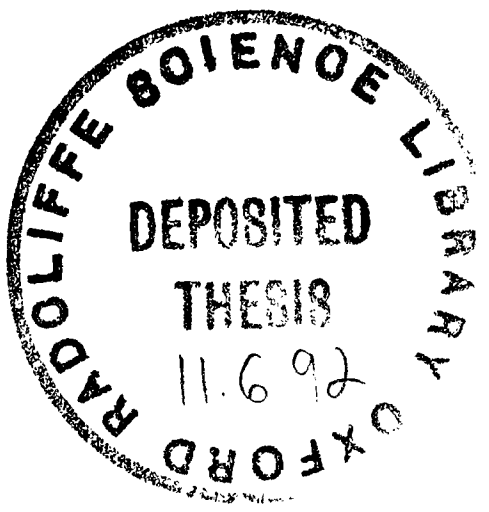
D.Phil Thesis
Towards Infrared Image Understanding[1]

Peter Foulkes[2]

Robotics Research Group
Department of Engineering Science
Oxford University

June 10, 1991



[1]This work has been carried out with the support of SERC and Procurement Executive Ministry of Defence

[2]Now a Junior Research Fellow at RSRE

Abstract

Peter Foulkes
Pembroke College, Oxford

Submitted for the degree of D.Phil.
Trinity Term, 1991.

Towards Infrared Image Understanding

An extensive literature survey has revealed that the majority of previous work in infrared image processing has ignored the processes leading to the formation of infrared images. Processing has normally either been restricted to simple low-level image enhancement convolutions or has consisted of algorithms copied from computer vision without regard for the inherent differences between infrared and visible images. In this thesis, we address the problem of infrared image formation and derive an irradiance equation for simple infrared scenes. We consider the complications caused by mutual illumination of one or more bodies and indicate how the infrared irradiance equation can also be specified for more complex scenes. The infrared irradiance equation we derive is solved in closed form for some simple geometries for both Lambertian and non-Lambertian surfaces. An infrared imager has been built and is described. Images taken with the imager of a variety of scene geometries show that the experimental results compare favourably with the theoretically derived equations, indicating the validity of the theoretical analysis. We describe how a knowledge of the formation of infrared images can be used to predict the image irradiance pattern of a particular object. We also show how, with a knowledge of the radiance properties and surface geometry of the object, it is possible to detect instances of that object in a scene. Examples are given of successful object detection based on an understanding of the image irradiance. We present a brief history of infrared imagers and a description of the principles on which modern infrared imagers are based. In addition to the survey of the literature published on infrared image processing, a brief summary of some techniques from the computer vision literature and their suitability to infrared image processing is given. A selection of vision techniques are applied to both infrared and visible images to verify conclusions reached in the thesis.

Acknowledgements

I should like to thank all the staff and students of the Robotics Research Group for their support, especially Henry Lau for the use of his motor driver and Will Dickson for his assistance with system problems.

Thanks also to my colleagues at RSRE, particularly Paul Manning for his help and advice in building the thermal imager and to Jim Sherlock who helped set up both the CASE award and the Research Fellowship which I now hold.

Above all, I would like to thank my supervisor Mike Brady for his enthusiasm and guidance throughout. Without his encouragement this work would never have been completed.

To my parents

Contents

List of Figures

Chapter 1

Introduction

"It is hard to find a black cat in a dark room - especially when the cat isn't there." *Confucius.*

Our eyes help us to interpret our environment. The process we call "seeing" involves the use of highly specialised sensors (our eyes) detecting energy which is transmitted in a narrow band of the electromagnetic spectrum. The portion of the electromagnetic spectrum which contains energy which our eyes can interpret is known as the visible part of the spectrum and energy transmitted in this part of the spectrum is known as visible light.

Visible light is not by any means the only useful energy in the electromagnetic spectrum. Radiation at much longer wavelengths is used to transmit radio and television signals in a way which is invisible to humans but can be detected and decoded using the correct apparatuses. In a radar system, electromagnetic waves are emitted by a device and the reflections of the emitted signal are monitored. The reflections indicate the direction and range of objects in the path of the

signal. Careful processing of these reflections can provide a description of the object which could be used for recognition. At much shorter wavelengths, X-rays which can penetrate many substances impervious to ordinary light, are used for examining human bones in medical diagnoses.

Devices like radar which send information into the environment are called active devices; while devices which take information from their environment but do not affect their environment, television cameras for example, are called passive devices. Electromagnetic radiation with many different wavelengths are utilized in different ways in both passive and active devices.

The subject of this thesis concerns the region of electromagnetic radiation known as the infrared. Infrared radiation is commonly used in active devices: in laser range finders for example, but it is devices which passively sense the infrared radiation emitted by a scene which are discussed in this thesis. It is not possible for humans to "see" unaided at infrared wavelengths, but with the correct sensing devices, it is possible to detect variations in intensity of energy transmitted at these wavelengths and to produce greyscale images similar to those produced by television cameras.

There are very good reasons for wishing to monitor energy in the infrared parts of the spectrum: for this is the region of the spectrum at which bodies at or around room temperature emit the majority of their energy. This means that a detector of infrared energy should be capable of detecting differences in temper-

ature between bodies or parts of bodies at or around human body temperature or detecting objects which either through the creation of their own heat or by reflecting heat from some other warm body differ in temperature from their environment. Because energy in the infrared part of the spectrum is emitted by warm bodies at all times of the day, devices sensing this form of energy will be useful at night as well as during the day.

It is because detection of infrared energy can lead to determination of the temperature of a body that the term "thermal imager" is used to describe a camera which takes images of the energy transmitted in these parts of the spectrum.

Since the information received when imaging at infrared wavelengths is different from the information received when imaging visual light, it can be used in applications where visual light does not provide the required information. For example, infrared detectors have been used for thermal surveys of houses (e.g. [2]); agricultural surveys (e.g. [60]); fire and explosion analyses (e.g. [40]); assisting medical diagnoses of such things as infections and cancers (e.g. [5]); assisting firemen locate bodies in buildings collapsed due to fires or earthquakes (e.g. [88]); even such unlikely uses as stress in beef cattle due to shipping [23] and identifying causes of male infertility [32] have been investigated using thermal imaging. Thermal imagers could even be used for detecting black cats in dark rooms .... Numerous military applications include night surveillance equipment; infrared rifle sights; tracking equipment for guided missiles and missile trackers.

As early as the 1950s, people were predicting the possible uses of thermal imaging (see, for example, [69]). Today we can see that some of those predictions were inaccurate, while some have been fulfilled. We in turn could speculate on possible future uses of thermal imaging and computer vision such as helping the blind or partially sighted to see clearly; but such suggestions at this moment in time would be pure speculation. We can say that computer vision has the potential to be useful in harsh or inhospitable environments where human life may be endangered; in areas where it is difficult or impossible for a human to reach or where communication to an "operator" is not possible. (e.g. nuclear plants; space; bomb diffusing; etc.) Such devices have already been used under human control. Autonomous devices will be faster and hopefully more accurate. We can say that thermal imaging has the potential to be a valuable diagnostic tool in the medical and other professions (indeed it already is used; and further uses are being explored.) Computer vision has already had its usefulness demonstrated in tasks where the task description is rigorously detailed. The more those constraints can be loosened by improving current techniques and developing new techniques, the more useful computer vision will be.

A CASE studentship (Cooperative Award in Science and Engineering) was set up by Oxford University and The Royal Signals and Radar Establishment (RSRE), Great Malvern, to investigate the nature of thermal images; to distinguish the differences between characteristics of thermal images and visual images;

to investigate how far algorithms developed for visual images can be used for thermal images; and based on these results, to determine how best to begin processing the thermal images. It was hoped that this research would lead to the formation of an image irradiance equation for thermal irradiance analogous to that developed by Horn [37] for visual light; and to a critical analysis of image processing techniques relevant to thermal imagery.

The literature survey of infrared image processing algorithms presented in chapter 8 reveals that most processing of infrared images takes little or no account of the underlying effects which cause the formation of the infrared images and furthermore that most work has been limited to fairly simple operations which are designed to work in real time without recourse to parallel computers.

In chapter 2 we describe the theoretical analysis we performed leading towards an infrared irradiance equation for different image geometries and surface types. A closed form solution for the irradiance equation is given for some simple image geometries.

In order to get empirical confirmation of the theoretical results, we built an infrared imager so that data could be acquired under experimental conditions. Details of the construction of the imager and of interfacing it with the Oxford computer system are presented in chapter 4. A brief history of infrared imagers and a review of the types of imagers currently available is given in chapter 3.

Problems specific to the type of imager we built and suggested solutions to

these problems are discussed in chapters 5 and 6. Chapter 5 deals with the early processing of images from multiple-element arrays while chapter 6 deals with some anomalies of the particular detector we built. Chapter 6 also describes a number of experiments we performed to confirm the analyses we made in chapter 2 and some experiments we performed to indicate how knowledge of infrared irradiance patterns can be used for detecting targets in a scene.

In conclusion, we show to what extent the original aim of the CASE award was achieved and indicate what direction we believe future work in this field should take.

Chapter 2

Theory of Infrared Radiation

2.1 Introduction

Infrared radiation was first noticed by Sir William Herschel in 1800. He repeated Newton's prism experiment, noticing a heating effect beyond the usual visual distribution of the spectrum. About a century later, due to the work of pioneers such as Maxwell and Hertz, Wien derived his equation of radiation using statistical methods and based on Maxwell's law of the distribution of velocity. A few years after that, starting from the law of equipartition of energy, Lord Rayleigh (1900) and Jeans (1905) proposed an alternative, contradictory equation.

The apparent discrepancy was explained by Max Planck who proposed that the radiation could be considered either as a wave or as packets of energy. Planck was able to explain the failure in Wien's and Rayleigh-Jeans's approaches to accurately predict the proper distribution of energy in the electromagnetic spectrum of a warm body. By suggesting that a particle of energy has a finitely small value, he exposed the failing in the classical statistics. Planck's work laid the foundations

of the theory of radiation as it is today. Further work, such as that by Einstein, who in 1905 presented his theory of the photoelectric effect, has led to the understanding we now have in radiation transfer, and to the development of thermal imaging devices.

Much of the work done in the physics of radiation, particularly quantum models, though interesting, is not directly relevant to this thesis. In this chapter we will consider only those equations which are needed for the development of our subsequent argument.

2.2 Radiation Laws

All bodies whose temperatures are above absolute zero emit energy. If a body is hotter than its surroundings, it will emit more energy to its environment than it receives from it. For bodies at or around 300K, the majority of the energy emitted by the body will be emitted in the infrared part of the spectrum, as heat.

The hotter the body gets, the easier it is to "feel" the heat: it appears to give off more radiation. If the body is warmed beyond about 800K, it begins to glow red. Further heating causes the colour to change to orange, yellow, and then a blue-white. The changes in colour are indicative of the fact that the frequency at which a body radiates depends on the temperature of the body.

In 1896, Wien noted that wavelength at which the peak radiation occurs is inversely proportional to the absolute temperature of the emitting body. In calculating the constant of proportionality, he derived "Wien's Displacement Law":

$\lambda_{max} = \frac{c}{T}$ where $c = 2898\mu deg$ (and where μ = microns). As most objects on earth are at a considerably lower temperature than 800K, we have a very strong incentive to conduct research into the detection of the radiation at infrared frequencies. Indeed, using Wien's Displacement law, we can see that for a body at 300K, the wavelength of the peak in the transmitted radiation is approximately 9.7μ which falls into one of the atmospheric windows in the infrared portion of the electromagnetic spectrum. The meaning of and reasons for atmospheric windows are discussed in section 2.6.

The amount of radiation emitted by a body depends not only on the temperature of the body, but also on the material of the body and on the nature of the surface of the body. In order to categorise the relative strengths of radiation, it was necessary to establish a datum. That datum arose based on work done by Kirchhoff. A *black body* was postulated. A black body was defined as a complete absorber of energy and the most effective radiator of energy. No real body behaves as a black body across the entire spectrum, but certain materials approach "blackness" over limited ranges of the spectrum. A body closely approximating a black body can be constructed using a cavity of opaque isothermal material with a small hole for energy to radiate into and out of.

From the definition of a black body, the energy emitted by any real body must be less than that of a black body. In 1860, Kirchhoff showed that the ratio of radiated energy of a body to the radiated energy of a black body: $\frac{W}{W_b}$, referred to

as the *emissivity* ϵ of the body, is a property of the body and its surface coating. Furthermore, the *absorptivity* α of a body, defined as the ratio of power absorbed by a body to the power incident on the body is also constant for a body at a particular temperature and wavelength; and for an opaque body is equal to the emissivity. Kirchhoff's law, which can be written $W = \epsilon W_b = \alpha W_b$, is valid for radiation at any wavelength, and hence for total radiation. A body which has a constant value of emissivity (between 0 and 1) across the spectrum is said to be a *gray body*, whereas a body in which the emissivity varies with wavelength $(0 \leq \epsilon(\lambda) \leq 1)$ is called a *selective radiator*, a *spectral body* or a *coloured body*.

In 1879, Stefan noticed that the quantity of radiation emitted from a black body was proportional to the area of the body and to the fourth power of the absolute temperature. This was proved for a perfectly black surface by Boltzmann in 1884. The Stefan-Boltzmann law: $W(T) = \sigma A T^4$, still widely used, relates the total energy emitted by a body (across the entire electromagnetic spectrum) to the area of the body and its temperature.

To see how emissivity varies with wavelength, we need to consider Planck's law, which relates amount W of energy radiated by a black body to the wavelength λ of radiated energy for a body at a particular temperature T:

$$W_\lambda(\lambda, T) = \frac{2\pi hc^2}{\lambda^5} \frac{1}{exp\frac{ch}{\lambda kT} - 1}$$

where $W_\lambda(\lambda, T)$ is given in $\frac{watts}{cm^3}$, h is Planck's constant, c is the speed of light in a vacuum and k is Boltzmann's constant. This is often rewritten in a more concise

form as:

$$W_\lambda(\lambda, T) = \frac{c_1}{\lambda^5} \frac{1}{exp\frac{c_2}{\lambda T} - 1}$$

where c_1 and c_2 are referred to as the first and second radiation constants. Graphs of radiated energy against wavelength for a variety of temperatures are a clear and compact way of illustrating the relationships between radiation, wavelength and temperature. Most text books explaining radiation laws include a graph of this type. (See for example [47], [81] or [39].)

Working from Planck's law, it is possible to see why the Rayleigh-Jeans law is only valid for long wavelengths. If we expand the exponential in Planck's law and disregard all terms except the first, we are left with the Rayleigh-Jeans law: $W_\lambda(\lambda, T) = \frac{c_1}{c_2} \frac{T}{\lambda^4}$. The expansion is only valid however if $\frac{c_2}{\lambda T} \ll 1$ i.e. $\lambda T \gg c_2 = 1.4 * 10^4 \mu m K$. In particular, for $T \simeq 300$, it is only valid if $\lambda \gg 47 \mu m$. It is therefore not valid for radiation of energy from room-temperature objects at infrared frequencies.

Integrating Planck's law from zero to infinity, we get $W(T) = \frac{2\pi^5 k^4}{15 c^2 h^3} T^4$, which we recognise as a version of the Stefan-Boltzmann law. The Wien displacement law can also be derived from Planck's law by differentiating it with respect to wavelength for a particular temperature and calculating the maximum.

In 1760, Lambert proposed that the radiation coming from a diffusely radiating body is highest in the direction of the surface normal of the body and falls off in proportion to the cosine of the angle from the normal in other viewing direc-

tions. Lambert's law, which expresses this mathematically is only valid for perfect diffusers (which in the case of infrared means black bodies), but serves as a useful tool for analysing approximately black surfaces.

A consequence of Lambert's law is that a true diffuse emitter appears equally bright from all view directions. This is because when viewed at an angle θ from the surface normal, the apparent area of a surface patch is *foreshortened* by a factor $\cos(\theta)$. The energy emitted in the direction θ is, from Lambert's law $cos(\theta)$ times the energy emitted in the direction of the surface normal. The perceived brightness is the same, therefore, since the amount of energy being radiated from a perceived unit area is equal. This means that an intensity profile of a uniform temperature warm convex object will indicate equal radiance across the surface. This has significant consequences which will be discussed later.

2.3 Radiation transfer between bodies

In considering the effect of radiation from one object to another or from an object to an imager, we need first to develop an equation relating the radiated energy transmitted from an element of one surface patch to an element of another surface patch when the two surfaces are at arbitrary orientations in space. In principle we can then calculate the radiant energy from any body to any other body by integrating the elemental radiances over the surfaces of the bodies.

Consider the elemental surface patches dS_1 and dS_2 of the two surfaces S_1 and S_2 depicted in figure 2.1. The line joining the elemental surface patches is of

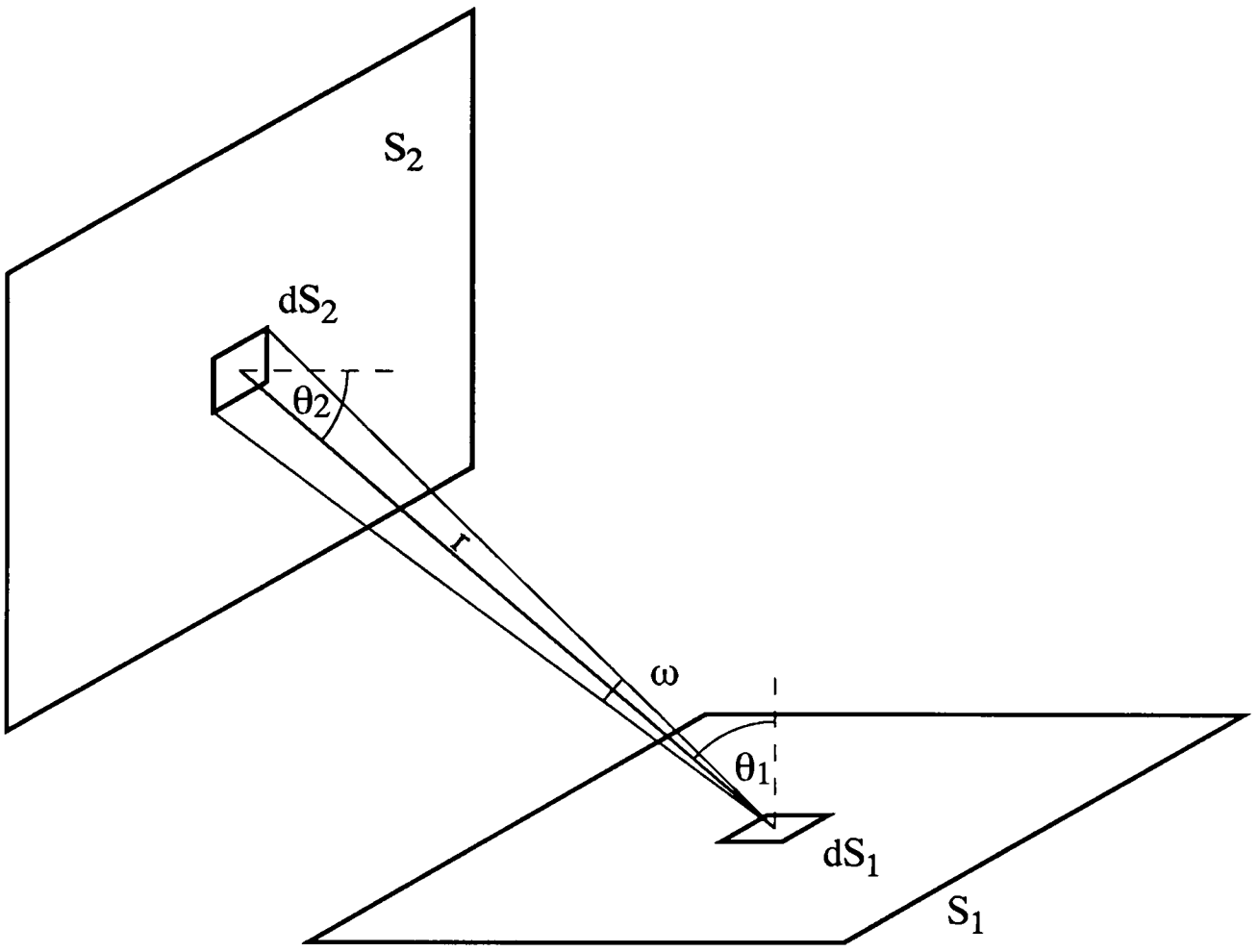


Figure 2.1: Exchange of radiated energy between two surfaces

length r and makes an angle θ_1 with the normal from dS_1 and an angle θ_2 with the normal from dS_2. When looking from point 1, the centre of the surface patch dS_1, the patch dS_2 subtends a solid angle $d\omega_2$, where $d\omega_2 = \frac{dS_2 cos(\theta_2)}{r^2}$. Following Stefan-Boltzmann, the radiant power emitted by dS_1 which is incident on element dS_2 is given by $dW_1 = \epsilon_1 \sigma T_1^4 cos(\theta_1) dS_1 d\omega_2$, where T_1 is the absolute temperature of S_1 and ϵ_1 is the emissivity of S_1. Substituting for $d\omega_2$, we have:

$$dW_1 = \frac{\epsilon_1 \sigma T_1^4}{r^2} cos(\theta_1) cos(\theta_2) dS_1 dS_2 \tag{2.1}$$

Integrating over the surface S_2, we find that the irradiation on the surface S_2 due to the surface patch dS_1 is

$$W_1 = \epsilon_1 \sigma T_1^4 dS_1 \int_{S_2} \frac{\cos\theta_1 \cos\theta_2}{r^2} dS_2 \tag{2.2}$$

The term "view factor" or "geometrical factor" of the surface S_2 from the surface S_1 is given to the integral: $F_{12} = \int_{S_2} \frac{\cos\theta_1 \cos\theta_2}{r^2} dS_2$. Substituting F_{12} into equation 2.2 gives $W_1 = \epsilon_1 \sigma T_1^4 dS_1 F_{12}$. The integral F_{12} can only be evaluated explicitly for simple geometric shapes. Integrating over the surface S_1 and assuming that the surfaces are opaque (and hence that $\alpha = \epsilon$), we have the total power absorbed by surface S_2 radiated by surface S_1:

$$W_{12} = \epsilon_1 \epsilon_2 \sigma T_1^4 \int_{S_1} F_{12} dS_1$$

It is clear that a similar calculation can be made for the energy radiated from surface S_2 incident on surface S_1. Noticing that $\int_{S_1} F_{12} dS_1 = \int_{S_2} F_{21} dS_2$ we can

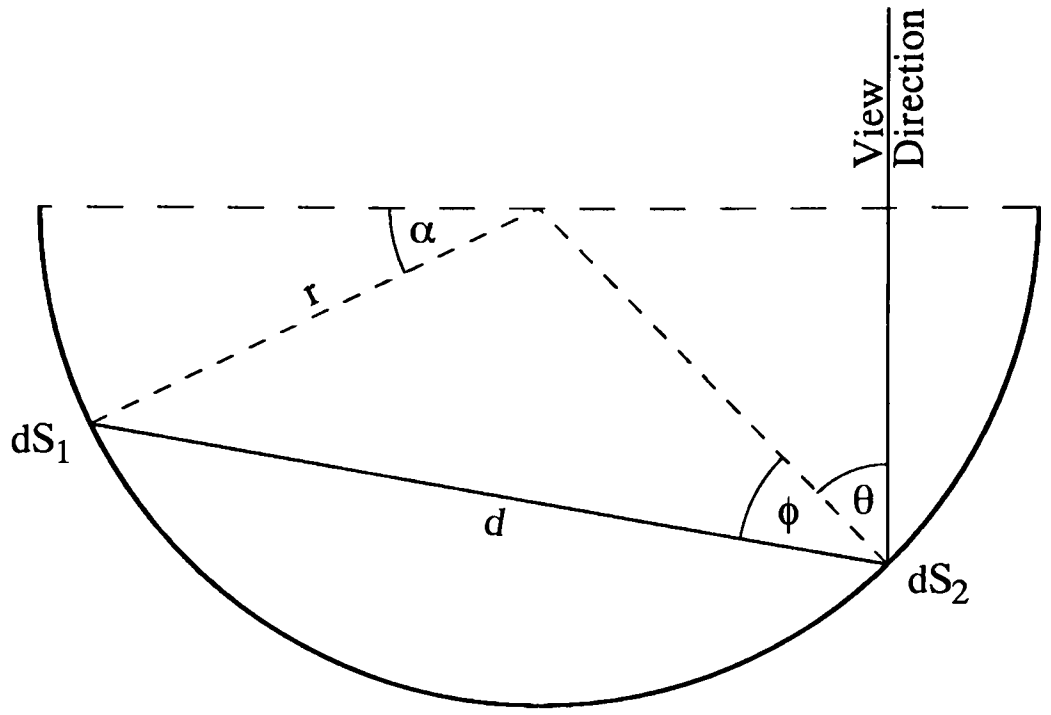


Figure 2.2: Geometry of self-irradiation of concave cylinder

get an expression for the net radiant power exchanged between the two surfaces:

$$W = \epsilon_1 \epsilon_2 \sigma (T_1^4 - T_2^4) \int_{S_1} F_{12} dS_1$$

2.4 Self irradiation

Consider now the effect of radiation of one part of a body affecting another part of the same body. This will occur whenever a body has concavities. As stated in the previous section, it is not possible to obtain explicit solutions for the view factor unless the geometry of the problem is quite simple. Here we begin by analysing a manageable problem: that of a Lambertian emitting concave semi-cylinder.

Figure 2.2 shows a section of a concave semi-cylinder viewed in the direction indicated. From the geometry of the cylinder, we note that the angle between the surface normal of element dS_1 and the line joining dS_1 to dS_2 is also equal to ϕ.

We make the reasonable assumption that emissivity is constant in the cylinder and we will assume that the temperature of the cylinder is constant throughout. The general problem described in the previous section is thus greatly simplified.

The distance d between point 1 and point 2 can be calculated by considering the geometry of the figure: $d = 2r\cos\phi$. As was the case in the previous section, the solid angle subtended by the patch dS_2 is given by $d\omega_2 = \frac{dS_2\cos\phi}{d^2}$. Substituting for d, we get $d\omega_2 = \frac{dS_2}{4r^2\cos\phi}$. The radiant power emitted by dS_1 for the general case is given by equation 2.1. Making the relevant changes for this specific case, we have:

$$dW_1 = \frac{\epsilon\sigma T^4}{4r^2} dS_1 dS_2. \tag{2.3}$$

We notice that dW_1 is not dependent on either θ or ϕ. To calculate the total irradiance at S_2 simply involves integrating with respect to α for $0 \leq \alpha \leq \pi$.

Now consider the hypothetical case of a cylinder which is a Lambertian emitter but a spectral reflector of infrared irradiation. To simplify further, assume that multiple reflections do not occur. (For a real surface, this is usually not unreasonable as the thermal reflectivity of most materials is quite low.) Figure 2.3 shows the ray trace of two rays emanating from the wall of the cylinder and reflecting from a different part of the wall towards the viewer.

For this spectral reflection case each position on the cylinder wall reflects radiation from only one other position on the wall in the direction of the viewer. The radius at the point of reflection bisects the angle between the line of sight

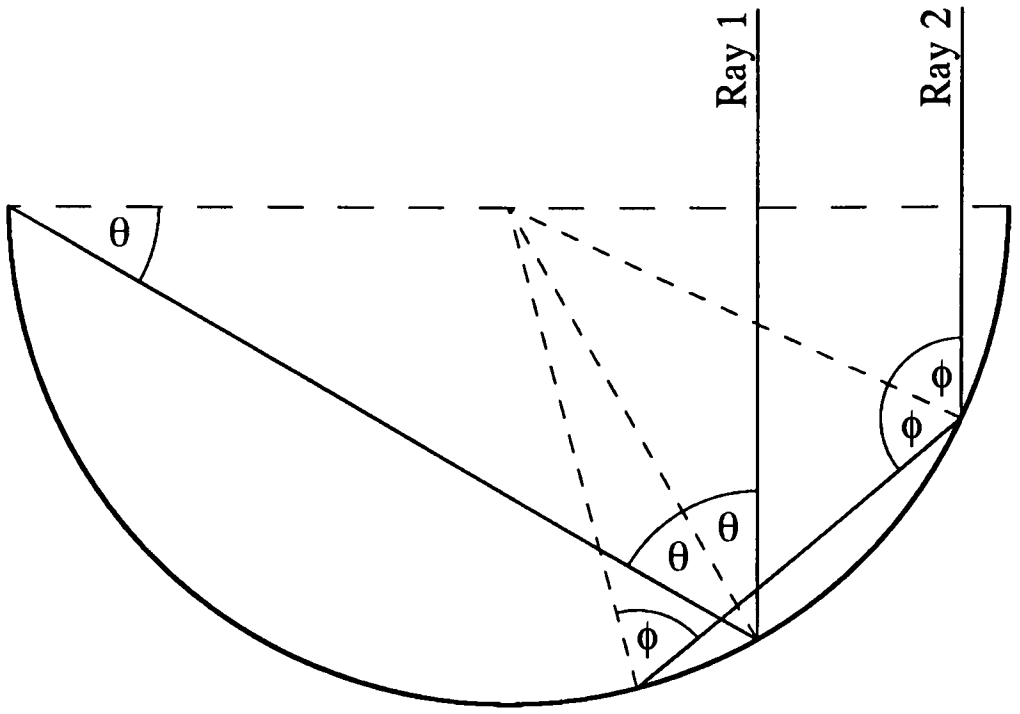


Figure 2.3: Geometry for specular reflection of irradiation

of the viewer and the line to the point of emission. Ray 2 is a typical example of a position on the surface of the cylinder reflecting the irradiation from another position on the surface of the cylinder.

For a Lambertian emitter, equation 2.3 tells us that the magnitude of the irradiation at dS_2 (figure 2.2) is independent of the angle ϕ. Given that this is so, then the irradiation at the viewer due to reflection will be the same for all points on the surface of the cylinder which reflect radiation in the direction of the viewer. Not all points on the surface will reflect in the direction of the viewer, however: only points which satisfy the geometry required for reflection. In figure 2.3, Ray 1 illustrates the limiting case: any points nearer to the centre of the cylinder will not reflect another point on the cylinder towards the viewer. For this limiting case, θ is easily calculated as $\frac{\pi}{6}$ radians.

The reflective component of the intensity profile will have two regions of equal intensity starting at the edges of the semi-cylinder, each one quarter of the diameter of the cylinder surrounding a region with the same intensity as the background.

Consider now a body which is not a Lambertian radiator of energy. Assume that the body has a radiation distribution which has its strongest component in the direction of the surface normal with intensity of radiation dropping away rapidly in other directions. This model fits many non-black bodies. If we were viewing a convex body with this type of radiation distribution, we would expect the irradiance to reduce as the surface normal turned further away from the viewer. We will show in chapter 6 that certain bodies do indeed give this type of intensity profile.

In the absence of any mutual radiation effects, we would expect a concave body to have a similar intensity profile to a convex body: parts of the body's surface directed towards the viewer would appear brighter. For bodies which have radiance patterns of this type and which also exhibit reflection effects, we would expect a radiance distribution which was a superposition of the radiance distributions for the spectral reflector and the non-Lambertian emitter. This might appear similar to the intensity profile illustrated in figure 2.4.

For a real body, we would not expect the transition between the region in which no reflection occurs to the region in which reflection occurs to be a step as it is in the figure. The assumption that reflection of irradiation will be only in

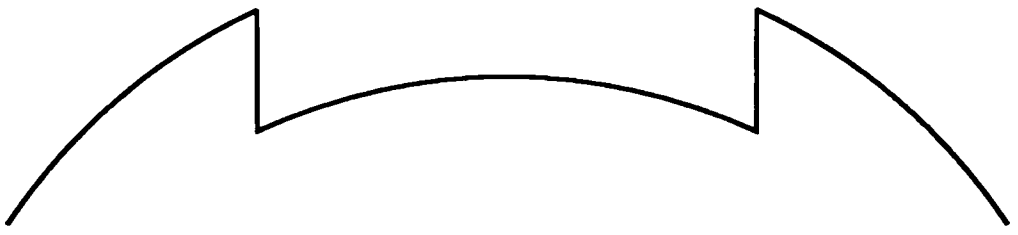


Figure 2.4: Predicted intensity profile of specularly reflecting concave non-Lambertian radiator

one direction is clearly unlikely to be true for real objects. If instead we postulate that the radiation will be scattered in a similar manner to the radiation pattern assumed for the non-Lambertian radiator, then the sharp corners in the expected intensity profile would be smoothed.

A model was proposed by Phong [71] for reconstructing the reflection characteristics of surfaces to generate a computer graphical representation of a surface coated with glossy white paint - a linear combination of a Lambertian component corresponding to matte white paint and a specular reflector for the lacquer cover. Horn [37] adapted the model and used it to predict the reflectivity function of a surface given the light distribution and viewing geometry. The equation used by Horn was:

$$\phi(I,E,G) = \frac{1}{2}s(n+1)(2IE-G)^n + (1-s)I \tag{2.4}$$

where s is the fraction of incident light reflected specularly and I, E, G are the cosines of the incident i, emitted e and phase g angles respectively. The value of n determines the sharpness of the specularity peak. We will now consider how this model can be applied to the problem described above.

Consider, again using the geometry of figure 2.2, a non-Lambertian emitter

which emits primarily in the direction normal to the surface. Let the energy emitted from dS_1 in the direction ϕ be proportional, not to $\cos(\phi)$ but to $\cos^2(\phi)$. Using the analysis leading to equation 2.3, we have

$$dW_1 = \frac{\epsilon\sigma T^4}{4r^2}\cos(\phi)dS_1dS_2. \tag{2.5}$$

The reflection of the ray from dS_1 at dS_2 is determined by the nature of the surface; by the incident angle ϕ; by the emitted angle θ and by the phase angle $(\phi+\theta)$ and can be approximated by equation 2.4. Due to the two dimensional nature of this problem, we can calculate $(2IE-G)$ quite simply:

$$\begin{aligned} 2\cos\phi\cos\theta - \cos(\phi+\theta) &= 2\cos\phi\cos\theta - (\cos\phi\cos\theta - \sin\phi\sin\theta) \\ &= \cos\phi\cos\theta + \sin\phi\sin\theta \\ &= \cos(\phi-\theta) \end{aligned}$$

Combining equations 2.4 and 2.5, the spectral component of the reflected beam in direction θ is given by: $\frac{1}{2}s(n+1)cos^n(\phi-\theta)\frac{\epsilon\sigma T^4}{4r^2}\cos(\phi)dS_1dS_2$. Integrating with respect to ϕ over the possible range of values and noticing that $dS_1 = 2rd\phi$, we get the total energy seen by the viewer from the point defined by θ due to reflection from other parts of the cylinder:

$$W_1 = s(n+1)\frac{\epsilon\sigma T^4}{4r}dS_2\int_{\phi_1}^{\phi_2}\cos^n(\phi-\theta)\cos(\phi)d\phi \tag{2.6}$$

We now evaluate the integral. Substituting $y=(\phi-\theta)$ gives $\int_{y_1}^{y_2}\cos^n y\cos(y+\theta)dy$. Expanding $\cos(y+\theta)$ and tidying up, the resulting integral is given by:

$\cos\theta \int_{y_1}^{y_2} \cos^{n+1}(y)dy - \sin\theta \int_{y_1}^{y_2} \cos^{n}(y)\sin(y)dy$. The second integral can be evaluated quite simply. Re-substituting $y = \phi - \theta$ into the solution we have:

$$\sin\theta \left[\frac{1}{1+n}\cos^{n+1}(\phi-\theta)\right]_{\phi_1}^{\phi_2} \tag{2.7}$$

We can get a closed form solution to the first integral for even values of n. Let $m = n/2$ and substitute $\cos^2(y) = (1 - \sin^2(y))$ then the first integral is given by $\cos\theta \int_{y_1}^{y_2}(1-\sin^2(y))^m \cos(y)dy$. We now expand $(1-\sin^2(y))^m$ as a binomial: $(1-\sin^2(y))^m = \sum_{k=0}^{m} \frac{(-1)^k m!}{k!(m-k)!}\sin^{2k}(y)$. We can now integrate the expression and the solution to the first integral for even n is given by

$$\cos\theta \left[\sum_{k=0}^{m} \frac{(-1)^k m!}{k!(m-k)!}\cdot\frac{1}{2k+1}\cdot\sin^{2k+1}(\phi-\theta)\right]_{\phi_1}^{\phi_2} \tag{2.8}$$

(re-substituting $y = (\phi - \theta)$).

The limits for ϕ can be calculated from the geometry of the figure for the parts of the figure to the left of dS_2: $\phi_1 = \frac{\pi-2\theta}{4}$ and $\phi_2 = \frac{\pi}{2}$. Because of the way the angles were defined, it was necessary to treat the portion of the cylinder to the right of dS_2 slightly differently to the portion to the left of dS_2. The mathematics is very similar, however and will not be repeated here. The spectral reflection component of the solution of the original equation is proportional to the sum of the results of equations 2.7 and 2.8.

Calculating the non-specular reflection component is much simpler. Using equations 2.4 and 2.5; and making the substitution $dS_1 = 2rd\phi$, the non-specular component is given by: $(1-s)\frac{\epsilon\sigma T^4}{2r}dS_2 \int_{\phi_1}^{\phi_2} \cos^2(\phi)d\phi$. The integral can be solved

to give:

$$\left[\frac{1}{2}\phi + \frac{1}{4}\sin(2\phi)\right]_{\phi_1}^{\phi_2} \tag{2.9}$$

Using the same angle definitions as before, for consistency, we calculated this integral in two parts: to the left and the right of dS_2 and summed the result.

The final component of radiation reaching the viewer using this model is that which is emitted directly from the body of the cylinder. This is also simple to calculate using the radiation distribution of the surface which was defined earlier to be proportional to $\cos^2(\theta)$.

The total solution for the irradiance at the viewer was calculated. The irradiance distribution depends on the radius of the cylinder: the smaller the radius, the greater the importance of the reflected components. Figure 2.5 shows the irradiance profile predicted by the model described above for a unit radius with n chosen for a "fairly specular" surface (as defined by Phong).

We can see that taking account of the assumptions we made about smoothing sharp corners of the profile we anticipated from our earlier discussion, and plotted in figure 2.4, the two approaches would give fairly similar results: of course the first approach was only intended to give a rough guide to the shape; while the second plot depends strongly on the accuracy of the assumptions made.

We show in chapter 6 that intensity profiles of irradiation from non-black body concave cylinders are indeed very similar to the intensity profiles we have predicted. For ease of reference, we reproduce one of the figures from chapter 6 in

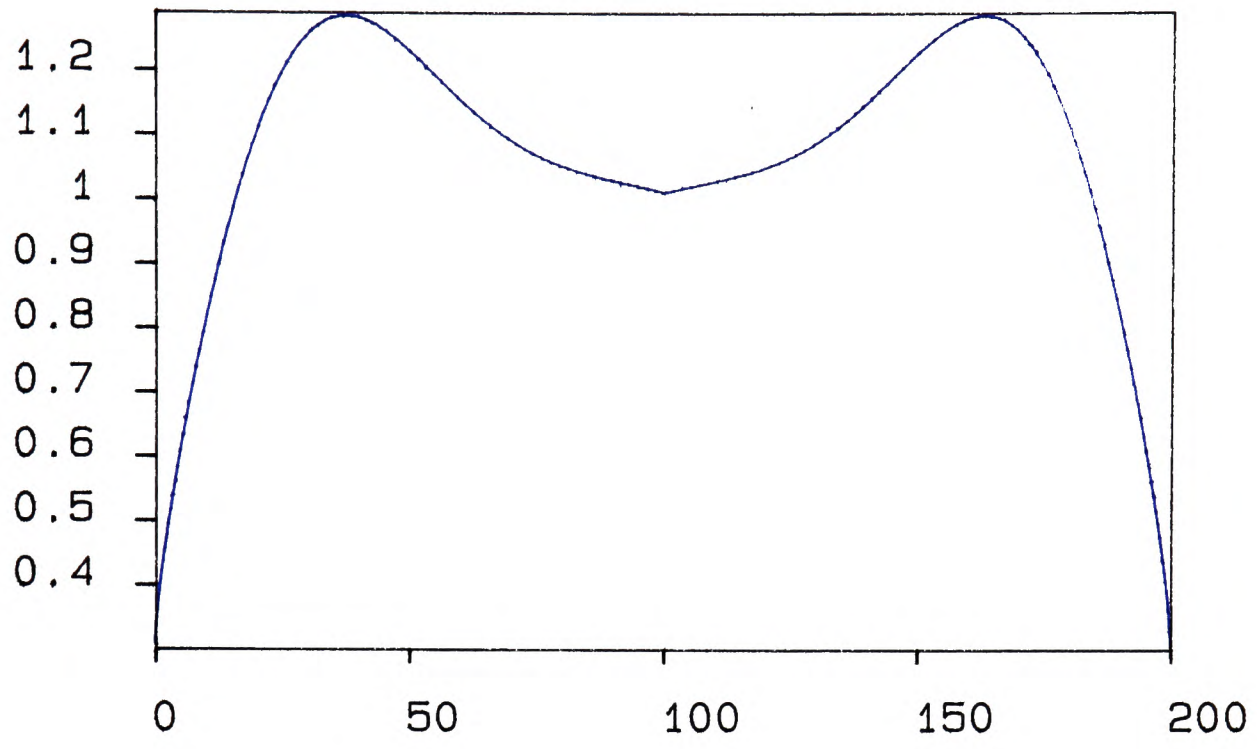


Figure 2.5: Predicted intensity profile of specularly reflecting concave non-Lambertian radiator

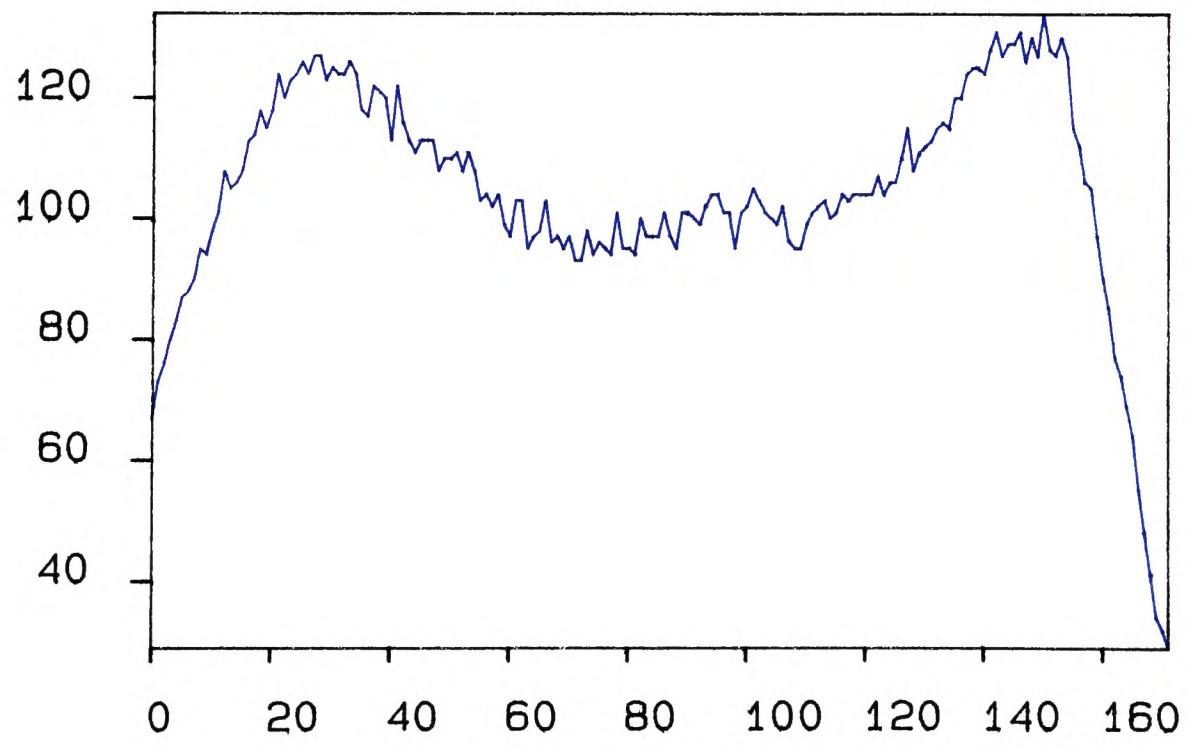


Figure 2.6: Observed intensity profile of specularly reflecting concave non-Lambertian radiator

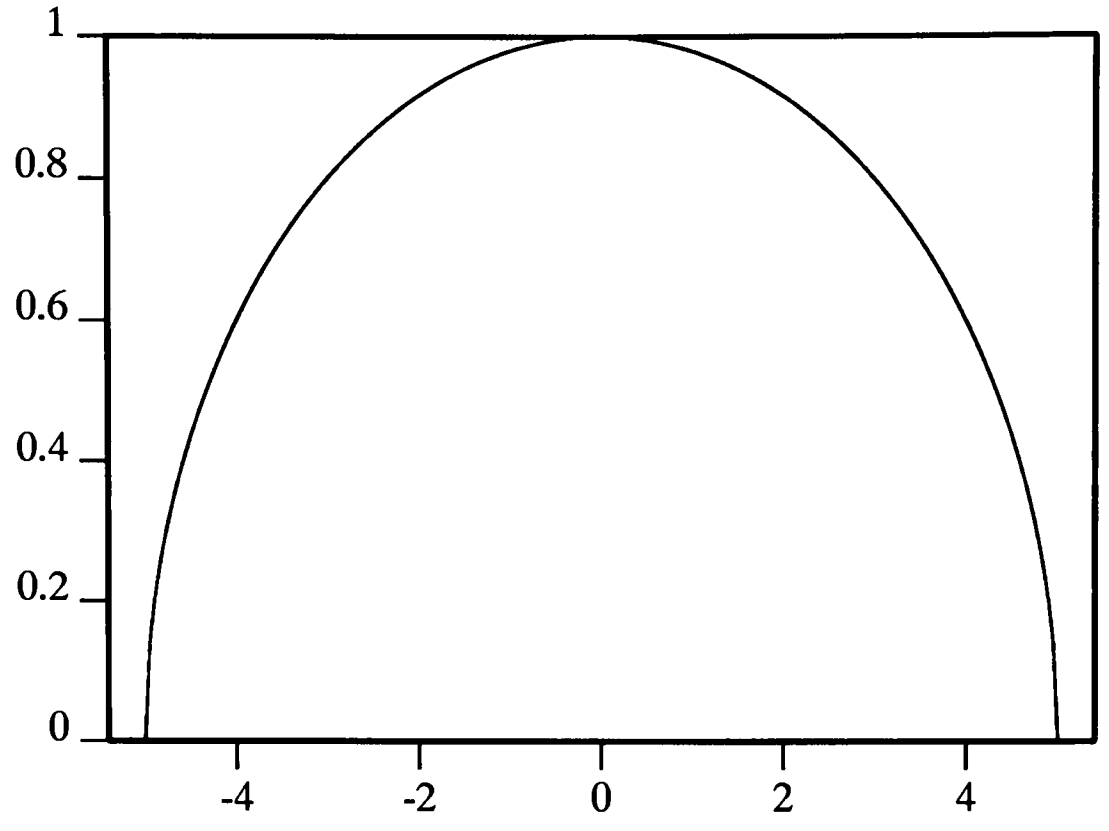


Figure 2.7: Predicted intensity profile of a convex non-Lambertian radiator

figure 2.6. The intensity profile in figure 2.6 is plotted for a white-is-hot image for comparison with figure 2.4 (In chapter 6 most of the images are shown black-is-hot since black-is-hot images reproduce better when printed). Clearly, the profile predicted by the model is very similar to the profile obtained experimentally.

We also show in figure 2.8 that the intensity profile of a convex non-Lambertian cylinder is similar to that which we would expect from the theory (see figure 2.7).

2.5 Self irradiation of an angle

In section 2.4, using a reflection model suggested by Horn, we derived a model for the irradiance profile across a non-Lambertian semi-cylinder. In this section, we repeat the procedure for a non-Lambertian angle.

Consider the equal ninety degree angle depicted in figure 2.9. Let the length

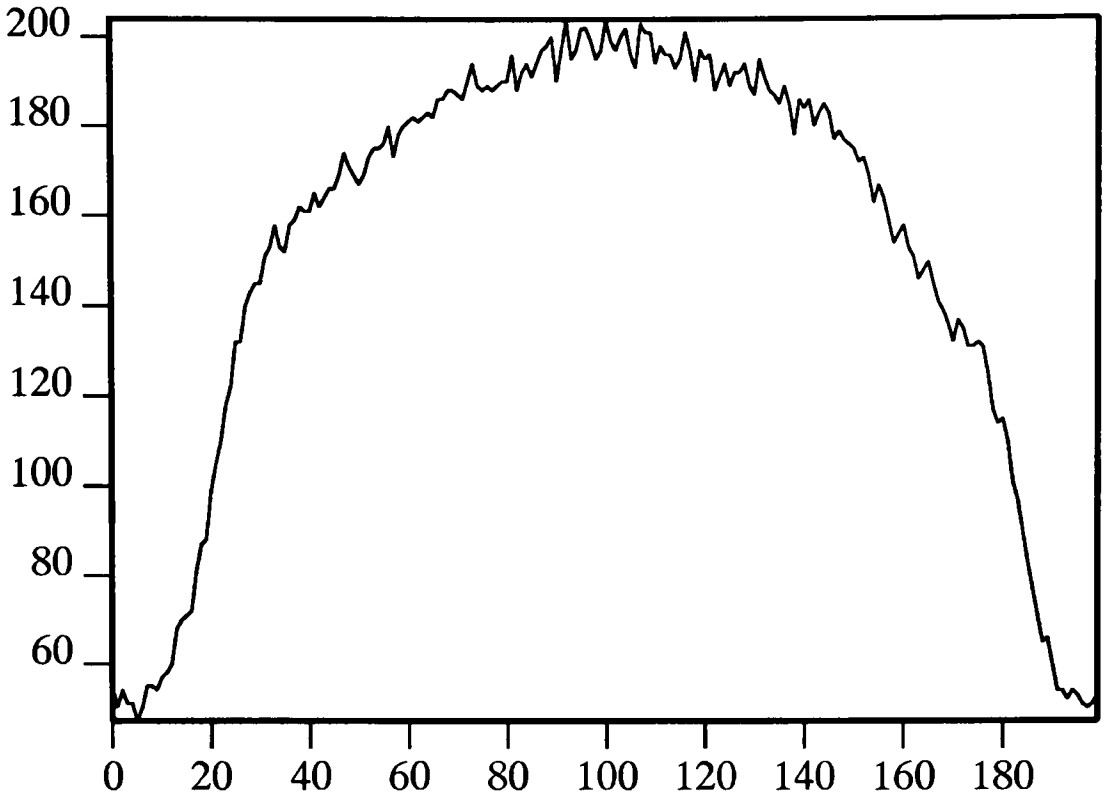


Figure 2.8: Observed intensity profile of a convex non-Lambertian radiator

of both sides of the angle be L; let the elemental patches at x and y be dx and dy respectively and let the solid angle subtended by dx when looking from point y be $d\omega$. The distance between the point at x and the point at y is r.

Now, $d\omega = \frac{dx\cos\theta}{r^2}$. Substituting for r, we have $d\omega = \frac{\cos\theta\sin^2\theta}{x^2}dx$. Now, if we assume the same radiance function as before (radiance proportional to cos^2), we have: $dW = \epsilon\sigma T^4\cos^2(\frac{\pi}{2} - \theta)d\omega dy$. Substituting for $d\omega$, we have: $dW = \epsilon\sigma T^4\frac{\sin^4\theta\cos\theta}{x^2}dxdy$.

Consider the effect caused by the radiation from side y of the angle onto the point x of the other side of the angle. For this situation, we can treat x as constant and consider the effect of varying y, or alternatively of varying θ. Notice that $\tan\theta = x/y$, so $dy = -\frac{x}{sin^2\theta}d\theta$. In terms of x and θ, we therefore have:

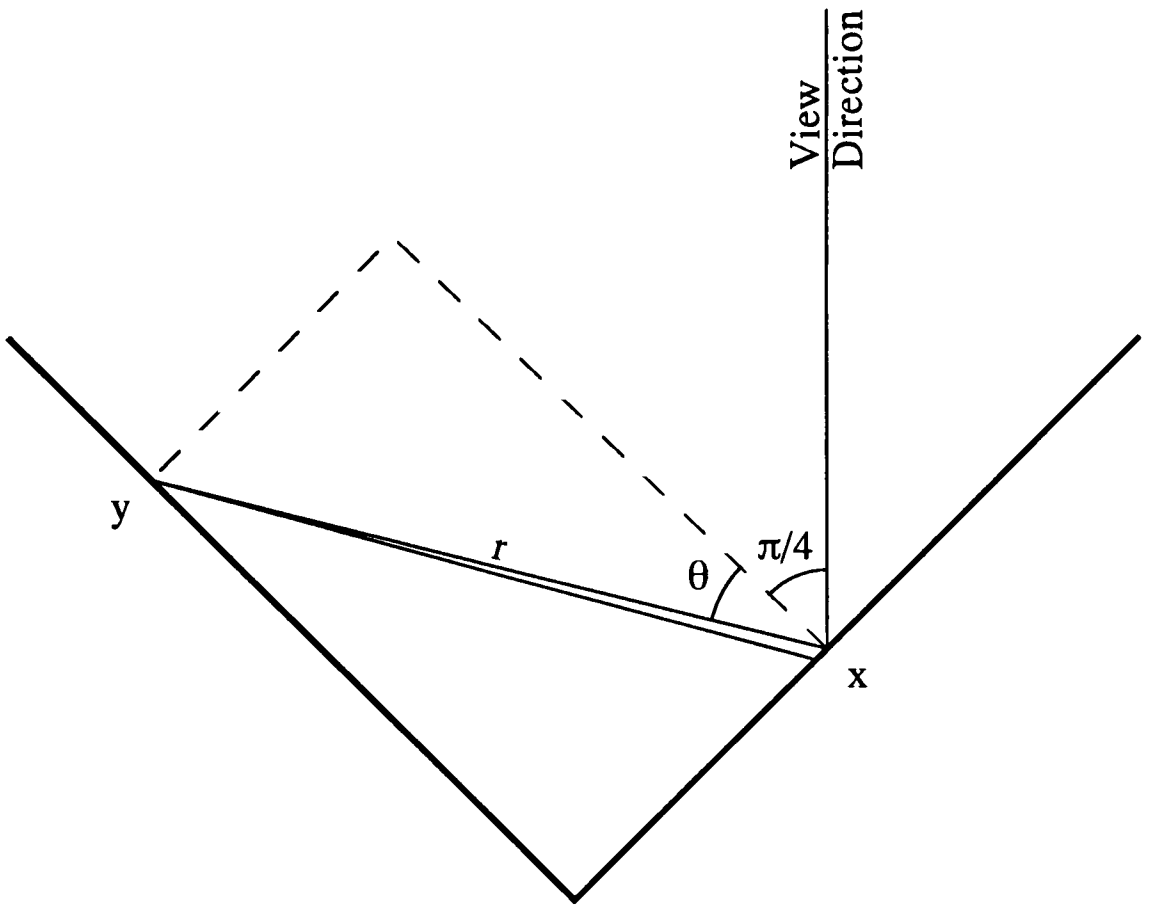


Figure 2.9: Geometry of equal angle

$$dW = -\epsilon\sigma T^4 \frac{cos\theta \sin^2\theta}{x} dx d\theta \tag{2.10}$$

As before, we can simplify the expression $(2IE - G)$ for the specularly reflective part of Horn's equation: $(2IE - G) = \cos(\theta - \frac{\pi}{4})$.

We now have the specular component of reflection in the direction of the viewer at point x due to dy as: $-\epsilon\sigma T^4 \frac{1}{2} s(n+1)\cos^n(\theta - \frac{\pi}{4}) \frac{\cos\theta \sin^2\theta}{x} dx d\theta$. The total specular component of reflection in the direction of the viewer at point x due to side y is therefore:

$$-\epsilon\sigma T^4 \frac{1}{2} s(n+1) \frac{dx}{x} \int_{\theta_1}^{\theta_2} \cos^n(\theta - \frac{\pi}{4}) \cos\theta \sin^2\theta d\theta \tag{2.11}$$

Solving the integral, I, we first substitute $z = \theta - \frac{\pi}{4}$ and notice that $\cos(z + \frac{\pi}{4}) = \cos\frac{\pi}{4}(\cos z - \sin z)$ and $\sin(z + \frac{\pi}{4}) = \sin\frac{\pi}{4}(\cos z + \sin z)$. We also notice that $\cos\frac{\pi}{4} = \sin\frac{\pi}{4}$. For convenience of notation, let $A = \cos\frac{\pi}{4}$. Making these substitutions and tidying up, we have:

$I = A^3 \int_{z_1}^{z_2} \cos^n z(2\cos^3 z - \cos z + 2\cos^2 z \sin z - \sin z)dz$

We now split the integral into four parts:

$$I = A^3(I_1 + I_2 + I_3 + I_4). \tag{2.12}$$

where:

$$I_1 = \int_{z_1}^{z_2} 2\cos^{n+3} z dz \tag{2.13}$$

$$I_2 = \int_{z_1}^{z_2} -\cos^{n+1} z dz \tag{2.14}$$

$$I_3 = \int_{z_1}^{z_2} 2\cos^{n+2} z \sin z dz \tag{2.15}$$

$$I_4 = \int_{z_1}^{z_2} -\cos^n z \sin z dz \tag{2.16}$$

Solving $(I_3 + I_4)$ is straight forward. Re-substituting $z = \theta - \frac{\pi}{4}$ into the result gives:

$$I_3 + I_4 = \left[\frac{-2\cos^{n+3}(\theta - \frac{\pi}{4})}{n+3} + \frac{\cos^{n+1}(\theta - \frac{\pi}{4})}{n+1}\right]_{\theta_1}^{\theta_2} \tag{2.17}$$

We can get a closed form solution for $I_2 = \int_{z_1}^{z_2} -\cos z \cos^n z dz$, for even values of n. Let $m = n/2$ then: $I_2 = \int_{z_1}^{z_2} -\cos z(1 - \sin^2 z)^m dz$. Expanding $(1 - \sin^2 z)^m$ as a binomial and solving the resulting integral gives (re-substituting $z = \theta - \frac{\pi}{4}$):

$$I_2 = \left[-\sum_{k=0}^{m} \frac{(-1)^k m!}{k!(m-k)!} \frac{\sin^{2k+1}(\theta - \frac{\pi}{4})}{2k+1}\right]_{\theta_1}^{\theta_2} \tag{2.18}$$

We can solve I_1 in a similar way to give:

$$I_1 = \left[2 \sum_{k=0}^{m+1} \frac{(-1)^k (m+1)!}{k!(m+1-k)!} \frac{\sin^{2k+1}(\theta - \frac{\pi}{4})}{2k+1} \right]_{\theta_1}^{\theta_2} \tag{2.19}$$

Using equations 2.11, 2.12 and 2.13 - 2.16 we have the specular component of reflection in the direction of the viewer from point x due to the side y as:

$$-\epsilon\sigma T^4 \frac{1}{2} s(n+1) \frac{dx}{x} A^3 (I_1 + I_2 + I_3 + I_4) \tag{2.20}$$

where the values of I_1, I_2, I_3 and I_4 are given in equations 2.17, 2.18 and 2.19.

We now consider the Lambertian component of reflection in the direction of the viewer at point x due to dy: $(1-s)\cos\theta dW$. The total Lambertian component of reflection in the direction of the viewer at point x due to side y is therefore:

$$\frac{-(1-s)\epsilon\sigma T^4 dx}{x} \int_{\theta_1}^{\theta_2} \cos^2\theta \sin^2\theta d\theta \tag{2.21}$$

We now solve the integral part, I. Noting that $\cos^2\theta = \frac{1}{2}(1+\cos 2\theta)$ and $\sin^2\theta = \frac{1}{2}(1-\cos 2\theta)$, we get: $I = \frac{1}{4}\int_{\theta_1}^{\theta_2} \sin^2 2\theta d\theta$. Simplifying, we get: $I = \frac{1}{8}\int_{\theta_1}^{\theta_2}(1-\cos 4\theta)d\theta$, which is easily solved. Combining the solution with equation 2.21, we have the Lambertian component of reflection in the direction of the viewer at point x due to side y:

$$\frac{-(1-s)\epsilon\sigma T^4 dx}{8x} \left[\theta - \frac{\sin 4\theta}{4} \right]_{\theta_1}^{\theta_2} \tag{2.22}$$

As with the semi-cylinder, we calculate the total solution for the irradiance at the viewer by combining the results from equations 2.20 and 2.22 with the component of radiation emitted by the body in the direction of the viewer. This

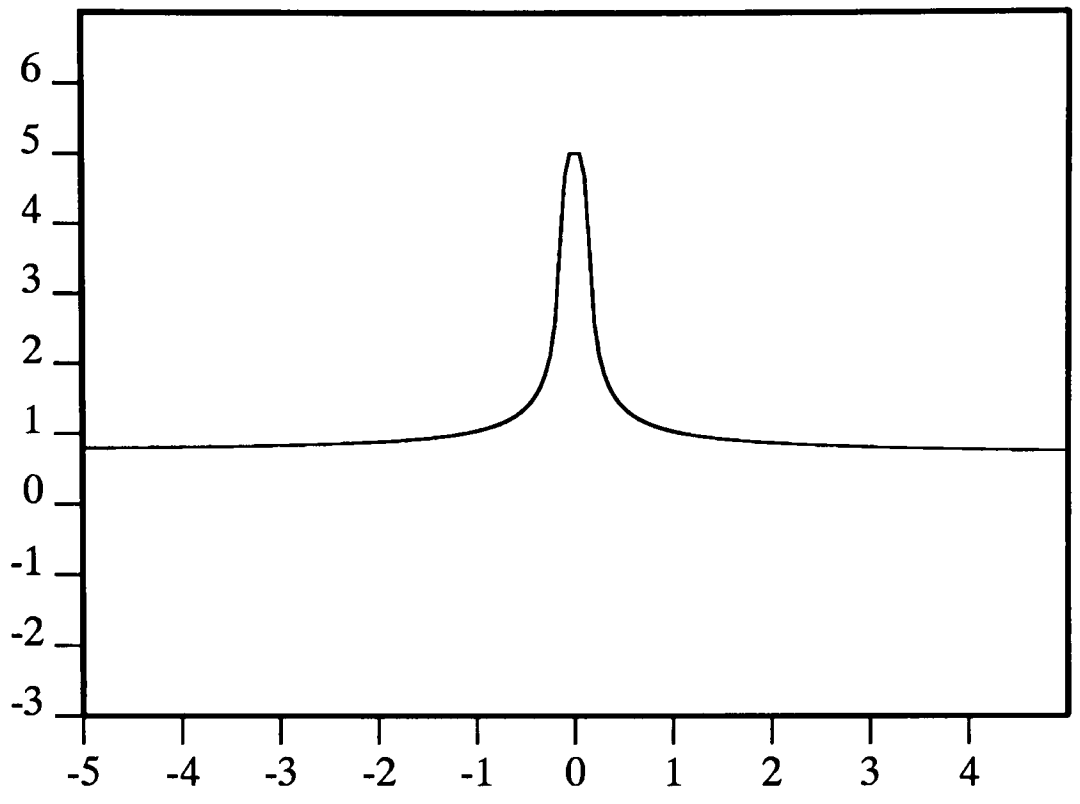


Figure 2.10: Predicted intensity profile of equal angle

theory predicts that there will be an infinite peak at the very centre of the angle. Obviously, this will not occur in real images: real edges will be rounded towards the very centre and when distances get very small, the wavelength of the radiation will affect the equation. Figure 2.10 shows the irradiance profile predicted by this model with the value at the very centre rounded off.

Figure 2.11 shows an image of a concave angle taken with a thermal imager. A cross-section of this image is shown in figure 2.12. Because of the large amount of noise in the image, the cross-section has been taken as an average of three lines of the image. Although a large amount of noise still remains, it is possible to see that the predicted intensity profile of figure 2.10 is very similar to the experimental results depicted in figure 2.12.

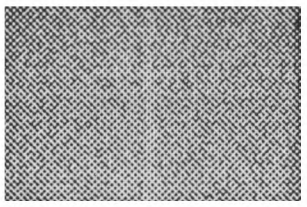


Figure 2.11: Image of a concave equal angle

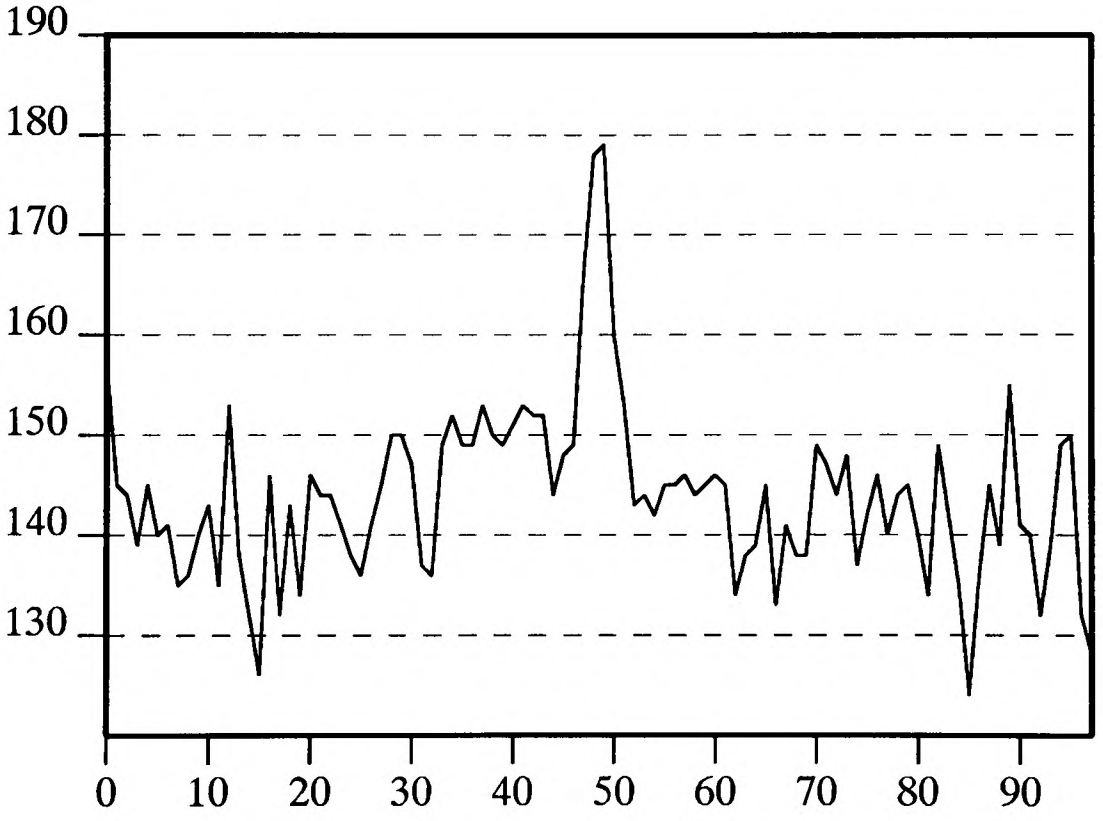


Figure 2.12: Actual intensity profile of an equal angle

2.6 Atmospheric Attenuation

When one body radiates energy towards another body, the energy has to pass through some medium; in most thermal imaging applications this medium will be the atmosphere. In passing through the atmosphere, the signal is attenuated and noise is introduced into the signal.

When radiation passes through a gas with molecules of a certain size, some of the radiation particles will hit the molecules and be scattered; i.e. deflected from their original paths. For some gases, radiation incident on the molecules will be absorbed and re-emitted: usually in a different direction and possibly with a different frequency. These two processes: absorption and scattering help to explain both the reduced signal strength and the background noise in thermal images.

The radiation incident on the detector will comprise not only of radiation from the radiating body, but also of scattered radiation from other bodies radiating large amounts of energy (e.g. the sun) and radiation emitted by particles in the atmospheric path between the source and detector. These additional sources of radiation will appear at the detector in the form of noise superimposed on the required signal.

In the lower atmosphere, those molecules which are the primary absorbers are (in order of precedence): Water, Carbon Dioxide, Ozone and Methane. Scattering is caused by molecular clusters, water (in solid, liquid and gaseous forms); and

particle suspensions such as fog, smoke or haze.

The amount of absorption of infrared radiation by each type of molecule is a function of the wavelength of the radiation. Absorption over much of the infrared part of the spectrum effectively blocks radiation at those wavelengths leaving only "atmospheric windows" through which radiation can pass. The two windows which are most important in thermal imaging are: $3.5-5\mu$ and $8-14\mu$. Thermal detectors are therefore designed to detect radiation at these frequencies.

The actual transmittance of radiation for a path length x is given by $e^{-(\alpha+\beta)x}$ where α and β are absorption coefficient and scattering coefficient respectively and depend on wavelength. The sum of the absorption and scattering coefficients is termed the extinction coefficient. The expression $\tau_A(\lambda) = exp(-\gamma(\lambda)x)$ is known as the Lambert-Beer la where τ_A is the atmospheric transmission and γ is the extinction coefficient.

The type of scattering that takes place depends on the size of particle and the wavelength of the radiation incident on it. For a particle with diameter $a \ll \lambda$, Rayleigh scattering occurs; where the scattering coefficient is proportional to λ^{-4}. Shorter wavelengths are affected by Rayleigh scattering far more than longer wavelengths. If $a \simeq \lambda$, Mie scattering occurs. For this type of scattering, the amount of scattering varies considerably with wavelength. Tables of Mie scattering data have been printed: some references can be found in [96]. If $a \gg \lambda$, Nonselective scattering occurs. Here, the scattering coefficient is almost independent of

wavelength. This type of scattering occurs in fog and cloud, where the size of water droplets varies between 5 and 100μ - often larger than the wavelength of the radiation. In fog or haze, scattering is the dominant factor in the reduction of transmission. A selection of particle sizes can be found in various places in the literature, e.g. [86].

Because the majority of the work in this thesis concerns thermal images taken at short ranges (typically one or two metres), atmospheric attenuation does not greatly affect the images and has been ignored. Consequently, this section is intended only as a cursory introduction to the problems faced when imaging at longer ranges; and to illustrate the reasons why thermal detectors only detect in two wavebands. The results presented in this thesis and the conclusions drawn from them are not invalidated by this assumption, however. Images at longer ranges will contain more noise and the signals will not be as strong, but we would still expect to see the same patterns in objects viewed at long ranges as at short ranges.

A more complete description of the relative effects of the absorption due to different atmospheric constituents is given in [31]. Diagrams showing atmospheric transmission as a function of wavelength for a variety of path lengths have been produced. Examples can be found in [96] or [47].

Chapter 3

Infrared Imaging Devices

3.1 Introduction

The first device for recording data of infrared radiation was developed initially by Sir John Herschel, son of Sir William Herschel (the discoverer of infrared). Much later, in the 1930s, it reappeared as the "evaporograph". The first sensitive *photon* detector, however, was not developed until 1920, and was capable of detecting only in the near infrared. The potential of infrared for detection of military targets such as ships, tanks and aircraft spurred large investment into infrared research during World War II and led to the development of the first infrared viewing systems. Since then, the development of improved extrinsic semiconductor materials and advances in electronics have led to infrared imaging devices with ever improving clarity and resolution.

In this section, the types of detector currently in use will be summarised. A brief description will be presented of how each type of detector operates. A detailed analysis of the theory and design of infrared detectors is beyond the

scope of this thesis: for such an analysis, refer to [42] or [83].

There are two classes of infrared detectors: photon detectors which detect incident photons; and thermal detectors which change their electrical properties when incident radiation causes their temperature to change. Both types of detectors have advantages and disadvantages making them less or more suitable for a particular application. A brief description of the theory of each is presented below.

3.2 Photon Detectors

A photon detector is sometimes referred to as a quantum detector. This is because it detects quanta of energy in the form of radiation. To explain how this occurs, we use a little simple semiconductor theory.

Electrons in a semiconductor lattice exist in discrete energy levels. To move an electron from one energy level to a higher energy level requires an input of energy; and similarly, an electron moving from a higher to a lower energy band will emit a packet of energy.

The outermost filled atomic level is known as the valence band. Electrons in this band are "tied" to a particular atom in the lattice. Electrons moving into the next energy level, the conduction band, however, are free to move about the lattice. When an electron moves into the conduction band, it leaves behind a positively charged "hole", which is also free to move within the lattice. These free electrons and holes (which can be considered as positively charged quasi-particles)

contribute to the conductivity of the material. The conductivity of the material can be increased therefore, by giving extra energy to the valence band electrons to move them into the conduction band. This energy can come from a photon incident on the material.

When a sufficiently energetic photon is incident on a piece of *photoconductive* material, an electron is released; thereby increasing the conductivity of the material. The amount of energy needed to release an electron is called the work function, ϕ. The work function is equal to the difference in energy between an electron in the valence band and one in the conduction band; and is a function of the material. If an incoming photon is to release an electron then its energy must exceed ϕ. The relationship $E = h\nu$ gives us an upper bound on the wavelength of radiation which will liberate an electron. It is for this reason that quantum detectors have a rigorously defined spectral cutoff response.

By doping a pure semiconductor with impurities, energy levels can be created between the conduction band and the valence band. A semiconductor doped with impurities causing an energy level to exist near to the conduction band is called an n-type semiconductor. The effect of this doping will be to create a "donor" band of electrons which can be promoted into the conduction band with a far smaller input of energy; leaving bound holes in the donor band. Another type of impurities causes an energy band near the valence band which can accept valence electrons; leaving positively charged free holes in the valence band. Semiconductors with

doping causing an "acceptor" band are called p-type semiconductors. Since the energy required to free a donor electron or to free a hole in p-type semiconductor is much smaller than in the intrinsic semiconductor, suitably doped semiconductor material can detect incoming radiation at much longer wavelengths.

Compounds of two elements (binary compounds) can display much more useful properties than pure compounds. Furthermore, by mixing two binary compounds with a common element, a ternary compound can be created with properties dependent on the ratio of the mixture. Ternary compounds can be manufactured which have very small band gaps and can consequently operate at comparatively long wavelengths. One such ternary compound used a great deal in thermal imaging is Cadmium-Mercury-Telluride (CdHgTe). By choosing the proportions of CdTe and HgTe, the response of the detector can be tailored to fit a required application.

One consequence of decreasing the energy gap required to liberate an electron is that the energy required can now be present as thermal energy in the lattice itself; in which case the "signal" due to incoming photons can become less obvious in the "noise" caused by self emission. To reduce the number of thermally emitted electrons, the detectors must be cooled - usually to the temperature of liquid nitrogen (77K).

Imagers which make use of the photoconductive nature of semiconductors are called *photoconductive detectors*. A simple photoconductive detector can be built

by connecting a battery to a piece of photoconductive material and monitoring the current flowing in the circuit. Increasing the voltage across the material causes a build up of electrons (or holes) on one of the contacts of the material and hence increases the recombination rate; limiting the output signal. To increase the lifetime of an electron, and therefore increase the signal, it is necessary to trap the excess holes. Unfortunately, this will also make the device slower. There is therefore a trade off between the magnitude and the speed of response.

The performance of a photoconductive detector is determined by its *quantum efficiency*. This is defined as the number of electrons which will flow for one incident photon. A typical value might be 0.25. The speed of response is determined by the time constant of the device; which depends on the product of the resistance and the capacitance of the detector and associated circuitry.

If a piece of n-type semiconductor is connected to a piece of p-type semiconductor, electrons diffuse across the junction, creating a region depleted of spare electrons and holes and a large electric field. When radiant energy falls on the junction, electrons and holes are excited and are accelerated across the junction. If a circuit is connected, a current will flow. This forms the basis for a second type of photon detector: a *photovoltaic detector*. Because the resistance and capacitance of the junction are high, the time constant is higher than in the photoconductive detector.

If a voltage is applied across the junction, the width of the depletion region will

increase. Carriers accelerated across the junction will now have more energy and will be capable of releasing more carriers and consequently increasing the current that will flow. A detector designed in this way is called an *avalanche photodiode.* The amount of multiplication of current is limited, however, because increasing the signal current also causes an increase in randomly generated currents; which appear as noise in the output of the detector. Applying the voltage decreases the capacitance: reducing the time constant. Avalanche photodiodes can therefore be both fast and sensitive devices. A lack of suitable materials for manufacturing these devices currently limits the wavelength range over which they can be used.

A further type of photon detector is the *photo-electro-magnetic detector.* Here, a magnetic field across the detector causes the free electrons to deflect to the sides of the device, setting up a lateral current across the device.

In a *vacuum photocell,* the target for the photons is a negatively charged cathode. The electron freed by the photon is accelerated towards an anode across a vacuum. If the electron frees several more electrons when it strikes the anode; and these electrons are accelerated towards an electrode with a higher potential, freeing more electrons when they strike it; we have a rapid multiplication of electrons. A device built in this way is called a *photomultiplier tube*; the electrodes are called dynodes.

For more details of photon detectors, refer to [48] or [98].

3.3 Thermal Detectors

In a thermal detector, incident radiation causes a change in the temperature of the detecting material. By determining the rise in temperature of the material, it is possible to calculate the amount of energy incident on the material and hence predict the strength of the target emitting the energy. In general, some property of the material provides a measure of the change in its temperature. The magnitude of this change determines the responsivity of the device.

Before examining particular types of devices, it is useful to consider the thermal equation of a general device.

Consider a thermal mass H attached to a heat sink of temperature T_S via a thermal conductance G. In the absence of incident radiation, the temperature of the detector, T, will equal $T_S \pm n$, where n is a noise fluctuation. This noise fluctuation, a *temperature noise* in the detector, limits the size of signal that can unambiguously be detected. If the fraction η of the incident radiation I falling on the detector is absorbed and causes a temperature rise $\theta = T - T_S$ then:

$$\eta I = H\frac{d\theta}{dt} + G\theta.$$

Consider incident radiation $I = I_0 + I_\omega \exp j\omega t$. Solving for θ in the above equation for a frequency ω gives θ with amplitude $\theta_\omega = \eta I_\omega (G^2 + \omega^2 H^2)^{-\frac{1}{2}}$ and phase $\phi = \arctan \frac{\omega H}{G}$. It is clear that for low frequencies, the magnitude of the signals will be larger if G is small. For $\omega H \ll G$, the amplitude of response is approximately constant with frequency; while at higher frequencies, the amplitude

will be inversely proportional to frequency. The thermal time constant is defined as $\tau_T = \frac{H}{G}$. Reducing the size of G to increase amplitude of response will also have the effect of increasing the time constant. Since the mass, H, is limited by the fragility of smaller detectors, and G cannot be vanishingly small, the operating limits of the device are determined by its physical properties.

The responsivity of the device will not reach the theoretical maximum because of several other factors. In general, $\eta \neq 1$ and losses in the signal occur because of reflection and absorption in the optics and protective windows and consequently not all the incident radiation is absorbed. The detector must be supported; and if a signal is to be read, electrical contacts must be made. The supports and contacts also conduct heat, increasing G and introducing a source of incoming heat additional to the radiation. Noise, ubiquitous in electronic circuits affects all types of thermal detectors; the most important sources for most types of thermal detector being Johnson noise, thermal noise and amplifier noise.

3.3.1 Thermal Detectors Measuring Change in Temperature

The first thermal imagers, as distinct from thermal detectors, used absolute material properties to determine the temperature of the material. Detectors using the rate of change of temperature as a measure of incident radiation will be discussed in the next section. In this section, detectors measuring absolute temperature change will be considered: pneumatic cells, bolometers and thermojunctions.

If an ideal gas is contained within a sealed container kept at constant pressure, then from the gas laws, the volume of the gas will be proportional to its temperature. This effect is the basis of a type of detectors called pneumatic cells. The most well-known usage of this theory is the *Golay Cell* first demonstrated by Golay in 1947. In the Golay cell, the expanding gas moves a silvered diaphragm. Optical methods are used to detect small changes in the position of the diaphragm and hence determine the extent of expansion in the gas. The absorption of incident radiation can be through blackened absorbing surfaces which transmit energy to the gas. Alternatively, a window can be put in the wall of the container so that the incident radiation heats the gas directly. In this latter approach, the properties of the gas determine the spectral response of the device. Golay cells have been used in this way to analyse gases. Golay cells are sensitive devices, but are fragile and have a slow response time. To avoid pressure changes affecting the detector output, the cell is sometimes given a small leak to the environment. This also limits the response of the device, since very slow changes will leak away before they are detected. Because of their fragility and sensitivity to vibration, Golay cells are only useful in laboratory environments.

In a *bolometer*, the detecting material has a resistance that is temperature dependent. The effect was first noticed by Langley in 1880 in certain metals. The design usually has one piece of material open to radiation, and one shielded - to account for effects due to changes in the ambient temperature. Initially, all

bolometers were constructed using metals, but more recently, since their development by Brattain and Becker at Bell Laboratories in 1946, semiconductors have been used. The semiconductor bolometers, or thermistor bolometers, have much higher responsivities with comparable amounts of noise. Cooling of the semiconductor bolometers greatly enhances their performance. At similar temperatures to those at which photon detectors work, bolometers have a comparable responsivity; although they cannot operate at such high frequencies. Because they have a wide spectral response, they are sometimes used in preference; when speed of response is not an important factor. Superconducting bolometers have also been investigated but the critical temperature control which they require limits their usefulness. Bolometers have been designed that rely on the change of state of a material to give a dramatic change in the output when the state change occurs. Examples include changes in the form of a crystal lattice, or the transition from superconducting to normal conducting mode. These devices must be accurately designed, however, and have not found much support.

To detect the change in the resistance of the material, a current must be passed through the material. This current will heat the material further. If the material has a negative temperature coefficient, the increase in temperature due to the current heating will cause a further drop in resistance. If the material has a constant voltage across it, the decrease in resistance will cause more current to flow and hence more heating of the material and a further decrease in resistance. This

effect, known as "thermal run away", is one limiting factor in the performance of bolometers. As with other imaging devices, noise is also a limiting factor. The main source of noise in bolometers is usually Johnson noise. Putley [74] discusses noise in bolometers in more detail.

If two different electrical conductors are connected in a loop; the two junctions being kept at different temperatures, an emf will appear. This *thermovoltaic* effect, first reported by Seebeck in 1822, is the effect which causes *thermojunctions* to work. In a thermojunction, one of the junctions is kept at a fixed temperature, the other is exposed to radiation. The radiation warms the junction and causes the emf in the loop to change. The value of the emf is a measure of the intensity of the radiation. The output of the thermojunctions is small, but Melloni noticed in 1835 that thermojunctions can be chained together, amplifying the output. This was the first *thermopile*. A device which produces a large change in voltage per degree change in junction temperature is said to have a high *thermoelectric* or *Seebeck* coefficient. The ideal thermojunction would have a high Seebeck coefficient, a high electrical conductivity to minimize Joulean heat loss and a small thermal conductivity to minimize the heat conduction between the junctions. For a fast response time, a small thermal mass is also necessary. No single material fulfills all these requirements. Some semiconductor materials have a very high performance, but making fine wires is impracticable; so metal thermopiles are still sometimes used.

3.3.2 Thermal Detectors Measuring Rate of Change of Temperature

This section deals with detectors which have a constant value output until a change in the temperature of the detector material occurs. The theory of these *pyroelectric detectors* and some of the reasons which have caused them to be one of the most important detectors today will be described.

Pyroelectric materials, possessing a low degree of symmetry, were first investigated shortly before the Second World War. These materials have an internal dipole moment which causes an external charge distribution to appear near the surface of the crystals. The magnitude of the dipole and hence the external charge is temperature dependent; but due to leakage, the charge lasts for only a brief period. When a circuit is connected across the material and the temperature of the material is changed, a burst of current appears. The magnitude of this current, or the magnitude of the voltage in open circuit mode, can be used to indicate the magnitude of the change in temperature.

To use pyroelectric detectors to image stationary scenes, the detectors are chopped. A chopper blade alternately opens the detector element to the scene and to a chopper at a reference temperature. The speed of chopper rotation is chosen carefully, considering the thermal and electrical time constants of the device and its electronics, and the changes in surface charge are measured. If the electrical time constant is longer than the thermal time constant, a chopper speed can be chosen such that in each open-close cycle of the chopper, the surface charge

distribution of the detector can be measured, indicating the rise in temperature of the material before leakage currents attenuate the signal.

The pyroelectric crystal is sliced into thin slithers for improved performance. Metal contacts are evaporated onto the pyroelectric, creating a capacitor from which the signal value is read. Modern multiple element detectors must then be connected to a silicon multiplexer. This is done by "bump bonding" the pyroelectric to the multiplexer. This *hybrid* technology has its critics, but modern techniques allow accurate positioning of molten lead/tin solder bumps (or cold soldering of indium bumps) to very high tolerances: small elemental pitches, approaching the minimum "reasonable" [91] pitch can already be confidently manufactured.

Porter shows [72] that the current responsivity of pyroelectric devices increases for low frequencies and remains constant for higher frequencies. He also shows that the voltage responsivity increases with frequency for low frequencies; is constant for frequencies $\frac{1}{\tau_T} < \omega < \frac{1}{\tau_E}$ where τ_T is the thermal time constant and τ_E is the electrical time constant; and falls with frequency for frequencies greater than $\frac{1}{\tau_E}$. The chopping frequency is chosen to maximise the data rate, while keeping a high detectivity. Putley [74] suggests that for most applications, it is reasonable to use the high frequency approximation for responsivity.

The detectivity of a pyroelectric device is expressed as the *pyroelectric coefficient.* Wright [98] defines the pyroelectric coefficient, Γ, as "The variation of

charge on a unit area of crystal when the temperature changes by one absolute degree". The value of the pyroelectric coefficient, which is more commonly denoted p, increases with temperature until the temperature approaches the *Curie* temperature; at which point it drops to zero.

Most pyroelectric materials are also ferroelectric materials; i.e. their polarisation can be reversed by applying a suitable electric field. The applied electric field is used to orient the separate domains of differently polarised material.

Noise is a severe problem in pyroelectric devices. This is due in part to the fact that they work at comparatively high temperatures. Putley suggests that the most important type of noise in pyroelectric detectors is Johnson noise. Astheimer and Schwartz [3] pointed out, however, that there are limits on the improvements which can be made in detectivity: ultimately, the unavoidable temperature noise will be the dominant noise source.

In addition to noise, pyroelectric detector arrays have a problem not faced by the majority of infrared imaging devices: that of microphony. The pyroelectric materials have piezoelectric properties which cause them to produce electrical signals when the pyroelectric material is strained. This strain could easily be produced if the imager were moved and the detector material were not mounted on a microphony-free mount. As typical applications for pyroelectric imagers involve the use of aerial and hand held imagers, the problem must be addressed. One solution which appears to eliminate microphony problems is *reticulation*: cutting

the detector material into discrete elements. Nowadays, this can be done to high accuracy using laser technology [91]. Reticulating the detector material has the added benefit of cutting down the amount of lateral diffusion of heat between detector elements and hence reducing the smoothing of image signal between elements.

3.4 Current Trends in Infrared Technology

Economic factors are determining the direction and amount of research conducted in infrared imaging. Historically, the majority of funding for the development of infrared detectors has been provided by the Military both in Britain and abroad. Recent easing of East-West relations means that the amount of funding for defence and defence related research may be reviewed. It is too early to say what form defence cuts will take. It may be that quantity is sacrificed, but quality retained in which case research would be largely unaffected. Baynham [6] argues that the future of infrared imaging is assured: military needs will remain important in surveillance roles. He also suggests that work for civil customers will become increasingly important, citing the automobile industry as one possible customer. He recognises that detectors designed for military use will not necessarily be suitable for use in civil projects and that market oriented research must be undertaken. The requirements of these projects will probably be different from the requirements of the military. High resolution may be a less important factor than low unit cost, simplicity of use or ease of installation. These factors will effect the

course of research done in this field and have been mentioned to justify some of the conclusions that are reached in this section.

There is no doubt that the quality of visible cameras is vastly superior to that of infrared devices: they are smaller, lighter and cheaper. Visible cameras have been mass produced bringing costs down. Costs could be reduced in infrared imaging devices if the devices were produced in large numbers, though perhaps not as low as the costs of visible cameras. The size of infrared devices cannot, however, be reduced to the size of visible cameras. The infrared devices image at longer wavelengths, so the optics must be larger and the infrared cameras must be bigger and heavier. Elliot [26] points out that detector sizes are reducing, larger arrays are replacing mechanical scanned systems and operating temperatures are increasing so simpler, more efficient cooling systems can be used. As mentioned previously, the improvement in performance in infrared devices has led to an increase in the potential market. The types of infrared detector which satisfy the most of the market's requirements will be the ones which will dominate in the coming decades.

Currently, the detector types which are receiving the most interest are Cadmium Mercury Telluride (CdHgTe or CMT); pyroelectric; and Platinum Silicide. Each satisfies some of the requirements of some of the potential markets. The leader in high resolution imagery at the moment, particularly for longer wavelengths, is CMT: a type of intrinsic photon detector. Large arrays have been

produced, though it is thought that with present technology, array size is limited by the size of material which can be produced (about 1cm square) to arrays of 256 by 256. Larger arrays can be produced by assembling more than one slice of detector material but this is a complicated process.

Devices made from Platinum Silicide, a photoemissive compound, have the advantage that they are compatible with silicon circuitry. It is possible therefore, that devices can be built with parallel computation done "on chip". Platinum Silicide devices have good detector uniformity but low quantum efficiency; and must perform under very stringent cooling requirements, particularly for detection at longer wavelengths. Research being conducted into techniques for growing CMT onto substrates compatible with silicon circuits, if successful, could cause Platinum Silicide devices lose their place in the market.

Thermal detectors have the important advantage that they can be produced at a much lower cost than any of the cooled detectors. The Pyroelectric Vidicon (see [88]) has for the last decade been an important tool for the rescue services. It has paved the way for future sales of pyroelectric devices. The vidicon has reached a mass market and proved that low-cost thermal imaging is feasible. It seems likely that two dimensional arrays of pyroelectric detectors with increasing size will lead the market for low-cost infrared imaging. The current performance of pyroelectric devices is inferior to that of some cooled devices however, so pyroelectric devices are currently not challenging CMT arrays in high resolution imaging.

Material improvements are not the only developments leading to increased performance in infrared imagers. The development of microscan has led to improved resolution without the need for increased array size. Microscan pans the detector array by a sub-pixel amount so that each field of the image is taken slightly offset from the first. In a two-by-two microscan, three additional fields are taken per frame; each being offset by half a pixel width in the horizontal and/or the vertical direction from the first pixel. When displaying the image, the data from all fields are displayed together, the positioning determined by the relationship of the fields in the microscan. In this way, a 64 square element detector with two-by-two microscan will give an image size of 128 square. A three-by-three microscan will give an image size of 192 square and so on. Normally microscan other than two-by-two or three-by-three will not be used as the improvements gained from increasing the level of microscan over three are not worth the additional complexity. Microscan does improve the appearance of images produced by small arrays. The problem of aliasing which is common in small arrays is reduced when microscan is used. In the short to medium term, microscan appears to be a useful aid for low resolution imagers. It requires additional circuitry to run it, however, and re-introduces a mechanical scanning technique. It is therefore unlikely to be used once larger arrays can be produced.

Chapter 4

Imager design

4.1 Introduction

The aims of the CASE studentship were listed in the introduction. In order to fulfil those aims, we needed to examine a variety of infrared images. We could either use images from currently available imagers at RSRE or we could build our own imager.

We chose to build our own thermal imager for a number of reasons. Perhaps the most important reason was that the line scan array imager available at RSRE would have required substantial modifications to allow a raw signal to be fed into a computer. That imager is a complete system, which processes the raw signal and produces a display on a monitor. Transferring this data to a computer introduced rather serious problems regarding pixel identification; and a large quantity of noise to the image.

Also, an important part of the DPhil was the examination of the effect of a variety of signal processing filters on the raw data rather than just the finished

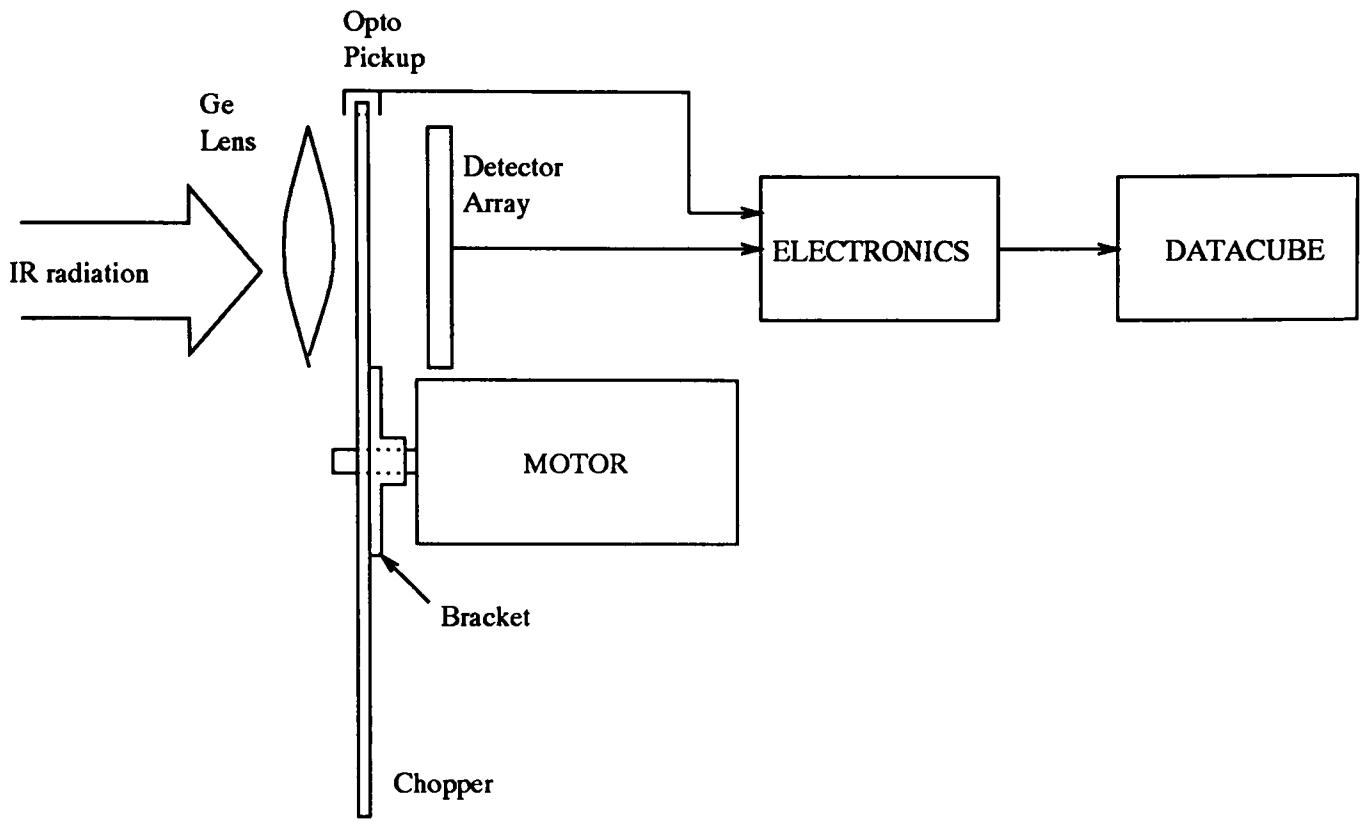


Figure 4.1: Schematic diagram of thermal imager

picture to determine whether the pre-processing of the data for the RSRE imager could be improved upon to yield a clearer image. To do this meant being able to conduct a variety of experiments taking images of deliberately chosen scenes. This necessitated ready access to an imager: further encouraging the construction of one for use in Oxford.

4.2 Overview of Imager

A schematic diagram of the imager we produced, based on the imager designed at RSRE is given in figure 4.1. A Germanium lens is used to focus the infrared radiation through a rotating focal plane chopper onto an array of pyroelectric detectors. The need for a chopper was explained in chapter 3; choppers will be

described in more detail in section 4.3. The chopper is driven by a DC motor. An opto-pickup straddling the chopper blade gives position feedback for the chopper and enables the motor speed to be phase locked with a master oscillator. The master oscillator also sets the speed for data collection and drives the clock to the electronics.

A logic circuit determines when data is recorded, based on the reference signal from the opto-pickup. The data is collected both during the open and closed fields of the chopper; multiplexed; digitised; and sent (along with relevant timing signals) to the digital input of a Maxscan board of a Datacube image processing system. The Datacube is controlled by a Sun host and is preprogrammed to read one frame of data, where a "frame" is a number of columns of data, selected a priori. A column consists of 64 data items in each of the open and closed chopper positions. Typically, a frame will be between 100 and 500 columns.

A polished aluminium mirror is situated immediately in front of the lens. The mirror is rotated as data is collected, thereby giving a second dimension to the image.

A diagrammatic overview of the electronics in the imager is given in figure 4.2.

4.3 Chopper

As pointed out in chapter 3, elements of the pyroelectric array respond only to *changes* in thermal irradiation. The function of the chopper therefore is to create

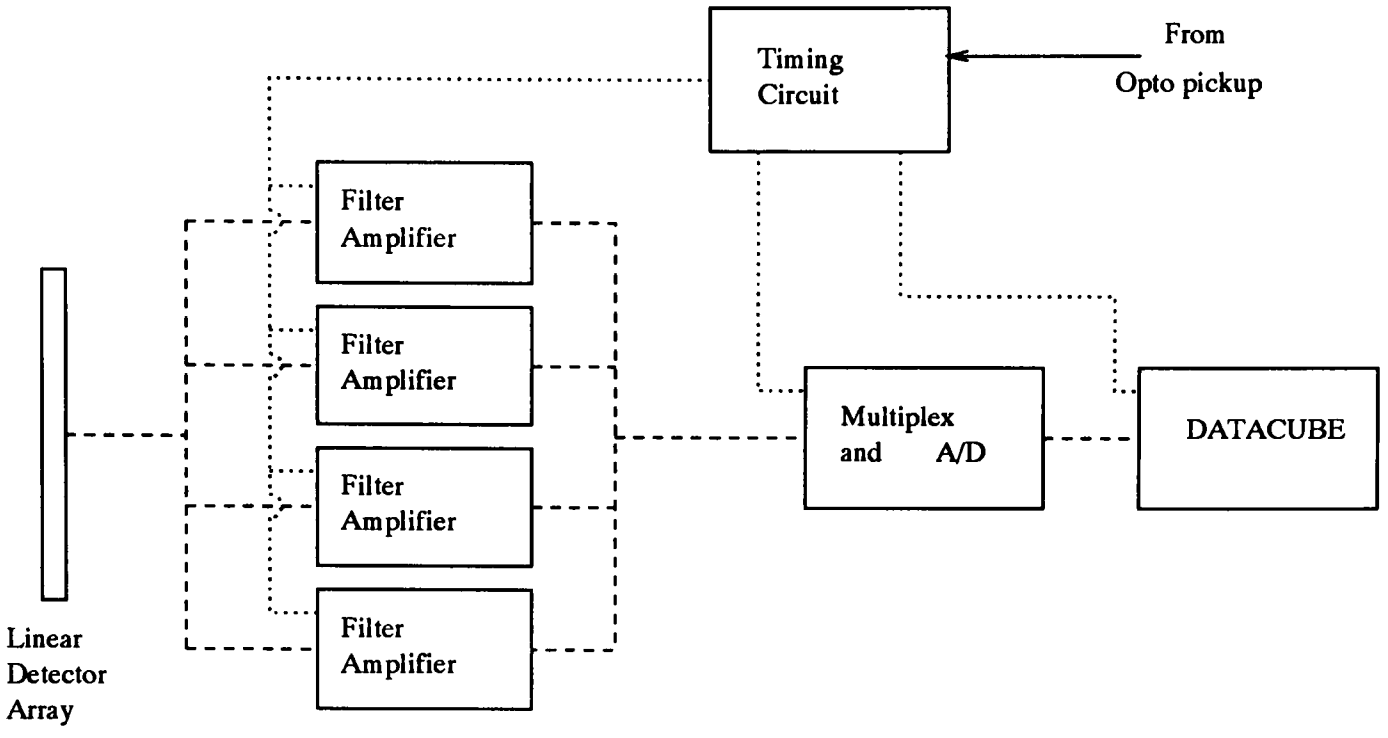


Figure 4.2: Electronics overview

a temperature difference between the imaged scene and a fixed (chopper) temperature. Without a chopper, the thermal imager would be capable of detecting changes in the scene it imaged and a stationary array might therefore be used for detecting an event, such as a car or person passing. With a chopper, however, the thermal detector is capable of imaging a stationary scene and consequently can provide far more (potentially) useful information.

A chopper can be anything which alternately "opens" and "closes" the detector array; that is allows thermal radiation, or prevents thermal radiation from falling on it.

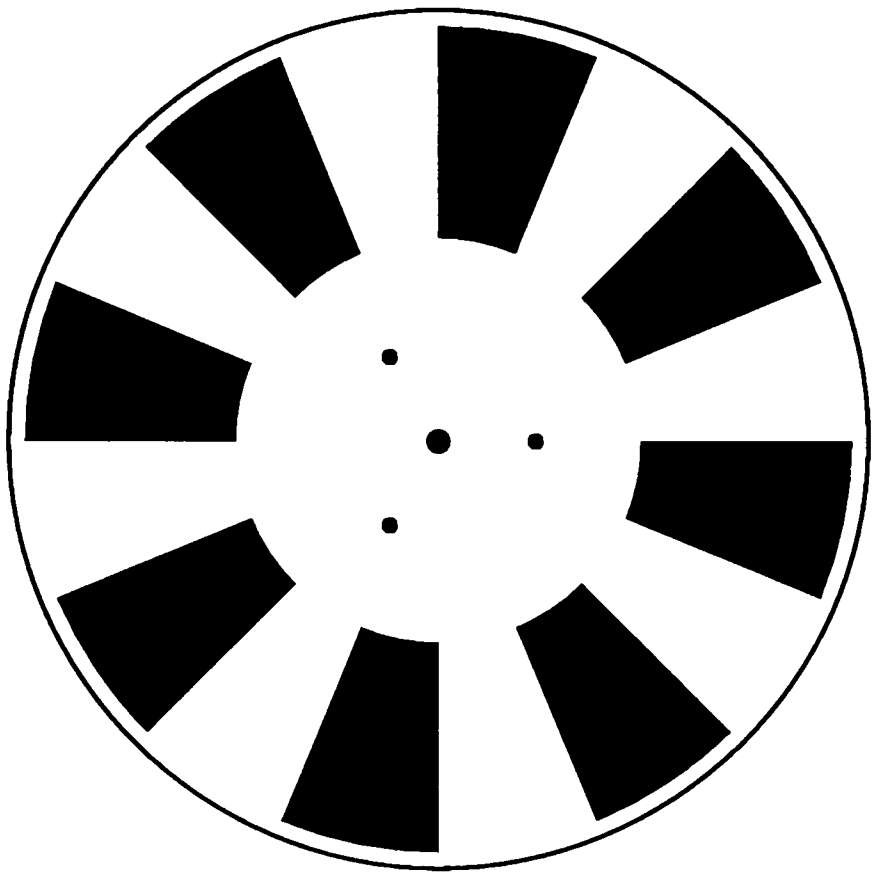
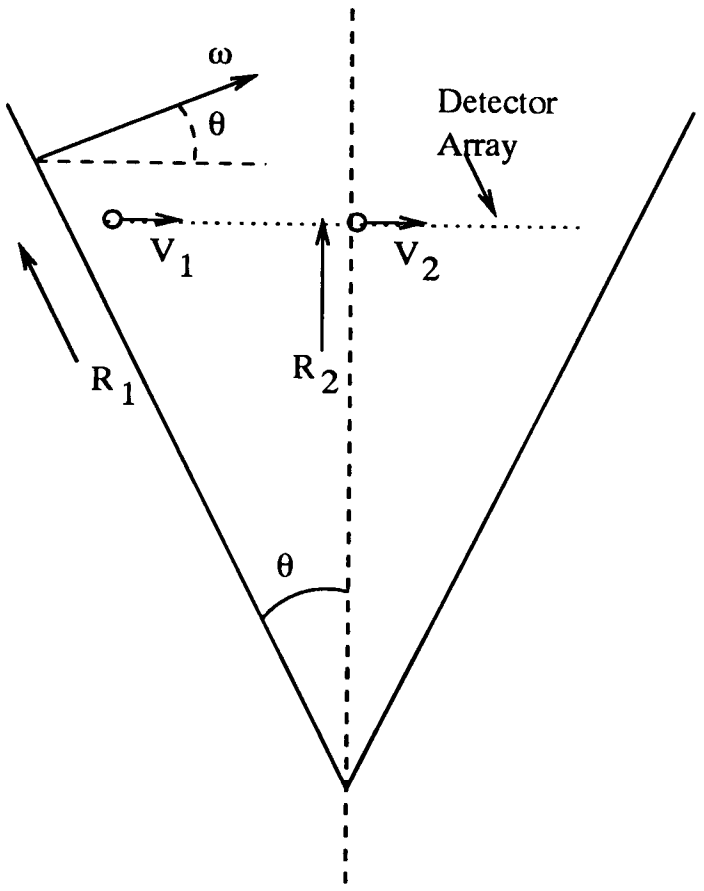


Figure 4.3: Design of chopper blade

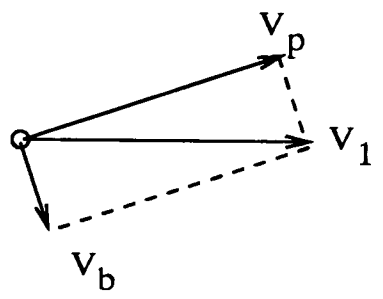
4.3.1 Chopper Blade

Several designs of chopper are possible. For this application, where one of the primary concerns is the need to adjust the system, a simple chopper design was used (see figure 4.3). The detector array is placed on a line approximately tangential to the rotating chopper at a distance r from the centre of rotation.

As the chopper rotates, it covers each of the array elements in turn. The speed at which the elements are covered is assumed constant. The error in that assumption can be calculated: consider a chopper blade passing across the detector array with angular frequency ω (see figure 4.4). From the diagram, $V_p = R_1\omega$ and $V_1 = \frac{R_1\omega}{cos\theta}$. Also, $V_2 = \omega R_2$ since at this point, the chopper is passing across the detector elements perpendicular to it. But, as $cos\theta = \frac{R_2}{R_1}$ then $V_1 = \frac{R_2\omega}{cos^2\theta} = \frac{V_2}{cos^2\theta}$.



(a)



(b)

Figure 4.4: (a) chopper blade passing across the detector array. (b) velocity diagram for point of blade passing across array.

The difference between the velocity at which the blade crosses the first element of the array and the velocity at which the blade crosses the central element of the array is therefore $V_1 - V_2 = V_2\left(\frac{1}{cos^2\theta} - 1\right) = V_2 tan^2\theta$. As the chopper has eight blades, the angle $\theta = \frac{\pi}{16}$. The percentage difference between the maximum and minimum velocities is therefore given by:

$$tan^2\frac{\pi}{16} \times 100 < 4\%.$$

The distance of the array from the centre of rotation can be varied. This could be useful for two reasons. Firstly, the lengths of different pyroelectric arrays can vary so if the array needed to be replaced, this could be accommodated. Secondly, it is required in this application to have a small period of settling time after the chopper blade has passed over the detector element. The duration of this settling time can be varied by moving the detector array along a line radial to the chopper (and therefore increasing or decreasing the relative proportion of the rotation time that the chopper edge is crossing the detector array). More will be said about this second point in section 4.5.

The size of the chopper was designed to accommodate all probable detector array sizes. The outer dimension was designed to be twenty five percent larger than the largest probable detector size: twenty five percent being the amount of blanking we required. We used a Pafec CAD system to design the chopper blade. The output plot was photographed and the negative was used as a template in a metal-etching process to produce the chopper blade. The final blade was not

ideal, due to innate problems with the metal etching technique, but it was perfectly adequate due to our circumspect design criteria.

The chopper was attached to the motor shaft with a light aluminium bracket.

4.3.2 Driving the Chopper

A 12 volt dc motor was used to drive the chopper. A drive circuit was designed to enable the motor to be actively speeded up and slowed down. Using this circuit it could be controlled to reach the required speed very quickly. The circuit required a quad op-amp: one amplifier to define a mid-rail voltage; one to drive one terminal of the motor from this mid-rail voltage and one to drive the other terminal of the motor, using the output of the phase-locked-loop as its input (see section 4.5). The extra op-amp was used to buffer the signal from the opto-pickup to the phase-locked-loop input - thereby saving circuit space.

The motor would not perform correctly, however, as the op-amps were only capable of supplying twelve milliamps - insufficient current to accelerate the motor from rest and only barely sufficient to keep it rotating at the required frequency. Two high current motor drivers were incorporated into the electronics to drive the motor, amplifying the current output from the op-amps.

4.4 Path of Data

4.4.1 Pre-amplifier Boards

The signals coming from the thermal array are first filtered to remove any DC components and high frequency noise. Each of the signals is filtered separately before being multiplexed and therefore 64 identical filters needed to be built.

The response of a pyroelectric detector falls with chopper frequency for frequencies greater than about twenty Hertz. The frequency response of the pre-amplifiers was designed to compensate for this and to give an output which is approximately constant regardless of chopper frequency. Figure 4.5 shows a Bode diagram which defines the required frequency response of the pre-amplifiers. The range of frequencies $20Hz - 150Hz$ covered all the frequencies at which we might possibly wish to run the detector. The drop-off in frequency for frequencies greater than $150Hz$ was designed to reduce the high frequency noise in the signal. The outputs of the detectors are small and therefore susceptible to noise. The amplification in the pre-amplifiers reduces this susceptibility. The amplification is done in two parts because it is desirable to have control over the magnitude of the gain of all the detectors. If a second stage is included after the multiplexor, the gains of the detectors can all be altered at the same time.

The bandpass filter for each of the thermal signals (as shown in figure 4.6) consisted of a standard op-amp, resistors and capacitors. The calculation of appropriate resistor and capacitor values is given in appendix A. Testing the filter

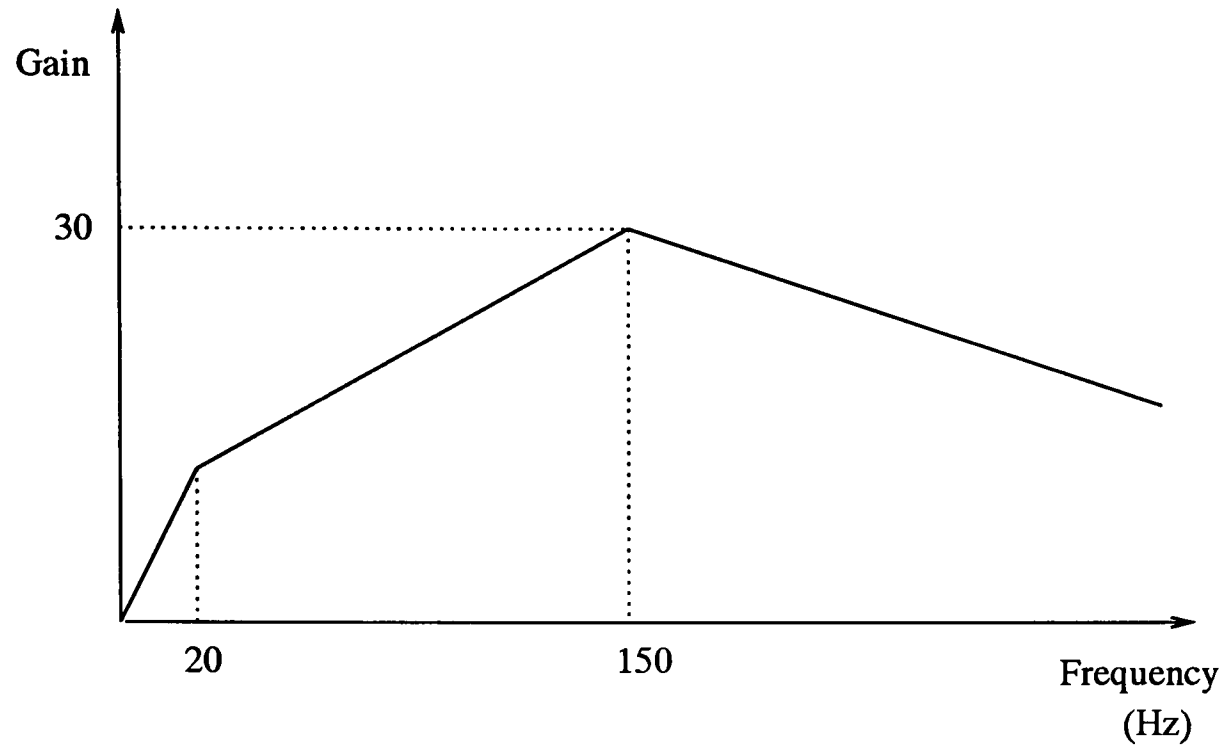


Figure 4.5: Desired Pre-amplifier Characteristics

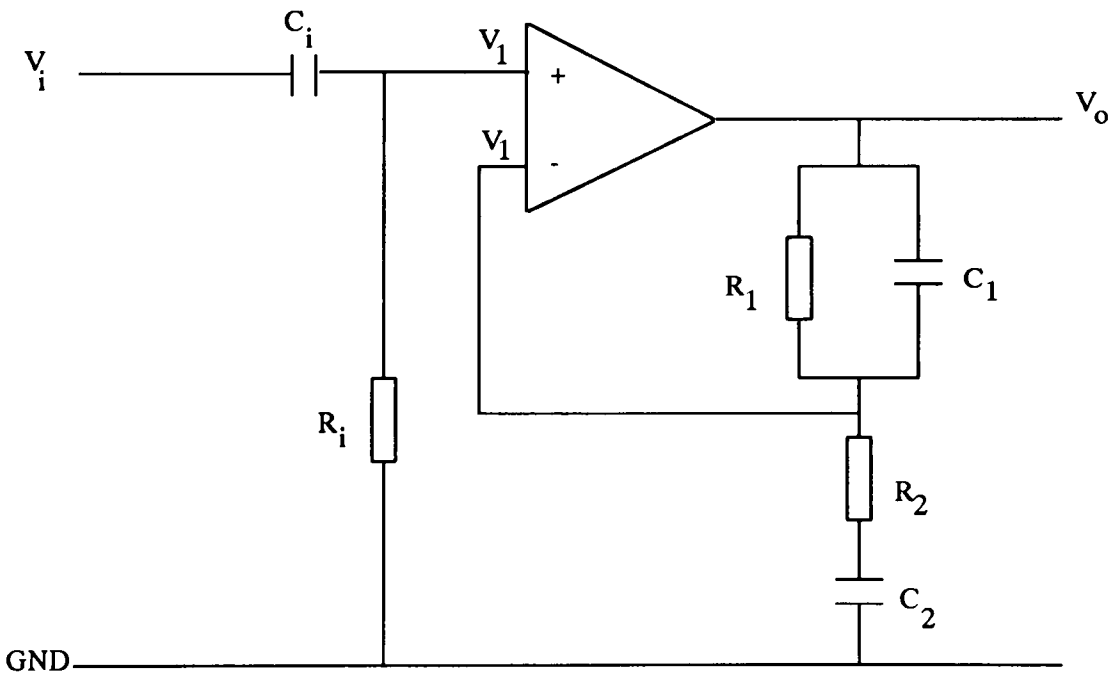


Figure 4.6: Pre-amplifier for one signal

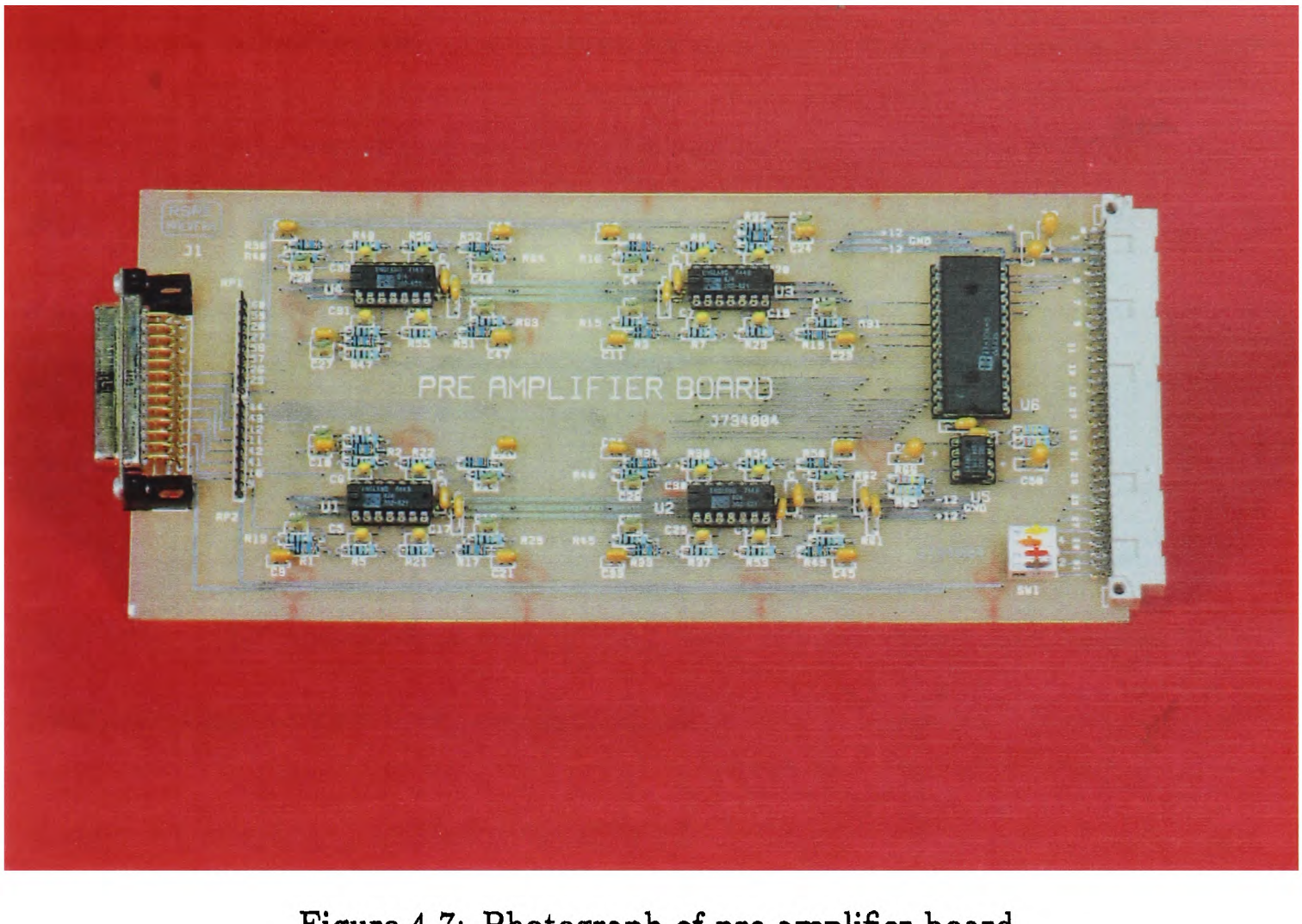


Figure 4.7: Photograph of pre-amplifier board

with a signal generator showed that the actual frequency response was very close to the desired response.

Once the basic design for one filter had been tested, the CAD package Futurenet was used to design four similar boards, each with sixteen bandpass filters, an analog multiplexor and associated inputs and outputs. The boards were produced in the CAD centre at RSRE.

A photograph of one of the pre-amplifier boards is given in Figure 4.7.

4.4.2 Digitisation

The sixty four pre-amplified data signals are multiplexed into a single stream of data. This multiplexed signal stream is then further amplified to give voltages which lie in the working range of an analog-to-digital converter. A sample and hold chip holds the voltage steady while it is sent to the analog-to-digital converter.

The twelve bit digital data produced by the analog-to-digital converter is then buffered to give a signal strong enough to be read by the Datacube (see section 4.6). The digital data and concomitant timing signals are passed to the Datacube over a thirty-four wire ribbon cable.

4.5 Timing control logic

4.5.1 Overview

Each of the elements in the data processing path must be clocked at the correct time for the data to be read properly. To achieve this, it is essential that they are all clocked by a signal derived from the same oscillator. The clocks for each of the ADC (analog-to-digital converter), the S/H (sample and hold) and the Datacube; and the enable signals for each of the multiplexors were all synchronized to a signal produced by a master oscillator. The speed of rotation of the chopper blade was phase-locked to a frequency derived from the frequency of the oscillator.

Figure 4.8 shows how the various timing units interacted.

Irregularities in the chopper blade, gyroscopic effects, or noise in the signal

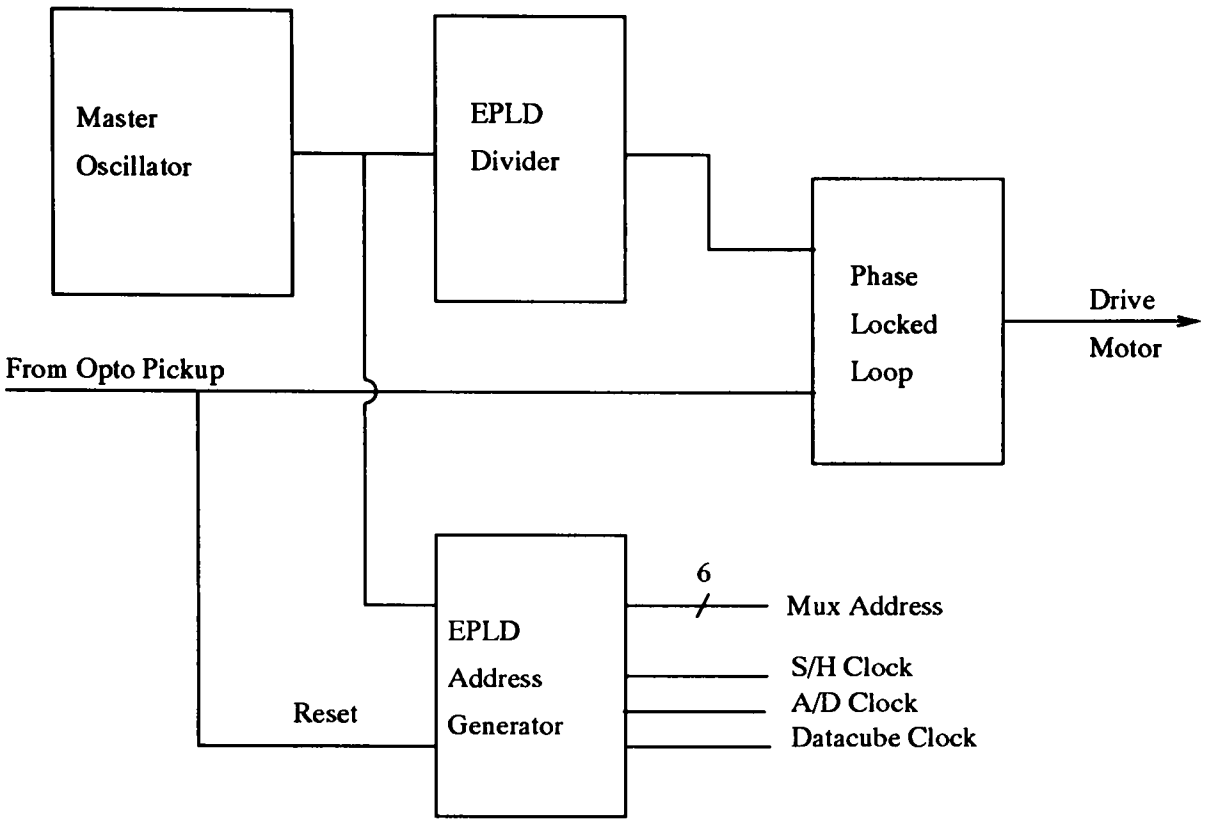


Figure 4.8: Overview of timing electronics

driving the motor could cause the motor speed to vary slightly during scanning. We therefore designed a short blanking period before and after reading the data for each line of the image.

To maximize the performance of the detector, each element should be read at a frequency of approximately 50Hz (see [50]). A blanking period of approximately 12% of the data time either side of reading the data was considered sufficient to overcome the problems mentioned above. We therefore required to read 80 data elements every $\frac{1}{50}$ second ($80 = 64 + 8 + 8$). For reasons explained in section 4.5.2, the oscillator needed to oscillate at eight times this frequency. As flexibility was an important consideration the early stages of construction, a master oscillator was designed with a variable frequency between 20kHz and 200kHz.

An electronic solution was initially sought for varying the position of the blanking, but this was later rejected in favour of a mechanical solution (see section 4.10).

4.5.2 Timing signals

The ADC, the S/H and the Datacube must be clocked at the correct times relative to each other to ensure that the data is processed correctly. The S/H must be given sufficient time to settle after taking a sample since the acquisition time is a fixed (and non-negligible) period; and the ADC must be allowed to complete its conversion before the data is sent to the Datacube. The start of the timing diagram describing this operation is given in figure 4.9.

Each time the S/H is activated, the correct element of the detector array must be read. This entails sending the digital value of 0, 1, 2, ..., 14, 15 to the multiplexors on the four pre-amplifier boards. A signal must also be sent to each of the four boards to enable them at the correct time and disable them for the remaining time to ensure that the 64 elements of the detector array are read in the correct sequence.

Since the S/H, the ADC and the Datacube must all be clocked each pixel time, the master oscillator must produce a frequency eight times the frequency of the pixel frequency as illustrated in figure 4.9.

To produce the required timing signals, an EPLD (erasable programmable logic device) was used. An EPLD is a chip which can be programmed such that its outputs are logical combinations of its inputs, outputs and internal variables.

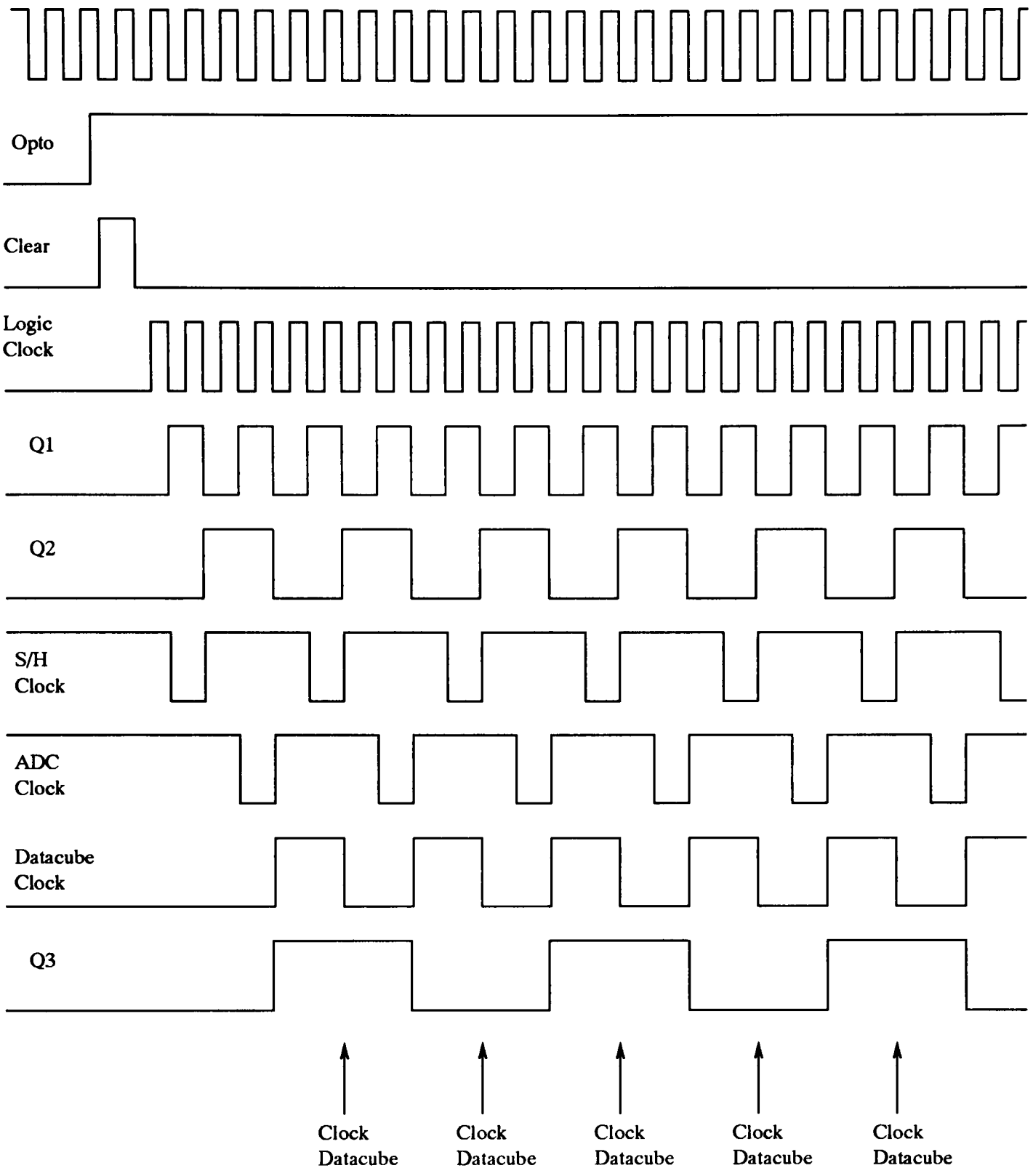


Figure 4.9: Timing diagram

Combinations of logic gates such as multi-input AND, OR, XOR, NOT etc. and D-type and T-type latches can be implemented on the same chip possibly creating a large and complex logic array. The chip can be programmed for specific pins to output specific logic values. Consequently, retaining pin-compatibility whilst changing the design of the logic is relatively simple. The use of EPLDs in this circuit obviated the need for a number of other logic chips and clumsy wiring.

Given inputs from the opto-pickup and the master oscillator, the EPLD was programmed to produce outputs to drive each of the multiplexors; the S/H; the ADC and the Datacube. A number of programs were tried before the final version was chosen. The final version, implemented on a 600 gate 24 pin Altera device, used synchronous logic and successfully clocked each of the components in the correct order and at correct intervals.

Following a change in the state of the opto-pickup, the sequence was started; and continued until all 64 data elements had been read. The count was then stopped. Care was taken to ensure that the device was left in the correct state at the end of each line scan. Because of a slight delay in the Datacube reading the data (due to the acquisition time of the S/H and the conversion time of the ADC), the Datacube clock continued slightly beyond the end of the count.

4.5.3 Phase-locked-loop

The object of the phase-locked-loop (PLL) was to drive the chopper blade at a fixed angular frequency with a fixed, known phase.

A PLL compares two input signals, $SIGN_{IN}$ (signal input) and $COMP_{IN}$ (comparator input). If $SIGN_{IN}$ has a higher frequency, the output of the PLL is held predominantly at VDD (5V) and in a high impedance state for the remainder of the time. If $COMP_{IN}$ has the higher frequency, then the output of the PLL is held predominantly at VSS (0V) and in a high impedance state for the remainder of the time. If the two signals have the same frequency but differ in phase then the PLL output remains at VDD or VSS for a time corresponding to the difference in phase.

If the output of the opto-pickup is directed to one of the inputs of the PLL and a divided version of the master oscillator to the other input, then the output of the PLL can be used to determine to what extent the two signals differ and to drive the chopper motor to bring the two signals to the same frequency and phase. If the frequency of rotation of the chopper is less than the frequency of the divided oscillator frequency, then the chopper drive circuitry will accelerate the chopper motor. Conversely, if the chopper is rotating at a higher frequency than the divided oscillator frequency, then the output of the PLL will act in such a way as to decelerate the chopper.

The required angular frequency of the chopper was 50/8Hz, which would produce a square wave from the opto-pickup at 50Hz. The frequency of the master oscillator therefore needed to be divided by 640. Purpose built divider chips were considered for this task, but rejected for a number of reasons. Instead, a second

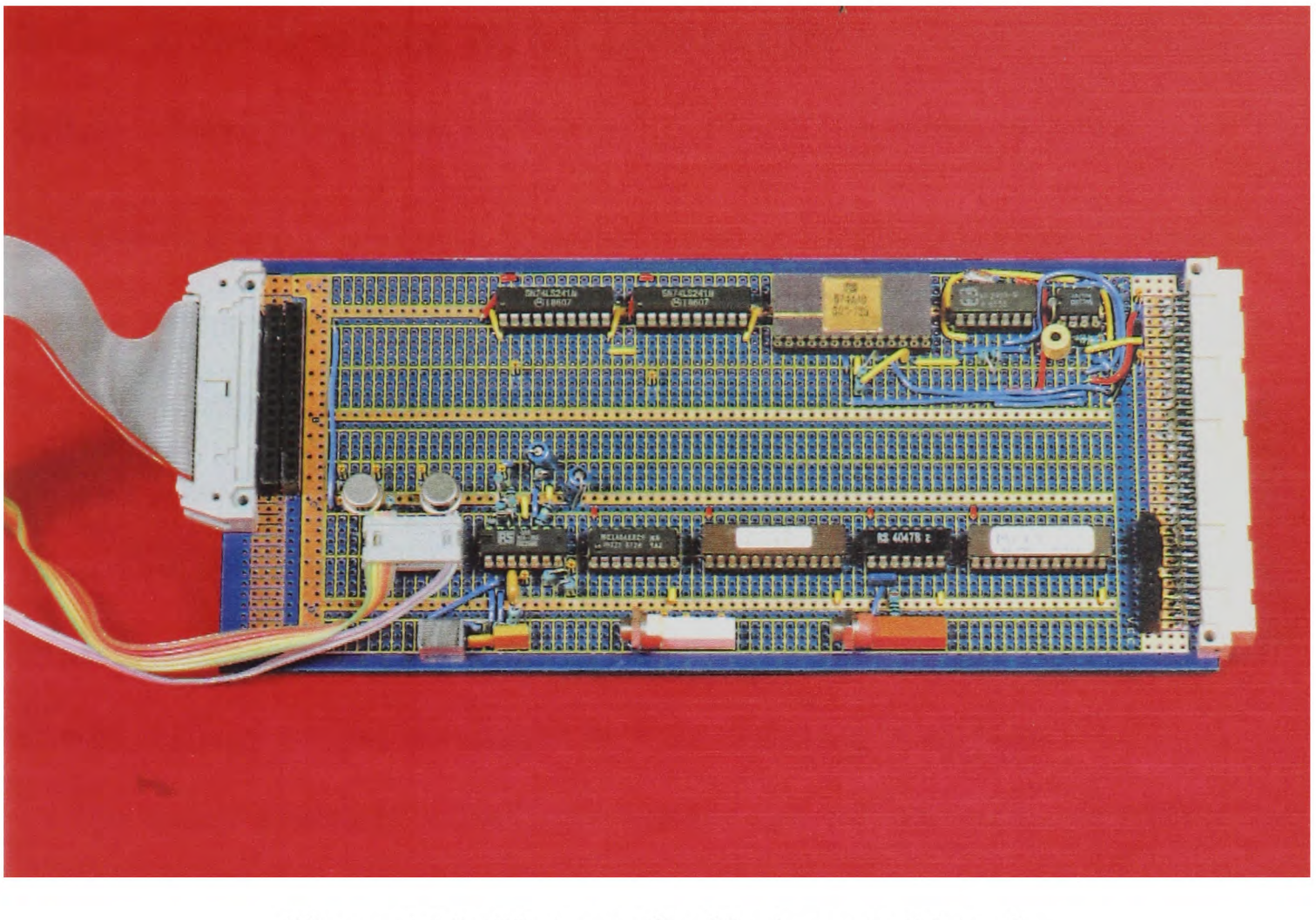


Figure 4.10: Photograph of logic control board

EPLD was used to do the division. Taking the oscillator as its input, it divided by 640 and output a square wave at this frequency.

The PLL performed correctly: the chopper was brought to the correct speed and phase quickly and continued to follow the input accurately thereafter.

Figure 4.10 shows the board containing the PLL, oscillator, EPLDs, power amplifiers, ADC, S/H and motor driver circuit. The grey ribbon cable carries data to the Datacube and the other ribbon cables control the chopper motor and the opto-pickup.

4.6 Interfacing with Datacube

4.6.1 Introduction

The data had to be transferred from the imager into a Sun computer. A number of options were available: each of which presented its own problems. The Sun itself, lacking a parallel port, was incapable of reading the data at the rates required.

The possibility of using a microprocessor to collect the data and then re-transmit it to the Sun was considered. All the available microprocessors, however, either lacked the ability to read data at the required rates, or did not possess enough memory to save the images we were expecting to produce.

The group possessed a Datacube system and was expecting to purchase a Maxscan board to use with it. The Datacube MaxVideo system is designed to be a versatile and high speed video acquisition and processing system. A number of boards, each designed to perform separate dedicated tasks can be pipelined together or, in some cases, used alone in either acquiring, processing or storing frames of data of programmable size and shape. The boards are connected together using a "MAXbus". Data is sent along the MAXbus and processed in a pipeline manner (i.e. processing on the target board can start when only a small portion of the image, typically 3 lines, has been received). Each board is also connected to a "VMEbus" which is used for communication both from and to the host computer.

Typical applications of the Datacube boards in the Robotics Research Group

in Oxford include adding or subtracting two images, local image processing (e.g. convolving with a three-by-three mask), acquiring images, resizing images and multiplying images. It has also been used to implement morphological filters. Some of the work done using the Datacube is described in [13].

A Maxscan board is a general purpose input interface designed to read data from non-standard devices into the Datacube. The Maxscan board was amply fast enough; could speak directly to a Sun host; and was part of a family of Datacube boards which could offer a degree of fiexibility for future work. It was therefore chosen as the interface between the imager and the Sun.

A Datacube is a piece of hardware attached to a host computer. To control it requires correctly connecting the various boards; initialising the boards as timing masters/slaves; selecting a region-of-interest bus and MAXbus; and setting the various software switches which control the data flow.

In the Maxscan board, software readable and/or writeable registers control which regions of the data are to be passed on to the next process. Information must be supplied to the board so that it knows when to begin reading data - this can be on a signal from the host computer, an internally produced signal, a signal from an external source, or it can be programmed to acquire data continually. The board must also be told what to do with the data: it can be sent via the VMEbus to the host computer or via the MAXbus to another board on the Datacube.

4.6.2 Design of test board

The decision to use a Datacube led to a number of problems. In particular, the Maxscan board was very poorly documented. It was not clear from the manual what sort of data Maxscan was expecting to be sent. We therefore decided to build a test board producing a known signal. Using this, we could examine how the Maxscan board functioned.

The board that was built consisted of an oscillator which drove a series of binary counters. Twelve outputs were taken from the counters and sent to twelve of the digital inputs on the Maxscan digital input port. (The other four digital inputs to Maxscan were tied to ground.) The board also provided a clock pulse, an inverted clock pulse, a line pulse and a frame pulse. CMOS chips were used throughout the test board.

The output of the board was a digital number, which started at zero and incremented in unit steps every pixel period until it reached 4095 (2^{12}-1). A line pulse was produced every 64 pixel periods and a frame pulse, lasting for 64 pixel periods, occurred every 4096 pixel periods (every frame). This data was output continuously.

The high level design of the board is depicted in figure 4.11.

When the test board was working correctly, we connected it to the Maxscan and attempted to read the data it was producing.

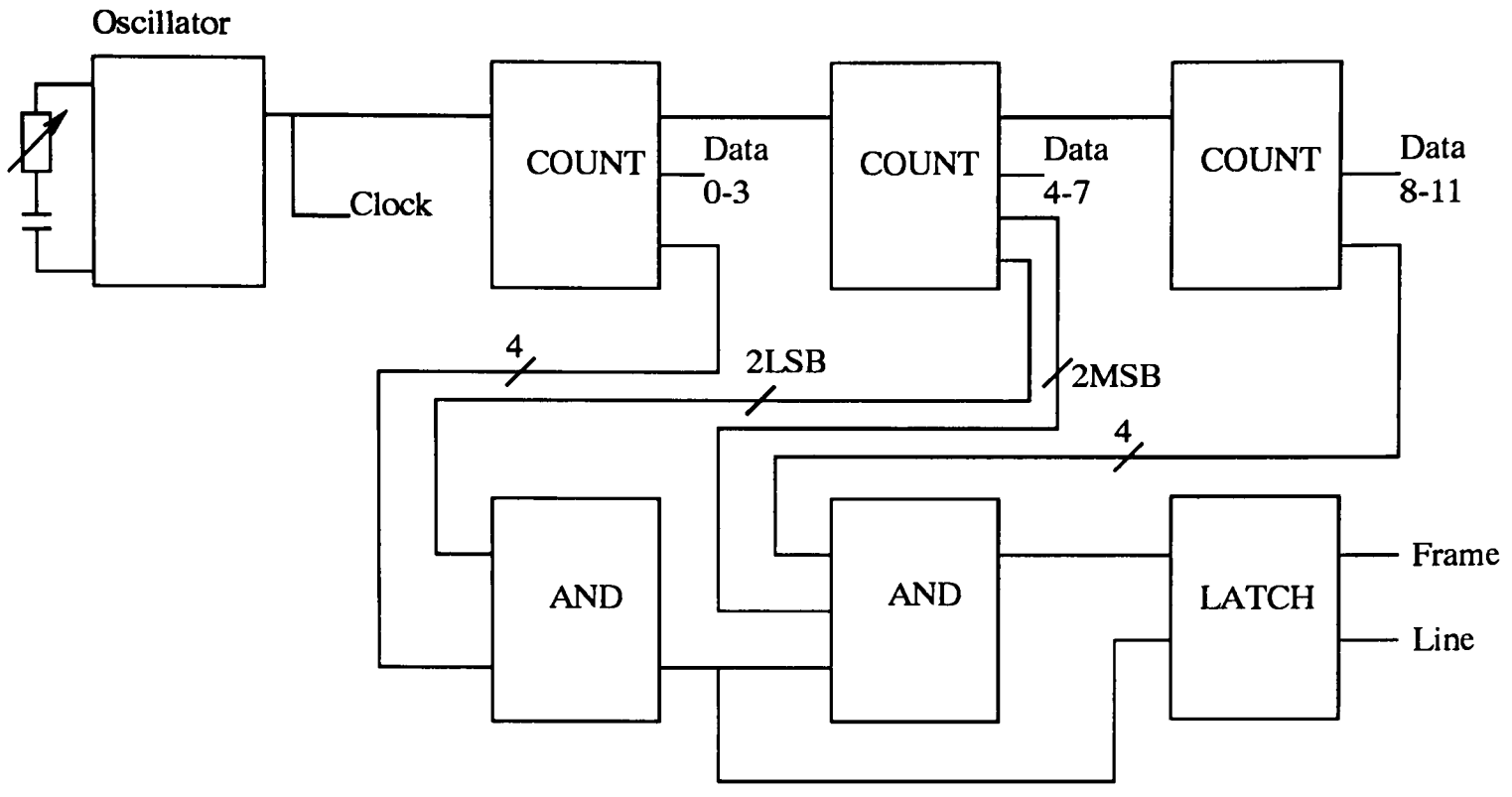


Figure 4.11: Design of test board for investigating Maxscan

4.6.3 Results

The most obvious and important result was the discovery of the type of signal which could be read by the Maxscan board. The manual explained how to set up Maxscan to accept timing signals from CMOS devices. No mention was made, however, of what signals the digital input could or could not read. Experimenting with the test board, it became clear that the CMOS devices were barely capable of producing sufficient strength signals to be read by the Maxscan. The signals being input to the Maxscan board were greatly attenuated; to such an extent that they could only be read when the input voltage to the test board was set within a very narrow range of values. It would therefore be necessary to build a TTL buffer or a voltage level converter on the end of the final imager, or to design the

imager using TTL chips. This latter course of action was not taken as work on the imager electronics had already been started. The outputs of the imager were therefore buffered before being sent to the Datacube. This proved to be sufficient to prevent the Maxscan from affecting the electronics of the imager and for the imager to be capable of driving the Maxscan.

Having set up Maxscan to read CMOS timing signals, these were read in without problem. In the final imager, the timing signals were also buffered and the Maxscan was reconfigured to read TTL data.

It was possible, by carefully adjusting the voltage driving the test board, to read data into the Maxscan. The test board could then be used to experiment with Maxscan to discover how to set it up to read a frame of data correctly.

At this stage, we were severely hampered by the manuals for Datacube's software package: "MaxWare". The manual lacked key information and at times was totally misleading in its advice. Different sections of the manual contradicted each other (sometimes even on the same page); while other sections, gave information that was, as we found by examining the software primitives, totally opposite to what should have been written. We made numerous changes to the manual, which have aided subsequent users of Maxscan.

Having overcome these difficulties, a program was written which successfully set up the Maxscan to read in a frame of data. It was then possible to examine the effect of setting or unsetting various software switches on Maxscan. Based on the

results obtained, we decided on a combination of switch settings which minimized the delay between the line/frame syncs and the data being read.

The delay in the frame sync, inherent in Maxscan, but poorly documented, would have caused an offset between the frame pulse and the start of the image being read. Though the use of the test board made it possible to take the delay into account, it was decided to use the extra data sent to Maxscan through one of the four previously unused digital inputs to eliminate any problems which may be caused by resetting any of the software switches on Maxscan. This "extra data", the buffered output of the opto-pickup, could be used to distinguish between open and closed chopper positions; and therefore to mark the beginning of a line of data.

The line syncs, which were also delayed with respect to the data, were only used to check that the correct number of data elements had been read. The image should always be 64 pixels high and any behaviour which caused the line syncs to be separated by any number of pixels other than 64 was reported by the program.

Interfacing the imager with the Datacube should not have been an overly arduous task. For the reasons outlined in this section, however, this apparently straightforward job took far longer than was planned.

4.7 The Second Dimension

4.7.1 Introduction

The linear array detector provides an image of "a line" in the scene which the detector is imaging. Obviously, a two dimensional image would provide far more

information and in most applications would be far more useful.

A two dimensional image can be created in three ways: by moving the scene, by moving the imager, or by having something between the scene and the detector which gives the effect of a moving scene. (e.g. a rotating mirror).

A moving scene in front of the imager (a person walking, a car driving past etc.) gives adequate images, but to image any stationary object in this way would entail moving the object at constant velocity across the scene. Clearly this is an unsatisfactory solution for a general purpose imager. How, for example could it image a building? For the experiments we wished to conduct, in which we wanted to examine a wide variety of stationary scenes, it was essential to provide a second dimension for inanimate objects either by moving the imager or by scanning the scene with a mirror or similar device.

Rotation as a means of scanning was preferred to translation because of simplicity of implementation. (Systems using translation as a means of scanning include, for example aircraft and vehicle mounted imagers.)

We were left therefore with the choice of panning the imager or panning a mirror in front of the sensor. (Incorporating mirrors inside the detector was rejected due to the complexity of design.) Panning a mirror was not ideal, since imperfections in the mirror would distort the image and the mirror would not reflect all of the incident radiation; but panning the imager was not feasible, due to its large mass, awkward size and fragility. Eventually we decided to use a

mirror to produce the panning because its light weight would make it far easier to rotate; and high accelerations could do no damage to it, therefore an image could be taken more rapidly.

A polished aluminium mirror was chosen, aluminium having reasonably good reflective properties in the thermal spectrum (see, e.g. [96]). Thin aluminium foil, with far better reflective properties was considered, but the problem of mounting it flat proved too difficult. A mount was built to connect the mirror to a motor shaft. The polished mirror has proven satisfactory in practice; though it does cause degraded performance compared to taking images without the mirror and moving the scene instead. This is due to the spreading of the infrared signal by the mirror which is by no means a perfect reflector and by the loss of some of the signal due to absorption by the mirror.

4.7.2 Driving the Mirror

The motor could be driven either by a stepper motor, or by a DC motor. A stepper motor would give a more rigid pixelisation of the scene and for that reason was preferable. We required very small steps, however, and geared stepper motors were unavailable without a delay in manufacture in excess of three months. We decided to use a DC motor instead. While this would in principle cause a slight skewing of the image, it has the advantage that it could be run at any speed.

We needed to drive the mirror at a constant velocity (possibly a range of velocities) with control over its starting and stopping. Although it would have

been possible to build or purchase a motor driver to do this, we decided to use a motor interface, developed in the laboratory by Henry Lau [44]. The interface was specifically designed to be used in projects of this kind. Although the motor was not ideal, being designed for position control rather than velocity control, the lightness of the load it was to drive made it perfectly acceptable.

The motor interface contains RAM into which code in S-Record Format (a form of raw binary) must be loaded each time it is turned on. The interface had been developed using the software development system OS-9 running on a 68020 machine. The S-Record format file had to be ported onto a Sun, therefore, so that the motor interface could be loaded without the need for the OS-9 system.

The interface, as given, was operated with one serial line to a dumb terminal and another to the OS9 system. Commands issued from the dumb terminal controlled the other serial line. We first "replaced" the OS9 system by a serial port on a Sun computer. Code was then written to emulate the dumb terminal so that a second port on the Sun could be used, "replacing" the dumb terminal. A program was written which loaded the recently turned on interface with the S-Record file and left it ready for use. Another program made it possible for commands to be issued from the Sun to start, change speed, reverse direction or stop the motor (for example). A version of this program was written into the imager driving software (see 4.8).

The motor, although reduced by a gear-box, could not rotate the mirror at

sufficiently slow speeds without encountering friction forces which affected the rotation speed. (The required rotation speed was approximately 0.0125Hz.) We calculated that a 2048:1 reduction gear-box would enable the motor to spin at a comfortable speed while rotating the mirror at the required rate. The gear-box was purchased and when we fitted it, we found that it solved the friction problem.

To find the exact speed to rotate the motor, an image was taken of an object of known size. The aspect ratio of the image was compared with the aspect ratio of the object and a new speed for the motor was calculated which would give the correct aspect ratio for the image. This speed was adjusted further by taking further images and comparing the aspect ratios in the same way. (Non-linearity in the motor/driver caused the calculations to be slightly inaccurate.) Images could now be taken which had the correct aspect ratio, though it was still possible, and indeed simple, to vary the speed of the imager if required.

The motor controller caused the motor to run at different speeds when it was actuated in different directions. To avoid both this problem, and the concomitant ambiguity introduced by running in different directions, we decided to restrict images to be taken with the mirror travelling in one direction only.

4.8 Sun Software

The imager was to be controlled by a Sun computer on which the Datacube drivers and application software resided. The Sun had to be programmed to control the Datacube; to control the motor interface to the mirror motor and to provide a user-

friendly interface for running the imager. To achieve the last of these three goals, we wrote a window-based program using the SunView [1] package. A screendump of the layout of the final version of the program is shown in figure 4.12.

Initially, a panel was created containing buttons to capture an image; save an image; save the raw-data and quit. At this stage, an image of fixed size was to be read in. Clicking the mouse on the "grab" button caused a command to be written to the Datacube to begin reading data on the next vertical timing signal. The twelve bit data from the imager for both open and closed fields was read into the digital port of the Maxscan, transferred into the FIFO and then read over a VME bus into the Sun and stored in an array. Clicking on "output-save" caused the contents of the array to be written to a file; while clicking on "image-save" wrote just the eight most significant bits of the difference between open and closed fields for each data item to an MIT [2] image file; which could then be viewed.

We found that invariably, the eight most significant bits of data were sufficient for examining the image. The imager was still being developed and the proliferation of noise in this development stage rendered the four least significant bits of data virtually useless.

We decided to read in a frame of data rather than adopt the RSRE approach of

[1]Sunview is a system supporting interactive graphics-based applications running within windows. (See [84])

[2]Images are stored in memory as a list of intensity values. A "header" is attached to the beginning of the file containing information about width, height and type of the image. An MIT header is one such header which is popular in the department and is compatible with almost all software in use in the Robotics Research Group.

continually reading data and reversing the direction of panning after each scan. To read in data continuously would not be difficult using the Datacube (by disabling external horizontal and vertical resets), but reversing the direction of panning introduced unwanted (and in our case unnecessary) ambiguity; in addition to the problem of eliminating effects of the speed difference discussed in section 4.7. An alternative approach might be rapidly to reposition the mirror ready for another scan after each scan was completed, but as we really wanted single images and could be sure of grabbing the image we wanted, we decided to read in only one frame for each image. Also, repositioning the motor accurately was non-trivial as the motor positioning was relative not absolute; so we could not be sure of beginning the new scan at the same position.

The program was adapted: the name of the file to be saved could now be specified by editing a panel text item; the number of lines of data to be read could be changed also by editing a panel text item; and clicking on another button caused the output image to be toggled between "black-is-hot" and "white-is-hot".

Control of the motor was then incorporated into the main program. Clicking on "grab" now started the motor before beginning to read data; and stopped the motor when a complete frame of data had been read. Buttons were added for positioning the mirror: clockwise, anticlockwise and stop. To tidy up the display, a pop-up panel was used for motor control. A panel text item was used to set the fine speed of the motor; and a panel cycle was used to set the coarse speed.

The button to change from "black-is-hot" to "white-is-hot" was replaced with a neater panel cycle and options were added to change an upper cut-off threshold (above which, pixels would be set to 255) and a grey offset (added or subtracted from all data values). The upper threshold and grey offset together histogram stretched the image, facilitating interpretation of the stored image. For most images, it was not necessary to adjust these parameters but they were extremely useful for viewing low contrast images or images whose temperatures were outside the normal viewed temperature range (e.g. cool images). The approach used in the RSRE imager, where it was not possible to interactively adjust the offsets, was to time average the data over several seconds and subtract from this the data values. Images produced in this way will not normally be as clear, but the RSRE approach has the advantage that the "average" is updated as the image is being taken so the imager adjusts with only a small delay.

To avoid the necessity of saving the image to view it, an image previewer was built into the program. By clicking the mouse on a "show image" button, a canvas was drawn and a raster operation filled the canvas with the pixel values. The image displayed by the previewer depended on the values of the grey offset, threshold and black/white cycle. The image could therefore be examined and, if necessary, adjusted before being saved.

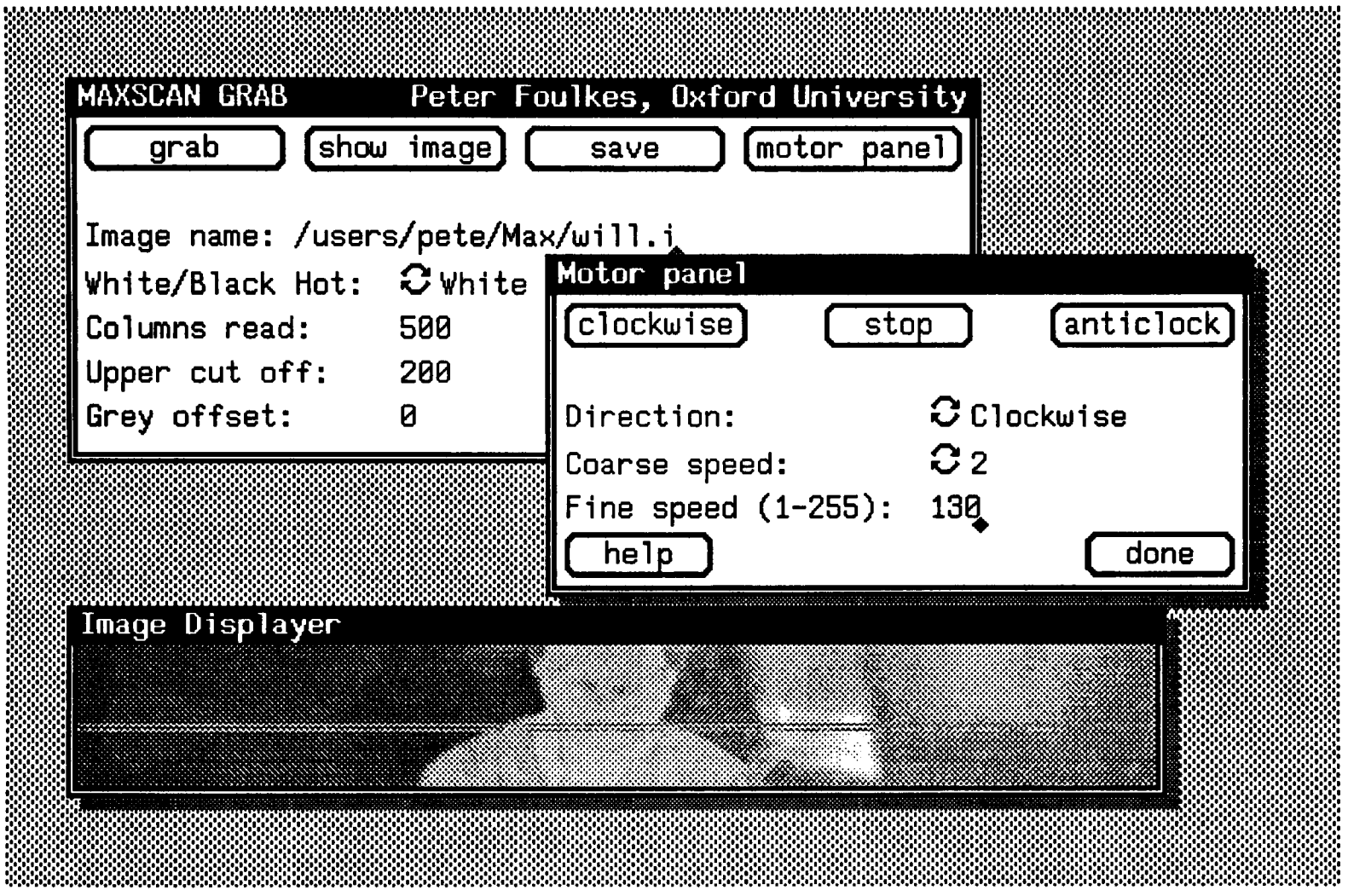


Figure 4.12: Screendump of the layout used by the final program

4.9 Noise Reduction

Throughout the construction of the imager, minimization of noise was a primary consideration. The electronics were designed to keep digital and analogue components and signals separated as much as possible; and care was taken to avoid creating earth loops in the system. Notwithstanding this, however, early images were rife with noise. A large amount of time was spent tackling the problem of noise: some of the problems are outlined below.

Two of the elements failed to produce correct outputs: either the detector elements themselves, or the pre-amplifier electronics was faulty. No fault could be found in the electronics, though one may have existed. Under some conditions, the non-functioning elements picked up noise at mains frequency with a high amplitude. This noise was also picked up by the three other detectors whose pre-amplifier electronics occupied similar positions on the other boards. This problem was solved when the electronics was tidied up, ensuring that no contact between the boards and no contact between the boards and the casing or the rack was allowed.

Noise at mains frequency and aliased versions of mains frequency were ubiquitous in the images we were taking. Changing the power supply; even moving the power supply with respect to the imager caused the magnitude of the noise to vary. The levels of noise were exacerbated by sampling the imager at a frequency close to the mains frequency we were using to drive the electronics: thereby introducing

additional beat frequencies into the noise.

A number of possible solutions to the noise problems were considered including optically isolating the imager from the Datacube (the source of noise could be due to a mains loop through the Datacube and the imager) and fixing the problem in software by Fourier transforming the image, removing the peaks in the Fourier transform corresponding to the mains noise, and taking the inverse Fourier transform of the result. The solution that was chosen, however, was to run the imager from batteries (the RSRE system was designed to be run from batteries), which would also break any mains loops. After a few teething problems, this proved to be extremely successful: eliminating the majority of the noise.

Figure 4.13 shows a cross section of an image taken of a constant temperature wall. The upper line shows a cross section when the imager was being powered by a power supply and the lower line shows a cross section when the imager was being powered by a battery. The x-scale corresponds to number of pixels across the image, while the y-scale corresponds to intensity. The difference in y-scale comes from setting a different grey level offset when taking the image; and does not affect the magnitude of the noise. The noise in the image taken using the battery has been reduced to about twenty percent of what it was in the image taken using the power supply.

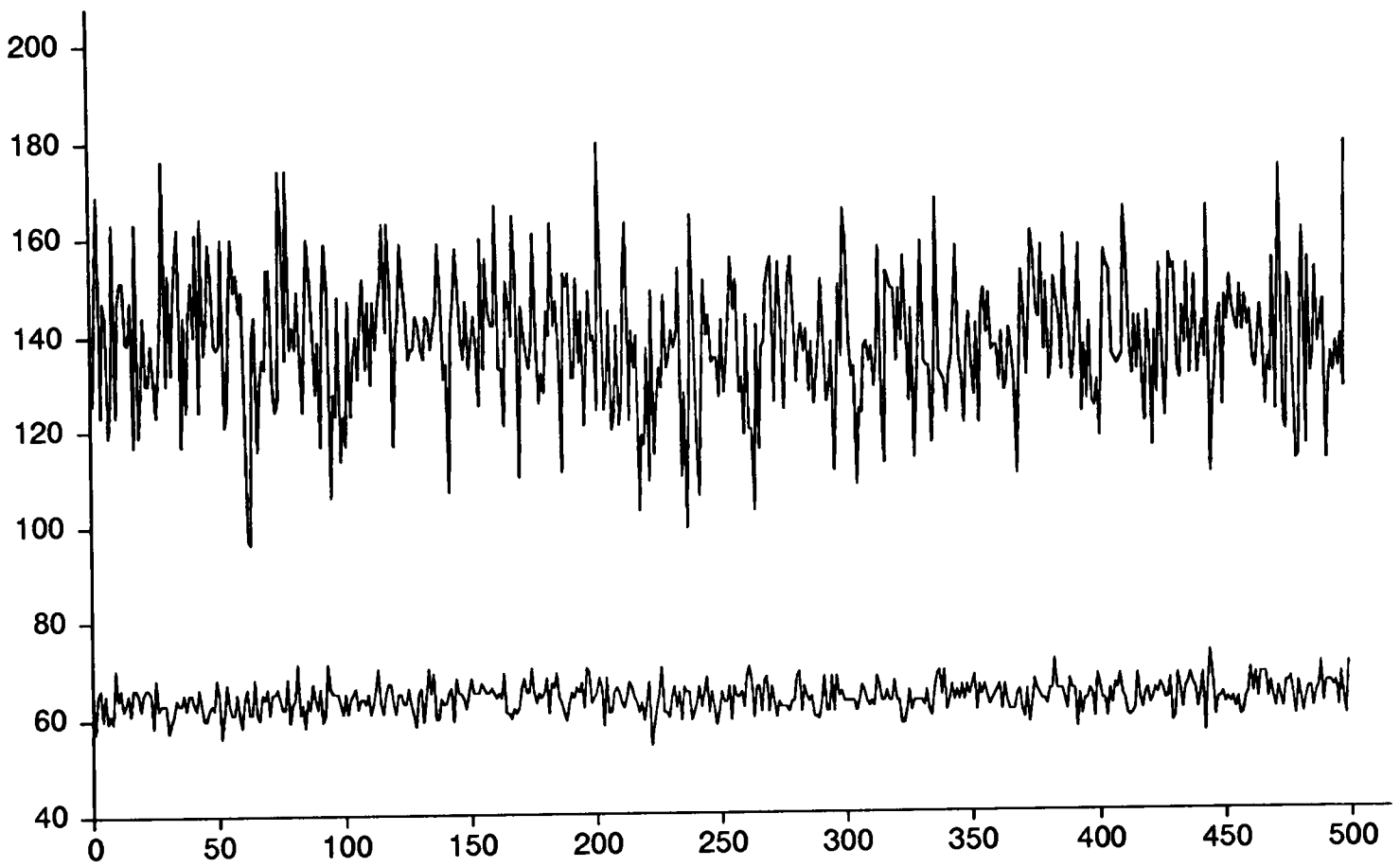


Figure 4.13: Cross sections of constant images

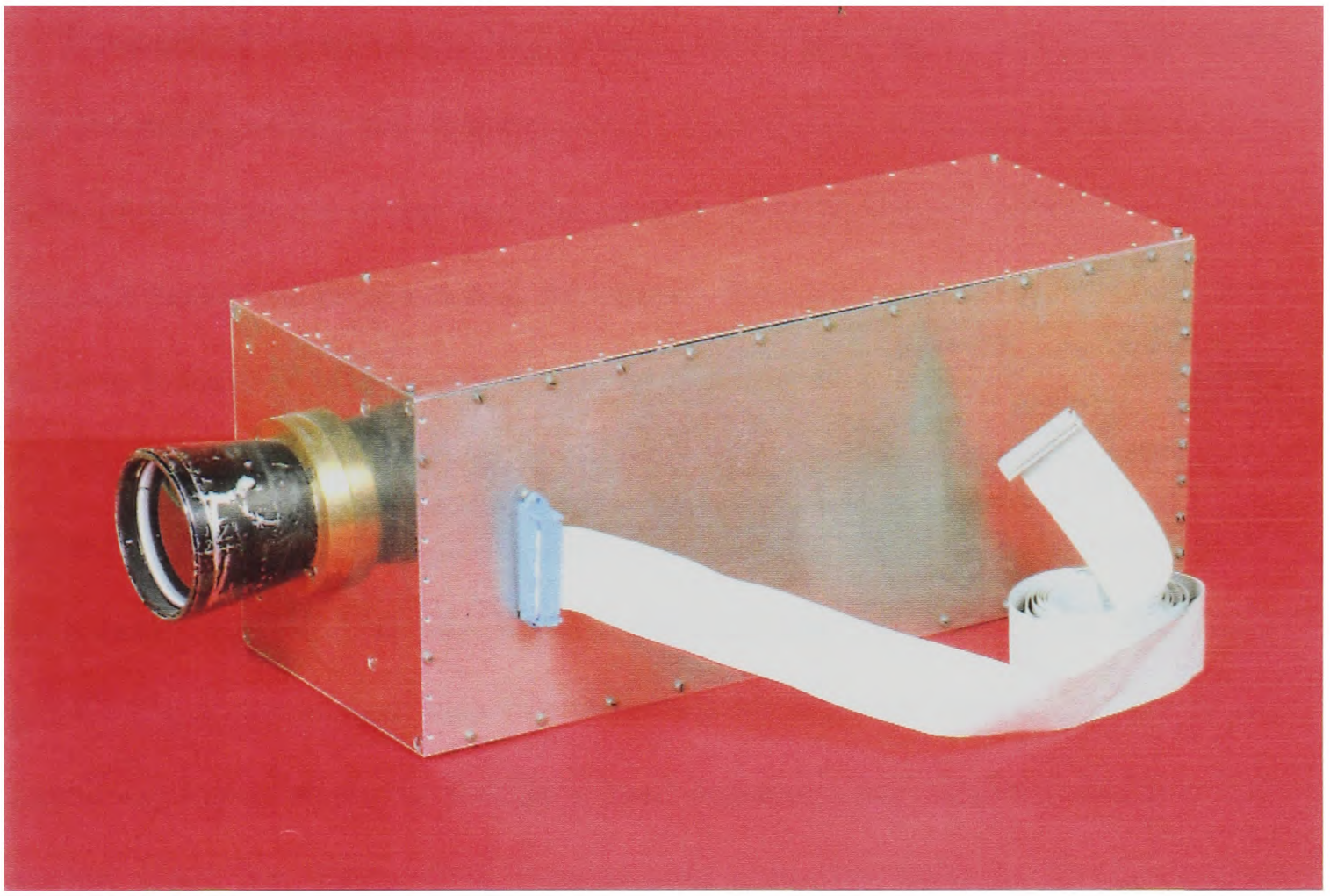


Figure 4.14: Photograph of imager housing

4.10 Housing the Electronics

Figure 4.14 shows the box containing the imager. The imager was housed in a box that was big enough to contain it leaving room for any possible modifications which at the time were being considered (e.g. a transformer for running the imager directly from the mains; or additional signal processing). The electronics were mounted on a Eurocard racking system which was positioned inside the main box. Power supply leads to the equipment were attached to sockets in the box and the output of the imager went to a 34 pin socket into which the ribbon cable from the Datacube was plugged.

The parts of the mounting which required accurate design were the front panel and the panel holding the chopper motor. The chopper needed to be rotated very close to the detector array; and the correct distance from it radially; while the detector needed to be positioned the correct distance from the lens for the image to be in focus at a sensible distance from the imager.

The lens, a fixed focus 50mm f0.7 germanium lens, was mounted on a Wreathall Mount on the front panel of the box. (A different focal length lens would have been preferable, but this was the only one available that RSRE could spare; and purchasing a germanium lens was infeasible for financial reasons.) The second panel was mounted on the first at a fixed distance using threaded spacers. This second panel (see figure 4.15) incorporated mounting holes to fix the chopper motor, mounting holes to attach the detector and a groove in which to locate the opto switch. The groove enabled the position of the opto-switch to be changed. Moving the opto switch changes the time at which the imager begins to read data relative to the time at which the chopper blade crosses the first element of the detector array. Altering the position of the opto-switch therefore increases or decreases the time that each detector element has to register the change in irradiation after switching from viewing the chopper to viewing the scene and vice-versa.

The detector itself was positioned behind the second panel using spacers. A sketch of a plan view of the mounting is given in figure 4.16.

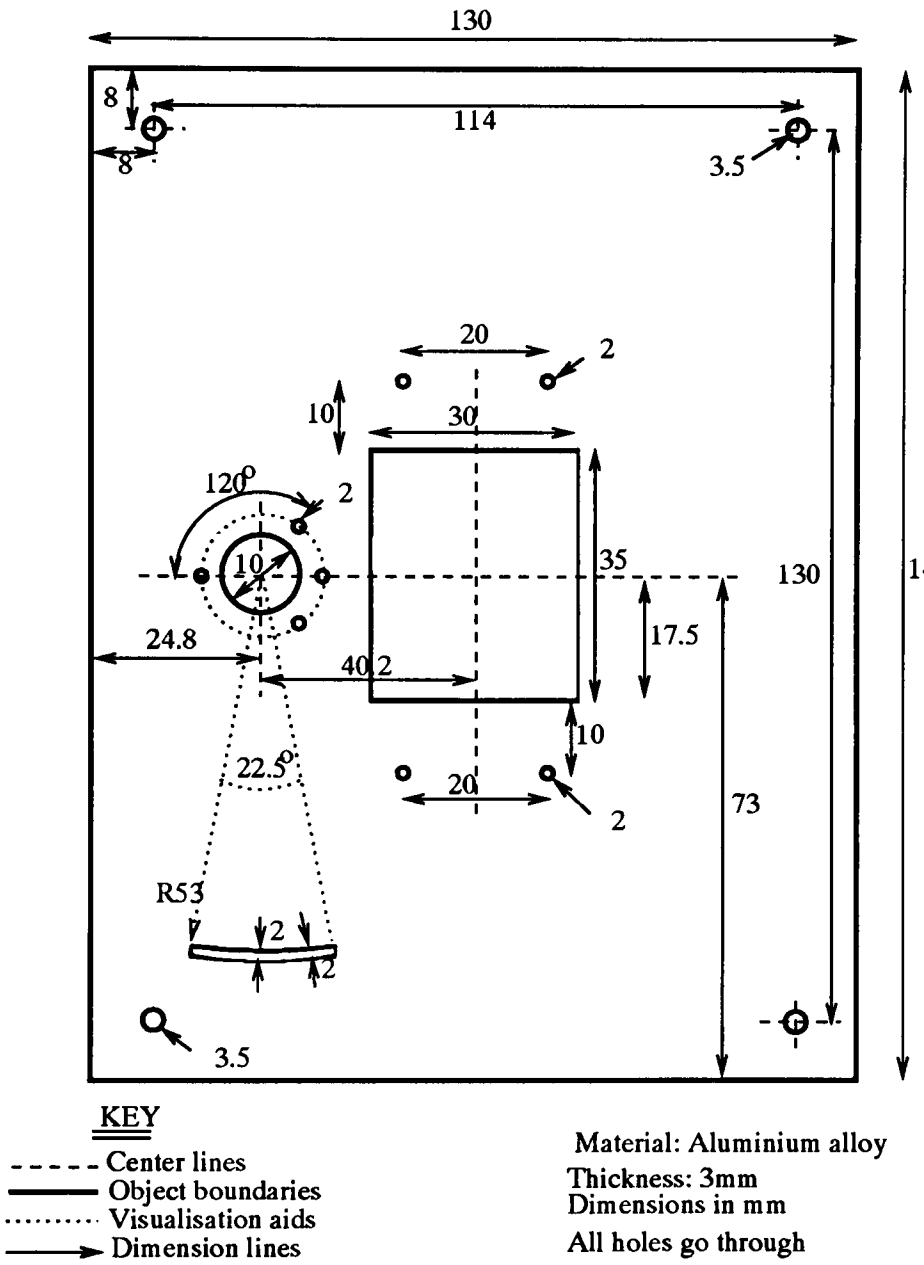


Figure 4.15: Second mounting panel

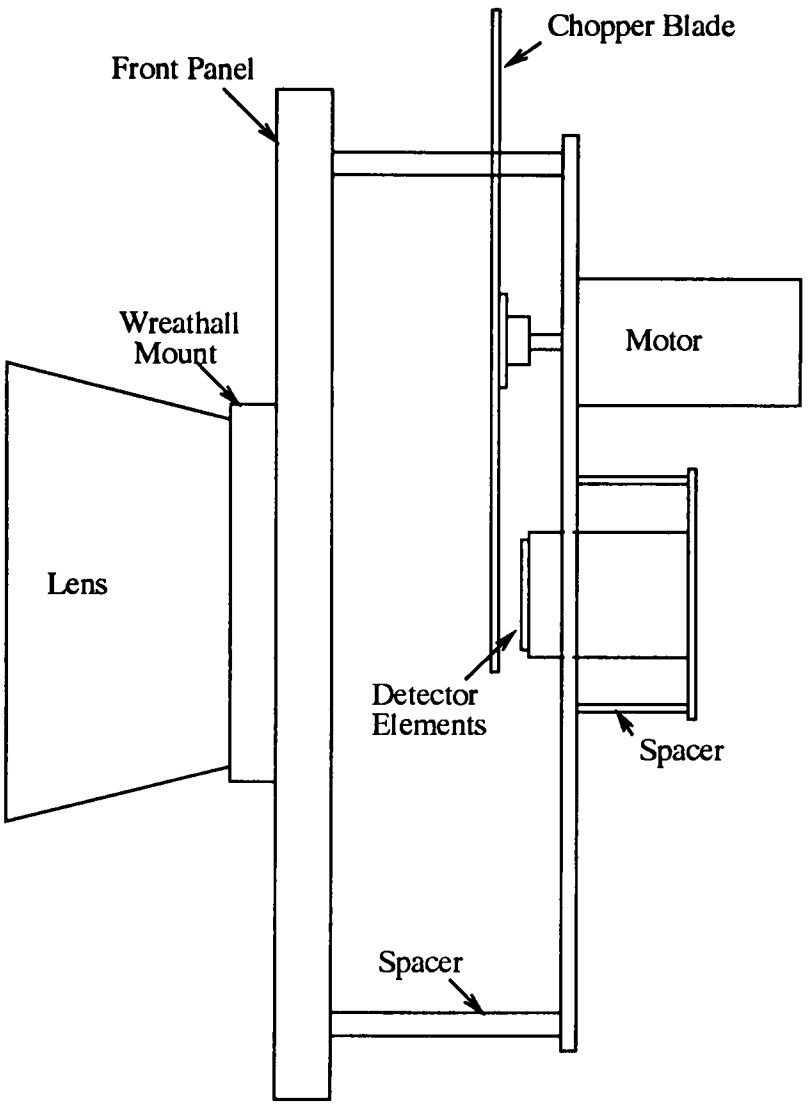


Figure 4.16: Plan View of Mountings

The calculations of the required focus distance proved to be slightly inaccurate due to the vagueness of the specification for the Wreathall mount. The distance between the detector elements and the lens was reduced by shortening the length of the spacers until the opto pickup was flush against the front panel of the box. Needing to move the detector even closer to the lens, a small area had to be milled out of the front panel. The new focus distance, approximately five feet, was ideal for taking images of small objects the size of a cup.

Figure 4.17 shows a photograph of the front panel on the right and the second mounting panel on the left with chopper, detector window and opto-pickup visible. Figure 4.18 shows a photograph of the assembled front panel mountings: the lens, Wreathall mount, front panel and second mounting panel with chopper motor, opto-pickup and detector elements mounted on it.

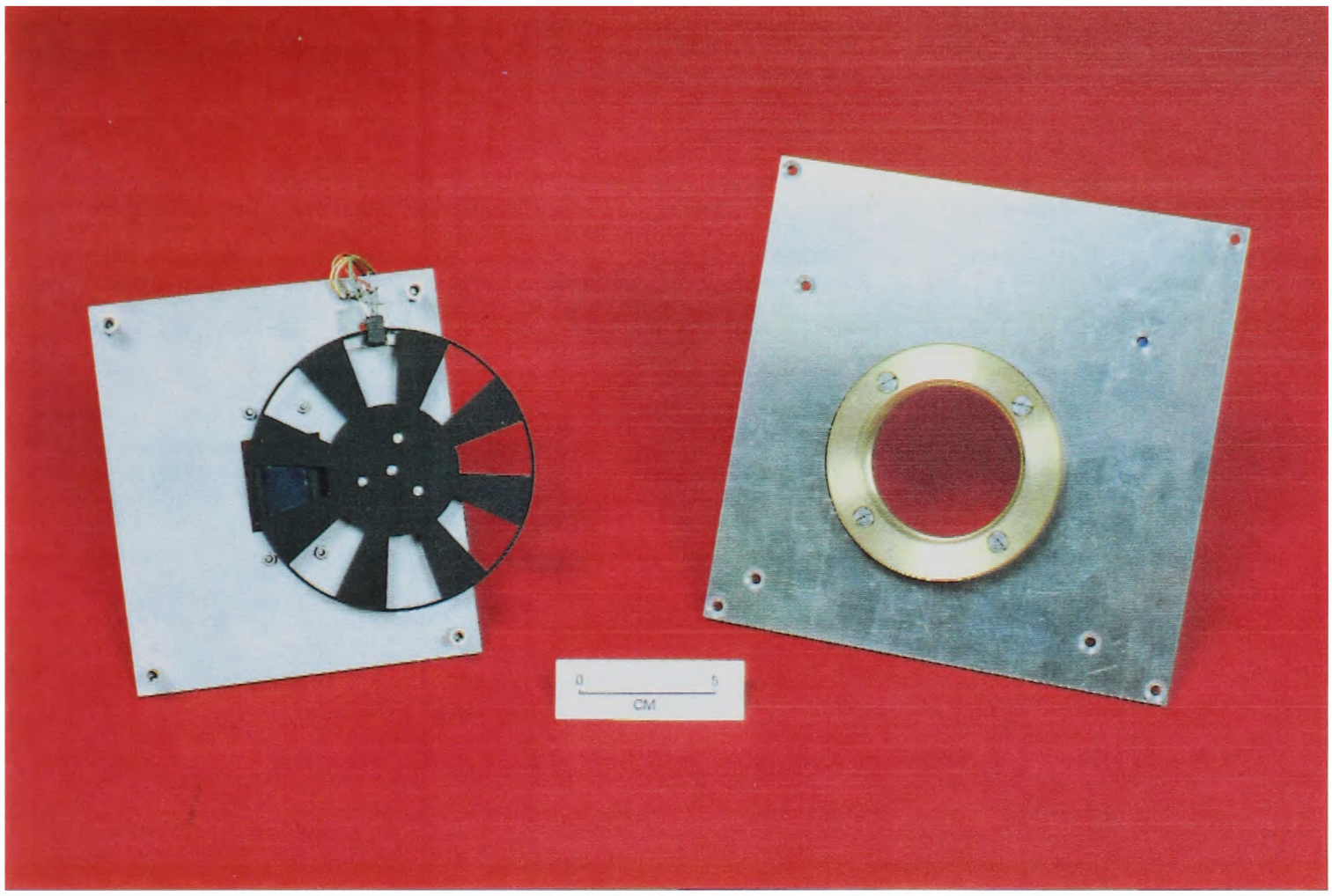


Figure 4.17: Photograph of mounting panels

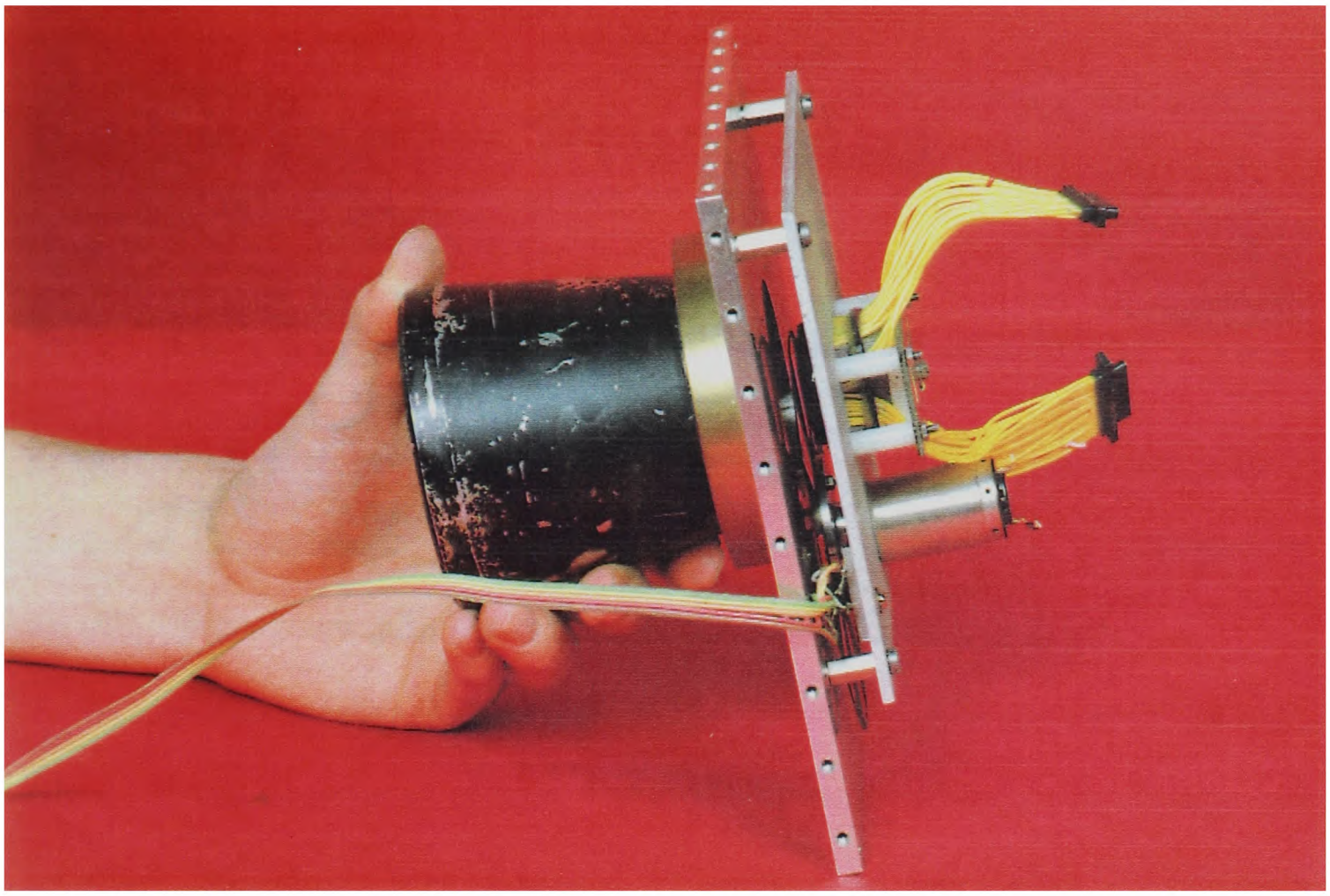


Figure 4.18: Photograph of front panel mountings

Chapter 5

Early Image Processing

5.1 Image Difference Processing

In section 4.3, it was pointed out that to image a stationary scene, a chopper is needed to alternately "open" and "close" the array to the scene. In this section, we will explain more fully the reasons for, and benefits of, this process.

As explained earlier, a pyroelectric detector responds to changes in temperature. In the absence of change, there will be no observable output from the detector. Following a step change in temperature on the surface of a detector element, the output of the element will change: initially it will increase exponentially, the rate being dependent on the thermal time constant of the detector. Subsequently, the charge across the detector caused by this change in temperature will leak away, and the output of the detector will fall once again to zero. It is therefore necessary to incorporate some form of chopping if constant temperature scenes are to be viewed.

The electrical time constant is much larger than the thermal time constant

(by orders of magnitude). The detector is chopped with a time period close to the thermal time constant so the effects of the electrical time constant on a chopped array are not noticeable, and will not therefore be considered further. When the detector is chopped, the step change in detector temperature from chopper temperature to scene temperature causes the detector output to change: it causes it to tend exponentially towards what in the absence of charge leakage would be an open field equilibrium value. At normal chopping frequencies, however, when the chopper closes, this change is arrested before it reaches this "equilibrium"; causing the output to tend exponentially towards the closed field "equilibrium" again.

The output of each detector element will change each time the chopper opens or closes and in a constant scene will vary between two values. For each detector element imaging the same temperature, the magnitude of the output for the open fields will vary quite considerably, as will the magnitude of the output for the closed fields. These differences in magnitude are usually larger than the differences in useful signal and if unattended would cause an image to have an appalling amount of fixed pattern noise. Their effect must be eliminated, or at least reduced.

Fortunately, simply subtracting the output of the closed field from the output of the open field removes most of the effect of the differences between the detectors. The signal remaining after the subtraction is proportional to the difference between the temperature of the scene and the temperature of the chopper. By

performing this operation, therefore, we have also established a temperature datum, i.e. the temperature of the chopper. The process of subtracting the open field data from the closed field data is known as Image Difference Processing (IDP) [50]. In this particular example, two fields are used to calculate the output, and hence the method is known as two field IDP. In practice, two field IDP works extremely well, but it does introduce a thermal lag causing the smearing of hot objects over several columns. The cause of this is described below.

While the imager is pointed at a stationary scene, the output values that the detector does reach will remain approximately constant, and will lie between the chopper and scene asymptotic equilibrium values. If the temperature of the scene rises, however, then the output of the detector will move towards a new equilibrium position. Because the temperature differences are larger, the gradients of the exponentials will increase; and the difference between the open and closed outputs will be larger. When the imager is settled at this new temperature, the output values which the detector reaches both in the open and closed fields will be higher than the corresponding values before the temperature increase.

The change from one settled state to the other will occur slowly, however, because the exponential curves are truncated before reaching equilibrium. Over a period of several columns, the offset of both the value returned in the closed field and the value returned in the open field will climb slowly upwards, as shown in figure 5.1.

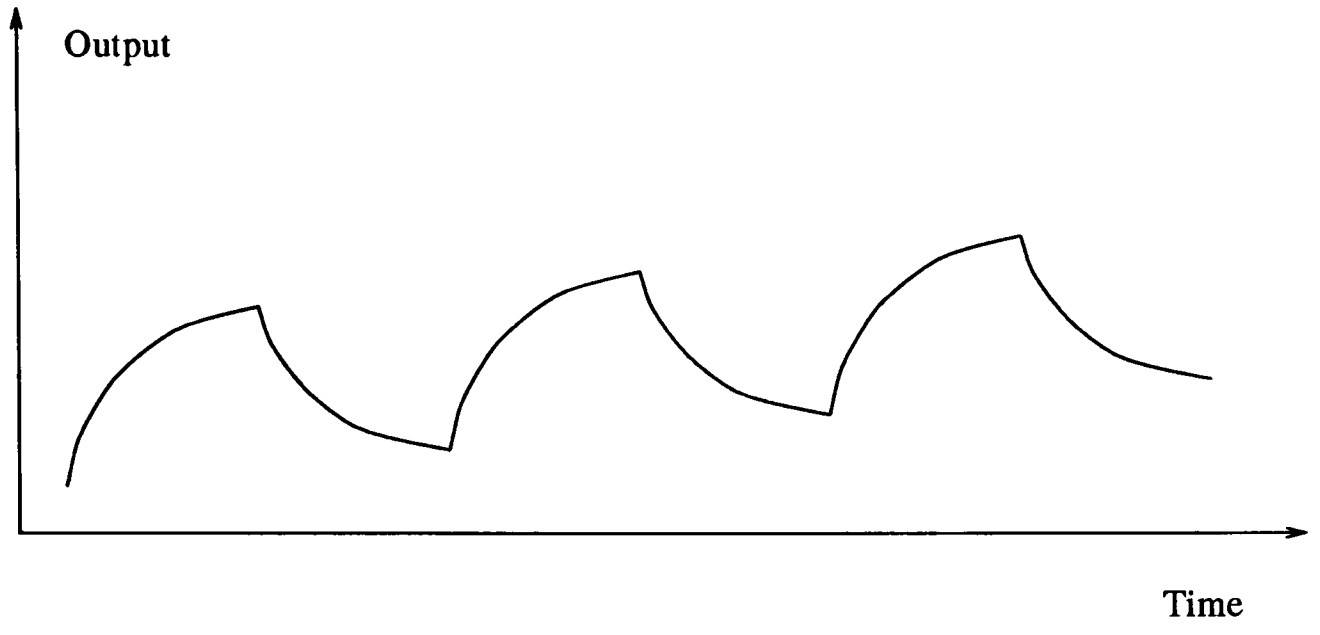


Figure 5.1: Output of detector element following a change in temperature

While this is happening, the output of a two field IDP will be inaccurate, since the ramping up of both open and closed field outputs causes the perceived difference between open and closed fields to appear less than it actually is. (Assuming the difference is formed by subtracting the closed field from the preceding open field.) A corresponding problem occurs when the image temperature is decreased; here the perceived difference is larger than the correct difference, causing "thermal trails" to follow hot objects.

This effect can be accounted for by using a three field IDP [90]. A three field IDP takes the values returned by the chopper in the fields on either side of the scene temperature in question, V_{C1} and V_{C2} and the value returned by the scene, V_S and gives an output value $(2V_S - V_{C1} - V_{C2})/2$. This eliminates the thermal lag apparent when using the two field IDP. Manning [49] has also shown that the

use of three point IDP additionally reduces the 1/f noise by 30% compared to the two point IDP.

5.2 Removing the effect of dead elements

It is not uncommon in pyroelectric thermal detectors for one or more of the detector elements to fail to operate correctly. When this occurs, the output of that detector element will usually remain constant, though the output of the element could also take some other form. It is possible, for example, for the lead to pick up signals from the mains: causing a sinusoidally varying signal to be output for that line or for a dead element to "flicker" giving a noisy output.

The output of each detector element is initially processed separately: each detector element has its own associated preamplifying electronics. A failure in any part of the electronics associated with a particular detector element could therefore block the data path, giving an effect similar to that caused by a dead element.

In the image, the effect of a dead element or faulty electronics will be to create a line of approximately constant intensity, or a line which varies in intensity in some unusual way. Whatever form the intensity profile of that line takes, it is certain that it will differ from the intensity profile of other lines which are functioning correctly.

Clearly, the presence of such a line will encumber any attempt to get information from the image, whether by a human observer who will be distracted by the

line; or by an automatic image recognition system. For example, consider a line in which the intensity differs greatly from the intensity of the lines surrounding it. This line will be identified immediately as an edge by any non-specific edge detection algorithm and will be classified as an interesting feature; whereas it is in fact of no interest whatsoever.

For any effective image processing to be done, the effect of dead elements must be eliminated.

To eliminate the effect of the dead elements, we must locate which of the elements are not functioning correctly and process the image appropriately. There are three ways we could approach this: we could filter the data as it arrives with no prior knowledge of the dead elements; we could locate the dead elements before processing and treat them separately when we read the data; or we could locate the dead elements before processing and update our position estimates as we acquire new data.

One of the simplest methods of removing the effect of the dead elements is to filter the columns of the image with a linear filter as the data are taken. Leaving open for a moment the question of what filter would be used, we consider the merits and demerits of this approach.

This approach will certainly be fast to implement: a convolution can be performed on the data as the data arrive; it requires no calibration stage; and it will adapt to changes in the imager. If for example an additional element dies, that

element will automatically be processed correctly. Unfortunately, it has a readily apparent disadvantage: the entire image will be affected by this method, while it would be preferable to process only the lines which contained useless data. The resultant image will have undergone a filtering designed to remove thin bars. Possibly useful high frequency data will also be removed by this process: clearly an unsatisfactory situation.

Because of the unacceptable shortcomings of this approach, we chose not to pursue it further, but to concentrate on identifying the dead elements and process only those elements in the first stage of processing.

The second approach is to calibrate the imager before use; locate the dead elements and then, when on-line, process those lines appropriately. This approach only fails when the characteristics of the imager change during operation. It is important, however, that the initialisation successfully locates all and only the dead elements since any mistakes will continue to be made until the imager is re-initialised.

The third approach is the most effective method. This is to keep a record of which elements are considered dead and update that record as new data arrive. In this way, a priori knowledge of the positions of the dead elements is used and it is not therefore necessary to recalculate them for every column of data read as with the first approach. Since the dead element positions are periodically updated, however, changes in the imager elements will be noticed and provision can be

made for correcting the image: this is not the case with the second approach.

This approach is, however, by far the most awkward to implement. The re-calculation of dead element positions would require looking at more than one column of data - it would not be a simple matter of performing a single convolution on the incoming data and would therefore require the design of special purpose electronics.

One way it might be implemented is by updating the current estimates of the dead elements using a predictor-corrector filter such as a Kalman-filter. In this way, the likelihood that an element is dead could be updated as new data arrive without having to consider data stored over a number of frames. As the re-calculation would only need to be done infrequently, not all the data would need to be included in this calculation: a selection would do.

Although live elements do die, it is not a regular occurrence, and considering the additional effort involved in updating the estimates of the dead elements, it is almost certainly not worth adopting this third approach.

It therefore appears that the most sensible approach is to pre-calculate the positions of the dead elements in a calibration stage. It would also be sensible to re-calibrate the imager occasionally to add any new dead elements to the dead element map. We must now decide how to detect the dead elements.

One approach we might adopt is to incorporate the dead element detection into the method described in section 5.3 for equalising the gains in the detector

elements. Several columns of data could be used to give an estimate of the average value of intensity for a particular line in the image of a constant scene. If the average of any element was very different to the average of the other elements, then we could conjecture that the element corresponding to that line in the image was dead. A similar method to this is used by Perkins and described in [56]. When calculating gain estimates, he decides that any element which has an output less than a user-selectable percentage of the mean output for the entire array is dead and does not include it in determining his gain estimates.

The additional processing time for this approach would in fact be minimal as the calculations would be combined with the processing to eliminate the gain differences. This is not a major consideration, however, since the processing is all completed off-line in a calibration phase.

An alternative is to detect the dead elements before proceeding to the gain estimate calculations. This is best done by processing the image locally, for reasons which will be described later. We have made the assumption that a dead element will appear in the image as a line which differs from the lines around it so we can detect the dead element by locally selecting lines which are "unusual". To define what is "usual" we could take the mean intensity value or the median intensity value of a group of neighbouring lines and see which lines differed from this average by some chosen threshold.

In this type of application, the mean is not really appropriate. Suppose, for

example, three neighbouring elements have values 100, 159, 104. The mean of the three values is 121 which is near neither 100 nor 159. In this example, a difference of 37 is significant, whereas a difference of 21 is not. Suppose also, that three other elements in the same image had the values 100, 132, 104: with a mean of 112. Now the middle element differs from the mean by only 20 which is not above the threshold we would have needed to use for the previous three elements. The median, however, would have been 104 in both cases: the choice of threshold in this case is much easier. Clearly the median is a better choice than the mean. For reasons discussed in chapter 7, it may be preferable to use a morphological filter to an ordinary median filter.

The median filter, as described above, will not be successful if two of the three neighbouring elements are dead. If this were true, then the "unusual" element would be the element which was not dead. Clearly this would be unsatisfactory. The probability of this undesirable occurrence depends on the number of dead elements in the detector. In the calculation of these probabilities (see appendix C), we have assumed that the intensity of the line read from the dead element differs from the intensity of the lines surrounding it.

If only one element is dead, obviously the problem does not occur. If two elements are dead, then there is a probability of $\frac{125}{2016} \simeq 0.06$ that they will be in the same group of three. If three elements are dead, the probability that two of them will be in the same group of three rises to $\frac{1861}{10416} \simeq 0.18$. This is quite a high

probability and, although we would hope otherwise, it is quite possible that three elements may be dead in an imager. We should therefore consider how to reduce the probability of failing to detect the dead elements.

One way in which we can reduce the probability of failing to detect a dead element is by saying that an element is not dead if the elements on either side of it appear to be dead. If there are two dead elements, this will reduce the probability of failing to detect the dead elements to $\frac{1}{32} \simeq 0.03$; and if there are three dead elements, it will reduce the probability of failing to detect a dead element to $\frac{31}{336} \simeq 0.09$. This method is somewhat ad hoc but would certainly work.

If a median filter is used with five elements instead of three, then the dead elements will always be detected if there are only two dead elements. The probability of failing to detect a dead element if there are three dead elements is the probability that three elements in any neighbouring group of five are dead; which is equal to $\frac{13}{1488} \simeq 0.009$.

Using a median filter with five elements, there is obviously a good improvement in performance. Unfortunately, the cost of performing a median filter of this size is quite large: the processing required is quite extensive. If the calculations are all done in a calibration stage, however, the cost of calculation is not really important.

It was stated earlier that when looking for dead elements, the processing should be done locally. This is necessary because of the possible presence of flickering elements. Because of the noise in the unprocessed image, the values at an instant

in any one detector line are not necessarily representative of that element. It is therefore better to take an average over several pixels. If an average is being taken of a flickering element, the average may well be within the thresholds set for an element being live, even though many of the individual values returned by the dead element are outside the threshold. By locally determining whether an element is outside the threshold for each reading and incrementing a "potential dead element" array if the element appears dead at an instant, then the flickering elements will be detected. An example output of a flickering element is given in chapter 7.

When a dead element is detected, the effect of it is removed by replacing the value of the dead pixel by the mean of the two elements on either side of it. If two or more elements are next to each other, the average taken must be of the two nearest live elements surrounding them. Taking an average is significantly more difficult to implement in electronics than simply copying the value of the previous detector into the dead element position. For this reason, it is likely that an electronic solution in a line scan array would choose this approach, despite the poorer performance. For an imager controlled by a computer, taking an average is just as simple as copying the element.

Our approach to detecting the dead elements was highly successful. By processing the image in the manner described above, we significantly improved the appearance of the images. The images of constant scenes now correctly appeared

homogeneous; while in images of "real" scenes, it was virtually impossible for a human observer to detect which line(s) had been dead.

5.3 Equalising differences between array elements

Once the effects of the dead elements have been removed, it is then possible to begin the process of eliminating the disparities caused by the differences between the detecting elements. Examining the images, it is readily apparent that each detector element behaves slightly differently from the others, causing the image to appear "streaked".

This is because in the raw data, each detector element has a characteristic offset and concomitant gain. The elimination of the differences in offset is accomplished almost completely by IDP, and is described in section 5.1. In this section, we will describe the method for eliminating the differences in gain.

Empirical evidence suggests that the gains do not vary greatly with time: the pattern of the streaking does not appear to change dramatically between frames. This is because the variations in responsivity are primarily a function of geometrical and material uniformity in the detector manufacture process. We shall investigate this further, but as a first approximation it seems reasonable to assume that the gains are fixed. If this is the case, we can use one image to calibrate the imager: calculating a fixed set of gains which we can apply to subsequent images to equalise the elements.

As a calibration image, we used an object which had a uniform temperature and placed it close to the lens such that the object would be completely out of focus. A number of different images of the object were taken at different times and with the object having a variety of temperatures both hot and cold with respect to the chopper temperature.

The gains are all relative. That is, the gain of element two might be twice that of element one; but provided the ratio is maintained, it does not matter what value of gain is chosen for element one. We chose to minimize the change in the data by choosing the average gain to be 1 (i.e. the output of an "average" element will not be changed.)

From the constant image, we calculated the average value of each line of data. If the image was genuinely constant and there were no gain differences between the elements, then the average values would be equal. As we have taken a constant image, any variation in the average values will be due to differences in gain. We are assuming here that the effect of noise in the image is negligible if a sufficient number of pixels are used to calculate the average. Five hundred pixels were used to calculate each average value and so this assumption should be quite accurate.

Calculating the gain differences is then simply a matter of choosing some value which we wish to have for the average intensity of each line and calculating $g[i] = \frac{A_d}{A[i]}$ where $g[i]$ is the gain of the i_{th} detector element, A_d is the desired intensity and $A[i]$ is the average value of the i_{th} detector.

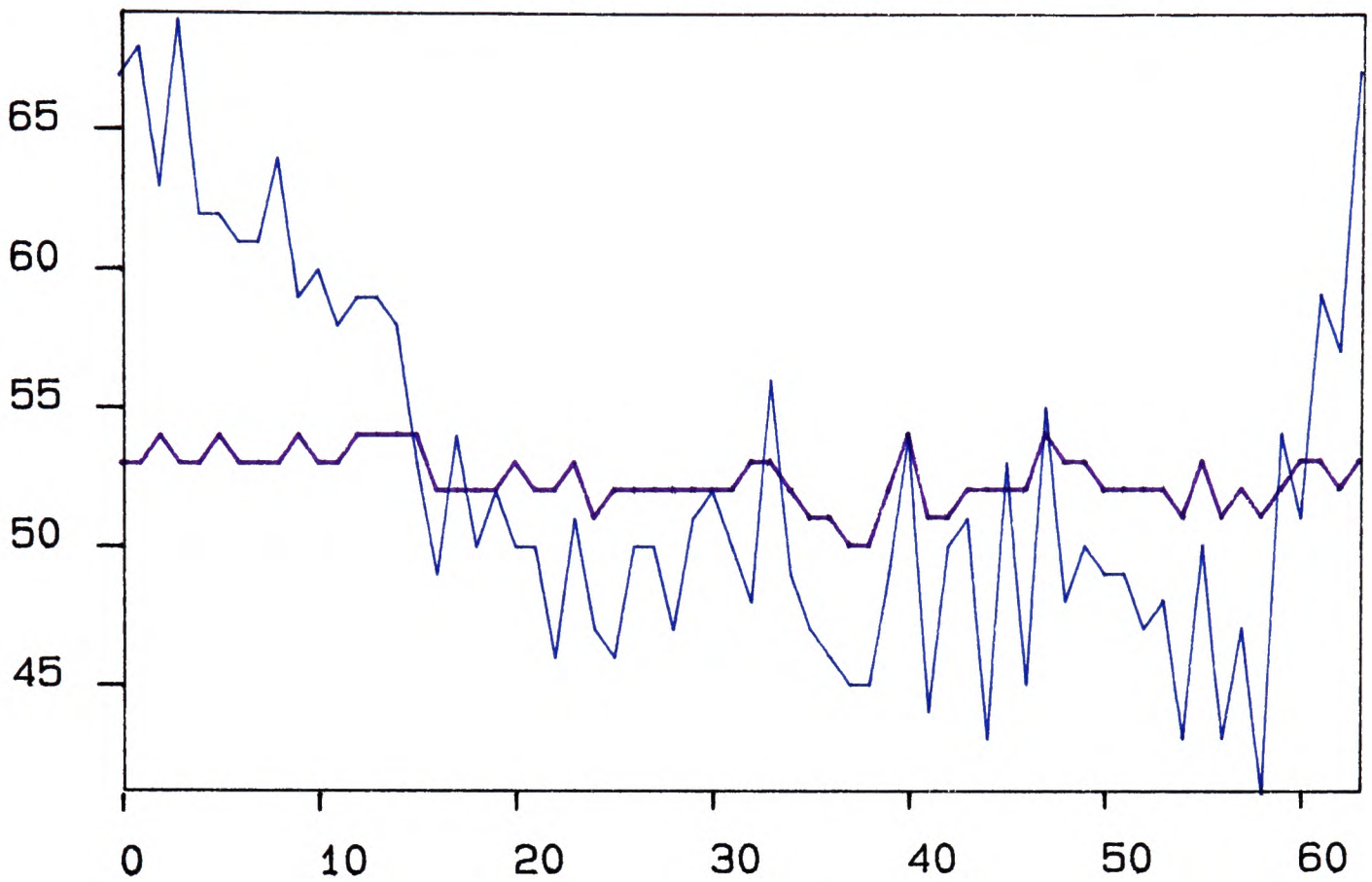


Figure 5.2: Cross section of outputs of detector array for a constant scene both with and without gain correction.

When an image is subsequently taken, the output of the I_{th} detector following IDP is multiplied by the gain $g[i]$ for that element. The result of this process is illustrated in figure 5.2, where the blue line is a plot of the average output of each of the 64 detectors averaged over 10 columns to eliminate the effect of noise; while the mauve line is a plot of the average output of the same pixels for the same data after a gain correction has been applied. The image which was used to calculate the gain offsets was taken at a different time to the image represented in figure 5.2 and was a different constant temperature scene.

To test the assertion that the gain remained constant, we took a number of images over a period of time; and at different temperatures; and calculated the

gain correction figures for each. We plotted the gains of the array for each image and found that the curves superimposed extremely accurately.

5.4 Updating the estimates of offset and gain

In the previous sections, methods were discussed for estimating and then eliminating the offsets and the gains of each of the elements in a linear detector array. These methods provide an image much improved on the original data. Since they require calibration images to be taken, however, the imager must be taken off-line to calibrate it. While an initial calibration period is reasonable, it may be highly undesirable for the user of an imager to cease using the imager in order to recalibrate it, even if the calibration process is rapid; simple to perform; and need only be done infrequently.

As explained, however, empirical evidence suggests that the gains associated with each detector element remain approximately constant over time, regardless of scene and or detector temperature variations. It is therefore possible to precalculate the gains in the manner described above. Any small changes in the values of the gains will not noticeably affect the clarity of the output image.

The offsets of the detector elements do change with time: changes in scene or ambient temperature cause the offsets of the elements to alter with respect to the nominal datum. The effect of these changes in offset will be accounted for, however, in the image difference processing calculations described in the same section.

To summarise, we have found that through the use of image difference processing and pre-calculation of gains, we can accurately eliminate the differences between the detector elements over a period of time - even with changes in the scene and ambient temperatures.

5.5 Summary

In this section, we will summarise which of the preceding techniques we chose to use with our imager.

To equalise the effect of the detector offsets, three-point image-difference-processing as described in section 5.1 was used. This removed the large offsets between the elements.

The image was then analysed using a three element median filter as described in section 5.2 to detect the positions of the dead elements. We used a three element filter rather than a five element filter because for our array, we knew the positions of the dead elements in advance. The position of the dead elements will not always be known a priori, however, so in the general case we would advise the use of a five element median filter.

Having located the dead elements, we then used the techniques described in section 5.3 to equalise the differences between the gains of the detector elements. The outputs of the dead elements were not included in this calculation.

Finally, we replaced the values of the dead elements with the average of the values of the two gain-corrected elements on either side of them.

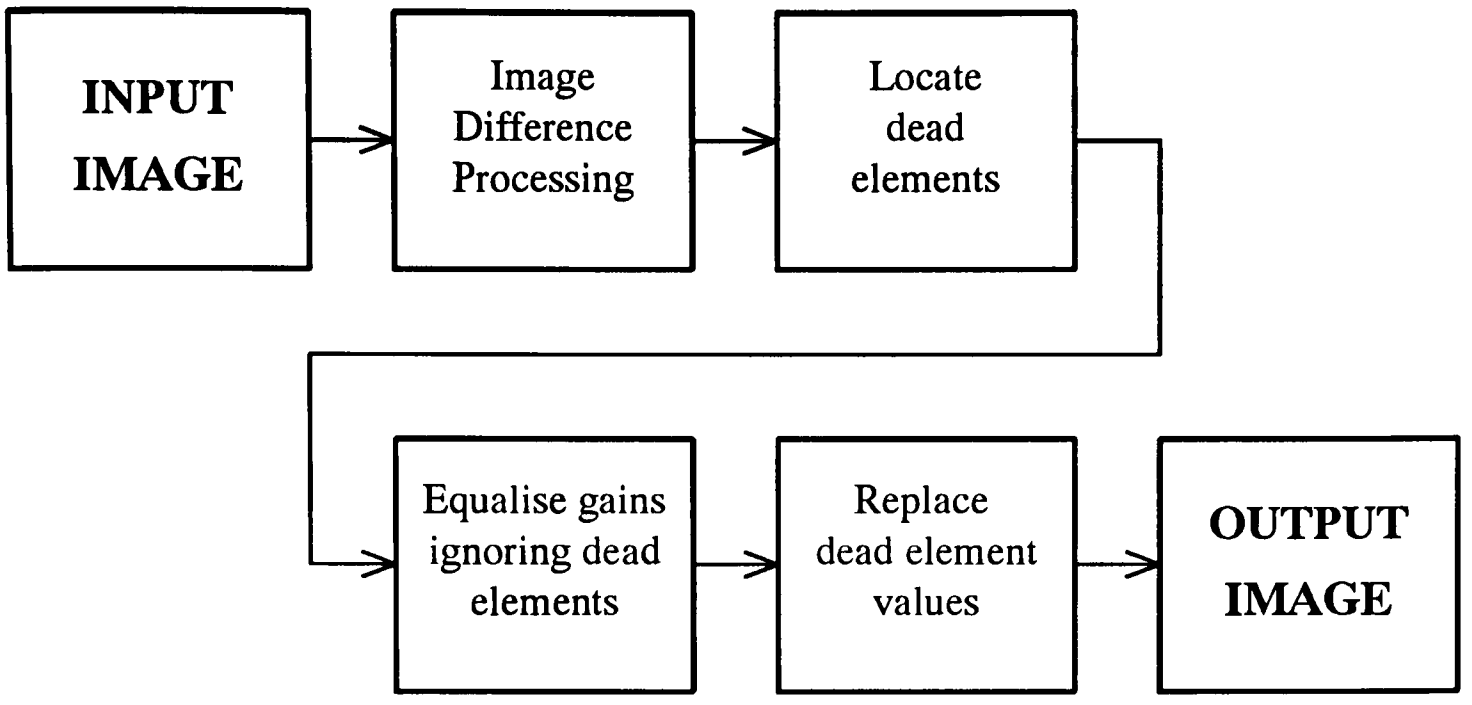


Figure 5.3: Summary of early image processing techniques

Figure 5.3 summarises the order in which the various algorithms were applied.

5.6 Extending gain and offset correction to two dimensions

The ideas discussed in this chapter were based on a line-scanned array detector. The future trend in imaging systems, however, is towards two dimensional arrays of detector elements. In this section, we will examine how the techniques developed in the previous sections can be used to improve the quality of two dimensional images.

The first stage of processing, the IDP, does not depend on the type of array used: it is applicable to two dimensional arrays and is consequently used in current imagers under development. In the two dimensional case, the main difference is

that each combination of two or more fields now corresponds to a complete frame of data rather than to a line of data.

The removal of dead elements and the correction of gain and offset by the methods discussed in this chapter so far do require alteration, however, if they are to work properly on two dimensional images. A simple implementation could consider each column in an image separately and treat a time sequence of each column in the same way that a line scan array was treated. While this would certainly work, it does not fully utilize the additional information present in the two dimensional image. A dead element will now differ not just from its nearest neighbours in the x-direction, but from its nearest neighbours in the y-direction too. A better approach when locating the dead element, therefore, is to consider all nearest neighbours of the element in question.

When replacing the dead elements with more reasonable values, we could either replace the value of the pixel with the value of the previous pixel or with an average of the live pixels surrounding it. This re-introduces the question discussed in section 5.2 of whether to use an awkward to implement but efficacious method; or a method with a poorer performance but which is easier to implement.

The method of averaging the surrounding pixels would be harder still to implement electronically if it were extended to two dimensions, but the effects caused by the other method can be rather bad. If a group of neighbouring pixels are dead, as is often the case due to manufacturing error, the correction approach of

copying the last pixel will fail badly when the last live pixel is on the edge of an intensity boundary: the intensity value on one side of the boundary will be copied into the region on the other side of the boundary. In many applications, the intensity boundaries are extremely important and any confusion caused by this kind of processing could cause severe difficulties. For the same problem, the averaging method would blur the edge, but would not produce a misleading image.

A compromise method could be implemented with only a very slight delay in data throughput by storing the last live value of data; waiting for the next live value of data on that line and averaging those two values. This would be far simpler to implement and would under almost all conditions cope with the problem described above.

One major problem introduced when changing from one dimensional arrays to two dimensional arrays is the huge increase in data. The number of data points arriving per frame in a sixty four squared array will of course be sixty four times as many as in a sixty four element linear array.

In order to carry out any analysis of the images over time on-board the device, it would be necessary to store and process multiple frames of data. To store only ten frames of a one-hundred by one-hundred array would mean storing one hundred thousand pixels. (Compare this with only one thousand pixels for ten frames of a one hundred element linear array.)

The time series data used for gain-correction as described earlier in this chapter

is only required in the calibration stage. It is therefore much simpler to read the data for the calibration stage into an off-line computer and store the gain values permanently into the memory of the device. This is the approach adopted in [56]. Initial dead element positions can also be stored in a dead element array in the calibration process.

Chapter 6

Experimental Results

6.1 Introduction

In this chapter, some of the results obtained when using the imager are presented.

The images shown are chosen to represent typical images taken with the imager: they are not the "best" images. Unless otherwise stated, no processing has been done on the images with the exception of that described in chapter 5.

The majority of the images are taken of objects positioned approximately five feet from the camera lens. The lens used for most of the time was a fixed focus lens which differed slightly from its specification, as described in section 4.10. Some images were taken with a different lens, however, which was on loan from RSRE for a couple of hours. A selection of these are presented to illustrate how choice of lens affects the type of images captured.

During the development of the imager, various problems caused a variety of unusual effects in the images. Only those effects which had some bearing on the design of the imager or subsequent software processing will be discussed; and in

those cases, the images or image descriptions will be presented.

The second part of this chapter deals with the interpretation of the images. The images are analysed to see if they produce the patterns which we would expect from the theoretical analysis in chapter 2.

6.2 Performance of Imager

Because the imager has a long frame time, objects which move in the scene are not immediately recognizable. An object moving in the same angular sense as the mirror will appear much larger than it actually is, while an object moving in the opposite direction will appear smaller. This problem will occur in any imager which involves a scanning mechanism unless the angular scan velocity is very much larger than angular velocity of the moving object. For this reason, the majority of the images taken with the imager were of stationary objects.

As mentioned earlier, the lens used for the vast majority of the time was a fixed focus lens. Some attempt was made, when necessary, to vary the focal length of the device by moving the lens in and out on the Wreathall mount and by changing the length of the pillars supporting the imager mount, but the majority of the images were taken at a fixed distance and fixed magnification. This reduced the choice of objects which were suitable for imaging. Objects about the size of a cup were about the right size, so when suitable objects were not available, it was possible to manufacture the required shape without too much difficulty. Variable focus infrared lenses are available but are rather expensive. A variable focus zoom

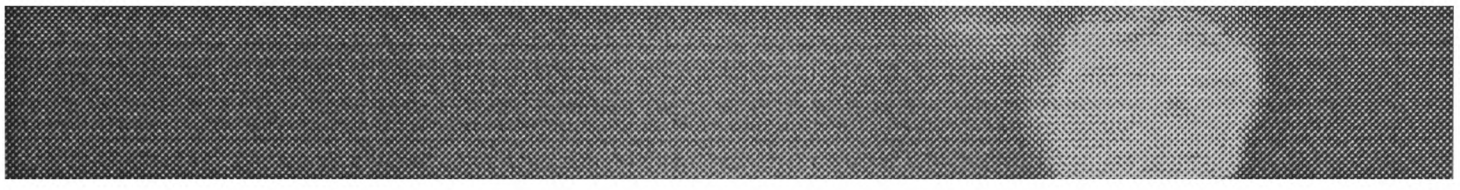


Figure 6.1: Head image

lens would have been very useful for this research work, as it would have been advantageous to examine images of objects with a variety of sizes and at a variety of ranges. For most applications, however, the focal length will be fixed: the majority of military applications require the imager to be focussed at infinity; while many inspection processes require a constant distance between the imager and the objects under inspection.

Figure 6.1 shows an image of a human head taken using the normal lens. Only a part of the head is contained within the image and because the person was not exactly the correct distance from the lens, the image is slightly out of focus. Compare this image with the image taken using a different lens in figure 6.2 where a far larger scene is being viewed. We notice that the focus is better for this second image. Being further from the lens, the depth of field is greater so it is easier to position the object where it will be in focus. The noise is worse in figure 6.2 but this has nothing to do with the lens: the camera had not been fully tuned when this image was taken.

In section 5.2, dead elements were described and methods for removing them were discussed. Even after tidying up the electronics of the imager, two dead elements remained. In addition to this, on occasions, one or more of the other

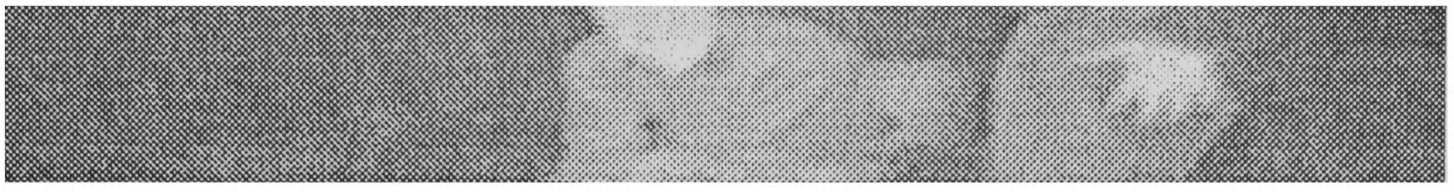


Figure 6.2: Body image

elements would appear dead due to solder joints drying up, components failing, or loose connections appearing. Most of these "dead elements" behaved in the same way as the real dead elements, but sometimes the element flickered: a problem discussed in section 5.2 but rarely addressed.

Figure 6.3 shows one image which was taken when an element was flickering. The plot in figure 6.4 shows the horizontal cross section of the intensity of the background part of the image along the flickering element and a cross section of the image along the pixel neighbouring the flickering element. The standard deviations of the two plots in figure 6.4 are 60.38 (flickering element) and 3.29.

At different times, flickering elements appeared producing different magnitudes of flicker. Figure 6.5 shows an intensity profile of a flickering element before and after the image was filtered with the program described in chapter 5.2. This intensity profile shows far less variation about the "correct intensity" (notice the different vertical axes on graphs 6.4 and 6.5) but, as can be seen from figure 6.5, the program still found the dead element and corrected its output. In this image, the standard deviations of the two plots were lower than in the previous image: 10.39 for the flickering plot and 1.81 for the corrected plot. We would expect the standard deviation of the corrected plot to be lower than the standard deviation

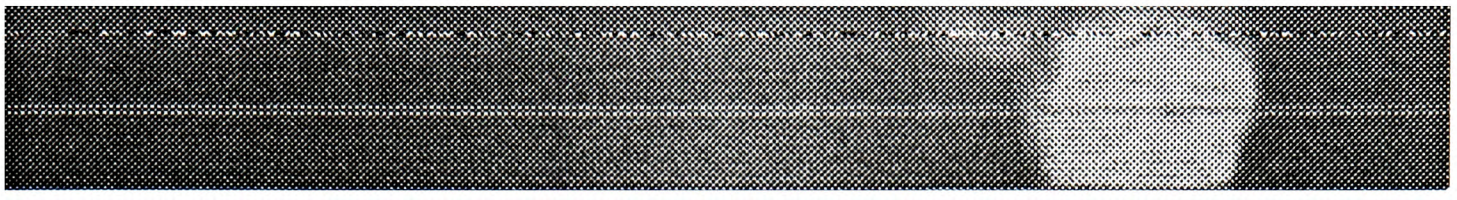


Figure 6.3: Image containing flickering element

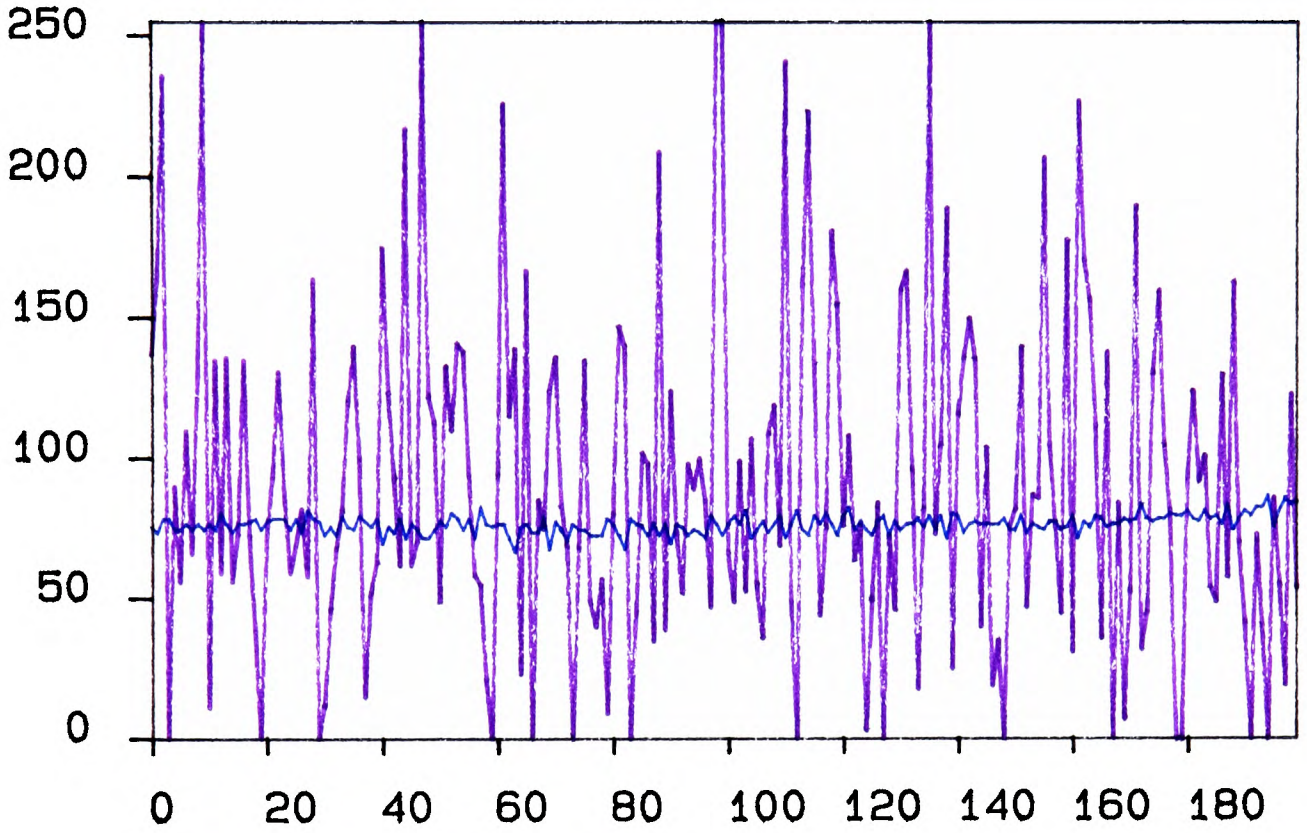


Figure 6.4: Cross section along flickering element. The violet plot is the flickering element; the blue plot is a neighbouring element.

of the uncorrected non-flickering plot because it is the average of two plots. We should also note that standard deviation could not be used to detect the presence of dead elements from non-calibration images, since the standard deviation of a noise free image containing two different intensities would be high for both flickering and non-flickering elements. The result of filtering the image in figure 6.3 to remove dead elements is displayed in figure 6.1.

Clearly the program has successfully removed the effect of both the real dead

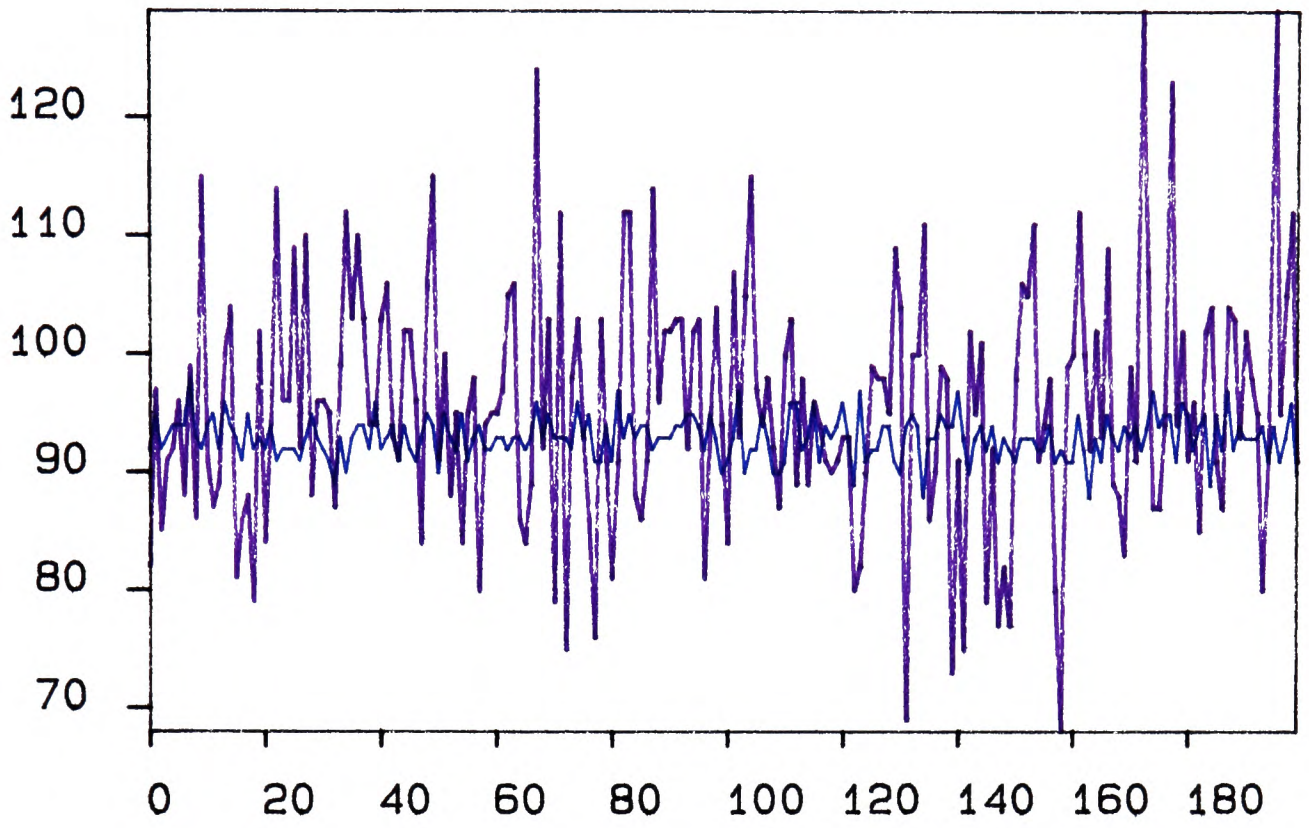


Figure 6.5: Cross section across image before (violet) and after (blue) dead element correction

elements, and the flickering dead element. Removing the flickering element as well as the constant dead elements significantly improves the appearance of the image and removes one of the large obstacles preventing good automatic recognition of the images. Although flickering elements are quite rare, this approach to dead element removal is no more computationally expensive than other reasonable approaches and can be implemented in parallel. As it performs better without undue computational cost, we opted to use this approach rather than any other.

When plotting a cross-section of an image, we noticed that one aspect of the fixed pattern noise in the unprocessed images was that the detector elements higher in the array tended to give intensity values higher than those lower in

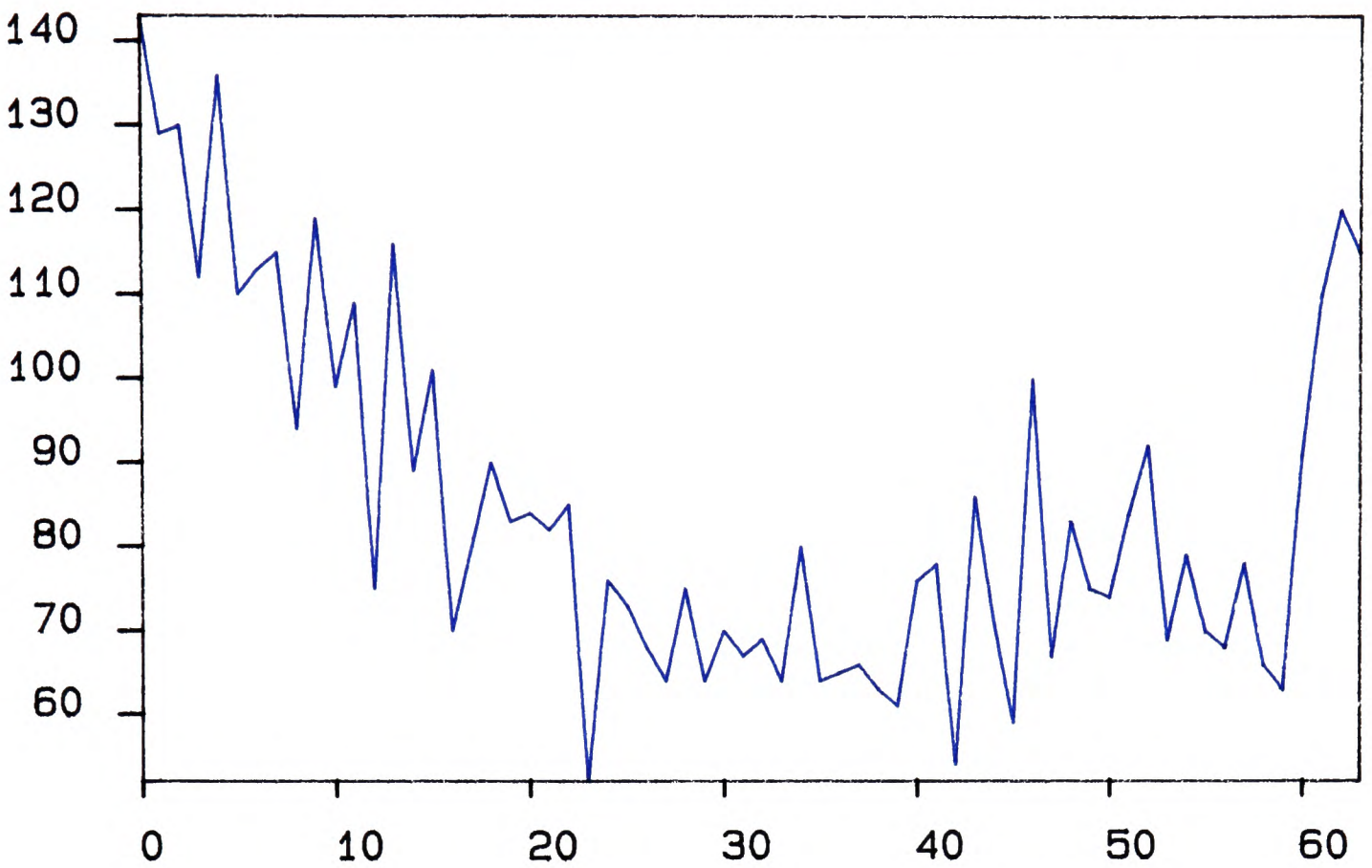


Figure 6.6: Variation of intensity across the detector array

the array: the change between these two being approximately linear. Figure 6.6 illustrates this. This figure shows the intensity of each detector element in a randomly chosen column of a constant intensity image when no processing has been done on the image. Ignoring for the moment the local fixed pattern noise on the signal, it appears at first sight that the plot can be divided into three regions: a region where the intensity is falling; a region where the intensity is approximately constant; and a region close to the edge of the detector where the intensity rises quite sharply. Examining a number of intensity profiles, however, it appeared that the data was better described by an exponential decrease followed by a sharp increase in intensity.

When this effect was first noticed, it was assumed that the cause was the variation in temperature of the target. Further images were taken ensuring that the object had a constant temperature throughout. The effect was still apparent. It was possible that the lens was providing more radiation throughput selectively or that radiation was entering the electronics housing via another route and falling on the detector array. Neither rotating the lens nor altering the packing inside the housing removed the effect, however, so we discounted these possibilities. Having eliminated the obvious causes, the most likely cause was either the difference in response between the detector elements or a timing problem in the electronics.

Because of the global nature of the effect, it seemed more likely that it was being caused electronically rather than by differences in the detector elements; although it is possible that when the detector was manufactured, uneven pressure across the array caused the detectivity of the elements to vary approximately linearly across the array. Analysing the electronics, it seemed possible that the choice of position of the opto-pickup could affect the output in this way if the radial position of the detector array was not accurate. This would cause a small difference between the period that the detector elements spent open to radiation.

For the reasons described in section 4.5, the detector array is open to radiation for a period longer than the time it requires to read the radiation. If the position of the detector array is accurate, the system will work correctly: regardless of where in the chopper cycle the electronics begins the read routine, the elements

will all receive the same amount of radiation. If the array is not positioned, correctly, however, the amount of radiation falling on each of the detectors will vary linearly. Although the mountings for the chopper and the detector element were constructed carefully to the prescribed sizes, the specification for the size of the detector array was a little vague. If the size of the array was not exactly as given in the specification, this would explain some of the effect we noticed. The error could be caused by the pitch of the detector elements not agreeing with the specification; or by the array of elements not occupying the centre of the detector.

If the amount of radiation falling on the elements varies linearly, then the signal provided by those elements will vary exponentially as described in chapter 5. We would therefore expect a cross section of a constant image to show properties similar to those displayed in figure 6.6 for pixels 0-55. This could also explain the sharp rise in intensity shown on the right of the plot if the opto-pickup is poorly positioned and the last few pixels are read after the chopper has closed them from radiation. If this were the case, the signal would decrease rapidly in an exponential fashion.

If this argument is correct, the position of the opto-pickup will affect the shape of the cross-sectional plot. If we move it forwards (in the same direction as the rotating chopper), the detector will be read later and so the effect of the exponential on the earlier elements will be lessened and the position of the cutoff will be moved further into the array. Similarly, if the opto-pickup is moved backwards,

we would expect to see a more noticeable effect of the exponential on the earlier elements and to see the cutoff move nearer the edge of the array; possibly to remove the sharp increase altogether.

Images were taken of a constant temperature object and the position of the opto-pickup was moved by a small amount both clockwise and anticlockwise from its original position. Figure 6.7 shows the cross-sectional plots of several images taken with the opto-pickup in different positions. The plots in this figure represent the whole range of images that could be taken by varying the position of the opto-pickup. If the opto-pickup was moved much beyond the most forward and most backward positions used in creating these plots, then the image became badly degraded. The plots have been offset from each other so that they do not overlap and positioned so as to retain the order in which the images were taken: with the top plot being a plot of the image taken when the opto-pickup was in its most forward position etc.. A local average has been done on the intensity plots to remove the worst of the noise but the plots are otherwise unaltered.

The first thing to notice is that the position of the exponential curve is changing as the opto-pickup is moved. To illustrate this more clearly, a graph was plotted of the estimated position on each curve where the plot appeared to be flattening out. Although it is somewhat difficult to measure this accurately, it is perhaps one of the only measurable properties of the plots; and does have a physical significance: it is the point on each graph at which the same amount of signal integration

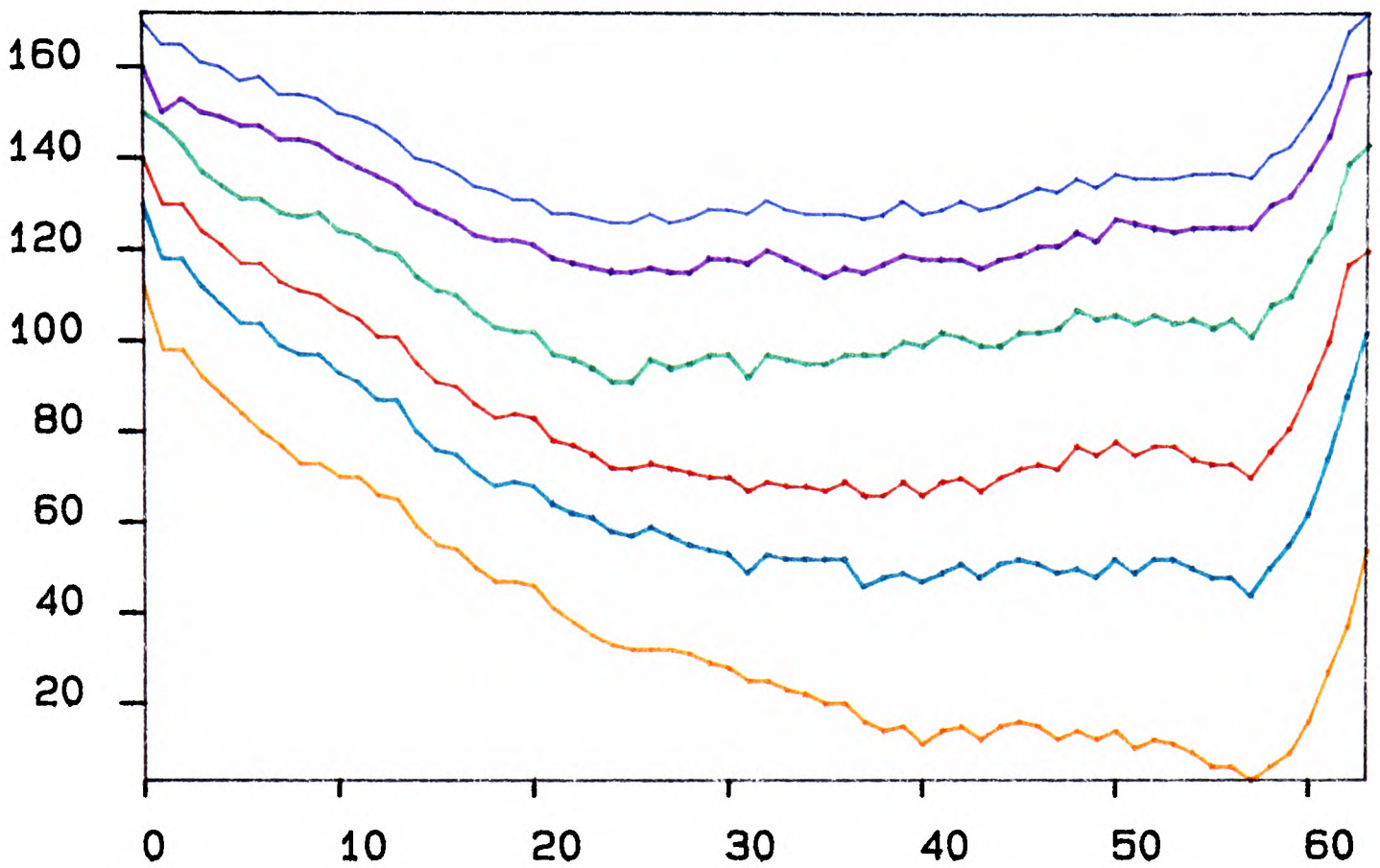


Figure 6.7: Variation of intensity across the detector array for different positions of the opto-pickup

has occurred. Figure 6.8 shows how the estimated position of the flattening of the curve varies with the position of the opto-pickup. If the above theory were correct, then we would expect that the output of the same detector for each plot would vary in intensity with respect to the others in an exponential fashion. We would also expect the point where the flattening of the curve takes place to vary exponentially with the position of the opto-pickup since the time each element is exposed will vary linearly with the position of the opto-pickup. Justification for this assertion is given below. The results plotted in figure 6.8 agree with the interpretation given above.

Figure 6.9 schematically illustrates a window of a chopper passing across a

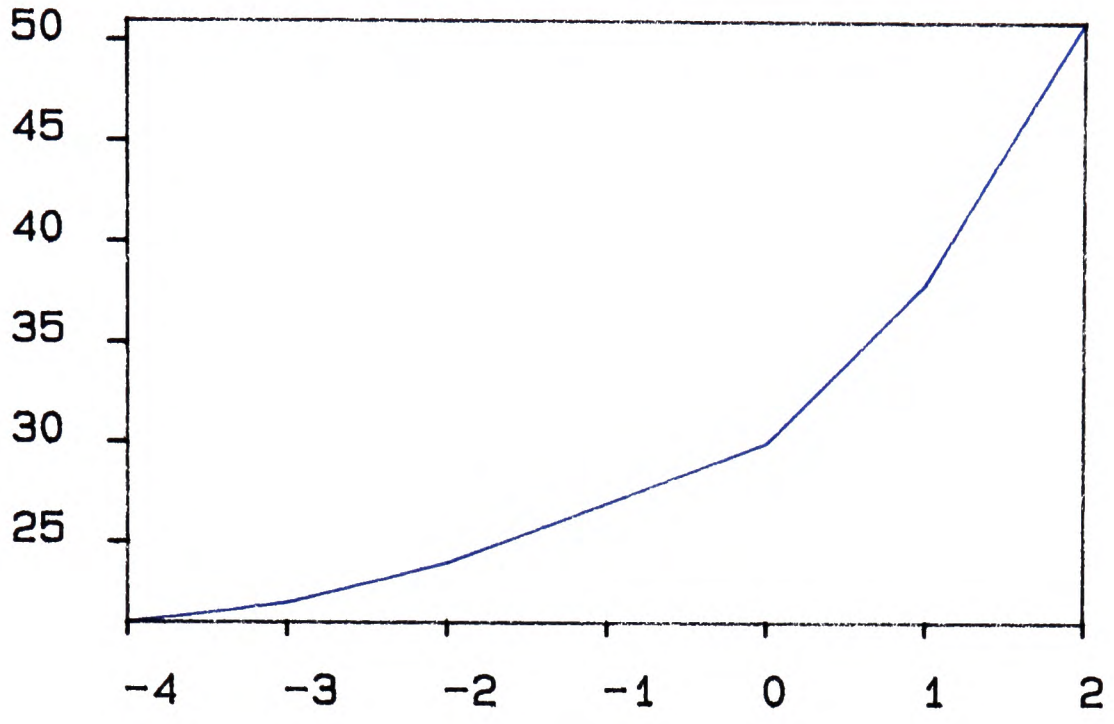


Figure 6.8: Variation of position of curve flattening against pickup position

detector array. In each diagram, the small circle represents the detector element being read at that particular time; the rectangle containing the circle is the detector array and the triangle represents the window in the chopper passing across the array. The upper two diagrams illustrate the position of the array when the first element is read: the position of the opto-pickup is different in these two so that the electronics begins reading data at different times. This is illustrated by the different lengths ae and be where e is the length of a detector element and a and b are numbers of elements. The lower left figure is of the first opto-pickup position a time t_1 later and the lower right figure is of the second opto-pickup position at a time t_2 later such that the element being read at that time has been exposed for the same length of time as the element in the lower left figure.

In these diagrams, the array is shown to be moving rather than the chopper. This makes the drawing and mathematics simpler, but does not affect the result.

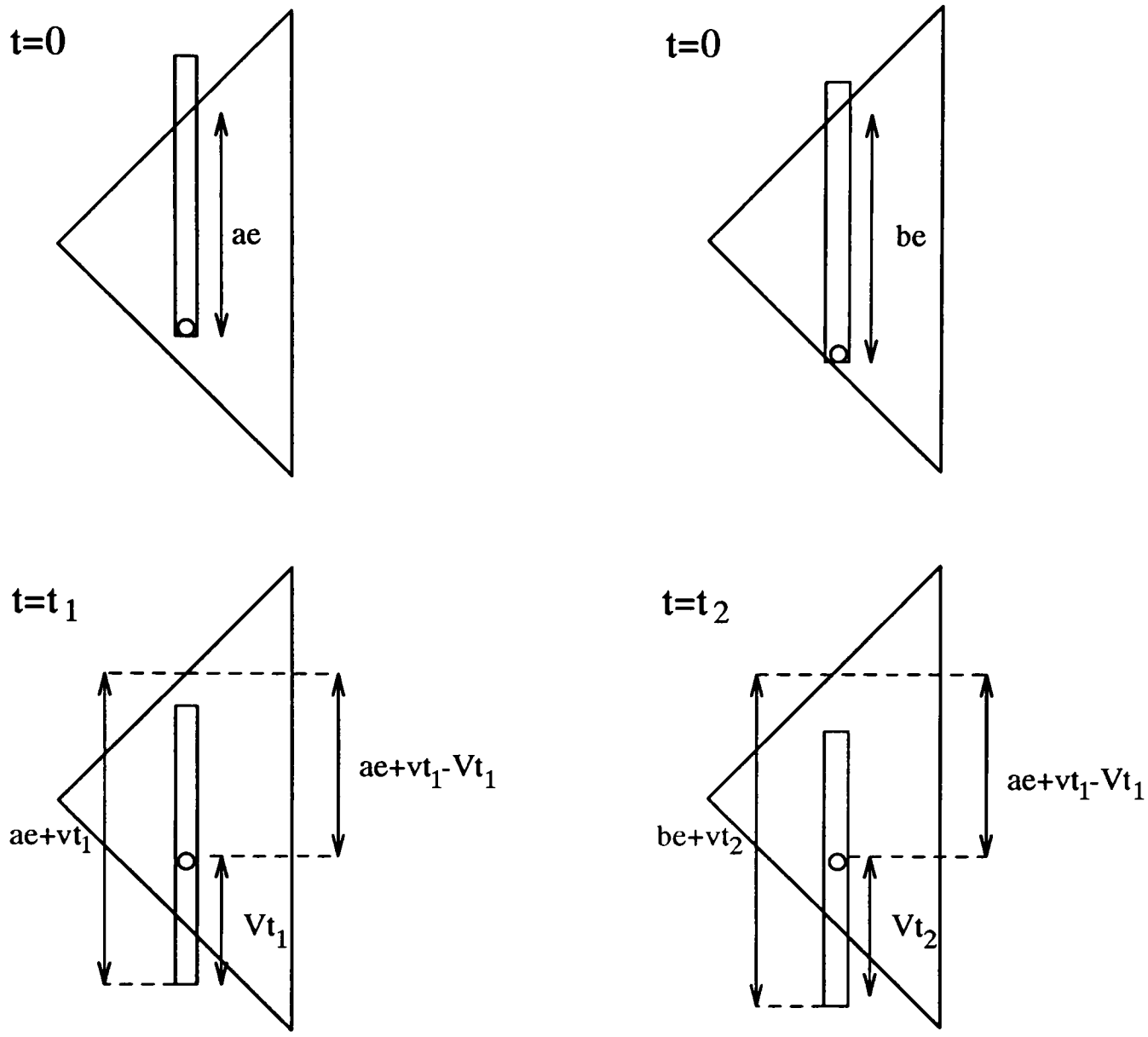


Figure 6.9: Illustration of chopper blade moving across detector array for two different opto-pickup positions

In each case, the array is assumed to be moving at a radial distance r with a velocity v, where $v = V\frac{r}{R}$ and V is velocity of an array in the correct radial position R. Equating the distances on the lower right diagram and tidying up, we have: $t_2 - t_1 = \frac{e(b-a)}{(V-v)}$. In this problem, we are considering an unalterable value of $(V-v)$, so $t_2 - t_1$ is proportional to $(b-a)$ the distance moved by the opto-pickup. We have therefore verified the assertion made earlier.

Clearly, the position of the opto-pickup is affecting the output of the early detector elements in a way consistent with the ideas expressed above. From the plots in figure 6.7, however, it does not appear that the position of the opto-pickup affects the point at which the sharp cut-off in intensity occurs.

From these results, we concluded that the sloping across the array was due to the timing error introduced by the mispositioning of the detector array and that the effects at the lower edge of the array were caused in the manufacturing process. Indeed, this explanation is consistent, since the manufacture of this device involved pushing the two components of the hybrid together. It is quite possible that at the edge, insufficient pressure or too much pressure was applied causing elements with inferior detectivities. If we were able to move the array radially with respect to the chopper motor, we could have compensated for the sloping of intensities across the image; but not for the sharp increase at the bottom of the array. It is of course possible to correct for this effect in software.

6.3 Observations and Analysis of Images

Cups were often used as test images for the imager. They proved to be highly suitable; not only because they were a good size for imaging, but also because their temperature could be controlled quite easily by filling them with water at different temperatures; and because it was possible to model accurately their geometry. Because cups are cylindrical, images of cups can be used to investigate how the intensity of the radiation from the surface varies with surface orientation.

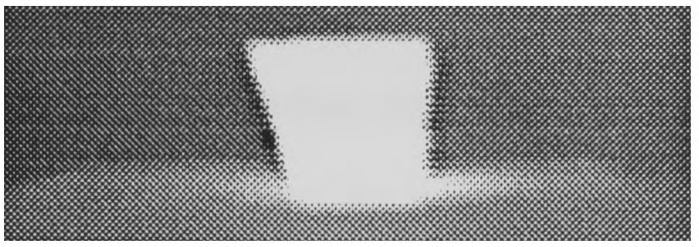


Figure 6.10: Plastic cup

Figure 6.11: Metal cup

Figure 6.10-6.12 show a selection of images of cups. The figures illustrate the difference in appearance between different cups at similar temperatures which have been left to stand for similar periods of time. The handle of the metal cup stands out very clearly from its background, whereas the handle of the ceramic cup is far less visible: showing the difference in conductivity of the different materials. The base around the plastic cup (a cup dispensed from a coffee machine) shows evidence of large heat losses from the cup due to radiation and conduction. This is indicative of the lack of insulation in this type of cup (making them somewhat unsuitable for the purpose for which they are used).

Figure 6.12: Ceramic cup

A graph was plotted of the intensity profile of an image of the ceramic cup to investigate how the intensity of irradiation varied with the direction of the surface normal. The graph was quite noisy, however, and the cross-section of the cup was quite narrow, so interpreting the graph was difficult. Further images were taken of the same cup but with the mirror of the imager moving very much slower so that the cup appeared elongated in the image. A cross section of one of these images is given in figure 6.13. Although this graph is still fairly noisy, the plot is wide enough for us to see that over the central part of the cup, the intensity is approximately constant. At the edges, the intensity falls off quite rapidly. This agrees with the theoretical analysis in chapter 2. In these images, care was taken to shield the cup from any external heat sources as thermal reflections would have distorted the results. We chose shields which did not reflect thermal radiation themselves, as this would have introduced problems of the cup being irradiated by its own radiation.

We wished to examine the intensity profile across a concave cylinder and compare it with that of the cup (a convex cylinder). It was not possible to view the cup in this manner, so a target was made from a section of brass piping which could be viewed either as a concave cylinder from one side or a convex cylinder from the other. Images were taken of the target at different temperatures with both the convex side facing the imager and the concave side facing the imager. The surface of the target was then made black and the measurements were repeated.

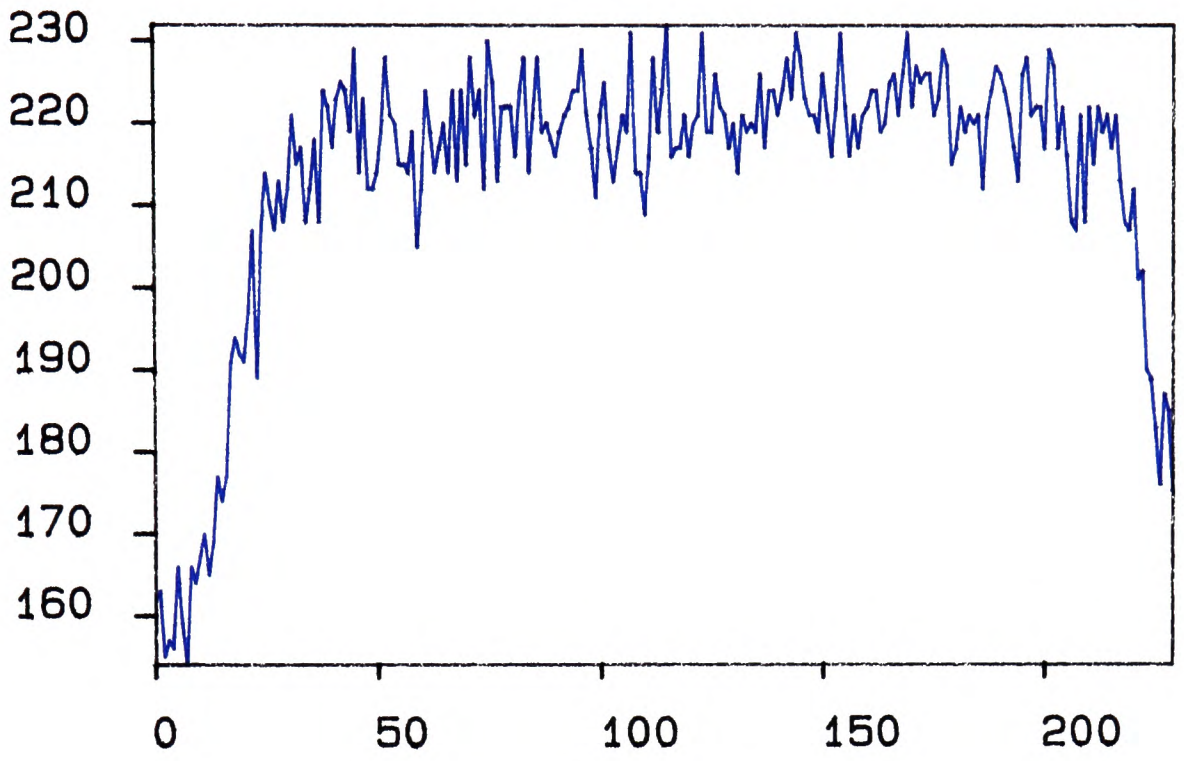


Figure 6.13: Cross-section of elongated image of ceramic cup

Figure 6.14 shows cross-sections of the intensity of the target when made black from the convex side. The blue line is with the target at $20°C$, the mauve line is with the target at $29°C$, the green line is with the target at $33°C$ and the red line is with the target at $46°C$.

Figure 6.15 shows cross-sections of the intensity of the target when made black from the concave side. The blue line is with the target at $20°C$, the mauve line is with the target at $29°C$, the green line is with the target at $33°C$ and the red line is with the target at $41°$. The intensity profiles of the two hotter objects in both figures have been offset by 220 intensity units with respect to the cold objects so that both can be displayed with reasonable resolution on the same axes. The background temperature for these experiments was approximately $31°C$.

Figures 6.14 and 6.15 show that the theory described in chapter 2 for black

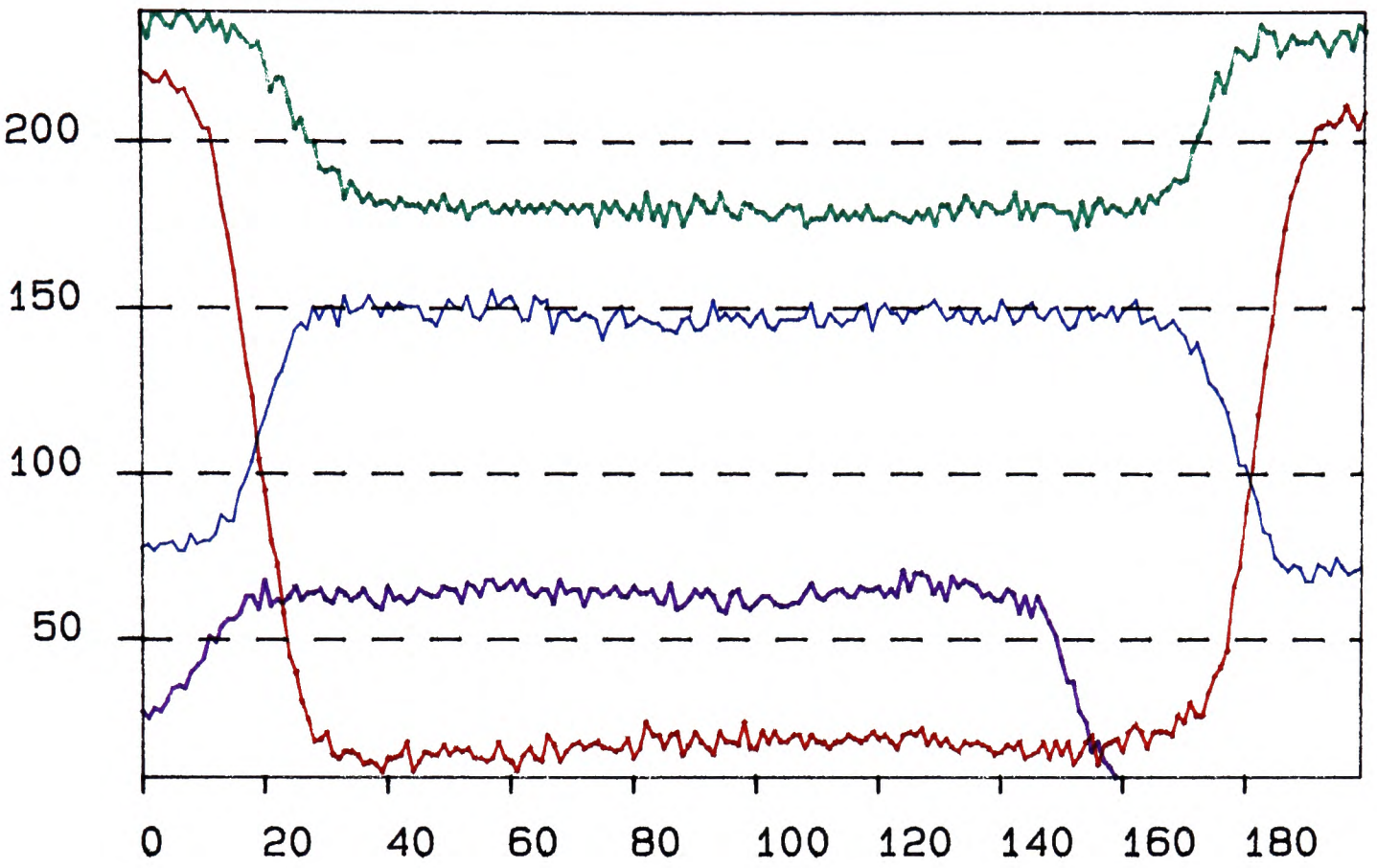


Figure 6.14: Cross-sections of convex black body

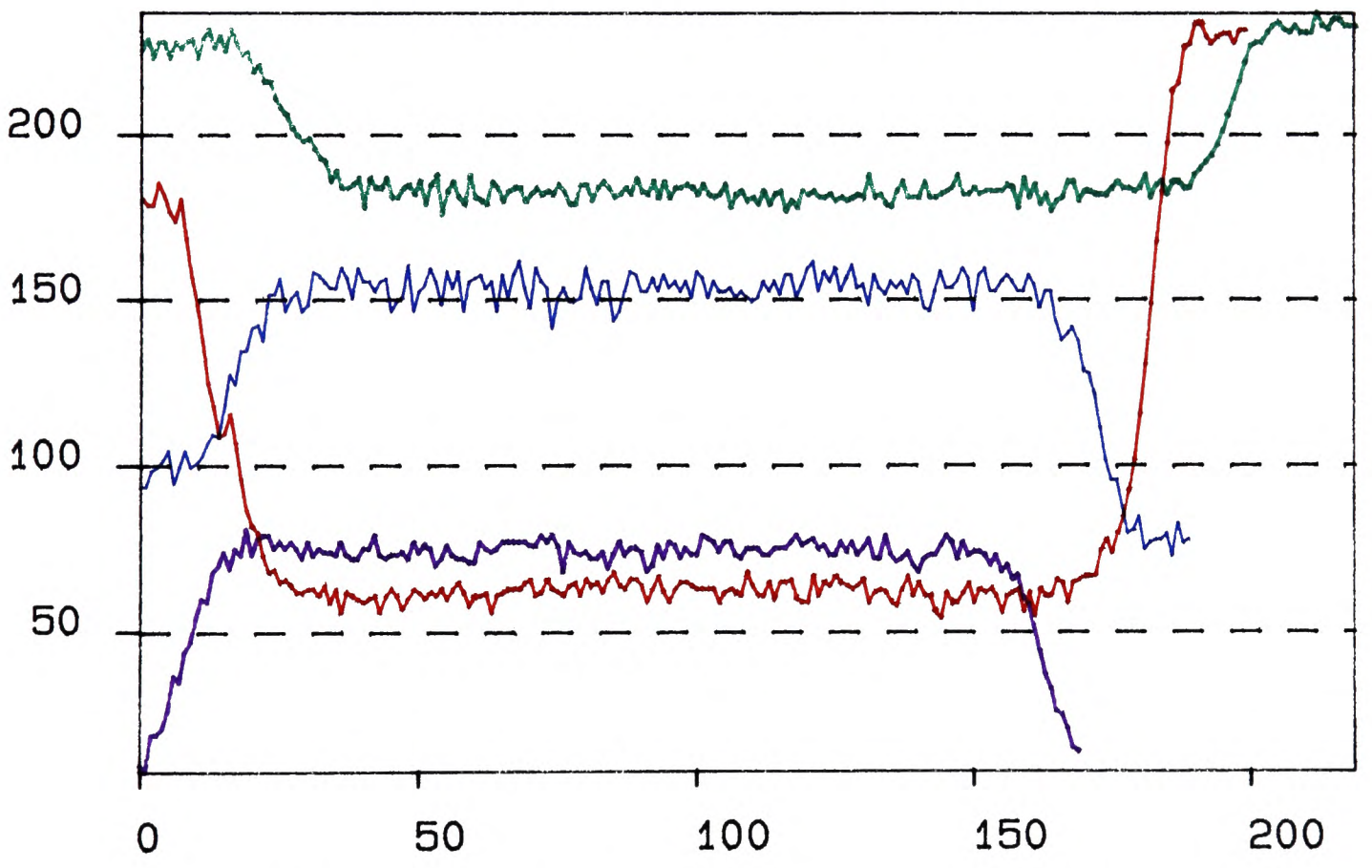


Figure 6.15: Cross-sections of concave black body

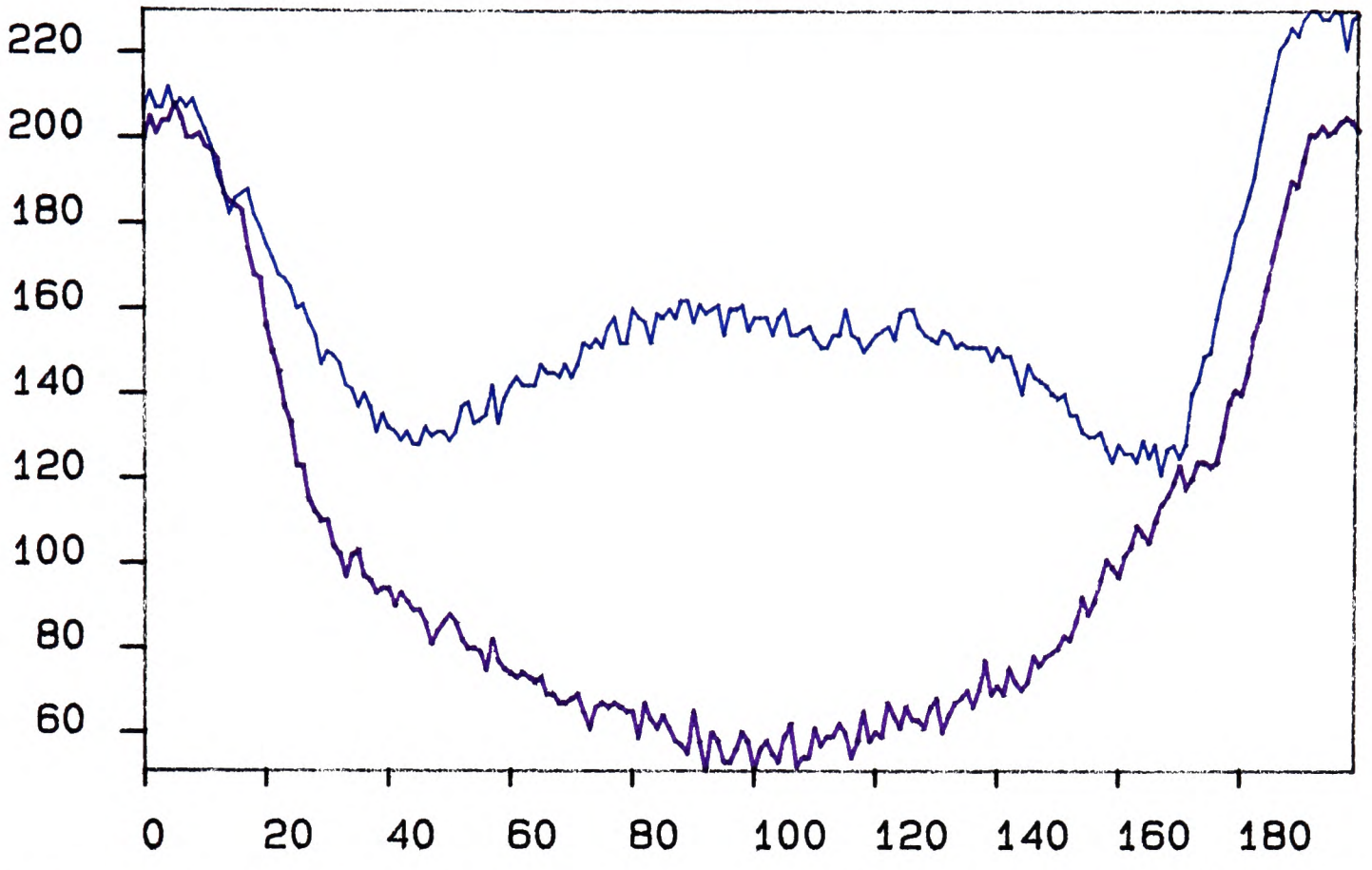


Figure 6.16: Cross-section of convex (mauve) and concave (blue) non-black bodies

bodies does appear to be valid: the intensity profiles are very nearly flat for both concave and convex objects indicating that the irradiation at the imager is independent of the surface orientation *for black bodies.*

The intensity profiles of the images taken of the target before it was made black are very different from those of the black target. In this case, there was a specular reflective component in the radiation from the surface of the body. Figure 6.16 gives typical intensity profiles for the non-black targets. The mauve line is the intensity profile across the image of the convex target at $46°C$ and the blue line is the intensity profile of the image of the concave target when it was at $41°C$. The difference between the two profiles is due to the mutual illumination effects discussed in chapter 2 and confirms the theory presented there.

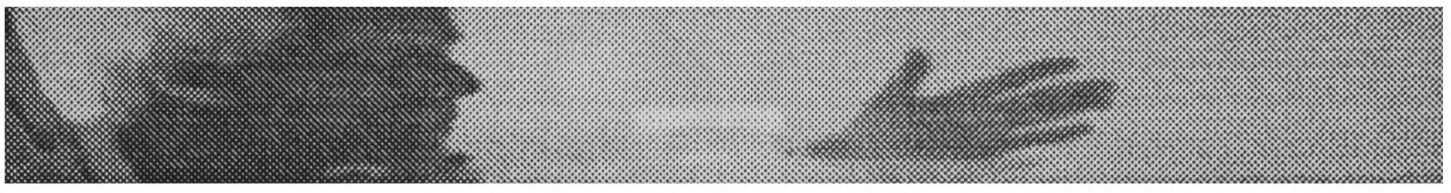


Figure 6.17: Image containing convex cylinder

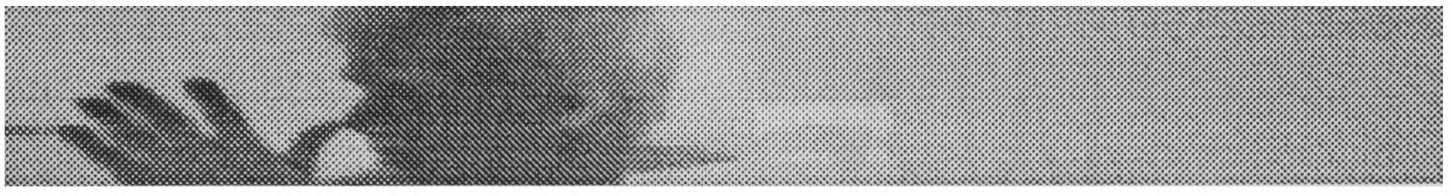


Figure 6.18: Image containing convex cylinder

6.4 Image recognition

In the previous section, we demonstrated that certain different object shapes and surfaces have characteristic intensity profiles. In this section, we will demonstrate how knowledge of these intensity profiles can be used to recognise the existence of a particular object.

A series of images was taken containing cool convex non-black cylinders along with other objects which were both warmer and cooler than the environment. Figures 6.17 and 6.18 are two typical examples. In both of these images, a black body cylinder at the same temperature as the non-black cylinder lies immediately above the cylinder indicating the position of the cylinder (just to the left of the hand in figure 6.17 and to the right of the head in figure 6.18).

The non-black body cylinder has a radiance pattern similar to that of the plot of the intensity of the convex body in figure 6.16, while the intensity profile of the black body is approximately constant. Figure 6.19 shows horizontal intensity

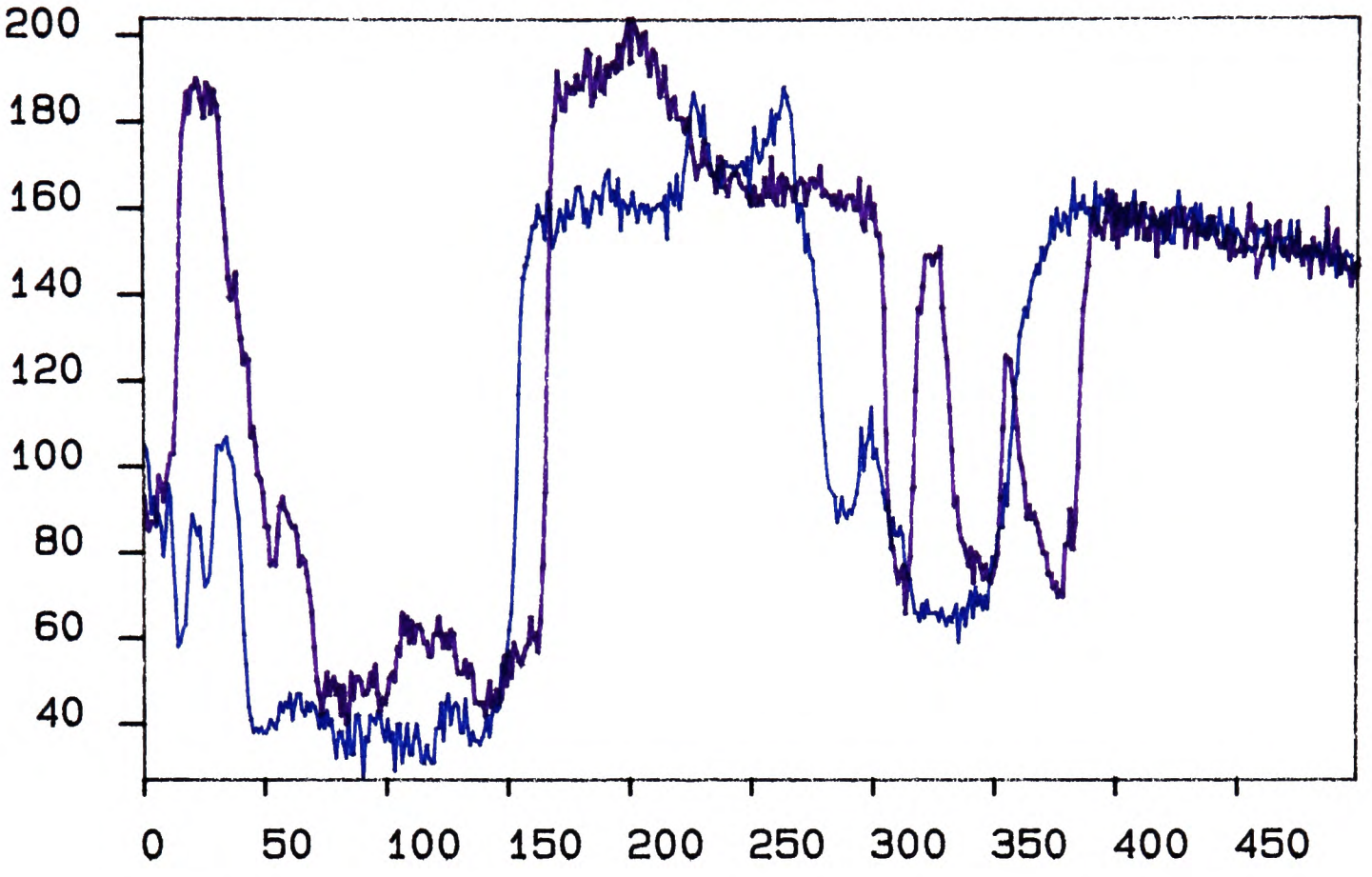


Figure 6.19: Two intensity cross-sections from the image in figure 6.17

profiles of two lines of the image in figure 6.17. The blue plot is a cross section which crosses the non-black cylinder, while the mauve plot is a cross section from a part of the image above the black cylinder.

The purpose of the demonstration described in this section was to identify the cylinder which occurs in the blue plot in about the centre of the image using only information about its shape characteristics. Examining the plots in figure 6.19, and other cross-sectional plots, we decided that it would be possible to estimate the position of the cylinder by detecting the position of a "cylinder-like" profile in the one dimensional signal. The algorithm which is briefly described below was not designed to be useful (without modification) for any other problem: the sole purpose was to demonstrate the feasibility of object detection using intensity

information from different surface properties within a body.

The first stage of the processing was to filter the image to reduce the amount of noise. This was done using a simple Gaussian filter. We removed the background by convolving the image with a simple edge detector and eliminating points with zero edge strength.

We then searched the remaining points for occurrences of pairs of edge gradients: a negative edge gradient separated from a positive gradient by a distance within a specified range. Once a negative edge gradient had been encountered, a search was made for a neighbouring positive edge gradient. Starting at a point beyond the maximum width of a cylinder, the pixel values were searched backwards until a positive gradient was found with an intensity value not too different from the intensity value of the point with the negative edge gradient. If such a point was found before the distance between the points became less than the minimum expected width of a cylinder, the edge was retained as a possible cylinder edge; otherwise the point was removed from the search space. A fairly generous size margin was permissible between the edges of the cylinder to allow for variations in mirror speed, distance of target from the imager etc..

Further constraints were then applied to the remaining set of possible candidates. A particular set of intensity values could only be classified as a cylinder if the intensity of all points between the negative gradient edge and the positive gradient edge in the smoothed image had an intensity which was lower than the

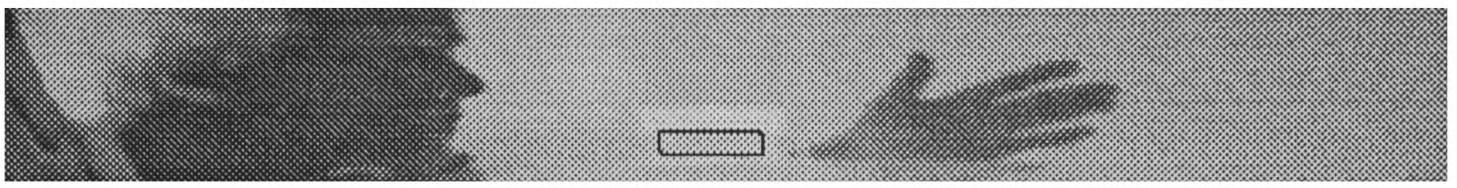


Figure 6.20: Convex cylinder located

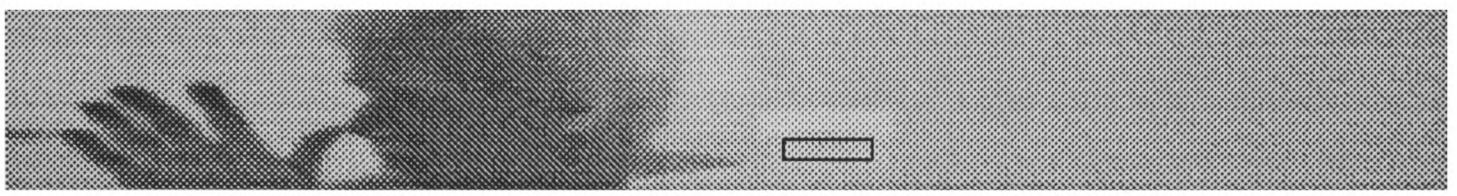


Figure 6.21: Convex cylinder located

higher intensity value of the surrounding edge points. Similarly as the characteristic cross section of intensity values from a particular type of cylinder would not vary greatly if the temperature of the cylinder fell within some reasonable range, a region in the image would not be classified as a cylinder unless the intensities of points between the two suspected cylinder boundaries were neither too far from nor too near to the average intensity value of the surrounding points.

Initially, the thresholds which were set on all the parameters mentioned were very weak. We found that using this algorithm, we could detect the required target with only a few false targets. By tightening the thresholds, it was possible to reduce the number of false targets reported. Figure 6.20 shows the target from figure 6.17 located with no false targets and figure 6.21 shows the target in figure 6.18 located, again with no false targets.

To demonstrate the robustness of the algorithm, we ran it on a variety of images with different background conditions and different target temperatures. To detect

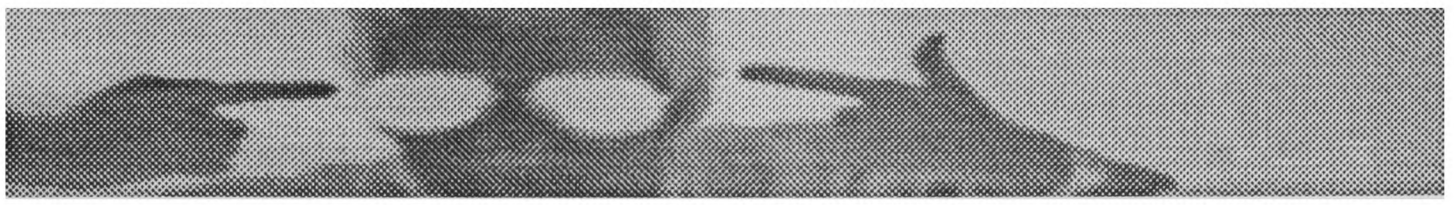


Figure 6.22: Image containing convex cylinder

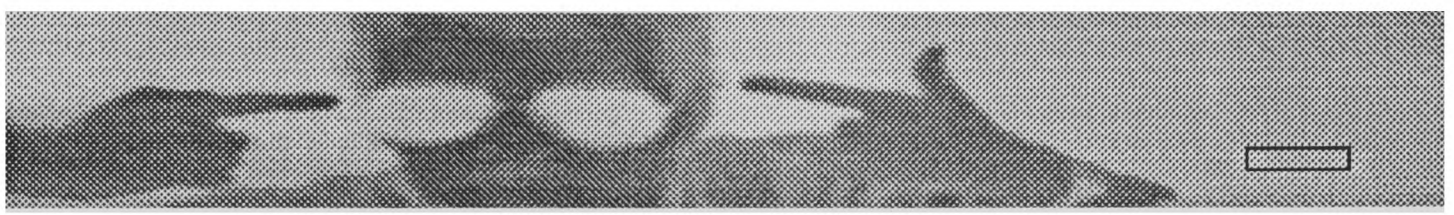


Figure 6.23: Convex cylinder located

all targets in all of the images, it was necessary to weaken the thresholds again, but even with the weaker thresholds, only one or two false targets were reported. We found that over a reasonably large range of target temperatures, we could detect the target in the scene with no false target detection. Figure 6.22 shows an image where the target is harder to see (it is to the right of the person). In this image, using the same thresholds which were used to detect the targets in figures 6.20 and 6.21, the target was successfully detected, again with no false targets. The result is shown in figure 6.23.

We do not claim that this algorithm is anything more than an empirically derived set of instructions for detecting a very limited set of possible targets. It is important, however, since it demonstrates that internal infrared intensity values can be used to distinguish between shapes without recourse to temperature thresholding.

Chapter 7

Computer Vision

7.1 Introduction

When the first attempts at processing images with the intention of automatically recovering information were made in the early 1960s, the quality of the cameras was very poor. The images that they produced were very noisy. The results of processing them were therefore unlikely to be too encouraging. Notwithstanding this, the early work in image processing did lay the foundations of what became the fast moving science of computer vision. As cameras developed, images became better: higher resolution, with better signal to noise ratios. The targets that people were setting for automatic recognition became bolder, while the achievements became more prolific. Modern CCD cameras can provide high resolution, low cost, high speed images: quality that is now considered a prerequisite for the computer algorithms which have been developed.

While computer vision has proved highly successful for images taken in the visible spectrum, less attention has been paid to processing images taken in the

infrared part of the spectrum. This is due in a large part to the fact that the development of high resolution, low noise infrared imagers has lagged behind the development of the visible light cameras. Consequently, the quality of images produced by the average infrared imager is very much lower than the quality of image produced by the average CCD camera. The raw material of computer vision, the image, dictates the level of results that can be expected. With infrared images, the expectations must be lower.

The sparsity of work done on automatic recognition of infrared images can also be attributed in part to the nature of the work for which the imagers have been designed. Most of the funding for infrared imagers has come from the military. The aim has often been to equip troops with a tool which will help them to see in the dark, or pilots with a night navigation aid. Images produced by cameras designed for these purposes may require some image enhancement, but more sophisticated techniques, such as those used in computer vision until recently, were unlikely to be significantly useful. More recently, the projects in infrared have changed, as devices have become more efficient and more robust, but in the past, when the military had vast human resources, the need for automatic recognition was less pressing.

Other fields in which infrared imaging was used included projects where an "expert" used the images which were produced and provided an explanation of the evidence. Examples include doctors examining medical thermograms, or art-

experts inspecting forgeries. Here again, the onus of the work is on the human operator, not the computer. Few examples can be quoted of problems similar to the production-line applications to which computer vision was so readily applicable.

Nowadays, the infrared imaging community has access to all the advanced work that has been done in computer vision. The temptation has been to use this work without much alteration for infrared images. This is likely to be unsuccessful for two reasons: the quality of the images is poorer, so the algorithms designed to work on low-noise images will fail; and more significantly (and this is the basis for the work done in this thesis), the images are intrinsically different so the feature points sought in the algorithms may not even exist in the infrared images.

This chapter addresses these problems. The basic ideas of computer vision are discussed. Particular attention is paid to early vision, which has more relevance to the practical work in this thesis; though some relevant work in intermediate level vision is also described. Mention is made of instances where the theory derived for ordinary images is not applicable to infrared images, or where the algorithms developed for use on visible images are unsuitable for use on infrared images. Finally, some recommendations are made on how processing of infrared images should proceed in the future.

7.2 Low-level vision

The first stage in any computer vision problem is always done at the pixel level. The term "low-level processing", analogous to "low-level programming" refers to the level at which the problem is addressed not in any way to the complexity of the problem.

In computer vision, the first stage of processing is usually designed to segment an image into two or more regions. The processes by which this is achieved vary, but typically will include some form of noise reduction followed by either a region-finding operation or an edge-finding operation. This segmentation can be based on simple differences between two regions or more complex, more subtle differences depending on the problem. In the simplest case, a region of constant intensity will be segmented from another region which has a different but also constant intensity.

A simple "edge" in an image is a part of the image where the intensity is varying locally in a way different to the global variation of intensity around it. The process of finding edges in images therefore requires locating these positions of local change. The edge might be a sudden change in intensity as described above; it might be a gradual change in the intensity followed by a levelling off at the new intensity; it might be a thin region where the intensity varies from the surrounding intensity; or it might be the point at which the gradient of a slowly varying intensity changes. These different types of edges, known respectively as

step, *ramp*, *bar* and *roof* edges require different treatment to locate them.

It is important to consider, particularly in the context of infrared imaging, what underlying physical events are the cause of each type of edge apparent in an image. Binford [11] discusses the types of edges that can occur in a visible scene and how they appear in the image. He suggests that edges in the image can be categorised as: true edges (edges of a surface), limbs (e.g. the edge of a cylinder), surface reflectivity discontinuities (e.g. surface marks) and illumination discontinuities (e.g. shadow edges). More will be said about this later in this thesis.

The earliest attempts at edge finding sought to locate step edges by convolving the image with a filter. Roberts used two by two masks to approximate $\frac{\partial I}{\partial(x+y)}$ and $\frac{\partial I}{\partial(x-y)}$. The three by three masks proposed by Sobel and Prewitt to estimate $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$ are still used today in the early stages of many more complex edge detection algorithms. Other authors chose to use convolution masks which approximated higher order differentials. These higher order differentials are considered better at localising the edge; but errors due to noise in images grow when the image is differentiated so the higher order differentials produce poor results in the presence of noise. Details of these convolution masks and the justification for them can be found in most books on image processing: for example [73].

In addition to these edge detection masks, early edge detection also required the image to be smooth, i.e. not to contain much noise. The reason for this is

obvious: the operators are local and will respond to a noise spike just as if they were responding to an edge. If no filtering were done to remove the noise, the *edge map* formed from convolving the image with the edge filter would be cluttered with responses from the noise. Without further processing, it would be impossible to distinguish the output produced by the noise from the required output produced by edges.

To reduce the amount of noise present in an image, the image is usually convolved with a centre-weighted averaging filter. The most common filter used for noise reduction is the *Gaussian* filter. This is a separable filter and can therefore be applied very quickly; it behaves well for both high and low standard deviations and is readily differentiated and integrated. The degree of smoothing of the Gaussian filter is determined by its support. Horn [38] shows that for a Gaussian filter with support N, the standard deviation of the noise, σ, in the convolved image is reduced to $\sigma' = \frac{\sigma}{\sqrt{\frac{N}{2}}}$. Increasing the size of the mask will therefore further decrease the amount of noise in the image. This will, however, also cause smaller features to be lost when the mask averages them into the background.

The Sobel and Prewitt masks were adopted by many people attempting to locate edges. Some argued, however, that as edges were not necessarily oriented horizontally and vertically in the image, then the edge detector should not be given a bias in these directions either. Marr and Hildreth proposed the use of a non-directional edge detector: the Laplacian of a Gaussian, which they implemented

as a *Difference of Gaussians* filter: a close approximation. (This is described in [53].) While this had a certain amount of theoretical justification, in practice it appeared to perform worse than equally complex algorithms employing uni-directional filters. This is probably because in a man made world, there is a preponderance of vertical and horizontal edges so the uni-directional filters, tuned to detecting these, are more suitable. The cost of combining the responses of two or more filters is not excessive: increased computing power and early parallel processing make light of this problem which Marr and Hildreth argued to be one advantage of their non-directional filter over a set of uni-directional filters.

The magnitude of an edge is the rate of change of the intensity perpendicular to the edge. Figure 7.1 illustrates an edge from an image. Vectors show the directions of the normal and tangential components of the edge. A non-directional edge detector will output the vector sum of the normal and tangential components. If two directional edge detectors are used to approximate $\frac{dI}{dx}$ and $\frac{dI}{dy}$ then the magnitude of the edge is given by the vector sum of these two values, which is equal to the vector sum of the normal and tangential components. The direction of the edge is given by $\arctan(\frac{dy}{dx}) = \arctan(\frac{dI}{dx}/\frac{dI}{dy})$. More than two directional edge detectors can be used to give more accuracy when predicting the edge direction. The true magnitude of the edge is the magnitude of the vector in the direction normal to the edge. The magnitude of the vector tangential to the edge contributes to noise in the estimation of the magnitude of the edge.

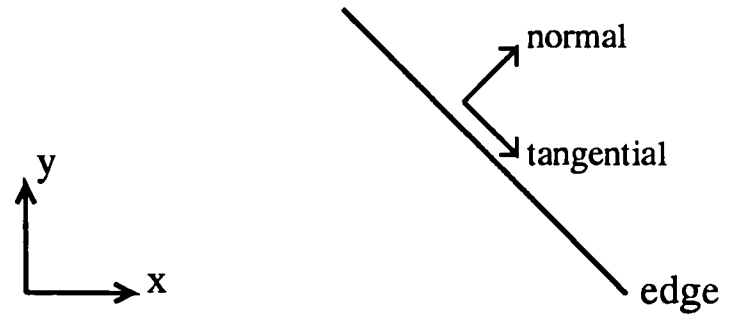


Figure 7.1: Directions of edge components

The important criteria by which edge detectors are judged are: the completeness of description (how few edges are missed); the false alarm rate (how many false edges are detected); the richness of description (the description of the edge as a true edge/limb/etc.); the robustness of the algorithm to noise; and the accuracy in localisation of the algorithm (how near to the real edge the prediction given by the algorithm is). It is also important that if an algorithm is to run in real time, it must run in parallel. It is therefore necessary that algorithms are written with that in mind: early stages of processing must consist predominantly of local operations which, with the correct hardware could be performed synchronously.

The concept of an "optimal" edge detector was introduced: one which satisfied an *optimality criterion.* This optimality criterion, defined mathematically, described the requirements of the detector. One such criterion, defined by Canny for example [18], attempted to optimise the product of a figure of merit for detection and one for localisation. (i.e. locating as many real edges as possible in their correct positions while marking as few false edges as possible.) This optimisation was defined on a one-dimensional signal: the intensity profile across an edge.

The figures of merit he chose to use, the way he chose to combine these factors,

and the weightings he chose to give them determined the form of his edge detector. Canny further constrained his calculations by specifying that there should be only one response for each edge. Having formulated a mathematical expression for his optimal edge detector, Canny then chose to approximate this expression with a filter which was easily implementable. He calculated the effect that this choice of sub-optimal filter would have and decided that the benefits of reduced calculation outweighed the cost in terms of performance. The results of Canny's edge detector were extremely good. Variants of his detector are still used.

Starting with the same sort of objectives, Spacek [82] defined an edge detector which also attempted to optimise signal-to-noise ratio (the figure of merit used by Canny to measure detectivity) and localisation; while suppressing as much noise as possible. Spacek used a different weighted product of the measures and arrived at a different optimality criterion. He also chose to implement different filters: he used a Laplacian as the first stage of his edge detector, whereas Canny had used directional masks: his operator was therefore more suited to viewing natural scenes. This illustrates why an edge detector cannot claim to be optimal for all possible images: the detector is only optimal when used under the conditions for which it was defined mathematically; and those conditions will not be optimal for all (or maybe any) images in the real world.

Although Canny and Spacek both rigorously defined the first stages of their edge detectors mathematically, the performance of their detectors was determined

by the way they chose to implement their theory. For example, Canny's choice of directional filters meant that his detector performed much better on horizontal and vertical edges.

Probably the most significant steps in Canny's detector were the non-linear stages of his algorithm which followed the optimal filtering. These stages, in contrast to the early stages, were ill-understood heuristic steps which appeared to work. They adjusted the edge map by suppressing non-maximal edges and extending contours with a hysteresis thresholding: the thresholds were chosen heuristically based on a histogram of edge magnitudes. It was these stages of his algorithm which differed greatly from what had been done before. Early edge detectors usually involved a number of linear filters: Gaussian, Sobel etc. followed perhaps by thresholding. Canny realised that non-linear processing was needed to improve the performance of these early detectors. More will be said about non-linear processing later in this section.

If an image is smoothed with a Gaussian or other averaging filter, high frequency noise will be removed but so will the high frequency parts of the signal. The more the image is smoothed, the more noise and information is removed. What remains is low frequency information and noise. The information at this low frequency describes the slowly varying effects in an image: it can be used to roughly segment the image into different regions. The high frequency data that has been removed by the smoothing contains the detail of the image including the

fine positioning of edges, surface marks, thin objects etc.. Without this high frequency data, only a rough interpretation of the image can be drawn. Interpreting an edge map formed without any smoothing, however, is extremely difficult due to the large number of edge-elements or edgels.

Information at different frequencies is said to be at different *scales*. The concept of information existing at different scales is not a well understood one. Smoothing with a Gaussian filter is a way of removing noise from an image. It is not mathematically the same thing as changing the scale of an image but it can be used to approximate a scale operator. Because of its ease of implementation, it is invariably used for this purpose rather than any other method. Marr [52] stated that no one scale is more interesting or important than another; meaning that unless information at all scales is considered, information will have been lost. Marr and Hildreth [53] argued that intensity is localised both in the spatial and the frequency domains but that the two localising requirements are conflicting. They also pointed out that a Gaussian filter simultaneously optimises localisation in both domains.

Witkin [94] introduced *scale space filtering* where a continuously varying scale parameter sweeps out a *scale space image*. He constructed a scale space description of the image through which edges could be tracked for increasing or decreasing σ. The *identity assumption* asserts that a common zero contour in scale space arises from a single underlying event. If an edge can be tracked from coarse

to fine through scale space therefore, then the zero contour at the finest scale accurately gives the position of the edge noticed at a coarser scale. Witkin used a heuristic method based on observations of human observers to prune the scale space tree. Canny intended to incorporate information at different scales into his edge-detector, but he never implemented this.

When itemizing the events which give rise to edges in images, Binford [11] does not consider edges due to texture regions. If the orientations and/or sizes of the edge segments show a correlation with each other, a texture pattern probably exists in the image. When discussing textures, edge segments regularly spaced or of similar size or orientation are referred to as *texture elements* or texels. Consider for example a brick wall. An edge map of the wall will contain a large number of horizontal and a large number of vertical edge segments, but the edges of interest will probably be the bounding contours of the wall. If the wall is bounded on one side by a different design of brick wall, it will be extremely hard to segment the two walls simply by examining the edge segments. If the texels in each segment can be grouped and distinguished from those in other segments, however, the two regions can be segmented: one wall segment might contain a large number of regularly spaced short vertical texels and long horizontal texels for example, whereas the other might additionally contain a large number of short horizontal texels.

Much of the work on texture segmentation has involved a large amount of

statistical analysis of the texels. For a review see [34]. Vilnrotter et. al. [87] detected arrays of edges in an image which shared a similar orientation and spacing; creating "edge repetition arrays". Once detected, the texels were categorised by their average intensity and size. Texture analysis is a method of segmenting an image that relies on edge detectors as its raw material, but the output of the edge detectors themselves do not segment the scene. It is a region finding rather than an edge finding approach to image segmentation. Fleck [28] argues that smoothing a textured image with a Gaussian filter can cause a loss of information about texture. If the texels are only two or three pixels large, then the application of a smoothing filter will indeed destroy the texture information.

A very different approach to finding features in images is through using the frequency domain. Any one dimensional signal can be expressed as an appropriately weighted sum of sinusoids. This is the basis of the frequency dependent approach of Fourier analysis. A signal is converted into the frequency domain by taking its Fourier transform. Peaks in the frequency domain indicate the presence in the image of regularly occurring events such as textures which have regular structures. Detecting these peaks in the frequency domain is therefore one method of detecting the presence of a regularly occurring texture. The locations of the peaks determine the nature of the texture: the size (and in the two dimensional case the orientation) of the various texture elements in the texture.

Feature detection using the Hough transform approach works on a different

type of frequency image. Here, the intention is to detect features of a particular kind: maybe straight lines or circles. Considering the problem of detecting circles as an example, the approach is as follows. If an edge segment is consistent with being on the perimeter of a circle of radius r with centre (x, y) then increment the entry in an array of possible radius and centre locations (r, x, y). If a circle does exist at a particular location and with a particular radius, there will be a peak in this array at that point. Each peak in the array therefore indicates the presence of a circle.

Although Fourier analysis and Hough transforms have been quite popular, they have limited use: they are most suitable for solving particular problems such as the circle locating problem described above. Changes in the image are changes which occur in space not in frequency and therefore for more general problems, techniques using information in the spatial domain are more suitable than techniques using information in the frequency domain.

Non-linear processing is becoming increasingly important in edge detection. The first non-linear processing step to be used was thresholding: initially applied directly to the image intensity values in a crude attempt at segmentation; then later on edge-maps to eliminate weak edges and leave the edge map with a more tractable density of edges. Later, median and other rank-order filters were suggested as a replacement for the Gaussian or other mean-averaging type filters for removing noise in an image. Rank-order filters are particularly good at removing

noise spikes from an image while preserving edge sharpness (though not necessarily edge position) but there are several difficulties in implementing them: they cannot be implemented with a simple convolution in the same way as a Gaussian; they cannot be implemented in parallel except using special purpose hardware and their implementation is therefore slow - particularly for filter masks larger than three-by-three; and their output is less predictable than that of Gaussian filters.

Working from the mathematics of digital topology, Fleck [28] designed a test for edges. Having roughly segmented an image into regions separated by candidate edges, she proceeded to test the edges using her $*$-convexity algorithm. To test a candidate edge, paths through homogeneous texture regions were examined for consistency with the initial estimations of edges and regions. If successful, the edge was marked as a real edge, but if the test failed, the edge was marked as a false edge. This non-linear algorithm was successful at correctly segmenting finely textured regions. It did not perform as well when the texture was at a coarser scale; and although it was essentially a local process which in theory could be run in parallel, implementing it in parallel would be very difficult. The algorithm was rather slow.

Filters based on morphological opening and closing transformations have for several years been used for suppressing noise. Like rank order filters, morphological filters are particularly suitable for suppressing impulse noise; but unlike rank

order filters, they also preserve two dimensional image structure. Morphological filters further differ from median filters in that they are idempotent meaning that re-application of the filter does not further change the structure of the signal. Noble [66] shows that morphological filters will outperform any rank-order filter. Furthermore, morphological filters can be implemented efficiently in parallel using appropriate hardware and can therefore be made to run quickly.

Noble and Brady [67] have used morphological filters in detecting edges of texture regions. Following the application of close-open and open-close morphological filters, a Mann-Witney test is used to determine whether different regions have significant statistical differences. (The Mann-Witney test is similar to the Student t-test but for non-parametric data.)

Morphological filters seem to be highly suitable for infrared image data which typically has a large amount of impulse noise. This will be discussed further later in this chapter.

7.3 Intermediate-level vision

Low-level vision and intermediate-level vision are terms used to categorise roughly the type of image processing. The border between these is somewhat arbitrary. Here, we choose to distinguish between them by saying that low-level algorithms operate on pixels, edgels or texels usually with the intention of segmenting an image; whereas intermediate level algorithms build on the output of low-level algorithms to provide more useful descriptions of the image. This follows the

modular approach described in [12]. Because of the vagueness of the definitions for low and intermediate levels, some of the material in this section could perhaps be included in the previous section and vice-versa.

Getting intermediate level information from an edge map or region map entails some form of grouping of edge segments or regions to form connected edges or regions and forming a description of the shape of that connected edge or region. This is then used to provide a description of the shape of the object in the scene giving rise to it. The key point here is that features in the edge map or region map are interpreted as scene events.

Binford [11] makes a number of assumptions about the various types of edges he defines (as described in the previous section), which constrain their interpretation when examining the edge map. He assumes for example that straight lines in images come from straight lines in space; that if a line in an image is consistent with it being vertical, then it is vertical; that if two lines are consistent with being parallel, then they are parallel; and so on. This type of reasoning constrains the possible interpretation of the edge map, but it does not provide a description of the object in the image.

Work on determining shape from contour has been attempted. Here, as with other "shape from" algorithms, the aim is to provide a description of the object in the scene. A full description might be a bounding contour of the object and an estimate of the surface normal at each point on the surface of the object.

When estimating surface normal from contours, it is necessary to make the prediction that surfaces are locally planar. The normal direction is then estimated for each local planar area and local facets are produced lying tangential to the surface. Connecting these facets provides a piecewise planar estimate to a surface reconstruction of the object.

An alternative to shape from contour is shape from texture. If an image contains a regularly occurring texture, then the size and orientation of that texture can be used to estimate the depth and orientation of the surface which gave rise to it. The assumption underlying this approach is that the facets on the surface of the object project onto the image plane in some recognisable way. If this is so, then the texels produced by some texture detector can be used to predict the nature of the surface.

Horn [37] has worked on understanding an image from the intensities of the pixels. The object of this work is to understand the physical laws which cause the formation of the image. It is highly relevant to some of the work done in this thesis on the modeling of infrared signals. A summary of Horn's work on understanding image intensities is now presented.

Horn begins by assuming that opaque bodies exist in a transparent medium and light reflected from them are projected under a perspective projection onto an image plane. He further assumes that the viewer is far from the objects so that orthographic projection is a valid approximation. Light is assumed to come from

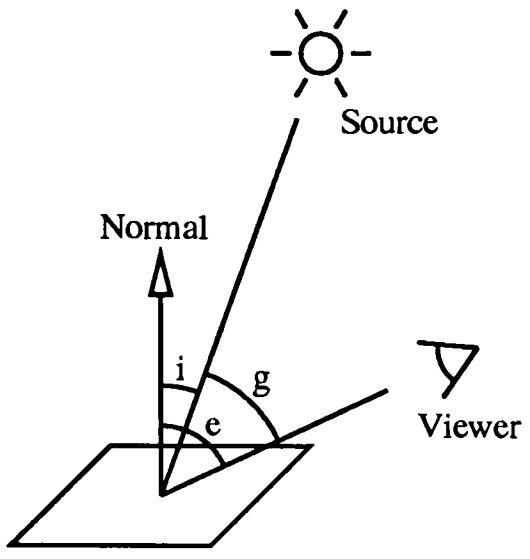


Figure 7.2: Geometry of reflection for single light source

a single source. To deal with more complex lighting patterns, the results from each of the single light sources are superimposed.

The geometry of reflection for the single light source case is defined by three angles, as shown in figure 7.2. The angle of incidence, i is defined as the angle between the source and the surface normal; the angle of emittance, e is defined as the angle between the viewer and the surface normal; and the phase angle, g is defined as the angle between the angle of incidence and the angle of emittance. The three directions are not necessarily coplanar.

The surface is divided into surface patches, each with a corresponding surface normal. An equation of the tangent plane of the surface patch has two degrees of freedom. These are used to define the *gradient* at that point: local normal $= (z_x, z_y, -1) = (p, q, -1)$ then gradient $= (p, q)$; where the surface is defined as $z = f(x, y)$ and z_x and z_y are the partial derivatives of z with respect to x and y. By constructing a surface normal for every surface patch in the image and

mapping them onto the point (p, q) in *gradient space*, a dual representation of the image is built up. This gradient space image is used to investigate how image intensity varies with surface orientation.

Horn suggests that for most surfaces the reflectance is uniquely defined so that for a given surface orientation, the image intensity will also be uniquely defined. This assumption underlies the dual-space approach adopted by Horn: if surfaces differed wildly from this behaviour then the work on gradient space would have little value. The amount of light reflected by the surface is dependent therefore on the nature of the surface, the orientation of the surface and also the distribution of the light sources around the surface. If intensity is measured at a point, then the possible orientations which the surface at that point can take is limited to a known subset as defined by the gradient space image. The orientation is not, however, uniquely defined.

Two types of surface are considered: Lambertian and specular. For the Lambertian reflector, Horn shows that for a light source near the viewer, the reflectance plotted in gradient space is given by a set of concentric circles of iso-brightness. For the more general case, when the light source is not near the viewer, the plot in gradient space consists of a family of conic sections. He suggests that for more general reflectivity functions, the locus of possible normals constitutes a more general figure: a Monge cone.

The other type of surface considered by Horn is the specular surface. Surfaces

which are smooth on a microscopic scale are usually specular reflectors. While many surfaces exhibit some degree of specularity when they reflect light, few are pure specular reflectors. One example quoted and analysed by Horn is white paint which can be thought of as a combination of Lambertian and specular components. An equation is given which relates the amount of incident light reflected specularly and the amount which penetrates the surface and is then reflected in a Lambertian fashion.

Horn also begins to address the problem of mutual illumination. Realising that the general case is difficult, he restricts his attention to the relatively straightforward task of examining the effect of two semi-infinite planes, joined at right angles and illuminated by a distant source. The illumination of each plane depends on the angle of incidence of the light source. He shows that for light incident along the angle bisector of the two planes and with planes having reflectivity of 1, then the illumination of the planes is twice that which would be apparent in the absence of mutual illumination effects. He also shows that as the angle between the planes gets more acute, the illumination is further increased. This analysis is done for Lambertian reflectors. For planes which have a large specular component, the effect is much less as most of the light is bounced back to the source after two reflections. It is suggested that for planes of finite dimension, the analysis is still approximately valid, but intensity falls off as a function of distance from the corner. No closed form solution is offered for this problem, though an approximation

given by numerical methods does agrees with this assertion.

More recently, work by Forsyth and Zisserman [29] has extended the understanding of mutual illumination. They use numerical methods: finite element analysis to predict the effect of mutual illumination for a selection of shapes and surface properties. They show experimentally that the effects predicted theoretically do occur in real images.

They argue in [30] that information from mutual illumination can be used to help in the process of estimating shape from shading; and indeed that if mutual illumination effects are not taken into consideration then the results of shape from shading will under some circumstances be incorrect.

Horn went on to interpret the profiles produced by edge detectors. By examining the image intensity profile across two faces of a corner and by examining the corresponding image in gradient space, he asserted that an edge profile with a peak shape or step with a peak superimposed on it is most likely to correspond to a convex edge in the scene. This does not magically solve the problem of detecting a convex edge, however, as convex edges do not necessarily give rise to peaks in the edge profile. Considering the effects of mutual illumination, we might expect that a roof shaped edge profile or a step with a roof shape superimposed on it is a good indication that a concave edge exists. Roof profiles can also be caused for other reasons, however, for example if the bandwidth of the detector is not very high, so the existence of a roof edge does not guarantee the existence of a

concave edge or a mutual illumination effect caused by other objects in the scene. Step edges are indicative of objects obscuring each other; though they can also be found with concave and convex edges and with changes in surface colouring or type. A negative peak associated with a step or peak edge strongly suggests a surface obscuring a self-shadowed region: here the line connecting the two points in gradient space corresponding to the two surfaces, passes through the terminator - the line separating lighted from shadowed regions.

It is possible to determine the shape of the object from the intensity measurements by considering the gradient space image and the edge profiles. The ambiguities in surface angle inherent in the gradient space approach can be resolved by assuming that local gradient does not change rapidly except at edges. This constrains the surface orientation at each point. A separate technique has also been proposed for solving the ambiguity: photometric stereo. Here, the object is illuminated using a single light source and calculations of possible surface orientations are made as before. The light source is then replaced by a second light source in a different position. This provides a separate set of iso-brightness contours in gradient space. By superimposing the two gradient space images, the surface orientation is constrained to be at one of two possible angles. A third light source removes any remaining ambiguity.

A large amount of work has been done on recovering shape from stereo. If two images are taken of the same scene from slightly different positions then the

minor disparities between the images can be used to calculate depth in the image, thereby allowing three-dimensional reconstruction of the object. Stereo is the technique which allows humans to perceive depth. In computer vision, the main problem is determining how to match points in the two images. The basic approaches are region matching and edge matching. For images containing a large number of features, edge matching performs well while region matching performs well for smoothly shaded regions and for textured regions. Other approachs include matching feature points: points which have previously been determined as "important"; and matching features: objects or parts of objects in the image. Matching at these higher levels is more likely to produce accurate matches, but producing the higher level description requires more calculation.

Matching involves locating a feature, region or edge in the left image and pairing it with the same feature in the right image. This may not always be possible: the feature may be beyond the edge of the right image, it might be obscured by another object or it may not have been detected. Determining which feature to match - the so called "correspondence problem" is more difficult when there is a large number of possible feature points; though the larger the number of matched features, the richer the description of the scene will be. Matching can be done at different scales - refining the estimates from a coarse scale match with the information provided by the finer scales. A number of scene events can confuse the matching process, such as objects moving in the scene or cylindrical objects

with limb edges. A large amount of work is being conducted in this field: seeking the best possible matching criteria and attempting to solve the various problems of stereo.

Another approach to scene segmentation and image reconstruction: "structure from motion" uses similar techniques to stereo. Here, one camera is used to provide a sequence of images. Two consecutive images are then compared. As with stereo, the purpose is to match points or features in one image with corresponding points or features in the other image. It is readily apparent that the matched points give an indication of the motion of the feature point between the two frames. Essentially, a time derivative is being taken of the image.

The velocity direction and magnitude is plotted for each matched point illustrating the motion of each region in the image. The pattern thus produced is know as the *optic flow* pattern. Clearly if a neighbouring region of feature points share similar velocity magnitude and direction then they can be assumed to belong to the same object. This simple approach is only valid however, if the motion is purely translational. If the object additionally has rotational motion, the velocity vectors will have different directions and magnitudes. Interpreting such a velocity field requires more intelligent processing. Like stereo, a major difficulty in estimating motion is the correspondence problem. In early optical flow work, the camera was assumed stationary and the aim was to estimate the (predominantly translational) motion in the scene. More recently, work has been

done on detecting rotational motion. Work has also been done with data taken by moving cameras: for example the DROID work done by Plessey [19] on estimating driveable regions and ego-motion from a sequence of image frames taken from a moving vehicle.

7.4 Infrared image processing

This section is not a review of previous work on infrared image processing: that can be found in chapter 8. It is instead a comment on the applicability of image processing in general to infrared images. Some of the points made in this section will be referred to when examining the work of previous authors.

Before discussing what image processing should be applied to infrared images, it is essential to state what sort of features we can expect to find in the images and how they will appear. The work on computer vision directed at image understanding almost exclusively assumes the model of a single distant light source and considers its effect on Lambertian and specular reflectors. This model is totally unsuitable for infrared images. The source of energy in infrared scenes can come from the sun: outside scenes of stationary cars, for example, can provide many solar reflections; but invariably the majority of the signals will be produced by internal emission of energy from the objects: people, car engines, buildings, electrical appliances all produce heat internally. Scenes of walls and roads can have reflected components, but usually the energy emitted by a building is by absorption and re-emission rather than by reflection. Similarly, vegetation, trees

and grass do not reflect solar radiation: they absorb and re-emit it. Consequently, the work done on interpretation of edge profiles or shape from shading cannot be used on infrared images without modification.

Whereas in visual images, surface orientation is an important cause of intensity changes, this is not the case in infrared images: for many material types and surface conditions, the orientation of the material does not greatly affect the intensity: convex cylindrical blackbody objects appear to have a flat intensity profile (see chapter 6 for an example) and internal edges of convex objects are not always visible. As explained in chapter 2, most bodies are Lambertian emitters of infrared radiation and hence radiate equally in all directions.

Extremal boundaries are present in infrared images: segmenting using extremal boundaries is relatively straightforward in tasks where the objects to be segmented differ from their surrounding environment as is the case in many infrared tasks. It is also possible to detect changes in depth: two similar surfaces with different distances from the camera will give different intensities because of the attenuation caused by the atmosphere and the spreading out of the signal.

Two similarly shaped objects at the same temperature but made of different materials will emit different amounts of infrared radiation and can therefore be distinguished. Surface type, both texture and colour, also greatly affects the way in which a body will emit radiation.

Sources of heat are of course another detectable feature in infrared images.

The sources could be internal, such as the engine inside a car, or external: a laser shining onto a material for example. The external heat source could cause an area of local heating or it could give a specular reflection depending on the nature of the surface of the material.

In visual images, the image is formed usually by the illumination of a scene by a single light source. The objects in the scene are distinguished from each other by the way they reflect the light to the viewer, by the shadows caused and to a lesser extent by the way they reflect light on each other (mutual illumination). For infrared images, the analogue of "light source" is the "heat source". Of course the sun is one very important heat source, but by no means the only one. Indeed, unlike with visible images, it is very often the source of energy which is the interesting feature in the image. Very little work has been done by the vision community on the detection of light sources since they are rare and usually of no interest. Ullman [85] is one exception: he suggested that there is a light source in the image if the ratio of intensity of two points near an intensity boundary is different from the ratio of the intensity gradients.

With more sources of energy, the amount of mutual illumination can also be large. The mutual illumination may be either from reflection or from absorption and re-emission depending on the surface type. The effect of mutual illumination on simple shapes is explained in chapter 2. Analysing the various mutual illumination interactions in a complex scene is, however, an extremely difficult

task.

Because infrared images tend to contain more noise than visible ones, the algorithms for early edge detection will have a degraded performance when applied to infrared images. The best performances will be from those algorithms which are designed to work in the presence of noise. When considering noise reduction, the most suitable filters appear to be those designed to reduce impulse noise: specifically rank order filters and morphological filters. Morphological filters, being easier and faster to implement and being superior in performance are the obvious choice.

The particular imager described in chapter 4 relies on a scanning mirror to provide the second dimension. Because the mirror is not a perfect reflector, it causes a blurring of the signal. Perfect step edges are therefore rare in images taken using the mirror; but if an edge detector with sufficient support is used, boundary detection using an algorithm similar to Canny's should be possible.

In principle, there is no reason why frequency based techniques cannot be used just as well with infrared images as with visible images. Indeed, chapter 8 includes a description of one attempt at using a Hough transform on infrared images.

Photometric stereo is not a technique which can be used for infrared images. Pixel intensities rely more on internal radiation of energy than on reflection of external sources, which is the method of image formation in photometric stereo. Also, the success of photometric stereo relies on providing an image in gradient

space. If information about orientation is not available from the images, as is the case with infrared images, the technique will fail.

Inferring structure from motion in infrared images will still be possible provided a sufficient number of feature points can be found and matched correctly and provided that the noise is not too great. A similar argument can be made for inferring structure from binocular stereo. Neither motion nor stereo could be performed on the images produced by the imager described in chapter 4 of course, but this is because it is a line scan imager not because it is an infrared imager. One further problem which may be faced when attempting stereo with infrared imagers is that the infrared imagers are less consistent than the visual ones: responsivity and pixel pitch might be different in two imagers making the problem of matching harder still.

To summarise, this chapter has listed some of the low-level and intermediate-level computer vision work completed or in progress. Comparisons have been made between visual imagery and infrared imagery to illustrate the relevance of the computer vision work to infrared images.

Chapter 8

Infrared Image Processing

8.1 Previous work

8.1.1 Introduction

In chapter 7, a brief outline of work done in computer vision was presented and its relevance to infrared imaging was assessed. In this chapter, a selection of previous authors' work on processing of infrared images is reviewed. The work is drawn from a large variety of disparate sources but the majority comes from the defence and medical fields. Some of the work done for the military is classified and cannot be found in the open literature. All the papers reviewed in this section describe unclassified work.

Much of the infrared image processing literature is concerned with improving the appearance of a televised output from a particular device for viewing by a human observer. Typically, these papers describe the manufacture of the device and outline the electronics used to create the display. This information is of limited interest in the context of general purpose image processing and papers

of this kind will therefore not be included in this review. Similarly, papers from the remote sensing field concerning satellite images have little direct relevance to work in the previous chapters and will also not be considered here.

Unlike computer vision, where a large amount of work has been done in generalised algorithms, work in infrared image processing has been predominantly problem-driven. Some authors have returned to basic physics and derived models of the infrared radiation, but the majority have uncritically imported the algorithms developed for use with visible imagery. Examples of both types of work will be described in the first section of this chapter.

None of the papers reviewed here follows a similar approach to the one we discussed in chapter 2. This section has been included to give a review of the kind of work being done by the majority of authors in infrared image processing.

8.1.2 Literature review

In applications where the object temperature can be accurately controlled, the need for image processing is minimized. Gallo et al [32], for example, describe a medical application where the patient must spend fifteen minutes at room temperature before images are taken. Here, simple intensity averaging and maximum extent shape processing follows a manual segmentation of the images. The results are used to aid diagnosis. Similarly, Clarke and Leonard [20] work with a controlled scene containing two black body sources of (different) known temperatures. The radiance for each scene location is given by linearly interpolating the

intensities for each pixel in the image. As with the work of Gallo et al, this work is only valid in situations where the scene can be completely controlled, however, and is therefore rather limited.

Histogram analysis is used by Radford [75]. A Sobel edge detector gives an edge map which is scaled to six bit accuracy. This is concatenated with the six most significant bits of the grey scale image to form a twelve bit image. All regularly occurring intensities in the histogram of this twelve bit image are discarded. Remaining pixels which are not isolated or pairs are used to form a mask of the "interesting" parts of the image. The algorithm detects targets which are small with respect to the image size and which differ in intensity from the background.

The relative merits of median and mean filters are discussed by Holst [36]. He notes that median filters have been reported to remove impulse noise without affecting edges: for a $2P+1$ filter, signals with width $P+1$ or more will pass through the filter. He argues that as analytical expressions cannot be used to describe a median filter, that its performance must be analysed by implementing it on a known test image. Images of bars with a variety of widths are used to compare the performance of mean and median filters. Holst argues that the poorer performance of median filters on the image of very thin bars is justification for preferring the mean filter. He also points out that the theoretical noise reduction of a median filter ($\sqrt{\frac{\pi}{\pi+2N-2}}$) is less than that of a mean filter ($\frac{1}{\sqrt{N}}$). He concludes that because of their unpredictability, nonlinear filters should only be used when

image details such as resolution and noise characteristics are known. His argument fails to address the problems of mean filtering, however, such as loss in edge position information, and he does not consider morphological filters - non-linear filters which can be described mathematically. In almost all practical applications, the minimum width of the feature of interest will be known so mask can be chosen which does not cause a loss of object information.

An overview of a variety of infrared image processing techniques is summarised in [70]. This book deals with all aspects of image formation from radiation equations through detector types, optics and scene modelling to signal processing. Edge detection, line detection, skeletonisation, texture, scene segmentation and scene matching are all covered in the chapter of signal processing, but only superficially.

Layuan and Haiming [46] describe a method of contour enhancement to overcome the problem of thermal spread (blurring) inherent in infrared images. They assume that the image taken at time $t = 0$ contains no thermal spread and can be expressed as a Taylor's expansion of the image at time $t = \tau$. Making several simplifying assumptions, they derive a digital approximation to a mask which will recreate the image at $t = 0$. Layuan [45] improves the accuracy of the algorithm and derives a $7 * 7$ mask for contour enhancement. Only one example of his results is plotted and that is for a one-dimensional signal. As this takes no account of any orientation effects, it is not clear how well the algorithm will perform on two

dimensions. The result does indicate, however, that at least for one-dimensional signals, the method may work quite well.

A number of different filters are proposed by Nack [61]. The performances of them vary, but the best results are achieved with: $G(x,y) = (I_{x+1,y} - I_{x-1,y})^2 + (I_{x,y-1} - I_{x,y+1})^2$. This convolution performed better than such algorithms as Sobel and Laplacian when applied to a range of images. This is perhaps somewhat surprising, however, since $(I_{x+1,y} - I_{x-1,y}) \simeq \frac{\partial I}{\partial x}$ and therefore $G(x,y) \simeq \nabla^2$, so the filter is a (poor) approximation to a Laplacian filter.

A different filter for image enhancement is described by Schau [77]. He suggests the use of a third central moment filter to reduce clutter in an image while enhancing objects or their gradients. The third central moment, defined by: $X_L = \frac{1}{A}\sum(X_j - \overline{X})^3$ is an unconstrained filter (requires no prior knowledge of the image). Its performance on a sequence of infrared images is compared with that of Sobel and Laplacian operators and is shown to be superior when the signal to noise ratio is high. As we show in section 8.2, however, the Sobel operator is a poor quality edge detector: very much better algorithms are available.

Hayden et al. [35] describe an edge-based image segmenter using data from a number of consecutive frames. They use a Prewitt edge detector to calculate edge magnitudes and phases. (The phase information is quantised to eight directions and used to guide the thinning algorithm.) The edge points are extracted from the edge magnitude data using a threshold calculated for each frame: $T = m + c * \frac{\sigma}{1 - \frac{1}{m}}$,

where m is the mean, σ is the standard deviation and c is an empirically chosen constant which controls the strictness of the threshold.

Data from previous frames are incorporated by applying an alpha filter to every pixel in the image: $y_t = (1-\alpha)x_t + (\alpha)y_{t-1}$ with $(0 \leq \alpha \leq 1)$. For white noise, this reduces the noise by a factor $\frac{(1-\alpha)^2}{(1-\alpha^2)}$. Because of motion of objects in FLIR images, if α is large and many frames are integrated, then the images will appear blurred. Objects within frames are therefore registered using a minimum-difference correlation tracker before the information is combined. In a target tracking application, where the current frame is the most important, it is necessary to combine time history data only in a postprocessing stage.

The thresholded edge map of Hayden's system is tidied up using a number of local operations in a $3 * 3$ neighbourhood. Isolated points are deleted; non-edge points with two non-connected edge neighbours are reclassified as edges; and "jetting points" (edge points which are one pixel extensions orthogonal to a straight edge) are reclassified as non-edges. End points of edge lines after this cleaning operation denote gaps in boundaries. These gaps are filled by correlating the data from the previous frames. Once a connected boundary has been found, it is preserved between frames. The approach suggested by Hayden et al. is only suitable for tracking shapes which move by small amount between frames: distant objects and slow moving objects and where imager movement is not allowed. In these cases, optic flow is well conditioned and motion compensation could be used

for enhancement. The results that they present indicate that it works reasonably well under their specified conditions for values of SNR greater than 2. Because the processing is restricted to convolution and thresholding, it is possible to implement the algorithm in hardware.

Cussons [22] built a system which located aircraft in images, suppressing false targets and tracked the targets in real time. A 5×1 vertical median filter was first used to eliminate the striping characteristic of the particular imager. This had the undesirable effect of removing targets with height of two pixels or less. A Sobel operator was chosen to extract edge information because it performed better on low contrast data. Connected boundaries, formed by examining nearest neighbours, were used to give shape descriptions of the possible targets. Each potential target is assigned a probability that it is a real target. The probability is increased if a target is found in the same location the following frame; and decreased if the target is absent for several frames. If the probability gets sufficiently high, the potential target is classed as a definite target. This type of temporal updating of probabilities is suitable for problems where noise causes measurements to be uncertain.

Burton and Benning [17] compare four different target detection algorithms: "contrast box", "spoke filter", "superslice" and "double-gated filter". A database of 256 infrared images is used to test the ability of the algorithms to detect targets with sizes ranging from 5 to 500 pixels. The contrast box uses coarse range data

to predict the size of the target. A convolution mask consists of two concentric rectangles: the inner rectangle being the approximate shape of the target. The mean, μ and standard deviations, σ of the pixels in the inner and outer box (target and background) are calculated and then the contrast box metric is given by: $c = \frac{(\mu_T - \mu_B)^2 + \sigma_T^2}{\sigma_B}$. This metric indicates how the target differs from the background and hence gives a measure of likelihood of the presence of a target.

The double gated filter (DGF) bears some similarity to the contrast box filter except that the background filter is divided into N subregions and local thresholding is performed on each. The DGF metric is then given by: $c = \prod_{k=1}^{N}(I_T - I_B^k)$, where I_T is the mean of the target gate pixels exceeding the threshold and I_B^k is the mean of the pixels in subregion k exceeding the threshold. (If no such pixels exist, then the mean of the region is set equal to the background gate mean.)

The spoke filter is a generalisation of a Hough circle detector. (See section 8.2.3 for a description of the implementation of a Hough circle detector.) An edge map is formed using a Sobel operator. A spoke filter is applied to the resulting *phase angle* image. A directional matching criterion then locates potential targets.

The superslice algorithm combines the output of a connected components algorithm run over a sequence of images with different thresholds with the output of a difference operator which has been thinned using non-maximal suppression. The threshold at which the best edge/perimeter match is found is used as a threshold to segment the image.

The contrast box algorithm was the only algorithm designed to detect bimodal targets and the only one requiring range data. In order to compare the performances of the different algorithms, the other three algorithms were modified to additionally detect cold or bimodal targets; and pseudo range data was given to them to improve their performances. Comparing the performances of the detectors over a selection of targets, the contrast box algorithm was shown to perform best when the target size is small (less than 150 pixels). This is perhaps not surprising given the modifications made to the other algorithms and considering that the superslice and spoke algorithms were not designed to detect targets less than 20 pixels in size. For large targets (over 150 pixels), the spoke filter (the only one of the algorithms with any shape representation in it) gave the best performance. We notice that for very small targets, if the target and backgrounds are both constant, the contrast box becomes quite similar to the Laplacian operator.

The contrast box algorithm is used by McWilliams and Srinath [58] for detecting surface targets in an air-launched or ground-launched guided missile application. The specification for the missile states that within thirty seconds of launch it must lock on to a target. The problem is broken down into two sections. In the first stage, the "seeker" stage, possible targets are identified using the contrast box. The hot spot of the most likely targets is passed onto the second stage, the "ranker", which uses intensity features, edge features and binary segmentation features from the contrast box to choose the most likely target. The missile is

aimed at the target with the highest ranking score. The algorithm will always choose a target even if no real target exists.

The paper does not discuss how the target is tracked once it has been located nor whether the processing is done entirely on the first image taken. If this is the case, the missile may have travelled a long way before the decision of which target to seek is made.

Allen and Northfield [1] have studied locating and then tracking targets for an air-to-air guided missile application. In the first stage of processing, they locate possible targets from large distances using a thresholded output from a Laplacian filter. The test images used were background cloud scenes with a uniform square patch superimposed upon them to simulate a target. Region growth is performed on the image of potential targets if the connected pixels are higher than a certain threshold. The method for choosing which one of a number of possible targets to acquire and follow is not discussed.

Once acquired, the target is tracked. The method for tracking is not described, but some of the problems associated with tracking targets are outlined - the target may get temporarily lost in high background clutter or may present a reduced radiance level due to its orientation changing or atmospheric effects. The final stage of the process is directing the missile at the target from close range. Again, this is only discussed - no solutions are suggested for the problems. When the missile is near to its target, it is possible to extract additional information from

the image, such as whether the target is an enemy or friendly, and if it is an enemy, it is possible to maximise the "lethality" of the missile by directing it to hit a specific part of the target.

A multinational cooperative research project is described in [80]. It considered the problem of discriminating and classifying military targets from natural background scenes from a common data base. Comparing their results, they found that locally adaptive segmenters which assumed that the background was composed of large uniform areas performed better than globally adaptive segmenters; but no improvement was noticed when using object adaptive segmenters which adjusted parameters for each object. The algorithms which performed best were those which combined the results of a number of different feature detectors to reduce the number of false alarms. They also found that the more complex algorithms tended to be better at classifying the targets. The performance of the algorithms was compared with that of a trained human observer and found at every stage to be almost as good. It was suggested that other types of sensor data (e.g. radar) could be combined as might temporal and contextual information. None of these suggestions was implemented, however.

Recently, work by Radford has been extended slightly and written up by Bilsby [10]. This work concerns locating instances of tanks in infrared scenes. The first step is to extract the noise in the background without affecting objects in the foreground. This is done by calculating the variance of a 3×3 region around

every pixel. If the variance is below a threshold, mark the pixel. Join all marked pixels and replace their value with the mean grey level value of the region. This joins the background pixels. A binary mask of the image is then created by thresholding using the background intensity as the threshold value.

Blobs of reasonable target size (18-1600 pixels) are then labelled; all other blobs are removed. The remaining blobs are merged together if their maximum horizontal or vertical extents overlap. Blobs are denoted as targets, possible targets or false targets. A history file of locations of labelled blobs is updated each frame time and the certainty that a blob is a target is modified: a blob remaining in approximately the same place for several consecutive frames is likely to be a target, while a previously found target must be absent for several frames before it is re-classified. Finally, cross hairs are placed on the hot-spot in the target and on the centroid. This algorithm is sensitive to movements of the camera.

Three methods for extracting features from infrared images were assessed by Series et al. [79]. Each was applied to the problem of detecting car wheels in a scene. The first method, a Hough transform technique used a non-maximally suppressed, thinned output from a Sobel detector, thresholded to remove rough lines as its input. Sharp peaks in the Hough transformed image indicated the presence of circles in the image. We present results of our own Hough transform code in section 8.2.3. We see that while it occasionally is successful, it is much less robust in noisy infrared data than visible imagery. The other two methods described

by Series et al. are dynamic programming based methods. The second method, "edge list search" (ELS) matches a template with edge data from the scene. The edge data is first decomposed into a list of isolated lines and closed loops (node points are deleted). Any set of observed segments which could correspond to the reference segment are given a goodness of fit score: those collections of edges with the highest scores are labelled as possible targets. The third method, "full image search" (FIS) is similar to the ELS but any pixel (not just edges) can be matched. One problem with both the ELS and FIS algorithms is that the two best solutions might lie almost on top of each other - both coming from one wheel. The Hough transform method is insensitive to fragmentation, but they claim that it is a non-trivial task to generalise it to non-circular image segments.

A reduced form of optical flow is utilized by Markham [51] for estimating the "time to go" to reach a target. The procedure is to measure the movement of points on the target and from them derive the real world motion. The tracked points are either hot spots in the image of the target or are manually chosen. At least two such points must be tracked over time to get an estimate of the distance of the target. To decouple the motion of the camera, either a Kalman filter is used or points in the background are tracked and used as a reference. The algorithm improves in performance as the target is neared and the relative motion of the tracked points becomes greater. As is often the case, accurate estimates of time to go were achieved when the algorithm was tested using synthetic images. The

algorithm is not robust, however, to changes in the image, such as the target temporarily being obscured or the shape of the target changing (for example if a vehicle was being tracked and turned to face the tracker).

Target detection algorithms were used by Aykroyd [4] not to detect targets but to characterise backgrounds by examining the rate of detection of "false targets". This work was part of a project to improve camouflage. The procedure was to try to match a template to each part of the image allowing for rotations and changes of scale (matching in four dimensions). The number of false targets detected was counted. The experiment was repeated with p different algorithms and on n different images. An $n \times p$ matrix of false target rates summarised the results. Cluster analysis and principle component analysis techniques were used to show which images were similar and which algorithms produced similar results. (The principle components analysis gave slightly better results.) It was suggested that background characterisation could not be achieved by using the results of one target detection algorithm alone. They demonstrated, however, that combining information from a number of algorithms was a successful method for distinguishing between different background types. The cluster analysis showed that a reduced number of algorithms could be used to give a description of the background; that the description was similar for similar images and discriminated between different images. Robustness to noise, rotation and scaling were also demonstrated.

Work has been done by Nandhakumar and Aggarwal [62, 63] on the integration of data from visible and infrared images to give information which cannot be obtained by processing one of these types of images alone. They argue that in most instances, the energy emitted by a body is much larger than the energy reflected by the sun or from other bodies. There are of course many exceptions to this; including specular reflectors in the scene and mutual illumination effects, but they claim that these form only small regions in the image.

Incident radiation is estimated using empirical methods based on the time of day and the time of the year. (This is clearly dependent on the consistency of the weather and is not at all robust.) The energy detected by an infrared imager and the ambient temperature are used to estimate the temperature of the surface; and wind speed is used to estimate the heat convected by the surface. An energy sum is then calculated at the surface to give the heat conducted into the surface. The ratio of heat conducted to the energy absorbed by the surface is used to distinguish between different classes of objects.

Information about surface orientation used in estimating the absorptivity of the surface, comes from the visible image. The image is first segmented manually. They assume that the image is composed of uniform Lambertian reflectors. Points of known orientation in each segment of the image are used to calculate local orientations from the image intensity values. Errors are caused when specular reflectors are present in the scene or when surfaces are not uniform.

The problem of registering the thermal and visible images is not addressed, nor is the problem of segmentation. The ideas are only valid when the scene is illuminated by sunlight: those areas in the scene which are in shadow cause errors. A lot of assumptions are made, at least some of which will be invalid in most imaging situations, but the work does address some of the problems in infrared imaging that have rarely been looked at by other authors.

8.1.3 Summary

For ease of reference, in table 8.1 we present a summary of some of the strengths and weaknesses of the algorithms discussed in section 8.1.

The previous section describes some of the work published on image processing of infrared images and related topics. It is clear by comparing the work here with the overview of computer vision techniques given in chapter 7 that many of the ideas of the vision community have not been considered at all by the infrared community. This is primarily for two reasons. Firstly the poorer performance of infrared imagers makes processing them that much more difficult. Some techniques developed for visible images will fail to overcome the noise problems in the infrared images. Secondly, for most of the applications in infrared imaging, it has been necessary to produce real time images so it has not been possible to date to implement complicated algorithms. Authors have preferred to use simple but poor quality edge detectors such as Sobel's.

The real time requirement of infrared problems does not, however, invalidate

Paper(s)	Comments
Simple image processing	
[32, 20]	This work relies on the scene being completely controlled.
[36]	Criticizes median but does not consider problems of mean filters.
[46, 45]	Only one result is presented (for a 1D signal). It is not clear if the technique would work on 2D data.
[61]	The best performances of a selection of techniques come from a poor approximation to a Laplacian.
[77]	Only performs well for a high S/N ratio.
[75]	Suitable only for finding small bright targets.
[1]	Discusses problems of tracking targets in infrared images. A number of questions are left unanswered.
[58]	Reasonable target detection rate only if several targets exist in scene. Does not discuss how targets are tracked.
Tracking with time	
[35]	Reasonable results if S/N ratio is fairly high and target only moves a small amount between frames.
[22]	Uses simple techniques and compares images between frames. Algorithm has been implemented in real time.
[10]	A combination of simple processing techniques with a temporal analysis gives reasonable target tracking results.
[51]	Some tracked points are manually selected. Algorithm is not robust to temporary obscuration or rotation of target.
Comparisons of techniques	
[17]	Adapts three techniques to compare them with a fourth. The unmodified technique performs quite well.
[80]	A variety of algorithms is tested on a large data set. Results indicate that a combination of algorithms gives the best performance.
[79]	Three fairly complicated techniques are compared. All the techniques require extensive processing time.
[4]	Considers the problem of infrared camouflage. Suggests that one algorithm alone is insufficient to classify the background.
Data integration	
[62, 63]	A bold attempt to model all processes involved in heat transfer. Unanswered problems include how to segment images and how to register thermal and visible images. Many of their assumptions will be invalid in many practical applications.

Table 8.1: Table summarising merits and shortcomings of a selection of papers from the infrared image processing literature

the investigation of the applicability of state of the art computer vision solutions to infrared imaging problems. Parallel architectures are being developed which make it possible to perform complicated vision tasks at speeds approaching real time. The vision algorithms should not be imported into infrared problems uncritically, however, as has been done in much of the previous work on infrared, because of the different nature of infrared image data as discussed in chapter 2 and in the next section.

8.2 Experimental Results

In this section, we aim to show that the assertions that we have made about thermal images are valid and that algorithms developed for vision may not be successful with thermal images. We show that infrared images contain more noise than visible images; that features prominent in visible images are often less prominent in infrared images; and indeed that they are sometimes not apparent at all.

We have selected a representative sample of thermal images with our sensor and recorded CCD images of the same or similar scenes with similar resolutions. We have applied a selection of image processing algorithms to both sets of images in order to demonstrate how infrared images differ from visible images.

Figure 8.1 shows a typical pair of input images: the two images in the left column are thermal, while those in the right column are visible light images of similar scenes. We note that in choosing the threshold values required to get a reasonable contrast in the thermal image of a person drinking tea, low intensities

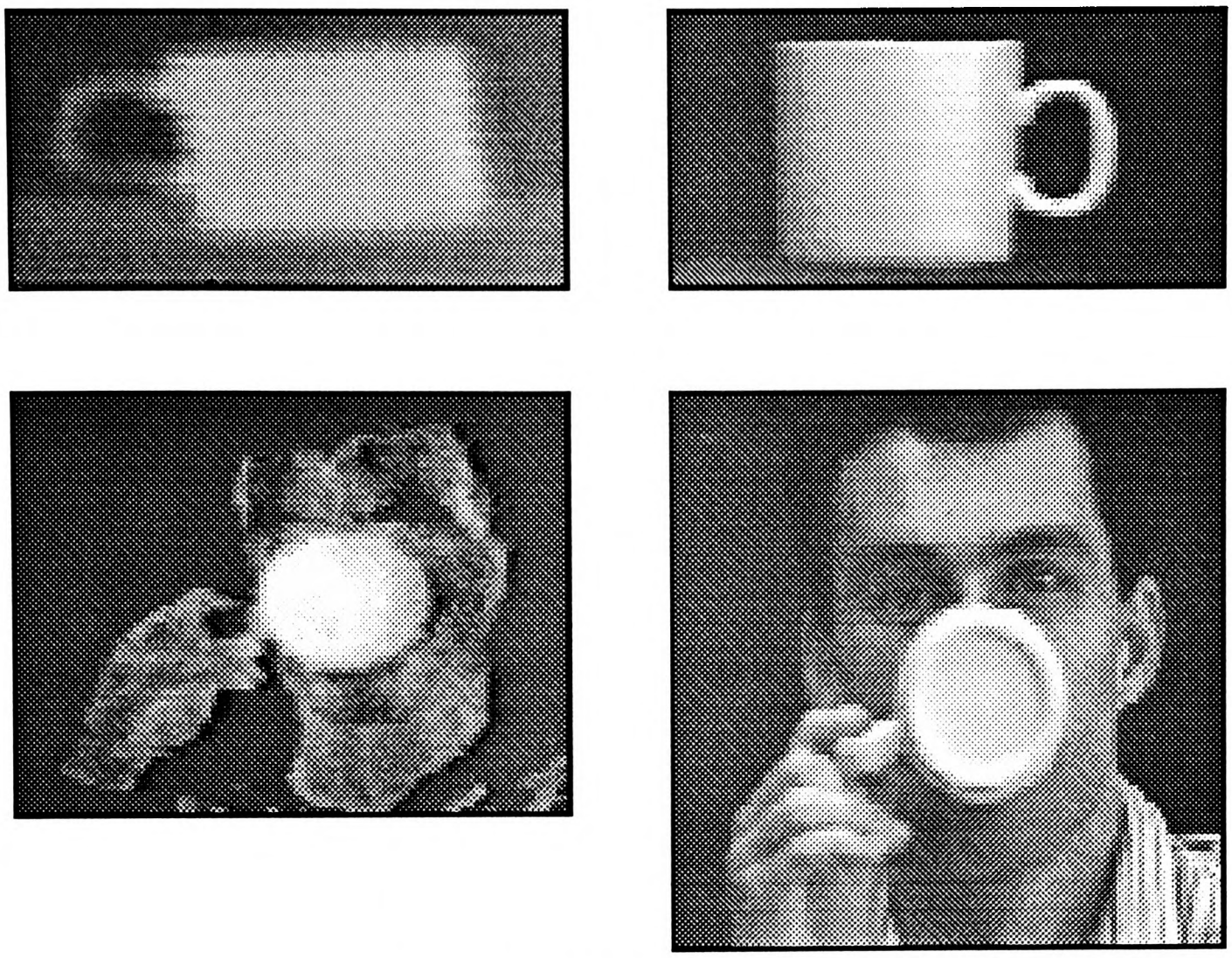


Figure 8.1: Input images

were reduced to zero and the background became uniformly dark.

8.2.1 Noise

We have suggested that thermal images typically contain more noise than visible CCD images. In this section, we present figures to support this argument.

To illustrate how thermal images contain more noise than visible images, we have plotted an intensity cross-section across the thermal image of the cup. The result is presented in figure 8.2. Figure figure 8.3 depicts an intensity cross-section across the visible image of a cup. Examining the plots, it is clear that the cross-

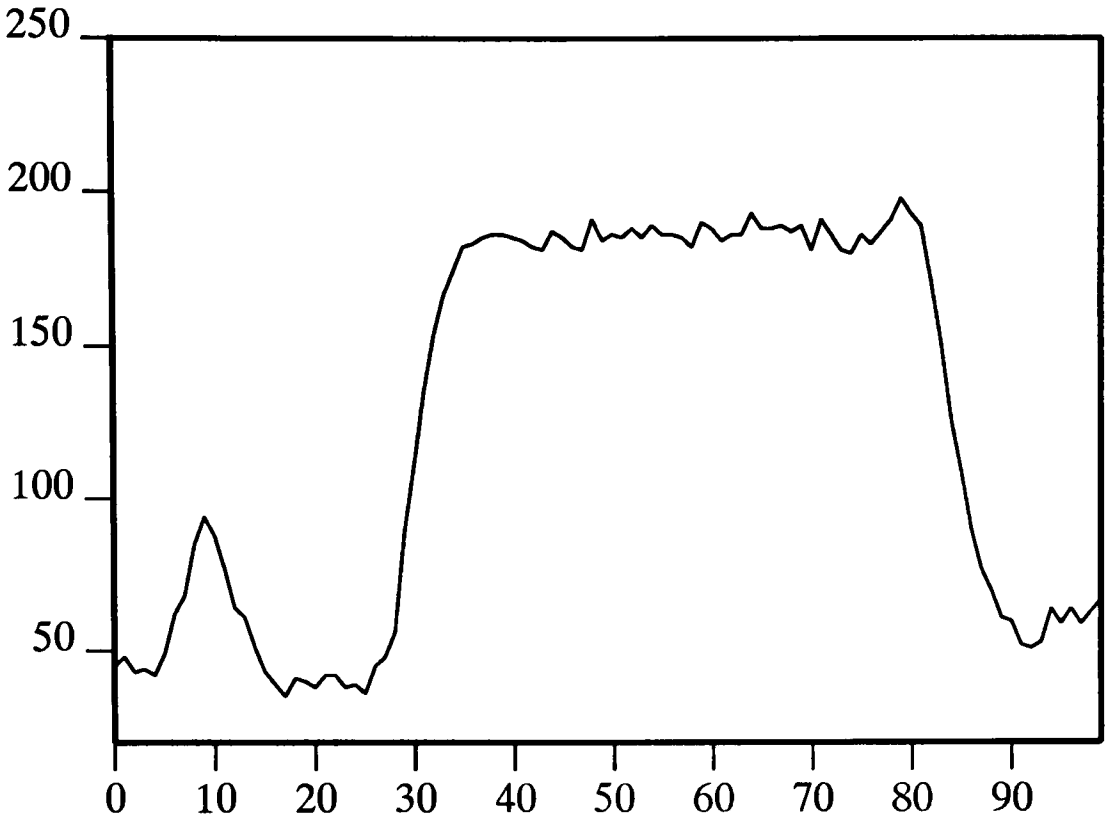


Figure 8.2: Intensity profile across thermal image of cup

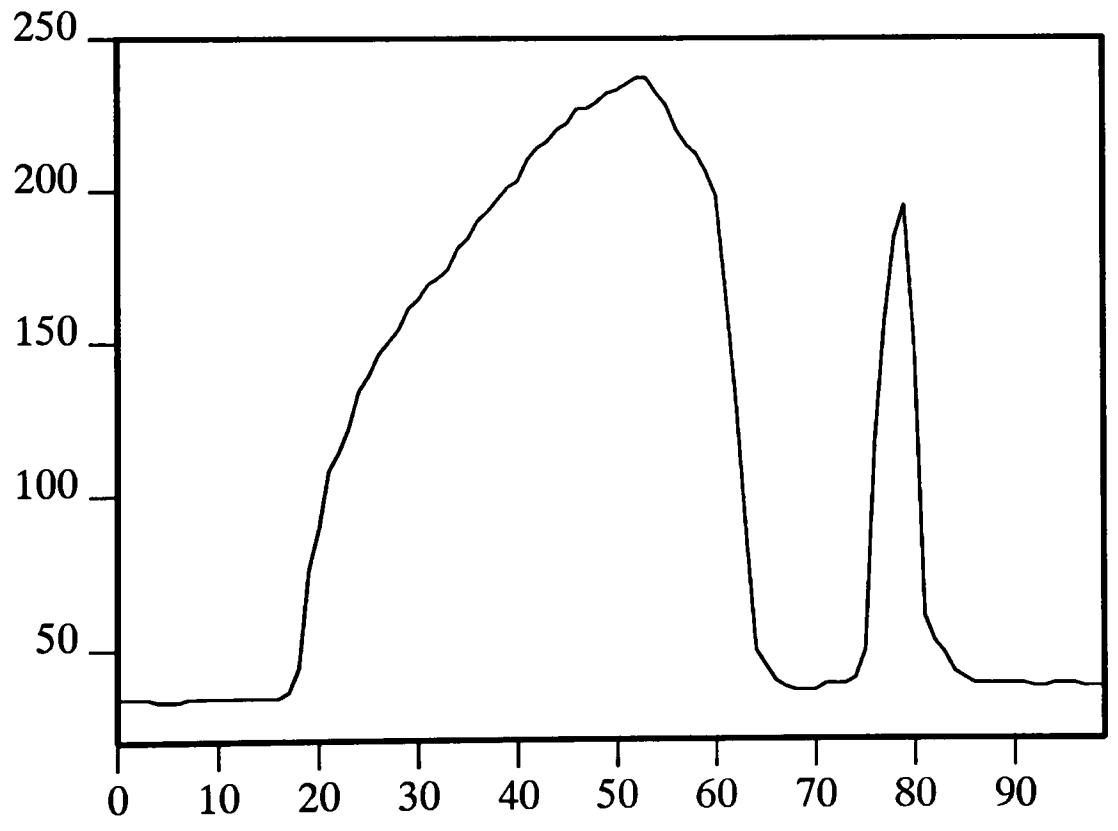


Figure 8.3: Intensity profile across visible image of cup

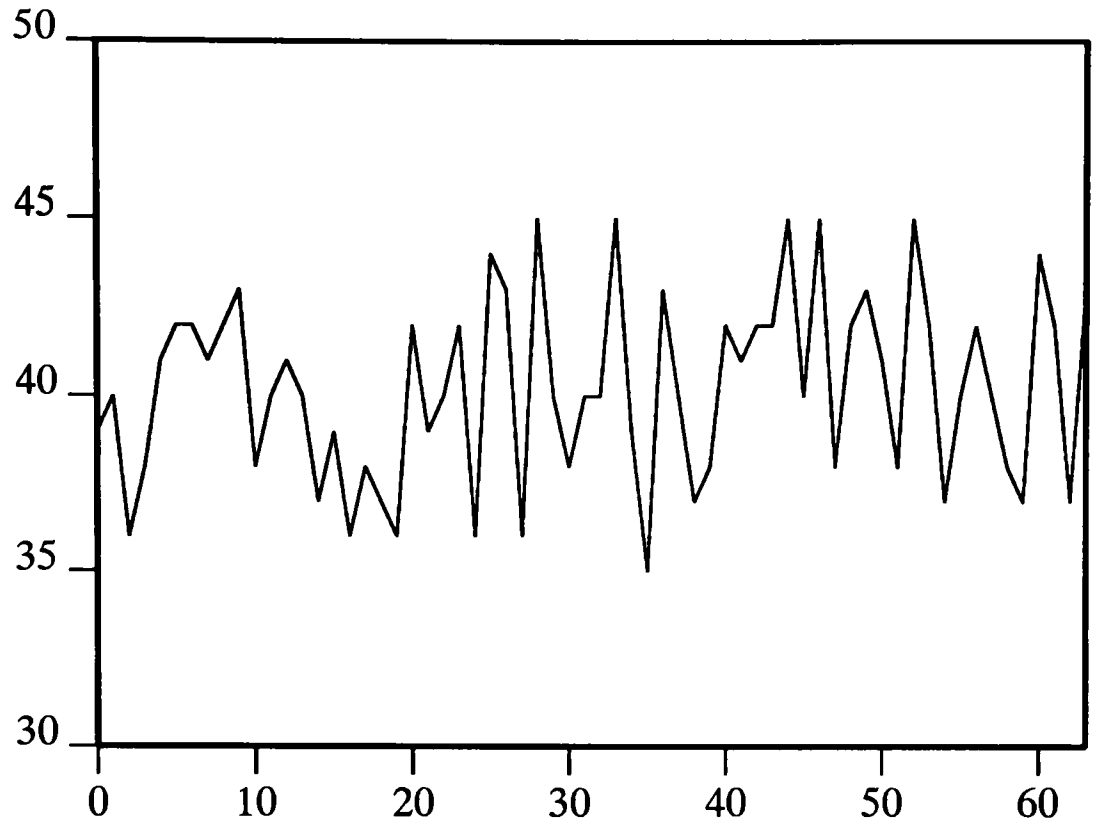


Figure 8.4: Intensity profile across background portion of thermal image

section across the thermal image is noisier than the cross-section across the visible image.

The cup images in figure 8.1 are sub-images of larger images which contain areas of constant background. We took sub-images of the background in each case. In figures 8.4 and 8.5 we present cross-sections across the thermal and visible background images respectively. Once again, it is clear that there is more noise in the cross-section of the thermal image than in the visible image.

For each background image, we subdivided the image into non-overlapping ten pixel by ten pixel patches and calculated the standard deviation of the intensity values in each patch. The results for the thermal image are printed in table 8.2 and for the visible image in table 8.3.

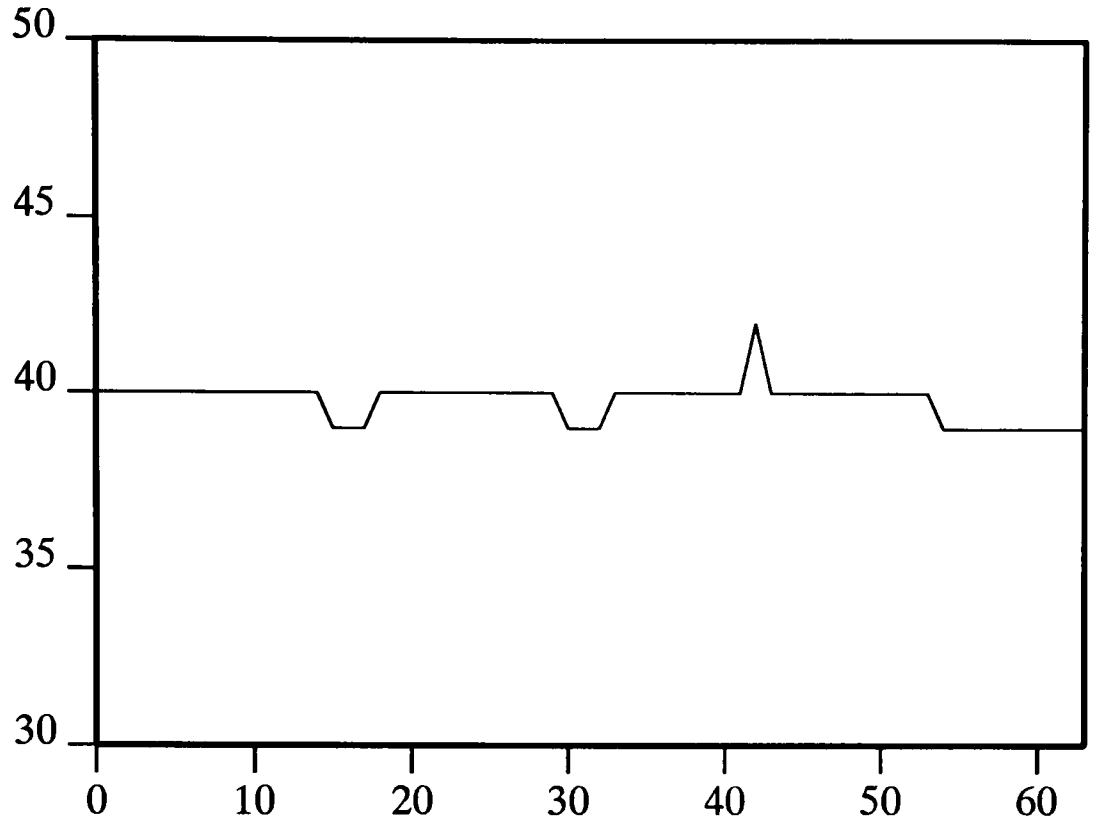


Figure 8.5: Intensity profile across background portion of visible image

	0-10	10-20	20-30	30-40	40-50	50-60
0-10	2.852	3.462	3.257	3.635	3.224	3.719
10-20	2.862	3.301	3.598	3.667	3.494	3.357
20-30	2.710	3.092	2.897	3.315	3.051	3.261
30-40	3.785	4.009	3.766	4.176	3.762	3.698
40-50	4.524	5.393	4.970	4.790	4.639	4.942
50-60	3.479	3.747	3.949	3.789	3.670	3.548

Table 8.2: Standard deviations: thermal image

	0-10	10-20	20-30	30-40	40-50	50-60
0-10	0.307	0.383	0.365	0.426	0.368	0.336
10-20	0.378	0.414	0.421	0.483	0.463	0.564
20-30	0.301	0.313	0.304	0.355	0.486	0.297
30-40	0.410	0.426	0.481	0.445	0.402	0.397
40-50	0.362	0.415	0.379	0.409	0.371	0.384
50-60	0.444	0.570	0.391	0.305	0.429	0.271

Table 8.3: Standard deviations: visible image

The results in the tables clearly indicate that, as we expected, the noise (as estimated by the standard deviation) in the thermal image is much greater (by an order of magnitude) than the noise in the visible image.

8.2.2 Simple edge detection

In this section, we consider one of the simplest filters for edge detection: specifically the Sobel operator which has been used by a number of researchers in the infrared field - usually in the first stage of processing (e.g. [75], [22], [58], [79]).

In figure 8.6, we show the output of a Sobel operator on the images in figure 8.1. The first column illustrates the output of the Sobel mask as an inverted grey scale image where darker values are high and lighter values are low. The other three columns show the effect of thresholding the output of the Sobel operator with increasing thresholds of 10, 25 and 40.

We argued in the previous section that the Sobel edge detector in general performs poorly. The results presented in figure 8.6 show that this is indeed the case. For both thermal and visible images, the output of the Sobel detector is

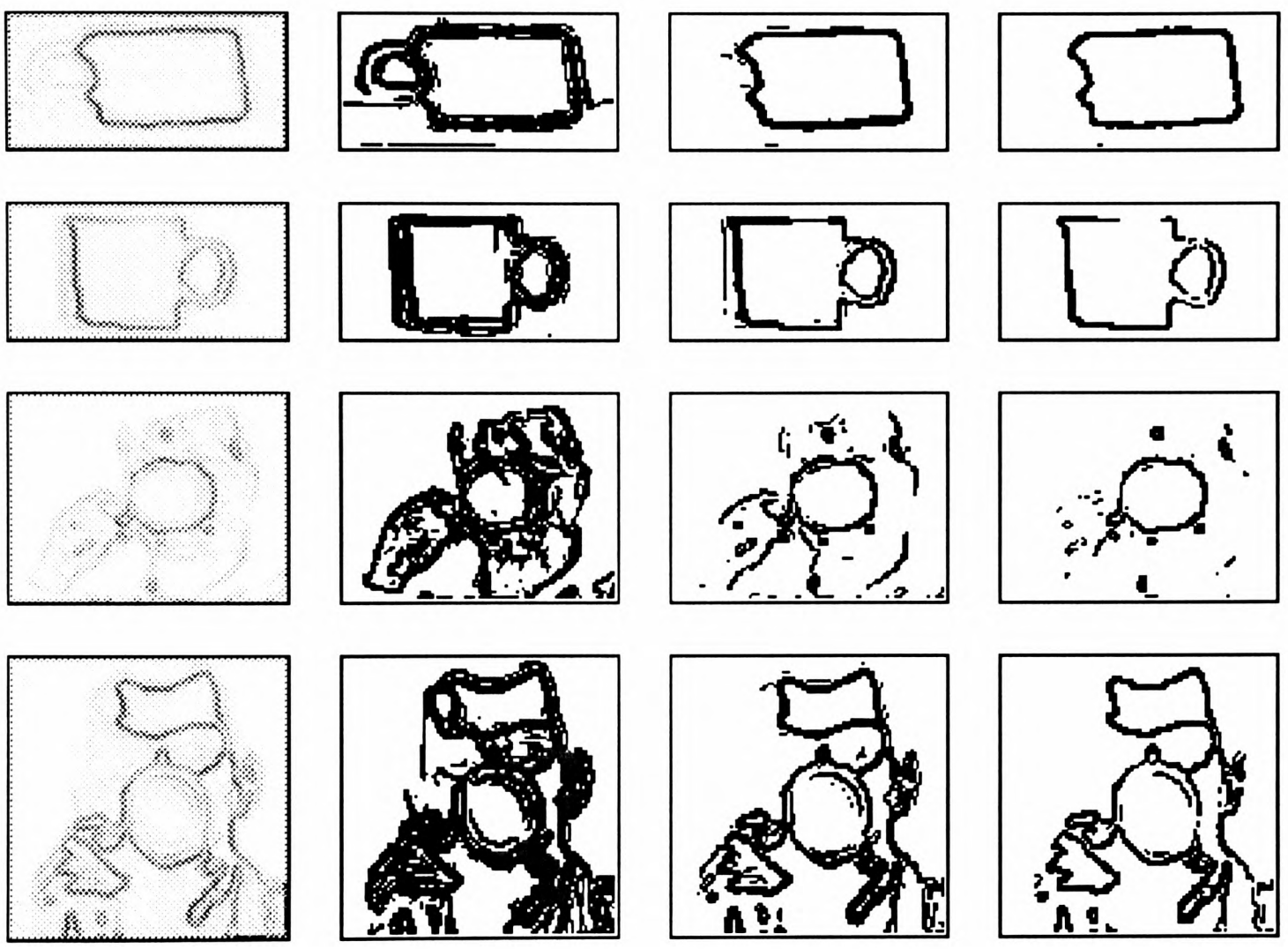


Figure 8.6: Output of Sobel operator

unsatisfactory. It is highly cluttered if a low threshold is used; while important edges are lost if higher thresholds are used. Even the image of the cup in the visible image, which has quite a high contrast with its background, illustrates how some edges are lost when the threshold is made high enough to remove multiple responses from an edge. Indeed, in the resulting thresholded Sobel output, the left edge of the cup has been lost and the edge detected is due to sharp shading changes on the cup itself. The effect of thresholding is more pronounced still in the thermal images: the handle of the cup in the cup image is lost even with a medium threshold, as are the edges of the face.

The handle of the cup in the thermal image is lost because it is not as warm as the body of the cup and therefore does not provide such high intensity changes. This is not true of the handle of the cup in the visible image.

A different effect is noticed in the images of the faces. The shaded side of the face in the visible image is lost for low thresholds but the illuminated side remains a clear edge even for high thresholds. In the thermal image of a face, we note that no real difference is present between the edge strengths on either side of the face even though the image was taken in similar lighting conditions to the visual image of a face.

Both the cup handle effect and the face effect can be explained by remembering that the intensity in the visual images depends on reflected light while the intensity in the thermal images depends on emitted heat. As the handle of the cup has

similar reflective properties to the rest of the cup, the intensity of the handle in the visual image is similar to that of the main body of the cup and as the face emits heat approximately equally across its surface, the strength of the edges on both sides is similar.

The cross-section plots in figures 8.2 and 8.3 illustrate the effect which illumination shading has on the the two types of image. The cross-section of the cup from the thermal image is approximately constant, while the cross-section of the visual image shows a very marked shading effect.

Clearly, as we have stressed earlier, the thermal images are different from visible images and simply importing without modification algorithms developed for visible images, which often depend on the output of a simple edge detector such as Sobel's, may not be successful for thermal images.

8.2.3 Hough transform

In this section, we show that, given the right problem, some algorithms developed for computer vision can be applied to infrared images. To illustrate this, we have written a program to perform a Hough transform on an image. The program has been run on the set of images depicted in figure 8.1.

The particular form of Hough transform program we wrote was a circle detector, using edge information. (The effectiveness of a Hough circle detector was examined by Series et al [79]. Their results are summarised in section 8.1.) We used a Sobel edge detector with a medium threshold (see section 8.2.2) to give edge

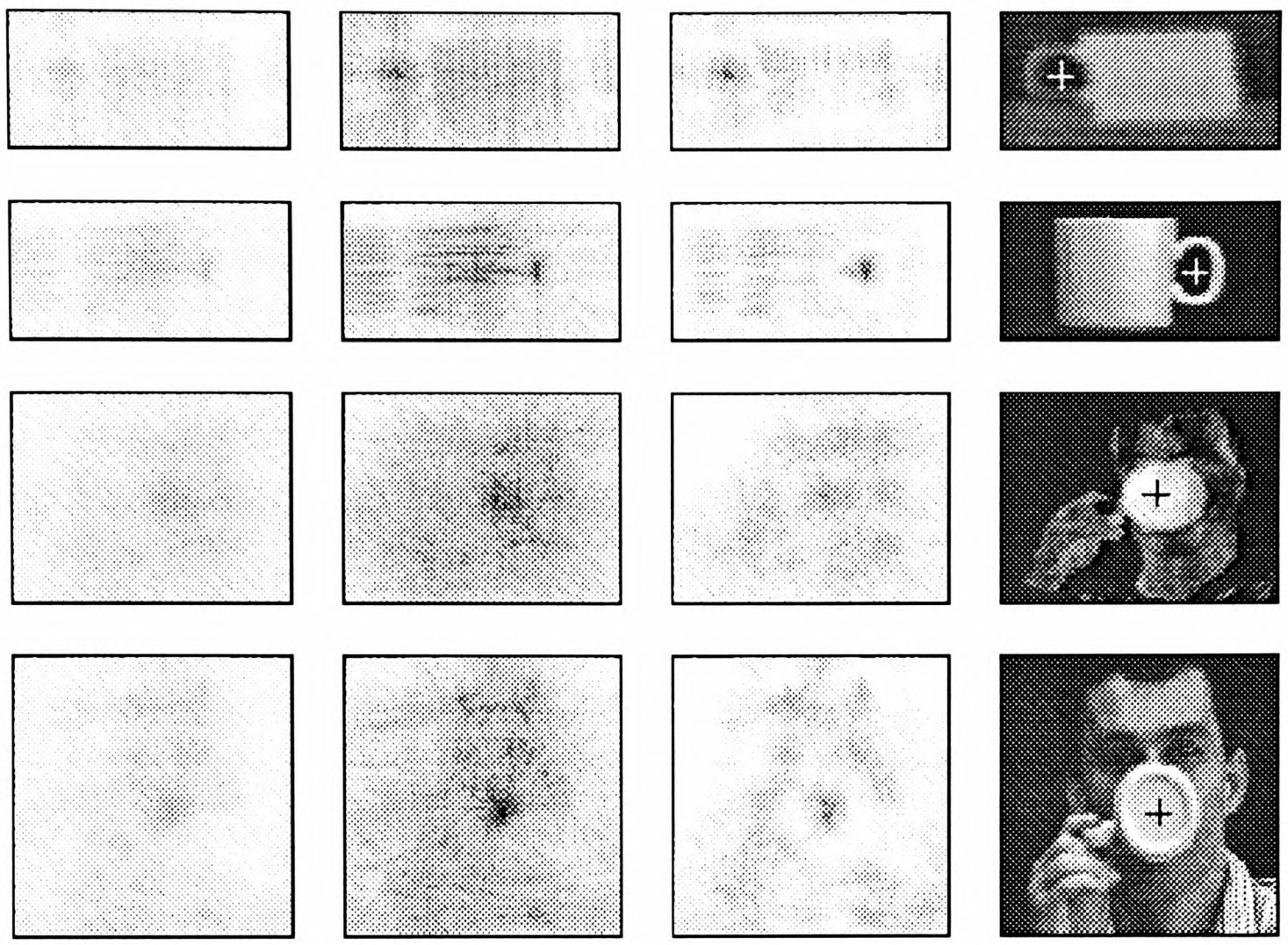


Figure 8.7: Output of Hough transform

magnitude and direction for every pixel in the image. We used the edge magnitude to threshold the image. For all edge points, we incremented the Hough array for all points which were on a line perpendicular to the edge direction. The resulting Hough images had peaks in locations which were centres of circles.

A summary of the results from the program is presented in figure 8.7. In all the Hough images, the darker the plot, the higher the count is in the Hough array. The left most column is the actual output of the Hough transform. To make the Hough images easier to interpret, we have contrast stretched all the values in the

second and third columns (by the same amounts), setting the highest value to maximum blackness. The second column is simply a scaled version of the first column.

For all of the images, the Hough transform correctly located the approximate centre of the circles contained in the images (the base of the cup in the two images of faces and the centre of the handle of the cup in the two cup images). To tune the program for this particular problem, we specified that only circles with radii between 5 and 15 pixels were of interest. The outputs from this tuned program are given in the third column of figure 8.7.

The predicted locations of the centres given by the peak in the Hough image for the tuned and non-tuned programs are actually quite similar. We note however, that using the tuned program it is very much easier to detect the required circle centre from the false circle centres, particularly for the cup images. The tuned algorithm would therefore be far less susceptible to noise.

We also note from the face images, that the peak in Hough space due to the base of the cup in the visible spectrum is much higher and sharper than that for the thermal image. This is because the base of the cup has a circular rim on it which is detected by the visible imager, but not the thermal imager. (Hough circle detectors are particularly good at locating concentric circles.)

In the fourth column of figure 8.7, we reproduce the images from figure 8.1 with a cross on each indicating the location of the most likely centre of a circle

predicted by the Hough transform. The centres used were those predicted by the tuned version of the program, though these only differed by one or two pixels from those predicted by the non-tuned version.

The results from the Hough transform indicate that for the images chosen, the Hough transform can be used successfully with thermal images. We would therefore expect that it would be suitable for any infrared circle recognition task where the target stands out well from its background: hot wheels from a recently driven car for example.

8.2.4 Canny operator

In this section, we consider a popular edge detector due to Canny and examine the difference in its performance when applied to both thermal and visible images.

We applied Canny's edge detector to the set of images depicted in figure 8.1 with a large range of different values for the three variable input parameters: the scale of smoothing; and the upper and lower cut-off thresholds of the hysteresis. A selection of the results is presented in figure 8.8.

The images in the first column are the results when the default settings of smoothing and lower and upper hysteresis thresholds (s=1, l=2, u=6) were used. The second column shows the effect of increasing the smoothing (to 2) while leaving the other parameters unaltered. The third column shows the effect of raising the hysteresis thresholds to 5 and 15 while keeping a smoothing value of 1. The final column shows the effect of a smoothing of 2; and lower and upper

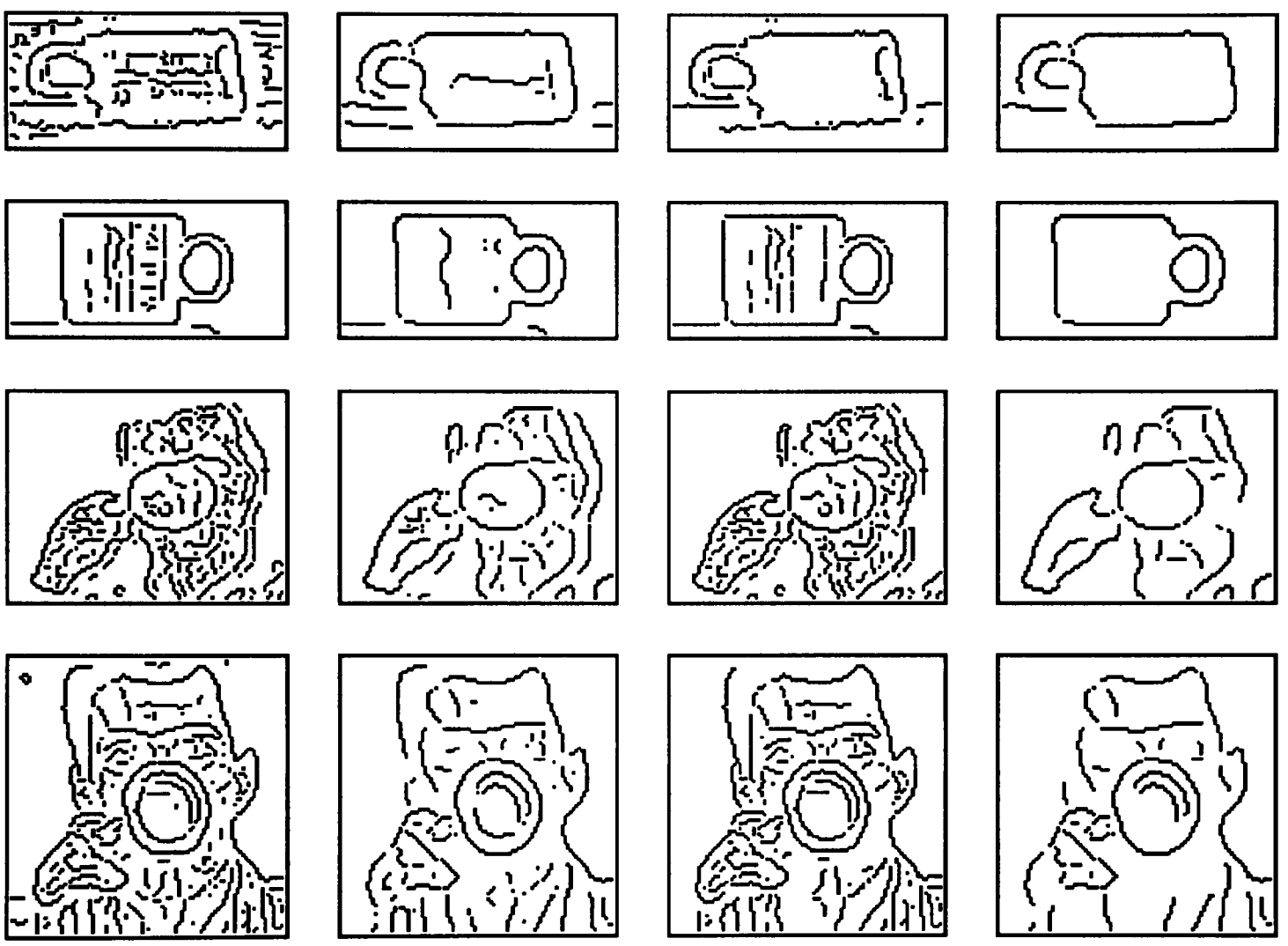


Figure 8.8: Output of Canny edge detector

hysteresis thresholds of 10 and 30.

Looking at the results for the visible face image in all columns, we notice that facial features are found quite well: eyes, eyebrows and ears are highlighted by the edge finder (especially for lower hysteresis thresholds). In the thermal images, however, very little recognizable facial detail is present. This is because human skin is approximately a black body emitter so any variations are due primarily to surface differences caused, for example, by perspiration, facial hair or to internal heat sources (e.g. veins).

In the thermal face image, the only recognizable detail comes from the boundaries of the face, the hand and the cup. Although some other edges are detected in the thermal face image as a consequence of variable temperatures on different parts of the face, the object boundaries would be the most useful edges for object recognition as they are more typical of the object.

Examining the cup images, we see that for high thresholds, the boundary of the cup in the visible image has been extracted extremely well. As we would expect from the results in section 8.2.2, the handle of the cup in the thermal image has been detached from the body of the cup. We notice that once the effect of the noise is ameliorated by smoothing, the performance of the detector on the two types of image is quite similar.

The cup images in figure 8.1 were of a plain white cup. We tested the Canny algorithm on a different cup: one with a pattern on it. The left hand column of

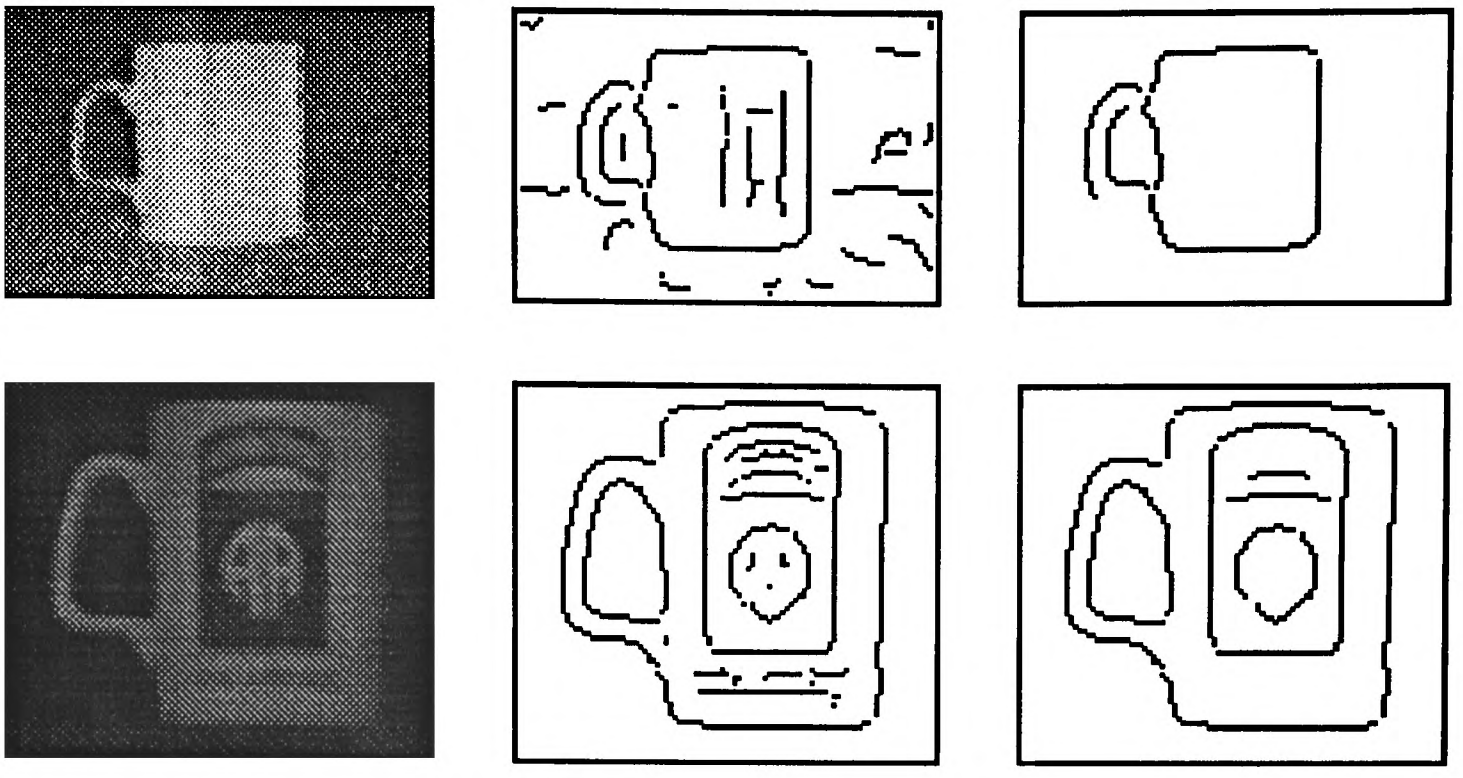


Figure 8.9: Output of Canny edge detector on patterned cup

figure 8.9 shows an image of the cup taken with the thermal and visible cameras. We notice immediately that the thermal camera has failed to detect any detail of the pattern in the image. The second and third columns of figure 8.9 show the output of the Canny edge finder on the two images, with smoothing 2 in both cases and with hysteresis thresholds of 2 and 6 in the middle column and 10 and 30 in the right hand column. We can see that the broad outline of the pattern has been detected in the Canny output from the visible image, while no useful detail is detected in the output of Canny on the thermal image, with the exception of the cup and handle boundaries.

It is not always the case that surface markings are not detected by thermal imagers. We showed in chapter 6 that painting the surface of a metal cylinder with matt black paint dramatically changed its radiance properties. With the patterned cup image, however, the object was an approximate black body emitter

with or without the surface pattern so the pattern could not be detected.

8.3 Conclusions

In the second half of this chapter, we have seen that infrared images are indeed different from visible images. In section 8.2.1 we showed that, at least for the examples chosen, thermal images are noisier than visible images. We demonstrated in section 8.2.2, that the noise affects the performance of simple edge detectors, illustrating our case with the example of a Sobel detector.

We have shown in section 8.2.4 that it is not always possible to extract edges from thermal images, even with high quality edge detectors.

In section 8.2.3, using a Hough transform as an example, we have shown that for certain problems image processing algorithms can be successful when run on infrared images.

Throughout section 8.2, we have illustrated how the features present in infrared images are not necessarily those which would be apparent in visible images. We have shown that for the majority of objects, the dominant features in infrared images will be bounding contours of objects; internal heat sources; and for certain surfaces, changes in surface properties. Most importantly perhaps, we would not expect to find shading due to the illumination of the sun. Recalling the work in chapter 2, we would also expect mutual illumination to be important in the formation of infrared images, but only for non-Lambertian reflectors and emitters of infrared energy.

Chapter 9

Conclusions

An extensive literature survey of infrared image processing (chapter 8) has revealed that the vast majority of work published on infrared image processing takes little or no account of the fundamental laws which give rise to the infrared images. It additionally shows that processing of infrared images has primarily been restricted to low-level image enhancement; while more complicated algorithms are often imported without adaptation from computer vision.

This thesis critically examines the assumption that infrared images can be processed in the same way as visible images and points the way towards an understanding-based approach towards infrared image processing. A comparison of the performance of a selection of conventional image processing algorithms run on both infrared and visible images is presented in chapter 8. The experiments are chosen to highlight some of the differences between infrared images and visible images.

We have shown in chapter 2 that the irradiance pattern from black bodies

gives no indication of surface orientation: that the only information which can be derived from it is the bounding contours of the bodies and the location of any internal heat sources. For non-black bodies, we have shown that for a constant surface type and temperature, the irradiance from a surface element can be used to indicate the orientation of the surface. For a variable surface type or temperature, the irradiance will be a complex function of surface emissivity, surface geometry and surface temperature. Furthermore, in the presence of other emitting surfaces, the radiance pattern of those surfaces and the geometry of the scene will also influence the radiance of the surface being viewed.

To facilitate experimental work, a line-scanned array pyroelectric thermal imager was built and interfaced with the computer system in Oxford. The design and construction of the imager are described in chapter 4. Images taken with this imager demonstrate that the intensity profiles we predicted in chapter 2 do indeed occur in real images.

We have shown both theoretically (chapter 2) and experimentally (chapter 6) that mutual illumination effects give rise to distinctive intensity profiles in concave objects and we have demonstrated in chapter 6 that the intensity profile of a body can be used as a matching template for locating the presence of a particular type of object in a scene.

This work lays the foundations for further study into understanding and predicting the intensity profiles of infrared images and hence towards the development

of detection and recognition algorithms based on meaningful target signatures. The procedure used in the thesis was to find closed form solutions to the infrared image irradiance equation for simple geometric shapes. Using more complicated mathematics, such as finite element analysis or other numerical methods, it would be possible to analyse more complex image shapes. The work was done on only one dimensional data due primarily to the restriction of the imager type, but the principles apply just as well in two dimensions.

Most of the algorithms described in this thesis operate on local neighbourhoods in an image so it would be possible to implement the techniques in parallel. In this way, they could be made to run in real time thereby satisfying the requirements of many of the likely infrared imaging applications.

Appendix A

Calculation of preamplifier component values

Calculating resistor & capacitor values

With reference to figure 4.6, the values of the resistors and capacitors were calculated as follows.

Assuming the op-amp has infinite input impedance, then the current flowing through R_1 plus the current flowing through C_1 is equal to the current flowing through R_2 and C_2. Hence

$$\frac{V_0}{V_1} = 1 + \frac{Z_1}{Z_2} \tag{A.1}$$

where $Z_1 = \frac{R_1}{1+j\omega R_1 C_1}$ and $Z_2 = R_2 + \frac{1}{j\omega C_2}$. This configuration has the property that for low frequencies, $\frac{V_0}{V_i}$ is approximately proportional to frequency, f; while for high frequencies, $\frac{V_0}{V_i}$ is approximately proportional to $\frac{1}{f}$.

If for low frequencies we simply wanted the gain to be proportional to the frequency, with a gain of 30 at $f = 150$; by making the assumption that the impedance of a capacitor is very high at low frequencies, we could say: $\left|\frac{V_0}{V_i}\right| =$

$\frac{f}{5} = \left|1 + j\omega R_1 C_2\right|$; hence:

$$R_1 C_2 \simeq \frac{1}{10\pi} \tag{A.2}$$

Similarly, if for high frequencies we wanted the gain to be inversely proportional to the frequency; making the assumption that the impedance of a capacitor is negligible, we could say: $\left|\frac{V_o}{V_i}\right| = \frac{4500}{f} = \left|1 + \frac{1}{j\omega C_1 R_2}\right|$; hence:

$$R_2 C_1 \simeq \frac{1}{9\pi * 10^3} \tag{A.3}$$

At the frequency corresponding to maximum gain, we wish to match the impedances. We therefore have: $R_1 = \left|\frac{1}{j\omega C_1}\right|$ and $R_2 = \left|\frac{1}{j\omega C_2}\right|$; hence:

$$R_1 C_1 = \frac{1}{300\pi} = R_2 C_2 \tag{A.4}$$

Combining equations A.2 to A.4 and choosing a value for R_1, we have component values for C_1, C_2, and R_2.

These calculations are approximately accurate for high frequencies and for low frequencies, but for all other frequencies, the assumptions made about the impedance of the capacitors will clearly not be valid. For frequencies close to 150Hz, the gain will be reduced from the value assumed in the calculations. One simple way to compensate for this is to raise the value of the "desired" gain. Repeating the calculations with a required gain of 100 gave component values of a filter which quite accurately followed the required filter characteristics. Accuracy in the exact value of gain was not crucial as gain could be altered later in the data path.

Now, consider R_i and C_i. From the design specification, we want a cut off frequency at around 20Hz ($\omega = 40\pi$). Given that $R_i C_i = \tau = \frac{1}{\omega}$, R_i and C_i were calculated, thus completing the specification of the component values.

Appendix B

Component values

This appendix contains component numbers and values of the components used to construct the thermal imager described in chapter 4.

B.1 Optics and Detector Array Dimensions

The detector array was a sixty four element pyroelectric array designed to operate in the $8-14\mu$ waveband. The detector elements are approximately 160μ square and have a 200μ pitch. The array is mounted behind a germanium filter window which is 13mm by 19mm and mounted on a microphony reduction block which has dimensions 24mm by 31mm and is approximately 25mm deep. The microphony reduction block is attached to a mounting base which has dimensions 30mm by 55mm.

The lens used with the imager was a fixed focus 50mm f0.7 germanium lens for $8-14\mu$ infrared radiation.

B.2 Master Oscillator

The master oscillator used was a 4047B monostable/astable multivibrator used in its astable configuration. A 100pF capacitor, C_t was connected between pins 1 and 3. (It is recommended that $C_t \geq 100pF$, and we wanted a small C_t to give a fast frequency.)

The equation defining the time constant of the multivibrator is $t_A = 4.4R_tC_t$ where R_t is the external resistor connected between pins 2 and 3. To get a frequency range of approximately $20kHz \leq f \leq 200kHz$, we chose R_t to be a $10k\Omega$ resistor in series with a $100k\Omega$ resistor. (It is recommended that $10k\Omega \leq R_t \leq 1M\Omega$.)

To operate the multivibrator in its astable mode, pins 4,5,6 and 14 were tied to V_{DD} and pins 7,8,9 and 12 were tied to ground.

B.3 Preamplifiers

The design of the preamplifiers is described in appendix A. The chips used were 074 quad op-amps. The component values were: $R_1 = 1M\Omega$, $R_2 = 10k\Omega$, $C_1 = 1nF$, $R_i = 1M\Omega$ and $C_i = 8nF$.

B.4 Multiplexers

The multiplexers used were HI-506 sixteen channel single-ended multiplexers. The four-bit address of the required channel was passed along the back board along

with four enable signals, one of which was switched on for each of the multiplexers.

B.5 Amplifier

The multiplexed signal was amplified before arriving at the sample and hold. An OP27 op-amp was used with a $10k\Omega$ resistor between the negative input and the output; and a 180Ω resistor in series with a $5k\Omega$ potentiometer connecting the negative input to ground. By adjusting the potentiometer, the gain of this amplifier could be varied between 2 and 55.

B.6 Sample and Hold

An HA-2425 fast sample and hold chip was used to give a steady input to the ADC. The typical unity gain circuit described in the Harris data book was used. $47k\Omega$ resistors joined pins 3 and 4 to pin 9 ($V+$). A $1nF$ capacitor was used as the hold capacitor joining pins 11 and 13. The control signal from the EPLD arrived at pin 14, the input from the amplifier on pin 2 and pins 1,7,10 and 11 were connected to the output.

B.7 A to D Converter

The analogue to digital converter used was an RS574AJD twelve bit A to D. A short negative pulse from the EPLD to the READ/$\overline{\text{CONVERT}}$ pin begins the conversion of the analogue signal input on pin 13. 50Ω resistors are connected between pins 8 and 10 and between pins 8 and 12. Pins 1,2, and 6 are held high,

while pins 3 and 4 are held low. This sets the ADC to begin conversion on a low pulse from the EPLD. The twelve digital outputs of the ADC (pins 16-27) are buffered before being sent to the datacube.

B.8 Buffers

Two SN74LS241 octal buffers were used to buffer the twelve bit output of the ADC before sending it to the Datacube. One of the spare buffers was used to buffer the output of the opto-switch, which was also sent to the datacube to give information about whether the imager was detecting the scene or the chopper.

The buffered outputs, along with timing signals from the EPLD were sent along a 32 wire ribon cable to the datacube.

B.9 Datacube

The following section of code describes the switches required to set up the maxscan board to read data from the imager.

```
scRegSet(scfd);                    /* sets all physical registers to zero */
scRoiXEnMd(scfd, DQ_DISABLE);      /* disable ROI and fifo transfers */
scMaster(scfd);                    /* Max scan is master interupter */
scDecMd(scfd, 0);                  /* no decimation */
scILevel(scfd, 0);                 /* disables interupts */
scOCMMd(scfd,0);                   /* something to do with the OCM */
scClockMd(scfd, 0);                /* Zero delay for A/D clock bits */
scInpMd(scfd, SC_DIG);             /* Select digital input */
scMHzMd(scfd, SC_10MHZ);           /* 10 MHz mode */
scRLEMd(scfd, DQ_DISABLE);         /* Disable run length mode */
scDSPByp(scfd, DQ_ENABLE);         /* Enable DSP bypass */
scRoiMasMd(scfd, DQ_SET);          /* set max_scan as ROI master */
```

```
scMaxBus(scfd, DQ_OFF);              /* select P4 ROI timing bus */
scRoiBus(scfd, 0);                   /* select ROI bus 0 */
scGPSMd(scfd, DQ_DISABLE);           /* Disable gain per sample mode */
scRoiXferMd(scfd, SC_SSHOT);         /* Single_shot mode (no entry in primitives) */
scRFifoPort(scfd, DQ_SET);           /* FIFO read port = VME bus */
scIntlMd(scfd, SC_NORMAL);           /* Non interlaced */
scS1(scfd, DQ_UNSET);                /* ext vert input to sync conditioner from PS7 */
scS2(scfd, DQ_UNSET);                /* Horizantal clock selected by S17 and PS8 */
scK3S5(scfd, 0x0A);                  /* S4 output - reference clock directly */
scS6(scfd, DQ_SET);                  /* Camera clock return (after PS1) */
scS7(scfd, DQ_UNSET);                /* Time clock (not clock/2) clocks dig. proc */
scS9(scfd, 1);                       /* Selected horiz. reset enable */
scS10(scfd, DQ_UNSET);               /* HLUT reset disabled */
scS12(scfd, DQ_SET);                 /* Vertcal resets from host computer enabled */
scS13(scfd, DQ_UNSET);               /* Even field P3 Vsyncs ignored */
scS14(scfd, DQ_SET);                 /* Ext. vertical resets through PS7 enabled */
scS15(scfd, DQ_UNSET);               /* Vert resets from lowpass&edg_det ignored */
scS16(scfd, DQ_UNSET);               /* Vertical resets from VLUT not enabled */
scS17(scfd, 2);                      /* Select Ext. horiz input */
scS18(scfd, DQ_UNSET);               /* Normal time clock (1 = clk/2) */
scS19(scfd, DQ_UNSET);               /* clk ret after ps1 */
scS23(scfd, DQ_UNSET);               /* Maxbus odd field vert not used */
scPS1(scfd, DQ_SET);                 /* dont invert camera return clock */
scPS3(scfd, DQ_UNSET);               /* uninverted time clock for horiz resyc */
scPS6(scfd, DQ_UNSET);               /* data clk uninverted for data sync clk */
scPS7(scfd, DQ_UNSET);               /* EXTVIN not inverted, high resets counter */
scPS8(scfd, DQ_SET);                 /* EXTHIN not inverted, rising edge used */

/* Set Horizontal and Vertical look up tables */
scHSetC(scfd, 0, 0x0000, 0x1fff);    /* clear HLUT */
scVSetC(scfd, 0, 0x0000, 0x1fff);    /* clear VLUT */
scHSet(scfd, 4, 0, 63);              /* horizontal aquire enable for pixels 0-63 */
scHSet(scfd, 1, 0, 0);               /* clock vert timing at start of each line */
scVSet(scfd, 4, 0, 100);             /* vertical aquire enable for lines 0-100 */
sleep(1);                            /* wait before reseting fifo */
scFifoCl(scfd);                      /* Reset fifo */
scRoiXEnMd(scfd, DQ_ENABLE);         /* Enable ROI transfer */
```

B.10 EPLDs

The EPLDs were 24 pin Altera EP600 devices. One of the EPLDs was configured with 1 input (the clock from the master oscillator) and 12 outputs; while the other, clocked by the master oscillator, was configured with no inputs and only one output to the phase locked loop.

B.11 Opto-pickup

A slotted opto switch with logic (RS stock number 304-560) was used to detect the position of the chopper. This device, consisting an LED, photodiode, amplifier, voltage regulator and Schmitt trigger, needed to be filed down slightly before insertion into the housing because of the nearness of the chopper to the front panel.

B.12 Phase-Locked-Loop

A 4046B Phase-Locked-Loop chip was used to lock the chopper speed to a divided version of the master oscillator frequency. The output of the opto-pickup was fed into the $COMP_{IN}$ input of the PLL (pin 3) and the output of the divider EPLD was fed into the $SIGN_{IN}$ input (pin 14). The inhibit input on pin 5 was held high to disable the VCO and the source follower. No external components were required.

B.13 Chopper motor and drive circuit

The motor used to drive the chopper was an Escap 23L21 which when driven from its maximum of twelve volts was capable of rotating at 7800 rpm. We wanted a speed of 50/8Hz (375rpm) so only a small voltage difference between the motor terminals was needed to keep the motor rotating. It was possible to spin the motor up to speed quickly by applying a higher voltage and rapidly to correct any errors in motor speed which might occur. The standard RSRE chopper control circuit was used to provide the drive values for the chopper motor.

B.14 Current amplifiers

The op-amps used in the chopper drive circuit were not capable of supplying sufficient current to start the motor from rest so two LH0002 high current amplifiers were used to buffer the inputs to the motor.

Appendix C

Calculation of probabilities

If only two elements are dead, then the probability that the two elements are next to each other or are separated by only one pixel is given by:

$$p[1 \text{ near } 2] = 2 \times \left[\left(\frac{62}{64} \times \frac{2}{63}\right) + \left(\frac{1}{64} \times \frac{1}{63}\right)\right] = \frac{125}{2016} \qquad \text{(C.1)}$$

Now consider the problem of three dead elements. The probability that any one is next to or within one pixel of one of the others is given (by equation C.1) as $\frac{125}{2016}$. The probability that the first dead element is within a pixel of the second and is also within a pixel of the third is given by:

$$p[(1 \text{ near } 2) \cap (1 \text{ near } 3)] = \left(\frac{60}{64} \times \frac{4}{63} \times \frac{3}{62}\right) + \left(\frac{2}{64} \times \frac{3}{63} \times \frac{2}{62}\right) + \left(\frac{2}{64} \times \frac{2}{63} \times \frac{1}{62}\right) = \frac{23}{7812} \qquad \text{(C.2)}$$

The probability that all three dead elements are within a pixel of each other (i.e. three in a row) equals three times the probability that one dead element is surrounded by the other two. The probability that the three elements are in a

row is therefore given by:

$$p[\text{3 elements in a row}] = 3\left(\frac{62}{64} \times \frac{2}{63} \times \frac{1}{62}\right) = \frac{1}{672} \tag{C.3}$$

Combining the results of equations C.1, C.2 and C.3 we get the probability that two dead elements will be near each other if three elements are dead:

$$p[\text{2 near with 3 dead}] = \frac{3 \times 125}{2016} - \frac{3 \times 23}{7812} + \frac{1}{672} = \frac{1861}{10416} \tag{C.4}$$

The probability that two elements are next to each other given that there are only two dead elements is given by:

$$p[\text{2 next with 2 dead}] = 2\left(\frac{63}{64} \times \frac{1}{63}\right) = \frac{1}{32} \tag{C.5}$$

The probability that two elements are next to each other given that there are three dead elements equals three times the probability that two elements are next to each other given that there are two dead elements minus the probability that all three elements are in a row. This is given (using the results from equations C.5 and C.3) by:

$$p[\text{2 next with 3 dead}] = 3\left(\frac{1}{32}\right) - \frac{1}{672} = \frac{31}{336} \tag{C.6}$$

The probability for a mask size of five and with three dead elements, that all three elements will fall inside the mask is given by:

$$\begin{aligned} p[\text{3 dead in 5 mask}] = 3\left(\frac{60}{64} \times \frac{4}{63} \times \frac{3}{62}\right) + 3\left(\frac{1}{64} \times \frac{3}{63} \times \frac{2}{62}\right) \\ + 3\left(\frac{1}{64} \times \frac{2}{63} \times \frac{1}{62}\right) = \frac{13}{1488} \end{aligned} \tag{C.7}$$

Appendix D

Bibliography

[1] P.D. Allen and J. Northfield. Algorithm development for infra-red air-to-air guidance systems. In *Proc. 1st International Conference on Advanced Infrared Detectors and Systems*, pages 102–111, October 1981.

[2] Sharon Allen. Mobile infrared thermographic surveys of buildings within a community. *Proc. Society of Photo-Optical Instrumentation Engineers*, 934:16–18, 1988.

[3] Astheimer, R.W. and Schwarz, F. Thermal imaging using pyroelectric detectors. *Applied Optics*, 7:1687–1695, September 1968.

[4] R.G. Aykroyd. Background characterisation using false target rate. Technical report, Department of Mathematical Sciences, University of Durham, November 1989.

[5] R. Bowling Barnes. Diagnostic thermography. *Applied Optics*, 7(9):1673–1685, 1968.

[6] A.C. Baynham. Are the Technologies User Friendly? (Keynote Address). In *Proc. 4th Int. Conference of Advanced Infrared Detectors and Systems*, pages 1–3, IEE, Savoy Place, London, June 1990. IEE conference publication no, 321.

[7] R.C. Beckmann and D.A. Pritchard. A view for security: thermal imagers. Technical Report SAND 90-495C, United States Government, January 1990.

[8] N. Ben-Yosef and K. Wilner. Temporal behaviour of thermal images. *Applied Optics*, 24(2):284–286, January 1985.

[9] N. Ben-Yosef, K. Wilner, S. Simhony, and G. Feigin. Measurement and analysis of 2-d infrared natural background. *Applied Optics*, 24(14):2109–2113, July 1985.

[10] D.C.M. Bilsby. Automatic tank recognition, targeting and tracking. Unpublished RSRE report, 1990.

[11] Thomas O. Binford. Inferring surfaces from images. *Artificial Intelligence*, 17:205–244, 1981.

[12] Michael Brady. Computational approaches to image understanding. *Computing Surveys*, 14(1):3–71, March 1982.

[13] Brady, J.M. and McIvor, A., Brint, A., Foulkes, P., Scott, G., Dickson, W. Vision and the Oxford AGV. In *Proceedings of Image Processing '88*, pages 267–286, London, November 1988.

[14] G. Brown. Real time adaptive processing of thermal images. In *Proc. 2nd International Conference on Advanced Infrared Detectors and Systems*, pages 95–99, October 1983.

[15] J.C. Brown. Detector noise cross correlation measurements. *Proc. Society of Photo-Optical Instrumentation Engineers*, 890:167–170, 1988.

[16] D.E. Burgess, P.A. Manning, and R. Watton. The theoretical and experimental performance of a pyroelectric array imager. Unpublished RSRE report, 1985.

[17] M. Burton and C. Benning. Comparison of imaging infrared detection algorithms. *Proc. Society of Photo-Optical Instrumentation Engineers*, 302:26–32, 1981.

[18] John Francis Canny. Finding Edges and Lines in Images. Technical Report Tech.Rep. 720, Massachusetts Inst. Technol., June 1983.

[19] Charnley, D. and Blisset, R. Surface Reconstruction from Outdoor Image Sequences. In *Proc. 4th Alvey Vision Conference*, pages 153–158, University of Manchester, September 1988.

[20] Clarke, F.J.J. and Leonard, J.K. Radiometric image restoration and calibration for thermal and other video images. *Journal of Physics E (Scientific Instruments)*, 22(10):841–849, October 1989.

[21] C. Corsi, G. Salcito, M. Sommacal, and L. Fiorentino. Image processing and analysis applied to a state-of-the-art ir biomedical system "adir". In *Proc. 2nd International Conference on Advanced Infrared Detectors and Systems*, pages 100–105, October 1983.

[22] S. Cussons. Automatic acquisition of aircraft targets using infra-red imaging sensors. In *Proc. 1st International Conference on Advanced Infrared Detectors and Systems*, pages 29–30, October 1981.

[23] Gary B. Desroches. Stress affected livestock as seen by thermography. *Proc. Society of Photo-Optical Instrumentation Engineers*, 129:120–125, April 1988.

[24] Michael J. Duggin. Descrimination of targets from background of similar temperature, using two-channel data in the 3.5-4.1 μm and 11-12 μm regions. *Applied Optics*, 25(7):1186–1195, April 1986.

[25] C.T. Elliot. Detectors of thermal infrared radiation. *Proc. Society of Photo-Optical Instrumentation Engineers*, 915:9–19, 1988.

[26] C.T. Elliot. Future Infrared Detector Technologies. In *Proc. 4th Int. Conference of Advanced Infrared Detectors and Systems*, pages 61–68, IEE, Savoy Place, London, June 1990. IEE conference publication no, 321.

[27] M.A. Fischler, J.M. Tenenbaum, and H.C. Wolf. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Computer graphics and image processing (USA)*, 15(3):201–223, March 1981.

[28] M.M. Fleck. Representing space for practical reasoning. *Image and Vision Computing*, 6(2):75–86, 1988.

[29] David Forsyth and Andrew Zisserman. Mutual illumination. In *Proc. Conference on Pattern Recognition and Image Processing*, pages 466–473, 1989.

[30] David Forsyth and Andrew Zisserman. Shape from shading in the light of mutual illumination. In *Proc. 5th Alvey Vision Conference*, pages 193–198, University of Reading, September 1989.

[31] P. W. Foulkes. Early Vision Algorithms Adapted to Actively Sensed Non-Visible Modalities. Change of status thesis, University of Oxford, Department of Engineering Science, Oxford, 1987.

[32] Gallo, L.M., Bösiger, P., Rageth, Chr.J., and Stucki, D. Quantitative infrared thermography to identify varicoceles as the cause of male infertility. *Biomedizinische Technik*, 30(11):284–290, November 1985.

[33] D. Gerson, E. Khedouri, and P. Gaborski. Detecting the gulf stream from digital infrared data pattern recognition. *Belle W Baruch Library in Marine Science*, 12:19–39, 1982.

[34] R.M. Haralick. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67:786–804, 1979.

[35] C.H. Hayden, R.C. Gonzalez, and A. Ploysongsang. A temporal edge-based image segmentor. *Pattern Recognition*, 20(3):281–290, 1987.

[36] G.C. Holst. Digital median filters. *Proc. Society of Photo-Optical Instrumentation Engineers*, 890:158–163, January 1988.

[37] B.K.P. Horn. Understanding image intensities. *Artificial Intelligence*, 8(2):201–231, April 1977.

[38] B.K.P. Horn. *Robot Vision.* McGraw-Hill, NY, 1986.

[39] Max Jacob. *Heat Transfer*, volume 1. John Wiley and Sons, inc., New York, 1949.

[40] Charles C. Roberts Jnr. The application of infrared thermography in fire and explosion investigation. *Proc. Society of Photo-Optical Instrumentation Engineers*, 934:2–9, April 1988.

[41] R.B. Johnson. Relative merits of the 3-5 μm and 8-12μm spectral bands. *Proc. Society of Photo-Optical Instrumentation Engineers*, 917:102–111, 1988.

[42] R. J. Keyes, editor. *Optical and Infrared Detectors*, volume 19 of *Topics in Applied Physics.* Springer-Verlag, 1980.

[43] N. Krishnakumar, S. Sitharama Iyengar, R. Holyer, and M. Lybanon. Feature labelling in infrared oceanographic images. *Image and Vision Computing*, 8(2):142–147, May 1990.

[44] Ying Kei Henry Lau. Intelligent servo motor interface. Undergraduate Project 3YP/88/172, University of Oxford, Department of Engineering Science, Oxford, April 1988.

[45] Li Layuan. Digital contour enhancement of infrared image. *International Journal of Infrared and Millimeter Waves*, 7(1):137–145, January 1986.

[46] Li Layuan and Zou Haiming. Enhancement techniques of infrared image. *Proc. Society of Photo-Optical Instrumentation Engineers*, 572:46–49, August 1985.

[47] J. M. Lloyd. *Thermal Imaging Systems.* Optical Physics and Engineering. Plenum Press, New York, 1975.

[48] D. Long. *Photovoltaic and Photoconductive Infrared Detectors*, volume 19 of *Topics in Applied Physics*, pages 101–147. Springer-Verlag, 1980.

[49] P.A. Manning. Modelling of thermal detector arrays with image difference processing (addendum). Internal RSRE Memorandum.

[50] P.A. Manning, D.E. Burgess, and R. Watton. A linear pyroelectric array IR sensor. *Proc. Society of Photo-Optical Instrumentation Engineers*, 590:2–10, 1985.

[51] K.C. Markham. A multiple point correlation tracker for time-to-go estimation. In *Proc. 4th Int. Conference of Advanced Infrared Detectors and Systems*, pages 138–144, IEE, Savoy Place, London, June 1990. IEE conference publication no, 321.

[52] Marr, D. *Vision.* W.H. Freeman, San Fransisco, CA, 1982.

[53] Marr, D. and Hildreth, E. Theory of Edge Detection. *Proc. R Soc. London B*, 207:187–217, 1980.

[54] P. Marrucci and G. Sgarzi. Ir surveillance: improved signal processing techniques in automatic target detection. In *Proc. 2nd International Conference on Advanced Infrared Detectors and Systems*, pages 89–94, October 1983.

[55] V.I. Matveev. Radiative contrast of heated bodies in the atmospheric transmission windows. *Soviet Journal of Optical Technology*, 54(2):70–73, 1987.

[56] R.K. McEwan, A.D. MacLeod, P.A. Manning, and M.C.J. Perkins. An Uncooled Solid State Thermal Imager. In *Proc. 4th Int. Conference of Advanced Infrared Detectors and Systems*, pages 93–99, IEE, Savoy Place, London, June 1990. IEE conference publication no, 321.

[57] R. McKendall and M. Mintz. Models of sensor noise and optimal algorithms for estimation and quantization in vision systems. pages 101–116, February 1987.

[58] J.K. McWilliams and M.D. Srinath. Performance analysis of a target detection system using infrared imagery. *IEEE Transactions on Aerospace and Electronic Systems*, AES-20(1):38–49, January 1984.

[59] E.R. Meyer and R.C. Gonzalez. Spatial techniques for digital image enhancement and restoration. In *Proc. 1st South African Symposium on digital image processing*, pages 137–182, Durban, July 1983.

[60] Victor I. Myers and William A. Allen. Electrooptical remote sensing methods as nondestructive testing and measuring techniques in agriculture. *Applied Optics*, 7(9):1819–1838, 1968.

[61] M.L. Nack. Temporal registration of multispectral digital satellite images using their edge images. In *AAS/AIAA Astrodynamics Specialist Conference*, pages 26–33, Nassau, Bahamas, July 1975. Paper no. AAS 75–104.

[62] Nandhakumar, N. and Aggarwal, J.K. Integrating information from thermal and visual images for scene analysis. *Proc. Society of Photo-Optical Instrumentation Engineers*, 635:132–142, 1986.

[63] Nandhakumar, N. and Aggarwal, J.K. Integrated analysis of thermal and visual images for scene interpretation. *IEEE Trans. Pattern Analysis Machine Intell.*, 10(4):469–481, July 1988.

[64] D.G. Nichol. Autonomous extraction of an eddy-like structure from infrared images of the ocean. *IEEE Transactions of Geoscience and Remote Sensing*, GE-25(1):24–34, January 1987.

[65] L.W. Nichols. Conversion of infrared images to visible in color. *Applied Optics*, 7:1757–1762, September 1968.

[66] J. Alison Noble. *Descriptions of image surfaces.* PhD thesis, Oxford University, September 1989.

[67] J.A. Noble and J.M. Brady. Estimating Asymmetric Noise. In *Proc. 5th Alvey Vision Conference*, pages 157–162, University of Reading, September 1989.

[68] Claes Ohman. Practical methods for improving thermal measurements. *Proc. Society of Photo-Optical Instrumentation Engineers*, 313:204–212, 1981.

[69] Ovrebo, P.J., Sawyer, R.R., Ostergren, R.H., Powell, R.W., and Woodcock, E.L. Industrial Technical and Medical Applications of Infrared Techniques. *Proceedings of the IRE*, pages 1629–1645, September 1959.

[70] Pau, L.F. and Nahas, M.Y. *An Introduction to Infrared Image Acquisition and Classification Systems.* Research Studies Press Ltd., John Wiley and Sons inc., 1983.

[71] Bui Tuong Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975.

[72] S. G. Porter. A brief guide to pyroelectric detectors. *Ferroelectrics*, 33:193–206, 1981.

[73] W.K. Pratt. *Digital Image Processing.* John Wiley and Sons, New York, 1978.

[74] E. Putley. *Thermal Detectors*, volume 19 of *Topics in Applied Physics*, pages 71–100. Springer-Verlag, 1980.

[75] C.J. Radford. Infra-red image processing using a bit plane array processor. In *Proc. 2nd International Conference on Advanced Infrared Detectors and Systems*, pages 84–88, October 1983.

[76] Leonard E. Ravich. Thermal imaging: A review. *Laser Focus/Electro-Optics*, 22(2):98–110, February 1986.

[77] H.C. Schau. Statistical filter for image feature extraction. *Applied Optics*, 19(13):2182–2190, July 1980.

[78] John R. Schott. Image processing of thermal infrared images. *Photogrammetric Engineering and Remote Sensing*, 55(9):1311–1321, September 1989.

[79] R.W. Series, C.J. Radford, M.J. Varga, P. Fretwell, and A.C. Sleigh. Comparison of approaches to feature detection. In *Proc. 5th Alvey Vision Conference*, pages 145–150, University of Reading, September 1989.

[80] L. Sevigny, G. Hvedstrup-Jensen, M. Bohner, E. Østevold, S. Grinaker, and J. Dehne. Discrimination and classification of vehicles in natural scenes from thermal imagery. *Computer Vision, Graphics, and Image Processing*, 24:229–243, 1983.

[81] Ivan Simon. *Infrared Radiation.* D. Van Nostrand Company, inc, Princeton, New Jersey, 1966.

[82] Libor A. Spacek. Edge Detection and Motion Detection. Technical report, University of Essex, 1985.

[83] Allen Stimson. *Photometry and Radiometry for Engineers.* John Wiley & Sons Inc., New York, 1974.

[84] Sun Microsystems. *SunView 1 Programmer's Guide*, May 1988.

[85] Shimon Ullman. On visual detection of light sources. Technical Report AI Memo No. 333, MIT AI Lab, May 1975.

[86] Vasko, Antonín. *Infra-red Radiation.* Iliffe Books Ltd., London, 1968. Translated from Czechoslovakian.

[87] F. Vilnrotter, R. Nevatia, and K. Price. Structural analysis of natural textures. *IEEE Trans. Pattern Analysis Machine Intell.*, 8(1):76–89, January 1986.

[88] R. Watton. Infrared television: Thermal imaging with the pyroelectric vidicon. *Phys. Technol.*, 11:62–66, 1980.

[89] R. Watton. Ferroelectrics for Infrared Detection and Imaging. In *Proc. 6th Int. Symposium on Applications of Ferroelectrics*, pages 172–181, Lehigh Univ., Bethlehem, PA, USA, June 1986.

[90] R. Watton and P.A. Manning. The design of low noise arrays of MOSFETS for pyroelectric array readout, (LAMPAR). *Proc. Society of Photo-Optical Instrumentation Engineers*, 807:98–105, 1987.

[91] R. Watton, M.A. Todd, and J.P. Gillham. Materials and Technology Research in Uncooled Pyroelectric IR Detector Arrays. In *Proc. 4th Int. Conference of Advanced Infrared Detectors and Systems*, pages 69–77, IEE, Savoy Place, London, June 1990. IEE conference publication no, 321.

[92] D. B. Webb. Thermal imaging via cooled detectors. *The Radio and Electronic Engineer*, 52(1):17–30, January 1982.

[93] P.R. White, P.M. Clarkson, and I.C. Carmichael. The application of adaptive noise cancellation to pyroelectric detectors. *Infrared Physics*, 27(4):267–273, July 1987.

[94] Andrew P. Witkin. Scale space filtering. In *Proc. 8th Int. Joint Conference on Artificial Intelligence*, pages 1019–1022, Karlsruhe, West Germany, 1983.

[95] W. L. Wolfe and G.J. Zissis, editors. *The Infrared Handbook.* Environmental Research Institute of Michigan, Ann Arbor, 1978.

[96] William L. Wolfe, editor. *Handbook of Military Infrared Technology.* Naval Research Laboratory, Washington, D.C., 1965.

[97] Eric M. Wormser. Sensing the invisible world. *Applied Optics*, 7(9):1667–1671, September 1968.

[98] H. C. Wright. *Infrared Techniques.* Monographs in Electrical and Electronic Engineering. Oxford University Press, Ely House, London W1, 1973.

[99] S. Young and A.C. Prior. Sea Owl - a passive identification device for the Royal Navy Lynx helicopter. In *Proc. 4th Int. Conference of Advanced Infrared Detectors and Systems*, pages 158–165, IEE, Savoy Place, London, June 1990. IEE conference publication no, 321.