

THE MANIPULATION OF IMAGE STATISTICS USING ELECTRONIC HARDWARE SYSTEMS

THESIS

presented by

ANTONIO CANAS WILKINSON

for the degree of

DOCTOR OF PHILOSOPHY

in the

UNIVERSITY of LONDON

and the

DIPLOMA OF MEMBERSHIP OF THE IMPERIAL COLLEGE

Optics Section,
Blackett Laboratory,
IMPERIAL COLLEGE
LONDON SW7 2BZ

October 1983

THE MANIPULATION OF IMAGE STATISTICS USING ELECTRONIC HARDWARE SYSTEMS.

ABSTRACT OF THESIS :

The image processing described in this work is based on the use of fast, low cost and highly interactive hardware systems which can modify the statistics of an input image. As an introduction to the field the various technologies for the acquisition, storage and processing of pictures are reviewed. The information storage capacity of standard Landsat 70 mm transparencies is investigated by means of densitometry with the aim of characterising the effects of noise on classification. A second approach to this problem makes use of bitplane images and their autocorrelations.

The design and construction of a digital system for nonlinear contrast enhancement is presented. This system operates on digitised video images in real time and is controlled by a microprocessor. The choice of stretching function is discussed and various means of user interaction are presented. Examples of the operation of this system are given for some trial images including infra red linescan data and radiographic pictures.

The use of the Principal Components Transformation for data compression in multispectral work is described together with the normal methods of application. A novel method for the realisation of this process is presented which makes use of high speed electronic circuits in the acquisition of statistics and in the transformation itself. The design and construction of digital image stores to hold multispectral data are described together with the transformation electronics. The system is able to carry out the PCT within a few seconds.

Finally a review of the work is given and directions for possible future research are indicated.

TABLE OF CONTENTS

ABSTRACT
CONTENTS
ACKNOWLEDGEMENTS

Chapter / section	page
<u>CHAPTER 1 : IMAGES</u>	
1.1 Introduction	1
1.2 Images	3
1.3 Image acquisition systems	7
1.4 Storage and display systems	26
1.5 Image processing systems	33
1.6 Precis	36
<u>CHAPTER 2 : NOISE AND INFORMATION IN PHOTOGRAPHIC MEDIA</u>	
2.1 Objectives	39
2.2 The photographic process	41
2.3 Densitometry of film samples	42
2.4 Factors affecting noise in practical systems	55
2.5 Bitplanes and digital storage	62
2.6 Images and information	75
<u>CHAPTER 3 : STATISTICALLY BASED ELECTRONIC CONTRAST ENHANCEMENT</u>	
3.1 Contrast enhancement	76
3.2 Transfer functions and image statistics	77
3.3 Digitisation and transformation	84
3.4 Video signal reconstruction	93
3.5 Histogram acquisition	95
3.6 TV camera interface	100

Chapter / section	page
3.7 The system microcomputer	101
3.8 System software	105
3.9 System summary	108

CHAPTER 4 : INTERACTIVE CONTRAST ENHANCEMENT

4.1 Tactical image processing	109
4.2 Experimental considerations	112
4.3 Basic processing	113
4.4 Manual stretch specification	121
4.5 Contrast rearrangement	129
4.6 Conclusions	136

CHAPTER 5 : THE MANIPULATION OF FEATURE SPACE

5.1 Information in feature space	137
5.2 The principal components analysis	138
5.3 Noise sensitivity	142
5.4 Contrast enhancement of PC channels	144
5.5 Covariance acquisition by electronic hardware	145
5.6 Eigenvector determination	148
5.7 Feature space rotation	150
5.8 Digital framestores	153
5.9 The buffer unit	159
5.10 The PCA system - a summary	164

CHAPTER 6 : PCA BY ELECTRONIC HARDWARE

6.1 A practical PCA system	166
6.2 A 2-dimensional PCA example	171
6.3 System performance	177
6.4 A 4-dimensional Landsat PCA	182
6.5 Post processing and interpretation	192

Chapter / section	page
<u>CHAPTER 7 : DISTILLATION AND COMMENT</u>	
7.1 Introduction	202
7.2 Photographic storage	204
7.3 Nonlinear contrast enhancement	206
7.4 The PCA system	209
7.5 Conclusions	212
List of References	214

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to the many people who have offered advice and help during the course of this work. Thanks are due to Dr M. E. Barnett for his supervision, and to William Hossack of Q.E.C., London for his invaluable computing help for the section on bitplane autocorrelations in chapter 2. Also worthy of particular note are Dr Chris Bird for his comments and proof reading, and Dr Lawrence Robinson for practical and moral support. The author wishes to acknowledge the financial support of the SERC during this research.

This thesis has been produced using the Format-80 word processor on an Apple II computer using a Theme 11 point PS typeface.

TO THE AGÉD P's AND D.D.D.

CHAPTER 1 : IMAGES

1.1 Introduction

The most important medium for the storage, transmission and presentation of information is the image. The worth of a modern image in terms of information content is measured in tens of millions of words, a significant advance over the proverbial thousand, and the proliferation of image sources presents a difficult problem for the archivist and for the would-be user alike. A typical evening's television viewing may present the viewer with 45 gigapixels of information, each one coded for brightness and colour, during a three hour period (a pixel is an elemental sample of an image, a two-dimensional array of which can be used to compose any picture). This number, which is so large as to be almost without meaning, is actually severely limited by the constraints of channel bandwidth, since the commercial television system is a relatively low resolution display medium. A 35 mm monochrome negative may contain 10 megapixels, assuming a 10 μ m pixel size on fine grained film, which compares with the 180 kilopixels of a TV frame. In the face of such enormous quantities of data the common tendency is to take images for granted and belittle the scale of the problem, but if a serious attempt at image interpretation is required, or if an image must be acquired, transmitted and stored, then a full appreciation of images and their characteristics is essential.

The TV picture and the 35 mm negative are only two of a large number of image storage and display systems, each with varying properties, which confront the research worker in image science. Often an image must be acquired by a system under adverse conditions and in an inaccessible location, as in the case of space probes or medical endoscopy, and the image may be the only source of data about the area or object of interest. Unfortunately it is rare to find an image which completely satisfies every user, especially as each one may have an interest in a different aspect of the scene and the limitations of the environment and the system always conspire to degrade the required details. Faced with a "poor" image, the user must apply some type of enhancement in an attempt to render the image interpretable in his particular application.

Historically, image processing to aid interpretation arose first when artificial aids were used to supplement the performance of the naked eye as an

acquisition system. Magnification and light gathering limitations were lessened by the introduction of the telescope by Galileo Galilei in the 16th century, the use of reflective optics by Newton in the 17th century and the invention of the microscope by Hans and Zacharias Jansen of Holland [1] at about the same time. These optical systems all share two common properties in terms of their effects on an image: firstly, they process the entire image at once, and secondly, they operate in real time, with no significant delay. These instruments are thus, in modern terminology, highly parallel in operation.

The next significant advance in image processing came with the use of photography to store an image and reproduce it by a printing process. The taking of a photograph is another parallel operation and the film or plate is a relatively fast storage system, although a development process is required before a useable image can be seen (the film can store a latent or undeveloped image simply by exposure, and development may be delayed as long as the film is kept cool). The printing process allows some simple yet powerful processing to be carried out, such as contrast modification, gamma correction [2], and edge detection by unsharp masking [3]. These processes are all based on a series of chemical reactions whose conditions must be carefully controlled, which are slow and difficult to reproduce, and whose exact operation is usually hard to quantify. These limitations do not prevent much useful work from being undertaken photographically, but a more immediate and accessible (at the parameter level) method of processing is required.

The use of electrical signals to represent images allows just such access with considerable speed and flexibility (compare the adjustment of contrast and colour balance of a photograph with the adjustment controls of a TV set), but it requires that the image be acquired in some electrical format, and be available for conversion to a displayable form for viewing. The most important device for this purpose is the television camera tube, developed in several forms during the early part of this century, and now perhaps the most common means of image acquisition after the film camera, although it is rapidly catching up. Once the image can be represented as an electrical signal, many processes become possible including that of processing by means of a computer, and high speed transmission of large amounts of image data e.g. by satellite.

This chapter will outline some of the properties of images and their acquisition, storage and display systems. Brief details of some modern image

processing systems will be given and some ideas as to the philosophy behind the design and use of these systems will be presented, taking into account the requirements which must be met by such systems.

1.2 Images

An image is the representation, with appropriate scaling, of an intercepted energy field. Usually, for most real images (as opposed to computer synthetic images), the scaling allows the image to be represented as a two-dimensional function with positive values between maximum and minimum limits such as black to white or some similar colour coding. The image function represents the spatial or in some cases spatial-temporal distribution of some parameter. Most often this is an electromagnetic field intensity distribution, such as visible, ultra-violet (UV) or infra-red (IR), but more recently radio wavelength images have become available both from astronomy and from imaging radar systems. In fact it is possible to obtain images over nearly all the electromagnetic spectrum which spans nearly twenty orders of magnitude in wavelength. The EM field which is measured may be generated by the imaging system itself, as in the case of Synthetic Aperture Radar which has been used both from aircraft and satellites for terrain mapping, or it may be a natural field such as reflected sunlight or thermal emissions from hot objects. In most cases, the precise values of the field intensity are not in themselves of interest, but rather the relative values and the way in which the field is modulated by some parameter of interest in the scene. An example of this is the monitoring of soil toxicity by plant health: the chlorophyll of a green plant absorbs visible light for photosynthesis but reflects in the IR region. The wavelength at which the absorption / reflection transition occurs and the strength of the absorption are affected by the health of the plant, which is in turn affected by soil conditions. Airborne or spaceborne imagers can be used to measure the spectrum of reflected light and thus determine soil conditions indirectly. Other parameters which are often measured in this indirect way include mineralisation and its associated alteration of rock reflectance in the visible and IR regions of the spectrum, and, more recently, the distribution of certain metabolites in the body by nuclear magnetic resonance imaging (NMR).

Whatever the parameter to be measured, the data which is presented pictorially is always a real-valued function of two real variables defined over

a finite region and limited to a finite range of values. This will always hold true even if the function values do not themselves represent a real parameter such as a light intensity. An image function is only defined inside a specified boundary of some standard size and shape, usually square, and is assumed to be zero outside. The value of the image function at any point is a "grey level" or "brightness value", and it must be non-negative and bounded, i.e. if $P(x,y)$ is a picture function, then

$$0 \leq P(x,y) \leq M \quad \text{for all } (x,y)$$

where M is a maximum or upper bound. The picture function may be operated upon by mathematical procedures such as Fourier Transformation or Edge Enhancement, but to utilise any of these techniques analytically is usually impractical due to the complexity of $P(x,y)$. For this reason, it is usual to apply a numerical computation algorithm to an approximate form of $P(x,y)$ in order to produce the desired effect, but this approximation will inevitably lead to some degradation of the results. The gains which offset this loss are a reduction in the storage space required for the image and, often, an acceleration and simplification of procedure.

One common technique is to derive a "digital picture function" or "digital image" from the original data. This is a discrete array of numbers (a matrix) whose elements represent samples of the picture $P(x,y)$ defined in a similar way to the following: given an $n \times m$ matrix with row and column indices i,j such that $1 \leq i \leq n$ and $1 \leq j \leq m$, a series of square areas A_{ij} are defined by

$$i-1 \leq x < i$$

$$j-1 \leq y < j$$

and a constant value (d_{ij}) is assigned to each square, these forming the elements of the digital matrix D . The value (d_{ij}) may correspond to the value of the picture function P at the centre of the square A_{ij} or it may be an average value of P over A , i.e.

$$(d_{ij}) = \frac{1}{A_{ij}} \iint_{A_{ij}} P(x,y) \, dx \, dy \quad (1.1)$$

This integration may be performed by the acquisition system rather than being explicitly applied to the raw data. As well as discretising the spatial coordinates, a quantisation is normally applied to the picture values, thus allowing only a finite number of grey levels. The result is an approximation of $P(x,y)$ in the form of a matrix whose elements are known as pixels or pels (from

picture elements), and the values of n and m are usually chosen to be equal and an integer power of two for computational simplicity. The matrix $\underline{D}$ can be made indistinguishable from the original picture function $P(x,y)$ provided that the spatial and brightness quantisations are sufficiently fine.

Image processing consists in the careful analysis of initial and desired image properties and the elucidation of an appropriate transformation process, rather than the arbitrary application of mathematical functions. Complex processes may require or assume certain properties of the input image function such as continuity or integrability, i.e. that the function is "well behaved". Such assumptions may be rationalised if not proven, and a simple treatment is found in [4].

This section began by employing the concept of an image as a two-dimensional brightness function representing some type of radiation field. This is only one of a number of application-dependent approaches to the treatment of images, each of which can lead to different types of image transformation. An alternative approach is to consider the image to be a collection of objects which together compose a scene. These objects can be assigned properties or features the exact nature of which will depend on the application in hand. The objects may now be sorted into classes on the basis of these features, each class being defined by a "significant difference" in one or more of the features of an object. This approach is fundamental to the study of pattern recognition, but is much more widely useful.

The features chosen are arbitrary, and may include colour, length, shape or texture. If numerical values are assigned to these features (possibly quantised) then the image may be represented in an alternative form: Feature Space. This is a multi-dimensional hypothetical space whose axes are defined by the features in use (the number of features defining the dimensionality of the space), and within which objects are located by "feature vectors". A feature vector is a multi-dimensional vector whose components are the feature values along the basis vectors of the space. An illustration of this concept arises in the use of multispectral images such as those obtained by the Landsat series of satellites. Four monochrome images are obtained for each area of ground, each one corresponding to the reflected intensity within a narrow spectral band. Thus each point on the ground has four spectral samples associated with it which can be used as features of that point on the ground. The image is composed of

different ground cover types, each with different "spectral signatures", and a common requirement is to pick out a particular cover type. The power of the feature space concept now becomes evident: the four spectral samples are formed into a feature vector whose tip position can be measured. Since similar cover types (with similar spectral signatures) will generate similar feature vectors, the operation is reduced to detecting which vectors point into a subvolume of feature space, and this may be accomplished automatically. The number of clusters of feature vector tips is equal to the number of classes in the image, and the density of a cluster is a measure of the number of points which fall in that particular class, although cluster definition is somewhat arbitrary.

An image cannot conveniently be displayed as a feature space distribution since in general there are more than two dimensions or features involved, and multiple two-dimensional projections must be used. It must be remembered that, assuming all the useful features of the objects in a scene are described, a feature space representation is equivalent to the original image, but displays the available data in a different form.

Another useful approach to images is to consider them as samples of a statistical process with spatial and temporal variables. This process is described by a probability density function (PDF) which is usually unknown and which must be approximated from a study of the image(s) involved. Some simple modelling may be attempted by reference to the physics of the processes which gave rise to the image, as in the case of noise analysis from sensors, but such studies are in general complex. The simplest way in which to characterise a statistical process is by the measurement of the ensemble averages, i.e. the mean, variance, autocorrelation etc. The term "ensemble" signifies the concept of a process which is characterised by samples, all of which taken together form the ensemble. The statistical process is said to be stationary if the moments of the PDF are invariant under coordinate or time translation, or, more practically, if the mean is constant and the image autocorrelation function is independent of particular coordinate values [5]. As in all statistical studies, great caution must be exercised when selecting samples for analysis, for, as will be shown in later chapters, statistics may vary widely within an image.

A statistical approach allows the effects of processing, acquisition and storage systems to be evaluated on the basis of their modification of image statistics. If a model can be constructed for the system process, and a

transfer function derived, then for any given input statistics, the output statistics can be computed. A major use of this type of study is in the field of image encoding and data compression, where an image must be transmitted or stored in the minimum of time or bandwidth without significant loss of fidelity. Along this road lies the complex area of information content in images, which will be considered further in a later chapter. The statistical approach forms the basis for much of the practical work presented in this thesis.

1.3 Image acquisition systems

The purpose of an image acquisition system is to gather pictorial data and make it available for subsequent processing or storage in a suitable form. The exact manner in which the system operates in acquiring the data determines the way in which the picture function is represented and approximated. In designing such a system many factors must be considered, including the type of data to be gathered, accessibility of the acquisition system, speed of data reception and transmission and the volume of data required (the image resolution). One parameter whose value results from such factors is the degree of parallelism of the system i.e. the number of separate but simultaneous data channels through which the image is sensed. A broad range of values exists, from very highly parallel film cameras through intermediate multiple detector scanners such as the Landsat multispectral scanner (MSS) to single detector scanners such as an IR LineScan unit (IRLS).

The most common method by which the image function is gathered is some form of scanning, as referred to above, in a similar way to a domestic TV screen. The image is divided into "scan lines", and the image function is sampled along a parallel set of these lines until the entire image has been sampled. Further spatial quantisation may occur along these lines to form pixels as detailed in section 1.2, or a continuous analogue signal may be output for each scan line as in the case of a TV camera. Grey level quantisation may then be effected by digitisation of the detector signal after suitable conditioning.

In considering all image acquisition, processing or display systems it must be remembered that the final image will have to be interpreted by the eye and its associated processing system. For this reason it may be useful to match the image characteristics to the eye to aid interpretation, and this will be covered

in a later chapter. The eye is a highly parallel system which employs two types of detector: rod cells which are panchromatic and three classes of cone cells which are colour-sensitive. The rods, numbering 75 to 100 million, are used for low light levels and peripheral motion detection with low resolution, and are spread over the retinal surface. This dim light vision is termed "scotopic" as opposed to the "photopic" or bright light vision of the cones. These cells are concentrated around the optical centre of the retina (the fovea centralis) and number approximately 7 million. They are capable of fine resolution at high light levels and of giving colour information since the three classes are sensitive to red, green and blue light. The optical image is formed by a two stage lens system composed of the cornea and the lens, the lens acting as a variable focal length "fine tuning" element. Since the system is refractive, chromatic aberration is present and there is a 2 dioptre difference in power between blue and red light. The focal length is approximately 17 mm and the aperture ranges from 1.5 mm to 8 mm depending on light levels.

An important parameter when discussing imaging systems is the sensitivity to changes of light level or colour. In the case of the visual system this is a complex area, but sensitivity must not be confused with dynamic range. The eye can detect light levels over a range of 10 orders of magnitude, and the perceived brightness is a logarithmic function of the light incident intensity as is shown in figure 1.1. The eye cannot utilise the entire range at once but adapts to an average level B_a shown on the figure and discriminates brightness values in the range $B_a - B_b$. Any levels below B_b are black and the presence of levels much above B_a causes the average or adaption level to rise.

Thus the instantaneous dynamic range is around two orders of magnitude. The contrast sensitivity is measured by comparing the subjective brightnesses of two adjacent target areas surrounded by a constant background level. The difference of brightness which is just perceived (ΔB) divided by the target brightness (B) is known as the Weber ratio, and is found to vary with the background level (therefore with the adaption level) as shown in figure 1.2 [6].

As the target brightness B differs from the background B_0 the sensitivity changes and the target's apparent brightness also changes. The narrow sensitivity range suggested by these results is made worse in a real image due to a lack of sharp brightness changes and boundaries, and only about 20 levels are distinguishable over any small area of a complex image. Display systems should be capable of more resolution than this since the eye adapts very quickly to different average levels depending on the area viewed and can even adapt

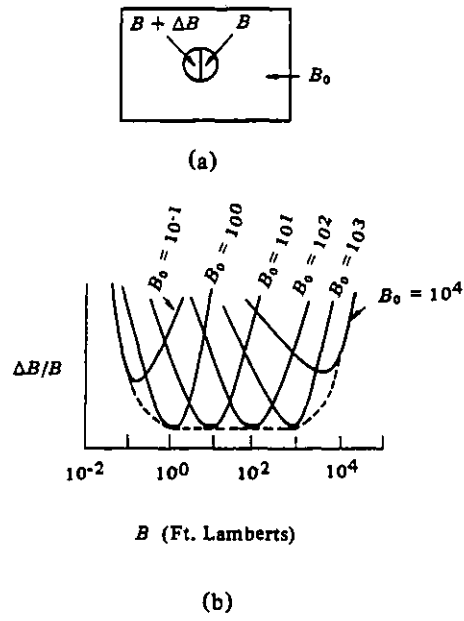


Figure 1.2 : Perceived brightness vs. intensity

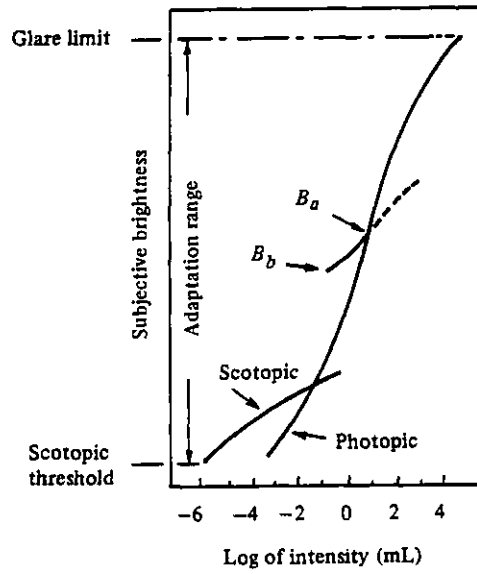


Figure 1.1 : Contrast sensitivity of the eye

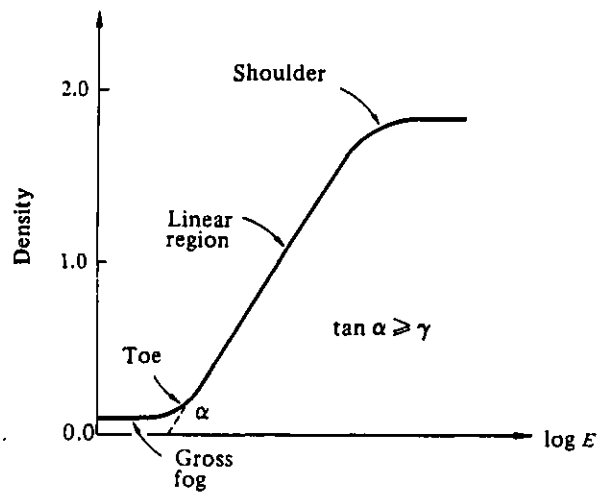


Figure 1.3 : Hurter and Driffield (H.D.) film response curve

differently across the retina, thus giving a total range of intensity discrimination approaching 100 levels in any image. Since the visual system is complex other factors may influence the perceived brightness of a target, such as the spatial frequencies or edge overshoot effects (the Mach Band) in which a uniform area appears variable near boundaries with other levels. The question of available grey levels will be addressed further in chapter 2 with reference to film responses.

Although the eye suffers from many distortions and aberrations, the brain can "remove" or obscure the defects to a large extent, and the overall system is very versatile. The data rate is large even though some preliminary compression and processing is carried out in the retina, and in some ways it is a pity that there is no 'external output' socket! In a film camera the optical defects can be reduced by the use of corrected lens systems and an irregular array of silver halide crystals replaces the retina. This system can extend the spectral range to the far UV and with sensitising dyes to the near IR. The data rate is again high, but as was noted in section 1.1 the image is not immediately accessible because of development. The grain size of modern films is small, around 400 Å, with typical resolutions up to 6000 lines per mm in a holographic or precision lithography emulsion, but the image will be blurred because of "mottle" effects caused by nonuniformities in the grain distribution. This causes the film to yield noisy images, and a study of this noise forms the basis for chapter 2. The recording process gives a spatial density variation over the film area corresponding to the exposure or integrated intensity at each point. Optical density is defined as the negative logarithm of the transmittance, and it is found that the relation between density and Log exposure is linear over a wide range of exposures, the exact transfer function being referred to as the Hurter and Driffeld (H and D) curve as in figure 1.3. The slope of the linear region is the "gamma" of the film, and indicates the contrast. Note that at high exposures the curve is non-linear due to saturation, and that there is a residual "fog" level even at zero exposure. These nonlinearities are termed "reciprocity failure". Films can be made to resolve colour by the use of suitable sensitising dyes, but although three primary colours are used as in the eye, the "detectors" are placed in separate layers in the film rather than being side by side.

The film is used as the recording medium for some type of image-forming optical system which intercepts the desired radiation and scales the image to

the required size and shape. This optical system does not give a faithful reproduction of the object field but degrades it according to the parameters of the system. Apart from aberrations, the limiting effect is due to diffraction, and the system can be characterised either by a Point Spread Function (PSF) or by a Modulation Transfer Function (MTF). The PSF is obtained by performing a Fourier Transform (FT) on the entrance aperture, and the degradation of the final image is calculated by performing a convolution of the PSF with the perfect image function. Alternatively, if the spatial frequency spectrum of the input image is known or computed, the output image is formed by transforming the product of this spectrum and the MTF. These statements are expressed in a mathematical form as follows:

$$I_o(x,y) = I_i(x,y) * P(x,y) \quad (1.2)$$

$$I_o(x,y) = F^{-1}[M(u,v) \cdot F[I_i(x,y)]] \quad (1.3)$$

where u,v are spatial frequency variables, $*$ represents convolution and $F[]$ denotes the application of a two-dimensional Fourier Transform. These equations are equivalent since application of the convolution theorem to 1.2 together with

$$M(u,v) = F[P(x,y)] \quad (1.4)$$

gives (1.3) because the PSF and MTF form a Fourier pair. In the case of the common circular aperture it is possible to calculate the form of the PSF, known as the Airy function [7] and this is given in equation (1.5) :

$$A(r,\theta) = I_o \left[2J_1(kaw) / (kaw) \right]^2 \quad (1.5)$$

One important derivative of the circular aperture is the Cassegrain-type of reflecting telescope which has an annular aperture. The PSF for this type of telescope can be derived from equation (1.5) by invoking Babinet's principle and has the form shown in equation (1.6) where a is the outer radius, ea is the inner or obstructing radius ($0 \leq e \leq 1$), w is the angular radius, k is the wavenumber and J_1 denotes a Bessel function of the first kind and order

$$I(p) = I_o(1-e^2)^2 \left[\frac{2J_1(kaw)}{kaw} - \frac{2e^2 J_1(kaew)}{kaew} \right]^2 \quad (1.6)$$

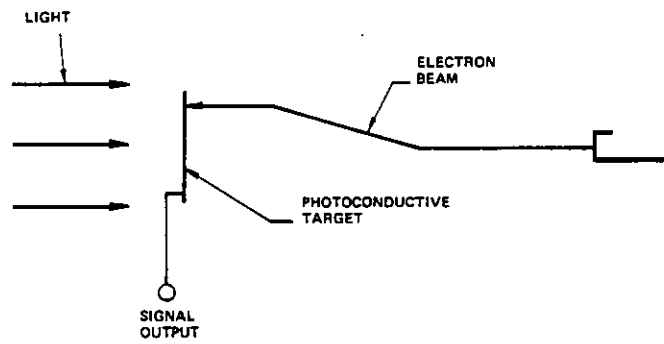


Figure 1.4 : VIDICON TV camera tube

I_0 is the central intensity of the pattern given by

$$I_0 = C \frac{2}{\pi} \frac{2}{a^4} (1 - e^{-2})^2 \quad (1.7)$$

where C is a (complex) pupil transmittance function which is assumed constant over the radius range $a \gg r \gg ea$. Although circular and annular apertures are the most common, some scanning systems utilise complex optics and analytic computation of their PSFs becomes difficult.

If an electrical output is required from the imaging system then some form of detector must be placed in the image plane of the optics. In almost all cases this involves some form of scanning, either electrical or mechanical, to yield an essentially serial stream of data. Varying degrees of parallelism are possible but on all but the smallest images it is impractical to acquire the whole picture at once. If a stationary image is used then electrically controlled scanning can be used as in the case of a TV camera tube or the newer Charge Coupled Device (CCD) diode array cameras, and this is known as Image Plane Scanning. Alternatively it is possible to place an oscillating mirror in front of the optics to sweep the image across a central detector or group of detectors (Object Plane Scanning). This last approach has the advantages of working near to the optical axis of the system where distortions are less and of using a small number of detectors whose responses may readily be matched.

The most common method of image plane scanning by electrical means is the TV camera tube. This device uses a photoconductive material as a detector and the input optics form an image on one surface of this target. Scanning of the image is accomplished by means of a suitably deflected (either electrostatically or magnetically) electron beam impinging on the rear surface of the target as in figure 1.4. The front of the target is covered with a conductive transparent signal electrode operated positive with respect to the rear surface which is near to the cathode potential. The beam charges the rear surface to cathode potential but the incident light of the image causes the photoconductor to "leak" charge and a pattern of more or less positive charge is formed on the rear surface according to the image intensity. When the electron beam scans the target again it deposits electrons on the positive areas and gives a capacitively coupled signal on the target electrode. If the image is very bright two passes of the beam may be required to recharge the target. Various target materials are used depending on sensitivity and spectral range and

include antimony trisulfide, lead oxide and selenium-arsenic-tellurium alloys. This type of camera tube is called a "Vidicon". Other types of tube work in the "Return Beam" mode and the primary photosensitive element is a photocathode as shown in figure 1.5. Photoelectrons emitted from the front of the tube are focussed magnetically onto a thin target and build up a positive charge pattern by secondary electron emission. The scanning beam deposits charge on the target as before but the unused portion of the beam is reflected and collected by an electron multiplier at the back of the tube which amplifies the signal. This "Orthicon" type tube can cope with high contrast ranges and is sensitive. A variant of this type of tube employs the backscattered electrons from the scanning beam rather than a reflected beam and is called an "Isocon" [8] and is depicted in figure 1.6. The form of scanning is arbitrary and may depend on the application but the most common type is a rectangular raster composed of parallel lines with an aspect ratio of 4:3 as in commercial broadcast systems. The number of lines is variable but is usually in the range of 500 to 1500, and the scan rate is usually arranged so as not to show visible flicker when the picture is displayed. To reduce the flicker it is common practice to use two interleaved scanning fields per frame [9].

As in the case of the optical system the output image is a degraded form of the input illumination. This degradation arises because the electron beam forms a spot of finite size on the target usually with a Gaussian intensity profile. The resultant image is the convolution of the input with the Gaussian profile, and the effect is that of low pass filtering with a very sharp cut-off slope along the scan lines. The effect on resolution perpendicular to the scan lines is more complex to interpret and involves sampling theory. If it is required to image a horizontal bar pattern of width and spacing D using scan lines of the same separation then the contrast of the image depends critically on the relative position of the scan lines and the bar pattern since sampling is being attempted at the Nyquist limit. It is found by experiment that the bar pattern will always be visible if the scan line spacing is $K.D$ where K (the Kell factor) has a value of approximately 0.7. The output image is presented in the form of an analogue video signal for each line and synchronisation pulses are added to mark the start of each line and frame [10].

There exist many electro-mechanical scanners which perform image scanning by means of a mechanical motion of the image or sensor(s), ranging from small ground based systems to airborne multispectral IR and spaceborne resource

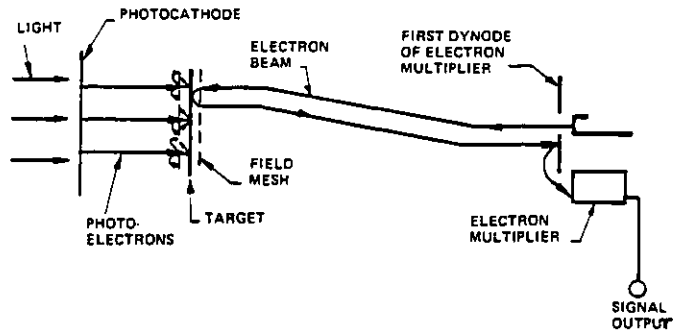


Figure 1.5 a : Image ORTHICON TV camera tube

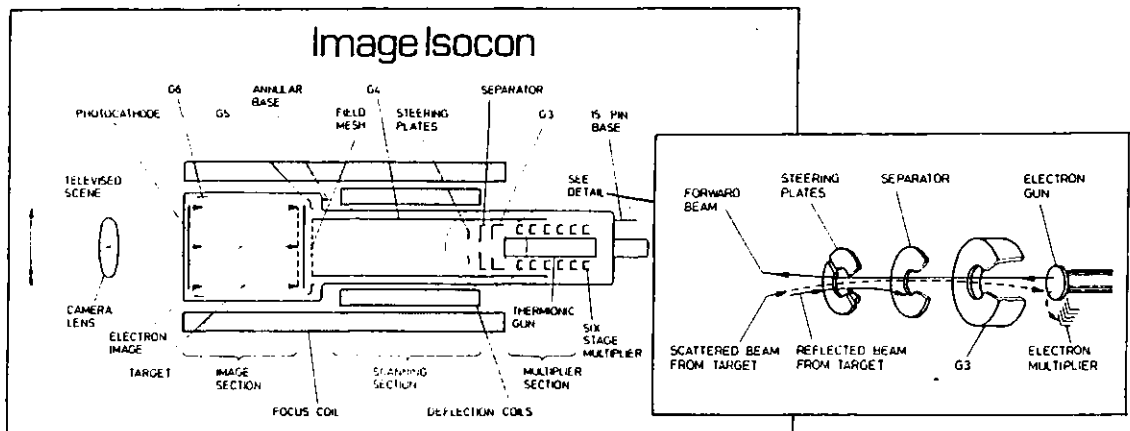


Figure 1.5 b : Image ISOCON TV tube

monitoring packages and a review of the Landsat or Earth Resources Technology Satellite (ERTS) multispectral scanners serves to indicate the most important features of such systems. This series of four satellites launched since 1972 has been designed for resource mapping applications and each is equipped with a four-band multispectral scanner (MSS) and various return beam TV tubes as described above. Their near polar orbits are designed for repetitive ground coverage with a period of about 18 days at a nominal altitude of 920 km (these figures are different for Landsat 4 which carries the higher resolution Thematic Mapper at an altitude of some 700 km). Images are acquired with a ground coverage of 185 km square and a resolution of around 80m and adjacent images are overlapped to allow mosaicing and continuous coverage. The major details of the MSS system are set out below in table 1.1:

TABLE 1.1 LANDSAT MSS SPECIFICATIONS

Telescope optics	22.9cm aperture Ritchey-Cretien with 8.9cm diameter secondary
Focal length	82.3cm
f number	3.6
Scanning method	Oscillating flat mirror $\pm 2.9^\circ$ at 13.62 Hz
Cross track field of view	11.6°
Optical fibre core size	68.3 μ m square
Blur circle diameter	27 μ m
Spectral bands and detectors (PMT = PhotoMultiplier Tube)	Band 4: 0.5 - 0.6 μ m PMT Band 5: 0.6 - 0.7 μ m PMT Band 6: 0.7 - 0.8 μ m PMT Band 7: 0.8 - 1.1 μ m Si diode
Lines per scan per band	6

Figure 1.6 shows the MSS in section together with the arrangement of the detectors. In order to economise on scan rate a group of six detectors is used for each of the four bands so that 6 "horizontal" lines of image are collected during each mirror cycle. The orbital motion of the satellite provides the "vertical" scan motion. The detectors and spectral filters are mounted behind the focal plane of the telescope and the image intensity is transferred by means of an array of 24 clad optical fibres, and this cladding gives rise to a separation of adjacent scan lines on the ground. The spatial resolution of the scanner is difficult to define, but there are four major approaches to its characterisation [11]:

- (1) Geometrical properties of the imaging system
- (2) Resolution of point targets
- (3) Measurement of periodically varying objects
- (4) Spectral signatures of small targets

The first measure is calculated from the focal length of the optics, the altitude and the detector (fibre) size, and the result is a ground-projected "footprint" of the detector. The Instantaneous Field Of View (IFOV) is computed from the following equation:

$$\text{IFOV} = H \cdot d / f \quad (1.8)$$

where H is the altitude, d is the detector size and f is the focal length. Note that if an object smaller than the IFOV falls in the centre of the footprint it may be detected if it has a significantly different intensity from the surrounding area but its size will always appear to be one IFOV.

The above IFOV takes no account of diffraction effects in the optics, and some reference to the system PSF must be used. In the case of the MSS the PSF has the form of equation (1.6) due to the annular aperture and the usual figure quoted is the angular radius at which the PSF falls to 50% of its central value. Compared to the Airy pattern for a circular aperture this PSF has a narrower central peak but more energy appears in the side lobes. Values for the diffracted energy distribution have been computed for various obscuration ratios (e) and source sizes [12], and the conventional Rayleigh resolution criterion is often modified to the threshold case for complete resolution for greater radiometric accuracy. It is found that if the source size is small then an

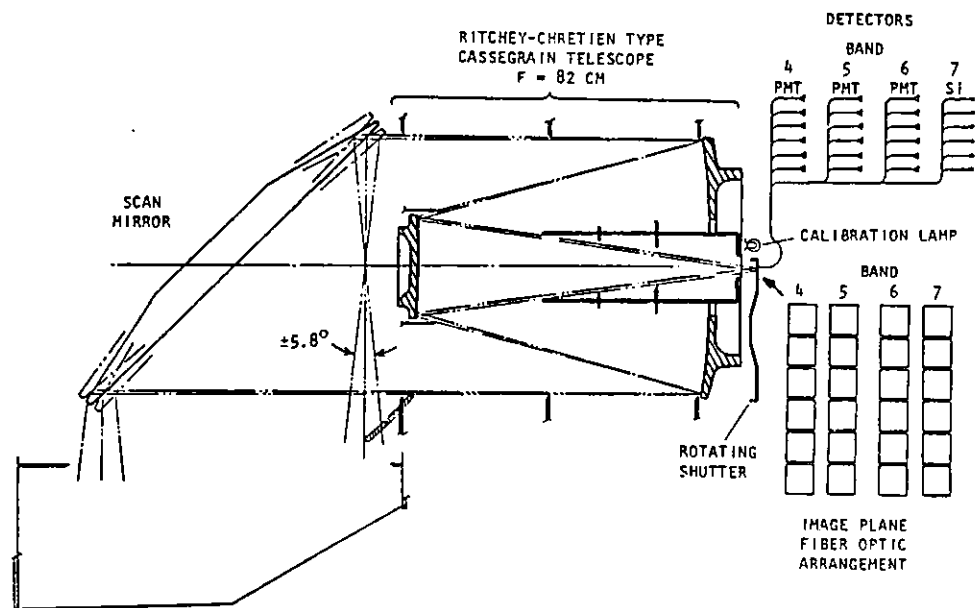


Figure 1.6 : Landsat MultiSpectral Scanner (MSS)

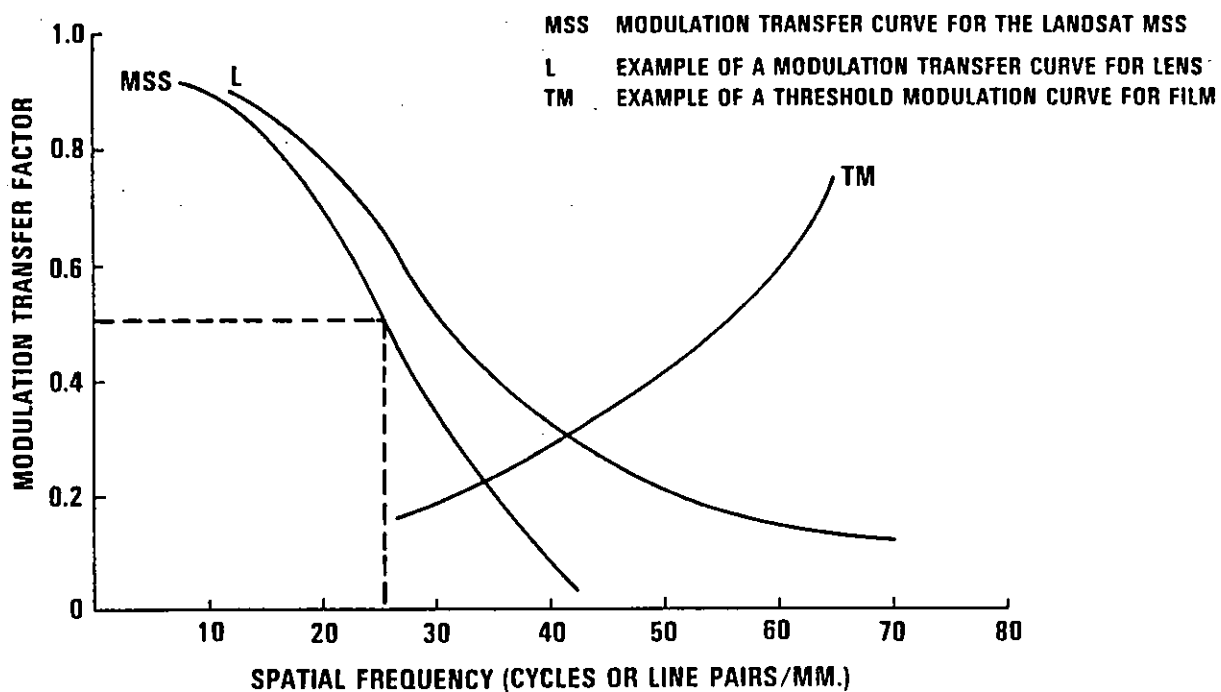


Figure 1.7 : Spatial frequency response of the MSS

increase in the detector size does not incur a proportionate loss of resolution and since the signal to noise ratio for many detectors depends on the square of the detector size this may be advantageous. As an example an increase of detector size from one to two Airy diameters in a system with $e = 0.4$ gives a 30% loss in resolution and a 200% increase in S/N. This type of resolution calculation is ideally suited to an astronomical observer dealing with point objects but the situation becomes much more complex when the targets are not point-like and are surrounded by background features, and the computation of image energy spread attracts many workers [13].

If the above approach is used in the spatial frequency domain as by application of equation (1.4) then an MTF approach to resolution may be used. The exact criterion used to define the effective IFOV (EIFOV) is arbitrary but it is common to specify the spatial frequency at which the modulation transfer ratio T is 50%, where

$$T = M(\text{image}) / M(\text{object}) \quad (1.8)$$

for contrast modulations defined as

$$M = (I_{\max} - I_{\min}) / (I_{\max} + I_{\min}) \quad (1.9)$$

The parameters I are the image or object intensities.

It is useful to consider the MTF of the recording and viewing systems and to plot the MTF together with the minimum modulation required to give a detectable response (the Threshold Response Function or TRF) and the point of intersection of the two curves defines the limiting spatial frequency of the system [14] as in figure 1.7.

The main purpose of multi-spectral scanners such as the Landsat MSS is the monitoring of resources by their spectral signatures and the classification of ground cover types as described briefly in section 1.2. To this end perhaps the most relevant measure of resolution is the minimum size of target required for accurate classification. The resolution is quoted as the minimum size of target for which an intensity measurement is within a stated tolerance of the correct value for a certain proportion of the readings e.g. within 5% for 95% of the readings [15]. Often greater classification accuracy can be obtained by the

use of lower resolution imagery as nonuniformities in a required cover type are obscured (e.g. individual trees or clearings in a "woodland" cover type) but this "scene noise" depends on the number of classes required. The accuracy of an intensity reading depends on how much of an IFOV is filled by the desired target and on the contrast between the target and the surround, and for this reason the target must be at least a few pixels on a side.

The data from a scanner is primarily in the form of an analogue signal and may be recorded as such but often it is digitised for ease of transmission, and the resolution of the digitisation will affect the output data rate. In deciding the number of bits per sample it is useful to estimate the signal to noise ratio of the data and a brief example of this is presented below for the use of photodiode and PMT detectors [16]. Both these detectors are photon or quantum sensitive as opposed to a thermal or absorption detector (thermopile) and are high speed and sensitive. They differ in their noise characteristics in that the PMT is dominated by photoelectron shot noise whereas the diode shows thermal, shot and dark current noise. The noise currents may be written as follows:

$$I_n \text{ (PMT)} = (2N^* a^2 A_o t_o \bar{R} k_e e \Delta f_n)^{1/2} \quad (1.10)$$

$$I_n \text{ (PD)} = [(i_{ns}^2 + i_{na}^2 + i_{np}^2 + 4kT/R_L) \Delta f_n]^{1/2} \quad (1.11)$$

and the signal current is the same for both :

$$I_s \text{ (PMT/ PD)} = (N^* - N') F(x, a) a^2 A_o t_o \bar{R} \quad (1.12)$$

where	N^*, N'	are the apparent radiances of a bar test pattern in the optical pass band of the system
	$F(x,a)$	is the spatial frequency response as a function of the bar width x and IFOV a
	A_o	is the entrance aperture area
	t_o	is the optical efficiency including obscuration
	$\bar{R}$	is the mean detector responsivity over the optical pass band
	k_e	is the noise multiplication facor from the electron multiplier of the PMT
	e	is the electron charge

Δf_n	is the effective noise bandwidth
i_{ns}	is the diode signal shot noise current
i_{na}	is the dark current noise
i_{np}	is the preamplifier noise current
$4kT/R_L$	is the load resistor thermal noise current

The effective bandwidth can be written as

$$\Delta f_n = k_f (v/h) \theta / 2k_s n a^2 \quad (1.13)$$

where k_f	is the ratio of the filter effective noise bandwidth to the information bandwidth
v, h	are the scanner velocity and height
θ	is the across-track swath width
k_s	is the scanning efficiency factor
n	is the number of parallel detectors per channel

The PMT signal current is defined as the difference in the number of photoelectrons emitted when viewing the bright and dark bars of the test pattern, and the noise is the square root of the number of photoelectrons emitted from a single IFOV of background (dark bars) from Poisson statistics. In the case of the PD the signal current is the same but the noise is comprised of several factors as shown in equation (1.11). Since the scanner bandwidth is fixed then R_L is fixed and usually small, thus dominating the noise together with preamplifier noise. At low currents dark current and shot noise may be neglected and the PD noise may be rewritten as

$$I_n (PD) = i_n \Delta f_n^{1/2} \quad (1.14)$$

where i_n is the root sum square of thermal and preamplifier noise components. The signal to noise ratios for the two cases are found by dividing equation (1.12) by equations (1.10) and (1.11), and if the aperture area is replaced by an expression in terms of the diameter D the S/N ratios have the forms

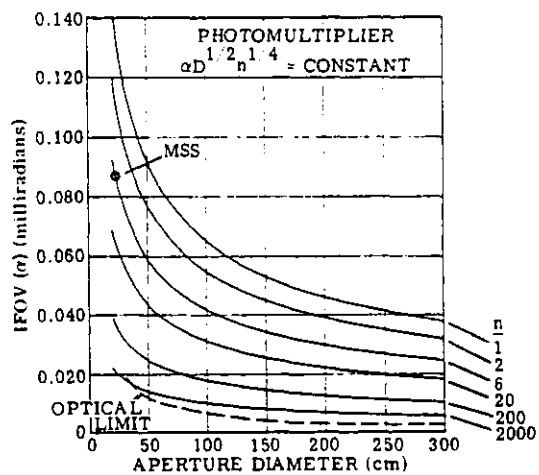
$$SNR (PMT) = a D^{1/2} n^{1/4} K_{PMT} () \quad (1.15)$$

$$SNR (PD) = a D^{2/3} n^{1/6} K_{PD} () \quad (1.16)$$

where K_{PMT} and K_{PD} are functions of the other parameters involved and may be found explicitly by the above method. These equations serve to show the effects of altering the major system parameters on the signal quality for any general scanner. In both cases a decrease in angular IFOV causes data rate [17] and stability problems but also requires an increase in D which in turn increases the system mass and bulk. Figure 1.8 (a) and (b) show the relationships graphically. Another factor which may cause problems is the optical quality of the telescope which must give diffraction limited images over an increasingly wide range of wavelengths. If the equations (1.15) and (1.16) are rearranged to include an explicit dependence on bandwidth it is found that in both cases the SNR scales as the square root of the bandwidth. An increasing requirement is being generated for fine multi-band spectral resolution as in the case of the "Daedalus" airborne IRLS and the Thematic Mapper, and if quality is to be maintained then improvements in D , n and optics are required as demonstrated in figure 1.8 (c). The future of scanner devices probably lies with multi-element detector arrays and very high performance data processing systems.

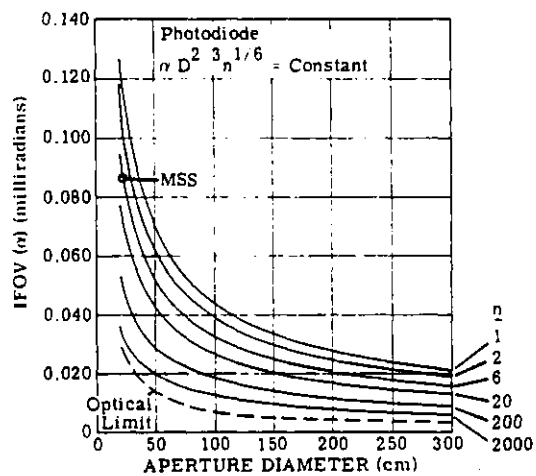
A final area to be considered in image acquisition systems is image conversion. It is often inconvenient to detect directly the desired image information due to a lack of detectors or poor detector characteristics and in this case some form of conversion is required. One of the most important areas in which image conversion can be used is in medical X-rays where conventional film methods are slow and require relatively high dosages. It is possible to use very low levels of X-rays in real time using a MicroChannel Plate (MCP) converter/ intensifier and a TV camera. The MCP acts in the same way as a PMT but is able to preserve spatial information, and amplifies a photocurrent from a photocathode. These devices may be very small [18] and are capable of a 35% efficiency. Other applications include the detection of particles from accelerators and image intensification.

This section has attempted to outline some of the major forms of image acquisition device currently in use and to show how general performance parameters may be derived. Great care must be exercised in the choice of system and both the purpose of the imagery and the method of processing must be carefully considered. A particularly confusing area is that of resolution where the number of quoted values for any system is proportional to the number of users. The future trend in image acquisition lies in electronically scanned systems such as CCD's and promises large volumes of interesting data.



Aperture diameter required to achieve increased resolution with scanners employing PMT's (relative to ERTS-1 MSS).

Figure 1.8 a



Aperture diameter versus resolution for scanners employing photodiode detectors (relative to MSS).

Figure 1.8 b

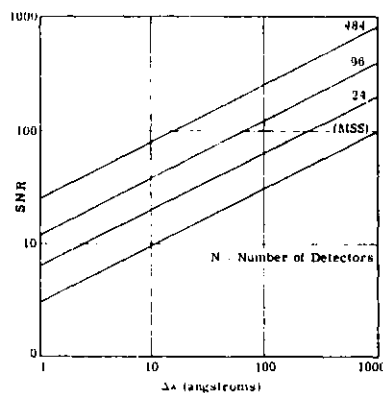


Fig. 12. Calculated SNR versus $\Delta\lambda$ for various values of n with MSS-type scanner using PMT's.

Figure 1.8 c

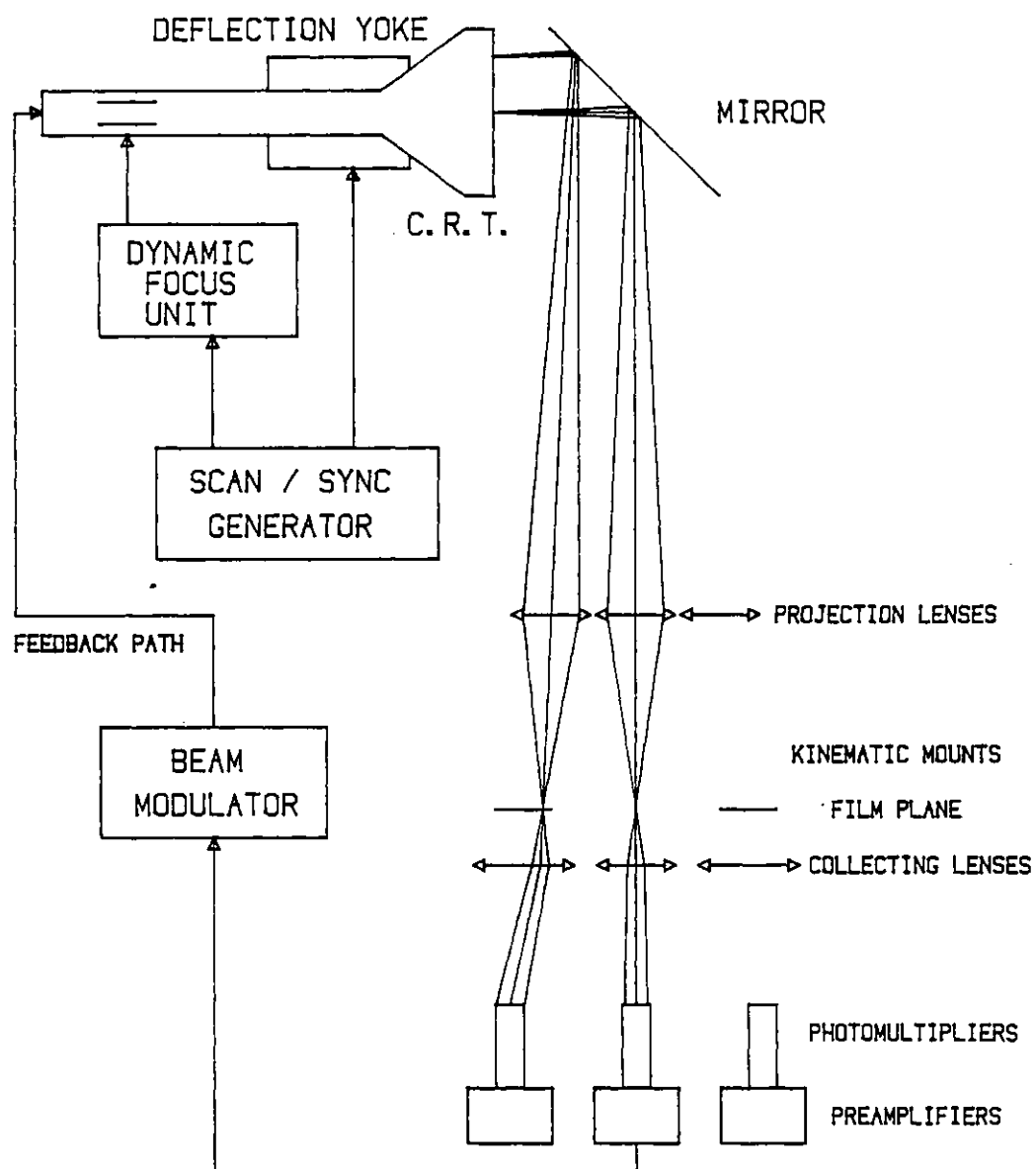


Figure 1.9 : Flying Spot Scanner (FSS)

1.4 Storage and display systems

The output from various scanners or processing systems must be capable of easy storage and viewing for interpretation, and this requires the use of specialised archiving and display systems. Memory devices and computer-related magnetic storage systems are becoming cheaper and more capacious and are relieving the pressure on the image user but it must be remembered that the amount of data in even small images is vast as noted in section 1.1: a Landsat MSS image in four spectral bands occupies 35 megabytes of storage or 1200 feet of high density Computer Compatible Tape (CCT) and has 3780 by 2286 pixels.

There are two main technologies in use for long term image storage, these being photographic and magnetic. The desirable characteristics of any storage system are high capacity, low cost, high density and ease of storage/ retrieval. Photographic negatives fulfil all these criteria except for ease of storage where development is involved. Retrieval may be accomplished simply by viewing the transparency or negative and colour images may be used with suitable film. A disadvantage of film if it is to be used to store images from a scanner which are in the form of electronic signals is the need for a film writing device. The same problem arises if electronic signals are required from a film image and a reading device must be employed to perform the transformation. This conversion of image format may be accomplished by using a scanner such as a TV camera as in the previous section with suitably designed optics, but it is more usual to employ a special purpose system. One such system is the Flying Spot Scanner (FSS) which is commonly used to screen films in commercial broadcasting (telecine machine). A diagram of a simple FSS appears in figure 1.9. The output is in the form of a TV standard video signal and is generated as follows: a blank (unmodulated) raster is synthesised electronically and displayed on a flat-face precision Cathode Ray Tube (CRT) similar to an oscilloscope screen. This raster, which appears as a uniformly lit area on the CRT, is actually formed from a small (gaussian profile) "flying" spot of light, and if the CRT screen is imaged onto the film this spot will scan the image in the same way as a TV camera. The amount of light passing through the film is proportional to the film transmission and the spot brightness, and a collecting lens is used to direct the light onto a photomultiplier tube. The final signal current then carries the image information and after buffering it has the synchronisation pulses added from the scan generator unit. The CRT used must have several special characteristics [19] : the phosphor must have a short decay time i.e.

the light output must drop sharply after the electron beam has moved away from any point. This is necessary to prevent blurring of the picture since, once again, the output is the result of convolving the image with the spot profile. The spot size should be gaussian to introduce the least possible filtering and should be sufficiently fine to resolve the expected image detail. The light output should be enough to 'clear' the darkest regions of the image.

Selective scanning of subimages is easily accomplished by changing the scale and position of the raster on the CRT electronically but care must be exercised when operating with small rasters for extended periods as phosphor ageing may occur. This effect results in a reduced light output over an area of phosphor which has been subjected to excessive beam current density. Typically a CRT of this type is operated at 25 to 30 kV and 30 to 120 μ A beam current. A tube with an aged phosphor is permanently damaged and will give a patchy image signal. To overcome this and any natural nonuniformity across the CRT screen it is good practice to use a feed-back loop in the system which monitors the CRT light output via a beam splitter and modulates the beam current in such a way as to maintain a constant light output. The basic FSS may be extended to work with several channels of multispectral data if more lenses are used, each one off axis, and provision is made for kinematic film mounts to allow registration. Nonuniformities in the photocathode are avoided by arranging the collecting lenses so that they image the projection lenses onto the PMTs: this ensures that the CRT spot output is spread over the entire photocathode surface at all times. The same FSS system may be converted to write images if a video signal is used to modulate the raster and an unexposed film is used. The PMTs and collecting lenses are then redundant but the feedback loop can still be employed to ensure a faithful output. The spot size of a typical CRT may be 50 μ m or less and the best resolution is about 1000 pixels square even when the demagnification by the projection lenses is included. This is only just adequate for small images especially when the Kell factor (section 1.3) is taken into account.

The FSS described above employed electrical scanning in both directions which is convenient for control purposes but has the disadvantage of being bulky and of using high voltages. An alternative method for reading film is to use a CCD linear diode array to scan lines across the film and a precision mechanical motion to provide the vertical scan either by moving the film or by moving an image of it. Diode arrays are available with more than 2000 elements thus providing high resolution and are relatively simple to operate using ordinary

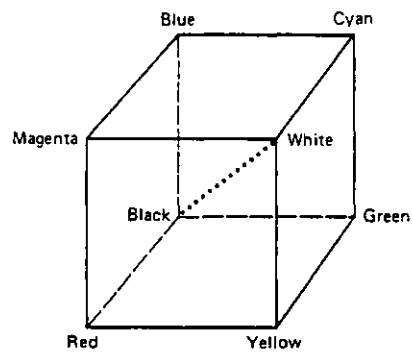
digital and analogue techniques. These devices also exhibit nonuniformities in response along their length which must be compensated for if accurate data is to be gathered. This is most easily accomplished by storing calibration data for each diode of the array in a solid state memory and performing the correction immediately as each diode is sampled. Two dimensional arrays are also available but they offer limited resolution and correction becomes more difficult as the number of diode elements increases.

Recording of image data on film may be performed in a variety of ways if the source is an electrical signal rather than a real optical image. One method is to wrap unexposed film around a cylinder which is made to spin at a controlled rate. A small glow tube or Light Emitting Diode (LED) is used as a modulated light source and a spot of light is projected onto the film with a microscope objective. Scanning in the vertical direction is accomplished by moving the objective/ source assembly along a precision leadscrew. This type of filmwriter is often used in the graphic arts trade and is the machine used to write all Landsat images in RAE Farnborough. In the all but the most recent machines the light source is a glow tube or gas discharge source which suffers from short lifetime and a decaying response with age. Some modern systems use laser diode sources which are powerful and fast, thus improving writing speeds. An alternative method is to use a modulated and deflected laser source such as a He-Ne gas laser or Argon ion laser. Modulation may be accomplished by means of a Pockels or Faraday cell and deflection by means of an acousto-optic deflector. The first two cells utilise the electro-optic and magneto-optic effects [20] respectively to change the plane of polarisation of a light beam according to a modulating voltage or magnetic field. The deflector relies on diffraction from an ultrasonically generated grating in a liquid or solid medium. A piezo-electric transducer can be used to set up a standing compression wave in the cell which modulates the refractive index and thus produces a phase grating whose spacing is proportional to the frequency applied. Such deflection systems suffer losses due to diffraction into other orders and low efficiency but they are faster and more easily controlled than moving mirror systems. An example of a commercial film writing device is detailed in [21].

If image data is available in electrical form it may be stored on a magnetic medium such as a tape or disk. A familiar example of this is the domestic video recorder which records analogue signals from a TV camera in the form of scan lines. Similar systems are used to store data from commercial

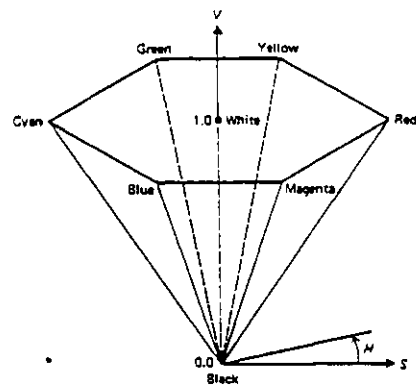
scanners such as the IRLS but the analogue tape has the disadvantages of noise and loss of fidelity with use due to abrasion of the magnetic layers. For this reason, and also because many processing systems are digital, it is common practice to record High Density Digital Tapes (HDDTs) which have a much greater noise immunity. The format of the data on such tapes is often dependent on the particular system, and computer processors require a standard format CCT. The information density on a CCT is commonly 1600 bytes per inch of tape although modern systems can run in excess of 6250 bytes per inch. Tape systems tend to be slow in retrieval and the serial or line structure of the data make the reading of subscenes slow and complex, and for this reason many computer processor systems utilise hard disk storage which offers fast and random access to the data with capacities up to Gigabytes.

Image display is dominated by the CRT although there have been some recent developments which use electro-optic Spatial Light Modulators (SLMs) in TV projection systems [22]. CRT's can be used to display line drawings using vector display generators which can move a bright spot in any required direction to draw lines and specialised tubes have been developed which can store the images produced as a charge pattern on a target [23] as in the case of Tektronix™ displays. The most common usage of CRTs is for the display of continuous tone raster scanned images such as TV pictures. The simplest form is a monochrome monitor where image data is presented as a voltage signal for each scan line and modulates the screen brightness accordingly. Often, particularly in the case of multispectral imagery, it is desirable to use a colour display and a special CRT tube is required. Many forms of colour tube exist but they all share two common features: firstly they display any colour by a suitable combination of three primary colour phosphors and secondly they all use a spatial multiplexing technique to combine the three coloured images. A full treatment of colour theory in TV displays may be found in [24]. The essential result is that most colours may be visually matched by a suitable combination of three "primary" colours, red, green and blue. The intensities of the three primaries can be used to define a coordinate system, referred to as RGB colour space, (as for feature space in section 1.2) which is cartesian and bounded as shown in figure 1.10 (a). Note that all colours along the main diagonal are different intensities of white. If the colour of a point is separated from its intensity by requiring that the components of R, G and B are normalised then another colour coordinate system results- Hue Lightness Saturation space (HLS) [25] or the similar Hue Saturation Value (HSV) space shown in figure 1.10 (b)



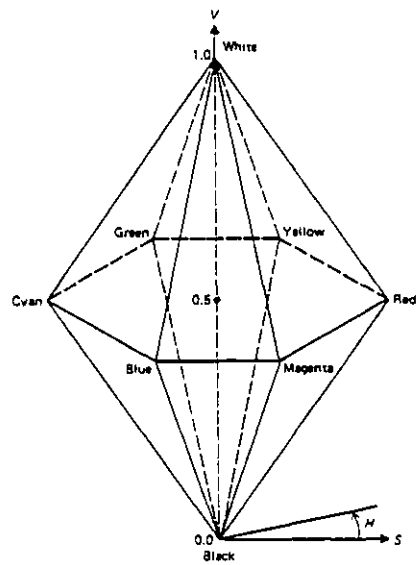
RGB color cube. Grays are on the dotted main diagonal.

Figure 1.10 a



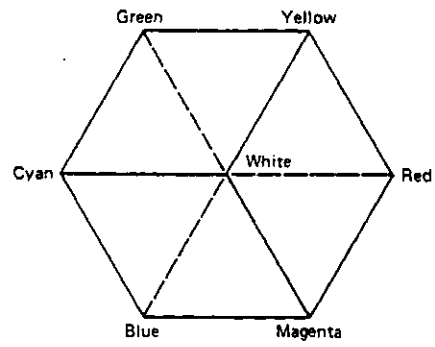
Single hexcone HSV color model.

Figure 1.10 b



Double hexcone HLS color model.

Figure 1.10 c



RGB color cube viewed along principal diagonal. Visible edges of cube are solid, while the invisible ones are dashed.

Figure 1.10 d

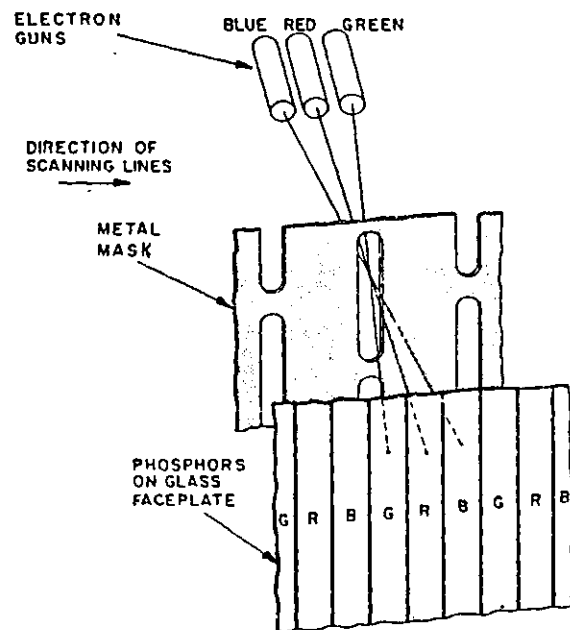


Figure 1.11 : TRINITRON colour TV display tube

and (c) respectively. These latter systems are visually oriented in that they match the properties of the eye as opposed to the RGB system which is convenient from the hardware point of view. Hue is the perceived colour of an object, saturation is the purity or amount of white diluting a colour and lightness or value is the intensity or brightness. The relation between RGB and HSV can be explained by noting that the top surface of the HSV hexcone has the same appearance as the RGB cube when viewed along its leading diagonal as shown in figure 1.10 (d). The usefulness of the HSV system is described in [26].

The system used to combine three coloured images on a single screen employs a spatially selective electron beam system. The three phosphors are deposited on the screen in adjacent stripes and bombarded by three separate electron beams through an aperture grille or shadow mask as shown in figure 1.11. The three electron beams come from a single gun and are horizontally aligned. They are focussed through a single electron lens onto the screen but the beam directions are such as to prevent any electron beam from exciting the wrong phosphors i.e. the "red" beam always strikes red phosphor. Each slit in the aperture grille has three phosphor stripes behind it and all three beams pass through simultaneously. This type of tube, the Trinitron™ (developed by Sony), suffers from two shortcomings: the spatial multiplexing causes a loss in resolution- the minimum phosphor spacing being 0.2mm at present and 0.6mm usually- and the aperture grille causes a loss of electrons and consequent dimming of the picture. The resolution is being improved but the tubes are expensive, and the efficiency is being combatted by the use of a Beam Indexing Tube in which a single electron beam carries all three images in a Time Division Multiplexed (TDM) format i.e. the beam carries red information while it is pointing at red phosphor. This tube requires efficient switching and careful alignment.

Other display technologies are becoming available including liquid crystal and plasma panel systems and integrated LED arrays may arise soon if the problems of manufacturing blue sources are overcome. Such devices may offer a smaller and lower power alternative to the CRT but they will have to be very high quality and low cost to displace it entirely. Many of these devices employ arrays of separately addressed display cells or pixels and they suffer problems of power dissipation, physical space for signal lines and uniformity of pixel responses which are the subjects of current research.

1.5 Image processing systems

Image processing systems have the curious property of inspiring a high degree of chauvinism in their users, particularly when comparisons are attempted. It is common to distinguish processing systems into two classes: analogue or digital, but a better distinction can be made on the basis of hardware or software processing. If an image can be acquired in suitable form it may be fed into a general purpose computer and processed as an array of digital numbers. The computer can perform complex operations many times but it is limited in the speed of calculation by its very generality of purpose. It is also limited by the size of its memory which, although modern devices are falling in price, must be very large even for small images. Some relief from this may be afforded by the use of virtual memory systems such as the VAX series from Digital Equipment Corporation (DEC) in which a fast disk unit is used to swap blocks of data rapidly (page faulting) to give the appearance of continuous memory at the cost of disk access times. A computer-based system needs only a display unit consisting of a CRT with an associated screen-mapped memory or framestore and a bulk storage device: all operations are then coded in a suitable language such as FORTRAN, C or PASCAL. The alternative is to construct a special purpose system which is less versatile but is capable of higher speeds in its chosen tasks than a computer. It is in this area that the analogue/digital distinction becomes more relevant. Optical techniques can be used to perform such functions as Fourier transformation and spatial filtering with an analogue image as the input (a transparency). Analogue techniques may also be applied electrically to video signals in Closed Circuit TV (CCTV) systems [27], where the operations are essentially serial and point-invariant e.g. density slicing, false colour compositing, multispectral classification. Digital techniques may be applied in a similar fashion, but the use of a digital picture memory can allow some local operator functions such as edge detection or smoothing. The current trend is towards a mixture of the two main philosophies in which purpose-built hardware is supervised and controlled by distributed computing power in the form of fast microprocessors which are in turn supervised by a host computer. The main aim of any such system must be rapid processing to allow user interaction because the most important part of any processing system is the operator/ interpreter. It is in this field that the general purpose computer tends to be at a disadvantage. The main features of some current systems are described below:

The heart of any processing system is an image memory. This may be a semiconductor digital memory or a transparency image. Often the display is memory-mapped, i.e. a well defined relationship exists between the position of a point on the display and the address in the memory, and such a system is referred to as refresh memory since its contents are continually accessed to 'refresh' the display. A typical size of memory is 512 by 512 pixels with an 8 bit binary word (byte) of intensity data for each pixel, giving a total of a quarter megabyte (262,144 bytes, $= 2^{18}$) which is convenient for a TV display. Some very high resolution displays can show 1280 by 1024 pixels, but each pixel only lasts for 18 nS and this is approaching the limit of digital logic speed. If colour displays are required then several of these framestores must be used, and they can also serve as a temporary workspace by the system computer. Input to these memories may come synchronously with their display accessing from a TV camera or asynchronously from the computer. Many systems use several micro-processing units around the main memory, a popular choice being the DEC LSI 11 chip set with both permanent software in Read Only Memory (ROM) and spare Random Access Memory (RAM) to accept downloaded software from a host computer. Another set of memory is usually provided for the display of graphical information as line drawings or graphs and this is termed 'Overlay' memory. Users interact with the system host computer via a range of input devices and via a Visual Display Unit (VDU) or the image display overlays. Some form of special purpose hardware is usually provided for operations such as pixel addition or logical functions and this is sometimes designed to work with the same data stream which is being sent to the display. Software is run on the main computer which is capable of handling the user interaction and controlling the microprocessors to accomplish high level operations. Functions supported fall into several groups [28] such as input/ output, arithmetic (image addition, division etc.), geometrical correction, transformation (Fourier, Hadamard, Cosine) and related operations (spatial filtering, smoothing, edge detection, convolution), image classification and statistical operations such as grey level histogram display. Examples of systems like this are the GEMS [29] processor from the Computer Aided Design Centre, Cambridge and the I²S system (International Imaging Systems), in California. These systems require a host computer (IBM PRIME 750, VAX 11/750 or 780 and PDP 11/ series are popular choices) and feature extensive processing software but many of the higher functions require the intervention of the host computer and most require the microprocessor, so that such processors are not as fast as a purely hardwired machine, although the complexity of new hardware devices is rapidly increasing.

A different type of system is the Multichannel Electronic Image Classifier (MEIC) Mk 1 and 2 which have been developed at Imperial College, London for use with Landsat and IRLS imagery in transparency form. These systems use the transparency as a refresh memory and data is continuously read out by a FSS as described in section 1.4. The systems use analogue electronics to process the image video signals in an essentially serial manner, and such operations as addition, division/ multiplication, mixing, linear contrast enhancement, additive colour viewing and classification are possible. Geometric correction and zoom/ pan operations can be achieved by a suitable distortion and scaling of the raster in the FSS and this can be controlled digitally. Many of the controls of the system are directly available to the user and are instantaneous in their effect, since they operate on the image data as it is displayed, thus providing great ease of use and instant response to user ideas. In the Mk 2 version a small microprocessor handles a user operation console, but has only limited participation in the processing operations in contrast to the systems described above. A full description of the design, capabilities and use of the MEIC systems is given in [27] and [31], but a comparison of one important operation performed on the MEIC and a system such as the GEMS is useful here. The classification of ground cover types in Landsat multispectral data is very easily performed on the MEIC. A simple x-y display allows the video signals from any two bands to be plotted against each other, giving a two dimensional projection of feature space as in section 1.2. Clustering can be observed directly and instantaneously for any image and the location of specific cover types in feature space can be achieved by the use of a cursor on the image and a corresponding bright-up in feature space. The operator has two 'window' level controls for each band and the display can be coloured to indicate if the four video voltages associated with any point on the image lie within their respective windows simultaneously. The user selects a training area and adjusts the controls until it is just picked out by the colour: any other points on the image which are also coloured show a similar spectral signature and may be of the same cover type. A box on the feature space display corresponding to the window levels shows the extent of the classified areas in feature space.

The same operation on a digital system requires the operator to choose a training area and the computer reads out the four intensity values for each pixel in this area. The microprocessor must then access each pixel in the image and test its values against those found in the training area. If a match is found then a coloured point is set up in a graphics overlay memory. A feature

space display must be acquired by building up a histogram in two dimensions, storing the data in a framestore after suitable scaling and displaying the contents on the screen. This process involves much memory access and processing which is relatively slow and the feature space display cannot be viewed at the same time as the image. Another feature is that only a finite number of pixel values are picked out, and they are chosen from a square box in the image: this leads to errors which the MEIC avoids by allowing direct access to the window levels.

The 'optimum system' is a nebulous entity whose form depends upon the person considering it, but it will probably include a large amount of semiconductor storage since these devices are rapidly falling in cost, and this will be surrounded by several powerful microprocessors. Special purpose hardware will be designed for very fast computations of specific types, and the entire system will be highly user interactive. Another part of the system has yet to be fully designed, and this is the 'intelligent operator'. At present the user community is dominated by two species: those who have an application which they think may be usefully subjected to image processing and those who design and build the systems and software and understand the various operations. It is rare to find a user who combines these qualities, and interaction between the two camps is to be encouraged. One of the most important parts of any system is the development of an operational strategy, but there is a great danger, especially in commercial environments, for image processing to be regarded as a source of confirmation rather than hard results, and to regard any results from a shiny new image processor as true without interpretation. In so complex a field this is perhaps to be expected, and must be guarded against.

1.6 Precis

It is hoped that this section will serve to indicate the form and contents of this thesis and to whet the potential reader's appetite! The contents are divided into two sections, the first of which presents a study of noise from film transparencies which are used to store Landsat images and form the input data to the MEIC system described above. The presence of noise in an analogue system is often cited as the main disadvantage of such systems in comparison to digital processors. This is not necessarily the case however since it is now possible to design very low noise circuitry using modern integrated circuits.

It must be remembered that the act of digitising an image (or any signal) itself introduces a noise signal: if the quantisation interval is d , then a random error of $\pm d/2$ maximum is introduced, and if the mean square error is calculated it is found that it has a value of $d^2/12$. This noise is introduced at the start of the process and is additional to any noise in the signal due to the detector or scanner. The author's limited experience with users of image processing equipment prompts the following observation: if two images with equal resolution are presented, but one of them is blurred in the fine detail while the other shows distinct pixels, then some form of psychological effect appears to render the latter the more acceptable in terms of 'quality'.

The problems addressed in chapter 2 are concerned with the quantification of the noise due to 'grain' and its effects on the resolution of the classifier of the MEIC. An attempt is made to describe the information content of a transparency and the results of a statistical analysis of film samples are presented which show how the noise varies with the signal recorded. The area of information content is also approached from a digital viewpoint in which the number of significant bits required adequately to represent pixel intensities without visible loss of quality is assessed.

The second main section deals with the design and construction of fast image processing systems following the philosophy outlined in the previous section of using hardwired electronics wherever possible. Other constraints placed on this work are high speed, user friendliness and low cost. The systems described are designed to perform statistically-based operations on images for enhancement purposes, both as a precursor to further processing and as an aid to the interpretation of results. Chapter 3 describes the design and development of a contrast modification system which has been tentatively named CONDOR 1 (Contrast Optimisation by Nonlinear Digital Operations in Real time). This system is capable of acquiring the statistics of an image and modifying them to any desired form. It is of use in enhancing low contrast images such as X-ray shadowgraphs and in modifying input signals so that they make the best possible use of the available image range. Details of the operating software which has been written to run the system are also given and some examples of its use are provided.

The CONDOR system works with real time video signals from a TV camera or FSS, and this is also true of the second system described in this section, but as

it is intended to be the basis for more complex operations in the future, and for convenience of operation, digital picture memories have been constructed to store one four-channel multispectral image. The object of the system is to investigate the use of data compression by means of the Hotelling transform also known as the Karhunen-Loeve Transform (KLT) or Principal Components Analysis (PCA). Modern scanner systems are proposed which will give many channels of multispectral data, and this makes the conception and manipulation of feature space difficult. Landsat's four bands give 6 two-dimensional projections of feature space which must be assimilated by the operator for classification, and a seven channel system as carried by the new Landsat 4 gives 21 projections which are a daunting prospect for even the experienced operator. In practice it is found that the data from a multispectral scanner is often highly correlated i.e. bright areas in one spectral band are bright in others, especially if topographic shading occurs in the image. The PCA aims to produce decorrelated images by a suitable combination of the input images. The feature space distribution is analysed and directions of maximum variance are selected with the constraint that they are orthogonal and the data is projected onto these new axes. Although the same number of images result from this process, it is found that most of the interesting information is contained in only a few, and the others may be discarded without great loss of information. The system described in this thesis is designed to perform the analysis of the feature space statistics and the transformation with the aid of special-purpose circuitry using a combination of analogue and digital techniques, and to provide a significant speed enhancement over the more usual software approach found in many processors. Some examples of the process are presented, including some colour composite images formed from the PC images.

Throughout the work the emphasis has been on speed and ease of use, but since all the circuitry and software has been designed and constructed by the author, insufficient time has been allowed for the important "strategy-forming" stage outlined in the previous section. Further work would enable the full potential of these systems to be realised. It must be remembered that this thesis can only hope to cover a small area of the field, and that "To explain all nature is too difficult a task for any one man or even for any one age. 'Tis much better to do a little with certainty, and leave the rest for others that come after you, than to explain all things" [30].

2.1 Objectives

This chapter examines the effects encountered when a photographic transparency is used as an image store. An example of a processing system employing film as a picture memory is the MEIC, described in section 1.4, which uses a flying spot scanner to produce electronic signals for manipulation. Chapter 1 hinted at some of the difficulties associated with photographic processes and this chapter aims to describe the noise properties of film and the way in which they affect subsequent image processing operations, in particular that of box classification in multispectral feature space. The fundamental questions which are posed in this study concern the amount of useful data which can be stored and retrieved from film when it is used as an analogue image memory. This is an important distinction since film can be used as a digital store with potentially high capacity, but the data format does not allow immediate visual interpretation of a picture. When used as a computer memory, the film is divided into small regions which are either exposed or not according to the data which is in binary form. This gives the data a high degree of noise immunity in the same way as electronic logic circuits which operate with two distinct voltage levels, and which allow considerable variation of the signal before an error occurs.

In recording an analogue signal a continuous grey level or transparency range is used, but with this comes a much lower tolerance to noise signals. In order to preserve the same error probability as a binary signal of the same data rate the power in the multilevel signal must be much greater relative to the noise, and scales as the square of the number of levels used. An obvious area for investigation is the exact number of levels available in a film image, a topic which is often glossed over. This chapter will present numerical estimates for the number of grey levels in a film, but these numbers must never be used out of the context of the accompanying conditions and assumptions. The number depends on the exact definition of a 'distinct grey level', and two main approaches have been taken in this work. A major operation in the use of Landsat imagery is multispectral box classification (section 1.2) and the effect of noise is obviously to degrade the accuracy of results. It has been shown [31] that in certain circumstances a noise signal can serve to indicate fine

differences in classified ground cover types. Although various processes (e.g. ratioing) can be employed in an attempt to improve classifications, noise is still a problem, and should be well understood by any potential user. The first approach taken in this work is to specify an acceptable error tolerance (considering only the noise in the film and not sensor or ground induced errors) and to calculate from measured noise statistics how many levels are available within the film's dynamic range. The results of statistical measurements on samples of Landsat 70 mm transparencies are presented which indicate the variance of the noise signal and its dependence on the recorded density level.

The second approach has been to ask how large a memory is required to store an image without significant loss of fidelity. Since modern semiconductor memories store digital pictures with quantised intensity values, this is equivalent to determining the required quantisation level, or the number of binary bits per pixel. The section describing this method introduces the basic architecture of the framestores whose construction is described in a later chapter, and the important concept of the Bit Plane image. Briefly, this is a black and white (i.e. no grey) picture formed by constructing a digital array from the different binary digits or 'Bits' of each pixel's digital intensity word. This process is equivalent to performing a contouring operation on the image with 1, 2, 4, 8, etc. contours equally placed throughout the intensity range. The important observation is that as successively less significant digits or bits are displayed, the picture becomes more random, until no visible structure exists. Further bits are insignificant in their contribution to the image data and need not be stored. The number of grey levels is then equal to two raised to the power of the number of bits per pixel. This work accordingly seeks to measure the structure in these bit plane images by a study of their auto- and cross-correlation functions.

The numbers of levels given by these two approaches do not coincide exactly, but this is to be expected since different criteria are used to define a separable grey level with different aims or applications in mind: the more structural or bit plane approach is not necessarily well suited to the needs of the image classifier. Indeed, as with other branches of the field of image processing, confusion and ready argument abound when discussing the information content of pictures, and several definitions exist each with its own followers. It must always be remembered that the photographic process is very complex and variable, and that although many workers have devoted much time to its study a

complete understanding of film is extremely difficult to attain.

2.2 The photographic process

The photographic emulsion is one of the most sensitive and versatile image recording media available. The photosensitive 'detectors' are small grains or crystals of a silver halide compound (usually a bromide / iodide mixture) which vary in size down to a few hundred Angstroms in fine-grained film. These crystals are dispersed in a matrix of gelatin which is coated onto a prepared substrate such as glass or acetate film, thus forming a thin, randomly scattered layer. The crystals are cubic, and are ionic lattices, but their sensitivity arises from the presence of impurities and imperfections: if a photon of sufficient energy (about 2.5 eV) strikes a grain and is absorbed it can liberate a photoelectron which may in turn be trapped by one of the defects. The photoelectron is raised to the conduction band of the crystal along with some thermally excited electrons but the defects provide trapping energy levels just below the conduction band which temporarily bind the electron. Silver ions which are present in the lattice are attracted and combine with the electron to form silver atoms. When three or four silver atoms have been formed at the defect site the grain has been exposed and a stable 'latent' image is formed which is revealed by development. The developing process is a chemical reaction in which the emulsion is treated with an organic reducing agent which converts the halide grains into silver metal. All the grains will be reduced given sufficient time but those which have been successfully exposed will be reduced faster, giving a pattern of black silver metal in the emulsion. The excess halide is removed by the fixing agent to render the emulsion insensitive to further exposure. Note that each grain is either exposed and reduced or removed by the fixer, thus acting as a two-state detector, and grey areas are produced by the combined effects of many grains. The development process gives a large amplification of the latent image, from a few silver atoms to a whole grain, and therein lies the great sensitivity of the process, with an increase in the number of silver atoms of the order of 10^9 . The density of a point on a film may be affected by the exposure of immediately adjacent areas, giving rise to distortions, especially at edges and in small dots. These adjacency effects are due to variations in development conditions across the film surface caused by local depletion of developer by highly exposed areas and the production of inhibitory by-products. The characteristic curve for large areas of film is

then no longer valid for small objects, and will affect the modulation transfer function of the film. These degradations may be minimised by the use of suitable developers and a high degree of agitation during processing to ensure uniform distribution of the chemicals.

It must be remembered that the image is composed of a random array of opaque grains which have a range of sizes and sensitivities so that some noise is inevitable. The exact form of the characteristic or H.D. curve (section 1.3) will be influenced by these parameters, and a full treatment of the various models and results may be found in [32]. The point to be emphasised here is that the characteristics of photographic emulsions vary widely and require carefully controlled standard conditions for their manufacture and use if measurements are to be meaningful. Most photographic emulsions have a useful dynamic range of around 2 density units (i.e. down to 1% transmission), and sensitivity is usually quoted as the reciprocal of the amount of exposure required to give a specified density level. The use of a D versus Log E characteristic curve is convenient in that it closely relates to the perceived image contrast because of the logarithmic eye response: the slope of the curve (the gamma) is usually equated with the image contrast.

2.3 Densitometry of film samples

There is some debate as to the best way in which to present experimental results obtained from film samples, either in terms of transmission or density. Both of these quantities have been used in the past in connection with noise or "granularity" measurements, and a comparative study may be found in [33]. A common feature of all measurements on emulsions is the use of an aperture of some type to scan small areas of the film. If this aperture is made too small then it may be filled by a single grain and this will give rise to a high noise figure. Similarly a large aperture tends to smooth out local fluctuations in density, and the size and shape of the area scanned is a most important parameter in all measurement work. A simple model of the photographic emulsion due to Nutting [34] in 1913 allows the density to be related to the number of image grains n_A in an area A by the equation

$$D = \log_{10} e \cdot n_A \bar{a} / A \quad (2.1)$$

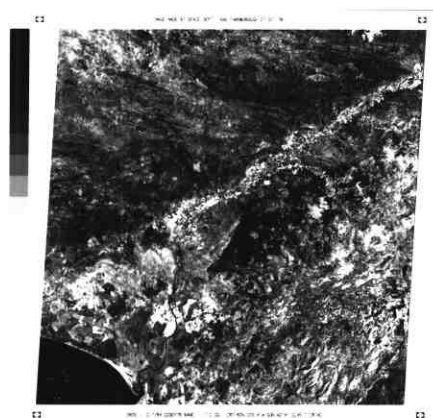


Figure 2.1 : Print of a typical 70 mm Landsat transparency

where $\bar{a}$ is the mean projected area of a grain. This equation is derived by assuming the emulsion to be composed of layers one grain thick in which the grains are randomly distributed with Poisson statistics, and grain overlap within each layer is assumed to be negligible. Each layer is further assumed to scatter and diffuse transmitted light so much that overlapping effects between adjacent layers are also negligible. Equation (2.1) holds well even for high image densities, but must be corrected for the area of a grain which tends to depend on the development conditions. In the present study, photographic density has been chosen as the basic parameter for measurements following the Nutting formula.

The mean density and the variations of local densities from the mean may be defined by the standard equations for a statistical process, with the assumption that the process is stationary and ergodic:

$$\bar{D} = \lim_{x,y \rightarrow \infty} (1/4xy) \int_{-x}^x \int_{-y}^y D(x,y) dx dy \quad (2.2)$$

and

$$\sigma_D^2 = \lim_{x,y \rightarrow \infty} (1/4xy) \int_{-x}^x \int_{-y}^y [D(x,y) - \bar{D}]^2 dx dy \quad (2.3)$$

where $\bar{D}$ represents the mean density and σ_D^2 is the variance of the density distribution. Equations (2.2) and (2.3) are the first and second central moments of the random noise Probability Density Function (PDF) which is unknown. If the scanning aperture is made large enough then because the measured density is due to the effects of many grains it is possible to invoke the Central Limit Theorem and to assume that the density variations follow a normal or Gaussian distribution. The validity of this assumption is tested for the apertures used in this work and has been found to hold true.

Measurements of film noise were carried out on Landsat 70 mm film images obtained from the EROS central data facility in Sioux Falls, Dakota since this is a commonly used form of data, particularly for the MEIC. These images, an example of which is shown in figure 2.1, include a grey level scale at the edge of the image data for calibration purposes, each step of which is exposed uniformly to a preset transmittance level. These steps form a convenient means of measuring the standard deviation of the density variations and also for

investigating any variation of the noise with mean density level. The instrument used for the measurements was a Joyce Loebel dual beam scanning microdensitometer equipped with a digital recording unit. A review of the construction of densitometers is given in [34], but a brief description of the operation of this instrument follows (see figure 2.2). The film is scanned using a small and adjustable spot of light from a primary source which is focussed by a microscope objective. A second objective collects the light and relays it to a detector (a photomultiplier tube) via a scanning aperture and a chopper. The PMT alternately receives light from the film and a reference beam which is modulated by a linear density wedge. The wedge is moved by a servo motor until the reference beam has the same intensity as the film beam, and the wedge movement representing the density is mechanically coupled to a pen. The film is mounted on a scanning stage which is linked to the paper chart, thus producing a trace of image density along a line of film. The wedge position is also recorded by a precision potentiometer which is linked to a digitiser and magnetic recorder so that numerical data is available for analysis. The speed with which the film is scanned affects the bandwidth of the PMT signal, and in order to compensate for the relatively slow mechanical response a differential scan control is available which slows the film down when a sharp change in the density is detected and allows the wedge to track more closely. The wedge movement servo system is also provided with a variable damping control to limit overshoot and ringing.

The results obtained from the densitometer are critically dependent on the aperture size and the illumination / collection optics of the system. The illumination of the film must be uniform, and the instrument is adjusted to the Köhler criterion in which all light from any point on the source is incident on the film with the same angle. The size of the aperture stop determines how much of the source is used and therefore the coherence of the illumination. In order to avoid unwanted effects from partial coherence, and to allow a simple treatment of the optics using an Optical Transfer Function (OTF) the Numerical Aperture of the condensing objective and the aperture stop must be chosen carefully. The imaging objective must have sufficient magnification to allow a convenient size of scanning aperture to be used and a sufficiently high NA to allow collection of all the light from the condenser. The objectives together with the control electronics determine the MTF of the densitometer and thus the smallest resolvable grain size. The choice of scanning aperture is dependent on several factors including those noted above. This aperture, which is a

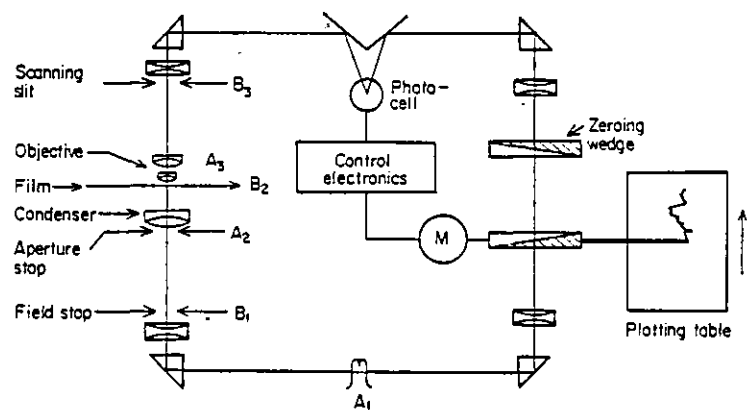


Figure 2.2 : The Joyce Loebel Scanning Microdensitometer

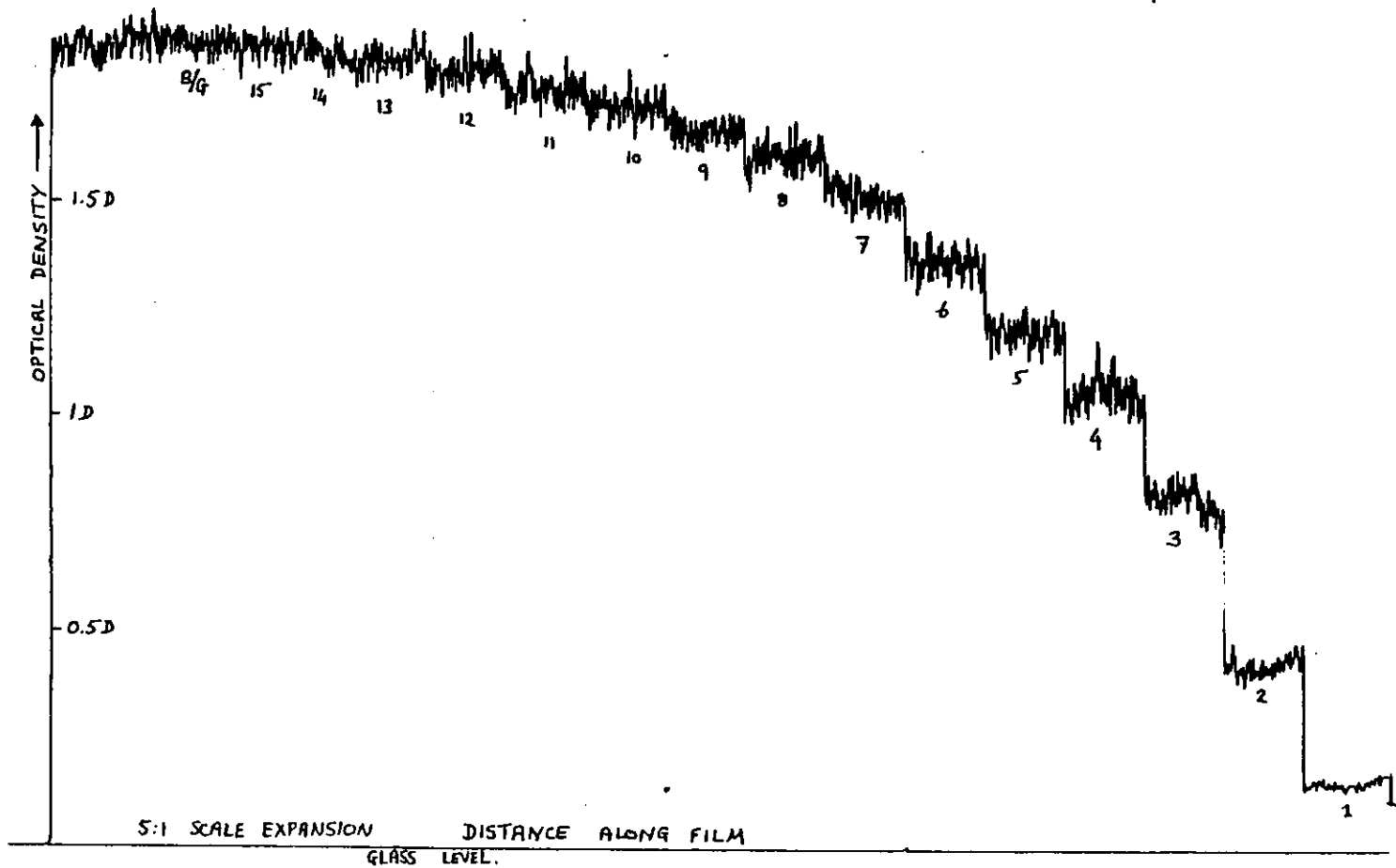


Figure 2.3 a : Densitometer trace of 70 mm grey scale

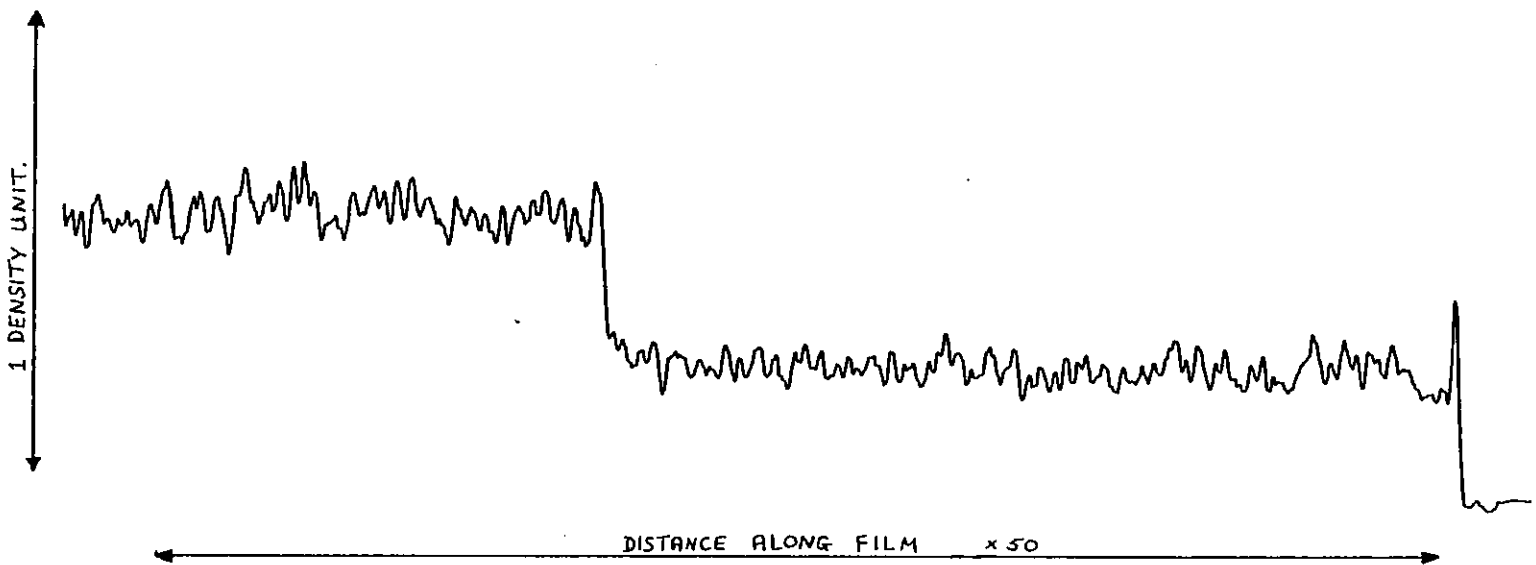


Figure 2.3 b : Magnified portion of figure 2.3 a

rectangular slit of variable width and length, determines the spatial frequency response of the system, but a small aperture chosen for high resolution will lead to high noise levels. The area of the aperture may not be the only factor which affects the noise signal however, since if the grains (or clumps of grains which are commonly found in emulsions) are relatively coarse compared to this area, the shape of the aperture becomes important: a circular aperture may be obscured by a single clump of grains but a thin rectangular aperture of the same area will perform an averaging function and reduce the noise fluctuations. Apart from complying with the illumination requirements the aperture used in this work was chosen to match the probable size and shape of a high resolution FSS which could be used to read Landsat transparencies, and was a square 23 μm on a side. A real FSS would use a circular spot with a Gaussian intensity profile so that the square is only a rough approximation.

Figure 2.3 (a) shows a typical scan along the entire grey scale of a 70 mm transparency and figure 2.3 (b) shows a magnified portion of the scan covering one grey level step. There is some evidence of edge effects at the boundaries between steps which causes 'overshoots' or local distortions in the mean densities, and in order to prevent these degradations in the numerical data all subsequent scans were carried out in the central regions of each step. The entire scan shown in figure 2.3 (a) shows the logarithmic shape which results from a linear transmission wedge which has been written on the film. The grey levels become closely spaced at high densities and are hard to distinguish since the noise is often of similar amplitude to the level separation. In such cases levels may be resolved visually only when large areas are viewed. In order to obtain statistical data the densitometer was set up to scan within each grey level step on four samples of film, acquiring 1024 data values per step per scan. This number was chosen as a compromise between the requirement for a large, statistically valid sample and the problems of data errors in the recording equipment. A FORTRAN program was written to compute the mean and standard deviation for each block of data and to sort it into classes for a histogram. The relative populations of these classes were then compared to those obtained from a hypothetical Gaussian distribution having the same mean, standard deviation and class boundaries. The hypothesis that the spread of densities about the mean follows a Normal distribution was then tested using a Chi-Squared calculation for each grey level. Up to eleven grey level steps were used for each film sample, and many scans were repeated several times in order to check accuracy. Histograms of the results were plotted on a HP-85 micro

computer and plotter and some results are shown in figure 2.4. A good agreement with the Gaussian distribution is evident and is supported by the Chi-squared test, although a more thorough study with much larger samples would be required to confirm the accuracy of this hypothesis. This result confirms that the aperture used to scan the film is large compared to the granularity (not only the individual grains but clusters and clumps of grains).

The standard deviations of the density are seen to vary from around 0.01 D to 0.05 D (D = Density units) with the lowest values occurring at both low and high mean densities. In order to check this further the standard deviation values obtained in each scan were plotted against their respective mean density values (ranging from 0.1 D to 2.5 D) as shown in figure 2.5. Values obtained from multiple scans of the same level have been grouped together into single points. The solid curve was obtained by performing a least squares curve fit with Chebychev polynomials, and represents the lowest order curve that gave a significant reduction in the Residual Sum of Squares (RSS). Higher order curves begin to oscillate between the points and do not represent an improvement in fitting. The curve used is of the fourth order, and it should be noted that extrapolation beyond the minimum and maximum density values used in the fit is invalid, since the Chebychev polynomials are not defined in these regions. At the low density region extrapolation would be limited by the residual fog level of the film which is not far different from the lightest grey level step. The form of the Standard deviation - density curve may be explained as follows: as the exposure is increased, the larger grains in the emulsion will be exposed first because of their greater photon capture area. Their size also makes their contribution to the noise signal relatively greater than for small grains, so that the standard deviation rises quickly as in the experimental curve. Finer grains then begin to be exposed in greater numbers as the number of available quanta increases, and they tend to fill in the gaps between the larger grains without such a large contribution to the noise. As the film approaches full exposure the source of the noise changes from a random array of large to medium grains in a clear matrix to a random array of smaller gaps in an opaque 'matrix' of silver. These gaps decrease in size and contribute more to the higher spatial frequencies, until at full exposure (saturation) only residual gaps remain due to the random nature of the grain distribution in the gelatin. The noise thus tends to decrease at higher densities. This proposed 'change' in the noise source is a continuous process, and is complicated by random distributions of quanta, overlapping and scattering effects and adjacency effects in the

FILM NOISE DENSITY DISTRIBUTIONS

FILM# 1 APERTURE 22.7 * 22.7 MICRONS

1024 DATA POINTS

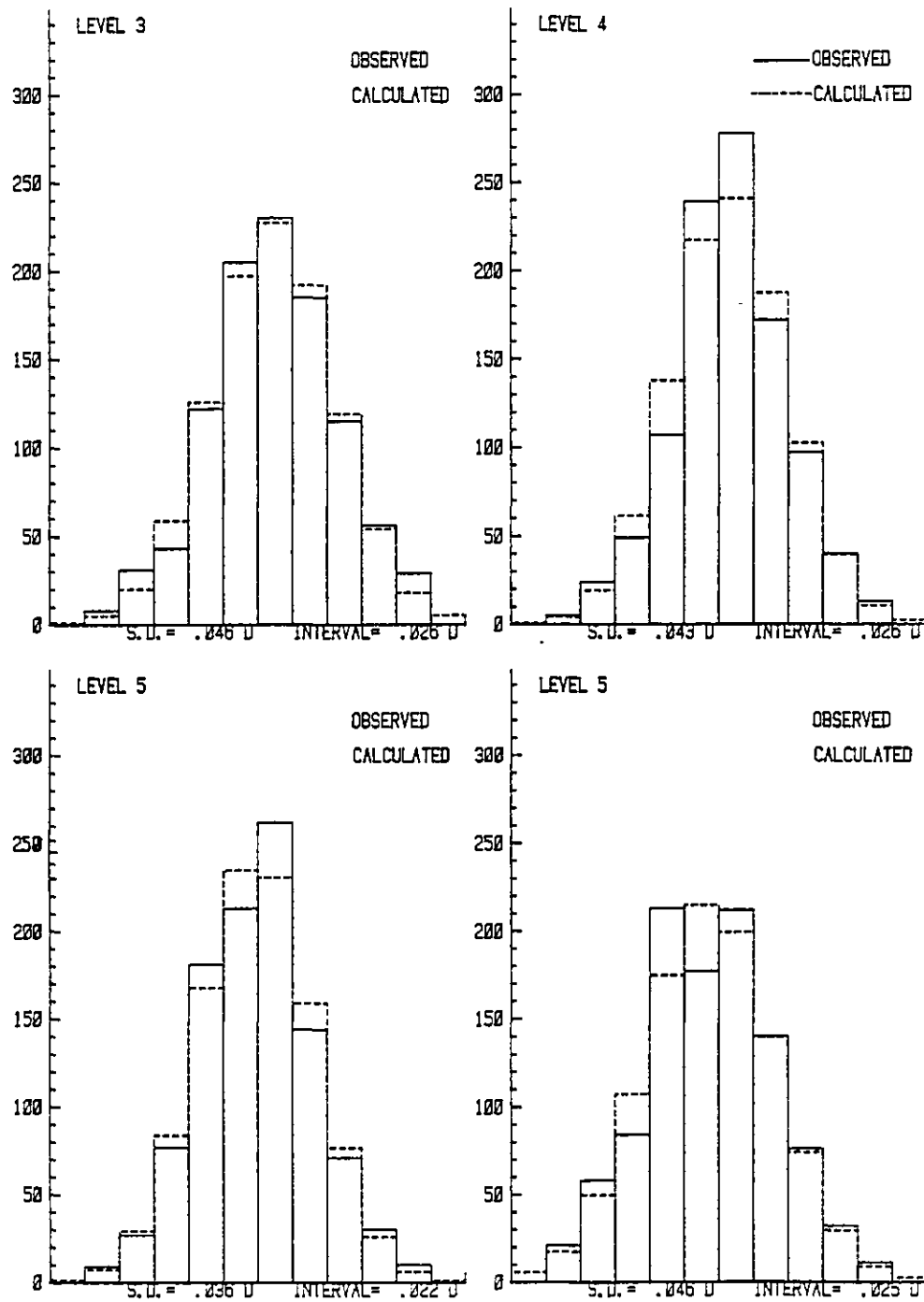


Figure 2.4 : Typical grey level distributions

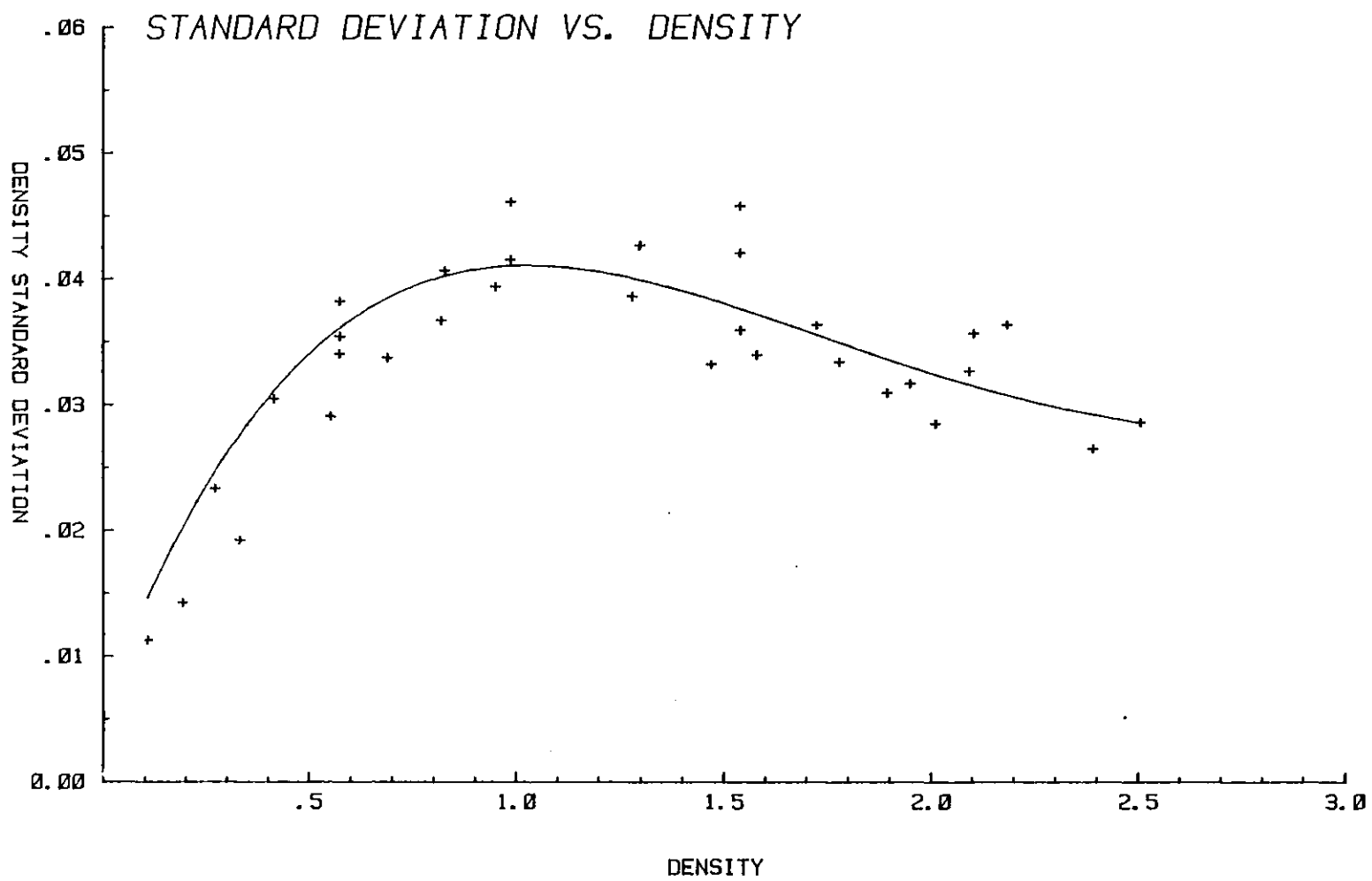


Figure 2.5 : Density standard deviation against density

development stage. Detailed study of such effects is beyond the scope of the present work.

Accepting a Gaussian noise distribution now permits the calculation of the numbers of levels available within the grey level range of the film for the given film samples. In order to proceed it is necessary to specify an error tolerance with which each level must be detected from a spot measurement. A set of grey levels is constructed within the dynamic range separated by density values which are related to the standard deviation of the density about the mean. The separation between levels is calculated with reference to the normal distribution and its Cumulative Distribution Function (CDF) sometimes called the error function. Figure 2.6 illustrates several adjacent grey levels and the Gaussian spread of measured densities around their means. To a first order approximation the overlapping areas represent the error probabilities (assuming that correct normalisation is applied). The error function $\text{erf}(x)$ measures the area under the Normal distribution from $-\infty$ to a specified level x above the mean, and the complementary error function $\text{erfc}(x)$ is defined as $1 - \text{erf}(x)$. Overlaps from next nearest neighbours will also make a contribution but this will be small and may be neglected. The errors are thus given by

$$P_1 = \text{erfc}(x_2/\sigma_2) \quad (2.4 \text{ a})$$

$$P_2 = \text{erfc}(x_1/\sigma_1) + \text{erfc}(x_3/\sigma_3) \quad (2.4 \text{ b})$$

$$P_3 = \text{erfc}(x_2/\sigma_2) + \text{erfc}(x_4/\sigma_4) \quad (2.4 \text{ c})$$

$$P_i = \text{erfc}(x_{i-1}/\sigma_{i-1}) + \text{erfc}(x_{i+1}/\sigma_{i+1}) \quad (2.4 \text{ d})$$

where P_i represents the error tolerance, σ_i are the standard deviations and x_i are the widths of each level as shown in figure 2.6. Since the standard deviations are not constant with density analytical derivation of the positions and widths of each level is complex, so a further FORTRAN program was written which handled the problem numerically. This program places a first level so that its mean satisfies

$$D_1 - D_{\min} = k \cdot \sigma(D_1) \quad (2.5)$$

where $D_{\min}$ is the minimum density for which the fitted curve of figure 2.5 is valid and $\sigma(D_1)$ is the standard deviation at a mean density D_1 . The factor k is specified by the error tolerance, in this case for an imaginary level with

OVERLAPPING GAUSSIAN GREY LEVELS

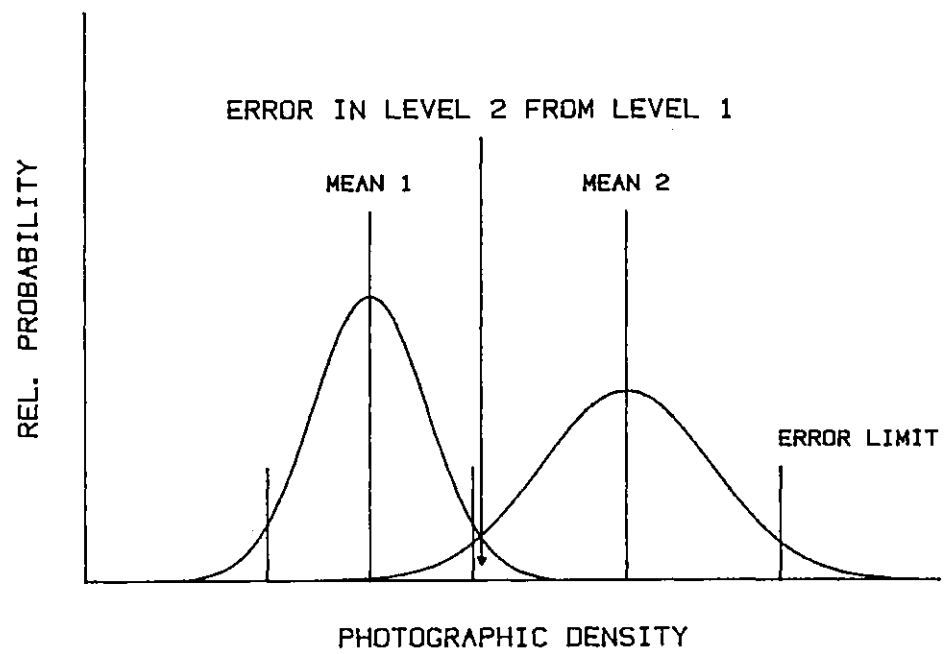


Figure 2.6 : Overlapping Gaussian grey levels

a boundary at $D_{\min}$. The next level is then placed at $k\sigma(D_1)$ above the first and is slid up the density range in small steps until the separation between the two levels is given by

$$D_2 - D_1 = k.\sigma(D_1) + k.\sigma(D_2) \quad (2.6)$$

taking account of the variable deviations. If an overshoot occurs the program retraces its steps using a finer increment. This process continues until the entire range of density is filled. No extrapolation is attempted at either low or high levels. When this routine is applied with different error tolerances the numbers of distinguishable grey levels varies as shown in table 2.1.

Table 2.1 DISTINGUISHABLE GREY LEVELS

ERROR TOLERANCE	NUMBER OF LEVELS
0.5%	13
1.0%	15
5.0%	19
10.0%	22
20.0%	29
50.0%	54

It must be remembered that these are raw data which take no account of variations in aperture size, filtering in electronic amplifiers in a practical FSS or externally generated noise.

2.4 Factors affecting noise in practical systems

When a photographically recorded image is read using a real FSS several factors combine to modify the noise characteristics and hence the number of grey levels which are available. In addition, Landsat multispectral images represent four channels of data, each with its own noise signal, and the above calculation for the number of grey levels must be extended and interpreted in terms of feature space. This section will outline some of the areas which contribute to a greater number of resolvable grey levels, but it must be remembered that they all lead to a loss in spatial resolution in the image.

The first parameter which affects the noise level is the scanning spot size and intensity profile. Any aperture used to scan a film acts as a spatial frequency filter whose characteristics may be determined by performing a two dimensional Fourier transform on the aperture shape function. This shape function includes any intensity weighting across the aperture such as may be obtained with an apodising filter. The spot used to scan the film samples in the previous section was square and evenly illuminated across the entire area, and the corresponding spatial frequency characteristic is a SINC function which passes low frequencies and cuts out higher frequencies. This frequency response shows 'side lobes' because of the subsidiary maxima of the SINC function, giving an indistinct cut-off frequency, and practical FSSs do not use a square spot. The spot on a CRT may be approximated by a Gaussian distribution of intensity because of the electron optics involved [36] and this gives almost the ideal low pass filter characteristic, showing no 'ringing' whilst having a very sharp cut-off. The spatial cut-off frequency may be found as indicated above, and if the standard deviation is used as a width measure, the cut-off occurs at a frequency f_c where

$$f_c = \text{S.D. of response} = 1 / 2\sigma_r \quad (2.7)$$

where σ_r is the radius of the spot at one standard deviation from the centre, assuming a circularly symmetrical spot. Note the inverse relation between pass bandwidth and spot size. The exact definition of f_c is arbitrary, and has been specified variously as the half power, one-on-e or one-on-e-squared point. In order to determine the effects of any spot on the signal it is necessary to multiply the signal spectrum by the spot response or in spatial coordinates to convolve the signal with the spot distribution. In the case of photographic films it can be shown that as the spot size is increased the standard deviation of the noise signal decreases with the square root of the spot area: doubling the spot dimensions (quadruple area) halves the noise. Obviously a larger spot also reduces the resolution of the scanner, but this may be offset to some extent by using large format imagery (e.g. 240 mm transparencies instead of 70 mm) since the graininess will remain of comparable size but the smallest resolvable element of the image will be larger. This is equivalent to reducing the image spatial frequency bandwidth, allowing the use of a narrower pass band filter (a larger spot) which passes less of the noise spectrum. On a real FSS the spot may not be entirely symmetrical because of the combined effects of scanning and phosphor response time. As was noted in section 1.4 the phosphor

must have a short decay time or the spot will leave a 'tail' of light which illuminates parts of the film which have already been sampled by the main spot. In TV rate FSS scanners the phosphors employed in the CRT have decay times measured in nanoseconds, and the streaking effect is negligible.

A second parameter which must be considered is the frequency response of the amplifiers following the PMTs or other detectors and the response time of the detectors themselves. The signal bandwidth which must be accommodated by the system depends on the scanning rate, and slower scans place a lighter burden on the electronics. A slow scan may be convenient where bandwidth is a limitation (e.g. Slow Scan TV transmission by radio amateurs) but it gives a low data rate and is inconvenient for display purposes because of flicker. It is common practice to scan at standard broadcast TV rates, with the added advantage of compatibility with commercial video recording and transmission systems. The frame scan rate is typically 25 Hz (two fields interleaved at 50 Hz to reduce flicker) and each frame consists of 625 scanned lines. The line rate is thus 15.625 kHz, and about 575 lines actually carry picture information. Each line is 'active' for 52 μ S out of a total period of 64 μ S. The typical bandwidth encountered in a broadcast system is about 4 MHz, but high definition cameras are available with up to 40 MHz bandwidth. The PMT has a very fast response, and if silicon diode detectors are used then PIN types are to be preferred for their sensitivity and speed (around 0.5 nS rise time). As in the paragraph on spot size, the effect of limited bandwidth in amplifiers is best characterised by the frequency response, but most electronic circuits exhibit a slower cut-off than the Gaussian filter, and attempts to sharpen the pass band usually results in ringing or resonance effects and an uneven response. Using the scanning rate and the frequency response, an 'effective spot shape' may be found for the electronics which can then be convolved with the incoming signal if required.

The design of a FSS system is constrained by many factors such as spot size and demagnification, focal lengths of lenses, image dimensions and the number of channels, but one limitation stems from the detector itself. Often the amount of light passing through a dense area of film may be small and the sensitivity of the PMT and associated preamplifier must be high. Unfortunately the system can be limited by a lack of photons, and the quantum nature of the detected signal manifests itself as noise. Simply increasing the gain of the PMT does not help, and the spot-forming optics must be carefully designed in order to

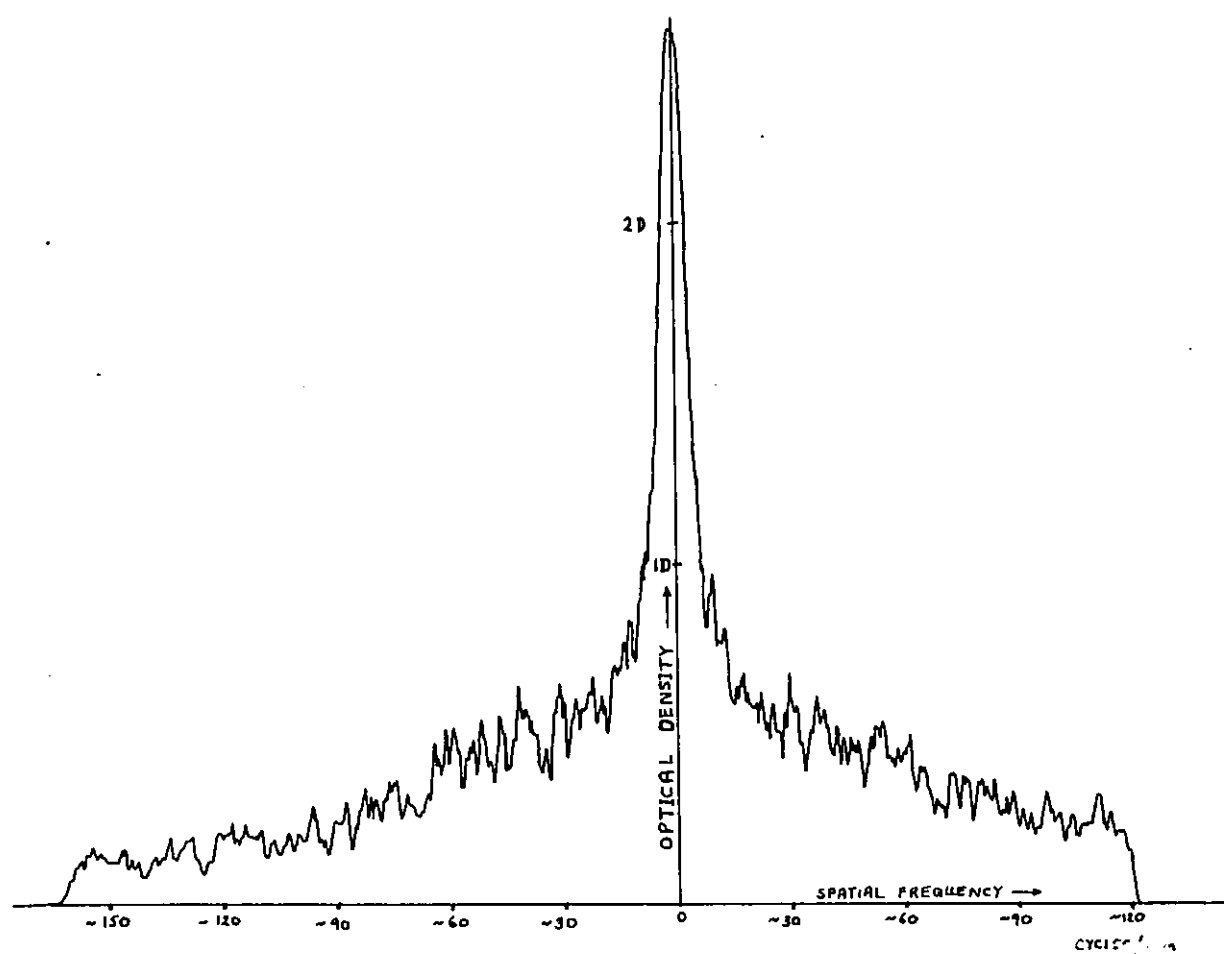


Figure 2.7 : Densitometer scan of optical diffraction pattern

collect as much light as possible from the CRT. This factor may limit the dynamic range of the image transparencies which can be accommodated in the system.

The various parameters noted in this section must all be taken into account in any calculation of the number of distinct grey levels available from film. Obviously it is necessary to characterise the raw film noise as in the previous section which used the widespread and 'standard' 70 mm Landsat transparencies, but the final results are highly dependent on the subsequent hardware. The formula by which an estimate may be reached runs as follows: firstly an aperture size must be chosen on the basis of resolution requirements and a rough estimate of the film graininess. Secondly the density noise must be measured and a Gaussian distribution verified - if it is not then the aperture is probably too small and contains too few grains. Thirdly a level fitting program must be run if the noise level shows a significant variation with mean density. Fourthly and lastly a combined filtering response must be measured and/or estimated for the reading system and applied to the signals. Various models for film grain noise suggest a reasonably flat spatial frequency spectrum for the primary grain or reduced halide crystals [38], but it is found that clumping and adjacency effects in development tend to cause a distinct drop at higher spatial frequencies. Much work has been done in this area [39] but it is usually necessary to measure spectra for each particular film. In the present work a coherent optical processor was used to measure the spectrum of a sample of grey level steps. The film specimens were mounted in a liquid gate in order to index match to the surfaces and reduce diffraction from phase effects. The samples were illuminated with blue light from an argon ion laser and the diffracted light was collected and focussed by a 700 mm focal length lens. Recording was carried out on sheets of film placed in the focal plane of the lens, and the density of the film (which is assumed proportional to exposure and therefore to the spectral energy distribution) was measured with the Joyce Loebel scanning microdensitometer described in section 2.3. Figure 2.7 shows a typical scan across the centre of the diffraction pattern, with the central dense peak corresponding to the Airy maximum due to the mean transmission of the sample. The beginning of section 2.3 indicated the direct relation between density and the number of grains within an aperture, thus explaining the use of density rather than transmission in most of this chapter. The spectrum recorded by optical means is one of transmission, and was used in preference to one of density for the following reasons: the optical spectrum averages statistical

variations over a large area of film (several mm square) and gives a picture of the entire two dimensional spectrum immediately. The alternative is to perform multiple scans on the densitometer and apply a digital FFT to each scan. The slit shape determines the form of the results and the smaller number of samples leads to less accurate results (the limiting factor in the optical case is the lens PSF which, for a large aperture, well corrected lens is very small). The relationship between the two types of spectra is shown in [39] to be

$$F_D(w) = [(\log_{10} e)^2 / T^2] F_T(w) \quad (2.8)$$

where F_D and F_T are the two spectra, w is spatial frequency and T is the mean transmittance of the area of film under consideration. This equation shows that the trace of figure 2.7 is proportional to $F_D(w)$. The overall shape is almost an exponential decay to higher frequencies, with no distinct boundary between the spectra of the primary grain and the clumps of grains. If the half power point of this spectrum is compared with that of the entire reading system an estimate of the noise reduction may be obtained assuming the two spectra are similar in shape. If this is not the case then the areas under the spectra must be compared. Application of the above techniques to the MEIC (chapter 1) which employs the same type of transparency as the samples used here gives an estimated 60 distinct grey levels given 5% error.

The specification of error tolerance is essential in any study of this type. A simple number of levels is not sufficient alone. The work so far described uses only one channel of image data, but in the case of multispectral work four or seven channels are more common. In this case the above methods may be applied to each channel, and the result is that the multidimensional feature space is partitioned into notional boxes whose edges are defined by the grey level boundaries in each channel and whose number is the product of the number of levels per channel (for given error within each channel) and the number of channels. These boxes are not all equal in size or shape but change in sympathy with the noise levels, and the overlap widths will depend on error tolerances. The total error in measuring a feature vector tip position (the classification error due to the film alone) will be the product of the errors in each channel. If an overall tolerance of 5% error is set, then this product must equal 0.95, and the number of distinct 4-D boxes can be found from table 2.1 (each channel has about 1% error). When the effects of frequency response are included, the final figure is approximately 4 million.

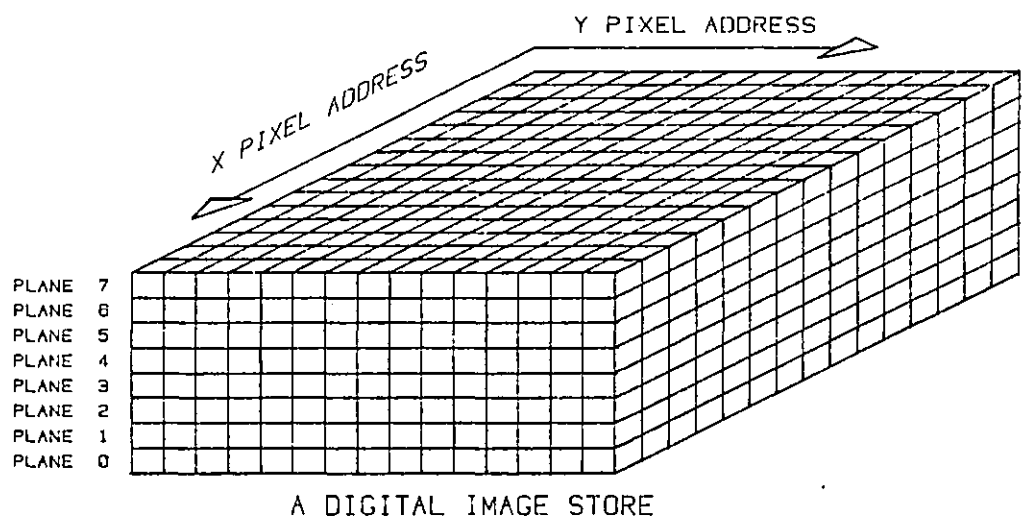


Figure 2.8 : Digital image storage

2.5 Bitplanes and digital storage

This section describes the second approach to the estimation of the number of grey levels on film, and the approach can be widened to estimate the amount of information in any digitally coded image. The digital picture function was described in section 1.2, and in practical storage systems each pixel is coded as a binary word whose length (the number of binary digits or bits per word) determines the maximum intensity resolution of the system. The most common word size is 8 bits (commonly called a BYTE) which allows intensity values up to 2^8 or 256. Figure 2.8 shows an array of bytes arranged as in a digital display in rows and columns. The divisions in the 'depth' of the diagram represent the different bits within each word, and the entire array may be thought of as being composed of 8 slices or planes of these bits. Each plane makes a weighted contribution to the pixel numbers, the weighting factor being proportional to two raised to the power of the plane's number. The most significant bit thus makes a relative contribution of 2^7 (from plane 7) and the least significant only 2^0 . Section 2.1 indicated that each plane of bits could be treated as an image in itself, and since each bit is a binary digit, the images thus formed are also binary or black and white (with no grey).

The bit plane image may be derived in another way, namely contouring. If the available grey level range is split into two halves, and all intensity values below and above the centre are made black and white respectively, the result is the most significant bitplane image. This is only a crude approximation to the original image. The second bitplane image can be obtained by repeating the contouring so that the intensity range is split into four, coded black-white-black-white. If these two bitplane images are now added together with the binary weighting described above the result is a better approximation to the original with four possible shades of grey. This process of successive approximation can be continued indefinitely, but after a certain point further bitplanes do not make a significant contribution. The most significant bitplane image shows much of the gross structure of the original, and this structure becomes finer and more random as the bitplanes become less significant until they resemble random fields of pixels. This forms the basis for the present study: if the number of bitplanes required to form an acceptably faithful copy of the original image can be determined then the number of grey levels will be approximately two to the power of this number. For any given digital image the optimum amount of storage space can then be decided, and for

large images the savings possible can be significant.

To obtain digitised images for the present work the image on a 70 mm Landsat transparency whose grey scale had been used in the previous sections was used. Digital data was transferred onto CCT from a Joyce Loebel Scandig 3 rotating drum densitometer and Data General Nova computer system. This densitometer was set up so that the available dynamic range was filled by the image so as to collect the maximum signal possible from the film. The scanning aperture selected was approximately the same size as the effective aperture of the MEIC FSS system. The digital data was then read off to disk storage on a Digital Equipment Corp. VAX 11/780 computer. A FORTRAN 77 routine was written which enabled the extraction of bitplanes from a subscene 64 pixels on a side. The size of the subscene was chosen as a compromise between statistical reliability and ease of handling. The particular area chosen is part of the Medway estuary in North Kent taken through a red filter by Landsat 3 in 1973 (figure 2.9).

In order to determine the usefulness of any bitplane it was decided to examine the corresponding autocorrelation functions. In a significant bitplane there are large areas of the image in which adjacent pixels are likely to have the same value (black or white), whereas a less significant bitplane is more 'broken' in appearance. This shows up in the width of the autocorrelation functions, and gives a similar measure to a visual appraisal of each plane since the eye seems to perform a similar calculation automatically. Autocorrelations were computed by first applying a Fast Fourier Transform (FFT) to the required bitplane (or bitplanes if a cross correlation is required) and then squaring the results before a reverse transformation. In order to prevent aliasing or wrap-around errors in the results the FTs were reduced to zero mean arrays by subtraction of an appropriate constant and then padded around the outside with an array of zeros. The two FTs were combined by forming the product of one array and the complex conjugate of the other. The entire process required about a minute to run, and the correlations were presented as isometric 'surface mesh' plots on a graphics terminal and plotter. Each meshplot was reduced from 128 pixels square to 64 pixels square by averaging for ease of viewing and reduction in plotting time. Hard copy of the bitplane images was obtained on a commercial camera system attached to a monitor. These images and their corresponding autocorrelation plots are shown in figure 2.10 (a) to (h).



Figure 2.9 : Original image (part of North Kent), band 5



Figure 2.10 a : Bitplane 7 (most significant)

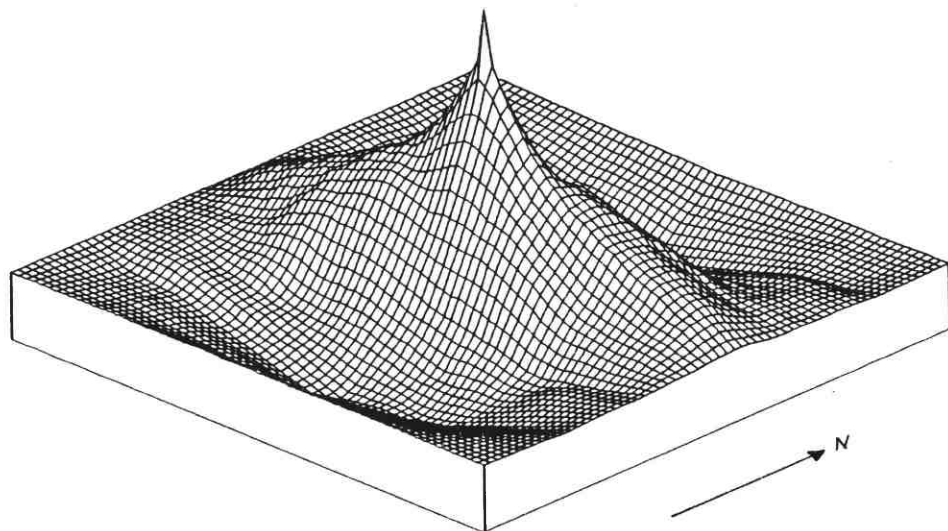


Figure 2.10 a : Autocorrelation plot for bitplane 7



Figure 2.10 b : Bitplane 6

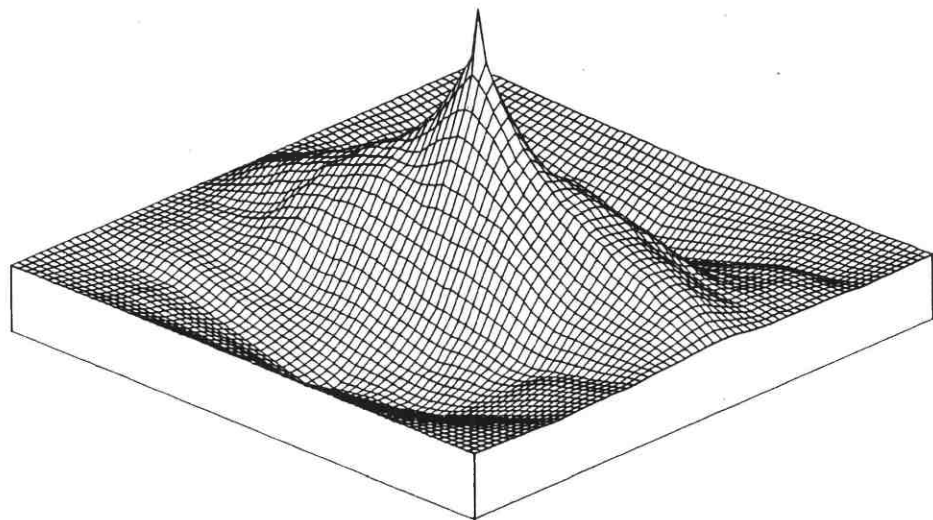


Figure 2.10 b : Autocorrelation plot of bitplane 6



Figure 2.10 c : Bitplane 5

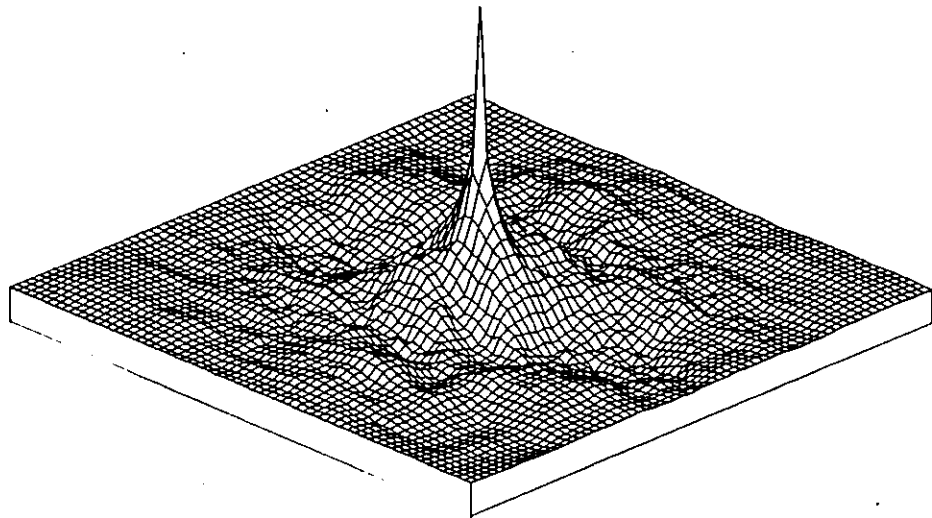


Figure 2.10 c : Autocorrelation plot for bitplane 5



Figure 2.10 d : Bitplane 4

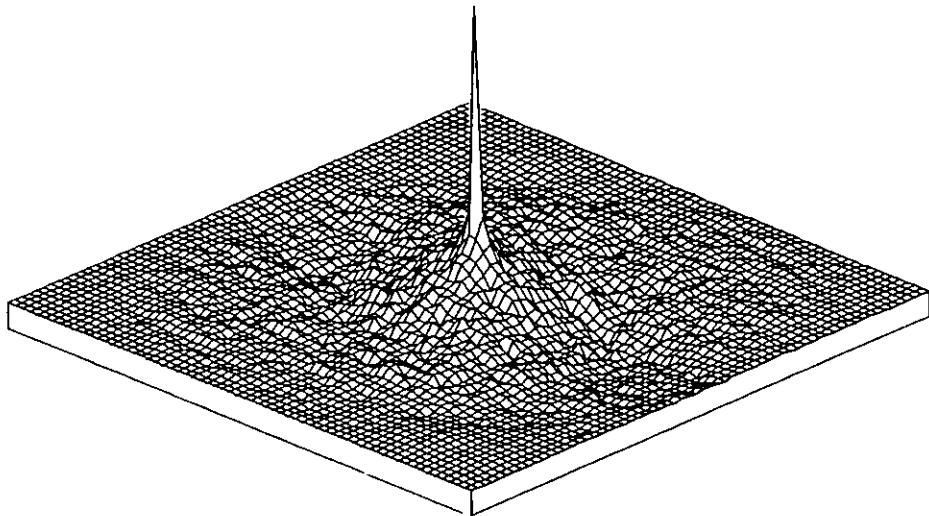


Figure 2.10 d : Autocorrelation plot of bitplane 4



Figure 2.10 e : Bitplane 3

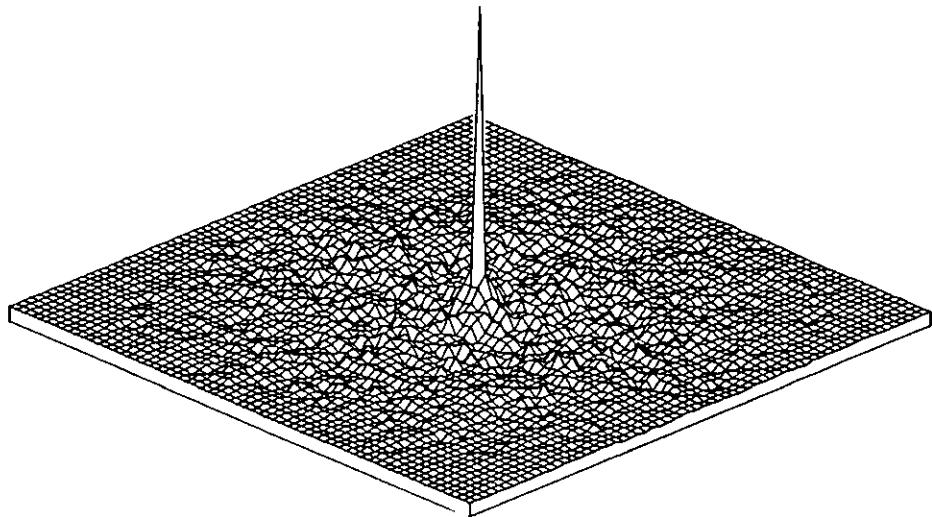


Figure 2.10 e : Autocorrelation plot for bitplane 3

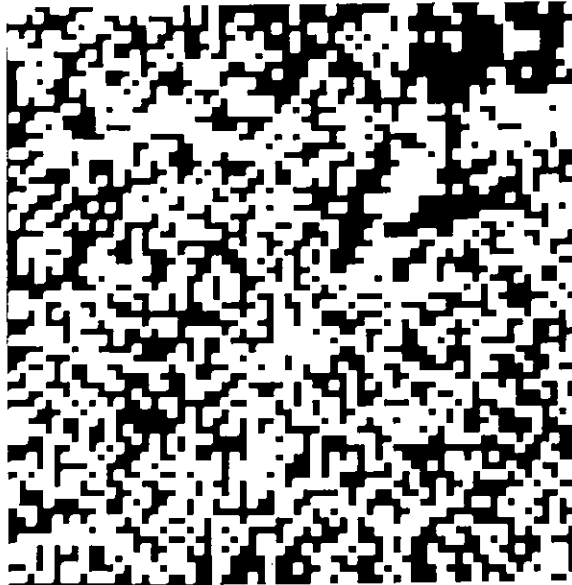


Figure 2.10 f : Bitplane 2

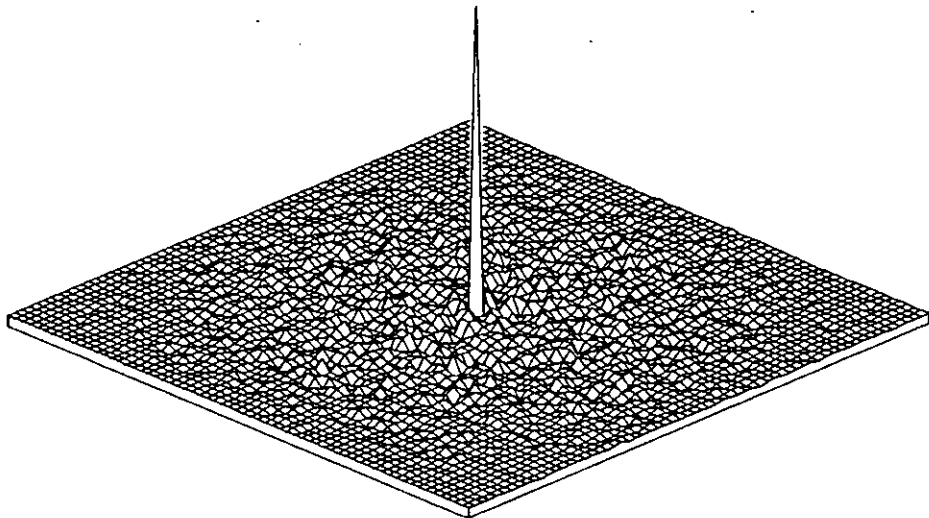


Figure 2.10 f : Autocorrelation plot for bitplane 2

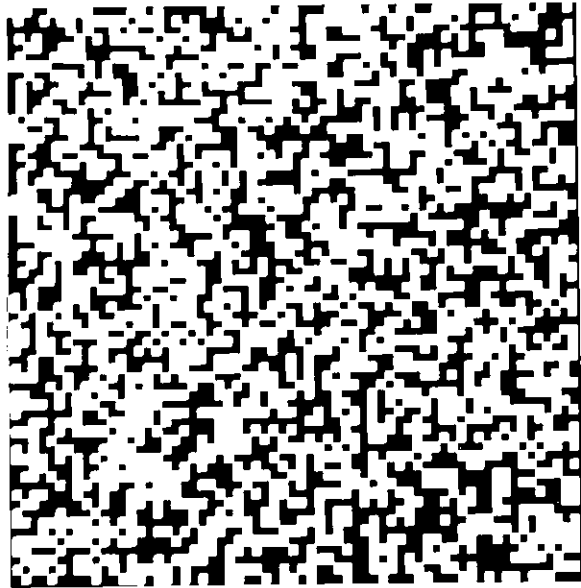


Figure 2.10 g : Bitplane 1

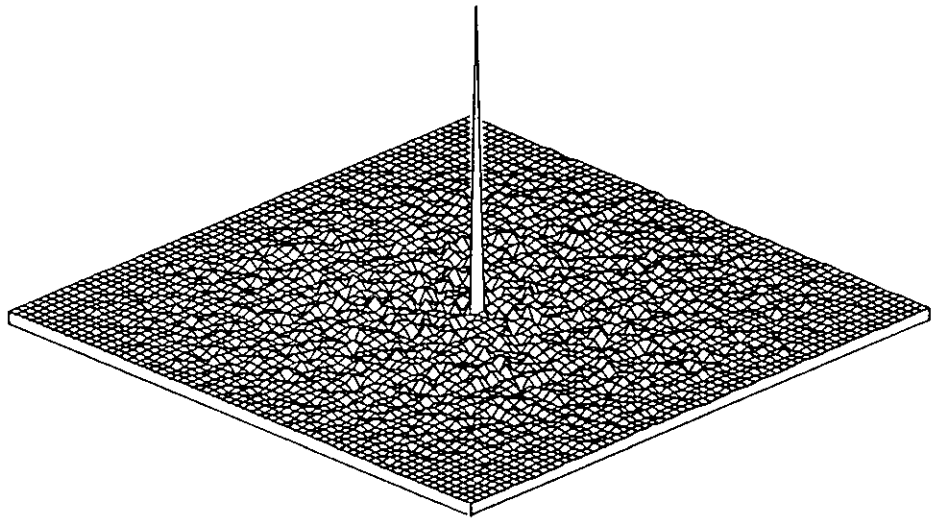


Figure 2.10 g : Autocorrelation plot of bitplane 1

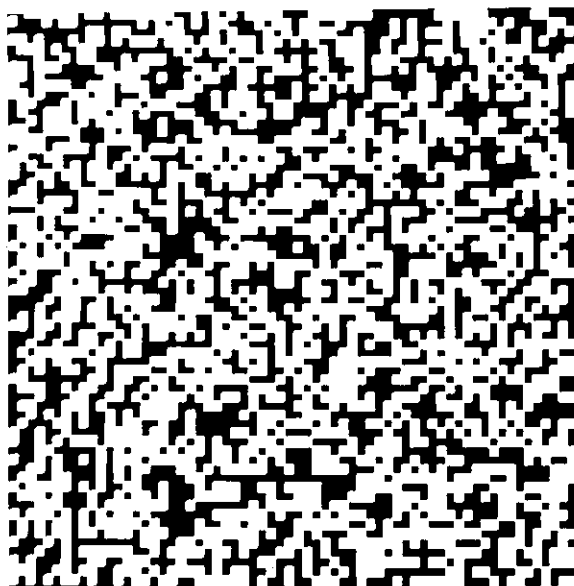


Figure 2.10 h : Bitplane 0 (least significant)

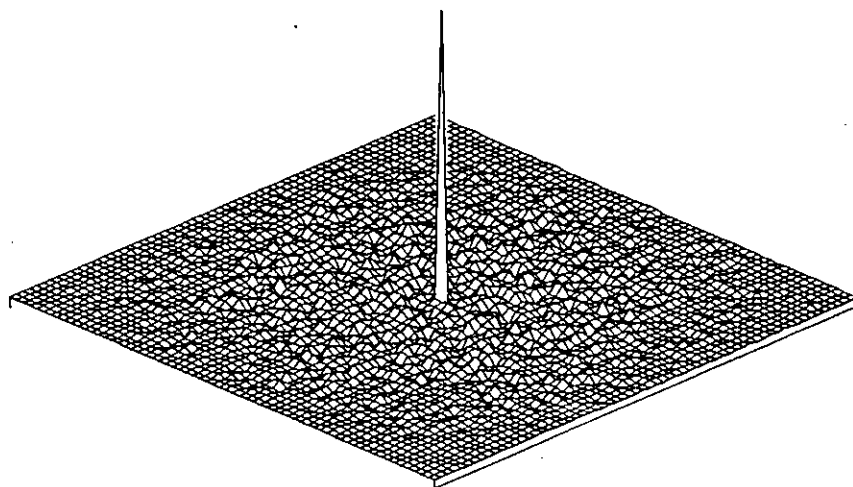


Figure 2.10 h : Autocorrelation plot of bitplane 0

The bitplane images show several interesting properties and they are worthy of careful viewing in comparison with the original image. The most obvious general trend as the bitplane number decreases (becoming less significant) is a loss in gross features and a randomising effect. In fact planes 7 and 6 are nearly negatives of each other except for a few pixels, a situation which is not usual but which does not affect the interpretation. Both of these images pick out the coastline including the sand bar in the centre and some features inland corresponding to the lighter areas of the original. Neither image shows any variation over the sea surface and both show a uniform area of land in the lower left. The presence of a uniform area in a bitplane image signifies that the original picture intensities in that area were confined to a range smaller than the contouring interval of the particular bitplane. The autocorrelation function for these two bitplanes shows a wide but directional form, with the major correlation occurring in a direction broadly parallel to the coastline as may be expected. The negative 'valleys' indicate the lack of correlation between the land and sea areas of the image. Because the images are almost identical save for an inversion the two autocorrelations are also the same.

The image of bitplane 5 presents a different appearance from the previous pair. The coastline is no longer so clearly defined and the image contains less extensive uniform areas. In comparison with the original image however the features of plane 5 are readily 'seen' in the original. The dark band along the landward side of the coastline is visible, as are other features on the land. Of particular note is the division of the sea into two broad areas. Physically this division corresponds to a separation between shallow silted areas and deeper water, as can be seen on the original image. The autocorrelation function shows a much narrower peak in the centre but there is still distinct structure in the wings. The influence of the major division in the sea area has caused the directionality of the autocorrelation to shift around almost perpendicular to the previous two plots.

Bitplanes 4 and 3 show increasing randomness, particularly over the land area where the original image is complex, and their corresponding autocorrelation plots show a more circularly symmetric form as the large uniform areas break up. There is still structure present in both these images, with features similar to the silting in the sea being visible. Bitplane 2 shows less structure but if it is compared to plane 3 or 4 then similarities become evident and some indication of the silt patterns is still present. The corresponding

autocorrelation is a very sharp peak with only a small amount of information indicated in the surrounding ripples. Finally bitplanes 1 and 0 show very little gross structure or relation to the original image although a long look at bitplane 1 may reveal some small correlation. The images are almost random fields of black and white pixels as is confirmed by their autocorrelations. The autocorrelation of a completely random signal whose transitions are constrained to occur at regular intervals (i.e. the pixel boundaries) may be calculated analytically [40], and has the form of a triangular peak one pixel wide, closely resembling the plot for bitplane 0.

The interpretation of these images and autocorrelations is very difficult and open to discussion, as is the entire question of the information content of an image. It can be argued that every bitplane should be saved regardless of its random appearance because over an area of complex texture little correlation is to be expected in the less significant bitplanes. This effect can be seen to some extent over the land areas of the images used here. Set against this proposal are considerations of the noise in the sensor and recording medium which have been indicated in the earlier parts of this chapter. If the image signal is corrupted with random noise which cannot be filtered out conveniently (perhaps by integrating over many images) then there is little point in digitising the data to finer limits than the noise allows. The complex textural areas will be corrupted as much as more uniform areas, so that the best criterion by which to judge the required quantisation limit would be the level at which a notionally uniform area (such as the sea in the present case) becomes randomised in its least significant bit plane. Obviously it is seldom possible to find a uniform area in a practical image, but experience will usually allow a reasonable estimate of the number of bits required, and it is normal practice to digitise to one or two more bits of accuracy, thus allowing a requantisation at a later stage. The image used in the present study was a Landsat band 5 image which was quantised with logarithmic intervals to 6 bit accuracy. When the data is processed this logarithmic quantisation is transformed into 7 bit linear which is then written onto film [41]. The densitometer used in this study was equipped with a 50 μ m scanning aperture (approximately twice that used in the film study of section 2.3) and the bitplane images indicate that data with 7 significant bits can just be recovered, and 6 bit data is available with ease. Remembering that the number of grey levels may be obtained by raising two to the power of the quantisation level, this indicates a dynamic range of between 64 and 128 levels.

2.6 Images and information

The previous sections have put forward two strategies for the determination of the number of distinguishable grey levels in a photographically recorded image. The first approach has been a direct measurement of the noise signals on uniform areas of film followed by various corrections which take account of the frequency response of the reading system, whereas the second approach sought to estimate the quantisation level beyond which no further useful information could be obtained. These two studies are not confined to the photographically recorded image, and should not be looked upon as the source of a definitive 'true' number of levels, but rather as methods to be applied to each individual situation, taking due account of the uses for which each system is intended. In most practical applications of the work other factors will influence the type of film or the quantisation levels used. An example of this occurs in the design of digital picture stores where, if the results of section 2.5 are applied, each store would have 7 bit resolution. The internal architecture of many modern microcomputers caters for 8 bit parallel data, and it is often more convenient to design a store with some redundancy. The total amount of storage will depend upon the spatial sampling frequency used, which will also determine the resolution, but sampling at a spacing smaller than the aperture size does not increase the information obtained.

In conclusion this chapter has investigated the properties of standard issue Landsat 70 mm transparencies as information stores and has shown them to be capable of giving many distinguishable grey levels. The important points which must be remembered are that a distinguishable grey level must be defined by an error tolerance, particularly in the case of multispectral classification, and that the characteristics of the entire system must be taken into account. Early attempts to arrive at a 'granularity' coefficient which completely describes film grain are not useful in this type of study. The storage space required for the same images in terms of digital words has been estimated and the idea of information in terms of significant bits has been introduced. A final point to note is that although there may be about 60 levels available, the distribution of image data among the levels may not be uniform or such as to make the image easily interpretable. This will be covered in the next chapter.

3.1 Contrast enhancement

This chapter describes a purpose-built electronic system designed to allow an operator to modify the contrast of a TV video image with reference to the image intensity statistics. The image intensity statistics represent the way in which the image function values are distributed throughout the available dynamic range of the system in use, and is usually presented in the form of a histogram in the case of a digital picture function. This histogram can be interpreted, given suitable normalisation, as a probability density function for the value of any pixel, but it contains no information about the conformational relationships between pixel values such as those used in edge detection, and does not represent the entire information content of the image. This chapter shows how the grey level or intensity histogram can be obtained from the electronic representation of a chosen image and how it may be modified to effect a contrast enhancement. The requirement for a change in the contrast qualities of an image may arise from many factors but some of the most important include preprocessing of raw data which has low contrast due to sensor or environment and postprocessing of results to render them more amenable to visual interpretation. The presence of atmospheric haze in a satellite image or a background fog in a photograph can cause the image to become immersed in a sea of noise, and some technique is required to extract a useable image from the raw data. Whilst the system described in this chapter can produce seemingly dramatic improvements in contrast it does not possess a magic property which extracts information which was not present in the original image. It can render subsequent processing operations more effective, and can give results a better appearance but in many cases a long hard look with a well practised eye would extract the same information. One of the main objectives of the system is therefore to provide a very fast, interactive and flexible device which allows both the experienced and the less experienced operator to experiment without undue strain or cost. The exact form of the enhancement used depends upon what the user is looking for in the image, and the system must be able to perform a variety of operations to cope with different needs, but it is possible to take into account the intensity response of the eye to some extent and to match the intensity statistic. the image to the visual system using a suitable model.

The design philosophy of the system has been that of fast digital hardware

to carry out image transformations and to gather image statistics, controlled by a microprocessor which can perform statistical calculations at reasonably high speed and can provide the user with a convenient and intelligent interface. This strategy is somewhat similar to that employed by the MEIC which has been described in earlier chapters but the present system works on a digital picture function in place of an analogue signal. The advantages of the specialised hardware are many and include low cost, extreme speed and flexibility, a property which has traditionally been attributed to software-based processors. The design of flexible hardware presents a difficult challenge, but the use of a controlling microprocessor allows the functions of the component parts to be modified with ease. The digital image function is derived from a video picture as will be described in a later section, and the use of a real time source of data requires the response time of the system to be short, of the order of a few tens of nanoseconds. The ability to deal with such fast data is the main reason why the microprocessor is shielded from direct contact with the picture signals by hardware: although it can perform a simple machine instruction in a few microseconds it is several orders of magnitude too slow for direct use.

Unlike the MEIC, the system described in this chapter is not limited to linear contrast operations such as those found on a domestic TV set (brightness and contrast or gain) and can use nonlinear contrast 'stretches' to good advantage. The definition of a 'contrast stretch' will be dealt with in the next section, but the system has been tentatively christened CONDOR, standing for Contrast Optimisation by Nonlinear Digital Operations in Real time.

3.2 Transfer functions and image statistics

All the transformations described in this chapter are spatial domain operations which operate on each pixel separately, i.e. without reference to near neighbours, but whose forms may be influenced by the image statistics. The law which relates the input pixel value to its corresponding output value is a contrast transfer function or input / output curve. Since the possible range of values for pixel intensities is limited (chapter 1) the mapping which defines any contrast modification must result in a set of pixel values which are restricted to the same range. The curve may thus be drawn in a square box whose axes represent input and output intensities, and the curve can take any possible shape within this perimeter. The use of a system such as this requires a feel for those curves which are useful for an application and not destructive.

The simplest form of contrast stretch is a linear function whose lift and gain are adjusted such that the darkest point on the image to be stretched is mapped to black and the lightest to white. If the original image only occupied a narrow range of grey levels this transformation will ensure that the final image will occupy the entire range, although not necessarily in a uniform manner. This stretch function has the advantage that only two parameters relating to the intensity histogram are required: the minimum and maximum values. In practical work however, only a restricted sub-interval of the image's intensity range may be of interest, in which case a full representation of the histogram will be required, and the linear stretch should be applied to enhance only the areas of interest. This selection of appropriate ranges of interest highlights an important feature of all contrast modification processes, namely that information can be compressed or even destroyed in some areas while others are enhanced, and that the image may be interpreted best by a collection of different stretches.

The use of a nonlinear transfer function allows many of the limitations of the linear stretch to be overcome, but the penalty comes in the form of increased complexity, both in the choice and in the specification of the transfer function. It is important to have a clear idea of the object of the stretch, and to have some means of characterising its effects. This problem may be approached by considering the effect of any stretch on the intensity histogram or PDF of the original image. In order to facilitate the calculation it is convenient to restrict the possible form of the transfer function to one of those which increases monotonically over the range of input grey levels. This condition ensures that the relative order of the grey levels is preserved and also allows the calculation of an inverse transformation function. Transfer functions which do not conform to this specification can be used to good advantage but their effects are harder to describe mathematically. Assuming that the histogram is correctly normalised and can be treated as the PDF, then the output PDF can be written as [44]:

$$P_y(y) = [P_x(x) \, dx/dy]_{x=T^{-1}(y)} \quad (3.1)$$

where $P_x(x)$ and $P_y(y)$ are the input and output PDFs respectively and $T(x)$ is the transfer function whose inverse is $T^{-1}(y)$. Equation (3.1) may be used in two ways, either to predict the form of P_y given a measured P_x or to determine $T(x)$ from a given P_x and a specified P_y . The first approach would

be of use if a satisfactory transformation had already been achieved, perhaps by trial and error, and the resulting PDF is required. The second approach allows an image to have its PDF moulded to any desired form, as long as that form can be specified mathematically. There are several standard transformations for which this is possible, such as a Gaussian stretch (which aims to give an image with a Gaussian PDF) or a hyperbolic stretch designed to match the eye, but in the majority of cases it is not possible or convenient (in terms of speed and complexity) to give equations. This difficulty highlights the problem of data entry and user interaction which will be described in later sections.

For the above reasons it is convenient to approach any desired PDF by way of a uniform or flat PDF in which the entire intensity range is occupied with equal probability. If an appropriate transformation which produces this effect can be found it represents a 'best use of dynamic range' transfer function in the absence of any specific areas of interest in the image. If, in equation (3.1), $P_y(y)$ is set equal to 1 over the whole range of intensities (0 to 1 say) the the slope of $T(x)$ must be given by

$$dx/dy = 1/ P_x(x) \quad (3.2)$$

and if the differential equation is solved it is found that $T(x)$ is given by the Cumulative Distribution Function (CDF) of $P_x(x)$, i.e. by

$$T(x) = \int_0^x P_x(z) dz \quad (3.3)$$

where z is a dummy variable for the integration. The CDF obviously satisfies the restrictions which were assumed at the start since it is monotonically increasing and bounded, with a maximum value of 1. The effect of this histogram equalisation or flattening process is depicted in figure 3.1 for a hypothetical image histogram. The most important feature of the transfer function is that it shows a greater slope in those intensity ranges where the histogram has a peak. Thus the greatest contrast stretching occurs in large areas of low contrast in the original image (which give rise to the peaks). In parts of the range where the histogram is sparsely populated the contrast may be compressed, but the output image always occupies the full range from black to white as in the simple linear stretch. In some cases the top and bottom tails of the input PDF are cut off prior to the calculation of the CDF because they are regarded as noise, but usually they make little difference to the form of the CDF.

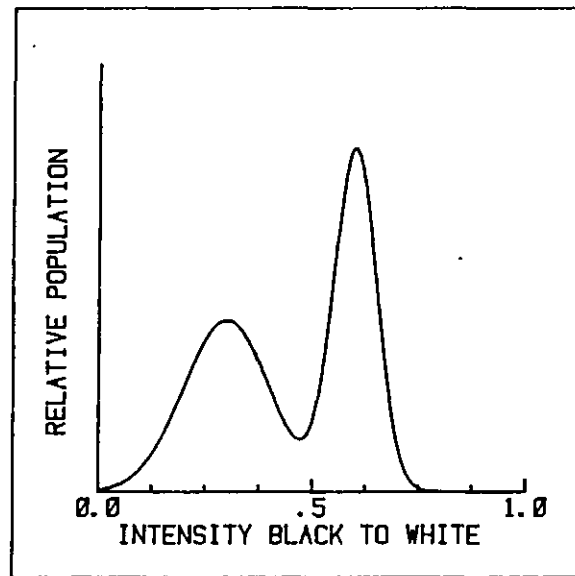


Figure 3.1 a : A hypothetical image intensity histogram

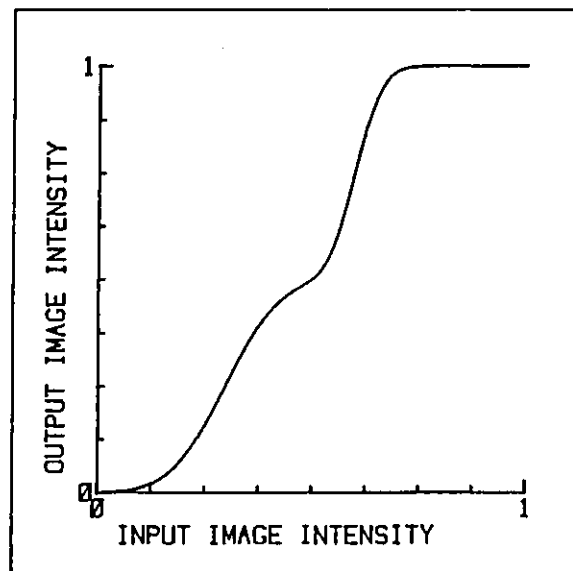


Figure 3.1 b : The corresponding CDF

If the above method is applied to a specified histogram function, the transformation which will give a flat distribution function can be found. If the intermediate intensity variable is z and the specified histogram is $P_y(y)$ then we have

$$U(z) = F [P_y(y)] \quad ; \quad U = 1 \text{ for all } z \quad (3.4a)$$

with $U(z)$ representing the uniform or flat distribution and F the transfer function. Similarly the original image histogram $P_x(x)$ can be flattened by a transformation G :

$$U(z) = G [P_x(x)] \quad (3.4b)$$

Now if G is applied to $P_x(x)$ followed by the inverse of F , the result is an image whose histogram is identical to $P_y(y)$:

$$P_y(y) = F^{-1} [G [P_x(x)]] \quad (3.5)$$

This is the route by which image histograms may be specified directly, assuming that some convenient method can be found to enter the target histogram shape and to calculate the required inverse transformation F^{-1} . In a practical case the two transformations can be combined into a single operation, and inverse transformations are computed numerically rather than analytically.

The system described in the present work uses a digital image function whose intensity values are quantised to 8 bit accuracy (256 levels of grey). This quantisation affects the above theory and prevents the attainment of perfectly flattened histograms. In the discrete case the definition of the CDF given in equation (3.2) must be replaced by a summation in which the discrete output intensities y_i are related to the input intensities x_i by

$$y_i = T(x_i) = \sum_j n_j / n \quad ; \quad 0 \leq i \leq N \quad (3.6)$$

where n is the total number of pixels in the image, N is the number of grey levels, and n_j is the number of pixels with intensity j . The PDFs are similarly redefined in terms of discrete probabilities, and are now true histograms rather than continuous functions. The important point to note is that the transformation T reassigns each input pixel value x_i to a single

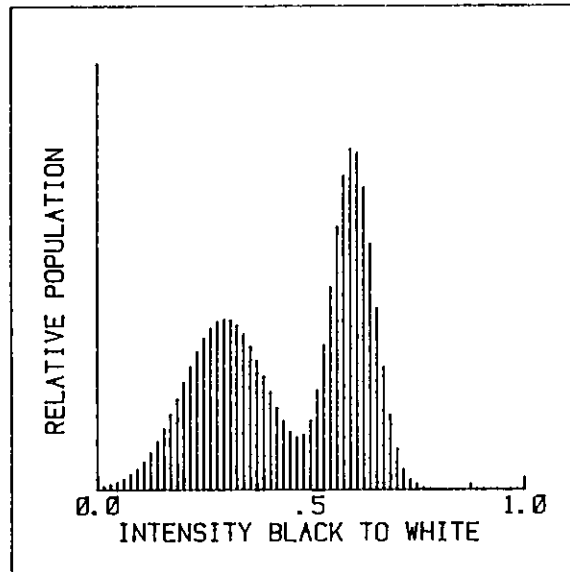


Figure 3.2 a : Quantised image histogram

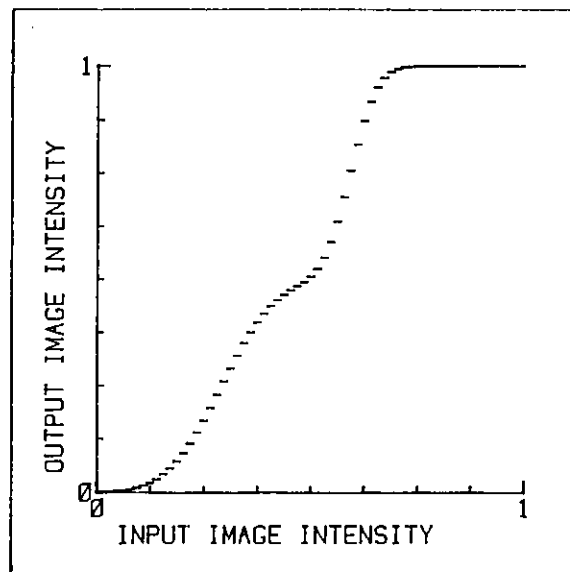


Figure 3.2 b : Corresponding discrete CDF

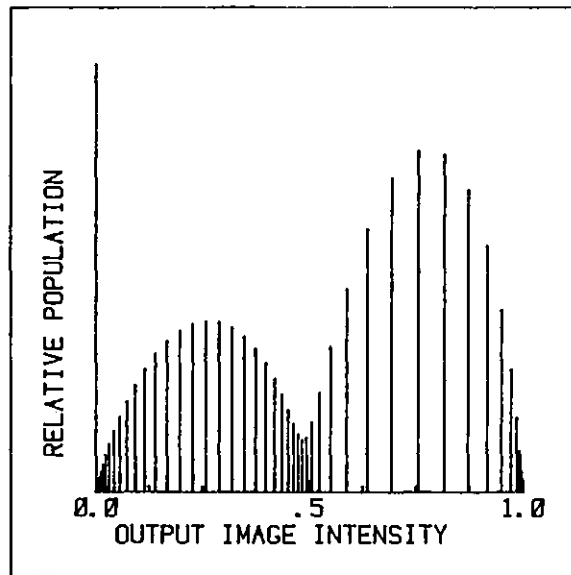


Figure 3.2 c : Final image histogram

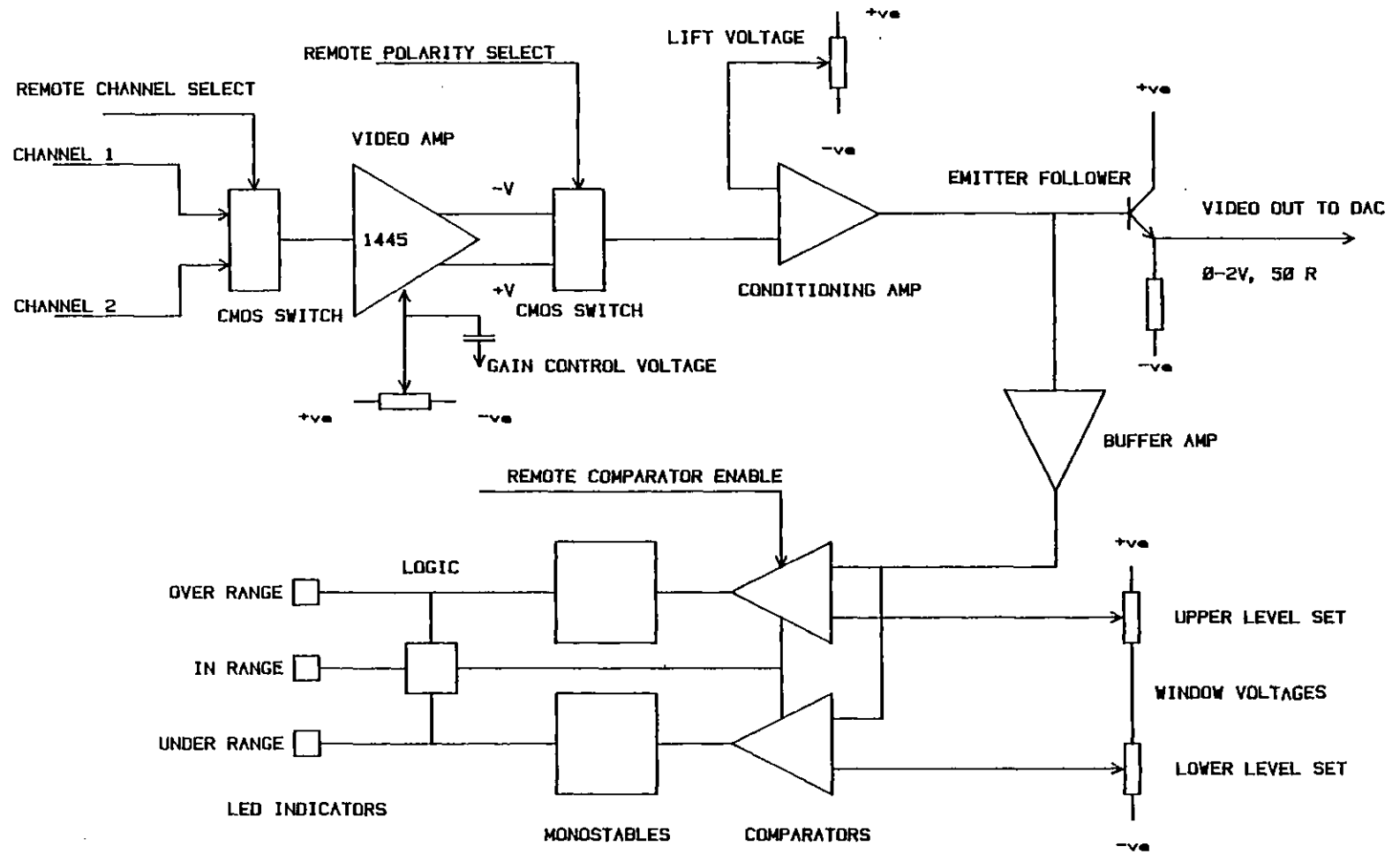
corresponding output value y_i and does not allow any redistribution of pixels within the histogram. The effect of this is illustrated in figure 3.2 which shows the discrete equalisation of the same histogram (or PDF) as figure 3.1. The result is now no longer flat, and the population in each bin is almost the same. The main difference lies in the positions of the bins which are spaced according to their relative populations. The exact value of each bin may change if the transform function has a low slope and rounding errors cause two adjacent bins to be added together. The figure shows a coarse digitisation, and the results can be expected to approach those of the continuous case as the digitisation becomes finer. Some workers [45] have developed strategies for equalisation which do involve a redistribution of pixel values within the histogram, and the usual approach is to consider the area surrounding a pixel to determine the optimum transformation. This method requires that the image be available in a stored form, and while it can be useful for visual interpretation purposes, images which have been processed in this way are not useful for classification.

This section has outlined the basic strategies employed by the CONDOR system for contrast enhancement by non-linear stretching. All of these methods require a degree of operator experience, and are especially powerful when the results can be viewed in after a very short processing delay. The following sections describe the way in which the CONDOR system implements the contrast modifications and acquires the image statistics from a TV standard video signal.

3.3 Digitisation and transformation

The raw data on which the CONDOR system operates is an analogue video signal derived from a VIDICON TV camera or a FSS. As has been noted in earlier sections, non-linear operations are difficult in analogue electronics, and accordingly a digitisation unit has been constructed to provide 8-bit parallel data in real time from the analogue signal. The major element in the digitiser is an Analogue to Digital Converter (ADC) type ADC TV8 B (manufactured by Analog Devices) which is capable of data conversion at rates up to 20 MHz with a 75 nS conversion delay. The input required by this device is a 0 to 2V analogue signal with a 50 Ohm impedance, and since this is not a standard video range (0 to 0.6V) an input amplifier and buffer was constructed according to the circuit shown in figure 3.3. The specifications of this amplifier must be carefully defined, particularly in terms of frequency response, since the quality of the

Figure 3.3 : The video amplifier



CONDOR DAC VIDEO AMPLIFIER

digital data used in subsequent stages is determined here. The amplifier is designed to accept video voltage signals from one of two sources without synchronisation pulses (i.e. blanked video), and it is assumed that the sources are locked to a master timing generator. Connections are made via miniature coaxial cable (type RG174/U) and since this may have significant length it is terminated with a 68 Ω resistor at each input socket to prevent reflections. Selection between the two input channels is accomplished by means of a CMOS analogue switch (type CD4052B) which has a low ON resistance of approximately 80 Ω and a wide bandwidth. The logical selection signal for this device is generated from a toggle switch mounted on the front panel of the amplifier module. This indirect mode of switching allows all high frequency video signals to be restricted to the board rather than appearing on switch contacts which can cause interference and cross-talk, and which can become unreliable for low level signals. The selected video signal is then fed to a wideband variable gain video amplifier IC (type 1445) via an input resistor which comprises both a discrete device and the CMOS switch resistance. The input impedance to this amplifier is arranged so that it is insignificant compared to the termination resistor, thus avoiding excess loading of the video source. The amplifier IC provides a variable gain facility via a control voltage which is derived from a potentiometer across the power supply rails. This control voltage is carefully decoupled to ground with high frequency ceramic capacitors since any variation here will affect the image contrast. The use of a DC control voltage for gain control avoids unnecessary diversions of the video signal as in the case of the selector switch, and gives cleaner, more reliable operation. A second CMOS analogue switch (in the same package as the first) allows the user to choose between inverted and noninverted video signals as appropriate with a single control. The video signals have now been buffered and partially amplified, and a final fixed gain stage brings them to the specified limits. This stage consists of an operational amplifier (LM318H) configured as an inverting adder which combines the video signal with a decoupled DC offset voltage to provide a lift or brightness control. Finally, since the ADC has a low input impedance (50 Ω) and is connected via coaxial cable, a driver transistor is used in emitter follower mode so as not to load the operational amplifier.

The amplifier as described is all that is required, but an extra stage has been incorporated to allow the video voltages to be set up correctly without the use of an oscilloscope monitor. This part of the unit consists of a second LM318 acting as a buffer followed by a pair of fast comparators (NE527K) which

act as a window detector. Two preset levels specified by means of trimming potentiometers correspond to the maximum and minimum permissible ranges of video voltage. The comparators detect any excursions outside this range and their outputs cause a pair of monostables to trigger. These monostables produce a 25 mS pulse (approximately one TV frame) which lights a pair of Light Emitting Diodes (LEDs) indicating over or under range. The monostables allow the detection of very short pulses which would not otherwise be visible on the LEDs. A small amount of logic drives an 'in range' LED whenever neither monostable is triggered. Although this window detector circuit is isolated from the main amplifier by a buffer, and the comparators are fitted with capacitors on their inputs to give hysteresis (and thus prevent high frequency oscillations on slowly varying input signals), some interference is possible from the digital signals. As a precaution an override switch is provided which inhibits the comparators for normal use as opposed to during setting up.

The entire video amplifier is constructed on a single 3U short Eurocard circuit board (100 by 160 mm) fitted with a collander ground plane on the top surface. This provides a stable ground for the analogue signals and prevents circuit instabilities. All power supply connections including ground are made via separate cables direct to the power supplies and are decoupled to ground at each device with ceramic capacitors to prevent noise on the signals. All connections are short except where unavoidable, in which cases screened cables are used. The entire circuit is surrounded by earthed metal screens to prevent interference from the digital sections of the equipment.

The amplifier is used as the input stage for a digitiser, and due account of sampling theory must be taken in designing its performance characteristics. The digitiser is driven at a frequency of 10.25 MHz from a master clock which is also used to synchronise the video source and the rest of the system. This conversion rate provides pixels on the screen which are about 100 nS long and are too small to be seen easily, with a resolution of approximately 530 by 570 over the screen. The 100 nS pixel time is long enough to allow fast computation but it imposes a limit on the maximum temporal frequency which can be resolved by the system, given by the Nyquist Criterion. This specifies that the highest frequency which can be passed by the system is half the sampling rate or in this case 5.125 MHz. If any frequencies above this limit are digitised they cause aliasing errors due to effective frequency shifting by the sample clock. In order to avoid this the amplifier frequency characteristic is made to drop

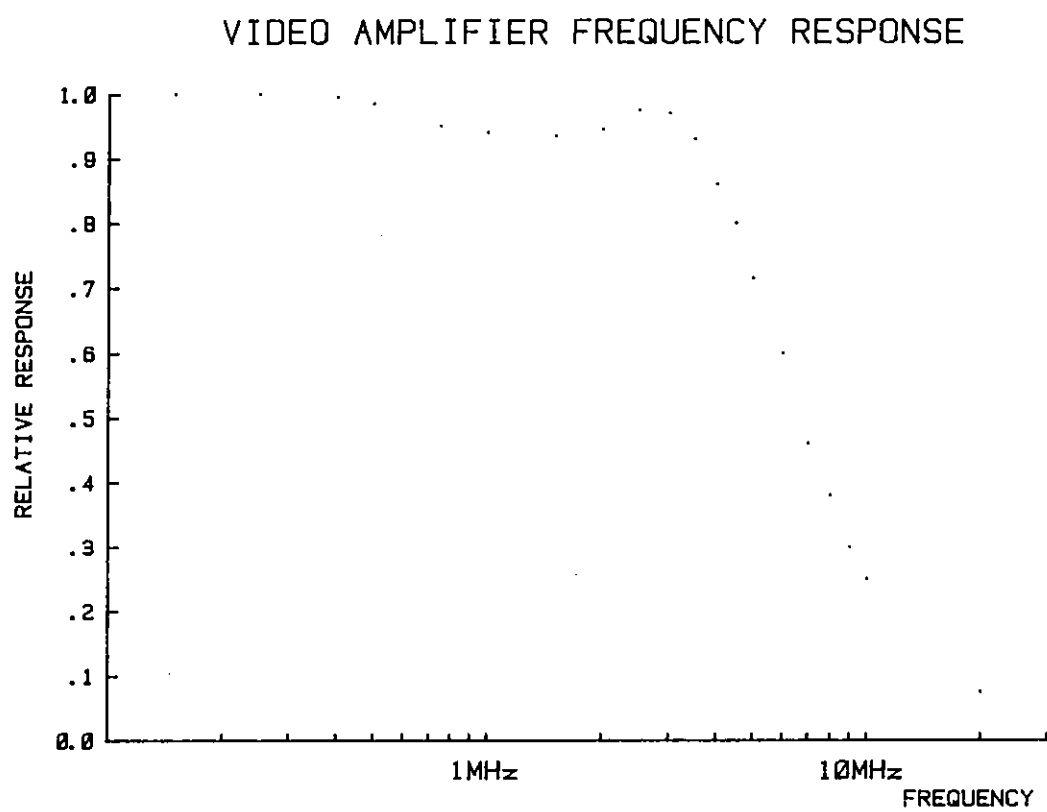
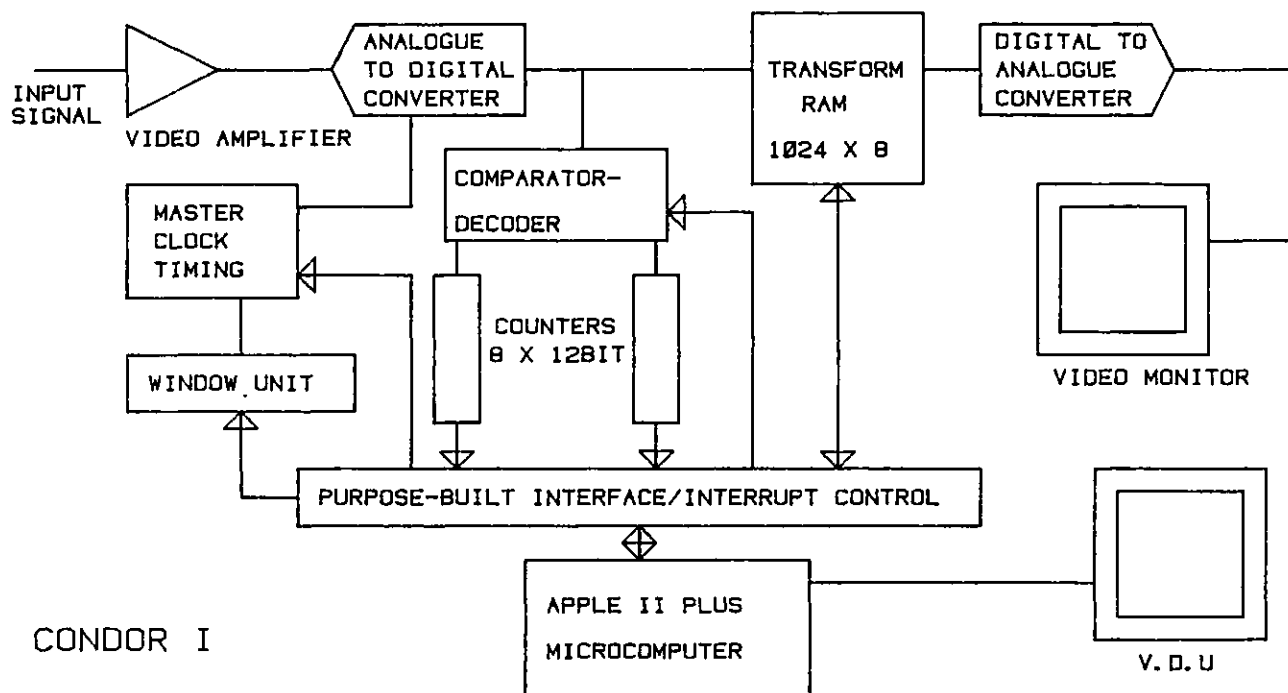


Figure 3.4 : Video amplifier frequency response

sharply above 5 MHz. This is achieved by the use of compensation components in the LM318H OP AMP [46] which introduce a slight resonance at 3 to 4 MHz with a steeper roll-off slope than can be achieved with a simple low pass filter. Another factor which contributes to the avoidance of aliasing is the form of the video signal spectrum which has relatively few components in the high frequency ranges. A graph of the amplifier frequency response is shown in figure 3.4.

The digitiser requires a Start Conversion pulse to initiate a digitisation cycle, and this is supplied from the master clock rising edge via a pulse forming circuit. The pulse causes the ADC to cease tracking the input signal and go into hold mode for a conversion. After 75 nS an End Of Conversion (EOC) pulse is generated by the ADC to indicate the completion of a cycle and the presence of valid 8-bit data at the TTL-compatible outputs (TTL is Transistor-Transistor Logic, a family of digital circuits used in this equipment). The conversion technique is a hybrid between flash and successive approximation, and further details are to be found in [46]. All subsequent operations are synchronised to the falling edge of the EOC pulse or to the inverted sampling clock's rising edge which occurs a little later. The Start Conversion pulse is also a low impedance input, and requires the use of a high power TTL driver to match the line.

An overall block diagram of the CONDOR system is shown in figure 3.5, and the section described above is indicated at the start of the diagram. The resulting digital data stream which represents the video signals is fed to two main circuits: the histogram acquisition section and the transform RAM. The transform RAM is the 'central processor' of the system and performs all the pixel intensity transformations with great speed. An expansion of the system diagram is provided in figure 3.6 which shows this section in more detail. The RAM (or Random Access Memory) is a type MK4801A-55 1k by 8 bit device with a typical address access time of 55 nS. This means that data is available on the Data I/O pins 55 nS after a stable digital address is applied to the Address pins. Since the video data is 8 bits wide, with 256 possible grey levels, the corresponding output data must also be of the same format, thus allowing the entire transform to be specified by a 256 byte 'page' of values. The transform is calculated by the system microprocessor in advance and is stored in the transform RAM. If the digital picture data is applied to the address pins (thus selecting the input grey level) a new intensity value will emerge from the data



CONDOR I

Figure 3.5 : CONDOR system block diagram

pins as an 8-bit byte. This type of operation is sometimes referred to as a 'look-up table' and it lends itself to very rapid computation in situations where the number of possible input and output values is limited. Since the RAM is a 1k device, four 'pages' or transformation curves can be stored at any time, and selection between them can be accomplished by means of the two most significant address lines A8 and A9 as in figure 3.6 where they are controlled by the microprocessor via a latch. The arrangement of buffers on the lower 8 address lines is designed to allow the microprocessor to gain access to the RAM without contention from the digitised video data. The buffers are tristate (i.e. their outputs can be made to assume a high impedance state as well as the normal digital low or high states) and the output enable pins are driven from a control latch as above. The inverter between the two buffers ensures that one of them is always off or in a high impedance state.

In operation the sequence runs as follows: firstly the address lines are isolated from the ADC and connected to the microprocessor using the buffer arrangement. Secondly the buffer on the data lines is enabled, connecting the RAM to the microcomputer's data bus. Data is now written into the RAM one page at a time according to the state of the upper address lines. Lastly the ADC is reconnected to the RAM and contrast modification recommences. This circuit allows any of the four possible transformations to be selected at will simply by writing an appropriate digital word to the latch controlling A8 and A9. This switching process can occur at any time to allow comparisons to be made, and may even be switched at selected times inside a TV frame to give different enhancements to different parts of the picture. During the microprocessor's access to the RAM the data lines carry apparently 'random' picture data which changes states at a rate of 1 MHz which is the computer clock frequency, and this appears as noise on the viewing screen. This noise could be replaced by a blank field by adding a tristate buffer on the output of the circuit and disabling it at the appropriate moments, but this would still cause the screen to flash. In order to overcome this problem all access to the RAM by the microprocessor is restricted to the vertical blanking time within the TV field when the picture is normally blank. Since blanking is applied at a later stage and does not require a set of zero picture values, no flicker is seen on the screen. Writing speed is impaired somewhat, but a single page of 256 bytes can be written in four blocks of 64 bytes in four field periods or 2/25 of a second. If a second RAM were inserted into the system with an identical transfer function it would be possible to write to one RAM while the other was addressed

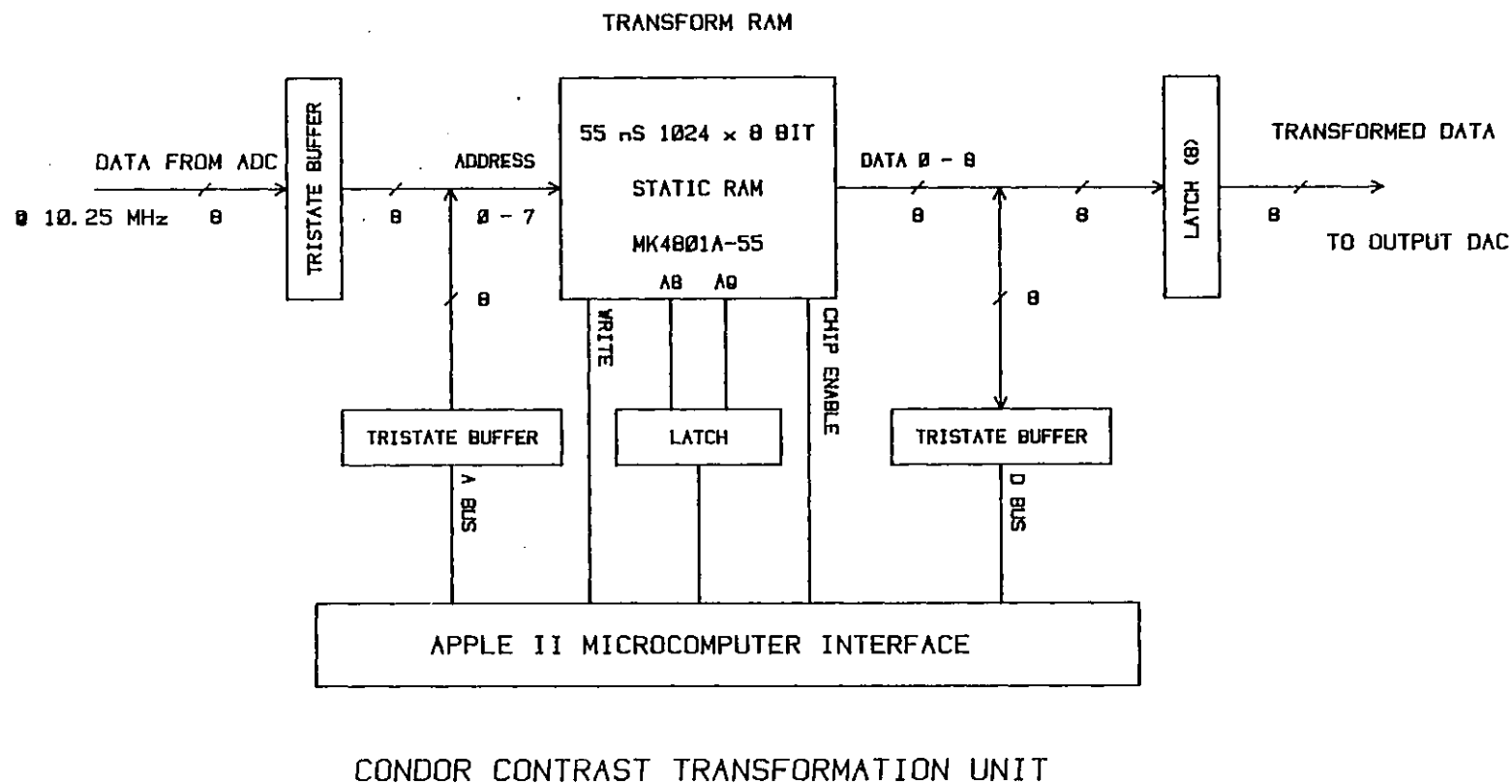


Figure 3.6 : CONDOR contrast transformation circuit

by the video data but this strategy increases the complexity of the circuit and does not give extra space for different transforms, so that in practice the speed improvement is not significant. The total delay through the system as it has been described so far is a combination of the ADC conversion time and the RAM access time, and amounts to approximately 130 nS.

3.4 Video signal reconstruction

The overall block diagram in figure 3.5 shows that the data output from the transform RAM is fed to Digital to Analogue Converter (DAC) so that a normal video voltage can be recovered for display purposes. The simplified circuit for this section appears in figure 3.7. Data from the RAM is clocked in with a very fast TTL latch manufactured using ALS (Advanced Low power Schottky TTL) technology as opposed to LS (used elsewhere in the system). The reason for this latch is to deskew the data prior to conversion. Deskewing causes all changes in the data between pixel values to occur synchronously even though the ADC and its connecting cables may have introduced random delays between the different bits of each word. If the data is not deskewed then a late (or early) transition on some of the data lines can cause the DAC to output a 'glitch' of erroneous signal for the time between the first and last transitions, and since the DAC employed here is very fast (under 25 nS settling time for a full scale step) these glitches can be of significant amplitude. A similar effect can occur in the internal switches of the DAC and a good quality device must be used. In this case the DAC HF8 B from Datel is used [47]. In many commercially obtainable DAC systems it is common practice to use a sample / hold device immediately following the converter to eliminate the effects of such spikes, but since this equipment was primarily intended for viewing rather than as a preprocessor, and since the level of glitches was found to be insignificant such a device was not used. In any future version of the system which is intended for classification work a DAC with an integral sample/hold would be employed.

The DAC gives an output which is the product of the reference source with the digital input word as a fraction of 256, and this signal is buffered by an OP AMP which scales it to standard video ranges. The signal has no blanking or synchronisation added since the digitiser is run continuously to avoid instabilities and drifting. These signals are derived from the timing generator via divider chains and are added to the video signal after suitable scaling. The blanking signal cannot be added directly as can the sync. pulses, so a CMOS

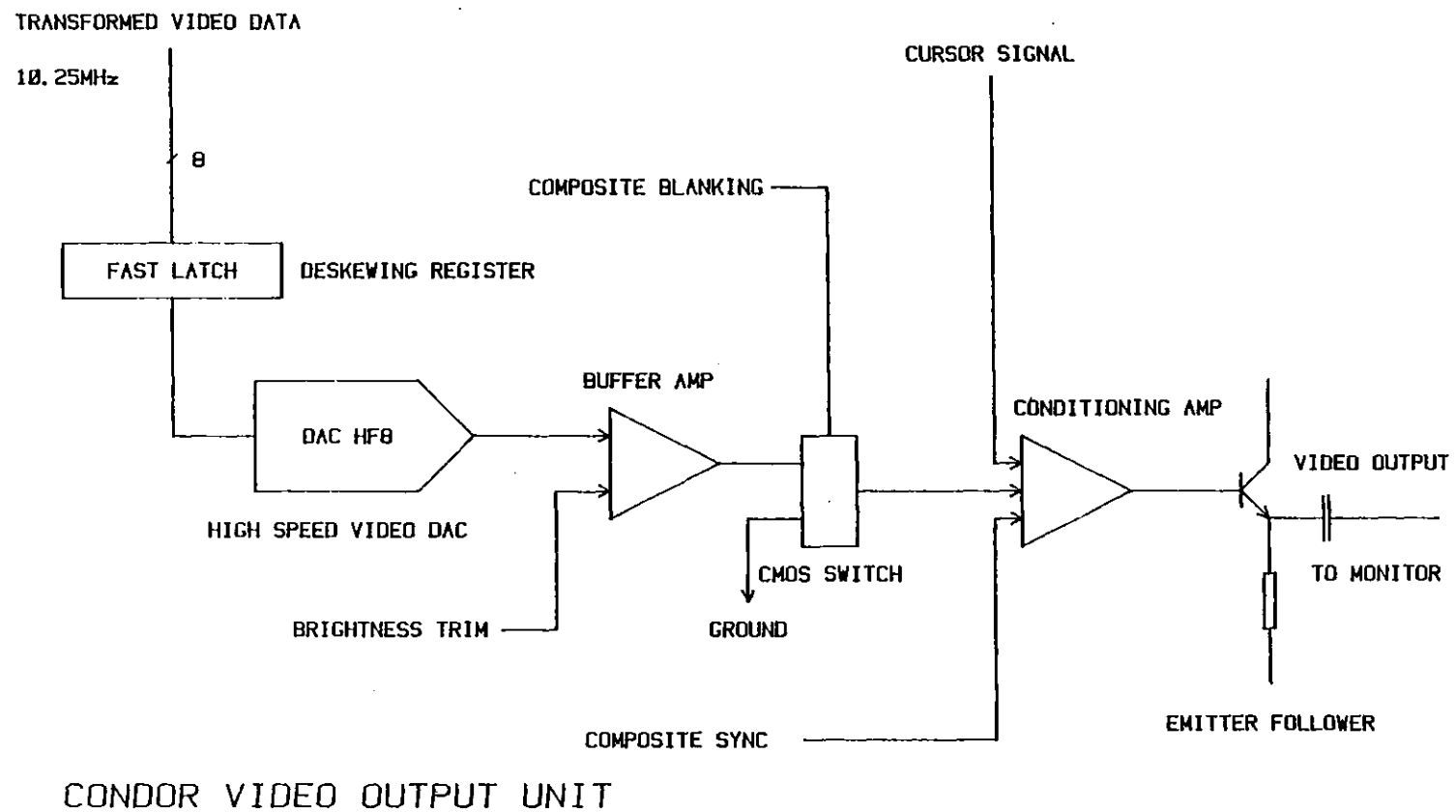


Figure 3.7 : CONDOR Video signal reconstruction circuit

switch is used to chop the video signal, switching the input of the operational adder alternately between signal and ground using the blanking signal as the switch control. A further CMOS switch allows synchronisation pulses to be added or not as desired in case external sync. is to be used. The same operational adder is used to mix in a window signal or cross hairs from a generator which will be described below, thus allowing a cursor to be overlaid on the video image. The output gain and lift can be trimmed with preset potentiometers on the board and an emitter follower transistor is used to provide drive to connecting cables as in the video amplifier described above. These cables must be correctly terminated to avoid any unwanted reflections.

Construction is again on a ground plane Eurocard placed in a screened enclosure, but particular care must be exercised in routing the digital signals on the board so that they do not cause any interference in the analogue output. Analogue and digital grounds are kept separate and great care is taken to decouple all power supplies at each device, especially the DAC HF8 itself. The frequency response of the circuit is primarily limited by the performance of the operational amplifiers, and these are given compensation to extend their range to about 5 MHz, giving a response similar to that of the video amplifier.

3.5 Histogram acquisition

The histogram collection circuit forms the second major part of the CONDOR system after the transform RAM. In the block diagram (figure 3.5) it is indicated by a comparator / decoder and a bank of counters. Before a detailed description is given it is worth considering the possible design strategies for this section. The major factors which must be taken into account are the sampling time, sampling density or size, circuit speed and complexity and the possibility of windowing. Taking the extremes of sampling time, it is possible to construct 256 counters and to increment one of these according to the value of each pixel. This arrangement would acquire the histogram in one TV frame ($1/25$ S), but it is expensive on counters. The counters could be replaced by a wide word memory device and the incrementing operation would be modified to a read-increment-write cycle (sometimes called a read-modify-write cycle). The length of each counter would have to be such that a very low contrast picture would not cause overflow, and both the memory and the incrementing circuit would need to be very fast and thus expensive. At the other extreme it is possible to use a single counter which scans for a single grey level in each frame, taking

a total of 256 frames or about 11 seconds to collect the entire histogram. This is somewhat slow and is also vulnerable to drift and noise in the picture which might not be steady for such a time. The speed restrictions on any counter system are relieved if sampling is carried out for every other pixel or by some other similar rule. This approach is valid as long as enough pixels are gathered to form a valid statistical sample, but the form of the histogram changes over the image and a sparse sampling strategy does not easily allow the use of a window on a picture to pick out areas of special interest. Obviously these extremes also represent a considerable variation in complexity from a simple single counter to a large fast memory system.

In the CONDOR system a compromise solution has been adopted which aims to scan the picture for groups of 8 adjacent grey levels per frame, requiring 32 frames or just over one second to acquire the entire histogram. The circuit accordingly has 8 counters which are incremented by a special pixel value detector circuit. The values on the counters are read into the memory of the microprocessor at various times during the frame thus reducing the required size of counter to a manageable 12 bits (maximum count 4095). This compromise achieves an acceptably short sampling time and a reasonable immunity from noise although it can be affected by 'roll bars' as will be described later.

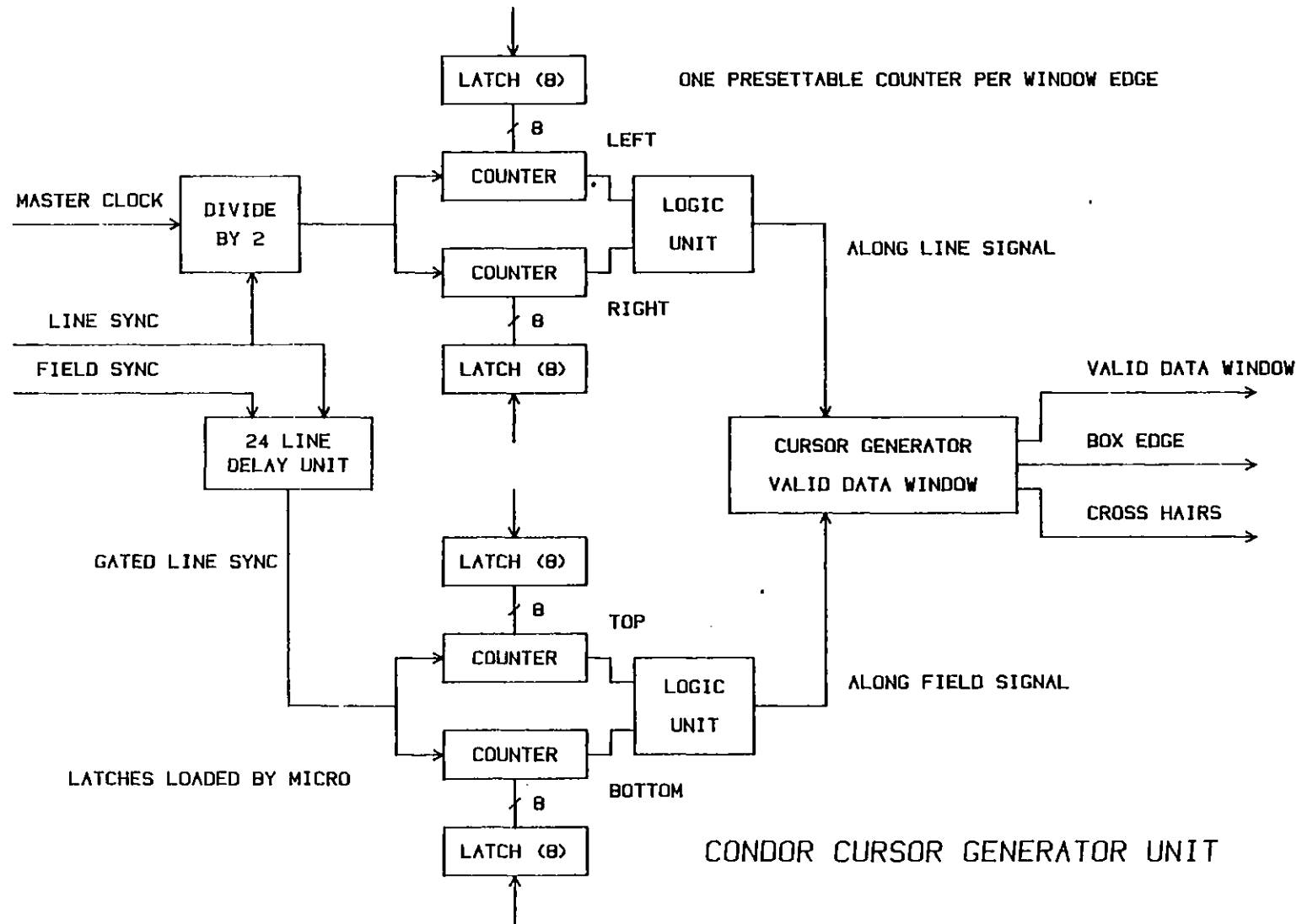
A simplified block diagram of the histogram acquisition circuit is shown in figure 3.8, but only one quarter of the counters are depicted for clarity since the other sections are identical. The circuit requires the active participation of the microprocessor for its operation and forms a mixture of fast logic and computing power. The microprocessor first supplies a 5-bit mask word and stores it in a latch. This word forms one of the inputs to a 5-bit digital magnitude comparator formed from a 4-bit device and an exclusive OR gate, and the other input comprises the 5 most significant bits of the digital picture data from the ADC. The mask word is started at zero and incremented after each frame (or two fields) up to a maximum of 31. The magnitude comparator provides a pulse whenever both of its input words are equal, and this determines the value of any pixel to within a range of 8 levels, with the lowest three bits giving the exact level. The equality signal is used to enable a 3 to 8 decoder which is fed with these three data bits. This device decodes the 3-bit binary word formed by the data bits and gives a pulse on one of 8 output lines according to its value. These lines are connected to the count inputs of the 8 counter chains. The decoder is enabled also by the window signal which limits the area of valid

data, and by the DAC conversion clock to ensure correct synchronisation. The total delay between a pixel value becoming valid and a count pulse is only about 30 ns which is well within the 100 ns pixel duration. In order to economise on chips the counters are formed from 4-bit asynchronous devices which come two to a package and are not presettable. Three such circuits are used for each counter and the clock inputs of the higher counters are chained from the previous devices. The figure shows two 12-bit counters (occupying 3 counter packages) and three 8-bit data latches which collect the results from the counters, with one latch being split between two counters.

At the start of a sampling cycle the microprocessor clears a 256 word by 12 bit data buffer in main memory and initiates counting. After each 8 valid lines the micro is interrupted by an interface circuit which also latches the data from the counters. During the next 8 lines the micro unscrambles the mixed data from each set of three latches and performs eight 24-bit addition operations to update the histogram data in the buffer. If the particular set of eight lines was completely outside the specified window boundaries then this stage is skipped and the next set of lines is examined. After one frame the mask word is incremented so that data on the next eight grey levels can be acquired. The operating software for this process is written in machine or assembler code for speed and is just fast enough to process each set of data before the next set becomes available.

The histogram acquisition circuit described above uses all the pixels to contribute to the results and does not operate a sparse sampling scheme. It can however be restricted to gathering data from a window or subscene on the image, and this is accomplished by means of a cursor generator circuit of the form depicted in figure 3.9. The screen is divided into 256 lines each with 256 columns for this circuit and the user can define a rectangular box or window by specifying the positions of its sides. These positions are stored in data latches which provide the preset inputs for two pairs of counters, one pair for the top and bottom and the other for the left and right sides of the box. At the start of each field the top / bottom counters are preset and they begin counting line pulses from the end of the vertical blanking period. In a similar way the left / right counters are preset at the start of each video line and they count at half the pixel rate from the end of line blanking. When the counters reach zero they output a 'borrow' pulse which is fed to some logic whose purpose is to generate an 'in box' signal. This signal indicates the

Figure 3.9 : CONDOR window / cursor generator



presence of valid video data to the histogram circuit of figure 3.8 by enabling the decoder. The same logic can be used to generate both a cross hair cursor and a 'window edge' signal which is displayed as a bright perimeter on the viewing monitor, thus allowing the operator to locate the window accurately. The operation of the window generator is delayed for 24 video line periods from the field pulse (approximately the vertical blanking time), preventing the acquisition of data from blank lines and the loss of the top of the window on the video display. The board is fitted with a signal selector circuit which can display any of the above signals through the DAC circuit of figure 3.7 as required. The microprocessor performs any resetting of the window generator's latches immediately after the field blanking pulse so that the data in the latches is always stable when the counters are reset. This avoids the display of broken lines or the loading of spurious box coordinates.

3.6 TV camera interface

The source of the video data is a VIDICON-based TV camera manufactured by Link Electronics. The various timing signals required by the CONDOR system were derived from the internal circuitry of the camera by suitable transistor drivers and connections were made via a multiway plug and miniature coaxial cable. The signals used are mixed blanking, synchronising pulses, line and field drives and blanked video. The entire camera was originally locked to a 50 Hz frequency by means of a phase sensitive detector which used the A.C. supply as a reference for internal clocks. This circuit was disabled and a master clock based on a 20.5 MHz crystal oscillator was installed. By means of suitable division this clock provides the 10.25 MHz sampling frequency, the 5.125 MHz window unit frequency and a 31250 Hz signal (twice the video line frequency) which is fed into the camera's internal dividers to generate the field drive and blanking. The entire camera circuit is operated with reference to a -12 V ground rail, and the video signal is output via a capacitive coupling, thus necessitating the use of level shifting. The clock drive to the camera was fed through an opto isolator, and all the camera signals were shifted to standard TTL levels by means of FET input stages and Schmitt trigger gates.

The use of an external reference clock allows the entire system to be locked in frequency but it emphasises the shortcomings of the camera's video circuitry in terms of power supply rejection. The supply contains a small amount of 100 Hz ripple from the rectifying section, and this appears on the

video signal. In the unmodified camera this is not objectionable since it occurs in synchronism with the field frequency of 50 Hz, and it is not easily visible. When the external clock is used it generates an accurate 50 Hz signal which is not locked to the mains supply, and the ripple appears as a 'roll bar' moving up or down the display. This variation in the base level of the video signal will affect the accuracy of any histogram which is collected in both cases, but when using an external clock the effects are not spatially constant. As the sampling circuit takes data for successive groups of intensity levels the roll bar can 'follow' its progress and artificially raise or lower the video signal used for the next set of eight levels. This will produce a blurring of the histogram, and in exceptional cases where a large peak is being detected from a noisy signal it can lead to the acquisition of two peaks. The effect of a roll bar is not constant, and drifts with changing supply frequency. In an attempt to reduce this problem after the supply to the camera had been better regulated, a circuit similar to a valve 'hum dinger' was constructed. This circuit consists of a low frequency amplifier which is fed with the offending 100 Hz signal and which injects an inverted portion of this signal into the video amplifier to cancel out the ripple. The required components were mounted on a small board attached to the video amplifier and the gain was adjusted for optimum rejection using an oscilloscope. The roll bar signals were considerably reduced and the resulting histograms are more accurate.

It must be noted that the VIDICON tube itself does not give an even response over its entire sensitive area, with reductions in sensitivity occurring at the edges of the field and in areas where the target has been 'burnt' by exposure to excessively intense images. The use of small windows around the screen must therefore be treated with caution, but the normalisation and quantisation that occurs in a process such as equalisation can render small errors insignificant as long as the overall histogram shape is preserved.

3.7 The system microcomputer.

As has been noted in the above sections the CONDOR system is controlled by a microprocessor which enables flexible operation of the hardware circuits. Several factors must be taken into account in choosing the best system for this purpose, and some of these are set out below. It is important that the system should be readily amenable to expansion and interfacing, with all the relevant signals available easily. The processor must be accessible for programming in

both a high level language such as BASIC for operator interaction and higher mathematical operations and assembly language, since many of the functions require extreme speed of response. The system should support reasonable quality high resolution graphics to allow the displaying of results and it must also support a disk drive for speed of storage / retrieval. It would be possible to design and construct such a system from scratch, but this almost amounts to a separate project, and for this reason it was decided to purchase a commercial unit.

The system chosen was an APPLE][PLUS, equipped with 48 kBytes of RAM and a single disk drive unit. This machine supports Applesoft floating point BASIC and a useful monitor program which allows ready access to machine code programmers. Expansion capability is provided in the form of seven peripheral board connectors which carry the processor clocks, address and data busses as well as decoded address signals specific to each slot. Graphics is available at 280 by 192 resolution in two pages although these are mapped into the centre of user RAM space. The microprocessor unit is an 8 bit type R6502 manufactured by Rockwell or MOS Technology, and although it does not have some of the features of more recent processors such as the 6809 or the 16-bit chips it does have versatile addressing modes and is relatively easy to program.

The interface to the CONDOR system is constructed on a prototyping card and incorporates a bidirectional 8-bit data bus transceiver / buffer to drive signals along ribbon cable connections. Address decoding is provided which gives selection signals down to the individual byte level. In order to economise on connections, some of this decoding is situated on the CONDOR main boards. Control of the system is accomplished mainly by data latches which can be written to or read from by the microprocessor. The bytes contained in these latches may be entire data words or individual bits may be used to control different functions. The other type of control is the 'soft switch' which consists of a flip flop whose state can be toggled or cleared by accessing certain addresses in memory space. This type of interface is particularly suited to data selection and routing applications. Since the microprocessor is required to operate in synchronism with the video signals during histogram acquisition and transform curve writing for example, some form of timing signal is required. This could be achieved by programming the microprocessor to read a memory location to assess the status of the system, but this is both slow and time-wasting in that the processor could be otherwise occupied while it is

reading the status and waiting for a timing signal. It is possible that the timing signal could be missed entirely, and in order to avoid these problems it was decided to run the system by interrupts. An interrupt is a signal which is applied directly to the microprocessor chip and which causes program execution to jump to a specified address on completion of the current instruction. The address contains a pointer or vector to an appropriate interrupt handling routine which can initiate the correct action depending on the source of the interrupt. The 6502 has one maskable interrupt request line (i.e. the interrupt can be ignored by the computer if so programmed), and an external data selector has been constructed to allow the use of two or more interrupt sources. The two sources are the field drive and the output of the 8-line counter which signals that the histogram counters should be read. These signals, after selection, drive a monostable which gives a uniform 10 to 15 μ S pulse to the micro. Interrupts may be further inhibited by clearing the data selector. In operation the micro is performing a task or waiting in a loop when it is interrupted, and the current program counter and status register are saved. On completion of the interrupt servicing routine the processor can resume its previous task.

User input to the system can come via several methods including a standard QWERTY keyboard on which commands or responses to prompts can be typed, and push button inputs designed for video games. In order to provide a more useful interface, particularly for the direct specification of histogram shapes or transfer functions (section 3.2), a special purpose switch box was constructed. This consists of a small sloping console attached to the computer by a multiway cable and equipped with a set of single pole double throw centre off switches and some light emitting diodes to act as indicators. The pole contact of each switch is earthed and the two outer contacts are connected to two inputs of an 8-bit tristate data latch. Four switches are attached to each latch in this way, and five latches are used, thus giving 20 switch positions. If a switch is in the centre off position, both latch inputs will register a logical '1' but if the switch is moved one of the bits will be a logical '0'. The five bytes corresponding to the data latches are read by the microprocessor which checks the data to determine the positions of each switch and writes the results into a table in memory. In order to aid the operator in assessing the status of the system quickly several LED's are provided, and these are run from the individual bits of a sixth data latch. One LED indicates that the switch box is active, and others indicate the need for input and which switches to use. Seventeen switches are allocated to the input of histogram or transfer function data and

the other three control the window / cursor and general input. The switches can be sampled quickly and give fast responses, and since they do not feed directly into counter circuitry they do not require hardware debouncing.

The input of data from switches is convenient and relatively simple, but their use for incrementing or decrementing a parameter stored in memory implies the use of a software counter. An initial design for this data entry device employed sliding potentiometers, but these give analogue signals which have to be digitised before use, and they are not very linear or accurate. In use a movement of the slider can be detected but it must be allowed to settle before a reliable position can be established, and the quick jumping from one position to another can be simulated in software that reads switches.

The APPLE computer does support graphics as was noted above but these graphics have several disadvantages which slow down the operation of the system, although they were useful for starting with. Each graphics page occupies 8 kBytes of storage which is situated in the centre of the available address space, thus limiting the possible length of any BASIC program. The resolution is only medium and it forces all graphs to be scaled down vertically. The complex nature of the graphics circuitry which is designed to produce colour displays on a domestic TV set results in somewhat slow graphics routines and the production of some spurious artefacts on the display. In an effort to overcome these limitations and to gain experience for the video framestores described in a later chapter it was decided to construct a video interface board with separate video memory. This board provides 512 by 512 pixels with one bit intensity resolution (on-off) contained in 32 kBytes of dynamic RAM, and is controlled by a Graphics Display Processor (GDP) type EF9365 [49], manufactured by Thomson-EFCIS. This chip, which has recently become available, contains circuitry for the refreshing of the dynamic RAM chips and for writing vectors and characters into the display at a peak rate of approximately 1.5 million pixels per second. Vectors and characters can be freely mixed on the display and the GDP is autonomous in operation, requiring only simple commands from the APPLE, and it signals the completion of the current command so that another can be initiated. The entire graphics system occupies only 16 bytes of memory space and releases the two old graphics screens for programs or data storage. In comparison with the APPLE graphics, a 256 bin histogram can be written within a frame period as against several seconds for the old method, and four such displays can be viewed at the same time, thus giving the operator the facility

for quick comparisons. A more detailed description of the operation of framestore memory systems will be given in a later chapter.

3.8 System software

This section outlines the overall form of the software which runs the CONDOR system and detailed descriptions of the major functions will be given in a later chapter. The operator interacts primarily with an Applesoft BASIC program which provides various processing options in the form of a menu. At each stage the operator can choose the appropriate course of action. In future versions of the software it may be possible to implement a simple command interpreter which controls the processing according to instruction codes input by the operator but this is more complex and harder to adapt to. The operating program has an input stage for the major options and a set of subroutines which perform the desired tasks. BASIC is used mostly for higher mathematical functions and for user interaction via the text display, but since the GDP board has been commissioned these tasks have become increasingly amenable to machine code implementation. The BASIC program can affect the status of the machine directly by means of memory 'POKEs' (writing a given value to a designated address), and this is useful for slow, single changes such as turning on the window display.

This form of control is not satisfactory for more demanding tasks, particularly those for which synchronous working with the video signals is necessary. In such cases programs are coded in machine or assembly language with the aid of an assembler program. This allows the production of annotated listings written in mnemonic codes which can be edited prior to transformation into binary files. Interaction between the BASIC controlling program and machine code routines is achieved with the aid of a subroutine-type CALL statement which transfers control to a machine code routine at a specified address. All the machine code is written in the form of subroutines, some of which are not designed for direct access from BASIC but rather for calling from inside other machine code subroutines. The functions supported include data acquisition, normalisation and display together with user interaction (e.g. reading and decoding the switch box). The machine code routines are located at the start of the second graphics screen and extend upward in memory, occupying currently in excess of 2.8 kBytes ($1 \text{ kByte} = 1024 \text{ bytes} = 2^{10}$). They are protected from overwriting by the BASIC routines and since this graphics screen

is not used they will not be erased by any display functions.

The first major routine is concerned with the acquisition of histogram data from the hardware described in section 3.5. A brief description of the operation was given in the above section, but there are some points worthy of note here. The acquisition cycle which reads out the data every 8 lines down the TV frame must be extremely fast since it has to unscramble the mixed 12-bit data from the latches and perform eight 24-bit additions within 512 μ S as well as keeping track of the number of frames to sample and the position of the window. Prior to entering the data gathering routine the window edge positions are tested to determine which of the 8-line interrupt signals will be followed by valid data and which should be ignored. This information is written in a table of vectors in memory which are written in turn into the offset byte of a relative branch instruction situated at the start of the data collection routine. If the data is valid the vector causes a branch to the next location and data is processed, but if not it points to the end of the routine, and processing is skipped. The program is thus self-modifying, with the decisions made before the actual acquisition.

After the entire 256 intensity levels have been scanned (each for two fields) the resulting data is in the form of a three-byte wide table of 256 words and it is unnormalised. Normalisation requires a division operation and the isolation of the maximum value in the histogram, and this is achieved in two stages. Firstly the data is repetitively multiplied by two by shifting each 24-bit word one digit to the left in memory. This is a rapid process and is stopped when at least one data value has a logic 1 in its Most Significant Bit (MSB). At the same time the maximum value is extracted. In order to make the data available to BASIC routines the area of memory used for the storage of variables is filled directly with the results, thus saving a slow BASIC data reading routine. A normalisation factor is calculated by a division routine and the data is normalised by multiplying it by this factor. The entire process takes only a fraction of a second when coded in machine language, and the delay is insignificant compared to the actual data collection time.

The second major routine performs a numerical integration on the histogram data to give the corresponding CDF. This is then normalised by the same routines as above with the aim of setting the last byte equal to 255. The routine simply consists of a repeated 24-bit addition program, and the execution

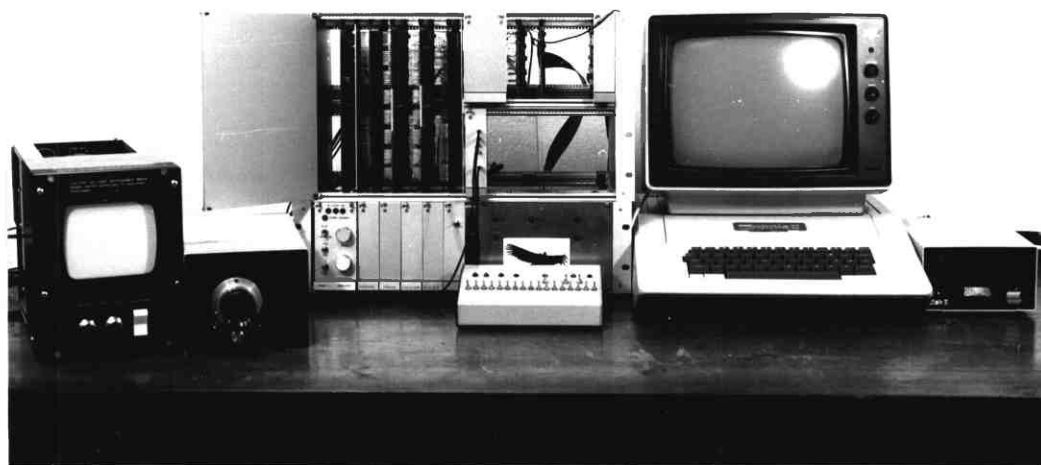


Figure 3.10 : The CONDOR system (lower rack)

time is again minimal. Other routines include various utilities which permit the transformation RAM to be loaded or unloaded during vertical blanking, and which produce a linear transfer function which does not modify the input video data. Various graphics routines have been written which allow histograms and curves to be displayed conveniently.

The third major set of routines is concerned with the direct specification of histograms and transfer functions for interactive contrast stretching, and the use of these programs will be described with examples in the next chapter. Transfer functions and histogram shapes are approximated by straight line segments joining points specified by the switch box. In order to define transfer functions it is necessary to generate the coordinates of points on these line segments, and an algorithm originated by J. Bresenham [50] has been implemented which gives the most accurate approximations. Routines have also been written to invert any given transfer function prior to histogram shape specification.

3.9 System summary

This chapter has described the design principles and the construction of a nonlinear contrast stretching system which operates on digital video data at TV rates. The concept of a transfer function linking input and output intensities has been introduced together with some of the methods by which they may be derived. The transformation has been implemented in the form of a fast memory with space for four different contrast stretches, thus allowing fast comparisons to be made. The design of a subsystem for the collection of image intensity statistics has been described and the operation of the system software in processing the data has been considered.

The entire CONDOR system is shown in figure 3.10, and consists of a 19" Eurocard rack frame containing power supplies, the video digitiser, histogram and transformation circuitry. External to this rack are the GDP and associated graphics system, the switch box and the APPLE computer with attached disk unit. Apart from the analogue signal boards, all construction has made use of wire-wrapping techniques and fast TTL logic chips mounted on high density boards. The following chapter will present some examples of contrast stretching on CONDOR as an indication of its as yet little tested capabilities.

4.1 Tactical image processing

This chapter presents some examples of images which have been processed interactively using the CONDOR system described in the previous chapter. Before demonstrating the capabilities of the system it is useful to consider the motivations and goals behind any such processing so that the combination of machine and operator may be disposed to the best advantage. Once the aims of the processing have been defined, the capabilities of the system can be fitted with those of the operator to give a useful result. In the present case, an original image is presented for processing (i.e. a contrast stretch) and a modified image is output which, it is hoped, shows some improvement. The exact nature of the improvement depends on the particular image and that part of the information contained therein which is of use to the operator. It may be that the modified image shows up some detail more clearly or that the results of some subsequent process are improved when the modified image is used as an input. In the present case no information is gained, and it may indeed be lost by the compression of certain grey levels since the processing consists of a redistribution or re-presentation of intensity information. The image which is subjected to the processing has usually been judged to be deficient in some way, and must be enhanced, but perhaps the most important underlying point about any image enhancement is that there exists no 'best' image, only images which are useful to their users or interpreters. The enhancement applied, even to the same image, depends critically on the interests of the interpreter. In dealing with processed images great caution must be exercised in the choice of stretch, and any blind following of 'popular' enhancement operations is to be avoided. A further caveat must be given here: an image which is presented for enhancement is usually of poor quality (perhaps noisy or with poor contrast) and does not have good 'cosmetic quality'. Cosmetic quality or prettiness is not the aim of the serious image processor, but is a welcome secondary bonus after the required information has been extracted.

Much effort has been expended in the search for automatic processing machines which can be set up, fed with images and left to crunch data without supervision. In the face of enormous quantities of data from ever more sophisticated imaging systems this approach cannot be overlooked, but unless the machines can be made highly intelligent their results are likely to be less than

perfect, and their processing inflexible. Such machines often lack one of the most important steps in the image processing cycle - the interpreter. A trained operator or interpreter who has a reasonably well defined aim represents a very high level of computing ability, although every effort must be made to supply supportive data (e.g. statistical data, intermediate results) in order to maintain his effectiveness. Probably the best model of the general process of image enhancement is cyclic and iterative. The procedure for the CONDOR system is set out as follows, but it is readily generalised to any other system. To begin with the image must be observed, features of interest noted as well as their environment, and any other information collected, including intensity histograms from the whole or selected parts of the image. Based on this information a preliminary contrast enhancement can be defined roughly and applied to the image. The resulting image must then be subjected to careful scrutiny and tested against the objectives decided upon at the start (which may alter as the process continues). The next step is one of refinement in which the enhancement is trimmed or modified in an effort to improve the results, and the cycle repeats. Occasionally a different type of stretching function can be introduced, perhaps on the basis of previous experience, and the same iterative method then applies. Often it is necessary to process different parts of an image with different 'tailored' stretches in order to achieve the best results.

Throughout the above procedure, the important thinking or interpretation stages may be hindered by excessive delays in the execution of a contrast stretch or the acquisition of a histogram. The periods during which the machine is active represent potential bottlenecks in the process which prevent efficient operation. These delays must be reduced as far as possible by the design of fast and flexible hardware and software as described in the previous chapter. Access to the system should also be simple yet flexible and comprehensive, although not necessarily for the complete novice since this type of processing relies upon and assumes prior training or experience, and a sense of direction. If the system is made to respond quickly then the operator is able to follow up ideas and perhaps discount unproductive stretches without undue time penalties. A very important point which is often missed in image processing communities is that the availability of a faster machine does not (or should not) necessarily imply an increased speed of image processing. The primary effect of a fast machine must be to allow the burden of the work to shift from the hardware to the operator, with a consequent improvement in the quality of the results. Increased throughput can then arise as the operator becomes more efficient.

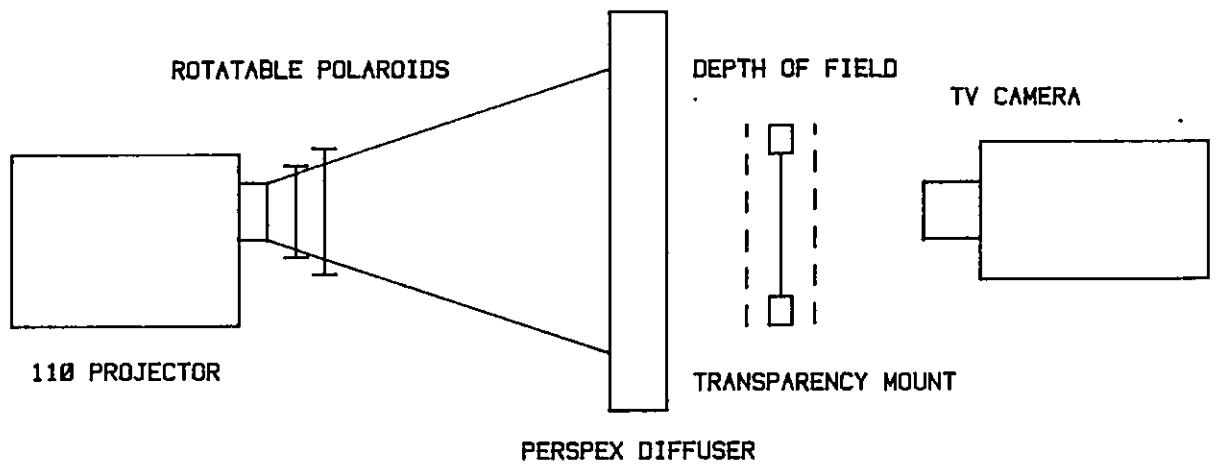


Figure 4.1 Transparency illumination system.

4.2 Experimental considerations.

The following sections will show how the CONDOR system can be used to enhance images with examples, but this section describes the experimental arrangement with which the images were input and output. As noted in chapter 3 the input device is a vidicon TV camera. This is fitted with a 25 mm close focusing lens which allows small transparencies or negatives to be viewed at close range. The transparencies can be of any size, but in the present cases they are mostly 35 and 70 mm. The images must be illuminated evenly, and this is achieved by a projector and diffuser arrangement as shown in figure 4.1. A 110 size projector is fitted with an aperture in the position normally occupied by the slide and adjusted so as to image it on a diffuser, thus giving a restricted area of uniform illumination. The diffuser consists of a thick sheet of opalescent or translucent 'milky' perspex, and it prevents any imaging of the projector lamp by the camera lens. The transparencies are held in front of this diffuser such that when correctly adjusted, the TV camera cannot focus the surface of the diffuser, thus reducing the effects of any scratches or other nonuniformities. The negatives are then seen against a uniformly illuminated background, the intensity of which can be regulated by a combination of the camera lens aperture and a pair of rotatable polaroid filters placed between the projector and the diffuser. The lens aperture is not enough on its own because of the great range of densities in the images and because it affects the depth of focus.

The video signal which is seen by the CONDOR system consists mainly of image information, but it is modified by nonuniformities in the illumination and in the response of the vidicon across the target area. The illumination system has been made almost uniform, but the camera target shows some shading at the edges. This shading is most evident when the vidicon is left in near darkness and a very harsh contrast stretch is applied to the resulting 'blank' field, but in normal use the video signal is dominated by the image information. Variations in the image far outweigh those created by the camera tube which become statistically insignificant, and selective sampling from the edges and centre of the field when no image is present (i.e. only the diffuser) show no significant differences in illumination. As a precaution, the TV camera is allowed a generous warming up period, and the image is left stationary for a period before any results are taken to allow 'burnt in' highlights from previous images to decay away fully (i.e. areas of the target which have been saturated

with light and which require some time to recover their normal characteristics). As an aid to this latter precaution, and to prevent permanent damage to the vidicon, the intensity of illumination is carefully controlled by filters.

Output from the system is collected from a flat faced monochrome TV monitor with high linearity and resolution which is photographed with a 35 mm camera fitted with a 50 mm macro lens. This monitor is designed for photography, and carries some pincushion distortion which is designed to compensate for barrel distortion in the corresponding polaroid camera lens, and this shows up in the results since the camera used does not suffer this distortion. Results were recorded on FP4 fine grain film processed in the normal way. In photographing the screen a reasonable time for exposure must be allowed so that the shutter speed does not interfere with the TV scan rate, and typical times were 1 S to 1/8 S. In some cases where a harsh stretch is applied over some part of the intensity range two adjacent grey levels in the original image may become widely separated in the final image. Any noise on the input signal tends to become amplified by this discontinuous transfer function (due to quantisation) but the time exposure averages this noise somewhat and makes it less harsh. All images were recorded with the same contrast settings on the monitor.

4.3 Basic processing.

The following sections will demonstrate some examples of contrast enhancement on different images as an illustration of the processing methods outlined in this thesis. The first image which has been processed by the CONDOR system appears in figure 4.2 a. This image is a neutron transmission image of a potted plant, and the phenomenon which causes contrast in such images is selective neutron absorption. Unlike X rays, neutron absorption does not depend on the density of an absorber in a simple way (e.g. lead) but is more erratic, with some elements showing very high capture cross sections. One element which is particularly opaque to neutrons is hydrogen, and this is the major absorber in the plant. Neutron images are very noisy and usually of low contrast because of the difficulties involved in detecting the transmitted radiation- the detector is usually quite transparent! In practice images can be formed by intercepting collimated neutron beams with phosphor screens which are viewed with an integrating TV camera. The integration is necessary because of the weak signals involved, and comprises some form of picture store which can accumulate the sum of many images, thus cancelling out the noise. Unfortunately, some

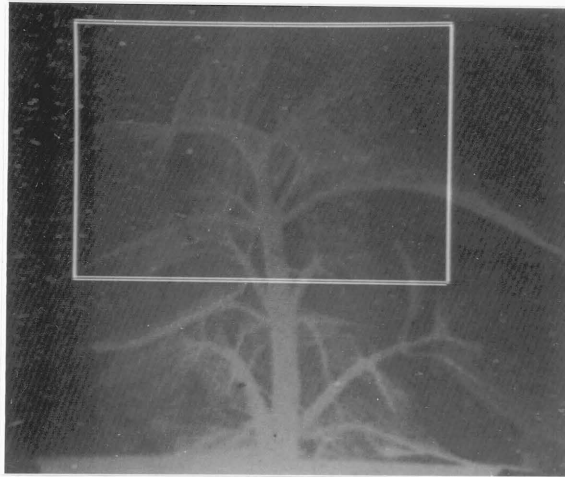


Figure 4.2 a : Original plant image



Figure 4.2 c : Linear contrast stretch with suppressed background



Figure 4.2 d : Equalisation with suppressed background

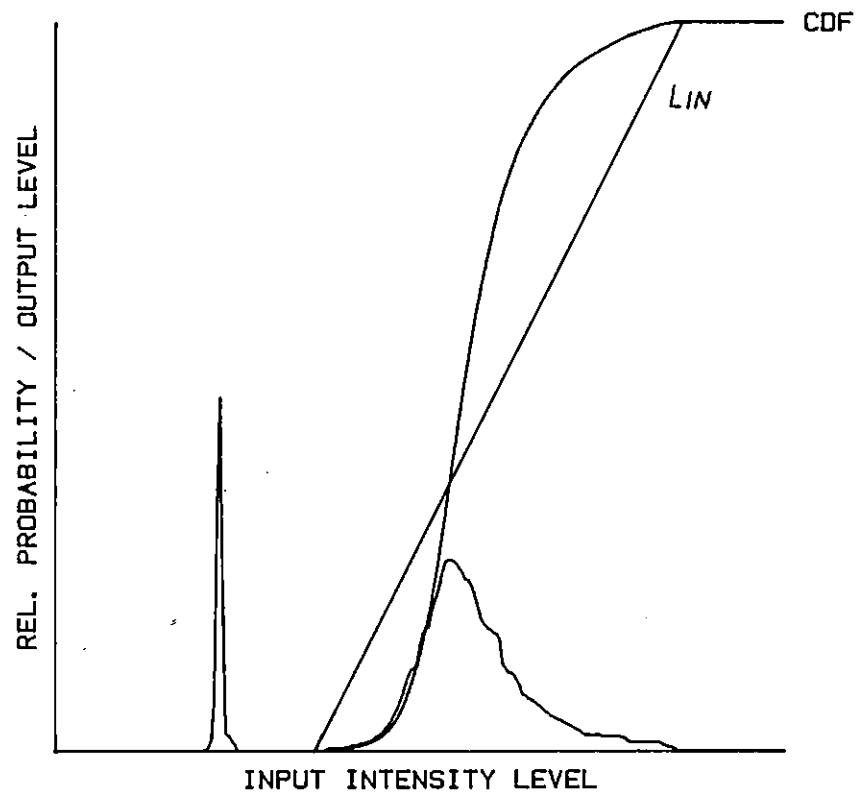


Figure 4.2 b : Entire image histogram and stretching functions

interesting objects such as the plant still give relatively poor quality images and are thus prime candidates for enhancement, the aim being a general increase of contrast and the exposing of fine details which are otherwise hazy.

As a first step, a visual appraisal of the image is made, aided by the acquisition of histograms from different areas to gain a feel for the image statistics. The image is obviously of low contrast, even after the linear gain afforded by the camera and video amplifier which has improved upon the original. The background has a range of intensities, and almost blends into the upper leaves. The lower part of the image shows a greater contrast (the thicker plant stems) and also an increased average intensity in the background. The window placed over the image refers to a data sampling area which will be used shortly. The histogram of figure 4.2 b was acquired by the system using the entire image with the exception of the flower pot at the base. The histogram has two major parts, a sharp peak at low intensity and a wider, almost Gaussian peak further up the intensity range. The sharp peak is due in the main to the background, and does not convey much useful image information. The main peak is skewed somewhat, with a short tail off to higher intensities, indicating that the image is mainly grey with small areas of whiter content. This second peak is not smooth, but shows a slight kink on its upper edge, the origin of which will become clear below. Primary processing for this image consists of a linear stretch whose slope and position is shown superimposed on the histogram of figure 4.2 b. Notice that under this mapping the peak of the histogram which corresponds to background has been suppressed, and the line's slope has been arranged so that the minimum and maximum values in the output image are black and white rather than different shades of grey. The result of this process appears in figure 4.2 c and shows a reasonable improvement in contrast, particularly at the centre and top of the image.

The simple linear stretch gives a useful result in this case, probably because of the low contrast of the original, but it was decided to apply a nonlinear stretch to the image for comparison purposes. The process chosen was equalisation as described in chapter 3. The transformation curve is calculated from the histogram by a numerical integration and is shown in figure 4.2 b along with the linear stretch above. The curve rises from zero at the start of the main peak of the histogram since the integrating software was instructed to omit the background peak for the same reasons as above. The curve shows a skewed S shape, tailing off more slowly at high intensity values. When considering the

effects of a transfer function such as this, it is important to remember that the contrast at any grey level will be proportional to the slope of the transfer function (and thus to the height of the histogram at the same level) but that the process of quantisation may cause either gaps in or compression of the intensities in the output image. The equalised image appears in figure 4.2 d and does show compression, particularly at the base of the plant towards the higher intensity levels since the transfer function has low slope in this region. The image appears with characteristic harshness and the background is much more black. Overall this processing is not quite as effective as the linear stretch above, but it can be improved upon by selective sampling.

The equalisation process described above used a histogram acquired from the entire image in order to calculate the transfer function. This is a common and even 'standard' process in much image enhancement work, but as is obvious from the above image, it is not always the optimum course of action. The variation in average grey level from top to bottom of this image was noted in the visual appraisal stage, and this observation can now be used to good effect. The image is split into two areas of interest comprising the leaves at the top and the lower stems, and these are processed separately. To begin with histograms are acquired from two appropriately placed windows which are designed to eliminate as much of the background as possible. The original image appears again in figure 4.3 a together with the first window, and the histograms are shown in figure 4.3 b. The two equalised images appear in figures 4.3 c and 4.3 d. The windows are specified interactively to the CONDOR system and can be modified quickly to allow rapid optimisation of the results. Both histograms consist of fairly narrow peaks which differ in shape and mean levels. The histogram gained from the upper foliage is narrow (range of about 65 out of 256 grey levels) and is nearly symmetrical. The transfer function is a sigmoid curve similar to the error function and tends to expand the central levels most. The other histogram is not so simple in shape and has a higher mean intensity value. It is more asymmetrical, with a tail at high intensities which causes its transfer function to rise more slowly in the corresponding region. Some faint sub-peaks are evident on the upper side also, probably arising from the different individual stems. The resulting images are quite different in appearance, and enhance their respective window areas quite well. The upper foliage in figure 4.3 c shows veins distinctly which are hard to see in the original and which the simple equalisation of figure 4.2 d nearly lost. The lower portion of the image has been degraded, emphasising the difference in image statistics between the two



Figure 4.3 a : Original plant image

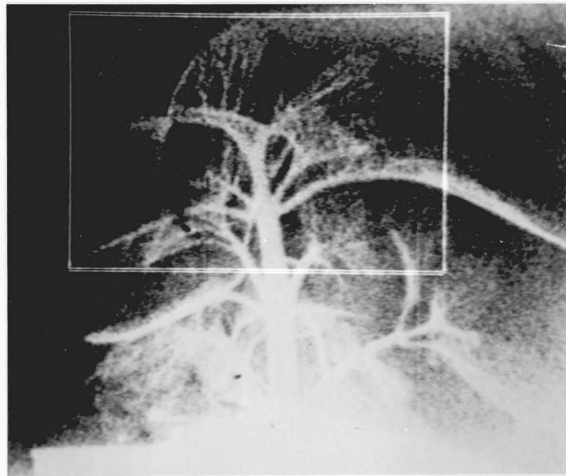


Figure 4.3 c : Equalisation over window indicated

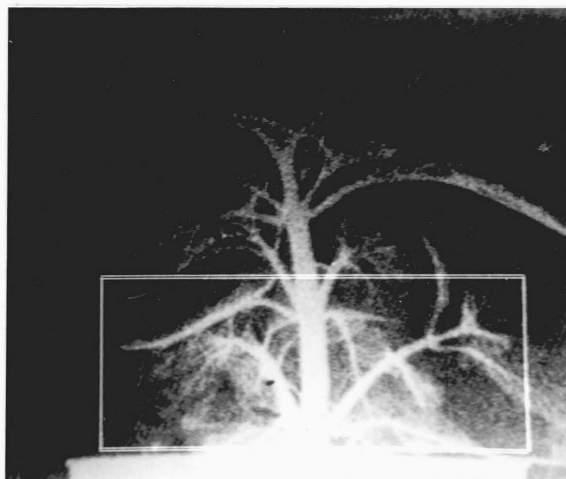


Figure 4.3 d : Equalisation over window indicated

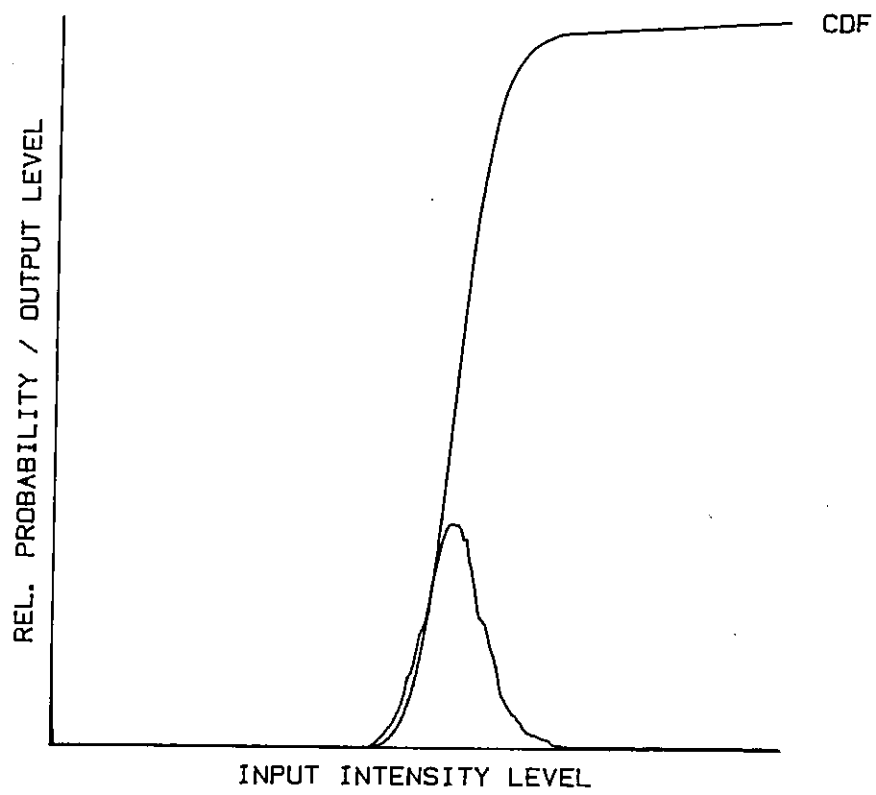
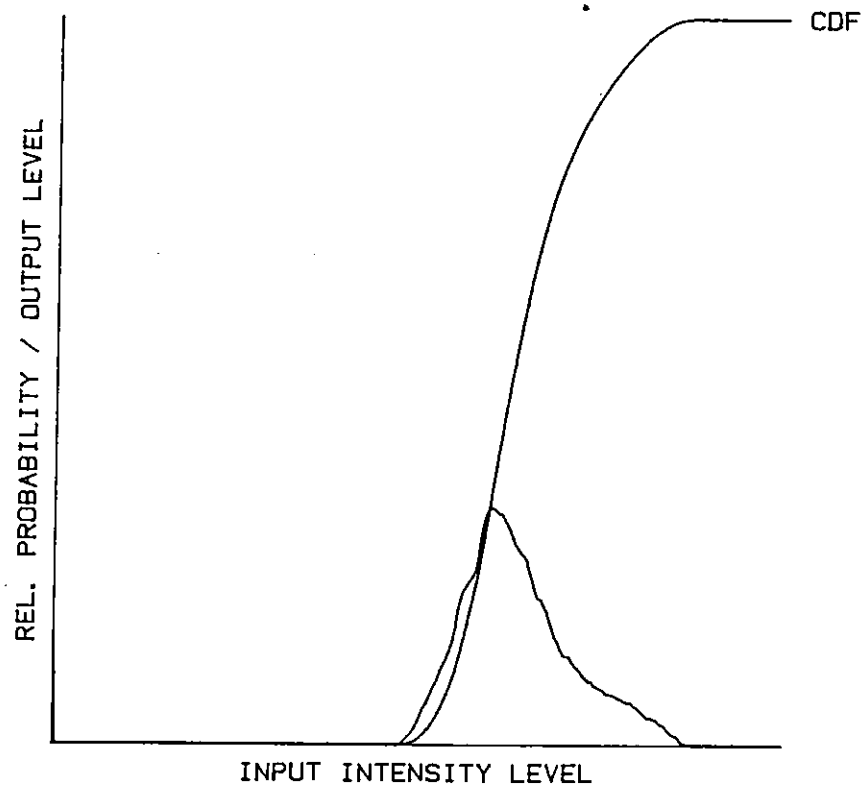


Figure 4.3 b : Histogram and transfer function from top window area



Histogram and transfer function from lower window area

areas, but the stems are much clearer in figure 4.3 d. This image has demonstrated the basic approach to contrast enhancement for low contrast noisy images, but any comparison between the results must be undertaken with care. There is no new information in any of the processed images which is not present in the original, and in some cases a loss has occurred. No single image represents the 'best' enhancement, but the whole set should be used together. The results also show the different levels to which the processing can be pursued, from a simple overall linear stretch for easy viewing to detailed enhancements of small areas when specific detail is required such as leaf structure.

As a further illustration of the use of selective sampling and the local variations in image statistics, a second image was subjected to similar processing. The original image is shown in figure 4.4 a and it depicts an X ray view of a weld between two steel plates. This type of image is common in industrial manufacturing of large steel structures such as pressure vessels or oil rigs and they are routinely taken to check the quality of deep welds. The X ray film is usually taped to the outside of the weld and a small but powerful source is arranged to cast a shadow of the weld onto it. The image is used to look for cracks, nonuniformities in the weld metal (gas or slag inclusions) and for flaws in the parent metal on both sides of the weld. Primary reconnaissance shows that the weld (running north-south in the centre) is darker at the top as in the plant image. This may be due to the weld metal or the placing of the X ray source, but it can be overcome by selective sampling, and three regions have been chosen for this study. These regions are the top and bottom of the weld area and the parent metal to the right of the weld. The three histograms appear in figure 4.4 b and the three processed images (all equalised according to the transfer functions shown) appear in figures 4.4 c to e. Figure 4.4 c shows the top of the weld, and it has enhanced a line of slag inclusions running down the left of the weld (dark spots) which are common in automatic welding machines. Also visible are several darker curved bands crossing the weld, and these are probably due to variations in the metal flow during the weld. The histogram corresponding to the upper section of the weld is fairly narrow (about 80 levels) but it shows some sharp peaks within it. The highest sub-peak is probably due to the brighter bands in the weld metal, and the lower peaks encompass the background area of parent metal and the inclusions. By contrast, the lower part of the weld shows a much wider histogram with two major peaks widely separated, and the image has more gross structure. The bright areas of

the weld are contributing to the upper peak and the parent metal gives most of the lower peak. The transfer function for this sub-image is much longer and shows two regions of high slope (at each sub-peak) separated by a region of near constant slope. When bimodal distributions of this type arise the relative heights of the two peaks will determine how the available grey level range is 'shared out' in the final image. If the first peak is large it will cause the integrated transfer function to rise by a large proportion of the full range, leaving a smaller range of intensities for the second peak. This can lead to compression, but it may be overcome by deliberately leaving out portions of the histogram during the integration.

The lower image (figure 4.4 d) has brought out details of possible fractures and inclusions which were lost in the first image (figure 4.4 c), but neither of these two images shows much detail in the parent metal. In order to remedy this situation, the last image was produced (figure 4.4 e), and it reveals a north-south nonuniformity as well as a round inclusion half way up the image. The corresponding histogram reflects the dark upper areas at the peak and tails off at higher intensities where the inclusion and lower area makes a contribution. This image also shows up a long crack on the left to good effect.

In the industrial environment many thousands of X ray images such as this are produced routinely, and their value depends upon the skill of highly trained interpreters. Whilst the practiced eye could probably pick out most of the features in this image without enhancement, the availability of a simple and quick means of improving images could be a great advantage. It would allow fine detail to be followed up and would perhaps reduce errors caused by operator fatigue which at present are avoided by using several interpreters independently for each image.

4.4 Manual stretch specification.

The processing described above has been quite simple in that a standard algorithm has been applied (linear stretch, equalisation), but the benefits to be gained from selective sampling have been demonstrated. The problem of training with such a system is minimised to a few basic concepts, and success relies on experience on the part of the user. If there is more time available and the user is willing to experiment, good results can be obtained from the use of manually specified nonlinear stretching functions. Some attempts have been

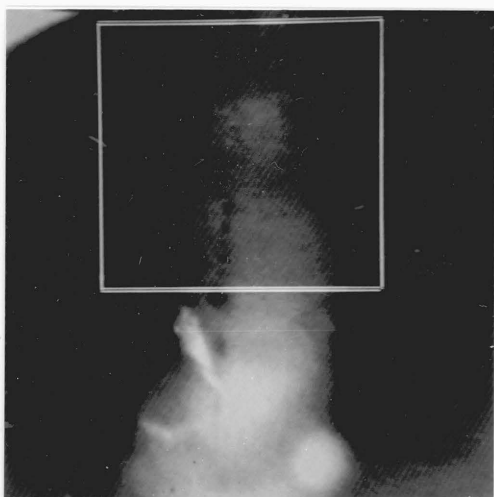


Figure 4.4 a : Original

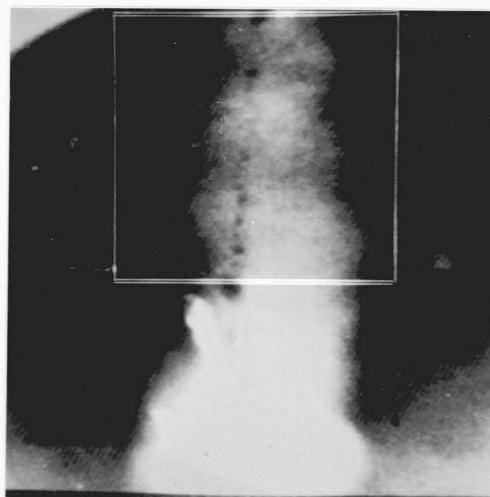


Figure 4.4 c : Equalised upper window

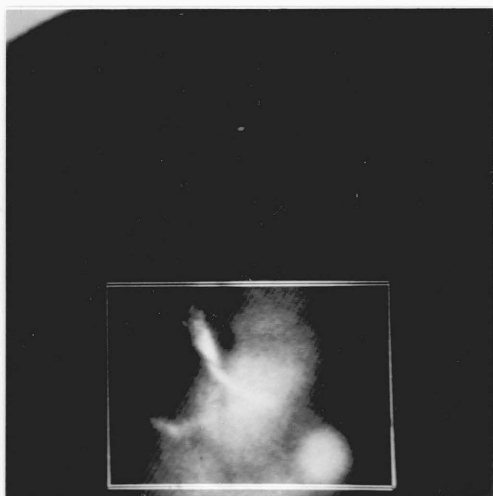


Figure 4.4 d : Equalised lower window

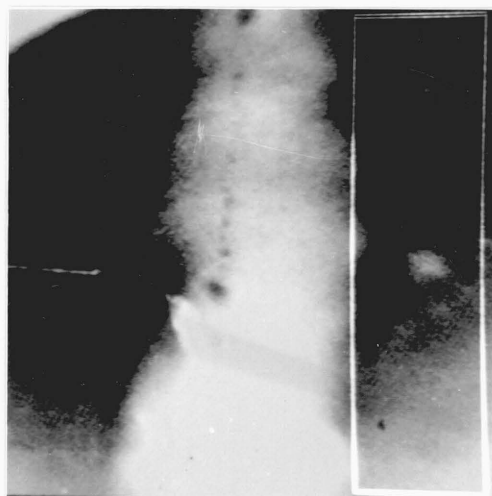
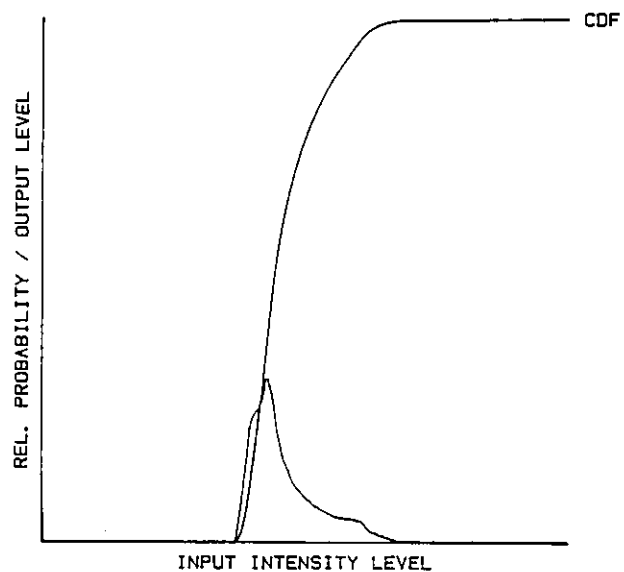
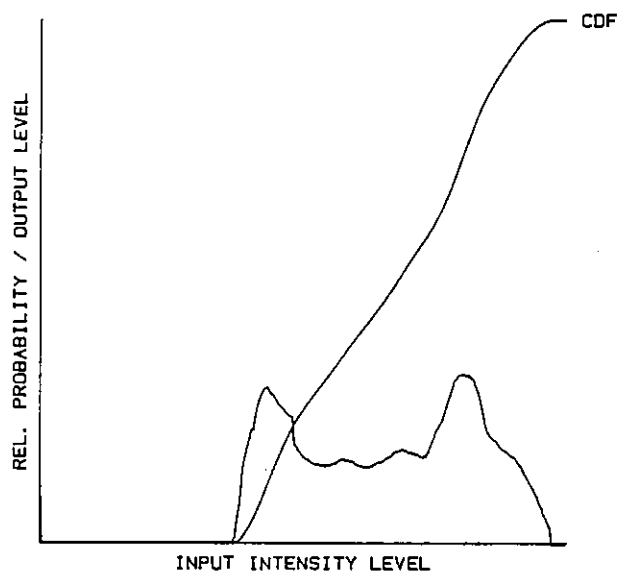


Figure 4.4 e : Equalised parent metal

Top of weld :



Bottom of weld :



Parent metal :

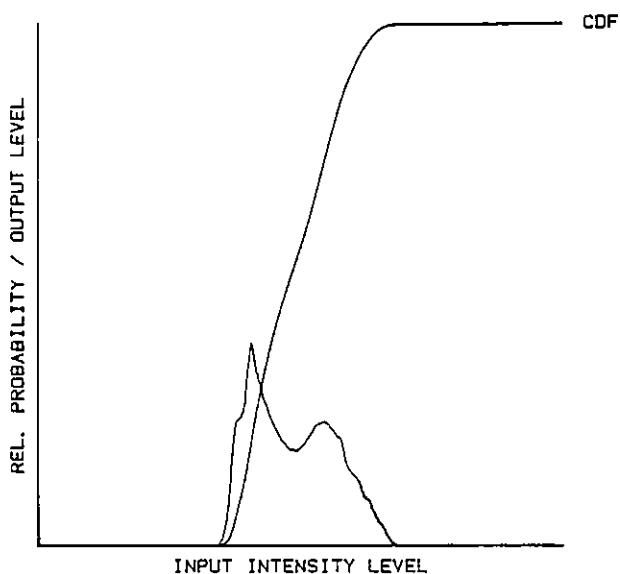


Figure 4.4 b : Histograms and transfer functions for weld images

made to achieve this but they have been somewhat limited in their capabilities, and such functions are not usually found on large image processing systems, perhaps because it requires significant work on the part of the operator! The basic system whereby transfer functions may be specified on the CONDOR system was described in chapter 3, and consists of a switch box equipped with 17 centre-biased toggle switches. Any curve may be specified interactively in conjunction with a computer display by altering a set of ordinate values, one for each switch, and performing a linear interpolation between them. The user is presented with a display consisting of two sets of axes, one showing the histogram of the image or the last transfer function, and the other showing the manually set transfer function which is initially linear. The lower and upper intensity levels between which the curve is specified can be selected by positioning indicator arrows along the x axis. Positions for the intermediate intensities corresponding to each switch are computed automatically or may be specified by the operator. As an example of the possible effects of such a process, it was applied to the neutron transmission image of a plant used above. The original image appears again in figure 4.5 a together with the transfer function in figure 4.5 b and the processed image in figure 4.5 c. The transfer function diagram is taken from the computer graphics display whose construction is described in chapters 3 and 5. The small bright spots on the curve indicate the sample positions corresponding to the switches, and the curve on the left represents the equalising curve used to produce the image in figure 4.2 c. The curve has been constructed in an attempt to overcome the difficulties caused by the average intensity shift over the image. A sharp and nearly linear curve is used to map the low intensity levels over the entire output range, thus enhancing faint details such as the leaves. Immediately following this a sharp negative slope serves to begin the mapping process over again with a second positive-sloped section. This begins by mapping mid greys back to black (e.g. in the central stem) and working up to white (at the base of the plant) for a second time. The negative slope section does not contain too much information so that negative areas in the image are hard to see. The resulting image has had both low and high intensity ranges amplified to cover the full intensity range, but an offset has been removed at the high end in order to accommodate the excess intensity. Such a 'wrap-around' technique must be used with care and with reference to the image features. In this case it highlights some edges (e.g. on the long leaf to the right) and it details the intensity variations on the main stem. This stem is smooth, i.e. without too much texture, and the two ranges of intensity correspond almost entirely to well defined and separate

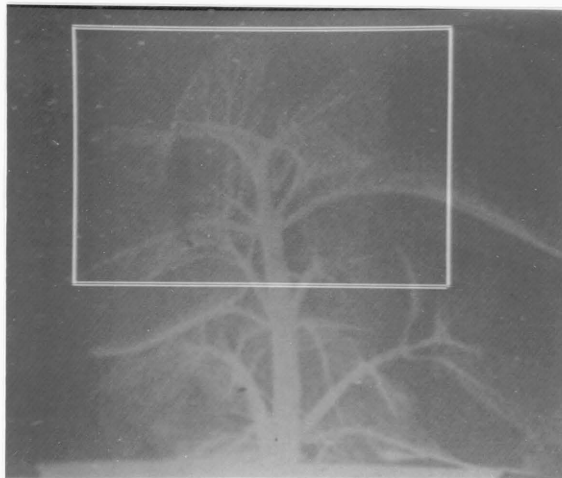


Figure 4.5 a : Original plant image

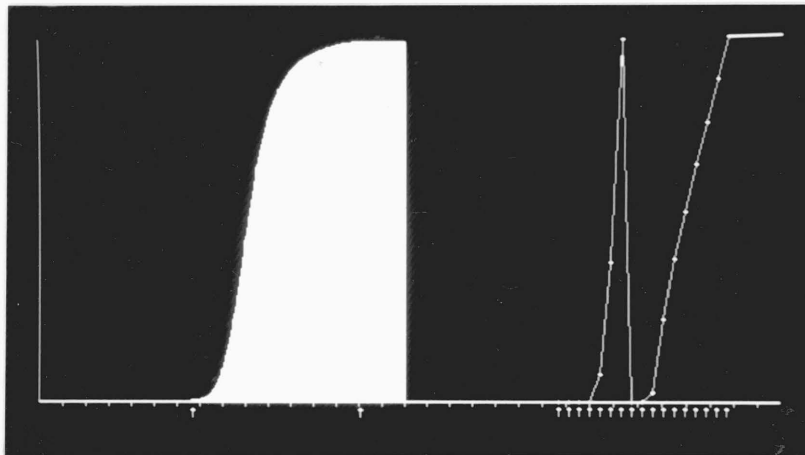


Figure 4.5 b : Manually specified transfer function



Figure 4.5 c : Stretched image using curve above

areas of the image. If this were not the case the result may become excessively complicated and noisy, hindering rather than aiding interpretation.

As a second example of this technique a new image was tried. This image is also a neutron transmission image obtained as before, but this time the object is a domestic kettle (figure 4.6 a). The steel side walls of the kettle appear nearly transparent but the hydrogen-rich plastic of the handle and the water inside appear opaque. The thicker metal of the element and socket fixings can be seen at the left, and the lid in the centre. The histogram of this image is shown in figure 4.6 b and was acquired over the indicated window. The shape of the histogram is the most complex so far, with a bimodal main section followed by a substantial tail at higher intensities. The lowest sub-peak corresponds to the background outside the kettle, the second peak to the area within the kettle, and the tail to the densest parts, the water and the handle. As a comparison with previous processing, figure 4.6 c shows the effect of equalising the image using the transfer function shown. Since the curve has a large section with low slope, detail within the water and the handle are relatively unaffected, but the image looks better and detail is beginning to emerge within the lid and the body of the kettle. Above the spout a light trail leaning towards the handle indicates escaping water vapour. The second and more irregular transfer function shown in figure 4.6 c was used to produce the image of figure 4.6 d. The overall shape of the curve is broadly similar to the equalising curve, although a greater slope has been chosen at the start of the tail to enhance the handle more effectively. The small peak affects the body of the kettle and reveals a broad band of nonuniformity as well as highlighting the edges along the top. The slope at the high intensity range has been arranged so that some detail becomes visible within the water. The element can just be seen running diagonally down into the water at the left, and two short near-vertical lines also appear above the body of the element.

The above example shows how a transfer function can be tailored to suit the particular image to good effect. The basic requirement is a feel for the relationships between areas of interest and the form of the histogram and transfer function. As an aid to this somewhat elusive skill the system responds to changes apparently instantaneously, and this allows experimenting and 'tweaking' to effect the best results. Unfortunately it is not easy to give general guidelines as to the effects of different shapes of transfer function on a general image because of the great variations possible, and as can be seen

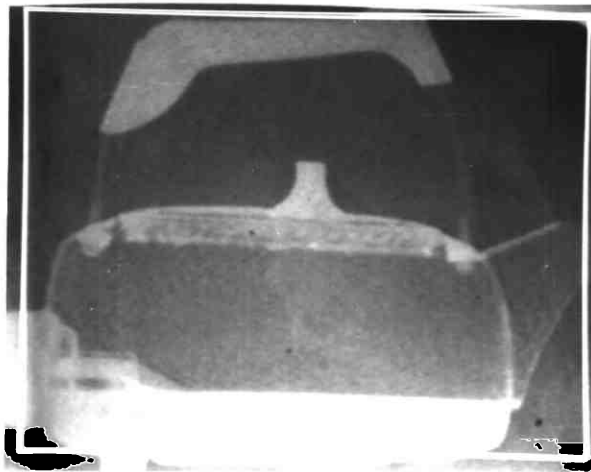


Figure 4.6 a : Raw neutron image of a kettle

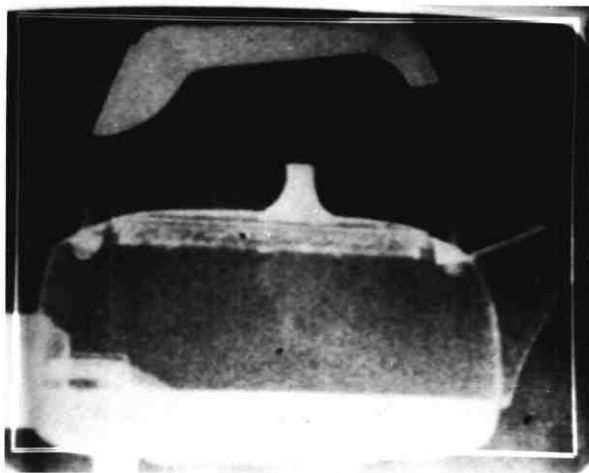


Figure 4.6 c : Equalised image of kettle

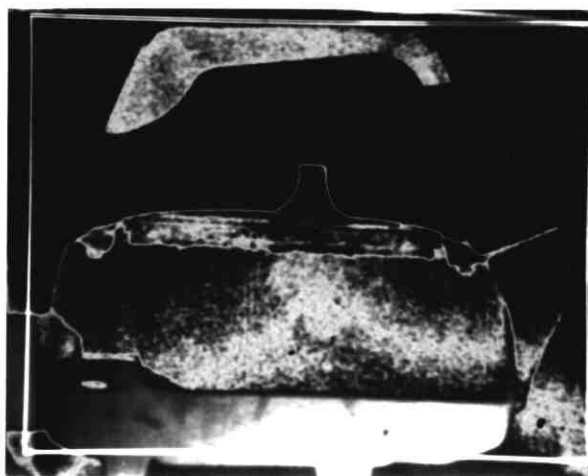


Figure 4.6 d : Manual stretch function applied to kettle image

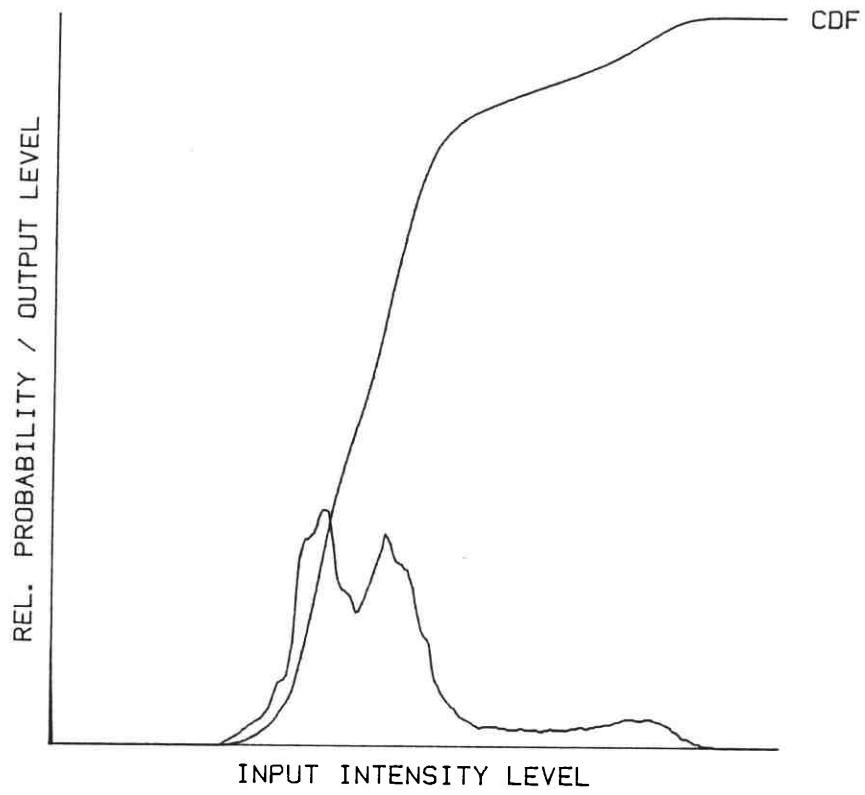
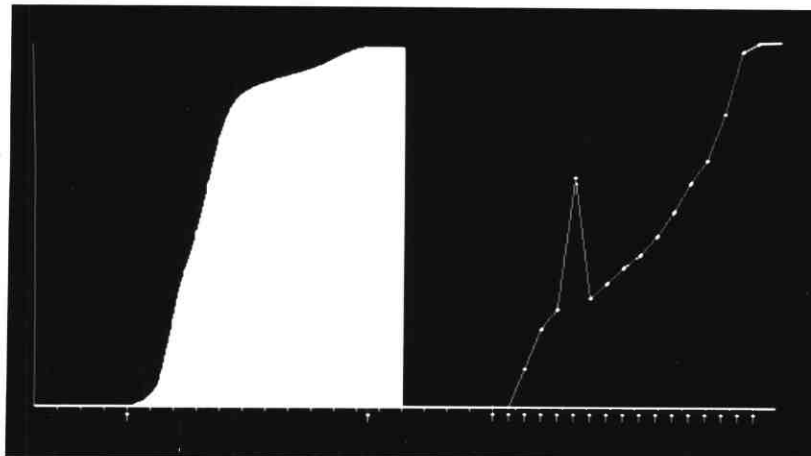


Figure 4.6 b : Histogram and transfer function for kettle image



from the above examples, the technique is not easily suited to automatic execution by an unsupervised machine (the software would be most complicated!). This must not be seen as a disadvantage however, because the computer/software combination that operates the hardware can be very potent given the right facilities.

Both the transformation functions used in this section have not preserved the correct order of the grey levels from the input image to the output, i.e. they have not been monotonically increasing single valued functions such as the equalisation curves. This fact means that the use of such images in an image classifying system is very difficult without great care. The main area for which such images are useful is in visual interpretation, but if monotonic functions are specified this problem will not arise. The number of different curves available is almost limitless, and the use of this facility of the CONDOR system requires much more exploration. One further example of the use of manually-specified transfer functions appears in the next section.

The same hardware and similar software can be used to specify a target histogram shape as described in chapter 3. The required shape is input using the switches, but instead of a linear interpolation between the ordinates, an integration and normalisation step is performed to produce an equalising function which is then inverted. The raw data is first equalised (according to its histogram) and then this inverse transform is applied. The process is not quite as fast as manual transfer function specification, but it still operates within a second and can be used interactively. This type of process is usually applied as an aid to multispectral classification of similar images where the histograms are made to conform to each other as closely as possible. Often the effects are not immediately obvious, and no examples are included in this thesis although it is within the capabilities of the system. In a future version of the system a four channel unit could be used for colour composite work and the composite transfer function described above could be replaced with two look-up tables in series for faster operation, the first performing the equalisation and the second the inverted equalisation.

4.5 Contrast rearrangement

The images used in previous sections to demonstrate the capabilities of the system were generally of poor contrast to begin with. In many situations this

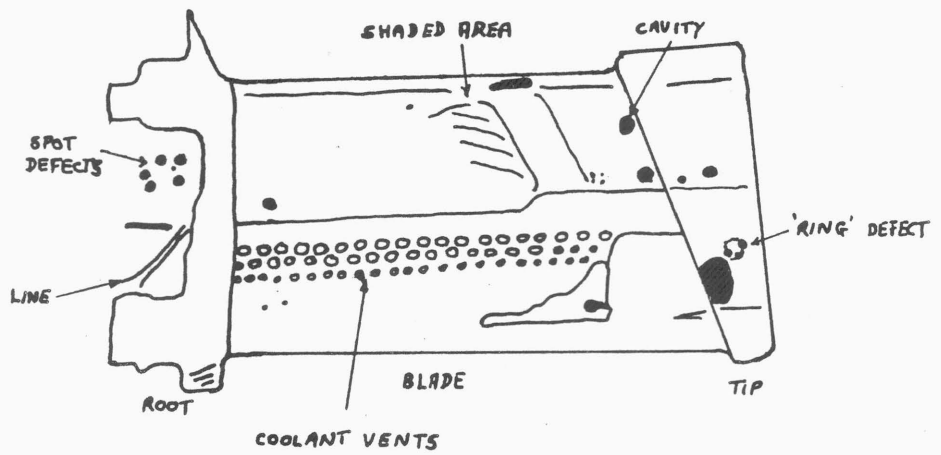
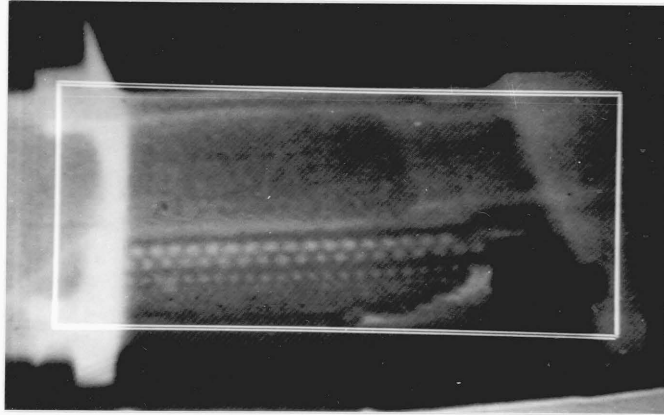


Figure 4.7 a : Original turbine blade X ray image

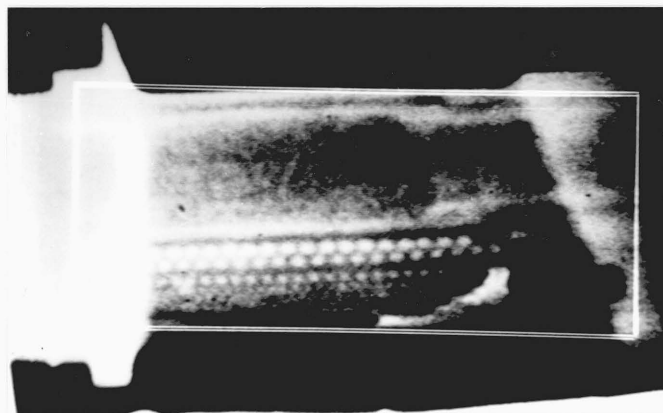


Figure 4.7 c : Equalised turbine blade image

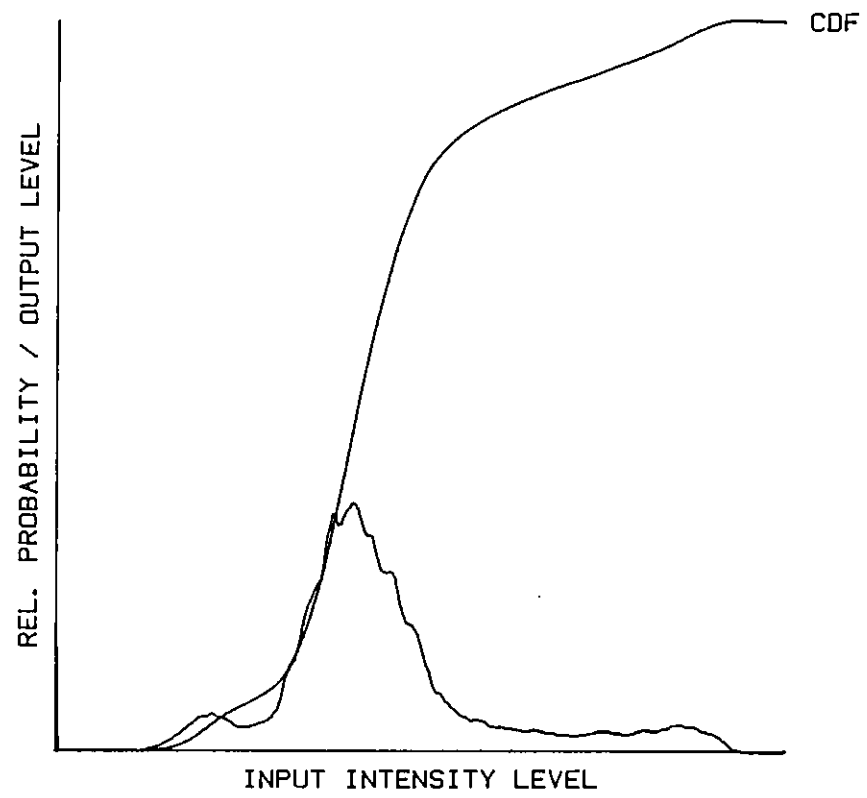


Figure 4.7 b : Histogram of turbine blade within window

is a common occurrence, particularly with 'soft' objects in X ray images or when a satellite views the ground through a hazy, scattering atmosphere. The lack of contrast can be deduced by reference to the shape of the appropriate histogram, which will tend to show a distinctly bunched form with a relatively narrow range of grey levels. Of the above mentioned images, the neutron picture of a kettle had the widest histogram because of the distinct contrast between the separate areas of the image (i.e. the background, the body and the handles / water). In this section one further image is presented in order to show that contrast modification can still be of use even if the image has superficially a 'good visual contrast'. The image appears in figure 4.7a together with a line drawing which indicates the major features. The image depicts a turbine blade of the type found in jet engines as seen in X ray transmission. The blade is made of specially processed steel, and presents a fairly dense object to the X ray source. The blade would normally be emplaced in a wheel or hub and secured by the root section which has locating slots. Above this comes the airfoil which is the working part, seen here in profile (i.e. from above in transmission) and this is topped by the blade tip (seen here as a somewhat triangular area) which serves to fit into the turbine housing. The rows of dots spaced along the blade section represent coolant ducts and outlets which carry high pressure gas. The blade is kept cool by this flow to minimise thermal stresses and the gas can serve to modify the performance of the blade. Examination of turbine blades is a very important operation for safety reasons and makes significant demands on the inspectors who have to spot even the smallest defects. When many blades have to be checked with high accuracy, some form of machine aid can be useful.

On the original image (fig. 4.7a) there are several details worthy of note before any processing is attempted. The upper part of the blade has some faint shading present as indicated in the drawing and a line running parallel to the tip. The tip itself shows some dark spots, and there is a defect near the lower edge of the blade. A final area of possible interest is the root where a diagonal line can be seen along with some shading. After the preliminary visual appraisal which serves to indicate the targets for investigation, a histogram was acquired using data from within the box shown. The box is designed to fit over most of the blade and tip since an excessive area of background would tend to obscure any fine detail. This detail is 'lost' when the data are normalised for display, and whilst unnormalised data could be displayed in sections, this may not be convenient when an overall quick view is required. Furthermore, since the background is not useful, the problem can be eliminated by a suitable

choice of window. The result is depicted in figure 4.7b, and it shows the wide range of contrast in the image. Most of the image appears as a mid grey as indicated by the central peak, but there is a darker peak, probably due to the defect near the tip of the blade, and a long tail up to the white end which arises from the root area, the coolant vents and some of the tip. The curve which is drawn over the histogram represents the equalising transfer function which has been applied to give figure 4.7c. The shape of the histogram, and thus of the curve explains why much of the information in the root area has been compressed. It must be remembered that the transfer function as applied to the image data is quantised, and that regions of low slope in the transfer function can cause loss of information as well as compression. Since the blade occupied most of the window, it has received the most stretching, and the defects noted above have become more visible as a result. The gradual shading along the blade is enhanced as are some of the dark spots, two becoming visible near the diagonal line on the blade which were quite hard to see before. These spots may represent defects such as inclusions in the steel, or they may have arisen during the processing of the X ray film (less likely), but their visibility allows a more detailed, possibly ultrasonic investigation to be attempted. A second area where a feature has 'emerged' from the image is just above the large dark spot in the tip of the blade. This takes the form of a ring or small circle right at the edge of the window with some smaller spots around its edge.

The root area of the blade is most important for the correct operation of the turbine since it must bear the greatest loads. In order to investigate the root more closely a second window was placed over it as the original image in figure 4.8a shows, and the histogram acquired. This histogram (figure 4.8b) also shows a wide range of intensity levels to be present, but as may be expected from the look of the image, the histogram is concentrated at higher intensity levels where there are several distinct sub-peaks, probably due to the differing thicknesses of the steel in this region. Again the equalising transfer function is superposed on the histogram and the result, somewhat magnified and cropped, appears in figure 4.8c. Once again the diagonal line is visible in the lower half of the image, along with a darker band running along the line of the blade which may represent a fixing groove. The light area represents thick metal, and as such does not show much detail except at the top where it appears to be a little thinner. A careful look at both the original and the processed image will reveal the presence of a cluster of spots in the

upper half of the root cavity, these appearing slightly darker than the surround. They are situated in an area of relatively low intensity, somewhere in the low tail of the histogram above the initial small peak (which is itself caused by the background areas in the window. The transformation curve in figure 4.8b has low slope and thus has not picked out these levels since the brighter parts of the image have dominated the transfer function.

In order to reveal these areas several options are possible, perhaps the simplest of which would be to confine the part of the histogram which contributes to the curve to the low and mid grey region. This is equivalent to setting the limits for the integration and is easily performed by the software. A possible disadvantage of this method is that much of the rest of the image would be compressed beyond usefulness by such a harsh transfer function. Accordingly it was decided to produce an image which enhanced most of the features described above at the same time by specifying the curve manually. In this way an overall image which portrays the information quickly and efficiently, even to people other than the primary interpreter can be produced. Such an image is most important as a method of communication as long as it is accompanied by some guidelines as to its capabilities. The transfer function was constructed using the switch box as described in earlier sections. The lower limit of the curve was chosen so as to miss out the small area of background, and the upper limit was set to white so that the entire histogram could be accommodated. Each sample position was set automatically within the range to give uniform coverage.

The design of the transfer function has to take into account the intensity levels of the features of interest as well as their range of intensities. As an example, in order to reveal the defects noted above in the root area, the curve should be steep at low intensities. Also any variations in the tip region or the whiter parts of the root will require a similar steep slope at high intensities. Such strategies as these are usually mixed with intuition and experiment to synthesise the final result, and it is shown in figure 4.9a with the transfer function in figure 4.9b. This may not be the optimum image, and a puritan may argue that no single image can ever be adequate completely to describe the features present, but it does show them all to a reasonable degree, and when interpreting time is short and expensive, it is worth having. Unfortunately, the constraints on the curve and the limited dynamic range of both the photography and the eye have forced the use of a discontinuous curve

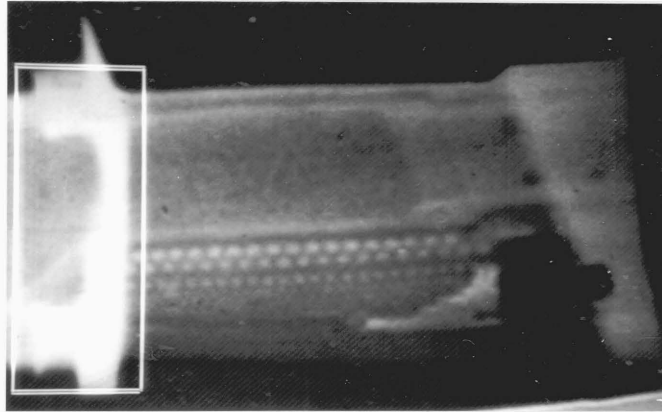


Figure 4.8 a : Original turbine blade image (root area window)

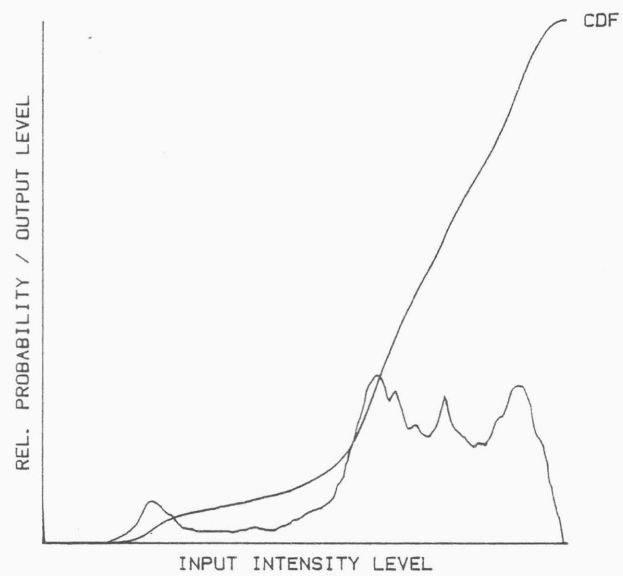


Figure 4.8 b : Histogram of above within window

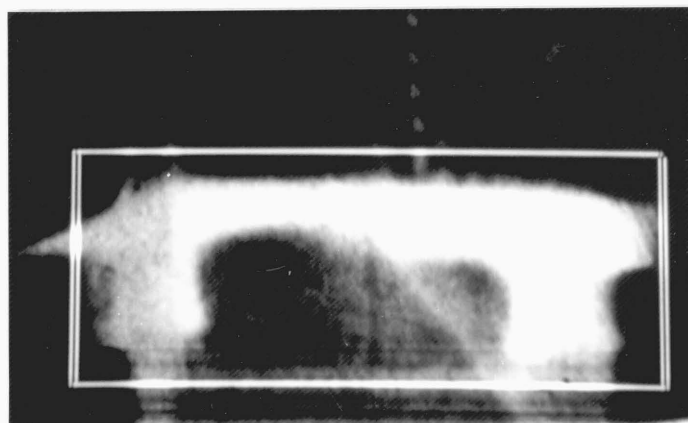


Figure 4.8 c : Equalisation within window

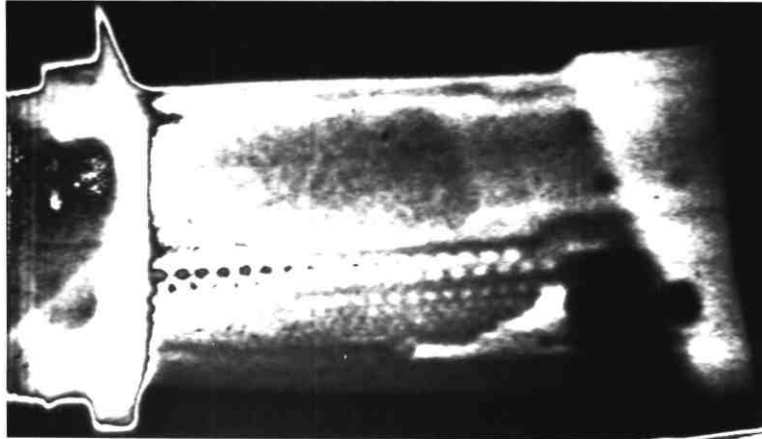


Figure 4.9 a : "Contrast rearrangement" of turbine blade

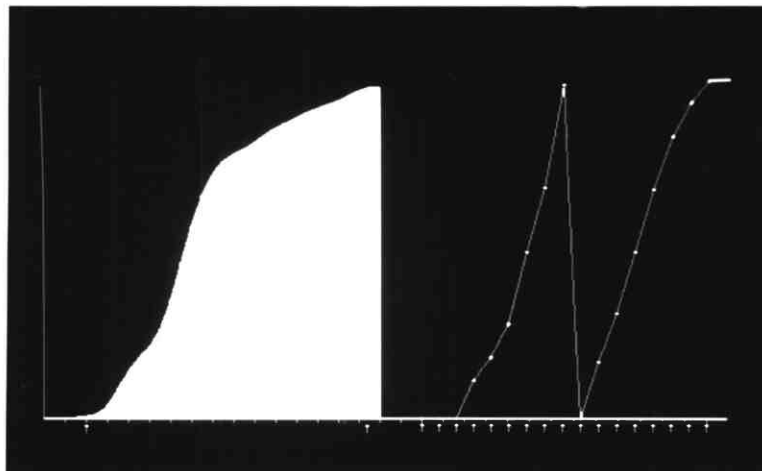


Figure 4.9 b : Nonlinear transfer function used above

again, but it is not without its advantages. The narrow band of negative slope gives a type of contour in the image, highlighting the edges of the root fixings where the metal is thinner as seen from this angle, and also the variation in blade thickness from root to tip by the varying shading of the coolant vents. The broad shading on the blade has been preserved by the lower portion of the curve, and the defects in the root area are very well shown by an appropriate section of steep slope. At the same time the slight variations in the tip have been preserved along with the 'ring' defect which is just visible.

4.6 Conclusions

This chapter has described the use of the CONDOR system as applied to several images both for contrast enhancement and rearrangement. The use of high speed functions such as histogram acquisition, selective sampling and quick transformation allows the user to concentrate on the enhancement problem in hand without undue worry or delay caused by the system. As with any complex machine having 'shiny' new features, some strategy for its effective use must be formed, and the limitations or capabilities of each process explored. The strategy adopted in this case has been demonstrated by example, and has several (not necessarily distinct) stages, i.e visual appraisal of the image, analysis of histogram shapes from different areas, application of controlled standard stretches, and finally manual stretch design. It is always vital to be sure of the aim of the processing, and the effective use of the system relies quite heavily on thought by the operator. The strategy described above has been developed by the designer of the system to be as general as possible, but specific types of image or problem may cause the evolution or modification of these ideas in future.

One major feature of this system which is only possible with a fast hardware and software combination is the facility to specify transformations by hand in real time and see the results instantly. If an effective feel for the shape of the histograms and the overall form of the image can be gained, this is a very useful technique for enhancement which allows fast modifications to be examined and easy experimentation by the operator. The use of a switch box is a convenient method of input, but it is hoped to modify the software to allow other devices such as tracker balls, joysticks and graphics tablets to be used.

The overall system is compact and relatively cheap, and it could be coupled to any standard TV camera or flying spot scanner which is capable of external synchronisation. It would also be possible to use diode array cameras which are becoming available, although some resolution might be lost in the image. The resolution is controlled by the clock rate and if extra high definition is required, as many as 1200 pixels could be accommodated along a TV line. The system computer could be reduced to a more dedicated form for extra compactness and reduced cost. The system is versatile, yet simple to use, and untrained operators can produce reasonable results quickly, learning as they go. As a routine aid in many fields where images must be interpreted and visual techniques are used at present, the CONDOR system might be a useful aid.

5.1 Information in feature space

The concept of a multidimensional space which could be used to represent the spectral characteristics of ground cover types has been introduced in earlier chapters with reference to the operation of box classification. This operation is one of the most common processes applied to multispectral data such as that from Landsat, and any method which can improve its accuracy or simplify its application is always welcomed. It has been noted that the distribution of feature vector tips (defined with respect to the four Landsat bands) is not uniform throughout feature space, but shows local clustering. The extent of a cluster reflects the statistical distribution of various samples of a cover type about the ideal spectral signature. The clusters may be thought of as defining a four-component 'colour' in the same way as three colours are used to form a TV picture, and a classification is thus a method for discriminating fine colour differences. Clusters are usually elliptical in shape, as can be seen in a two-dimensional projection of feature space, and in general their major axes are not aligned with any of the defining axes (the four bands). The classification method employed by the MEIC can only pick out points within a rectangular 4-D box whose sides are parallel with the axes, and this can lead to errors as shown in figure 5.1a. Classifiers which operate via software can be programmed to analyse the pattern of clusters and to classify any point in feature space by deciding to which cluster it is most likely to belong. This requires the measurement of cluster widths and positions which can be a time-consuming process. Other methods which are sometimes employed to open up or separate the clusters include dividing two similar images in the hope that the ratio image will enhance any small differences between the original bands [51].

Figure 5.1a shows that the feature space distribution is roughly oriented along a line 45° to the axes, and thus indicates that the two bands shown in this 2-D projection have some correlation between them. This is a common occurrence, especially for bands which are closely spaced in their wavelengths, and it indicates that the Landsat system is producing some redundancy of information in four bands, and that perhaps fewer bands might suffice if carefully chosen. One option which is relatively simple to implement in an electronic form is a rotation of the feature space distribution to bring the

major axis of a cluster parallel to one of a new set of axes which are also used to define the classifier box as shown in figure 5.1b. This reduces the inaccuracies described above and has been used to advantage in practical work to improve classifications. Unfortunately the selection of the correct rotation parameters for a four dimensional distribution requires skill on the part of the operator who has to interpret six two-dimensional projections in order to find his bearings. It is unlikely that any automated process will completely remove this burden from the operator, but an indication of the best rotation required would certainly save time and effort. The operator will still be needed to check results and make adjustments, since the 'best' rotation may change according to his exact objectives. The future sensor systems such as Landsat 4 will operate with seven spectral bands instead of four, and this increases the complexity of the problem dramatically. The redundancy between bands on such a system may allow some degree of data compression, perhaps to a more familiar four bands, thus simplifying classification, but such compression must be carried out carefully in case the information of interest is lost. The main method which is used to accomplish some of the aims set out above is the Principal Components Analysis or Transform (PCA), which is described in the next section.

5.2 The principal components analysis

This method has been discovered and used by several workers, sometimes in unrelated fields, and is variously referred to as the Hotelling Transform, the Eigenvector Transform or the Karhunen-Loeve Transform (KLT). The aim of the method is to redistribute the data in the input channels between the same number of new channels so that they are uncorrelated, and the first few bands contain most of the information, thus allowing compression. The method is applicable wherever a process is described by many parameters, some of which may be correlated, and it is required to form new parameters from the old which define the process more simply. In the example shown in figure 5.1a,b the major axis of a roughly elliptical cluster was rotated so that it was parallel to one of the new axes, and this is equivalent to defining the new axis along the direction of maximum variance of the cluster. The range of values occupied by this cluster along the new axis is larger than before, but the range on the perpendicular axis (the minor axis of the ellipse) is smaller, thus showing a redistribution of image variance between the two bands. The PCA produces new bands whose axes point along the directions of the maximum variance of the data

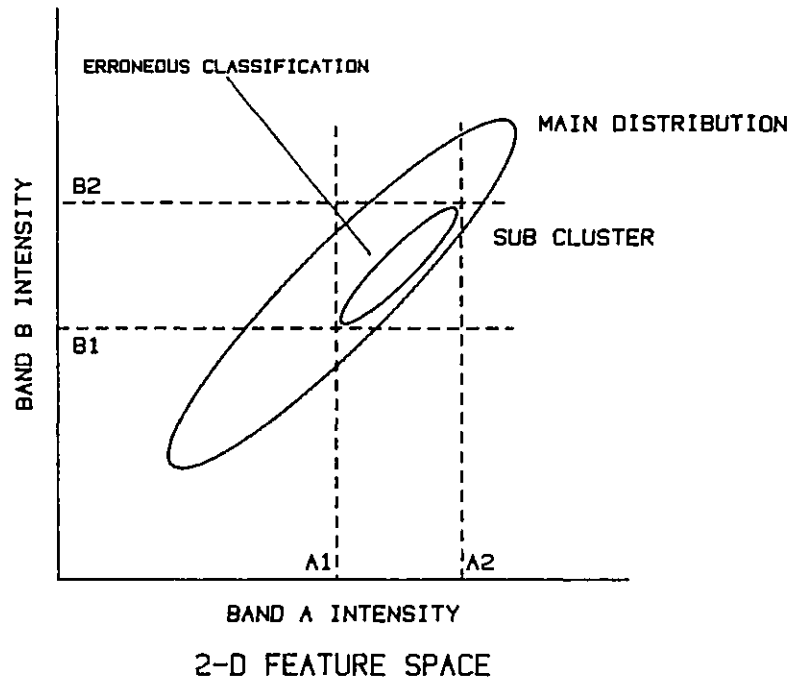


Figure 5.1 a : A typical 2-D feature space plot

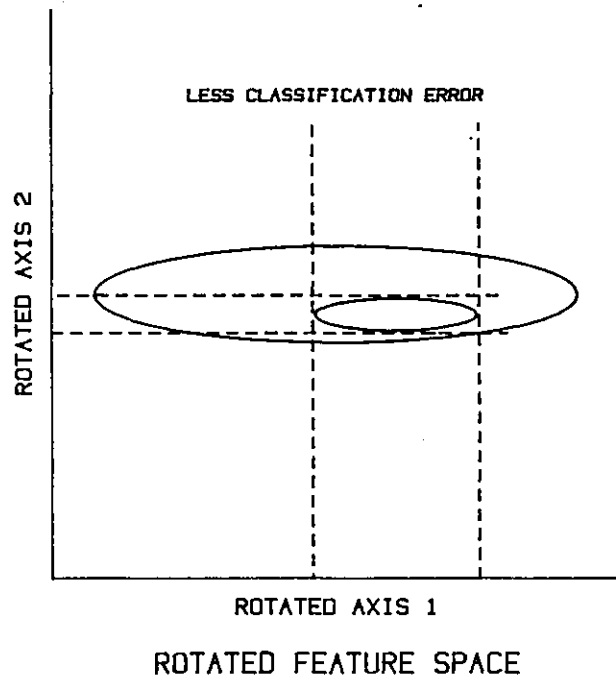


Figure 5.1 b : Rotated feature space plot

subject to the constraint that the new axes must be orthogonal. The calculation of the transformation depends on an analysis of the variance of the feature space distribution.

Since the aim of the PCA is to produce decorrelated bands, the best way to characterise the data is via its covariance matrix C . The data, which has usually been digitised, is in the form of feature vectors x defined with respect to a set of orthogonal basis vectors (the original bands) r . This can be written as follows:

$$x = \sum_{i=1}^d x_i r_i \quad (5.1)$$

where the x_i are the components along each basis vector r_i and d is the dimensionality of the space (the number of bands). Digitisation restricts the values of x_i to be integers of given precision. Each pixel thus has associated with it a vector of the form of equation (5.1), with the x_i representing the intensities in each spectral band of the original data. The image can be thought of as a vector random variable X defined over the same basis vectors, and the covariance matrix C is then given by a vector equation as [52]:

$$C = E[(X - E(X)) * (X - E(X))] \quad (5.2)$$

where $E()$ represents mathematical expectation and $*$ represents a tensor product. Equation (5.2) is highly compact but bears some explanation for clarity. C is a square matrix with d rows and columns (equal to the number of bands) whose elements are as follows: each on-diagonal element is the variance of one of the original bands, and each off-diagonal element is the covariance of the appropriate pair of bands. The variance may be defined together with the covariance to give an equation for the elements of C :

$$c_{ij} = \frac{1}{n} \sum_{i,j} (x_i - \mu_i)(x_j - \mu_j) \quad (5.3)$$

where $\mu_{i,j}$ are the respective means of the two bands $x_{i,j}$. Because equation (5.3) is symmetric in i and j , it follows that $c_{ij} = c_{ji}$ and that the matrix C is diagonally symmetric. Since this is the case it can be shown [53] that it is always possible to find a set of real eigenvalues and eigenvectors by solving the appropriate secular equation

$$C - \lambda I = 0 \quad (5.4)$$

where the eigenvalues λ_i are roots of equation (5.4) and I represents the identity matrix. For the given form of C it is always possible to find a transformation matrix T which diagonalises C when used as in equation (5.5):

$$C' = T C T^{-1} \quad (5.5)$$

with the resulting matrix having the eigenvalues of C along its leading diagonal. The matrix T has as its rows the d normalised eigenvectors of C . Equation (5.5) defines the covariance matrix of the principal components, and it can be seen that since all off-diagonal elements of C' are zero, the PCs are uncorrelated, but this must not be confused with independence which can only arise if the distribution of the original vector random variable X was normal. The original points x_i in feature space now have corresponding coordinates relative to the new axes found by transformation by matrix T as

$$y_i = \sum_j T_{ij} x_j \quad (5.6)$$

and these must be requantised to integer values as before. In general, the intensities in the principal components may be written as the quantised values

$$p_i = Q \left\{ a_i \sum_j T_{ij} x_j + b_i \right\} \quad i=1 \text{ to } d \quad (5.7)$$

where p_i are the intensities in principal component band i , Q represents the operation of quantisation, and a_i , b_i are constants which change the contrast and brightness of the final results. The amount of variance in each raw PC image is proportional to the corresponding eigenvalue, and it is common to find that over 90% of all the image variance can be accounted for in the first two PC channels for Landsat data. If different scaling factors a_i are chosen, this may not hold true unless the ratios between the variances are preserved.

The usual method by which the PCA is accomplished in large digital image processors is software. The image bands must be held in memory and accessed by a computer which calculates the covariance matrix elements. This must then be diagonalised and the eigenvectors extracted, and finally the images must be

reprojected along the new axes. This process is slowed down significantly by the need for multiple accesses to the stored image data, and is not interactive. The observation that most of the image variance occurs in the first few PC bands prompts the idea of data compression, since the original bands can be recovered from the PC bands by an inverse transformation of only these channels with little loss of fidelity. The exact interpretation of the contents of the PC bands in terms of recognisable features is very complex and will be touched upon in the next chapter, but instances have been reported in which valuable information was present in one of the supposedly dispensable channels. Great care must be excersised in defining information content, and equation with image variance is not always satisfactory.

5.3 Noise sensitivity

The operation of the principal components transformation obviously depends on the acquisition of valid covariances from the original channels, but the raw data is almost invariably polluted with noise. This noise tends to obscure the boundaries of clusters in the untransformed feature space, causing inaccurate classifications, a problem which can be especially acute when only small areas of ground information are available. It has been shown that classifications carried out on principal component bands are more accurate when the first few PCs are used than for the same number of original bands [54], but that the two cases produce similar results when all the available bands are utilised. This suggests that the first few PC channels have an improved signal to noise ratio over the original bands. The transformation redistributes the image variance, giving most to those PCs with the largest eigenvalues, and often the noise seems to appear in one of the last PC bands.

The effect of a random noise signal on the calculation can be assessed simply as follows. The noise signal is assumed to be completely random, and to give a vector in feature space (the noise is multispectral) which has no correlation between its components. The noise is further assumed to be additive so that the data presented for the calculation are now given by vectors

$$\mathbf{y} = \sum_{i=1}^d (x_i + n_i) \mathbf{r}_i \quad (5.8)$$

where the n_i are the noise components and all other parameters are as in equation (5.1). The noise has a covariance matrix which C_n which is simply the identity matrix scaled by the noise variance σ_n^2 . The covariance matrix calculated for the noisy data set will be the sum of the covariance matrices of the noise and the original clean data. The transformation will calculate the eigenvalues of this matrix from the characteristic equation, i.e.

$$(C_x + C_n) y = \lambda_y y \quad (5.9)$$

where y represent the eigenvectors of the composite covariance matrix and λ_y are the eigenvalues. Because of the form of C_n this equation can be rearranged to give

$$C_x y = (\lambda_y - \sigma_n^2) y \quad (5.10)$$

which indicates that the resulting eigenvectors y are identical to those derived from the clean data and that the eigenvalues have been increased by the noise variance. The derivation of the transformation matrix T in the presence of additive random noise results in the same eigenvectors but modified eigenvalues, so that valid results can still be obtained. The noise variance adds to the PC variances, and since the last PC has nearly zero variance in the noiseless case, the calculated value can be used as a measure of the noise present in the data. When the transformation is carried out, the noise distribution is unaffected since it has a spherical distribution in feature space, but the image data are rotated as before. Since the variances are redistributed as noted above, the signal to noise ratios in each PC band alter, being given by the ratio of the appropriate eigenvalue to the noise variance. The first eigenvalue is always larger than any of the original band variances and the signal to noise is thus improved in the first PC. This improvement may also extend to other PCs according to their variances.

The results derived above indicate that the PC transformation is relatively insensitive to noise. The use of selective sampling which has been proposed as a means of distinguishing class types of interest more effectively [55] must be treated with caution since a small data sample will not characterise the covariance matrix accurately, and this type of "sampling noise" is dangerous.

5.4 Contrast enhancement of PC channels

The image vectors which result from the PC transformation have been defined by equation (5.7) above which included two constants which affect contrast and brightness in the final image. The choice of the constant values will depend on several factors such as visual appearance, the display device and any special requirements of further processing operations. Obviously the image values must fall within the allowed range of black to white which, in the present work is a video voltage from 0 to about 0.6 V. Negative values are not allowed, so a lift may be required via the constants b_i . In order to prevent the PC images from taking values above the maximum bound the gain factors may have to be adjusted. The furthest point from the origin in feature space is one whose components are all maximum values, i.e. the tip of the diagonal of the multidimensional cube which bounds feature space. If this diagonal becomes one of the PC axes, the resultant data point will overshoot the maximum, and the gains a_i must be set to a value equal to the inverse ratio of the diagonal length and the length of a side of the hypercube given by

$$a_i = 1/\sqrt{d} \cdot m \quad \text{for } i = 1, d \quad (5.11)$$

where d is the dimensionality of the space and m is the length of the cube side (i.e. the maximum value of any component). This gain adjustment obviously preserves the relationships between the variances of the PC channels. Perhaps the simplest method for choosing the gain and lift is to provide the operator with an interactive means of altering the image to suit his requirements. All that is required is a monitor and a waveform or feature space display to check on possible excursions outside the required ranges. This approach is extremely powerful, but is very seldom found on commercial systems since they rely on software calculation to change each pixel in a framestore.

A further possibility is that of passing the unenhanced PC data through a machine similar to the CONDOR system described in a previous chapter, thus allowing analysis of the intensity distribution of each channel for the optimum contrast stretch. Any of the transformations described for the CONDOR system can be implemented, and if each PC channel is provided with a separate look-up table circuit different transformations are possible with great speed. Note that any PC band can be replaced by the corresponding negative if this improves visual interpretation since the senses of the eigenvectors are arbitrary.

5.5 Covariance acquisition by electronic hardware

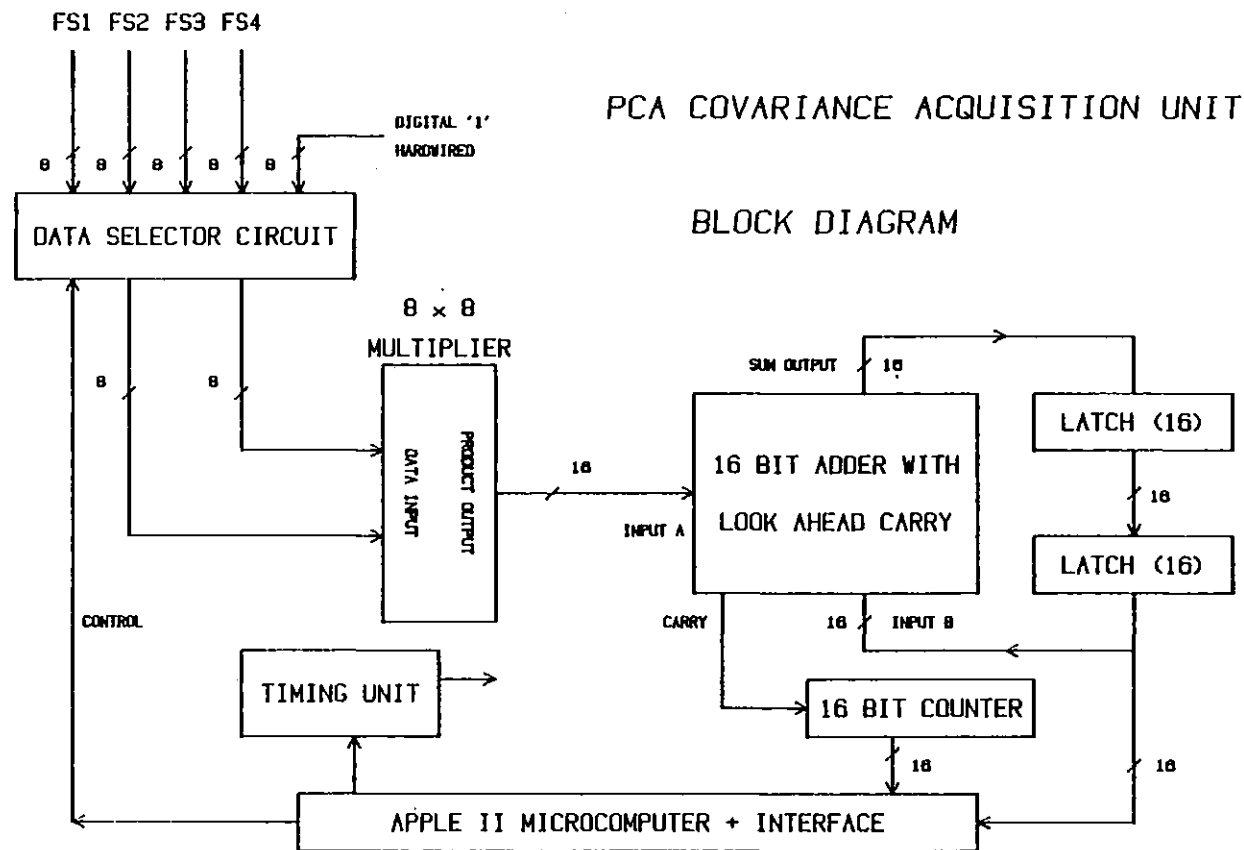
In the present work, execution time is an important parameter and must be minimised. In order to achieve this aim circuitry has been designed which allows the collection of all the elements of the covariance matrix from a fast-flowing digital data stream. The use of specialised mathematical logic eliminates the need for direct processor access to the image data which is slow, and allows the processor to act more as a controller than a calculator, a task to which it is better suited. The system described in this thesis utilises four specially designed digital framestores, each of which holds and displays a 256 pixel square image with 8 bits of intensity resolution. The design of these units will be described in a later section, but for the present they are the source of synchronised four-dimensional image data both for the special logic and for viewing. All calculations are carried out simultaneously with the TV display cycle at a pixel data rate of 7 MHz, several orders of magnitude faster than is possible with a microprocessor.

The elements of the covariance matrix have been defined above by equation (5.3), and the component terms of this equation are calculated by the hardware. The microprocessor is again used to control the flow of data and to read results. In order to understand the mechanism by which this circuit operates it is necessary to rearrange equation (5.3) to give

$$c_{ij} = (1/N) \sum_{i,j} x_i x_j - \mu_i \mu_j \quad (5.8)$$

where the symbols are defined as before. The basic operation in calculating the covariance elements is one of multiplication and accumulation, and since the matrix is symmetric, only half the matrix has to be acquired (i.e. the upper or lower diagonal elements). A simplified diagram of the covariance circuit appears in figure 5.2. The first stage consists of a data selection network which chooses two data streams from the four framestore outputs according to which matrix element is to be calculated. The data is read out of the framestores automatically under the control of the framestore timing circuits, and all mathematical operations occur synchronously with the display cycle. This method allows very fast access to the stored data (within 1/50 S) and avoids clashes with display times which on larger systems can cause the screen to flash or which confine active data periods to the blanking periods. The two selected data streams (which may be identical if a diagonal element is required,

Figure 5.2 : The PCA covariance acquisition circuit



i.e. the "autocovariance") are fed to the input ports of a fast multiplier device type MPY8 HUI from TRW inc. [56]. This is a bipolar VLSI (Very Large Scale Integration) chip which can calculate the 16-bit product of two 8-bit unsigned magnitude numbers (as opposed to two's complement format) in typically 45 nS. The next part of the circuit, built around two 16-bit data latches and a 16-bit addition unit, is concerned with the accumulation or integration process outlined in equation (5.8). When a valid result is available from the multiplier chip, it is applied to one input of the addition circuit. This circuit is constructed from an array of fast 4-bit arithmetic logic units used as adders, and a full carry look-ahead system is implemented to reduce the addition time. Look-ahead carry is achieved by the use of auxiliary circuits which anticipate any carry operations which may arise from each 4-bit addition before the additions have been performed, and modify the results accordingly.

The addition time for two 16-bit binary words is typically under 26 nS. The second input to the adder comes from a data latch which is in turn fed from the adder output via an intermediate data latch in a ring configuration. This is the heart of the digital integrator: Results from the addition are stored by the first latch one addition time after the multiplier data becomes valid. These data are then transferred to the inputs of the adder at the same time as the next pixel product appears on the other inputs, so that the new data is added to a cumulative sum of all the old data. Two data latches are necessary (and a two-stage data transfer) because of detailed timing considerations.

The summation takes place for each of the pixels in the stored image, giving 256^2 operations or 2^{16} . The possible maximum value of any such summation could be 2^{16} times the maximum possible product of two 8-bit numbers, which is also $\approx 2^{16}$, giving a 32-bit number. This number may not be reached in any of the products, but the result is likely to be more than 16-bits in length. This could be accommodated by constructing a 32-bit addition system, but this is slower and more expensive than the 16-bit version. Since the products which are added each time are 16-bits in length, the maximum possible result at each cycle is a 17-bit number (from two 16-bit numbers), and the carry will be signalled by the most significant adder. The upper 16 bits of the adder are therefore replaced with a counter which adds up the total number of carry operations, and saves much complex circuitry. When the process is complete the results are read by the microprocessor (most significant word from the counter, least significant word from the data latch), and the system is reset for the

next matrix element. The microprocessor is synchronised to the operation by means of interrupts at the start of each TV field which are selected by an interrupt control circuit. In the case of a 4-dimensional image data set as used in this system, it is necessary to acquire 10 matrix elements, and the time required is $10/50$ S (the framestores do not give interlaced fields, so that the time is 10 field times rather than 10 frame times). This data rate is high above the maximum achievable with a software system accessing the framestores sequentially. At the end of this acquisition stage the computer can calculate the matrix elements by simple arithmetic, and the eigenvalues and vectors can then be determined.

Equation (5.4)¹² shows the need for the mean values to be acquired, and this is achieved by a slight modification to the same circuit rather than by building a separate system. At the data selector section an extra channel of input data is provided which always contains a binary number 1 (preset or hardwired). This causes the multiplier output to equal the value of the other input, and the addition circuit simply accumulates the pixel values throughout the field. The multiplier is not bypassed in this section because the timing would become excessively complex, and the constant data source is treated just as a fifth framestore data stream.

The circuit described above can perform a single multiply - add - store cycle within the 140 nS pixel time, requiring approximately 70 nS actively. If it were required to use the same system on a larger framestore with shorter pixel times some modification of the timing would be required, but the same strategy still holds good. The system is moderately pipelined as described above (i.e. final calculation of one result occurs while the next is being started), and this could be extended. In extreme circumstances the system could be replicated, and each circuit so produced could be fed with alternate pixels.

5.6 Eigenvector determination

This section describes the method by which the new axes for feature space are calculated by the system. Several approaches are possible to this problem, but all have the common feature that they require computation time, and cannot easily be performed by special purpose hardware. Perhaps the most simple minded approach is to solve the secular equation (5.4) to find the eigenvalues, and then to substitute these values into the characteristic eigenvalue equation

The eigenvectors can then be determined by a Gaussian Elimination method (i.e. the system of equations (5.9) are solved numerically). The simultaneous equations can be solved with reasonable ease, with the use of row and column pivoting to reduce numerical errors, but the solution of the secular equation is not so easy. In the present case the equation takes the form of a fourth-order polynomial whose roots are the eigenvalues. This is the highest order of polynomial which can be solved analytically [57], by first finding one root of a related cubic and then substitution into two derived quadratics. The terms of the quartic are however highly complicated and tedious to compute, and the method is useless in the general case where the number of bands in the image is not four, and may be well over seven. Numerical methods are available for finding the roots of high order equations but they are not simple to use and are usually slow.

Fortunately there exist several iterative methods for approximating the eigenvalues and eigenvectors of real and symmetric matrices such as the covariance matrix, some of which have arisen from the requirements of quantum mechanics. The method adopted in this work is due to Jacobi, and a full treatment is to be found in [58]. The orthogonal transformation matrix T of equation (5.5) is computed as a product of special orthogonal matrices derived as follows: In the original matrix C , a search is made for the largest magnitude off-diagonal element c_{ik} . This serves to define a two by two matrix whose elements are c_{ii} , c_{ik} , c_{ki} and c_{kk} . This matrix can be rendered diagonal easily by a rotation of angle θ , whose value can be determined analytically. Application of this rotation matrix R as in equation (5.5) leaves a diagonal two by two submatrix, and R is applied to the whole matrix C by embedding it into a full-sized identity matrix. This process is repeated several times, and a cumulative transformation matrix is assembled from the individual rotations R . The process can be halted when the largest off-diagonal element of the matrix derived from C (i.e. after repeated application of equation (5.5) with embedded rotations R) is below a given limit. At this stage the resulting matrix has as its diagonal elements the eigenvalues of C . The composite rotation matrix then has as its columns the corresponding eigenvectors of C . The method has been checked for convergence, and always approaches the correct eigenvalues after a large enough number of iterations.

An initial coding of this method has been performed using floating point BASIC on the system Apple][and it has been found that a typical run requires

15 to 20 iterations to reduce the largest off-diagonal element to less than 10^{-5} . The execution time is on average around 10 seconds, which is useable in a working situation. This timescale is beginning to be approached by the fastest software methods, and wastes much time in the mechanics of the BASIC language. It is proposed to implement the method in machine code entirely which will give a large speed increase. This translation is simplified by the fact that no trigonometrical functions are required (the individual rotations θ need not be calculated explicitly, only their sines and cosines are required and these can be found directly), and the most complex operation is that of determining square roots. In order to enhance the speed of operation further it is also proposed to install a special purpose arithmetic coprocessor chip which can perform such operations under the control of the microprocessor but independently, and such a device would have beneficial effects on all the arithmetical operations of the system.

The final result of this stage of the process is a set of normalised eigenvectors derived from the covariance matrix C which describe the rotation required to produce the principal component channels of the image. These vectors have components which define the relative proportions of each original band which must be mixed together to give each new band, and the transformation operation itself is described in the next section.

5.7 Feature space rotation

The system as described so far enables covariance data to be acquired and suitable rotations to be computed on the basis of these data. Equation (5.7) indicates the usual digital method of performing the actual transformation on the raw data, and it involves the calculation of a weighted sum of the input bands, multiplication by a constant and addition of another constant. In software - driven systems this process is carried out pixel by pixel and the resulting principal component images are stored in framestores, sometimes wiping out the raw data in the process if storage is limited. It might be possible to construct hardware circuitry to achieve the same effect digitally but at a much higher data rate such as the TV display rate, although such a system would be expensive and complex, requiring many adders, multipliers and normalisers. In the present work it has been decided to implement the transformation process in a mixture of analogue and digital electronics, using each to its best advantage.

PCA DIGITALLY CONTROLLED VIDEO MIXER

BLOCK DIAGRAM

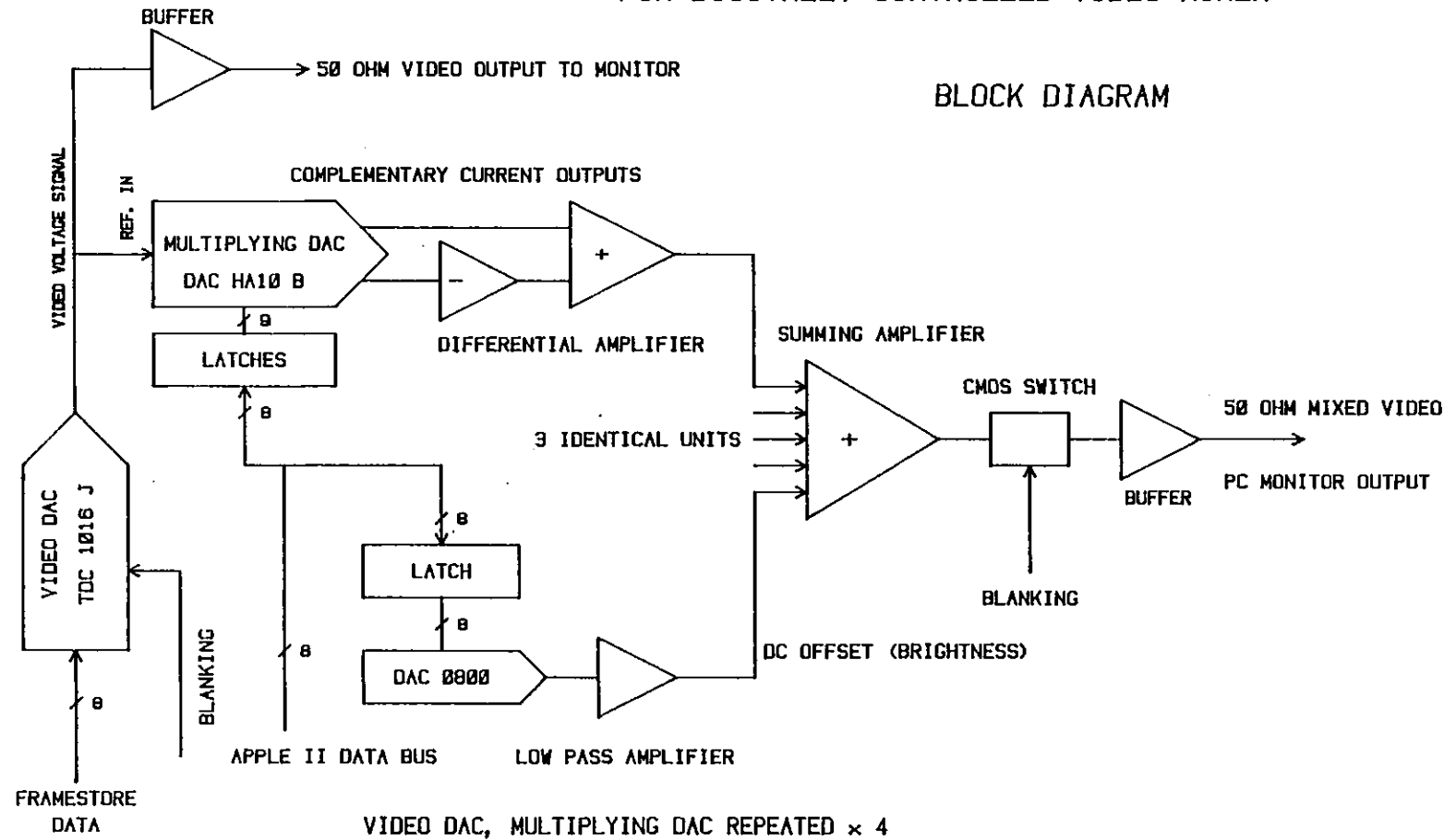


Figure 5.3 : PCA feature space rotation mixer circuit

The system appears in the diagram of figure 5.3, and the first stage is one of digital to analogue conversion. This process is required anyway in order to view the contents of the framestores on a video monitor. To achieve the conversion a high speed and quality DAC is used, type TDC 1016J from TRW inc. [59]. This device is able to accept pixel rates up to 20 MHz and includes deskewing registers and a deglitching sample and hold section as described in chapter 3 (the CONDOR output circuit). It can deliver a standard video signal directly into a 75 Ω load and is relatively simple to use. The result of the DAC section is a set of four analogue video signals which represent the images stored in the framestores. These signals must be multiplied by weighting factors derived from the eigenvectors and then added with an appropriate gain control factor and brightness to form the final principal components. The weighting factors are constant once the eigenvector calculation is completed, and the mixing can conveniently be handled by the use of four multiplying DACs. These devices are similar to normal DACs in that their output signal is proportional to the digital input number as a fraction of a maximum value multiplied by a reference voltage or current. The special feature is that the reference can be an A.C. signal such as a video voltage which can have any amplitude between specified positive and negative limits. This is achieved by the use of CMOS switches internally to control the paths of the signals rather than bipolar transistors which cannot switch very small signals without the addition of a D.C. offset. The devices used here are DAC HA10BCs from Analog Devices [60] which offer a 20 MHz reference bandwidth, and they are configured as four quadrant multipliers i.e. both the analogue input and the digital number may take positive or negative values. Four identical circuits are used, each one being fed with appropriate mixing coefficients in the form of 9-bit digital data, the most significant bit being the sign and the lower 8 bits carrying the magnitude. The outputs of these circuits are fed to an operational adder circuit built around a LM318H amplifier. Also at this stage a variable DC offset can be introduced by means of a low speed DAC attached to a data latch. The output is buffered by an emitter follower transistor prior to display by a video monitor. In order to produce all of the principal component images simultaneously it is necessary to replicate the transforming circuit four times, but initial work in this thesis has been carried out with only one unit because of the relatively high cost of the multiplying DACs.

The system described above has several advantages over the equivalent software process. Firstly it is extremely fast in operation as described and

it is applied in real time to the stored images as they are displayed. Secondly the system is cheaper than the software equivalent, especially in terms of the amount of storage required. If it should be required to gain access to numerical principal component data at a later stage it can be generated in machine code by the Apple, but this is rare. Thirdly the actual transformation process does not require the requantisation operator shown in equation (5.7) since it operates with analogue signals. This can in some instances result in cosmetically cleaner images since the quantisation may cause the generation of unpopulated intensity levels which are blurred in the analogue system by the inevitable noise. Finally, the system as described is more powerful than a simple principal component machine since the output circuit can be used interactively by an operator in conjunction with a feature space display to effect special - purpose rotations according to the task in hand. The principal component transformation gives only a guide as to the best rotation, and operator knowledge may indicate the need for modifications to this basic algorithm. In such cases the speed of the mixing process is paramount since it allows experimentation on the part of the operator without annoying delays.

This section concludes the description of the dedicated PCA hardware used in this study, and the next section describes the design and construction of the framestore units which are used to store the image data. The raw data which has formed the basis for the above sections consists of vectors in a multi-dimensional feature space, with the components of each vector being spectral reflectance samples. The same transformations may be applied in real coordinate space to rotate objects in an image, with two or more dimensions. In this case the coordinates of the object are regarded as a vector random variable, and all the subsequent processing is identical. The rotation orients the object(s) of interest so that the new axes are aligned with the directions of maximum variance. Thus the Principal Components Transformation can be used in pattern recognition applications as well as for multispectral work.

5.8 Digital Framestores

The data source for the PCA system described above is a set of four digital image memories which can store the four Landsat raw images as noted in section 5.5. The design strategies used in the implementation of these stores and their parameters form the substance of this section. The data source for the framestores is the Flying Spot Scanner (FSS) which forms the basis of the MEIC

system which has been described in earlier chapters. The FSS gives four registered video signals as outputs derived from four transparency images. The PCA covariance circuitry requires four digital data streams, and the most obvious method for obtaining these would be to use some high speed digitisers fed with the appropriate analogue video signals. Unfortunately digitisers which are capable of the required data rates, although convenient to use, are highly expensive. In the present case at least two units would be required to give simultaneous digital data for the covariance circuit along with some involved software controlled switching to enable all the signals to be analysed. If intermediate storage can be provided quickly and cheaply it becomes possible to use only one digitiser, and the possibility of downloading digital data directly from a large computer without the need for film images is made available. This would eliminate some of the degradations associated with photographic images, and could give access to many images without the need to perform a manual alignment or registration each time. The use of image stores also permits the study of other processes which require access to digital stored data such as edge detection and texture analysis.

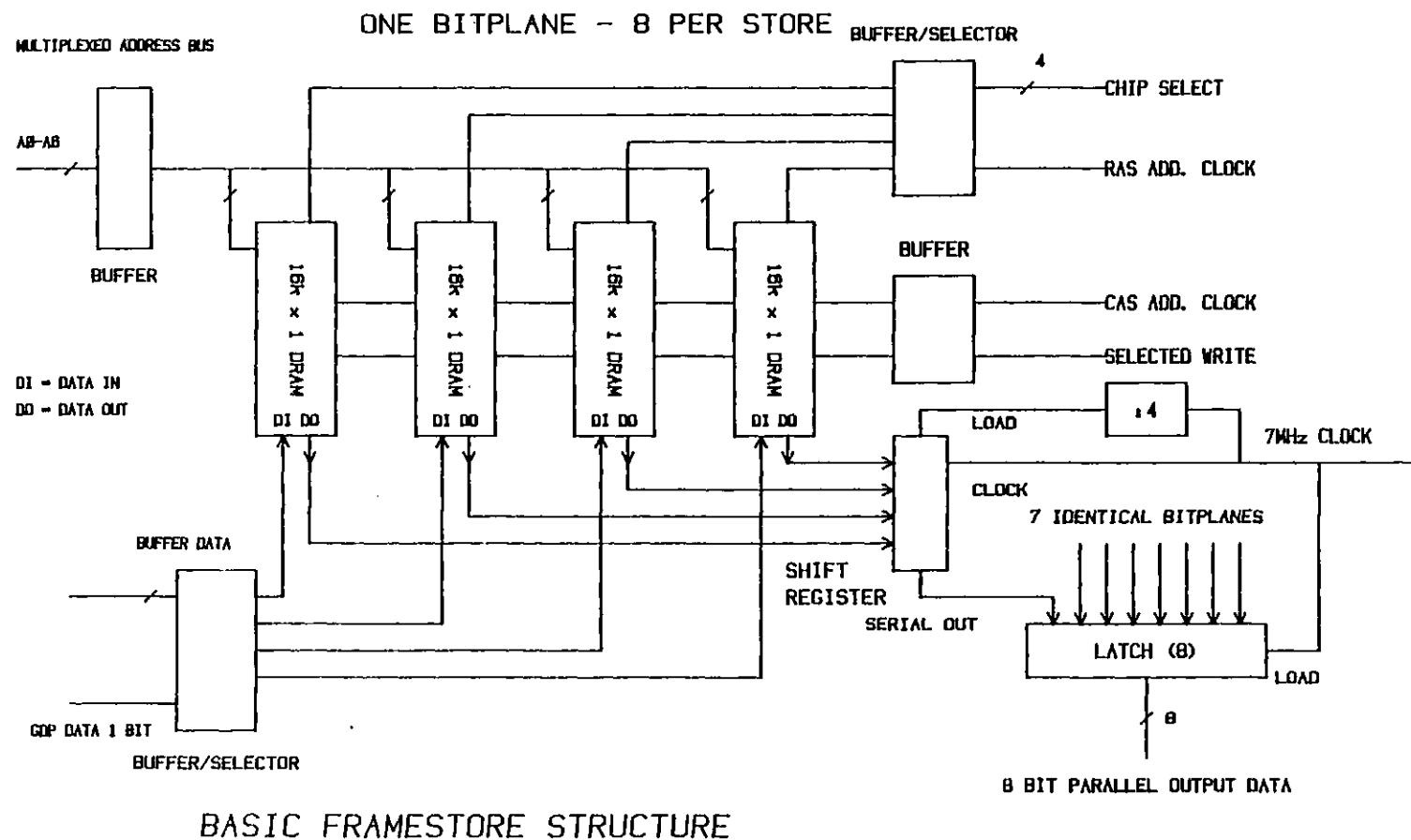
In the light of the above reasoning it was decided to design and construct four digital image stores to form the basis of the PCA system. The size of these stores was set at 256 pixels square as a compromise between resolution, speed of operation and memory considerations. It was felt that a store with fewer than 256 elements on a side would be incapable of sufficient spatial resolution, and that sizes above 512 elements square would incur excessive memory costs and timing difficulties. The other major consideration which determined the final image size was that of memory size. It was decided to give each pixel a full 8 bits of intensity resolution in each spectral band. The results of chapter 2 suggest that at least 7 bits would be required, but the extra bit allows for the possibility of using better quality data (perhaps from an image integrator). A further option which is available is to use one of the bits as a graphics overlay plane on which to display information relating to calculations or system status. The size of store described above requires the use of 256×256 bytes of memory or 64 kbytes, equivalent to the normal address capability of an 8-bit microprocessor. Thus the total memory size is one quarter megabyte for the four stores.

The memory used must obviously be capable of being both read from and written to, i.e. Random Access Memory (RAM), and it must have sufficient speed

to allow its contents to be displayed at standard TV rates as a video picture. Because of the size of memory needed, and to promote compact construction, Dynamic (as opposed to Static) RAM is employed. This type of RAM offers very high storage densities, response times as low as 120 nS for fast devices and low cost, but their use entails complex timing considerations since stored data must be continuously refreshed if it is not to be lost. This fact arises from the compact design of the memory elements on the chip which take the form of leaky capacitors with time constants of about 2 mS. Unlike the logically stable latches of static RAMs the data must be restored at each location within the refresh period. Modern DRAM devices offer 64 kbits on a single chip, so that 8 chips might suffice per store, but the great "depth" of these devices and their relatively high cost compared to the older 16k kbit chips proves disadvantageous in this application.

In order to comply with the speed requirements of an image or framestore the memory structure must be carefully designed. The simplest method would be to read out a byte each time a new pixel value is to be displayed. Since 256 pixels must be displayed in each TV linetime (52 μ S) this implies a minimum data rate of about 5 MHz, or 200 nS per pixel. Although DRAMs are available with read access times of 150 nS and less, they require a recovery time after each read or write operation, thus giving a total cycle time of around 250 to 300 nS. The simple approach which uses a single contiguous block of memory is thus not practical, and some way of increasing the effective memory bandwidth is required. In chapter 2 a description of a digital image store was given which introduced the concept of bitplanes, or framestores with one bit of intensity resolution per pixel, a combination of which could be used to form a complete store with many bits per pixel. This structure is the basis on which the framestores described in this chapter have been implemented. Each framestore is composed of 8 single bit stores of the same size run in parallel, and the outputs are combined to form the required data stream. A schematic diagram appears in figure 5.4 which shows four 16k x 1 bit DRAMS (type 4116N-3, 200 nS access time [61]) which are used in a 16k x 4 bit configuration. Each bit represents information for a single pixel, and there are 64 k bits in these four chips (note 1 k bit = 1024 bits = 2^{10}). The data output pins of the DRAMS are connected to the parallel loading inputs of a 4-bit shift register. This device is clocked at the required pixel rate which in this case is 7 MHz (140 nS per pixel), but new data is only loaded into it every fourth clock cycle. In this way the reading frequency of the DRAMS is reduced by a factor of four to

Figure 5.4 : A single framestore bitplane (256 by 256 by 1 bit)



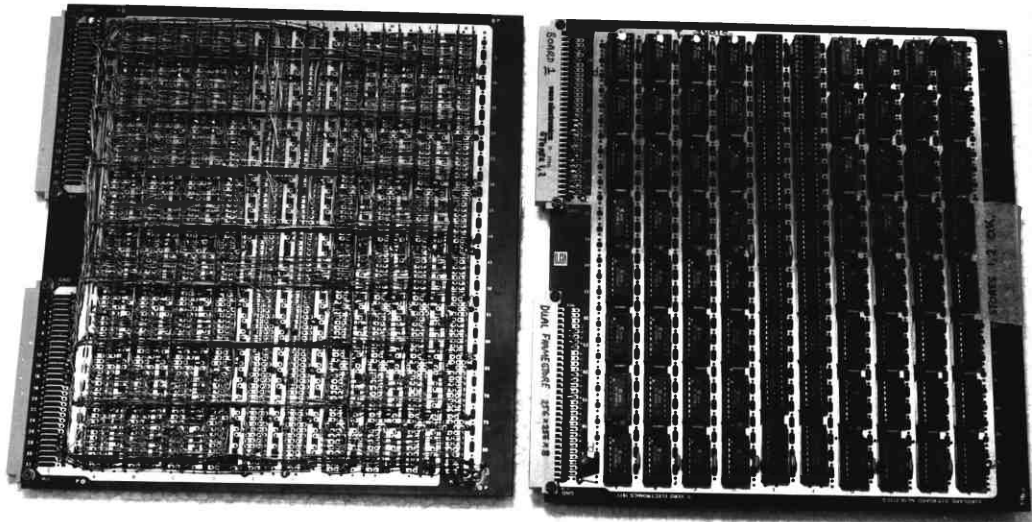


Figure 5.5 : The two framestore boards

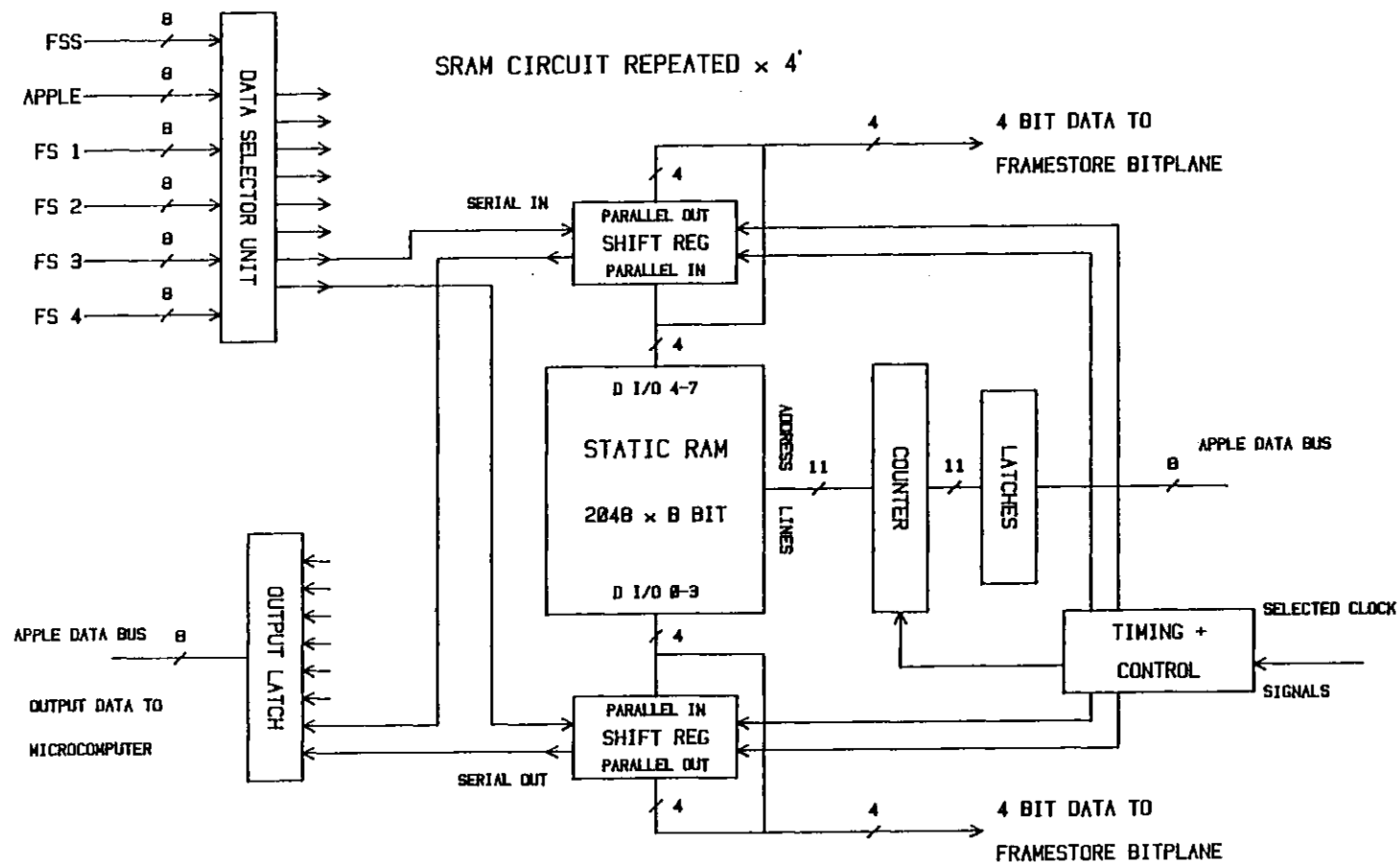
1.75 MHz, with about 570 nS between accesses, which is well within the capability of the memory chips. The graphics display board which was described briefly in chapter 3 uses the same principle but the shift register is 8 bits long since the display size is four times larger (512 by 512 pixels). The basic memory unit of four chips is repeated 8 times in each framestore, with one unit holding the data corresponding to a particular binary digit of each pixel's intensity word. The shift register outputs are connected to a fast octal latch (type 74ALS574 in the figure) which deskews the data before it is output to other parts of the system, thus eliminating any differences in response time between the various shift registers and memories.

The DRAMS used in this circuit are very compact, having only 16 pins, four of which are power supplies. For this reason all address information is multiplexed and supplied in two phases, controlled by two clock signals labelled RAS and CAS in the figure. These signals are not symmetrical square waves, and are generated from a master clock running at 14 MHz by reading out a small permanent memory (PROM) in which they are stored. The address information is generated by the master control chip which is a Graphics Display Processor (GDP) type EF 9365J made by Thomson-EFCIS [62],[63]. This chip provides all the necessary memory control functions to cater for refresh and display of the stored data and has in addition the capability to draw characters and vectors in the memory under the control of the system microprocessor. It is designed to handle a single plane of memory (i.e. a one bit store equivalent to one four chip unit in figure 5.4), but with suitable external logic and buffering it has been made to service 32 such stores which comprise the four image memories. The process of reading out the data for display achieves the necessary refreshing during the active part of a TV field period, and the GDP performs three subsidiary bursts of refreshing during the vertical blanking interval to ensure data retention. Two modes of access are possible to the memory, these being a display access in which all the DRAMs are read out simultaneously and a GDP write access in which only selected chips are addressed, the selection being performed by signals which gate the RAS clock as shown in the diagram. GDP writing can occur outside the display and refresh cycles at peak rates up to 1.5 million pixels / S. The use of a GDP eliminates much of the complex logic which would otherwise be required to service the memory, and also allows graphics characters to be written in selected bit planes. Data may be written into the memories by the GDP or by a special purpose buffer which will be described in the next section.

The framestores are constructed on 6U x 220 mm Eurocards which fit into a rack case, and two stores are accommodated on each card. Connections are made with wire wrapping cable with very high densities. The two boards are shown in figure 5.5, one pin side up, together with a scale. Each board contains two banks of 32 DRAM chips (the outer 4 columns on each side) and the two centre columns house the 16 output shift registers and the two deskewing latches at the top. Liberal use is made of power supply decoupling capacitors in order to eliminate transients which arise from the dynamic nature of the operation of the memories. Three supply voltages are provided (+5V output supply, -5V bias and +12 V for the main parts of the chips), and they are distributed in grids to provide the shortest possible path from any chip to the decoupling capacitors. Power is available from a high performance switched mode supply which also serves all the TTL logic in the rest of the system. Connections are made via two 64-way indirect edge connectors to the backplane of the rack which distributes all signals as required. The GDP is situated on a smaller card elsewhere in the rack. All address and clock signals are distributed after passing through 22 ohm resistors. These serve as series terminations which minimise the effects of reflections from signal lines which are sometimes very long and terminated only by high impedance NMOS input stages in the DRAMs. This form of termination at source attempts to absorb reflected energy after one pass down the line without the excessive current drain which would be incurred by a parallel termination of 50 to 100 ohms. The RAS clock line in particular is susceptible to reflections and a further termination consisting of a series resistor and capacitor was used to give an AC trap at the end of the lines on each board. The use of a capacitor again cuts out unnecessary DC current drain.

5.9 The buffer unit

The framestores as described above can only be written to by the GDP under the control of the microprocessor. Selection logic is provided to enable writing to selected bitplanes and stores, but the loading of a complete four band image by this means would be painfully slow and complex. The raw data is derived from the MEIC via the digitiser which forms part of the CONDOR system after suitable amplification. The MEIC contains a master clock whose frequency is 10.25 MHz, and it is not a simple matter to synchronise it with the GDP clocks. For this reason it was decided to build an intermediate buffer unit equipped with static RAMs which could hold a portion of image from the MEIC and then load it into the framestores at a later time. The writing into the



BUFFER UNIT BLOCK DIAGRAM

Figure 5.6 : Framestore buffer unit

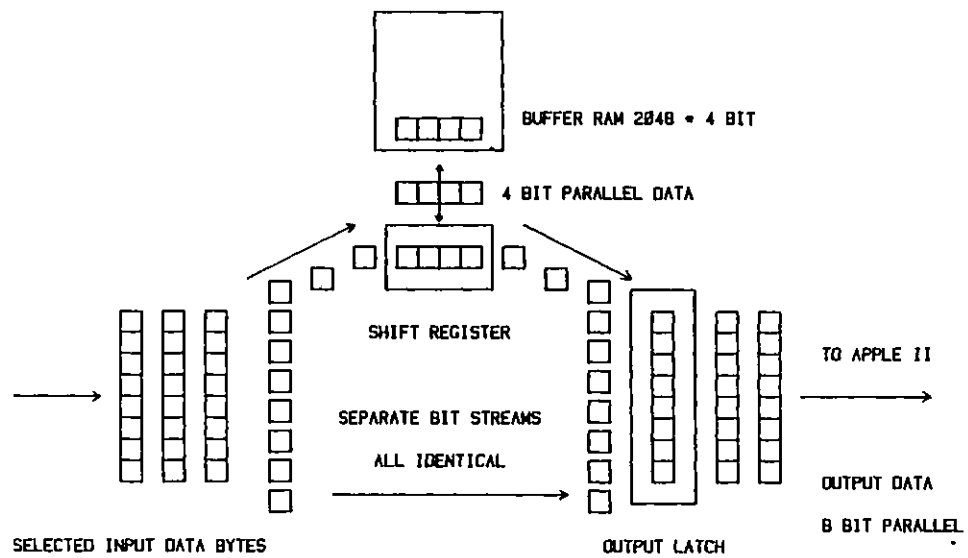


Figure 5.7 a : Data input to buffer and output to Apple II

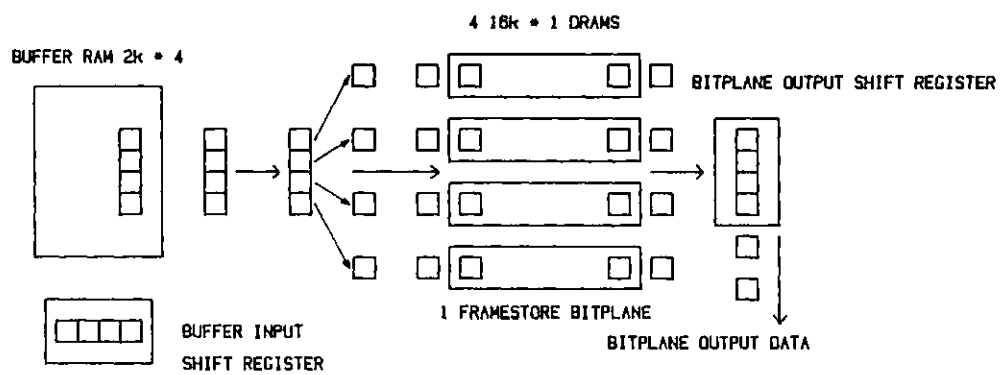


Figure 5.7 b : Data output from buffer to framestore

framestores is achieved synchronously with the display cycle. When the display cycle is underway the correct memory addresses will be available to cause the reading out of groups of four pixels ready for the shift registers. If at the same time the required input data is applied to the data in pins and the write control line is activated, the framestores will be loaded. In principle, with a large enough buffer memory, the entire 64 k bytes could be downloaded inside one TV field (1/50 S), but in practice static RAMs are relatively expensive both in cost and space. As a compromise it was decided to provide a buffer capable of storing 32 TV lines of data, using 8 kbytes of SRAM in four devices.

A simplified diagram of the buffer unit appears in figure 5.6, but only one of the static RAM chips is shown for clarity as the remaining chips are connected in an identical fashion. The SRAM device is a 2k x 8 bit chip type MK4802-1 [64] with a read or write access time of 120 nS, and it is configured as two four bit wide memories, each one holding data about one particular bit in the pixel intensity bytes. Data is input from one of several sources via a selector circuit which also gates through the appropriate clocking signals. The sources of data are the four framestores, the FSS via the digitiser and the Apple][computer. These sources all give picture information at different rates, ranging from 7 MHz from the framestores to less than 1 MHz from the microprocessor. The incident information is split to form 8 serial bitstreams from the single 8-bit parallel source, and the serial data is fed to 8 shift registers, two of which are shown on the figure. The buffer loading process is almost exactly the reverse of the framestore output process, and a similar reduction in the memory speed requirement is achieved. As each set of four pixel bytes is clocked into the shift registers a counter circuit generates a writing signal which stores the data on the SRAM and advances an address counter to access the next memory location. The data is stored in the RAM in a composite form, with the two halves of each byte being derived from different bits of the image data. The flow of information is depicted graphically in figure 5.7a. The shift register clock signals are gated by the appropriate video blanking signals and by a 32 line window signal. This signal is derived from a pair of counters and determines which block of 32 lines out of a field of 256 lines will be sampled. The window location is controlled by the system microprocessor during the loading program. The data from the FSS is in the form of interlaced video signals, and to minimise distortions only data from even fields is acquired. The video information from the framestores is non interlaced so that it is always valid.

The buffer unit is designed to output data both to the microprocessor and to the framestores under software control. The computer interface has been designed for data transfer at rates under 1 MHz by machine code programs since it was felt that sufficient speed would be gained by this means. An alternative would be the use of Direct Memory Access (DMA) in which the transfer is mediated by a special purpose DMA controller while the micro is disabled. This option may be considered further in future if greater demand for buffer I/O arises. The data is read out into the same shift registers in parallel from the RAMs and it is converted into serial form. An 8-bit latch deskews the separate streams of data and reforms them into a parallel stream which is read by the computer. This stage of the data flow is also depicted in figure 5.7a. Output to the framestores is somewhat different as described above, and occurs synchronously with the TV display cycle. Instead of passing through the shift registers the data (still in 8 4-bit parallel streams) passes through a selector circuit to the input pins of the framestores. The selector determines which of the stores is to be written to, and whether the source of the data is to be the buffer or the GDP (when vectors are drawn). A similar selection is performed on the framestore write signal which is activated in synchronism with the DRAM access cycle. Again the validity of the data is indicated by the 32 line window circuit which is now synchronised to the framesore timing signals. The flow of data for this case is shown in figure 5.7b. Data can be loaded into each framestore in 8 TV field times ($4/25$ S) and the buffer can be loaded from the FSS in 8 TV frame times ($8/25$ S - interlaced video) resulting in a very fast transfer. Again the philosophy of controllable high speed circuits instead of direct microprocessor intervention allows much more interactive operation and reduces the 'dead time' during which the operator is idle.

Throughout the various loading and unloading cycles the buffer's logical configuration is altered by the microprocessor to suit the task in hand. This strategy allows much of the same circuitry to be used but with different timing signals, thus saving valuable space and chips. Even so the buffer unit is probably the most complex part of the framestore system since it must accept at least three different sets of timing signals with different frequencies and waveforms. Construction is as before, on a double height Eurocard with wire wrapping connections and two 64 way edge connectors to the backplane. The image displayed from the framestores is square, and placed centrally on the monitor screen whereas the FSS image is rectangular with an aspect ratio of 4:3. The transfer process introduces some spatial distortion (a horizontal compression)

which must be taken into account in interpreting the images, but it is not a disadvantage. The problem of synchronising the buffer to external sources of video data has been approached in this system by the use of asynchronous loading and unloading which, because of the requirement to reduce memory access frequencies has proved convenient. In a commercially manufactured system a closer synchronism between the various components would be designed in from the beginning, and image distortions might be avoided. It is worth noting that most software operated image processing systems use 512 x 512 pixel square framestore memories and overcome aspect ratio problems by sub- or resampling algorithms.

5.10 The PCA system - a summary

This chapter has described some possible methods by which multispectral or similar data may be manipulated with reference to its feature space distribution so as to produce decorrelated images. The methods are based on the rotation of the axes of feature space by means of a suitable matrix which rearranges the data. The purposes for which this transformation is attempted may include improved classification accuracies, data compression or reduction, noise removal or suppression or simply improved visual clarity. The rotation process is carried out in real time on simultaneous video signals representing the raw images and is controlled by a microprocessor. As an aid to the efficient determination of the required rotation in a multidimensional space the Principal Components Analysis or Karhunen- Loeve transformation has been implemented using a combination of fast digital circuits and software. Analysis of the form of the feature space distribution in terms of covariances is performed at high speed by a dedicated circuit which eliminates the time consuming acquisition of this data by the computer which would otherwise be required. The correct rotation is determined by an iterative method due to Jacobi in software, and this process will be accelerated by the use of an arithmetic processor in the near future. The aim of the system is to enable an interactive use of the principal components analysis and related image rotations which is not possible at present on software driven processing systems.

As the PCA system described above requires access to digital image information, four framestore units have been designed and constructed which are each capable of holding a 256 pixel square image with 8 bits of intensity resolution per pixel per store. Image information may be obtained from the Flying Spot Scanner which forms part of the MEIC or it can come from the

microprocessor memory. Images may be stored on disks or tapes and downloaded from larger computers via terminal lines if required. All data transfers to and from the framestores are mediated by a special purpose buffer unit which performs the necessary format conversions and timing alterations. The availability of four synchronous sources of digital image information allows the possibility of future work on such data also at high speeds using digital circuits.

The next chapter will show some examples of image processing carried out with the aid of this system, and the results of the PCA applied to some test images. A review of the various attempts at interpreting the contents of the various PC bands will be presented. It must be remembered that the PCA is not a process which is limited to Landsat or other multispectral imagery, and it can be used wherever multiple data set images have to be interpreted or compressed.

6.1 A practical PCA system

Before any results obtained by using the PCA system described in the previous chapter are presented, it is useful both to review the method and to consider the effects of contrast modification in the original data set. This section will restate the aims of a principal components analysis and illustrate its effects on feature space distributions.

The type of PCA considered in the present work has as its raw material a multispectral or similar set of images, each one measuring some different property of the scene in question, such as the spectral reflectance in a given band of wavelengths, and the data is usually correlated. Correlation in this context implies that the raw images are similar or contain largely the same information with only minor differences between them. Two possible motives are usually given for the use of a PCA, these being (a) a reduction of the dimensionality of the data set and (b) a separation of the data into distinct, uncorrelated variables. Followers of the first motive seek to rearrange the information in such a way that only a few images can be used to represent the data without significant loss of fidelity. They are able to accomplish this in most cases because the higher order PC channels contain information relating to progressively finer differences between the raw images, and they contain progressively smaller proportions of the total information present. Followers of the second motive may use some of the aims of the first motive, but they are usually more interested in the exact nature of the fine differences which become apparent in the PC channels, and they seek to describe their origins.

The PCA method is based upon the use of the variance-covariance matrix to characterise the correlations between the raw images. As described in the previous chapter, this matrix has as its elements the variances or covariances of the images either collected over the entire image or sometimes estimated. It is important to remember that variances and covariances in this work are defined with reference to the mean of the distribution concerned (i.e. the feature space distribution), and in the case of a multidimensional data set, this means the centroid of the distribution. The method consists in determining the eigen vectors of the covariance matrix, and thus a new set of orthogonal axes in

feature space oriented in the directions of maximum variance in feature space. The transformation then consists of a rotation of the feature space distribution onto the new axes (or equivalently, a resampling or reprojection onto the new axes). The process is illustrated simply in figure 6.1 which depicts a simplified 2-dimensional feature space distribution. The first step is to determine the centroid and the directions of maximum variance as shown. The second step is to perform the rotation to give two new bands whose feature space distribution appears in figure 6.1 b. These two bands are uncorrelated (their covariance coefficient is zero), but they are not necessarily independent. In order to be independent, the distribution would have to be separable into two 1-dimensional distributions of random variables with normal or Gaussian shape, and while this may be true of some clusters or feature space plots, it cannot be said to hold in the general case. The final stage is one of shifting the origin of the axes so that the resulting images do not have negative values. This step is required since the definition usually applied to an image specifies positive values only, with black displayed for a zero value and white for a positive maximum value. This step can be viewed as a redefinition of the display coding so that black now corresponds to some maximum negative value and white to a corresponding positive value, mid grey signifying a zero. The production of negative values is inevitable in the PCA process and often causes worry about their meaning, but as the above argument shows, it is merely an artefact of the process, and a shift in origin does not affect the results. In a practical situation, a TV monitor may not be able to display negative video voltages, so an offset is added.

An important point to notice is the effect of applying a contrast enhancement to the axes before and after the PC transformation. The raw bands can be stretched until they each occupy the full intensity range available (this corresponds to enlarging the ellipse in the figure). Such treatment does not alter the shape of the ellipse, which may still be very narrow if the images are well correlated. After the PC transformation, it is possible to apply stretches independently along the major and minor axes of the ellipse, and the shape can thus be distorted in such a way as almost to fill the entire feature space. This gives a much wider range of colours in a composite image, and separates out any clusters which may be present, allowing more accurate classification to be achieved.

Often the processing chain between the image acquisition system and the

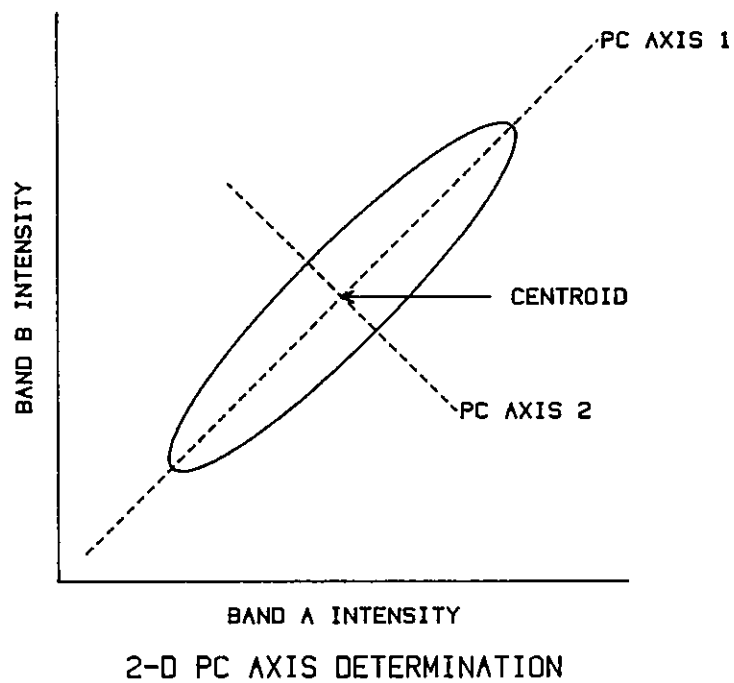


Figure 6.1 a : Original feature space plot

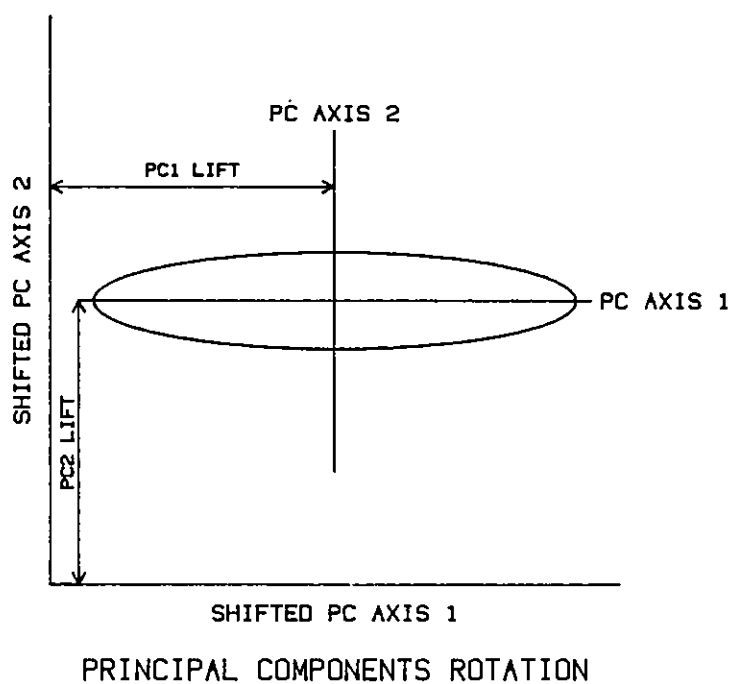


Figure 6.1 b : Transformed feature space plot

PCA system is long and tortuous, and many modifications to image contrast and lift can occur, all of which affect the exact form of the feature space distribution. The effects of these changes on the PC channels can be deduced quite simply by reference to the diagrams shown in figure 6.2. The first diagram shows the effect of a DC shift on the raw image channels which shifts the entire feature space distribution without affecting its shape. The original PC axes are translated by a distance equal to the DC shift on each axis, but the eigenvector directions and magnitudes remain unchanged. No DC shift is observed in the PC channels if the inputs are referred to the centroid of the distribution, but in the present equipment which mixes video signals, an extra DC shift has to be applied in order to compensate. Since this brightness correction is under the interactive control of the user, this presents no problem.

The next diagram in figure 6.2 shows the effect of a linear contrast stretch applied to the input images, not necessarily the same to both. The feature space distribution is extended by the stretch, and the centroid moves. The change in shape produced by non-equal stretches of the input channels causes the directions of the eigenvectors to change, but the new vectors still point in the direction of maximum variance, and the final results will be identical except that they will have a greater or lesser contrast since the relative positions of image classes in the feature space distribution are still the same—only the distances between them are altered. The situation becomes much more complex if the raw images are subjected to a nonlinear contrast stretch such as those demonstrated in chapter four. In such a case the shape of the feature space distribution may become grossly distorted, and exact characterisation of the result requires the specification of the stretching functions and the feature space distribution. In general, as long as the overall relationships between the various image classes are maintained, the results will be similar to the unstretched case except for a variation in gain and lift. Thus the many unavoidable changes to the raw data which occur during preprocessing stages need not be disastrous.

The effect of noise was treated in the previous chapter, where it was shown that additive, uncorrelated and image-independent noise did not alter the computed directions of the eigenvectors, only their magnitude. The important parameters which must be computed are the directions of the eigenvectors. Once they are known, the results are only cosmetically affected by changes of

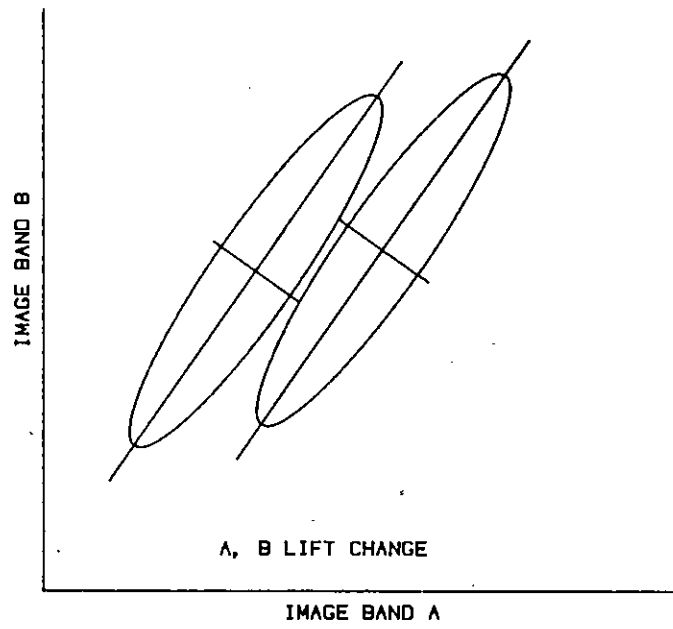


Figure 6.2 a : Lift or brightness change

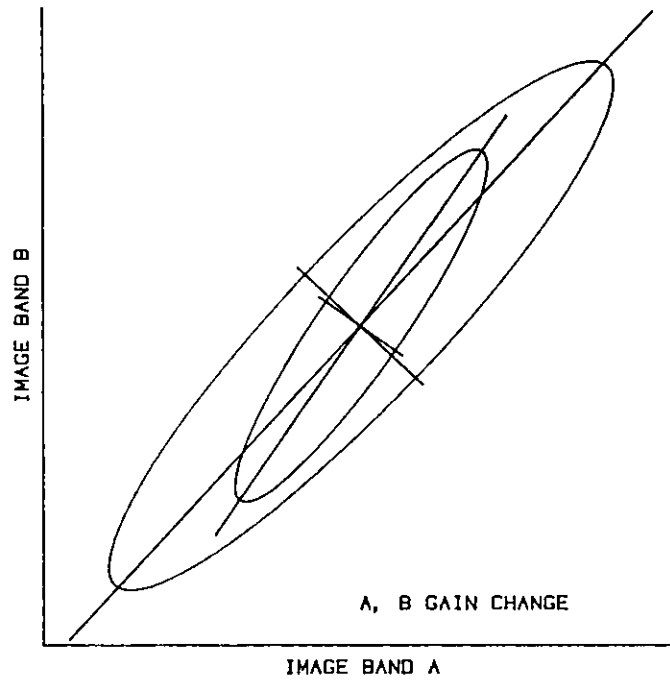


Figure 6.2 b : Gain change- centroid position preserved here

contrast or brightness, and negative directions can be used freely without changing the results of any post-processing. The present system is capable of performing linear contrast stretches on the resulting PC channels, either to give maximum contrast or scaled according to the appropriate eigenvalue, but it is possible to perform many stretching operations on the new images by applying some of the techniques described in chapters 3 and 4. Some examples of this type of post processing will be described in a later section.

Succeeding sections will illustrate the use of the PCA system which was described in the previous chapter using both 2- and 4- dimensional data sets, and some attempt at the interpretation of PC channels will be presented. The performance of the system will be reviewed and possible means of improvement will be given.

6.2 A 2-dimensional PCA example.

This section will demonstrate the PCA process with reference to a two-dimensional example to clarify its effects. The raw data comprises the two near-IR band images (6 and 7) from a Landsat frame of the south west part of Spain which will be used further in a later section. The raw images, which are stored in the framestores described previously, appear in figure 6.3, and a feature space display is shown in figure 6.4. The feature space distribution is obtained by plotting the video signal from one image against that from the other image using an oscilloscope as an x-y display. Variations in intensity arise not from any beam modulation, but from the integration of the camera which records the picture. Unfortunately, it is very hard to represent the subtle detail in such a display photographically, but in this example, it is only the broad form of the distribution which is relevant, not the positions of each sub-cluster in feature space. Points of interest are the bright tip of the distribution, caused by a significant area which is light in both images, this being due to vegetation (in the centre, next to the river), and a bright area towards the lower left which corresponds to the sea (dark in both images). The overall shape is not exactly elliptical, but the narrow spread about the diagonal indicates close correlation between the images (both raw images have been stretched to full range before use). A careful look at the original images will reveal several small differences such as at the edge of the dark area just above the coast where a triangular area pointing upwards is much more distinct

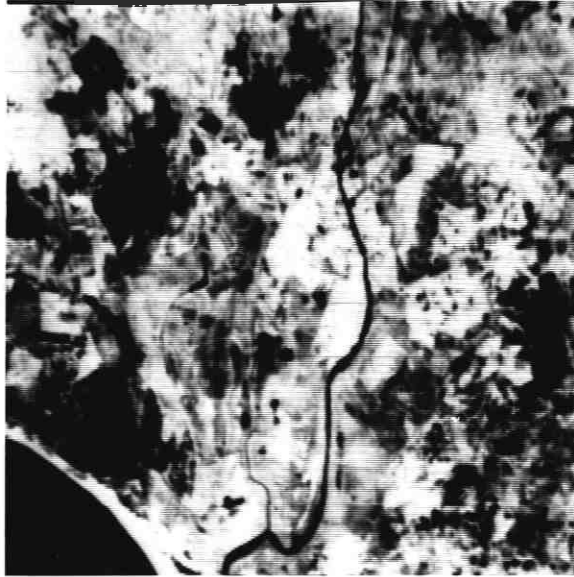


Figure 6.3 a : Band 6 original image



Figure 6.3 b : Band 7 original image

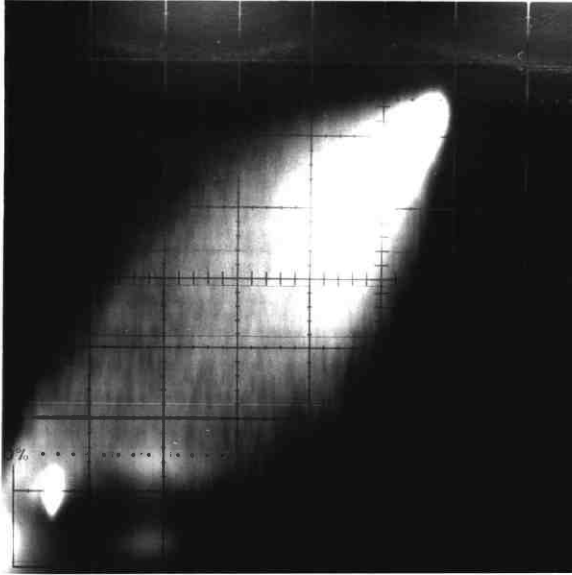


Figure 6.4 : Band 6 vs. Band 7 feature space plot

in band 6 than in band 7. Since this image will appear in a later section, a detailed interpretation of the scene will not be attempted here, rather the general properties of the PCA will be demonstrated.

The first step is the acquisition of the image statistics under software control using the special purpose hardware described in the previous chapter. The results are shown in table 6.1, and they reflect the nature of the feature space display in that the covariance term (band 6 against band 7) is smaller than the individual variances, thus indicating the presence of correlation.

TABLE 6.1 : Image statistics

Mean of band 6 = 133.39 (range 0 to 255)
Mean of band 7 = 141.61 (range 0 to 255)

Covariance matrix :

$$\begin{aligned} \sigma_{11}^2 &= 693.04 & \sigma_{12}^2 &= 505.20 \\ & & \sigma_{22}^2 &= 706.05 \\ \text{Trace} &= 1399.09 \text{ (sum of variances)} \end{aligned}$$

Eigenvalues :

$$\begin{aligned} \lambda_1 &= 1204.79 \\ \lambda_2 &= 194.30 & \text{Sum of eigenvalues} &= 1399.09 \end{aligned}$$

Eigenvectors :

$$\begin{aligned} \mathbf{e}_1 &= (+0.703, +0.712) \\ \mathbf{e}_2 &= (+0.712, -0.703) \end{aligned}$$

These data show several features worthy of note: firstly, the sum of the raw image variances (1399.09) is equal to the sum of the eigenvalues which are obtained as the result of the PCA. This indicates that no loss of image variance has occurred. Secondly, the first eigenvalue is much larger than the second, indicating that the first principal component image will contain most of the available variance. The two eigenvectors are orthogonal (their dot product is 0), and they lie broadly along and perpendicular to the leading diagonal of

the distribution (roughly at 45 degrees to the horizontal). Recalling the shape of the feature space distribution, the origin of the large variance of the first PC is now obvious, since the first vector points along a direction of high variance. Note that the image contrast is proportional to the square root of variance (i.e. to standard deviation) so that the difference is not as large as it may at first appear. The next step is to apply the eigenvectors to the mixer unit to produce the PC images, and these appear in figure 6.5. The first image is a weighted sum of the raw images (0.703 times band 6 + 0.712 times band 7), and this can be seen by close inspection (e.g. the triangular area noted previously is grey whereas it is black in band 6 and light in band 7). The large vegetated area in the centre remains white as it was in both raw images. It must be remembered that the image is only composed of 256 pixels per line and 256 lines, so that the resolution is not very high. Nevertheless, some degradation in resolution is apparent as 'blurring' or smoothing, indicating that the mixer unit does not possess sufficient bandwidth. The second PC image is a weighted difference of the raw images, and it has been contrast stretched to the maximum capability of the system (i.e. not following the relationship between the eigenvalues). The vegetated area now appears as mid grey, but the area below and between the two arms of the river appears significantly darker than the average, and observation of the original images will show that there were some small differences in this area. The triangular area noted above now appears lighter than average, again indicating an area of difference between the raw images. The 'relief' effect in the 2nd PC image is due to poor mixer performance causing an effective differentiation of the image, and should be ignored - only broad areas of lighter or darker shade than the average indicate differences between the two raw images. The problems associated with the mixer and their solution will be covered in the next section since the presence of false artifacts obscures detail and hinders interpretation.

The entire process of acquiring statistics, calculating the eigenvectors and producing the images requires only a few seconds. The operator can modify the contrast and the brightness of the images interactively, and negatives images can be displayed instantly. This simple example has served to show the operation of the PCA in a simple way, and to introduce a feel for the process. As in all image processing it is (or should be) vital to know what is happening, and to be able to check the form of the results from the statistical data available.

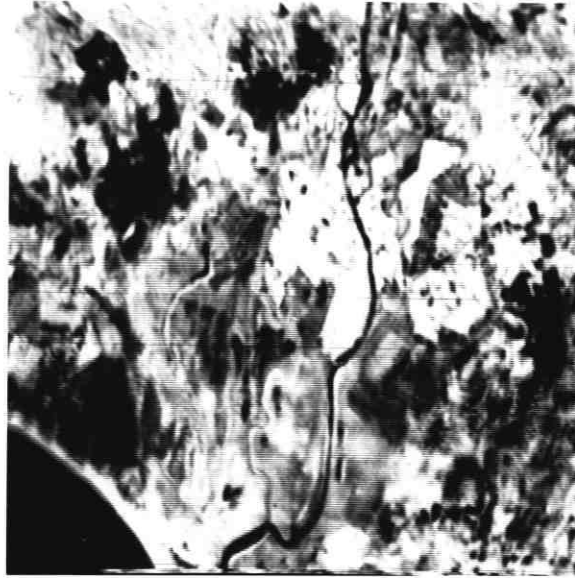


Figure 6.5 a : The first Principal Component image

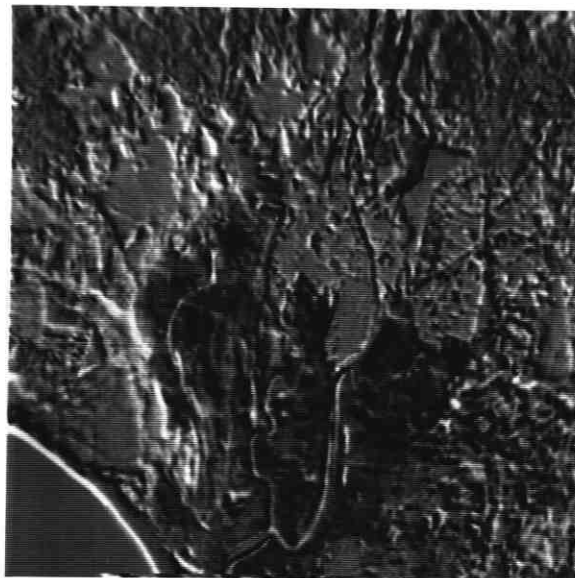


Figure 6.5 b : The second Principal Component image

6.3 System performance

This section describes the performance of the PCA system, in particular the digitally controlled video mixer unit which actually performs the feature space rotations. The 2-dimensional example of section 6.2 presented some test results from the system which demonstrated that PC's can be performed quickly and easily. Unfortunately, the performance of the mixer leaves something to be desired both in terms of basic design and bandwidth. The raw images from the framestore are composed of 256 video lines, each with 256 pixels, and the pixel data rate is 7MHz. The video DAC which gives the analog output signal corresponding to the stored digital image is capable of producing signals with a bandwidth in excess of 15MHz, but the operational amplifier stage which buffers and inverts the video signals restricts this to about 7MHz. The effect of this is to smooth out the sharp pixel edges somewhat, and this bandwidth is perfectly adequate (sharp pixels do not imply a higher resolution image and the maximum image frequency is 3.5 MHz by the Nyquist criterion). The area at fault is indicated in figure 5.3 by the output amplifier stage following the multiplying DAC (the two OP AMP triangles). The DAC works by dividing the input video signal by means of a resistor network which is controlled by switches. The positions of these electronic CMOS switches are determined by the digital data presented to the device. The output is in the form of two complementary currents whose sum is constant (fixed by the input video signal and the network total resistance) but whose relative proportions vary according to the digital data. When the most significant bit is a '1' and all the others are '0' (i.e. the digital number is exactly half full scale) then the two output currents are equal, and if one is subtracted from the other the result will be zero.

The differential amplifier circuit is designed to perform this subtraction, but a single operational amplifier cannot be used as a subtractor. This is because the input impedances seen by the two currents would not be equal since one would be feeding into a virtual earth (the inverting input) whilst the other sees a resistor to ground. To overcome this problem a second amplifier is employed to invert one of the current signals before it is added to the other current in an operational adder.

Several areas of possible difficulty arise in this circuit which conspire to degrade its performance. The first is due to feedthrough capacitance in the multiplying DAC itself which effectively couples a small proportion of the input

signal to the output even when the digital data is such as to multiply the input by zero. This error increases with frequency but it is not serious when viewing images as the multiplied signal usually far exceeds any feed through error. Stray capacitance must be minimised by careful layout, but since the DAC HA10 is one of very few suitable devices available, better performance is hard to achieve. A much more serious problem arises in the choice of amplifier and the circuit itself. Firstly, amplifiers with low input bias currents should be used for fast operation and exact cancelling of the two currents, but this is a feature which becomes increasingly hard to find as the amplifier bandwidth goes up (some devices require several microamps). A solution to this problem is the use of fast FET input amplifiers which take only a few tens of picoamps. Speed of response is the cause of the second and worst problem. The circuit shown works very well at low frequencies, but at high frequencies the delay time caused by the propagation of one signal through the inverting amplifier whilst the other signal travels directly to the summing amplifier gives rise to errors. The errors take the form of spikes or short pulses wherever there is a very sharp and high amplitude change in the video signal. The effect on the image is similar to that produced by differentiating the signal- edges are 'enhanced', and the output cannot be reduced to zero under digital control. As an example figure 6.6 shows an image multiplied by zero and displayed with maximum gain so as to show the effect.

The effect can be minimised by reducing the bandwidth of the amplifiers at the expense of image resolution. A second improvement can be achieved by matching the DAC output capacitance with a trimmer across the feedback resistor which minimises overshoot at sharp edges (i.e. spurious ringing). It might be possible to improve the situation further by using amplifiers with much shorter propagation delays, but the problem will still be present and will reemerge in future when it is required to work with higher resolution images. The best solution is to eliminate the problem altogether, and this can be achieved by using a 2- quadrant multiplying circuit in place of the 4- quadrant circuit. Such a circuit is shown in figure 6.7 and it uses only one of the current outputs with a single fast FET amplifier. The propagation delay problem is now eliminated, but as the name suggests, a positive and negative input video signal can now only be multiplied by a positive digital number (rather than a positive or negative number). Thus the circuit is now a digital attenuator with 8 bit resolution, and the ninth bit has been discarded. In order to restore full 4- quadrant capability, a phase splitter circuit is incorporated following the DAC

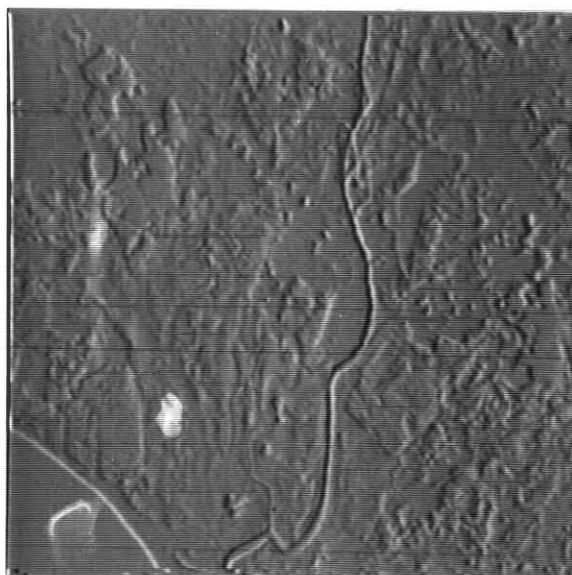
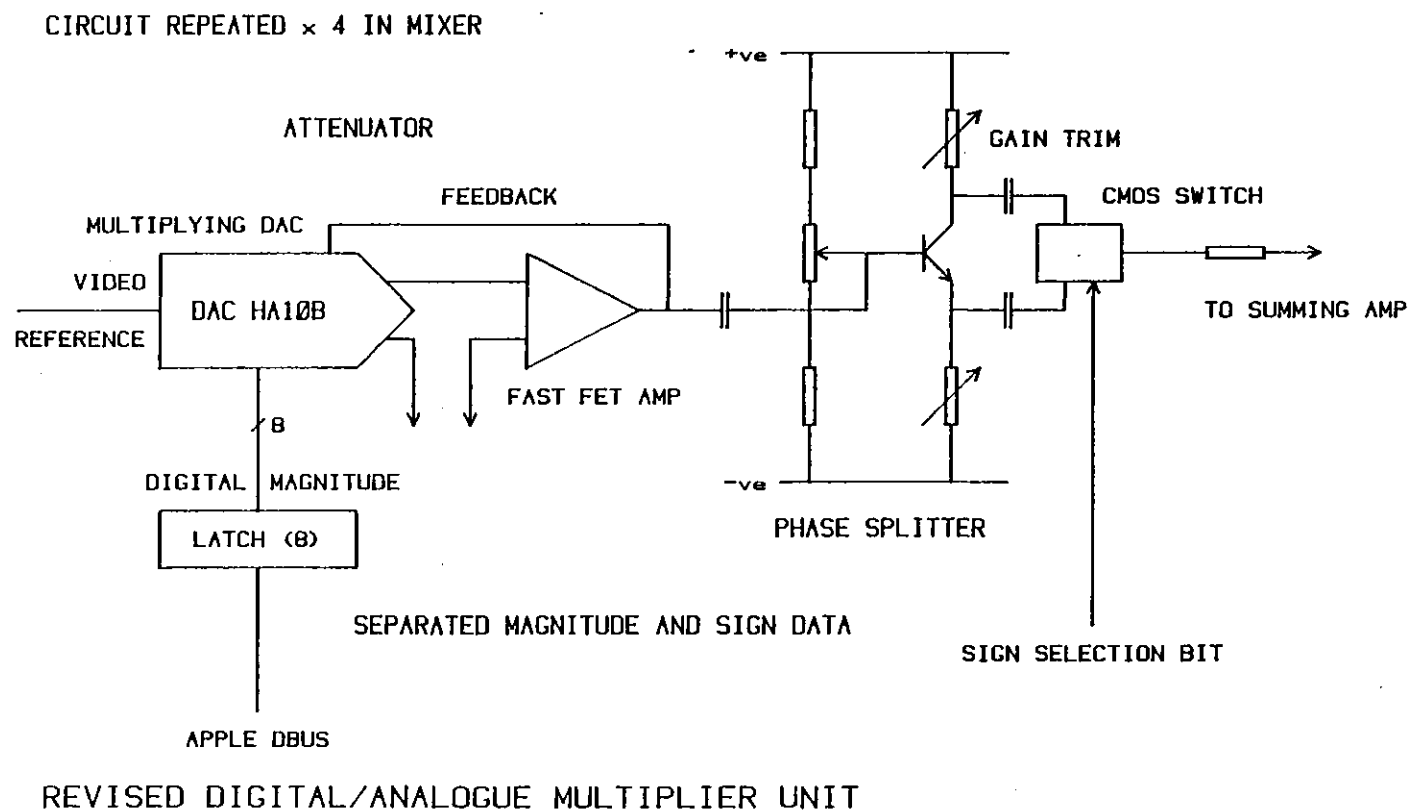


Figure 6.6 : A 'zero contrast' image from the first mixer

Figure 6.7 : The improved mark 2 mixer unit



amplifier. This could consist of a pair of matched OP AMPs, one inverting and the other not, but there is no guarantee that their propagation delays would be equal. A far simpler solution is a single transistor with equal emitter and collector resistors. As the input signal causes a change in collector current, the collector and emitter voltages change in exact antiphase with the same amplitudes (except for the small additional base current which flows in the emitter resistor). A CMOS digitally controlled analogue switch can then be used to select between the positive and negative signals before the final summing amplifier, and the missing ninth bit can be used as the sign control signal. In practice, the finite input impedance of the summing amplifier can cause the collector signal to be reduced when the CMOS switch selects it. This problem does not arise in the emitter signal because of the lower output impedance, but it is easily overcome by varying the collector resistor. The trimming capability also allows the four mixer channels to be matched precisely.

The circuit shown gives DC offsets to the two signals because of the base-emitter and base-collector quiescent voltages, but these are removed by capacitive coupling of the A.C. signal. The faster FET amplifier gives improved bandwidth, and the circuit should be more efficient. One problem which potentially exists in both the old and the new circuits is that of matching the propagation delays through the four DAC attenuator / multiplier circuits so that no phase error exists at the final summing amplifier. Experience with the old circuit indicates that this is not so great a problem as the differential amplifier.

In this section the various shortcomings of the mixer unit have been described, along with the possible solution which has now been implemented. If the input image is multiplied by a digital zero as before, the result is an almost blank screen even under maximum gain, and the accuracy is limited only by the feedthrough capacitance. In addition, the faster amplifiers provide improved bandwidth of over 7MHz, and overshoots can be eliminated at edges by means of the trimmer capacitors in the feedback circuit. It should be remembered that the method itself was not at fault, rather that the skill and technology was found wanting. The problem could be circumvented entirely by computing the PC transformations in software and storing them in the framestores, but this would take very much longer and would be most inconvenient. The achievement of a fast real time PC viewer is a worthy goal, but as in all such cases, it requires great care and attention to perfect. The

following section will show the computation of a 4-dimensional PCA using the improved mixer circuit along with a simple attempt at the interpretation of the results.

6.4 A four dimensional Landsat PCA

Images obtained from the Landsat satellite system often exhibit inter-band correlations, and many workers have made use of the PCA in an attempt to reduce or remove this redundancy of information. The eventual aim may be to reduce the dimensionality of the data by compressing the information into a few bands (thus reducing storage requirements) or it may be to improve classification results by making better use of feature space. The post processing and interpretation of PC images will be covered in this and the following sections using the system described above with a full four dimensional data set. All PCA computations to date have been software-based, and have required large computers. The advent of a 7-dimensional Landsat system and multichannel airborne scanners is placing an increasing strain on such processors, and a rapid hybrid approach is fast becoming more attractive.

The image chosen for this brief study is taken from a Landsat frame covering the south west part of Spain, and was acquired in September 1973. Spanish images tend not to show poor contrast arising from haze or cloud, and the generally arid ground gives clear indications of cover types. The four spectral channels comprising the image appear in figure 6.8 a-d and a sketch map is shown in figure 6.9. The area is fairly low-lying and encloses the old age section of a large river (i.e. a marshy area with many wandering water channels and mud flats). Some of the features are detailed in the key attached to the map, and it is interesting to compare the reflectances of the different cover types in each of the four bands. In a simple study such as this, detailed analysis of the structure of feature space is both difficult and inappropriate, and visual analysis of the differences between bands will be used instead.

The first major feature of note is the vegetated area at the top center of the scene (key No. 1). This is an area of rice cultivation which takes advantage of the water supply and the marshy terrain. The rice shows up dark in bands 4 and 5 (the green and red spectral bands) because of the spectral properties of chlorophyll which absorbs visible radiation. Conversely the infra

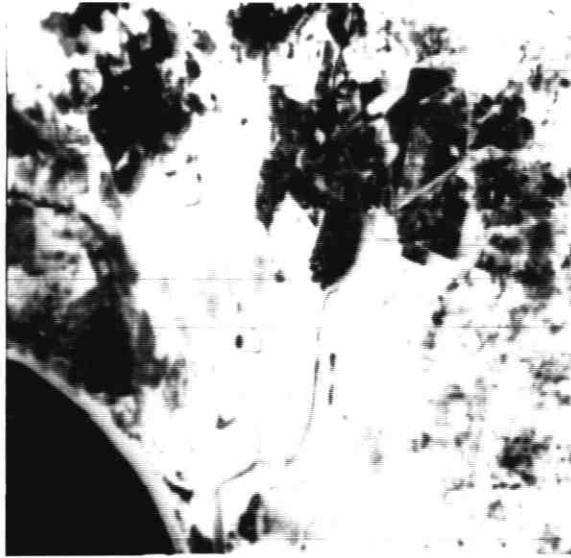


Figure 6.8 a : original band 4 image (green)

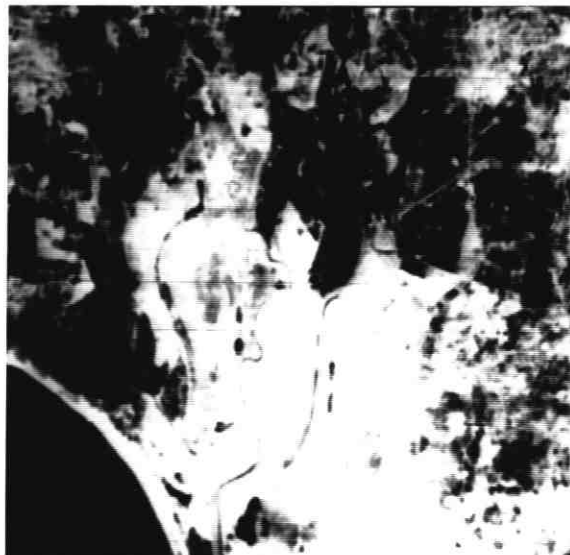


Figure 6.8 b : original band 5 image (red)

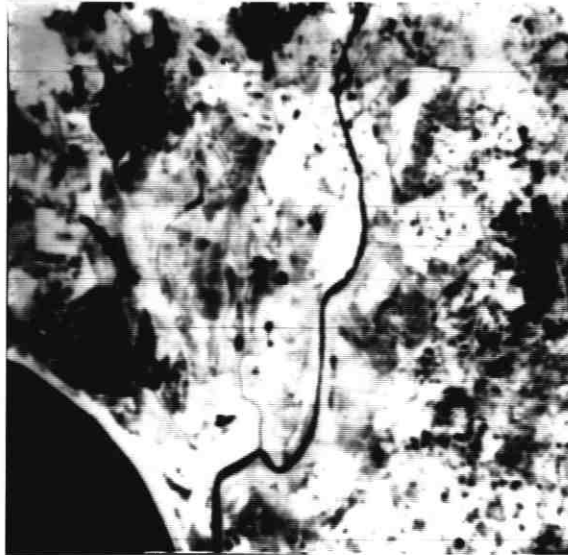


Figure 6.8 c : original band 6 image (near IR)

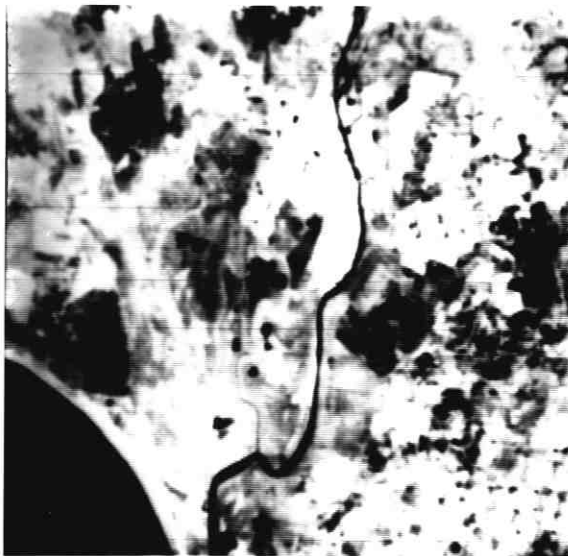


Figure 6.8 d : original band 7 image (near IR)

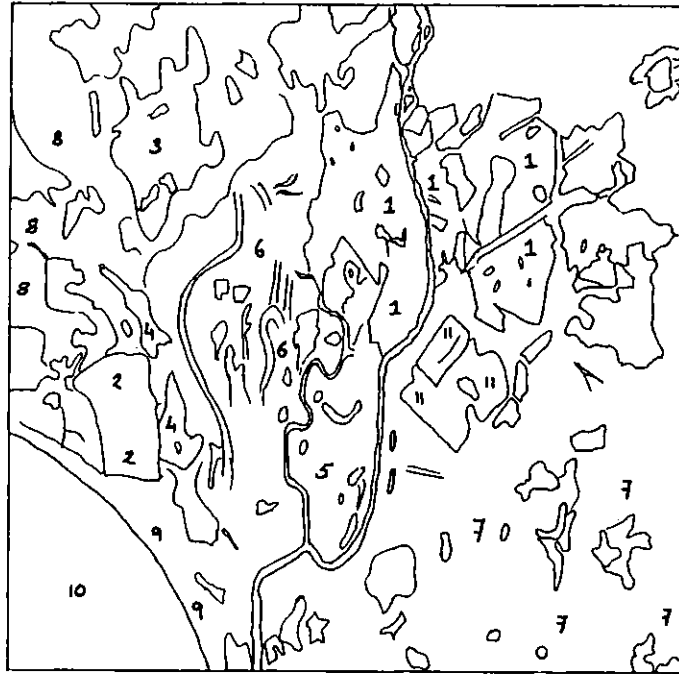


Figure 6.9 : Key to sketch map of Spain image subscene.

- 1) Area of rice cultivation
- 2) Pine woods (umbrella pine near sea)
- 3) Pine woods (fir trees inland)
- 4) Scrub for ast, reeds and willows at edge of a river channel
- 5) Central area of marsh with small lakelets
- 6) Upper part of marsh showing old river channels
- 7) Rising ground, some olives, vines, small farms and open country
- 8) Olive groves, vines and some orange trees, farm land
- 9) Coastal strip, sand dunes and sparse vegetation
- 10) Atlantic ocean
- 11) Area under drainage for cultivation

red images (bands 6 and 7) show a high reflectance for the vegetation. In fact the two visible bands tend to be well correlated as do the two infra red bands, and this is evident from their appearance. Not all vegetation has the same spectral signature as a relatively uniform field of rice however, as for example the wooded area at the lower left of the scene (key 2, 3). This is mainly a very dry type of pine forrest on sandy soils, and both the darker colour of the leaves and the texture of the trees combine to reduce the reflectivity. Close to the area of woods marked No. 2 on the map is the triangular region noted in section 6.2 (key No. 4). This area shows variable reflectivity across the bands and lies on the edge of the main marsh (key No. 5). The ground cover is mainly tall reeds and scrub with smaller willow-type trees. Within the upper main marsh (key No. 6) the various channels of the river are visible among the reed beds and (dried) mud flats, and the dark spots are small lakes of variable size and depth. The river itself is most clearly visible on the IR bands since the visible bands tend to show up silt in suspension and the presence of sandbanks or shallows.

The area to the lower right (key No. 7) is made up of rising ground (the marsh does not rise above 10m) which has scattered olive groves and vines planted on it. On the left (key No. 8) is another area of olives and vines between the pine woods, and near a small tributary river lies a diagonal belt of scrub visible most clearly on band 6. The coastal region (key No. 9) is edged by dunes rising to about 20m, and the large dark region (key No. 10) is the Atlantic ocean. At the river outflow just off the bottom of the image a pattern of silt is visible in the sea on the band 4 image. A final area of interest is situated just below the main vegetation area (key No. 11). This region is under drainage / irrigation by small canals and pipes in a rectangular grid and it is backed by more cultivated ground (olives etc.). Evidence of habitation can be seen in the pattern of roads in the rice area but there are no major towns.

After the visual appraisal of the image, aided by the use of maps, the next step is to collect the variance / covariance statistics. As can be seen from the raw images, the overall contrast is high, and this has been achieved by the application of a linear contrast stretch to the video signals from the FSS before the framestores were loaded. The relative clarity of the images implies that the use of feature space is quite wide, particularly when a plot of bands 4 or 5 against bands 6 or 7 is viewed, and this is confirmed in practice by viewing the appropriate displays.

TABLE 6.2 : Four dimensional statistics

Mean band 4 = 129.29 (range 0-255)

Mean band 5 = 106.63

Mean band 6 = 104.71

Mean band 7 = 121.80

Covariance matrix :

$$\begin{array}{cccc}
 \sigma_{11}^2 = 5192.16 & \sigma_{12}^2 = 3866.65 & \sigma_{13}^2 = 2722.83 & \sigma_{14}^2 = 1094.98 \\
 & \sigma_{22}^2 = 3781.67 & \sigma_{23}^2 = 2520.50 & \sigma_{24}^2 = 1462.00 \\
 & & \sigma_{33}^2 = 4806.52 & \sigma_{34}^2 = 3652.66 \\
 & & & \sigma_{44}^2 = 3927.78
 \end{array}$$

Sum of variances = 17708.13

Eigenvalues :

$$\begin{array}{ll}
 \lambda_1 = 12252.33 \\
 \lambda_2 = 4464.40 \\
 \lambda_3 = 639.82 \\
 \lambda_4 = 351.58 & \text{Sum of eigenvalues} = 17708.13
 \end{array}$$

Eigenvectors :

$$\begin{array}{l}
 \mathbf{e}_1 = (+0.542, +0.483, +0.559, +0.401) \\
 \mathbf{e}_2 = (-0.563, -0.356, +0.403, +0.627) \\
 \mathbf{e}_3 = (-0.370, +0.682, -0.504, +0.379) \\
 \mathbf{e}_4 = (+0.503, -0.418, -0.521, +0.549)
 \end{array}$$

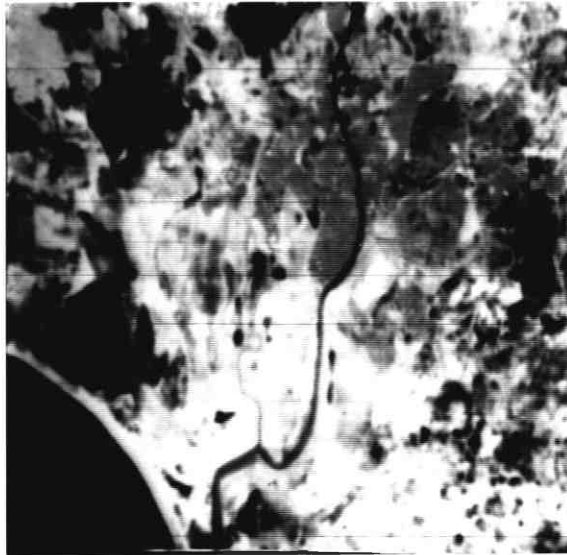


Figure 6.10 a : First Principal Component image

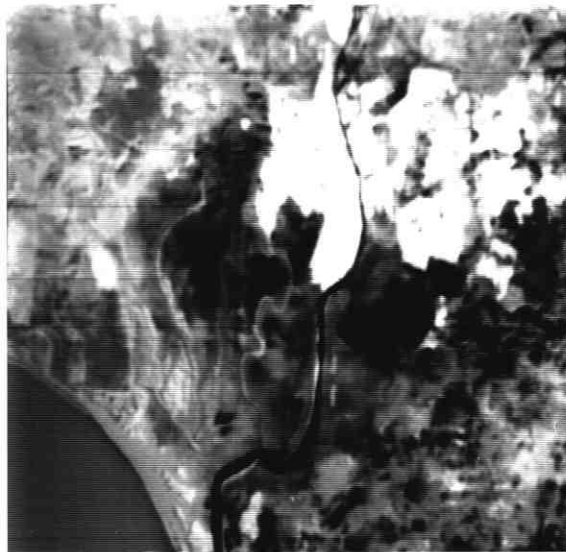


Figure 6.10 b : Second Principal Component image

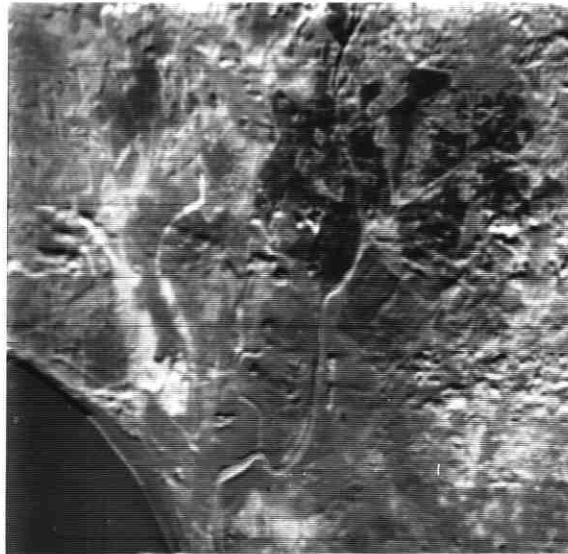


Figure 6.10 c : Third Principal Component image

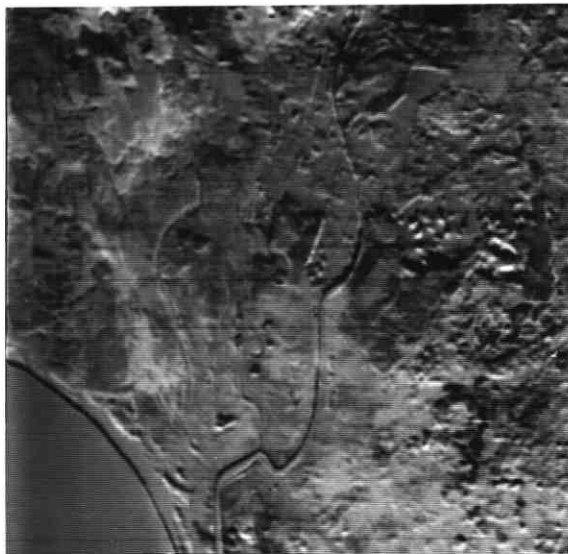


Figure 6.10 d : Fourth Principal Component image

The raw covariance matrix confirms the impression that the visible bands are not as well correlated to the infrared bands as they are to each other since the covariance terms are smaller (e.g. $\sigma_{14}^2 = 1094.98$ for bands 4 and 7 as compared to $\sigma_{12}^2 = 3866.65$ for bands 4 and 5). All the terms on the leading diagonal of the matrix are the 'autocovariances' of the raw pictures, and their sum can be related loosely to the amount of information in the four bands. As in the previous example, the sum of the variances of the principal component images has the same value, indicating no net loss of information. When the matrix is subjected to the principal components analysis and is diagonalised, the available variance is redistributed to the values shown (i.e. the four eigenvalues). It is clear that the first two principal components have been assigned most of the variance and thus most of the contrast or 'information'. The exact percentages of the total variance are 69.19%, 25.21%, 3.61%, and 1.99% respectively into PC channels 1, 2, 3 and 4, and the contrast of the PC images will be proportional to the square root of the variances.

In order to view the PC images the mixer unit is programmed with the coefficients contained in the eigenvectors, each component multiplying the appropriate raw video signal. All the vectors have unit length and are mutually orthogonal in the four dimensional feature space (all dot products are 0). The exact values of these components will depend on the original images and their processing prior to the analysis but some general observations can be made with regard to the signs. PC1 is obviously a weighted sum of all four bands, and as such it corresponds to a panchromatic view of the scene. Indeed, this image could be seen by viewing the scene through a suitably manufactured spectral filter. PC2 is somewhat different in that it has some negative weights which cannot be realised physically so easily, but it will clearly show up differences between the visible and infrared parts of the spectrum. PC3 and PC4 are more complex mixtures and are much more difficult to interpret, but the pattern of their signs is similar to that found in many other Landsat PC images.

The resulting images appear in figure 6.10 a-d, but they are not strictly true representations of the eigenvectors given above since the lower PC images have been contrast enhanced for better visibility. In order to gain a better feel for the process it is useful to note the appearance of some of the features noted above in the PC images. The clearest example is the area of rice vegetation: this appears mid grey in PC1 but white in PC2. Referring back to the original images it can be seen that this area was dark in the visible and

light in the infrared, and the weightings have tended to enhance the difference in PC2. PC1 shows similarities with all the original images as expected (e.g. the pine woods are still dark and the marsh is still light), and it can be said to reflect the broad spectral features of the scene. PC2, reflecting visible / IR differences, shows much more detail in the marsh area, including some evidence of roadways in the pine forrest (including firebreaks). The upper part of the marsh appears darker near the rice, and this corresponds to a drier part where water is not always present (particularly at the end of summer). Note that both of these images are of better quality than those produced with the old mixer unit and that the horizontal 'relief' effect is not present.

PC3 is much more complex to interpret, but some features do stand out, even though it is more noisy because of the large amplification which has been applied. The area noted above as being reeds and willows adjoining the lower pine woods is brought out as being light. In the original images this area shows some minor differences from band to band, being darkest in bands 5 and 6. The 3rd PC image has picked out this difference in vegetation which is associated with the drainage of the area. Another interesting detail which has emerged is the difference between the two sets of pine woods (lower and upper): the upper woods, being inland and not so subject to salt and sea winds, is slightly different and has come out darker. This fact is not easily visible on the originals, even allowing for the photographic reproduction. PC4, sometimes called a mathematical 'nonsuch' shows some interesting detail in the drained area just below the rice in which the rectangular structure can be seen. In general it seems to bring out the texture of the scene, as does PC3 to a lesser extent, and some good examples of this are to be found near the left hand edge of the image. There is some hint of a relief effect in a vertical direction (e.g. on the edges of lakes), and this may arise from several sources. Optimistically, and particularly around the river, it may be due to differences in the original images such as silt banks in the river, but realistically it is probably due to some slight misregistration of the original images. The effect is not constant across the image, and this may indicate a distortion in the original transparency introduced during its manufacture. To conclude this section, it must be noted that there is no clear relation between the contents of the original images and those of the PC images. The process is very abstract, and any attempt at interpretation must be empirical in nature, and based upon a feel for the structure of feature space. Without a full and lengthy analysis this is very hard to achieve.

6.5 Post processing and interpretation

This section will outline some of the possible processes which can be applied to PC images once they have been computed. The first and most simple process is one of contrast enhancement as has been described in the previous section. The images which were shown above were produced in such a way as to stretch the low contrast channels relative to PC1. They represent a 'maximum contrast' option which may be useful for interpretation, but for a true reflection of the relative information contents of the PC images, this simple approach must be modified. Apart from the method of chapter 5, one alternative is to make the gain applied to each image proportional to the square root of the variance (i.e. the standard deviation), thus giving a truer feel for the redistribution of variance which has occurred. Figure 6.11 a-c shows the effect of such a transformation on PC images 2 to 4. PC1 is not modified as it is taken as the reference image and all variances have been scaled relative to its eigenvalue. The lack of contrast in PC3 and 4 is now obvious, and it can be seen why some workers discard PC4 as being insignificant when data compression is required. In all these images the brightness is arbitrary and adjusted for convenience, and only the gain has been altered (in a linear fashion). It is possible to apply nonlinear stretches to PC images exactly as in chapters 3 and 4, and some useful results can be gained in this way, but the present equipment and software does not possess this capability. The overall effect of such stretches is to open out the feature space distribution by stretching along the major and minor axes independently (see back to section 6.1), and to improve cluster separation.

One very important process which is commonly applied to conventional Landsat and other images is colour composition. A false colour composite image is made by using the intensity levels from three component images to modulate the intensities of three primary colours on a photograph or a colour TV display. In the case of Landsat the 'standard' false colour composite assigns band 4 to blue, band 5 to green and band 7 to red. Obviously only three bands can be used and the choice must attempt to encompass as much of the available information as possible, hence the use of a visible and an IR band. Since band 5 is usually more different to band 4 than band 6 is to band 7, band 6 is omitted. Image processing workers have built up experience of this type of image and can use colour to interpret the scene, although the eye is not always very sensitive to subtle colour differences. As an example, figure 6.12 shows a false colour

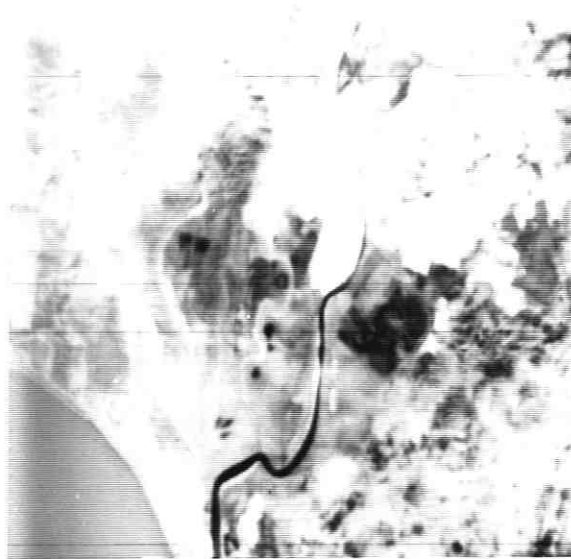


Figure 6.11 a : Proportionately stretched PC2 image



Figure 6.11 b : Proportionately stretched PC3 image

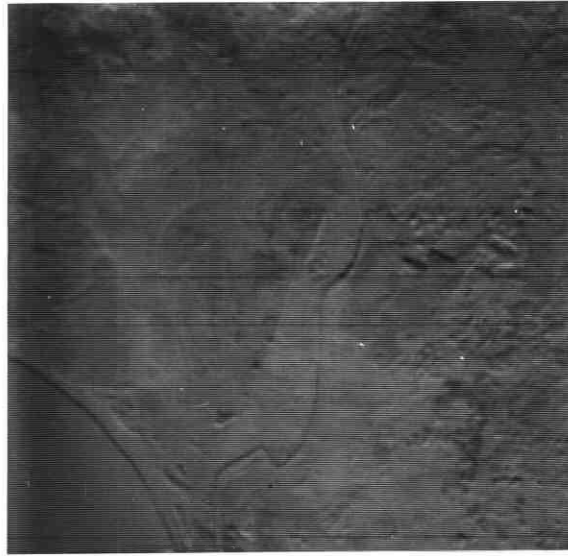


Figure 6.11 c : Proportionately stretched PC4 image

composite image made up as described above and stretched to maximum contrast. The vegetation (rice) is immediately visible as a red patch, underlining the visible-IR difference, and the marsh area is mostly a blue green colour. Light areas denote mostly bare sand or rock with sparse cover or 'burnt' (sun dried) vegetation. The forested areas appear as dark brown or black. Using the variance figures from the last section, this image contains approximately 72.9% of the available image variance in three bands which are partially redundant (i.e. correlated). The next figure (6.13) shows a colour composite made up from the first three PC images in the same way, with PC1 assigned to red, PC2 to green and PC3 to blue. The PC images have not been stretched as in the previous section, and the lower contrast of PC3 shows up as an overall blue green wash or cast to the image, a common occurrence with raw PC composite images. The image does however contain over 98% of the available variance in three uncorrelated components. As an aid to viewing, figure 6.14 shows the same PC composite after linear contrast enhancement (each PC component set to the same contrast), and the improvement is well worthwhile. The rice is now green-blue, and the marshy area has a wider range of colours within it, enabling much finer interpretation. A cursory glance at the PC composite may not reveal the increased information content, but if it is compared with the original several differences become apparent. The difference between the two areas of pine woods noted above can be seen as a colour variation, and there is more variation in the upper part of the marsh area. This arises from the band 6 information which was absent from the original composite image. Also visible is a green border to the lower pine wood area which differentiates the area of marshy scrub which was noted in the previous section much more clearly than the original.

The choice of colours used to represent the three PC images was arbitrary, and it may not be the best option. In this case the marsh area is reddish in colour, and the eye is at its least sensitive to colour differences when viewing red/orange colours. Many other choices are available, and their applicability will depend on the type of image and the aim of the user, but as an example figure 6.15 shows an alternative colour scheme (PC1 = green, PC2 = red, PC3 = blue) which makes the marsh green-blue, where the eye is most sensitive to subtle changes. As a practical note, the colour images presented in this thesis are not produced by photographing a colour monitor but by building up the composite image on colour slide film from three monochrome images using filters. This approach is adopted for several reasons, not the least of which is that a single mixer cannot be used to give three colour signals simultaneously. Other

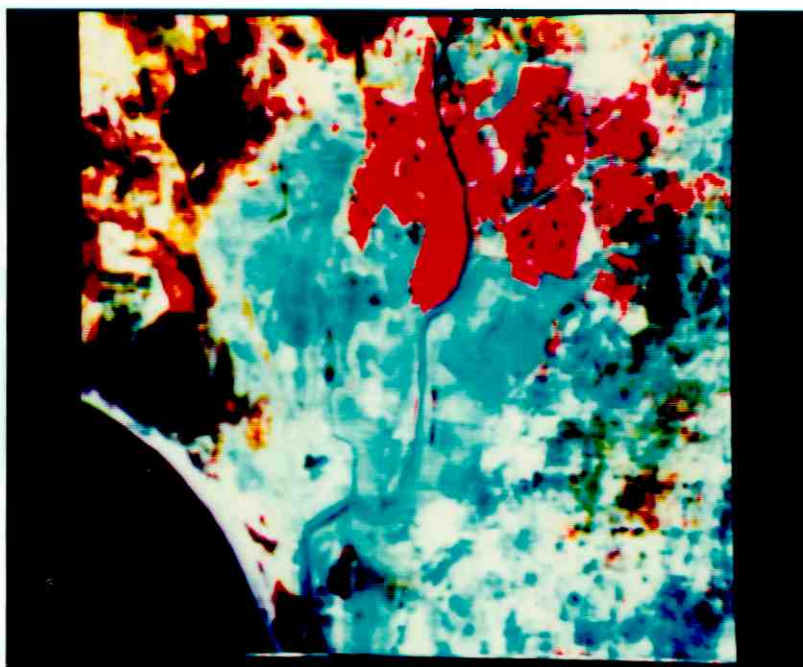


Figure 6.12 : A standard Landsat false colour composite



Figure 6.13 : An unstretched Principal component colour composite

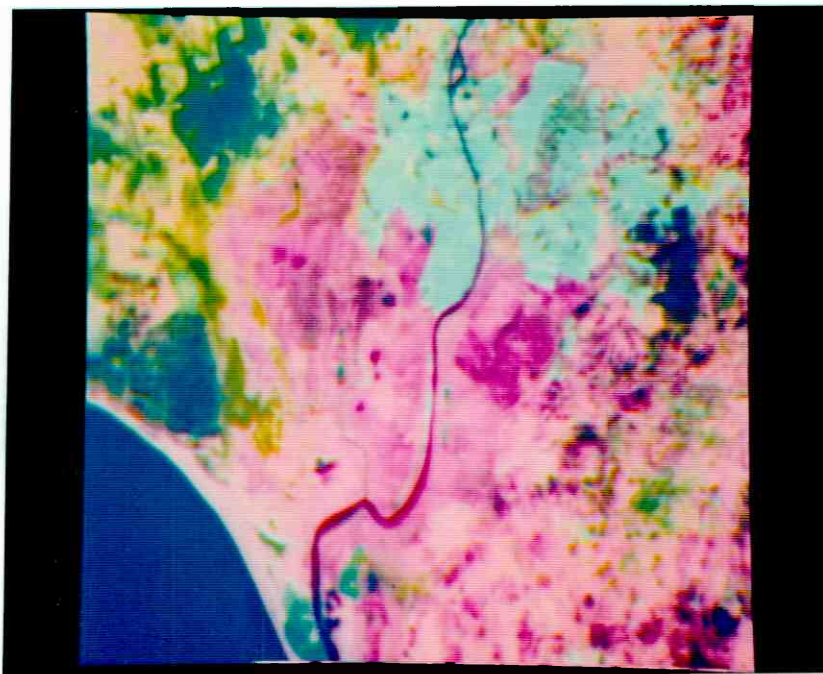


Figure 6.14 : Linearly stretched PC colour composite (equal variances)

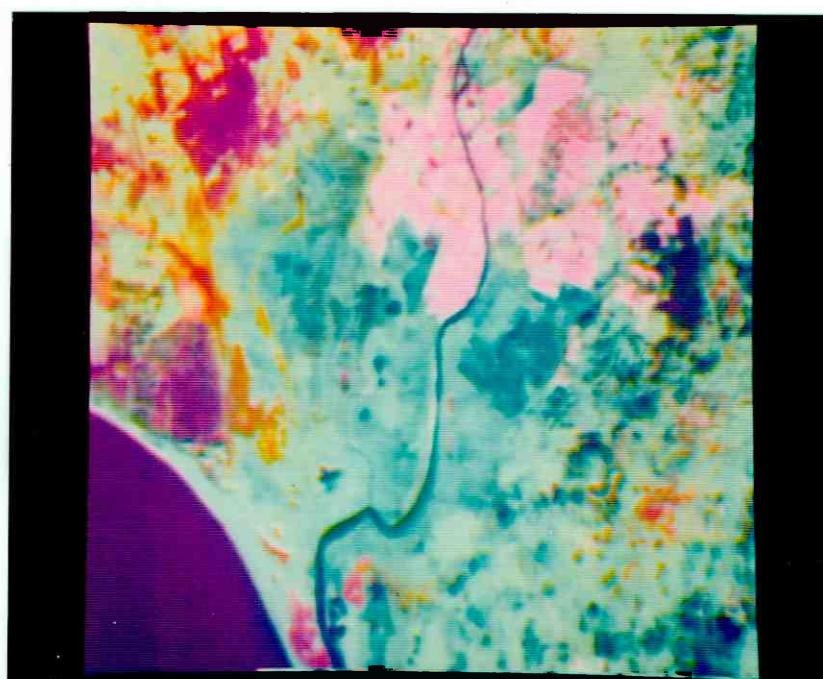


Figure 6.15 : An alternative choice of colours

factors include image degradation which arises from the presence of a shadow mask in a colour TV display tube (this gives a dotted or barred structure to the image which reduces resolution), and colour degradation. Unfaithful colour rendition arises because of the spectral output of the colour phosphors which does not match the spectral sensitivity of the film dyes, but the use of suitable colour filters and exposure times from a white phosphor can give very good results. The computer is used to control the mixer during exposure and it can display any image for a given number of complete fields, thus ensuring uniform results.

There is one major difference between the original colour composite and the PC composite, and this is the lack of white areas in the PC image. In the original, a white area occurs whenever the three component images have similar intensity values. One of the effects of the decorrelation which occurs when the PCA is performed is that the PC channels tend not to have similar intensity values at any point. This leads to previously white areas such as those in the marsh being shown as coloured, and in general it is common to find white (or grey) areas changed to different colours.

From the point of view of the experienced image interpreter, the advent of a PC colour composite is a mixed blessing, since although it embodies much more information, it requires him to learn a new set of colour to ground type relations. If the PC image could be processed so as to give a colour scheme similar to the original yet with the benefits of the PC process preserved, his job would be much easier. The key to this problem lies in the observation that the PCA gives uncorrelated but not independent images which can be stretched until they are virtually independent. If a Gaussian stretch is applied to each of the four PC images so that their histograms resemble Gaussian distributions which fill their available ranges, then the feature space distribution will be essentially spherical, filling almost the entire 4-D volume. This distribution now looks the same wherever the axes are pointed, and it is possible to perform the exact inverse of the original PC rotation without reintroducing any correlation, thus returning to the original colour relationships. The stretch can be made linear if the variances of the four PC channels are made equal before the reverse rotation as this achieves a very similar effect. Since the contrast enhancement has been applied to uncorrelated bands, the result contains much more of the original information in a wider range of colours. To perform this operation requires the availability either of four video mixers to produce

all the PC channels simultaneously plus an extra set of mixers to give the reverse rotation, or access to stored versions of the PC images. At present this facility is not available on the equipment described here, but it will be implemented in the future by using the four framestores to hold PC images computed by the Apple computer or a special purpose hardware board.

All the above sections on PC images have given only sketchy ideas on the interpretation of the data, and although it may be frustrating it is for good reason. The PC process is very abstract and its effects are closely linked to the detailed structure of the feature space under investigation. This structure or shape is in general multidimensional and in turn critically dependent on the exact image scene, acquisition and processing combination which precedes the analysis. It does not appear to be possible to give a precise general interpretation of any results obtained, and the only way in which predictions can be made is with the aid of a feel both for the feature space distributions which are likely to occur and for the overall characteristics of the sensor system. Many workers have attempted to ascribe meanings to PC images, but such meanings must never be divorced from their context. It is possible to give some very rough attempts at interpretation if assumptions about the sensor and scene are built in. If it is assumed that the scene is of low contrast, possibly obscured by haze or mist, and that the major contributing factor to image contrast is the illumination (e.g. the sunlight in a Landsat scene), then PC1 will always be a positive weighted sum of the constituent bands or in other words a panchromatic image. This is because any change in illumination will tend to affect all the spectral bands more or less equally, so that the major axis of the feature space is always located in the first 4-D 'hyper-octant'. An implicit assumption which should be noted in practice is that the bands are narrow enough and closely enough spaced to preclude the influence of some strong spectral feature in the scene. For the reasons given above it is common to find that in a mountainous scene, PC1 will tend to show up differences in the illumination caused by topography whilst the other PC images show up spectral differences. This observation has been used by several workers in attempts to remove topographic effects from such images. PC2 must be orthogonal to PC1 and will tend to show up the most significant spectral differences within the scene. In the case of Landsat this is most often the vegetation pattern because of the visible-IR difference in chlorophyll absorption. PC's 3 and 4, under the constraint of being orthogonal, then contain progressively less significant spectral differences, and PC4 in particular often contains artefacts such as

striping caused by the sensor. So far, all analysis of the PC images has been confined to visual methods, either using monochrome PC images or colour composites. Although the PC composite image has a much wider range of colours than a raw composite, eye limitations and the fact that one PC image has been left out (without much information loss), it may be advantageous to apply classification techniques to the PC images similar to those used by the MEIC described in earlier chapters. The PCA and subsequent contrast enhancements will have opened out feature space and hopefully will allow more accurate classifications.

If the image contains some recognisable areas of known ground cover then it may be possible to separate these cover types by using the technique of Canonical Analysis. The basis for this approach is to restrict the areas of the original image which contribute to the covariance statistics by selective sampling. The effect is to distill out of the raw feature space only those clusters of interest, and then to rotate feature space using a conventional PCA so that these clusters are decorrelated and separated in the final images. Successful application of the process depends upon a good knowledge of the terrain in the scene and the use of statistically valid samples, but it can be very powerful in the preparation of land use maps. The next section will outline one possible method by which Canonical analysis could be implemented on the present system. The previous chapter indicated that the mixer unit could be used interactively by the operator to explore different mixing algorithms using the PCA as a guide. This facility has not been implemented so far but its potential is very great and would require another thesis for its exploitation.

'Wild rumors and legends' abound amongst users of the PCA about the value of the least significant PC images to the effect that they show up the valuable ore deposits with great clarity and should not on any account be discarded, but when pressed, interpreters usually phrase their conclusions with a generous dose of caution. As always the value of any process must depend upon the user, and, in this case particularly, on the scene. The serious user must be very wary of the salesmen who sell the PCA as a 'magic' process which always finds the goldmine, but at the same time it is a powerful aid in the interpretation of complex multidimensional imagery. This chapter has shown two examples of the process together with some of the possible post processing options which are available. The shortcomings of the original mixer design have been described along with one possible solution which has been implemented. Limitations of

space preclude a deeper investigation of the capabilities of this interesting process which will have to be left to future work, but it is hoped that this chapter has given a useful introduction and demonstration of the Principal Components Analysis.

7.1 Introduction

As an aid to the digestion of this work, a brief review of the major areas covered will be presented together with some possible extensions and ideas for further research. For simplicity the contents of the thesis have been divided into three, covering photographic storage and noise, interactive contrast enhancement and fast Principal Components Analysis. Each topic will be covered in a separate section, and it is hoped that some clearer insight into the motivations and design philosophies underlying the research may be imparted.

Before embarking on the main review, a brief distillation of the first, introductory chapter is necessary. The aim of the first chapter is to define the meaning of an image and to convey a feel for the various methods of collecting, storing and processing images. A common response to the increasing flood of data in pictorial form is one of brief curiosity and "taking for granted", and the magnitude of the tasks involved in the acquisition and processing of such images is often not appreciated. As image processors become more powerful and fast it is easy to forget the size of the data manipulation problems involved, and the first chapter serves as a reminder. A further consequence of improvements in processing technology is the advent of a new type of user who regards image processing as an all-powerful method of achieving his aims (albeit through the use of mysterious 'occult' processing algorithms). An appreciation of the characteristics and limitations of imaging systems is vitally important if any validity is to be ascribed to the results of image processing, and chapter 1 seeks to introduce some of the basic ideas by examples such as the Landsat resource monitoring system whose data appear in a later chapter.

The introduction commences with some of the mathematical aspects of images, treating them both as 2-D functions or arrays of numbers and as collections of objects which have certain defined properties or features. Following from this comes the very important concept of feature space and its particular application to the treatment of multispectral images. Such descriptions introduce the statistical nature of images and they allow the effects of processing operations on highly complex scenes to be characterised more easily. Indeed, some

operations such as classification, contrast enhancement and feature space rotation cannot be utilised to their best effect without an appreciation of the statistical nature of images.

The acquisition and storage of large volumes of image data requires the careful use of sensor technology, and almost always some form of scanning to cover the scene of interest. Parallel systems such as film cameras are used in some cases, but they are not well suited to the provision of electrical output signals, the usual raw material of modern image processors. In both cases it is important to note the resolution of the system, but in scanners additional problems are introduced by the sampling process, and quoted pixel sizes "on the ground" must be used with care. Once again, the exact meaning of resolution is application dependent (i.e. clarity is in the eye of the interpreter). The process of sampling gives rise to very high data rates, and the scanner must be designed so as to accommodate this and minimise it as far as possible whilst retaining the desired (high) resolution.

Once the image data has been gathered and stored (photographically or electronically), it can be subjected to processing and interpretation. One of the important contentions of this work is that the principal component of any image processing system must be an experienced operator, with any hardware or software playing a secondary and supportive role. Unfortunately it is very hard to acquire such operators and much easier to rely too heavily on the machine. Among image processing machines there exist several different types, some computer based, some optical and some hardwired electronics (analogue and digital), but even with all the capabilities presented by this large array of options users tend to stick with one type, sometimes ignoring more efficient alternative approaches. This effect is most common with large computer based processors which rely for their considerable power on their speed of number crunching. The enormity of the data content of an image becomes apparent when some operations are attempted (e.g. Fourier transformation requires several minutes for a 512 pixel square image whereas it can be performed instantaneously with an optical system for much larger images). For the above reasons, and also because of cost and convenience, the author has adopted a dedicated hardware approach. The usual criticism leveled at hardware is that it lacks the flexibility of a software driven machine, but the systems designed in this thesis have been made as flexible as possible, with the added advantages of ease of use and high speed which is not always the case with large computers.

7.2 Photographic storage

The first of the three sections covers the characteristics of film as an image storage medium, and in particular it seeks to investigate the noise levels associated with grain. The motivation for this study arose from a practical need to estimate the accuracy of classifications carried out on the MEIC using transparencies as image stores. The basic aim was to determine how many distinct grey levels were available from the film. In attempting an answer, the exact method by which the film is viewed or read must be taken into account and a more precise definition given to the term 'distinct grey level'.

The noise in a photographic record is mainly a consequence of mottle or granularity noise, caused by the clumping together of the small grains of silver which make up the image. This clumping is in turn caused by many effects, some in the manufacturing process and some by the exposure and chemical development stages, but the nett effect is to produce a random noise signal all over the film. The effect of exposure will be to make the average density of the film larger or smaller, thus moving the average of the noise distribution, and as has been shown in chapter 2, there is some effect on the standard deviation of the noise signal.

The approach to the problem is practical, and consists in measuring the distribution of the noise at a set of different average exposure levels throughout the dynamic range of the film and using these to estimate the number of levels available. The first point which had to be checked was the shape of the noise distribution, and it was found to be Gaussian to a good approximation. Using this assumption, it is possible to define the distinct grey level by a probability of error in detecting any particular grey level due to noise from an adjacent level. If a tolerance is specified, such as 5% error, then the Error function can be used to deduce the spacing which must be left between levels in units of standard deviations. The measurements were made using a densitometer, and involved the processing of about 50000 data samples overall. An interesting observation was made concerning the value of the noise standard deviation as it varied with increasing average film density. The results showed a variation (fitted by a fourth order Chebychev polynomial) which gave a higher noise width at medium exposures and less noise at low or high exposures. This result may be specific to the particular film type used for Landsat images (further checking is required), but it appears to be real and has not been included in any

previous estimates. The explanation put forward gives an insight into the exposure process and the interaction of light with the transparency. The model of the emulsion is a random array of randomly sized grains, the larger of which are exposed first due to their increased capture area. A lightly exposed film thus presents a sparse array of large grains which give an increasing noise signal as the exposure increases. At higher exposures, when many of the grains have been exposed, the model can be revised to one in which the emulsion consists of a random array of smaller gaps which are progressively filled in by the smaller grains. This stage mirrors the low exposure region where the noise is less, and leads to a decrease in noise at high densities. If an unexposed film did not show background fog and a fully exposed film was completely opaque, the noise at these extremes would be zero.

The estimated numbers of levels for different errors depends critically on the reading system (in this case the FSS), and particularly on the size of spot used to scan the film. A spot of finite size and intensity profile acts as a filter which may reduce the effective noise signals by reducing the effective system bandwidth. This effect must be taken into account in the final estimates because it causes degradation of resolution as well as noise reduction. In conclusion, the above method is relatively simple yet practical and direct, and it gives useful results. It can be extended to any given system but the results may vary according to the film type used.

The method described above is not the only option available, and it was decided to supplement the study by trying another approach. The second method is not so intuitive or obvious, but it is useful because it relates to digital image storage, a topic which is covered in a later chapter and which is becoming increasingly widespread in image processing systems. In a digital system, each element of the image (pixel) is represented as a binary number with a given precision, commonly 8 binary digits or bits. The number of discrete intensity levels which can be represented in a binary word is 2 to the power of the number of bits (in this case $2^8 = 256$). If all but the top or most significant bit are discarded, only two levels are available, and the array of top bits can be used to form an image. The image is either black or white depending on whether the original image intensity was below or above half of the available range at any point. The image thus formed represents a binary slice or contour, and is the most primitive approximation to the original possible. The same operation can be performed with each of the bits in turn, and the images represent

finer corrections which, if correctly weighted and added to all previous bits, improve the approximation of each number to the original intensity values. As less significant bits are taken, the resulting image becomes more chaotic and random, bearing less resemblance to the original, until eventually only random fields of black and white pixels remain. In a perfect, noise free signal, these random fields of pixels would be significant no matter how many bits had been used in the digitisation, but in the presence of noise they can be discarded. The criterion by which any of these bitplane images can be judged is the amount of structure contained within it, and this can be characterised by examining the autocorrelation function. A random image gives an almost flat autocorrelation except for a triangular peak two pixels wide at the centre, and if there is any structure this peak will be wider and the surround will not be so flat.

Digital image data was acquired from a densitometer using the same transparency type as above and a wider scanning aperture. The results indicate that up to 6 bits are significant (thus giving about 64 levels), and that the extra two bits could be discarded without significant loss of fidelity. To be safe, it would be prudent to use 7 bits, and in many cases it is convenient to use 8 bits because of the internal architectures of computers. The aperture used is larger than that employed previously, but when the filtering effects in both cases are taken into account, the derived numbers of grey levels are in good agreement. The autocorrelation method can be generalised to any system where quantisation occurs or there is noise present, and it does not require the explicit specification of an error tolerance. The accuracy of the method depends on the definition of a random field of pixels, and it may be possible in future to derive some parameter from the form of the autocorrelation function which can be tested for randomness. The use of autocorrelations is very similar to the operation carried out by the visual system when viewing a random field of pixels and searching for structure, and even very faint indications of structure can be picked up visually.

7.3 Nonlinear contrast enhancement

The second major section of the thesis describes a system which can perform nonlinear contrast modifications in real time on standard video signals and which can provide the user with detailed information about the image statistics. The aim of the system is to give a powerful, flexible, interactive and easy to

use tool for the operator and to act as a stand-alone processor or as a preprocessor for raw data ready for another system. Existing systems such as the MEIC can be used to perform linear contrast stretches (i.e. gain and lift adjustment), but this is not always sufficient to give good results. If a nonlinear stretch can be applied, the possibility of selective enhancements of certain features becomes available, along with an almost endless variety of nonlinear transfer functions. The only problems then facing the user are specifying the transfer function required and gaining a feel for the effects of different contrast stretches.

The main part of the system deals with digital data derived from a standard video signal by means of a converter. The signal source can be a TV camera, a FSS or any other source as long as suitable synchronising signals can be obtained. The sampling rate determines the resolution of the image horizontally and this has been chosen to be 10.25 MHz, giving a maximum resolvable video frequency of 5.125 MHz (by the Nyquist criterion), and this is comparable with broadcast TV standards. The contrast stretch is characterised by a transfer function which describes the relationship between the input grey levels and the output grey levels. This transfer function is calculated once for all possible input intensities and then stored in a memory which serves as a look-up table. The transformation then occurs very fast and continuously as each input intensity value is used to 'look up' its corresponding output value within 55 nS. If the same process were to be attempted directly by the system micro-processor, the total bandwidth would be limited to about 25 kHz at best, a factor of 400 slower! The idea of a look up table has been applied before, and it is becoming increasingly common in larger image processing systems where it is used to give false colour to monochrome images. Such systems do not offer the same flexibility or ease of access as the CONDOR system, and they are usually very difficult to 'get at' because of layers of protective software.

The operator presented with a new system such as CONDOR would soon become lost without some method of assessing the effects of his transformations and deciding upon appropriate actions. The key to successful processing is knowledge of the intensity statistics of the image. In the CONDOR system, access to the intensity histogram is provided by a very fast acquisition unit as described in chapter 3. This hardwired board, together with some fast controlling software, gives CONDOR a great advantage over a software driven system with the image stored in a framestore. The reason is that the process of

reading out every pixel's intensity value and accumulating the histogram in core memory is relatively slow, even with large computers. The speed is further reduced in most cases because access to the framestore has to be shared with the display circuitry, effectively restricting the time available to the vertical blanking interval. The alternative is a corrupted display, and this is usually avoided for cosmetic reasons.

Perhaps the most powerful aspect of the CONDOR system is the facility for the specification of generalised nonlinear transfer functions using an interactive switch box. When using this mode, the operator can rearrange the contrast of an image in a very complex manner and the results are available instantaneously as he operates the switches. Accompanying the greater flexibility comes the inevitable penalty of greater complexity and difficulty in interpreting the effects of the process. The operator is aided by computer graphic displays of the transfer function and the image histogram, and the results of chapter 4 show that some useful effects can be achieved. This aspect of the system is worthy of further expansion, and both new methods of user input and new operating software are proposed.

Several opportunities exist for the future expansion of the entire system, the first of which is the construction of a video synchronising circuit which would allow the CONDOR to be locked to any external source of video signals instead of a specially modified camera. Once this has been achieved, the range of image sources is increased dramatically, and the system would be much more portable. A second stage of expansion is the extension of the system to colour operation by replicating the transforming circuit three (or four for Landsat MSS data) times, one for each primary colour. In this way a source of RGB colour signals such as the FSS could be used, and enhanced colour composites could be produced in real time. A third stage could consist of improvements to the operating software, especially making better use of computer graphics and input devices such as tracker balls, joysticks and graphics tablets. One further useful addition would be a digital plotter for the provision of hard copy of histograms for reports etc. A final option might be the addition of either a higher performance microprocessor such as the 6809 or some form of fast arithmetic processor to further enhance the software speed.

This section has outlined the basic operation of the CONDOR system and some of the extensions proposed for the future. It is the belief of the author that

a small, interactive and flexible system such as this could find many uses both in industry and in research, and that the approach that it represents is still valid even in the face of increasing computer power. To the author's knowledge, there are no such systems commercially available at present, and one has to look to very expensive software driven machines before even part of the same capability becomes available. A commercial version of CONDOR could be made for a few thousand pounds or perhaps less by mass producing printed circuit boards, and given the large number of small microcomputers available, it could be adapted to run with many different controllers.

7.4 The PCA system

The third section of the thesis is concerned with the design and operation of a second hardwired system, used for the Principal Components Analysis. Similar design philosophies to those used in the CONDOR system have been adopted to give speed and flexibility, but the processing is accomplished by a hybrid analogue-digital purpose built computer instead of purely digital logic.

The PCA is of use in efforts to improve classification accuracies and visual interpretability when dealing with multispectral data sets. Various other techniques are available such as ratioing (dividing one image by another to enhance small differences), but the PCA is the only one which is based on an analysis of the image statistics. The major purpose or effect is to eliminate correlation between the spectral images by redistributing the image information, and a secondary effect is one of data compression. Chapter 5 describes the process as it is carried out on large software based systems, and introduces a description which involves the effect on the corresponding feature space. This method of characterising the PCA is very important for the development of a feel for the process, which would otherwise remain somewhat abstract in nature.

The same advantages of hardware over software obtain in the case of the PCA system as in CONDOR except that because the volume of statistical data to be processed is much larger in this case the advantages are magnified considerably. The process can be divided into three main sections comprising a statistical analysis, some mathematical operations on the statistical data and a mixing operation. The first and last of these stages are very intensive in their use of peripheral access in a computer system, especially framestore access which is

relatively slow. In order to collect the statistics (in the form of the matrix of covariances), it is necessary to access each pixel in the image and accumulate both the sum of the pixel values and the sum of the interband products of these values. These parameters are required in the calculation of the covariance matrix elements, and they involve large additions and multiplications (i.e. involving long binary words). Apart from the delays involved in reading the pixel values into the processor, arithmetic operations of this type are also relatively slow and cumbersome unless a very expensive arithmetic unit is fitted to the computer. The system described in this thesis avoids these problems by making use of a hardwired arithmetic unit that has as its input the video data which is being continuously read out for display. The entire set of data can be acquired within about one third of a second, and the multiplication and addition operations are performed by special purpose logic at the rate of seven million products (8 bit by 8 bit) and additions (16 bit plus 16 bit) per second, thus equalling a very large computer working directly from internal memory (i.e. without framestore access). The images used are of medium resolution (256 pixels square), but the hardware design could cope with 512 pixels square with no difficulty, and even with higher resolutions at the cost of a doubling of processing time (i.e. alternate pixels could be processed in two passes).

The second step is one of mathematical operations on the covariance matrix, and this is at present the weakest part of the system since it is coded in floating point BASIC on the system microprocessor. This operating program is only a first attempt, and the relatively slow operation (5 to 10 seconds) can be improved upon in several ways. One option would be to code the software in assembly language (machine code), and an immediate increase of perhaps a factor of 10 to 100 would be gained assuming an efficient piece of code. A further improvement could be gained by using a special purpose arithmetic processor chip to speed up floating point operations, and possibly a more efficient microprocessor. It should be remembered that since this system carries out both the first and the last stages almost instantaneously, it is still several times faster than any software driven system available, some of which can take several minutes, and that the process is quite conveniently interactive as it stands (although a two second operation would be well worthwhile!).

The final stage is one of mixing according to the algorithm given by the eigenvectors of the covariance matrix. Once again, a software driven system has

to read in four pixel values, multiply them by suitable constants (the eigenvectors), add and scale them before writing the result into a second image store. This entire cycle must be repeated for each pixel to gain one PC image, and then repeated for all the other PC images. The system developed here accomplishes the same effect using a digitally controlled analogue computer in the form of a video mixer unit. The pixel information is operated upon in the form of an analogue voltage signal (i.e. a video signal) which is derived from the framestores continuously by a converter (this signal is the same one that is used to display the framestore contents on a video monitor). The weighting is carried out by a digitally controlled attenuator and the four signals (in this case) are added using a wideband operational adder circuit before being output to a monitor for display. When the attenuators have been programmed (within a few microseconds), the PC image is available instantaneously and continuously.

Criticisms have in the past been leveled at the use of analogue circuits on the grounds of inaccuracy, noise and instability, but such problems can now largely be overcome using modern integrated circuits. The defence of the method lies in arguments concerning the speed of the system and the great convenience of the analogue mixer. The image quality from the present system may not be the best ever produced, but this is the fault of the non-optimal design and lack of experience on the part of the author, and it could be improved upon by a good designer. A second major factor is the low cost of the system, as both the mixer unit and the arithmetic unit cost only about a hundred pounds and the framestores were constructed from readily available microcomputer memories costing 75p each (total cost about 120 pounds).

Despite the obvious advantages of using the hybrid approach, attention is being given to the problem of designing an equivalent digital unit to perform the same functions using very fast logic circuits. It should be noted that the hybrid approach allows the system to be expanded to any number of channels with very little modification. The only changes are a slightly longer statistics acquisition time (about 6/50 second extra for five channels and 14/50 second for 6 channels), a longer mathematical stage and the extension of the mixer to include some extra digital attenuators. If more than one PC image is required simultaneously, then more mixer units will be required, and if mass produced on printed circuit boards this is simple to achieve. Conversely, the four or more channel addition is quite hard to achieve digitally since most adders are two input devices only, and it may be necessary to form the PC images in a

framestore by adding weighted pairs of images in several passes, but this would still leave a scaling problem to prevent overflow. If a four input adder could be designed with a very fast response time, this problem could be removed. A third aspect is the speed and flexibility of the mixer section which can be controlled interactively and which could cope with higher bandwidth images.

Future areas where expansion could be attempted include the development of larger framestores (up to 512 by 512 pixels) and up to 7 channels of input data to accommodate the new Thematic Mapper data from Landsat 4. A further possibility is to incorporate some of the ideas of the CONDOR system, for example a histogram collection unit and a contrast enhancer for each framestore to allow the raw data to be improved prior to the PCA. If a digital mixer can be designed, then a second contrast enhancer unit to stretch the resulting PC images would also be useful, and it would allow the inverse rotation described in the previous chapter to be implemented. As an aid to the achievement of Canonical Analysis with selective sampling, a single bit framestore could be constructed which would contain a black and white image of the areas to be sampled. The sample area can be entered interactively by the user by means of a light pen or similar input device and the input data would be displayed superimposed on the raw images. The contents of this store would then gate data through to the arithmetic circuit, thus achieving the selective sampling in a very convenient manner.

7.5 Conclusions

The work covered in this thesis is of necessity somewhat shallower than the ideal mainly because of time restrictions, but it has covered some interesting aspects of statistically-based image processing. A longer study would be able to investigate in depth some of the interesting side-avenues (such as photographic grain), most of which merit a complete thesis to themselves. Useful results about photographic storage have been derived and a much better feel for the quality of a transparency image has been gained. Two image processing systems have been developed, for contrast enhancement and the PCA, and both have been made to operate correctly with useful results. The processes achieved by these systems are not new, but the methods of implementation are novel, as is some of the design philosophy. These systems could form the basis of useful commercial devices, and they fill a gap at the low cost end of the

image processor market. This does not mean that they are limited in their capabilities- in fact the opposite could be said to be true. The CONDOR system allows very fast contrast enhancements to be performed with ease, and the full potential of the generalised non-linear transfer function (specified by switches) has yet to be realised. The PCA system is even more remarkable in that there is no cheap or fast machine capable of this operation available at present (to the author's knowledge) since it is 'traditionally' implemented on large computer driven systems.

Several areas for possible improvements and extensions have been put forward, and it is hoped that they may be followed up soon since the machines are only research prototypes and they would benefit from a tidying up operation. An important lesson which has been learned by the author in a field where cosmetically pretty images abound is always to question the significance of the results being presented and to understand the exact means by which they have been achieved. Although the work contained in this thesis has been laced with many caveats, descriptions of hardware design and mathematical processes, it is the hope of the author that the reader has not lost his interest in image processing nor his appreciation of the many beautiful images which form the raw material of the field.

List of References

- [01] Harris, R. H. (ed.) "Secrets of the Microscope", Hamlyn
- [02] Dean, K. G. and Spencer, J. P. "Evaluation of the photographic enhancement of Landsat Imagery", J. Remote Sensing of the Environment 12, 5, 1982
- [03] Rosenfeld, A. "Picture Processing by Computer", A.P. 1969
- [04] As [03], chapter 2
- [05] Pratt, "Digital Image Processing"
- [06] Gonzalez, R. and Wintz, P. "Digital Image Processing", p16
- [07] Born, M. and Wolf E. "Principles of Optics", Pergamon 1959
- [08] R.C.A inc. Electronic Imaging Devices data sheets
- [09] Smol, G., Hamer, M. and Hills, M. "Telecommunications - A Systems Approach", p161
- [10] As [09], p168
- [11] Townshend, J. R. G. "The Spatial Resolving Power of Earth Resources Satellites: a review", NASA Technical Memo 82020, Sept. 1980
- [12] Goldberg, I. and McCulloch, A. "Annular Aperture Diffracted Energy Distributions", A.O. vol 8, 7, July 1979
- [13] Otterman, J. "Diffraction Limited Resolution for Geoscene Imagery", A.O. vol 8, 9, Sept. 1969
- [14] As [11]
- [15] Colvocoresses. "The Effective Resolution Element of Remote Sensors", USGS Memo 1979, see [11]
- [16] Mundie et al. "The Design of Scanners for Remote Sensing Application" Proc. IEEE 63, 1, Jan. 1975
- [17] Slater, P. N. "Remote Sensing and Optical Systems" Addison Wesley 1980
- [18] Schagen, P. "Image Tubes with Channel Electron Multiplication", Advances in Electro Optical Devices
- [19] Ferranti Ltd. Data on High Resolution CRTs.
- [20] Hecht and Zajac, "Optics", Addison Wesley 1974
- [21] Optronics inc. Data Sheet on Filmwriting System
- [22] As [18]
- [23] Foley, J. D. and Van Dam, A. "Fundamentals of Interactive Computer Graphics", Addison Wesley 1982, p108
- [24] As [09], p207
- [25] As [23], p613
- [26] Electro Optical Systems Design, March 1980

- [27] Barnett, M. E., Williams, T. H. and Lee, R. M. "Television Techniques in Image Processing", O.A. vol 8, 1978
- [28] International Imaging Systems inc., System 500 and Model 70 image processor data
- [29] Computer Aided Design Centre, GEMS image processor data
- [30] Newton, I. in Bronowski, J. "The Ascent of Man", BBC
- [31] Bird, A. C. "The Electronic Image Classifier and its Applications to Multispectral Analysis", Ph.D Thesis 1981 Imperial College
- [32] Dainty, J. C. and Shaw, "Image Science", A.P. 1974, Chapter 2, 3
- [33] Selwyn, "A Theory of Graininess", Photogr. J. 75, 571
- [34] Nutting, "On the Absorption of Light in Heterogeneous Media", Phil. Mag. 26, 423
- [35] As [32], Chapter 9
- [36] Barnett, M. E. and Klemperer, O. "Electron Optics", C.U.P. 1971
- [38] Picinbono, "Modele Statistique Sugere par la Distribution des Grains d'Argent dans les Films Photographiques", Comptes Rendus 240, 2206, 1955
- [39] As [32], Chapter 8
- [40] Barnett, M. E., "Detection and Noise", Applied Optics MSc lecture course notes, 1981
- [41] Landsat data users handbook, NASA
- [44] Woods, R. and Gonzalez, R., "Real Time Digital Image Enhancement", Proc. IEEE, 60, 828, 1977
- [45] Frei, W. "Image Enhancement by Histogram Hyperbolisation", Computer Graphics and Image Processing 6, 286, 1977
- [46] Analog Devices inc. "ADC TV 8 B Video Digitiser Data Sheet"
- [47] Intersil Datel inc. "DAC HF 8 BMC Ultra High Speed DAC Data Sheet"
- [49] Thomson EFCIS "EF 9365/6 Graphics Display Processor data sheet"
- [50] Bresenham, J. "An Algorithm for the Computer Control of Digital Plotters", IBM Syst. J. 4(1), 1965
- [51] As [31], chapter 4
- [52] Santisteban, A. and Munoz, L. "Principal Components of a Multispectral Image: Application to a Geological Problem", IBM J. Res. Develop. Vol. 22, 5, Sept. 1978
- [53] Froberg, C. E. "An Introduction to Numerical Analysis", Ad. Wes. 1965, Chapter 6, p 120

- [54] Ready, P. J. and Wintz, P. A. "Information Extraction, SNR Improvement and Data Compression in Multispectral Imagery", IEEE Trans. Comm. 21, 10, Oct. 1973
- [55] Mulder, N. J. and Hempenius, S. A. "Data Compression and Data Reduction Techniques for the Visual Interpretation of Multispectral Images", ITC Journal 1974-3
- [56] TRW inc. VLSI Multiplier Device MPY8 HUJ data sheet
- [57] Abramowitz, M. and Stegun, I. "Handbook of Mathematical Functions", Dover books, p17
- [58] As [53]
- [59] TRW inc. High Speed Video DAC Type TDC 1016J data sheet
- [60] Analog Devices inc. CMOS Multiplying DAC Type DAC HA10 B data sheet
- [61] Mostek inc. Memory Data Book 1981, MK4116N-3 Dynamic RAM data sheet
- [62] Thomson- EFCIS, Graphics Display Controller EF9365/6 data sheet
- [63] Thomson- EFCIS, Graphics Display Controller EF9365/6 Applications Note
- [64] As [61], MK4802-1 2k Static RAM data sheet