75

An Investigation into the Design of a Satellite based Stereo Imaging Sensor and the use of Automatic Image Matching in the Production of Digital Elevation Models

By Peter Olaf Moon

Submitted to the University of Cape Town in fulfilment of the requirements for the degree of Master of Science in Engineering.
Cape Town, September 1993

Declaration

I declare that this dissertation is my own work. It is being submitted for the Degree of Master of Science in Engineering at the University of Cape Town. It has not been submitted before for any degree or examination at this or any other university.

P. O. Moon
(Signature of Candidate)

Acknowledgements

I would like to thank my supervisor, Professor Gerhard de Jager, for his guidance and enthusiasm.

I would also like to thank:

- My parents, for their continued support and encouragement.
- The Foundation for Research Development (FRD), for their financial assistance.
- The University of Stellenbosch, for their financial assistance.
- The Anglo-American Corporation, for their financial assistance.
- Wayne Borchardt, Sean Borman, Greg Cox, Robert Crida, Fred Hoare, Anthony Kanfer, Jia Liu, Andre le Roux, and Brendt Wohlberg, who helped create the stimulating working environment.

Abstract

Two problems are addressed in this dissertation. They are the design of a micro-satellite based stereo imaging sensor and the automatic matching of digital stereo images for automatic cartography applications. The two problems are related; they are both components of a stereo vision system.

The research was initially motivated by the decision of the Department of Electrical and Electronic Engineering of the University of Stellenbosch to develop and build an experimental micro-satellite, SUNSAT. The proposed payload included a high resolution multi-spectral stereo imaging sensor. The second problem was motivated by the desire to use an automatic matching system to process the images produced by the sensor.

The investigation, into the sensor design, was divided into two parts. The first part investigated the feasibility of the sensor and the second part dealt with the development of a design specification.

The investigation, into automatic matching, dealt with the degree to which a set of requirements could be met. These requirements relate to the accuracy, reliability, generality, predictability and complexity of the matching system. The effect of scene characteristics was also investigated.

The results showed that it is possible to build a micro-satellite based stereo imaging sensor. The recommended sensor design included three spectral bands, an 8 bit analogue-to-digital converter and a focal length of 535 mm. Furthermore, it was found that a sub-pixel accuracy matching requirement can be met and that a matching reliability of 89.6 % can be achieved. Finally, it was found that the best matching results are obtained in areas of high image variance and low disparity variance.

Nomenclature

Symbol	Definition
$\mathbb{Z}$	The set of integers.
$\mathbb{R}$	The set of real numbers.
$*$	The convolution operator.
I_1	A digital image. It is regarded as an integer array and is obtained from a continuous image by a process of digitisation.
I_2	A digital image.
$g_1(\vec{m})$	The intensity function of image I_1 defined on the discrete Cartesian grid of integer coordinate pairs $\vec{m} = (r, c) \in \mathbb{Z}^2$.
$g_2(\vec{m})$	The intensity function of image I_2 defined on the discrete Cartesian grid of integer coordinate pairs $\vec{m} = (r, c) \in \mathbb{Z}^2$.
T_G	A geometric spatial-mapping function. It reflects the knowledge about the geometrical relation between two images.
p_G	A vector of parameters specifying T_G.
T_I	An intensity-mapping function. It reflects the knowledge about the intensity relation between two images.
p_I	A vector of parameters specifying T_I.
W	A subset of $\mathbb{Z}^2$ defining the coherent collection of pixels used for matching.
$f_1(\vec{m})$	An intensity function defined over W.
$f_2(\vec{m})$	An intensity function defined over W.
$\Phi(f_1, f_2, W)$	Normalised cross-correlation over W.
$\Phi_0(f_1, f_2, W)$	Zero mean normalised cross-correlation over W.
$\rho(\vec{k})$	A correlation function defined on a discrete grid.
$\psi(x)$	A prototype or mother wavelet.
Q_e	Radiant energy (J), the energy associated with electromagnetic radiation.
Φ_e	Radiant flux (W), the rate of radiant energy transfer.
E_e	Radiant incidance or irradiance (Wm^{-2}), the radiant flux incident upon a real or imaginary surface of unit area.
M_e	Radiant exitance or emittance (Wm^{-2}), the radiant flux which is emitted by a surface of unit area.
L_e	radiant sterance or radiance ($\mathrm{Wm}^{-2}\mathrm{sr}^{-1}$), the radiant flux density transmitted from a small area and viewed through a solid angle.

Symbol	Definition
Q_λ	Spectral radiant energy ($J\mu m^{-1}$).
Φ_λ	Spectral radiant flux ($W\mu m^{-1}$).
E_λ	Spectral irradiance ($Wm^{-2}\mu m^{-1}$).
M_λ	Spectral exitance ($Wm^{-2}\mu m^{-1}$).
L_λ	Spectral radiance ($Wm^{-2}sr^{-1}\mu m^{-1}$).
ρ	Reflectance.
τ	Transmittance.
α	Absorptance.
$\rho(\lambda)$	Spectral reflectance.
$\tau(\lambda)$	Spectral transmittance.
$\alpha(\lambda)$	Spectral absorptance.
f	Focal length of a lens.
N	F/number of a lens.
D	Entrace pupil diameter.
λ	Wavelength of electromagnetic radiation.
T	Absolute temperature.
BRDF	Bidirectional refectance distribution function.
$\tau_o(\lambda)$	Spectral transmittance of optical system.
$\tau_u(\lambda)$	Spectral transmiitance of the atmosphere along the atmospheric path between the sensor and the ground object.
IFOV	Instantaneous field of view, the area on the ground covered by the image of a detector element as magnified by the ratio sensor altitude to the focal length of the optics of the sensor system.
MTF(ν)	Modulation transfer function, it describes the spatial resolution of the imaging sensor in terms of its spatial frequency response.
$MTF_S(\nu)$	MTF of imaging sensor.
$MTF_L(\nu)$	MTF of the lens (optical system).
$MTF_I(\nu)$	MTF due to image smear.
$MTF_D(\nu)$	MTF of detector.
ν	Spatial frequency.
EIFOV	Effective instantaneous field of view, the spatial half wavelength for which the modulation of the distribution of radiance had been reduced by one-half for a sinusoidal distribution of radiance.
h	Planks constant (6.63×10^{-34} J s).
c	Speed of light (3×10^8 m s^{-1}).
CCD	Charge coupled device.
η	Quantum efficiency of a CCD.
$r_\oplus$	Mean radius of Earth (6370 km).
h_s	Orbital altitude of satellite.
$\dot{\theta}_s$	Angular velocity of satellite.
θ_s	Angular position of satellite.
t	Time (s).
T_s	Orbital period of satellite.
θ_o	Offset of sensor from vertical.
α_g	Grazing angle.

Symbol	Definition
l_s	Slant range, distance between the sensor and ground object.
B/H	Base-to-height ratio.
RMSME	Root mean square matching error.
$\vec{D}$	Reference or exact disparity.
$\vec{d}$	Estimated disparity.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background and Motivation for Research

Remote sensing of the Earth's resources, from space-based sensors, has evolved from a scientific experiment to a commonly used technological tool [30]. The Landsat and SPOT satellites have provided a wealth of information about the Earth at a competitive cost [14]. The term "remote" sensing in this context refers to imaging sensors which produce high resolution[1] multi-spectral digital images of the Earth, using the reflective portion of the optical spectrum (0.4 μm to 2.4 μm). Applications include: land-usage classification, crop analysis, automatic cartography, hydrological studies and geological studies [14,1]. Both the Landsat and SPOT satellites provide high resolution multi-spectral digital images, while SPOT also offers stereo image pairs [41,57,7] (see Chapter 4).

The Department of Electrical and Electronic Engineering of the University of Stellenbosch decided to develop and build an experimental 50 kg, 45 cm cubic micro-satellite, SUNSAT. The proposed payload includes a high resolution multi-spectral stereo imaging sensor and an Amateur Radio packet communications system. The design also includes the required support systems. An attitude control system to provide the required stability, a thermal control system to maintain the internal temperature of the satellite, a computer system to provide storage and data processing, and a communications system to provide a channel to transport the image data back to Earth [44].

The use of micro-satellite based imaging sensor results in a significant cost reduction compared to satellite missions such as SPOT and Landsat (see Chapter 4). This includes the manufacture and launching costs. The total cost developing and building SUNSAT will be approximately R 2.1 million and the launching costs could amount to R 1.4 million[2] [45].

An application of the stereo image pairs obtained from the proposed imaging sensor would

[1] In this dissertation the term *high resolution* refers to sensors that have a spatial resolution (see Section 3.3.1) of less than 100 m.

[2] These estimates are based on an exchange rate of 3 SA rands per US dollar.

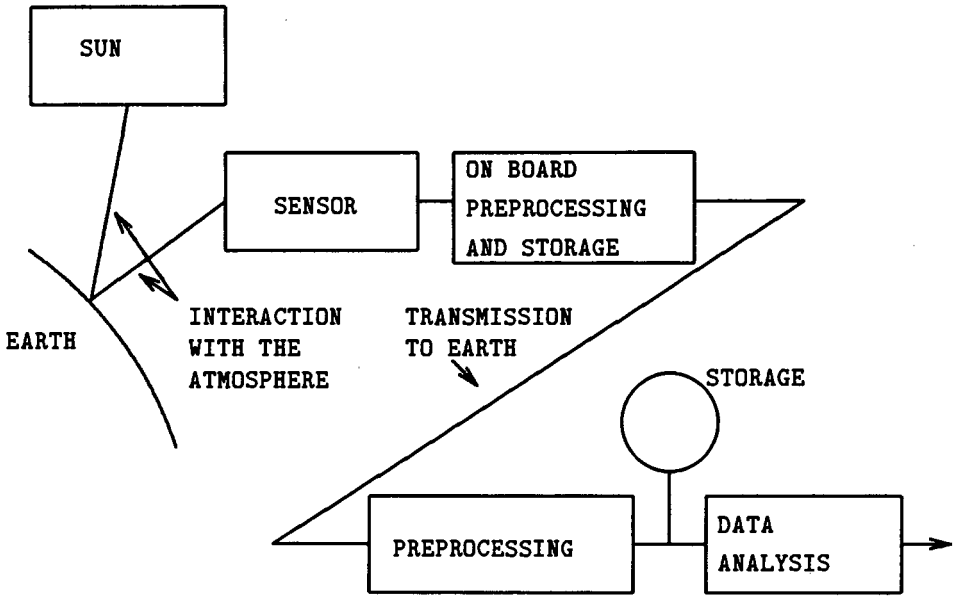


Figure 1.1: A block diagram of a remote sensing system.

be automatic cartography (the production of digital elevation models). This application requires that the stereo image pairs be "matched" (see Section 2.3.2). A trained operator can do this. It is, however, desirable to use an automatic digital stereo image matching system [54]. The applications of automatic digital stereo image matching also include autonomous vehicle navigation, industrial automation and modelling of human visual system [4]. However, the computational problems involved in developing such a system have not yet been completely solved. The large number of proposed techniques and their limited domains demonstrates this (see [15,4]).

In summary, the research that led to this dissertation was motivated by: (a) the proven usefulness of remote sensing data, (b) the opportunity to fly the imaging sensor, if it proved feasible, (c) the cost advantages of micro-satellites, and (d) the requirement for an automatic digital stereo image matching technique.

1.2 Definition of the Problem

The problems addressed in this dissertation are defined in this section. System requirements, study objectives, study scope and study limitations are all discussed.

Figure 1.1 depicts a block diagram of a remote sensing system. It shows the path that the electromagnetic radiation takes from the Sun to the sensor and the conversion of radiation detected by the sensor into a useful product. During the passage from the Sun to the Earth

and from the Earth to the sensor the radiation interacts with the atmosphere. The sensor detects, measures and converts a fraction of the radiation reflected by the surface of the Earth into digital data. Preprocessing on board, the satellite prepares the data for transmission back to Earth. The result of further processing on Earth, including preprocessing, storage and data analysis, is a useful product.

The first problem to be addressed is that of an investigation into the feasibility and design of a sensor that meets the following requirements:

- It must produce multi-spectral images.
- It must produce a stereo image pair in a single pass.
- It must produce high resolution images.
- It must be possible to fly it on a micro-satellite like, SUNSAT.

The concept design [44], motivated by the design and method of operation of the Landsat and SPOT satellites, limited the scope of the investigation. The spectral range of interest was limited to the visible to near infrared region of the electromagnetic spectrum (0.4 μm to 1.0 μm). The light sensing devices (detectors) were assumed to be line-array charge coupled devices (CCDs) (see Section 3.5). Furthermore, the sensor will operate as a so called "pushb-room" scanner. It will use the satellite motion to build up a two-dimensional image. Further discussion of the concept design and its motivation can be found in Section 5.2.

The scope of the investigation was also limited to a theoretical investigation. Practical implementation of the design is beyond the scope of the study.

The second problem addressed in this dissertation is that of an investigation into automatic digital stereo image matching for automatic cartography applications. If an automatic digital stereo image matching system is to find application in automatic cartography it would have to meet the following requirements:

- Accuracy. It must match to sub-pixel accuracy.
- Reliability. It must not make gross matching errors.
- Generality. It must be applicable to as wide a range of images as possible.
- Predictability. It must have a predictable performance.
- Complexity. It must have a minimal computational complexity.

The objectives of the investigation are: (a) to identify the degree to which the different requirements can be satisfied, (b) to investigate techniques that improve the reliability, and (c) to identify the effect of scene characteristics on performance.

The scope of the investigation was limited: (a) to the application of automatic cartography, and (b) to techniques that do not make direct use of the camera model (see Section 2.2.2). Experimental work concentrated on techniques that improve the reliability of the results obtained. Investigation of the computational complexity of the techniques was limited.

The primary limitation was the limited amount of data available for experimentation. Only digitised aerial photographs could be obtained. This limited the investigation into the degree to which the requirements of generality and predictability were satisfied.

In summary, the objectives of the study are: (a) to investigate the feasibility of the imaging sensor, (b) to develop a specification for the imaging sensor, and (c) to investigate the use of automatic digital stereo image matching. The objectives of the dissertation are: (a) to present the theoretical background to the problems, (b) to present the theoretical approach used in the investigation, (c) to present the results of the investigation, and (d) to present the conclusions and recommendations.

1.3 Plan of Development

This section presents the plan of development for this dissertation. It is designed to meet the objectives presented in Section 1.2 and is divided into three parts.

First, the relevant theory and other results from the literature are presented. Chapter 2 deals with the theoretical background to computational stereo vision and automatic digital stereo image matching in particular. Topics discussed include: the computational stereo vision paradigm of Barnard and Fischler [4], area and feature matching, the digital image matching problem, correlation based measures of match and mismatch, the search for corresponding points, estimating the position of the corresponding points to sub-pixel accuracy, and the wavelet transform and multiresolution analysis.

Chapter 3 presents the theoretical background to optical remote sensing systems. Topics discussed include: radiometric theory, resolution, the quantum nature of light, charge coupled devices, quantisation noise, orbital mechanics and satellite sensor geometry.

Remote sensing satellite missions, in particular the Landsat and SPOT series, are reviewed in Chapter 4. The relevant design specifications and principles of operation are also discussed.

Second, the theoretical approaches to the problems are discussed. Chapter 5 deals with the development of a model of the imaging sensor. The model relates parameters of the sensor, the satellite and the environment to various output characteristics of the sensor. Chapter 6 deals with the determination of the feasibility of the sensor and with the design of the sensor.

The investigation into automatic digital stereo image matching is described in Chapters 7 and 8. Chapter 7 deals primarily with the development of a hierarchical area correlation matching algorithm. Other tropics discussed include: the preparation of the experimental data, techniques for evaluating the results, and the applicability of the wavelet transform.

Chapter 8 deals with techniques for improving the reliability and with the effect of scene characteristics. Back-matching, the removal of inconsistencies and the detection of occlusion points are described.

Finally, the conclusions and recommendations are presented in Chapter 9.

Chapter 2

Computational Stereo Vision and Image Matching Theory

2.1 Introduction

Stereo vision is a technique for recovering the information about the three-dimensional structure of a scene, which is not contained in any single image. Computational stereo vision deals with those aspects that make stereo vision computationally possible. The theory of computational stereo vision deals not ónly with artificial stereo vision systems that are presently implemented using digital computer technology, but also with the human visual system.

This chapter presents the theoretical background of computational stereo vision required for this dissertation, including relevant results from the literature. The sections of the theory that are of the most interest to this dissertation are: (a) how the design of the imaging system affects the accuracy of any measurements, (b) how the design of the imaging system affects the performance of automatic digital stereo image matching systems, and (c) automatic digital stereo image matching techniques.

The human visual system has played an important role in the development of computational stereo vision theory (see [40]). The reason for this importance is that it shows that stereo vision is both physically and computationally possible. Proof that the human visual system uses the stereo vision technique to recover three-dimensional information, lies in experiments that have shown that the human visual system can interpret random dot stereograms, which contain no monocular or high-level depth clues [40] (see Section 2.2.1).

Conceptually stereo vision is a relatively simple application of triangulation. Figure 2.1 illustrates the concept in the binocular case.[1] Consider a scene point S, two cameras represented by their optical centres C_1 and C_2, and image planes P_1 and P_2.

[1]Trinocular stereo systems have also been proposed and developed (see [3, 63, 9]). The advantage of using trinocular systems is that they introduce redundancy. However, most systems are binocular and the rest of the chapter will assume the binocular case, unless otherwise stated.

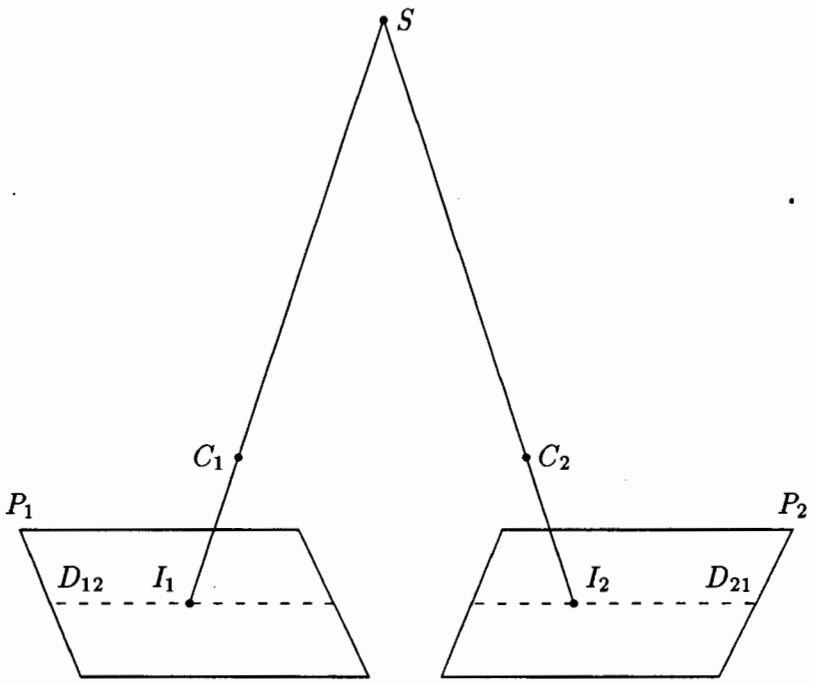


Figure 2.1: The concept of triangulation applied to stereo vision. The images I_1 and I_2 of the scene point S is the projection of S onto the image planes P_1 and P_2 through optical centres C_1 and C_2 respectively. Lines I_1C_1 and I_2C_2 intersect at the scene point S. A pair of conjugate epipolar lines is depicted by D_{12} and D_{21}.

The stereo vision concept is characterised by the following three steps:

1. A point in three-dimensional space is projected by cameras onto two images. In Figure 2.1, the image I_i of the scene point S by camera i is given by the intersection of line SC_i with plane P_i, $i = 1,2$.

2. The projections of the scene point in the two images are located. In Figure 2.1, they are points I_1 and I_2 in planes P_1 and P_2, respectively.

3. The positions of the image points and parameters of the imaging geometry are used to estimate the three-dimensional coordinates of the scene points. In Figure 2.1, the intersection of lines I_1C_1 and I_2C_2 give the three-dimensional coordinates of S.

Section 2.2 will present the computational stereo vision paradigm of Barnard and Fischler [4]. It will highlight the issues of stereo vision system design that affects the computational aspects of the problem. It will also identify the major computational problem of locating the corresponding points, stereo image matching (step 2 above). The theory dealing with automatic digital stereo image matching will be presented in Sections 2.3 and 2.4. Section

2.3 will present the digital image matching theory relevant to this dissertation. Section 2.4 will review the theory of the wavelet transform and multiresolution analysis, in particular the construction of a sequence of images at dyadic resolutions.

2.2 Computational Stereo Vision Paradigm

The computational stereo vision paradigm introduced by Barnard and Fischler [4] will be described in this section. It was a consequence of the considerable amount of research directed towards solving the computational stereo vision problem. The objective of this section is to give the reader a basic overview of the steps involved in computational stereo vision. The paradigm involves the following steps:

- Image acquisition,
- Camera modelling,
- Feature acquisition,
- Image matching,
- Depth/height determination, and
- Interpolation.

A brief description of each step will be given in the sections that follow. However, the opportunity will also be taken to highlight those aspects of the design of a stereo vision system that affect the results that can be obtained.

2.2.1 Image Acquisition

The method used to acquire the images is largely determined by the application, the nature of the scene and other physical constraints. Two applications and their associated methods will be discussed in this section. The first application is that of the computational modelling of the human stereo vision system, where synthetic random dot stereograms are used. The second application is that of automated cartography, where aerial photographs and satellites images are used.

Important aspects of image acquisition, which affect the results that can be obtained, include:

- The timing, whether the images are recorded simultaneously, nearly simultaneously, or at different times,
- The camera orientations, whether they are slightly or radically different,

- The photometry and geometry of the camera,
- The image resolution,
- The illumination of the scene,
- The atmospheric response, and
- The scene characteristics, including the geometrical and reflectance properties.

Synthetic random dot stereograms have been employed for research into the computational modelling of the human stereo vision system. A random dot stereogram consists of two synthetic images of the same virtual surface taken from different view points. Each image consists of uncorrelated dots and contains no monocular or high-level depth clues. When the two images are viewed stereoscopically, however, the three-dimensional virtual surface is readily perceived. They have the following advantages over natural images:

- They contain no monocular depth clues,
- The exact correspondence is known, and
- Parameters such as noise and density can be controlled.

The images used for automated cartography are generally obtained in one of two ways. The first method is aerial photography where stereo images are taken from an aircraft using a frame camera. The second method is to use a line scanning imaging system carried by a satellite. The geometry of the two methods is different, due to the motion of the satellite during image acquisition. In both methods, the time between the recording of the images is kept to a minimum.

The content of the images, used for automated cartography significantly, affects the performance of the stereo vision system. In this context the presence of cultural features such as buildings and roads is important. These cultural features present special problems that natural objects and surfaces, such as mountains, flat or "rolling" terrain, foliage and water, do not. Periodic structures such as windows of buildings and road grinds also confuse. The relative abundance of occlusion edges in a city scene can also cause problems because large portions of the images may be unmatched.

The image resolution is another important factor affecting system performance. In low resolution images the cultural features have less of an effect. The resolution also places a limit on the accuracy of any measurements. This will be discussed in Section 5.7.

2.2.2 Camera Modelling

Camera modelling is the step of estimating the camera attributes. These attributes are required to map the corresponding points[2] into three-dimensional scene locations. A camera model is a representation of the geometrical and physical attributes of the camera. The geometrical attributes include the position and attitude of the camera.

The camera model can also be used to constrain the search for corresponding points, using the principle of *epipolar lines*. This is illustrated in Figure 2.1. A scene point S and the optical centres of two camera C_i and C_j define an *epipolar plane* Q_{ij}. The intersection of this epipolar plane with the image plane P_i is the epipolar line D_{ij}, while its intersection with image plane P_j is the epipolar line D_{ji}. D_{ij} and D_{ji} are called conjugate epipolar lines (see Figure 2.1). The corresponding point I_j, of the image point I_i on the line D_{ij}, will be on the line D_{ji} [3,4,42]. Therefore, the search for corresponding points can be limited to conjugate epipolar lines, a one dimensional search. The applicability of the epipolar concept in practice depends on the accuracy of the camera model, if it is known.

The parameters of the camera model are determined using the concept of orientation. *Orientation* refers to the determination of the position and attitude of a camera, image or model in a space relative to a system of coordinate reference [19]. It is divided into interior and exterior orientation. Exterior orientation is further divided into relative and absolute orientation.

Interior orientation refers to the determination of the perspective geometry of the camera. These parameters are generally known *a priori* and can be determined by precise calibration. They are known as the elements of interior orientation, which usually include: (a) the focal length, (b) the position of the principle point in the image plane of the camera, and (c) the geometric distortion characteristics of the lens system.

Exterior orientation refers to the determination of two sets of parameters: (a) the location of the perspective centre, and (b) the angular orientation of the camera axis or the image-plane.

Relative orientation refers to the determination of the relative positions of the two images in a stereo pair with respect to each other. It alone can be used to determine conjugate epipolar lines. Absolute orientation refers to the establishment of the model coordinate system with respect to the ground coordinates.

Complete solution of the orientation problem requires that ground control points be located. Determination of the camera model is beyond the scope of this dissertation and the interested reader is referred to texts on photogrammetry (for example [19]).

[2]If the points P_1 and P_2 in the images I_1 and I_2, respectively, are the projections of the same scene point, then they are called *corresponding points*.

2.2.3 Feature Acquisition

The need for feature acquisition stems from the fact that different regions of the image can be matched with different degrees of confidence. For example, areas of nearly homogeneous brightness cannot be matched with confidence. The objective of feature acquisition is to identify "features" that can be matched with the greatest degree of confidence. The method and feature type is closely coupled to the matching strategy (see Section 2.2.4).

Methods and feature types that have been applied include:

- "Interest operators". They locate points in one image that can be matched with confidence to corresponding points in the second image of a stereo pair. These operators identify regions of high image-intensity variance. It is also desirable that the variance is not due to brightness changes in the direction perpendicular to the epipolar line. A review of the different operators can be found in [4].
- Point-like features are often used when the camera model is unknown. They have the advantage that they can be matched in any direction without concern for perspective distortion.
- Edge elements are used since they often correspond to physical object boundaries. Various models have been proposed. Some only used the binary edge/no-edge, while others use edge "magnitude" and "direction" (see [40,24]).

2.2.4 Image Matching

Image matching is the step of identifying the corresponding points. It is the step that presents that greatest challenge in solving the computational stereo vision problem. An overview of the image matching problem and the two basic approaches that have been applied will be presented in this section. A detailed discussion of digital image matching theory will be presented in Section 2.3.

In the case of stereo image matching, a match is sought between images, and not between an image and an abstract model (template matching). Furthermore, a dense grid of matching points is required.

The important differences between the images result from a change in view point, and not, from changes in the scene. The degree of change is dependent on the stereo angle (the angle between the rays joining the scene point to the corresponding image points). Therefore, the most significant changes occur in the appearance of nearby objects.

Changes can also be introduced in both the geometry and photometry of the film development and/or scanning steps. If the images are not taken simultaneously, changes in the scene and lighting conditions can also introduce significant effects.

Marr and Poggio [40] identified two important physical constraints on the matching process: (1) a given point on a physical surface has a unique position in space at any given time,

and (2) matter is cohesive, it is separated into objects, and the surfaces of objects are generally smooth compared with their distance from the viewer. Marr and Poggio [40] translated these constraints into two rules:

1. *Uniqueness.* Each item from each image may be assigned at most one disparity value.[3]
2. *Continuity.* Disparity values change smoothly, except at a few surface discontinuities.

The information content in the intensity value of a single pixel it too low for unambiguous matching [4]. Therefore, coherent collections of pixels are matched. The collections are determined and matched in two distinct ways:

- Area Matching, and
- Feature Matching.

A comparison of the two basic methods is presented in Table 2.1. The advantages and disadvantages of each method show that area and feature matching are to a degree complementary. This had led to experiments that have shown that combining the two techniques leads to better results [26,29].

The sensitivity of area matching to absolute intensity, contrast and illumination can be reduced by:

- preprocessing the images to have the same mean and variance, and/or
- by using correlation measures that are insensitive to absolute intensity (see Section 2.3.5).

It is possible to match to sub-pixel accuracy, which is important for automatic cartography, using either method. This is achieved with feature matching by locating the position of the feature to sub-pixel accuracy. Area matching techniques achieve it by interpolating the correlation function (see Section 2.3.7).

2.2.5 Depth/Height Determination

Once an accurate match has been found, it is a relatively simple matter of triangulation to determine the distance to a scene point. In the case of remote sensing and cartography the height above a reference plane is often required. However, close examination of this step, reveals various steps that can be taken to improve the accuracy of the depth/height estimate.

As a first approximation, the error in a depth measurement is directly proportional to the matching error and inversely proportional to the stereo baseline. The stereo baseline is the

[3]The disparity is the amount by which the projection of a scene point in an image moves due to the change in the view point. If P_1 and P_2 are corresponding points, and have coordinates (r_1, c_1) and (r_2, c_2) in I_1 and I_2, respectively, then the disparity is given by $(r_1 - r_2, c_1 - c_2)$.

Property	Area Matching	Feature Matching
Matching Units	Regularly sized neighbourhoods of a pixel.	"Semantic features" (with known physical properties and/or spatial geometry. "Intensity anomaly features" (isolated anomalous intensity patterns not necessarily having any physical significance).
Measure of Match	Correlation.	Symbolic classification. Correlation.
Advantages	Robust, with a wide range of applicability over image types [26]. Produces a dense disparity map.	Good delineation of disparity discontinuities [29]. Less sensitive towards changes in contrast and ambient lighting [15]. Faster, because fewer points (features) are considered and the matching of symbolic features is simpler [42,15].
Disadvantages	Sensitive to distortions as a result of changes in viewing position (perspective) and changes in absolute intensity, contrast, and illumination [15,42]. Confused by the presence of occluding boundaries, giving erroneous depth estimates [15]. Require the presence of a detectable texture within each correlation window [42]. Confused by the presence of surface discontinuity in the correlation window [42]. Confused by rapidly changing depth fields [42].	Generate noisy and sparse disparity maps that have to be interpolated [26,29]. Extraction of edges results in loss of information [26]. Confounded by densely textured areas or moderate levels of noise [26].

Table 2.1: A comparison of area and feature matching.

distance between the optical centres of the cameras. Therefore, the accuracy of the depth measurements can be improved by decreasing the error in the matches and/or by increasing the stereo baseline.

It is possible to match to sub-pixel accuracy, using either area or feature matching. The accuracy is limited by the signal-to-noise ratio, depth changes and geometric distortions.

The second method of increasing the stereo baseline is not without its problems. The stereo baseline cannot just be increased without bound. Increasing the stereo baseline complicates matching and decreases the matching accuracy in two ways:

1. It increases the range of disparity values; and
2. The difference in appearance of the images increases.

The designer of a stereo system has to take account of the nature of these errors and how they are related, to minimise the error of the final depth/height estimate.

2.2.6 Interpolation

Stereo applications usually demand a depth/height map that is more dense than that provide by the depth/height determination step. The map is usually more sparse if feature matching is used. Consequently, an interpolation step is usually required.

Techniques that have been applied, include:

- The fitting a continuous function to the sparse map; and
- The fitting a geometric model to the sparse map.

Further discussion of interpolation techniques is beyond the scope of this dissertation.

2.3 Digital Image Matching Theory

The overview of stereo image matching, presented in Section 2.2.4, will be extended in this section to form the theoretical basis for the investigation into automatic digital image matching (Chapters 7 and 8). The scope of this section is limited to the required theory. A review of field can be found in Barnard and Fischler [4], and Dhond and Aggarwal [15].

An explanation of the term "digital" in digital image matching will be given in Section 2.3.1. The formulation of the digital image matching problem will be presented in Section 2.3.2. The classification of the matching errors made by digital image matching systems is described in Section 2.3.3. The component parts of digital image matching techniques are described in Section 2.3.4. Correlation based ways of measuring the degree of match or mismatch between two coherent collections of pixels are described in Section 2.3.5. Search strategies for finding the corresponding points will be described in Section 2.3.6. Methods

of interpolating the correlation surface to estimate the position of corresponding points to sub-pixel accuracy are described in Section 2.3.7.

2.3.1 Digital Image Terminology

The use of the word *digital* in the term digital stereo image matching refers to the use of digital images. An image can be represented by a function of two spatial variables $f(x, y)$. The value of this function at a point is called the *gray level*, intensity or brightness of the image at that point.[4] A digital image is obtained from the continuous image by a process of digitisation. This involves two steps [52]: (1) A sampling process is used to extract from the image a discrete set of real numbers ("samples"). (2) A quantisation process is applied to these samples to yield numbers having a discrete set of possible values (see Section 3.6 for a discussion of quantisation noise). A digital image so obtained can be regarded as an integer array. The elements of a digital image or picture are called picture elements, pixels, pels or "points".

2.3.2 Formulation of the Digital Image Matching Problem

The formulation of the digital image matching problem and the basic assumptions made will be described in this section. A model of image matching, which is based on the work of Haralick and Shapiro [27], will also be developed.

Consider two images I_1 and I_2. Let the points P_1 and P_2 of images I_1 and I_2 have coordinates $\vec{m}_1 = (r_1, c_1)$ and $\vec{m}_2 = (r_2, c_2)$, and intensities $g_1(\vec{m}_1)$ and $g_2(\vec{m}_2)$, respectively. If $\vec{m}_1$ and $\vec{m}_2$ are corresponding points, their coordinates can be related by

$$\vec{m}_1 = T_G(\vec{m}_2; p_G) \tag{2.1}$$

where T_G is a specified geometric spatial-mapping function reflecting the knowledge about the geometrical relation between the images, and p_G is a vector of unknown parameters. The intensities can be related by

$$g_1(\vec{m}_1) = T_I(g_2(\vec{m}_2); p_I) \tag{2.2}$$

where the intensity-mapping function T_I contains knowledge about the intensity relation between the images, and p_I is a vector of unknown parameters.

Equations 2.1 and 2.2 are combined to obtain the model of image matching:

$$g_1(T_G(\vec{m}_2; p_G)) = T_I(g_2(\vec{m}_2); p_I). \tag{2.3}$$

The mapping functions T_G and T_I may be deterministic, stochastic, and/or piecewise continuous.

[4] A multi-spectral image can be represented by a set of functions, one for each spectral band.

The model can be further extended by including a noise component. This can be done in two ways. Either one or none of the images is assumed to be perfectly known. In the case where image I_2 is assumed to be perfectly known, the model of image matching becomes:

$$g_1(T_G(\vec{m}_2; p_G)) = T_I(g_2(\vec{m}_2); p_I) + n_1(T_G(\vec{m}_2; p_G)) \tag{2.4}$$

where n_1 is the observational noise component of image I_1. It usually assumed to be zero mean Gaussian noise. In the case where neither image is assumed to be perfectly known, the model of image matching becomes:

$$g_1(T_G(\vec{m}_2; p_G)) = T_I(g_2(\vec{m}_2) - n_2(\vec{m}_2); p_I) + n_1(T_G(\vec{m}_2; p_G)) \tag{2.5}$$

where n_1 is the observational noise component of image I_1 and n_2 is the observational noise component of image I_2.

The problem of image matching now has two solutions:

1. All the corresponding points, and
2. The parameters p_G and p_I of the mapping functions T_G and T_I.

Most approaches attempt to obtain solution (1). From these corresponding points the parameters p_G and p_I of the mapping functions T_G and T_I can be estimated, if required. In practice, however, it is not possible to find all corresponding points. It is thus assumed that the limited number of corresponding points obtained is sufficient to reconstruct the three-dimensional structure of the scene (see Section 2.2.6).

Image matching techniques make assumptions regarding the form of the mapping functions T_G and T_I. If T_G and T_I were unconstrained then any two points could be made to match. The success of a technique is often dependent on the degree to which the assumptions made are valid.

In the case of digital images, the intensity functions $g_1(\vec{m})$ and $g_2(\vec{m})$ are only defined on the discrete Cartesian grid of integer coordinate pairs $\vec{m} = (r, c) \in \mathbb{Z}^2$. Thus determination of the corresponding points to sub-pixel accuracy will require the interpolation of g_1 and g_2, or of a function derived from them. Section 2.3.7 will describe techniques for interpolating the correlation surface.

2.3.3 Classification of Matching Errors

The classification of matching errors in the case of no occlusions is discussed in this section. Based on the manner in which the matching errors are made, they can be classified into two classes:

1. Errors due to non-corresponding objects being matched, and

2. Errors in the estimation of the position of the corresponding objects.

The errors due to non-corresponding objects being matched usually result in large errors in the depth or height estimates (see Section 2.2.5). Therefore, the term "gross" error is used to describe this type of error. These errors have to be eliminated or reduced to an acceptable level, if the matching technique is to be used in an automatic system.

Feature matching techniques reduce the number of "gross" errors by only trying to find corresponding points for points that can be matched with a high degree of confidence (see Section 2.2.3). This is in contrast to area matching techniques. They generally attempt to match all the pixels.

The errors in the estimation of the position of the corresponding objects are limited by the characteristics of the scene. This includes the signal-to-noise ratio, the image resolution, and the disparity pattern (see [49]). It is possible to achieve position estimates accurate to less than one pixel (sub-pixel accuracy) (see Section 2.3.7). These errors finally limited the accuracy of three-dimensional scene reconstruction (depth or height estimates) (see Section 2.2.5).

2.3.4 Component Parts of Digital Image Matching Techniques

Digital image matching techniques can be divided into various component parts. A measure of match or mismatch is used to determine the degree to which two points are similar. A basic assumption that is made, is that the corresponding points are similar. A search method is used to find the potential corresponding points, that is, points where the measure of match/mismatch is a maximum/minimum. This procedure is often subject to other control strategies aimed at producing a result that is globally consistent and with a minimal number of "gross" errors. The three component parts are closely related.

Measure of Match or Mismatch

A measure of match is used to measure to degree to which two points match. To measure this degree of match, assumptions have to be made about T_I and T_G. The two basic methods for measuring the degree of match or mismatch are correlation and symbolic classification (see Table 2.1).

Correlation measures are used by area matching and some feature matching techniques. They use regularly sized neighbourhoods of the pixel and will be further described in Section 2.3.5.

Symbolic classification is used by some feature matching techniques. It uses symbolic features, such as edges, corners and intensity anomalies [4,15].

Search Method

The search method is used to find the optimal match based on the measure of match. The search strategy is design to minimise the computation require to find the optimal solution subject to the control strategies. In certain cases the search strategy will make indirect assumptions about the nature of T_G.

The search methods used in conjugation with correlation based measures of match or mismatch will be described in Section 2.3.6.

Control Strategies

Control strategies are used to apply additional constraints. Techniques that have been applied include:

- The rules of Marr and Poggio [40] (see Section 2.2.4) to produce a result that is physically consistent.
- Back-matching [26]. It validates each match.
- Multiresolution techniques [4,15,40] to reduce the search space and the number of false targets.

2.3.5 Correlation based Measures of Match or Mismatch

It is stated in Section 2.2.4 that the information content in the intensity value of a single pixel is too low for unambiguous matching and, hence, the requirement to use coherent collections of pixels for matching. This section describes various correlation based measures of match or mismatch between two such collections of pixels.

It is assumed that the degree of match/mismatch between two coherent collections of pixels surrounding corresponding points will be higher/lower than the degree of match/mismatch between two coherent collections of pixels surrounding non-corresponding points. This assumption is central to all techniques that use correlation based measures of match or mismatch. Techniques for finding these corresponding points, which entail finding the maximum/minimum degree of match/mismatch, will be described in Section 2.3.6. Techniques for estimating the positions these corresponding point of to sub-pixel accuracy will be described in Section 2.3.7.

To be able to measure the degree of match or mismatch, assumptions have to be made regarding the form of the transforms T_I and possibly the vector of unknown parameters p_I. A plausible assumption that is made is that $T_G(x; p_G) = x$. A consequence of this assumption is the sensitivity to changes in absolute intensity, contrast and illumination stated in Table 2.1. Techniques for reducing this sensitivity are stated in Section 2.2.4.

An assumption that is not always explicitly stated is that the disparity is assumed to be constant over the coherent collection of pixels. A consequence of this assumption is the sensitivity to rapidly changing depths and surface discontinuities stated in Table 2.1.

Consider two digital images I_1 and I_2 with intensity functions $g_1(\vec{m})$ and $g_2(\vec{m})$, respectively. Since I_1 and I_2 are digital images, g_1 and g_2 are defined on the discrete Cartesian grid of integer coordinate pairs $\vec{m} = (r, c) \in \mathbb{Z}^2$. Let W, a subset of $\mathbb{Z}^2$, define the coherent collection of pixels used for matching. This collection of pixels is called the *correlation window*. A rectangular collection is most often used.

Consider two intensity functions $f_1(\vec{m})$ and $f_2(\vec{m})$, sub-images of I_1 and I_2 respectively, defined over W. Under the assumptions stated above, two *metrics* have been used as a measure of the mismatch between f_1 and f_2 [52,39,49]. They are the l_1 error norm

$$\|f_1 - f_2\|_1 = \sum_{\vec{m} \in W} |f_1(\vec{m}) - f_2(\vec{m})| \tag{2.6}$$

and the l_2 error norm, also called the *sum of squared differences* (SSD), (assuming f_1 and f_2 are real)

$$\|f_1 - f_2\|_2 = \sqrt{\sum_{\vec{m} \in W} (f_1(\vec{m}) - f_2(\vec{m}))^2}. \tag{2.7}$$

From the errors norms above, which are used as a measure of mismatch, a measure of match can be derived. In the case of the l_1 norm Maragos [39] derived a *morphological cross-correlation* metric to measure the degree of match.

From the l_2 error norm a measure of match can also be derived. The l_2 error norm squared can be written as:

$$\sum_{\vec{m} \in W} (f_1(\vec{m}) - f_2(\vec{m}))^2 = \sum_{\vec{m} \in W} f_1(\vec{m})^2 + \sum_{\vec{m} \in W} f_2(\vec{m})^2 - 2 \sum_{\vec{m} \in W} f_1(\vec{m}) f_2(\vec{m}). \tag{2.8}$$

Therefore, a possible measure of match is $\sum_W f_1 f_2$, which is large if f_1 and f_2 are similar. A normalisation factor is introduced and the measure of match becomes:

$$\Phi(f_1, f_2, W) = \frac{\sum_{\vec{m} \in W} f_1(\vec{m}) f_2(\vec{m})}{\sqrt{\sum_{\vec{m} \in W} f_1(\vec{m})^2 \sum_{\vec{m} \in W} f_2(\vec{m})^2}}, \tag{2.9}$$

which is normalised cross-correlation.

The normalisation factor bounds the measure of match Φ to be in the range from -1 to 1. A perfect match is indicated by $\Phi = 1$, i.e. $\Phi(f, f, W) = 1$, and a perfect mismatch by $\Phi = -1$, i.e. $\Phi(f, -f, W) = -1$. The normalised cross-correlation measure also has the property that

$$\Phi(f_1, f_2, W) = \Phi(f_1, a f_2, W), \tag{2.10}$$

where $a \in \mathbb{R}^+$.

A common modification is to subtract the local mean of each function from itself. If $\overline{f}_1$ is the mean of f_1 over W and $\overline{f}_2$ is the mean of f_2 over W, then the zero mean normalised cross-correlation is given by

$$\Phi_0(f_1, f_2, W) = \frac{\sum_{\vec{m}\in W}(f_1(\vec{m}) - \overline{f}_1)(f_2(\vec{m}) - \overline{f}_2)}{\sqrt{\sum_{\vec{m}\in W}(f_1(\vec{m}) - \overline{f}_1)^2 \sum_{\vec{m}\in W}(f_2(\vec{m}) - \overline{f}_2)^2}}. \tag{2.11}$$

The zero mean normalised cross-correlation measure has the additional property that

$$\Phi_0(f_1, f_2, W) = \Phi_0(f_1, af_2 + b, W), \tag{2.12}$$

where $a \in \mathbb{R}^+$ and $b \in \mathbb{R}$.

A consequence of the property stated in Equation 2.12 is that the zero mean normalised cross-correlation measure of match is less sensitive to absolute intensity, contrast and illumination.

2.3.6 Searching for Corresponding Points

Based on the assumptions stated in Section 2.3.5, the search for potential corresponding points entails finding the position of sub-images for which the measure of match/mismatch is a maximum/minimum. The technique that is used is related to the measure of match or mismatch. For example, in the case of the l_2 error norm, a differential displacement searching technique[5] is often used (see [49,27]). However, this technique would be inappropriate for a normalised cross-correlation measure. Techniques for finding the potential corresponding points in the case of a normalised cross-correlation measure will be described in this section.

In the case of a normalised cross-correlation measure an exhaustive search technique is used. A *search window* S, a subset of $\mathbb{Z}^2$, defines the sub-images that will be searched. It is usually rectangular in nature. The size of the search window (the number of elements in S) is determined by the range of disparities present in the scene. Multiresolution techniques allow the size of the search window to be reduced. This has two advantages:

- The computational requirements are reduced, and
- The number of false targets is reduced.

Consider two digital images I_1 and I_2 with intensity functions g_1 and g_2, respectively. Since I_1 and I_2 are digital images g_1 and g_2 are only defined on the discrete Cartesian grid of integers pairs. Let a pair corresponding points P_1 and P_2 of images I_1 and I_2 have coordinates $\vec{m}_1$ and $\vec{m}_2$, respectively. Assume that it is required to find the corresponding point of P_1 and that $\vec{m}_1 \in \mathbb{Z}^2$, i.e. the problem is that of finding the coordinates P_2 given the coordinates of P_1, and the images I_1 and I_2.

[5] It is an iterative procedure in which the matching error and the image differential are used to obtain the next disparity estimate. The procedure usually converges in a few steps.

The first step in finding the coordinates of P_2 is to calculate the correlation function ρ for sub-image defined by the search window. Thus, for all $\vec{k} \in S$ the following equation is evaluated:

$$\rho(\vec{k}) = \Phi(T[g_1, \vec{m}_1], T[g_2, \vec{m}_1 + \vec{k} + \vec{k_0}], W), \tag{2.13}$$

where Φ is the cross-correlation measure of match (see Section 2.3.5), T is a translation operator used to extract the sub-images, W is the correlation window (see Section 2.3.5) and $\vec{k}_0 \in \mathbb{Z}^2$ is an initial estimate of the disparity at $\vec{m}_1$. The translation operator T is defined by:

$$T(f(\vec{m}), \vec{k}) = f(\vec{m} + \vec{k}). \tag{2.14}$$

The second step is to find the local maxima in the correlation function ρ. A point $\vec{u}$ is a local maximum if $\rho(\vec{u})$ is greater than the value of ρ at all the nearest neighbours of $\vec{u}$. A point $\vec{l} = (l_1, l_2)$ is a nearest neighbour of the point $\vec{u} = (u_1, u_2)$, if and only if

$$\max\{|l_1 - u_1|, |l_2 - u_2|\} = 1. \tag{2.15}$$

Each local maximum corresponds to a potential corresponding point. The position of the local maxima with the largest correlation coefficient is generally selected as the position of the corresponding point. Techniques for estimating the position of the corresponding points to sub-pixel accuracy will be discussed in Section 2.3.7.

2.3.7 Estimation of the Positions of the Corresponding Points to Sub-pixel Accuracy

In the case of a normalised cross-correlation measure, a local maximum in the correlation function is assumed to indicate the position of a potential corresponding point. The sub-pixel accuracy requirement implies that the position of the correlation maxima must be determined to sub-pixel accuracy. Two techniques that have been used to achieve this will be described in this section.

The first technique is an extension of the one dimensional solution. In the case of a one dimensional correlation, the position of the correlation maxima can be estimated to sub-pixel accuracy using quadratic interpolation of $\rho(u)$ [27, pg. 317]. This leads to the following estimate for the position of the local maximum, $\hat{u}$ [46, pg. 139]:

$$\hat{u} = u_0 - \frac{1}{2} + \frac{\rho(u_0) - \rho(u_0 - 1)}{2\rho(u_0) - \rho(u_0 - 1) - \rho(u_0 + 1)} \tag{2.16}$$

where u_o is position of the maxima on the discrete grid, and ρ is the discrete correlation function.

This can be extended to two dimensions by estimating the horizontal and vertical position components using Equation 2.16. The estimated position of the correlation peak $(\hat{u}, \hat{v})$ is given by:

$$\hat{u} = u_0 - \frac{1}{2} + \frac{\rho(u_0, v_0) - \rho(u_0 - 1, v_0)}{2\rho(u_0, v_0) - \rho(u_0 - 1, v_0) - \rho(u_0 + 1, v_0)} \tag{2.17}$$

$$\hat{v} = v_0 - \frac{1}{2} + \frac{\rho(u_0, v_0) - \rho(u_0, v_0 - 1)}{2\rho(u_0, v_0) - \rho(u_0, v_0 - 1) - \rho(u_0, v_0 + 1)} \tag{2.18}$$

where (u_0, v_0) is the position of the peak on the discrete grid, and ρ is the two-dimensional correlation function.

An alternative approach is to model the correlation surface by a second order two-dimensional polynomial [54]:

$$\rho(u, v) = a_0 + a_1 u + a_2 v + a_3 u^2 + a_4 uv + a_5 v^2. \tag{2.19}$$

The coefficients can be obtained by least squares adjustment using a 3 by 3 window around the discrete peak in the correlation. Using the coefficients of the polynomial the position of the correlation peak can be estimated by:

$$\hat{u} = \frac{a_2 a_4 - 2a_1 a_5}{4a_3 a_5 - a_4^2} \tag{2.20}$$

$$\hat{v} = \frac{a_1 a_4 - 2a_3 a_2}{4a_3 a_5 - a_4^2}. \tag{2.21}$$

2.4 The Wavelet Transform and Multiresolution Analysis

Multiresolution techniques have been widely applied to digital stereo image matching (see [4,15,40]). The wavelet transform can be viewed as a multiresolution decomposition technique. These decompositions provide a representation that is intermediate between a spatial and a Fourier representation [38]. This section reviews the theory of the wavelet transform and other multiresolution techniques, in particular the construction of a sequence of images at different resolutions.

The use of multiresolution techniques is motivated by two observations. First, the structures within an image have very different sizes and hence, it not possible to define, *a priori,* an optimal resolution for analysing images. Secondly, multiresolution models have been particularly successful in explaining some low-level biological vision processes [38].

The window Fourier transform is the best-known decomposition that is intermediate between a spatial and a frequency representation, and has found wide application in signal processing. The wavelet transform is designed to overcome some the shortcomings of the window Fourier transform. The wavelet transform expands a signal into a family of functions that are the dilations and translations of a unique function $\psi(x)$. The wavelet transform can be interpreted as a decomposition into a set of frequency channels having the same bandwidth on a logarithmic scale, i.e. constant Q.

A multiresolution decomposition of an image produces a sequence of images at increasing resolutions $(r_j)_{j \in \mathbb{Z}}$. The detail of an image at resolution r_j is defined as the difference of information between its approximation at resolution r_j and its approximation at the lower

resolution r_{j-1} [37]. The Laplacian/Gaussian pyramid is an example of such a representation that has been applied to image coding [5] and digital stereo image matching [26].

The theoretical background to the production of the Laplacian/Gaussian pyramid of Burt and Adelson [5] and a wavelet based image pyramid will be discussed in this section. The definition of the integral wavelet transform will be presented in Section 2.4.1. More detail of the mathematical formulation can be found in text on the subject, for example [8,13]. General pyramidal multiresolution decompositions will be discussed in Section 2.4.2. The case of the Gaussian pyramid of Burt and Adelson [5] will be discussed in Section 2.4.3 and the case of the wavelet pyramid in Section 2.4.4.

2.4.1 Definition of the Integral Wavelet Transform

The integral wavelet transform (IWT) is defined in terms of dilations and translations of a prototype or *mother* wavelet $\psi(x)$. The IWT maps a one-dimensional function $f(x)$ onto a time-scale space by

$$W_f(a,b) = \int_{-\infty}^{\infty} f(x)\psi_{ab}^*(x)dx \tag{2.22}$$

where

$$\psi_{ab}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right), \tag{2.23}$$

a is the dilation or scale parameter, b is the translation parameter and $\psi_{ab}^*(x)$ is the complex conjugate of $\psi_{ab}(x)$.

The function $\psi(x)$ must satisfy the *admissibility* condition:

$$C_\psi = \int_0^{\infty} \frac{|\hat{\psi}(\omega)|}{w}dw < \infty, \tag{2.24}$$

where $\hat{\psi}(\omega)$ is the Fourier transform of $\psi(x)$. If the admissibility condition is satisfied then the transform is invertible and it is given by

$$f(x) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_0^{\infty} W_f(a,b)\psi_{ab}(x)\frac{dadb}{a^2}. \tag{2.25}$$

The admissibility condition implies that the DC gain of $\psi(x)$ is zero and hence, $\psi(x)$ is bandpass in nature.

2.4.2 Pyramidal Multiresolution Decompositions

General pyramidal multiresolution decompositions are described in this section. The special case of the Gaussian pyramid will be discussed in Section 2.4.3 and the case of the wavelet pyramid in Section 2.4.4. These sections are largely based on the work of Mallat [38], and Burt and Adelson [5].

A pyramidal multiresolution decomposition produces a sequence of approximations of a signal at dyadic resolutions, i.e. the resolution is an integer power of two. The resolution r

refers to the number of samples per unit length used to approximate the original signal $f(x)$. These samples are computed by uniform sampling at a rate r the signal $f(x)$ smoothed by a lowpass filter whose bandwidth is proportional to r. These lowpass filters are derived from a unique function $\theta(x)$ which is dilated by the resolution factor r: $\theta_r(x) = \sqrt{r}\theta(rx)$. Hence, the discrete approximation of a signal $f(x)$ at resolution 2^j is given by

$$A_{2^j} f = (f * \theta_{2^j}(\frac{n}{2^j}))_{n \in \mathbf{Z}} \tag{2.26}$$

The lowpass filter function $\theta(x)$ is chosen such that its Fourier transform is of the form

$$\hat{\theta}(\omega) = \prod_{p=1}^{\infty} U(e^{-i2^{-p}\omega}) \tag{2.27}$$

where $U(e^{-i\omega})$ is the transfer function of a discrete lowpass filter. It can be shown that the discrete approximation of $f(x)$ at resolution 2^j can be calculated by filtering $A_{2^{j+1}} f$ with the discrete lowpass filter $U = (u_n)$ and sub-sampling by a factor of two.

The signal that is provided by a measuring device has a finite resolution. Its resolution is usually normalised to one. A set of discrete approximations can be derived from it.

2.4.3 Gaussian Pyramid

Burt and Adelson [5] devised an algorithm for generating such a set of approximations and called it a Gaussian pyramid. They also devised an algorithm for calculating the difference in information between two resolutions. The resulting structure was called the Laplacian pyramid.

2.4.4 Wavelet Pyramid

In the case of the wavelet model an orthonormal projection of $f(x)$ on the vector space V_{2^j} is used. The vector space V_{2^j} can be viewed as the set of all possible approximations of functions at resolution 2^j. The sequence of vector spaces $(V_{2^j})_{j \in Z}$ is called a *multiresolution approximation* of $L^2(R)$.

To compute this approximation an orthonormal basis of V_{2^j} is required. It can be shown [8,13] that such an orthonormal basis can be build by dilating and translating a particular function $\phi(x)$ called a *scaling function.* The Fourier transform of $\phi(x)$ is characterised by

$$\hat{\phi}(\omega) = \prod_{p=1}^{\infty} H(e^{-i2^{-p}\omega}) \tag{2.28}$$

where $H(e^{-i\omega})$ is the transfer function of a discrete filter. Furthermore, $H(e^{-i\omega})$ satisfies the condition

$$|H(e^{-i\omega})|^2 + |H(-e^{-i\omega})|^2 = 1. \tag{2.29}$$

The difference of information between the approximations of a function at resolutions 2^j and 2^{j+1} is extracted as follows. The approximation at resolution 2^{j+1} must be a better estimate of $f(x)$ than the approximation at resolution 2^j. Hence, the vector spaces V_{2^j} and $V_{2^{j+1}}$ should satisfy

$$V_{2^j} \subset V_{2^{j+1}}. \tag{2.30}$$

The difference of information between the approximations at resolutions 2^j and 2^{j+1} is therefore equal to the orthogonal projection of $f(x)$ on the orthogonal complement of V_{2^j} in $V_{2^{j+1}}$. It can be shown [8,13] that an orthonormal basis of the orthogonal complement vector space O_{2^j} can be built by dilating and translating a particular wavelet $\psi(x)$.

A description of the algorithms used to generate the wavelet pyramid can be found in the work of Daubechies [13]. The applicability of the wavelet pyramid to image matching will be discussed in Section 7.6.

Chapter 3

Optical Remote Sensing Theory

3.1 Introduction

Remote sensing is the acquisition of information about an object without coming in physical contact with it, i.e. sensing at a distance. This involves the detection and measurement of the interaction of the object with surrounding fields, such as electromagnetic, acoustic, or potential (gravitational or magnetic) fields.

Optical remote sensing refers to remote sensing techniques that measure the electromagnetic radiation emanating from an object in the optical region of the spectrum. The optical region of the spectrum is generally considered to extend from 0.4 μm to 1000 μm, but it is further restricted to atmospheric transmission windows within 0.4 μm to 15 μm [1]. Only the reflective portion of the optical spectrum (0.4 μm to 2.4 μm) is of interest in this dissertation.

An optical remote sensing data-acquisition system can be considered to have four component parts: the radiation source, the atmospheric path, the target object and the sensor [60,11] (see Figure 3.1). Passive systems[1] use the Sun as their radiation source. It radiates energy throughout the electromagnetic spectrum with a peak in the visible region. The energy propagates through space, and then through the atmosphere, where its intensity and spectrum distribution is modified. The energy then interacts with the target object. It is reflected, transmitted, and/or absorbed by the target object. The energy that is absorbed in one region of the spectrum may be emitted in another region of the spectrum. The reflected/emitted energy again passes through the atmosphere and again is subject to modification by the atmosphere. Finally, the energy reaches the sensor where it is measured and converted into data for subsequent processing.

This chapter presents the theoretical background to optical remote sensing systems. Section 3.2 presents the relevant radiometric theory to estimate the radiation that reaches the sensor. Section 3.3 discusses the terms: spatial resolution, spectral resolution, radiometric

[1] A system is considered *passive* if it does not provide the radiation that is finally detected by the sensor. The rest of this chapter will assume a passive system unless otherwise stated.

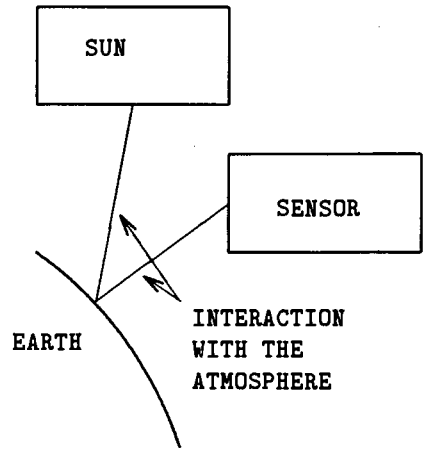


Figure 3.1: Component parts of an optical remote sensing data-acquisition system.

resolution, and temporal resolution. These terms refer to the ability of the sensor to differentiate between objects. Section 3.4 deals with the quantum nature of light (electromagnetic radiation) and its effect on the noise performance of all light sensing devices. Section 3.5 discusses the use of charge coupled devices (CCDs) as imaging devices (the detector). The noise characteristics and spatial frequency response will be discussed. Section 3.6 will describe the effect of quantisation noise on the signal-to-noise ratio of the imaging sensor. Quantisation is the step of converting the signal produced by the detector into digital data. This is required for transmission to Earth and subsequent processing. Section 3.7 deals with the orbital mechanics of a satellite in a circular, sun-synchronous, near-polar orbit. It is the orbit used by satellites that carry high resolution imaging sensors (see Section 4.2). A model of satellite sensor geometry will be presented in Section 3.8. It will allow the coordinates of the observed points to be estimated.

3.2 Radiometric Theory

The objective of this section is to present the radiometric theory that will enable an estimate to be made of the radiation which reaches the sensor. Issues that affect the design and performance of the system will also be discussed. This will be achieved by presenting an overview of radiometric terms and geometric optics, and by considering the four component parts of data-acquisition system described in Section 3.1.

Basic radiometric terms and concepts will be presented in Section 3.2.1 and basic geometrical optic terms will be discussed in Section 3.2.2. Section 3.2.3 will discuss thermal radiation sources, in particular the Sun. The interaction of the radiation with the atmosphere and with the surface of the Earth will be discussed in Sections 3.2.4 and 3.2.5, respectively. Finally,

an equation for the irradiance received by the sensor will be presented in Section 3.2.6. The material in these sections is derived from the work of Slater [57], Goetz and Rowan [20], Suits [59], and Mather [41].

3.2.1 Radiometric Terminology

This section will present the basic radiometric terms and concepts. The radiation quantities, spectral radiation quantities, and concepts of reflectance, transmittance and absorptance will be introduced.

The radiation quantities are used to describe quantitative measurements. *Radiant energy*, Q_e, is the energy associated with electromagnetic radiation. *Radiant flux* or *radiant power*, Φ_e, is the rate of radiant energy transfer from one place to another. It is measured in *watts* (W). *Radiant flux density* is the magnitude of the radiant flux that is incident upon or, conversely, is emitted by a surface of unit area. It is measured in *watts per square meter* (W m^{-2}). *Radiant incidance* or *irradiance*, E_e, is used in place of radiant flux density if radiant energy is incident on a real or imaginary surface. *Radiant exitance* or *emittance*, M_e, is used in place of radiant flux density if the flow of energy is away from a surface. *Radiant sterance* or *radiance*, L_e, is the radiant flux density transmitted from a small area and viewed through a solid angle. It is measured in *watts per square meter per steradian* (W m^{-2} sr^{-1}).

Spectral radiometric quantities describe the spectral distribution of the radiation. This includes the *spectral radiant energy*, Q_λ, the *spectral radiant flux*, Φ_λ, the *spectral irradiance*, E_λ, the *spectral exitance*, M_λ, and the *spectral radiance*, L_λ.

The concepts of reflectance, transmittance and absorptance describe the degree to which irradiance that is incident on a plane of material is reflected, transmitted and absorbed by the plane, respectively. The *reflectance*, ρ, is defined by the ratio of the reflected exitance of a plane of material to the irradiance on that plane:

$$\rho = M_{reflected}/E. \tag{3.1}$$

The *transmittance*, τ, is defined by the ratio of the transmitted exitance, leaving the opposite side of the plane, to the irradiance:

$$\tau = M_{transmitted}/E. \tag{3.2}$$

The *absorptance* α is defined by:

$$\alpha = 1 - \tau - \rho. \tag{3.3}$$

The corresponding *spectral reflectance*, $\rho(\lambda)$, *spectral transmittance*, $\tau(\lambda)$, and *spectral absorptance*, $\alpha(\lambda)$, are defined by:

$$\rho(\lambda) = M_{reflected}(\lambda)/E(\lambda), \tag{3.4}$$

$$\tau(\lambda) = M_{transmitted}(\lambda)/E(\lambda), \tag{3.5}$$

$$\alpha(\lambda) = 1 - \tau(\lambda) - \rho(\lambda). \tag{3.6}$$

3.2.2 Geometric Optics Terms

This section will introduce the basic terms and relationships of geometrical optics. A detailed discussion of geometrical optics can be found in [33,25,28]. These terms and relationships will be used in latter sections. The thin lens equation [25],

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}, \tag{3.7}$$

relates the distance from the lens to the object o, the distance from the lens to the image i and the focal length of the lens f. The lateral magnification of a lens is given by

$$m = \frac{i}{o}. \tag{3.8}$$

The F/number of a lens N is defined as the ratio of the focal length of a lens f to the diameter of the entrance pupil D,

$$N = \frac{f}{D}. \tag{3.9}$$

3.2.3 Thermal Radiation Sources

A discussion of thermal radiation sources, in particular the Sun, is presented in this section. The Sun provides the electromagnetic energy that is finally detected by the sensor. The quantity and spectral distribution of the radiation emitted by the Sun are required to estimate the radiation that reaches the detector.

As a first approximation the Sun can be modelled as a *blackbody*, an ideal thermal emitter, at a temperature of about 5900 K [57]. The spectral exitance, M_λ, of a blackbody is given by Planck's formula [59],

$$M_\lambda = \frac{c_1 \lambda^{-5}}{e^{\frac{c_2}{\lambda T}} - 1}, \tag{3.10}$$

where

$c_1 = 3.74 \times 10^{-16}$ W m^2,

$c_2 = 1.44 \times 10^{-2}$ m K,

λ is the wavelength in metres, and

T is the absolute temperature in Kelvin.

The Earth is also a thermal emitter, but with the temperature of only about 290 K. The spectral exitance of a blackbody at 5900 K and 290 K is shown in Figure 3.2.

The Sun, however, is not an ideal blackbody. A better approximation of the quantity and spectral distribution of the radiation, which reaches the Earth, can be obtained by measurement. The spectral solar irradiance that does reach the Earth is shown in Figure 3.4.

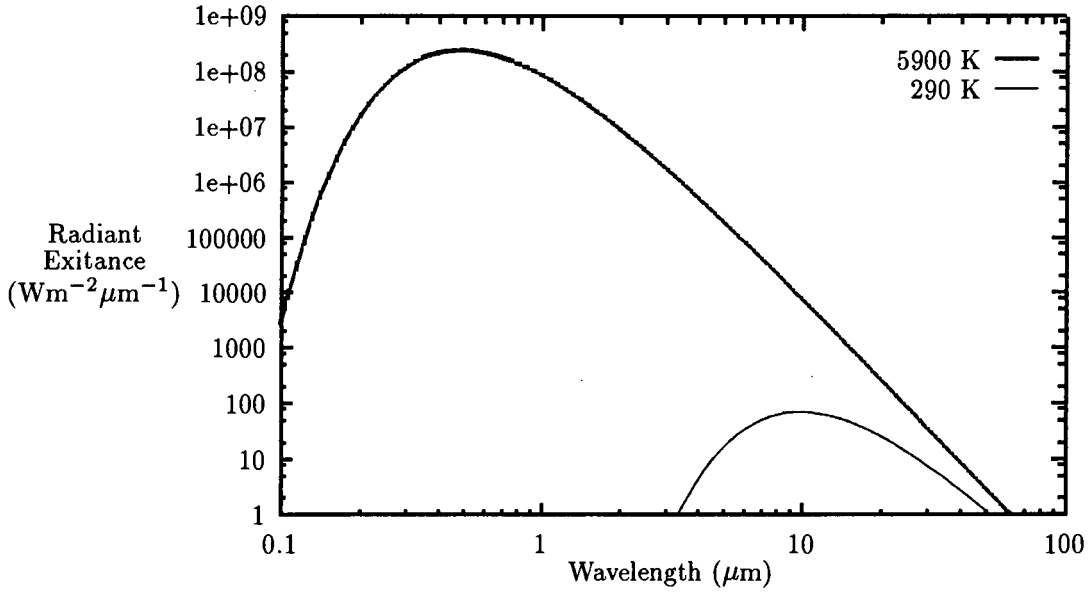


Figure 3.2: Spectral exitance for a blackbody at 5900 K and 290 K, the approximate temperature of the Sun and the Earth.

3.2.4 Interaction with the Atmosphere

The electromagnetic radiation emitted by the Sun passes through the atmosphere before and after it interacts with the surface of the Earth. During its passage through the atmosphere it interacts with the particulate matter suspended in the atmosphere and with the molecules of the constituent gases.

This interaction is in the form of *scattering* and *absorption*. The combined effect is to reduce the contrast of the image and the amount of radiation that reaches the sensor. This reduction is termed *attenuation* or *extinction*. Detailed discussions of scattering and absorption mechanisms can be found in [6,34].

Scattering is the deflection of the radiation from its path. The mechanisms of scattering are beyond the scope of this dissertation. However, it is important to note that shorter wavelengths are more seriously affected by scattering. It is the reason why the blue end of the optical spectrum (less than 0.5 μm) is not often used for remote-sensing.

Absorption is the conversion of the energy present in the electromagnetic radiation into the internal energy of the absorbing molecule. This energy is often emitted in another region of the electromagnetic spectrum.

Both absorption and scattering vary in their effect from one part of the spectrum to another. Remote sensing is impossible in those parts of the spectrum that are seriously affected by scattering and/or absorption. Spectral atmospheric transmittance for the optical region of the electromagnetic spectrum is depicted in Figure 3.3. It shows the atmospheric

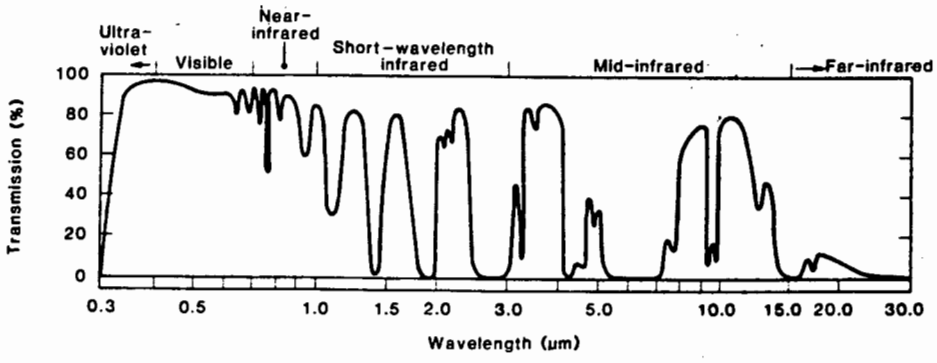


Figure 3.3: Spectral atmospheric transmittance for the optical region of the electromagnetic spectrum (from [20]).

transmission windows.

The effect of the atmosphere on the irradiance, which reaches the surface of the Earth, is shown in Figure 3.4. It depicts the spectral irradiance of the solar radiation at the top of the atmosphere and at sea level.

The degree of absorption and scattering is dependent on time and position. This introduces a noise component into the signal detected by the sensor. The noise due to atmospheric interaction can be divided into two components [57]: long-term variations (over periods of minutes or hours) of atmospheric transmittance, and short-term variations (in the frequency range 0 to 1 kHz). Since the observation interval is relatively short, the short-term variations are of the most interest. Salter [57] stated that a reasonable estimate for $\sigma(E)/\overline{E}$, for short term fluctuations, is of the order of 1% of the average values, where E is the radiant incidence on the surface of the Earth. Hence, the signal reaching the detector has a significant noise component.

3.2.5 Interaction with the Surface of the Earth

The interaction of electromagnetic radiation with the surface of the Earth is in the form of *reflection*, *transmission*, and *absorption*. The energy absorbed may be *emitted* in a different region of the spectrum (usually the thermal infrared). Reflection is of interest, due to the use of the reflective portion of the optical spectrum.

Two types of reflection can occur at a surface. The first type is *specular* refection in which the energy leaves the surface without being scattered, i.e. with the angle of incidence being equal to the angle of reflectance. Surfaces that reflect specularly are smooth relative to the wavelength of the incident energy. Secondly, *diffuse* reflection in which the incident energy is scattered in all directions. This occurs when the reflecting surface is rough relative to the

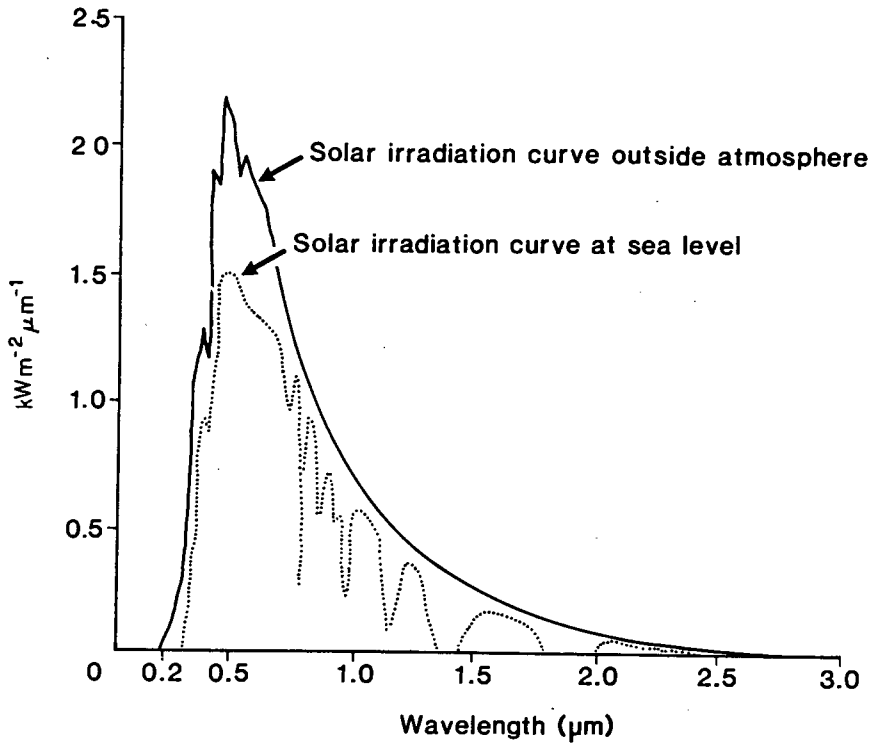


Figure 3.4: Spectral irradiance at the top of the atmosphere (solid line) and at sea level (dotted line) (from [41]).

wavelength of the incident energy.

The reflective properties of a surface are best described by the *bidirectional reflectance distribution function* or BRDF (see [57]). The BRDF takes all possible observation and illumination angles into account.

A simpler reflectance model is that of a *Lambertian* surface. It has the property that the radiance emitted by the surface is constant for any observation angle θ measured from the surface normal [33, pg. 207]. If no other information is available about the reflectance characteristics, it is best to consider the surface of the Earth to be a Lambertian surface.

The spectral reflectance characteristics of different Earth-surface materials allow them to be differentiated. A discussion of the spectral reflectance of Earth-surface materials can be found in [41, pp. 27-34].

3.2.6 Image-Plane Irradiance

The image-plane irradiance determines the signal received by the sensor. A general equation for the image-plane irradiance will be presented in this section.

From basic geometrical optics and radiometric considerations, it can be shown [57] that

the spectral irradiance in the image plane of a system imaging an extended Lambertian surface (see Section 3.2.5) is given by

$$E_{i\Delta\lambda} = \frac{(1-A)^2 K_N(\alpha)\cos^n(\alpha)}{4N^2(1-m)^2}\int_{\lambda_1}^{\lambda_2}\left[E_\lambda \rho(\lambda)\tau_u(\lambda) + \pi L_u(\lambda)\right]\tau_o(\lambda)d\lambda + \int_{\lambda_1}^{\lambda_2} E_{\lambda f}\, d\lambda, \quad (3.11)$$

where

$E_{i\Delta\lambda}$ = image irradiance in wavelength range $\Delta\lambda = \lambda_2 - \lambda_1$,

A = ratio of the diameter of the central obscuration to that of the primary mirror,

$K_N(\alpha)$ = vignetting factor,

$\cos^n(\alpha)$ = off axis reduction in image irradiance (n is usually 4),

α = the off axis angle,

N = F/number of the lens,

m = lateral magnification (see Equation 3.8),

E_λ = object plane irradiance,

$\rho(\lambda)$ = spectral reflectance of the ground object,

$\tau_u(\lambda)$ = spectral transmittance of the atmosphere along the atmospheric path between the sensor and the ground object,

$L_u(\lambda)$ = upward radiance along the atmospheric path between the ground and the sensor (due to atmospheric scattering),

$\tau_o(\lambda)$ = spectral transmittance of the optical system, and

$E_{\lambda f}$ = flare or veiling glare of the optical system.

Equation 3.11 can be simplified by ignoring the effect of parameters that are not well known, i.e. by setting $K_N(\alpha) = 1$, $\cos^n(\alpha) = 1$, $L_u(\lambda) = 0$ and $E_{\lambda f} = 0$. Then Equation 3.11 becomes

$$E_{i\Delta\lambda} = \frac{(1-A)^2}{4N^2(1-m)^2}\int_{\lambda_1}^{\lambda_2} E_\lambda \rho(\lambda)\tau_u(\lambda)\tau_o(\lambda)d\lambda, \quad (3.12)$$

or as a function of λ

$$E_i(\lambda) = \frac{(1-A)^2}{4N^2(1-m)^2} E_\lambda \rho(\lambda)\tau_u(\lambda)\tau_o(\lambda). \quad (3.13)$$

Equation 3.13 will be used with characteristics of the detector to determine the output signal of the sensor under different conditions (see Section 5.5).

3.3 Resolution of Imaging Sensors

The major characteristics of imaging sensors can be described by their *spatial*, *spectral*, *radiometric* and *temporal* resolutions. This section introduces these four terms.

3.3.1 Spatial resolution

The spatial resolution of an imaging sensor is not an easy concept to define. It is measured in a number of different ways, depending on the user's purpose. This section will introduce some ways of defining the spatial resolution of an imaging sensor.

The *instantaneous field of view* or IFOV is defined as the area on the ground covered by the image of a detector element as magnified by the ratio of the sensor altitude to the focal length of the optics of the sensor system. It is given by

$$\text{IFOV} = \frac{od}{i} = \frac{d}{m} \tag{3.14}$$

where o is the distance from the lens to the object, i is the distance from the lens to the image, d is the detector size and m is the lateral magnification (see Section 3.2.2). It can also be expressed as an angle. The IFOV is not always a very useful measure of the spatial resolution, since it does not take the properties of the optical components into account.

The *point spread function* or PSF is the imaged intensity distribution of a single point source. The Rayleigh criterion states that two equally-bright point sources will be distinguishable on the image if the bright central disc of one falls on the first dark ring of the PSF of the other [41].

The *modulation transfer function* or MTF describes the spatial resolution of the imaging sensor in terms of its spatial frequency response. The MTF at a given spatial frequency is given by the ratio

$$\text{MTF}(\nu) = \frac{M_i(\nu)}{M_o(\nu)} \tag{3.15}$$

where $M_i(\nu)$ and $M_o(\nu)$ are the modulation of the image and object as a function of spatial frequency ν. The modulation M is defined as

$$M = \frac{E_{max} - E_{min}}{E_{max} + E_{min}} \tag{3.16}$$

where E_{max} and E_{min} are the maximum and minimum radiance values.

The MTF can be applied not only to the imaging sensor, but also to the individual components of the imaging sensor, including the optics, the detector and image smear due to motion. If the sensor components are assumed to be linear and shift invariant then the modulation transfer function of an imaging sensor can be written as [56]

$$\text{MTF}_S(\nu) = \text{MTF}_L(\nu)\text{MTF}_I(\nu)\text{MTF}_D(\nu), \tag{3.17}$$

where

$\mathrm{MTF}_S(\nu)$ is modulation transfer function of the system,

$\mathrm{MTF}_L(\nu)$ is modulation transfer function of the lens,

$\mathrm{MTF}_I(\nu)$ is modulation transfer function due to image smear, and

$\mathrm{MTF}_D(\nu)$ is modulation transfer function of the detector.

The term *effective instantaneous field of view* (EIFOV) was defined and formulated, during a working-group meeting, in 1973, between NASA and user community representatives [57, pp. 26-27]. The EIFOV is defined, for sinusoidal distribution of radiance, as the spatial half wavelength for which the modulation of the distribution of radiance has been reduced by one-half due to the MTF of the system.

3.3.2 Spectral resolution

Many imaging sensors are referred to as *multi-spectral* or *multi-band* systems, which means they can record an image of the same scene separately in each of several spectral bands. The term, spectral resolution, refers to the width of these spectral bands. The width of the spectral bands is not the only important parameter. The position in the spectrum and the number of bands is also important. A discussion of the positioning of spectral bands can be found in [41, pp. 27-34].

3.3.3 Radiometric resolution

The data collected by the sensor is quantised into a finite number of levels. The *radiometric resolution* refers to the number of levels used. An increase in the number of levels results in an increase in the detail of information.

The radiometric resolution, however, cannot simply be increased [41]. The signal-to-noise ratio of the sensor and the effect of quantisation on the signal-to-noise ratio has to be considered when determining the number of levels or radiometric resolution to be used (see Section 3.6).

3.3.4 Temporal resolution

The temporal resolution of an imaging sensor refers to its ability to detect change with time. For example, images taken at 1-week to 1-month intervals could be used to detect crop development. The repeat cycle of the satellite is the primary factor determining the temporal resolution. This is determined by the orbit of the satellite and the swath width. The swath width is the total width of the strip of the Earth in field of view of the sensor. It is given by the IFOV multiplied by the number of detector elements (assuming "pushbroom" operation).

3.4 The Quantum Nature of Light and Shot Noise

This section analyses the effects of the quantum nature of light on the noise performance of light-sensing devices. This leads to the concept of shot noise that represents an inherent, natural barrier to noiseless operation of all light-sensing devices, including CCDs (see Section 3.5) and the human eye.

Light can be considered to consist of discrete quanta called *photons*. The energy per photon is given by

$$E = h\frac{c}{\lambda}, \tag{3.18}$$

where E is the energy per photon, h is Planks constant (6.63×10^{-34} J s), c is the speed of light and λ is the wavelength of the light.

The radiant flux incident on a surface can be modelled as a stream of randomly incident photons. The average rate of incidence a, in photons per second, is related to the radiant flux Φ, and the wavelength of the light λ by

$$a = \frac{\Phi\lambda}{hc}. \tag{3.19}$$

If the actual number of photons incident on the surface is carefully observed, it will be found to fluctuate. This fluctuation constitutes an AC noise signal superimposed on the DC, or average value. The magnitude of the AC noise depends on the period of observation and the average photon current. This fluctuation is called shot noise.

To calculate the magnitude of the shot noise, the statistical properties of shot noise have to be investigated. The statistical analysis of shot noise, given below, is based on the work of Schreiber [55, pp. 54-57]. The probability of getting K photons in a period of t seconds, if the average photon current is a photons per second, is calculated by dividing t into n intervals. The probability of getting a photon in one of these intervals is at/n and if at/n is small enough, the probability of more than one is negligible, i.e. if n is very large. The probability of getting exactly K photons in t seconds can now be modelled by the binomial distribution [43, pg. 57] and thus it is given by

$$P_n(K,t) = \frac{n!}{K!(n-K)}\left(\frac{at}{n}\right)^K\left(\frac{1-at}{n}\right)^{n-K}. \tag{3.20}$$

As $n \to \infty$ the binomial distribution tends towards the Poisson distribution [43, pp. 79-80] and, hence, the probability of getting exactly K photons in t seconds becomes

$$P(K,t) = \frac{(at)^K}{K!}e^{-at}. \tag{3.21}$$

For a Poisson distribution the mean of the distribution, μ, is equal to the variance of the distribution, σ^2 [43, pg. 82]. Hence, the variance or mean square deviation of K is

$$\sigma_K^2 = \overline{K} = at. \tag{3.22}$$

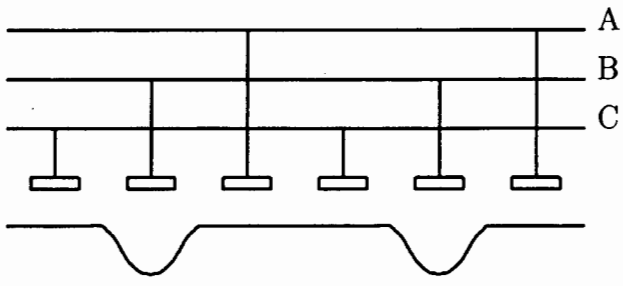


Figure 3.5: Potential wells of a CCD formed by applying a higher voltage to line B than to lines A and C. (from [17]).

Hence, the RMS shot noise is equal to the square root of the average number of photons detected by the light-sensing device in the time t. Shot noise places a lower limit on the light levels that can be used.

3.5 Charge Coupled Devices

The theoretical background to the operation of *charge coupled devices* (CCDs) as electro-optical imaging detectors will be presented in this section. A discussion of other electro-optical imaging detectors used for remote sensing can be found in [48].

The discussion will be limited to the case of line-array (one-dimensional) CCDs. Area-array (two-dimensional) CCDs are beyond the scope of this dissertation.

The principles of operation of CCDs will be described in Section 3.5.1. The noise characteristics will be discussed in Section 3.5.2. Finally, the spatial frequency response will be discussed in Section 3.5.3.

3.5.1 Principles of Operation

The basic principles of operation of line-array CCDs as electro-optical imaging detectors will be described in this section.

A line-array, CCD contains a linear array of pixels. The connection of the pixels to several address lines and the voltages applied create potential wells in the CCD substrate, a doped semiconductor. This is illustrated in Figure 3.5 where line B has a higher voltage than lines A and C. Charge can be stored in these potential wells and moved, as a packet, across the surface by translating the potential minima. This is done by appropriate clocking of the address lines.

The CCD operates as a photon detector. An absorbed photon activates an electron that is then stored in the corresponding potential well. A clocking scheme controls this photo-

electron collection. During the charge integration interval, electrons are collected and stored in the potential wells. During the charge transfer interval, the electron packets are transferred to the CCD output.

The expected number of electrons to be accumulated in a potential well is equal to the time and space integral of the irradiance as modified by the quantum efficiency of the CCD material. Assuming the irradiance is constant over the active area of the pixel and during the integration period, the expected number of electrons, n, is given by:

$$n = \frac{E_\lambda \lambda}{hc} A T_i \eta \tag{3.23}$$

where λ is the wavelength of the incident radiation, E_λ is the spectral irradiance, h is Planks constant, c is the speed of light, A is the active area of the pixel, T_i is the integration period and η is the quantum efficiency. The quantum efficiency is the probability that an incident photon will result in the addition of an electron to the corresponding potential well.

The physical principles of operation of CCDs ensure excellent linearity. This has been verified experimentally (see [12]). The potential wells can only store a finite number of electrons, usually about 10^5 electrons [35].

Noise sources and nonlinearities include: dark current noise, shot noise, fixed pattern noise and blooming. Shot noise is due to the quantum nature of light (see Section 3.4). Dark current noise is due to current that is present without any light. It is temperature dependant and its effects can be reduced by cooling the CCD. Fixed pattern noise is due to different pixels having different quantum efficiencies (gains) and dark currents. Blooming is an overflow of charge out of the potential wells.

3.5.2 CCD Noise Characteristics

The noise characteristics of CCDs are described in this section. CCD noise can be divided into two different sources. Readout noise that is additive in nature, and shot noise is dependent on the exposure of the CCD.

The RMS amplitude of the readout noise of a CCD n_o at room temperature is usually of the order of 100 electrons [35]. It is due to various sources, dark current uncertainties, reference level uncertainties, etc. The shot noise is due to the quantum nature of light (see Section 3.4). From Equation 3.22 the shot noise power n_s^2 is given by

$$n_s^2 = n_e \tag{3.24}$$

where n_e is the expected number of electrons in the potential well of the CCD. The total noise power n_{total}^2 is given by

$$\begin{aligned} n_{total}^2 &= n_o^2 + n_s^2 \\ &= n_o^2 + n_e. \end{aligned} \tag{3.25}$$

If n_{max} is the maximum number of electrons that can be stored in a potential well, then the peak signal-to-noise ratio is given by

$$\left(\frac{S}{N}\right)_{pk} = \frac{n_{max}^2}{n_{total}^2} = \frac{n_{max}^2}{n_o^2 + n_e}. \tag{3.26}$$

Thus, the peak signal-to-noise ratio is intensity dependant. The noise power is lower in darker regions of an image than in the lighter regions.

3.5.3 The Spatial Frequency Response of CCDs

The term modulation transfer function (MTF) was introduced in Section 3.3.1 to describe the spatial resolution of an imaging sensor. The effect of a line-array CCD on the MTF of an imaging sensor will be described in this section. The sampling nature of a CCD introduces a lack of shift invariance (the MTF is dependent on the alignment between the CCD and the input signal as shown in Figure 3.6). Hence, Equation 3.17 cannot be used to calculate the MTF of the imaging sensor. The description presented in this section is largely based on the work of Feltz [17], however, a different approach is taken. This approach allows both the optical characteristics and motion effects to also be included in the MTF.

The effect of the CCD on the imaging system MTF can be divided into two parts. The first effect, due to the spatial integration, can be multiplied with other components of the imaging system MTF, as in Equation 3.17. The second effect, due to sampling nature of the CCD, introduces a dependence on the relative alignment of the input signal and the CCD.

The spatial integration component of the MTF (due to the finite size of each detector element), MTF_D, is calculated as follows. Consider the input function

$$I(s) = \frac{1}{2}(1 + \cos(\nu s)), \tag{3.27}$$

where ν is the spatial frequency. The function in Equation 3.27 has a modulation of 1. The MTF due to spatial integration is then given by

$$\text{MTF}_D(\nu) = \frac{E_{max}(\nu) - E_{min}(\nu)}{E_{max}(\nu) + E_{min}(\nu)}, \tag{3.28}$$

where E_{min} and E_{max} are the minimum and maximum radiance values after spatial integration.

The output after spatial integration is dependent on the CCD pixel active width, T_α, and is given by

$$B(s) = I(s) * f(s), \tag{3.29}$$

where

$$f(s) = \begin{cases} \frac{1}{T_\alpha} & -T_\alpha/2 < s < T_\alpha/2, \\ 0 & \text{elsewhere.} \end{cases} \tag{3.30}$$

Evaluating Equation 3.29 gives

$$B(s) = \frac{1}{2} + \frac{1}{\nu T_\alpha}\cos(\nu s)\sin\left(\frac{\nu T_\alpha}{2}\right). \tag{3.31}$$

Hence, the MTF due to spatial integration is now given by

$$\mathrm{MTF}_D(\nu) = \frac{\sin\left(\frac{\nu T_\alpha}{2}\right)}{\frac{\nu T_\alpha}{2}}. \tag{3.32}$$

The effect of the sampling nature of the CCD is to introduce a dependence on the alignment of the CCD and the input signal. Consider the input to the optical system, the object plane irradiance, being

$$I(s) = \frac{1}{2}(1 + \cos(\nu s)), \tag{3.33}$$

where ν is the spatial frequency. The MTF of the complete imaging system at the spatial frequency ν is then given by

$$\mathrm{MTF}(\nu) = \frac{B_{max}(\nu) - B_{min}(\nu)}{B_{max}(\nu) + B_{min}(\nu)}, \tag{3.34}$$

where B_{max} and B_{min} are the maximum and minimum outputs of the CCD.

The signal that the CCD samples can be obtained by applying the definition of the MTF to Equation 3.33 and it is given by

$$I_S(s) = \frac{1}{2}k(1 + \mathrm{MTF}_S(\nu)\cos(\nu s)), \tag{3.35}$$

where k is a constant dependant on the system design, and $\mathrm{MTF}_S(\nu)$ is given by Equation 3.17. $\mathrm{MTF}_S(\nu)$ includes the effect of the optical system, the sensor motion and CCD spatial integration.

The sampling rate is determined by the CCD pixel pitch, T_β. The alignment of the CCD and the input signal $I_S(s)$, illustrated in Figure 3.6, is defined by two parameters. The offset between a maximum in the input signal and the nearest sampling point, ψ_H, and the offset between a minimum in the input signal and the nearest sampling point, ψ_L.

The alignment parameters are limited to the following ranges

$$-\frac{T_\beta}{2} \leq \psi_H \leq \frac{T_\beta}{2}, \tag{3.36}$$

$$-\frac{T_\beta}{2} \leq \psi_L \leq \frac{T_\beta}{2}. \tag{3.37}$$

Hence, B_{max} and B_{min} are given by $I_S(s)$ evaluated at ψ_H and $\frac{\pi}{\nu} + \psi_L$ respectively. Evaluating $I_S(s)$ at gives

$$B_{max} = \frac{1}{2}k(1 + \mathrm{MTF}_S(\nu)\cos(\nu\psi_H)) \tag{3.38}$$

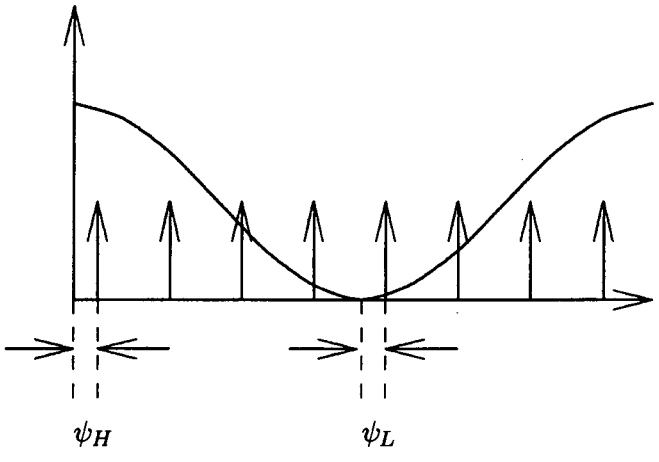


Figure 3.6: The alignment of the CCD and the input signal.

and

$$B_{min} = \frac{1}{2}k(1 + \text{MTF}_S(\nu)\cos(\pi + \nu\psi_L)) \tag{3.39}$$

Substituting into Equation 3.34 gives

$$\text{MTF}(\nu) = \frac{\text{MTF}_S(\nu)(\cos(\nu\psi_H) - \cos(\pi + \nu\psi_L))}{2 + \text{MTF}_S(\nu)(\cos(\nu\psi_H) + \cos(\pi + \nu\psi_L))} \tag{3.40}$$

Note that if $\psi_H = 0$ and $\psi_L = 0$ then $\text{MTF}(\nu) = \text{MTF}_S(\nu)$, i.e. if the input signal and CCD are in perfect alignment then the MTF is the same as it would without sampling. At the Nyquist sampling rate under worst alignment condition the MTF is zero.

3.6 Quantisation Noise

This section analyses the effect of quantisation noise, which is introduced by a quantiser, such as an analogue-to-digital converter, on the signal-to-noise ratio of a signal. The analysis is based on the work of Oppenheim and Schafer [50, pp. 114-123], and Stremler [58, pp. 509-511,525].

A quantiser is a nonlinear system, which transforms an input sample $x[n]$ into one of a finite set of prescribed values. The quantised sample is referred to as $\tilde{x}[n]$. This operation can be represented as

$$\tilde{x}[n] = Q(x[n]). \tag{3.41}$$

Quantisers can be defined with either uniformly or non-uniformly spaced quantisation levels and with bipolar or unipolar quantisation levels; however, this analysis assumes unipolar uniformly spaced quantisation levels.

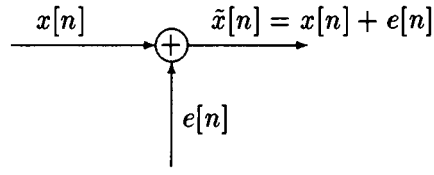


Figure 3.7: An additive noise model of a quantiser.

Generally the number of quantisation levels will be a power of 2, since these 2^B quantisation levels can be coded with a B-bit binary code. The step size of the quantiser a is dependent on the full scale amplitude of the quantiser X_m and is given by

$$a = \frac{X_m}{2^B}. \tag{3.42}$$

Now the distribution of quantisation levels is assumed to be $0, a, 2a, ...(2^B - 1)a$.

The quantised sample $\tilde{x}[n]$ will in general be different from the true sample $x[n]$. The difference between them is the *quantisation error*, defined as

$$e[n] = \tilde{x}[n] - x[n]. \tag{3.43}$$

A useful model of a quantiser is depicted in Figure 3.7. In this model the quantisation error samples are thought of as an additive noise signal. This model is exactly equivalent to a quantiser if $e[n]$ is known. In most cases $e[n]$ is not known, and a statistical model based on Figure 3.7 is used to represent the effects of quantisation. The statistical representation of quantisation errors is based on the following assumptions [50, pg. 120]:

1. The error sequence $e[n]$ is a sample sequence of a stationary random process.
2. The error sequence is uncorrelated with the sequence $x[n]$.
3. The random variables of the error process are uncorrelated; i.e. the error is a white-noise process.
4. The probability distribution of the error process is uniform over the range of quantisation error.

If the quantiser rounds the sample value to nearest quantisation level, then the amplitude of the quantisation noise is in the range

$$-a/2 < e[n] < a/2. \tag{3.44}$$

Now, if $e[n]$ is a random variable uniformly distributed from $-a/2$ to $a/2$, the probability density function for the quantisation error is

$$p(e) = \begin{cases} 1/a & -a/2 < e < a/2, \\ 0 & \text{elsewhere.} \end{cases} \tag{3.45}$$

The mean value of $e[n]$ is zero and the variance is [58, p. 511]

$$\begin{aligned} \sigma_e^2 &= \int_{-\infty}^{\infty} e^2\, p(e)\, de \\ &= \int_{-a/2}^{a/2} \frac{1}{a} e^2\, de \\ \sigma_e^2 &= \frac{a^2}{12}. \end{aligned} \tag{3.46}$$

Let the signal variance of the sequence $x[n]$ be denoted by σ_x^2 and the noise variance by σ_n^2. The signal-to-noise ratio of the sequence $x[n]$ is then given by

$$\frac{S_i}{N_i} = \frac{\sigma_x^2}{\sigma_n^2}. \tag{3.47}$$

Expressed in decibels, this becomes

$$\left(\frac{S_i}{N_i}\right)_{\text{dB}} = 10 \log \frac{\sigma_x^2}{\sigma_n^2}. \tag{3.48}$$

The signal-to-noise ratio of the sequence $\tilde{x}[n]$ is now given by

$$\frac{S_o}{N_o} = \frac{\sigma_x^2}{\sigma_n^2 + \sigma_e^2}. \tag{3.49}$$

Equation. 3.47 can be expressed in terms of σ_n^2,

$$\sigma_n^2 = \frac{\sigma_x^2}{S_i/N_i}. \tag{3.50}$$

Substituting Equation 3.50 into Equation 3.49 gives

$$\frac{S_o}{N_o} = \frac{\sigma_x^2}{\frac{\sigma_x^2}{S_i/N_i} + \sigma_e^2}. \tag{3.51}$$

Expressed in decibels, this becomes

$$\left(\frac{S_o}{N_o}\right)_{\text{dB}} = 10 \log \frac{\sigma_x^2}{\sigma_x^2\, 10^{-\frac{(S_i/N_i)_{\text{dB}}}{10}} + \sigma_e^2} \tag{3.52}$$

The variance of the sequence $x[n]$ is in general not known. The peak signal-to-RMS noise ratio of the quantised sequence $\tilde{x}[n]$, can be related to the peak signal-to-RMS noise ratio of the input sequence $x[n]$ and the number of quantisation levels $m = 2^B$. This is done by substituting $\sigma_x = m - 1$ into Equation 3.52 and it becomes

$$\left[\left(\frac{S_o}{N_o}\right)_{\text{pk}}\right]_{\text{dB}} = 10 \log \frac{(m-1)^2}{(m-1)^2\, 10^{-\frac{(S_i/N_i)_{\text{dB}}}{10}} + \frac{1}{12}}. \tag{3.53}$$

The effect of quantisation on the peak signal-to-RMS noise ratio of a sequence is shown in Figure 3.8 for 12, 10, 8, and 6 bit quantisers.

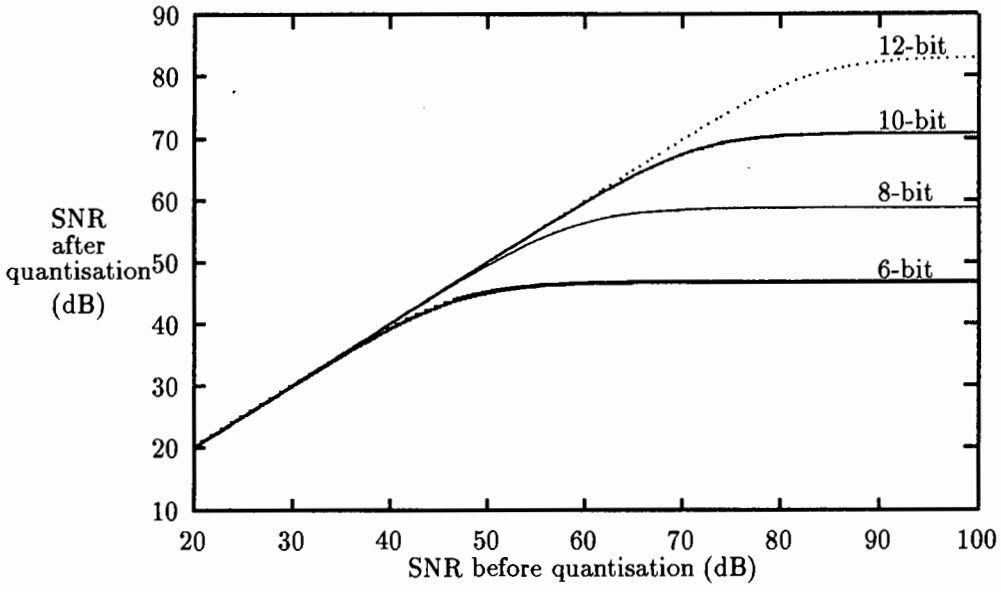


Figure 3.8: The effect of quantisation on the SNR of a sequence.

3.7 Orbital Mechanics

The coordinates of the sensor's perspective centre are constrained by the orbit of the satellite. This section describes the orbital mechanics of a satellite in circular, sun-synchronous, near-polar orbit. It is the orbit used by satellites that carry high resolution imaging sensors (see Section 4.2). The description is based on the work of Duck and King [16].

The assumption of a circular orbit can never be exactly satisfied. A circular orbit is a special case of an elliptical orbit with eccentricity equal to zero. A discussion of elliptical orbits can be found in [16].

Circular orbits will be discussed first in Section 3.7.1. The concept of orbit precession will be introduced in Section 3.7.2. The orbit precession and inclination required for a sun-synchronous orbit will also be discussed. The concept of a nodal day, which is used to determine the repeat pattern of the satellite, will be discussed in Section 3.7.3. The nadir trace pattern of circular orbits will be discussed in Section 3.7.4. Repeating coverage patterns will also be introduced.

3.7.1 Circular Orbits

A circular orbit is specified most simply in polar coordinates (r_s, θ_s):

$$r_s = r_\oplus + h_s \tag{3.54}$$

$$\theta_s = \dot{\theta}_s t \tag{3.55}$$

where

$r_\oplus$ = radius of the earth (6370 km),

h_s = orbital altitude,

$\dot{\theta}_s$ = angular velocity of satellite, and

t = time.

The angular velocity of the satellite $\dot{\theta}_s$ is given by

$$\dot{\theta}_s = \sqrt{\frac{\mu}{r_s^3}} \tag{3.56}$$

where μ is Earth's gravitational constant (3.98601×10^{14} m^3s^{-2}) and r_s is given above.

The orbital period is now given by

$$T_s = \frac{2\pi}{\dot{\theta}_s} = 2\pi\sqrt{\frac{r_s^3}{\mu}}. \tag{3.57}$$

The speed of the satellite track on the surface of the Earth v_s is given by

$$v_s = \dot{\theta}_s r_\oplus. \tag{3.58}$$

3.7.2 Orbit Precession

The precession of an orbit is the rotation of the orbital plane about the polar axis of the earth. This is primarily due to the Earth's oblateness. It is described as the rate of change of the nodal longitude Ω, the orientation of the node line (the intersection of the orbital and equatorial planes) relative to the vernal equinox.

It is given by [16]

$$\dot{\Omega} = -\frac{3}{2}\mathrm{J}_2 r_\oplus^2 \sqrt{\mu}(r_\oplus + h_s)^{-\frac{7}{2}} \cos i \tag{3.59}$$

where J_2 is the dimensionless number 0.00108263 derived from the Earth's oblateness and i is the orbital inclination angle, the angle between the orbital and equatorial planes.

To maintain a sun-synchronous orbit the node-to-sun angle must remain constant. This implies that the nodal-rate relative to the sun, $(\dot{\Omega})_\odot$, must equal zero to maintain a sun-synchronous orbit. The sun-relative scale $(\dot{\Omega})_\odot$ differs from the inertial[2] $\dot{\Omega}$ scale by one rev/year, corresponding to the one rev/year apparent motion of the sun.

Hence, for a sun-synchronous orbit $\dot{\Omega}$ must equal 1 rev/year or 1.991×10^{-7}rad/seconds. Now for a given orbital altitude h_s, the orbital inclination required for a sun-synchronous orbit is given by

$$i = \arccos \frac{1.991 \times 10^{-7}}{-\frac{3}{2}\mathrm{J}_2 r_\oplus^2 \sqrt{\mu}(r_\oplus + h_s)^{-\frac{7}{2}}}. \tag{3.60}$$

[2]Relative to space-fixed reference frame (e.g. star background).

3.7.3 Nodal Day

The nodal day, D_n, is the Earth's rotation period relative to the orbital plane or node. The concept of the nodal day is used to determine the repeat pattern of the satellite and the temporal resolution of the sensor (see Section 3.3.4). The rotational rate of the earth relative to the orbital plane is given by

$$\dot{\Lambda}_n = \dot{\Lambda} - \dot{\Omega} \tag{3.61}$$

where $\dot{\Lambda}$ is the inertial earth rotation rate and $\dot{\Omega}$ is the orbital precession rate (see section 3.7.2). The nodal day is now given by

$$D_n = \frac{2\pi}{\dot{\Lambda}_n} = \frac{2\pi}{\dot{\Lambda} - \dot{\Omega}}. \tag{3.62}$$

The inertial earth rotational rate can be rewritten in terms of the sidereal or inertial day, $D_s = 86\ 164.09$ seconds, as

$$D_s = \frac{2\pi}{\dot{\Lambda}}. \tag{3.63}$$

The solar time change per nodal day δt is the difference between the nodal day D_n and the solar day D, i.e.

$$\delta t = D_n - D. \tag{3.64}$$

It is zero for a sun-synchronous orbit.

3.7.4 Nadir Trace Pattern

The nadir trace pattern of satellites in circular orbits will be discussed in this section. The concept of the pattern repetition will be introduced.

The nadir trace is generated by the combined action of satellite orbital motion and the Earth rotation relative to the orbit plane. The nadir trace is specified by the geodetic coordinates (longitude λ and latitude ϕ) of any point $S'(t)$ on the trace. The coordinates are given by [16]:

$$\lambda_s = \tan^{-1}(\cos i \tan \theta_s) + \dot{\Lambda}_n t \tag{3.65}$$

$$\phi_s = \sin^{-1}(\sin i \sin \theta_s) \tag{3.66}$$

where i is the orbital inclination, $\theta_s = \dot{\theta}_s t$ is the angular position of the satellite and $\dot{\Lambda}_n$ the relative rotational rate.

The relative rotation causes a steady westward creep of the trace that is proportional to the ratio of the angular rates:

$$Q = \frac{\dot{\theta}_s}{\dot{\Lambda}_n}. \tag{3.67}$$

The rotation produced creep accumulated in one satellite revolution is

$$S = \frac{360}{Q} \qquad \text{degrees longitude.} \tag{3.68}$$

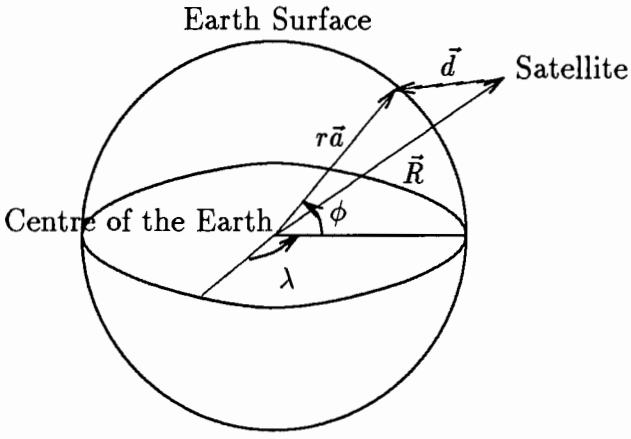


Figure 3.9: Geometrical configuration of the satellite and the Earth.

S is often called the "step", i.e. the interval (longitude) between consecutive crossings (in the same direction) of the equator or other parallel latitude.

The rate ratio Q determines the development of the nadir trace pattern. If Q is an integer then the pattern is a one-day repeater. Multi-day repeating patterns are produced when an integral number (R) of steps occupy exactly the same angular interval as another integral number (N) of 360 degree earth rotations.

A desired repeating pattern specified by N and R can be converted into a specification for the orbital period:

$$T = D_n \frac{N}{R}. \tag{3.69}$$

3.8 A Model Satellite Sensor Geometry

This section describes the geometrical relationship between the Earth, a satellite and its sensor. The development is based on the work of Sawada *et al.* [53]. The relationship is solved to the yield a direction vector of the observed point, whose origin is at the Earth's centre. Finally, the direction vector is expressed using the geocentric coordinate system. This will allow the geometry of the problem to be investigated. The model is simplified to include only circular orbits.

The basic geometrical configuration of the satellite and the Earth is shown in Figure 3.9 and can be expressed by

$$\frac{\vec{R} - r\vec{a}}{|\vec{R} - r\vec{a}|} = -\vec{d} \tag{3.70}$$

where r, $\vec{a}$, $\vec{R}$ and $\vec{d}$ are the local Earth radius, the direction vector of the observed point, the radius vector of the satellite and the direction cosine of the sensor, respectively. From

Equation 3.70, the direction vector $\vec{a}$ is given by [53]

$$\vec{a} = \frac{1}{r}\left[\vec{R} - \vec{d}\left\{\vec{R}\cdot\vec{d} + \sqrt{r^2 - R^2 + (\vec{R}\cdot\vec{d})^2}\right\}\right] \tag{3.71}$$

where one possible root is selected from the quadratic equation. The relationship between the direction vector $\vec{a} = (a_x, a_y, a_z)$ and the geodetic coordinates of latitude ϕ and longitude λ is expressed by:

$$\phi = \sin^{-1} a_z \tag{3.72}$$

and

$$\lambda = \tan^{-1}\frac{a_y}{a_x}. \tag{3.73}$$

The satellite orbit vector, $\vec{R}$, can be simply expressed as

$$\vec{R} = \begin{pmatrix} 0 \\ 0 \\ -(r+h) \end{pmatrix} \tag{3.74}$$

in the *orbital coordinate system*. It is a right-hand coordinate system with its origin at the centre of mass of the satellite. The x-axis is in the direction of the satellite flight direction and z-axis is in the direction of the Earth's centre. The direction cosine, $\vec{d}$, can be expressed as

$$\vec{d} = \frac{1}{\sqrt{1 + \tan^2\alpha_s(I) + \tan^2\beta_s(I)}} \begin{pmatrix} \tan\alpha_s(I) \\ \tan\beta_s(I) \\ 1 \end{pmatrix} \tag{3.75}$$

where $(\alpha_s(I), \beta_s(I))$ is the line-of-sight angle of pixel I in the *sensor coordinate system*. It is a right-hand coordinate system with its origin at the perspective centre of the sensor and xy plane parallel to the image plane. The optical axis is normal to the image plane. The perspective centre of the objective lens lies on the optical axis. Its distance from the image plane is the focal length, f. The x-axis is in the direction of flight of the satellite.

In order to obtain the components of $\vec{a}$ in the geocentric coordinate system, it is necessary to transform the components of vectors $\vec{d}$ and $\vec{R}$ to the geocentric coordinate system. Following Sawada *et al.* [53] this transformation is sequentially performed by matrix multiplication as follows:

$$T_s^g = T_i^g T_o^i T_b^o T_s^b \tag{3.76}$$

where subscripts and superscripts g, i, o, b and s indicate the geocentric, inertial, orbital, body and scanner coordinate systems, respectively. The coordinate transformation is taken from the lower coordinate system to the upper one. The transformation from the orbital to geocentric system is

$$T_o^g = T_i^g T_o^i. \tag{3.77}$$

Each of these transformations and the coordinate systems will be described in turn.

The transformation from the inertial coordinate system to the geocentric coordinate system is given by

$$T_i^g = \begin{pmatrix} \cos\gamma_e & \sin\gamma_e & 0 \\ -\sin\gamma_e & \cos\gamma_e & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.78}$$

where

$$\gamma_e = \dot{\Lambda}_n t \tag{3.79}$$

and $\dot{\Lambda}_n$ is the angular velocity of the Earth relative to the orbital plane (see Equation 3.61). Note that $\dot{\Lambda}_n$ is fixed for a sun-synchronous orbit.

The transformation from the orbital coordinate system to the inertial coordinate system is given by

$$T_o^i = \begin{pmatrix} -\sin\theta_s & 0 & -\cos\theta_s \\ \cos\theta_s\cos i & \sin i & -\sin\theta_s\cos i \\ \cos\theta_s\sin i & -\cos i & -\sin\theta_s\sin i \end{pmatrix} \tag{3.80}$$

where θ_s is given by Equation 3.55 and i is the orbital inclination (see Section 3.7.2).

The satellite body coordinate system is defined in terms of the orbital coordinate system; its x, y and z directions are obtained from the orientation angles of roll, pitch and yaw relative to the orbital coordinate system. The transformation from the orbit coordinate system to the satellite body coordinate system is given by

$$T_o^b = \begin{pmatrix} \cos\beta\sin\gamma & \sin\alpha\sin\beta\cos\gamma + \cos\alpha\sin\gamma & -\cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma \\ -\cos\beta\sin\gamma & -\sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \cos\alpha\sin\beta\sin\gamma + \sin\alpha\cos\gamma \\ \sin\beta & -\sin\alpha\cos\beta & \cos\alpha\cos\beta \end{pmatrix} \tag{3.81}$$

where α, β and γ are the roll, pitch and yaw angles, respectively. The transformation from the satellite body coordinate system to orbit coordinate system T_b^o is the transpose of T_o^b.

The transformation of the satellite body coordinate system to the sensor coordinate system is a rotation of azimuth angle, ψ_b, and zenith angle, θ_b, and is given by

$$T_b^s = \begin{pmatrix} \cos\psi_b & \sin\psi_b & 0 \\ -\sin\psi_b\cos\theta_b & \cos\psi_b\cos\theta_b & \sin\theta_b \\ \sin\psi_b\sin\theta_b & -\cos\psi_b\sin\theta_b & \cos\theta_b \end{pmatrix}. \tag{3.82}$$

The transformation from the sensor coordinate system to the satellite body coordinate system T_s^b is given by the transpose T_b^s.

The model of satellite sensor geometry will be used in Sections 5.6 and 6.3.3 as part of the investigation into the stereo geometry (see Section 5.2).

Chapter 4

Remote Sensing Satellite Missions

4.1 Introduction

An overview of the design and principles of operation of high resolution multi-spectral imaging sensors carried by remote sensing satellites is presented in this section. Three imaging sensors, and the satellites that carry them, will be described. They are the *Multi-spectral Scanner* (MSS) that was on board all five Landsat satellites, the *Thematic Mapper* (TM) that was on board Landsat-4 and -5, and the *High Resolution Visible* (HRV) sensor that was on board SPOT-1 (Satellite Pour l'Observation da la Terre) and -2 [41,57].

The principles of operation of these imaging sensors will be discussed in Section 4.2. The Landsat satellite series will be described in Section 4.3, and SPOT satellite series in Section 4.4.

4.2 Principles of Operation

The principles of operation of the satellite based imaging sensors is presented in this section. These principles form the basis for the concept design outlined in Section 5.2.

Satellite based, high resolution imaging sensors generally use a scanning system to build up a two-dimensional image, i.e. they operate as so called "pushbroom" scanners. The primary motivation for this is the reduction in the number of detectors required. It does, however, result in an increase in complexity.

Modern imaging sensors use line-array charge coupled device (CCD) detectors, which have no moving mechanical parts, in place of mechanical scanning systems (e.g. scanning mirrors, disk choppers, or mechanical modulators). This is due to the advantages that CCD detectors have [7]:

- The "exposure" time for each ground point images is automatically maximised, and
- Excellent photogrammetric quality along the linescan axis is ensured.

As stated in Section 3.7, the orbit of satellites, which carry high resolution imaging sensors, are generally near-polar, circular, repetitive and sun-synchronous. This is motivated by the requirement to maintain a constant image scale and illumination geometry. The near-polar nature of the orbit also ensures maximum coverage of the surface of the Earth. The constant image scale is provided by the circular nature of the orbit and to maintain the illumination geometry of the scenes as constant as possible the orbit is sun-synchronous, i.e. the satellite passes over the area being imaged at approximately the same local sun time. The repetitive nature of the orbit permits repetitive viewing of a particular geographic area at fixed, specified intervals.

4.3 The Landsat Satellites

The first Landsat satellite, then called ERTS-1 (Earth Resources Technology Satellite 1) was launched on July 23rd, 1972. It was the first satellite to carry imaging sensors designed specifically to provide for those scientists working in the area of terrestrial resources. The orbit was chosen to be near-polar, circular, repetitive and sun-synchronous. The near-polar orbit gave it coverage of most of the land and sea surface of the Earth (between 82°N and 82°S) from a nominal altitude of 913 km. The orbital pattern gave it a repeat cycle of 18 days. The equatorial crossing time on the descending (N-S) node of its orbit was approximately 0930 local sun time.

The next two Landsat satellites, numbers 2 and 3, were launched into an orbit similar to Landsat-1, in 1975 and 1978 respectively, and carried similar MSS sensors. Landsat-4 was launched in July 1982, into a near-polar circular repetitive sun-synchronous orbit with a repeat period of 16 days at a nominal altitude of 705 km. The equatorial crossing time on the descending (N-S) node of its orbit was approximately 0945 local sun time. It also carried the new TM sensor. Problems with Landsat-4's power supply led to an earlier launch of Landsat-5 in March 1984, into an orbit similar to Landsat-4.

4.3.1 The Multi-spectral Scanner

The Multi-spectral Scanner was carried by all five Landsat satellites. It is a line scanner that sweeps the ground below the satellite and builds up an image of the Earth's surface as the satellite's forward motion moves it along. The swath width was 185 km. The scanner has four spectral bands (Table 4.1). A fifth band, operating in the thermal infrared region, was included in the MSS carried by Landsat-3; however, it failed soon after launch.

The MSS used a 22.9 cm aperture with a 8.9 cm aperture secondary mirror. The focal length was 82.3 cm and the F/number 3.6. The IFOV was nominally 79 m. The radiometric resolution was 6 bits.

Channel[1]	Spectral Band (μm)
4	0.5-0.6
5	0.6-0.7
6	0.7-0.8
7	0.8-1.1

Note:

1. The channels on Landsat-4 and -5 MSS are numbered 1-4.

Table 4.1: Multi-spectral Scanner Spectral Bands.

Channel	Spectral Band (μm)
1	0.45-0.52
2	0.52-0.60
3	0.63-0.69
4	0.75-0.90
5	1.55-1.75
6	10.4-12.5
7	2.08-2.35

Table 4.2: Thematic Mapper Spectral Bands.

4.3.2 The Thematic Mapper

The Thematic Mapper (TM) sensor carried by Landsat-4 and -5 is also a line scanner. It has 7 spectral bands (Table 4.2) and the IFOV is 30 m except band 6 (the thermal infrared band), where it is 120 m. The radiometric resolution is 8 bits. The diameter of the lens aperture is 41 cm and the focal length 2.28 m, giving a F/number of 5.6.

4.4 SPOT

This section gives an overview of the design on the SPOT series of high resolution imaging remote sensing satellites. SPOT 1 was successfully launched on February 22nd, 1986 by the ARIANE launcher. SPOT 2 was successfully launched on January 22, 1990 [2]. Further satellites in the series are being build and developed [2]. Chevrel *et al.* [7] presented a discussion on the design motivation of the SPOT satellite mission. The information for this section was largely obtained from this document.

Its orbit was near-polar, circular, repetitive and sun-synchronous with a nominal height of 832 km. The long repetition cycle of 26 days was compensated to some extent by the pointable sensors. The equatorial crossing time of 1030 gives imagery of higher solar elevation angles

Channel	Spectral Band (μm)
1	0.50-0.59
2	0.61-0.68
3	0.79-0.89

Table 4.3: High Resolution Visible Sensor Spectral Bands.

than Landsat. The pointable sensors also provide a means to obtain imagery of the target area from different view angles on *successive* days. Stereoscopic analysis can be performed on these different views to obtain elevation data (see Section 2.2).

4.4.1 The High Resolution Visible Sensor

Due to the high ground resolution required by the HRV, mechanical scanning systems, like those used in the Landsat sensors, were not considered. Line-array charge coupled device (CCD) detectors were used.

The High Resolution Visible (HRV) sensor carried by SPOT has two modes of operations, a multi-spectral mode and a panchromatic mode. In the panchromatic mode a single wide spectral band (510 nm to 730 nm) is used with an IFOV of 10 m. In the multi-spectral mode an IFOV of 20 m is used and observations are made in three bands (Table 4.3). The radiometric resolution is 6 bits in panchromatic mode and 8 bits in multi-spectral mode.

The optical system of the HRV consisted of the following elements [7]:

- A front end mirror that can be rotated about the roll axis, allowing selection of target scenes,
- A folded pseudo-Schmidt telescope, with F/number 3.5 and focal length 1082 mm,
- Three dichroic prisms located at the focus for spectral separation, and
- Four beam splitting prisms, one for each band, providing precise optical butting for four CCD arrays into a single scan line.

Chapter 5

Modelling of Imaging Sensor

5.1 Introduction

The design of the imaging sensor (Chapter 6) requires that the performance of the sensor be simulated. To achieve this, a model of the imaging sensor was developed. The development of the model is described in this chapter. The model is based on the concept design of the imaging sensor (Section 5.2) and theory presented in the previous chapters.

The model relates parameters of the sensor, the satellite and the environment to various output characteristics of the sensor. Environmental parameters include: the spectral irradiance incident on the Earth, the reflectivity of the Earth, the atmospheric transmission characteristics and other physical properties of the Earth. Sensor and satellite parameters include the orbit of the satellite, the focal length of the lens, the entrance pupil diameter, the obscuration of the lens, the spectral bands, the CCD characteristics, the offset of the sensor from the vertical and the A/D resolution. The output characteristics include: geometrical parameters (the distance from sensor to object and the base-to-height ratio), the MTF, the CCD output voltage, the signal-to-noise ratio, the spatial resolution, the ground track, the accuracy of stereo height measurements and the data rate.

The concept design of the imaging sensor is described in Section 5.2. It forms the basis for the relationships developed in the rest of the chapter. Section 5.3 deals with the geometric relationships between Earth and the imaging sensor from a static point of view. The estimation of the spatial resolution of the sensor is described in Section 5.4. Section 5.5 deals with estimating the output of the detector (a CCD) and the signal-to-noise ratio of the sensor. Section 5.6 deals with estimating the ground track (the ground in the field of view of the sensor), which is largely determined by the orbit of the satellite. This extends the static geometric model described in Section 5.3. Section 5.7 deals with the development of the relationships that govern the accuracy of the height measurements obtained from the stereo pairs produce by the sensor. Section 5.8 presents a relationship to estimate the data rate of the sensor. Finally, Section 5.9 deals with the implementation of the model, to enable simulations to be

performed.

5.2 Concept Design of the Imaging Sensor

This section presents the concept design of the imaging sensor and its motivation. The concept design defines the basic method of operation and technologies to used. An overview of the design concept was presented in Section 1.2.

The imaging sensor can be divided into three component parts: (a) The optical system collects the radiant energy, splits it into the various spectral bands and focuses it onto the detectors; (b) The detectors, one for each spectral band, convert the incident radiant energy into electrical signals; and (c) The analogue-to-digital converters (A/D) convert the electrical signals to digital data.

The sensor is required to produce a two-dimensional image of the Earth. This will be achieved by using a one-dimensional detector array and the forward motion of the satellite, i.e. by operating the sensor as a "pushbroom" scanner (see Section 4.2). The same procedure is used by the HRV of SPOT, and TM and MSS of Landsat (see Section 4.2).

The proposed optical system will use either a refractive or reflective lens system to focus the light. Dichroic prisms will be used to split the incident radiation into the required spectral bands. The choice of lens system will depend on the focal length required. At short focal lengths the refractive lens will be smaller and lighter, but at long focal lengths the reflective lens will be smaller and lighter. The reflective lens type has the additional disadvantage of a central obscuration.

Line-array (one-dimensional), charge coupled devices (CCDs) (see Section 3.5) will be used as the detectors, one per spectral band. They are also used by the HRV of SPOT-1 and -2. Furthermore, one-dimensional CCDs are available with more than 3000 detector elements. One disadvantage is that the most commonly available CCDs do not operate well beyond 0.9μm. This will limit the choice of spectral bands.

The imaging sensor will produce a stereo image pair in a single pass. This is different to the SPOT satellites that require at least two passes. The advantage of using a single pass is that the time between the taking of the images is kept to a minimum. Figure 5.1 illustrates how the single pass stereo imaging will be achieved. The satellite "looks" forward when it is between A and B, to image the Earth's surface between E and F. The same section of the Earth's surface is imaged while the satellite is between C and D, but the satellite "looks" backward. Therefore, a stereo image pair, two images of the same scene taken from different orientations, is obtained.

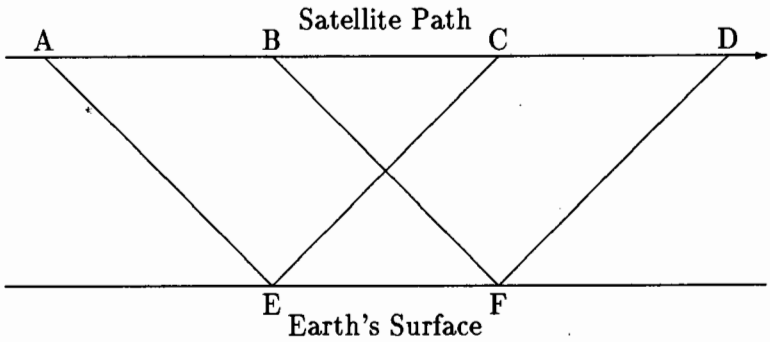


Figure 5.1: The operation of single pass stereo imaging. Note that a flat Earth model is used to simplify the diagram.

5.3 Geometric Relationships

The relationships governing the viewing geometry are presented in this section. The parameters involved in these relationships are the base-to-height (B/H) ratio, the slant range (the distance from the object to the sensor), the sensor offset from the vertical, the grazing angle and the orbital altitude.

To simplify the model, it is assumed that the Earth is a perfect sphere with a radius of 6370 km and that the orbit of the satellite is perfectly circular.

Under these assumptions the viewing geometry is depicted in Figure 5.2, where $r_\oplus$ is the radius of the Earth (6370 km), h_s is the orbital altitude of the satellite, θ_o is the offset of the sensor from the vertical, B/H is the base-to-height ratio, l_s is the slant range and α_g is the grazing angle.

Using Figure 5.2, the following relationships can be derived. The base-to-height is given by:

$$B/H = 2\tan\alpha_g. \tag{5.1}$$

The slant range, in terms of θ_o, is given by:

$$l_s = (r_\oplus + h_s)\cos\theta_o - \sqrt{(r_\oplus + h_s)^2\cos^2\theta_o - h_s^2 - 2r_\oplus h_s}, \tag{5.2}$$

or, in terms of α_g, it is given by:

$$l_s = -r_\oplus\cos\alpha_g + \sqrt{r_\oplus^2\cos^2\alpha_g + 2r_\oplus h_s + h_s^2}. \tag{5.3}$$

The viewing angles α_h and θ_o are related by:

$$\sin\theta_o = \sin\alpha_g\frac{r_\oplus}{r_\oplus + h_s} \tag{5.4}$$

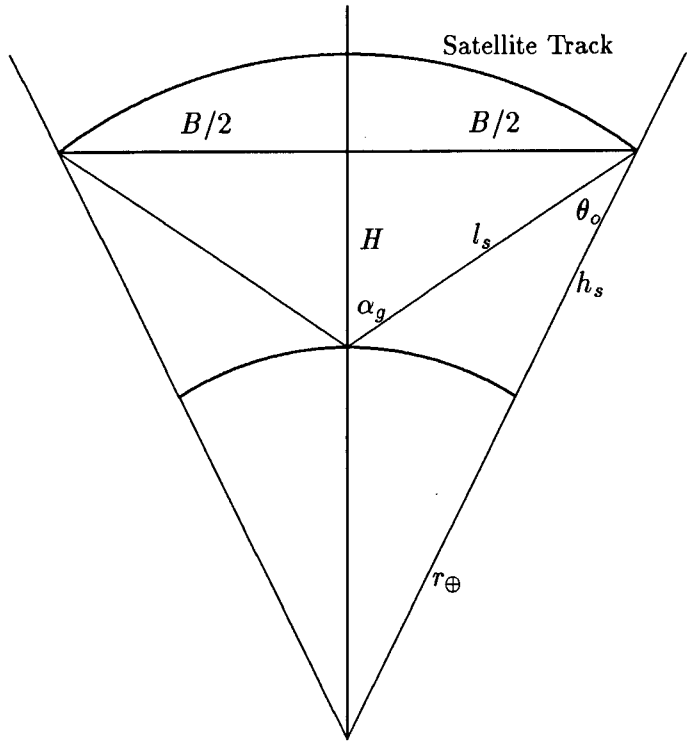


Figure 5.2: Viewing geometry of a satellite based imaging sensor (from [63]).

and θ_o is related to the base-to-height ratio by:

$$\sin\theta_o = \frac{B/H}{\sqrt{4+(B/H)^2}} \frac{r_\oplus}{r_\oplus + h_s}. \tag{5.5}$$

These relationships will be used to estimate the spatial resolution of the sensor in the next section.

5.4 Spatial Resolution

The term, spatial resolution, was discussed in Section 3.3.1. Estimates of the IFOV (Section 5.4.1) and MTF (Section 5.4.2) will be derived in this section. The IFOV provides a simple estimate of spatial resolution while the MTF provides a more complete description. The assumptions stated in Section 5.3 are again used.

5.4.1 The IFOV

By substituting Equation 3.7 into Equation 3.14 then IFOV is now given by

$$\text{IFOV} = l_s d(\frac{1}{f} - \frac{1}{l_s}) \tag{5.6}$$

where l_s is the slant range (see Section 5.3), f is the focal length of the lens and d is the detector size. This equation can now be simplified by observing that the slant range is much greater than the focal length[1] and, hence, $\frac{1}{f} - \frac{1}{l_s} \approx \frac{1}{f}$. Now Equation 5.6 becomes

$$\text{IFOV} = \frac{d}{f} l_s. \tag{5.7}$$

Thus the IFOV is directly proportional to the slant range and detector size, and inversely proportional to the focal length.

5.4.2 The MTF

The MTF is a more complete description of the spatial resolution. The geometrical factors that determine the MTF of the imaging sensor will be modelled in this section.

The MTF of the lens (optical system) will be discussed first. The MTF of a perfect, diffraction-limited circular lens is a function of the wavelength and the F/number of the lens for incoherent radiation. The MTF is given by [51]

$$\text{MTF}_L(\nu_i) = \frac{2}{\pi}\left[\cos^{-1}(\lambda N \nu_i) - \lambda N \nu_i \sqrt{1 - (\lambda N \nu_i)^2}\right] \tag{5.8}$$

where ν_i is the spatial frequency with reference to the image plane, N is the F/number of the lens (see Equation 3.9) and λ is the wavelength of the radiation. The MTF of a real lens will be less than that of the diffraction-limited case.

The effect of a central obscuration of 55 percent on the MTF of a lens is illustrated in Figure 5.3. It shows the MTF of a diffraction-limited lens with and without a central obscuration. In both cases the focal length is 400 mm, the entrance pupil diameter is 10 cm and the wavelength is 700 nm.

The component of the MTF due to the image smear, as a result of constant forward motion of the satellite, MTF_I is given by

$$\text{MTF}_I(\nu_o) = \frac{\sin(\pi \nu_o \gamma)}{\pi \nu_o \gamma} \tag{5.9}$$

where ν_o is the spatial frequency with reference to the object plane and γ is the distance the satellite track moved during the integration period. The distance moved γ is given by

$$\gamma = T_i v_s \tag{5.10}$$

where T_i is the integration period and v_s is the speed of the satellite track on the surface of the Earth (see Equation 3.58).

The effect of the CCD on the MTF of the imaging sensor was discussed in Section 3.5.3. Incorporation of Equations 5.8 and 5.9 into Equations 3.17 and 3.40 will allow the MTF of the sensor to be estimated. This will be done in Section 6.3.2.

[1] The slant range is a minimum of about 800 km and the focal length a maximum of 1.1 m

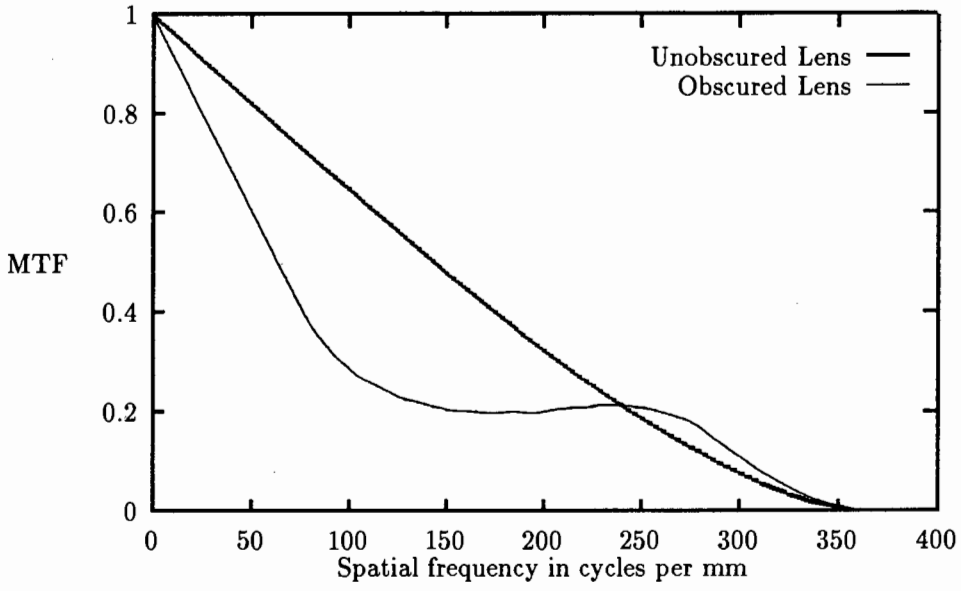


Figure 5.3: The MTF of a diffraction-limited lens with and without a central obscuration. The focal length is 400 mm. The entrance pupil diameter is 10 cm. The wavelength is 700 nm. The central obscuration was 55 percent of the diameter. The data for the obscured case was obtained from [22].

The spatial frequency, with reference to the image plane, ν_i and, with reference to the object plane, ν_o are related by the lateral magnification of the lens m:

$$\frac{\nu_i}{\nu_o} = \frac{1}{m}. \tag{5.11}$$

5.5 Detector Output and Sensor Signal-to-Noise Ratio

This section deals with the radiometric modelling of the imaging sensor. A relationship is developed to estimate the output voltage of the CCD. This is then used to develop relationships governing the signal-to-noise ratio of the sensor. The digital output of the sensor can also be estimated.

Under the assumptions given in Section 3.2.6 the image-plane irradiance as a function of wavelength is given by Equation 3.13. The exposure of the CCD S is now given by

$$S(\lambda) = E_i(\lambda)T_i, \tag{5.12}$$

where T_i is the integration period of the CCD. The output voltage of the CCD is given by the integral of product of the CCD exposure and sensitivity over the spectral band of interest (see Section 3.5.1 and Appendix B), i.e.

$$V_{\Delta\lambda} = \int_{\lambda_1}^{\lambda_2} S(\lambda)R(\lambda)d\lambda$$

$$= \frac{(1-A)^2 T_i}{4N^2(1-m)^2} \int_{\lambda_1}^{\lambda_2} E_\lambda \rho(\lambda) \tau_u(\lambda) \tau_o(\lambda) R(\lambda) d\lambda, \tag{5.13}$$

where

T_i = integration period,

$R(\lambda)$ = sensitivity of CCD (given in $\mathrm{V}/(\mu\mathrm{J}/\mathrm{cm}^2)$),

$\lambda_1 - \lambda_2$ is the spectral band of interest,

A = ratio of the diameter of the central obscuration to that of the primary mirror,

N = F/number of the lens,

m = lateral magnification (see Equation 3.8),

E_λ = object plane irradiance,

$\rho(\lambda)$ = spectral reflectance of the ground object,

$\tau_u(\lambda)$ = spectral transmittance of the atmosphere along the atmospheric path between the sensor and the ground object, and

$\tau_o(\lambda)$ = spectral transmittance of the optical system.

The signal-to-noise ratio of the system cannot be estimated without making assumptions regarding the nature of the signal. For the purposes of estimating the signal-to-noise ratio, the signal is assumed to be sinusoidal. As stated in Section 3.5.1, there is a limit to the number of electrons that can be stored in a potential well and, hence, there is a maximum or saturation output voltage.

If V_{sat} is the saturation voltage of the CCD, the peak sinusoidal output signal is

$$V_{out}(s) = V_{sat}\left(\frac{1}{2} + \frac{1}{2}\sin(\nu s)\right), \tag{5.14}$$

where s is the spatial variable and ν is the spatial frequency. The signal variance σ_s^2 is then given by:

$$\sigma_s^2 = \frac{V_{sat}^2}{8}. \tag{5.15}$$

The signal-to-noise ratio is now given by:

$$\left(\frac{S}{N}\right) = \frac{\sigma_s^2}{\sigma_n^2} \tag{5.16}$$

where σ_n^2 is the noise power or variance.

The model of the noise power includes two noise sources. The CCD readout and shot noise (see Section 3.5.2), and quantisation noise (see Section 3.6). The variance of the quantisation

SNR Before Quantisation	Number of Bits	SNR After Quantisation	Loss in SNR due to Quantisation
51.0 dB	6	37.7 dB	13.3 dB
51.0 dB	8	47.4 dB	3.6 dB
51.0 dB	10	50.6 dB	0.4 dB
51.0 dB	12	50.9 dB	0.1 dB
40.6 dB	6	36.0 dB	4.6 dB
40.6 dB	8	40.1 dB	0.5 dB
40.6 dB	10	40.5 dB	0.1 dB
40.6 dB	12	40.6 dB	0.0 dB

Table 5.1: The signal-to-noise ratio as a function the number of bits used for quantisation.

noise is given by Equation 3.46 and the total noise power of the CCD by Equation 3.25. Hence, noise power is given by:

$$\sigma_n^2 = n_o^2 + n_e + \sigma_e^2 \tag{5.17}$$

where σ_n^2 is the total noise power, n_o^2 is readout noise power, n_e is shot noise power, it is equal to the expected number of electrons in the potential well, and σ_e^2 is the quantisation noise power.

The maximum number of electrons, which can be stored in a potential well, is typically 10^5 and the readout noise n_o is typically 100 electrons (see Section 3.5.1). The shot noise component is dependent on the expected signal level. Under worst case conditions it is equal to the maximum output signal and under best case conditions it is zero. This gives a signal-to-noise ratio of 40.6 dB and 51.0 dB, respectively. The quantisation noise power can be rewritten as (see Equations 3.42 and 3.46):

$$\sigma_e^2 = \frac{1}{12}\frac{V_{sat}^2}{4^B} \tag{5.18}$$

where B is the number of bits used to quantise the signal, by assuming the full scale amplitude of the quantiser is given by V_{sat}. The expected signal-to-noise ratios when using 6, 8, 10 and 12 bit quantisers is illustrated in Figure 5.1.

The choice of A/D resolution (number of bits used for quantisation) will be discussed in Section 6.4.2. It should be noted that if the full dynamic range of CCD is not used then the signal-to-noise ratio will be reduced.

5.6 Ground Track

The estimation of the ground track (the position of points in the field of view of the sensor) will be discussed in this section. It is assumed that the satellite is in a circular, sun-synchronous,

near-polar orbit (see Section 3.7). The ground track can be used for estimating the temporal resolution of the sensor and for analysing the stereo geometry (see Section 6.3.3).

The temporal resolution is the minimum number of days between which the same point of the Earth can be imaged. Under optimal conditions the number passes required to image the entire equator (assuming no overlap) is given by the circumference of the Earth at the equator divided by the swath width of the satellite. The time required can then be obtained by multiplying the number of passes by the orbital period. This can be determined from the analysis presented in Section 3.7.4.

Analysis of the stereo geometry requires that off vertical "look" directions be used. The model of satellite sensor geometry of Sawada *et al.* [53] (see Section 3.8) was implemented to allow the ground track to be determined under these conditions.

5.7 Stereo Height Accuracy

In this section a relationship is developed to estimate the height accuracy that can be achieved. The developments in this section are based on the work of Welch [61,62].

The difference in x parallax (Δp) can be used to estimate the height difference (Δh) between two objects. This is illustrated in Figure 5.4. The height difference measurements are closely approximated by [61]:

$$\Delta h = \frac{\Delta p}{2 \tan \alpha_g} SF \tag{5.19}$$

where $\Delta p = x_1 - x_2$ is the difference in x parallax, α_g is the grazing angle (see Figure 5.2), and SF is the image scale factor (it is the inverse of the lateral magnification given by Equation 3.8).

Since, the heights are a function of x parallax, the measurement error ($\sigma \Delta p$) is given by [62]:

$$\sigma \Delta p = \sqrt{\sigma_t^2 + \sigma_b^2} \tag{5.20}$$

where σ_t and σ_b are the error in the top and bottom parallax measurements, respectively.

The height measurement error $\sigma \Delta h$ (in pixels) can be related to the measurement error in the x parallax (in pixels) by:

$$\sigma \Delta h = \frac{\sigma \Delta p}{2 \tan \alpha_g}, \tag{5.21}$$

or in terms of the base-to-height ratio (see Equation 5.1):

$$\sigma \Delta h = \frac{\sigma \Delta p}{B/H}. \tag{5.22}$$

The relationship between $\sigma \Delta h$ and the base-to-height ratio (Equation 5.22) is depicted in Figure 5.5 for $\sigma \Delta p$ values of 1.4, 0.7, 0.35 and 0.14 pixels. It should be noted that the B/H ratio cannot be increased without bound as this will complicate the matching process and increase $\sigma \Delta p$ (see Section 2.2.5).

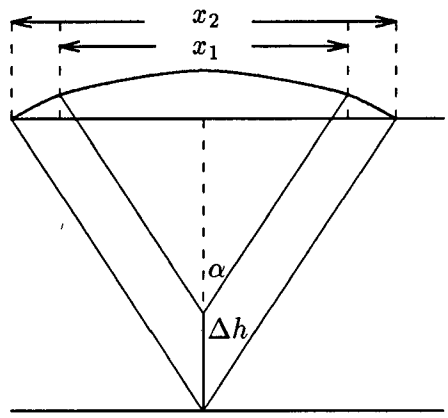


Figure 5.4: Illustration of the stereo geometry. It shows the relationship between the height difference and the x parallax difference.

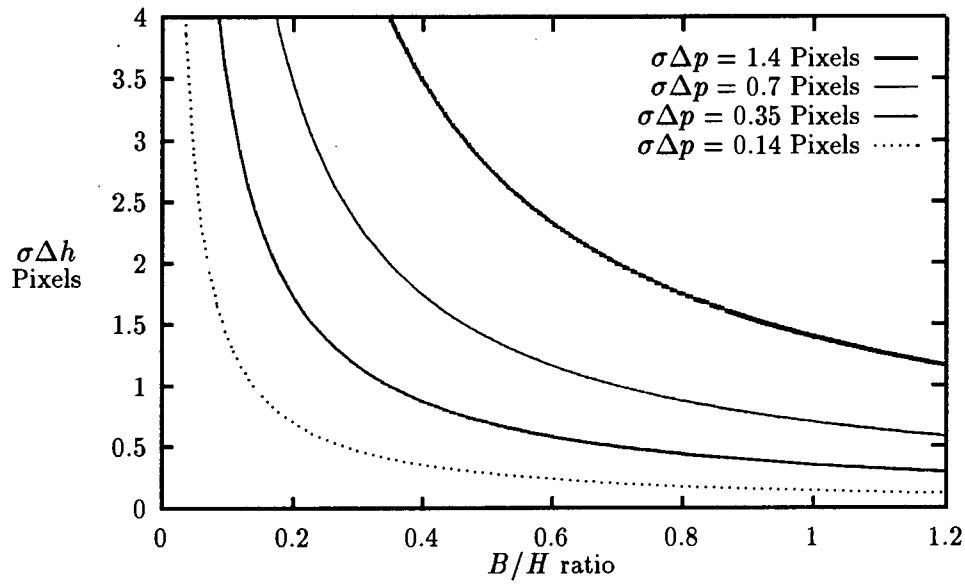


Figure 5.5: Height difference error in pixels as a function of base-to-height ratio, for $\sigma\Delta p$ values of 1.4, 0.7, 0.35 and 0.14 pixels, respectively.

5.8 Sensor Data Rate

A relationship is presented in this section to estimate the rate at which the sensor produces digital data. This will determine the bandwidth required for the communication channel which transports the data back to Earth.

The data rate in bits per second is given by:

$$B = \frac{N_s b N_d}{T_s} \tag{5.23}$$

where N_s is the number of spectral bands, b is the number of bits used to quantise the signal, N_d is the number of detector elements and T_s is the period between scan line samples.

5.9 Implementation of the Model

Issues related to the implementation of the model are discussed in this section. This includes estimates of environmental parameters and sensor parameters.

5.9.1 Environmental Parameters

The environmental parameters that have to be estimated are: (a) the spectral solar irradiance, (b) the transmittance of the atmosphere, and (c) the reflectivity of the Earth's surface.

The estimate of spectral solar irradiance incident of the Earth's surface E_λ was obtained from Coulson [10, pp. 313-317] and is tabulated in Appendix A for the spectral band, 0.4 μm to 0.9 μm.

The estimate of the atmospheric transmittance $\tau_u(\lambda)$ was obtained from Figure 3.3. Over the spectral range of interest 0.4 μm to 0.9 μm τ_u is assumed to be a constant 0.9.

The reflectivity of the Earth's surface $\rho(\lambda)$ is assumed to be constant over the spectral band of interest.

5.9.2 Sensor Parameters

The estimation of the transmittance of the optical system and the percentage obscuration of the lens is discussed in this section.

The estimate of the transmittance of the optical system is based on assumptions of the design of the optical system. The proposed optical system to be used in SUNSAT will require 3 reflections and 12 refractions. The expected reflection of a mirror surface is 88 % and transmission of a refractive surface is 99 % [23]. It is proposed that dichroic prisms be used for spectral separation. The expected transmission in the pass band will be about 95 % [44]. Thus, τ_o is assumed to be a constant 0.57.

Focal Length (mm)	Percentage Obscuration
200	55
280	55
400	55
560	50
800	42
1100	37

Table 5.2: Percentage obscuration as a function of focal length. The data was obtained from [22].

The percentage obscuration of the proposed lens is a function of the focal length [22]. The percentage obscuration of the proposed lens for focal length of 200, 280, 400, 560, 800 and 1100 mm is given in Table 5.2.

Chapter 6

Design of Imaging Sensor

6.1 Introduction

The design of the imaging sensor is discussed in this chapter. It is based on the model presented in Chapter 5. The design procedure was aimed at producing an imaging sensor that met the requirements stated in Section 1.2 and had the best possible performance.

This was achieved with a three-stage procedure. First, certain sensor and satellite parameters were fixed or limited to certain values (Section 6.2). This was done to reduce the scope of the design investigation and to ensure the design requirements were met. The second stage involved determining if it is possible to design such an imaging sensor (the feasibility investigation) (Section 6.3). Finally, the third stage involves the determination of the sensor specification (Section 6.4).

6.2 Sensor and Satellite Parameters

The sensor and satellite parameters that were fixed or limited to certain values are described in this section. The restriction of these parameters was motivated by the need to ensure that the design requirements (see Section 1.2) are met and to reduce the scope of the investigation.

The parameters affected are: (a) the entrance pupil diameter, (b) the number of spectral bands, (c) the choice of spectral bands, (d) the choice of CCD, (e) the integration period, (f) the base-to-height ratio, and (g) the satellite orbit. Each of these parameters will be discussed in turn.

The requirement of being able to fly the imaging sensor on a micro-satellite like SUNSAT places limits on the size and mass of the sensor. This can be translated into a limitation on the diameter of the entrance pupil D of 10 cm and a limit of three spectral bands in SUNSAT's case [44]. Hence, an entrance pupil diameter of 10 cm was used and number of spectral bands was limited to three.

The spectral range that these bands could occupy was limited to the range from 0.5 μm

Orbital Altitude	800 km
Orbital Inclination	98.6 degrees
Angular Velocity	1.038×10^{-3} rad/s
Linear Velocity	7 452 m/s
Ground track speed	6 621 m/s
Angular velocity of the Earth relative to the orbital plane	7.272×10^{-5} rad/s

Table 6.1: The orbital parameters used in simulations.

to 0.9 μm. This was motivated by the spectral characteristics of CCDs (see Section 5.2) and atmospheric scattering characteristics (see Section 3.2.4).

For the purposes of this investigation, bands 4, 5 and 6 of the MSS (see Table 4.1), bands 2, 3 and 4 of the TM (see Table 4.2), and bands 1, 2 and 3 of the HRV (see Table 4.3) were considered. This was motivated by the observation that potential users would be familiar with these bands.

The TC104 CCD produce by Texas Instruments was assumed to be used as the detector. It was selected to be used in the imaging sensor to be carried by SUNSAT [44]. Thus parameters such as CCD sensitivity, saturate voltage and noise power could be estimated. The data sheet of the TC104 CCD is given in Appendix B.

The length of the integration period of the CCD is limited by the requirement to have nominally the same pixel spacing in the in-track and cross-track directions. Thus, the maximum integration period is given by:

$$T_i = \frac{T_\beta}{m v_s} \tag{6.1}$$

where T_β is the detector pitch (distance between the centre of detector elements), m is the lateral magnification (see Equation 3.8) and v_s is the speed of the satellite track on the surface of the Earth (see Equation 3.58).

The base-to-height ratio was limited to the range from 0 to 1. This was motivated by: (a) The increased matching difficulty associated with high base-to-height ratios (see Section 2.2.5); and (b) The range of base-to-height ratios achievable by the SPOT satellites is in this range.

The orbit of the satellite is assumed to be circular, near-polar and sun-synchronous (see Sections 3.7 and 4.2). The nominal orbital altitude was fixed at 800 km (the proposed altitude of SUNSAT) for the investigation. The orbital parameters used in the simulations are shown in Table 6.1.

Surface Reflectivity	Required IFOV (m)								
	MSS Bands			TM Bands			HRV Bands		
	4	5	6	2	3	4	1	2	3
0.1	35.9	36.0	39.4	38.5	43.0	40.0	37.4	41.0	47.3
0.2	28.5	28.6	31.2	30.5	34.1	31.7	29.7	32.5	37.5
0.3	24.9	25.0	27.3	26.7	29.8	27.7	25.9	28.4	32.8
0.4	22.6	22.7	24.8	24.2	27.1	25.2	23.5	25.8	29.8
0.5	21.0	21.0	23.0	22.5	25.2	23.4	21.8	24.0	27.6
0.6	19.7	19.7	21.7	21.2	23.7	22.0	20.5	22.6	26.0
0.7	18.6	18.6	20.5	20.0	22.5	20.9	19.4	21.4	24.7
0.8	17.7	17.8	19.6	19.1	21.5	19.9	18.5	20.4	23.6
0.9	17.0	17.0	18.7	18.3	20.6	19.1	17.7	19.6	22.7
1.0	16.4	16.4	18.1	17.6	19.9	18.3	17.1	18.9	21.9

Table 6.2: IFOV required to saturate the TC104 CCD for various spectral bands and surface reflectivities.

6.3 Feasibility Investigation

The feasibility investigation is described in this section. The objective of the feasibility investigation is to determine if it is possible to design an imaging sensor based on the concept design (see Section 5.2) which meets the requirements stated in Section 1.2.

The feasibility investigation was divided into three parts. The first part of the investigation was aimed at determining is sufficient radiation will reach the detector, given the requirement of high spatial resolution (small IFOV). The second part of the investigation entailed determining if the spatial frequency response of the system will be satisfactory at high spatial resolutions. Finally, the third part of the investigation entails determining if the stereo geometry will result in sufficient stereo overlap.

6.3.1 Radiant Energy

The investigation into determining if sufficient radiant energy reaches the sensor was done by estimating the IFOV required to saturate the CCD for the various spectral bands and surface reflectivities. The saturation of the CCD would allow full use of its dynamic range. From these values it could be determined if the requirement of high spatial resolution can be met.

The grazing angle (see Figure 5.2) was assumed to be zero degrees and an obscured lens was assumed to have been used. Equation 5.13 was used to estimate the CCD output voltage.

The IFOV required to saturate the CCD, assuming a saturation voltage of 400 mV (the minimum saturation voltage of the TC104), for various spectral bands and surface reflectivities are shown in Table 6.2. It shows that the requirement of high spatial resolution can be met, while full use is made of the dynamic range of the CCD.

6.3.2 Spatial Frequency Response

The investigation into the spatial frequency response of the imaging sensor is described in this section. The objective of the investigation was to determine if the spatial frequency response is limited by the lens for high spatial resolutions. This was achieved by estimating the MTF of the sensor for various spatial resolutions.

The calculation of the MTF is based on the following assumptions: (a) The maximum allowed integration period is used; (b) The lens has a central obscuration (MTF data was obtained from [22]); (c) The grazing angle is zero degrees; (d) The wavelength is 700 nm; and (e) The alignment of the CCD relative to the input signal is "perfect" (see Section 3.5.3), i.e. $\phi_H = 0$ and $\phi_L = 0$.

The MTF of the sensor can be broken down into three components (see Sections 3.3.1, 3.5.3 and 5.4.2). If it is assumed that the integration period is given by Equation 6.1, i.e. the maximum period allowed, then relative to the sampling frequency the components due to the detector and image smear are constant. However, this is not so for the component due to the lens.

The MTF of a perfect, diffraction limited circular lens as a function of spatial frequency at the *image plane* is given by Equation 5.8. This can be rewritten as a function of spatial frequency at the *object plane* by using Equation 5.11:

$$\mathrm{MTF}_L(\nu_o) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{\lambda\nu_o l_s}{D}\right) - \frac{\lambda\nu_o l_s}{D}\sqrt{1-\left(\frac{\lambda\nu_o l_s}{D}\right)^2}\right] \tag{6.2}$$

where ν_o is the spatial frequency at the object plane, λ is the wavelength of the radiation, D is the diameter of the entrance pupil and l_s, the slant range, the distance from the lens to the object plane. Thus, the MTF due to a perfect, diffraction limited circular lens as seen at the object plane is not a function of focal length or, hence, IFOV (see Equation 5.7). The same is true for an obscured lens (assuming the degree obscuration is constant).

The MTF of the sensor and the components due to the lens and detector for focal lengths of 200 mm and 1100 mm are shown in Figure 6.1 and 6.2, respectively. The range of spatial frequency is from DC to the Nyquist rate (half the sampling frequency). The focal lengths of 200 mm and 1100 mm correspond to IFOVs of 42.8 m and 7.8 m, respectively. The MTF due to the image smear is not shown, since it is equal to the MTF due to the detector, under the assumptions stated above.

Figures 6.1 and 6.2 illustrate the following points: (a) The MTF due to the detector and image smear is constant relative to the sampling frequency; and (b) The MTF of the sensor is not limited by the lens, which is limited to an entrance pupil diameter of 10 cm. The MTF due to the detector and image smear is significant.

The MTF of the sensor for focal length of 200, 400, 800 and 1100 mm is shown in Figure 6.3. It shows the increased spatial resolution associated with increasing the focal length.

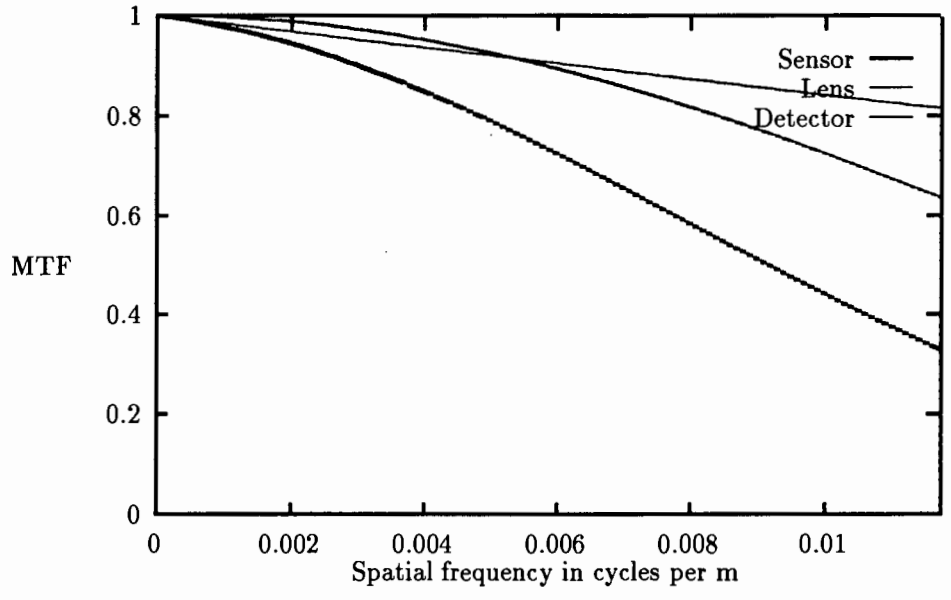


Figure 6.1: MTF of the sensor, lens and detector for a focal length of 200 mm.

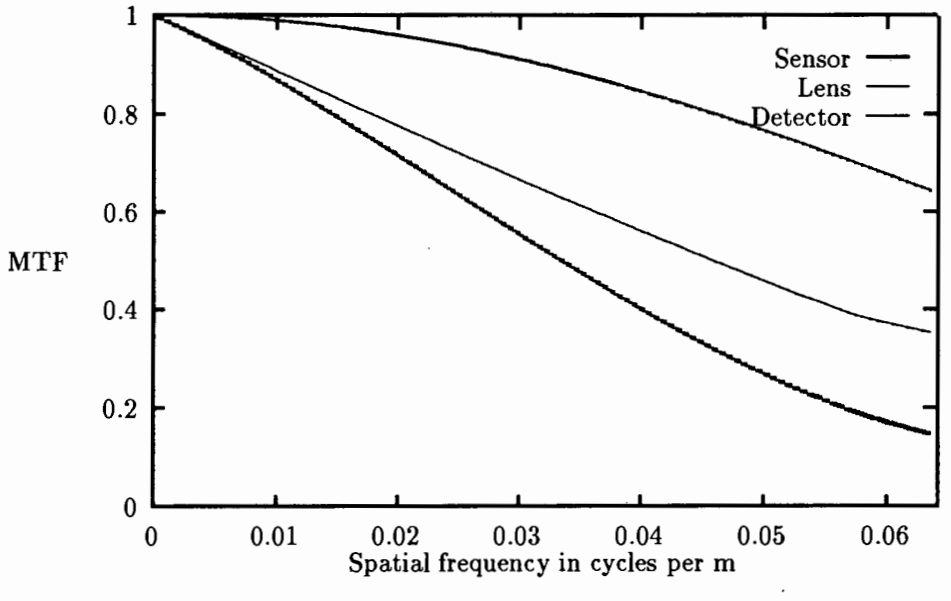


Figure 6.2: MTF of the sensor, lens and detector for a focal length of 1100 mm.

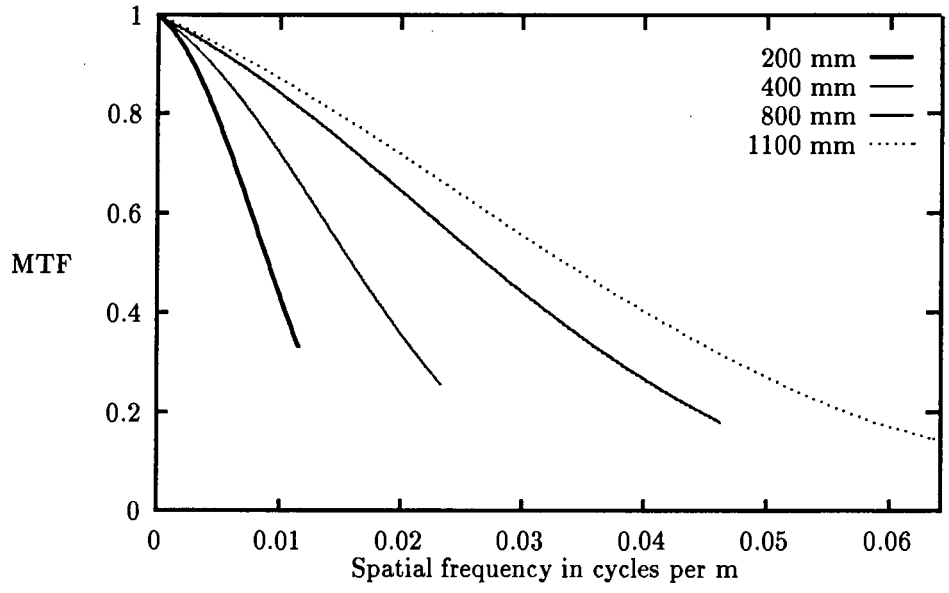


Figure 6.3: MTF of the sensor for focal lengths of 200, 400, 800 and 1100 mm.

6.3.3 Stereo Geometry

The investigation into the stereo geometry of the concept design entailed determining the degree of overlap of stereo pairs. This investigation was done by using the model of satellite sensor geometry described in Section 3.8.

The position of the ground track was determined when looking "forward" and "backward" in the same pass, the fore and aft tracks, respectively. It was assumed that the roll, yaw and pitch angles of the satellite were zero. The azimuth and zenith angles of the sensor were also assumed to be zero. The central CCD element is assumed to be at the centre of the image plane. From these tracks the degree of overlap could be determined as a function of latitude, B/H ratio and swath width.

The directional cosine of the sensor $\vec{d}$ was determined from the "look" direction of the sensor and the sensor element position. The line-of-sight angle of pixel I is given by:

$$\alpha_s(I) = \begin{cases} \theta_o & \text{look direction forward} \\ -\theta_o & \text{look direction backward} \end{cases} \tag{6.3}$$

and

$$\beta_s(I) = \tan^{-1} \frac{(I - 0.5I_{max})T_\beta}{f} \tag{6.4}$$

where θ_o is shown in Figure 5.2 and is given by Equation 5.4, I_{max} is the total number of detector elements on the CCD, T_β is the CCD pixel pitch and f is the focal length.

From the fore and aft tracks the percentage overlap was calculated. It was found to be a function of latitude, base-to-height ratio and swath width.

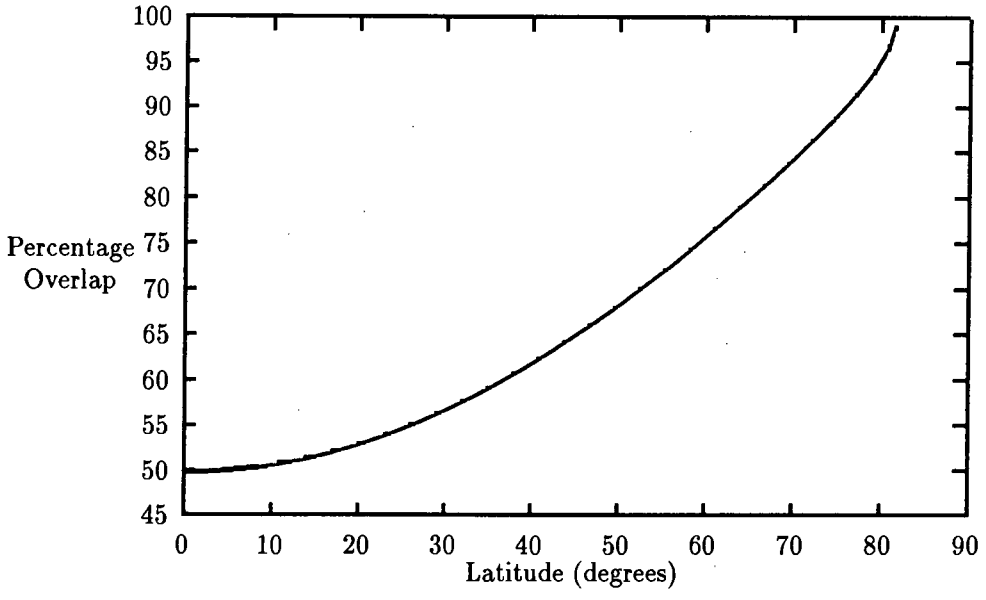


Figure 6.4: Percentage overlap of fore and aft tracks as a function of latitude. The base-to-height ratio was 0.85 and the swath width 86.4 km.

The percentage overlap of the fore and aft ground tracks as a function latitude is shown in Figure 6.4. A swath width of 86.4 km (corresponding to an IFOV of 25 m with 3456 detector elements) and a base-to-height ratio of 0.85 was used. It shows that the overlap is a minimum at the equator and reaches nearly 100 percent at the poles.

The percentage overlap of the fore and aft ground tracks at the equator as a function of the base-to-height ratio is shown in Figure 6.5. A swath width of 86.4 km was used. It shows the decrease in overlap as the base-to-height ratio is decreased.

The percentage overlap of the fore and aft ground tracks as a function of swath width is shown in Figure 6.6. A base-to-height ratio of 0.85 and zero degrees latitude was used. It shows that increasing the swath width increases the percentage overlap.

The results show that some overlap is obtained without special control of the satellite attitude. A higher percentage overlap can be obtained by controlling of the pitch, roll and yaw of the satellite. This will be discussed further in Section 6.4.5.

6.4 Sensor Specification

The determination of a specification for the imaging sensor is described in this section. The sensor parameters which have to be specified are: (a) the spectral bands (Section 6.4.1), (b) the A/D resolution (Section 6.4.2), (c) the focal length (Section 6.4.3), and (d) the offset of the sensor from the vertical (Section 6.4.4). The implications of the design on attitude

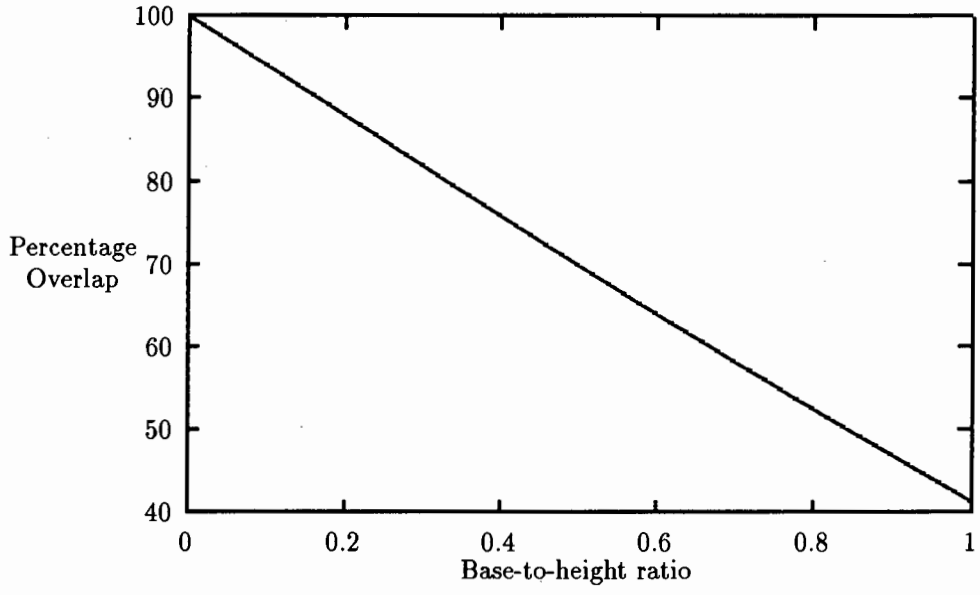


Figure 6.5: Percentage overlap of fore and aft tracks as a function of the base-to-height ratio at the equator. The swath width was 86.4 km.

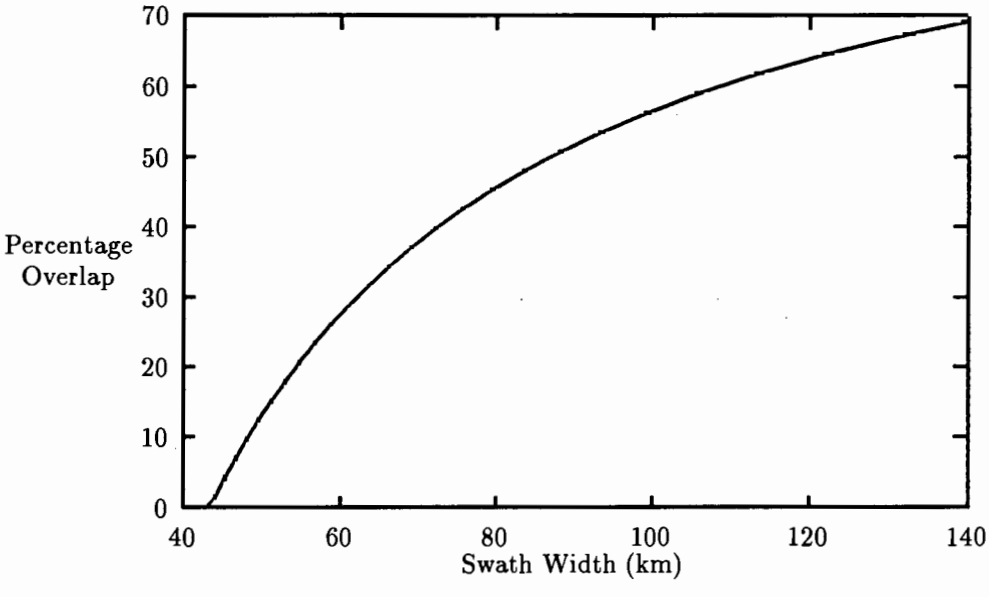


Figure 6.6: Percentage overlap of fore and aft tracks as a function of swath width. A base-to-height ratio of 0.85 and zero degrees latitude was used.

control are discussed in Section 6.4.5 and a summary of the sensor characteristics is presented in Section 6.4.6.

6.4.1 Spectral Bands

The specification of the spectral bands is described in this section. Three sets of spectral bands were identified in Section 6.2 for investigation. The choice of spectral bands is, hence, between these three sets.

The following desirable characteristics of the spectral bands were identified: (a) They must have the same spectral width; (b) They must be contiguous[1]; and (c) They must give an equal output for the same surface reflectance.

The spectral bands 4, 5 and 6 of the MSS meet these characteristics to the greatest degree. Each of these bands is 100 nm wide. They are contiguous and from Table 6.2 it can be seen the range of IFOV values required to saturate of the CCD is the smallest in the case of the MSS bands.

Therefore, the three spectral bands that are considered optimal are: 0.5 μm to 0.6 μm, 0.6 μm to 0.7 μm and 0.7 μm to 0.8μm.

6.4.2 A/D resolution

This section describes the procedure used to determine the appropriate choice of A/D resolution, the number of bits to used to quantise the signal. The theoretical basis for this investigation was presented in Section 3.6.

The signal-to-noise ratio of the imaging sensor is determined mainly by three factors, the noise and uncertainties in the surface irradiance (see Section 3.2.4), the noise introduced by the detector (a CCD) (see Section 3.5.2) and the quantisation noise introduced by the analogue-to-digital converter (see Section 3.6). The effect of quantisation was analysed in Section 3.6. Equation 3.52 shows that final signal-to-noise ratio is a function of the signal-to-noise of the input signal and the number of bits used for quantisation. Thus, the choice of A/D resolution must be made after the noise characteristics of the signal to be quantised has been analysed.

Short-term variations in atmospheric transmission of the order of 1% (40 dB) were described in Section 3.2.4. These variations should place an upper limit on the radiometric resolution of any imaging sensor. However, an accurate estimate of these short-term variations is not available, hence, it is assumed that these uncertainties are not limiting the signal-to-noise ratio.

The radiometric performance of the sensor was modelled in Section 5.5. This included the effect of quantisation, but not variations in atmospheric transmission. Table 5.1 shows the

[1] A set of three spectral bands that is the division of a larger band can be split with two dichroic prisms. A more complicated set of bands will require more hardware.

effect of quantisation on the signal-to-noise ratio. It can be seen that a 12 bit A/D has only minimal advantage over a 10 bit A/D. A 6 bit A/D results in a significant reduction in the signal-to-noise ratio. Hence, a choice has to be made between an 8 bit and a 10 bit A/D.

The MSS, TM and HRV all used 8 bits or less (see Sections 4.3.1, 4.3.2 and 4.4.1). In some bands, for example, the panchromatic band of the HRV, only 6 bits were used.

The analysis in Section 5.5 assumes that the full dynamic range of the CCD was used. However, this will not necessarily be the case. At high spatial resolutions insufficient radiation will reach the detector. This will result in a reduction in the signal-to-noise ratio and, hence, the benefit that a 10 bit A/D has over an 8 bit A/D.

The major disadvantage, of a 10 bit A/D compared to an 8 bit A/D, is the increase in the data rate of 25 percent (see Section 5.8). Thus, if a 10 bit A/D is to be justified it must result in an increase in the usefulness of the data.

For the reasons stated above an 8-bit A/D is considered to be optimal.

6.4.3 Focal Length

The specification of the focal length of the lens is described in this section. The focal length will determine the spatial resolution of the sensor and will affect numerous sensor characteristics.

An upper and a lower limit for the focal length was identified. The lower limit is determined by the focal length that will result in the CCD saturating (this requires an assumption to be made about the maximum surface reflectivity). The upper limit is determined by the focal length at which the MTF due to the lens is equal to zero at the Nyquist frequency.

At a wavelength of 900 nm the upper limit is 2370 mm, corresponding to an IFOV of 3.6 m. This will result in an output of only 5 mV from the CCD with a surface reflectivity of 1.0. This is not acceptable, hence, a shorter focal length is required. It should be noted that at this focal length the spatial frequency response would be limited by the lens.

Table 6.2 can be used to determine the lower limit of the focal length, give a maximum surface reflectivity. The choice of maximum surface reflectivity should be based on the reflectivity properties of Earth surface materials. Assuming a maximum surface reflectivity of 100 % the lower limit of the focal length is 522 m (IFOV 16.4 m).

Based on these considerations a focal length of 535 mm (IFOV 16 m) is considered to be optimal. The maximum output voltage would be 376 mV at a surface reflectivity of 100 % and 188 mV at a surface reflectivity of 50 %. A variable gain amplifier between the CCD and the A/D resolution should be used to optimise the dynamic range of the A/D under different scene conditions.

6.4.4 Sensor Offset

The specification of the sensor offset from the vertical is described in this section. The sensor offset determines the base-to-height ratio, which has a major effect on the stereo performance of the sensor. Thus, the determination of the sensor offset is done via the specification of the base-to-height ratio.

A base-to-height ratio of 1.0 is considered optimal for the following reasons:

- It offers a compromise between height difference errors (see Section 5.7) and matching difficulty (see Section 2.2.5).
- A base-to-height ratio of 1 has been proposed for used by Mapsat [9].
- A base-to-height ratio of 1 is the maximum that can be achieved by SPOT [7].

The required offset of the sensor from the vertical is 23.4 degrees. This gives a grazing angle of 26.6 degrees.

6.4.5 Attitude Control and Measurement

The implications of the design specifications on the attitude control system of the satellite are discussed in this section. Estimates of the attitude as a function of time are required for stereo reconstruction and a controlled yaw has to be introduced to maximise stereo overlap.

The production of digital elevation models from the stereo image pairs obtained from the sensor will required that a camera model be build (see Section 2.2.2). This will require that the roll, pitch and yaw of the satellite be controlled and measured during the imaging process. Estimates of roll, pitch and yaw must be encoded with the image data.

Errors in the measurement of the roll $\sigma\alpha$, pitch $\sigma\beta$ and yaw $\sigma\gamma$ will result in ground positioning errors. The attitude errors in radians multiplied by the satellite altitude will give the position errors on the ground [36]. For these position errors to be comparable to the ground pixel size, the attitude error should be of the order of 4 sec (for the design specified in the previous sections).

The design specification implies a swath width of only 55.3 km. This will result in a reduction in the stereo overlap as compared to the results presented in Section 6.3.3 unless corrective action is taken.

A baseline yawing algorithm was proposed by Welch and Marko [63] to achieve maximum stereo overlap. It can be expressed as

$$\gamma = \gamma_{max} \cos\phi \tag{6.5}$$

where ϕ is the satellite angle past the ascending node equator crossing and γ_{max} is the yaw required at the equator. In the case of the design specified in the previous sections the yaw required at the equator γ_{max} is 3.0 degrees.

Spectral Bands	0.5 μm to 0.6 μm
	0.6 μm to 0.7 μm
	0.7 μm to 0.8 μm
A/D resolution	8 bits
Entrance pupil diameter	10 cm
F/number	5.35
Focal length	535 mm
IFOV	16.0 m
Swath width	55.3 km
Integration period	2.4 ms
Base-to-height ratio	1.0
Sensor offset	23.4 degrees
Orbital Altitude	800 km

Table 6.3: Summary of sensor characteristics.

6.4.6 Summary of Sensor Characteristics

A summary of the imaging sensor characteristics is presented in this section. The characteristics were determined using the model developed in Chapter 5 and the specification developed in the previous sections.

The summary of sensor characteristics is tabulated in Table 6.3.

Chapter 7

Hierarchical Area Matching

7.1 Introduction

This chapter presents the first stage of the investigation into automatic digital stereo image matching. It deals primarily with the development of a *hierarchical area correlation matching algorithm* that is used as the basis for the second stage of the investigation (Chapter 8). The motivation for the algorithm was obtained from the literature review (Chapter 2). The stereo image pairs used for experimentation, the development tools and the evaluation of the results are also discussed.

The multiresolution (hierarchical) approach was motivated by: (a) The reduced computational burden (see Section 2.3.6); (b) The reduction in the number of false targets (see Section 2.3.6); and (c) Its wide use in digital stereo image matching (see Section 2.4).

The use of a correlation based measure of match, in particular, zero mean normalised cross-correlation (see Section 2.3.4), was motivated by: (a) Its wide use by the remote-sensing community (see [46,26,54,47]); (b) Its use by both area and feature matching techniques (see Section 2.3.4); (c) Its reduced sensitivity to absolute intensity, contrast and illumination (see Section 2.3.5); (d) Its ability to produce sub-pixel accuracy matches (see Section 2.3.7); and (e) Its robustness and wide range of applicability over image type [26].

The need for a dense disparity map motivated the use of an area matching technique (see Sections 2.2.3 and 2.2.4). Furthermore, the objective of feature selection is to only select points that can be matched with a high degree of confidence (see Section 2.2.3). Hence, techniques intended to improve reliability would only be tested on high confidence points if feature matching was used.

The synthetic and natural stereo image pairs used for experimentation are described in Section 7.2. The development tools (computer hardware and software) are described in Section 7.3. Section 7.4 deals with the evaluation of the matching results.

The development and implementation of the hierarchical area correlation matching algorithm is discussed in Section 7.5. Previous implementations of such algorithms (e.g. [26]) have

used the Gaussian pyramid of Burt and Adelson [5]. An investigation into the applicability of using an image pyramid based on the wavelet transform is described in Section 7.6.

7.2 Experimental Data

Both synthetic images, in the form of random dot stereograms, and natural images, in the form of digitised stereo aerial photographs, were used in the investigation. This section deals with the conditions the stereo images have to meet and the procedure used to obtain them.

The stereo image pairs had to meet certain conditions. This simplified the implementation of the matching algorithm, in particular the image pyramid algorithms, and reduced the scope of the investigation. Both images of the stereo pair had to have the same resolution (sampling rate), and the same size. Furthermore, they had to be square (the number of rows had to be equal to the number of columns) with the number of rows/columns a power of two. A final condition was that the stereo pair be in approximate correspondence.[1]

The technique used to generate the synthetic random dot stereograms is described in Section 7.2.1. The procedure used to obtain the digital stereo images from the aerial photographs is described in Section 7.2.2.

7.2.1 Synthetic Random Dot Stereograms

Synthetic random dot stereograms were introduced in Section 2.2.1. They were used due to the advantages they have over natural images. In particular, the fact that the exact correspondence is known and that parameters such as noise and density can be controlled.

A program was written that could generate a random dot stereogram with a central rectangular region at a different disparity to the rest of the stereogram. The stereogram also met the requirements stated above. The disparity of the central region, the density (probability of a pixel being white) and the size could be specified. Noise could also be added.

The structure of the virtual surface was motivated by the difficulties that area matching techniques have with depth discontinuities and occluding boundaries (see Table 2.1).

An example of such a random dot stereogram is shown in Figure 7.1. It has a density of 50 percent. The disparity of the central region is five pixels and it has a size of 128 by 128 pixels. The reference disparity map for this random dot stereogram is shown in Figure 7.2. The light area has a disparity of five pixels, the grey area a disparity of zero and the black area has an undefined disparity (this is due to occlusions).

[1]The peak disparity had to be small compared to the image size and the relative rotation of the images had to be small. The sensitivity to relative rotation can be reduced by using a rotationally invariant measure of match (see [21]).

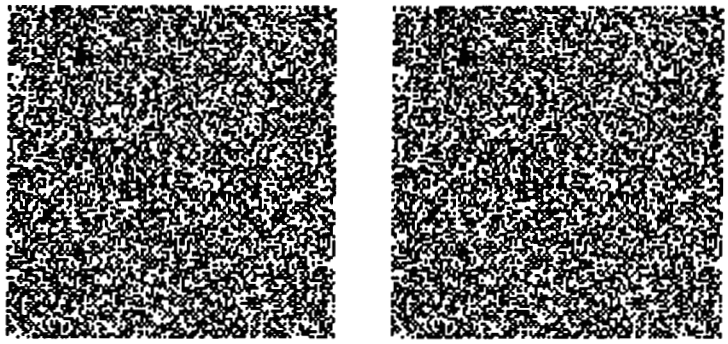


Figure 7.1: An example of a synthetic random dot stereogram. The disparity of the central region is five pixels. The density is 50 percent and the size is 128 by 128 pixels.

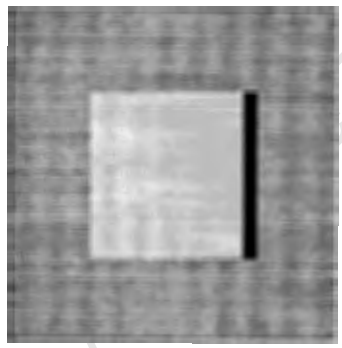


Figure 7.2: The reference disparity map of the random dot stereogram shown in Figure 7.1. The solid black region has an undefined disparity. This is due to occlusions. The light and grey regions have a disparity of five and zero pixels, respectively.

7.2.2 Stereo Aerial Photographs

A set of four aerial photographs of approximately the same area was obtained. It was not possible to obtain satellite stereo image data. The aerial photographs and the method used to obtain digital stereo image pairs from them are described in this section.

The four aerial photographs were digitised with a sampling rate of 300 samples per inch and 8 bits per sample (see Section 2.3.1). From the digitised aerial photographs a set digital stereo image pairs was extracted that met the requirements stated in Section 7.2.

The photographs contain regions of both natural and cultural features (see Section 2.2.1). This allowed stereo images pairs with different scene characteristics to be selected. Two digital stereo image pairs of 256 by 256 pixels were selected for experimental purposes. Test image pair #1 was completely covered by cultural features, roads and buildings, and is shown in Figure 7.3. It has a disparity range of 8.6 pixels. Test image pair #2, shown in Figure 7.4, consisted mainly of natural terrain and has a disparity range of 31.3 pixels.

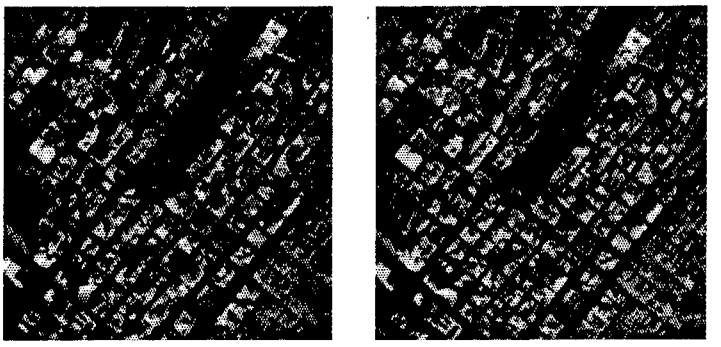


Figure 7.3: Test image pair #1.

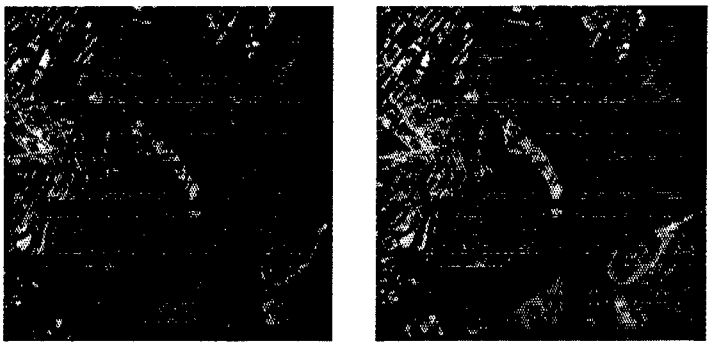


Figure 7.4: Test image pair #2.

7.3 Development Tools

The development tools (computer hardware and software) used in the development and implementation of the algorithms described in Chapters 7 and 8 are described in this section.

All the required software was written in the C programming language and runs on Sun workstations. The workstations run SunOS 4.1.3 and Openwindows V3.0.

The Khoros image processing package [31,32] was used to perform the basic image processing required, including the extraction of stereo image pairs and the visualisation of results. The software was integrated into the Khoros package as a "Toolbox". This made it possible to use the development tools and software libraries that are part of the Khoros package.

7.4 Evaluation of Matching Results

The procedures used to evaluate the matching results are discussed in this section. They are based on the requirements of an automatic stereo image matching system (see Section 1.2), in particular accuracy and reliability, and on the classification of matching errors presented in Section 2.3.3.

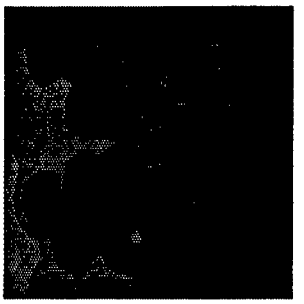


Figure 7.5: The horizontal reference disparity map for test image pair #1 (see Figure 7.3). The disparity range is from -8.9 pixels to -0.3 pixels.

A distinction must be made between the two image types, synthetic and natural. In the case of synthetic random dot stereograms the exact correspondence is known and in the case of natural images it is not. It is possible in the case of natural images for several points to be matched by a human operator. A dense reference disparity map can then be obtained by interpolation.

The procedure used to obtain a reference disparity map in the case of natural images will be described in Section 7.4.1. The procedure used to evaluate the results, with the aid of an exact or reference disparity map, will be described in Section 7.4.2.

7.4.1 Reference Disparity Maps

In the case of stereo image pairs for which the exact correspondence is not known the results have to be evaluated with the aid of a human operator. The procedure used to perform this task is based on the work of Hsieh *et al.* [29] and Niblack [46, pp. 137-140].

The procedure consists on three steps: (a) The disparity at points on a regular grid is determined by a human operator; (b) A dense reference disparity map is generated by interpolation; and (c) The reference disparity map is compared to the disparity map being evaluated (see Section 7.4.2).

The disparity values determined by the human operator were estimated to sub-pixel accuracy using the procedure described in Section 2.3.7 [46]. A correlation window radius of 7 pixels was used. Bilinear interpolation was used to interpolate the disparity map (see [46, pg. 144]). The number of points is dependent on the image size and disparity complexity. For the test image pairs (see Section 7.2), 256 points were used.

Since the images were not required to be in exact vertical correspondence two disparity maps were obtained, one for the vertical component of the disparity and one for the horizontal component of the disparity. The horizontal reference disparity for test image pair #1 is shown in Figure 7.5 and the horizontal reference disparity map for test images pair #2 is shown in Figure 7.6. In both figures undefined disparities are shown in white.

Figure 7.6: The horizontal reference disparity map for test image pair #2 (see Figure 7.4). The disparity range is from -21.1 pixels to 10.2 pixels.

7.4.2 Comparison of Disparity Maps

If the exact disparity map or a reference disparity map is available then the matching error[2] for each point in a disparity map can be determined. The presence of occlusions, which result in the disparity being undefined, must also be considered. To distinguish between the two matching error classes described in Section 2.3.3, the magnitude of the matching error was used. If the magnitude of the matching error is greater than one pixel then a "gross" matching error is assumed to have occurred, otherwise the error is attributed to the errors in estimating the position the corresponding points. The same classification was used by Hsieh *et al.* [29].

Each disparity estimate can be classified into one of five classes. They are depicted in Table 7.1. The features used for classification are: (a) the presence of an occlusion, (b) the detection of an occlusion, and (c) the matching error. It should be noted that if an occlusion is present or has been detected then the matching error cannot be determined. The number of disparity estimates in each class indicate the degree of success of the match. In the ideal case all disparity estimates would be in classes (1) and (2). The number of "gross" matching errors is given by the number of disparity estimates in class (3). Classification into class (4) indicates the failure to find a corresponding point and classification into class (5) indicates the failure to detect an occlusion.

Using these five classes measures of performance can be defined. The percentage of "correct" matches is given by the ratio of the number of disparity estimates in class (1) to the number in classes (1) and (3). The percentage of "occlusions" detected is given by the ratio of the number of disparity estimates in class (2) to the number in classes (2) and (5). These measures are used to determine the degree to which the requirement of reliability is satisfied.

The root mean square matching error (RMSME) is used to measure the degree to which

[2]The matching error is the difference between the exact/reference disparity and the estimated disparity. If is exact/reference disparity is $\vec{D}$ and the estimated disparity is $\vec{d}$ then the matching error is given by $\vec{D} - \vec{d}$.

Class	Occlusion Present	Occlusion Detected	Matching Error (pixels)
1	No	No	< 1
2	Yes	Yes	-
3	No	No	≥ 1
4	No	Yes	-
5	Yes	No	-

Table 7.1: The five classes into which a disparity estimate can be classified by comparing it to the exact or reference disparity.

the requirement of accuracy is satisfied. It is defined as:

$$e_d = \sqrt{\frac{\sum_{i \in I} |\vec{D}_i - \vec{d}_i|^2}{N}} \tag{7.1}$$

where I is the set of points in classes (1) and (3), $\vec{D}_i$ is the exact or reference disparity at point i, $\vec{d}_i$ is the estimated disparity at point i and N is the number of points in classes (1) and (3).

7.5 Hierarchical Area Correlation Matching Algorithm

The development and implementation of the hierarchical area correlation matching algorithm used as the basis for the investigation is described in this section. The basis for this algorithm was described and motivated in Section 7.1. The conditions that the stereo images have to meet, to be used as input to this algorithm were described in Section 7.2.

The algorithm is similar to that of Hannah [26], except that no feature selection is done. The hierarchical (multiresolution) nature of the algorithm implies that a dyadic image pyramid must be generated for each image of the stereo pair. Two techniques for generating such image pyramids were described in Section 2.4, namely the Gaussian and wavelet pyramids. Both techniques were implemented. An investigation into the applicability of the wavelet pyramid is described in Section 7.6.

A description of the hierarchical area correlation matching algorithm is given below. The description is divided into two parts: (a) the hierarchical matching algorithm, and (b) area correlation algorithm. The hierarchical matching algorithm determines when each pyramid level is used, the matching order of the pixels and how the initial disparity estimates for each match are determined. The area correlation algorithm improves the disparity estimate for a pixel. The number of images of different resolution in the image pyramid is depicted by L and the resolution of an image at level l is 2^l.

Hierarchical Matching Algorithm

(1) Generate an initial disparity map at level $-L$.

(2) Set l to $-L+1$.

(3) Estimate the disparity map at level l from disparity map level $l-1$.

(4) For all pixels for which the disparity is defined, use the area correlation algorithm to improve the disparity estimate.

(5) For all pixels for which the disparity is undefined, but is defined for at least one of its nearest neighbours, use the area correlation algorithm to estimate the disparity. The average disparity of the nearest neighbours for which it is defined is used as the initial estimate.

(6) Increase l by 1.

(7) If l is less than or equal to zero, go to step (3).

Area Correlation Algorithm

(1) If boundary conditions are not satisfied, exit with the disparity undefined. The boundary conditions require that the search and correlation windows remain within the bounds of the image pair.

(2) Calculate correlation coefficients for each pixel in the search window centred at the initial disparity estimate, using the zero mean normalised cross-correlation measure.

(3) Find the maximum of the correlation coefficients, excluding the pixels on the boundary of the search window.

(4) If that maximum is not greater than all its nearest neighbours, including the pixels on the boundary of the search window, exit with the disparity undefined.

(5) Estimate the position of the maximum to sub-pixel accuracy.

(6) Calculate the disparity estimate from the estimated position of the maximum.

The position of the correlation maxima is estimated to sub-pixel accuracy using Equations 2.17 and 2.18, and the procedure described in Section 2.3.7.

The initial disparity map at level $-L$ is generated by setting the disparity to zero for all points in the disparity map (step (1) of the hierarchical area correlation algorithm). This procedure is based on the approximate correspondence required stated in Section 7.2. The selection of the number of pyramid levels L is based on the expected peak disparity and the search radius.

The procedure used to estimate the disparity map at level l from the disparity map at $l-1$ (step (3) of the hierarchical area correlation algorithm) consists of two steps. First, all point in the disparity map at level l, d_l, are set to *undefined*. The second step is to set $d_l(2\vec{u}) = 2d_{l-1}(\vec{u})$ for all $\vec{u}$ in the disparity map at level $l-1$ for which the disparity, $d_{l-1}(\vec{u})$, is defined. It should be noted that the disparity map at level l is twice the size of that at level $l-1$.

The major implementation problems were related to boundary effects. The boundary condition of the area correlation algorithm (step (1)) resulted in a set of points along the border of the images that could not be matched. The width of this border became larger after resolution level increased. To overcome this problem step (5) of the hierarchical matching algorithm being iterated until all possible pixels had been matched or until a maximum number of iterations had been completed.

7.6 Applicability of the Wavelet Pyramid

Two techniques for generating the required image pyramids were described in Section 2.4, namely the Gaussian and wavelet pyramids. The investigation into determining the applicability of wavelet pyramid is described in this section. The Gaussian pyramid had been previously used for stereo image matching (e.g. [26]), therefore it is assumed to be suitable.

The procedure was based on the observation that in both cases the final matching is done on the original stereo image pair. Hence, differences in the final disparity maps are due to matching errors made at the lower levels of the pyramid. To determine the applicability of the wavelet pyramid, the performance of the hierarchical area correlation matching algorithm was evaluated when using a wavelet pyramid and a Gaussian pyramid.

The investigation used both synthetic images (random dot stereograms) and natural images (digitised aerial photographs). The case of random dot stereograms is described in Section 7.6.1 and case of natural images in Section 7.6.2.

7.6.1 Synthetic Images

Two random dot stereograms were generated. Both had a density of 50 percent and a size of 128 by 128 pixels. The disparity of the central region was 4 and 5 pixels, respectively. Each random dot stereogram was matched using the hierarchical area correlation matching algorithm described in Section 7.5 with a search window radius of 4 pixels and with two image pyramid levels.

A correlation window radius of 3, 5 and 7 pixels was used. Both the wavelet pyramid and the Gaussian pyramid were used to generate the required image pyramids. Four compactly supported orthonormal wavelets of Daubechies [13] were used in the wavelet pyramid case. Wavelets #1, #2, #3 and #4 correspond to $N = 1$, $N = 2$, $N = 3$ and $N = 4$ in Table 6.1

Image Pyramid Type	Correlation Window Radius	Class 1	Class 2	Class 3	Class 4	Class 5
Gaussian	3	10565	0	346	0	325
Wavelet #3	3	10488	0	423	0	325
Wavelet #2	3	10430	0	481	0	325
Wavelet #1	3	10353	0	558	0	325
Wavelet #4	3	10325	0	586	0	325
Gaussian	5	10360	0	551	0	325
Wavelet #4	5	10258	0	653	0	325
Wavelet #3	5	10241	0	670	0	325
Wavelet #1	5	10169	0	742	0	325
Wavelet #2	5	10169	0	742	0	325
Gaussian	7	10212	0	699	0	325
Wavelet #4	7	10035	0	876	0	325
Wavelet #3	7	9996	0	915	0	325
Wavelet #2	7	9966	0	945	0	325
Wavelet #1	7	9925	0	986	0	325

Table 7.2: Results of matching a random dot stereogram with density of 50 percent and peak disparity of 5 pixels. The basic hierarchical area correlation matching algorithm was used. For each correlation window size the results are in order of performance (number of matches in classes (1) and (2)).

of Daubechies [13, pg. 195], respectively.

A border of 11 pixels wide was not considered for classification. This border allowed for the search window radius of 4 pixels and the maximum correlation window radius of 7 pixels. Thus, the total number of matches classified per random dot stereogram was 11236.

The results of classifying the matches using the procedure described in Section 7.4.2 are depicted in Tables 7.2 and 7.3. The disparity of the central region was 5 and 4 pixels in the case of Tables 7.2 7.3, respectively. The results are in order of performance (the number of matches in classes (1) and (2)) for the different correlation window radii.

The overall performance of the basic hierarchical area correlation algorithm was low. It failed to detect most occlusion points. This is indicated by the high number of points in class (5) compared to the number points in class (2). Many "gross" matching errors (class (3)) were also made, which increased with the correlation window radius. The number of points for which a match (class (4)) was not found was low.

In the case of a disparity of 5 pixels (Table 7.2) the Gaussian pyramid outperformed all the wavelet pyramid types. However, in the case of a disparity of 4 pixels (Table 7.3) the Gaussian pyramid did not perform best.

The results indicated that neither the wavelet pyramid nor the Gaussian pyramid has an overall advantage when matching random dot stereograms.

Image Pyramid Type	Correlation Window Radius	Class 1	Class 2	Class 3	Class 4	Class 5
Wavelet #2	3	10757	1	218	1	259
Wavelet #1	3	10754	1	220	2	259
Wavelet #3	3	10752	1	223	1	259
Gaussian	3	10739	1	236	1	259
Wavelet #4	3	10730	1	245	1	259
Wavelet #1	5	10698	0	278	0	260
Wavelet #2	5	10675	0	301	0	260
Wavelet #3	5	10674	0	302	0	260
Wavelet #4	5	10623	0	353	0	260
Gaussian	5	10623	0	353	0	260
Wavelet #1	7	10598	0	378	0	260
Wavelet #2	7	10570	0	406	0	260
Wavelet #3	7	10551	0	425	0	260
Wavelet #4	7	10532	0	421	0	260
Gaussian	7	10525	0	451	0	260

Table 7.3: Results of matching a random dot stereogram with density of 50 percent and peak disparity of 4 pixels. The basic hierarchical area correlation matching algorithm was used. For each correlation window size the results are in order of performance (number of matches in classes (1) and (2)).

7.6.2 Natural Images

The two test image pairs described in Section 7.2.2 were used. Each image pair was matched using the hierarchical area correlation matching algorithm described in Section 7.5 with a search window radius of 3 pixels and with three image pyramid levels.

A correlation window radius of 3, 5 and 7 pixels was used. Both the wavelet pyramid and the Gaussian pyramid were used to generate the required image pyramids. The same four compactly supported orthonormal wavelets of Daubechies [13] were used in the wavelet pyramid case.

The reference disparity maps described in Section 7.4.1 were used to evaluate the results obtained. The results obtain for test image pair #1 are presented in Table 7.4 and those for test image pair #2 are presented in Table 7.5. In both cases the percentage of "correct" matches obtain for various image pyramid types and correlation window radii are shown.

The results again indicate that the dominant factors determining the performance is the correlation window radius and the image characteristics.

The results of both the synthetic and natural images show that the wavelet pyramid can be applied to digital stereo image matching. In Chapter 8 techniques for improving the reliability of the matching algorithm and the effect of scene characteristics will be discussed.

Image Pyramid Type	Correlation Window Radius	Percentage of "Correct" Matches
Wavelet #1	3	61.3 %
Wavelet #2	3	61.3 %
Wavelet #3	3	61.3 %
Wavelet #4	3	61.3 %
Gaussian	3	61.2 %
Wavelet #1	5	75.0 %
Wavelet #2	5	75.0 %
Wavelet #3	5	75.0 %
Wavelet #4	5	75.0 %
Gaussian	5	74.9 %
Wavelet #1	7	85.2 %
Wavelet #2	7	85.2 %
Wavelet #4	7	85.1 %
Wavelet #3	7	85.0 %
Gaussian	7	85.0 %

Table 7.4: Results of matching test image pair #1 with different image pyramid types. The basic area correlation matching algorithm was used.

Image Pyramid Type	Correlation Window Radius	Percentage of "Correct" Matches
Gaussian	3	62.1 %
Wavelet #1	3	60.0 %
Wavelet #3	3	59.5 %
Wavelet #2	3	58.7 %
Wavelet #4	3	57.5 %
Wavelet #4	5	76.9 %
Wavelet #3	5	76.2 %
Wavelet #1	5	75.6 %
Wavelet #2	5	73.5 %
Gaussian	5	73.5 %
Gaussian	7	77.9 %
Wavelet #4	7	76.8 %
Wavelet #1	7	76.1 %
Wavelet #3	7	73.8 %
Wavelet #2	7	73.4 %

Table 7.5: Results of matching test image pair #2 with different image pyramid types. The basic area correlation matching algorithm was used.

Chapter 8

Reliability of Matching Techniques

8.1 Introduction

The second stage of the investigation into automatic digital stereo image matching is presented in this chapter. It deals with techniques for improving the reliability of the hierarchical area correlation matching algorithm developed in Chapter 7 and with the effect of scene characteristics on performance. The accuracy of the results is also considered.

The investigation into the applicability of the wavelet pyramid (see Section 7.6) gave an indication of the performance of the hierarchical area correlation matching algorithm (see Section 7.6). In the case of random dot stereograms it failed to detect the majority of the occlusion points and in the case of the digitised aerial photographs the percentage of "correct" matches was between 57.5 % and 85.2 %. The techniques investigated in this chapter are aimed at improving these results.

Back-matching (Section 8.2), the removal of inconsistencies (Section 8.3) and double matching (Section 8.4) were investigated. Back-matching checks each potential corresponding point by matching it in the reverse direction. The removal of inconsistencies produces a result that is physically consistent. Double matching finds occlusion points by validating each match produced by matching the stereo pair in two directions.

The objective of the investigation into the effect of scene characteristics on performance is to determine the *generality* and *predictability* of the various techniques. It is described in Section 8.5. A discussion of the matching results is presented in Section 8.6.

8.2 Back-matching

Back-matching was introduced in Section 2.3.4 as a control strategy that validates each potential match. It is based on the assumption that a pair of corresponding points will result in a local maximum in the correlation surface whatever the matching direction. The matching algorithm, described in Section 7.5, only matched in one direction, i.e. one image is always

Random-Dot Pair	Correlation Window Radius	With Back-Matching: Percentage "Correct" Matches	With Back-Matching: Percentage Occlusion Detected	Without Back-Matching: Percentage "Correct" Matches	Without Back-Matching: Percentage Occlusion Detected
#1	3	98.1 %	1.5 %	98.0 %	0.4 %
#2	3	95.1 %	2.8 %	94.9 %	0.0 %
#1	5	97.5 %	0.0 %	97.5 %	0.0 %
#2	5	93.3 %	1.5 %	93.2 %	0.0 %
#1	7	96.6 %	0.0 %	96.6 %	0.0 %
#2	7	91.0 %	0.3 %	91.0 %	0.0 %

Table 8.1: Results of back-matching random dot stereograms.

used as the template.

Back matching requires additional steps that validate each potential match by matching in the reverse direction and only accepting those matches that have a local maximum in the correlation surface. These steps are inserted between steps (4) and (5) of the area correlation algorithm.

The additional steps are:

(4a) Calculate correlation coefficients for each pixel in a search window of 5 by 5 pixels centred at the pixel currently being matched. The template image is centred at position of match being validated.

(4b) Check that a local maximum is present in the central 3 by 3 region of the correlation coefficients.

(4c) If a local maximum is not present, exit with the disparity undefined.

To determine the benefit of back-matching a set of experiments was performed. Both natural images and synthetic images were used. The performance of the hierarchical area correlation matching algorithm with and without back-matching was compared. The wavelet pyramid, using wavelet #1, was used to generate the required image pyramids.

In the case of the random dot stereograms the search window radius was four pixels and the number of image pyramid levels two. Two stereograms were used. Pair #1 has a central disparity of 4 pixels and pair #2 a central disparity of 5 pixels. Both pairs have a density of 50 % and a size of 128 by 128 pixels. A correlation window radius of 3, 5 and 7 pixels was used.

The results for the random dot stereograms are shown in Table 8.1. They show that the use of back-matching resulted in a minor increase in the percentage of "correct" matches and in the percentage of occlusion points detected.

Image Pair	Correlation Window Radius	Percentage "Correct" Matches		RMSME (Pixels)	
		With Back-matching	Without Back-matching	With Back-matching	Without Back-matching
#1	3	66.5 %	61.3 %	1.33	1.56
#2	3	66.7 %	60.0 %	4.38	7.16
#1	5	79.6 %	75.0 %	0.86	1.13
#2	5	80.2 %	75.6 %	1.97	3.83
#1	7	89.0 %	85.2 %	0.58	0.81
#2	7	83.1 %	76.1 %	1.63	6.99

Table 8.2: Results of back-matching the test image pairs.

In the case of natural images, the two test image pairs described in Section 7.2.2 were used. The search window radius was three pixels and the number of image pyramid levels three. A correlation window radius of 3, 5 and 7 pixels was used.

The results of matching the test image pairs are shown in Table 8.2. The improvement in the percentage of "correct" matches was significant, between 3.8 and 7.0 percentage points. An improvement in the RMSME was also found. The improvement in the RMSME can be attributed to the rejection of matches with large matching errors (greater than one pixel).

8.3 Removal of Inconsistencies

This section discusses the motivation, development and implementation of the module that detects and removes inconsistent matches. A set of experiments aimed at determining its performance are also described. It is designed to produce a result that is physically consistent.

The basis of the module is the definition of a matching inconsistency. The definition is motivated by the uniqueness and continuity rules on Marr and Poggio [40] (see Section 2.2.4). A pair of matches is declared inconsistent if the pair of matches violate the uniqueness and continuity rules. This implies that at least one match is incorrect, subject to the assumptions made about the nature of the scene.

The horizontal and vertical components of the disparity are considered separately. In the case of the horizontal component, a pair of matches x_i and x_j on the same horizontal row are inconsistent if

$$x_i - x_j > j - i \tag{8.1}$$

where x_i is the horizontal component of the disparity at pixel i, x_j is the horizontal component of disparity at pixel j, and $j > i$. In the case of the vertical component, a pair of matches y_i and y_j on the same vertical column are inconsistent if

$$y_i - y_j > j - i \tag{8.2}$$

where y_i is the vertical component of the disparity at pixel i, y_j is the vertical component of disparity at pixel j, and $j > i$. Note that application of this definition to the disparity map of the random dot stereogram shown in Figure 7.1 will not produce any inconsistencies.

The next problem is that of identifying which match of an inconsistency pair is more likely to be incorrect. The basis for this determination is a measure of confidence associated with each match. The correlation coefficient is used as the measure of confidence.

An iterative procedure is used. It is based on the assumption that the lower the confidence measure the more likely the match is incorrect, and that the larger the number of times a point is part of an inconsistency pair the more likely that the match is incorrect. The iteration is terminated when no more inconsistencies are detected.

Each iteration consists of three steps.

(1) All inconsistencies are detected and the number of times each match is part of an inconsistency pair is recorded.

(2) For all single pair inconsistencies the match of the pixel with the lower confidence is removed (the disparity is set to undefined).

(3) The match of the pixel with the largest number of inconsistencies associated with it is removed (the disparity set to undefined).

When this module is used as part of the hierarchical area correlation algorithm, it is called at the end of each iteration of step (5) of the hierarchical matching algorithm.

To determine the effect of including this module in the hierarchical area correlation matching algorithm a set of experiments was performed. The following three cases were compared: (1) inconsistency removal only, (2) inconsistency removal and back-matching, and (3) the standard hierarchical area correlation matching algorithm. The same test image pairs and parameter settings were used as in Section 8.2.

The results of matching the random dot stereograms are shown in Table 8.3. Only cases (1) and (2) are shown (case (3) is given in Table 8.1). They show a dramatic increase in the percentage of occlusion points detected and a minor increase in the percentage of "correct" matches. The use of back-matching with inconsistency removal has a minor advantage of inconsistency removal alone.

The results of matching the test image pairs are shown in Table 8.4. They show that the removal of inconsistencies results in a significant improvement in the percentage of "correct" matches. Again the use of inconsistency removal and back-matching results in the highest percentage of "correct" matches. The RMSME is reduced by the removal of inconsistencies. The use of inconsistency removal results in a RMSME of less than one pixel for a correlation window of 5 and 7 pixels.

Random-Dot Pair	Correlation Window Radius	Inconsistency Removal Only: Percentage "Correct" Matches	Inconsistency Removal Only: Percentage "Occlusion" Detected	Inconsistency Removal and Back-Matching: Percentage "Correct" Matches	Inconsistency Removal and Back-Matching: Percentage "Occlusion" Detected
#1	3	99.0 %	71.5 %	99.1 %	73.5 %
#2	3	98.3 %	75.1 %	98.5 %	75.1 %
#1	5	98.7 %	66.5 %	98.7 %	66.5 %
#2	5	95.5 %	75.1 %	95.6 %	72.0 %
#1	7	98.3 %	66.5 %	98.3 %	66.2 %
#2	7	94.2 %	66.8 %	94.2 %	66.8 %

Table 8.3: Results of inconsistency removal on the matching of random-dot stereograms.

Image Pair	Correlation Window Radius	Percentage "Correct" Matches: Case (1)	Case (2)	Case (3)	RMSME (pixels): Case (1)	Case (2)	Case (3)
#1	3	64.6 %	68.6 %	61.3 %	1.18	1.05	1.56
#2	3	72.9 %	74.7 %	60.0 %	1.37	1.14	7.16
#1	5	76.3 %	80.2 %	75.0 %	0.91	0.73	1.13
#2	5	81.6 %	83.6 %	75.6 %	0.74	0.75	3.83
#1	7	85.7 %	89.1 %	85.2 %	0.71	0.55	0.81
#2	7	83.8 %	85.8 %	76.1 %	0.75	0.68	6.99

Table 8.4: Results of inconsistency removal on the matching of the test image pairs. Case (1) is inconsistency removal only. Case (2) is inconsistency removal and back-matching. Case (3) is the standard hierarchical area correlation matching algorithm.

8.4 Detection of Occlusion Points

This section describes the use of a double matching procedure to detect occlusion points (points present in one image). It is based on the work of Fleck [18]. The requirement for such a procedure was the poor performance in this respect of previous algorithms. Not more than 75 % of the occlusion points were detected.

The computational cost of the procedure is high as it requires that the stereo image pair be matched twice, once with reference to right image and once with reference to the left image. This is done by applying the matching algorithm twice.

The two disparity maps, d_l (with reference to the left image) and d_r (with reference to the right image), are then analysed. For each point $\vec{x}_l$ in $\vec{d}_l$ for which the disparity is defined the following steps are performed:

(1) The point corresponding to $\vec{x}_l$ is given by $\vec{x}_l + \vec{d}_l(\vec{x}_l)$.

Random-Dot Pair	Correlation Window Radius	With Double Matching: Percentage "Correct" Matches	With Double Matching: Percentage "Occlusion" Detected	Without Double Matching: Percentage "Correct" Matches	Without Double Matching: Percentage "Occlusion" Detected
#1	3	99.5 %	89.6 %	99.1 %	73.5 %
#2	3	99.0 %	80.3 %	98.5 %	75.1 %
#1	5	99.2 %	78.1 %	98.7 %	66.5 %
#2	5	96.5 %	74.8 %	95.6 %	72.0 %
#1	7	98.5 %	83.5 %	98.3 %	66.2 %
#2	7	94.7 %	68.6 %	94.2 %	66.8 %

Table 8.5: Results of double matching random dot stereograms.

Image Pair	Correlation Window Radius	Percentage "Correct" Matches: With Double Matching	Percentage "Correct" Matches: Without Double Matching	RMSME (Pixels): With Double Matching	RMSME (Pixels): Without Double Matching
#1	3	71.8 %	68.6 %	0.86	1.05
#2	3	81.0 %	74.7 %	0.73	1.14
#1	5	81.4 %	80.2 %	0.67	0.73
#2	5	86.7 %	83.6 %	0.59	0.75
#1	7	89.6 %	89.1 %	0.53	0.55
#2	7	87.8 %	85.8 %	0.56	0.68

Table 8.6: Results of double matching the test image pairs.

(2) The point nearest to $\vec{x}_l + \vec{d}_l(\vec{x}_l)$ on the discrete grid is determined and assigned to $\vec{x}_r$.

(3) If the disparity at $\vec{x}_r$ is undefined, set the disparity at $\vec{x}_l$ to undefined and exit.

(4) The point corresponding to $\vec{x}_r$ is given by $\vec{x}_r + \vec{d}_r(\vec{x}_r)$.

(5) If $\vec{x}_r + \vec{d}_r(\vec{x}_r)$ not within one pixel of $\vec{x}_l$, set the disparity at $\vec{x}_l$ and $\vec{x}_r$ to undefined.

The output of the procedure is a new disparity map with reference to the left image.

To determine the performance improvement, of double matching, a set of experiments was performed. The same test image pairs and parameters were used as in Section 8.2. Both matches were performed with back-matching and inconsistency removal. The results were compared to the use of back-matching and inconsistency removal alone.

The results of double matching the random dot stereograms are shown in Table 8.5. It shows an improvement in the percentage of "correct" matches and in the percentage of occlusion points detected. A maximum of 89.6 % of the occlusion points was detected.

Image Variance Range	Percentage "Correct" Matches		
	7	5	3
0 to 5	- (0.0%)	21.1 % (0.1 %)	34.2 % (0.4 %)
5 to 10	- (0.0 %)	60.6 % (0.1 %)	58.0 % (0.7 %)
10 to 20	95.7 % (0.1%)	79.3 % (0.6 %)	64.0 % (1.5 %)
20 to 70	82.1 % (1.2%)	76.0 % (2.1 %)	75.3 % (4.5 %)
70 to 300	93.7 % (5.4 %)	85.1 % (8.1 %)	73.7 % (14.3 %)
300 to ∞	88.9 % (93.2 %)	79.9 % (89.0 %)	67.7 % (78.6 %)

Table 8.7: The percentage of "correct" matches as a function of local image variance for test image #1. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used. The percentage of points within each variance range is shown in brackets.

The results of matching the test image pairs are shown in Table 8.6. They show an improvement in the percentage of "correct" matches and in the RMSME. With double matching the RMSME was less than one pixel for all cases shown in Table 8.6.

8.5 The Effect of Scene Characteristics

This section describes the investigation into the effect of scene characteristics. The scope of this investigation is limited by the available image data. Various aspects of image acquisition and scene characteristics that effect the results obtain were described in Section 2.2.1. It was not possible to investigate the effect of the photometry and geometry of the camera, the image resolution, the illumination of the scene and atmospheric response. The effect of the disparity and image variance was investigated.

The percentage of "correct" matches and RMSME as a function of the local image and disparity variance was determined for the two test image pairs (see Section 7.2.2). The result obtained from the use of both back-matching and inconsistency removal was used for this analysis.

In the case of the image variance, this was done by classifying the matches for points that have a local variance with a specified range. The window size used to calculate the variance was the same as the correlation window size. The percentage of "correct" matches and RMSME within these ranges was determined. The percentage of "correct" matches is shown in Tables 8.7 and 8.8 for test images #1 and #2, respectively. The percentage of points with each range is shown in brackets.

The results show that the higher the image variance the higher the percentage of "correct" matches. This can be explained by the observation that the higher the local image variance the higher the local signal-to-noise ratio will be. This will result in higher correlation coefficient

Image Variance Range	Percentage "Correct" Matches		
	7	5	3
0 to 5	70.9 % (16.2 %)	68.8 % (22.7 %)	56.1 % (32.5 %)
5 to 10	84.8 % (20.3 %)	81.9 % (20.7 %)	74.2 % (19.6 %)
10 to 20	86.8 % (13.9 %)	86.1 % (13.8 %)	82.5 % (11.7 %)
20 to 70	85.6 % (13.3 %)	85.7 % (13.8 %)	84.0 % (12.5 %)
70 to 300	91.2 % (22.2 %)	92.9 % (17.7 %)	90.8 % (15.2 %)
300 to ∞	95.6 % (14.0 %)	95.3 % (11.4 %)	93.1 % (8.5 %)

Table 8.8: The percentage of "correct" matches as a function of local image variance for test image #2. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used. The percentage of points within each variance range is shown in brackets.

Image Variance Range	RMSME (Pixels)		
	7	5	3
0 to 5	-	1.62	1.89
5 to 10	-	1.53	1.03
10 to 20	0.62	0.71	1.04
20 to 70	0.50	0.72	0.93
70 to 300	0.55	0.65	0.89
300 to ∞	0.55	0.74	1.07

Table 8.9: The RMSME as a function of the local image variance for test image #1. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used.

that will reduce the probability of a false match.

It should be noted that image pair #1 has a higher variance than image pair #2. This can be seen by comparing Tables 8.7 and 8.8, and Figures 7.3 and 7.4. This could account for the better performance when matching image #1.

The RMSME as a function of the local image variance is shown in Tables 8.9 and 8.10 for test image pair #1 and #2, respectively. They show a decrease in the RMSME as the local image variance increases. This can be explained by increase in the percentage "correct" matches.

A similar procedure was used to determine the effect of the disparity variance on the performance of the matching algorithm. A disparity variance map was calculated for each test image pair. This was done by calculating the variance of a set of points in a rectangular window centred at the point of interest. The variance maps obtained for the horizontal and vertical components of the disparity were summed together to obtain the disparity variance map. A window size equal to correlation window size was used.

Image Variance Range	RMSME (Pixels)		
	7	5	3
0 to 5	1.13	1.10	1.66
5 to 10	0.65	0.72	1.08
10 to 20	0.59	0.62	0.70
20 to 70	0.61	0.67	0.67
70 to 300	0.49	0.50	0.55
300 to ∞	0.38	0.40	0.47

Table 8.10: The RMSME as a function of the local image variance for test image #2. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used.

Disparity Variance Range	Percentage "Correct" Matches		
	7	5	3
0 to 0.05	94.7 % (16.8 %)	86.2 % (53.5 %)	71.2 % (86.0 %)
0.05 to 0.10	91.4 % (40.9 %)	77.5 % (33.1 %)	48.9 % (10.6 %)
0.10 to 0.20	85.1 % (24.0 %)	62.3 % (9.7 %)	41.8 % (2.6 %)
0.20 to 0.50	68.5 % (7.3 %)	43.8 % (3.1 %)	24.1 % (0.7 %)
0.5 to 1.0	46.9 % (1.0 %)	20.3 % (0.6 %)	53.3 % (0.1 %)

Table 8.11: The percentage of "correct" matches as a function of disparity variance for test image pair #1. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used. The percentage of points within each disparity variance range is shown in brackets.

The disparity of the points with their disparity variance within specified ranges was classified. The percentage of "correct" matches and RMSME was then calculated. The results obtained, from using the hierarchical area correlation matching algorithm with back-matching and inconsistency removal, were used.

The percentage of "correct" matches as a function of the disparity variance is shown in Tables 8.11 and 8.12 for test image pair #1 and #2, respectively, for a correlation window radius of 7, 5 and 3 pixels. The percentage of points within the specified disparity variance range is shown in brackets.

The results indicate that the best performance is obtained in regions of low disparity variance. This is the expected result. It is stated in Section 2.3.5 that the disparity is assumed to be constant over the correlation window. Therefore, the further the deviation from this assumption the worse the performance.

The RMSME as a function of the local disparity variance is shown in Tables 8.13 and 8.14, respectively. They show a decrease in the RMSME as the disparity variance decreases.

Disparity Variance Range	Percentage "Correct" Matches		
	7	5	3
0 to 0.05	100 % (0.5 %)	94.7 % (3.8 %)	87.8 % (24.5 %)
0.05 to 0.10	98.2 % (4.4 %)	94.9 % (14.2 %)	82.0 % (34.6 %)
0.10 to 0.20	95.9 % (16.8 %)	91.2 % (34.5 %)	64.0 % (27.4 %)
0.20 to 0.50	91.2 % (48.0 %)	77.7 % (36.3 %)	52.5 % (13.3 %)
0.5 to 1.0	76.9 % (30.3 %)	61.5 % (11.2 %)	80.5 % (0.2 %)

Table 8.12: The percentage of "correct" matches as a function of disparity variance for test image pair #2. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used. The percentage of points within each disparity variance range is shown in brackets.

Disparity Variance Range	RMSME (Pixels)		
	7	5	3
0 to 0.05	0.42	0.62	1.02
0.05 to 0.10	0.50	0.77	1.02
0.10 to 0.20	0.65	0.94	1.38
0.20 to 0.50	0.80	1.19	2.19
0.5 to 1.0	1.38	2.07	1.80

Table 8.13: The RMSME as a function of the local disparity variance for test image #1. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used.

This can be explained by the increase in the percentage "correct" matches.

In summary, the effect of the image variance and the disparity variance on the performance of the matching algorithm is as follows:

- The higher the local variance the better the performance. This can be explained by a higher local signal-to-noise ratio.
- The higher the local disparity variance the worse the performance. This can be explained by the violation of the assumption of constant disparity (see Section 2.3.5).

8.6 Discussion of image matching results

A discussion of the image matching results is presented in this section. An attempt is made to determine the degree to which the requirements of an automatic digital stereo image matching system are met (Section 8.6.1). A discussion of the implications for stereo height accuracy of the imaging sensor is also presented (Section 8.6.2).

Disparity Variance Range	RMSME (Pixels)		
	7	5	3
0 to 0.05	0.18	0.44	0.64
0.05 to 0.10	0.32	0.45	0.73
0.10 to 0.20	0.38	0.53	1.49
0.20 to 0.50	0.54	0.85	1.75
0.5 to 1.0	0.87	1.20	0.81

Table 8.14: The RMSME as a function of the local disparity variance for test image #2. A correlation window radius of 3, 5 and 7 pixels was used. The hierarchical area correlation matching algorithm with back-matching and inconsistency removal was used.

8.6.1 Image Matching Requirements

In Section 1.2 a set of requirements for an automatic digital stereo image matching system was presented. Each of these requirements will be discussed in turn.

Accuracy

Sub-pixel matching accuracy was required. A RMS matching error of less than one pixel was achieved with back-matching and inconsistency removal. Thus the requirement of sub-pixel matching accuracy can be met.

Reliability

The reliability requirement implied that gross matching errors must not be made. The maximum matching reliability, as measured by the percentage of "correct" matches, achieved was 89.6 %. If this is acceptable would depend on the application.

Generality

It was required that the matching system be applicable to as wide a range of images as possible. The degree to which this required could be met was limited by the available data (see Section 7.2.2). However, the investigation did show that the use of back-matching and inconsistency removal reduced the variance in accuracy and reliability.

Predictability

Predictable performance was required. The investigation into scene characteristics (Section 8.5) did show that the effect of scene characteristics could be explained by assumptions made.

Complexity

Minimal computational complexity was required. However, improvement in accuracy and reliability was always at the expense of complexity. The tradeoff between complexity, and accuracy and reliability will be application dependent.

8.6.2 Implications for Stereo Height Accuracy

In Section 5.7 a relationship was developed to estimate the height accuracy that can be achieved by a stereo imaging sensor. The height accuracy was a function of the base-to-height ratio and the x parallax measurement error ($\sigma\Delta p$). Assuming the results presented in this Chapter are applicable and that the x parallax measurement error is only due to matching errors then an estimate can be made of the height accuracy of the sensor design in Chapter 6 (see Table 6.3).

The lowest RMS matching error obtained was 0.53 pixels (see Table 8.6). Equation 5.20 and a RMS matching error of 0.53 pixels gives a x parallax measurement error $\sigma\Delta p$ of 0.75 pixels. The pixel spacing at the object plane of the sensor is 16 m and the base-to-height ratio 1.0. Therefore, an estimate for the height measurement error $\sigma\Delta p$ of the sensor is 12 m.

Chapter 9

Conclusions and Recommendations

9.1 Summary of the Work

The first problem addressed in this dissertation was an investigation into the feasibility and design of a satellite based imaging sensor. The sensor had to produce high resolution multi-spectral digital images. Furthermore, it had to produce a stereo image pair in a single pass. Finally, it had to be possible to fly it on a micro-satellite, like SUNSAT.

A concept design, motivated by the design and method of operation of the Landsat and SPOT satellites, limited the scope of the investigation. It defined the basic method of operation and technologies to be used. Furthermore, only a theoretical investigation was done, practical implementation was beyond the scope of the investigation.

The objectives of the investigation were to determine the feasibility of the imaging sensor, and to develop a specification for the imaging sensor.

The investigation was divided into three stages. First, a review of the relevant theory and literature was undertaken. This was used as the theoretical basis for the next two stages of the investigation. The radiometric theory required to estimate the radiation which reaches the sensor was reviewed. The terms, spatial resolution, spectral resolution, radiometric resolution and temporal resolution, were defined. The quantum nature of light and its effect on the noise performance of all light sensing devices was investigated. The characteristics of charge coupled devices, including the noise performance and spatial frequency response, were identified. The effect of quantisation on the signal-to-noise ratio of a signal was determined. The orbital mechanics of a satellite in a circular, sun-synchronous, near-polar orbit was examined. A model of satellite sensor geometry, developed by Sawada *et al.* [53], was identified. Finally, the Landsat and SPOT remote sensing satellite missions were investigated.

The second stage of the investigation dealt with the development of a model of the imaging sensor. The model related parameters of the sensor, the satellite and the environment to various output characteristics of the sensor. The environmental parameters included the spectral irradiance incident on the Earth, the reflectivity of the Earth, the atmospheric transmission

characteristics and other physical properties of the Earth. Sensor and satellite parameters included the orbit of the satellite, the focal length of the lens, the entrance pupil diameter, the obscuration of the lens, the spectral bands, the CCD characteristics, the offset of the sensor from the vertical and the A/D resolution. The output characteristics included geometrical parameters (the distance from sensor to object and the base-to-height ratio), the MTF, the CCD output voltage, the signal-to-noise ratio, the spatial resolution, the ground track, the accuracy of stereo height measurements and the data rate.

The final stage of the investigation dealt with the determination of the sensor feasibility and with the development of the sensor specification. This was done in a three-stage process. First, certain sensor and satellite parameters were fixed or limited to certain values. This reduced the scope of design investigation and ensured that the design requirements were met. Second, the feasibility of the imaging sensor was investigated. The results showed that sufficient radiation will reach the sensor for the requirement of high resolution to be satisfied. Furthermore, it was found that the spatial frequency response of the sensor is not limited by lens. Finally, it was found the proposed sensor geometry will allow stereo image pairs to be produced in a single pass. In the final stage of the process, the imaging sensor specifications were developed. These specifications included three spectral bands, a focal length 535 mm and a base-to-height ratio of 1.0.

Thus, the objectives were achieved. It was shown that the imaging sensor is feasible and a design specification was developed.

The second problem to be addressed in this dissertation was that of an investigation into automatic digital stereo image matching for cartography applications. An automatic digital stereo image matching system would have to meet accuracy, reliability, generality, predictability and complexity requirements.

The objectives of the investigation were: (a) to identify the degree to which the different requirements can be satisfied, (b) to investigate techniques that improve the reliability, and (c) to identify the effect of scene characteristics on performance.

The investigation was divided into three stages. First, a review of the relevant theory and literature was undertaken. The computational stereo vision paradigm of Barnard and Fischler [4] was reviewed. A comparison of area and feature matching was made. The digital image matching problem was formulated. Correlation based measures of match and mismatch were identified. The search for corresponding points was examined. Techniques for estimating the position of the corresponding points to sub-pixel accuracy were investigated. Finally, the wavelet transform and multiresolution analysis was reviewed.

The second stage of the investigation dealt primarily with the development of a hierarchical area correlation matching algorithm. It was used as the basis for the final stage of the investigation. This stage of the investigation also dealt with the preparation of experimental data, the development of a technique for evaluating the results, and an investigation into the

applicability of the wavelet transform.

The final stage of the investigation dealt with techniques for improving the reliability of the results obtained and with the effect of scene characteristics. Attention was also paid to the degree to which the requirements can be satisfied. Back-matching, the removal of inconsistencies and the detection of occlusion points were investigated. It was found that the use of all three techniques together gave the best results. The effect of the image and disparity variance was investigated. It was found that the best results were obtained in region of high image variance and low disparity variance. The results showed that the accuracy requirements could be met. A reliability of 89.6 % was achieved. Furthermore, the results could be explained by the assumptions made. Finally, it was found that there is a tradeoff between complexity, and accuracy and reliability.

Thus, the objectives were achieved. The degree to which the requirements can be satisfied was found. Techniques for improving the reliability were investigated and the effect of scene characteristics, in particular the image and disparity variance, was identified.

9.2 Conclusions

Based on the findings of this dissertation, the following conclusions may be drawn:

1. Imaging Sensor
   (a) It is possible to build an imaging sensor that meets the stated requirements.
   (b) The following three spectral bands are optimal: 0.5 μm to 0.6 μm, 0.6 μm to 0.7 μm and 0.7 μm to 0.8 μm.
   (c) An 8-bit analogue-to-digital converter would be optimal.
   (d) A focal length of 535 mm would be optimal.
   (e) A base-to-height ratio of 1 would be optimal.
   (f) The stereo overlap can be maximised by a baseline yawing algorithm.
2. Image Matching
   (a) The wavelet transform can be applied to digital stereo image matching.
   (b) The accuracy requirement can be met.
   (c) A reliability of 89.6 % can be achieved.
   (d) Performance can be explained by the assumptions made.
   (e) There is a tradeoff between computational complexity, and accuracy and reliability.
   (f) Best results are achieved in regions of high image variance and low disparity variance.
   (g) The use of all three techniques for improving reliability is best.

9.3 Recommendations

As a result of the findings and conclusions of this dissertation, the following recommendations are made regarding the imaging sensor design:

1. The following three spectral bands should be used: 0.5 μm to 0.6 μm, 0.6 μm to 0.7 μm and 0.7 μm to 0.8 μm.
2. The A/D resolution should be 8-bits.
3. The focal length of the lens should be 535 mm.
4. The sensor offset from the vertical should be 23.4 degrees.
5. A baseline yawing algorithm should be used to maximise stereo overlap.

Regarding automatic image matching, it is recommended that the following topics should be investigated further:

- The estimation of the position of the corresponding points.
- The reduction of the computational complexity.
- The use of satellite stereo images.

Bibliography

[1] John B. Wellman Alexander F. H. Goetz and William L. Barnes. Optical remote sensing of the earth. *Proceedings of IEEE*, 73(6):950–969, June 1985.

[2] Michel Arnaud and Marc Leroy. SPOT 4: a new generation of SPOT satellites. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46:205–215, 1991.

[3] Nicholas Ayache and Francis Lustman. Trinocular stereo vision for robotics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(1):73–85, January 1991.

[4] Stephen T. Barnard and Martin A. Fischler. Computational stereo. *Computing Surveys*, 14(4):553–572, December 1982.

[5] Peter J. Burt and Edward H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, Com-31(4):532–540, April 1983.

[6] Moustafa T. Chahine. Interaction mechanisms within the atmosphere. In R. N. Colwell and D. S. Simonett, editors, *Manual of Remote Sensing*, volume 1, chapter 5. American Society of Photogrammetry, second edition, 1983.

[7] M. Chevrel, M. Courtois, and G. Weill. The SPOT satellite remote sensing mission. *Photogrammetric Engineering and Remote Sensing*, 47(8):1163–1171, August 1981.

[8] C. K. Chui. *An Introduction to Wavelets.* Academic Press, 1992.

[9] Alden P. Colvocoresses. An automated mapping satellite system (Mapsat). *Photogrammetric Engineering and Remote Sensing*, 48(10):1585–1591, October 1982.

[10] K. L. Coulson. *Solar and Terrestrial Radiation.* Academic Press, 1975.

[11] Paul J. Curran. *Principles of Remote Sensing.* Longman Scientific & Technical, 1986.

[12] S. Curray. Calibration of an array camera. *Photogrammic Engineering and Remote Sensing*, 52:627–636, May 1986.

[13] I. Daubechies. *Ten Lectures on Wavelets.* Number 61 in CBMS-NSF Series in Applied Mathematics. SIAM Publications, Philadelphia, 1992.

[14] Jean Denègre. Role of satellite remote sensing in the geographic information economics in France. *ISPRS Journal of Photogrammetry and Remote Sensing*, 46:359–370, 1991.

[15] Umesh R. Dhond and J. K. Aggarwal. Structure from stereo – a review. *IEEE Transactions on Systems, Man and Cybernetics*, 19(6):1489–1510, November/December 1989.

[16] K. I. Duck and J. C. King. Orbital mechanics for remote sensing. In R. N. Colwell and D. S. Simonett, editors, *Manual of Remote Sensing*, volume 1, chapter 16. American Society of Photogrammetry, Falls Church, Virgina, second edition, 1983.

[17] John C. Feltz. Development of the modulation transfer function and contrast transfer function for discrete systems, particularly charge coupled devices. *Optical Engineering*, 29(8):893–904, August 1990.

[18] Margaret M. Fleck. A topological stereo matcher. *International Journal of Computer Vision*, 6(3):197–226, 1991.

[19] Sanjib Kumar Ghosh. *Analytical Photogrammetry.* Pergamon Press, second edition, 1988.

[20] A. F. M. Goetz and L. C. Rowan. Geologic remote sensing. *Science*, 211(20):781–791, February 1981.

[21] Ardeshir Goshtasby. Template matching in rotated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(3):338–344, May 1985.

[22] Derek Griffith, November 1991. Private Communication.

[23] Derek Griffith, August 1992. Private Communication.

[24] W. Eric L. Grimson. Computational experiments with a feature based stereo algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(1):17–34, January 1985.

[25] David Halliday and Robert Resnick. *Fundumentals of Physics.* John Wiley & Sons, 1986.

[26] Marsha Jo Hannah. A system for digital stereo image matching. *Photogrammetric Engineering and Remote Sensing*, 55(12):1765–1770, December 1989.

[27] Robert M. Haralick and Linda G. Shapiro. *Computer and robot vision*, volume 2. Addison-Welsey, 1993.

[28] H. Hecht. *Optics.* Addison-Wesley, 1987.

[29] Yuan C. Hsieh, David M. McKeown, and Frederic P. Perlant. Performance evaluation of scene registration and stereo matching for cartographic feature extraction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):214–237, February 1992.

[30] John P. Kerekes and David A. Landgrebe. An analytical model of earth-observational remote sensing systems. *IEEE Transactions on Systems, Man and Cybernetics*, 21(1):125–133, January/February 1991.

[31] The Khoros Group, Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque. *Khoros Manual - Volume I: User's Manual.*

[32] The Khoros Group, Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque. *Khoros Manual - Volume II: Progammer's Manual.*

[33] Miles V. Klein and Thomas E. Furtak. *Optics.* John Wiley & Sons, 1986.

[34] K. YA. Kondratyev. *Radiation in the Atmosphere.* Academic Press, 1969.

[35] Reimar Lenz. Image data acquisition with CCD cameras. In A. Gruen and H. Kahmen, editors, *Optical 3-D Measurement Techniques*, pages 22–34. Herbert Wichman, 1989.

[36] Donald L. Light. Characteristics of remote sensors for mapping and earth science applications. *Photogrammetric Engineering and Remote Sensing*, 56(12):1613–1623, December 1990.

[37] S. G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

[38] Stephane G. Mallat. Multifrequency channel decompositions of images and wavelet models. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(12):2091–2110, December 1989.

[39] Petros Maragos. Optimal morphological approaches to image matching and object detection. In *First International Conference on Computer Vision*, pages 695–699, 1988.

[40] D. Marr and T. Poggio. A computational theory of human stereo vision. *Proc. R. Soc. Lond. B.*, 204:301–328, 1979.

[41] Paul M. Mather. *Computer Processing of Remotely-Sensored Images: An Introduction.* John Wiley & Sons, 1987.

[42] Gerard Medioni and Ramakant Nevatia. Segment-based stereo matching. *Computer Vision, Graphics and Imaging Processing*, 31:2–18, 1985.

[43] Irwin Miller and John E. Freund. *Probability and statistics for engineers.* Prentice-Hall, third edition, 1985.

[44] G. W. Milne. Personal Communication, 1992.

[45] G. W. Milne, P. J. Bakkes, A. Schoonwinkel, J. J. du Plessis, J. H. R. Enslin, W. H. Steyn, S. Mostert, K. D. Palmer, and D. M. Weber. SUNSAT, Stellenbosch Univesity and SA-AMSAT's remote sensing and packet communications microsatellite. To appear in the Seventh Annual AIAA/USU Conference on Small Satellites, August 1993.

[46] Wayne Niblack. *An Introduction to Digital Image Processing.* Prentice/Hall International, 1986.

[47] F. Raye Norvelle. Stereo correlation: Window shaping and DEM corrections. *Photogrammetric Engineering and Remote Sensing*, 58(1):111–115, January 1992.

[48] Virgina T. Norwood and Jr Jack C. Lansing. Electro-optical imaging sensors. In R. N. Colwell and D. S. Simonett, editors, *Manual of Remote Sensing*, volume 1, chapter 8. American Society of Photogrammetry, second edition, 1983.

[49] M. Okutomi and T. Kanade. A locally adaptive window for signal matching. In *Third International Conference on Computer Vision*, pages 190–199, 1990.

[50] Alan V. Oppenheim and Ronald W. Scafer. *Discrete-time Signal Processing.* Prentice-Hall, 1989.

[51] RCA Corporation. *Electro-Optics Handbook*, 1974.

[52] Azriel Rosenfeld and Avinash C. Kak. *Digital Picture Processing.* Academic Press, 1976.

[53] Nobuo Sawada, Masatsugu Kidode, Hidenori Shinoda, Haruo Asada, Masayuki Iwanaga, Sadakazu Watanabe, Ken-Ichi Mori, and Minoru Akiyama. An analytic correction method for satellite MSS geometric distortions. *Photogrammetric Engineering and Remote Sensing*, 47(8):1195–1203, August 1981.

[54] Toni Schenk, Jin-Cheng Li, and Charles Toth. Towards an autonomous system for orienting digital stereopairs. *Photogrammetric Engineering and Remote Sensing*, 75(8):1057–1064, August 1991.

[55] William F. Schreiber. *Fundamentals of Electronic Imaging Systems.* Springer-Verlag, 1986.

[56] Philip N. Slater. Use of MTF in the specification and first order design of electro-optical and photographic imaging and radiometric systems. *Optica Acta*, 22(4):277–290, 1975.

[57] Philip N. Slater. *Remote sensing, optics and optical systems.* Addison-Wesley, 1980.

[58] Ferrel G. Stremler. *Introduction to Communication Systems.* Addison-Wesley, second edition, 1982.

[59] Gwynn H. Suits. The nature of electromagnetic radiation. In R. N. Colwell and D. S. Simonett, editors, *Manual of Remote Sensing*, volume 1, chapter 16. American Society of Photogrammetry, Falls Church, Virgina, second edition, 1983.

[60] Philip H. Swain and Shirley M. Davis, editors. *Remote Sensing. The Quantitative Approach.* McGraw-Hill, 1978.

[61] R. Welch. Measurements from linear array camera images. *Photogrammetric Engineering and Remote Sensing*, 46(3):315–318, March 1980.

[62] R. Welch. Impact of geometry on height measurements from MLA digital image data. *Photogrammetric Engineering and Remote Sensing*, 49(10):1437–1441, October 1983.

[63] R. Welch and Wayne Marko. Cartographic potential of a spacecraft line-array camera system: Stereosat. *Photogrammetric Engineering and Remote Sensing*, 47(8):1173–1185, August 1981.

Appendix A

Spectral Solar Irradiance Data

Wavelength μm	Solar Spectral Irradiance W m^{-2} μm^{-1}
0.400	1429
0.405	1644
0.410	1751
0.415	1774
0.420	1747
0.425	1693
0.430	1639
0.435	1663
0.440	1810
0.445	1922
0.450	2006
0.455	2057
0.460	2066
0.465	2048
0.470	2033
0.475	2044
0.480	2074
0.485	1976
0.490	1950
0.495	1960
0.500	1942
0.505	1920
0.510	1882
0.515	1833
0.520	1833

Appendix A: Spectral Solar Irradiance Data

Wavelength μm	Solar Spectral Irradiance W m^{-2} μm^{-1}
0.525	1852
0.530	1842
0.535	1818
0.540	1783
0.545	1754
0.550	1725
0.555	1720
0.560	1695
0.565	1705
0.570	1712
0.575	1719
0.580	1715
0.585	1712
0.590	1700
0.595	1682
0.600	1666
0.605	1647
0.610	1635
0.620	1602
0.630	1570
0.640	1544
0.650	1511
0.660	1486
0.670	1456
0.680	1427
0.690	1402
0.700	1369
0.710	1344
0.720	1314
0.730	1290
0.740	1260
0.750	1235
0.800	1107
0.850	988
0.900	889

Appendix B

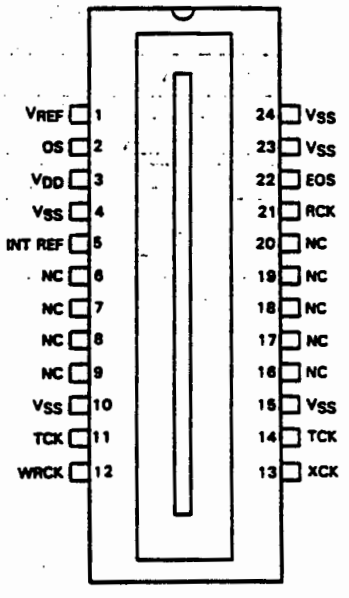
TC104 Datasheet

TC104
3456 × 1 CCD LINEAR IMAGE SENSOR

D2687, FEBRUARY 1983

- 3456 x 1 Sensor Element Organization
- Virtual-Phase N-Channel Silicon MOS Technology
- High Quantum Efficiency
- Enhanced Blue Response
- Output Signal Approximately 0.6 Volt Peak-to-Peak
- Dynamic Range Relative to Peak-to-Peak Noise Typically 1000:1
- End-of-Scan Signal
- Internal Black and White References
- Simple and Stable Operation
- OPTIONAL FEATURE: Internal Reference Voltage

TC104 . . . DUAL-IN-LINE PACKAGE
(TOP VIEW)



NC — No internal connection.

description

The TC104, a 3456-element CCD line image sensor, functions in high-resolution image scanning applications such as document reading and optical character recognition. The TC104 incorporates virtual-phase MOS technology, which provides simplified operation and high reliability. The 3456 sensor elements provide 400 points-per-inch resolution across 8.5 inches.

This device is supplied in a 24-pin dual-in-line ceramic side-braze package designed for insertion in mounting-hole rows on 15.2-mm (0,600-inch) centers. The glass window may be cleaned by wiping with a cotton swab soaked in alcohol.

Caution. These devices have limited built-in gate protection. The leads should be shorted together or the device placed in conductive foam during storage or handling to prevent electrostatic damage to the MOS gates. Avoid shorting either OS or EOS to V_{SS} during operation to prevent damage to the amplifiers.

virtual phase technology

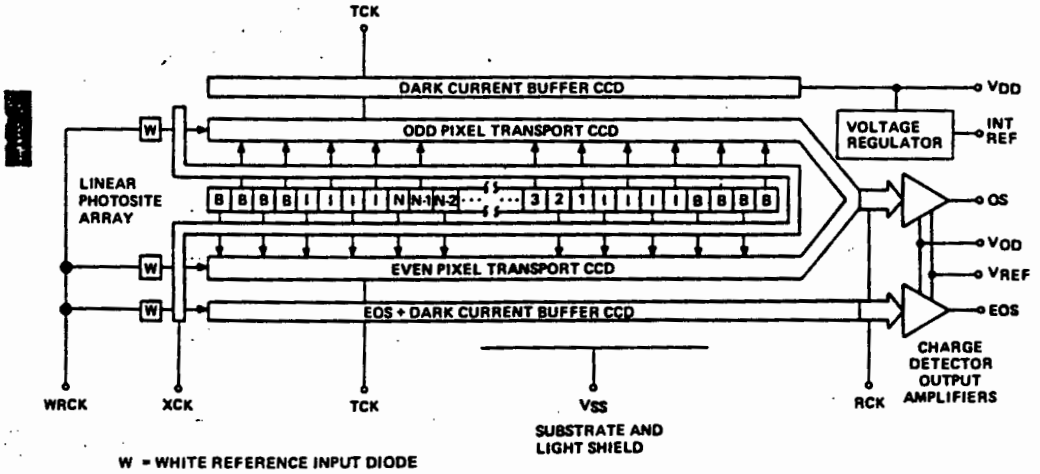
This patented design results in simplified clocking circuits, reduced noise, and greater light sensitivity. The virtual phase utilizes a junction-gate region at the substrate dc potential. This accomplishes the same gating and transport function as a separate gate electrode requiring multiple layers and multiple process steps common in other device designs. The resulting simplicity of process and ease of operation will increase performance and reliability for the user.

TEXAS INSTRUMENTS
POST OFFICE BOX 655012 • DALLAS, TEXAS 75265

TC104

3456 × 1 CCD LINEAR IMAGE SENSOR

functional block diagram



W = WHITE REFERENCE INPUT DIODE
B = BLACK REFERENCE ELEMENT
I = ISOLATION ELEMENT
N = 3456 SENSOR ELEMENTS

PIN FUNCTIONAL DESCRIPTION

PIN NUMBER	SIGNATURE	NAME	DESCRIPTION
1	V_{REF}	Reference Voltage	Bias input for the output amplifiers and internal reference.
2	OS	Output Signal	Video output from a cascaded source-follower MOS amplifier.
3	V_{DD}	Supply Voltage	Output amplifier supply voltage.
4, 10, 15, 23, 24	V_{SS}	Substrate	All voltages are referenced to the substrate.
5	INT REF	Internal Reference	Potential derived internally for operational reference voltage.
6, 7, 8, 9, 16 17, 18, 19, 20	NC		No internal connection.
11, 14	TCK	Transport Clock	Drives the CCD transport registers.
12	WRCK	White Reference Clock	Injects a controlled charge into the white reference CCD shift register elements to become white-reference and end-of-scan pulses.
13	XCK	Transfer Clock	Controls the transfer of charge packets from sensor elements to shift registers. The interval between pulses of the transfer clock determines the exposure time.
21	RCK	Reset Clock	Controls recharging of the charge-detection diodes in the output amplifiers, and clocks the output shift registers where the odd and even signals have been merged.
22	EOS	End-of-Scan Pulse	Indicates that all charge packets have been shifted out of the transport registers.

functional description

image sensor elements

The line of sensor elements (also called photosites or pixels) consists of 3456 photo-sensitive areas, 10.7 micrometers (0.42 milliinches) square and approximately 10.7 micrometers from center to center. Image photons create electron/hole pairs in the single-crystal silicon. The electrons are collected in the sensor elements, and the holes are swept into the substrate. The amount of charge accumulated in each element is a linear function of the incident light and the exposure time. The output signal charge will vary in an analog manner from a thermally generated noise background at zero illumination to a maximum at saturation under bright illumination.

transfer gate

This structure is adjacent to the line of image sensor elements. The charge packets accumulated in the image sensor elements are transferred into the transfer gate storage well when the transfer gate voltage goes high. When the transfer gate voltage goes low, the charge is transferred into the CCD transport shift registers. The transfer gate also controls the exposure time for the sensor elements and permits charges to enter the end-of-scan (EOS) shift registers to create the end-of-scan waveform. In addition, the transfer gate permits entry of charge packets to the transport CCD shift register to create the white reference signals.

shift registers

There are two CCD transport registers, one on each side of the line of image sensor elements and outside of the transfer gate. Alternate charge packets are transferred to the CCD transport shift registers and moved serially to the output amplifier. The phase relationship of the reset clock and the transport clock and the geometric layout of the paths provide for alternate delivery of charge packets to re-establish the original sequence of the linear image data. The two outer buffer CCD shift registers protect the signal charges in the inner transport CCD shift registers from peripherally generated dark current noise.

black and white reference elements

Four additional sensor elements at each end of the sensor element array (labelled "B" in the block diagram) are covered by opaque metallization. They provide a black (no illumination) signal reference that is delivered at each end of the linear image output signal. Also included on the transport CCD shift register, at the opposite end from the amplifier, is an input diode that provides two white reference pulses in the output signal. The reference pulses are useful as inputs to external dc restoration and/or automatic exposure control circuitry. The white reference pulse amplitude is approximately 100% of the maximum output signal amplitude.

output signal amplifier

The charge packets are transported to a precharge diode whose potential changes linearly in response to the amount of the signal charge delivered. This potential is applied to the input gate of an N-channel MOS double-source-follower amplifier to produce an output signal (OS). A reset transistor, driven by the reset clock (RCK), recharges the charge-detector-diode capacitance before the arrival of each new signal charge packet from the CCD shift registers. Reference voltage (V_{REF}) is applied to the drain of the reset transistor and acts to bias the OS and EOS amplifiers. A current sink is used as an on-chip load for the amplifier output, so no external current sink is needed. The output signal on pin 2 is a series of negative-going pulses on a dc level.

internal reference voltage

An internal reference voltage (INT REF) is available on the chip to provide the V_{REF} voltage. The required connections appear in Figure 3. If the internal reference voltage is not used, an external voltage is connected directly to pin 1. Pin 5 is then left unconnected.

resolution

The modulation transfer function decreases at longer wavelengths. (See Figures 7 and 8.) If optimum resolution is required with a light source that has a significant infrared component, then the designer must use appropriate filters to restrict the optical pass band to shorter wavelengths.

end-of-scan amplifier

The EOS amplifier is similar to the OS amplifier. XCK transfers charge from the input diode into the EOS register where it is transported at the TCK clock frequency to the EOS amplifier. This EOS pulse is coincident with the first of the two white reference pulses that pass through the odd and even transport CCDs, respectively. The EOS output can be used to alert the external circuitry that the linear image data readout has been completed.

clocks

The transfer clock (XCK) pulse controls the exposure time of the sensor elements. The minimum exposure time is the time required to shift the entire contents of the transport registers to the output signal amplifier and equals 3489 multiplied by the RCK period. The maximum exposure time is determined by the tolerable level of dark signal.

The transport clock (TCK) transports the linear image signal charge from the sensor element region to the output amplifier.

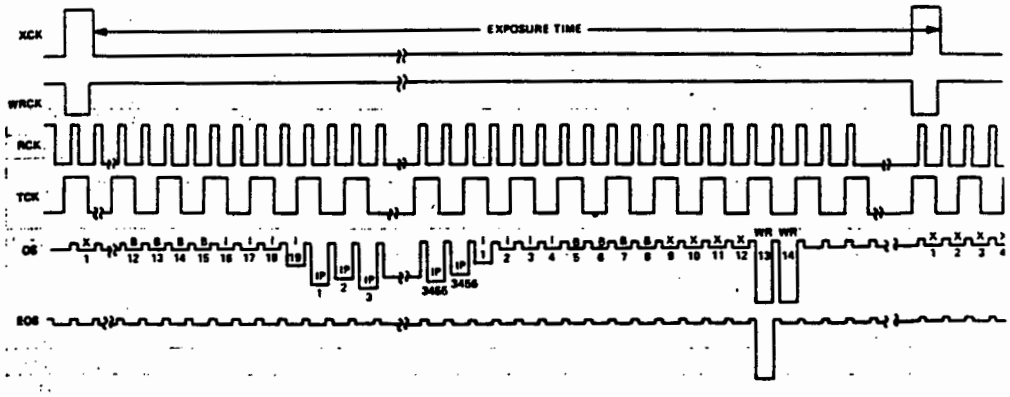
The reset clock (RCK) operates at twice the transport clock frequency so as to recombine the signal charge in the original sequence and present the charge to the output amplifier. The data rate is equal to the reset clock frequency.

The white reference clock (WRCK) runs at the transfer clock frequency and generates the white reference and the end-of-scan pulses. These pulses can be eliminated by connecting WRCK to V_{DD}. Transients on WRCK going below zero volts will cause charge injection resulting in an increase in apparent dark signal.

Figure 3 presents a suggested circuit for generating the clock waveforms. The RCK clock generator runs continuously. A binary divider halves the frequency to create TCK. After all signal charges have been transported to the output amplifier, TCK continues to run to keep thermally generated charges from accumulating in the transport registers.

The XCK and WRCK clock frequencies are submultiples of the TCK frequency. Figure 2 details the timing relationships among the different clock pulses.

TEXAS INSTRUMENTS
POST OFFICE BOX 655012 • DALLAS, TEXAS 75265



Output Signal (OS) pulse identification: I = isolation pixel, IP = image pixel, B = Black reference pixel, WR = White reference pixel, X = empty pixel.

FIGURE 1—OPERATING INPUT AND OUTPUT VOLTAGE WAVEFORMS

absolute maximum ratings over operating free-air temperature range (unless otherwise noted) (see Note 1)

Amplifier drain voltage (V_{DD})	−0.3 V to 30 V
Transfer clock (XCK) voltage	−25 V to 5 V
Transport clock (TCK) voltage	−25 V to 5 V
Reset clock (RCK) voltage	−25 V to 5 V
White reference clock (WRCK) voltage	−0.3 V to 30 V
Storage temperature	−25°C to 125°C
Operating free-air temperature	−25°C to 70°C

NOTE 1: Voltage values are with respect to V_{SS}.

recommended operating conditions at T_A = 25°C (see Note 1)

		MIN	NOM	MAX	UNIT
V_{DD}	Amplifier supply voltage	13	14	15	V
$V_{IH(X)}$	Transfer clock high-level input voltage	3	4	5	V
$V_{IL(X)}$	Transfer clock low-level input voltage	−15†	−14	−13	V
$V_{IH(T)}$	Transport clock high-level input voltage	3	4	5	V
$V_{IL(T)}$	Transport clock low-level input voltage	−15†	−14	−13	V
$V_{IH(R)}$	Reset clock high-level input voltage	3	4	5	V
$V_{IL(R)}$	Reset clock low-level input voltage	−15†	−14	−13	V
$V_{IH(WR)}$	White reference clock high-level input voltage	13	14	15	V
$V_{IL(WR)}$	White reference clock low-level input voltage	6	7	8	V
f_{RCK}	Reset clock frequency (output data rate)			8	MHz

†The algebraic convention, where the most negative limit is designated as minimum, is used in this data sheet for clock voltage levels only.
NOTE 1: Voltage values are with respect to V_{SS}.

TC104
3456 × 1 CCD LINEAR IMAGE SENSOR

electrical characteristics at 25°C free-air temperature ¶

PARAMETER		MIN	TYP	MAX	UNIT
Dark-signal amplitude	Average		0.5	10	mV
	Low frequency component		0.5	5	mV
	Nonuniformity relative to average of adjacent pixels		4	20	
Sensitivity		1.4	2	3.5	V/(μJ/cm²)
Output amplitude variation (PRNU)‡	Peak-to-peak		30	60	mV
	Adjacent pixels from alternate registers (imbalance)		10		
Peak-to-peak noise			0.6		mV
Equivalent exposure§ of peak-to-peak noise			0.3		nJ/cm²
Saturation exposure§			300		nJ/cm²
Saturation output amplitude		400	600	800	mV
Dynamic range relative to peak-to-peak noise†		500:1	1000:1		
Charge transfer efficiency			0.99999		
White reference amplitude		400	600		mV
End-of-scan amplitude		200	350		mV
Output offset (dc) voltage			6		V
Output impedance			1		kΩ
Resistance to V_{SS}	Transfer gate		150		kΩ
	Transport gate		700		
	Reset gate		700		
Amplifier reference voltage, V_{REF}			7		V
Capacitance to V_{SS}	Transfer gate		400		pF
	Transport gate		900		
	Reset gate		16		
Amplifier supply current			8	12	mA
Total power dissipation			112		mW

†Dynamic range = saturation output amplitude/standard deviation peak-to-peak noise.
‡Measured at 400 mV output amplitude with an f/2.8 lens.
§Exposure = intensity × time
¶Test conditions are f_{RCK} = 0.5 MHz, t_{exp} = 10 ms, tungsten light source operating at color temperature of 2854 K with 2.0-mm-thick Fish-Schurman HA-11 IR-absorbing filter, and all operating voltages at nominal recommended values using the internal reference voltage.

TEXAS INSTRUMENTS
POST OFFICE BOX 655012 • DALLAS, TEXAS 75265

timing requirements

		MIN	NOM	MAX	UNITS
t_{THXH}	Time delay from the transport clock rising edge to the transfer clock rising edge.	0		50	ns
t_{THWL}	Time delay from the transport clock rising edge to the white reference clock falling edge.	0		50	ns
t_{THRH}	Time delay from the transport clock rising edge to the reset clock rising edge.	0			ns
$t_{w(RH)}$	Pulse duration of the high state for the reset clock.	40			ns
t_{TLXL}	Time delay from the transport clock falling edge to the transfer clock falling edge.	50			ns
t_{TLWH}	Time delay from the transport clock falling edge to the white reference clock rising edge.	0		50	ns
t_{XLTH}	Time delay from the transfer clock falling edge to the rising edge of the next transport clock pulse.	50			ns
t_r	Rise time (all clocks)	15			ns
t_f	Fall time (all clocks)	5			ns

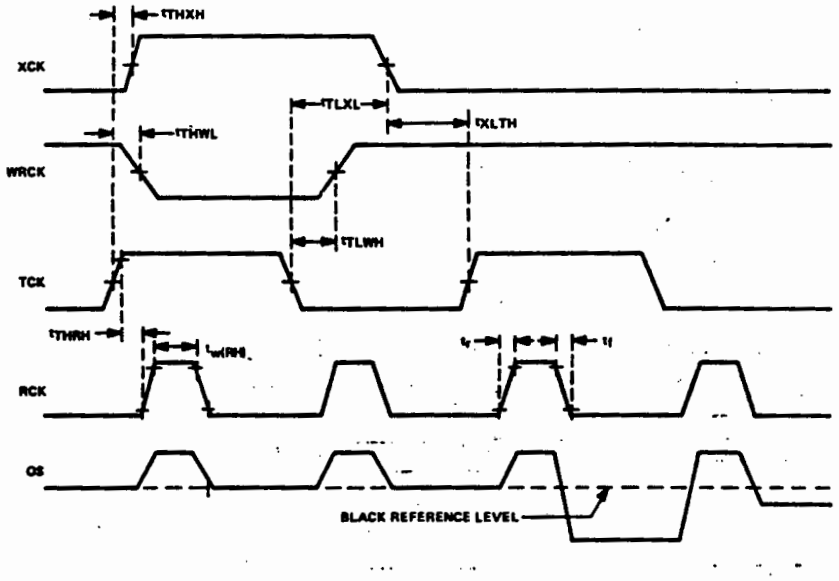
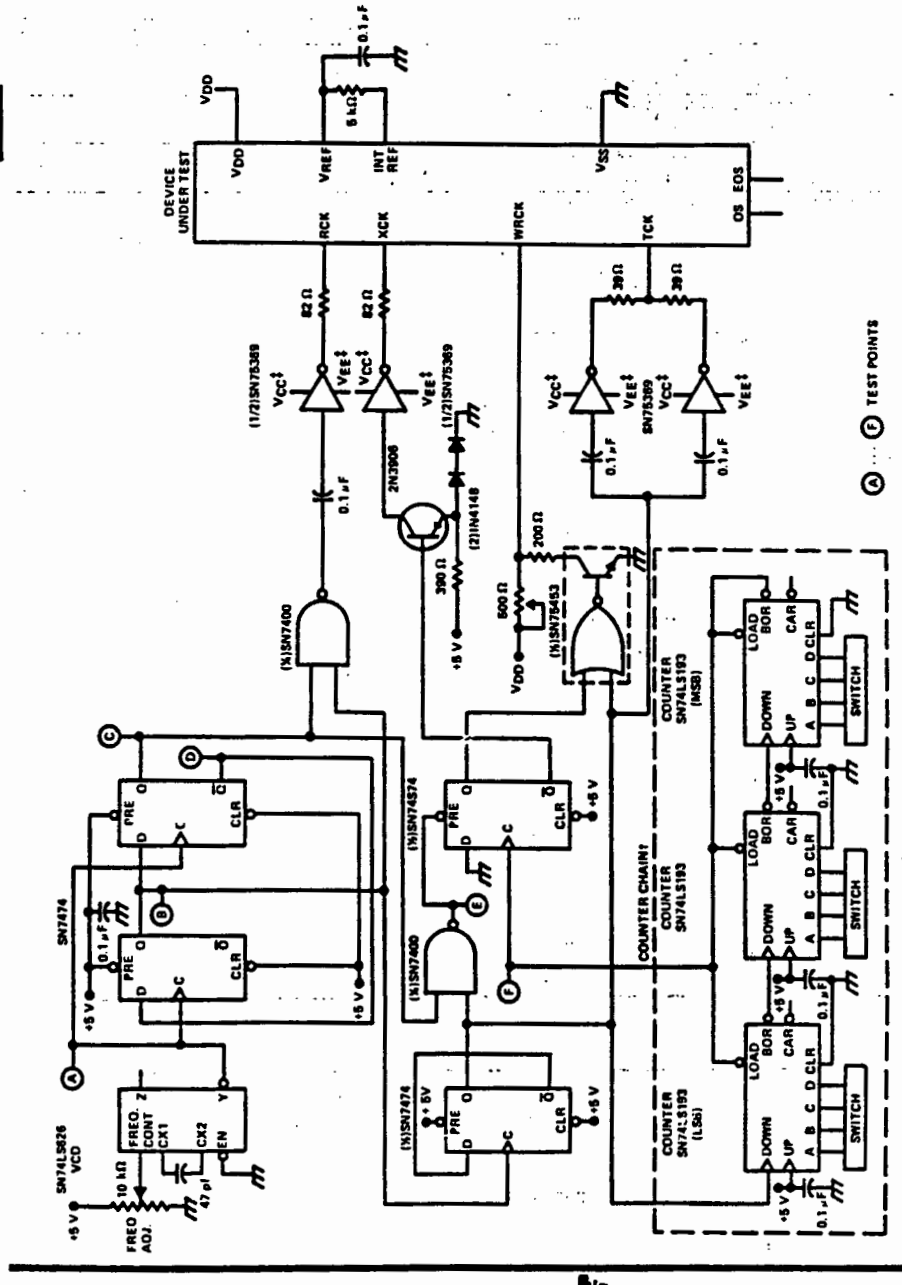


FIGURE 2—DEVICE TIMING REQUIREMENTS

TC104
3456 × 1 CCD LINEAR IMAGE SENSOR

PARAMETER MEASUREMENT INFORMATION



†This counter chain counts transport clock periods to generate the exposure time interval. The data rate is twice the count rate.
‡V_{CC} and V_{EE} are the voltages that will produce the desired values of V_{IH} and V_{IL}, respectively, at the RCK, XCK, and TCK inputs.

FIGURE 3–DRIVER CIRCUIT FOR TESTING IMAGE SENSOR

TEXAS INSTRUMENTS
POST OFFICE BOX 655012 • DALLAS, TEXAS 75265

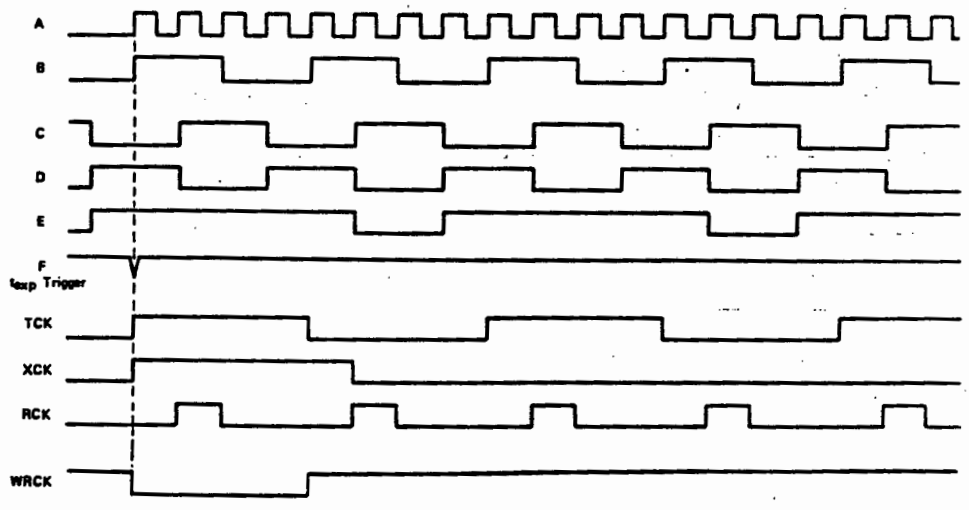


FIGURE 4 — WAVEFORMS IN DRIVER CIRCUIT

TYPICAL CHARACTERISTICS

(In the circuit of Figure 3 with T_A = 25 °C, f_{RCK} = 0.5 MHz, t_{exp} = 10 ms, and all operating voltages at nominal recommended values, unless otherwise noted)

SENSITIVITY
vs
WAVELENGTH OF INCIDENT LIGHT

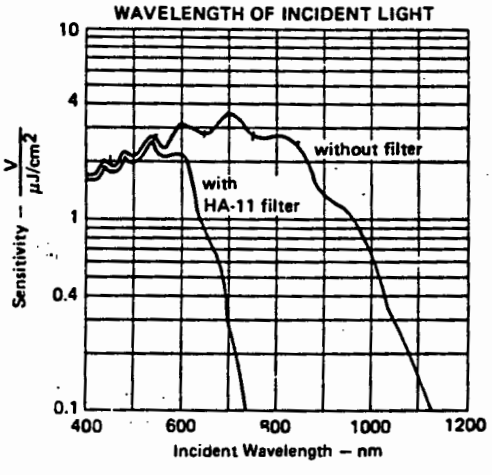


FIGURE 5

RESPONSIVITY
vs
WAVELENGTH OF INCIDENT LIGHT

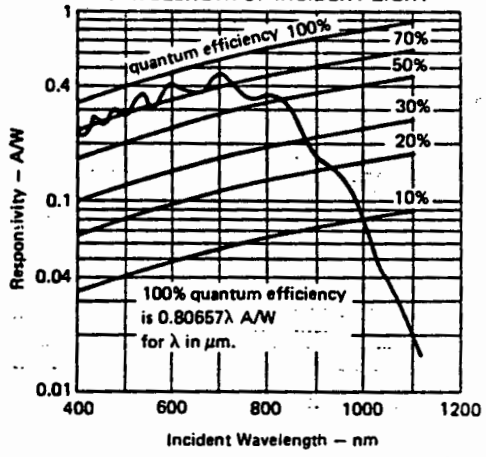


FIGURE 6

TEXAS INSTRUMENTS

TC104
3456 × 1 CCD LINEAR IMAGE SENSOR

TYPICAL CHARACTERISTICS

(In the circuit of Figure 3 with T_A = 25°C, f_{RCK} = 0.5 MHz, t_{exp} = 10 ms, and all operating voltages at nominal recommended values, unless otherwise noted)

MODULATION TRANSFER FUNCTION
vs
SPATIAL FREQUENCY

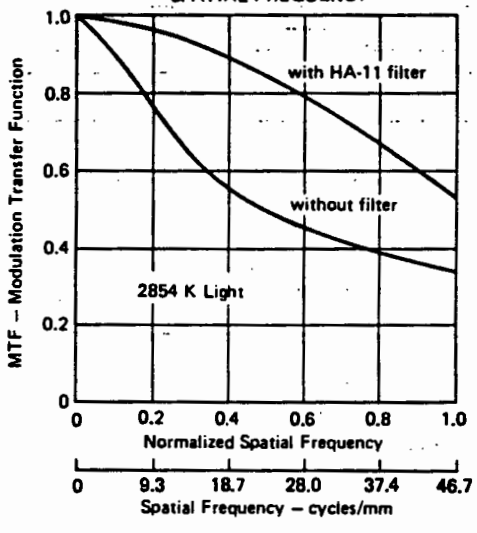


FIGURE 7

MODULATION TRANSFER FUNCTION
vs
SPATIAL FREQUENCY

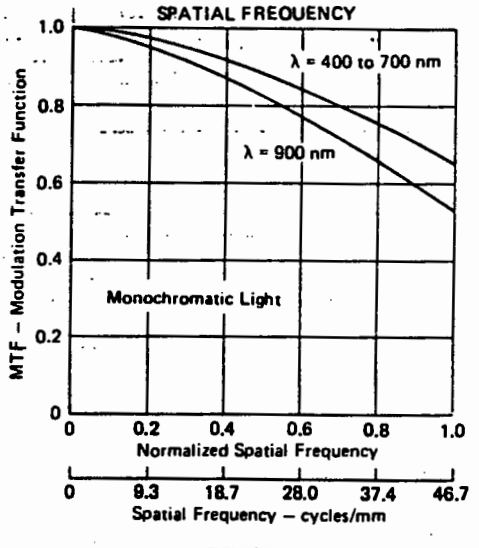


FIGURE 8

AVERAGE AND LOW FREQUENCY
DARK SIGNAL
vs
EXPOSURE TIME

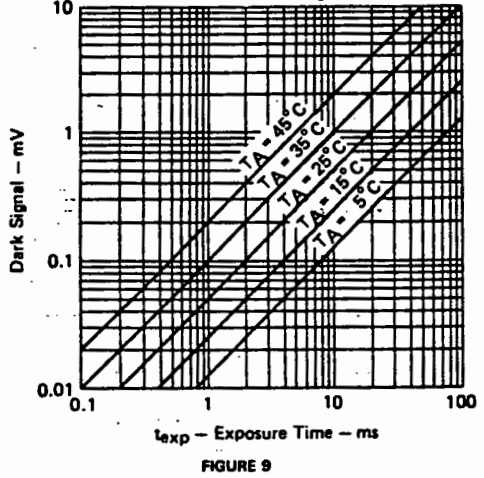


FIGURE 9

OUTPUT SIGNAL VOLTAGE RELATIVE TO
SATURATED OUTPUT VOLTAGE
vs
EXPOSURE TIME

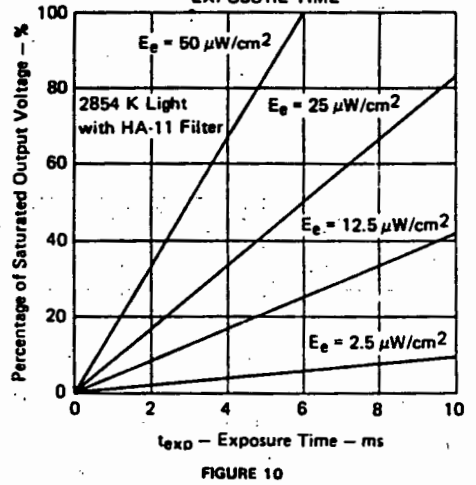
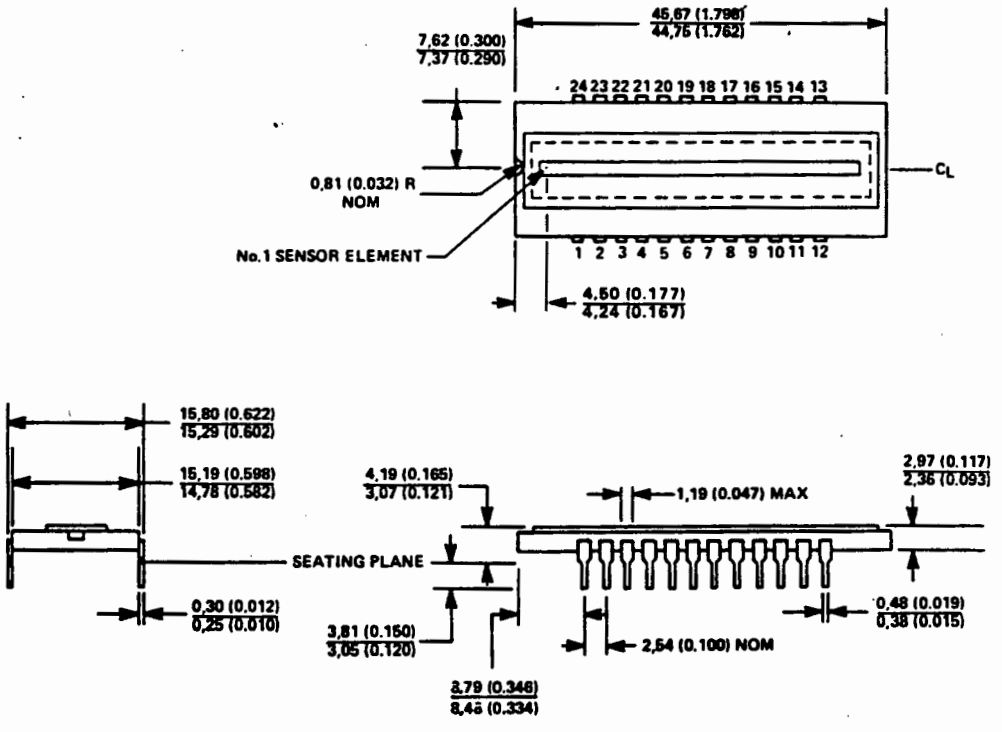


FIGURE 10

TEXAS

MECHANICAL DATA



ALL DIMENSIONS ARE IN MILLIMETERS AND PARENTHETICALLY IN INCHES.

NOTE 1: The distance between the top surface of the window and the surface of the sensor is nominally 0,89 (0.035). This is determined by observing the vertical motion of a microscope focused first at one plane, then at the other.