lambda m})(NH_{\lambda m}F_{\lambda m} + E_{\lambda m})}$$  \text{(7.9)}

$$= \frac{NH_{\lambda m}F_{\lambda m}F_{\lambda m} + F_{\lambda m}E_{\lambda m}}{N^2H_{\lambda m}F_{\lambda m}F_{\lambda m} + E_{\lambda m}E_{\lambda m} + (NH_{\lambda m}F_{\lambda m}E_{\lambda m} + NH_{\lambda m}F_{\lambda m}E_{\lambda m})}$$  \text{(7.10)}

and now by stipulating the cross-correlation of $F_{\lambda m}$ and $E_{\lambda m}$ is zero, the terms containing the cross products will vanish so that
\[ R_{\omega m} = \frac{N \bar{H}_{\omega m}}{N^2 \bar{H}_{\omega m} \bar{H}_{\omega m} + \bar{e}_{\omega m} \bar{e}_{\omega m} / \bar{f}_{\omega m}} \quad \text{for } \bar{f}_{\omega m} \neq \phi \] 7.11

The condition that \( \bar{f}_{\omega m} \) should not be zero is required for the evaluation of the denominator of equation 7.11. This usually presents no problems on a digital computer when floating point arithmetic is used, as the probability of \( \bar{f}_{\omega m} \) being exactly zero is very small. When \( \bar{f}_{\omega m} \) approaches zero (at high frequencies), \( R_{\omega m} \) will also approach zero.

If \( \phi_{\omega m} \) represents the image power and \( \phi_{e_{\omega m}} \) the noise power, then

\[ R_{\omega m} = \frac{N \bar{H}_{\omega m}}{N^2 \bar{H}_{\omega m} \bar{H}_{\omega m} + \bar{e}_{\omega m} \bar{e}_{\omega m} / \bar{f}_{\omega m}} \] 7.12

from which it follows

\[ \hat{f}_{\omega m} = \frac{N \bar{H}_{\omega m} \bar{G}_{\omega m}}{N^2 \bar{H}_{\omega m} \bar{H}_{\omega m} + \bar{e}_{\omega m} \bar{e}_{\omega m} / \bar{f}_{\omega m}} \] 7.13

Equation 7.12 is in the generally accepted form of the Wiener filter (see for example Helstrom (1966) or Slepian (1967)). The above derivation is mathematically elegant, although of course it is only applicable to periodic data.

Wiener filtering therefore requires a priori information about the impulse response and estimates of the image signal and noise power or auto-correlation functions. This amount of information required for the restoration is a major drawback of the Wiener approach. A
further criticism of the method is the requirement that the cross correlation of the object and noise data should be zero. This requirement is not just a useful assumption for simplifying the derivation of the Wiener filter but the very essence of it, without which one returns to an inverse filter as a solution, which has already been shown to be of little practical use. It remains an open question as to whether another practical filter could be derived using instead the assumption of Poisson distributed, signal dependent noise.

7.3 ESTIMATION OF SIGNAL TO NOISE POWER

The power spectrum of the object data plus noise is readily obtained from the discrete Fourier transform of $g_{jk}$. However it is another matter to try and separate the relative portions of the object and noise spectral components. Indeed if the object spectrum was known all would be known about the radioactive distribution other than its phase. Obviously what is required is an estimate of the power in the object distribution. The simplest approach would be to accept that the signal power components between all scintigrams possess the same features; in this case a model may be used to generate a noise free power spectrum by convolving it with the impulse response. Hence a typical or average power
spectrum may be computed which can be used to implement the Wiener filter. The main drawback of this technique is that the restoration filter function will always be based on the features of this average power spectrum. This will inevitably reduce the detectability of barely visible abnormalities in an image undergoing restoration, so defeating the purpose of deconvolution.

A second approach, which was the one used in this study, is to try and estimate the noise power of the measured distribution, \( g_{jk} \), and so deduce an estimate of the object power by taking the difference between the power spectrum of \( g_{jk} \) and the noise power estimate. The assumption is made that the scintigraphic imaging system behaves as a Poisson process, such that the noise, or variance of data, may be modelled as additive. Because it is impossible to evaluate noise exactly it will be estimated by considering the expected values of the auto-correlation function.

An estimate of the observed data, \( g_{jk} \), may be expressed as

\[
E\{g_{jk}\} = \hat{g}_{jk}
\]

where \( E \) means expected value, and \( \hat{g}_{jk} \) is the theoretical or mean value of \( g_{jk} \); in other words, as more and more
samples of $g_{jk}$ are taken their mean value will approach $g_{jk}$. The variance of $g_{jk}$ is defined as

$$\text{var} \{ g_{jk} \} = \mathbb{E} \{ (g_{jk} - \mathbb{E} g_{jk})^2 \}$$  \hspace{1cm} 7.15$$

from which it can be shown that for a Poisson distribution

$$\text{var} \{ g_{jk} \} = g_{jk}$$  \hspace{1cm} 7.16$$

and

$$\mathbb{E} \{ g_{jk}^2 \} = g_{jk}^2 + g_{jk}^2$$  \hspace{1cm} 7.17$$

An estimate of the circular auto-correlation, $R_{gg}$, of $g_{jk}$ is defined as

$$\mathbb{E} \{ R_{gg_{p,q}} \} = \mathbb{E} \left\{ \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} g_{jk} g_{j+p,k+q} \right\}$$  \hspace{1cm} 7.18$$

where the suffices are taken as modulo $N$.

$$\mathbb{E} \{ R_{gg_{p,q}} \} = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left[ \mathbb{E} \{ g_{jk} \} \mathbb{E} \{ g_{j+p,k+q} \} + \mathbb{E} \{ g_{jk} \} \delta_{p,q} \right]$$  \hspace{1cm} 7.19$$

where

$$\delta_{p,q} = \begin{cases} 1 & \text{for } p = 0, \quad q = 0 \\ \phi & \text{otherwise} \end{cases}$$

At this point it should be noted that the underlying assumption has been made that any two measurements of $g_{jk}$ are independent of each other. However in equation 7.19 the case of $p=0$ and $q=0$ has to be considered separately because this corresponds to zero shift in the auto-correlation summation; in this case of course the data are not independent and the second term involving a
delta function is introduced so that equation 7.17 is satisfied. The reader is referred to Metz's (1969) account for a fuller treatment of the noise process. However Metz does use Ergodic theory so his treatment differs slightly from the approach used here.

By using the definition in equation 7.14, equation 7.19 may be written as

\[
E\{R_{jk}\} = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left[ g_{jk} g_{j+p,k+q} + g_{jk} \delta_{p,q} \right] \tag{7.20}
\]

The power spectrum of \( g_{jk} \) is obtained by taking the discrete Fourier transform of \( R_{jk} \). If \( \Phi_{jk} \) represents the power spectrum with frequency components \( l \) and \( m \) then

\[
E\{\Phi_{jk}\} = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left[ E\{R_{jk}\} \exp\left\{-\frac{2\pi i}{N}(pl + qm)\right\} \right] \tag{7.21}
\]

\[
= \frac{1}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left[ \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \left( g_{jk} g_{j+p,k+q} + g_{jk} \delta_{p,q} \right) e^{-\frac{2\pi i}{N}(pl + qm)} \right] \exp\left\{-\frac{2\pi i}{N}(pl + qm)\right\} \tag{7.22}
\]

\[
= \frac{1}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} \left( g_{jk} g_{j+p,k+q} + g_{jk} \delta_{p,q} \right) e^{-\frac{2\pi i}{N}(pl + qm)} \tag{7.23}
\]

\[
= \frac{1}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} g_{jk} g_{j+p,k+q} e^{-\frac{2\pi i}{N}(pl + qm)} + \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} g_{jk} \delta_{p,q} e^{-\frac{2\pi i}{N}(pl + qm)} \tag{7.24}
\]

The first term of equation 7.24 is just the power spectrum of the observed data, which could be obtained by considering noise free data in which \( g_{jk} \) would equal \( g_{jk} \). The second term in equation 7.24 represents the contribution to the power spectrum of noise, so that
\[ \Phi_{\Omega_{\text{pm}}} = \frac{1}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \sum_{j'=0}^{N-1} \sum_{k'=0}^{N-1} \delta_{p_{\text{pm}}, q_{\text{pm}}} \exp \left\{ -\frac{2\pi i}{N} (p_{\text{pm}} + q_{\text{pm}}) \right\} \]  
7.25

\[ \Phi_{g_{\Omega_{\text{pm}}}} = \frac{1}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} g_{jk} \]  
7.26

\[ \Phi_{\Omega_{\text{pm}}} = T / N^2 \]  
7.27

where \( T \) is the total number of radioactive events in the array \( g_{jk} \). The important aspect of this result is that the noise power is white; it is a constant value independent of frequency. An estimate of the object power is obtained by taking the discrete Fourier transform of \( g_{jk} \) using the method described in section 4.5 and subtracting from it \( \Phi_{\Omega_{\text{pm}}} \).

7.4 COMPUTER IMPLEMENTATION

The Wiener filtering technique requires calculation of the FFTs of the image and the impulse response and estimates of the image and noise power spectra. A step by step description of the PDP 12 computer program is now given; as for other deconvolution methods the program could be made much more efficient if run on a large computer with the program and data permanently resident in memory.

STEP 1: Computation of the FFT of the image data and the impulse response. The FFT implementation is by the method
described in section 4.4. This process produces two complex arrays, $G_{lm}$ and $H_{lm}$, which are stored as four disc files.

STEP 2: Computation of the noise power, $\Phi_{e_{lm}}$. This is calculated using equation 7.27. Inspection of equation 4.11 shows that the required sum, T, can be derived from the FFT of $g_{jk}$. By using equation 5.24

$$T = NG_{oo}$$  

so that

$$\Phi_{e_{lm}} = G_{oo}/N$$  

STEP 3: Computation of the signal plus noise power spectrum, $\Phi_{g_{lm}}$. It is calculated by evaluating the square of the modulus of $G_{lm}$ at each frequency $l,m$. It is stored as a temporary disc file.

STEP 4: An estimate of the object power spectrum is obtained using the results from steps 2 and 3. In order to smooth out any discontinuities in $\Phi_{g_{lm}}$, it is filtered using a simple nine point smoothing process. This ensures that $\Phi_{g_{lm}}$ is reasonably well behaved. Discontinuities in $\Phi_{g_{lm}}$ can produce similar disturbances in the Wiener restoring filter which will lead to unpredictable estimates of $\hat{f}_{jk}$. The object power spectrum is then estimated by subtracting $\Phi_{e_{lm}}$ from the smoothed spectrum. If the resulting object power values are negative for any
1,m they are set equal to zero. Negative power components may occur because $\hat{\Phi}_{e_{lm}}$ is a statistically estimated quantity. This process will automatically discriminate against those components of the object power spectrum which are likely to be noisy. Theoretically this process does not really estimate the object power but the power contained in $g_{jk}$. Although a more precise algorithm can be derived it has been found to be ill-conditioned with genuine data. A complete analysis of the computational problems involved is complex. The premise that the system is modelled as purely additive is probably a contributing factor. Ideally one would prefer to make repeated estimates of $\hat{\Phi}_{f_{lm}}$ to obtain an average, but this is not possible when only one data set, $g_{jk}$, is available.

STEP 5: Computation of the complex products of $\bar{H}_{lm} \cdot G_{lm}$ and $H_{lm} \cdot \bar{H}_{lm}$ which are required for implementation of equation 7.13. The results of these calculations are stored as temporary disc files, which physically occupy the same space as $G_{lm}$ and $H_{lm}$.

STEP 6: Evaluation of $\hat{F}_{lm}$ using equation 7.13. The result is held as two disc files which physically occupy the same space as $G_{lm}$.

STEP 7: Use the inverse FFT to produce the final image, $\hat{f}_{jk}$.
7.5 RESULTS WITH THE PHANTOM MODEL

The result of subjecting the William's model to the Wiener filter is shown in figure 7.1. Comparison with the constrained optimisation (figures 5.3 or 5.4) shows that for this restoration the Wiener filter is the less aggressive. The restoration filter peaks rather more than the constrained optimisation filter, so that higher frequencies are amplified more greatly with the result that the restoration appears slightly more noisy. In this particular case a check was made to ensure that the estimated noise power was approximately equal to that actually generated using the Central Limit theorem. When the noise sequence was being generated (see section 4.8) in order to create the original Williams model, the noise values themselves were saved as an auxiliary disc file. The power spectrum of the noise data was computed and the average of the resultant data found. In fact the two noise powers agreed to within 7%. Some workers (for example Kirch (1973)) arbitrarily increase the estimated noise figure to increase the smoothing and thereby increase the statistical confidence of the resultant image. The result of increasing the estimated noise figure by a factor of four is shown in figure 7.3 and the associated Wiener filter in figure 7.4. The small hot and cold areas in the phantom have been blurred out; a
FIGURE 7.2 Wiener filter response used to obtain the result in figure 7.1.

The peak in the response has a value of 2.09 at approximately 0.4 cycles per cm.
IDENTIFICATION: MODP9  (WIENER FILTER, NOISE FACTOR OF 4)

MAX COUNT 230  (ROW 18  COL 38)
TOTAL BGD SUB 0
TOTAL COUNT 180629
MULTI-CYCLES 1
THRESHOLD (X) 100

FIGURE 7.3
FIGURE 7.4 Modified Wiener filter response after the noise power has been increased by a factor of 4. The peak in the response has a value of 1.61 at approximately 0.3 cycles per cm.
tradeoff has occurred between the amount of noise and the sharpness of the image.

The astute reader will have noticed that there is a very close similarity between the solutions for the constrained optimisation (equation 5.14) and the Wiener restoration filter (equation 7.13). The solutions become equivalent if

\[ N \gamma C_{2m} \bar{C}_{2m} = \frac{\Phi_{2m}}{\Phi_{f2m}} \quad 7.33 \]

Therefore the question arises: which of the solutions is better. The answer must lie in the amount of a priori information available. The Wiener method requires spectral estimates of both signal and noise which have already been shown to be tedious functions to evaluate. On the other hand the only quantitative measurement required for the constrained optimisation is an estimate of \( \mathcal{E} \). The only other stipulation is that the solution should be smooth. The constrained optimisation therefore requires less a priori information and is by far the easier of the two restoration methods to implement. It is surprising that the constrained approach has not been applied to scintigraphy before, where there is an abundance of literature available on the Wiener filter. The answer probably lies in the fact that the Wiener technique has frequently been applied in other fields of
study; the mathematical foundations have been laid down for the past thirty years.

7.6 SUMMARY

Although many workers have claimed to use Wiener filtering few (if any) have given a plausible account of the implementation of such filters in scintigraphy. This chapter has cleared up some of the apparent anomalies of the M.M.S.E. criterion for filter design, and given a relatively straightforward account of the derivation of the filter in terms of the discrete Fourier transform. Chapter 10 will cite further examples of the use of the Wiener filter along with other designs described in the text.
CHAPTER 7
APPENDIX

DIFFERENTIATION OF REAL FUNCTIONS OF A COMPLEX VARIABLE

A generalised form of equation 7.5 is

\[ y = (ax + b)(\overline{ax + b}) \]

where \( a, b \) and \( x \) are complex numbers. To differentiate \( y \) with respect to \( x \), the following steps may be taken.

Let

\[ x_r = \text{Re}[x] \]
\[ x_\lambda = \text{Im}[x] \]

and similarly for \( a_r, a_\lambda, b_r, b_\lambda \). Then \( y \) may be expressed as

\[ y = a\overline{a}x^2 + b\overline{b} + a\overline{b}x + ab\overline{x} \]

\[ = a\overline{a}(x_r^2 + x_\lambda^2) + b\overline{b} + (a_r + ia_\lambda)(b_r - ib_\lambda)(x_r + ix_\lambda) \]
\[ + (a_r - ia_\lambda)(b_r + ib_\lambda)(x_r - ix_\lambda) \]

Differentiating with respect to the real numbers \( x_r \) and \( x_\lambda \)

\[ \frac{dy}{dx_r} = 2a\overline{a}x_r + 2a_r b_r + 2a_\lambda b_\lambda \]
\[ \frac{dy}{dx_\lambda} = 2a\overline{a}x_\lambda - 2a_\lambda b_r + 2a_r b_\lambda \]
Thus if one seeks the stationary points of $y$ then

$$\frac{dy}{dx_r} = 0 \quad \text{and} \quad \frac{dy}{dx_i} = 0$$

or alternatively

$$\frac{dy}{dx_r} + i\frac{dy}{dx_i} = 0$$

The stationary points are given by

$$\frac{dy}{dx_r} + i\frac{dy}{dx_i} = 2a\bar{a}x_r + 2arbr + 2aib_i$$
$$\quad \quad \quad \quad \quad \quad \quad \quad + 2ia\bar{a}x_i - 2ia_ib_r + 2ia_rb_i$$
$$= 2a\bar{a}x + 2\bar{a}b$$
$$= 2\bar{a}(ax + b)$$
$$= 0$$

The result bears an obvious resemblance to the usual rules for differentiating real functions of real variables. It is in this sense that the statement: "$y = |ax + b|^2$ has a stationary point if $\frac{dy}{dx} = 2\bar{a}(ax + b) = 0$" is allowed. There is no question of a true "complex" differentiation.
CHAPTER 8

POWER SPECTRUM EQUALISATION

8.1 INTRODUCTION

This author became interested in applying non-linear signal processing techniques because of the apparently successful implementation of homomorphic restoration systems in the fields of audio signal processing and image processing. In particular Oppenheim (1968) has applied algebraically linear transformations to systems between the time, or space domain, and the frequency domain. In mathematical jargon such systems are termed homomorphic. Homomorphic representations of systems are particularly useful for describing systems which may be modelled as multiplicative or convolutional. For example the spatial convolution of two functions may be mapped into addition by taking the logarithms of the Fourier transforms of the two functions.

Oppenheim applied the technique to multiplicative filtering of audio signals in which he analysed certain audio waveforms by considering them as being composed of products of two components. Such a model would be applicable if a note from a musical instrument were
amplitude modulated, as indeed, would normally occur during a performance. He also considered two dimensional multiplicative filtering of images. The justification here is based on the fact that the observed brightness of a scene is determined by a multiplicative process between illumination and reflection of objects. Cannon (1974) produced a thesis on digital image deblurring by homomorphic filtering, and later, with his colleagues Stockham and Ingrebretson (1975) described methods of "blind" deconvolution. Blind in this context refers to deconvolution when the form of the blur is unknown. Their paper describes applications to the restoration of old voice recordings; the method was used to restore an old gramophone recording made in 1907 of Enrico Caruso singing a Verdi aria. The basic steps of their approach will be described as they will be relevant to the method used here for image restoration.

If \( s(t) \) is the singing waveform, and \( p(t) \) the playback signal then

\[
p(t) = s(t) * h(t) + e(t) \tag{8.1}
\]

where \( h(t) \) represents the impulse response of the recording system and \( e(t) \) the surface noise from the record. A solution, \( s(t) \), to equation 8.1 is required even though both \( s(t) \) and \( h(t) \) are unknown. Furthermore
the form of \( h(t) \) will vary from one gramophone recording to another. Stockham (1975) chose to split the recording, \( p(t) \), into a series of equal subsections and averaged them such that

\[
\frac{1}{Q} \sum_{\xi=1}^{Q} \log |P_\xi(u)| = \frac{1}{Q} \sum_{\xi=1}^{Q} \log |S_\xi(u)H_\xi(u) + E_\xi(u)| \tag{8.2}
\]

where \( Q \) is the number of subsections made from the recording. \( P(u) \) is the Fourier transform of \( p(t) \), and likewise for \( S(u), H(u) \) and \( E(u) \). An estimate of the average log spectrum of \( S(u) \) was made by taking the average log spectrum of a modern recording of the same aria sung by Jussi Bjoerling, the assumption being made that the modern recording is effectively distortion free. A prototype average log spectrum can be formulated such that

\[
\log |V(u)| = \frac{1}{Q} \sum_{\xi=1}^{Q} \log |S_\xi(u)| \tag{8.3}
\]

In order to estimate the log spectrum of \( RH(u) \), the homomorphic restoration filter, it is necessary to subtract from the prototype the average log spectrum of the playback signal

\[
\log |V(u)| - \frac{1}{Q} \sum_{\xi=1}^{Q} \log |P_\xi(u)| = \frac{1}{Q} \sum_{\xi=1}^{Q} \log |S_\xi(u)| - \frac{1}{Q} \sum_{\xi=1}^{Q} |S_\xi(u)H_\xi(u) + E_\xi(u)| \tag{8.4}
\]

It can be shown that equation 8.4 can be further simplified such that
\[ \log |V(\omega)| - \log |P(\omega)| = \log |S(\omega)| - \log |S(\omega).H(\omega) + E(\omega)| \quad 8.5 \]

where \( V, P, S, H \) and \( E \) now represent the results of the averaging process of equation 8.4. Hence the restoration filter can be expressed as

\[ \log |R_h(\omega)| = \log \left| \frac{V(\omega)}{P(\omega)} \right| \div \log |S(\omega)| - \log |S(\omega).H(\omega) + E(\omega)| \quad 8.6 \]

so that

\[ |R_h(\omega)| \div \frac{|S(\omega)|}{|S(\omega).H(\omega) + E(\omega)|} \quad 8.7 \]

To express \( R_H(\omega) \) in terms of power spectra the following relationships are required

\[ |S(\omega)| = (S(\omega) \cdot \overline{S(\omega)})^{1/2} \quad 8.8 \]

and

\[ \Phi_S = |S(\omega)|^2 \quad 8.9 \]

By substitution equation 8.7 becomes

\[ R_H^2(\omega) \div \left[ \frac{|S(\omega)|}{|S(\omega).H(\omega) + E(\omega)|} \right] \left[ \frac{|S(\omega)|}{|S(\omega).H(\omega) + E(\omega)|} \right] \quad 8.10 \]

Providing the noise is uncorrelated with the signal, equation 8.10 can be further simplified so that

\[ |R_h(\omega)| \div (H(\omega).\overline{H(\omega)} + \Phi_e/\Phi_S)^{-\frac{1}{2}} \quad 8.11 \]

The filter is said to be homomorphic because in its derivation the time domain convolution was mapped into addition in the frequency domain. Stockham claims that
the homomorphic restorations of old recordings are very striking. In particular, the prominent surges in volume caused when the pitch of the voice resonates with the recording horn are almost entirely gone, so that the megaphone quality is completely eliminated. The success of their venture is further substantiated by the fact that Dr. Stockham has recently set up a company to restore old recordings.

A similar two dimensional derivation can be obtained for the blind deconvolution of images. There are problems in implementing the technique in scintigraphy. The main problem lies in splitting the observed data into many sub-sections to calculate average power spectra. Cannon has experimentally found that averaging over 50 sub-sections produces successful restorations. This becomes physically impossible to do when the original data has only a 64 X 64 matrix. Furthermore, it is implicit in the averaging process that any sub-section should contain basically the same frequency components as any other. This situation rarely occurs in scintigraphy because normally the data are located within the central area of the field of view of the detector system. However, it will now be shown that a similar two dimensional result can be derived, when the impulse response can be estimated a priori, merely by considering power spectrum
equalisation.

8.2 POWER SPECTRUM EQUALISATION

Like the elegant derivation of the Wiener filter, it is possible to obtain a similar derivation for the power spectrum equalisation restoration. The stipulation to be made is simply that the power spectrum of the restored image should be approximately equal to that of the object data.

From section 4.3 it follows that

$$g_{jk} = h_{jk} * f_{jk} + e_{jk} \quad 8.12$$

and

$$G_{jm} = NH_{jm} F_{jm} + E_{jm} \quad 8.13$$

The power spectrum of $G_{jm}$ can be expressed as

$$G_{jm} \overline{G_{jm}} = (NH_{jm} F_{jm} + E_{jm})(NH_{jm} F_{jm} + E_{jm}) \quad 8.14$$

$$= N^2 H_{jm} F_{jm} F_{jm} + NH_{jm} F_{jm} E_{jm} + NH_{jm} F_{jm} E_{jm} + E_{jm} E_{jm} \quad 8.15$$

and if the correlation of signal and noise is negligible then
\[ G_{\Delta m} \tilde{G}_{\Delta m} = N^2 H_{\Delta m} \overline{H}_{\Delta m} F_{\Delta m} \overline{F}_{\Delta m} + E_{\Delta m} \overline{E}_{\Delta m} \]  

that is

\[ \Phi_g = N^2 H_{\Delta m} \overline{H}_{\Delta m} \Phi_f + \Phi_e \]  

The requirement of power equalisation is that

\[ \Phi_f = \Phi_f \]  

where \( \hat{f}_{jk} \) is the restored estimate of \( f_{jk} \). If \( R_{H_{\Delta m}} \) represents the restoring filter then

\[ \Phi_f = R_{H_{\Delta m}} \overline{R}_{H_{\Delta m}} \Phi_g \]  

It follows that

\[ \Phi_f^2 = R_{H_{\Delta m}} \overline{R}_{H_{\Delta m}} (N^2 H_{\Delta m} \overline{H}_{\Delta m} \Phi_f + \Phi_e) \]  

\[ R_{H_{\Delta m}} \overline{R}_{H_{\Delta m}} = \frac{\Phi_f}{N^2 H_{\Delta m} \overline{H}_{\Delta m} \Phi_f + \Phi_e} \]  

or

\[ |R_{H_{\Delta m}}| = \frac{1}{N} \left( H_{\Delta m} \overline{H}_{\Delta m} + \frac{\Phi_e}{N^2 \Phi_f} \right)^{-\frac{1}{2}} \]  

The similarities between the homomorphic filter of equation 8.11 and equation 8.22 are obvious. The scaling factor, \( N^2 \), in equation 8.22 occurs because of the discrete implementation of the Fourier transform.
It follows that \( \hat{F}_{\lambda m} \) is evaluated from

\[
\hat{F}_{\lambda m} = \frac{G_{\lambda m}}{N(H_{\lambda m} H_{\lambda m} + \Phi_{\lambda} / N^2 \Phi_f)^{1/2}}
\]

and the inverse Fourier transform will give a solution, \( \hat{f}_{\lambda} \).

8.3 COMPUTER IMPLEMENTATION

The power spectrum equalisation restoration requires the same functions to be evaluated as the minimum mean square error filter. For this reason the computer program used to implement the deconvolution is very similar to that already described in section 7.4. The programs only differ at step 6 (section 7.4) when the evaluation of \( \hat{F}_{\lambda m} \) is done using equation 8.23.

8.4 RESULTS WITH THE PHANTOM MODEL

Figure 8.1 shows the effect of restoring the William's model phantom using the power equalisation filter. It is less aggressive than the constrained optimisation and Wiener filters. This suggests that the filter follows more closely the response of the inverse filter with the result that higher frequencies are better preserved. The actual filter response used to produce
PROGRAM: TIMJ (REV N)  
DATE: 17-4-78 
IDENTIFICATION: MDP6 (POWER SPECTRUM EQUALISATION)  

MAX COUNT  223  
TOTAL BGD SUB  0  
TOTAL COUNT  180746  
MULTI-CYCLES  1  
THRESHOLD (X)  100  

FIGURE 8.1
figure 8.1 is shown in figure 8.2. For comparison figure 8.3 shows the restoration after increasing the noise value by a factor of four, and figure 8.4 contains the associated restoration filter. It can be readily shown that the power equalisation filter is the geometric mean of the Wiener and the inverse filter. Remembering that

\[ R_{W,bm} = \frac{H_{bm}}{N \left( H_{bm} H_{bm} + \frac{\Phi_e}{N^2 \Phi_f} \right)} \]  

then the product of \( R_{W,bm} \) with the inverse filter is

\[ \frac{H_{bm}}{N^2 H_{bm} \left( H_{bm} H_{bm} + \frac{\Phi_e}{N^2 \Phi_f} \right)} \]

\[ \Rightarrow \left| R_{H,bm} \right|^2 \]

On a log scale the power equalisation filter response will lie half way between that of the Wiener and the inverse filter. It is therefore not surprising that the resolution obtained by the power equalisation restoration is better than that obtained using the Wiener filter. However in areas where noise prevails and the impulse response frequency components are small, then

\[ \lim_{|H_{bm}| \to 0} \{ R_{H,bm} \} = \left[ \frac{\Phi_e}{\Phi_f} \right]^{1/2} \]

whereas both the Wiener and constrained optimisation filter responses approach zero. At the other extreme the inverse filter will approach infinity as \( H_{bm} \) tends to zero. At low frequencies, where noise may be considered
FIGURE 8.2 Power spectrum equalisation filter response used to obtain the result in figure 8.1

The maximum in the response has a value of 2.30 at approximately 0.4 cycles per cm.
PROGRAM: TIMJ (REV N)   DATE: 17-4-78

IDENTIFICATION: MODP7 (POWER SPECTRUM EQUALISATION, NOISE FACTOR OF 4)

MAX COUNT  225   (ROW 18, COL 37)
TOTAL BGD SUB  0
TOTAL COUNT  189923
MULTI-CYCLES  1
THRESHOLD (X)  100

FIGURE 8.3
FIGURE 8.4 Modified power spectrum equalisation filter response after the noise power has been increased by a factor of 4.

The maximum in the response has a value of 1.81 at approximately 0.4 cycles per cm.
negligible, all three filters approach the inverse filter.

8.5 SUMMARY

Chapter 8 concludes a set of four chapters which have looked at Fourier methods of deconvolution. In each method the emphasis has been on developing algorithms that work practically (ie. with real noisy data), that have built in automatic optimisation of one form or another, and which can be implemented on a mini-computer. The mathematical foundations of the techniques have been indicated and potential problems with each method, where appropriate, have been identified. All the methods have been shown to produce plausible solutions when used with the model data. The next chapter comes rather as an interlude for it introduces digital filtering, per se. The intention here is to bring deconvolution into perspective alongside "simple" digital filtering methods. The scene is then set to bring together all the separate Fourier processing methods by applications to patient data.
CHAPTER 9

DIGITAL FILTERING

9.1 INTRODUCTION

Throughout the text the terms restoring function and filter have been used synonymously. This simply arises because the implementation of all the deconvolution techniques results in a filtering operation on the spectral components of the observed data. Strictly speaking a filter will only control the amount of each frequency component propagated by it, whereas the deconvolution methods described all show marked amplification or enhancement of certain frequencies.

The preceding chapters have shown that all the Fourier deconvolution methods require a considerable amount of computer time. It would therefore be instructive to compare the deconvolved images with those obtained by standard digital filtering. In this case the computer time required will normally be significantly less than that required for deconvolution. The filter response may be shaped not so much to produce optimum image restoration for a given set of data, but rather to allow enhancement of specific characteristics of the data. In
this context filtering may be considered a subtractive process whereby certain frequency components are suppressed so that other frequencies appear enhanced because the observer is presented with less information. Of course it is perfectly legitimate to use a cascaded process of deconvolution and simple filtering should this be desired. For example selective bandpass filtering of a restored image could be used to enhance "small" lesions. The treatment here has veered towards a comparison of the two operations.

Digital filtering is not a new subject; there is a copious amount of literature available which often causes confusion to the newcomer. Rabiner and Gold (1975) give accounts of several digital filter designs and Mersereau and Dudgeon (1975) present a concise text on two dimensional digital filtering.

9.2 TYPES OF DIGITAL FILTER

Two basic types of filter are normally required for image processing; these are the lowpass and bandpass filters. The lowpass filter will ideally pass all spatial frequency components of the signal up to a predefined cut-off frequency. Similarly the bandpass filter will propagate all frequencies between a lower and an upper
frequency limit. In practical filter design the cut-off frequencies of a filter are not so clearly defined. The region of the filter response between the pass and the stopband is called the transition band.

Ideally the transition band should be as sharp as possible, so that the response might take the form of a rectangular function. However if a signal is sharply band limited then it is effectively convolved with an oscillating sinc function in the spatial domain. In order to describe fully the sinc function an infinite number of non zero Fourier coefficients would be required to represent the frequency characteristic exactly. Since this is not possible an approximation is made by using a truncated Fourier series. The effect of truncation is described by the Gibb's phenomenon which manifests itself as overshoot and ringing before and after the transition band. Conversely if a block of data is space limited by isolation, the data are convolved with an oscillating sinc function in the frequency domain.

Digital filters may be designed as recursive or non-recursive. The non-recursive filter produces an output which depends solely on a given number of input data values, whereas the recursive filter represents a summation of terms which relate to both the input and
output data sequences. In this study non-recursive finite impulse response filters have been implemented. There are two main reasons for using this approach. The first is that such filters offer zero or linear phase characteristics, resulting in minimal distortion to the signal waveshape compatible with a given gain characteristic. This is an important consideration in image processing (Andrews 1972). The second reason lies in the fact that finite impulse response filters are inherently stable. They are analogous to passive filters used in electronic signal processing. The recursive filter on the other hand may or may not have an impulse response which decays asymptotically to zero. The possibility of recursive filters being unstable may be due to poor design or to arithmetic overflow, within for example the computer floating point accumulator, due to an unbounded output sequence. An essential feature, therefore, of non-recursive filters is that the impulse response is of finite length. In effect the filter has finite "memory" so that it is bound to be stable. The main disadvantage of the non-recursive filter is that compared with recursive designs it generally requires more computer memory and computation time. However the use of the FFT to implement non-recursive filters may substantially reduce the computation required.
9.3 DESIGN OF NON-RECURSIVE FILTERS

There are three synthesis procedures which may be used to achieve a desired frequency response for a non-recursive filter. The techniques are (i) windowing, (ii) Chebyshev and Butterworth approximations and (iii) designs by successive approximation methods. The window method has been adopted here because the design is relatively straightforward and will generally produce a satisfactory frequency response.

It has already been pointed out that a filter with a rectangular frequency response has associated with it an infinitely oscillating impulse response. Basically the incorporation of a window involves lowpass filtering the filter spectral estimates, via shaping of the rectangular response, to blunt discontinuities. This is readily achieved by controlling the convergence of a Fourier series with a window weighting function to modify the Fourier coefficients. Since multiplication of Fourier coefficients by a window corresponds to convolution of the original frequency response with the Fourier transform of the window, a design criterion for windows is to find a finite window whose Fourier transform has relatively small side lobes. Some optimal designs, which have been catalogued by Rabiner and Gold, are rectangular, Blackman,
Hamming and Kaiser windows. It is not necessary here to describe the characteristics of all the non-recursive filters designed with different window shapes. This author has investigated several window functions and chosen the Hamming window for incorporation into all the filters used in this study.

The discrete one dimensional Hamming window function is defined as

\[ H_n = 0.54 + 0.46 \cos\left(\frac{2\pi n}{W}\right), \quad -\frac{W}{2} \leq n \leq \frac{W}{2} \]

where \( W \) represents the window width. The function, \( H_n \), takes the form of a cosine bell. The Hamming window filter offers intermediate performance. It does not offer the short transition bandwidth of the rectangular window, but it does offer relatively good stopband attenuation of greater than 50dB. Conversely the Blackman and Kaiser windows offer very good stopband attenuation (typically greater than 70dB suppression) at the expense of wider transition bands. The Hamming window filter therefore offers a frequency response which is fairly flat in the pass and stopbands, reasonable transition bandwidths and attenuation in the stopband. One further advantage is that the window function is mathematically simple and therefore easily programmed on a digital computer.
9.4 REALISATION OF TWO DIMENSIONAL NON-RECURSIVE FILTERS

The implementation of a non-recursive filter by the window method can be broken down into several steps. A diagrammatic representation of this process for a one dimensional case is shown in figure 9.1. The design procedure starts with the ideal frequency response (figure 9.1A), which for the time being is assumed to be known analytically. This frequency response can be directly transformed to the impulse response using the inverse FFT (figure 9.1B). The coefficients for the impulse response consist of a complex array of samples. The defined smoothing window (figure 9.1C) is multiplied point by point with the filter samples (figure 9.1D). By taking the FFT of the resultant complex data an approximated frequency response is produced (figure 9.1E). If the gain characteristic does not represent a close enough approximation to the desired response, the whole procedure can be repeated with different window function parameters. For example, for a constant filter bandwidth the effect of window width will be to alter the transition band in the filter frequency response; the wider the window (which allows more terms in the impulse response to be utilised) the shorter the transition band. In the filters used here the transition band typically falls over three samples; this offers a reasonably short transition band without
IDEAL FREQUENCY RESPONSE

FIGURE 9.1A
FIGURE 9.1B

IDEAL IMPULSE RESPONSE
HAMMING WINDOW FUNCTION

FIGURE 9.1C
MODIFIED IMPULSE RESPONSE

FIGURE 9.1D
MODIFIED FREQUENCY RESPONSE

FIGURE 9.1E
excessive ripple.

One last step is required to convert the filter response into two dimensions suitable for image processing. Several approaches have been investigated. One that gives a circularly symmetrical frequency response can be obtained from the one dimensional sequence. If $F_{1q}$ represents the one dimensional filter containing $N/2$ values (only $N/2$ points are required because the filter response is symmetrical about $N/2$), then a two dimensional sequence, $F_{2lm}$, may be generated such that

$$F_{2lm} = F_{1q}$$

where

$$q = (l^2 + m^2)^{1/2} \quad \text{for} \quad 0 \leq l, m \leq N/2 - 1$$

$$q = ((N-l)^2 + m^2)^{1/2} \quad \text{for} \quad \begin{cases} \frac{N}{2} \leq l \leq N-1 \\ 0 \leq m \leq N/2 \end{cases}$$

$$q = (l^2 + (N-m)^2)^{1/2} \quad \text{for} \quad \begin{cases} 0 \leq l \leq N/2 \\ \frac{N}{2} \leq m \leq N-1 \end{cases}$$

$$q = ((N-l)^2 + (N-m)^2)^{1/2} \quad \text{for} \quad \frac{N}{2} \leq l, m \leq N-1$$

and finally if $q$ is greater than $(N/2 - 1)$, it is set equal to $(N/2 - 1)$. If necessary $q$ is truncated to an integer value and will always lie in the range 0 to $(N/2 - 1)$. The constraint that $q$ is less than or equal to
(N/2 - 1) is necessary for valid referencing of the filter coefficients \( F_{lq} \). The consequence of the constraint is insignificant for it only modifies the extreme high frequency diagonal components of the two dimensional filter response where for all practical purposes the data being filtered contain no frequency components. The main advantage of this filter design is one of computational convenience. The method does not require the evaluation of a two dimensional convolution to produce the windowed filter response, so it is rapidly evaluated within a generalised image filtering program. This method of filter synthesis has worked very effectively both for lowpass and bandpass designs.

9.5 A METHOD FOR OPTIMISING THE LOWPASS FILTER BANDWIDTH

Although the filter synthesis method described can be used very efficiently to construct a particular filter with defined cut-off frequencies, so far no criterion of fixing these frequencies has been given.

In the case of lowpass filtering only one upper cut-off frequency has to be chosen for a particular image being filtered. Usually the object of lowpass filtering is to remove high frequency noise without unnecessarily removing valuable signal information from the picture.
One possible approach is to evaluate the FFT of the observed data and from this compute both the power spectrum and the expected noise value using the method described in section 7.3, and then define the cut-off frequency as that frequency at which the power spectral components are approximately equal to the noise value. This is only a rough guide for a possible cut-off frequency because usually power spectra are not circularly symmetric.

An alternative approach, which has been adopted in this study, is to compute the variance of the observed data and equate this figure with the area under the high frequency tail of the power spectrum of the data. A cut-off frequency is chosen such that the area under the tail of the power spectrum from this cut-off frequency to the high frequency limit is approximately equal to the variance. By using Parsavel's relationship the variance of the data can be carried through to the Fourier domain, and the assumption is then made that those frequency components which contain the lower signal to noise ratios and largest statistical uncertainty will be associated with the high frequency end of the spectrum. The result is that noisy scintigrams will be more heavily smoothed than higher quality, statistically more accurate ones. Generally the lowpass filtering technique produces
superior results to those obtained, for example, from the
nine point smoothing methods used on many commercially
available nuclear medicine data processing systems. The
ability of this method to cope with varying degrees of
noise is demonstrated in figures 9.2 to 9.6. Figure 9.2
shows the Williams model (figure 4.13) after being
processed by the lowpass filtering program. The program
produced a cut-off frequency of 0.43 cycles per cm. based
upon the criterion already described, and the resultant
windowed frequency response is shown in figure 9.3.
Figure 9.4 is a nine point smoothed version of the same
data. The latter image is significantly more blotchy than
that of figure 9.2, and also distorts the symmetrical
shapes of the "lesions" to a greater extent. In other
words the optimised lowpass filter produces an image which
is a more faithful representation of the original data
while introducing a moderate degree of smoothing. To
demonstrate the effect of reducing the statistical quality
of the data on the filtering process an alternative
Williams model was generated by reducing the counts per
pixel by a factor of ten and then subjecting this data to
the Poisson noise generator already described in section
4.8. The resultant model data is shown in figure 9.5.
Using this data the lowpass filter program produced a
cut-off frequency of 0.35 cycles per cm. The program has
detected the poorer signal to noise ratio of the data and
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: MODP18 (OPTIMISED LOWPASS FILTER)

MAX COUNT  213  (ROW 17  COL 37)  
TOTAL BGD SUB  0  
TOTAL COUNT  178558  
MULTI-CYCLES  1  
THRESHOLD (x)  100

FIGURE 9.2
FIGURE 9.3 Lowpass filter response used to obtain the result in figure 9.2.

The transition frequency is at 0.43 cycles per cm.
FIGURE 9.4
PROGRAM: TIMJ (REV N)  DATE: 17- 4-78
IDENTIFICATION: MODP15  (OPTIMISED LOWPASS FILTER)

MAX COUNT  23  (ROW 17  COL 38)
TOTAL BGD SUB  0
TOTAL COUNT  17912
MULTI-CYCLES  1
THRESHOLD (%)  100

FIGURE 9.6
reduced the filter bandwidth accordingly. The result of lowpass filtering is shown in figure 9.6. By contrast a nine point smoothed version is shown in figure 9.7. In the author's opinion figure 9.6 is to be preferred to figure 9.7. Close inspection of the two images shows that both have comparable resolution but the gross blotchiness in figure 9.7 of the uniform activity areas is undesirable, and indeed may be distinctly misleading - a situation which may lead to a false positive interpretation in an equivalent clinical scintigram.

A relatively simple method of choosing the upper cut-off frequency of a filter frequency response has been described. It seems a better approach to take than nine point smoothing where no consideration is given to the statistical quality of the data. The lowpass filter is adaptive, and as such, the computer program needs no intervention on the part of the operator to choose any filter parameter.

9.6 BANDPASS FILTERING

In some circumstances it may be considered an advantage to bandpass filter an image. Usually the observer is looking for small lesions which may be masked by relatively high background activity throughout the
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: MODCAN AFTER NINE POINT SMOOTH  

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX COUNT</td>
<td>23</td>
</tr>
<tr>
<td>ROW</td>
<td>15</td>
</tr>
<tr>
<td>COL</td>
<td>37</td>
</tr>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>16143</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
</tr>
<tr>
<td>THRESHOLD (x)</td>
<td>100</td>
</tr>
</tbody>
</table>
organ being imaged. A typical example of a bandpass filter is the unsharp mask (Schreiber 1970), which in scintigraphy has been termed Canterbury filtering (Corfield 1975). The Canterbury filter is frequently used to enhance suspected brain abnormalities.

Once it has been decided to lowpass filter an image it is a trivial step to convert the computed response into a bandpass one. Basically, all that is required is that the window function should modify both ends of the ideal rectangular frequency response. However, unlike the lowpass response, it is more difficult to choose automatically the low frequency cut-off of the bandpass filter. This stems principally from the fact that the cut-off is inversely related to the size of the lesion that one wishes to visualise. The bandpass filter therefore has to be used by the operator with discretion, and even then, the results of filtering will normally be compared along side the unprocessed data. For completeness the effects of different bandpass filters on the Williams model are shown in figures 9.8 and 9.9. As one would expect the results show that the more uniform activity areas of the image have been suppressed and those areas containing higher spatial frequencies appear enhanced. The upper cut-off frequency in each case is the same as that of figure 9.3. A typical windowed bandpass
PROGRAM: TIMJ (REV N)  DATE: 17- 4-78

IDENTIFICATION: MODP11 (LOW CUT-OFF 0.08; HIGH CUT-OFF 0.43 CYCLES/CM)

<table>
<thead>
<tr>
<th>MAX COUNT</th>
<th></th>
<th>TOTAL BGD SUB</th>
<th></th>
<th>TOTAL COUNT</th>
<th>112527</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUB TOTAL</td>
<td>COUNT</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>THRESHOLD (%)</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 9.8
IDENTIFICATION: MODP14 (LOW CUT-OFF 0.20; HIGH CUT-OFF 0.43 CYCLES/CM)

MAX COUNT 62 (ROW 30 COL 37)
TOTAL BGD SUB 0
TOTAL COUNT 21284
MULTI-CYCLES 1
THRESHOLD (X) 100
frequency response is shown in figure 9.10.

9.7 SUMMARY

Fourier methods of implementing two dimensional non-recursive filters have been demonstrated. A relatively simple method of choosing the lowpass filter bandwidth has been described. The question arises whether deconvolution of scintigraphic images offers any advantage over that of lowpass filtering. The question may be answered in the affirmative. With the model data the true distribution of the data is known. It would therefore seem appropriate to compare each processing method by a simple criterion of looking at which images have the sharpest outlines at the known locations of each "lesion" consistent with a faithful representation of the uniform activity areas of the model. With the Williams model data the deconvolution techniques generally do offer some improvement on image sharpness. The results of any deconvolution method show that "lesion" boundaries contain less activity levels on the grey scale display, implying a greater rate of change of activity per pixel at the "lesion" boundary. Also the grey scale level in the centre of each "lesion" is more consistent with the expected grey level. For example, the second largest hot area usually reaches the darkest grey tone on the
FIGURE 9.10 Bandpass filter response used to obtain the result in figure 9.9.

The transition frequencies are at 0.20 and 0.43 cycles per cm.
deconvolved data (for an example see figure 5.3) which does not happen for the lowpass filtered data (figure 9.2) or the original data (figure 4.13). All the data processing methods demonstrate some smoothing out of the statistical fluctuations present in the unprocessed data.

The model data of course do not possess any artefacts other than Poisson noise. In a clinical situation other forms of noise may be introduced which may impede the resolution enhancement hoped for from the deconvolution process. The next chapter shows some of the results of deconvolving clinical data and attempts to analyse various problems associated with refocussing real data.
CHAPTER 10

CLINICAL EXAMPLES

10.1 INTRODUCTION

This chapter presents some preliminary results of deconvolving clinical data. For each example the results of deconvolution by the four optimised deconvolution processes are available for comparison. The clinical studies are broadly divided between data originating from morphic and amorphous organs. Generally, deconvolution of data from morphic organs is relatively straightforward, whereas deconvolution of data originating from amorphous structures may present much more difficult and often complex problems. The chapter concludes by identifying some of the problems the author experienced in deconvolving certain types of scintigram.

10.2 PROCESSING DATA FROM MORPHIC STRUCTURES

The term "morphic structure" is used to describe a part of the body which has a well defined form. Examples of morphic structures are the skeleton and the brain. On a macroscopic scale the structures may be considered to have a constant shape which is independent of the effects
of pulmonary and cardiac activity. From the point of view of scintigraphy the photons may be considered to originate from a volume of constant shape during the data collection period. An example of an amorphous organ is the liver whose shape is continually modulated by the diaphragm due to the effects of respiration. It will be shown in section 10.4.1 that deconvolution of scintigrams of the liver can produce unpredictable results.

A brain and a bone scan will be used to demonstrate the effects of processing scintigraphic data from morphic structures. Figure 10.1 is a technetium pertechnetate scan of a left lateral brain. An experienced radiologist would immediately interpret the scan as abnormal. A lesion can be seen near the centre of the image. The other patchy areas of high activity in the lower part of the picture are probably due to the action of the salivary, lachrymal and parotid glands and are quite normal. The results of deconvolving the brain scan by constrained optimisation, binomial series expansion, Wiener filtering and power spectrum equalisation are shown in figures 10.2 to 10.5 respectively. All the processed data clearly show a well of activity in the base and top of the mouth. The lesion is also more apparent. Comparison with the unprocessed data demonstrates that deconvolution localises activity very effectively. The
PROGRAM: TIMJ (REV N)       DATE: 10- 4-76

IDENTIFICATION: ROUR4

MAX COUNT  600  (ROW 14  COL 26)
TOTAL BGD SUB  0
TOTAL COUNT  336328
MULTI-CYCLES  1
THRESHOLD (%)  100

FIGURE 10.1
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: ROURPI  (CONSTRAINED OPTIMISATION)  

DATE: 18-4-78

MAX COUNT 918  (ROW 14  COL 24)
TOTAL BGD SUB 0
TOTAL COUNT 343496
MULTI-CYCLES 1
THRESHOLD (x) 70

FIGURE 10.2
PROGRAM: TIMJ (REV N)  DATE: 18-4-78
IDENTIFICATION: ROURP2  (BINOMIAL SERIES EXPANSION)

MAX COUNT 923  (ROW 14  COL 24)
TOTAL BGD SUB 0
TOTAL COUNT 343548
MULTI-CYCLES 1
THRESHOLD (x) 70

FIGURE 10.3
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: ROJRP3  (WIENER FILTER)  

MAX COUNT  800  (ROW 14  COL 25)  
TOTAL BGD SUB  0  
TOTAL COUNT  342396  
MULTI-CYCLES  1  
THRESHOLD (%)  70  

FIGURE 10.4
<table>
<thead>
<tr>
<th>Max Count</th>
<th>850</th>
<th>(ROW 14 COL 24)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total BGD Sub</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Total Count</td>
<td>343819</td>
<td></td>
</tr>
<tr>
<td>Multi-Cycles</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Threshold (x)</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 10.5**
display threshold on the processed data has been reduced to 70% so that texture in the lower activity areas can also be discerned. It is interesting to note that all the deconvolution techniques produce very similar results. In particular figures 10.2, 10.3 and 10.4 show the lesion in the left parietal region with decreased activity in the centre. The halo effect is consistent with a metastatic tumour with a necrotic centre. This particular patient was diagnosed as having secondaries following bronchogenic carcinoma.

For comparison purposes the unprocessed data from the brain scan have also been bandpass filtered (figure 10.6). The optimised filter program produced an upper cut-off frequency of 0.66 cycles per cm., and a low frequency cut-off of 0.08 cycles per cm. was chosen to emphasise high activity areas. The data are processed in a similar way to that produced by the unsharp mask (Corfield 1975) which is frequently used for processing brain scintigrams (Bits and Pixides 1975). The reader should note that high activity areas in particular are not resolved as well as can be achieved with the deconvolution methods for data displayed using the same threshold value. This is not unexpected because the bandpass filter only attempts to remove low frequency detail and control high frequency noise. There is no refocussing process.
PROGRAM: TIMJ (REV N)  DATE: 18-4-78

IDENTIFICATION: ROURP5 (BANDPASS FILTER)

MAX COUNT  541 (ROW 14  COL 24)
TOTAL BGD SUB  0
TOTAL COUNT  211993
MULTI-CYCLES  1
THRESHOLD (X)  70

FIGURE 10.6
An abnormal bone scan is shown in figure 10.7. This is a picture of a left lateral head scan which was obtained after injection of technetium polyphosphate. This scan has been chosen for demonstration purposes because the extreme metastatic disease in the cranium, lower mandible and cervical regions is well localised so that the data should contain a relatively large proportion of high frequency "energy". This dissertation has already shown that mathematical instability during deconvolution is associated with the high frequency end of the power spectrum. The results of deconvolution are presented in figures 10.8 to 10.11. All the deconvolution methods give results which show significant enhancement of the metastatic areas.

The maxima in the processed data are typically a factor of two greater than those contained in the unprocessed data. The implication of this is that after deconvolution, counts within the unprocessed metastatic regions are contained within smaller areas. This is of course consistent with the refocussing of the data.

Careful study of figures 10.8 to 10.11 shows that all the deconvolved data exhibit varying degrees of ringing. The ringing becomes more apparent near lesion boundaries
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX COUNT</td>
<td>958</td>
</tr>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>233850</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
</tr>
<tr>
<td>THRESHOLD (X)</td>
<td>100</td>
</tr>
</tbody>
</table>

**FIGURE 10.7**
<table>
<thead>
<tr>
<th><strong>PROGRAM</strong>: TIMJ (REV N)</th>
<th><strong>DATE</strong>: 19-4-78</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IDENTIFICATION</strong>: BO6P1 (CONSTRAINED OPTIMISATION)</td>
<td></td>
</tr>
<tr>
<td><strong>MAX COUNT</strong></td>
<td>2440</td>
</tr>
<tr>
<td><strong>TOTAL BGD SUB</strong></td>
<td>0</td>
</tr>
<tr>
<td><strong>TOTAL COUNT</strong></td>
<td>276282</td>
</tr>
<tr>
<td><strong>MULTI-CYCLES</strong></td>
<td>1</td>
</tr>
<tr>
<td><strong>THRESHOLD (%)</strong></td>
<td>50</td>
</tr>
</tbody>
</table>

**FIGURE 10.8**
PROGRAM: TIMJ (REV N)  DATE: 19- 4-78
IDENTIFICATION: BO6P2  (BINOMIAL SERIES EXPANSION)

MAX COUNT 1884 (ROW 49  COL 29)
TOTAL BGD SUB 0
TOTAL COUNT 248272
MULTI-CYCLES 1
THRESHOLD (X) 50

FIGURE 10.9
PROGRAM: TIMJ (REV N)  DATE: 19-4-78

IDENTIFICATION: 806P3 (WIENER FILTER)

MAX COUNT  1565 (ROW 49  COL 29)
TOTAL BGD SUB   0
TOTAL COUNT    247264
MULTI-CYCLES   1
THRESHOLD (X)  50

FIGURE 10.10
PROGRAM: TIMJ (REV N)   DATE: 19-4-78
IDENTIFICATION: B06P4   (POWER SPECTRUM EQUALISATION)

MAX COUNT: 1928 (ROW 49, COL 29)
TOTAL BGD SUB: 0
TOTAL COUNT: 248796
MULTI-CYCLES: 1
THRESHOLD (%): 50

FIGURE 10.11
because the data contain large high frequency components in these areas. In this particular example constrained optimisation demonstrates ringing rather more than the other methods. This is a consequence of the low value of 0.2441E-05 for \( \gamma \). It should be noted that ringing can produce negative pixel values around the high count regions in the data. Although negative data are non-physical they do not particularly distort the appearance of the displayed data because these pixels take values which are only a few percent of the maximum pixel count. For display purposes the images presented here have negative values included within the lowest activity level. Although the processed data contain artefacts it is encouraging to see how well the deconvolution processes cope with data that is inherently difficult to restore.

The brain and bone scans demonstrate that deconvolution may help to localise clinical abnormalities and give more fine structure to the data. It can be argued that restoration could also make discernible lesions within a scintigram that would otherwise be judged as normal.

10.3 PROCESSING DATA FROM AMORPHOUS ORGANS
Clinical interpretations of scans of amorphous organs, such as the liver, generally require considerable radiological expertise. The variability of liver shape and size in physiologically normal individuals produces a class of scintigrams which covers a wide variety of image structure. Furthermore the shape of the liver may not be constant during the scanning procedure. This variability of shape is usually due to the effects of respiration in which motion of the diaphragm effects the relative positioning of the liver within the abdomen. Obviously the deconvolution methods described here can only attempt to refocus data originating from a stationary source. In the author's opinion one of the main problems in scintigraphy is the fact that in many in vivo procedures the object distribution itself is not stationary. Although deconvolution of liver scans will be demonstrated the techniques should be used with caution.

Figure 10.12 is a gamma camera picture of a clinically normal anterior liver scintigram obtained using technetium labelled sulphur colloid. The results of deconvolution are shown in figures 10.13 to 10.16. Figures 10.13 to 10.15 are similar in that the data appear as less noisy versions of the original images. The power spectrum equalisation method produces superior high frequency enhancement which can be seen in figure 10.16.
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: WISH01

MAX COUNT 716  (ROW 32  COL 42)
TOTAL BGD SUB 0
TOTAL COUNT 335192
MULTI-CYCLES 1
THRESHOLD (X) 100

FIGURE 10.12
PROGRAM: TIMJ (REV N)

IDENTIFICATION: WISHP1 (CONSTRAINED OPTIMISATION)

MAX COUNT 723 (ROW 33 COL 41)
TOTAL BGD SUB 0
TOTAL COUNT 336452
MULTI-CYCLES 1
THRESHOLD (%) 100

FIGURE 10.13
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX COUNT</td>
<td>717</td>
</tr>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>336184</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
</tr>
<tr>
<td>THRESHOLD (x)</td>
<td>100</td>
</tr>
</tbody>
</table>

FIGURE 10.14
**PROGRAM:** TIMJ (REV N)  
**DATE:** 19- 4-78  
**IDENTIFICATION:** WISHP3 (WIENER FILTER)

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX COUNT</td>
<td>723</td>
<td>(ROW 33 COL 41)</td>
</tr>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>337189</td>
<td></td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>THRESHOLD (X)</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 10.15**
PROGRAM: TIMJ (REV N)  DATE: 19-4-78
IDENTIFICATION: WISHP4  (POWER SPECTRUM EQUALISATION)

MAX COUNT  744  (ROW 32  COL 41)
TOTAL BGD SUB  0
TOTAL COUNT  338221
MULTI-CYCLES  1
THRESHOLD (X)  100

FIGURE 10.16
A significant amount of extra detail is apparent in the central area of the right lobe of the liver; the crescent shapes that are resolved in the processed data are only hinted at in the original scan. For completeness nine point smoothed and lowpass filtered versions of the data appear in figures 10.17 and 10.18. The optimised lowpass filter cut-off at 0.51 cycles per cm; the data are not filtered as heavily as by the nine point smooth. For this liver scan the nine point smooth produces a result with less detail than for any of the other processing methods.

As a contrast figure 10.19 presents the unprocessed data from an abnormal anterior liver scan recorded after injection of technetium labelled sulphur colloid. The patchiness of the right lobe and poor uptake in the left lobe are indicative of multiple metastases. The deconvolved data are shown in figures 10.20 to 10.23. It is interesting to note that a considerable amount of structure is contained within the processed images while a reasonable amount of smoothing has been obtained. The power spectrum equalisation restoration offers the most high frequency emphasis. This is expected because the restoration filter follows the inverse response to a closer degree than can be obtained using the other deconvolution techniques. Of course the data appear slightly more noisy than the other processed results, but
IDENTIFICATION: WISH01 AFTER NINE POINT SMOOTH

<table>
<thead>
<tr>
<th>MAX COUNT</th>
<th>689 (ROW 34 COL 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL BSD SUB</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>333772</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
</tr>
<tr>
<td>THRESHOLD (x)</td>
<td>100</td>
</tr>
</tbody>
</table>

FIGURE 10.17
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: WISHP5 (OPTIMISED LOWPASS FILTER)  

MAX COUNT   690 (ROW 29 COL 43)  
TOTAL BGD SUB 0  
TOTAL COUNT 335338  
MULTI-CYCLES 1  
THRESHOLD (%) 100  

FIGURE 10.18
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: S602  

<table>
<thead>
<tr>
<th>MAX COUNT</th>
<th>240</th>
<th>(ROW 20 COL 48)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>195004</td>
<td></td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>THRESHOLD (T)</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 10.19
PROGRAM: TIMJ (REV N)  DATE: 19- 4-78
IDENTIFICATION: S682P1 (CONSTRAINED OPTIMISATION)

MAX COUNT: 232 (ROW 24 COL 46)
TOTAL BGD SUB: 0
TOTAL COUNT: 197213
MULTI-CYCLES: 1
THRESHOLD (X): 100

FIGURE 10.20
MAX COUNT 230 (ROW 24 COL 46)
TOTAL BGD SUB 0
TOTAL COUNT 196956
MULTI-CYCLES 1
THRESHOLD (x) 100

FIGURE 10.21
PROGRAM: TIMJ (REV N)  DATE: 19- 4-78
IDENTIFICATION: S682P3 (WIENER FILTER)

MAX COUNT: 224  (ROW 24  COL 45)
TOTAL BGD SUB: 0
TOTAL COUNT: 197330
MULTI-CYCLES: 1
THRESHOLD (X): 100

FIGURE 10.22
**PROGRAM: TIMJ (REV N)**

**DATE:** 19-4-78

**IDENTIFICATION:** S682P2 (POWER SPECTRUM EQUALISATION)

<table>
<thead>
<tr>
<th>MAX COUNT</th>
<th>240 (ROW 39 COL 47)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>198816</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
</tr>
<tr>
<td>THRESHOLD (%)</td>
<td>100</td>
</tr>
</tbody>
</table>

**FIGURE 10.23**
still less noisy than the unprocessed data. The optimised lowpass filtered version of the data (figure 10.24) contains the least fine structure.

10.4 PROBLEMS OF DECONVOLVING CERTAIN TYPES OF SCINTIGRAM

So far it would appear that there are no great problems in deconvolving patient data. Indeed if the mathematical theory is thorough and accommodates all the characteristics of genuine data then there should be no computational problems. However to the author's consternation results were sometimes unpredictable. For example, processed data sometimes contained parallel striations across the image. Eventually these problems were shown to be associated with artefacts in the raw data which may be physiological or instrumental in origin.

10.4.1 PHYSIOLOGICAL ARTEFACTS

Physiological artefacts are only apparent when data are collected from the rectilinear scanner, and are generally due to organ motion occurring as the detector traverses the patient. An example of this effect is shown in figure 10.25 which is an anterior scan of a liver containing technetium labelled sulphur colloid. The horizontal striations are due to the effects of
PROGRAM: TIMJ (REV N)  DATE: 19-4-78
IDENTIFICATION: S682P5 (OPTIMISED LOWPASS FILTER, BANDWIDTH 0.6 CYCLES/CM)

MAX COUNT  218  (ROW  23  COL  46)
TOTAL BGD SUB  0
TOTAL COUNT  196438
MULTI-CYCLES  1
THRESHOLD (X)  100

FIGURE 10.24
PROGRAM: TIMJ (REV N)  DATE: 19-4-78
IDENTIFICATION: S341

MAX COUNT 300 (ROW 37 COL 53)
TOTAL BGD SUB 0
TOTAL COUNT 221841
MULTI-CYCLES 1
THRESHOLD (X) 100

FIGURE 10.25
respiration in which the motion of the diaphragm continually moved the relative position of the liver with respect to the scanner frame. The effect is more marked at the liver boundary because here the organ may leave the field of view of the detector collimator. Consequently the data appear to have discontinuities which are perpendicular to the scanning direction because the patient's respiration cycle is not synchronised to the scanning period of the data collection system.

Obviously these striations, or artefacts, are not related to liver function. If the Fourier transform of the observed data is computed the resultant complex data will contain frequency components which originate from these discontinuities. The effect will be a manifestation of erroneous frequency components throughout the Fourier domain. It is also possible that these observed frequency components will have been aliased because there is no way of determining whether discontinuities occurred at a spatial frequency that was outside the maximum frequency propagated by the detector system. The frequency response may therefore be distorted on two accounts, the first being due to physiological artefacts and the second to aliasing of the data. It will be appreciated that all of the deconvolution methods rely heavily on the data, \( G_{lm} \), the frequency components of the observed data. Generally
$G_{km}$ will be over estimated or alternatively, the signal to noise ratio will be over estimated particularly at the high frequency end of the spectrum. Unfortunately this has the effect of relaxing the mathematical constraints in the deconvolution processes so that some instability may occur and the result may show "enhancement" of the striations. Figure 10.26 demonstrates this effect. The data in figure 10.25 have been deconvolved using the constrained optimisation technique. The striations in the deconvolved data may be consistent with the observed data in that the results show sharper discontinuities than those seen in the original data. The author admits that the explanation for this mathematical instability is intuitive rather than theoretical. However a further test which substantiates the above argument has been carried out by the following method. The data in figure 10.25 were nine point smoothed using a standard 4,2,1 weighting matrix so that the image in figure 10.27 resulted. This process helps to smooth out discontinuities in the unprocessed data. The data in figure 10.27 were then used as input to the constrained optimisation program. The result is shown in figure 10.28. The data now show none of the instabilities of figure 10.26. The result is meaningless as an interpretable image but it does clearly demonstrate that the mathematical instabilities are associated with the high frequency discontinuities in the
PROGRAM: TIMJ (REV N)  DATE: 19-4-78
IDENTIFICATION: S341P1 (CONSTRAINED OPTIMISATION, GAMMA=0.98E-05)

MAX COUNT  389  (ROW 34 COL 53)
TOTAL BGD SUB  0
TOTAL COUNT  232133
MULTI-CYCLES  1
THRESHOLD (X)  100

FIGURE 10.26
PROGRAM: TIMJ (REV N)  
IDENTIFICATION: S341 AFTER NINE POINT SMOOTH  

MAX COUNT  280  (ROW  29  COL  10)  
TOTAL BGD SUB  0  
TOTAL COUNT  219109  
MULTI-CYCLES  1  
THRESHOLD (x)  100  

FIGURE 10.27
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>MAX COUNT</td>
<td>273</td>
</tr>
<tr>
<td>TOTAL BGD SUB</td>
<td>0</td>
</tr>
<tr>
<td>TOTAL COUNT</td>
<td>220801</td>
</tr>
<tr>
<td>MULTI-CYCLES</td>
<td>1</td>
</tr>
<tr>
<td>THRESHOLD (x)</td>
<td>100</td>
</tr>
</tbody>
</table>

**FIGURE 10.28**
170

object data.

10.4.2 INSTRUMENTAL ARTEFACTS

A similar sort of high frequency distortion can arise for instrumental reasons. Figure 10.29 shows a normal anterior technetium pertechnetate brain scan recorded using the J and P Engineering scanner. The relatively high uptake areas at the bottom of the scintigram are due to salivary and lachrymal activity. The radiographer terminated the scan once the detector was beyond the base of the cortex so that an abrupt discontinuity in count rate is apparent on the displayed scintigram. In mathematical terms the data exhibit a discontinuity in the y direction so that a frequency analysis of the data will show relatively large components in this direction. Figure 10.30 shows the power spectrum of the brain scan; a ridge running in the y direction is clearly visible. The results of deconvolution using constrained optimisation are shown in figure 10.31. The relatively large amounts of high frequency energy cause relaxation in the constraint producing a low value of \( \gamma \) of only 0.7629E-07. The result is amplification of noise and ringing in the deconvolved data particularly at the position of the discontinuity in the object data.
PROGRAM: TIMJ (REV N)  DATE: 19-4-78

IDENTIFICATION: S329

MAX COUNT  236  (ROW 36  COL 38)
TOTAL BGD SUB  0
TOTAL COUNT  140560
MULTI-CYCLES  1
THRESHOLD (X)  100

FIGURE 10.29
FIGURE 10.30  Power spectrum of the data shown in figure 10.29
PROGRAM: TIMJ (REV N)  DATE: 19-4-78
IDENTIFICATION: S329P2 (CONSTRAINED OPTIMISATION, GAMMA=0.7629E-07)

MAX COUNT 1668 (ROW 29 COL 29)
TOTAL BGD SUB 0
TOTAL COUNT 471835
MULTI-CYCLES 1
THRESHOLD (%) 100

FIGURE 10.31
One method of showing that the mathematical instability in the deconvolution is due to the presence of large high frequency components is achieved by artificially removing the discontinuity. In the case of the brain scan the simplest way of doing this is to reflect the data about the centre row in the image. Figure 10.32 shows the effect of this process and figure 10.33 shows the power spectrum of the modified data. Because the discontinuity has been removed the ridge in the power spectrum has been suppressed. Constrained optimisation is now well behaved and produces the result shown in figure 10.34. The data in the lower half of the image have been set to zero prior to display. The value of $\gamma$ was $0.7812E-02$.

It is important to appreciate that both physiological and instrumental artefacts would not have been observed if the data had been collected using a gamma camera. These artefacts are not seen on the gamma camera display because the whole field of view of the organs being visualised is simultaneously exposed. However it is futile to try and deconvolve data that is from a non stationary source. The situation is analogous to restoring photographic images when the subject in the picture is known to have been moving during the exposure. With the ever increasing use of gamma cameras and the less frequent use of scanners for
PROGRAM: TIMJ (REV N)                                       DATE: 19- 4-7B
IDENTIFICATION: S329P3

MAX COUNT 108 (ROW 33 COL 49)
TOTAL BGD SUB 0
TOTAL COUNT 227,143
MULTI-CYCLES 1
THRESHOLD (X) 100

FIGURE 10.32
FIGURE 10.33 Power spectrum of the data shown in figure 10.32
```markdown
| MAX COUNT | 185  | (ROW 32 COL 49) |
| TOTAL BGD SUB | 0    |
| TOTAL COUNT | 1126B4 |
| MULTI-CYCLES | 1    |
| THRESHOLD (%) | 100  |

FIGURE 10.34
```
routine imaging it is not surprising that little attention has been paid to whether or not the data are stationary. All the workers cited in this thesis who have been involved in medical scintigraphic deconvolution do not consider such problems, presumably because they have not encountered them, or because the problems are camouflaged by the nature of the gamma camera data collection method.
CHAPTER 11

SUMMARIES

11.1 HARDWARE

The adaptation of a general purpose digital computer as a nuclear medicine data processing system has been described. In view of the fact that the resolution of isotope imaging equipment has improved significantly in recent years it would be preferable to use a 16 bit word computer. This would facilitate the collection of data up to a 256*256 format. The most demanding piece of electronic equipment constructed was the television display which was built at a cost of less than one tenth of commercially available systems. When the project started there were no commercially available systems and to the author's knowledge there is no television display marketed for the PDP 8 and PDP 12 range of computers. The success of the prototype television display is substantiated by the number of people within Edinburgh who have frequently used the system. A second display using a finer matrix structure, real time interpolation and including an alpha-numeric character generator is currently being developed.
The electrostatic plotter has also proved popular among many workers. The algorithms for the various grey tone methods are easily understood, and the results are believed to be amongst some of the best produced by electrostatic plotters. The ease with which the television and electrostatic plotter may be used in a Fortran based system has proved popular with inexperienced programmers.

11.2 SOFTWARE

Fourier deconvolution methods have been developed which have been tested on model and clinical data. Constrained optimisation should be the method of choice because it employs sound mathematical principles which satisfy the statistical properties of photon data. The binomial series expansion generally gives similar results to constrained optimisation. The main drawback of the binomial series expansion is that convergence of equation 6.24, for estimating n, may be slow. For reasons outlined in section 7.2 true Wiener filtering is fraught with mathematical assumptions that are difficult to justify for Poisson distributed data. Similar assumptions are associated with the power spectrum equalisation method, although this restoration filter does offer a closer approximation to the inverse filter. When the impulse
response is known the power spectrum equalisation restoration has been shown to be similar to the homomorphically derived restoration filter. Other workers have suggested that homomorphic deconvolution systems have produced some of the best results yet at restoring blurred photographic images. Admittedly in most cases the quality of restoration has only been assessed subjectively, but the general consensus of opinion appears to be that most observers prefer to look at slightly noisy, but better focussed, images than smoother blurred images. Indeed the aim of deconvolution is to refocus data not to smooth it; in other words it is not a prerequisite of restoration that the solution should necessarily look less noisy. This is a nicety, not a fundamental requirement. Of course it is debatable whether similar conclusions would be drawn from observers looking at restored scintigraphic images.

For the model data and for data from morphic structures the results of all the deconvolution methods are encouraging. The localisation of activity in abnormal brain and bone scans is particularly gratifying. Intuitively it would seem that if deconvolution can refocus and therefore localise lesions in abnormal scans then perhaps the method may make undetectable lesions in apparently "normal" scans more discernible. This
undoubtedly would be an area of further study, although the practicalities of proving this with patient data are formidable. Ideally, lesions have to be verified by autopsy or surgery, but this alone may present ethical problems particularly as equivocal scintigrams require the closest scrutiny. R.O.C. plots may be used but the analysis will need to include many (at least 100) studies if statistically significant conclusions are to be drawn.

When the author commenced this study he did not appreciate that biological artefacts from the motion of amorphous organs were so gross. The liver has been used to highlight the problems although similar problems have been observed in deconvolution of thyroid scans. In the latter case artefacts can be induced by the act of swallowing as the detector head traverses the thyroid area. Lung scintigrams may be subject to perturbations induced by respiration and cardiac activity. Deconvolution of such organs can apparently greatly "enhance" these artefacts or discontinuities in the data, rendering it very difficult for the clinician to estimate the true distribution of radioactivity within an organ. This phenomenon should not be treated as a deficiency of the deconvolution methods but as an indication of the sensitivity of the methods to locate fine detail in the object data. In this investigation mathematical
instability in deconvolved scintigrams has been associated with instrumental or physiological artefacts within the data. Usually it is in retrospect (after poor deconvolved results) that these artefacts are first noticed in the unprocessed data. Indeed poor deconvolved results may suggest that the unprocessed data should be interpreted with the utmost caution.
ACKNOWLEDGMENTS

I would like to acknowledge my supervisor, Dr. P. Tothill, for his support and advice throughout the project. I am indebted to Prof. J.R. Greening for making the facilities at the Department of Medical Physics and Medical Engineering available. I wish to thank Mr. S.H.C. Hughes for his critical advice and help on mathematical aspects of the project.

My thanks go to Miss. J. Donaldson for help on the construction of the gamma camera interface and the preparation of figure 2.9, and to Mr. P. Martin for help on the construction of the television system.

In addition I wish to acknowledge the Medical Research Council for financing the purchase of the basic PDP 12 computer system, and to the University of Edinburgh and the Scottish Home and Health Department for providing extra peripheral equipment connected to the computer.

Finally I thank my wife, Christine, for giving moral support and encouragement particularly during the preparation of this script.
REFERENCES

"Image processing by digital computer". I.E.E.E.
Spectrum V9 pp20-32.

Bits and Pixides (1975)
Newsletter No4 of Nuclear Medicine Computer Users
Group.

Boardman A.K. (1976)
"An inexpensive computer colour television system
for the presentation of radionuclide scintigrams".

Bone D., Pullan B.R. and Taylor C.J. (1973)
"Reconstructing pictures from sampled data". Proc.
of 3rd conference on information processing in
scintigraphy (Cambridge, Massachusetts), pp125-136.

Bracewell R. (1965)
"The Fourier transform and its applications". McGraw
Hill.

Brownell G.L. and Chesler D.A. (1971)
"The basic concepts of resolution enhancement". In
Quantitative Organ Visualisation in Nuclear
Medicine, pp509-530. Edited by P.J.Kenny. Miami
Press.

Budinger T.F. (1972)
"Clinical and research quantitative nuclear medicine
system". Medical Radioisotope Scintigraphy Vol 1

Clifton J.S. (1971)
"Analog displays for quantitative visualisation". In Quantitative Organ Visualisation in Nuclear Medicine, pp 679-694. Edited by P.J.Kenny, Miami Press.


Corfield J.R. (1975)


Helstrom C.W. (1967)


Hunt B.R. (1971)

Hunt W.A. (1970)
"Optimum sample size in digital radioscintigraphy". IBM Publication 70.09.004.

I.A.E.A. (1972)


Iinuma T.A. (1971)
"Image enhancement by the iterative approximation method". In Quantitative Organ Visualisation in Nuclear Medicine, pp549-580. Edited by P.J.Kenny, Miami Press.


"Nonlinear frequency domain techniques for enhancement of radionuclide images". Proc. of 3rd conference on information processing in scintigraphy (Cambridge, Massachusetts), pp102-114.

Levinson N (1946)
"The Wiener RMS (root mean square) error criterion in filter design and prediction". J. Math and Physics V25, No4, pp261-278.

Levinson N. (1947)
"A heuristic exposition of Wiener's mathematical
theory of prediction and filtering". J. Math and Physics V26, No2, pp110-119.

Limb J.O. (1969)


Lynn P.A. (1973)
"The analysis and processing of signals". Macmillan

MacIntyre W.J., Fedoruk S.O., Harris C.C., Kuhl D.E. and Mallard J.R. (1968)

Mersereau R.M. and Dudgeon D.E. (1975)

Metz C.E. (1969)


"Digital signal processing". Prentice Hall

Palmer L.. (1973)
"A pseudo random number generator". DECUS program library 8-690.

Phillips D.L. (1962)

Pistor P. (1972)
"The Heidelberg scintigraphic imaging system". Proc. of 2nd. conference on sharing of computer programs and technology in nuclear medicine (Oak Ridge, Tennesse), pp411-427.

Pistor P. (1972)
"Digital processing of scintigraphic images by two
dimensional recursive Wiener filters". I.B.M. publication 70.12.006.

Pullan B.R. (1976)
"Principles of image processing". Paper given at Hospital Physicists' Association annual conference.

Rabiner L.R. and Gold B. (1975)
"Theory and application of digital signal processing". Prentice Hall

"Wirephoto quality improvement by unsharp masking". Pattern Recognition V2, pp117-121.

Slepian D. (1967)

Stockham T.G. (1972)


Todd-Pokropek A.E. (1972)
Todd-Pokropek A.E. and Pizer S.M. (1976)

Twomey S. (1965)

Van Cittert P.H. (1931)
"Zum einfluß der spaltbreite auf die intensitatsverteilung in spektrallinien". Zeittschrift fur Physik V69, pp298-308.

Vander-Lugt A (1968)
MATHEMATICAL SYMBOLS

\[ i = \sqrt{-1} \]

\[ \star \quad \text{Convolution} \]

\[- \quad \text{Complex conjugate (Eg. } \overline{A} \text{ is the complex conjugate of } A) \]

\[ \hat{\ } \quad \text{Computed estimate (Eg. } \hat{A} \text{ is the computed estimate of } A) \]

\[ | | \quad \text{Modulus (Eg. } |A| \text{ is the modulus of } A \) \]

\[ | \quad \text{Modulo (Eg. } A \mod N \text{ is the value of } A \text{ to modulo } N) \]

Other symbols are explained in the text.